Math ECG 2. 2023-2024

Mathématiques Appliquées - F. Gaunard
http://frederic.gaunard.com
ENC Bessières, Paris 17e.

Chapitre 7. Couples de variables aléatoires finies ou discrètes

Dans tout ce chapitre, les variables aléatoires considérées sont dites finies ou discrètes, c'est à dire que leur univers image $X(\Omega)$ est un (sous-) ensemble fini ou discret de $\mathbb{R}$. Un sous-ensemble discret D de $\mathbb{R}$ est un ensemble dont tous les éléments sont "espacés" les uns des autres : pour tout $x \in D$, il existe toujours un voisinage dans lequel le seul élément de D est x. Il n'y a pas besoin de comprendre cette définition. Pour nous, dans la plupart des cas $X(\Omega)$ sera une partie de $\mathbb{N}$.

Toutes les variables aléatoires considérées sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ qu'on ne cherche, comme d'habitude, pas à expliciter.

1 Couples aléatoires discrets

Définition

On appelle **couple aléatoire discret** une application $V : \Omega \mapsto \mathbb{R}^2$.
Ainsi, on peut écrire $V = (X, Y)$, où X et Y sont deux variables aléatoires discrètes, avec :

$$\forall \omega \in \Omega,\ V(\omega) = (X(\omega), Y(\omega))\,.$$

1.1 Loi conjointe

Définition

La **loi conjointe** du couple (X, Y) est la donnée de la fonction :

$$\begin{array}{rccl} \varphi_{X,Y} : & X(\Omega) \times Y(\Omega) & \to & [0;1] \\ & (i,j) & \mapsto & P\left([X=i] \cap [Y=j]\right). \end{array}$$

Exercice 1. Une urne contient 3 boules numérotées de 1 à 3. On tire successivement et sans remise deux boules dans cette urne et on note X le numéro de la première boule et Y le numéro de la deuxième boule.

Déterminer la loi conjointe de (X, Y).

Exercice 2. On lance une pièce qui, à chaque lancer, donne *Pile* avec probabilité p et *Face* avec probabilité $q = 1 - p$ (avec $p \in]0;1[$). On note X le rang d'apparition du premier *Pile*, Y le rang d'apparition du premier *Face* et Z le rang d'apparition du second *Pile*.

(1) Déterminer la loi conjointe de (X, Z).
(2) Déterminer la loi conjointe de (X, Y).

☞ Ces deux premiers exercices ont permis de voir notamment que $V(\Omega) \subsetneq X(\omega) \times Y(\Omega)$.

À retenir!

Pour obtenir la loi du couple (X, Y), on commence par déterminer $X(\Omega)$ et $Y(\Omega)$ puis on donne les valeurs de $P([X = i] \cap [Y = j])$ pour tout $i \in X(\Omega)$ et tout $j \in Y(\Omega)$ en décrivant l'évènement correspondant à l'observation *simultanée* de X et Y.
Lorsque $X(\Omega)$ et $Y(\Omega)$ sont finis, on peut résumer cette loi dans un tableau à double entrée.

☠ Attention, en toute généralité

$$P([X = i] \cap [Y = j]) \neq P([X = i]) \cdot P([Y = j]).$$

☞ On peut parfois utiliser la formule des probabilités composées

$$P([X = i] \cap [Y = j]) = P_{[Y=j]}([X = i])P([Y = j]).$$

Propriété

On a :

$$\sum_{i \in X(\Omega)} \sum_{j \in Y(\Omega)} P([X = i] \cap [Y = j]) = \sum_{j \in Y(\Omega)} \sum_{i \in X(\Omega)} P([X = i] \cap [Y = j]) = 1.$$

1.2 Lois marginales

Définition

(1) La première loi marginale du couple (X, Y) est la loi de la variable aléatoire X.
(2) La seconde loi marginale du couple (X, Y) est la loi de la variable aléatoire Y.

À retenir!

Soit (X, Y) un couple de variables aléatoires discrètes dont on connait la loi conjointe.

- On obtient la loi de X grâce à la formule des probabilités totales avec le s.c.e $\{[Y = j] : j \in Y(\Omega)\}$.

$$\begin{aligned} P([X = i]) &= \sum_{j \in Y(\Omega)} P([X = i] \cap [Y = j]) \\ &= \sum_{j \in Y(\Omega)} P([Y = j])P_{[Y=j]}([X = i]) \end{aligned}$$

- On obtient la loi de Y grâce à la formule des probabilités totales avec le s.c.e $\{[X = i] : i \in X(\Omega)\}$.

$$\begin{aligned} P([Y = j]) &= \sum_{i \in X(\Omega)} P([X = i] \cap [Y = j]) \\ &= \sum_{i \in X(\Omega)} P([X = i])P_{[X=i]}([Y = j]) \end{aligned}$$

☞ Si la loi du couple est résumée dans un tableau à double entrée, les lois marginales s'obtiennent en sommant les éléments de chaque ligne et de chaque colonne.

Exercice 3. Donner les lois marginales des variables aléatoires X, Y et Z de l'Exercice 2.

1.3 Lois conditionnelles

Définition

Soit (X,Y) un couple de variables aléatoires discrètes.

- Soit $j \in Y(\Omega)$ fixé tel que $P([Y=j]) \neq 0$.
  On appelle **loi conditionnelle de** X **sachant** $[Y=j]$ la donnée des valeurs, pour tout $i \in X(\Omega)$
  $$P_{[Y=j]}([X=i]).$$
- Soit $i \in X(\Omega)$ fixé tel que $P([X=i]) \neq 0$.
  On appelle **loi conditionnelle de** Y **sachant** $[X=i]$ la donnée des valeurs, pour tout $j \in Y(\Omega)$,
  $$P_{[X=i]}([Y=j]).$$

Exercice 4. On disposes de n urnes numérotées de 1 à n. Pour chaque $k \in [\![1,n]\!]$, l'urne k contient des boules, indiscernables au toucher, numérotées de 1 à k.
On choisit une urne au hasard et on pioche une boule dans cette urne. On note X le numéro de l'urne choisie et Y le numéro de la boule piochée.

(1) Reconnaître la loi de X. Rappeler la valeur de $E(X)$.
(2) Pour tout $k \in X(\Omega)$, déterminer la loi conditionnelle de Y sachant $[X=k]$.
(3) En déduire la loi marginale de Y. On exprimera le résultat sous forme d'une somme qu'on ne cherchera pas à calculer.
(4) Calculer $E(Y)$.
(5) Écrire une fonction Python d'en-tête `def simul_XY(n) :` qui renvoie une simulation du couple (X,Y).

À retenir!

On peut obtenir une loi conditionnelle de différentes manières suivants les situations:

- Si on connait la loi du couple (donc les probabilités $P([X=i]\cap[Y=j])$), on utilise la formule des probabilités composées. Par exemple, la loi conditionnelle de X sachant $[Y=j]$ est donnée par :
  $$P_{[Y=j]}([X=i]) = \frac{P([X=i]\cap[Y=j])}{P([Y=j])}, \quad \forall i \in X(\Omega).$$
- Si on ne connait pas la loi du couple, on peut souvent reconnaitre pour la loi conditionnelle une loi connue en utilisant l'information contenue dans l'évènement qui conditionne.

Exercice 5. On lance une pièce donnant pile avec probabilité $p \in]0;1[$. On note X le rang d'apparition du premier pile et si $[X=n]$, on note Y le nombre de faces obtenus lors de n lancers supplémentaires.

(1) Déterminer la loi conditionnelle de Y sachant $[X=n]$.
(2) Préciser $Y(\Omega)$ puis montrer que pour tout $k \in \mathbb{N}$,

$$P(Y=k) = \left(\frac{q}{p}\right)^{k-1} \sum_{n=k}^{+\infty} \binom{n}{k} (pq)^n.$$

2 Indépendance de variables aléatoires discrètes

2.1 Cas de deux variables

Définition

On dit que deux variables aléatoires discrètes X et Y sont **indépendantes** si et seulement si :

$$\forall (i,j) \in X(\Omega) \times Y(\Omega), \quad P([X=i] \cap [Y=j]) = P([X=i]) \times P([Y=j])$$

À retenir!

☞ Pour montrer que deux variables aléatoires X et Y **ne sont pas** indépendantes, il suffit de trouver un contre-exemple à la relation précédente.

On cherche alors souvent des évènements $[X=i]$ et $[Y=j]$ de probabilité non nulle et tels que l'évènement $[X=i] \cap [Y=j]$ est impossible donc de probabilité nulle.

Exercice 6. Montrer que les variables aléatoires (X, Y) de l'Exercice 1 ne sont pas indépendantes.

Exercice 7. On lance une pièce équilibrée n fois et on considère les variables aléatoires X et Y correspondant respectivement au nombre de *Pile* et de *Face* obtenus. X et Y sont-elles indépendantes?

2.2 Cas de n variables

Définition

- On dit que les variables aléatoires discrètes $X_1, \ldots, X_n$ sont **mutuellement indépendantes** si et seulement si
$$\forall (x_1, \ldots, x_n) \in \mathbb{R}^n, \quad P([X_1 = x_1] \cap \cdots \cap [X_n = x_n]) = \prod_{k=1}^{n} P([X_k = x_k]).$$
- On dit que $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires mutuellement indépendantes si pour tout $n \in \mathbb{N}$ les variables aléatoires $X_0, \cdots, X_n$ sont mutuellement indépendantes.

Question type de sujet de concours

On considère une suite (X_n) de variables mutuellement indépendantes de même loi géométrique $\mathcal{G}(p)$. On note $Z_n = \min(X_1, ..., X_n)$.

(1) Pour tout $k \in \mathbb{N}^*$, pour tout $j \in \mathbb{N}$, montrer que $P(X_j > k) = q^k$.

(2) Justifier que $P(Z_n > k) = P\left(\bigcap_{j=1}^{n} [X_j > k]\right)$.

(3) En déduire la loi de Z_n.

2.3 Lemme des coalitions

Propriété

Lemme des colations.

- Si X et Y sont deux variables aléatoires discrètes indépendante et si f et g sont deux fonctions numériques définies respectivement sur $X(\Omega)$ et $Y(\Omega)$ alors $f(X)$ et $g(Y)$ sont indépendantes.
- Soient $X_1, ..., X_n$ mutuellement indépendantes et soit $p \in [\![2, n-1]\!]$. Alors toute variable aléatoire fonction des variables $X_1, \cdots, X_p$ est indépendante de toute variable aléatoire fonction des variables $X_{p+1}, \cdots, X_n$.

Exemple

- Si X_1, X_2, X_3, X_4, X_5 sont 5 variables aléatoires discrètes discrètes mutuellement indépendantes alors les variables $X_1 - X_2 + 2X_4^2$ et $X_3 - X_5^2$ sont indépendantes.
- Si X^2 et Y^2 ne sont pas indépendantes, alors X et Y ne le sont pas non plus.

3 Fonctions de deux variables aléatoires discrètes

Dans cette section, on étudie des variables aléatoires du type $Z = g(X, Y)$ où (X, Y) est un couple de variables aléatoires et g une fonction de $X(\Omega) \times Y(\Omega)$ dans $\mathbb{R}$.

Exercice 8. On se place à nouveau dans le cadre de l'Exercice 1, c'est à dire qu'on pioche successivement et sans remise deux boules dans une urne qui en contient 3, numérotées de 1 à 3. La variable aléatoire X (resp. Y) correspond au numéro de la première (resp. deuxième) boule tirée.

(1) Déterminer la loi de $S = X + Y$.
(2) Déterminer la loi de $D = X - Y$.

Exercice 9. Soient $X \hookrightarrow \mathcal{B}(p)$ et $Y \hookrightarrow \mathcal{B}(q)$ deux variables de Bernoulli indépendantes. Déterminer la loi de XY.

À retenir!

On obtient la **loi de la somme** (ou de la différence) deux variables aléatoires discrètes X et Y par application de la formule des probabilités totales avec au choix le s.c.e associé à X ou à Y.

$$\begin{aligned} P(X+Y=k) &= \sum_{i \in X(\Omega)} P([X+Y=k] \cap [X=i]) \\ &= \sum_{i \in X(\Omega)} P([Y=k-i] \cap [X=i]) \end{aligned}$$

et on utilise la loi conjointe de (X, Y) pour poursuivre le calcul.

☞ La probabilité $P([X=i] \cap [Y=k-i])$ est nulle dès que $k-i$ n'appartient pas à $Y(\Omega)$, ce qui a pour conséquence de restreindre les indices de la somme.

Exercice 10. Soient X et Y deux variables indépendantes suivant toutes deux la loi géométrique $\mathcal{G}(p)$.

(1) Déterminer, en utilisant un s.c.e associé à X, $P(X=Y)$.
(2) Montrer, en utilisant un s.c.e associé à X que pour tout $n \geq 2$: $P(X+Y=n)=(n-1)p^2q^{n-2}$.

Propriété

Stabilité par somme de variables aléatoires indépendantes usuelles.

- **Somme de lois de Bernoulli.** Si $X_1, X_2, ..., X_n$ sont n variables aléatoires mutuellement indépendantes suivant une même loi de Bernoulli de paramètre p, alors
$$\sum_{k=1}^{n} X_k \hookrightarrow \mathcal{B}(n,p).$$
- **Stabilité des lois binomiales.** Soient $X \hookrightarrow \mathcal{B}(n,p)$ et $Y \hookrightarrow \mathcal{B}(m,p)$ deux lois binomiales indépendantes. Alors,
$$X+Y \hookrightarrow \mathcal{B}(n+m,p).$$
- **Stabilité des lois de Poisson.** Soient $X \hookrightarrow \mathcal{P}(\lambda)$ et $Y \hookrightarrow \mathcal{P}(\mu)$ deux lois de Poisson indépendantes. Alors,
$$X+Y \hookrightarrow \mathcal{P}(\lambda+\mu).$$

Remarque

Pour généraliser ces résultats à des sommes de n variables (mutuellement) indépendantes et de même loi usuelle, on procède par récurrence avec notamment l'aide cruciale du **lemme des coalitions**.

Exercice 11. Soient $X_1, \ldots, X_k$ des variables aléatoires mutuellement indépendantes. Déterminer la loi de $X_1+\cdots+X_k$ lorsque $X_1 \hookrightarrow \mathcal{P}(\lambda_1)$, $\ldots$, $X_k \hookrightarrow \mathcal{P}(\lambda_k)$.

4 Calculs d'espérance

4.1 Espérance de $Z=g(X,Y)$ - Théorème de transfert

Le théorème de transfert, outil fondamental terriblement sous-estimé et incompris par un trop grand nombre d'étudiant.e.s, permet de calculer l'espérance d'une variable aléatoire de la forme $Z=g(X,Y)$ sans connaître la loi de Z mais uniquement à partir de la loi du couple (X,Y).

Propriété

Théorème de transfert.

Soit (X,Y) deux variables aléatoires discrètes. Alors, on a (sous réserve de convergence absolue) :
$$E(Z)=E(g(X,Y))=\sum_{i\in X(\Omega)}\sum_{j\in Y(\Omega)} g(i,j)P([X=i]\cap[Y=j]).$$

Remarque

Si $X(\Omega)$ et $Y(\Omega)$ sont deux ensembles finis, il n'y a pas de problème de convergence (absolue), puisqu'il s'agit alors d'une somme (double) finie.

Dans le cas où (au moins) un des deux ensembles est infini, le caractère licite de l'écriture de la somme ci-dessus est conditionné par la convergence absolue d'une série à double indice et présente donc un problème assez subtil. Les sujets des problèmes de concours faisant intervenir ce type de somme proposent donc un guidage étape par étape pour en justifier l'existence d'une telle somme ou bien admettent carrément la convergence de celle-ci.

4.2 Espérance d'une somme

Propriété

Linéarité de l'espérance

- Soient $X_1, \dots, X_n$ des variables aléatoires discrètes admettant une espérance, $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$. Alors :
$$E(\lambda_1 X_1 + \cdots + \lambda_n X_n) = \sum_{i=1}^{n} \lambda_i E(X_i).$$
- Lorsque les constante sont égales à 1, on obtient le cas particulier important :
$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i).$$

Exercice 12. Soient $X_1, \dots, X_n$ des variables aléatoires de même loi admettant toutes une (même) espérance μ. Déterminer l'espérance de la variable aléatoire
$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

4.3 Espérance d'un produit

Règle(s) de calcul

Soit (X, Y) un couple de deux variables aléatoires discrètes. Alors, le théorème de transfert permet d'écrire, sous réserve de convergence absolue,
$$E(XY) = \sum_{i \in X(\Omega)} \sum_{j \in Y(\Omega)} ij P([X = i] \cap [Y = j]).$$

Propriété

Si X et Y sont **indépendantes** et admettent une espérance, alors XY admet une espérance, et on a
$$E(XY) = E(X)E(Y).$$

☞ On peut utiliser le théorème précédent pour montrer que deux variables aléatoires discrètes **ne sont pas** indépendantes en vérifiant que $E(XY) \neq E(X)E(Y)$.

Exercice 13. Soient X_1, X_2, et X_3, indépendantes, suivant des lois de Bernoulli de paramètre p. Montrer que les variables aléatoires $Y_1 = X_1 X_2$ et $Y_2 = X_2 X_3$ ne sont pas indépendantes.

5 Variance et covariance

5.1 Covariance

Définition

Soient (X, Y) deux variables aléatoires discrètes admettant des moments d'ordre 2.
La **covariance** de X et Y est définie par :

$$\operatorname{cov}(X, Y) = E\left((X - E(X))\,(Y - E(Y))\right).$$

Règle(s) de calcul

Formule de König-Huygens

Soient (X, Y) deux variables aléatoires discrètes admettant des moments d'ordre 2. Alors :

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

En particulier, on a :

$$\operatorname{cov}(X, X) = V(X)$$

Exercice 14. Déterminer la covariance des variables X et Y de l'Exercice 1.

Exercice 15. Une urne contient n boules blanches numérotées de 1 à n et 2 boules noires numérotées 1 et 2. On tire une à une, sans remise, toutes les boules de l'urne et on définit deux variables aléatoires X et Y correspondant respectivement au rang d'apparition de la première boule blanche et de la boule numérotée 1.

(1) Déterminer les lois de X et Y.
(2) Les variables X et Y sont-elles indépendantes? Calculer $\operatorname{cov}(X, Y)$.

Propriété

Propriétés de la covariance.

- Symétrie : $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$.
- Linéarité à gauche : $\forall(\lambda, \mu) \in \mathbb{R}^2 \quad \operatorname{cov}(\lambda X_1 + \mu X_2\,,\, Y) = \lambda \operatorname{cov}(X_1\,,\, Y) + \mu \operatorname{cov}(X_2\,,\, Y)$.
- Linéarité à droite : $\forall(\lambda, \mu) \in \mathbb{R}^2 \quad \operatorname{cov}(X\,,\, \lambda Y_1 + \mu Y_2) = \lambda \operatorname{cov}(X\,,\, Y_1) + \mu \operatorname{cov}(X\,,\, Y_2)$.

Exercice 16. Soient X et Y deux variables aléatoires discrètes admettant un moment d'ordre 2.

(1) Soit $t \in \mathbb{R}$. Montrer que

$$V(X + tY) = V(X) + 2\operatorname{cov}(X, Y)t + V(Y)t^2$$

(2) On note $P : \; t \mapsto V(X + tY)$. Justifier que, pour tout $t \in \mathbb{R}$, $P(t) \geq 0$. En déduire, à l'aide d'une considération sur un discriminant, que

$$\left|\frac{\operatorname{cov}(X, Y)}{V(X)V(Y)}\right| \leq 1.$$

À connaître sur le bout des doigts

Si X et Y sont indépendantes, alors : $\operatorname{cov}(X, Y) = 0$.

☠ **La réciproque est fausse !**

5.2 Variance d'une somme

Propriété

- Si X et Y admettent une variance, alors $X+Y$ également, et :
$$V(X+Y)=V(X)+V(Y)+2\cdot\operatorname{cov}(X,Y).$$
- Si X et Y sont indépendantes, et admettent une variance, alors $X+Y$ aussi, et on a :
$$V(X+Y)=V(X)+V(Y).$$
- Si les variables $X_1,\dots,X_n$ sont mutuellement indépendantes, et admettent une variance, alors la variable $X_1+\cdots+X_n$ admette une variance et on a :
$$V(X_1+\cdots+X_n)=V(X_1)+\cdots+V(X_n).$$

Exercice 17. Soient $X_1,\dots,X_n$ des variables aléatoires indépendantes admettant toutes une même variance σ^2. Déterminer la variance de la variable aléatoire
$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i.$$

5.3 Coefficient de corrélation linéaire

Définition

Soit (X,Y) un couple de variables aléatoires discrètes admettant des variances non nulles. Le **coefficient de corrélation linéaire** de X et Y est la quantité définie par $\rho_{X,Y}=\dfrac{\operatorname{cov}(X,Y)}{\sigma(X)\sigma(Y)}$.

Propriété

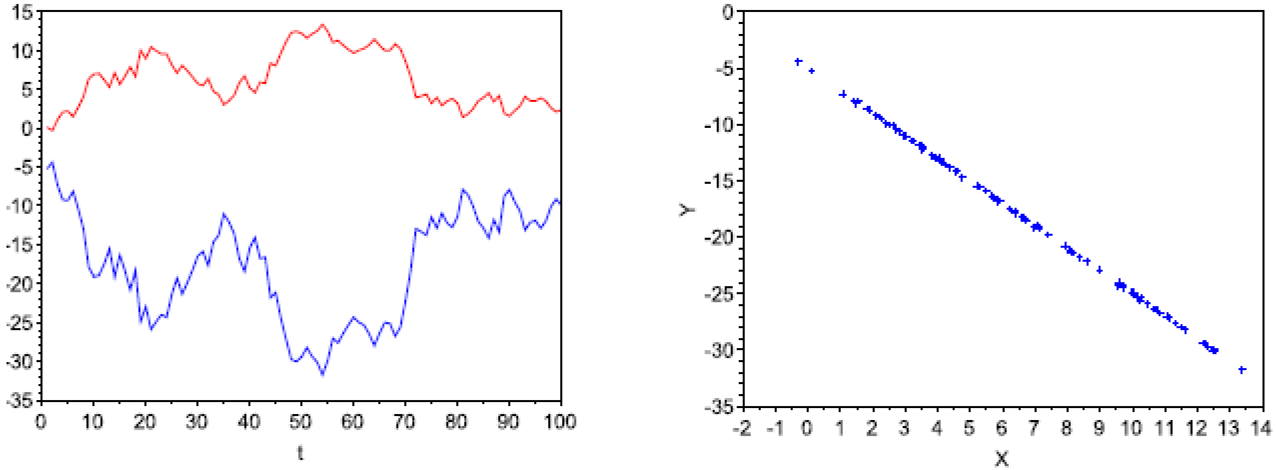
On a obtenu dans l'Exercice 16 une inégalité qui se reformule comme suit.
Le coefficient de corrélation linéaire de X et Y est un réel compris entre -1 et 1 :
$$-1\leq\rho_{X,Y}\leq 1.$$

Propriété

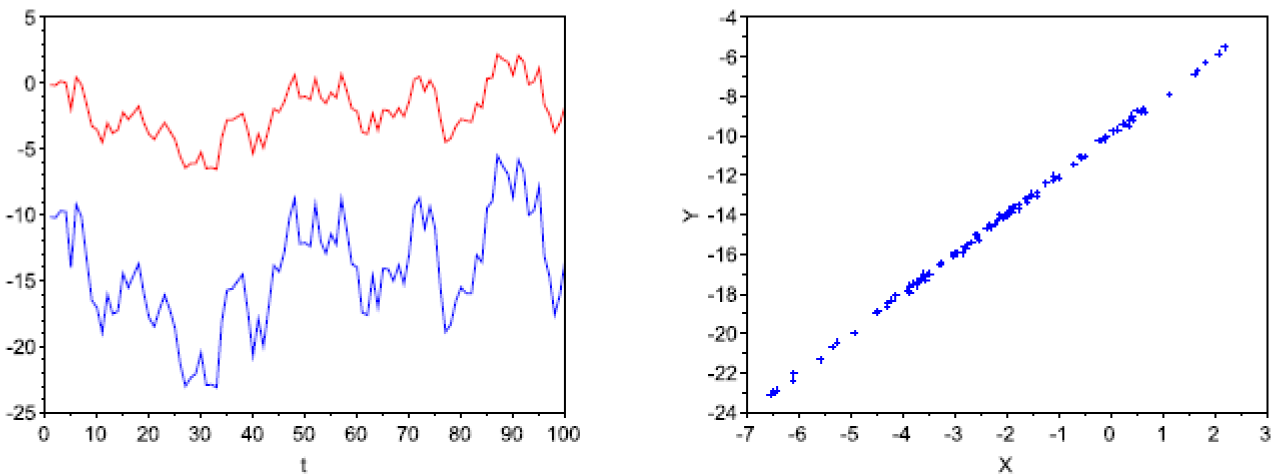
Le coefficient de corrélation linéaire de X et Y est un réel compris entre -1 et 1 :
$$-1\leq\rho_{X,Y}\leq 1$$

- Il est égal à 1 dans le cas où l'une des variables est fonction affine croissante de l'autre variable.
- Il est égal à -1 dans le cas où l'une des variables est fonction affine décroissante de l'autre variable.
- Les valeurs intermédiaires renseignent sur le degré de dépendance linéaire entre les deux variables.
  - Plus le coefficient $\rho_{X,Y}$ est proche des valeurs extrêmes -1 et 1, plus la corrélation entre les variables est forte.
  - Deux variables aléatoires dont la covariance est nulle (et donc le coefficient de corrélation linéaire est nul) sont dites non corrélées.
  - Une corrélation positive ($\rho_{X,Y}>0$) indique que les variables X et Y varient dans le même sens.
  - Une corrélation négative ($\rho_{X,Y}<0$) indique que les variables X et Y varient en sens inverse.

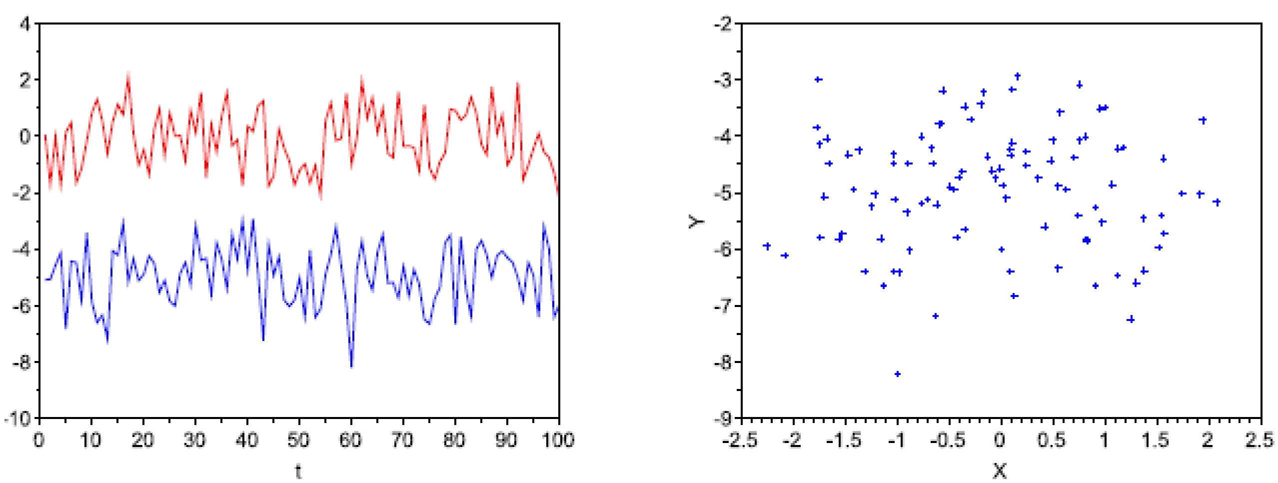
☞ Ces notions seront illustrées notamment au cours du TP 2.



Corrélation négative



Corrélation positive



Corrélation nulle

6 Sélection d'exercices - Travaux dirigés

Exercice 18. Soient X et Y deux variables aléatoires dont on suppose que la loi du couple est donnée par le tableau

$X \backslash Y$	1	2	3
1	2β	3β	3β
2	3β	2β	3β
3	3β	3β	2β

(1) Déterminer la valeur du réel β pour que ce tableau représente effectivement la loi d'un couple.
(2) Expliciter les lois marginales de X et Y. En déduire la valeur de $E(X)$ et $E(Y)$.
(3) Calculer $E(XY)$ puis vérifier que

$$\operatorname{Cov}(X,Y) = -\frac{1}{12}.$$

Les variables X et Y sont-elles indépendantes?

Exercice 19. Une urne U_1 contient n boules numérotées de 1 à n. Une urne U_2 contient des boules rouges en proportion p. On tire une boule au hasard dans U_1 et on note X la variable aléatoire correspondant au numéro de la boule obtenue. Si $X = k$, on tire k fois, avec remise, une boule dans l'urne U_2 et on appelle Y le nombre de boules rouges tirées.

(1) Déterminer la loi de X, son espérance et sa variance.
(2) Déterminer la loi conditionnelle de Y sachant $X = k$, pour $k \in X(\Omega)$.
(3) Déterminer la loi du couple (X, Y).
(4) Déterminer l'espérance et la variance de Y.

Exercice 20. Soient X et Y telles que $X \hookrightarrow \mathcal{B}(n,p)$ et telles que la loi conditionnelle de Y, sachant $X = k$ (où $k \in X(\Omega)$) est une loi binomiale de paramètre $n-k$ et p. Quelle est la loi de Y?

Exercice 21. On place au hasard deux paires de chaussettes de couleurs différentes dans deux tiroirs notés t_1 et t_2. Un tiroir peut donc contenir de 0 à 2 paires.

On note X_1 et X_2 respectivement les variables aléatoires égales au nombre de paires de chaussettes contenues dans les tiroirs t_1 et t_2 ainsi que N le nombre de tiroirs occupés.

On note alors T_i le numéro du tiroir choisi pour la paire de chaussettes numéro i. Les paires de chaussettes étant rangées dans les tiroirs indépendamment les unes des autres, on suppose donc que les variables T_1 et T_2 sont indépendantes.

(1) Préciser les lois de T_1 et T_2.
(2) À l'aide des variables T_i, déterminer la loi conjointe du couple (N, X_1) que l'on présentera en recopiant et complétant le tableau suivant.

$X_1 \backslash N$	0	1	2
1			
2			

(3) En déduire les lois marginales de N et X_1. Que dire de l'indépendance de N et X_1?
(4) Calculer $E(N)$, $E(X_1)$ puis $\operatorname{cov}(N, X_1)$.
(5) Que peut-on dire du couple (N, X_2) ?

Exercice 22. (Extrait de **DM n°4, Automne 2021**)

Un immeuble de p étages est équipé d'un ascenseur. N personnes montent dans l'ascenseur au rez de chaussée et descendent chacune à un étage au hasard et de façon indépendante.
Soit X le nombre d'arrêts de l'ascenseur.
On note X_i la variable valant 1 si l'ascenseur s'arrête à l'étage i et 0 sinon et on note E_k la variable aléatoire qui prend la valeur de l'étage où descend la $k-$ième personne.

(1) Pour tout $k \in [\![1; N]\!]$, quelle est la loi de E_k?
(2) En utilisant les variables aléatoires E_k, déterminer $P(X_i = 0)$.
(3) Calculer $E(X)$.
(4) Déterminer pour i différent de j : $P(X_i = 0 \cap X_j = 0)$.
En déduire $P(X_iX_j = 0)$.
(5) (*) Déduire de la question précédente $\mathrm{Cov}(X_i, X_j)$ puis calculer alors $V(X)$.

Exercice 23. Une urne contient des boules numérotées de 1 à n (où $n \in \mathbb{N}, n \geq 2$). On effectue deux tirages **sans remise** dans cette urne et on note X (resp. Y) le plus petit (resp. grand) numéro obtenu.

(1) Déterminer les lois de X et Y ainsi que $E(X)$ et $E(Y)$.
(2) Montrer que

$$P(X = i \cap Y = j) = \begin{cases} \frac{2}{n(n-1)}, & \text{si } 1 \leq i < j \leq n \\ 0, & \text{sinon} \end{cases}$$

(3) Montrer que

$$E(XY) = \frac{(n+1)(3n+2)}{12}$$

et en déduire la $\mathrm{Cov}(X, Y)$.
(4) Écrire une fonction Python qui permette de simuler le couple (X, Y). On commencera par simuler la pioche successive sans remise de deux boules.

Exercice 24. Un binôme de deux personnes nommées A et B participent à une épreuve physique. Ces deux personnes doivent grimper une corde. Une fois que l'une des deux personnes a réussi, elle doit attendre que l'autre personne en fasse de même. On considère que

- A et B disposent chacun de leur propre corde.
- A et B ont droit à autant d'essais qu'ils le souhaitent.
- Les essais sont indépendants.
- Chaque essai, qu'il soit réussi ou non, dure une minute et est réussi avec probabilité p.

On note X_1 (resp. X_2) le nombre d'essais nécessaires à A (resp. B) pour grimper la corde, et Y la variable aléatoire réelle égale à $|X_1 - X_2|$.

(1) Quelle est la loi de X_1? De X_2? Donner leur espérance et leur variance.
(2) Que représente l'événement $[Y = 0]$? Déterminer sa probabilité.
(3) Montrer que pour tout on a, pour tout $k \in \mathbb{N}^*$,

$$P(Y = k) = \frac{2p(1-p)^k}{2-p}.$$

(4) Écrire un programme Python permettant de simuler la variable aléatoire Y.
(5) Pour quelles valeurs de p les deux personnes s'attendront-elles en moyenne moins de 5 minutes ?

Exercice 25. Un auto-stoppeur attend au péage d'une autoroute pendant une certaine période. On admet que le nombre de véhicules franchisant le péage pendant cette période est une variable aléatoire suivant la loi de Poisson de paramètre $\lambda > 0$. A chaque fois qu'un véhicule franchit le péage, il lance une pièce truquée qui donne pile avec la probabilité p. On pose $q = 1 - p$.

On note N, X et Y respectivement le nombre de lancers, le nombre de piles obtenus et le nombre de face obtenus.

(1) (a) Pour tout couple (i,j) d'entiers naturels, calculer la probabilité $P_{(N=j)}(X=i)$.
(b) En déduire la loi du couple (X,N).
(c) Montrer que X suit une loi de Poisson de paramètre λp.

(2) Expliquer sans aucun calcul pourquoi Y suit une loi de Poisson de paramètre λq.

(3) Montrer que X et Y sont indépendantes.

(4) (a) Montrer que : $V(Y) = V(N) + V(X) - 2\text{cov}(N,X)$.
(b) En déduire $\text{cov}(N,X)$.
(c) Donner de même $\text{cov}(N,Y)$.

Exercice 26. (Extrait de **EML 2013**)

On considère une urne $\mathcal{U}$ contenant n boules numérotées de 1 à n et indiscernables au toucher ($n \geq 2$).On effectue une suite de tirages d'une boule avec remise de la boule dans l'urne $\mathcal{U}$.

Soit k un entier supérieur ou égal à 1. Pour tout $i \in [\![1;n]\!]$, on note X_i la variable aléatoire égale au nombre d'obtentions de la boule numéro i au cours des k premiers tirages.

(1) Soit $i \in [\![1;n]\!]$. Donner la loi de X_i. Rappeler l'espérance et la variance de X_i.

(2) Les variables aléatoires $X_1, X_2, \ldots, X_n$ sont-elles indépendantes?

(3) Soit $(i,j) \in [\![1;n]\!]^2$ tel que $i \neq j$.
(a) Déterminer la loi de la variable $X_i + X_j$. Rappeler la variance de $X_i + X_j$.
(b) En déduire la covariance du couple (X_i, X_j).

Exercice 27. (Un exercice **incontournable**).

On effectue une succession infinie de lancers indépendants d'une pièce donnant pile avec probabilité p. On note $q = 1 - p$.
On dit que la première série est de longueur $n > 1$ si les n premiers lancers ont amené le même côté de la pièce et le $(n+1)$-ième l'autre côté. De même la deuxième série commence au lancer suivant la fin de la première série et se termine au lancer précédant un changement de côté.

Soient X_1 et X_2 les variables aléatoires égales aux longueurs de la première et deuxième série.

(1) Déterminer la loi de X_1.

(2) Montrer que X_1 admet une espérance et que $E(X_1) = \dfrac{p}{q} + \dfrac{q}{p}$.

(3) Déterminer la loi du couple (X_1, X_2).

(4) En déduire la loi de X_2.

(5) Montrer que X_2 admet une espérance et que $E(X_2) = 2$.

(6) Montrer que $E(X_1X_2)$ existe et que $E(X_1X_2) = \dfrac{1}{q} + \dfrac{1}{p}$.

(7) On suppose que $p = 1/2$. Montrer que X_1 et X_2 sont indépendantes.

(8) On suppose que $p \neq 1/2$.
   (a) En considérant $P([X_1 = 1] \cap [X_2 = 1])$ montrer que X_1 et X_2 ne sont pas indépendantes.
   (b) Vérifier que $\text{cov}(X_1, X_2) = 4 - \dfrac{1}{pq}$ puis en déduire une nouvelle preuve que X_1 et X_2 ne sont pas indépendantes.

Exercice 28. (Extrait de **ECRICOME 2016**)

Soient X et Y deux variables aléatoires définies sur le même espace probabilisé et à valeurs dans $\{1, ..., N\}$. On dit que X et Y sont *échangeables* si, pour tous entiers $i, j \in \{1, ..., N\}$

$$P((X = i) \cap (Y = j)) = P((X = j) \cap (Y = i)).$$

(1) Montrer que si X et Y sont échangeables, alors elles suivent la même loi.

(2) Montrer que si X et Y sont indépendantes et de même loi, alors elles sont échangeables.

(3) L'exemple suivant vise à montrer que la réciproque est fausse. Soient n, b, c trois entiers strictement positifs. Une urne contient initialement n boules noires et b boules blanches. On pioche une boule dans l'urne. On définit une première variable aléatoire X comme suit. Si la boule est est noire, alors $X = 1$ sinon $X = 2$.

   (a) Écrire en langage Python, une fonction simulant la variable X.
   (b) Déterminer la loi de X.

   On remet la boule tirée dans l'urne ainsi que c boules supplémentaires de la couleur tirée. On tire une seconde boule et on introduit la variable aléatoire Y qui vaut 1 si la nouvelle boule tirée est noire et 2 sinon.

   (c) Déterminer la loi de Y.
   (d) Montrer que X et Y sont échangeables.
   (e) Sont-elles indépendantes?

Exercice 29. On considère une suite (X_n) de variables de Bernoulli, $X_n \hookrightarrow \mathcal{B}(p_n)$, un réel $\lambda \in]0; 1[$ et on suppose que

$$P_{[X_n=1]}(X_{n+1} = 1) = p_n, \qquad P_{[X_n=0]}(X_{n+1} = 1) = \lambda p_n.$$

(1) Montrer que

$$p_{n+1} = (1 - \lambda)p_n^2 + \lambda p_n.$$

(2) Calculer $\text{Cov}(X_n, X_{n+1})$ en fonction de p_{n+1} et p_n, puis en fonction de p_n et λ.
(3) Les variables X_n et X_{n+1} sont-elles indépendantes?

Exercice 30. On met n jetons numérotées au hasard dans n boites numérotées (un seul jeton dans chaque boite).
Soit X_k la variable valant 1 si le k-ième jeton est dans la k-ième boite et 0 sinon.
Soit X le nombre de coïncidences entre les numéros des jetons et des boites.

(1) Déterminer la loi de X_k.
(2) Pour i différent de j, déterminer la loi de X_iX_j
(3) Exprimer alors $\text{Cov}(X_i, X_j)$.
(4) (**) Déterminer l'espérance et la variance de X.

Exercice 31. (Extrait de **DM n°5**, Automne 2022. Inspiré par HEC/ESSEC sujet 0)

Dans tout l'exercice n désigne un entier supérieur ou égal à 2 et p un réel de l'intervalle $]0;1[$.
Soit $S = \{s_1, ..., s_n\}$ un ensemble fini de n sommets. On s'intéresse dans cet exercice à des graphes aléatoires construits à partir de l'ensemble S.
Plus précisément, pour tout couple $(i,j) \in [\![1,n]\!]^2$ avec $i < j$, on introduit les variables aléatoires $T_{i,j}$ mutuellement indépendantes de même loi $\mathcal{B}(p)$. Les arêtes du graphe sont les paires de sommets $\{s_i, s_j\}$ pour lesquelles $T_{i,j} = 1$.
On dit qu'un sommet est isolé s'il n'y a aucune arête incidente à ce sommet.
On introduit N_n la variable aléatoire égale au nombre d'arêtes du graphe. Pour $k \in [\![1,n]\!]$, on note D_k la variable aléatoire qui prend pour valeur le degré du sommet s_k (c'est à dire le nombre d'arêtes incidentes à ce somment).
Enfin, on note X_k la variable aléatoire qui vaut 1 si le sommet s_k est isolé et 0 sinon puis Z_n celle égale au nombre de sommets isolés du graphe.

(1) Recopier et compléter la fonction Python ci-dessous qui renvoie la liste d'adjacence d'un tel graphe aléatoire en prenant pour argument la liste des sommets `S` et p.

```
import numpy as np
import numpy.random as rd

def list_adj(S, p):
    n=.....
    l=[ [ ] for k in range(n)]
    for i in range(n-1):
        for j in range(i+1, n):
            if rd.random()<p :
                l[i].append(.....)
                l[j].append(.....)
    return l
```

(2) On exécute alors la commande suivante qui génère l'affichage suivant

```
list_adj('abcdef', 1/3)
```

Affichage Python

```
>>>
[['b', 'e', 'f'], ['a', 'c', 'f'], ['b'], [], ['a'], ['a', 'b']]
```

(a) Donner une représentation graphique du graphe aléatoire généré par cette exécution.
(b) Ce graphe contient-il des sommets isolés? Si oui, combien ?
(c) Expliciter la matrice d'adjacence de ce graphe. Quelle propriété a-t-elle ? Pourquoi ?

(3) Écrire une fonction Python d'en-tête `def nb_som_is(l):` qui, prenant en argument la liste `l` d'adjacence d'un graphe renvoie le nombre de sommets isolés de celui-ci.

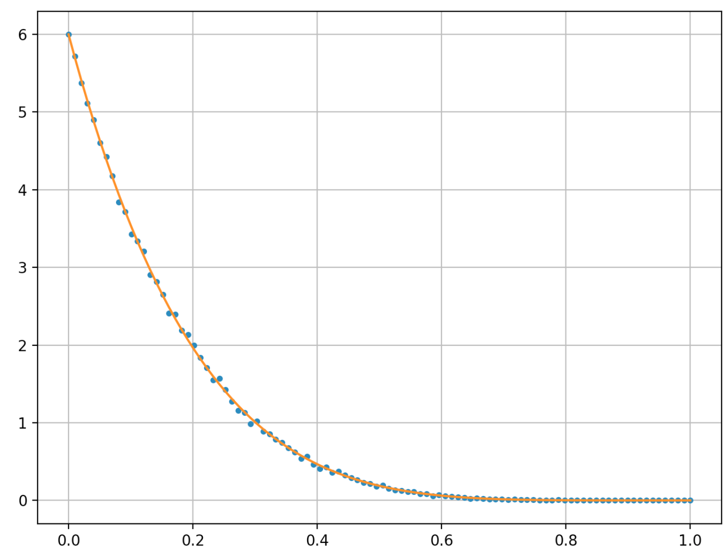
(4) On dispose de la fonction mystère suivante

```
def mystere(S, p):
    return np.mean([nb_som_is(list_adj(S,p)) for k in range(1000)])
```

(a) Que renvoie-t-elle ?
(b) On ajoute les instructions suivantes dont l'exécution produit la figure ci-contre. Interpréter et émettre une conjecture.

```
S='abcdef'
x=np.linspace(0,1, 100)
y=[mystere(S, p) for p in x]
w=[len(S)*(1-p)**(len(S)-1) for p in x]
plt.grid()
plt.plot(x,y, '.')
plt.plot(x,w)
plt.show()
```

Affichage Python



(5) (a) Justifier que $N_n(\Omega) \subset \llbracket 0, \binom{n}{2} \rrbracket$.

(b) Montrer que $P(N_n = 0) = P\left(\bigcap_{i=1}^{n-1} \bigcap_{j=i+1}^{n} [T_{i,j} = 0]\right) = (1-p)^{n(n-1)/2}$.

(c) Que vaut $P\left(N_n = \binom{n}{2}\right)$?

(6) (a) Justifier que, pour tout $k \in \llbracket 1, n \rrbracket$, on a

$$D_k = \sum_{i=1}^{k-1} T_{i,k} + \sum_{i=k+1}^{n} T_{k,i}.$$

En déduire la loi de D_k.

(b) Montrer que, pour $1 \leq k < \ell \leq n$, on a

$$\operatorname{cov}(D_k, D_\ell) = p(1-p).$$

Les variables D_k et D_ℓ sont-elles indépendantes ?

(c) Déterminer $E(Z_n)$. Est-ce cohérent avec la conjecture précédente?

(7) (a) Montrer que, pour tous $i < j$, on a

$$P([X_i = 1] \cap [X_j = 1]) = (1-p)^{2n-3}.$$

(b) En observant que

$$Z_n^2 = \sum_{i=1}^{n} X_i + 2 \sum_{1 \leq i < j \leq n} X_i X_j,$$

montrer que
$$E(Z_n^2) = n(1-p)^{n-1} + n(n-1)(1-p)^{2n-3}.$$

7 Une première annale

Cet exercice fait partie du **Concours Blanc** *que j'ai posé à l'automne 2019.*

Soit n un entier naturel supérieur ou égal à 3.
Une urne contient une boule noire non numérotée et $n-1$ boules blanches, dont $n-2$ portent le numéro 0 et une porte le numéro 1. On extrait ces boules au hasard, une à une, **sans remise**, jusqu'à l'apparition de la boule noire.

Pour chaque i de $[\![1, n-1]\!]$, on note B_i l'événement : "le i-ème tirage donne une boule blanche", et on pose $\overline{B_i} = N_i$. Enfin, on note X la variable aléatoire égale au rang d'apparition de la boule noire et Y la variable aléatoire qui vaut 1 si la boule numérotée 1 a été piochée lors de l'expérience précédente, et qui vaut 0 sinon.

Partie I - Simulation informatique

(1) Compléter la fonction Python suivant afin qu'il simule l'expérience aléatoire décrite dans cet exercice et renvoie les valeurs prises par les variables X et Y.
On admettra que la boule noire est codée tout au long de ce script par le nombre `nB+1`, où `nB` désigne le nombre de boules blanches.

```
import numpy as np
import numpy.random as rd

def simul_XY(n) :
    nB=n-1
    x=1
    y=...............
    u=rd.randint(1,nB+2)
    while u<nB+1
            nB=..........
            if u==1 :
                y=.............
            u=rd.randint( ... , ... )
            x=............
    return ....
```

(2) Écrire une suite de commandes permettant, pour un n quelconque, d'obtenir la *covariance empirique* d'un échantillon de taille 1000 du couple (X, Y).
(3) En rentrant $n = 4$, le programme affiche `0.449475`. Que peut-on conjecturer?

Partie II - Loi conjointe de X et Y

(5) Donner l'ensemble $X(\Omega)$ des valeurs que peut prendre la variable X.

(6) (a) Pour tout i de $[\![2, n-1]\!]$, justifier que
$$P_{B_1 \cap \ldots \cap B_{i-1}}(B_i) = \frac{n-i}{n-i+1}.$$
(b) Utiliser la formule des probabilités composées pour trouver $P(X = k)$, pour tout k de $X(\Omega)$.
(c) Reconnaître la loi de X et donner son espérance et sa variance.

(7) (a) Pour tout k de $X(\Omega)$, montrer, toujours grâce à la formule des probabilités composées, que

$$P\Big([X=k]\cap[Y=0]\Big)=\frac{n-k}{n(n-1)}.$$

(b) En déduire $P(Y=0)$ puis $P([X=k]\cap[Y=1])$.

(c) Reconnaître la loi de Y et donner son espérance et sa variance.

(d) Les variables X et Y sont-elles indépendantes?

Partie III - Un calcul de covariance

(8) (a) On considère deux nombres entiers naturels k et n tels que $n \geqslant k$. Rappeler le lien entre les nombres

$$\binom{k}{j},\qquad \binom{k+1}{j+1}\quad \text{et}\quad \binom{k}{j+1}.$$

(b) Établir alors la formule suivante:

$$\sum_{k=j}^{n}\binom{k}{j}=\binom{n+1}{j+1}$$

(c) En faisant $j=2$, en déduire une expression factorisée de la somme suivante:

$$\sum_{k=2}^{n}k(k-1).$$

(9) Montrer que

$$E(XY)=\sum_{k=1}^{n}kP([X=k]\cap[Y=1])=\frac{n+1}{3}.$$

(10) En déduire la valeur de $\operatorname{cov}(X,Y)$.

8 Et une autre

J'ai posé ce problème dans le ***DS n°3****, à l'Automne 2022.*

Une urne contient des boules blanches (en proportion p), des boules noires (en proportion q) et des boules rouges (en proportion r). On a donc $p+q+r=1$ (on suppose que $p,q,r\in]0;1[$). On effectue des tirages successifs **avec remise** dans cette urne. Pour $i\in\mathbb{N}^*$, on note T_i la variable qui vaut 1 si la $i-$ème boule tirée est blanche, -1 si elle est noire et 0 si elle est rouge.
Les tirages étant effectués avec remise, les variables (T_i) sont donc mutuellement indépendantes.

Partie A

On note ensuite X_1 la variable aléatoire égale au numéro du tirage qui amène pour la première fois une boule blanche et X_2 celui correspondant au tirage où sort pour la première fois la **deuxième** boule blanche.

Par exemple, si les premiers tirages donnent *Noire, Rouge, Noire, Blanche, Rouge, Blanche, ...*, on a $X_1=4$ et $X_2=6$.

(1) Expliciter, pour tout $i\in\mathbb{N}^*$, la loi de T_i. Calculer son espérance et sa variance.

(2) Reconnaître la loi de X_1. Rappeler son espérance et sa variance.

(3) Compléter la fonction Python suivante, afin qu'elle simule les variables X_1 et X_2. Pourquoi la fonction ne prend-elle en argument que a proportion p de boules blanches?

```python
import numpy as np
import numpy.random as rd

def simul_X1_X2(p):
    x1=.......
    while ................ :
        x1=.............
    x2=..........
    while ................ :
        x2=................
    return [x1, x2]
```

(4) (a) Expliciter la loi conjointe du couple (X_1, X_2).
(b) En déduire la loi marginale de X_2.
(c) Montrer que X_2 admet une espérance et que $E(X_2) = \dfrac{2}{p}$.

(5) On note $U_2 = X_2 - X_1$.

(a) Pour $j \in \mathbb{N}^*$, déterminer $P(U_2 = j)$.
En déduire que X_1 et U_2 suivent la même loi, puis que U_2 admet une variance et préciser sa valeur.
(b) Montrer que U_2 est indépendante de X_1.
(c) Exprimer X_2 en fonction de U_2 et X_1 et en déduire que X_2 admet une variance qu'on explicitera.
(d) Que vaut $\text{cov}(X_1, X_2)$? Les variables X_1 et X_2 sont-elles indépendantes?
(e) On rajoute les instructions Python suivantes à la suite de la fonction précédente. Que peut-on prévoir quant à l'affichage après exécution de celles-ci ?

```python
p=1/4;
L=[ ]; M=[ ];
for k in range(10000):
    [X1,X2]=simul_X1_X2(p)
    L=L.append(X1)
    M=M.append(X2)
U=[M[k]-L[k] for k in range(10000)]
print(np.mean([L[k]*U[k] for k in range(10000)])-np.mean(L)*np.mean(U))
```

Partie B

On note W la variable correspondant au nombre de boules rouges obtenues avant l'obtention de la première boule blanche.

(6) Pour $i \in \mathbb{N}^*$, déterminer la loi conditionnelle de W sachant $[X_1 = i]$.
(7) En déduire que la loi de W est donnée par la somme, pour $k \in \mathbb{N}$,

$$P(W = k) = p\left(\frac{r}{q+r}\right)^k \sum_{i=k+1}^{+\infty} \binom{i-1}{k} \left(\frac{q}{q+r}\right)^{i-1-k} (1-p)^{i-1}$$

(8) Vérifier que, pour tout $i \in \mathbb{N}^*$ et pour tout $k \in [\![0, i]\!]$,

$$k\binom{i}{k} = i\binom{i-1}{k-1}$$

(9) En admettant qu'il est licite de permuter les sommes

$$\sum_{k=0}^{+\infty} \sum_{i=k+1}^{+\infty} \cdots = \sum_{i=1}^{+\infty} \sum_{k=0}^{i-1} \cdots ,$$

et que W admet une espérance, montrer que

$$E(W) = \frac{r}{p}$$

Partie C

On note, pour $n \in \mathbb{N}^*$, $S_n = T_1 + ... + T_n$.

(10) Quelle est la loi de S_1? Préciser son espérance et sa variance.
(11) Expliciter l'espérance et la variance de S_n.
(12) Soit $t > 0$. On pose $V_n = t^{S_n}$.

  (a) Déterminer la loi, l'espérance et la variance de V_1.
  (b) En déduire l'expression de $E(V_n)$.