Resumos MPEI:

Probabilidade: Medida do grau de certeza associado a um resultado proveniente de um fenómeno do acaso / aleatório.

Espaço de Amostragem (S): conjunto de todos os resultados possíveis de uma experiência aleatória

- Discreto → se for contável, ou seja, tem um número finito de elementos
- Contínuo → se não for contável
- Os seus elementos designam-se por resultados

Acontecimento: subconjunto de S

Como determinar uma probabilidade?

- Através da medição
- Através da construção de modelos probabilísticos

Diferentes Abordagens para o cálculo de probabilidades:

→ Teoria Clássica (Laplace)
→ Frequencista
→ Teoria Matemática.

Teoria clássica

$$p(A) = \frac{\text{nº casos Fav}}{\text{nº cPos}}$$

Regras:

- $P(\bar{A}) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Noção frequencista

→ medida empírica de probabilidade

- $f(A) = \frac{\text{nº ocorrência de A}}{\text{Nº de experiências}}$
- $p(A) = \lim_{N \to +\infty} \frac{\text{nº ocorrências de A}}{\text{Nº de experiências}}$
- $0 \leq f(A) \leq 1$
- $\sum_{i=1}^{K} f(A_i) = 1$

Probabilidade condicional:

- Probabilidade de A, sabendo (dado) B é dada por:

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)} , \quad P(B) \neq 0}$$

Exemplo:

- Uma família tem 2 filhos.
- $p(\text{rapaz}) = 0,5$
- Sabendo que um é rapaz, qual a probabilidade de o outro ser também rapaz?

Casos possíveis

1 2
MF
MM
FF
FM

$$p(x) = \frac{P(R \cap R)}{P(R_1)} = \frac{1/4}{3/4} = \frac{1}{3}$$

pelo menos 1 rapaz

Regra da Cadeia/multiplicação:

$$P(A.B.C) = P(A) \cdot P(B|A) \cdot P(C|A.B)$$

Ex: p("Universidade de Aveiro") = p("Universidade") · P("de" | "Universidade") · P("Aveiro" | "Universidade de").

Ex:

- existem 2 caixas (X e Y) com bolas brancas (B) e pretas (P)
- Caixa (X) → 4 brancas + 5 pretas
- Caixa (Y) → 3 brancas + 6 pretas)

Escolhendo uma urna ao acaso, qual a probabilidade de sair uma branca?

$$p(\text{branca na } x) = P(X) \cdot P(B|X) = \frac{1}{2} \cdot \frac{4}{9} = \frac{4}{18}$$

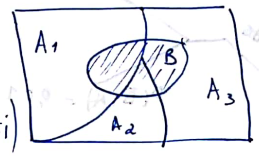
$$p(\text{branca na } Y) = P(Y) \cdot P(B|Y) = \frac{1}{2} \cdot \frac{3}{9} = \frac{3}{18}$$

$$p(\text{branca}) = \frac{4}{18} + \frac{3}{18} = \frac{7}{18}$$

Scanned by CamScanner

Lei da probabilidade total:

$$P(B) = P(A_1)\cdot P(B|A_1) + P(A_2)\cdot P(B|A_2) + P(A_3)\cdot P(B|A_3) \Bigg\} \; P(B) = \sum_i P(A_i)\cdot P(B|A_i)$$



Condicionamento inverso / Regra de Bayes:

• E se quisermos $P(\underbrace{\text{"caixa X"}}_{\text{causa}} \mid \underbrace{\text{"bola branca"}}_{\text{efeito}})$

$$P(\text{causa} \mid \text{efeito}) = \frac{P(\text{efeito} \mid \text{causa}) \times P(\text{causa})}{P(\text{efeito})}$$

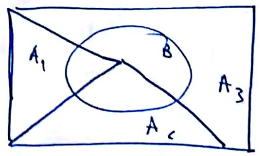
$$P(\text{caixa X} \mid \text{bola branca}) = \frac{P(\text{branca} \mid X) \cdot P(X)}{P(\text{branca})} =$$

$$= \frac{4/9 \cdot \frac{1}{2}}{\frac{7}{18}} = \frac{4/18}{\frac{7}{18}} = \frac{4}{7}$$

Regra de Bayes

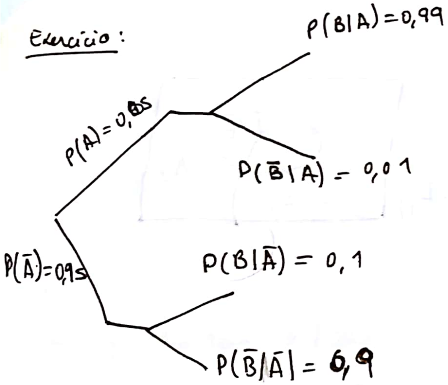
$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)\cdot P(B|A_i)}{P(B)}$$

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



Scanned by CamScanner

Exercício:

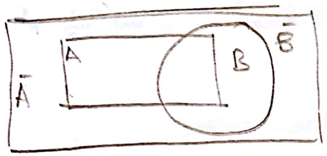


$$\boxed{P(A|B) = ?}$$

Pela regra de Bayes:

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

$$= \frac{0{,}99 \cdot 0{,}05}{P(B)} =$$



- $P(B) = P(B|\bar{A}) \cdot P(\bar{A}) + P(B|A) \cdot P(A)$

$$= 0{,}1 \cdot 0{,}95 + 0{,}99 \cdot 0{,}05 =$$

$$= 0{,}095 + 0{,}0495 =$$

$$= 0{,}1445$$

$$\boxed{P(A|B) = \frac{0{,}0495}{0{,}1445} = 0{,}3426}$$

- Voltamos à família de 2 filhos:

Sabendo que 1 rapaz abre a porta, qual a probabilidade de outro ser também rapaz?

$$P(P_R|A_R) = \frac{P(A_R \cap P_R)}{P(A_R)} = \frac{1/4}{\frac{1}{2}} = \frac{2}{4} = \frac{1}{2}$$

Abre	Porta
M	M
M	F
F	M
F	F

Scanned by CamScanner

Acontecimentos independentes:

Se, e só se: $\boxed{P(A\cap B) = P(A)\cdot P(B)}$

$\Rightarrow$ implica que $P(A|B) = P(A$, mas não é definição

Independência 2 a 2 vs Independência:

- 2 lançamentos de moeda
- A: moeda$_1$ cara (Head)
- B: moeda$_2$ cara
- C $\Rightarrow$ moeda$_1$ = moeda$_2$

$C_p = \{HT, \underline{HH}, TH, \underline{TT}\}$

$P(A) = \frac{1}{2} = P(B)$

$P(C) = \frac{2}{4} = \frac{1}{2}$

- $P(C\cap A) = 1/4 = \underbrace{P(C)\cdot P(A)}$

C e A independente

- $P(C\cap B) = \frac{1}{4} = \frac{1}{2}\times\frac{1}{2}$ ✓

C e B são independentes

- $P(A\cap B) = \frac{1}{4} = \frac{1}{2}\cdot\frac{1}{2}$ ✓

A e B são independentes

$P(C\cap B\cap A) = \frac{1}{4}$

$P(C)\cdot P(B)\cdot P(A) = \frac{1}{8} \neq P(C\cap B\cap A)$

Logo: Independência 2 a 2 não significa independência

Acontecimentos mutuamente exclusivos

- 2 acontecimentos mutuamente exclusivos que têm probabilidade $\neq 0$ não podem ser independentes

$P(A\cap B) = 0 \xrightarrow{\text{implica}} P(A)\cdot P(B) = 0$, o que não é verdade pois nenhum pode ter probabilidade nula

Experiência de Bernoulli:

Realização de uma experiência e registo se há sucesso, ou se se falha

Probabilidade de K sucessos em M ensaios independentes?

$p(\text{sucesso}) = p$

$p(\text{falha}) = 1 - p$

• Probabilidade de K sucessos e $M-K$ falhas é $\Big\}$ $p^K \cdot (1-p)^{M-K}$

• Os K sucessos em M, podem ocorrer de C^M_K maneiras

$$\boxed{P_M(K) = C^M_K \cdot p^K \cdot (1-p)^{M-K}}$$

$$C^M_K = \binom{M}{K} = \frac{M!}{K! \cdot (M-K)!} = \binom{M}{M-K}$$

Ponto de vista frequencista:

$$P(A|B) \approx \frac{K_{A\,e\,B}/N}{K_B/N} = \frac{K_{A\,e\,B}}{K_B}$$

→ número de ocorrência de "A e B" = f absoluta (AB) = f_{AB}

Axiomas:

• logicamente independentes

• compatíveis

① $P(A) \geq 0$

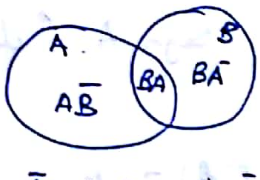
② $P(S) = 1$

③ Se $P(A) \cap P(B) = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

Teoremas ⇒ obtidos dos axiomas, por dedução.

Teoremas:

① $P(\bar{A}) = 1 - P(A)$



② $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, vem de $A \cup B = A\bar{B} \cup AB \cup \bar{A}B$

③ $P(A \cup C | B) = P(A|B) + P(C|B) - P(A \cap C | B)$

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

$$P(A \cup C | B) = \frac{P((A \cup C) \cap B)}{P(B)} =$$

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

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

$$= P(A|B) + P(C|B) - P(AC|B)$$

c.q.d

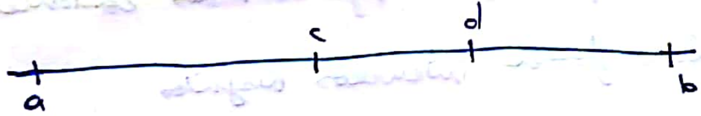
- Axiomática, verificada por:
  - Lei de Laplace
  - Definição frequencista

Probabilidade em espaços de amostragem não contáveis:

- Há experiências em que os resultados possíveis ocorrem no domínio de IR (infinito)

Assim, uma das situações mais simples é considerar intervalos em que qualquer valor é igualmente provável. ex: $S = [a, b]$

- Exemplo:



A → número de $[a, b]$ pertencer a $[d, c]$
$=]d, c]$

$$P(A) = \frac{d - c}{b - a}$$

$P(\text{qlq ponto } x \in [a, b]) = 0$!

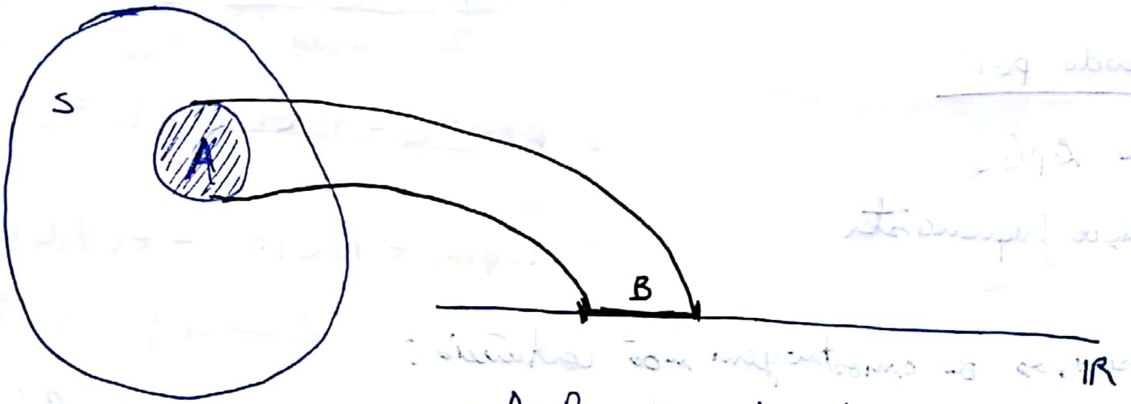
Exemplo: Atraso $= [0,90]$

Seja A. "chegar dentro da tolerância" $= [0,15[$

$$P(A) = \frac{15-0}{90-0} = 0,17$$

Variáveis aleatórias:

- funções que mapeiam o espaço amostral na reta real
- é o resultado numérico dos nossos experiências aleatórios



- A e B são acontecimentos equivalentes.

- Dividem-se em:

① Discretas: valores finitos, ou infinitos mas contáveis (ex 1, 2, ...)

② Contínuas: valores formam intervalos disjuntos.

③ Mistas: Discretas e Contínuas

Função (massa) de probabilidade:

- $P_X(x_i) = P(X=x_i)$, onde $P(X=x_i) = P(w: X(w)=x_i)$, ou seja, A probabilidade do evento cujos resultados w satisfazem $X(w)=x_i$

- $P_X(x_i) \geq 0$

- $\sum_i P(x_i) = 1$

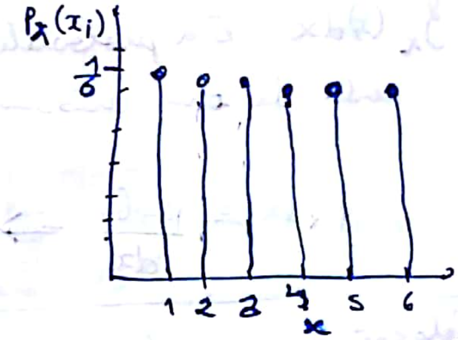
Scanned by CamScanner

Ex. de função de probabilidade:

Lançamento de dado equilibrado e X igual ao nº que sai
└→ variável aleatória discreta.

Função de probabilidade ⇒ $x_i = \{1, 2, 3, 4, 5, 6\}$

$P_X(x_i) = 1/6$

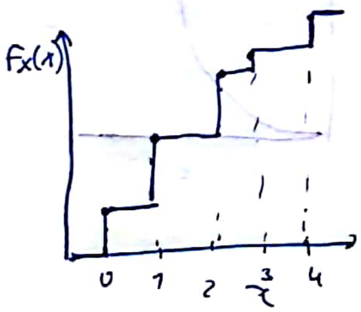


Função distribuição acumulada (discreta)

- $F_X(x) = P_X(X \leq x) = \sum_{i: x_i \leq x} P_X(x_i)$

- $\lim_{x \to -\infty} F_X(x) = 0$ / $\lim_{x \to +\infty} F_X(x) = 1$

- função escada ⟶



Variáveis aleatórias contínuas:

- também podem ser especificadas pela sua função distribuição acumulada:

$F_X(x) = Prob(X \leq x)$
- → contínua
- → ≥ 0
- → ≤ 1

- $a < b \Rightarrow F_X(a) \leq F_X(b)$
- $P[a < X \leq b] = F_X(b) - F_X(a)$

- Podem ser especificadas pela função de densidade de probabilidade ⟶ $f_X(x) = \frac{d\,F_X(x)}{dx}$

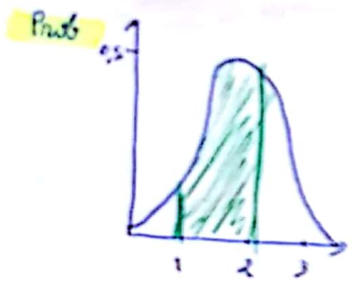
Função de densidade de probabilidade:

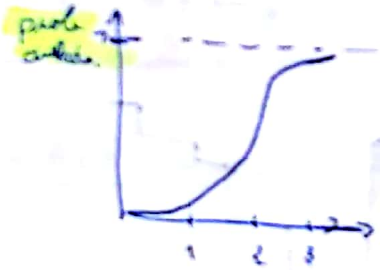
- $f_X(x) \neq$ probabilidade
- $P(a < X < b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx$
- $f_X(x)\,dx$ é a probabilidade da variável $X \in$ ao intervalo $(x, x+dx)$ nesse dx de comprimento infinitesimal

$$f_X(x) = \frac{\text{prob}}{dx} \longrightarrow \text{densidade}$$

Relação com a função de distribuição:

- $F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$





- $P(1 \leq x \leq 2) =$ Área a verde
- Área total da curva $= 1$

Média:

- $\sum_i p(x_i) \cdot x_i$ ⟶ Para um lançamento de 1 dado
  $p(1)\times 1 + p(2)\times 2 + \dots + p(6)\times 6$

Valor esperado:

- O valor esperado de X é o valor médio de X ao repetirmos as experiência indefinidamente

$$= \lim_{n \to +\infty} \frac{X_1 + X_2 + \dots + X_n}{n}$$

, se o limite não existir então não há valor esperado.

Só existe se X_i tiver limite inferior e superior finito (ex: peso ≥ 0)

Scanned by CamScanner

Seja x_i os m diferentes valores que X_i pode assumir e $K_{i,n}$ o nº de vezes que ocorre cada x_i, tem-se

- $\lim_{n \to \infty} \frac{x_1 K_{1,n} + x_2 K_{2,n} + \dots + K_{m,n}}{n}$

- $\sum_{i=1}^{m} x_i \, p(X = x_i)$

Valor esperado ⟹ enganador, pois na realidade não é algo que devemos esperar que ocorra, sendo que muitas vezes é improvável/impossível.
(ex: valor esperado do lançamento de um dado de 6 faces é 3,5)

→ designa-se por $E[X]$
- contínuo → $E[X] = \int_{-\infty}^{+\infty} x f_X(x)\,dx$
- discreto → $E[X] = \sum_i x_i \, p(x_i)$

→ operador linear ⟹
- $E[aX] = a\,E[X]$
- $E[X+Y] = E[X] + E[Y]$
- $E[X+c] = E[X] + c$

Exemplo:

x_i	$p_X(x_i)$	$x_i \, p(x_i)$
-1	0,1	-0,1
0	0,2	0
1	0,4	0,4
2	0,2	0,4
3	0,1	0,3

$E[X] = 0{,}4 + 0{,}4 + 0{,}3 - 0{,}1 = 1$

Variância:

Mede-nos a dispersão.

- Ideia base: Usar a diferença dos valores da variável para fazer a sua média

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

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

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

Desvio padrão:

- σ
- $\sigma = \sqrt{Var(X)}$

Slide 37

Variância - propriedades:

X é uma variável aleatória
c é uma constante

- $Var(X+c) = Var(X)$
- $Var(cX) = c^2 Var(X)$

E[X] → Valor médio de X
→ centro de gravidade da função massa de probabilidade (discreta) ou da função densidade de probabilidade

Desvio padrão/variância:

→ dá a dispersão de uma variável aleatória
→ Se for 0, então não existe var. aleatória (todos valores = média)

...tos de ordem n:

Caso discreto $\longrightarrow$ $m_n = E[X^n] = \sum_i x_i^n \, p_X(x_i)$

Exemplo dos dados:

$$E[X^2] = 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + \dots + 6^2 * \frac{1}{6} = 15,17$$

Momentos centrados de ordem n:

- Resulta na generalização da variância

$$E[(X - E[X])^n] = \sum_i (x_i - E[X])^n \, p(x_i)$$

- Variância $\longrightarrow$ momento centrado de 2ª ordem

Distribuições Discretas

- Funções massa e densidade de probabilidade podem assumir $\neq$ formas:

Distribuições:

1. Discretas:
   - Bernoulli
   - Binomial
   - Poisson
2. Contínuas:
   - Uniforme
   - Normal

Distribuições Discretas:

Variável de Bernoulli (I_A)

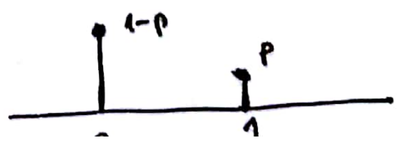
$$\underbrace{I_A(w)}_{\text{indicadora de ocorrência}} = \begin{cases} 1, & \text{se } w \in A \\ 0, & \text{se } w \notin A \end{cases}$$

Exemplo:

- $S_I = \{0, 1\}$
- $p = P_r(A)$

$\longrightarrow$ $P_I(1) = p$

$P_I(0) = 1 - p$



Valor esperado:

- $E[I] = \sum_i x_i \, p(x_i) = \cancel{0 \times (1-p)} + 1 \times p = p$

Variância:

$$Var(I) = E[I^2] - (E[I])^2 = 0^2 \cdot (1-p) + 1^2 \cdot p - (p)^2 =$$

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

$\boxed{\text{Variável Binomial}}$

- $X = \sum_{j=1}^{n} I_j \longrightarrow S_X = \{0, 1, 2, \ldots, n\}$

X ↓ número de sucessos em n experiências

- $p_X(k) = P_r(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

- $F_X(x) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$

Média:

$$E[X] = E[\sum X_i] = \sum E[X_i] = \underline{\underline{n p}}$$

Variância

$$Var(\sum X_i) = \sum Var(X_i) = \underline{\underline{n}} \, p(1-p)$$

Exemplos:

- Nº de peças defeituosas num lote de 50 peças

...plo:

- De acordo com os censos, 6% dos trabalhadores de Aveiro estão desempregados
- Ao fazer um inquérito via telefone, random, qual a probabilidade de termos 2 ou menos trabalhadores desempregados, numa amostra de 20 pessoas?

$p(\text{desemprego}) = 0{,}06$

$n = 20$

$p(\text{emprego}) = 1 - 0{,}06 = 0{,}94$

x = nº sucesso

$P(x \leq 2) = p(x=0) + p(x=1) + p(x=2)$

$P(x=0) = \binom{20}{0}(0{,}06)^0 * (0{,}94)^{20} = \frac{20!}{0!(20-0)!} \cdot 1 \cdot (0{,}94)^{20} = 0{,}94^{20} = 0{,}2901$

$p(x=1) = \ldots$

$p(x=2) = \ldots$

$P(x \leq 2) = 0{,}2901 + 0{,}3703 + 0{,}2246 = 0{,}8850$ //

Variáveis geométricas

$X \rightarrow$ nº de vezes que é necessário repetir-se uma experiência de Bernoulli até obter um sucesso

- $p(\text{sucesso}) = p$
- $P_X(k) = p(1-p)^{k-1}$
- $F_X(x) = \sum_{k=0}^{x} p(1-p)^{k-1}$

Média $\Rightarrow E[X] = \frac{1}{p}$

Variância $\Rightarrow (1-p)/p^2$

Exemplo:

- $p(\text{sucesso}) = 0{,}1$
- $p(\text{necessitar menos de 3 chamadas para ter sucesso}) = ?$

$$P(<3) = p(1) + p(2) = p(1-p)^{1-1} + p(1-p)^{2-1} = p + p(1-p) =$$

$$= p + p - p^2 = 2p - p^2 = 0{,}2 - 0{,}01 = 0{,}19$$

Distribuição de Poisson

- Surgiu devido à precariedade da fórmula binomial para valores de n elevados

Consideremos que existe uma variável Binomial, n cresce, p decresce e $np \to \lambda > 0$
Podemos aproximar, para n grande:

$$\boxed{p \cong \frac{\lambda}{n}} \quad \boxed{1 - p \simeq 1 - \frac{\lambda}{n}}$$

$$\boxed{P_X(K) = \frac{\lambda^K e^{-\lambda}}{K!}}$$

→ Função de massa de probabilidade da distribuição de Poisson, com $K = 0, 1, 2, \ldots$

Distribuição para os vários valores de λ:

Função probabilidade → $P_X(K) = P_r(X=k) = \dfrac{\lambda^K e^{-\lambda}}{K!}$

Média → $E[X] = \lambda$

Variância → $V(X) = \lambda$

Slide 30

Variável de Poisson

- foca-se no nº de ocorrências (discreto), num intervalo, ou contínuo
- Esta distribuição não tem um nº de experiências (n)

Distribuição de Poisson → usada quando n é grande e p é pequeno (eventos raros), aproximando a distribuição binomial

Regra: Se $n > 20$ e $np <= 7$ → usa-se a aproximação de Poisson !

$\boxed{\text{Como aproximar a distribuição binomial por Poisson?}}$

① Calcular média binomial: $\mu = np$

② $\lambda = E[x] = \mu$

③ usar: $P_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}$

Exemplo:

Chegam 3,2 clientes em 4 minutos, em média.
Probabilidade de chegarem mais de 7 clientes em 4 minutos?

$P($"$X > 7$ clientes / 4 minutos"$) = ?$

$\lambda = ?$

$\lambda = 3,2$ [clientes em 4 minutos]

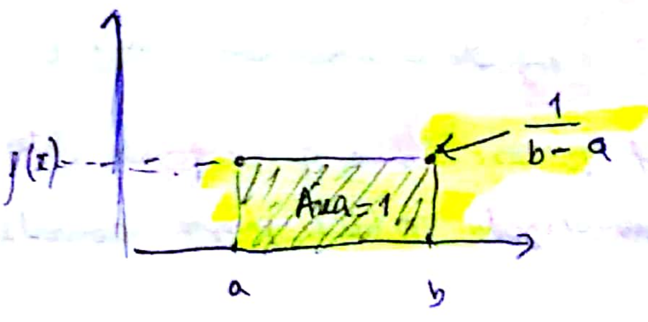
$$P(>7) = 1 - \sum_{k=0}^{7} \frac{0,32^k \cdot e^{-0,32}}{k!} = 0,0169$$

Distribuições Contínuas

Variável aleatória uniforme:

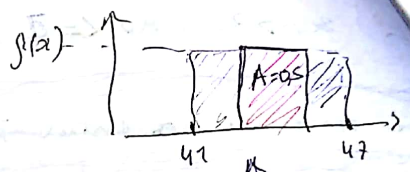
$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{caso contrário} \end{cases}$$

$$F_X(x) = \begin{cases} 0, & x < a \\ \frac{1}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$



$E[X] = \frac{a+b}{2}$

$Var(X) = \frac{(b-a)^2}{12}$



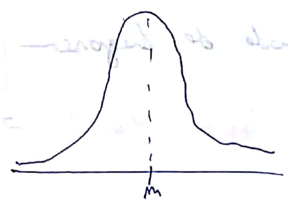
Exemplo:

$P(42 \leq X \leq 45)$, com $U(41, 47)$

$$P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b-a} = \frac{45-42}{47-41} = \frac{1}{2}$$

Variável aleatória Normal/gaussiana:

Se: $\boxed{f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}}$

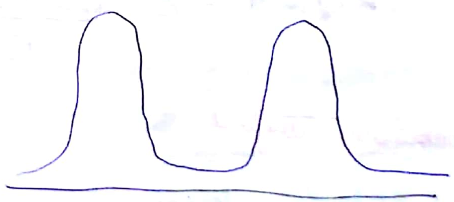


- Notação $\rightarrow \boxed{N(m, \sigma^2)}$

m = média

σ = desvio

- $E[X] = m = \mu$
- $Var[X] = \sigma^2$



- Nota: se variamos 2 parâmetros $\longrightarrow$

gaussiana normalizada

- Número infinito de combinações de m e σ $\longrightarrow$ família infinita de curvas
- Há um mecanismo que permite obter uma distribuição única (gaussiana normalizada $N(0,1)$) que converte qualquer distribuição norm

...lo: subtrair a média e dividir pelo desvio padrão

$$\boxed{Z = \frac{X - m}{\sigma}}$$

Função densidade de probabilidade $\Longrightarrow$ $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{(x)^2}{2}\right)}$

Função distribuição acumulada $\longrightarrow$ $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{(t)^2}{2}} dt$

↳ pode, também, ser expressa:

$F_X(x) = \Phi\left(\frac{x - m}{\sigma}\right)$ $\quad Q(x) = 1 - \Phi(x)$

slide 12

Exemplo: Uma empresa tem uma procura anual X, que segue uma distribuição normal $N(75, 10^2)$. $P(78 < X < 80) = ?$

$m = 75$
$\sigma = 10$

$$P[78 < X < 80] = P\left[\frac{78 - 75}{10} \leq U_X \leq \frac{80 - 75}{10}\right] =$$

$$= P(0{,}3 \leq U_X \leq 0{,}5) = Q(0{,}3) - Q(0{,}5) = 0{,}074$$

Distribuição Binomial e a normal

$m = np$
$\sigma = n \cdot p(1 - p)$, com n não muito pequeno e n elevado.
Aproximado por:

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(k - m)^2}{2\sigma^2}} \longrightarrow$$ Distribuição normal, desde que $m = np$ e variância $= np(1-p)$

Distribuição exponencial:

- $f_X(x) = \begin{cases} 0, & \text{se } x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases}$

- $F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$

slide 18

- $E[X] = \frac{1}{\lambda}$
- $Var(X) = \frac{1}{\lambda^2}$

Área → probabilidade

Exemplo: A vida útil de um robô é uma variável aleatória com distribuição exponencial de valor médio 10 horas.
Qual a probabilidade de um robô, selecionado ao acaso, ter vida inferior a 4 horas

$$\lambda = \frac{1}{10} = 0{,}1$$

$$P[X<4] = \int_0^4 0{,}1\, e^{-0{,}1 x}\, dx = F_X(4) = 0{,}33$$

Distribuição normal vs distribuição exponencial

Distribuição normal	Distribuição exponencial
• características humanas (altura, peso, ...)	• problemas de filas de espera e fiabilidade (ex: tempo até um PC avariar)
• Surge quando vários efeitos acumulados e independentes se sobrepõem	• Relacionada com a distribuição (discreta) de Poisson

...o distribuição:

$\boxed{\text{Lei de Benford}} \longrightarrow P(d) = \log_{10}\left(1 + \frac{1}{d}\right)$

→ lei dos primeiros digitos

→ aponta suspeitos de fraudes, erros de digitação, etc

$\boxed{\text{Lei de Zipf}} \longrightarrow p(n) = \frac{K}{n} \longrightarrow$ constante dependente da língua

→ prevê que a ocorrência de uma palavra, num texto, está ligada à sua ordem

• Utilizado na área da segurança — slide 26

Variáveis aleatórias multidimensionais:

ex: Relação altura peso

• 2 tipos de casos: ① Experiência aleatória com vários saídas

② Repetição de experiência aleatória com apenas 1 saída

Vector aleatório → $X = (\underbrace{X_1 \quad X_2 \; ... \; X_n}_{\text{variáveis aleatórias}})$

Exemplo: $X = (H \; W \; A)$

$H(\varepsilon)$ = altura de ε

$W(\varepsilon)$ = peso de ε

$A(\varepsilon)$ = idade de ε

Funções de distribuição conjuntas:

obtêm-se através da extensão das definições para 1 variável

Função probabilidade de massa conjunta

Para 2 variáveis discretas X e Y:

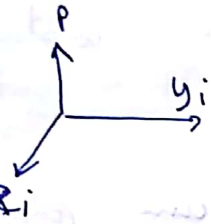
$$P_{X,Y}(i,j) = P(X=i \wedge Y=j)$$

→ X → dado 1
Y → dado 2

$P_{X,Y}(1,1) = P_{X,Y}(1,2) = \dots = P_{X,Y}(6$

$= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

→ representação 3D



Função de distribuição acumulada conjunta

$$F_{X,Y}(x,y) = P(X \le x \wedge Y \le y)$$

Distribuição de cada uma das variáveis:

$$F_X(a) = P(X \le a) = P(X \le a, Y < \infty) = F_{X,Y}(a, \infty)$$

Função de massa de probabilidade de cada uma das variáveis:

$$P_X(x) = \sum_y P_{X,Y}(x,y)$$
$$P_Y(y) = \sum_x P_{X,Y}(x,y)$$

→ somam-se as linhas apropriadas da tabela

	B (y_2)			
A (y_1)		0	1	P(y_1)
	0	0,3	0,4	0,7
	1	0,2	0,1	0,3
	P(y_2)	0,5	0,5	1

...lizosoõ :

↳ Caso de n variáveis discreta

Função Massa de probabilidade conjunta → $p_{X_1,\dots,X_n}(x_1,\dots,x_n) = P(X_1 = x_1, \dots, X_n = x_n)$

Função de probabilidade marginal para X_1 → $p_{X_1}(x_1) = \sum_{x_2,\dots,x_n} p_{X_1,\dots,X_n}(x_1,\dots,x_n)$

Função de probabilidade marginal para X_1 e X_2 → $p_{X_1,X_2}(x_1,x_2) = \sum_{x_3,\dots,x_n} p_{X_1,\dots,X_n}(x_1,x_2,\dots,x_n)$

Independência :

X e Y são independentes se, para qlq a e b :

- $P(X \leq a, Y \leq b) = P(X \leq a)\, P(Y \leq b)$, ou seja
- $E_a = \{X \leq a\}$ e $E_b \{Y \leq b\}$ são independentes

Caso discreto :

$$p(x,y) = p_X(x) \times p_Y(y)$$

Caso contínuo :

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

Momentos centrais

· momentos centrais conjuntos de ordem j, k das variáveis X, Y :

$$E[X^j Y^k] = [(X - E[X])^j (Y - E[Y])^k]$$

j	k	
2 0	0 2	variâncias
0 1	1 0	valores médios de X, Y
1	1	Correlação

Quando $j = k = 1 \Rightarrow \underbrace{E[X.Y]}$ é a $\boxed{\text{correlação das Variáveis}}$ X e Y.

→ se $E[X.Y] = 0 \Rightarrow$ variáveis ortogonais

Se X e Y independentes:

$$E[XY] = E[X] \cdot E[Y]$$

$\boxed{\text{Covariância de } X, Y}$

Quando $j = k = 1$

- $Cov(X,Y) = E[(X - E[X])(Y - E[Y])] =$
$$= E[XY] - E[X]E[Y]$$

- $Var(X) = Cov(X,X)$

Relação não linear ⟶ covariância pode não ser sensível à relação

Se $Cov(X,Y) = 0 \longrightarrow X$ e Y independentes

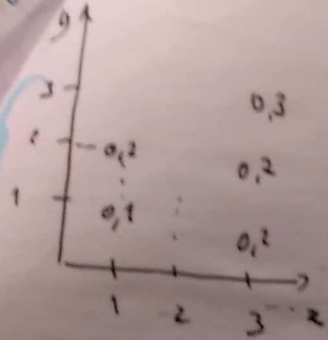
$\boxed{\text{propriedades da covariância}}$

- $Cov(X,X) = Var(X)$
- $Cov(X,Y) = Cov(Y,X)$
- $Cov(cX, Y) = c\, cov(Y,X)$
- $Cov(X, Y+Z) = Cov(Y,X) + Cov(Y,Z)$

Se tivermos $Y = (Y_1, Y_2, \dots, Y_m) \longrightarrow$ vetor de v.a.

$$(Y) = \begin{bmatrix} Cov(Y_1,Y_1) & \dots & Cov(Y_1,Y_m) \\ \vdots & \ddots & \vdots \\ Cov(Y_m,Y_1) & \dots & Cov(Y_m,Y_m) \end{bmatrix} = \begin{bmatrix} Var(Y_1) & \dots & Cov(Y,Y_m) \\ \vdots & \ddots & \vdots \\ Cov(Y_1,Y_m) & \dots & Var(Y_m) \end{bmatrix}$$

há uma troca



$E[X] = 1(0{,}1+0{,}2) + 3(0{,}2+0{,}2+0{,}3) = 2{,}4$

$E[Y] = 1(0{,}3) + 2 \cdot 0{,}4 \cdot 3 \times 0{,}3 = 2$

$Cov(X,Y) = E[(X-E[X]) \cdot (Y-E[Y])] =$

$= (1-2{,}4)(1-2{,}0)\times 0{,}1 + (1-2{,}4)(2-2{,}0)\times 0{,}2 +$

$+ (3-2{,}4)(1-2{,}0)\times 0{,}2 + (3-2{,}4)(2-2{,}0)\cdot 0{,}2 +$

$+ (3-2{,}4)(3-2{,}0)\cdot 0{,}3 = 0{,}2$

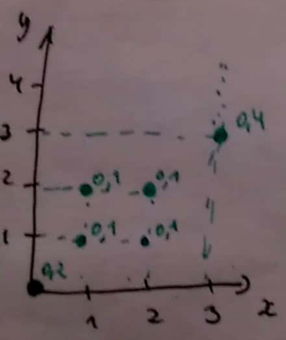
Coeficiente de correlação de X e Y:

$$\boxed{\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}} \quad \text{e} \quad -1 \leq \rho_{X,Y} \leq 1$$

→ Se for $= 0$ → variáveis descorrelacionadas

- X e Y independentes → $Cov(X,Y) = 0$, mas o contrário não é verdadeira

Exemplo



x	y	$P(x,y)$	$xy\,P(x,y)$	$x\,P(x)$	$y\,P(y)$	$x^2\,P(x)$	$y^2\,P(y)$
0	0	0,2	0	0	0	0	0
1	1	0,1	0,1	0,1	0,1	0,1	0,1
1	2	0,1	0,2	0,1	0,2	0,1	0,4
2	1	0,1	0,2	0,2	0,1	0,4	0,1
2	2	0,1	0,4	0,2	0,2	0,4	0,4
3	3	0,4	3,6	1,2	1,2	3,6	3,6
SUM		1	4,5	1,8	1,8	4,6	4,5

$Var(X) = E[X^2] - (E[X])^2 = 4{,}6 - 3{,}24 = 1{,}36$

$Var(Y) = 1{,}36$

$Cov(X,Y) = E[XY] - E[X]\cdot E[Y] = 4{,}5 - (1{,}8)(1{,}8) = 1{,}26$ | $\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{1{,}26}{(\sqrt{1{,}36})^2} = 0{,}926$