



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

**ESTIMAÇÃO DO TOTAL DE VEÍCULOS EM
POLOS GERADORES DE VIAGENS
UTILIZANDO O ESTIMADOR DE
HORVITZ-THOMPSON**

JOÃO PAULO VIEIRA COSTA

10/25236

Brasília

2013

JOÃO PAULO VIEIRA COSTA

10/25236

**ESTIMAÇÃO DO TOTAL DE VEÍCULOS EM
POLOS GERADORES DE VIAGENS
UTILIZANDO O ESTIMADOR DE
HORVITZ-THOMPSON**

Relatório apresentado à disciplina Estágio Supervisionado II do curso de graduação em Estatística, Departamento de Estatística, Instituto de Exatas, Universidade de Brasília, como parte dos requisitos necessários para o grau de Bacharel em Estatística.

Orientador: Prof. Dr. Alan Ricardo da Silva

Brasília

2013

A família e amigos.

João Paulo Vieira Costa

Agradecimentos

Agradeço ao Sindicato dos Supermercados do Distrito Federal, por ter disponibilizado informações que facilitaram a realização do trabalho.

Agradeço aos meus amigos que sempre estiveram por perto nos momentos em que precisei.

Aos meus pais e à criação que me deram.

Agradeço ao professor Alan por toda a orientação e dedicação durante a realização desse trabalho.

Resumo

Os Polos Geradores de Viagens são locais que, devido ao seu empreendimento, atraem uma grande quantidade de veículos diariamente, influenciando negativamente o trânsito de seu entorno. Por isso muitos estudos na área dos PGV são realizados, a fim de determinar as medidas a serem tomadas para que a influência negativa gerada por eles seja a menor possível.

Uma estimativa do total de viagens geradas por um determinado tipo de PGV em um município ajudaria a determinar as medidas a serem tomadas naquela região além de possibilitar uma comparação entre municípios a respeito de onde seria mais adequada a instalação de um PGV. No entanto, uma estimativa desse tipo tem cada PGV como unidade amostral e uma amostra significativa para uma estimativa suficientemente precisa seria muito custosa.

Este trabalho propõe a utilização do estimador de Horvitz-Thompson para realizar essa estimativa, o que gerou intervalos de confiança menores que o estimador de média simples utilizando o mesmo tamanho de amostra. O estudo foi aplicado a supermercados e hipermercados no Distrito Federal.

Lista de Tabelas

2.1	Classificação dos Empreendimentos Varejistas segundo a ASBRA . . .	8
3.1	Exemplo Estimador de Horvitz-Thompson - População	12
3.2	Exemplo Estimador de Horvitz-Thompson	12
4.1	Totais de viagens geradas no dia pico nos empreendimentos selecionados na amostra	16
4.2	Estimativas das distâncias entre vértices e distâncias reais do Estádio Olímpico João Havelange	20
5.1	Estimativas das áreas dos supermercados/hipermercados	24
5.2	Correlações das covariáveis com o total de viagens de automóveis . .	27
5.3	Estimativas do total de viagens no Distrito Federal utilizando a covariável área do empreendimento	28
5.4	Estimativas do Total de Viagens em Taguatinga utilizando a covariável área do empreendimento	28
5.5	Estimativas do total de viagens no Distrito Federal utilizando a covariável quantidade de vagas	29

5.6	Estimativas do Total de Viagens em Taguatinga utilizando a co-	
	variável quantidade de vagas	29

Lista de Figuras

3.1	Comportamento Normal Assimétrica	14
4.1	Vértices das áreas de vendas dos Supermercados	18
4.2	Vértices da área do campo de futebol do Estádio Olímpico João Ha- velange	20
5.1	Localizações das Redes de Supermercado do Distrito Federal	23

Sumário

Resumo	iv
1 INTRODUÇÃO	1
1.1 OBJETIVOS	2
2 POLOS GERADORES DE VIAGENS	4
2.1 INTRODUÇÃO	4
2.2 GERAÇÃO DE VIAGENS	5
2.3 SUPERMERCADOS E HIPERMERCADOS COMO POLOS GE- RADORES DE VIAGENS	7
3 ESTIMADOR DO TOTAL DE HORVITZ-THOMPSON	9
3.1 INTRODUÇÃO	9
3.2 AMOSTRAGEM SEM REPOSIÇÃO EM POPULAÇÕES FINITAS .	9
3.3 PROPRIEDADES DO ESTIMADOR DE HORVITZ-THOMPSON .	10
3.4 ESTIMAÇÃO INTERVALAR PARA O ESTIMADOR DE HORVITZ-THOMPSON	13
4 MATERIAL E MÉTODOS	15
4.1 INTRODUÇÃO	15

4.2	CONTAGEM VOLUMÉTRICA DO TRÁFEGO	15
4.3	ESTIMATIVA DA ÁREA DE VENDAS ATRAVÉS DAS COORDE- NADAS GEOGRÁFICAS	17
5	ANÁLISE DOS RESULTADOS	22
5.1	INTRODUÇÃO	22
5.2	INFORMAÇÕES COLETADAS	22
5.3	ESTIMATIVAS DE HORVITZ-THOMPSON	26
5.4	ESTIMATIVA HT UTILIZANDO A COVARIÁVEL ÁREA DO EM- PREENDIMENTO	27
5.5	ESTIMATIVA HT UTILIZANDO A COVARIÁVEL QUANTIDADE DE VAGAS	28
5.6	COMPARAÇÃO ENTRE OS RESULTADOS	29
6	CONCLUSÕES	31
	REFERÊNCIAS	33

Capítulo 1

INTRODUÇÃO

Locais de diferentes naturezas que atraem uma quantidade significativa de viagens, tais como: *Shoppings centers*, hospitais, supermercados, entre outros, geram consequências em áreas do seu entorno devido ao aumento no fluxo de veículos e, assim, aumentando congestionamentos, quantidade de estacionamento irregulares, acidentes de trânsito etc.

Esses locais são chamados de Polos Geradores de Viagens (PGV). Saber as influências causadas por diferentes PGV é importante para que medidas sejam tomadas para diminuir o impacto causado por estes. Em estudos sobre influências de PGV, o conhecimento sobre o total de viagens geradas em cada polo é de grande importância, pois, em geral, os modelos de gerações de viagens são funções do número total de viagens em relação a características dos PGV, como: área bruta, quantidade de vagas, área de vendas etc.

Para a estimação do total de viagens, toma-se o próprio PGV como a unidade amostral a ser coletada na amostra, de forma que se torna difícil a coleta de uma amostra significativamente grande desses dados para gerar estimativas precisas. Silva (2006) coletou amostras em PGV, com contagens de todos os veículos que trafegas-

sem com origem ou destino nesses PGV no período de 7h00 às 0h00. Conhecendo o total populacional dessas características, o estimador do total de Horvitz Thompson é aplicável ao total de viagens realizadas com origem ou destino em PGV.

O estimador de Horvitz Thompson pode ser usado quando em uma amostragem probabilística sem reposição haja conhecimento sobre as probabilidades de seleção dos elementos da população. É dado por:

$$\hat{Y}_{HT} = \sum_{i=1}^n \frac{y_i}{\pi_i} \quad (1.1)$$

onde π_i é a probabilidade do i -ésimo elemento estar presente na amostra.

O estimador de Horvitz-Thompson do total populacional possui variância dada por:

$$V(\hat{Y}_{HT}) = \sum_i^N \frac{1 - \pi_i}{\pi_i} y_i^2 + 2 \sum_i^N \sum_{j>i}^N \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j} y_i y_j \quad (1.2)$$

sendo esta menor do que a variância do estimador da média simples.

Buscar-se-á nesse trabalho utilizar o estimador de Horvitz-Thompson para estimar o total de veículos atraídos pelos supermercados e hipermercados do Distrito Federal.

1.1 OBJETIVOS

O objetivo geral do trabalho é estimar o total de veículos atraídos pelos PGV (no caso, supermercados e hipermercados) no Distrito Federal.

Os objetivos específicos são:

- mapear onde se localizam os PGV;

- comparar os resultados com o estimador da Amostragem Aleatória Simples (AAS).

Capítulo 2

POLOS GERADORES DE VIAGENS

2.1 INTRODUÇÃO

Os Polos Geradores de Viagens são empreendimentos de grande porte que, devido às suas ofertas de bens e serviços atraem uma quantidade elevada de deslocamentos e, conseqüentemente, causam impactos negativos ao sistema viário do seu entorno. Exemplos de PGV são: *Shoppings center*, Supermercados e Universidades (Grando, 1986). Anteriormente denominado “Polos Geradores de Tráfego (PGT)”, o termo “Polos Geradores de Viagens (PGV)” passou a ser utilizado, pois o impacto causado vai além do observado simplesmente no tráfego. Entre os efeitos de um PGV estão o impacto na segurança de veículos, condutores e pedestres, e na estrutura urbana, influenciando diretamente na qualidade de vida. Assim, o termo “Polos Geradores de Viagens ” é mais adequado por dar um sentido mais amplo aos problemas causados por esses empreendimentos (Silva, 2006).

Diante dos problemas apresentados, torna-se necessária a análise do impacto causado por PGV, para que se possa conhecer a intensidade do problema e, assim, tomar as devidas providências. Em grandes cidades, tem-se adotado medidas legais

para que sejam licenciados projetos envolvendo os estudos de PGV. No Brasil, São Paulo foi a primeira cidade a fazer estudos desse tipo, através da Companhia de Engenharia de Tráfego de São Paulo - CET que criou o Manual de Polos Geradores de Viagens que define os seguintes objetivos para a análise de impacto de um PGV.

- Garantir a melhor inserção possível do empreendimento proposto na malha viária;
- Diminuir ao máximo a perturbação do tráfego de passagem em virtude do tráfego gerado pelo empreendimento;
- Viabilizar a absorção pela edificação de toda a demanda por estacionamento gerada pelo empreendimento;
- Assegurar que as operações de carga e descarga ocorram nas áreas internas da edificação;
- Reservar espaços seguros para circulação e travessia de pedestres;
- Assegurar a existência e vagas de estacionamento específicas para deficientes físicos e motocicletas.

Diferentes metodologias são utilizadas para a análise do impacto e para o tratamento das PGV.

2.2 GERAÇÃO DE VIAGENS

Uma das etapas mais importantes no estudo sobre PGV é a *Geração de Viagens*, que define a quantidade total de veículos atraídos por aquele Polo Gerador de

Viagens. O conhecimento sobre a geração de viagens é importante, pois a partir dele torna-se capaz dimensionar a infraestrutura necessária na região do entorno da PGV para comportá-lo, tal como a largura das vias, quantidade de vagas etc.

Ary (2002) define que os fatores que podem influenciar na quantidade de viagens geradas por um PGV, podem ser relacionados às características dele assim como da região de sua implantação. Entre esses fatores, podem ser citados:

- Área de Vendas.
- Tipo de itens ofertados.
- Localização.
- Porte do empreendimento.
- Existência de concorrência.
- Características socioeconômicas da região.

De modo geral, nos estudos sobre os impactos de um PGV, busca-se uma melhor forma de prever a geração de viagens que o empreendimento provocará para as medidas a serem tomadas para a sua implantação que sejam dimensionadas corretamente. Para isso, usa-se os *Modelos de Geração de Viagens* que buscam, através de informações como as citadas, estimar a geração de viagens de um PGV. Geralmente a previsão da quantidade de viagens é feita através de técnicas de regressão relacionando-a com as características do PGV e assim com base em PGV semelhantes, pode-se estimar o total de viagens que serão geradas por outro que será implantado.

2.3 SUPERMERCADOS E HIPERMERCADOS COMO POLOS GERADORES DE VIAGENS

A ASBRA (Associação Brasileira de Supermercados) classifica os supermercados com base em sua área de vendas e quantidade de itens ofertados por um empreendimento varejista. A Tabela 2.1 mostra a classificação desses empreendimentos segundo a ASBRA.

A grande quantidade de itens ofertados por empreendimentos desse porte atrai muitas viagens diariamente, influenciando no sistema viário em seu entorno. Segundo Sesso Filho (2003), os supermercados e hipermercados comercializam a maior parte dos produtos adquiridos pela população brasileira. Por esse motivo, são considerados potenciais Polos Geradores de Viagens, e necessitam de atenção na análise de seus impactos.

Há uma carência de estudos voltados para os supermercados e hipermercados como Polos Geradores de Viagens, visto que são encontradas poucas referências nacionais sobre esse assunto exclusivamente, segundo Portugal and Goldner (2003).

Tabela 2.1: Classificação dos Empreendimentos Varejistas segundo a ASBRA

Classificação	Área de vendas	Nº de itens oferecidos	Outras restrições
Loja de conveniência	-	Aprox. 1000 itens	Ausência de produtos não alimentícios e carece de qualquer tipo de serviço.
Loja de sortimento limitado	$A < 400 \text{ m}^2$	Aprox. 700 itens	Opera com até quatro caixas.
Supermercado convencional	$700 \text{ m}^2 < A < 2.500 \text{ m}^2$	Aprox. 9.000 itens	Opera com no mínimo cinco seções: mercearia, carne, frutas e verduras, frios e laticínios e produtos não alimentícios.
Supermercado compacto	$300 \text{ m}^2 < A < 700 \text{ m}^2$	Aprox. 4.000 itens	-
Superloja	$A < 4.500 \text{ m}^2$	Mais de 14.000 itens	-
Hipermercado	$A < 5.000 \text{ m}^2$	Aprox. 35.000 itens	Departamentos de: mercadoria, perecíveis, têxtil, bazar e eletrodomésticos.
Supercenter	$A < 19.000 \text{ m}^2$	Aprox. 50.000 itens	Loja de departamentos com descontos pelo auto-serviço.
Loja-depósito	$4.000 \text{ m}^2 < A < 7.000 \text{ m}^2$	Mais de 7.000 itens	Loja com aparência despojada e de baixo investimento (transparece a imagem de preços baixos).

Capítulo 3

ESTIMADOR DO TOTAL DE HORVITZ-THOMPSON

3.1 INTRODUÇÃO

Horvitz and Thompson (1952) propuseram uma forma de lidar com amostras probabilísticas sem reposição em uma população finita, no caso onde a probabilidade de seleção de cada um dos elementos da população é conhecida e utilizada no momento da seleção da amostra. Além de ser um estimador linear não viesado para o total populacional, assim como da média da variável de interesse, o estimador de Horvitz-Thompson possui menor variância que uma estimação desses pela utilização de estimadores de média simples.

3.2 AMOSTRAGEM SEM REPOSIÇÃO EM POPULAÇÕES FINITAS

Em uma população U , composta por N elementos $(u_1, u_2, u_3 \dots, u_N)$, retira-se uma amostra de tamanho n sem que haja reposição. Por ser uma amostragem sem reposição, a probabilidade de seleção deve ser recalculada a cada retirada para os elementos restantes, podendo ser calculada com base nas probabilidades de seleção

iniciais ou sem nenhuma relação (Horvitz and Thompson, 1952).

A probabilidade de seleção do i -ésimo elemento da população na m -ésima retirada é descrita como p_{i_m} $i = (1, 2, 3, \dots, N)$, e $m = (1, 2, 3, \dots, n)$ sendo que:

$$\sum_{i=1}^{N-m+1} p_{i_m} = 1 \quad e \quad p_{i_m} \geq 0 \quad (3.1)$$

onde $N - m + 1$ a quantidade de elementos restantes na população de tamanho N após m retiradas.

Quando se conhece a probabilidade de seleção de cada retirada, a probabilidade de seleção *a priori* de cada u_i em uma amostra de tamanho n pode ser calculada.

A notação para essa probabilidade será denotada por π_i , sendo que:

$$\sum_{i=1}^N \pi_i = n \quad (3.2)$$

Por essa não ser uma soma de eventos mutuamente exclusivos, a não ser no caso onde a amostra tem tamanho 1.

3.3 PROPRIEDADES DO ESTIMADOR DE HORVITZ-THOMPSON

Horvitz and Thompson (1952) introduziu uma generalização do estimador proposto por Hansen and Hurwitz (1943), que foi direcionado a amostras aleatórias de uma população finita, com reposição e, por isso, a probabilidade de seleção de cada indivíduo na população é sempre a mesma em qualquer amostra.

O estimador de Horvitz-Thompson pode ser usado em casos com e sem reposição em casos em que se conhece a probabilidade de inclusão π_i dos elementos da po-

pulação. É um estimador não viesado do total populacional e é dado por:

$$\hat{Y}_{HT} = \sum_{i=1}^n \frac{y_i}{\pi_i} \quad (3.3)$$

onde y_i são os valores observados da amostra coletada. A estimativa da média pode ser calculada dividindo-se \hat{Y}_{HT} por N .

A variância associada a esse estimador é dada por:

$$V(\hat{Y}_{HT}) = \sum_i^N \frac{1 - \pi_i}{\pi_i} y_i^2 + \sum_i^N \sum_{j \neq i}^N \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j} y_i y_j \quad (3.4)$$

onde π_i é a probabilidade de que o elemento i esteja na amostra, e π_{ij} a probabilidade de que os elementos i e j estejam na amostra. No estimador de Horvitz-Thompson as probabilidades de seleção associadas a cada elemento podem ser calculadas com base em variáveis associadas à variável de interesse.

Além disso, existem dois estimadores mais conhecidos para a variância do estimador de Horvitz-Thompson, que serão representados por v_1 e v_2 nas equações 3.5 e 3.6 (Cochran, 1977):

$$v_1(\hat{Y}_{HT}) = \sum_i^n \frac{(1 - \pi_i)}{\pi_i^2} + 2 \sum_i^n \sum_{j > i}^n \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j \pi_{ij}} y_i y_j \quad (3.5)$$

$$v_2(\hat{Y}_{HT}) = \sum_i^n \sum_{i > j}^n \frac{(\pi_i \pi_j - \pi_{ij})}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2 \quad (3.6)$$

Dependendo das probabilidades de seleção associadas às amostras selecionadas, os estimadores v_1 e v_2 podem resultar em estimativas negativas da variância.

A Tabela 3.2 é um exemplo que mostra todas as possíveis amostras de $n=2$ que podem ser selecionadas para estimar o número de vendas dos quatro supermercados de uma pequena cidade apresentados na Tabela 3.1 (Lohr, 1999) utilizando o

estimador de Horvitz-Thompson. As estimativas foram feitas utilizando a área dos supermercados como covariável, sendo ela apresenta uma correlação positiva com o número de vendas.

Tabela 3.1: Exemplo Estimador de Horvitz-Thompson - População

Supermercado	Vendas	Área
A	11	100
B	20	200
C	24	300
D	245	1000

Tabela 3.2: Exemplo Estimador de Horvitz-Thompson

Elemento 1	Elemento 2	HT	AAS	prob	varht2	varsrs
11	20	111,9	62	0.00833	47,06	162
20	11	111,9	62	0.00893	47,06	162
11	24	102,4	70	0.01250	502,81	338
24	11	102,4	70	0.01442	502,81	338
11	245	330,1	512	0.04167	7939,75	109512
245	11	330,1	512	0.10417	7939,75	109512
20	24	98,5	88	0.02679	232,72	32
24	20	98,5	88	0.02885	232,72	32
20	245	326,1	530	0.08929	5744,06	101250
245	20	326,1	530	0.20833	5744,06	101250
24	245	316,7	538	0.14423	3259,78	97682
245	24	316,7	538	0.31250	3259,78	97682

O objetivo é mostrar que é possível encontrar amostras onde a variância do estimador de Horvitz-Thompson é maior que a variância do estimador de média simples mesmo que populacionalmente a variância do estimador de Horvitz-Thompson seja menor. Isso mostra que mesmo que, em alguns casos, o erro padrão do estimador de Horvitz-Thompson seja maior na amostra coletada não significa dizer que haja um prejuízo no uso desse estimador.

3.4 ESTIMAÇÃO INTERVALAR PARA O ESTIMADOR DE HORVITZ-THOMPSON

Dominoni (2012) propôs um intervalo de confiança para o estimador de Horvitz-Thompson, levando em consideração a assimetria gerada pela adequação às diferentes probabilidades de seleção da distribuição de probabilidade do estimador de Horvitz-Thompson. Esse intervalo de confiança foi criado com base na distribuição Normal Assimétrica (*Skew Normal*) (Azzalini, 1985) e também a distribuição t assimétrica (*Skew t*) (Azzalini and Capitanio, 2003).

Os limites do intervalo de confiança gerado a partir do estimador de Horvitz-Thompson no caso da distribuição Normal Assimétrica são:

$$\text{Limite inferior: } \hat{Y}_{HT} + z_{\alpha/2} \sqrt{\text{var}(Y_{HT})}$$

$$\text{Limite superior: } \hat{Y}_{HT} + z_{1-\alpha/2} \sqrt{\text{var}(Y_{HT})}$$

E no caso da distribuição t Assimétrica:

$$\text{Limite inferior: } \hat{Y}_{HT} + t_{\alpha/2} \sqrt{\text{var}(Y_{HT})}$$

$$\text{Limite superior: } \hat{Y}_{HT} + t_{1-\alpha/2} \sqrt{\text{var}(Y_{HT})}$$

onde \hat{Y}_{HT} é o total estimado pelo estimador de Horvitz-Thompson, $\text{var}(Y_{HT})$ é a variância populacional do estimador de Horvitz-Thompson, z e t são os quantis das distribuições Normal Assimétrica e t Assimétrica respectivamente a um nível de significância α .

O parâmetro de assimetria das distribuições citadas é calculado através da distribuição dos $\pi_i s$, sendo que a assimetria calculada para eles será a mesma a ser utilizada nas distribuições Normal e T assimétricas. No caso onde o parâmetro de

assimetria é igual a 0, os quantis da distribuições Normal e T assimétricas são os mesmos da distribuição Normal e T-Student. As distribuições Normal e T-Student são casos particulares da distribuição Normal Assimétrica e T Assimétrica respectivamente.

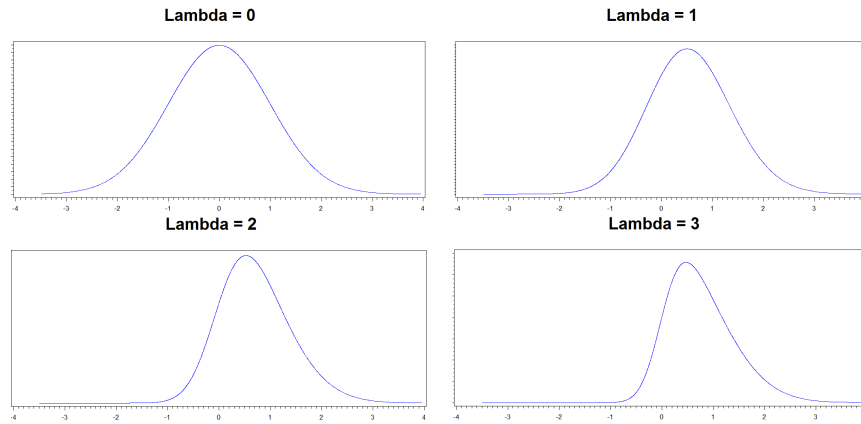


Figura 3.1: Comportamento Normal Assimétrica

A Figura 3.1 ilustra o comportamento da distribuição Normal Assimétrica diante de diferentes parâmetros de assimetria. A distribuição Normal assimétrica apresenta um comportamento diferente das distribuições assimétricas mais conhecidas, de modo que a variação do parâmetro de assimetria causa uma mudança muito menor em uma das caudas. Dessa forma os quantis utilizados nos intervalos de confiança tendem a se aproximar quando o parâmetro de assimetria se distancia de zero.

Capítulo 4

MATERIAL E MÉTODOS

4.1 INTRODUÇÃO

Como visto, o estimador de Horvitz-Thompson necessita de covariáveis para ser calculado. Quanto maior for a correlação entre a variável de interesse e a covariável utilizada, mais precisas serão as estimativas.

No caso, a melhor forma de se obter essas informações seria consegui-las diretamente com os empreendimentos. Por isso, foram enviadas cartas aos empreendimentos solicitando as informações que possivelmente poderiam ser boas covariáveis, todas não sigilosas. É possível, também, coletar as informações de área e número de vagas de estacionamento por meio de imagens de satélite, no caso do conhecimento de vagas subterrâneas deve ser feita uma contagem dessas vagas *in loco*.

4.2 CONTAGEM VOLUMÉTRICA DO TRÁFEGO

A amostra utilizada nesse trabalho é a mesma utilizada por (Silva, 2006), que em sua pesquisa coletou, em sete empreendimentos (sendo três em Goiânia-GO, dois em Anápolis-GO e dois em Taguatinga-DF), as seguintes informações:

- Informações de questionários aplicados aos clientes dos supermercados;
- Informações gerais sobre os supermercados através de suas gerências.
- Informações acerca de sua região de entorno;
- Contagem volumétrica do tráfego entrante ao estabelecimento no dia pico da semana.

No presente trabalho, será utilizada a informação obtida na etapa de contagem volumétrica do tráfego entrante ao estabelecimento no dia de pico, apenas dos dois casos em que o empreendimento localiza-se no Distrito Federal. A contagem foi realizada durante o horário de funcionamento dos estabelecimentos, por pesquisadores posicionados nos acessos aos estacionamentos e também registrando veículos em estacionamento irregulares, quando necessário.

A informação utilizada como total de viagens é a quantidade de veículos atraídos no dia de pico para os estabelecimentos localizados no Distrito Federal. Por motivos de sigilo da informação coletada nos supermercados pesquisados, eles serão referenciados como S1 e S2 quando necessário mencioná-los. Esses valores são os apresentados na Tabela 4.1:

Tabela 4.1: Totais de viagens geradas no dia pico nos empreendimentos selecionados na amostra

Supermercado/Veículos	Automóveis	Motos	Automóveis e motos	Caminhões
S1	504	15	519	14
S2	2856	198	3054	23

O número de veículos foi registrado a cada 15 minutos para cada estabelecimento, durante uma semana entre os meses de setembro e outubro de 2005, para que fatores

como férias, festividades e outros não influenciassem na quantidade de viagens geradas. No presente trabalho, serão utilizados somente os automóveis como viagens geradas, sendo que eles representam mais de 90% do total de viagens geradas pelo PGV nessa amostra. Para informações mais detalhadas sobre a forma de coleta e seleção da amostra, Silva (2006) pode ser consultado.

4.3 ESTIMATIVA DA ÁREA DE VENDAS ATRAVÉS DAS COORDENADAS GEOGRÁFICAS

Contendo as dimensões dos empreendimentos é possível calcular a sua área aproximada pela Equação 4.1 da área do retângulo:

$$A = b * h \quad (4.1)$$

sendo A a área calculada, b e h a base e a altura dos retângulos formados pelas superfícies dos empreendimentos.

Para que esta área seja calculada, é necessário obter os comprimentos dos lados das bases dos empreendimentos. Com as coordenadas geográficas dos vértices dos empreendimentos, encontradas pelo *Google Earth* através dos seus endereços, é possível aplicar a lei dos cossenos para calcular a distância entre dois vértices do empreendimento.

A Equação 4.2 descreve como calcular distância entre os pontos da superfície terrestre calculada pela lei dos cossenos no sistema de coordenadas latitude/longitude.

$$d = \arccos(\sin(lat1) \cdot \sin(lat2) + \cos(lat1) \cdot \cos(lat2) \cdot \cos(long2 - long1)) \cdot r \quad (4.2)$$

onde r é o raio da Terra (aproximadamente 6.371 km). Utilizando essas distâncias calculadas na Equação 4.1, obtém-se as áreas aproximadas empreendimentos localizados.

A Figura 4.1 mostra como foram referenciados os vértices dos empreendimentos que foram utilizados para calcular suas áreas construídas.

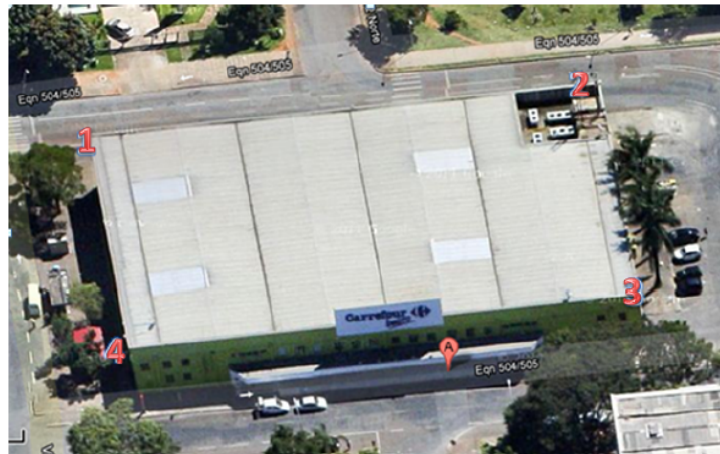


Figura 4.1: Vértices das áreas de vendas dos Supermercados

A programação abaixo, utilizada no Software SAS (2008) através do *PROC IML*, aplica a lei dos cossenos mencionada e calcula a área dos empreendimentos. Segue o exemplo do cálculo das áreas dos supermercados Supermaia do Guará I e II respectivamente:

```
data spmaia_guara; /*QE 17*/
input id y x;
datalines;
1 15.825990 47.979141
1 15.825478 47.979328
1 15.825400 47.979059
1 15.825895 47.978868
;
data spmaia_guara2; /*QE 13*/
input id y x;
datalines;
1 15.839456 47.967149
```

```

1 15.839402 47.966734
1 15.839653 47.966697
1 15.839723 47.967113
;

%macro area(base);
proc iml;
use &base;
read all var{x y} into COORD;
n=nrow(coord);
d=j(1,3,0);
nome={"idi" "idj" "d"};
create dist from d[colname=nome];
do i=1 to n;
do j=i+1 to n;
dif=abs(COORD[i,1]-COORD[j,1]);
raio=arccos(-1)/180;
if dif=0 then arco=0;
else
/* Law of Cosines */
arco=arccos(sin(COORD[i,2]*raio)*sin(COORD[j,2]*raio)+
cos(COORD[i,2]*raio)*cos(COORD[j,2]*raio)*cos(dif*raio));
d[1]=i;
d[2]=j;
d[3]=arco*6371 /*Earth's Radius = 6371 (approximately)*/;
append from d;
end;
else do;
d[1]=i;
d[2]=j;
d[3]=sqrt((COORD[i,1]-COORD[j,1])**2+(COORD[i,2]-COORD[j,2])**2);
append from d;
end;
end;
end;
close dist;
use dist;
read all into dist;
area=(dist[1,3]*1000)*(dist[3,3]*1000);
print "Supermercado &base",area;
quit;
%mend;
%area(spmaia_guara);
%area(spmaia_guara2);

```

A fim de verificar a precisão da estimativa feita para essas áreas, o procedimento foi aplicado à área já conhecida do campo de futebol do Estádio Olímpico João

Havelange (Engenhão). A Tabela 4.2 contém as estimativas das distâncias entre os vértices do campo, os valores reais e o erro na estimativa. Os vértices citados na Tabela 4.2 são os mostrados na Figura 4.2.



Figura 4.2: Vértices da área do campo de futebol do Estádio Olímpico João Havelange

Tabela 4.2: Estimativas das distâncias entre vértices e distâncias reais do Estádio Olímpico João Havelange

Vértices	Distância Estimada	Distância Real	Erro
1 x 2	105,372	105	0,372
3 x 4	105,788	105	0,788
1 x 4	67,504	68	-0,496
1 x 2	67,927	68	-0,073

Nesse caso, o erro da área estimada para a área real seria aproximadamente $27m^2$ e é esperado que o erro na estimativa das áreas dos supermercados não seja muito discrepante do que foi observado. O erro observado na estimativa da área é relativamente pequeno, por isso a técnica pôde ser aplicada às áreas dos supermercados/hipermercados sem que haja grandes prejuízos.

Com essas informações coletadas, torna-se possível calcular a estimativas para

o total de viagens geradas por Supermercado e Hipermercados no Distrito Federal através do estimador de Horvitz-Thompson e compará-lo com a estimativa feita por média simples, além de comparar as estimativas feitas com diferentes covariáveis.

Capítulo 5

ANÁLISE DOS RESULTADOS

5.1 INTRODUÇÃO

A obtenção das informações citadas como possíveis covariáveis para a estimativa através da disponibilização por todos os supermercados do Distrito Federal se mostrou muito difícil. As cartas enviadas às matrizes dos supermercados assim como visitas aos empreendimentos não obtiveram sucesso, mesmo com todas as informações sendo não sigilosas.

O Sindicato dos Supermercados do Distrito Federal forneceu uma lista com as principais redes de supermercados do Distrito Federal e a partir dessas informações foram listados todos os supermercados e hipermercados pertencentes a essa rede.

Como os dois supermercados selecionados na amostra localizam-se em Taguatinga-DF, além da estimativa do total de viagens geradas por supermercados no DF, a estimativa foi feita para os supermercados localizados em Taguatinga-DF.

5.2 INFORMAÇÕES COLETADAS

A partir da lista fornecida pelo Sindicato dos Supermercados do Distrito Federal, buscou-se os endereços de todos os supermercados nos sites de cada uma dessas redes,

dessa forma foram mapeados todos esses supermercados a fim de posteriormente aplicar a metodologia citada na seção anterior para o cálculo da área e realizar a contagem das vagas de estacionamento.

A Figura 5.1 mostra a localização de todos os supermercados que foram mapeados no Distrito Federal.

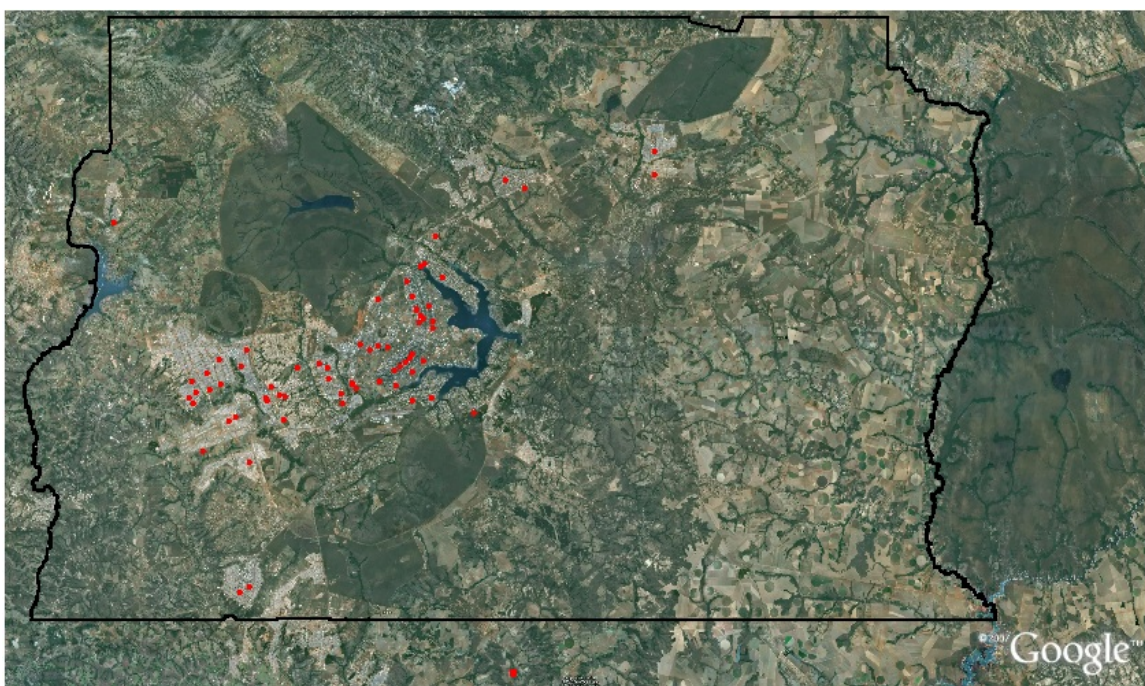


Figura 5.1: Localizações das Redes de Supermercado do Distrito Federal

A partir do mapeamento apresentado na Figura 5.1, após realizar a contagem das vagas de estacionamento e realizar o cálculo para obter a área dos empreendimentos. A Tabela 5.1 mostra a lista de todos os supermercados/hipermercados que foram utilizados nesse trabalho.

Tabela 5.1: Estimativas das áreas dos supermercados/hipermercados

Supermercado	Área	Vagas
Supercei Ceilândia EQNN 07/09	363,33	10
Supercei Gama	528,41	65
Supercei Ceilândia EQNP 10/14	599,38	20
Supercei Ceilândia P.Sul EQNP 10/14	431,58	20
Supercei Ceilândia P.Sul EQNP 26/30	565,69	20
Supercei Ceilândia P.Sul EQNP 30/34	118,99	10
Supercei Ceilândia Sul EQNM 07/09	3419,18	177
Supercei Ceilândia Sul EQNN 28	2807,49	210
Supercei Taquari	1131,23	32
Supercei Recanto das Emas	659,63	35
Supercei Riacho Fundo	1124,97	36
Supercei Samambaia Sul	1425,57	21
Supercei EPTG	1686,59	36
Supercei Taguatinga Norte EQNL 09/11	667,03	15
Supercei Taguatinga Norte QNG 02	1010,53	22
Supercei Taguatinga Norte EQNL 17/19	1808,59	20
Supercei Brazlândia	1127,84	20
Supermaia 508 Sul	713,14	57
Supermaia Cruzeiro Novo	1078,46	50
Supermaia QE 13 Guará II	1341,15	34
Supermaia QE 17 Guará II	1396,01	40
Supermaia Núcleo Bandeirante	1450,00	80
Supermaia Sobradinho	1768,28	79
Supermaia Sudoeste	1899,08	50
Supermaia Taguatinga	1952,77	14
Supermaia Águas Claras	15531,66	58
Supermaia Planaltina	1562,41	48
Supermaia 306 Norte	638,42	10
Supermaia Gama	3647,47	60
Supermaia Lago Norte	500,46	20
Supermaia Arapoanga	1808,59	55
Tokio Águas Claras	532,01	14
Tokio Av. Samdu	727,31	10
Tokio Guará II	1313,15	39
Tokio Taguatinga Sul	266,86	6
Tokio Taguatinga Norte	483,09	10
Tokio Taguatinga Cent	179,93	12
Tatico Ceilândia	4000,00	270
Tatico Samambaia Sul	2180,00	58

Tabela 5.1 : (cont.) Estimativas das áreas dos supermercados/hipermercados

Supermercado	Área	Vendas
Tatico Águas Lindas	5000,00	180
Tatico Santa Maria	6000,00	280
Tatico Av. Recanto	4000,00	170
Da Terra Guará	408,24	14
Da Terra Lago Sul	400,77	35
Caique Smambaia	1861,39	17
Caique Ceilândia	598,26	25
Big Box 402/403 Asa Norte	2159,38	57
Big Box Sudoeste	2290,95	52
Big Box 106 Norte	909,12	10
Big Box 412/413 Sul	1880,16	57
Big Box 408/409 Norte	1962,84	26
Big Box Lago Norte SHIN EPPN	1102,56	66
Big Box Vicente Pires	2157,34	98
Big Box Águas Claras	1198,30	50
Big Box Lago Sul	680,32	30
Big Box 503 Sul	951,38	10
Big Box Lago Norte SHIN CA	1053,08	132
Carrefour Norte	2451,89	1500
Carrefour Taguatinga	2538,51	640
Carrefour Ceilândia Sul	2717,99	1230
Carrefour Bairro 512/513 Norte	2826,53	57
Carrefour Bairro 504 Norte	2878,58	25
Carrefour Bairro 402 Sul	2952,00	25
Carrefour Bairro 310 Sul	1649,36	57
Carrefour Bairro ST SHI/SUL QI 13	3643,63	75
Comper 307/307 sul	3087,34	25
Comper Sobradinho	3247,23	140
Comper Taguatinga	3702,94	208
Comper Ceilândia	12486,70	188
Comper Gama	1304,78	48
Comper Samambaia	1865,62	64
Comper Rua Águas Claras	1458,58	18
Pão de Açúcar 308/309 Sul	3725,38	57
Pão de Açúcar Guará	10506,50	90
Pão de Açúcar 404/405 Norte	1158,60	30
Pão de Açúcar Lago Norte	14091,06	186
Pão de Açúcar Águas Claras	15338,32	81

Tabela 5.1 : (cont.) Estimativas das áreas dos supermercados/hipermercados

Supermercado	Área	Vagas
Pão de Açúcar 304/305 Sul	1283,26	57
Pão de Açúcar 406/407 Sul	2174,25	57
Pão de Açúcar Lago Sul	3209,85	64
Pão de Açúcar Sudoeste	1933,07	43
Pão de Açúcar 516 Sul	2957,92	96
Pão de Açúcar 508/509 Norte	3138,31	57
Extra	18121,73	232
Extra SMAS	18989,60	477
Extra SIA	19136,17	816
Extra Taguatinga	29945,70	1130

No caso dos supermercados Tatico, as informações de quantidade de vagas e área do empreendimento já são disponibilizadas em seu site na internet, em todos os outros supermercados a área apresentada é estimada conforme a metodologia da seção 4.3 As informações apresentadas na Tabela 5.1 foram utilizadas para calcular as estimativas apresentadas nas seções seguintes.

5.3 ESTIMATIVAS DE HORVITZ-THOMPSON

As estimativas do total de viagens foram feitas utilizando como variável o total de viagens de automóveis geradas pelos PGV apresentadas na Tabela 4.1. Essa variável foi escolhida por representar mais de 90% de todas as viagens geradas e estar mais relacionada com as covariáveis utilizadas.

Como dito, quanto maior a correlação com a covariável mais precisa será a estimativa gerada pelo estimador de Horvitz-Thompson. A Tabela 5.2 mostra o coeficiente correlação de Pearson das informações que poderiam ser utilizadas como covariáveis com o total de viagens de automóveis. Essa correlação foi calculada com base na informação obtida dos sete supermercados em Silva (2006).

Tabela 5.2: Correlações das covariáveis com o total de viagens de automóveis

Covariável	ρ
Itens ofertados	0,986
Área do Empreendimento	0,973
Quantidade de vagas	0,953

Entre essas covariáveis a que apresentou maior correlação foi o número de itens ofertados, porém a única forma de se obter essa informação seria pelo fornecimento dela pelos próprios empreendimentos. As outras covariáveis também apresentaram correlações altas e podem ser consideradas boas covariáveis. Dessa forma, as estimativas do total de viagens utilizando o estimador de Horvitz-Thompson serão feitas com as covariáveis “Área do Empreendimento” e “Quantidade de Vagas”.

5.4 ESTIMATIVA HT UTILIZANDO A CO-VARIÁVEL ÁREA DO EMPREENDIMENTO

As estimativas para o total de viagens gerada por supermercados e hipermercados no Distrito Federal e na cidade de Taguatinga utilizando suas áreas calculadas como covariáveis apresentou os resultados apresentados nas Tabelas 5.3 e 5.4, respectivamente.

Tabela 5.3: Estimativas do total de viagens no Distrito Federal utilizando a co-variável área do empreendimento

	Horvitz-Thompson	AAS
Estimativa	148.934	146.160
Erro Padrão	75.958,67	101.128,63
$IC_{95\%}$ (Normal)]126.745 ; 319.248[] -52.052 ; 344.372[
$IC_{95\%}$ (T-Student)]101.840 ; 2.032.590[] -1.138.801 ; 1.431.121[
Quantis (Normal)	(-0,29;2,24)	(-1,96;1,96)
Quantis (T-Student)	(-0,62;24,80)	(-12,71;12,71)
Assimetria	3,00	0

Tabela 5.4: Estimativas do Total de Viagens em Taguatinga utilizando a covariável área do empreendimento

	Horvitz-Thompson	AAS
Estimativa	36.510	28.560
Erro Padrão	22.518,52	18.779,21
$IC_{95\%}$ (Normal)]29.914 ; 86.983[] -8.247 ; 65.367[
$IC_{95\%}$ (T-Student)]10.163 ; 594.916[] -210.052 ; 267.172[
Quantis (Normal)	(-0,47;2,24)	(-1,96;1,96)
Quantis (T-Student)	(-1,17;24,24)	(-12,71;12,71)
Assimetria	2,13	0

Utilizando a área dos supermercados como covariável para estimar o total de viagens geradas no Distrito Federal, o erro padrão do estimador de Horvitz-Thompson é menor que o estimador de média simples. No caso da estimativa para Taguatinga o menor erro padrão é o da estimativa por média simples. Sendo o $Def f = 54,42\%$ para o Distrito Federal e $Def f = 143,79\%$ para Taguatinga, calculados com v_2 por não apresentar estimativas negativas da variância.

5.5 ESTIMATIVA HT UTILIZANDO A CO-VARIÁVEL QUANTIDADE DE VAGAS

Ao estimar o total de viagens gerados por supermercados e hipermercados no Dis-

trito Federal e na cidade de Taguatinga utilizando o estimador de Horvitz-Thompson com a quantidade de vagas como covariável, os resultados obtidos foram os das Tabelas 5.5 e 5.6.

Tabela 5.5: Estimativas do total de viagens no Distrito Federal utilizando a covariável quantidade de vagas

	Horvitz-Thompson	AAS
Estimativa	263.886	146.160
Erro Padrão	122.372,38	101.128,63
$IC_{95\%}$ (Normal)]234.516 ; 538.000[] -52.052 ; 344.372[
$IC_{95\%}$ (T-Student)]216.161 ; 3.298.433[] -1.138.801 ; 1.431.121[
Quantis (Normal)	(-0,21;2,24)	(-1,96;1,96)
Quantis (T-Student)	(-0,39;25,03)	(-12,71;12,71)
Assimetria	3,78	0

Tabela 5.6: Estimativas do Total de Viagens em Taguatinga utilizando a covariável quantidade de vagas

	Horvitz-Thompson	AAS
Estimativa	48.663	28.560
Erro Padrão	28.970,26	18.779,21
$IC_{95\%}$ (Normal)]40.177 ; 113.597[] -8.247,25 ; 65.367,25[
$IC_{95\%}$ (T-Student)]23.459 ; 767.057[] -210.052 ; 267.172[
Quantis (Normal)	(-0,38;2,24)	(-1,96;1,96)
Quantis (T-Student)	(-0,87;24,53)	(-12,71;12,71)
Assimetria	2,50	0

A estimativa feita pelo estimador de Horvitz-Thompson com a covariável quantidade de vagas possui uma variância maior que a do estimador de média simples. Sendo o $Def f = 146,52\%$ para o Distrito Federal e $Def f = 237,99\%$ para Taguatinga calculados com v_2 por não apresentar estimativas negativas da variância.

5.6 COMPARAÇÃO ENTRE OS RESULTADOS

Como mostrado na Seção 3.3, apesar de populacionalmente o estimador de

Horvitz-Thompson possuir uma variância menor que as estimativas de médias simples, é possível encontrar amostras em que o erro padrão observado seja maior no estimador de Horvitz-Thompson, como ocorreu em várias das estimativas realizadas.

Nota-se que a amplitude dos intervalos de confiança do estimador de Horvitz-Thompson é menor que o de média simples com a mesma confiança mesmo em casos onde o erro padrão estimado é maior. Isso ocorreu por causa da distribuição assimétrica desse estimador, que tornou os quantis da probabilidade associada a ele mais próximos que os quantis da distribuição normal do estimador de média simples.

Capítulo 6

CONCLUSÕES

Em alguns tipos de estudos, a coleta de uma amostra suficientemente grande para gerar estimativas precisas torna-se inviável por diversos motivos. O Estimador de Horvitz-Thompson, por apresentar uma menor variância, mostra-se uma alternativa interessante em estudos em que haja necessidade de se trabalhar com pequenas amostras, como é o caso do problema de transportes apresentado nesse trabalho. No entanto, sua maior dificuldade é a necessidade da utilização de covariáveis que estejam definidas para toda a população de interesse.

Uma covariável que se mostrou interessante para esta estimativa foi a quantidade de itens ofertados pelos empreendimentos, porém, pela dificuldade de se coletar essa informação populacionalmente, as únicas covariáveis utilizadas foram a área do empreendimento e a quantidade de vagas no estacionamento. A utilização do estimador de Horvitz-Thompson para estimar o total de viagens geradas por supermercados e hipermercados no Distrito Federal, quando auxiliado por essas covariáveis, gerou intervalos de confiança menores que os que são gerados pelos estimadores convencionais utilizando uma amostra com o mesmo tamanho e a mesma confiança.

O resultado é importante no caso de supermercados, de modo que um dono

de um empreendimento poderia comparar em qual localidade seria mais lucrativo instalar um novo empreendimento, visto que da forma como os dados foram coletados contabilizou-se apenas os veículos que realmente tinham destino naquele PGV. Sendo assim, a quantidade de clientes está ligada à quantidade de viagens geradas. A técnica aplicada a supermercados e hipermercados pode ser aplicada a qualquer outro tipo de PGV em qualquer região, facilitando as tomadas de decisões quanto a medidas a serem tomadas com respeito àquele tipo de PGV naquela localidade.

Referências Bibliográficas

- Ary, M. B. (2002). Análise da demanda de viagens atraídas por shopping centers em fortaleza. Master's thesis, Programa de mestrado em Engenharia de Transportes-UFC.
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*. pages 171 - 178.
- Azzalini, A. & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. *J. Roy. Statist. Soc.* pages 367-389, institution = Journal of the American Statistical Association,.
- Cochran, W. G. (1977). *Sampling Techniques*, (3rd ed.). John Wiley & Sons.
- Dominoni, T. D. B. (2012). Análise das propriedades do estimador de horivtz-thompson.
- Grando, L. (1986). A interferência dos pólos geradores de viagens no sistema viário: análise e contribuição metodológica para shopping centers. Master's thesis, COPEE-UFRJ.
- Hansen, M. H. & Hurwitz, W. N. (1943). On the theory of sampling from finite population. pp. 333-362.
- Horvitz, D. G. & Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. pp. 663-685.
- Lohr, S. L. (1999). *Sampling: Design and Analysis*. Duxbury Press.
- Portugal, L. S. & Goldner, L. G. (2003). *Estudo de Pólos Geradores de Viagens e de seus impactos nos sistemas viários e de transportes*. Editora Edgard Blucher Ltda.
- SAS (2008). *SAS software, Versão 9.2 do SAS System para Windows. Licenciado para o Departamento de Estatística - Universidade de Brasília. Cary - Carolina do Norte*.

Sesso Filho, U. A. (2003). O setor supermercadista no brasil nos anos 1990.

Silva, L. R. (2006). Metodologia de delimitação da área de influência dos pólos geradores de viagens para estudos de geração de viagens - um estudo de caso nos supermercados e hipermercados. Master's thesis, ENC-FT-UnB.