Le raisonnement probabiliste

Rentrée Cogmaster 2009

M. Cozic

Paris 12, IHPST & DEC

introduction

raisonner dans l'incertain

- la plupart de nos raisonnements, dans la vie ordinaire et dans l'activité scientifique, comportent une part d'incertitude
- exemple: qui gagnera la L1 cette année ?

raisonner dans l'incertain

- la plupart de nos raisonnements, dans la vie ordinaire et dans l'activité scientifique, comportent une part d'incertitude
  - exemple: qui gagnera la L1 cette année ?
- la logique (déductive) donne peu d'outils pour raisonner dans l'incertitude

raisonner dans l'incertain

- la plupart de nos raisonnements, dans la vie ordinaire et dans l'activité scientifique, comportent une part d'incertitude
- exemple: qui gagnera la L1 cette année ?
- la logique (déductive) donne peu d'outils pour raisonner dans l'incertitude
- la théorie qui prend en charge le raisonnement dans l'incertitude est la théorie des probabilités

notions élémentaires

espace d'états et événements

- espace d'états S = ensemble des issues possibles d'une certaine expérience
  - exemple : lancer de dé
    $S = \{1, 2, 3, 4, 5, 6\}$

espace d'états et événements

- espace d'états S = ensemble des issues possibles d'une certaine expérience
  - exemple : lancer de dé
    $S = \{1, 2, 3, 4, 5, 6\}$
- on peut être intéressé par (ou avoir des informations sur) non pas des états de S mais des événements
  - exemple : événement "le dé est tombé sur une face paire"
    $E = \{2, 4, 6\}$

espace d'états et événements

- espace d'états S = ensemble des issues possibles d'une certaine expérience
  - exemple : lancer de dé
    $S = \{1, 2, 3, 4, 5, 6\}$
- on peut être intéressé par (ou avoir des informations sur) non pas des états de S mais des événements
  - exemple : événement "le dé est tombé sur une face paire"
    $E = \{2, 4, 6\}$
- les événements sont donc des ensembles d'états

opérations ensemblistes

- opérations ensemblistes

(i) complémentation:

- l'événement E^c = "le dé n'est pas tombé sur une face paire" = $S - E = \{1, 3, 5\}$

opérations ensemblistes

- opérations ensemblistes

(i) complémentation:

- l'événément E^c = "le dé n'est pas tombé sur une face paire" = $S - E = \{1, 3, 5\}$

(ii) union:

- E' le dé est tombé sur une face ≤ 3"
  l'événément $E \cup E'$ = "le dé est tombé sur une face paire ou ≤ 3" = $\{1, 2, 3, 4, 6\}$

opérations ensemblistes

- opérations ensemblistes

(i) complémentation:

- l'événement E^c = "le dé n'est pas tombé sur une face paire" = $S - E = \{1, 3, 5\}$

(ii) union:

- E' le dé est tombé sur une face ≤ 3"
  l'événement $E \cup E'$ = "le dé est tombé sur une face paire ou ≤ 3" = $\{1, 2, 3, 4, 6\}$

(iii) intersection:

- l'événement $E \cap E'$ = "le dé est tombé sur une face paire et ≤ 3" = $\{2\}$

probabilités

- la représentation probabiliste de l'incertitude = assigner un poids aux événements en respectant certains axiomes

probabilités

- la représentation probabiliste de l'incertitude = assigner un poids aux événements en respectant certains axiomes
- on peut assigner ce poids à l'ensemble de tous les événéments possibles $\wp(S)$, ou seulement à une algèbre d'événements déterminée $\mathcal{E} \subseteq \wp(S)$

probabilités

- la représentation probabiliste de l'incertitude = assigner un poids aux événements en respectant certains axiomes
- on peut assigner ce poids à l'ensemble de tous les événéments possibles $\wp(S)$, ou seulement à une algèbre d'événements déterminée $\mathcal{E} \subseteq \wp(S)$
- axiomes des probabilités:

(A1) $P(S) = 1$

(A2) $P(\emptyset) = 0$

(A3) $P(E \cup E') = P(E) + P(E')$ si E et E' disjoints (i.e. $E \cap E' = \emptyset$)

exemple

- exemple : la probabilité uniforme sur $S = \{1, 2, 3, 4, 5, 6\}$
  $P(\{i\}) = 1/6$
  $P(E) = \sum_{i \in E} P(\{i\})$

exemple

- exemple : la probabilité uniforme sur $S = \{1, 2, 3, 4, 5, 6\}$
  $P(\{i\}) = 1/6$
  $P(E) = \sum_{i \in E} P(\{i\})$
- la probabilité n'est pas nécessairement uniforme
- pièce biaisée en faveur de face F: $S = \{F, P\}$, $P(F) = 0,8$

la probabilité conditionnelle

la probabilité conditionnelle

- $P(A|B)$: probabilité que A soit réalisé sachant que B s'est réalisé
  - exemple: probabilité que le dé soit tombé sur une face paire sachant qu'il est tombé sur une face ≤ 3

la probabilité conditionnelle

- $P(A|B)$: probabilité que A soit réalisé sachant que B s'est réalisé
  - exemple: probabilité que le dé soit tombé sur une face paire sachant qu'il est tombé sur une face ≤ 3
- comment évaluer cette nouvelle probabilité ?

la probabilité conditionnelle

- $P(A|B)$: probabilité que A soit réalisé sachant que B s'est réalisé
  - exemple: probabilité que le dé soit tombé sur une face paire sachant qu'il est tombé sur une face ≤ 3
- comment évaluer cette nouvelle probabilité ?
  - exemple
    - les états où B n'est pas le cas sont écartés et reçoivent une probabilité nulle
    - parmi les B-états (1, 2, 3), il y a le A-état 2
    - le poids relatif des A-états parmi les B-états est 1/3

la probabilité conditionnelle

- soient deux événements $A, B \in \mathcal{E}$ et $P(B) > 0$
  $P(A|B) =_{df} P(A \cap B)/P(B)$

la probabilité conditionnelle

- soient deux événements $A, B \in \mathcal{E}$ et $P(B) > 0$
  $P(A|B) =_{df} P(A \cap B)/P(B)$
  $P(A)$ = probabilité a priori de A
  $P(A|B)$ = probabilité a posteriori de A

la probabilité conditionnelle

- soient deux événements $A, B \in \mathcal{E}$ et $P(B) > 0$
  $P(A|B) =_{df} P(A \cap B)/P(B)$
  $P(A)$ = probabilité a priori de A
  $P(A|B)$ = probabilité a posteriori de A
- exemple: $P(B|A) = (1/6)/(1/2) = 1/3$

théorème des probabilités totales

- $A = AB \cup AB^c$ donc par additivité $P(A) = P(AB) \cup P(AB^c)$ soit

$$\boxed{P(A) = P(A|B) \cdot P(B) + P(A|B^c) \cdot P(B^c)}$$

théorème des probabilités totales

- $A = AB \cup AB^c$ donc par additivité $P(A) = P(AB) \cup P(AB^c)$ soit

$$\boxed{P(A) = P(A|B) \cdot P(B) + P(A|B^c) \cdot P(B^c)}$$

- de manière générale, si $(B_i)_{i \in I}$ est une partition de S,

$$\boxed{P(A) = \sum_{i \in I} P(A|B_i) \cdot P(B_i)}$$

exemple

- la maladie M est présente chez 1/1000 individu
- test pour détecter la maladie M: quand l'individu est malade (M), le test est positif 9/10 ; quand l'individu est sain, le test est négatif 8/10

exemple

- la maladie M est présente chez 1/1000 individu
- test pour détecter la maladie M: quand l'individu est malade (M), le test est positif 9/10 ; quand l'individu est sain, le test est négatif 8/10
- quelle est la probabilité qu'un test se révèle positif ?

exemple

- la maladie M est présente chez 1/1000 individu
- test pour détecter la maladie M: quand l'individu est malade (M), le test est positif 9/10 ; quand l'individu est sain, le test est négatif 8/10
- quelle est la probabilité qu'un test se révèle positif ?
- $P(POS) = P(POS|M) \cdot P(M) + P(POS|M^c) \cdot P(M^c)$
  $P(POS) = 9/10 \cdot 1/1000 + 2/10 \cdot 999/1000 \approx 0,2$

théorème de Bayes

- un autre théorème fondamental pour le raisonnement probabiliste, le théorème de Bayes:

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

théorème de Bayes

- un autre théorème fondamental pour le raisonnement probabiliste, le théorème de Bayes:

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

- quel intérêt d'exprimer $P(A|B)$ en termes de $P(B|A)$

théorème de Bayes

- un autre théorème fondamental pour le raisonnement probabiliste, le théorème de Bayes:

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

- quel intérêt d'exprimer $P(A|B)$ en termes de $P(B|A)$
- il arrive souvent que l'on soit intéressé par $P(A|B)$ mais que l'on connaisse plutôt $P(B|A)$
- exemple: diagnostic
  $P(maladie|symptome) =$
  $P(symptome|maladie) \cdot P(maladie)/P(symptome)$
  pourquoi connaîtrait-on mieux $P(symptome|maladie)$?

théorème de Bayes

- un autre théorème fondamental pour le raisonnement probabiliste, le théorème de Bayes:

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

- quel intérêt d'exprimer $P(A|B)$ en termes de $P(B|A)$
- il arrive souvent que l'on soit intéressé par $P(A|B)$ mais que l'on connaisse plutôt $P(B|A)$
  - exemple: diagnostic
    $P(maladie|symptome) =$
    $P(symptome|maladie) \cdot P(maladie)/P(symptome)$
    pourquoi connaîtrait-on mieux $P(symptome|maladie)$?
    $P(symptome|maladie)$ peut être invariant d'une population à l'autre tandis que $P(maladie|symptome)$ dépend aussi de $P(maladie)$ donc (en gros) de la fréquence de la maladie dans la population du patient

exemple: vrais positifs

- exemple: M touche 1/1000 individu ; si un individu est malade, le test est positif 999/1000 ; s'il est sain, le test est négatif 999/1000
  quelle est la probabilité pour qu'un individu testé positif soit malade ?
  $P(M|POS) = P(POS|M) \cdot P(M)/P(POS)$

exemple: vrais positifs

- exemple: M touche 1/1000 individu ; si un individu est malade, le test est positif 999/1000 ; s'il est sain, le test est négatif 999/1000
  quelle est la probabilité pour qu'un individu testé positif soit malade ?
  $P(M|POS) = P(POS|M) \cdot P(M)/P(POS)$
- comment déterminer $P(POS)$?
  $P(POS) = P(POS \cap M) + P(POS \cap M^c)$
  $P(POS) = P(POS|M) \cdot P(M) + P(POS|M^c) \cdot P(M^c)$
  $P(POS) = 999/1000 \cdot 1/1000 + 1/1000 \cdot 999/1000$
  donc
  $P(M|POS) = (999/1000 \cdot 1/1000)/(999/1000 \cdot 1/1000 + 1/1000 \cdot 999/1000) = 1/2$

indépendance

- deux événements A et B sont indépendants si $P(A \cap B) = P(A) \cdot P(B)$
- quand $P(B) > 0$, ceci est équivalent à dire $P(A|B) = P(A)$
- autrement dit: A et B sont indépendants quand, être informé de l'un des événements ne nous conduit pas à changer les probabilités de l'autre.
- on le note $A \perp B$.

indépendance, exemple

- exemple: on lance indépendamment deux pièces non biaisées
  $S = \{PP, FF, PF, FP\}$
  B = la première pièce tombe sur face = $\{FF, FP\}$
  A = les deux pièces tombent du même côté = $\{FF, PP\}$
- A et B sont-ils indépendants ?
  $P(A|B) = P(A \cap B)/P(B)$
  $P(A|B) = P((FF \cup PP) \cap (FF \cup FP))/P(FF \cup FP)$
  $P(A|B) = P(FF)/P(FF \cup FP) = 1/2$

indépendance conditionnelle

- A et B sont conditionnellement indépendants étant donné C si

$$\boxed{P(A|BC) = P(A|C)}$$

- idée: une fois que je sais que C est réalisé, apprendre que B ne provoque pas de changement de la probabilité de A

exemple

- soit les événements suivants:
  B le baromètre chute
  P la pression atmosphérique chute
  O il y a de l'orage

exemple

- soit les événements suivants:
  B le baromètre chute
  P la pression atmosphérique chute
  O il y a de l'orage
- ces événements ne sont (intuitivement) pas indépendants: $P(B|P) \neq P(B)$, $P(O|P) \neq P(O)$, $P(B|O) \neq P(B)$
- mais B et O ne sont pas conditionnellement indépendants étant donné P: l'info. que la pression atmo. chute ne m'amène pas à estimer la probabilité que le baromètre chute de manière différente que ne l'aurait fait l'info. que la pression atmo. chute et qu'il y a de l'orage

causalité et indépendance conditionnelle

- dans l'ex. précédent, il y a corrélation (proba.) entre B et O mais pas de lien de causalité
- la corrélation s'explique par une cause commune P
- idée générale: utiliser l'indépendance conditionnelle pour discriminer les corrélations causales des corrélations non-causales (Reichenbach, "P screens off B from C"