



**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CAMPUS 1  
CENTRO DE CIÊNCIA E TECNOLOGIA - CCT  
DEPARTAMENTO ESTATÍSTICA  
CURSO DE ESTATÍSTICA**

**CLEVIA BENTO DE OLIVEIRA**

**AMOSTRAGEM ALEATÓRIA SIMPLES**

**CAMPINA GRANDE - PB  
2020**

**CLEVIA BENTO DE OLIVEIRA**

**AMOSTRAGEM ALEATÓRIA SIMPLES**

Trabalho apresentado a Ísis Milane na disciplina de Amostragem I, curso de Estatística da Universidade Estadual da Paraíba, como requisito parcial à obtenção de nota de reposição.

**CAMPINA GRANDE  
2020**

## **SUMÁRIO**

- 1. INTRODUÇÃO**
- 2. ETAPAS PARA SELECIONAR UMA AMOSTRA**
- 3. VANTAGENS**
- 4. DESVANTAGENS**
- 5. CUIDADOS COM A PESQUISA**
- 6. EXEMPLOS**
- 7. CONCLUSÃO**
- 8. BIBLIOGRAFIA**

## INTRODUÇÃO

Amostragem Aleatória Simples, também chamada simples ao acaso, aleatória, casual, simples, elementar, randômica, etc., é equivalente a um sorteio lotérico. Nela, todos os elementos da população têm igual probabilidade de pertencer à amostra, e todas as possíveis amostras têm também igual probabilidade de ocorrer.

Sendo  $N$  o número de elementos da população e  $n$  o número de elementos da amostra, cada elemento da população tem probabilidade  $n/N$  de pertencer à amostra. A essa relação  $n/N$  denomina-se fração de amostragem. Por outro lado, sendo a amostragem feita sem reposição, a suposição que fazemos em geral, é que existem  $(n)$  possíveis amostras, todas igualmente prováveis.

Na prática, a amostragem simples ao acaso pode ser realizada numerando-se a população de  $1$  a  $N$ , sorteando-se, a seguir, por meio de um dispositivo aleatório qualquer,  $n$  números dessa seqüência, os quais corresponderão aos elementos sorteados para a amostra.

Um instrumento útil para realizar o sorteio acima descrito é a tabela de números ao acaso. Essa tabela é simplesmente constituída por inúmeros dígitos que foram obtidos por algum processo equivalente a um sorteio.

## ETAPAS PARA SELECIONAR UMA AMOSTRA

1. Definir a população alvo.
2. Identificar um quadro de amostragem atual da população-alvo ou desenvolva um novo.
3. Avaliar o quadro de amostragem para a falta de cobertura, cobertura excessiva, cobertura múltipla e agrupação.
4. Determinar um número único para cada elemento da trama.
5. Criar o tamanho da amostra.
6. Selecionar aleatoriamente o número específico de elementos da população.

Para selecionar o número de elementos da população é possível recorrer ao método de loteria. Isto é, uma tabela de números aleatórios é gerados aleatoriamente através de um programa de computador. Como resultado, com o sistema de sorteio, os números que representam cada um dos elementos da população alvo são colocados em chips. Após isso, os chips são colocados em um recipiente e misturados.

Em seguida, selecionam as fichas a partir do recipiente até que se tenha obtido o tamanho da amostra desejada. As desvantagens desse método de seleção é que consome muito tempo e se limitada a pequenas populações.

Em uma tabela de números aleatórios não se segue um padrão particular. Eles podem ser lidos de qualquer maneira, ou seja, horizontal, vertical, diagonal, para frente ou para trás.

O número de dígito utilizado deve corresponder ao tamanho total da população-alvo e o número que o investigador encontra que não combina com o atribuído ao elemento da população-alvo é ignorado.

Também é possível usar um software estatístico ou planilha de cálculos para gerar números aleatórios, e os elementos de populações cujos números correspondem aos números atribuídos gerados pelo software são incluídos na amostra.

A chave para uma seleção de amostra adequada é obter e manter uma lista atualizada de todos os indivíduos ou itens a partir dos quais a amostra será extraída. Tal lista é conhecida como a estrutura de população (universo) ou simplesmente população.

A lista da população irá servir como população-alvo, de modo que se muitas amostras probabilísticas diferentes forem retiradas desta lista, é de se esperar que cada amostra seja a representação em miniatura da população e que ela produza estimativas razoáveis de suas características.

Se a lista for inadequada porque certos grupos de indivíduos ou itens na população não foram incluídos apropriadamente, as amostras de probabilidade aleatórias irão apenas oferecer estimativas das características da população-alvo (e não da verdadeira população), os resultados sofrerão tendenciosidade.

Dois métodos básicos poderiam ser utilizados para selecionar a amostra: a amostra poderia ser obtida **com reposição** ou **sem reposição** da população finita. O método aplicado deve ser claramente estabelecido, uma vez que várias formas utilizadas posteriormente com o propósito de inferência estatística são dependentes do método de seleção.

Quando se realiza amostragem com reposição, a chance de que um membro em particular da população seja selecionado na primeira retirada da cesta é de  $1/N$ .

Independentemente de quem seja selecionado na primeira retirada, informações pertinentes são registradas num arquivo principal, e então este cartão específico é recolocado na cesta (amostragem com reposição).

Os  $N$  cartões na cesta são então bem embaralhados e o segundo cartão é retirado. Uma vez recolocado o primeiro cartão, a chance de seleção de cada membro em particular na segunda retirada incluindo aquele primeiro indivíduo, independentemente de ele ter sido previamente selecionado é ainda  $1/N$ .

Novamente, as informações pertinentes são registradas num arquivo principal e o cartão é recolocado para se preparar para a terceira retirada. Tal processo é repetido até que  $n$ , o tamanho desejado da amostra, seja alcançado. Assim, ao realizar amostragem com reposição, cada indivíduo ou item, em cada retirada, terá sempre a mesma chance  $1/N$  de ser selecionado.

No entanto, será que desejaríamos ter o mesmo indivíduo ou item possivelmente selecionado mais de uma vez? Ao compor amostras com seres humanos, é em geral mais apropriado ter uma amostra com pessoas diferentes do que permitir medições repetidas da mesma pessoa. Assim sendo, empregariamos o método de amostragem sem reposição, de modo que, uma vez retirado determinado indivíduo, o mesmo não poderia ser selecionado novamente.

Como antes, ao fazer amostras sem reposição a chance de que qualquer membro da população seja selecionado na primeira retirada da cesta é de  $1/N$ . Quem quer que seja selecionado, as informações pertinentes são registradas em um arquivo principal, e então o cartão específico é colocado de lado, em vez de recolocado na cesta (amostragem sem reposição). Os  $N-1$  cartões remanescentes na cesta são então bem embaralhados e o segundo cartão é retirado. A chance de qualquer indivíduo não selecionado previamente ser escolhido na segunda retirada é agora de  $1$  sobre  $N-1$ .

Este processo de selecionar um cartão, registrar a informação em um arquivo principal, embaralhar os cartões remanescentes e então retirar novamente continua até que seja obtida a amostra desejada de tamanho  $n$ .

## **VANTAGENS**

A vantagem está na possibilidade de produzir amostra representativa e permitir o uso de estatística inferencial nos dados. Sabemos que os procedimentos estatísticos necessários para analisar os dados e calcular os erros são mais fáceis do que os exigidos em outros procedimentos de amostragem probabilística.

Na amostragem sistemática, as chances de ser selecionado não são independentes entre si. Em outras palavras é em geral mais fácil do que em outros métodos de amostragem probabilística (tais como amostragem por aglomerados) para compreender e comunicar com os outros.

Comparado com outros métodos de amostragem probabilística, cada combinação possível de amostragem tem igual probabilidade de ser selecionado, é mais fácil de compreender e comunicar a outros, tende a produzir amostras representativas e os procedimentos estatísticos necessários para analisar erros de dados e de software de estatística são mais fáceis.

## **DESVANTAGENS**

A desvantagem está na necessidade de ter um quadro com elementos de uma população, esse quadro deve ser apropriado. Caso não exista, criar um à partir do zero pode ser uma tarefa difícil e nada viável. Com isso pode-se afirmar que a amostragem aleatória simples é uma boa escolha para os estudos que requerem uma análise comparativa de pequenas categorias de uma população.

Porém a amostragem aleatória simples pode não produzir um número suficiente de elementos de pequenos subgrupos.

Comparado com outros métodos de amostragem probabilística, as desvantagens são: Não tirar proveito do conhecimento que o pesquisador poderia ter sobre a população. Ter erros maiores e com menos precisão do que outros projetos de amostragem com o mesmo tamanho da amostra. Subgrupos da população tem interesses particulares e não podem ser incluídos com um número suficiente na amostra. Com a população dispersa, em que os custos por coleta de dados podem ser mais altos do que de outros modelos da amostra



## **CUIDADOS COM A PESQUISA**

1. Sub abrangência: Quando alguns grupos da população são deixados fora do processo de escolha da amostra

2. Não resposta: não há meios de se obter a observação da característica de interesse. Quando o indivíduo escolhido para

integrar a amostra não pode ser contatado ou se recusa a cooperar

3. Questões mal formuladas: questões confusas ou dirigidas podem introduzir viés (tendência) aos resultados, e mesmo,

pequenas modificações no fraseado podem mudar o resultado da pesquisa

4. Erro amostral: é a diferença entre o resultado amostral e o verdadeiro resultado populacional; tais erros resultam de

flutuações amostrais aleatórias. Os erros amostrais ocorrem única e exclusivamente em função do número de elementos da

amostra e do processo de seleção desses elementos.

5. Erros não amostrais: são os cometidos durante o processo de pesquisa de marketing que não sejam oriundos do tamanho e

do processo de seleção da amostra. Não são mensuráveis e tendem a crescer com o aumento do tamanho da amostra.

Ex. 1: escolha de amostra não aleatória e tendenciosidade

Ex. 2: utilização de instrumentos de mensuração defeituoso

Ex. 3: questão formulada de modo tendencioso

Ex. 4: não respostas

Ex. 5: definição errada do problema de pesquisa

Ex. 6: entrevistadores com nível inadequado para a pesquisa

Ex. 7: definição errada da população de pesquisa

## EXEMPLOS

### Amostragem por sorteio

Neste método, o que se tem que fazer primeiro é elaborar uma lista dos elementos da população, numerados de acordo com a quantidade de elementos, para então serem sorteados. Todo o número tem a mesma probabilidade de ser sorteado e não há repetição.

#### Exemplo:

Uma cidade turística tem 30 hotéis de três estrelas. Pretende-se conhecer o custo médio da diária para apartamento de casal. Os valores populacionais consistem nos seguintes preços diários (em dólares):

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.

Extraia uma amostra aleatória simples de tamanho 10 desta população por sorteio.

R: Escrevemos os valores em papéis, então os colocamos em uma urna, misturamos e sortecemos a amostra de  $n=10$

Resultado obtido:  $n = (20, 24, 22, 28, 23, 24, 21, 20, 25, 27)$  Esse processo não é muito prático para grandes populações, nesse caso é preferível utilizar uma tabela de números aleatórios.

#### Pelo uso da tabela de números aleatórios:

Utilizando o exemplo anterior. 1º passo:

Elaborar a relação dos dados brutos da população, ordenando os números com uma numeração aleatória.

Como dispomos de um conjunto de elementos de 30 números começaremos pelo 00 até o 29, usando dois dígitos, caso tivéssemos 1000 elementos, iniciaremos pelo 000 até o 999, e assim sucessivamente, usando então três dígitos.

N.º Hotel Custo	N.º Hotel Custo	N.º Hotel Custo	N.º Hotel Custo	N.º Hotel Custo
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

2º Passo: Agora iremos sortear o valor de n, aqui num tamanho igual a 10, utilizando a tabela de números aleatórios.

Utilizaremos a tabela agrupando 2 em 2 números pois nossa amostra é de dois dígitos, começando de qualquer ponto na vertical ou na horizontal, até conseguirmos sortear o tamanho de n existente.

09 - \$24	13 - \$19	06 - \$25	56 -
11 - \$25	67 -	21 - Repet.	43 -
51 -	20 - \$24	19 - \$24	37 -
86 -	21 - \$22	46 -	86 -
35 -	13 - Repet.	93 -	32 -
12 - \$20	33 -	80 -	70 -
25 - \$25	62 -	89 -	96 -
37 -	61 -	37 -	61 -
59 -	60 -	62 -	18 - \$28

3º Passo: Acima estão os números sorteados, os que não têm na amostra são descartados, e no nosso caso como não utilizaremos as repetições, pois queremos um sorteio sem reposição, então também serão descartadas as repetições. Nossa Amostra então será:

(24, 25, 20, 25, 19, 24, 22, 25, 24, 28)

Importante:

Nosso espaço amostral era de  $n(A) = 100$  números de 2 dígitos (00 a 99).

A probabilidade de sortear um hotel, seguindo a tabela de números seria de  $30/100$  ou  $0,3$ . Tivemos que percorrer um espaço amostral, na tabela com  $n(A) = 36$  (n.ºs de 2 dígitos) para que pudéssemos encontrar um conjunto evento com.  $N(E) = 10$  hotéis + 2 repets. = 12, então a probabilidade foi de  $12/36 = 0,33$ .

Como  $0,30$  e  $0,33$  não estão muito distantes, podemos afirmar que os números da tabela de números aleatórios, usados, são EQUIPROVÁVEIS.

## **CONCLUSÃO**

Concluimos que, uma amostra aleatória simples é usada para representar toda a população de dados. A facilidade de uso representa a maior vantagem da amostragem aleatória simples. Ao contrário dos métodos de amostragem mais complicados, como amostragem aleatória estratificada e amostragem probabilística, não há necessidade de dividir a população em subpopulações ou tomar outras medidas adicionais antes de selecionar aleatoriamente membros da população.

## **BIBLIOGRAFIA**

### **Livros:**

MOORE, D. S. Estatística básica e sua prática. Rio de Janeiro: Livros Técnicos e Científicos, 3. Ed., 2004.

BUSSAB, W.; BOLFARINE, H. Elementos de amostragem, São Paulo: Edgar Blucher, 2005.

### **Sites:**

<http://www.somatematica.com.br/estat/basica/pagina3.php>

[http://www.triangulomarketing.com.br/Deficao\\_AAS.htm](http://www.triangulomarketing.com.br/Deficao_AAS.htm)