

Resumo de estatística

Assunto: medidas de dispersão

- Mediana com intervalo:

Para cálculos de mediana com intervalo, deve-se fazer etapas diferentes. Como o exemplo abaixo:

1ª etapa: Organizar os dados de valores iniciais e a frequência ;

2ª etapa: Somar todos os intervalos e dividi-lo por 2, assim achará a mediana nas próximas etapas;

vendas	n° de vendedores
0 ---- 10.000	1
10.000 ---- 20.000	12
20.000 ---- 30.000	27
30.000 ---- 40.000	31
40.000 ---- 50.000	10

Soma: $1+12+27+31+10 = 81$

Dividir: $81/2 = 40,5$

Então, **40,5** é o valor da mediana a ser encontrado.

3ª etapa: Após a mediana encontrada, deve-se somar as frequências até “achar” a mediana calculada no passo anterior. Caso não encontre, mas consiga encontrar o intervalo de vendas, passaremos para a próxima etapa.

vendas	n° de vendedores
0 ---- 10.000	1
10.000 ---- 20.000	+ 12
20.000 ---- 30.000	27 /40
30.000 ---- 40.000	31
40.000 ---- 50.000	10

- Vá somando até chegar perto da mediana no passo anterior, **NÃO PODE ULTRAPASSAR.**

Soma: $1+12+27 = 40$

4ª etapa: Logo após encontrar o intervalo de frequência, deve-se organizar uma regra de três, como:

Intervalo \rightarrow n° de vendedores

$Md - Xi \rightarrow Md_{\text{encontrado}} - n^{\circ} \text{ de vendedores (próximo a } Md \text{ encontrado)}$

Xi: início do intervalo

Md: mediana

$40.000 - 30.000 \rightarrow 31$

$Md - 30.000 \rightarrow 40,5 - 40$



$31 * (md - 30.000) = (40.000 - 30.000) * 0,5$

$Md = \frac{(10.000 * 0,5) + 30.000}{31}$

Md = 30.161,29

- Mediana de maior frequência (com intervalo):

1ª etapa: Encontrar a maior frequência (maior número) e calcular a mediana de cada um (mesmo se aparecer várias vezes), como no exemplo abaixo:

notas	nº de alunos
1,5 --- 3,0	6
3,0 --- 4,5	6
4,5 --- 6,0	6
6,0 --- 7,5	9
7,5 --- 9,0	9

2ª etapa: Após encontrar, calcule a mediana somando as maiores frequências (devem ser nº iguais) e divide por 2, como:

$$\text{Md1} = \frac{6,0+7,5}{2} = 6,75$$

$$\text{Md2} = \frac{7,5+9,0}{2} = 8,25$$

- Desvio padrão sem intervalo:

1ª etapa: Organizar os dados do problema.

MÊS	jan	fev	Mar	abr	mai	jun	jul	ago	set	out	nov	dez
Xi	69	53	41	46	50	40	41	40	42	38	42	46

2ª etapa: Calcular Xi^2 e organizar junto aos demais dados:

MÊS	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
Xi	69	53	41	46	50	40	41	40	42	38	42	46
Xi^2	4761	2809	1681	2116	2500	1600	1681	1600	1764	1444	1764	2116

3ª etapa: Some todos os valores de Xi, depois todos os de Xi^2 :

Soma – xi ($\sum xi$) = 548

Soma – xi^2 ($\sum xi^2$) = 25836

N - 12

4ª etapa: Calcule o desvio padrão pela fórmula:

$$S^2 = \frac{\sum xi^2}{n} - \left(\frac{\sum xi}{n} \right)^2$$

$$S^2 = \frac{25836}{12} - \left(\frac{548}{12} \right)^2$$

$$S^2 = 2153 - 2085,75$$

$$S^2 = 67,25$$

Xi: valor correspondente à variável

Xi²: valor correspondente à variável elevado a 2

N: n° total de casos

S: Desvio Padrão $\rightarrow S = \sqrt{s}$

S²: Variância, ou seja, \sqrt{s}

- Média com intervalo:

1ª etapa: organizar as informações, adicionar em uma tabela f_i e i , e calculando x_i :

aumento de peso em KG	n° de animais
0 --- 1	1
1 --- 2	5
2 --- 3	35
3 --- 4	37
4 --- 5	28

intervalo	frequencia(f_i)	X_i
0 --- 1	1	0,5
1 --- 2	5	1,5
2 --- 3	35	2,5
3 --- 4	37	3,5
4 --- 5	28	4,5

$$X_i = (L_i + L_f)/2$$

L_i : início do intervalo

L_f : final do intervalo

$$X_1 = (0+1)/2 = 0,5$$

$$X_2 = (1+2)/2 = 1,5$$

$$X_3 = (2+3)/2 = 2,5$$

$$X_4 = (3+4)/2 = 3,5$$

$$X_5 = (4+5)/2 = 4,5$$

2ª etapa: calcular $X_i * f_i$:

intervalo	frequencia(f_i)	X_i	$X_i * f_i$
0 --- 1	1	0,5	0,5
1 --- 2	5	1,5	7,5
2 --- 3	35	2,5	87,5
3 --- 4	37	3,5	129,5
4 --- 5	28	4,5	126

3ª etapa: Calcule a média usando a fórmula:

$$X = \frac{\sum(X_i * f_i)}{\sum f_i} \quad \Rightarrow \quad X = \frac{351}{106} = 3,311\text{kg}$$

- Coeficiente de variação

1ª etapa: Organizar os dados, conforme dito no problema, em uma tabela.

	Forn A	Forn B
média	107,5	108
variância	3	1,8

2ª etapa: Usar a formula de coeficiente de variação para calcular o resultado final

$$CV = \frac{S}{x} * 100$$

Onde:

S: variância

X: média

$$CVA = \frac{3}{107,5} * 100 = 1,66\%$$

$$CVB = \frac{1,8}{108} * 100 = 2,79\%$$

Resumo de estatística

Assunto: probabilidade de binominal

1ª etapa: Organizar os dados conforme a logística do problema. Separando a probabilidade que queremos para resolver o problema (**p**), e a que não queremos (**q**), como também, a **condição**. Exemplo – QUESTÃO 6 – LISTA 3:

50 bolas $\left\{ \begin{array}{l} 20 \text{ pretas} \\ 30 \text{ brancas} \end{array} \right. \rightarrow 25 \text{ bolas retiradas com reposição}$

- 2 bolas pretas

N: 25

K: 2

P: 20/50 ou 2/5 ou 0,4

Q: 30/50 ou 3/5 ou 0,6

Onde:

N: quantidade de repetição do ciclo

K: n° de casos

P: Sucesso

Q: Fracasso

- No mínimo, 4 bolas pretas

N: 25

K \geq 4

P: 20/50 ou 2/5 ou 0,4

Q: 30/50 ou 3/5 ou 0,6

2ª etapa: Organizar os dados na fórmula, e calcular:

$$P(k=x) = \binom{n}{k} * P^k * Q^{n-k}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- 2 bolas pretas

$$P(k=2) = \binom{25}{2} * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{25-2}$$

$$\binom{25}{2} = \frac{25!}{2!(25-2)!} \rightarrow \frac{25 \cdot 24 \cdot 23!}{2! \cdot 23!} \rightarrow \frac{600}{2} \rightarrow 300$$

$$P(k=2) = 300 * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{23}$$

$$P(k=2) = 300 * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{23}$$

$$P(k=2) = 300 * 0,16 * 0,000007897$$

$$P(k=2) = 0,0003790$$

- No mínimo, 4 bolas pretas

$$P(k \geq 4) = \binom{n}{k} * P^k * Q^{n-k}$$

$$P(k=0) = \binom{25}{0} * \left(\frac{2}{5}\right)^0 * \left(\frac{3}{5}\right)^{25-0}$$

$$P(k=1) = \binom{25}{1} * \left(\frac{2}{5}\right)^1 * \left(\frac{3}{5}\right)^{25-1}$$

$$P(k=2) = \binom{25}{2} * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{25-2}$$

$$P(k=3) = \binom{25}{3} * \left(\frac{2}{5}\right)^3 * \left(\frac{3}{5}\right)^{25-3}$$

$$\binom{25}{0} = \frac{25!}{0!(25-0)!} \rightarrow \frac{25!}{25!} \rightarrow 1$$

$$\binom{25}{1} = \frac{25!}{1!(25-1)!} \rightarrow \frac{25.24!}{1!24!} \rightarrow 25$$

$$\binom{25}{2} = \frac{25!}{2!(25-2)!} \rightarrow \frac{25.24.23!}{2!23!} \rightarrow \frac{600}{2} \rightarrow 300$$

$$\binom{25}{3} = \frac{25!}{3!(25-3)!} \rightarrow \frac{25.24.23.22!}{3!22!} \rightarrow \frac{600}{2} \rightarrow 4600$$

$$P(k=0) = 1 * \left(\frac{2}{5}\right)^0 * \left(\frac{3}{5}\right)^{25}$$

$$P(k=0) = 1 * 1 * \left(\frac{3}{5}\right)^{25}$$

$$P(k=0) = 0,000002843$$

$$P(k=1) = 25 * \left(\frac{2}{5}\right)^1 * \left(\frac{3}{5}\right)^{24}$$

$$P(k=1) = 25 * \left(\frac{2}{5}\right)^1 * \left(\frac{3}{5}\right)^{24}$$

$$P(k=1) = 25 * 0,4 * 0,000004738$$

$$P(k=1) = 0,00004738$$

$$P(k=2) = 300 * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{23}$$

$$P(k=2) = 300 * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{23}$$

$$P(k=2) = 300 * 0,16 * 0,000007897$$

$$P(k=2) = 0,0003790$$

$$P(k=3) = 4600 * \left(\frac{2}{5}\right)^3 * \left(\frac{3}{5}\right)^{22}$$

$$P(k=3) = 4600 * \left(\frac{2}{5}\right)^3 * \left(\frac{3}{5}\right)^{22}$$

$$P(k=3) = 4600 * 0,064 * 0,000013162$$

$$P(k=3) = 0,003874$$

3ª etapa: Usar a ultima fórmula para descobrir a probabilidade:

$$P(\text{desejado}) + P(\text{não desejado}) = 1$$

- 2 bolas pretas

$$P(k=2) = 300 * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{23}$$

$$P(k=2) = 300 * \left(\frac{2}{5}\right)^2 * \left(\frac{3}{5}\right)^{23}$$

$$P(k=2) = 300 * 0,16 * 0,000007897$$

$$P(k=2) = 0,0003790$$

- No mínimo, 4 bolas pretas:

$$P(\text{desejado}) + P(\text{não desejado}) = 1$$

$$P(k \geq 4) + P(k < 4) = 1$$

$$P(k \geq 4) = 1 - (P(k=0) + P(k=1) + P(k=2) + P(k=3))$$

$$P(k \geq 4) = 1 - (0,000002843 + 0,00004738 + 0,0003790 + 0,003874)$$

$$P(k \geq 4) = 1 - 0,004303223$$

$$P(k \geq 4) = 0,9956$$

Resumo de estatística

Assunto: Técnicas de amostragem

- Amostragem aleatória simples

Exemplo 1: em uma empresa há 90 computadores. Obtenha uma amostra correspondente a 15%.

OBS: use a tabela de números aleatórios, a partir da 18ª linha, da esquerda para a direita, para obter os elementos dessa amostra.

18º: 610201817392606673585334426826383403274496044665

População: o todo	amostra: parte do todo
-------------------	------------------------

1ª etapa: é necessário fazer a relação entre as informações dos problemas, como:

Temos 90 pc – então a numeração de relação seria de 01 à 90.

2ª etapa: Por meio de uma regra de 3, descobriremos com a numeração de relação, a porcentagem da amostra, como:

$$\begin{array}{ll} 90 \rightarrow 100\% & 100x = 90 \cdot 15 \\ X \rightarrow 15\% & X = \frac{90 \cdot 15}{100} \rightarrow x = 13,5 \text{ ou } 14 \end{array}$$

Obs: caso tenha o resultado quebrado, pode-se arredondar.

3ª etapa: Usamos a tabela de n° aleatórios para chegarmos a nossa amostragem, pegamos a quantidade de casas da nossa população e agrupamos esses números. Temos que ter cuidado para não selecionarmos algum n° que ultrapasse o limite da nossa população. E, “pegaremos” a quantidade da nossa amostra (que obtemos na regra de 3 – passo anterior, lembrando sempre do agrupamento de casas, como:

Ou seja, pegamos 14 números dessa tabela de n° aleatórios e sem ultrapassar 90

Resultado: 61 02 01 81 73 60 66 73 58 53 34 42 68 26 38

Resumo de estatística

Assunto: Técnicas de amostragem

- Amostragem estratificada

Exemplo 2: em uma empresa há 90 computadores, dos quais 36 são Windows e os demais são Linux. Obtenha uma amostra contendo 9 computadores.

OBS: use a tabela de números aleatórios, a partir da 18ª linha, da esquerda para a direita, para obter os elementos dessa amostra.

18º: 610201817392606673585334426826383403274496044665

1ª etapa: Saber a quantidade proporcional do todo e da amostra. Separando as informações.

90 pc { 36 windows
54 linux

Duas possibilidades para 2ª etapa, sendo:

2ª etapa - 1: Pela amostra, devemos separar para cada proporção do problema e usamos regra de 3, para saber: quantidade da mostra separada por grupo.

p/ Windows, temos:

$$90 \rightarrow 9 \quad 90x = 9 \cdot 36$$

$$36 \rightarrow x \quad x = \frac{9 \cdot 36}{90} \rightarrow x = 3,6 \text{ ou } 4$$

p/ linux, temos:

$$90 \rightarrow 9 \quad 90x = 9 \cdot 54$$

$$54 \rightarrow x \quad x = \frac{9 \cdot 54}{90} \rightarrow x = 5,4 \text{ ou } 5$$

*Escolha qual dos dois você vai arredondar para + e qual para -, caso você some ambos os resultados, e ultrapasse o valor da amostragem de problema.

2ª etapa - 2: podemos calcular a porcentagem da amostragem de CADA parte do problema e após achar, “tirar” a quantidade proporcional dessa porcentagem encontrada. Como:

$$90 \rightarrow 100\% \quad 90x = 9 \cdot 100$$

$$9 \rightarrow x \quad x = \frac{9 \cdot 100}{90} \rightarrow x = 10\% \text{ ou } 0,1$$

$$\left\{ \begin{array}{l} 36 w \rightarrow 10\% = 3,6 \text{ ou } 4 \\ 54 l \rightarrow 10\% = 5,4 \text{ ou } 5 \end{array} \right.$$

3ª etapa: usamos a tabela de n° aleatórios, agrupamos pela quantidade do todo (população). E separando a nossa relação de n° do todo , com base em suas proporções e seu intervalo do todo, como:

- A amostra vai de até 90 , então:

Amostra_(01 à 90) { *win* (01 à 36) → 02, 01, 34, 36
 linux (37 à 90) → 61, 81, 73, 60, 66

*Com base no intervalo, selecionamos cada n° dos n° aleatórios da tabela, respeitando: o limite do todo(população), o limite de cada proporção do problema e o limite calculado na etapa 2.

Resumo de estatística

Assunto: Técnicas de amostragem

- Amostragem sistemática

1ª etapa: retirar as informações do problema. Calcular o intervalo entre os números da sequência, como:

População: 300 clientes

Amostra: 15 clientes

Calculo do intervalo:

$$\text{Intervalo} = \frac{\text{população}}{\text{amostra}}$$

$$\text{Intervalo} = \frac{300}{15} = 20$$

2ª etapa: caso a questão não decida e informe o n° que esta “dentro” da amostra, você pode escolher aleatoriamente e somar ao intervalo. Deve-se completar o ciclo da população com a quantidade da amostra, como:

Numeração: 01 à 300

Limite de n°: 20

N° escolhido: 69

Amostra:

69,89,109,129,149,169,189,209,229,249,269,289,009,029,049.