

# Inférence Bayésienne

Bruno Bouzy

[bruno.bouzy@u-paris.fr](mailto:bruno.bouzy@u-paris.fr)

Janvier 2022

# Des oiseaux noirs... ou gris...

## Des corneilles et des corbeaux...



Des corbeaux et des corneilles :  
Plutôt grises ou plutôt noires ?

Un oiseau : plutôt une corneille ou plutôt un corbeau ?

Un oiseau au plumage noir :  
plutôt une corneille ou plutôt un corbeau ?

Images récupérées de <https://lemagdesanimaux.ouest-france.fr/dossier-219-differences-corbeau-corneille.html>

Inférence bayésienne

# Généralités

- Problème de classification d'objets
  - Un objet possède une(des) caractéristique(s), disons  $X$  (la couleur du plumage)
  - Il appartient à une classe, disons  $C$  (corneille ou corbeau)
- Valeurs
  - Une caractéristique  $X$  prend des valeurs  $X_1, X_2, \dots, X_l, \dots, X_L$ . (noir ou gris)
  - Les classes sont  $C_1, C_2, \dots, C_k, \dots, C_K$ .
- Contexte probabiliste
  - Etant donné un objet dont la caractéristique vaut  $X_l$ , il a une probabilité  $P(X_l)$  d'être observé.
  - Etant donné un objet de la classe  $C_k$ , il a une probabilité  $P(C_k)$  d'être observé.

# Types de probabilités

- Probabilité à priori  $P(C_k)$ 
  - Probabilité que l'objet soit dans la classe  $C_k$  sans rien connaître d'autre sur l'objet. (un oiseau est-il une corneille ou un oiseau?)
- Probabilité jointe  $P(C_k, X_l) = P(X_l, C_k)$ 
  - Probabilité que l'objet soit dans la classe  $C_k$  avec une caractéristique  $X_l$ . (proba de voir un corbeau noir par exemple)
- Probabilités conditionnelles  $P(C_k | X_l)$  ou  $P(X_l | C_k)$ 
  - $P(C_k | X_l)$  est la probabilité que l'objet soit dans la classe  $C_k$  sachant que sa caractéristique vaut  $X_l$ . (sachant que l'oiseau est noir, quelle est la proba que ce soit un corbeau?)
  - $P(X_l | C_k)$  est la probabilité que l'objet ait la caractéristique  $X_l$  sachant qu'il appartient à la classe  $C_k$ . (sachant que l'oiseau est une corneille, quelle est la proba qu'elle soit noire?)

# Formules probabilistes (1/2)

- Formules de la proba conditionnelle :
  - $P(C_k, X_l) = P(C_k | X_l) P(X_l)$  (1)
  - $P(X_l, C_k) = P(X_l | C_k) P(C_k)$  (2)
- Formule de Bayes :
  - $P(C_k | X_l) P(X_l) = P(X_l | C_k) P(C_k)$  (3)
- Formule de la probabilité à postériori :
  - $P(C_k | X_l) = P(X_l | C_k) P(C_k) / P(X_l)$  (4)
- Avec ces formules, on fait tout... ou presque !

# Formules probabilistes (2/2)

- Deux formules évidentes :

- $\sum_k P(C_k, X_l) = P(X_l) \quad (5)$

- $\sum_l P(C_k, X_l) = P(C_k) \quad (6)$

- Une formule intéressante :

- $P(X_l) = \sum_k P(X_l | C_k) P(C_k) \quad (7)$

- Formule de la proba conditionnelle « bis » :

- $P(C_k | X_l) = P(X_l, C_k) / \sum_k P(C_k, X_l) \quad (8)$

# En pratique

- Contexte probabiliste et empirique
  - Il faut être capable de transformer un jeu de données en probabilités empiriques
    - Être capable de quantifier  
(étant donné un niveau de gris de plumage, décider si l'oiseau est considéré comme noir ou gris)
    - Être capable de compter les objets
      - Appartenant à une classe  $C_k$
      - Ayant  $X_i$  pour caractéristique
      - Les deux en même temps
      - Ou l'un sachant l'autre
  - Que ces compteurs sont suffisamment grands pour que les probabilités soit significatives

Pour que la méthode marche :

- Avoir suffisamment de données
- Plusieurs objets par couple  $(X_i, C_k)$

Inférence bayésienne

# Prendre le problème par un bout

- Si je peux compter tous les objets ayant la caractéristique  $X_l$  et appartenant à la classe  $C_k$ 
  - Alors je peux estimer les  $P(X_l, C_k)$
  - Avec la formule (8) je peux trouver la proba à postériori  $P(C_k|X_l)$
- Si je peux compter les objets de manière à estimer empiriquement les probas  $P(C_k)$  et  $P(X_l | C_k)$ 
  - Alors je peux estimer  $P(X_l)$  avec (7)
  - Et  $P(C_k|X_l)$  avec (8)



# Et la classification dans tout ça ?

- Les données sur la caractéristique  $X$  sont incorrectes ou incomplètes :
  - Les données ne permettent que d'estimer  $P(C_k)$
  - Ne connaissant rien sur l'objet, choisir la classe maximisant  $P(C_k)$  (c'est mieux que rien).
  - Maximisation à priori (j'observe un oiseau mais je ne connais pas sa couleur)
- Les données sur la caractéristique  $X$  sont correctes et complètes
  - Maximisation à postériori
  - L'objet et  $X_l$  sont connus : choisir la classe  $C_k$  maximisant :
    - La probabilité à postériori  $P(C_k|X_l)$
    - La vraisemblance  $P(X_l|C_k)$

Selon les domaines maximiser  $P(C_k|X_l)$  ou maximiser  $P(X_l|C_k)$  est discutable...

- Faut-il choisir la classe  $C_k$  qui maximise la proba que :
  - l'objet appartienne à  $C_k$  sachant  $X_l$  ?
  - $X$  vale  $X_l$  si on suppose que l'objet appartient à  $C_k$  ?
- Là est toute la question !!

# Retour sur les oiseaux !



Des corbeaux et des corneilles :  
Plutôt grises ou plutôt noires ?

Un oiseau au plumage noir :  
plutôt une corneille ou plutôt un  
corbeau ?

Images récupérées de <https://lemagdesanimaux.ouest-france.fr/dossier-219-differences-corbeau-corneille.html>

Inférence bayésienne

# Observations et probabilités

	Gris	Noir	total
Corneille	5	45	50
Corbeau	30	70	100
total	35	115	150

$$\begin{aligned} P(\text{Corneille}) &= 50/150 = 0.333 \\ P(\text{Gris}) &= 35/150 = 0.233 \end{aligned}$$

$$\begin{aligned} P(\text{Corbeau}) &= 100/150 = 0.667 \\ P(\text{Noir}) &= 115/150 = 0.767 \end{aligned}$$

$$\begin{aligned} P(\text{Gris} \mid \text{Corneille}) &= 5/50 = 0.1 \\ P(\text{Gris} \mid \text{Corbeau}) &= 30/100 = 0.3 \end{aligned}$$

$$\begin{aligned} P(\text{Noir} \mid \text{Corneille}) &= 45/50 = 0.9 \\ P(\text{Noir} \mid \text{Corbeau}) &= 70/100 = 0.7 \end{aligned}$$

$$\begin{aligned} P(\text{Corneille, Gris}) &= 5/150 = 0.033 \\ P(\text{Corbeau, Gris}) &= 30/150 = 0.2 \end{aligned}$$

$$\begin{aligned} P(\text{Corneille, Noir}) &= 45/150 = 0.3 \\ P(\text{Corbeau, Noir}) &= 70/150 = 0.467 \end{aligned}$$

$$\begin{aligned} P(\text{Corneille} \mid \text{Gris}) &= 5/35 = 0.142 \\ P(\text{Corneille} \mid \text{Noir}) &= 45/115 = 0.391 \end{aligned}$$

$$\begin{aligned} P(\text{Corbeau} \mid \text{Gris}) &= 30/35 = 0.858 \\ P(\text{Corbeau} \mid \text{Noir}) &= 70/115 = 0.609 \end{aligned}$$

Inférence bayésienne

# Probabilités et raisonnement incorrect

$$P(\text{Corneille}) = 50/150 = 0.333$$

$$P(\text{Corbeau}) = 100/150 = 0.667$$

Il y a plus de corbeaux que de corneilles.

$$P(\text{Gris}) = 35/150 = 0.233$$

$$P(\text{Noir}) = 115/150 = 0.767$$

Il y a plus d'oiseaux noirs que d'oiseaux gris.

$$P(\text{Gris} \mid \text{Corneille}) = 5/50 = 0.1$$

$$P(\text{Noir} \mid \text{Corneille}) = 45/50 = 0.9$$

$$P(\text{Gris} \mid \text{Corbeau}) = 30/70 = 0.3$$

$$P(\text{Noir} \mid \text{Corbeau}) = 70/100 = 0.7$$

La probabilité qu'une corneille soit noire est plus grande que celle qu'un corbeau soit noir.

Raisonnement incorrect:

Les corneilles ont plus de chances d'être noires que les corbeaux. ==> Il y a plus de chances qu'un oiseau noir soit une corneille qu'un corbeau.

$$P(\text{Corneille}, \text{Gris}) = 5/150 = 0.033$$

$$P(\text{Corneille}, \text{Noir}) = 45/150 = 0.3$$

$$P(\text{Corbeau}, \text{Gris}) = 30/150 = 0.2$$

$$P(\text{Corbeau}, \text{Noir}) = 70/150 = 0.467$$

$$P(\text{Corneille} \mid \text{Gris}) = 5/35 = 0.142$$

$$P(\text{Corbeau} \mid \text{Gris}) = 30/35 = 0.858$$

$$P(\text{Corneille} \mid \text{Noir}) = 45/115 = 0.391$$

$$P(\text{Corbeau} \mid \text{Noir}) = 70/115 = 0.609$$

Correct : Il est plus probable qu'un oiseau noir soit un corbeau plutôt qu'une corneille.

Inférence bayésienne

# Merci de votre attention !!