



www.relainep.ufpr.br



A PROPOSAL OF STRATIFIED SAMPLING FOR ORIGIN E DESTINATION (O/D) RESEARCH

UMA PROPOSTA DE AMOSTRAGEM ESTRATIFICADA PARA PESQUISA DE ORIGEM E DESTINO (O/D)

Rodrigo R. Freitas^{1✉}, Aryane C. D. Moura¹, Thiago R. Sagawe¹, Fabrício G. Ribeiro¹

¹Centro Federal de Educação Tecnológica, Rio de Janeiro, RJ, Brasil

✉ rodrigo.freitas@cefet-rj.br

Recebido: 17 julho 2020 / Aceito: 24 setembro 2020 / Publicado: 14 dezembro 2020

ABSTRACT

The statistical sampling to affirm a null hypothesis in the degree of confidence should follow precise and correct methodological criteria and processes, because in the occurrence of estimation errors, besides being controlled, it may increase errors in the verification of results. The research aims to evaluate the sampling collected at km 193 of BR-393 in the municipality of Paraíba do Sul, about 154 km from the capital of Rio de Janeiro. The purpose was to verify whether the sampling of the 4 different vehicle groups among 29 different vehicle types had sufficient sample to affirm, whether research of origin and destination - O/D had statistical significance in the degrees of confidence of the research. The result presents three scenarios to simulate the best method. The stratified sampling was the most indicated, as it allows the grouping of homogeneous data in heterogeneous population data. The research data presented problems of underestimation in the number of people interviewed and in the execution time, so much in the applied method and in the simulation of the Arena software.

Keywords: Traffic Volume Count, Toll road, Origin and destination research, Stratified sampling.

RESUMO.

A amostragem estatística para determinar a hipótese de grau de confiança deve seguir os processos e métodos metodológicos precisos e corretos, caso ocorram erros de aplicação, além de controlados, podem aumentar os erros na verificação dos resultados. A proposta da pesquisa foi avaliar a amostragem coletada no km 193 da BR-393 no município da Paraíba do Sul, cerca de 154 km da capital do Rio de Janeiro. O objetivo era verificar, se a amostragem de 4 grupos de veículos distintos entre 29 categorias de veículos possuía um nível suficiente para o teste, se a pesquisa de origem e destino - O/D tinha uma estatística significativa nos graus de confiança da pesquisa. O resultado apresenta três cenários, de modo a simular o melhor método. A amostragem estratificada foi a mais indicada, pois permite o agrupamento de dados homogêneos em dados populacionais heterogêneos. Os dados da pesquisa mostram problemas de subestimação no número de pessoas entrevistadas e no tempo de execução, tanto no método aplicado, quanto na simulação do software Arena.

Palavras-chave: Contagem volumétrica, Praça de pedágio, Pesquisa de origem e destino, Amostragem estratificada.



1 INTRODUÇÃO

A pesquisa de origem e destino (O/D) reproduz informações para o deslocamento de pessoas e transporte de carga, mensurando a quantidade de produtos, o motivo da viagem e a frequência, de modo que foi realizado uma pesquisa em ambos os sentidos para verificar o deslocamento dos diferentes modos de transporte que circulam pela região. A pesquisa foi realizada no km 193 da BR-393 próximo a cidade de Paraíba do Sul, a 13 km do centro da cidade e 154 km da capital do estado do Rio de Janeiro. O IPEA (2016) detalha que a pesquisa de origem e destino reproduz informações para carga transportada e passageiros, pois para o primeiro o mais importante é o produto e sua quantidade transportada, o segundo seria o motivo da viagem e a frequência. Muitos dos casos observados, a pesquisa não satisfaz a veracidade das informações extraídas do universo pesquisado. Uma pesquisa de campo possui elevados custos logísticos, por isso a fase de planejamento deve ser elaborada com maior precisão, a fim de evitar amostragem com resultados insuficiente e interpretações equivocadas.

O fundamento da pesquisa por amostragem é possibilitar o menor custo, assim como o menor número de indivíduos pesquisados, a simulação com cenários é importante para possibilitar a comparação entre as amostras e o nível de confiança. Pesquisas com menor probabilidade de erro, possui um custo elevado para execução, demandando mais tempo e estrutura para a obtenção do resultado. A separação por grupos de veículos possibilita conhecer a quantidade para cada estrato, a fim de reduzir o tempo e o custo de pesquisa ao nível de confiança planejado. A comparação dos métodos de amostragem infinito e finito é importante para verificar o volume de veículos que passam pelo local de pesquisa, pois o método infinito exige um grande esforço para obter um alto nível de confiança, ao modo que o método finito por estratificação, exige menos esforços, contudo não tem a garantia que a quantidade de veículos necessária para cada estrato vai ser suficiente no período de planejamento da pesquisa.

Segundo Chen et al. (2019) e Ekström et al. (2016) ao compararem vários modelos para verificar a acurácia da pesquisa, a amostragem pode obter bom desempenho, se utilizando parte representativa da população, de forma rápida e com baixo custo, porém deve-se observar a existência de erros para apresentá-los como vantagem para o resultado (SPIEGEL, 1975). Deve-se preocupar com três fatores indispensáveis na pesquisa



(MATTAR, 1999). Precisão: buscando a exatidão dos resultados de medições obtidas na amostra em relação à população, levando em consideração o erro amostral, sabendo-se que quanto menor o erro mais precisa será a amostra. Eficiência: quando se refere à medida de comparação entre diversos projetos amostrais, um projeto é mais eficiente do que o outro se, em condições específicas, reportar resultados mais confiáveis, ao mesmo tempo em que produzirá resultados com maior precisão e menor custo. Correção: pela ausência de viés não amostrais, principalmente quando as medidas superestimadas e as subestimadas forem compensadas entre os membros da amostra, possibilitando para que não existam erros sistemáticos numa amostra correta.

Gu et al. (2019) argumenta que na testagem de uma amostragem, a alternância dos preditores pode demonstrar incertezas que antecipam os problemas relacionados com os erros. Desta forma, há importância na construção de cenários para observar o comportamento das variáveis utilizadas e prováveis erros. A amostragem bem elaborada pode equilibrar as ações na direção da certeza probabilística, pois é importante obter pontos de melhora, sugerindo uma oportunidade de inputs de equilíbrio entre a amostragem e a população (FORRESTER et al., 2008). Há formas de verificar o tipo de veículos é a utilização de tecnologias nas estradas ou até mesmo nos veículos. Lim et al. (2017) utilizou sensores e GPS nos pedágios para verificar os tipos de veículos que passavam de forma instantânea. Esse sistema foi utilizado para reduzir os custos de operação e aumentar os requisitos de precisão.

Antes de identificar a população da pesquisa, deve-se verificar os horários e os dias escolhidos, de forma poder comparar com outras pesquisas (VANDENBERGHE et al., 2012; DE FABRITIIS et al., 2008; LI et al., 2009; SHI e LIU, 2010; DRIRA et al., 2016), pois a coleta de informação deve ser discutida para evitar externalidades que possam comprometer os dados (ZHANG et al., 2016; TURCANU et al., 2016; BAIOCCHI et al., 2015; HE e ZHANG, 2016; WONGDEETHAI e SIRIPONGWUTIKORN, 2016).

A pesquisa demanda de modelos probabilísticos que permitam testar um conjunto de dados semelhantes, dentro de conjuntos discrepantes. O modelo de amostragem estratificado é o mais indicado, pois o processo metodológico divide a população em vários tipos de camadas, extrai dados aleatórios para camadas semelhantes (BOLFARINE e BUSSAB, 2005). A amostragem estratificada possui várias vantagens em relação a amostragem aleatória, podendo melhorar a precisão das estimativas para população, assim como para a subpopulação. Contudo, deve-se respeitar as características heterogêneas intrínseca da



população pesquisada (SHI, 2015). Neste caso, a escolha permite: i) uma representação próxima de todos os grupos da população alvo na composição final da amostragem, ii) a população alvo deve ser subdividida em grupos (estratos) com base em características importantes para o estudo, iii) os indivíduos em cada estrato devem ser selecionados aleatoriamente (várias amostras simples aleatórias), iv) diminuir os efeitos (sub ou superestimação) de uma amostra simples aleatória, e v) deve-se conhecer antecipadamente a composição da população em relação a essas características (GRISI e BORBA, 2016).

Cochran (1953) afirma que quando os dados são heterogêneos e não tem como controlar o tipo de veículo por horário e sua origem e destino, é indicada a amostragem estratificada. Para isso, foi estabelecido a estratificação em 4 grupo de veículos entre 29 tipos diferentes. A pesquisa utilizou 4 etapas para determinar a estratificação dos dados (VARÃO, BATISTA e MARTINHO, 2005). Assim, deve-se identificar e definir a população, identificar as variáveis e os estratos para os quais se pretende a representação adequada, classificar os membros da população como membro de cada estrato identificado e selecionar uma amostra em cada um dos estratos da população inquirida.

O objetivo foi verificar se a amostra extraída da pesquisa de O/D possui significância estatística para determinar se o universo de veículos que transitaram é suficiente para determinar a hipótese verdadeira, de modo a formar um conjunto de métodos que auxiliam no planejamento, organização e na verificação dos resultados (SPIEGEL, 1975; TRIOLA, 2009).

A apuração dos dados foi realizada pelo Departamento Nacional de Infraestrutura de Transportes (DNIT), com auxílio do Exército Brasileiro. O local selecionado foi o posto de contagem volumétrica e classificação na BR-393, a antiga “Rio-Bahia” a 154 km da capital do estado do Rio de Janeiro.

2 MÉTODO

O tipo de amostragem será proporcional ao tamanho do estrato pesquisado. Três hipóteses foram destacadas, onde cada unidade tem probabilidade igual de ser selecionada e representativa da população alvo. 1) para amostragem infinitas por proporção populacional, 2) para amostragem finita por proporção populacional e 3) por amostragem estratificada.

1º etapa: Identificar e definir a população: A população pesquisada compreende todos os veículos que passaram pelo posto de contagem e classificação no dia 04 de



julho de 2019, quinta-feira no horário entre 06h:29 até 17h:46 (em um intervalo de 11h:17). A contagem volumétrica e classificação (CVC) é a apuração de todos os veículos que passaram no período mencionado acima. A equipe classificou os 29 veículos diferentes entre transporte de passageiros e carga. A pesquisa foi estratificada em 4 grupos distintos para facilitar o enquadramento estatístico dos dados. Os grupos estão separados em veículos leves, ônibus, caminhão leve e caminhão pesado. Conforme a tabela 1.

TABELA 1 - CLASSIFICAÇÃO DOS VEÍCULOS CONFORME A PESQUISA

Leves (nL)	Ônibus (nO)	Caminhões Leves (nCL)	Caminhões Pesados (nCP)
Quantidade = 1.742	Quantidade = 72	Quantidade = 686	Quantidade = 813
Amostragem = 519	Amostragem = 17	Amostragem = 101	Amostragem = 164
>5% de 29,73% do total	>5% de 23,11% do total	>5% de 14,72% do total	>5% de 20,17% do total

FONTE: os autores (2019).

A quantidade total dos dados foi de 3313 veículos com CVC de 801 veículos na pesquisa de O/D dentro do intervalo mencionado, o tempo médio foi de 2:03s por veículo, e todos os dados possuem uma amostragem acima de 5% da população em cada grupo pesquisado.

2º etapa: Identificar as variáveis e os estratos para os quais se pretende a representação adequada: os 29 tipos de veículos pesquisados foram separados em 4 grupos para melhor enquadramento do processo metodológico. O enquadramento em si, não exclui a necessidade de uma abordagem heterogênea, pois os veículos são semelhantes e não possuem as mesmas características. A preocupação nesta fase era o agrupamento dos veículos em pesos semelhantes, excluindo outras variáveis como eficiência energética, tipo de transporte entre outros.

3º etapa: Classificar a população como membro de cada estrato identificado: automóveis e motos (veículos leves), ônibus, caminhões leves e caminhões pesados. Deve-se observar que a unidade amostral tem probabilidade igual de ser selecionada e ser representativa da população alvo, onde para cada grupo do CVC, houve uma separação por hora.

4º etapa: Selecionar uma amostra em cada um dos estratos da população inquirida: essa etapa é fundamental para escolha do método, de forma a satisfazer a inferência estatística. Segundo Ortúzar e Willumsen (2011) a amostra infinita n para as entrevistas de tráfego podem ser definidas a partir da Eq. 1. O tamanho da amostra com



base na estimativa da média populacional e número de indivíduos é representada valor crítico $Z_{\alpha/2}$ que corresponde ao grau de confiança desejado, desvio-padrão populacional σ da variável estudada, margem de erro ou erro máximo estimado E que identifica a diferença máxima entre a média amostral e a verdadeira média populacional.

$$n = (Z_{\alpha/2} * \sigma / E) \quad (1)$$

O Tamanho da amostra com base na estimativa da média populacional N (população finita < 10000 com amostragem maior 5% da população) ocorre para populações restritas, pois cada unidade dessa população representa uma significativa fração do universo amostral finito de acordo com o Eq. 2 (Miot, 2011).

$$n = N * \sigma^2 (Z_{\alpha/2})^2 / (N - 1)E^2 + \sigma^2 (Z_{\alpha/2})^2 \quad (2)$$

A eq. 3 é orientada para tamanho de amostra em populações finitas com média populacional estratificada e amostragem maior que 5% da população, em que os dados são heterogêneos. Segundo Cochran (1953) e Silva, Morais e Costa (2009) os passos são:

- i) Dividir a população em subgrupos, supondo que exista heterogeneidade entre eles e homogeneidade dentro deles. Utiliza-se a variância fixada em:

$$V = d^2 / Z_{\alpha/2}^2 \quad (3)$$

Onde V é a variância; d^2 é a tolerância; $Z_{\alpha/2}^2$ é a significância obtida utilizando a tabela da distribuição normal.

- ii) Para obtenção do tamanho da amostra, a estimativa inicial:

$$n_0 = (\sum N_h S_h)^2 / V \quad (4)$$

Onde N_h é o tamanho da população e S_h é o desvio padrão, ambos para cada estrato.

- iii) Para populações finitas pode utilizar a correção, observando a fração amostral dada por:



$$f_i = n/N \quad (5)$$

iv) Se f_i for superior a 0,05, a correção da estimativa é dada por:

$$n = n_0/1 + 1/V\sum N_h S_h^2 \quad (6)$$

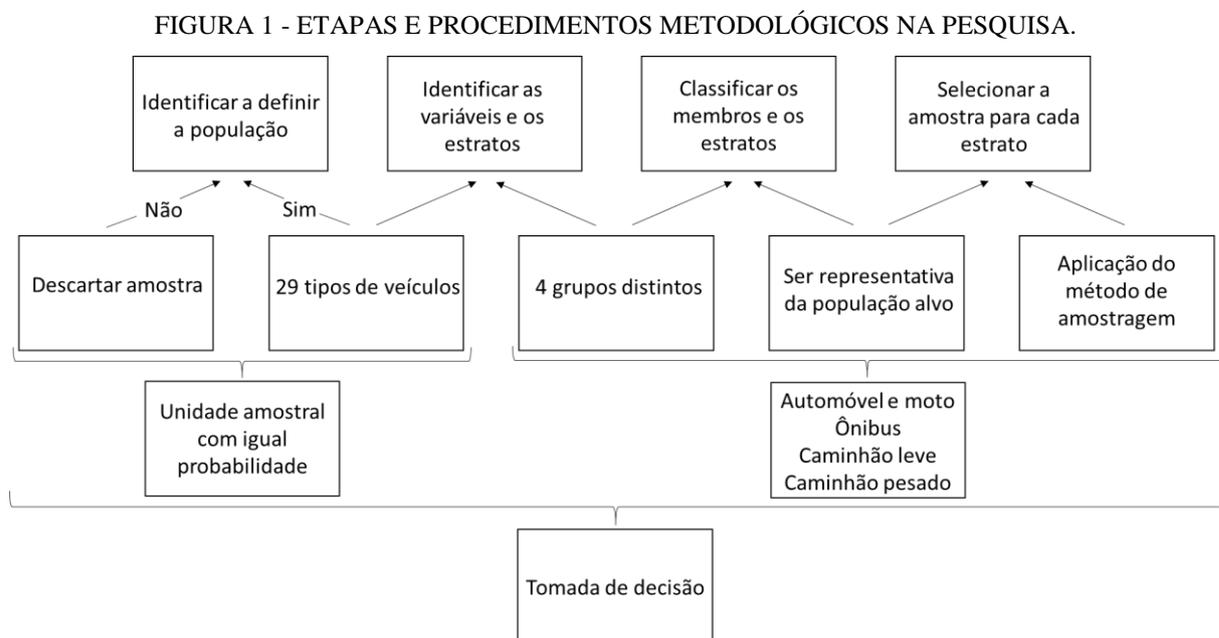
v) Após a obtenção do tamanho da amostra deve-se calcular o peso de cada estrato, denotado por w_h , e distribuir a amostra proporcionalmente:

$$w_h = N_h/N \quad (7)$$

vi) Assim, pode-se calcular a amostra para cada grupo:

$$n_h = w_h n \quad (8)$$

A figura 1 apresenta o procedimento para aplicação das etapas da pesquisa, assim como o fluxo de informação em cada fase.



FONTE: os autores (2019).



3. RESULTADO E DISCUSSÃO

A pesquisa demonstra três cenários, alternando o grau de confiança em todos os grupos pesquisados, a amostragem foi superior a 5% da população. Número de veículos registrados foram de 3313 e veículos parados foram 801. A tabela 2 exemplifica os valores críticos ao nível de confiança na amostra.

TABELA 2 - VALORES CRÍTICOS AO GRAU DE CONFIANÇA NA AMOSTRA

Nível de confiança	$\alpha = E$	Valor crítico
90%	0,10	1,645
95%	0,05	1,96
99%	0,01	2,575

FONTE: os autores (2019).

- Cenário ‘a’ para dados populacionais infinitos:

TABELA 3 - QUANTIDADE NECESSÁRIA DE PESQUISA PARA MÉTODO INFINITO

Número de pessoas entrevistadas			
Cenário a	Nível de confiança		
	90%	95%	99%
	0,08	0,46	19,86

FONTE: os autores (2019).

- Cenário ‘b’ para dados populacionais finitos até 10000:

TABELA 4 - QUANTIDADE NECESSÁRIA DE PESQUISA PARA MÉTODO FINITO >5% E ATÉ 10.000 PESSOAS

Número de pessoas entrevistadas			
Cenário b	Nível de confiança		
	90%	95%	99%
	3301	3311	3313

FONTE: os autores (2019).

- Cenário ‘c’ para dados populacionais infinitos e amostragem estratificada:

TABELA 5 - QUANTIDADE NECESSÁRIA DE PESQUISA PARA MÉTODO FINITO >5% E POR AMOSTRAGEM ESTRATIFICADA

Número de pessoas entrevistadas						
Itens do cenário c	Taxa estratificada	Nível de confiança			Quantidade necessária	Quantidade pesquisada
		Número mínimo necessário				
		90%	95%	99%		
nL	0,527498323	1359	1372	2298	421	519
nO	0,075117371	193	195	195	17	17
nCL	0,192152918	495	499	500	166	101
nCP	0,205231388	529	533	534	197	164
Total	1	2576	2596	2601	801	801

FONTE: os autores (2019).



O cenário ‘a’ demonstra imperfeições para o número populacional da pesquisa, como é definida para população infinita, o resultado para os três graus de confiança não foi necessário pesquisar 20 pessoas para obter o maior nível de confiança. Em simulação computacional, o modelo respondeu para *input* de dados em milhares.

O cenário ‘b’ para dados finitos $>5\%$ e população até 10000, obteve uma pequena resposta próxima da população pesquisada, porém o método não verifica amostragem por agrupamento, podendo apresentar problemas de desbalanceamento.

O cenário ‘c’ é o mais apropriado para a pesquisa proposta, pois separa os dados a serem pesquisados por grupo homogêneos, apontando o número mínimo de dados para obtenção de significância estatística. Porém como é um método estratificado, o número de dados foi bem superior a simulação por amostragem infinita, e a alternância do nível de confiança alteraram pouco o resultado, conforme a tabela 6 em 90%, 95% e 99% de nível de confiança, a comparação da quantidade estratificada para realizada na pesquisa, os grupos nCL e nCP não obtiveram amostragem necessária para satisfazer a hipótese verdadeira.

3.1 SIMULAÇÃO DO MODELO NO SOFTWARE ARENA

As simulações dos dados ocorreram no software Arena versão 14.0, com objetivo de verificar se os *inputs* permitem uma visualização do funcionamento da pesquisa. Os dados de configuração são próprios do sistema utilizado com as seguintes premissas: seria utilizado somente três pesquisadores, a escolha por tipo de veículo seria aleatória e o tempo de duração seria de aproximadamente 11 horas e 30 minutos, e não seria permitida pesquisa no período noturno. No final, os *outputs* deveriam responder as quantidades de veículos suficiente para satisfazer a quantidade estratificada em cada agrupamento.

Na tabela 6, a simulação no software Arena apresentou um *output* de 41400 segundos em atendimento na pesquisa O/D com passagem total de 2982 veículos entre os grupos pesquisados. A simulação obteve um atendimento total 769 veículos (334 sentidos AB e 435 no sentido BA), com tempo médio de atendimento de 2,07s. Esses dados foram inferiores a quantidade pesquisa, logo as premissas não foram suficientes para testar confiabilidade da pesquisa ao nível de confiança de 95%.



TABELA 6: NÚMERO DE VEÍCULOS PARADOS PARA A PESQUISA POR GRUPO

Número de veículos			
Grupo de veículos	Quantidade	Taxa estratificada	Quantidade necessária
nL	1573	0,527498323	830
nO	224	0,075117371	17
nCL	573	0,192152918	110
nCP	612	0,205231388	126
Total	2982	1	1082

FONTE: os autores (2019).

A quantidade simulada exigiu a participação de 1082 veículos, comparando com a quantidade pesquisada por agrupamento, o nL e o nCL não satisfizeram a amostragem mínima. Observa-se que o levantamento dos dados fora prejudicado pela falta de planejamento, pois os parâmetros não foram adequados para o modelo de pesquisa.

4 CONCLUSÃO

A pesquisa apresenta três cenários para a tomada de decisão: o modelo do cenário ‘a’ não contribui estatisticamente, pois pelos dados populacionais e amostrais, não é necessário pesquisar nenhuma pessoa para obter 90% de certeza e apenas 20 pessoas para obter 99%. O modelo do cenário ‘b’, assim como o anterior, não contribui para a pesquisa, pois não satisfaz a condições de número mínimo de veículos por agrupamento. O modelo do cenário ‘c’ apresenta uma amostragem estratificada, pois é mais indicado por separar em camadas heterogêneas, em que dentro de cada camada, os dados são homogêneos e aleatórios.

Na pesquisa, o número de entrevistados, nL e nCL não satisfaz a quantidade mínima proposta pela amostragem estratificada, na tabela 6, a comparação da simulação com a quantidade pesquisada apresenta falhas de planejamento, em que as premissas utilizadas não foram suficientes para atestar a confiabilidade. A utilização do software Arena, na mesma proporção de tempo realizado *in loco*, obteve uma quantidade de veículos próxima da quantidade pesquisada, entretanto, não obteve o número mínimo de entrevistados para a amostragem estratificada. O problema não está no tempo total disponibilizado para pesquisa, sabendo que, o tempo médio por entrevistado na simulação foi de 2,07s, e *in loco* foi de 02:03s.

A pesquisa de campo deveria dispor um número maior de pesquisadores para o intervalo de tempo disponível, justamente para satisfazer o nível de confiança mínimo de 95%. Não existia um plano contingencial e houve problemas na fase de planejamento, pois



não poderia ocorrer nenhum procedimento no período noturno ou intempéries climáticas desfavoráveis.

Para futuras contribuições, orienta-se uma simulação prévia da pesquisa, pois o método por amostragem estratificado exige um grande esforço para obter um número mínimo de entrevistados para cada estrato, a fim de evitar o desnivelamento de veículos em cada agrupamento.

REFERÊNCIAS

- BAIOCCHI, A. et al. Vehicular ad-hoc networks sampling protocols for traffic monitoring and incident detection in intelligent transportation systems. **Transportation Research Part C: Emerging Technologies** 56. 2015. <http://dx.doi.org/10.1016/j.trc.2015.03.041>.
- BOLFARINE, H., BUSSAB, W.O. **Elementos de amostragem**. São Paulo: Blücher, 2005.
- CHEN, X. et al. Surrogate-based optimization of expensive-to-evaluate objective for optimal highway toll charges in transportation network. **Computer-Aided Civil and Infrastructure Engineering**, 29(5), 359–381. 2014. <http://dx.doi.org/10.1111/mice.12058>.
- COCHRAN, W.G. *Sampling Techniques*. New York: Editora Walter A. Shewhart, 1953.
- DE FABRITIIS, C. et al. **Traffic estimation and prediction based on real time floating car data, in: Intelligent Transportation Systems**, 2008. ITSC 2008. 11th International IEEE Conference on, IEEE. pp. 197–203. 2008. <http://dx.doi.org/10.1109/ITSC.2008.4732534>.
- DRIRA, W. et al. **Development and testing of a 3g/lte adaptive data collection system in vehicular networks**. IEEE Trans. Intelligent Transportation Systems 17, 240–249. 2016. <http://dx.doi.org/10.1109/TITS.2015.2464792>.
- EKSTRÖM, J. et al. Surrogatebased optimization of cordon toll levels in congested traffic networks. **Journal of Advanced Transportation**, 50(6), 1008–1033. 2016. <http://dx.doi.org/10.1002/atr.1386>.
- FORRESTER, A. I. J. et al. **Engineering design via surrogate modelling: A practical guide**. Chichester. England: John Wiley & Sons, 2008.
- GRISI, R.; BORBA, M. **Métodos de amostragem. Cálculo do tamanho de amostra**. Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, SP, 2015.
- GU, S., KELLY, B., XIU, D. Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies* 33. 2020. 2223-2273. <http://dx.doi.org/10.1093/rfs/hhaa009>.
- HE, Z., ZHANG, D. Cost-efficient traffic-aware data collection protocol in (VANET). **Ad Hoc Networks**. 2016. <http://dx.doi.org/10.1016/j.adhoc.2016.09.021>.



www.relainep.ufpr.br



- LI, X. et al., 2009. **Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring**. Vehicular Technology, IEEE Transactions on 58, 1647–1653. 2009. <http://dx.doi.org/10.1109/TVT.2008.2005775>.
- LIM, J. H., CHOI, K. H., CHO, J., LEE, H. K. Integration of GPS and monocular vision for land vehicle navigation in urban area. International Journal of Automotive Technology 18. 2017. 345–356. <https://doi.org/10.1007/s12239-017-0035-3>.
- MATTAR, F.N. **Pesquisa de marketing: metodologia, planejamento (5º ed.)**. Atlas, São Paulo, 1999.
- MIOT, H.A. Tamanho da amostra em estudos clínicos e experimentais, **Jornal. Vascular Brasileiro**, Vol. 10, Nº 4. 2011.
- ORTÚZAR, J.D.; WILLUMSEM, L.G. **Modelling Transport (4º ed.)**. Wiley, 2011.
- Instituto de pesquisa Econômica e Aplicada – IPEA. **Relatório de pesquisa. Desenho da pesquisa origem-destino do transporte rodoviário no Brasil**. Ministério do Planejamento, Desenvolvimento e Gestão, DF, 2016.
- SHI, F. Study on a Stratified Sampling Investigation Method for Resident Travel and the Sampling Rate. **Discrete Dynamics in Nature and Society**, Volume 2015, Article ID 496179, 7 pages. 2015. <http://dx.doi.org/10.1155/2015/496179>.
- SHI, W., LIU, Y. Real-time urban traffic monitoring with global positioning system-equipped vehicles. **IET intelligent transport systems** 4, 113–120. 2010. <http://dx.doi.org/10.1049/iet-its.2009.0053>.
- SILVA, A.H.A., MORAIS, R. M., COSTA, M. J; C. **Plano amostral utilizando amostragem estratificada juntamente com amostragem sistemática para aplicação do Inquérito Nutricional do município de João Pessoa, Paraíba**. Anais da II Escola de Amostragem e Metodologia de Pesquisa, (ESAMP), 2009.
- SPIEGEL, M.R. **Estatística**, Coleção Schaum, (1ª ed.). São Paulo: Editora McGraw-HILL do Brasil, 1975.
- TRIOLA, M.F. **Introdução à estatística. (10ª ed.)** LTC, Rio de Janeiro. pp. 2-31, 2008.
- TURCANU, I. et al. An integrated vanet-based data dissemination and collection protocol for complex urban scenarios. **Ad Hoc Networks** 52, 28–38. 2016. <http://dx.doi.org/10.1016/j.adhoc.2016.07.008>.
- VANDENBERGHE, W. et al. Feasibility of expanding traffic monitoring systems with floating car data technology. **Intelligent Transport Systems, IET** 6, 347–354. 2012. <http://dx.doi.org/10.1049/iet-its.2011.0221>.
- VARÃO, C. et al. **Métodos de Amostragem**. Departamento de Educação FCUL, Universidade de Lisboa, Lisboa, 2005.
- WONGDEETHAI, S., SIRIPONGWUTIKORN, P. Collecting road traffic information using vehicular ad hoc networks. **EURASIP Journal on Wireless Communications and Networking**, 9. 2016. <http://dx.doi.org/10.1186/s13638-015-0513-0>.
- ZHANG, Q. et al. **An autonomous information collection and dissemination model for large-scale urban road networks**. IEEE Trans. Intelligent Transportation Systems 17, 1085–1095. 2016. <http://dx.doi.org/10.1109/TITS.2015.2497338>.