Méthodes à retenir :

- Il est utile de savoir représenter les marginales dans un tableau pour un couple de variables aléatoires qui ne prennent qu'un nombre fini de valeurs.

I. Applications directes du cours

Exercice 1 ☆

On considère le couple (X, Y) de variables aléatoires à valeurs dans $\{0, 1\}$, dont la loi est donnée par :
$\mathbf{P}((0,0)) = 1/8$, $\mathbf{P}((0,1)) = 1/4$,
$\mathbf{P}((1,0)) = 3/8$, $\mathbf{P}((1,1)) = 1/4$.

1. Déterminer les lois marginales $\mathbf{P}_X$ et $\mathbf{P}_Y$, à l'aide d'un tableau.
2. X et Y sont-elles indépendantes ?

Exercice 2 ☆☆

Soient X et Y deux variables aléatoires indépendantes suivant des lois de Bernoulli $b(1/2)$. Les variables aléatoires $X+Y$ et $X-Y$ sont-elles indépendantes ?

Exercice 3 ☆☆ *Sondage*

Une population de personnes présente une propriété donnée avec une proportion inconnue $p \in]0,1[$.
On choisit un échantillon de n personnes et l'on pose $X_i = 1$ si le i-ème individu présente la propriété étudiée, 0 sinon. On considère que les variables aléatoires X_i ainsi définies sont indépendantes et suivent toute une loi de Bernoulli de paramètre p.
a) Quelle est la loi suivie par

$$S_n = X_1 + \cdots + X_n\,?$$

b) Déterminer espérance et variance de S_n/n.
c) Soit $\varepsilon > 0$. Etablir

$$\mathbf{P}\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) \leqslant \frac{1}{4n\varepsilon^2}$$

d) Pour $\varepsilon = 0,05$, quelle valeur de n choisir pour que S_n/n soit voisin de p à ε près avec une probabilité supérieure à 95 % ?

Exercice 4 *approximation binomiale-Poisson* ☆☆

Des machines fabriquent des plaques de tôle destinées à être empilées ; on estime à 0.1% la proportion de plaques inutilisables. L'utilisation de ces plaques consiste à en empiler n, numérotées de 1 à n en les prenant au hasard.

1. Pour $n = 2000$, quelle est la loi suivie par la variable aléatoire N "nombre de plaques inutilisables parmi les 2000" ? (on utilisera une loi de probabilité adaptée) ;
2. Quelle est la probabilité pour que N soit strictement inférieure à 3 ?
3. quelle est la probabilité pour que N soit inférieure ou égale à 3 ?

II. A savoir rédiger

Exercice 5 ☆☆☆ *Pièce*

On lance une pièce équilibrée plusieurs fois de suite. X est le rang pour lequel on obtient pour la 2ème fois « Pile ». On tire alors une boule dans une urne contenant $X-1$ boules numérotées de 1 à $X-1$ et on note Y le numéro de la boule tirée.

1. Montrer que X admet une variance et une espérance.
2. Déterminer la loi de Y à partir de la loi de Y sachant $\{X = k\}$.
3. Montrer que Y admet une variance et une espérance.
4. X et Y sont-elles indépendantes ?
5. Déterminer le coefficient de corrélation linéaire de X et Y.

Exercice 6 ☆☆

Soit X et Y deux variables aléatoires réelles sur l'espace probabilisé $(\Omega, \mathbf{P})$.
Montrer que :

$$|\mathrm{Cov}(X,Y)| \leqslant \sqrt{\mathbb{V}(X)\mathbb{V}(Y)}$$

Exercice 7 ✩
On note S la somme de deux dés équilibrés.

1. Déterminer les lois des variables aléatoires D_1 et D_2 représentant les résultats respectifs des deux dés.
2. Déterminer la loi de S.
3. Déterminer l'espérance de S.
4. Déterminer la variance de S.

Exercice 8 ✩✩✩ *Pièce*

On lance une pièce équilibrée plusieurs fois de suite. X est le rang pour lequel on obtient pour la 2ème fois « Pile ». On tire alors une boule dans une urne contenant $X-1$ boules numérotées de 1 à $X-1$ et on note Y le numéro de la boule tirée.

1. Montrer que X admet une variance et une espérance.
2. Déterminer la loi de Y à partir de la loi de Y sachant $\{X=k\}$.
3. Montrer que Y admet une variance et une espérance.
4. X et Y sont-elles indépendantes ?
5. Déterminer le coefficient de corrélation linéaire de X et Y.

Exercice 9 ✩✩ *CCP PSI 2015*
Soit $a>0$. On considère une variable aléatoire discrète X à valeurs dans $\mathbb{N}^*$ telle que :

$$\forall n \in \mathbb{N}^*,\ \mathbf{P}[X=n] = \frac{a}{n(n+1)}$$

.

1. Déterminer la valeur de a.
2. X admet-elle une espérance ? Une variance ?
3. Déterminer la série génératrice.

III. Exercices

Exercice 10 ✩✩
On dispose de n boîtes numérotées de 1 à n . La boîte k contient k boules numérotées de 1 à k.

On choisit au hasard une boîte puis une boule dans cette boîte.

Soit X le numéro de la boîte et Y le numéro de la boule.

1. Déterminer la loi du couple (X,Y).
2. Calculer $\mathbf{P}(X=Y)$.
3. X et Y sont-elles indépendantes ?

Exercice 11 ✩✩✩
Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$.
On suppose que la loi conjointe de X et Y vérifie

$$\mathbf{P}\left(X=j, Y=k\right) = \frac{a}{2^{j+1}\ k!} \text{ avec } a \in \mathbb{R}$$

a) Déterminer la valeur de a.
b) Reconnaître les lois marginales de X et Y.
c) Les variables X et Y sont elles indépendantes ?

Exercice 12 ✩✩
Soit X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que X suit une loi de Poisson de paramètre $\lambda>0$ et que la loi de Y sachant $X=n$ est binomiale de paramètres n et $p \in]0,1[$.
a) Déterminer la loi conjointe de (X,Y).
b) Reconnaître la loi de Y.

Exercice 13 ✩✩✩ *Pièce*
On lance une pièce équilibrée plusieurs fois de suite. X est le rang pour lequel on obtient pour la 2ème fois « Pile ». On tire alors une boule dans une urne contenant $X-1$ boules numérotées de 1 à $X-1$ et on note Y le numéro de la boule tirée.

1. Montrer que X admet une variance et une espérance.
2. Déterminer la loi de Y à partir de la loi de Y sachant $\{X=k\}$.
3. Montrer que Y admet une variance et une espérance.
4. X et Y sont-elles indépendantes ?
5. Déterminer le coefficient de corrélation linéaire de X et Y.

Exercice 14 ✩✩✩ *Modèle d'assurance*

Des sinistres surviennent sur une année avec une probabilité $p \in]0,1[$. Chaque sinistre à un coût fixe de C pour l'assureur. Chaque client paye une prime annuelle d'assurance π.

1. Pourquoi peut-on modéliser par une famille $(X_i)_{1\leq i\leq N}$ de N variables aléatoires de loi $b(p)$ les N clients d'un assureur ?
2. Quelle est l'espérance de S le solde annuel des sinistres pour l'assureur ?
3. Quelle est la variance de S ?
4. Donner une valeur de π la prime annuelle d'un assuré pour que l'assureur soit rentable une année avec une probabilité supérieure à 99%.

Exercice 15 ☆☆

Une urne contient n boules blanches et n boules rouges. On tire simultanément n boules dans celle-ci et on note X le nombre de boules rouges obtenues lors de ce tirage.
Quelle est la loi de X, son espérance, sa variance ?

Exercice 16 ☆☆ *Loi de Poisson*

Un insecte pond des oeufs suivant une loi de Poisson $P(\lambda)$. Chaque oeuf à une probabilité d'éclore avec une probabilité p , indépendante des autres oeufs. Soit Z le nombre d'oeufs qui ont éclos. Donner la loi de Z et en déduire son espérance

Exercice 17 ☆☆☆ *Approximation Binomiale Poisson*

Soit $p \in]0,1[$ fixé.
Pour tout $k \in \mathbb{N}$ fixé, vérifier que

$$\binom{n}{k} p^k (1-p)^{n-k} \underset{n\to+\infty}{\sim} \frac{e^{-np}(np)^k}{k!}$$

En déduire, pour $n \in \mathbb{N}$ et $\lambda = np$ que

$$\lim_{n\to+\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$$

Interprétation : La relation $\lambda_n = np$ montre que pour n grand, la probabilités $\binom{n}{k} p^k(1-p)^{n-k}$ qu'une variable de loi $\mathcal{B}(n,p)$ soit égale à k est très proche de la probabilités $e^{-\lambda}\frac{\lambda^k}{k!}$ qu'une variable de loi $\mathcal{P}(\lambda)$ soit égale à k : on dit que la loi de Poisson est la loi des évènements "rares".

IV. Pour aller plus loin

Exercice 18 ☆☆

On s'intéresse aux nombres de clients arrivant à l'un des deux guichets (A et B) d'une banque. Pour cela modélisons le problème de la manière suivante : Le nombre de client quittant la file d'attente, noté X , est modélisé par une loi de Poisson $P(\lambda)$. Le choix de guichet A par le i ème client est modélisé par une loi Bernoulli $b(p)$. Ainsi Y_i vaut 1 avec probabilité p si le $i$ème client choisit bien le guichet A et 0 avec probabilité $1-p$ sinon. On supposera de plus que toutes les variables (X_i, Y_i) sont mutullement indépendantes. Posons de plus $S = \sum_{k=1}^{X} Y_k$.

1. Que représente S ?
2. Que vaut $\mathbf{P}_{\{X=n\}}(\{S=k\})$?
3. Montrer que S suit une loi de Poisson.
4. En déduire $\mathbf{P}_{\{S=k\}}(\{X=n\})$, pour $k \in S(\Omega)$.

Exercice 19 ☆☆☆ *estimateurs statistiques*

Soient $X_1, \ldots, X_n$ des variables mutuellement indépendantes suivant une même loi d'espérance m et de variance σ^2.

a) On pose

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Calculer espérance et variance de $\bar{X}_n$.
Interpréter à l'aide de la loi des grands nombres.

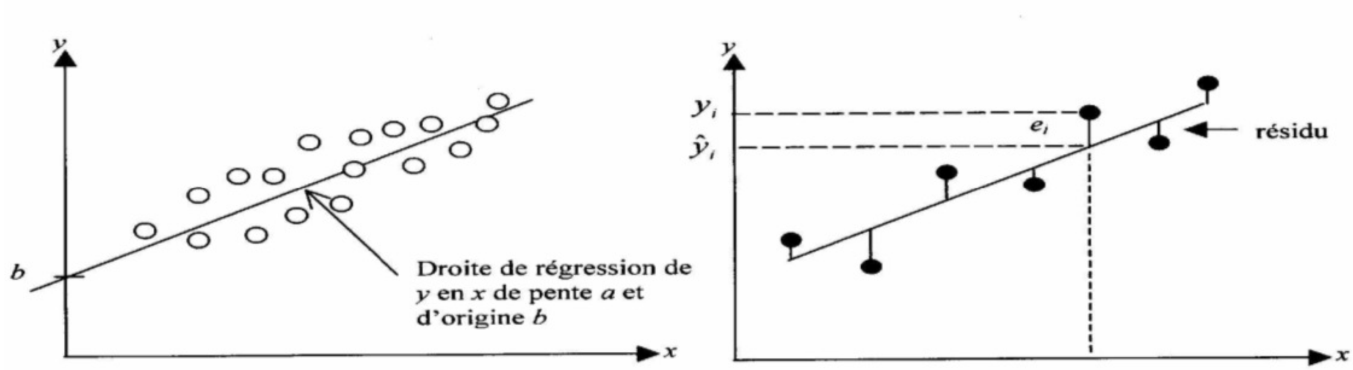
b) On pose

$$\bar{V}_n = \frac{1}{n-1}\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2$$

Calculer l'espérance de $\bar{V}_n$.

Exercice 20 ☆☆☆☆ *Modélisation par moindres carrés*

$$\sum_{i=0}^{n} \varepsilon_i^2 = \sum_{i=0}^{n} (y_i - (ax_i + b))^2.$$



Etant donné un n échantillon de mesures physiques $((x_i, y_i))_{1 \leq i \leq n}$, on cherche à savoir si cet échantillon peut-être assimilé à des réalisations aléatoires de couples $((x_i, ax_i + b))_{1 \leq i \leq n}$, pour des paramètres a et b de régression linéaire, selon une relation linéaire de la forme

$$y = a\ x + b$$

On note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, et $U = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathfrak{M}_{n,1}(\mathbb{R})$

1. Par définition, la droite de régression linéaire associée à ces mesures est la droite d'équation $y = \hat{a}x + \hat{b}$, où $(\hat{a}, \hat{b}) \in \mathbb{R}^2$ telle que $\delta_n(\hat{a}, \hat{b}) = \sum_{i=1}^{n} |y_i - \hat{a}\ x_i - \hat{b}|^2$ réalise le minimum de la fonction $(a, b) \mapsto \sum_{i=1}^{n} |y_i - a\ x_i - b|^2$.
2. Justifier que résoudre ce problème revient à minimiser une quantité de la forme

$$f(V) = \|MV - Y\|_2^2 \qquad (\star)$$

avec $M = [X, U] \in \mathfrak{M}_{n,2}(\mathbb{R})$ et $V = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathfrak{M}_{2,1}(\mathbb{R})$, et $Y \in \mathfrak{M}_{n,1}(\mathbb{R})$ à préciser.
3. On note $\hat{Y}$ le projeté orthogonal de Y sur le sous-espace vectoriel $\mathrm{Im}(M)$.
   (a) Pourquoi a-t-on $\mathrm{Im}(M) = \mathrm{Vect}(X, U)$?
   (b) Justifier que ce maximum est atteint et est unique.
   on utilisera la formule de projection sur un sous-espace vectoriel de dimension finie, muni d'une base orthonormée
   (c) Pourquoi a-t-on $\hat{Y} - Y \perp MV,\ \forall V \in \mathfrak{M}_{2,1}(\mathbb{R})$?
   (d) On pose $\hat{V}$ tel que $M\hat{V} = \hat{Y}$. En déduire que : $\forall\ V \in \mathfrak{M}_{2,1}(\mathbb{R}),\ V^T M^T (Y - M\hat{V}) = 0$
   (e) En déduire que : $M^T(Y - M\hat{V}) = 0$
   (f) Montrer que $\det(M^T M) = n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i^2)^2$
   (g) Etablir que $\hat{V} = (M^T M)^{-1}\ M^T\ Y$.
4. On rappelle que l'espérance de X (resp. Y) est $\mathbb{E}(X) = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ (resp. $\mathbb{E}(Y) = \dfrac{1}{n}\sum_{i=1}^{n} y_i$).
   On rappelle que la variance de X (resp. Y) est $\mathbb{V}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \mathbb{E}(X))^2$ (resp. $\mathbb{V}(Y) = \mathbb{E}(Y - \mathbb{E}(Y))^2 = \dfrac{1}{n}\sum_{i=1}^{n}(y_i - \mathbb{E}(Y))^2$).
   On appelle covariance de X et Y le nombre : $\mathrm{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \mathbb{E}(X))(y_i - \mathbb{E}(Y))$
   On vérifie que :

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left(\begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & \dots & 1 \\ x_1 & \dots x_n & \end{pmatrix} Y$$

$$= \frac{1}{n^2\mathbb{V}(X)} \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}.$$

En déduire que l'équation de la droite des moindres carrées est obtenue pour les paramètres : $\hat{b} = \mathbb{E}(Y) - \hat{a}\ \mathbb{E}(Y)$ et $\hat{a} = \dfrac{\mathrm{Cov}(X,Y)}{\mathbb{V}(X)}$.

5. La droite des moindres carrées et obtenue pour des paramètres $\hat{a}$ et $\hat{b}$ et a pour équation cartésienne

$$y = \hat{a}x + \hat{b}$$

Interpréter l'alignement des points sur cette droites dans le cas d'une covariance nulle.

Notes

[1] correction :

$\mathbf{P}_X(0) = 1/8 + 2/8 = 3/8$, $\mathbf{P}_X(1) = 3/8 + 2/8 = 5/8$, $\mathbf{P}_Y(0) = 1/8 + 3/8 = 1/2$, $\mathbf{P}_Y(1) = 2/8 + 2/8 = 1/2$

Non, X et Y ne sont pas indépendantes car $\mathbf{P}((0,0)) = 1/8 \neq 3/8 * 1/2 = 3/16$