



Universidade Estadual de Santa Cruz

Pavilhão de Ciências Exatas e Tecnológicas

Curso de Ciência da Computação

NOÇÕES DE AMOSTRAGEM

Métodos Probabilísticos

CET083 – Probabilidade e Estatística.

Prof. José Cláudio Faria.

Discentes: Levi Freitas e Eduardo Oliveira.

2023.2.

Roteiro da apresentação

Téoria:

Apresentação sobre o tema:

- Termos importantes 1 e 2;

Tipos de amostragem:

- Métodos de amostragem;
- Seleção de amostragem;
- Tamanho da amostra 1, 2 e 3;
- Tipos de amostras;

Prática:

Exemplos de amostragens:

- Amostragem Aleatória Simples;
- Amostragem Sistemática;
- Amostragem Estratificada;
- Amostragem por Conglomerados;
- Amostragem por Área;

Códigos de exemplos

Termos importantes - 1

População

A população é o grupo completo de elementos que se deseja estudar. Por exemplo, quando queremos realizar um estudo sobre a idade de todas as pessoas em uma cidade, a população é a população total dessa cidade.

Amostra

A amostra é um subconjunto da população, selecionado de forma aleatória, que representa a população.

População x Amostra



Termos importantes - 2

Censo

Um censo envolve a coleta de dados de todos os membros da população. Isso proporciona uma visão completa e precisa de toda a população, eliminando a incerteza associada à amostragem. No entanto, isso pode ser demorado, caro e, às vezes, inviável. Contudo, os resultados de um censo são considerados precisos, desde que a coleta de dados seja completa e precisa.

Amostragem

São coletados dados de um subconjunto representativo da população, sendo assim mais eficiente em termos de tempo, recursos e custos. O que permite que você obtenha uma estimativa razoavelmente precisa da população com um esforço menor. Com isso é possível fazer inferências e generalizações sobre toda a população.

Censo x Amostragem



Censo



Amostragem

Métodos probabilísticos

Definição:

A amostragem probabilística é um método de seleção de amostras de uma população, garantindo que cada elemento tenha uma chance conhecida e não nula de ser escolhido. Isso garante que a amostra seja representativa da população, permitindo fazer inferências estatísticas com base nos resultados da amostra.

Quando usar:

Ela é a escolha adequada sempre que a representatividade e a confiabilidade dos resultados são cruciais. Assim, ela ajuda a evitar o viés na seleção da amostra e permite que se façam inferências estatísticas válidas sobre a população.

Alguns cenários onde ela pode ser utilizada são:
Pesquisas Científicas, estudos estatísticos, estudos de saúde pública e avaliações educacionais

Vantagens

Representatividade

Garante que cada elemento da população tenha uma chance conhecida e não nula de ser selecionado

Viés Reduzido

O risco de viés na seleção da amostra é minimizado.

Reprodutibilidade

Ao repetir o processo de seleção de amostra, é provável que obtenha resultados semelhantes.

Métodos não probabilísticos

Definição:

A seleção dos elementos da população para compor a amostra depende ao menos em parte do julgamento do pesquisador ou do entrevistador no campo.

Quando usar:

Métodos não probabilísticos de seleção de amostra são recomendados em situações em que a amostragem probabilística não é viável, prática ou apropriada. Embora a amostragem probabilística seja geralmente preferida devido às suas vantagens em termos de representatividade e validade estatística, há momentos em que métodos não probabilísticos podem ser úteis, como nos casos de: Recursos Limitados, acesso restrito aos dados, estudos exploratórios ou em pesquisas piloto.

Vantagens

Facilidade

Para pessoas leigas que realizam pesquisas, esse método é o mais prático.

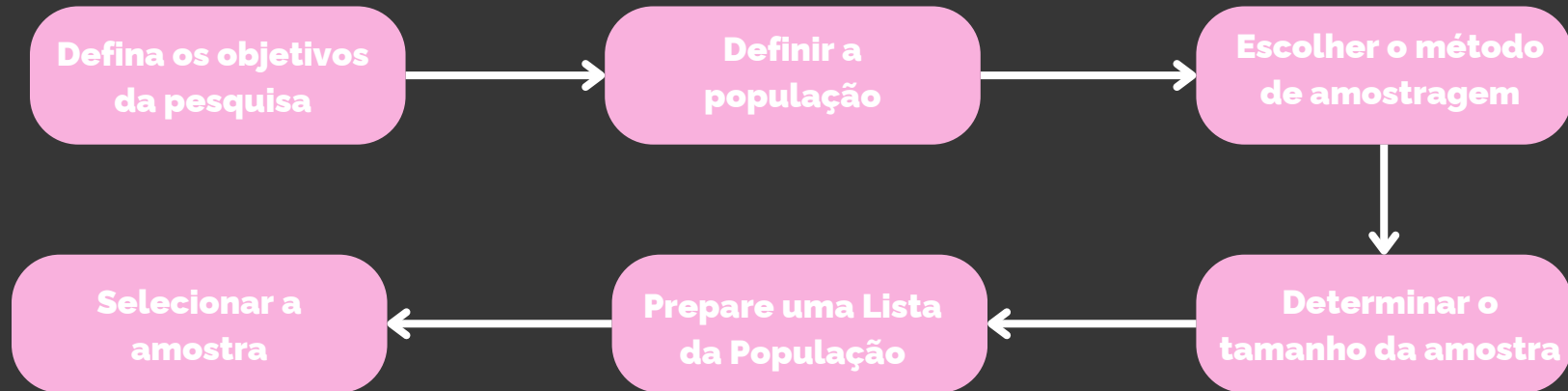
Rapidez

Fator que pode ser de extrema importância em determinados cenários.

Economia de recursos

Requerem menos recursos financeiros, de tempo

Seleção de uma amostragem



Como definir o tamanho da amostra? - 1

A determinação do tamanho da amostragem é um aspecto crítico da seleção de amostragem e depende de vários fatores, dentre eles:

- **Definir a margem de erro (Máximo tolerável):** expressa como uma porcentagem da população.
- **Escolha do nível de confiança:** é a probabilidade de que a verdadeira média ou proporção da população esteja dentro da margem de erro da estimativa da amostra.
- **Variabilidade na População:** Pode ser obtido por meio de dados históricos, estudos anteriores ou informações sobre a população.

Como definir o tamanho da amostra? - 2

Uma vez que os parametros tenham sido estabelecidos, torna-se possível utilizar a **Fórmula de Tamanho da Amostra**, dada por:

$$n = \frac{Z^2 * p * (1-p)}{E^2}$$

Onde:

- n é o tamanho da amostra necessário.
- Z é o valor crítico associado ao nível de confiança escolhido.
- p é a estimativa da proporção na população (se estiver calculando proporções) ou a estimativa da variância na população (se estiver calculando médias).
- E é a margem de erro desejada.

Como definir o tamanho da amostra? - 3

Contudo, ela não considera o tamanho da população, pois é usada quando a população é grande o suficiente. No entanto, quando a população é relativamente pequena em comparação com o tamanho da amostra, devemos ajustá-la. Com a seguinte fórmula, conhecida como **Correção de População Finita**:

$$n_f = \frac{n}{1 + (n-1) / N}$$

Onde:

- n_f é o tamanho da amostra corrigido.
- n é o tamanho da amostra calculado usando a fórmula padrão (sem considerar o tamanho da população).
- N é o tamanho da população.

Tipos de amostragem

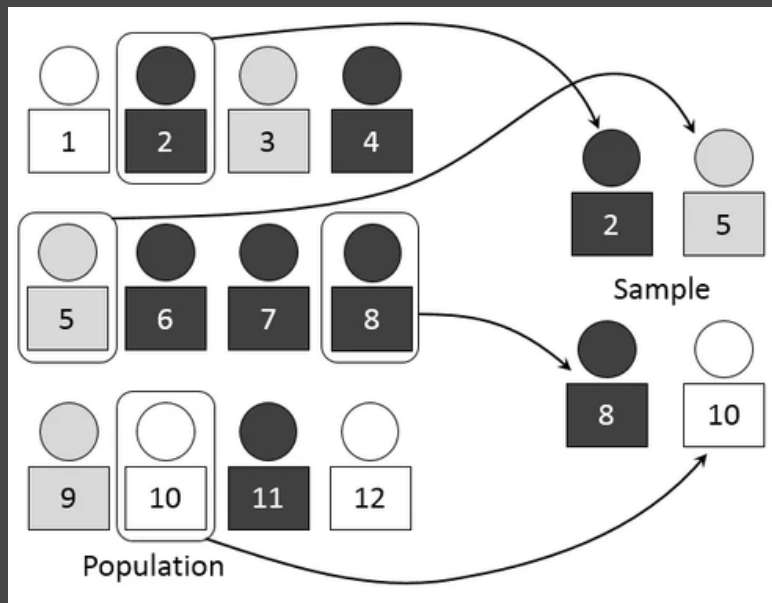
Probabilísticas

- Amostragem Aleatória simples;
- Amostragem Sistemática;
- Amostragem Aleatória estratificada;
- Amostragem Conglomerado;
- Amostragem de Área;

Não Probabilísticas

- Amostragem Conveniência;
- Amostragem Intencional;
- Amostragem Bola de neve;
- Amostragem Quota;

Amostragem aleatória simples



Amostragem aleatória simples

Definição:

A AAS é a escolha de uma amostra de uma população em que cada elemento da população tem a mesma probabilidade de ser selecionado.

Como utilizar:

- 1º Passo:** Definir claramente a população da qual deseja extrair uma amostra.
- 2º Passo:** Definir o tamanho da amostra.
- 3º Passo:** Atribuir um único número a cada elemento da população.
- 4º Passo:** Com a utilização de um gerador de número aleatórios, selecionar os números dos elementos da população que farão parte da amostra.
- 5º Passo:** Coletar os dados da amostra.

Vantagens

Representatividade

Garante que cada elemento da população tenha uma chance conhecida e não nula de ser selecionado.

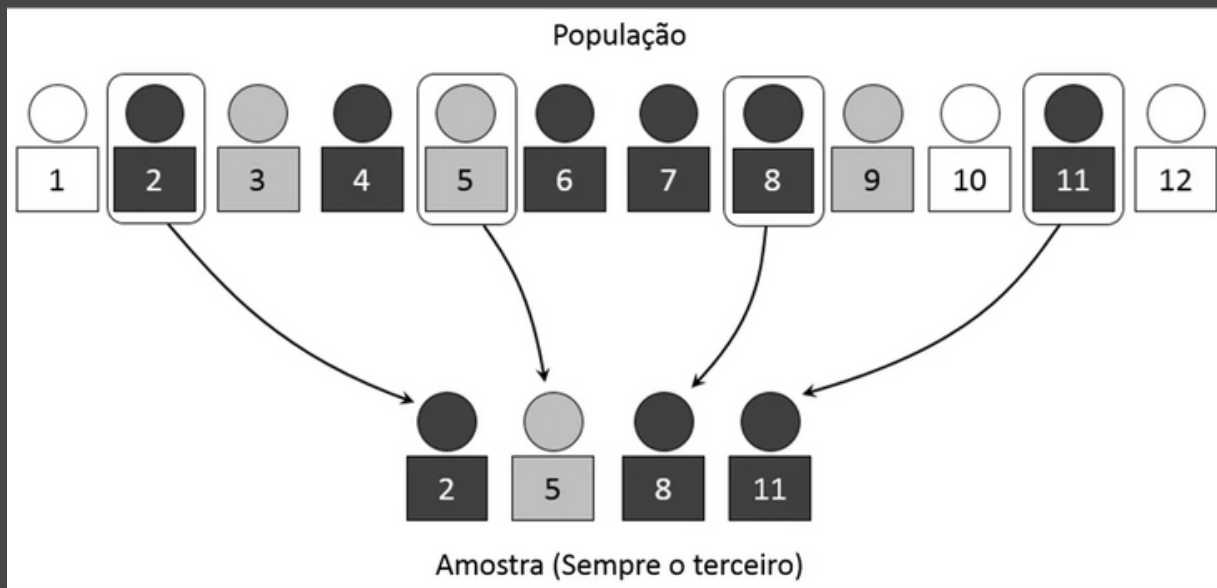
Simplicidade

É fácil de entender e implementar.

Base

Facilita a aplicação de técnicas estatísticas para fazer inferências sobre a população.

Amostragem sistemática



Amostragem sistemática

Definição:

Nesse método é selecionado uma unidade aleatória inicial e, em seguida, coleta dados de cada "n-ésima" unidade subsequente.

Como utilizar:

1º Passo: Definir claramente a população alvo (N).

2º Passo: Definir o tamanho da amostra (n).

3º Passo: Os elementos da população devem ser ordenados de acordo com algum critério, como ordem alfabética, ordem temporal, ordem geográfica, etc.

4º Passo: Calcule o intervalo de amostragem (k), que é o número de elementos entre as seleções sucessivas da amostra, dado pela fórmula $k = N / n$.

5º Passo: Escolher entre os primeiros k elementos, para se tornar o primeiro elemento da amostra.

Vantagens e desvantagem

Eficiência

Pode ser mais eficiente do que a AAS em termos de tempo e recursos.

Simplicidade

Fácil de implementar, especialmente quando a população está ordenada.

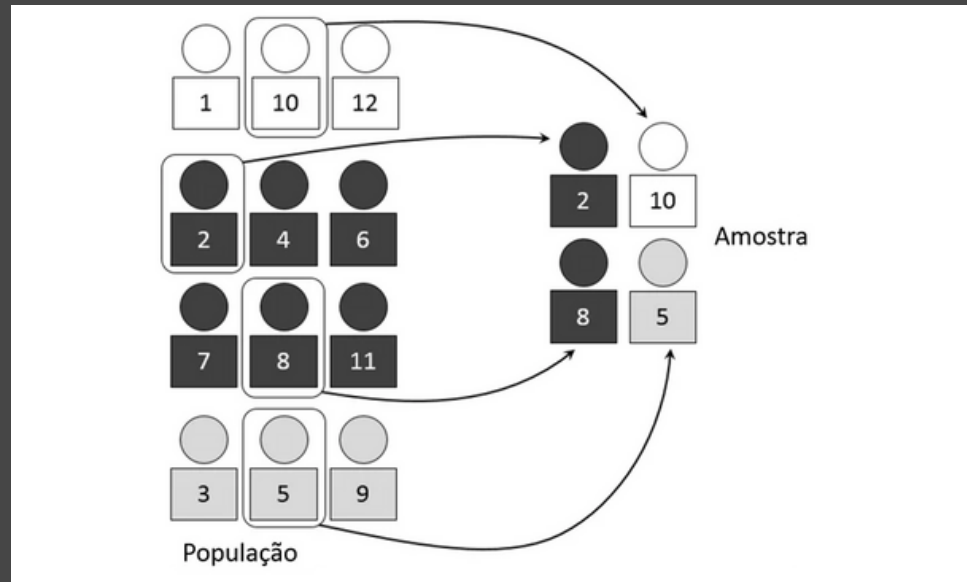
Pouca variação

É capaz de reduzir a variação dos resultados em comparação com a AAS.

Cíclica

Não recomendada para fenômenos com comportamentos cíclicos, pois mantem o mesmo intervalo de análise dos dados

Amostragem estratificada



Amostragem estratificada

Definição:

A amostragem estratificada é um método de seleção de amostras que envolve a divisão da população em subgrupos distintos chamados estratos. É utilizado quando a população pode ser naturalmente dividida em grupos homogêneos em relação à característica de interesse.

Como utilizar:

- 1º Passo:** Definir claramente a população alvo (N).
- 2º Passo:** Definir o tamanho da amostra (n).
- 3º Passo:** Determinar os estratos da população (os subgrupos homogêneos com relação à característica de interesse).
- 4º Passo:** Atribuir cada elemento da população a um estrato específico com base na característica de estratificação.
- 5º Passo:** Determinar o tamanho desejado da amostra para cada estrato. Levando em consideração sua proporção ou não.
- 6º Passo:** Selecionar aleatoriamente as amostras independentemente em cada estrato.
- 7º Passo:** Realizar a coleta de dados.

Amostragem estratificada

Agora que conhecemos os passos necessários para sua aplicação, vamos observar em mais detalhes como de fato a utilizar.

Para realizar a amostragem aleatória simples em cada um de seus estratos (6º Passo), deve-se levar em consideração:

$$f = \frac{\text{tamAmostra}}{\text{numPop}} \quad \text{e} \quad nEi = Ei * f$$

Onde:

- f = fração da amostragem;
- $numPop$ = número de elementos da população;
- $tamAmostra$ = tamanho da amostra a ser selecionada;
- Ei = número de elementos do estrato, onde i é o número do estrato correspondente;
- nEi = o número de elementos a serem sorteados em cada estrato;

Vantagens

Representatividade

Cada subgrupo (estrato) é representado na amostra.

Precisão

Resulta em estimativas precisas, pois considera a variação dentro dos estratos e ajusta o tamanho da amostra.

Eficiência

É eficiente quando há grandes variações nas características de interesse entre os estratos.

Desafios e Considerações

Identificação

A identificação precisa dos estratos é crucial, e é importante garantir que cada elemento seja atribuído ao estrato correto.

Custo

Em comparação com a amostragem aleatória simples, a amostragem estratificada pode ser mais cara e complexa.

Detalhamento

É necessário ter informações detalhadas sobre a população para estratificar efetivamente.

Amostragem por áreas



Amostragem por áreas

Definição:

É um método de seleção de amostras que envolve a divisão da área de estudo em unidades geográficas distintas chamadas áreas de amostragem e a seleção aleatória de áreas para coleta de dados. Uma vez que, listas que identificam cada membro da população nem sempre estão disponíveis aos pesquisadores.

Como utilizar:

1º Passo: Definir área geográfica que está sendo estudada..

2º Passo: Definir o tamanho da amostra (n).

3º Passo: Dividir a área de estudo em unidades geográficas menores chamadas áreas de amostragem.

4º Passo: Continuar a divisão até chegarmos em uma única unidade, as casas.

6º Passo: Realizar a seleção aleatória da amostra usando métodos como a AAS.

7º Passo: Realizar a coleta de dados.

Vantagens

Eficiência

É eficiente quando a área de estudo é extensa, como em estudos que abrangem cidades, estados ou países.

Econômica

É mais econômica e rápida do que a realização de um censo.

Facilita logística

Pois as áreas selecionadas podem ser acessadas de maneira mais eficiente.

Desafios e Considerações

Heterogeneidade

Pode haver heterogeneidade dentro das áreas selecionadas, exigindo uma consideração cuidadosa da variabilidade.

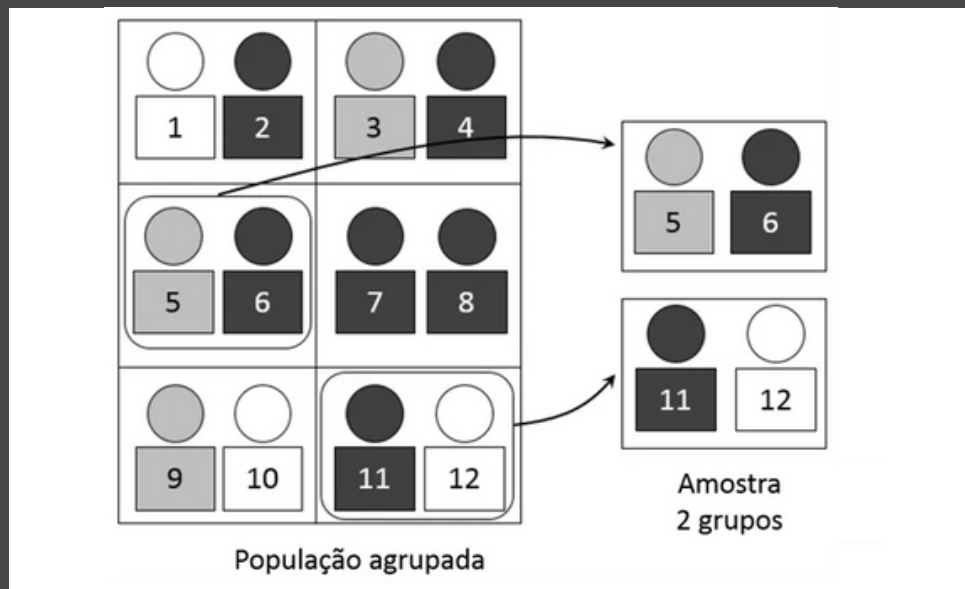
Interpretação

Deve levar em conta a representatividade das áreas selecionadas em relação à área de estudo como um todo.

Possível Viés

Se a distribuição das características de interesse não for homogênea.

Amostragem por conglomerados



Amostragem por conglomerados

Definição:

A população é dividida em grupos com base em sua localização geográfica. Em seguida, uma amostra aleatória de cada grupo é selecionada para a pesquisa. É importante que cada grupo seja o mais heterogêneo possível, selecionando indivíduos diferentes entre si, mas que representem seu grupo. Além disso, os grupos devem representar a população-alvo como um todo.

Como utilizar:

- 1º Passo:** Definir a população.
- 2º Passo:** Divisão da população em conglomerados.
- 3º Passo:** Determine o número desejado de conglomerados que comporão a amostra.
- 4º Passo:** Seleção aleatória de conglomerados (AAS).
- 6º Passo:** Seleção dos elementos dos conglomerados que farão parte da amostra.
- 7º Passo:** Realizar a coleta de dados.

Vantagens

Eficiência

É eficiente quando a população é grande e dispersa geograficamente.

Custo e Tempo

É mais econômica e rápida do que a realização de um censo e outros tipos de amostragem.

Facilita logística

Pois os conglomerados podem ser acessados de maneira mais eficiente do que elementos individuais dispersos.

Desafios e Considerações

Variação

Se houver muita variação entre os conglomerados, pode ser necessário aumentar o número de conglomerados selecionados.

Possível Viés

A seleção aleatória dos conglomerados é crucial para minimizar esse viés. Casos os conglomerados não sejam representativos.

Possível Ineficiência

Se os conglomerados forem homogêneos, a amostragem por conglomerados pode ser menos eficiente do que outros métodos.

Roteiro da apresentação

Téoria:

Apresentação sobre o tema:

- Termos importantes 1 e 2;

Tipos de amostragem:

- Métodos de amostragem;
- Seleção de amostragem;
- Tamanho da amostra 1, 2 e 3;
- Tipos de amostras;

Prática:

Exemplos de amostragens:

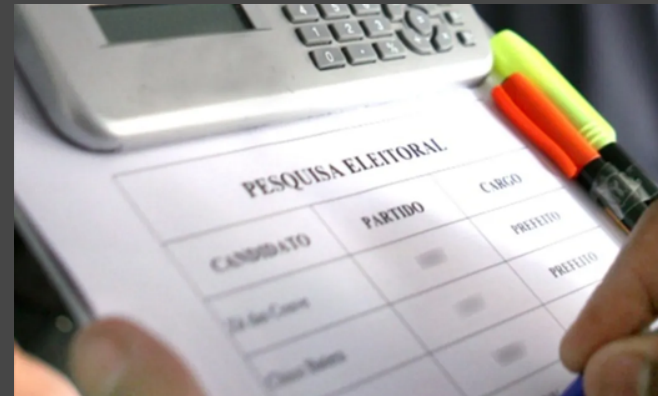
- Amostragem Aleatória Simples;
- Amostragem Sistemática;
- Amostragem Estratificada;
- Amostragem por Conglomerados;
- Amostragem por Área;

Códigos de exemplos

Exemplo de AAS

Exemplo: Estudo de Opinião Eleitoral

Suponha que você deseje realizar uma pesquisa para medir a preferência eleitoral em uma cidade com 100.000 eleitores. Você atribui a cada eleitor um número único de 1 a 100.000 e utiliza um método de seleção aleatória para escolher aqueles que farão parte da amostra. Esta é uma amostragem aleatória simples, onde cada eleitor tem, a priori, a mesma probabilidade de ser escolhido.



Exemplo de AAS

Exemplo: Mega-sena

No caso da Mega-sena, a população é composta por 60 elementos numerados de 01 a 60. O jogador deve escolher entre 6 e 25 números que ele acredita representar uma possível amostra a ser sorteada. O sorteio ocorre de forma que, entre os 60 números, um é sorteado por vez até que o tamanho da amostra desejado (6 números) seja atingido.



Exemplo Amostragem Sistemática

Exemplo: Avaliação da Qualidade do Ar em uma Cidade

Suponha que você deseje avaliar a qualidade do ar em uma cidade com medições de dióxido de nitrogênio. Você organiza as medições em uma ordem sequencial de localização geográfica e, em seguida, realiza uma amostragem sistemática, escolhendo, por exemplo, uma medição a cada 10 pontos para criar uma amostra representativa da cidade.



Exemplo Amostragem Sistemática

Suponha um condomínio com 1000 casas, das quais deseja-se obter uma amostra com 40 casas.

PROCEDIMENTO:

1. $1000/40 = 25$. A cada 25 casas, uma será observada.
2. Faz uma Amostragem Aleatória Simples de 1 a 25. Supondo que foi sorteado o número 4, temos que:
3. Somando 25 e determinando as outras casas, formando assim um sistema previamente estabelecido: As casas que compõe a amostra são: 4 – 29 – 54 – 79
4. Por acaso, alguma casa da amostra estiver fechada, pesquisa-se a próxima e depois retorna para a sequência normal

Exemplo Amostragem Estratificada

Exemplo: Desempenho Acadêmico em uma Escola

Imagine que você está interessado em avaliar o desempenho acadêmico dos alunos em uma escola com diferentes séries. Você divide a população em estratos (por exemplo, alunos do ensino fundamental e alunos do ensino médio) e, em seguida, realiza uma amostragem aleatória simples dentro de cada estrato para obter uma amostra representativa de ambos os grupos.



Exemplo Amostragem Estratificada

Exemplo: Amostragem Estratificada proporcional Numa escola com 100 alunos, 25 deles tem entre 5 e 9 anos, 35 deles tem entre 10 e 13 anos, e 40 alunos tem entre 14 e 17 anos. Extrair uma amostra de 20 alunos desta escola, por meio do método de amostragem estratificada proporcional.

tamPop = 100 alunos;
L = 3 estratos;

ES1 = 25 alunos enumerados de 1 a 25;
ES2 = 35 alunos enumerados de 1 a 35;
ES3 = 40 alunos enumerados de 1 a 40;

tamAmostra = 20 alunos

Dessa forma, para formar a amostra, deve-se sortear 5 alunos entre 5 e 9 anos, 7 alunos entre 10 e 13 anos, e 8 alunos entre 14 e 17 anos.

$$f = \text{tamAmostra} / \text{tamPop} = 20 / 100 = 0,2$$


$$nES1 = \text{Estrato1} \times f = 25 \times 0,2 = 5 \text{ alunos}$$

$$nES2 = \text{Estrato2} \times f = 35 \times 0,2 = 7 \text{ alunos}$$

$$nES3 = \text{Estrato3} \times f = 40 \times 0,2 = 8 \text{ alunos}$$

$$nES1 + nES2 + nES3 = \text{tamAmostra} \quad 5 + 7 + 8 = 20$$

Exemplo Amostragem Estratificada



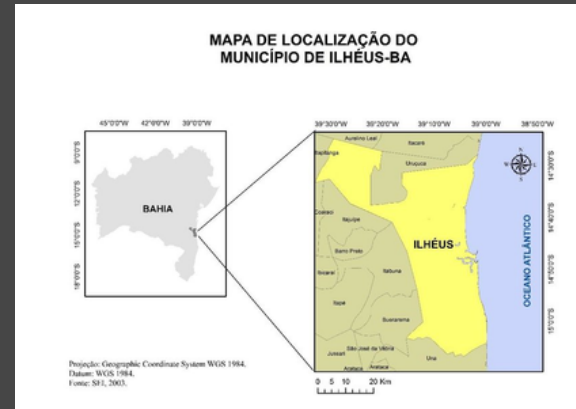
Exemplo: no estudo multicêntrico “Hipertensão Arterial e Diabetes Mellitus entre trabalhadores da saúde: associação com hábitos de vida e estressores ocupacionais” (Novaes Neto, Araújo, and Sousa 2020) (CC BY), a amostra dos participantes no estudo foi realizada da seguinte forma: a partir de listas nominais de todos os trabalhadores em atividade nos serviços de saúde da atenção básica e da média complexidade, fornecidas pelas Secretarias de Saúde dos municípios estudados, estratificou-se a amostra em três níveis: área geográfica, nível de assistência (atenção básica e média complexidade) e grupo ocupacional. A composição da amostra foi definida com base na participação percentual de cada grupo por nível de estratificação estabelecido, seguindo-se o sorteio dos trabalhadores para comporem a amostra (o sorteio foi feito com base em listagem de números aleatórios).

Exemplo Amostragem por Áreas

Exemplo: Grau de alfabetização das pessoas, com idade entre 18 a 45 anos, na cidade de Ilhéus.

PROCEDIMENTO:

1. A cidade de Ilhéus foi dividida em 100 áreas menores.
2. Através da amostragem aleatória simples, foram sorteados 20 dentre as 100 áreas
3. E então das 20 áreas, foram sorteados 30 pessoas alfabetizadas, com idade entre 18 e 45 anos para comporem a amostra.



Exemplo Amostragem por Grupos

Exemplo: Deseja-se descobrir a carreira profissional que alunos das escolas da Bahia desejam seguir. ●●●


PROCEDIMENTO:

1. Elabora-se uma lista de escolas baianas, e realiza o sorteio;
2. Das escolas sorteadas, cria-se uma lista de turmas e realiza o sorteio;
3. A partir das turmas sorteadas, faz-se uma lista de alunos de cada turma e realiza o sorteio, selecionando-os.

Dessa forma, foi realizada uma amostragem por conglomerados em 3 estágios.



Exemplo Amostragem por Grupos



Exemplo: no estudo “Associação da depressão com as características sociodemográficas, qualidade do sono e hábitos de vida em idosos do Nordeste brasileiro: estudo seccional de base populacional” ([Lopes et al. 2015](#)) (CC BY), a população-alvo constou de aproximadamente 40 mil idosos, no ano de 2010, residentes na zona urbana da cidade de Campina Grande-PB. O tamanho da amostra de idosos foi estimada em 205 idosos. Esse estudo realizou uma amostragem por conglomerados, na qual o primeiro estágio consistiu na seleção aleatória de unidades básicas de saúde (UBS) em quatro distritos sanitários urbanos. No segundo estágio, foi realizada uma amostragem sistemática das residências nas ruas das unidades básicas de saúde selecionadas. Por fim, dentro das unidades residenciais selecionadas, todos seus elementos, que faziam parte do público alvo fizeram parte da amostragem.

Referências

- http://leg.ufpr.br/~fernandomayer/aulas/ce001e-2016-2/01_introducao_e_amostragem/01_Introducao_a_Estatistica_e_amostragem.pdf
- https://lec.pro.br//download/faria/apostilas/CET018_10ed_1pf.pdf
- <https://pt.surveymonkey.com/mp/amostragem-por-conglomerado/>
- https://www.lampada.uerj.br/arquivosdb/_book/delineamentos.html#:~:text=5.3.1.2%20Amostragem%20estratificada&text=Por%20exemplo%2C%20a%20popula%C3%A7%C3%A3o%20de,faixa%20et%C3%A1ria%20na%20popula%C3%A7%C3%A3o%20alvo

Referências

- http://leg.ufpr.br/~fernandomayer/aulas/ce001e-2016-2/01_introducao_e_amostragem/01_Introducao_a_Estatistica_e_amostragem.pdf
- <http://www.de.ufpb.br/~luiz/Adm/Aula9.pdf>
- <https://www.scielo.br/j/rbso/a/bMg5nzYYqSBGWCjZrzjhYPb/>
- <https://www.scielo.br/j/rbgg/a/PtkLqMxhrj8b7tQwBVgnr8z/abstract/?lang=pt>



Gratos pela atenção de todos
Tenham um bom dia!

Levi Silva Freitas: lsfreitas.cic@uesc.br
Eduardo Oliveira Freire: eofreire.cic@uesc.br