



UNICAMP

Trabalho ME415

255677 - Gabriella de Oliveira Argenton
204274 - Gabriela Namie Hidaka

Campinas
2025

1 Resumo

O presente estudo tem como objetivo comparar a Amostragem Aleatória Simples Sem Reposição (AASs) e a Amostragem Estratificada (AE) na estimativa da média de ocupação de vagas de estacionamento em três institutos da Unicamp: IMECC, IFGW e IQ. A coleta de dados foi realizada em três períodos do dia às 10h (manhã), 14h (tarde) e 18h (noite) de segunda a sexta-feira. A hipótese do estudo é que a AE, ao considerar as diferenças entre horários (estratos), resulte em estimativas mais precisas e com menor desvio-padrão do que a AASs. Como era esperado, os resultados demonstraram que quando há discrepância entre os estratos, a AE produz médias mais precisas: próximas da média populacional e com menor desvio padão, mostrando-se mais eficiente para contextos com variação entre grupos.

2 Introdução

Os estacionamentos universitários costumam apresentar grande variação na ocupação ao longo do dia. Em horários de pico, a alta demanda por vagas causa congestionamentos e dificuldades de mobilidade, enquanto em outros períodos os espaços não são aproveitados ao máximo. Entender esse comportamento é essencial para aprimorar o planejamento interno e a gestão dos estacionamentos dentro da Unicamp, contribuindo para uma melhor utilização dos espaços e organização do fluxo de veículos no campus.

Neste estudo, a população considerada é composta pelos estacionamentos de três institutos da Unicamp, escolhidos como unidades representativas do campus. Embora medir todos esses estacionamentos seja viável, a amostragem estatística foi utilizada com o intuito de simular um cenário real de pesquisa, em que a medição de toda a população seria inviável. Assim, foi possível comparar diferentes planos amostrais e avaliar seu impacto nas estimativas obtidas.

O estudo compara dois métodos de amostragem: a Amostragem Aleatória Simples sem Reposição (AASs) e a Amostragem Estratificada (AE), buscando analisar como a escolha do plano amostral influencia a precisão das estimativas.

A hipótese principal é que a AE, ao considerar as variações entre horários (estratos), produzirá médias mais precisas e com menor desvio padrão do que a AASs. O objetivo geral é comparar o desempenho dos dois planos amostrais na estimativa da média de ocupação dos estacionamentos. Especificamente, busca-se avaliar o desvio padão das médias obtidas em cada método, calcular o ganho de eficiência da AE e ilustrar a aplicação prática de conceitos de amostragem em um contexto real.

3 Método

3.1 Desenho Experimental

O experimento foi realizado nos estacionamentos dos institutos IMECC, IFGW e IQ, localizados no campus da Unicamp (Campinas). A coleta de dados ocorreu de segunda a sexta-feira, nos períodos da manhã (10h), tarde (14h) e noite (18h). Em cada local e horário, foram registrados:

- O número total de vagas disponíveis;
- O número de vagas ocupadas;
- O dia da semana;
- O horário do dia.

A variável dependente é a proporção de ocupação das vagas (número de vagas ocupadas dividido pelo total de vagas). A variável independente é o plano amostral utilizado (AASs ou AE). As variáveis controladas incluem o dia da semana e o horário, para reduzir a influência de fatores externos.

A amostragem foi conduzida de duas formas:

Amostragem Aleatória Simples Sem Reposição(AASs): as observações foram selecionadas de maneira completamente aleatória dentro da população total.

Amostragem Estratificada (AE): os dados foram divididos em estratos definidos por horário. As médias foram calculadas dentro de cada estrato e ponderadas conforme o tamanho de cada grupo.

Esse delineamento permitiu comparar diretamente o desempenho dos dois planos amostrais, minimizando vieses e garantindo representatividade entre os períodos do dia.

3.2 Materiais e Equipamentos

Foram utilizados os seguintes materiais e equipamentos ao longo do experimento:

- Planilha para registro de dados (Google Sheets);
- Computador com software estatístico R para análise dos dados.

3.3 Procedimento Experimental

O experimento foi conduzido nos estacionamentos dos institutos IMECC, IFGW e IQ, todos localizados no campus da Unicamp (Campinas/SP). Inicialmente, foram mapeados os estacionamentos de cada instituto, identificando o número total de vagas disponíveis e delimitando as áreas observadas para garantir consistência entre as medições.

A coleta de dados ocorreu ao longo de segunda a sexta-feira, em três horários fixos representando diferentes períodos do dia: 10h (manhã), 14h (tarde) e 18h (noite). Em cada sessão de coleta, foram anotadas as seguintes informações:

- Número total de vagas existentes no estacionamento;
- Número de vagas ocupadas no momento da observação;
- Dia da semana e horário da coleta;

As observações foram registradas em uma planilha digital (Google Sheets), garantindo organização e rastreabilidade das informações. Para minimizar possíveis erros de contagem, as medições foram realizadas de forma direta por observação em cada local.

Após o término da coleta, os dados foram organizados e processados em ambiente computacional(linguagem R). Com base nas informações de ocupação, foram simuladas duas estratégias de amostragem:

Amostragem Aleatória Simples (AASs): as observações foram sorteadas pelo R de forma totalmente aleatória a partir do conjunto total de dados disponíveis, sem considerar a origem (dia da semana, instituto ou horário).

Amostragem Estratificada (AE): os dados foram divididos em estratos por horário de coleta (cujos níveis são manhã, tarde e noite). Para cada estrato, calculou-se a média e o desvio padrão da proporção de ocupação, ponderando os resultados de acordo com o tamanho relativo de cada grupo.

Por fim, as médias e variâncias obtidas em cada método foram comparadas a fim de avaliar a precisão das estimativas e o ganho de eficiência da AE em relação à AASs. Essa etapa permitiu verificar empiricamente se a consideração das diferenças entre horários realmente resulta em estimativas mais estáveis e representativas da ocupação geral de alguns dos estacionamentos da Unicamp.

3.4 Metodologia e análise de dados

A hipótese a ser investigada foi que a amostragem estratificada apresenta maior precisão e menor desvio padão do que a AASs na estimação da média populacional.

Foi utilizado um conjunto de dados reais contendo 45 observações, sendo 15 para cada período do dia. A variável de interesse foi a proporção de vagas ocupadas em cada momento de observação.

As variáveis foram padronizadas no R, convertendo-se Dia e Instituto em fatores, Horário em fator ordenado (manhã, tarde, noite) e Quantidade em valor numérico correspondente à proporção de ocupação.

Inicialmente, foi calculada a média populacional da proporção de vagas ocupadas, que serviu como referência (“verdade populacional”) para avaliar o desempenho dos

estimadores. Também foram obtidas a variância e o desvio-padrão populacionais. Além disso, os estratos foram definidos segundo o período do dia e calculados seus respectivos tamanhos, médias, desvio padão e pesos relativos.

Adotou-se um plano com 5000 repetições, utilizando dois métodos de estimação: AAS sem reposição, com tamanho amostral total n=6

$$\text{Media da AASs : } \bar{y}_{AASs} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$\text{Desvio - Padrao da AASs : } s_{AASs} = \sqrt{\left(1 - \frac{n}{N}\right) \frac{S^2}{n}}$$

- n : tamanho da amostra;
- N : tamanho da população;
- S^2 : variância populacional;
- $(1 - \frac{n}{N})$: fator de correção para população finita.

Amostragem estratificada proporcional, alocando $n = 2$ observações em cada um dos três estratos (manhã, tarde e noite)

$$\text{Media da AE : } \bar{y}_{AE} = \sum_{h=1}^H W_h \bar{y}_h$$

$$\text{Desvio - Padrao da AE : } s_{AE} = \sqrt{\sum_{h=1}^H W_h^2 s_h^2}$$

- H : número total de estratos (ex.: manhã, tarde e noite);
- $W_h = \frac{N_h}{N}$: peso do estrato h na população;
- \bar{y}_h : média amostral do estrato h ;
- N_h : tamanho populacional do estrato h ;
- $W_h = \frac{N_h}{N}$: peso do estrato h ;
- n_h : tamanho da amostra selecionada no estrato h ;

- S_h^2 : variância dentro do estrato h ;
- $\left(1 - \frac{n_h}{N_h}\right)$: fator de correção finita dentro de cada estrato.

Para cada repetição, foi gerada uma amostra segundo cada plano amostral e calculada a média estimada.

Foram computadas e comparadas entre os planos amostrais:

Média das estimativas; Desvio Padrão; Erro Quadrático Médio (EQM); Viés; Eficiência relativa ($RE = \text{Var(AAS)}/\text{Var(AE)}$), Variâncias teóricas e empíricas

Gráficos de boxplot e densidade foram construídos para ilustrar visualmente a distribuição das estimativas e sua dispersão em relação à média populacional.

4 Resultados

Parâmetros populacionais

A população do estudo possui $N=45$ observações. A proporção média populacional observada foi:

Média populacional: 0,6333

Desvio padrão populacional: 0,2598

Observou-se que o período da tarde apresentou maior ocupação média (0,814), seguido da manhã (0,756) e, por fim, da noite (0,331), confirmando uma clara heterogeneidade entre estratos — reforçando o uso da amostragem estratificada.

Desempenho dos estimadores

A tabela abaixo resume os resultados da simulação com $B=5000$ réplicas e $n=6$:

Tabela 1: Desempenho dos estimadores com $B = 5000$ e $n = 6$

Plano Amostral	Média	Desvio Padão	EQM	Viés	RE
AAS (s/ reposição)	0.636	0.0984	0.00969	0.00251	1.00
Estratificada	0.635	0.0555	0.00309	0.00118	3.14

Os resultados mostram que:

Ambas as metodologias produziram estimativas sem viés relevante, próximas da média verdadeira (0,6333).

O desvio padrão do estimador estratificado (0,0555) foi aproximadamente duas vezes menor que a da AAS (0,0984).

O EQM acompanhou essa diferença, sendo também consideravelmente menor para a AE.

A eficiência relativa ($RE = 3,14$) indica que a amostragem estratificada foi 314% mais eficiente do que a AASs, ou seja, produziria a mesma precisão com uma amostra 3 vezes menor.

Tabela 2: Comparação entre variâncias teóricas e empíricas

Plano	Var. Empírica	Var. Teórica
AAS (s/ reposição)	0.00969	0.00975
Estratificada	0.00308	0.00304

Os valores teóricos e empíricos foram praticamente idênticos, validando as fórmulas analíticas e o processo de simulação.

Os gráficos gerados (boxplot e curvas de densidade) evidenciaram que as estimativas da amostragem estratificada são mais concentradas em torno da média populacional, com menor dispersão, confirmando visualmente sua superior precisão.

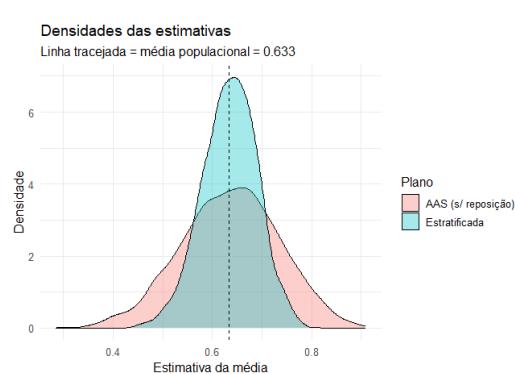


Figura 1: *
Figura 1: Densidades das estimativas

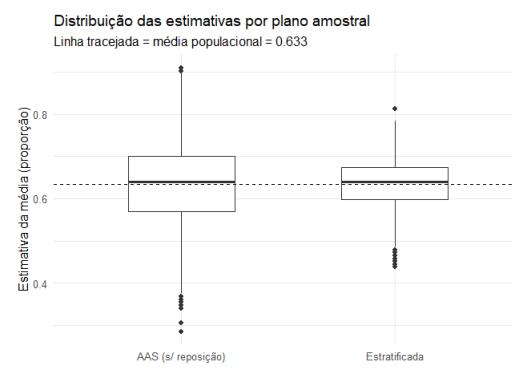


Figura 2: *
Figura 2: Boxplot das estimativas por plano amostral

5 Conclusão

O objetivo deste estudo foi comparar o desempenho da Amostragem Aleatória Simples sem reposição (AASs) e da Amostragem Estratificada (AE) na estimativa da proporção média de vagas ocupadas nos estacionamentos da Unicamp. A partir de dados reais coletados nos períodos da manhã, tarde e noite, e utilizando um experimento de simulação com 5000 repetições, foi possível avaliar a precisão e a variabilidade dos estimadores associados a cada método amostral.

Os resultados demonstraram que, embora ambos os métodos tenham produzido estimativas não viesadas e próximas da média populacional, a amostragem estratificada apresentou desempenho superior. Especificamente, a AE apresentou desvio padrão quase duas vezes menor do que a AASs, além de menor Erro Quadrático Médio (EQM) e maior eficiência relativa. Esses resultados confirmam empiricamente a teoria estatística, que estabelece que quando há heterogeneidade entre os estratos — como observado entre os diferentes horários de coleta — a estratificação tende a produzir estimativas mais precisas.

Assim, conclui-se que a estratificação por horário foi eficaz em reduzir a variabilidade da estimativa, evidenciando que a Amostragem Estratificada é mais adequada e eficiente para estudos desse tipo, onde grupos naturais apresentam padrões distintos de comportamento. Portanto, o presente trabalho reforça a importância de um planejamento amostral adequado e da escolha de técnicas que aproveitem a estrutura da população, quando conhecida.

6 Apêndice

Link para o código das análises (feitas no RStudio)

<https://github.com/g255677-gabriella/metodo-cientifico-codigos.git>