Bases Indispensables des Mathématiques

Chapitre 2 : Bases des Probabilités (Partie 3).

Julian Tugaut

Télécom Saint-Étienne

Sommaire

Plan

Notion de probabilité conditionnelle - 1

Soient A et B deux évènements associés à une même expérience aléatoire e. On note Ω l'espace fondamental associé. On suppose $\mathbb{P}(A) \neq 0$.

Notion de probabilité conditionnelle - 1

Soient A et B deux évènements associés à une même expérience aléatoire e. On note Ω l'espace fondamental associé. On suppose $\mathbb{P}(A) \neq 0$.

La probabilité conditionnelle de B par rapport à l'évènement A est la probabilité que B soit réalisé lorsque l'on sait que A est réalisé.

Notion de probabilité conditionnelle - 1

Soient A et B deux évènements associés à une même expérience aléatoire e. On note Ω l'espace fondamental associé. On suppose $\mathbb{P}(A) \neq 0$.

La probabilité conditionnelle de B par rapport à l'évènement A est la probabilité que B soit réalisé lorsque l'on sait que A est réalisé.

Notation

La probabilité conditionnelle de B en sachant A est notée $\mathbb{P}_A(B)$ ou $\mathbb{P}\left(B \mid A\right)$. On préfère la seconde notation. En effet, par la suite, $\mathbb{P}_X$ représentera la loi de probabilité de toute variable aléatoire X. Néanmoins, si aucune confusion n'est possible, on s'autorisera de temps en temps la première notation.

Notion de probabilité conditionnelle - 2

Exemple

On jette un dé. Soient les évènements $A :=$"on obtient une face paire" et $B :=$"on obtient une face supérieure ou égale à trois". L'univers $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ est muni de l'équiprobabilité. Ici, $A = \{\omega_2, \omega_4, \omega_6\}$ d'où $\mathbb{P}(A) = \frac{3}{6} = \frac{1}{2}$. Et, $B = \{\omega_3, \omega_4, \omega_5, \omega_6\}$ d'où $\mathbb{P}(B) = \frac{4}{6} = \frac{2}{3}$. Puis, $A \cap B = \{\omega_4, \omega_6\}$ d'où $\mathbb{P}(A \cap B) = \frac{2}{6} = \frac{1}{3}$.

Notion de probabilité conditionnelle - 3

On jette le dé. Supposons que l'on sache que l'on a obtenu une face paire. (Pour ce faire, il suffit de relancer le dé jusqu'à ce que la face obtenue soit paire.) Quelle est la probabilité qu'on obtienne une face supérieure ou égale à trois ? Dans notre exemple, l'univers est A et la probabilité que B soit réalisé est donc la probabilité que $A \cap B$ soit réalisé pour l'univers A muni de l'équiprobabilité associée $\mathbb{P}_A$ avec $\mathbb{P}_A(\omega_2) = \mathbb{P}_A(\omega_4) = \mathbb{P}_A(\omega_6) = \frac{1}{3}$. Il vient ainsi $\mathbb{P}_A(A \cap B) = \frac{\#A\cap B}{\#A} = \frac{2}{3} = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)}$. Conséquemment, on a $\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(A)}$.

Définition

Soit A un évènement de probabilité $\mathbb{P}(A)$ non nulle. On définit l'application $\mathbb{P}\left(. \mid A\right)$ qui, à tout évènement B fait correspondre le nombre

$$\mathbb{P}\left(B \,|\, A\right) := \frac{\mathbb{P}\left(A \cap B\right)}{\mathbb{P}(A)}\,.$$

On commence par établir le résultat suivant :

Définition

Soit A un évènement de probabilité $\mathbb{P}(A)$ non nulle. On définit l'application $\mathbb{P}(. \mid A)$ qui, à tout évènement B fait correspondre le nombre

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

On commence par établir le résultat suivant :

Théorème

$\mathbb{P}(. \mid A)$ est une probabilité sur Ω.

Remarques

Remarque

Tous les résultats portant sur les probabilités peuvent donc s'appliquer à l'application $B \mapsto \mathbb{P}(B \mid A)$. Notamment :

Remarques

Remarque

Tous les résultats portant sur les probabilités peuvent donc s'appliquer à l'application $B \mapsto \mathbb{P}(B \mid A)$. Notamment :

- la probabilité conditionnelle par rapport à A de l'évènement contraire de B est 1 moins la probabilité conditionnelle de l'évènement B,

Remarques

Remarque

Tous les résultats portant sur les probabilités peuvent donc s'appliquer à l'application $B \mapsto \mathbb{P}(B \mid A)$. Notamment :

- la probabilité conditionnelle par rapport à A de l'évènement contraire de B est 1 moins la probabilité conditionnelle de l'évènement B,
- la probabilité conditionnelle par rapport à A de l'ensemble vide est 0,

Remarques

Remarque

Tous les résultats portant sur les probabilités peuvent donc s'appliquer à l'application $B \mapsto \mathbb{P}(B \mid A)$. Notamment :

- la probabilité conditionnelle par rapport à A de l'évènement contraire de B est 1 moins la probabilité conditionnelle de l'évènement B,
- la probabilité conditionnelle par rapport à A de l'ensemble vide est 0,
- la probabilité conditionnelle par rapport à A de B est inférieure à celle de C si $B \subset C$,

Remarques

Remarque

Tous les résultats portant sur les probabilités peuvent donc s'appliquer à l'application $B \mapsto \mathbb{P}(B \mid A)$. Notamment :

- la probabilité conditionnelle par rapport à A de l'évènement contraire de B est 1 moins la probabilité conditionnelle de l'évènement B,
- la probabilité conditionnelle par rapport à A de l'ensemble vide est 0,
- la probabilité conditionnelle par rapport à A de B est inférieure à celle de C si $B \subset C$,
- le théorème des probabilités totales est encore valide.

Remarques

Remarque

Tous les résultats portant sur les probabilités peuvent donc s'appliquer à l'application $B \mapsto \mathbb{P}(B \mid A)$. Notamment :

- la probabilité conditionnelle par rapport à A de l'évènement contraire de B est 1 moins la probabilité conditionnelle de l'évènement B,
- la probabilité conditionnelle par rapport à A de l'ensemble vide est 0,
- la probabilité conditionnelle par rapport à A de B est inférieure à celle de C si $B \subset C$,
- le théorème des probabilités totales est encore valide.

Exercice

Vérifiez qu'en toute généralité, si A et B sont deux évènements associés à une même expérience aléatoire, alors
$\mathbb{P}[A \mid B^c] \neq 1 - \mathbb{P}[A \mid B]$.

ATTENTION

Remarque

On définit une probabilité conditionnelle mais en aucun cas on ne définit des évènements conditionnels. En effet, on ne considère que des éléments de la tribu. Ainsi, la probabilité conditionnelle de B sachant A est plutôt la probabilité conditionnelle, sachant A, de B.

Théorème des probabilités composées

Théorème des probabilités composées

Soient deux évènements A et B de probabilités non nulles. Alors, on a

$$\mathbb{P}\left(A \cap B\right) = \mathbb{P}\left(B \mid A\right) \mathbb{P}(A) = \mathbb{P}\left(A \mid B\right) \mathbb{P}(B)\,.$$

Théorème de Bayes - 1

Voyons maintenant le théorème de Bayes, lequel porte aussi le nom de théorème de la probabilité des causes.

Théorème de Bayes - 1

Voyons maintenant le théorème de Bayes, lequel porte aussi le nom de théorème de la probabilité des causes.

Soit $(A_1, \cdots, A_n)$ un système complet d'évènements de Ω. Soit B un évènement de Ω de probabilité non nulle. On suppose que l'on connaît $\mathbb{P}(A_1), \cdots, \mathbb{P}(A_n)$ et que ces n probabilités sont strictement positives. On suppose que l'on connaît également $\mathbb{P}(B \mid A_1), \cdots, \mathbb{P}(B \mid A_n)$. On cherche à calculer $\mathbb{P}(A_k \mid B)$ pour tout $1 \leq k \leq n$.

Théorème de Bayes - 2

On utilise alors le théorème des probabilités composées :

$$\mathbb{P}(A_k \mid B) = \frac{\mathbb{P}(B \cap A_k)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A_k)\mathbb{P}(A_k)}{\mathbb{P}(B)} .$$

Théorème de Bayes - 2

On utilise alors le théorème des probabilités composées :

$$\mathbb{P}(A_k \mid B) = \frac{\mathbb{P}(B \cap A_k)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A_k)\mathbb{P}(A_k)}{\mathbb{P}(B)} .$$

Or, $(A_1, \cdots, A_n)$ est un système complet d'évènements de Ω. Ainsi, d'après le Théorème de décomposition, il vient

$$\mathbb{P}(B) = \sum_{k=1}^{n} \mathbb{P}\left(B \cap A_k\right) .$$

Théorème de Bayes - 3

On utilise à nouveau le théorème des probabilités composées et l'on a alors

$$\mathbb{P}(A_k \mid B) = \frac{\mathbb{P}\left(B \mid A_k\right)\mathbb{P}(A_k)}{\mathbb{P}\left(B \mid A_1\right)\mathbb{P}(A_1) + \cdots + \mathbb{P}\left(B \mid A_n\right)\mathbb{P}(A_n)}\,.$$

Théorème de Bayes - 3

On utilise à nouveau le théorème des probabilités composées et l'on a alors

$$\mathbb{P}(A_k \mid B) = \frac{\mathbb{P}\,(B \mid A_k)\,\mathbb{P}(A_k)}{\mathbb{P}\,(B \mid A_1)\,\mathbb{P}(A_1) + \cdots + \mathbb{P}\,(B \mid A_n)\,\mathbb{P}(A_n)}\,.$$

Théorème de Bayes

Soit $(A_1, \cdots, A_n)$ un système complet d'évènements de Ω. Soit B un évènement de Ω de probabilité non nulle. Alors :

$$\mathbb{P}(A_i \mid B) = \frac{\mathbb{P}\,(B \mid A_i)\,\mathbb{P}(A_i)}{\sum_{j=1}^{n} \mathbb{P}\,(B \mid A_j)\,\mathbb{P}(A_j)}\,.$$

Théorème de Bayes - 4

On a notamment le cas particulier avec (A, A^c) si $\mathbb{P}(A) \in]0;1[$:

Théorème de Bayes - 4

On a notamment le cas particulier avec (A, A^c) si $\mathbb{P}(A) \in]0;1[$:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B \mid A)\,\mathbb{P}(A) + \mathbb{P}(B \mid A^c)\,\mathbb{P}(A^c)}\,.$$

Théorème de Bayes - 5

Exercice

Une machine fabrique des composants électroniques. Une proportion $R = 5\%$ de ces composants est hors des tolérances. L'entreprise peut repérer les composants défectueux par un test. On suppose que si le composant est hors des tolérances alors le test est positif dans 95% des cas (sensibilité du test). Si le composant n'est pas défectueux, on suppose que le test est négatif dans 90% des cas (spécificité du test).

Théorème de Bayes - 5

Exercice

Une machine fabrique des composants électroniques. Une proportion $R = 5\%$ de ces composants est hors des tolérances. L'entreprise peut repérer les composants défectueux par un test. On suppose que si le composant est hors des tolérances alors le test est positif dans 95% des cas (sensibilité du test). Si le composant n'est pas défectueux, on suppose que le test est négatif dans 90% des cas (spécificité du test).

On suppose que tous les composants sont testés et que les composants testés positifs sont jetés. Quelle proportion de composants est jetée ? Parmi les composants jetés, quelle est la proportion de composants effectivement hors des tolérances ?

Théorème de Bayes - 5

Exercice

Une machine fabrique des composants électroniques. Une proportion $R = 5\%$ de ces composants est hors des tolérances. L'entreprise peut repérer les composants défectueux par un test. On suppose que si le composant est hors des tolérances alors le test est positif dans 95% des cas (sensibilité du test). Si le composant n'est pas défectueux, on suppose que le test est négatif dans 90% des cas (spécificité du test).

On suppose que tous les composants sont testés et que les composants testés positifs sont jetés. Quelle proportion de composants est jetée ? Parmi les composants jetés, quelle est la proportion de composants effectivement hors des tolérances ?

On suppose maintenant que l'on fait deux tests (indépendamment) et l'on jette les composants testés positifs deux fois. Quelle proportion de composants est jetée ? Parmi les composants jetés, quelle est la proportion de composants effectivement hors des tolérances ?

Théorème de Bayes - 6

Réponses : 14.25% des composants sont jetés si l'on procède à un unique test et seulement un tiers de ces composants est effectivement hors des tolérances. En procédant à deux tests indépendants, on ne jette que 5.4625% des composants et environ 82.6% des composants jetés sont effectivement hors des tolérances.

Théorème de Bayes - 7

Théorème de Bayes - 7

Exemple

Paradoxe de Monty Hall.

Plan

1. Probabilité conditionnelle

2. Indépendance d'évènements
   - Indépendance de deux évènements
   - Indépendance mutuelle de n évènements

Heuristique

Soient A et B deux évènements associés à une même expérience aléatoire e à laquelle on associe l'univers Ω. Rappelons l'approche intuitive des probabilités par les fréquences. La fréquence de réalisation de l'évènement A lorsque l'on répète l'expérience e converge vers $\mathbb{P}(A)$, qui quantifie "les chances de voir A réalisé". Cette approche est validée par les lois des grands nombres.

Heuristique

Soient A et B deux évènements associés à une même expérience aléatoire e à laquelle on associe l'univers Ω. Rappelons l'approche intuitive des probabilités par les fréquences. La fréquence de réalisation de l'évènement A lorsque l'on répète l'expérience e converge vers $\mathbb{P}(A)$, qui quantifie "les chances de voir A réalisé". Cette approche est validée par les lois des grands nombres.

Supposons maintenant qu'on sache que l'évènement B est réalisé. Les chances de voir A réalisé vont *a priori* changer et être quantifiées par le nombre $\mathbb{P}(A \mid B)$. Toutefois, il se peut que le fait de savoir que B est réalisé ne donne aucune information quant à la réalisation de A. On dit alors que A est indépendant de B.

Indépendance de deux évènements - 1

Définition

On dit que A est indépendant de B si l'on a $\mathbb{P}(A \mid B) = \mathbb{P}(A)$.

Indépendance de deux évènements - 1

Définition

On dit que A est indépendant de B si l'on a $\mathbb{P}(A \,|\, B) = \mathbb{P}(A)$.

Exemple

On jette deux dés à six faces. L'un est rouge et l'autre est bleu. On considère les évènements $A :=$"le dé rouge indique une face paire" et $B :=$"le dé bleu indique une face paire". Alors A est indépendant de B : $\mathbb{P}(A \,|\, B) = \mathbb{P}(A) = \frac{1}{2}$. De même, B est indépendant de A.

Indépendance de deux évènements - 1

Définition

On dit que A est indépendant de B si l'on a $\mathbb{P}(A \mid B) = \mathbb{P}(A)$.

Exemple

On jette deux dés à six faces. L'un est rouge et l'autre est bleu. On considère les évènements $A :=$"le dé rouge indique une face paire" et $B :=$"le dé bleu indique une face paire". Alors A est indépendant de B : $\mathbb{P}(A \mid B) = \mathbb{P}(A) = \frac{1}{2}$. De même, B est indépendant de A.

Contre-exemple

On tire deux cartes sans remise dans un jeu de 32 cartes. On considère les évènements $A :=$"la première carte est l'as de cœur" et $B :=$"la deuxième carte est l'as de cœur". Alors A n'est pas indépendant de B : $\mathbb{P}(A \mid B) = 0$ tandis que $\mathbb{P}(A) = \frac{1}{32}$. De même, B n'est pas indépendant de A.

Indépendance de deux évènements - 2

Théorème

La relation d'indépendance est symétrique : si A est indépendant de B alors B est indépendant de A. De plus, A et B sont indépendants si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Indépendance de deux évènements - 2

Théorème

La relation d'indépendance est symétrique : si A est indépendant de B alors B est indépendant de A. De plus, A et B sont indépendants si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

ATTENTION

Deux évènements incompatibles et de probabilités non nulles ne sont pas indépendants. L'indépendance et l'incompatibilité sont deux notions radicalement différentes.

Indépendance de deux évènements - 3

Proposition

On a équivalence entre les quatre assertions suivantes :

- A et B sont indépendants,
- A et B^c sont indépendants,
- A^c et B sont indépendants,
- A^c et B^c sont indépendants.

Approche intuitive - 1

Exemple

Une urne contient quatre jetons : un bleu, un blanc, un rouge et un bleu-blanc-rouge. On en tire un au hasard. Considérons les trois évènements

- $A := \{\text{le jeton tiré contient du bleu}\}$,
- $B := \{\text{le jeton tiré contient du blanc}\}$,
- $C := \{\text{le jeton tiré contient du rouge}\}$.

Il est clair que $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{2}{4} = \frac{1}{2}$.

Approche intuitive - 2

D'autre part, on a

$$\mathbb{P}(A \cap B) = \mathbb{P}\left(\{\text{on tire le jeton tricolore}\}\right) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B)\,.$$

Et, de même $\mathbb{P}(B \cap C) = \frac{1}{4} = \mathbb{P}(B)\mathbb{P}(C)$ et $\mathbb{P}(A \cap C) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(C)$. Ainsi, les évènements A, B et C sont deux à deux indépendants.

Approche intuitive - 2

D'autre part, on a

$$\mathbb{P}(A \cap B) = \mathbb{P}\left(\{\text{on tire le jeton tricolore}\}\right) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B)\,.$$

Et, de même $\mathbb{P}(B \cap C) = \frac{1}{4} = \mathbb{P}(B)\mathbb{P}(C)$ et $\mathbb{P}(A \cap C) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(C)$. Ainsi, les évènements A, B et C sont deux à deux indépendants.

Cependant, $\mathbb{P}(A \mid B \cap C) = 1$ car $B \cap C = \{\text{on tire le jeton tricolore}\}$. Donc la connaissance simultanée de B et C modifie notre information sur A.

Indépendance mutuelle de n évènements

Définition

Soient n évènements $A_1, \cdots, A_n$ associés à une même expérience aléatoire e à laquelle on associe l'univers Ω ; c'est-à-dire que $A_1, \cdots, A_n$ sont des sous-ensembles de Ω. On dit que les évènements $A_1, \cdots, A_n$ sont mutuellement indépendants si pour toute sous-famille $A_{i_1}, A_{i_2}, \cdots, A_{i_k}$ avec $1 \leq i_1 < \cdots < i_k \leq n$ où $k \in [\![1; n]\!]$, on a :

$$\mathbb{P}\left(A_{i_1} \cap \cdots \cap A_{i_k}\right) = \mathbb{P}(A_{i_1}) \times \cdots \times \mathbb{P}(A_{i_k})\,.$$

Indépendance mutuelle de n évènements

Définition

Soient n évènements $A_1, \cdots, A_n$ associés à une même expérience aléatoire e à laquelle on associe l'univers Ω ; c'est-à-dire que $A_1, \cdots, A_n$ sont des sous-ensembles de Ω. On dit que les évènements $A_1, \cdots, A_n$ sont mutuellement indépendants si pour toute sous-famille $A_{i_1}, A_{i_2}, \cdots, A_{i_k}$ avec $1 \leq i_1 < \cdots < i_k \leq n$ où $k \in [\![1; n]\!]$, on a :

$$\mathbb{P}\left(A_{i_1} \cap \cdots \cap A_{i_k}\right) = \mathbb{P}(A_{i_1}) \times \cdots \times \mathbb{P}(A_{i_k}) .$$

L'indépendance mutuelle implique évidemment l'indépendance deux à deux (en considérant toutes les sous-familles de deux évènements) et la réciproque est fausse comme le montre l'Exemple.

Propriétés de l'indépendance mutuelle

Proposition

Soient n évènements $A_1, \cdots, A_n$ associés à une même expérience aléatoire e à laquelle on associe l'univers Ω. On suppose que les évènements $A_1, \cdots, A_n$ sont mutuellement indépendants. Alors toute famille obtenue en remplaçant certains des A_i par leurs complémentaires est encore mutuellement indépendante.

Propriétés de l'indépendance mutuelle

Proposition

Soient n évènements $A_1, \cdots, A_n$ associés à une même expérience aléatoire e à laquelle on associe l'univers Ω. On suppose que les évènements $A_1, \cdots, A_n$ sont mutuellement indépendants. Alors toute famille obtenue en remplaçant certains des A_i par leurs complémentaires est encore mutuellement indépendante.

Corollaire

Soient n évènements $A_1, \cdots, A_n$ associés à une même expérience aléatoire e à laquelle on associe l'univers Ω. On suppose que les évènements $A_1, \cdots, A_n$ sont mutuellement indépendants. Alors, on a

$$\mathbb{P}\left(A_1 \cup \cdots \cup A_n\right) = 1 - \left(1 - \mathbb{P}\left(A_1\right)\right) \times \cdots \times \left(1 - \mathbb{P}\left(A_n\right)\right) .$$