CHAPITRE PREMIER

DEUX SIMPLES DEVINETTES POUR ÉBRANLER NOS CERTITUDES

Où le lecteur désarçonné constate qu'il est très facile de commettre des erreurs en toute probité quand on raisonne avec des probabilités.

Ce qui nous attire des ennuis, ce n'est pas notre ignorance, ce sont nos fausses certitudes[1].

Mark Twain

Afin de vous faire entrer de plain-pied, cher lecteur, dans le domaine ébouriffant des pièges probabilistes, voici deux charmantes devinettes qui seront plus éloquentes qu'un long discours.

Devinette 1 : Le médecin

Supposons que vous soyez un médecin et que vous disposiez du test Elisa pour détecter si une personne est atteinte du SIDA. Ce test est connu pour avoir une sensibilité de l'ordre de 99 %, autrement dit sur 100 personnes porteuses de la maladie, seule une personne échappera au test.

1. La splendide citation originale est : « *The greatest obstacle to discovery is not ignorance –it is the illusion of knowledge.* »

Figure 2. La devinette au médecin

Dans le cadre d'un programme de dépistage systématique pour les femmes enceintes, vous effectuez le test sur Madame X. Il se révèle positif.

Quelle probabilité Madame X a-t-elle d'être porteuse de la maladie ?

Pour simplifier, veuillez donner une estimation sur l'échelle suivante en vous fiant à votre bon sens (indiquer l'intervalle où selon vous se situe la probabilité) :

100 %

80 %

60 %

40 %

20 %

0 %

?

Détendez-vous, prenez quelques secondes avant de choisir votre réponse, vous l'avez ?

Devinette 2 : Le juge d'instruction

Vous êtes à présent dans la peau d'un juge d'instruction.

Un meurtre a été commis et du sang appartenant à un inconnu a été trouvé sur la victime. Nous supposons que le sang sur la victime ne peut qu'appartenir à l'assassin quel qu'il soit.

Vous ordonnez qu'un test ADN soit effectué sur la population vivant à proximité du lieu où habitait la victime. Ce test se révèle positif sur Monsieur X, autrement dit le profil génétique de Monsieur X est déterminé par le test comme étant identique à celui de l'inconnu.

Vous savez qu'il n'y a qu'une chance sur 1 milliard de trouver un autre individu sur Terre ayant un profil génétique identique à celui de l'inconnu.

Quelle est selon vous la probabilité que Monsieur X soit le coupable (utilisez l'échelle précédente) ?

Là encore, prenez quelques secondes et fiez-vous à votre intuition. Vous y êtes ?

Lorsque vous êtes prêt, rendez-vous sur la page suivante pour les solutions.

Figure 3. La devinette à l'enquêteur

➲ **SOLUTION pour la devinette 1**

Avec une telle fiabilité, ne pensez-vous pas qu'un résultat positif indique presque sûrement que la personne est atteinte de la maladie ?

Eh bien *détrompez-vous* !

En fait, contrairement à ce que l'intuition suggère à la plupart d'entre nous, la probabilité que la personne soit porteuse de la maladie est seulement de 17 %, soit moins d'une chance sur cinq.

➲ **SOLUTION pour la devinette 2**

Ici les chiffres sont tellement impressionnants que l'on peut être pratiquement sûr que la personne suspecte est coupable au-delà de tout doute raisonnable. Et pourtant...

Là encore, *détrompez-vous* !

Contrairement à ce que l'intuition suggère là encore à la plupart d'entre nous, la probabilité que la personne soit bien l'auteur du crime est seulement de **9 %** (au maximum), moins d'une chance sur dix, soit tout à fait insuffisante pour condamner Monsieur X à la perpétuité ou à la peine de mort suivant les pays. Si l'on se base uniquement sur ces données factuelles, Monsieur X ne peut être tenu pour coupable au-delà du doute raisonnable !

Nous allons maintenant expliquer pourquoi, sur des exemples apparemment si simples, le raisonnement humain *intuitif* est pris en défaut avec autant de facilité.

Pour cela, il nous faudra au préalable faire un détour par le raisonnement bayésien et utiliser une technique originale issue des recherches en intelligence artificielle.

Les solutions aux devinettes pourront alors être exposées en pleine lumière à la fin du chapitre qui vient.

CHAPITRE DEUXIÈME

FAISONS CONNAISSANCE AVEC LES RÉSEAUX BAYÉSIENS

➡ Où le lecteur comprend le grand intérêt à faire coopérer humains et machines en toute intelligence.

2.1. QU'EST-CE QU'UN RAISONNEMENT PROBABILISTE ?

Les faits sont têtus. Il est plus facile de s'arranger avec les statistiques.

Mark Twain

Un raisonnement probabiliste opère toujours dans un contexte donné, donc en tenant compte d'un ensemble de connaissances. On cherche par conséquent, à un instant donné, à délivrer le raisonnement probabiliste le plus précis, compte tenu de toutes les connaissances pertinentes du moment.

Quelle que soit la connaissance, il est possible d'attribuer une probabilité à chaque valeur possible d'une variable.

Supposons que vous n'ayez aucune connaissance particulière sur une variable X qui peut prendre les valeurs a, b et c. La probabilité *a priori* dans un tel cas sera simplement donnée par $P(X=a) = P(X=b) = P(X=c) = 1/3$

L'une des conceptions les plus simples de la probabilité est celle d'une simple proportion au sein d'une population ou d'un groupe.

Si $P(A)$ dénote une probabilité inconditionnelle (la proportion des membres de la population totale possédant l'attribut A), $P(A|B)$ dénote la **probabilité conditionnelle** de A sachant B, autrement dit la proportion dans le sous-groupe qui possède l'attribut B de ceux qui possèdent également l'attribut A.

Plus généralement, étant donné les attributs A et B que peuvent posséder les individus d'une population de référence, les probabilités obéiront toujours aux conditions suivantes :

1. $0 \leq P(A) \leq 1$
2. Si A est possédé par tous les membres de la population, $P(A) = 1$
3. Si A et B ne peuvent être présents ensemble, $P(A \text{ ou } B) = P(A) + P(B)$
4. $P(A \text{ et } B) = P(A|B) \times P(B)$

Ces simples règles forment en fait la base de la théorie mathématique des probabilités et continuent de s'appliquer aux probabilités basées sur d'autres interprétations que celle purement statistique présentée ici.

Cela signifie que, si nous désirons manipuler des probabilités, quelle que soit leur origine ou interprétation, nous pouvons toujours faire comme si elles avaient été déterminées en tant que proportions dans un groupe et ses sous-groupes, comme dans l'approche statistique, et utiliser cette représentation pour résoudre nos problèmes.

D'une façon générale, comme nous le verrons à maintes reprises dans ce livre, il faut toujours se méfier des évidences lorsqu'on raisonne dans un contexte incertain.

Qu'est-ce qu'un **raisonnement bayésien** ?

C'est une inférence probabiliste (donc le calcul d'une probabilité) qui met en jeu le théorème de Bayes que nous expliciterons plus loin et qui permet de réviser les probabilités en fonction de la connaissance disponible. En pratique, cela consiste souvent à démarrer à partir d'observations pour remonter jusqu'aux causes.

Dans un contexte donné, il peut arriver que l'affirmation la plus probable (y compris au-delà du doute raisonnable) soit en réalité fausse. Il n'en reste pas moins que l'approche rationnelle nous impose de choisir l'alternative la plus vraisemblable, compte tenu de la totalité de la connaissance disponible à un moment donné, quitte à réviser la conclusion lorsque la connaissance est modifiée.

Dans l'approche bayésienne, une probabilité n'est pas interprétée nécessairement comme le passage à la limite d'une fréquence, mais comme l'évaluation d'un état de connaissance, autrement dit le degré de confiance attribué à une hypothèse. Rien n'interdit néanmoins d'y utiliser les probabilités avec le point de vue fréquentiste (c'est même souvent le cas).

Par rapport aux approches statistiques classiques, la statistique bayésienne est très bien adaptée aux domaines pour lesquels on manque de données significatives qui permettraient d'élaborer un modèle fréquentiste robuste. Les résultats obtenus par ces différentes approches convergent dans le cas où l'on dispose d'un grand nombre d'observations dans le domaine considéré.

2.2. RÉSEAUX BAYÉSIENS : PRÉSENTATION ET DÉFINITIONS

> *Les réseaux bayésiens sont par nature des modèles permettant de traiter l'information incomplète. Un réseau bayésien peut calculer la probabilité de n'importe laquelle de ses variables, conditionnellement à la connaissance d'un sous-ensemble quelconque de variables observées*[1].
>
> *Le domaine des réseaux bayésiens a comme particularité d'allier deux champs différents des mathématiques dans le but de représenter l'incertitude : la théorie des graphes, d'une part, qui fournit le cadre nécessaire pour une modélisation qualitative des connaissances ; et la théorie des probabilités, d'autre part, qui permet d'introduire une information quantitative dans ces connaissances.*
>
> Patrick Naïm *et al.*

Le but de ce paragraphe est d'initier le lecteur aux réseaux bayésiens en lui donnant la définition la plus claire possible et en montrant en quoi on peut les utiliser efficacement pour tout type de raisonnement probabiliste.

Tout d'abord, il me faut rappeler quelques termes qui seront essentiels pour la suite.

La probabilité jointe est la probabilité que deux événements (ou plus) se produisent en conjonction, donc que l'un et l'autre soient observés.

La probabilité jointe de A et B s'écrit en adoptant l'une des conventions suivantes :

$P(A \text{ et } B)$ ou $P(A \cap B)$ ou $P(A, B)$ ou encore $P(A \& B)$.

1. Les citations en tête de ce paragraphe sont extraites du livre « Réseaux bayésiens » cité en bibliographie.

La probabilité marginale ou **probabilité *a priori*** est la probabilité d'un événement donné, en l'absence de toute autre connaissance dans le contexte considéré.

La probabilité marginale de A s'écrit P (A), et la probabilité marginale de B s'écrit P (B).

Il est important de noter que ces définitions n'imposent aucune relation causale ou temporelle entre les événements A et B.

Il se peut que A précède B ou vice-versa ou ils peuvent se produire en même temps. A peut être la cause de B ou vice-versa ou bien il peut ne pas exister du tout de lien causal entre les deux événements.

Le conditionnement des probabilités, autrement dit leur mise à jour en tenant compte d'une éventuelle connaissance supplémentaire est réalisé soit à l'aide du théorème de Bayes, soit en utilisation la règle de chaînage des probabilités conditionnelles (qui intervient dans la formule générale pour une loi jointe qui sera donnée un peu plus loin), soit une combinaison des deux.

La **probabilité conditionnelle** est la probabilité d'un événement A, sachant que l'événement B s'est produit. Nous l'avons déjà rencontrée au paragraphe précédent mais nous allons donner ici quelques précisions.

Elle s'écrit $P(A \mid B)$ et se lit « probabilité de A, sachant B ».

Étant donné deux événements A, B avec $P(B) > 0$, la probabilité conditionnelle de A sachant B est définie par

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

Deux événements A and B sont statistiquement indépendants si et seulement si $P(A \cap B) = P(A)\,P(B)$.

Ainsi, si A et B sont indépendants, alors leur probabilité jointe peut être exprimée comme un simple produit de leurs probabilités individuelles.

Pour deux événements A et B indépendants, on a :

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

En d'autres termes, si A et B sont indépendants, alors la probabilité conditionnelle de A sachant B est simplement identique à la probabilité de A seul ; similairement, la probabilité conditionnelle de B sachant A est simplement identique à la probabilité de B seul.