Probabilités, Statistiques, Combinatoire, CM 10

Adrian Tanasa

Université de Bordeaux – Licence Informatique

Notions vu le dernier cours - variables aléatoires

- expérience de la somme de 2 dés
- définition
- exemples
- variables aléatoires particulières
- loi d'une variables aléatoires ; exemples

Plan du cours d'aujourd'hui

variable aléatoire = v. a.

- loi jointe de deux v. a.
- indépendance de v. a.
- description d'une v. a. en fonction d'autres v. a.
- lois de Bernoulli, géométrique, binomiale, Poisson

Loi jointe de deux variables

- Soit X et Y des v.a. à valeurs dans E.
  On peut considérer le couple (X, Y) comme une v.a. à valeurs dans E^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)

- 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 v. a.

- **Indépendance** de deux v.a. : X et Y sont dites indépendantes, si on a, pour tout a et pour tout b,

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

- **Conséquence :** si X et Y sont indépendantes, tout événement $\{X \in A\}$ est indépendant de tout événement $\{Y \in B\}$ (même si A et B ne sont pas réduits à un seul élément chacun)

- Inversement, pour montrer que deux variables X et Y ne sont **pas** indépendantes, il suffit de trouver deux valeurs a et b telles que $\{X = a\}$ et $\{Y = b\}$ ne soient pas indépendants.

- **Cas particulier :** si on a un a et un b tels que $\mathbb{P}(X = a) > 0$, $\mathbb{P}(Y = b) > 0$, mais $\mathbb{P}(X = a \text{ et } Y = b) = 0$, X et Y ne sont pas indépendantes.

Description d'une variable en fonction d'autres

- **Situation fréquente :** on a une variable aléatoire X dont on connaît la loi, et une fonction f ; on définit une nouvelle variable Z par $Z = f(X)$.

- **Problème :** décrire la loi de Z.

Exemple avec une variable

- X uniforme sur $[[1, 6]]$; $Z = X^2 - 3X + 2$.
- Ici, $f(x) = x^2 - 3x + 2$.
- Les valeurs possibles de X sont les entiers de 1 à 6 ; donc les valeurs possibles de Z sont $f(1)$, $f(2)$, $f(3)$, $f(4)$, $f(5)$ et $f(6)$, soit : $\{0, 2, 6, 12, 20\}$.
- $\{Z = 20\} = \{X = 6\}$: $\mathbb{P}(Z = 20) = 1/6$.
- $\{Z = 12\} = \{X = 5\}$: $\mathbb{P}(Z = 12) = 1/6$.
- $\{Z = 6\} = \{X = 4\}$: $\mathbb{P}(Z = 6) = 1/6$.
- $\{Z = 2\} = \{X = 3\}$: $\mathbb{P}(Z = 2) = 1/6$.
- $\{Z = 0\} = \{X = 1\} \cup \{X = 2\}$: $\mathbb{P}(Z = 0) = 1/3$.

Exemple avec deux variables : somme

- X et Y deux variables aléatoires, **indépendantes**.
- $\mathbb{P}(X = 0) = 1/4$, $\mathbb{P}(X = 1) = 1/2$, $\mathbb{P}(X = 2) = 1/4$.
- $\mathbb{P}(Y = 0) = 1/2$, $\mathbb{P}(Y = 1) = 1/2$.
- Indépendantes, donc on a la loi jointe : $\mathbb{P}(X = i, Y = j) = \mathbb{P}(X = i)\mathbb{P}(Y = j)$.
- On pose $Z = X + Y$; les valeurs possibles de Z sont 0, 1, 2, 3.
- Pour décrire la loi de Z, on doit calculer $\mathbb{P}(Z = k)$, pour $k \in \{0, 1, 2, 3\}$
- $\{Z = 0\} = \{X = 0\} \cap \{Y = 0\}$: $\mathbb{P}(Z = 0) = 1/4 \cdot 1/2 = 1/8$.
- $\{Z = 1\} = (\{X = 0\} \cap \{Y = 1\}) \cup (\{X = 1\} \cap \{Y = 0\})$: $\mathbb{P}(Z = 1) = 1/4 \cdot 1/2 + 1/2 \cdot 1/2 = 3/8$.
- $\{Z = 2\} = (\{X = 1\} \cap \{Y = 1\}) \cup (\{X = 2\} \cap \{Y = 0\})$: $\mathbb{P}(Z = 2) = 1/2 \cdot 1/2 + 1/4 \cdot 1/2 = 3/8$.
- $\{Z = 3\} = \{X = 2\} \cap \{Y = 1\}$: $\mathbb{P}(Z = 3) = 1/4 \cdot 1/2 = 1/8$.
- On peut vérifier que les probabilités se somment à 1, **ou** on peut calculer l'une d'entre elles à partir des trois autres.

Loi d'une variable aléatoire

- Rappel : déterminer la **loi** d'une variable aléatoire X, cela revient à déterminer :
  1. l'ensemble V des valeurs que peut prendre X
  2. pour chaque $v \in V$, la probabilité de l'événement $(X = v)$:

$$p(v) = \mathbb{P}(X = v)$$

- Parmi l'infinie variété de lois possibles, il y en a certaines qui reviennent très souvent, parce qu'elles sont utiles pour modéliser beaucoup de situations

Lois de Bernoulli

- La variable aléatoire la plus simple possible à part les variables constantes : on a un événement A, et on pose $X = \mathbf{1}_A$ (indicatrice de A).
- Une telle variable a deux valeurs possibles : 1 (pour $t \in A$), et 0 (pour $t \notin A$).
- La loi est tout aussi simple à décrire : si on pose $p = \mathbb{P}(A)$, on a $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p$.
- Le **paramètre** p $(0 \leq p \leq 1)$ détermine complètement une loi de Bernoulli.
- **Terminologie :** "X suit la loi de Bernoulli de paramètre p" ; on dit aussi "X est une Bernoulli de paramètre p".

Lois binomiales

- situation classique : on a un nombre **fixe** d'événements **indépendants**, chacun ayant la **même probabilité** de se produire ; on s'intéresse au **nombre** de ceux qui se produisent.
- **Exemple :** je lance 3 dés, et je marque 1 point par dé qui donne un résultat pair ; loi du nombre de points ?
- **Exemple :** dans un jeu de 32 cartes, dont j'ai séparé les 8 cartes de chaque couleur, je tire aléatoirement une carte de chaque couleur ; loi du nombre d'As obtenus ?
- **Exemple :** je vais Place de la Victoire un jeudi soir, et je questionne 10 personnes "au hasard" ; je compte le nombre d'étudiants de l'université parmi elles.
- *A priori*, il y a deux paramètres importants : le *nombre* d'événements, et la probabilité (commune) de chacun d'eux ; (n, p)
- C'est la loi de la somme de n Bernoulli indépendantes de même paramètre p.

Détermination de la loi binomiale

- Les valeurs possibles : tous les entiers de 0 à n : $V = [[0, n]]$.
- Il faut donc déterminer, pour chaque $k \in [[0, n]]$, la probabilité $\mathbb{P}(X = k)$.
- On décompose l'événement en union disjointe d'événements dont on sait calculer la probabilité.
- Fixons un ensemble $I \subset [[1, n]]$, de cardinal k ; on définit l'événement
$$F_I = (\cap_{i \in I} E_i) \cap \left(\cap_{j \notin I} \overline{E_j} \right)$$
("les E_i, pour $i \in I$, se produisent, les autres non" ; $F_I \subset \{X = k\}$)
- Chaque E_i a probabilité p, chaque $\overline{E_j}$ a probabilité $1 - p$; et tous sont indépendants ; donc $\mathbb{P}(F_I) = p^k (1 - p)^{n-k}$.
- **De plus**, $\{X = k\}$ est l'union (disjointe) de tous les événements F_I, pour toutes les parties I **de cardinal** k de $[[1, n]]$; et il y a $\binom{n}{k}$ telles parties.
- **Donc**, $\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Lois géométriques

- Une autre situation classique : on a une *suite* (infinie) de Bernoulli indépendantes $B_1, B_2, \ldots$, de même paramètre ; et on s'intéresse à l'**indice** de la première qui vaut 1.
- Formellement,

$$X = \min \{n > 0 : B_n = 1\}$$

- **Exemple :** je joue à pile ou face jusqu'à tirer Pile pour la première fois ; je regarde le nombre de fois que je lance ma pièce.
- **Exemple :** Chaque jour, M. Malotru prend un sens interdit ; chaque jour, il a une chance sur 50 de prendre une contravention. Quelle est la loi du nombre de sorties jusqu'à l'amende ?
- Il y a un seul paramètre : la probabilité p, à chaque coup, de **s'arrêter** (le paramètre commun des Bernoulli).

Détermination de la loi géométrique

- Ensemble des valeurs possibles : tous les entiers strictement positifs, $\mathbb{N} - \{0\}$.
- (plus, possiblement, $+\infty$? Au cas où toutes les B_i seraient nulles)
- On se donne donc un entier $k > 0$, et on détermine la probabilité de l'événement $\{X = k\}$.
- On peut écrire l'événement $\{X = k\}$ en terme des B_i :
  - Pour que $X = k$, il faut que chacune des B_i, pour $i < k$, soit nulle.
  - Il faut aussi qu'on ait $B_k = 1$.
  - Et ça suffit.
  - Comme les B_i sont indépendantes, la probabilité se calcule facilement
- Donc, $\mathbb{P}(X = k) = (1 - p)^{k-1} p$ (pour tout entier $k > 0$)
- On calcule $\sum_{k=1}^{+\infty} (1 - p)^{k-1} p = 1$; donc $\mathbb{P}(X = \infty) = 0$ et on peut considérer X comme une variable aléatoire à valeurs dans $\mathbb{N}$.

Propriétés de la loi géométrique

- On considère une variable G, géométrique de paramètre p ; et pour un entier $k > 0$, on note A l'événement $\{G > k\}$.
- On regarde alors la **loi de G sachant A**.
- $\mathbb{P}(A) = (1-p)^k$ ("les k premières Bernoulli sont nulles")
- Pour un entier m, on calcule $\mathbb{P}_A(G = m)$.
  - si $m \leq k$: $\{G = m\} \cap A = \emptyset$, donc $\mathbb{P}_A(G = m) = 0$;
  - si $m > k$: $\{G = m\} \subset A$, donc $\mathbb{P}_A(G = m) = \mathbb{P}(G = m)/\mathbb{P}(A)$, ce qui donne

$$\mathbb{P}_A(G = m) = (1-p)^{m-k-1}p = \mathbb{P}(G = m - k)$$

- Au final : pour tout n, $\mathbb{P}_A(G = k + n) = \mathbb{P}(G = n)$; cette propriété est **l'absence de mémoire** de la loi géométrique.
- **Interprétation :** si j'attends le "premier 6" dans une séquence de lancers de dé, et que j'ai déjà lancé 10 fois le dé sans obtenir de 6, le nombre de lancers qu'il me reste à faire suit toujours la loi géométrique.

Lois de Poisson

- Pas d'interprétation élémentaire.
- **Une situation naturelle :** on sait que, *grosso modo*, il y a environ m voitures par heure, en moyenne, qui passent par un carrefour ; mais il y a des variations aléatoires, ce n'est pas exactement m toutes les heures.
- Si on considère que les voitures ne s'influencent pas les unes les autres, on peut modéliser le nombre de celles qui passent pendant une période donnée d'une heure, par une loi de Poisson de paramètre m.
- **Par exemple :** il y a N (un grand nombre) voitures dans les parages, et chacune a probabilité m/N (petite donc) de passer chaque heure ; et les passages des voitures sont indépendants.
- On pourrait utiliser la loi binomiale de paramètres N et m/N, mais N est grand et on ne le connaît pas forcément, on connaît seulement m.

Loi de Poisson (suite)

- Les valeurs possibles sont tous les entiers positifs ou nuls ($\mathbb{N}$).
- Pour tout $k \geq 0$, $\mathbb{P}(X = k) = e^{-m}\frac{m^k}{k!}$, pour $m > 0$ fixé (le "paramètre").
- (Il se trouve que $\sum_{k=0}^{+\infty} \frac{x^k}{k!} = e^x$, donc pour la loi de Poisson la somme fait bien 1)
- Le nombre de clients qui entrent en une minute dans une boutique ; le nombre d'accidents d'avion dans une année ; le nombre de meurtres au cours d'une année, peuvent souvent être modélisés par des lois de Poisson.
- **Propriété (TD) :** la somme d'une variable X de Poisson, de paramètre x, et d'une variable Y de Poisson, indépendante de X, de paramètre y, est une Poisson de paramètre $(x + y)$.

Récapitulatif

- **Bernoulli :** variable aléatoire X valant 0 ou 1 ; $p = \mathbb{P}(X = 1)$.
- **Géométrique :** nombre d'essais jusqu'au premier succès dans une séquence d'essais indépendants et de même probabilité de succès ; $\mathbb{P}(G = k) = (1 - p)^{k-1}p$ pour $k > 0$.
- Absence de mémoire pour la géométrique.
- **Binomiale :** nombre d'essais réussis parmi n essais indépendants et de même probabilité de succès ; $\mathbb{P}(B = k) = \binom{n}{k}p^k(1 - p)^{n-k}$ pour $0 \leq k \leq n$.
- **Poisson :** (genre de limite $n \to \infty$ pour les binomiales $(n, x/n)$) $\mathbb{P}(P = k) = e^{-x}x^k/k!$
- La somme de Poisson indépendantes, est une Poisson de paramètre, la somme des paramètres.

Un peu de curiosité ne nuit pas...

Dans certains cas, quand on somme des variables aléatoires indépendantes suivant des lois connues, la somme a aussi une loi connue...

- X binomiale (n, p), et Y Bernoulli p ; $Z = X + Y$?
- X binomiale (n, p), et Y binomiale (m, p) ; $Z = X + Y$?
- X Poisson de paramètre m, Y Poisson de paramètre ℓ ; $Z = X + y$?
- X et Y géométriques de même paramètre p ; $Z = X + Y$ suit une loi qu'on n'a pas nommée, mais qui a une forme intéressante.
- X et Y géométriques de paramètres p et q ; $Z = \min(X, Y)$?

Organisation DS

DS : vendredi 25 novembre à 15h30

il faudra arriver en amphis avant 15h30, afin de pouvoir débuter le DS à l'heure !

documents autorisés : 1 feuille A4 manuscrite recto-verso

répartition :

- amphi Darwin : groupes A3, A5 et A6
- amphi Edison : groupes A1, A2 et A4
- salle 201, bât. A22 : les étudiants tiers-temps