

Introduction à la statistique bayésienne

Eric PARENT

Enseignant-chercheur AgroParisTech

Ingénieur en chef des Ponts, Eaux et Forêts

Equipe MOdélisation, Risque, Statistique, Environnement de l'UMR MIA 518
INRA/AgroParisTech (Math. Info. App. 518)

Paris, 26 janvier 2012,

Séminaire JEM SFES

Pavillon Dauphine

Place du Maréchal de Lattre de Tassigny

75016 Paris



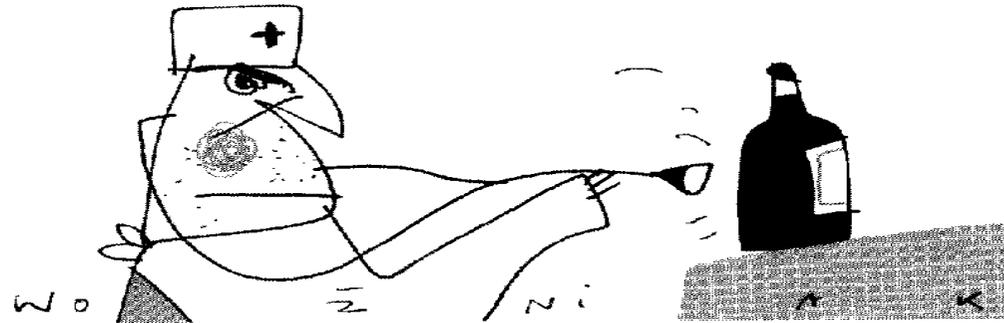
Pinard de vivre

ON va perdre notre triple A, mais voilà une bonne nouvelle : des scientifiques viennent de prouver que le vin, notre breuvage national, allonge l'espérance de vie. De quoi laver l'affront du dernier rapport de l'Institut national du cancer, qui assurait tout de go qu'un verre de vin par jour fichait en l'air la santé. Vous allez dire que cette nouvelle étude en faveur du pinard c'est, comme chaque fois, grâce à des publications orientées : on nous vante le truc « bon pour la santé », le chocolat, le café, le vin, le foie gras. Mais cette fois-ci c'est du sérieux.

Des chercheurs italiens en épidémiologie ont passé au tamis 97 publications scientifiques sur les effets de l'alcool sur l'organisme, pour n'en garder que 18 qui tenaient la route, celles irréprochables sur le plan scientifique. Une « meta-analyse », dans le jargon. Et sur le sujet c'est une première mondiale. Ce travail, qui vient d'être publié dans une revue scientifique de référence (« European Journal of Epidemiology », 2011), est formel : jusqu'à une bouteille par jour, les

amateurs de vin sont mieux protégés des affections cardio-vasculaires, en particulier des crises cardiaques et des accidents vasculaires cérébraux, que ceux qui ne lèvent jamais le coude. Et, si on soustrait les cirrhoses, certains cancers et accidents de la route qui peuvent frapper les disciples de Bacchus, on découvre que, entre deux et trois verres par jour, siroter du vin ferait même grimper l'espérance de vie.

Pour la petite histoire, ce sont les brasseurs italiens qui, sans le vouloir, ont mis du baume au cœur des viticulteurs. C'est en effet leur association professionnelle qui a décidé de lancer une étude pour vérifier les effets de l'alcool sur le cœur et les artères. Avec le secret espoir de communiquer sur les bienfaits de la bière. Sauf qu'en fin de compte c'est le vin qui a rafé la mise. Une divine surprise pour le premier producteur mondial de pinard que nous sommes, avec 44 millions d'hectolitres par an. Même si la crise ne nous laisse plus un radis, il restera toujours une feuille de vigne pour s'habiller...



Le Canard Enchaîné
Mercredi 4 Janvier 2012

Bayes, une formule *bien* connue...

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

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(A|B) = \frac{P(B|A)P(A)}{\sum_i P(B|A_i)P(A_i)}$$

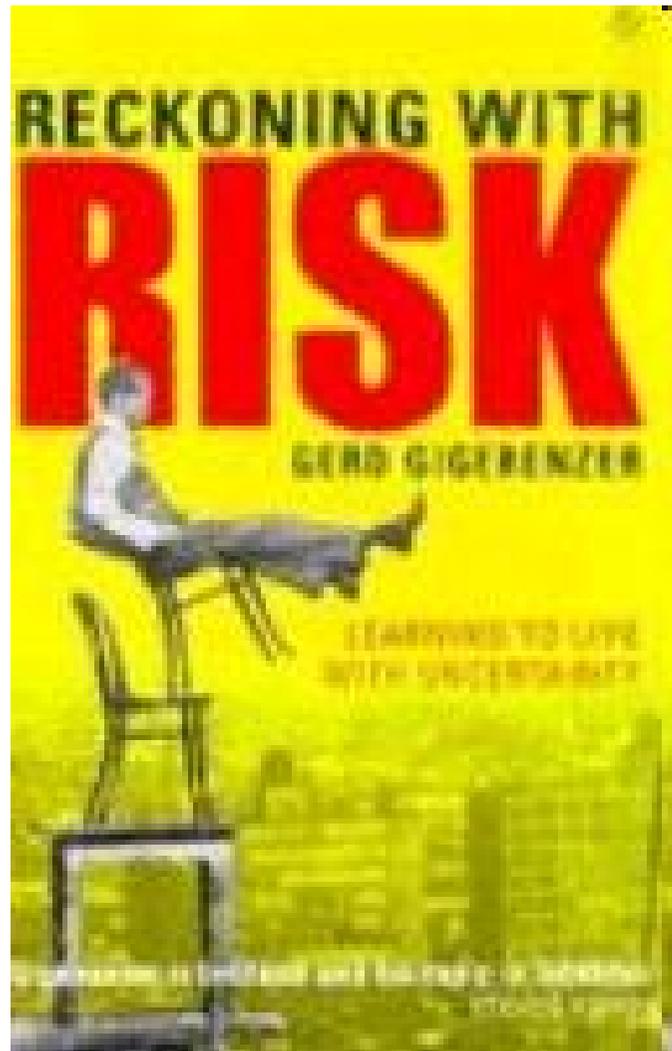
A : cause, état, maladie

B: conséquence, manifestation, symptôme

Bayes, une formule *difficile* ...

The probability that a women of age 40 has breast cancer is about 1%. If she has breast cancer, the probability that she tests positive on a screening mammogram is 90%. If she does not have breast cancer, the probability that she nevertheless tests positive is 9%.

What are the chances that a women who tests positive actually has breast cancer?



Gerd Gigerenzer
Reckoning with risk (2002):
learning to live with uncertainty
Penguin books



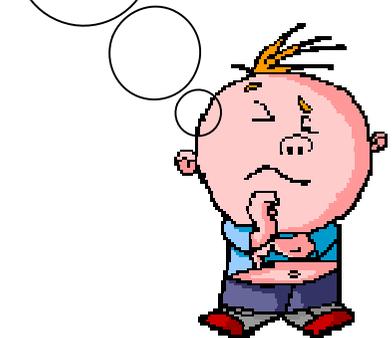
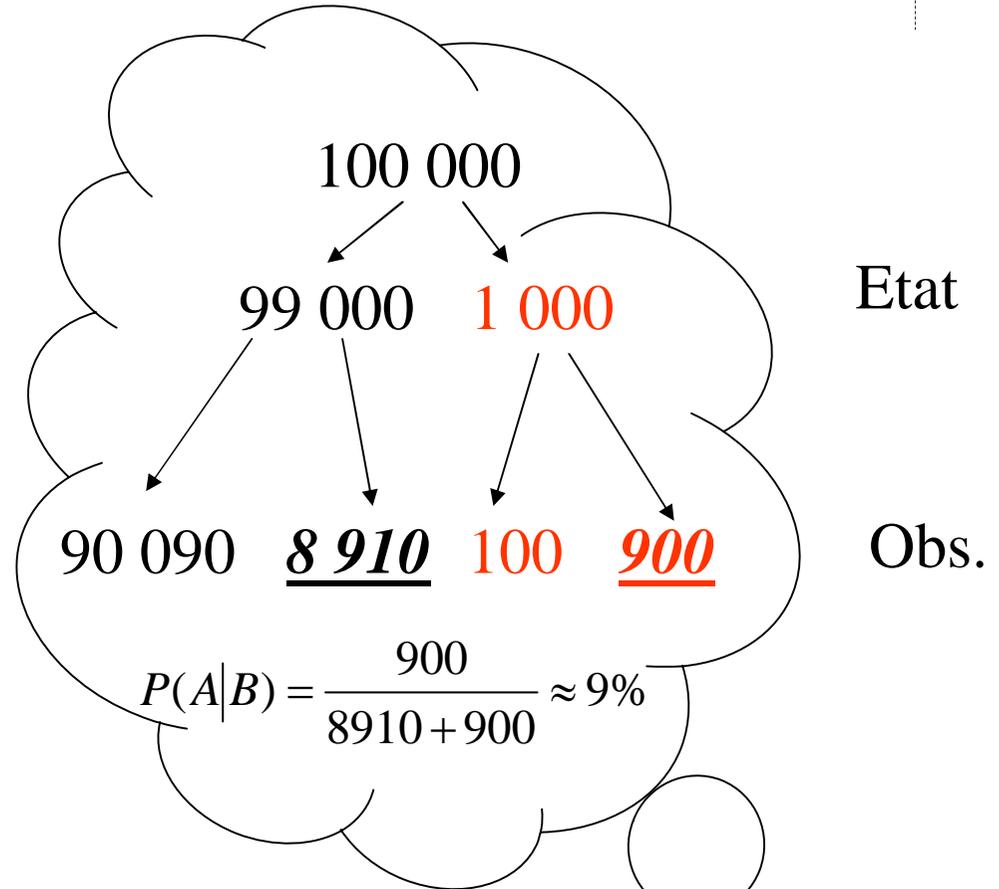
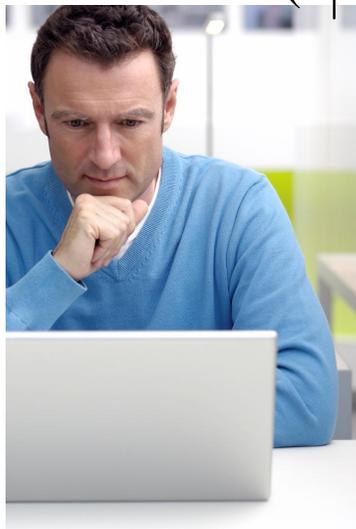
<http://www.informedmedicaldecisions.org/video-popup.php?video=1>

Changer le point de vue



$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(A|B) = \frac{P(B|A)P(B)}{P(A)}$$



Autre explication...

Changer l'énoncé : Si la dame est venue consulter (événement C) ..., et doute raisonnable

$$P(A|C) = 10\% \text{ au lieu de } P(A) = 1\%$$

$$P(A|B, C) = 52\%$$

$$P(A|C) = 50\% \text{ au lieu de } P(A) = 1\%$$

$$P(A|B, C) = 91\%$$

Autre application...

LET'S SEE...
WE DID IT 5 TIMES THIS PAST MONTH.
YOU COUNTED?!

SO THAT MEANS I HAVE A 16.6% CHANCE OF HAVING SEX TONIGHT.
YOU'RE FORGETTING ABOUT BAYESIAN INFERENCE.

LET "A" BE THE EVENT THAT YOU GET LAID. LET "B" BE THE EVENT THAT YOU COMPLIMENT HER EYES.
OUT OF THE FIVE TIMES YOU GOT LAID, HOW MANY TIMES DID YOU COMPLIMENT HER EYES?

HMM...
FROM THOSE 5 TIMES, I COMPLIMENTED HER EYES 3 TIMES.
BUT I COMPLIMENT HER EYES ONCE A WEEK.

SO...
 $P(B|A) = 3/5$
 $P(A) = 5/30$
 $P(B) = 4/30$
THUS,
$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

$$= 3/4$$

SO, GIVEN THAT YOU COMPLIMENT HER EYES, YOU HAVE A 75% CHANCE OF GETTING SEX TONIGHT.

SWEET!
BAYESIAN INFERENCE IS GETTING ME SEX TONIGHT!

PROBABLY.

spikedmath.com
© 2010

Le « Bayes », qu'est ce que c'est?



- Le détective inférentiel : ...*Bayesian statistics is difficult in the sense that thinking is difficult*, Berry (1997)
- La probabilité comme seule mesure de l'incertitude, une épreuve pédagogique? *Bayesian statistics is not a branch of statistics, it is a way of looking at the whole of statistics*, Lindley (1971)
- La statistique = apprentissage stochastique, une version intuitive?



De la stat *bayésienne* aux axiomes des probabilités

Définition : Probabilité Subjective

On appelle $\Pr(A)$ le prix que **Vous** êtes disposé à payer ou à vendre un pari qui rapporte : 1 euro si l'événement A se produit, 0 sinon.

Propriétés (de **cohérence**)

Pour éviter d'être perdant systématique, l'opérateur \Pr **doit** obéir aux règles fondamentales :

$$\Pr(A) \geq 0$$

$$\Pr(E) = 1$$

$$\Pr(A) + \Pr(B) = \Pr(A \cup B) \text{ si } A \text{ et } B \text{ disjoints}$$

Remarque : Il faut imaginer un marché avec des opérateurs pour lesquels prix de vente = prix d'achat. ...sinon voir construction dite des probabilités imprécises (Walley, 1990; Dempster, 1968; Shafer, 1976)



Rationalité bayésienne = calcul des probabilités

Conclusions :

la Probabilité Subjective $Pr(A)$ obéit aux règles usuelles du calcul des probabilités.

Remarques :

Dans le contexte d'une algèbre d'événements discrets, la seule condition de cohérence est de refuser d'être perdant systématique. (Homme rationnel?)

Dans le cas de cardinal d'événements infini, dénombrable ou non, la définition en terme de pari doit être adaptée (cf. convergence des séries ou de l'intégrale en théorie de la mesure)

Des individus différents, avec une même information a priori, peuvent valuer des paris différemment. La probabilité bayésienne est subjective et conditionnelle (à un état d'information)!

On montre que : Le théorème de Bayes implique que des individus avec des priors différents (subjectifs mais non *définitifs*) arriveront à s'accorder (asymptotiquement) sur une probabilité a posteriori commune (objective?)
En utilisant toute l'information disponible, les règles de Bayes (et leur limites) fournissent toutes les décisions avec le taux d'erreur le plus petit possible.



Rationalité bayésienne : les (re)fondateurs

Suivant Bayes/Price (1763) puis Laplace (1774) Mémoire sur la Probabilité des Causes par les Événements í .

Bruno DeFinetti. (1937). La Prévision: ses lois logiques, ses sources subjectives. Annales de l'Institut Henri Poincaré.

Leonard Jimmie Savage (1954). Foundations of Statistics. John Wiley and Sons.

Dennis V. Lindley. (1965). Introduction to Probability and Statistics from a Bayesian Viewpoint, 2 volumes, Cambridge.

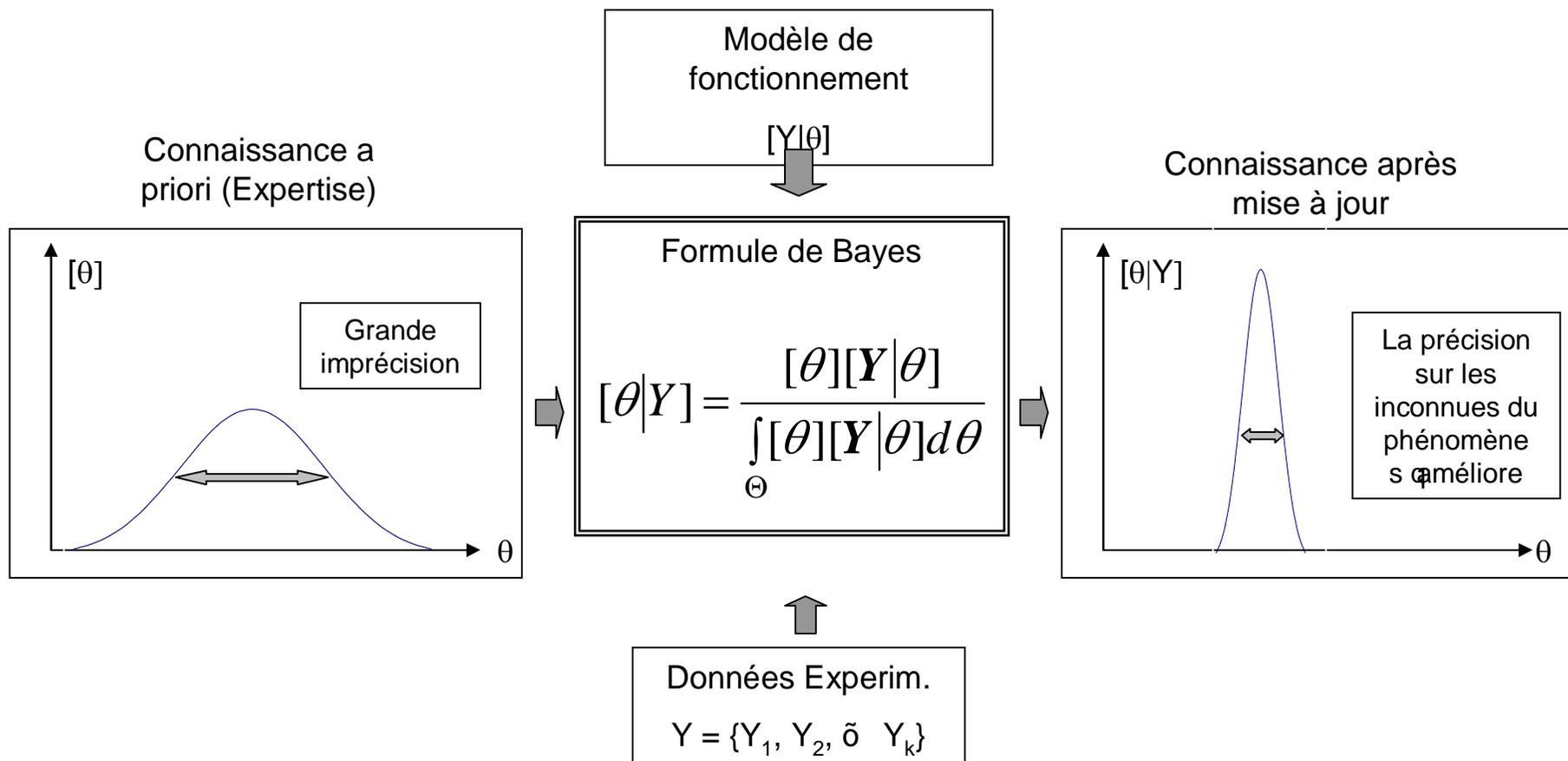
John W. Pratt, Howard Raiffa and Robert Schlaifer (1964). The Foundations of Decision Under Uncertainty: An Elementary Exposition. JASA, 59(306).

Jaynes, Edwin T. (2003). Probability theory: the logic of science. Cambridge University Press.

Dennis V. Lindley. (2006). Understanding uncertainty. John Wiley and Sons.

Joseph B. Kadane (2011) Principles of Uncertainty (Chapman & Hall/CRC Texts in Statistical Science)

Mise à jour bayésienne de la connaissance



Mise à jour à partir de quoi?

Question de prior

Prise en compte du savoir expert,
exigeant un contact plus étroit avec
les scientifiques du domaine

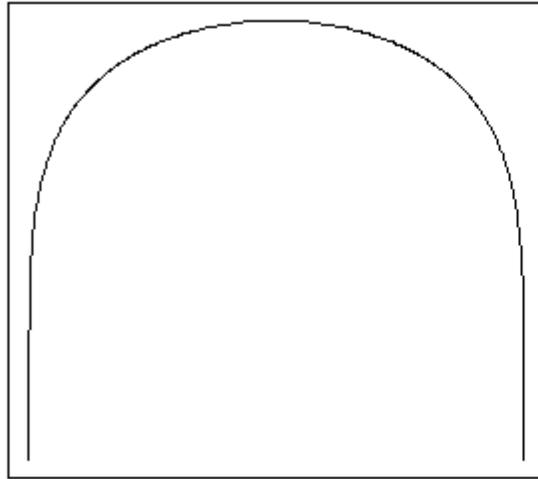
öUncertainty is a personal matter; it is not the uncertainty but *your* uncertainty.ö

Dennis Lindley, Understanding Uncertainty (2006)

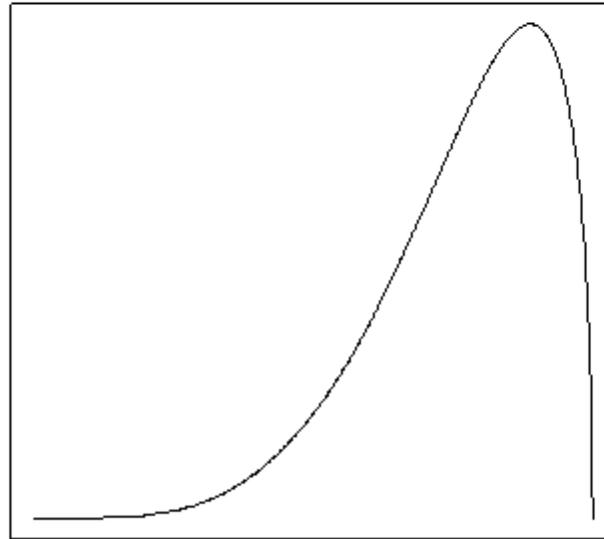
Question de prior...

- Une buveuse de thé au lait, affirme qu'elle est capable de dire si c'est le lait ou le thé qui ont été versé en premier dans la tasse. Sur 6 essais, elle détermine correctement le premier ingrédient
- Un amateur de musique annonce qu'il est capable de distinguer Haydn de Mozart. Sur 6 morceaux écoutés, il trouve le bon auteur.
- Une nuit fort arrosée, votre collègue se prétendait devin. Lançant 6 fois une pièce en l'air, il prédit correctement les occurrences de pile ou face

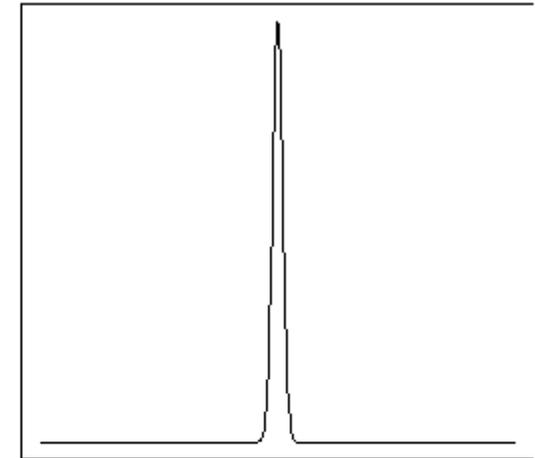
?



Case 1 :
Lady drinking tea



Case 2 :
Music lover

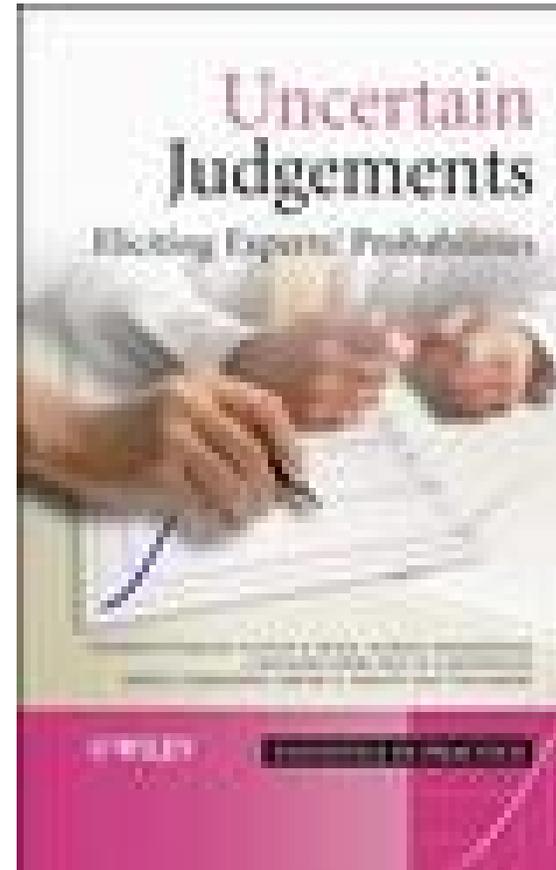


Case 3 :
Flipping a coin

Encodage probabiliste d'expertise ou *élicitation*

- Prior informatif
- Prior vague
- Convenance mathématique
- Principe d'entropie
- ...

- Refus par principe en stat classique.



La vision prédictive

Le langage naturel du scientifique

PROBABILITY DOES NOT EXIST

The abandonment of superstitious beliefs about the existence of the Phlogiston, the Cosmic Ether, Absolute Space and Time, . . . or Fairies and Witches was an essential step along the road to scientific thinking. Probability, too, if regarded as something endowed with some kind of objective existence, is no less a misleading misconception, an illusory attempt to exteriorize or materialize our true probabilistic beliefs.

(De Finetti, 1974)

La vision prédictive

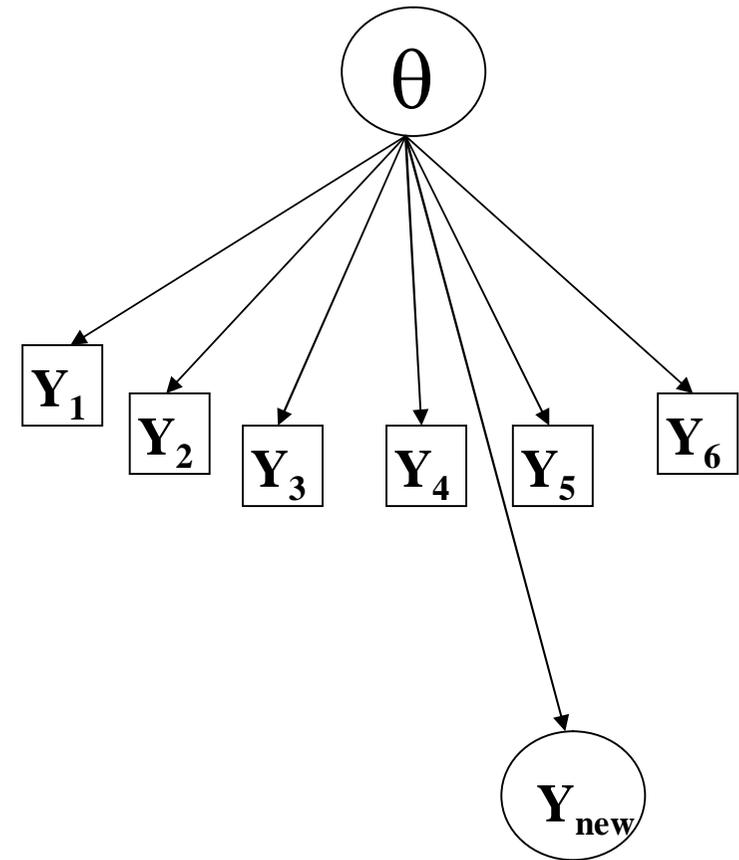
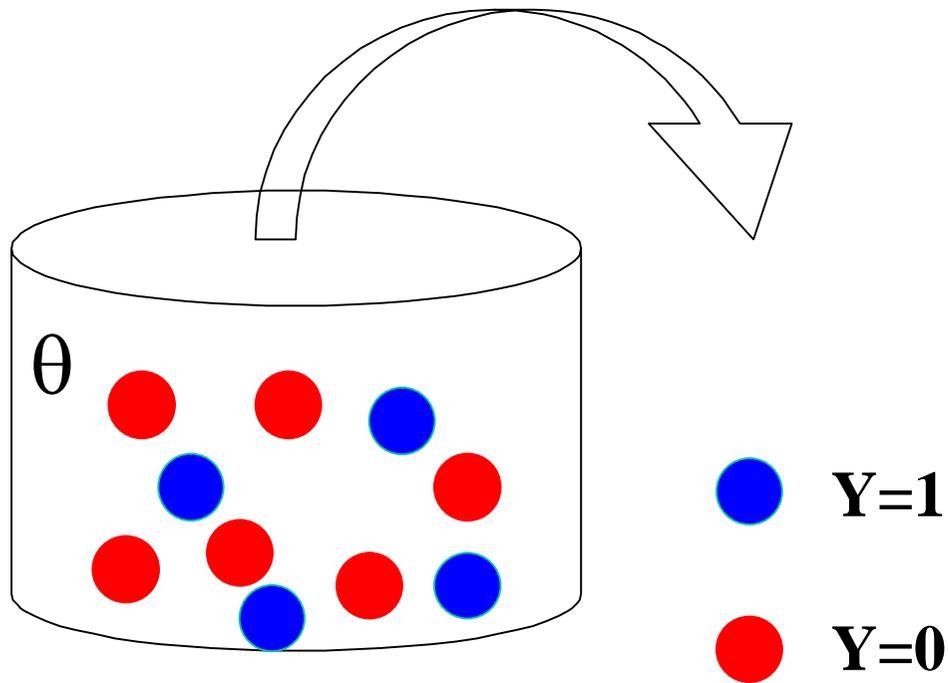
sur observations binaires $Y=1$ (succès) ou $Y=0$ (échecs):

Les 3 séquences sont elles différentes?

- 1 0 1 1 1 1 1 ? Jeu de pile ou face avec pièce équilibrée
- 1 0 1 1 1 1 1 ? Naissance attendue dans une famille avec six garçons et une fille
- 1 0 1 1 1 1 1 ? Contrôle de la qualité dans une usine : une seule pièce correcte sur sept échantillonnées

La vision prédictive

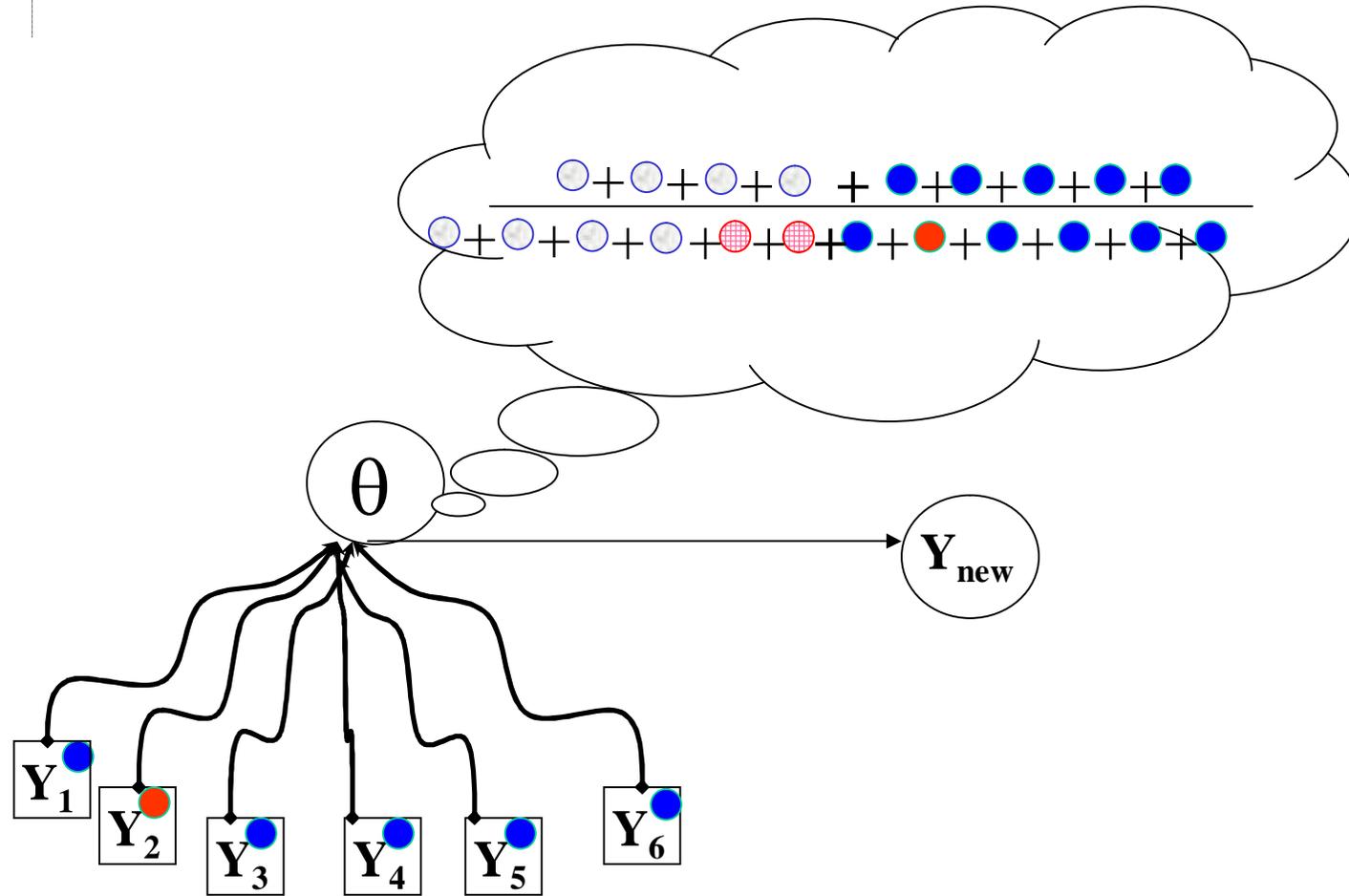
Même modèle d'orne



Données

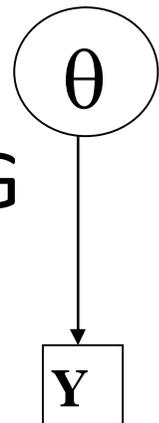


La vision re-constructive



La vision prédictive

- observations binaires: 1 0 1 1 1 1 1 ?
Prédictive=pari probabiliste naturel
- L'expertise, codée par le prior, est équivalente à un jeu de données virtuelles
- La probabilité est conditionnelle
- Indépendance ou indépendance conditionnelle
- Faire apparaître le conditionnement => DAG



Modélisation hiérarchique et DAG

L' utilisation de graphes acycliques orientés permettant une meilleure interface avec le spécialiste de terrain

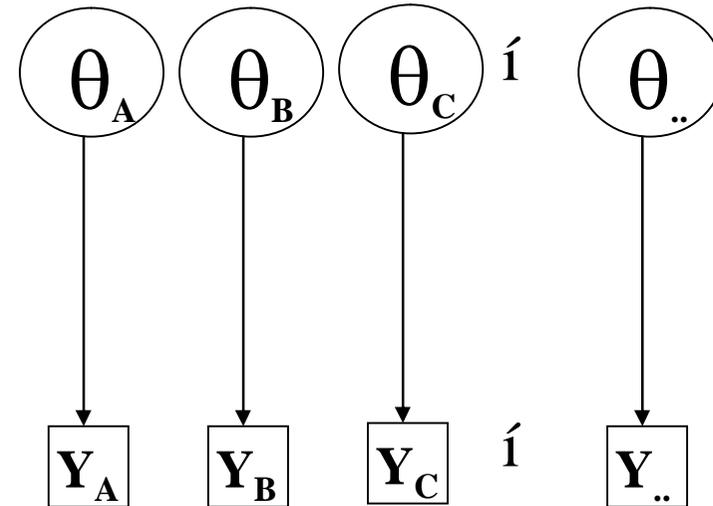
Décès dans les hopitaux (Thanks to WinBugs)

- Quelle est la probabilité de mourir durant une intervention chirurgicale dans un hôpital en UK?
- Le nombre y_i de décès suit une loi binomiale de paramètre θ_i et N_i essais (i indice de l'hôpital)
- Le prior sur θ_i est pris dans une loi beta

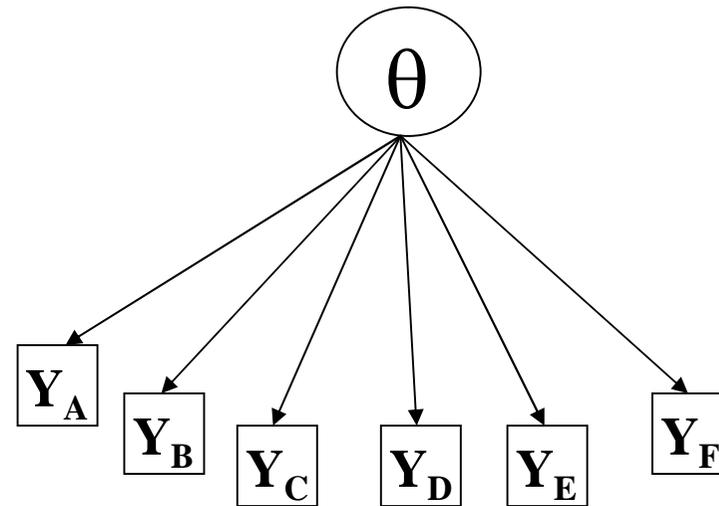
Hopital	Opérations	Décès
A	47	0
B	148	18
C	119	8
D	810	46
E	211	8
F	196	13
G	148	9
H	215	31
I	207	14
J	97	8
K	256	29
L	360	24

Décès dans les hopitaux UK

- Répéter l'analyse?

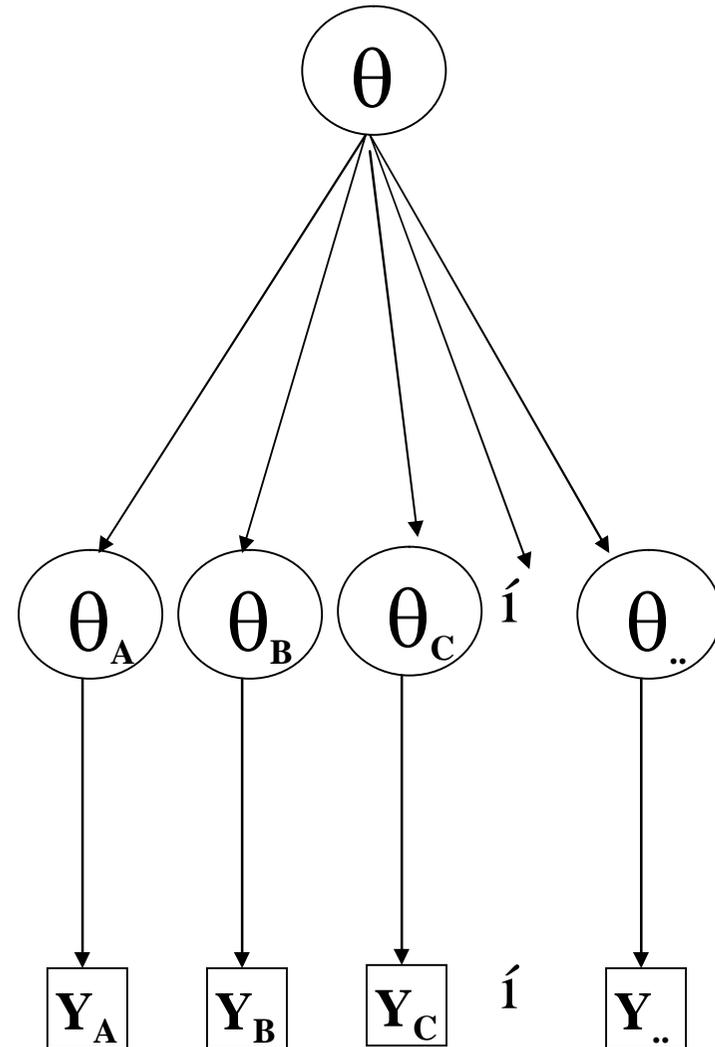


- Mélanger les données?

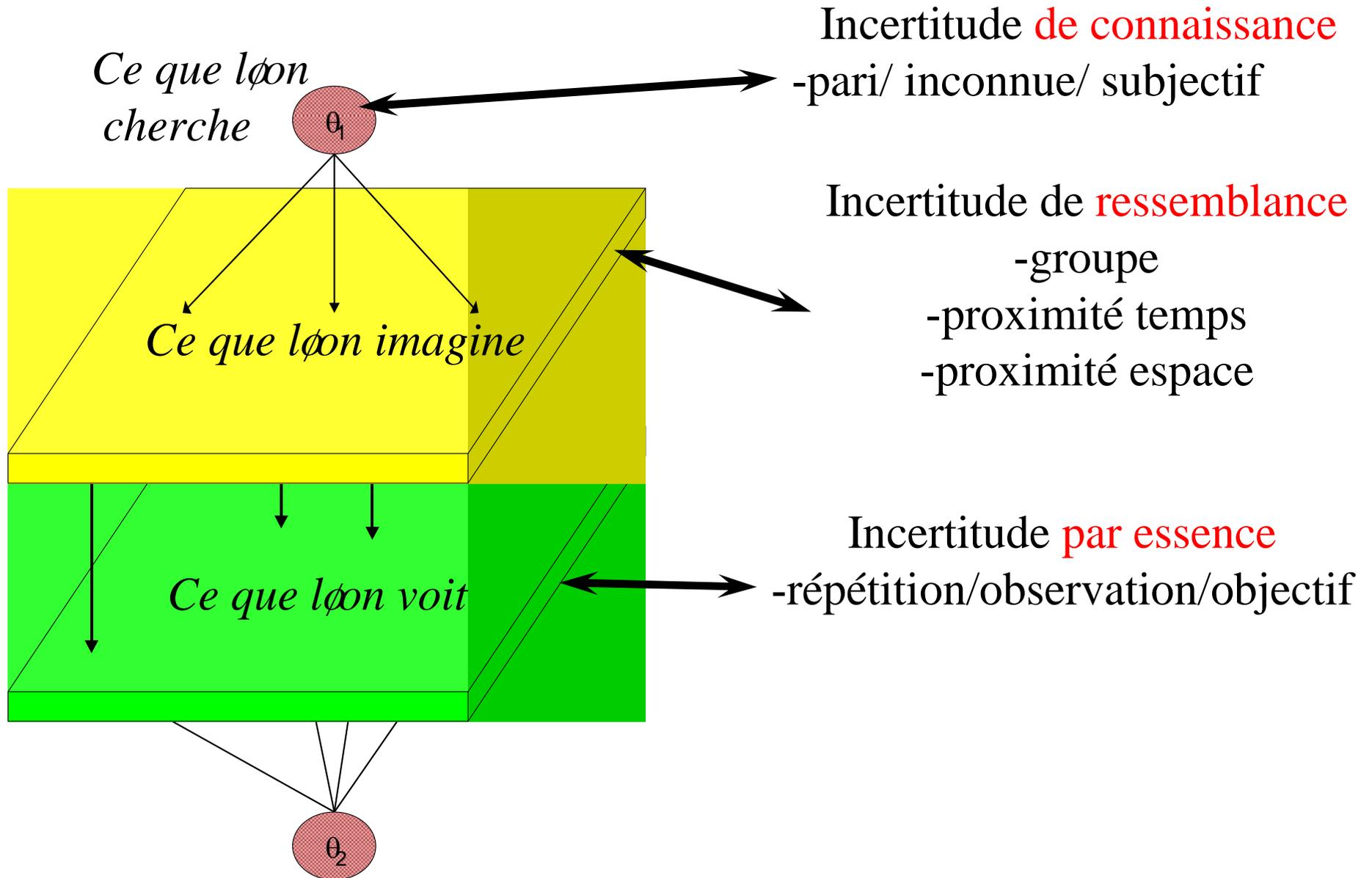


Décès dans les hopitaux UK Modèle structuré

- Exprimer la + ou –grande ressemblance entre établissements
- Autorise le transfert d'information entre établissements
- Permet de tempérer les estimations optimistes ou pessimistes quand n petit

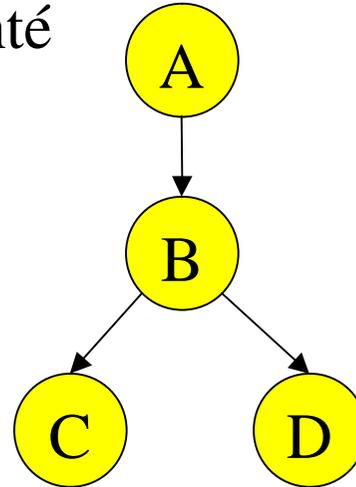


Encodage probabiliste de l'incertitude



Modèle graphique ou réseau bayésien

1. Un graphe acyclique orienté



2. Les lois conditionnelles pour chaque noeud

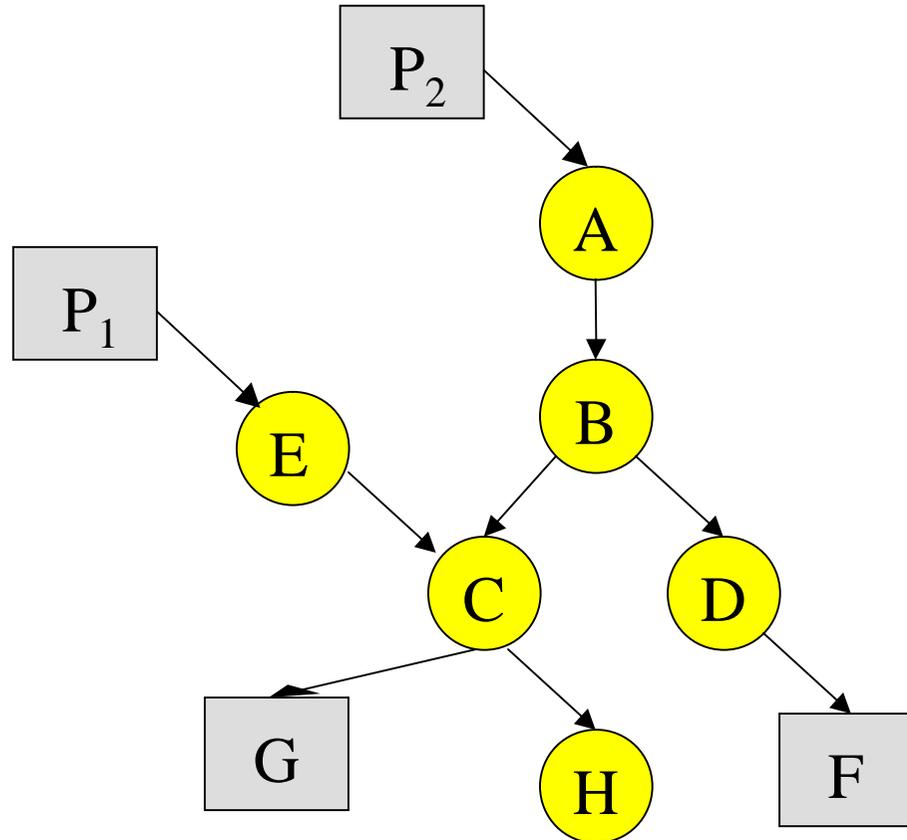
A	P(A)
Faux	0.6
Vrai	0.4

A	B	P(B A)
Faux	Faux	0.01
Faux	Vrai	0.99
Vrai	Faux	0.7
Vrai	Vrai	0.3

B	D	P(D B)
Faux	Faux	0.02
Faux	Vrai	0.98
Vrai	Faux	0.05
Vrai	Vrai	0.95

B	C	P(C B)
Faux	Faux	0.3
Faux	Vrai	0.7
Vrai	Faux	0.9
Vrai	Vrai	0.1

DAG: Outil d'interface



INFERENCE BAYESIENNE=

Chercher la loi CONDITIONNELLE et jointe de certains nœuds sachant d'autres nœuds

Inférence = apprentissage statistique via les probabilités conditionnelles

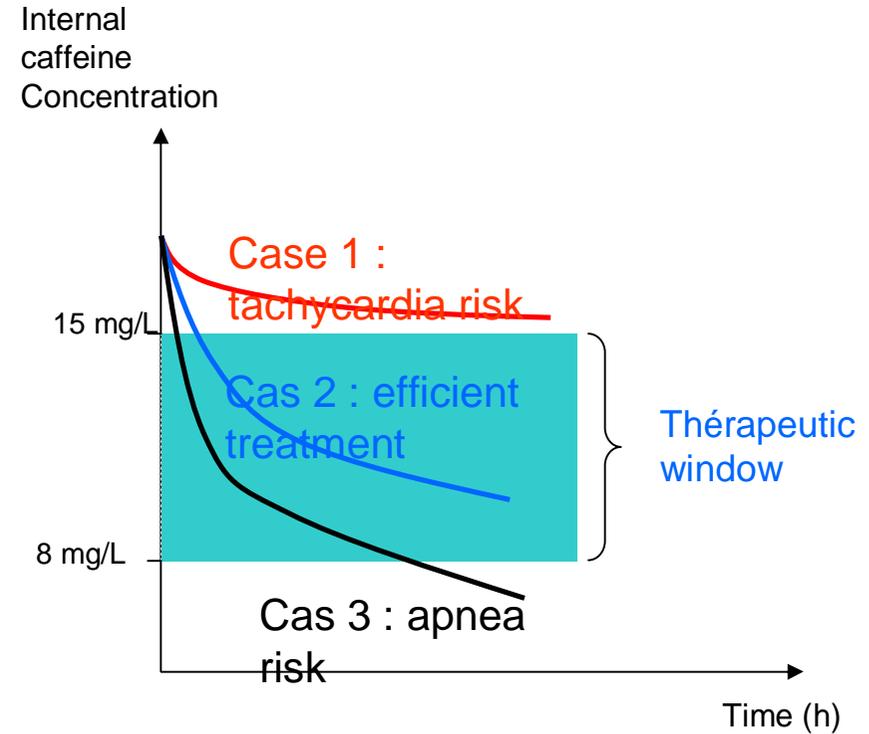
Théorie de la décision et risque

Aproche décisionnelle

Le calcul de risque de l'ingénieur



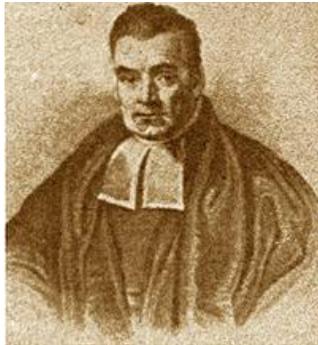
Optimisation d'un protocole d'injection de caféine chez les prématurés (Hpt Amiens)



Outils d'inférence bayésienne

De nombreux outils logiciels libres

Depuis Tom B. jusqu'aux nouvelles vagues MC(MC)



1762

Reverend
Thomas Bayes
(1702-1761)

Bayes formula

$$P(A|B) = \frac{P(A) \times P(B|A)}{P(B)}$$

$$[\theta|y] = \frac{[\theta] \times [y|\theta]}{[y]}$$

- Conjugate prior
- Non informative prior
- Bayesian decision theory

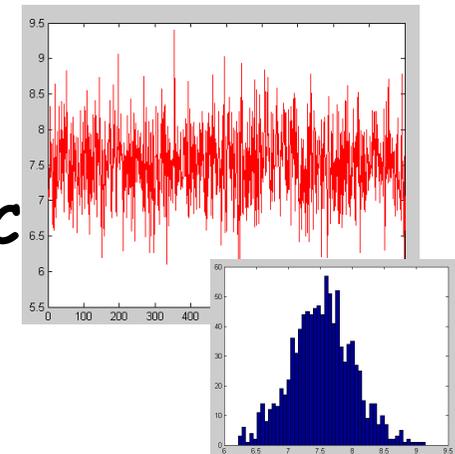
1950s

$$[\theta_1|y] = \int \int \int_{\theta_2} [\theta_1, \theta_2, \dots, \theta_n | y] d\theta_2 \dots d\theta_n$$

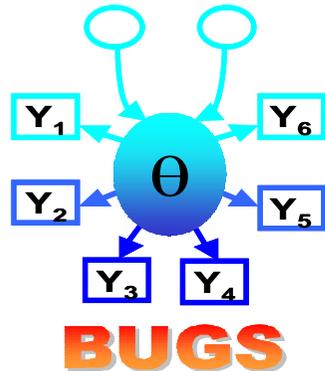
1990s

MCMC

- Gibbs sampling
- Hasting-Metropolis
- Interactive MCMC's
- Approximate Variational Bayes
- ABC



Friendly softwares, e.g. WinBUGS[®]



“Bayesian inference Using Gibbs Sampler”

Medical Research Council, Biostatistics Unit, Cambridge, UK
Imperial College, London, UK

É Bayesian statistical modeling

É Very popular

É Sampling based methods using Gibbs sampling (eventually hybrid)

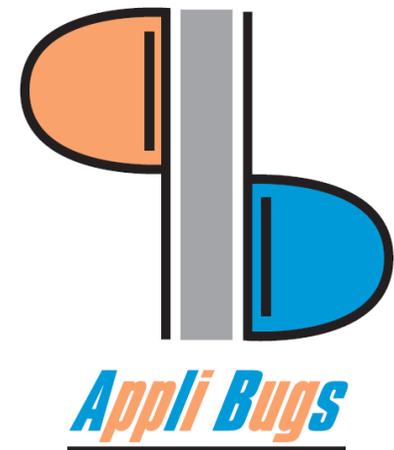
É Models can be described graphically

É Freeware, available at : www.mrc-bsu.cam.ac.uk/bugs/

É Extensions : $\left\{ \begin{array}{l} \text{Spatial models : GeoBUGS}^{\text{®}} \\ \text{Convergence : Coda}^{\text{®}}, \text{Boa}^{\text{®}} (\text{R} \text{ :kages}) \end{array} \right.$

É Coupling with 

<http://w3.jouy.inra.fr/unites/miaj/public/matrisq/Contacts/applibugs.welcome.html>



Ouvrages Bayes

**Un essor impressionnant du nombre d'ouvrages
de statistique bayésienne appliquée ... *avec
souvent les programmes R ou WinBugs associés***

Une approche innovante pour l'analyse
des risques environnementaux

Statistique pour l'environnement

Traitement bayésien de l'incertitude

Jacques Bernier Éric Parent Jean-Jacques Boreux

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int_0 f(x|\theta)\pi(\theta)d\theta}$$



Nouveauté

Editions
TEC
&
DOC

Collection
Statistique
et probabilités
appliquées

Dirigée par
Yadolah Dodge

COMITÉ EDITORIAL :
Christian Genest
Université Laval, Québec

Marc Hallin
Université libre de Bruxelles,
Belgique

Ludovic Lebart
INIST, Paris

Stephan Morgenthaler
EPFL, Lausanne

Gilbert Saporta
CNAM, Paris

Éric Parent, Jacques Bernier

Raisonnement bayésien Modélisation et inférence

Cette collection met à la disposition du public intéressé par la statistique (étudiants, enseignants, chercheurs) des ouvrages qui concilient effort pédagogique et travail permanent de mise à jour. Cette démarche implique de prendre en compte de façon sélective et critique les renouvellements des concepts, des champs d'application et des outils de traitement. Seules une compréhension profonde et une appropriation des connaissances permettent de s'adapter aux évolutions qui n'ont pas fini de bouleverser cette discipline.

Cet ouvrage expose de façon détaillée la pratique de l'approche statistique bayésienne à l'aide de nombreux exemples choisis pour leur intérêt pédagogique. Il insiste particulièrement sur l'emploi du raisonnement conditionnel qui fonde la cohérence profonde des méthodes de la modélisation et de l'inférence statistique sous le paradigme bayésien.

La première partie donne les principes généraux de modélisation statistique permettant d'encadrer mais aussi de venir au secours de l'imagination de l'apprenti modélisateur. En examinant des exemples de difficulté croissante, le lecteur forge les clés pour construire son propre modèle. La seconde partie présente les algorithmes de calcul les plus utiles pour estimer les inconnues du modèle. Chaque méthode d'inférence est présentée et illustrée par de nombreux cas d'applications. Si nécessaire, un rappel de leur cadre théorique essentiel est présenté sans démonstration.

Le livre cherche ainsi à dégager les éléments clés de la statistique bayésienne, en faisant l'hypothèse que le lecteur possède les bases de la théorie des probabilités et s'est déjà trouvé confronté à des problèmes ordinaires d'analyse statistique classique. Il peut servir de support à un cours de modélisation ou de statistique appliquée dans un programme de Master ou d'École d'Ingénieur. Il s'adresse également aux chercheurs et utilisateurs désireux de s'assurer de la pertinence des méthodes qu'ils emploient. Le débutant, au prix d'un investissement intellectuel acceptable, saura bien que le spécialiste, pourront y trouver les informations fondamentales pour comprendre et mettre en œuvre des modèles répondant à leurs besoins spécifiques.

ISBN : 978-2-287-39066-6



Springer.com

Collection
Statistique
et probabilités
appliquées

É. Parent, J. Bernier

Raisonnement bayésien Modélisation et inférence

Éric Parent, Jacques Bernier

Raisonnement bayésien Modélisation et inférence



$$VPP = \frac{(P \times Ss) + (1-P) \times (1-Sp)}{(P \times Ss) + (1-P) \times (1-Sp)}$$

Springer

Collection
Statistique
et probabilités
appliquées

Dirigée par
Yadolah Dodge

COMITÉ EDITORIAL :
Christian Genest
Université Laval, Québec

Marc Hallin
Université libre de Bruxelles,
Belgique

Ludovic Lebart
Télécom ParisTech, Paris

Stephan Morgenthaler
EPFL, Lausanne

Gilbert Saporta
CNAM, Paris

Jean-Jacques Boreux,
Éric Parent, Jacques Bernier

Pratique du calcul bayésien

Cette collection met à la disposition du public intéressé par la statistique (étudiants, enseignants, chercheurs) des ouvrages qui concilient effort pédagogique et travail permanent de mise à jour. Cette démarche implique de prendre en compte de façon sélective et critique les renouvellements des concepts, des champs d'application et des outils de traitement. Seules une compréhension profonde et une appropriation des connaissances permettent de s'adapter aux évolutions qui n'ont pas fini de bouleverser cette discipline.

Pratique du calcul bayésien est né de l'expérience acquise lors des cours donnés en sciences de l'environnement, tant à l'université de Liège (Arlon), qu'à la grande école AgroParisTech (Paris). Son fil conducteur peut se résumer par la formule « de la plume à la souris », tournée cependant à un opuscle retraçant la vie d'une école fréquentée jadis par le premier auteur.

La première partie privilégie les modèles statistiques paramétriques calculables « à la plume » et dépendent très riches, tant du point de vue de la présentation des concepts fondateurs du paradigme bayésien, que de leurs applications opérationnelles, notamment en matière d'aide à la décision. Des le premier chapitre, la représentation du modèle par un graphique cyclique orienté permet de distinguer clairement la phase où la créativité de chercher s'exprime de celle où il calcule. À cette fin, le logiciel libre WINBUGS sera très utile à l'apprenti modélisateur.

La seconde partie présente des applications réelles, plus sophistiquées, qui nécessitent souvent d'introduire une couche de variables latentes entre les observables et les paramètres. Conduire une inférence bayésienne sur ces modèles hiérarchiques implique un nouveau retour aux méthodes modernes de calcul et mobilise donc « la souris » de l'ordinateur.

45 € TTC

ISBN : 978-2-287-99666-5



Springer.com

Collection
Statistique
et probabilités
appliquées

J.-J. Boreux, É. Parent, J. Bernier

Pratique du calcul bayésien

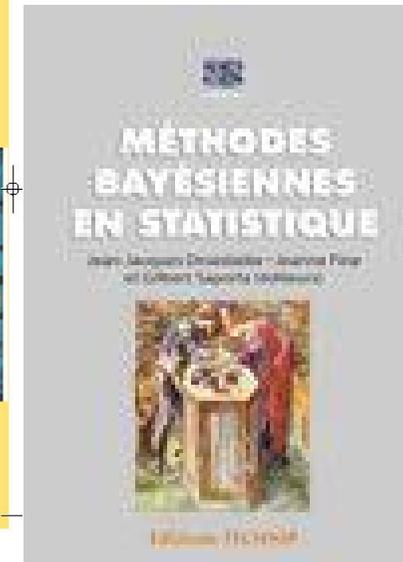
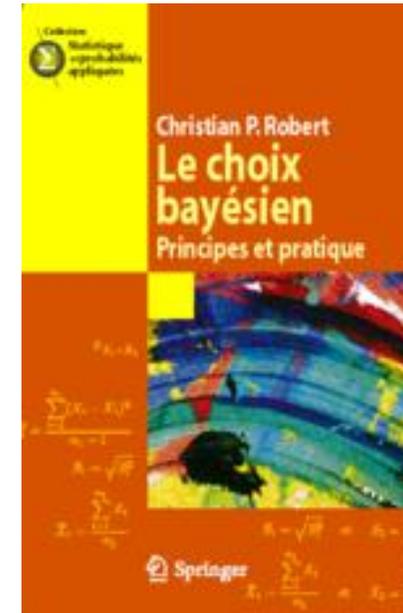
Jean-Jacques Boreux,
Éric Parent, Jacques Bernier

Pratique du calcul bayésien



$$VPP = \frac{(P \times Ss) + (1-P) \times (1-Sp)}{(P \times Ss) + (1-P) \times (1-Sp)}$$

Springer





Réf. - Statistique bayésienne (en anglais)

Box G.E.P. et Tiao G.C. 1992. Bayesian inference in statistical analysis. John Wiley and Sons, N-Y and London. Col. Wiley Classics Library. 588p.

Gelman, A., Carlin J.B., Stern H. and Rubin D.B. 2006. Bayesian data analysis. Second Edition. Chapman & Hall, London, New-York. Col. Texts in Statistical Science, 668 pp.

Hoff, P. D. 2009. A First Course in Bayesian Statistical Methods. Springer

Marin J.M., Robert, C.P. 2007. Bayesian Core. Springer, 2007.

Congdon P. 2001. Bayesian Statistical Modelling. Wiley. Wiley series in probability and statistics.

Berger, J.O. 1985. Statistical decision theory and Bayesian analysis. 2nd edition. Springer. Springer series in statistics.

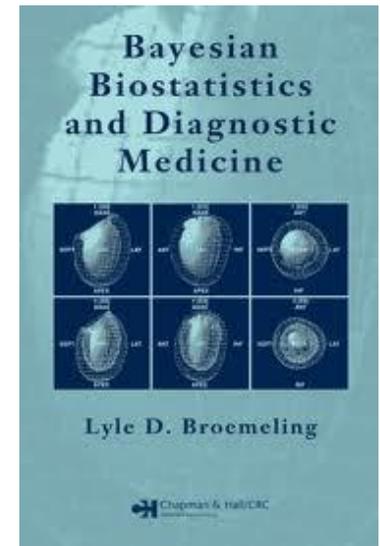
Kadane, J. B. 2011. Principles of Uncertainty. CRC Press. Chapman & Hall

Bernardo, J. M. and Smith, A. F. M. 1994. Bayesian theory. Wiley, London, 1 edition

Robert, C. P. and Casella, G. 1998. Monte Carlo Statistical Methods. Springer-Verlag

Un champ en expansion dans le domaine médical

Fifth Annual Bayesian Biostatistics Conference. The University of Texas MD Anderson Cancer Center Houston, Texas USA, January 23 - 25, 2012



Researchers are invited to submit abstracts for consideration as contributed presentations at the conference.

All topics in Bayesian biostatistics are acceptable, including clinical trial design and analysis, health policy, epidemiology, health economics, clinical decision making, comparative effectiveness, and bioinformatics.

Keynote presentation (Day 1): : *Bayesian Adjustment for Multiplicity and Subgroup Analysis* James O. Berger

Session Topics of Day 2: *Bayesian Methods for Evidence Synthesis to Drive Decision Making in Drug Development ; Bayesian Methods in Trial Design and Analysis; Incorporating Biological Structure in Bayesian Modeling*