Aula 2

Aula passada

- Logística, programação, regras do jogo
- Três problemas
- *Monte Carlo to the rescue*
- História das origens
- Monte Carlo na Computação

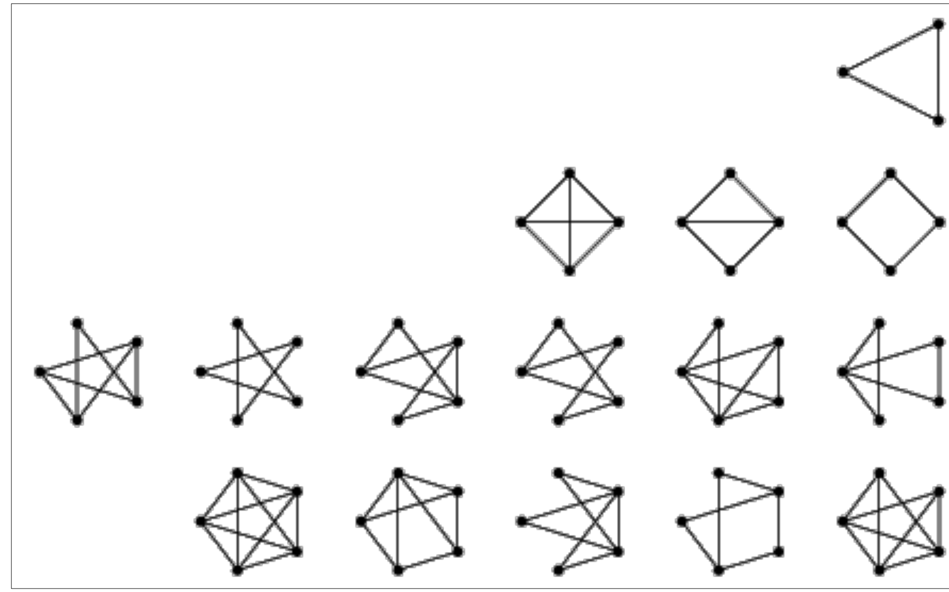
Aula de hoje

- Espaço amostral
- Probabilidade
- Eventos
- Independência
- Exclusão mútua
- Probabilidade total
- Regra de Bayes
- Variável aleatória
- Função de distribuição
- Distribuição de Bernoulli
- Sequência de v.a.
- Distribuição Binomial

Espaço Amostral

- Conjunto de objetos, enumerável (pode contar)

Frutas



Grafos

Letras

- S é o conjunto
  - ex. $S = \{a,b,c,\ldots,z\}$
- $|S|$ é sua cardinalidade
  - ex. $|S| = 26$

Figueiredo 2018

Probabilidade

- Função que associa a cada elemento de S um valor entre 0 e 1
  - $p : S \rightarrow [0, 1]$
- Restrição: soma sobre todos elementos vale 1

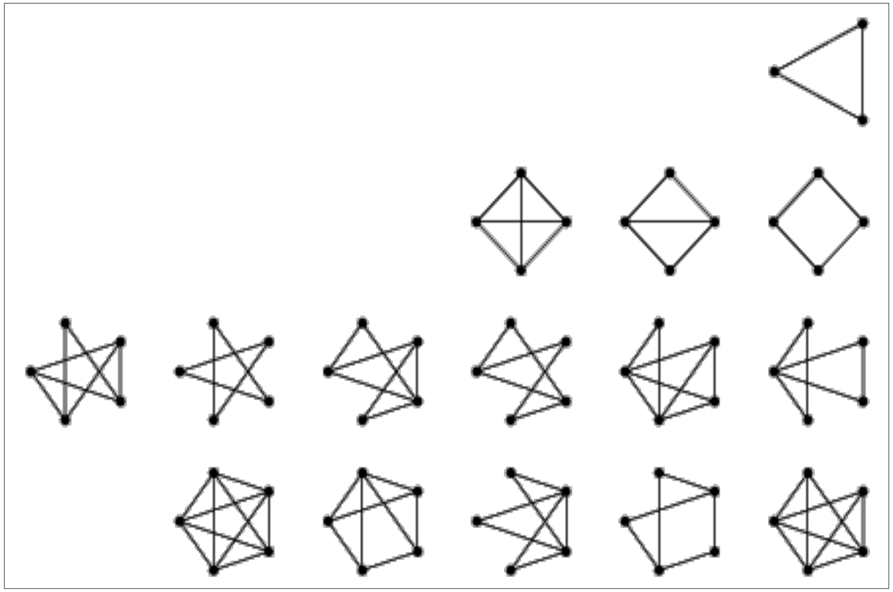
Letras

- S = alfabeto, $|S| = 26$
- $p_x = 1/26$ para qualquer letra x
- $p_x = 1/10$ para qualquer x vogal, e $1/42$ para qualquer x consoante

Eventos

- Subconjunto do espaço amostral

- S = alfabeto, $|S| = 26$
- $A = \{a,b,c,d\}$
- B = todas as consoantes
- C = todas as letras depois de Q



- S = grafos ao lado, $|S| = 15$
- A = todos as cliques
- B = todas os grafos com menos de quatro vértices
- C = todas as árvores

Figueiredo 2018

Probabilidade de Eventos

- Soma das probabilidades dos elementos que definem o evento

$$P[A] = \sum_{e \in A} p_e$$

- S = alfabeto, $|S| = 26$, $p_x = 1/26$
- $A = \{a,b,c,d\}$
- B = todas as consoantes
- C = todas as letras depois de Q
- $P[A] = 2/13$
- $P[B] = 21/26$
- $P[C] = 9/26$

Manipulando Eventos

- Eventos são conjuntos, podem ser manipulados usando lógica de conjuntos
- Operações básicas
  - união, interseção, complemento

- A = {a,b,c,d}
- B = todas as consoantes
- C = todas as letras depois de Q

$$A \cup B = ?$$

$$A \cap B = ?$$

$$B^c = ?$$

$$A \cap B \cap C = ?$$

Manipulando Eventos

- Probabilidade de eventos resultantes seguem a mesma definição
  - ex. $Y = A \cap B \longrightarrow P[Y] = \sum_{e \in Y} p_e$

- S = alfabeto, $|S|$ = 26, p_x = 1/26
- A = {a,b,c,d}
- B = todas as consoantes
- C = todas as letras depois de Q

$$P[A \cup B] = ?$$

$$P[A \cap B \cap C] = ? \qquad P[B^c] = ?$$

Ponto de Confusão

- Operadores lógicos são usados no lugar de operadores de conjunto
  - “ou”, “+” ao invés de “união”
  - “e”, “.” ao invés de “interseção”
  - “negação”, “not” ao invés de “complemento”
  - origem de grande confusão!
- Exemplos

$$A \cup B = A + B = A \vee B$$

$$A \cap B = A \, . \, B = A \wedge B$$

$$B^c = \overline{B} = \neg B$$

Iremos usar esta notação, que é a mais usual

Independência

- Dois eventos *A* e *B* são independentes sse

$$P[A \cap B] = P[A \wedge B] = P[A]P[B]$$

- Ou seja, a probabilidade da conjunção (interseção) é igual ao produto das probabilidades

- S = alfabeto, $|S| = 26$, $p_x = 1/26$
- $A = \{a,b,c,d\}$
- B = todas as consoantes
- *A* e *B* são independentes?
- A = todas as letras antes de N
- $B = \{a, z\}$
- *A* e *B* são independentes?

Exclusão Mútua

- Dois eventos *A* e *B* são mutuamente exclusivos sse

$$P[A \cup B] = P[A \vee B] = P[A] + P[B]$$

- Ou seja, a probabilidade da disjunção (união) é igual a soma das probabilidades

- S = alfabeto, $|S| = 26$, $p_x = 1/26$
- $A = \{a,b,c,d\}$
- B = todas as consoantes
- *A* e *B* são mutuamente exclusivos?
- *A* = todas as letras antes de N
- $B = \{x, y, z\}$
- *A* e *B* são mutuamente exclusivos?

Probabilidade Condicional

- Probabilidade do evento *A* dado a ocorrência do evento *B*
  - ao saber que *B* ocorreu, o espaço amostral para ocorrência de *A* passa a ser o conjunto *B*

$$P[A|B] = \frac{P[A \wedge B]}{P[B]}$$

- S = alfabeto, $|S| = 26$, $p_x = 1/26$
- $A = \{a,b,c,d\}$
- B = todas as consoantes
- $C = \{x, y, z\}$

$A \cap B = \{b,c,d\} \rightarrow P(A \cap B) = 3/26$

$P(B) = \frac{21}{26}$ Logo, $P[A|B] = 3/21$

$$P[A|B] = ? \quad P[B|A] = ? \quad P[C|B] = ?$$

Lei da Probabilidade Total

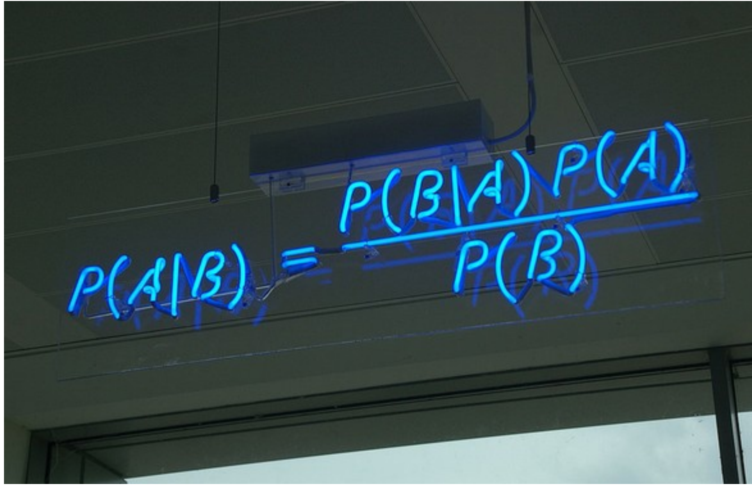
- Particionamento de um conjunto *S*
  - conjunto de subconjuntos de *S* tal que todo elemento de *S* aparece em exatamente um subconjunto
- ex. S = alfabeto: O_1 = todas vogais, O_2 = todas consoantes

- Relação entre probabilidade de um evento e probabilidade condicional com eventos de uma partição

$$P[A] = \sum_i P[A \wedge B_i] = \sum_i P[A|B_i]\,P[B_i]$$

- onde B_i é qualquer partição do espaço amostral
- Regra muito usada para calcular probabilidade de um evento

Regra de Bayes

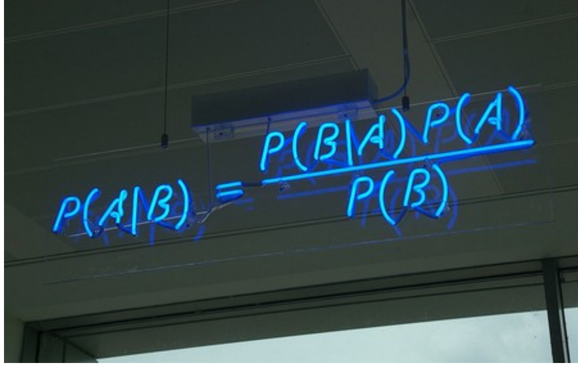


- Relação fundamental entre probabilidades condicionais

$$P[A|B] = \frac{P[B|A]\,P[A]}{P[B]}$$

- Talvez a mais importante relação em probabilidade
  - usada em muitos, muitos problemas
- Razão
  - queremos P[A|B], mas calcular isto pode ser difícil
  - calcular P[B|A] pode ser mais fácil, assim como P[A] e P[B]
  - e muitas vezes calcular P[B] não é necessário

Exemplo



- Filtro de spam usando regra de Bayes
- *S* = todos os e-mails de um grande universo
- *M* = subconjunto de e-mails que são spam
- *V* = subconjunto de e-mails que possuem a palavra “viagra”

$P[M|V] \longleftarrow$ Probabilidade de um e-mail ser spam dado que possui a palavra “viagra”

- Como calcular esta probabilidade? Regra de Bayes!

$$P[M|V] = \frac{P[V|M]\,P[M]}{P[V]}$$

- *P*[*V* | *M*] é mais fácil de calcular empiricamente, assim como *P*[*V*] e *P*[*M*]

Figueiredo 2018

Variável Aleatória

- Até agora, não temos nada aleatório!
  - vamos introduzir uma "variável" cujo valor que ela assume será "aleatório"
- Variável aleatória *X* é uma função que mapeia o espaço amostral nos inteiros (v.a. discreta)
  - permite trabalhar com números, independente dos elementos do espaço amostral

$$X : S \rightarrow \mathbb{Z}$$

- Usaremos *X* para definir eventos em função de seus valores

$$A = \{X > 5\} = \{e \in S : X(e) > 5\}$$

Exemplo de Variável Aleatória

- S = alfabeto, $|S| = 26$, $p_x = 1/26$
- X é uma v.a. tal que $X(a)=1$, $X(b)=2$, $X(c)=3$, …, $X(z)=26$
- Y é uma v.a. tal que Y(vogal)=1, Y(consoante)=2
- Z é uma v.a. tal que Z(primeiras 10 letras) = 1, Z(últimas 10 letras) = 2, Z(outras letras) = 3

- Usamos v.a. para definir eventos
- $\{X > 13\} = \{n,l,\ldots,z\}$, $\{X \text{ ímpar}\} = \{a,c,e,g,\ldots,y\}$
- $\{Y = 1\} = \{a,e,i,o,u\}$
- $\{Z = 3\} = \{k,l,m,n,o,p\}$

Figueiredo 2018

Probabilidade de V.A.

- Probabilidade associada é dada pela probabilidade do evento definido pela v.a.
- $\{X > 13\} = \{n,l,\ldots,z\}$, $\{X \text{ ímpar}\} = \{a,c,e,g,\ldots,y\}$
- $\{Y = 1\} = \{a,e,i,o,u\}$
- $\{Z = 3\} = \{k,l,m,n,o,p\}$

- $P[X>13] = ?$
- $P[Y=1] = ?$
- $P[Y=1 \text{ e } Z=3] = ?$

  $= P[Y=1]\, P[Z=3]$?

Cuidado! *Y* e *Z* são v.a. diferentes mas não necessariamente independentes

Manipulando V.A.

- Variáveis aleatórias também assumem valores de sua imagem
  - chamado de realização da v.a.
- Por isto podem ser manipuladas algebricamente (com cuidado)
  - simplificação: $2X > 4 = X > 2$
  - atribuição: $Z = 2X - 3Y$

Pergunta: $X + X = 2X$

- **Depende!**
  - Sim, se as duas ocorrências de *X* se referem a mesma realização (instância) da v.a.
  - Não, se as ocorrências de *X* se referem a realizações diferentes
- Exemplo: *X* é o valor da face de um dado

V.A. Indicadora

- Tipo mais comum de variável aleatória
  - assume apenas dois valores (1 ou 0), indica a ocorrência de um evento de interesse
  - $X = 1$, subconjunto do espaço amostral que possui evento de interesse
  - $X = 0$, caso contrário
- Exemplo: S = todos os grafos com n vértices
  - $X = 1$: todos os grafos conexos, X=0 c.c.
  - $Y_k = 1$: todos os grafos com diâmetro k, Y_k=0 c.c.
  - $Z_k = 1$: todos os grafos com k componentes conexas
  - P[X = 1] = P[X] = probabilidade do grafo escolhido aleatoriamente ser conexo

Novas Regras!

- Probabilidade associada ao valor de uma v.a. está relacionada a probabilidade dos eventos definidos por este valor
  - $P[\ X = 1\ ]$
- **Novas regras!** Iremos definir a probabilidade associada ao valor de uma v.a. diretamente
  - sem pensar muito nos eventos definidos por este valor e na probabilidade destes eventos
  - $P[\ X = 1\ ] = p_1$
- Iremos determinar o valor p_1 a partir de uma função
  - função de probabilidade (ou função de massa de probabilidade ou função de distribuição)

Função de Distribuição

- Seja *X* uma v.a. e *x* um de seus possíveis valores

$$f_X(x) = P[X = x]$$ ← Função de probabilidade

$$F_X(x) = P[X \le x]$$ ← Função cumulativa

- Uma pode ser definida usando a outra

$$F_X(x) = \sum_{y \le x} f_X(y)$$ ← Soma por exclusão mútua dos eventos X = x

- Restrição

$$0 \le f_X(x) \le 1, \forall x \in O_X \qquad \sum_{x \in O_x} f_X(x) = 1$$

- Onde O_X é a imagem da v.a. *X* (valores que ela pode assumir)

Figueiredo 2018

Distribuição de Bernoulli

- Uma v.a. *X* que possui distribuição de Bernoulli assume apenas dois valores

$$f_X(1) = P[X=1] = p$$

$$f_X(0) = P[X=0] = 1-p$$

- Possui um único parâmetro 0< p < 1
- Usada como distribuição de qualquer v.a. indicadora
  - ex. cara ou coroa, sim ou não, verdade ou falso
- Notação: *X*~Bernoulli(*p*)
  - *X* é uma v.a. que possui distribuição de Bernoulli com parâmetro *p*

Sequência de V.A.

- Considere uma sequência de *n* v.a.

$$X_{1,}\ X_{2,} \ldots, X_n$$

- Dizemos que a sequência é i.i.d. (independente e identicamente distribuída) se
  - v.a. são independentes
  - v.a. possuem a mesma função de distribuição
  - v.a. são distintas
- Exemplos de sequências iid
  - jogar um mesmo dado *n* vezes: X_i é o valor observado na *i*-ésima jogada do dado

Distribuição Binomial

- Considere uma sequência iid de *n* v.a. de Bernoulli

$$X_{1,}\ X_{2,} \ldots,\ X_n$$

  - onde X_i ~ Bernoulli(p) para i = 1,2,…, n

- Seja Z a soma destas v.a.

$$Z = \sum_{i=1}^{n} X_i$$

- Z possui distribuição Binomial, com parâmetros n e p
  - $Z \sim \text{Bin}(n, p)$
- Que valores que Z pode assumir ? $\{0, 1, \ldots, n\}$

Distribuição Binomial

- Qual a expressão para a distribuição Binomial?
  - probabilidade da soma ser igual a *i*

$$f_Z(i)=P[Z=i]$$

prob. de i variáveis terem saído 1

i saídas 1

n-i saídas 0

- Quanto vale?

$$f_Z(i)=P[Z=i]=\binom{n}{i}p^i(1-p)^{n-i}$$

combinação

- Podemos obter expressão a partir de princípios elementares
  - ver exercício da lista
- Exemplos
  - número de caras ao jogar uma moeda 20 vezes
  - grau do vértice em um grafo com *n* vértices onde cada aresta incidente ocorre com prob *p* (modelo G(*n*,*p*))

Figueiredo 2018