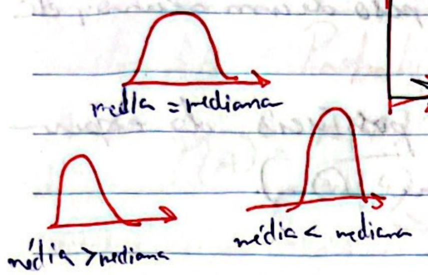
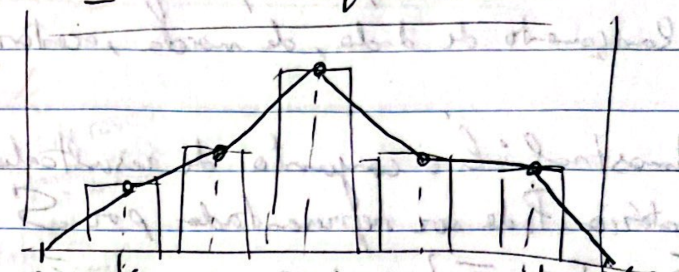
★ Histograma, polígono

↳ gráfico de barras justapostos resultantes da Representação da Tabela de Distribuição de Frequência





o polígono pode ser construído a partir do histograma

★ Média da Amostra $= \bar{x} = \dfrac{\sum x_i \times f_i}{n}$

★ Desvio-padrão da Amostra $= \sqrt{\dfrac{(n \cdot \sum x_i^2 \cdot f_i) - (\sum x_i \cdot f_i)^2}{n(n-1)}}$

★ Diagrama em Caixa (BOXPLOT): descreve simultaneamente várias características importantes de um conjunto de dados, como centro, dispersão, desvio da simetria e, principalmente, identificação de observações não usuais, ou OUTLIERS

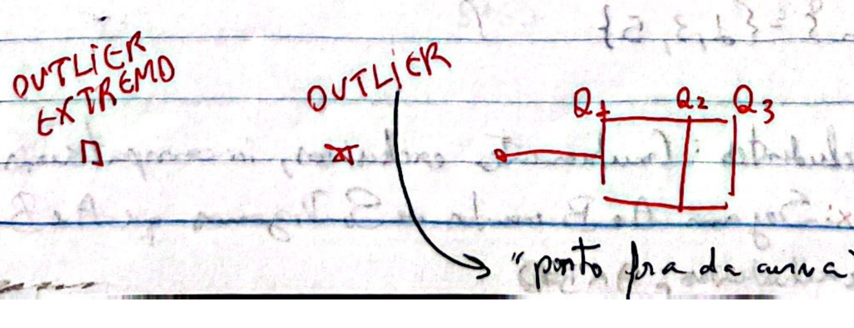
$IQR = Q_3 - Q_1$

↳ etapas:

- colocar os dados em ordem crescente
- determinar Q_1, Q_2, Q_3, menor valor, maior valor, IQR.
- calcular:

$Q_1 - (1,5 \cdot IQR)$ $\quad$ $Q_3 + (1,5 \cdot IQR)$

$Q_1 - (3 \times IQR)$ $\quad$ $Q_3 + (3 \cdot IQR)$



Probabilidade

- E → experimento aleatório (obtenção de uma observação)
- S → espaço amostral (todos os resultados possíveis)
- A → evento (subconjunto gerado a partir de um espaço amostral)
- Probabilidade do evento → entre 0 e 1

$$P(A) = \frac{N(A)}{N(S)}$$

N(A) → nº de eventos favoráveis a A

N(S) → nº de eventos possíveis (S)

▷ Abordagem axiomática:

i) a probab. varia entre 0 e 1

ii) a prob. do evento certo é 1

iii) se A e B forem mutuamente excludentes → $P(A \cup B) = P(A) + P(B)$

▷ Operações com eventos:

- evento união → ocorre se A ocorre, B ocorre ou A e B ocorrem simultaneamente.

$(A \cup B)$

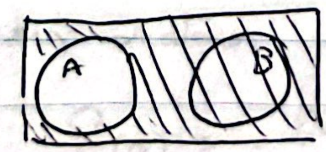
- evento interseção → ocorre se A e B ocorrem simultaneamente

$(A \cap B)$

- evento mutuamente exclusivo → não podem ocorrer na mesma tentativa

$P(A \cap B) = 0$

- evento complementar → A^c → ocorre se A não ocorrer.



A^c

Regra da Adição (Soma):

- $P(A \cup B) = P(A) + P(B) - P(A \cap B) \Rightarrow$ A e B eventos quaisquer
- $P(A \cup B) = P(A) + P(B) \Rightarrow$ A e B disjuntos

Regra da Multiplicação:

- $P(A \cap B) = P(A) \cdot P(B/A) \Rightarrow$ A e B eventos quaisquer
  - $P(B/A)$ → probabilidade condicional
- $P(A \cap B) = P(A) \cdot P(B) \Rightarrow$ A e B independentes

- probabilidade condicional → quando mudamos o nº de eventos possíveis na equação devido à uma condição dada previamente
  - ↳ ex: $\frac{5}{12}$ carros defeituosos.
    prob. do 2º ser defeituoso sendo que sabemos que um defeituoso já foi selecionado $\rightarrow \frac{4}{11} = 36\%$

Probabilidade Condicional:

- probabilidade de um evento B ocorrer, dado que outro evento A já ocorreu (A e B são eventos dependentes)

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

binomial → 2 resultados possíveis (dicotomia)
↳ independência

NCR
↳ combinação

↳ hipergeométrica → pode ter a dicotomia tb
sem reposição

poisson → repetida
uma média

Teorema de Bayes

▷ Partição do Espaço Amostral:

- seja B um evento qualquer, assim:

$$P(B) = \sum P(A_i) \cdot P(B/A_i)$$

▷ O teorema de bayes: considerando a probabilidade de ocorrência de A_i, dado a ocorrência de B:

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)} \qquad \boxed{P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{P(B)}}$$

Variáveis Aleatórias

↳ É quando uma função X que associa um espaço amostral não-numérico a um espaço numérico. = Variável Aleatória X.

- discretos
  - assume números finitos ou infinitos valores enumeráveis
  - é obtida por meio de contagem
- contínuos
  - nºs finitos ou infinitos não-enumeráveis
  - é obtida por meio de medição.

Discretos:

▷ Função ou distribuição de Probabilidade

- determina a probabilidade de uma variável aleatória
- representa variáveis aleatórias
- equação modelo:

$$P_{(x)} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

- dada uma v. a discreta temos que o Valor Esperado, Esperança ou Média são dados por:

$$E(X) = \mu = \sum x \cdot P(x) \quad (x \cdot p)$$

- variância de X é dada por:

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

onde: $E(x^2) = \sum x^2 P(x)$

- desvio padrão de X é dado por:

$$DP(x) = \sigma = \sqrt{Var(x)}$$

Distribuição Binomial:

$(n; p)$

- experimento c/ dois resultados possíveis (dicotomia)
- n realizações independentes do experimento
- ex: lançar um dado 5 vezes e aparecer a face 2 (sucesso), 3 vezes.

$n = 5$, $P = 1/6$ e $X = 3$ → nº de sucessos nas n realizações (variável aleatória).

(combinações) →

$$P(SSSFF) = (1/6)(1/6)(1/6)(5/6)(5/6) = (1/6)^3 (5/6)^2 = 0,003215$$

- p/ o nº de combinações:

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

- p/ cada uma das maneiras:

$$p^x (1-p)^{n-x}$$

- assim, a fórmula é dada por:

$$P(X=x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

Distribuição Hipergeométrica

(N, A, n, x)

- considera n relações dependentes de um experimento
- sem reposição
- ex: retiram, sem reposição, 5 lâmpadas de uma caixa, que contém 10 lâmpadas defeituosas e 30 perfeitas.

experimento
população (N) — sucesso pop. (A) — sucesso na amostra (x) — $\binom{A}{x}$
— fracasso pop. $(N-A)$ — fracasso na amostra $(n-x)$ — $\binom{N-A}{n-x}$

- fórmula geral:

$$P(X=x) = \frac{\overbrace{\binom{A}{x} \cdot \binom{N-A}{n-x}}^{\text{combinações de sucesso e fracasso}}}{\underbrace{\binom{N}{n}}_{\text{nº de amostras possíveis}}}$$

- ex: retira sem reposição 5 lâmpadas de uma caixa, que contém 10 defeituosas e 30 perfeitas ($N-A$). Qual a prob. de encontrar 2 defeituosas?

$N=40$; $A=10$; $n=5$; $x=2$

$$\frac{\binom{10}{2} \cdot \binom{30}{3}}{\binom{40}{5}}$$

outro ex p/ entreinar:

com suspeita → $N = 50$ $n = 10$ (estão sendo testados)

infectados → $A = 20$ $x = 5$ → prob que quero descobrir p/ esse nº de infectados

$$\Rightarrow \frac{\binom{20}{5}\binom{30}{5}}{\binom{50}{10}} = \frac{15504 \cdot 142506}{10272278170} = 0,2151$$

para números inteiros

Distribuição de Poisson

(λ)

- realizações independentes de um experimento onde estamos interessados em observar apenas o sucesso.
- X é o número de vezes que um evento (sucesso) ocorre num intervalo contínuo
- a probabilidade de que o evento ocorra é a mesma em cada intervalo (em geral tempo, área ou espaço). (taxa constante)
- λ = taxa ou média de sucesso no intervalo considerado.

★ Assim, a prob. de x ocorrências em um dado intervalo é:

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

e (Euller) = número irracional $\cong 2{,}71828$

Variáveis Aleatórias Contínuas

- nºs finitos ou infinitos não enumeráveis
- obtida por medição
- ex: tempo de ligação de marketing; notas de 0 a 10; valor total recebido

Contínuas:

1) $0 \leq f(x) \leq 1$

2) $\int f(x)\,dx = 1$

- Esperança/Média:

$$E(X) = \mu = \int x\, f(x)\,dx$$

- Variância:

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

$$E(X^2) = \int x^2 f(x)\,dx$$

- Desvio Padrão:

$$DP(X) = \sigma = \sqrt{Var(X)}$$

Distribuição Exponencial
(λ)

(por Poisson)

$$f(x) = \lambda e^{-\lambda x}$$

$$P(X < x) = F(x) = 1 - e^{-\lambda x}$$

Distribuição Normal
(μ, σ)

pode ter números decimais ("quebrados")

- descrição de diversos processos e fenômenos
- pode ser utilizada p/ aproximar distribuições discretas
- é a base da estatística inferencial

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad Z = \frac{X - \overset{\text{média}}{\mu}}{\underset{\text{desvio padrão}}{\sigma}}$$

μ = média (o valor central da distribuição)
σ = desvio padrão (medida de dispersão que indica o quão espalhados estão os dados em torno da média.
Z = z-score (indica quantos desvios padrão um valor específico está da média).

▷ se a qst pedir prob de um valor $< X$: a tabela normal já entrega a área à esquerda do valor de Z.
▷ se pedir o valor $> x$: precisamos fazer $1 -$ (probabilidade da área à esquerda de Z).