Variables Aléatoires
sur un univers fini

MPSI Prytanée National Militaire

Pascal Delahaye - D'après le manuel "Mathématiques tout-en-un" MPSI

21 juin 2017

Dans ce chapitre comme dans le précédent, on ne considère que des univers Ω finis.

1 Variables aléatoires

Dans une expérience aléatoire, on s'intéresse rarement à l'espace probabilisé dans son ensemble, mais plus souvent à un aspect particulier du résultat. Par exemple :

1. le nombre de boules rouges obtenues lors d'un tirage de n boules.
2. la somme des faces abtenues lors d'un lancer de 2 dés.

Dans les deux cas précédents, ce qui nous intéresse est en fait une grandeur X qui peut prendre différentes valeurs entières selon les événements élémentaires qui se produisent. Il s'agit donc d'une application $X : \Omega \to \mathbb{N}$.

Définition 1 : **Variable aléatoire**
Soit Ω un univers fini et E un ensemble.
Toute application $X : \Omega \to E$ est appelée une *variable aléatoire* ⚠. On note : $\boxed{X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}}$
Lorsque $E = \mathbb{R}$ on dit que X est une *variable aléatoire réelle*.

Notations :

1. Pour tout $A \subset E$, $X^{-1}(A)$ est l'ensemble des événements élémentaires ω tels que $X(\omega) \in A$.
   En pratique, cet ensemble $X^{-1}(A)$ est l'événement noté :
   $$(X \in A) \quad \text{ou} \quad \{X \in A\}$$
2. Plus généralement, si P est une propriété pouvant être vérifiée par les éléments de E, alors l'ensemble des éléments $\omega \in \Omega$ tels que "$X(\omega)$ vérifie P" est noté :
   $$\{X \text{ vérifie } P\}$$
   Ainsi :
   - $\{\omega \in \Omega \mid X(\omega) = k\}$ est noté : $\{X = k\}$ ou $(X = k)$.
   - $\{\omega \in \Omega \mid X(\omega) \leq k\}$ est noté : $\{X \leq k\}$ ou $(X \leq k)$.

Remarque 1.

1. En pratique, nous verrons que dans la grande majorité des situations, X est à valeurs dans $\mathbb{N}$.
2. ⚠ Attention à la dénomination *variable* : X est en fait une application !

`Exemple 1.` $(*)$

1. On lance une pièce à trois reprises et on s'intéresse au nombre de fois où l'on a obtenu "Face".
2. On tire n boules au hasard dans une urne qui contient p boules rouges et on s'intéresse au nombre de boules rouge obtenues.
3. On distribue une main de 8 cartes et on s'intéresse au nombre de coeurs que contient cette main.

Remarque 2. On n'a pas besoin de la notion de probabilité pour définir une variable aléatoire.
En revanche, nous aurons besoin d'une probabilité dès lors que l'on veut calculer $P(\{X \in A\})$.

Exercice : 1

$(*)$ Une urne contient N boules numérotées de 1 à N.
L'expérience consiste à tirer au hasard et simultanément $p \leq N$ boules numérotées. On note :
- X la variable aléatoire qui a tout élément de Ω associe le plus petit numéro tiré
- Y la variable aléatoire qui a tout élément de Ω associe le plus grand numéro tiré

1. Que prendre comme univers Ω ?
2. Déterminer $X(\Omega)$ et $Y(\Omega)$.

Exercice : 2

$(*)$ Soit X une variable aléatoire à valeurs dans E ainsi que A et B deux parties de E.
Exprimer les événements $\{X \in A \cup B\}$, $\{X \in A \cap B\}$ et $\{X \in E \backslash A\}$ en fonction des événements $\left\{ \begin{array}{l} \{X \in A\} \\ \{X \in B\} \end{array} \right.$.

Proposition 1 : **Décomposition en partition de $\{X \in A\}$**
Soit $X : \Omega \to E$ une variable aléatoire et $A \subset E$. On a :

$$\{X \in A\} \quad = \quad \bigcup_{a \in A} \{X = a\} \quad = \quad \bigcup_{a \in A \cap X(\Omega)} \{X = a\}$$

Décomposition en partition de l'événement $\{X \in A\}$

`Exemple 2.` $(*)$ On lance deux dés et on s'intéresse à la variable aléatoire donnant la somme des chiffres obtenus.

1. Quel univers Ω peut-on choisir ?
2. Déterminer précisément les événements :

   (a) $A_1 = \{X = 7\}$ (b) $A_2 = \{X \leq 4\}$ (c) $A_3 = \{8,5 \leq X \leq 10,1\}$

Définition 2 : **Trois variables aléatoires particulières**
Soit Ω un univers fini, $A \subset \Omega$ et $a \in E$.

1. $X : \Omega \to E$ définie pour tout ω de Ω par $X(\omega) = a$ est appelée *variable aléatoire constante.*
2. $X : \Omega \to \{0,\ 1\}$ est appelée une *variable de Bernoulli.*
3. $X : \Omega \to \{0,\ 1\}$ définie par $\left\{ \begin{array}{l} X(\omega) = 1 \text{ si } \omega \in A \\ X(\omega) = 0 \text{ si } \omega \notin A \end{array} \right.$ est appelée *variable indicatrice de A.*
   Cette variable est notée 1_A et c'est bien entendu une variable de Bernoulli particulière.

Remarque 3. Avec les notations introduites précédemment, nous avons bien entendu : $\left\{ \begin{array}{l} \{1_A = 1\} = A \\ \{1_A = 0\} = \bar{A} \end{array} \right.$.

PROPOSITION 2 : Le produit de deux variables de Bernoulli est une variable de Bernoulli.

Preuve 2 : Immédiat puisqu'un tel produit ne peut également prendre que les valeurs 0 et 1.

2 Loi de probabilité d'une variable aléatoire

2.1 Loi d'une variable aléatoire

La variable aléatoire X permet de définir un nouvel univers $X(\Omega)$.
Si l'univers initial Ω est muni d'une probabilité P, la variable aléatoire X induit également une nouvelle probabilité P_X sur cet univers $X(\Omega)$.

DÉFINITION 3 : **Loi de probabilité P_X associée à une variable aléatoire X**
Soit P une probabilité sur un univers fini Ω et X une variable aléatoire définie sur Ω.
Ω étant de cardinal fini, $X(\Omega)$ est aussi de cardinal fini.
La fonction suivante est alors une probabilité sur $X(\Omega)$:

$$\begin{array}{rccc} P_X : & \mathcal{P}(X(\Omega)) & \longrightarrow & [0,\ 1] \\ & A & \mapsto & P(X \in A) \end{array}$$

Cette probabilité est appelée la *loi de X*.

Remarque 4.

1. Une variable aléatoire X se définit indépendamment de la notion de probabilité.
2. La loi de probabilité de X est quant à elle, déterminée à partir de la probabilité choisie sur Ω.
3. On allègera les notations en notant $P(X = x)$ la probablité $P(\{X = x\})$

$(\Omega,\ P)$ **induit un nouvel espace probabilisé** $(X(\Omega),\ P_X)$

DÉFINITION 4 : **Système complet d'événements associé à X**
Soit X une variable aléatoire sur un univers Ω.
La famille $(\{X = x\})_{x \in X(\Omega)}$ forme un système complet d'événements appelé le *système complet d'événements associé à X*. On a donc :

$$\sum_{x \in X(\Omega)} P(X = x) = 1$$

PROPOSITION 3 : **Determination de la loi P_X**
Soit X une variable aléatoire sur un univers probabilisé $(\Omega,\ P)$.
La loi P_X est déterminée de façon unique par les valeurs de $P(X = x)$ pour tout $x \in X(\Omega)$.

En effet, pour tout $A \subset X(\Omega)$ on a alors : $P_X(A) = \sum_{x \in A} P(X = x)$

`Exemple 3.` ($*$) Dans l'exemple précédent où X est la variable aléatoire donnant le nombre de "face" obtenu après avoir lancé 3 fois une pièce équilibrée, la probabilité P_X est définie par :

1. $P(X=0)=\frac{1}{8}$
2. $P(X=1)=\frac{3}{8}$
3. $P(X=2)=\frac{3}{8}$
4. $P(X=3)=\frac{1}{8}$

Loi d'une variable aléatoire X :

Déterminer la loi d'une variable aléatoire X consiste donc à :

1. Déterminer l'ensemble $X(\Omega)$
2. Déterminer les valeurs $P_X(x)=P(\{X=x\})$ pour tout $x \in X(\omega)$.

Lorsque $X(\Omega)$ est fini, la loi de probabilité P_X de X est souvent présentée sous forme de tableau.
Avec l'exemple de la pièce de monnaie lancée trois fois et X = nombre de face obtenus, cela donne :

$X(\Omega)$	$x_1=0$	$x_2=1$	$x_3=2$	$x_4=3$	Total
$P(X=x_k)$ ou $P_X(\{x_k\})$	$p_1=\dfrac{1}{8}$	$p_2=\dfrac{3}{8}$	$p_3=\dfrac{3}{8}$	$p_4=\dfrac{1}{8}$	1

`Exemple 4.` ($*$) Déterminer la loi de la variable aléatoire donnant la somme des deux chiffres obtenus par le lancer de deux dés équilibrés.

Exercice : 3

($*$) Une urne contient N boules numérotées de 1 à N.
L'expérience consiste à tirer au hasard et simultanément $p \leq N$ boules numérotées. On note :

— X la variable aléatoire qui a tout élément de Ω associe le plus petit numéro tiré
— Y la variable aléatoire qui a tout élément de Ω associe le plus grand numéro tiré

Déterminer les lois de probabilités de X et Y.

Remarque 5. Deux variables aléatoires peuvent être différentes et avoir la même loi de probabilité.
Considérer par exemple le lancer d'un dé équilibré, X le nombre de "pile" et Y le nombre de "face".

RUSE !

Soit X une variable aléatoire sur une espace probabilisé $(\Omega,\ P)$ telle que $X(\Omega) \subset \mathbb{Z}$.
Lorsque le calcul de $P(X \leq k)$ ou de $P(X \leq k)$ est plus simple que le calcul de $P(X=k)$, on peut envisager d'utiliser l'une des deux formules suivantes :

$$P(X=k)=\begin{cases} P(X \leq k)-P(X \leq k-1) \\ P(X \geq k)-P(X \geq k+1) \end{cases}$$

Exercice : 4

($**$) Un joueur prélève successivement n boules avec remise dans une urne contenant N boules numérotées de 1 à N.
On considère :

— X la variable aléatoire égale au plus grand numéro tiré
— Y la variable aléatoire égale au plus petit numéro tiré

1. Déterminer un espace probabilisé $(\Omega,\ P)$ rendant compte de l'expérience.
   Déterminer $X(\Omega)$ et $Y(\Omega)$.
2. Déterminer $P(X \leq k)$ pour tout $k \in [\![1, N]\!]$. En déduire la loi de X.
3. Déterminer $P(Y \geq k)$ pour tout $k \in [\![1, N]\!]$. En déduire la loi de Y.

Proposition 4 : Construction d'une variable aléatoire et de sa loi
Soit Ω un ensemble fini et E un ensemble non vide tel que $\text{Card}(\Omega) \geq \text{Card}(E) = n$.

Soient $x_1, \ldots, x_n \in E$ distincts deux à deux et $p_1, \ldots, p_n \in \mathbb{R}^+$ tels que $\sum_{k=1}^{n} p_k = 1$.

Il existe alors une variable aléatoire X (non unique) sur Ω et une probabilité P telles que :

$$X(\Omega) = \{x_1, \ldots, x_n\} \quad \text{et} \quad \forall k \in [\![1, n]\!],\ P(X = x_k) = p_k$$

Preuve 4 : Pas de difficulté.

`Exemple 5.` $(*)$ Construisez vous-même une variable aléatoire associée à une loi de probabilité.

Remarque 6.

Pour que la probabilité P précédente soit bien définie, il suffit de s'assurer que $p_1, \ldots, p_n \in \mathbb{R}^+$ et que $\sum_{k=1}^{n} p_k = 1$.

`Exemple 6.` $(*)$ Pour quelles valeurs de $p_2 \in \mathbb{R}$, existe-t-il une variable aléatoire X à valeurs dans $[\![0, 5]\!]$ telle que $P(X = 2) = p_2$ et que la suite $(P(X = k))_{0 \leq k \leq 5}$ soit arithmétique ?

2.2 Opérations sur les variables aléatoires

Définition 5 : La variable aléatoire $u(X)$
Soit X est une variable aléatoire sur Ω et $u \in \mathcal{F}(X(\Omega),\ E)$.
$u \circ X$ est une nouvelle variable aléatoire qui sera notée $u(X)$.
La loi de la variable $u(X)$ est donnée par :

$$\forall y \in u(X(\Omega)) : \qquad P(u(X) = y) \quad = \sum_{x \in u^{-1}(\{y\})} P(X = x)$$

Remarque 7. La notation $u(X)$ est cohérente avec l'appellation "variable" utilisée pour parler de X.

`Exemple 7.` $(*)$ On pose $Y = u(X)$.

1. Lorsque u est une application injective, si on note $\forall x \in X(\Omega),\ y = u(x)$, on a $P(Y = y) = P(X = x)$.
2. Lorsque u est définie par $u(x) = x^2$ et que $X(\Omega) = [\![-n, n]\!]$, alors $Y(\Omega) = \{k^2 \mid k \in [\![0, n]\!]\}$ et :

$$P(Y = 0) = P(X = 0) \quad \text{et} \quad \forall k \in [\![1, n]\!] : \quad P(Y = k^2) = P(X = k) + P(X = -k)$$

Exercice : 5

$(*)$ Un joueur lance successivement deux fois un dé équilibré.
Soit X la variable aléatoire égale à la différence entre le résultat du premier et du deuxième lancer.
Déterminer les lois de X, $|X|$ et X^2.

Remarque 8. L'ensemble des variables aléatoires réelles est $\mathcal{F}(\Omega,\ \mathbb{R})$.
Or, nous savons que cet ensemble a une structure de $\mathbb{R}$-ev et admet une multiplication interne.
Il sera donc possible de faire des *Combinaisons Linéaires* et des *Produits* de variables aléatoires.

Proposition 5 : Lois de $X + Y$ et XY
Soient X et Y deux variables aléatoires réelles sur un même espace probabilisé $(\Omega,\ P)$.

1. On a : $(X + Y)(\Omega) = \{(X + Y)(\omega) \mid \omega \in \Omega\}$ et en notant $I_z = \{(x,\ y) \in X(\Omega) \times Y(\Omega) \mid z = x + y\}$ on a :

$$\forall z \in (X + Y)(\Omega), \qquad \boxed{P(X + Y = z) = \sum_{(x,\ y) \in I_z} P(\{X = x\} \cap \{Y = y\})}$$

2. On a : $(XY)(\Omega) = \{X(\omega)Y(\omega) \mid \omega \in \Omega\}$ et en notant $J_z = \{(x,\ y) \in X(\Omega) \times Y(\Omega) \mid z = xy\}$ on a :

$$\forall z \in (XY)(\Omega), \qquad \boxed{P(XY = z) = \sum_{(x,\ y) \in J_z} P(\{X = x\} \cap \{Y = y\})}$$

Preuve 5 : Il suffit dans les deux cas, de partitionner correctement les événements $\{X+Y=z\}$ et $\{XY=z\}$.

Remarque 9. Les sommes précédentes contiennent des termes dont la probabilité est nulle.

`Exemple 8.` $(*)$ Si X est le nombre de fois où l'on a obtenu "Face" après avoir lancé n fois une pièce, alors on peut écrire $X = X_1 + X_2 + \cdots + X_n$ où X_k est la variable aléatoire qui vaut 1 si on a obtenu "pile" au $k$ème lancer et 0 sinon.

Exercice : 6

$(*)$ On lance à deux reprises un dé à 3 faces numérotées de 1 à 3.
Soit X et Y les variables aléatoires telles que :

$$\begin{cases} X = 0 \text{ si la face est paire au 1er lancé} \\ X = \text{le numéro de la face sinon} \end{cases} \quad \text{et} \quad \begin{cases} Y = 0 \text{ si la face est impaire au 2eme lancé} \\ Y = \text{le numéro de la face sinon} \end{cases}$$

Déterminer les lois des variables aléatoires $X+Y$ et XY.

3 Espérance, variance et écart type d'une variable aléatoire

Soit Ω un univers fini correspondant à une expérience aléatoire et P une probabilité sur Ω.
Soit X une variable aléatoire $\boxed{\text{réelle}}$ définie sur Ω. $X(\Omega)$ est alors de cardinal fini.

3.1 L'espérance

3.1.1 Définition et calcul de l'espérance

Définition 6 : **Espérance de X**
L'espérance de X est le réel défini par : $$E(X) = \sum_{x \in X(\Omega)} x.P(X=x)$$

Remarque 10.

1. Il s'agit de la moyenne des valeurs prises par X pondérées par les probabilités d'obtenir ces valeurs.
   L'espérance représente donc la valeur que l'on atteint en moyenne lorsqu'on réalise l'expérience aléatoire.
2. Lorsque X représente le gain du joueur à un jeu de hasard, $E(X)$ représente l'espoir de gain moyen par partie, lorsqu'il joue un grand nombre de fois. Si $E(X) > 0$ (resp. $E(X) < 0$) alors le jeu est avantageux (resp. désavantageux) pour le joueur. Si $E(X) = 0$ alors le jeu est dit équitable.
3. L'espérance ne dépendant que de la loi. Ainsi, deux variables aléatoires distinctes ayant la même loi auront la même espérance.

Remarque 11. Pour être utilisée, la formule de définition nécessite la connaissance de la loi de la VA X. Nous verrons par la suite deux situations permettant de calculer l'expérance d'une VA sans connaître sa loi :

1. la décomposition $X = X_1 + \cdots + X_n$
2. la "formule de transfert pour l'espérance de $u(X)$.

`Exemple 9.` $(*)$ On lance deux dés.
Calculer l'espérance de X, la variable aléatoire égale à la somme des chiffres apparus.

Exercice : 7

$(*)$ Un joueur tire simultanément n boules dans une urne contenant N boules numérotées de 1 à N.
Déterminer l'espérance de la variable aléatoire X donnant le plus petit numéro tiré.

Aide : on utilisera la formule de Pascal généralisée : $\displaystyle\sum_{k=n}^{N} \binom{k}{n} = \binom{N+1}{n+1}$

Proposition 6 : **Cas particuliers**
Soit (Ω, P) un espace probabilisé.

1. La variable aléatoire constante égale à a a pour espérance : $E(X) = a$.
2. L'espérance de la variable aléatoire indicatrique de A est : $E(1_A) = P(A)$.

Preuve 6 : Immédiat.

Lemme 7 : **Autre expression de l'espérance**
Soit X une variable aléatoire sur un espace probabilisé fini $(\Omega,\ P)$.
L'espérance est aussi donnée par la formule :

$$E(X) = \sum_{\omega\in\Omega} X(\omega).P(\omega)$$

Preuve 7 : Immédiat.

`Exemple 10.` $(*)$ Dans l'exemple du lancer d'une pièce à trois reprises, lorsque X donne le nombre de "face" obtenu, la formule précédente donne :

$$E(X) = \frac{1}{8}\sum_{\omega\in\Omega} X(\omega) = \frac{1}{8}(X(PPP)+X(PPF)+X(PFP)+X(PFF)+X(FPP)+X(FPF)+X(FFP)+X(FFF))$$

$$E(X) = \frac{1}{8}(0+1+1+2+1+2+2+3) = \frac{3}{2}$$

Théorème Fondamental 8 : **Théorème de Transfert**
Soit X une variable aléatoire réelle sur une espace probabilisé fini $(\Omega,\ P)$ et $u \in \mathcal{F}(X(\Omega),\ \mathbb{R})$.
L'espérance de $u(X)$ est donnée par la formule :

$$E(u(X)) = \sum_{x\in X(\Omega)} u(x)P(X=x)$$

Preuve 8 : On utilise la formule de définition de l'espérance et celle donnant la loi de probabilité de $u(X)$.
On conclut en remarquant que $\displaystyle\sum_{y\in u(X)(\Omega)} \Big(\sum_{x\in u^{-1}(\{y\})}\Big) = \sum_{x\in X(\Omega)}$.

Remarque 12. L'intérêt de cette formule est qu'elle permet de calculer l'espérance de $u(X)$ sans avoir besoin de connaître sa loi de probabilité. Seule la loi de X intervient dans la formule.

Exercice : 8

$(*)$ Soit X une variable aléatoire sur une espace probabilisé $(\Omega,\ P)$ telle que $\begin{cases} X(\Omega) = [\![1,n]\!] \\ \forall k\in[\![1,n]\!],\ P(X=k) = \frac{1}{n} \end{cases}$.
Donner les espérances de X^2 et de e^X.

3.1.2 Linéarité de l'espérance - Application

Proposition 9 : **Linéarité de l'espérance**
Soit X et Y deux variables aléatoires réelles sur un espace probabilisé fini $(\Omega,\ P)$. Soit $\lambda,\ \mu\in\mathbb{R}$.
On a alors :

$$E(\lambda X+\mu Y) = \lambda E(X) + \mu E(Y)$$

Ainsi, si $a\in\mathbb{R}$, on a : $E(aX+b) = aE(X)+b$.

Preuve 9 : Pas de difficulté avec la formule donnée dans le lemme 6.

Remarque 13. La fonction espérance est donc une forme linéaire non nulle de l'ev des variables aléatoires. Son noyau est donc un hyperplan de cet ev.

RUSE !

Lorsqu'on cherche à calculer l'espérance d'une variable aléatoire X "compliquée", il est parfois possible de décomposer cette variable en somme de variables "plus simples" :

$$X = X_1 + X_2 + \cdots + X_n$$

On peut alors calculer l'espérance de X en appliquant la formule : $E(X) = E(X_1) + E(X_2) + \cdots + E(X_n)$

`Exemple 11.` On lance 4 dés, et on note S la somme des résultats obtenus. Calculer $E(S)$.

`Exemple 12.` $(*)$ Une urne contient N boules dont b blanches.
On tire simultanément $n \in [\![1, N]\!]$ boules et l'on considère la variable aléatoire X égale au nombre de boules blanches obtenues. Calculer l'espérance de X.

Exercice : 9

$(*)$ Le problème des rencontres.
Une urne contient n boules numérotées de 1 à n.
On les extrait successivement et sans remise. On dit qu'il y a "rencontre au i-ème tirage" lorsque la boule tirée porte le numéro i. Déterminer le nombre moyen de rencontres.

Exercice : 10

$(*)$ Soit $n \in \mathbb{N}^*$. On range n objets dans n tiroirs (chaque tiroir pouvant contenir de 0 à n objets).
Calculer le nombre moyen de tiroirs vides et déterminer sa limite lorsque $n \to +\infty$.

Exercice : 11

$(*)$ On considère r boules numérotées de 1 à r et n tiroirs numérotés de 1 à n.
On lance au hasard chacune des r boules dans l'un des n tiroirs.
On note :

— V la variable aléatoire égale au nombre de tiroirs restés vides

— T la variable aléatoire égale au nombre de boules placées dans le tiroir 1.

Déterminer l'espérance des variables V et T.

3.1.3 Autres propriétés de l'espérance

Définition 7 : **Variable aléatoire centrée**
On dira que X est une variable aléatoire centrée lorsque : $E(X) = 0$.

Proposition 10 : **Centrage d'une variable aléatoire**
Pour tout variable aléatoire réelle X, la variable $Y = X - E(X)$ est centrée.

Preuve 10 : Pas de difficulté.

Proposition 11 : **Positivité de l'espérance**
Soit X une variable aléatoire réelle **positive ou nulle** sur une espace probabilisé fini $(\Omega,\ P)$.
On a alors :

1. $E(X) \geq 0$
2. $E(X) = 0$ si et seulement si $P(X = 0) = 1$.
   Dans ce cas, on dit que la variable X est presque sûrement nulle.

Preuve 11 : Pas de difficulté. On pourra prendre $X(\Omega) = \{x_1,\ \dots,\ x_n\}$ avec $x_1 < \dots < x_n$.

Remarque 14.

1. Cette propriété est analogue aux propriétés connues sur l'intégrale d'une fonction positive.
2. Dans un espace probabilisé fini, une variable aléatoire positive dont l'espérance est nulle sera nulle.

Proposition 12 : **Croissance de l'espérance**
Soit X et Y deux variables aléatoires réelles sur une espace probabilisé fini $(\Omega,\ P)$.
Alors :

$$X \leq Y \quad \Rightarrow \quad E(X) \leq E(Y)$$

On dit que $X \leq Y$ lorsque $\forall \omega \in \Omega,\ X(\omega) \leq Y(\omega)$.

Preuve 12 : Pas de difficulté en utilisant la linéarité et la positivité de l'espérance.

Théorème 13 : **Inégalité de Markov**
Soit X une variable aléatoire réelle positive sur une espace probabilisé fini $(\Omega,\ P)$ et $a > 0$.
On a alors :

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Preuve 13 : Il suffit de comparer les variables $1_{\{X\geq a\}}$ et $\dfrac{X}{a}$.

3.2 La variance

DÉFINITION 8 : **Moment d'ordre r**
Soit X une variable aléatoire réelle et r un entier naturel.
Le *moment d'ordre* r de la variable X est :

$$m_r(X) = E(X^r)$$

Remarque 15. Le moment d'ordre 0 est égal à 1 et le moment d'odre 1 est l'espérance de X.

PROPOSITION 14 : **Formule du moment**
Soit X une variable aléatoire réelle sur un espace probabilisé fini $(\Omega,\ P)$ et r un entier naturel.
On a :

$$m_r(X) = \sum_{x\in X(\Omega)} x^r P(X=x)$$

Preuve 14 : Immédiat à l'aide du théorème de transfert.

DÉFINITION 9 : **Variance**
Soit X une variable aléatoire réelle sur un espace probabilisé fini $(\Omega,\ P)$.
On appelle *variance* de X le réel :

$$V(X) = E((X-E(X))^2)$$

Remarque 16.

1. La formule montre que la variance est une mesure de la dispersion de X par rapport à $E(X)$.
2. Il aurait semblé plus naturel de prendre $E(|X-E(X)|)$ comme mesure de dispersion, mais cette expression ne possèdent pas les propriétés intéressantes (vues plus loin) de la variance.

PROPOSITION 15 : **Formule 1 de la variance**
Soit X une variable aléatoire réelle sur un espace probabilisé fini $(\Omega,\ P)$ et r un entier naturel.
La variance de X est donnée par la formule suivante :

$$V(X) = \sum_{x\in X(\Omega)} (x-E(X))^2.P(X=x)$$

Preuve 15 : Immédiat à l'aide du théorème de transfert.

`Exemple 13.` $(*)$ Calculer la variance de la variable aléatoire X donnant la somme des chiffres apparus lors du lancer de deux dés équilibrés.

THÉORÈME 16 : **Formule 2 de la variance : Koenig-Huyghens**
La variance d'une variable aléatoire X peut aussi se calculer avec la relation suivante :

$$V(X) = E(X^2) - [E(X)]^2$$

Preuve 16 : D'après la linéarité de l'espérance :
$V(X) = E((X-E(X))^2) = E(X^2 - 2XE(X) + E(X)^2) = E(X^2) - 2E(X)E(X) + E(X)^2E(1)$.
Ainsi : $V(X) = E(X^2) - [E(X)]^2$.

Remarque 17. La variance est donc aussi l'écart entre l'espérance des carrés et le carré de l'espérance.

Pour le calcul de la variance :
On préférera l'emploi de la formule de Koenig-Huyghens plutôt que celle donnée en définition.
En effet, outre le fait que les calculs sont beaucoup plus simples à mener, la formule de Koenig-Huyghens est surtout plus fiable lorsque l'espérance $E(X)$ ne tombe pas juste. En effet, dans la formule de la définition, l'erreur due à l'arrondi de $E(X)$ se propage tout au long du calcul alors qu'elle n'apparaît que dans le dernier terme dans la formule de Koenig-Huyghens.

`Exemple 14.` On lance 3 fois une pièce de monnaie et on note X le nombre de "face" obtenu. Calculer $V(X)$.

Exercice : 12

$(*)$ Soit $\lambda \in \mathbb{R}^+$, $n \in \mathbb{N}^*$ et X une variable aléatoire à valeurs dans $[\![1, n]\!]$ telle que $\forall k \in [\![1, n]\!],\ P(X = k) = \lambda k$.
Déterminer λ, $E(X)$ et $V(X)$.

Exercice : 13

$(*)$ Soit X une variable aléatoire sur un espace probabilisé fini.
Déterminer le minimum de la fonction φ définie par $\varphi(m) = E((X - m)^2)$.

RUSE Pour le calcul de $E(X^2)$

Parfois, la valeur de $E(X^2)$ est plus simple à obtenir en calculant $\left\{\begin{array}{l} E(X(X+1)) \\ \text{ou} \\ E(X(X-1)) \end{array}\right.$ puis en utilisant la linéarité.

Exercice : 14

$(**)$ On tire simultanément n boules dans une urne contenant N boules numérotés de 1 à N.
On considère X la variable aléatoire égale au plus grand des numéros tirés.
On admettra que la loi de X est donnée par : $\forall k \in [\![1, N]\!],\ P(X = k) = \dfrac{\binom{k-1}{n-1}}{\binom{N}{n}}$.

1. Montrer que $E(X) = \frac{n(N+1)}{n+1}$.
2. Calculer $E(X(X+1))$.

   Aide : on pourra utiliser la formule de Pascal généralisée : $\displaystyle\sum_{k=n}^{N} \binom{k}{n} = \binom{N+1}{n+1}$
3. En déduire que $V(X) = \frac{n(N+1)(N-n)}{(n+2)(n+1)^2}$.

Proposition 17 : **Propriétés de la variance**
Soit X une variable aléatoire réelle. On a :

1. $V(X) \geq 0$
2. $V(X) = 0 \iff \exists m \in \mathbb{R} \mid P(X = m) = 1$ (c'est à dire X est presque sûrement constante).
3. $V(aX + b) = a^2 V(X)$ lorsque $a,\ b \in \mathbb{R}$ (⚠ la variance n'est pas linéaire!)

Preuve 17 : Pas de difficulté avec la formule de Koenig-Huyghens compte-tenu des propriétés de l'espérance.

Remarque 18. En particulier : $V(aX) = a^2 V(X)$ et $V(X + b) = V(X)$.

Définition 10 : **Ecart-type**
Soit X une variable aléatoire réelle sur un espace probabilisé fini $(\Omega,\ P)$.
On appelle *écart-type* de X le réel :

$$\sigma(X) = \sqrt{V(X)}$$

Comme la variance, il s'agit d'une mesure de dispersion de la variable X.

Remarque 19. En particulier, si $a,\ b \in \mathbb{R}$, on a : $\sigma(aX) = |a|\sigma(X)$ et $\sigma(X + b) = \sigma(X)$.
Une translation n'a pas d'incidence sur la variance ou l'écart type d'une variable aléatoire (en effet, le degré de dispersion n'est pas modifié).

Définition 11 : **Variable centrée réduite**

Une variable aléatoire X telle que $E(X) = 0$ est dite *centrée*.
Une variable aléatoire X telle que $\sigma(X) = 1$ est dite *réduite*.
Une variable aléatoire X telle que $E(X) = 0$ et $\sigma(X) = 1$ est dite *centrée réduite*.

Proposition 18 : Soit X une variable aléatoire réelle telle que $V(X) \neq 0$.

La variable aléatoire $X^* = \dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Preuve 18 : Aucune difficulté.

Théorème Fondamental 19 : **Inégalité de Bienaymé-Tchebychev**
Toute variable aléatoire X vérifie la proposition :

$$\forall \varepsilon > 0, \quad P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

Preuve 19 : Pas de difficulté en utilisant l'inégalité de Markov.

Exercice : 15
($*$) **Inégalité de Bienaymé-Tchebytchev**
Soit X une variable aléatoire d'espérance μ et d'écart-type σ, et $\alpha > 0$.
Montrer que l'on a : $P(\mu - \alpha\sigma < X < \mu + \alpha\sigma) \geq 1 - \frac{1}{\alpha^2}$

Exercice : 16
($* * *$) **Inégalité de Markov**
Soit X une variable aléatoire d'espérance μ et d'écart-type σ, et $\alpha > 0$.
En utilisant la variable aléatoire $Y = (\alpha(X - \mu) + \sigma)^2$ que l'on a : $P(X \geq \mu + \alpha\sigma) \leq \frac{1}{1+\alpha^2}$.

Remarque 20. Les majoration et minoration obtenues par les inégalités de Markov et de Bienaymé-Tchebychev sont assez grossières.

4 Les lois usuelles

4.1 La loi uniforme

Modèle : Une variable aléatoire suit une loi uniforme lorsque toutes les valeurs de cette variable sont équiprobables.

Définition 12 : Une variable aléatoire X prenant $n \in \mathbb{N}^*$ valeurs toutes équiprobables suit *une loi uniforme.*
Lorsque $X(\Omega) = [\![1, n]\!]$, on a alors :

$$\forall k \in [\![1, n]\!], \quad P(X = k) = \frac{1}{n}$$

Plus généralement, lorsque $\begin{cases} a,\ b \in \mathbb{N} \\ X(\Omega) = [\![a, b]\!] \end{cases}$, on a : $\forall k \in [\![a, b]\!], \quad P(X = k) = \dfrac{1}{b - a + 1}$.

Remarque 21. Lorsque X suit une loi uniforme sur $[\![a, b]\!]$, on note : $X \hookrightarrow \mathcal{U}([\![a, b]\!])$.

`Exemple 15.` ($*$)

1. La valeur obtenue lors du lancer d'un dé.
2. La face obtenue lors du lancer d'une pièce.
3. Le numéro de la boule tirée parmi n toutes indiscernables au touché.

Proposition 20 : **Espérance et Variance (loi uniforme)**
Si $X \hookrightarrow \mathcal{U}([\![1, n]\!])$, alors :

$$E(X) = \frac{n+1}{2} \quad \text{et} \quad V(X) = \frac{n^2 - 1}{12}$$

Preuve 20 : Aucune difficulté.

Remarque 22. On constate que $E(X)$ est tout simplement la valeur médiane de $[\![1, n]\!]$.

Exercice : 17

($*$) Calculer l'espérance et la variance d'une variable X suivant une loi uniforme sur $[\![a, b]\!]$.

On pourra remarquer que dans ce cas : $Y = X - a + 1 \quad \hookrightarrow \quad \mathcal{U}([\![1, b - a + 1]\!])$.

4.2 La loi de Bernoulli

Modèle : On obtient une variable suivant une loi de Bernoulli lorsqu'on ne s'intéresse qu'au fait qu'un événement se produise ou non.

Définition 13 : Soit $p \in [0,\ 1]$.
On dit qu'une variable aléatoire X suit une *loi de Bernoulli* de paramètre p lorsque :

$$X(\Omega) = \{0,\ 1\} \quad \text{et} \quad \begin{cases} p(X = 1) = p \\ P(X = 0) = 1 - p \quad \text{souvent notée } q \end{cases}$$

Remarque 23. Lorsque X suit une loi de Bernoulli de paramètre p, on note : $X \hookrightarrow \mathcal{B}(p)$.
On parle alors d'une *variable de Bernoulli*.

`Exemple 16.` ($*$)

1. Toute expérience aléatoire ne comportant que deux issues peut être représentée par une variable de Bernoulli. Une telle expérience est appelée une *épreuve de Bernoulli*.
2. La variable indicatrice d'un événement est une variable de Bernoulli.
3. Le produit de deux variables de Bernoulli est une variable de Bernoulli.

Proposition 21 : **Espérance et Variance (loi de Bernoulli)**
Si $X \hookrightarrow \mathcal{B}(p)$, alors :

$$E(X) = p \quad \text{et} \quad V(X) = p(1 - p) = pq$$

Preuve 21 : Aucune difficulté.

Exercice : 18

($*$) Quelle est la valeur maximale de la variance d'une variable de Bernoulli.

Se ramener à une variable de Bernoulli

Lorsque qu'une variable aléatoire ne prend que 2 valeurs, on peut facilement obtenir son espérance et sa variance en se ramenant à une variable de Bernoulli.

Lorsque X est telle que $\begin{cases} X(\Omega) = \{a,\ b\} \\ P(X = b) = p \end{cases}$, on peut poser $Y = \dfrac{X - a}{b - a}$.

Ainsi on a $Y \hookrightarrow \mathcal{B}(p)$ et $X = (b - a)Y + a$ et donc : $\begin{cases} E(X) = (b - a)p + a \\ V(X) = (b - a)^2 p(1 - p) \end{cases}$

`Exemple 17.` ($*$) Déterminer l'espérance et la variance de la variable X telle que $\begin{cases} X(\Omega) = \{-1,\ 1\} \\ P(X = 1) = 1/3 \end{cases}$

4.3 La loi Binomiale

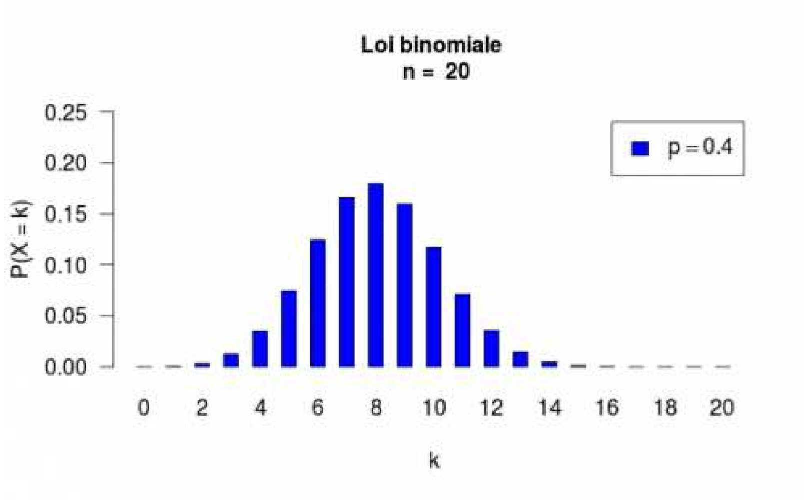
Modèle : On obtient une variable suivant une loi binomiale lorsque lorsqu'on s'intéresse au nombre de succès obtenu lors de la réalisation de n épreuves de Bernoulli indépendantes.
On obtient alors :

$$X(\Omega) = [\![0, n]\!] \quad \text{et} \quad \forall k \in [\![0, n]\!], \quad P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Une telle variable X s'écrit aussi $X = X_1 + \dots X_n$ où X_i prend la valeur 1 si le résultat de la ième expérience est un succès et 0 sinon.

DÉFINITION 14 : Soit $p \in [0,\ 1]$.
On dit qu'une variable aléatoire X suit une *loi binomiale* de paramètres $(n,\ p)$ lorsque :

$$X(\Omega) = [\![0, n]\!] \quad \text{et} \quad \forall k \in [\![0, n]\!], \quad P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$



Remarque 24.

1. Lorsque X suit une loi binomiale de paramètres $(n,\ p)$, on note : $X \hookrightarrow \mathcal{B}(n,\ p)$.
2. Les lois $\mathcal{B}(1,\ p)$ et $\mathcal{B}(p)$ ont la même signification.

Exemple 18. $(*)$

1. On lance indépendamment n pièces équilibrées et on s'intéresse à X, le nombre de "face" obtenu. On a alors $X \hookrightarrow \mathcal{B}(n,\ \frac{1}{2})$.
2. On effectue n tirages avec remise dans une urnes contenant une proportion p de boules blanches. Le nombre X de boules blanches tirées vérifie $X \hookrightarrow \mathcal{B}(n,\ p)$.

PROPOSITION 22 : **Espérance et Variance (loi binomiale)**
Si $X \hookrightarrow \mathcal{B}(n,\ p)$, alors :

$$E(X) = np \quad \text{et} \quad V(X) = np(1-p) = npq$$

Preuve 22 : On pourra décomposer la variable X en somme de variables de Bernoulli.
On admet que lorsque les variables $X_1, \dots, X_n$ sont indépendantes, on a $V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n)$

Exercice : 19

$(*)$ Montrer que si $X \hookrightarrow \mathcal{B}(n,\ p)$ alors $n - X \hookrightarrow \mathcal{B}(n,\ 1-p)$.

Exercice : 20

$(**)$ On lance n fois un dé équilibré à six faces.
Déterminer un entier naturel n_0 tel que pour tout $n \geq n_0$, on ait plus d'une chance sur 2 d'obtenir une fréquence d'apparition de 1 qui s'écarte de moins de 10^{-2} de la valeur $\frac{1}{6}$.

Remarque 25. **Loi hypergéométrique**
Parmi les lois usuelles qui ne sont pas au programme, on trouve aussi la *loi hypergéométrique*.
Lorsqu'une population E est constituée de N éléments répartis en deux groupes, les "bons" et les "méchants", les bons représentant un pourcentage p de la population. On sélectionne au hasard un échantillon de la population comprenant n éléments de E (tirage simultanée sans remise). La variable aléatoire X égale au nombre de bons éléments sélectionnés suit une loi hypergéométrique (voir exercice de TD).

5 Couples de variables aléatoires

5.1 Définition

Définition 15 : **Couple de variables aléatoires**
Soit Ω un univers fini et X et Y deux variables aléatoires à valeurs dans E et E' sur cet univers.

L'application : $\begin{array}{rrcl}(X,\ Y) : & \Omega & \longrightarrow & E \times E' \\ & \omega & \mapsto & (X(\omega),\ Y(\omega))\end{array}$ est appelée *couple de variables aléatoires sur* Ω

On note $(X,\ Y)$ ce couple de variables aléatoires.
Si $E = E' = \mathbb{R}$, on parle de couple de variables aléatoires réelles.

Remarque 26.
1. Il s'agit en fait d'une nouvelle variable aléatoire à valeurs dans $E \times E'$ telle que $(X,\ Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$.
2. L'ensemble des valeurs prises par $(X,\ Y)$ est seulement inclus dans $X(\Omega) \times Y(\Omega)$.

`Exemple 19.`
1. On lance deux dés parfaitement équilibrés et on considère X le plus petit des nombres obtenus et Y le plus grand des nombres obtenus.
2. On tire p boules dans une urnes contenant des boules de couleurs et on considère X le nombre de boules blanches et Y le nombre de boules noires obtenues.

Proposition 23 : Soit $(X,\ Y)$ un couple de variables aléatoires sur un univers fini Ω et $(x,\ y) \in X(\Omega) \times Y(\Omega)$.

- L'événement $(X,\ Y) = (x,\ y)$ est le même que $\{X = x\} \cap \{Y = y\}$ et est également noté $\{X = x,\ Y = y\}$.
- La famille $(\{X = x,\ Y = y\})_{(x,\ y) \in X(\Omega) \times Y(\Omega)}$ est un système complet d'événements de Ω.

5.2 Loi conjointe

Définition 16 : **Loi conjointe d'un couple de variables**
Soient X et Y deux variables aléatoires sur un espace probabilisé fini $(\Omega,\ P)$.
La loi conjointe de X et Y est la loi associée à la variable aléatoire $(X,\ Y)$ d'univers $(X,\ Y)(\Omega)$.
Cette loi est déterminée par les valeurs de :

$$(P(X = x,\ Y = y))_{(x,\ y) \in X(\Omega) \times Y(\Omega)}$$

- Lorsque $(x,\ y) \notin (X,\ Y)(\Omega)$, on obtient naturellement $P(X = x,\ Y = y) = 0$.
- On a bien entendu $\displaystyle\sum_{(x,\ y) \in X(\Omega) \times Y(\Omega)} P(X = x,\ Y = y) = 1$

Remarque 27. Cette loi est souvent (lorsque $X(\omega)$ et $Y(\omega)$ comportent peu de valeurs) représentée par un tableau à double entrée avec les valeurs de $X(\Omega)$ en ligne et les valeurs de $Y(\Omega)$ en colonne. Voir l'exemple suivant...

`Exemple 20.` $(*)$ On lance deux dés bien équilibrés. On note S la somme des résultats obtenus et P le produit.
Donner, sous forme de tableau la loi de probabilité du couple $(S,\ P)$.

Dressons tout d'abord deux petits tableaux donnant les différentes possibilités de sommes et de produits :

S	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

P	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

Loi du couple $(S,\ P)$:

S/P	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36	Loi de S
2	$\frac{1}{36}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{36}$
3	0	$\frac{2}{36}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{36}$
4	0	0	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{3}{36}$
5	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{4}{36}$
6	0	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	0	0	0	0	0	0	$\frac{5}{36}$
7	0	0	0	0	0	$\frac{2}{36}$	0	0	$\frac{2}{36}$	$\frac{2}{36}$	0	0	0	0	0	0	0	0	$\frac{6}{36}$
8	0	0	0	0	0	0	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	0	0	$\frac{5}{36}$
9	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	0	0	0	0	$\frac{4}{36}$
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{36}$	$\frac{1}{36}$	0	0	$\frac{3}{36}$
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{2}{36}$	0	$\frac{2}{36}$
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
Loi de P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

Exercice : 21

$(*)$ On lance deux dés à 6 faces parfaitement équilibrés et on considère X le plus petit des nombres obtenus et Y le plus grand des nombres obtenus. On munit $\Omega = [\![1, 6]\!]$ de la probabilité uniforme.
Donner la loi de probabilité de $(X,\ Y)$.

5.3 Lois marginales

Définition 17 : **Lois marginales**
Pour tout couple de variables aléatoires $(X,\ Y)$ sur un univers fini $(\Omega,\ P)$:
- la loi de X est appelée la première loi marginale du couple
- la loi de Y est appelée la deuxième loi marginale du couple

Comme l'indique le thèorème suivant, on peut déduire les lois marginales à partir de la loi du couple.

Proposition 24 : Soient X et Y deux variables aléatoires sur un espace probabilisé fini $(\Omega,\ P)$.
On a alors :

1. Pour tout $x \in X(\Omega)$, $\quad P(X = x) = \sum_{y \in Y(\Omega)} P(X = x,\ Y = y)$
2. Pour tout $y \in Y(\Omega)$, $\quad P(Y = y) = \sum_{x \in X(\Omega)} P(X = x,\ Y = y)$

Preuve 24 : Il suffit de remarquer que :
$P(X = x,\ Y = y) = P(\{X = x\} \cap \{Y = y\})$ et que $\{\{X = x \text{ et } Y = y\}\}_{y \in Y(\Omega)}$ est une partition de $\{X = x\}$.

`Exemple 21.` $(*)$ Retrouver les lois marginales de l'exemple précédent.

Détermination des lois marginales :
Lorsque la loi conjointe est représentée sous forme de tableau, on retrouve les lois marginales en sommant sur chaque ligne et sur chaque colonne.

`Exemple 22.` $(*)$ Une urne contient 3 boules blanches et 5 boules noires. On tire successivement deux boules dans l'urne sans remise. Soit X (resp. Y) la variable aléatoire qui vaut 1 si la première (resp. deuxième) boule tirée est blanches et 0 sinon.
Déterminer la loi conjointe du couple $(X,\ Y)$ et les deux lois marginales.

Remarque 28. En revanche, comme le montre l'exemple suivant, les lois marginales ne suffisent pas en général (sauf dans le cas de variables aléatoires indépendantes) à retrouver la loi conjointe.

Exercice : 22

$(*)$ Une urne contient 3 boules numérotées de 1 à 3. On en tire 2 et on note X le premier numéro tiré et Y le second.

Déterminer la loi conjointe et les lois marginales dans le cas d'un tirage avec remise et d'un tirage sans remise.
Que peut-on en conclure ?

5.4 Lois conditionnelles

DÉFINITION 18 : **Lois conditionnelles**
Pour tout couple de variables aléatoires $(X,\ Y)$ sur un univers fini $(\Omega,\ P)$:

— Pour tout $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$, la loi de X sachant $\{Y = y\}$ est la loi de X dans l'espace probabilisé $(\Omega,\ P_{\{Y=y\}})$. Elle est donnée pour tout $x \in X(\Omega)$ par :

$$P_{\{Y=y\}}(X = x) = \frac{P(\{X = x\} \cap \{Y = y\})}{P(\{Y = y\})}$$

— Pour tout $x \in X(\Omega)$ tel que $P(X = x) \neq 0$, la loi de Y sachant $\{X = x\}$ est la loi de Y dans l'espace probabilisé $(\Omega,\ P_{\{X=x\}})$. Elle est donnée pour tout $y \in Y(\Omega)$ par :

$$P_{\{X=x\}}(Y = y) = \frac{P(\{Y = y\} \cap \{X = x\})}{P(\{X = x\})}$$

Remarque 29. La loi de X sachant $\{Y = y\}$ est aussi appelée la *loi conditionnelle à* $\{Y = y\}$ *(sous entendu "de X")*

`Exemple 23.` $(*)$ On lance deux dés à 6 faces parfaitement équilibrés et on considère X le plus petit des nombres obtenus et Y le plus grand des nombres obtenus. On munit $\Omega = [\![1, 6]\!]$ de la probabilité uniforme.
Déterminer les différentes lois conditionnelles associées à X et Y.

Exercice : 23
$(**)$ Une urne contient 2 boules blanches, 3 boules rouges et 4 boules bleues.
On extrait 4 boules de l'urne, successivement avec remise.
On note X le nombre de boules blanches et Y le nombre de boules rouges obtenues.

1. Déterminer la loi conjointe de $(X,\ Y)$.
2. En déduire les deux lois marginales. Commenter le résultat.
3. Déterminer les lois conditionnelles. Commenter le résultat.

PROPOSITION 25 : **Lien avec les autres lois**
Soit $(X,\ Y)$ un couple de variables aléatoires sur un univers fini $(\Omega,\ P)$.
On suppose que pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$, on a $P(X = x) \neq 0$ et $P(Y = y) \neq 0$.
Pour tout $(x,\ y) \in X(\Omega) \times Y(\Omega)$, on a alors :

1. $P(X = x,\ Y = y) = P_{\{Y=y\}}(X = x)P(Y = y)$ et $P(X = x,\ Y = y) = P_{\{X=x\}}(Y = y)P(X = x)$
2. $P(X = x) = \displaystyle\sum_{y \in Y(\Omega)} P_{\{Y=y\}}(X = x)P(Y = y)$
3. $P(Y = y) = \displaystyle\sum_{x \in X(\Omega)} P_{\{X=x\}}(Y = y)P(X = x)$

Preuve 25 : Immédiat d'après les formules sur les probabilités conditionnelles.

Remarque 30. Ainsi, on retrouve la loi conjointe et les lois marginales à partir de la loi marginale d'une des variables et de la loi conditionnelle de l'autre variable par rapport à celle-ci.

Exercice : 24
$(*)$ Soit $n \in \mathbb{N}^*$ et $p,\ p' \in]0,\ 1[$.
On considère deux variables aléatoires X et Y à valeurs dans $[\![0, n]\!]$ telles que $X \hookrightarrow \mathcal{B}(n,\ p)$ et que pour tout $k \in [\![0, n]\!]$, la loi conditionnelle de Y par par rapport à $X = k$ est $\mathcal{B}(k,\ p')$.

1. Détermine la loi conjointe du couple $(X,\ Y)$.
2. Déterminer la loi de Y.
   Vous vérifierez que Y vérifie une loi binômiale.

5.5 Généralisation aux n-uplets de variables aléatoires

Définition 19 : **n-uplet de variables aléatoires**
Soit Ω un univers fini et $X_1, \ldots, X_n$, n variables aléatoires à valeurs dans $E_1, \ldots, E_n$ sur cet univers.

L'application : $\begin{array}{rccc} : & \Omega & \longrightarrow & E_1 \times \cdots \times E_n \\ & \omega & \mapsto & (X_1(\omega), \ldots, X_n(\omega)) \end{array}$ est appelée *n-uplet de variables aléatoires sur Ω*

On note $(X_1, \ldots, X_n)$ ce n-uplet de variables aléatoires.
Si $E_1 = \cdots = E_n = \mathbb{R}$, on parle de n-uplet de variables aléatoires réelles.

Remarque 31. $(X_1, \ldots, X_n)$ est une variable aléatoire sur Ω à valeurs dans $E_1 \times \cdots \times E_n$.

Définition 20 : **Loi conjointe et lois marginales d'un n-uplet de variables**
Soit $X_1, \ldots, X_n$, n variables aléatoires sur un espace probabilisé fini (Ω, P).

1. La *loi conjointe* de $X_1, \ldots, X_n$ est la loi associée à la variable aléatoire $(X_1, \ldots, X_n)$.
   Cette loi est déterminée par les valeurs de $(P(\{X_1 = x_1\} \cap \cdots \cap \{X_n = x_n\}))_{(x_1, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)}$.

2. Les *lois marginales* du n-uplet $(X_1, \ldots, X_n)$ sont les différentes lois associées aux variables $X_1, \ldots, X_n$.

Remarque 32. Les évenements $\{X_1 = x_1\} \cap \cdots \cap \{X_n = x_n\})$ et $(X_1, \ldots, X_n) = (x_1, \ldots, x_n)$ étant identiques, la probabilité $P(\{X_1 = x_1\} \cap \cdots \cap \{X_n = x_n\})$ est également notée $P(X_1 = x_1, \ldots, X_n = x_n)$.

Proposition 26 : **Formule d'obtention des lois marginale :**
Soit $X_1, \ldots, X_n$ des variables aléatoires sur un même espace probabilisé fini (Ω, P).
Pour tout $k \in [\![1, n]\!]$ et $x_k \in X_k(\Omega)$, on a :

$$P(X_k = x_k) = \sum_{(x_i \in X_i(\Omega))_{i \neq k}} P(X_1 = x_1, \ldots, X_k = x_k, \ldots, X_n = x_n)$$

Preuve 26 :
Il suffit d'appliquer la formule des probabilités totales en remarquant que :
$\{X_1 = x_1, \ldots, X_k = x_k, \ldots, X_n = x_n\}_{(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)}$ forme une partition de $(X_k = x_k)$.

`Exemple 24.` $(*)$ Soit n et N des entiers tels que $1 \leq n < N$.
Une urne contient N boules numérotées de 1 à N : on en tire n simultanément.
Les numéros obtenus sont notés $Y_1 < Y_2 < \cdots < Y_n$.
Déterminer la loi conjointe et les lois marginale du n-uplet $(Y_1, \ldots, Y_n)$ dans l'hypothèse où les $\binom{N}{n}$ tirages sont équiprobables.

6 Indépendance de variables aléatoires

6.1 Indépendance de 2 variables

Définition 21 : **Indépendance de variables aléatoires**
Soient X et Y deux variables aléatoires définies sur un univers Ω telles que $X(\Omega)$ et $Y(\Omega)$ soient finis.
On dira que les variables X et Y sont indépendantes lorsque pour tout événement $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants, c'est à dire lorsqu'on a :

$$P(\{X \in A\} \cap \{Y \in B\} = P(X \in A).P(Y \in B)$$

Remarque 33. Avec cette définition, il y a $2^{\mathrm{Card}(X(\Omega)) + \mathrm{Card}(Y(\Omega))}$ vérifications à effectuer. C'est beaucoup!!

`Exemple 25.` $(*)$ Reprenons l'exemple où l'on lance deux dés bien équilibrés et on s'intéresse à X la somme des résultats obtenus et Y le produit. Les variables aléatoires X et Y ne sont pas indépendantes.

En effet, en prenant $\left\{\begin{array}{l} A=\{2\} \\ B=\{2\}\end{array}\right.$, on constate que $\left\{\begin{array}{l} P(\{X=2\}\cap\{Y=2\})=0 \\ P(X=2).P(Y=2)=\dfrac{1}{36}\times\dfrac{2}{36}\neq 0\end{array}\right.$.

Théorème 27 : Caractérisation de l'indépendance
Soient X et Y deux variables aléatoires définies sur un univers Ω telles que $X(\Omega)$ et $Y(\Omega)$ soient finis.
On dira que les variables X et Y sont indépendantes lorsque pour tout $(x,\ y)\in X(\Omega)\times Y(\Omega)$ on a :

$$P(\{X=x\}\cap\{Y=y\})=P(X=x).P(Y=y)$$

Preuve 27 : L'une des implications est évidente.
Pour la deuxième, on écrit $\left\{\begin{array}{l} \{X\in A\} \\ \{Y\in B\}\end{array}\right.$ comme la réunion d'événements de la forme $\left\{\begin{array}{l} \{X=x\} \\ \{Y=y\}\end{array}\right.$.

Remarque 34. Dans ce cas, il y a $\text{Card}(X(\Omega))\times\text{Card}(Y(\Omega))$ vérifications à effectuer.

A retenir !!

Lorsque les variables sont indépendantes, la donnée des lois marginales permet de retrouver la loi conjointe.

Remarque 35. La notion d'indépendance n'est pas une notion ensembliste : elle dépend de la probabilité choisie.

`Exemple 26.` $(*)$ On considère $\Omega=[\![1,2]\!]^2$. On note X et Y la valeur du premier et du second élément du couple. Etudier l'indépendance des variables X et Y dans le cas où :

1. Ω est muni de la probabilité uniforme.
2. Ω est muni de la probabilité P définie par $P((i,\ j))=\left\{\begin{array}{l} \frac{1}{3} \text{ si } i=j \\ \frac{1}{6} \text{ si } i\neq j\end{array}\right.$.

`Exemple 27.` $(*)$ Soit X et Y deux variables aléatoires indépendantes sur $(\Omega,\ P)$ et $(x_0,\ y_0)\in X(\Omega)\times Y(\Omega)$. Prouver que les deux variables $1_{\{X=x_0\}}$ et $1_{\{Y=y_0\}}$ sont indépendantes.

Exercice : 25

$(*)$ On considère deux variables de Bernoulli $\left\{\begin{array}{l} X \\ Y\end{array}\right.$ de paramètres $\left\{\begin{array}{l} p \\ q\end{array}\right.$ définies sur le même espace $(\Omega,\ P)$.
Montrer que X et Y sont indépendantes si et seulement si les événements $\{X=1\}$ et $\{Y=1\}$ sont indépendants.

Proposition 28 : Autre caractérisation de l'indépendance
Soit $(X,\ Y)$ un couple de variables aléatoires sur un univers probabilisé fini $(\Omega,\ P)$.
Il y a équivalence entre :

1. Les variables X et Y sont indépendantes.
2. Pour tout $x\in X(\Omega)$ tel que $P(X=x)\neq 0$, on a $\forall y\in Y(\Omega)$: $P_{\{X=x\}}(Y=y)=P(Y=y)$.
3. Pour tout $y\in Y(\Omega)$ tel que $P(Y=y)\neq 0$, on a $\forall x\in X(\Omega)$: $P_{\{Y=y\}}(X=x)=P(X=x)$.

Preuve 28 : Immédiat d'après les formules sur les probabilités conditionnelles.

Proposition 29 : Fonctions de deux variables indépendantes
Soit $(X,\ Y)$ un couple de variables aléatoires indépendantes sur un univers probabilisé fini $(\Omega,\ P)$.
Soit $f,\ g$ deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$.
Les variables $f(X)$ et $g(Y)$ sont alors indépendantes.

Preuve 29 :
$P(\{f(X)\in A\}\cap\{g(Y)\in B\})=P(\{X\in f^{-1}(A)\}\cap\{Y\in g^{-1}(B)\})=P(X\in f^{-1}(A)).P(Y\in g^{-1}(B))$

Remarque 36. Ainsi, si X et Y sont indépendantes, alors X^n et Y^p sont aussi indépendantes pour tout $n,\ p\in\mathbb{N}^*$.

PROPOSITION 30 : **Somme de deux variables indépendantes**
Soit $(X,\ Y)$ un couple de variables aléatoires indépendantes sur un univers probabilisé fini $(\Omega,\ P)$.
La loi de $X+Y$ est alors donnée par :

$$\forall z \in (X+Y)(\Omega), \qquad P(X+Y=z) = \sum_{\substack{(x,\ y)\in X(\Omega)\times Y(\Omega)\\ x+y=z}} P(X=x)P(Y=y)$$

Lorsque $\left\{\begin{array}{l} X(\Omega)\subset\mathbb{N}\\ Y(\Omega)\subset\mathbb{N}\end{array}\right.$ la formule s'écrit : $\forall k\in\mathbb{N},\quad P(X+Y=k) = \displaystyle\sum_{i=0}^{k} P(X=i)P(Y=k-i).$

Preuve 30 : Immédiat !

Remarque 37. Plus généralement, si $Z = f(X,\ Y)$ avec f à valeurs dans E, alors on obtient la loi de Z en remarquant que :

$$\{f(X,\ Y)=z\} \quad = \bigcup_{\substack{(x,\ y)\in X(\Omega)\times Y(\Omega)\\ f(x,y)=z}} \{X=x\}\cap\{Y=y\}$$

Exercice : 26
$(*)$ Soit X et Y deux variables indépendantes sur le même espace probabilisé et suivant une loi uniforme sur $[\![1,n]\!]$.
Déterminer les lois de $X+Y$ et de $X-Y$.

6.2 Indépendance de n variables

DÉFINITION 22 : **Indépendance deux à deux de n variables**
On dira que n variables $X_1,\ \ldots,\ X_n$ sont *indépendantes deux à deux* lorsque pour tout entiers $i<j\in[\![1,n]\!]$, X_i et X_j sont indépendantes.

DÉFINITION 23 : **n variables mutuellement indépendantes**
On dira que n variables $X_1,\ \ldots,\ X_n$ à valeurs dans $E_1,\ \ldots,\ E_n$ sont *mutuellement indépendantes* (ou plus simplement *indépendantes*) lorsque pour toutes parties $A_1\subset E_1,\ \ldots,\ A_n\subset E_n$, les événements $\{X_1\in A_1\},\ \ldots,$ $\{X_n\in A_n\}$ sont (mutuellement) indépendants.
Cette notion d'indépendance est indépendante de l'ordre des variables.

Remarque 38. Des variables mutuellement indépendantes sont aussi indépendantes deux à deux.

PROPOSITION 31 : **Caractérisation de l'indépendance mutuelle**
Soit n variables $X_1,\ \ldots,\ X_n$ sur un univers probabilisé $(\Omega,\ P)$.

$X_1,\ \ldots,\ X_n$ sont indépendantes
si et seulement si
$\forall(x_1,\ \ldots,\ x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega)$ les événements $\{X_1=x_1\},\ \ldots,\ \{X_n=x_n\}$ sont indépendants.

Preuve 31 : Généralisation de la démonstration faite pour 2 variables aléatoires

Exercice : 27
$(*)$ Soit X et Y deux variables indépendantes suivant la loi de Bernoulli de paramètre $p=\frac{1}{2}$ et $Z=|X-Y|$.
Etudier l'indépendance deux à deux et l'indépendance mutuelle des variables X, Y et Z.
On pourra utiliser ce que l'on a vu sur l'indépendance de deux variables de Bernoulli.

Proposition 32 : **Autres variables indépendantes**
Soit n variables X_1, ..., X_n indépendantes sur un univers probabilisé $(\Omega,\ P)$.
Alors :

- toute sous-famille de $\{X_1, \dots,\ X_n\}$ est constituée de variables indépendantes.
- $f_1(X_1)$, ..., $f_n(X_n)$ sont indépendantes lorsque pour tout $i \in [\![1, n]\!]$, $f_i \in \mathcal{F}(X_i(\Omega),\ E_i)$.
- pour tout $p \in [\![1, n-1]\!]$, les "variables" $(X_1,\ \dots,\ X_p)$ et $(X_{p+1}, \dots,\ X_n)$ sont indépendantes.
- $f(X_1,\ \dots,\ X_p)$ et $g(X_{p+1}, \dots,\ X_n)$ sont indépendantes lorsque f, g sont des fonctions adaptées.
- Lemme des coalitions :
  Soit $k \in \mathbb{N}^*$ et I_1, ..., I_k des sous-ensembles non vides et disjoints de $[\![1, n]\!]$.
  Pour tout $i \in [\![1, k]\!]$, on considère une variable aléatoire Y_j fonction des variables X_i où $i \in I_j$.
  Les variables Y_1, ..., Y_k sont alors indépendantes.

Remarque 39. Ainsi :
— Si X_1, ..., X_n sont indépendantes alors $X_1 + \cdots + X_p$ et $X_{p+1} + \cdots + X_n$ sont indépendantes.
— Si X, Y, Z et T sont indépendantes, alors XY, ZT sont indépendantes, ainsi que X, $Y + Z$ et T^2.

Remarque 40. On peut utiliser une liste de variables indépendantes pour modéliser la succession de n expériences aléatoires dont les résultats sont indépendants.
Par exemple, la répétition de n lancers de "pile ou face" aux résultats indépendants sera décrite par une liste $(X_1,\ \dots,\ X_n)$ de n variables de Bernoulli (mutuellement) indépendantes où X_k représente le résultat de la $k$ème épreuve.

6.3 Somme de variables indépendantes suivant une loi binomiale

Théorème 33 : **Stabilité de la loi binomiale pour la somme**

1. Si X et Y sont des variables aléatoires **indépendantes** suivants des lois binomiales de paramètres respectifs $(m,\ p)$ et $(n,\ p)$ alors la somme $X + Y$ suit une loi binomiale de paramètres $(m + n,\ p)$.
2. Plus généralement, si X_1, ..., X_n sont des variables aléatoires **indépendantes** suivants des lois binomiales de paramètres $(n_i,\ p)$ alors la somme $X_1 + \cdots + X_n$ suit une loi binomiale de paramètres $(n_1 + \cdots + n_n,\ p)$.

Preuve 33 :

1. Simple calcul à l'aide de la formule de Vandermonde : $\displaystyle\sum_{i=0}^{k} \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}$
2. Par récurrence sur n.

Corollaire 34 : si X_1, ..., X_n sont des variables aléatoires **indépendantes** suivants des lois de Bernoulli de paramètre p alors la somme $X_1 + \cdots + X_n$ suit une loi binomiale de paramètres $(n,\ p)$.

Preuve 34 : Immédiat car $\mathcal{B}(p) = \mathcal{B}(1, p)$.

Exercice : 28
$(*)$ Soit $(X_1,\ \dots,\ X_n)$, un n-uplet de variables indépendantes suivants la même loi à valeurs dans E, et $x \in E$.
On note N_x la variable aléatoire égale à $\mathrm{Card}\{k \in [\![1, n]\!] \mid X_k = x\}$, c'est à dire définie par :

$$\forall \omega \in \Omega, \quad N_x(\omega) = \mathrm{Card}\{k \in [\![1, n]\!] \mid X_k(\omega) = x\}$$

Déterminer la loi de N_x, son espérance et sa variance.
Aide : utiliser le théorème précédent en introduisant les variables $1_{\{X_k = x\}}$.

7 Covariance

7.1 Définition et généralités

DÉFINITION 24 : **Covariance**
Soit X et Y deux variables aléatoires réelles sur le même espace probabilisé fini.
On appelle *covariance* de X et Y (ou du couple $(X,\ Y)$) le réel :

$$\mathrm{Cov}(X,\ Y) = E((X - E(X))(Y - E(Y)))$$

Il s'agit de l'espérance du produit des variables centrées.
On remarque que $\cos(X, X) = V(X)$.

`Exemple 28.` $(**)$ Montrer que $V(X) = 0 \Rightarrow \mathrm{Cov}(X,\ Y) = 0$.
On pourra utiliser le fait qu'une variable dont la variance est nulle est quasi-constante.

THÉORÈME 35 : **Formule de la covariance**
La covariance du couple $(X,\ Y)$ est aussi donnée par la formule :

$$\mathrm{Cov}(X,\ Y) = E(XY) - E(X)E(Y)$$

Preuve 35 : Simple calcul utilisant la linéarité de l'espérance...

PROPOSITION 36 : **Formule de calcul de $E(XY)$**
Lorsque X et Y sont deux variables aléatoires réelles sur un espace probabilisé fini $(\Omega,\ P)$, on obtient $E(XY)$ par la formule suivante :

$$E(XY) = \sum_{(x,\ y)\in X(\Omega)\times Y(\Omega)} xyP(X = x,\ Y = y)$$

Preuve 36 : On pourra introduire la variable $Z = XY$ et utiliser le théorème de transfert.

PROPOSITION 37 : **Propriétés**
Soient X, X', Y, Y' des variables aléatoires sur une même espace probabilisé fini et $\lambda,\ \mu \in \mathbb{R}$.
La covariance est bilinéaire, symétrique et positive.

1. $\mathrm{Cov}(X,\ X) = V(X) \geq 0$
2. $\mathrm{Cov}(X,\ Y) = \mathrm{Cov}(Y,\ X)$
3. $\mathrm{Cov}(\lambda X + \mu X',\ Y) = \lambda\mathrm{Cov}(X,\ Y) + \mu\mathrm{Cov}(X',\ Y)$
4. $\mathrm{Cov}(X,\ \lambda Y + \mu Y') = \lambda\mathrm{Cov}(X,\ Y) + \mu\mathrm{Cov}(X,\ Y')$

Preuve 37 : Pas de difficulté.

Remarque 41. La covariance n'est pas un produit scalaire car elle n'est pas *définie*.
En effet, une variable aléatoire constante vérifie $\mathrm{Cov}(X,\ X) = 0$ sans être nécessairement nulle.

Exercice : 29
$(**)$ On considère r boules numérotées de 1 à r et n tiroirs numérotés de 1 à n.
On lance au hasard chacune des r boules dans l'un des n tiroirs.
On note :
- V la variable aléatoire égale au nombre de tiroirs restés vides
- T la variable aléatoire égale au nombre de boules placées dans le tiroir 1.

Montrer en la calculant, que la covariance de V et T est nulle.
On pourra commencer par décomposer V et T en somme de variables de Bernoulli.

Théorème 38 : **Variance d'une somme**
Soit X et Y deux variables aléatoires réelles . On a :

$$V(X+Y) = V(X) + 2\text{Cov}(X,\ Y) + V(Y)$$

Plus généralement, lorsque $a,\ b \in \mathbb{R}$: $V(aX+bY) = a^2V(X) + 2ab\text{Cov}(X,\ Y) + b^2V(Y)$.
La généralisation à n variables aléatoires donne :

$$V(X_1 + \cdots + X_n) = \sum_{k=1}^{n} V(X_k) + 2 \sum_{1\leq i<j\leq n} \text{Cov}(X_i,\ X_j)$$

Preuve 38 : Simple calcul...

`Exemple 29.` (∗∗) Une urne contient des boules blanches, noires ou rouges dans les proportions p_1, p_2 et p_3.
On extrait n boules de l'urne avec remise. On note X, Y et Z le nombre de boules blanches, noires et rouges obtenues.
Déterminer la covariance du couple $(X,\ Y)$. *On pensera à utiliser la formule précédente...*

Exercice : 30
(∗∗) Le problème des rencontres.
Une urne contient n boules numérotées de 1 à n. On les extrait successivement et sans remise. On dit qu'il y a "rencontre au i-ème tirage" lorsque la boule tirée porte le numéro i.
On note X le nombre de rencontres. Montrer que $V(X) = 1$.
On pensera à décomposer X en somme de variables de Bernoulli X_i et on vérifiera que $E(X_i) = \frac{1}{n}$.

7.2 Le coefficient de corrélation (affine)

Théorème 39 : **Majoration de la covariance**
Si X et Y sont deux variables aléatoires réelles sur $(\Omega,\ P)$, alors :

$$|\text{Cov}(X,\ Y)| \leq \sigma(X)\sigma(Y)$$

On a $|\text{Cov}(X,\ Y)| = \sigma(X)\sigma(Y)$ ssi les deux variables sont presque sûrement liées par une relation affine.

Preuve 39 : On s'inspire de la démonstration de Cauchy-Schwarz en utilisant f définie par $f(t) = V(tX+Y)$.

Définition 25 : **Coefficient de corrélation**

- Si $V(Y) \neq 0$ et $V(X) \neq 0$, on peut définir : $\rho(X,\ Y) = \dfrac{\text{Cov}(X,\ Y)}{\sigma(X)\sigma(Y)}$

  $\rho(X,\ Y) \in [-1,\ 1]$ et est appelé le *coefficient de corrélation* entre X et Y.

- Deux variables réelles X et Y sont dites *non corrélées* lorsque : $\text{Cov}(X,\ Y) = 0$.

Remarque 42. **Mesure de dépendance affine**
$|\rho(X,\ Y)|$ est minimale ($= 0$) lorsque les variables X et Y sont indépendantes.
$|\rho(X,\ Y)|$ est maximale ($= 1$) lorsque les variables sont presque sûrement liées par une relation affine.
Le coefficient de corrélation mesure donc, en quelque sorte, le degré de dépendance affine des variables X et Y.

Corrélation (affine) entre 2 variables aléatoires

7.3 Cas de variables indépendantes

Théorème 40 : **Formules dans les cas de 2 variables indépendantes**
Soit X et Y deux variables aléatoires réelles indépendantes sur le même espace probabilisé $(\Omega,\ P)$.
On a alors :

1. $E(XY) = E(X)E(Y)$
2. $\text{Cov}(X,\ Y) = 0$
3. $V(X+Y) = V(X) + V(Y)$

Preuve 40 :
On calcule $E(XY)$ à partir de la définition en remarquant que $P(\{X = x\} \cap \{Y = y\}) = P(X = x).P(Y = y)$.
Les deux autres sont des conséquences immédiates de la première.

Exercice : 31

$(*)$ On considère 3 variables aléatoires mutuellement indépendantes X, Y et Z suivant la même loi de Bernoulli de paramètre p. Déterminer la covariance du couple $(X+Y,\ Y+Z)$ et celle du couple $(XY,\ YZ)$.

Lien entre la notion d'indépendance et celle de non corrélation affine

- Deux variables réelles indépendantes sur le même espace probabilisé fini sont non corrélées.
- ⚠ Comme le montre l'exemple suivant, la réciproque est fausse.

`Exemple 30.` $(*)$ On considère $\begin{cases} \text{la variable } X \text{ qui suit une loi uniforme sur } \{-1,\ 0,\ 1\} \\ \text{la variable } Y \text{ indicatrice de l'événement } \{X = 0\} \end{cases}$.

1. Montrer que ces variables se sont pas indépendantes.
2. Montrer que ces variables sont non corrélées.

Remarque 43. On a cependant l'équivalence pour des variables de Bernoulli.
On montre ce résultat en utilisant le fait que 2 variables de Bernoulli X et Y sont indépendantes si et seulement si les événements $\{X = 1\}$ et $\{Y = 1\}$ sont indépendants.

Exercice : 32

$(*)$ On considère une suite de N lancers de "pile ou face" indépendants, la probabilité d'obtenir "pile" étant $p \in]0,\ 1[$.
Pour $n \in \mathbb{N}^*$, on note S_n le nombre de succès obtenus lors des n premiers lancers.
Montrer que $\text{Cov}(S_m,\ S_n) = \mathfrak{M}(()n, m)p(1-p)$ pour $m,\ n \in [\![1, N]\!]$.

Proposition 41 : Lorsque $X_1, \dots,\ X_n$ sont n variables aléatoires réelles sur un même espace probabilisé, deux à deux indépendantes, alors :

$$V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n)$$

Preuve 41 : Pas de difficulté en utilisant la formule donnant la variance d'une somme en fonction des covariances.

Remarque 44.

1. Ce résultat est a fortiori vrai lorsque les n variables sont mutuellement indépendantes.
2. Cette formule doit nous inciter à décomposer une variable aléatoires en somme de variables aléatoires indépendantes plus simples.
   Retrouver ainsi la variance d'une variable suivant une loi binomiale.

Proposition 42 : Lorsque $X_1, \dots,\ X_n$ sont n variables aléatoires réelles indépendantes sur un même espace probabilisé, alors :

$$E(X_1 \times \cdots \times X_n) = E(X_1) \times \cdots \times E(X_n)$$

Preuve 42 : Par récurrence en utilisant le lemme des coalitions.

Exercice : 33

($*$) Soit $X_1, \dots,\ X_n$, des variables aléatoires réelles indépendantes sur $(\Omega,\ P)$, toutes d'espérance m et de variance 1. Déterminer l'espérance et la variance de $Y = \prod_{k=1}^{n} X_k$.

Feuille d'exercices MPSI : Variables Aléatoires

Codage :

1. Les exercices avec des coeurs $\heartsuit$ sont à traiter en priorité.
2. Le nombre d'étoiles $*$ ou de coeurs $\heartsuit$ correspond à la difficulté des exercices.

I] Espérance

1. L'espérance d'une VA X est donnée par la formule : $E(X) = \sum_{x\in X(\Omega)} x.P(X = x)$.
2. La formule de transfert permet de trouver l'espérance de $u(X)$: $E(u(X)) = \sum_{x\in X(\Omega)} u(x)P(X = x)$
3. C'est un opérateur linéaire et croissant.
4. Pour calculer l'espérance d'une VA X, il peut être utile de décomposer $X = X_1 + \cdots + X_n$.
5. La loi de Markov nous dit que : $P(X \geq a) \leq \frac{E(X)}{a}$

Exercice de TD : 1

($\heartsuit$) Soit $n \in \mathbb{N}^*$ et soit X une variable aléatoire finie dont la loi est donnée par :

$$X(\Omega) = [\![1, n]\!] \quad \text{et} \quad \forall k \in [\![1, n]\!], \quad P(X = k) = ak(n + 1 - k)$$

1. Trouver a.
2. Déterminer $E(X)$. Le résultat est-il surprenant ?
3. Déterminer $E(1/X)$.

Exercice de TD : 2

($\heartsuit\heartsuit$) Une urne est composée de n boules numérotées de 1 à n.
On effectue des tirages successifs et avec remise dans cette urne et on note $b_1, \ldots, b_k$ les numéros des boules prélevées. Les tirages s'arrêtent dès que $b_k \geq b_{k-1}$. Soit X_n la variable aléatoire égale au nombre de tirages effectués.

1. Déterminer la loi de probabilité de X_n.
   Aide : on pourra commencer par calculer $P(X_n > k)$ pour tout $k \in X_n(\Omega)$.
2. En déduire $E(X_n)$ (on pourra faire apparaître une somme télescopique) puis la limite de $E(X_n)$ lorsque $n \to +\infty$.

Exercice de TD : 3

($\heartsuit\heartsuit$) **Temps d'attente**.
Une urne contient $n \in \mathbb{N}^*$ boules numérotés de 1 à n. On retire, l'une après l'autre, toutes les boules de l'urne.

1. Quelle est la probabilité que les boules 1, 2 et 3 sortent successivement et dans cet ordre ?
2. Quelle est la probabilité que les boules 1, 2 et 3 sortent dans cet ordre (successivement ou pas) ?
3. On note X_n la variable aléatoire égale au nombre de tirages nécessaires pour obtenir les boules 1, 2 et 3. Déterminer la loi de X_n, son espérance et la méthode permettant de calculer sa vériance.

Exercice de TD : 4

(♡♡) Une urne contient $n-2$ boules blanches et 2 boules rouges.
On la vide en retirant les boules une à une et on note X_i le rang d'apparition de la $i$ème boule rouge.
Déterminer les espérances de X_1 et de X_2 en calculant successivement des lois de X_1, $X_2 - X_1$ et $n+1-X_2$.

Exercice de TD : 5

($*$) Soit X une variable aléatoire réelles et $g : \mathbb{R}^+ \to \mathbb{R}^+$ une fonction strictement croissante.
Montrer que $\forall a > 0$, on a : $P(|X| \geq a) \leq \dfrac{E(g(|X|))}{g(a)}$

II] Variance

1. La variance peut s'obtenir à l'aide de deux formules :
   - $V(X) = \displaystyle\sum_{x \in X(\Omega)} (x - E(X))^2.P(X = x)$
   - $V(X) = E(X^2) - [E(X)]^2$
2. C'est un opérateur positif mais non linéaire : $V(X+Y) \neq V(X) + V(Y)$ et $V(aX+b) = a^2V(X)$
3. Inégalité de Bienaymé-Tchebychev : $P(|X - E(X)| \geq \varepsilon) \leq \dfrac{V(X)}{\varepsilon^2}$

Exercice de TD : 6

($*$) Soit $n \in \mathbb{N}^*$ et soit X une variable aléatoire telle que : $X(\Omega) = [\![1, n]\!]$ et $\forall k \in [\![1, n]\!]$, $P(X = k) = \dfrac{1}{n}$.

1. Vérifier que l'on a bien défini une variable aléatoire. Déterminer $E(X)$ et $V(X)$.
2. Calculer le moment d'ordre 3 de X
3. Calculer $E(\frac{1}{X(X+1)})$.

Exercice de TD : 7

(♡♡) Soit $n \in \mathbb{N}^*$ et $a \in \mathbb{R}$.
Soit X une variable aléatoire à valeurs dans $[\![0, n]\!]$ telle que $\forall k \in [\![0, n]\!]$: $P(X = k) = \dfrac{a}{k+1}\dbinom{n}{k}$.
Déterminer la valeur de a puis calculer l'espérance et la variance de X.

Il pourra être intéressant de calculer $E(X+1)$ et $E(X(X+1))$.

Exercice de TD : 8

(♡) On lance simultanément 2 dés à 6 faces.
On appelle Z la variable aléatoire égale à la valeur absolue de la différence entre les 2 valeurs obtenues.

1. Déterminer la loi de probabilité de Z.
2. Calculer son espérance et sa variance.

III] Loi Binomiale

1. Une variable aléatoire suit une loi binomiale de paramètres n, p lorsque :
$$X(\Omega) = [\![0, n]\!] \quad \text{et} \quad P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$
2. C'est la probablilité d'obtenir k succès lorsqu'on répète n fois une épreuve de Bernoulli de probabilité p.
3. Si $X \hookrightarrow \mathcal{B}(n,\ p)$, alors :
$$E(X) = np \quad \text{et} \quad V(X) = np(1-p) = npq$$

Exercice de TD : 9

(♡♡) Soit $k \in [\![0, n]\!]$ et $X \hookrightarrow \mathcal{B}(n,\ p)$.

1. Déterminer ε de telle sorte que $(X \leq k) \subset (|X - E(X)| \geq \varepsilon)$.
2. En déduire que $P(X \leq k) \xrightarrow[n \to +\infty]{} 0$.

Exercice de TD : 10

(♡) Soit X une variable aléatoire suivant la loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in]0,\ 1[$.

1. Calculer l'espérance de la variable aléatoire $Y = \frac{1}{1+X}$.
2. On suppose que $p = \frac{1}{2}$ et que $a > 0$. Calculer l'espérance de $Z = \frac{e^X}{2n}$

Exercice de TD : 11

(♡♡) Un mobile évolue de façon aléatoire le long d'un axe gradué. A l'instant $t = 0$, il est en O (l'origine de l'axe). A chaque instant entier t appartenant $\mathbb{N}^*$, son abscisse varie de $+2$ avec la probabilité p et de 1 avec la probabilité $q = 1 - p$. Pour tout $n \in \mathbb{N}^*$, on note X_n son abscisse au temps $t = n$ et Y_n le nombre de déplacements de $+2$ effectués.

1. Soit n un entier naturel non nul fixé. Donner la loi de Y_n. Donner, sans calcul, $E(Y_n)$ et $V(Y_n)$.
2. Soit $n \in \mathbb{N}^*$.
   (a) Exprimer X_n en fonction de Y_n et en déduire la loi de X_n.
   (b) Calculer $E(X_n)$ puis $V(X_n)$.
3. Dans cette question, on suppose que $n = 3$ et que $p = 1/2$.
   Tracer le graphe de la fonction de répartition de X_3.

Exercice de TD : 12

(♡♡) Dans une ville, une proportion p de la population est atteinte par un virus contagieux. Si une personne saine est en contact avec une personne contaminée, il y a 2 chance sur 3 pour qu'elle soit elle-même contaminée. Un représentant de commerce, en parfaite santé, décide de rendre visite à n habitants de cette ville.

1. Soit N la variable aléatoire égale au nombre de malades rencontrés par le représentant.
   Quelle est la loi de N ?
2. Quelle est la probabilité pour que le représentant soit contaminé à l'issu de sa tournée ?

Exercice de TD : 13

(♡♡) On se propose d'analyser le sang d'une population de N individus pour déceler la présence éventuelle (résultat du test positif) d'un virus dont on sait qu'il affecte une personne donnée avec la probabilité p. On dispose pour cela de deux méthodes :

> **Méthode 1 :** On analyse le sang de chacune de N personnes.
>
> **Méthode 2 :** On regroupe la population en g groupes de n individus. On collecte le sang des n individus de chaque groupe dans une même éprouvette. Si le résultat d'un groupe est positif, on procède alors à une analyse individuelle de ses membres.

1. Quelle est la loi de la variable X égale au nombre de groupes positifs ?
2. Soit Y la variable égale au nombre d'analyses dans la deuxième méthode.
   Calculer $E(Y)$ en fonction de N, n et p.
3. Comparer les deux méthodes lorsque $N = 1000$, $n = 100$ et $p = 0,01$.

Exercice de TD : 14

(∗∗) Soit n un entier pair strictement positif et soit $X \hookrightarrow B(n,\ 1/2)$.
On note A l'événement "X est un nombre pair".

1. Pour tout $k \in [\![0, n/2]\!]$ donner $P(X = 2k)$.
2. Calculer $P(A)$.
3. Soit B l'événement "X est un nombre impair". Calculer $P(B)$

Exercice de TD : 15

(∗∗) Un autre exercice classique du même style. Soit X une variable aléatoire réelle suivant une loi binomiale de paramètres n et $1/2$. On note A l'événement "X est un multiple de 3". Calculer $P(A)$.

Exercice de TD : 16

(♡) Une population de personnes présente une propriété donnée avec une proportion inconnue $p \in]0,\ 1[$.
On choisi un échantillon de n personnes et l'on appelle S_n la variable aléatoire donnant le nombre de personnes vérifiant la propriété étudiée.

1. Quelle est la loi suivie par la variable S_n ?
2. Déterminer l'espérance et la variance de $\frac{S_n}{n}$
3. Soit $\varepsilon > 0$. Montrer que : $P(|\frac{S_n}{n} - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$
4. Pour $\varepsilon = 0,05$, quelle valeur de n choisir pour que $\frac{S_n}{n}$ soit voisin de p à ε près avec une probabilité supérieure à 95% ?

Exercice de TD : 17

(♡♡) **Inégalité de Hoeffding**

1. Soit X une variable aléatoire réelle positive sur un univers Ω fini et $a > 0$.
   (a) Prouver l'inégalité de Markov : $P(X \geq a) \leq \dfrac{E(X)}{a}$.
   (b) En déduire l'inégalité de Bienaymé Tchebytchev : $P(|X - E(X)| \geq a) \leq \dfrac{V(X)}{a^2}$.
2. On suppose que $X \hookrightarrow \mathcal{B}(n,\ p)$ sur $\Omega = [\![0, n]\!]$ avec $n > 0$ non multiple de 5 et $p = \frac{1}{10}$.
   (a) Prouver que $E(X) = np$ et que $V(X) = np(1-p)$.
   (b) Application de l'inégalité de Bienaymé Tchebytchev :
      i. Quelle valeur donner à n pour avoir $P(X \leq \frac{n}{5}) \geq 0,95$?
      ii. Quelle valeur donner à n pour avoir $P(X \leq \frac{n}{5}) \geq 0,99$?
   (c) L'inégalité de Hoeffding nous donne la relation : $P(|X - E(X)| \geq a) \leq 2\exp(-2\frac{a^2}{n})$.
      i. Quelle valeur donner à n pour avoir $P(X \leq \frac{n}{5}) \geq 0,95$?
      ii. Quelle valeur donner à n pour avoir $P(X \leq \frac{n}{5}) \geq 0,99$?
   (d) Constatez-vous une amélioration par rapport à l'inégalité de Bienaymé Tchebytchev ?

IV] Loi hypergéométrique

Même si cette loi n'est pas officiellement au programme de MPSI, il est important d'en traiter des exemples.

Exercice de TD : 18

(♡♡) **La loi hypergéométrique**

Une population E est constituée de N éléments répartis en deux groupes, les "bons" et les "méchants", les bons représentant un pourcentage p de la population. On sélectionne au hasard un échantillon de la population comprenant n éléments de E (tirage simultanée sans remise). On note X la variable aléatoire égale au nombre de bons éléments sélectionnés.

1. Déterminer la loi de X.
2. En déduire son espérance et sa variance.
   On reconnaitra dans l'expression de l'espérance la formule de Vandermonde.

On dit que la variable aléatoire X suit une loi hypergéométrique de paramètre N, n et p.

Exercice de TD : 19

(♡♡) **La loi hypergéométrique**
Dans un bocal de vinaigre blanc ont été placés C cornichons et G oignons blancs avec $(C,\ G) \in \mathbb{N}^{*2}$.
On pose $N = C + G$, $p = \frac{C}{N}$ et $q = \frac{G}{N}$. Ainsi, on a $p,\ q \in \mathbb{R}^+$ et $p + q = 1$.
On prend simultanément n condiments dans le bocal et on note X le nombre de cornichons recueillis.

1. Quelles sont les valeurs prises par X ?
2. Expliquer la formule : $P(X = k) = \dfrac{\binom{C}{k}\binom{N-C}{n-k}}{\binom{N}{n}}$.
3. Démontrer que $E(X) = np$.
4. Calculer $E(X(X+1))$ puis $V(X)$.

On dit que X suit une loi hypergéométrique de paramètres n, p, N notée $\mathcal{H}(n,\ p,\ N)$.

V] Couples de variables aléatoires

1) Lois conjointe, marginales et conditionnelles

1. Bien connaître les notions liées aux couples de VA :
   (a) loi conjointe
   (b) lois marginales
   (c) loi conditionnelle
   (d) Indépendance de deux VA
2. Lorsque deux variables aléatoires ne sont pas indépendantes, il est impossible de déterminer la loi conjointe à partir des lois marginales. On procédera alors toujours dans l'ordre suivant :

   Détermination de la loi conjointe → puis → Détermination des lois marginales

2) Covariance

1. La covariance s'obtient à l'aide de l'une des deux formules suivantes :
   $$\mathrm{Cov}(X,\ Y) = E((X - E(X))(Y - E(Y))) \quad \text{ou} \quad \mathrm{Cov}(X,\ Y) = E(XY) - E(X)E(Y)$$
2. La covariance est une "forme" bilinéaire symétrique positive (non définie).
3. On a la formule : $V(X+Y) = V(X) + 2\mathrm{Cov}(X,\ Y) + V(Y)$

Exercice de TD : 20

(♡) Soient X et Y deux variables de Bernoulli indépendantes et de même paramètre $p \in]0,\ 1[$ sur le même espace probabilisé.
Soit $U = X + Y$ et $V = X - Y$. Déterminer :

1. La loi du couple $(U,\ V)$.
2. la covariance de U et V.
3. Les variables U, V sont-elles indépendantes ? Conclusion ?

3) Variables Indépendantes

1. Deux VA X et Y sont indépendantes lorsque pour tout $(x,\ y) \in X(\Omega) \times Y(\Omega)$ on a :

$$P(\{X = x\} \cap \{Y = y\}) = P(X = x).P(Y = y)$$

2. Dans le cas où X et Y sont indépendantes, on a :

$$E(XY) = E(X)E(Y), \quad \text{Cov}(X,\ Y) = 0 \quad \text{et donc} \quad V(X+Y) = V(X) + V(Y)$$

Exercice de TD : 21

($\heartsuit$) Soient X et Y deux variables aléatoires suivant les lois binômiales $\mathcal{B}(n,\ \frac{1}{2})$ et $\mathcal{B}(m,\ \frac{1}{2})$.
Déterminer la probabilité de l'événement $(X = Y)$.

Exercice de TD : 22

($**$) Soit X_0, X_1 et X_2 trois variables aléatoires deux à deux indépendantes telles que $\left\{ \begin{array}{l} X_0 + X_1 \\ X_0 + X_2 \end{array} \right.$ soient indépendantes.
Montrer que la probabilité de ($X_0 =$ constant) vaut 1 (cad que $V(X_0) = 0$).
Vous penserez à utiliser la covariance.

Exercice de TD : 23

($\heartsuit\heartsuit$) On lance deux dés honnêtes. On note X_1 le numéro donné par le dé numéro 1 et X_2 le numéro donné par le dé numéro 2. Les variables aléatoires réelles X_2 et X_2 sont donc indépendantes. On note également X le plus grand des numéros obtenus et Y le plus petit.

1. Reconnaître les lois de X_1 et de X_2. Donner sans calcul $E(X_1)$ et $V(X_1)$.
2. Déterminer les lois de X et Y.
3. Calculer $E(X)$ et $E(Y)$. Comparer ces espérances et commenter.
4. Calculer $V(X)$ et $V(Y)$.
5. Représenter graphiquement la fonction de répartition de X.

Exercice de TD : 24

($\heartsuit\heartsuit - \heartsuit\heartsuit\heartsuit$) **Marche aléatoire**

1. Marche dans $\mathbb{Z}$.
   Un objet A se trouve en position 0 sur l'axe des réels.
   Chaque seconde, il a la probabilité $\left\{ \begin{array}{l} p \in]0,\ 1[\text{ de se déplacer de 1 vers la droite} \\ 1-p \text{ de se déplacer de 1 vers la gauche} \end{array} \right.$.
   On note X_n sa position après $n \in \mathbb{N}$ secondes et $Z_n = X_{n+1} - X_n$.
   (a) Déterminer la loi de Z_n, son espérance et sa variance.
   (b) En déduire la loi de X_n, son espérance et sa variance.

2. Marche dans $\mathbb{Z}^2$.
   $M_0 = (0,\ 0) = O$.
   Pour tout $n \in \mathbb{N}$, on pose $M_{n+1} = M_n + A_n$ où A_n suit une loi uniforme sur $\{(1,\ 0),\ (-1,\ 0),\ (0,\ 1),\ (0,\ -1)\}$.
   On pose également $M_n = (X_n,\ Y_n)$.

   (a) En posant $U_n = X_{n+1} - X_n$, déterminer $E(X_n)$ et $V(X_n)$.
   En déduire l'espérance du carré de la distance euclidienne de M_n à O.
   (b) Les variables X_n et Y_n sont-elles indépendantes ?
   (c) Prouver que $P(M_{2n} = O) = \dfrac{1}{16^n}\binom{2n}{n}^2$ puis en donner un équivalent lorsque $N \to +\infty$.