

Probabilités, Statistiques, Combinatoire

Philippe Duchon

Université de Bordeaux – Licence Informatique

<http://www.labri.fr/perso/duchon/Enseignements/Probas/>

2019-2020

Jusqu'ici. . .

- ▶ Jusqu'à présent, on a vu comment modéliser une expérience, et (seulement) comment calculer une probabilité, ou une probabilité conditionnelle.
- ▶ Très souvent, dans une expérience aléatoire, il y a une ou plusieurs grandeurs numériques (des nombres aléatoires) dont la ou les valeurs nous intéressent.
- ▶ **Exemple** : On lance deux dés ; on s'intéresse à la **somme** des résultats des deux dés.
- ▶ **Exemple** : On lance 5 dés ; on s'intéresse au **nombre de dés qui donnent le résultat 6**.
- ▶ **Exemple** : On tire 10 fois consécutivement à pile ou face ; on s'intéresse à *plusieurs* nombres :
 - ▶ le nombre de Face ;
 - ▶ le nombre Pile qu'on obtient avant le premier Face (s'il y en a)
 - ▶ le nombre maximum de Pile consécutifs, par forcément avant le premier Face.

Jusqu'ici (suite)

- ▶ **Exemple** : On choisit un individu dans une population donnée, et on s'intéresse à son *âge*, sa *taille*, son *nombre de frères et sœurs*.
- ▶ Dans tous ces exemples, on a une **grandeur numérique** (souvent un entier, mais pas toujours) qui **dépend** du résultat de l'expérience.

Somme de deux dés

On prend l'expérience consistant à lancer deux dés, qu'on modélise par l'ensemble $\Omega = [[1, 6]]^2$ avec la loi de probabilités uniforme ; et on s'intéresse à la somme des deux dés.

On peut ainsi définir 11 événements pertinents :

- ▶ S_2 : “la somme vaut 2”, $S_2 = \{(1, 1)\}$; sa probabilité vaut $1/36$.
- ▶ S_3 : “la somme vaut 3”, $S_3 = \{(1, 2), (2, 1)\}$; sa probabilité vaut $2/36$.
- ▶ ... S_7 : “la somme vaut 7”,
 $S_7 = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$; sa probabilité vaut $6/36$.
- ▶ ... S_{12} : “la somme vaut 12”, $S_{12} = \{(6, 6)\}$; sa probabilité vaut $1/36$.
- ▶ On a ainsi 11 événements deux à deux disjoints (la somme des deux dés ne peut pas prendre deux valeurs à la fois), et dont l'union représente l'univers tout entier (la somme des deux dés est forcément un nombre entier entre 2 et 12).

Notion de “variable aléatoire”

- ▶ La notion de variable aléatoire recouvre *toute quantité numérique que l'on peut déterminer en connaissant entièrement le résultat d'une expérience.*
- ▶ **Exemple 1 :** dans l'expérience consistant à lancer deux dés, on peut définir tout un tas de variables aléatoires :
 - ▶ S : somme des résultats des dés (entier entre 2 et 12)
 - ▶ D : différence entre le résultat du premier dé et celui du second (entier entre -5 et 5)
 - ▶ I : nombre de dés donnant un résultat impair (0, 1 ou 2)
 - ▶ M : maximum des résultats des deux dés (entier entre 1 et 6)
 - ▶ B : “ π si on obtient (1, 1), 0 sinon”
 - ▶ (et littéralement une infinité d'autres possibilités)
- ▶ **Exemple 2 :** dans l'expérience consistant à tirer au dé de manière répétée jusqu'à obtenir un 6, on peut définir aussi tout un tas de variables aléatoires :
 - ▶ N : nombre de tirages nécessaires pour obtenir un 6
 - ▶ S : somme des résultats obtenus avant de s'arrêter, sans compter le 6 final
 - ▶ N_3 : nombre de fois où on obtient 3 avant d'obtenir un 6

Variable aléatoire

- ▶ **Quelle condition** faut-il vérifier pour être sûr d'avoir une variable aléatoire ?
- ▶ À notre niveau, *pas grand chose* : on veut seulement que la valeur du “nombre aléatoire” soit **entièrement déterminée par le résultat de l'expérience**.
- ▶ C'est ce qui est capturé par la notion usuelle (mathématique) de **fonction** définie sur l'univers Ω .
- ▶ **Définition** : si $(\Omega, \mathbb{P})$ est un espace de probabilités, une *variable aléatoire* sur Ω , est n'importe quelle fonction $X : \Omega \rightarrow \mathbb{R}$.
- ▶ (Dans un cours plus complet de probabilités, il faudrait que X soit une fonction “mesurable” ; dans notre cadre, où Ω est fini ou dénombrable, on peut considérer que toute fonction est mesurable)
- ▶ **Note** : la terminologie est trompeuse : une **variable aléatoire** n'est **pas** une variable, c'est plutôt une fonction ; et elle n'est pas vraiment aléatoire (ne dépend pas de la loi de probabilités)...

Conventions

- ▶ On parle de “variable aléatoire à valeurs dans E ” pour dire que la variable aléatoire ne prend que des valeurs dans l'ensemble E ; le cas le plus général est $E = \mathbb{R}$, mais le plus souvent on prendra $E = \mathbb{N}$ ou $E = \mathbb{Z}$ (“variable aléatoire à valeurs entières”).
- ▶ De même qu'on prend l'habitude de nommer les événements par des lettres majuscules ($A, B, C \dots$), on utilise aussi généralement des lettres majuscules pour désigner des variables aléatoires, mais plutôt avec la fin de l'alphabet : $X, Y \dots$
- ▶ Éventuellement, on ajoute des indices : $X_1, X_2 \dots$ de peur de manquer de lettres.
- ▶ Et de même que souvent on ne décrit pas forcément les éléments de l'univers de probabilité, très souvent on n'écrit pas la “variable” (élément de Ω) quand on décrit une variable aléatoire, pas plus qu'on ne le fait pour un événement.
- ▶ (Sauf dans les diapos qui suivent, pour donner des exemples)

Exemples

- ▶ Avec une expérience “lancer de deux dés” modélisée par un espace $\Omega = [[1, 6]]^2$, on peut définir des variables aléatoires :
 - ▶ résultat du premier dé : $X_1(x, y) = x$;
 - ▶ résultat du second dé : $X_2(x, y) = y$;
 - ▶ somme des deux dés : $S(x, y) = x + y = X_1(x, y) + X_2(x, y)$;
ce qu'on note plus simplement : $S = X + Y$
 - ▶ nombre de dés donnant un résultat pair :
 $P(x, y) = [[x \in \{2, 4, 6\}]] + [[y \in \{2, 4, 6\}]]$ (**Notation :**
[[cond]] = 1 si *cond* est vraie, 0 sinon)

Combinaison de variables aléatoires

- ▶ **Remarque fondamentale** : on peut **combinaison** les variables aléatoires pour en construire de nouvelles. . .
 - ▶ $S = X_1 + X_2$
 - ▶ $D = X_1 - X_2$
 - ▶ $Y = f(X)$, si X est une variable à valeurs dans $\mathbb{R}$ et f est une fonction définie sur $\mathbb{R}$, à valeurs dans $\mathbb{R}$; $Z = f(X, Y)$ si X et Y sont deux variables aléatoires, et f une fonction de **deux** variables.

Variables aléatoires particulières

On suppose que Ω est défini.

- ▶ **constantes** : on peut définir la “fonction 1” comme étant la fonction constante qui vaut toujours 1 : pour tout $t \in \Omega$, $\mathbf{1}(t) = 1$. (On peut considérer n'importe quelle constante comme une variable aléatoire d'un genre particulier)
- ▶ **Variable indicatrice d'un événement** : si $A \subset \Omega$ est un événement, on peut définir la variable aléatoire $\mathbf{1}_A$: pour tout $t \in \Omega$, $\mathbf{1}_A(t) = [[t \in A]]$ (vaut 1 sur A , 0 ailleurs)
- ▶ On peut alors exprimer $\mathbf{1}_{A \cap B}$ ou (plus difficile) $\mathbf{1}_{A \cup B}$ en fonction de $\mathbf{1}_A$ et de $\mathbf{1}_B$. . . (solution diapo suivante)

Solutions

- ▶ $\mathbf{1}_{A \cap B}$, c'est la fonction qui vaut 1 sur $A \cap B$ et 0 ailleurs ; en d'autres termes, la fonction qui vaut 1 lorsque $\mathbf{1}_A$ et $\mathbf{1}_B$ valent 1, et 0 sinon :

$$\mathbf{1}_{A \cap B} = \mathbf{1}_A \cdot \mathbf{1}_B$$

(produit de fonctions ; on utilise le fait que $1 \times 1 = 1$ et $0 \times 0 = 0$)

- ▶ $\mathbf{1}_{A \cup B}$ peut s'écrire $\mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_{A \cap B}$ (si on oublie le dernier terme, on obtient 2 sur $A \cap B$) ; ce qui peut se réécrire en

$$\begin{aligned}\mathbf{1}_{A \cup B} &= \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A \cdot \mathbf{1}_B \\ &= 1 - (1 - \mathbf{1}_A)(1 - \mathbf{1}_B)\end{aligned}$$

À partir d'une variable aléatoire

(Le retour des probabilités !)

- ▶ Une fois définis un univers de probabilités $(\Omega, \mathbb{P})$ et une variable aléatoire X sur Ω , que peut-on calculer comme probabilités ?
- ▶ Une notation comme $\mathbb{P}(X)$ **n'a aucun sens** : les seules choses dont la probabilité est définie, ce sont les événements (les parties de Ω) ; or X n'est pas une partie de Ω .
- ▶ **En revanche**, on peut utiliser X pour définir tout un tas d'événements, dont on pourra calculer la probabilité. . .
- ▶ Si a est une valeur possible de X , on peut définir l'événement

$$\{t \in \Omega : X(t) = a\} = X^{-1}(\{a\})$$

(c'est l'ensemble des $t \in \Omega$ pour lesquels X prend la valeur a)

- ▶ Plus généralement, si A est une partie de l'ensemble des valeurs possibles de X , on peut définir

$$\{t \in \Omega : X(t) \in A\} = X^{-1}(A)$$

(ensemble des $t \in \Omega$ pour lesquels X prend une valeur dans A)

Notations (importantes)

On suppose que X est une variable aléatoire déjà définie sur notre univers de probabilité Ω . a désigne une valeur possible pour X , et A un ensemble de valeurs possible.

- ▶ On note $\{X = a\}$, l'événement

$$\{t \in \Omega : X(t) = a\}$$

- ▶ On note $\{X \in A\}$, l'événement

$$\{t \in \Omega : X(t) \in A\}$$

- ▶ Il est important de comprendre qu'il s'agit **d'événements** : dans une expérience, cela peut se produire (être vrai) ou non.
- ▶ Dans la notation d'une probabilité, on s'autorise à ne pas écrire les accolades : on écrira $\mathbb{P}(X = 1)$ au lieu de $\mathbb{P}(\{X = 1\})$ par exemple.
- ▶ Répétons-le : la notation $\mathbb{P}(X)$ n'a pas de sens ; seuls les événements ont une probabilité, et une variable aléatoire n'est pas un événement.

Loi d'une variable aléatoire

Définition/Théorème (loi d'une variable aléatoire)

Soit $(\Omega, \mathbb{P})$ un espace de probabilités, et X une variable aléatoire sur Ω , à valeur dans E . La **loi** de X est par définition la loi de probabilités sur E , notée parfois $\mathbb{P}_X$, définie par : pour toute partie $A \subset E$,

$$\mathbb{P}_X(A) = \mathbb{P}(\{t \in \Omega : X(t) \in A\}).$$

(**c'est** une loi de probabilités sur E)

Digérons cette définition. . .

- ▶ L'ensemble $\{t \in \Omega : X(t) \in A\}$ est une partie de Ω : c'est bien un événement sur Ω
- ▶ On le note souvent $X^{-1}(A)$ **Attention** : c'est une notation, il n'existe pas de fonction X^{-1} .
- ▶ On le note également $(X \in A)$ (on fait disparaître la variable t de l'écriture ; **l'espace de probabilités n'est pas important, on est plus intéressé par les lois des variables aléatoires qu'on y définit**). Notamment, c'est ce qui se cache derrière la notation $\mathbb{P}(X \in A)$.
- ▶ Quand $A = \{a\}$ (A est un singleton), on accepte pour le même événement, la notation $(X = a)$; encore une fois, cette notation est juste une façon commode de noter un événement sur Ω .
- ▶ **Remarque** : la variable aléatoire X ne dépend pas de la loi de probabilités $\mathbb{P}$ qu'on met sur Ω ; sa loi $\mathbb{P}_X$, elle, en dépend !
- ▶ On préfère généralement la notation $\mathbb{P}(X \in A)$ qui ne risque pas d'être confondue avec une probabilité conditionnelle.

Identifier la loi d'une variable aléatoire

Très souvent, on définit une variable aléatoire *via* une description d'expérience. Mais ce qui nous intéresse pour faire des calculs de probabilités, c'est d'identifier sa loi, qui permet (au moins en théorie) de “tout” calculer.

Déterminer la loi d'une variable aléatoire X , cela veut généralement dire

- ▶ décrire l'ensemble des valeurs possibles pour la variable X (formellement : décrire l'ensemble $X(\Omega)$).
- ▶ pour chaque valeur possible a , calculer la probabilité que la variable X prenne la valeur a (calculer la probabilité de l'événement $\{X = a\}$)
- ▶ parfois on a de la chance, et on identifie la loi comme étant une “loi connue” (qui porte un nom, parfois avec un ou plusieurs paramètres).

Un exemple de calcul de loi

Prenons une expérience consistant à lancer deux dés équilibrés : on peut modéliser ça par l'espace $\Omega = [[1, 6]]^2$, avec la probabilité uniforme. Calculons la loi de la variable aléatoire “maximum des deux dés”.

- ▶ On définit, pour $(x, y) \in \Omega$, $M(x, y) = \max(x, y)$, à valeur dans $E = \{1, 2, 3, 4, 5, 6\}$.
- ▶ Pour décrire la loi de M , il suffit de décrire la probabilité $\mathbb{P}_M$ pour les singletons de E , i.e., $\mathbb{P}(M = i)$ pour chaque $i \in E$:
- ▶ pour $i = 1$, c'est facile, $(M = 1) = \{(1, 1)\}$ et $\mathbb{P}(M = 1) = 1/36$;
- ▶ pour $i = 2$: $(M = 2) = \{(1, 2), (2, 1), (2, 2)\}$ et $\mathbb{P}(M = 2) = 3/36$;
- ▶ pour $i = 3$: $(M = 3) = \{(1, 3), (2, 3), (3, 1), (3, 2), (3, 3)\}$ et $\mathbb{P}(M = 3) = 5/36$;
- ▶ $\mathbb{P}(M = 4) = 7/36$, $\mathbb{P}(M = 5) = 9/36$ et $\mathbb{P}(M = 6) = 11/36$; on peut vérifier que la somme fait bien 1.

Loi jointe de deux variables

- ▶ Si X et Y sont des v.a. à valeurs dans $\mathbb{R}$, on peut considérer le couple (X, Y) comme une v.a. à valeurs dans $\mathbb{R}^2$; c'est la fonction définie par : pour tout $t \in \Omega$,

$$(X, Y)(t) = (X(t), Y(t))$$

- ▶ Ce qu'on appelle la *loi jointe* de X et Y , c'est la loi du couple (X, Y)
- ▶ Là où la loi de X est définie par la donnée des $\mathbb{P}(X = a)$ pour les différentes valeurs de a , la loi de (X, Y) va être définie par la donnée des $\mathbb{P}(X = a, Y = b)$ pour les différentes valeurs du **couple** (a, b)
- ▶ **Indépendance** (de deux v.a.) : X et Y sont dites indépendantes, si on a, pour tout a et pour tout b ,

$$\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b)$$