Chapitre 7

Méthodes inductives

1. L'induction
2. Formule de Bayes
3. Maximum de vraisemblance

L'induction

Exemple 1 : tir au canon

1. Méthode déductive :
si on connaît les lois du mouvement : on résoud les équations, on calcule la trajectoire, on détermine le point d'impact et l'erreur probable

2. Méthode inductive :
si on ne connaît pas les lois du mouvement : on procède à des tirs sans modifier la position du canon, on mesure le point d'impact moyen et la dispersion, on répète l'expérience pour plusieurs positions du canon, on essaie de trouver une loi empirique

Exemple 2 : choix aveugle d'une boule dans un sac

1. Si on connaît le contenu du sac : on peut calculer les probabilités

2. Si on ne connaît pas le contenu du sac : on peut estimer les probabilités, avec une confiance qui croît avec le nombre d'essais

Une part importante de la science est basée sur l'induction.

En logique déductive, une proposition est soit vraie, soit fausse. On peut faire correspondre le nombre 0 à une proposition fausse et le nombre 1 à une proposition vraie

$$\Rightarrow \text{algèbre booléenne}$$

Pour rendre l'induction quantitative, il faut faire correspondre des nombres aux propositions.

Une proposition peut être confirmée à des degrés divers, depuis la preuve logique jusqu'à la réfutation. On peut donc faire correspondre au degré de confirmation des nombres appartenant à un intervalle fini.

Soient :

p les données de base

q une proposition

a le nombre qui exprime le degré de confirmation de q par les données p

$$\Rightarrow \text{on écrit } P(\,q \mid p\,) = a$$

Le degré de confirmation doit obéir à un certain nombre d'axiomes.

Axiome 1 :

Etant donné p, si q et r sont deux propositions, soit q est plus probable que r, soit q est moins probable que r, soit q et r sont aussi probables l'une que l'autre. Deux de ces affirmations ne peuvent être vraies simultanément.

Axiome 2 :

La relation «*plus probable que*» est transitive : étant donné p, si q est plus probable que r et r plus probable que s, alors q est plus probable que s.

Axiome 3 :

Etant donné p,
(1) si q et r s'excluent mutuellement,
(2) si q' et r' s'excluent mutuellement,
(3) si q et q' sont confirmées au même degré,
(4) si r et r' sont confirmées au même degré,
alors $q \vee r$ et $q' \vee r'$ sont confirmées au même degré.

Règle d'addition :

Etant donné p, si q et r s'excluent mutuellement, alors les nombres correspondant aux degrés de confirmation de q, r et $q \vee r$ satisfont à la règle :

$$P(q \vee r \mid p) = P(q \mid p) + P(r \mid p)$$

Conséquence :

$$0 \leq P(q \mid p) \leq \beta$$

On adopte $\beta = 1$

Axiome 4 :

$$P(q \,.\, r \mid p) = P(q \mid p)\, P(r \mid q \,.\, p)$$

(= règle de multiplication)

$$\Rightarrow \textit{degrés de confirmation} = \textit{probabilités}$$

Formule de Bayes

Soient

p les données initiales

r des données supplémentaires

$q_1, \ldots, q_n$ un ensemble d'hypothèses

$$P(\, q_i \,.\, r \mid p\,) = P(\, q_i \mid p\,)\, P(\, r \mid q_i \,.\, p\,)$$

$$P(\, q_i \,.\, r \mid p\,) = P(\, r \mid p\,)\, P(\, q_i \mid r \,.\, p\,)$$

$$\Rightarrow \frac{1}{P(\, r \mid p\,)} = \frac{P(\, q_i \mid r \,.\, p\,)}{P(\, q_i \mid p\,)\, P(\, r \mid q_i \,.\, p\,)}$$

est indépendant de i

Si les q_i sont mutuellement exclusifs et exhaustifs :

$$\Rightarrow P(\, q_i \mid r \,.\, p\,) = \frac{P(\, q_i \mid p\,)\, P(\, r \mid q_i \,.\, p\,)}{\Sigma_{j=1}^{n} P(\, q_j \mid p\,)\, P(\, r \mid q_j \,.\, p\,)}$$

Formule de Bayes :

$$P(q_i \mid r \,.\, p) = \frac{P(q_i \mid p)\, P(r \mid q_i \,.\, p)}{\Sigma_{j=1}^{n} P(q_j \mid p)\, P(r \mid q_j \,.\, p)}$$

$P(q_i \mid p)$ = **probabilité a priori**
(avant les données supplémentaires r)

$P(q_i \mid r \,.\, p)$ = **probabilité a posteriori**
(après considération de r)

$P(r \mid q_i \,.\, p)$ = **vraisemblance** = « probabilité directe »
(= probabilité du résultat r sous l'hypothèse q_i et les données de base p)

$\Rightarrow$ Probabilité a posteriori $\propto$
Probabilité a priori $\times$ Vraisemblance

Exemples :

(1) Si r a une probabilité faible sur toutes les hypothèses sauf une (q_k) : (test crucial)

$$P(r \mid q_k \,.\, p) \gg P(r \mid q_i \,.\, p) \quad \forall i \neq k$$
$$\Rightarrow P(q_k \mid r \,.\, p) \simeq 1$$

(2) Si $P(r \mid q_k \,.\, p) = 1$ et si $-r$ est vérifié

$$\Rightarrow P(q_k \mid -r \,.\, p) = 0$$

l'échec d'un test crucial entraîne le rejet d'une hypothèse *a priori* plausible

(3) Si $P(r \mid q_k \,.\, p)$ est la même $\forall k$

$$\Rightarrow P(q_k \mid r \,.\, p) = P(q_k \mid p)$$

les nouvelles données n'apportent rien

(4) Si $P(q_k \mid p) = 0$ alors $P(q_k \mid r \,.\, p) = 0$

De nouvelles données ne modifient rien à une hypothèse déjà connue pour être fausse

Maximum de vraisemblance

On désire trouver le modèle qui rend le mieux compte des données

Soient
p les données initiales
r des données supplémentaires
$q_1, \ldots, q_n$ un ensemble de modèles

Nous choisirons le modèle le plus probable, compte tenu des connaissances de base p et des données à modéliser r

Il faut donc trouver le modèle qui maximise la probabilité *a posteriori*

$$P(\, q_i \mid r \,.\, p\,)$$

Il est souvent plus commode de trouver le maximum du logarithme de P :

$$\ln P(\, q_i \mid r \,.\, p\,) = \ln P(\, q_i \mid p\,) + \ln P(\, r \mid q_i \,.\, p\,) + C$$

Dans le cas où tous les modèles considérés sont ***a priori*** équiprobables, il suffit de trouver le modèle qui maximise

$$\ln P(\, r \mid q_i \,.\, p\,)$$

$\Rightarrow$ *maximum de vraisemblance*