



**By @kakashi\_copiador**

## APRESENTAÇÃO DO MATERIAL

Queridos alunos!!

Sabemos que os **resumos** das disciplinas **são fundamentais para fixação de conteúdos** e, também, para **realização de revisões**. Um resumo bem feito garante que os principais pontos de cada matéria sejam revisados de forma rápida, **aumentando a produtividade dos estudos e a eficiência das revisões**.

Além disso, sabemos que, principalmente para os grandes concursos, o número de matérias cobradas no edital é muito grande. Dessa forma, além de revisar os pontos marcados em seus materiais, um bom resumo pode encurtar o tempo de revisão, garantindo, assim, que todo o material possa ser revisado em um período de tempo mais curto.

Com isso em mente, apresentamos a vocês o **Resumo de Estatística - Medidas Separatrizes ou Quantis**. Trata-se de um material pensado para lhe ajudar em todo esse processo, visando, inclusive, uma economia de tempo de confecção de materiais, tempo que é o bem mais precioso de um concurseiro, não é mesmo?

Esperamos poder ajudá-los!

Conte sempre com o Estratégia em sua caminhada!

**Estratégia Concursos**



*Esse é um material resumido. Em momento algum ele substitui o estudo do material completo. Trata-se de um complemento aos estudos e um facilitador de revisões!*

## RESUMO DE ESTATÍSTICA

### Medidas Separatrizes

**As separatrizes são medidas que dividem (ou separam) uma série ordenada em duas ou mais partes, cada uma contendo a mesma quantidade de elementos.**

Nesse caso, o nome da medida separatriz é definido de acordo com a quantidade de partes em que a série é dividida:

### MEDIANA

Divide uma série ordenada em duas partes iguais, cada uma contendo **50%** dos valores da sequência

### QUARTIS

Dividem uma série ordenada em quatro partes iguais, cada uma contendo **25%** dos valores da sequência.

### QUINTIS

Dividem uma série ordenada em cinco partes iguais, cada uma contendo **20%** dos valores da sequência

### DECIS

Dividem uma série ordenada em dez partes iguais, cada uma contendo **10%** dos valores da sequência

### PERCENTIS

Dividem uma série ordenada em cem partes iguais, cada uma contendo **1%** dos valores da sequência

## Mediana

A mediana é, simultaneamente, uma **MEDIDA DE POSIÇÃO, de TENDÊNCIA CENTRAL e SEPARATRIZ**. Ela separa uma série de valores em **duas** partes de **tamanhos iguais**, cada uma contendo o **mesmo número de elementos**. Sendo representada pelos símbolos ***Md*** ou, em menor frequência, ***x<sub>~</sub>***

- **Mediana para dados não-agrupados**

→ A **mediana** é o elemento que ocupa a **POSIÇÃO CENTRAL** de uma série de observações ordenadas segundo suas grandezas. Para estabelecer que a mediana de um conjunto composto por  $n$  elementos ordenados de forma crescente ou decrescente temos:

Se  $n$  for ímpar:

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

Se  $n$  for par:

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

- **Mediana para dados agrupados sem intervalos de classe**

→ A **mediana** é um valor que dividirá a distribuição de frequências em **duas partes** contendo o **mesmo número de elementos**. Para estabelecer que a mediana de uma tabela de frequências composta por um total de  $n$  elementos:

Se  $n$  for ímpar:

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

Se  $n$  for par:

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

- **Mediana para dados agrupados em classes**

→ Para calcular a mediana de dados que estão agrupados por intervalo de classes, **identificamos a classe mediana, que corresponde à frequência acumulada imediatamente igual ou superior à metade da frequência total, ou seja, metade da soma das frequências simples,  $\sum f_i/2$ .**

→ Conhecendo a classe mediana, podemos aplicar a fórmula da mediana:

$$M_d = l_{inf} + \left[ \frac{\left( \frac{\sum f_i}{2} \right) - f_{ac_{ant}}}{f_i} \right] \times h$$

→ **Método de interpolação linear** consiste em utilizar valores conhecidos para estimar valores desconhecidos de forma linear, isto é, por meio de uma reta

- **Propriedades da Mediana**

### 1ª Propriedade

- Somando-se (ou subtraindo-se) uma constantes  $c$  a todos os valores de uma variável, a mediana do conjunto fica aumentando (ou diminuída) dessa constante.

### 2ª Propriedade

- Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante  $c$ , a mediana do conjunto fica multiplicada (ou dividida) por esta constante.

### 3ª Propriedade

- A soma dos desvios absolutos de uma sequência de números, em relação a um número  $\alpha$ , é mínima quando  $\alpha$  é a mediana dos números.

## QUARTIL, DECIL E PERCENTIL

O **nome da medida separatriz** é atribuído de acordo com a **quantidade de partes em que a série é dividida**:

### QUARTIS

- Valores de uma série que a dividem em **QUATRO PARTES IGUAIS**, isto é, quatro partes contendo o mesmo número de elementos (25%).
- **Q1**: o primeiro quartil corresponde à separação dos primeiros **25%** de elementos da série.
- **Q2**: o segundo quartil corresponde à separação de metade dos elementos da série, coincidindo com a mediana (**Q2 = Md**)
- **Q3**: o terceiro quartil corresponde à separação dos primeiros **75%** de elementos da série, ou dos últimos **25%** de elementos da série.



- **Quartil para dados não-agrupados**

→ O cálculo do **quartil para dados não-agrupados** será realizado por meio de três etapas

→ 1ª etapa: determinar a posição do quartil, por meio da expressão:

$$P_{Q_k} = \frac{k \times n}{4} \quad (k = 1, 2, 3)$$

→ 2ª etapa: identificar a posição mais próxima do rol;

→ 3ª etapa: verificar o valor que está ocupando essa posição.

- **Quartil para dados agrupados sem intervalos de classe**

→ O cálculo do **quartil para dados agrupados sem intervalos de classe** será realizado por meio de três etapas:

→ 1ª etapa: determinar a posição do quartil, por meio da expressão:

$$P_{Q_k} = \frac{k \times \sum fi}{4} \quad (k = 1, 2, 3)$$

- 2ª etapa: identificar a posição do **quartil** na coluna de frequências acumuladas, isto é, **a frequência acumulada imediatamente igual ou superior à posição do quartil**;
- 3ª etapa: verificar o valor da variável que corresponde a essa posição.

- **Quartil para dados agrupados em classes**

O cálculo do quartil para dados agrupados em classes será realizado por meio das seguintes etapas:

- 1ª etapa: determinar a **posição do quartil**, por meio da expressão:

$$P_{Q_k} = \frac{k \times \sum fi}{4} \quad (k = 1, 2, 3)$$

- 2ª etapa: identificar a posição do **quartil** na coluna de frequências acumuladas, isto é, **a frequência acumulada imediatamente igual ou superior à posição do quartil**;
- 3ª etapa: verificar as informações referentes à classe correspondente a essa posição; e

- 4ª etapa: calcular o valor do **quartil** por meio da fórmula apresentada a seguir, que **consiste em uma variação da fórmula da mediana para dados agrupados em classes**:

$$Q_k = l_{inf Q_k} + \left[ \frac{\frac{k \times \sum f_i}{4} - f_{acant}}{f_{Q_k}} \right] \times h_{Q_k}$$

## DECIS

- Valores de uma série que a dividem em **DEZ PARTES IGUAIS**, isto é, dez partes contendo o mesmo número de elementos (**10%**).
- **D1**: o primeiro decil corresponde à separação dos primeiros **10%** de elementos da série;
- **D5**: o quinto decil corresponde à separação de metade dos elementos da série, **coincidindo com a mediana (D5 = Md)**;
- **D9**: o nono decil corresponde à separação dos primeiros **90%** de elementos da série, ou dos últimos 10% de elementos da série.

- **Decil para dados não-agrupados**

→ O cálculo do **decil para dados não-agrupados** será realizado por meio de três etapas:

→ 1ª etapa: determinar a posição do decil, por meio da expressão:

$$P_{D_k} = \frac{k \times n}{10} \quad (k = 1, 2, \dots, 9)$$

→ 2ª etapa: identificar a posição mais próxima do rol;

→ 3ª etapa: verificar o valor que está ocupando essa posição.

- **Decil para dados agrupados sem intervalos de classe**

→ O cálculo do **decil para dados agrupados sem intervalos de classe** será realizado por meio das seguintes etapas:

→ 1ª etapa: determinamos a **posição do decil**, por meio da expressão:

$$P_{D_k} = \frac{k \times \sum fi}{10} \quad (k = 1, 2, \dots, 9)$$

→ 2.a etapa: identificamos a posição do **decil** na coluna de frequências acumuladas, isto é, **a frequência acumulada imediatamente igual ou superior à posição do decil;**

→ 3.a etapa: verificamos o valor da variável que corresponde a essa posição.

- **Decil para dados agrupados em classes**

→ O cálculo do **decil para dados agrupados em classes** será realizado por meio das seguintes etapas:

→ 1ª etapa: determinamos a **posição do decil**, por meio da expressão:

$$P_{D_k} = \frac{k \times \sum fi}{10} \quad (k = 1, 2, \dots, 9)$$

→ 2ª etapa: identificamos a posição do **decil** na coluna de frequências acumuladas, isto é, **a frequência acumulada imediatamente igual ou superior à posição do decil;**

→ 3ª etapa: verificamos as informações referentes à classe correspondente a essa posição; e

- 4.a etapa: calculamos o valor do **decil** por meio da fórmula apresentada a seguir, que **consiste em uma variação da fórmula da mediana para dados agrupados em classes**:

$$D_k = l_{inf_{D_k}} + \left[ \frac{\frac{k \times \sum f_i}{10} - f_{ac_{ant}}}{f_{D_k}} \right] \times h_{D_k}$$

## PERCENTIS

- Valores de uma série que a dividem em **CEM PARTES IGUAIS**, isto é, cem partes contendo o mesmo número de elementos (**1%**)
- **P1**: o primeiro percentil corresponde à separação do primeiro **1%** de elementos da série;
- **P50**: o quinquagésimo percentil corresponde à separação de metade dos elementos da série, **coincidindo com a mediana (P50 = Md)**;
- **P99**: o nonagésimo nono percentil corresponde à separação dos primeiros **99%** de elementos da série, ou do último **1%** de elementos da série.

- **Percentil para dados não-agrupados**

→ O **cálculo do Percentil para dados não-agrupados** será realizado por meio de três etapas:

→ 1ª etapa: determinar a posição do percentil, por meio da expressão:

$$P_{P_k} = \frac{k \times n}{100} \quad (k = 1, 2, \dots, 99)$$

→ 2ª etapa: identificar a posição mais próxima do rol;

→ 3ª etapa: verificar o valor que está ocupando essa posição.

- **Percentil para dados agrupados sem intervalos de classe**

→ O cálculo do **percentil para dados agrupados sem intervalos de classe** será realizado por meio das seguintes etapas:

→ 1ª etapa: determinar a **posição do percentil**, por meio da expressão:

$$P_{P_k} = \frac{k \times \sum fi}{100} \quad (k = 1, 2, \dots, 99)$$

- 2ª etapa: identificar a posição do **percentil** na coluna de frequências acumuladas, isto é, **a frequência acumulada imediatamente igual ou superior à posição do percentil;**
- 3ª etapa: verificar o valor da variável que corresponde a essa posição.

- **Percentil para dados agrupados em classes**

- O cálculo do **percentil para dados agrupados em classes** será realizado por meio das seguintes etapas:
- 1ª etapa: determinar a **posição do percentil**, por meio da expressão:

$$P_{P_k} = \frac{k \times \sum fi}{100} \quad (k = 1, 2, \dots, 99)$$

- 2ª etapa: identificar a posição do **percentil** na coluna de frequências acumuladas, isto é, **a frequência acumulada imediatamente igual ou superior à posição do percentil;**
- 3ª etapa: verificar as informações referentes à classe correspondente a essa posição; e

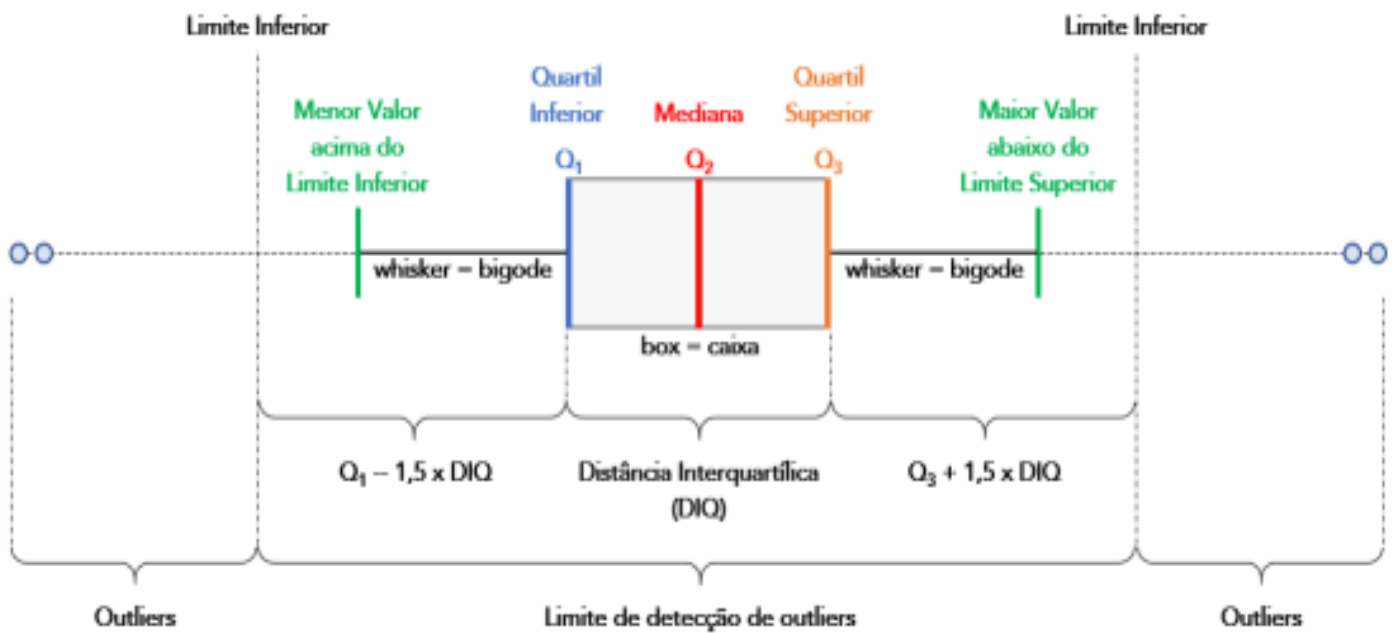


→ 4.a etapa: calculamos o valor do **percentil** por meio da fórmula apresentada a seguir, que **consiste em uma variação da fórmula da mediana para dados agrupados em classes**:

$$P_k = l_{inf P_k} + \left[ \frac{\frac{k \times \sum f_i}{100} - f_{ac_{ant}}}{f_{P_k}} \right] \times h_{P_k}$$

## BOX PLOT

Um **boxplot** ou **box-and-whisker plot** é uma ferramenta gráfica frequentemente utilizada na análise exploratória de dados que permite visualizar a distribuição dos dados e os valores discrepantes (outliers). Essa ferramenta **resume cinco medidas descritivas** de um conjunto de dados, incluindo: **o valor mínimo, o primeiro quartil, a mediana, o terceiro quartil e o valor máximo**.



A **distância interquartílica** (ou amplitude interquartílica, ou intervalo interquartílico) é calculada pela fórmula:

$$DIQ = Q_3 - Q_1$$