Iwatarax

Probabilités conditionnelles

Probabilités conditionnelles

Petit mot

Salut à toutes et à tous et welcome pour votre deuxième cours de probabilités de l'année (miam 😋) !!

Cette fois-ci, nous quittons le monde des probabilités élémentaires pour rejoindre celui des probabilités conditionnelles (que je trouve plus intéressant).

Si vous avez fait des maths au lycée, ce cours ne devrait pas vous choquer puisque nous parlerons notamment d'arbres de probabilité, du théorème de Bayes et de la notion d'indépendance.

Introduction

Définition

Soient *A* et *B* deux événements quelconques d'un ensemble fondamental *Ω* auxquels on associe respectivement les probabilités $P(A)$ et $P(B)$.

On s'intéresse à l'effet que peut avoir la réalisation d'un événement sur la probabilité de réalisation du second, autrement dit on étudie comment peut varier la probabilité d'un événement sachant qu'un autre s'est déjà réalisé.

La **probabilité de *A* sachant *B*** est définie par la relation : $P_B(A) = \frac{P(A \cap B)}{P(B)}$

PS : $P_B(A)$ peut également s'écrire $P(A \mid B)$

Attention

Il ne faut pas confondre les probabilités suivantes :
- ***$P_B(A)$*** : "probabilité de *A* sachant *B*", elle désigne la probabilité que l'événement ***A*** se réalise au sein de l'ensemble ***B***
- ***$P(A \cap B)$*** : "probabilité de *A* inter *B*", elle désigne la probabilité que l'événement *A* **ET** l'événement *B* se réalisent **simultanément** au sein de l'ensemble fondamental ***Ω***

Théorème de la multiplication

Reprenons la définition d'une **probabilité conditionnelle** :

$$P_B(A) = \frac{P(A \cap B)}{P(B)} \qquad P_A(B) = \frac{P(A \cap B)}{P(A)}$$

Nous obtenons la relation : $P(A \cap B) = P(B) \times P_B(A) = P(A) \times P_A(B)$

La formule généralisée est la suivante :

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \times P_{A_1}(A_2) \times \ldots \times P_{A_1 \cap \ldots \cap A_{n-1}}(A_n)$$

Exemple de situation (1)

Dans une urne, on retrouve 20 pins dont 5 sont abîmés. On vous propose de tirer 3 pins au hasard. En notant *A* l'événement "le pin est abîmé", quelle est la probabilité pour que les 3 pins tirés au sort soient abîmés ?

Ici, il faut comprendre que l'on cherche la probabilité $P(A_1 \cap A_2 \cap A_3)$.

D'après le théorème de la multiplication, nous avons la relation suivante :

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P_{A_1}(A_2) \times P_{A_1 \cap A_2}(A_3)$$

Au 1er tirage, la probabilité de tirer un pin abîmé vaut 5/20.

Sachant qu'un pin abîmé est tiré au 1er tirage, la probabilité de tirer un deuxième pin abîmé vaut 4/19.

Enfin, sachant qu'un pin abîmé est tiré au 1er et au 2ème tirage, la probabilité de tirer un troisième pin abîmé vaut 3/18.

Au final, on a donc : $P(A_1 \cap A_2 \cap A_3) = \frac{5}{20} \times \frac{4}{19} \times \frac{3}{18} = \frac{1}{114}$

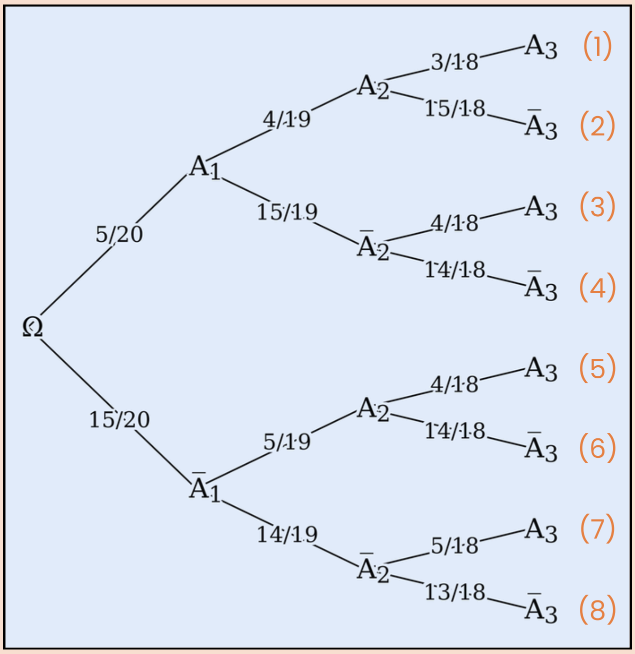
Arbre de probabilité

On considère une séquence finie d'expériences présentant un nombre fini de résultats possibles dont les probabilités dépendent du **résultat précédent**.

Un **arbre de probabilité** (ou diagramme en arbre) permet de représenter une telle situation (probabilités conditionnelles) et de faciliter les calculs.

Exemple de situation (2)

La situation décrite dans l'exemple précédent peut être représentée par l'arbre de probabilité suivant :



La probabilité qu'un chemin particulier de l'arbre se réalise est, d'après le théorème de la multiplication, le <u>produit</u> des probabilités de chaque branche du chemin.

Exemple : la probabilité d'avoir 3 pins abîmés se calcule en faisant le produit des probabilités des branches du chemin 1

Puisque les chemins s'excluent <u>mutuellement</u>, il est possible de les <u>additionner</u> pour déterminer certaines probabilités.

Exemple : la probabilité d'avoir exactement 2 pins abîmés est égal à la somme des probabilités des chemins 2, 3 et 5 (cf. arbre)

Formule et théorème de Bayes

Formule de Bayes

Le **théorème de multiplication** nous donne la relation suivante :

$$P(A \cap B) = P(B) \times P_B(A) = P(A) \times P_A(B)$$

Bayes reprend cette relation pour donner les “**formules de Bayes**” :

$$P_B(A) = \frac{P(A) \times P_A(B)}{P(B)} \qquad P_A(B) = \frac{P(B) \times P_B(A)}{P(A)}$$

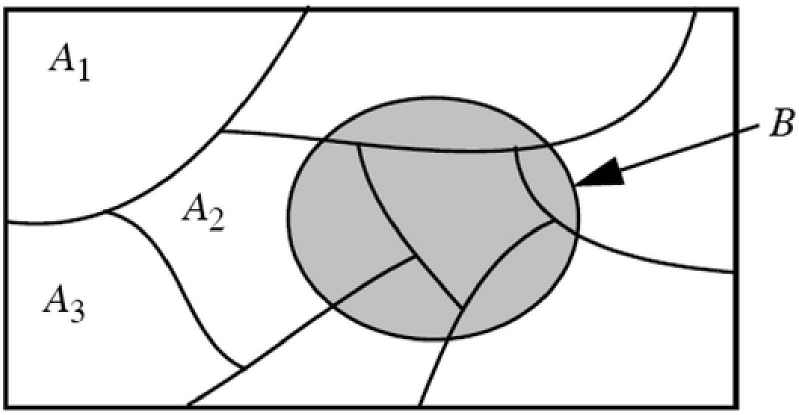
PS : elles peuvent paraître toutes bêtes comme ça, mais ces formules sont fondamentales en statistique

Théorème de Bayes

Soient A_1 , A_2 , ... , A_n des événements tels qu'ils forment une partition de l'ensemble fondamental Ω.

PS : pour rappel, une partition d'un ensemble E correspond à une subdivision de E en sous-ensembles disjoints dont la réunion forme E

Considérons *B* un événement quelconque tel que :



Sachant que $\Omega = A_1 \cup A_2 \cup ... \cup A_n$ et que $B = B \cap \Omega$, nous avons :

$$B = B \cap (A_1 \cup A_2 \cup \ldots \cup A_n)$$

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)$$

Comme les événements A_i sont deux à deux disjoints, nous en déduisons que les événements $B \cap A_i$ sont également deux à deux disjoints et forment donc une **partition de *B***.

D'après le théorème des probabilités totales, nous avons :

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_n)$$

En appliquant le théorème de la multiplication, nous obtenons :

$$P(B) = P(A_1) \times P_{A_1}(B) + P(A_2) \times P_{A_2}(B) + \ldots + P(A_n) \times P_{A_n}(B)$$

Pour chaque A_i, la <u>formule de Bayes</u> nous donne la relation suivante :

$$P_B(A_i) = \frac{P(A_i) \times P_{A_i}(B)}{P(B)}$$

Dès lors, nous obtenons le <u>théorème de Bayes</u> :

$$P_B(A_i) = \frac{P(A_i) \times P_{A_i}(B)}{P(A_1) \times P_{A_1}(B) + P(A_2) \times P_{A_2}(B) + \ldots + P(A_n) \times P_{A_n}(B)}$$

Application du théorème de Bayes

Soit une population au sein de laquelle une maladie grave affecte **une** personne sur **10 000**.

On se propose de tester un tout nouveau test de dépistage tel que :
- Si une personne est malade, le test est positif à **99%**
- Si une personne est saine, le test est positif à **0,5%**

Même si ces chiffres paraissent incroyables, ce qui nous intéresse vraiment c'est la probabilité qu'une personne soit réellement malade lorsque son test est positif.

On définit alors 2 événements :
- **Evénement M^+** (resp. **M^-**) : être **malade** (resp. **sain**)
- **Evénement T^+** (resp. **T^-**) : avoir un test **positif** (resp. **négatif**)

On cherche la probabilité d'être malade sachant que l'on a un test positif, autrement dit $P_{T+}(M^+)$.

Les informations dont nous disposons sont les suivantes :
- $P(M^+) = 0{,}0001$ et $P(M^-) = 0{,}9999$
- $P_{M+}(T^+) = 0{,}99$ et $P_{M-}(T^+) = 0{,}005$

D'après le <u>théorème de Bayes</u>, nous avons la relation suivante :

$$P_{T^+}(M^+) = \frac{P(M^+) \times P_{M^+}(T^+)}{P(M^+) \times P_{M^+}(T^+) + P(M^-) \times P_{M^-}(T^+)}$$

En remplaçant par nos valeurs, on trouve une probabilité de **1,9%**, ce qui signifie que lorsque l'on obtient un test positif, il n'y a en réalité qu'environ 2% de chance d'être véritablement malade (c'est peu…).

Indépendance

Définition

Soient A et B deux événements d'un espace probabilisé.

Les événements A et B sont dits "**indépendants**" si et seulement si :

$$P(A \cap B) = P(A) \times P(B)$$

Concrètement, cela signifie que la probabilité que A soit réalisé <u>n'est pas modifiée</u> par le fait que B se produise, ce qui nous donne l'égalité suivante :

$$P_B(A) = P(A)$$

PS : cette formule est également valable appliquée à B, l'indépendance se faisant aussi bien dans un sens que dans l'autre

Le fait que A et B soient indépendants implique plusieurs choses :
- A et $\bar{B}$ sont indépendants
- $\bar{A}$ et B sont indépendants
- $\bar{A}$ et $\bar{B}$ sont indépendants

La notion d'indépendance peut être généralisée à <u>plus de 2</u> événements.

Ainsi, pour 3 événements A, B et C, on parle d'indépendance :
- **Si** A, B et C sont indépendants deux à deux
- **ET si** $P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$

PS : la seconde condition <u>n'est pas</u> une conséquence de la précédente

Inclusion et exclusion

Lorsque $A \subset B$, on se retrouve avec $A \cap B = A$ et donc $P(A \cap B) = P(A)$.

En appliquant la <u>formule de Bayes</u>, nous obtenons les relations suivantes :

$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} \qquad P_A(B) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$$

Par conséquent, cela signifie que, lorsqu'un événement est <u>inclus</u> dans un autre, les événements concernés <u>ne peuvent pas</u> être **indépendants**.

Lorsque *A* et *B* sont <u>disjoints</u>, on se retrouve avec $A \cap B = \emptyset$ et $P(A \cap B) = 0$.

*PS : on parle d'événements "**exclusifs**" ou "**incompatibles**" (synonymes)*

Comme $P(A \cap B) = 0$, on se retrouve avec la relation suivante :

$$P_B(A) = P_A(B) = 0$$

Par conséquent, cela signifie que lorsque des événements sont <u>disjoints</u>, les événements concernés <u>ne peuvent pas</u> être **indépendants**.

Incompatibilité et indépendance

Les notions d'**incompatibilité** et d'**indépendance** sont trop souvent confondues, pourtant elles ne s'adressent pas du tout aux mêmes situations.

D'une part, on retrouve les <u>événements incompatibles</u> :
- Ils <u>ne font pas</u> intervenir leur probabilité
- Ils <u>ne peuvent pas</u> se réaliser en même temps
- Ils sont caractérisés par la relation : $\boldsymbol{P(A \cup B) = P(A) + P(B)}$

D'autre part, on retrouve les événements indépendants :
- Ils sont <u>liés</u> à leur probabilité
- Ils peuvent se produire <u>en même temps</u>, mais la réalisation de l'un <u>n'a pas d'influence</u> sur la réalisation du second
- Ils sont caractérisés par la relation : $\boldsymbol{P(A \cap B) = P(A) \times P(B)}$

FIN