

Probabilité conditionnelle

PST

4 - Probabilité conditionnelle

Résumé du document

Definition

Table des matières

1. Probabilité conditionnelle 2

1.1. Concept 2

1.1.1. Remarques 2

1.2. Théorème de multiplication 2

1.3. Théorème des probabilités totales 2

2. Théorème de Bayes 3

2.1. Version simplifiée 3

2.2. Version composée 3

2.3. Filtre bayésien anti-spam 3

2.3.1. Exemple 3

2.3.2. Formules utiles 4

3. Indépendance 5

1. Probabilité conditionnelle

1.1. Concept

La probabilité conditionnelle nous permet de calculer la probabilité d'un événement en fonction d'une condition.

L'opération permettant de calculer la probabilité conditionnelle est la suivante:

A = probabilité que l'événement A se passe

B = événement qui s'est réalisé

Nous cherchons donc la chance que l'événement A se passe en sachant que l'événement B s'est réalisé:

$$P(A \mid B)$$

La formule de base permettant de calculer cette probabilité est:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

1.1.1. Remarques

$$P(B \mid B) = 1$$

si A est inclus dans B , alors $A \cap B = A$ et donc

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

1.2. Théorème de multiplication

En utilisant l'inverse de la formule présentée au point 2 nous pouvons retrouver $P(A \cap B)$, pour cela nous aurons la formule suivante:

$$\begin{aligned} P(A \cap B) &= P(A \mid B) * P(B) \\ &= P(B \mid A) * P(A) \end{aligned}$$

1.3. Théorème des probabilités totales

Soient A et B deux événements quelconques. Comme B et \overline{B} forment une partition de Ω , on aura selon le théorème des probabilités totales,

$$\begin{aligned} P(A) &= P(A \mid B) * P(B) + P(A \mid \overline{B}) * P(\overline{B}) \\ &= P(A \mid B) * P(B) + P(A \mid \overline{B}) * (1 - P(B)) \end{aligned}$$

2. Théorème de Bayes

Le théorème de Bayes qui fait appel aux théorèmes de multiplication et de probabilités totales est très important. Par exemple, il donna naissance à une autre approche de la statistique. Nous présenterons d'abord la version simple du théorème puis sa version composée.

2.1. Version simplifiée

Supposons que A et B soient deux événements d'un ensemble fondamental Ω , avec $P(B) \neq 0$. Alors,

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

2.2. Version composée

Soient une partition H_1, H_2, \dots, H_k et un événement B d'un ensemble fondamental Ω , avec $P(B) \neq 0$. Pour tout indice $1 \leq j \leq k$, on aura,

$$\begin{aligned} P(H_j | B) &= \frac{P(H_j \cap B)}{P(B)} \\ &= \frac{P(B | H_j) * P(H_j)}{P(B | H_1) * P(H_1) + \dots + P(B | H_k) * P(H_k)} \end{aligned}$$

2.3. Filtre bayésien anti-spam

p_i : probabilité qu'un mot choisi au hasard dans un message électronique est le mot i en sachant que le mot est un spam.

q_i : probabilité qu'un mot choisi au hasard dans un message électronique est le mot i en sachant que le message n'est pas un spam

2.3.1. Exemple

M_i : le mot choisi au hasard dans le message électronique est le mot i ;

S : le message électronique est un spam

Ainsi,

$$p_i = P(M_i | S) \text{ et } q_i = P(M_i | \bar{S})$$

Pour illustrer le fonctionnement du filtre bayésien, supposons que la proportion de messages spam d'une certaine compagnie vaut 0.9 et que pour le mot "hypothèque" noté 1, $p_1 = 0.05$ et $q_1 = 0.001$.

Pour ce mot, on a alors

$$p_1 = P(M_1 | S) = 0.05 \text{ et } q_1 = P(M_1 | \bar{S}) = 0.001$$

Un nouveau message électronique vient d'arriver et le mot "hypothèque" y apparaît exactement une fois. En appliquant le théorème de Bayes, la probabilité que le message électronique soit un spam est

$$P(S | M_1) = \frac{P(M_1 | S) * P(S)}{P(M_1 | S) * P(S) + P(M_1 | \bar{S}) * P(\bar{S})} = 0.998$$

2.3.2. Formules utiles

Dans le cas où les événements M_1 et M_2 sont dit **indépendants** nous aurons,

$$\begin{aligned}
 P(S \mid M_1 \cap M_2) &= \frac{P(M_1 \cap M_2 \mid S) * P(S)}{P(M_1 \cap M_2 \mid S) * P(S) + P(M_1 \cap M_2 \mid \bar{S}) * P(\bar{S})} \\
 &= \frac{P(M_1 \mid S) * P(M_2 \mid S) * P(S)}{P(M_1 \mid S) * P(M_2 \mid S) * P(S) + P(M_1 \mid \bar{S}) * P(M_2 \mid \bar{S}) * P(\bar{S})}
 \end{aligned}$$

3. Indépendance

L'événement A est **indépendant** de l'événement B si le fait de savoir que B s'est déroulé n'influence pas la probabilité de A .

Nous aurons donc

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

Or, par définition d'une probabilité conditionnelle,

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

Ainsi, on obtient

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

On peut donc dire que A est indépendant de B si

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

Deux événements sont dépendants s'ils ne sont pas indépendants.