

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS
Bacharelado em Ciências da Computação**

**ANDRÉ LUIZ DA SILVA JUNIOR
ANGELO MONTE SERRAT BONINI
DAVI AUGUSTO NEVES LEITE
LUIS FELIPE MUNIZ DE ANGELO
THOMAS TOBIAS DOS SANTOS**

**TRABALHO SOBRE AMOSTRAGEM PROBABILÍSTICA 1: AMOSTRAGEM CASUAL
SIMPLES E AMOSTRAGEM SISTEMÁTICA**

**Bauru
2020**

ANDRÉ LUIZ DA SILVA JUNIOR
ANGELO MONTE SERRAT BONINI
DAVI AUGUSTO NEVES LEITE
LUIS FELIPE MUNIZ DE ANGELO
THOMAS TOBIAS DOS SANTOS

**TRABALHO SOBRE AMOSTRAGEM PROBABILÍSTICA 1: AMOSTRAGEM CASUAL
SIMPLES E AMOSTRAGEM SISTEMÁTICA**

Trabalho apresentado ao Curso de
Bacharelado em Ciências da
Computação, como parte do requisito
para a aprovação na disciplina de
Probabilidade e Estatística.
Docente: Prof. Me. Neylan Leal Dias

Bauru
2020

SUMÁRIO

1 INTRODUÇÃO	4
2 DESENVOLVIMENTO	5
3 CONCLUSÃO.....	9
REFERÊNCIAS.....	10

1 INTRODUÇÃO

A coleta e análise de dados têm como principal objetivo auxiliar na tomada de decisões ou comprovar fatos de problemas da sociedade. Tem-se exemplos como: pesquisas científicas, de mercado ou determinações de taxas (econômicas ou sociais). Além disso, em grande parte dos casos, é ideal e necessário recorrer à coleta dos dados diretamente com a chamada população de interesse, ou seja, todos os indivíduos relevantes ao estudo. Contudo, torna-se inviável o acesso e a obtenção desses dados (haja visto a quantidade de sujeitos envolvidos) sendo, portanto, utilizado um conjunto de técnicas para a aquisição de dados da chamada amostra, ou seja, uma parcela da população de interesse.

Sendo assim, para que a pesquisa fique comprometida com os fatos é necessário que a amostra seja representativa da população. Para tal, a amostragem (processo de obtenção de amostras) pode ser realizada em dois tipos: a probabilística, em que os elementos da população têm probabilidade conhecida e diferente de zero, de tal forma que cada elemento tem chance igual e independente de ser selecionado para a amostra; e a não probabilística, em que não é possível generalizar os resultados para a população, haja visto que as amostras selecionadas não garantem a representatividade do todo.

A partir disso, o intuito deste trabalho consiste em explicar dois tipos de amostragem probabilística, sendo elas: casual simples (ou aleatória) e sistemática. Para isso, além das definições, serão apresentados vantagens, desvantagens e exemplos para cada tipo de amostragem.

2 DESENVOLVIMENTO

- **Amostragem Casual Simples**

Segundo Salgado (2019), Valgas (2007) e Filho (2013), a amostragem casual simples (ou aleatória) se fundamenta no princípio de que todos os elementos da população têm a mesma probabilidade de serem incluídos na amostra, sendo o método mais simples e utilizado. Além disso, possui duas técnicas para se obter uma determinada amostra, sendo elas: **o método do sorteio**, em que são escolhidos um a um, de maneira aleatória e sem repetição, até completar o máximo desejado; e o **método da tabela de números aleatórios**, em que os elementos são atribuídos com números e sorteados em uma tabela, sendo selecionados conforme a amostra pretendida.

A principal vantagem consiste na facilidade de compreensão, haja visto que possui métodos fáceis de serem realizados. Contudo, resulta em menor precisão e com maiores erros quando relacionado com outras técnicas de amostragem, podendo resultar em uma amostra não representativa da população.

No método do sorteio, os elementos da população são organizados em uma lista, numerados a partir da quantidade total de elementos e, enfim, sorteados. Dessa forma, ocorre a garantia de uma mesma probabilidade de um elemento ser sorteado e não haver repetição. Por exemplo: considere que um pesquisador deseja realizar a média da idade de 10 pessoas que residem em um bairro com 30 pessoas no total. As idades dessas 30 pessoas são: 25, 20, 35, 21, 22, 24, 25, 30, 38, 24, 20, 20, 25, 20, 19, 25, 23, 24, 28, 24, 24, 22, 28, 26, 23, 25, 22, 27, 25, 23.

Para a obtenção de uma amostra aleatória simples, utilizando o método do sorteio, basta pegar um elemento aleatoriamente e sem a repetição deste (excluindo-o da população após a escolha), a fim de se completar o máximo desejado pelo pesquisador ($n = 10$). Sendo assim, após a aplicação do método, tem-se: amostra = {20, 24, 22, 28, 23, 24, 21, 20, 25, 27}. Lembrando que o valor do elemento, neste caso a idade, pode-se repetir; mas não há a escolha do mesmo elemento (da mesma posição) posteriormente. Por exemplo: suponha que três pessoas X, Y e Z têm idades de 24 anos e deseja-se selecionar apenas duas delas para realizar uma pesquisa. O pesquisador, ao aplicar o método do sorteio, obterá dois valores de 24 anos, porém poderia haver combinações diferentes, tais como: X e Y, X e Z, Z e X, Y e Z, dentre outros. Ou seja: ele pode repetir o valor em sua amostra final, dependendo do que deseja obter, mas deve excluir o elemento da população inicial após a escolha.

Contudo, o método do sorteio demonstra-se ineficiente e não prático para grandes populações, sendo utilizado o método da tabela de números aleatórios. Esse método consiste em: atribuir uma chave (número) aleatoriamente aos elementos da população de interesse; separar em uma tabela em ordem numérica da chave; sortear e agrupar pela quantidade de dígitos da chave (na vertical ou horizontal), excluindo as chaves repetidas.

Utilizando o exemplo anterior, tem-se a seguinte tabela (chave – elemento da população):

00 – 25	06 – 25	12 – 20	18 – 28	24 - 23
01 – 20	07 – 30	13 – 19	19 – 24	25 – 25
02 – 35	08 – 38	14 – 25	20 – 24	26 - 22
03 – 21	09 – 24	15 – 23	21 – 22	27 – 27
04 – 22	10 – 20	16 – 20	22 – 28	28 – 25
05 – 24	11 – 25	17 – 24	23 – 26	29 - 23

Após isso, utilizando o $n = 10$, sorteia-se 2 a 2 (pela utilização de chave com 2 dígitos) qualquer número, começando de qualquer ponto na vertical ou horizontal. Somente deve-se selecionar para a amostra final, as chaves atribuídas a algum elemento da população.

Desta forma: **09, 11**; 51, 86; 35, 59; **12, 25**; ...; **13, 20**; **21**, 13 (este segundo é descartado pela repetição); até se obter a amostra desejada (neste caso, 10 elementos). Com a utilização do processo anterior, obteve-se: amostra = {24, 25, 20, 25, 19, 24, 22, 25, 24, 28}.

- **Amostragem Sistemática**

Segundo Salgado (2019), Machado (2012) e Varão, Batista e Martinho (2005), a amostragem sistemática se fundamenta em uma população homogênea, tendo os dados organizados e identificados por posição. Desta forma, retira-se os elementos periodicamente com base em uma razão, dada por: $R = N/n$, em que “N” é o tamanho da população e “n” é o tamanho da amostra.

Ou seja, primeiramente é organizado e ordenado o conjunto de dados, por meio de uma identificação da posição. Após isso, é realizado o cálculo para encontrar a **razão** em que os dados serão colhidos (com base no número desejado da amostra) de maneira periódica. Após a obtenção da razão, sorteia-se um número entre 1 e R (inclusivo) e, a partir disso, obter a amostra com a seguinte sequência (em que X é o número sorteado): X (entre 1 e R); X + R; X + 2R; e assim sucessivamente.

Uma vantagem desse método está na facilidade de sua execução, necessitando de menos tempo do que a amostragem casual simples. Contudo, esse tipo de amostragem não seleciona os elementos ao acaso (exceto o primeiro), haja visto que existe um intervalo fixo.

Um exemplo de aplicação desse tipo de amostragem pode ser, considerando uma população com 100 idades (N) que vão de 1 até 100 e desejando-se obter uma amostra com 10 idades: primeiramente, é necessário verificar a organização dos dados e identificação de suas respectivas posições. Como, neste caso, os dados estão organizados de maneira crescente e cada posição equivale a idade, basta encontrar o valor da razão R: $R = 100/10 = 10$. Sendo assim, escolhe-se um número aleatório entre 1 e 10 (inclusivo). Sendo 2 o número sorteado, tem-se o seguinte espaço amostral: 2, 12, 22, 32, 42, 52, 62, 72, 82, 92.

3 CONCLUSÃO

Em síntese, as técnicas de amostragem apresentadas possibilitam a coleta de dados de forma estável e eficiente. Desta forma, é conferido igual probabilidade na escolha dos elementos ou indivíduos que serão utilizados para a pesquisa. Contudo, as técnicas de amostragem casual simples e a amostragem sistemática podem comprometer à representatividade final da amostra com a população de interesse, quando relacionados a um espaço populacional muito grande e variado. Ainda assim, as técnicas apresentadas conferem simplicidade e velocidade na obtenção da amostra, conferindo desta forma que os resultados não sejam comprometidos no processo de análise de dados.

REFERÊNCIAS

FILHO, Luiz Medeiros de Araújo Lima. **Amostragem**. Universidade Federal da Paraíba. 2013. Disponível em: <http://www.de.ufpb.br/~luiz/Adm/Aula9.pdf>. Acesso em: 22/08/2020.

MACHADO, Nivea da Silva Matuda. **Amostragem Sistemática (AS)**. Universidade Federal do Paraná. 2012. Disponível em: <https://docs.ufpr.br/~niveam/ce070/exemplo%20geral%20AS.pdf>. Acesso em: 22/08/2020.

SALGADO, Manuel Henrique. Amostragem - Distribuições Amostrais. In: SALGADO, Manuel Henrique. **Probabilidade e Estatística**. 2019. cap. 4, p. 54-55.

VARÃO, Carla; BATISTA, Cláudia; MARTINHO, Vânia. Métodos de amostragem. **Departamento de Educação FCUL. Metodologia de Investigação I**, v. 2006, 2005. Disponível em: <http://www.educ.fc.ul.pt/docentes/ichagas/mi2/metodosamostragemt2.pdf>. Acesso em: 22/08/2020.

VALGAS, Ricardo. **Amostragem Aleatória Simples**. 2007. Disponível em: <https://docs.ufpr.br/~ricardo.valgas/amostragem/aleatoria.pdf>. Acesso em: 22/08/2020.