AGA 0505 - Análise de Dados em Astronomia

2. Probabilidades

Laerte Sodré Jr.

1o. semestre, 2025

aula de hoje:

1. o que são probabilidades
2. distribuições de probabilidades
3. probabilidades condicionais e conjuntas
4. as regras fundamentais das probabilidades e o teorema de Bayes
5. aplicações simples do teorema de Bayes

A teoria das probabilidades é o senso comum reduzido ao cálculo

Pierre-Simon Laplace

dados e probabilidades

- descrevemos os dados (medidas, observações) como *distribuições de probabilidades*
- dados podem ser complexos, variáveis e, sobretudo, têm incertezas
  ex. de incerteza: os erros nas medidas
- a teoria das probabilidades provê um meio de modelar os dados e quantificar as incertezas

- duas visões sobre a natureza das probabilidades:
  - *frequentista (clássica)*: probabilidade como medida de frequência
  - *bayesiana*: probabilidade como medida de plausibilidade
- podemos considerar probabilidade como medida de frequência como um caso particular de medida de plausibilidade

o que são probabilidades

- probabilidade frequentista:
  - medida da frequência de eventos (em vários experimentos ou conjunto de sistemas estatisticamente equivalentes)
  - $P(x)$: o número de vezes que um evento ocorre dividido pelo número total de tentativas, no limite de um número infinito de tentativas
  - $P(x)$: número entre 0 e 1 que mede a *frequência* com que a proposição x aparece em uma população ou amostra
- problemas:
  - não comporta eventos únicos ou não repetíveis
  - lida com propriedades assintóticas (“no limite...”)

- probabilidade bayesiana:
  medida da plausibilidade de um evento
  - $P(x)$: número entre 0 e 1 que mede a *plausibilidade* de um dado valor de x (com 0 impossível, falso e 1 certo, verdadeiro)
  - este conceito permite atribuir probabilidades para qualquer evento
- problemas:
  - nem sempre se consegue definir um modelo probabilístico adequado
  - em particular, pode-se ter uma certa ambiguidade devido à escolha de diferentes priores
    (veremos isso mais tarde)

distribuições de probabilidades

- x: variável *aleatória contínua ou discreta*
  - aleatória: variável cujos valores são resultados de um processo aleatório ou estocástico, obedecendo a uma certa distribuição de probabilidades, $P(x)$
  - contínua: assume valores reais
  - discreta: assume valores inteiros

- notação: $x \sim P(x)$
  x é uma variável aleatória com distribuição que obedece a $P(x)$
- $P(x)$: número entre 0 e 1 que mede a incidência da variável x ou o grau com que uma proposição x é verdadeira
- $P(x)$ pode ser uma função discreta ou contínua
- x pode ser multidimensional

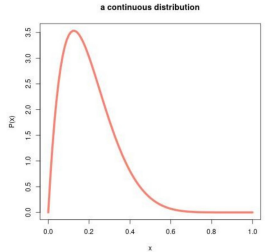
- distribuições cumulativas
  - $P(x)$: função de distribuição de probabilidades (FDP)
  - FDP cumulativa:

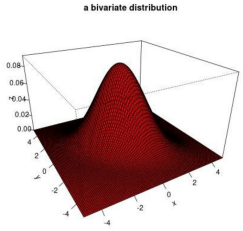
$$C(x) = \int_{-\infty}^{x} P(x')dx'$$

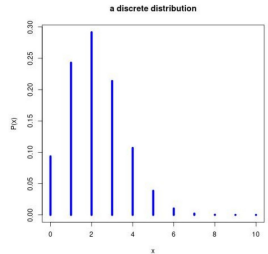
$$C(\infty) = \int_{-\infty}^{\infty} P(x')dx' = 1$$

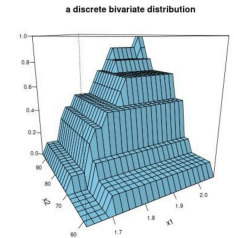
  probabilidades são normalizadas!
  - se o intervalo onde a variável x é definida é finito, os limites da integral são os mesmos da variável

distribuições de probabilidades









probabilidades condicionais e conjuntas

- probabilidades são, normalmente, condicionais, isto é, dependem ou podem depender de outras proposições
  exemplo: qual é a probabilidade de uma galáxia ser espiral se ela for vermelha? $P(espiral|vermelha)$?

  $P(x|y)$: lê-se probabilidade de x dado y
  $P(x|y)$ mede a probabilidade de x dado o valor da variável y
  $P(x|y,z,w)$: a probabilidade de x é condicional a tudo o que estiver à direita da barra "|"

  exemplo: distribuição gaussiana

$$P(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big[-\frac{(x-\mu)^2}{2\sigma^2}\Big] \sim N(x|\mu,\sigma) \sim N(\mu,\sigma)$$

- probabilidade conjunta: $P(x,y)$
  probabilidade conjunta de duas variáveis x E y
- $P(x,y|z)$: probabilidade conjunta de x e y dado z (podemos ter coisas como $P(x,y,z,w|a,b,c)$)

- $P(x) = P(x|H)$: toda probabilidade depende de hipóteses (H), implícitas ou explícitas
  por exemplo: a distribuição é gaussiana

as regras fundamentais das probabilidades

- **as regras fundamentais das probabilidades**:
  - probabilidades satisfazem
    $$0 \leq P(x) \leq 1$$
  - **regra da soma**: para dois eventos x e y,
    $$P(x, y) = P(x) + P(y) - P(x \cap y) \qquad \cap : \text{intersecção}$$
    se x e y são mutuamente exclusivos (i.e., $x \cap y = \emptyset$),
    $$P(x, y) = P(x) + P(y)$$
    esses resultados podem ser generalizados: para proposições mutuamente exclusivas e exaustivas
    $$\int P(x)dx = 1 \qquad \sum_i P(x_i) = 1 \qquad \text{(probabilidades são normalizadas)}$$
  - **regra do produto**:
    $$P(x, y) = P(x|y) \times P(y) = P(y|x) \times P(x)$$

as regras fundamentais das probabilidades: algumas consequências

- **marginalização:**

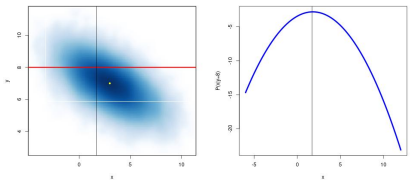
$$P(x) = \int P(x, y)dy = \int P(x|y)P(y)dy$$

ou

$$P(x) = \sum_k P(x, y_k) = \sum_k P(x|y_k)P(y_k)$$

este resultado é conhecido como
lei das probabilidades totais

- exemplo: marginalização de uma gaussiana bivariada



$$P(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_x}{\sigma_x} \right)^2 \right.\right.$$

$$\left.\left. -2\rho \left(\frac{x-\mu_x}{\sigma_x} \right) \left(\frac{y-\mu_y}{\sigma_y} \right) + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 \right] \right\}$$

$$P(x) = \int P(x, y)dy \sim N(\mu_x, \sigma_x)$$

as regras fundamentais das probabilidades: algumas consequências

aplicações da regra do produto:

$$P(x, y) = P(x|y) \times P(y)$$

- vamos supor que x e y são **variáveis independentes**:

  $$P(x|y) = P(x) \text{ e } P(y|x) = P(y)$$

  portanto,

  $$P(x, y) = P(x|y)P(y) = P(x)P(y)$$

  a probabilidade de duas variáveis independentes é o produto das probabilidades de cada uma

- outra aplicação da regra do produto: *cadeias de probabilidades*:

  $$P(x, y, z) = P(x|y, z)P(y, z)$$

  $$P(y, z) = P(y|z)P(z)$$

  logo,

  $$P(x, y, z) = P(x|y, z)P(y|z)P(z)$$

as regras fundamentais das probabilidades: algumas consequências

combinando distribuições

- muitas vezes conhecemos a distribuição de uma variável mas queremos saber a distribuição de uma quantidade derivada, $y = f(x)$
  ex: temos a distribuição de magnitudes e queremos a distribuição de fluxos
- $P(x)$: distribuição de probabilidades de x
- $P(y)$: distribuição de probabilidades de y
- *a densidade de probabilidades é conservada*,

$$P(x)dx = P(y)dy, \quad \text{logo} \quad P(y) = \left|\frac{dx}{dy}\right| P(x)$$

(o módulo é para assegurar $P \geq 0$)

- exemplo:
  suponha que $P(x) = \exp(-x)$ (com $x > 0$)
  e queremos $P(y)$, onde $y = \ln x$
- como $x = \exp(y)$,

$$P(y) = P(x)/|dy/dx| = x\exp(-x)$$

ou

$$P(y) = \exp(y - \exp(y))$$

- esta técnica se torna difícil de aplicar para mais de uma variável
- cuidado se $f(x)$ não é monotônica!

as regras fundamentais das probabilidades: algumas consequências

o teorema de Bayes
(Bayes,∼1740/1763; Laplace, 1774)

- da regra do produto vem que:

  $P(x, y) = P(x|y) \times P(y) \qquad P(y, x) = P(y|x) \times P(x)$

- como $P(x, y) = P(y, x)$, temos o teorema de Bayes:

  $$P(x|y) = \frac{P(y|x) \times P(x)}{P(y)}$$

  o teorema de Bayes relaciona $P(x|y)$ com $P(y|x)$
  probabilidade inversa

aplicações simples do teorema de Bayes (1)

temos uma amostra de galáxias de vários tipos e massas

- seja A: a galáxia tem formação estelar
  B: a galáxia é massiva
- sabemos que, nessa amostra,
  - P(A)=0.7
  - P(B)=0.2
  - também sabemos que 10% das galáxias da amostra com formação estelar são massivas: P(B|A)=0.1
- *qual é a probabilidade da galáxia ter formação estelar se ela for massiva? P(A|B)?*

- teorema de Bayes

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} = \frac{0.1 \times 0.7}{0.2} = 0.35$$

nessa amostra, a probabilidade da galáxia ter formação estelar se ela for massiva é 35%

(SDSS)

aplicações simples do teorema de Bayes (2)

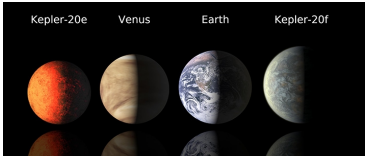
em uma galáxia muito, muito distante...

- 5% de seus planetas têm água líquida e 1% tem alguma forma de vida
- a probabilidade de vida em um planeta com água líquida é 10%
- qual é a probabilidade de haver água líquida em um planeta com vida?
- dados: P(água) = 0.05, P(vida) = 0.10, P(vida|água) = 0.10

- teorema de Bayes

$$P(\text{agua}|\text{vida}) = \frac{P(\text{vida}|\text{agua}) \times P(\text{agua})}{P(\text{vida})} =$$

$$= \frac{0.10 \times 0.05}{0.01} = 0.5$$



aplicações simples do teorema de Bayes (3)

- Num levantamento de 200 galáxias, deseja-se verificar se a presença de formação estelar depende do ambiente em que a galáxia está: se em um aglomerado ou no "campo". O evento A indica se a galáxia está num aglomerado e o B se ela tem formação estelar. Metade das galáxias da amostra está em cada ambiente e a fração com formação estelar nos aglomerados é 40%, contra 80% no campo.
- Qual é a probabilidade de que uma galáxia tenha formação estelar e esteja num aglomerado, $P(B, A)$?
  - seja: A: galáxia em aglomerado, $\bar{A}$: galáxia no campo, B: galáxia com formação estelar, $\bar{B}$: galáxia quiescente (sem formação estelar)
  - temos que
    $P(A) = P(\bar{A}) = 1/2$
    $P(B|A) = 0.4$
    $P(B|\bar{A}) = 0.8$
  - logo,
    $P(B, A) = P(B|A)P(A) = 0.4 \times 0.5 = 0.2$

- Qual é a probabilidade de se ter uma galáxia com formação estelar nessa amostra?
  $P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) = 0.4 \times 0.5 + 0.8 \times 0.5 = 0.6$

aplicações simples do teorema de Bayes (4)

- em uma certa amostra, observamos que
  - 1% dos objetos são Nebulosas Planetárias (NP)
  - 90% das NP são classificadas como NP (PV: positivos verdadeiros)
  - 7% são classificados como falsos positivos (FP)
- se um objeto é classificado como NP, qual é a probabilidade de ele ser de fato uma NP?
- seja NP: o objeto é uma NP
  $\bar{NP}$: o objeto não é uma NP
  $P(NP) = 0.01$; logo $P(\bar{NP}) = 0.99$
- seja c: o objeto foi classificado como NP
  sabemos que $P(c|NP) = 0.9$ e $P(c|\bar{NP}) = 0.07$
- *queremos* $P(NP|c)$:

$$P(NP|c) = \frac{P(c|NP)P(NP)}{P(c)}$$

onde

$$P(c) = P(c|NP)P(NP) + P(c|\bar{NP})P(\bar{NP})$$

- então,

$$P(NP|c) = \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.07 \times 0.99} = 0.115$$

a probabilidade de um objeto classificado como NP ser de fato uma NP é 11.5%

aplicações simples do teorema de Bayes (5)

- Um astrônomo pede a um colega para revisar um artigo que ele escreveu. A probabilidade do colega não aceitar é 1/3. A probabilidade do artigo ser aceito se o colega revisar é 1/2, e 1/4 no caso contrário.

- Qual é a probabilidade do artigo ser aceito, $P(A)$?
  - seja A: aceito, $\bar{A}$: rejeitado, R: revisou, $\bar{R}$: não revisou
  - temos que
    $P(\bar{R}) = 1/3 \longrightarrow P(R) = 2/3$
    $P(A|R) = 1/2$
    $P(A|\bar{R}) = 1/4$
  - logo,
    $P(A) = P(A|R)P(R) + P(A|\bar{R})P(\bar{R}) = 1/2 \times 2/3 + 1/4 \times 1/3 = 5/12$

aplicações simples do teorema de Bayes (6)

- Um astrônomo pede a um colega para revisar um artigo que ele escreveu. A probabilidade do colega não aceitar é 1/3. A probabilidade do artigo ser aceito se o colega revisar é 1/2, e 1/4 no caso contrário.

- Suponha que o artigo foi rejeitado. Qual é a probabilidade do colega não ter revisado o artigo, $P(\bar{R}|\bar{A})$? da página anterior:

$$P(\bar{R}) = 1/3, \quad P(R) = 2/3, \quad P(A|R) = 1/2, \quad P(A|\bar{R}) = 1/4, \quad P(A) = 5/12$$

  - probabilidade conjunta de o colega não ter revisado e o artigo não ter sido aceito:

$$P(\bar{R}, \bar{A}) = P(\bar{R}|\bar{A})P(\bar{A}) = P(\bar{A}|\bar{R})P(\bar{R})$$
$$P(\bar{R}|\bar{A}) = P(\bar{A}|\bar{R})P(\bar{R})/P(\bar{A})$$
$$P(\bar{A}) = P(\bar{A}|R)P(R) + P(\bar{A}|\bar{R})P(\bar{R})$$

  - note que $P(\bar{A}|R) = 1 - P(A|R)$ e $P(\bar{A}|\bar{R}) = 1 - P(A|\bar{R})$ e, portanto,

$$P(\bar{R}|\bar{A}) = P(\bar{A}|\bar{R})P(\bar{R})/P(\bar{A}) = \frac{3/4 \times 1/3}{7/12} = 3/7 \simeq 0.43$$

exercícios sobre probabilidades

- probabilidade de A ou B:
  dados 2 eventos, A e B, qual é a probabilidade de se ter A OU B?

$$P(A \text{ ou } B) = P(A) + P(B) - P(A \text{ e } B)$$

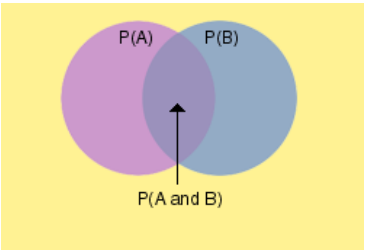


diagrama de Venn

- exemplo: qual é a probabilidade de se tirar uma carta vermelha ou um 3 em um baralho de 52 cartas (4 naipes de 13 cartas)?
  $P(V) = 26/52; \quad P(3) = 4/52; \quad P(V, 3) = 2/52$

  e, portanto,
  $P(V \text{ ou } 3) = 26/52 + 4/52 - 2/52 =$
  $= 28/52 = 7/13 \simeq 0.538...$

- eventos mutuamente exclusivos: se um ocorre o outro não ou os dois não podem ocorrer ao mesmo tempo
  $P(A \text{ e } B) = 0$

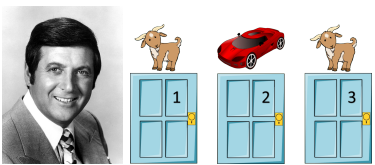
o problema de Monty Hall

um programa de auditório:

- você tem 3 portas na sua frente: uma contém uma Ferrari e as duas outras um bode em cada uma
- MH pede para você escolher uma delas: você escolhe, por exemplo, a 1
- antes de abrir a porta que você escolheu, MH (que sabe onde o carro está), escolhe uma das portas com um bode e a abre; sobra assim uma outra porta fechada

- MH então te pergunta: quer trocar de porta ou não quer? Você quer trocar a porta que escolheu por esta que sobrou ou não?

- o que é mais vantajoso fazer?

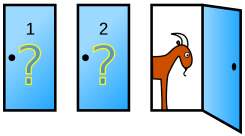


o problema de Monty Hall

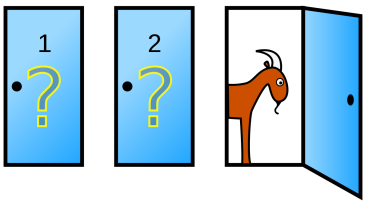
- digamos que voce escolheu a porta 1 e a Ferrari está na 2
- MH então abre a 3, mostrando um bode
- você, sem saber onde está a Ferrari, continua com a porta 1 ou muda para a 2?
- vamos ver com qual ação você tem maior probabilidade de ganhar a Ferrari!
- $P(ci)$: probabilidade a priori de que o carro esteja atrás da porta i:
  $P(c1) = P(c2) = P(c3) = 1/3$
- $P(i)$: probabilidade de que MH abra a porta i
- $P(ci|j)$: probabilidade de que o carro esteja atrás da porta i se MH abriu a porta j
- como MH abriu a porta 3, precisamos comparar $P(c1|3)$ com $P(c2|3)$

- probabilidade de o carro estar na porta 1 dado que MH abriu a porta 3:
  $P(c1|3) = P(3|c1)P(c1)/P(3)$
- probabilidade de o carro estar na porta 2 dado que MH abriu a porta 3
  $P(c2|3) = P(3|c2)P(c2)/P(3)$
  onde
  $P(3) = P(3|c1)P(c1) + P(3|c2)P(c2) + P(3|c3)P(c3)$

- qual é maior? $P(c1|3)$ ou $P(c2|3)$?



o problema de Monty Hall



- qual é maior? $P(c1|3)$ ou $P(c2|3)$?
- probabilidades relevantes:

$P(c1) = P(c2) = P(c3) = 1/3$

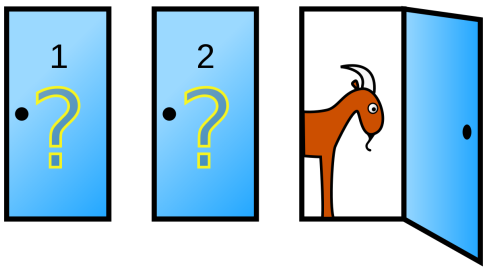
$P(c1|3) = P(3|c1)P(c1)/P(3)$

$P(c2|3) = P(3|c2)P(c2)/P(3)$

$P(3) = P(3|c1)P(c1) + P(3|c2)P(c2) + P(3|c3)P(c3)$

- mas
  $P(3|c1) = 1/2$: ele tinha duas portas para abrir
  $P(3|c2) = 1$: como o carro estava em 2, ele só podia abrir a 3
  $P(3|c3) = 0$: 3 tem um bode, por isso MH a abre
- logo,
  $P(c1|3) =$
  $(1/2 \times 1/3)/(1/2 \times 1/3 + 1 \times 1/3 + 0 \times 1/3) = 1/3$
- se trocar de porta:
  $P(c2|3) = P(3|c2)P(c2)/P(3) = 2/3$

é vantajoso trocar!!

o problema de Monty Hall

- quando você escolheu uma das portas, a chance de ter um bode atrás dela era 2/3
- se você muda de porta, a chance de ter uma Ferrari atrás dela é 2/3
- logo, convém mudar de porta!