

**1) Responda as questões abaixo:**

a) Quais as 3 grandes áreas de estudo da Estatística? Explique a diferença entre elas apresentando as características de cada área de estudo.

**Descritiva**

Esta é a área da estatística que diz respeito à coleta, caracterização e resumo dos dados. Ao se fazer análises estatísticas o ideal é coletar 100% dos dados e analisar eles como um todo, porém nem sempre isto é possível, como por exemplo: é impossível contar, catalogar e descrever todas bactérias em uma colônia. O objetivo principal da estatística descritiva é organizar e projetar resumos dos dados que representem a população real de forma fidedigna.

**Probabilidade**

É área orientada ao estudo e quantificação numérica de eventos aleatórios.

**Inferência**

É a área destinada a retirar conclusões, não necessariamente matemáticas, sobre os dados analisados.

b) Como podemos classificar as variáveis na Estatística? Explique cada uma das classificações.

**Quantitativas**

Formadas por dados numéricos

Discretas:

O domínio dos dados é formado por um número contável de elementos.

Contínuas:

O domínio dos dados é gradiente em um intervalo real e pode ter incontáveis valores.

**Qualitativas**

Representam categorias dos dados.

Nominal

Não possuem uma ordenação natural entre elas

Ordinal

Possuem uma ordenação natural entre elas

c) O que é um Histograma?

É um tipo de gráfico que mostra como os dados estão distribuídos.

d) O que é um Box-plot? Quais as grandezas que podem ser representadas por um Box-plot?

É um tipo de gráfico usado para a visualização da mediana, assim como a assimetria, variabilidade e os outliers.

e) Explique os 3 padrões de simetria/assimetria em histogramas.

### **Simétrica**

Quando os dados têm moda, mediana e média iguais. Ambas caudas são iguais.

### **Assimétrica negativa.**

Cauda à esquerda (de quem vê o histograma) é maior do que à direita. A média é menor do que a mediana e a mediana é menor do que a moda.

### **Assimétrica positiva**

Cauda à direita. A média é maior do que a mediana e a mediana é maior do que a moda.

### **f) Cite e explique os 3 padrões de curtose.**

**Leptocurtose:** Quando a curva é mais protuberante que a curva Normal

**Mesocurtose:** Quando a curva se assemelha à curva Normal

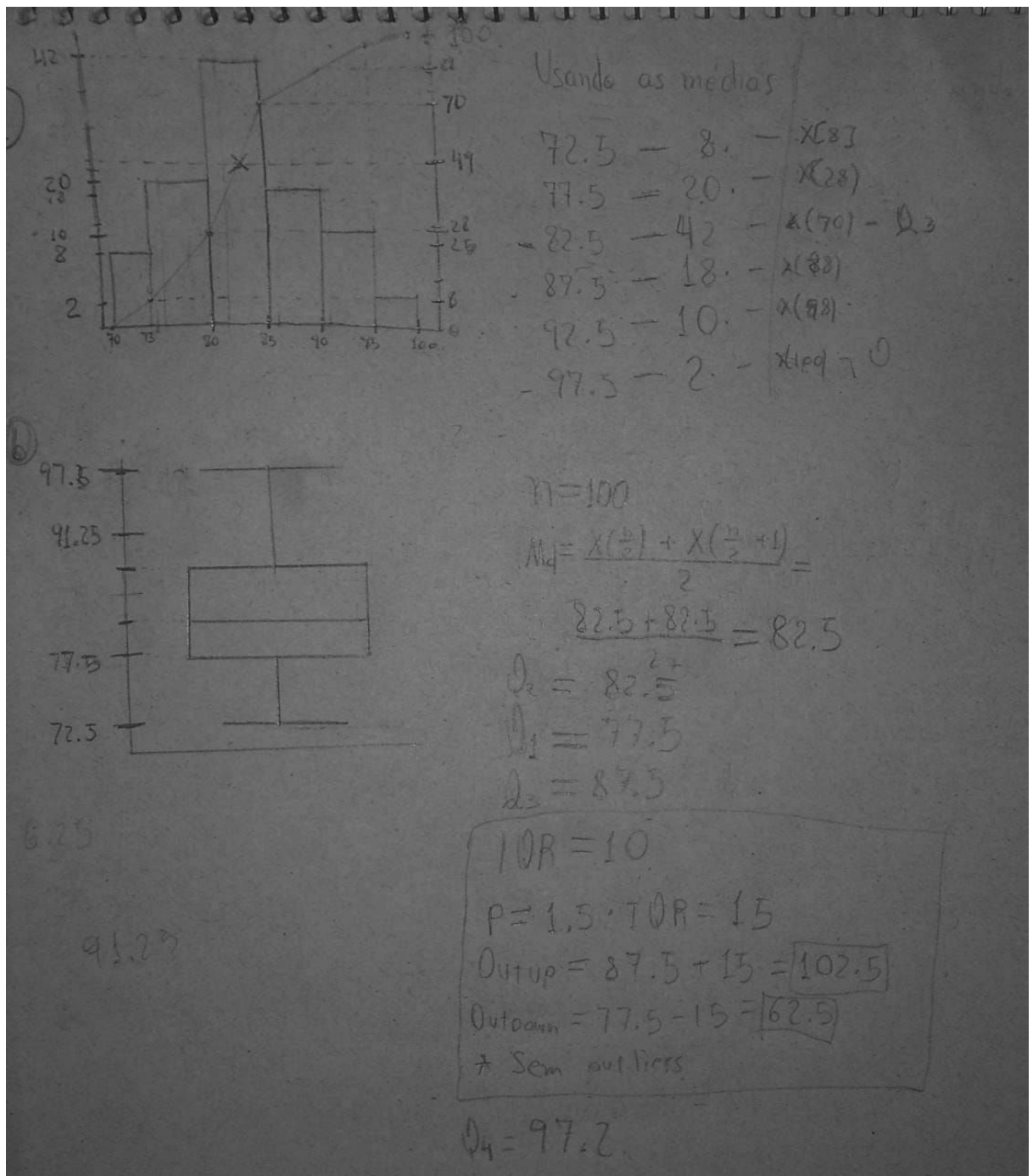
**Platicurtose:** Quando a curva é mais achatada do que a curva Normal

2) (1.16) O índice de germinação é um dos principais fatores para definir a qualidade de sementes. Ele é determinado em experimento científico conduzido pelo fabricante e regulamentado pelos órgãos fiscalizadores. Um fabricante afirma que o índice de germinação de suas sementes de milho é de 85%. Para verificar tal afirmação, uma cooperativa de agricultores sorteou 100 amostras, cada uma de tamanho 100 e anotou a porcentagem de germinação em cada amostra.

Germnação (%)	Frequência
70 + 75	8
75 + 80	20
80 + 85	42
85 + 90	18
90 + 95	10
95 + 100	2

a. Faça uma representação gráfica da tabela

b. Com as devidas suposições, construa o box-plot.



**c. Comente a afirmação do fabricante.**

Mais da metade (mediana) das amostras tem o índice de germinação que ele afirmou ter. E 25% das amostras tiveram germinação entre 72.5%~78.5%, ou seja, um quarto das amostras tiveram 25% de não aproveitamento. O produtor deu uma boa forçada de barra ao dizer que as sementes dele tinha 85% de germinação, mas apenas 30% das amostras tiveram germinação dentro do aproveitamento indicado.

**3) (2.7) Em uma cidade do interior de São Paulo, estima-se que cerca de 20% dos habitantes têm algum tipo de alergia. Sabe-se que 50% dos alérgicos praticam**

esporte, enquanto que essa porcentagem entre os não alérgicos é de 40%. Para um indivíduo escolhido aleatoriamente nessa cidade, obtenha a probabilidade de:

- Não praticar esporte.
- Ser alérgico dado que não pratica esportes.

$P(A) = \frac{20}{100}$  (I)  
 $P(E|A) = \frac{50}{100}$  (II)  
 $P(E|A^c) = \frac{40}{100}$  (III)  
 $P(E^c) = ? \leftarrow (a)$   
 $P(A|E^c) = ? \leftarrow (b)$

Em  $|S| = 100$ , temos

$P(A) = 0.2 \Rightarrow$

	$\neg A$	$A$
$\neg E$	80	20

$P(E|A) = 0.5 \Rightarrow$

	$\neg A$	$A$
$\neg E$		10
$E$		10
	80	20

$P(E|A^c) = 0.4 \Rightarrow$

	$\neg A$	$A$
$\neg E$	48	10
$E$	32	10
	80	20

$P(\neg E) = 0.58$   
 $P(A|E^c) = \frac{10}{38}$

4) (3.16) Uma vacina contra a gripe é eficiente em 70% dos casos. Sorteamos, ao acaso, 20 dos pacientes vacinados e pergunta-se a probabilidade de obter:

Esse é um problema que envolve uma distribuição binomial com  $p = 0.7$

$$q = 0.3$$

$$n = 20$$

$$P(k, n, p) = \frac{n!}{k!(n-k)!} * p^k * (1-p)^{n-k}$$

**a) Pelo menos 18 imunizados**

Temos aqui que k é no mínimo 18, ou seja, [18,20]

Isso implica que a probabilidade nesse caso é a soma das probabilidades de P em [18,20]

$$\sum_{n=18}^{20} P(n, 20, 0.7) = 0.0354831322985$$

**b. No máximo 4 imunizados**

Aqui temos k sendo no máximo 4, ou seja, k está em [0,4]

Podemos usar a mesma lógica aplicada no item a)

$$\sum_{n=0}^4 P(n, 20, 0.7) = 0.0000055502530783$$

**c. Não mais que 3 não imunizados.**

Mesma coisa, mudando só o k que agora está em [0,3]

$$\sum_{n=0}^3 P(n, 20, 0.7) = 5.4269474679 \times 10^{-7}$$

**5) (4.6) O Departamento de atendimento ao consumidor de uma concessionária de veículos recebe, por telefone, as reclamações dos clientes. O número de chamadas dos últimos 30 dias foram anotados e os resultados foram:**

**3,4,5,4,4,5,6,9,4,4,5,6,4,3,6,7,4,5,4,5,7,8,8,5,7,5,4,5,7 e 6.**

**a. Construa a tabela de frequência.**

Chamadas	Freq
3	2
4	9
5	8
6	4
7	4
8	2
9	1

**b. Calcule a média e o desvio padrão.**

$$md = (3+4+5+4+4+5+6+9+4+4+5+6+4+3+6+7+4+5+4+5+7+8+8+5+7+5+4+5+7+6)/30 = 5.3$$

$$dp = \sqrt{\frac{1}{30}(2(3-5.3)^2 + 9(4-5.3)^2 + 8(5-5.3)^2 + 4(6-5.3)^2 + 4(7-5.3)^2 + 2(8-5.3)^2 + (9-5.3)^2)}$$

$$= \sim 1.5$$

**c. Admitindo que cada telefonema acarreta serviços sob garantia avaliados em R\$50,00 por chamada, calcule a média e o desvio padrão das despesas oriundas do atendimento ao consumidor.**

Pegamos a quantidade de chamada em cada mês e multiplicamos pelo preço

Para calcular a média podemos fazer:

$$v = [3,4,5,4,4,5,6,9,4,4,5,6,4,3,6,7,4,5,4,5,7,8,8,5,7,5,4,5,7,6]$$

$$md = \frac{50}{30} * \sum_{n=1}^{30} v[n] = 265$$

Já o desvio padrão podemos fazer:

$$\sqrt{\sum_{n=1}^{30} \frac{1}{30} (v[n] - md)^2} = 76.73285$$