Modèles probabilistes

C. Lalanne, J.-B. Poline

Cogmaster 2006–2007, Bloc thématique 4

Objet de ce cours

- Introduction aux modèles probabilistes et à la théorie de l'information
  - Notions fondamentales
  - Applications

Organisation

- Cours (3h)
  - Objet de la modélisation
  - Rappels de probabilités
  - Introduction à la théorie de l’information
  - Applications
- TP (3h)
- Synthèse (1h)

Plan

1. Principes de la modélisation

§ Rappels de probabilités

§ Théorie de l'information

§ Applications

Qu'est-ce qu'un modèle ?

- Une définition *opérationnelle* :

  Un modèle est une théorie orientée vers l'action qu'elle doit servir.

  ⇒ connotation pragmatique

- Une définition *mathématique* :

  « La modélisation mathématique est le processus par lequel un problème du monde réel est interprété et représenté en termes de symboles abstraits. La description abstraite faisant intervenir une formulation mathématique est appelée *Modèle Mathématique* associé au problème de départ. Le dit problème, issu du réel, peut être alors traité uniquement en termes mathématiques. » (Y Cherruault. *Modèles et méthodes mathématiques pour les sciences du vivant*, PUF, 1998)

  ⇒ idée de formalisation

Principes de la modélisation (1)

- L'activité de modélisation consiste à dégager une version réduite de la réalité à l'aide d'un **modèle abstrait/formel**, qui permet
  - la simulation (informatique) du comportement du système,
  - la prédiction de son évolution possible,
  - la comparaison avec des données issues de l'expérimentation

Principes de la modélisation (2)

- On a besoin de
  - concepts permettant de décrire le phénomène ou système considéré
  - relations fonctionnelles liant ces concepts

  ⇒ Outils probabilistes pour étudier un modèle formel, en tenant compte de données réelles.

- Démarche ascendante : complexification croissante du modèle initial

Pourquoi modéliser ?

- La modélisation est utile pour mieux comprendre les systèmes complexes
- Cas des **Sciences Cognitives** : étude et compréhension des mécanismes de la pensée
  - raisonnement
  - langage
  - habiletés sensorimotrices
  - etc.

De la réalité au modèle

- La description d’un phénomène à l’aide d’un modèle implique une réduction de la réalité, donc une **perte d’information** (approximations)
- La modélisation permet également l’extrapolation de certaines propriétés ou la prédiction de certains comportements
- Dans tous les cas, il y a perte d’information et présence d’incertitude (ou bruit)

Exemples

- Modèle proie-prédateur
- Modèle de diffusion
- Modèle de détection du signal (observateur idéal)
- Modèle d'intégration de sources multiples d'informations
- Modèle d'évolution du marché (macro-/micro-économie)

Lien avec la démarche statistique

- Objectif : minimiser incertitude et erreur dans le modèle
- Compromis variance minimale / erreur minimale lors de l'estimation des paramètres du modèle
  - ⇒ On préfèrera généralement, quand on a le choix, minimiser la variance

Plan

1. Principes de la modélisation

§ Rappels de probabilités

§ Théorie de l'information

§ Applications

L'approche probabiliste

- Permet de prendre en compte
  - les erreurs de mesure,
  - l'incertitude liée à la connaissance d'un phénomène suite à une observation unique,
  - des connaissances préalables dans l'estimation des caractéristiques d'un système
- Nécessaire à la réalisation de simulations de systèmes complexes

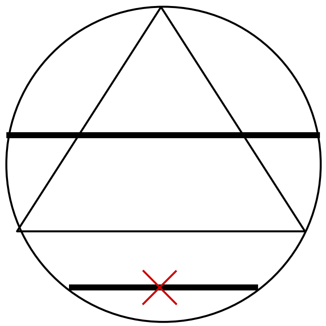
Qu'est-ce que le hasard ?

- Multiplicité des définitions
- e.g. « la rencontre de deux séries causales indépendantes »

Le paradoxe de Bertrand (1)

- On considère un triangle équilatéral et son cercle circonscrit. On tire une corde au hasard.

  Quelle est la probabilité que sa longueur soit supérieure à celle du côté du triangle ?

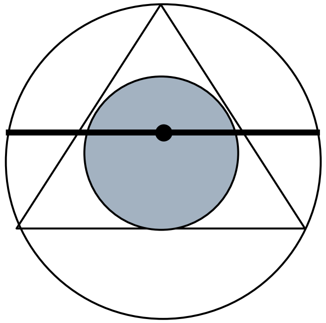


Le paradoxe de Bertrand (2)

- **1ère solution** : choix aléatoire du milieu de la corde

Alors, P(« corde plus longue que le côté ») = P(« milieu de la corde $\in$ cercle inscrit »)

Soit $$P = \frac{\pi(\rho/2)^2}{\pi\rho^2} = \frac{1}{4}$$

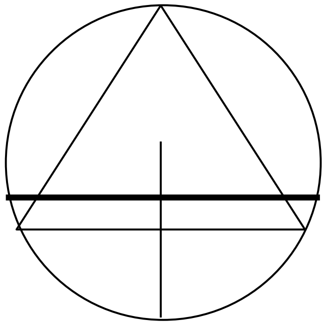


Le paradoxe de Bertrand (3)

- **2ème solution** : distance du milieu de la corde au centre du cercle

On considère que le milieu de la corde est pris aléatoirement sur un rayon donné du cercle (symétrie).

La corde sera plus longue que le côté du triangle inscrit si son milieu est à une distance du centre < r/2



D'où $P = \frac{1}{2}$

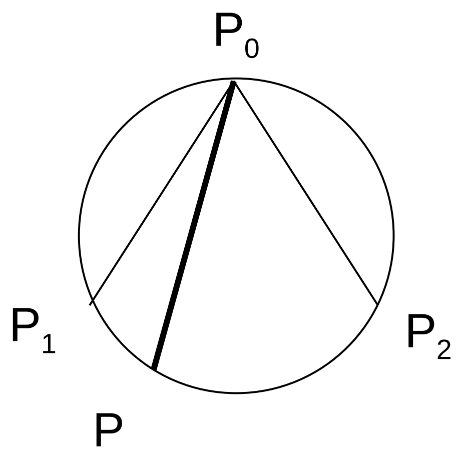
Le paradoxe de Bertrand (4)

- **3ème solution** :
  choix aléatoire d'une extrémité de la corde, l'autre étant fixée

Soit P_0 l'extrémité choisie. Si on admet que la probabilité que l'autre extrémité P tombe sur un arc donné de la circonférence est proportionnelle à la longueur de cet arc, alors P_0P est plus grand que le côté du triangle inscrit quand P se trouve sur $\overset{\frown}{P_1P_2}$.

Donc, la longueur est le 1/3 de celle de la circonférence et

$$P = \frac{1}{3}$$



Le paradoxe de Bertrand (5)

- En résumé :
  - Trois hypothèses de répartition également réalisables
  - Même algèbre d'événements
  - Trois probabilités différentes
- Conclusion :
  - problème du « choix du hasard »
  - limite de la conception objectiviste

- Pour en savoir plus
  - http://www-ensps.u-strasbg.fr/enseignants/harthong/Hist/BERTRAND.HTM
  - http://perso.orange.fr/therese.eveilleau/pages/paradoxe/textes/bertrand.htm
  - http://www-irem.univ-fcomte.fr/bulletins/067/067-article1-paradoxe-Bertrand.html
  - http://www.trigofacile.com/maths/curiosite/index.htm

Le concept de probabilité (1)

- Historique :
  - *Pascal* – répartition de la mise après interruption d'un jeu de pari en 3 parties

    $\Rightarrow$ Le gain attribué à un des 2 joueurs = moyenne pondérée des gains possibles si le jeu se termine

    $\Rightarrow$ à partir de cette espérance mathématique, on en déduit la probabilité (de gain)
  - Depuis l'axiomatisation de *Kolmogorov* (1930), on fait l'inverse (probabilité $\rightarrow$ espérance), et la probabilité est considérée comme une **mesure**

Le concept de probabilité (2)

- Définition :

  Si A, B, C, … sont des ensembles munis d'une fonction *m* dont les propriétés sont

  - à valeurs positives

  - croissante relativement à l'inclusion

  $$A \subset B \Rightarrow \mu(A) \leq \mu(B)$$

  - additive pour la réunion de 2 ensembles disjoints

  $$\mu(A \text{ Ç } B) = m(A) + m(B)$$

  alors A, B, C, … sont dits mesurables (par *m*).

Le concept de probabilité (3)

- Vocabulaire
  - expérience aléatoire,
  - espace des possibles (Ω),
  - événement (élémentaire, ou sous-ensemble de Ω)

Avant → expérience aléatoire → **Après**

ensemble de résultats connus (Ω)

1 résultat, élément de l'ensemble

Le concept de probabilité (4)

- Exemple :
  - Jeu pile/face
    - exp. aléa. = jeter une pièce
    - $\Omega = \{P;F\}$
    - un événement = observer « pile »
  - Jeu de dés
    - exp. aléa. = lancer un dé
    - $\Omega = \{1;2;3;4;5;6\}$
    - un événement = observer « 1 »

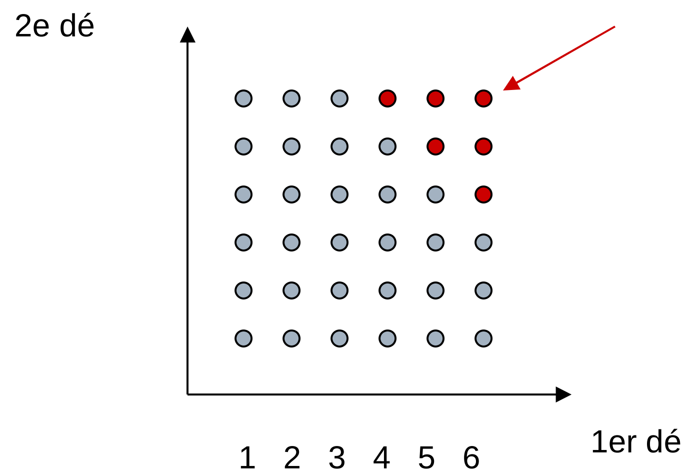
Le concept de probabilité (5)

- Exemple :

  Lancer de 2 dés (réguliers)

  E : « la somme est > 9 »

  $\Rightarrow$ événement réalisé par 6 résultats



Définition (1)

- Définition (classique et idéale) de la probabilité :

$$P(A) = \frac{\nu_A}{\nu}$$

$\Rightarrow$ Si tous les événements observables (Ω) sont également possibles, la probabilité d'observer A est égale au nombre de fois où A a été observé rapporté à l'ensemble des événements observés.

- Exemple :

  jeu infini de pile ou face ou lancer de dé

  - P(« Face ») = 1/2
  - P(« {1;2} ») = 1/3

Définition (2)

- Définition **fréquentiste** (objective) :

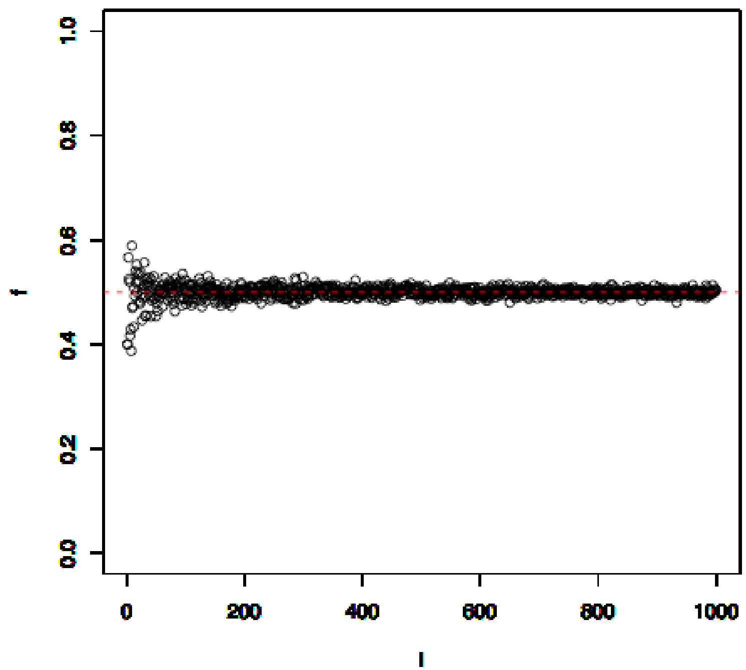
$$P(A) = \lim_{\nu \to \infty} \frac{\nu_A}{\nu}$$

$\Longrightarrow$ Lorsque l'expérience est répétée un (très) grand nombre de fois, la fréquence relative de l'événement approche une valeur constante

- Loi des grands nombres (faible/forte)

Définition (3)

Simulation d'un lancer de pièce (n=1000) :



Définition (4)

- Définition **bayésienne** (subjective) :
  - La définition fréquentiste implique la possibillité de répéter indéfiniment *et* à l’identique l’expérience : peu réaliste !
  - Probabilité d’un événement = mesure du degré de croyance (personnelle) dans l’occurrence de celui-ci
  - Rôle des connaissances *a priori* (‘prior distribution’ dans le processus d’inférence bayésienne en statistique)

Définition (5)

- Événement A
  - (1) valeur que j'attribue à A
  - (2) Valeur intrinsèque de A, indépendante de (1)
  - (1) $\leq$ (2)

P(A) = valeur attribuée / valeur intrinsèque

- Analogie avec les paris : « ce que je suis prêt à parier » *vs.* « ce que je devrais parier si j'étais sûr que A se produit »

Axiomes des probabilités (1)

- Propriétés fondamentales (Axiomes) :

$$P_i \geq 0$$

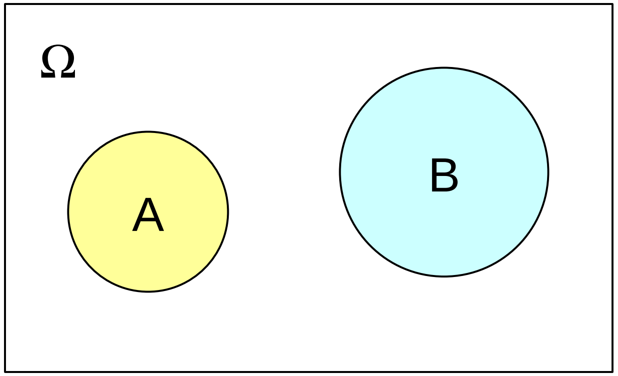
$$P(\Omega) = 1$$

si A et B disjoints, $P(A \cup B) = \Pi(A) + \Pi(B)$

gŽnŽralisation :

si $A_1, A_2 \mathrm{K}\ A_k$ disjoints,

$$P(\mathrm{U}A_i) = \sum_{\iota=1}^{\kappa} \Pi(A_\iota)$$



Axiomes des probabilités (2)

- Conséquences :

$$A \subset B \Rightarrow \Pi(A) \leq \Pi(B) \qquad (*)$$

$$P(A) \leq 1 \; \chi\alpha\rho \; A \subset \Omega$$

$$P(A) + \Pi(\bar{A}) = 1$$

$$\left.\begin{array}{l} P(A \cup B) = \Pi(A) + \Pi(B) - \Pi(A \cap B) \quad (*\,*) \\ P(\emptyset) = 1 - \Pi(\Omega) = 0 \end{array}\right\} \textbf{si } A \cap B = \emptyset, \text{ ον ρετουπε Αξ. 3}$$

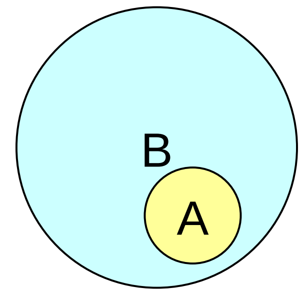
Axiomes des probabilités (3)

- Preuves :
  - $(*)$

    $B = A \cup (B - A)$

    δ϶ο $\Pi(B) = \Pi(A) + \Pi(B - A)$ (Aξ. 3)

    αϖεχ $\Pi(B - A) \geq 0$ (Aξ. 1)



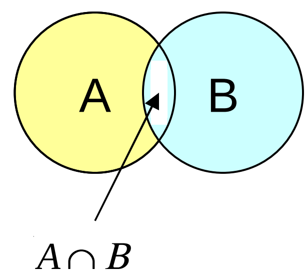
  - $(**)$

    $A \cup B = A \cup (B - A \cap B)$

    δ϶ο $\Pi(A \cup B) = \Pi(A) + \Pi(B - (A \cap B))$

    ορ $[B - (A \cap B)] \cup [A \cap B] = B$

    δ϶ο $\Pi(B - (A \cap B)) + \Pi(A \cap B) = \Pi(B)$



Probabilité conditionnelle

- La probabilité de l’événement A peut être influencée par l’information apportée par la connaissance de l’événement B :

$$P(A \mid B) = \frac{\Pi(A \cap B)}{\Pi(B)}$$

- A et B sont dits indépendants si l’observation d’un événement n’influence pas la probabilité d’observer l’autre :

$$P(A \mid B) = \Pi(A)$$
$$(\Pi(B \mid A) = \Pi(B))$$

Indépendance mutuelle

- Indépendance mutuelle :

$$P(A \cap B) = \Pi(A).\Pi(B)$$

- Exemples :
  - Obtenir un 6 et un 2 en lançant 2 dés : P=1/36
  - A : « voter Chirac », P(A) = 1/3 (pop. totale)

    B : « famille Dupont » contient 3 votants, et 1 personne a voté Chirac, P(A|B) = 1/3

    $\Rightarrow$ A et B indépendants

Variable aléatoire

- VA = fonction $X:\Omega \to \Upsilon$
  $\Rightarrow$ permet de quantifier les événements

- Exemple :
  exp. aléa. = lancer deux fois une pièce
  Ω = {(P,F);(F,P);(P,P);(F,F)}

  $\Rightarrow$ le nombre de « pile » est une VA définie comme suit :
  X((F,F)) = 0 ; X((P,F)) = X((F,P)) = 1 ; X((P,P)) = 2

Lois de probabilités

- V.A. discrètes ou continues
- Fonction de répartition, fonction de densité
- Paramètres des lois
- Exemples :
  - Lois discrètes : binomiale, poisson, …
  - Lois continues : laplace-gauss, gamma, exponentielle, …

Lois de probabilités

- Cas discret (Ω fini, dénombrable)
  - Fonction de probabilité :

$$P(X = \xi)$$

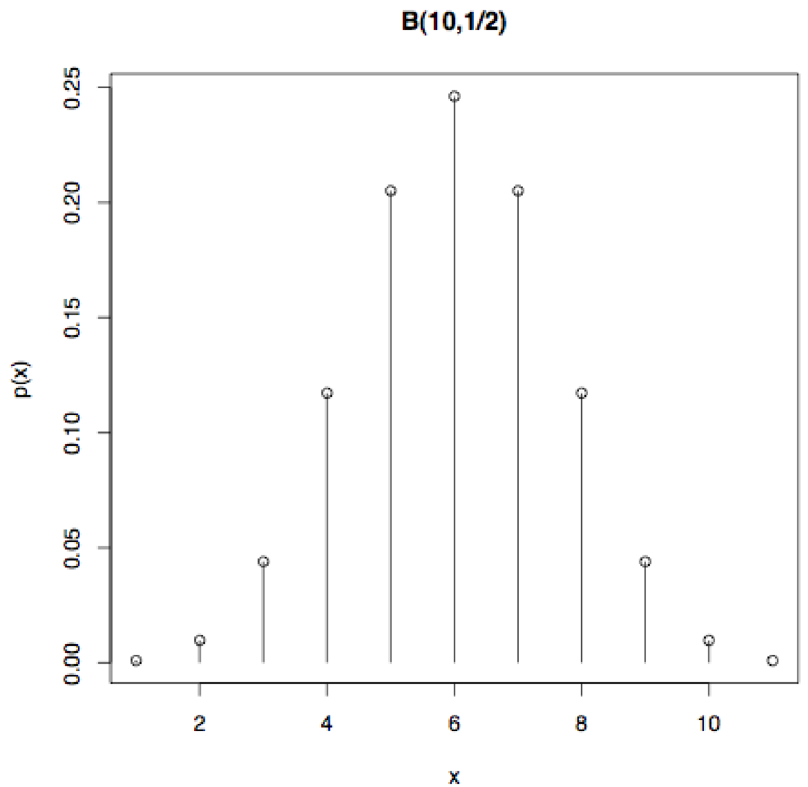
- Cas continu
  - Fonction de densité :

$$P(a < \Xi < \beta) = \int_{\alpha}^{\beta} \phi(\xi)\delta\xi$$

Exemples de lois discrètes

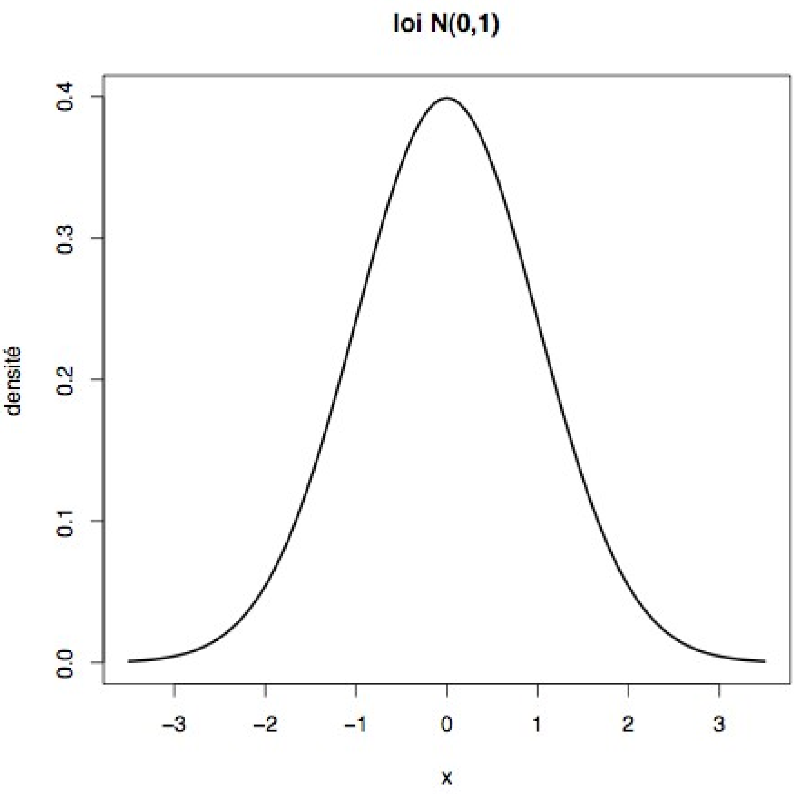
- **Loi binomiale**
  - Loi à 2 paramètres (n et p), symétrique
  - Modèle de l'urne ou du tirage Pile-Face
- Exemple :

$$\textbf{si } p = 0.5,\ \Pi(\Xi = 0) = \Pi(\Xi = 1) = 0.5$$



Exemples de lois continues

- **Loi de Laplace-Gauss** (loi « normale »)
  - Loi à 2 paramètres (m et σ), symétrique
  - « loi de l'erreur »
  - Très utilisée en statistique inférentielle
- Caractéristiques remarquables :
  - $P(-1<X<1) \approx 67\ \%$
  - $P(-2<X<2) \approx 95\ \%$
  - $P(-3<X<3) \approx 99\ \%$



Espérance mathématique (1)

- Valeur attendue, moyenne (en probabilité)
- Cas discret

$$E(X)=\sum_{\xi}\xi(\pi(\xi)$$

- Cas continu

$$E(X)=\int_{-\infty}^{+\infty}\xi\phi(\xi)\delta\xi$$

Espérance mathématique (2)

- Exemples :
  - Lancer de dés ({(P,F);(F,P);(P,P),(F,F)}

$$E(X) = 1/2 \cdot 1 + 1/4 \cdot 2 + 1/4 \cdot 0 = 1$$

  - Soit une VA dont la densité est donnée par

$$f(x) = \begin{cases} \dfrac{1}{2} & \sigma\iota\; 0 \leq \xi \leq 2 \\ 0 & \sigma\iota\nu o\nu \end{cases}$$

alors $$E(X) = \int_{-\infty}^{+\infty} \frac{1}{2}\, \delta\xi = 1$$

Entropie (1)

- Entropie = quantité moyenne d'information apportée par l'observation de la valeur prise par une VA
- Information contenue dans une réalisation :

$$H(X=\xi) = \ln \frac{1}{\pi(\xi)} = -\ln \pi(\xi)$$

Entropie (2)

- Exemple :
  - Pièce faussée, P(« pile ») = 0,9,
    X(« pile ») = 1 ; X(« face ») = 0

$$H(X=1) = -\lambda\nu\, 0,9 = 0,105$$

⇒ Observer « pile » apporte peu d’information puisque P(« pile ») = 0,9 : on est presque sûr d’observer « pile » lors d’un tirage

Entropie (3)

- Si l'on répète l'expérience (n grand) :

$$H(X) = -\sum_{\iota} \pi(\xi_\iota)\lambda\nu\, \pi(\xi_\iota)$$

$\Rightarrow$ Quantité d'information moyenne générée par les événements observés

- Exemple :
  - Lancer de pièces (infini) :

$$H(X) = -(0,5\,\lambda\nu\, 0,5 + 0,5\,\lambda\nu\, 0,5) = 0,693$$

Loi de probabilité conditionnelle

- (Cas discret)

$$P(x\,|\,Y) = \Pi(\Xi = \xi\,|\,\Psi) = \frac{\Pi(\{\Xi = \xi\} \cap \Psi)}{\Pi(\Psi)}$$

- Exemple :
  - On jette 2 dés et on note les n° sortis. Soit X la somme de ces deux nombres (e.g. X((3,2))=5).
    On considère également l'événement Y : les deux nombres sont pairs. On a donc

$$P(Y) = \frac{1}{4} \; \varepsilon\tau \; \Pi(\{\Xi = 6\} \cap \Psi) = \frac{1}{18}$$

  - Quelle est la probabilité d'observer X=6 sachant Y ?

$$P(X = 6\,|\,\Psi) = \frac{1/18}{1/4} = \frac{2}{9}$$

Loi des probabilités totales (1)

- Dans certains cas, la probabilité d'un événement ne peut être calculée directement, mais seulement à travers un ensemble de probabilités conditionnelles connues.

$$\textbf{Soit } B_1, B_2, \ldots, B_k \text{ une partition de } \Omega \; \left(B_i \cap B_j = \emptyset \;\; \forall i \neq j \text{ et } \bigcup_{i=1}^{k} B_i = \Omega\right),$$

$$P(A) = \sum_{i=1}^{k} P(A \mid B_i) P(B_i)$$

Loi des probabilités totales (2)

- Exemple :
  - 3 boules blanches + 1 boule rouge dans une urne. On choisit une boule au hasard, puis une seconde (tirage sans remise).
    Quelle est la probabilité que la seconde boule soit rouge ?

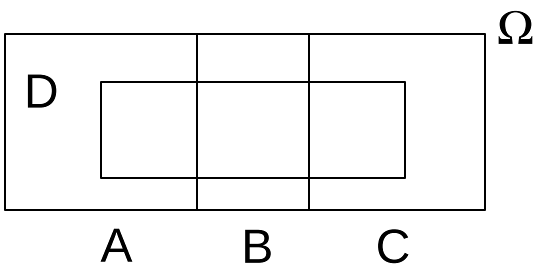
$$
\begin{aligned}
P(\text{"2-rouge"}) &= P(\text{"2-rouge"} \mid \text{"1-blanche"})\, p(\text{"1-blanche"}) \\
&\quad + P(\text{"2-rouge"} \mid \text{"1-rouge"})\, p(\text{"1-rouge"}) \\
&= 1/3 \cdot 3/4 + 0 \cdot 1/4 = 1/4
\end{aligned}
$$

Règle de Bayes (1)

- La probabilité peut être modifiée par l'observation ou la connaissance préalable de l'événement.
- Principe de Bayes :

$$P(B_i \mid A) = \frac{\Pi(A \mid B_i)\,\Pi(B_i)}{\sum_{\varphi} \Pi(A \mid B_{\varphi})\,\Pi(B_{\varphi})}$$

Règle de Bayes (2)



- On a une partition de Ω
- On connaît $P(B_i)$ et $P(A|B_i)$
- Problème : ayant observé l'événement A, quelle est la nouvelle probabilité associée à B_i

probabilité a priori

$$P(A \mid D) = \frac{\Pi(A \cap \Delta)}{\Pi(\Delta)} = \frac{\Pi(A)\Pi(\Delta \mid A)}{\Pi(\Delta)}$$

probabilité a posteriori

coefficient de réévaluation

Règle de Bayes (3)

- Exemple :
  - Diagnostic médical (présence/absence d'un virus ayant une prévalence de 0,1%) :

$$P(D)=0,001 \;\varepsilon\tau\; \Pi(\bar{\Delta})=0,999$$

	T^+	T^-
Δ	$0,95$	$0,05$
$\bar{\Delta}$	$0,02$	$0,98$

(probabilité *a priori*)

On note T^+ un test positif et T^- un test négatif.

Quelle est la probabilité que le patient soit réellement infecté par le virus si le test est positif ?

$$P(D\,|\,T^+)=\frac{\Pi(T^+\,|\,\Delta)\Pi(\Delta)}{\Pi(T^+\,|\,\Delta)\Pi(\Delta)+\Pi(T^+\,|\,\bar{\Delta})\Pi(\bar{\Delta})}=\frac{0,95\cdot 0,001}{0,95\cdot 0,001+0,02\cdot 0,999}=0,045$$

soit env. 5 chances sur 100 d'être réellement infecté !

Modélisation probabiliste (1)

- Modèle probabiliste d’un phénomène :
  - nécessite le plus souvent une **estimation** de paramètres (e.g. estimation de la distribution de probabilité a priori)
  - nécessite une bonne **adéquation** avec les données issues de l’expérience

Modélisation probabiliste (2)

- Méthodes classiques d'estimation
  - paramétrique vs. non-paramétrique
  - MCO, EMV
  - Simulation Monte Carlo

Conclusion

- Qu’est-ce qu’on retient de tout ça ?
  - On dispose d’outils permettant de modéliser le comportement non-déterministe d’un système
  - On peut prévoir son évolution, et quantifier la valeur prise par certains de ses paramètres caractéristiques (e.g. évolution de la réponse d’une population de neurones, choix d’un individu en fonction de ses réponses précédentes, évolution d’un marché financier, etc.)
  - Intérêt des modèles bayésiens : permettent d’incorporer des connaissances additionnelles, éventuellement fonction du temps (liées par exemple à une expérience ou à des observations préalables)

Plan

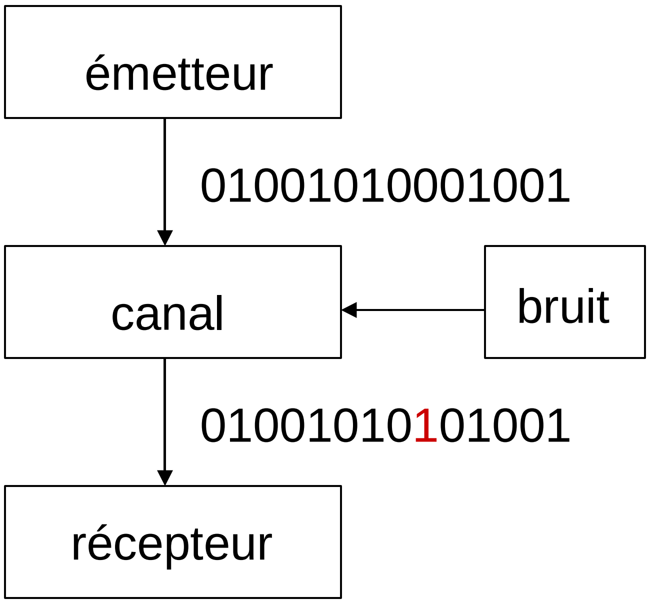
- Principes de la modélisation
- Rappels de probabilités
- Théorie de l'information
- Applications

Théorie de l'information

- Un schéma et une théorie de la communication (C. Shannon, 1949)
- Une formalisation et une quantification de l'information circulant entre un système (ouvert) et son environnement
- Exemples : télécommunications, langage naturel, transmission de l'information neuronale, etc.

Schéma de la communication

- **Idée de base** : Transmission d'un message codé (signal) entre un émetteur et un récepteur, éventuellement en présence de bruit sur le canal de transmission.
- **Signal** = phénomène physique discriminable
- **Code** = bijection, optimalité



Quantité d’information (1)

- Idée de base :
  - Information = sélection de possibilités (e.g. suite de choix binaires, 0/1)
  - Être informé = apprendre quelque chose de nouveau ⇔ diminution de l’incertitude
- **Quantité d’information** assimilée à la *probabilité de survenue d’un état par rapport à tous ses états possibles*.

Quantité d'information (2)

- Bit (binary digit) : unité de mesure élémentaire
- **Bit** = quantité d'information apportée par un événement dont la probabilité de survenue vaut 1/2.
- Exemples :
  - choix binaire, Q(I) = 1 bit
  - 4 possibilités, Q(I) = 2 bits
  - Assigner un nombre différent à chaque être humain : > 32 bits

Intérêts et limites

- Intérêt :
  - Réduction de la masse d'information (redondance, compression, correction)
  - Notion de codage optimal
  - Maximiser quantité d'information
- Limitation :
  - Suppression de la richesse sémantique du message (langage naturel ≠ code)

Information et entropie

- En physique, l'entropie d'un système isolé caractérise son degré de désordre.
- Entropie :
  - « mesure de l'ignorance »
  - mesure de la probabilité d'apparition d'une information sur un ensemble d'informations possibles
- 2nde loi de la thermodynamique :

basse entropie → haute entropie

Organisation et auto-organisation

- Auto-organisation
  - création d'ordre par le bruit (von Foerster)
  - complexification par le bruit (Atlan)
- Comprendre la nature de la complexité d'un système, car son adaptation dépend de sa capacité à distinguer dans son environnement l'aléatoire du régulier

Organisation d'un système

- Organisation = système structuré par un certain ordre
  - Tout ordre possède un caractère redondant et une fiabilité qui lui assure sa stabilité dans le temps.
  - Redondance d'un système ≠ 0, sinon le hasard régnerait et aucune structure ne saurait se maintenir en équilibre
  - Régularités = tendent à stabiliser les événements lorsqu'elles se reproduisent (dynamique temporelle)

Auto-organisation d’un système (1)

- Auto-organisation
  - lorsque les facteurs de bruit de l’environnement ou les fluctuations internes permettent d’augmenter Q(I)
  - $\Leftrightarrow$ augmenter sa capacité de produire plus d’états intérieurs possibles (mesurables en nombre de bits)

Auto-organisation d'un système (2)

- Réponse d'un système à des perturbations aléatoires
  - Modification de son organisation interne pour produire des réponses originales
  - Complexification progressive de sa structure

$\Rightarrow$ Apparition de propriétés nouvelles

Complexité d’un système (1)

- Idée : **mesurer** ce processus de complexification permettant d’aboutir à la formation de systèmes complexes
- Types de complexité (brute, effective, algorithmique)

  $\Rightarrow$ fonction de la longueur du message permettant de décrire les états d’un système

Complexité d'un système (2)

- Complexité algorithmique $\Rightarrow$ **contenu d'information algorithmique** (Kolmogorov, Chaitin & Solomonoff)
  - = longueur d'un programme informatique (chaîne de bits) en fonction d'un matériel et d'un logiciel donné
  - constitue une mesure de l'aléatoire
  - propriété : *incalculabilité* (aucune procédure a priori ne permet de prédire qu'un algorithme ne permettrait pas de compresser davantage la chaîne de bits considérée)

Complexité d'un système (3)

- Selon la théorie de l'information, la complexité d'un système se définit par
  - le nombre des états qu'il peut prendre
  - $\Leftrightarrow$ la quantité d'information qu'il contient

- C.I.A. : entre ordre et désordre
  - proche de 0 $\Rightarrow$ très grande régularité dans la chaîne de données (e.g. 111111…) ou désordre parfait
  - complexité effective = C.I.A. ni trop haut, ni trop bas
  - $\Leftrightarrow$ système ni trop ordonné, ni trop désordonné

Perspectives « chaotiques » (1)

- Tendance actuelle : associer la notion de complexité à celle de **chaos**
- Exemples :
  - *Neurobiologie*

    1) transmission quantique des signaux entre un neurone et son voisinage

    $\Rightarrow$ communication non garantie mais d'une richesse très grande

    2) interactions modulable par l'apprentissage

    $\Rightarrow$ forte valeur adaptative de cette modulation probabiliste pour le comportement

Perspectives « chaotiques » (2)

- *Économie*

  analyse des krachs financiers : remise en question de l'idée que ce sont des phénomènes aléatoires imprévisibles

  ⇒ krach = ordre total

  tout le monde vend en même temps ses actions ; cette phase est précédée d'une « bulle spéculative » dans laquelle les agents interagissent dans une boucle de rétroaction positive (imitation) ; celle-ci se renforce jusqu'à ce que le marché devienne instable et soit « corrigé » par un krach

  ⇒ émergence de l'ordre à partir du désordre

Perspectives « chaotiques » (3)

- Éthologie

  évolution des colonies de fourmis

  malgré le désordre apparent (certaines fourmis construisent des structures que d'autres détruisent aussitôt), certaines régularités sont observables dans l'organisation globale de leurs comportements

  ⇒ au niveau collectif, l'ordre émerge de processus auto-organisés (apprentissage et compétition, communication)

Perspectives « chaotiques » (4)

- Au niveau individuel, la modélisation est plus problématique
- Exemple : modélisation du trafic routier
  - Modèle statistique tenant compte des variations observées sur 10 ans : ok
  - Modèle de chaos déterministe (désordre auto-ordonné)

    $\Rightarrow$ il manque un paramètre essentiel : le comportement humain, or c'est bien l'interaction conducteur-voiture-environnement qui permet d'expliquer les phénomènes d'embouteillage

Plan

- Principes de la modélisation
- Rappels de probabilités
- Théorie de l’information
- Applications

1. Intégration multisensorielle (1)

- Contexte théorique
  - Traitement multimodal de l'information sensorielle
  - Intégration adaptative des informations issues de 2 canaux de traitement, e.g. vision et audition, vision et modalité haptique (proprioception + mouvements actifs exploratoires de la main)

    (≠ simple procédure de vote)

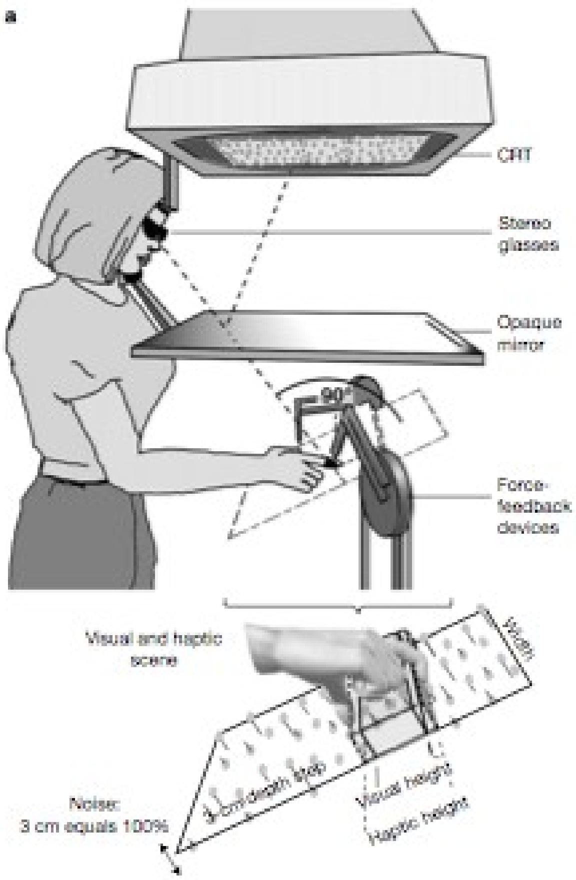
1. Intégration multisensorielle (2)

- Référence

MO Ernst and MS Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415:429-433, 2002.

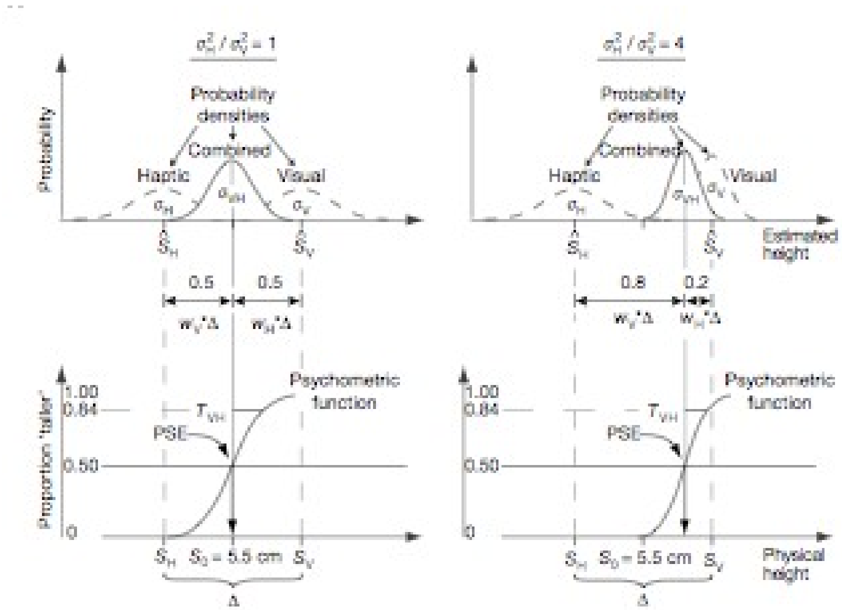
When a person looks at an object while exploring it with their hand, vision and touch both provide information for estimating the properties of the object. Vision frequently dominates the integrated visual-haptic percept, for example when judging size, shape or position, but in some circumstances the percept is clearly affected by haptics. Here we propose that a general principle, which minimizes variance in the final estimate, determines the degree to which vision or haptics dominates. This principle is realized by using maximum-likelihood estimation to combine the inputs. To investigate cue combination quantitatively, we first measured the variances associated with visual and haptic estimation of height. We then used these measurements to construct a maximum-likelihood integrator. This model behaved very similarly to humans in a visual-haptic task. Thus, the nervous system seems to combine visual and haptic information in a fashion that is similar to a maximum-likelihood integrator. Visual dominance occurs when the variance associated with visual estimation is lower than that associated with haptic estimation.

1. Intégration multisensorielle (3)



1. Intégration multisensorielle (4)

Principe d'intégration visuo-haptique par estimation MV

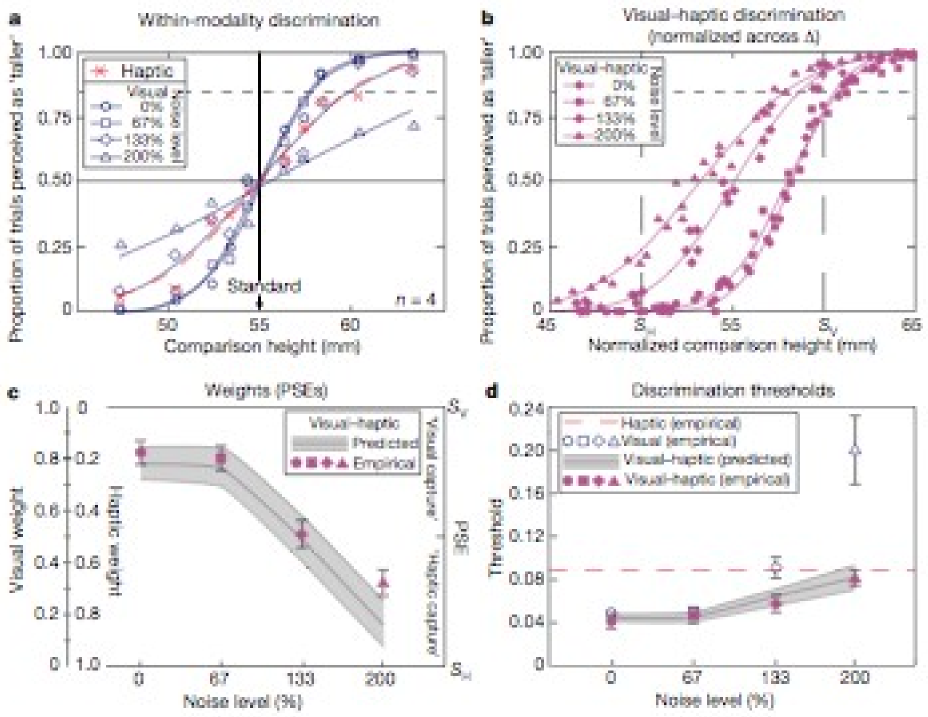


1. Intégration multisensorielle (5)

- En résumé :
  - Prendre en compte les deux estimations unimodales
  - Pondérer ces estimations par leur précision relative ($1/\sigma^2$)
  - Combiner ces estimations

1. Intégration multisensorielle (6)

Comparaison valeurs prédites / valeurs observées



1. Intégration multisensorielle (7)

- Conclusion :
  - Bonne adéquation des données au modèle

    ⇒ Adéquation du modèle
  - Filtrage adaptatif de l'information afférente
    - ⇒ Utilisation optimale de la spécificité/sélectivité de chacune des modalités sensorielles lorsqu'elles sont recrutées simultanément

2. Contrôle moteur (1)

- Contexte théorique
  - Utilisation optimale des informations sensorielles pour le contrôle de l'action (e.g. calibration d'un geste d'atteinte ou de saisie)
  - Boucles sensorimotrices
  - Modèle de contrôle moteur :
    - Modèle direct ('forward model')
    - Modèle inverse ('inverse model')

    (≠ servo-contrôle)

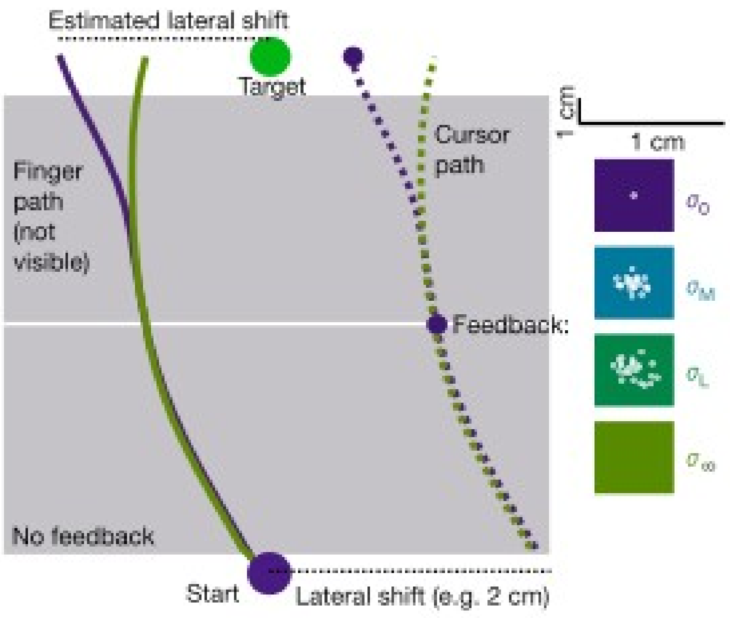
2. Contrôle moteur (2)

- Référence

Körding, S Ku, and DM Wolpert. Bayesian integration in force estimation. *Journal of Neurophysiology*, 92:3161-3165, 2004.

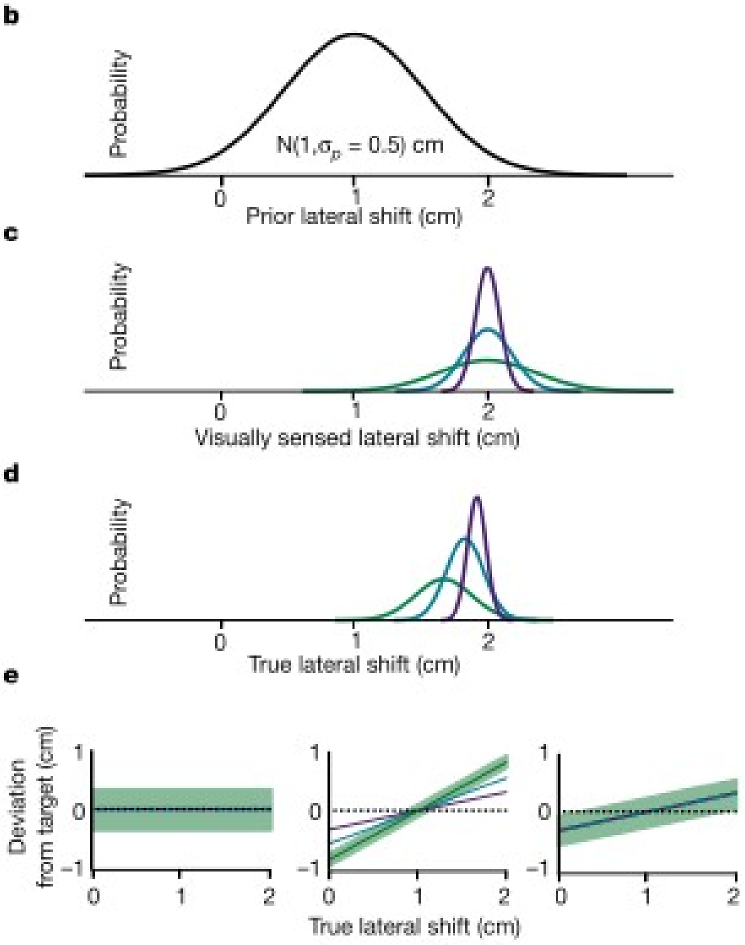
When we interact with objects in the world, the forces we exert are finely tuned to the dynamics of the situation. As our sensors do not provide perfect knowledge about the environment, a key problem is how to estimate the appropriate forces. Two sources of information can be used to generate such an estimate: sensory inputs about the object and knowledge about previously experienced objects, termed prior information. Bayesian integration defines the way in which these two sources of information should be combined to produce an optimal estimate. To investigate whether subjects use such a strategy in force estimation, we designed a novel sensorimotor estimation task. We controlled the distribution of forces experienced over the course of an experiment thereby defining the prior. We show that subjects integrate sensory information with their prior experience to generate an estimate. Moreover, subjects could learn different prior distributions. These results suggest that the CNS uses Bayesian models when estimating force requirements.

2. Contrôle moteur (3)



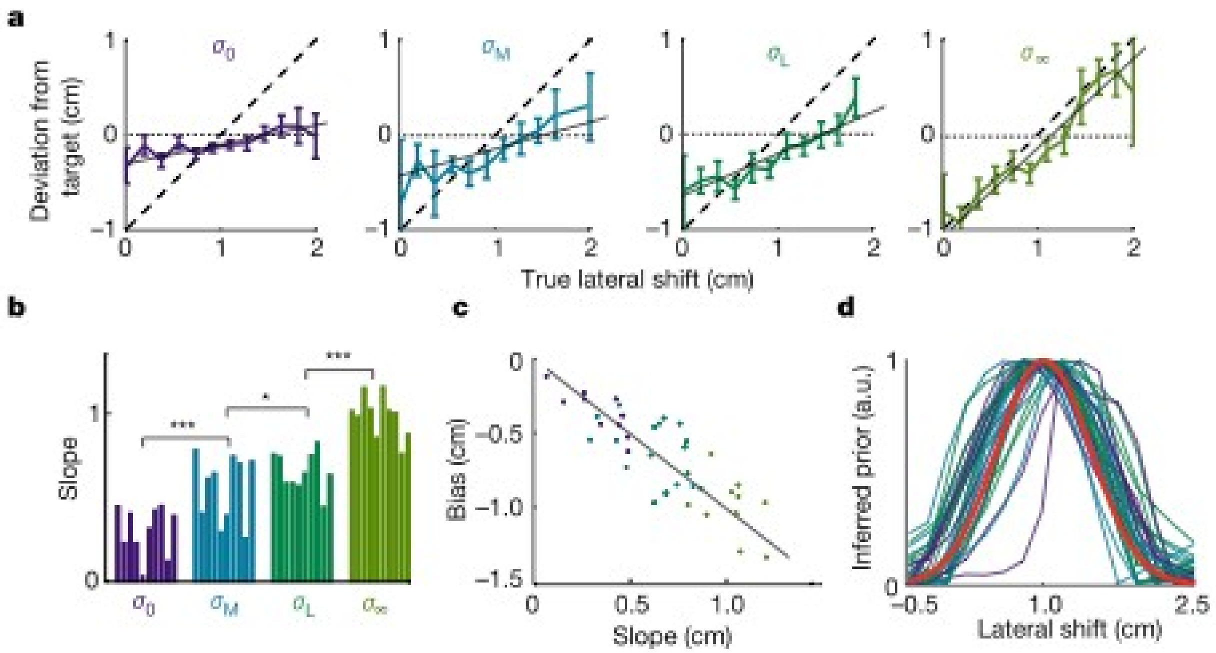
2. Contrôle moteur (4)

Combinaison bayésienne des informations efférentes et afférentes



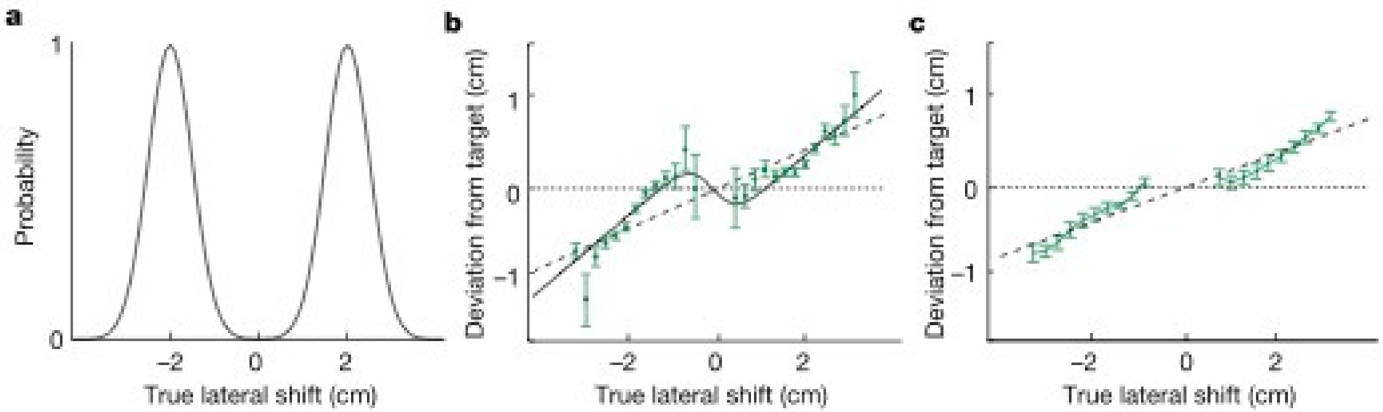
2. Contrôle moteur (5)

Résultats en assumant une distribution gaussienne



2. Contrôle moteur (6)

Résultats en assumant un mélange de distributions



2. Contrôle moteur (7)

- Conclusion :
  - utilisation progressivement optimale des informations sensorielles afférentes
  - apprentissage sensorimoteur modélisé comme un processus bayésien

Conclusion générale

- Intérêt des modèles probabilistes
- Intérêt d'une démarche alliant la modélisation et l'expérimentation