Resumo Provas

ESTATÍSTICA

(beeem resumido)

Medidas e Box Plot:

Desvio Padrão

$$DV = \sqrt{\frac{n(\sum x^2) - (\sum x)^2}{n(n-1)}}$$

$$CV = \frac{DV}{média} \cdot 100\%$$

↳ Coefic. de Variação

▸ **Boxplot:** resume a distribuição de um conjunto de dados.

↳ ordena os valores e calcula os quartis (mediana (Q_2), divide pela mediana → a mediana do inferior é Q_1 e do superior é Q_3).

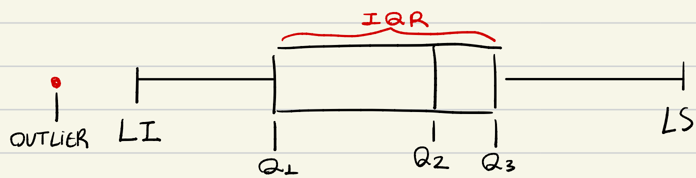
↳ identifica o menor valor, maior valor e $IQR = Q_3 - Q_1$

↳ limite superior = $LS = Q_3 + 1{,}5 \cdot IQR$

↳ limite inferior = $LI = Q_1 - 1{,}5 \cdot IQR$

↳ acima de LS ou abaixo de LI = OUTLIERS

↳ menor valor é o menor do conjunto que está dentro do Limite Inf. e o mesmo p/ o maior valor.



▸ **Histograma:** representação da Tabela de Distribuição de Frequência

"assimétrica à direita"

Probabilidade:

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

$N(A)$ → nº de eventos favoráveis a A

$N(S)$ → nº de eventos possíveis

▸ $P(A \cup B) = P(A) + P(B)$

↳ "ou" / mutuamente exclusivos

▸ $P(A \cap B) = P(A) \cdot P(B)$

↳ "e" / eventos independentes

▸ **Teorema de Bayes:** considera a probabilidade de A_i, dado a ocorrência de B.

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

prob. de A_i ser verdadeira dado que B foi observada

prob. de observar B dado que A_i é verdadeiro (verossimilhança)

Variáveis Aleatórias

↳ uma V.A. associa valores numéricos aos resultados de um experimento.

V.A.

- → **discretos** - conjunto de resultados finito ou enumerável (inteiros)
- → **contínuos** - valores são intervalos ou união de números reais (decimais)

ex: peso [40, 60] tem os pontos nesse intervalo

Contínuos vs Discretos

Contínuos

• 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:

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

- Dist Exponencial (por Poisson)
- Dist Normal

Discretos

dados por:

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

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

↳ onde: $E(X^2) = \sum_i x^2 P(x)$

é dado por:

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

- Dist. Binomial
- Dist. Hipergeométrica
- Dist. de Poisson

//

DISCRETAS

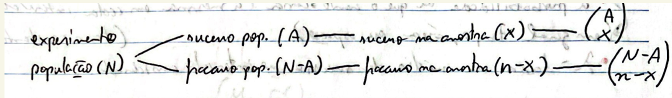
▲ **Binomial:** duas respostas possíveis
tentativas independentes

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

n → nº de tentativas

x ↳ nº de sucessos desejado

p ↳ prob. de sucesso em cada tentativa



Hipergeométrica: nº de sucessos em uma amostra retirada de uma população finita ==sem reposição==.

sucessos na população (elementos da categoria de interesse)

$$P(X=x) = \frac{\binom{A}{x} \cdot \binom{N-A}{n-x}}{\binom{N}{n}}$$

nº de sucessos que quero na amostra

tamanho total da população

tamanho da amostra (os que serão testados)

$\binom{a}{b}$ = coeficiente binomial

Poisson: nº de eventos que ocorrem em um ==intervalo fixo de tempo== ou espaço. Sendo que os eventos ocorrem c/ uma ==taxa constante== e independentemente.

taxa ou média de sucesso no intervalo considerado

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

euller ($\cong 2{,}718$)

x ocorrências em um dado intervalo

//

CONTÍNUAS

ex: tempo até a falha de um equipamento/ tempo de espera até o próx. cliente.

Exponencial: modela o tempo entre eventos em um processo de ==Poisson==, onde os eventos ocorrem ==continuamente== e independentemente

FDP = função densidade de prob. $= f(x) = \underbrace{\lambda e^{-\lambda t}}_{\text{p/ o exato ponto } t\text{:}}, \quad x \geq 0$

euller

taxa de ocorrências por un. de tempo (inverso da média)

$p/\ x < t = f(x) = 1 - e^{-\lambda t}$

$p/\ x > t = f(x) = e^{-\lambda t}$

▶ Uniforme Contínua: todos os valores dentro de um intervalo específico têm a mesma probabilidade de ocorrer.

$$FDP = f(x) = \frac{1}{b-a} \quad , \quad a \le x \le b$$

limite superior ↗ ↘ limite inferior

▶ Normal (ou Gaussiana): modela fenômenos em que a maioria dos valores estão próximos da média c/ uma queda simétrica (formando o sino).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

valores encontrados na tabela da normal p/ os valores de z.

$$Z = \frac{x - \mu}{\sigma}$$

→ média

Z → z-score

σ → desvio padrão

▶ prob valor < X : a tabela normal já entrega a área da esquerda de z

▶ prob valor > X : fazemos 1-(prob. da área à esquerda de z).

μ = valor central da distribuição

σ = o quão espalhados os dados estão da média

Z = indica quantos desvios-padrão um valor está da média

▶ se pedir c/ uma probabilidade, olhamos o z dessa probabilidade na tabela para o valor da probabilidade + 0,5% (pegando o intervalo simétrico ao redor da média ✓)

↳ ex: p/ 25% → 0,255 na tabela → z = -0,6