Chapitre 2 : Variables aléatoires discrètes

Les variables aléatoires qui apparaissent dans ce chapitre (et son appendice) sont des exemples de variables aléatoires discrètes.

I- Variables aléatoires.

1. Un exemple

On lance deux dés non truqués et on appelle X le point indiqué par le premier dé et Y le point indiqué par le second. Posons $S = X + Y$.
On veut étudier la loi de la quantité S sans avoir à considérer en détail X et Y.
On observe d'abord que S prend ses valeurs dans $\{2, ..., 12\}$. On trouve $S = 2$ ssi $X = Y = 1$, ce qui arrive avec une probabilité $\left(\frac{1}{6}\right)^2$ (puisque les dés sont indépendants).
Continuons : $S = 3$ ssi $(X, Y) = (1, 2)$ ou bien $(2, 1)$. Chacun de ces couples sort avec une probabilité $\left(\frac{1}{6}\right)^2$. Comme les deux événements sont incompatibles, $\mathbb{P}(S = 3) = \frac{2}{36}$.
Notation. Pour gagner du temps, on n'écrit pas dans une probabilité les accolades qui devraient encadrer $S = 3$.

On présente habituellement les valeurs $p_k = \mathbb{P}(S = k)$ dans un tableau :

k	2	3	4	5	6	7	8	9	10	11	12
p_k	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

On verra plus loin que connaître la loi de S équivaut à connaître son tableau.
On voit que la symétrie par rapport à la colonne $S = 7$ se traduit par $p_{14-k} = p_k$ pour tout k dans $\{2, ..., 12\}$.

2. Variables aléatoires.

Définition. Soit Ω un ensemble fini $\{x_1, ..., x_n\}$ muni d'une probabilité $\mathbb{P}$. On appelle variable aléatoire toute fonction définie sur l'ensemble Ω.

Des variables aléatoires d'un autre type (type continu) sont étudiées dans le chapitre 4.
Notation. En général, on note une variable aléatoire par une lettre majuscule, et la variable s'appelle ω plutôt que x. Donc l'image par la variable aléatoire X de l'élément ω se note $X(\omega)$.

Exemples

a) Soit la fonction constante qui associe à tout x_i le nombre réel fixé. C'est donc une variable aléatoire constante !
b) S^2 est une variable aléatoire parce que S est une variable aléatoire.

En fait, il existe une autre notion plus délicate : celle de loi d'une variable aléatoire.

3. Définition de la loi d'une variable aléatoire X.

On appelle loi de X la probabilité $\mathbb{Q}$ sur l'ensemble image $X(\Omega) = \{X(x_1), ..., X(x_n)\}$ telle que si y est un élément de $X(\Omega)$, $\mathbb{Q}(\{y\}) = \mathbb{P}(\{\omega$ tels que $X(\omega) = y\})$.

$\mathbb{Q}(\{y\})$ est donc la somme des probabilités des différents antécédents de y par la fonction X.

En fait, l'ensemble des ω tels que $X(\omega) = y$ est noté souvent $\{X = y\}$.

Ainsi dans l'exemple du début, $\mathbb{Q}(\{3\} = \mathbb{P}(\{S = 3\}) = \mathbb{P}(\{(X, Y) = (1, 2)$ ou$(2, 1)\})$.

En supprimant les accolades pour gagner du temps, la loi $\mathbb{Q}$ est telle que

$$\mathbb{Q}(\{y\}) = \mathbb{P}(X = y).$$

Remarque pratique. Comme on l'a vu au chapitre I, il est commode de décrire une probabilité par un tableau. Comme la loi d'une variable aléatoire est une probabilité, elle peut se décrire par un tableau. C'est justement ce qui a été fait dans l'exemple du 1.

Définition. On dit que deux variables aléatoires ont la même loi ssi elles ont le même tableau. On verra que la plupart des calculs ne nécessitent pas la connaissance des variables aléatoires, mais seulement celle de leurs lois.
Ainsi dans l'exemple du 1., S et $14 - S$ ont la même loi. Bien entendu, S et $14 - S$ ne sont pas des variables aléatoires égales, nous aurions dans ce cas l'absurdité $S = 14 - S$, donc $S = 7$ (???).

4. Fonction de répartition.

♡ **Définition.** On appelle fonction de répartition de la variable aléatoire X la **fonction** de $\mathbb{R}$ dans $\mathbb{R}$:

$$t \longrightarrow \mathbb{P}(X \leq t) = F(t).$$

Dessin du graphe d'une fonction de répartition d'une va. à valeurs dans $\{0, 1, 2, 3\}$.

Propriétés.

Si F est la fonction de répartition d'une variable aléatoire X discrète alors,

- F est une fonction en escalier, **continue à droite**

- F est croissante (au sens large)
- $\lim_{t\to-\infty} F(t) = 0$ et $\lim_{t\to+\infty} F(t) = 1$.

Remarque.
Si a et b sont deux réels tels que $a < b$, $\mathbb{P}(a < X \leq b) = F(b) - F(a)$.

5. Exemples de lois de probabilité.

a) Loi uniforme

♡ **Définition.** On appelle loi uniforme sur un ensemble de n éléments (distincts) $\Omega = \{x_1, ..., x_n\}$, la loi de probabilité telle que pour tout $1 \leq k \leq n$, $\mathbb{P}(\{x_k\}) = \dfrac{1}{n}$.

♡ **b) Loi binômiale**

Définition. Soient n un entier plus grand que 1 et p un réel dans $[0,1]$.
On dit qu'une variable aléatoire X possède la loi binômiale $\mathcal{B}(n,p)$ ssi X prend ses valeurs dans $\{0, ..., n\}$ avec

$$\mathbb{P}(X = k) = C_n^k\, p^k (1-p)^{n-k} \quad \text{pour} \quad 0 \leq k \leq n.$$

Interprétation. On répète n fois une épreuve conduisant au succès avec une probabilité p ou bien à l'échec avec une probabilité $1 - p = q$. On suppose en outre que les essais sont indépendants. Alors le nombre de succès X suit la loi $\mathcal{B}(n,p)$.

Remarques
a) $\mathcal{B}(n,p)$ est le nom de la loi. Le premier paramètre est le nombre d'épreuves, le second est la probabilité de réussite.
b) Le nombre d'échecs (qui est $n - X$) suit une loi $\mathcal{B}(n,q)$ si $q = 1 - p$.
c) Comme p ne dépend pas du numéro de l'épreuve, la loi binômiale ne peut modéliser qu'un **tirage avec remise**.

c) Un cas particulier de la loi binômiale : la loi de Bernoulli

Définition. C'est la loi $\mathcal{B}(1,p)$ telle que $n = 1$. Autrement dit, $p_0 = q$ et $p_1 = p$.

♡ **d) La loi hypergéométrique**

Il s'agit maintenant de modéliser un **tirage sans remise** : on prélève $n \leq b + r$ boules dans une urne contenant b boules bleues et r boules rouges. La probabilité de tirer une boule bleue dépend des tirages antérieurs. Appelons X le nombre de boules bleues obtenues après n tirages sans remise. On dira que X suit la loi hypergéométrique $\mathcal{H}(n,b,r)$.

Définition
Soient r, b et n des entiers plus grands que 1 tels que $n \leq r + b$.
On dit qu'une variable aléatoire X possède la loi hypergéométrique $\mathcal{H}(n,b,r)$ ssi X prend ses valeurs dans $\{0, ..., n\}$ avec

$$\mathbb{P}(X = k) = \frac{C_b^k C_r^{n-k}}{C_{b+r}^n} \quad \text{pour } 0 \leq k \leq b \text{ et } 0 \leq n - k \leq r.$$

Remarques

a) Les inégalités servent à assurer l'existence des combinaisons C_b^k et C_r^{n-k}.

b) **Approximation binômiale**. Quand k est très petit devant b et r, $\mathbb{P}(X = k)$ diffère peu de la formule binômiale avec $p = \frac{b}{b+r}$.

II- Espérance et variance d'une variable aléatoire.

1. Définition de l'espérance.

♡ Si X est une variable aléatoire prenant les valeurs $x_1, ..., x_n$ avec les probabilités respectives $p_1, ...p_n$, on appelle espérance de X le nombre réel $\sum_{k=1}^{n} p_k x_k$.

On pose alors $\mathbb{E}(X) = \displaystyle\sum_{k=1}^{n} p_k x_k$.

C'est la valeur moyenne de X au sens des probabilités. On verra au chapitre suivant que sous certaines hypothèses, la moyenne arithmétique de n réalisations indépendantes converge vers l'espérance. L'appellation « espérance mathématique » est désuète. Au XIXe siècle, on voulait distinguer $\mathbb{E}(X)$ d'un concept oublié aujourd'hui, appelé l'espérance morale.

Remarque. Analogie avec la mécanique : les probabilités p_k sont analogues à des masses, les x_k à des distances. $\mathbb{E}(X)$ est analogue à un moment.

Exemples :

a) Si X est constante, $n = 1$, $x_1 = c$ et $p_1 = 1$, on trouve par définition $\mathbb{E}(X) = p \times c = c$.

b) Si X prend deux valeurs 1 et 3 avec les probabilités respectives p et $1 - p = q$, $\mathbb{E}(X) = p \times 1 + q \times 3$. Si on prend maintenant $p = 0,5$ on trouve comme espérance $\frac{1+3}{2} = 2$, moyenne au sens usuel.

Attention, ce n'est pas la valeur la plus probable de la variable aléatoire. Ici $\mathbb{P}(X = 2) = 0$.

c) Soit le jeu suivant : on parie 5 francs sur une liste ordonnée de 3 chevaux parmi 10. Si on gagne dans l'ordre on touche 10.000 francs, par contre si on gagne dans le désordre on touche 450 francs. Quelle est l'espérance du gain?

Notons $p_1 = \mathbb{P}($ gagner dans l'ordre$)$ et $p_2 = \mathbb{P}($ gagner dans le désordre et pas dans l'ordre $)$.

L'espérance de gain (algébrique, en tenant compte de la dépense initiale de 5 francs, bilan entre ce qui sort du porte-monnaie et ce qui y rentre éventuellement) est

$p_1 \times 9995 + p_2 \times 445 - \mathbb{P}($ ne rien gagner du tout $) \times 5 = -5 + p_1 \times 10000 + p_2 \times 450$.

En effet les 5 francs pariés sont de toute façon partis du porte-monnaie.

d) Si X possède une loi de Bernoulli $\mathcal{B}(1, p)$, $\mathbb{E}(X)$ est le paramètre p.

e) Si X possède la loi uniforme sur $\{x_1, ..., x_n\}$ (de cardinal n), alors $\mathbb{E}(X) = \dfrac{\sum_{i=1}^{n} x_i}{n}$.

C'est la moyenne au sens usuel.

f) Soit A un événement et X la variable aléatoire $\mathbb{1}_A$. Alors $\mathbb{E}(X) = \mathbb{P}(A)$.

Preuve : X possède la loi de Bernoulli de paramètre $p = \mathbb{P}(A)$.

2. Centrage

$\heartsuit$ **Définition.** Une variable aléatoire est **centrée** ssi son espérance est nulle.
Centrer une variable aléatoire, c'est retrancher son espérance. On pose donc $\tilde{X} = X - \mathbb{E}(X)$.
D'après les propriétés du paragraphe suivant, on trouve que $\mathbb{E}(\tilde{X}) = 0$. Cette transformation a fait apparaître une variable aléatoire centrée.

Commentaire : revenons à notre exemple du jeu de hasard. On dit qu'il est équitable ssi l'espérance du gain algébrique G est nulle.
Par contre si $\mathbb{E}(G) < 0$, il est défavorable (déséquilibré en direction de la caisse de l'organisateur).
D'après la loi des grands nombres (qui sera énoncée dans le chapitre suivant), on peut montrer que la fortune de l'organisateur tend presque sûrement vers l'infini quand le jeu dure. Il lui suffit d'être patient pour s'enrichir.

Analogie avec la mécanique : équilibre d'une balance par égalité des moments.

$\heartsuit$ 3. Propriétés de l'espérance

1. Si X est la variable aléatoire constante $= c$, $\mathbb{E}(X) = c$ (déja vu).
2. Si λ est un réel (qui ne dépend pas du hasard), la variable aléatoire λX a l'espérance $\lambda \mathbb{E}(X)$.
3. Si X et Y sont deux variables aléatoires définies sur Ω, $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ (admis).

Corollaires. En combinant les propriétés précédentes :

4. **Linéarité de l'espérance.**
Si λ et μ sont deux réels (qui ne dépendent pas du hasard), $\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$.
On peut l'étendre par récurrence à la somme de n variables aléatoires.
5. Si X est une variable aléatoire à valeurs positives, $\mathbb{E}(X)$ est positif.
6. Plus généralement si toutes les valeurs de X sont comprises entre des réels a et b, alors $a \leq \mathbb{E}(X) \leq b$.

C'est une propriété que toute moyenne raisonnable doit vérifier : on voit mal comment la note moyenne d'un groupe de TD pourrait ne pas être comprise entre 0 et 20.
Encore plus généralement si deux variables aléatoires X et Y sont telles qu'on a toujours $X \leq Y$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

7. Comme $\mathbb{E}(X)$ ne dépend que des nombres p_k et x_k, l'espérance ne dépend que de la loi de X. Par conséquent, si deux variables aléatoires ont la même loi, alors elles ont la même espérance.

Un exemple d'application de la propriété 3. : pour calculer simplement l'espérance de S du début du chapitre, il est commode de remarquer que $\mathbb{E}(S) = \mathbb{E}(X) + \mathbb{E}(Y)$.
Mais X et Y ayant la même loi, c'est $2\,\mathbb{E}(X)$ (prop 7).
La loi de X étant uniforme sur $\{1, ..., 6\}$, elle a pour espérance, $\frac{1}{6} \times (1 + ... + 6) = \dfrac{7}{2}$.
Finalement, $\mathbb{E}(S) = 7$.

Encore plus astucieusement, S et $14-S$ ayant la même loi, $\mathbb{E}(S) = \mathbb{E}(14-S)$. Mais par linéarité, $\mathbb{E}(14-S) = \mathbb{E}(14) - \mathbb{E}(S) = 14 - \mathbb{E}(S)$. En résolvant $\mathbb{E}(S) = 14 - \mathbb{E}(S)$, on retrouve le même résultat : $\mathbb{E}(S) = 7$.

♡ 4. Formule de transfert

Enoncé. *Soient f une fonction définie sur $\Omega = \{x_1, ..., x_n\}$ et X une variable aléatoire définie sur Ω. Alors $Y = f(X)$ est une variable aléatoire et son espérance $\mathbb{E}(Y)$ se calcule par la formule*

$$\sum_{k=1}^{n} p_k f(x_k).$$

Exemple : soit X la variable aléatoire du b) et f la fonction carré sur $\{1, 3\}$. On trouve $\mathbb{E}(X^2) = p \times 1^2 + q \times 3^2$.
Cela donne le même résultat que si nous utilisions la formule définissant l'espérance à la variable aléatoire X^2 qui vérifie $\mathbb{P}(X=1) = p$ et $\mathbb{P}(X=9) = q$.

Autre exemple : si X possède une loi de Bernoulli $\mathcal{B}(1,p)$, $\mathbb{E}(X^2) = p \times 1^2 + q \times 0^2 = p$. De même pour tout entier positif l, $\mathbb{E}(X^l) = p$. C'était évident a priori puisque $X^l = X$ (car X prend ses valeurs dans $\{0, 1\}$).

Attention ! Il ne faut surtout pas élever au carré les probabilités : on voit mal comment $p^2 + q^2$ pourrait être encore égal à 1. De même, en poursuivant l'analogie avec la mécanique, on voit mal le sens physique à l'élévation au carré des masses (=probabilités). Ce sont les distances qui s'élèvent au carré dans le calcul des moments d'inertie.

5. Variance et écart-type

Le but est de mesurer simplement la dispersion de X autour de son espérance. On utilise un carré car les carrés sont faciles à développer. Une autre raison sera vue dans le chapitre 3 : la variance apparaît dans le théorème central limite.

♡ Définition

On appelle variance de X (notée $\mathrm{Var}(X)$) le nombre réel

$$\mathrm{Var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2 \right).$$

On peut donc écrire $\mathrm{Var}(X) = \mathbb{E}(\tilde{X}^2)$.

C'est un nombre dont l'unité de mesure est le carré de celle de X.

♡ Propriétés

1. C'est un nombre positif ou nul.
2. La variance est nulle si et seulement si X est presque constante, c'est à dire qu'il existe un réel b tel que $\mathbb{P}(X=b) = 1$ (en fait ce réel b est forcément $\mathbb{E}(X)$).
3. Si α est un réel (qui ne dépend pas du hasard), $\mathrm{Var}(\alpha X) = \alpha^2 \mathrm{Var}(X)$.
4. Si c est une constante, $\mathrm{Var}(X+c) = \mathrm{Var}(X)$.

En général, pour effectuer le calcul de la variance, on se sert plutôt de la formule de Koenigs:

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

♡ **Définition.** On appelle écart-type d'une variable aléatoire la racine carrée de sa variance. On la note en général σ (sigma). Ce nombre a la même unité de mesure que X.

Exemples
a) Soit X de loi de Bernoulli $\mathcal{B}(1,p)$. Alors $\mathrm{var}(X) = p - p^2 = pq$.

Preuve : $\mathbb{E}(X) = p$ et $\mathbb{E}(X^2) = p$, donc d'après la formule de Koenigs, $\mathrm{Var}(X) = p - p^2$.
b) On verra plus loin la formule célèbre

$$\mathrm{var}(X) = npq$$

en supposant que X possède la loi $\mathcal{B}(n,p)$.

♡ **Définition**
On dit qu'une variable aléatoire est **réduite** ssi elle possède une variance égale à 1.

Pour **réduire** une variable aléatoire (c'est à dire la rendre de variance 1), on la divise par son écart-type σ.

III- Couples de variables aléatoires numériques

Dans ce paragraphe, on cherche à étudier deux variables aléatoires à la fois. Commençons par un exemple.

1. Un exemple

On tire simultanément trois boules dans une urne qui contient 3 boules rouges, 1 boule bleue et 2 boules noires. On définit les variables aléatoires X, Y et Z respectivement égales au nombre de boules bleues, noires ou rouges.
On s'aperçoit qu'il est inutile de les étudier toutes les trois à la fois puisque $X + Y + Z = 3$. X prend ses valeurs dans $\{0,1\}$, Y prend ses valeurs dans $\{0,1,2\}$. Comme a priori le couple (X,Y) peut prendre toutes les valeurs possibles dans $\{0,1\} \times \{0,1,2\}$, pour décrire complètement le couple de variables aléatoires, il faut calculer les six probabiliés $p_{i,j} = \mathbb{P}(X = i \ \& \ Y = j)$ quand i décrit $\{0,1\}$ et j décrit $\{0,1,2\}$. On remplit pour cela les six cases intérieures du tableau, l'intersection de la ligne i et de la colonne j contenant le nombre $p_{i,j}$.

X / Y	0	1	2	
0	$\frac{1}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	$\frac{10}{20}$
1	$\frac{3}{20}$	$\frac{6}{20}$	$\frac{1}{20}$	$\frac{10}{20}$
	$\frac{4}{20}$	$\frac{12}{20}$	$\frac{4}{20}$	1

Détail de certain résultats :
L'espace de probabilité est Ω constitué des 3-parties de l'ensemble des six boules. Il est donc de cardinal $C_6^3 = 20$.
Le numérateur de $p_{0,0}$ est 1 puisque seul le tirage des trois boules rouges convient.
Le numérateur de $p_{0,1} = \mathbb{P}(X = 0 \,\&\, Y = 1)$ est $C_2^1 C_3^2$. Donc $p_{0,1} = \frac{2 \times 3}{20}$.
Il est commode de remplir aussi cinq cases supplémentaires. La dernière ligne contient q_1, q_2 et q_3 où $q_j = \mathbb{P}(Y = j)$. De même la dernière colonne contient les $p_i = \mathbb{P}(X = i)$.
On peut calculer directement ces probabilités : par exemple, en regroupant les boules noires et rouges en une seule catégorie (les non bleues), X possède une loi hypergéométrique : $\mathbb{P}(X = i) = \dfrac{C_1^i C_5^{3-i}}{C_6^3}$ pour i dans $\{0, 1\}$.
De même, Y possède une loi hypergéométrique : $\mathbb{P}(Y = j) = \dfrac{C_2^j C_4^{3-j}}{C_6^3}$ pour j dans $\{0, 1, 2\}$.

On peut aussi observer que les nombres de la dernière ligne sont les sommes des probabilités de leur colonne. De même, les $\frac{1}{2}$ de la dernière colonne correspondent à la somme de leurs lignes respectives.

2. Notations et définitions

Soient m et n deux entiers plus grands que 1 et deux ensembles $E = \{x_1, ..., x_m\}$ et $F = \{y_1, ..., y_n\}$ possédant respectivement m et n éléments. On considère alors l'ensemble $E \times F$ de tous les couples possibles.
La loi d'une variable aléatoire (X, Y) à valeurs dans $E \times F$ est caractérisée par la donnée des $m \times n$ nombres $p_{i,j} = \mathbb{P}(X = x_i \,\&\, Y = y_j)$ pour $1 \leq i \leq m$ et $1 \leq j \leq n$.
Les probabilités $p_{i,j}$ doivent être des nombres positifs ou nuls dont la somme doit être égale à 1 :

$$\sum_{i=1}^{m} \sum_{j=1}^{n} p_{i,j} = 1.$$

♡ On appelle **première loi marginale** la loi de la première composante X. Elle est caractérisée par les m nombres $p_i = \mathbb{P}(X = x_i)$ pour $1 \leq i \leq m$.

♡ On appelle **seconde loi marginale** la loi de la seconde composante Y. Elle est caractérisée par les n nombres $q_j = \mathbb{P}(Y = y_j)$ pour $1 \leq j \leq n$.

Relations entre les $p_{i,j}$ et les p_i.

Fixons i dans $\{1, ..., m\}$. Comme les événements $\{(X,Y) = (x_i, y_j)\}$ $(1 \leq j \leq n)$ sont disjoints deux à deux de réunion $\{X = x_i\}$, on a

$$p_i = \mathbb{P}(X = x_i) = \sum_{j=1}^{n} \mathbb{P}((X,Y) = (x_i, y_j)) = \sum_{j=1}^{n} p_{i,j}.$$

Bien entendu, il existe des relations analogues entre les $p_{i,j}$ et les q_j.

Conclusions
Si i est le numéro de la ligne, p_i est la somme de cette ligne.
En général, la donnée des nombres p_i , $i \leq m$ et des q_j $j \leq n$ ne permet pas de reconstituer les p_i, j. **Sauf dans des cas particuliers, connaître séparément les lois de X et de Y ne permet pas de retrouver la loi du couple (X,Y).**

3. Indépendance de deux variables aléatoires

♡ **Définition.** : Avec les notations du paragraphe précédent, on dit que deux variables aléatoires sont indépendantes, si et seulement si pour tous les $1 \leq i \leq m$ et les $1 \leq j \leq n$, $p_{i,j} = p_i \times q_j$, c'est à dire $\mathbb{P}(X = x_i \ \& \ Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j)$.

Par conséquent, pour prouver que deux variables aléatoires ne sont pas indépendantes, il suffit d'exhiber au moins un couple d'indices (i, j) pour lequel $p_{i,j} \neq p_i \times q_j$.

Ainsi, les variables aléatoires X et Y du 1. ne sont pas indépendantes, puisque
$\mathbb{P}((X,Y) = (0,0)) \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 0)$.

Théorème III.3. *Si X et Y sont indépendantes, alors pour tout événement A inclus dans $E = \{x_1, ..., x_m\}$ et pour tout événement B inclus dans $F = \{y_1, ..., y_n\}$, $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants.*

4. Espérance d'un produit de deux variables aléatoires indépendantes

♡ **Théorème III-4.** *Si X et Y sont indépendantes, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.*

Preuve : il suffit de remarquer que

$$\sum_{i=1}^{m} \sum_{j=1}^{n} p_{i,j} x_i y_j = \left(\sum_{i=1}^{m} p_i x_i \right) \left(\sum_{j=1}^{n} q_j y_j \right)$$

puisque $p_{i,j} = p_i \times q_j$.

IV- Covariance de deux variables aléatoires

1. Définition. Soient deux variables aléatoires définies sur le même espace de probabilité. On appelle covariance de X et de Y le nombre (**de signe quelconque**)

$$\mathrm{cov}(X,Y) = \mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)\right)$$

Remarques

a) Avec les notations du paragraphe II.2, $\tilde{X} = X - \mathbb{E}(X)$ (X rendue centrée) et $\tilde{Y} = Y - \mathbb{E}(Y)$, $\mathrm{cov}(X,Y) = \mathbb{E}(\tilde{X}.\tilde{Y})$.

b) $\mathrm{cov}(X,X) = \mathrm{var}(X)$ qui est positif.
Penser aux analogues $x.x$ (carré d'un réel x) ou $u.u$ produit scalaire d'un vecteur u avec lui même.

♡ En pratique, on calcule une covariance par **la formule de Koenigs** :

$$\mathrm{cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Evidemment si $X = Y$, c'est la formule de Koenigs déjà connue.

♡ **2. L'utilité principale de la covariance**

Soient X et Y deux variables aléatoires définies sur le même espace de probabilité. Alors

$$\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\ \mathrm{cov}(X,Y).$$

C'est l'analogue de $(x+y)^2 = x^2 + y^2 + \ 2\ xy$ ou de $\parallel u+v \parallel^2 = \parallel u \parallel^2 + \parallel v \parallel^2 + 2\ u.v$ où u et v sont deux vecteurs du plan euclidien.

3. Propriétés de la covariance

En fait, la plupart des propriétés de la covariance sont analogues à celles du produit de deux réels ou du produit scalaire de deux vecteurs. En termes savants, cela s'appelle la bilinéarité.

Soient X, Y, X_1, X_2, Y_1 et Y_2 des variables aléatoires définies sur le même espace de probabilité. Soient c_1, c_2, α, β, α_1, α_2, β_1 et β_2 des nombres réels qui ne dépendent pas du hasard.

a) $\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$ *(symétrie)*

b) $\mathrm{cov}(\alpha_1 X_1 + \alpha_2 X_2, Y) = \alpha_1 \mathrm{cov}(X_1,Y) + \alpha_2 \mathrm{cov}(X_2,Y)$
c) $\mathrm{cov}(X, \beta_1 Y_1 + \beta_2 Y_2) = \beta_1 \mathrm{cov}(X,Y_1) + \beta_2 \mathrm{cov}(X,Y_2)$
d) En regroupant b) et c)
$\mathrm{cov}(\alpha_1 X_1 + \alpha_2 X_2, \beta_1 Y_1 + \beta_2 Y_2) = \alpha_1\beta_1 \mathrm{cov}(X_1,Y_1) + \alpha_1\beta_2 \mathrm{cov}(X_1,Y_2) + \alpha_2\beta_1 \mathrm{cov}(X_2,Y_1) + \alpha_2\beta_2 \mathrm{cov}(X_2,Y_2)$.

e) *Propriété nouvelle par rapport au produit scalaire.*
$\mathrm{cov}(X+c_1, Y+c_2) = \mathrm{cov}(X,Y)$.

4. Variables aléatoires non corrélées

♡ **Définition.** On dit que deux variables aléatoires sont non corrélées ssi leur covariance est nulle.

♡ ♡ **Théorème IV-1.** *Si deux variables aléatoires sont indépendantes, alors elles sont non corrélées.*

Preuve : c'est simplement le théorème III-4.

La RÉCIPROQUE EST FAUSSE, IL EXISTE BEAUCOUP DE VARIABLES ALÉATOIRES NON CORRELÉES, QUOIQUE NON INDÉPENDANTES !!!

Elles abondent dans les exercices et les sujets d'examen.

Voici un exemple : on tire au hasard un point dans le plan de coordonnées X et Y qui sont des variables aléatoires.
On suppose qu'on a le choix équiprobable entre les quatre points (0,1), (1,0), (-1,0) et (0,-1).

On trouve $\mathbb{E}(X) = \mathbb{E}(Y) = 0$ et aussi $\mathbb{E}(XY) = 0$ puisque XY est toujours nul. Donc la covariance de X et de Y est nulle.
Par contre, X et Y **ne sont pas indépendantes**. En effet,
$\mathbb{P}(X = 1) = \frac{1}{4}$ et $\mathbb{P}(Y = 1) = \frac{1}{4}$ alors que $\mathbb{P}(X = 1 \;\&\; Y = 1) = 0 \neq (\frac{1}{4})^2$.

Remarque 4.2 : si les variables aléatoires X et Y sont indépendantes, d'après le théorème 4.1, $\mathrm{cov}(X, Y) = 0$. Alors la formule (utilisation principale de la covariance) devient

$$\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y).$$

5. Cas des sommes de variables aléatoires indépendantes de même loi

Proposition IV-3. *Soient X_1,...,X_n n variables aléatoires indépendantes et de même loi. On pose alors*

$$S_n = \sum_{i=1}^{n} X_i = X_1 + ... + X_n.$$

- *a)* $\mathbb{E}(S_n) = n\,\mathbb{E}(X_1)$ *et donc* $\mathbb{E}(\frac{S_n}{n}) = \mathbb{E}(X_1)$.
- *b)* $\mathrm{var}(S_n) = n\mathrm{var}(X_1)$ *et* $\mathrm{var}(\frac{S_n}{n}) = \frac{\mathrm{var}(X_1)}{n}$
- *c) L'écart-type de* $\frac{S_n}{n}$ *est* $\frac{1}{\sqrt{n}}\sigma(X_1)$ *si* $\sigma(X_1)$ *est l'écart-type de* X_1.

Par conséquent, l'écart-type de $\frac{S_n}{n}$ tend vers 0 quand n croît vers l'infini.
Cela veut dire que $\frac{S_n}{n}$ a tendance à se concentrer autour de sa valeur moyenne (qui est $\mathbb{E}(X_1)$).

6. Complément : coefficient de corrélation

Soient X et Y deux variables aléatoires possédant des variances non nulles. On appelle alors coefficient de corrélation (noté $\rho_{X,Y}$) entre les variables aléatoires X et Y le nombre réel

$$\rho_{X,Y} = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}}.$$

Remarque : d'après l'inégalité de Cauchy-Schwarz, $-1 \leq \rho_{X,Y} \leq 1$.

V. Propriétés de la loi binômiale

♡ **1. Écriture comme somme.** Soit S une variable aléatoire de loi binômiale $\mathcal{B}(n,p)$. On sait qu'elle représente le nombre de succès lors de n épreuves indépendantes conduisant ou bien à un échec ou bien à un succès.
Posons $X_i = \mathbb{1}_{\{\text{la i-ème épreuve est un succès}\}}$, c'est à dire que cette variable aléatoire vaut 1 (si c'est un succès) ou bien 0 (si c'est un échec).
Il est clair que la somme de nombres valant 0 ou bien 1 est un entier qui est le nombre de fois que 1 apparaît dans la somme.
Par conséquent, $S = \sum_{i=1}^{n} X_i$. Les variables aléatoires X_i sont indépendantes puisque les épreuves sont indépendantes. De plus, elles possèdent la même loi, qui est une loi de Bernoulli $\mathcal{B}(1,p)$.

♡ **2. Calcul de l'espérance et de la variance**

♡ Si S possède la loi $\mathcal{B}(n,p)$, alors $\mathbb{E}(S) = np$ et $\text{var}(S) = np(1-p) = npq$.

Démonstration. D'après la représentation comme somme et la linéarité de l'espérance, $\mathbb{E}(S) = \sum_{i=1}^{n} \mathbb{E}(X_i) = np$ *(ici l'indépendance ne sert pas)*.
On calcule facilement la variance de la somme comme au paragraphe IV.5 *(par indépendance)*.

3. Propriété de stabilité

Théorème. *Supposons que X possède la loi $\mathcal{B}(n,p)$ et que Y possède la loi $\mathcal{B}(m,p)$ (avec le même p !). Alors* **si X et Y sont indépendantes**, *$X+Y$ possède la loi $\mathcal{B}(m+n,p)$.*

Preuve : il suffit d'écrire en vertu du point 1., $X = \sum_{i=1}^{n} X_i$ et $Y = \sum_{i=n+1}^{m} X_i$, les $n+m$ variables aléatoires X_i $(1 \leq i \leq m+n)$ étant indépendantes de même loi de Bernoulli $\mathcal{B}(1,p)$. C.Q.F.D

Appendice : complément sur les variables aléatoires à valeurs dans $\mathbb{N}$.

Ces variables peuvent donc prendre un nombre infini de valeurs entières.

A1. Un exemple.

On lance une pièce de monnaie qui donne pile avec une probabilité p et face avec une probabilité $q = 1 - p$. On s'intéresse au premier instant (aléatoire) où pile sort pour la première fois. On suppose pour le moment $0 < p < 1$.

A priori, il n'est pas évident que T soit forcément fini. En fait, on va montrer que $\{T = +\infty\}$ a une probabilité nulle.

D'abord T ne peut prendre que des valeurs ≥ 1 et $\mathbb{P}(T = 1) = p$.
$T = 2$ n'arrive que si F sort d'abord suivi d'un pile. Donc par indépendance des lancers, $\mathbb{P}(T = 2) = qp$. De même $\mathbb{P}(T = 3) = q^2 p$ et finalement pour tout $k \geq 1$, $\mathbb{P}(T = k) = q^{k-1}p$.
En admettant que T est une variable aléatoire, on pose $p_k = q^{k-1}p$.

Par analogie avec le cas fini, on voudrait savoir si la somme de tous les p_k est bien 1, comme pour toute loi de probabilité. Pour cela, on doit introduire la notion de série (somme de longueur infinie).

A2. Série convergente.

Définition. Si la suite des sommes partielles $\sum_{k=1}^{n} p_k$ possède une limite finie l quand n croît vers l'infini, on pose $\sum_{k=1}^{\infty} p_k = l$.
(Fin de l'exemple). On rappelle que si ρ est un réel $\neq 1$, la progression géométrique ρ^m (avec $m \geq 0$) a pour somme de 0 à n, $\sum_{m=0}^{n} \rho^m = \frac{\rho^{n+1}-1}{\rho-1}$.
Donc $\sum_{k=1}^{n} p_k = p \sum_{k=1}^{n} q^{k-1}$ en sortant le p qui ne varie pas avec k. En posant $\rho = q$ et $m = k - 1$, on obtient $\sum_{k=1}^{n} p_k = p\frac{q^n-1}{q-1} = 1 - q^n$ puisque $q - 1 = -p$.
Comme $p \neq 0$, p appartient à $]0, 1[$ et alors q^n converge vers 0 quand n tend vers l'infini.

On peut donc en conclure que $\sum_{k=1}^{\infty} p_k = 1$.

En admettant que les propriétés des probabilités sur un ensemble fini sont encore valides, on devrait avoir $\mathbb{P}(T = \infty) = 1 - \mathbb{P}($ T prend une valeur finie$) = 1 - \sum_{k=1}^{\infty} p_k = 0$.

A3. Variables aléatoires à valeurs dans $\mathbb{N}$.

Théorème admis. *Soit p_k une suite quelconque de nombres réels* **positifs** *ou nuls tels que $\sum_{k=0}^{\infty} p_k = 1$. Alors on peut construire une probabilité $\mathbb{P}$ (vérifiant les propriétés de la définition 2 du chapitre 1) et une fonction T telle pour tout k entier $\mathbb{P}(T = k) = p_k$. On dit encore que T est une variable aléatoire.*

On peut utiliser toutes les propriétés usuelles des variables aléatoires à condition d'interpréter tous les symboles sigma par des sommes de longueur infinie. (sous réserve d'existence)

Par exemple, on dit que T a une espérance ssi la somme $\sum_{k=0}^{+\infty} kp_k$ existe.

A4. La loi géométrique.

♡ **Définition.** Soit p un paramètre dans $]0,1[$. On dit qu'une variable aléatoire à valeurs dans $\mathbb{N}$ telle que
a) $\mathbb{P}(T=0)=0$
b) alors que $\mathbb{P}(T=k)=pq^{k-1}$ pour tout $k \geq 1$ suit la loi géométrique de paramètre p.

On peut noter en abrégé ce nom « loi géométrique de paramètre p » par le symbole $\mathcal{G}^*(p)$. Le * est pour rappeler que $p_0 = 0$ (de la même façon que $\mathbb{N}^*$ est privé de 0).

Cas limites : si $p = 0$, on ne peut jamais observer de pile, donc T n'est pas défini. Par contre si $p = 1$, on obtient $p_1 = 1$, T est la va constante égale à 1.

A5. La loi de Poisson.

♡ **Définition.** Soit λ un paramètre strictement positif. On dit qu'une variable aléatoire à valeurs dans $\mathbb{N}$ telle que pour tout k dans $\mathbb{N}$

$$\mathbb{P}(T=k)=\exp(-\lambda)\frac{\lambda^k}{k!}$$

suit la loi de Poisson de paramètre λ (d'après Denis Poisson, mathématicien français).

Un résultat admis (voir un cours d'analyse). Pour tout réel y, on a

$$\sum_{k=0}^{+\infty}\frac{y^k}{k!}=\exp(y).$$

Par conséquent, en prenant $y = \lambda$, on trouve que $\sum_{k=0}^{+\infty} p_k = 1$. Les p_k vérifient bien les conditions du 3.

La loi de Poisson sert à modéliser le nombre de certains événements rares qui se produisent pendant un intervalle de temps donné, par exemple le nombre de désintégrations de particules.

A6. Calcul d'espérances et de variances

Définition.
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On pose alors pour tout entier k, $p_k = \mathbb{P}(X = k)$. Si la somme de longueur infinie $\sum_{k=0}^{+\infty} k\ p_k$ existe, on dit que X possède une espérance, qui est le nombre défini par cette somme $\sum_{k=0}^{+\infty} k\ p_k$.
Exemple : on dit que $\mathbb{E}(X^2)$ existe ssi la somme $\sum_{k=0}^{+\infty} k^2\ p_k$ existe.

Définition.
Si $\mathbb{E}(X)$ et $\mathbb{E}(X^2)$ existent, on pose comme d'habitude, $\mathrm{var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

Résultats.
a) Soit X une variable aléatoire qui possède la loi de Poisson de paramètre λ. Alors

$$\mathbb{E}(X)=\mathrm{var}(X)=\lambda.$$

b) Soit T une variable aléatoire qui possède la loi géométrique de paramètre p. Alors

$$\mathbb{E}(T) = \frac{1}{p} \quad \text{et} \quad \operatorname{var}(T) = \frac{q}{p^2}.$$

Démonstation :
a) On constate qu'ici $k\,p_k = k \times \exp(-\lambda)\frac{\lambda^k}{k!} = \exp(-\lambda)\frac{\lambda^k}{(k-1)!}$ pour $k \geq 1$.
Donc $\sum_{k=0}^{+\infty} k\,p_k = \exp(-\lambda)\sum_{k=1}^{+\infty} kp_k = \exp(-\lambda)\sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} = \exp(-\lambda)\sum_{i=0}^{+\infty} \frac{\lambda^{i+1}}{i!}$
en posant $i = k-1$.
Finalement d'après le résultat de seconde année, il reste $\exp(-\lambda)\lambda\exp(\lambda) = \lambda$ C.Q.F.D

Par la même méthode, puisque $k(k-1)p_k = k(k-1) \times \exp(-\lambda)\frac{\lambda^k}{k!} = \exp(-\lambda)\frac{\lambda^k}{(k-2)!}$ pour $k \geq 2$, on obtient $\mathbb{E}\left(X(X-1)\right) = \lambda^2$.
Donc $\mathbb{E}(X^2) = \mathbb{E}\left(X(X-1)\right) + \mathbb{E}(X) = \lambda^2 + \lambda$.
Alors $\operatorname{var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$. C.Q.F.D

b) Limitons nous au calcul de l'espérance de T. Sous réserve d'existence,

$$\mathbb{E}(T) = \sum_{k=1}^{+\infty} kp_k = \sum_{k=1}^{+\infty} kpq^{k-1} = p\sum_{k=0}^{+\infty} kq^{k-1}.$$

Soit la fonction f définie sur $[0,1]$ par

$$f(x) = \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}.$$

Alors $f'(x) = \frac{1}{(1-x)^2} = \sum_{k=1}^{+\infty} kx^{k-1}$. Donc $\mathbb{E}(T) = pf'(q) = \frac{1}{p}$ puisque $1-q = p$.

A7. Deux propriétés de la loi de Poisson.

♡ **a) Stabilité.**

Théorème. *Si X et Y sont deux variables aléatoires indépendantes telles que X possède la loi $\mathcal{P}(\lambda)$ et Y possède la loi $\mathcal{P}(\mu)$ où λ et μ sont deux nombres réels strictement positifs quelconques. Alors la variable aléatoire $X+Y$ possède la loi de Poisson de paramètre $\lambda+\mu$.*

Remarque : cet énoncé n'est pas contradictoire avec les calculs de l'espérance et de la variance de $X+Y$: en effet par linéarité, $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) = \lambda + \mu$ ce qui permet d'identifier le paramètre de $X+Y$ si on sait déja que $X+Y$ possède une loi de Poisson.
De plus par indépendance, $\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y) = \lambda + \mu$.

Idée de la preuve : Fixons n. $\mathbb{P}(X+Y=n) = \sum_{k=0}^{n} \mathbb{P}(X=k)\mathbb{P}(X+Y=n/X=k)$
$= \sum_{k=0}^{n} \mathbb{P}(X=k)\mathbb{P}(Y=n-k) = \sum_{k=0}^{n} \exp(-\lambda)\frac{\lambda^k}{k!}\exp(-\mu)\frac{\mu^{n-k}}{(n-k)!}$
$= \exp(-(\lambda+\mu))\sum_{k=0}^{n} \frac{C_n^k}{n!}\lambda^k\mu^{n-k} = \exp(-(\lambda+\mu))\frac{(\lambda+\mu)^n}{n!}$. C.Q.F.D.

♡ b) Approximation Poissonienne d'une loi binômiale $\mathcal{B}(n,p)$

Règle utilisée en statistiques :
Soit X une variable aléatoire possèdant la loi $\mathcal{B}(n,p)$. Si p est petit (disons ≤ 0.10) et n grand, les valeurs numériques des probabilités $\mathbb{P}(X=k)$ sont raisonnablement voisines de $\exp(-np)\frac{(np)^k}{k!}$.
Cela signifie que X se comporte approximativement comme si elle avait une loi de Poisson de paramètre np.

Un exemple numérique.

On lance 100 fois deux dés non truqués et on appelle X le nombre de fois qu'un double six est sorti. La loi exacte de X est $\mathcal{B}(100, \frac{1}{36})$ qui conduit à des calculs compliqués.
Dans le tableau suivant, on indique les valeurs numériques exactes à 10^{-4} près des p_k (obtenues par un programme informatique) et les approximations $p'_k = \exp(-\frac{100}{36})\frac{(\frac{100}{36})^k}{k!}$.

k	0	1	2	3	4	5
p_k	0.0596	0.1705	0.2414	0.2255	0.1564	0.0858
p'_k	0.0620	0.1725	0.2397	0.2221	0.1544	0.0858

Compléments hors programme:

1. Calcul de l'espérance d'une variable aléatoire de loi hypergéométrique

L'espérance est la même que dans le cas binômial :
nombre de tirages $\times$ probabilité de tirer une boule bleue au premier coup.

On reprend les notations du paragraphe I-5 d). On pose $N = r + b$. Numérotons de 1 à b les boules bleues et définissons pour tout k compris entre 1 et b l'événement:
$E_k = \{$ on a tiré parmi les n boules la boule bleue k $\}$.
Comme le nombre total X de boules bleues tirées est $X = \sum_{k=1}^{b} \mathbb{1}_{E_k}$, par linéarité de l'espérance, $\mathbb{E}(X) = b\,\mathbb{P}(E_1)$.
Évaluons maintenant $\mathbb{P}(E_1)$. En passant au complémentaire,

$$\mathbb{P}(\overline{E_1}) = \frac{C_{N-1}^n}{C_N^n} = \frac{(N-1)!}{n!(N-1-n)!}\frac{n!(N-n)!}{N!} = \frac{N-n}{N}$$

qui est la probabilité de ne jamais tirer une boule donnée.
Donc $\mathbb{P}(E_1) = 1 - \frac{N-n}{N} = \frac{n}{N} = \frac{n}{r+b}$.

On en conclut que $\mathbb{E}(X) = \frac{nb}{r+b}$. C.Q.F.D.

2. Fonction génératrice d'une va. à valeurs dans $\mathbb{N}$

♡ **Définition.** Si N est une variable aléatoire qui prend ses valeurs dans $\mathbb{N}$, sa fonction génératrice est la fonction g de $[0,1]$ dans $\mathbb{R}^+$ telle que $s \longrightarrow g(s) = \mathbb{E}(s^N)$.

Remarques. a) On trouve donc $g(s) = \sum_{n=0}^{\infty} \mathbb{P}(N = n)s^n$. Il est clair que cette série à termes positifs converge puisqu'on s'est restreint aux s dans $[0,1]$.
D'ailleurs par définition $g(1) = 1$.
b) Si on veut rappeler qu'elle dépend de N, on note cette fonction g_N.
c) On rencontre aussi la « fonction génératrice de moments » de la va réelle X. C'est une autre fonction : $s \longrightarrow \mathbb{E}(\exp(sX))$.

Exemples. a) Si N suit la loi $\mathcal{B}(n,p)$, $g(s) = (ps+q)^n$.
b) Si N suit la loi $\mathcal{P}(\lambda)$, $g(s) = \exp(-\lambda\,(1-s))$.
c) Si N suit la loi $\mathcal{G}^*(p)$, $g(s) = \frac{ps}{1-qs}$.

Proposition 1. La fonction génératrice ne dépend que de la loi de N. Si deux vas à valeurs dans $\mathbb{N}$ ont la même fonction génératrice, alors elles ont la même loi.

Théorème 2. Si N_1 et N_2 sont indépendantes, la fonction génératrice de $N_1 + N_2$ est le produit $g_{N_1} \times g_{N_2}$.

Corollaire. On retrouve ainsi les propriétés de stabilité V-3. et V-A7a) en identifiant le produit des fonctions génératrices.

Théorème 3. a) Si N possède une espérance finie, $g'(1) = \mathbb{E}(N)$.
b) Si $\mathbb{E}(N^2)$ est finie, $g''(1) = \mathbb{E}(N(N-1))$.

Application. On peut retrouver ainsi les espérances et les variances des lois binômiales, géométriques et de Poisson. C'est la même idée que la méthode de dérivation vue page 5 du chapitre 0.