

ESTATÍSTICA DESCRITIVA ASSIMETRIA E CURTOSE III

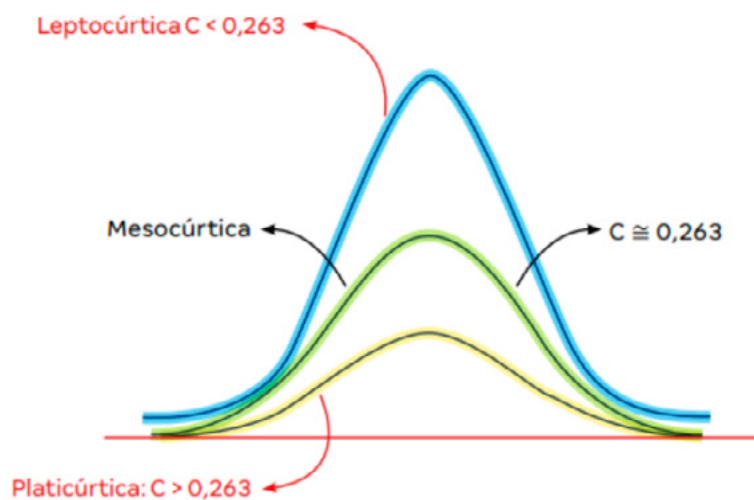
A análise estatística descritiva prossegue com o estudo da curtose, após a abordagem da simetria.

A simetria refere-se ao deslocamento da distribuição de frequência em direção à direita ou à esquerda, a depender dos valores da média, da moda e da mediana. Trata-se, portanto, de um movimento horizontal, relacionado às extremidades da curva.

Foram também apresentados os coeficientes de assimetria e os métodos de cálculo correspondentes.

CURTOSE

- Denomina-se curtose ao grau de “achatamento” de uma distribuição de frequências, geralmente unimodal, medido em relação ao de uma distribuição normal (de Gauss) que é tomada como padrão.



- Muito embora seja comum explicar a curtose como o “grau de achatamento” de uma distribuição de frequências, o que as medidas de curtose buscam indicar realmente é o grau de concentração de valores da distribuição em torno do centro desta distribuição. Numa distribuição unimodal, quanto maior for a concentração de valores em torno do centro da mesma, menor será o valor da sua curtose (achatamento). Gráficamente isto será associado a uma curva com a parte central mais afilada, mostrando um pico de frequência simples mais destacado, mais pontiagudo, caracterizando a moda da distribuição de forma mais nítida.

Isso ocorre porque o achatamento está diretamente relacionado à forma como os dados se distribuem em torno da média.

Quando a curva da distribuição é empurrada para baixo, os valores tendem a se dispersar. Ao contrário, quando a curva é elevada, os valores aproximam-se do centro da distribuição. Dessa forma, a análise da curtose envolve não apenas a visualização do achatamento da curva, mas, sobretudo, a observação da concentração dos dados centrais.

A distribuição pode apresentar diferentes graus de curtose, conforme o valor do coeficiente percentílico correspondente. Considera-se como referência o valor de aproximadamente 0,263. Assim, o coeficiente de curtose pode ser:

- Aproximadamente igual a 0,263;
- Superior a 0,263;
- Inferior a 0,263.

Esses valores indicam o grau de achatamento da curva. Quanto mais achatada for a curva, maior será o valor do coeficiente de curtose. Por exemplo, ao achatar progressivamente uma curva, o valor do coeficiente, que inicialmente era inferior a 0,263, poderá se igualar a esse valor e, posteriormente, superá-lo, conforme o achatamento aumenta.

Diz-se que uma distribuição de frequências é:

- **Mesocúrtica** – quando apresenta uma medida de curtose igual à da distribuição normal.
- **Platicúrtica** – quando apresenta uma medida de curtose menor que a da distribuição normal.
- **Leptocúrtica** – quando apresenta uma medida de curtose maior que a da distribuição normal.

A denominação dessas categorias depende do valor do coeficiente de curtose, que expressa o grau de concentração dos dados na região central da distribuição.

Quando o coeficiente de curtose é inferior a 0,263, a distribuição é classificada como **leptocúrtica**. Esse tipo de curva apresenta menor achatamento e maior concentração de valores no centro da distribuição.

Ao aumentar o achatamento da curva, o coeficiente aproxima-se do valor de referência. Quando o valor do coeficiente atinge aproximadamente 0,263, a distribuição passa a ser denominada **mesocúrtica**. Trata-se da **curva normal**, também chamada de **curva de Gauss**, que representa o padrão de referência para distribuições de frequência. O seu achatamento considerado adequado é representado por esse coeficiente.

Caso o achatamento continue a aumentar, o valor do coeficiente também se eleva. Nesse cenário, a distribuição é classificada como **platicúrtica**, caracterizada por menor concentração de valores centrais e maior dispersão.



5m

Quanto maior o achatamento, maior a dispersão dos dados. Quanto menor o achatamento, maior a concentração dos valores em torno do centro da distribuição.

COEFICIENTE PERCENTÍLICO DE CURTOSE

- Este coeficiente é definido como o quociente entre a amplitude semi-interquartílica e a amplitude entre o 10º e o 90º percentis.

O **coeficiente percentílico de curtose** é um indicador utilizado para mensurar o grau de achatamento de uma curva de distribuição. Esse coeficiente é calculado por meio de um quociente, que representa a razão entre a **amplitude semi-interquartílica** e a **amplitude entre os percentis 90 e 10**.

A amplitude semi-interquartílica é obtida a partir da diferença entre o terceiro quartil (Q3) e o primeiro quartil (Q1), dividida por dois:

$$\text{Amplitude semi-interquartílica} = \frac{Q3 - Q1}{2}$$

Já a amplitude entre os percentis 90 e 10 é dada pela diferença entre o nonagésimo percentil (P90) e o décimo percentil (P10):

$$\text{Amplitude entre os percentis} = P90 - P10$$

Em algumas representações, os percentis podem ser substituídos pelos decis correspondentes: D9 no lugar de P90 e D1 no lugar de P10. A substituição é válida, pois ambos representam os mesmos valores, diferindo apenas quanto à nomenclatura (percentual relativa ou absoluta).

A fórmula do coeficiente percentílico de curtose é, portanto:

$$C_p = \frac{\frac{Q3 - Q1}{2}}{P90 - P10}$$

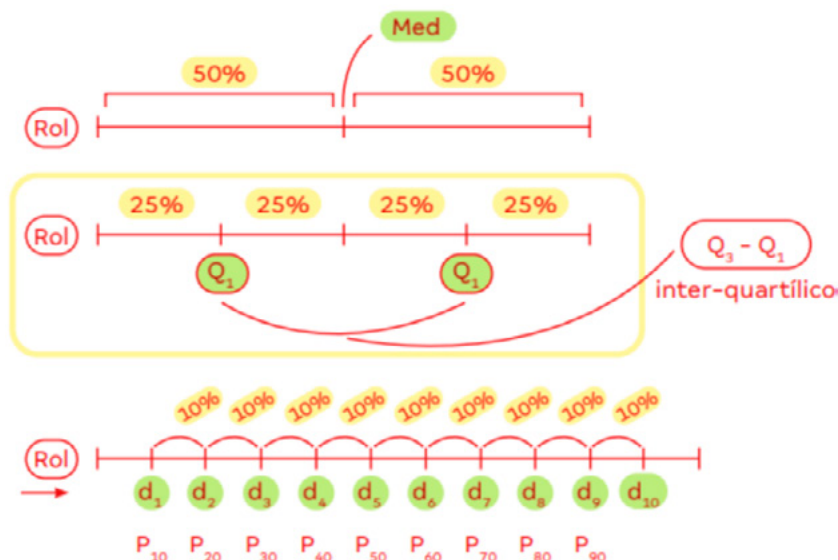
ou, alternativamente:

$$\frac{\frac{Q3 - Q1}{2}}{D9 - D1}$$

O valor deste coeficiente para a curva normal é 0,26367... Assim sendo, ao calcular o coeficiente percentílico de curtose de uma distribuição qualquer tem-se:

- Quando $C_p \cong 0,263 \rightarrow$ diz-se que a distribuição é mesocúrtica.
- Quando $C_p < 0,263 \rightarrow$ diz-se que a distribuição é leptocúrtica.
- Quando $C_p > 0,263 \rightarrow$ diz-se que a distribuição é platicúrtica.

A mediana é uma das principais separatrizes e corresponde ao valor central que divide a distribuição em duas partes iguais, ou seja, 50% dos dados estão abaixo da mediana e 50% estão acima.



Além da mediana, existem outras separatrizes importantes, como os **quartis**, **decis** e **percentis**:

Quartis: dividem a distribuição em quatro partes iguais. Os principais são:

- Q1 (primeiro quartil);
- Q2 (segundo quartil ou mediana);
- Q3 (terceiro quartil).

A **amplitude interquartílica** é obtida pela diferença entre Q3 e Q1:

$$\text{Amplitude interquartílica} = Q3 - Q1$$

Caso se deseje a **semi-amplitude interquartílica**, basta dividir essa diferença por dois:

$$\text{Semi-amplitude interquartílica} = \frac{Q3 - Q1}{2}$$

- **Decis:** dividem a distribuição em dez partes iguais, numerados de D1 a D9. Utilizam-se D1 e D9 para fins de cálculo do coeficiente percentílico de curtose. Note-se que não se inclui D10, pois, ao se dividir a distribuição em dez partes, são necessárias apenas nove divisões internas (D1 a D9).
- **Percentis:** dividem a distribuição em cem partes iguais. Utilizam-se, nesse contexto, os percentis P10 e P90.

A fórmula do **coeficiente percentílico de curtose** pode utilizar tanto os decis quanto os percentis, sendo expressa da seguinte forma:

$$\text{Coeficiente percentílico de curtose} = \frac{\left(\frac{Q3 - Q1}{2}\right)}{P90 - P10} \text{ ou } \frac{\left(\frac{Q3 - Q1}{2}\right)}{D9 - D1}$$

Obs.: o uso da fórmula não é altamente recorrente em provas, mas seu conhecimento pode ser útil, especialmente em questões de verdadeiro ou falso que testem a identificação correta de expressões estatísticas. É fundamental lembrar que:

- O numerador é a **semiampplitude interquartílica**.
- O denominador é a **diferença entre os percentis 90 e 10**, ou entre os decis 9 e 1.



DIRETO DO CONCURSO

01. (FGV/2024/TJ-MS/TÉCNICO DE NÍVEL SUPERIOR/ESTATÍSTICO/ESTATÍSTICA) Um estatístico realizou uma análise exploratória de dados sobre o desempenho acadêmico dos alunos de uma instituição de ensino. Ele coletou informações sobre as notas de uma avaliação e construiu um histograma para visualizar a distribuição das notas. Após a construção, o profissional observou que a distribuição era simétrica e concentrada em torno da moda. A medida descritiva utilizada para aferir o grau de achatamento da distribuição das notas é:

- assimetria;
- curtose;
- mediana;
- média;
- moda.



Diante desse enunciado, conclui-se que a medida descritiva que avalia o **grau de achatamento** da distribuição é a **curtose**.

A curtose é responsável por indicar o grau de concentração dos dados em torno da média, ou seja, o quão “achatada” ou “pontuda” é a curva da distribuição.

01. Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.



15m

	distribuição dos tempos gastos assistindo televisão (T, em horas)
1.º quartil	2
2.º quartil	4
3.º quartil	8
1.º decil	1
9.º decil	10

Tendo como referência essas informações, julgue o seguinte item. A amplitude total dos tempos T é igual ou superior a 9 horas.



Foram informados os valores do primeiro quartil (Q1), do segundo quartil (Q2) e do terceiro quartil (Q3). Ressalta-se que o segundo quartil (Q2) corresponde à **mediana**, uma vez que tanto a mediana quanto o Q2 dividem os dados em duas partes iguais: 50% dos valores abaixo e 50% acima.

Além dos quartis, o enunciado também apresenta decis, e não percentis. Caso seja necessário calcular o coeficiente de curtose baseado em percentis, os decis podem ser utilizados como referência.

Com base nas informações fornecidas, foi proposto o julgamento do seguinte item: “A amplitude total dos tempos é igual ou superior a 9 horas.”

Para a verificação desse item, considera-se que a amplitude total pode ser estimada por meio da diferença entre o **nono decil (D9)** e o **primeiro decil (D1)**, uma vez que esses valores delimitam o intervalo entre os extremos dos dados centrais.

Considerando que cada intervalo representado na escala do gráfico equivale a uma hora, e que o intervalo entre D1 e D9 abrange nove dessas unidades, conclui-se que a diferença entre D9 e D1 é igual a **9 horas**. Como esse valor não contempla as **extremidades da distribuição** (ou seja, os valores abaixo de D1 e acima de D9), infere-se que a **amplitude total** — que considera todo o intervalo de variação dos dados — é **superior a 9 horas**.

Portanto, o item está correto, pois a amplitude total dos tempos gastos assistindo à televisão é igual ou superior a 9 horas.

02. Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.

distribuição dos tempos gastos assistindo televisão (T, em horas)	
1.º quartil	2
2.º quartil	4
3.º quartil	8
1.º decil	1
9.º decil	10

A distribuição dos tempos T possui assimetria positiva.



Quando uma distribuição é simétrica, os valores da média, mediana e moda são aproximadamente iguais. No entanto, em distribuições assimétricas, a posição relativa dessas medidas varia conforme o tipo de simetria:

Na **simetria positiva** (ou assimetria à direita), a cauda da distribuição está voltada para a direita. Nessa situação:

- A média é maior que a mediana;
- A mediana é maior que a moda.

Na **simetria negativa** (ou assimetria à esquerda), a cauda da distribuição está voltada para a esquerda. Nessa situação:

- A média é menor que a mediana;
- A mediana é menor que a moda.

Em uma distribuição simétrica, as três medidas de tendência central coincidem:

Média = Mediana = Moda

No caso analisado, o segundo quartil (Q2) equivale à mediana, cujo valor é 4 horas. Já a média observada é 6 horas. Como a média é superior à mediana, conclui-se que a distribuição apresenta simetria positiva.

Portanto, a assertiva que afirma que a distribuição dos tempos possui simetria positiva está correta.

Por fim, em algumas distribuições, não há moda definida, o que caracteriza uma distribuição amodal. Nesses casos, mesmo na ausência da moda, é possível identificar o tipo de simetria da distribuição a partir da comparação entre a média e a mediana.

Se a média for maior que a mediana, a distribuição é assimétrica positivamente (ou com simetria positiva). No exemplo analisado, a média é igual a 6 e a mediana é igual a 4, portanto a distribuição apresenta simetria positiva.

03. Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.

distribuição dos tempos gastos assistindo televisão (T, em horas)	
1.º quartil	2
2.º quartil	4
3.º quartil	8
1.º decil	1
9.º decil	10

O índice percentílico de curtose foi superior a 0,4, o que sugere que a distribuição dos tempos T seja leptocúrtica



Inicialmente, percebe-se que, nesta questão, será necessária a utilização da fórmula.

$$C_p = \frac{\frac{Q3 - Q1}{2}}{P90 - P10}$$

Substituindo os valores fornecidos:

$$C_p = \frac{\frac{8 - 2}{2}}{10 - 1} = \frac{3}{9} \approx 0,333...$$

Portanto, o resultado aproxima-se de 0,33, que é inferior a 0,4.

04. Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.

distribuição dos tempos gastos assistindo televisão (T, em horas)	
1.º quartil	2
2.º quartil	4
3.º quartil	8
1.º decil	1
9.º decil	10

O desvio quartílico dos tempos T foi igual a 3.



Primeiramente, é fundamental compreender a equivalência entre os termos desvio quartílico e amplitude semi-interquartílica. Ambos expressam a mesma ideia e utilizam a mesma fórmula:

$$\text{Desvio quartílico} = \text{amplitude semi-interquartílica} = \frac{Q3 - Q1}{2}$$

Apesar de a expressão “desvio quartílico” poder dar a entender, equivocadamente, que se trata apenas de $Q3 - Q1$, trata-se, na verdade, da diferença entre o terceiro e o primeiro quartis dividida por dois, ou seja, a amplitude interquartílica dividida ao meio.

No caso da questão, pode-se substituir o valores da fórmula, sendo:

$$Q3 = 8 \text{ e } Q1 = 2.$$

$$\frac{8 - 2}{2} = \frac{6}{2} = 3$$

Portanto, o valor do desvio quartílico (ou da amplitude semi-interquartílica) é **3**.

Essa é uma observação importante, pois a confusão terminológica pode levar ao erro. Há autores que não abordam explicitamente essa equivalência, o que reforça a necessidade de atenção. O termo “semi-interquartílico” indica, justamente, a divisão da amplitude interquartílica por dois.

05. (PROVA: CESPE/CEBRASPE/2022/PC-RO/TÉCNICO EM NECROPSIA) Com base no conjunto de dados mostrado no quadro acima, tendo como medida de assimetria a expressão e representa a média amostral, M denota a mediana amostral e s é o desvio-padrão amostral, então o valor de A é igual a

1	0	1	2	1	1
---	---	---	---	---	---

- a) -0,5.
- b) 0.
- c) 0,1.
- d) 0,25.
- e) 1,6.



Dada a sequência de dados 1,0,1,2,1,1 a questão 5 propõe o cálculo da medida de assimetria utilizando a seguinte fórmula:

$$A = 3 \frac{\bar{x} - med}{S}$$

Onde:

- $\bar{x}$ é a média amostral,
- m é a mediana amostral,
- S é o desvio padrão amostral.

Para calcular a média, somam-se os valores da sequência:

$$\frac{1 + 0 + 1 + 2 + 1 + 1}{6} = \frac{6}{2} = 1$$

Logo, a média é 1.

Para calcular a mediana, é necessário ordenar os valores da sequência:

0, 1, 1, 1, 1, 2

Como o número de elementos é par, a mediana é a média dos dois valores centrais:

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

Logo, a mediana é 1.

O desvio padrão não precisa ser calculado nesse caso, pois a diferença entre a média e a mediana é 0. O produto de 0 por 3 também resulta em 0, o que implica que o valor da assimetria será 0. Portanto, o desvio padrão não afeta o cálculo.

Portanto, o valor da assimetria é 0.

Observações finais sobre curtose:

- Leptocúrtica: A curva é mais acentuada, com os valores mais concentrados. O coeficiente de curtose é maior que 0.
- Mesocúrtica: A curva segue a forma normal, com o coeficiente de curtose aproximado de 0,263.
- Platicúrtica: A curva é mais achatada, com os valores mais dispersos. O coeficiente de curtose é menor que 0.

Esses conceitos são importantes para compreender a distribuição dos dados e a forma da curva de distribuição.

GABARITO**01. b****02. C****04. C****01. C****03. E****05. b**

Este material foi elaborado pela equipe pedagógica do Gran Concursos, de acordo com a aula preparada e ministrada pelo professor Josimar Padilha.

A presente gravação tem como objetivo auxiliar no acompanhamento e na revisão do conteúdo ministrado na videoaula. Não recomendamos a substituição do estudo em vídeo pela leitura exclusiva deste material.
