

# De l'interprétation statistique des données à une modélisation probabiliste actualisée par les données : la synthèse bayésienne. Principes et exemples.

Jean-Baptiste Denis\*, Isabelle Albert\*\*

\*MIAJ - INRA, F78352 Jouy-en-Josas

Jean-Baptiste.Denis@Jouy.Inra.Fr,

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

\*\*Met@risk - INRA, F75231 Paris cedex 05

Isabelle.Albert@Paris.Inra.Fr

[http://www.paris.inra.fr/metarisk/members/albert\\_isabelle](http://www.paris.inra.fr/metarisk/members/albert_isabelle)

**Résumé.** Statisticiens plongés dans un domaine d'application où les données sont rares et hétérogènes, nous proposons d'utiliser successivement deux approches bayésiennes complémentaires : les réseaux bayésiens et la statistique bayésienne. Dans une première partie, les deux approches sont brièvement rappelées pour montrer qu'à part le théorème de Bayes, elles n'ont, dans nos acceptations des deux concepts, rien en commun. S'appuyant ensuite sur la modélisation du nombre de campylobactérioses en France liées à la consommation de poulets, un nouveau point de vue est suggéré pour l'interprétation des données. Il s'agit (1) de modéliser en soi le phénomène d'intérêt à l'aide d'un réseau bayésien ; puis (2) de l'étendre pour définir la vraisemblance des données disponibles et extraire l'information qu'elles contiennent par conditionnement, c'est-à-dire en appliquant le principe de la statistique bayésienne. R et les logiciels de la famille BUGS se révèlent bien adaptés pour la réalisation pratique de cette proposition.

## 1 Introduction

Dans certains domaines, comme celui de la sécurité sanitaire microbiologique liée à l'alimentation humaine dans lequel nous travaillons, les approches statistiques standard ne sont guère aisées à pratiquer. Alors que les questions sont précises et les réponses lourdes de conséquence, les données sont peu nombreuses et celles qui sont disponibles sont très hétérogènes. Par exemple : si on permet à une ville de 10000 habitants de consommer une eau de source dont les deux derniers prélèvements ont révélé la présence d'oocystes de cryptosporidium, quelles peuvent en être les conséquences sur la santé de la population en termes de maladies, voire de décès ?

Traditionnellement, les décisions sont prises par les politiques après consultation d'experts qui synthétisent implicitement et subjectivement un certain nombre de situations similaires soit vécues soit relatées dans la littérature du domaine. Mais cette démarche pragmatique est de moins en moins prise en compte car d'une part les citoyens exigent de plus en plus les justifications des