



UNIVERSIDADE FEDERAL DO CEARÁ
FACULDADE DE ECONOMIA ADMINISTRAÇÃO ATUÁRIAS E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA APLICADA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

ABEL BRASIL RAMOS DA SILVA

**ESTIMAÇÃO DE CURVAS DE INDIFERENÇA DE ACESSIBILIDADE URBANA VIA
MODELOS ORDENADOS E KRIGAGEM**

FORTALEZA

2013

ABEL BRASIL RAMOS DA SILVA

ESTIMAÇÃO DE CURVAS DE INDIFERENÇA DE ACESSIBILIDADE URBANA VIA
MODELOS ORDENADOS E KRIGAGEM

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Economia, da Faculdade de Economia Administração Atuárias e Contabilidade da Universidade Federal do Ceará, como requisito parcial para obtenção do Título de Mestre em Economia.

Orientador: Prof. Dr. José Raimundo de Carvalho Júnior

FORTALEZA

2013

ABEL BRASIL RAMOS DA SILVA

ESTIMAÇÃO DE CURVAS DE INDIFERENÇA DE ACESSIBILIDADE URBANA VIA
MODELOS ORDENADOS E KRIGAGEM

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Economia, da Faculdade de Economia Administração Atuárias e Contabilidade da Universidade Federal do Ceará, como requisito parcial para obtenção do Título de Mestre em Economia.

Aprovada em: Data de Aprovação 04/03/2013

BANCA EXAMINADORA

Prof. Dr. José Raimundo de Carvalho Júnior (Orientador)
Pós-graduação em Economia (CAEN-UFC)

Prof. Dr. Emerson Marinho
Pós-graduação em Economia (CAEN-UFC)

Prof. Dr. Juvêncio Santos Nobre
Departamento de Estatística e Matemática Aplicada (DEMA-UFC)

AGRADECIMENTOS

Mais uma etapa da minha vida está se encerrando! Fico feliz por isso! Pois, sabemos que depois de uma etapa concluída, novos caminhos irão se abrir, novas pedras terei que quebrar, ainda maiores que as de então! Mas, com a bagagem adquirida, em vez de quebrar, um “desvio da pedra” facilitará meu caminho. Tenho que agradecer a um “bucado” de gente que me ajudou a construir este trabalho.

Primeiramente, agradeço a Deus pela vida e a saúde que tenho.

Agradeço à minha Mãe e ao meu Pai, Conceição e Assis, e minha irmã, Amanda, por serem minha fortaleza em todos os dias de minha vida, que me ajudam e apoiam em tudo. E, também, a “família Brasil e Ramos” que estão sempre perto de mim.

À minha amiga e namorada Josy, por acreditar no meu potencial, que sempre me incentivará e me fortalecerá, para que eu concluisse mais esta jornada de minha vida.

Ao Prof. Dr. José Raimundo Carvalho, pela contribuição acadêmica e profissional; pela paciência de ensinar, de orientar e incentivar para que o trabalho seja feito com excelência. Pela longa parceria desde do início da bolsa de iniciação científica em 2008 à conclusão desta dissertação, além de ser um amigo.

Certamente, não posso deixar de agradecer aos professores do Departamento de Estatística da UFC que passaram 4 anos e meio me ensinando junto aos meus colegas de turma a nos tornar excelentes Estatísticos que somos hoje. Em especial, ao professor Mauricio pelo incentivo acadêmico e a paixão pela profissão e ao professor Juvêncio por ter me indicado ao CAEN-UFC como bolsista. Agradeço, também, aos profissionais que trabalham nos bastidores, em especial, Carmem e Marjorie, do CAEN e do DEMA, respectivamente, sempre tentando ajudar o aluno no que for necessário.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), pelo apoio financeiro fundamental para a efetivação deste trabalho de dissertação.

Aos meus amigos que contribuíram de alguma forma e que agora festejam mais uma etapa vencida. Em especial aos meus amigos sobreviventes da turma “Éramos desinteressados” pelo conhecimento, discussões e divertimento: Tato (saber, saber, não!), Max (“*I should have sknown better*”), Rômulo, Leandro (“Estatística: Bem Molin, mah!”) e Giuan; também aos parceiros de futebol Mario e Juanito (é “ru” professor!). Aos meus amigos de sala de mestrado! “Suponha” e “Suponha” que todos os nomes estão aqui (Tony, Iran, Luis Carlos, Lívia Menezes, Denise, Arley, Maurício (preto é transparente?), Rafael, Carla, Uyara e Lívia Rabelo),

rsrsrs. Aos meus amigos de colégio: Débora, Raquel, Sérgio e Whallen. Aos meus amigos de infância: Camilo, Tancredo, Frankson e Diego.

RESUMO

O processo de urbanização, crescimento das cidades e estruturação urbana ocorrido nas últimas décadas nas grandes cidades brasileiras vem colocando a questão da acessibilidade como fator relevante na qualidade de vida da população. Neste contexto, analisar rigorosamente o nível de acessibilidade e o bem-estar dos indivíduos a partir do momento que deixam suas residências até o ponto de execução de atividades ou satisfação de consumo torna-se uma questão de grande importância científica ainda pouco explorada de maneira rigorosa. Nesta dissertação buscamos analisar e modelar acessibilidade considerando uma perspectiva teórica baseada na metodologia da maximização da utilidade e na estimação de modelos econométricos. Para tanto, este estudo está dividindo em dois eixos de pesquisa: o primeiro, analisa a acessibilidade com o uso de modelos ordenados generalizados através de uma base inédita de micro dados geo-referenciados coletada na cidade de Fortaleza, Brasil. Os resultados mostram que variáveis como renda, posse de automóveis, distância, entre outras, são importantes para explicar a acessibilidade dos indivíduos. O segundo eixo de análise propõe e desenvolve, de maneira pioneira, uma superfície de utilidade espacial através de técnicas de krigagem. Os resultados mostram que a distância entre o domicílio e o ponto de destino possui uma relação bastante heterogênea com a acessibilidade, revelando um padrão espacial influenciado pela desigualdade econômica da cidade. Esse resultado coloca em dúvida suposições simplistas tradicionais que assumem uma relação linear ou polinomial entre distância e acessibilidade.

Palavras-Chave: Acessibilidade. Utilidade. Krigagem.

ABSTRACT

The process of urbanization, growth of cities and urban structuring in recent decades among large Brazilian cities revealed the issue of accessibility as a relevant factor in quality of life. In this sense, analyzing the level of accessibility and welfare of individuals from where they leave their homes up to the point of execution of activities or consumer satisfaction becomes a matter of great scientific importance, yet to be explored in rigorously way. Thus, in this dissertation we analyze and model urban accessibility considering a theoretical perspective based on the methodology of utility maximization and estimation of econometric models. Therefore, this study is divided into two lines of research. The first one analyzes the accessibility using generalized ordered models through a new geo-referenced micro data set collected in the city of Fortaleza, Brazil. Our results show that variables such as income, car ownership, distance, and others are important for explaining accessibility of individuals. The second line of inquiry proposes and develops, in a pioneering way, a surface of spatial utility by means of Kriging techniques. The results point to the fact that the distance between home and destination has a very heterogeneous relationship with accessibility, revealing a spatial pattern greatly influenced by the prevailing economic inequalities all over the city. This result puts into question simplistic traditional assumptions that assume a linear or polynomial relation between distance and accessibility.

Keywords: Accessibility. Utility. Kriging.

SUMÁRIO

LISTA DE FIGURAS	9
LISTA DE TABELAS	10
1 INTRODUÇÃO	11
2 CONCEITOS E MEDIDAS DE ACESSIBILIDADE	15
2.1 Introdução	15
2.2 A Classificação de Geurs e van Wee	16
2.2.1 Medidas baseadas na infraestrutura	17
2.2.2 Medidas baseadas na localização	17
2.2.3 Medidas baseadas no indivíduo	20
2.2.4 Medidas baseadas na utilidade	21
2.3 Utilidade e Acessibilidade	22
2.4 Componentes de Acessibilidade	25
2.4.1 <i>Transporte</i>	27
2.4.2 <i>Uso do solo</i>	28
2.4.3 <i>Tempo</i>	28
2.4.4 <i>Individual</i>	29
2.5 Critérios para Medidas de Acessibilidade	30
3 SURVEY DE MODELOS DE ESCOLHA DISCRETA	30
3.1 Introdução	30
3.2 Modelos de Escolha Binária	30
3.3 Modelos Lineares de Probabilidade	31
3.4 Modelos de Resposta Binária: Probit e Logit	32
3.4.1 Transformação em $x\beta$	32
3.4.2 Utilidade para modelos de escolha binária	33
3.4.3 Fundamentos teóricos para variável latente	34
3.5 Modelos de Escolha Ordenada	35
3.5.1 Modelos logit cumulativos	38
3.6 Modelos de Escolha Ordenada Generalizados	40
3.6.1 Testando a hipótese de regressão paralela	42

4	BASE DE DADOS	42
5	UM MODELO ECONOMETRICO DE ACESSIBILIDADE URBANA	47
5.1	Introdução	47
5.2	Modelo Econométrico	47
5.3	Destino: Hospital de Emergência - Instituto José Frota (IJF)	48
5.4	Destino: Estádio Castelão	52
5.5	Destinos: Rodoviária Central, Aeroporto e Mercado Central	55
5.6	Criando uma Superfície de Utilidade	56
5.6.1	Introdução	57
5.6.2	Krigagem: fundamentação teórica	58
5.6.3	Krigagem ordinária	61
5.6.4	Criação das superfícies de indiferença de acessibilidade	63
5.6.5	Resultados	64
6	CONSIDERAÇÕES FINAIS	70
A	GRÁFICOS	72
	REFERÊNCIAS	83
	REFERÊNCIAS	83

LISTA DE FIGURAS

1	Mapa de Fortaleza	44
2	Opinião dos entrevistados por destino	45
3	Parâmetros do variograma	60
4	Processo de interpolação: amostras e aproximação da superfície por uma grade regular	63
5	Superfície interpolada da utilidade estimada para o IJF	65
6	Superfície interpolada de utilidade estimada para a Rodoviária	66
7	Superfície interpolada de utilidade estimada para o Castelão	67
8	Superfície interpolada de utilidade estimada para o Aeroporto	68
9	Superfície interpolada de utilidade estimada para o Mercado	69
10	Destino IJF: Distribuição empírica dos efeitos marginais individuais	72
11	Destino Rodoviária de Fortaleza: Distribuição empírica dos efeitos marginais individuais	74
12	Destino Castelão: Distribuição empírica dos efeitos marginais individuais	76
13	Destino Aeroporto: Distribuição empírica dos efeitos marginais individuais	78
14	Destino Mercado Central: Distribuição empírica dos efeitos marginais individuais	80
15	Probabilidade Estimada para o Nível de Acesso ao Destinos pela Distância, mantendo as outras variáveis em suas médias	82

LISTA DE TABELAS

1	Relação entre componentes e as medidas de acessibilidade	26
2	Estatística descritiva das variáveis	46
3	Estatística descritiva para as variáveis dependentes (em %)	47
4	Modelo logit de escolha ordenada estimado via regressão paralela (destino: IJF)	48
5	Teste da hipótese de regressão paralela (destino: IJF)	49
6	Modelo logit de escolha ordenada estimado via regressão generalizada (destino: IJF)	50
7	Efeito Marginal Médio (destino: IJF)	51
8	Modelos Logit de Escolha Ordenado Estimado via Regressão Paralela (destino: Castelão)	53
9	Teste da Hipótese de Regressão Paralela (destino: Castelão)	53
10	Modelos Logit de Escolha Ordenado Estimado via Regressão Generalizada (destino: Castelão)	54
11	Efeito Marginal Médio (destino: Castelão)	54
12	Coeficientes estimados dos destinos Rodoviária, Aeroporto e Mercado via re- gressão paralela	55
13	Efeito marginal médio para os destinos Rodoviária, Aeroporto e Mercado . . .	56

1 INTRODUÇÃO

O processo de estruturação urbana ocorrido nas últimas décadas vem influenciando no cotidiano dos indivíduos à medida que estes necessitam locomover-se constantemente para realizarem as mais diversas funções e atividades, afim de adquirir bens e serviços que se encontram distribuídos por toda a cidade e são fundamentais no seu dia-a-dia. Dessa forma, a distribuição espacial de cada serviço pode facilitar ou dificultar esses deslocamentos intra-urbanos.

Nos últimos anos, em nosso país, estamos passando por um processo de crescimento econômico em todos os setores o que favorece, por exemplo, a qualidade de vida de milhares de habitantes. O aparecimento do PAC (Programa de Aceleração do Crescimento) na economia contribuiu para o aumento da infraestrutura no Brasil nas áreas de saneamento, de habitação, transporte, energia e recursos hídricos, entre outros. É notável em nosso cotidiano a crescente quantidade de novas habitações sendo construídas, bem como a modernização do transporte público de uma forma em geral. No entanto, esse crescimento quase sempre se processa de maneira desordenada.

O crescimento descontrolado das grandes cidades ocasionou a segregação socioeconômica da população no espaço urbano, sendo as camadas mais pobres deslocadas para as áreas periféricas, onde existem menos oportunidades e maiores custos relacionados ao deslocamento (Vieira e Haddad, 2012). Daí, a urbanização da população e o crescimento das cidades coloca a questão da acessibilidade como fator relevante na qualidade de vida da população.

Vale salientar que a rede de transporte público, atualmente, padece da tecnologia, do conforto e da facilidade (detalhes, veja Henrique (2004)). Sendo assim, é interessante analisar o cenário atual com foco no conforto e facilidade no deslocamento dos habitantes da sua origem ao seu destino. De uma maneira mais abrangente, analisar o bem-estar dos indivíduos a partir do momento que deixam suas residências ao ponto da execução da atividade torna-se interessante e necessário. Ou seja, por mais que o Brasil esteja passando pelo processo de modernização da rede de transporte, construção de postos de saúde, hospitais, lazer e etc, tais atividades que envolvem o uso desse serviço não necessariamente se localizam em pontos de fácil acesso tal que a população usufrua de maneira fácil e simples.

Com esse embasamento, o estudo do tema “acessibilidade” torna-se muito importante no cenário atual. Com efeito, a acessibilidade é muitas vezes confundida com mobilidade e algumas vezes acessibilidade é vista na perspectiva da facilidade de deslocamento de pessoas com necessidades especiais ou pelo uso de serviços públicos por elas mesmas. Porém, esta visão

é muito restritiva. Não há como definir acessibilidade de forma única. Cada autor a define de maneira diferente, dependendo do objetivo de estudo do trabalho. Para muitos, acessibilidade está ligada a distância do deslocamento que o indivíduo executa ao sair de sua residência ao local de realizar a sua atividade.

O conceito acessibilidade é muito flexível, por exemplo, podemos empregá-lo no planejamento do transporte urbano, cujo interesse seria a distribuição espacial do número de viagens de ônibus ofertadas pelo serviço público (Henrique, 2004); outro exemplo seria na área da saúde com ênfase na distribuição geográfica (Luo e Wang, 2003), o trabalho desenvolvido por Talen e Anselin (1998) faz uso de “medidas tradicionais” como o número de serviços dentro do setor censitário, modelo potencial (baseado no modelo gravitacional) ponderado pela distância, uma medida de custo de viagem e, por último, distância mínima entre origem e o destino considerado. Em seguida, o autor faz uso da metodologia da estatística espacial com o auxílio da correlação de Moran. Guagliardo (2004) criou duas superfícies baseadas na metodologia da densidade de Kernel via ArcGis, uma fornecendo a superfície dos serviços de saúde ofertados, e a outra, a distribuição populacional. A proposta do autor para a acessibilidade de cuidados básicos de saúde é criar uma nova superfície pela razão serviços/população. Podemos, também, utilizar acessibilidade na otimização do caminho de um ponto de origem ao ponto de destino (Ferreira e Martins, 2011). Dessa forma, a aplicabilidade do conceito é bastante vasta.

No presente estudo, buscamos definir acessibilidade numa perspectiva bastante geral, considera-se o tempo de percurso, tipo de transporte, conforto da jornada, condições das ruas, estradas etc que são incorporadas pelo indivíduo quando o mesmo está diante de uma estrutura de escolha. Ou seja, utilizaremos a metodologia da maximização da utilidade Ben-Akiva e Lerman (1978) para explicar o bem estar do indivíduo no processo de deslocamento ao sair de sua residência até a chegada ao destino. Portanto, não levaremos em consideração aspectos relacionados ao usufruto do serviço do destino propriamente dito, apenas características próprias dos indivíduos, por exemplo, uso de carro próprio no percurso.

Para este fim, este trabalho possuirá duas linhas de pesquisas que seguem como motivações. A primeira é explicar “acessibilidade” na perspectiva do comportamento do indivíduo com relação a uma estrutura de escolha ordenada, ou seja, através das características individuais (renda, educação, sexo e etc) e características do destino (distância). Buscamos entender como o indivíduo classifica a sua acessibilidade para cinco destinos de Fortaleza/CE, através de modelos logit ordenado. Foi utilizada uma base inédita coletada na cidade de Fortaleza/CE no

período de 2012, com um total de 4030 entrevistados em todos os bairros, envolvendo questões sobre análises socioeconômicas, criminológicas e de segurança pública (Carvalho, 2012). Os resultados alcançados mostram que, em geral, a população se sente desfavorecida/prejudicada com o processo de deslocamento de sua residência aos locais estratégicos (Hospital IJF, Rodoviária Central, Estádio Castelão, Aeroporto e Mercado Central) em Fortaleza. Para medir o nível de bem-estar do indivíduo, foi utilizada como *proxy* de utilidade a resposta sobre a satisfação desse processo de deslocamento.

Na segunda parte, utilizaremos o recurso da geoestatística para entender como se distribui a utilidade estimada na primeira parte, através de isoquantas de curva de utilidade. Dados, dois domicílios equidistantes do local de destino analisado, eles podem diferir quanto ao bem-estar. Ou seja, seja dois residentes A e B, o residente A utiliza transporte público e usufruindo de melhores condições das ruas e trânsito poderá estar em uma curva de bem-estar maior, do que o residente B, que possui carro, porém, enfrenta trânsito com maior fluxo de carros em seu trajeto. O interessante desta análise é evidenciar que alguns fatores não mensuráveis dos indivíduos podem suprir a questão da distância. Em linhas gerais, esses fatores não mensuráveis são incorporados pelo indivíduo quando este está diante de uma estrutura de escolha, isto é, quando o entrevistador pergunta ao entrevistado como ele classifica o acesso ao destino tal, o entrevistado não responde instantaneamente, ele visualiza/analisa por poucos segundos: onde o destino se localiza, qual o meio de transporte que irá utilizar, o nível de trânsito, faz-se avaliação da condição das vias e vários outros fatores. Para alcançar este objetivo, a interpolação por *krigagem* é fundamental na criação de uma superfície de utilidade.

Os resultados mostraram-se satisfatórios evidenciando que características dos indivíduos são fundamentais para suprir a questão da distância. A distância ainda continua como um fator fundamental na decisão do entrevistado para classificar o seu acesso ao destino. Porém encontramos, por exemplo (considerando o acesso ao Estádio Castelão) que homens possuem 12% de probabilidade a mais que as mulheres de escolher as opções Razoável ou Excelente. Além disso, níveis educacionais mais elevados tendem a estimular a escolha das categorias Razoável ou Excelente.

Outro resultado se refere a variável renda. Era de se esperar que quanto maior o nível de renda mais o indivíduo possuirá bens que facilitam seu deslocamento. Veremos que para todos os destinos, exceto aeroporto, o nível de renda provoca a escolha das opções Péssimo ou Razoável.

Com o auxílio da krigagem, veremos que em algumas regiões de Fortaleza existem locais mais afastados do destino que possuem um nível de utilidade igual aqueles mais próximos. Por exemplo, veremos que indivíduos residindo a uma distância de 8km do destino estarão na mesma curva de utilidade daqueles que residem em áreas mais próximas (no raio de 3 km) do destino. Aqueles indivíduos que se encontram equidistantes do destino podem diferir de sua utilidade (bem-estar), pois um dos residentes se encontra em uma área mais estratégica do que o outro.

Além deste capítulo introdutório, este trabalho está organizado em mais cinco capítulos. O Capítulo 2 apresentará conceitos e medidas de acessibilidade abordados pela literatura com foco nos trabalhos desenvolvido por [Geurs e van Eck \(2001\)](#) e [Geurs e van Wee \(2004\)](#) que servirão de fundamento aos conceitos de acessibilidade. No capítulo 3, apresentamos um *survey* de modelos logit clássico, ordenado e generalizado a fim de embasar a metodologia deste trabalho. Em seguida, no Capítulo 4, descreveremos a base de dados utilizada neste pesquisa, bem como as descritivas das variáveis. O Capítulo 5 é dedicado à apresentação dos principais achados empíricos de alguns destinos em Fortaleza/CE, bem como o resultado para a superfície da utilidade estimada com o objetivo de detectar o nível de acessibilidade. Finalmente, no Capítulo 6 discutimos algumas conclusões alcançadas.

2 CONCEITOS E MEDIDAS DE ACESSIBILIDADE

2.1 Introdução

De acordo com o dicionário Aurélio (Ferreira, 2001) a palavra “acessibilidade” indica qualidade ou caráter do que é acessível, do que tem acesso, de acesso fácil (coisa ou pessoa). O conceito de acessibilidade pode ser empregado em vários ramos do conhecimento: na área de planejamento urbano, cujo interesse seria a distribuição espacial do número de viagens de ônibus ofertadas pelo serviço público (Henrique, 2004); Na área da saúde, com ênfase em aspectos geográficos, acessibilidade poderia ser estudada como o grau de impedância (ou dificuldade) dos usuários em atingir os serviços médicos em uma dada localidade (Rosero-Bixby, 2004). Porém, muitas vezes acessibilidade é mal definida e medida erroneamente (Geurs e van Wee, 2004).

Por ser aplicado em diversas áreas do conhecimento, o conceito de acessibilidade possui formas distintas de mensuração. Por exemplo, Niemeier (1977) define acessibilidade como: “...a facilidade como que desejado destino pode ser atingido e é geralmente medido como uma função das oportunidades disponíveis moderado por alguma medida de impedância...”.

Não obstante, Burns (1979) a define como: “...a liberdade dos indivíduos em decidir se participa ou não em diferentes atividades. Restrições que limitam a liberdade do indivíduo, ou acessibilidade, são organizadas em três componentes: transporte, tempo e espacial.”

Por outro lado, Talen e Anselin (1998) utilizaram conceitos de equidade espacial e acessibilidade conjuntamente, e explicitaram que “...estes conceitos são construídos para avaliar a distribuição espacial ou padrão espacial dos serviços públicos (...). A análise de equidade espacial está preocupada com a comparação da distribuição locacional das facilidades ou serviços à distribuição de diferentes grupos socioeconômicos.”

Portanto, qualquer tentativa de definir o significado de acessibilidade deverá necessariamente abranger várias e diferentes perspectivas teóricas. No entanto, antes de prosseguirmos na busca de uma melhor conceituação de acessibilidade é importante discutirmos diferenças em relação ao conceito de mobilidade. O trabalho de Henrique (2004) conceitua as diferenças entre “mobilidade” e “acessibilidade” com ênfase no transporte público: “A mobilidade e a acessibilidade dos usuários do transporte público são fenômenos intimamente relacionados. O primeiro, que diz respeito às características dos indivíduos, refletindo a sua capacidade de efetuar deslocamento, é altamente influenciado pelo segundo, que retrata as condições de serviço

ofertadas pelo sistema”. Além disso, Lindemann *et al.* (1998) *apud* Henrique (2004) mencionam que “o conceito de mobilidade é associado e, muitas vezes, confundido com o conceito de acessibilidade, principalmente quando os dois estão relacionados tão somente com a facilidade de deslocamento”.

Definições mais abrangentes podem ser encontradas na literatura de Ferreira e Martins (2011). Para a população em geral, acessibilidade é vista como uma questão social, com foco em deficientes físicos, cujo acesso a meios de transportes, serviços públicos e ambientes físicos é limitado. Porém, esta visão é muito restritiva.

Geralmente, acessibilidade é estudada como a facilidade em atingir um dado destino. Tal facilidade é medida por função da distância, tempo de viagem, nível de impedância, entre outros, ou seja, características físicas mensuráveis no processo de deslocamento. No presente estudo buscamos definir acessibilidade numa perspectiva bastante geral, levando em consideração o tempo de percurso, tipo de transporte, conforto da jornada, condições das ruas, estradas etc, que o indivíduo incorpora na sua utilidade. Ou seja, utilizaremos uma metodologia baseada no paradigma econômico da maximização da utilidade para explicar o bem-estar do indivíduo no processo de deslocamento ao sair de sua residência até a chegada ao destino.

Seguindo esse raciocínio, segundo Geurs e van Wee (2004), o conceito de acessibilidade é bastante difícil e complexo, portanto, é fundamental conhecer os indicadores de acessibilidade elaborados por esses autores. Com esse intuito, revisaremos os grupos de indicadores na seção seguinte, baseados na classificação de Geurs e van Wee (2004).

2.2 A Classificação de Geurs e van Wee

Medidas de acessibilidade são baseadas em critérios teóricos, operacionais e interpretativos. Contudo, na prática, nem sempre uma medida de acessibilidade conterá todas estas informações, pois isso tornaria essa medida complexa. Assim, as vantagens e desvantagens de utilizar uma medida de acessibilidade devem obedecer os critérios que serão mencionados posteriormente. Seguindo de perto o desenvolvimento lógico de Geurs e van Eck (2001) e Geurs e van Wee (2004), encontramos quatro perspectivas diferentes para medidas de acessibilidade: **Medidas baseadas na Infraestrutura, Medidas baseadas na localização, Medidas baseadas no Indivíduo e Medidas baseadas na Utilidade.**

2.2.1 *Medidas baseadas na infraestrutura*

De acordo com [Geurs e van Eck \(2001\)](#), medidas baseadas na infraestrutura tais como tempo de jornada, congestionamento e velocidade de operação na rede rodoviária, desempenham um papel em políticas de transporte relacionadas à acessibilidade. Medidas de acessibilidade baseadas na infraestrutura fornecem informações valiosas sobre o nível de serviços de infraestrutura na área, mas falham em reconhecer que destinos de interesse podem estar longe daquela área. Segundo [Geurs e van Wee \(2004\)](#) as vantagens deste tipo de acessibilidade estão relacionadas ao critério de operacionalização, a necessidade de dados e modelos (transporte) estarem prontamente disponíveis e medidas são fáceis de interpretar pelos pesquisadores e *policy makers*. No entanto, ainda segundo os autores, estas medidas não satisfazem muitos dos critérios teóricos.

2.2.2 *Medidas baseadas na localização*

Estudaremos a seguir medidas baseadas na localização usadas para mensurar acessibilidade. Existem três formas de medir: Medidas de distância, Medidas de Contorno e Medidas Potenciais.

Medidas de distância

Relative accessibility ([Ingram, 1971](#)) é definido como o grau de conexão entre dois lugares na mesma área. O caso mais simples que conhecemos é a distância em linha reta entre dois pontos. No entanto, medidas de infraestrutura entre dois pontos como por exemplo, tempo médio de viagem e velocidade média podem ser interpretadas como *relative accessibility*.

Medidas de *Relative accessibility* são geralmente usadas no planejamento do uso do solo e estudos geográficos como medidas padrões para tempo máximo de viagem ou distância a uma dada localização e a infraestrutura de transporte; Por exemplo, cada habitante deve ser capaz de alcançar um hospital no intervalo de 30 minutos de percurso, ou cada habitante deve ter um ponto de parada de ônibus no intervalo de 500 metros com a origem em suas residências.

Medidas de contorno

Dado que a medida de distância considera apenas um destino, a medida de contorno vem preencher tal lacuna ao considerar mais de dois destinos a serem analisados. A medida de contorno, também chamada de “medida isocrômica”, “oportunidade acumulativa”, “contagem das proximidades” ou “acessibilidade diária”, tem por objetivo quantificar o número de oportunidades a serem alcançadas dentro de um intervalo de tempo ou distância. É de se esperar que quanto maior o número de oportunidades a serem alcançadas no intervalo de tempo ou distância maior será o índice de acessibilidade. Por exemplo, melhorias na infraestrutura que diminuam o tempo de viagem (facilidade de atingir os serviços) ou mais serviços a serem atingidos (mudanças na componente uso do solo). Podemos distinguir três tipos, segundo [Breheny \(1978\)](#) *apud* [Geurs e van Eck \(2001\)](#): a) Custos fixos: número de oportunidades acessíveis dentro de um limite de custo; b) oportunidades fixas: o tempo (médio ou total) ou custo requerido para acessar um número de oportunidades fixas; c) população fixa: a média (sobre a população) do número de oportunidades disponíveis dentro de vários custos fixos de viagem.

Alguns estudos modelaram acessibilidade como função acumulada das oportunidades que poderiam ser atingidas dentro de um intervalo de tempo, como é visto no trabalho de [Rosero-Bixby \(2004\)](#). O autor utiliza a proporção acumulada da população de acordo com a distância aos serviços de cuidados médicos em ambulatórios e hospitais mais próximos.

As vantagens das medidas de distância e contorno estão relacionadas aos critérios de operacionalização e interpretabilidade. Estas medidas não são exigentes de dados e são fáceis de interpretar pelos pesquisadores e *policy makers*. No entanto, estas medidas não satisfazem critérios teóricos ([Geurs e van Eck, 2001](#)).

Medidas potenciais

Também conhecida como Medidas *Gravity-Based*, representa a interação entre uma zona i e todos os serviços (ou outras zonas) n na qual é ajustada/ponderada pelo grau de impedância da distância ou tempo (veja exemplos em: [Guagliardo \(2004\)](#), [Talen e Anselin \(1998\)](#), [Luo e Wang \(2003\)](#) e [Liu e Zhu \(2004\)](#)). Para cada zona i o escore caracteriza a oferta de serviços, ou seja, quanto maior o escore maior a disponibilidade de oferta, por exemplo:

$$A_i = \sum_{j=1}^n D_j d_{ij}^{-\beta} \quad (1)$$

em que A_i é uma medida de acessibilidade na zona i (por exemplo, a residência ou a centróide¹ de um bairro ou área de interesse), D_j representa o número de serviços na zona j , d_{ij} é a distância entre i e j , e β é um coeficiente de declínio da gravidade, alguma vezes referido como o coeficiente de atrito do percurso. O parâmetro β representa a mudança na dificuldade do percurso como tempo de viagem ou distância, ou seja, a acessibilidade na zona i melhora se aumentarmos o número de serviços presente na própria zona ou em diminuir a dificuldade (ou impedância) em atingir a zona. De acordo com [Geurs e van Wee \(2004\)](#), para encontrar resultados plausíveis, a forma funcional desta função deverá ser escolhida com cuidado, e os parâmetros devem ser estimados usando dados empíricos recentes.

Uma maneira de generalizar medidas de acessibilidade potencial pode ser vista como:

$$A_i = \sum_j^n D_j F(c_{ij}), \quad (2)$$

tal que c_{ij} é o custo generalizado (distância ou tempo de viagem) entre i e j e $F(c_{ij})$ é a função de impedância. Na literatura (ver [Geurs e van Eck \(2001\)](#)) podemos encontrar várias formas para a função $F(c_{ij})$ que são usadas em estudos de acessibilidade:

- *Função poder negativa* ($F(c_{ij}) = c_{ij}^{-\beta}$);
- *Função exponencial negativa* ($F(c_{ij}) = e^{-\beta c_{ij}}$);
- *Versão modificada da função normal ou Gaussiana* ($F(c_{ij}) = 100 \cdot e^{-\frac{c_{ij}^2}{u}}$), onde u é uma constante;
- *Função logística modificada* ($F(c_{ij}) = 1 + e^{a+b \cdot \ln(c_{ij})}$), com a e b constantes.

A equação 2 é conhecida como “medida básica de acessibilidade potencial”. A função exponencial negativa (equação 3) é a mais usada na literatura por causa da proximidade com a teoria do comportamento da viagem ([Handy e Niemeier, 1997](#)). Uma ideia geral da forma funcional da medida potencial é mostrada a seguir:

$$A_i = \sum_{j=1}^n D_j e^{-\beta c_{ij}} \quad (3)$$

¹É o ponto localizado no centro geométrico

Segundo Geurs e van Wee (2004) a medida cobre algumas das lacunas teóricas da medida contorno: a medida avalia o efeito combinado dos elementos uso do solo e transporte, e incorpora hipóteses sob uma percepção do transporte da pessoa pelo uso de função decaimento da distância. As medidas são apropriadas como indicador social para analisar o nível de acesso a oportunidades econômicas e sociais para diferentes grupos socioeconômicos. Medidas potenciais têm vantagens práticas e eles podem ser facilmente computadas usando dados de uso do solo e transporte e/ou modelos tradicionalmente empregados como entradas para estimar medidas de infraestrutura. As desvantagens da medida potencial estão relacionadas à dificuldade de interpretação.

Outras desvantagens são encontradas na literatura como a auto-potencial (o número de oportunidades dentro da zona i ponderada pelo tempo médio ou distância naquela zona i), ou seja, quando $d_{ij} = 0$. Como já mencionado anteriormente, a medida potencial tem por característica avaliar a acessibilidade entre uma zona i e todas as outras zonas n . Contudo, essa medida não leva em conta as características dos indivíduos, ou seja, todos os indivíduos na mesma zona i têm o mesmo nível de acessibilidade. Por exemplo, se moradores de uma área estão próximos a um grande número de oportunidades de trabalho, mas eles não têm habilidade ou educação para se qualificar para tal emprego, eles terão ainda baixo nível de acessibilidade ao trabalho, pois são menos qualificados. Obviamente, se desagregarmos de acordo com as características dos indivíduos esse problema pode ser superado.

2.2.3 Medidas baseadas no indivíduo

Medidas baseadas na acessibilidade da pessoa são encontradas na literatura como *space-time geography* (ver, Miller (1991) e Miller (1999)). A componente temporal de acessibilidade envolve a disponibilidade de atividades em diferentes momentos do dia e o número de vezes em que o indivíduo participa daquela atividade específica. Na perspectiva *space-time geography*, as componentes de acessibilidade uso do solo e tempo são consideradas igualmente importantes. As medidas de acessibilidade do ponto de vista do indivíduo incorpora restrições espaciais e temporais, isto é, como atividades individuais podem ser alcançadas dadas restrições do tempo.

Segundo Geurs e van Eck (2001), medidas baseadas no indivíduo possuem ótimas vantagens teóricas: elas satisfazem talvez todos os critérios teóricos. Dessa forma os fatores de transporte, uso do solo, indivíduo, temporal influenciam a habilidade da pessoa em parti-

cipar de atividades desejadas ou necessárias, e portanto ao acesso a oportunidades. Ou seja, é uma metodologia desagregada que leva em conta características da pessoa. No entanto, essa desagregação torna-se a principal desvantagem desta medida, pois descreve as características dos indivíduos dada a sua restrição temporal. Assim, requer maiores quantidades de dados e esforço. A desvantagem mais forte está relacionada a operacionalização e interpretabilidade. Apesar das vantagens no GIS e modelagem espacial, operacionalização de medidas baseadas no indivíduo ainda enfrentam muitas dificuldades: detalhamento dos dados dos indivíduos sobre *activity-travel*, sua computação intensiva e algoritmos factíveis (Kwan, 1998).

2.2.4 Medidas baseadas na utilidade

As medidas de acessibilidade abordadas nas seções anteriores não são baseadas na teoria comportamental. O problema com estas medidas é a falta de uma relação entre escolha individual e acessibilidade (Niemeier, 1977). Todos os indivíduos de uma mesma região têm o mesmo nível de acessibilidade independentemente das alternativas de escolha da viagem (Ben-Akiva e Lerman, 1978). Ainda segundo o autor, a escolha da forma funcional é relativamente arbitrária e a medida de atração entre as regiões é geralmente o nível populacional ou emprego. Nesta seção descreveremos a classe de medidas de acessibilidade baseada no trabalho seminal de Ben-Akiva e Lerman (1978) e Ben-Akiva e Bierlaire (1999) que é consistente com a aplicação de modelos econômicos de utilidade aleatória para a decisão do indivíduo. A medida proposta é derivada da maximização da utilidade esperada que o consumidor obtém de uma dada escolha.

Estas medidas são usadas para analisar os benefícios econômicos que as pessoas derivam do acesso às atividades distribuídas espacialmente, por exemplo, o acesso a um shopping ou supermercado etc. Esta classe de indicadores de acessibilidade é uma outra alternativa para incluir características individuais (por exemplo: renda, sexo e idade) em modelos de acessibilidade. Esses indicadores tiveram suas origens na modelagem da demanda por viagem (Ben-Akiva e Lerman, 1978).

Por se tratar da perspectiva teórica a ser utilizado na nossa modelagem discutiremos mais detalhadamente o assunto na próxima seção.

2.3 Utilidade e Acessibilidade

As primeiras hipóteses desta abordagem são encontradas em [Koenig \(1980\)](#):

- a) Pessoas associam uma utilidade cardinal para cada alternativa que eles têm à disposição e escolhem a alternativa associada com a máxima utilidade adquirida.
- b) Como não é possível avaliar todos fatores, esta utilidade pode ser representada como a soma de uma componente não aleatória (determinística) e uma componente aleatória (estocástica).

O que se observa desta metodologia são as decisões que o indivíduo toma e que envolvem: escolhas do destino, escolha do modo de viagem, escolha do tempo de partida e escolha da rota ([Ben-Akiva e Bierlaire, 1999](#)). Ainda segundo os autores, a abordagem de utilidade aleatória pode ser vista como um tópico de modelos de escolha discreta.

Modelos de escolha discreta podem ser apresentados por um conjunto de definições e hipóteses ([Ben-Akiva e Bierlaire, 1999](#)). Em linhas gerais, um modelo de escolha discreta é composto por:

1. Tomador de Decisões: definindo a entidade de tomada de decisão e suas características. O tomador de decisão é assumido ser um indivíduo;
2. Alternativas: determinando as opções disponíveis ao tomador de decisões. O conjunto das alternativas é chamado de escolhas, por exemplo, a escolha do modo de transporte;
3. Atributos: medindo os benefícios e custos de uma alternativa ao tomador de decisões. Cada alternativa no conjunto de escolhas é caracterizada por um conjunto de atributos, por exemplo: carro, moto e ônibus são atributos do modo de transporte;
4. Regra de Decisão: descrevendo o processo usado pelo tomador de decisão para escolher a alternativa. A regra de decisão é o processo usado para avaliar o atributo das alternativas no conjunto de escolhas e determinar a escolha. Muitos modelos usados são baseados na teoria da utilidade esperada, na qual assume-se que a preferência do indivíduo (considerado neste trabalho como o tomador de decisão) para uma alternativa é capturado por um valor, chamado de utilidade, e o indivíduo seleciona a alternativa do conjunto com a maior utilidade. Modelos de utilidade aleatória, usados na econometria e em análise do

comportamento das viagens, são baseados em regras de decisão determinísticas, em que utilidades são representadas por variáveis aleatórias.

Modelos de utilidade aleatória assumem, em analogia a teoria do consumidor, que o tomador de decisão tem uma capacidade de discriminação perfeita (Ben-Akiva e Bierlaire, 1999). Em termos gerais, a utilidade que o indivíduo n associa à alternativa i no conjunto de escolhas C_n é dado por:

$$U_{in} = V_{in} + \varepsilon_{in}, \quad (4)$$

em que V_{in} é a parte determinística da utilidade, e ε_{in} é o termo aleatório, capturando a incerteza. A alternativa com a maior utilidade é então escolhida.

O termo determinístico V_{in} de cada alternativa é a função do atributo da própria alternativa e a característica do indivíduo. Isto é,

$$V_{in} = V(z_{in}, S_n)$$

z_{in} é o vetor de atributos percebido pelo indivíduo n para a alternativa i , e S_n é o vetor de características do indivíduo n . Isto pode ser simplificado usando uma nova função h dos atributos de z_{in} e S_n ,

$$x_{in} = h(z_{in}, S_n) \Rightarrow V_{in} = V(x_{in})$$

A escolha de $h()$ é geral, tomando uma função linear nos parâmetros, isto é,

$$V_{in} = \sum_k \beta_k x_{ink} \quad (5)$$

O termo determinístico da utilidade é portanto totalmente especificado pelo vetor de parâmetros β (Ben-Akiva e Bierlaire, 1999).

Podemos encontrar na literatura a equação 4 tomando outra perspectiva encontrada no trabalho de Koenig (1980). Neste caso, a utilidade U_{ij}^t associada por um indivíduo t , morando na região i , com destino na região j , pode ser representada por:

$$U_{ij}^t = V^t - C_{ij}^t \quad (6)$$

em que V^t é a utilidade bruta de atingir o destino j pelo indivíduo t (variável aleatória) e C_{ij}^t é custo gerado pela viagem ou tempo de percurso da região i a região j pelo indivíduo t (considerado pelo autor como preditivo e não-aleatório)

Como ressalta o autor, o problema é determinar uma função de probabilidade para a variável aleatória V^t , a fim de calcular o valor esperado da maximização da utilidade entre a utilidade U_{ij}^t associada a cada um dos destinos. Este valor esperado representa o benefício médio derivado a partir de uma viagem pelo indivíduo t e leva em conta ambas a desejabilidade dos destinos e o custo de viagem ou tempo requerido para atingi-los.

Encontramos na literatura vários métodos (detalhes veja [Ben-Akiva e Bierlaire \(1999\)](#), [Koenig \(1980\)](#), [Greene \(1951\)](#) e [Wooldridge \(2001\)](#)): modelos probit, modelos logit, modelos logit aninhado, modelo logit multinomial, modelos hixex, modelos logit aninhado cruzado, entre outros.

A fim de motivar a discussão, descreveremos a idéia da aplicação e escolha de modelos de escolha ordenado neste trabalho. A nossa principal evidência empírica que dará suporte a nossa análise é o nível de satisfação do entrevistado com relação a acessibilidade para alguns destinos (Estádio Castelão, Hospital Instituto José Frota, Aeroporto Pinto Martins, Rodoviária Central de Fortaleza e Mercado Central de Fortaleza) na cidade de Fortaleza/CE. Com efeito, o entrevistado é questionado sobre como o mesmo avalia seu acesso ao destino. Em seguida, escolhe uma entre as alternativas: Péssimo, Ruim, Razoável, Bom e Excelente. Nota-se que existe um *raking* nesta variável (ou seja, é de caráter ordinal), portanto, neste trabalho buscaremos, através de modelo logit ordenado, medir o efeito da acessibilidade baseado na metodologia proposto nesta seção. Descreveremos com mais detalhes sobre modelos de escolha ordenado no capítulo 3, bem como a existência de uma variável latente (não mensurável) que o indivíduo maximiza para escolher uma alternativa. A seleção das variáveis que podem influenciar as respostas do entrevistado foi baseada nas componentes descritas no trabalho de [Geurs e van Wee \(2004\)](#) e [Geurs e van Eck \(2001\)](#).

Com esse intuito, este trabalho emprega como *proxy* para a acessibilidade o nível de satisfação do entrevistado com relação a acessibilidade numa perspectiva de bem-estar do consumidor. Para este fim, o seguinte questionamento foi feito:

“Gostaria agora que o(a) Sr(a) pensasse sobre o que acontece quando é preciso ir a locais para resolver problemas ou aproveitar o tempo. Considerando que acessibilidade envolve tudo aquilo que é importante para ir e voltar (tempo de percurso,

custo de transporte, conforto da jornada, horário de funcionamento, condições das ruas, estradas, entre outros”.

Em linhas gerais, como aspectos individuais (renda, sexo, idade, educação, uso de carro próprio ou não) e aspectos de custos urbanos, como por exemplo, a distância, podem influenciar a probabilidade da escolha de resposta. Ou seja, interpretaremos essa resposta como o valor máximo de utilidade obtido pelo indivíduo após escolher de maneira ótimo os atributos do deslocamento. Em outras palavras, a teoria da utilidade aleatória propõe que os indivíduos escolhem alternativas que produzem a maior utilidade e, assim, a probabilidade de selecionar uma alternativa aumenta com o aumento da utilidade associada (exemplos dessa metodologia, veja Niemeier (1977), Junior (2010), Burton *et al.* (2001)).

Descreveremos na seção seguinte as componentes teóricas que constituem as medidas de acessibilidade que servirão como base para a estimação do modelo proposto.

2.4 Componentes de Acessibilidade

A seção atual descreverá cada componente de acessibilidade para melhor entendimento do assunto. Cada uma das medidas de acessibilidade, como citado anteriormente, é constituída de quatro componentes de acessibilidade: transporte (*transportation*), uso do solo (*land-use*), tempo (*temporal*), indivíduo (*individual*).

1. Componente transporte reflete a distância, o tempo de viagem, custo e esforço para viajar entre uma origem e local de destino.
2. Componente uso do solo consiste em: (a) na magnitude, qualidade e na característica da distribuição (oferta) espacial das atividades de cada destino (por exemplo, trabalho, escolas, shopping, hospitais) (b) e a demanda por estas oportunidades (por exemplo, trabalhadores, onde habitantes moram), (c) a confrontação entre oferta e demanda em que resulta em atividades competitivas gerando restrições tais como vagas de emprego e escolas e leitos nos hospitais.
3. Componente tempo reflete: (a) a disponibilidade da atividade em diferentes tempos do dia ou semana (exemplo horas abertas no shopping), e (b) o tempo disponível para indivíduos participar de certas atividades (exemplo trabalho, lazer).

4. Componente individual reflete a necessidade (dependendo da idade, riqueza, nível educacional, situação da habitação, etc), habilidades (dependendo das condições físicas das pessoas, disponibilidade de modos de viagens e etc) e oportunidades (depende da riqueza da pessoa, restrições orçamentária, nível educacional, etc) dos indivíduos. Estas características influenciam o acesso ao modo de transporte (por exemplo, ser capaz de dirigir e usar um carro) e oportunidades distribuídas espacialmente (por exemplo, ter alguma habilidade ou educação para se qualificar em empregos perto de sua residência) e pode influenciar fortemente a acessibilidade agregada total.

A relação entre as componentes e as medidas de acessibilidade é vista na tabela 1 (Geurs e van Wee, 2004).

Tabela 1: Relação entre componentes e as medidas de acessibilidade

	Componente Transporte	Componente Uso do Solo	Componente Tempo	Componente Individual
Medidas baseadas na infraestrutura	Velocidade do percurso; perda e horas em congestionamento	-	Período de horário de pico; período de 24 hrs	Viagem baseada em estratificação, e.x, casa ao trabalho, negócios
Medidas baseadas na localização	Tempo de viagem e/ou custos entre localização de atividades	Quantidade e distribuição espacial da demanda e/ou oferta de oportunidade	Tempo de viagem e custos podem diferir, e.x, entre horas dos dias, entre dias da semana, ou estações	Estratificação da população (e.x por renda, nível educacional)
Medidas baseadas no indivíduo	Tempo de viagem entre localização das atividades	Quantidade e distribuição espacial das oportunidades ofertadas	Retrição temporal por atividades e tempo disponível para atividade	Acessibilidade é analisada em nível individual
Medidas baseadas na Utilidade	Custo da viagem entre localização das atividades	Quantidade e distribuição espacial das oportunidades ofertadas	Tempo e custo da viagem podem diferir, e.x., entre horas do dias, entre dias da semana	Utilidade é derivada no indivíduo ou população homogênea

Fonte: Tabela extraída de Geurs e van Wee (2004)

Uma medida de acessibilidade deve considerar todas as componentes mencionadas anteriormente, segundo Geurs e van Wee (2004). Porém, na prática isso não é fácil. Caso a medida adotada contenha toda essa informação, resultará complexa e de difícil interpretação. Portanto, nem toda medida de acessibilidade conterá todas essas informações. Por exemplo, a medida de infraestrutura não inclui uma componente de uso do solo, isto é, tal medida não é sensível a mudanças na distribuição espacial das atividades. Medidas baseadas na utilidade e indivíduo geralmente focam na componente individual, analisando acessibilidade em um nível micro². A seguir abordaremos com detalhes os componentes de acessibilidade com base nos

²Medidas baseadas na utilidade são sensíveis a ambos componentes uso do solo e transporte, pois, a acessibilidade em uma localização pode melhorar com o aumento de atividades, redução do tempo de viagem, ou ambos.

trabalhos de [Geurs e van Eck \(2001\)](#) e [Geurs e van Wee \(2004\)](#).

2.4.1 Transporte

Descreve o sistema de transporte em si, e possui três elementos: a) a oferta da infraestrutura, sua localização e características (por exemplo, velocidade máxima de viagem, número de linhas, horários do transporte público, custo de viagem, tempo de viagem, nível de conforto), b) a demanda por passageiros e viagens de carga, c) choque entre oferta da infraestrutura e demanda por viagem, resultando na distribuição de tráfego, e tempo de viagem, custo e esforço em alcançar um destino.

As facilidades de deslocamento do indivíduo do ponto de origem ao destino, podem se dividir em três elementos: a) o valor do tempo de viagem, b) custo da viagem e c) esforço da viagem. Por exemplo, o tempo de viagem para uma típica viagem de carro tem os seguintes elementos de tempo: caminhar ao local de estacionamento, tempo de viagem de carro (que por sua vez depende da distância, limite de velocidade, nível de congestionamento), tempo de tráfego, encontrar um estacionamento próximo ao local e caminhar ao destino. O total de custo da viagem de carro consiste de custo fixo (por exemplo, licença de motorista, compra do carro, seguro) e custo variável (combustível e manutenção). O esforço da viagem, caracterizado pelo conforto, confiança, nível de estresse e risco de acidente. Claramente, tais medidas de esforço são difíceis de ser mensurados, por exemplo, nível de conforto e stress.

De acordo com [Geurs e van Eck \(2001\)](#) “a interação entre duas localidades diminui com o aumento da desutilidade (distância, tempo e custo) entre elas”. A perspectiva de distância entre origem e destino difere de acordo com:

- i)** Modo de transporte: carro, transporte público, modos não motorizados;
- ii)** Propósito de viagem: casa-trabalho;
- iii)** Características da família: riqueza, nível educacional;
- iv)** Características do destino: exclusividade e atratividade (estabelecimentos comerciais grandes no centro da cidade têm um caimento diferente da distancia que supermercados)

2.4.2 Uso do solo

A princípio pelo próprio uso do nome podemos pensar que o objetivo deste componente seria o de incorporar a distribuição espacial das oportunidades (ou serviços) influenciando o nível de acessibilidade. Pelo princípio da equidade pode-se acreditar que, por exemplo, se todos os trabalhos e residências são igualmente distribuídos sob um determinado espaço (ou certa área), conclui-se que todos os habitantes possuem o mesmo nível de acessibilidade ao trabalho. De forma contrária, caso todos os locais de trabalhos estejam de certa forma aglomerados nas áreas centrais da cidade, então os moradores que residem próximo a esta área são mais privilegiados e, portanto possuem maiores níveis de acessibilidade ao trabalho do que pessoas que residem em áreas mais periféricas (ou afastada da área central) das cidades, que por sua vez terá menor nível de acessibilidade.

Com base na informação citada anteriormente a componente uso do solo pode se dividir em dois elementos: a) distribuição espacial ofertada de cada destino e suas características, tais como: o local de trabalho, escolas, hospitais, estádios e seus atrativos, capacidade, qualidade, etc. e b) a distribuição espacial da demanda por estes serviços e suas características, tais como localização das residências e seus habitantes.

A competição entre oferta e demanda é um aspecto relevante para análise de acessibilidade, por exemplo, a população demanda serviços tais como trabalho, escolas e hospitais. Porém, tais serviços são limitados, ou seja, a distribuição espacial da demanda por serviços também influencia a acessibilidade, e também fica claro que a distribuição das oportunidades (lado da oferta) também influencia o nível de acessibilidade.

2.4.3 Tempo

Restringe-se aos efeitos de tempo, isto é, a) a disponibilidade das atividades em diferentes tempos do dia ou semana, anos, etc. e b) tempo disponível para indivíduos participar em certas atividades (por exemplo, trabalho e lazer). De acordo com [Geurs e van Eck \(2001\)](#) “As componentes tempo e uso do solo são interdependentes porque indivíduos podem somente estar em um local em um dado tempo e a viagem consome tempo”.

2.4.4 Individual

Para finalizar, o componente individual desempenha um papel importante no quesito nível de acessibilidade social e econômico. Em estudos psicológicos, como cita [Vlek e Steg \(1996\) apud Geurs e van Eck \(2001\)](#), três características são abordadas na mobilidade dos passageiros: Necessidades, Habilidades e Oportunidades. A seguir descrevemos tais características, de acordo com [Geurs e van Eck \(2001\)](#):

- i) Necessidades que pessoas têm para viajar e acesso a oportunidades dependem de tais características como idade, riqueza, nível educacional, situação familiar. Por exemplo, familiar com crianças terão uma necessidade para acesso às escolas, pessoas com maiores níveis educacionais terão necessidade de uma oportunidade de trabalho específico, pessoas idosas terão maior demanda por acesso a cuidados médicos, etc.;
- ii) Habilidades de pessoas estão relacionadas ao nível de capacidade física (por exemplo, cognitiva, sensorial, intelectual ou necessidades especiais) e habilidades específicas necessárias para acessar um modo de transporte (exemplo, qualificação a dirigir um carro). Por exemplo, pessoas sem carteira de motorista somente terão acesso a oportunidades de trabalho que podem ser alcançadas por transporte público ou modos não motorizados, e pessoas com deficiência física precisam de serviços específicos para usar transporte público;
- iii) Oportunidades de pessoas estão relacionadas a renda e restrição orçamentária, por exemplo, 80% do grupo de rendimentos mais baixos (menos que 15,000 Euros por ano) na Holanda não possuem um carro, e portanto, dependem do serviço público e modos não motorizados de transporte;

Para mais detalhes sobre a relação entre medidas e componentes de acessibilidade veja a figura 2.1 que encontra-se no trabalho de [Geurs e van Wee \(2004\)](#).

Revisaremos a seguir conceitos sobre os diferentes tipos de critérios que podemos utilizar para classificar as medidas de acessibilidade. Tais critérios são importantes pois podemos avaliar a utilidade e as limitações de cada medida na literatura. De acordo com [Geurs e van Wee \(2004\)](#) classificaremos as medidas através de: (1) critérios teóricos; (2) operacionalização (ou execução) e (3) interpretabilidade;

2.5 Critérios para Medidas de Acessibilidade

Os **critérios teóricos** são basicamente mencionados na seção 2.4. Portanto, uma medida de acessibilidade deve ser flexível a mudanças no sistema de transporte (componente transporte), ao sistema de uso do solo (componente uso do solo), a restrições temporais das atividades (componente tempo) e as necessidades, habilidades e oportunidades dos indivíduos (componente indivíduo).

O critério da **operacionalização** é a facilidade da execução da medida criada, por exemplo, disponibilidade de dados, modelos e técnicas, e tempo e orçamento.

Com base na proposta da medida, pesquisadores, planejadores e *policy makers* deverão entendê-la e interpretá-la, caso contrário tal medida não terá aplicabilidade na avaliação de estudos ou desenvolvimento político - critério da **Interpretabilidade**.

Após essa discussão conceitual sobre acessibilidade, abordaremos de que maneira esse conceito pode ser medido e operacionalizado. Através dos modelos de escolha discreta ordenada.

3 SURVEY DE MODELOS DE ESCOLHA DISCRETA

3.1 Introdução

Em muitos trabalhos econômicos a variável de interesse, Y , não é contínua, mas sim discreta, tomando valores finitos. Por exemplo, escolher consumir entre os produtos A ou B, nível de satisfação da saúde na escala de 0 a 10, aceita-se ou não a oferta de emprego, se trabalha ou não e etc. Dessa forma a variável de interesse, Y , pode classificar-se nos modelos: modelos binários, modelos politômicos ou modelos ordinais. Para cada possibilidade o interesse está em modelar a probabilidade do evento em questão em função das variáveis explanatórias (ou simplesmente: independentes).

3.2 Modelos de Escolha Binária

Para este modelo a variável dependente assume os valores 0 e 1, onde $Y = 1$ indica a presença do atributo e $Y = 0$ caso contrário. Por exemplo, $Y = 1$ a pessoa está empregada, $Y = 0$ caso contrário; $Y = 1$ se a pessoa está doente, $Y = 0$ caso contrário. Na literatura, geralmente, encontramos $Y = 1$ para o sucesso e $Y = 0$ para o fracasso. Ou seja, o interesse consiste em modelar a probabilidade de sucesso ($Y = 1$). Dessa maneira considere que a esperança não condicional da variável binária Y é por definição uma probabilidade:

$$E(Y) = P(Y = 1)$$

além disso, seja um conjunto de variáveis independentes $\mathbf{x} \equiv (x_1, x_2, \dots, x_k)$ influenciando Y , podemos, então, reescrever a equação na forma condicional como:

$$\mathbb{E}(Y|\mathbf{x}) = \mathbb{P}(Y = 1|\mathbf{x})$$

levando em consideração a abordagem de análise de regressão, então podemos escrever qualquer função de regressão da seguinte forma:

$$\mathbf{y} = H(\mathbf{x}, \boldsymbol{\beta}) + \epsilon$$

onde $H()$ é uma função qualquer de \mathbf{x} e $\boldsymbol{\beta}$, aplicando a esperança condicional (supondo, inicialmente, $\mathbb{E}(\epsilon|\mathbf{x}) = 0$), têm-se:

$$\mathbb{E}(\mathbf{y}|\mathbf{x}) = H(\mathbf{x}, \boldsymbol{\beta}) + \mathbb{E}(\epsilon|\mathbf{x}) = H(\mathbf{x}, \boldsymbol{\beta})$$

Note que a função $H(\mathbf{x}, \boldsymbol{\beta})$ é nada mais que a representação da esperança condicional de \mathbf{y} dado \mathbf{x} . Portanto, se a variável dependente é binária, então a função $H()$ estará diretamente ligada a probabilidade condicional de ocorrência para $Y = 1$. Assim, modelos de escolha binária dependerão da forma de especificar a função de regressão $H(\mathbf{x}, \boldsymbol{\beta})$ (Duncan, 2001). Aplicaremos a seguir a metodologia de mínimos quadrados para Modelos Lineares de Probabilidade.

3.3 Modelos Lineares de Probabilidade

Considere \mathbf{y} e $\mathbf{x}_{1 \times k}$, como sendo os vetores de variáveis dependentes e independentes, respectivamente. Sabendo que o modelo linear possui a forma:

$$\mathbb{P}(\mathbf{y}|\mathbf{x}) = H(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{x}\boldsymbol{\beta} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k \quad (7)$$

onde $\beta_{k \times 1}$ são parâmetros desconhecidos. Considerando o erro no modelo, reescrevemos a equação 7 da seguinte forma:

$$\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \epsilon$$

por suposição, $f(\epsilon)$ é a densidade com $\mathbb{E}(\epsilon|\mathbf{x}) = 0$. Como a função é linear podemos aplicar mínimos quadrados ordinários (MQO), produzindo estimativas não viesadas e consistentes.

Porém, existem alguns razões para não aplicar o método MQO (Greene, 1951), como:

1. O erro não possui distribuição normal:

De acordo com o modelo $\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \epsilon$, o erro teria apenas dois valores:

$$\epsilon_i = 1 - \mathbf{x}_i\boldsymbol{\beta} \text{ com probabilidade } f(Y = 1)$$

$\epsilon_i = -\mathbf{x}_i\beta$ com probabilidade $1 - f(Y = 1)$

2. o erro é heterocedástico:

$$\text{Var}(\epsilon_i|\mathbf{x}_i) = \mathbb{E}(\epsilon_i^2|\mathbf{x}_i) = (-\mathbf{x}_i\beta)^2 \cdot [1 - f(Y = 1)] + (\mathbf{x}_i\beta)^2 \cdot [f(Y = 1)] = \mathbb{P}(y_i = 1|\mathbf{x}_i) \cdot \mathbb{P}(y_i = 0|\mathbf{x}_i) = \mathbf{x}_i\beta(1 - \mathbf{x}_i\beta)$$

3. A esperança não estará, necessariamente, entre 0 e 1:

$$\mathbb{E}(y_i|\mathbf{x}_i) = \mathbb{P}(y_i = 1|\mathbf{x}_i) = \mathbf{x}_i\beta$$

Ou seja, não é possível assumir a hipótese de homocedasticidade e de normalidade. Além disso, como o objetivo de estudo é modelar valores discretos, e portanto, as probabilidades, o método de MQO ignora essa característica da variável, Y . De acordo com (Wooldridge, 2001) para alguns valores de x_1, x_2, \dots, x_k as estimativas de \hat{y} poderão estar fora do intervalo unitário. Mesmo assim, ainda segundo o autor, MLP pode ser visto como uma aproximação conveniente. Para aprimorar tais problemas em utilizar MQO, existe uma outra metodologia baseada na distribuição acumulada como será visto na seção 3.4.

3.4 Modelos de Resposta Binária: Probit e Logit

Como vimos na seção 3.3 podemos escrever o valor esperado de Y como:

$$\mathbb{E}(Y|\mathbf{x}) = \mathbb{P}(Y = 1|\mathbf{x}) = H(\mathbf{x}, \beta)$$

Utilizando a metodologia do modelo linear encontramos problemas como não normalidade, heterocedasticidade e $\mathbb{E}(Y|\mathbf{x})$ os quais não impedem a aplicação do MQO. Para contornar esse problema aplicaremos transformações em $\mathbf{x}\beta$ de tal forma que o resultado esteja no intervalo $[0,1]$. Em muitas aplicações $H()$ é uma função de distribuição acumulada.

3.4.1 Transformação em $\mathbf{x}\beta$

Para a transformação Probit, façamos:

$$H(\mathbf{x}, \beta) = \Phi(\mathbf{x}\beta)$$

onde $\phi(z) = (2\pi)^{-\frac{1}{2}} \cdot \exp(-\frac{z^2}{2})$ e $\Phi(z) = \int_{-\infty}^z \phi(t)dt$ representam a densidade de probabilidade e distribuição acumulada da normal padrão, respectivamente.

Para a transformação Logit, façamos:

$$H(\mathbf{x}, \beta) = \Lambda(\mathbf{x}\beta)$$

onde $\Lambda(z) = \frac{\exp(z)}{1 + \exp(z)} = \frac{1}{1 + \exp(-z)}$ representa a distribuição acumulada da função logística.

Como as funções $\Phi(z)$ e $\Lambda(z)$ são monótonas crescentes, tem-se:

$$\begin{aligned} H(\mathbf{x}\beta) &\rightarrow 0 \quad \text{quando } \mathbf{x}\beta \rightarrow -\infty; \\ H(\mathbf{x}\beta) &\rightarrow 1 \quad \text{quando } \mathbf{x}\beta \rightarrow \infty \end{aligned}$$

Qual distribuição usar é uma questão natural. A distribuição logística é similar a normal, exceto na calda, que são mais pesadas (parecida com a distribuição t). Portanto, para valores no intervalo entre -1.2 e +1.2, as duas distribuições tendem a fornecer probabilidades similares [Greene \(1951\)](#). Ainda segundo Greene, a distribuição logística tende a fornecer probabilidade maior a $Y = 0$ quando $\mathbf{x}\beta$ é extremamente pequeno que a distribuição normal. Existem razões práticas para favorecer um ou outra em alguns casos por conveniência matemática, mas é difícil justificar a escolha de uma distribuição ou outra nos fundamentos teóricos. Por outro lado, utilizar a distribuição normal pode ser motivada pelo recurso do teorema do limite central. A distribuição logística tem provado ser uma função útil para modelar por várias décadas ([Greene, 1951](#)).

Independente da forma funcional usada, é importante notar que os parâmetros do modelo, não são necessariamente os efeitos marginais como no modelo linear clássico. Portanto,

$$\begin{aligned} \frac{\partial \mathbb{E}[\mathbf{y}|\mathbf{x}]}{\partial \mathbf{x}} &= \left[\frac{dH(\mathbf{x}\beta)}{d(\mathbf{x}\beta)} \right] \beta \\ &= h(\mathbf{x}\beta)\beta \end{aligned} \tag{8}$$

onde $h(\cdot)$ é a função de densidade da correspondente função acumulada definida anteriormente.

3.4.2 Utilidade para modelos de escolha binária

Na economia, tratar o modelo de escolha binária como variável latente é comum nesses estudos. Dessa forma, uma outra alternativa metodológica para interpretar estes modelos binários é derivar a estrutura baseando no comportamento do indivíduo, ou seja, existe uma variável latente, que não podemos mensurá-la, y^* , cujo intervalo é $(-\infty, \infty)$. Essa variável latente é não observada, cuja representação matemática é:

$$\mathbf{y}^* = \mathbf{x}\beta + \epsilon \tag{9}$$

Assume-se que ϵ representa uma componente estocástica com densidade simétrica $f(\cdot)$ sob zero e função acumulada $F(\cdot)$. Tal função acumulada por ser a distribuição normal ou logística. Então, para categorizar a variável latente, determinados pontos de cortes que definem a ocorrência dos dois valores possíveis de y :

$$\begin{aligned}
 y &= 1 \quad \text{se } y^* > 0 \\
 y &= 0 \quad \text{se } y^* \leq 0
 \end{aligned}$$

É fácil notar o significado do termo variável latente para y^* , pois:

$$\begin{aligned}
 \mathbb{E}(\mathbf{Y}|\mathbf{x}) &= \mathbb{P}(Y = 1|\mathbf{x}) \\
 &= \mathbb{P}(y^* > 0|\mathbf{x}) \\
 &= \mathbb{P}(\mathbf{x}\boldsymbol{\beta} + \epsilon > 0|\mathbf{x}) \\
 &= \mathbb{P}(\epsilon > -\mathbf{x}\boldsymbol{\beta}|\mathbf{x}) \\
 &= 1 - F(-\mathbf{x}\boldsymbol{\beta}) \\
 &= F(\mathbf{x}\boldsymbol{\beta})
 \end{aligned}$$

Se ϵ é uma distribuição normal padrão, então teremos o modelo Probit, enquanto se ϵ for logística, teremos o modelo logit. Assumi-se, por conveniência, que para o modelo probit $\sigma = 1$ e para o modelo logit $\sigma = \frac{\pi}{\sqrt{3}}$.

É importante saber como interpretar β_j . Para simplificar o entendimento supomos que x_j seja contínua, o efeito fica:

$$\frac{\partial \mathbb{P}(Y = 1|\mathbf{x})}{\partial x_j} = f(\mathbf{x}\boldsymbol{\beta})\beta_j \quad (10)$$

Portanto, o efeito parcial de x_j em $\mathbb{P}(Y = 1|\mathbf{x})$ depende de \mathbf{x} através de $f(\mathbf{x}\boldsymbol{\beta})$. Se $F()$ é estritamente crescente, como caso do probit e logit, $f(z) > 0$ para todo z . Portanto, o sinal do efeito é dado pelo sinal de β_j (Wooldridge, 2001).

3.4.3 Fundamentos teóricos para variável latente

Para melhor entendimento desta abordagem, seja Y representando a escolha na participação da força do trabalho ($Y = 1$ se a mulher trabalha, 0 caso contrário) e sejam dois resultados, trabalhar ou não trabalhar, descrito pela utilidade U_y^* ;

$$\begin{aligned}
 U_{y=1}^* &= \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \\
 U_{y=0}^* &= \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0
 \end{aligned}$$

onde \mathbf{x} representa um conjunto de variáveis independentes comum, $\boldsymbol{\beta}_0$ e $\boldsymbol{\beta}_1$ são vetores de

paramêtros desconhecidos e ϵ_0 e ϵ_1 representa componentes não observados. Dadas estas informações, um indivíduo escolheria participar no mercado de trabalho se a utilidade de trabalhar $U_{y=1}^*$ exceder a utilidade de ficar sem trabalho $U_{y=0}^*$. Ou seja, a escolha para participar é feita quando $U_{y=1}^* \geq U_{y=0}^*$. Definindo $\mathbf{1}(\cdot)$ como a função indicador, temos que:

$$\begin{aligned}
 y &= \mathbf{1}(U_{y=1}^* \geq U_{y=0}^*) \\
 &= \mathbf{1}(\mathbf{x}\beta_1 + \epsilon_1 \geq \mathbf{x}\beta_0 + \epsilon_0) \\
 &= \mathbf{1}(\epsilon_1 - \epsilon_0 \geq -\mathbf{x}(\beta_1 - \beta_0)) \\
 &= \mathbf{1}(\mathbf{x}(\beta_1 - \beta_0) + (\epsilon_1 - \epsilon_0) \geq 0) \\
 &= \mathbf{1}(\mathbf{x}\beta + \epsilon \geq 0)
 \end{aligned}$$

Em outras palavras, a metodologia de variável latente para modelo de escolha binária pode ser derivada a partir de um modelo econômico. Isto é bastante conveniente já que a nossa perspectiva teórica para entender acessibilidade será baseada em medidas baseadas na utilidade como descrito na seção 2.2.4.

3.5 Modelos de Escolha Ordenada

Existem algumas aplicações onde o modelo de escolha binária é inapropriado. Por exemplo, na área da saúde poderíamos analisar se o indivíduo está satisfeito ou não com o serviço de saúde pública. Para tal aplicação, o modelo escolhido seria o mesmo abordado na seção 3.4. No entanto, em outros casos o objetivo a ser estudado seria o nível de satisfação, cujas categorias poderiam ser classificadas como: (1) Péssimo, (2) Ruim, (3) Razoável, (4) Bom e (5) Excelente. Claramente observamos que a variável dependente, Y , possui mais de duas respostas possíveis e essas respostas apresentam um ordenamento. A aplicação de modelos ordenados probit e logit para este problema permite que a variável dependente possua um ordenamento natural (*ranking*). Logo, o objetivo é modelar estas respostas ordenadas como função de outras variáveis explicativas.

Este modelo foi proposto inicialmente por [Aitchison e Silvey \(1957\)](#) que utilizaram um modelo probit ordenado para analisar experimentos nos quais a resposta dos sujeitos a vários níveis de estímulos eram divididas em classes ranqueadas. [Snell \(1964\)](#) sugeriu o uso do modelo logístico no lugar do modelo de distribuição normal por uma questão de simplificação matemática. A primeira aplicação do modelo probit nas ciências sociais foi desenvolvida por [McKelvey e Zavoina \(1975\)](#) que generalizou o modelo de Aitchison e Silvey para mais de uma variável dependente.

[McCullagh \(1980\)](#) desenvolveu independentemente o chamado modelo cumulativo na literatura estatística. Ele modelou diretamente as probabilidades cumulativas da resposta ordenada como uma transformação monotônica crescente de um preditor linear dentro de um

intervalo unitário, assumindo uma ligação logit ou probit. Esta especificação resulta na mesma função de probabilidade do modelo de McKelvey e Zavonaina, sendo, portanto observacionalmente equivalente.

Formalmente, seja a variável dependente, Y , com J possibilidades com ordenamento natural (ou seja, a escolha $j + 1$ é melhor que j). Assim Y toma os valores $1, 2, \dots, J$. O modelo de escolha ordenado para Y pode ser derivado a partir de um modelo de variável latente (ou economicamente, a utilidade y^*). Assume-se que a variável latente y^* é determinada por:

$$y^* = \mathbf{x}\boldsymbol{\beta} + \epsilon$$

seja \mathbf{x} o conjunto de covariáveis (nessa abordagem \mathbf{x} não conterá uma constante) cuja dimensão seja $1 \times k$ e considere $\boldsymbol{\beta}$ como vetor de parâmetros desconhecidos $k \times 1$. As J escolhas se relacionam com a variável latente da seguinte forma:

$$\begin{aligned} y = 1 & \text{ se } \mu_0 \leq y^* \leq \mu_1 \\ y = 2 & \text{ se } \mu_1 < y^* \leq \mu_2 \\ & \vdots \\ y = J & \text{ se } y^* > \mu_{J-1} \end{aligned}$$

ou simplesmente,

$$y = j \text{ se } \mu_{j-1} \leq y^* \leq \mu_j \text{ para } j = 1, 2, \dots, J$$

onde $\mu_0 < \mu_1 < \dots < \mu_J$ são os pontos de cortes desconhecidos (ou *threshold*), com $\mu_0 = -\infty$ e $\mu_J = \infty$. Então, a probabilidade condicional de observar a j -ésima categoria pode ser escrita como:

$$\begin{aligned} \mathbb{P}(y = j|\mathbf{x}) &= \mathbb{P}(\mu_{j-1} \leq y^* \leq \mu_j) \\ &= \mathbb{P}(\mu_{j-1} \leq \mathbf{x}\boldsymbol{\beta} + \epsilon \leq \mu_j) \end{aligned}$$

rearranjando os termos, tem-se

$$\begin{aligned} \mathbb{P}(y = j|\mathbf{x}) &= \mathbb{P}(\mu_{j-1} - \mathbf{x}\boldsymbol{\beta} \leq \epsilon \leq \mu_j - \mathbf{x}\boldsymbol{\beta}) \\ &= \mathbb{P}(\epsilon \leq \mu_j - \mathbf{x}\boldsymbol{\beta}) - \mathbb{P}(\epsilon \leq \mu_{j-1} - \mathbf{x}\boldsymbol{\beta}) \end{aligned}$$

vemos que o modelo descreve a probabilidade do resultado. Ou seja, não descreve a relação entre y e as covariáveis \mathbf{x} , portanto não há relação de regressão óbvia entre a variável depen-

dente observada, y , e as covariáveis \mathbf{x} . No entanto, em geral, algumas normalizações são necessárias para identificar os parâmetros do modelo. Isto leva a questão da interpretação de β (Greene e Hensher, 2009).

Devemos supor alguma distribuição para ϵ . A literatura recomenda a escolha da distribuição normal padrão (modelo probit) ou a distribuição logística (média zero e variância $\pi^2/3$), que produz o modelo ordenado logit. Historicamente, a preferência para a distribuição logística tem sido baseada na conveniência matemática e o fato do “*Odds Ratios*” em uma forma fechada (Greene e Hensher, 2009). Tomando como exemplo a distribuição normal, a probabilidade condicional poderá ser derivada:

$$\mathbb{P}(y = j|\mathbf{x}) = \Phi(\mu_j - \mathbf{x}\beta) - \Phi(\mu_{j-1} - \mathbf{x}\beta) \quad (11)$$

Para derivar os efeitos marginais dos regressores \mathbf{x} vamos supor que existam apenas três categorias. Assim, temos:

$$\begin{aligned} \mathbb{P}(y = 1|\mathbf{x}) &= \Phi(\mu_1 - \mathbf{x}\beta) \\ \mathbb{P}(y = 2|\mathbf{x}) &= \Phi(\mu_2 - \mathbf{x}\beta) - \Phi(\mu_1 - \mathbf{x}\beta) \\ \mathbb{P}(y = 3|\mathbf{x}) &= 1 - \Phi(\mu_2 - \mathbf{x}\beta) \end{aligned}$$

Logo, o efeito marginal da mudança nos regressores são:

$$\begin{aligned} \frac{\partial \mathbb{P}(y = 1)}{\partial \mathbf{x}} &= -\phi(\mu_1 - \mathbf{x}\beta)\beta \\ \frac{\partial \mathbb{P}(y = 2)}{\partial \mathbf{x}} &= [\phi(\mu_2 - \mathbf{x}\beta) - \phi(\mu_1 - \mathbf{x}\beta)]\beta \\ \frac{\partial \mathbb{P}(y = 3)}{\partial \mathbf{x}} &= \phi(\mu_2 - \mathbf{x}\beta)\beta \end{aligned}$$

percebe-se que nem o sinal nem a magnitude do coeficiente é informativo sobre o resultado. Então a interpretação da direção dos coeficientes é fundamentalmente ambígua. A implicação deste resultado é que os efeitos de uma mudança em uma das variáveis no modelo depende de todo os parâmetros do modelo, os dados e qual probabilidade da categoria é de interesse (Greene e Hensher, 2009).

O efeito parcial dá o impacto na probabilidade por mudanças de unidade do regressor. Nem o sinal nem a magnitude dos coeficientes são diretamente interpretáveis no modelo de escolha ordenada (Greene e Hensher, 2009). Assim, é necessário computar o efeito parcial (marginal) para interpretar o modelo de forma significativa.

Para definir a função de Máxima Verossimilhança para o modelo probit, defina a

variável $z_{im} = \mathbf{1}(y_i = j)$. Então, a função de verossimilhança para a i -ésima observação pode ser derivada como:

$$\begin{aligned} l_i &= \prod_{j=1}^J \mathbb{P}(y_i = j | \mathbf{x}_i)^{z_{ij}} \\ &= \prod_{j=1}^J [\Phi(\mu_j - \mathbf{x}_i \boldsymbol{\beta}) - \Phi(\mu_{j-1} - \mathbf{x}_i \boldsymbol{\beta})]^{z_{ij}} \end{aligned}$$

agregando todos os indivíduos, temos:

$$L(\mu, \boldsymbol{\beta}) = \prod_{i=1}^n \prod_{j=1}^J [\Phi(\mu_j - \mathbf{x}_i \boldsymbol{\beta}) - \Phi(\mu_{j-1} - \mathbf{x}_i \boldsymbol{\beta})]^{z_{ij}}$$

tomando o logaritmo, encontra-se:

$$L(\mu, \boldsymbol{\beta}) = \sum_{i=1}^n \sum_{j=1}^J z_{ij} [\Phi(\mu_j - \mathbf{x}_i \boldsymbol{\beta}) - \Phi(\mu_{j-1} - \mathbf{x}_i \boldsymbol{\beta})]$$

o ajustamento do modelo ordinal é feito iterativamente através do método mínimo quadrado ponderados (*iteratively reweighted least squares* (IRLS)³) e algoritmos de *Newton-Raphson* (Christensen, 2012).

3.5.1 Modelos logit cumulativos

Existe outra vertente de pensamento sobre modelos ordenados. Podemos encontrar, principalmente na literatura sobre análise de dados categorizados, modelos ordenados pertencentes a uma poderosa classe de Modelo de link Cumulativo. O nome modelos de link cumulativos (*cumulative link models*) é adotado por Agresti (2002). Estes modelos são também conhecidos como *modelos de regressão ordinal* embora esse termo será algumas vezes usado para outros modelos de regressão para resposta ordinal tal como *continuation ratio models* (detalhes em Agresti (2002)). Outros nomes são *modelo logit ordenado* e *modelo probit ordenado* (veja Greene e Hensher (2009)) para funções de ligações logit e probit, respectivamente. Além disso, o *cumulative link models* com link logit é amplamente conhecido como o *proportional odds model* devido a McCullagh (1980).

Um *cumulative link models* é um modelo para variável de resposta ordinal, y_i , que possui uma categoria das $j = 1, \dots, J$ categorias⁴. Então y_i segue uma distribuição multinomial com parâmetro π onde π_{ij} denota a probabilidade que a i -ésima observação caia na categoria resposta j (Christensen, 2012). Definimos a probabilidade acumulada como:

³O pacote VGAM utiliza o método IRLS

⁴Onde $J \geq 2$. Se $J = 2$ modelos binários também se aplicam

$$\gamma_{ij} = \mathbb{P}(y_i \leq j) = \pi_{i1} + \cdots + \pi_{ij} \quad (12)$$

considerando o link logit, a função logit é definida como $\text{logit}(\pi) = \log[\pi/(1 - \pi)]$ e logit's acumulados são definidos:

$$\text{logit}(\gamma_{ij}) = \text{logit}(\mathbb{P}(y_i \leq j)) = \log\left(\frac{\mathbb{P}(y_i \leq j)}{1 - \mathbb{P}(y_i \leq j)}\right), \quad j = 1, \dots, J - 1 \quad (13)$$

de modo que o logit's acumulados são definidos para todas com exceção da última categoria⁵.

Um *cumulative link model* com função de ligação logit, ou simplesmente, *cumulative logit model* (ou *proportional odds model* devido a [McCullagh \(1980\)](#)) é um modelo de regressão para logit's acumulados:

$$\text{logit}(\gamma_{ij}) = \mu_j - \mathbf{x}_i\boldsymbol{\beta}, \quad j = 1, \dots, J - 1 \quad (14)$$

isto é, matematicamente idêntico ao modelo de escolha ordenado discutido anteriormente. Através de notação recentes na literatura, temos:

$$\mathbb{P}(y \leq j) = \Lambda(\mu_j - \mathbf{x}\boldsymbol{\beta}) = \frac{\exp(\mu_j - \mathbf{x}\boldsymbol{\beta})}{1 + \exp(\mu_j - \mathbf{x}\boldsymbol{\beta})} \quad j = 1, \dots, J - 1 \quad (15)$$

Para cada j , existe um logit binário que divide o resultado em categorias inferiores e superiores (por exemplo, categoria 1 *versus* categorias superiores; categorias 1 e 2 *versus* categorias superiores e assim por diante). Para melhor entendimento, suponha que uma variável x_k possui efeito positivo, se x_k aumenta, a probabilidade de categorias inferiores aumenta, enquanto que categorias superiores diminuem. Nota-se que temos várias regressões binárias com diferentes interceptos, μ_j , mas com inclinações idênticas. Isto é conhecido na literatura como a *hipótese de regressão paralela*.

Segundo [Long \(2012\)](#), como consequência da inclinação idêntica, pode-se combinar categorias adjacentes⁶ e obter estimativas consistentes e ineficientes dos β_k 's. Ainda segundo o autor: "Precisão é perdida uma vez que informação é perdida através da combinação de categorias".

Segundo [Jones e Westerland \(2006\)](#) *proportional odds model* na equação 14 possui algumas características atrativas. Como os parâmetros da regressão $\boldsymbol{\beta}$ são invariantes para

⁵Uma vez que $j = J$ o denominador seria $1 - \mathbb{P}(y_i \leq J) = 1 - 1 = 0$ e assim a fração não é definida.

⁶De acordo com [Long \(2012\)](#) a hipótese da Independência de Alternativas Irrelevantes (IIA) é uma propriedade aplicada a modelos logit multinomial e não será aplicada nesse estudo.

as categorias (não são indexadas por j), os odds ratios são os mesmos para as $j - 1$ probabilidade cumulativas (Liu e Agresti, 2005) *apud* (Jones e Westerland, 2006). Definindo o odds ratios como:

$$\frac{\mathbb{P}(y \leq j | \mathbf{x}, x_k + \delta) / \mathbb{P}(y > j | \mathbf{x}, x_k + \delta)}{\mathbb{P}(y \leq j | \mathbf{x}, x_k) / \mathbb{P}(y > j | \mathbf{x}, x_k)}$$

$$\frac{\mathbb{P}(y \leq j | \mathbf{x}, x_k + \delta) / \mathbb{P}(y > j | \mathbf{x}, x_k + \delta)}{\mathbb{P}(y \leq j | \mathbf{x}, x_k) / \mathbb{P}(y > j | \mathbf{x}, x_k)} = \frac{\exp[-(x_k + \delta)\beta]}{\exp(-x_k\beta)} = \exp(-\delta\beta) \quad (16)$$

que é independente de j e proporcional a diferença entre os valores de \mathbf{x} (Greene e Hensher, 2009). Por causa desta proporcionalidade, o modelo denomina-se “*proportional odds*”.

Podemos interpretar o *odds ratio* como: para um aumento de δ em x_k , o *odds* (a “chance”) de uma categoria inferior, comparado a uma categoria superior, muda por um fator de $\exp(-\delta\beta)$, mantendo todas outras variáveis constantes.

Tipos de Links Cumulativos

Tomando a perspectiva de Agresti (2002) podemos analisar modelos de escolha ordenado em uma estrutura de Modelos Lineares Generalizados. O autor ressalta a importância deste tipo de estrutura, permitindo a aplicação de outras funções de ligação. Considere a equação 15, vista de uma outra maneira:

$$G^{-1}[P(Y \leq j | \mathbf{x})] = \mu_j - \mathbf{x}\beta \quad (17)$$

o modelo 17 liga a probabilidade cumulativa ao preditor linear. A função de ligação logit $G^{-1} = \log[u/(1 - u)]$ é a inversa da função de distribuição acumulada da logística padrão. Segundo o autor, o modelo 17 deriva da medida discreta y^* a partir de uma família de parâmetros de locação⁷ tendo distribuição acumulada $G(y^* - \mathbf{x}\beta)$. Os parâmetros μ_j são pontos de cortes das categorias em uma versão padronizada da escala latente. Modelos de link cumulativos são modelos de regressão, pois usam um preditor linear $\mathbf{x}\beta$ para descrever efeitos da variável explanatórias na medição ordinal de y^* . Assim, a escolha da função G é vasta, por exemplo: Normal padrão, distribuição de valor extremo e outras.

3.6 Modelos de Escolha Ordenada Generalizados

Os modelos vistos na seção 3.5.1 impõe restrições de parâmetros β fixos entre as equações. Essa formulação é conhecida na literatura como Hipótese das Regressões Paralelas para modelos ordenados. No entanto, esta hipótese é muitas vezes violada. Assim, precisamos

⁷Os valores de y^* variam em torno de um parâmetro de locação η (tal como a média) que depende de \mathbf{x} através de $\eta(\mathbf{x}) = \mathbf{x}\beta$

linhas paralelas seria muito restritivo e estimar via modelos generalizados produzirá muitos parâmetros a serem interpretados. No entanto, para solucionar esse impasse o modelo *partial proportional odds* permite que alguns β 's sejam o mesmo para todo valor de j , enquanto outros podem diferir. Seja o exemplo de Williams (2006), considere que os β 's para \mathbf{x}_1 e \mathbf{x}_2 são os mesmos para todo j , mas os β 's de \mathbf{x}_3 são livres, então temos:

$$P(y > j) = \frac{\exp(\mu_j + \mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \mathbf{x}_3\beta_{3j})}{1 + \exp(\mu_j + \mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \mathbf{x}_3\beta_{3j})}, j = 1, \dots, J - 1$$

3.6.1 Testando a hipótese de regressão paralela

Uma maneira geral para comparar modelos na estatística é através do teste da razão de verossimilhanças. Considere dois modelos, m_0 e m_1 , onde m_0 é um submodelo do modelo m_1 , isto é, m_0 é mais simples que m_1 e m_0 está contido em m_1 . A estatística da razão de verossimilhança para a comparação de m_0 e m_1 é

$$LR = -2(l_0 - l_1) \quad (21)$$

onde l_0 é o log da verossimilhança de m_0 e l_1 é o log da verossimilhança de m_1 . A estatística da razão de verossimilhanças mede a evidência nos dados para a complexidade extra em m_1 relativo a m_0 . A estatística da razão de verossimilhança segue assintoticamente a distribuição χ^2 com grau de liberdade igual a diferença no número de parâmetros entre m_0 e m_1 . O teste da razão de verossimilhança é geralmente mais preciso que a estatística de Wald. Modelos de escolha ordenada podem ser comparados por meio do teste da razão de verossimilhança, mais precisamente a hipótese de regressão paralela ou modelos generalizados. Assim a hipótese nula do teste é que não existe diferença nos parâmetros β entre as categorias (Christensen, 2012).

4 BASE DE DADOS

Os dados utilizados referem-se ao projeto de pesquisa “Montagem de uma Base de Dados Longitudinal de Vitimização do Ceará: Aspectos Sócio-Econômicos e Espaciais”, realizada em 2012. O projeto teve como participantes a Universidade Federal do Ceará, através do CAEN/Pós-graduação em Economia, e a Secretaria de Segurança Pública e Defesa Social do Estado do Ceará (SSPDS/CE). O projeto surgiu em um contexto paralelo a outro projeto⁹ (também financiado pela FUNCAP) que consistia em desenvolver um laboratório de análise estatística e criminal no âmbito da SSPDS/CE sob a coordenação do Prof. Dr. José Raimundo A. Carvalho (CAEN/UFC). Este levantamento contemplou 118 bairros de Fortaleza/CE. Tal pesquisa foi realizada com indivíduos de ambos os sexos com idade mínima de 16 anos, tota-

⁹Desenvolvimento de Infraestrutura de Análise Estatística e Criminal na SSPDS/CE.

lizando uma amostra de 4030 pessoas que responderam a um questionário com 123 perguntas com temas de análises socioeconômicas, criminológicas e de segurança pública.

Especificado a base metodológica neste trabalho, faremos agora uma breve descrição dos dados antes de prosseguirmos para a apresentação dos resultados. A variável dependente considerada neste trabalho é a resposta do nível de satisfação do entrevistado com relação a acessibilidade para cinco destinos de Fortaleza de forma independente. Vale salientar como descrito na seção 2.2.4 que a variável dependente mensura a utilidade do indivíduo com relação a seguinte argumentação:

“Gostaria agora que o(a) Sr(a) pensasse sobre o que acontece quando é preciso ir a locais para resolver problemas ou aproveitar o tempo. Considerando que acessibilidade envolve tudo aquilo que é importante para ir e voltar (tempo de percurso, custo de transporte, conforto da jornada, horário de funcionamento, condições das ruas, estradas, entre outros), como o(a) Sr(a) classificaria o acesso aos seguintes locais”.

Os locais considerado neste estudo são: Hospital IJF-Instituto José Frota, Rodoviária Central de Fortaleza, Estádio Castelão, Aeroporto Internacional Pinto Martins e Mercado Central de Fortaleza. Para melhor entendimento da distribuição espacial desses destinos, a figura 1 facilitará a visualização:

Considerou-se como a escala de classificação ao acesso: 1-Péssimo, 2-Ruim, 3-Razoável, 4-Bom e 5-Excelente. As variáveis explicativas/independentes utilizadas são:

[SEXO] Sexo do Entrevistado: 0-Feminino e 1-Masculino

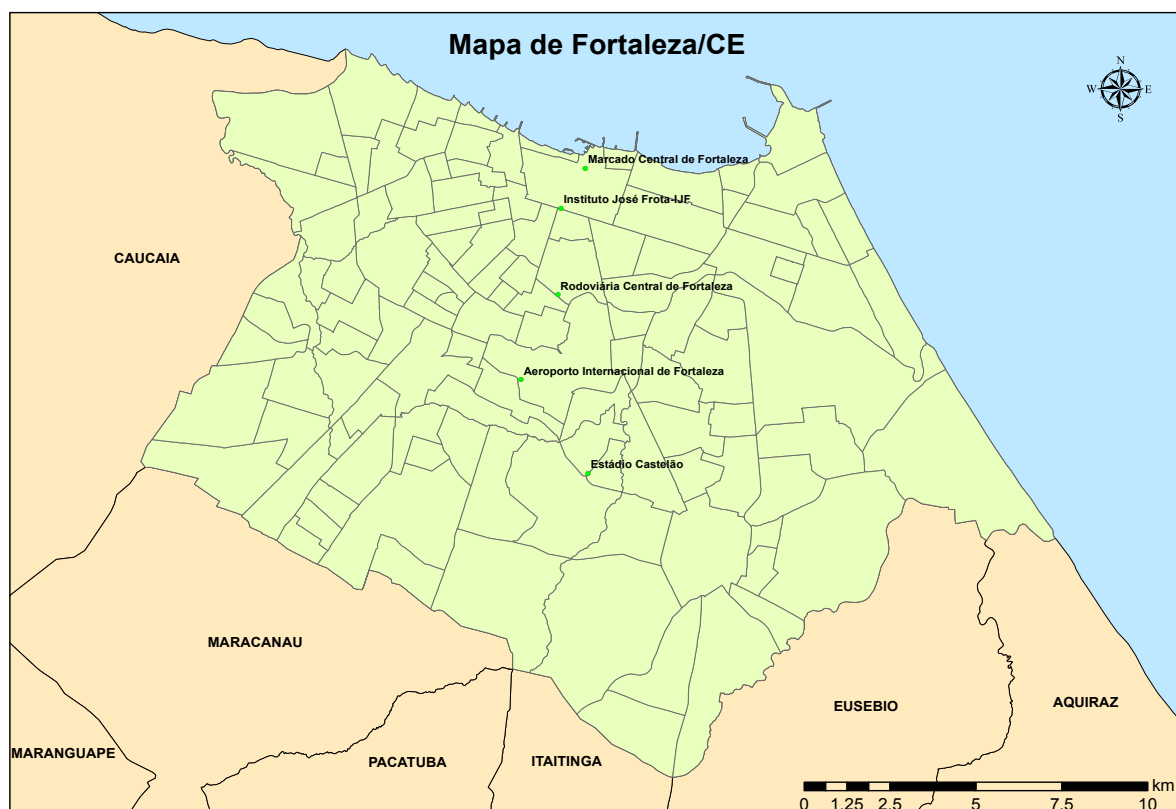
[IDADE] Idade do Entrevistado

[EDUCACAO] Qual é o curso mais elevado que o(a) Sr(a) já concluiu: 1-Sem instrução, 2-Ensino Fundamental incompleto (1º Grau incompleto), 3-Ensino Fundamental completo (1º Grau completo), 4-Ensino médio incompleto (2º Grau incompleto), 5-Ensino médio completo (2º Grau completo), 6-Superior incompleto (Universidade / Faculdade incompleto), 7-Superior completo (Universidade / Faculdade completo) e 8-Pós-Graduação;

[TRANSPORTE] Qual o principal meio de transporte que o(a) Sr(a) costuma utilizar em seu dia-a-dia: 1-A pé, 2-Bicicleta, 3-Moto, 4-Carro (particular ou carona) e 5-Transporte Público;

[RENDA.IND] Contempla a faixa de renda do entrevistado em termos de salários mínimos com os seguintes resultados: 1-Não possui renda, 2-Até 1 salário mínimo (R\$ 545), 3-Mais de 1 até 2 salários mínimos (R\$ 545 a R\$ 1.090), 4-Mais de 2 até 3 salários mínimos (R\$ 1.090 a R\$ 1.635), 5-Mais de 3 até 5 salários mínimos (R\$ 1.635 a R\$ 2.725), 6-Mais de 5 até 10 salários mínimos (R\$ 2.725 a R\$ 5.450),

Figura 1: Mapa de Fortaleza



Fonte: Dados gerados pelo autor

7-Mais de 10 até 15 salários mínimos (R\$ 5.450 a R\$ 8.175), 8-Mais de 15 até 20 salários mínimos (R\$ 8.175 a R\$ 10.900), 9-Mais de 20 salários mínimos (Maior que R\$ 10.900);

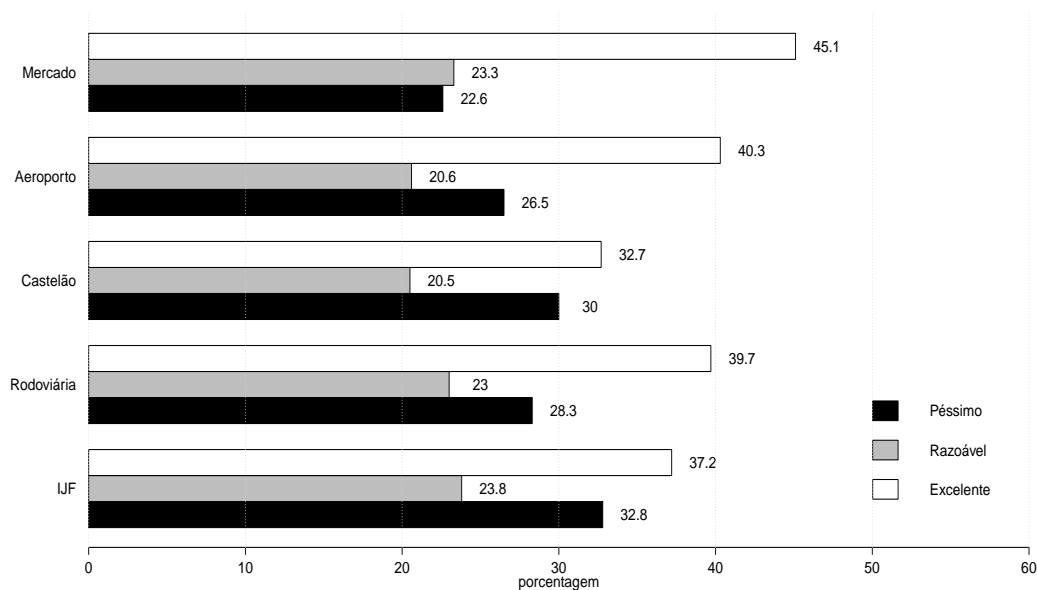
[**DIST.DESTINO**] Distância (em quilômetros) da residência do entrevistado ao local de destino: onde DESTINO é substituído por: IJF (Hospital IJF-Instituto José Frota), RODOVIARIA (Rodoviaria Central de Fortaleza), ESTADIO (Estádio Castelão), AEROPORTO (Aeroporto Internacional Pinto Martins), MERCADO (Mercado Central de Fortaleza).

As variáveis escolhidas são baseadas na metodologia proposta no Capítulo 2, onde medidas de utilidades são compostas de componentes de acessibilidade. Como mostrado, as componentes de acessibilidade são: transporte, uso do solo, tempo e individual. No estudo, a componente transporte é representada pela distância entre a residência e o destino e a componente individual é representada pelas características socioeconômicas do indivíduo, não utilizaremos componentes de uso do solo e tempo. Vale salientar que não utilizou-se nenhum critério de seleção das variáveis, porém, analisou-se as principais variáveis socioeconômicas do entrevistado.

As análises estatísticas foram feitas utilizando o *software* estatístico R¹⁰ para a es-

¹⁰R Development Core Team (2012)

Figura 2: Opinião dos entrevistados por destino



Fonte: Dados gerados pelo autor

timação¹¹ e SPSS para manipulações da base de dados. Como no presente estudo utilizou-se a distância entre a residência e o destino, os dados coletados foram geocodificados, restando 3324 entrevistados (dos 4030) adequados para seguir no processo de modelagem dos dados.

Foram estimados cinco modelos, independentes, baseados nos destinos citados anteriormente. Em cada modelo houve perda de informação (*missing*), havendo variação da quantidade de dados analisados para cada modelo.

Considerando os 3324 entrevistados, as principais características dos entrevistados são (ver Tabela 2): sexo masculino (45%) e (55%) feminino. Sobre o nível de instrução dos entrevistados teremos: Sem instrução (5,5%), Ensino Fundamental incompleto (33%), Ensino Fundamental completo (11,4%), Ensino médio incompleto (12,5%), Ensino médio completo (28,5%), Superior incompleto (4,2%), Superior completo (3,7%) e Pós-Graduação (1,1%).

Com relação ao principal meio de transporte utilizado: A pé (19,8%), Bicicleta (8,8%), Moto (7,5%), Carro (14,9%) e Transporte Público (48,9%). Sobre a faixa de renda os entrevistados se distribuem: Não possui renda (28,1%), Até 1 salário mínimo (43,6%), Mais de 1 até 2 salários mínimos (17,6%), Mais de 2 até 3 salários mínimos (5,8%), Mais de 3 até 5 salários mínimos (2,5%), Mais de 5 até 10 salários mínimos (1,6%), Mais de 10 até 15 salários mínimos (0,3%), Mais de 15 até 20 salários mínimos (0,2%), Mais de 20 salários mínimos (0,2%).

Sobre as variáveis dependentes temos a seguinte Tabela 3 e Figura 2 mostrando a distribuição das respostas para cada destino.

Com o intuito de facilitar a interpretação do modelo proposto, optamos por agregar às categorias 1 e 2 (respectivamente “péssimo” e “ruim”) em “Péssimo” e também às categorias

¹¹Tivemos auxílio do pacote VGAM, ver Yee (2010)

Tabela 2: Estatística descritiva das variáveis

	[Amostra] n=3324
SEXO:	
0: Feminino	1843 (55.4%)
1: Masculino	1481 (44.6%)
EDUCACAO:	
Sem instrução	183 (5.51%)
1º Grau Incompleto	1096 (33.0%)
1º Grau Completo	380 (11.4%)
2º Grau Incompleto	417 (12.5%)
2º Grau Completo	947 (28.5%)
Superior Incompleto	140 (4.21%)
Superior Completo	123 (3.70%)
Pós-Graduação	38 (1.14%)
TRANSPORTE:	
1: A pé	658 (19.7%)
2: Bicicleta	293 (8.81%)
3: Moto	252 (7.59%)
4: Carro	494 (14.9%)
5: Transporte Público	1625 (48.9%)
<i>Missing</i>	2 (0.06%)
RENDA.IND:	
Não possui renda	922 (27.7%)
Até 1 salário mínimo (R\$ 545)	1432 (43.1%)
Mais de 1 até 2 salários mínimos	576 (17.3%)
Mais de 2 até 3 salários mínimos	192 (5.7%)
Mais de 3 até 5 salários mínimos	82 (2.4%)
Mais de 5 até 10 salários mínimos	54 (1.6%)
Mais de 10 até 15 salários mínimos	12 (0.36%)
Mais de 15 até 20 salários mínimos	6 (0.18%)
Mais de 20 salários mínimos	6 (0.18%)
<i>Missing</i>	42 (1.26%)
Var. Contínuas	Média(dp)
IDADE	39.4 (16.9)

Fonte: Dados gerados pelo autor

4 e 5 (respectivamente “bom” e “excelente”) em “Excelente” e mantivemos a categoria 3 como “Razoável”.

Tabela 3: Estatística descritiva para as variáveis dependentes (em %)

Resposta	IJF	Rodoviária	Castelão	Aeroporto	Mercado
1-Péssimo	32.8	28.3	30.0	26.5	22.6
2-Razoável	23.8	23.0	20.5	20.6	23.3
3-Excelente	37.2	39.7	32.7	40.3	45.1

Fonte: Dados gerados pelo autor

5 UM MODELO ECONOMETRICO DE ACESSIBILIDADE URBANA

5.1 Introdução

Por meio do modelo logit ordenado proposto na equação 22, estimamos os parâmetros para os cinco modelos para cada destino considerado nesta análise. A estratégia econométrica por trás de um modelo de utilidade, proposta na seção 2.2.4 busca avaliar que fatores influenciam a escolha do entrevistado com relação ao seu acesso. Com esse intuito descreveremos a seguir os principais achados para este trabalho.

5.2 Modelo Econométrico

Com o intuito de estimar a probabilidade do nível de acessibilidade, em categorias ordenadas, entre o domicílio e o destino a ser considerado e definidas as variáveis a serem utilizadas na presente pesquisa, as equações estimadas para os cinco modelos logit ordenado serão:

$$\begin{aligned}
U_{IJF}^* &= \mu_j^{IJF} + \beta_1 \text{Sexo} + \beta_2 \text{Idade} + \beta_3 \text{Educação} + \beta_4 \text{Transporte} + \beta_5 \text{RendaInd} + \\
&\quad \beta_6 \text{DistIJF} + \epsilon_{IJF} \\
U_{rodoviaria}^* &= \mu_j^{rod} + \delta_1 \text{Sexo} + \delta_2 \text{Idade} + \delta_3 \text{Educação} + \delta_4 \text{Transporte} + \delta_5 \text{RendaInd} + \\
&\quad \delta_6 \text{DistRodoviaria} + \epsilon_{rodoviaria} \\
U_{castelao}^* &= \mu_j^{cast} + \gamma_1 \text{Sexo} + \gamma_2 \text{Idade} + \gamma_3 \text{Educação} + \gamma_4 \text{Transporte} + \gamma_5 \text{RendaInd} + \\
&\quad \gamma_6 \text{DistCastelao} + \epsilon_{castelao} \\
U_{aeroporto}^* &= \mu_j^{aerop} + \lambda_1 \text{Sexo} + \lambda_2 \text{Idade} + \lambda_3 \text{Educação} + \lambda_4 \text{Transporte} + \lambda_5 \text{RendaInd} + \\
&\quad \lambda_6 \text{DistAeroporto} + \epsilon_{aeroporto} \\
U_{mercado}^* &= \mu_j^{merc} + \phi_1 \text{Sexo} + \phi_2 \text{Idade} + \phi_3 \text{Educação} + \phi_4 \text{Transporte} + \phi_5 \text{RendaInd} + \\
&\quad \phi_6 \text{DistMercado} + \epsilon_{mercado}
\end{aligned} \tag{22}$$

de modo que:

U^* é a variável latente representando o comportamento do indivíduo (nível de utilidade) com relação ao nível de acesso que será categorizada em três ordenamentos de respostas: 1-Péssimo, 2-Razoável e 3-Excelente;

μ_j representa o *threshold* para cada resposta j , onde $j = 1, 2$;

ϵ_i é o termo de erro aleatório.

5.3 Destino: Hospital de Emergência - Instituto José Frota (IJF)

Iniciaremos esta seção com o modelo proposto ao Hospital de Emergência - Instituto José Frota (IJF), discutiremos as estimativas do modelo, bem como suas interpretações. Nas seções seguintes abordaremos outros destinos. Por meio da proposta metodológica descrita na seção 3.5, estimou-se os parâmetros das variáveis explicativas via Máxima Verossimilhança. Os resultados se encontram na Tabela 4. Vale salientar que a variável independente transporte contém como categoria de referência a resposta “a pé” e as demais serão vistas como: Transporte:Bicicleta, Transporte:Moto, Transporte:Carro, Transporte:Trans.Público.

Tabela 4: Modelo logit de escolha ordenada estimado via regressão paralela (destino: IJF)

	Coefficiente	Erro padrão	z	Valor-P	Odds-Ratio	
threshold:1	1.190	0.174	6.837	< 0.001	-	***
threshold:2	0.096	0.173	0.557	0.578	-	
Sexo(masculino)	0.058	0.075	0.775	0.438	1.060	
Idade	0.013	0.002	5.583	< 0.001	1.013	***
Educação	0.016	0.024	0.659	0.510	1.015	
Transporte:Bicicleta	-0.144	0.142	-1.021	0.307	0.865	
Transporte:Moto	-0.237	0.150	-1.578	0.114	0.789	
Transporte:Carro	-0.288	0.126	-2.284	0.022	0.749	*
Transporte:Público	-0.199	0.093	-2.142	0.032	0.819	*
Renda Indiv.	-0.074	0.035	-2.122	0.034	0.928	*
Distância-IJF	-0.120	0.011	-11.427	< 0.001	0.886	***

- - -

*** indica significância de 0.1%; ** de 1%; * de 5%; · de 10%

N=3010

Log-likelihood: -3164.02

Fonte: Dados gerados pelo autor

Como podem ser verificados nos resultados da Tabela 4, os coeficientes das variáveis binária de transporte para carro e transporte público, renda do indivíduo e distância do IJF para o domicílio do indivíduo são negativos e estatisticamente significativos. A variável idade do entrevistado foi também significativa e com sinal positivo. Por outro lado, os coeficientes das variáveis sexo (masculino como categoria de referência) e educação foram estatisticamente insignificantes com relação ao nível de acessibilidade para o IJF.

Nota-se que pessoas mais velhas são mais favoráveis ao nível de acessibilidade positivos. Porém, a variável binária para carro e transporte público (a resposta “a pé” foi considerada como categoria de referência), nível de renda e a distância tendem a influenciar resultados menos favoráveis para o nível de acessibilidade. No entanto, podemos supor que o aumento do nível da renda dos indivíduos permitirá que os mesmos não frequentem o IJF, dado que eles possuem plano mais satisfatório, então, o uso fica restrito a pessoas com baixo nível de renda. Para

o meio de transporte, observamos que o uso de bicicletas e uso de moto, com relação ao deslocamento a pé, não produziu diferença no nível de acesso, já o uso de carro e transporte público afetam o bem-estar do entrevistado negativamente. Com relação ao *Odds Ratios*, algumas vezes é conveniente falar sobre mudanças percentuais que é computada como $ORP_k = 100(OR_k - 1)$, daí, podemos notar que o acréscimo de dez anos de idade aumenta a chance em 14% de escolher alternativas superiores *versus* alternativas inferiores, mantendo as outras variáveis constantes.

Todavia, os valores dos parâmetros estimados não fornecem informações precisas dos efeitos ou elasticidades. Além disso, os valores dos coeficientes não podem ser interpretados como uma influência quantitativa sobre a probabilidade, uma vez que sua relação é não linear. Antes de calcularmos e discutirmos as elasticidades do modelo, procedemos à análise da especificação do mesmo, principalmente no que diz respeito à hipótese de “regressão paralela”.

Como foi mostrado na seção 3.5, é interessante notar que, uma vez definidas as três respostas ordenadas (“1:Péssimo”, “2:Razoável” e “3:Excelente”) da variável dependente, têm-se dois interceptos. Contudo, esses interceptos não podem ser interpretados diretamente, até porque eles dependem dos valores das variáveis independentes.

O modelo estimado (tabela 4) sob H_0 assume que todos os parâmetros de inclinação são iguais (i.e que as curvas estimadas são paralelas) enquanto que o modelo generalizado assume que a H_1 : Existe pelos menos um parâmetro diferente entre as categorias.

Estes resultados, apesar de bastante razoáveis, tendem a suscitar dúvidas, pois implicitamente se assume a hipótese H_0 de que os β 's são os mesmos para as diferentes categorias enquanto que o modelo generalizado assume que a hipótese H_1 : existe pelo menos um β diferente entre as diferentes categorias. Para verificar tal suposição utilizaremos o teste da razão de verossimilhança $LR = -2(l_0 - l_1)$, que segue distribuição χ^2 , para cada parâmetro separadamente e globalmente. Ou seja, o modelo logit ordenado assume que os coeficientes que descrevem a relação entre uma resposta categórica, por exemplo, “Péssimo”, com todas as outras (“Razoável” e “Excelente”) são as mesmas daqueles que descrevem a relação entre “Péssimo” e “Razoável” *versus* “Excelente” (detalhes veja seção 3.5).

Tabela 5: Teste da hipótese de regressão paralela (destino: IJF)

	χ^2	gl	$p > \chi^2$
Global	29.070	9	<0.001
Sexo	4.000	1	0.046
Idade	9.800	1	0.002
Educação	13.860	1	<0.001
Transporte	9.530	4	0.049
Renda	1.760	1	0.184
Distância	1.120	1	0.289

Fonte: Dados gerados pelo autor

Assim, de acordo com a tabela 5, pode-se verificar que a hipótese de regressões

paralelas foi rejeitada ($\chi^2_9 = 29,07$ e $p < 0,001$). Logo, emprega-se o modelo logit ordenado generalizado permitindo que os β 's sejam flexíveis para cada categoria, menos para as variáveis sexo, transporte, renda e distância (seremos criteriosos nesta análise utilizando nível de significância de 1%). A seguir é apresentado o resultado da estimação para o modelo ordenado generalizado, nota-se que a variável idade e educação possuem 2 (duas) estimativas, a primeira estimativa vem da regressão “Péssimo” *versus* as categorias superiores e a segunda estimativa uni-se as categorias “Péssimo” e “Razoável” *versus* a resposta “Excelente”.

Tabela 6: Modelo logit de escolha ordenada estimado via regressão generalizada (destino: IJF)

	Coefficiente	Erro padrão	z	Valor-P
threshold:1	1.167	0.189	6.190	< 0.001 ***
threshold:2	0.125	0.183	0.680	0.496
Sexo(masculino)	0.061	0.075	0.806	0.420
Idade:1	0.010	0.003	3.839	< 0.001 ***
Idade:2	0.015	0.003	5.959	< 0.001 ***
Educação:1	0.053	0.027	1.976	0.048 *
Educação:2	-0.015	0.026	-0.568	0.570
Transporte:Bicicleta	-0.137	0.142	-0.967	0.333
Transporte:Moto	-0.236	0.150	-1.571	0.116
Transporte:Carro	-0.288	0.126	-2.282	0.022 *
Transporte:Público	-0.198	0.093	-2.126	0.033 *
Renda Indiv.	-0.074	0.035	-2.127	0.033 *
Distância-IJF	-0.121	0.011	-11.488	< 0.001 ***

- - -

*** indica significância de 0.1%; ** de 1%; * de 5%; · de 10%

N=3010

Log-likelihood: -3154.48

Fonte: Dados gerados pelo autor

O modelo generalizado é mais difícil de interpretar que o modelo de regressão paralela anterior, porém fornece percepções que foram omitidas anteriormente. Basicamente, a metodologia por trás disso é a estimação de várias regressões logísticas binárias (Williams, 2006), i.e, a primeira regressão confronta a categoria 1 (péssimo) com as categorias 2 (razoável) e 3 (excelente), na segunda regressão confronta categorias 1 e 2 com a categoria 3. Assim, coeficientes positivos indicam que valores altos na variável independente possuem efeito maior e tornam mais provável que o entrevistado estará em uma categoria maior de U que a categoria atual, enquanto coeficientes negativos indicam que maiores valores na variável explanatória aumentam a chance de estar na categoria atual ou menor. A questão de interpretação a partir daqui será apenas uma questão de grau.

Na Tabela 6 detectamos, por exemplo, que pessoas mais velhas tendem a ser mais favoráveis ao acesso para o IJF. Mas o maior efeito da idade foi na segunda regressão. Existe um coeficiente de inclinação maior na segunda regressão do que na primeira, ou seja, na segunda regressão a idade tem mais peso. Dessa forma é mais provável (ou existe uma chance maior)

que idosos escolham a categoria excelente do que as demais. Com relação a educação, o efeito é observado somente na primeira regressão, assim, quanto maior o nível educacional maior será a chance do entrevistado optar pelas categorias Razoável ou Excelente. Já as variáveis como renda, tipo de transporte, sexo e distância não apresentam nenhuma mudança em relação aos resultados obtidos na tabela 4, mantendo-se sua interpretação.

Após verificar a significância estatística dos parâmetros, deve-se levar em conta que os coeficientes estimados não refletem os efeitos marginais das variáveis, dessa forma, estas análises devem ser feitas em separado. Com efeito, calculou-se o efeito parcial médio¹² (*Average Partial Effect*) para facilitar a interpretação, juntamente com auxílio do *software* STATA na geração dos erros padrões fundamentais para a significância.

Tabela 7: Efeito Marginal Médio (destino: IJF)

	Péssimo	Razoável	Excelente
Sexo	-0.0132	-0.0005	0.0137
Idade	-0.0022***	-0.0012**	0.0034***
Educação	-0.0115*	0.0148**	-0.0033
Transporte:Bicicleta	0.0297	0.0012	-0.0310
Transporte:Moto	0.0511	0.0021	-0.0532
Transporte:Carro	0.0624*	0.0026	-0.0651**
Transporte:Público	0.0429*	0.0018	-0.0447*
Renda Ind.	0.0161*	0.0006	-0.0167*
Distância	0.0262***	0.0011**	-0.0273***

Fonte: Dados gerados pelo autor

Em números, por exemplo na variável idade, existe aumento de 0.3% na probabilidade de escolher Excelente com aumento de um ano de idade, que reforça o resultado visto na tabela 6 na qual existe um efeito maior na segunda regressão do que na primeira.

Outro exemplo na variável educação, quanto maior o nível educacional dos entrevistados haverá um aumento em 1% a probabilidade de responder a opção Razoável. Contudo, como mencionamos anterioremos, quanto maior o nível educacional o indivíduo tende a estar nas opções Razoável e Excelente. Analisando a tabela 7, observamos que existe efeito negativo na probabilidade de escolha das categorias Excelente e Péssimo, que seria uma contradição à afirmação prestada. No entanto isso não acontece pois a primeira regressão da regressão generalizada engloba as categorias Razoável e Excelente como uma categoria agregada, assim, o indivíduo prefere escolher a categoria conjunta Razoável ou Excelente do que a Péssimo. Mas, considerando as três categorias de forma distinta o entrevistado tenderia somente a escolha da opção Razoável ao invés da Péssimo ou Excelente.

O efeito marginal da variável renda individual mostra a dependência da categoria na interpretação da variável. Note que, com o aumento da renda existe aumento na probabilidade

¹²Para ver detalhes da distribuição empírica dos efeitos por categoria veja anexo A

da categoria péssimo em 1.6%, ou seja, tal resultado leva a geração da desutilidade em pessoas com maiores níveis de renda ao IJF, preferindo hospitais particulares.

Com relação a distância o efeito é notável, quanto maior a distância do IJF as probabilidades de escolher as categorias péssimo aumentam, ao contrário das categorias excelente que diminuem.

As variáveis categóricas de transportes todas, exceto uso de bicicleta e modo, com relação ao deslocamento a pé, têm a mesma interpretação. Existe aumento na probabilidade da escolha da categoria péssimo. Com relação ao sexo, homens possuem probabilidade de 1.3% maior de escolher a categoria excelente do que as mulheres.

Uma outra maneira de explicar os efeitos é plotando a probabilidade estimada ($\hat{P}(y = j)$) por meio de mudanças na variável quantitativa. Por exemplo, seja $\bar{x}_{[-distancia]}$ contendo as médias de todas as variáveis exceto distância. Aplicando as equações 19 e $\hat{\beta}$, encontramos as probabilidades estimadas individuais para cada categoria. Para ressaltar os resultados já mencionados, na figura 15, observa-se que o aumento da distância tende o indivíduo a optar pela categoria péssimo, enquanto que a probabilidade de escolha da categoria razoável e excelente diminuem com o aumento da distância. Entretanto, a uma distância de aproximadamente 7 km os indivíduos têm mais probabilidade de optar pela categoria excelente.

Neste trabalho estimamos cinco modelos independentes representando cada destino. Apesar de nesta seção termos descrito de maneira completa a interpretação de modelos ordenados para o destino IJF, nas seções posteriores resumiremos os comentários em relação aos resultados das estimações para os outros destinos.

5.4 Destino: Estádio Castelão

Continuando com a interpretação dos modelos propostos neste trabalho, discutiremos a seguir o modelo estimado para avaliar o nível de acesso da população com relação ao Estádio Castelão. A tabela 8 apresenta as primeiras estimativas.

A maioria dos coeficientes foram significativos, mais precisamente, a variável binária para uso de transporte público, renda e a distância do Estádio Castelão para a residência do entrevistado foram negativos. A variável educação e sexo foi também significante, porém, com sinal positivo.

Nota-se, por exemplo, que pessoas com maiores níveis educacionais e homens são mais favoráveis ao acesso ao Castelão. Porém, o uso de transporte público e a distância tendem a influenciar o bem-estar do entrevistado provocando opiniões negativas com relação ao seu acesso. Com relação ao *Odds Ratios*, o acréscimo de dez anos de idade aumenta a chance em 3% de escolher alternativas superiores *versus* alternativas inferiores, mantendo as outras variáveis constantes.

Assim, de acordo com a Tabela 9, pode-se verificar que a hipótese de regressões paralelas foi rejeitada ($\chi^2(9) = 17,38$ com $p < 0,043$). Porém, analisaremos somente a variável

Tabela 8: Modelos Logit de Escolha Ordenado Estimado via Regressão Paralela (destino: Castelo)

	Coefficiente	Erro padrão	z	Valor-P	Odds-Ratio	
threshold:1	1.691	0.189	8.940	< 0.001	-	***
threshold:2	0.603	0.187	3.233	0.001	-	**
Sexo(masculino)	0.419	0.080	5.221	< 0.001	1.521	***
Idade	0.003	0.003	1.210	0.226	1.003	
Educação	0.052	0.025	2.068	0.039	1.053	*
Transporte:Bicicleta	-0.184	0.150	-1.224	0.221	0.831	
Transporte:Moto	-0.238	0.159	-1.496	0.135	0.788	
Transporte:Carro	-0.188	0.136	-1.384	0.166	0.828	
Transporte:Público	-0.296	0.102	-2.910	0.004	0.743	**
Renda Indiv.	-0.056	0.037	-1.505	0.132	0.945	
Distância-Castelão	-0.172	0.013	-13.019	< 0.001	0.841	***

- - -

*** indica significância de 0.1%; ** de 1%; * de 5%; · de 10%

N=2667

Log-likelihood: -2779.12

Fonte: Dados gerados pelo autor

Tabela 9: Teste da Hipótese de Regressão Paralela (destino: Castelo)

	χ^2	gl	p > χ^2
Global	17.380	9	0.043
Sexo	8.060	1	0.004
Idade	2.970	1	0.085
Educação	2.050	1	0.152
Transporte	4.960	4	0.291
Renda	0.020	1	0.880
Distância	1.550	1	0.213

Fonte: Dados gerados pelo autor

sexo, já que individualmente todas as demais variáveis aceitam a suposição.

Note que, através da Tabela 10, somente a variável sexo tem duas estimativas e as demais somente uma. De fato, homens dão mais peso a primeira regressão (Péssimo *versus* Razoável e Excelente), ou seja, a chance do homem escolher a opção Excelente e Razoável é maior do que somente Excelente *versus* Razoável e Péssimo.

Com o efeito marginal, na tabela 11, podemos detectar que homens possuem aproximadamente 11.5% de escolher a opção Razoável ou Excelente do que as mulheres, as quais teriam 11.4% de escolher a resposta Péssimo.

Pessoas que costumam usar o transporte público diariamente possuem aproximadamente 6% a mais, do que pessoas que constumam se deslocar “a pé”, de escolher a opção

Tabela 10: Modelos Logit de Escolha Ordenado Estimado via Regressão Generalizada (destino: Castelão)

	Coefficiente	Erro padrão	z	Valor-P
threshold:1	1.637	0.190	8.616	< 0.001 ***
threshold:2	0.654	0.188	3.488	< 0.001 ***
Sexo(masculino):1	0.530	0.090	5.916	< 0.001 ***
Sexo(masculino):2	0.315	0.088	3.571	< 0.001 ***
Idade	0.003	0.003	1.206	0.228
Educação	0.052	0.025	2.062	0.039 *
Transporte:Bicicleta	-0.179	0.150	-1.192	0.233
Transporte:Moto	-0.236	0.159	-1.483	0.138
Transporte:Carro	-0.185	0.136	-1.360	0.174
Transporte:Público	-0.290	0.102	-2.850	0.004 **
Renda Indiv.	-0.055	0.037	-1.486	0.137
Distância-Castelão	-0.172	0.013	-13.011	< 0.001 ***

- - -

*** indica significância de 0.1%; ** de 1%; * de 5%; · de 10%

N=2667

Log-likelihood: -2775.09

Fonte: Dados gerados pelo autor

Tabela 11: Efeito Marginal Médio (destino: Castelão)

	Péssimo	Roçoável	Excelente
Sexo	-0.1146***	0.0444**	0.0702***
Idade	-0.0006	-0.0001	0.0006
Educação	-0.0113*	-0.0004	0.0116*
Transporte:Bicicleta	0.0387	0.0012	-0.0399
Transporte:Moto	0.0509	0.0016	-0.0526
Transporte:Carro	0.0399	0.0012	-0.0412
Transporte:Público	0.0627**	0.0019	-0.0647**
Renda Ind.	0.0119	0.0004	-0.0123
Distância	0.0372***	0.0011.	-0.0384***

Fonte: Dados gerados pelo autor

péssimo.

Com relação a distância o efeito é notável, quanto maior a distância para o estádio as probabilidade de escolher as categorias péssimo aumentam, ao contrário das categorias excelente que diminuem. Com até 8 km de distância para o Castelão, entrevistados optam pela categoria excelente (veja 15). A partir daqui, para não ser repetitivo, condensaremos as estimativas dos modelos restantes na representação de tabelas únicas.

5.5 Destinos: Rodoviária Central, Aeroporto e Mercado Central

Abordaremos de forma mais resumida os modelos restantes. As interpretações são análogas as vistas nas seções 5.3 e 5.4.

Tabela 12: Coeficientes estimados dos destinos Rodoviária, Aeroporto e Mercado via regressão paralela

	Rodoviária ($\exp(\delta_k)$)		Aeroporto ($\exp(\lambda_k)$)		Mercado ($\exp(\phi_k)$)	
threshold:1	1.459	***	1.397	***	2.276	***
threshold:2	0.356	*	0.357	.	1.105	***
Sexo(masculino)	0.225	(1.25)**	0.282	(1.32)***	0.121	(1.12)
Idade	0.008	(1.01)**	0.002	(1.00)	0.006	(1.00)*
Educação	0.047	(1.04).	0.051	(1.05)*	-0.030	(0.97)
Transporte:Bicicleta	-0.171	(0.84)	-0.060	(0.94)	-0.220	(0.80)
Transporte:Moto	-0.251	(0.77)	-0.144	(0.86)	-0.717	(0.48)***
Transporte:Carro	-0.331	(0.71)*	-0.012	(0.98)	-0.467	(0.62)***
Transporte:Público	-0.345	(0.70)***	-0.387	(0.67)***	-0.306	(0.73)**
Renda Indiv.	-0.052	(0.94)	0.052	(1.05)	-0.104	(0.90)**
Distância	-0.141	(0.86)***	-0.138	(0.87)***	-0.108	(0.89)***

*** indica significância de 0.1%; ** de 1%; * de 5%; . de 10%

$N_{rodoviaria}=2910$; $N_{aeroporto}=2797$; $N_{mercado}=2925$

Log-likelihood: -3047.68;-2898.06;-2961.63

Fonte: Dados gerados pelo autor

Verifica-se que, de acordo com a tabela 12, a existência de algumas interpretações equivalentes diferindo apenas no valor dos coeficientes. Por exemplo, homens tendem a opinar de forma positiva com relação a questão do acesso destes destinos (Rodoviária Central, Aeroporto e Mercado Central). Outro resultado, a idade não é um fator decisivo na questão do acesso ao aeroporto, porém, idosos tendem a ser favoráveis ao acesso para a Rodoviária e ao Mercado. A educação possui os mesmos resultados encontrados para o IJF e Castelão, quanto maior o nível educacional do indivíduo maiores as chances do entrevistador opinar positivamente com relação ao processo de deslocamento para a Rodoviária e Aeroporto. Somente no Mercado essa variável não conseguiu explicar o acesso. Novamente, as interpretações mantiveram-se as mesmas (para aqueles significativos) com relação ao meio utilizado no dia-a-dia, os resultados mostram que o uso de algum meio (bicicleta, moto, carro e transporte público) é sempre pior que o deslocamento a pé, segundo o modelo. Com relação a renda individual, obtivemos somente coeficiente positivo para o destino Aeroporto, ao contrário de todos os outros destinos considerados. A distância é sempre vista como um fator de dificuldade no deslocamento de dois pontos quaisquer. Dessa forma, em todos os modelos estudados o coeficiente é negativo, mostrando que o indivíduo considera a distância como ponto negativo do trajeto para seu bem-estar.

Os efeitos marginais podem ser vistos na Tabela 13.

Tabela 13: Efeito marginal médio para os destinos Rodoviária, Aeroporto e Mercado

		Rodoviária	Aeroporto	Mercado Central
Péssimo	Sexo	-0.07336 ***	-0.05744 ***	-0.02180
	Idade	-0.00101 .	0.00025	-0.00111 *
	Educação	-0.00982 *	-0.01029 *	-0.00156
	Transporte:Bicicleta	0.03345	-0.00537	0.03907
	Transporte:Moto	0.05008	-0.01553	0.12643 ***
	Transporte:Carro	0.06682 *	-0.06841 .	0.08362 ***
	Transporte:Público	0.06921 ***	0.05728 ***	0.05451 **
	Renda	0.01049	-0.01112	0.01892 **
	Distância	0.02924 ***	0.02841 ***	0.02312 ***
Razoável	Sexo	0.04289 *	-0.00960 ***	-0.00742
	Idade	-0.00130 **	-0.00111 **	-0.00038 *
	Educação	-0.00139 .	-0.00172 *	0.01417 **
	Transporte:Bicicleta	0.00474	0.03150	0.01331 .
	Transporte:Moto	0.00710 **	0.07937	0.04306 ***
	Transporte:Carro	0.00948 ***	0.11280 ***	0.02848 ***
	Transporte:Público	0.00982 ***	0.04718 ***	0.01856 **
	Renda	0.00149	-0.00186	0.00644 **
	Distância	0.00415 ***	0.00475 ***	-0.00018
Excelente	Sexo	0.03048	0.06704 ***	0.02922
	Idade	0.00231 ***	0.00085	0.00149 *
	Educação	0.01121 *	0.01201 *	-0.01261 *
	Transporte:Bicicleta	-0.03820	-0.02613	-0.05238
	Transporte:Moto	-0.05718	-0.06384	-0.16949 ***
	Transporte:Carro	-0.07630 *	-0.04440	-0.11209 ***
	Transporte:Público	-0.07902 ***	-0.10446 ***	-0.07307 **
	Renda	-0.01197	0.01298	-0.02537 **
	Distância	-0.03339 ***	-0.03316 ***	-0.02294 ***

*** indica significância de 0.1%; ** de 1%; * de 5%; . de 10%

Fonte: Dados gerados pelo autor

5.6 Criando uma Superfície de Utilidade

Como segunda análise deste trabalho, entenderemos o processo de criação de uma superfície de utilidade. Esta análise será fundamental a fim de explorar os resultados do modelo ordenado descritos na seção anterior. Com efeito, utilizaremos o valor estimado \hat{U}^* obtido das equações estimadas 22. No entanto, para representarmos uma superfície de utilidade dispoendo apenas de um número finito de \hat{U}_i^* devemos utilizar algum processo que extrapole esses valores para uma superfície no \mathbb{R}^2 . Escolhemos utilizar a técnica de krigagem *Druck et al. (2004)*. As próximas subseções explicam de maneira sucinta a metodologia de krigagem, antes de detalhar

a nossa implementação de superfícies de utilidade.

5.6.1 Introdução

Muitos trabalhos hoje em dia limitam-se a modelar fenômenos da natureza em função de outras variáveis através de metodologias tradicionais conhecidas na literatura. Suponha que estamos interessados em explicar o número de roubos na cidade de Fortaleza. Pesquisadores tenderiam a aplicar análises sofisticadas de regressão para explicar este fenômeno. Porém, alguns fatores importantes não são levados em consideração, os chamados efeitos espaciais.

De uma forma geral existem dados que estão dispostos na forma de coordenadas geográficas. E para utilizá-los de forma efetiva em um ambiente de geoprocessamento, necessitamos de técnicas que aplicam-se em fenômenos distribuídos espacialmente.

Uma das operações mais utilizadas em sistemas de informação geográfica é a apresentação espacial de variáveis como, por exemplo, total da população ou renda por bairro (ou município), ou medidas mais sofisticadas, tais como índice de criminalidade de uma cidade. Grande parte dos trabalhos científicos limitam seu uso de SIG (Sistema de Informação Geográfica) a operações de visualização, tirando conclusões intuitivas. Mas é possível ir muito além. Quando visualizamos um padrão espacial, é muito útil traduzi-lo em considerações objetivas: o padrão que observamos é aleatório ou apresenta uma agregação definida? Esta distribuição pode ser associada a causas mensuráveis? (Almeida, 2004).

Os métodos “convencionais” de análise de dados georreferenciados utilizam dados no formato vetorial, se valendo de ponto, linhas, e polígonos para tentar representar as características físicas/geográficas da área de estudo. O interesse primário ao analisar padrões de distribuição de pontos é determinar se os eventos observados exibem algum padrão sistemático, ao invés de estar distribuídos aleatoriamente. Uma alternativa simples para analisar o comportamento de padrões de pontos é ajustar uma função bi-dimensional através do estimador de *Kernel*. A análise de padrões de área diz respeito a distribuição de eventos cuja localização está associada a áreas (delimitadas por polígonos), assim como no caso de distribuição de pontos, nosso objetivo de análise será determinar a existência de um padrão espacial nos valores observados. Uma das formas de detecção de similaridade entre áreas é através do índice global de Moran I.

Existe uma outra perspectiva de abordagem conhecida como Análise Espacial de Superfícies que, através de alguns valores representativos do fenômeno estudado no mapa, produz uma superfície que aproxima o fenômeno estudado de forma realista. Assim, é necessário modelar a variabilidade espacial que objetivam fazer extrapolações desses valores para regiões não mensuráveis, por meio do fenômeno de autocorrelação espacial em que observações próximas no espaço possuem valores similares. Esse processo é conhecido como Primeira Lei da Geografia, onde valores mais próximos são mais parecidos do que valores mais distantes.

Com base nessa metodologia o objetivo desta capítulo é introduzir ao leitor um

melhor método para analisar a utilidade dos entrevistados no âmbito espacial. Nesse sentido, iniciaremos apresentando técnicas estatísticas para tratamento e análise de dados em superfícies.

5.6.2 *Krigagem: fundamentação teórica*

Os modelos que visam gerar superfícies a partir de procedimento de interpolação, de forma geral, representam a variável em estudo como uma combinação da variabilidade em larga e pequena escala (Druck *et al.*, 2004). Ainda segundo os autores, existem três abordagens:

- Modelos determinísticos de efeitos locais: cada ponto da superfície é estimado apenas a partir da interpolação das amostras mais próximas, utilizando funções, por exemplo, como inverso do quadrado da distância. Neste caso, não é feita qualquer hipótese estatística sobre a variabilidade espacial (Druck *et al.*, 2004).

As funções encontradas na literatura para gerar uma superfície bidimensional são: vizinho mais próximo, média simples, média ponderada e estimadores de densidade não-paramétricos (estimadores de *kernel*). Segundo os autores, os estimadores de *kernel* são uma alternativa viável a métodos mais sofisticados de interpolação, pois não requerem a parametrização da estrutura de correlação espacial, porém, as desvantagens destes estimadores são a forte dependência no raio de busca e a excessiva suavização da superfície.

- Modelos determinísticos de efeitos globais: a suposição implícita nesta classe de interpoladores é que, para a caracterização do fenômeno em estudo, predomina a variação em larga escala, e que a variabilidade local não é relevante.
- Modelos estatísticos de efeitos locais e globais (krigagem): cada ponto da superfície é estimada apenas a partir da interpolação das amostras mais próximas, utilizando um estimador estatístico. Segundo Druck *et al.* (2004), esses estimadores apresentam propriedades de não ser tendenciosos e de procurar minimizar os erros inferenciais.

Como parte de modelo estatísticos, a krigagem compreende um conjunto de técnicas de estimação e predição de superfícies baseada na modelagem da estrutura de correlação espacial. A hipótese implícita no procedimento geoestatístico é que o processo estudado é estacionário (média constante e covariância depende somente de h). A krigagem é considerada uma boa metodologia de interpolação de dados. Utilizando o princípio da Primeira Lei de Geografia, que diz que unidades de análise mais próximas entre si são mais parecidas do que unidades mais afastadas, a krigagem utiliza funções matemáticas e matrizes de variância espacial para acrescentar pesos maiores nas posições mais próximas aos pontos amostrais e pesos menores nas posições mais distantes, e criar assim os novos interpolados com base nessas combinações lineares de dados. Além disso, a krigagem também fornece estimadores com propriedades de não tendenciosidade e eficiência (Druck *et al.*, 2004).

Uma variável regionalizada pode ser expressa pela soma de três componentes, a relação pode ser expressa como: $Z(x) = \mu(x) + \gamma(x) + \varepsilon$ representando a média constante, a correlação espacial e o erro residual, respectivamente.

A hipótese mais simples sobre o comportamento da variável regionalizada é que a média do fenômeno, $\mu(x)$, seja constante na região de estudo, o que implica em não haver variação significativa na larga escala. Esta hipótese dá origem aos interpoladores de *Krigagem ordinária*. No caso de se querer modelar uma tendência, há vários métodos disponíveis: *Krigagem Universal*, *Funções Aleatórias Intrínsecas de Ordem k*, não discutidos neste capítulo.

Segundo [Druck et al. \(2004\)](#), na hipótese da *Krigagem ordinária*, $\mu(x)$ é constante e denotada por m . Deste modo, o valor esperado da função aleatória Z nas posições x e $x + h$ são iguais a m . Isto implica que o valor esperado da diferença entre os valores observados em x e $x + h$, separados por um vetor de distância h , é nulo:

$$E[Z(x) - Z(x + h)] = 0$$

Admite-se também que o fenômeno considerado seja *estacionário de segunda ordem*, isto é, a covariância entre dois pares quaisquer $Z(x)$ e $Z(x + h)$, separados por um vetor distância h , existe e depende somente de h . Então:

$$C(h) = COV[Z(x), Z(x + h)] = E[Z(x) \cdot Z(x + h)] - m^2$$

Adicionalmente, a estacionariedade da covariância implica na estacionariedade da variância:

$$Var(Z(x)) = E[Z(x) - m]^2 = E[Z^2(x)] - 2E[Z(x)] \cdot m + m^2 = E[Z^2(x)] - m^2 = C(0) \quad (23)$$

Deste modo, verifica-se que as hipóteses de média constante e estacionariedade da covariância implicam que a determinação da função $C(h)$ é suficiente para caracterizar a *variável regionalizada*. Para determinar $C(h)$, utiliza-se uma função auxiliar, chamada de função *variograma* $2\gamma(h)$, definida por:

$$2\gamma(h) = E[Z(x) - Z(x + h)]^2 \quad (24)$$

desenvolvendo 24 obtemos a seguinte relação:

$$2\gamma(h) = E[Z^2(x)] - 2E[Z(x)Z(x + h)] + E[Z^2(x + h)] \quad (25)$$

da equação 23 tem-se:

$$E[Z^2(\mathbf{x})] = E[Z^2(\mathbf{x} + \mathbf{h})] = C(0) + m^2 \quad (26)$$

e,

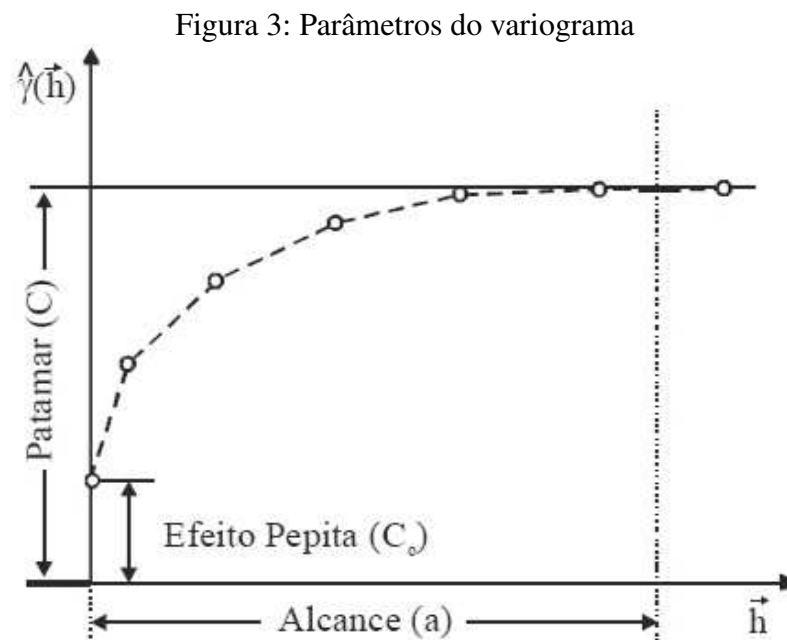
$$E[Z(\mathbf{x})Z(\mathbf{x} + \mathbf{h})] = C(\mathbf{h}) + m^2 \quad (27)$$

fazendo as substituições de 26 e 27 em 25, obtemos:

$$2\gamma(\mathbf{h}) = 2C(0) - 2C(\mathbf{h}) \text{ ou } \gamma(\mathbf{h}) = C(0) - C(\mathbf{h}) \quad (28)$$

onde:

$\gamma(\mathbf{h})$ representa o *semivariograma*, que é a metade do *variograma*. A relação 28 pode ser vista no gráfico a seguir:



Fonte: Figura extraída do trabalho de Druck *et al.* (2004)

O semivariograma, como pode ser visto na figura 3, mostra a medida do grau de dependência espacial entre amostras ao longo de um suporte específico. Para sua construção, são usados simplesmente as diferenças ao quadrado dos valores obtidos, assumindo-se uma estacionaridade nos incrementos. Isso significa que o semivariograma é uma medida da variabilidade espacial condicionada pela distância.

Os parâmetros do semivariograma são:

- *Alcance (a)*: distância máxima onde existe autocorrelação espacial das variáveis.

- *Patamar* (C): é o valor do semivariograma correspondente a seu *alcance* (a). Neste ponto não existe mais nenhuma correlação espacial entre as amostras, sendo assim a variância do conjunto de dados.
- *Efeito Pepita* (C_0): o *efeito pepita* é o valor da semivariância para a distância zero e representa a componente da variabilidade espacial que não pode ser relacionado com uma causa específica (variabilidade ao acaso).

Ao ajustarmos o semivariograma empírico é necessário ajustá-lo a um modelo teórico, ou seja, uma função que definirá os parâmetros do semivariograma, “efeito pepita”, “alcance” e “patamar”. Existem três formas de modelar o semivariograma, os modelos mais utilizados são: Modelo esférico, Modelo exponencial e Modelo gaussiano.

Em resumo, os semivariogramas dos modelos básicos são definidos como:

- Modelo Esférico de Semivariograma

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[\frac{3}{2} \left(\frac{|h|}{a} \right) - \frac{1}{2} \left(\frac{|h|}{a} \right)^3 \right], & 0 < |h| < a \text{ e } C_1 = C - C_0 \\ C, & |h| > a \end{cases}$$

- Modelo Exponencial de Semivariograma

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[1 - \exp \left(-\frac{|h|}{a} \right) \right], & |h| \neq a \end{cases}$$

- Modelo Gaussiano de Semivariograma

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[1 - \exp \left(-\frac{|h|}{a} \right)^2 \right], & |h| \neq a \end{cases}$$

5.6.3 Krigagem ordinária

O método de krigagem foi desenvolvido para solucionar problemas de mapeamentos geológicos, mas seu uso expandiu-se com sucesso no mapeamento de solos, mapeamento hidrológico, mapeamento atmosférico e outros campos correlatos.

A metodologia geoestatística apresenta diversas técnicas de estimativas disponíveis e a mais usual é a krigagem ordinária (Landim, 2006). Além disso, segundo Issaks e Srivastava (1989) apud Jakob (2002): “A krigagem ordinária é geralmente associada como BLUE (*best linear unbiased estimator*). A krigagem ordinária é linear porque suas estimativas são combinações lineares ponderadas dos dados disponíveis; é não-viciada porque busca o valor de erro ou resíduo média igual a 0; e é melhor porque minimiza a variância dos erros”.

A diferença entre a krigagem e outros métodos de interpolação é a maneira como os pesos são atribuídos às diferentes amostras (Druck *et al.*, 2004). A krigagem, entendida como um estimador que se baseia em técnicas de análise de regressão, procura minimizar a variância estimada a partir de um modelo prévio que leva em conta a dependência estocástica entre os dados distribuídos no espaço.

Considere uma superfície sobre a qual se observe algum fenômeno a ser estudado, Z , em n pontos distintos, com coordenadas representadas pelo vetor \mathbf{x} . Assim, tem-se um conjunto de valores $\{z(\mathbf{x}_i), i = 1, 2, \dots, n\}$, onde \mathbf{x}_i representa uma posição em duas dimensões (x_i, y_i) . Nós estamos interessados em estimar o valor de $Z(\mathbf{x}_0)$ no ponto \mathbf{x}_0 , através de uma combinação linear dos n valores observados:

$$Z^*(\mathbf{x}_0) = \sum_{i=1}^n \lambda_i Z(\mathbf{x}_i) \quad (29)$$

O objetivo da krigagem é procurar pelo conjunto ótimo de ponderadores de modo que a variância do erro de estimativa seja a menor possível sujeita a $\sum \lambda_i = 1$ (Landim, 2006). Ou seja, minimizando a variância do erro ($\text{Var}[Z(X_0) - Z^*(X_0)]$), sujeito a $\sum_{i=1}^n \lambda_i = 1$. Podemos obter o seguinte sistema de equações, denominado sistema de krigagem ordinária (Druck *et al.*, 2004):

$$\begin{cases} \sum_{j=1}^n \lambda_j C(x_i, x_j) - \alpha = C(x_i, x_0) \text{ para } i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases} \quad (30)$$

onde, $C(x_i, x_j)$ e $C(x_i, x_0)$ são, respectivamente, a semivariância entre os pontos x_i e x_j e entre os pontos x_i e x_0 . O multiplicador de Lagrange, α , é necessário para a minimização da variância do erro.

Por simplificação, podemos tratar o sistema 30 como:

$$[X][A] = [Y]$$

assumindo que $[X]$ seja invertível, então:

$$[X]^{-1}[X][A] = [X]^{-1}[Y] \Rightarrow [A] = [X]^{-1}[Y]$$

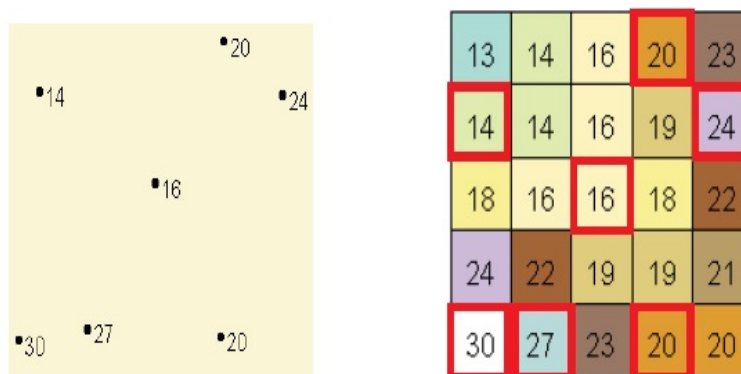
Dessa forma encontramos a seguinte relação:

$$\lambda = C^{-1}c = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} c_{10} \\ c_{20} \\ \vdots \\ c_{n0} \\ 1 \end{bmatrix}$$

os elementos das matrizes contém os valores obtidos no semivariograma referentes às distâncias entre as amostras¹³.

¹³Para melhor entendimento da elaboração do semivariograma, bem como o procedimento do cálculo dos valores de $C(x_i, x_j)$ e $C(x_i, x_0)$ veja o trabalho de Landim (2006)

Figura 4: Processo de interpolação: amostras e aproximação da superfície por uma grade regular



Fonte: Figura extraída de ESRI (2001)

Ou seja, o estimador obtido via krigagem se baseia em técnicas de análise de regressão, procura minimizar a variância estimada a partir de um modelo prévio que leva em conta a dependência estocástica entre os dados distribuídos no espaço (Landim, 2006).

Para exemplificar o entendimento do método via krigagem, considere a Figura 4. A figura a esquerda contém os valores mensurados. Note que dispomos de alguns valores sob a superfície que desejamos estimá-la. Com base nas informações preliminares e através de técnicas da geoestatística (krigagem) podemos estimar valores desconhecidos. Note que a figura a direita, contém as informações reais e, também, as informações obtidos por meio dessas técnicas. Veja que os valores estimados são próximos dos valores reais, justificando a primeira lei da geografia e a correlação espacial.

5.6.4 Criação das superfícies de indiferença de acessibilidade

Como segunda parte desta pesquisa descreveremos o processo de criação de uma superfície de utilidade. O processo de construção deve-se ao auxílio do *software* ArcGis. Como foi descrito no capítulo 4, a amostra consiste em 4030 entrevistados, que responderam a um questionário. Tivemos alguns erros com o endereços destes entrevistados, sendo estes retirados da amostra. Dos 4030 entrevistados restaram 3324 pessoas com seus endereços corretamente geocodificadas, sendo esses nosso público alvo para fazer as estimações via modelos ordenados. Vale lembrar, como mencionamos no capítulo 4, que o tamanho da amostra variará para cada modelo em função dos *missing*.

Utilizamos o primeiro modelo estimado baseado na hipótese da regressão paralela para estimar a utilidade com o objetivo de facilitar a interpretação. Através desta estimacão para os cinco modelos propostos, encontramos através de $\hat{U}_i = \mathbf{x}_i \hat{\beta}$ a utilidade estimada. Com essa utilidade estimada e as coordenadas dos endereços, construímos a superfície via a metodologia de *Krigagem*, através do modelo Esférico no qual é o mais comum utilizado na literatura. Os principais achados com relação ao uso da geoestatística para a estimacão de uma superfície de utilidade em torno da cidade de Fortaleza/CE será descrita a seguir.

5.6.5 Resultados

O intuito desta análise é verificar onde são os maiores níveis de utilidade dos entrevistados com relação ao cinco destinos considerado nesta pesquisa, bem como constatar que o indivíduo possui bons acessos aos destinos estudados, mesmo esses situando a uma distância maior do que aqueles que residem em áreas mais próximas. Ou seja, o interessante desta análise é perceber que a distância é um fator importante, isto é evidente, mas indivíduos que residem em áreas mais afastas, por possuir características mais favoráveis (características podem ser socioeconômicas) e algumas medidas que não são medidas nesse estudo (como estrutura das ruas e nível de trânsito) possuirão o mesmo nível, ou maior, de bem-estar daqueles que estão próximo do destino considerado, pois essas medidas são incorporadas pelos indivíduos quando estão diante de uma estrutura de escolha ordenada.

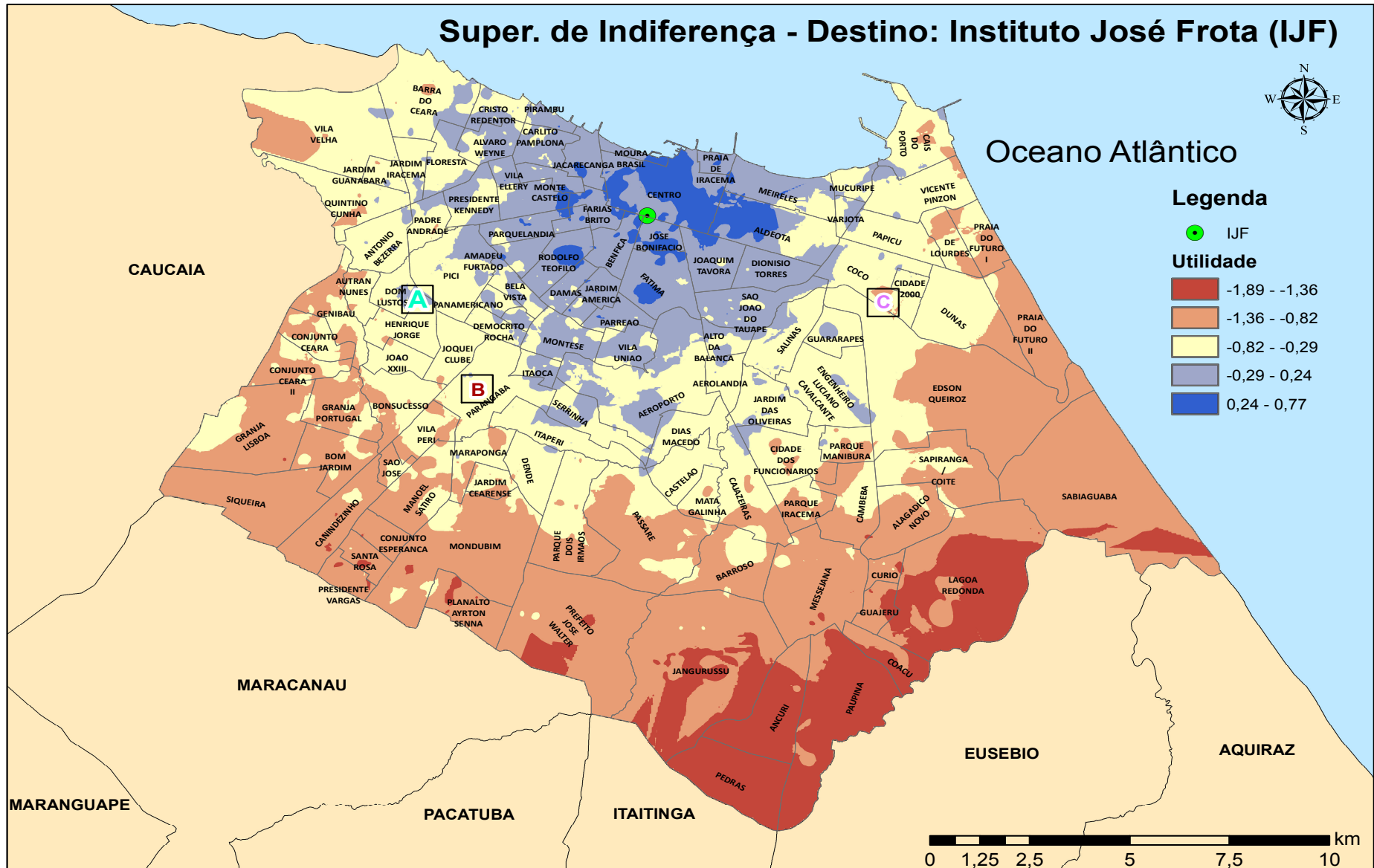
Com efeito, vemos claramente, através das figuras a seguir, que quanto mais próximo¹⁴ do local de destino maior é a utilidade da pessoa com relação ao processo de deslocamento de sua residência a esse ponto. Obviamente, a distância é um dos principais motivos para levarmos em consideração quando estamos falando de acesso. Nota-se que pessoas morando nas zonas de fronteira da cidade possuem os piores acessos. Não obstante, o intuito deste trabalho não é somente mostrar que a distância é importante, mas que outros fatores (características dos indivíduos) podem explicar o quanto que o indivíduo classifica esse deslocamento. Ou seja, indivíduos que moram mais afastados do destino, mas que possuem uso de carro próprio, possuem alto nível de renda e residindo em áreas mais privilegiadas, estes possuem o mesmo nível de utilidade daqueles que moram próximos, como nota-se nas figuras a seguir.

Para melhor entendimento, veja a figura 5, observe os três pontos “A”, “B” e “C”. Esses três pontos estão equidistantes, aproximadamente 6.3km, com relação ao hospital IJF. Mas, observamos que os três pontos estão em curvas de utilidade distintas. O ponto “A” está na curva superior a “B”, e “B” está na curva superior a “C”. Dessa maneira, observamos que alguns fatores não mensuráveis (por exemplo, fluxo de carros e condição das vias) e alguns mensuráveis (por exemplo, o uso de carro próprio) são incorporados pelo indivíduo para que ele classifique o nível de sua utilidade com relação ao processo de deslocamento ao IJF.

Considere agora a figura 7, os dois pontos “A” marcado no mapa que se encontram a uma distância de, aproximadamente, 6.5km do Estádio Castelão e o outro ponto “B” situado a uma distância de, aproximadamente, 2.5km do mesmo. Esses três pontos marcados no mapa, estão diferindo na distância, mas possuindo o mesmo nível de utilidade.

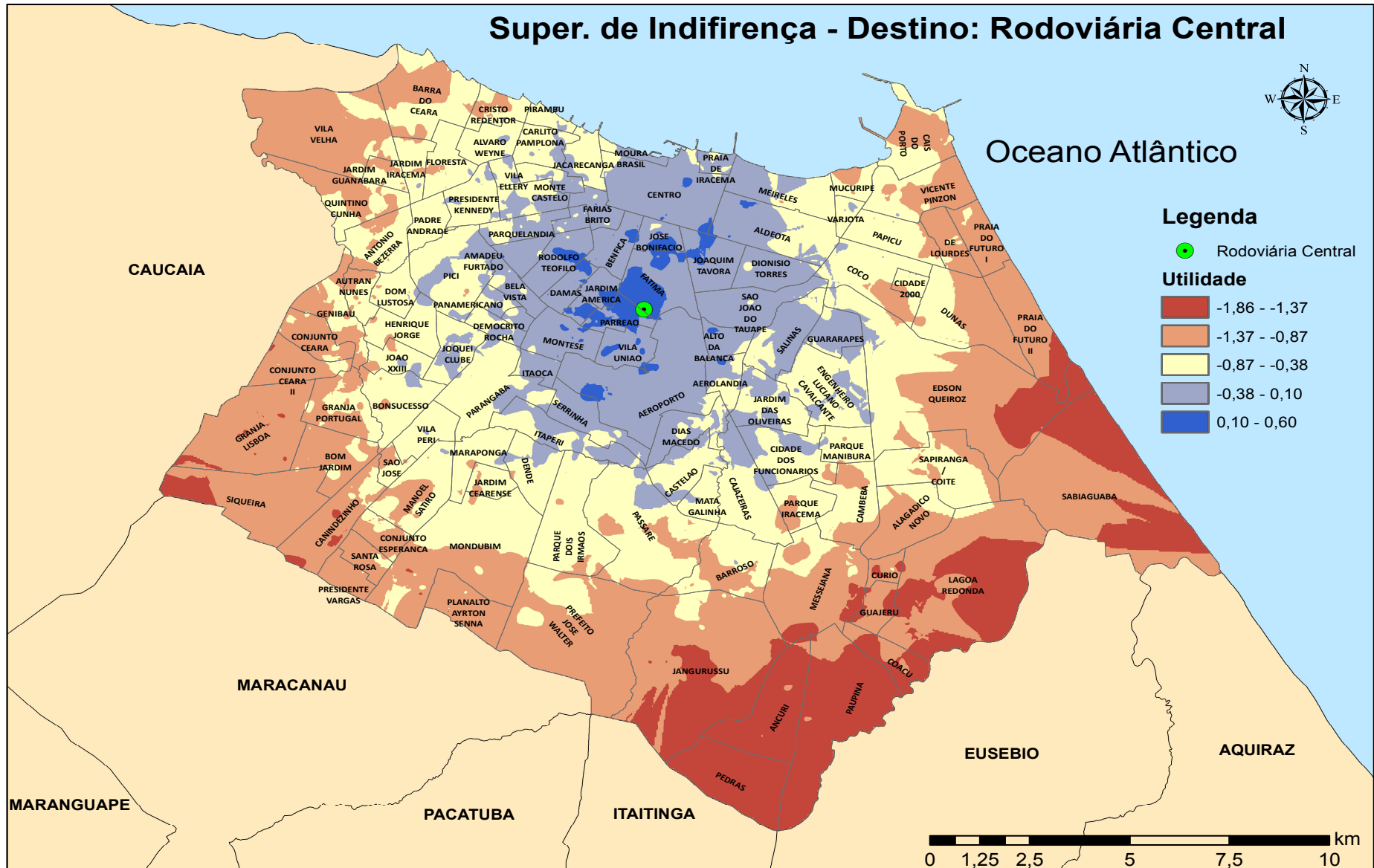
¹⁴Quanto mais azul melhor será o nível de utilidade, quanto mais vermelho pior o nível de utilidade

Figura 5: Superfície interpolada da utilidade estimada para o IJF



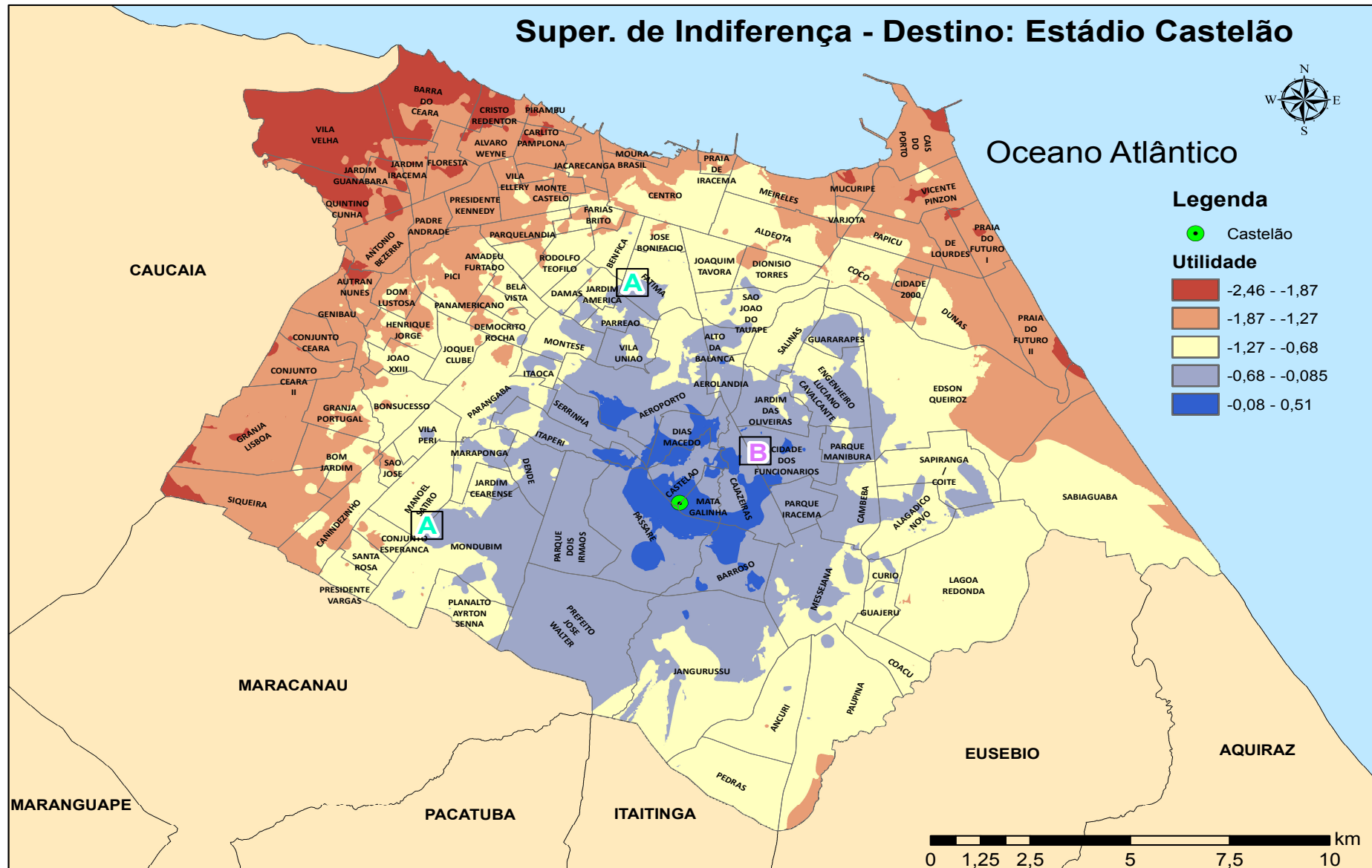
Fonte: Dados gerados pelo autor

Figura 6: Superfície interpolada de utilidade estimada para a Rodoviária



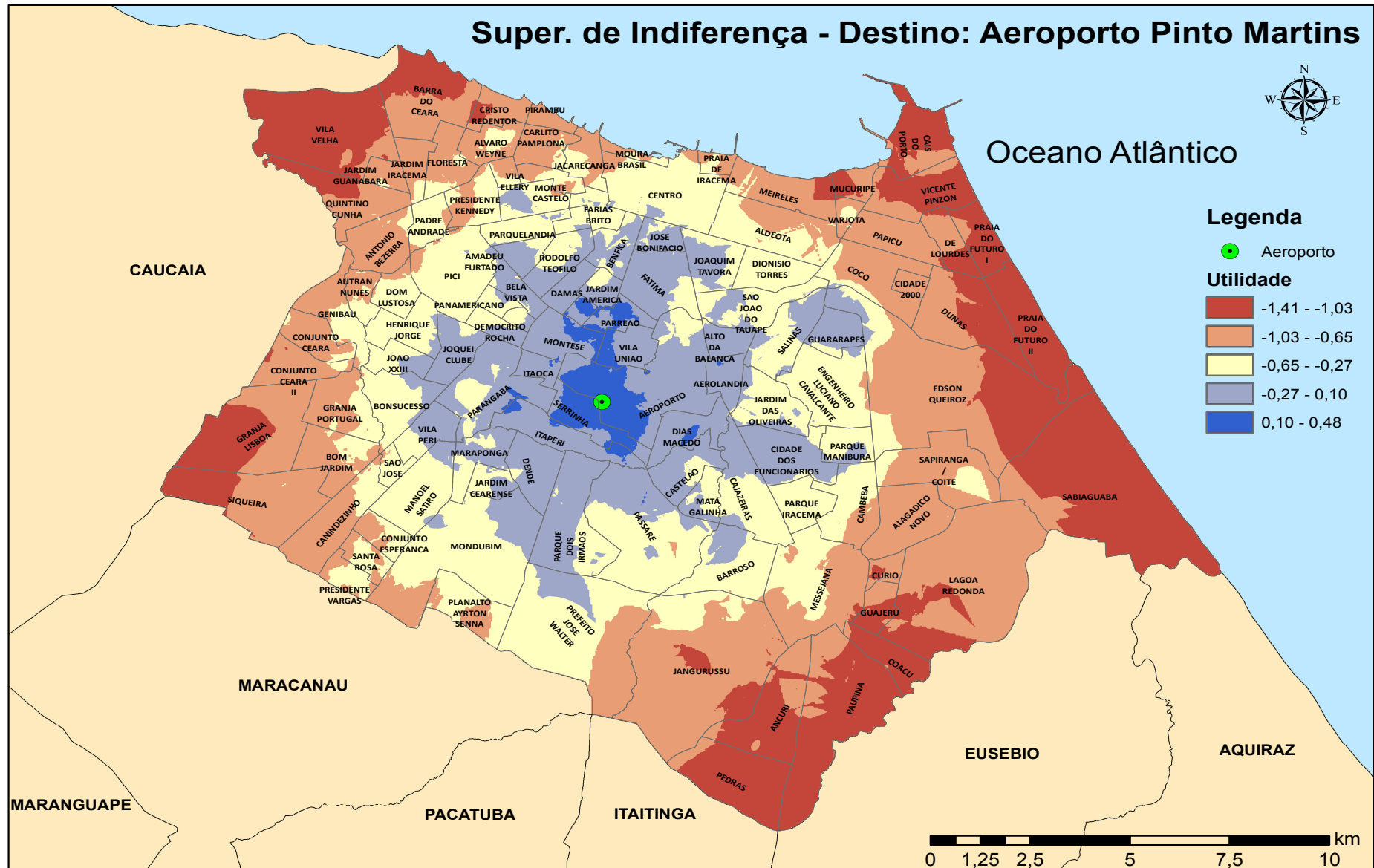
Fonte: Dados gerados pelo autor

Figura 7: Superfície interpolada de utilidade estimada para o Castelão



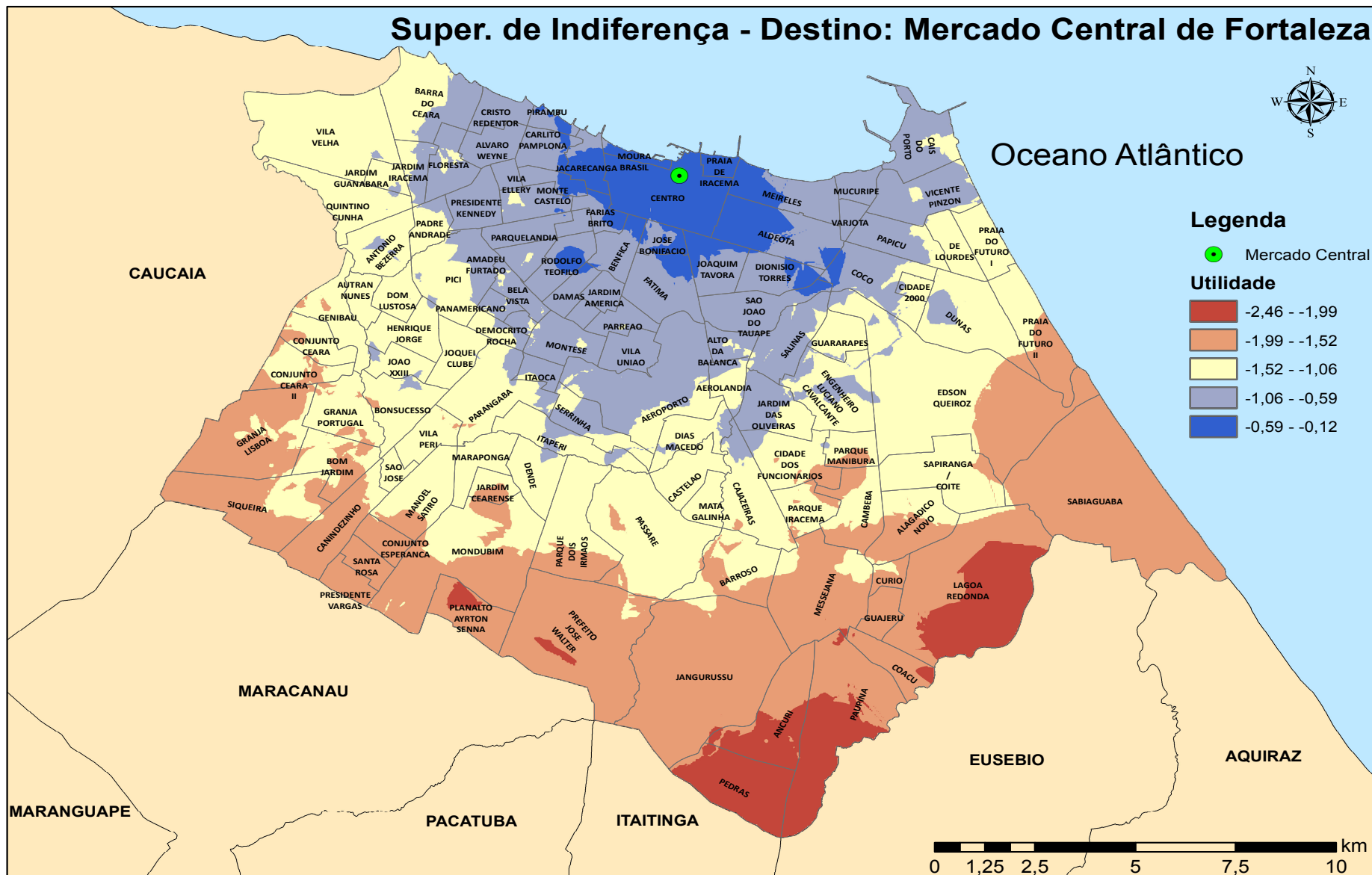
Fonte: Dados gerados pelo autor

Figura 8: Superfície interpolada de utilidade estimada para o Aeroporto



Fonte: Dados gerados pelo autor

Figura 9: Superfície interpolada de utilidade estimada para o Mercado



Fonte: Dados gerados pelo autor

6 CONSIDERAÇÕES FINAIS

O presente estudo analisou a opinião dos entrevistados com relação a acessibilidade na cidade de Fortaleza/CE observando aspectos microeconômicos que afetam o bem-estar do indivíduo ao se deparar com seu trajeto para alguns locais estratégicos. Buscamos oferecer contribuições acerca das características individuais e locais na probabilidade da pessoa ter um custo de acessibilidade por respostas ordenadas entre sua residência e o local de destino. Para este fim, utilizou-se como *proxy* de avaliação da acessibilidade a variável descrita no banco de dados da pesquisa realizada com apoio da FUNCAP junto ao centro CAEN/UFC no desenvolvimento do projeto intitulado: “Montagem de uma Base de Dados Longitudinal de Vitimização do Ceará: Aspectos Socioeconômicos e Espaciais” em 2012.

É de ser considerado a falta de pesquisas nacionais e internacionais para avaliar a acessibilidade no aspecto micro nas perspectivas de modelos de utilidade. Como já mencionado e ressaltado nos trabalhos de [Ben-Akiva e Lerman \(1978\)](#) e [Koenig \(1980\)](#) a maioria dos estudos sobre acessibilidade é voltada aos aspectos de infraestrutura e da distribuição espacial dos serviços ou aspectos de transporte público, não levando em consideração o comportamento do indivíduo. O uso de modelos que buscam medir como indivíduos se comportam diante de um conjunto de escolhas é fundamental nesses estudos. Nesse sentido a classe de modelos de escolha discreta é fundamental.

Buscamos entender como os indivíduos analisam, através de modelos de utilidades propostos por [Ben-Akiva e Lerman \(1978\)](#), o trajeto ao sair de suas casas até a chegada à porta do serviço. Com efeito, buscamos entender como o indivíduo analisa o processo de deslocamento, ou seja, ao sair de sua residência o indivíduo: desloca-se muito até a parada de ônibus, ao subir no ônibus o entrevistado usufrirá do sistema de transporte com relação a conforto e a lotação, o sistema de trânsito nesse trajeto é relativamente bom ou existe congestionamento, e, para aqueles que usam transporte próprio, será que o trajeto por onde o indivíduo passa é confortável com relação a estrada. Então todas essas características estão sendo contabilizadas quando o entrevistado responde a esta pergunta: “Gostaria agora que o(a) Sr(a) pensasse sobre o que acontece quando é preciso ir a locais para resolver problemas ou aproveitar o tempo. Considerando que acessibilidade envolve tudo aquilo que é importante para ir e voltar (tempo de percurso, custo de transporte, conforto da jornada, horário de funcionamento, condições das ruas, estradas, entre outros), como o(a) Sr(a) classificaria o acesso aos seguintes locais”. Para este fim, analisou-se cinco destinos estratégicos de Fortaleza/CE: Hospital IJF-Instituto José Frota, Rodoviária Central de Fortaleza, Estádio Castelão, Aeroporto Internacional Pinto Martins e Mercado Central de Fortaleza. Considerou-se inicialmente, como a escala de classificação ao acesso: 1-Péssimo, 2-Ruim, 3-Razoável, 4-Bom e 5-Excelente. Porém, os resultados não foram satisfatórios e tivemos dificuldade de interpretar optando pela agrupamento das categorias em: 1:Péssimo (correspondente as categorias péssimo e ruim), 2-Razoável, 3-Excelente (correspondente as categorias bom e excelente).

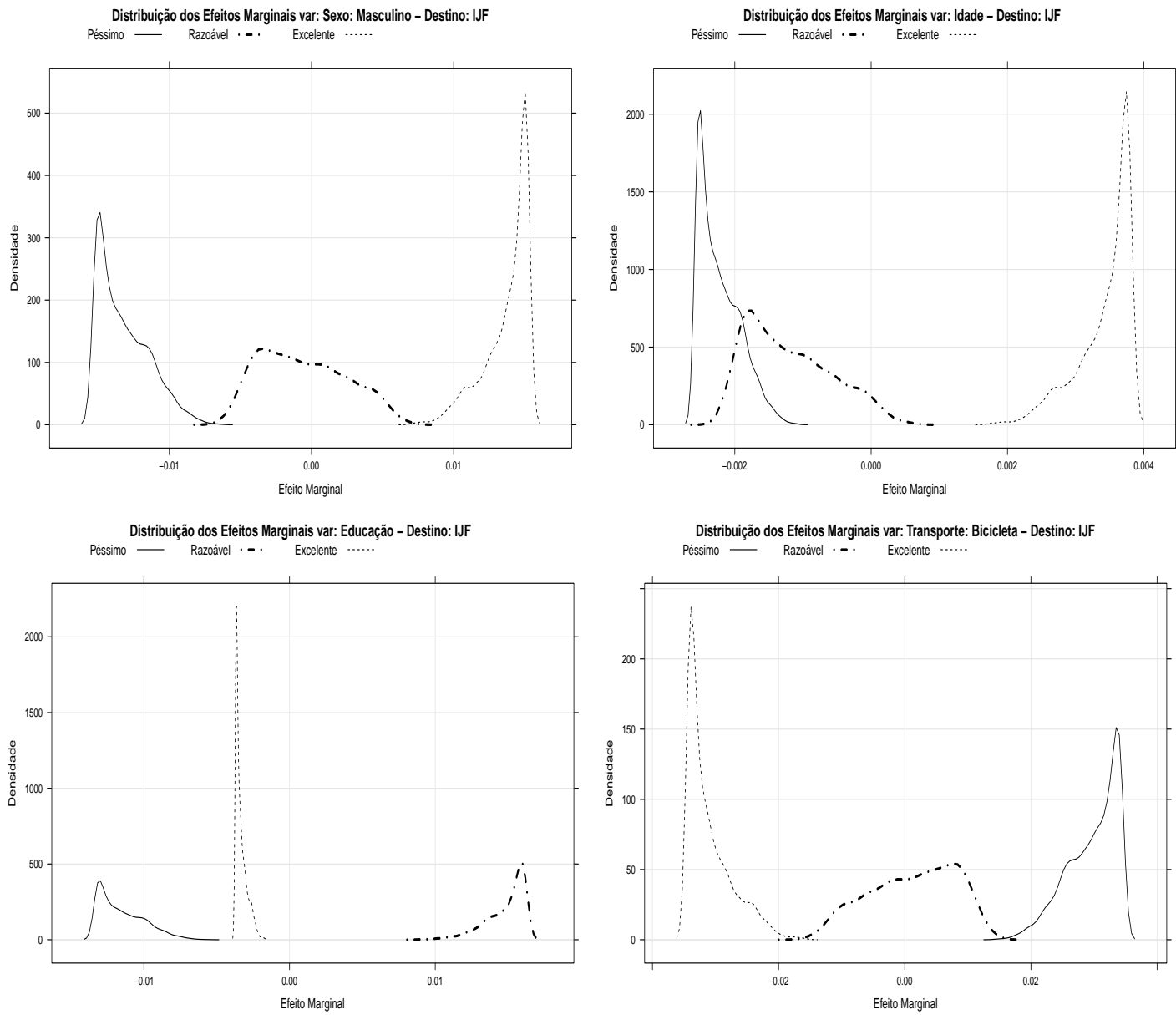
Os resultados encontrados nesse estudo mostram que a distância é um fator crucial no nível de acessibilidade dos entrevistados. Para ver melhor esse resultado, utilizamos a metodologia da análise de interpolação via krigagem encontrada na geoestatística. Para encontrar esse resultado utilizamos o valor preditivo da estimação da utilidade que cada indivíduo possui. Dada a informação georeferenciada das residências do entrevistado e dos destinos, podemos criar uma superfície dessa utilidade em Fortaleza. Com efeito, podemos ver que os maiores níveis de utilidade são para aqueles pontos que estão próximos ao destino, enquanto que áreas mais afastadas do destino analisado tendem a ter menores níveis de utilidade. Mas, podemos encontrar outros resultados importantes através desta análise. Podemos ver que em algumas regiões de Fortaleza existem locais mais afastados do destino que possuem um nível de utilidade igual aqueles mais próximos. Tal resultado pode ser explicado pela análise do trajeto e pelas características dos entrevistados, pois esses indivíduos possuem acesso mais simples (seja pelo uso de transporte próprio ou pela ótima via que trafegam), trânsito sem congestionamentos ou pela ótima localização onde mora.

Por exemplo, em relação ao acesso ao Castelão, vemos que os bairros vizinhos ao estádio estão em melhores condições de acesso, porém, vemos que bairros mais distantes (por exemplo, Guararapes, aproximadamente 7 km do destino) tem a mesma curva de utilidade daqueles que estão a aproximadamente 3 km.

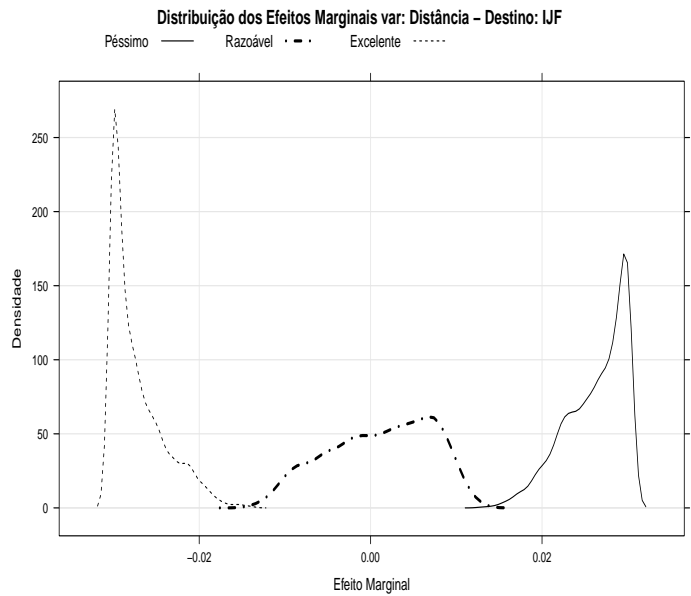
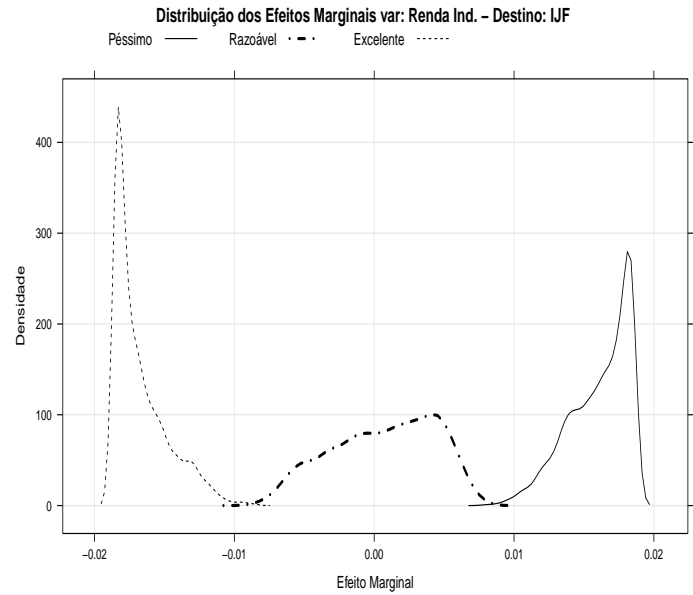
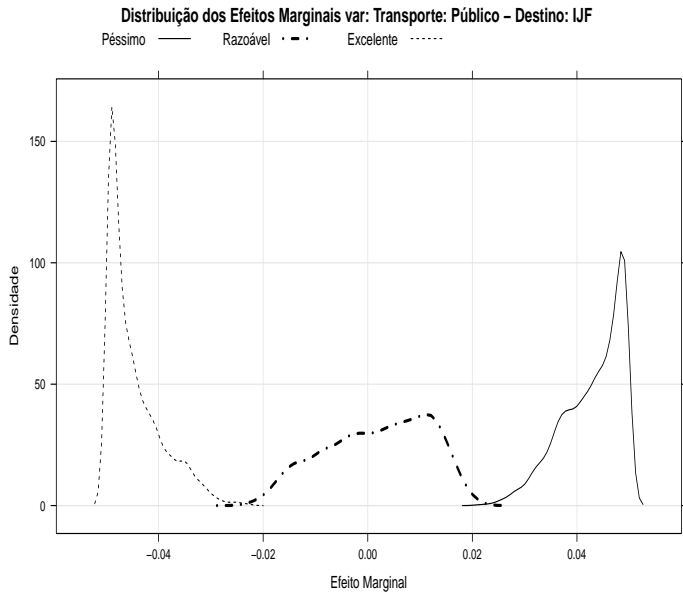
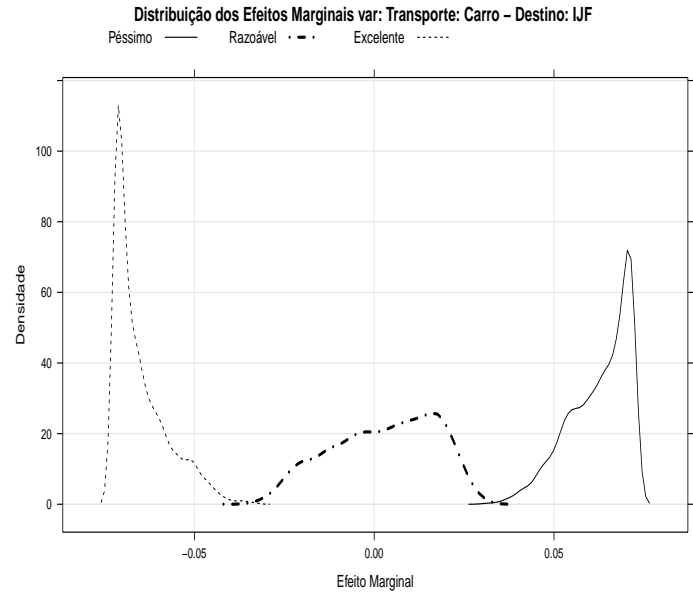
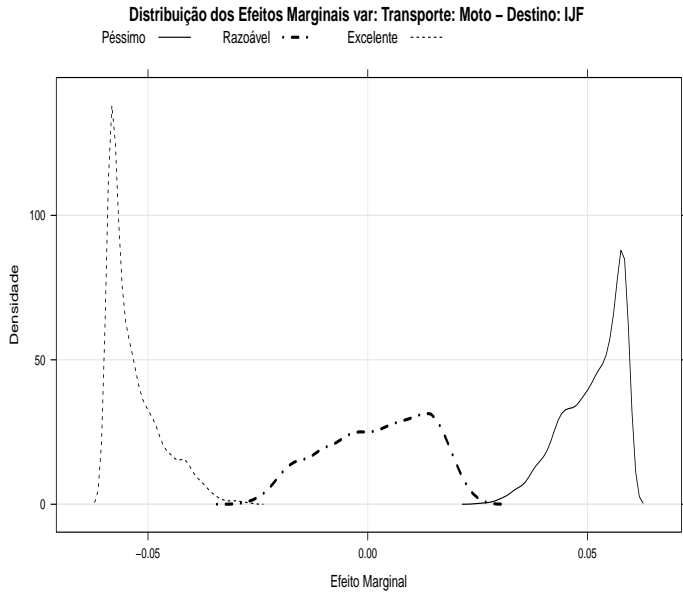
Para pesquisa futuras pretendemos utilizar modelos bayesianos espaciais. Utilizando como evidência empírica essa base de dados juntamente com esta metodologia complexa, tentaremos comparar/melhorar o entendimento da acessibilidade. Outra possível extensão é através dos modelos mencionados na seção 2.2.3, mais precisamente os modelos *space-time* propostos por Miller (1991). A fim de analisar a acessibilidade em perspectivas tradicionais e modelos amplamente utilizados na literatura tentaremos entender melhor o processo de modelos gravitacionais propostos por Hansen (1959).

A GRÁFICOS

Figura 10: Destino IJF: Distribuição empírica dos efeitos marginais individuais

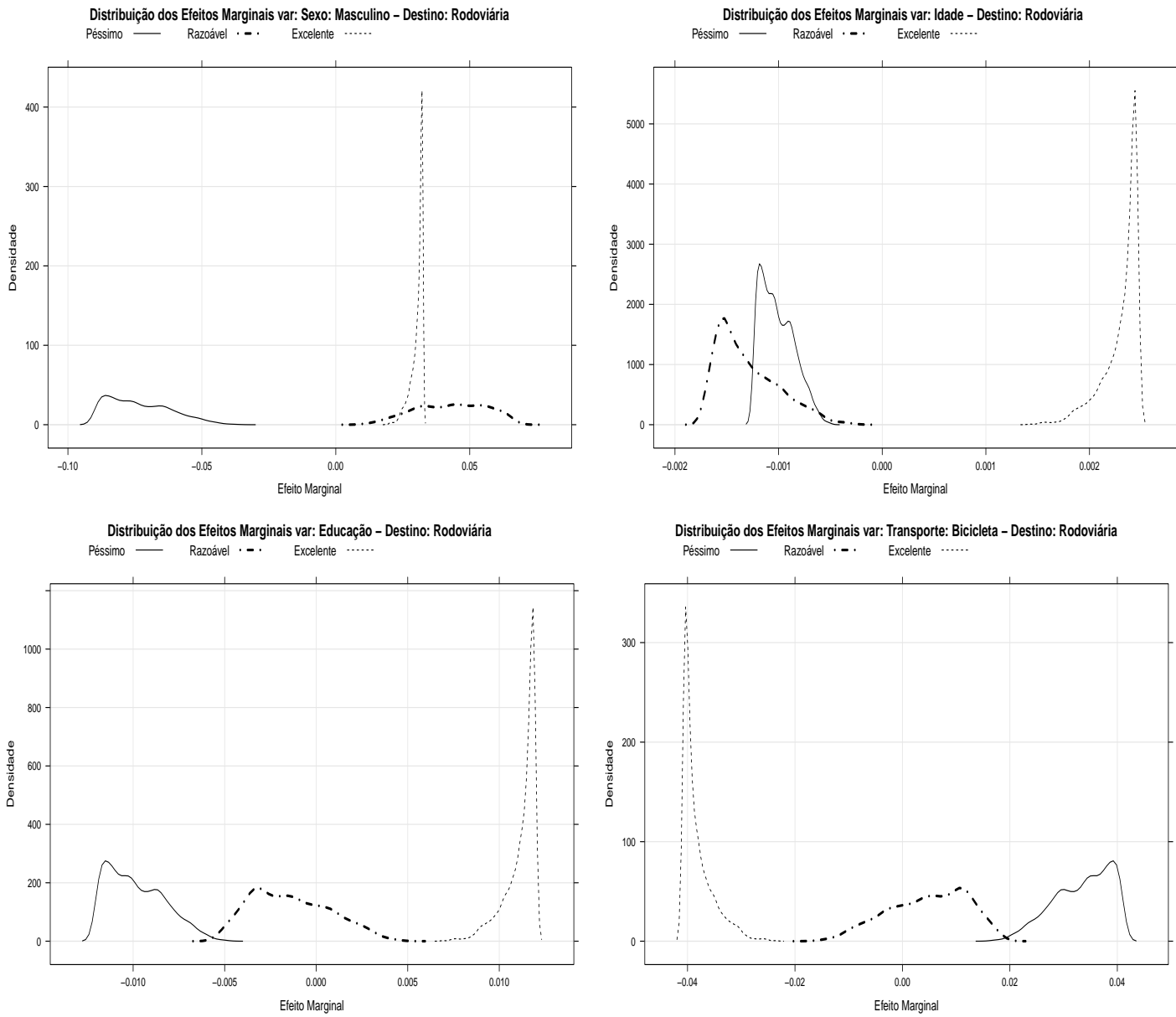


Fonte: Dados gerados pelo autor

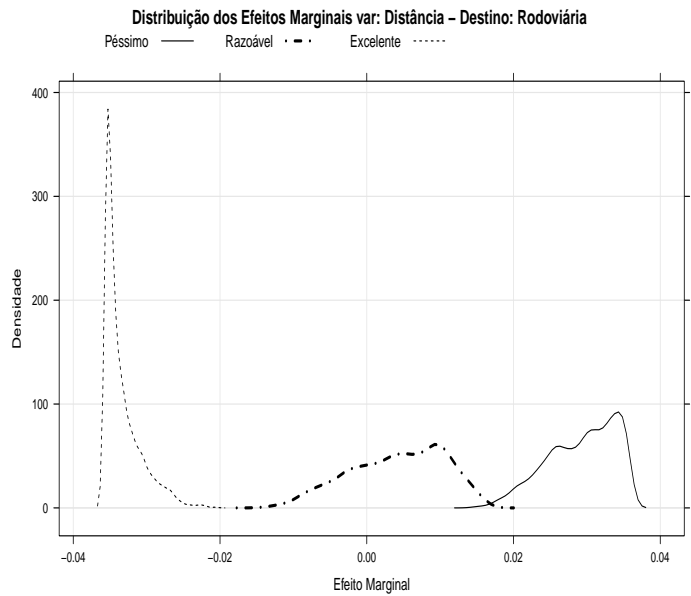
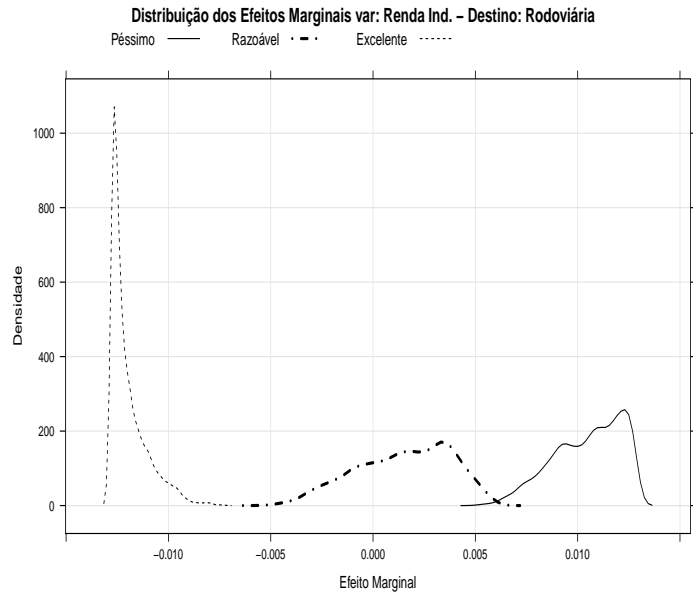
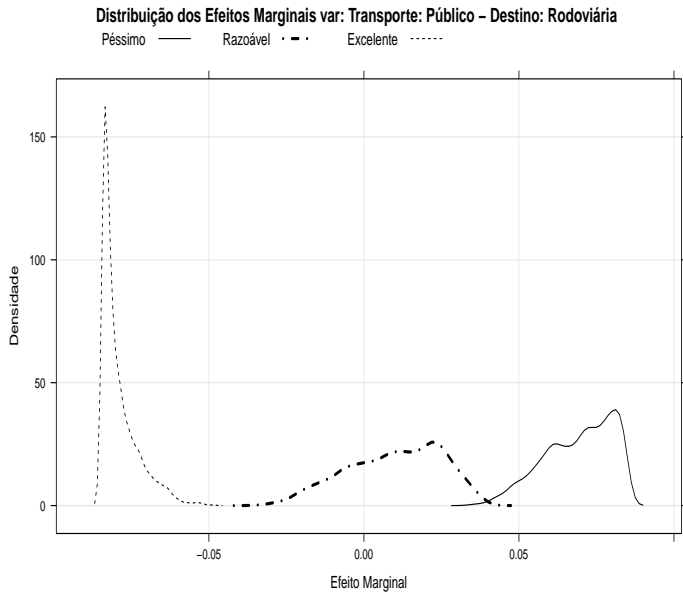
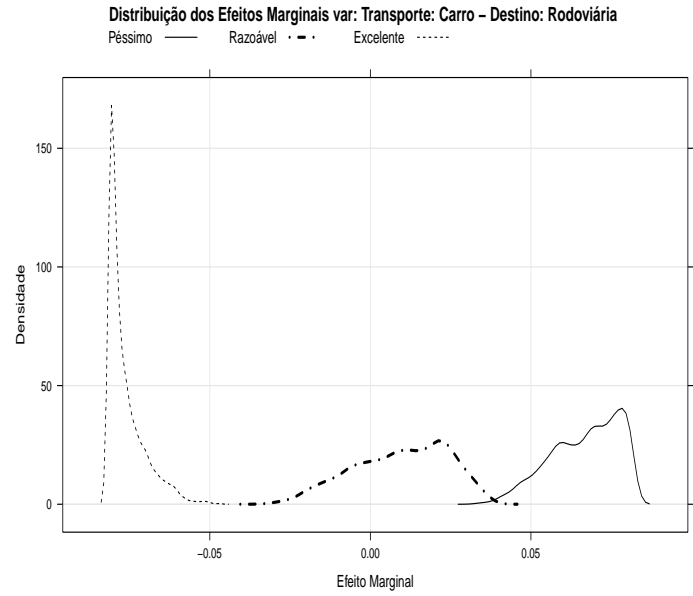
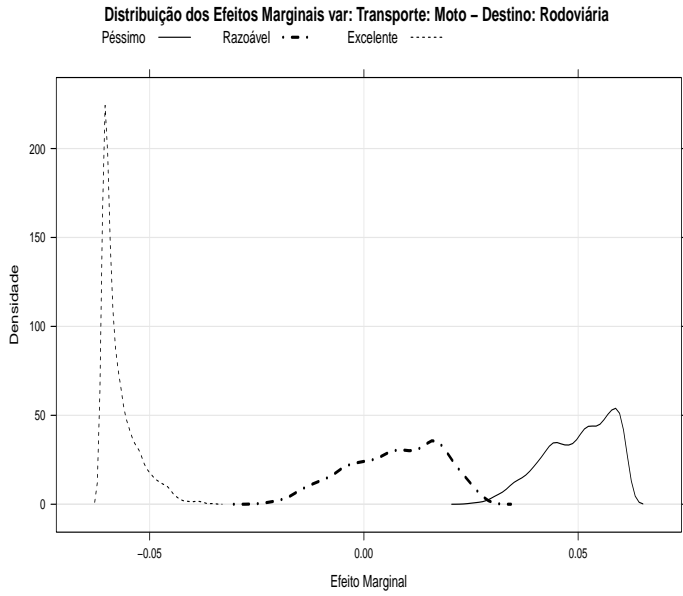


Fonte: Dados gerados pelo autor

Figura 11: Destino Rodoviária de Fortaleza: Distribuição empírica dos efeitos marginais individuais

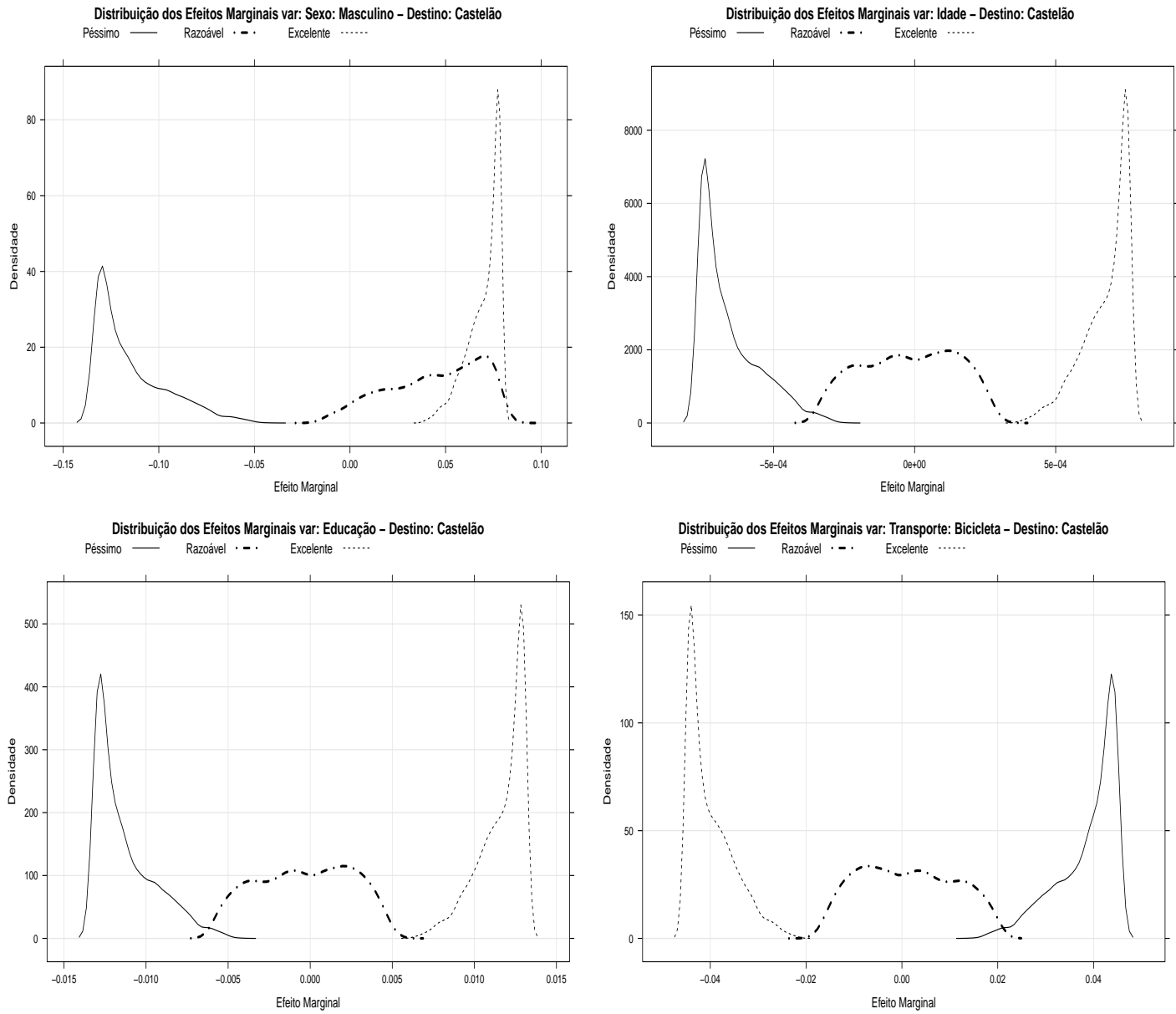


Fonte: Dados gerados pelo autor



Fonte: Dados gerados pelo autor

Figura 12: Destino Castelão: Distribuição empírica dos efeitos marginais individuais



Fonte: Dados gerados pelo autor

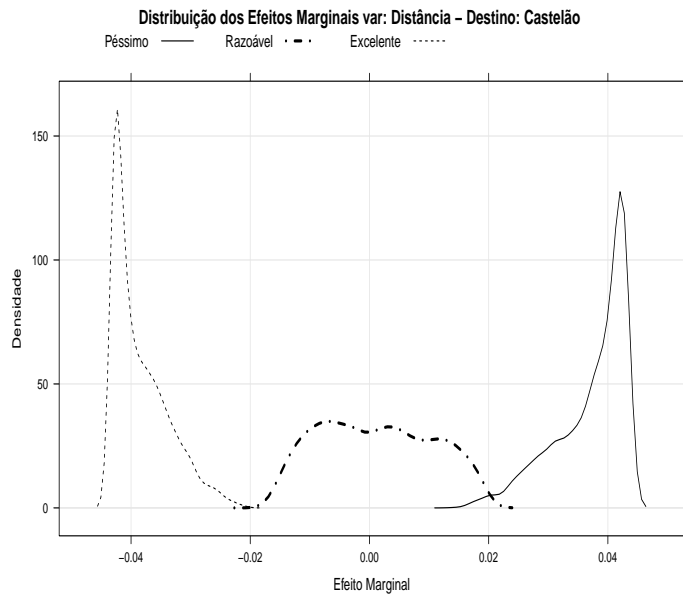
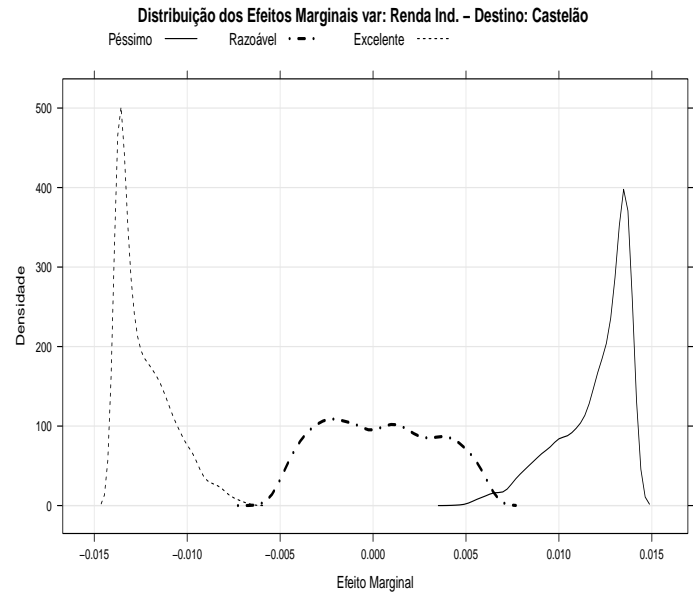
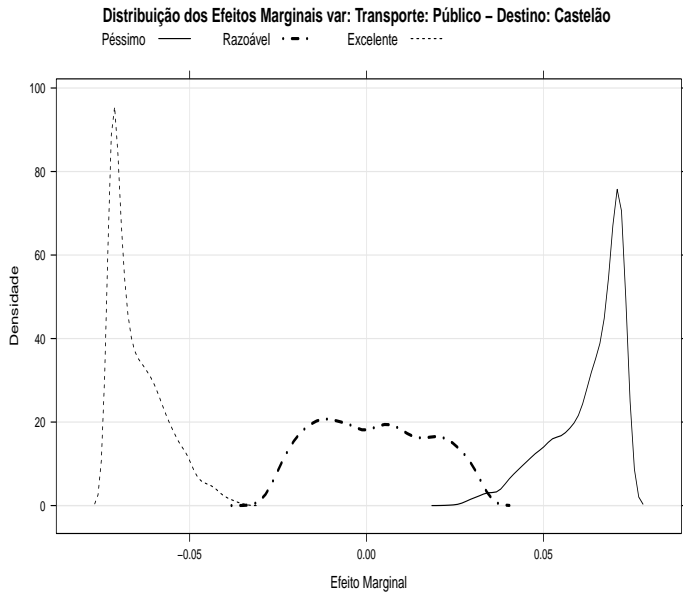
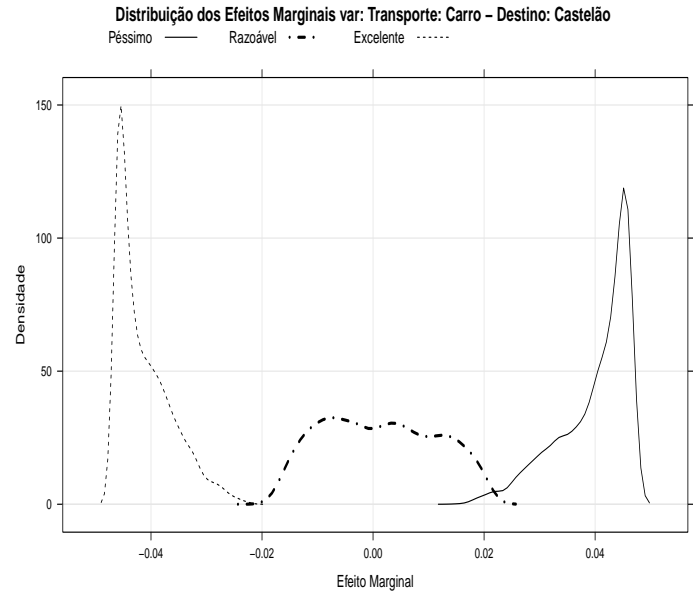
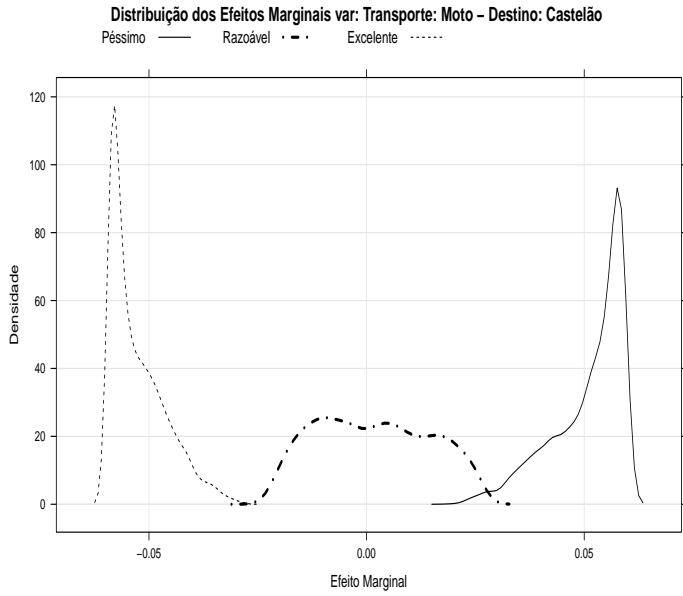
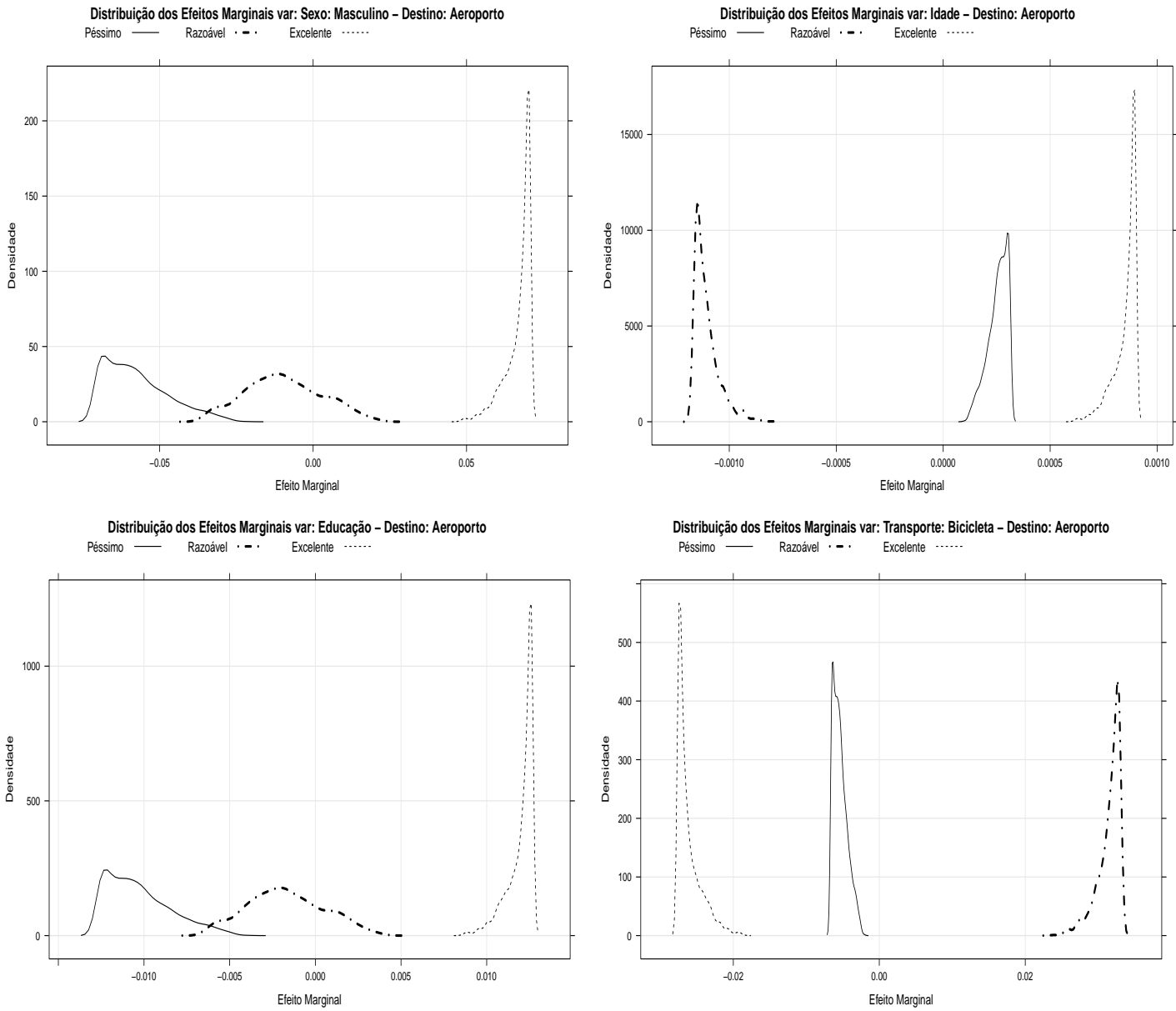
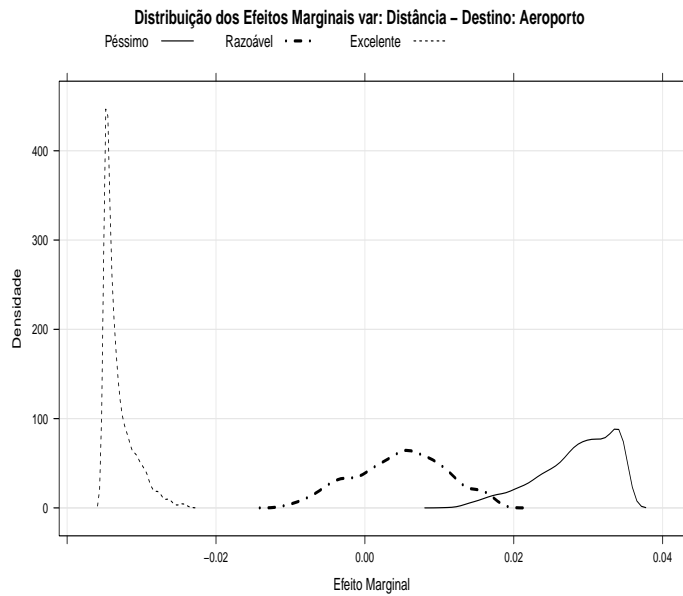
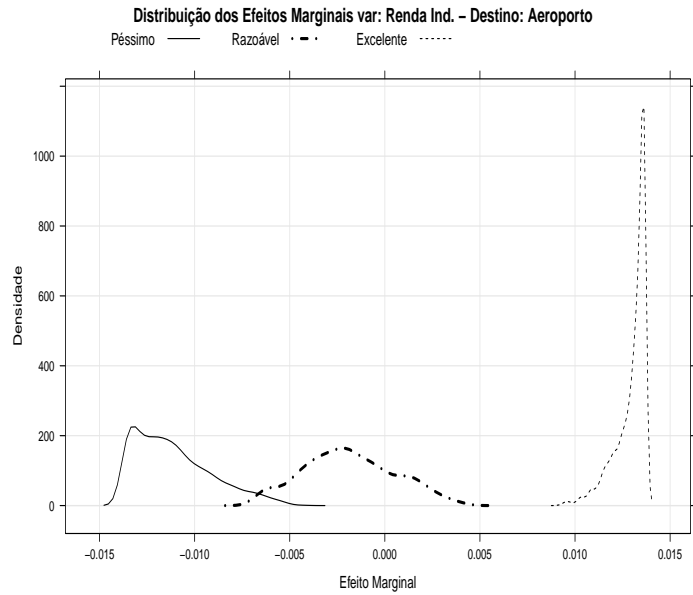
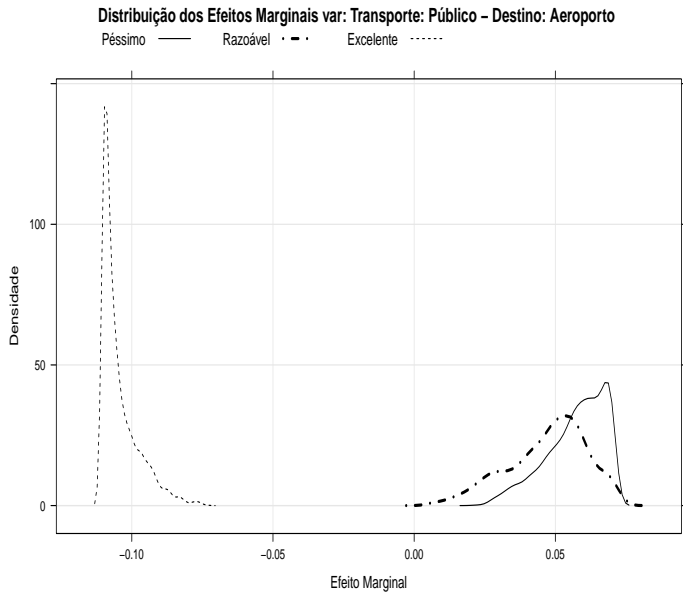
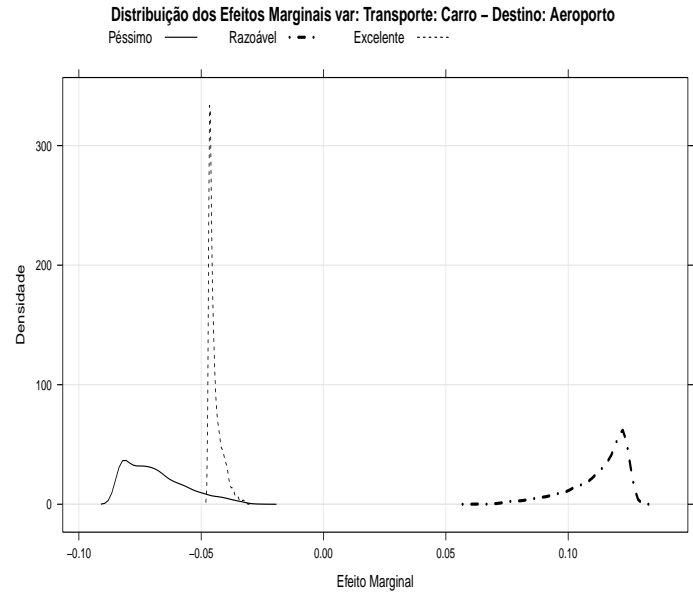
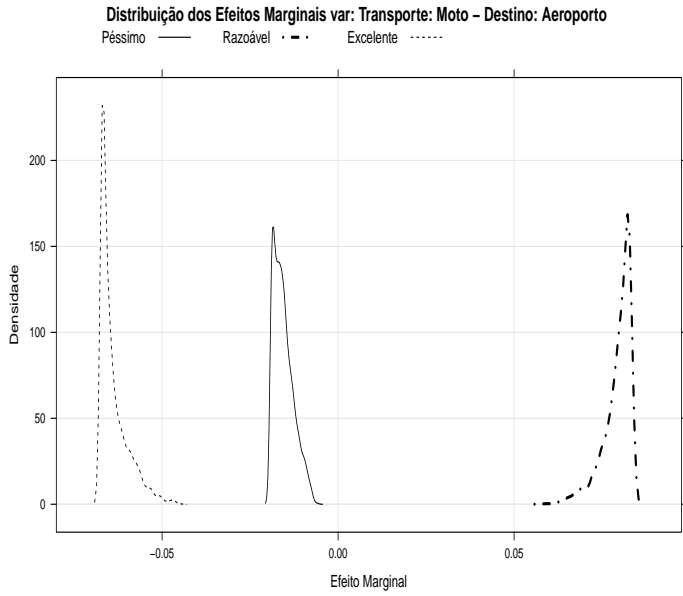


Figura 13: Destino Aeroporto: Distribuição empírica dos efeitos marginais individuais

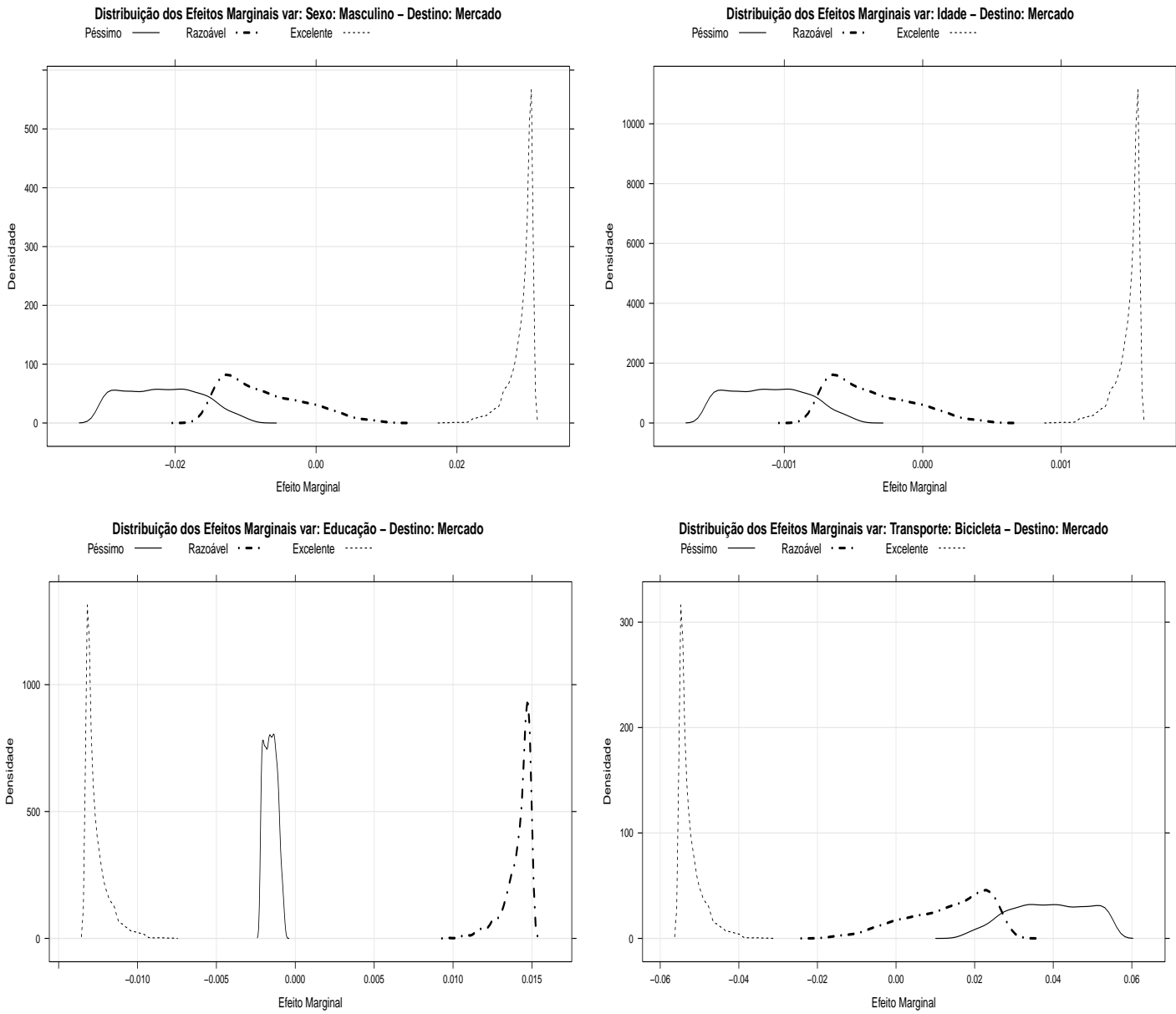


Fonte: Dados gerados pelo autor

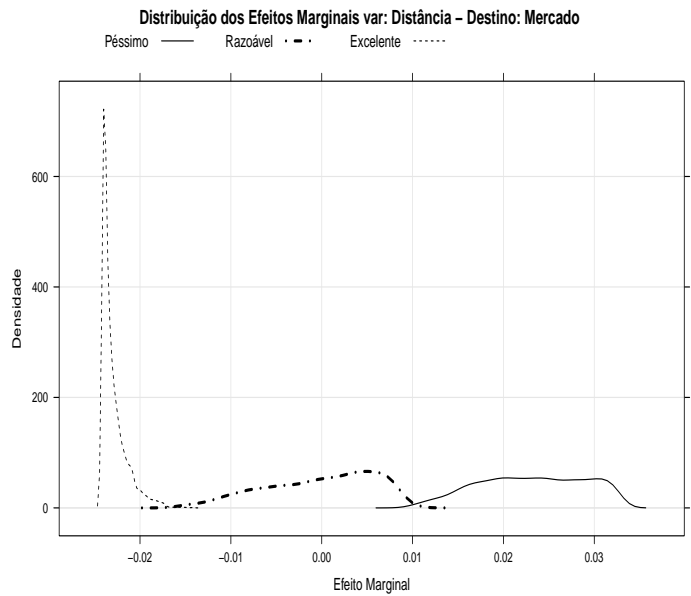
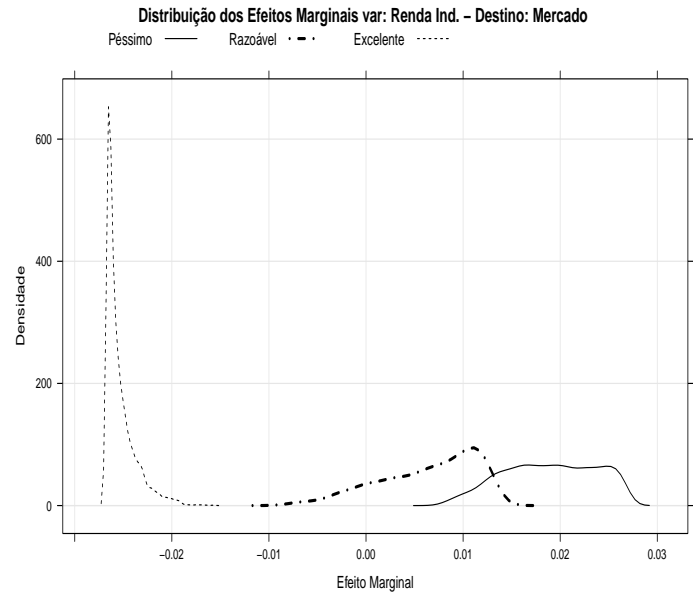
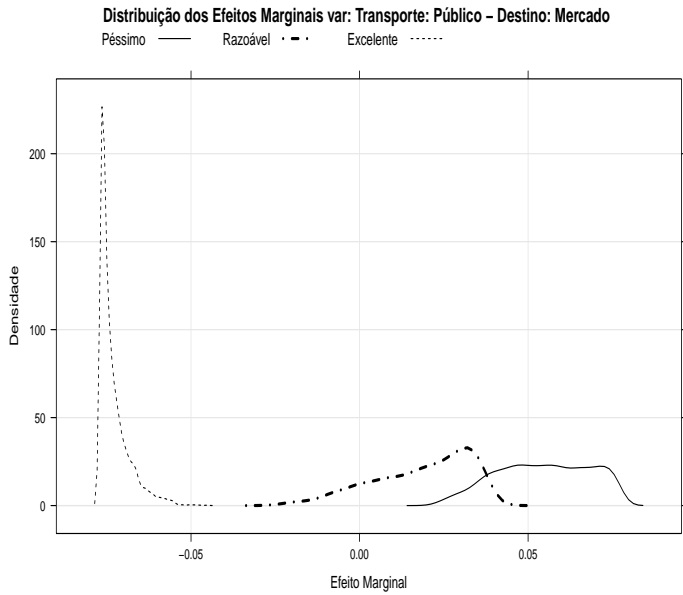
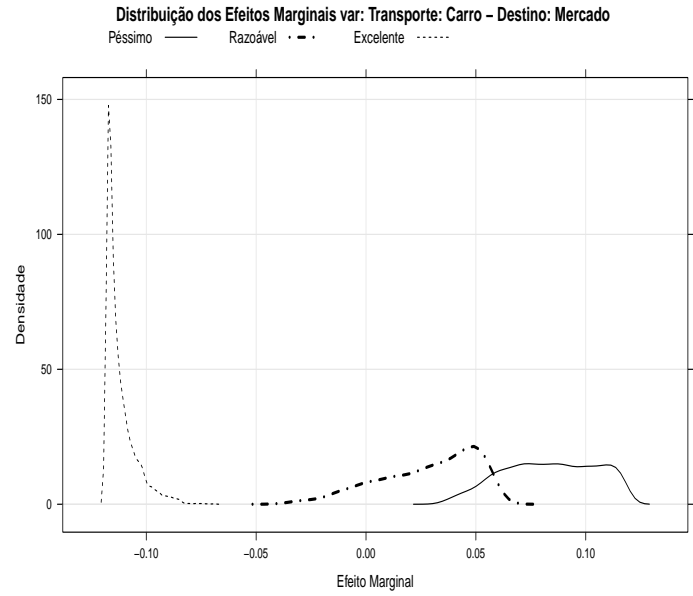
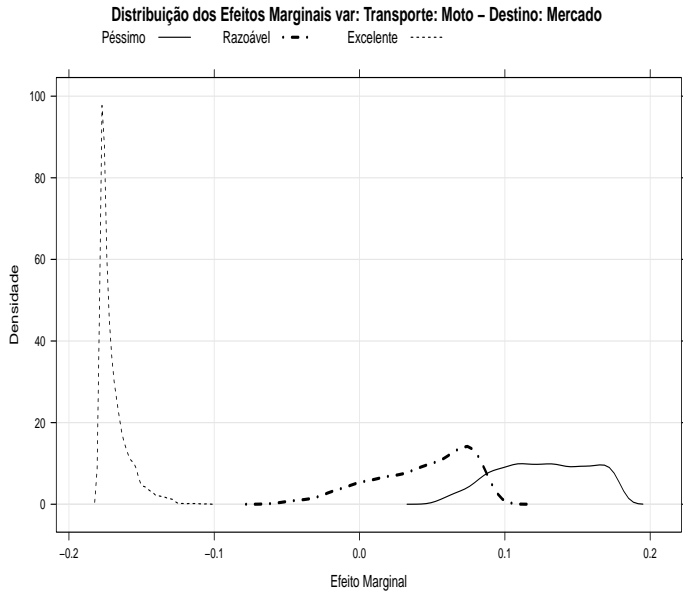


Fonte: Dados gerados pelo autor

Figura 14: Destino Mercado Central: Distribuição empírica dos efeitos marginais individuais

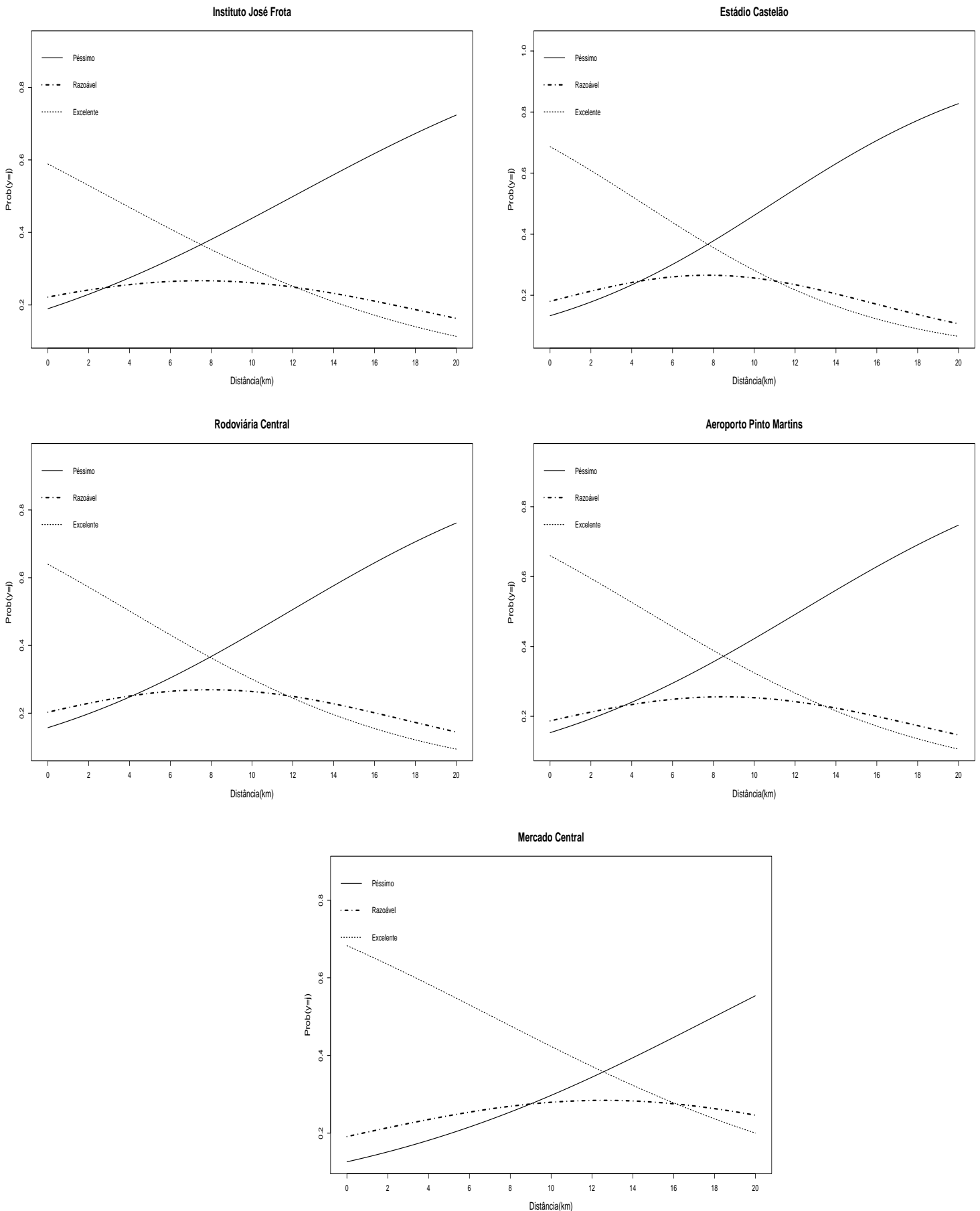


Fonte: Dados gerados pelo autor



Fonte: Dados gerados pelo autor

Figura 15: Probabilidade Estimada para o Nível de Acesso aos Destinos pela Distância, mantendo as outras variáveis em suas médias



REFERÊNCIAS

- Agresti (2002)** Alan Agresti. *Categorical Data Analysis*. Citado na pág. 38, 40
- Aitchison e Silvey (1957)** J. Aitchison e S. D. Silvey. The generalization of probit analysis to the case of multiple responses. *Biometrika*, 44:131–140. Citado na pág. 35
- Almeida (2004)** Eduardo Simões Almeida. *Curso de econometria espacial aplicada*. Citado na pág. 57
- Anderson (1984)** J. A. Anderson. Regression and ordered categorical variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 46-1:1–30. Citado na pág. 41
- Ben-Akiva e Bierlaire (1999)** Moshe Ben-Akiva e Michel Bierlaire. Discrete choice methods and their applications to short-term travel decisions. Em *Handbook of Transportation Science, Operations Research and Management Science*, páginas 5–34. Kluwer. Citado na pág. 21, 22, 23, 24
- Ben-Akiva e Lerman (1978)** Moshe Ben-Akiva e Steven R. Lerman. *Disaggregate travel and mobility choice models and measures of accessibility*. Behavioural Travel Modelling. URL <http://infoscience.epfl.ch/record/88153/files/handbook-final.pdf>. Citado na pág. 12, 21, 70
- Breheeny (1978)** M. J. Breheeny. The measurement of spatial opportunity in strategic planning. *Regional Stu*, 12:463–479. Citado na pág. 18
- Burns (1979)** L.D. Burns. Transportation, temporal and spatial components of accessibility. Citado na pág. 15
- Burton et al. (2001)** Michael Burton, Dan Rigby, Trevor Young e Sallie James. Consumer attitudes to genetically modified organisms in food in the uk. *Eur Rev Agric Econ*, 28 (4): 479–498 doi:10.1093/erae/28.4.479. Citado na pág. 25
- Carvalho (2012)** Jose Raimundo Araújo Carvalho. Montagem de uma base de dados longitudinal de vitimização do ceará: Aspectos sócio-econômicos e espaciais. Outubro 2012. Citado na pág. 13
- Christensen (2012)** Rune Haubo B Christensen. *Analysis of ordinal data with cumulative link models - estimations with the R-package ordinal*, September 2012. Citado na pág. 38, 42
- Druck et al. (2004)** S. Druck, M.S. Carvalho, G. Câmara e A.V. M. Monteiro. *Análise Espacial de Dados Geográficos*. Brasília, EMBRAPA (ISBN: 85-7383-260-6). URL www.dpi.inpe.br/gilberto/livro/analise. Citado na pág. 56, 58, 59, 60, 61, 62
- Duncan (2001)** Alan Duncan. *Cross-section and panel data econometrics. Lecture 1: Binary Choice Models*. School of Economics MSc/Ma Programme in Economics., 2001. Citado na pág. 31
- ESRI (2001)** ESRI. *ArcGis 9: Using ArcGIS Geostatistical Analyst*, 2001. Citado na pág. 63
- Ferreira (2001)** Aurélio Buarque Holanda Ferreira. *Miniaurélio Século XXI: O minidicionário da língua portuguesa*. Rio de Janeiro: Nova Fronteira, 5^a ed. rev. ampliada ed. Citado na pág. 15

- Ferreira e Martins (2011)** Dora Isabel Rodrigues Ferreira e Natalino Martins. Indicadores de acessibilidade. contributos para a síntese de conhecimento, Junho 2011. URL http://www.dpp.pt/Lists/Pesquisa%20Avanada/Attachments/3219/Indicadores_Acessibilidade.pdf. Citado na pág. 12, 16
- Geurs e van Eck (2001)** Karst T. Geurs e J.R. Ritsema van Eck. *Accessibility measures: review and applications Evaluation of accessibility impacts of land-use transport scenarios, and related social and economic impacts*. Urban Research Centre, Utrecht University. Citado na pág. 14, 16, 17, 18, 19, 20, 24, 27, 28, 29
- Geurs e van Wee (2004)** Karst T. Geurs e Bert van Wee. Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, 12: 127–140. Citado na pág. 14, 15, 16, 17, 19, 20, 24, 26, 27, 29
- Greene (1951)** William H. Greene. *Econometric Analysis*. 3^o edição ed. Citado na pág. 24, 31, 33
- Greene e Hensher (2009)** William H. Greene e David A. Hensher. Modeling ordered choices:, 2009. URL <http://pages.stern.nyu.edu/~wgreene/DiscreteChoice/Readings/OrderedChoiceSurvey.pdf>. Citado na pág. 37, 38, 40
- Guagliardo (2004)** Mark F Guagliardo. Spatial accessibility of primary care: concepts, methods and challenges. *International Journal of Health Geographics*, 3. Citado na pág. 12, 18
- Handy e Niemeier (1997)** S.L. Handy e D.A. Niemeier. Measuring accessibility: an exploration of issues and alternatives. *Environment and Planning A*, 29:1175–1194. Citado na pág. 19
- Hansen (1959)** W Hansen. How accessibility shapes land use. *Journal of the American Institute of Planners*, 2:73–76. Citado na pág. 71
- Henrique (2004)** Camila Soares Henrique. Diagnóstico espacial da mobilidade e da acessibilidade dos usuários do sistema integrado de transporte de fortaleza. Dissertação de Mestrado, Universidade Federal do Ceará. Citado na pág. 11, 12, 15, 16
- Ingram (1971)** D.R. Ingram. The concepts of accessibility: A search for an operational form. *Regional Studies*, 5:101–107. Citado na pág. 17
- Issaks e Srivastava (1989)** Edward H. Issaks e R. Mohan Srivastava. *An Introduction to Applied Geostatistics*. Citado na pág. 61
- Jakob (2002)** Alberto Augusto Eichman Jakob. A krigagem como método de análise de dados demográficos. Trabalho apresentado no XIII Encontro da Associação Brasileira de Estudos Populacionais, realizado em Ouro Preto, Minas Gerais, Brasil, 2002. URL http://www.abep.nepo.unicamp.br/docs/anais/pdf/2002/GT_SAU_ST3_Jakob_texto.pdf. Citado na pág. 61
- Jones e Westerland (2006)** Bradford S. Jones e Chad Westerland. *Order Matters (?): Alternatives to Conventional Practices for Ordinal Categorical Response Variables*. Department of Political Science - University of Arizona, 2006. Citado na pág. 39, 40
- Junior (2010)** Admir Antonio Betarelli Junior. Custo de acessibilidade entre residência e trabalho: um enfoque das características individuais, familiares e locais. Texto para discussão Nº 407, 2010. Citado na pág. 25

- Koenig (1980)** J. G. Koenig. Indicators of urban accessibility: Theory and application. *Transportation* 9, páginas 145–172. Citado na pág. 22, 23, 24, 70
- Kwan (1998)** M-P Kwan. Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis*, 30 (3):191–216. Citado na pág. 21
- Landim (2006)** Paulo M. Barbosa Landim. Sobre geoestatística e mapas. 2(1):19–33. URL <http://www.ige.unicamp.br/terraeducativa/>. Citado na pág. 61, 62, 63
- Lindemann et al. (1998)** F. Lindemann, V.B.G. Campos e A.F.M. Gonçalves. Método de avaliação de acessibilidade viária e sua relação com o uso do solo. Em *Anais do XII Congresso de Pesquisa e Ensino em Transporte - ANPET*, páginas 457–468. Citado na pág. 16
- Liu e Agresti (2005)** Ivy Liu e Alan Agresti. The analysis of ordered categorical data: An overview and survey of recent developments. *Sociedad de Estadística e Investigación Operativa*, 14:1–30. Citado na pág. 40
- Liu e Zhu (2004)** Suxia Liu e Xuan Zhu. Accessibility analyst: an integrated gis tool for accessibility analysis in urban transportation planning. *Environment and Planning B: Planning and Design*, 31:105–124. Citado na pág. 18
- Long (2012)** J. Scott Long. *Regression Models for Nominal and Ordinal Outcomes*. Indiana University, 2012. Citado na pág. 39
- Luo e Wang (2003)** Wei Luo e Fahui Wang. Measures of spatial accessibility to health care in a gis environment: synthesis and a case study in the chicago region. *Environment and Planning B: Planning and Design*, 30:865–884. Citado na pág. 12, 18
- McCullagh (1980)** P. McCullagh. Regression models for ordinal data. *Journal of the Royal Statistical Society, Series B* 42:109–142. Citado na pág. 35, 38, 39
- McKelvey e Zavoina (1975)** Richard D. McKelvey e William Zavoina. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, 4: 103–120. Citado na pág. 35
- Miller (1991)** Harvey J. Miller. Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Systems*, 5 (3): 287–301. Citado na pág. 20, 71
- Miller (1999)** Harvey J. Miller. Measuring space-time accessibility benefits within transportation networks: Basic theory and computational procedures. *Geographical Analysis*, 31 (2): 187–212. Citado na pág. 20
- Niemeier (1977)** Debbie A. Niemeier. Accessibility: an evaluation using consumer welfare. *Transportation*, 24:377–396. Citado na pág. 15, 21, 25
- R Development Core Team (2012)** R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL <http://www.R-project.org/>. ISBN 3-900051-07-0. Citado na pág. 44
- Rosero-Bixby (2004)** Luis Rosero-Bixby. Spatial access to health care in costa rica and its equity: a gis-based study. *Social Science & Medicine*, 58:1271–1284. Citado na pág. 15, 18

- Snell (1964)** E. J. Snell. A scaling procedure for ordered categorical data. *Biometrika*, 20:592–607. Citado na pág. 35
- Talen e Anselin (1998)** E Talen e L Anselin. Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds. *Environment and Planning A*, 30:585–613. Citado na pág. 12, 15, 18
- Vieira e Haddad (2012)** Renato S. Vieira e Eduardo A. Haddad. Índice de acessibilidade para são paulo, 2012. URL http://www.anpec.org.br/encontro/2012/inscricao/files_I/i9-3a9177bde138e0856209565459aa85a7.pdf. Citado na pág. 11
- Vlek e Steg (1996)** C Vlek e L Steg. Societal reasons, conditions and policy strategies for reducing the use of motor vehicles. *Towards sustainable transportations, Vancouver, Canada*, páginas 24–27. Citado na pág. 29
- Williams (2006)** Richard Williams. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *The Stata Journal*, 6:58–82. Citado na pág. 41, 42, 50
- Wooldridge (2001)** Jeffrey M. Wooldridge. *Econometric analysis of cross section and panel data*. The MIT Press. Citado na pág. 24, 32, 34
- Yee (2010)** Thomas W. Yee. The vgam package for categorical data analysis. *Journal of Statistical Software*, 32-10. URL <http://www.jstatsoft.org/v32/i10/paper>. Citado na pág. 45