

Estatística Básica

Amostragem

Professora Ma. Tainara Volan
tainaravolan@gmail.com

Conceitos básicos

População

Ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos População estatística ou Universo estatístico.

Amostra

É o subconjunto finito de uma população.

Amostragem

É uma técnica especial para recolher amostras, que garante, tanto quanto possível, o acaso na escolha.

Exemplos

Pesquisa epidemiológica

A população pode ser definida como todas as pessoas da região em estudo, no momento da pesquisa. O principal parâmetro a ser avaliado deve ser a porcentagem de pessoas contaminadas.

Pesquisa eleitoral

A três dias de uma eleição municipal, a população pode ser definida como todos os eleitores com domicílio eleitoral do município. Os parâmetros podem ser % de votos de cada candidato à prefeitura.

Representatividade da amostra

Para que a amostra (parte) possa representar a população (todo), esta precisa ser representativa, isto é, **deve possuir as mesmas características básicas da população** em relação ao fenômeno que queremos estudar.

Aleatoriedade da amostra

- Uma amostra aleatória é um subconjunto de uma população selecionada por um processo que torna igualmente provável a ocorrência de todas as amostras de um tamanho especificado.
- As amostras devem ser aleatórias para eliminar o vício da seleção. O vício da seleção significa que alguns sujeitos apresentam maior probabilidade de estar na amostra do que outros sujeitos (**todos os elementos devem ter a mesma chance**).

Tamanho de amostra via erro amostral

- **ERRO AMOSTRAL** – é a diferença entre uma estatística e o parâmetro que se quer estimar.
- Para a determinação do tamanho de uma amostra, o pesquisador precisa especificar o erro amostral tolerável, ou seja, o quanto se admite errar.

Ex: pesquisa eleitoral. 2% de erro.

Candidato com 20% (18-22%)

Tamanho de amostra

n = nº de elementos da amostra

N = nº de elementos da população

E = erro amostral

$$n_o = \frac{1}{E_o^2}$$

$n = n_o$ se N é muito grande ou desconhecido

$n = \frac{N \cdot n_o}{N + n_o}$ se N não for muito grande e for conhecido

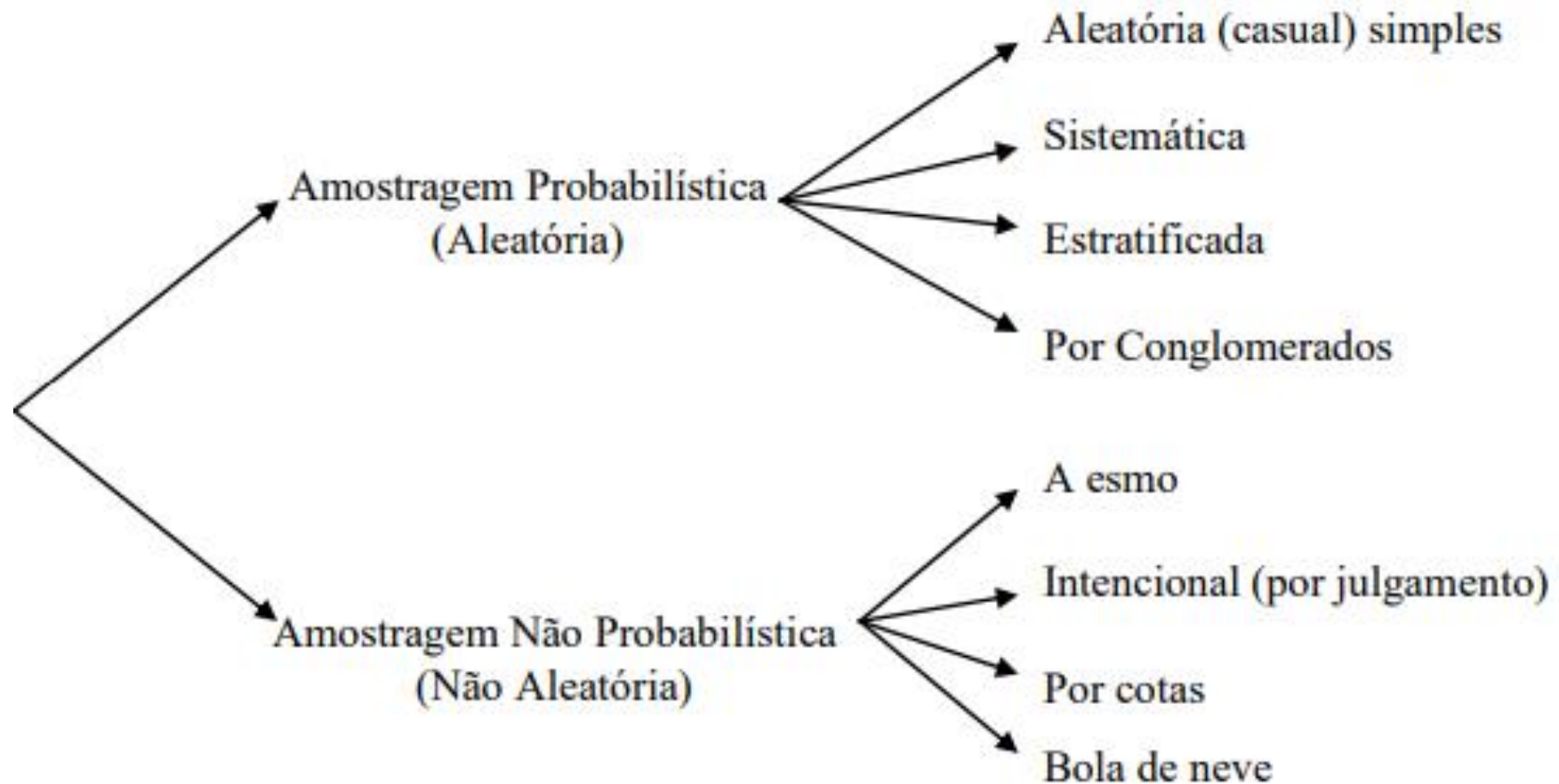
Cálculo do Tamanho de amostra - exemplo

Planeja-se um levantamento por amostragem para avaliar diversas características da população das $N=200$ famílias moradoras de um certo bairro. Qual deve ser o tamanho mínimo da amostra aleatória simples, tal que possamos admitir, com alta confiança, que os erros amostrais não ultrapassem 4% ($E_o = 0,04$)?

$$n_o = \frac{1}{E_o^2}, \text{ então } n_o = \frac{1}{(0,04)^2} = 625$$

$$n = \frac{N \cdot n_o}{N + n_o}, \text{ então } n = \frac{200 \cdot 625}{200 + 625} = \frac{125000}{825} = 152 \text{ famílias}$$

Plano de amostragem



Amostras não probabilísticas

Quando não é possível designar uma probabilidade a cada elemento, dizemos que a amostragem é não probabilística.

- Este processo de amostragem é subjetivo e depende do conhecimento que o pesquisador tem a respeito da população que está estudando.

Exemplos de Amostras Não Probabilística:

- amostragem restrita aos elementos que se tem acesso (ex: drogados)
- escolha a esmo (ex: coelhos numa gaiola)
- impossibilidade de sorteio (ex: sangue)
- amostragem intencional, sem sorteio (ex: escolha de elementos “típicos”)
- voluntários (ex: testes de vacina)

Amostras probabilísticas

Cada elemento da população tem uma probabilidade (não nula) de ser escolhido (segue um método).

Ser probabilística refere-se a procedimentos que utilizam alguma forma de seleção aleatória dos seus membros. Para isto há que se ter garantia de que as diferentes unidades da população têm probabilidades iguais de serem escolhidas.

Amostra Casual ou Aleatória Simples

Cada elemento da população é selecionado de maneira aleatória.

Exemplo:

Pesquisa em uma empresa com 5000 empregados, deseja-se selecionar uma amostra de 100 pessoas

Amostra Casual ou Aleatória Simples

É equivalente a um sorteio, é utilizada quando a população encontra-se desordenada, sendo que por essa técnica, qualquer elemento tem a mesma chance de ser sorteado.

- Quando a população é relativamente pequena (até 30, por exemplo), pode-se numerar os elementos e em seguida, realizar o sorteio.

Casual ou aleatória simples

- Tratando-se de uma quantidade grande, o processo de numeração torna-se trabalhoso. Para tanto, utiliza-se a **TABELA DE NÚMEROS ALEÁTORIOS (TNA)**.
- A leitura dessa tabela é feita, após escolhido o ponto de início, da esquerda para a direita e vice-versa, de cima para baixo e vice versa, na diagonal, etc. A opção deve ser feita antes de iniciado o processo.

Tabela de números aleatórios

57 72 00 39 84	84 41 79 67 71	40 21 13 97 56	49 86 54 08 93	29 68 74 54 83
28 80 53 51 59	09 93 98 87 58	70 27 71 77 17	06 32 02 78 62	16 74 69 65 17
92 59 18 52 87	30 48 86 97 48	35 25 18 88 74	03 62 98 38 58	65 86 42 41 03
90 38 12 91 74	30 19 75 89 07	50 64 15 59 71	88 13 74 95 30	52 78 30 11 75
80 91 16 94 67	58 60 82 06 66	90 47 56 18 46	45 11 12 35 32	45 50 41 13 43
22 01 70 31 32	96 91 92 75 40	16 54 29 72 74	99 00 95 97 61	00 98 24 30 07
56 24 10 04 30	20 46 29 90 53	53 11 05 84 41	21 64 79 19 76	29 51 62 60 66
79 44 92 62 02	96 86 64 30 00	94 56 69 30 20	59 87 87 35 44	22 50 97 78 19
53 99 66 45 08	89 78 50 77 53	37 25 77 41 27	62 38 02 23 57	62 01 41 60 35
18 92 87 35 88	56 05 21 36 51	39 28 50 14 66	85 79 30 19 79	72 66 64 31 45
53 08 58 96 63	05 61 25 70 22	50 41 28 96 62	66 43 63 06 63	01 32 79 85 22
03 58 80 29 28	76 89 51 18 24	88 89 46 47 48	59 19 29 87 03	10 33 99 67 12
27 07 81 88 65	69 49 98 00 28	04 70 51 30 01	47 18 97 33 21	85 82 45 43 24
05 21 08 59 01	06 22 24 98 91	81 17 55 44 66	16 07 73 07 66	10 12 31 78 58
40 36 13 27 84	30 82 33 36 39	69 42 05 58 64	61 12 33 89 27	89 52 66 71 93
54 60 25 28 85	88 20 00 10 59	61 05 36 61 33	72 01 01 19 01	61 10 51 20 91
71 51 63 40 76	71 11 73 73 52	37 31 60 45 88	92 73 43 71 28	04 98 09 02 48
61 02 01 81 73	92 60 66 73 58	53 34 42 68 26	38 34 03 27 44	96 04 46 65 93
82 55 93 13 46	30 95 26 55 06	96 17 65 91 72	39 79 96 12 49	52 80 63 26 99
89 98 54 14 21	74 13 57 68 19	86 28 60 89 47	33 15 26 28 77	45 38 48 08 08
00 99 84 84 14	67 95 13 77 58	90 14 50 79 42	73 63 31 06 60	43 40 12 55 04
62 41 50 78 20	48 05 88 43 52	98 03 19 93 92	03 04 97 25 84	95 95 03 63 31
94 27 90 69 24	68 09 92 11 86	07 63 83 19 32	99 51 15 55 71	09 27 02 67 00
44 89 29 28 84	36 28 25 15 82	87 74 18 97 25	76 10 63 26 76	02 26 74 53 28
97 30 76 95 33	21 10 54 26 95	66 65 52 04 99	36 58 48 03 08	93 63 58 17 96
39 16 58 04 44	80 15 59 59 83	90 95 54 66 81	84 39 60 85 38	88 66 33 35 69
60 78 11 03 26	67 50 34 09 61	31 30 20 76 93	66 30 83 51 09	33 83 64 76 05
03 19 23 47 62	89 57 77 91 33	88 47 60 59 37	54 39 48 77 67	49 85 38 43 91
41 28 52 67 56	25 39 59 96 65	51 36 90 32 22	39 33 05 22 99	03 39 97 96 99
77 54 98 50 39	25 37 42 52 97	10 03 56 04 92	81 66 86 70 01	48 89 55 82 10
28 63 41 61 91	64 24 83 81 37	34 48 83 27 96	38 71 69 73 06	77 50 25 64 60
74 24 48 85 40	12 33 59 67 50	14 98 14 26 42	79 79 13 52 89	69 78 80 44 71
00 24 03 37 96	46 68 75 05 32	42 16 63 33 28	97 26 36 47 27	73 65 38 34 46
05 41 47 69 69	45 36 16 71 18	95 51 97 22 04	13 23 96 58 60	03 69 48 79 83
62 69 84 97 97	47 23 66 51 56	13 08 69 11 52	75 59 26 86 81	80 43 00 98 92

Exemplo

Suponhamos que uma amostra deverá ter 12 elementos de uma população total de 90 indivíduos, e que se tenha escolhido começar na primeira linha da Tabela de Números Aleatórios (TNA), partindo da esquerda para a direita. O primeiro número escolhido seria 57, o segundo 72, e a sequência seria:

57 - 72 - 00 - 39 - 84 - 84 - 41 - 79 - 67 - 71 - 40 - 21 - 13 - 97 - 56...

Destes números sorteados seriam utilizados os 12 primeiros:

57 - 72 - 39 - 84 - 41 - 79 - 67 - 71 - 40 - 21 - 13 - 56

Se o procedimento escolhido fosse da direita para esquerda, os elementos utilizados:

83 - 54 - 74 - 68 - 29 - 93 - 08 - 54 - 86 - 49 - 56 - 97 - 13 - 21 - 40

Eliminando-se os repetidos e fora do limite:

83 - 54 - 74 - 68 - 29 - 08 - 86 - 49 - 56 - 13 - 21 - 40

Amostragem Sistemática

Novamente é feito o sorteio, sendo que nessa amostragem os elementos da população já se encontram ordenados e, nesses casos, não é necessário se construir um sistema de referência (TNA).

Exemplos de populações ordenadas: fichas individuais de empregados (alfabética), casas de uma rua (número), notas fiscais (data), etc.

Define-se um intervalo constante para escolherem-se os itens.

Ex.: Suponhamos que uma empresa tenha 720 colaboradores em determinado setor, dentre os quais deseja-se uma amostra formada por 30 empregados

Amostragem Sistemática - exemplo

1. Determinar o intervalo de amostragem

$$\text{Intervalo} = \frac{\text{população}}{\text{amostra}} = \frac{720}{30} = 24 \text{ (arredondar sempre para inteiro)}$$

2. Escolhemos, por sorteio, um número de 01 a 24.

Este número indicará o primeiro elemento da amostra.

3. Se o primeiro número sorteado for o 5, escolhemos os demais colaboradores relacionados com o primeiro elemento da amostra:

$$2^{\circ} \rightarrow 5 + 24 = 29^{\circ}$$

$$3^{\circ} \rightarrow 29 + 24 = 53^{\circ}$$

$$4^{\circ} \rightarrow 53 + 24 = 77^{\circ}$$

Os demais elementos serão escolhidos, periodicamente, em intervalos de 24 em 24.

Amostragem Estratificada proporcional

Utilizada quando a população encontra-se dividida em estratos (ou camadas, faixas, intervalos, etc).

Exemplos de populações divididas em estratos:

- sexo (homem e mulher);
- idade (criança, adolescente, adulto e idoso);
- setores de uma empresa (administração, vendas, tesouraria, serviços gerais, etc).
- cursos de uma faculdade (C. Contábeis, administração, Direito, Enfermagem, etc);
- Faixa salarial (até 1 SM, de 1 a 2 SM, de 2 a 4 SM, acima de 4 salários-mínimos).

Estratificada proporcional - exemplo

Obs.: após a escolha da quantidade de elementos por estrato, será utilizado o sorteio (por ex TNA) para determinar os indivíduos que comporão a amostra.

Ex.: Será realizada uma pesquisa, a partir de uma amostra, de 12 pessoas. Essas pessoas compõem um grupo de 94 que farão parte de uma expedição na Amazônia, sendo: 45 argentinos, 18 bolivianos e 31 colombianos. Determinar a quantidade de pessoas de cada nacionalidade que responderá a pesquisa.

Estratificada proporcional - exemplo

Calcula-se primeiro o percentual da amostra:

$$\% amostra = \frac{amostra}{população}$$

$$Taxa = \frac{12}{94} = 0,127$$

Estrato (nacionalidade)	População	Cálculo proporcional	Valor	Amostra
Argentina	45	45 x 0,127	5,7	6
Bolívia	18	18 x 0,127	2,2	2
Colômbia	31	31 x 0,127	3,9	4
Total	94	-	-	12

Exemplo

Obter o tamanho mínimo de uma amostra aleatória simples, com um erro amostral máximo de 2%, supondo que a população tenha:

- a) 200 elementos
- b) 200 000 elementos

Exemplo

Obter o tamanho mínimo de uma amostra aleatória simples, com um erro amostral máximo de 2%, supondo que a população tenha:

a) 200 elementos

$$n_0 = \frac{1}{E_0^2} = \frac{1}{(0,02)^2} = 2500$$

b) 200 000 elementos

$$n = \frac{N \times n_0}{N + n_0} = \frac{200 \times 2500}{200 + 2500} = 185,185 \cong 186$$

$$n = \frac{N \times n_0}{N + n_0} = \frac{200000 \times 2500}{200000 + 2500} = 2469,136 \cong 2470$$