

Amostragem

Gilberto Pereira Sassi

Universidade Federal Fluminense
Instituto de Matemática e Estatística
Departamento de Estatística

07 de junho de 2016

Conceitos

População	O conjunto de todos os elementos alvo do estudo
Amostra	Uma parte da população
Variável	Uma característica de um elemento da população
Parâmetro	Uma característica da população
Estimativas	Os valores estimados / calculados dos parâmetros a partir da amostra

Exemplos – nomenclatura

Exemplo

Em pesquisa de intenção de voto para as eleições de 2018, a população são todos os 100 milhões eleitores brasileiros, a amostra é o conjunto de pessoas entrevistada no estudo, variável de interesse é a intenção do eleitor e os parâmetros são a porcentagem de votos de cada candidato.

Exemplo

No estudo dos chefes de família no bairro Saco Grande II, a população são todos os chefes de família no bairro Saco Grande II, a amostra são os 120 chefes de família entrevistados no estudo, a variável de interesse é o grau de instrução e os parâmetros são a porcentagem de chefes com Nenhum Grau Completo, a porcentagem de chefes com Primeiro Grau Completo e a porcentagem de chefes com o Segundo Grau Completo.

Amostragem: usar ou não usar

Razões para usar

Economia Mais rápido e barato estudar apenas parte da população

Tempo Leva menor para levantar os dados

Operacionalidade É mais fácil organizar um estudo pequeno

Razões para não usar

População pequena Se a população é pequena, talvez seja necessário amostrar quase toda população

Variável fácil de medir Pode ser que seja fácil medir uma variável compense estudar toda população

Necessidade de alta precisão Pode ser necessário ter alta precisão no estudo. Por exemplo, o IBGE realiza a cada 10 anos o censo demográficos para estudar diversas variáveis da população e as políticas públicas nacionais são projetadas usando essas informações.

Plano de Amostragem

Em um plano de amostragem, precisamos definir:

Unidade de amostragem que podem ser os próprios elementos da população ou outras unidades que sejam fácil de selecionar.

Exemplos:

- 1) No estudo dos chefes de família no bairro Saco Grande II, podemos primeiro selecionar as residência para chegar ao chefe de família.
- 2) Em uma manifestação, os pesquisadores escolhem uma região e conta o número de manifestantes.

Forma de seleção dos elementos da amostra.

Exemplo: Podemos sortear as unidades amostrais.

Tamanho da amostra

O foco desse curso será o estudo de técnicas de amostragem em que as unidades amostrais são selecionadas por sorteio.

Amostragem Aleatória Simples

Para obter uma Amostra Aleatória Simples precisamos ter uma lista completa das unidades de amostragem e selecionamos as unidades de amostragem por meio de um sorteio, sem reposição.

Características da Amostra Aleatória Simples:

- Os elementos da população podem ser sorteados no máximo uma vez
- Qualquer subconjunto da população tem a mesma probabilidade de ser amostra colhida
- Todos elementos da população tem a mesma probabilidade de fazer parte da amostra

Uso de Tabelas de Número Aleatórios

A tabela de número aleatórios consiste de algarismos 0, 1, 2, ..., 9 sorteados aleatoriamente com igual probabilidade e com reposição. Na Tabela 1, exibimos uma tabela de número aleatórios.

2	5	6	3	6	9	7	4	2	5	6	8	7	3	1	1
7	8	5	9	7	9	2	3	8	2	0	8	0	1	4	6
4	5	1	8	5	6	3	1	0	1	3	7	4	2	9	3
8	4	8	6	5	7	2	1	4	7	2	4	2	6	8	6
9	9	2	6	2	0	1	2	7	8	8	7	3	4	8	3
0	4	0	9	1	4	4	4	1	3	4	6	6	7	1	1

Tabela 1: Tabela de Números Aleatórios.

Exemplo 1

Com o objetivo de estudar o perfil socio-econômico de uma certa empresa Z com 15 funcionários vamos extrair uma amostra de tamanho 5. Na tabela Z, listamos todos os funcionários da empresa Z.

Aristóteles	Anastácia	Arnaldo
Bartolomeu	Bernadino	Cardoso
Carlito	Cláudio	Emílio
Ercílio	Ernestino	Endeivaldo
Francisco	Felício	Fabício

Tabela 2: Funcionário da empresa Z.

Exemplo 1

Abaixo, realizamos os três passos para sortear os funcionários da empresa Z.

- 1) Atribuímos um número inteiro a cada funcionário da empresa Z conforme Tabela 3.

1 – Aristóteles	2 – Anastácia	3 – Arnaldo
4 – Bartolomeu	5 – Bernadino	6 – Cardoso
7 – Carlito	8 – Cláudio	9 – Emílio
10 – Ercílio	11 – Ernestino	12 – Endevaldo
13 – Francisco	14 – Felício	15 – Fabrício

Tabela 3: Funcionário da empresa Z.

- 2) Na Tabela 1, de dois em dois pego os primeiros de 01 a 15 andando da esquerda pra direita e de cima pra baixo. Em nosso caso, os números foram: 11, 08, 01, 12, 04.
- 3) Então, os funcionários selecionados foram Ernestino, Cláudio, Aristóteles, Endevaldo e Bartolomeu.

Depois de selecionados os funcionários coletamos as características ou variáveis importantes para o estudo, como, por exemplo:

- Grau de Instrução

Exemplo 2

Extrai uma amostra aleatória simples de 4 letras do alfabeto da língua portuguesa. Na Tabela 4 apresentamos todas as letras do alfabeto.

A	E	I	M	Q	U	Y
B	F	J	N	R	V	Z
C	G	K	O	S	W	
D	H	L	P	T	X	

Tabela 4: Letras do alfabeto.

Exemplo 2

Abaixo, realizamos os três passos para sortear as 4 letras do alfabeto.

- 1) Atribuímos um número a cada letra do alfabeto conforme Tabela 5.

1 – A	5 – E	9 – I	13 – M	17 – Q	21 – U	25 – Y
2 – B	6 – F	10 – J	14 – N	18 – R	22 – V	26 – Z
3 – C	7 – G	11 – K	15 – O	19 – S	23 – W	
4 – D	8 – H	12 – L	16 – P	20 – T	24 – X	

Tabela 5: Letras do alfabeto numerado.

- 2) Na Tabela 1, de dois em dois pego os 4 primeiros de 01 a 26 andando da esquerda pra direita e de cima pra baixo. Em nosso caso, os números foram: 25, 11, 23, 08.
- 3) Então as letras selecionadas foram Y, K, W e H.

Exemplo

Divida aleatoriamente em dois grupos os números inteiros entre 1 e 8.

Para dividir esse números, primeiro criamos um grupo com 4 número aleatórios para o grupo 1 e os números que restarem constituem o grupo 2.

- 1) Selecionamos quatro números na Tabela 1 entre 1 e 8: buscamos os número de um em um da esquerda pra direita de cima para baixo na tabela. Os números selecionados foram: 2, 5, 6, 3. Ou seja, O grupo 1 é composto por 2, 5, 6, 3.
- 2) O grupo 2 é composto pelos números que não estão no grupo 1, isto é, o grupo 2 é composto por 1, 4, 7, 8.

Tamanho da Amostrai

Objetivo: Determinar o tamanho mínimo da amostra aleatória simples.

Erro Amostral diferença entre estimativa e parâmetro

Erro Amostral Tolerável erro amostral máximo, em proporção, aceitável pelo pesquisador.

Notações:

N número de elementos da população

n número mínimo de elementos da amostra

E_0 erro amostral tolerável (em proporção)

Tamanho Amostral:
$$n = \frac{N}{N \cdot E_0^2 + 1}$$

Exemplo

Planeja-se um levantamento para avaliar a porcentagem que moram em casa próprias em população em um bairro com $N = 200$ famílias. Qual deve ser o tamanho mínimo amostral para termos um erro amostral tolerável de 4%?

Desejamos achar n , conhecendo $N = 200$, $E_0 = \frac{4}{100} = 0,04$. O tamanho amostral mínimo é

$$n = \frac{200}{200 \cdot (0,04)^2 + 1} = 151,52.$$

Logo, o tamanho mínimo amostral é $n = 152$.

Note que $\left(\frac{152}{200}\right) \cdot 100 = 76\%$, ou seja, o tamanho mínimo amostral corresponde a 76% dos elementos da população. Este é um caso em que a amostragem pode não ser vantajoso.

Exemplo

Suponha agora que $N = 200000$, então

$$n = \frac{200000}{200000 \cdot (0,04)^2 + 1} = 623,053.$$

Logo, $n = 624$.

Note que $\frac{624}{200000} 100 = 0,312\%$, ou seja, o tamanho amostral corresponde a 0,312% dos elementos da população. Aqui, vemos claramente a vantagem de usar amostragem.

Observações:

- Com o mesmo erro amostral tolerável, usamos diferentes proporções da população.
- O erro amostral tolerável não é a proporção que devemos coletar da população.

Exemplo

Considero a população composta de 200 crianças do sexo masculino (representados por $H1, \dots, H200$) e 100 crianças do sexo feminino (representados por $M1, \dots, M100$).

- a) Qual deve ser o tamanho da amostra para que o erro amostral tolerável seja 0,1?

$$n = \frac{N}{N \cdot E_0^2 + 1} = \frac{300}{300 \cdot (0,1)^2 + 1} = 75$$

- b) Retire uma amostra aleatória simples do tamanho $n = 5$.
Primeiro atribuímos um número a cada elemento da população:

H1	...	H200	M1	...	M100
1	...	200	1	...	100

Sorteamos 5 números usando a tabela de números aleatórios:
256, 178, 013, 147, 242.

A amostra aleatória simples é: $M56, H178, H13, H147, M242$.

Exemplo

Qual o erro amostral tolerável para uma amostra de tamanho $n=10$.

$$\begin{aligned} E_0 &= \sqrt{\frac{N-n}{N \cdot n}} \\ &= \sqrt{\frac{300-10}{300 \cdot 10}} \\ &= \sqrt{\frac{290}{300 \cdot 10}} \\ &\approx 0,31 \end{aligned}$$