Estratégia
CONCURSOS

Aul

Estatística p/ SEDEST-DF (Estatística) Pós-Edital

Professor: Gustavo Menezes Santana, Vitor Menezes

APRESENTAÇÃO PESSOAL

Este curso será ministrado a “quatro mãos”.

Eu sou **Vítor Menezes**, professor de lógica, matemática, matemática financeira, e estatística para concursos. Também sou Auditor Federal de Controle Externo pelo TCU. Fui também aprovado nos concursos do ICMS-SP, ICMS-MG (cheguei a exercer o cargo por 1 ano e meio), Polícia Federal e MPU. Sou engenheiro eletrônico pelo ITA.

E eu sou o **Gustavo Santana**, professor de raciocínio lógico, matemática, matemática financeira e estatística para concursos. Sou mestre em Ciências e Engenharia de Petróleo, pela Unicamp. Graduado em Engenharia Mecânica, pela mesma instituição. Atuei como Engenheiro de Lean Manufacturing na Thyssenkrupp.

Atualmente estamos atuando tanto aqui no Estratégia Concursos como no Tec Concursos, site parceiro do Estratégia. Juntos, já comentamos mais de 12 mil questões de exatas dentro do site do Tec, o que nos dá uma visão bem prática do que realmente cai em prova e de como cai em prova.

APRESENTAÇÃO DO CURSO – ESTATÍSTICA PARA SEDEST-DF

O nosso edital traz o seguinte conteúdo:

“I) Noções sobre amostragem e distribuições amostrais. 2. Estimação pontual e por intervalo. 3. Testes de hipóteses para médias, com variância conhecida e desconhecida. 4. Testes qui-quadrado: aderência, homogeneidade e independência. 5. Testes de comparação de médias. 6. Teste para variância. 7. Regressão e correlação. II) 1. Introdução: estatística descritiva e inferência estatística, tipos de dados, bancos de dados, ordem de grandeza, precisão e arredondamento de dados quantitativos, proporções e porcentagens, taxas e números índices, sugestões para construção e apresentação de gráficos e tabelas. 2. Representação gráfica e tabular da distribuição de dados: tabelas de frequências, gráficos de barras e do tipo "torta", histogramas, função de distribuição empírica. 3. Medidas resumo: medidas de posição, de dispersão, de assimetria e curtose, gráficos do tipo "boxplot". 4. Modelos para distribuições de frequências: gráficos de probabilidade, estatística de Pearson para avaliação de aderência. 5. Associação entre variáveis qualitativas: tabelas de contingência, coeficientes de associação, sensitividade e especificidade, risco relativo, razão de chances. 6. Associação entre uma variável quantitativa e uma variável qualitativa: homogeneidade de distribuições (análise de variância com um fator), gráficos de médias, gráficos de perfis, regressão logística. 7. Associação entre variáveis quantitativas: gráficos de dispersão, correlação linear, concordância, gráficos de perfis, regressão linear simples, suavização. 8. Associação entre três ou mais variáveis: tabelas de contingência de múltiplas entradas, distribuições multivariadas, matrizes

de covariâncias, análise de variância com dois ou mais fatores, regressão linear múltipla. 9. Outros tópicos: elaboração de relatórios técnicos, análise de sobrevivência e tábuas atuariais, modelos estatísticos (parametrização e identificabilidade). III) 1. Noções de modelagem de bancos de dados: projeto conceitual, lógico e físico de bancos de dados. 2. Modelo Entidade Relacionamento. 3. Modelo Relacional. 4. Técnicas para mapeamento do Modelo Entidade Relacionamento para o Modelo Relacional. 5. Linguagens do modelo relacional: álgebra relacional, cálculo relacional e SQL. 6. Linguagem SQL: detalhes dos comandos de definição de dados e de manipulação de dados. 7. Introdução a bancos de dados multidimensionais. 8. Ferramentas de apoio para o desenvolvimento do projeto lógico e físico de bancos de dados. 9. Principais aspectos para especificação e implementação de um banco de dados com seus procedimentos de atualização e consulta para um determinado sistema de bancos de dados para análises estatísticas. IV) 1. Probabilidade subjetiva. 2. O método Bayesiano. 3. Qualidade de uma inferência: métodos clássicos, inferência como um problema de decisão. 4. Inferência conjugada: modelo normal, modelos discretos. 5. Intervalos de credibilidade. 6. Testes de hipóteses. 7. Métodos computacionais. 8. Aplicações: modelos lineares, análise de dados discretos e outras. V) 1. Gráficos multivariados. 2. Análise de agrupamentos. 3. Análise de componentes principais. 4. Análise fatorial. 5. Discriminação e classificação. 6. Biplot. 7. Análise Procrustes 8. Escalonamento multidimensional. 9. Análise de regressão multivariada. 10. Análise de correlação canônica. 11. Análise de correspondência. 12. Modelos de equações estruturais. VI) 1. Álgebra matricial, formas quadráticas. 2. Distribuição multivariada e momentos, distribuição normal multivariada. 3. Teorema limite central multivariado. 4. Distribuição de Wishart e de Hotteling. 5. Coeficiente de correlação múltipla. 6. Coeficiente de correlação parcial. 7. Estimação: vetor de médias e matriz de covariâncias. 8. Distribuição do coeficiente de correlação amostral. 9. Inferência sobre vetores de médias: região de confiança, intervalos de confiança simultâneos e testes de hipóteses. 10. Testes não paramétricos."

O nosso curso não abordará os seguintes temas do edital:

Precisão e arredondamento de dados quantitativos. Sensividade, especificidade, risco relativo e razão de chances. Gráficos de médias, de perfis e regressão logística. Associação entre três ou mais variáveis. Outros tópicos (item II.9 do edital). Itens III a VI do edital.

Teremos as seguintes aulas:

Aulas	Conteúdo
0	Introdução
1	Formas de apresentação de dados
2	Média
3	Quantis e Moda
4	Medidas de dispersão

5	Somatório, proporção amostral e momentos
6	Assimetria, Box Plot e Curtose. Análise bidimensional
7	Números índices
8	Probabilidade
9	Variável discreta. Covariância. Correlação linear entre variáveis aleatórias. Variáveis contínuas
10	Principais distribuições discretas e contínuas
11	Momentos (inferência). Distribuições conjuntas, marginais e condicionais: caso discreto e caso contínuo. Amostragem
12	Estimadores pontuais e distribuições amostrais. Características dos estimadores
13	Intervalos de confiança
14	Testes de hipóteses
15	Análise de variância
16	Regressão simples. Regressão Múltipla

Todas as aulas terão listas de exercícios, com uma mescla das principais bancas do país.

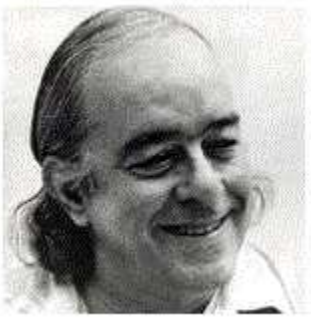
Sumário

1 – Ramos da Estatística 2
1.1 - Amostra 7
2 - Tipos de Variáveis 8
2.1 – Variável quantitativa discreta 11
2.2 – Variável quantitativa contínua 11
2.3 – Variável qualitativa ordinal 12
2.4 – Variável qualitativa nominal 12
2.5 - Tipos de variáveis e cálculo de medidas estatísticas 13
2.6 - Polêmica envolvendo variáveis dicotômicas 15
3 – Escalas de medida 18
3.1 – Escala nominal 19
3.2 – Escala Ordinal 21
3.3 – Escala Intervalar 22
3.4 – Escala de razão 23
4 – Séries estatísticas 24
4.1 – Séries cronológicas ou temporais 24
4.2 – Séries geográficas ou de localização 25
4.3 – Séries específicas 26
4.4 – Séries conjugadas ou mistas 26
5 – Caderno no Tec Concursos 27

1 – RAMOS DA ESTATÍSTICA

A estatística cobrada em concursos é usualmente dividida em duas partes: a descritiva e a inferencial. Para melhor entendimento, vejamos um exemplo.

Nas eleições presidenciais, os candidatos Antônio Carlos Jobim e Vinícius de Moraes vão disputar o segundo turno. Todos os jornais divulgam as fotos e a agenda dos candidatos, que saem em passeatas pelas principais cidades. Eles fazem discursos inflamados, com inúmeras promessas de campanha. Um promete mais claves de sol, o outro, mais rimas cruzadas.





Às vésperas da eleição, o jornal encomenda uma pesquisa de intenção de voto, numa tentativa de antecipar o resultado da eleição.

Neste exemplo, o conjunto de todos os eleitores do país é a nossa **população**.

POPULAÇÃO = TODO

Existem duas maneiras de nos referirmos à população.

Podemos dizer que a população é o conjunto de todos os **eleitores**. Seria formada pelo João, pelo José, pela Maria, etc. Nesse primeiro caso, dizemos que a população é formada por elementos que possuem determinada característica em comum. No caso, a característica é: ter um título de eleitor e poder votar para presidente.

População: {João, Maria, José, Vitor, Paula, Cláudia, Mário, etc}

Podemos também nos referir aos **atributos** que essas pessoas têm. Nesse segundo caso, a população seria formada pelo conjunto das intenções de voto:

{Antônio, Antônio, Vinícius, Antônio, Branco/Nulo, Vinícius, Branco/Nulo, Antônio, Antônio, Branco/Nulo, Vinícius...}

As duas formas de nos referirmos à população estão corretas.

Seria ótimo se o Instituto de Pesquisa pudesse entrevistar todos os eleitores do país. Se ele fizesse isso, teríamos um **censo**. O censo consiste na análise de todos os elementos da população. O grande problema é que o censo geralmente é caro e demorado. Imaginem o tanto de gente que o Instituto de Pesquisa teria que contratar para entrevistar todos os eleitores do país. E o tempo que ia demorar...

Para evitar esse trabalhão todo, o que se faz geralmente é uma **amostragem**. Uma amostra é um subconjunto da população. É um pedaço da população.

AMOSTRA = PARTE

E é isso que o Instituto de Pesquisa faz. Ele escolhe um pedacinho da população (ou seja, alguns eleitores), entrevista, e faz sua pesquisa.

Muito bem. Suponha então que o Instituto fez a tal da pesquisa. Segue um trechinho do resultado:

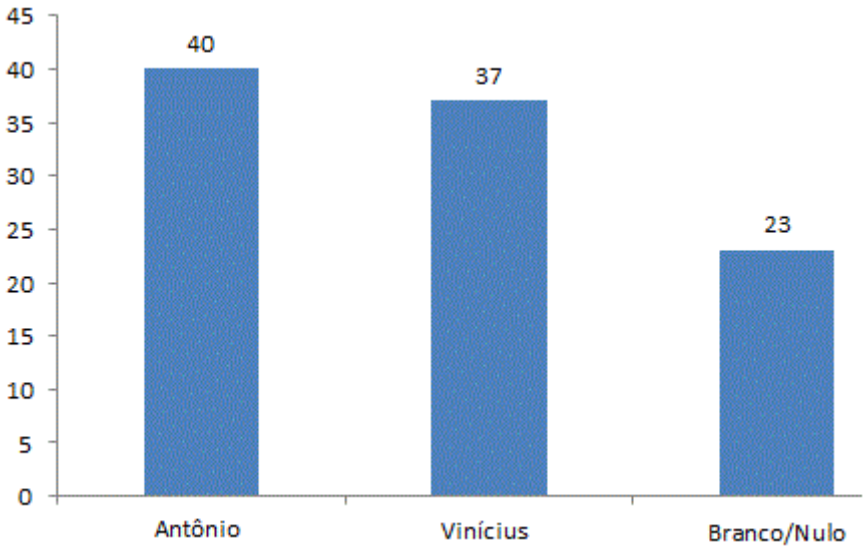
{Antônio, Antônio, Vinícius, Nulo, Antônio, Branco, Vinícius, ...}

Simplesmente pegar essa listagem e entregar para o telejornal nada vai adiantar. O que é que o apresentador vai fazer com isso? Ele vai ler cada intenção de voto? Já pensou no apresentador lendo uma por uma todas as intenções de voto? Não, assim não dá.

Antes de tudo, o Instituto precisa dar um jeito de **apresentar** o resultado de sua pesquisa. Pode-se fazer isso usando gráficos, tabelas, diagramas. Podemos também utilizar medidas que descrevem, resumidamente, o conjunto de dados. Tudo isso é o objeto de estudo da **estatística descritiva**.

Assim, o Instituto poderia dizer que, para a pesquisa feita, o Antônio teve 40% das intenções de voto e o Vinícius teve 37% das intenções de voto. Esses dois percentuais descrevem, resumidamente, todo o resultado da pesquisa.

Poderia também ter sido feito um gráfico:



Agora sim, o apresentador já vai ter o que mostrar no telejornal. Ele poderá apresentar esse gráfico e dizer que Antônio teve 40% das intenções de voto.

Ou seja, a estatística descritiva, como o próprio nome indica, busca descrever um conjunto de dados.

Só que a pesquisa não para por aí. Feita a amostragem, calculadas as proporções de intenções de voto para Antônio e Vinícius, é preciso inferir o que ocorre na população. Nesse momento, utilizamos ferramentas de **inferência estatística**. Buscamos fazer generalizações. Generalizar o resultado da amostra para toda a população.

O Instituto calcula então que, na população, com 95% de confiança, Antônio tem 40% das intenções de voto, com margem de erro de 2% para mais ou para menos. Ou seja, na população ele deve ter de 38% a 42% das intenções de voto. Cálculos como esse são objeto de estudo da estatística inferencial.

Podemos agora ver algumas questões de concurso:

HORA DE PRATICAR!

(Fundação Carlos Chagas) O processo estatístico que consiste em uma avaliação direta de um parâmetro, utilizando-se todos os componentes da população, denomina-se:

a) amostragem

b) estimação

c) censo

d) parametrização

e) correlação

Resolução:

Quando temos acesso a todos os valores da população, estamos realizando um censo.

Gabarito: C.

(Fundação Universa) Julgue os itens a seguir.

I - Uma cidade possui 1.000 habitantes. Um estatístico, necessitando fazer uma determinada pesquisa, entrevistou 200 pessoas. É correto dizer que, nesse exemplo específico, de uma amostra de 1.000 pessoas, o estatístico entrevistou uma população de 200 indivíduos.

II - Um estudante tinha 1 moeda, 1 folha de papel em branco e 1 caneta e, com esse material, resolveu fazer uma experiência. Arremessou uma moeda 20 vezes seguidas. Em cada uma das vezes, ele verificava se a face sorteada era “cara” ou “coroa”. Caso fosse “cara”, ele escrevia o número 1 no papel. Caso fosse “coroa”, ele escrevia o número 2 no mesmo papel. No final da experiência, o estudante obteve 7 “coroas” e somou todos os números existentes no papel. Esse resultado foi atribuído a uma variável X. Com isso, o resultado encontrado para X foi 27.

III - Uma fábrica produz 100.000 lâmpadas por mês. São sorteadas 100 lâmpadas, e essas são mantidas acesas até queimarem, com o objetivo de calcular a vida média desse tipo de lâmpada. A experiência, que utiliza um subconjunto de um grupo para calcular determinado parâmetro e admite que esse parâmetro é válido para todo o grupo, é um problema estudado pela estatística inferencial.

Assinale a alternativa correta.

(A) Nenhum item está certo.

(B) Apenas os itens I e II estão certos.

(C) Apenas os itens I e III estão certos.

(D) Apenas os itens II e III estão certos.

(E) Todos os itens estão certos.

Resolução.

Item I: a questão inverteu a utilização dos termos “população” e “amostra”. A população corresponde ao todo, ao conjunto de todos os elementos que possuem certa característica. No caso, a população é formada pelos 1.000 habitantes. A amostra é qualquer subconjunto da população. No caso, a amostra é formada pelos 200 indivíduos entrevistados. O item está errado.

Item II:

Foram 7 coroas e 13 caras. A soma dos pontos obtidos fica:

$$7 \times 2 + 13 \times 1 = 27$$

Realmente o resultado encontrado é 27. O item está certo.

Item III: Quando generalizamos um resultado obtido em uma amostra para toda a população, utilizamos ferramentas de inferência estatística. O item está certo.

Gabarito: D

(Cespe) Julgue os seguintes itens.

1. Um censo consiste no estudo de todos os indivíduos da população considerada.

2. Como a realização de um censo tipicamente é muito onerosa e(ou) demorada, muitas vezes é conveniente estudar um subconjunto próprio da população, denominado amostra.

Resolução:

1. item está perfeito! Quando analisamos todos os elementos de uma população, temos um censo. Contrariamente, quando analisamos apenas uma parte da população, temos uma amostra. Raramente se faz um censo, por razões de custo (é um procedimento caro) e tempo (é um procedimento demorado).

Item certo

2. Item correto

Para melhor entendimento, considere o censo realizado pelo IBGE. É necessário contratar muitas e muitas pessoas para percorrer o país inteiro entrevistando as pessoas. Ou seja, é algo caro. Além disso, demora para conseguir abarcar todas as famílias do país (é demorado).

Por isso é tão difícil realizar um censo. Um procedimento alternativo é analisar só um pedaço da população, chamado de amostra. A partir da amostra obtida é possível ter ideia do que ocorre na população.

Note que o item traz a expressão "subconjunto próprio". Nós estudaremos este conceito mais abaixo.

Gabarito: certo, certo

1.1 - AMOSTRA

Agora que vimos uma ideia geral dos conceitos iniciais da matéria, podemos detalhar melhor o conceito de "amostra".

Estudamos que a população é o conjunto de **todos** os elementos que possuem determinada característica. Vimos também que a amostra é um pedaço da população.

De forma mais técnica, dizemos que a amostra é um **subconjunto próprio** e **não vazio** da população.

Vamos ver cada uma destas palavras isoladamente.

Subconjunto: Se interpretarmos a população como um conjunto e a amostra como um segundo conjunto, veremos que a amostra está "contida" dentro da população. Por isso é dita um subconjunto.



Próprio: Lá na matemática estudamos que qualquer conjunto A é um subconjunto de si mesmo.

"A" está contido no próprio "A"

Deste modo, se a definição de amostra fosse apenas "*a amostra é um subconjunto da população*", teríamos uma situação meio estranha. Isto porque a população é um subconjunto de si mesma. Com isso, a população mesma seria uma amostra, algo totalmente sem sentido.

Para evitar este problema, usamos a palavra "próprio". Quando dizemos que B é um subconjunto **próprio** de A estamos garantindo que B está contido em A e, além disso, B é diferente de A. Ou seja,

a palavrinha "próprio" garante que B nunca terá todos os elementos que A tem. Ou ainda: B é menor do que A. Ou usando um termo mais técnico: B está estritamente contido em A.

Assim, ao dizermos que a amostra é um subconjunto próprio da população, garantimos que a amostra é menor que a população.

Não vazio: Com este termo garantimos que a amostra sempre apresenta elementos.

Concluindo: A amostra é um subconjunto próprio e não vazio da população. Ou seja, a amostra é um pedaço da população, algo menor que a população. E algo que sempre apresenta elementos (pois é não vazio).

2 - TIPOS DE VARIÁVEIS

Nesse capítulo nosso intuito é estudar os diferentes **tipos** de **variável**.

Vejam bem o que eu falei: vamos estudar tipos.... *tipos do quê?* Tipos de uma coisa chamada de "variável".

Bem, então nada mais lógico que, antes de classificar essa coisa, a gente defina o que ela é, concordam? Não faz muito sentido irmos direto classificando algo que nem definimos ainda.

O problema é que, se eu apresentar a definição de "variável' já de cara no começo do curso, todo mundo arranca os cabelos e desiste da matéria, =)

Por ora, vamos ficar com uma ideia simplista mesmo: pense em variável como algo que varia.

Exemplo: se estivermos estudando a umidade relativa do ar de uma cidade, observaremos que todos os dias ela assume valores diferentes. Um dia pode estar na casa dos 60%. Outro dia pode estar na casa dos 30%. No outro dia pode estar na casa dos 20%. Ela vai variando conforme a estação do ano, ou conforme o horário de medição. Em síntese: a umidade relativa do ar é uma **variável**.

Vista então uma ideia geral do que é essa tal de variável, vamos direto para o que interessa no momento: classificar as variáveis.

O primeiro grande passo é classificar entre: **quantitativa versus qualitativa**

Os próprios nomes já dão uma boa noção de como funciona essa classificação:

- quantitativa lembra quantidade → quantidade lembra número → são as variáveis numéricas
- qualitativa lembra qualidade → qualidade lembra atributo → são variáveis não-numéricas, que indicam atributos

Vamos ver na prática a diferença entre ambas.

Para tanto, vou adaptar um exemplo constante do livro “Estatística Básica”, dos autores Bussab e Morettin.

Uma grande empresa quer conhecer melhor seus funcionários, e aplica um questionário, que contém os seguintes campos:

Formulário - Empresa Alfa - Setor de Recursos Humanos

1) Grau de instrução	( ) fundamental ( ) médio ( ) superior
2) Estado civil	( ) solteiro ( ) casado
3) Número de filhos	
4) Região de procedência	( ) Centro-oeste ( ) Sudeste ( ) Nordeste ( ) Norte ( ) Sul
5) Altura	
6) Sexo	( ) masculino ( ) feminino

Todos esses seis campos correspondem a variáveis.

As possíveis respostas para o campo “3” (número de filhos) são números. A pessoa pode ter 0 filhos. Ou pode ter 1 filho. Ou pode ter 2 filhos. Ou pode ter 3 filhos. Ou 4 filhos. Ou até mais. Tudo isso é número. Uma variável cujas realizações são numéricas é dita **quantitativa.**

As possíveis respostas para o campo "5" (altura) também são números. A pessoa pode ter 1,70 metros, ou 1,84 metros, ou 1,90 metros, etc. Tudo isso é número. Trata-se de outra variável **quantitativa**.



Variável **quantitativa**: a resposta para ela é sempre um número.

Já o campo "2" (estado civil) tem como possíveis respostas: solteiro / casado. Nada disso é número. Portanto, dizemos que se trata de uma variável **qualitativa**, já que suas respostas são atributos.

O campo "4" (região de procedência) tem como possíveis respostas: Centro-Oeste, Sudeste, Nordeste, Sul, Norte. Novamente, as respostas são atributos e não número. Novamente se trata de uma variável **qualitativa**.



Variável qualitativa: a resposta para ela é sempre um atributo, nunca uma quantidade.

Ok, já estudamos então a principal divisão: variáveis quantitativas versus qualitativas.

Agora, dentro de cada uma delas, podemos ver mais classificações.

2.1 – VARIÁVEL QUANTITATIVA DISCRETA

Uma variável quantitativa é **discreta** quando apresenta valores que correspondem a um conjunto enumerável de pontos da reta real.

Hein, como assim? O que é isso professor?

É o seguinte. Quando uma variável é discreta, nós conseguimos enumerar seus valores. Por exemplo, volte ao nosso formulário que será aplicado aos funcionários da empresa e atente para o campo "3" (número de filhos). Os possíveis valores são:

1º valor: zero filhos

2º valor: 1 filho

3º valor: 2 filhos

4º valor: 3 filhos

E assim por diante.

Nós conseguimos listar os possíveis valores. Conseguimos dizer qual é o primeiro valor possível, qual o segundo valor possível etc. Sabendo um dado valor, nós conseguimos determinar o próximo.

2.2 – VARIÁVEL QUANTITATIVA CONTÍNUA

Observem atentamente agora o campo "5" do formulário (altura). Ele corresponde a uma variável **contínua**. Isso ocorre porque nós não conseguimos enumerar seus **possíveis** valores. Dada uma altura, não conseguimos identificar qual a altura que viria a seguir. Isto ocorre porque ela pode assumir qualquer valor num intervalo real.

Para melhor entendimento, considere a altura de 1,70 m. Qual a altura que viria logo após este valor?

Uma pessoa poderia dizer que é 1,71 m.

Aí outra pessoa poderia dizer que é 1,701 m.

Outra pessoa diria que é 1,7001 m.

E assim por diante. Para qualquer número que você pensar, é possível determinar outro que esteja ainda mais próximo de 1,70. O resultado é que simplesmente não é possível determinar quem seria

o próximo número logo após o 1,70. É exatamente por esse motivo que dizemos que a variável altura é contínua.

Para a maioria das aplicações práticas que nos interessam (resolver questões de prova), podemos entender que as variáveis contínuas são aquelas que assumem valores em determinado intervalo da reta real. Ou seja, associamos variável contínua ao conjunto dos números reais (ou a um subconjunto, um intervalo real qualquer).

2.3 – VARIÁVEL QUALITATIVA ORDINAL

O campo "1" do formulário (grau de instrução) corresponde a uma variável **qualitativa**. Suas possíveis realizações não são números. São um atributo, ou uma qualidade. Apesar de suas possíveis realizações não serem numéricas, é possível ordená-las. Dizemos que se trata de uma variável qualitativa **ordinal**.

É possível ordena-las? Como assim?

Nós conseguimos estabelecer uma ordem. Por exemplo, começando do grau de instrução inferior para o superior:

1º: nível fundamental

2º: nível médio

3º: nível superior.

2.4 – VARIÁVEL QUALITATIVA NOMINAL

Existem outras variáveis qualitativas que não podem ser ordenadas. São as variáveis qualitativas **nominais**. Um exemplo é a variável associada aos campos "2" e "4". Você não consegue ordenar as regiões de procedência da forma como fizemos para o grau de escolaridade.

Resumindo, os tipos de variáveis são:

- quantitativas discretas. Exemplo: quantidade de filhos: 0, 1, 2, 3, 4, ...
- quantitativas contínuas. Exemplo: altura: pode assumir qualquer valor entre, digamos, 0,4 metros (recém-nascido) e 2,60 metros (homem mais alto do mundo)
- qualitativas ordinais. Exemplo: grau de instrução: nível fundamental / nível médio / nível superior
- qualitativas nominais. Exemplo: região de procedência



(Fundação João Goulart) Os dados de um determinado estudo representam muitas variáveis para cada uma das pessoas que se submeteram ao estudo. Uma variável considerada qualitativa é a seguinte:

a) idade

b) altura

c) sexo

d) peso

Resolução:

As variáveis "idade", "altura" e "peso" correspondem a quantidades. São todas variáveis quantitativas.

Por exemplo, podemos ter 45 anos de idade, 1,80 metros de altura e 87 kg. Tudo isso é número.

A variável "sexo" não é quantitativa. Ela tem duas realizações, correspondentes a dois diferentes atributos: "masculino" e "feminino". Trata-se de uma variável qualitativa.

Gabarito: C

2.5 - Tipos de variáveis e cálculo de medidas estatísticas

Ao longo de um curso de estatística várias medidas diferentes são estudadas. Nós ainda não vimos nenhuma delas, mas vamos já trazer alguns aspectos, de forma bem resumida, só para entendermos como elas se comportam perante os diversos tipos de variáveis.

Considere cinco pessoas, com as seguintes informações:

	Mônica	Paulo	Maria	Ricardo	Júlia
Idade	14	22	24	25	25
Sexo	feminino	masculino	feminino	masculino	feminino
Escolaridade	fundamental	médio	superior	superior	superior
UF	RJ	RJ	SP	MG	GO

Nós ainda não passamos pela aula de "média aritmética", mas você provavelmente sabe como calcular. Basta somar todos os valores e dividir pelo número de observações.

Portanto, se quisermos calcular a idade **média** desse grupo de pessoas, basta somar as idades e dividir por 5, já que são 5 pessoas:

$$\frac{14 + 22 + 24 + 25 + 25}{5} = 22$$

A idade **média** é de 22 anos.

Veja que, para calcular a média, foi preciso somar os valores. Só podemos somar algo que seja numérico, não dá para somar atributos.

Para deixar mais claro, como você calcularia a média da variável "sexo"? Simplesmente não dá pra fazer, concorda? Ficaríamos com algo do tipo:

$$\frac{\text{feminino} + \text{masculino} + \text{feminino} + \text{masculino} + \text{feminino}}{5}$$

O que não faz sentido algum.

Portanto, só dá para calcular a média para variáveis quantitativas.

Outra medida que vamos estudar é a **moda**. A moda é o termo que mais ocorre.

No nosso exemplo, conseguimos calcular a moda para todas as variáveis. Para começar, nas idades, dizemos que a moda vale 25, pois foi a idade que mais ocorreu. Na variável sexo, a moda é "feminino", pois foi o sexo que mais ocorreu. Na variável "escolaridade", a moda é "superior", pois foi o que mais ocorreu. Na variável "UF", a moda é RJ, pois foi o termo que mais ocorreu.

Notem que dá para determinar a moda para qualquer tipo de variável, já que basta verificarmos qual termo ocorreu mais vezes.

Outra medida que vamos estudar é a **mediana**. A mediana é o termo central da série. Para determiná-la, precisamos inicialmente ordenar os valores, para só depois determinar quem é o central.

Assim, na variável idade, conseguimos ordenar os valores assim:

14, 22, **24**, 25, 25.

O termo central é 24. Logo, a mediana vale 24.

Na variável "escolaridade", conseguimos ordenar assim:

fundamental / médio /**superior** / superior / superior.

O termo central é "superior", que é nossa mediana.

Na variável "UF", sequer conseguimos ordenar as observações, pois trata-se de uma variável qualitativa nominal. Então não dá para determinar a mediana.

Resumindo tudo: para determinar a mediana precisamos que a variável, no mínimo, possa ser ordenada. Logo, não dá para apontar a mediana no caso de variável qualitativa nominal.

	Média	Mediana	Moda
Quantitativa discreta	ok	ok	ok
Quantitativa contínua	ok	ok	ok
Qualitativa ordinal	-	ok	ok
Qualitativa nominal	-	-	ok

2.6 - POLÊMICA ENVOLVENDO VARIÁVEIS DICOTÔMICAS

Vamos modificar um pouco o formulário que usamos como exemplo lá no começo dessa aula:

Formulário - Empresa Alfa - Setor de Recursos Humanos

1) Grau de instrução	0) fundamental 1) médio 2) superior
2) Estado civil	0) solteiro

	1) casado
3) Número de filhos	
4) Região de procedência	0) Centro-oeste 1) Sudeste 2) Nordeste 3) Norte 4) Sul
5) Altura	
6) Sexo	0) masculino 1) feminino

Perceba que agora nós atribuímos um valor numérico a cada atributo.

Como exemplo, na Região de procedência, a pessoa pode responder que vem da região 0 (centro-oeste), ou da região 1 (sudeste), ou da 2 (nordeste), etc.

Em outras palavras, a resposta para o campo "4 - Região de procedência" agora é um número. Mas isso não faz com que a variável passe a ser quantitativa. Ela continua sendo qualitativa, já que esses números 0, 1, 2, 3 e 4 são apenas "nomes" que foram dados às regiões, não significam absolutamente nada concreto.

Não faz sentido algum eu juntar uma pessoa que veio do Centro-oeste (0) com uma do Nordeste (2), e querer calcular a média:

$$\frac{0+2}{2} = 1$$

Olha o absurdo - estaríamos afirmando que a média entre Centro-oeste e Nordeste seria Sudeste, algo totalmente despropositado. Isso ocorre porque esses valores aí foram aleatoriamente escolhidos, são desprovidos de qualquer significado, são meros nomes dados aos atributos. Em essência, a variável continua sendo qualitativa.

O grande problema surge para o campo "6 - sexo". Quando os valores relacionados aos atributos forem 0 e 1, dizemos que a variável originada é chamada de **dicotômica**.



Ao atribuirmos os valores 0 e 1, criamos uma nova variável, chamada de **dicotômica**.

O problema é que essa variável dicotômica não só apresenta um significado, como é largamente usada em estatística. Ela dá origem a uma distribuição de probabilidade chamada de Bernoulli, para a qual podemos calcular média e variância. A distribuição de Bernoulli, por sua vez, dá origem à distribuição binomial. Que no limite para infinitos experimentos dá origem à distribuição de Poisson. Que por sua vez é empregada na teoria das filas (processos de nascimento e morte). E por aí vai.

Ou seja: a variável dicotômica pode ser tranquilamente tratada como uma variável quantitativa.

Para deixar mais claro o porquê disso, considere um grupo formado por 4 mulheres e 1 homem. Esse grupo apresentaria os seguintes valores da variável dicotômica:

$$1, 1, 1, 1, 0$$

Se calcularmos a média desse conjunto, o resultado será:

$$\frac{1+1+1+1+0}{5} = 0{,}8$$

Aí você pergunta: *qual o significado desse valor?*

E há uma resposta: ele nos indica que 80% das pessoas do grupo são mulheres.

Quando a atribuição de valores 0 e 1 é feita de tal forma que o resultado 1 represente os casos de interesse (casos de sucesso) e o resultado 0 represente os demais casos (fracasso), a média nos dá a proporção de casos favoráveis na amostra.

E aqui vem a polêmica: essa tal de variável dicotômica é quantitativa ou qualitativa?

Infelizmente, não tenho como dar uma resposta fechada. Há autores que dizem que é quantitativa [1], outros dizem que é qualitativa [2]. Há questões que dizem que é quantitativa [3]. Outras dizem que é qualitativa [4].

Como é um assunto pouco cobrado, ainda não dá para estabelecer diretrizes de como cada banca se comporta.

Notas:

[1] Joseph F Hair Jr [et al.]. **Análise Multivariada de Dados**. 6. ed - Porto Alegre: Bookman, 2009

[2] Bussab, Wilton de O. ; Morettin, Pedro A. **Estatística Básica**. 5ª ed - São Paulo: Saraiva, 2002

[3] Questões:

https://www.tecconcursos.com.br/conteudo/questoes/2691

https://www.tecconcursos.com.br/conteudo/questoes/2693

[4] Questão:

https://www.tecconcursos.com.br/conteudo/questoes/54954

3 – ESCALAS DE MEDIDA

Mensurar alguma coisa é atribuir um rótulo ou um número a determinado item em estudo.

Exemplo: se estivermos mensurando a altura de pessoas, seria possível atribuir ao indivíduo “A” a altura de 1,70 m, ao indivíduo “B” a altura de 1,85 m, e assim por diante.

Na hora de atribuir estes rótulos, nós usamos uma determinada regra (escala de medida). Nesta parte da aula nós vamos estudar as características das escalas de medida.

Antes de iniciar, vamos comparar duas variáveis diferentes: temperatura de uma cidade ao longo de um ano, e massa de diferentes objetos.

A temperatura é uma variável quantitativa contínua, pois pode assumir qualquer valor num intervalo real. Exemplo: pode assumir valores 23,5666... ºC, ou π ºC.

A massa também é uma variável quantitativa contínua, pois também pode assumir qualquer valor num intervalo real. Exemplo: pode assumir valores 13,444... kg, ou $\sqrt{2}$ kg.

Ou seja, temos variáveis do mesmo tipo: ambas são quantitativas e contínuas.

No entanto, suas escalas têm comportamento diferente - e é disso que vamos tratar agora.

Para o caso da temperatura, não há qualquer sentido físico na razão entre dois valores observados.

Se um dia tem temperatura de 20ºC e outro dia tem temperatura de 40ºC, não podemos dizer que o segundo dia é duas vezes mais quente que o primeiro. Isto ocorre porque a escala "temperatura" tem origem arbitrariamente escolhida, desprovida de sentido físico. Se você converter essas mesmas temperaturas para qualquer outra escala, como a escala Kelvin, verá que um valor não é mais o dobro do outro.

Já para a "massa", tudo muda de figura. Um objeto de 2 kg tem o dobro da massa de um de 1 kg. E isso valerá sempre, mesmo que você use outras unidades de medida - tais como grama, libra, ou qualquer outra.

Assim, muito embora ambas as variáveis sejam do tipo quantitativo contínuo, a forma como atribuímos os valores a cada uma delas muda. No caso da "massa" temos a chamada **escala de razão,** pois há sentido físico na razão entre duas observações diferentes. No caso da temperatura temos a chamada **escala intervalar,** pois não há sentido físico na razão entre duas observações diferentes.

3.1 – ESCALA NOMINAL

Na escala de medida nominal, os dados são divididos em categorias, sem que seja possível estabelecer qualquer relação de ordem ou grandeza. Não dá para dizer que um é melhor ou pior que o outro, maior ou menor, mais pesado ou mais leve, etc.

O máximo que se pode fazer é dizer que um objeto é diferente dos demais.

Exemplo:

Número de Técnicos Administrativos - USP 2013

Homens	8.940
Mulheres	8.510
Total	**17.450**

Suponha que a gente atribua o valor 1 a cada homem e o valor 2 a cada mulher.

Neste caso, 1 e 2 não têm significado físico algum, são apenas dois rótulos que usamos para dizer que homem (1) é diferente de mulher (2).

$$1 \neq 2$$

Só isso.

Na matemática, sabemos que:

$$2 > 1$$

$$1 + 1 = 2$$

$$2 - 1 = 1$$

$$2 \times 1 = 2$$

Mas nenhuma destas relações matemáticas pode ser trazida para o nosso caso. Não podemos dizer que mulher é maior que homem, apesar de 2 ser maior que 1. Nem dá para falar que um homem somado com outro homem resulta numa mulher. Ou que uma mulher, subtraída de um homem, resulta em um homem. Ou que dobrando um homem o resultado é uma mulher. Nada disso tem sentido algum.

O máximo que podemos dizer é que homem é diferente de mulher. Portanto estamos diante de uma escala de medida nominal.

Repare que para dados em escala nominal não faz sentido falar em média aritmética por exemplo. Na tabela acima, vamos tentar calcular a média:

Sexo	Valor atribuído	Quantidade	Produto
Homens	1	8.940	8.940
Mulheres	2	8.510	17.020
Total		**17.450**	**25.960**

Nesta aula nós não estudamos como se calcula a média de um conjunto de dados já tabelado. Então fiquem com a informação de que, para calcular a média, nós pegamos cada valor (segunda coluna), multiplicamos pela quantidade de vezes em que ele aparece (terceira coluna), obtendo a quarta coluna (produto).

Em seguida, totalizamos a quarta coluna (25.960).

Depois basta dividir pelo número total de objetos (17.450):

$$\frac{25.960}{17.450} = 1{,}487679$$

Ou seja, entre os técnicos administrativos da USP em 2013, o "sexo médio" é igual a 1,487679.

Pergunta: isso faz algum sentido?

Isso não tem sentido algum. É completamente absurdo pensar em "sexo médio", isso não tem nenhum significado matemático. É mais uma característica da escala de medida nominal: não dá para trabalhar com nenhuma medida estatística que exija operações matemáticas, como é o caso da média, do desvio padrão, da variância.

3.2 – Escala Ordinal

A escala ordinal é muito similar à nominal. A única diferença é que, aqui, já somos capazes de ordenar os objetos.

Exemplo: vamos atribuir valores a cada grupo etário.

Pirâmide etária - Brasil / 2010

Grupo	Valor atribuído	Quantidade
Crianças	1	45.932.295
Jovens	2	34.236.060
Adultos	3	89.996.845
Idosos	4	20.590.599
Total		**190.755.799**

Continuamos sem poder afirmar que um jovem vale duas crianças, mesmo sabendo que $2 = 2 \times 1$.

Os problemas são exatamente os mesmos verificados para a escala nominal.

No entanto, podemos sim ordenar os objetos de diferentes categorias. Se tomarmos uma criança (1), um jovem (2), um adulto (3) e um idoso (4), conseguimos ordená-los por idade, assim:

1, 2, 3, 4

Em síntese, na escala ordinal já é possível tanto dizer que objetos das diversas classes são diferentes entre si, quanto também é possível ordenar tais objetos.

Em outras palavras os rótulos já nos indicam que existe uma ordem entre as diferentes categorias.

3.3 – ESCALA INTERVALAR

Nessa escala já existe sentido em efetuar operações de adição e subtração.

O exemplo clássico, presente em todos os livros de estatística, é o da temperatura.

Considere que, em determinado local, a temperatura máxima medida ao longo de uma semana tenha sido a seguir:

Temperatura Máxima na cidade X

Dia	Temperatura (ºC)
Segunda	15
Terça	18
Quarta	21
Quinta	25,45
Sexta	29,3
Sábado	29,6
Domingo	30
Média	**24,050**

Alguns fatos interessantes:

(1) De segunda para terça a temperatura aumentou 3ºC. E de terça para quarta aumentou mais 3ºC. Ou seja:

$$21 - 18 = 18 - 15 = 3$$

Isso faz sentido matemático e também faz sentido físico. A diferença é constante porque a temperatura aumentou sempre do mesmo tanto. Ou seja, faz sim sentido somar e subtrair temperaturas.

Resultado: a **distância** entre duas observações tem sentido físico, tem sentido real. A distância entre as temperaturas de segunda e terça feira é 3ºC. A distância entre as temperaturas de segunda e

quarta feira é igual a 6ºC. Uma distância é o dobro da outra, pois, realmente, num intervalo a temperatura aumentou duas vezes mais que no outro.

(2) Faz sentido falar em média. Ao dizer que temperatura média foi de 24,05 ºC sabemos que as temperaturas desta semana giraram em torno deste valor.

(3) A temperatura do domingo é o dobro da temperatura de segunda. Mas isso não tem significado físico algum. Não dá para falar que domingo foi duas vezes mais quente que segunda.

Para deixar mais claro isso, basta mudar a escala da temperatura, de graus Celsius para Fahrenheit

Temperatura Máxima na cidade X

Dia	Temperatura (ºC)	Temperatura (F)
Segunda	15	59
Terça	18	64,4
Quarta	21	69,8
Quinta	25,45	77,81
Sexta	29,3	84,74
Sábado	29,6	85,28
Domingo	30	86
Média	**24,050**	**75,290**

Bastou mudar a unidade de medida para ver que, agora, a temperatura do domingo (86 F) não é mais o dobro da temperatura de segunda (59 F).

Ou seja, a **razão** entre valores diferentes não tem significado algum. Isso ocorre porque uma escala intervalar tem origem arbitrária. Seu ponto 0 é escolhido arbitrariamente.

3.4 – Escala de razão

Nesta escala temos tudo o que já existia na escala intervalar. E com um adicional: finalmente faz sentido calcular a razão entre dois valores diferentes.

Exemplo: massa dos objetos.

Considere 5 itens, com as seguintes massas:

Item	Massa (kg)
A	2
B	3
C	4
D	6,8
E	8
Média	**4,760**

Podemos dizer tranquilamente que o objeto C "pesa" o dobro do objeto A. Isto é, se juntarmos dois objetos do tipo A teremos exatamente o mesmo "peso" do objeto C.

A razão entre dois valores tem significado físico. Isto porque a origem do "peso" não é arbitrária. Para qualquer escala que a gente adotar (quilograma, grama, libra), um objeto com massa 0 terá exatamente isso: massa 0.

0 quilogramas = 0 gramas = 0 libras

4 – SÉRIES ESTATÍSTICAS

4.1 – SÉRIES CRONOLÓGICAS OU TEMPORAIS

São séries de dados em que uma variável é medida em diferentes instantes de tempo, para um mesmo fato e um mesmo local.

Ou seja:

- fato e local = constantes
- tempo = variável

Exemplo: Índice de Preços ao Consumidor (IPC) para os meses de março e dezembro de 2014:

Índice de Preços ao Consumidor - Brasil

Período	Inflação (%)
mar/14	6,154
abr/14	6,28
mai/14	6,374
jun/14	6,523
jul/14	6,502
ago/14	6,513
set/14	6,746
out/14	6,587
nov/14	6,555
dez/14	6,409

Notem que:

- o fato é sempre o mesmo = inflação
- o local é fixo = Brasil
- o período varia: de março a dezembro de 2014

4.2 – SÉRIES GEOGRÁFICAS OU DE LOCALIZAÇÃO

Nesse tipo de série, o local é variável, mas o tempo e o fato ficam fixos.

Exemplo: PIB dos países da América do Sul em 2012:

Produto Interno Bruto da América do Sul
Por Países, 2012

(Em Valores Correntes - US$ Milhões)

Países	2012
Argentina	475.502
Bolívia	27.035
Brasil	2.252.664
Chile	269.869
Colômbia	369.606
Equador	84.040
Guiana e Guiana Francesa	2.851
Paraguai	25.502
Peru	203.790
Suriname	5.012
Uruguai	49.920
Venezuela	381.286

Fonte: Banco Mundial (favereiro 2014)
OBS: Em 2012, a Venezuela tornou-se o quinto membro a integrar o Mercosul.

Observem que:

- O fato é sempre o mesmo: PIB em milhões de dólares
- O período é fixo: 2012
- O local é que muda: Argentina, Bolívia, Brasil, etc.

4.3 – SÉRIES ESPECÍFICAS

Agora o fato varia, mas o local e o período ficam fixos. Exemplo:

Número de Técnicos Administrativos - USP 2013

Nível superior	4.174
Nível técnico	8.104
Nível básico	5.149
Outros	24
Total	**17.451**

O local é fixo (Universidade de São Paulo). O tempo é fixo (2013). O fato varia: na primeira linha quantificamos técnicos administrativos de nível superior; na segunda, de nível técnico, e assim por diante.

Observem que acima tivemos na verdade um fato maior (técnico administrativo) que por sua vez é decomposto em categorias.

4.4 – SÉRIES CONJUGADAS OU MISTAS

Uma série é conjugada quando é uma combinação entre os diferentes tipos de séries estudados acima, ou seja, há mais de um item variando - lugar, fato e tempo.

Como exemplo, trago a tabela a seguir, extraída do site www.insper.edu.br, **referente à cidade de São Paulo**:

Parcela dos entrevistados que foram vitimizados ao longo da vida, por tipo de ocorrência

	Parcela dos entrevistados vitimizados		
	2003	2008	2013
Teve um bem roubado/furtado*	34,9%	32,7%	33,2%
Ameaçado com arma de fogo	23,6%	19,3%	17,0%
Teve veículo roubado/furtado	16,7%	19,5%	16,8%
Teve casa invadida	18,3%	15,3%	13,6%
Sofreu agressão física	11,7%	10,7%	11,0%
Ameaçado com outro tipo de arma	8,2%	7,3%	9,3%
Alguém disparou arma de fogo contra a pessoa	4,1%	3,1%	2,3%
Usou/mostrou arma para se defender	3,2%	2,3%	2,2%
Foi ferido por arma de fogo	1,2%	1,5%	1,3%
Ferido com outro tipo de arma	1,8%	1,3%	1,2%

*Não inclui roubo/furto de veículo e casa.

Notem que o local é fixo (cidade de São Paulo). Mas temos variação no ano e no fato, logo, temos uma série conjugada.

5 – Caderno no Tec Concursos

Logo abaixo traremos nossa lista de questões comentadas.

Para os alunos que quiserem uma ênfase maior na resolução de exercícios, deixamos também o link de compartilhamento de um caderno contendo questões dos assuntos tratados nesta aula.

https://tec.ec/s/QUbFB

Sumário

Lista de exercícios 2
Ramos da estatística; Censo e amostra 2
Tipos de variáveis. Medidas estatísticas relacionadas a cada tipo de variável. 3
Escalas de medida 12
Séries Estatísticas 13
Gabaritos 14
Questões comentadas 15
Ramos da estatística; Censo e amostra 15
Tipos de variáveis. Medidas estatísticas relacionadas a cada tipo de variável. 19
Escalas de medida 34
Séries Estatísticas 36

LISTA DE EXERCÍCIOS

RAMOS DA ESTATÍSTICA; CENSO E AMOSTRA

01. (CESPE / Sefaz-AL – 2002)

Julgue o item a seguir.

Um censo consiste no estudo de todos os indivíduos da população considerada.

02. (CESPE / Sefaz-AL – 2002)

Julgue o item a seguir.

Como a realização de um censo tipicamente é muito onerosa e(ou) demorada, muitas vezes é conveniente estudar um subconjunto próprio da população, denominado amostra.

03. (CESPE / TC-DF – 2002)

Julgue o item seguinte, relativo a estatística e matemática financeira.

Por Estatística Descritiva entende-se um conjunto de ferramentas, tais como gráficos e tabelas, cujo objetivo é apresentar, de forma resumida, um conjunto de observações.

04. (VUNESP/ CM-SC – 2013)

Trata-se da estatística que tem as atribuições de obtenção, organização, redução e representação de dados estatísticos, bem como a obtenção de algumas informações que auxiliam na descrição de um fenômeno observado. Esta estatística é denominada

a) Coletora.
b) Celetista.
c) Populacional.
d) Amostral.
e) Descritiva.

05. (CESPE / Agente Penitenciário Federal – 2015)

O diretor de um sistema penitenciário, com o propósito de estimar o percentual de detentos que possuem filhos, entregou a um analista um cadastro com os nomes de 500 detentos da instituição para que esse profissional realizasse entrevistas com os indivíduos selecionados.

A partir dessa situação hipotética e dos múltiplos aspectos a ela relacionados, julgue o item, referente a técnicas de amostragem.

A diferença entre um censo e uma amostra consiste no fato de esta última exigir a realização de um número maior de entrevistas.

06. (CONSULPLAN / Técnico Nível Superior/Engenheiro – Patos de Minas – 2015)

População em termos estatísticos serve para descrever um grande conjunto de unidades que tem algo em comum. As razões que levam os pesquisadores a trabalhar com amostras são diversas. São alternativas referentes às amostras, EXCETO:

a) Populações muito grandes.

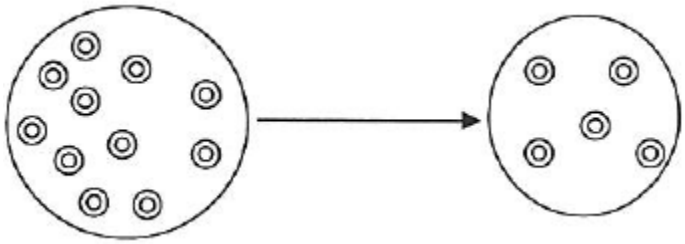
b) Custo e rapidez dos censos.

c) Impossibilidade física de examinar toda a população.

d) Comprovado valor científico das informações coletadas por meio de amostras.

07. (INSTITUTO MAIS/ CM – Osasco - 2016)

Observe os conjuntos abaixo.



Pode-se afirmar que se trata de uma representação para ilustrar, estatisticamente.

a) média aritmética simples e média aritmética ponderada.
b) moda e mediana.
c) quartis e decis.
d) população e amostra

TIPOS DE VARIÁVEIS. MEDIDAS ESTATÍSTICAS RELACIONADAS A CADA TIPO DE VARIÁVEL.

08. (CESPE/Auditor Federal de Controle Externo – 2008)

Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (*X*) e vendidos (*Y*) em determinado município, nos anos de 2005 a 2007.

ano	número de imóveis	
	ofertados (*X*)	vendidos (*Y*)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correio Braziliense, 29/4/2008, p. 17 (com adaptações).

Com respeito ao texto, considere que cada imóvel ofertado em determinado ano seja classificado como vendido ou não-vendido, e, a um imóvel e classificado como vendido seja atribuído um valor $Z = 1$, e, ao imóvel classificado como não-vendido, seja atribuído um valor $Z = 0$. Supondo-se que as classificações dos imóveis como vendido ou não-vendido em um dado ano possam ser consideradas como sendo realizações de uma amostragem aleatória simples, julgue os itens a seguir.
A variável *Z* é classificada como variável qualitativa nominal, pois representa o atributo do imóvel como vendido ou não-vendido.

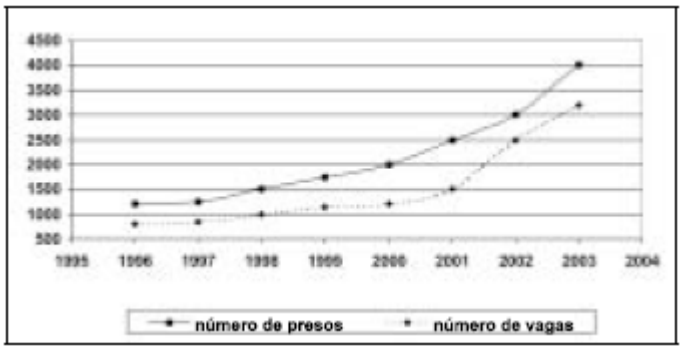
09. (CESPE/ Auditor Federal de Controle Externo – 2008)

tipo	2006	2007
funcionários	70	74
estagiários	16	16
prestadores de serviço	14	10
total	**100**	**100**

Demonstrações contábeis da CAIXA – Exercício de 2007.
Internet: <www.caixa.gov.br> (com adaptações).

Considerando a tabela acima, que apresenta a distribuição do quadro de colaboradores da CAIXA, em mil pessoas, no final dos anos de 2006 e 2007, julgue os itens seguintes.
Se uma variável *X* registra a classificação de cada pessoa do quadro como funcionária, estagiária ou prestadora de serviço, então *X* é uma variável qualitativa.

10. (CESPE /Perito Criminal Federal - 2004)



Nos últimos oito anos, a população carcerária em uma unidade da Federação cresceu de 1.200 presos (1996) para 4.000 presos (2003). Essa população carcerária é formada por presos nas casas penais, seccionais e delegacias. Por causa desse crescimento, foram construídas novas cadeias públicas, penitenciárias e novos blocos carcerários. Mesmo assim, não foi possível resolver o problema de superlotação. Em 1996, a capacidade total de lotação das casas penais, seccionais e delegacias era de apenas 800 vagas. Após a inauguração das novas instalações em 2003, o número de vagas aumentou para 3.200, o que resulta em *deficit* de 800 vagas. O gráfico acima apresenta a evolução temporal da população carcerária (linha contínua) e do número de vagas (linha pontilhada) de 1996 a 2003.

Com base na situação hipotética e no gráfico apresentados ao lado, julgue o item a seguir.

A capacidade total de lotação das casas penais, seccionais e delegacias (número de vagas) em 2000 é uma variável aleatória contínua.

11. (CESPE /Perito Criminal Federal – 2004)

Um projeto de serviços de assistência social foi desenvolvido para ser implementado em todas as delegacias e plantões policiais de um estado brasileiro. Porém, antes da sua aplicação em todo o estado, ele foi implementado em 10 municípios, em caráter experimental, por 12 meses. Esses municípios foram escolhidos aleatoriamente entre os 250 municípios do estado. Nesse período experimental, foram registradas 48.000 ocorrências nos 10 municípios selecionados. Em 25% dessas ocorrências, as pessoas envolvidas foram encaminhadas aos assistentes sociais. A partir dessas ocorrências, os 100 assistentes sociais envolvidos nesse projeto atenderam, em média, 500 pessoas por mês. Os resultados obtidos foram positivos, observando-se uma queda na reincidência de denúncias e ocorrências registradas nesses municípios após a implementação do projeto.
A partir dos dados apresentados no texto acima, julgue o item subseqüente.
O número de ocorrências registradas durante o período experimental de 12 meses nos 10 municípios selecionados (48.000) é a realização de uma variável aleatória contínua.

12. (CESPE/ Perito Criminal Federal – 1997)

A tabela seguinte apresenta informações relativas ao recolhimento de imposto de renda e salário mensal de 16 funcionários de duas microempresas no ano de 1997.

	Empresa A		Empresa B	
Funcionário	Recolhimento de Imposto*	Salário Mensal**	Recolhimento de Imposto*	Salário Mensal**
1	1	2,0	0	2,2
2	0	1,0	0	1,1
3	0	2,0	0	2,2
4	0	2,0	0	2,2
5	1	5,0	0	5,5
6	1	10,0	0	11,0
7	0	2,0	0	2,2
8	1	2,0	0	2,2
9	1	2,0	1	2,2
10	1	5,0	1	5,5
11	1	5,0	1	5,5
12	1	5,0	1	5,5
13	0	50,0	1	55,0
14	0	1,0	1	1,1
15	0	1,0	1	1,1
16	1	5,0	1	5,5

*Recolhimento de Imposto: 0 = NÃO, 1 = SIM
**Salário Mensal: Número de salários mínimos

Com base na tabela, julgue o item

A média e a moda são medidas descritivas adequadas para avaliar a variável Recolhimento de Imposto.

13. (CESPE/ Perito Criminal Federal – 1997)

A tabela seguinte apresenta informações relativas ao recolhimento de imposto de renda e salário mensal de 16 funcionários de duas microempresas no ano de 1997.

	Empresa A		Empresa B	
Funcionário	Recolhimento de Imposto*	Salário Mensal**	Recolhimento de Imposto*	Salário Mensal**
1	1	2,0	0	2,2
2	0	1,0	0	1,1
3	0	2,0	0	2,2
4	0	2,0	0	2,2
5	1	5,0	0	5,5
6	1	10,0	0	11,0
7	0	2,0	0	2,2
8	1	2,0	0	2,2
9	1	2,0	1	2,2
10	1	5,0	1	5,5
11	1	5,0	1	5,5
12	1	5,0	1	5,5
13	0	50,0	1	55,0
14	0	1,0	1	1,1
15	0	1,0	1	1,1
16	1	5,0	1	5,5

*Recolhimento de Imposto: 0 = NÃO, 1 = SIM
**Salário Mensal: Número de salários mínimos

Com base na tabela, julgue o item

A mediana é adequada para avaliar a variável Salário Mensal porque é resistente a valores extremos, e para avaliar a variável Recolhimento de Imposto porque é nominal.

14. (SMA-RJ (antiga FJG) / Fiscal de Rendas do Munícipio do RJ – 2002)

Os dados de um determinado estudo representam muitas variáveis para cada uma das pessoas que se submeteram ao estudo. Uma variável considerada qualitativa é a seguinte:

a) idade

b) altura

c) sexo

d) peso

15. (CESPE/ TRE – ES - 2011)

cargo	candidatos	candidatos aptos	eleitos
presidente da República	9	9	1
governador de estado	170	156	27
senador	272	234	54
deputado federal	6.021	5.058	513
deputado estadual/distrital	15.268	13.076	1.059
total	21.640	18.533	1.658

Internet: <www.tse.gov > (com adaptações).

Com base na tabela acima, referente às eleições de 2010, que apresenta a quantidade de candidatos para os cargos de presidente da República, governador de estado, senador, deputado federal e deputado estadual/distrital, bem como a quantidade de candidatos considerados aptos pela justiça eleitoral e o total de eleitos para cada cargo pretendido, julgue o item a seguir.
A variável "cargo" classifica-se como uma variável qualitativa ordinal.

16. (CESPE/ TER - ES – 2011)

quantidade de eleitores	quantidade de municípios
0 ⊢ 2.000	364
2.000 ⊢ 4.000	1.000
4.000 ⊢ 6.000	3.000
6.000 ⊢ 8.000	1.000
8.000 ⊢ 10.000	200
total	5.564

A tabela acima apresenta uma distribuição hipotética das quantidades de eleitores que não votaram no segundo turno da eleição para presidente da República bem como os números de municípios em que essas quantidades ocorreram. Com base nessa tabela, julgue o item segunite, relativo à análise exploratória de dados.

Na tabela de frequências, o uso de intervalos de classe permite concluir que a variável em questão é contínua.

17. (FUNIVERSA/ PC – GO - 2015)

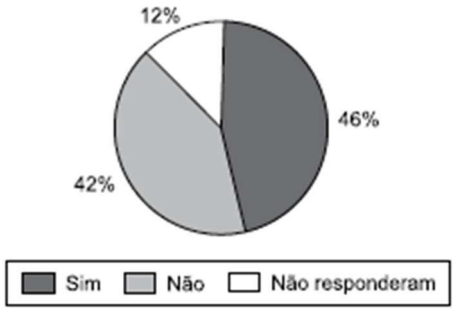
A delegacia virtual é um serviço de registro de ocorrências, via Internet, por meio do site da Polícia Civil de Goiás. O objetivo é fornecer, para o cidadão, comodidade no registro de alguns tipos de ocorrências, otimizando os recursos e melhorando o atendimento. Depois de verificada a veracidade da informação pela Polícia Civil e autorizado pela autoridade policial, o usuário receberá, pelo correio eletrônico (*e-mail*) fornecido no registro, as informações sobre como obter uma cópia do Boletim de Ocorrência e qual delegacia de Polícia está responsável por apurar o fato. A qualquer momento, poderá ser consultado o andamento do seu registro, acessando a delegacia virtual. O serviço está disponível, sem interrupções, a qualquer dia e em qualquer horário. Pela delegacia virtual, o usuário poderá registrar as seguintes ocorrências: roubo/furto de objetos e(ou) valores, roubo/furto de celular, roubo/furto de documentos, perda ou extravio de documentos e(ou) objetos, desaparecimento de pessoa e acidente de trânsito sem vítimas.

Internet:<www.policiacivil.go.gov.br> (com adaptações).

Para a Estatística, a variável "tipo de ocorrência" é classificada como
a) qualitativa nominal.
b) quantitativa discreta.
c) quantitativa contínua.
d) qualitativa discreta.
e) qualitativa contínua.

18. (VUNESP/ Analista/Informações, Cultura e Desporto/Biblioteconomia - Pref SP - 2015)

Considere o gráfico a seguir, que representa as respostas à enquete "Você já leu pelo menos dois livros no primeiro semestre desse ano?", pergunta essa feita a um grupo de 60 pessoas pesquisadas.

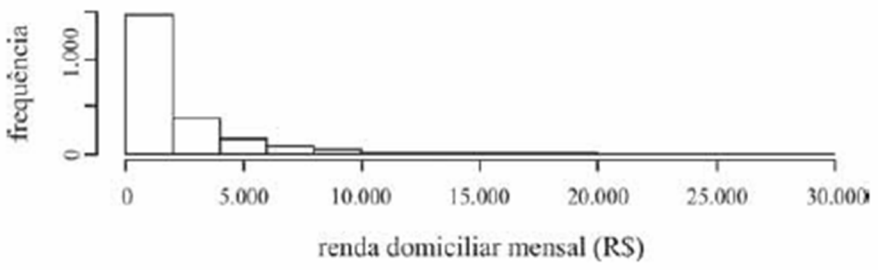


Levando-se em conta as medidas centrais média, moda e mediana, assim como a variável envolvida na estatística dessa enquete, é correto afirmar que pode(m) ser determinada(s), dentre essas três medidas,

a) a moda e a mediana, apenas.
b) a moda, apenas.

c) a média, apenas.
d) a média e a mediana, apenas.
e) a média, a moda e a mediana.

19. (CESPE/ TELEBRAS – 2015)

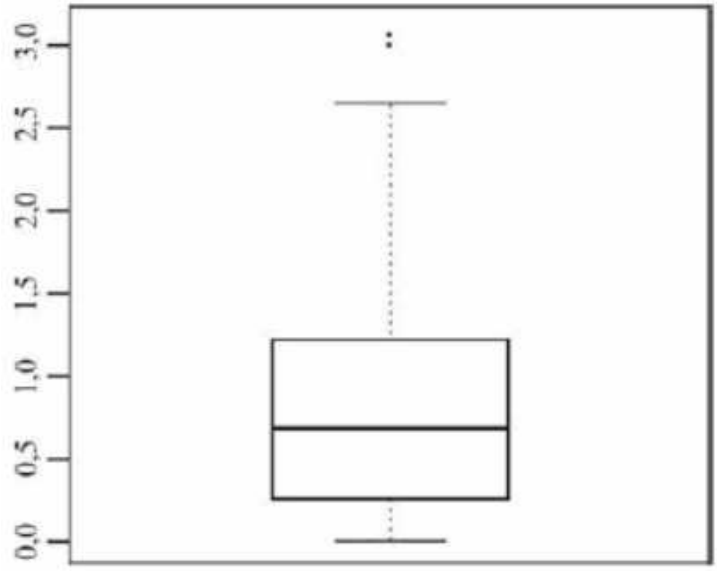


Uma empresa coletou e armazenou em um banco de dados diversas informações sobre seus clientes, entre as quais estavam o valor da última fatura vencida e o pagamento ou não dessa fatura. Analisando essas informações, a empresa concluiu que 15% de seus clientes estavam inadimplentes. A empresa recolheu ainda dados como a unidade da Federação (UF) e o CEP da localidade em que estão os clientes. Do conjunto de todos os clientes, uma amostra aleatória simples constituída por 2.175 indivíduos prestou também informações sobre sua renda domiciliar mensal, o que gerou o histograma apresentado.

Com base nessas informações e no histograma, julgue o item a seguir.

O CEP da localidade dos clientes e o valor da última fatura vencida são variáveis quantitativas.

20. (CESPE/ TCE - PA – 2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho *X* permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

X representa uma variável qualitativa ordinal.

21. (CESPE/ TCE - PA – 2016)

número diário de denúncias registradas (X)	frequência relativa
0	0,3
1	0,1
2	0,2
3	0,1
4	0,3
total	**1,0**

A tabela precedente apresenta a distribuição de frequências relativas da variável X, que representa o número diário de denúncias registradas na ouvidoria de determinada instituição pública. A partir das informações dessa tabela, julgue o item seguinte.

A variável X é do tipo qualitativo nominal.

ESCALAS DE MEDIDA

22. (CESPE / Antaq - 2009)

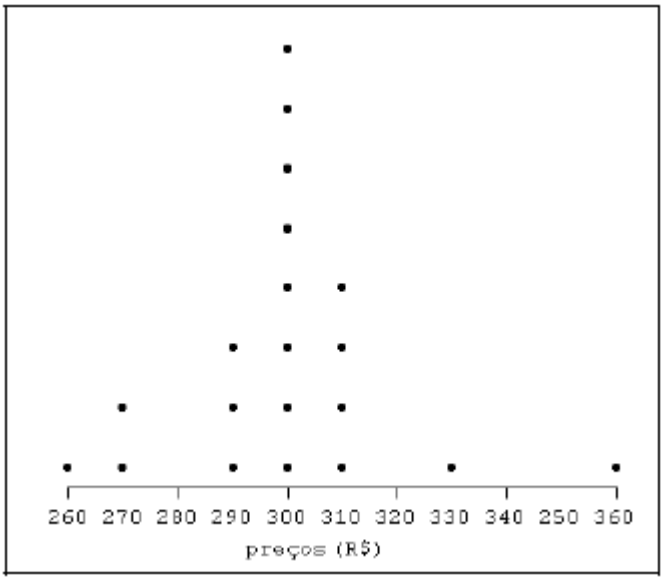
	variável	2003	2004	2005	2006	2007
exportação	X	40	46	50	52	54
importação	Y	20	21	22	24	27
total	X + Y	60	67	72	76	81

Considerando a tabela acima, que apresenta a movimentação anual de cargas no porto de Santos de 2003 a 2007, em milhões de toneladas/ ano e associa as quantidades de carga movimentadas para exportação e importação às variáveis X e Y, respectivamente, julgue o item subsequente.

X e Y são variáveis qualitativas em escala ordinal.

23. (CESPE / Antaq – 2005)

Em uma pesquisa realizada para verificar os preços cobrados em 2002 para o embarque de contêineres em um porto brasileiro, foram apurados 20 preços, que estão representados na figura abaixo (diagrama de pontos).



Com base nas informações do texto e da figura acima, julgue o item a seguir.

O preço cobrado para o embarque de contêineres em 2002 pode ser considerado uma variável discreta e qualitativa em escala nominal.

SÉRIES ESTATÍSTICAS

24. (CESPE/ Auditor Federal de Controle Externo – 2008)

Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (*X*) e vendidos (*Y*) em determinado município, nos anos de 2005 a 2007.

ano	número de imóveis	
	ofertados (*X*)	vendidos (*Y*)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correio Braziliense, 29/4/2008, p. 17 (com adaptações).

Considerando as informações do texto, julgue os itens subsequentes.

A variável *X* forma uma série estatística denominada série temporal.

GABARITOS

1	CERTO	13	ERRADO
2	CERTO	14	C
3	CERTO	15	CERTO
4	E	16	ERRADO
5	ERRADO.	17	A
6	B	18	B
7	D	19	ERRADO.
8	ERRADO	20	ERRADO.
9	CERTO	21	ERRADO.
10	ERRADO	22	ERRADO.
11	ERRADO	23	ERRADO.
12	ERRADO	24	CERTO

QUESTÕES COMENTADAS

RAMOS DA ESTATÍSTICA; CENSO E AMOSTRA

01. (CESPE / Sefaz-AL – 2002)

Julgue o item a seguir.

Um censo consiste no estudo de todos os indivíduos da população considerada.

Comentários

O item está perfeito!

Quando analisamos todos os elementos de uma população, temos um censo. Contrariamente, quando analisamos apenas uma parte da população, temos uma amostra.

Raramente se faz um censo, por razões de custo (é um procedimento caro) e tempo (é um procedimento demorado).

Gabarito: CERTO

02. (CESPE / Sefaz-AL – 2002)

Julgue o item a seguir.

Como a realização de um censo tipicamente é muito onerosa e(ou) demorada, muitas vezes é conveniente estudar um subconjunto próprio da população, denominado amostra.

Comentários

Item correto

Para melhor entendimento, considere o censo realizado pelo IBGE. É necessário contratar muitas e muitas pessoas para percorrer o país inteiro entrevistando as pessoas. Ou seja, é algo caro. Além disso, demora para conseguir abarcar todas as famílias do país (é demorado).

Por isso é tão difícil realizar um censo. Um procedimento alternativo é analisar só um pedaço da população, chamado de amostra. A partir da amostra obtida é possível ter ideia do que ocorre na população.

Gabarito: CERTO

03. (CESPE / TC-DF – 2002)

Julgue o item seguinte, relativo a estatística e matemática financeira.

Por Estatística Descritiva entende-se um conjunto de ferramentas, tais como gráficos e tabelas, cujo objetivo é apresentar, de forma resumida, um conjunto de observações.

Comentários

Perfeito. A estatística descritiva visa descrever um conjunto de dados, ou seja, representá-los de forma sucinta. Isso pode ser feito por meio de tabelas, de gráficos, ou medidas "resumo", como medidas de posição, de dispersão, de assimetria, entre outras.
Item certo.

Gabarito: CERTO

04. (VUNESP/ CM-SC – 2013)

Trata-se da estatística que tem as atribuições de obtenção, organização, redução e representação de dados estatísticos, bem como a obtenção de algumas informações que auxiliam na descrição de um fenômeno observado. Esta estatística é denominada

a) Coletora.
b) Celetista.
c) Populacional.
d) Amostral.
e) Descritiva.

Comentários

Pelo texto, no fundo estamos interessados em descrever um fenômeno:

“Trata-se da estatística que tem as atribuições de obtenção, organização, redução e representação de dados estatísticos, bem como a obtenção de algumas informações que auxiliam na **descrição** de um fenômeno observado. Esta estatística é denominada”

Estamos diante da estatística **descritiva**.

A estatística descritiva busca descrever um conjunto de dados por meio de medidas como média, mediana, moda, desvio padrão, variância, coeficiente de assimetria, coeficiente de curtose.

Todas essas medidas buscam descrever, de forma resumida, todo o conjunto de dados, segundo alguma de suas características.

Vejam que, para calcular uma média, precisamos primeiro obter os dados. Em seguida, organizamos (exemplo: em uma tabela, agrupados por valor). A própria tabela já é também uma forma de apresentar os dados (apresentação). Em seguida calculamos a média (resumimos os dados).

Gabarito: E

05. (CESPE / Agente Penitenciário Federal – 2015)

O diretor de um sistema penitenciário, com o propósito de estimar o percentual de detentos que possuem filhos, entregou a um analista um cadastro com os nomes de 500 detentos da instituição para que esse profissional realizasse entrevistas com os indivíduos selecionados.

A partir dessa situação hipotética e dos múltiplos aspectos a ela relacionados, julgue o item, referente a técnicas de amostragem.

A diferença entre um censo e uma amostra consiste no fato de esta última exigir a realização de um número maior de entrevistas.

Comentários

O censo consiste em entrevistar todos os elementos da população.

A amostra consiste em entrevistar um subconjunto da população.

Portanto, é o censo quem exige mais entrevistas.

Gabarito: ERRADO.

06. (CONSULPLAN / Técnico Nível Superior/Engenheiro – Patos de Minas – 2015)

População em termos estatísticos serve para descrever um grande conjunto de unidades que tem algo em comum. As razões que levam os pesquisadores a trabalhar com amostras são diversas. São alternativas referentes às amostras, EXCETO:

a) Populações muito grandes.

b) Custo e rapidez dos censos.

c) Impossibilidade física de examinar toda a população.

d) Comprovado valor científico das informações coletadas por meio de amostras.

Comentários

Alternativa A - CORRETO. Quando a população é muito grande, fica inviável analisar todos os seus elementos. Exemplo: pesquisa de intenção de votos. Não dá para os institutos de pesquisa

entrevistarem todos os eleitores, isso seria extremamente caro e demorado. Então o que eles fazem é tomar elementos por conveniência.

Alternativa B - INCORRETO. Censo são caros e demorados. Pensa no censo realizado pelo IBGE. Ter que colocar pessoal na rua para rodar o país inteiro, passando de casa em casa, isso demora muito e é caro.

Alternativa C - CORRETA. Considere um estudo sobre o nível de proteção de os carros dão para os passageiros em caso de colisão frontal. O teste consiste em colocar bonecos dentro do carro, acelerá-lo contra uma parede, assim:



Simplesmente não dá para testar 100% dos carros produzidos, se não a fábrica não vai ter nenhum deles para comercializar e o teste perderia completamente o sentido. Para quê testar algo que não será comercializado? Então é testada apenas uma amostra.

Alternativa D - CORRETA. Se é viável usar amostras, bem mais baratas e rápidas, e ainda assim obter resultados satisfatórios, mais um motivo para empregá-las. No caso de amostras bem feitas, os resultados podem ser muito bons.

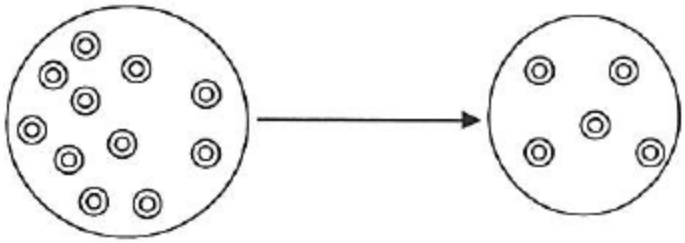
O problema que surge no imaginário das pessoas é fruto de amostras mal feitas, muitas vezes intencionalmente. As pesquisas de intenção de voto por exemplo, elas, em vários casos, são feitas com o intuito de influenciar o eleitor e não para de fato tirar um retrato da população. Depois o

resultado da urnas vem e as manchetes são no sentido de "resultado inesperado". E aí a pobre coitada da Estatística fica desacreditada.

Gabarito: B

07. (INSTITUTO MAIS/ CM – Osasco - 2016)

Observe os conjuntos abaixo.



Pode-se afirmar que se trata de uma representação para ilustrar, estatisticamente.

a) média aritmética simples e média aritmética ponderada.
b) moda e mediana.
c) quartis e decis.
d) população e amostra

Comentários

Partimos de um conjunto "maior" (a população) e extraímos um pedaço (a amostra). **Resposta: D**.

Média, mediana e moda tentam apontar para o centro de um conjunto de dados. Isso não tem nada a ver com o desenho. Erradas as letras A e B.

Quartis e decis dividem o conjunto de dados em partes com mesma quantidade de elementos. Exemplo: os quartis separam a distribuição em quatro partes, cada uma delas com 25% das observações. Isso também não tem nada a ver com o desenho. Errada a letra C.

Gabarito: D

TIPOS DE VARIÁVEIS. MEDIDAS ESTATÍSTICAS RELACIONADAS A CADA TIPO DE VARIÁVEL.

08. (CESPE/Auditor Federal de Controle Externo – 2008)

Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (*X*) e vendidos (*Y*) em determinado município, nos anos de 2005 a 2007.

ano	número de imóveis	
	ofertados (*X*)	vendidos (*Y*)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correio Braziliense, 29/4/2008, p. 17 (com adaptações).

Com respeito ao texto, considere que cada imóvel ofertado em determinado ano seja classificado como vendido ou não-vendido, e, a um imóvel e classificado como vendido seja atribuído um valor $Z = 1$, e, ao imóvel classificado como não-vendido, seja atribuído um valor $Z = 0$. Supondo-se que as classificações dos imóveis como vendido ou não-vendido em um dado ano possam ser consideradas como sendo realizações de uma amostragem aleatória simples, julgue os itens a seguir. A variável *Z* é classificada como variável qualitativa nominal, pois representa o atributo do imóvel como vendido ou não-vendido.

Comentários

A variável Z é uma variável dicotômica.

Interessante observar que as variáveis dicotômicas são utilizadas para permitir que variáveis qualitativas sejam analisadas como se fossem quantitativas.

A variável que assume os valores "vendido" e "não-vendido" é nominal. No entanto, quando atribuímos a cada uma dessas realizações um valor numérico (0 ou 1), obtemos números e podemos passar a tratar tal variável como quantitativa. Isso pode ser feito porque a atribuição de valores numéricos foi feita de modo a fornecer algo passível de interpretação. No caso, esse tipo de variável dá origem a distribuições discretas de probabilidade muito utilizadas: a de Bernoulli e a Binomial.

A cada imóvel analisado, temos sucesso ou fracasso (se Z = 1 ou 0) e, somando todos os valores de Z, temos, na verdade, a contagem de quantos imóveis em uma dada amostra foram vendidos. Assim, a variável Z é quantitativa. Nesse ponto, citamos o livro "Análise multivariada de dados", dos autores Hair, Black, Babin, Anderson e Taham:

"Uma variável dicotômica é aquela que foi convertida para uma distribuição métrica e representa uma categoria de uma variável independente não-métrica."

HAIR, J, et al, *Análise multivariada de dados.* Tradução Adonai Schlup Sant'Anna - 6ª edição. Porto Alegre: Bookman, 2009

Item errado.

No entanto, é pertinente registrar a opinião de Bussab e Morettin, segundo os quais, apesar de a variável dicotômica apresentar realizações representadas por números (0 e 1), ela não deixa de ser qualitativa:

"Existe um tipo de variável qualitativa para a qual essa quantificação é muito útil: a chamada variável dicotômica. Para essa variável só podem ocorrer duas realizações, usualmente chamadas de sucesso e fracasso."

BUSSAB, W. e MORETTIN, P. *Estatística Básica.* 5ª edição. São Paulo: Saraiva, 2002.

Gabarito: ERRADO

09. (CESPE/ Auditor Federal de Controle Externo – 2008)

tipo	2006	2007
funcionários	70	74
estagiários	16	16
prestadores de serviço	14	10
total	**100**	**100**

Demonstrações contábeis da CAIXA – Exercício de 2007.
Internet: <www.caixa.gov.br> (com adaptações).

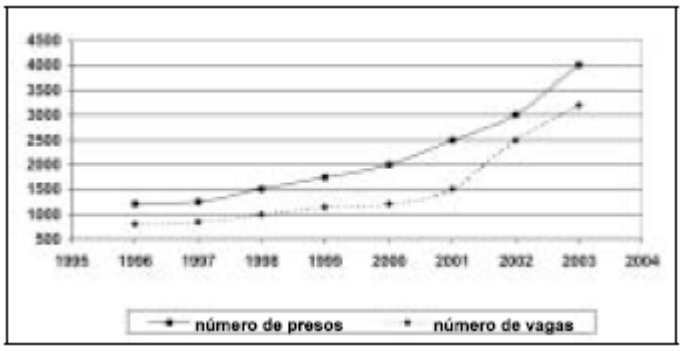
Considerando a tabela acima, que apresenta a distribuição do quadro de colaboradores da CAIXA, em mil pessoas, no final dos anos de 2006 e 2007, julgue os itens seguintes.
Se uma variável *X* registra a classificação de cada pessoa do quadro como funcionária, estagiária ou prestadora de serviço, então *X* é uma variável qualitativa.

Comentários

De fato, as possíveis realizações de X não são números, são atributos. Temos uma variável qualitativa.

Gabarito: CERTO

10. (CESPE /Perito Criminal Federal - 2004)



Nos últimos oito anos, a população carcerária em uma unidade da Federação cresceu de 1.200 presos (1996) para 4.000 presos (2003). Essa população carcerária é formada por presos nas casas penais, seccionais e delegacias. Por causa desse crescimento, foram construídas novas cadeias públicas, penitenciárias e novos blocos carcerários. Mesmo assim, não foi possível resolver o problema de superlotação. Em 1996, a capacidade total de lotação das casas penais, seccionais e delegacias era de apenas 800 vagas. Após a inauguração das novas instalações em 2003, o número de vagas aumentou para 3.200, o que resulta em *deficit* de 800 vagas. O gráfico acima apresenta a evolução temporal da população carcerária (linha contínua) e do número de vagas (linha pontilhada) de 1996 a 2003.

Com base na situação hipotética e no gráfico apresentados ao lado, julgue o item a seguir.

A capacidade total de lotação das casas penais, seccionais e delegacias (número de vagas) em 2000 é uma variável aleatória contínua.

Comentários

A capacidade total de lotação não pode assumir qualquer valor em um intervalo real. Ela assume apenas valores inteiros, como: 1, 2, 3, 4, ... Logo, é uma **variável discreta**.

Além disso, se tomarmos seu valor exclusivamente no ano de 2000, temos uma observação, algo fixo, constante, que não varia. Ou seja, seu valor para o ano 2000 é uma constante, e não variável. **Item errado.**

Gabarito: ERRADO

11. (CESPE /Perito Criminal Federal – 2004)

Um projeto de serviços de assistência social foi desenvolvido para ser implementado em todas as delegacias e plantões policiais de um estado brasileiro. Porém, antes da sua aplicação em todo o estado, ele foi implementado em 10 municípios, em caráter experimental, por 12 meses. Esses

municípios foram escolhidos aleatoriamente entre os 250 municípios do estado. Nesse período experimental, foram registradas 48.000 ocorrências nos 10 municípios selecionados. Em 25% dessas ocorrências, as pessoas envolvidas foram encaminhadas aos assistentes sociais. A partir dessas ocorrências, os 100 assistentes sociais envolvidos nesse projeto atenderam, em média, 500 pessoas por mês. Os resultados obtidos foram positivos, observando-se uma queda na reincidência de denúncias e ocorrências registradas nesses municípios após a implementação do projeto.
A partir dos dados apresentados no texto acima, julgue o item subseqüente.
O número de ocorrências registradas durante o período experimental de 12 meses nos 10 municípios selecionados (48.000) é a realização de uma variável aleatória contínua.

Comentários

O número de ocorrências, sendo fruto de uma contagem, pode apenas assumir valores inteiros não negativos, tais como: 0, 1, 2, 3, 4, 5, 6, 7.... Logo, não pode assumir qualquer valor em um dado intervalo real. Deste modo, é uma variável **discreta**, e não **contínua**.

Item errado.

Gabarito: ERRADO

12. (CESPE/ Perito Criminal Federal – 1997)

A tabela seguinte apresenta informações relativas ao recolhimento de imposto de renda e salário mensal de 16 funcionários de duas microempresas no ano de 1997.

	Empresa A		Empresa B	
Funcionário	Recolhimento de Imposto*	Salário Mensal**	Recolhimento de Imposto*	Salário Mensal**
1	1	2,0	0	2,2
2	0	1,0	0	1,1
3	0	2,0	0	2,2
4	0	2,0	0	2,2
5	1	5,0	0	5,5
6	1	10,0	0	11,0
7	0	2,0	0	2,2
8	1	2,0	0	2,2
9	1	2,0	1	2,2
10	1	5,0	1	5,5
11	1	5,0	1	5,5
12	1	5,0	1	5,5
13	0	50,0	1	55,0
14	0	1,0	1	1,1
15	0	1,0	1	1,1
16	1	5,0	1	5,5

*Recolhimento de Imposto: 0 = NÃO, 1 = SIM
**Salário Mensal: Número de salários mínimos

Com base na tabela, julgue o item

A média e a moda são medidas descritivas adequadas para avaliar a variável Recolhimento de Imposto.

Comentários

Gabarito: Errado

Sabendo que o gabarito foi "errado", vamos entender a interpretação da banca. A banca considerou que a variável "recolhimento de imposto" é qualitativa. Logo, a média não é uma medida adequada para sua avaliação. A média depende de operações aritméticas: somamos todos os dados e dividimos pela quantidade de observações. Assim, ela é adequada para variáveis quantitativas.

Já a moda, que só depende de uma contagem, pode sim ser utilizada para variáveis qualitativas.

Resposta: ERRADO

No entanto, cabe destacar que, quando estamos diante de uma variável qualitativa e, a cada atributo fazemos corresponder um número (no caso, os números são 0 e 1), e fazemos isso de tal forma que a atribuição de valores tenha utilidade e seja passível de interpretação, podemos sim tratar essa variável como se quantitativa fosse.

No caso da atribuição dos valores 0 e 1, damos origem a uma variável aleatória com distribuição de Bernoulli. Para essa variável é sim possível calcular uma média e uma variância. Caso essa fosse a interpretação da banca, o item seria "certo".

13. (CESPE/ Perito Criminal Federal – 1997)

A tabela seguinte apresenta informações relativas ao recolhimento de imposto de renda e salário mensal de 16 funcionários de duas microempresas no ano de 1997.

	Empresa A		Empresa B	
Funcionário	Recolhimento de Imposto*	Salário Mensal**	Recolhimento de Imposto*	Salário Mensal**
1	1	2,0	0	2,2
2	0	1,0	0	1,1
3	0	2,0	0	2,2
4	0	2,0	0	2,2
5	1	5,0	0	5,5
6	1	10,0	0	11,0
7	0	2,0	0	2,2
8	1	2,0	0	2,2
9	1	2,0	1	2,2
10	1	5,0	1	5,5
11	1	5,0	1	5,5
12	1	5,0	1	5,5
13	0	50,0	1	55,0
14	0	1,0	1	1,1
15	0	1,0	1	1,1
16	1	5,0	1	5,5

*Recolhimento de Imposto: 0 = NÃO, 1 = SIM
**Salário Mensal: Número de salários mínimos

Com base na tabela, julgue o item

A mediana é adequada para avaliar a variável Salário Mensal porque é resistente a valores extremos, e para avaliar a variável Recolhimento de Imposto porque é nominal.

Comentários

Para variáveis nominais não podemos determinar a mediana, eis que elas não são passíveis de ordenação.

Se não é possível ordenar as observações, não há como saber quem ocupa a posição central, ou seja, quem é a mediana.

Gabarito: ERRADO

14. (SMA-RJ (antiga FJG) / Fiscal de Rendas do Munícipio do RJ – 2002)

Os dados de um determinado estudo representam muitas variáveis para cada uma das pessoas que se submeteram ao estudo. Uma variável considerada qualitativa é a seguinte:

a) idade

b) altura

c) sexo

d) peso

Comentários

As variáveis "idade", "altura" e "peso" correspondem a quantidades. São todas variáveis quantitativas.
Por exemplo, podemos ter 45 anos de idade, 1,80 metros de altura e 87 kg. Tudo isso é número.

A variável "sexo" não é quantitativa. Ela tem duas realizações, correspondentes a dois diferentes atributos: "masculino" e "feminino". Trata-se de uma variável qualitativa.

Gabarito: C

15. (CESPE/ TRE – ES - 2011)

cargo	candidatos	candidatos aptos	eleitos
presidente da República	9	9	1
governador de estado	170	156	27
senador	272	234	54
deputado federal	6.021	5.058	513
deputado estadual/distrital	15.268	13.076	1.059
total	21.640	18.533	1.658

Internet: <www.tse.gov > (com adaptações).

Com base na tabela acima, referente às eleições de 2010, que apresenta a quantidade de candidatos para os cargos de presidente da República, governador de estado, senador, deputado federal e deputado estadual/distrital, bem como a quantidade de candidatos considerados aptos pela justiça eleitoral e o total de eleitos para cada cargo pretendido, julgue o item a seguir.
A variável "cargo" classifica-se como uma variável qualitativa ordinal.

Comentários

De fato é uma variável qualitativa, pois suas realizações são atributos, e não números.

Além disso, podemos ordenar estas variáveis segundo, por exemplo, o número de cargos disponíveis. É exatamente o que fez a tabela dada no enunciado, iniciando por presidente (um único cargo disponível), depois indo para governador (27 cargos) e assim por diante. Logo, é uma variável qualitativa e ordinal.

Gabarito: CERTO

16. (CESPE/ TER - ES – 2011)

quantidade de eleitores	quantidade de municípios
0 ⊢ 2.000	364
2.000 ⊢ 4.000	1.000
4.000 ⊢ 6.000	3.000
6.000 ⊢ 8.000	1.000
8.000 ⊢ 10.000	200
total	5.564

A tabela acima apresenta uma distribuição hipotética das quantidades de eleitores que não votaram no segundo turno da eleição para presidente da República bem como os números de municípios em que essas quantidades ocorreram. Com base nessa tabela, julgue o item segunite, relativo à análise exploratória de dados.

Na tabela de frequências, o uso de intervalos de classe permite concluir que a variável em questão é contínua.

Comentários

A quantidade de eleitores é discreta, pois assume valores num conjunto enumerável de pontos da reta real: 0, 1, 2, 3, 4...

Apenas agrupamos os dados em faixas, dando origem às classes. Isso, no entanto, não altera a natureza da variável, não a torna contínua.

Gabarito: ERRADO

17. (FUNIVERSA/ PC – GO - 2015)

A delegacia virtual é um serviço de registro de ocorrências, via Internet, por meio do site da Polícia Civil de Goiás. O objetivo é fornecer, para o cidadão, comodidade no registro de alguns tipos de ocorrências, otimizando os recursos e melhorando o atendimento. Depois de verificada a veracidade da informação pela Polícia Civil e autorizado pela autoridade policial, o usuário receberá, pelo correio eletrônico (*e-mail*) fornecido no registro, as informações sobre como obter uma cópia do Boletim de Ocorrência e qual delegacia de Polícia está responsável por apurar o fato. A qualquer momento, poderá ser consultado o andamento do seu registro, acessando a delegacia virtual. O serviço está disponível, sem interrupções, a qualquer dia e em qualquer horário. Pela delegacia virtual, o usuário poderá registrar as seguintes ocorrências: roubo/furto de objetos e(ou) valores, roubo/furto de celular, roubo/furto de documentos, perda ou extravio de documentos e(ou) objetos, desaparecimento de pessoa e acidente de trânsito sem vítimas.

Internet:<www.policiacivil.go.gov.br> (com adaptações).

Para a Estatística, a variável "tipo de ocorrência" é classificada como
a) qualitativa nominal.
b) quantitativa discreta.
c) quantitativa contínua.
d) qualitativa discreta.
e) qualitativa contínua.

Comentários

Uma variável é quantitativa quando seus resultados são numéricos. Exemplo: massa. A massa assume valores numéricos, tais como 79,84 kg. Isso é um número, logo, a variável correspondente é quantitativa.

Já podemos descartar todas as alternativas que falam em variável quantitativa (letras B e C).

Uma variável é qualitativa quando duas realizações não são numéricas, mas sim atributos. É justamente o caso da questão.

Uma variável qualitativa pode ser:

- nominal: ela é nominal quando suas observações não podem ser ordenadas segundo determinado critério. É justamente o caso da questão, não há uma ordem clara entre os diversos tipos de ocorrência.
- ordinal: a variável é ordinal quando há um critério claro segundo o qual as realizações podem ser ordenadas. Exemplo: nível de escolaridade (analfabeto, fundamental, médio, superior). Podemos ordenar essas realizações segundo o tempo de estudo.

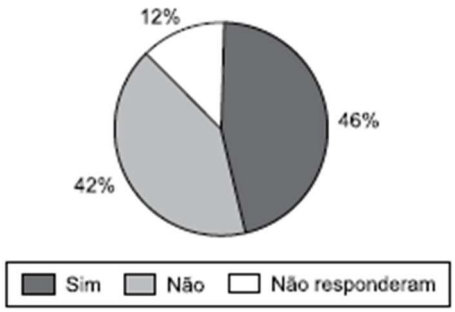
No nosso caso, estamos diante de uma variável qualitativa nominal. **Resposta: A**

Nota: há alternativas que falam em variável qualitativa discreta e qualitativa contínua. Isso não faz o menor sentido. As características “discreta” e “contínua” se referem apenas a variáveis quantitativas, e não qualitativas

Gabarito: A

18. (VUNESP/ Analista/Informações, Cultura e Desporto/Biblioteconomia - Pref SP - 2015)

Considere o gráfico a seguir, que representa as respostas à enquete “Você já leu pelo menos dois livros no primeiro semestre desse ano?”, pergunta essa feita a um grupo de 60 pessoas pesquisadas.



Levando-se em conta as medidas centrais média, moda e mediana, assim como a variável envolvida na estatística dessa enquete, é correto afirmar que pode(m) ser determinada(s), dentre essas três medidas,

a) a moda e a mediana, apenas.
b) a moda, apenas.
c) a média, apenas.
d) a média e a mediana, apenas.
e) a média, a moda e a mediana.

Comentários

A grandeza em estudo é do tipo "qualitativa", pois assume atributos ("sim", "não", "sem resposta") em vez de quantidades. Logo, não é possível calcular a média. Isto ocorre porque a média depende de cálculos: devemos somar todas as observações e depois dividir pelo total das frequências.

a) a moda e a mediana, apenas.
b) a moda, apenas.
~~**c)** a média, apenas.~~
~~**d)** a média e a mediana, apenas.~~
~~**e)** a média, a moda e a mediana.~~

A mediana depende de uma ordenação dos valores. Não há uma ordenação clara entre "sim", "não" e "sem resposta", não sendo possível calcular a mediana.

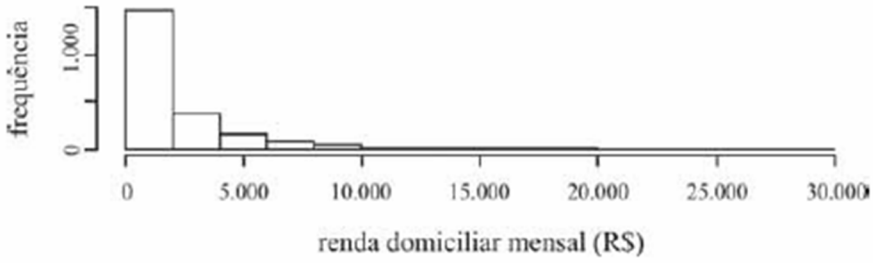
~~a) a moda e a mediana, apenas.~~
b) a moda, apenas.

Quanto à moda, basta determinar o termo com maior frequência. No caso, a resposta "sim" ocorreu mais vezes (46%). Portanto:

Moda = "sim"

Gabarito: B

19. (CESPE/ TELEBRAS – 2015)



Uma empresa coletou e armazenou em um banco de dados diversas informações sobre seus clientes, entre as quais estavam o valor da última fatura vencida e o pagamento ou não dessa fatura. Analisando essas informações, a empresa concluiu que 15% de seus clientes estavam inadimplentes. A empresa recolheu ainda dados como a unidade da Federação (UF) e o CEP da localidade em que estão os clientes. Do conjunto de todos os clientes, uma amostra aleatória simples constituída por 2.175 indivíduos prestou também informações sobre sua renda domiciliar mensal, o que gerou o histograma apresentado.

Com base nessas informações e no histograma, julgue o item a seguir.

O CEP da localidade dos clientes e o valor da última fatura vencida são variáveis quantitativas.

Comentários

O valor da última fatura de fato é quantitativo, indica uma quantidade, o montante de dinheiro gasto com telefonemas.

Já o CEP, embora seja representado por números, não indica quantidade alguma. Como exemplo, considere dois CEPs abaixo

12.300-460

74.790-800

Estes dois "valores" não correspondem a nenhuma quantidade. Nada está sendo medido ou contado. São só atributos de diferentes ruas. Não podemos dizer que a segunda rua tem "mais CEP" do que a primeira, nem qualquer outra coisa parecida com isso. Por isso, o CEP não é variável quantitativa.

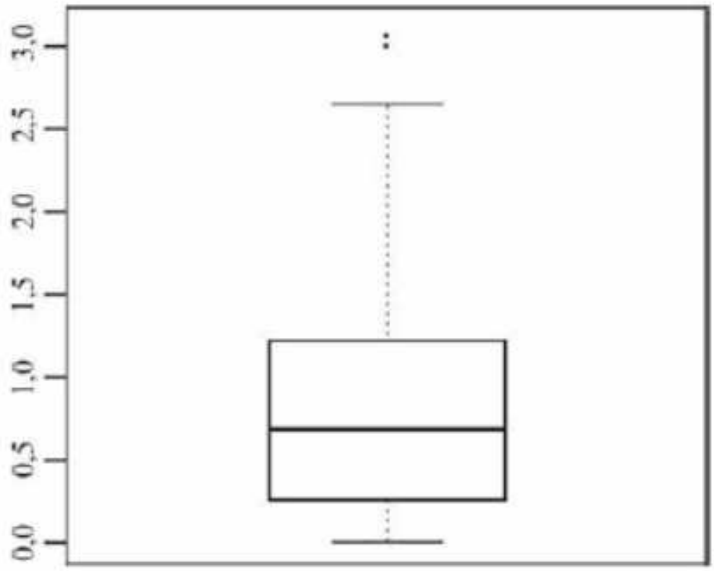
ITEM ERRADO.

Uma redação correta seria:

O CEP da localidade dos clientes é uma variável qualitativa; o valor da última fatura vencida é variável quantitativa.

Gabarito: ERRADO.

20. (CESPE/ TCE - PA – 2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho *X* permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

X representa uma variável qualitativa ordinal.

Comentários

Note que X assume valor máximo igual a 3,10 e valor mínimo igual a 0. Tudo isso é número, é quantidade, valor numérico. Logo, trata-se de uma variável quantitativa, e não qualitativa.

Gabarito: ERRADO.

21. (CESPE/ TCE - PA – 2016)

número diário de denúncias registradas (X)	frequência relativa
0	0,3
1	0,1
2	0,2
3	0,1
4	0,3
total	**1,0**

A tabela precedente apresenta a distribuição de frequências relativas da variável X, que representa o número diário de denúncias registradas na ouvidoria de determinada instituição pública. A partir das informações dessa tabela, julgue o item seguinte.

A variável X é do tipo qualitativo nominal.

Comentários

X corresponde à **quantidade** diária de denúncias, e por isso assume valores **numéricos** fruto de uma contagem. Logo, é uma variável quantitativa, e não qualitativa.

Gabarito: ERRADO.

ESCALAS DE MEDIDA

22. (CESPE / Antaq - 2009)

	variável	2003	2004	2005	2006	2007
exportação	X	40	46	50	52	54
importação	Y	20	21	22	24	27
total	X + Y	60	67	72	76	81

Considerando a tabela acima, que apresenta a movimentação anual de cargas no porto de Santos de 2003 a 2007, em milhões de toneladas/ ano e associa as quantidades de carga movimentadas para exportação e importação às variáveis X e Y, respectivamente, julgue o item subsequente.

X e Y são variáveis qualitativas em escala ordinal.

Comentários

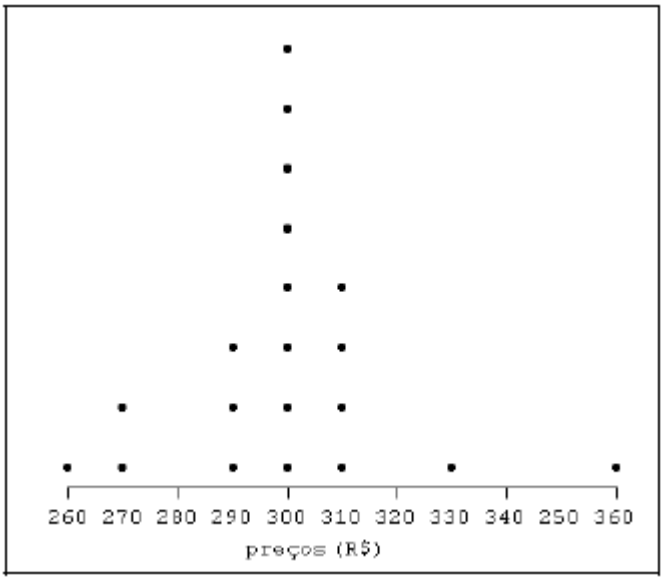
Uma variável é qualitativa ordinal quando representa um atributo que pode ser ordenado. Exemplo: nível de escolaridade (analfabeto / fundamental / médio / superior / pós-graduação).

No caso acima, temos as quantidades movimentadas. Ou seja, a resposta é um número, um valor, uma quantidade. Trata-se, portanto, de uma variável **quantitativa,** e não qualitativa, como foi afirmado.

Gabarito: ERRADO.

23. (CESPE / Antaq – 2005)

Em uma pesquisa realizada para verificar os preços cobrados em 2002 para o embarque de contêineres em um porto brasileiro, foram apurados 20 preços, que estão representados na figura abaixo (diagrama de pontos).



Com base nas informações do texto e da figura acima, julgue o item a seguir.

O preço cobrado para o embarque de contêineres em 2002 pode ser considerado uma variável discreta e qualitativa em escala nominal.

Comentários

O preço cobrado é uma variável **quantitativa,** e não qualitativa como foi dito na questão. É uma variável que indica a quantidade de unidades monetárias correspondentes ao serviço de embarque de contêineres.

Gabarito: ERRADO.

SÉRIES ESTATÍSTICAS

24. (CESPE/ Auditor Federal de Controle Externo – 2008)

Uma agência de desenvolvimento urbano divulgou os dados apresentados na tabela a seguir, acerca dos números de imóveis ofertados (*X*) e vendidos (*Y*) em determinado município, nos anos de 2005 a 2007.

ano	número de imóveis	
	ofertados (*X*)	vendidos (*Y*)
2005	1.500	100
2006	1.750	400
2007	2.000	700

Correio Braziliense, 29/4/2008, p. 17 (com adaptações).

Considerando as informações do texto, julgue os itens subsequentes.

A variável *X* forma uma série estatística denominada série temporal.

Comentários

Segundo Gilberto de Andrade Martins e Denis Donaire, no livro "Princípios de Estatística", séries estatísticas são as tabelas nas quais existe um critério distinto que as especifica e diferencia. Por este critério, as séries podem ser:

- cronológicas
- geográficas
- específicas
- conjugadas

Segundo os mesmos autores, a série temporal, também denominada série cronológica, tem como característica o fato de o tempo ser variável, enquanto que o local e o fato permanecem fixos. Como exemplo, citam uma pesquisa que traz a quantidade de unidades escolares de ensino de 1º grau (mesmo fato), no Brasil (mesmo local), ao longo dos anos de 1971 a 1974.

No caso desta questão temos exatamente isso: o fato é o mesmo (número de imóveis ofertados), em determinado município (mesmo local), e variamos apenas os instantes de tempo (anos de 2005 a 2006). Trata-se de uma série cronológica ou temporal.

Gabarito: CERTO

ESSA LEI TODO MUNDO CONHECE:
PIRATARIA É CRIME.

Mas é sempre bom revisar o porquê e como você pode ser prejudicado com essa prática.

1. 

Professor investe seu tempo para elaborar os cursos e o site os coloca à venda.

2. 

Pirata divulga ilicitamente (grupos de rateio), utilizando-se do anonimato, nomes falsos ou laranjas (geralmente o pirata se anuncia como formador de "grupos solidários" de rateio que não visam lucro).

3. 

Pirata cria alunos fake praticando falsidade ideológica, comprando cursos do site em nome de pessoas aleatórias (usando nome, CPF, endereço e telefone de terceiros sem autorização).

4. 

Pirata compra, muitas vezes, clonando cartões de crédito (por vezes o sistema anti-fraude não consegue identificar o golpe a tempo).

5. 

Pirata fere os Termos de Uso, adultera as aulas e retira a identificação dos arquivos PDF (justamente porque a atividade é ilegal e ele não quer que seus fakes sejam identificados).

6. 

Pirata revende as aulas protegidas por direitos autorais, praticando concorrência desleal e em flagrante desrespeito à Lei de Direitos Autorais (Lei 9.610/98).

7. 

Concurseiro(a) desinformado participa de rateio, achando que nada disso está acontecendo e esperando se tornar servidor público para exigir o cumprimento das leis.

8. 

O professor que elaborou o curso não ganha nada, o site não recebe nada, e a pessoa que praticou todos os ilícitos anteriores (pirata) fica com o lucro.



Deixando de lado esse mar de sujeira, aproveitamos para agradecer a todos que adquirem os cursos honestamente e permitem que o site continue existindo.