By @kakashi_copiador

VARIÁVEIS ALEATÓRIAS DISCRETAS

Conceitos Iniciais

Introdução à Estatística Inferencial

Com esta aula, estamos iniciando os estudos em **Estatística Inferencial ou Inferência Estatística**, que é o ramo da Estatística que nos ajuda a tirar conclusões a respeito de um **todo** (que chamamos de **população**) a partir das observações feitas em uma **parte** dessa população (que chamamos de **amostra**). A inferência é uma técnica importante, pois normalmente **não** é possível conhecer a informação **exata** (por exemplo, altura, idade, salário, etc.) para **toda a população**.

Por exemplo, para conhecer a **altura média dos brasileiros**, os órgãos responsáveis por essa estatística NÃO verificam a altura de TODOS os brasileiros para conhecer exatamente a sua média. Em vez disso, é selecionada uma **amostra** de brasileiros, a partir da qual são feitas as **medições** necessárias. Por fim, com base nessas **medições da amostra**, são feitos os cálculos necessários para tirar **conclusões** a respeito da altura média da **população** de brasileiros.

Vamos destacar os seguintes conceitos do processo que acabamos de descrever:

- A **característica numérica da população** que se deseja conhecer (no nosso exemplo, a altura média da população de brasileiros) é chamada de **parâmetro populacional**;
- A medida correspondente feita na **amostra** (no caso, a altura média dos brasileiros da amostra selecionada) é chamada de **parâmetro de estimativa** ou **estatística da amostra**;
- As **conclusões** a respeito da população feitas a partir da amostra são chamadas de **inferência**.

As etapas desse processo de inferência estão representadas abaixo:

Hoje, vamos começar a estudar a base que permeia o processo.

A **inferência** do parâmetro populacional **não** é uma informação **exata**, a qual seria obtida somente se **toda** a população fosse verificada. Como os cálculos são feitos a partir de uma amostra, somente, os resultados vão **depender** da amostra selecionada. Ou melhor, vão **variar** de acordo com a amostra. Por isso, dizemos que as medidas obtidas a partir da amostra são **variáveis aleatórias**!

Variáveis Aleatórias

A definição de **variável aleatória**, ou simplesmente **v.a.**, é uma função que associa um **número real** a cada **ponto amostral**, isto é, a cada elemento do Espaço Amostral. Com isso, passamos a ter uma **caracterização numérica** do resultado de um experimento ou fenômeno aleatório.

Quando utilizamos o **número 0** (zero) para representar a face CARA e o **número 1** para representar a face COROA, criamos justamente uma **variável aleatória**! Outro exemplo de variável aleatória é atribuir o **número** indicado na face superior do dado {1,2,3,4,5,6} ao resultado do seu lançamento.

Assim como os resultados dos experimentos e fenômenos que representam, os valores das variáveis aleatórias também são **incertos** (variáveis que apresentam valores **certos**, com resultados **previamente** conhecidos, **não** são variáveis aleatórias). Apesar dessa incerteza, os resultados das variáveis aleatórias apresentam certo **padrão**, o que torna possível lhes atribuir **probabilidades**.

Por exemplo, se denotarmos a variável que representa o lançamento de uma moeda equilibrada por **X**, então a probabilidade de termos **X = 0** (o que equivale à probabilidade de obtermos a face CARA) é:

$$P(X=0)=\frac{resultados\ favoráveis\ (CARA)}{resultados\ possíveis\ (CARA, COROA)}=\frac{1}{2}$$

Note que mudamos um pouco a forma que escrevemos a probabilidade, pois agora estamos associando-a a uma **variável aleatória** e não mais a um **evento**. Por isso, em vez de escrevê-la como **P(A)**, para representar o evento A, estamos utilizando a forma **P(X = 0)**. Alternativamente, podemos utilizar a forma **P(0)**.

Da mesma forma, a probabilidade de termos **X = 1** (face COROA) é:

$$P(X=1)=\frac{resultados\ favoráveis\ (COROA)}{resultados\ possíveis\ (CARA, COROA)}=\frac{1}{2}$$

Se denotarmos a variável que representa o resultado do lançamento de um dado equilibrado por Y, então a probabilidade de termos Y = 1 é:

$$P(Y=1)=\frac{resultados\ favoráveis}{resultados\ possíveis}=\frac{1}{6}$$

A probabilidade de qualquer outro resultado do dado é a mesma:

$$P(Y=2)=\frac{1}{6} \qquad P(Y=3)=\frac{1}{6} \qquad P(Y=4)=\frac{1}{6} \qquad P(Y=5)=\frac{1}{6} \qquad P(Y=6)=\frac{1}{6}$$

Nesses dois exemplos, as variáveis aleatórias são **discretas**, pois a **quantidade de valores** que elas podem assumir é **enumerável** (isto é, contável). No caso da moeda, há **2** possíveis resultados; para o dado, há **6** possíveis resultados. Normalmente, os valores possíveis de uma variável discreta são números **racionais**.

> O conjunto dos números racionais ($\mathbb{Q}$) engloba os números inteiros e decimais (finitos ou infinitos com dízima periódica), ou seja, todos os números que podem ser descritos em forma de **fração**.
>
> **Não** estão englobados os números **irracionais** ($\mathbb{I}$), isto é, os números **infinitos sem dízima periódica**, como as constantes π = 3,1415... e o número neperiano *e* = 2,718...
>
> Juntos, os racionais e irracionais formam o conjunto dos números reais ($\mathbb{R}$).

Por outro lado, há variáveis aleatórias **contínuas**, cujos resultados **não são enumeráveis**. Essas podem assumir **quaisquer valores** dentro de um **intervalo** (ou conjunto de intervalos). Por exemplo, suponha que a variável Z represente a **quantidade de água** que um brasileiro ingere em um dia. Assim, Z pode assumir qualquer valor desde 0L até 5L (imagino). O intervalo de uma variável contínua pode ser, ainda, infinito (por exemplo, todos os números reais).

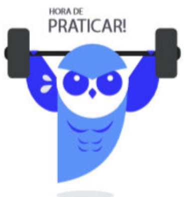
Para variáveis **contínuas**, **não** podemos atribuir **probabilidades** aos seus resultados. Ou seja, não podemos calcular a probabilidade de um brasileiro beber **exatamente** 2L de água em um dia, isto é, **exatamente** 2,000000... litros, nem um milésimo de litro a mais nem a menos. Podemos apenas atribuir **probabilidades** a um **intervalo**, por exemplo, entre 1,95L e 2,05L e **mensurar** os resultados obtidos.

Para variáveis **contínuas**, há **infinitos** valores possíveis. Porém, **não** é essa característica que distingue os dois tipos de variável aleatória. Isso porque as variáveis **discretas** também **podem** assumir um número **infinito** de valores, desde que tais valores sejam **enumeráveis**.

Por exemplo, suponha que uma variável aleatória represente o número de carros que chegam em um pedágio. Essa variável poderá assumir os valores 1, 2, 3, ... Não há um limite para a variável, pois sempre será possível chegar mais um carro. Portanto, há **infinitos** valores possíveis. Entretanto, esses valores são **enumeráveis** (contáveis), pois nunca aparecerá meio carro, ou qualquer outra parcela de carro.

Em resumo, as Variáveis Aleatórias Discretas e Contínuas possuem as seguintes características:

Variáveis Aleatórias Discretas	• Quantidade de valores possíveis é **enumerável** (finito ou não) • Atribuímos **probabilidades a resultados específicos**
Variáveis Aleatórias Contínuas	• Assumem qualquer valor dentro de um **intervalo** • Os resultados possíveis são **infinitos e não enumeráveis** • Não atribuímos probabilidade a resultados específicos, apenas a intervalos



(CESPE/2005 – Secretaria de Educação/MG) A identificação do tipo de variável é um requisito importante para a escolha do teste estatístico mais adequado. Acerca das variáveis, julgue o seguinte item.

Os valores das variáveis quantitativas discretas não podem ser contados, mas apenas mensurados.

Comentários:

Os valores das variáveis discretas podem ser contados, enquanto os valores das variáveis contínuas não podem ser contados, apenas mensurados.

Gabarito: Errado.

(2017 – SEDUC/MT) Sobre as variáveis serem discretas ou contínuas, analise as afirmativas abaixo, dê valores Verdadeiro (V) ou Falso (F).

( ) A contagem do número de alunos dentro de uma sala de aula só pode ser uma variável discreta, pois é um número inteiro racional e positivo.

( ) A contagem da quilometragem de um corredor em uma pista circular é uma variável contínua, pois este valor pode assumir qualquer valor dentro do intervalo real, no caso múltiplos de π (pi).

( ) O caso do termômetro analógico (de mercúrio), a variável representada nele é uma variável discreta, pois aceita todos os valores intermediários entre duas temperaturas a e b.

Assinale a alternativa que traga, de cima para baixo, a sequência correta.

a) V, V, F

b) V, V, V

c) V, F, V

d) F, F, V

e) F, V, F

Comentários:

Em relação ao primeiro item, o número de alunos dentro de uma sala é, de fato, uma variável discreta, pois os alunos podem ser enumerados. O número de alunos é um número racional e positivo. Portanto, a afirmativa é verdadeira.

Em relação ao segundo item, a distância percorrida por um corredor é uma variável contínua, pois pode assumir quaisquer valores reais, dentro de determinado intervalo, inclusive valores múltiplos de π, por se tratar de uma circunferência. Portanto, a afirmativa é verdadeira.

Em relação ao terceiro item, por aceitar todos os valores intermediários entre quaisquer duas temperaturas, os resultados de um termômetro analógico correspondem uma variável contínua. Portanto, a afirmativa é falsa.

Assim, a sequência correta é V, V, F.

Gabarito: A

VARIÁVEL ALEATÓRIA DISCRETA

Aqui, trataremos apenas de variáveis aleatórias discretas. Como já sabemos, para essas variáveis, podemos atribuir **probabilidades a resultados específicos**. É possível calcular a probabilidade de a face superior do dado lançado ser **especificamente** igual a 1, por exemplo.

Ou seja, sendo x um resultado possível para a variável X (no nosso exemplo, temos $x = 1$), então podemos calcular a probabilidade de a variável X assumir tal resultado, isto é, $X = x$:

$$P(X = x)$$

No lançamento de um dado, a probabilidade de obter a face $X = 1$ (ou qualquer outra face) é:

$$P(X = 1) = \frac{1}{6}$$

Agora, vamos supor que tenhamos testado um medicamento em uma amostra de 100 pessoas, das quais 80 pessoas apresentaram melhora em seu quadro. Então, podemos representar os resultados obtidos pela variável Y e dizer que Y = 1 representa o resultado de melhora e Y = 0 representa o resultado em que não houve melhora. Assim, a probabilidade de ter Y = 1 é:

$$P(Y = 1) = \frac{80}{100} = 0{,}8$$

Em outras palavras, podemos definir a probabilidade dessa variável como:

$$\boxed{P(Y = y) = \frac{n(Y = y)}{n(U)}}$$

Em que $n(Y = y)$ representa a **frequência** do resultado $Y = y$ (nesse exemplo, calculamos a probabilidade de Y = 1); e $n(U)$ representa a quantidade de **todos os resultados possíveis**.

Se estivéssemos observando os resultados em uma **amostra**, $n(Y = y)$ representaria o número de vezes em que foi observado o resultado y; e $n(U)$ representaria o número total de observações da amostra, chamado de **tamanho amostral**.

Chamamos essa função que atribui uma probabilidade a um resultado de uma variável aleatória discreta de **função de probabilidade**.

Por se tratar de uma probabilidade, essa função satisfaz aos **axiomas de probabilidade**: $\boldsymbol{P \geq 0}$ e $\boldsymbol{P(U) = 1}$.

TOME NOTA!

i)	$P(X = x) \geq 0$, pois **não** há **probabilidade negativa**;
ii)	Somatório das probabilidades de **todos** os possíveis resultados é igual a 1, pois a probabilidade de todo o **Espaço Amostral** é P(U) = 1.

Vimos que para o lançamento de um dado, a probabilidade de todos os resultados é $P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = P(X = 5) = P(X = 6) = \frac{1}{6}$, o que atende à primeira condição, por se tratar de um valor **positivo**, e também à segunda condição, pois:

$$P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6) =$$

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

No exemplo dos medicamentos, observamos que a frequência de pessoas que apresentaram melhora foi $P(Y = 1) = 0{,}8$. Então, para que o somatório de todos os possíveis resultados seja P(U) = 1, então a frequência de pessoas que **não** apresentaram melhora é:

$$P(Y = 0) = 1 - P(Y = 1) = 1 - 0{,}8 = 0{,}2$$

Assim, observamos que esse exemplo atende às 2 condições da função de probabilidade, isto é, todos os valores y apresentam valores de probabilidade não negativos, $P(Y = y) \geq 0$, e a soma das probabilidades de todos os possíveis valores é P(U) = 1.

Em vez de apresentar resultados fixos, como os que acabamos de ver, a função de probabilidade $P(X = x)$ também pode ser definida como uma **função** do valor de x, como no exemplo a seguir.

EXEMPLIFICANDO

A função de probabilidade pode ser definida como:

$$P(x = i) = \frac{1}{3^{i-1}} \text{ para } x = 2, 3, 4, \ldots$$

Ou seja, para **x = 2**, temos:

$$P(x = 2) = \frac{1}{3^{2-1}} = \frac{1}{3^1} = \frac{1}{3}$$

Para **x = 3**, temos:

$$P(x = 3) = \frac{1}{3^{3-1}} = \frac{1}{3^2} = \frac{1}{9}$$

A função de probabilidade pode ser qualquer função que atenda às **2 condições**: não haver valores negativos; e o somatório de todos os resultados possíveis ser igual a 1.

EXEMPLIFICANDO

Vamos definir a seguinte função de probabilidade, para um valor de k que ainda não conhecemos:

$$P(X = x) = \frac{x}{k} \text{ para } x = 1, 2, 3 \text{ e } 4$$

Ou seja, as probabilidade de X = 1, X = 2, X = 3 e X = 4 são:

$$P(X = 1) = \frac{1}{k}, \qquad P(X = 2) = \frac{2}{k}, \qquad P(X = 3) = \frac{3}{k}, \qquad P(X = 4) = \frac{4}{k}$$

E como descobrimos o valor de k? Considerando a condição de que o somatório de todas as probabilidades é igual a 1:

$$P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = 1$$

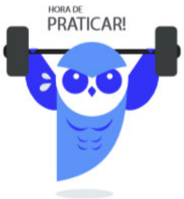
$$\frac{1}{k} + \frac{2}{k} + \frac{3}{k} + \frac{4}{k} = 1$$

$$\frac{1+2+3+4}{k} = \frac{10}{k} = 1$$

$$k = 10$$

Agora, podemos conhecer todas as probabilidades:

$$P(X = 1) = \frac{1}{10}, \qquad P(X = 2) = \frac{2}{10}, \qquad P(X = 3) = \frac{3}{10}, \qquad P(X = 4) = \frac{4}{10}$$



(2014 – Fundação João Pinheiro/MG) A fórmula P(x) = 3k/x representa a distribuição de probabilidade de uma variável aleatória, para x = 1,7,9. Portanto P(1 ≤ x ≤ 7) é igual a

a) 27/237

b) 23/63

c) 1/3

d) 216/237

e) 210/23

Comentários:

Sabendo que a x pode assumir 1, 7 ou 9, a soma dessas probabilidades deve ser igual a 1:

$$P(x = 1) + P(x = 7) + P(x = 9) = 1$$

$$\frac{3k}{1} + \frac{3k}{7} + \frac{3k}{9} = 1$$

$$\frac{3k \times 63 + 3k \times 9 + 3k \times 7}{63} = \frac{3k \times (63 + 9 + 7)}{63} = \frac{3k \times (79)}{63} = 1$$

$$3k \times 79 = 63$$

$$3k = \frac{63}{79}$$

Portanto, a probabilidade de x é:

$$P(X = x) = \frac{3k}{x} = \frac{63}{79.x}$$

Agora, podemos calcular o valor de $P(1 \leq X \leq 7)$:

$$P(1 \leq X \leq 7) = P(X = 1) + P(X = 7)$$

$$P(X = 1) = \frac{63}{79 \times 1} = \frac{63}{79}$$

$$P(X = 7) = \frac{63}{79 \times 7} = \frac{9}{79}$$

A soma será, portanto:

$$P(1 \leq X \leq 7) = \frac{63}{79} + \frac{9}{79} = \frac{72}{79}$$

Multiplicando o numerador e o denominador desse resultado por 3, obtemos a resposta:

$$P(1 \leq x \leq 7) = \frac{72}{79} = \frac{216}{237}$$

Gabarito: D

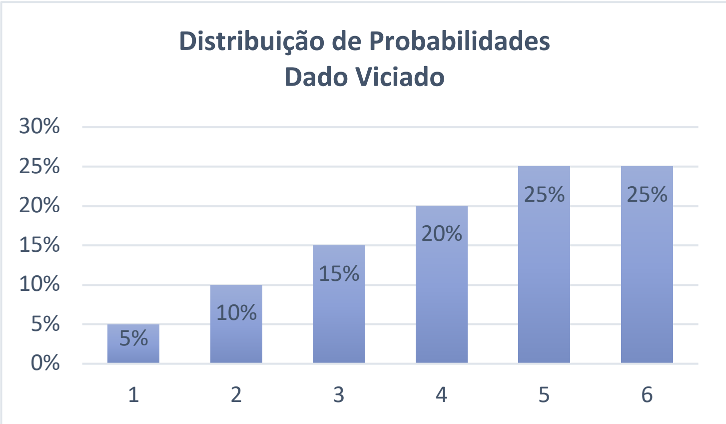
O **conjunto dos pares** $(x, P(X = x))$, isto é, o valor da variável e sua respectiva probabilidade, é chamado de **distribuição de probabilidade** da variável.

Uma forma de representá-la é por meio de uma **tabela** relacionando o valor da variável com a sua probabilidade. A seguir, apresentamos a distribuição de probabilidade para o lançamento de um **dado equilibrado**, em forma de tabela.

x_i	$P(X = x_i)$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$
$\sum_i P(X = x_i)$	**1**

Também é possível utilizar um **gráfico de barras**, com os valores da variável no eixo das abcissas (eixo horizontal) e os respectivos valores de probabilidade no eixo das ordenadas (eixo vertical).

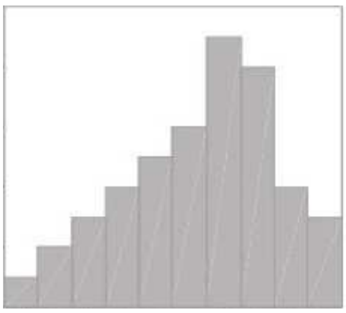
Para permitir uma melhor visualização da ferramenta, vamos supor um dado viciado (não equilibrado), com a distribuição de probabilidades apresentada no gráfico a seguir:



Podemos observar que a distribuição de probabilidades de cada face do dado viciado, representada na tabela acima, corresponde à distribuição representada na tabela abaixo:

x_i	$P(X = x_i)$
1	$5\% = 0{,}05$
2	$10\% = 0{,}10$
3	$15\% = 0{,}15$
4	$20\% = 0{,}20$
5	$25\% = 0{,}25$
6	$25\% = 0{,}25$
$\sum_i P(X = x_i)$	**1**

Pontue-se que o **gráfico de barras** é **diferente** do **histograma**, ilustrado abaixo.



O **histograma** apresenta uma distribuição de frequências de dados para categorias de **variáveis contínuas**, como faixas de altura de uma população, por exemplo, enquanto o **gráfico de barras** representa **variáveis discretas**. Por isso, **não há espaço entre as barras** de um **histograma**, enquanto **há espaço** em um **gráfico de barras**.

FIQUE ATENTO!

> Dizemos que duas ou mais variáveis aleatórias são **independentes e identicamente distribuídas** (**i.i.d.** ou **IID**) se todas as variáveis forem mutuamente independentes entre si e tiverem a **mesma distribuição de probabilidade**, isto é, mesmos $(x_i, P(x_i))$.

Vamos supor, por exemplo, uma variável Y com a seguinte distribuição de probabilidades:

$\boldsymbol{y_i}$	$\boldsymbol{P(Y = y_i)}$
1	0,05
2	0,10
3	0,15
4	0,20
5	0,25
6	0,25
$\boldsymbol{\sum_i P(Y = y_i)}$	**1**

Podemos observar que a variável X referente ao dado viciado e esta variável Y apresentam os mesmos resultados possíveis {1, 2, 3, 4, 5, 6} e as mesmas probabilidades associadas a cada resultado possível: P(1) = 0,05; P(2) = 0,10; P(3) = 0,15; P(4) = 0,20; P(5) = 0,25 e P(6) = 0,25. Portanto, podemos dizer que X e Y apresentam a **mesma distribuição de probabilidade**.

Assim, se X e Y forem independentes, então concluímos que essas variáveis são independentes e identicamente distribuídas (i.i.d.).

(CESPE/2015 – Telecomunicações Brasileiras S.A.) Considerando que Y_1, Y_2, ..., Y_n, ... sejam variáveis aleatórias independentes que satisfazem $P(Y_j = j) = P(Y_j = -j) = 1/2$ para j = 1, 2, ..., julgue o item que se segue.

As variáveis aleatórias Y_1, Y_2, ..., Y_n, ... são identicamente distribuídas.

Comentários:

Para Y_1, temos:

$$P(Y_1 = 1) = P(Y_1 = -1) = ½$$

Para Y_2, temos:

$$P(Y_2 = 2) = P(Y_1 = -2) = ½$$

O que segue indefinidamente. Para um Y_n qualquer, temos:

$$P(Y_n = n) = P(Y_n = -n) = ½$$

Ou seja, os valores de probabilidade são os mesmos, mas os valores que a variável assume (que normalmente chamamos de x, mas o enunciado chamou de j), **não** são os mesmos. Portanto, as variáveis **não** são identicamente distribuídas.

Gabarito: Errado.

MEDIDAS DE TENDÊNCIA CENTRAL

Neste tópico, veremos três **medidas de tendência central** relevantes para distribuições de probabilidade, quais sejam a esperança, a moda e a mediana. O objetivo dessas medidas é **resumir** a posição central das variáveis aleatórias, no intuito de **facilitar a análise** dos seus resultados.

Esperança Matemática

A esperança matemática de uma variável corresponde ao seu **valor médio**, podendo ser chamada também de **expectância, valor esperado** ou **média.**

Para ilustrar esse conceito, vamos supor que Maria enfrente **trânsito** de sua **casa até o trabalho**. Depois de algum tempo fazendo esse trajeto, Maria terá alguma noção de quanto tempo ela **costuma** levar para chegar no trabalho, isto é, uma **média** do tempo que ela leva.

Essa noção de quanto tempo se "costuma" ou se "espera" levar é justamente a **esperança** da variável. Neste último exemplo, a esperança corresponde ao tempo médio que a pessoa leva de casa ao trabalho; e no exemplo da altura dos brasileiros, a esperança corresponde à média de altura dos brasileiros.

Sendo X uma variável aleatória, a sua esperança é indicada por $\boldsymbol{E(X)}$ ou $\boldsymbol{\mu_X}$.

Para os exemplos dos lançamentos de moedas ou dados, em que os resultados são **equiprováveis**, a esperança corresponde à **média aritmética** dos resultados. Assim, para o lançamento de uma moeda, com faces 0 e 1, temos:

$$E(X) = \frac{0+1}{2} = 0{,}5$$

Para o lançamento de um dado, com faces de 1 a 6, temos:

$$E(Y) = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3{,}5$$

Contudo, no caso geral, para **qualquer variável aleatória discreta,** a esperança é calculada multiplicando-se cada valor da variável pela sua respectiva probabilidade, e, em seguida, somando-se todos os resultados.

TOME NOTA!

$$\boldsymbol{E(X) = \sum x.P(X = x)}$$

Aplicando essa fórmula geral[1] para a variável que representa o lançamento de uma moeda, em que $P(X=0)=\frac{1}{2}$ e $P(X=1)=\frac{1}{2}$, temos:

$$E(X)=0\times\frac{1}{2}+1\times\frac{1}{2}=\frac{1}{2}$$

Para o lançamento de um dado, em que todas as probabilidades são de $\frac{1}{6}$, temos:

$$E(Y)=1\times\frac{1}{6}+2\times\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+6\times\frac{1}{6}$$

$$E(Y)=\frac{1+2+3+4+5+6}{6}=\frac{21}{6}=3{,}5$$

Podemos efetuar esse mesmo cálculo, a partir da **tabela de distribuição de probabilidade**, criando **uma coluna** com o produto de x_i e $P(x_i)$. Assim, a esperança corresponderá à **soma** de **todas as linhas dessa nova coluna**.

x_i	$P(x_i)$	$x_i.P(x_i)$
1	$^1/_6$	$1\times{}^1/_6={}^1/_6$
2	$^1/_6$	$2\times{}^1/_6={}^2/_6$
3	$^1/_6$	$3\times{}^1/_6={}^3/_6$
4	$^1/_6$	$4\times{}^1/_6={}^4/_6$
5	$^1/_6$	$5\times{}^1/_6={}^5/_6$
6	$^1/_6$	$6\times{}^1/_6={}^6/_6$
Soma das Colunas	**1**	$\boldsymbol{E(X)=3,5}$

[1] A rigor, representamos o somatório junto a um índice i:

$$E(X)=\sum_{i=1}^{n}x_i.P(X=x_i)$$

Esse índice deixa claro que estamos somando para os diferentes valores de i, desde $i=1$ até $i=n$, ou seja, para $x_1, x_2, \ldots, x_n$.

Porém, nesta aula, daremos preferência à notação sem o índice, no intuito de simplificar a fórmula visualmente.

(2017 – Secretaria de Saúde/RO) Uma variável aleatória discreta X tem valores possíveis 0, 1, 2 e 3 com probabilidades respectivamente iguais a 0,2, 0,4, 0,3 e 0,1. A média de X é igual a

a) 1,0.

b) 1,3.

c) 1,5.

d) 1,8.

e) 1,9.

Comentários:

Vamos inserir as informações do enunciado na tabela de distribuição de probabilidade a seguir:

x	$P(x)$	$x.P(x)$
0	0,2	$0 \times 0{,}2 = 0$
1	0,4	$1 \times 0{,}4 = 0{,}4$
2	0,3	$2 \times 0{,}3 = 0{,}6$
3	0,1	$3 \times 0{,}1 = 0{,}3$

Assim, a esperança é a soma da última coluna:

$$E(X) = 0 + 0{,}4 + 0{,}6 + 0{,}3 = 1{,}3$$

Gabarito: B.

(VUNESP/2019 – Prefeitura de Campinas/SP) Sabe-se que as probabilidades de um carro transportar 1, 2, 3, 4 ou 5 pessoas são de 0,05, 0,20, 0,40, 0,25 e 0,10, respectivamente. Se em uma cidade chegaram 400 carros, a estimativa de pessoas que chegaram é de

a) 1400.

b) 1600.

c) 1260.

d) 2000.

e) 1320

Comentários:

A estimativa do número de pessoas transportadas **por carro** corresponde à **esperança** dessa distribuição. Inserindo as informações do enunciado na tabela de distribuição, temos:

x	$P(x)$	$x.P(x)$
1	0,05	$1 \times 0{,}05 = 0{,}05$
2	0,20	$2 \times 0{,}2 = 0{,}4$
3	0,40	$3 \times 0{,}4 = 1{,}2$
4	0,25	$4 \times 0{,}25 = 1$
5	0,10	$5 \times 0{,}1 = 0{,}5$

Assim, a esperança é a soma da última coluna:

$$E(X) = 0{,}05 + 0{,}40 + 1{,}20 + 1{,}00 + 0{,}50 = 3{,}15$$

Logo, são transportadas, em média, 3,15 pessoas por carro. Sabendo que chegam 400 carros, então a estimativa de pessoas é de:

400 x 3,15 = 1260

Gabarito: C

(FGV/2022 – PC/AM) Suponha que X, uma variável aleatória discreta, assuma a seguinte distribuição de probabilidade:

X	Proba(X)
0	0
1	1/4
2	1/4
3	K

O valor de K e o valor esperado de X são, respectivamente:

a) 0 e 3/4

b) 1/4 e 3/2

c) 1/2 e 3/4

d) 1/2 e 3/2

e) 1/2 e 9/4

Comentários:

Para resolver essa questão, o primeiro passo é calcular o valor de K, considerando que a soma das probabilidades dos possíveis resultados é igual a 1 (probabilidade de todo o Espaço Amostral):

$$P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 1$$

$$0 + \frac{1}{4} + \frac{1}{4} + K = 1$$

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

$$K = 1 - \frac{1}{2} = \frac{1}{2}$$

E para calcular a esperança, somamos os produtos dos valores de x com as respectivas probabilidades

$$E(X) = \sum x \times P(X = x)$$

$$E(X) = 0 \times 0 + 1 \times \frac{1}{4} + 2 \times \frac{1}{4} + 3 \times \frac{1}{2} = \frac{1}{4} + \frac{1}{2} + \frac{3}{2} = \frac{1 + 2 + 6}{4} = \frac{9}{4}$$

Gabarito: E

(CESPE/2010 – Agência Brasileira de Inteligência) Sabendo que X é variável aleatória discreta que pode assumir valores inteiros não negativos, julgue o próximo item.

A média de X é não negativa.

Comentários:

A média da variável é a soma dos produtos dos resultados pelas respectivas probabilidades:

$$E(X) = \sum x.P(X = x)$$

Sabemos que a **probabilidade** é sempre **maior ou igual a zero** (condição necessária). Então, se os valores da variável são **não negativos**, o produto de ambos **será sempre maior ou igual a zero** (não negativo). Consequentemente, a média, que corresponde à **soma** desses valores será **maior ou igual a zero**.

Gabarito: Certo.

(CESPE/2015 – Telecomunicações Brasileiras S.A.) Considerando que Y_1, Y_2, ..., Y_n, ... sejam variáveis aleatórias independentes que satisfazem $P(Y_j = j) = P(Y_j = -j) = 1/2$ para j = 1, 2, ..., julgue o item que se segue.

O valor esperado para a variável aleatória Y_j é nulo para todo número natural positivo j

Comentários:

Para ilustrar o cálculo da esperança para Y_j, vamos primeiro calcular a esperança para j = 1 e j = 2.

Sabemos que para Y_1, temos $P(Y_1 = 1) = P(Y_1 = -1) = ½$. Logo, a esperança de Y_1 é:

$$E(Y_1) = -1 \text{ x } ½ + 1 \text{ x } ½ = -½ + ½ = 0$$

Para Y_2, temos $P(Y_2 = 2) = P(Y_1 = -2) = ½$, então a esperança dessa variável é:

$$E(Y_2) = -2 \text{ x } ½ + 2 \text{ x } ½ = -1 + 1 = 0$$

Ou seja, para um Y_j qualquer, temos $P(Y_j = j) = P(Y_j = -j) = ½$ e sua esperança é dada por:

$$E(Y_j) = -j \text{ x } ½ + j \text{ x } ½ = 0$$

Logo, a esperança é nula para qualquer j = 1, 2, ...

Gabarito: Certo.

Propriedades da Esperança

Nesta seção, veremos **propriedades da esperança**. Vale adiantar que essas propriedades valem também para a esperança de variáveis aleatórias contínuas.

Nos enunciados abaixo, consideramos X e Y variáveis aleatórias e k uma **constante real** qualquer.

i) $\boldsymbol{E(k.X) = k.E(X)}$

De acordo com essa propriedade, a esperança de uma variável aleatória X, cujos valores foram multiplicados por uma constante k, é igual a k vezes a esperança da variável aleatória X.

Podemos considerar, como exemplo, um grupo de funcionários com salários distintos, de modo que a média seja R$ 5.000. Segundo essa propriedade, se os salários de todos os funcionários forem **dobrados**, então a **média** também será **dobrada** (passará para R$ 10.000).

Mas vamos "confirmar" isso. Digamos que os funcionários tenham os seguintes valores de salário, em mil reais:

X = {1, 2, 2, 2, 3, 5, 7, 7, 10, 11}

Sorteando um desses 10 funcionários ao acaso, o valor esperado do salário do funcionário sorteado pode ser calculado pela média aritmética de todos os 10 salários:

$$E(X) = \frac{1 + 2 + 2 + 2 + 3 + 5 + 7 + 7 + 10 + 11}{10} = \frac{50}{10} = 5$$

Duplicando os salários de todos os funcionários, temos:

Y = 2.X = {2, 4, 4, 4, 6, 10, 14, 14, 20, 22}

Nesse caso, o valor esperado será:

$$E(Y) = \frac{2 + 4 + 4 + 4 + 6 + 10 + 14 + 14 + 20 + 22}{10} = \frac{100}{10} = 10$$

Também podemos multiplicar uma variável por uma constante e representar a distribuição resultante utilizando a **tabela de distribuição de probabilidade**.

Para ilustrar, vamos considerar o mesmo exemplo dos salários. Considerando que há 10 elementos no total, a tabela de distribuição para X é:

x	P(X = x)
1	1/10
2	3/10
3	1/10
5	1/10
7	2/10
10	1/10
11	1/10

E a média (ou esperança) pode ser calculada como:

$$E(X) = 1 \times \frac{1}{10} + 2 \times \frac{3}{10} + 3 \times \frac{1}{10} + 5 \times \frac{1}{10} + 7 \times \frac{2}{10} + 10 \times \frac{1}{10} + 11 \times \frac{1}{10}$$

$$E(X) = \frac{1}{10} + \frac{6}{10} + \frac{3}{10} + \frac{5}{10} + \frac{14}{10} + \frac{10}{10} + \frac{11}{10} = \frac{50}{10} = 5$$

Para obter a tabela de distribuição de Y = 2X, **duplicamos os valores** da primeira coluna (referentes aos elementos da distribuição) e **mantemos as mesmas probabilidades**:

y	P(Y = y)
2	1/10
4	3/10
6	1/10
10	1/10
14	2/10
20	1/10
22	1/10

Calculando a esperança de Y, observamos que ela é o dobro da esperança de X:

$$E(Y) = 2 \times \frac{1}{10} + 4 \times \frac{3}{10} + 6 \times \frac{1}{10} + 10 \times \frac{1}{10} + 14 \times \frac{2}{10} + 20 \times \frac{1}{10} + 22 \times \frac{1}{10}$$

$$E(X) = \frac{2}{10} + \frac{12}{10} + \frac{6}{10} + \frac{10}{10} + \frac{28}{10} + \frac{20}{10} + \frac{22}{10} = \frac{100}{10} = 10$$

A mesma propriedade vale quando **dividimos** a variável pela constante, isto é:

$$\boldsymbol{E\left(\frac{X}{k}\right) = \frac{E(X)}{k}}$$

Ou seja, se cada salário fosse **dividido** por **2**, então a **média** também seria **divida** por **2**: passaria de R$ 5.000 para R$ 2.500.

Na verdade, essa é a **mesma propriedade** da multiplicação, pois ao invés de pensarmos que estamos dividindo por 2, podemos pensar que estamos multiplicando por $k = \frac{1}{2}$.

ii) $\boldsymbol{E(X + k) = E(X) + k}$

A esperança de uma variável aleatória X, sendo esta somada a uma constante k, é igual a k mais a esperança de X.

Ou seja, se todos os funcionários do nosso exemplo, cuja média salarial era de R$ 5.000, tiverem um aumento de R$ 2.000, então a média desse grupo passará para R$ 7.000, segundo essa propriedade.

Vamos "verificar" essa propriedade novamente. Sabendo que os salários originais eram X = {1, 2, 2, 2, 3, 5, 7, 7, 10, 11}, então, com um aumento de 2 mil reais, eles passarão a ser:

Y = X + 2 = {3, 4, 4, 4, 5, 7, 9, 9, 12, 13}

Assim, a esperança será:

$$E(Y) = \frac{3 + 4 + 4 + 4 + 5 + 7 + 9 + 9 + 12 + 13}{10} = \frac{70}{10} = 7$$

A mesma propriedade vale quando **subtraímos** a variável pela constante, isto é:

$$\boldsymbol{E(X - k) = E(X) - k}$$

Ou seja, se cada funcionário recebesse uma **redução** de R$ 2.000 do seu salário, então a **média** também seria **reduzida** em R$ 2.000: passaria de R$ 5.000 para R$ 3.000.

Na verdade, essa é a **mesma propriedade** da adição, pois ao invés de pensarmos que estamos subtraindo 2, podemos pensar que estamos somando k = -2.

iii) $\boldsymbol{E(X + Y) = E(X) + E(Y)}$

Por essa propriedade, temos que a esperança da soma de duas variáveis, X e Y, é igual à soma da esperança de X com a esperança de Y.

Digamos que um grupo de homens receba um salário médio E(X) = 5.000; e que um grupo de mulheres receba, em média, E(Y) = 4.000. Ao selecionar uma mulher e um homem, o valor do **salário somado** será, em média:

$$E(X + Y) = E(X) + E(Y) = 5.000 + 4.000 = 9.000$$

Para verificar essa propriedade, vamos utilizar um exemplo menor, pois envolve encontrar **todas as possíveis combinações** dos elementos dos dois grupos e **somá-los**. Suponha, então, o experimento de lançamento de duas moedas (X_1 e X_2), com faces representadas por 0 e 1.

Ao lançarmos ambas as moedas, temos os seguintes valores possíveis:

$$(X_1,X_2) = \{(0,0), (0,1), (1,0), (1,1)\}$$

Ao **somarmos** esses valores, temos:

$$X_1 + X_2 = \{(0 + 0), (0 + 1), (1 + 0), (1 + 1)\} = \{0,1,1,2\}$$

A média é então:

$$E(X_1 + X_2) = \frac{0+1+1+2}{4} = \frac{4}{4} = 1$$

Sabendo que o valor esperado dos dois lançamentos é $E(X_1) = E(X_2) = \frac{1}{2}$, então confirmamos que, de fato:

$$E(X_1 + X_2) = E(X_1) + E(X_2) = \frac{1}{2} + \frac{1}{2} = 1$$

A mesma propriedade vale para a **subtração** de variáveis:

$$\boldsymbol{E(X_1 - X_2) = E(X_1) - E(X_2)}$$

iv) $\boldsymbol{E(k) = k}$

Ou seja, o valor esperado de uma constante é igual à própria constante.

Digamos que em um grupo de funcionários, o salário individual de todos seja igual a k = 5 mil reais. Selecionando um funcionário ao acaso, qual será o salário esperado? Certamente, 5 mil reais.

Também podemos obter esse resultado, a partir da fórmula da esperança. Sendo k = 5, então para um grupo de n funcionários, a esperança é:

$$E(k) = \frac{n \times 5}{n} = 5$$

v) Se X e Y são independentes, então $\boldsymbol{E(X.Y) = E(X).E(Y)}$

Se X e Y são variáveis aleatórias independentes, então a esperança do produto de X e Y é igual ao produto da esperança de X com a esperança de Y.

Supondo novamente o lançamento das duas moedas, em que (X_1,X_2) = {(0,0), (0,1), (1,0), (1,1)}. Assim, o produto X_1 x X_2 corresponde ao conjunto:

$$X_1 \text{ x } X_2 = \{0,0,0,1\}$$

A média de X_1 x X_2 é, portanto:

$$E(X_1 \times X_2) = \frac{0+0+0+1}{4} = \frac{1}{4}$$

Sabendo que o valor esperado dos dois lançamentos é $E(X_1) = E(X_2) = \frac{1}{2}$, então, confirmamos que, de fato:

$$E(X_1 \times X_2) = E(X_1) \times E(X_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

No entanto, para isso é essencial que as variáveis X e Y sejam **independentes** (como é o caso de lançamentos distintos).

Por outro lado, é possível verificar a identidade E(X.Y) = E(X).E(Y) e as variáveis **não** serem independentes. Ou seja, se as variáveis são **independentes**, **então** podemos concluir que E(X.Y) = E(X).E(Y); mas se E(X.Y) = E(X).E(Y), não sabemos se as variáveis são ou não independentes.

RESUMINDO

i)	$\boldsymbol{E(kX) = k.E(X)}$
ii)	$\boldsymbol{E(X \pm k) = E(X) \pm k}$
iii)	$\boldsymbol{E(X \pm Y) = E(X) \pm E(Y)}$
iv)	$\boldsymbol{E(k) = k}$
v)	Se X e Y forem **independentes**, então $\boldsymbol{E(X \times Y) = E(X) \times E(Y)}$

Essas propriedades podem ser utilizadas **conjuntamente** e para qualquer número de variáveis aleatórias. Por exemplo, sendo X, Y e Z variáveis aleatórias, então:

$$E\left(3.X + \frac{Y}{4} - 2.Z + 1\right) = E(3.X) + E\left(\frac{Y}{4}\right) - E(2.Z) + E(1)$$

$$E\left(3.X + \frac{Y}{4} - 2.Z + 1\right) = 3.E(X) + \frac{E(Y)}{4}. -2.E(Z) + 1$$

(2016 – Instituto Federal de Educação/BA – *Adaptada*) Sendo X uma variável aleatória, com média μ, então a esperança matemática da função $Y = a + bX$, com a e b $\in$ R, é

a) E(Y) = a + b

b) E(Y) = a

c) E(Y) = $b\mu$

d) E(Y) = $a + b\mu$

e) E(Y) = a^2

Comentários:

Pelas propriedades da esperança, temos E(Y) = E(a + bx) = a +b.E(X). Como μ = E(X), temos:

$$E(Y) = a + b.\mu$$

Gabarito: D.

(FGV/2017 – IBGE – *Adaptada*) Para o caso de variáveis aleatórias quaisquer, existem diversas propriedades que se aplicam diretamente à esperança matemática. Dentre essas propriedades está:

a) E(X.Y) = E(X).E(Y)

b) E(X+a) = a

c) $E(X \pm Y) = E(X) \pm E(Y)$

d) E(a.X) = E(X), sendo a uma constante qualquer

Comentários:

Em relação à alternativa A, podemos afirmar que E(X.Y) = E(X).E(Y) **somente se** X e Y forem **independentes**. Como o enunciado não menciona que X e Y são independentes, então a alternativa A está incorreta.

Em relação à alternativa B, temos a seguinte propriedade da esperança:

$$E(X + a) = E(X) + a$$

Portanto, a alternativa B está incorreta.

Em relação à alternativa C, temos de fato a seguinte propriedade, para **quaisquer** variáveis aleatórias X e Y:

$$E(X + Y) = E(X) + E(Y)$$

$$E(X - Y) = E(X) - E(Y)$$

Portanto, a alternativa C está correta. Em relação à alternativa D, sabemos que: E(a.X) = a.E(X). Portanto, a alternativa D está incorreta.

Resposta: C.

Moda

A moda de uma variável aleatória é o seu valor **mais provável**, isto é, o **valor com maior probabilidade**.

No exemplo de lançamento de duas moedas, X_1 e X_2, a soma das variáveis resulta no seguinte conjunto de resultados possíveis:

$$X = X_1 + X_2 = \{0,1,1,2\}.$$

Assim, a probabilidade de cada resultado é:

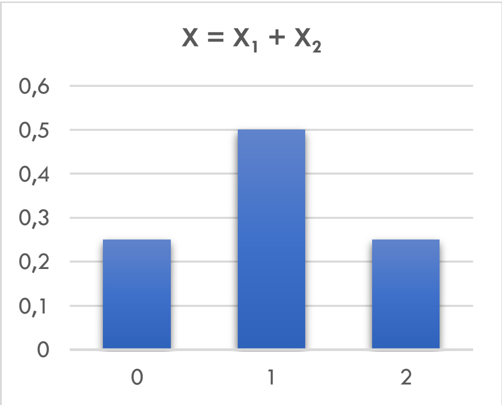
$$P(X = 0) = \frac{1}{4} = 0,25$$

$$P(X = 1) = \frac{2}{4} = 0,5$$

$$P(X = 2) = \frac{1}{4} = 0,25$$

Como a probabilidade de X = 1 é maior que as demais, então concluímos que a **moda** dessa variável é **X = 1**.

No gráfico de barras que representa a distribuição de probabilidades de uma variável, a moda pode ser identificada visualmente, pois estará associada à coluna mais alta, como representado abaixo.

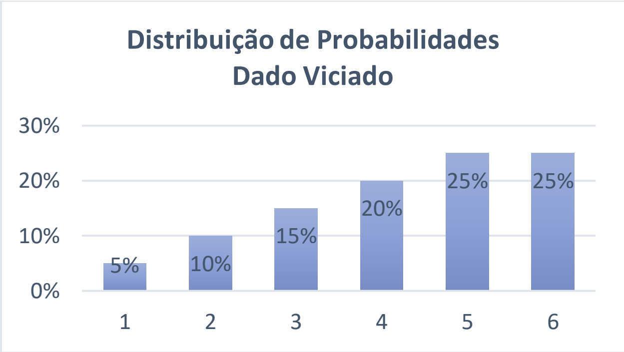


A moda é um valor da **variável aleatória** e **não** a sua **probabilidade**. Assim, no gráfico de barras, a moda estará no **eixo horizontal**.

Se estivermos lidando com uma amostra, a moda, chamada de **moda amostral**, corresponde ao valor da variável obtido com **maior frequência**.

É possível haver **mais de uma moda**, quando a **maior probabilidade** estiver associada a **mais de um resultado**.

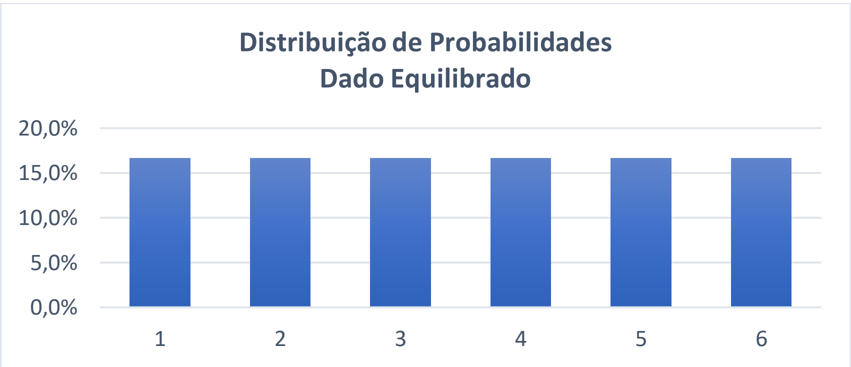
No exemplo anterior do dado viciado, tanto a face 5 quanto a face 6 apresentavam a maior probabilidade, de 25%.



Com 2 modas, a distribuição é chamada **bimodal**.

Uma distribuição **trimodal** apresenta 3 modas e uma distribuição **multimodal** apresenta múltiplas modas.

Por outro lado, também é possível que uma distribuição **não tenha moda**, o que ocorre todos os resultados da variável são **equiprováveis**, como é o caso do lançamento de uma moeda equilibrada e de um dado equilibrado, como ilustrado abaixo.



Nessa situação, a distribuição é dita **amodal**.

Propriedades da Moda

Nesta seção, veremos **propriedades da moda**, sendo X uma variável aleatória e k uma **constante real**.

i) $\boldsymbol{Mo(k.X) = k.Mo(X)}$

Quando uma variável X é multiplicada por uma constante k, a sua moda é igual k vezes a moda de X. Considerando o exemplo dos salários dos funcionários, se os salários dobram, a nova moda também será o dobro da moda anterior.

Para ilustrar, vamos considerar novamente os seguintes valores de salário, em mil reais:

X = {1, 2, 2, 2, 3, 5, 7, 7, 10, 11}

Podemos observar que a moda é igual a 2 mil reais. **Duplicando** os salários de todos os funcionários, temos:

2.X = {2, 4, 4, 4, 6, 10, 14, 14, 20, 22}

E a nova moda é igual a 4 mil reais, que é o dobro da moda anterior.

A mesma propriedade vale quando **dividimos** a variável pela constante, isto é:

$$\boldsymbol{Mo\left(\frac{X}{k}\right) = \frac{Mo(X)}{k}}$$

Ou seja, se cada salário fosse **dividido** por **2**, então a **moda** também seria **dividida** por **2**: passaria de 2 mil reais para mil reais.

ii) $\boldsymbol{Mo(X + k) = Mo(X) + k}$

Quando somamos uma constante k a uma variável X, a sua moda é acrescida da mesma constante k.

Em relação ao exemplo dos salários dos funcionários, se há um aumento de 3 mil reais no salário, a moda também terá esse mesmo aumento: passará de 2 mil reais para 5 mil reais.

Para ilustrar, sabendo que os salários originais eram X = {1, 2, 2, 2, 3, 5, 7, 7, 10, 11}, com um aumento de 3 mil reais, eles passarão a ser:

X + 2 = {4, 5, 5, 5, 6, 8, 10, 10, 13, 14}

Podemos observar que a nova moda é, de fato, de 5 mil reais.

A mesma propriedade vale quando **subtraímos** a variável pela constante, isto é:

$$\boldsymbol{Mo(X - k) = Mo(X) - k}$$

Ou seja, se cada funcionário recebesse uma **redução** de mil reais do seu salário, então a **moda** também seria **reduzida** em mil reais: passaria de 2 mil reais para mil reais.

RESUMINDO

i) $Mo(k.X) = k.Mo(X)$

ii) $Mo(X \pm k) = Mo(X) \pm k$

HORA DE PRATICAR!

(FCC/2018 – Prefeitura de Macapá/AP – Adaptada) A medida de tendência central que representa o valor com maior frequência na distribuição de uma amostra é a

a) média amostral.

b) variância.

c) amplitude total.

d) mediana.

e) moda amostral.

Comentários:

A moda de uma amostra é o valor obtido com maior frequência.

Gabarito: E.

(CESPE/2015 – Agente do Departamento Penitenciário Nacional)

quantidade diária de incidentes (N)	frequência relativa
0	0,1
1	0,2
2	0,5
3	0,0
4	0,2
total	1

Considerando os dados da tabela mostrada, que apresenta a distribuição populacional da quantidade diária de incidentes (N) em determinada penitenciária, julgue o item que se segue.

A moda da distribuição de N é igual a 4, pois esse valor representa a maior quantidade diária de incidentes que pode ser registrada nessa penitenciária.

Comentários:

A moda é o valor com maior frequência relativa, que é igual a 0,5.

Essa frequência está associada a N = 2, então a moda é N = 2.

Gabarito: Errado.

(CESPE/2018 – Departamento de Polícia Federal)

	dia				
	1	2	3	4	5
X (quantidade diária de drogas apreendidas, em kg)	10	22	18	22	28

Tendo em vista que, diariamente, a Polícia Federal apreende uma quantidade X, em kg, de drogas em determinado aeroporto do Brasil, e considerando os dados hipotéticos da tabela precedente, que apresenta os valores observados da variável X em uma amostra aleatória de 5 dias de apreensões no citado aeroporto, julgue o próximo item.

A moda da distribuição dos valores X registrados na amostra foi igual a 22 kg.

Comentários:

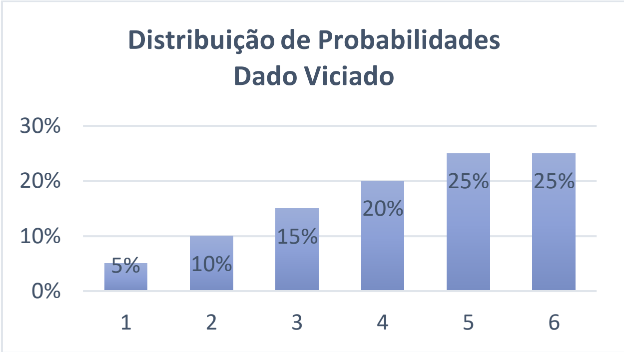
A moda de uma amostra é o valor obtido com maior frequência. No caso, foram apreendidos 22kg em 2 dias, enquanto as demais quantidades foram apreendidas em um único dia somente.

Gabarito: Certo

Mediana

A mediana de uma variável é o valor que divide a distribuição em **duas partes com mesma probabilidade**, de modo que a probabilidade dos valores menores ou iguais à mediana é igual a **50%** e a probabilidade dos valores maiores ou iguais à mediana é igual a **50%**.

Para ilustrar, o gráfico replicado a seguir representa a distribuição de probabilidades do dado viciado que vimos anteriormente.

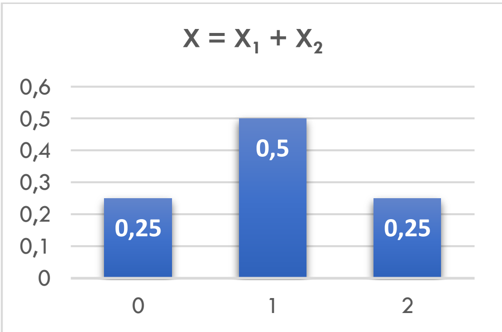


Podemos observar que a mediana está entre X = 4 e X = 5, pois a probabilidade associada aos valores menores ou iguais a 4 é de 50% e a probabilidade associada aos valores maiores ou iguais a 5 é igual a 50%.

Por convenção, quando a mediana está entre 2 valores, consideramos a média aritmética desses valores:

$$Md = \frac{4+5}{2} = 4{,}5$$

Agora, vamos considerar o exemplo em que somamos os resultados do lançamento de duas moedas, conforme o gráfico replicado a seguir:



Neste caso, temos 25% menor que 1 e 75% menor ou igual a 1; por outro lado, também temos 25% maior que a e 75% maior ou igual a 1. Ainda assim, a mediana será igual a 1. *Afinal, não faz sentido ela ser igual a 0 e nem igual a 2*.

Então, vamos ajustar a definição: a probabilidade dos valores menores ou iguais à mediana é de **pelo menos** 50%; e a probabilidade dos valores são maiores ou iguais à mediana também é de **pelo menos** 50%.

Propriedades da Mediana

As **propriedades da mediana** que veremos nesta seção são as mesmas das propriedades da moda que vimos anteriormente.

i) $\boldsymbol{Md(k.X) = k.Md(X)}$

Quando uma variável X é multiplicada por uma constante k, a sua mediana é igual k vezes a mediana de X.

Considerando o exemplo dos salários dos funcionários, se os salários dobram, a nova mediana também será o dobro da mediana anterior.

Para ilustrar, vamos considerar o mesmo exemplo dos salários, mas agora vamos analisar a tabela de distribuição de probabilidade da variável X = {1, 2, 2, 2, 3, 5, 7, 7, 10, 11}:

X	P(X = x)
1	10%
2	30%
3	10%
5	10%
7	20%
10	10%
11	10%

Podemos observar que a mediana está entre 3 e 5, pois 50% dos valores são menores ou iguais a 3 e 50% dos valores são maiores ou iguais a 5, de modo que a mediana é, por convenção:

$$Md(X) = \frac{3+5}{2} = 4$$

Duplicando os salários de todos os funcionários, temos s seguinte tabela de distribuição de probabilidade:

2.X	P(X = x)
2	10%
4	30%
6	10%
10	10%
14	20%
20	10%
22	10%

Agora, a mediana está entre 6 e 10, pois 50% dos valores são menores ou iguais a 6 e 50% dos valores são maiores ou iguais a 10, de modo que a nova mediana é:

$$Md(2.X) = \frac{6+10}{2} = 8$$

Que é o dobro da mediana anterior.

A mesma propriedade vale quando **dividimos** a variável pela constante, isto é:

$$\boldsymbol{Md\left(\frac{X}{k}\right) = \frac{Md(X)}{k}}$$

Ou seja, se cada salário fosse **dividido** por **2**, então a **mediana** também seria **dividida** por **2**: passaria de 4 mil reais para 2 mil reais.

ii) $\boldsymbol{Md(X+k) = Md(X) + k}$

Quando somamos uma constante k a uma variável X, a sua mediana é acrescida da mesma constante k.

Em relação ao exemplo dos salários dos funcionários, se há um aumento de 3 mil reais no salário, a mediana também terá esse mesmo aumento: passará de 4 mil reais para 7 mil reais. Vejamos:

X + 3	P(X = x)
4	10%
5	30%
6	10%
8	10%
10	20%
13	10%
14	10%

E a nova mediana está entre 6 e 8:

$$Md(X+3) = \frac{6+8}{2} = 7$$

A mesma propriedade vale quando **subtraímos** a variável pela constante, isto é:

$$\boldsymbol{Mo(X-k) = Mo(X) - k}$$

Ou seja, se cada funcionário recebesse uma **redução** de mil reais do seu salário, então a **mediana** também seria **reduzida** em mil reais: passaria de 4 mil reais para 3 mil reais.

RESUMINDO

i) $\boldsymbol{Md(k.X) = k.Md(X)}$
ii) $\boldsymbol{Md(X \pm k) = Md(X) \pm k}$

(CESPE/2020 – ME) Considerando que R representa uma variável quantitativa cuja média, mediana e variância são, respectivamente, iguais a 70, 80 e 100, e que $U = \frac{R}{10} - 7$, julgue o próximo item, acerca das variáveis U e R.

A mediana de U é negativa.

Comentários:

De acordo com as propriedades da mediana que vimos, a mediana de $U = \frac{R}{10} - 7$ é dada por:

$$Md(U) = \frac{R}{10} - 7 = \frac{Md(R)}{10} - 7$$

O enunciado informa que a mediana de R é Md(R) = 80, logo a mediana de U é:

$$Md(U) = \frac{80}{10} - 7 = 8 - 7 = 1$$

Ou seja, a mediana é **positiva**.

Gabarito: Errado.

FUNÇÃO DE DISTRIBUIÇÃO ACUMULADA

A **função de distribuição acumulada** de uma variável aleatória (ou simplesmente **f.d.a** ou **função de distribuição**) apresenta a **probabilidade acumulada** de todos os valores **menores ou iguais** a determinado valor.

TOME NOTA!

$$\boldsymbol{F(x) = P(X \leq x)}$$

Ou seja, equivale à **soma** de **todas as probabilidades menores ou iguais** ao valor em questão.

Por exemplo, no experimento de lançar um dado, a probabilidade de cada uma das faces, numeradas de 1 a 6, é $\frac{1}{6}$. Assim, o valor da função de distribuição acumulada para X = 1 equivale à probabilidade de X = 1, uma vez que não há valor menor:

$$F(1) = P(X \leq 1) = P(X = 1) = \frac{1}{6}$$

Para X = 2, a f.d.a. corresponde à soma das probabilidades de X = 1 e de X = 2:

$$F(2) = P(X \leq 2) = P(X = 1) + P(X = 2)$$

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

Para X = 3, temos a somar das probabilidades de X = 1, X = 2 e X = 3:

$$F(3) = P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3)$$

$$F(3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6}$$

O mesmo raciocínio se aplica para F(4), F(5) e F(6):

$$F(4) = P(X \leq 4) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$F(4) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{4}{6}$$

$$F(5) = P(X \leq 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$$

$$F(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{5}{6}$$

$$F(6) = P(X \leq 6) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5) + P(X = 6)$$

$$F(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6}$$

Para facilitar a visualização da f.d.a., podemos **incluir uma coluna** na **tabela da distribuição de probabilidade**. Para preenchê-la, basta somarmos o valor da função acumulada **acima** (valor de **X anterior**), com o valor da **probabilidade** da linha em questão (valor de **X atual**), como ilustrado pelas setas para **F(3)**.

$\boldsymbol{x}$	$\boldsymbol{P(X = x)}$	$\boldsymbol{F(x) = P(X \leq x)}$
1	$1/6$	$1/6$
2	$1/6$	$2/6$
3	$1/6$	$3/6$
4	$1/6$	$4/6$
5	$1/6$	$5/6$
6	$1/6$	$6/6$

Uma questão de prova também poderia apresentar essa f.d.a. da seguinte forma:

$$F(x) = \begin{cases} 0, se\ x < 1 \\ 1/6, se\ 1 \leq x < 2 \\ 1/3, se\ 2 \leq x < 3 \\ 1/2, se\ 3 \leq x < 4 \\ 2/3, se\ 4 \leq x < 5 \\ 5/6, se\ 5 \leq x < 6 \\ 1, se\ x \geq 6 \end{cases}$$

Também podemos calcular a função de **distribuição acumulada para uma amostra**, chamada de **distribuição amostral ou empírica acumulada**, a partir das **frequências relativas** observadas na amostra.

De maneira geral, a função acumulada de uma variável aleatória X (discreta ou contínua) apresenta as seguintes **características**:

i) F é **não decrescente**, porque as probabilidades são sempre **somadas**.

ii) Por ser uma **probabilidade**, a f.d.a. também assume valores somente entre 0 e 1:

$$\mathbf{0 \leq F(x) \leq 1}$$

A função de distribuição acumulada $\boldsymbol{F(x)}$ é **definida** em **toda a reta real**, ou seja, ela pode ser calculada para **qualquer valor de** $\boldsymbol{x}$.

Isso significa que não é só para X = 1, X = 2, X = 3, ..., X = 6 que podemos calcular o valor da função de distribuição acumulada. Ela também pode ser calculada, por exemplo, para X = 0,5; X = 8,1; X = 4,7; X = 5,3, mesmo para o exemplo do dado.

Para X = 0,5, em que X representa os resultados possíveis do lançamento de um dado, a f.d.a. corresponde à soma das probabilidades de todos os valores menores ou iguais a 0,5. Como o **menor valor possível é X = 1**, então a probabilidade acumulada até X = 0,5 é nula:

$$F(0{,}5) = P(X \leq 0{,}5) = 0$$

De forma geral, podemos dizer que, para os valores de x **menores** do que o **menor valor possível da variável** (no caso, para x < 1), o valor da f.d.a. é $\boldsymbol{F(x) = 0}$.

Para X = 8,1, a f.d.a. corresponde à soma das probabilidades de todos os valores menores ou iguais a 8,1. Como o **maior valor possível é X = 6**, então a probabilidade acumulada até X = 8,1 é igual à probabilidade acumulada até X = 6, isto é, à probabilidade de todo o Espaço Amostral:

$$F(8{,}1) = P(X \leq 8{,}1) = P(X \leq 6) = 1$$

De forma geral, podemos dizer que, para os valores de x **maiores ou iguais** ao **maior valor possível da variável** (no caso, para $x \geq 6$), o valor da f.d.a. é $\boldsymbol{F(x) = 1}$.

Para X = 4,7, a f.d.a. corresponde à soma das probabilidades de todos os valores menores ou iguais a 4,7. Como as faces do dado são valores **inteiros**, a probabilidade acumulada até X = 4,7 é igual à probabilidade acumulada até X = 4:

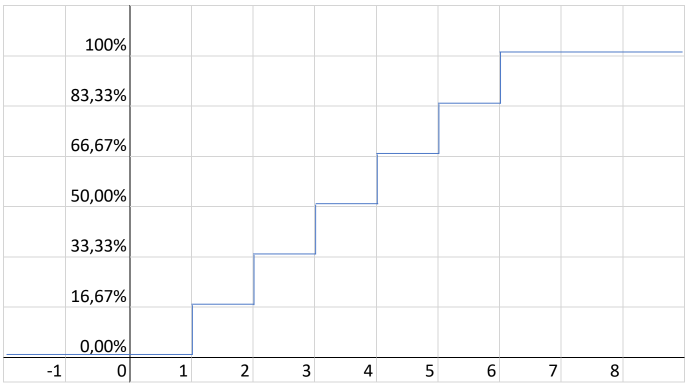
$$F(4{,}5) = P(X \leq 4{,}5) = P(X \leq 4) = \frac{4}{6}$$

Similarmente, para X = 5,3, a f.d.a. corresponde à probabilidade acumulada até X = 5:

$$F(5{,}3) = P(X \leq 5{,}3) = P(X \leq 5) = \frac{5}{6}$$

A f.d.a. pode ser calculada para **qualquer** valor de x, mas os seus valores serão **alterados** somente para **determinados valores de x**, quais sejam aqueles cuja probabilidade P(X = x) for diferente de zero.

Para o nosso exemplo do dado, a f.d.a. os valores da f.d.a. serão alterados somente para X = 1, X = 2, X = 3, ..., X = 6. Veja, no gráfico abaixo, como a f.d.a. dá saltos para esses valores de x.

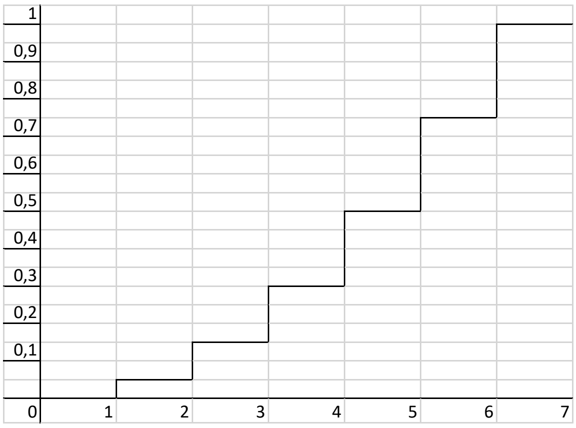


O tamanho dos "**saltos**" no gráfico da f.d.a. depende justamente do valor da **probabilidade** em cada ponto x, $P(X = x)$. Nesse exemplo, os "saltos" (ou diferenças) são todos iguais, pois os valores de X são todos equiprováveis.

Para o exemplo do dado viciado, temos a seguinte f.d.a.:

y	$P(Y = y)$	$F(y) = P(Y \leq y)$
1	0,05	0,05
2	0,10	0,15
3	0,15	0,30
4	0,20	0,50
5	0,25	0,75
6	0,25	1,00

Assim, o gráfico da f.d.a. para o dado viciado é:



Observe que os "saltos" apresentam tamanhos diferentes, uma vez que as probabilidades não são todas iguais.

Dessa forma, é possível percorrer o caminho **inverso**, ou seja, calcular a **função de probabilidade $P(X = x)$, a partir da função de distribuição acumulada $F(x)$**. Para isso, calculamos a **diferença** entre o valor da f.d.a. **no ponto** e o valor da f.d.a. **no ponto anterior**.

Por exemplo, para um dado equilibrado, a probabilidade da face $X = 4$ pode ser calculada a partir dos valores da f.d.a. nos pontos X = 4 e X = 3:

$$P(X = 4) = F(4) - F(3) = \frac{4}{6} - \frac{3}{6} = \frac{1}{6}$$

Para o nosso exemplo do dado viciado, temos:

$$P(Y = 4) = F(4) - F(3) = 0{,}5 - 0{,}3 = 0{,}2$$

É possível calcular, ainda, a **probabilidade de um intervalo de valores**, a partir da f.d.a. Para o nosso exemplo do dado equilibrado, a probabilidade do intervalo $P(2 < X \leq 5)$ pode ser calculada pela diferença entre a f.d.a. para x = 5 e a f.d.a. para x = 2:

$$P(2 < X \leq 5) = F(5) - F(2) = \frac{5}{6} - \frac{2}{6} = \frac{3}{6}$$

Todavia, para isso, é importante atentar-se principalmente se a **igualdade** está ou não **contemplada** em cada extremo do intervalo. A diferença entre a f.d.a. no ponto X = 5 e a f.d.a. no ponto X = 2 fornece a probabilidade de a face do dado ser **maior que** 2 e **menor ou igual** a 5.

$$F(5) - F(2) = P(X \leq 5) - P(X \leq 2) = P(2 < X \leq 5)$$

Isso porque, no ponto X = 5, a f.d.a. contempla a probabilidade para todo valor menor ou igual a 5 e no ponto X = 2, ela contempla a probabilidade para todo valor **menor ou igual** a 2. Logo, quando subtraímos uma da outra, teremos a probabilidade de todo valor menor ou igual a 5 e maior que 2.

TOME NOTA!

De maneira geral, para $a < b$, temos:

$$\boldsymbol{F(b) - F(a) = P(a < X \leq b)}$$

Se precisarmos de um intervalo de uma forma diferente, precisaremos **adaptá-lo** para que fique dessa forma.

Se o extremo inferior estiver contemplado no intervalo, ou seja, se o intervalo for da forma:

$$P(a \leq X \leq b)$$

fazemos a adaptação buscando um valor **menor** que a como **novo extremo inferior**, o qual **não** estará contemplado no novo intervalo.

Para o nosso exemplo do dado, se a face 2 estiver **contemplada** no intervalo, ou seja, se o intervalo for $2 \leq X \leq 5$, então fazemos a adaptação, **subtraindo** 1 unidade. Logo, o novo extremo inferior será 2 – 1 = 1:

$$P(2 \leq X \leq 5) = P(1 < X \leq 5) = F(5) - F(1)$$

$$P(2 \leq X \leq 5) = \frac{5}{6} - \frac{1}{6} = \frac{4}{6}$$

Por outro lado, se extremo superior **não** estiver **contemplado** no intervalo, ou seja, se o intervalo for da forma $a < X < b$, então fazemos a adaptação buscando um valor **menor** que b como **novo extremo inferior**, o qual estará contemplado no novo intervalo.

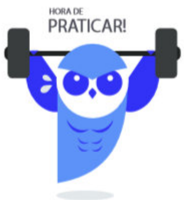
Por exemplo, se a face 5 **não** estiver contemplada, ou seja, se o intervalo for $2 < X < 5$, então fazemos a adaptação, **subtraindo** 1 unidade. Logo, o novo extremo superior será 5 – 1 = 4:

$$P(2 < x < 5) = P(2 < x < 4) = F(4) - F(2)$$

$$P(2 < x < 5) = \frac{4}{6} - \frac{2}{6} = \frac{2}{6}$$

Se o intervalo for da forma $a \leq x < b$, faremos **as duas adaptações**, ou seja, buscamos um valor menor que a **e** um valor menor que b:

$$P(2 \leq x < 5) = P(1 < X \leq 4) = F(4) - F(1) = \frac{4}{6} - \frac{1}{6} = \frac{3}{6}$$



(FGV/2018 – Analista Legislativo/RO) Uma variável aleatória discreta X tem função de probabilidade dada por:

x	-2	-1	0	1
p(x)	0,1	0,2	0,3	0,4

Se F(x) representa a função de distribuição de X, ∀ x real, então F(-0,8) é igual a

a) 0,3.

b) 0,4.

c) 0,5.

d) 0,6.

e) 1,0.

Comentários:

A função acumulada F(-0,8) corresponde à probabilidade P(X ≤ -0,8), que nesse caso é igual à soma das probabilidades P(X = –2) + P(X = –1):

$$F(-0{,}8) = P(X \leq -0{,}8) = P(X = -2) + P(X = -1) = 0{,}1 + 0{,}2 = 0{,}3$$

Gabarito: A

(FGV/2015 – TJ/RO) A função distribuição de probabilidade acumulada da variável “número de anos de experiência de magistrados” de um dado tribunal é dada por:

Anos (X)	0	5	10	15	25	35
F(x)	0	0,30	0,48	0,69	0,85	1

Então, a probabilidade de que um magistrado escolhido ao acaso tenha experiência maior do que cinco anos e menor ou igual a 15 anos é igual a:

a) 0,39.

b) 0,45.

c) 0,48.

d) 0,57.

e) 0,61.

Comentários:

A probabilidade de $P(5 < x \leq 15)$ é igual à diferença entre os valores da função acumulada F(15) – F(5), uma vez que o intervalo não precisa ser adaptado, pois o extremo superior está contemplado e o inferior, não:

$$P(5 < x \leq 15) = F(15) - F(5) = 0{,}69 - 0{,}30 = 0{,}39$$

Gabarito: A

(FGV/2017 – MPE/BA) Considere a variável aleatória do tipo discreta(X), relativa às fases de andamento de um processo podendo assumir apenas três valores numéricos 1, 2 ou 3, conforme o mesmo esteja em conhecimento, liquidação ou execução, respectivamente. Se F(.) é a função distribuição acumulada correspondente, com F(1,17) = 0,15 e F(2,76) = 0,45. Então é verdadeiro que

a) $P(X > 1{,}9) = 0{,}75$ e $P(X < 2{,}5) = 0{,}60$.

b) $P(X < 2{,}70) < 0{,}45$ e $P(X > 1{,}5) = 0{,}85$.

c) $P(X = 1) = 0{,}15$ e $P(X = 2) = 0{,}30$.

d) $P(X = 3) = 0{,}55$ e $E(X) = 2{,}70$.

e) $P(1{,}44 < X < 3) = 0{,}85$ e $Mo(X) = 3$.

Comentários:

O enunciado informa que há apenas três valores possíveis: X = 1, X = 2 ou X = 3.

Para conhecermos as probabilidades de cada valor, o enunciado informa os valores que a função de distribuição acumulada assume:

- $F(1{,}17) = P(X \leq 1{,}17) = 0{,}15$.

  O único valor menor ou igual a 1,17 é o valor X = 1. Ou seja:

$$P(X \leq 1{,}17) = P(X = 1) = 0{,}15$$

Assim, concluímos que:

$$P(X = 1) = 0{,}15$$

- $F(2{,}76) = P(X \leq 2{,}76) = 0{,}45$.

Os valores menores ou iguais a 2,76 são $X = 1$ e $X = 2$, ou seja:

$$P(X \leq 2{,}76) = P(X = 1) + P(X = 2) = 0{,}45$$

Sabemos que $P(X = 1) = 0{,}15$, logo:

$$P(X \leq 2{,}76) = 0{,}15 + P(X = 2) = 0{,}45$$

$$P(X = 2) = 0{,}30$$

Isso nos permite concluir que a alternativa C está correta, mas vejamos as demais alternativas.

Em relação à alternativa A, o valor de $P(X > 1{,}79)$ pode ser calculado como:

$$P(X > 1{,}79) = 1 - P(X \leq 1{,}79) = 1 - P(X = 1) = 1 - 0{,}15 = 0{,}85$$

E o valor de $P(X < 2{,}5)$ é:

$$P(X < 2{,}5) = P(X = 1) + P(X = 2) = 0{,}15 + 0{,}30 = 0{,}45$$

Logo, a alternativa A está incorreta.

Em relação à alternativa B, o valor de $P(X < 2{,}70)$ é:

$$P(X < 2{,}70) = P(X = 1) + P(X = 2) = 0{,}15 + 0{,}30 = 0{,}45$$

Ou seja, $P(X < 2{,}70) = 0{,}45$ (não $< 0{,}45$).

E o valor de $P(X > 1{,}5)$ pode ser calculado como:

$$P(X > 1{,}5) = 1 - P(X \leq 1{,}5) = 1 - P(X = 1) = 1 - 0{,}15 = 0{,}85$$

A segunda parte está correta, mas a alternativa B está incorreta, porque a primeira parte é falsa.

Em relação à alternativa D, o valor de $P(X = 3)$ é:

$$P(X = 3) = 1 - P(X = 1) - P(X = 2) = 1 - 0{,}15 - 0{,}30 = 0{,}55$$

E o valor da esperança é:

$$E(X) = 1x0{,}15 + 2x0{,}30 + 3x0{,}55 = 0{,}15 + 0{,}60 + 1{,}65 = 2{,}40$$

A primeira parte está correta, mas a alternativa D está incorreta, porque a segunda parte é falsa.

Em relação à alternativa E, o valor de $P(1{,}44 < X < 3)$ é:

$$P(1{,}44 < X < 3) = P(X = 2) = 0{,}30$$

E a moda de X (valor com maior probabilidade) é, de fato, $Mo(X) = 3$.

A segunda parte está correta, mas a alternativa E está incorreta, porque a primeira parte é falsa.

Gabarito: C

Quartis e Mediana

A partir da função de distribuição acumulada, podemos calcular a **mediana** da distribuição. A mediana é, assim como a esperança e a moda, uma medida de tendência central. Ela divide a distribuição em **2 partes iguais**, **de forma que 50% das observações fiquem abaixo dessa medida e 50% das observações fiquem acima.** Portanto, a **mediana** é o valor de X para o qual a **f.d.a. é igual a 50%.**

TOME NOTA!

$$\boldsymbol{F(x_{Mediana}) = 50\% = 0,5}$$

Também é possível calcular os **quartis** da distribuição (Q1, Q2 e Q3), os quais dividem a distribuição em **4 partes iguais**.

O primeiro quartil (Q1) deixa 25% das observações abaixo e 75%, acima; o segundo quartil (Q2) deixa 50% das observações abaixo e 50%, acima; e o terceiro quartil (Q3) deixa 75% das observações abaixo e 25%, acima. Note que a **mediana** equivale ao **segundo quartil**!

Assim, os valores da f.d.a. no primeiro, no segundo e no terceiro quartis são:

$$F(x_{Q1}) = 25\% = 0{,}25$$

$$F(x_{Q2}) = F(x_{Med}) = 50\% = 0{,}5$$

$$F(x_{Q3}) = 75\% = 0{,}75$$

Para as variáveis discretas, pode não ser possível encontrar valores para x que separem a distribuição **exatamente** nesses percentuais. Nesses casos, devemos encontrar os valores de X para os quais a **f.d.a.** apresenta valores **maiores** (ou iguais) a esses percentuais:

$$F(x_{Q1}) \geq 25\%$$

$$F(x_{Q2}) = F(x_{Med}) \geq 50\%$$

$$F(x_{Q3}) \geq 75\%$$

Para ilustrar, replicamos, a seguir, a tabela com os valores da f.d.a. (em percentual) para o lançamento do dado:

x	$F(x) = P(X \leq x)$
1	16,7%
2	33,3%
3	50%
4	66,7%
5	83,3%
6	100%

Podemos observar que não há um valor de X para o qual a f.d.a. seja exatamente igual a 25%. Por isso, escolhemos a probabilidade imediatamente **superior**, qual seja, de 33,3%, associada a X = 2. Com isso, concluímos que o **primeiro quartil** é:

$$x_{Q1} = 2$$

Também observamos que não há um valor de X para o qual a f.d.a. seja exatamente igual a 75%. A probabilidade imediatamente **superior** é de 83,3%, associada a X = 5. Logo, o terceiro quartil é

$$x_{Q3} = 5$$

Em relação à mediana, podemos observar que a f.d.a. para X = 3 é exatamente $F(3) = 50\%$. Isso significa que $x_{Med} = 3$ é um **possível** valor para a mediana. Entretanto, esse **não é o único** valor.

Lembra que a função acumulada é definida para qualquer valor de x? Pois é! Para **todos** os valores de x que pertencem ao intervalo [3,4)[1], temos $\boldsymbol{F(x) = 50\%}$. Sendo assim, podemos dizer que **quaisquer desses valores são a mediana**, inclusive o próprio 4. Entretanto, por convenção, normalmente consideramos o **valor médio** desse intervalo como a mediana:

$$x_M = \frac{3+4}{2} = 3,5$$

[1] Utilizamos o **parêntesis**, ou o **colchete voltado para fora**, para indicar que o intervalo é **aberto** naquele extremo, o que significa que o extremo **não** está incluído no intervalo.

Utilizamos o **colchete voltado para dentro** para indicar que o intervalo é **fechado** naquele extremo, o que significa que o extremo está **incluído**.

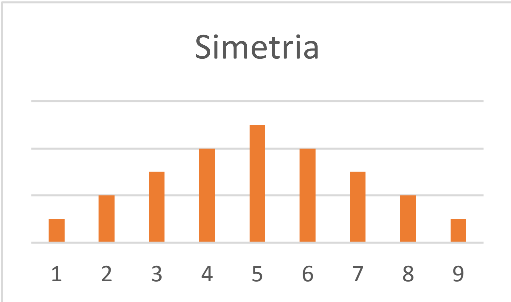
Assim, a expressão [3,4) equivale a [3,4[e representa o intervalo $3 \leq X < 4$.

Além dos quartis, há outros **quantis**, que dividem a distribuição em partes **iguais**, como os decis, que dividem em 10 partes iguais, e os percentis, que dividem em 100 partes iguais. Calculamos esses quantis da mesma forma, a partir da função de distribuição acumulada.

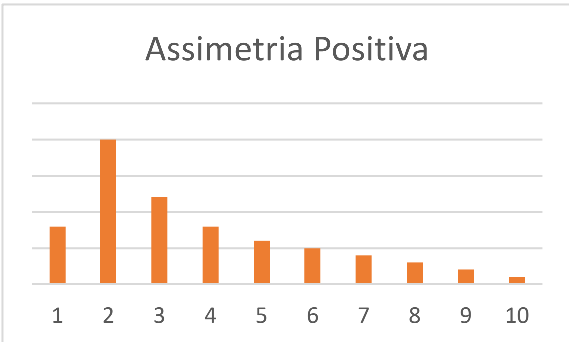
Essas medidas podem ser utilizadas para calcular a **assimetria** da distribuição probabilística, que representa o quanto a distribuição se **afasta** de uma distribuição **simétrica**. Em distribuições simétricas, temos:

Distribuição Simétrica: Média = Mediana = Moda

Consequentemente, toda a distribuição se afasta desse valor, tanto para cima, quanto para baixo, da **mesma forma**, como ilustrado abaixo.



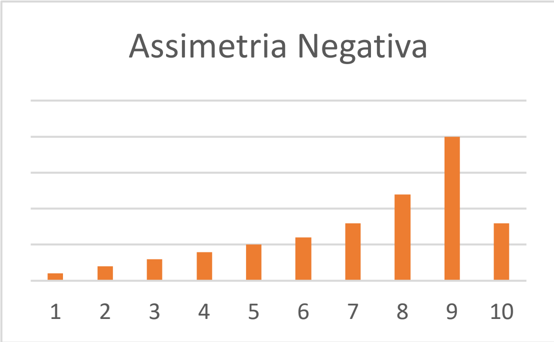
Em uma distribuição **positivamente assimétrica**, a distribuição é **alongada à direita** (cauda direita mais longa). Isso significa que há uma quantidade maior de valores **superiores** à **moda.** Observe, no gráfico abaixo, que há mais valores superiores à moda X = 2 (quais sejam, de X = 3 a X = 10) do que inferiores (apenas X = 1).



Nessa situação, temos a seguinte relação entre moda, mediana e média:

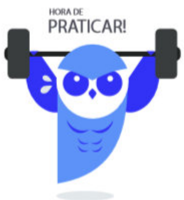
Distribuição Positivamente Assimétrica: Moda < Mediana < Média

Em uma distribuição **negativamente assimétrica**, a distribuição é **alongada à esquerda** (cauda esquerda mais longa). Isso significa que há uma quantidade maior de valores **inferiores** à moda. Observe, no gráfico a seguir, que há mais valores inferiores à moda X = 9 (quais sejam, de X = 1 a X = 8) do que superiores (apenas X = 10).



Nessa situação, temos a seguinte relação entre moda, mediana e média:

Distribuição Negativamente Assimétrica: Moda > Mediana > Média.



(CESPE/2015 – Agente do Departamento Penitenciário Nacional)

quantidade diária de incidentes (N)	frequência relativa
0	0,1
1	0,2
2	0,5
3	0,0
4	0,2
total	1

Considerando os dados da tabela mostrada, que apresenta a distribuição populacional da quantidade diária de incidentes (N) em determinada penitenciária, julgue o iten que se segue.

O segundo quartil da distribuição das quantidades diárias de incidentes registradas nessa penitenciária é igual a 2.

Comentários:

A mediana é calculada a partir da função de distribuição acumulada, cuja tabela consta abaixo:

x	P(x)	F(x)
0	0,1	0,1
1	0,2	0,3
2	0,5	0,8
3	0,0	0,8
4	0,2	1

Nesse exemplo, não temos nem um valor de X para o qual o valor de F(x) seja exatamente igual a 0,5. O valor imediatamente superior a 0,5 é de 0,8, o qual está associado ao valor X = 2. Logo, o segundo quartil (ou mediana) é $x_{Med} = 2$.

Gabarito: Certo.

(FCC/2014 – TRT 19ª Região) Seja F(x) a função de distribuição da variável X que representa o número de trabalhadores por domicílio em uma determinada população. Se:

$$F(x) = \begin{cases} 0{,}00 & \text{se } x < 0 \\ 0{,}10 & \text{se } 0 \leq x < 1 \\ 0{,}25 & \text{se } 1 \leq x < 2 \\ 0{,}50 & \text{se } 2 \leq x < 3 \\ 0{,}80 & \text{se } 3 \leq x < 4 \\ 1{,}00 & \text{se } x \geq 4 \end{cases}$$

então, o número médio de trabalhadores por domicílio subtraído do número mediano de trabalhadores por domicílio é igual a

a) 0,15

b) 0,10

c) 0,25

d) -0,15

e) -0,50

Comentários:

A média (ou valor esperado) de trabalhadores por domicílio é dada pela fórmula:

$$E(X) = \sum_i x_i . P(x_i)$$

Para aplicá-la, precisamos dos valores de probabilidade, a serem calculados a partir dos valores da função de distribuição acumulada:

$$P(X = 0) = F(0) = 0{,}10$$

$$P(X = 1) = F(1) - F(0) = 0{,}25 - 0{,}10 = 0{,}15$$

$$P(X = 2) = F(2) - F(1) = 0{,}50 - 0{,}25 = 0{,}25$$

$$P(X = 3) = F(3) - F(2) = 0{,}80 - 0{,}50 = 0{,}30$$

$$P(X = 4) = F(4) - F(3) = 1 - 0{,}8 = 0{,}20$$

Assim, a média é dada por:

$$E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) + 2 \times P(X = 2) + 3 \times P(X = 3) + 4 \times P(X = 4)$$

$$E(X) = 1 \times 0{,}15 + 2 \times 0{,}25 + 3 \times 0{,}30 + 4 \times 0{,}20$$

$$E(X) = 0{,}15 + 0{,}50 + 0{,}90 + 0{,}80 = 2{,}35$$

A mediana é calculada a partir da função de distribuição acumulada. Pelo enunciado, podemos observar que $F(x) = 50\%$ para $x \in [2{,}3)$. Por convenção, temos:

$$x_M = \frac{2 + 3}{2} = 2{,}5$$

A diferença entre a média e a mediana é:

$$E(x) - x_M = 2{,}35 - 2{,}5 = -0{,}15$$

Gabarito: D

VARIÂNCIA E DESVIO PADRÃO

Variância

Nesta seção, veremos medidas de **dispersão** (ou **variabilidade**), que representam o quanto os elementos **desviam** em relação à **média**.

Para calcular o **desvio** de um elemento x em relação à média do conjunto de elementos μ, fazemos:

$$Desvio = x - \mu$$

Entretanto, ao somar os desvios de todos os elementos do conjunto, os **desvios positivos** ($x > \mu$) **anulariam os desvios negativos** ($x < \mu$), de modo que o resultado seria **zero**, tendo vista a própria definição de média.

Nesse sentido, vamos supor o seguinte conjunto de números {1, 1, 1, 1, 3, 7, 7}. A média desse conjunto é:

$$\mu = \frac{1+1+1+1+3+7+7}{7} = \frac{21}{7} = 3$$

A soma dos desvios em relação à média é:

$$Soma\ Desvios = (1-3)+(1-3)+(1-3)+(1-3)+(3-3)+(7-3)+(7-3)$$

$$Soma\ Desvios = -2-2-2-2+0+4+4 = \mathbf{0}$$

Assim, para que possamos **somar** os desvios, precisamos elevá-los ao **quadrado** antes.

$$Desvio\ quadrado = (x-\mu)^2$$

A **variância** é, portanto, definida como a **média do quadrado dos desvios**. Assim, em um conjunto de elementos, somamos o quadrado de todos os desvios e dividimos pela quantidade N de elementos, para obtermos a variância, indicada por σ^2:

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N}$$

Para o nosso exemplo, temos:

$$\sigma^2 = \frac{(1-3)^2+(1-3)^2+(1-3)^2+(1-3)^2+(3-3)^2+(7-3)^2+(7-3)^2}{7}$$

$$\sigma^2 = \frac{(-2)^2+(-2)^2+(-2)^2+(-2)^2+(0)^2+(4)^2+(4)^2}{7} = \frac{4+4+4+4+0+16+16}{7} = \frac{48}{7}$$

Observe que, em relação aos elementos repetidos, podemos **multiplicá-los pela sua frequência**. Em outras palavras, outra forma de obter a variância é somando os produtos dos desvios quadrados multiplicados pela sua frequência relativa:

$$\sigma^2 = (1-3)^2 \times \frac{4}{7} + (3-3)^2 \times \frac{1}{7} + (7-3)^2 \times \frac{2}{7} = (-2)^2 \times \frac{4}{7} + (0)^2 \times \frac{1}{7} + (4)^2 \times \frac{2}{7}$$

$$\sigma^2 = 4 \times \frac{4}{7} + 0 + 16 \times \frac{2}{7} = \frac{16+32}{7} = \frac{48}{7}$$

Genericamente, essa segunda forma de obter a variância pode ser representada pela seguinte fórmula, em que f_r é a frequência relativa:

$$\sigma^2 = \sum (x-\mu)^2 \times f_r$$

Para uma **variável aleatória**, a variância é calculada de maneira similar, porém, em vez da frequência relativa, utilizamos a **probabilidade** para cada valor de $X = x$.

TOME NOTA!

$$\boxed{\boldsymbol{\sigma^2 = \sum (x-\mu)^2 \times P(X=x)}}$$

A variância da variável aleatória X pode ser denotada também por $\boldsymbol{V(X)}$ ou $\boldsymbol{Var(X)}$.

Vamos, então, calcular a variância para o nosso exemplo do dado equilibrado, em que os valores da variável são $\{1, 2, 3, 4, 5, 6\}$ com probabilidade $P(X = x) = \frac{1}{6}$ para todos os elementos e média $\mu = 3{,}5$.

$$\sigma^2 = (1-3{,}5)^2.\frac{1}{6} + (2-3{,}5)^2.\frac{1}{6} + (3-3{,}5)^2.\frac{1}{6} + (4-3{,}5)^2.\frac{1}{6} + (5-3{,}5)^2.\frac{1}{6} + (6-3{,}5)^2.\frac{1}{6}$$

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

$$\sigma^2 = 6{,}25 \times \frac{1}{6} + 2{,}25 \times \frac{1}{6} + 0{,}25 \times \frac{1}{6} + 0{,}25 \times \frac{1}{6} + 2{,}25 \times \frac{1}{6} + 6{,}25 \times \frac{1}{6} = 17{,}5 \times \frac{1}{6} = \frac{35}{12}$$

INDO MAIS FUNDO!

A esperança de uma variável X é a soma dos produtos de cada valor de X pela sua probabilidade:

$$E(X) = \sum x \times P(X = x)$$

Se substituirmos x por $(x - \mu)^2$, obtemos justamente a fórmula da variância:

$$E(X - \mu)^2 = \sum (x - \mu)^2 \times P(X = x)$$

Por isso, dizemos que a variância de uma variável aleatória é definida como:

$$\boldsymbol{\sigma^2 = E(X - \mu)^2}$$

A variância também pode ser calculada da seguinte forma:

TOME NOTA!

$$\boldsymbol{\sigma^2 = E(X^2) - \mu^2}$$

Considerando que $\mu = E(X)$, podemos escrever essa fórmula como:

$$\boldsymbol{\sigma^2 = E(X^2) - [E(X)]^2}$$

O termo $E(X^2)$ representa a esperança dos valores da variável aleatória X, **elevados ao quadrado**, isto é, o produto de x^2 pela sua probabilidade:

$$\boldsymbol{E(X^2) = \sum x^2 \times P(X = x)}$$

Agora, vamos calcular a variância para o exemplo do lançamento do dado, utilizando a segunda fórmula:

$$\sigma^2 = E(X^2) - \mu^2$$

O primeiro termo dessa fórmula é:

$$E(X^2) = \sum x^2 \times P(X = x)$$

$$E(X^2) = (1)^2.\frac{1}{6} + (2)^2.\frac{1}{6} + (3)^2.\frac{1}{6} + (4)^2.\frac{1}{6} + (5)^2.\frac{1}{6} + (6)^2.\frac{1}{6}$$

$$E(X^2) = 1.\frac{1}{6} + 4.\frac{1}{6} + 9.\frac{1}{6} + 16.\frac{1}{6} + 25.\frac{1}{6} + 36.\frac{1}{6} = \frac{91}{6}$$

O segundo termo dessa fórmula (em fração) é:

$$\mu^2 = (3{,}5)^2 = \left(\frac{7}{2}\right)^2 = \frac{49}{4}$$

A variância é dada pela **diferença**:

$$\sigma^2 = \frac{91}{6} - \frac{49}{4} = \frac{182}{12} - \frac{147}{12} = \frac{35}{12}$$

Para essa **segunda** forma de cálculo, podemos utilizar, como apoio, a tabela de distribuição de probabilidade, com os valores de x e $P(X = x)$, acrescentando duas colunas, uma com o valor da variável ao quadrado, x^2, e outra com o produto $x^2.P(X = x)$.

A **soma** da última coluna será o resultado de $\boldsymbol{E(X^2)}$.

$\boldsymbol{x}$	$\boldsymbol{P(X = x)}$	$\boldsymbol{x^2}$	$\boldsymbol{x^2.P(X = x)}$
1	$^1/_6$	1	$^1/_6$
2	$^1/_6$	4	$^4/_6$
3	$^1/_6$	9	$^9/_6$
4	$^1/_6$	16	$^{16}/_6$
5	$^1/_6$	25	$^{25}/_6$
6	$^1/_6$	36	$^{36}/_6$
$\boldsymbol{E(X^2)} =$			$\boldsymbol{^{91}/_6}$

RESUMINDO

Para calcular a variância, seguimos os seguintes passos:

i) Calcular a **média**: $\mu = E(X) = \sum_i x_i . P(x_i)$;

ii) Elevar a **média ao quadrado**: μ^2;

iii) Elevar os valores de **X ao quadrado** e multiplicá-los pela **probabilidade**: $x^2 . P(X = x)$;

iv) **Somar** os resultados do passo iii para calcular $E(X^2) = \sum_i (x_i)^2 . P(x_i)$;

v) Calcular a variância pela **diferença** (iv) – (ii): $\sigma^2 = E(X^2) - \mu^2$.

Atente-se para não esquecer o último passo. Em outras palavras, não pense que o resultado do passo iv, $E(X^2)$, é a variância.

INDO MAIS FUNDO!

Chamamos $E(X^2)$ de **segundo momento** (ou **momento de segunda ordem**) da variável aleatória. Também podemos chamar a **variância** de **segundo momento central** (ou **momento central de segunda ordem**) da variável aleatória.

HORA DE PRATICAR!

(2017 – DPE/PR)

Tabela – Distribuição da variável aleatória X.

X	P (X)
1	0,42
2	0,25
3	0,18
4	0,08
5	0,07

Seja X uma variável aleatória discreta, sua esperança e variância são respectivamente:

a) Esperança = 2,00 e Variância = 2,13.

b) Esperança = 2,13 e Variância = 1,53

d) Esperança = 1,00 e Variância = 1,53

e) Esperança = 2,13 e Variância = 2,53

Comentários:

Vamos utilizar novamente a tabela de distribuição de probabilidade fornecida para calcular os valores de $\mu = E(X) = \sum_i x_i . P(x_i)$ e de $E(X^2) = \sum_i (x_i)^2 . P(x_i)$:

x	P(x)	x.P(x)	x^2	x^2.P(x)
1	0,42	0,42	1	0,42
2	0,25	0,50	4	1,00
3	0,18	0,54	9	1,62
4	0,08	0,32	16	1,28
5	0,07	0,35	25	1,75
Total	**1,00**	**2,13**	**-**	**6,07**

Portanto, temos $\mu = E(X) = 2{,}13$, então, $\mu^2 \cong 4{,}54$; e $E(X^2) = 6{,}07$. Assim, a variância é:

$$\sigma^2 = E(X^2) - \mu^2 \cong 6{,}07 - 4{,}54 = 1{,}53$$

Gabarito: B

(VUNESP/2014 – TJ/PA) Em uma locadora de automóveis a demanda diária é uma variável aleatória com a seguinte distribuição de probabilidades:

Automóveis X_i	0	1	2	3	4	5
Probabilidade	0,10	0,10	0,30	0,30	0,10	0,10

A variância da demanda diária é:

a) 1,85.

b) 1,5.

c) 1,25.

d) 1,0.

e) 0,85

Comentários:

A variância pode ser calculada como:

$$V(X) = E(X^2) - [E(X)]^2$$

Vamos utilizar a tabela para calcular o valor de E(X):

X_i	0	1	2	3	4	5
$P(X_i)$	0,1	0,1	0,3	0,3	0,1	0,1
$X_i.P(X_i)$	0x0,1=0	1x0,1=0,1	2x0,3=0,6	3x0,3=0,9	4x0,1=0,4	5x0,1=0,5

Sabendo que E(X) é a soma de $X_i.P(X_i)$, temos:

$$E(X) = 0 \times 0{,}1 + 1 \times 0{,}1 + 2 \times 0{,}3 + 3 \times 0{,}3 + 4 \times 0{,}1 + 5 \times 0{,}1$$

$$E(X) = 0 + 0{,}1 + 0{,}6 + 0{,}9 + 0{,}4 + 0{,}5 = 2{,}5$$

Logo, o quadrado de E(X) é:

$$[E(X)]^2 = (2{,}5)^2 = 6{,}25$$

Agora, vamos calcular $E(X^2)$:

X_i	0	1	2	3	4	5
$(X_i)^2$	$0^2=0$	$1^2=1$	$2^2=4$	$3^2=9$	$4^2=16$	$5^2=25$
$P(X_i)$	0,1	0,1	0,3	0,3	0,1	0,1
$(X_i)^2.P(X_i)$	0x0,1=0	1x0,1=0,1	4x0,3=1,2	9x0,3=2,7	16x0,1=1,6	25x0,1=2,5

Sabendo que $E(X^2)$ é a soma de $X_i^2.P(X_i)$, temos:

$$E(X^2) = 0^2 \times 0{,}1 + 1^2 \times 0{,}1 + 2^2 \times 0{,}3 + 3^2 \times 0{,}3 + 4^2 \times 0{,}1 + 5^2 \times 0{,}1$$

$$E(X^2) = 0 \times 0{,}1 + 1 \times 0{,}1 + 4 \times 0{,}3 + 9 \times 0{,}3 + 16 \times 0{,}1 + 25 \times 0{,}1$$

$$E(X^2) = 0 + 0{,}1 + 1{,}2 + 2{,}7 + 1{,}6 + 2{,}5 = 8{,}1$$

Assim, a variância é:

$$V(X) = 8{,}1 - 6{,}25 = 1{,}85$$

Gabarito: A

(2019 – IF-PA) Uma variável aleatória discreta Z tem função de probabilidade dada por:

Z	-2	-1	0	1	2
P(Z)	0,3	0,1	0,2	0,3	0,5

Pode-se afirmar que a variância da variável aleatória Z é igual a:

a) 2,64

b) 3,00

c) 3,24

d) 4,64

e) 2,84

Comentários:

Vamos utilizar a tabela de distribuição de probabilidade fornecida para calcular os valores de $\mu = E(Z) = \sum_i z_i.P(z_i)$ e de $E(Z^2) = \sum_i (z_i)^2.P(z_i)$. Atente-se que a tabela abaixo está em um formato diferente daquele que vimos antes (está transposta), para acompanhar a tabela do enunciado.

Z	-2	-1	0	1	2	**Total**
P(Z)	0,3	0,1	0,2	0,3	0,5	**1**
Z.P(Z)	-0,6	-0,1	0	0,3	1	**0,6**
Z^2	4	1	0	1	4	**-**
Z^2.P(Z)	1,2	0,1	0	0,3	2	**3,6**

Portanto, temos $\mu = E(Z) = 0{,}6$, então, $\mu^2 = 0{,}36$; e $E(Z^2) = 3{,}6$. Assim, a variância é:

$$\sigma^2 = E(Z^2) - \mu^2 = 3{,}6 - 0{,}36 = 3{,}24$$

Gabarito: C

(FGV/2022 – SEFAZ/AM) Uma variável aleatória X tem a seguinte função de probabilidade, sendo k uma constante:

x	-2,0	-1,0	0,0	1,0	2
p(x)	0,2	0,1	0,4	0,1	k

A variância de X é igual a:

a) 1,8

b) 2,0

c) 2,2

d) 2,4

e) 2,6

Comentários:

Para calcular a variância, precisamos do valor de k e da esperança. Sabendo que a soma das probabilidades é igual a 1, o valor de k é dado por:

$$0{,}2 + 0{,}1 + 0{,}4 + 0{,}1 + k = 1$$

$$k = 1 - 0{,}8 = 0{,}2$$

Assim, a esperança é:

$$E(X) = (-2) \times 0{,}2 + (-1) \times 0{,}1 + 0 \times 0{,}4 + 1 \times 0{,}1 + 2 \times 0{,}2$$

$$E(X) = -0{,}4 - 0{,}1 + 0 + 0{,}1 + 0{,}4 = 0$$

Agora, precisamos calcular os desvios de cada valor de x em relação à média (que será igual ao próprio valor da variável) e elevar cada desvio ao quadrado. Em seguida, multiplicamos cada quadrado pela respectiva probabilidade e somamos todos os resultados:

$$V(X) = \sum [X - E(X)]^2 \times P(X = x)$$

Esses cálculos constam na tabela a seguir:

x	-2,0	-1,0	0,0	1,0	2
p(x)	0,2	0,1	0,4	0,1	0,2
x - E(X)	-2	-1	0	1	2
$[x - E(X)]^2$	4	1	0	1	4
$[x - E(X)]^2.p(x)$	0,8	0,1	0	0,1	0,8

A variância é a soma dos resultados da última linha:

$$V(X) = 0{,}8 + 0{,}1 + 0 + 0{,}1 + 0{,}8 = 1{,}8$$

Gabarito: A

(2012 – Empresa de Pesquisa Energética) Sejam X e Y variáveis aleatórias independentes. Sabendo-se que: $E(X) = 2$; $E(X^2Y) = 8$; $E(XY^2) = 6$ e $E((XY)^2) = 24$, conclui-se que o valor da variância de Y, Var (Y), é

a) 48

b) 24

c) 10

d) 3

e) 2

Comentários:

Sendo X e Y variáveis independentes, então vale a propriedade multiplicativa da esperança:

$$E(XY) = E(X).E(Y)$$

Sabendo que $E(X) = 2$ e que $E(XY^2) = 6$, então:

$$E(X.Y^2) = E(X).E(Y^2) = 2.E(Y^2) = 6$$

$$E(Y^2) = 3$$

Considerando esse resultado e sabendo que $E((XY)^2) = 24$, então:

$$E((XY)^2) = E(X^2.Y^2) = E(X^2).E(Y^2) = E(X^2).3 = 24$$

$$E(X^2) = 8$$

Considerando esse resultado e sabendo que $E(X^2Y) = 8$, então:

$$E(X^2Y) = E(X^2).E(Y) = 8.E(Y) = 8$$

$$E(Y) = 1$$

Sabendo que $E(Y^2) = 3$ e $E(Y) = 1$, podemos calcular a variância, por:

$$VAR(Y) = E(Y^2) – [E(Y)]^2 = 3 – 1 = 2$$

Gabarito: E

Desvio Padrão

Ao utilizarmos os quadrados dos desvios, perdemos um pouco a noção de grandeza. Se estivermos interessados na altura dos brasileiros, por exemplo, a média adulta masculina seria algo em torno de 173cm e a variância, 400cm^2. Analisando somente esses números, não entendemos muito bem o que 400cm^2 querem dizer.

Por isso, existe o conceito do **desvio padrão**, indicado por $\boldsymbol{\sigma}$, $\boldsymbol{D}(\boldsymbol{X})$ ou $\boldsymbol{DP}(\boldsymbol{X})$, definido como a **raiz quadrada da variância**:

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

Nesse exemplo hipotético, o desvio padrão seria de:

$$\sigma = \sqrt{400} = 20cm$$

Ora, esse resultado é bem mais palatável – ele indica que uma boa parcela da população adulta masculina tem altura entre 153cm e 193cm. Agora, o quanto uma "boa parcela" representa depende de alguns fatores. Então, aguarde cenas dos próximos capítulos!

Para o nosso exemplo da moeda equilibrada, em que calculamos a variância $\sigma^2 = \frac{35}{12}$, o desvio padrão é a raiz quadrada desse valor:

$$\sigma = \sqrt{\frac{35}{12}} \cong 1{,}7$$

Variância e Desvio Padrão Amostrais

Podemos calcular a variância e o desvio padrão a partir de **amostras**, isto é, utilizar os dados obtidos em amostras para **estimar** a variância ou o desvio padrão da **população** de interesse.

Para isso, considerando uma amostra de tamanho n (isto é, com n observações), primeiro calculamos a **média** da amostra, que denotamos por $\bar{x}$:

$$\bar{x} = \frac{\sum x}{n}$$

Para calcular a estimativa da variância, que denotamos por s^2, dividimos a soma do quadrado dos desvios $\sum(x - \bar{x})^2$ por $n - 1$:

$$\boldsymbol{s^2 = \frac{\sum(x-\bar{x})^2}{n-1}}$$

Observe que essa fórmula é bastante similar à variância populacional, com a seguinte diferença: no cálculo da variância populacional, dividimos pelo total da população N e, no cálculo da variância amostral (isto é, para estimar a variância, a partir dos dados da amostra), dividimos por $n - 1$.

Vamos supor que o conjunto que vimos anteriormente, {1, 1, 1, 1, 3, 7, 7}, represente os números observados em uma **amostra**, obtidos a partir de uma população X. Assim, para **estimar a variância** da população, a **partir dessa amostra**, utilizamos a fórmula:

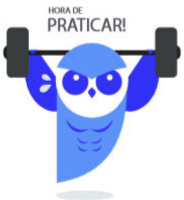
$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

Já calculamos a média desse conjunto: $\bar{x} = 3$. Então a estimativa da variância é:

$$s^2 = \frac{(1-3)^2 + (1-3)^2 + (1-3)^2 + (1-3)^2 + (3-3)^2 + (7-3)^2 + (7-3)^2}{7-1}$$

$$s^2 = \frac{(-2)^2 + (-2)^2 + (-2)^2 + (-2)^2 + (0)^2 + (4)^2 + (4)^2}{6}$$

$$s^2 = \frac{4+4+4+4+0+16+16}{6} = \frac{48}{6} = 8$$



(FCC/2015 – SEFAZ/PI – *Adaptada*) Julgue a seguinte afirmativa:

As amostras I e II dadas abaixo possuem a mesma variância amostral igual a 10.

Amostra I: 1 3 5 7 9 Amostra II: 11 13 15 17 19

Comentários:

Para calcular a variância amostral da Amostra I, primeiro calculamos a média da amostra:

$$\overline{x_1} = \frac{\sum x}{n} = \frac{1+3+5+7+9}{5} = \frac{25}{5} = 5$$

Logo, a variância amostral da Amostra I é dada por:

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1}$$

$${s_1}^2 = \frac{(1-5)^2 + (3-5)^2 + (5-5)^2 + (7-5)^2 + (9-5)^2}{4} = \frac{16+4+0+4+16}{4} = \frac{40}{4} = 10$$

Para calcular a variância amostral da Amostra II, começamos pelo cálculo da média:

$$\overline{x_2} = \frac{\sum x}{n} = \frac{11+13+15+17+19}{5} = \frac{75}{5} = 15$$

A variância amostral da Amostral II é, portanto, dada por:

$$s_2{}^2 = \frac{(11-15)^2 + (13-15)^2 + (15-15)^2 + (17-15)^2 + (19-15)^2}{4} =$$

$$s_2{}^2 = \frac{16+4+0+4+16}{4} = \frac{40}{4} = 10$$

Portanto, as amostras I e II possuem a mesma variância amostral, igual a 10.

Resposta: Certo.

Propriedades

Agora, veremos as propriedades da variância e do desvio padrão, que são aplicáveis tanto a variáveis aleatórias discretas, quanto a variáveis contínuas.

Nos enunciados a seguir, consideramos X e Y variáveis aleatórias e k uma constante real qualquer.

i) $\boldsymbol{V(X+k) = V(X)}$

A variância de uma variável aleatória X, sendo esta somada a uma constante k, se **mantém igual** à variância de X.

Por exemplo, para k = 2 e $V(X) = \frac{35}{12}$, a variância de X + 2 será:

$$V(X+2) = V(X) = \frac{35}{12}$$

INDO MAIS FUNDO!

Vamos entender o porquê disso, com base no exemplo do dado. Sabendo que a média é $\mu = 3{,}5$, vamos replicar o início do cálculo da variância, pela primeira fórmula:

$$V(X) = \sum (x-\mu)^2 \times P(X=x)$$

$$V(X) = (1-3{,}5)^2.\tfrac{1}{6} + (2-3{,}5)^2.\tfrac{1}{6} + (3-3{,}5)^2.\tfrac{1}{6} + (4-3{,}5)^2.\tfrac{1}{6} + (5-3{,}5)^2.\tfrac{1}{6} + (6-3{,}5)^2.\tfrac{1}{6}$$

$$V(X) = (-2{,}5)^2.\tfrac{1}{6} + (-1{,}5)^2.\tfrac{1}{6} + (-0{,}5)^2.\tfrac{1}{6} + (0{,}5)^2.\tfrac{1}{6} + (1{,}5)^2.\tfrac{1}{6} + (2{,}5)^2.\tfrac{1}{6}$$

Agora, vamos supor que Y represente um dado igualmente equilibrado, cujas faces variam de Y = 3 até Y = 8. Ou seja, somamos k = 2 às faces do dado X:

Y = X + 2 = {3, 4, 5, 6, 7, 8}

A média (ou esperança) de Y será:

$$E(Y) = 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} + 7 \times \frac{1}{6} + 8 \times \frac{1}{6} = \frac{33}{6} = 5{,}5$$

E a variância será calculada como:

$$V(Y) = (3 - 5{,}5)^2 . \frac{1}{6} + (4 - 5{,}5)^2 . \frac{1}{6} + (5 - 5{,}5)^2 . \frac{1}{6} + (6 - 5{,}5)^2 . \frac{1}{6} + (7 - 5{,}5)^2 . \frac{1}{6} + (8 - 5{,}5)^2 . \frac{1}{6}$$

$$V(Y) = (-2{,}5)^2 . \frac{1}{6} + (-1{,}5)^2 . \frac{1}{6} + (-0{,}5)^2 . \frac{1}{6} + (0{,}5)^2 . \frac{1}{6} + (1{,}5)^2 . \frac{1}{6} + (2{,}5)^2 . \frac{1}{6}$$

Observe que os desvios $y - E(Y)$ são exatamente os mesmos de $x - E(X)$, e as probabilidades $P(Y = y)$ são as mesmas de $P(X = x)$. Como o cálculo é o mesmo, o resultado, isto é, a variância, será igual.

Isso acontece porque, ao somarmos a constante $k = 2$ aos valores de y, a média também sofre esse mesmo acréscimo (essa é uma propriedade da esperança):

$$E(Y) = E(X + 2) = E(X) + 2$$

Consequentemente, os desvios $y - E(Y)$ são:

$$y - E(Y) = (x + 2) - [E(X) - 2] = x + 2 - E(X) - 2 = x - E(X)$$

Como as probabilidades não são alteradas pela soma da constante, $P(X = x) = P(Y = y)$, então o produto dos desvios de Y pelas respectivas probabilidades são iguais aos de X e, consequentemente, as variâncias são iguais!

Essa propriedade vale também quando **subtraímos** uma constante k (trata-se da mesma propriedade, pois podemos considerar que estamos **somando** $-\boldsymbol{k}$):

$$\boldsymbol{V(X - k) = V(X)}$$

Então, para esse mesmo exemplo de $V(X) = \frac{35}{12}$, a variância de X – 4 será:

$$V(X - 4) = V(X) = \frac{35}{12}$$

Por ser a raiz quadrada da variância, que não se altera, o desvio padrão também permanece o mesmo:

$$\boldsymbol{D(X + k) = D(X)}$$

$$\boldsymbol{D(X - k) = D(X)}$$

ii) $\boldsymbol{V(k.X) = k^2.V(X)}$

A variância de uma variável aleatória X, sendo esta multiplicada por uma constante k, é igual à variância de X multiplicada pelo quadrado de k.

Ou seja, para k = 2 e $V(X) = \frac{35}{12}$, a variância de 2.X será:

$$V(2.X) = 2^2.V(X) = 4.\frac{35}{12} = \frac{35}{3}$$

INDO MAIS FUNDO!

Vamos, novamente, entender o porquê disso, mas agora vamos supor que Y represente os valores de X multiplicados por 3:

Y = 3.X

Pelas propriedades da esperança, a média de Y será multiplicada por 3:

$$E(Y) = E(3.X) = 3 \times E(X)$$

E os desvios serão multiplicados por 3:

$$y - E(Y) = (3.x) - [3.E(X)] = 3.[x - E(X)]$$

Ao elevarmos esses desvios ao quadrado, os resultados serão multiplicados pelo quadrado de 3, $(3)^{\mathbf{2}}$:

$$\left(y - E(Y)\right)^2 = (3.[x - E(X)])^2 = (3)^{\mathbf{2}}.([x - E(X)])^2$$

Como as probabilidade probabilidades $P(Y = y)$ são as mesmas de $P(X = x)$, então, ao multiplicarmos esses desvios pelas respectivas probabilidades obtemos o mesmo resultado multiplicado por $(3)^{\mathbf{2}}$.

$$\left(y - E(Y)\right)^2 \times P(Y = y) = (3)^{\mathbf{2}}.([x - E(X)])^2 \times P(X = x)$$

Quando somamos os produtos para todos os valores de y, obtemos a mesma variância, multiplicada por $(3)^2$:

$$\sum\left(y - E(Y)\right)^2 \times P(Y = y) = \sum(3)^{\mathbf{2}}.([x - E(X)])^2 \times P(X = x)$$

$$V(Y) = (3)^{\mathbf{2}}.V(X)$$

Essa propriedade também vale para a divisão por uma constante k (podemos considerar que estamos multiplicando por $\frac{1}{k}$):

$$V\left(\frac{X}{k}\right) = \frac{V(X)}{k^2}$$

Note que não importa se k é positivo ou negativo (seja para a multiplicação por k, seja para a divisão por k), pois o seu quadrado será sempre positivo. Por exemplo, para $Y = -\frac{X}{2}$ e $V(X) = \frac{35}{12}$, o resultado também será igual a:

$$V(Y) = V\left(\frac{X}{-2}\right) = \frac{V(X)}{(-2)^2} = \frac{1}{4}.\frac{35}{12} = \frac{35}{3}$$

Como o desvio padrão é a raiz quadrada da variância, o desvio padrão do produto $k.X$, é:

$$D(k.X) = \sqrt{V(k.X)} = \sqrt{k^2.V(X)} = |k|.D(X)$$

ESCLARECENDO!

Como a **raiz** de um número é **sempre um número positivo**, então a raiz de k^2 é o módulo de k, denotado por $|k|$:

$$\sqrt{k^2} = |k|$$

O módulo de k, $|k|$, é uma "**versão positiva**" do número k, ou seja:

$$|k| = \begin{Bmatrix} k, se\ k \geq 0 \\ -k, se\ k < 0 \end{Bmatrix}$$

Por exemplo, se $k = 3$, então $|k| = 3$ e se $k = -3$, então $|k| = 3$.

Portanto:

$$\boldsymbol{D(k.X) = k.D(X)}, \textbf{ para } \boldsymbol{k \geq 0}$$

$$\boldsymbol{D(k.X) = -k.D(X)}, \textbf{ para } \boldsymbol{k < 0}$$

$$\boldsymbol{D\left(\frac{X}{k}\right) = \frac{D(X)}{k}}, \textbf{ para } \boldsymbol{k \geq 0}$$

$$\boldsymbol{D\left(\frac{X}{k}\right) = -\frac{D(X)}{k}}, \textbf{ para } \boldsymbol{k < 0}$$

Por exemplo, para Y = 3.X, o desvio padrão será:

D(Y) = D(3.X) = 3.D(X)

E para $Y = -\frac{X}{2}$, o desvio padrão será:

$$D(Y) = D\left(\frac{X}{-2}\right) = \frac{D(X)}{2}$$

iii) $\boldsymbol{V(k) = 0}$

A variância de uma **constante** qualquer é **zero**. Por exemplo, a variância da constante $k = 3$ é:

$$\boldsymbol{V(3) = 0}$$

Pelas propriedades da esperança, a média de uma constante k:

$$\mu = E(k) = \boldsymbol{k}$$

Assim, o desvio $k - \mu$ será:

$$desvio = k - \mu = \boldsymbol{k} - \boldsymbol{k} = \boldsymbol{0}$$

Portanto, a variância será **0**.

Consequentemente, a variância e o desvio padrão de uma constante são iguais a zero:

$$\boldsymbol{D(k) = 0}$$

Por exemplo, $D(3) = 0$.

iv) **Se X e Y são independentes, então $\boldsymbol{V(X+Y) = V(X) + V(Y)}$**

Somente se X e Y forem variáveis aleatórias **independentes**, poderemos concluir que a variância da soma das variáveis é igual à **soma das variâncias** (propriedade aditiva).

EXEMPLIFICANDO

Vamos supor que **X** represente os resultados do lançamento de um **dado normal** (equilibrado, com faces de 1 a 6) e que **Y** represente os resultados do lançamento do dado equilibrado com **faces de 3 a 8**. Assim, se lançarmos **os dois dados** ao mesmo tempo, qual será a variância da distribuição dos resultados?

Já calculamos as variâncias de X e Y em exercícios anteriores:

$$V(X) = \frac{35}{12}$$

$$V(Y) = V(X) = \frac{35}{12}$$

Por estarmos lançando os dois dados, teremos a distribuição de X + Y, sendo X e Y variáveis independentes, pois um lançamento não influencia no outro. Logo, a variância da distribuição X + Y, é:

$$V(X) + V(Y) = \frac{35}{12} + \frac{35}{12} = 2.\frac{35}{12} = \frac{35}{6}$$

Além disso, se X e Y forem **independentes**, então a **variância** da **diferença** X – Y também é a **soma** das variâncias:

$$\boldsymbol{V(X - Y) = V(X) + V(Y)}$$

Para variáveis independentes, temos $V(X+Y) = V(X) + V(Y)$, porém o contrário não é necessariamente verdadeiro. Ou seja, é **possível** verificar a identidade $V(X+Y) = V(X) + V(Y)$ e as variáveis **não** serem independentes.

RESUMINDO

i)	$\boldsymbol{V(X \pm k) = V(X)}$
ii)	$\boldsymbol{V(k.X) = k^2.V(X)}$
iii)	$\boldsymbol{V(k) = 0}$
iv)	Se X e Y forem **independentes**, então $\boldsymbol{V(X \pm Y) = V(X) + V(Y)}$

(2016 – Instituto Federal de Educação/BA – *Adaptada*) Sendo X uma variável aleatória, com variância σ^2, então a variância da função Y = a + bX, com a e b ∈ R, é

a) $V(Y) = b^2$

b) $V(Y) = a + b$

c) $V(Y) = \sigma^2$

d) $V(Y) = b^2\sigma^2$

e) $V(Y) = a^2 + b^2$

Comentários:

Pelas propriedades da variância, temos:

$$V(Y) = V(a + bx) = b^2.V(X)$$

Como a variância de X é $V(X) = \sigma^2$, então a variância de Y é:

$$V(Y) = b^2\sigma^2$$

Gabarito: D.

(FGV/2017 – IBGE – *Adaptada*) Para o caso de variáveis aleatórias quaisquer, existem diversas propriedades que se aplicam diretamente à esperança matemática e ao momento central de segunda ordem. Dentre essas propriedades está:

a) $Var(X) > E(X^2)$

b) $Var(X \pm Y) = Var(X) \pm Var(Y)$

c) DP(a) = 0, sendo a uma constante qualquer

d) Var(a.X) = a.Var(X), sendo a uma constante positiva

e) DP(a.X) = a.DP(X), sendo a uma constante qualquer

Comentários:

Em relação à alternativa A, podemos calcular a variância como:

$$Var(X) = E(X^2) - [E(X)]^2$$

Como $[E(X)]^2 > 0$ para qualquer variável X, então, temos:

$$Var(X) < E(X^2)$$

Portanto, a alternativa A está incorreta.

Em relação à alternativa B, se X e Y forem **independentes**, então podemos afirmar que:

$$Var(X \pm Y) = Var(X) + Var(Y)$$

Logo, a alternativa B está incorreta por 2 motivos:

i) Não pode considerar a propriedade aditiva da variância, $Var(X \pm Y) = Var(X) + Var(Y)$, pois o enunciado **não** afirmou que X e Y são independentes.

ii) Ainda que X e Y fossem independentes, a variância da diferença de X e Y seria igual à soma das variâncias: $Var(X - Y) = Var(X) + Var(Y)$

Em relação à alternativa C, sabemos que a variância e o desvio padrão de uma constante "a" **qualquer** são iguais a zero.

$$DP(a) = 0$$

Portanto, a alternativa C está correta.

Em relação à alternativa D, sabemos que para uma constante "a" **qualquer**:

$$Var(a.X) = a^2.Var(X)$$

Portanto, a alternativa D está incorreta.

Em relação à alternativa E, sabemos que para uma constante "a" **qualquer**, temos:

$$DP(a.X) = |a|.DP(X)$$

Assim, sendo a uma constante **positiva** então:

$$DP(a.X) = a.DP(X)$$

Sendo a uma constante **negativa** então:

$$DP(a.X) = -a.DP(X)$$

Portanto, temos equações **distintas** para constantes positivas e negativas, logo a alternativa E está incorreta.

Resposta: C.

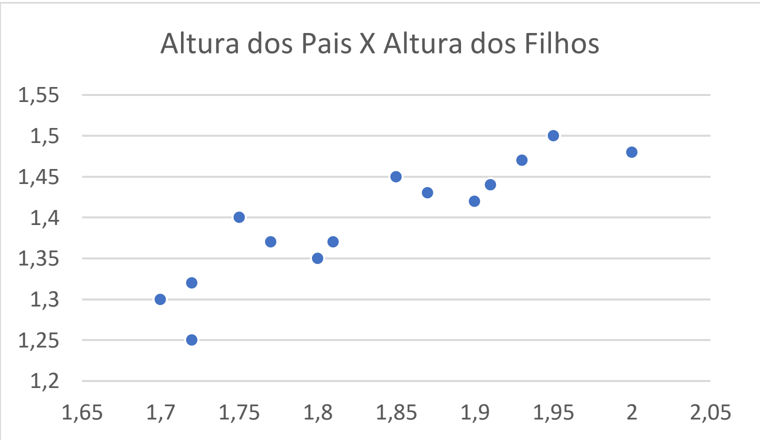
Covariância e Correlação

As medidas que estudaremos nesta seção representam a **relação** entre **duas variáveis aleatórias**.

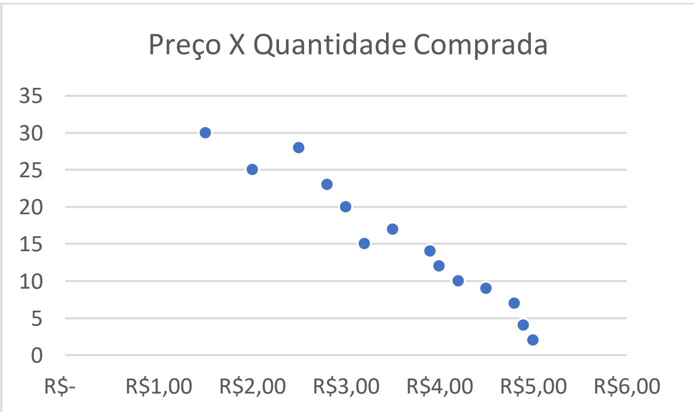
Existem diversos exemplos de variáveis **independentes**, como os resultados de dois lançamentos de dados ou moedas. Mas também há variáveis **relacionadas**, em que o resultado de uma influencia de alguma forma o resultado de outra. Por exemplo, a **altura de uma criança** é dependente da **altura de seus pais**; o **volume de água** em uma caixa d'água se relaciona diretamente com o seu **peso**; a **demanda** de certo produto varia de acordo com o seu **preço**.

Há variáveis **fortemente** relacionadas, como o volume e o peso de água, e outras **nem tanto**, como a altura dos pais e a altura dos filhos. Também existem variáveis que se relacionam em um **mesmo sentido** (quanto mais altos são os pais, mais altos os filhos tendem a ser) e variáveis que se relacionam em **sentidos opostos** (quanto maior o preço, menor a demanda).

No gráfico abaixo, ilustramos um exemplo hipotético de duas variáveis que se relacionam no **mesmo sentido**, como a altura dos pais e a altura dos filhos.



No gráfico a seguir, ilustramos um exemplo hipotético de duas variáveis que se relacionam em **sentidos opostos**, como o preço de um produto e a quantidade adquirida.



A covariância e a correlação caracterizam tanto a **força** da relação entre duas variáveis, quanto a sua **orientação** (se variam no **mesmo sentido** ou em **sentidos opostos**).

A covariância entre duas variáveis aleatórias X e Y, representada por $Cov(X,Y)$, é, por **definição**:

$$Cov(X,Y) = E[(X - \mu_X).(Y - \mu_Y)]$$

Nessa expressão, μ_X corresponde à média (esperança) da variável X, $\mu_X = E(X)$, e μ_Y corresponde à média de Y, $\mu_Y = E(Y)$. Essa expressão equivale à seguinte:

TOME NOTA!

$$\boxed{\boldsymbol{Cov(X,Y) = E(X.Y) - E(X).E(Y)}}$$

Nessa fórmula, $E(X.Y)$ corresponde ao seguinte:

$$E(X.Y) = \sum x.y.p(x,y)$$

Ou seja, multiplicamos os possíveis valores das variáveis pelas probabilidades correspondentes. Se os valores forem **igualmente prováveis**, como em uma amostra, podemos calcular $E(X.Y)$ como:

$$E(X.Y) = \frac{\sum x.y}{N}$$

Por exemplo, vamos considerar uma parte dos dados hipotéticos do gráfico Preço X Quantidade Comprada, conforme indicado na tabela abaixo.

Para calcular a covariância, podemos criar uma **nova coluna** com o **produto** das duas variáveis, o que permitirá calcular E(X.Y).

	X: Preço	**Y: Qtdade**	**X.Y**
i	1,50	30	45
ii	2,00	25	50
iii	3,00	20	60
iv	5,00	2	10
Total	**11,50**	**77**	**165**

Como esses valores são igualmente prováveis, o valor de E(X.Y) pode ser calculado como:

$$E(X.Y) = \frac{\sum x.y}{N} = \frac{165}{4} = 41{,}25$$

Agora, calculamos E(X) e E(Y), isto é, a média de X e de Y:

$$E(X) = \frac{\sum x}{N} = \frac{11,50}{4} = 2,875$$

$$E(Y) = \frac{\sum y}{N} = \frac{77}{4} = 19,25$$

A covariância será, portanto:

$$Cov(X,Y) = 41,25 - 2,875 \times 19,25 \cong -14$$

Nesse caso, obtivemos uma **covariância negativa**. Isso ocorreu porque as variáveis se relacionam em **sentidos opostos** (**relação negativa**), isto é, quando uma aumenta, a outra diminui, em média.

Quando a **covariância** das variáveis é **positiva**, elas variam **no mesmo sentido** (**relação positiva**), isto é, quando uma aumenta, a outra também aumenta, em média.

RESUMINDO

Para calcular a **covariância**, podemos seguir os seguintes passos:

i) **Multiplicar** os valores de X e Y;
ii) **Somar** os produtos X.Y e **dividir por N** para obter $E(X.Y) = \frac{\sum x.y}{N}$;
iii) Calcular as **médias** $E(X) = \frac{\sum x}{N}$ e $E(Y) = \frac{\sum y}{N}$ e multiplicá-las;
iv) **Subtrair** o resultado de ii pelo resultado de iii para obter a covariância.

Entretanto, a **força** da relação entre duas variáveis é **difícil** de interpretar a partir da covariância. Em relação ao nosso exemplo, uma covariância de -14 indica uma forte relação negativa ou uma fraca relação?

Para isso, há o conceito de **correlação** (ou **coeficiente de correlação**), indicado por $\boldsymbol{\rho}$, em que **dividimos** a **covariância** pelo **desvio padrão** de ambas as variáveis.

TOME NOTA!

$$\boldsymbol{\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y}}$$

Para calcular o desvio padrão (populacional) para as variáveis do nosso exemplo, vamos utilizar a seguinte fórmula da variância:

$$V(X) = E(X^2) - [E(X)]^2$$

Para isso, vamos criar uma coluna com os valores de X^2 e Y^2:

	X: Preço	**Y: Qtdade**	**X^2**	**Y^2**
i	1,50	30	2,25	900
ii	2,00	25	4	625
iii	3,00	20	9	400
iv	5,00	2	25	4
Total	**11,50**	**77**	**40,25**	**1929**

Os valores de $E(X^2)$ e $E(Y^2)$ são, portanto:

$$E(X^2) = \frac{\sum x^2}{N} = \frac{40,25}{4} \cong 10,06$$

$$E(Y^2) = \frac{\sum y^2}{N} = \frac{1929}{4} = 482,25$$

Sabendo que $E(X^2) = 10,06$ e que $E(X) = 2,875$, então a variância de X é:

$$V(X) = E(X^2) - [E(X)]^2 \cong 10,06 - [2,875]^2 \cong 10,06 - 8,26 \cong 1,80$$

E o desvio padrão de X é:

$$\sigma_X \cong \sqrt{1,80} \cong 1,34$$

Em relação a Y, sabendo que $E(Y^2) = 482,25$ e que $E(Y) = 19,25$, então a variância e desvio padrão são:

$$V(Y) = E(Y^2) - [E(Y)]^2 = 482,25 - [19,25]^2 \cong 482,25 - 370,56 = 111,69$$

$$\sigma_Y \cong \sqrt{111,69} \cong 10,57$$

Portanto, o coeficiente de correlação para o nosso exemplo é:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y} \cong \frac{-14}{1,34 \times 10,57} \cong -0,99$$

Com base nesse valor, podemos concluir que a relação negativa entre as variáveis é **muito forte**, pois o valor do coeficiente de correlação é próximo de -1.

A fórmula do coeficiente de correlação também pode ser representada como:

$$\rho(X,Y) = \frac{\sum x.y - n.\bar{x}.\bar{y}}{\sqrt{\sum x^2 - n.\bar{x}^2} \times \sqrt{\sum y^2 - n.\bar{y}^2}}$$

Em que o numerador é igual à covariância multiplicada por n e o denominador é igual ao produto dos desvios padrão, também multiplicado por n.

Vamos verificar isso! A covariância pode ser representada como:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y) = \frac{\sum x.y}{n} - \bar{x}.\bar{y} = \frac{1}{n}\left(\sum x.y - n.\bar{x}.\bar{y}\right)$$

E os desvios padrão são a raiz quadrada da variância:

$$\sigma_X = \sqrt{V(X)} = \sqrt{E(X^2) - [E(X)]^2} = \sqrt{\frac{\sum x^2}{n} - \bar{x}} = \sqrt{\frac{1}{n}\left(\sum x^2 - n.\bar{x}\right)}$$

$$\sigma_Y = \sqrt{V(Y)} = \sqrt{E(Y^2) - [E(Y)]^2} = \sqrt{\frac{\sum y^2}{n} - \bar{y}} = \sqrt{\frac{1}{n}\left(\sum y^2 - n.\bar{y}\right)}$$

Dividindo a covariância pelo produto dos desvios padrão, obtemos a fórmula acima!

O coeficiente de correlação mede a **força** e a **orientação** com que duas variáveis se relacionam **linearmente**, podendo assumir valores no intervalo **[-1,1]**.

Assim, como para a covariância, **valores positivos** do coeficiente de correlação indicam uma relação entre as variáveis **no mesmo sentido** (**relação positiva**) e **valores negativos** indicam relação **em sentidos opostos** (**relação negativa**).

Além disso, quando $\boldsymbol{\rho = 1}$, há uma **correlação linear perfeita positiva**, ou seja, as variáveis apresentam uma relação linear entre si da forma $Y = aX + b$, sendo a e b reais e $a > 0$. É o caso do volume de água e do peso da caixa d'água.

Quando $\boldsymbol{\rho = -1}$, há uma **correlação linear perfeita negativa**, ou seja, as variáveis apresentam uma relação linear entre si da forma $Y = aX + b$, sendo a e b reais e $a < 0$. Um exemplo dessa relação seria o peso da caixa d'água e o seu espaço disponível.

Vejamos agora qual valor a covariância assume quando as variáveis são **independentes**. Sendo X e Y variáveis independentes, sabemos que:

$$E(X.Y) = E(X).E(Y)$$

Portanto, a **covariância** de duas variáveis **independentes** é:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y) = E(X).E(Y) - E(X).E(Y) = 0$$

Consequentemente, o valor do coeficiente de correlação para variáveis **independentes** também é $\boldsymbol{\rho = 0}$.

Porém, é possível ter $Cov = 0, \rho = 0$ e as variáveis **não** serem independentes.

Existe uma **exceção** para essa regra!

Para **variáveis binárias**, isto é, que assumem apenas 2 valores, a covariância nula **implica** na independência dessas variáveis! Em outras palavras, se a covariância entre 2 variáveis binárias for nula, podemos garantir que essas variáveis são **independentes**.

$Cov(X,Y) = E(X.Y) - E(X).E(Y)$

$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y}, \rho \in [-1,1]$

Variáveis X e Y Independentes → $Cov(X,Y) = 0, \rho(X,Y) = 0$

$Cov(X,Y) > 0, \rho(X,Y) > 0 \leftrightarrow$ Relação positiva (X e Y variam no mesmo sentido)

$Cov(X,Y) < 0, \rho(X,Y) < 0 \leftrightarrow$ Relação negativa (X e Y variam em sentidos opostos)

$\rho(X,Y) = 1 \leftrightarrow$ relação linear perfeita positiva $\leftrightarrow Y = aX + b, a > 0$

$\rho(X,Y) = -1 \leftrightarrow$ relação linear perfeita negativa $\leftrightarrow Y = aX + b, a < 0$

(2018 – UFRGS) A análise de ____________ permite estudar a relação entre dois conjuntos de valores e quantificar o quanto um está relacionado com o outro, no sentido de determinar a intensidade e a direção dessa relação. Isto é, essa análise indica se, e com que intensidade, os valores de uma variável aumentam ou diminuem enquanto os valores da outra variável aumentam ou diminuem.

Assinale a alternativa que completa corretamente a lacuna do texto acima.

a) correlação

b) dispersão

c) classificação

d) agrupamento

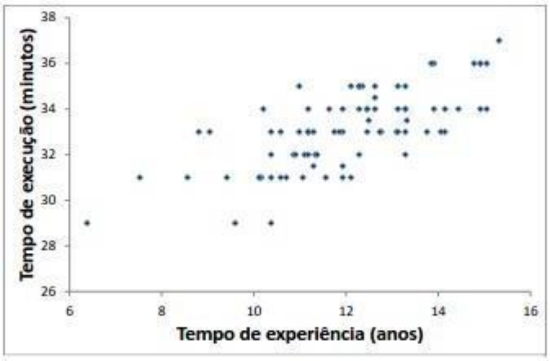
e) regressão

Comentários:

O conceito que estuda a relação entre duas variáveis, representando tanto a força dessa relação quanto o seu sentido, é a correlação.

Gabarito: A

(2015 – PC/GO) Com o intuito de avaliar possíveis correlações entre variáveis, um gráfico de dispersão pode ser um aliado na tomada de decisão. Esse gráfico, elaborado no eixo cartesiano, plota resultados das variáveis estudadas a fim de representá-las conjuntamente. Sejam x e y variáveis referentes a "tempo de experiência" e "tempo de execução de tarefa", respectivamente, e analisando o gráfico de dispersão apresentado, assinale a alternativa correta.



a) É observada uma correlação positiva perfeita entre as variáveis.

b) É observada uma correlação positiva entre as variáveis.

c) É observada uma correlação nula entre as variáveis.

d) É observada uma correlação negativa entre as variáveis.

e) É observada uma correlação negativa perfeita entre as variáveis..

Comentários:

Pelo gráfico, observamos que as variáveis se relacionam em um mesmo sentido, portanto a correlação é **positiva**. Porém, essa relação não é perfeitamente linear, por isso a correlação não é perfeita.

Gabarito: B

(FCC/2015 – SEFAZ/PI – *Adaptada*) Julgue as seguintes afirmações:

I – Se r é o coeficiente de correlação linear entre duas variáveis, então $-1 \leq r \leq 1$.

II – Se duas variáveis X e Y apresentam correlação linear inversa, o coeficiente de correlação linear entre elas será um número negativo menor do que −1.

Comentários:

Em relação à afirmação I, o coeficiente de correlação varia entre [-1,1]. Portanto, a afirmação I está correta.

Em relação à afirmação II, se X e Y se relacionam de forma **inversa**, então o coeficiente de correlação é **negativo**. Porém, como o menor valor para o coeficiente é −1, o coeficiente será um valor negativo maior ou igual a −1, não menor do que −1. Portanto, a afirmação II está incorreta.

Resposta: I – Certo; II – Errado.

(CESPE/2016 – TCE/PR) Se satisfação no trabalho e saúde no trabalho forem indicadores com variâncias populacionais iguais a 8 e 2, respectivamente, e se a covariância populacional entre esses indicadores for igual a 3, então a correlação populacional entre satisfação no trabalho e saúde no trabalho será igual a

a) 0,8125.

b) 1.

c) 0,1875.

d) 0,30.

e) 0,75.

Comentários:

Sabendo que V(X) = 8, V(Y) = 2 e Cov(X,Y) = 3, então a correlação é dada por:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X . \sigma_Y} = \frac{3}{\sqrt{8}.\sqrt{2}} = \frac{3}{4} = 0{,}75$$

Gabarito: E

(2018 – FUNPAPA) Um pesquisador suspeita que existe uma correlação entre o número de promessas que um candidato político faz e o número de promessas que são cumpridas uma vez que o candidato é eleito. Ele acompanha vários políticos proeminentes e registra as promessas feitas (X) e as promessas mantidas (Y). Utilizando os seguintes dados sumarizados, calcule o coeficiente de correlação entre as promessas feitas e as promessas mantidas e assinale a alternativa correta.

$$\sum_{i=1}^{7} x_i = 280, \quad \sum_{i=1}^{7} y_i = 28, \quad \sum_{i=1}^{7} x_i y_i = 940, \quad \sum_{i=1}^{7} x_i^2 = 12400 \quad \text{e} \quad \sum_{i=1}^{7} y_i^2 = 140.$$

a) O coeficiente de correlação entre as promessas feitas e as promessas mantidas indicam uma correlação forte e positiva.

b) O coeficiente de correlação entre as promessas feitas e as promessas mantidas indicam uma correlação fraca e negativa.

c) O coeficiente de correlação entre as promessas feitas e as promessas mantidas indicam uma correlação forte e negativa.

d) O coeficiente de correlação entre as promessas feitas e as promessas mantidas indicam uma correlação fraca e positiva.

e) O coeficiente de correlação entre as promessas feitas e as promessas mantidas indicam uma correlação $r \approx 0{,}5$.

Comentários:

O coeficiente de correlação é calculado por:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X . \sigma_Y}$$

A covariância pode ser calculada por:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

O enunciado informa que $\sum x.y = 940$ e $n = 7$, logo:

$$E(X.Y) = \frac{\sum x.y}{n} = \frac{940}{7}$$

O valor de E(X) pode ser calculado a partir da informação de que $\sum x = 280$, logo:

$$E(X) = \frac{\sum x}{n} = \frac{280}{7}$$

O valor de E(Y) pode ser calculado a partir da informação de que $\sum y = 28$, logo:

$$E(Y) = \frac{\sum y}{n} = \frac{28}{7} = 4$$

Assim, o valor da covariância é:

$$Cov(X,Y) = \frac{940}{7} - \frac{280}{7} \times 4 = \frac{940}{7} - \frac{1120}{7} = -\frac{180}{7} \cong 25{,}7$$

Para calcular o coeficiente de correlação, vamos primeiro calcular a variância de X utilizando a seguinte fórmula:

$$V(X) = E(X^2) - [E(X)]^2$$

O enunciado informa que $\sum x^2 = 12400$, logo:

$$E(X^2) = \frac{\sum x^2}{n} = \frac{12400}{7}$$

Sabendo que $E(X) = \frac{280}{7} = 40$, então a variância de X é:

$$V(X) = \frac{12400}{7} - 40^2 = \frac{12400}{7} - \frac{11200}{7} = \frac{1200}{7}$$

E o desvio padrão de X é:

$$\sigma_X = \sqrt{\frac{1200}{7}} = 10\sqrt{\frac{12}{7}} \cong 13{,}1$$

O enunciado informa que $\sum y^2 = 140$, logo:

$$E(Y^2) = \frac{\sum y^2}{n} = \frac{140}{7} = 20$$

Sabendo que $E(Y) = 4$, então a variância de Y é:

$$V(Y) = 20 - 4^2 = 20 - 16 = 4$$

E o desvio padrão de Y é:

$$\sigma_X = \sqrt{4} = 2$$

Assim, o coeficiente de correlação é:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y} \cong \frac{-25{,}7}{13{,}1 \times 2} \cong -0{,}98$$

Como -0,98 é muito próximo de -1, há uma correlação forte e negativa.

Gabarito: C

Propriedades

Veremos agora propriedades da covariância e da correlação, que valem tanto para **variáveis discretas**, quanto para **variáveis contínuas**. Nesta seção, deduziremos algumas propriedades para que você possa escolher se prefere deduzi-las ou memorizá-las.

A seguir, consideramos X, Y e Z variáveis aleatórias e k uma constante real qualquer.

i) $\boldsymbol{Cov(X,Y) = Cov(Y,X)}$

A covariância é considerada uma medida **simétrica**, pois não importa qual é a variável que aparece primeiro.

De fato, a fórmula da covariância é composta por **produtos** e, por isso, a ordem das variáveis é **indiferente** (a ordem dos fatores não altera o produto):

$$Cov(X,Y) = E(X.Y) - E(X).E(Y) = E(Y.X) - E(Y).E(X) = Cov(Y,X)$$

Por exemplo, se a covariância entre X e Y for Cov(X,Y) = 6, então a covariância entre Y e X também será Cov(Y,X) = 6.

Pelo mesmo motivo, o coeficiente de correlação também é **simétrico**:

$$\boldsymbol{\rho(X,Y) = \rho(Y,X)}$$

Afinal, ele é a razão entre a covariância e o **produto** dos desvios padrão:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y} = \frac{Cov(Y,X)}{\sigma_Y.\sigma_X} = \rho(Y,X)$$

ii) $\boldsymbol{Cov(X,X) = V(X)}$

A covariância de **uma mesma variável** é igual à sua **variância**.

Por exemplo, sendo X uma variável aleatória com variância V(X) = 4, então a covariância dessa variável com ela mesma é igual à própria variância: Cov(X,X) = 4.

ESCLARECENDO!

Podemos obter esse resultado, pela fórmula da covariância:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

Assim, o valor de $Cov(X,X)$ é:

$$Cov(X,X) = E(X.X) - E(X).E(X) = E(X^2) - [E(X)]^2$$

Essa é **exatamente** a fórmula da variância!

Dessa forma, o coeficiente de correlação de uma mesma variável é:

$$\rho(X,X) = \frac{Cov(X,X)}{\sigma_X.\sigma_X} = \frac{Var(X)}{{\sigma_X}^2} = 1$$

$$\boldsymbol{\rho(X,X) = 1}$$

Ou seja, não importa qual é a variável, o seu coeficiente de correlação com ela mesma é igual a **1**.

iii) $\boldsymbol{Cov(k,X) = 0}$

A covariância de uma **constante** e uma variável é igual a **zero**.

Ou seja, a covariância de uma variável X com uma constante k = 5, por exemplo, é Cov(X,5) = 0. Essa propriedade vale para qualquer variável X e qualquer constante k.

ESCLARECENDO!

Vejamos o porquê desse resultado. Pela fórmula da covariância, temos:

$$Cov(k, X) = E(k.X) - E(k).E(X)$$

Sabemos que $E(k.X) = k.E(X)$ e $E(k) = k$. Substituindo esses resultados, temos:

$$Cov(k, X) = k.E(X) - k.E(X) = 0$$

Como a covariância $Cov(k, X) = 0$, então a correlação também é igual a 0:

$$\boldsymbol{\rho(k, X) = 0}$$

iv) $\boldsymbol{Cov(X \pm a, Y \pm b) = Cov(X, Y)}$

A covariância não se altera quando somamos ou subtraímos constantes às variáveis. Por exemplo, sendo Cov(X,Y) = 6, então a covariância entre a variável X + 5 e a variável Y – 4 será a igual:

Cov(X + 5, Y – 4) = Cov(X,Y) = 6

INDO MAIS FUNDO!

Podemos verificar essa propriedade, pela fórmula de covariância:

$$Cov(X + a, Y + b) = E[(X + a).(Y + b)] - E(X + a).E(Y + b)$$

Aplicando a distributiva na primeira expressão, temos:

$$Cov(X + a, Y + b) = E(X.Y + b.X + a.Y + a.b) - E(X + a).E(Y + b)$$

Pelas propriedades da esperança, temos:

$$Cov(X + a, Y + b) = E(X.Y) + b.E(X) + a.E(Y) + a.b. -[E(X) + a].[E(Y) + b]$$

Aplicando a distributiva no segundo termo:

$$= E(X.Y) + b.E(X) + a.E(Y) + a.b. - [E(X).E(Y) + b.E(X) + a.E(Y) + a.b]$$

$$= E(X.Y) + b.E(X) + a.E(Y) + a.b. - E(X).E(Y) - b.E(X) - a.E(Y) - a.b$$

$$Cov(X + a, Y + b) = E(X.Y) - E(X).E(Y)$$

Essa é justamente a fórmula da covariância $Cov(X, Y)$!

Dessa forma, o coeficiente de correlação também não se altera quando somamos ou subtraímos constantes às variáveis:

$$\boldsymbol{\rho(X \pm a, Y \pm b) = \rho(X, Y)}$$

v) $\boldsymbol{Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z)}$

A covariância da soma de variáveis aleatórias X + Y e uma outra variável Z é igual à **soma** da covariância entre X e Z com a covariância entre Y e Z.

Por exemplo, vamos supor que a covariância entre as variáveis X e Z seja Cov(X,Z) = 1 e que a covariância entre as variáveis Y e Z seja Cov(Y,Z) = 2. Supondo que a variável S represente a soma S = X + Y, então podemos calcular a covariância entre S e Z:

Cov(S,Z) = Cov(X + Y, Z) = Cov(X,Z) + Cov(Y,Z) = 1 + 2 = 3

INDO MAIS FUNDO!

Novamente, podemos verificar essa identidade, pela fórmula de covariância:

$$Cov(X + Y, Z) = E[(X + Y).Z] - E(X + Y).E(Z)$$

Aplicando a distributiva na primeira expressão, temos:

$$Cov(X + Y, Z) = E(X.Z + Y.Z) - E(X + Y).E(Z)$$

Pela propriedade aditiva da esperança, temos:

$$Cov(X + Y, Z) = E(X.Z) + E(Y.Z) - [E(X) + E(Y)].E(Z)$$

Aplicando a distributiva no segundo termo:

$$Cov(X+Y,Z) = E(X.Z) + E(Y.Z) - [E(X).E(Z) + E(Y).E(Z)]$$

$$Cov(X+Y,Z) = E(X.Z) + E(Y.Z) - E(X).E(Z) - E(Y).E(Z)$$

Reorganizando esses termos:

$$Cov(X+Y,Z) = E(X.Z) - E(X).E(Z) + E(Y.Z) - E(Y).E(Z)$$

$$Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$$

A mesma propriedade pode ser aplicada para a **subtração** de variáveis:

$$\boldsymbol{Cov(X-Y,Z) = Cov(X,Z) - Cov(Y,Z)}$$

Ou seja, para o nosso exemplo em que Cov(X,Z) = 1 e seja Cov(Y,Z) = 2, supondo D = X – Y, então a covariância entre D e Z é:

Cov(D,Z) = Cov(X – Y, Z) = Cov(X,Z) – Cov(Y,Z) = 1 – 2 = -1

vi) $\boldsymbol{Cov(kX,Y) = Cov(X,kY) = k.Cov(X,Y)}$

A covariância de duas variáveis aleatórias, sendo **qualquer** uma delas multiplicada por uma **constante**, é igual ao **produto da constante pela covariância** das variáveis.

Considerando que a covariância entre X e Y é Cov(X,Y) = 6, então, supondo W = 5.Y, a covariância entre X e W será:

Cov(X,W) = Cov(X,5.Y) = 5.Cov(X,Y) = 5x6 = 30

E se definíssemos a variável H = 5.X, então a covariância entre H e Y seria:

Cov(H,Y) = Cov(5.X,Y) = 5.Cov(X,Y) = 30

Teríamos o mesmo resultado! Ou seja, não importa qual é a variável que está sendo multiplicada pela constante, pois o resultado será o mesmo: a covariância será multiplicada pela constante.

O mesmo vale para quando estamos dividindo por uma constante k (pois é o mesmo que multiplicar pela constante $\frac{1}{k}$). Por exemplo, para $G = \frac{Y}{3}$, a covariância entre X e W será:

$$Cov(X,G) = Cov\left(X,\frac{Y}{3}\right) = \frac{1}{3}.Cov(X,Y) = \frac{1}{3}.6 = 2$$

ESCLARECENDO!

Essa propriedade também pode ser verificada, a partir da fórmula da covariância e das propriedades da esperança.

$$Cov(k.X,Y) = E(k.X.Y) - E(k.X).E(Y)$$

$$Cov(k.X,Y) = k.E(X.Y) - k.E(X).E(Y) = k.[E(X.Y) - E(X).E(Y)]$$

$$Cov(k.X,Y) = k.Cov(X,Y)$$

Podemos deduzir que, se **ambas** as variáveis estiverem multiplicadas pela constante, então:

$$\boldsymbol{Cov(k.X,k.Y) = k.k.Cov(X,Y) = k^2.Cov(X,Y)}$$

Por exemplo, sendo W = 5.Y e H = 5.X, a covariância entre H e W é:

Cov(H,W) = Cov(5.X,5.Y) = 5x5.Cov(X,Y) = 25x6 = 150

E se as constantes forem **diferentes**, teremos:

$$\boldsymbol{Cov(k.X,l.Y) = k.l.Cov(X,Y)}$$

Por exemplo, para $G = \frac{Y}{3}$ e H = 5.X, a covariância entre H e G é:

$$Cov(H,G) = Cov\left(5.X,\frac{Y}{3}\right) = 5.\frac{1}{3}.Cov(X,Y) = 5.\frac{1}{3}.6 = 10$$

E como fica o coeficiente de correlação?

Se as variáveis estiverem multiplicadas por duas constantes quaisquer, k e l, o coeficiente de correlação se manterá o mesmo se as constantes tiverem o mesmo sinal ($kl > 0$) e terá sinal contrário se as constantes tiverem sinais diferentes ($kl < 0$):

$$\boldsymbol{\rho(k.X,l.Y) = \rho(X,Y), \text{ se } kl > 0}$$

$$\boldsymbol{\rho(k.X,l.Y) = -\rho(X,Y), \text{ se } kl < 0}$$

Assim, **o coeficiente de correlação não varia**, **em módulo**, ao multiplicarmos as variáveis aleatórias por **constantes** reais.

ESCLARECENDO!

Para obter o coeficiente de correlação, dividimos a covariância pelos desvios padrão:

$$\rho(k.X, l.Y) = \frac{Cov(k.X, l.Y)}{\sigma_{k.X}.\sigma_{l.X}} = \frac{k.l.Cov(X,Y)}{\sigma_{k.X}.\sigma_{l.X}}$$

Pelas propriedades do desvio padrão, sabemos que $\sigma_{k.X} = |k|.\sigma_X$ e $\sigma_{l.Y} = |l|.\sigma_Y$:

$$\rho(k.X, l.Y) = \frac{k.l.Cov(X,Y)}{|k|.\sigma_X.|l|.\sigma_Y} = \frac{kl}{|kl|} \times \frac{Cov(X,Y)}{\sigma_X.\sigma_X}$$

Sabemos que $\frac{Cov(X,Y)}{\sigma_X.\sigma_X} = \rho(X,Y)$, logo:

$$\rho(k.X, l.Y) = \frac{\boldsymbol{kl}}{\boldsymbol{|kl|}} \times \rho(X,Y)$$

Assim, se o produto das constantes for positivo, $kl > 0$, o que ocorre quando as constantes possuem o mesmo sinal, então teremos $kl = |kl|$ e o mesmo valor para o coeficiente de correlação:

$$\rho(k.X, l.Y) = \frac{\boldsymbol{kl}}{\boldsymbol{|kl|}} \times \rho(X,Y) = 1.\rho(X,Y)$$

Se o produto das constantes for negativo, $kl < 0$, o que ocorre quando as constantes possuem sinal contrário, então teremos $kl = -|kl|$ e o coeficiente de correlação terá sinal contrário:

$$\rho(k.X, l.Y) = \frac{\boldsymbol{kl}}{\boldsymbol{|kl|}} \times \rho(X,Y) = -1.\rho(X,Y)$$

Por exemplo, sendo A = 5.X e B = $\frac{1}{3}$.Y, o produto entre os coeficientes k = 5 e l = $\frac{1}{3}$ é **positivo**:

$$5.\frac{1}{3} = \frac{5}{3} > 0$$

Portanto, o coeficiente de correlação entre X e Y se mantém o **mesmo**:

$$\rho(A,B) = \rho\left(5.X, \frac{1}{3}.Y\right) = \rho(X,Y)$$

Similarmente, se tivermos C = -5.X e D = $-\frac{1}{3}$.Y, o produto também será **positivo**:

$$-5 \times -\frac{1}{3} = \frac{5}{3} > 0$$

Logo, o coeficiente de correlação também se mantém o **mesmo**:

$$\rho(C, D) = \rho\left(-5.X, -\frac{1}{3}.Y\right) = \rho(X, Y)$$

Porém, se tivermos A = 5.X e D = $-\frac{1}{3}$.Y, o produto entre os coeficientes é **negativo**:

$$5 \times -\frac{1}{3} = -\frac{5}{3} < 0$$

Por isso, o coeficiente de correlação terá **sinal oposto**:

$$\rho(A, D) = \rho\left(5.X, -\frac{1}{3}.Y\right) = -\rho(X, Y)$$

Se houver apenas **uma constante** k multiplicando uma das variáveis, temos um caso **específico** dessa propriedade, para $l = 1$. Nesse caso, o coeficiente de correlação será o mesmo se $k > 0$ e terá sinal contrário se $k < 0$:

$$\rho(k.X, Y) = \rho(X, Y), \text{ se } k > 0$$

$$\rho(k.X, Y) = -\rho(X, Y), \text{ se } k < 0$$

RESUMINDO

Propriedades da Covariância

- i) **Simetria**: $Cov(X, Y) = Cov(Y, X)$
- ii) **Mesma variável**: $Cov(X, X) = Var(X)$
- iii) Com uma **constante**: $Cov(k, X) = 0$
- iv) **Soma/Subtração de uma constante**: $Cov(X \pm a, Y \pm b) = Cov(X, Y)$
- v) **Soma/Subtração de variáveis**: $Cov(X \pm Y, Z) = Cov(X, Z) \pm Cov(Y, Z)$
- vi) **Produto de constantes**: $Cov(k.X, l.Y) = k.l.Cov(X, Y)$

(FGV/2015 – Prefeitura de Recife/PE) Uma variável aleatória X tem média igual a 2 e desvio padrão igual a 2. Se Y = 6 – 2X, então a média de Y, a variância de Y e o coeficiente de correlação entre X e Y valem, respectivamente,

a) -2, 4 e 1.

b) -2, 16 e 1.

c) 2, 16 e -1.

d) 10, 2 e -1.

e) 2, 4 e -1.

Comentários:

Pelas propriedades da esperança, temos:

$$E(Y) = E(6 - 2X) = 6 - 2.E(X)$$

O enunciado informa que a média de X é E(X) = 2, logo:

$$E(Y) = 6 - 2.2 = 2$$

Pela propriedade da variância, temos:

$$V(Y) = V(6 - 2X) = (-2)^2.V(X) = 4.V(X)$$

O enunciado informa que o desvio padrão de X é DP(X) = 2. Assim, a variância é $V(X) = 2^2 = 4$:

$$V(Y) = 4.4 = 16$$

Como Y = 6 – 2X, o coeficiente de correlação de X e Y é:

$$\rho(X, Y) = \rho(X, 6 - 2X)$$

Sabemos que a soma de constantes não altera o coeficiente de correlação, logo:

$$\rho(X, Y) = \rho(X, -2X)$$

Aqui, temos uma constante negativa multiplicando uma das variáveis: k = -2 < 0. Logo, o coeficiente de correlação terá sinal contrário:

$$\rho(X, Y) = -\rho(X, X)$$

Sabemos que $\rho(X, X) = 1$, logo:

$$\rho(X, Y) = -1$$

Gabarito: C

Variância da Soma e da Diferença

No **caso geral**, a **variância da soma** é dada pela seguinte fórmula:

$$\boxed{V(X+Y) = V(X) + V(Y) + 2.Cov(X,Y)}$$

Por exemplo, vamos supor que a variância de X é V(X) = 3, que a variância de Y é V(Y) = 4 e a covariância entre X e Y é Cov(X,Y) = 1. Então, a variância da soma das variáveis S = X + Y será:

$$V(S) = V(X + Y) = V(X) + V(Y) + 2.Cov(X,Y) = 3 + 4 + 2x1 = 9$$

Para a **subtração das variáveis**, temos:

$$\boxed{V(X-Y) = V(X) + V(Y) - 2.Cov(X,Y)}$$

Para o mesmo exemplo, sendo D = X – Y, a variância de D será:

$$V(D) = V(X - Y) = V(X) + V(Y) - 2.Cov(X,Y) = 3 + 4 - 2x1 = 5$$

Observe que, tanto na fórmula da variância da soma V(S), quanto na fórmula da variância da subtração V(D), iremos **somar** as **variâncias** de X e Y.

A diferença entre as duas fórmulas está no **sinal da covariância**, multiplicada por 2. Para a variância da **soma** V(S), **somamos** o dobro da covariância e para a variância da **subtração** V(D), **subtraímos** o dobro da covariância.

Para ajudar a lembrar, observe a similaridade das fórmulas acima com os **produtos notáveis**:

$$(x+y)^2 = x^2 + y^2 + 2.x.y$$

$$(x-y)^2 = x^2 + y^2 - 2.x.y$$

Para variáveis **independentes** X, Y, temos $Cov(X,Y) = 0$ e, portanto, a variância da soma será igual à variância da diferença (propriedade aditiva da variância):

$$V(X+Y) = V(X) + V(Y) + 2.Cov(X,Y) = V(X) + V(Y)$$

$$V(X-Y) = V(X) + V(Y) - 2.Cov(X,Y) = V(X) + V(Y)$$

Por exemplo, se X e Y forem independentes, com V(X) = 3 e V(Y) = 4, então a variância da soma e da diferença serão iguais a:

$$V(X + Y) = V(X - Y) = V(X) + V(Y) = 7$$

INDO MAIS FUNDO!

Se as variáveis forem **multiplicadas por constantes** reais quaisquer k, l:

$$V(k.X + l.Y) = V(k.X) + V(l.Y) + 2.Cov(k.X, l.Y)$$

Sabemos que:

$$V(k.X) = k^2.V(X)$$

$$V(l.Y) = l^2.V(Y)$$

$$Cov(k.X, l.Y) = k.l.Cov(X, Y)$$

Então:

$$\boldsymbol{V(k.X + l.Y) = k^2.V(X) + l^2.V(Y) + 2.k.l.Cov(X, Y)}$$

Analogamente, temos:

$$\boldsymbol{V(k.X - l.Y) = k^2.V(X) + l^2.V(Y) - 2.k.l.Cov(X, Y)}$$

Perceba que a similaridade com os **produtos notáveis** se mantém:

$$(k.x + l.y)^2 = k^2.x^2 + l^2.y^2 + 2.(k.l).x.y$$

$$(k.x - l.y)^2 = k^2.x^2 + l^2.y^2 - 2.(k.l).x.y$$

E se houver mais de 2 variáveis? A variância da soma de 3 variáveis, por exemplo, é:

$$V(X + Y + Z) = V(X) + V(Y) + V(Z) + 2[Cov(X, Y) + Cov(X, Z) + Cov(Y, Z)]$$

Ou seja, precisamos somar as variâncias ao dobro das covariâncias entre **todas** as variáveis. É importante notar que consideramos a covariância entre duas variáveis **uma única vez**, em razão da sua simetria, isto é, $Cov(X, Y) = Cov(Y, X)$.

Para n variáveis $X_1, X_2, \dots, X_n$, podemos representar a variância da soma $\sum_{i=1}^{n} X_i$ como:

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2.\sum_{j>i}^{n}\sum_{i=1}^{n} Cov(X_i, X_j)$$

(2016 – IBGE) Se duas variáveis aleatórias, X e Y, têm correlação linear negativa, então:

a) Quanto menor for o valor de X, menor será o valor de Y.

b) A soma dos valores esperados de X e Y é menor do que o valor esperado de X + Y.

c) O produto dos valores esperados de X e Y é menor do que o valor esperado do produto X.Y.

d) A soma das variâncias de X e Y é igual ou menor do que as variâncias de X + Y.

e) A soma das variâncias de X e Y é estritamente maior do que a variância de X + Y.

Comentários:

A questão informa que a correlação linear entre X e Y é negativa.

Em relação à alternativa A, como a covariância é negativa, então X e Y se relacionam em sentidos opostos. Assim, quanto menor for o valor de X, maior será o valor de Y (em média).

Portanto: alternativa A incorreta.

Em relação à alternativa B, a soma dos valores esperados E(X) + E(Y) é **igual** ao valor esperado E(X+Y), para **quaisquer** variáveis X e Y.

Portanto: alternativa B incorreta.

Em relação à alternativa C, o valor de E(X.Y) pode ser calculado a partir da covariância:

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

$$E(X.Y) = Cov(X,Y) + E(X).E(Y)$$

Como a correlação entre X e Y é negativa, então $Cov(X,Y) < 0$. Dessa forma:

$$E(X.Y) < E(X).E(Y)$$

Ou seja, o produto dos valores esperados E(X).E(Y) é **maior** que o valor esperado do produto E(X.Y).

Portanto: alternativa C incorreta.

Em relação às alternativas D e E, a variância de X + Y é:

$$V(X + Y) = V(X) + V(Y) + 2.Cov(X,Y)$$

Como $Cov(X,Y) < 0$, então:

$$V(X + Y) < V(X) + V(Y)$$

Ou seja, A soma das variâncias de V(X) + V(Y) é **maior** que a V(X + Y).

Portanto: alternativa D incorreta e alternativa E correta.

Gabarito: E.

(FGV/2015 – TJ/RO) Seja X = número de anos de condenação e Y = nível de renda do condenado (mil reais). São fornecidas ainda as seguintes informações:

Var(X) = 25; Var (Y) = 16 e Var (X+Y) = 21

Assim sendo, a correlação entre X e Y é igual a:

a) 0,20

b) 0,25

c) 0,50

d) -0,50

e) -0,10

Comentários:

A correlação é dada por:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X . \sigma_Y}$$

O valor de Cov(X,Y) pode ser obtido pela fórmula da **variância da soma**:

$$V(X + Y) = V(X) + V(Y) + 2.Cov(X,Y)$$

O enunciado informa que V(X + Y) = 21, V(X) = 25 e V(Y) = 16, logo:

$$21 = 25 + 16 + 2.Cov(X,Y)$$

$$2.Cov(X,Y) = 21 – 25 – 16 = -20$$

$$Cov(X,Y) = -10$$

Os valores dos desvios padrão são a raiz quadrada das variâncias. Sendo V(X) = 25, então:

$$\sigma_X = \sqrt{25} = 5$$

Sendo V(Y) = 16, então:

$$\sigma_Y = \sqrt{16} = 4$$

Portanto, o coeficiente de correlação é:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X . \sigma_Y} = \frac{-10}{5 \times 4} = -0{,}5$$

Gabarito: D

Resumo da Aula

Função de Distribuição Acumulada

$$F(x) = P(X \leq x)$$

Esperança Matemática (média)

$$E(X) = \sum x.P(X = x)$$

Propriedades:

- $E(kX) = k.E(X)$
- $E(X + k) = E(X) + k$
- $E(X \pm Y) = E(X) \pm E(Y)$
- $E(k) = k$
- Se X e Y forem **independentes**, então $E(X \times Y) = E(X) \times E(Y)$

Moda: valor de **X** com maior probabilidade

Mediana: divide a distribuição em duas partes iguais

$$F(x_{Med}) = 0{,}5$$

Variância

$$V(X) = \sum (x - \mu)^2 \times P(X = x)$$

Propriedades:

- $V(X + k) = V(X)$
- $V(k.X) = k^2.V(X)$
- $V(k) = 0$
- Se X e Y forem **independentes**, então $V(X \pm Y) = V(X) + V(Y)$

Desvio Padrão

$$\sigma = \sqrt{V(X)}$$

Covariância

$$Cov(X,Y) = E(X.Y) - E(X).E(Y)$$

Correlação

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X.\sigma_Y}$$

Se X e Y forem **independentes**, então $Cov(X,Y) = 0, \rho(X,Y) = 0$

Variância da Soma e da Diferença

$$\boldsymbol{V(X+Y) = V(X) + V(Y) + 2.Cov(X,Y)}$$

$$\boldsymbol{V(X-Y) = V(X) + V(Y) - 2.Cov(X,Y)}$$

Coeficiente de Variação

$$\boldsymbol{C_V = \frac{\sigma}{\mu}}$$

Coeficiente de Variação e Variância Relativa

A definição de **coeficiente de variação** (também chamado de **desvio padrão relativo** ou, ainda, de **coeficiente de variabilidade**), C_V, é:

$$\boldsymbol{C_V = \frac{\sigma}{\mu}}$$

Podemos dizer que esse parâmetro representa uma **normalização** do **desvio padrão** pela **média**. A normalização tem como objetivo transformar determinados valores para uma **escala comum**, permitindo assim uma análise comparativa. Nesse caso, dividimos o desvio padrão pela média para que possamos comparar a dispersão de variáveis com **médias distintas**.

Por exemplo, vamos supor que a variável aleatória X apresente média $\mu_X = 100$ e desvio padrão $\sigma_X = 20$; e que a variável aleatória Y apresente média $\mu_Y = 10$ e desvio padrão $\sigma_Y = 5$.

Nesse caso, não poderíamos afirmar que a dispersão de X é maior que a de Y, simplesmente porque $\sigma_X > \sigma_Y$. Para efetuarmos essa comparação, precisamos do **Coeficiente de Variação**. Para esse exemplo, temos:

$$C_{V_X} = \frac{\sigma_X}{\mu_X} = \frac{20}{100} = 0{,}2$$

$$C_{V_Y} = \frac{\sigma_Y}{\mu_Y} = \frac{5}{10} = 0{,}5$$

Portanto, $C_{V_Y} > C_{V_X}$. Como a média e o desvio padrão consideram a mesma unidade de medida (que dependem da variável aleatória), o coeficiente de variação é **adimensional**, isto é, **não possui unidade de medida**, sendo apenas um **número**.

A **variância relativa**, V_R, também apresenta o mesmo objetivo, qual seja, de permitir comparações entre variáveis com **ordens de grandeza distintas**.

Por definição, a **variância relativa** é o **quadrado do coeficiente de variação**:

$$\boldsymbol{V_R = (C_V)^2 = \frac{\sigma^2}{\mu^2} = \frac{V(X)}{\mu^2}}$$

Ou seja, a variância relativa é o **quociente entre a variância e o quadrado da média**.

Para o nosso exemplo, em que $C_{V_X} = 0{,}2$, a variância relativa de X é:

$$V_R = (C_V)^2 = (0{,}2)^2 = 0{,}04$$

(2015 – Analista de Planejamento e Orçamento) O coeficiente de correlação de duas variáveis aleatórias x e y é igual 0,7, ou seja: δ (x , y) = 0,7. O coeficiente de variabilidade de x é 0,3 – por γ_x =0,3. O coeficiente de variabilidade de y é 0,5 – γ_y =0,5. Com essas informações sobre as variáveis x e y, pode-se, corretamente, afirmar que:

a) à medida que x cresce, em média y decresce.

b) a variabilidade absoluta de x é maior que a variabilidade absoluta de y.

c) o desvio-padrão de x é 30% menor do que sua média.

d) o desvio-padrão de y é 50% de sua média.

e) o desvio-padrão de y é 50% maior do que sua média.

Comentários:

A questão informa que o coeficiente de correlação entre X e Y é 0,7: ρ(X,Y) = 0,7; que o coeficiente de variabilidade (ou de variação) de X é 0,3:

$$C_{V_X} = \frac{\sigma_X}{\mu_X} = 0{,}3$$

E que o coeficiente de variabilidade (ou de variação) de Y é 0,5:

$$C_{V_Y} = \frac{\sigma_Y}{\mu_Y} = 0{,}5$$

Em relação à alternativa A, como o coeficiente de correlação é positivo, à medida que x cresce, em média, y também cresce.

Portanto, a alternativa A está incorreta.

Em relação à alternativa B, não é possível calcular as médias μ_X e μ_Y com as informações fornecidas, assim, não é possível afirmar algo sobre as variabilidades absolutas (isto é, os desvios padrão ou as variâncias) das variáveis.

Portanto, a alternativa B está incorreta.

Em relação à alternativa C, podemos afirmar que o desvio padrão de X é 30% da sua média:

$$\sigma_X = 0{,}3 \times \mu_X$$

Portanto, a alternativa C está incorreta.

Em relação às alternativas D e E, podemos afirmar que o desvio padrão de Y é 50% da sua média:

$$\sigma_Y = 0{,}5 \times \mu_Y$$

Portanto, a alternativa D está correta e a alternativa E está incorreta.

Gabarito: D.

Resumo

Função de Distribuição Acumulada: $F(x) = P(X \leq x)$

Esperança Matemática (média): $E(X) = \sum x . P(X = x)$

- $E(kX) = k. E(X)$
- $E(X + k) = E(X) + k$
- $E(X \pm Y) = E(X) \pm E(Y)$
- $E(k) = k$
- Se X e Y forem **independentes**, então $E(X \times Y) = E(X) \times E(Y)$

Moda: valor de **X** com maior probabilidade

Mediana: divide a distribuição em duas partes iguais, $F(x_{Med}) = 0{,}5$

Variância: $V(X) = \sum (x - \mu)^2 \times P(X = x)$

- $V(X + k) = V(X)$
- $V(k. X) = k^2. V(X)$
- $V(k) = 0$
- Se X e Y forem **independentes**, então $V(X \pm Y) = V(X) + V(Y)$

Desvio Padrão: $\sigma = \sqrt{V(X)}$

Covariância: $Cov(X, Y) = E(X. Y) - E(X). E(Y)$

Correlação: $\rho(X, Y) = \frac{Cov(X,Y)}{\sigma_X . \sigma_Y}$

- Se X e Y forem **independentes**, então $Cov(X, Y) = 0, \rho(X, Y) = 0$

Variância da Soma e da Diferença

$$\boldsymbol{V(X + Y) = V(X) + V(Y) + 2. Cov(X, Y)}$$

$$\boldsymbol{V(X - Y) = V(X) + V(Y) - 2. Cov(X, Y)}$$

Coeficiente de Variação: $\boldsymbol{C_V = \frac{\sigma}{\mu}}$

Variância Relativa: $\boldsymbol{V_R = (C_V)^2 = \frac{V(X)}{\mu^2}}$