

Probabilités et Statistiques pour l'Informatique (2)

Akka ZEMMARI
LaBRI - Université de Bordeaux

Indépendance

Définitions

- Deux événements A et B sont dit *indépendants*, si :

$$Pr(A \cap B) = Pr(A)Pr(B).$$

- Une famille $A_i, i = 1, \dots, n$, d'événements est dite *indépendante dans son ensemble*, si pour tout sous-ensemble $J \subset \{1, \dots, n\}$:

$$Pr\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} Pr(A_i).$$

Remarque. Pour démontrer l'indépendance d'une famille $A_i, i = 1, \dots, n$, il ne suffit pas de prouver que les A_i sont 2-à-2 indépendants, comme le montre l'exemple suivant.

Exemple. Soit $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$ l'espace associé au tirage d'un dé à 8 faces avec la probabilité uniforme $\frac{1}{8}$. Soient :

$$A_1 = \{1, 2, 3, 4\}$$

$$A_2 = \{1, 2, 5, 6\}$$

$$A_3 = \{3, 4, 5, 6\}$$

On a :

$$Pr(A_1 \cap A_2) = Pr(\{1, 2\}) = \frac{1}{4} = Pr(A_1)Pr(A_2)$$

$$Pr(A_1 \cap A_3) = Pr(\{3, 4\}) = \frac{1}{4} = Pr(A_1)Pr(A_3)$$

$$Pr(A_2 \cap A_3) = Pr(\{5, 6\}) = \frac{1}{4} = Pr(A_2)Pr(A_3)$$

Or :

$$Pr(A_1 \cap A_2 \cap A_3) = Pr(\emptyset) = 0 \neq Pr(A_1)Pr(A_2)Pr(A_3) = \frac{1}{8}.$$

Probabilités Conditionnelles

Il arrive que la donnée d'un événement modifie la probabilité initialement affectée à un autre. La notion de **probabilité conditionnelle** est introduite afin de formaliser le concept de la probabilité de l'occurrence d'un événement B sachant qu'un autre A s'est produit.

Notons tout de suite que B conditionné par A *n'implique pas* que A précède nécessairement B (dans l'ordre chronologique).

Exemple. Un dé uniforme est lancé et nous savons que le point obtenu est **pair** (événement A). Quelle est la probabilité pour que le point **soit au moins 4** (événement B) ?

Il est clair que le nouvel espace (conditionné par l'événement A) sur lequel les événements élémentaires sont à définir est $\Omega' = A$, qui contient 3 éléments. Or, la portion de A qui est en même temps favorable à B en contient 2. Il est donc raisonnable de définir la probabilité de B conditionné par A par le ratio $\frac{2}{3}$ (à comparer avec $Pr(B) = \frac{1}{2}$).

Définitions. Soit (Ω, Pr) un espace probabilisé discret et soit A un événement de probabilité non nulle. On définit sur $\mathcal{P}(\Omega)$, l'application $Pr(. / A)$ à valeurs dans $[0, 1]$ par :

$$Pr(B/A) = \frac{Pr(A \cap B)}{Pr(A)}, \quad \forall B \in \mathcal{P}(\Omega).$$

On appelle $Pr(B/A)$ *probabilité conditionnelle de B sachant A* (ou *probabilité de B sachant A*).

Nous avons alors :

$$Pr(A \cap B) = Pr(B/A)Pr(A).$$

Nous disons alors que l'on a *effectué un conditionnement par A* .

Dans le cas où $Pr(A) \neq 0$, l'indépendance de A et B est équivalente à $Pr(B/A) = Pr(B)$; ce qui est en accord avec sa définition intuitive.

Proposition. Le couple $(A, Pr(. / A))$ est un espace probabilisé discret.

Proposition. Soit $A_1, \dots, A_n$ une partition de Ω . Si chacun de ces ensembles est de probabilité non nulle, alors :

$$Pr(B) = \sum_{i=1}^n Pr(A_i)Pr(B/A_i), \quad \forall B \in \mathcal{P}(\Omega).$$

Exemple. On lance deux dés uniformes.

- Quelle est la probabilité d'avoir obtenu un double sachant que la somme des points vaut 8 ?
- Quelle est la probabilité d'avoir obtenu un double sachant que la somme des points vaut au moins 10 ?

Solution : Chacune de ces probabilités conditionnelles est définie par une fraction dont le numérateur est la probabilité de l'intersection et le dénominateur celle de la condition. On trouve facilement :

- $Pr(\text{double}/S = 8) = \frac{1/36}{5/36} = \frac{1}{5}.$
- $Pr(\text{double}/S \geq 10) = \frac{2/36}{6/36} = \frac{1}{3}.$

Exemple 2. Le jeu du Monty Hall s'énonce comme suit :

Le jeu oppose un présentateur à un candidat (le joueur). Ce joueur est placé devant trois portes fermées. Derrière l'une d'elles se trouve une voiture et derrière chacune des deux autres se trouve une chèvre. Il doit tout d'abord désigner une porte. Puis le présentateur doit ouvrir une porte qui n'est ni celle choisie par le candidat, ni celle cachant la voiture (le présentateur sait quelle est la bonne porte dès le début). Le candidat a alors le droit ou bien d'ouvrir la porte qu'il a choisie initialement, ou bien d'ouvrir la troisième porte.

- Quelle est la meilleure stratégie à adopter pour augmenter la probabilité de gain de la voiture

Probabilité des Causes

Soit l'événement B causé par l'un des événements $A_1, A_2, \dots, A_n$, tous de probabilité non nulle. On connaît les probabilités $Pr(A_i)$ de ces derniers événements et aussi les probabilités conditionnelles $Pr(B/A_i)$. Comment trouver les probabilités des causes sachant que B s'est produit, i.e. les probabilités $Pr(A_i/B)$?

Théorème (Formule de Bayes). Soit $A_1, \dots, A_n$ une partition de Ω telle que $Pr(A_i) \neq 0, \forall i = 1, \dots, n$. Soit B un événement de probabilité non nulle. Nous avons :

$$Pr(A_i/B) = \frac{Pr(A_i)Pr(B/A_i)}{\sum_{j=1}^n Pr(A_j)Pr(B/A_j)}, \quad i = 1, 2, \dots, n$$

En particulier, pour deux événements A et B de probabilité non nulle, nous avons :

$$Pr(A/B) = \frac{Pr(A)Pr(B/A)}{Pr(B)}.$$

- $Pr(A)$: probabilité *a priori* de A
- $Pr(A/B)$: probabilité *a posteriori* de A

Exemple (F. Dress). Dans un jeu télévisé, un candidat doit choisir une question de remplacement en tirant au hasard parmi 3 papiers. Il y a :

- une question facile avec 3 chances sur 4 de donner la réponse correcte,
- une question moyenne avec 2 chances sur 5 de donner la réponse correcte, et
- une question difficile avec 1 chance sur 5 de donner la réponse correcte.

Sachant que le candidat a donné la réponse **correcte**, quelle est la probabilité pour qu'il s'agisse de la question **facile** ?

Solution : Désignons par F , M et D les événements tirage de la question facile, de la question moyenne et de la question difficile respectivement. Soit, par ailleurs, l'événement "réponse correcte" noté par C .

D'après la formule de Bayes, nous avons :

$$Pr(F/C) = \frac{Pr(F)Pr(C/F)}{Pr(C/F)Pr(F) + Pr(C/M)Pr(M) + Pr(C/D)Pr(D)},$$

ce qui vaut :

$$\frac{\frac{1}{3} \cdot \frac{3}{4}}{\frac{1}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{2}{5} + \frac{1}{3} \cdot \frac{1}{5}} = \frac{5}{9}.$$

Loi Conjointe

Dans la suite, nous considérons un couple de v.a., mais l'étude s'étend à un nombre quelconque de v.a.

Définitions. Soient X et Y deux v.a. réelles sur le même espace Ω . Le couple (X, Y) peut être considéré comme un **vecteur** aléatoire sur Ω . Sa loi est appelée *loi conjointe* de (X, Y) . La connaissance de cette dernière permet de retrouver la loi de X et celle de Y individuellement. En effet, pour tout ensemble $A \subset X(\Omega)$ et tout ensemble $B \subset Y(\Omega)$:

$$Pr_1(X \in A) = Pr(X \in A, Y \in \mathbb{R})$$

et

$$Pr_2(Y \in B) = Pr(X \in \mathbb{R}, Y \in B).$$

X et Y sont dites *indépendantes*, si pour tout ensemble $A \subset X(\Omega)$ et tout ensemble $B \subset Y(\Omega)$:

$$Pr(X \in A, Y \in B) = Pr_1(X \in A)Pr_2(Y \in B).$$

Les distributions Pr_1 et Pr_2 sont appelées *lois marginales* du vecteur (X, Y) . Dans la suite, nous supprimons les indices de Pr lorsqu'il n'y a pas le danger de confusion. La connaissance des lois marginales **ne permet pas** de reconstruire la loi conjointe.

Exemple. On lance une pièce authentique une première fois. Pour le 2e lancer, on considère deux possibilités :

- A.** On reprend la pièce et on la repose sur la même face.
- B.** On la relance indépendamment du premier lancer.

Les lois conjointes sont données par :

A. $Pr(\{F, F\}) = Pr(\{P, P\}) = \frac{1}{2}$ et $Pr(\{F, P\}) = Pr(\{P, F\}) = 0$.

B. $Pr(\{F, F\}) = Pr(\{P, F\}) = Pr(\{F, P\}) = Pr(\{P, P\}) = \frac{1}{4}$.

Il s'agit donc de deux lois distinctes

Les lois marginales dans les deux cas sont données par :

$$Pr(\{P\}) = Pr(\{F\}) = \frac{1}{2}.$$

Pour deux lois conjointes distinctes, on a les mêmes lois marginales.

Espérance Conditionnelle

Soit X une v.a. et soit A un événement de probabilité non nulle.
L'espérance de X sachant A est définie par :

$$\mathbb{E}(X/A) = \mathbb{E}_{Pr(. / A)} X$$

ou :

$$\mathbb{E}(X/A) = \frac{1}{Pr(A)} \sum_{\omega \in A} X(\omega) Pr(\{\omega\}).$$

Exemple. Quelle est l'espérance d'un dé uniforme sachant que le nombre sorti est inférieur ou égal à 3 ?

Solution : Soit X la v.a. associée au lancer du dé. L'événement A associé à la condition est l'ensemble $\{1, 2, 3\}$. La probabilité de la condition A est $\frac{1}{2}$. L'espérance conditionnelle recherchée vaut :

$$\mathbb{E}(X/A) = \frac{1}{\frac{1}{2}} \left[\frac{1}{6} + \frac{2}{6} + \frac{3}{6} \right] = 2.$$

Exemple (F. Dress)

Une urne contient a boules rouges et b boules vertes. On tire une boule au hasard puis on la remet dans l'urne en ajoutant c boules de la même couleur ; on tire enfin une boule de l'urne dans sa nouvelle composition.

On appelle X_i , $i = 1, 2$ la variable de Bernoulli égale à 1 si la boule tirée au i -ème tirage est rouge, 0 si elle est verte.

Dresser le tableau de la loi conjointe de X_1 et X_2 , et des lois marginales. Calculer les espérances et les variances.

Loi des Grands Nombres

Il arrive souvent que l'on répète un grand nombre de fois une épreuve de façon indépendante. Soit A un événement qui peut se réaliser avec une probabilité p . Que peut-on dire du lien entre la fréquence observée de A et sa probabilité p ? Il arrive que p soit inconnue ; peut-on en faire une approximation fondée sur la fréquence de A ?

Avant de répondre à ces questions, considérons le cas d'une v.a. Sa variance étant introduite pour mesurer sa dispersion autour de l'espérance. Peut-on, en termes de ce paramètre, confirmer avec une confiance relativement grande, que l'écart entre l'espérance et la valeur prise par la v.a. ne sera pas trop grand ?

Proposition (Inégalité de Bienaymé-Tchebychev). Soit X un v.a. d'espérance μ et de variance V . Alors, pour tout $\epsilon > 0$:

$$Pr(|X - \mu| \geq \epsilon) \leq \frac{V}{\epsilon^2}.$$

Exemple. Soit X_n une v.a. binomiale de paramètres $p = \frac{1}{10}$ et $n \in \mathbb{N}$.

- Peut-on confirmer, que pour $n = 100$ l'écart entre $\frac{X_n}{n}$ et son espérance $\frac{1}{10}$ ne dépasse pas 0.1 avec une probabilité supérieure à 0.9 ?
- Pour quelles valeurs de n , peut-on confirmer, qu'avec une probabilité ≥ 0.99 , cet écart ne dépasse pas 0.05 ?

Solution

- La variance de $\frac{X_n}{n}$ pour $n = 100$ vaut :

$$npq \times \frac{1}{n^2} = \frac{9}{10000}.$$

D'après l'inégalité de Bienaymé-Tchebychev, nous avons :

$$Pr \left(\left| \frac{X_{100}}{100} - \frac{1}{10} \right| \geq 0.1 \right) \leq \frac{9}{10000} \times 100 = 0.09 < 0.1.$$

La réponse est donc positive.

- Nous avons :

$$Pr \left(\left| \frac{X_n}{n} - \frac{1}{10} \right| \geq 0.05 \right) \leq \frac{9}{100} \times \frac{1}{n} \times \frac{1}{0.05^2}.$$

D'où, pour que la probabilité recherchée soit au moins égale à

0.99 il suffit que n soit assez grand pour qu'on ait $\frac{9}{100} \times \frac{1}{n} \times \frac{1}{0.0025} \leq 0.01$

c'est-à-dire $n \geq 3600$.

Théorème (Loi faible des grands nombres). Soit X une v.a. admettant l'espérance μ et la variance V et soit $X_1, X_2, \dots, X_n$ une suite de v.a. indépendantes chacune suivant la même loi que X . Désignons par S_n la valeur moyenne de la suite :

$$S_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Alors S_n converge en probabilité vers son espérance μ ; i.e. pour tout $\epsilon > 0$, on a :

$$\lim_{n \rightarrow \infty} \Pr(|S_n - \mu| \geq \epsilon) = 0.$$

Exemple. On lance une paire de dés authentiques n fois et l'on calcule la moyenne arithmétique des produits des deux points. Nous avons vu que l'espérance du produit vaut $\frac{49}{4}$. Puisque cette v.a. admet une variance, la loi des grands nombre permet de confirmer que la moyenne des produits converge en probabilité vers $\frac{49}{4}$.