Notas de aula: Fundamentos Estatísticos para Ciência dos Dados

Ricardo Pagoto Marinho

3 de abril de 2018

- 13/03

  $P(\cup A_i) \leq \sum P(A_i) \rightarrow$ é igual quantos os $A_i s$ forem disjuntos.

- 15/03

  $P(A|B) = P(B) \rightarrow$ quando A ocorre e não tem nenhuma influência sobre B_0.

- 20/03

  Variável aleatória: Lista de valores possíveis e lista de probabilidades associadas

  ω dentro de um Ω. Exemplo: $\Omega =$ todos e-mails enviados.

  - $\omega_0 =$ é spam?
  - $\omega_1 =$ número de caracteres.
  - $\cdots$

  Elementos em uma mesma linha (ω_n), são correlacionados.

  - Atribuir valores de probabilidades a uma V.A. $\rightarrow$ contar quantos elementos no Ω possuem aquela característica.
    $P(X = 3) = P(A)$ onde A=$\{\omega \in \Omega / \omega \ tem\ 3\ caras\}$ em $\Omega =$ lançamento de 6 moedas.
  - Esperança matemática $E(X)$
    $E(X) = \sum_i x_i p(xi) \approx \sum_i xi \times \frac{N_i}{N}$
  - Distribuição Binomial
    $P(X = 0) = (1 - \theta)^n$
    $[X = 0] = \{\omega \in \Omega : X(\omega) = 1\} = \{\omega \in \Omega : \omega \in \{(\neg c, \neg c, \neg c, \cdots, \neg c)\}\} =$
    $P(\neg c\ no\ 1^o) \times P(\neg\ c\ no\ 2^o) \times \cdots = (1 - \theta) \times (1 - \theta) \cdots = (1 - \theta)^n$

- 27/03

  $P(Y \in (y_0 \pm \frac{\delta}{2})) = \int_{y_0 - \frac{\delta}{2}}^{y_0 + \frac{\delta}{2}} f^*(y) dy \approx f^*(y_0) 2 \times \frac{\delta}{2} = f^*(y_0) \times \delta$

  Teste de Kolmogorov:

  $\sqrt{n}(D_n) \rightarrow K$, onde K é uma Variável Aleatória contínua.

  Se o modelo é o verdadeiro, quando comparado com os dados, a distância entre eles multiplicado por $\sqrt{n}$ vai cair dentro da densidade de K. Se não cair, provavelmente seu modelo não é adequado. Quanto maior o número de dados, mais confiável o resultado.

- 03/04 Variáveis aleatórias: Lista de valores possíveis + probabilidades associadas

	Discretas	Contínuas
Valores	0,1,2,...	[0,1] ou [0,∞)
Probabilidades	$p_0, p_1, p_2, ...$	Densidade sob a curva

	$E(X)$
Discreta	$\sum_i x_i \times \; P(X = x_i)$
Contínua	$\int x \times \; f(x)dx$