



## Conteúdos

- Estatística - Generalidades*
- Organização e interpretação de caracteres estatísticos (qualitativos e quantitativos)*
- Medidas de tendência central e medidas de dispersão*
- Distribuições bidimensionais (abordagem gráfica e intuitiva)*
- Reta de regressão*
- Variáveis discretas e contínuas*
- Função massa de probabilidade*
- Distribuição de probabilidades*
- Modelo Binomial*
- Modelo de Poisson*
- Modelo Normal*

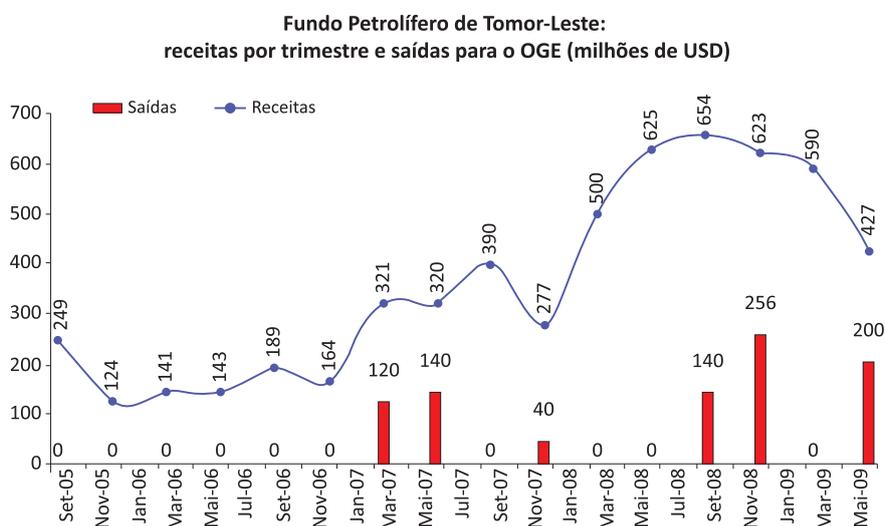


## Subtema 2 - Estatística descritiva e indutiva

### Introdução

A recolha de dados, o seu tratamento e análise assim como a interpretação dos mesmos fazem parte de qualquer estudo estatístico. Todos os dias são publicadas notícias baseadas em estudos estatísticos de maior ou menor relevância, as quais pretendem transmitir informações em diferentes domínios da atualidade.

Existem entidades que se dedicam à realização de diferentes estudos estatísticos.



Nota: além das saídas para o OGE há saídas para pagamento da comissão de gestão pela ABP e variações do valor de mercado dos títulos em carteira, as quais podem ser negativas

Fonte: relatórios trimestrais do Fundo Petrolífero publicados pela ABP

### Recenseamento e Sondagem

A direção Nacional de Estatística (DNE) tem a seu cargo efetuar censos da população Timorense. O recenseamento permite conhecer diferentes características da população timorense, tais como situação civil, habitacional, agregados familiares, rendimentos e muitos outros aspetos. Este conhecimento permite aos governantes tomar decisões relativas à saúde, educação, habitação etc.

A população sobre a qual realizamos um estudo estatístico não é necessariamente a população de um país, podemos realizar estudos sobre parte de uma população ou ainda sobre instituições ou objetos.

Consideramos **População**, ou **Universo**, a uma coleção de objetos com uma característica comum.

A cada elemento da população damos o nome de **Indivíduo** ou **unidade estatística**.

Um **censo ou recenseamento** é um estudo estatístico que implica a observação de **todos** os elementos da população.

A realização de censos é por vezes, muito dispendiosa pelo facto de terem que ser recolhidos dados de todos os indivíduos da população.

Também não é viável realizar um censo quando o estudo em causa implica a destruição de todas as unidades, por exemplo, se pretendemos analisar a resistência a um choque frontal de determinado modelo automóvel a realização de um censo é impossível, a fábrica iria a falência destruiria toda a sua produção no estudo de qualidade.

Por este motivo, muitos estudos estatísticos são realizados observando só uma parte da população. Um estudo estatístico deste tipo é uma sondagem.

Uma **amostra** é um subconjunto de elementos extraídos da população em estudo utilizando uma metodologia adequada.

Uma **sondagem** é um estudo estatístico de uma população a partir de uma amostra.

A fiabilidade de uma sondagem depende das metodologias usadas para construir a amostra. A este procedimento de seleção chama-se amostragem.

As amostragens mais usuais são:

#### **Amostragem aleatória simples:**

Cada indivíduo da população tem a mesma probabilidade de ser selecionado.

#### **Amostragem estratificada**

A população está dividida em diferentes estratos, devendo a proporção de indivíduos escolhidos em cada estrato ser aproximadamente igual a proporção do estrato no total da população.

#### **Amostragem sistemática**

Os elementos da população são escolhidos a partir de uma regra previamente definida.

#### **Referência Histórica**

A estatística e os métodos a ela associados sofreram um notável desenvolvimento durante o século XX devido aos trabalhos de matemáticos como Pearson, Fisher, Gosset e Kolmogorov, entre outros, sobre tudo pelo notável desenvolvimento da tecnologia. Mas os métodos estatísticos sempre foram usados.

Nas civilizações antigas, como a chinesa, a egípcia, a assíria e a grega eram feitos periodicamente inquéritos sobre a produção anual de produtos essenciais como o trigo.

Contagem de populações, estabelecimentos de impostos e recolha de dados para fins militares foram algumas das primeiras aplicações dos métodos estatísticos.

### Tarefa 36

Indica, justificando, das situações seguintes aquelas em que é mais conveniente estudar toda a população ou apenas uma amostra:

- A duração de uma marca de pilhas produzidas por uma fábrica.
- A altura de todos os turistas estrangeiros durante o ano 2011 em Timor-Leste.
- A altura de um grupo de oito amigos.
- A disciplina preferida dos alunos de uma turma.

### Tarefa 37

Numa escola existem 800 alunos, dos quais 610 são raparigas. Pretende-se escolher dessa população 20 alunos para integrar uma amostra. Se a amostra for estratificada quantas raparigas devem ser seleccionadas?

## Estatística Descritiva e Estatística Indutiva

Depois de recolhidos os dados, seja um censo ou uma sondagem, procede-se ao tratamento dos mesmos observando o valor mais frequente, a média dos valores obtidos e outros aspetos que descrevem as características da população em estudo.

A Estatística Descritiva tem por finalidade descrever certas propriedades relativas a uma amostra ou população.

Mas quando realizamos uma sondagem para além de organizar os dados e descrever as características da amostra, inferimos propriedades para toda a população.

A Estatística Indutiva procura inferir características da população a partir das características observadas na amostra.

## Atributos Estatísticos

As características comuns de uma população, em determinado estudo estatístico, que podem tomar valores ou modalidades diferentes de indivíduo para indivíduo chama-se atributo ou carácter estatístico.

Podemos analisar diferentes tipos de atributos.

Um **atributo qualitativo** é uma qualidade, uma característica ou uma categoria, que pode assumir várias modalidades que não são possíveis de medir.

Um **atributo quantitativo** é uma característica que se pode medir, ou seja que se pode exprimir por um número.

Muitas vezes usamos o termo “**variável estatística**” para designar um atributo quantitativo.

Se uma **variável quantitativa** toma apenas **um número finito de valores**, em qualquer intervalo de variação limitado, diz-se que é uma **variável discreta**.

Se a **variável quantitativa** pode tomar **todos os valores numéricos**, compreendidos no seu intervalo de variação, diz-se que é uma **variável contínua**.

Em resumo:



## Organização de dados

Os resultados obtidos por observação de um atributo, seja quantitativo ou qualitativo, chamam-se **dados estatísticos**.

### Tabelas de frequências

Quando se recolhem dados estatísticos é essencial organizá-los de forma a facilitar a sua leitura e interpretação, por exemplo recorrendo a tabelas diagramas ou gráficos.

Uma das questões mais importantes é que a variável estatística em estudo seja definida de forma clara, bem como a população ou amostra sobre a qual incidem os dados.

### Distribuições Estatísticas

Na turma do 10º ano de uma determinada escola realizou-se um estudo sobre as idades dos alunos da turma. Os dados obtidos estão registados na seguinte tabela:

16	15	16	15	18	16	17	15
17	15	14	17	15	16	15	16

A variável estatística em estudo é a variável “Idade” sendo a população em estudo a turma do 10º.

Neste caso os valores que a variável estatística pode tomar são 14; 15; 16; 17; 18.

### Tarefa 38

Justifica porque é que as seguintes situações representam uma má amostra.

- Para saber qual o candidato, mais votado para as eleições de 1º Ministro em Timor Leste, ouviu-se a opinião dos simpatizantes de determinado partido político.
- Para conhecer a situação financeira das empresas petrolíferas em Timor Leste, verificou-se a situação das empresas que tiveram, no último ano, maior volume de exportações.
- Para avaliar a qualidade dos fósforos produzidos numa fabrica, acenderam-se todos os fósforos fabricados num dia.
- Para saber a opinião das pessoas sobre música rock, questionou-se um grupo de pessoas que saiam de um concerto de música rock.

### Tarefa 39

Das seguintes características de uma população, indica as que são qualitativas e as que são quantitativas, indicando se são discretas ou contínuas.

- a) Número de calçado.
- b) Naturalidade.
- c) Estado civil.
- d) Preferências desportivas.
- e) Número de horas a ver televisão.
- f) Idade dos empregados de uma fábrica.
- g) Duração de uma lâmpada
- h) A cor do cabelo.
- i) O consumo de gasolina por cada 100 Km.

### Tarefa 40

O género de vinte pessoas escolhidas ao acaso foi:

M F F M M M F F M  
M F M F F M M M M  
F M

Elabora a tabela de frequências absolutas associada.

Temos uma **distribuição estatística** sempre que se conhece o valor da variável estatística para cada individuo ou elemento da população ou amostra em estudo. Designamos por  $x_1, x_2, x_3, \dots$  os diferentes valores que a variável estatística  $X$  pode tomar.

- No caso da **variável estatística  $X$  ser quantitativa** e tomar apenas um número limitado de valores  $x_1, x_2, x_3, \dots, x_n$  é necessário ordenar os valores que a variável assume por ordem crescente.

$$x_1 < x_2 < x_3 < \dots < x_n$$

No exemplo anterior temos

$$x_1 = 14 \quad x_2 = 15 \quad x_3 = 16 \quad x_4 = 17 \quad x_5 = 18$$

### Frequências Absolutas

Quando se pretende organizar os dados numa tabela, na coluna da esquerda colocam-se os diferentes valores  $x_i$  **que a variável em estudo pode tomar**.

Na coluna seguinte regista-se o número de vezes que cada valor  $x_i$  da variável aparece na população ou amostra estudada.

### Exemplo:

Os alunos do 12º ano de uma escola foram inquiridos relativamente a sua preferência musical. Os resultados obtidos foram os seguintes:

Tipo de musica	Pop	Hip-Hop	Reggae	Rock	Techno	Clássica
Nº de alunos	122	86	92	70	66	22

A **frequência absoluta** de um dado estatístico é o número de vezes que este se repete e representa-se por  $f_i$ .

A tabela anterior será então representada da seguinte forma:

$x_i$	$f_i$
Pop	122
Hip-Hop	86
Reggae	92
Rock	70
Techno	66
Clássica	22
Total 458	

Pretendemos agora comparar as preferências destes alunos com os alunos do 11º ano, desta mesma Escola, é necessário fazer um inquérito aos alunos do 11º ano sobre o seu tipo de música preferida, depois de aplicado sabemos que 75 preferem Rock.

Esta informação só por si, não permite fazer comparação, não sabemos qual é o peso deste 75 alunos no total dos alunos do 11º ano.

Observa que se os alunos do 11º ano forem 150, então metade dos alunos gostariam de Rock mas se os alunos do 11º ano forem 300, então só 25% gostariam de Rock.

A razão entre a frequência absoluta e o total das observações permite uma melhor comparação entre distribuições e também permite conhecer a importância relativamente a totalidade das observações.

A frequência relativa de um dado estatístico é o quociente entre a frequência absoluta e o número total de dados e representa-se:

$$fr_i = \frac{f_i}{n}$$

A nossa tabela teria agora o seguinte aspeto:

$x_i$	$f_i$	$fr_i$
Pop	122	$\frac{122}{458} = 0,27$
Hip-Hop	86	0,19
Reggae	92	0,20
Rock	70	0,15
Techno	68	0,15
Clássica	20	0,04
Total	458	1

Observa que se pretendemos obter a percentagem de cada estilo musical é suficiente multiplicar cada frequência relativa por 100.

Por exemplo 15% dos alunos inquiridos prefere Rock.

A **frequência relativa** expressa em termos de **percentagem** se obtém ao multiplicar a frequência relativa por 100.

A informação dada desta forma permite comparar as preferências musicais dos alunos do 11º e 12º ano.

#### Tarefa 41

Ao lançar um dado obtiveram-se os seguintes pontos:

1;2;4;6;2;3;4;4;5;6;1;2;2;3

- Constrói uma tabela de frequências relativamente ao lançamento do dado.
- Qual a percentagem de vezes que saiu a face com o número 6?
- Qual é a face que saiu com mais frequência?