



INSTITUTO FEDERAL
Rio Grande do Sul

Probabilidade e Estatística

Estatística descritiva:

- Introdução
- Tipos de variáveis

Prof. Ruana Maíra Schneider
Ruana.Schneider@farroupilha.ifrs.edu.br



Estatística Descritiva

Estatística é a ciência que trata da coleta, organização, análise e interpretação dos dados para a tomada de decisões.

Estatística Descritiva: tem o objetivo de organizar, resumir, analisar e representar observações e dados disponíveis



Estatística inferencial: é o ramo da estatística que envolve o uso de uma amostra para chegar a conclusões sobre uma população. Uma ferramenta básica no estudo da estatística inferencial é a probabilidade. (estudaremos mais adiante)



Definições

Dado: é o resultado de investigação, cálculo ou pesquisa

Variável: é toda característica que pode assumir diversos valores conforme pessoa ou objeto

As respostas de uma pesquisa (dados) ou experimento são a matéria-prima da análise estatística.



Conjuntos de dados

Uma **população** é a coleção de todos os resultados, respostas, medições ou contagens que são de interesse.

Uma **amostra** é um subconjunto ou parte de uma população.



Observações:

Os dados **amostrais** devem ser coletados usando-se um método apropriado.

Quando os dados amostrais são coletados usando-se um método **inapropriado**, eles **não podem ser usados para tirar conclusões** sobre a população.



Um **censo** consiste em obter os dados de interesse de uma população inteira. Mas, a menos que a população seja pequena, é normalmente impraticável obter todos os dados da população.



Tipos de dados estatísticos

Qualitativos: São características de uma população que não pode ser medidas.

Ordinais – Ex: dia da semana, grau de escolaridade, grau de satisfação

Nominais – Ex: cor dos olhos, Presença de um sintoma, gênero



Quantitativos: São características de uma população que pode ser quantificadas.

Discretos – Ex: Número de cirurgias (quantidades, números inteiros)

Contínuos – Ex: Altura, peso, (pode resultar em qualquer valor) (75,2kg 60,5kg...)



Exemplo 1: Você realizou uma pesquisa em seu trabalho para verificar a frequência de uso da internet. Suas perguntas foram:

- a) Você usa a internet durante o seu horário de trabalho ? [1] Sim [2] Não
- b) Qual a intensidade do uso da internet ? [1] Nenhuma [2] Pequena [3] Média [4] Grande
- c) Quantas vezes você usa a internet durante o expediente de trabalho ? _____vezes/dia
- d) Quanto tempo você usa a internet durante o dia ? _____

Classifique cada uma das variáveis estudadas.



Respostas

- A) qualitativa nominal
- B) qualitativa ordinal
- C) quantitativa discreta
- D) quantitativa contínua



Exemplo 2: Numa loja de departamentos foi realizado o levantamento dos clientes cadastrados no último mês, com os dados apresentados na tabela a seguir:

NOME	BAIRRO	GÊNERO O [1]M [2]F	IDADE [anos]	RENDA [R\$]	VALOR compras [R\$]
Ana	Centro	2	42	1.200,00	120,50
João	Vicentina	1	24	2.500,00	250,32
Paulo	São Luiz	1	23	2.100,00	210,08
Eduardo	Vicentina	1	31	1.250,00	60,40
Maria	Vicentina	2	38	980,00	28,50
Carolina	Centro	2	18	1.350,00	48,34
Anita	Centro	2	22	1.820,00	85,00
Evandro	São Luiz	1	25	2.670,00	189,99
Elis	Vicentina	2	20	1.110,00	93,00
Márcia	São Luiz	2	36	950,00	31,25

Classifique cada uma das variáveis:

- ORDEM: _____
- GÊNERO: _____
- NOME: _____
- IDADE: _____
- BAIRRO: _____
- RENDA: _____
- VALOR DA COMPRA: _____



Classifique cada uma das variáveis:

- GÊNERO: qualitativa nominal
- NOME: qualitativa nominal
- IDADE: quantitativa discreta
- BAIRRO: qualitativa nominal
- RENDA: quantitativa contínua
- VALOR DA COMPRA: quantitativa contínua



Coleta de dados e planejamento de experimentos



INSTITUTO FEDERAL
Rio Grande do Sul



1. Identifique a(s) variável(is) de interesse (o foco) e a população do estudo.
2. **Desenvolva um plano** detalhado para a coleta de dados. Se usar uma amostra, certifique-se de que **a amostra é representativa** da população.
3. Colete os dados.
4. Descreva os dados usando técnicas de estatística descritiva.
5. Interprete os dados e tome as decisões sobre a população usando estatística inferencial.
6. Identifique quaisquer erros possíveis.



Tipos de estudo estatístico

Estudo observacional: o pesquisador não influencia as respostas.

Experimento: o pesquisador, aplica um tratamento antes de observar as respostas. O tratamento é aplicado em uma parte da população, chamada de grupo de tratamento, e as respostas são observadas. Outra parte da população pode ser usada como um grupo controle, no qual nenhum tratamento é aplicado.



Técnicas de amostragem

O **tamanho da amostra**, que é o número de indivíduos em um estudo, é outra parte importante do planejamento experimental, bem como a **técnica** escolhida.



Amostragem aleatória simples ou casual: Cada elemento da população recebe um número distinto e então alguns elementos são sorteados.

Todos os elementos têm a mesma chance de serem sorteados.

OBS: formulário do google enviado para uma lista qualquer de e-mail NÃO É ALEATÓRIO
Entrevistar pessoas na praça NÃO É ALEATÓRIO



Uso de computador para gerar números aleatórios:

No Excel: função **=ALEATÓRIO()** - valores entre 0 e 1 ou

Função **=ALEATÓRIOENTRE(m;n)** inteiros entre m e n. (incluindo m e n)

OBS: Nomenclatura também utilizadas RAND, RANDOM, RANDI(apenas numeros inteiros)



Exercício: Como você faria para gerar uma amostra aleatória de 15 indivíduos de uma população de 300?



Amostragem sistemática: utilizada quando a população apresenta uma ordenação segundo algum critério. **Escolhe-se aleatoriamente** o primeiro indivíduo e segue-se o mesmo intervalo para escolha dos próximos.

Exemplos: em uma esteira, separar o 10º item na ordem em que aparece.

Em uma lista, escolher um nome a cada 6.



Amostragem sistemática:

Obs: se houver alguma regularidade ou chance de repetição de padrão na população, essa técnica deve ser evitada



Amostragem estratificada: consiste em dividir a população, supostamente heterogênea, em subgrupos que chamaremos de **estratos**.

Estes subgrupos deverão ser mais **homogêneos** que a população em relação a variável em estudo

Uma amostra é selecionada **aleatoriamente** de cada um dos estratos. O uso de uma amostra estratificada assegura que cada segmento da população está representado.



As variáveis de estratificação mais comuns são: classe social, idade, gênero, profissão, etc.

Obs.: Ao usar uma amostragem estratificada, alguns cuidados devem ser tomados, de modo a assegurar que todos os estratos forneçam **amostras proporcionais** às suas reais porcentagens de ocorrência na população.

Exemplo.: se 40% das pessoas da cidade pertencem ao grupo de renda mais baixa, então a amostra deve ter uma proporção de 40% desse grupo.



Amostragem por conglomerado: escolher subgrupos que ocorrem naturalmente na população. Ex.: bairros, código postal, cidades, turmas diferentes do mesmo curso etc.

Escolhe-se um ou mais (mas não todos) os conglomerados e coleta-se o dados de **todos os indivíduos desse(s) conglomerado.**

Obs.: Se houver alguma característica predominante em algum conglomerado, esse método não é representativo. (ex: bairro com habitantes de renda mais alta ou mais baixa)



Amostragem por julgamento: técnica não aleatória e consiste obter os elementos através de julgamento. Ou seja, o pesquisador seleciona aqueles elementos da população que julga representarem melhor a população.

Esta técnica pode apresentar **tendenciosidade**, dado que as características subjetivas de julgamento podem não ser adequadas para obter boa representatividade,



Obs: Uma amostra tendenciosa é a que não é representativa da população da qual é extraída.

Exemplo: uma amostra consistindo apenas em estudantes universitários entre 18 e 22 anos não seria representativa de toda a população entre 18 e 22 anos do país. **Por que?**



ATENÇÃO: Um tipo de amostragem que frequentemente leva a estudos tendenciosos (portanto, não é recomendada) é a **amostragem por conveniência**.

Uma amostra por conveniência consiste somente em membros da população que são fáceis de contatar. **TOTALMENTE TENDENCIOSA**

Ex: formulário para uma lista qualquer do google, pesquisa na praça, pessoas dispostas a responder, etc



Identificando a técnica

Você está realizando um estudo para determinar a opinião dos estudantes em sua escola sobre um determinado tema. Identifique a técnica de amostragem escolhida em cada caso:

1. Você divide a população de estudantes com relação às graduações e, aleatoriamente, seleciona e questiona alguns estudantes em cada curso de graduação.

2. Você designa um número para cada estudante e gera números aleatórios. Então, você questiona cada estudante cujo número é selecionado aleatoriamente.



3. Você questiona estudantes que estão no refeitório no horário do intervalo

4. Você seleciona uma turma aleatoriamente e questiona cada estudante.

5. Você designa um número para cada estudante e, depois de escolher um número inicial, questiona a cada 25 alunos.



Identificando a técnica

Respostas

1. Estratificada
2. Aleatória simples
3. Conveniência
4. Conglomerado.
5. Amostragem sistemática.



Perguntas teóricas:

1. Qual é a diferença entre aleatória simples e sistemática?
2. Qual é a diferença entre a técnica estratificada e por conglomerados?
3. Qual é a diferença entre um censo e uma amostragem?
4. Como você selecionaria uma amostra que não seja tendenciosa para uma pesquisa eleitoral para o cargo de prefeito de Farroupilha?



Respostas

1. Na aleatória simples são sorteados cada elemento da amostra, na sistemática é sorteado o primeiro elemento e os próximos ficam “automaticamente definidos” pois serão escolhidos a cada N elementos.

Aleatória simples: cada **elemento** tem a mesma chance de ser escolhido

Aleatória sistemática cada **AMOSTRA** tem a mesma chance de ser escolhida.



Respostas

2. Na estratificada são selecionados elementos representantes de cada “estrato” da população, de forma proporcional. Na amostragem por conglomerado, alguns grupos são escolhidos e todos os elementos desses grupo são selecionados.

3. **Censo:** Toda população participa da pesquisa
Amostragem: é escolhida alguma técnica para obter um grupo menor de elementos para a pesquisa que irá representar a população



Respostas

4. A técnica mais aleatória seria a aleatória simples, porém, caso você queira que haja representatividade de toda população segundo alguma característica (faixa etária escolaridade, etc) recomenda-se a amostragem estratificada.

Amostragens não tendenciosas: Aleatória simples,
aleatória sistemática, estratificada,
conglomerados,



Verdadeiro ou falso?

1. Usar amostras sistemáticas garante que elementos de cada grupo dentro de uma população serão amostrados.
2. Um censo é uma contagem de parte de uma população.
3. O método para seleção de uma amostra estratificada é ordenar uma população de alguma maneira e, então, selecionar elementos da população em intervalos regulares.
4. Para selecionar uma amostra por conglomerado, divide-se a população em grupos e, então, selecionam-se todos os elementos em pelo menos um (mas não todos) dos grupos.



Verdadeiro ou falso?

1. Falso
2. Falso
3. Falso
4. Verdadeiro

