



ESTATÍSTICA

Curso de Formação e Graduação de Sargentos (Ensino Superior)

ISBN: 978-65-994652-1-5



© Escola de Sargentos das Armas – 2022
1ª edição 2020.
2ª edição 2022.

ESTATÍSTICA

Grupo de trabalho responsável pela elaboração do Livro de Estatística do Curso de Formação e Graduação de Sargentos (CFGS) da Escola de Sargentos das Armas (ESA):

Órgão Gestor:

DETMil

General de Brigada Alexandre Oliveira **CANTANHEDE** Lago
Diretor de Educação Técnica Militar

Órgão Elaborador:

ESA

General de Brigada Reinaldo **SALGADO** Beato
Comandante da Escola de Sargentos das Armas

Órgão Executor:

ESA

Coronel **JADILSON** Tadeu da Silva dos Santos
Chefe da Divisão de Ensino

Supervisão Geral

Capitão QCO **ARACELI** Paula Naves
Chefe da Seção de Disciplinas Acadêmicas
Supervisão Geral

2º Tenente Priscila da Silva **ROCHA** Pinto
Adjunta da Seção de Disciplinas Acadêmicas
Revisão

2º Tenente OTT **DIONE** Aparecido Ferreira da Silva
Adjunto da Seção de Disciplinas Acadêmicas
Produção, Revisão e Diagramação geral

2º Tenente OTT **DINALVA** Ferreira da Silva
Adjunta da Seção de Disciplinas Acadêmicas
Produção, Revisão e Diagramação geral

2º Tenente OTT **THAMARA** Marques Rodrigues
Adjunta da Seção de Disciplinas Acadêmicas
Produção, Revisão e Diagramação geral

1º Sargento Inf Eduardo **LUINI** da Silva
Auxiliar da Seção de Disciplinas Acadêmicas
Produção, Revisão e Diagramação geral

Soldado Paulo Henrique **ARAÚJO CÉZAR** Leigo
Diagramação da capa



Ficha catalográfica:

Elaborada por Fabiane Neves Fonseca Andrade – Bibliotecária - CRB-6/3200.

Escola de Sargentos das Armas

E74e

Estatística / Dinalva Ferreira da Silva, Dione Aparecido Ferreira da Silva, Eduardo Luini da Silva, Thamara Marques Rodrigues. 2. ed. Três Corações: Escola de Sargentos das Armas, 2022.
114p.: il.

ISBN: 978-65-994652-1-5

1. Probabilidade e matemática aplicada. 2. Análises numéricas.
3. Análise de dados. 4. Estatística descritiva. 5. Inferência Estatística.

CDD: 519.5

CDD: 519.2



PREFÁCIO

O Decreto nº 9.171, de 17 de outubro de 2017, do Presidente da República, alterou o Regulamento da Lei de Ensino do Exército, a partir do qual a Portaria nº 504-EME, de 8 de Dezembro de 2017, emitiu as Diretrizes para a Equivalência de Estudos dos Cursos Destinados aos Sargentos e Subtenentes e a Implantação do Curso de Formação de Sargentos no Nível Superior de Tecnologia, dentre outras providências (EB20-D-01.059).

A Portaria nº 268-DECEX, de 12 de dezembro de 2018, aprovou as Instruções Reguladoras para a Execução e a Equivalência de Nível de Educação dos Cursos destinados aos Sargentos e Subtenentes (EB60-IR57.010) – 3ª Edição, e determinou que, na elaboração do novo currículo, deveria ser considerada a inclusão de disciplinas acadêmicas, notadamente, a disciplina de Raciocínio Lógico e Estatístico.

O parágrafo 5º, da Seção IV (Das Diretrizes Curriculares Específicas do Exército para os CFGS Tecnólogos), do Capítulo II, da Portaria 504-EME, intitulado “Dos Cursos de Formação de Sargentos”, determinou que, na elaboração do novo currículo, deveria ser considerada a inclusão de Raciocínio Lógico e Estatístico.

Dessa forma, o ensino de Raciocínio Lógico e Estatístico ganhou o merecido lugar na formação do líder de pequenas frações, uma vez que é capaz de organizar os pensamentos e assim obter bons resultados em suas atividades cotidianas.

O ensino de Raciocínio Lógico e Estatístico marcará um novo paradigma na formação dos sargentos enquanto líderes de pequenas frações, que poderão, por meio de abstrações, interpretações das informações e relações entre o que foi apresentado e os conhecimentos adquiridos, desenvolver a liderança que se espera num Exército em transformação, no qual além de executores perfeitos das ordens emanadas do Comando, devem também atuar como resolvedores de problemas e partícipes de decisões no amplo espectro dos conflitos.

Cap QCO **ARACELI** Paula Naves
Chefe da Seção de Disciplinas Acadêmicas – ESA



LISTA DE TABELAS

Tabela 1 - Apresentação genérica de uma tabela	27
Tabela 2 - Produção Café Brasil 1991 a 1995	28
Tabela 3 - Composição de uma tabela	29
Tabela 4 - Tabela Primitiva	30
Tabela 5 - Tabela de dados em Rol	31
Tabela 6 - Distribuição de Frequência 1	33
Tabela 7 - Distribuição de Frequência 2	34
Tabela 8 - Distribuição de Frequência 3	35
Tabela 9 - Distribuição de Frequência 4	36
Tabela 10 - Distribuição de Frequência 5	41
Tabela 11 - Distribuição de Frequência 6	42



LISTA DE QUADROS

Quadro 1 - Classificação das variáveis.....	13
Quadro 2 - Exemplos de variáveis.....	13



LISTA DE ABREVIATURAS E SIGLAS

ABNT - Associação Brasileira de Normas Técnicas
AIE - Agência Internacional de Energia
BCSv - Batalhão de Comando e Serviço
Cia - Companhia
C Inf - Curso de Infantaria
C Cav - Curso de Cavalaria
C Art - Curso de Artilharia
C Eng - Curso de Engenharia
C Com - Curso de Comunicações
CFGS - Curso de Formação e Graduação de Sargentos
EME - Estado-Maior do Exército
ENADE - Exame Nacional de Desempenho dos Estudantes
ESA - Escola de Sargentos das Armas
ESAG - Escola Superior de Administração e Gerência
IBGE - Instituto Brasileiro de Geografia e Estatística
IMBEL - Indústria de Material Bélico do Brasil
IMC - Índice de Massa Corpórea
NBR - Normas Brasileiras
PE - Polícia do Exército
TAF - Teste de Aptidão Física
UFMT - Universidade Federal do Mato Grosso
UD - Unidade Didática



SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO À ESTATÍSTICA	9
1.1 O que é Estatística?	9
1.2 A Estatística Descritiva e Indutiva	9
1.3 Importância e Aplicações.....	10
CAPÍTULO 2 – CONCEITOS BÁSICOS DE ESTATÍSTICA	11
CAPÍTULO 3 – FASES DO TRABALHO ESTATÍSTICO	21
3.1 Definição do Problema.....	21
3.2 Planejamento	22
3.3 Coleta de Dados	22
3.3.1 Espécies de Dados.....	23
3.3.2 Formas de Coleta de Dados	23
3.4 Crítica dos Dados	24
3.5 Apuração dos Dados.....	24
3.6 Exposição ou Apresentação dos dados	25
3.7 Análise e Interpretação dos resultados.....	25
CAPÍTULO 4 – TABELAS	27
4.1 Tabela	27
4.1.1 Composição da Tabela:.....	28
4.1.2 Sinais Convencionais.....	29
4.2 Distribuição de Frequências	29
4.2.1 Conceitos importantes.....	30
4.2.2 Distribuição de Frequências Sem Intervalos de Classe.....	31
4.2.3 Distribuição de Frequências Com Intervalos de Classe.....	36
CAPÍTULO 5 – REPRESENTAÇÕES GRÁFICAS	49
4.3 Gráficos Estatísticos	49
4.3.1 Tipos de Gráficos.....	49
CAPÍTULO 6 – MEDIDAS DE POSIÇÃO	61
5.1 Média Aritmética (\bar{x})	61
5.2 Média Aritmética Ponderada (MP).....	62
5.3 Mediana (Md)	62
5.4 Moda (Mo).....	64
5.5 Média Aritmética, Moda e Mediana a partir de uma Distribuição de Frequências Sem Intervalo de Classe	65
5.6 Média Aritmética, Moda e Mediana a partir de uma Distribuição de Frequências Com Intervalo de Classe	67



CAPÍTULO 7 – MEDIDAS DE DISPERSÃO.....	87
6.1 Desvios.....	87
6.2 Variância.....	88
6.3 Desvio Padrão.....	90
APÊNDICE 1 – MEDIDAS DE SEPARATRIZES.....	97
1.1 Quartis (Q_i).....	97
1.2 Decis (D_i).....	97
1.3 Centis (C_i).....	98
1.4 Procedimentos para os cálculos dos Quartis para dados não agrupados.....	98
1.5 Procedimentos para os cálculos dos Quartis para dados agrupados sem intervalo de classe.....	101
1.6 Procedimentos para os cálculos dos Quartis, Decis e Centis para dados agrupados com intervalo de classe.....	104
REFERÊNCIAS.....	114



CAPÍTULO 1 – INTRODUÇÃO À ESTATÍSTICA

1.1 O que é Estatística?

Segundo Freitas, Rennó e Júnior (2003, p. 11), “Estatística é a ciência que investiga os processos de obtenção, organização e análise de dados sobre uma população, e os métodos de tirar conclusões ou fazer predições com base nesses dados”.

A palavra “estatística” significa ‘análise de dados’. Como ciência, surgiu milênios antes de Cristo, sendo, no início, uma simples compilação de números, pois contar, enumerar e recensear, sempre foi uma preocupação de todas as culturas, de modo a sistematizar a coleta e a análise das informações e dos dados necessários a algum tipo de estudo.

Essa utilização da estatística como ferramenta para a tomada de decisões é tão antiga quanto a civilização, desde quando o homem sentiu necessidade de se reunir em grupos, seja para a gestão de alimentos para grandes populações, seja para a formação de exércitos para a defesa de cidades e para a quantificação da arrecadação de impostos, dentre outros.

Atualmente, estamos imersos numa sociedade que produz uma quantidade imensa de dados, os quais são registrados e armazenados em algum ambiente específico, para posterior análise e sistematização.

Desse modo, o estudo da estatística é um fator preponderante e indispensável na tarefa de analisar esta enorme quantidade de dados, com a finalidade de que um estudioso sobre o assunto possa entender e tomar decisões baseadas nos resultados.

1.2 A Estatística Descritiva e Indutiva

De acordo com Santos (2016, p.3), a Estatística está compreendida em duas partes:

a. Estatística Descritiva: reúne um conjunto de técnicas para sumarizar os dados (tabelas, gráficos) e medidas descritivas (medidas de tendência central ou dispersão), que permitem tirar muitas informações contidas nos dados. Sua finalidade é tornar os dados mais fáceis de serem entendidos, relatados e discutidos;

b. Estatística Inferencial ou Indutiva: é constituída por um conjunto de Métodos para a tomada de decisões, a fim de produzir afirmações sobre uma dada característica da população, na qual estamos interessados, a partir de informações colhidas de uma parte dessa população.



1.3 Importância e Aplicações

Conforme Rosa (2021, p.5), “[...] a Estatística lida com a coleta, a apresentação, a análise e o uso dos dados para tomar decisões, resolver problemas e planejar produtos e processos[...]”, de modo que os dados devem ser cuidadosamente coletados (observados), devidamente conhecidos e utilizados para analisar e interpretar a sua variabilidade, a fim de possibilitar uma correta resposta à hipótese em estudo.

Quanto às áreas de aplicação, pode-se citar algumas: social, saúde, tecnológica industrial, financeira e bancária, pesquisas científicas, etc, dentre as quais merecem destaque:

- ✓ na área médica - os estudos estatísticos fornecem a metodologia adequada que possibilita decidir sobre a eficiência de uma nova vacina/tratamento no combate à determinada doença, permitindo identificar situações críticas e, conseqüentemente, atuar em seu controle;
- ✓ na área industrial - a estatística pode atuar desde a fase de implantação de uma fábrica, até a avaliação das necessidades de expansão industrial; na pesquisa e desenvolvimento de técnicas, produtos e equipamentos; nos testes de produtos; no controle de qualidade e quantidade, dentre outros;
- ✓ na área demográfica - o estatístico estuda a evolução e as características da população, analisando taxas de natalidade e mortalidade, bem como fluxos migratórios, níveis e padrões para testes;
- ✓ na área educacional - o ensino da Estatística está atrelado à pesquisa científica, mediante o desenvolvimento de novas metodologias de análise estatística para os mais variados problemas práticos e teóricos, seja promovendo suporte científico para a tomada de decisões, seja auxiliando na escolha da metodologia científica a ser adotada, no planejamento da pesquisa, na escolha qualificada dos dados, na análise das respostas, etc.



CAPÍTULO 2 – CONCEITOS BÁSICOS DE ESTATÍSTICA

Bussab e Morettin (2010, p. 264) definem que “População é o conjunto de todos os elementos ou resultados sob investigação. Amostra é qualquer subconjunto da população.” Esses conceitos básicos são necessários para o desenvolvimento da Inferência Estatística. Sendo assim, passamos a melhor defini-los:

a. População Estatística e Amostra

População é entendida como um conjunto de objetos ou itens com alguma característica ou propriedade comum mensurável, ordenável ou comparável, a partir dos quais formam-se um conjunto de elementos que constituem o universo passível de observação.

Ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos **população estatística ou universo estatístico**.

Em alguns casos, seja por impossibilidade ou inviabilidade econômica ou temporal, há a ocorrência da necessidade de limitar as observações a apenas uma parte da população. A essa parte proveniente da população chamamos de Amostra.

Uma amostra é um subconjunto finito de uma população.

b. Indivíduo ou Objeto

Cada elemento que compõe a amostra é um **indivíduo** ou **objeto**.

Exemplo: intenções de voto numa eleição.

Nesse caso, os indivíduos da pesquisa são pessoas.

Exemplo: quando se consideram algumas marcas de fardamentos para testar a durabilidade.

Nesse outro caso, cada marca é um objeto da pesquisa.



c. Variável

É necessário, inicialmente, que se defina qual(is) a(s) característica(s) dos elementos que deverá(ão) ser verificada(s), ou seja, não se trabalha estatisticamente com os elementos existentes, mas com alguma(s) característica(s) desses elementos.

Quanto ao tipo, as variáveis podem ser classificadas em:

a) **Quantitativa**: quando seus valores são expressos em números. Subdividem-se em:

- ✓ **Contínua**: a variável assume, teoricamente, qualquer valor entre dois limites.

Exemplo: quantidade de leite consumido no café da manhã pelos alunos do CFGS.

- ✓ **Discreta**: a variável assume valores pertencentes a um conjunto enumerável. Os valores são obtidos pela simples contagem.

Exemplo: número de fuzis 7,62mm constantes da Reserva de Armamento da Cia PE/BCSv.

b) **Qualitativa**: quando seus valores são expressos por atributos, classificações ou registros não numéricos: sexo, cor da pele etc. Subdividem-se em:

- ✓ **Ordinal**: seguem um certo tipo de precedência, de escala crescente ou decrescente.

Exemplo: grau de queimadura (1º, 2º, 3º grau), pressão sanguínea (baixa, normal, alta).

- ✓ **Nominal**: expressa pelos atributos, com possibilidade de nomeação.

Exemplo: sexo, raça, região, grupo sanguíneo.



Resumindo, tem-se:

Quadro 1 – Classificação das variáveis.

Variável	Qualitativa	nominal
		ordinal
	Quantitativa	discreta
		contínua

Fonte: próprio autor.

Quadro 2 - Exemplos de variáveis.

Tipos de Variáveis			
Qualitativas		Quantitativas	
Valores expressos por atributos		Valores expressos por números	
Nominal	Ordinal	Discreta	Contínua
- Profissão - Sexo - Religião	- Escolaridade - Estágio de doença - Classe social	- N° de filhos - N° de acesso ao AVA	- Altura - Peso - Salário
Não permite uma ordenação	Permite uma ordenação	Números inteiros	Números “contínuos”

Fonte: próprio autor.



EXERCÍCIOS

Capítulos 1 e 2 – Introdução e Conceitos de Estatística

1. Classifique as variáveis em qualitativas nominais, qualitativas ordinais, quantitativas discretas ou quantitativas contínuas:

- a) número de militares do 1º Pelotão do C Inf.
- b) altura dos candidatos convocados no Concurso Público ao CFGS 2022/2023.
- c) cor do cabelo de uma militar (preto, ruivo, loiro...).
- d) número de defeitos observados num equipamento eletrônico do C Com.
- e) tipos de defeitos observados num fuzil 7,62mm.
- f) série/ano em que um aluno estuda nos Ensinos Fundamental e Médio.

Respostas:

- a) Quantitativa discreta
- b) Quantitativa contínua
- c) Qualitativa nominal
- d) Quantitativa discreta
- e) Qualitativa nominal
- f) Qualitativa ordinal

2. Um batalhão possui em seus efetivo 300 militares e fez uma pesquisa sobre a preferência de compra de agasalhos/abrigos em relação à “cor” (branco, vermelho ou azul), “preço”, “número de faixas na calça” (duas ou quatro) e “estado de conservação” (novo ou usado).

Foram consultados 210 militares.

Diante dessas informações, responda:

- a) Qual é o universo estatístico e qual é a amostra dessa pesquisa?
- b) Quais são as variáveis e qual a classificação de cada uma quanto ao tipo?
- c) Quais os possíveis valores da variável “cor” nessa pesquisa?



Respostas:

a) 300 militares; 210 militares, respectivamente.

b)

Cor (qualitativa nominal);

Preço (quantitativa contínua);

Número de faixas na calça (quantitativa discreta); e

Estado de Conservação (qualitativa ordinal).

c) Branco, vermelho ou azul.

3. Assinale V (verdadeiro) ou F (falso) para as alternativas abaixo:

a. () Variável qualitativa nominal é aquela em que observa-se um certo tipo de precedência, de escala crescente ou decrescente.

b. () Em um estudo sobre o número de dependentes dos alunos do C Art, foram feitas 30 entrevistas com parte desses alunos. Sendo assim, a amostra é composta por todos os alunos do C Art e a variável estudada é o número de dependentes.

c. () Cada elemento que compõe a amostra é um indivíduo ou objeto.

d. () Ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos Amostra.

Respostas:

F, F, V, F

Justificativa:

a. Falsa. O correto é: variável qualitativa ORDINAL é aquela em que observa-se um certo tipo de precedência, de escala crescente ou decrescente;

b. Falsa. O correto é: em um estudo sobre o número de dependentes dos alunos do C Art, foram feitas 30 entrevistas com parte desses alunos. Sendo assim, a POPULAÇÃO é composta por todos os alunos do C Art e a variável estudada é o número de dependentes.

c. Verdadeira.

d. Falsa. O correto é: ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos população estatística.



4. A partir das informações fornecidas por um alfaiate da cidade de Três Corações-MG, referente ao enxoval dos alunos do CFGS, que incluem dados básicos sobre os clientes e a avaliação dos produtos por estes (que variam de 1, péssimo, a 5, ótimo), classifique as variáveis em qualitativas (nominal ou ordinal) ou quantitativas (discreta ou contínua).

Alunos	Curso do CFGS	Altura (m)	Qtde de Materiais Comprados	Nº Idt Militar	Valor a pagar	Avaliação do Produto
Bruno	Infantaria	1,85	1	257561336	100,00	3
Marcelo	Artilharia	1,76	4	540902661	400,00	2
João	Engenharia	1,93	2	267832901	200,00	5
Marcos	Comunicações	1,82	3	640383772	300,00	3
Pedro	Cavalaria	1,69	5	123456789	500,00	4

Respostas:

Alunos: qualitativa nominal.

Curso do CFGS: qualitativa nominal.

Altura (m): quantitativa contínua.

Quantidade de Materiais Comprados: quantitativa discreta.

Nº Identidade Militar: qualitativa nominal.

Valor a pagar: quantitativa contínua.

Avaliação do Produto: qualitativa ordinal.

5. (UFU-Técnico em Estatística – 2019)

Considere as seguintes variáveis:

I. Tamanho de um objeto (pequeno, médio ou grande).

II. Volume de água em um rio.

III. Número de clientes numa fila.

IV. Número da seção de votação.

V. Comprimento de um inseto.

VI. Classe Social.

Com relação à classificação dos dados requeridos como variáveis de pesquisa, é correto afirmar que:



- a) as variáveis I, IV e VI são qualitativas.
- b) as variáveis III e V são quantitativas contínuas.
- c) as variáveis II e III são quantitativas discretas.
- d) a variável IV é qualitativa ordinal.

Resposta: letra a)

Justificativa:

- I. Tamanho de um objeto (pequeno, médio ou grande): qualitativa ordinal
- ii. volume de água em um rio: quantitativa contínua
- III. Número de clientes numa fila: quantitativa discreta
- IV. Número da seção de votação: qualitativa nominal
- V. Comprimento de um inseto: quantitativa contínua
- VI. Classe Social: qualitativa ordinal

6. (Prefeitura Municipal de Penalva - Assistente Administrativo (IMA - 2017))

Assinale a alternativa que apresenta o conceito de variável quantitativa discreta:

- a) É aquela que expressa o valor de uma contagem, por exemplo, idade, quantidade de televisores numa casa, quantidade de habitantes de uma cidade.
- b) É aquela que separa os indivíduos em classes com uma determinada ordem, por exemplo, nível de escolaridade: fundamental, médio e superior.
- c) É aquela que expressa uma medida como um valor real, por exemplo, peso e altura.
- d) É aquela que separa os indivíduos em classes, porém não é possível estabelecer uma ordem, por exemplo, sexo (masculino e feminino) e esporte praticado (futebol, basquete, ciclismo...).

Resposta: letra a)

Justificativa:

- a) Quantitativa discreta
 - b) Qualitativa ordinal
 - c) Quantitativa contínua
 - d) Qualitativa nominal
-



7. (Câmara Municipal de Itatiba/SP – Analista de Recursos Humanos – VUNESP)

Em estatística, a técnica que nos permite fazer inferências sobre uma população, a partir da análise de uma parte dela, denomina-se:

- a) Dedução.
- b) Amostragem.
- c) Probabilidade.
- d) Descrição.
- e) Extração.

Resposta: letra b)

Justificativa: Pois a Estatística Indutiva ou Inferencial tem por objetivo tirar conclusões sobre as populações, com base em resultados verificados em amostras retiradas dessa população.

Logo, a técnica é AMOSTRAGEM.

8. (EBSERH – Analista Administrativo – Estatística (HE-U FSCAR) – INSTITUTO AOCP)

Que parte da estatística se preocupa apenas em descrever determinada característica da população?

- a) Regressão estatística.
- b) Estatística contínua.
- c) Estatística descritiva.
- d) Estatística amostral.
- e) Estatística inferencial.

Resposta: letra c)

Justificativa: A Estatística Descritiva é relacionada à coleta, organização e descrição dos dados.



9. Considere as seguintes afirmações a respeito de certas variáveis, observadas numa Organização Militar:

- a) Um projétil de pistola possui 9,00 mm de diâmetro.
- b) Tipos de defeitos observados num armamento da Armaria do C Art.
- c) O peso dos alunos avaliados por ocasião da medição do IMC para o Teste de Aptidão Física.
- d) O nível do Rio Verde está 2 m acima do alerta de enchente.

Assinale a alternativa correta:

- a) Todas são variáveis de mesma natureza, exceto em (a).
- b) Todas são variáveis de mesma natureza exceto em (b).
- c) Todas são variáveis de mesma natureza exceto em (c).
- d) Todas são variáveis de mesma natureza exceto em (d).

Resposta: letra b)

Justificativa:

- a) Quantitativa contínua.
- b) Qualitativa nominal.
- c) Quantitativa contínua.
- d) Quantitativa contínua.

10. (CESGRANRIO)

No questionário socioeconômico que faz parte integrante do ENADE há questões que abordam as seguintes informações sobre o aluno:

- I – Unidade da Federação em que nasceu.
- II – Número de irmãos.
- III – Faixa de renda mensal da família.
- IV – Estado civil.
- V – Horas por semana de dedicação aos estudos.



São qualitativas APENAS as variáveis:

- a) I e III.
- b) I e IV.
- c) I, IV e V.
- d) II, III e V.
- e) I, II, IV e V.

Resposta: letra b)

Justificativa:

Vamos classificar cada uma das variáveis abaixo, conforme os conceitos vistos acima:

I – Unidade da Federação em que nasceu (qualitativa).

II – Número de irmãos (quantitativa).

III – Faixa de renda mensal da família (quantitativa).

IV – Estado civil (qualitativa).

V – Horas por semana de dedicação aos estudos (quantitativa).

Dessa forma, tem-se como dados qualitativos as afirmativas I e IV.



CAPÍTULO 3 – FASES DO TRABALHO ESTATÍSTICO

Segundo o Curso de Estatística do Centro Científico Conhecer, por ocasião da realização de estudo estatístico completo, existem diversas fases do trabalho que devem ser desenvolvidas para se chegar aos resultados finais do estudo. Essas etapas são chamadas Fases do Trabalho Estatístico e são de âmbito da Estatística Descritiva, estudada no Cap 1.

As fases do método estatístico serão especificadas sucintamente abaixo:

- ✓ Definição do problema;
- ✓ Planejamento;
- ✓ Coleta dos dados;
- ✓ Crítica dos dados;
- ✓ Apuração dos dados;
- ✓ Exposição ou Apresentação dos dados; e
- ✓ Análise e Interpretação dos resultados

3.1 Definição do Problema

É a primeira fase do trabalho estatístico e consiste na formulação correta do problema a ser estudado.

Indagação: qual será a pesquisa?

É a fase em que se define o porquê de usar o método estatístico e qual será a sua pesquisa, com a finalidade de delimitá-la e definir corretamente o problema.

Exemplo: Um fabricante de artigos militares e de segurança, que deseja lançar um novo modelo de coldre de pistola, poderia estar interessado em um estudo sobre as características dos atiradores deste tipo de armamento. Não havendo estudos semelhantes, ele deverá formular o problema com base em sua própria experiência. Uma lista de fatores relevantes deverá resultar dessa investigação preliminar: número de unidades adquiridas por ano, material utilizado na fabricação, tipo de uso (velado ou ostensivo), especificações do tipo de armamento em questão, de acordo com as especificações do fabricante (Taurus, IMBEL, GLOCK, etc.), e assim por diante.



3.2 Planejamento

Após a definição do problema, vem a fase do Planejamento.

Indagações: Como levantar dados? Que dados deverão ser obtidos? Como se deve obtê-los? Quem pode patrocinar minha pesquisa?

Esta fase consiste em determinar o procedimento necessário para resolver o problema e, em especial, como levantar informações sobre o assunto objeto do estudo, ou seja, como fazer o levantamento dos dados.

A maior preocupação consiste na escolha das perguntas, bem como na sua correta formulação, pois é nesta fase que será escolhido o tipo de levantamento a ser utilizado.

Também são tratados nesta fase o Cronograma das atividades, através do qual são fixados os prazos, os custos envolvidos, o exame das informações disponíveis, o delineamento da Amostra, e assim por diante.

3.3 Coleta de Dados

Este é o terceiro passo, que é essencialmente operacional.

Compreende a coleta das informações propriamente ditas, referindo-se à obtenção, reunião e registro sistemático de dados, com um objetivo determinado. Para tanto, é necessário considerarmos as *Espécies de Dados* e as *Formas de Coleta de Dados*.

Exemplo: Uma Organização Militar pode valer-se de diversas fontes ao utilizar os dados em seus trabalhos estatísticos. Assim, os seus registros de pessoal podem conter muitas informações úteis para diversos fins, uma vez que se encontram arquivados em seu Suporte Documental. Entretanto, pode ocorrer que os registros da própria organização não proporcionem toda a informação necessária para resolver determinado problema, podendo ter alguma outra organização militar que reúna os dados e os distribua em forma de publicações, orientações normativas, dentre outras. Dessa forma, os registros internos não serão suficientes nesta situação, necessitando a Organização Militar de recorrer a fontes externas para obter as informações desejadas. Nessa senda, caso a OM necessitasse de saber a opinião dos seus militares sobre algum aspecto particular de sua atividade, um exame direto sobre ele seria recomendável.



3.3.1 Espécies de Dados

Sendo assim, pode-se inferir que é possível distinguir dois tipos de fontes externas, as quais dão origem a duas espécies de dados: os dados primários e os dados secundários.

a. Dados primários: são primários quando são publicados ou comunicados pela própria pessoa ou organização que os tenha recolhido, isto é, são obtidos a partir da realização de pesquisas diretas ou pela observação experimental.

Exemplos: questionários, pesquisas de campo.

b. Dados secundários: quando são publicados ou comunicados por outra organização, ou seja, são elaborados a partir da existência de dados primários já disponibilizados em outros bancos de dados, livros, periódicos, etc.

Exemplo: Censo demográfico.

3.3.2 Formas de Coleta de Dados

Ainda no que tange à Coleta de Dados, ela pode ser realizada de duas maneiras: direta ou indiretamente. Observe:

a. Coleta direta – é aquela obtida diretamente da fonte, como no caso da Organização Militar que realiza uma pesquisa para saber a opinião dos seus militares quanto a algum aspecto da caserna. Há três tipos de coleta direta:

- ✓ Coleta contínua – é aquela no qual os dados são obtidos ininterruptamente, automaticamente e durante um determinado período.

Exemplo: os registros de nascimento, de casamento e de óbito.

- ✓ Coleta periódica – é aquela realizada em períodos curtos e determinados.

Exemplo: recenseamento demográfico, o qual é feito de 10 em 10 anos.



- ✓ Coleta ocasional – é aquela onde dados são colhidos esporadicamente, atendendo a uma conjuntura qualquer ou a uma determinada emergência.

Exemplo: coleta de dados em um surto epidêmico

b. Coleta indireta – é aquela coleta realizada a partir dos elementos conseguidos pela coleta direta. É permeada por deduções e conjecturas, podendo ser realizada de 4 maneiras:

- ✓ por analogia – quando o conhecimento de um fenômeno é induzido a partir de outro que com ele guarda relações de casualidade;
- ✓ por proporcionalização – quando o conhecimento de um fato se induz das condições quantitativas de uma parte dele;
- ✓ por indícios – quando são escolhidos fenômenos sintomáticos para discutir um aspecto geral da vida social; e
- ✓ por avaliação – quando, através de informações fidedignas ou estimativas cadastrais, se presume o estado quantitativo de um fenômeno.

3.4 Crítica dos Dados

Após a obtenção dos dados, eles são analisados para verificar a ocorrência de erros, a fim de que não ocorram distorções que interfiram no resultado final. Em seguida, os dados devem ser cuidadosamente criticados, à procura de possíveis falhas e imperfeições, a fim de não incorrer erros grosseiros ou de certo vulto, que possam influir sensivelmente nos resultados.

A Crítica pode ocorrer de 2 formas:

- a. Externa: quando visa às causas dos erros por parte do informante, por distração ou má interpretação das perguntas que lhe foram feitas; e
- b. Interna: quando visa observar os elementos originais dos dados da coleta.

3.5 Apuração dos Dados

A Apuração é realizada com a finalidade de proporcionar um tratamento prévio aos dados obtidos, a fim de torná-los mais expressivos. É também chamada de *Sumarização* e consiste em resumir os dados, por meio de sua contagem e agrupamento. Consiste num trabalho



de condensação e de tabulação dos dados, que chegam ao analista de forma desorganizada, tornando impossível a tarefa de apreender todo o significado pela simples leitura.

As formas de realizar a apuração são as mais diversas: manual, mecânica, eletromecânica e eletrônica. Veja:

- a. Manual: sem o auxílio de máquinas mecânicas;
- b. Mecânica: com o auxílio de máquinas mecânicas;
- c. Eletromecânica: com o auxílio de máquinas distintas da anterior, pelo fato de suas engrenagens internas serem movidas a energia elétrica; e
- d. Eletrônica: Essas máquinas efetuam as operações através de impulsos elétricos, dispensando qualquer tipo de engrenagem.

É por meio da Apuração que se tem a oportunidade de condensar os dados, de modo a obter um conjunto compacto de números, o qual possibilita distinguir melhor o comportamento do fenômeno na sua totalidade. A apuração nada mais é que a soma e o processamento dos dados obtidos.

3.6 Exposição ou Apresentação dos dados

Há duas formas de apresentação dos dados:

- a. Apresentação tabular: é uma apresentação numérica dos dados, em forma de tabela. Consiste em dispor os dados em linhas e colunas distribuídas de modo ordenado, seguindo algumas regras práticas adotadas pelos diversos sistemas estatísticos.
- b. Apresentação gráfica: é uma apresentação geométrica. Possui a vantagem de facilitar ao analista obter uma visão rápida, fácil e clara do fenômeno e sua variação.

3.7 Análise e Interpretação dos resultados

É a última fase do trabalho estatístico e é a parte mais importante e sensível de todo o processo, cuja finalidade é proporcionar as inferências sobre as conclusões, as quais auxiliarão o pesquisador a resolver seu problema. Ela está essencialmente ligada ao cálculo de medidas, cuja finalidade principal é descrever o fenômeno. Dessa forma, o significado exato dos valores obtidos, por meio do cálculo das várias medidas estatísticas, deve ser bem interpretado.



Quadro-Resumo das Fases do Método Estatístico

Consoante Fonseca (2015, p. 17), pode-se resumir as Fases do Método Estatístico, de acordo com o seguinte quadro:

Quadro 3: Quadro-Resumo das Fases do Método Estatístico.

Etapa	Descrição
Definição	Consiste na formulação correta do problema a ser estudado.
Planejamento	Nesta fase é determinado o objetivo do estudo e são escolhidos os métodos que serão utilizados durante o estudo.
Coleta de dados	Nesta fase é realizada a coleta dos dados. Pode ser do tipo direta ou indireta. A direta ocorre quando os dados são coletados pelo próprio pesquisador através de questionários ou quando é feito sobre elementos informativos de registro obrigatório (nascimentos, óbitos, compra e venda de mercadorias). A do tipo indireta pode ser classificada quanto ao fator de tempo como contínua, periódica e ocasional.
Crítica dos dados	Nesta fase os dados obtidos devem ser criticados à procura de falhas no planejamento, aquisição e armazenamento dos dados.
Apuração dos dados	Nesta fase, os dados são processados mediante critérios de classificação que foram definidos na fase de planejamento.
Exposição ou apresentação dos dados	Nesta fase, segundo Crespo (1999), os dados são apresentados através de tabelas ou gráficos, seguindo os critérios determinados no planejamento e utilizados no processamento dos dados. A exposição dos dados tem o objetivo de facilitar a análise daquilo que é objeto do estudo estatístico.
Análise	Esta última etapa do processo estatístico consiste em tirar conclusões sobre os dados levantados e processados, inferindo conclusões sobre o todo (população) a partir de dados coletados de uma parte representativa da população (amostra).

Fonte: próprio autor.



CAPÍTULO 4 – TABELAS

Análise Descritiva dos Dados

Uma das preocupações da Estatística reside no fato de analisar os dados, os quais necessitam de, previamente, estarem organizados e resumidos, o que é possível em virtude da tabulação dos dados, a qual deverá seguir a norma NBR 14724, de 2011. Dessa forma, a Análise Descritiva dos Dados consiste, basicamente, num processo de organização e descrição dos dados, em que os elementos básicos são as tabelas, gráficos e as medidas numéricas.

4.1 Tabela

Consiste num quadro que resume os dados de maneira segura, oferecendo uma visão geral do comportamento do fenômeno que se propõe analisar.

A tabela apresenta-se, genericamente, da seguinte forma:

Tabela 1 - Apresentação genérica de uma tabela

Título da Tabela
Corpo da Tabela
Rodapé

Fonte: próprio autor.



Exemplo:

Tabela 2 – Produção de Café Brasil 1991 a 1995

Tabela – Produção de Café Brasil – 1991 a 1995	
Anos	Produção (1.000 t)
1991	2.535
1992	2.666
1993	2.122
1994	3.750
1995	2.007

Fonte: IBGE

4.1.1 Composição da Tabela:

- ✓ **Título:** é a indicação que precede a tabela. Um bom título deverá conter de forma clara a referência ao fenômeno que é descrito, ao local onde ocorreu e a época à qual se refere;
- ✓ **Cabeçalho:** é a parte superior da tabela, a qual especifica o conteúdo das colunas.
- ✓ **Corpo da tabela:** é o espaço que contém as informações sobre o fenômeno observado. Há ainda alguns elementos complementares à tabela localizadas, de preferência, no rodapé, que são:
 - ✓ **Fonte:** é a indicação da entidade responsável pelo levantamento dos dados.
 - ✓ **Notas:** é o texto que esclarece o conteúdo estudado, podendo ser de caráter geral ou específico;
 - ✓ **Chamadas:** símbolo remissivo atribuído a algum elemento da tabela que necessita de uma nota específica;



Tabela 3 - Composição de uma tabela

Tabela – Produção de Café Brasil – 1991 a 1995		Título
Cabeçalho	Anos	Produção (1.000 t)
Coluna indicadora	1991	2.535
	1992	2.666
	1993	2.122
	1994	3.750
	1995	2.007
Rodapé	Fonte: IBGE	

Fonte: IBGE

4.1.2 Sinais Convencionais

De acordo com a Resolução nº 886-IBGE, nas casas ou células devemos realizar, as seguintes substituições padronizadas de determinada informação constante da tabela:

- ✓ **Traço (-)**: dado numérico igual a zero;
- ✓ **(...)**: ausência de dados;
- ✓ **(?)**: presença de dúvida na informação; e
- ✓ **Zero (0)**: quando a informação for muito pequena.

4.2 Distribuição de Frequências

Neste momento, serão aprofundados os conceitos de tabulação de dados, estudando as distribuições de frequências como técnica para a realização de análises estatísticas de grandes volumes de dados.



4.2.1 Conceitos importantes

a. Dados Brutos

São os dados que ainda não foram numericamente ordenados.

Suponha um levantamento de dados simples, em que os resultados foram dispostos da forma com que foram coletados, referentes às idades de 20 alunos da equipe de judô de uma escola de formação militar, para fins de classificação dentro das categorias da competição das Olimpíadas Escolares:

18, 23, 23, 19, 20, 19, 19, 21, 21, 21, 18, 22, 18, 20, 22, 23, 18, 23, 21, 22

b. Tabela Primitiva

Baseado nos dados coletados acima, a Tabela Primitiva consiste numa tabela cujos elementos não foram numericamente organizados.

Tabela 4 - Tabela Primitiva

18	23	23	19
20	19	19	21
21	21	18	22
18	20	22	23
18	23	21	22

Fonte: próprio autor.

c. Rol

É o arranjo dos Dados Brutos em ordem crescente ou decrescente. Pelo exemplo supracitado, tem-se:

18, 18, 18, 18, 19, 19, 19, 20, 20, 21, 21, 21, 21, 22, 22, 22, 23, 23, 23, 23



No caso de organizar o Rol em forma de Tabela (da esquerda para a direita e de cima para baixo), tem-se:

Tabela 5 - Tabela de dados em Rol

18	18	18	18
19	19	19	20
20	21	21	21
21	22	22	22
23	23	23	23

Fonte: próprio autor.

d. Amplitude Amostral ou Range (A)

É a magnitude de variação existente entre os dados levantados. É calculada pela diferença entre o valor máximo e o valor mínimo da amostra ou do Rol. Sendo assim, considerando o levantamento das idades acima, tem-se:

$$A = 23 - 18 = 5$$

4.2.2 Distribuição de Frequências Sem Intervalos de Classe

Distribuição de Frequências é uma tabela que reúne um conjunto de dados, conforme a frequência de seus valores, referentes a algum fato, local e época de determinado fenômeno.

Diante da tarefa de tabular uma quantidade muito grande de dados, antes de os analisar mesmos, faz-se necessário que eles sejam organizados e agrupados em subconjuntos que apresentem características similares. Esses agrupamentos são chamados de classe e o processo de realizar o agrupamento são chamados de distribuição de frequências.

Dessa forma, pode-se dizer que Distribuição de Frequências é uma forma de apresentação resumida de dados constantes numa tabela, na qual são agrupados em classes, possibilitando o fornecimento de quantidades ou percentuais de dados em cada classe e, conseqüentemente, a análise destes e a formulação de conclusões sem levar em conta os valores individuais.



Aqui, primeiramente, definiremos os diversos elementos constantes de uma distribuição, ao passo que construiremos uma Distribuição de Frequências Sem Intervalos de Classe, propriamente dita.

a. Elementos de uma Distribuição de Frequências Com Intervalos de Classe

1) Frequência Absoluta ou Simples (f_i)

Frequência Simples ou Absoluta é o número de vezes que uma determinada variável assume esse valor.

Consideremos, à título de exemplo, o Rol já estudado anteriormente, a saber:

18, 18, 18, 18, 19, 19, 19, 20, 20, 21, 21, 21, 21, 22, 22, 22, 23, 23, 23, 23

Verifica-se, pelos dados fornecidos, que a idade de 18 anos repete 4 vezes, isto é, há 4 alunos com a idade de 18 anos, enquanto que a idade de 19 anos repete 3 vezes, ou seja, há 3 alunos com a idade de 19 anos.

Assim, tem-se:

- a Frequência Absoluta ou Simples (f_i), referente à idade (x_1) = 18 anos, corresponde a 4, ou seja, $f_1 = 4$;

- a Frequência Absoluta ou Simples (f_i), referente à idade (x_2) = 19 anos, corresponde a 3, ou seja, $f_2 = 3$;

E assim por diante.

Logo, pode-se condensar tais resultados, para os dados discretos da amostra (Rol) supracitada (Idades dos alunos da equipe de Judô), de acordo com a seguinte distribuição de frequência:



Tabela 6 - Distribuição de Frequência 1

Idade (x_i)	Frequência Absoluta (f_i)
18	4
19	3
20	2
21	4
22	3
23	4
TOTAL	$n = \sum f_i = 20$

Fonte: próprio autor.

2) Frequência Absoluta Relativa (f_R)

Frequência Absoluta Relativa é a Frequência Absoluta expressa em percentagem.

Para o cálculo da percentagem, deve-se comparar a Frequência Absoluta (f_i) com o Somatório das Frequências ($\sum f_i = 20$). Sendo assim,

- para $x_1 = 18$, cuja Frequência Absoluta é $f_1 = 4$, tem-se:

$$f_R = \frac{4}{20} = 0,2 = 20\%$$

- para $x_2 = 19$, cuja Frequência Absoluta é $f_2 = 3$, tem-se:

$$f_R = \frac{3}{20} = 0,15 = 15\%$$

E assim por diante.



Logo, efetuando os demais cálculos, obtem-se a seguinte tabela:

Tabela 7 - Distribuição de Frequência 2.

Idade (x_i)	Frequência Absoluta (f_i)	Frequência Absoluta Relativa (f_R)
18	4	$4/20 = 0,20 = 20\%$
19	3	$3/20 = 0,15 = 15\%$
20	2	$2/20 = 0,10 = 10\%$
21	4	$4/20 = 0,20 = 20\%$
22	3	$3/20 = 0,15 = 15\%$
23	4	$4/20 = 0,20 = 20\%$
TOTAL	$n = \sum f_i = 20$	$20/20 = 1 = 100\%$

Tabela 7: Distribuição de Frequência 2.

3) Frequência Acumulada (F_{Ac})

Frequência Acumulada de um valor é o número de vezes que uma variável assume um valor inferior ou igual a esse valor.

Para a obtenção dos valores da Frequência Acumulada (F_{Ac}), observe que:

- para a idade $x_1 = 18$, tem-se que $F_{Ac1} = f_1 = 4$;
- para a idade $x_2 = 19$, tem-se que $F_{Ac2} = f_1 + f_2 = 4 + 3 = 7$;
- para a idade $x_3 = 20$, tem-se que $F_{Ac3} = f_1 + f_2 + f_3 = F_{Ac2} + 2 = 7 + 2 = 9$;

E assim por diante.



Logo, efetuando os demais cálculos, obtem-se a seguinte tabela:

Tabela 8 - Distribuição de Frequência 3.

Idade (x_i)	Frequência Absoluta (f_i)	Frequência Absoluta Relativa (f_R)	Frequência Acumulada (F_{Ac})
18	4	$4/20 = 0,20 = 20\%$	4
19	3	$3/20 = 0,15 = 15\%$	$4+3=7$
20	2	$2/20 = 0,10 = 10\%$	$7+2=9$
21	4	$4/20 = 0,20 = 20\%$	$9+4=13$
22	3	$3/20 = 0,15 = 15\%$	$13+3=16$
23	4	$4/20 = 0,20 = 20\%$	$16+4=20$
TOTAL	$n = \sum f_i = 20$	$20/20 = 1 = 100\%$	-

Fonte: próprio autor.

4) Frequência Acumulada Relativa (F_{AcR})

Frequência Acumulada Relativa é a Frequência Acumulada expressa em porcentagem.

Para o cálculo da porcentagem, referente à Frequência Acumulada Relativa (F_{AcR}), deve-se comparar a respectiva Frequência Acumulada (F_{Ac}) com o Somatório das Frequências ($\sum f_i$). Sendo assim, tem-se:

- para $x_1 = 18$, cuja Frequência Acumulada é $F_{Ac1} = 4$, tem-se:

$$F_{AcR} = \frac{4}{20} = 0,2 = 20\%$$

- para $x_2 = 19$, cuja Frequência Acumulada é $F_{Ac2} = 7$, tem-se:

$$F_{AcR} = \frac{7}{20} = 0,35 = 35\%$$

E assim por diante...



Logo, efetuando os demais cálculos, obtem-se a seguinte tabela:

Tabela 9 - Distribuição de Frequência 4.

Idade (x_i)	Frequência Absoluta (f_i)	Frequência Absoluta Relativa (f_R)	Frequência Acumulada (F_{Ac})	Frequência Acumulada Relativa (F_{AcR})
18	4	$4/20 = 0,20 = 20\%$	4	$4/20 = 0,20 = 20\%$
19	3	$3/20 = 0,15 = 15\%$	7	$7/20 = 0,35 = 35\%$
20	2	$2/20 = 0,10 = 10\%$	9	$9/20 = 0,45 = 45\%$
21	4	$4/20 = 0,20 = 20\%$	13	$13/20 = 0,65 = 65\%$
22	3	$3/20 = 0,15 = 15\%$	16	$16/20 = 0,80 = 80\%$
23	4	$4/20 = 0,20 = 20\%$	20	$20/20 = 1,00 = 100\%$
TOTAL	$n = \sum f_i = 20$	$20/20 = 1 = 100\%$	-	-

Fonte: próprio autor.

Observação Importante!

À tabela construída acima denominamos **Distribuição de Frequências Sem Intervalos de Classe!**

4.2.3 Distribuição de Frequências Com Intervalos de Classe

Diante da análise das tabelas apresentadas anteriormente, verifica-se que o processo é um pouco inconveniente, pois acaba exigindo muito espaço, mesmo quando o número de valores da variável é de tamanho razoável. Consequentemente, surge a necessidade de agruparmos esses valores em vários intervalos, os quais, em Estatística, serão denominados de Intervalos de Classes.

Sendo assim, chamando de frequência de uma classe o número de valores da variável pertencente ao intervalo de classe, em que os dados podem ser dispostos como na tabela denominada Distribuição de Frequência com Intervalo de Classe.



a. Elementos de uma Distribuição de Frequências Com Intervalos de Classe

1) Classe (i)

Classes de frequência ou, simplesmente, classes, são intervalos de variação da variável.

- ✓ As Classes representadas simbolicamente por i , sendo $i = 1, 2, 3, 4, \dots, k$, onde k é o número total de classes da distribuição;
- ✓ O número total de valores da variável classes é simbolizado por n .

2) Limites do Intervalo de Classe (l_i e L_i)

Utilizando-se dos conceitos de intervalos, observa-se que o menor valor do Intervalo de Classe é denominado **limite inferior** (l_i) e o maior é denominado **limite superior** (L_i).

No que tange ao intervalo de classe, observa-se:

- ✓ o termo fechado significa que o valor participa da classe;
- ✓ as classes, na medida do possível, deverão ter amplitudes iguais;
- ✓ o número de intervalos não deve ultrapassar 20, nem ser menor do que 5;
- ✓ ao construir o histograma, cada retângulo deve ter área proporcional à frequência relativa correspondente; e
- ✓ é representado graficamente com as seguintes simbologias:

$l_i \text{ -----} L_i$	Intervalo fechado à direita e aberto à esquerda.
$l_i \text{-----} L_i$	Intervalo fechado à direita e à esquerda
$l_i \text{-----} L_i$	Intervalo fechado à esquerda e aberto à direita

3) Amplitude Total da Distribuição (AT)

É a diferença entre o limite superior da última classe (limite superior máximo) e o limite inferior da primeira classe (limite inferior mínimo).

$$AT = L_{(max)} - l_{(min)}$$



4) Amplitude Amostral (AA)

É a diferença entre o valor máximo e o valor mínimo da amostra.

$$AA = x_{(max)} - x_{(min)}$$

5) Ponto Médio de Uma Classe (x_i)

É o ponto que divide o intervalo de classe em duas partes iguais, o qual é obtido pela metade da soma dos limites (inferior e superior) da classe.

Obs: O ponto médio de uma classe é o valor que a representa.

$$x_i = \frac{l_i + L_i}{2}$$

6) Número de Classes (i)

Para a determinação do Número de Classes de uma Distribuição de Frequências Com Intervalo de Classe deve-se lançar mão da **REGRA DE STURGES**, a qual permite exprimir o seu cálculo em função da variável:

$$i \cong 1 + 3,3 \cdot \log n$$

Onde:

i = número de classes

log = logaritmo

n = número de observações

Obs: lembrar de arredondar o valor obtido!

7) Amplitude de um Intervalo de Classe (h_i)

É a medida do intervalo que define a classe e é obtida pela diferença entre os limites do seu intervalo.



$$h_i = L_i - l_i$$

Neste ponto, após o cálculo do Número de Classes pela Regra de Sturges, compete agora resolver o problema da determinação da Amplitude de um Intervalo de Classe (h_i), o que se consegue dividindo a Amplitude Amostral pelo Número de Classes:

$$h_i = \frac{\text{Amplitude Amostral}}{\text{Número de Classes}}$$

Obs: lembrar de arredondar o valor obtido para o 1º número inteiro a maior!

b. Processo prático de Construção de uma Distribuição de Frequência Com Intervalo de Classe.

i - Inicialmente, suponhamos uma coleta de dados relativos à estatura de 40 alunos do Curso de Comunicações da Escola de Sargentos das Armas, que compõem um amostra desse curso, resultando na seguinte Tabela Primitiva:

166	160	161	150	162	160	165	167	164	160
162	161	168	163	156	173	160	155	164	168
155	152	163	160	155	155	169	151	170	164
154	161	156	172	153	157	156	158	158	161

ii - Organizando a Tabela em Rol (da esquerda para a direita e de cima para baixo), observa-se:

150	151	152	153	154	155	155	155	155	156
156	156	157	158	158	160	160	160	160	160
161	161	161	161	162	162	163	163	164	164
164	165	166	167	168	168	169	170	172	173



III – Cálculo da Amplitude Amostral (AA)

$$AA = x_{(max)} - x_{(min)}$$

Logo, obtem-se:

$$AA = 173 - 150 = 23$$

IV – Cálculo do Número de Classes (i)

Utilizar-se-á a Regra de Sturges:

$$i \cong 1 + 3,3 \cdot \log n$$

Agora, considerando $\log 40 \cong 1,6$ e $n = 40$ (pois foram coletados dados de 40 alunos), obtem-se:

$$i \cong 1 + 3,3 \cdot \log 40$$

$$i \cong 1 + 3,3 \cdot 1,6$$

$$i \cong 1 + 5,28$$

$$i \cong 6,28$$

Agora, arredondando o valor obtido, tem-se:

$$i = 6$$

V – Cálculo da Amplitude de um Intervalo de Classe (h_i)

$$h_i = \frac{\text{Amplitude Amostral}}{\text{Número de Classes}}$$

Efetando os cálculos, tem-se:

$$h = \frac{23}{6} = 3,83$$

Logo, fazendo o arredondamento para o 1º número inteiro a maior (por convenção), tem-se que $h = 4$.



Sendo assim, pode-se concluir que a Distribuição de Frequência Com Intervalos de Classe terá 6 (seis) Classes com intervalos de 4 (quatro).

VI – Construção, propriamente dita, da Distribuição de Frequência com Intervalo de Classe

Sabendo, então, que a distribuição deverá possuir 6 (seis) classes (i) e que cada classe deverá ter intervalo de 4 (quatro), obtem-se, a partir do dado 150, os seguintes intervalos de classe:

$i = 1$: 150 |----- 154;

$i = 2$: 154 |----- 158;

$i = 3$: 158 |----- 162;

$i = 4$: 162 |----- 166;

$i = 5$: 166 |----- 170;

$i = 6$: 170 |----- 174.

Logo, tem-se a seguinte tabela:

Tabela 10 - Distribuição de Frequência 5.

i	Estaturas (cm)	Frequência Absoluta (f_i)
1	150 ----- 154	4
2	154 ----- 158	9
3	158 ----- 162	11
4	162 ----- 166	8
5	166 ----- 170	5
6	170 ----- 174	3
-	-	$\sum f_i = 40$

Fonte: próprio autor.



Resumindo a Distribuição de Frequência com Intervalo de Classe com TODAS as frequências estudadas, obtem-se:

Tabela 11 - Distribuição de Frequência 6.

i	Estaturas (cm)	Frequência Absoluta (f_i)	Ponto Médio (x_i)	Frequência Absoluta Relativa (f_R)	Frequência Acumulada (F_{Ac})	Frequência Acumulada Relativa (F_{AcR})
1	150 ----- 154	4	$\frac{150 + 154}{2} = 152$	$4/40=0,10=$ 10%	4	$4/40=0,10=$ 10%
2	154 ----- 158	9	$\frac{154 + 158}{2} = 156$	$9/40=0,225=$ 22,5%	13	$13/40=0,325=$ 32,5%
3	158 ----- 162	11	$\frac{158 + 162}{2} = 160$	$11/40=0,275=$ 27,5%	24	$24/40=0,60=$ 60%
4	162 ----- 166	8	$\frac{162 + 166}{2} = 164$	$8/40=0,20=$ 20%	32	$32/40=0,80=$ 80%
5	166 ----- 170	5	$\frac{166 + 170}{2} = 168$	$5/40=0,125=$ 12,5%	37	$37/40=0,925=$ 92,5%
6	170 ----- 174	3	$\frac{170 + 174}{2} = 172$	$3/40=0,075=$ 7,5%	40	$40/40=1=$ 100%
Total	-	$\sum f_i = 40$	-	$40/40=1=$ 100%	-	-

Fonte: próprio autor.

Observações Importantes!

Mediante a análise da Distribuição de Frequência com Intervalo de Classe construída acima, é possível a obtenção e verificação dos vários tipos de frequência, de modo a permitir a resposta a muitas questões com relativa facilidade, como as especificadas abaixo:



- ✓ Quantos alunos têm estatura entre 154 cm, inclusive, e 158 cm?

Resposta:

Basta analisar os valores da variável que formam a 2ª Classe (154 |----- 158).

Como $f_2 = 9$, a resposta é 9 alunos.

- ✓ Qual é a porcentagem de alunos cujas estaturas são inferiores a 154 cm?

Resposta:

Analisando a 1ª Classe (150 |----- 154), como sua Frequência Absoluta Relativa $f_R = 10\%$, logo a porcentagem de alunos é 10%.

- ✓ Quantos alunos têm estatura abaixo de 162 cm?

Resposta: verificando a distribuição, percebem-se que as estaturas consideradas são aquelas constantes das 1ª, 2ª e 3ª Classes.

Assim, o número de alunos é dado pela Frequência Acumulada (F_{Ac}) na 3ª Classe, que é 24, ou seja,

$$f_1 + f_2 + f_3 = \sum_{i=1}^3 f_i = F_{Ac\ 3} = 24$$

Portanto, 24 alunos têm estatura abaixo de 162 cm.



EXERCÍCIOS

Capítulo 4 – Tabelas

1. (UFMT - adaptado)

Em um estudo sobre o potencial de germinação de sementes de algodão, dividiu-se uma área em 40 parcelas com a mesma área, tipo de solo, iluminação, etc. Em cada uma destas parcelas foram plantadas 4 sementes e verificou-se o número de sementes que germinaram. Os dados obtidos são apresentados a seguir:

2	0	0	4	3	0	0	1	0	0
2	1	1	1	1	1	1	0	0	0
0	0	2	0	0	1	1	2	0	2
1	3	0	3	0	0	0	0	0	3

a) Especifique a variável estudada, classificando-a quanto ao Tipo de Variável. Justifique a sua resposta.

Resposta:

A variável estudada é “Nº de sementes que germinaram” e é classificada, quanto ao tipo, como quantitativa discreta, pois ela é enumerável e foi obtida a partir de um processo de contagem.

b) Represente tabularmente os dados supracitados, de acordo com a Distribuição de Frequências Sem Intervalos de Classe abaixo, completando as células em branco:



Nº de Germinações	Frequência Absoluta (f_i)	Frequência Absoluta Relativa (f_R)	Frequência Acumulada (F_{Ac})	Frequência Acumulada Relativa (F_{AcR})
Total			-	-

Resposta:

Nº de Germinações	Frequência Absoluta (f_i)	Frequência Absoluta Relativa (f_R)	Frequência Acumulada (F_{Ac})	Frequência Acumulada Relativa (F_{AcR})
0	20	50%	20	50%
1	10	25%	30	75%
2	5	12,5%	35	87,5%
3	4	10%	39	97,5%
4	1	2,5%	40	100,00%
Total	40	100%	-	-

c) Qual a percentagem de parcelas em que germinaram no máximo 2 sementes?

Resposta:

Analisando a tabela acima, basta verificar a Frequência Acumulada Relativa (F_{AcR}) com número de germinações igual a 2, a qual corresponde a 87,5%.



2. (UFMT – adaptado)

Os dados seguintes representam 20 observações relativas aos índices pluviométricos de determinados municípios do Estado:

Milímetros de Chuva									
144	152	159	160	160	151	157	146	154	145
141	150	142	146	142	141	141	150	143	158

a) Construa a Distribuição de Frequência Com Intervalo de Classe, completando as células em branco.

i	Milímetros de Chuva (mm)	Frequência Absoluta (f_i)	Ponto Médio (x_i)	Frequência Absoluta Relativa (f_R)	Frequência Acumulada (F_{Ac})	Frequência Acumulada Relativa (F_{AcR})
1	-----					
2	-----					
3	-----					
4	-----					
5	-----					
Total		$\sum f_i =$	-		-	-



Resposta:

1º Passo) Colocar os dados da Tabela Primitiva fornecida em Rol (ordem crescente da esquerda para a direita e de cima para baixo):

Milímetros de Chuva									
141	141	141	142	142	143	144	145	146	146
150	150	151	152	154	157	158	159	160	160

2º Passo) Calcular a Amplitude Amostral (AA)

$$AA = x_{(max)} - x_{(min)}$$

Logo, tem-se:

$$AA = 160 - 141 = 19$$

3º Passo) Calcular o Número de Classes (i), utilizando a Regra de Sturges,

$$i \cong 1 + 3,3 \cdot \log n$$

Considerando $\log 20 \cong 1,3$ e $n = 20$ (pois foram coletados dados de 20 municípios), obten-se:

$$i \cong 1 + 3,3 \cdot \log 20$$

$$i \cong 1 + 3,3 \cdot 1,3$$

$$i \cong 1 + 4,29$$

$$i \cong 5,29$$

Agora, arredondando o valor obtido, tem-se que:

$$i = 5$$



4º Passo) Calcular a Amplitude de um Intervalo de Classe (h_i)

$$h_i = \frac{\text{Amplitude Amostral}}{\text{Número de Classes}}$$

Efetuada os cálculos, tem-se:

$$h = \frac{19}{5} = 3,8$$

Logo, fazendo o arredondamento para o 1º número inteiro a maior (por convenção), tem-se:

$$h = 4.$$

Sendo assim, podemos concluir que a Distribuição de Frequência com Intervalos de Classe terá 5 (cinco) Classes, com intervalos de 4 (quatro)!

5º Passo) Construção da Distribuição de Frequência com Intervalo de Classe

Sabendo, então, que a distribuição deverá possuir 5 (cinco) classes (i) e que cada classe deverá ter intervalo de 4 (quatro), obtem-se:

i	Milímetros de Chuva (mm)	Frequência Absoluta (f_i)	Ponto Médio (x_i)	Frequência Absoluta Relativa (f_R)	Frequência Acumulada (F_{Ac})	Frequência Acumulada Relativa (F_{AcR})
1	141 ----- 145	7	143	$7/20=0,35=35\%$	7	35%
2	145 ----- 149	3	147	$3/20=0,15=15\%$	10	50%
3	149 ----- 153	4	151	$4/20=0,20=20\%$	14	70%
4	153 ----- 157	1	155	$1/20=0,05=5\%$	15	75%
5	157 ----- 161	5	159	$5/20=0,25=25\%$	20	100%
Total	-	$\sum f_i = 20$	-	100%	-	-



CAPÍTULO 5 – REPRESENTAÇÕES GRÁFICAS

4.3 Gráficos Estatísticos

O gráfico estatístico é uma forma de apresentação de dados estatísticos, cujo objetivo é permitir ao leitor uma impressão mais rápida e dinâmica do fenômeno em estudo.

O gráfico é um instrumento que possibilita transmitir significados de forma mais rápida e direta que textos ou tabelas complexas, devendo, portanto, conhecer o tipo de informação que se deseja transmitir para, então, criar um determinado tipo de gráfico.

O profissional da estatística se utiliza de gráficos, pois eles se propõem a diferentes usos: exploração, descoberta, interpretação, diagnóstico, comunicação, apresentação, destaque, clareza e até contar uma história.

O gráfico atende às mais diversas situações, sendo assim, ele é diretamente ligado a cada tipo de informação que se deseja transmitir e deve obedecer a alguns requisitos, a saber:

- ✓ **Simplicidade:** deve ser destituído de detalhes de importância secundária;
- ✓ **Clareza:** possibilitar uma interpretação correta dos valores; e
- ✓ **Veracidade:** expressar a verdade sobre o fenômeno em estudo.

4.3.1 Tipos de Gráficos

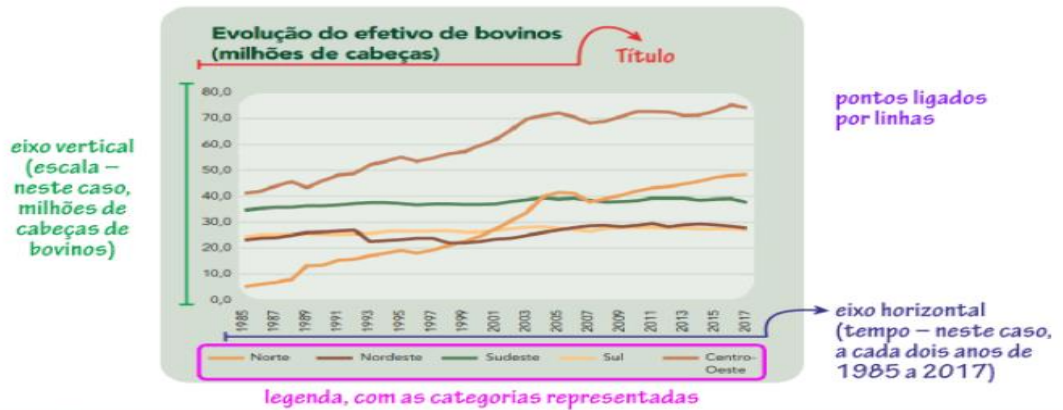
Quanto ao tipo de gráfico, os mais utilizados são:



a. Linha ou Curva

Gráfico que utiliza uma linha poligonal para representar uma série estatística, constituindo-se de uma aplicação da representação das funções no Sistema de Coordenadas Cartesianas Ortogonais.

Figura 1 - Gráfico de Linhas

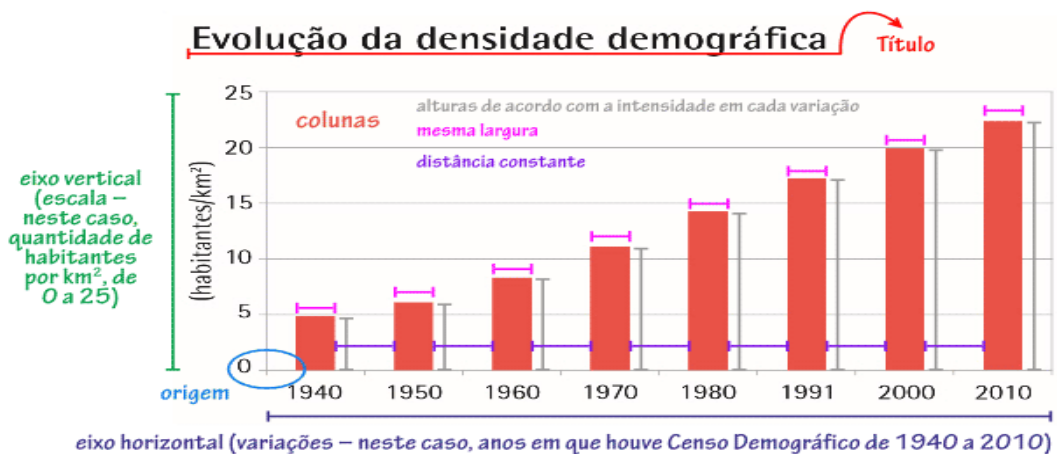


Fonte: IBGE

b. Colunas ou Barra

Representam séries estatísticas por meio de retângulos, dispostos verticalmente (em colunas) ou horizontalmente (em barras), de modo que todos os retângulos devem apresentar a mesma largura, ficando os seus comprimentos proporcionais aos respectivos dados.

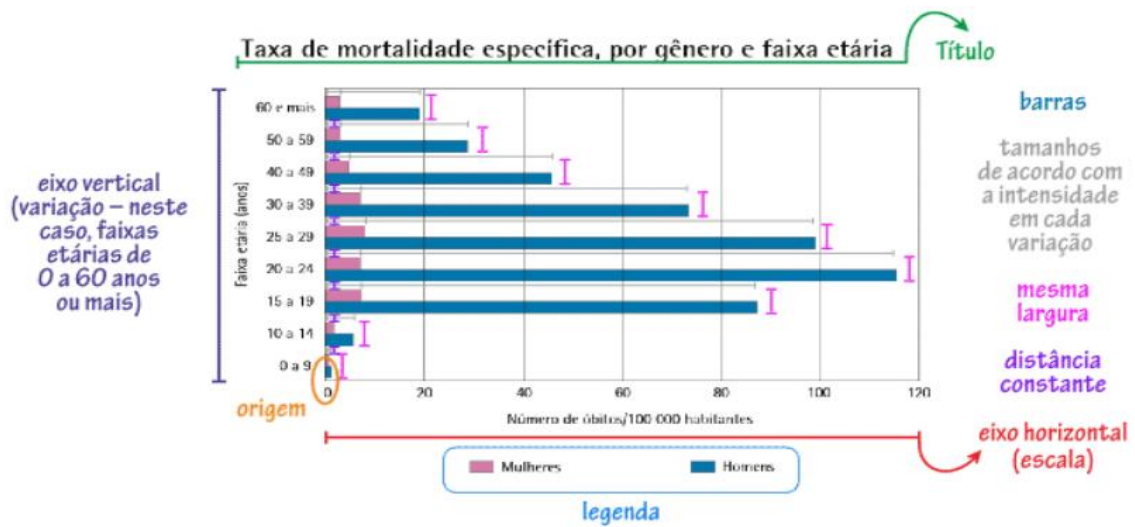
Figura 2 - Gráfico de Colunas



Fonte: IBGE



Figura 3 - Gráfico de Barras

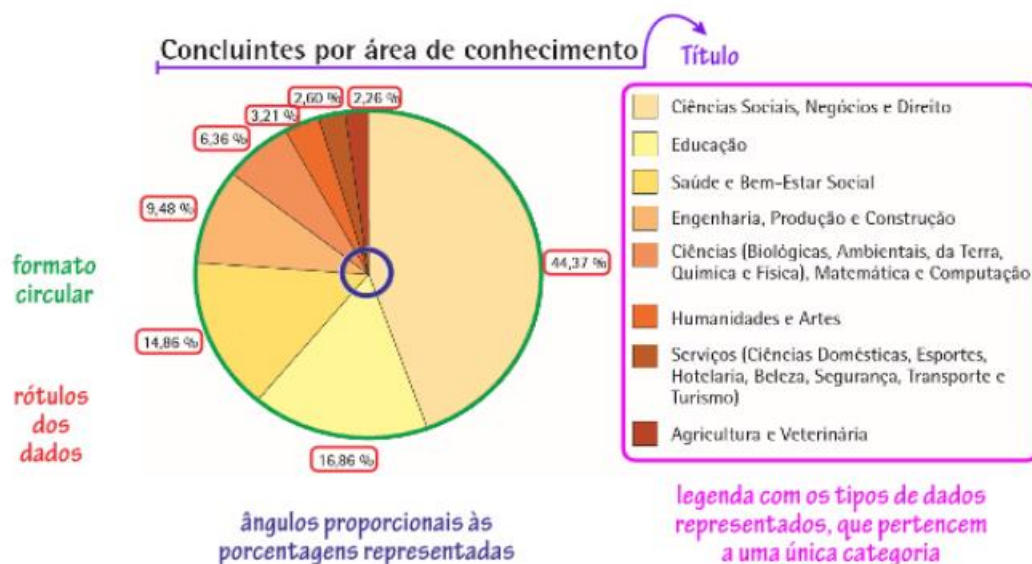


Fonte: IBGE

c. Setores

São representados por meio de setores de um círculo (Pizza) e tem a finalidade de representar um fenômeno e todas as partes em que ele se subdivide. É muito útil para agrupar ou organizar quantitativamente dados considerando um total.

Figura 4 - Gráfico de Setores



Fonte: IBGE



d. Pictóricos ou Pictogramas

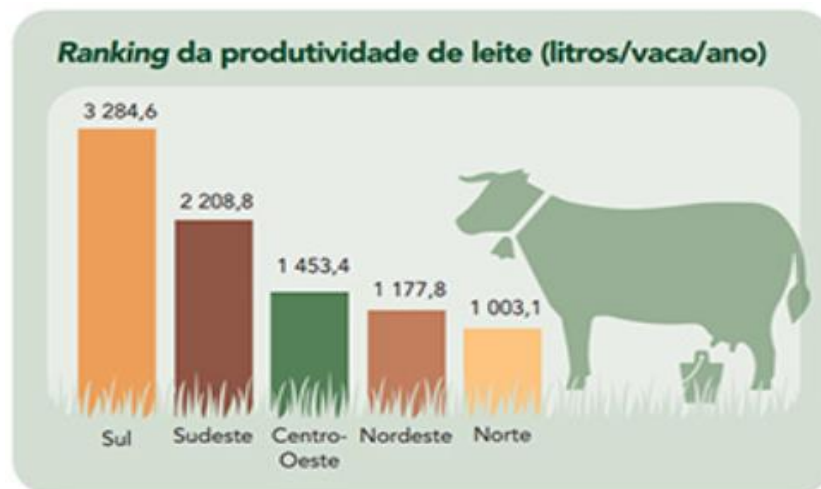
São gráficos que simbolizam, através de figuras, fatos estatísticos, ao mesmo tempo que indicam proporcionalidades.

Algumas regras fundamentais:

- ✓ Os símbolos devem explicar-se por si próprios;
- ✓ As quantidades maiores são indicadas por meio de um número maior de símbolos e não por um símbolo maior;
- ✓ Os símbolos comparam quantidades aproximadas, não detalhes minuciosos; e
- ✓ Os gráficos pictóricos só devem ser usados para comparações, nunca para afirmações detalhadas ou isoladas.

Exemplo: Ranking da produtividade de leite (litros/vaca/ano)

Figura 5 - Pictograma



Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Agropecuária, Pesquisa da Pecuária Municipal 2017.

Fonte do gráfico: IBGE, Produção da Pecuária Municipal 2017, pág. 4.

Fonte: IBGE



e. Cartogramas

É uma representação sobre cartas geográficas, sendo que pontos ou legendas representam as quantidades.

Exemplo: Distribuição da população/Concentração da população

Figura 6 - Cartograma

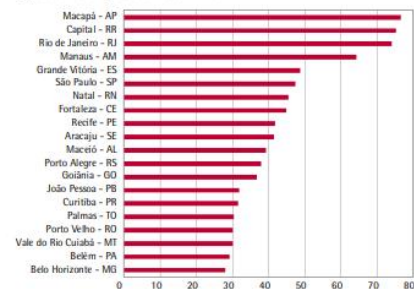
Distribuição da população 2017



Concentração da população Faixa litorânea



Regiões metropolitanas - 2017



Fontes: 1. Estimativas da população residente no Brasil e para as unidades da federação com data de referência em 1o de julho de 2017. Rio de Janeiro, IBGE, 2018. Disponível em: <<https://www.ibge.gov.br/estatisticas-novoportal/sociais/populacao/9103-estimativas-de-populacao.html?edicao=1638561=resultados>>. Acesso em: out. 2018. 2. Regiões metropolitanas, aglomerações urbanas e regiões integradas de desenvolvimento 2010-2019. Rio de Janeiro: IBGE, 2018. Disponível em: <<https://www.ibge.gov.br/geociencias-novoportal/organizacao-do-territorio/estrutura-territorial/18354-regioes-metropolitanas-aglomeracoes-urbanas-e-regioes-integradas-de-desenvolvimento.html?t=download>>. Acesso em: out. 2018.

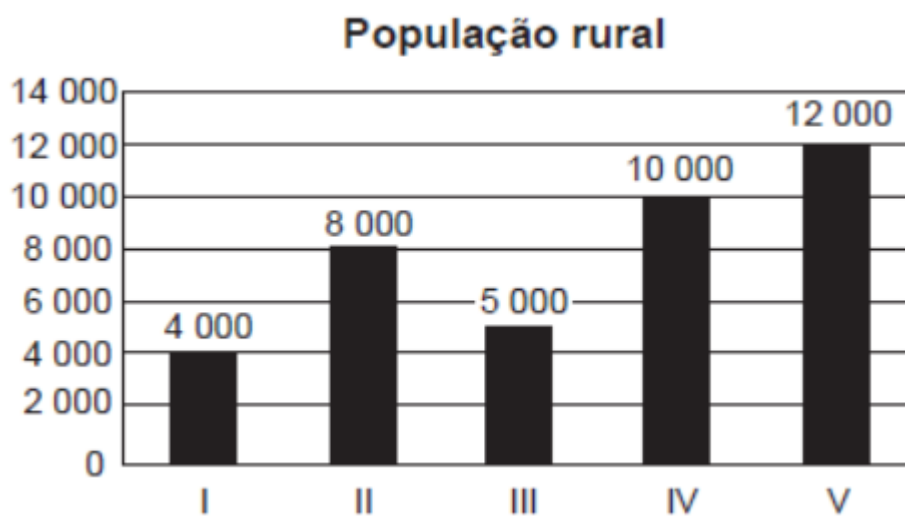
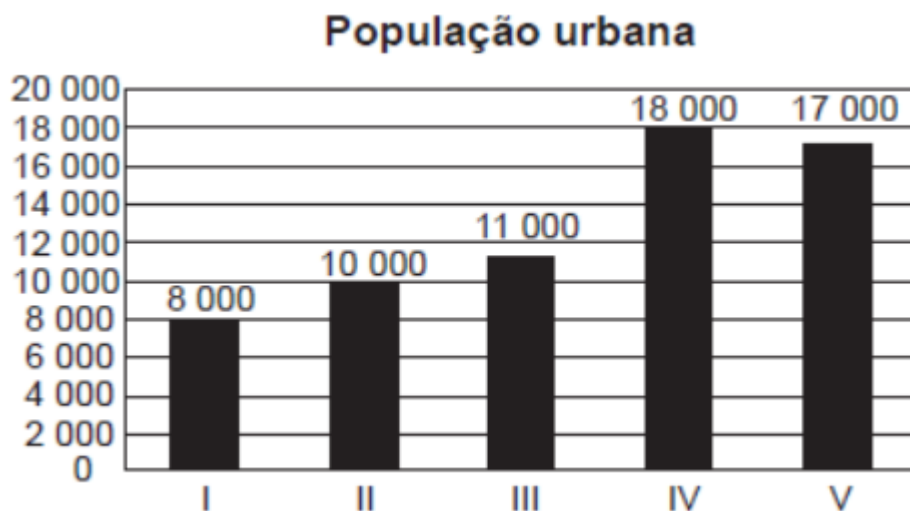


EXERCÍCIOS

Capítulo 5 – Representações Gráficas

1. (ENEM 2019)

A taxa de urbanização de um município é dada pela razão entre a população urbana e a população total do município (isto é, a soma das populações rural e urbana). Os gráficos apresentam, respectivamente, a população urbana e a população rural de cinco municípios (I, II, III, IV, V) de uma mesma região estadual. Em reunião entre o governo do estado e os prefeitos desses municípios, ficou acordado que o município com maior taxa de urbanização receberá um investimento extra em infraestrutura.





Segundo o acordo, qual município receberá o investimento extra?

- a) I
- b) II
- c) III
- d) IV
- e) V

Resposta: letra c) III

Solução:

Calculando as taxas de urbanização, obtém-se:

I) $8000/(8000 + 4000) = 0,666\dots$

II) $10000/(10000 + 8000) = 0,555\dots$

III) $11000/(11000 + 5000) = 0,6875\dots$

IV) $18000/(18000 + 10000) = 0,6428\dots$

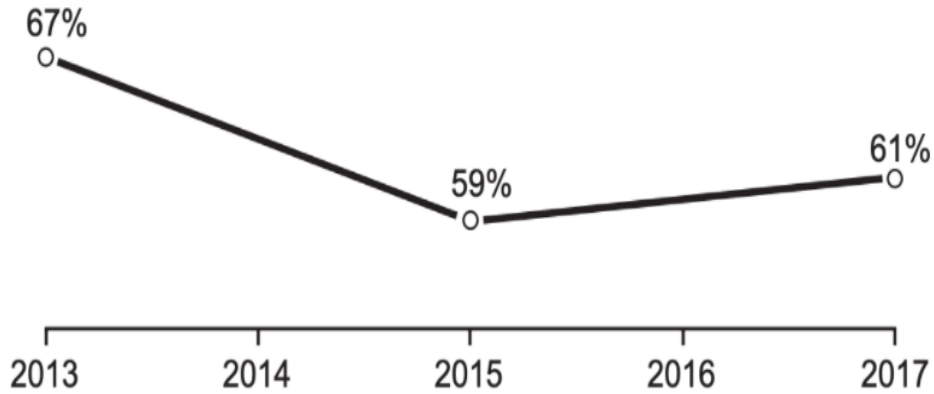
V) $17000/(17000 + 12000) = 0,5862\dots$

Logo, deverá ser o município III.

2. (ENEM 2018)

A raiva é uma doença viral e infecciosa, transmitida por mamíferos. A campanha nacional de vacinação antirrábica tem o objetivo de controlar a circulação do vírus da raiva canina e felina, prevenindo a raiva humana.

O gráfico mostra a cobertura (porcentagem de vacinados) da campanha, em cães, nos anos de 2013, 2015 e 2017, no município de Belo Horizonte, em Minas Gerais. Os valores das coberturas dos anos de 2014 e 2016 não estão informados no gráfico e deseja-se estimá-los. Para tal, levou-se em consideração que a variação na cobertura de vacinação da campanha antirrábica, nos períodos de 2013 a 2015 e de 2015 a 2017, deu-se de forma linear.



Disponível em: <http://pni.datasus.gov.br>. Acesso em: 5 nov. 2017.

Qual teria sido a cobertura dessa campanha no ano de 2014?

- a) 62,3%
- b) 63,0%
- c) 63,5%
- d) 64,0%
- e) 65,5%

Resposta: letra b) 63,0%

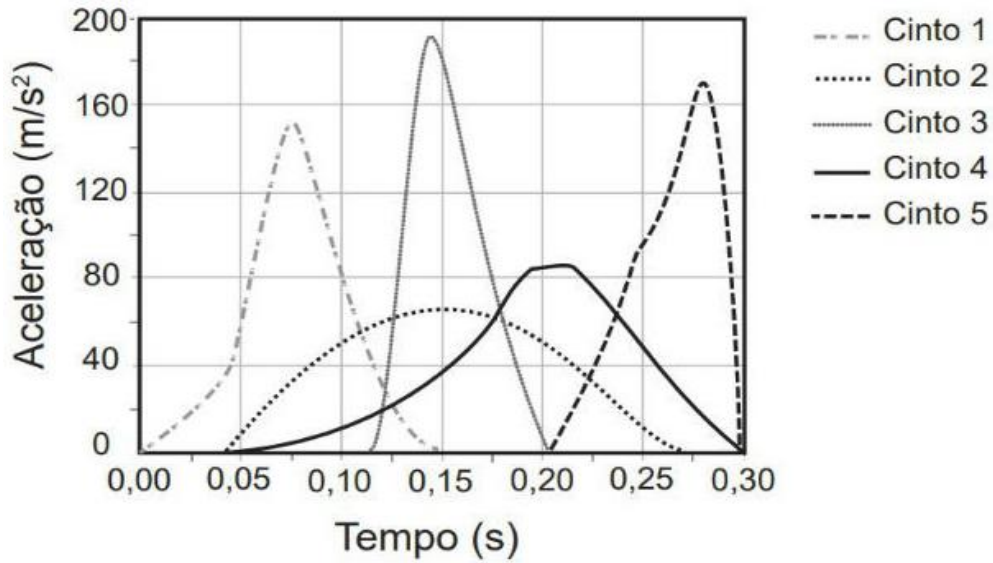
Solução:

A cobertura dessa campanha no ano de 2014 foi $(67\% + 59\%)/2 = 63\%$

3. (ENEM 2017)

Em uma colisão frontal entre dois automóveis, a força que o cinto de segurança exerce sobre o tórax e abdômen do motorista pode causar lesões graves nos órgãos internos.

Pensando na segurança do seu produto, um fabricante de automóveis realizou testes em cinco modelos diferentes de cinto. Os testes simularam uma colisão de 0,30 segundo de duração, e os bonecos que representavam os ocupantes foram equipados com acelerômetros. Esse equipamento registra o módulo da desaceleração do boneco em função do tempo. Os parâmetros como massa dos bonecos, dimensões dos cintos e velocidade imediatamente antes e após o impacto foram os mesmos para todos os testes. O resultado final obtido está no gráfico de aceleração por tempo.



Qual modelo de cinto oferece menor risco de lesão interna ao motorista?

- a) 1
- b) 2
- c) 3
- d) 4
- e) 5

Resposta: letra b)

Solução:

O risco de lesão interna está ligado à intensidade da força de frenagem e, portanto à intensidade da aceleração do carro.

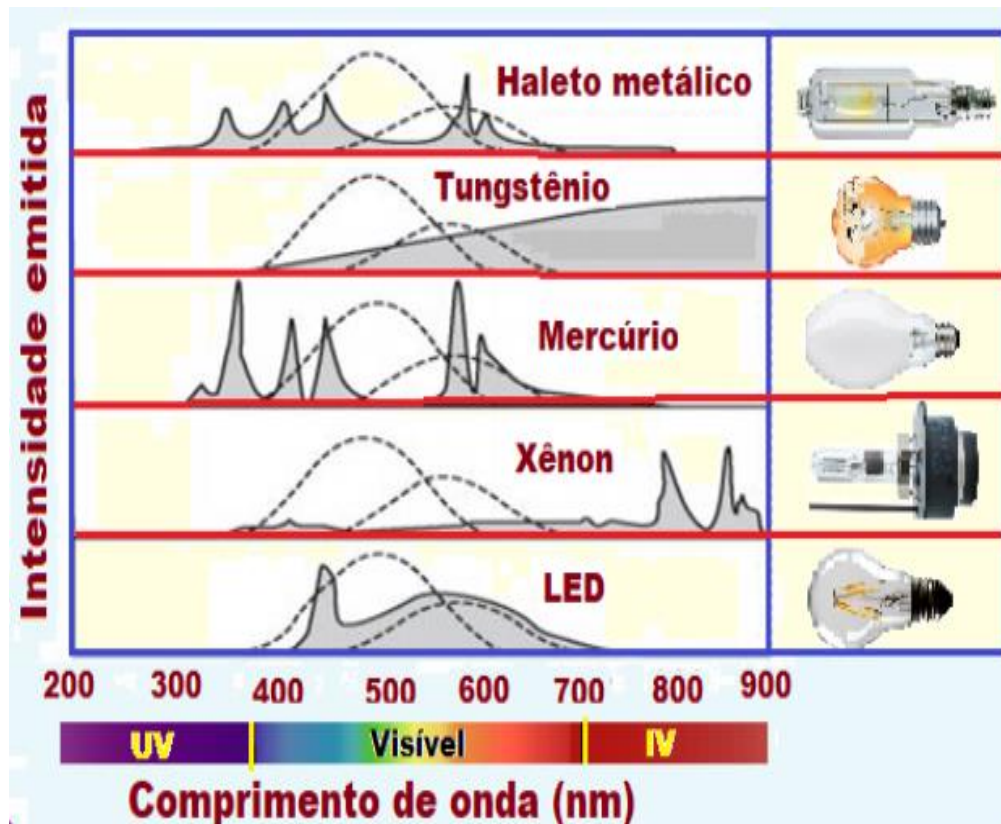
O cinto 2 corresponde ao de menor aceleração máxima e, portanto, o de menor risco de lesão.

4. (ENEM 2017)

A figura mostra como é a emissão de radiação eletromagnética para cinco tipos de lâmpada: haleto metálico, tungstênio, mercúrio, xênon e LED (diodo emissor de luz). As áreas marcadas em cinza são proporcionais à intensidade da energia liberada pela lâmpada. As linhas pontilhadas mostram a sensibilidade do olho humano aos diferentes comprimentos de onda. UV e IV são as regiões do ultravioleta e do infravermelho, respectivamente.



Um arquiteto deseja iluminar uma sala usando uma lâmpada que produza boa iluminação, mas que não aqueça o ambiente.



Qual tipo de lâmpada melhor atende ao desejo do arquiteto?

- a) Haletos metálicos
- b) Tungstênio
- c) Mercúrio.
- d) Xênon
- e) LED

Resposta: letra e)

Solução:

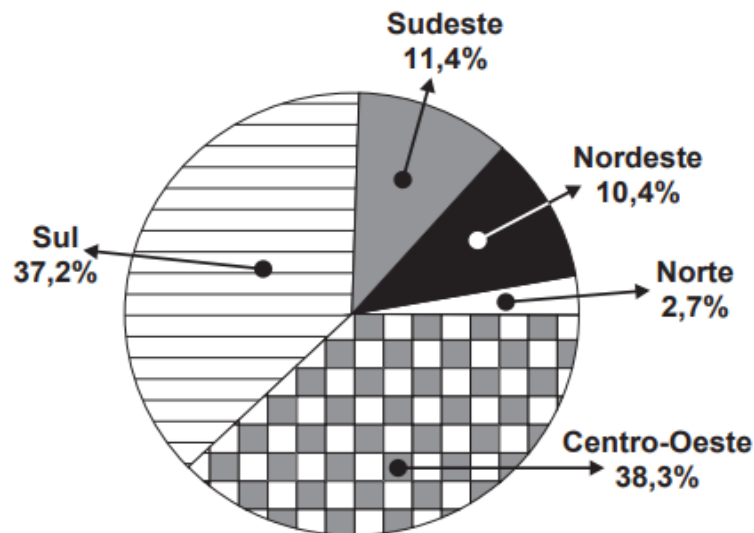
A lâmpada LED, dentre as opções apresentadas, é a que mais emite na faixa do visível. Além disso, a intensidade emitida na faixa do IV é desprezível, o que melhor atende o desejo do arquiteto, que é de não aquecer em demasia o ambiente.



5. (ENEM 2019)

Considere que a safra nacional de cereais, leguminosas e oleaginosas, em 2012, aponte uma participação por região conforme indicado no gráfico.

Em valores absolutos, essas estimativas indicam que as duas regiões maiores produtoras deveriam produzir juntas um total de 119,8 milhões de toneladas em 2012.



De acordo com esses dados, a produção estimada, em milhão de tonelada, de cereais, leguminosas e oleaginosas, em 2012, na Região Sudeste do país, foi um valor mais aproximado de:

- a) 11,4.
- b) 13,6.
- c) 15,7.
- d) 18,1.
- e) 35,6

Resposta: letra d)

Solução:

"As duas regiões maiores produtoras deveriam produzir juntas um total de 119,8 milhões de toneladas em 2012."

Sendo, Centro-Oeste: 38,3% e Sul: 37,2%.

Logo, somando = $38,3\% + 37,2\% = 75,5\%$.

Agora, deve ser feita uma regra de três simples para calcular o valor absoluto da Região Sudeste que possui 11,4% da safra em 2012.



Logo, tem-se que:

75,5% \rightarrow 119,8 (milhões de toneladas)

11,4% \rightarrow x (milhões de toneladas)

Sendo assim,

$x = (119,8 \cdot 11,4) / 75,5$ (Para simplificar os cálculos ajusta-se 119,8 para 200).

Portanto, $x \cong 18,1$ (milhões)



CAPÍTULO 6 – MEDIDAS DE POSIÇÃO

No estudo de uma série estatística é conveniente o cálculo de algumas medidas que a caracterizam. Essas medidas, quando bem interpretadas, podem fornecer informações muito valiosas relativas à série, sendo uma delas a Medida de Tendência Central.

As principais Medidas de Tendência Central ou de Posição são:

- ✓ Média;
- ✓ Mediana; e
- ✓ Moda.

5.1 Média Aritmética (\bar{x})

A Média Aritmética é a razão do somatório de todos os termos pelo número total de termos.

É considerada uma Medida de Tendência Central, pois focaliza valores médios dentre os maiores e menores. É utilizada no intuito de expressar, por meio de um único valor, a ideia principal de um grupo de valores.

Pode-se representar a Média Aritmética (para dados não agrupados) pela seguinte expressão:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Exemplo:

Dados os números 10, 8, 15 e 9, qual o valor da média aritmética simples?

$$\bar{x} = \frac{10 + 8 + 15 + 9}{4} = \frac{42}{4} = 10,5$$



5.2 Média Aritmética Ponderada (M_P)

A Média Aritmética Ponderada é aquela que possibilita a correção de distorções apresentadas pela Média Aritmética Simples, quando um ou mais valores for muito diferente do conjunto, atribuindo-se peso ou importância diferente a cada valor.

Pode-se representar a Média Aritmética Ponderada (para dados não agrupados) pela seguinte expressão:

$$M_P = \frac{p_1 \cdot x_1 + p_2 \cdot x_2 + \dots + x_n \cdot p_n}{p_1 + p_2 + \dots + p_n}$$

Exemplo:

Calcule a média dos números 10, 15 e 20, sabendo que seus pesos são 3, 4 e 3, respectivamente.

$$M_P = \frac{10 \cdot 3 + 15 \cdot 4 + 20 \cdot 3}{3 + 4 + 3} = \frac{30 + 60 + 60}{10} = \frac{150}{10} = 15$$

Resposta: 15

5.3 Mediana (M_d)

A Mediana é uma medida de posição do centro da distribuição dos dados.

Inicialmente, para definir o valor da Mediana (para dados não agrupados) deve-se colocar os elementos da amostra em Rol, ou seja, ordem crescente ou decrescente. Importante destacar que ela pertencerá ou não à amostra, dividindo-a ao meio, ou seja, 50% dos elementos da amostra são menores ou iguais à mediana e os outros 50% são maiores ou iguais à mediana.

Sendo assim, tem-se 2 casos distintos para o cálculo da mediana de uma amostra de n elementos (para dados não agrupados):



a. Se n é ímpar, então a mediana é o elemento central da Amostra.

Exemplo: suponhamos a seguinte Amostra, já em Rol.

1, 3, 4, 5, 7

Visualmente, pelo fato de termos uma quantidade muito pequena de elementos, é possível verificar que o elemento central da amostra é o **4**, ou seja, o **3º elemento da amostra**; entretanto, para uma amostra com uma quantidade de termos muito grande, fica inviável ou imperceptível a visualização do termo central, ou seja, da mediana.

Dessa forma, deve-se proceder determinando a posição do elemento central (e) de uma amostra com (n) elementos que, no caso do exemplo acima, n sendo **ímpar**, é obtida da seguinte forma:

$$e = \frac{(n + 1)}{2}$$

$$e = \frac{(5 + 1)}{2}$$

$$e = \frac{(6 + 1)}{2}$$

$$e = \frac{6}{2}$$

$$e = 3^{\text{a}} \text{ posição.}$$

Logo, a mediana será definida pelo elemento que ocupa a 3ª posição do Rol da amostra, isto é, $M_d = 4$.

b. Se n é par, então a mediana é a média aritmética dos dois elementos centrais da amostra.

Exemplo: suponhamos a seguinte Amostra, já em Rol.

1, 3, 5, 6, 7, 9

Sendo assim, deve-se proceder determinando as posições (e) e ($e+1$) dos elementos médios de uma amostra com (n) elementos que, no caso do exemplo acima, n sendo **par**, são obtidas da seguinte forma:



I - Para a definição dos elementos médios, obtem-se:

$$e = \frac{n}{2}$$

$$e = \frac{6}{2}$$

$$e = 3^{\text{a}} \text{ posição}$$

Logo, os elementos médios serão os que se encontram na 3ª posição (e) e também na 4ª posição do Rol ($e+1$), que no caso são os elementos 5 e 6.

II - Agora, como n é par, deve-se calcular a média aritmética entre os 3º e o 4º elementos, isto é:

$$M_d = \frac{\textit{elemento da 3ª posição} + \textit{elemento da 4ª posição}}{2}$$

Dessa forma, obtem-se:

$$M_d = \frac{5 + 6}{2} = \frac{11}{2} = 5,5$$

Logo, $M_d = 5,5$.

5.4 Moda (M_o)

A Moda é o valor que possui maior frequência, se os dados forem discretos, ou o intervalo de classe com maior frequência, se os dados forem contínuos.

Exemplo: suponhamos que os dados brutos abaixo correspondam à idade de 20 militares do C Cav que integrarão a turma de atletas para a competição de Pólo:

22, 21, 22, 23, 22, 21, 23, 22, 22, 21, 24, 23, 23, 22, 21, 19, 18, 18, 20, 20

Como a idade de 22 anos é o dado que mais se repete, neste caso, 6 (seis) vezes, então $M_o = 22$, pois é a idade que possui maior frequência.



Observação: um conjunto de dados pode ser Amodal (sem Moda), Bimodal (com 2 Modas), ou Multimodal (com mais de 3 Modas).

5.5 Média Aritmética, Moda e Mediana a partir de uma Distribuição de Frequências Sem Intervalo de Classe:

Acompanhe o seguinte exemplo:

Exemplo:

Considere a tabela abaixo, correspondente a uma Distribuição de Frequências Sem Intervalo de Classe, referente à pesquisa sobre o número de irmãos de cada aluno de um pelotão do Curso de Infantaria da ESA, para o cálculo da Média Aritmética, da Moda e da Mediana, a saber:

Número de irmãos (x_i)	Frequência Absoluta (f_i)
0	8
1	15
2	12
3	5
TOTAL	$n = \sum f_i = 40$

a. Cálculo da Média Aritmética:

$$\bar{x} = \frac{8 \cdot 0 + 15 \cdot 1 + 12 \cdot 2 + 5 \cdot 3}{8 + 15 + 12 + 5} = \frac{54}{40} = 1,35$$

b. Cálculo da Moda:

Como a maior Frequência Absoluta é o valor 15, a qual corresponde ao número de irmãos igual a 1, então, a Moda é 1 irmão, ou seja, $M_o = 1$.



c. Cálculo da Mediana:

Como o total de frequências é 40 (número par), os valores centrais são o 20º e o 21º elementos, cuja média aritmética simples definirá a Mediana.

Sendo assim, colocando os dados em ordem crescente, obtem-se:

- 8 valores correspondentes a “0” irmão;
- 15 valores correspondentes a “1” irmão; e assim por diante.

(0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, **1, 1**, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3)

Então, o 20º e o 21º valores serão, ambos, 1 irmão.

Logo, a média aritmética entre os 2 valores é:

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

Portanto, a Mediana é 1.



5.6 Média Aritmética, Moda e Mediana a partir de uma Distribuição de Frequências Com Intervalo de Classe:

Acompanhe o seguinte exemplo:

Exemplo:

Consideremos a tabela abaixo, correspondente a uma Distribuição de Frequências Com Intervalos de Classe, referente à pesquisa sobre a massa corporal (em quilogramas) de um grupo de alunos de um pelotão do Curso de Comunicações da ESA, para o cálculo da Média Aritmética, da Moda e da Mediana, a saber:

<i>i</i>	Massa Corporal (kg)	Frequência Absoluta (f_i)
1	40 ----- 44	1
2	44 ----- 48	3
3	48 ----- 52	7
4	52 ----- 56	6
5	56 ----- 60	3
Total	-	$\sum f_i = 20$

Antes de iniciar os cálculos, deve-se, a partir da tabela inicial acima, anexar uma nova coluna, com seus respectivos Pontos Médios (x_i). Sendo assim, observe:

<i>i</i>	Massa Corporal (kg)	Frequência Absoluta (f_i)	Ponto Médio (x_i)
1	40 ----- 44	1	42
2	44 ----- 48	3	46
3	48 ----- 52	7	50
4	52 ----- 56	6	54
5	56 ----- 60	3	58
Total	-	$\sum f_i = 20$	-

Agora, pode-se efetuar os devidos cálculos.



a. Cálculo da Média Aritmética:

O cálculo da Média Aritmética para dados agrupados com intervalos de classe é:

$$\bar{x} = \frac{1.42 + 3.46 + 7.50 + 6.54 + 3.58}{1 + 3 + 7 + 6 + 3} = \frac{42 + 138 + 350 + 324 + 174}{20} = \frac{1028}{20} = 51,40$$

Logo, a Média Aritmética é de 51,40.

b. Cálculo da Moda:

Das várias fórmulas utilizadas para o cálculo aproximado da Moda em uma distribuição de frequências, a mais utilizada é a **Fórmula de Czuber**, a saber:

$$M_0 = L + h \cdot \frac{d_1}{d_1 + d_2}$$

Onde:

- M_0 é o valor da Moda;
- L é limite inferior da classe modal (classe que possui a maior frequência);
- h é o intervalo de classe;
- d_1 é a diferença absoluta entre as frequências das classes modal e pré-modal; e
- d_2 é a diferença absoluta entre as frequências das classes modal e pós-modal.

Sendo assim, de acordo com a tabela fornecida, pode-se extrair os seguintes dados para o cálculo:



i	Massa Corporal (kg)	Frequência Absoluta (f_i)
1	40 ----- 44	1
2	44 ----- 48	3
3	48 ----- 52	7
4	52 ----- 56	6
5	56 ----- 60	3
Total	-	$\sum f_i = 20$

- Classe Modal: 48 |----- 52, pois é a classe que possui a maior frequência ($f_3 = 7$);
- L é limite inferior da classe modal: $L = 48$;
- h é o intervalo de classe: $h = 52 - 48 = 4$;
- d_1 : diferença entre as frequências das classes modal e pré-modal: $d_1 = 7 - 3 = 4$; e
- d_2 : diferença entre as frequências das classes modal e pós-modal: $d_2 = 7 - 6 = 1$.

Substituindo tais valores na Fórmula de Czuber, obtem-se:

$$M_0 = L + h \cdot \frac{d_1}{d_1 + d_2}$$

$$M_0 = 48 + 4 \cdot \frac{4}{4 + 1}$$

$$M_0 = 48 + \frac{16}{5} = \frac{256}{5} = 51,20$$

Logo, a moda é de 51,20.



c. Cálculo da Mediana:

Para dados agrupados, utiliza-se a seguinte fórmula:

$$M_d = L + h \cdot \frac{\frac{n}{2} - F}{f_m}$$

Onde:

- M_d é o valor da Mediana;
- L é limite inferior da classe mediana (classe que contém os valores centrais da amostra);
- h é o intervalo de classe;
- n é o número de observações da população ou amostra;
- F é a soma de frequências até a classe mediana, sem contar com a mesma; e
- f_m é a frequência da classe mediana.

Sendo assim, de acordo com a tabela fornecida, deve-se, primeiramente, criar a coluna das Frequências Acumuladas, a partir da qual extrai-se os seguintes dados para o cálculo:

i	Massa Corporal (kg)	Frequência Absoluta (f_i)	Frequência Acumulada (F_{Ac})
1	40 ----- 44	1	1
2	44 ----- 48	3	4
3	48 ----- 52	7	11
4	52 ----- 56	6	17
5	56 ----- 60	3	20
Total	-	$\Sigma f_i = 20$	-



- Classe Mediana: como o total de frequências é 20 (número par), os valores centrais são o 10º e o 11º elemento. Sendo assim, pela coluna da Frequência Acumulada, verifica-se que estes elementos encontrar-se-ão na 3ª Classe, de intervalo **48 |----- 52**, o qual denomina-se de Classe Mediana (ou Intervalo da Mediana);
- L é limite inferior da classe mediana: **48**;
- h é o intervalo de classe: $h = 52 - 48 = 4$;
- n é o número de observações da população ou amostra: **20**;
- F é a soma de frequências até a classe mediana, sem contar com a mesma: **4** e
- f_m é a frequência da classe mediana: **7**.

Substituindo tais valores na fórmula, obtém-se:

$$M_d = L + h \cdot \frac{\frac{n}{2} - F}{f_m}$$

$$M_d = 48 + 4 \cdot \frac{\frac{20}{2} - 4}{7}$$

$$M_d = 48 + 4 \cdot \frac{10 - 4}{7} = 48 + 4 \cdot \frac{6}{7} = 48 + \frac{24}{7} = 48 + 3,43 \cong 51,43$$

Logo, a mediana é de, aproximadamente, 51,43.



EXERCÍCIOS

Capítulo 6 – Medidas de Posição

1. (FGV 2010 - adaptado)

Seja um conjunto de quatro números cuja média aritmética simples é 2,5.

Se incluirmos o número 8 neste conjunto, quanto passará a ser a nova média aritmética simples?

- a) 3,6
- b) 2,5
- c) 4,5
- d) 2,0
- e) nda

Resposta: Letra a)

Solução:

Como ficou observado, a média aritmética é a soma de um conjunto de números dividida pela quantidade de elementos deste conjunto. Então obtem-se:

$$\bar{x} = \frac{Soma}{4} = 2,5$$

$$Soma = 2,5 \cdot 4 = 10$$

Ao incluir o número 8 neste conjunto de números, a soma destes mesmos passará de 10 para 18 e, como agora terão 5 números ao invés de 4, a média deles será 18 dividido por 5 que é igual a 3,6:

$$\bar{x} = \frac{Soma + 8}{4 + 1} = \frac{10 + 8}{4 + 1} = \frac{18}{5} = 3,6$$



2. (2016/IADES/PC-DF/PERITO CRIMINAL)

A média das idades dos 45 empregados de uma corporação é de 32 anos. Para os próximos meses, estão previstas as aposentadorias de cinco empregados cuja média de idades é de 62 anos. Considerando essa situação hipotética, é correto afirmar que, após a efetivação de todas as aposentadorias, a média das idades da corporação passará a ser a seguinte:

- a) 25,11 anos.
- b) 26 anos.
- c) 28,25 anos.
- d) 30,75 anos.
- e) 36 anos.

Resposta: Letra c)

Solução:

Chamando a soma das idades dos 45 empregados de x , e considerando que a média de suas idades é 32, tem-se:

$$\frac{x}{45} = 32$$

Logo, $x = 1440$.

Agora, chamando a soma das idades dos 5 empregados que aposentarão de y , e considerando que a média de suas idades é 62, tem-se:

$$\frac{y}{5} = 62$$

Logo, $y = 310$.



Por fim, após as 5 aposentadorias, tem-se $45 - 5 = 40$ funcionários, cuja soma das idades será de:

$$1440 - 310 = 1130$$

Logo, a nova média, que será chamada de z , será:

$$z = \frac{1130}{40} = 28,25.$$

3. (FGV)

A medida de um conjunto ordenado de dados, que divide este conjunto em duas partes de igual número de observações, denomina-se:

- a) média.
- b) moda.
- c) mediana.
- d) desvio padrão.
- e) variância.

Resposta: Letra c)

Solução:

Essa questão trata do conceito de mediana, que é uma medida de posição do centro da distribuição dos dados apresentados em rol.

4. (CESGRANRIO)

Para resolver **as três questões abaixo**, utilize as informações a seguir.

O número de ausentes, por sala, em um dos prédios de aplicação de certo exame de proficiência aplicado em certa região, segue a seguinte distribuição:



Ausência	Frequência
0	1
1	2
2	3
3	5
4	4
5	2
6	3
7	3
9	1
10	1

4.1. (CESGRANRIO)

O número total de salas de aplicação neste prédio foi:

- a) 5
- b) 10
- c) 11
- d) 25
- e) 47

Resposta: Letra d)

Solução:

O número total de salas é dado pelo somatório da frequência absoluta.

$$n = \sum f_i = 1 + 2 + 3 + 5 + 4 + 2 + 3 + 3 + 1 + 1 = 25$$



4.2. (CESGRANRIO)

Qual foi o número médio aproximado de ausentes por sala?

- a) 10,7
- b) 5,0
- c) 4,7
- d) 4,3
- e) 2,5

Resposta: Letra d)

Solução:

Para encontrar a média aritmética, deve-se multiplicar o valor de cada variável de estudo pela frequência e depois divide-se pelo somatório da frequência.

$$\bar{x} = \frac{0.1 + 1.2 + 2.3 + 3.5 + 4.4 + 5.2 + 6.3 + 7.3 + 9.1 + 10.1}{25}$$

$$\bar{x} = \frac{107}{25} = 4,28 \cong 4,3$$

4.3 (CESGRANRIO)

A mediana do número de ausentes neste prédio é

- a) 3,0
- b) 3,5
- c) 4,0
- d) 4,5
- e) 5,0

Resposta: Letra c)



Solução:

1ª maneira)

I - colocar os valores em rol, em ordem crescente:

0, 1, 1, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 5, 5, 6, 6, 6, 7, 7, 7, 9, 10

II - Como são 25 valores (n é ímpar), a mediana é o elemento central, logo deve-se fazer:

$$e = \frac{(n + 1)}{2} = \frac{(25 + 1)}{2} = \frac{26}{2} = 13^{\text{o}} \text{ elemento}$$

Como $e = 13^{\text{o}} \text{ elemento}$, então $M_d = 4$.

2ª maneira)

Utilizando-se do conhecimento de frequência acumulada.

Para encontrar o décimo terceiro termo, utiliza-se a tabela e cria-se mais uma coluna com a frequência acumulada.

Ausência	Frequência	Frequência Acumulada
0	1	1
1	2	3
2	3	6
3	5	11
4	4	15
5	2	17
6	3	20
7	3	23
9	1	21
10	1	25

Portanto, a linha da tabela indica que o valor de 4 ausências se repete da décima segunda posição até a décima quinta posição.

Como a mediana está na 13ª posição, pode-se inferir que a mediana são 4 ausências.

Logo, $M_d = 4$.



5. Num Curso de Educação Física, o instrutor do estabelecimento de ensino anotou a altura de um grupo de discentes.

Considerando que os valores medidos foram: 1,54 m; 1,67 m, 1,50 m; 1,65 m; 1,75 m; 1,69 m; 1,60 m; 1,55 m e 1,78 m, qual o valor da mediana das alturas desses alunos?

Resposta: 1,65 m

Solução:

Primeiramente, deve-se colocar os valores em ordem (Rol):

1,50; 1,54; 1,55; 1,60; 1,65; 1,67; 1,69; 1,75; 1,78

Como o conjunto é formado por 9 elementos, que é um número ímpar, então deve-se definir a posição da mediana, que será igual a:

$$M_d = \frac{9 + 1}{2} = 5$$

logo, será o 5º elemento.

Assim, admitindo o Rol:

1,50; 1,54; 1,55; 1,60; 1,65; 1,67; 1,69; 1,75; 1,78,

tem-se que $M_d = 1.65m$.

6. Calcule o valor da mediana da seguinte amostra de dados (32, 27, 15, 44, 15, 32).

Resposta: 29,5

Solução:

Primeiramente, deve-se colocar os valores em ordem (Rol):

15, 15, 27, 32, 32, 44



Como essa amostra é formada por 6 elementos, que é um número par, a mediana será igual a média dos elementos centrais, ou seja:

$$e = \frac{6}{2} = 3^{\circ} \text{ elemento}$$

Logo, $e + 1 = 4^{\circ} \text{ elemento}$.

Logo, $3^{\circ} \text{ elemento} = 27$ e $4^{\circ} \text{ elemento} = 32$.

$$M_d = \frac{27 + 32}{2} = \frac{59}{2} = 29,5$$

7. (CESPE - adaptado)

Um pequeno levantamento para melhorar a qualidade da produção de textos em um escritório colheu dados sobre o número de erros de digitação por página.

A tabela abaixo apresenta o resultado desse levantamento.

Número de erros por página	Frequência
0	10
1	20
2	10
3	5
4	5

A partir das informações apresentadas no texto, assinale a opção incorreta.

- a) O número mediano de erros de digitação por página é igual a 1 erro por página.
- b) A moda do número de erros de digitação é igual a 1 erro por página.
- c) A amplitude total do número de erros de digitação é igual a 4 erros por página.
- d) Em média, o número de erros por página é igual a 2.1.

Resposta: letra d)



Solução:

- a) Verdadeira.
- b) Verdadeira, pois o número de erros por página igual a 1 é o que apresenta a maior frequência.
- c) Verdadeira, pois $AT = 4 - 0 = 0$.
- d) Falsa, pois $\frac{0.10+1.20+2.10+3.5+4.5}{10+20+10+5+5} = \frac{75}{50} = 1,4$
-

8. (ENEM 2010)

O quadro seguinte mostra o desempenho de um time de futebol no último campeonato.

A coluna da esquerda mostra o número de gols marcados e a coluna da direita informa em quantos jogos o time marcou aquele número de gols.

Gols marcados	Quantidade de partidas
0	5
1	3
2	4
3	3
4	2
5	2
7	1

Se X, Y e Z são, respectivamente, a média, a mediana e a moda desta distribuição, então:

- a) $X = Y < Z$.
- b) $Z < X = Y$.
- c) $Y < Z < X$.
- d) $Z < X < Y$.
- e) $Z < Y < X$.

Resposta: letra e) $Z < Y < X$.



Solução:

I – Para o cálculo da média deve-se somar o número total de gols (obtido multiplicando o número de gols marcados pela quantidade de partidas) e dividindo pelo número de partidas.

- Número total de gols = $0.5 + 1.3 + 2.4 + 3.3 + 4.2 + 5.2 + 7.1 = 45$

- Número total de partidas = 20.

Logo, a média de gols será igual a:

$$\bar{x} = \frac{45}{20} = 2,25$$

II - Para encontrar o valor da moda, verifica-se a quantidade de gols mais frequente.

Nesse caso, as partidas que tiveram 2 gols foram as mais frequentes (ao todo, 4 partidas).

Portanto,

$$M_o = 2$$

III - A mediana será encontrada colocando os números de gols em ordem (Rol).

Como o número de jogos foi igual a 20 que é um valor par, deve-se calcular a média entre os dois valores centrais do Rol, assim tem-se:

0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 5, 5, 7

$$M_d = \frac{2 + 2}{2} = \frac{4}{2} = 2$$

Com esses resultados, sabe-se que:

X (média) = $\bar{x} = 2,25$

Y (mediana) = $M_d = 2$

Z (moda) = $M_o = 2$

Portanto, $Z < Y < X$, isto é $(2 < 2 < 2,25)$.



9. O quadro abaixo representa os dados referentes à Pesquisa sobre a massa corporal (em quilogramas) de um grupo de militares, para fins de avaliação do IMC (Índice de Massa Corpórea), relativos ao Teste de Aptidão Física (TAF). O valor da Média Aritmética é:

i	Massa Corporal (kg)	Frequência Absoluta (f_i)
1	50 ----- 54	3
2	54 ----- 58	7
3	58 ----- 62	4
4	62 ----- 66	2
5	66 ----- 70	4
Total	-	$\sum f_i = 20$

- a) 59,0
- b) 59,2
- c) 59,4
- d) 59,6
- e) nda

Resposta: Letra

Solução:

Antes de iniciar os cálculos, deve-se, a partir da tabela inicial acima, anexar uma nova coluna, com seus respectivos Pontos Médios. Sendo assim, obtém-se:

i	Massa Corporal (kg)	Frequência Absoluta (f_i)	Ponto Médio (x_i)
1	50 ----- 54	3	52
2	54 ----- 58	7	56
3	58 ----- 62	4	60
4	62 ----- 66	2	64
5	66 ----- 70	4	68
Total	-	$\sum f_i = 20$	-



Agora, pode-se efetuar o cálculo.

$$\bar{x} = \frac{3.52 + 7.56 + 4.60 + 2.64 + 4.68}{3 + 7 + 4 + 2 + 4} = \frac{156 + 392 + 240 + 128 + 272}{20} = \frac{1188}{20} = 59,4$$

10. (ESAG/TJ-MA – Analista Judiciário/2005)

A tabela a seguir apresenta a distribuição de frequência por classes de pontos obtidos por candidatos em um concurso.

<i>i</i>	Pontos	Frequência Absoluta (<i>f_i</i>)
1	39 ----- 50	8
2	50 ----- 61	19
3	61 ----- 72	10
4	72 ----- 83	5
5	83 ----- 94	3
TOTAL		$\sum = 45$

Determinando-se as estimativas da Mediana (por interpolação) e Moda (modelo de Czuber) chega-se aos resultados:

- a) Mediana = 58,68 pontos; Moda = 55,50 pontos.
- b) Mediana = 66,50 pontos; Moda = 61,00 pontos.
- c) Mediana = 58,39 pontos; Moda = 56,05 pontos.
- d) Mediana = 55,50 pontos; Moda = 56,11 pontos.

Resposta: Letra c)



Solução:

I – Para o cálculo da Mediana, para uma Distribuição de Frequências com Intervalos de Classe, acrescida da coluna da Frequência Acumulada, tem-se os seguintes dados:

<i>i</i>	Pontos	Frequência Absoluta (<i>f_i</i>)	Frequência Acumulada <i>F_{Ac}</i>
1	39 ----- 50	8	8
2	50 ----- 61	19	27
3	61 ----- 72	10	37
4	72 ----- 83	5	42
5	83 ----- 94	3	45
TOTAL		$\sum = 45$	-

- Classe Mediana: Como o total de frequências é 45 (número ímpar), o valor central é o 13º elemento, pois $(45 + 1)/2 = 46/2 = 13$. Sendo assim, pela coluna da Frequência Acumulada, verifica-se que este elemento encontrar-se-á na 2ª Classe, de intervalo 50 |----- 61, o qual denomina-se de Classe Mediana (ou Intervalo da Mediana);

- *L* é limite inferior da classe mediana: 50;

- *h* é o intervalo de classe: $h = 61 - 50 = 11$;

- *n* é o número de observações da população ou amostra: 45;

- *F* é a soma de frequências até a classe mediana, sem contar com a mesma: 8; e

- *f_m* é a frequência da classe mediana: 19.

Substituindo tais valores na fórmula, obtém-se:

$$M_d = L + h \cdot \frac{\frac{n}{2} - F}{f_m}$$

$$M_d = 50 + 11 \cdot \frac{\frac{45}{2} - 8}{19}$$



$$M_d = 50 + 11 \cdot \frac{14,5}{19}$$

$$M_d = 50 + 11 \cdot 0,763 = 50 + 8,393$$

$$M_d \cong 58,39 \text{ pontos.}$$

Logo, a mediana é de 58,39 pontos.

II – Para o cálculo da Moda, para uma Distribuição de Frequências com Intervalos de Classe, utiliza-se a fórmula de Czuber, com os seguintes dados:

<i>i</i>	Pontos	Frequência Absoluta (<i>f_i</i>)
1	39 ----- 50	8
2	50 ----- 61	19
3	61 ----- 72	10
4	72 ----- 83	5
5	83 ----- 94	3
TOTAL		$\sum = 45$

- Classe Modal: 50 |----- 61, pois é a classe que possui a maior frequência;
- *L* é limite inferior da classe modal: *L* = 50;
- *h* é o intervalo de classe: *h* = 61 – 50 = 11;
- *d*₁: diferença entre as frequências das classes modal e pré-modal: *d*₁ = 19 – 8 = 11;
- *d*₂: diferença entre as frequências das classes modal e pós-modal: *d*₂ = 19 – 10 = 9.

Substituindo tais valores na fórmula de Czuber, obtem-se:

$$M_0 = L + h \cdot \frac{d_1}{d_1 + d_2}$$

$$M_0 = 50 + 11 \cdot \frac{11}{11 + 9} = 50 + 11 \cdot \frac{11}{20}$$



$$M_0 = 50 + \frac{121}{20} = 50 + 6,05 = 56,05$$

Logo, a moda é de 56,05.



CAPÍTULO 7 – MEDIDAS DE DISPERSÃO

As Medidas de Tendência Central estudadas anteriormente revelam informações apenas sobre o tamanho do valor central de um conjunto de valores, no entanto não há informação a respeito de como estão dispersas as observações em torno desse valor central. Nessa senda, para a possibilidade de tomada de decisões, deve-se calcular as chamadas Medidas de Dispersão, que vão indicar, com algum detalhe, de que forma o conjunto das observações se comporta em função dos valores de centralidade.

6.1 Desvios

Desvio, em um conjunto de valores, é a “distância” de cada uma dessas informações até a média aritmética delas e é obtido subtraindo cada um dos valores de um conjunto de informações da média aritmética desse conjunto.

Exemplo:

Um aluno do Curso de Infantaria obteve as seguintes notas na Avaliação da disciplina de Metodologia de Pesquisa Científica:

Nota 1: 7,0;

Nota 2: 8,5;

Nota 3: 9,5; e

Nota 4: 9,0.

Portanto, em relação à Média Aritmética (\bar{x}), que é de $\bar{x} = \frac{7,0+8,5+9,5+9,0}{4} = \frac{34}{4} = 8,5$, tem-se

os seguintes desvios:

- Desvio 1: $7,0 - 8,5 = -1,5$;

- Desvio 2: $8,5 - 8,5 = 0$;

- Desvio 3: $9,5 - 8,5 = 1,0$; e

- Desvio 4: $9,0 - 8,5 = 0,5$.

Observações:

- O sinal dos Desvios indica se a nota tirada é menor ou maior que a média;
- A soma dos Desvios é igual a zero.



6.2 Variância

Variância é a soma dos quadrados dos Desvios dividida pelo número de observações do conjunto menos uma, em caso de amostras.

É representada por σ^2 .

Num conjunto de valores, é a “distância” de cada uma dessas informações até às suas respectivas médias aritméticas e é obtida subtraindo cada um dos valores de um conjunto de informações da média aritmética desse conjunto.

Para calcular a **Variância Populacional** de todos os valores de um conjunto, utiliza-se a seguinte fórmula:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Onde:

Σ : é o símbolo de Somatório;

σ^2 : variância;

x_i : valor analisado;

\bar{x} : média aritmética do conjunto; e

n : número de dados do conjunto.

Para definir seu valor, deve-se:

- Primeiramente, devemos calcular a média aritmética do conjunto;
- Em seguida, subtrair de cada valor a média aritmética encontrada e elevar o resultado ao quadrado;
- Por fim, para obter a variância, soma-se todos os valores e divide-se pelo número de dados.

Obs: quando o conjunto de dados é muito extenso e se quer utilizar uma variável aleatória, deve-se empregar a fórmula da **Variância Amostral**:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$



Exemplo:

Os dados a seguir representam a quantidade de bússolas disponíveis numa reserva de armamento de um curso, nos seus 5 pelotões componentes: 6, 5, 8, 5 e 6. A variância da quantidade de bússolas desse curso é de quanto?

Solução:

Para definir seu valor, deve-se:

a. Primeiramente, devemos calcular a média aritmética do conjunto:

$$\bar{x} = \frac{6 + 5 + 8 + 5 + 6}{5} = \frac{30}{5} = 6$$

b. Em seguida, subtrair de cada valor a média aritmética encontrada e elevar o resultado ao quadrado:

$$- (6 - 6)^2 = 0^2 = 0$$

$$- (5 - 6)^2 = (-1)^2 = 1$$

$$- (8 - 6)^2 = (-2)^2 = 4$$

$$- (5 - 6)^2 = (-1)^2 = 1$$

$$- (6 - 6)^2 = 0^2 = 0$$

c. Por fim, para obtermos a variância, soma-se todos os valores e divide-se pelo número de dados:

$$\sigma^2 = \frac{0 + 1 + 4 + 1 + 0}{5} = \frac{6}{5} = 1,2$$



6.3 Desvio Padrão

Desvio Padrão é uma medida que expressa o grau de dispersão de um conjunto de dados, ou seja, ele indica o quanto um conjunto de dados é uniforme.

É representado por “ σ ”.

Quanto mais próximo de 0 for o desvio padrão, mais homogêneo são os dados.

É uma Medida de Dispersão que apresenta a vantagem sobre a Variância por permitir uma interpretação direta da variação do conjunto de dados, uma vez que é expresso na mesma unidade que a variável (kg, cm, etc...), permitindo uma melhor análise e comparação.

Como o desvio padrão é expresso pela raiz quadrada da variância, basta que seja extraída a raiz do resultado calculado pela fórmula anterior:

$$DP = \sigma = \sqrt{\sigma^2}$$

Para o cálculo do Desvio Padrão, deve-se utilizar a seguinte fórmula:

$$DP = \sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Sendo,

Σ : é o símbolo de somatório ;

x_i : valor analisado;

\bar{x} : média aritmética do conjunto; e

n : número de dados do conjunto.



Exemplo:

No 1º pelotão do Curso de Engenharia, as cinco melhores notas referentes à Avaliação de Raciocínio Lógico e Estatístico, foram 8, 5, 9, 6 e 7. Calcule o Desvio Padrão considerando-se uma população.

Solução:

Como o desvio padrão é expresso pela raiz quadrada da variância, para definir seu valor deve-se:

a. Primeiramente, calcular a média aritmética do conjunto:

$$\bar{x} = \frac{8 + 5 + 9 + 6 + 7}{5} = \frac{35}{5} = 7$$

b. Em seguida, subtrair de cada valor a média aritmética encontrada e elevar o resultado ao quadrado:

- $(8 - 7)^2 = (1)^2 = 1$
- $(5 - 7)^2 = (-2)^2 = 4$
- $(9 - 7)^2 = (2)^2 = 4$
- $(6 - 7)^2 = (-1)^2 = 1$
- $(7 - 7)^2 = 0^2 = 0$

c. Por fim, para obter a variância, soma-se todos os valores e divide-se pelo número de dados:

$$\sigma^2 = \frac{1 + 4 + 4 + 1 + 0}{5} = \frac{10}{5} = 2$$

$$DP = \sigma = \sqrt{\sigma^2} = \sqrt{2} = 1,4$$



EXERCÍCIOS

Capítulo 7 – Medidas de Dispersão

I. (CESGRANRIO - adaptado)

A tabela abaixo, representa as frequências acumuladas das idades de 20 jovens entre 14 e 20 anos:

Idades (anos)	Frequência Acumulada
14	2
15	4
16	9
17	12
18	15
19	18
20	20

Sabendo-se que m é a média aritmética dessas idades, qual a Variância das idades na população formada pelos 20 jovens?

- a) 0,15
- b) 0,20
- c) 1,78
- d) 3,20
- e) 3,35

Resposta: Letra d)

Solução:

Para que se possa calcular a variância, que é a média do quadrado dos desvios, deve-se encontrar, primeiramente, a média aritmética.

A tabela forneceu a frequência acumulada, logo é necessário encontrar suas respectivas frequências absolutas:



Idades (anos)	Frequência Absoluta	Frequência Acumulada
14	2	2
15	2	4
16	5	9
17	3	12
18	3	15
19	3	18
20	2	20

Sendo assim, temos que a Média Aritmética será de:

$$\bar{x} = \frac{2.14 + 2.15 + 5.16 + 3.17 + 3.18 + 3.19 + 2.20}{2 + 2 + 5 + 3 + 3 + 3 + 2} = \frac{340}{20} = 17$$

Agora, utilizando a fórmula da Variância, obtem-se:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

$$\sigma^2 = \frac{2.(14 - 17)^2 + 2.(15 - 17)^2 + 5.(16 - 17)^2 + 3.(17 - 17)^2 + 3.(18 - 17)^2 + 3.(19 - 17)^2 + 2.(20 - 17)^2}{20}$$

$$\sigma^2 = \frac{2.(-3)^2 + 2.(-2)^2 + 5.(-1)^2 + 3.(0)^2 + 3.(1)^2 + 3.(2)^2 + 2.(3)^2}{20}$$

$$\sigma^2 = \frac{2.9 + 2.4 + 5.1 + 3.0 + 3.1 + 3.4 + 2.9}{20}$$

$$\sigma^2 = \frac{18 + 8 + 5 + 0 + 3 + 12 + 18}{20}$$

$$\sigma^2 = \frac{64}{20}$$

$$\sigma^2 = 3,2$$



2. (FUNCAB/2013)

Em estatística, existe um conceito que é definido, para um conjunto de dados, como “a média dos quadrados das diferenças dos valores em relação à sua média”.

Essa definição conceitual:

- a) amplitude.
- b) desvio-padrão.
- c) coeficiente de variação.
- d) variância.
- e) estimador

Resposta: Letra d)

Solução:

É o próprio conceito de Variância visto na aula.

3. (ENEM – 2016)

O Procedimento de perda rápida de "peso" é comum entre os atletas dos esportes de combate. Para participar de um torneio, quatro atletas da categoria até 66 kg, Peso-Pena, foram submetidos a dietas balanceadas e atividades físicas e realizaram três "pesagens" antes do início do torneio.

Pelo regulamento do torneio, a primeira luta deverá ocorrer entre o atleta mais regular e o menos regular quanto aos "pesos".

As informações com base nas pesagens dos atletas estão no quadro.

Atleta	1ª pesagem (kg)	2ª pesagem (kg)	3ª pesagem (kg)	Média	Mediana	Desvio Padrão
I	78	72	72	72	72	4,90
II	83	65	71	71	65	8,49
III	75	70	70	70	70	4,08
IV	80	71	73	73	77	7,87



Após as três "pesagens", os organizadores do torneio informaram aos atletas quais deles se enfrentariam na primeira luta.

A primeira luta foi entre os atletas:

- a) I e III.
- b) I e IV.
- c) II e III.
- d) II e IV.
- e) III e IV

Resposta: Letra c)

Solução:

Para encontrar os atletas mais regulares deve-se usar o desvio padrão, pois essa medida indica o quanto que o valor desviou da média.

O atleta III é o com menor desvio padrão (4,08), logo é o mais regular.

O menos regular é o atleta II, com maior desvio padrão (8,49).

Portanto, a alternativa correta é a letra c) II e III.

4. (VUNESP/2014)

Um casal tem 4 filhos com idades de 1, 6, 8 e 15 anos, respectivamente. O desvio padrão das idades dos filhos é de, aproximadamente:

- a) 625
- b) 15
- c) 14
- d) 7
- e) 5

Resposta: Letra e)



Solução:

O Desvio padrão é a raiz quadrada da Variância.

I - A Média Aritmética é de:

$$\bar{x} = \frac{1 + 6 + 8 + 15}{4} = \frac{30}{4} = 7,5$$

Calculando a Variância, obtem-se:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Logo,

$$\sigma^2 = \frac{(1 - 7,5)^2 + (6 - 7,5)^2 + (8 - 7,5)^2 + (15 - 7,5)^2}{4}$$

$$\sigma^2 = \frac{(-6,5)^2 + (-1,5)^2 + (0,5)^2 + (7,5)^2}{4}$$

$$\sigma^2 = \frac{42,25 + 2,25 + 0,25 + 56,25}{4}$$

$$\sigma^2 = \frac{101}{4}$$

$$\sigma^2 = 25,25$$

Agora, como o Desvio Padrão é a raiz quadrada da Variância, tem-se:

$$DP = \sigma = \sqrt{\sigma^2}$$

$$\sigma = \sqrt{25,25}$$

$$\sigma \cong 5.$$



APÊNDICE A

MEDIDAS DE SEPARATRIZES

Medidas de Separatrizes são valores que separam o rol (os dados ordenados) em quatro (quartis), dez (decis) ou em cem (percentis) partes iguais.

1.1 Quartis (Q_i)

São valores que dividem o conjunto de dados ordenados (rol) em 4 (quatro) partes iguais.

- a. Primeiro Quartil (Q_1): valor situado de tal modo na série de dados que 25% das observações são menores que ele e 75% são maiores.
- b. Segundo Quartil (Q_2): valor situado de tal modo na série de dados que 50% das observações são menores que ele e 50% são maiores.
- c. Terceiro Quartil (Q_3): valor situado de tal modo na série de dados que 75% das observações são menores que ele e 25% são maiores.

25	25	25	25
Q_1		Q_2	Q_3

1.2 Decis (D_i)

Decis (D_1) são valores que dividem o conjunto de dados ordenados (rol) em 10(dez) partes iguais.

- a. Primeiro Decil (D_1): valor situado de tal modo na série de dados que 10% das observações são menores que ele e 90% são maiores.
- b. Segundo Decil (D_2): valor situado de tal modo na série de dados que 20% das observações são menores que ele e 80% são maiores.

[...]

- c. Nono Decil (D_9): valor situado de tal modo na série de dados que 90% das observações são menores que ele e 10% são maiores.



1.3 Centis (C_i)

Centis (C_1) são valores que dividem o conjunto de dados ordenados (rol) em 100 (cem) partes iguais.

a. Primeiro Percentil (P_1): valor situado de tal modo na série de dados que 1% das observações são menores que ele e 99% são maiores.

b. Segundo Percentil (P_2): valor situado de tal modo na série de dados que 2% das observações são menores que ele e 98% são maiores.

[...]

c. Nono Percentil (P_3): valor situado de tal modo na série de dados que 2% das observações são menores que ele e 98% são maiores.

[...]

d. Nonagésimo Nono Percentil (P_{99}): valor situado de tal modo na série de dados que 99% das observações são menores que ele e 1% são maiores.

1.4 Procedimentos para os cálculos dos Quartis para dados não agrupados.

Método prático: os três Quartis podem ser calculados utilizando-se a definição de Mediana:

- (Q_1): mediana da primeira metade dos elementos da série;
- (Q_2): mediana de todos os elementos da série;
- (Q_3): mediana da segunda metade dos elementos da série.

Exemplo 1: (para n ímpar)

Dado o conjunto de valores: 7, 13, 5, 12, 16, 4, 9, 15, 6. Calcule os Quartis (Q_1), (Q_2) e (Q_3).

1º passo: Colocar os dados em Rol: 4, 5, 6, 7, 9, 12, 13, 15, 16.

2º passo: Calcular (Q_1), (Q_2) e (Q_3).



a. Cálculo de (Q_1) :

- Basta calcular a mediana da 1ª metade do conjunto:

1ª metade do Rol	2ª metade do Rol
4, 5, 6, 7,	12, 13, 15, 16

Logo, Q_1 será a média aritmética entre os dois valores centrais:

$$(Q_1) = \frac{5 + 6}{2} = 5,5$$

b. Cálculo de (Q_2) :

Como (Q_2) é a mediana de todos os elementos da série (**n é ímpar**), então $e = \frac{9+1}{2} = 5^\circ$ elemento. Assim, analisando o Rol abaixo, tem-se que o 5º elemento corresponde ao dado 9.

Rol
4, 5, 6, 7, 9, 12, 13, 15, 16

Logo, $Q_2 = 9$.

c. Cálculo de (Q_3) :

- Basta calcular a mediana da 2ª metade do conjunto:

1ª metade do Rol	2ª metade do Rol
4, 5, 6, 7	12, 13, 15, 16

Logo,

$$(Q_3) = \frac{13 + 15}{2} = 14$$



Exemplo 2: (para n par)

Dado o conjunto de valores: 3, 11, 4, 16, 19, 2, 9, 10, 8, 12. Calcule os Quartis (Q_1), (Q_2) e (Q_3):

1º passo: Colocar os dados em Rol: 2, 3, 4, 8, 9, 10, 11, 12, 16, 19.

2º passo: Calcular (Q_1), (Q_2) e (Q_3).

a. Cálculo de (Q_1):

- Basta calcular a mediana da 1ª metade do conjunto:

1ª metade do Rol	2ª metade do Rol
2, 3, 4, 8, 9	10, 11, 12, 16, 19

Logo, Q_1 será o valor central da 1ª metade do Rol:

$$(Q_1) = 4$$

b. Cálculo de (Q_2):

Como (Q_2) é a mediana de todos os elementos da série (**n é par**), então devemos-se efetuar a média aritmética entre os elementos centrais (e) e ($e + 1$), ou seja, $e = \frac{n}{2} = \frac{10}{2} = 5^\circ$ elemento e ($e + 1$) = 6º elemento. Assim, analisando o Rol abaixo, obtem-se:

Rol
2, 3, 4, 8, 9, 10, 11, 12, 16, 19

Logo,

$$Q_2 = \frac{9 + 10}{2} = 9,5$$



c. Cálculo de (Q_3) :

- Basta calcular a mediana da 2ª metade do conjunto:

1ª metade do Rol	2ª metade do Rol
2, 3, 4, 8, 9	10, 11, 12, 16, 19

Logo,

$$(Q_3) = 12$$

1.5 Procedimentos para os cálculos dos Quartis para dados agrupados sem intervalo de classe.

No cálculo de quartis para dados agrupados sem intervalo de classe, utilizam-se técnicas semelhantes àquelas do cálculo da mediana. Isso significa que a localização da posição do quartil na série é verificada pela Frequência Acumulada. Vejamos o exemplo:

Exemplo:

Um curso preparatório para carreiras militares conta com turmas destinadas aos Concursos Públicos da ESA, EPCAr e EsPCEEx, cujos alunos têm as idades discriminadas na tabela abaixo:

Classes (i)	Idade (x_i)	Frequência Absoluta (f_i)
1	11	2
2	12	4
3	13	3
4	14	4
5	15	3
6	16	5
7	17	3
-	TOTAL	$n = \sum f_i = 24$

Calcule o valor dos Quartis (Q_1) , (Q_2) e (Q_3) e, em seguida, interprete-os.



Solução:

1º passo: Para posicionar os quartis, deve-se acrescentar à tabela a coluna das Frequências Acumuladas:

Classes (i)	Idade (x_i)	Frequência Absoluta (f_i)	Frequência Acumulada (F_{Ac})
1	11	3	3
2	12	4	7
3	13	3	10
4	14	4	14
5	15	3	17
6	16	5	22
7	17	3	25
-	TOTAL	$n = \sum f_i = 25$	

2º passo: Identificação da localização dos Quartis (Q_1), (Q_2) e (Q_3), pela fórmula

$$Q_k = \frac{k \cdot n}{4}$$

Onde:

k: representação do quartil (com $k = 1, 2$ ou 3); e

n: número de elementos da amostra.

- Posição do Quartil (Q_1):

$$Q_1 = \frac{1 \cdot 25}{4} = 6,25$$

- Posição do Quartil (Q_2):

$$Q_2 = \frac{2 \cdot 25}{4} = 12,50$$

- Posição do Quartil (Q_3):

$$Q_3 = \frac{3 \cdot 25}{4} = 18,75$$

3º Passo: Identificar as classes que contém os Quartis desejados, por meio da Frequência Acumulada (F_{Ac}).



Logo, pode-se identificar:

- Quartil Q_1 : 2ª Classe, pois observa-se que a localização da posição 6,25 ocorre na coluna da Frequência Acumulada correspondente à 2ª classe (F_{Ac} de 4 a 7).
- Quartil Q_2 : 4ª Classe, pois observa-se que a localização da posição 12,50 ocorre na coluna da Frequência Acumulada correspondente à 4ª classe (F_{Ac} de 11 a 14).
- Quartil Q_3 : 6ª Classe, pois observa-se que a localização da posição 18,75 ocorre na coluna da Frequência Acumulada correspondente à 6ª classe (F_{Ac} de 18 a 22).

4º passo: Identificar a variável correspondente à classe obtida:

- Na segunda classe, o valor da variável é 12 anos, ou seja:

$$Q_1 = 12 \text{ anos}$$

- Na quarta classe, o valor da variável é 14 anos, logo:

$$Q_2 = 14 \text{ anos}$$

- Na sexta classe, o valor da variável é 16 anos, isto é:

$$Q_3 = 16 \text{ anos}$$

Interpretação:

a. Primeiro quartil (Q_1): 25% dos alunos têm idade menor ou igual a 12 anos (Q_1) e 75% tem idade maior ou igual a 12 anos.

25%		75%	
25%	25%	25%	25%
$Q_1 = 12 \text{ anos}$	Q_2	Q_3	



b. Segundo quartil (Q_2): 50% dos alunos têm idade menor ou igual a 14 anos (Q_1) e 50% tem idade maior ou igual a 14 anos.

50%		50%	
25%	25%	25%	25%
Q_1	$Q_2=14 \text{ anos}$	Q_3	

c. Terceiro quartil (Q_3): 75% dos alunos têm idade menor ou igual a 16 anos (Q_1) e 25% tem idade maior ou igual a 16 anos.

75%			25%
25%	25%	25%	25%
Q_1	Q_2	$Q_3 = 16 \text{ anos}$	

1.6 Procedimentos para os cálculos dos Quartis, Decis e Centis para dados agrupados com intervalo de classe.

1º Passo) Determinar as Frequências Acumuladas (F_{Ac}) da distribuição.

2º Passo) Calcular a posição do Quartil, Decil ou Percentil desejado, conforme cada caso.

- Para a definição do Quartil, sendo ($i = 1, 2, 3$):

$$P_i = \frac{i \cdot \sum f_i}{4}$$

- Para a definição do Decil, sendo ($i = 1, 2, 3, 4, 5, 6, 7, 8, 9$):

$$P_i = \frac{i \cdot \sum f_i}{10}$$



- Para a definição do Percentil, sendo ($i = 1, 2, 3, \dots, 99$):

$$P_i = \frac{i \cdot \sum f_i}{100}$$

3º Passo) Identificar a que classe que contém o Quartil, Decil ou Percentil desejado, por meio da Frequência Acumulada (F_{Ac}).

4º Passo) Calcular o Quartil, Decil ou Percentil desejado por meio de uma das fórmulas:

- Para o Quartil:

$$Q_i = l_i + \left(\frac{P - f_{ant}}{f_Q} \right) \cdot h$$

Onde:

l_i : limite inferior da classe do quartil;

f_{ant} : frequência acumulada anterior à classe do quartil;

f_Q : frequência simples da classe do quartil; e

h : amplitude de classe do quartil.

- Para o Decil:

$$D_i = l_i + \left(\frac{P - f_{ant}}{f_D} \right) \cdot h$$

Onde:

l_i : limite inferior da classe do decil;

f_{ant} : frequência acumulada anterior à classe do decil;

f_Q : frequência simples da classe do decil; e

h : amplitude de classe do decil.



- Para o Centil:

$$C_i = l_i + \left(\frac{P - f_{ant}}{f_c} \right) \cdot h$$

Onde:

l_i : limite inferior da classe do percentil;

f_{ant} : frequência acumulada anterior à classe do percentil;

f_c : frequência simples da classe do percentil; e

h : amplitude de classe do percentil.

Exemplo:

Os números de diárias pagas a 160 militares de um batalhão, em razão da participação numa Operação de Garantia da Lei e da Ordem (GLO) estão distribuídos conforme a tabela a seguir:

i	Número de Diárias	Número de Militares (f_i)
1	1 ----- 3	20
2	3 ----- 5	40
3	5 ----- 7	60
4	7 ----- 9	30
5	9 ----- 11	10
TOTAL		$\sum = 40$

Calcule o Q_1 , D_4 e C_{85} e interprete os resultados.

Solução:

Seguindo o item 7.4 acima (Procedimentos para os cálculos dos Quartis, Decis e Centis), tem-se:

1º Passo) Determinar as Frequências Acumuladas (F_{Ac}) da distribuição.



i	Número de Diárias	Número de Militares (f_i)	Frequência Acumulada (F_{Ac})
1	1 ----- 3	20	20
2	3 ----- 5	40	60
3	5 ----- 7	60	120
4	7 ----- 9	30	150
5	9 ----- 11	10	160
TOTAL		$\sum = 40$	-

2º Passo) Calcular a posição do Quartil, Decil ou Percentil desejado, conforme cada caso.

- Cálculo da posição de Q_1 , ou seja P_1 :

$$P_i = \frac{i \cdot \sum f_i}{4}$$

$$P_1 = \frac{1 \cdot 160}{4} = 40^{\text{º}} \text{ elemento.}$$

- Cálculo de D_4 , ou seja, P_4 :

$$P_i = \frac{i \cdot \sum f_i}{10}$$

$$P_4 = \frac{4 \cdot 160}{10} = 64^{\text{º}} \text{ elemento.}$$

- Cálculo de C_{85} , ou seja, P_{85} :

$$P_i = \frac{i \cdot \sum f_i}{100}$$

$$P_{85} = \frac{85 \cdot 160}{100} = 136^{\text{º}} \text{ elemento.}$$

3º Passo) Identificar a que classe que contém o Quartil, Decil ou Percentil desejado por meio da Frequência Acumulada (F_{Ac}).

Logo, identificamos:

- Quartil: 2ª Classe
- Decil: 3ª Classe
- Centil: 4ª Classe



4º Passo) Calcular o Quartil, Decil ou Percentil desejado por meio de uma das fórmulas:

I - Para o Quartil, com $P_1 = 40^{\text{o}}$ elemento (obtido no 2º passo) e com a 2ª Classe (identificada no 3º passo), obtem-se:

$$Q_i = l_i + \left(\frac{P - f_{ant}}{f_Q} \right) \cdot h$$

Onde:

l_i : limite inferior da classe do quartil = 3;

f_{ant} : frequência acumulada anterior à classe do quartil = 20;

f_Q : frequência simples da classe do quartil = 40; e

h : amplitude de classe do quartil = $5 - 3 = 2$.

i	Número de Diárias	Número de Militares (f_i)	Frequência Acumulada (F_{Ac})
1	1 ----- 3	20	20
2	3 ----- 5	40	60
3	5 ----- 7	60	120
4	7 ----- 9	30	150
5	9 ----- 11	10	160
TOTAL		$\sum = 40$	-

Logo,

$$Q_1 = 3 + \left(\frac{40 - 20}{40} \right) \cdot 2 = 4 \text{ diárias.}$$



II - Para o Decil, com $P_1 = 64^{\text{o}}$ elemento (obtido no 2º passo) e com a 3ª Classe (identificada no 3º passo), obtem-se:

$$D_i = l_i + \left(\frac{P - f_{ant}}{f_D} \right) \cdot h$$

Onde:

l_i : limite inferior da classe do decil = 5;

f_{ant} : frequência acumulada anterior à classe do decil = 60;

f_D : frequência simples da classe do decil = 60; e

h : amplitude de classe do decil = 7-5=2.

i	Número de Diárias	Número de Militares (f_i)	Frequência Acumulada (F_{Ac})
1	1 ----- 3	20	20
2	3 ----- 5	40	60
3	5 ----- 7	60	120
4	7 ----- 9	30	150
5	9 ----- 11	10	160
TOTAL		$\sum = 40$	-

Logo,

$$D_4 = 5 + \left(\frac{64 - 60}{60} \right) \cdot 2 = 5,13 \text{ diárias}$$



III - Para o Centil, com $P_1 = 136^{\text{o}}$ elemento (obtido no 2º passo) e com a 4ª Classe (identificada no 3º passo), tem-se:

$$C_i = l_i + \left(\frac{P - f_{ant}}{f_c} \right) \cdot h$$

Onde:

l_i : limite inferior da classe do percentil = 7;

f_{ant} : frequência acumulada anterior à classe do percentil = 120;

f_c : frequência simples da classe do percentil = 30; e

h : amplitude de classe do percentil = $9 - 7 = 2$.

i	Número de Diárias	Número de Militares (f_i)	Frequência Acumulada (F_{Ac})
1	1 ----- 3	20	20
2	3 ----- 5	40	60
3	5 ----- 7	60	120
4	7 ----- 9	30	150
5	9 ----- 11	10	160
TOTAL		$\sum = 40$	-

Logo,

$$C_i = l_i + \left(\frac{P - f_{ant}}{f_c} \right) \cdot h$$

$$C_{85} = 7 + \left(\frac{136 - 120}{30} \right) \cdot 2 = 8,07 \text{ diárias}$$

Portanto, de acordo com os valores obtidos, pode-se interpretar:

- 25% dos militares receberam até 4 diárias ou 75% dos militares receberam mais de 4 diárias;
- 40% dos militares receberam até 5,13 diárias ou 60% dos militares receberam mais de 5,13 diárias;
- 85% dos militares receberam até 8,07 diárias ou 15% dos militares receberam mais de 8,07 diárias.



EXERCÍCIOS EXTRAS

Apêndice – Medidas de Separatrizes

1. Na fabricação de um tipo de material de uso militar,

i	Pesos das peças (kg)	Frequência Absoluta (f_i)
1	0 ----- 5	52
2	5 ----- 10	36
3	10 ----- 15	30
4	15 ----- 20	41
5	20 ----- 25	28
6	25 ----- 30	25
7	30 ----- 35	18
-	-	$\sum f_i = 230$

Qual o valor do quartil Q_1 e interprete o resultado?

- a) 5,78
- b) 5,76
- c) 5,74
- d) 5,72
- e) 5,70

Solução:

Seguindo o item 7.4 do Livro (Procedimentos para os cálculos dos Quartis, Decis e Centis), tem-se:

1º Passo) Determinar as Frequências Acumuladas (F_{Ac}) da distribuição:



i	Pesos das peças (kg)	Frequência Absoluta (f_i)	Frequência Acumulada (F_{Ac})
1	0 ----- 5	52	52
2	5 ----- 10	36	88
3	10 ----- 15	30	118
4	15 ----- 20	41	159
5	20 ----- 25	28	187
6	25 ----- 30	25	212
7	30 ----- 35	18	230
-	-	$\sum f_i = 230$	-

2º Passo) Calcular a posição do Quartil desejado.

- Cálculo da posição de Q_1 , ou seja P_1 :

$$P_i = \frac{i \cdot \sum f_i}{4}$$

$$P_1 = \frac{1 \cdot 230}{4} = 57,5^\circ \text{ elemento.}$$

3º Passo) Identificar a que classe que contém o Quartil, por meio da Frequência Acumulada (F_{Ac}).

Logo, identifica-se:

- Quartil: 2ª Classe: **5 |----- 10**

4º Passo) Calcular o Quartil desejado por meio de sua fórmula:

I - Para o Quartil, com $P_1 = 57,5^\circ$ elemento (obtido no 2º passo) e com a 2ª Classe (identificada no 3º passo), tem-se:

$$Q_i = l_i + \left(\frac{P - f_{ant}}{f_Q} \right) \cdot h$$



Onde:

l_i : limite inferior da classe do quartil = 5;

f_{ant} : frequência acumulada anterior à classe do quartil = 52;

f_Q : frequência simples da classe do quartil = 36; e

h : amplitude de classe do quartil = $10 - 5 = 5$.

i	Pesos das peças (kg)	Frequência Absoluta (f_i)	Frequência Acumulada (F_{Ac})
1	0 ----- 5	52	52
2	5 ----- 10	36	88
3	10 ----- 15	30	118
4	15 ----- 20	41	159
5	20 ----- 25	28	187
6	25 ----- 30	25	212
7	30 ----- 35	18	230
-	-	$\sum f_i = 230$	-

Logo,

$$Q_1 = 5 + \left(\frac{57,5 - 52}{36} \right) \cdot 5 = 5 + \left(\frac{5,5}{36} \right) \cdot 5 = 5 + (0,15) \cdot 5 = 5 + 0,76 = 5,76 \text{ kg}$$

Interpretação do resultado: 25% das peças (do lote analisado) têm peso menor que 5,76 kg (Q_1).

Isto significa que os 75% restantes têm peso acima de 5,76 kg.



REFERÊNCIAS

BUSSAB, Wilton de O; MORETTIN, Pedro A. **Estatística Básica**. 6ª edição. 2010. Editora Saraiva.

CURSO de Estatística Básica – Módulo 1. Centro Científico Conhecer, 2020. Disponível em: <http://www.conhecer.org.br/download/estatistica>. Acesso em 26 jul. 2021.

FONSECA, Sammys César Chaves. **Fundamentos de Estatística**. Centro Educacional de Anápolis – GO. Universidade Federal de Mato Grosso (UFMT). 2015.

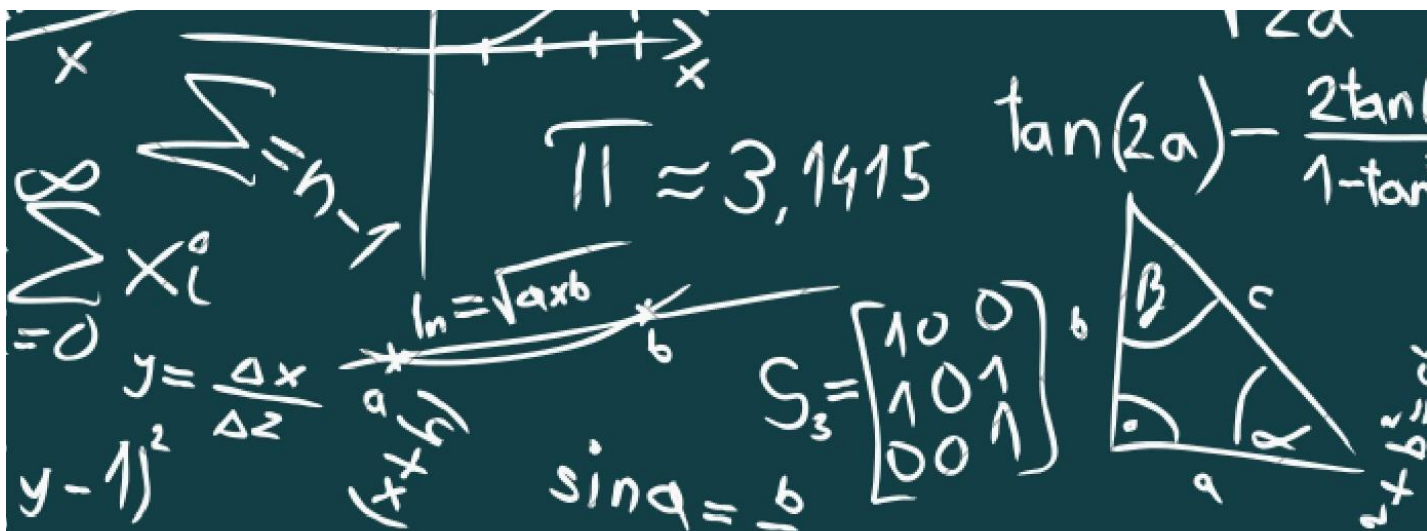
FREITAS, Corina da Costa; RENNÓ, Camilo Daleles; JÚNIOR, Manoel Araújo Souza. **Estatística – Curso 1**. Ministério de Ciência e Tecnologia. Instituto Nacional de Pesquisa Espaciais (INPE). São José dos Campos –SP. 2003.

ROSA, Daniele Toniolo Dias F. **Fundamentos de Estatística**. Universidade Tecnológica Federal do Paraná.

Disponível em: <http://www.fotoacustica.fis.ufba.br/daniele/planejamento/Aula2.pdf>. Acesso em: 26 jun 21.

SANTOS, Alcione Miranda dos Santos. **Introdução à Estatística**. Núcleo de Estatística e Informática. Disponível em:

<https://docente.ifsc.edu.br/gianpaulo.medeiros/MaterialDidatico/M%c3%a9todos%20Est%c3%a1tisticos/>. Acesso em: 19 jul. 2021



ESTATÍSTICA

ESTA OBRA FOI ESCRITA POR DOCENTES DA ESCOLA DE SARGENTOS DAS ARMAS (ESA), RESPONSÁVEIS PELAS DISCIPLINAS DE RACIOCÍNIO LÓGICO E ESTATÍSTICO E METODOLOGIA DE PESQUISA NO ENSINO SUPERIOR E TEM POR OBJETIVO PROPORCIONAR AO LEITOR O FUNDAMENTO INTRODUTÓRIO NECESSÁRIO PARA O ESTUDO DA ESTATÍSTICA, ALÉM DE SUPRIR UMA DEFICIÊNCIA BIBLIOGRÁFICA DE TEXTOS VOLTADOS AO COTIDIANO DA CASERNA.

SENDO ASSIM, ESTE LIVRO ABORDA ALGUNS ASPECTOS INTRODUTÓRIOS, OS CONCEITOS BÁSICOS E AS FASES DO TRABALHO ESTATÍSTICO, APÓS, SÃO ABORDADOS ASSUNTOS REFERENTES A TABELAS, REPRESENTAÇÕES GRÁFICAS E ÀS MEDIDAS DE POSIÇÃO, DISPERSÃO E, COMO APÊNDICE, SEPARATRIZES.

BIBLIOGRAFIA DOS AUTORES:

- DINALVA FERREIRA DA SILVA - ESPECIALISTA EM GERENCIAMENTO DE PROJETOS - PUC, LICENCIADA EM MATEMÁTICA - UNIFRAN E BACHAREL EM CIÊNCIAS ECONÔMICAS - PUC
- DIONE APARECIDO FERREIRA DA SILVA: MESTRE EM ENSINO DA MATEMÁTICA - UFRJ, ESPECIALISTA EM METODOLOGIA DO ENSINO DA MATEMÁTICA - FAEL;
- EDUARDO LUINI DA SILVA - MESTRE EM MATEMÁTICA - UFGG, ESPECIALISTA EM GESTÃO ESCOLAR: SUPERVISÃO, ORIENTAÇÃO E INSPEÇÃO ESCOLAR - IF SUL DE MINAS; E
- THAMARA MARQUES RODRIGUES: MESTRE EM MODELAGEM COMPUTACIONAL E SISTEMAS - UNIMONTES, ESPECIALISTA EM DIDÁTICA E METODOLOGIA DO ENSINO SUPERIOR - UNIMONTES.

ESA

