Estratégia
Concursos

Aula 01

Receita Federal (Analista Tributário) Bizu Estratégico - 2022 (Pré-Edital)

Autor:

Amanda Alves, Camila Damázio, Cíntia Bócoli, Diogo Matias das Neves, Elizabeth Menezes de Pinho Alves, Fernanda Harumi Amaral Jo, Guilherme Carvalho, Heloísa Tondinelli, Jefferson de Souza Correia, Kauê Salvaterra,

20 de Dezembro de 2021

BIZU ESTRATÉGICO DE ESTATÍSTICA (RECEITA FEDERAL)

Olá, prezado aluno. Tudo certo?

Neste material, traremos uma seleção de *bizus* da disciplina de **Estatística** para o concurso da **Receita Federal.**

O objetivo é proporcionar uma revisão rápida e de alta qualidade aos alunos por meio de tópicos que possuem as maiores chances de incidência em prova.

Todos os *bizus* destinam-se a alunos que já estejam na fase bem final de revisão (que já estudaram bastante o conteúdo teórico da disciplina e, nos últimos dias, precisam revisar por algum material bem curto e objetivo).

Diogo Matias

@oprimoconcursado

Leonardo Mathias

@profleomathias

ANÁLISE ESTATÍSTICA

Pessoal, segue abaixo uma análise estatística dos assuntos mais exigidos pelas Bancas **FCC, FGV e Cebraspe**, no âmbito da disciplina de **Estatística**, em concursos da **Área Fiscal.**

Estatística (Foram encontradas 83 questões)		
Assunto	**Quantidade de questões**	**% de cobrança**
Distribuições Teóricas de Probabilidade	32	38,55%
Medidas de Posição	18	21,69%
Probabilidade	15	18,07%
Medidas de Variabilidade	9	10,84%
Distribuição de Frequências	3	3,61%

*** Análise realizada em provas aplicadas de 2012 até 2022.**

Com essa análise, podemos verificar quais são os temas mais exigidos pelas bancas **FCC, FGV e Cebraspe** e, através disso, focaremos nos principais pontos em nossa revisão!

A disciplina **Estatística** no último edital do concurso da **Receita Federal** para o cargo de **Analista Tributário** abordou o seguinte conteúdo programático:

-Probabilidade e Estatística Descritiva.

Estatística – Receita Federal		
Assunto	**Bizus**	**Caderno de Questões**
Distribuições Teóricas de Probabilidade	1 a 3	http://questo.es/z9djtk
Medidas de posição	4 a 9	http://questo.es/0nf4jw
Probabilidade	10 a 15	http://questo.es/ktz5dz
Medidas de Variabilidade	16 a 19	http://questo.es/ediwal
Distribuição de frequências	20 a 24	http://questo.es/i16zfb

Apresentação

É com imensa satisfação que terei o privilégio de acompanhar a sua jornada rumo à aprovação. Antes de mais nada, permita-me uma breve apresentação:

Meu nome é ***Diogo Matias das Neves***, tenho 29 anos, sou formado em Administração pela Universidade Católica de Pernambuco (2013) e sou natural de Recife/PE.

Atualmente, moro em São Paulo em virtude do exercício do cargo de ***Auditor de Controle Externo*** no Tribunal de Contas do Estado de São Paulo (***TCE-SP***), tendo sido aprovado no último certame realizado em 2017.

Também fui aprovado nas vagas no último concurso da ***Polícia Federal*** para o cargo de Agente de Polícia Federal, além das aprovações em 30° para ***Auditor do Estado do RS (CAGE-RS)*** e também 30° no de ***Auditor de Controle Externo do TCM-BA.***

Tentarei utilizar da minha experiência de mais de 5 anos estudando para concursos e conquistando aprovações em diversas áreas para auxiliá-lo(a) na preparação desse almejado concurso.

Diogo Matias das Neves

Distribuições Teóricas de Probabilidade

1) Distribuições Uniformes

✓ Para distribuições uniformes, havendo um total de N elementos, a probabilidade de cada valor $X = x$ é calculada como:

$$P(X = x) = \frac{1}{N}$$

✓ A esperança matemática da distribuição uniforme corresponde à média aritmética dos valores de x

$$\boldsymbol{E(X) = \frac{\sum x}{N}}$$

Por exemplo, no caso de um dado, a esperança é:

$$E(X) = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = \frac{7}{2}$$

Para calcular a variância dessa distribuição, vamos lembrar a fórmula da variância:

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

Sabemos calcular E(X), então basta elevá-lo ao quadrado para calcular $[E(X)]^2$. Já o valor de $E(X^2)$ é definido como:

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

Como $P(X = x) = \frac{1}{N}$ para uma distribuição uniforme, então para essa distribuição, temos:

$$E(X^2) = \sum x^2 . \frac{1}{N} = \frac{\sum x^2}{N}$$

Ou seja, $E(X^2)$ é a **média aritmética** dos valores de X^2.

Para o exemplo do dado, o valor de $E(X^2)$ é:

$$E(X^2) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} = \frac{1 + 4 + 9 + 16 + 25 + 36}{6} = \frac{91}{6}$$

Logo, a variância será a diferença:

$$V(X) = E(X^2) - [E(X)]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182 - 147}{12} = \frac{35}{12}$$

2) Distribuição de Bernoulli

✓ Uma variável aleatória discreta X com distribuição de Bernoulli assume apenas 2 valores possíveis, 0 ou 1, em um experimento realizado uma única vez. Esse experimento é chamado de Ensaio de Bernoulli.

Distribuição de Bernoulli (p)

1 experimento: **Ensaio de Bernoulli**

2 resultados possíveis: sucesso ($X = 1$) ou fracasso ($X = 0$)

Probabilidade de sucesso: $P[X = 1] = \boldsymbol{p}$

Probabilidade de fracasso: $P[X = 0] = \boldsymbol{q} = \mathbf{1} - \boldsymbol{p}$

Esperança: $E(X) = \boldsymbol{p}$; Variância: $V(X) = \boldsymbol{p}.\boldsymbol{q}$

Vamos calcular a esperança matemática da distribuição de Bernoulli. A fórmula geral da esperança é:

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

$$E(X) = 0 \times q + 1 \times p$$

$$\boldsymbol{E(X) = p}$$

Para calcular a **variância**, primeiro calculamos $E(X^2)$:

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

$$E(X^2) = 0^2 \times q + 1^2 \times p$$

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

Logo, a variância é:

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

$$V(X) = p - p^2 = p(1 - p)$$

$$\boldsymbol{V(X) = p.q}$$

3) Distribuição Binomial

- ✓ Quando repetimos um mesmo Ensaio de Bernoulli (isto é, o experimento com 2 resultados possíveis), damos origem à **Distribuição Binomial.**

- ✓ Mais precisamente, n repetições independentes de um Ensaio de Bernoulli, com a mesma probabilidade de sucesso, resultam na Distribuição Binomial. O conceito de repetições independentes significa que o resultado de um experimento não afeta o resultado de outro

$$\boldsymbol{P(X = k) = C_{n,k} \times p^k \times q^{n-k}}$$

É possível calcular também a probabilidade de um **intervalo** da variável binomial, por exemplo, obter 1 **OU** 2 sucessos, em 3 lançamentos de uma moeda (com p = q = 0,5). Como são eventos exclusivos, devemos **somar** as probabilidades:

$$P(X = 1 \cup X = 2) = P(X = 1) + P(X = 2)$$

A probabilidade de obter 1 sucesso é:

$$P(X = 1) = C_{3,1} \times 0{,}5^1 \times 0{,}5^2 = 3 \times 0{,}5 \times 0{,}25 = 0{,}375$$

A probabilidade de obter 2 sucessos é:

$$P(X = 2) = C_{3,2} \times 0{,}5^2 \times 0{,}5^1 = 3 \times 0{,}25 \times 0{,}5 = 0{,}375$$

Então, a probabilidade de obter 1 OU 2 sucessos é:

$$P(X = 1 \cup X = 2) = 0{,}375 + 0{,}375 = 0{,}75$$

Também podemos calcular a probabilidade de obter **pelo menos um** sucesso. Normalmente, nesses casos, é mais fácil calcular a probabilidade do complemento:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0)$$

A probabilidade de obter 0 sucesso é:

$$P(X = 0) = C_{3,0} \times 0{,}5^3 \times 0{,}5^0 = 1 \times 0{,}125 \times 1 = 0{,}125$$

Logo, a probabilidade de obter pelo menos um sucesso é:

$$P(X \geq 1) = 1 - 0{,}125 = 0{,}875$$

Medidas de posição

4) Média Aritmética Simples

- ✓ Essa medida é definida como o quociente entre a soma de todos os elementos e o número deles. A propriedade principal da média é preservar a soma dos elementos de um conjunto de dados.

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Ex: Calcule a média aritmética dos números 8, 16, 26 e 30.

Para responder a essa questão, somaremos os quatro números e, em seguida, dividiremos o
resultado por quatro:

$\overline{x}$ = 8 + 16 + 26 + 30

4 = 80

4 = 20

Portanto, 20 é o valor da média aritmética dos números 8, 16, 26 e 30.

Sempre que a questão não especificar qual o tipo de média, faremos o cálculo da média aritmética.

Propriedades da Média Aritmética

1ª Propriedade: Dado um conjunto com $n \geq 1$ elementos, a média aritmética sempre existirá e será única.

2ª Propriedade: A média aritmética $\overline{x}$ de um conjunto de dados satisfaz a expressão $m \leq \overline{x} \leq M$, em que m e M são, respectivamente, os elementos que representam o valor mínimo e o valor máximo desse conjunto.

$$mínimo \leq \overline{x} \leq Máximo$$

3ª Propriedade: Somando-se (ou subtraindo-se) uma constante c de todos os valores de uma variável, a média do conjunto fica aumentada (ou diminuída) dessa constante.

$$\overline{y} = \overline{x} + c \quad \text{ou} \quad \overline{y} = \overline{x} - c$$

4ª Propriedade: Multiplicando-se (ou dividindo-se) uma constante c de todos os valores de uma variável, a média do conjunto fica multiplicada (ou dividida) por esta constante.

$$\bar{y} = \bar{x} \times c \quad \text{ou} \quad \bar{y} = \bar{x} \div c$$

5ª Propriedade: A soma algébrica dos desvios em relação à média é nula.

$$\sum_{i=1}^{n}(x_i - \bar{x}) = 0$$

6ª Propriedade: A soma dos quadrados dos desvios da sequência de números $\{x_i\}$, em relação a um número a, é mínima se a for a média aritmética dos números.

$$\sum_{i=1}^{n}(x_i - a)^2 \geq \sum_{i=1}^{n}(x_i - \bar{x})^2$$

5) Média Ponderada

- ✓ Uma média ponderada é a média de um conjunto de dados cujos valores possuem pesos variados. Ela é calculada pela igualdade a seguir, em que p é o peso de cada valor de x

$$\bar{x} = \frac{\sum_{i=1}^{n}(x_i \times p_i)}{\sum_{i=1}^{n} p_i}$$

- ✓ Observe que no numerador cada valor será multiplicado pelo seu respectivo peso, enquanto no denominador teremos a soma de todos os pesos.

- ✓ Ex. Suponha que um candidato tenha prestado um concurso público para o cargo de Auditor Fiscal, alcançando as seguintes notas:

Disciplina	Nota (x_i)	Peso (p_i)	$x_i \times p_i$
Língua Portuguesa	4,0	1	4,0 x 1 = 4,0
Direito Administrativo	4,0	2	4,0 x 2 = 8,0
Direito Constitucional	4,0	2	4,0 x 2 = 8,0
Direito Tributário	6,0	3	6,0 x 3 = 18,0
Legislação Tributária	6,0	3	6,0 x 3 = 18,0
Contabilidade	7,0	3	7,0 x 3 = 21,0
Auditoria	7,0	3	7,0 x 3 = 21,0

Nesse ponto, temos uma lista contendo todos os produtos de notas e pesos. Então, a média aritmética ponderada é dada por:

$$\bar{x} = \frac{x_1 \times p_1 + x_2 \times p_2 + \cdots + x_n \times p_n}{p_1 + p_2 + \cdots + p_n}$$

$$\bar{x} = \frac{4{,}0 + 8{,}0 + 8{,}0 + 18{,}0 + 18{,}0 + 21{,}0 + 21{,}0}{1 + 2 + 2 + 3 + 3 + 3 + 3} = \frac{98}{15} \cong 6{,}53$$

6) Média Geométrica

Essa medida é definida, para o conjunto de números positivos, como a raiz $\boldsymbol{n}$-ésima do produto de $\boldsymbol{n}$ elementos de um conjunto de dados. A propriedade principal dessa média é preservar o produto dos elementos de um conjunto de dados.

$$\boxed{\boldsymbol{G = \sqrt[n]{x_1 \times x_2 \times \cdots \times x_n}}}$$

7) Média Harmônica

Essa medida é definida, para o conjunto de números positivos, como o inverso da média aritmética dos inversos. A propriedade principal dessa média é preservar a soma dos inversos dos elementos de um conjunto de números.

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$$

Como vimos no início, muitas vezes, **a média harmônica é descrita como o inverso da média aritmética dos inversos**. Isso porque a fórmula acima também pode ser escrita na forma mostrada a seguir, em que $\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}\right)/n$ corresponde à média aritmética dos inversos.

$$H = \frac{1}{\frac{\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}\right)}{n}}$$

8) Moda

- ✓ A moda é uma medida de posição e de tendência central que descreve o valor mais frequente de um conjunto de observações, ou seja, o valor de maior ocorrência dentre os valores observados.

- ✓ A definição evidencia que um conjunto de valores pode possuir uma ou mais modas, ou não possuir nenhuma. Assim, dizemos que um conjunto é unimodal, bimodal, trimodal ou plurimodal, de acordo com o número de modas que apresenta. A ausência de uma moda caracteriza o conjunto como amodal.

- ✓ Em geral, a moda é utilizada em distribuições nas quais o valor mais frequente é o mais importante da distribuição. A moda também é útil para a determinação da medida de posição de variáveis qualitativas nominais, ou seja, variáveis não-numéricas que não podem ser ordenadas.

9) Moda para dados não-agrupados

- ✓ Para determinarmos a moda de um conjunto ordenado de valores não agrupados em classes, **basta identificarmos** o elemento (ou elementos) de maior frequência **no conjunto**. Diferentemente das outras medidas de **tendência central,** a moda nem sempre existirá em um conjunto de valores. Além disso, em certas situações, poderemos ter uma, duas, ou várias modas no mesmo conjunto.

Com relação ao número de modas, o conjunto de pode ser classificado como:

- ✓ amodal: **quando** todos os elementos apresentam a mesma frequência, isto é, **quando todos aparecem o mesmo número de vezes**:

$$\{1, 2, 3, 5, 6, 8, 10\}$$

- ✓ unimodal: **quando** a frequência de um elemento é maior que as frequências dos demais elementos. **Isto é, quando o conjunto tem** uma única moda**.** No conjunto a seguir, o elemento **2** repete-se **cinco** vezes, enquanto o elemento **3** aparece **duas** vezes. Logo, $M_o = 2$.

$$\left\{\underbrace{2, 2, 2, 2, 2}_{5\ elementos}, \underbrace{3, 3}_{2\ elementos}\right\}$$

- ✓ bimodal: **quando** as frequências de dois elementos são iguais e maiores que as frequências dos demais elementos. **Isto é, quando o conjunto tem um** duas modas**.** No conjunto a seguir, os elementos **2** e **3** repetem-se **cinco** vezes, enquanto o elemento **4** aparece **duas** vezes. Logo, $M_o = 2$ e $M_o = 3$.

$$\left\{\underbrace{2, 2, 2, 2, 2}_{5\ elementos}, \underbrace{3, 3, 3, 3, 3}_{5\ elementos}, \underbrace{4, 4}_{2\ elementos}\right\}$$

- ✓ multimodal ou plurimodal: **quando** as frequências de três ou mais elementos são iguais e maiores que as frequências dos demais elementos. **Isto é, quando o conjunto tem** três ou mais modas**.** No conjunto a seguir, os elementos 2, 3 e 4 repetem-se cinco vezes, enquanto o elemento 5 aparece duas vezes. Logo, Mo = 2, Mo = 3 e Mo = 4.

$$\left\{\underbrace{2,2,2,2,2}_{5\ elementos},\underbrace{3,3,3,3,3}_{5\ elementos},\underbrace{4,4,4,4,4}_{5\ elementos},\underbrace{5,5}_{2\ elementos}\right\}$$

Probabilidade

10) Conceitos

✓ A Teoria da Probabilidade é o ramo da Estatística que estuda experimentos e fenômenos aleatórios, isto é, cujos resultados são incertos.

✓ Os Experimentos/Fenômenos aleatórios:
i) Podem ser repetidos indefinidamente, sob condições inalteradas;
ii) Apresentam resultado incerto, porém com um padrão conhecido.

<u>Espaço Amostral</u>

O Espaço Amostral de um experimento/fenômeno aleatório é o conjunto de todos os resultados possíveis. Podemos chamar o Espaço Amostral de Universo e denotá-lo por U ou Ω.

No lançamento de uma moeda, por exemplo, o Espaço Amostral é o conjunto:

U_M = {CARA, COROA}

Para o lançamento de um dado (com 6 faces), o Espaço Amostral é o conjunto:

U_D = {1,2,3,4,5,6}

Se o experimento for o lançamento de 2 moedas, o Espaço Amostral é dado por:

U2M = {(CARA,CARA), (CARA,COROA), (COROA,CARA), (COROA,COROA)}

Podemos, ainda, chamar cada resultado possível de ponto amostral. No lançamento de 2 moedas que acabamos de ver, por exemplo, há 4 pontos amostrais.

<u>Evento</u>

✓ Um evento é todo e qualquer **subconjunto** do Espaço Amostral. No exemplo do lançamento de 2 moedas, podemos chamar de evento A aquele em que ambas as moedas apresentam o mesmo resultado para a face superior. Portanto, o evento A é o subconjunto:

A = {(CARA,CARA), (COROA,COROA)}

Evento **simples** ou **elementar** → n(B) = 1
Evento **impossível**: C = ∅ → n(C) = 0
Evento **certo**: D = U → n(D) = n(U)

11) Definição Clássica

✓ A probabilidade representa as chances de um evento ocorrer. Sendo U o Espaço Amostral, a probabilidade de ocorrer o evento A é (definição clássica):

$$P(A) = \frac{n(A)}{n(U)}$$

Ou seja, a probabilidade de um evento é a **razão** entre o **número de elementos do Evento**, $n(A)$, e o **número de elementos do Espaço Amostral**, $n(U)$.

✓ Vale pontuar que, em diversos casos, será necessário utilizar as técnicas de análise combinatória, que vimos na aula passada, para calcular o número de elementos do evento e/ou o número de elementos do Espaço Amostral.

Vamos supor haja 5 peças amarelas e 6 peças verdes dentro de um saco e que teremos que retirar 2 peças sem olhar. Qual é a probabilidade de retirar 2 peças amarelas?

Sabemos que a probabilidade é a razão entre o número de casos favoráveis e o número de casos totais:

$$P = \frac{casos\ favoráveis}{casos\ totais} = \frac{n(A)}{n(U)}$$

Os **casos favoráveis** são as maneiras de retirar 2 peças, dentre as 5 peças amarelas. Como a ordem não importa, temos a combinação 2, dentre 5 elementos:

$$n(A) = C_{5,2} = \frac{5!}{(5-2)!\times 2!} = \frac{5\times 4\times 3!}{3!\times 2!} = \frac{5\times 4}{2} = 10$$

Os **casos totais** são as maneiras de retirar 2 peças, de um total de 11 peças (entre amarelas e verdes), também sem importância de ordem:

$$n(U) = C_{11,2} = \frac{11!}{(11-2)!\times 2!} = \frac{11\times 10\times 9!}{9!\times 2!} = \frac{11\times 10}{2} = 55$$

Logo, a probabilidade de retirar 2 peças amarelas é: $P = \frac{10}{55}$

E se a ordem importasse?

Vamos supor, então, que há 5 mulheres e 6 homens, dos quais 2 serão escolhidos para ocupar a posição de presidente e vice-presidente do grupo. Qual seria a probabilidade de escolher mulheres para ambos os cargos?

A probabilidade é calculada pela mesma razão:

$$P = \frac{casos\ favoráveis}{casos\ totais} = \frac{n(A)}{n(U)}$$

Os **casos favoráveis** são as maneiras de escolher 2 mulheres, dentre 5, sendo que a ordem importa, por serem cargos distintos:

$$n(A) = A_{5,2} = \frac{5!}{(5-2)!} = \frac{5 \times 4 \times 3!}{3!} = 5 \times 4 = 20$$

Os **casos totais** são as maneiras de escolher 2 peças, de um total de 11 pessoas (entre mulheres e homens), também com importância de ordem:

$$n(U) = A_{11,2} = \frac{11!}{(11-2)!} = \frac{11 \times 10 \times 9!}{9!} = 11 \times 10 = 110$$

Logo, a probabilidade de escolher 2 mulheres é: $P = \frac{20}{110} = \frac{10}{55}$

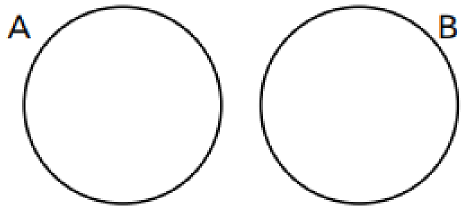
12) Combinação de Eventos e Probabilidade

Quando a **união** de eventos corresponde a **todo o Espaço Amostral**, dizemos que tais eventos são **exaustivos**.

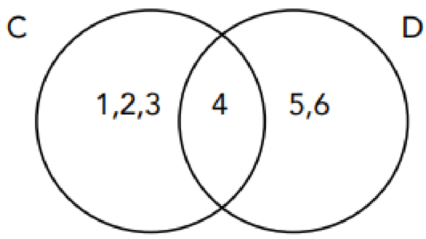
Eventos A e B **Exaustivos**: **A ∪ B = U**

Quando a **interseção** de eventos é um **conjunto vazio**, dizemos que tais eventos são **mutuamente excludentes** (ou exclusivos). Podemos dizer, ainda, que os conjuntos são **disjuntos.**

Eventos A e B **Mutuamente Excludentes**: **A ∩ B = ∅**



➢ Probabilidade da União (caso geral):



Sabendo que a probabilidade de um evento qualquer é $P(X) = \frac{n(X)}{n(U)}$, podemos dividir, por n(U), toda a equação referente ao número de elementos da união. Assim, obtemos:

$$\boxed{P(C \cup D) = P(C) + P(D) - P(C \cap D)}$$

Em resumo:

Eventos A e B **Exaustivos**: **P(A ∪ B) = P(U) = 1**

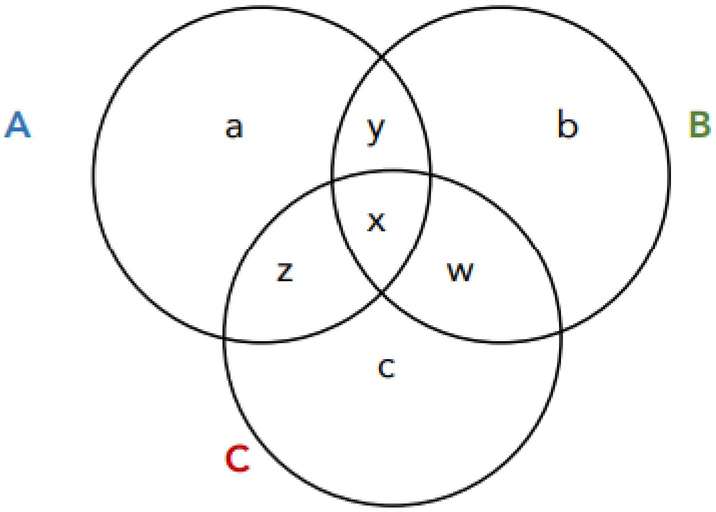
Eventos A e B **Mutuamente Excludentes**: **P(A ∩ B) = P(∅) = 0**

Probabilidade da **União** (caso geral): **P(C ∪ D) = P(C) + P(D) – P(C ∩ D)**

Probabilidade da **União** de **Eventos Excludentes**: **P(A ∪ B) = P(A) + P(B)**

13) União de 3 eventos

- A união de 3 eventos, A, B e C, pode ser representada pelo seguinte **Diagrama de Venn:**



$$\boxed{P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)}$$

14) Axiomas e propriedades

Axiomas

1. $P(A) \geq 0$
2. $P(U) = 1$
3. Se A e B são **mutuamente excludentes** então $P(A \cup B) = P(A) + P(B)$

Propriedades

i) Se $A = \varnothing$, então $P(A) = 0$
ii) $0 \leq P(A) \leq 1$
iii) $P(A - B) = P(A) - P(A \cap B)$
iv) Se A ⊆ B, então $P(A) \leq P(B)$
v) $P(A \cup B) = P(A) + P(B) - P(A)$
vi) $P(\bar{A}) = 1 - P(A)$

15) Probabilidade Condicional

- ✓ A probabilidade condicional trabalha com a probabilidade de um evento ocorrer, sabendo que outro já ocorreu. Por exemplo, vamos supor que, em um auditório, existam enfermeiros e dentistas, tanto homens quanto mulheres. Podemos calcular a probabilidade de uma pessoa escolhida aleatoriamente ser enfermeiro, dado que é homem.

- ✓ O fato de sabermos que a pessoa escolhida é homem corresponde a uma redução do universo de possibilidades – não estamos mais considerando todo o auditório, mas apenas os homens nesse auditório. Com esse "novo" universo, calculamos a probabilidade de esse homem ser enfermeiro.

$$P(E|H) = \frac{P(E \cap H)}{P(H)}$$

Medidas de Variabilidade

Medidas de Variabilidade

16) Amplitude Total

A **amplitude total** (ou simplesmente amplitude) é a diferença entre os valores extremos de um conjunto de observações, ou seja, a diferença entre o maior e o menor elemento desse conjunto

$$A = x_{máx} - x_{mín}$$

Por exemplo, vamos comparar os conjuntos e da tabela a seguir:

Conjunto	Média	Amplitude total
$A = 5, 7, 8, 9, 10, 11, 55$	$\bar{x} = 15$	$A = 55 - 5 = 50$
$B = 12, 13, 14, 15, 16, 17, 18$	$\bar{x} = 15$	$A = 18 - 12 = 6$

Amplitude Total para dados não-agrupados

Para dados não agrupados, o cálculo da amplitude total pode ser expresso pela seguinte fórmula:

$$\boxed{A = x_{máx} - x_{mín}}$$

Calcular a amplitude total dos conjuntos apresentados a seguir:

$$A = 50, 50, 50, 50, 50, 50, 50$$

$$B = 47, 48, 49, 50, 51, 52, 53$$

$$C = 20, 30, 40, 50, 60, 70, 80$$

Aplicando a fórmula anterior para esses dados, obtemos os seguintes resultados:

$$AT_A = x_{máx} - x_{mín} = 50 - 50 = 0$$

$$AT_B = x_{máx} - x_{mín} = 53 - 47 = 6$$

$$AT_C = x_{máx} - x_{mín} = 80 - 20 = 60$$

Amplitude Total para dados agrupados sem intervalos de classes

Para dados agrupados sem intervalos de classe, a fórmula usada para a identificação da amplitude total é similar à adotada para dados não-agrupados. A única diferença consiste na identificação dos valores
mínimo e máximo, que agora ocorre por meio de uma tabela de frequência.

Calcular a amplitude total da tabela de frequências apresentada a seguir.

x_i	f_i
1	10
3	15
5	10
7	8
9	7

Nesse caso, como 1 e 9 são os valores mínimo e máximo da variável x_i, temos o seguinte resultado:

$$A = x_{máx} - x_{mín}$$

$$A = 9 - 1 = 8$$

Amplitude Total para dados agrupados em classes

Para dados agrupados em intervalos de classe, podemos definir a amplitude total de duas formas:

1) pela diferença entre o limite superior da última classe (Lsup) e o limite inferior da primeira classe (Linf), conforme expresso na fórmula a seguir:

$$\boxed{\boldsymbol{A = L_{sup} - l_{inf}}}$$

2) pela diferença entre o ponto médio da última classe ($\boldsymbol{PM_{últ}}$) e o ponto médio da primeira classe ($\boldsymbol{PM_{pri}}$), conforme expresso na fórmula a seguir:

$$\boxed{\boldsymbol{A = PM_{últ} - PM_{pri}}}$$

Calcular a amplitude total da distribuição de frequências apresentada a seguir:

Classes	PM_i	f_i
$1 \vdash 5$	3	5
$5 \vdash 9$	7	10
$9 \vdash 13$	11	15
$13 \vdash 17$	15	10
$17 \vdash 21$	19	5
Total		**45**

Pelo primeiro método, temos que o limite superior da última classe é 21, enquanto o limite inferior da primeira classe é 1. Portanto, temos a seguinte amplitude:

$$A = L_{sup} - l_{inf}$$

$$A = 21 - 1 = 20$$

Pelo segundo método, temos que o ponto médio da última classe é 19, enquanto o ponto médio da
primeira classe é 3. Portanto, temos a seguinte amplitude:

$$A = PM_{\text{últ}} - PM_{pri}$$

$$A = 19 - 3 = 16$$

Propriedades da Amplitude Total

1ª Propriedade

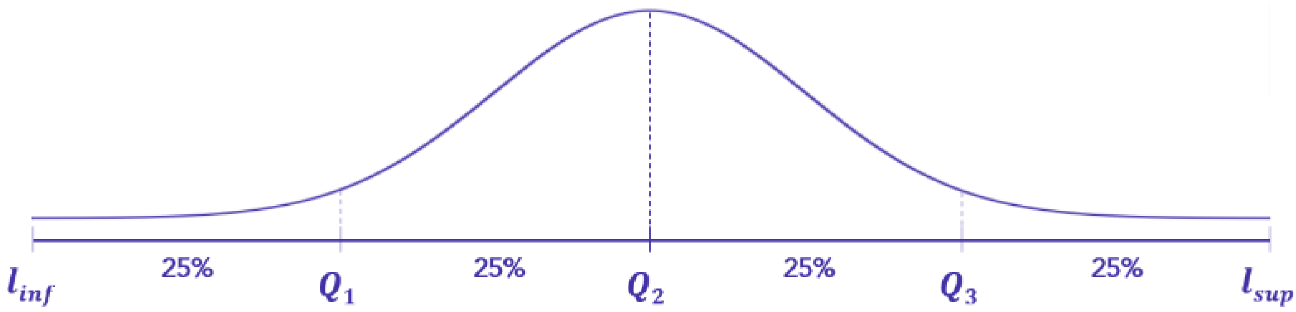
- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a amplitude do conjunto não é alterada.**

2ª Propriedade

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

17) Amplitude Interquartílica

Como já sabemos, denominamos de quartis os valores de uma série que a dividem em quatro partes iguais, isto é, quatro partes contendo o mesmo número de elementos (25%). A imagem a seguir mostra os quartis de uma distribuição hipotética:



Temos, então, 3 quartis ($Q1$, $Q2$ e $Q3$) para dividir uma série em quatro partes iguais:

• $Q1$: o primeiro quartil corresponde à separação dos primeiros 25% de elementos da série;

• $Q2$: o segundo quartil corresponde à separação de metade dos elementos da série, coincidindo com a mediana ($\boldsymbol{Q2} = \boldsymbol{Md}$);

• $Q3$: o terceiro quartil corresponde à separação dos primeiros 75% de elementos da série, ou dos últimos 25% de elementos da série

A amplitude interquartílica (ou distância interquartílica, ou intervalo interquartílico) é o resultado da subtração entre o terceiro quartil e o primeiro quartil:

$$\boldsymbol{A_{IQ} = Q_3 - Q_1}$$

A amplitude semi-interquartílica (ou desvio quartílico) é definida como a metade desse valor, sendo calculada pela expressão apresentada a seguir:

$$\boldsymbol{D_Q = \frac{Q_3 - Q_1}{2}}$$

Reparem que a fórmula da **amplitude interquartílica** (ou distância interquartílica) é muito parecida com a fórmula da **amplitude semi-interquartílico (ou desvio quartílico)**, podendo ser facilmente confundida.

Propriedades da Amplitude Interquartílica

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a amplitude interquartílica (e o desvio quartílico) do conjunto não é alterada.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a amplitude interquartílica (e o desvio quartílico) do conjunto fica multiplicada (ou dividida) por essa constante.**

18) Desvio Padrão

✓ O desvio padrão é definido como sendo a raiz quadrada da média aritmética dos quadrados dos desvios e, dessa forma, é determinado pela raiz quadrada da variância. É uma das medidas de variabilidade mais utilizadas porque é capaz de apontar de forma mais precisa a dispersão dos valores em relação à média aritmética

A fórmula para o cálculo do desvio padrão populacional é:

$$\boldsymbol{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(\boldsymbol{x_i} - \boldsymbol{\mu})^2}{\boldsymbol{n}}}$$

Para o desvio padrão amostral, a fórmula é a seguinte:

$$\boldsymbol{s} = \sqrt{\frac{\sum_{i=1}^{n}(\boldsymbol{x_i} - \overline{\boldsymbol{x}})^2}{\boldsymbol{n} - \boldsymbol{1}}}$$

Símbolo do desvio-padrão populacional:

σ

Símbolo do desvio-padrão amostral:

s

O desvio-padrão será igual a zero quando todos os elementos forem iguais. Se todos os elementos forem iguais, a média aritmética do conjunto será igual ao valor dos elementos e todos os desvios também serão iguais a zero. Logo, o desvio-padrão também será zero.

O desvio-padrão é sempre maior ou igual a zero, isto é, sempre tem valor positivo.

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, o desvio-padrão do conjunto não é alterado.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, o desvio-padrão do conjunto fica multiplicado (ou dividido) por essa constante.**

19) Variância

Fórmula da variância populacional:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} \quad \text{ou} \quad \sigma^2 = \overline{x^2} - (\bar{x})^2$$

Fórmula da variância amostral:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \quad \text{ou} \quad s^2 = \left[\overline{x^2} - (\bar{x})^2\right] \times \left(\frac{n}{n-1}\right)$$

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a variância do conjunto não é alterada.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a variância do conjunto fica multiplicada (ou dividida) pelo QUADRADO dessa constante.**

Distribuição de Frequências

20) Elementos de uma Distribuição de Frequências

✓ As classes são os intervalos nos quais o fenômeno é subdividido. Isto é, podemos dizer que as classes são os intervalos ou subdivisões dos elementos que compõem um conjunto de dados. Repare na tabela a seguir:

Classe	Intervalo	Frequência (f_i)
1ª Classe	$85 \leq x < 100$	5
2ª Classe	$100 \leq x < 115$	5
3ª Classe	$115 \leq x < 130$	12
4ª Classe	$130 \leq x < 145$	10
5ª Classe	$145 \leq x < 160$	7
6ª Classe	$160 \leq x < 175$	9
7ª Classe	$175 \leq x < 190$	2
		$\boldsymbol{n = 50}$

✓ Existem duas maneiras de determinar o número de classes, k, em função do número de dados da tabela, n. A primeira consiste em utilizar a fórmula de Sturges:

$$\boxed{\boldsymbol{k = 1 + 3,3 \times \log n}}$$

✓ Outra abordagem, quando o número de dados é menor ou igual a 50, é por meio da fórmula:

$$\boxed{\boldsymbol{k = \sqrt{n}}}$$

Vamos aproveitar para calcular o número de classes do nosso exemplo:

a) pela fórmula de Sturges:

$k = 1 + 3,3 \times log\ n$
$k = 1 + 3,3 \times log\ 50$
$k = 1 + 3,3 \times 1,7$
$k = 1 + 5,61 = 6,61$

b) pela outra fórmula:

$$k = \sqrt{n}$$

$$k = \sqrt{50} = 7{,}07$$

Amplitude de um intervalo de classe

A amplitude de um intervalo de classe, ou simplesmente intervalo de classe, é a distância entre os limites inferiores (ou superiores) de classes consecutivas. Ela é obtida pela diferença entre dois limites inferiores (ou superiores) consecutivos:

$$\boxed{h = l_{sup} - l_{inf}}$$

Amplitude total

A amplitude total é a diferença entre o limite superior da última classe (limite superior máximo) e o limite inferior da primeira classe (limite inferior mínimo). Portanto, corresponde à diferença entre o último e o primeiro elemento de um conjunto de dados ordenado de forma crescente:

$$\boxed{AT = l_{máx} - l_{mín}}$$

Note que, quando todas as classes possuem a mesma amplitude, também podemos determinar o valor da amplitude total multiplicando o valor do intervalo de classe (h) pela quantidade de classes da distribuição

(k):

$$\boxed{AT = h \times k.}$$

Em nosso exemplo, a amplitude total é calculada da seguinte maneira:

$AT = lmáx - lmín$
$AT = 190 - 85 = 105$

21) Frequência Absoluta Simples

- ✓ A frequência absoluta simples corresponde ao número de observações correspondentes a uma determinada classe ou a um determinado valor

i	Tempos (min)	Frequência (f_i)
1	$85 \vdash 100$	5
2	$100 \vdash 115$	5
3	$115 \vdash 130$	12
4	$130 \vdash 145$	10
5	$145 \vdash 160$	7
6	$160 \vdash 175$	9
7	$175 \vdash 190$	2

A frequência simples é simbolizada por fi . No exemplo anterior, temos: $f1 = 5$, $f2 = 5$, $f3 = 12$, $f4 = 10$, $f5 = 7$, $f6 = 9$ e $f7 = 2$. A soma de todas as frequências é igual ao número total de dados analisados:

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

Para a distribuição em análise, temos:

$$\sum_{i=1}^{7} f_i = f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 = 5 + 5 + 12 + 10 + 7 + 9 + 2 = 50$$

Agora, podemos incluir essa informação na representação tabular:

i	Tempos (min)	Frequência (f_i)
1	$85 \vdash 100$	5
2	$100 \vdash 115$	5
3	$115 \vdash 130$	12
4	$130 \vdash 145$	10
5	$145 \vdash 160$	7
6	$160 \vdash 175$	9
7	$175 \vdash 190$	2
		$\sum_{i=1}^{7} f_i = 50$

22) Frequência Absoluta Acumulada

- ✓ A frequência absoluta acumulada crescente (fac) é o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é: $fac4 = f1 + f2 + f3 + f4 = 5 + 5 + 12 + 10 = 32$, o que significa existem 32 alunos que estudam diariamente entre 85 e 145 minutos (limite superior da quarta classe).

$$\boxed{f_{ac_i} = f_1 + f_2 + f_3 + \cdots + f_i}$$

A frequência absoluta acumulada crescente é calculada de cima para baixo, da seguinte forma:

1) repetimos a frequência absoluta da primeira classe;
2) os demais valores da frequência absoluta são obtidos a partir da soma da frequência acumulada anterior
com a frequência absoluta da classe correspondente;
3) a frequência acumulada crescente sempre termina com o valor de n

i	Tempos (min)	Frequência (f_i)	Frequência Acumulada
1	85 ⊢ 100	5	5
2	100 ⊢ 115	5	10
3	115 ⊢ 130	12	22
4	130 ⊢ 145	10	32
5	145 ⊢ 160	7	39
6	160 ⊢ 175	9	48
7	175 ⊢ 190	2	50
		$\sum_{i=1}^{7} f_i = 50$	

A frequência absoluta acumulada decrescente (fad) é o total das frequências de todos os valores superiores ao limite inferior do intervalo de uma dada classe. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é: $fad4 = f4 + f5 + f6 + f7 = 10 + 7 + 9 + 2 = 28$, o que significa existem 28 alunos que estudam diariamente entre 130 (limite inferior da quarta classe) e 190 minutos.

$$\boxed{f_{ad_i} = f_i + f_{i+1} + f_{i+2} + \cdots + f_k}$$

A frequência absoluta acumulada decrescente é calculada de baixo para cima, da seguinte forma:
1) repetimos a frequência absoluta da última classe;
2) os demais valores da frequência acumulada decrescente são obtidos a partir da soma da frequência
acumulada anterior com a frequência absoluta da classe correspondente;
3) a frequência acumulada decrescente sempre termina com o valor de n.

23) Frequência Relativa Simples

✓ A frequência relativa simples corresponde à proporção de dados existentes em uma determinada classe. Para calcular a frequência relativa de uma classe, dividimos a frequência absoluta simples fi pela frequência total (isto é, dividimos a parte pelo todo)

$$F_i = \frac{f_i}{\sum f_i} = \frac{f_i}{n}$$

Em nosso exemplo, as frequências relativas são:

i	Tempos (min)	Frequência Absoluta (f_i)	Frequência Relativa (F_i)
1	85 ⊢ 100	5	$F_1 = \frac{5}{50} = 0{,}10$
2	100 ⊢ 115	5	$F_2 = \frac{5}{50} = 0{,}10$
3	115 ⊢ 130	12	$F_3 = \frac{12}{50} = 0{,}24$
4	130 ⊢ 145	10	$F_4 = \frac{10}{50} = 0{,}20$
5	145 ⊢ 160	7	$F_5 = \frac{7}{50} = 0{,}14$
6	160 ⊢ 175	9	$F_6 = \frac{9}{50} = 0{,}18$
7	175 ⊢ 190	2	$F_7 = \frac{2}{50} = 0{,}04$

Para representar esses valores em termos de porcentagem, basta multiplicarmos por 100%.

24) Frequência Relativa Acumulada

✓ A frequência relativa acumulada crescente (Fac) é a proporção de valores inferiores ao limite superior do intervalo de uma dada classe. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é: **$Fac4 = F1 + F2 + F3 + F4$ = 10% + 10% + 24% + 20% = 64%**, o que significa que 64% dos alunos estudam diariamente entre 85 e 145 minutos (limite superior da quarta classe).

$$F_{ac_i} = F_1 + F_2 + F_3 + \cdots + F_i$$

A frequência relativa acumulada crescente é calculada de cima para baixo, da seguinte forma:
1) repetimos a frequência relativa da primeira classe;
2) os demais valores são obtidos a partir da soma da frequência relativa acumulada anterior com a frequência
relativa da classe correspondente;

3) a frequência relativa acumulada sempre termina com o valor de 100%.

i	Tempos (min)	Frequência Absoluta (f_i)	Frequência Relativa (F_i)	Frequência Acumulada (F_{ac})
1	85 ⊢ 100	5	10%	10%
2	100 ⊢ 115	5	10%	20%
3	115 ⊢ 130	12	24%	44%
4	130 ⊢ 145	10	20%	64%
5	145 ⊢ 160	7	14%	78%
6	160 ⊢ 175	9	18%	96%
7	175 ⊢ 190	2	4%	100%
		$\sum_{i=1}^{7} f_i = 50$	$\sum_{i=1}^{7} F_i = 100\%$	

✓ A frequência relativa acumulada decrescente (Fad) a proporção de valores superiores ao limite inferior do intervalo de uma dada classe. No exemplo apresentado anteriormente, a frequência acumulada correspondente à quarta classe é: $Fad4 = F4 + F5 + F6 + F7$ = 20% + 14% + 18% + 4% = 56%, o que significa que 56% dos alunos estudam diariamente entre 130 (limite inferior da quarta classe) e 190 minutos.

$$\boxed{\boldsymbol{F_{ad_i} = F_i + F_{i+1} + F_{i+2} + \cdots + F_k}}$$

A frequência relativa acumulada decrescente é calculada de baixo para cima, da seguinte forma:
1) repetimos a frequência relativa da última classe;
2) os demais valoressão obtidos a partir da soma da frequência acumulada anterior com a frequência relativa da classe correspondente;

Vamos ficando por aqui.

Esperamos que tenha gostado do nosso Bizu!

Bons estudos!

Diogo Matias

@oprimoconcursado

Leonardo Mathias

@profleomathias

ESSA LEI TODO MUNDO CONHECE:
PIRATARIA É CRIME.

Mas é sempre bom revisar o porquê e como você pode ser prejudicado com essa prática.

1 Professor investe seu tempo para elaborar os cursos e o site os coloca à venda.

2 Pirata divulga ilicitamente (grupos de rateio), utilizando-se do anonimato, nomes falsos ou laranjas (geralmente o pirata se anuncia como formador de "grupos solidários" de rateio que não visam lucro).

3 Pirata cria alunos fake praticando falsidade ideológica, comprando cursos do site em nome de pessoas aleatórias (usando nome, CPF, endereço e telefone de terceiros sem autorização).

4 Pirata compra, muitas vezes, clonando cartões de crédito (por vezes o sistema anti-fraude não consegue identificar o golpe a tempo).

5 Pirata fere os Termos de Uso, adultera as aulas e retira a identificação dos arquivos PDF (justamente porque a atividade é ilegal e ele não quer que seus fakes sejam identificados).

6 Pirata revende as aulas protegidas por direitos autorais, praticando concorrência desleal e em flagrante desrespeito à Lei de Direitos Autorais (Lei 9.610/98).

7 Concurseiro(a) desinformado participa de rateio, achando que nada disso está acontecendo e esperando se tornar servidor público para exigir o cumprimento das leis.

8 O professor que elaborou o curso não ganha nada, o site não recebe nada, e a pessoa que praticou todos os ilícitos anteriores (pirata) fica com o lucro.

Deixando de lado esse mar de sujeira, aproveitamos para agradecer a todos que adquirem os cursos honestamente e permitem que o site continue existindo.