



# **Estatística Aplicada**

**Tronco Comum**

**2019**

***ENSINO ONLINE. ENSINO COM FUTURO***

## **Direitos de autor (copyright)**

Este manual de Estatística Aplicada é propriedade do Instituto Superior de Ciências e Educação a Distância (ISCED), e contém reservados todos os direitos. É proibida a duplicação ou reprodução parcial ou total deste manual, sob quaisquer formas ou por quaisquer meios (electrónicos, mecânico, gravação, fotocópia ou outros), sem permissão expressa de entidade editora (Instituto Superior de Ciências e Educação a Distância (ISCED)).

A não observância do acima estipulado o infractor é passível a aplicação de processos judiciais em vigor no País.

**Instituto Superior de Ciências e Educação a Distância (ISCED)**  
Coordenação do Programa de Licenciaturas  
Rua Dr.Lacerda de Almeida N° 211, 1º andar, Beira Ponta - Gea

Beira - Moçambique  
Telefone: +25823323501  
Cel: +258823055839  
Fax: 23323501  
Website: [www.isced.ac.mz](http://www.isced.ac.mz)

## **Agradecimentos**

Instituto Superior de Ciências e Educação a Distância, Coordenação do Programa das licenciaturas e o autor que elaborou o presente manual, agradecem a colaboração dos seguintes indivíduos e instituições na elaboração deste manual:

Pela Coordenação

Direcção Académica do ISCED

Pelo Design

Direcção de Qualidade e Avaliação do ISCED.

Financiamento e Logística

IAPED – Instituto Africano para a Promoção da Educação a Distância.

Pela Revisão final

Msc. Zacarias Mendes Magibire

### **Elaborado Por:**

Dr. Horácio Manuel Vunga – *Doutorando* em Ciências de Educação com Especialização em Inovação Currículo, pela Universidade Jean Piaget de Moçambique e Licenciado em Ensino de Matemática e Física Pela Universidade Pedagógica.

# Índice

Bem-vindo à Estatística .....	7
Objectivos do curso .....	7
Quem deveria estudar este módulo .....	8
Acerca dos ícones.....	10
Habilidades de estudo .....	11
Precisa de apoio? .....	13
Tarefas (avaliação e auto-avaliação) .....	14

## **TEMA I: ESTATÍSTICA BÁSICA** **TEMA II: PROBABILIDADES (Noções)**

---

### **TEMA I: ESTATÍSTICA BÁSICA**

#### **Unidade Temática N<sup>o</sup> 1.1. 10**

---

Estatística e sua aplicação.....	10
Introdução .....	10
Sumário .....	10
Exercícios .....	14

#### **Unidade Temática N<sup>o</sup> 1.2. 15**

---

As etapas do método estatístico .....	15
Introdução .....	15
Sumário .....	15
Exercícios .....	18

#### **Unidade Temática N<sup>o</sup> 1.3. 19**

---

População, Amostra, Senso e Sondagem. ....	19
Introdução .....	19
Sumário .....	19
Exercícios .....	28

#### **Unidade Temática N<sup>o</sup> 1.4. 31**

---

Sumário .....	31
Exercícios .....	39

#### **Unidade Temática N<sup>o</sup> 1.5. 42**

---

Frequências acumuladas e gráficos .....	42
Introdução .....	42

Sumário .....	42
Exercícios .....	46
<b>Unidade Temática N<sup>o</sup> 1.6.</b>	<b>48</b>
Somatórios.....	48
Introdução .....	48
Sumário .....	48
Exercícios .....	51
<b>Unidade Temática N<sup>o</sup> 1.7.</b>	<b>48</b>
Medidas de Posição ou tendência central (dados não agrupados).....	48
Introdução .....	48
Sumário .....	48
Exercícios .....	52
<b>Unidade Temática N<sup>o</sup> 1.8.</b>	<b>53</b>
Medidas de posição (dados agrupados em classes da mesma posição).....	53
Introdução .....	53
Sumário .....	53
Exercícios .....	56
<b>Unidade Temática N<sup>o</sup> 1.9.</b>	<b>58</b>
Outros tipos de médias.....	58
Introdução .....	58
Sumário .....	58
Exercícios .....	60
<b>Unidade Temática N<sup>o</sup> 1.10.</b>	<b>62</b>
Outras medidas de localização - quantis, quartis, decís e percentis.....	62
Introdução .....	62
Sumário .....	62
Exercícios .....	68
<b>Unidade Temática N<sup>o</sup> 1.11.</b>	<b>69</b>
Medidas de Dispersão ou variabilidade.....	69
Introdução .....	69
Sumário .....	69
Exercícios .....	76
<b>Unidade Temática N<sup>o</sup> 1.12.</b>	<b>79</b>
Medidas de Assimetria ou Curtose.....	79
Introdução .....	79

Sumário .....	79
Exercícios .....	83
<b>TEMA II: PROBABILIDADES (Noções)</b>	
<b>Unidade Temática N<sup>o</sup> 2.1.</b>	<b>85</b>
<hr/>	
Distribuições bidimensionais. ....	85
Introdução .....	85
Sumário .....	85
Exercícios .....	90
<b>Unidade Temática N<sup>o</sup> 2.2</b>	<b>94</b>
<hr/>	
Noção intuitiva e frequentista das probabilidades. ....	94
Introdução .....	94
Sumário .....	94
Exercícios .....	101
<b>Unidade Temática N<sup>o</sup> 2.3.</b>	<b>102</b>
<hr/>	
Acontecimentos equiprováveis. ....	102
Introdução .....	102
Sumário .....	102
Exercícios .....	104
<b>Unidade Temática N<sup>o</sup> 2.4.</b>	<b>107</b>
<hr/>	
Probabilidade condicional .....	107
Introdução .....	107
Sumário .....	107
Exercícios .....	116

## Visão geral

---

### Benvindo ao Módulo de Estatística

Etimologicamente a **Estatística** foi definida como a Ciência das coisas que **pertencem ao Estado** ou da coisa Pública.

Estatística é uma ciência exacta que visa fornecer subsídios ao analista para colectar, organizar (rol dedados), resumir, analisar, intepretar e apresentar dados. Trata de parâmetros extraídos da população, tais como média, frequência, mediana, variância, desvio padrão, etc.

A estatística fornece métodos ou técnicas para extrair informação de dados, que nos dão informação útil sobre o problema em estudo, sendo assim, é objectivo da Estatística extrair informação dos dados para obter uma melhor compreensão das situações que representam.

Quando se aborda uma problemática envolvendo métodos estatísticos, estes devem ser utilizados mesmo antes de se recolher a amostra, isto é, deve-se planejar a experiência que nos vai permitir recolher os dados, de modo que, posteriormente, se possa extrair o máximo de informação relevante para o problema em estudo, ou seja para a população de onde os dados provêm.

Quando já se está na posse dos dados, procura-se agrupá-los e reduzi-los, sob forma de amostra, deixando de lado a aleatoriedade presente. Seguidamente o objectivo do estudo estatístico pode ser o de estimar uma quantidade ou testar uma hipótese, utilizando-se técnicas estatísticas convenientes, as quais realçam toda a potencialidade da Estatística, na medida em que vão permitir tirar conclusões acerca de uma população, baseando-se numa pequena amostra, dando-nos ainda uma medida do erro cometido.

---

### Objectivos deste Módulo/Disciplina

Constituem objectivos deste Módulo, principalmente, os seguintes:

- Ensinar a extrair informação dos dados para obter uma melhor informação e compreensão das situações que representam.
- Ensinar a estimar uma quantidade ou testar uma hipótese, utilizando-se técnicas estatísticas convenientes, as quais realçam toda a potencialidade da Estatística.
- Habilitar e/ou potenciar aos investigadores e demais interessados, a tomar decisões fundamentadas, pelo que, adequadas a cada caso.

---

### **Quem deveria estudar este módulo**

Este Módulo de Estatística, por ser de Tronco Comúm, foi concebido para estudantes do 1º ano dos cursos de licenciatura do ISCED. Poderá ocorrer, contudo, que haja leitores que queiram se actualizar e consolidar seus conhecimentos nessa disciplina, esses serão bem vindos, não sendo necessário para tal se inscrever. Mas poderá adquirir o manual.

---

## Como está estruturado este Módulo

O módulo de Estatística, para estudantes do 1º ano dos cursos de licenciaturas do ISCED, está estruturado como se segue:

### Páginas introdutórias

- **Um índice** completo.
- Uma **visão geral detalhada** dos conteúdos do módulo, resumindo os aspectos-chave que você precisa conhecer para melhor estudar. Recomendamos vivamente que leia esta secção com atenção antes de começar o seu estudo, como componente de habilidades de estudos.

### Conteúdo desta Disciplina / módulo

Este módulo está estruturado em Temas. Cada tema, por sua vez comporta certo número de unidades temáticas ou simplesmente unidades,. Cada unidade temática se caracteriza por conter uma introdução, objetivos, conteúdos.

No final de cada unidade temática ou do próprio tema, são incorporados antes o sumário, exercícios de auto-avaliação, só depois é que aparecem os exercícios de avaliação.

Os exercícios de avaliação têm as seguintes características: Puros exercícios teóricos/Práticos, Problemas não resolvidos e actividades práticas algumas incluído estudo de caso.

### Outros recursos

A equipa dos académicos e pedagogos do ISCED, pensando em si, num cantinho, recôndito deste nosso vasto Moçambique e cheio de dúvidas e limitações no seu processo de aprendizagem, apresenta uma lista de recursos didácticos adicionais ao seu módulo para você explorar. Para tal o ISCED disponibiliza na biblioteca do seu centro de recursos mais material de estudos relacionado com o seu curso como: Livros e/ou módulos, CD, CD-ROOM, DVD. Para além deste material físico ou electrónico disponível na biblioteca, pode ter acesso a Plataforma digital moodle para alargar mais ainda as possibilidades dos seus estudos.

## **Auto-avaliação e Tarefas de avaliação**

Tarefas de **auto-avaliação** para este módulo encontram-se no final de cada unidade temática e de cada tema. As tarefas dos exercícios de auto-avaliação apresentam duas características: primeiro apresentam exercícios resolvidos com detalhes. Segundo, exercícios que mostram apenas respostas.

Tarefas de **avaliação** devem ser semelhantes às de auto-avaliação mas sem mostrar os passos e devem obedecer o grau crescente de dificuldades do processo de aprendizagem, umas a seguir a outras. Parte das tarefas de avaliação será objecto dos trabalhos de campo a serem entregues aos tutores/doceentes para efeitos de correcção e subsequentemente nota. Também constará do exame do fim do módulo. Pelo que, caro estudante, fazer todos os exercícios de avaliação é uma grande vantagem.

## **Comentários e sugestões**

Use este espaço para dar sugestões valiosas, sobre determinados aspectos, quer de natureza científica, quer de natureza didáctico-Pedagógica, etc, sobre como deveriam ser ou estar apresentadas. Pode ser que graças as suas observações que, em gozo de confiança, classificamo-las de úteis, o próximo módulo venha a ser melhorado.

---

## **Ícones de actividade**

Ao longo deste manual irá encontrar uma série de ícones nas margens das folhas. Estes ícones servem para identificar diferentes partes do processo de aprendizagem. Podem indicar uma parcela específica de texto, uma nova actividade ou tarefa, uma mudança de actividade, etc.

### **Acerca dos ícones**

Os ícones usados neste manual são símbolos africanos, conhecidos por adrinka. Estes símbolos têm origem no povo Ashante de África Ocidental, datam do século XVII e ainda se usam hoje em dia.

Pode ver o conjunto completo de ícones deste manual já a seguir, cada um com uma descrição do seu significado e da forma como nós interpretámos esse significado para representar as várias actividades ao longo deste módulo.

---

## Habilidades de estudo

Caro estudante, procure olhar para si mesmo em três dimensões, nomeadamente: o lado social, profissional e estudante. Daí ser importante planificar muito bem o seu tempo de estudos.

O principal **objectivo** deste campo é o de ensinar aprender a aprender. Afinal aprender aprende-se.

Durante a formação e desenvolvimento de competências, para facilitar a aprendizagem e alcançar melhores resultados, implicará empenho, dedicação e disciplina nos estudos. Isto é, os bons resultados apenas se conseguem com estratégias eficientes e eficazes. Por isso é importante saber **como, onde e quando** estudar. Apresentamos algumas sugestões com as quais esperamos que caro estudante possa rentabilizar o tempo dedicado aos estudos, procedendo como se segue:

1º Praticar a leitura. Aprender a Distância exige alto domínio de leitura.

2º Fazer leitura diagonal aos conteúdos (leitura corrida).

3º Voltar a fazer leitura, desta vez para a compreensão e assimilação crítica dos conteúdos (ESTUDAR).

4º Fazer seminário (debate em grupos), para comprovar se a sua aprendizagem confere ou não com a dos colegas e com o padrão.

5º Fazer TC (Trabalho de Campo), algumas actividades práticas ou as de estudo de caso se existirem.

**IMPORTANTE:** Em observância ao triângulo **modo-espaço-tempo**, respectivamente **como, onde e quando**...estudar, como foi referido no início deste item, antes de organizar os seus momentos de estudo reflecta sobre o ambiente de estudo que seria ideal para si: Estudo melhor em casa/biblioteca/café/outro lugar? Estudo melhor à noite/de manhã/de tarde/fins de semana/ao longo da semana? Estudo melhor com música/num sítio sossegado/num sítio barulhento!? Preciso de intervalo em cada 30 minutos, em cada hora, etc.

É impossível estudar numa noite tudo o que devia ter sido estudado durante um determinado período de tempo; Deve estudar cada ponto da matéria em profundidade e passar só ao seguinte quando achar que já domina bem o anterior.

Privilegia-se saber bem (com profundidade) o pouco que puder ler e estudar, que saber tudo superficialmente! Mas a melhor opção é juntar o útil ao agradável: Saber com profundidade todos conteúdos de cada tema, no módulo.

Dica importante: não recomendamos estudar seguidamente por tempo superior a uma hora. Estudar por tempo de uma hora intercalado por 10 (dez) a 15 (quinze) minutos de descanso (chama-se descanso à mudança de actividades). Ou seja que durante o intervalo não se continuar a tratar dos mesmos assuntos das actividades obrigatórias.

Uma longa exposição aos estudos ou ao trabalho intelectual obrigatório, pode conduzir ao efeito contrário: baixar o rendimento da aprendizagem. Por que o estudante acumula um elevado volume de trabalho, em termos de estudos, em pouco tempo, criando interferência entre os conhecimento, perde sequência lógica, por fim ao perceber que estuda tanto mas não aprende, cai em insegurança, depressão e desespero, por se achar injustamente incapaz!

Não estude na última da hora; quando se trate de fazer alguma avaliação. Aprenda a ser estudante de facto (aquele que estuda sistematicamente), não estudar apenas para responder a questões de alguma avaliação, mas sim estude para a vida, sobre tudo, estude pensando na sua utilidade como futuro profissional, na área em que está a se formar.

Organize na sua agenda um horário onde define a que horas e que matérias deve estudar durante a semana; Face ao tempo livre que resta, deve decidir como o utilizar produtivamente, decidindo quanto tempo será dedicado ao estudo e a outras actividades.

É importante identificar as ideias principais de um texto, pois será uma necessidade para o estudo das diversas matérias que compõem o curso: A colocação de notas nas margens pode ajudar a estruturar a matéria de modo que seja mais fácil identificar as partes que está a estudar e Pode escrever conclusões, exemplos, vantagens, definições, datas, nomes, pode também utilizar a margem para colocar comentários seus relacionados com o que

está a ler; a melhor altura para sublinhar é imediatamente a seguir à compreensão do texto e não depois de uma primeira leitura; Utilizar o dicionário sempre que surja um conceito cujo significado não conhece ou não lhe é familiar;

Lembre-se que o teu sucesso depende da sua entrega, você é o responsável pela sua própria aprendizagem e cabe a ti planificar, organizar, gerir, controlar e avaliar o seu próprio progresso.

---

## **Precisa de apoio?**

Caro estudante, temos a certeza que por uma ou por outra razão, o material de estudos impresso, lhe pode suscitar algumas dúvidas como falta de clareza, alguns erros de concordância, prováveis erros ortográficos, falta de clareza, fraca visibilidade, páginas trocadas ou invertidas, etc). Nestes casos, contacte os serços de atendimento e apoio ao estudante do seu Centro de Recursos (CR), via telefone, sms, E-mail, se tiver tempo, escreva mesmo uma carta participando a preocupação.

Uma das atribuições dos Gestores dos CR e seus assistentes (Pedagógico e Administrativo), é a de monitorar e garantir a sua aprendizagem com qualidade e sucesso. Dai a relevância da comunicação no Ensino a Distância (EAD), onde o recurso as TIC se torna incontornável: entre estudantes, estudante – Tutor, estudante – CR, etc.

As sessões presenciais são um momento em que você caro estudante, tem a oportunidade de interagir fisicamente com staff do seu CR, com tutores ou com parte da equipa central do ISCED indigetada para acompanhar as sua sessões presenciais. Neste período pode apresentar dúvidas, tratar assuntos de natureza pedagógica e/ou admibistrativa.

O estudo em grupo, que está estimado para ocupar cerca de 30% do tempo de estudos a distância, é muita importância, na medida em que permite lhe situar, em termos do grau de aprendizagem com relação aos outros colegas. Desta maneira ficar'a a saber se precisa de apoio ou precisa de apoiar aos colegas. Desenvolver habito de debater assuntos relacionados com os conteúdos programáticos, constantes nos diferentes temas e unidade temática, no módulo.

---

## Tarefas (avaliação e auto-avaliação)

O estudante deve realizar todas as tarefas (exercícios, actividades e auto-avaliação), contudo nem todas deverão ser entregues, mas é importante que sejam realizadas. As tarefas devem ser entregues duas semanas antes das sessões presenciais seguintes.

Para cada tarefa serão estabelecidos prazos de entrega, e o não cumprimento dos prazos de entrega, implica a não classificação do estudante. Tenha sempre presente que a nota dos trabalhos de campo conta e é decisiva para ser admitido ao exame final da disciplina/módulo.

Os trabalhos devem ser entregues ao Centro de Recursos (CR) e os mesmos devem ser dirigidos ao tutor/docente.

Podem ser utilizadas diferentes fontes e materiais de pesquisa, contudo os mesmos devem ser devidamente referenciados, respeitando os direitos do autor.

O plágio<sup>1</sup> é uma violação do direito intelectual do(s) autor(es). Uma transcrição à letra de mais de 8 (oito) palavras do texto de um autor, sem o citar é considerado plágio. A honestidade, humildade científica e o respeito pelos direitos autorais devem caracterizar a realização dos trabalhos e seu autor (estudante do ISCED).

---

## Avaliação

Muitos perguntam: Com é possível avaliar estudantes na modalidade de Ensino a Distância (EAD), estando eles fisicamente separados e muito distantes do docente/tutor!? E nós afirmamos: Sim é muito possível, talvez seja uma avaliação mais fiável e concistente.

Você será avaliado durante os estudos à distância que contam com um mínimo de 90% do total de tempo que precisa de estudar os conteúdos do seu módulo. Quando o

---

<sup>1</sup> Plágio - copiar ou assinar parcial ou totalmente uma obra literária, propriedade intelectual de outras pessoas, sem prévia autorização.

tempo de contacto presencial conta com um máximo de 10%) do total de tempo do módulo. A avaliação do estudante consta detalhada do regulamentada de avaliação.

Os trabalhos de campo por si realizaos, durante estudos e aprendizagem no campo, pesam 25% e servem para a nota de frequência para ir aos exames.

Os exames são realizados no final da cadeira disciplina ou modulo e decorrem durante as sessões presenciais. Os exames pesam no mínimo 75%, o que adicionado aos 25% da média de frequência, determinam a nota final com a qual o estudante conclui a cadeira.

A nota de 10 (dez) valores é a nota mínima de conclusão da cadeira.

Nesta cadeira o estudante deverá realizar pelo menos 2 (dois) trabalhos e 1 (um) (exame).

Algumas actividades praticas, relatórios e reflexões serão utilizados como ferramentas de avaliação formativa.

Durante a realização das avaliações, os estudantes devem ter em consideração a apresentação, a coerência textual, o grau de cientificidade, a forma de conclusão dos assuntos, as recomendações, a identificação das referências bibliográficas utilizadas, o respeito pelos direitos do autor, entre outros.

Os objectivos e critérios de avaliação constam do Regulamento de Avaliação.

## TEMA I. ESTATÍSTICA

## TEMA II: PROBABILIDADE (Noções)

---

### Unidade Temática: Estatística e sua aplicação

#### Introdução

Embora nas antigas civilizações (chinesa, egípcia e romana) já se fizessem inquéritos destinados a obter informações sobre as populações e as riquezas económicas, permitindo aos governantes fazer não só recrutamentos militares mas

também lançar impostos sobre as próprias populações, só bastante mais tarde (século XVIII) a palavra “ Estatística” foi usada pelo professor Godofredo Achenwal (economista alemão 1719 - 1772) da Universidade de Gottingen que a definiu como “a ciência das coisas que pertencem ao Estado”.

No entanto, o estudo da Estatística, com fundamentação matemática, só se conseguiria com a criação do Cálculo das Probabilidades e a sua aplicação aos fenómenos sociais. A estatística deixa de ser, então, um amontoado de dados para se transformar num instrumento de análise, de síntese e previsão das soluções nos casos mais diversos.

Com efeito, nos nossos dias, a sua utilização passou a ser imprescindível, quer a nível nacional e internacional quer a nível de empresas, quer a nível individual, procurando estudar situações e elaborar planos que permitem a tomada das decisões mais adequadas aos problemas apresentados.

### **O conceito estatística e a importância de estudar estatística**

A palavra **ESTATÍSTICA** vem do *STATUS* (Estados em Latim). Como é do conhecimento geral, a leitura de um simples jornal ou revista implica, hoje em dia, entender a linguagem dos gráficos e dos números.

Em todos os campos da actividade humana, a informação é essencial as decisões dos cidadãos, à vida das empresas, à sobrevivência dos estados.

Isto implica a profusão dos jornais, revista, livros e relatórios exibindo tabelas, mapas, gráficos, ... contendo variadíssima informação Estatística sobre os mais variados fenómenos e características da actividade de um país:

- Número de habitantes, de trabalhadores por profissão, de família com e sem casas, de pequenas, médias e grandes empresas, de importadores e exportadores.
- Distribuição de voto por região, reprovação e aprovação por nível e por disciplina, o nível de infecção de HIV/SIDA por região, número de professores e de escolas, etc.

Qualquer cidadão tem de ser capaz de compreender, tirar ilações, criticar e escolher o que lhe interessa, dessas informações que diariamente lhe chegam pelos meios de comunicação social.

Ao iniciar um estudo de natureza estatística deverá fazer-se o seguinte:

- Definir explicitamente o conjunto sobre o qual se vai fazer o inquérito, de tal forma que se possa dizer, sem ambiguidade, se um dado pertence ou não ao conjunto;
- Indicar com clareza os dados que se pretendem obter;
- Avaliar, na medida do possível, da veracidade dos dados recolhidos (de forma a não alterar os resultados da análise que se pretende);
- Ordenar convenientemente os dados recolhidos em tabelas, de utilização rápida e estudo simplificado.

Ao completar esta unidade, você será capaz de:



#### Objectivos

- Colaborar na resolução de problemas da comunidade em que se insere;
- Resolver problemas de natureza quantitativa no âmbito das ciências humanas;
- Comunicar com clareza, oralmente e por escrito;
- Avaliar e criticar afirmações de carácter estatístico;
- Interpretar e comparar distribuições estatísticas recorrendo às medidas de localização e de dispersão e gráficos;
- Indicar situações em que a estatística presta relevantes serviços, recorrer à calculadora para resolver problemas de estatísticas;

---

## Sumário

### Estatística descritiva e estatística Indutiva

Desde o século passado que a estatística constitui um ramo da Matemática estruturando cientificamente e em estreita ligação com o cálculo das probabilidades

Na estatística consideram-se dois ramos:

- **A Estatística Descritiva**, que visa descrever o real de forma a permitir entendê-lo melhor;

- **A Estatística Indutiva**, que a partir de uma amostra da população permite estender os resultados à população inteira.

### **Estatística Descritiva**

Neste módulo pretende-se apresentar os conceitos princípios relacionados com os métodos de recolha e apresentação de dados, bem como das medidas estatística próprias para análise e interpretação dos dados recolhidos. Em suma, pretende-se desenvolver cada um dos aspectos presentes na definição de Estatísticas Descritiva.

Consiste na recolha, apresentação, análise e interpretação de dados numéricos através da criação de instrumentos adequados: quadros, gráficos e indicadores numéricos.

Os métodos para recolher, classificar, sintetizar, apresentar e interpretar informação qualitativa constituem uma parte importante da teoria estatística; de facto, constituíram até a matéria, quase exclusiva, das primeiras obras desta área científica. Outros aspectos da teoria Estatística, cumpre assinalar, têm igual ou maior importância, como sejam os métodos de inferência Estatística que permitem retirar conclusões sobre um grupo determinado – população ou universo – a partir da informação recolhida para uma amostra.

Convém referir que o termo “ESTATÍSTICA” é utilizado para referir a dois conceitos diferentes, conforme se utiliza no singular ou no plural. Quando utilizado no plural, é sinónimo de factos ou dados numéricos, enquanto que no singular constitui um objecto de estudo, uma ciência, tal como o é a Matemática, e compreende, como foi referido anteriormente, um conjunto de princípios e métodos de recolha, classificação, síntese e apresentação de dados numéricos.

A utilidade da estatística pode ser resumida de seguinte modo:

- Permitir descrever e compreender relações entre variáveis: numa época em que a quantidade de informação aumenta tão rapidamente, os centros de decisão têm necessidade de se manterem actualizados e controlarem as grandes massas de dados com que são inundados quase diariamente; para tal é necessário que a informação lhe

seja apresentada de forma a possibilitar a sua interpretação imediata e a identificação das relações mais importantes;

- Permitir a tomada de melhores e mais rápidas decisões: porque é possível controlar mais informação num mais curto espaço de tempo;
- Facilitar a tomada de decisões para fazer face à mudança: num mundo em constante mudança, a planificação e a tomada de decisões deverá apoiar-se em bases sólidas, no conhecimento profundo das situações passadas e presentes e numa previsão fundamentada da evolução futura
- O objectivo da estatística descritiva é informar, prevenir e esclarecer.
- O campo de acção da estatística descritiva tem-se revelado quase ilimitado:
- Em diversos campos de carácter social sobre populações, as suas condições de vida, de trabalho e de saúde, de educação, de cultura.
- Em estudos de carácter económico, importações e exportações, consumo de turistas, emigração etc...
- Em estudos metrológicos (temperatura, precipitação...)
- Em estudos políticos (distribuição de votos por região, ..) etc.

### **Estatística Indutiva (ou Inferencial)**

Enquanto a estatística descritiva analisa todos os indivíduos de um dado conjunto e tira conclusões sobre esse conjunto no seu todo, a estatística indutiva trata de estabelecer conclusões relativas a um conjunto mais vasto de indivíduos (população) a partir da observação de uma parte dela (amostra) com base na estrutura matemática que lhe confere o cálculo das probabilidades.

Os resultados finais de uma eleição são objecto de estudo da Estatística Destrutiva.

As previsões feitas por ocasião das eleições, imediatamente depois do fecho das urnas, são feitas a partir de uma amostra utilizando a Estatística Indutiva.

A Estatística Indutiva desempenha um importante papel na investigação científica ,em diversos sectores como:

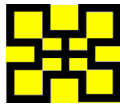
- A medicina, a farmácia, a sociologia, a psicologia, a química, a educação, a agricultura, a linguística, a biologia.

Permitindo estabelecer previsões sobre acontecimentos futuros, com margem de erro incrivelmente pequenas.

Os métodos estatísticos permitem hoje, em qualquer ciência, obter uma descrição da realidade física ou social e fornecem um meio de interpretação dessa realidade.

---

## Exercícios



Auto-avaliação

1. Como surgiu a Estatística?
2. Em quantas partes se divide a estatística e quais são as áreas de estudo de cada uma delas?
3. Qual a diferença entre os conceitos de estatística e estatísticas?
4. Porque razão é importante, no processo de tomada de decisões, recolher informações preliminares antes da definição específica do problema?
5. Diga em que domínio de estatísticas - descritiva ou inferência – incluiria as seguintes afirmações:
  - a) 30% dos estudantes de estatística do curso de gestão não conseguem fazer a cadeira em avaliação continua.
  - b) Os pneus da marca Rodamais duram 60.000 km.
  - c) Uma em cada dez empresas portuguesas tem dívidas à segurança social.

---

## Unidade Temática 1.2: As etapas do método

### Introdução

Nesta unidade vamos estudar as principais etapas do método estatístico pois elas são muito importantes para qualquer estudo ou investigação de natureza estatística.

Ao completar esta unidade, você será capaz de:



#### Objectivos

- Saber as etapas do método estatístico.
- Saber o objecto de estudo da estatística descritiva.

#### As etapas do método estatístico

O objectivo deste módulo é, precisamente, apresentar métodos estatísticos que satisfazem os três aspectos referidos e que se realizam nas diferentes etapas constantes da definição de estatística descritiva: recolha, apresentação e interpretação de dados.

São estas as etapas que definem o método estatístico de resolução de problemas:

#### 1. Identificação do problema ou situação

Deverá ser claro, desde o início do estudo, qual o problema a analisar e, uma vez conhecido, qual o tipo de decisões que se pretendem tomar. Esta etapa requer já algum conhecimento estatístico pois os métodos a aplicar não

são, de modo nenhum, independentes da informação que se pretende recolher. Uma identificação incorrecta do problema torna todas as etapas seguintes inúteis. Ainda nesta etapa e para tornar a identificação do problema mais fácil poderá ser utilizada alguma informação quantitativa já existente.

## 2. Recolha de dados

Uma vez identificado o problema, a etapa seguinte consiste na recolha dos dados necessários, apropriados, tão completos quanto possível e, sobretudo, pertinentes para a situação que se pretende analisar.

A recolha de toda a informação necessária pode ser feita directamente quando os dados são obtidos de fonte originária ou de forma indirecta quando os dados recolhidos provem já de uma recolha directa. Aos primeiros, que é possível encontrar em registos ou ficheiros, chamam-se dados primários enquanto que os valores não disponíveis nestas fontes e calculados a partir daquelas são dados secundários. Todos os dados resultantes de inquéritos feitos directamente a uma população ou a um grupo dessa população são dados primários. São ainda exemplos destes, todos os dados disponíveis nas estatísticas publicadas pelo I.N.E. – o número de nascimentos, casamentos e óbitos de cada região do ano de 2007, por exemplo, o número de desempregados em determinado sector de actividade económica, a distribuição das empresas do sector agrícola pela área de exploração, etc., etc. Dados secundários serão por exemplo, uma estimativa da esperança de vida à nascença nos valores observados nos últimos dez anos 2000 com base na inflação para o ano de 1995.

As fontes de dados podem ainda ser classificadas como **internas** ou **externas**. Por exemplo, os serviços de contabilidade, produção ou marketing de uma empresa constituem fontes internas de informação económica e comercial que deverão ser postas ao dispor dos órgãos de decisão da empresa. Informação externa à empresa é a proveniente dos organismos públicos, como o Governo, o Instituto Nacional de Estatística (I.N.E.) ou privados como os seminários económicos e revistas de especialidade.

No respeitante à periodicidade, a recolha dos dados pode ser classificada como:

- **Continua** -> Quando realizada permanentemente;
- **Periódica** -> Quando feita em intervalos de tempo;
- **Ocasional** -> Quando realizada de modo esporádico;

No processo de tomada de decisões na empresa, a todo momento nos deparamos com a necessidade de conhecer as características passadas e presentes da própria empresa e do seu meio envolvente: custos de produção, custos de aquisição da matéria-prima, custos de comercialização e publicidade, tempo de execução de determinadas tarefa, nível de escolaridade dos empregos, preços dos produtos concorrentes, procura de determinado produto, preferências dos consumidores, etc., etc.

Muitas vezes acontece não estar disponível toda a informação necessária ou porque não existe de todo ou porque se encontra desactualizada. Nestes casos, é necessário recolher nova informação, o que poderá ter vantagens e desvantagens. Vantagens porque permite uma definição precisa da informação a recolher e das suas escalas de medida, para que os dados verifiquem todas as propriedades necessárias para responder eficazmente ao problema em análise. Desvantagem, também, porque poderá tornar o estudo e a obtenção de resultados demasiados morosos e caros.

Existem vários métodos para recolha de nova informação. As entrevistas pessoais são uma prática corrente: um entrevistador faz ao inquirido perguntas retiradas, de preferência, de um questionário estruturado, e coloca as respostas nos espaços a elas reservadas. Mas as entrevistas podem também ser feita pelo telefone ou, quando por alguma razão se quer evitar a presença de um entrevistador, pode optar – se por enviar o questionário pelo correio.

### **3. Crítica dos dados**

Uma vez os dados recolhidos, quer sejam dados primários ou secundário é necessário proceder-se a uma revisão crítica de modo a suprimir valores estranhos ou eliminar erros capazes de provocar futuros enganos de apresentação e análise ou mesmo de enviesar as conclusões obtidos.

Esta crítica é tanto mais necessária quando toda ou parte da informação provem de fontes secundárias, sujeitas a erros de reprodução e que nem sempre explicitam como os dados foram recolhidos ou quais os limites à sua utilização.

### **4. Apresentação dos dados**

Após a recolha e a crítica, convém organizar os dados de maneira prática e racional, para um melhor entendimento do fenómeno que se pretende estudar.

Começa aqui o principal objectivo da Estatística Descritiva: criar os instrumentos necessários para classificar e apresentar conjuntos de dados numéricos de tal modo que a informação neles contida seja apreendida mais fácil e rapidamente.

O processo de classificação consiste na identificação de unidade de informação com características comuns e no seu agrupamento em classes. Para classificar é necessário utilizar um método, predeterminado, de codificação que torne possível a identificação abreviada das unidades de informação.

Uma vez classificados os dados, passa a ser possível sintetizar a informação neles contidos com a ajuda de quadros, gráficos e valores numéricos descritivos que ajudem a compreender a situação e a identificar relações importantes entre as variáveis.

## 5. Análise e interpretação de resultados

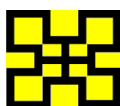
Por último é necessário interpretar os resultados encontrados. Esta interpretação estará tanto mais facilitada quando se tiverem concluído em etapas anteriores os instrumentos mais apropriados a representação de tipos de dados recolhidos.

Conclusões enviesadas podem ser propositadas ou não e ter diferentes causas. É suficientemente conhecido o exemplo de entidades que, para situações idênticas, retiram conclusões bastante divergentes: as taxas de infiltração e desemprego estimadas pelos órgãos governamentais e pelos sindicatos raramente coincidem. São exemplos de enviesamento propositado para servir fins políticos em que se torna difícil demonstrar, com rigor, qual delas está errada.

Mas, muitas vezes, o enviesamento não é propositado. Pode começar por ser o resultado de medidas de estatística descritiva pouco adequadas ao problema em causa, por diferentes escalas de medida ou ainda por bases de comparação pouco adequadas.

---

### Exercícios



Auto-avaliação

1. Quais as etapas do método estatístico de resolução de problemas?
2. Como classifica as fontes de dados estatísticos.
3. Como classifica a recolha de dados estatísticos no que respeita à periodicidade.
4. Quais os métodos de recolha de informação que conhece?
5. Em que consiste o processo de classificação dos dados?
6. Dê alguns exemplos de enviesamento dos resultados obtidos numa análise estatística.

---

## Unidade Temática 1.3: População, Amostra, Senso e Sondagem

### Introdução

Para trabalhar com segurança nos dados estatísticos, é preciso saber se está perante dados de uma amostra ou de uma população inteira. É nesta unidade onde podemos estudar as diferenças entre população e amostra, censo e sondagem.

Ao completar esta unidade, você será capaz de:



#### Objectivos

- Ter conhecimento dos conceitos população e amostra, censo e sondagem.
- Ser capaz de exemplificar cada conceito acima citado.

---

### População

É um conjunto de indivíduos ou objectos que apresentam pelo menos uma característica em comum. A população pode ser finita ou infinita. Na prática, quando uma população é finita, com um número grande de elementos, considera-se como população infinita.

### Amostra

Considerando-se a impossibilidade, na maioria das vezes, do tratamento de todos os elementos da população, retira-se uma parte da população (amostra).

Todo o subconjunto não vazio e com menor número de elementos do que o conjunto definido como população constitui, por definição, uma amostra dessa população.

Num estudo estatístico é sempre melhor usar uma população em vez de uma amostra, mas tal nem sempre é possível.

Algumas das causas que levam ao uso de uma amostra são:

- A população ser infinita;
- Economia de dinheiro e tempo;
- Comodidade (diminuição do número de documentos);
- Testes destrutivos (no estudo destroem-se os elementos, por exemplo: qualidade dos fósforos, de vinho, etc.).

É necessário ter muito cuidado na escolha da amostra. Se não for bem escolhida todo o estudo pode conduzir a conclusões erradas.

Na escolha de uma amostra deve ter-se em conta a imparcialidade, a representatividade e o tamanho.

As características da amostra devem aproximar-se tanto quanto possível da população.

No caso da população moçambicana, amostra deve conter por exemplo, indivíduos do norte, centro, e sul, do litoral e do interior, das cidades e do campo, homens e mulheres, jovens e adultos.

### **Censo**

Se todos os elementos da população são observados diz-se que se fez um levantamento exaustivo, ou recenseamento, ou apenas censo.

### **Sondagem**

Um estudo estatístico feito a partir de uma amostra chama-se sondagem ou amostragem.

As sondagens, estudos feitos a partir da amostra, são muito usadas, por serem mais fáceis de realizar, mais rápidas e mais económicas.

Alem disso, são indispensáveis quando a observação das unidades implica a sua destruição

## **Carácteres Estatísticos**

Chama-se carácter ou variável estatística a propriedade que vai ser estudada.

Entre os caracteres estatísticos temos:

- Carácteres quantitativos ou variáveis quantitativas – são os que podem medir ou referenciar usando números. Podem ser discretos ou contínuos.

O carácter ou a variável é discreta se o número de valores diferentes que podem tomar é finito, isto é, se só pode tomar valores isolados em pontos da recta real.

Exemplo<sub>1</sub>: número de erros em um livro, número de golos, número de irmãos, etc.

A variável é contínua se pode assumir teoricamente qualquer valor em certo intervalo da recta real.

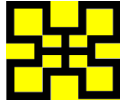
Exemplo<sub>2</sub>: Temperatura do ar, altura, peso de alunos, pressão arterial, etc. Pois teoricamente um indivíduo poderá ter como peso 50,5 kg, 50,572 kg, 50,585 kg ...

Carácteres qualitativos – são os que não podem medir ou referenciar numericamente.

Exemplo<sub>3</sub>: Nacionalidade, intenção de voto, profissão, raça, cor, sexo, grupo sanguíneo, ..

---

## Exercícios



### Auto-avaliação

1. Indique, nos casos seguintes, a população, o indivíduo, o carácter e o tipo de carácter em estudo.
  - a) Marca de cada um dos carros estacionados num parque num certo instante.
  - b) O número de espectadores com cartão-jovem num certo concerto.
  - c) O número de faltas de cada aluno de uma escola num trimestre.
  - d) A cor dos cabelos das senhoras com mais de 50 anos de idade.
2. Os estudos seguintes são feitos a partir de amostras.

Indica em cada caso, se as conclusões poderão ser válidas para toda a população em estudo. Justifica.

- a) Para investigar qual o desporto preferido pelos jovens entre 15 e 18 anos, inquiriram-se jovens dessas idades a saída de um jogo de basquetebol.
- b) Para estudar o interesse da população no programa televisivo Fama – Show entrevista-se 100 candidatos à academia do Fama no Caya – Cuanga.
- c) Para estimar quantos alunos de uma escola estariam interessados em usar a Internet inquiriram-se todos os alunos com o número de matrícula múltiplo de 5.

---

## Unidade Temática: Organização de dados e frequência

### Introdução

Na organização dos dados temos que ter em conta se vamos ou não agrupar em classes. Para os dois casos temos que saber os procedimentos para a composição das tabelas das frequências absolutas e relativas.

Ao completar esta unidade, você será capaz de:



#### Objectivos

- Saber compor a tabela das frequências absolutas e relativas.
- Saber determinados as amplitudes e o numero de classes nos dados agrupados.

---

### Sumário

#### Dados brutos e rol de uma amostra. Frequências absolutas e relativas

Como se observou anteriormente, a estatística tem como objectivo encontrar as leis de comportamento para todo o conjunto, por meio de sintetização dos dados numéricos, sob a forma de tabelas, gráficos e medidas. A seguir são apresentadas os procedimentos para a reprodução das distribuições de frequências.

As séries estatísticas são o resultado de recolha de dados que estão sujeitos no seu apuramento a determinadas leis:

### **Dados brutos**

O conjunto de dados numéricos obtidos após a crítica dos valores colectados constitui-se nos dados brutos, assim: 24, 23, 22, 28, 34, 35, 21, 23, 33, 34, 21, 25, 36, 26, 22, 30, 32, 25, 26, 33, 34, 21, 31, 25, 31, 26, 25, 35, 33, 31 são exemplos de dados brutos. (idades dos estudantes de uma turma).

### **Rol**

É o arranjo dos dados brutos em ordem crescente ou decrescente. Assim: 21, 21, 21, 22, 22, 23, 23, 24, 25, 25, 25, 25, 26, 26, 26, 28, 30, 31, 31, 32, 33, 33, 33, 34, 34, 34, 35, 35, 36.

### **Amplitude total “rengue” (R) ou (At)**

É a diferença entre o maior e o menor valores observados.

No exemplo dado,  $R = 36 - 21 = 15$  ou  $At = 36 - 21 = 15$

### **Frequência absoluta (fi)**

Frequência absoluta do valor  $x_i$  é o número de vezes que o elemento  $x_i$  aparece na amostra, ou é o número de elementos pertencentes a uma classe.

No exemplo anterior, 3 é a frequência absoluta do valor 21 ou  $f(21) = 3$ .

A soma de todas as frequências absolutas é igual a população.

$$\sum_{i=1}^n f_i = N$$

### **Distribuição de frequências**

(dados não agrupados)

É o arranjo de valores e suas respectivas frequências.

Para o exemplo dado será:

X <sub>i</sub>	21	22	23	24	25	26	28	30	31	32	33	34	35	36
F <sub>i</sub>	3	2	2	1	4	3	1	1	3	1	3	3	2	1

$$\sum_{i=1}^{15} f_i = 30$$

### Frequência relativa

A frequência relativa de um valor  $x_i$  é o quociente entre a frequência absoluta desse valor e o número total da população. A frequência relativa pode apresentar-se:

a) Por um número abstracto,  $fr = \frac{f_i}{N}$ , desta forma  $0 \leq fr \leq 1$

b) Ou em percentagem,  $fr = \frac{100f_i}{N} = \%$

No nosso exemplo (idade dos estudantes de uma turma) podemos calcular a frequência relativa, dos estudantes com 26 anos de idade.

$$fr(26) = \frac{3}{30} = 0,1 \text{ ou } fr(26) = \frac{3 \cdot 100}{30} = 10\%$$

A soma de todas as frequências relativas é igual a unidade ou a 100%.

$$\sum_{i=1}^n \frac{f_i}{N} = \sum_{i=1}^n fr_i = 1$$

### Dados agrupados em classe

Quando a variável é contínua ou quando o número de valores observados é grande, é conveniente fazer o seu agrupamento em **classes**.

### O número de classes (k)

Não há uma fórmula exacta para o cálculo do número de classes. Temos duas sugestões:

a)  $K = 5$  para  $N \leq 25$  e  $K \cong \sqrt{N}$  para  $N > 25$

b) Fórmula de Sturges  $K \cong 1 + 3,22 \cdot \lg N$  onde  $N$  é o tamanho da amostra.

### Amplitude da classe (h)

$$h \cong R : k$$

O número de classes (k) assim como a amplitude das classes (h), deve ser aproximado para o maior inteiro.

Assim, se  $K \cong 6,4$  usa-se  $K \cong 7$  ou se  $h \cong 1,7$  usa-se  $h \cong 2$ .

### Limite das classes

Existem diversas maneiras de expressar os limites das classes.

- a)  $[6,10]$  Compreende todos os valores de 6 a 10.
- b)  $[6,10[$  Compreende todos valores de 6 a 10 excluindo o 10.
- c)  $]6,10[$  Compreende todos valores de 6 a 10, excluindo os extremos.
- d)  $]6,10]$  Compreende todos os valores de 6 a 10, excluindo o 6, usa-se mais a forma b

### Ponto médio ou marca da classe

É a média aritmética entre o limite inferior e superior da classe.

#### Exemplo 4

Está registrado abaixo, o tempo em segundo que cada estudante aguenta sem inspirar nem expirar depois de encher os pulmões.

83	59	83	47	72
84	37	33	43	42
39	42	20	35	40
58	40	69	73	55
30	82	21	34	59
45	52	30	86	30
26	77	48	85	65
31	29	80	70	30
36	86	76	81	30
60	22	81	55	55

Componha a tabela indicando: as classes, os pontos médios e as frequências.

#### Resolução

Amplitude total  $R = 86 - 20 = 66$

Numero de classes  $K = 1 + 3,22 \lg 50 \cong 7$  ou  $K \cong \sqrt{50} \cong 7$

Amplitude de classes  $h = R : h = 66 : 7 \cong 10$

Vamos considerar 7 classes de amplitude 10, de 20 a 90.

Classes	Ponto médio ou marca da classe ( $x_i$ )	Frequência absoluta ( $f_i$ )	Frequência relativa ( $f_r$ )	$F_r$ (%)
[20;30[	25	5	0,1	10
[30;40[	35	12	0,24	24
[40;50[	45	8	0,16	16
[50;60[	55	7	0,14	14
[60;70[	65	3	0,06	6
[70;80[	75	5	0,1	10
[80;90[	85	10	0,2	20
Total		$N = \sum f_i = 50$	$\sum f_r = 1$	100

---

## Exercícios

**14.** Numa turma da 10ª classe, perguntou-se a cada aluno quantos irmãos tem. Segue-se o registo dos alunos:

0,1,6,3,2,3,1,1,0,1,1,12,0,2,2,4,2,1,0.

Construa a tabela das frequências.

**15.** De uma pauta onde estavam registados os resultados de um teste de estatística e cujas notas a atribuir variam de 7 a 14, registarem-se as seguintes classificações:

11,8,11,8,12,14,9,11,10,9,12,9,11,12,10,9,8,11,8,8,8,10,10,9,10,13,9,9,10,9,10,10,13,12,13,14,11,14,14,12,8,11,12,11,12,13,11,11,12 e 10.

Construa a tabela de frequências.

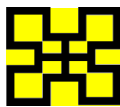
**16.** Na tabela abaixo estão os dados do inquérito feito aos docentes de uma escola secundária, referente a marca do telemóvel que usam.

<b>Marca do telemóvel</b>	<b>Efectivo (fi)</b>	<b>Frequência relativa (fr)</b>	<b>Fr (%)</b>
<i>Ericson</i>	10		
<i>Motorola</i>	10		
<i>Nokia</i>	12		
<i>Panasonic</i>	4		
<i>Samsung</i>	2		
<i>Siemens</i>	6		
<i>Sony</i>	8		

Complete a tabela (arredonde os dados para duas casas decimais).

**17.** Nas eleições para um distrito urbano estavam inscritos 4000 eleitores. No partido **A** votaram 34%, no partido **B** 1200 eleitores, 10% abstiveram-se e os restantes votaram no partido **C**

- a) Quantos leitores votaram?
- b) Calcule a percentagem dos eleitores que votaram no Partido B.
- c) Qual é o partido que venceu as eleições?



### Auto-avaliação

**18.** Num inquérito económico regional do nosso País, efectuaram-se levantamentos de dados sobre produção agrícola da produção de milho, numa determinada época, tendo-se apurado o seguinte resultado.

Regional	Efectivo da Produção (em toneladas)	Percentagem da produção
Norte		
Centro	850.000	35
Sul	580.000	
Total		

- Indique a variável em estudo e classifique-a;
- Complete o preenchimento da tabela acima apresentada;
- Sabendo que cada região gastou, em média, 350 litros de combustível (gasóleo) por tonelada de produção, quantos meticais o país desembolsou em combustível nas três regiões, atendendo que o litro custava 22,00Mt?

**19.** Num inquérito, feito a 30 alunos de uma classe de instrução primária sobre suas alturas foram obtidas os seguintes resultados, em centímetros:

126	124	130	136	124
128	128	131	127	130
131	130	127	124	127
130	125	136	129	135
127	132	132	133	123

- Indique a variável em estudo e classifique-a.
- Construa uma tabela de frequência.

---

## Tema: Frequências acumuladas e gráficos

### Introdução

As frequências absolutas e relativas podem ser acumuladas ascendentes e descendentes. Nesta unidade vamos compor as tabelas das frequências acumuladas ascendentes e descendentes, assim como os seus respectivos tipos de gráficos.

Ao completar esta unidade, você será capaz de:



#### Objectivos

- Saber calcular as frequências acumuladas ascendentes e descendentes;
- Ser capaz de reconhecer os gráficos: de barras e histogramas de diferentes tipos de gráficos.

---

### Sumário

#### Frequências absolutas acumuladas descendentes $fa(\downarrow)$

No exercício número 15 na unidade anterior falamos de dados referentes aos resultados de uma pauta de Estatística.

Queremos resolver a seguinte questão:

Qual é a frequência absoluta dos estudantes com classificações inferiores ou iguais a 10?

Podemos obter a resposta dessa questão somando as frequências das classificações de 7, 8, 9 e 10.

Representa-se por:

$$\begin{aligned}fa(x \leq 10) &= fa(x=7) + fa(x=8) + fa(x=9) + fa(x=10) = \\ &= 0 + 7 + 8 + 9 = 24.\end{aligned}$$

De igual modo podemos calcular a frequência dos estudantes com classificações iguais ou menores a 12.

$$fa(x \leq 12) = 0 + 7 + 8 + 9 + 10 = 42$$

Frequência acumulada descendente do valor  $x_i$ ,  $fa(\downarrow)$  é a soma de todas as frequências dos valores menores ou iguais a  $X_i$ .

### Frequência absoluta acumulada ascendente $fa(\uparrow)$

Agora queremos a frequência dos estudantes com notas maiores ou iguais a 10. Pede-se neste caso  $F(x \geq 10)$ .

Analogamente como na tarefa anterior, vamos somar todas as frequências dos valores maiores ou iguais a 10.

$$\begin{aligned} F(x \geq 10) &= fa(x=10) + fa + fa(x=11) + fa(x=12) + fa(x=13) + fa(x=14) = \\ &= 10 + 8 + 4 + 4 = 26 \end{aligned}$$

Frequência acumulada ascendente do valor  $X_i$ , é a soma de todas as frequência dos valores (menores?) maiores ou iguais a  $X_i$ .

Podemos construir uma tabela com todas as frequências absolutas e relativas, acumuladas descendentes e ascendentes:

$$fa(\downarrow), fa(\uparrow), fra(\uparrow)$$

Tabela das frequências do exercício número 2 da unidade 4.

$X_i$	$f_i$	$f_r$	$Fa(\downarrow)$	$Fa(\uparrow)$	$Fra(\downarrow)$	$fra(\uparrow)$
7	0	0	0	50	0	1
8	7	0,14	7	50	0,14	1
9	8	0,16	15	43	0,3	0,86
10	9	0,18	24	35	0,48	0,70
11	10	0,2	34	26	0,68	0,52
12	8	0,16	42	16	0,84	0,32

13	4	0,08	46	8	0,92	0,16
14	4	0,08	50	4	1,00	0,08
Total	50	1				

Através desta tabela podemos obter respostas rápidas das perguntas sobre as frequências acumuladas:

- A Frequência absoluta dos estudantes com notas inferiores a 11 é 24 (ou seja, 24 alunos têm uma classificação menor ou igual a 11 / 24 alunos têm uma classificação não superior a 11);
- A frequência absoluta dos estudantes com notas maiores ou iguais a 12 é 16 (ou seja, 16 alunos têm uma classificação maior ou igual a 12);
- A frequência relativa dos estudantes com notas maiores ou iguais a 10 é 0,48 ou 48%.
- A Percentagem dos estudantes com notas maiores que 9 é de 70% (ou seja, 70% dos alunos têm uma classificação não inferior a 9).

#### Formas de representação gráfica

- Gráfico de barras ou de colunas

Nestes gráficos os dados são representados por meio de colunas (ou barras) da mesma largura e com alturas correspondentes às frequências.

#### Exemplo 5

Fez-se um levantamento dos programas preferidos de 100 telespectadores de uma localidade.

Os resultados obtidos registaram-se na seguinte tabela:

Tipo de programas	Efectivos
Filmes (F)	20
Concursos (C)	15
Telenovelas (T)	30

Informação (I)	25
Outros (O)	10
Total	100

Vamos construir um gráfico de barras corresponde às frequências absolutas apresentadas na tabela.

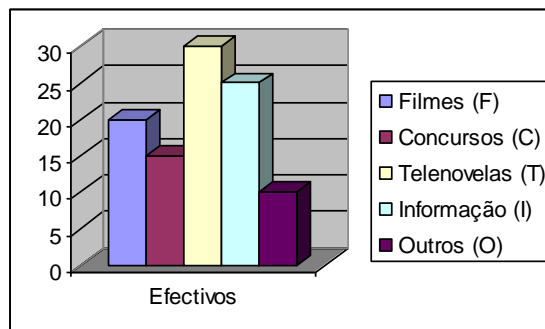


Gráfico de frequências absolutas

Para traçar um gráfico de barras, marca-se sobre o eixo das abcissas (de um sistema de dois eixos coordenados rectangulares) os valores da variável  $X_i$  e, sobre o eixo das ordenadas, os valores das frequências  $f_i$  ou  $f_a$ .

Por cada um dos pontos marcadores sobre o eixo das abcissas traçam-se segmentos de recta de comprimento igual à frequência respectiva,

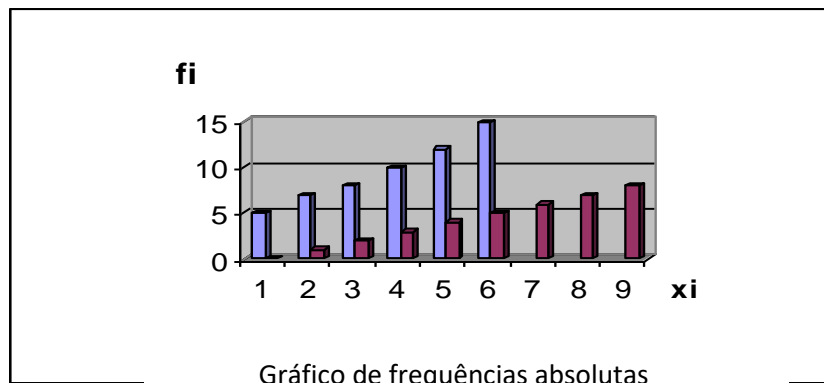
### Exemplo 6

Consideremos a seguinte tabela de frequências.

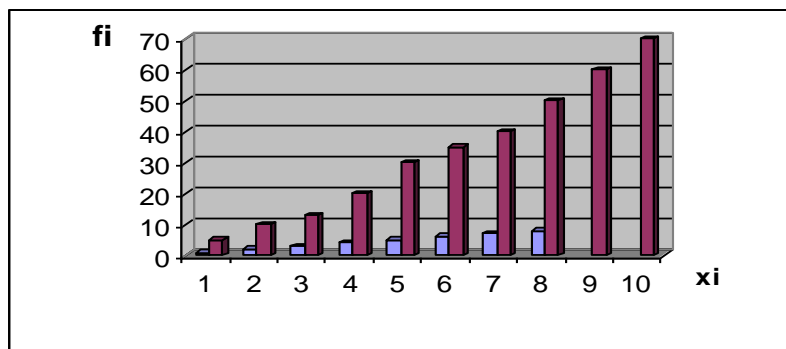
Variável ( $X_i$ )	Frequência absoluta ( $f_i$ )	Frequência acumulada ( $f_a$ )
1	5	5
2	8	13
3	10	23

4	7	30
5	5	35
6	8	43
7	12	55
8	15	70

Podemos fazer a sua representação gráfica utilizando um gráfico de barras.



Observação: Este tipo de diagrama é indicado para comparar dados qualitativos e quantitativos de tipo discreto.



### Sectogramas ou diagramas circulares

Este tipo de gráficos usa-se quando se pretende comparar diversas partes de um todo.

Divide-se um círculo em sectores de amplitude proporcionais às frequências absolutas.

Sabendo que ângulo ao centro correspondente a toda circunferência vale  $360^\circ$ , temos:

$$\frac{\textit{amplitude}}{f_r} = \frac{360^\circ}{1} \quad \text{donde amplitude} = 360^\circ \times f_r \quad \text{se } f_r \text{ está}$$

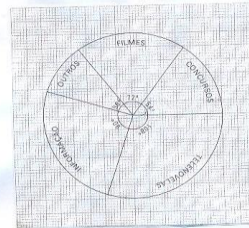
expresso em percentagem,  $\text{ampl} = 360^\circ \cdot f_r$

#### Exemplo 7

Retomemos os resultados apresentados na tabela do exemplo anterior:

$X_i$	$f_i$	fr (%)	Amplitude do sector
Filmes	20	20%	$3,6 \times 30 = 72^\circ$
Concursos	15	15%	$3,6 \times 15 = 54^\circ$
Telemóveis	30	30%	$3,6 \times 30 = 108^\circ$
Informações	25	25%	$3,6 \times 25 = 90^\circ$
Outros	10	10%	$3,6 \times 10 = 36^\circ$
Total	100	100%	$360^\circ$

O Sectograma correspondente a esta distribuição poderá ser o seguinte:



#### Pictogramas

São gráficos cuja característica principal é o uso de figuras alusivas ao fenómeno em estudo. Utiliza-se bastante em propaganda, dado o seu apelo visual permitir uma percepção imediata do que se está a tratar.

#### Exemplo 8

Uma fábrica produziu, durante três anos, os seguintes computadores:

Produção de computadores em	Efectivos ( $f_i$ )
2005	150000
2006	225000
2007	300000

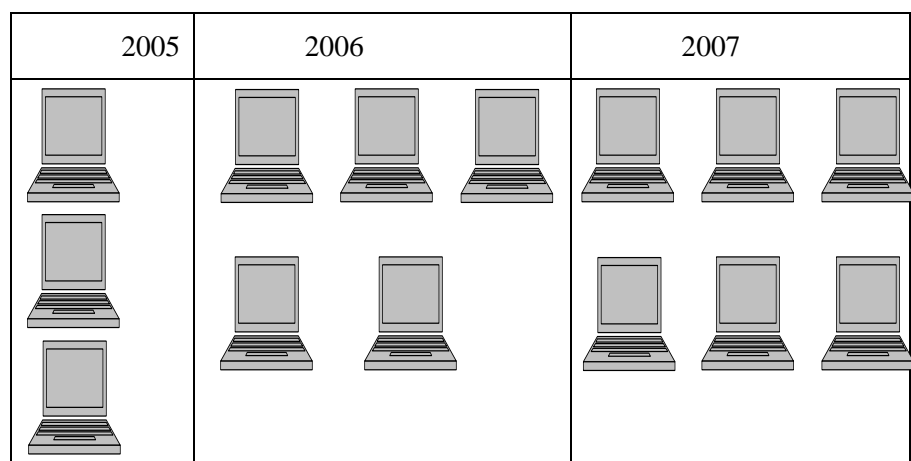
## Pictogramas


São gráficos cuja característica principal é o uso de figuras alusivas ao fenómeno em estudo. Utiliza-se bastante em propaganda, dado o seu apelo visual permitir uma percepção imediata do que se está a tratar.

### Exemplo 8

Uma fábrica produziu, durante três anos, os seguintes computadores:

Produção de computadores em	Efectivos ( $f_i$ )
2005	150000
2006	225000
2007	300000



Cada  representa 50 000 computadores

### Histogramas e polígonos de frequências

Os histogramas são gráficos que representam os dados de uma distribuição de valores agrupados em classes. São formados por rectângulos justapostos de bases correspondentes às amplitudes de cada classe, marcadas no eixo dos  $XX'$ , áreas proporcionais às respectivas frequências (simples ou acumuladas).

### Polígono de frequências simples

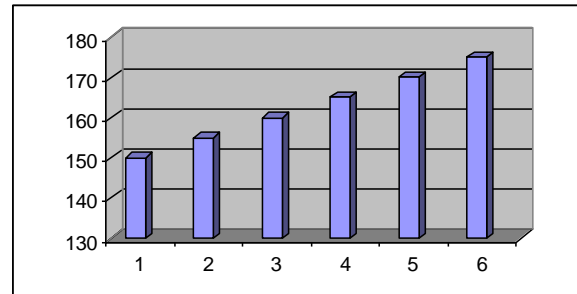
É a linha que une os vértices superiores de direitos dos rectângulos de um histogramas de frequências acumuladas.

### Exemplo 9

As alturas, em cm, de alunos de uma turma são dadas pela tabela:

$X_i$	$f_i$
[150,155[	4
[155,160[	7
[160,165[	8
[165,170[	6
[170,175[	5
Total	N = 30

Vamos construir um histograma correspondente a estes dados.

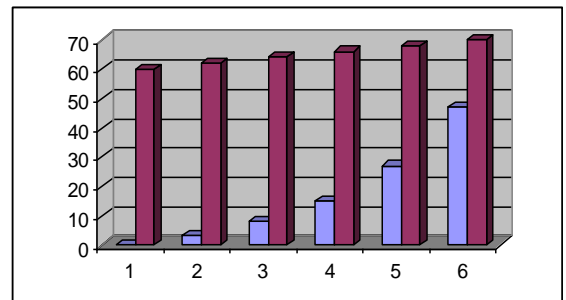


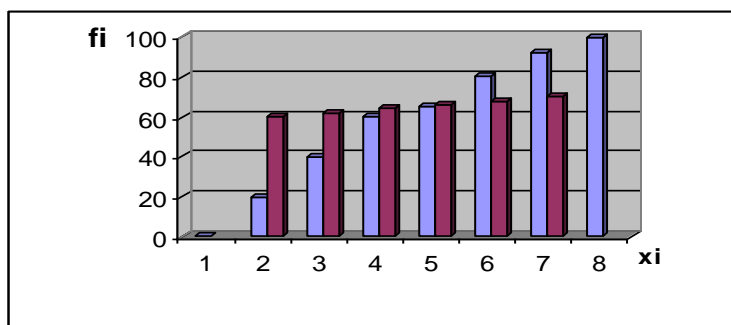
Altura dos alunos de uma turma

### Exemplo 10

Vamos construir o polígono de frequências, que diz respeito ao peso (em kg) dos alunos de uma escola cujos dados estão representados na tabela seguinte:

Peso (em kg)	Nº de alunos ( $f_i$ )	$f_a$
[60,62[	3	3
[62,64[	15	18
[64,66[	47	65
[66,68[	27	92
[68,70[	8	100
Total	N = 100	-





Histograma de frequências absolutas acumuladas

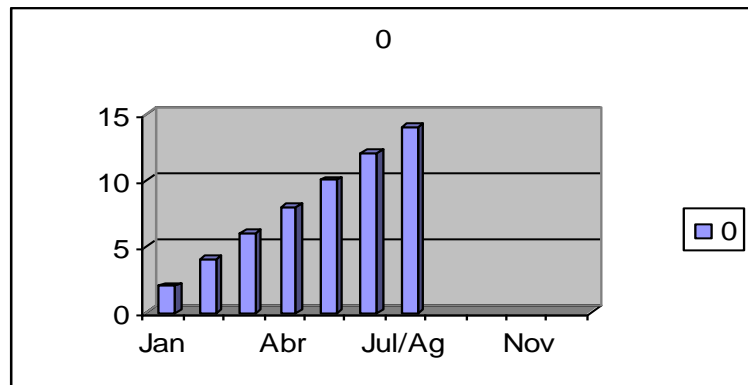
## Exercícios

**20.** As classificações dos alunos de uma turma na disciplina de filosofia no final do 1º trimestre, foram as seguintes:

10	10	12	11	9
14	8	9	10	12
15	13	8	13	14
9	11	10	12	17
14	9	8	12	10

- Elabore uma tabela de frequências absolutas e relativas, simples e acumuladas.
- Qual foi a classificação mais frequente?
- Determine a percentagem de alunos com classificação negativa.
- Quantos alunos tiveram nota superior a 12?
- Qual é o número de alunos que tiveram a nota inferior a 11?
- Qual a percentagem de alunos cuja classificação variou entre 10 e 13 (inclusive)?

**21.** O gráfico seguinte indica o número de máquinas expedidas em 1993 numa empresa.



- Qual a população em estudo?
- Indique a unidade estatística.
- Faça uma tabela de frequências absolutas simples e acumuladas que represente a distribuição apresentada no gráfico.
- Qual o mês em que houve maior expedição de máquinas?
- Quantas máquinas se expediram até Junho de 93 (inclusive)?

**22.** Num congresso sobre “Qualidade de ensino” estiveram presentes 10 oradores, 150 professores e 15 funcionários.

Represente esta informação através de um diagrama circular, fazendo as legendas respectivas.

**23.** Dos 835 mil contos atribuídos pelo M.E., em 1993, para o desporto de alta competição, 500 mil foram distribuídos de seguinte modo:

Atletismo – 128500

Andebol – 65000

Basquetebol – 50000

Futebol – 200000

Natação – 51000

Tiro – 5500

- Indique a população e amostra em estudo;

- b) Indique a percentagem da verba atribuída ao futebol, em relação à verba atribuída à amostra.
- c) Calcule a percentagem da verba atribuída à modalidade Basquetebol relativamente à dotação total da população;
- d) Construa um diagrama circular dos dados apresentados.
- e) Sabendo que foi de 3% a verba atribuída à Ginástica, determine o montante recebido por esta modalidade.

**25.** A tabela a seguinte indica a distância de casa à escola de 30 alunos de uma turma:

<b>Distância em km</b>	$[0,5[$	$[5,10[$	$[10,15[$
Efectivo	15	8	7

- a) Indique a amplitude de cada classe.
- b) Qual é a marca de cada classe?
- c) Represente os dados apresentados através de um histograma.
- d) Construa o polígono de frequências absolutas simples e de frequências absolutas acumuladas.

**26.** Considere a seguinte tabela de dados agrupados:

<b>Classes</b>	<b>Frequências</b>
$[40,50[$	5
$[50,60[$	10
$[60,70[$	16
$[70,80[$	12
$[80,90[$	8
$[80,90[$	5

Represente estes dados por meio de um histograma de frequências absolutas e desenhe o respectivo polígono de frequências.

---

## Tema: Somatórios

### Introdução

Para ler este manual não é necessário conhecimentos de matemática além do que é dado no 1º ciclo do Ensino Secundário Geral ou noutros níveis equivalentes.

Entretanto, nesta Unidade é dada uma noção sobre somatório que, apesar de simples, geralmente não é conhecido em tais níveis.

Ao completar esta unidade, você será capaz de:

- Conhecer o uso de Somatório e sua importância;
- Saber resolver expressões com somatórios



### Objectivos

---

### Somatórios

Muitas vezes é necessário indicar a soma de  $n$  valores.

Como exemplo, vamos supor que 20 alunos fizeram uma prova e existe interesse em determinar a média das notas obtidas. Devemos então somar todas as notas e dividir a soma das notas por 20.

Existe uma forma, bastante compacta, para indicar que deve ser feita uma operação de forma, conforme veremos aqui.

Suponhamos então que os nomes dos 20 alunos estão organizados em uma lista, por ordem alfabética.

Basta um  $X_1$  indicar a nota do aluno cujo o nome é o primeiro da lista,  $X_2$  indicar a nota do aluno é o segundo da lista e assim por diante, até  $X_{20}$ , que irá indicar a nota do aluno cujo o nome é o vigésimo da lista. Então os índices 1, 2, 3,... 20, corresponde à posição dos nomes da lista.

Estabelecida esta notação, podemos indicar a soma das notas dos 20 alunos como segue:

$X_1 + X_2 + \dots + X_{20}$  onde os pontos significam “e assim por diante”.

Entretanto, também podemos indicar esta soma de outra forma, bem mais compacta. Basta escrever  $\sum_{i=1}^{20} X_i$  que se lê “Somatório de x índice i, i variando de 1 a 20”. O símbolo que indica o somatório é,  $\sum$ , e é a letra grega sigma maiúscula.

Portanto, quando escrevemos  $\sum_{i=1}^n X_i$  estamos indicando que o índice i deve ser, sucessivamente, substituído por números inteiros, em ordem crescente, começando por 1 e terminando em n e depois deve ser feita a soma  $X_1 + X_2 + X_3 + \dots + X_n$ .

Podemos utilizar qualquer letra para indicar o índice. Entretanto, são mais frequentemente utilizadas as letras i, j e k.

Vejamos um exemplo. Sejam  $X_1 = 2$ ,  $x_2 = 4$ ,  $x_3 = 3$  e  $x_4 = 1$ . Para indicar a soma desses valores, escrevemos:

$$\sum_{i=1}^4 X_i \quad \text{Desse modo,} \quad \sum_{i=1}^4 X_i = 2+4+3+1 =$$

10

Em estatística, muitas vezes é necessário obter o quadrado da soma de  $X_1, X_2, \dots, X_n$ .

Ora, já vimos que a soma de  $x_1, x_2, \dots, x_n$  pode ser indicada por  $\sum_{i=1}^n X_i$ . Para indicar o quadrado dessa soma, isto é, para indicar  $(x_1+x_2+\dots, x_n)^2$  basta escrever

$(\sum_{i=1}^n X_i)^2$ . Vejamos um exemplo. Sejam  $X_1 = 3, x_2 = 4, x_3 = 1, x_4 = 2$  e  $x_5 = 3$ . É fácil ver que  $\sum_{i=1}^n X_i = 3+4+1+2+3 = 13$

Agora podemos obter  $(\sum_{i=1}^5 X_i)^2 = 13^2 = 169$

Muitas vezes também é necessário obter a soma dos quadrados dos valores  $X_1, X_2, \dots, X_n$ . Ora esta soma de quadrados pode ser indicada como segue:

$$x_1^2 + x_2^2 + \dots + x_n^2.$$

Entretanto, também podemos indicar essa soma de quadrados desses números escrevendo:

$\sum_{i=1}^n x_i^2$ . O valor dessa soma é

$$\sum_{i=1}^n x_i^2 = x_1^2 + x_2^2 + x_3^2 = 3^2 + 5^2 + 1^2 = 35$$

Suponhamos agora que temos dois conjuntos de números, isto é, o conjunto  $x_1, x_2, \dots, x_n$  e o conjunto  $y_1, y_2, \dots, y_n$ . Pode haver interesse em obter a soma dos produtos  $x_1 y_1, x_2 y_2, \dots, x_n y_n$ . Ora, esta soma pode ser indicada como segue:

$X_1 y_1 + x_2 y_2 + \dots + x_n y_n$  ou de uma forma mais compacta, como segue:  $\sum_{i=1}^n x_i y_i$

Vejamos um exemplo: Sejam  $X_1 = 2, X_2 = 3$  e  $X_3 = 0$  e sejam  $y_1 = 1, y_2 = 2$  e  $y_3 = 5$ . A soma dos produtos é indicada por

$$\sum_{i=1}^3 x_i y_i \text{ O valor dessa soma é}$$

$$\sum_{i=1}^3 x_i y_i = 2.1 + 3.2 + 0.5 = 8$$

Dado um conjunto de  $n$  números, se as somas que pretendemos obter se estendem a todos eles, podemos omitir o índice, por brevidade. Então muitas vezes escrevemos  $\sum x$ , em lugar de  $\sum_{i=1}^n x_i$ . Da mesma maneira podemos escrever  $\sum y, \sum x^2, \sum y^2, (\sum x)^2, (\sum y)^2$

### Algumas propriedades dos somatórios

$$\text{Aditiva: } \sum_{i=\kappa}^n (a_i + b_i) = \sum_{i=\kappa}^n a_i + \sum_{i=\kappa}^n b_i$$

$$\text{Homogênea: } \sum_{i=\kappa}^n (\alpha \cdot a_i) = \alpha \cdot \sum_{i=\kappa}^n a_i \quad (\alpha = \text{constante})$$

Vejamos alguns exemplos: a) vamos calcular, recorrendo às propriedades evocadas, o valor de  $\sum_{k=5}^9 (3k + 2)$ :

$$\begin{aligned} \text{Temos: } \sum_{k=5}^9 (3k + 2) &= \sum_{k=5}^9 (3k) + \sum_{k=5}^9 2 = 3 \cdot \sum_{k=5}^9 k + \sum_{k=5}^9 2 = \\ &= 3 \cdot (5 + 6 + 7 + 8 + 9) + \sum_{k=5}^9 2 = 105 + (2 + 2 + 2 + 2 + 2) = 115 \end{aligned}$$

Ou

$$\sum_{k=5}^9 (3k + 2) = \sum_{k=5}^9 (3k) + \sum_{k=5}^9 2 = 3 \cdot \sum_{k=5}^9 k + \sum_{k=5}^9 2 = 3 \cdot (5 + 6 + 7 + 8 + 9) + \sum_{k=5}^9 2 = 105 + (9 - 5 + 1) = 115$$

b) Dada a equação  $\sum_{i=0}^5 (3_i) - \sum_{i=0}^5 (2i - x) = 0$ , determinemos o valor de  $x$ :

Temos:

$$\sum_{i=0}^5 (3i) - \sum_{i=0}^5 (2i - x) = 0 \Leftrightarrow \sum_{i=0}^5 (3i - 2i + x) = 0$$

$$\Leftrightarrow \sum_{i=0}^5 (i + x) = 0 \Leftrightarrow \sum_{i=0}^5 i + \sum_{i=0}^5 x = 0 \Leftrightarrow$$

$$\Leftrightarrow (0+1+2+3+4+5) + (5-0+1)x = 0 \Leftrightarrow 15+6x = 0 \Leftrightarrow x = -\frac{5}{2}$$

---

## Exercícios

**26.** Represente sob a forma de somatório as seguintes expressões:

- a)  $1+2+3+4+5$
- b)  $1+2+3+4+5+\dots+n$
- c)  $1^2+2^2+3^2+\dots+20^2$
- d)  $4+4+4+\dots+4(12 \text{ paralelas})$
- e)  $0^2+1^2+2^2+\dots+(n-1)^2$
- f)  $1+3+5+7+9$
- g)  $2+4+6+8+10+12+14$
- h)  $2+4+6+8+\dots+(2n)$
- i)  $x_1 + x_2 + x_3 + x_4 + x_5$
- j)  $f_1x_1 + f_2x_2 + \dots + f_nx_n$

**27. Calcule:**

a)  $\sum_{i=1}^3 (2i);$       b)  $\sum_{j=0}^5 (-3j);$       c)  $\sum_{k=-1}^3 (k+1);$

d)  $\sum_{i=-4}^{\overline{0}} (1-3i)$

e)  $\sum_{\kappa=1}^{10} 2;$       f)  $\sum_{j=0}^{20} (5x);$       g)  $\sum_{i=2}^6 i^2$

h)  $\left( \sum_{i=2}^6 i \right)^2$

i)  $\sum_{k=1}^5 \kappa + 3$

28. Sabendo que  $\sum_{k=1}^{100} a_k = 25$  e  $\sum_{k=1}^{100} (a_k)^2 = 125$ , calcule:

a)  $\sum_{k=1}^{100} (a_k + 1)$ ; b)  $\sum_{k=1}^{100} (2 - 3.a_k)$ ; c)  $\sum_{k=1}^{100} (a_k + 1)^2$

29. Escreva sob a forma de um único somatório:

a)  $\sum_{k=-1}^8 (2k) + \sum_{k=1}^8 (3k)$ ; b)  $\sum_{i=0}^5 (3i) - \sum_{i=0}^5 (i=1)$ ; c)  $\sum_{i=1}^{15} (i-1) - \sum_{i=1}^{15} (2i) - 14$ ; d)  $\sum_{i=1}^{20} (3i) + \sum_{i=1}^{20} (1-i) + 2$

30. Resolva em ordem a x as equações

a)  $\sum_{i=1}^{20} (2i-1) + x = \sum_{i=1}^{20} (2i)$ ; b)  $\sum_{j=0}^{15} (1+2j) = \sum_{j=1}^4 x + \sum_{j=1}^{15} (1+2j)$ ;  
c)  $2x \cdot \sum_{i=-2}^{10} i^2 = \sum_{i=-2}^{10} i^2$ ; d)  $\sum_{k=0}^{100} k^2 = x + \sum_{k=1}^{100} (k^2 + 1)$

---

## **Tema: Medidas de Posição ou de Tendência central (dados não agrupados)**

### **Introdução**

As medidas de posição são muito importantes na interpretação dos dados estatísticos. Numa primeira fase, vamos falar da média, modas e mediana para dados não agrupados em classes.

Ao completar esta unidade, você será capaz de:



### **Objectivos**

- Calcular a mediana, moda e a média ponderada com as respectivas frequências.

---

### **Medidas de posição (dados não agrupados)**

Apesar das tabelas estatísticas e das representações gráficas nos darem uma ideia clara da distribuição de frequências da variável estudada, torna-se necessária simplificar ainda mais o conjunto de dados, de forma a caracterizar a distribuição por um número reduzido de medidas (parâmetros) que evidenciem o que demais significativo existe no conjunto.

Estes parâmetros podem agrupar-se em dois tipos:

- a) Medidas de posição ou de localização (ou de tendência central);

b) Medidas de dispersão ou de variabilidade;

Vamos estudar em primeiro lugar as medidas de posição ou de localização.

Essas medidas indicam-nos valores típicos a volta dos quais os dados se distribuem. Essas subdividem-se em duas partes, que são:

**Medida de tendência central** – Média, Moda e Mediana.

**Medida de separação (ou medidas de ordem)** – os quartis e decís.

### Média Aritmética ( $\bar{X}$ )

#### Dados não agrupados

Chama-se média aritmética de um conjunto de valores  $X_1, X_2, X_3, \dots, X_n$ , ao quociente que se obtêm da soma de todos os valores pelo efectivo no total  $N$ .

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{N}$$

Utilizando simbolo de somatório teremos  $\bar{X} = \frac{\sum_{i=1}^n x_i}{N}$  ou

simplesmente  $\bar{X} = \frac{\sum x}{N}$

#### Exemplo 11

A média aritmética simples de valores 3, 7, 8, 12, 15, é:

$$\bar{X} = \frac{3+7+8+12+15}{5} = \frac{45}{5} = 9$$

### Média Ponderada

Se  $x$  é uma variável discreta que toma os valores  $x_1, x_2, x_3, \dots, x_n$  com as frequências absolutas  $f_1, f_2, f_3, \dots, f_n$  respectivamente, a média aritmética será dada por

$$\bar{X} = \frac{f_1 x_1 + f_2 x_2 + f_3 x_3 + \dots + f_n x_n}{N} \text{ ou } \bar{X} = \frac{\sum_{i=1}^n f_i x_i}{N}$$

Neste caso, diz-se que  $\bar{X}$  é a média ponderada pelas respectivas frequências.

### Exemplo 12

Determine a média na seguinte distribuição:

$x_i$	1	2	3	4
$f_i$	1	3	5	2

Uma maneira mais prática de calcular é compor a tabela seguinte:

$x_i$	$f_i$	$X_i \cdot f_i$
1	1	1
2	3	6
3	5	15
4	2	8
Total	N=11	30

$$\bar{X} = \frac{\sum f_i \cdot x_i}{N} = \frac{30}{11} = 2,72$$

### Moda ( $M_o$ )

Chama-se Moda (ou valor modal) da distribuição de frequências ao valor da variável que corresponde a maior frequência.

Existem séries estatísticas com duas modas (bimodal), com três modas (trimodal), etc.

Também existem séries em que não existe a moda.

No nosso exemplo (exemplo 2) de dados não agrupados em classe a moda é o valor três porque tem a maior frequência absoluta.

Mais adiante vamos apresentar o cálculo da média para dados agrupados em classe.

### Mediana ( $M_e$ )

A mediana é a medida de posição que divide a série estatística em duas partes iguais, ou seja, é o valor da variável estatística precedido por 50% das observações.

Para a sua determinação é necessário que os dados estejam ordenados.

a) Variáveis discretas

- Número de valores observados ímpar.

$M_e$  = valor central depois da ordenação dos valores observados por ordem crescente.

- Número de dados observados par

$M_e$  = média aritmética dos dois valores mais centrais depois da ordenação dos valores observados por ordem crescente.

### Exemplo 13:

a) A mediana do conjunto de números 2, 3, 3, 4, 5, 6, 7, 8, 9 é

$M_e = 5$

b) Para o conjunto 2, 3, 3, 4, 5, 6, 7, 8, 9, 10, a mediana será:

$$M_e = \frac{x_5 + x_6}{2} = \frac{5 + 6}{2} = 5,5$$

c) Qual a mediana dos seguintes valores da tabela?

$x_i$	1	3	6	8	12	15
$f_i$	4	6	3	9	10	8

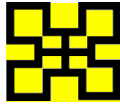
Para calcular a mediana começando por ordenar os dados para achar o valor do meio, será um processo mais laborioso. Vamos recorrer a tabela de frequências acumuladas.

$x_i$	$f_i$	$f_a$
1	4	4
3	6	10
6	3	13
8	9	22
12	10	32
15	8	40

Como o número de dados é par, a mediana é a média dos dois dados centrais,  $x_{20}$  e  $x_{21}$ . Pela tabela vê-se que o 20º e 21º dado tem valor 8; logo,  $M_e = 8$ .

A mediana é o valor a que corresponde a primeira frequência acumulada, maior do que  $\frac{N}{2}$  (metade do efectivo).

## Exercícios



Auto-avaliação

**31.** Mediram-se as alturas de 11 alunos de uma turma da 12ª classe e obteve-se os seguintes resultados:

1,75; 1,72; 1,70; 1,68; 1,68; 1,65; 1,65; 1,58; 1,56; 1,50; 1,49.

Determine:

a) A moda; b) A mediana; c) A média aritmética.

**32.** Mostre que a media aritmética pode ser dada pela fórmula  $\bar{X} = \sum x_i f_r$ , sendo  $f_r$  a frequência relativa.

**33.** Calcule a média aritmética para cada uma das distribuições:

a)

$x_i$	10	11	12	13
$f_i$	5	8	10	6

b)

$x_i$	2	3	4	5	6
$f_i$	3	9	19	25	28

c)

$x_i$	7	8	9	10	11
$f_i$	$\frac{1}{16}$	$\frac{5}{18}$	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{2}{48}$

d)

$x_i$	85	87	88	89	90
$f_i$	5	1	10	3	5

---

## Tema: Medidas de Posição (dados agrupados em classes de mesma amplitude)

### Introdução

Há uma ligeira diferença nos cálculos das medidas de tendência central para dados agrupados em classes. Para facilitar a compreensão das fórmulas da moda, mediana e média aritmética preferimos tratar separadamente dos dados não agrupados.

Ao completar esta unidade, você será capaz de:



### Objectivos

- Conhecer as fórmulas das medidas de tendência central para dados agrupados.
- Ser capaz de fazer a correção das classes quando temos amplitudes diferentes.

---

### Dados agrupados em classes

Vamos calcular a média, mediana e moda para dados agrupados em classes.

#### Média

Quando se trata de uma variável contínua, já agrupada em classes, a média aritmética obtém-se substituindo a variável de cada classe pela marca respectiva.

**Exemplo 14.** Determine a media da seguinte distribuição:

Peso em kg	Nº de pessoas (f <sub>i</sub> )	Pontomédio (x <sub>i</sub> )	x <sub>i</sub> ·f <sub>i</sub>
[40,44[	11	42	462
[44,48[	23	46	1058
[48,52[	20	50	1000
[52,56[	10	54	540
[56,60[	8	58	464
Total	72	-	3524

$$\bar{X} = \frac{\sum f_i \cdot x_i}{N} = \frac{3524}{72} = 48,9 \text{ kg}$$

### Mediana

Para o cálculo da mediana aplica-se a seguinte fórmula:

$$M_e = l_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{F_i} \cdot h_i$$

Onde:

$l_{i-1}$  é o limite inferior da classe mediana

$N$  é o total do efectivo da distribuição

$N_{i-1}$  é a frequência acumulada até a classe anterior à classe mediana.

$F_i$  é a frequência absoluta da classe mediana.

$h_i$  é a amplitude da classe mediana.

Usando os dados do exemplo anteriores (exemplo 14) podemos calcular a mediana :

Peso em kg	Nº de pessoas (f <sub>i</sub> )	Frequência acumulada
[40,44[	11	11
[44,48[	23	34

[48,52[	20	54
[52,56[	10	64
[56,60[	8	72
Total	$N = \sum f_i = 72$	-

Sendo  $\frac{N}{2} = \frac{72}{2} = 36$ , a mediana está na classe correspondente a primeira frequência acumulada superior a 36, que é a classe [48,52[.

Aplicando a fórmula dada,  $M_e = l_{i-1} + \frac{\frac{N}{2} - N_{i-1}}{F_i} \cdot h_i$

$$l_{i-1} = 48$$

$$N_{i-1} = 34$$

$$F_i = 20$$

$$h_i = 4$$

$$M_e = 48 + \frac{36 - 34}{20} \cdot 4 = 48 + \frac{1}{10} \cdot 4 = 48,4$$

### Moda

Se a variável ( $x_i$ ) é contínua e todos os dados estão agrupados em classes de iguais amplitudes, dá-se o nome de classe modal à classe de maior frequência.

Na maioria dos casos bastará saber o intervalo modal, mas querendo calcular com rigor o valor da moda ( $M_o$ ) poder-se-á utilizar a fórmula:

$$M_o = L_i + \frac{D_1}{D_1 + D_2} \cdot h$$

em que:

$L_i$  – Representa o limite inferior da classe modal.

$D_1$  – Representa a diferença entre a frequência absoluta da classe modal e a frequência absoluta da classe anterior.

$D_2$  – Representa a diferença entre a frequência absoluta da classe modal e a frequência da classe seguinte.

h – representa a amplitude da classe modal.

**Exemplo 15**

Na distribuição seguinte:

$X_i$	$F_i$	
$[0,2[$	3	
$[2,4[$	5	
$[4,6[$	10	
$[6,8[$	6	
$[8,10[$	3	
$[10,12[$	4	

A moda está na classe  $[4,6[$  por esta ser a de maior frequência; a esta chama-se classe modal

Calculando a moda temos

$$M_o = 4 + \frac{10 - 5}{(10 - 5) + (10 - 6)} \cdot 2 = 4 + \frac{5}{5 + 4} \cdot 2 \approx 5$$

**Exercícios**

**34.** considere a seguinte tabela que representa a distribuição das alturas dos alunos de uma turma.

Alturas em metros	Efectivos
$[1,40 - 1,45[$	3
$[1,45 - 1,50[$	10
$[1,50 - 1,55[$	12
$[1,55 - 1,60[$	5
$[1,60 - 1,65[$	2
$[1,65 - 1,70[$	4

- a) Classifique o carácter em estudo;
- b) Indique a amplitude da distribuição e de cada classe;
- c) Calcule a moda;
- d) Calcule a moda;
- e) Calcule a mediana;
- f) Calcule a média aritmética.

**35.** Dada a amostra:

28	33	27	30	31	30	33	30	33	29
27	33	31	27	31	28	27	29	31	24
31	33	30	32	30	33	27	33	31	33
23	29	30	24	28	34	30	30	18	17
18	15	16	17	17	18	19	19	20	29

- a) Agrupe os elementos em classe (use  $h = 5$ )
- b) Construa a tabela das frequências absolutas acumuladas descendentes e ascendentes.
- c) Represente no mesmo sistema de eixos, os polígonos das duas frequências acumuladas.

- d) Determine a mediana e visualize-a graficamente na sua figura.

**36.** A classificação final, em percentagem, de matemática de 80 estudantes de uma Universidade Estatal, é dada na tabela seguinte:

68	84	75	82	68	90	62	88	76	96
73	79	88	73	60	93	61	59	85	75
61	65	75	87	74	62	95	78	63	72
66	78	82	75	94	77	69	74	68	60
96	78	80	61	75	95	60	79	83	71
79	62	67	97	78	85	76	65	71	75
65	80	73	57	88	78	62	76	53	74
86	67	73	81	72	63	76	75	85	77

Determine:

- A amplitude da distribuição;
- O número de classes e as suas amplitudes (use a fórmula de Sturges)
- Constroi a tabela de frequências absolutas, relativas e acumuladas.
- Quantos estudantes receberam graus abaixo de 83?
- Qual é a percentagem dos estudantes que receberam graus entre 83 e 95?
- Qual é o valor da moda.

---

## Tema: Outros tipos de Médias

### Introdução

Nesta Unidade vamos falar doutros tipos de médias que são usadas na estatística, que são a media geométrica, mediana, quadrática e a média harmónica.

Ao completar esta unidade, você será capaz de:



#### Objectivos

- Saber calcular as médias geométricas, quadráticas e harmónicas;
- Conhecer a relação entre as diferentes tipos de médias.
- Ser capaz de aplicar as fórmulas do somatório;

---

### Sumário

#### Outros tipos de médias

##### Média Geométrica (Mg)

A média geométrica dos valores observados  $x_1, x_2, x_3, \dots, x_n$  é a raiz de índice  $n$  do produto desses números.

$$Mg = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \dots x_n}$$

Aplicando logaritmos temos

$$\lg(Mg) = \frac{1}{n} \sum_{i=1}^n \lg(x_i)$$

Se os valores  $x_1, x_2, x_3, \dots, x_n$  estiverem com as suas respectivas frequências absolutas  $f_1, f_2, f_3, \dots, f_n$ ; a média geométrica será dada por  $Mg = \sqrt[n]{x_1 \cdot f_1 \cdot x_2 \cdot f_2 \cdot x_3 \cdot f_3 \dots x_n \cdot f_n}$  donde resulta:

$$\lg(Mg) = \frac{1}{n} \sum_{i=1}^n f_i \cdot \lg(x_i)$$

##### Média harmónica (MH)

Ao inverso da média aritmética dos inversos dos valores  $x_1, x_2, x_3, \dots, x_n$  da variável dá-se o nome de média harmónica.

$$MH = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Se os valores  $x_1, x_2, x_3, \dots, x_n$  estiverem com as suas respectivas frequências  $f_1, f_2, f_3, \dots, f_n$ ; a média harmónica será dada por:

$$MH = \frac{n}{\frac{1}{x_1} f_1 + \frac{1}{x_2} f_2 + \dots + \frac{1}{x_n} f_n} = \frac{n}{\sum_{i=1}^n \frac{f_i}{x_i}}$$

### Média quadrática

Consideremos os valores observados,  $x_1, x_2, x_3, \dots, x_n$ . A média quadrática é a raiz quadrada da média aritmética dos quadrados da variável.

$$MQ = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}{n}} = \sqrt{\frac{\sum x_i^2 \cdot f_i}{n}}$$

No caso dos valores observados estiverem acompanhados com as suas respectivas frequências  $f_1, f_2, f_3, \dots, f_n$ ; a média quadrática será dada por:

$$MQ = \sqrt{\frac{x_1^2 \cdot f_1 + x_2^2 \cdot f_2 + x_3^2 \cdot f_3 + \dots + x_n^2 \cdot f_n}{n}} = \sqrt{\frac{\sum x_i^2 \cdot f_i}{n}}$$

Os diferentes tipos de medias relacionam-se entre si da seguinte maneira:  $Mh \leq Mg \leq M \leq MQ$

### Importância de cada uma das medidas de tendência central

As três medidas de tendência central média, moda e mediana, pretendem localizar os valores em torno dos quais os dados se agrupam.

#### Média

- É uma medida de precisão, cuja definição e propriedades são tratadas por meios algébricos e que intervém em cálculos estatísticos avançados.
- É uma medida complexa, isto é, faz intervir todos os dados, pelo que é muito sensível a qualquer alteração de um ou

mais valores. O seu valor não pertence, geralmente ao conjunto inicial de dados (pode não ter existência real).

- É a medida mais usada para comprar distribuição, serve de padrão de comparação.

### Mediana

- Permite situar um indivíduo na metade inferior ou superior da população quanto ao carácter em estudo;
- É um parâmetro “robusto” não afectado por flutuações dos externos é uma medida de posição.
- Usa-se em geral quando as distribuições são muito assimétricas, ou seja, desequilibradas nos externos.

### Moda

- Indica imediatamente o valor ou a modalidade de maior efectivo ou frequência. É especialmente importante em estudos de mercado.
- É a medida mais fácil e rápida de observar. Mas, com dados agrupados, seu cálculo é mais trabalhoso.
- Aplica-se tanto a caracteres quantitativos como qualitativos.

Nota: uma distribuição estatística chama-se simétrica se a moda, média e mediana coincidem.

### Exemplo 46

$x_i$	$f_i$	$x_i \cdot f_i$
2	3	6
3	6	18
4	10	40
5	6	30
6	3	18
Total	28	112

- $\bar{X} = \frac{\sum f_i \cdot x_i}{N} = \frac{112}{28} = 4$
- $Mg = \sqrt[n]{x_1 \cdot f_1 \cdot x_2 \cdot f_2 \cdot x_3 \cdot f_3 \dots x_n \cdot f_n}$
- $\lg(Mg) = \frac{1}{n} \sum_{i=1}^n f_i \lg(x_i) = \frac{16 \cdot 14}{28} = 0,582 \Leftrightarrow \lg(Mg) = 0,582$

então  $Mg = 3,82$

$$Mh = \frac{n}{\frac{1}{x_1} \cdot f_1 + \frac{1}{x_2} \cdot f_2 + \dots + \frac{1}{x_n} \cdot f_n} = \frac{n}{\sum \frac{f_i}{x_i}} = \frac{28}{7,7} = 3,6$$

---

## Exercícios

**37.** Dados os números 2, 5, 7, 10, 5, determine:

- a) A média aritmética
- b) A média geométrica
- c) A média harmónica.

**38.** Utilizando a tabela de frequências da tabela abaixo, calcule:

- a) Moda
- b) Mediana
- c) Média aritmética
- d) Média geométrica
- e) Média quadrática
- g) Qual é o valor do 11º dado? E do 32º?
- h) Quantos dados são menores que 8?

$x_i$	2	4	6	7	8	9	10	11
$f_i$	10	12	8	11	10	5	5	9

### T.P.C. – Primeiras sessões presencias

1. Ler e estudar todas as páginas do texto.
2. Resolver todos os exercícios de cada unidade.
3. Entregue os exercícios: 2, 6, 9, 12 b), 13 a), 15, 17, 18, 20, 23, 27 a), b) e 34.

---

## Tema: Outras medidas de localização – quantis, quartis, decís e percentis

### Introdução

Os quartis, decís e percentis, são medidas de localização muito usadas no estudo das séries estatísticas. É sobre esse assunto que vamos falar nesta unidade, quer para variável descrita, quer para contínua.

Ao completar esta unidade, você será capaz de:



### Objectivos

- Saber calcular os quartis, decís e percentis
- Ser capaz de interpretar o significado dessas medidas

---

### Quartis

Enquanto que a mediana divide um conjunto de dados estatísticos ordenados em duas partes iguais, os quartis dividem-no em quatro. São, por isso, em número de três e representam-se, respectivamente, por  $Q_1$ ,  $Q_2$  e  $Q_3$

O segundo quartil corresponde a mediana, ou seja,  $Q_2 = Me$ . Para a determinação dos quartis, teremos que atender ao tipo de distribuição apresentada.

O cálculo de  $Q_2 = Me$  já foi tratado anteriormente vejamos como determinar  $Q_1$  e  $Q_3$

### Para dados não agrupados

O primeiro quartil é a mediana da primeira metade da distribuição. Supondo que esta apresenta um número de observação igual a  $N'$ , ordem de  $Q_1$  será:

$$\frac{N+1}{2} \text{ se } N' \text{ é ímpar.}$$

$$\frac{\frac{N'}{2} + \frac{N'+1}{2} + 1}{2} \text{ Se } N' \text{ par}$$

O terceiro, quartil é a mediana da última metade da distribuição. A ordem de  $Q_3$  pode ser obtida do seguinte modo:

$$\frac{N}{2} + \text{Ordem de } Q_1 \text{ se } N \text{ par}$$

$$\frac{N+1}{2} + \text{Ordem de } Q_1 \text{ se } N \text{ ímpar.}$$

Nota: se alguma das ordens encontradas não for inteira, o quartil respectivo determina-se pela média aritmética de  $x_k$  e  $x_{k+1}$  (ou seja,  $\frac{x_k + x_{k+1}}{2}$ ), sendo  $k$  o número inteiro imediatamente inferior ao número encontrado.

### Exemplo 17

Na distribuição de dados simples:

13 14 16 18 20 21 22 24 27

$$Q_2 = M_e = \frac{x_{\frac{n+1}{2}}}{2} = x_5 = 20$$

Como  $N' = 4$ , então o índice de  $Q_1$  é:

$$\frac{\frac{N'}{2} + \frac{N'+1}{2} + 1}{2} = \frac{2+3}{2} = 2,5 \text{ logo } Q_1 = \frac{x_2 + x_3}{2} = \frac{14+16}{2} = 15$$

$$\text{A ordem de } Q_3 \text{ é: } \frac{N+1}{2} + 2,5 = 5 + 2,5 = 7,5$$

$$\text{Portanto, } Q_3 = \frac{x_7 + x_8}{2} = \frac{22+24}{2} = 23$$

### Exemplo 18

Consideramos, agora, a distribuição:

20 22 25 28 29 30

$$Q_2 = M_e = \frac{x_3 + x_4}{2} = 26,5$$

$$\text{Ordem de } Q_1: \frac{N'+1}{2} = \frac{3+1}{2} = 2$$

$$Q_1 = x_2 = 22$$

$$\text{Ordem de } Q_3: \frac{N}{2} + 2 = 3 + 2 = 5$$

$$Q_3 = x_5 = 29$$

### Exemplo 19

Seja a distribuição de dados não agrupados apresentada na tabela:

$x_i$	$f_i$	$f_a$
1	3	3
2	5	8
3	12	20
4	7	27
5	5	32
Total	N=32	-

$$Q_2 = M_e = \frac{x_{16} + x_{17}}{2} = \frac{3+3}{2} = 3$$

$$\text{Ordem de } Q_1: \frac{\frac{N'}{2} + \frac{N'+1}{2}}{2} = \frac{8+9}{2} = 8,5 \quad \text{logo}$$

$$Q_1 = \frac{x_8 + x_9}{2} = \frac{2+3}{2} = 2,5$$

$$\text{Ordem de } Q_3: \frac{N}{2} + 8,5 = 24,5 \quad \text{logo}$$

$$Q_3 = \frac{x_{24} + x_{25}}{2} = \frac{4+4}{2} = 4$$

### Exemplo 20

Dada a distribuição:

$x_i$	$f_i$	$f_a$
1	3	3
2	5	8
3	2	10
4	18	28
5	3	31
Total	N =31	-

$$Q_2 = M_e = \frac{x_{31+1}}{2} = x_{16} = 3$$

$$\text{Ordem de } Q_1 : \frac{N'+1}{2} = \frac{15+1}{2} = 8 \text{ logo } Q_1 = x_8 = 2$$

$$\text{Ordem de } Q_3 : \frac{N+1}{2} + 8 = 24 \text{ logo } x_3 = x_{24} = 4$$

### Para dados agrupados

As considerações feitas para a determinação da mediana de uma distribuição de dados agrupados aplicam-se também na determinação de  $Q_1$  e  $Q_3$ . Quer N seja par ou ímpar

$$Q_1 = X_{N/4} \quad Q_3 = X_{3N/4}$$

Tratando-se de dados agrupados em classes obtém-se os quartis através da formula.

### Exemplo 21

Seja dada a distribuição de dados agrupados apresentada na tabela:

Classes	$f_i$	$f_a$
$[1,55; 1,60[$	5	5
$[1,60; 1,65[$	11	16
$[1,65; 1,70[$	23	39
$[1,70; 1,75[$	8	47

[1,75;1,80[	3	50
Total	N = 50	-

$Q_1 = x_{\frac{N}{4}} = x_{12,5} \rightarrow$  o primeiro quartile encontra-se na classe [1,60;1,65[

$Q_3 = x_{\frac{3N}{4}} = x_{37,5} \rightarrow$  O terceiro quartile encontra – se na classe [1,65;1,70[

Tratando – se de dados agrupados em classes obtêm – se os quartis através da formula:

$$Q_c = l_{i-1} \cdot \frac{C \frac{N}{4} - N_{i-1}}{F_i} \cdot h_i$$

$$C = 1, 2, 3 \Rightarrow Q_1, Q_2, Q_3$$

As letras que entram nessa formula têm o mesmo significa que na mediana.

$$Q_c = l_{i-1} \cdot \frac{C \frac{N}{4} - N_{i-1}}{F_i} \cdot h_i$$

### Amplitude interquartis

Chama – se amplitude interquartis á diferença entre o terceiro e primeiro quartis.

$$A_Q = Q_3 - Q_1$$

### Diagrama de extremos e quartis

O diagrama de extremos e quartis é um esquema que nos permite ver se há concentração ou dispersão dos dados ao longo de toda a distribuição.

## Exemplo 22

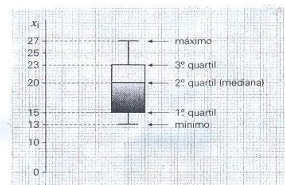
Seja a distribuição:

13 14 16 18 20 21 22 24 27 cujos valores dos quartis já foram determinados:

$$Q_1 = 15, Q_2 = M_e = 20, Q_3 = 23$$

O diagrama de extremos e quartis desta distribuição é o seguinte:

O diagrama de extremos e quartis desta distribuição é o seguinte:



Verifica-se que os dados estão mais concentrados antes de  $Q_1$  e entre  $Q_2$  e  $Q_3$ .

### Decis e Percentis

Os decis dividem o conjunto de dados em dez partes iguais e percentil divide em cem partes iguais.

Para dados não agrupados procede-se como os quartis.

Normalmente os decis e os percentis, só têm interesse para um grande número de valores, isto é, para dados agrupados.

A expressão que nos dá os decis é:

$$D_c = l_{i-1} \cdot C \frac{\frac{N}{10} - N_{i-1}}{F_i} \cdot h_i, \quad C = 1, 2, 3, \dots, 9 \Rightarrow D_1, D_2, D_3, \dots, D_9$$

A expressão que nos dá os pertences é dada por:

$$P_c = l_{i-1} \cdot \frac{C}{100} \frac{N - N_{i-1}}{F_i} \cdot h_i$$

$$C = 1, 2, 3, \dots, 99 \Rightarrow P_1, P_2, P_3, \dots, P_{99}$$

### Exemplo 23

É dada a distribuição de salários (semanais) de 100 funcionários duma empresa.

Salário em dolar	[5,10[	[10,15[	[15,20[	[20,25[	[25,30[
Número de funcionários	26	43	17	9	5

Calcule o  $P_{45}$  e interprete o resultado.

## Resolução

Primeiro vamos calcular a frequência acumulada para saber em que classe se localiza a mediana.

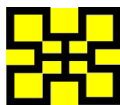
Salário em dólar ( $x_i$ )	$[5,10[$	$[10,15[$	$[15,20[$	$[20,25[$	$[25,30[$
Número de funcionário ( $f_i$ )	26	43	17	9	5
Frequência acumulada ( $f_a$ )	26	69	86	95	100

$$\frac{N}{2} = \frac{100}{2} = 50 \text{ logo a classe mediana é a } 2^{\text{a}} \text{ classe } [10,15[.$$

$$P_{45} = 10 \cdot \frac{45 \cdot \frac{100}{100} - 26}{43} \cdot 5 = 12,2. \text{ Isto significa que 45\% dos funcionários tem salários até 12,2 dólares.}$$

---

## Exercícios



Auto-avaliação

**38.** Determine a mediana e os quartis do seguinte conjunto de dados.

10, 11, 12, 13, 14, 15, 16, 20, 21, 22, 23, 24, 25, 26, 27.

**39.** Determine o 4º decil e o 62º percentil da seguinte distribuição:

Classes	$f_i$	$f_a$
$[4,9[$	8	
$[9,14[$	12	
$[14,19[$	17	
$[19,24[$	3	

---

## Introdução

As medidas de tendência central estudadas reduzem a série de dados a um só valor típico (media, mediana e moda). É necessário completar o estudo de uma distribuição estatística com outras medidas que permitem determinar o grau de dispersão dos dados em torno dos valores centrais.

Ao completar esta unidade, você será capaz de:



### Objectivos

- Conhecer as unidades de dispersão mais usadas.
- Ser capaz de usar e interpretar as medidas de dispersão.

---

## Unidade Temática: Medidas de Dispersão ou de Variabilidade

### Medidas de dispersão

As medidas de tendência central estudadas reduzem a série de dados a um só valor típico (media, mediana e moda).

O valor central mais usado é o valor médio, que contudo, nem sempre dá uma ideia suficiente da série estatística.

Ora vejamos o seguinte exemplo:

As notas de dois estudantes da 10ª classe no 1º trimestre são as seguintes:

9; 9,6; 10,6; 10,8; 11; 11,1; 11,4; 11,6; 12; 13

Y

5; 6,4; 8; 9; 10,6; 12; 12,2; 14,2; 15,6; 17

Calculando a média vimos que os dois estudantes têm o mesmo valor médio  $\bar{X} = 11$ . Mas a distribuição dos valores é muito diferente. Os valores do y são muito dispersos, enquanto que os de x são mais concentrados. Assim, a media é mais representativa para x do que para y.

É necessário completar o estudo de uma distribuição com outras medidas que permitem determinar o grau de dispersão dos dados em torno dos valores centrais.

As medidas de dispersão mais usadas são: Amplitude total, Desvio quartílico, desvio médio, variância e desvio padrão.

**Amplitude total (range)** – é a diferença entre o valor máximo e mínimo.

$$R = X_{m\acute{a}x} - X_{m\acute{i}n}$$

**Interquartil** – é a diferença entre o 3º e o 1º quartil, isto é,

$$Q_3 - Q_1$$

**Semi amplitude quartílica** é a metade do interquartil.

$$\frac{Q_3 - Q_1}{2}$$

**Desvio quartílico relativo**  $Q = \frac{Q_3 - Q_1}{Q_2}$

**Desvio de  $x_i$**  em relação a media  $\bar{x}$  é a diferença entre os valores observados e o valor médio.

$$d = x_i - \bar{x}$$

Chama-se **desvio médio** de uma distribuição ( $D_x$ ) a média aritmética dos valores absolutos dos desvios em relação a média.

$$D_{(x)} = \frac{\sum_{i=1}^n f_i \left( |x_i - \bar{x}| \right)}{N}$$

Chama-se **Variância**, a média aritmética dos quadrados dos desvios em relação a média aritmética.

$$V_{(x)} = \frac{\sum_{i=1}^n f_i \left( x_i - \bar{x} \right)^2}{N}$$

**Nota:** a variância é tanto maior quanto maior for a dispersão.

A representatividade dos valores centrais diminui quando a variância aumenta.

Chama-se **desvio padrão**, a raiz quadrada da variância.

$$\sigma = \sqrt{V_{(x)}} \quad \text{ou} \quad \sigma = \sqrt{\frac{\sum_{i=1}^n f_i \left( x_i - \bar{x} \right)^2}{N}}$$

A variância também pode ser dada pela fórmula.

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

### Importância do desvio padrão

- O desvio padrão informa sobre dispersão, isto é, sobre o afastamento dos dados em relação a média.

- O intervalo  $\left[ \bar{x} - \sigma, \bar{x} + \sigma \right]$  contém sempre mais de 50% dos dados; isto significa que mais de metade dos dados se situam a uma distância da média que é inferior a um desvio padrão.

### Coeficiente de variação ou de dispersão

O coeficiente de variação  $C_v$ , é uma medida de dispersão útil para a comparação em termos relativos, do grau de concentração em torno da média de séries distintas, e é

dado por: 
$$C_v = \frac{\sigma}{\bar{x}}$$

As vezes o  $C_v$  é dado em percentagem.

### Exemplo 24

Considere a distribuição estatística seguinte:  
-4 -2 0 1 3 5 7

- Determine a media aritmética dos valores apresentados
- Calcule o desvio médio.
- Determine os valores da variância e do desvio padrão

### Resolução

a) 
$$\bar{x} = \frac{-4 + (-2) + 0 + 1 + 3 + 5 + 7}{7} \approx 1,4$$

b) 
$$d = \frac{\sum_{i=1}^n f_i |x_i - \bar{x}|}{N}$$

$$|-4 - 1,4| = 5,4$$

$$|-2 - 1,4| = 3,4$$

$$|0 - 1,4| = 1,4$$

$$|1 - 1,4| = 0,4$$

$$|3 - 1,4| = 1,6$$

$$|5 - 1,4| = 3,6$$

$$|7 - 1,4| = 5,6$$

$$d = \frac{5,4 + 3,4 + 1,4 + 0,4 + 1,6 + 3,6 + 5,6}{7} \approx 3,06$$

c)

$x_i$	4	-2	0	1	3	5	7
$x_i - \bar{x}$	-5,4	-3,4	1,4	-0,4	1,6	3,6	5,6
$(x_i - \bar{x})^2$	29,16	11,56	1,96	0,16	2,56	12,96	31,36

$$V = \frac{\sum f_i(x_i - \bar{x})^2}{N} = \frac{89,72}{7} \approx 12,81$$

$$\sigma = \sqrt{12,81} \approx 3,58$$

### Exemplo 25

Considere a seguinte tabela de frequência absolutas:

$X_1$	9	10	12	13	15	16	17
$F_1$	4	5	10	12	4	4	1

a) Determine o valor da média aritmética

b) Calcule o desvio padrão.

### Resolução

O cálculo da média e do desvio padrão de uma distribuição do tipo da que se apresenta neste exercício pode ser feito com base no quadro seguinte:

$X_i$	$F_i$	$X_i \cdot F_i$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$F_i(X_i - \bar{X})^2$
9	4	36	-3,575	12,78	51,12
10	5	50	-2,575	6,63	33,15
12	10	120	-0,575	0,33	3,3
13	12	156	-0,425	0,18	2,16
15	4	60	2,425	5,88	23,52
16	4	64	3,425	11,73	46,58
17	1	17	4,425	19,58	19,58
Total	$N = 40$	503	-----	-----	179,75

$$a) \bar{X} = \frac{\sum_{i=1}^n f_i \cdot x_i}{N} = \frac{503}{40} = 12,575$$

$$b) \delta = \sqrt{\frac{\sum_{i=1}^n f_i (x_i - \bar{X})^2}{N}} = \sqrt{\frac{179,75}{40}} \approx 2,12$$

\* A variância é  $\delta^2 = \frac{179,75}{40} = 4,49375 \approx 4,49$

### Exemplo 26

Considere o seguinte histograma de uma distribuição estatística:

### Resolução

O cálculo da média e do desvio padrão de uma distribuição do tipo da que se apresenta neste exercício pode ser feito com base no quadro seguinte:

$X_i$	$F_i$	$X_i \cdot F_i$	$X_i \cdot X$	$(X_i - X)^2$	$F_i(X_i - X)^2$
9	4	36	-3,575	12,78	51,12
10	5	50	-2,575	6,63	33,15
12	10	120	-0,575	0,33	3,3
13	12	156	-0,425	0,18	2,16
15	4	60	2,425	5,88	23,52
16	4	64	3,425	11,73	46,58
17	1	17	4,425	19,58	19,58
Total	N = 40	503	-----	-----	179,75

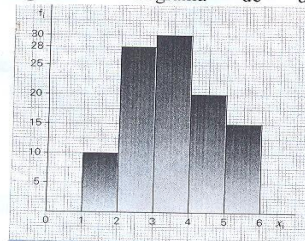
$$a) \bar{X} = \frac{\sum_{i=1}^n f_i \cdot x_i}{N} = \frac{503}{40} = 12,575$$

$$b) \delta = \sqrt{\frac{\sum_{i=1}^n f_i (x_i - \bar{X})^2}{N}} = \sqrt{\frac{179,75}{40}} \approx 2,12$$

\* A variância é  $\delta^2 = \frac{179,75}{40} = 4,49375 \approx 4,49$

### Exemplo 26

Considere o seguinte histograma de uma distribuição estatística:



- Com base no gráfico apresentado, elabore uma tabela de frequências absolutas, simples e acumuladas.
- Indique a classe modal e a classe mediana.
- Determine o valor da média e calcule o desvio padrão
- Calcule o coeficiente de variação.

### Resolução

a)

Classes	$f_i$	$f_{ac}$
$[1, 2[$	10	10
$[2, 3[$	28	38
$[3, 4[$	30	68
$[4, 5[$	20	88

[5,6[	15	103
Total	N = 103	-

a) A classe modal é [3,4[

b)  $\frac{N}{2} = \frac{103}{2} = 51,5 \rightarrow$  a classe modal é [3,4[

c)

Classes	fi	xi	Fi.xi	xi - $\bar{x}$	(xi - $\bar{x}$ ) <sup>2</sup>	Fi (xi - $\bar{x}$ ) <sup>2</sup>
[1,2[	10	1,5	15	-2	4	40
[2,3[	28	2,5	70	-1	1	28
[3,4[	30	3,5	105	0	0	0
[4,5[	20	4,5	90	1	1	20
[5,6[	15	5,5	82,5	2	4	60
Total	N = 103	-	362,5	0	-	148

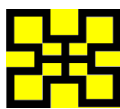
$$\bar{x} = \frac{\sum f_i \cdot x_i}{N} = \frac{362,5}{103} \approx 3,5$$

$$\sigma = \sqrt{\frac{148}{103}} \approx \sqrt{1,44} \approx 1,2$$

$$C_v = \frac{\sigma}{\bar{x}} = \frac{1,2}{3,5} \approx 0,34 \text{ ou } C_v \approx 34\%$$

---

## Exercícios



Auto-avaliação

40. Para a série 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 7, 8, 8, 8, 9, 9.

a) Construa uma tabela de frequências

b) Calcule a amplitude.

c) Determine o desvio médio.

d) Calcule a variância e o desvio padrão.

**41.** Considere as seguintes distribuições:

A: 2      4      5      6      7      9      11

B: 1,3      2,4      0,5      3      1,1      6      0,2      2,5

C: -3      4      -5      5      -1      -4      0      4  
10

Determine a mediana, o primeiro e o terceiro quartis dos dados apresentados.

**42.** Num inquérito efectuado a 1000 trabalhadores, referente à distância do local de trabalho à residência, obteve-se a seguinte distribuição de frequências:

Distância (km)	Número de trabalhadores
$[0,5[$	353
$[5,10[$	159
$[10,15[$	255
$[15,20[$	147
$[20,25[$	59
$[25,30[$	27
Total	N = 1000

a) Determine a média e indique a classe modal dos dados apresentados.

b) Localize os quartis.

c) Determine o valor da mediana.

d) Determine a percentagem de trabalhadores que percorre distâncias acima das da classe mediana.

**44.** Considere as seguintes distribuições:

A: 12, 6, 7, 3, 15, 18, 10,  
5.

B: 9, 3, 8, 8, 9, 8, 9,  
18.

Relativamente a cada uma delas, determine:

- a) A amplitude total.
- b) O desvio médio.
- c) O desvio padrão.

**45.** Considere a seguinte tabela de frequência absolutas:

$x_i$	$f_i$
0	12
1	16
2	27
3	20
4	16
5	9
TOTAL	N = 100

Determine:

- a) A amplitude total e a média da distribuição
- b) A moda e a mediana.
- c) O coeficiente de variação.

**46.** Considere a seguinte tabela de dados agrupados:

<b>Classes</b>	<b><math>f_i</math></b>
20-25	9
25-30	27
30-35	36
35-40	45
40-45	18
45-50	9
50-55	3
55-60	3
Total	N=150

Determine:

- a) A amplitude total.
- b) A media e a moda.
- c) A variância e o desvio padrão.

d) O coeficiente de variação.

---

## Unidade Temática: Medidas de Assimetria e Curtose

### Introdução

Assimetria é o grau de desvio em relação a uma distribuição simétrica. É nesta unidade onde vamos falar dos diferentes casos de assimetria e dos graus de achatamento – curtose de uma distribuição.

Ao completar esta unidade, você será capaz de:



#### Objectivos

- Conhecer os coeficientes de assimetria e de curtose.
- Ser capaz de classificar distribuições quanto ao grau de desvio e achatamento.

---

### Mediadas de Assimetria

Assimetria é o grau de desvio em relação a uma distribuição simétrica.

Ao representarmos graficamente uma distribuição de frequências, três casos são possíveis:

1º Caso:  $\bar{x} = M_e = M_o$ . Estamos perante uma distribuição simetria.

2º Caso  $\bar{x} < M_e < M_o$ . Estamos perante uma distribuição assimétrica

a esquerda ou negativa.

$3^o M_o < M_e < \bar{x}$  Estamos perante uma distribuição assimétrica a direita ou positiva.

### **Coefficiente de Assimetria ( $A_s$ )**

**Karl Pearson** propôs duas maneiras de avaliar o grau de achatamento ou de formação da curva de uma distribuição de frequências, cujo objectivo principal é justamente indicar a grandeza do afastamento em termos relativos.

#### **a) Primeiro coeficiente**

$$A_s = \frac{\bar{x} - M_o}{\sigma} \text{ Onde,}$$

$\bar{x}$  = Média Aritmética

$M_o$  = Moda

$\sigma$  = Desvio padrão da amostra.

#### **b) Segundo coeficiente**

$$A_s = \frac{3(\bar{x} - Med)}{\sigma} \text{ ou } A_s = \frac{3\bar{x} - 3Med}{\sigma}$$

onde, Med = Mediana

Em função dos resultados de ( $A_s$ ), é possível determinar o comportamento da curva de cada distribuição. Assim, se:

$A_s = 0$  A distribuição é simétrica.

$A_s > 0$  A distribuição é assimétrica positiva.

$A_s < 0$  A distribuição é assimétrica negativa.

### Exemplo 26

Consideremos a seguinte tabela de distribuição de frequências:

Notas	$f_i$
$[0,2[$	30
$[2,4[$	20
$[4,6[$	20
$[6,8[$	10
$[8,10]$	10
Total	$\sum f_i = 90$

Temos:

Para o primeiro coeficiente de assimetria

$$A_s = \frac{\bar{x} - M_o}{\sigma} = \frac{3,89 - 1,50}{2,32} = 0,885$$

Para o segundo coeficiente de assimetria

$$A_s = \frac{3(\bar{x} - Med)}{\sigma} = \frac{3(3,89 - 3,50)}{2,32} = 0,43 \text{ Logo, } A_s > 0 \text{ a}$$

distribuição é assimetria positiva.

### Exemplo 27

Consideremos a seguinte tabela de distribuição de frequências:

Notas	$f_i$
$[0,2[$	10
$[2,4[$	10
$[4,6[$	20
$[6,8[$	20
$[8,10]$	30
Total	$\sum f_i = 90$

Temos:

Para o primeiro coeficiente de assimetria

$$A_s = \frac{\bar{x} - M_o}{\sigma} = \frac{6,11 - 8,50}{2,70} = -0,885$$

$$A_s = \frac{3(\bar{x} - Med)}{\sigma} = \frac{(6,11 - 6,50)}{2,70} = -0,43 \text{ logo, } A_s > 0,$$

a distribuição é assimetria negativa.

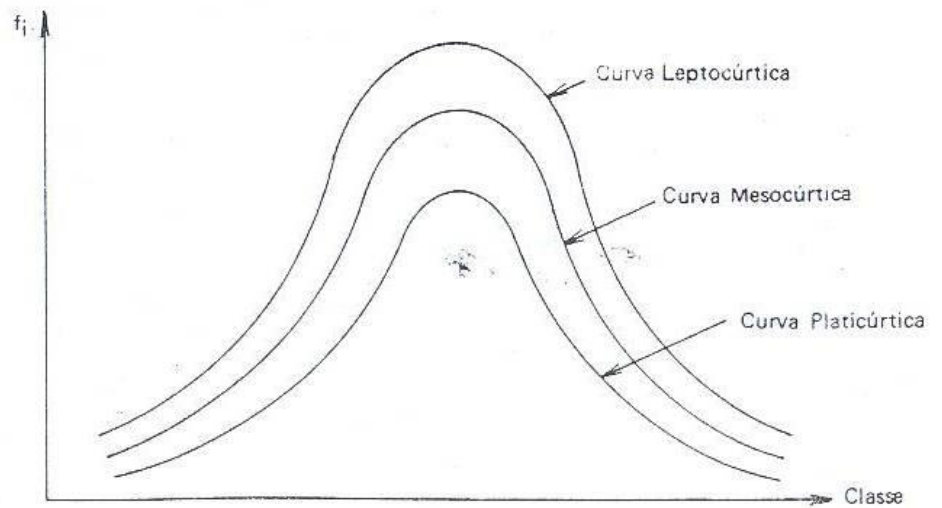
### **Medidas de Achatamento – Curtose**

Curtose nada mais é do que o grau de achatamento da curva de uma distribuição de frequências. Isto considerando que uma curva pode apresentar –se mais achatada ou mais afilada em relação a uma curva considerada curva – padrão ou curva normal.

Assim, dentro destas especificações, uma curva normal deve apresentar um coeficiente de curtose igual a 0,263 (mesocurtica) , e , se o mesmo coeficiente for menor que 0,263, a curva devera-se apresentar mais aguda que a normal (Leptocúrtica).

Graficamente, teremos:

Graficamente, temos:



145

O cálculo deste coeficiente é dado pela seguinte expressão:

$$C = \frac{Q_3 - Q_1}{C_{90} - C_{10}} \text{ ou } C = \frac{D_q}{C_{90} - C_{10}} \text{ onde } D_q = \text{desvio quartílico.}$$

### Exemplos 28

Para a tabela do exemplo 26, teremos:

$$Q_3 = 5,75$$

$$Q_1 = 1,50$$

$$C_{90} = 8,20$$

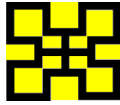
$$C_{10} = 0,60$$

$$\text{Logo, } C = \frac{D_q}{C_{90} - C_{10}} = \frac{2,125}{7,6} = 0,2796$$

Então,  $0,2796 > 0,263 =$  curva platicútica.

---

## Exercícios



Auto-avaliação

47. Dada a distribuição:

Classes	[2,6[	[6,10[	[10,14[	[14,18[	[18,22[	[22,26[
Frequências	2	3	8	10	12	5

- Calcula a moda, mediana e a média aritmética.
- Calcule o desvio padrão e o coeficiente de variação.
- Classifique a distribuição quanto a assimetria.
- Classifique a distribuição quanto ao achatamento (curtose).

48. Cronometrado o tempo para várias gincanas automobilísticas encontrou-se:

Equipa X:

40 Provas; tempo médio 45 segundos; variância 400.

Equipa y

<b>Tempo (em seg.)</b>	<b>20</b>	<b>40</b>	<b>50</b>	<b>80</b>
11º De pessoas	10	15	30	5

Qual é a equipa que apresenta resultados mais homogéneos?

Justifique.

**49.** Dada a amostra de 60 rendas em Euros, de uma dada região.

10	7	8	5	4	3	2	9	9	6
3	15	1	13	14	4	3	6	6	8
10	11	12	13	14	2	15	5	4	10
2	1	3	8	10	11	13	14	15	16
8	9	5	3	2	3	3	4	4	4
5	6	7	8	9	1	12	13	14	16

- Agrupar os elementos em classes. Faça  $K=6$  e  $h=3$
- Calcule a média e a mediana.
- Calcule a medida que deixa 25% das rendas.
- Determine  $Q_3$ ,  $Q_4$ ,  $P_{47}$  e, interprete o resultado.
- Determine o desvio padrão.
- Determine a variância e o desvio padrão.
- Qual é o valor do coeficiente de variação.
- A distribuição é simétrica?
- A distribuição é mesocúrtica?

## TEMA II: PROBABILIDADE (Noções) Unidade Nº 2.1.

---

**Unidade**      **Temática:**      **Distribuições**  
**bidimensionais**

### Introdução

Nesta unidade vamos falar das distribuições com duas variáveis, o modo como essas se relacionam mutuamente. Teremos que desenhar a nuvem de pontos, a recta de regressão e calcular o coeficiente de correlação.

Ao completar esta unidade, você será capaz de:



### Objectivos

- Conhecer os tipos de correlação existentes.
- Ser capaz de desenhar a nuvem de pontos e a recta de regressão.
- Ser capaz de calcular o coeficiente de correção.

---

### Distribuições Bidimensionais

O estado estatístico refere-se muitas vezes a dois caracteres da mesma população visando investigar em que medida eles se relacionam, isto é, de que modo a variação de um deles exerce influência na variação do outro.

Por exemplo:

1. Os gastos em publicidade e o volume de venda.
2. Altura média de um casal e a altura dos filhos.
3. Altura de um local e a respectiva pressão atmosférica.
4. Hora do dia e a temperatura atmosférica.
5. O peso dos alunos e a respectiva nota em filosofia.
6. Peso e o preço do telemóvel

As distribuições estatísticas envolvendo o estudo da relação entre duas variáveis chamam-se distribuições bidimensionais.

Há caracteres entre os quais é impossível descobrir qualquer relação, como por exemplo, o peso dos alunos e a respectiva nota em filosofia. Diz-se que estes são caracteres independentes.

Diz-se que há influência de um carácter neutro.

O maior ou menor grau de dependência estatística entre duas variáveis traduz-se por um número – o coeficiente de correlação que só varia de -1 a 1.

**Correlação positiva** – as variáveis evoluem em geral no mesmo sentido.

Exemplo: altura e o peso das pessoas.

**Correlação negativa** – as variáveis evoluem em sentido contrário quando uma aumenta a outra tende a diminuir.

Exemplo: intensidade da chuva e a temperatura do ar.

**Correlação nula** – não há influência de uma variável na outra, as variáveis são independentes.

Exemplo: a idade do pai e a nota de desenho de cada aluno.

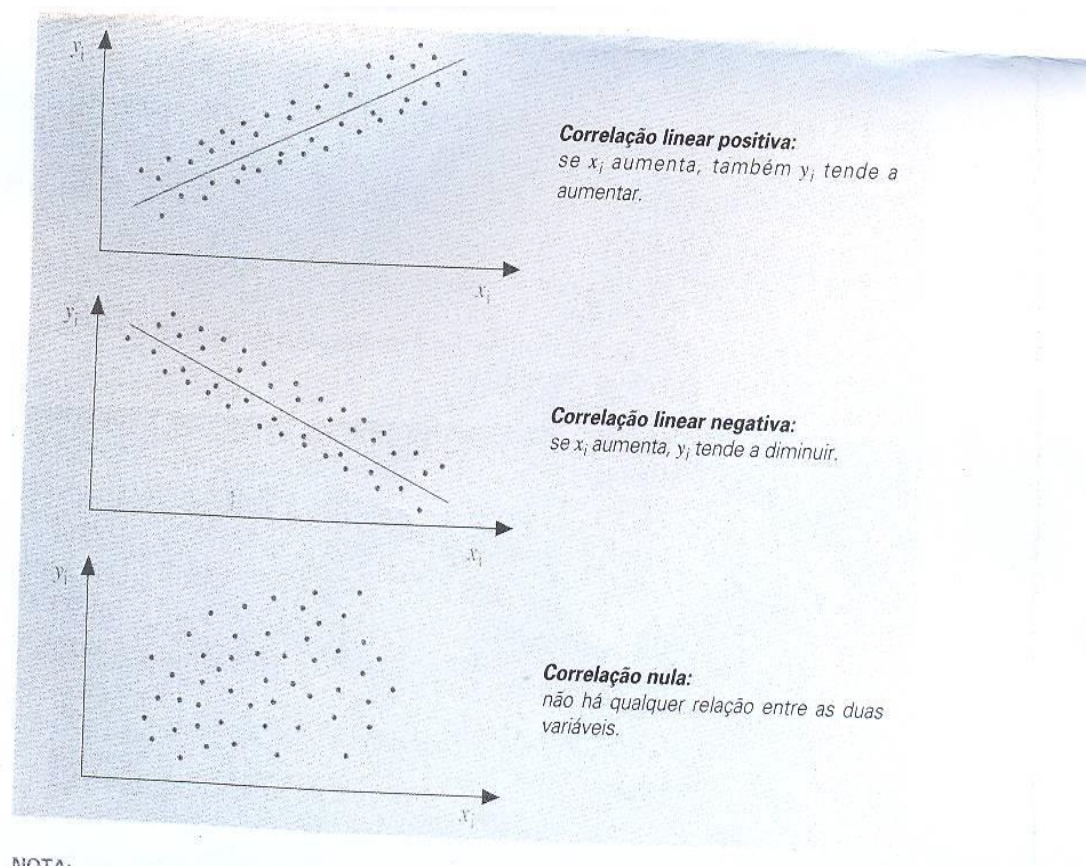
Se o coeficiente de correlação for igual a 1 ou a -1, então a dependência entre duas variáveis é tão rigorosa que se pode traduzir por uma lei matemática: diz-se que há uma **dependência funcional**.

A **dependência funcional** é um caso extremo da dependência estatística.

De uma distribuição bidimensional de variáveis  $x_i$  e  $y_i$  podemos fazer uma representação gráfica num sistema de eixo dos  $x'$  e os da outra variável no eixo dos  $y'$ .

A essa representação chama-se **diagrama de dispersão**.

Se todos os pontos desse diagrama se situar nas proximidades de uma recta (recta de regressão), a correlação diz-se linear.



NOTA:

### Coefficiente de Correlação

Intuitivamente podemos prever a existência de correlação entre duas variáveis. Para quantificar a relação usa-se o coeficiente de correlação linear de Pearson que se representa por  $\rho$ .

$$\rho = \frac{C_{x,y}}{\sigma_x \cdot \sigma_y} \text{ Onde:}$$

$$C_{x,y} = \frac{\sum_{i=1}^n f_i x_i \cdot y_i}{N} - \bar{x} \cdot \bar{y} \text{ (Co - variância ou variância conjunta das variáveis } x \text{ e } y)$$

- $\sigma_x =$  Desvio padrão da variável x.
- $\sigma_y =$  Desvio padrão da variável y.

O coeficiente de correlação  $\rho$  é um número do intervalo  $[-1,1]$

### Exemplo 29

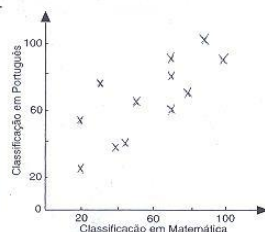
A tabela seguinte mostra a classificação de 12 alunos em dois testes cotados de 0 a 100.

Classificação em Matemática (x)	30	18	70	20	80	45	70	100	50	85	70	40
Classificação em Português (y)	75	24	60	54	70	40	90	90	65	100	80	37

- Desenhe o diagrama de dispersão.
- Calcule o coeficiente de correlação entre x e y.
- Interprete o resultado obtido em b).

### Resolução

a)



b)

$$\sum x = 678, \sum x^2 = 46074$$

$$\sum y = 785, \sum y^2 = 57411$$

$$\sum_{x,y} = 49492$$

$$C_{x,y} = \frac{\sum xy}{12} - \frac{\sum x}{12} \cdot \frac{\sum y}{12} = 428,3$$

$$\sigma_x = \sqrt{\frac{\sum x^2}{12} - \left(\frac{\sum x}{12}\right)^2} = 25,4$$

$$\sigma_y = \sqrt{\frac{\sum y^2}{12} - \left(\frac{\sum y}{12}\right)^2} = 22,5$$

$$\rho = \frac{C_{x,y}}{\sigma_x \cdot \sigma_y} = \frac{428,3}{25,4 \cdot 22,5} = 0,75$$

c) Há uma correlação positiva entre as variáveis.

### Exemplo 30

Os pares de valores  $(x,y)$  de uma variável bidimensional ocorrem com as frequências dadas pela seguinte tabela de dupla entrada.

$y \backslash x$	2	3	4
1	3	1	0
2	1	4	1
3	2	0	2

- Construa uma tabela simples a partir dos dados da tabela dada.
- Calcule o coeficiente de correlação entre  $x$  e  $y$ .

### Resolução

a)

X	Y	F
1	2	3
1	3	1
2	2	1
2	3	4
2	4	1
3	2	2
3	4	2

- Vamos construir uma tabela onde apareçam valores que ajudam ao cálculo do coeficiente de correlação.

x	y	f	f * y
1	2	3	6
1	3	1	3
2	2	1	4
2	3	4	24
2	4	1	8
3	2	2	12
3	4	2	24
		$\sum f = 14$	$\sum f \times y = 81$

Para a variável x, temos:

$$\bar{x} = 2, \sigma_x = 0,756$$

Para a variável y, temos:

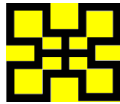
$$\bar{y} = 2,857; \quad \sigma_y = 0,833 \quad \text{Logo}$$

$$C_{x,y} = \frac{\sum f \times y}{\sum f} - \bar{x} \cdot \bar{y} = \frac{81}{14} - 2 \cdot 2,857 = 0,07$$

$$\text{Assim } \rho = \frac{C_{x,y}}{\sigma_x \cdot \sigma_y} = \frac{0,07}{0,756 \cdot 0,83} = 0,11$$

---

## Exercícios



Auto-avaliação

- 50.** Para cada uma das situações, diga se, na sua opinião, haveria interesse em realizar um estudo estatístico sobre a relação entre:
- O tráfego nas ruas que dão acesso a uma escola e o número de alunos que chegam atrasados a essa escola.
  - O número de televisores vendido por uma loja de uma cidade e o número de passageiros que viajam nos transportes públicos dessa cidade.
  - O número de pessoas que vivem num prédio e o custo da electricidade de todos os andares do prédio.
  - O número de flores de um jardim e o número de abelhas no mesmo jardim.
  - O número de horas de estudo da disciplina de Matemática e a nota em Matemática.
  - A marca da máquina de barbear que os alunos de uma turma usam e o número de pretendentes.

**51.** Construa uma tabela de dupla entrada a partir dos dados de cada uma das seguintes tabelas:

a)

$x_i$	$y_i$	$f_i$
1	4	3
2	5	5
3	6	2
		N=10

b)

$x_i$	$y_i$	$f_i$
[0,3]	[45,90]	1
[3,6]	[90,120]	3
[6,9]	[120,160]	4
		N = 8

c)

$x_i$	$y_i$	$f_i$
5	2	4
5	3	7
6	1	2
6	2	5
		N = 18

**52.** Construa uma tabela normal (simples) a partir dos dados de cada uma das seguintes tabelas de dupla entrada:

a)

$y_i \backslash x_i$	4	5	6	
1	-	2	-	2
2	1	1	2	4
	1	3	2	6

b)

$y_i \backslash x_i$	1	2	3	
-2	1	-	2	3
-3	1	1	-	2
-4	2	2	1	5
	4	3	3	10

**53.** Para cada uma das seguintes tabelas, represente o diagrama de dispersão.

a)

$x_i$	1	2	3	4	5
$y_i$	4	3	2	1	1

b)

$x_i$	1	2	3	4	5
$y_i$	3	2	4	1	3

**54.** Diga, na sua opinião, se será positiva, negativa ou nula a correlação entre:

- a) A altura de uma pessoa e a sua idade até aos vinte anos;
- b) O número de anos de um carro e o seu custo;
- c) O número de idas à praia e a medida do sapato;
- d) A temperatura de uma sala e a altura do mercúrio num termómetro colocado nessa sala.

**55.** A tabela seguinte mostra a classificação obtida por 18 alunos de uma turma no último teste das disciplinas de Matemática e Física.

Classificação em Matemática, %	40	55	20	70	60	25	55	70	85	35	45	90	10	45	100	10	60	65
Classificação em Física, %	60	65	40	65	80	50	65	85	80	50	50	95	30	65	95	40	85	70

- a) Desenhe, de acordo com os dados, o diagrama de dispersão.
- b) O António teve 55% no teste de Matemática mas faltou ao teste de Física. Com base no diagrama que desenhou em a) que percentagem ele poderia esperar obter no teste de Física?

56. Considere a seguinte variável bidimensional constituída pelas variáveis  $x_i$  e  $y_i$  dadas pelas tabelas:

$x_i$	2	3	4	5	6
$y_i$	3	4	5	3	6

Calcule:

- A média da variável  $x$  e a da variável  $y$ ;
- O desvio padrão da variável  $x$  e o da variável  $y$ ;
- A co-variância das variáveis  $x$  e  $y$ ;
- O coeficiente de correlação.

57. De um conjunto de 50 pares  $(x, y)$  sabe-se que:

$$\sum x = 100, \sum y = 900, \sum x^2 = 208, \sum y^2 = 16650 \text{ e } \sum xy = 1836$$

Calcule:

- $\bar{x}$  e  $\bar{y}$
- $\sigma_x$  e  $\sigma_y$
- $C_{x,y}$  e  $\rho$

**Segundas sessões presenciais**

**T. P. C.**

- 1.º Ler e 2.º estudar todas as páginas do referido texto.
2. Resolver todos os exercícios.

**3. Entregar os exercícios números: 43, 45, 46, 48, 54 e 56,**

---

## Unidade Temática: Noção intuitiva e frequencista das probabilidades

### Introdução

Existem duas noções até agora usadas para compreender melhor o conceito de probabilidades. Nesta primeira unidade do estudo das probabilidades vamos falar dessas duas noções e, faremos a menção dos principais termos usados na probabilidade.

Ao completar esta unidade, você será capaz de:

- Conhecer a noção intuitiva e frequencista da probabilidade.
- Ser capaz de calcular a probabilidade de um acontecimento.
- Ser capaz de distinguir acontecimentos incompatíveis ou contrários.



### Objectivos

---

## Sumário

### Noção intuitiva das probabilidades

Os jogos de azar que se praticam há milhares de anos, estão na origem das primeiras obras escritas sobre probabilidades e que começaram a surgir no século XVIII.

O dado é um instrumento dos jogos de azar por excelência, a tal ponto que duas das palavras mais usadas para lidar com o que é casual, ou fortuito, ou contingente são:

**Azar** de **az** – **zã** e **Aleatório** de **álea**, ambas com origem na palavra dado, respectivamente em árabe e latim.

Por esta sua origem concreta, a teoria das probabilidades utiliza uma linguagem própria que é preciso conhecer.

Uma experiência diz-se aleatória se pode ter vários resultados cujo aparecimento, em cada prova da experiência, é impossível prever.

Diz-se que o resultado depende do “acaso” ou da “sorte” ou do “azar”.

Os fenómenos aleatórios são o objecto de estudo da teoria das probabilidades.

**Acaso** significa ausência de causa conhecida.

São aleatórias todas as experiências que se fazem nos jogos:

- Tirar uma carta no baralho; atirar um dado; comprar uma rifa; rodar a roleta; jogar no totoloto...

Mas há muitos outros como:

- Abrir um livro e ver o número de página, perguntar a um aluno qualquer quantos irmãos tem...

As experiências do laboratório que se fazem na escola, cujo resultado já é conhecido e é sempre o mesmo, não são experiências aleatórias.

Para compreender melhor o conceito de probabilidade, há experiências aleatórias muito simples que podemos realizar com um grupo de trabalho, porque devem ser repetidas muitas vezes e em condições análogas; só assim é possível tirar conclusões significativas.

Vamos descrever algumas experiências e atribuir a probabilidade aos vários resultados que podem dar, introduzindo a terminologia adequada.

### **1ª Experiência – Lançamento de dado**

Lançamento de um dado sobre uma superfície lisa, plana e horizontal.

Qualquer das 6 faces pode ficar voltada para cima. Diz – se que esta experiência tem 6 resultados possíveis ou 6 casos possíveis.

O conjunto  $E = \{1, 2, 3, 4, 5, 6\}$  é de todos os resultados possíveis, é o espaço amostral desta experiência aleatória.

Qualquer elemento de  $E$ , é então um caso possível.  
 $3 \in E$  logo 3 é um caso possível.

A saída de 3 é um **acontecimento elementar** que se designa por  $\{3\}$ .

A saída de um número par também é um acontecimento mas não é elementar porque sair par equivale a sair 2, ou sair 4, ou sair 6. Tratando os acontecimentos em termos de conjunto de resultados temos:

Acontecimento sair número par  $\gg = \{2\} \cup \{4\} \cup \{6\}$ ; logo é uma reunião de acontecimentos elementares.

**Acontecimento** é qualquer subconjunto de  $E$ , incluindo  $A$  numa prova, significa que o resultado dessa prova é um elemento de  $A$ .

Nota: Diz-se que  $A$  se realizou quando o resultado é elemento  $A$ .

$A \cup B$  verifica – se quando ou se realiza  $A$  ou  $B$ .

$A \cap B$  verifica – se quando  $A$  e  $B$  se realizam simultaneamente.

### **Noção frequencista**

Suponha que o dado é perfeito, ou seja, rigorosamente simétrico, sem faces privilegiadas, então, é natural esperar que jogando o dado muitas vezes, o resultado de cada face seja aproximadamente o mesmo, visto que todas tem a mesma possibilidade de ficar para cima.

Faça essa experiência no seu grupo.

Componha uma tabela de frequências relativas dos resultados dos resultados do lançamento.

Acontecimento	Frequência Absoluta	Frequência Relativa
{1}		
{2}		
{3}		
{4}		
{5}		
{6}		

Se o número de provas for 250 ou 300, verás que as frequências relativas estarão próximas de  $\frac{1}{6} = 16,666\dots\%$  que é o resultado esperado a priori visto que são 6 acontecimentos com a mesma probabilidade de sair.

No entanto, há desvios bastante sensíveis.

#### **Lei dos Grandes Números (Lei de Bernoulli) – Definição frequencista de probabilidade.**

Quando o número de provas aumenta muito, frequências relativa de um acontecimento tende a estabilizar num determinado valor que se adopta como probabilidade teórica desse acontecimento

Podemos dizer que:

A probabilidade de um acontecimento associado a certa experiência é a frequência relativa esperada desse acontecimento quando o número de provas for muito elevado.

A probabilidade de um acontecimento é um número de 0 a 1 (em percentagem: de 0% a 100%)

$$0 \leq P(A) \leq 1$$

**Probabilidade = 1:** o acontecimento é certo nessa experiência que se realize sempre.

Exemplo: «sair menos de sete» = {1, 2, 3, 4, 5, 6}, no lançamento de um dado

**Probabilidade = 0:** o acontecimento nunca se dá nesta experiência, é um acontecimento impossível.

Exemplo: «sair zero» no lançamento de um dado.

Observa os resultados obtidos com um dado de madeira defeituoso, em 5000 provas.

Acontecimento	Frequência relativa
{1}	16%
{2}	9%
{3}	18%
{4}	14%
{5}	24%
{6}	19%

Podemos atribuir aos acontecimentos elementares as probabilidades da tabela.

Por exemplo:  $P(5) = 0,24$

Probabilidade de outros acontecimentos:

$A_1 = \text{«Sair múltiplo de 3»} = \{3\} \cup \{6\}$ .  $P(A_1) = 0,18 + 0,19 = 0,37$

$A_2 = \text{«Sair par»} = \{2\} \cup \{4\} \cup \{6\}$ .  $P(A_2) = 0,09 + 0,14 + 0,19 = 0,42$

$A_3 = \text{«Sair menos de 3»} = \{1\} \cup \{2\}$ .  $P(A_3) = 0,16 + 0,09 = 0,25$

Vemos que a probabilidade de um acontecimento é a soma das probabilidades dos acontecimentos elementares nele contidos.

Qual será agora a probabilidade de se dar  $A_1$  ou  $A_2$ ?

Sabendo que a disjunção “ou” corresponde a união de conjuntos teremos:

$P(A_1 \text{ ou } A_3) = P(A_1 \cup A_3) = P(\text{sair } 3, 6, 1, 2) = P(A_1) + P(A_3)$ ; porque os acontecimentos  $A_1$  e  $A_3$  são **incompatíveis**, isto é,  $A_1 \cap A_3 = \emptyset$  (não podem realizar – se simultaneamente).

Já não podemos dizer o mesmo de  $A_1$  e  $A_2$  porque a intersecção não é vazia.

$A_1 \cap A_2 = \{6\}$ , logo  $P(A_1 \cup A_2) \neq P(A_1) + P(A_2)$

$P(A_1 \cup A_2) = P(\text{sair } 3, 6, 2, 4) = 0,60$ ; enquanto  $P(A_1) + P(A_2) = 0,79$ .

Estes acontecimentos são compatíveis.

Um teorema importante ... o teorema da soma

Sejam dois acontecimentos A e B resultantes de uma mesma experiência aleatória. Então

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Se  $A \cap B = \emptyset$ , então  $P(A \cup B) = P(A) + P(B)$

No caso em que  $A \cap B = \emptyset$ , diz-se que os acontecimentos são incompatíveis

### **Acontecimentos contrários:**

Seja A = «sair menos 4» e B = «sair mais do que 5».

Temos  $A \cap B = \emptyset$

Seja A = «sair número par» e B = «sair número impar».

Vemos também que  $A \cap B = \emptyset$ ; mas este último exemplo tem uma diferença importante em relação aos anteriores. Se não se realiza A, então ocorre B (se não sai par, sai impar) o que significa que  $A \cup B$  engloba todos resultados possíveis desta experiência. Então os acontecimentos além de incompatíveis, dizem – se contrários.

Os acontecimentos A e B dizem – se contrários se a sua conjugação é impossível (não incompatíveis) e, simultaneamente a sua disjunção é certa.

$$A \cap B = \emptyset \wedge A \cup B = E$$

O acontecimento contrário de A designa – se por  $\bar{A}$  e é complementar de A.

### Definição Axiomática das Probabilidades

Acabamos de ver atrás intuitivamente, como a lógica dos conjuntos e noção da frequência relativa intervêm na construção do conceito de «**Probabilidade de um Acontecimento**». Foi com base nesses acontecimentos que **Kolmogrov** concebeu a primeira construção axiomática geral para a teoria das probabilidades. Em 1950, esta obra foi dada a conhecer a todo o Mundo científico pela tradução inglesa «**Foundations of the theory of Probability**» e a partir de então a teoria das probabilidades não parou de ganhar importância.

Como já dissemos, o espaço amostral **E** associado a uma experiência aleatória é o conjunto de todos os resultados possíveis dessa experiência. O espaço amostral pode ser discreto (por exemplo, caso da lotaria ou dado) ou contínuo (por exemplo, caso do tempo de esperar por um autocarro); no caso de ser discreto pode ser finito (lotaria ou dado) ou ter um número infinito de elementos (escolha de um número natural qualquer).

Lidaremos apenas com espaços discretos finitos.

Acontecimento é todo o subconjunto de E (incluindo o vazio  $\emptyset$  e o próprio E).

Espaço de acontecimento é o conjunto de todas as partes ou subconjuntos de E. Designa-se por  $\varphi(E)$ .

### Definição axiomática das Probabilidades

A cada acontecimento  $A$  de espaço de acontecimentos faz-se corresponder um número real que se chama probabilidade de  $A$ , escreve-se  $P(A)$ , o qual cumpre os seguintes **axiomas**:

$$A_1: P(A) \geq 0 \quad \forall A \in \mathcal{F}(E)$$

$$A_2: P(E) = 1 \text{ (acontecimento certo)}$$

$$A_3: \text{Se } A \cap B = \emptyset \text{ (A, B incompatíveis) então } P(A \cup B) = P(A) + P(B)$$

Os axiomas  $A_1$  e  $A_3$  fazem com que a probabilidade seja uma «**mediana**». O axioma  $A_2$  fixa a «**quantidade**» total da probabilidade em **1**, a semelhança das frequências relativas. Assim, podemos dizer que a probabilidade mede até que ponto se pode esperar que ocorra um acontecimento e que a medida máxima é **1**.

### Primeiras consequências dos axiomas

#### (propriedades)

$T_1$ : A probabilidade do acontecimento contrário a  $A$  é dada por  $P(\bar{A}) = 1 - P(A)$

Corolário 1:  $P(\emptyset) = 0$

A probabilidade de um acontecimento impossível é zero.

Corolário 2:  $P(A) \leq 1$ .

$T_2$ : Quaisquer que sejam os acontecimentos  $A, B, C, \dots$  Incompatíveis dois a dois tem-se  $P(A \cup B \cup C \cup \dots) = P(A) + P(B) + P(C) + \dots$

---

### Exercícios

**58.** Quando lançamos uma moeda sobre uma superfície plana, qual é o conjunto de todos os casos possíveis?

**59.** Se em 1000 lançamentos de um dado se obtêm 400 vezes a face 6, o dado deve ser viciado. Porque?

**60.** Uma urna contém várias bolas sendo umas vermelhas, outras amarelas e as restantes brancas. Sabe-se que a probabilidade de tirar bola vermelha é  $\frac{1}{4}$ .

- a) Qual é a probabilidade de tirar amarelas ou brancas?
- b) Qual é o espaço amostral?
- c) Sendo V, A, B os acontecimentos elementares e sendo  $P(B) = 2 P(A)$ , complete o quadro:

Acontecimentos	V	A	B
Probabilidade			

- d) Há acontecimentos equiprováveis; dá mais do que um exemplo.
- e) Calcula  $P(\bar{V})$ ,  $P(\bar{B})$  e  $P(\bar{A} \text{ ou } \bar{B})$ .

**61.** Supondo que para certo dado, a probabilidade de observar cada face é  $\frac{1}{6}$ .

- a) Qual a probabilidade de sair 4 ou 5 ou 6?
- b) De sair menos de 1.
- c) Lançando 200 vezes este dado é legítimo que o 6 saís quantas vezes?

**62.** Num mau dado de póquer, (dado de 6 faces) a probabilidade de sair rei é  $\frac{1}{5}$  e de sair às é de  $\frac{1}{4}$ .

Qual a probabilidade de num lançamento:

- a) Sair rei ou às?
- b) Sair rei e às?
- c) Não sair rei nem às?

**63.** Num dado viciado a probabilidade de obter a face 6 é dupla de qualquer das outras. Calcule a probabilidade da saída de cada uma das faces.

---

## Tema: Acontecimentos equiprováveis

### Introdução

Nesta unidade vamos falar da definição axiomática da probabilidade, a definição de **LAPLACE** que trabalha com casos favoráveis e casos possíveis.

Ao completar esta unidade, você será capaz de:

- Conhecer os casos de acontecimentos equiprováveis.
- Ser capaz de calcular a probabilidade usando a definição de **LAPLACE**.



### Objectivos

---

#### Acontecimentos equiprováveis

É bom lembrarmos que na maior parte das experiências aleatórias, a única forma de avaliar a probabilidade de cada acontecimento elementar consiste em repetir a experiência muitas vezes em condições análogas: por exemplo, experiências com objectos irregulares, acidentes na estrada, eleições e as experiências científicas com adubos e medicamentos. Só uma experimentação prolongada permite achar valores aproximados das suas probabilidades.

Mas pensemos agora nas experiências em que é possível calcular as probabilidades a priori, mesmo antes de realizar quaisquer provas. Em geral estão ligadas a jogos de azar que usam objectos regulares construídos com grande perfeição como dados, roletas, urnas com bolas, moedas, baralho de cartas, ... nestes casos a experimentação reiterada, limita-se a confirmar os cálculos feitos a priori.

Por exemplo no dado perfeito.

Se é perfeito,  $P(1) = P(2) = P(3) = P(4) = P(5) = P(6)$ , isto é, os acontecimentos elementares são equiprováveis.

Como são 6, a probabilidade de cada um deles é  $1/6$ .

Qualquer outro acontecimento é reunião de acontecimentos elementares equiprováveis.

### Definição clássicas de probabilidade

#### (ou lei de LAPLACE)

Se uma experiência aleatória tem  $n$  resultados possíveis incompatíveis e equiprováveis, a probabilidade dum acontecimento  $A$  é dada por:

$$P(A) = \frac{\text{n}^\circ \text{ de casos favoráveis a } A}{\text{n}^\circ \text{ de casos possíveis}} = \frac{\#A}{n} = \frac{\#A}{\#E}$$

Sendo os «casos favoráveis a  $A$ » os elementos de  $A$ .

( $\#A$  designa o número de elementos de  $A$ ).

#### Exemplo 31

Dum baralho de 52 cartas bem baralhados, extraiu-se uma carta ao acaso. Qual a probabilidade:

- a) de que seja figura?
- b) de que seja de espada?

#### Resolução

- a) Se o baralho estava correctamente baralhado e a extracção é feita «à sorte» temos 52 casos possíveis equiprováveis (e incompatíveis).

$\#E = \text{N}^\circ \text{ de casos possíveis} = 52$ .

Então a probabilidade de tirar uma dada carta é  $P(A)=1/52$ .

O acontecimento da a) é «saída de figura» e temos 12 figuras no baralho logo número de casos favoráveis = 12. Seja B: na extracção obter figura

A probabilidade de obter figura é  $P(B) = \frac{12}{52}$ .

b) Cada naipe tem 13 cartas. Seja C: a carta extraída é espada. Logo a probabilidade de tirar espadas é  $P(C) = \frac{13}{52}$ , pois existem 13 cartas de espada (coração preto).

### Exemplo 32.

Uma urna contém 5 bolas brancas e 3 bolas azuis. Retiramos ao acaso duas bolas. Qual a probabilidade de que as duas sejam brancas?

### Resolução

Calculando o numero de casos possíveis (escolha de 2 bolas das 8 disponíveis):

$$C_2^8 = \frac{8!}{2!(8-2)!} = \frac{8!}{2!6!} = 28$$

- Calculo do número de casos favoráveis (número de possibilidades de escolher 2 bolas brancas das 5 disponíveis):

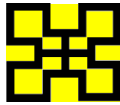
$$C_2^5 = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = 10$$

Então, se D representa o acontecimento “as duas escolhidas são brancas”, temos:

$$P(D) = \frac{10}{28} = \frac{5}{14}$$

---

## Exercícios



Auto-avaliação

**64.** Uma urna contém 20 bolas vermelhas e 30 bolas brancas.

Extrai-se uma ao acaso. Qual a probabilidade de que saia:

- a) Uma bola vermelha?
- b) Uma bola branca?
- c) Uma bola vermelha ou branca?

**65.** De um baralho de 52 cartas, retira-se uma delas. Determine a probabilidade dos seguintes eventos:

- a) a carta é um 5 de copas;
- b) a carta é um 5 qualquer.

**66.** Uma urna contém 10 bolas pretas, 11 bolas azuis e 12 bolas brancas. Retirando-se uma bola ao acaso, qual a probabilidade dela ser azul?

**67.** Uma urna contém 20 cartões numerados de 1 a 20. Um cartão é retirado ao acaso. Determine a probabilidade dos seguintes eventos:

- a) O cartão tem o número 11;
- b) O cartão tem um número maior que 15;
- c) O cartão tem um número múltiplo de 3;
- d) O cartão não tem o número 13.

**68.** Uma urna contém uma ficha para cada anagrama da palavra CARRO. Escolhendo uma ficha ao acaso, determine a probabilidade de que:

- a) O anagrama comece com C.
- b) O anagrama comece com R e termine com C.

**68.** De um baralho de 52 cartas, tiram-se 2 delas sem reposição. Determine a probabilidade dos seguintes eventos:

- a) As duas cartas serem “reis”;
- b) As duas cartas serem de “espada”;
- c) Termos um “rei” e um “oito”.

**70.** No lançamento simultâneo de 2 dados, determine a probabilidade de se obter:

- a) número nos dois lados;
- b) pontos iguais nos dois dados.

**71.** Seis casais estão em uma festa. Escolhendo duas pessoas ao acaso, determine a probabilidade de termos:

- a) um homem e uma mulher;
- b) um marido e sua esposa;
- c) duas mulheres;
- d) o casal Silva (um dos 6 casais presentes).

**72.** Uma urna contém 10 bolas: 5 azuis, 3 brancas e 2 pretas. Escolhendo-se 3 bolas ao acaso e sem reposição, determine a probabilidade de que:

- a) as três sejam brancas;
- b) apenas uma seja branca;
- c) pelo menos uma seja preta.

**73.** Numa caixa existem 6 parafusos e 6 porcas. Retira-se aleatoriamente 3 peças.

- a) qual é a probabilidade de que as três sejam porcas?
- b) qual é a probabilidade de termos retirado 3 peças iguais?

---

## Tema: Probabilidade Condicional

### Introdução

Muitos casos da probabilidade condicionada são vividos frequentemente na sociedade.

Nesta unidade vamos falar das possibilidades do cálculo.

Ao completar esta unidade, você será capaz de:

- Conhecer as regras do cálculo da probabilidade condicional e até que ponto uma informação adicional pode afectar o cálculo da probabilidade de um acontecimento
- Ser capaz de calcular a probabilidade da intersecção de acontecimentos.



### Objectivos

---

#### Probabilidade Condicionada

#### Acontecimentos independentes

#### *Probabilidade Condicional*

Parece razoável defender que a probabilidade de uma pessoa, escolhida ao acaso, gostar de Matemática, não é deve ser a mesma se a escolha for feita num grupo de Licenciatura em Ensino de Matemática ou num grupo de Licenciatura em ensino de Filosofia.

De que maneira uma informação adicional afecta a probabilidade de um acontecimento, é o que vamos tratar.

### Questão 1:

Sabemos que no lançamento de um dado perfeito a probabilidade de sair cada uma das faces é  $1/6$ .

Supõe, porém, que o dado foi lançado e que a face voltada para cima é um número par. Qual a probabilidade que a tal face é divisível por 3?

### Solução

No lançamento de um dado perfeito de 6 faces o espaço de resultados é

$E = \{1, 2, 3, 4, 5, 6\}$  e atribuímos a probabilidade  $1/6$  a cada elemento de  $E$ .

Consideremos os acontecimentos:

$A$ : “divisível por 3”  $\Rightarrow A = \{3, 6\}$ ;  $B$ : “par”  $\Rightarrow B = \{2, 4, 6\}$

O que pretendemos?

Pretendemos de que um elemento esteja em  $A$ , sabendo que pertence a  $B$ .

Dado que estamos unicamente interessados em resultados nos quais o número é par, podemos pôr de lado os resultados 1, 3, 5 resultando desse modo um novo espaço de resultados, que é  $B = \{2, 4, 6\}$ . Deste conjunto apenas  $\{6\}$  satisfaz a condição dada – ser divisível por 3.

Se todos os resultados de  $B$  são equiprováveis, então  $P(\{2\}) = P(\{4\}) = P(\{6\}) = 1/3$ . Logo, a probabilidade que se procura é  $1/3$ , isto é,

$P(A \text{ ocorrer dado que } B \text{ ocorreu}) = 1/3$ . Simbolicamente isto se escreve:

$$P(A/B) = 1/3.$$

Para resolvermos o problema, muito simplesmente mudamos o espaço de resultados de  $E$  para  $B$  e procedemos a uma nova atribuição de probabilidades.

Temos a liberdade de fazer isto de qualquer maneira consistente com a definição de medidas de probabilidades, mas para B não temos escolha a não ser  $P(B) = 1$ , pois B é o novo espaço amostral “reduzido”.

Sistematizemos como surgiu o resultado de  $P(A/B) = 1/3$

$$\left. \begin{array}{l} \#(A \cap B) = 1 \\ \#B = 3 \\ P(A/B) \end{array} \right\} \Rightarrow P(A/B) = \frac{\#(A \cap B)}{\#B}, \#B \neq 0 - \text{regra (1) para o} \\ \text{cálculo da probabilidade} \qquad \qquad \qquad \text{condicional}$$

Se de novo reparamos no espaço de resultados  $E = \{1, 2, 3, 4, 5, 6\}$ , vemos que

$$P(A \cap B) = 1/6; P(B) = 3/6$$

Para obtermos o resultado do problema proposto, basta considerar que:

$$P(A/B) = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{6} \times \frac{6}{3} = \frac{1}{3}$$

Portanto,  $P(A/B) = \frac{P(A \cap B)}{P(B)}$ ,  $P(B) \neq 0$ , já que B considera-se que ocorreu.

Suponhamos que os 40 alunos de uma turma da 11ª classe foram agrupados atendendo à cor do cabelo (louro e preto) e à cor dos olhos (azuis e castanhos), tendo-se obtido os seguintes resultados:

$$\text{Olhos azuis} \left\{ \begin{array}{l} \text{cabelos louros - 5} \\ \text{cabelos pretos - 2} \end{array} \right.$$

$$\text{Olhos castanhos} \left\{ \begin{array}{l} \text{cabelos louros - 10} \\ \text{cabelos pretos - 23} \end{array} \right.$$

Consideremos agora os seguintes acontecimentos:

$$A = \{ \text{ter cabelo louro} \}$$

$$\bar{A} = \{ \text{ter cabelo preto} \}$$

$$B = \{\text{ter olhos azuis}\}$$

$$A \cap B = \{\text{ter cabelo louro e olhos azuis}\}$$

$$B/A = \{\text{ter olhos azuis entre os que tem cabelo louro}\}$$

As probabilidades para cada um destes acontecimentos se verificar é:

$$P(A) = \frac{15}{40}; \quad P(B) = \frac{7}{40}; \quad P(\bar{A}) = \frac{25}{40};$$

$$P(\bar{B}) = \frac{33}{40}; \quad P(A \cap B) = \frac{5}{40}; \quad P(B/A) = \frac{5}{15}$$

(Uma vez que dos alunos de cabelos louros apenas 5 têm olhos azuis).

Comparando os resultados obtidos conclui-se:

$$P(B/A) = \frac{5}{15} = \frac{\frac{5}{40}}{\frac{15}{40}} = \frac{P(A \cap B)}{P(A)}$$

Observação:  $P(B/A)$  diz-se probabilidade de B se A ou probabilidade de B se realizar dado que A se realizou.

De um forma geral:

Sendo A e B dois acontecimentos relativos a uma experiência aleatória e  $P(A) \neq 0$ , chama-se probabilidade condicional de B, conhecendo A, ou mais simplesmente, probabilidade de B se A, ao quociente:  $P(B/A) = \frac{P(A \cap B)}{P(A)}$

Observação: De igual modo, se  $P(B) \neq 0$ ,  $P(A/B) = \frac{P(A \cap B)}{P(B)}$

Das relações anteriores obtém-se as formulas:

$$P(A \cap B) = P(B/A) \cdot P(A)$$

$$P(A \cap B) = P(A/B) \cdot P(B)$$

Conhecidas por leis da multiplicação.

A probabilidade do acontecimento intersecção ( $A \cap B$ ) é igual à probabilidade de um deles, suposta não nula, pela probabilidade do outro condicionada realização do anterior.

### Exemplo 33

Extraem-se sucessivamente, duas cartas de um baralho de 52 cartas.

Qual a probabilidade de obter dois valetes?

### Resolução

Usando a notação:

A – o acontecimento “sair valete” na primeira extracção.

B – o acontecimento “sair valete” na segunda extracção.

Temos:

$$P(A \cap B) = P(B/A) \cdot P(A)$$

$$P(A \cap B) = \frac{4}{52} \cdot \frac{3}{51} = \frac{4 \cdot 3}{52 \cdot 51} = \frac{1}{221} = 0,00452$$

Observação:  $P(B/A) = \frac{3}{51}$  porque uma vez extraído um valete apenas ficam 3 num conjunto de 51 cartas.

Se a ocorrência do acontecimento A não afecta a ocorrência do acontecimento B e vice-versa, diz-se que os

acontecimentos A e B são independentes, o que significa que:

$$P(B/A) = P(B) \text{ e } P(A/B) = P(A)$$

$$\text{Então } P(A \cap B) = P(A) \times P(B)$$

Ou generalizando:

Se os acontecimentos  $A_1, A_2, \dots, A_n$  são independentes teremos:

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2) \cdot \dots \cdot P(A_n)$$

### Exemplo 34

Consideremos uma urna com 15 bolas equiprováveis, sendo 7 brancas e 8 pretas.

Tira-se uma bola e, antes de tirar nova bola, repõe-se na urna a bola saída.

Qual a probabilidade de sair duas vezes bola branca em duas extracções sucessivas?

### Resoluções

Sejam os acontecimentos, A: “a primeira bola é branca” e B: “a segunda bola é branca”.

Dado que a extracção é com reposição, a probabilidade de extrair bola branca na primeira tentativa, não afecta a probabilidade de extrair bola branca na segunda tentativa. Então os acontecimentos A e B são independentes, logo, a probabilidade pedida será:

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P(A \cap B) = \frac{7}{15} \cdot \frac{7}{15} = 0,217$$

### Exemplo 35

Em 100 congressistas, 80 falam Francês e 40 falam Inglês. Qual a probabilidade de 2 congressistas, escolhidos ao acaso, se entenderem numa daquelas línguas?

### Resolução

Escolhido o primeiro congressista, ele pode falar em Francês ou Inglês.

Suponhamos que fala Francês.

A probabilidade do 1º congressista falar Francês é  $\frac{80}{100} = \frac{8}{10}$

A probabilidade do 2º congressista também falar Francês é  $\frac{79}{99}$

Então:

A probabilidade de 2 congressistas, escolhidos ao acaso, falarem Francês será:

$$\frac{8}{10} \cdot \frac{79}{99} = 0,639$$

A probabilidade de dois congressistas, escolhidos ao acaso, falarem Inglês será:

$$\frac{40}{100} \cdot \frac{39}{99} = \frac{4}{10} \cdot \frac{39}{99} = 0,157$$

Seja A: "os dois congressistas falam a mesma língua".

Então a probabilidade pedida é:

$$P(A) = 0,639 + 0,157 = 0,796$$

### Exemplo 36

No lançamento de dois dados, determinar a probabilidade que a soma dos resultados dos dois seja 5, supondo que no primeiro, o resultado tenha sido menor que 3.

#### Resolução

Vamos considerar:

Evento A: “a soma dos resultados dos dois dados igual a 5”

Evento E: “o resultado do primeiro dado menor que 3”

$$\text{Assim: } A = \{(1,4);(2,3);(3,2);(4,1)\}$$

$$E = \{(1,1);(1,2);(1,3);(1,4);(1,5);(1,6);(2,1);(2,2);(2,3);(2,4);(2,5);(2,6)\}.$$

$$A \cap E = \{(1,4);(2,3)\}$$

$$\text{Então: } P(A/E) = \frac{\#(A \cap E)}{\#E} = \frac{2}{12} = \frac{1}{6}$$

### Exemplo 37

De um baralho são retiradas, sucessivamente, duas cartas, sem reposição. Qual a probabilidade de que as duas cartas retiradas sejam damas?

#### Resolução

Vamos considerar:

- Evento A: obter dama na primeira carta.
- Evento B: obter dama na segunda carta.

Então, estamos procurando  $P(A \cap B)$ .  
 $P(A \cap B) = P(A) \cdot P(B/A)$ .

- Para a obtenção da primeira carta, existem 4 damas, num total de 52 cartas, então:  $P(A) = \frac{4}{52}$

- Para a obtenção da segunda carta, considerando que já saiu uma dama a primeira retirada, temos 3 damas num total de 51 cartas, então:  $P(B/A) = \frac{3}{51}$

$$\text{Assim, } P(A \cap B) = \frac{4}{51} \cdot \frac{3}{51} = \frac{1}{221}$$

### Exemplo 38

Uma moeda é lançada 3 vezes. Qual é a probabilidade de que apareça coroa nas três vezes?

#### Resolução

Espaço amostral:  $S = \{\text{cara, coroa}\}$

Sejam os acontecimentos:

A: saiu coroa no primeiro lançamento

B: saiu coroa no segundo lançamento

C: saiu coroa no terceiro lançamento.

Então:

- Probabilidade de sair coroa no 1º lançamento  
 $\rightarrow P(A) = \frac{1}{2}$
- Probabilidade de sair coroa no 2º lançamento  $\rightarrow P(B) = \frac{1}{2}$
- Probabilidade de sair coroa no 3º lançamento  $\rightarrow P(C) = \frac{1}{2}$

$$\text{Assim, } P(A \cap B \cap C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

### Exemplo 39

Uma urna contém bolas numeradas de 1 a 15. Escolhendo-se ao acaso uma bola, qual a probabilidade dela ser múltiplo de 2 ou múltiplo de 3?

### Resolução

Vamos considerar:

- $S = \{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15\}$
- Evento A: “ocorre múltiplo de 2”.

$$A = \{2,4,6,8,10,12,14,\}$$

- Evento B: “ocorre múltiplo de 3”.

$$B = \{3,6,9,12,15\}$$

Assim,  $A \cap B = \{6,12\}$ . Então, a probabilidade de que o número seja múltiplo de 2 ou múltiplo de 3,  $P(A \cup B)$

$$\text{é } P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{7}{15} + \frac{5}{15} - \frac{2}{15} = \frac{10}{15} = \frac{2}{3}$$

### Exemplo 40

Um lote de peças contém 15 peças boas, 4 com defeitos menores e 2 com defeitos graves.

Escolhendo uma peça ao acaso, ache a probabilidade de que:

- a) a peça seja boa;
- b) a peça não seja boa.

### Resolução

Se A é o evento “retirar uma peça boa”, então  $\bar{A}$  é o evento “não retirar uma peça boa”.

$$\text{a) } P(A) = \frac{15}{21} = \frac{5}{7}$$

$$\text{b) } P(A) + P(\bar{A}) = 1 \Rightarrow \frac{5}{7} + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = 1 - \frac{5}{7} = \frac{2}{7}$$

### Exemplo 41

Dois prêmios iguais serão sorteados entre 8 pessoas, sendo 3 homens e 5 mulheres. Sabendo que os dois prêmios não

podem ser ganhos pela mesma pessoa, qual é a probabilidade de ser premiada pelo menos uma pessoa?

### Resolução

Observe que, se  $P(A)$  é a probabilidade de que 2 homens sejam premiados, então  $P(\bar{A})$  é a probabilidade de que pelo uma mulher seja premiada.

- Calculando  $P(A)$ :

$$P(A) = \frac{3}{8} \cdot \frac{2}{7} = \frac{2}{28}$$

$$\text{Calculando } P(\bar{A}): P(A) + P(\bar{A}) = 1 \Rightarrow \frac{3}{28} + P(\bar{A}) = 1 \Rightarrow P(\bar{A}) = \frac{25}{28}$$

---

### Exercícios

**74.** Em uma bola há 20 fichas numeradas de 1 a 20.

Tira-se uma ficha que em seguida é recolocada na bolsa, logo após retira-se a outra ficha.

Qual é a probabilidade de se tirar duas vezes a mesma ficha?

**75.** Um grupo de 50 estudantes universitários é formado por 30 estudantes de Direito, 15 estudantes de Psicologia e 5 estudantes de Matemática. Escolhendo um desses estudantes ao acaso, qual é a probabilidade de que ele estude Psicologia ou Matemática?

**76.** Uma moeda é lançada, sucessivamente, três vezes.

Qual é a probabilidade de ocorrer pelo menos uma coroa, sabendo que saiu cara na primeira jogada?

**77.** Uma moeda é lançada quatro vezes seguidas.

Qual é a probabilidade de conseguirmos pelo menos uma cara, sabendo que no primeiro lançamento foi obtido cara?

**78.** Determine a probabilidade de um casal ter 3 filhos, todos do sexo feminino.

**79.** No lançamento de dois dados qual é a probabilidade da soma ser 8, supondo que saiu no número 5 no primeiro dado?

**80.** A probabilidade de um casal de namorados passear no sábado é 0,9 e quando saem, a probabilidade de jantar fora é 0,6. Qual é a probabilidade desse casal sair para jantar no sábado?

**81.** Em uma corrida de automóveis, as chances de um piloto ganhar são de 72%. Qual é a probabilidade dele perder a corrida?

**82.** Duas pessoas (A e B) vão resolver um exercício. As probabilidades de que consigam resolvê-lo são:

$P(A) = \frac{1}{3}$  e  $P(B) = \frac{2}{5}$ . Determine a probabilidade de que:

- a) Ambos resolvam o exercício;
- b) **A** resolva e **B** não consiga resolver;
- c) Ambos não resolvam o exercício;
- d) Pelo menos um deles resolva.

**83.** Uma urna contém 3 bolas: uma verde, uma azul e uma branca. Tira-se uma bola ao acaso, regista-se a cor e coloca-se a bola de volta à urna. Repete-se esta experiência mais duas vezes. Qual a probabilidade de serem registadas três cores distintas?

**84.** No problema anterior: Qual a probabilidade de obtermos três bolas da mesma cor?

**T.P.C.**

- 1.** Ler e estudar todas as páginas do referido texto.
- 2.** Resolver todos os exercícios.
- 3.** Entregar os exercícios números: 65, 67, 69, 72, 75, 78.

---

## BIBLIOGRÁFICA

1. Neves, M. (2008). *Introdução à Estatística e à Probabilidade*. Apontamentos de apoio às aulas [[Estatística Descritiva](#)] [[Teoria da Probabilidade](#)] [[Introdução à Inferência Estatística](#)] ---- *Colectânea de exames de Estatística do ISA com algumas resoluções* . AEISA.
2. BUSSAB, Wilton O.; Pedro A. Morettin (2006) - *Estatística Básica*, 5ª Ed., Saraiva, São Paulo
3. LARSON, Ron; Betsy Farber (2004) - *Estatística Aplicada*, 2ª Ed., Prentice Hall, São Paulo
4. SILVA, Sebastião Medeiros da, Et Al (1999) - *Estatística para os cursos de economia, administração e ciências contábeis*, Atlas, São Paulo
5. BEKAMAN, Otto Ruprecht; Pedro Luiz de Oliveira Costa Neto (1980) - *Análise Estatística da decisão*, Edgard Blucher, São Paulo