

## Lista 3: Fundamentos Estatísticos para Ciência dos Dados

Ricardo Pagoto Marinho

16 de março de 2018

1.
  - $P(A^C) = 1 - P(A)$   
 Seja  $P(A) = q$  e  $A^C$  os elementos que estão em  $\Omega$  mas não em  $A$ . Logo,  $A^C = \Omega - A$ , então  $P(A^C) = P(\Omega) - P(A) = 1 - P(A)$ .
  - Sabe-se que  $P(A) = 1$  e  $\forall A \subset \Omega, P(A) \leq 1$ . Logo,  $0 \leq P(A) \leq 1$
  - Suponha que  $A_1 \subset A_2$  e  $P(A_1) > P(A_2)$ . Logo, dentro do experimento, existem mais possibilidades de aparecer um elemento de  $A_1$  do que de  $A_2$ . Portanto,  $A_1 > A_2$ . Absurdo, já que  $A_1 \subset A_2$ .
  - Seja  $A_i$  uma coleção de conjuntos de  $\Omega$ .  $P(\cup_{n=1}^{\infty} A_i)$  cai em dois casos:
    - caso 1: Todos  $A_i$ s são disjuntos. Portanto,  $P(\cup_{n=1}^{\infty} A_i) = \sum_{n=1}^{\infty} P(A_i)$
    - caso 2: Alguns  $A_i$  se intersectam. Neste caso, deve-se tirar a probabilidade  $P(A_i \cap A_{i+1})$  para os  $A_i$  e  $A_{i+1}$  tais que  $A_i \cap A_{i+1} \neq \emptyset$ . Logo  $P(\cup_{n=1}^{\infty} A_i) < \sum_{n=1}^{\infty} P(A_i)$   
 Portanto  $P(\cup_{n=1}^{\infty} A_i) \leq \sum_{n=1}^{\infty} P(A_i)$ .
  - $P(A \cup B)$  é a probabilidade de um item selecionado estiver em  $A$  ou  $B$ . Contudo, caso  $A \cap B \neq \emptyset$ , a soma das probabilidades somará duas vezes os itens que estão em  $A \cap B$ . Logo,  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
2.
  - V
  - V
  - F
  - F
  - F
  - V
  - V
  - V
3. O contrário não pode ser dito.  
 $P(B|A) = \frac{P(B \cap A)}{P(A)}$ , caso  $A \supset B$ ,  $P(B \cap A) = P(B)$  e  $P(B|A) = \frac{P(B)}{P(A)}$ .  
 Como  $0 \leq P(A) \leq 1$ ,  $\frac{P(B)}{P(A)} \geq P(B)$ .
4. Essas probabilidades podem ser obtidas observando os pacientes ao longo de um ano e contabiliza os pacientes que se mantiveram vivos e os que não sobreviveram ao período.
5.
  - Significa que, quanto maior a taxa, mais confiável fica o teste, logo todos os positivos serão flagrados.

- A afirmação diz que quanto maior a sensibilidade, mais confiável o teste se torna.
- *Recall* é proporção de pacientes que o teste foi positivo e ele realmente estava doente.

6. To do.

7.    • F
- F
- F
- V

8. Para mostrar que um evento é independente de outro, temos que mostrar que  $P(A|B) = P(A) \times P(B)$ . Como  $P(A) = 0$ ,  $P(A|B) = 0 \times P(B) = 0, \forall B$ , logo A é independente de qualquer B quando  $P(A) = 0$ .

Sim, intuitivamente faz sentido, já que se não existe a possibilidade de A, os eventos seguintes não vão sofrer nenhuma influência dele.

9.  $P(A|A) = P(A) \times P(A) \iff P(A) = 1 \text{ ou } P(A) = 0$ .