TD2. Probabilité sur un ensemble dénombrable.

1. a. Soit $(\Omega, \mathscr{F}, \mathbb{P})$ un espace de probabilités. Soit $n \geq 1$. Soient $A_1, \ldots, A_n$ des événements. Montrer que

$$\mathbb{P}(A_1 \cup \ldots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \ldots < i_k \leq n} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k}\right).$$

C'est la formule d'*inclusion-exclusion.*

b. En appliquant cette formule à un espace de probabilités et à des événements bien choisis, calculer le nombre de surjections de $\{1, \ldots, p\}$ dans $\{1, \ldots, n\}$ pour tous n et p entiers.

Solution de l'exercice 1. a. On démontre cette formule par récurrence sur n. Pour $n = 1$, la formule dit que $\mathbb{P}(A_1) = \mathbb{P}(A_1)$. Pour $n = 2$, la formule s'écrit $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$ et c'est un énoncé démontré dans le cours. Supposons la formule établie pour $n-1$ événements, avec $n \geq 3$, et considérons n événements $A_1, \ldots, A_n$. On a

$$\begin{aligned}
\mathbb{P}(A_1 \cup \ldots \cup A_n) &= \mathbb{P}((A_1 \cup \ldots \cup A_{n-1}) \cup A_n) \\
&= \mathbb{P}(A_1 \cup \ldots \cup A_{n-1}) + \mathbb{P}(A_n) - \mathbb{P}((A_1 \cup \ldots \cup A_{n-1}) \cap A_n) \\
&= \mathbb{P}(A_1 \cup \ldots \cup A_{n-1}) + \mathbb{P}(A_n) - \mathbb{P}((A_1 \cap A_n) \cup \ldots \cup (A_{n-1} \cap A_n))
\end{aligned}$$

L'hypothèse de récurrence permet de calculer le premier et le troisième terme du membre de droite et on obtient :

$$
\begin{aligned}
\mathbb{P}(A_1 \cup \ldots \cup A_n) &= \sum_{k=1}^{n-1}(-1)^{k-1} \sum_{1\leq i_1<\ldots<i_k\leq n-1} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k}\right) + \mathbb{P}(A_n) \\
&\quad - \sum_{k=1}^{n-1}(-1)^{k-1} \sum_{1\leq i_1<\ldots<i_k\leq n-1} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k} \cap A_n\right) \\
\mathbb{P}(A_1 \cup \ldots \cup A_n) &= \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots \mathbb{P}(A_n) \\
&\quad + \sum_{k=2}^{n-1}(-1)^{k-1} \sum_{1\leq i_1<\ldots<i_k\leq n-1} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k}\right) \\
&\quad - \sum_{k=1}^{n-2}(-1)^{k-1} \sum_{1\leq i_1<\ldots<i_k\leq n-1} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k} \cap A_n\right) \\
&\quad - (-1)^{n-2}\mathbb{P}(A_1 \cap A_2 \cap \ldots A_n) \\
\mathbb{P}(A_1 \cup \ldots \cup A_n) &= \mathbb{P}(A_1) + \mathbb{P}(A_2) + \ldots \mathbb{P}(A_n) \\
&\quad + \sum_{k=2}^{n-1}(-1)^{k} \sum_{1\leq i_1<\ldots<i_k\leq n} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k}\right) \\
&\quad + (-1)^{n-1}\mathbb{P}(A_1 \cap A_2 \cap \ldots A_n)
\end{aligned}
$$

où dans la dernière inégalité on utilise que

$$
-\sum_{k=1}^{n-2}(-1)^{k-1} \sum_{1\leq i_1<..<i_k\leq n-1} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k} \cap A_n\right) = \sum_{k=2}^{n-1}(-1)^{k-1} \sum_{1\leq i_1<..<i_k\leq n | i_k=n} \mathbb{P}\left(A_{i_1} \cap \ldots \cap A_{i_k}\right)
$$

On obtient bien la formule souhaitée pour $\mathbb{P}(A_1 \cap \ldots \cap A_n)$.

b. Considérons $\Omega = \{1, \ldots, n\}^{\{1,\ldots,p\}}$ l'ensemble des applications de $\{1, \ldots, p\}$ dans $\{1, \ldots, n\}$. Munissons Ω de la tribu $\mathscr{P}(\Omega)$ et de la mesure de probabilité uniforme :

$$
\forall A \in \mathscr{P}(\Omega), \mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{n^p}.
$$

Pour tout $m \in \{1, \ldots, n\}$, soit A_m l'ensemble des applications dont l'image ne contient pas l'entier m. La réunion $A_1 \cup \ldots \cup A_n$ est l'ensemble des applications qui ne sont pas surjectives. Par ailleurs, pour tout $k \in \{1, \ldots, n\}$ et tous $i_1, \ldots, i_k$ avec $1 \leq i_1 < \ldots < i_k \leq n$, l'ensemble $A_{i_1} \cap \ldots \cap A_{i_k}$ est l'ensemble des applications dont l'image ne contient aucun des entiers $i_1, i_2, \ldots, i_k$. Le nombre de ces applications est le nombre d'applications de $\{1, \ldots, p\}$ dans $\{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$, c'est-à-dire $(n-k)^p$. Ainsi,

$$
\mathbb{P}(A_1 \cup \ldots \cup A_n) = \sum_{k=1}^{n}(-1)^{k-1} \sum_{1\leq i_1<\ldots<i_k\leq n} \frac{(n-k)^p}{n^p} = \sum_{k=1}^{n-1}(-1)^{k-1}\binom{n}{k}\frac{(n-k)^p}{n^p}.
$$

Il s'ensuit que le nombre $S_{p,n}$ d'applications surjectives de $\{1,\dots,p\}$ dans $\{1,\dots,n\}$, qui est le cardinal du complémentaire de $A_1 \cup \ldots \cup A_n$, vaut

$$S_{p,n} = n^p - \sum_{k=1}^{n-1}(-1)^{k-1}\binom{n}{k}(n-k)^p = \sum_{k=0}^{n-1}(-1)^k\binom{n}{k}(n-k)^p.$$

Notons que, bien que ce ne soit pas évident sur cette formule, $S_{p,n} = 0$ si $p < n$.

2. Dans une grande assemblée, on demande à chaque personne d'écrire son nom sur un bout de papier et de le mettre dans un chapeau. On agite le chapeau puis chacun tire un bout de papier. Quelle est la probabilité que personne ne tire le bout de papier portant son propre nom ?

Solution de l'exercice 2. Soit n le nombre de personnes présentes. On attribue à chaque personne un numéro entre 1 et n. Chaque personne tire un nom et un seul du chapeau et l'application qui au numéro d'une personne associe le numéro de la personne dont elle a tiré le nom est une bijection de l'ensemble $\{1,\dots,n\}$ dans lui-même. On prend pour Ω l'ensemble des permutations de l'ensemble $\{1,\dots,n\}$. On munit Ω de la tribu $\mathscr{P}(\Omega)$ et de la probabilité uniforme. On chercher à compter le nombre de permutations $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ telles que pour tout $i \in \{1,\dots,n\}$ on ait $\sigma(i) \neq i$. On appelle de telles permutations des permutations sans point fixe ou des *dérangements*. On va utiliser la formule d'inclusion-exclusion.

Pour tout $m \in \{1,\dots,n\}$, on note A_m l'ensemble des permutations qui fixent m, c'est-à-dire l'ensemble des permutations $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ telles que $\sigma(m) = m$. Pour tout $k \in \{1,\dots,n\}$ et tous $i_1,\dots,i_k$ avec $1 \leq i_1 < \ldots < i_k \leq n$, l'ensemble $A_{i_1} \cap \ldots \cap A_{i_k}$ est l'ensemble des permutations telles que $\sigma(i_1) = i_1, \dots, \sigma(i_k) = i_k$. Le nombre de telles permutations est le nombre de permutations de l'ensemble $\{1,\dots,n\} \setminus \{i_1,\dots,i_k\}$, c'est-à-dire $(n-k)!$. Ainsi, la probabilité d_n qu'une permutation tirée uniformément soit un dérangement, qui est la probabilité du complémentaire de $A_1 \cup \ldots \cup A_n$, vaut

$$d_n = 1 - \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}\frac{(n-k)!}{n!} = \sum_{k=0}^{n}\frac{(-1)^k}{k!}.$$

Lorsque n tend vers l'infini, cette série converge très rapidement vers $\frac{1}{e}$. Puisque l'assemblée est grande, on peut dire avec une excellente approximation que la probabilité que personne ne tire son propre nom est $\frac{1}{e} \simeq 37\%$.

3. Paradoxe des anniversaires Quelle est la probabilité pour que parmi N personnes, au moins 2 aient la même date d'anniversaire ? Pour quelle valeur de N cette probabilité est-elle supérieure à $1/2$? (*On négligera l'existence du 29 février*)

Solution de l'exercice 3. Attribuons un numéro de 1 à 23 à chaque personne. Les dates d'anniversaire de ces vingt-trois personnes constituent une application de $\{1,\dots,23\}$ dans l'ensemble $\{1, \ldots, 365\}$ des jours de l'année. On prend donc $\Omega = \{1, \ldots, 365\}^{\{1,\ldots,23\}}$, muni de la tribu $\mathscr{P}(\Omega)$ et de la probabilité uniforme. L'événement A dont nous voulons calculer la probabilité est l'ensemble des applications qui ne sont pas injectives. Il est plus facile de compter les applications injectives que les applications qui ne le sont pas. On a en effet

$$\mathbb{P}(A^c) = \frac{|A^c|}{|\Omega|} = \frac{\frac{365!}{(365-N)!}}{365^N} = \frac{365}{365}\frac{364}{365} \cdots \frac{365-N}{365} = \prod_{k=0}^{N-1}\left(1 - \frac{k}{365}\right).$$

On peut montrer que la $u_N = \prod_{k=0}^{N-1}\left(1 - \frac{k}{365}\right)$ est décroissante, et bien sûr on a $u_{365} = 0$, on peut donc déterminer numériquement la première valeur de N pour laquelle $u_N < \frac{1}{2}$. On trouve $N = 23$.

4. On tire deux cartes d'un jeu de 32. Quelle est la probabilité d'obtenir une paire ? Si l'on n'a pas obtenu une paire, on a le choix entre jeter l'une des deux cartes tirées et en retirer une parmi les 30 restantes, ou jeter les deux cartes tirées et en retirer deux parmi les 30 restantes. Quelle stratégie donne la plus grande probabilité d'avoir une paire à la fin ?

Solution de l'exercice 4. Il y a $C_4^2 = 6$ paires de chaque hauteur (sans l'ordre), donc $2 \times 6 \times 8 = 96$ en tout (avec l'ordre). Donc une probabilité de $\frac{96}{32*31} = \frac{3}{31} \simeq 0,097$.

Si on change une seule carte, on obtient une paire lorsqu'on tire une des 3 cartes de la même hauteur que celle qu'on a gardée en main. On a donc une probabilité $\frac{3}{30} = \frac{1}{10}$ d'obtenir une paire.

Si on change les deux cartes, il y a 30×29 tirages ordonnés possibles de 2 nouvelles cartes. Parmi eux, on a deux manières d'obtenir une paire : Considérons d'abord le cas où la première carte est l'une des 6 cartes de la même hauteur que l'une des deux cartes que l'on a jetées, alors il reste seulement 2 cartes parmi les 29 restantes permettant de compléter la paire. Dans l'autre cas, où la première carte est l'une des 24 autres cartes, il y a 3 secondes cartes possibles donnant une paire. Au final, on obtient $6 \times 2 + 24 \times 3 = 84$ tirages ordonnés donnant une paire. Soit une probabilité de $\frac{84}{30\times 29} = \frac{14}{145} < \frac{1}{10}$.

C'est donc la première stratégie qui donne la plus grandes possibilité d'obtenir une paire.

5. On considère un jeu de pile ou face infini. Soit $n \geq 0$ un entier. Calculer la probabilité que le premier temps auquel on obtient pile soit le temps n.

Soit $k \geq 1$ un entier. Calculer la probabilité que le k-ième temps auquel on obtient pile soit le temps n.

Solution de l'exercice 5. On note p la probabilité d'obtenir un pile. Celle d'obtenir un face est donc $1-p$. L'énoncé laisse entendre que le jeu est non biaisé et que $p = 1/2$ mais les calculs sont plus clairs en gardant la notation p.

La probabilité que le premier pile soit obtenu au temps n (et donc qu'on a donc obtenu face lors des $n-1$ premiers tirages) est $(1-p)^{n-1}p$.

Soit $k \geq 1$. Dire que le k-ième temps auquel on obtient pile est le temps n, revient a dire qu'on a obtenu exactement $k-1$ pile sur les $n-1$ premiers tirages, puis un pile encore au n-ième. Il y a C_{n-1}^{k-1} manières d'obtenir cela (chacune revient à choisir les positions des $k-1$ premiers pile parmi les $n-1$ premiers lancers). Chacun de ces tirages a la même probabilité $(1-p)^{n-k}p^k$. De plus ces événements sont disjoints, donc la probabilité totale (que le k-ième temps auquel on obtient pile soit le temps n) est $C_{n-1}^{k-1}(1-p)^{n-k}p^k$.

6. On lance un dé tétraédral dont les faces sont numérotées de 1 à 4 et un dé octaédral dont les faces sont numérotées de 1 à 8. Calculer la loi de la somme S, du produit P et du plus grand M des deux nombres obtenus.

Solution de l'exercice 6. Il y a 32 lancers équiprobables. Il faut compter le nombre de manière d'obtenir chacun des résultats. On obtient :

$\mathbb{P}(S=2) = \mathbb{P}(S=12) = 1/32$, $\mathbb{P}(S=3) = \mathbb{P}(S=11) = 1/16$,
$\mathbb{P}(S=4) = \mathbb{P}(S=10) = 3/32$,
$\mathbb{P}(S=5) = \mathbb{P}(S=6) = \mathbb{P}(S=7) = \mathbb{P}(S=8) = \mathbb{P}(S=9) = 1/8$.

$\mathbb{P}(P=1) = \mathbb{P}(P=5) = \mathbb{P}(P=7) = \mathbb{P}(P=9) = \mathbb{P}(P=10) = \mathbb{P}(P=14) = \mathbb{P}(P=15) = \mathbb{P}(P=18) = \mathbb{P}(P=21) = \mathbb{P}(P=28) = \mathbb{P}(P=32) = 1/32$,
$\mathbb{P}(P=2) = \mathbb{P}(P=3) = \mathbb{P}(P=16) = \mathbb{P}(P=24) = 1/16$,
$\mathbb{P}(P=4) = \mathbb{P}(P=6) = \mathbb{P}(P=8) = 3/32$, $\mathbb{P}(P=12) = 1/8$.

$\mathbb{P}(M=1) = 1/32$, $\mathbb{P}(M=2) = 3/32$, $\mathbb{P}(M=3) = 5/32$,
$\mathbb{P}(M=4) = 7/32$, $\mathbb{P}(M=5) = \mathbb{P}(M=6) = \mathbb{P}(M=7) = \mathbb{P}(M=8) = 1/8$.

7. Soient X, Y, Z trois variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que X et Y ont même loi. Soit $f : \mathbb{N} \to \mathbb{N}$ une fonction.

Est-il vrai que $f(X)$ et $f(Y)$ ont même loi ? Est-il vrai que $X+Z$ et $Y+Z$ ont même loi ?

Solution de l'exercice 7. $f(X)$ et $f(Y)$ ont bien la même loi, car

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(f(X)=k) = \mathbb{P}(X \in f^{-1}(\{k\})) = \mathbb{P}(Y \in f^{-1}(\{k\})) = \mathbb{P}(f(Y)=k).$$

En revanche, $X+Z$ et $Y+Z$ peuvent avoir des lois différentes. Par exemple, on peut déterminer les valeurs des 3 variables aléatoires en fonction du résultat d'un même lancer d'une pièce de monnaie équilibrée :

$$\mathbb{P}(X=0, Y=1, Z=2) = 1/2, \quad \mathbb{P}(X=1, Y=0, Z=5) = 1/2.$$

X et Y suivent alors la même loi, mais $\mathbb{P}(X+Z=2) = 1/2$ tandis que $\mathbb{P}(Y+Z=2) = 0$.

8. Un chimpanzé tape à la machine à écrire en appuyant chaque seconde sur une touche choisie au hasard. Quelle est la probabilité qu'il parvienne à écrire *Hamlet*, c'est-à-dire qu'à un certain moment il écrive d'une traite le texte de cette pièce ?

Solution de l'exercice 8. Soit n la longueur, en caractères, de la pièce *Hamlet*. La probabilité p qu'il tape la pièce du premier coup est faible, mais strictement positive. Pour tout entier naturel k, on définit une variable aléatoire X_k qui vaut 1 si les caractères $nk+1$ à $n(k+1)$ correspondent au texte de la pièce. Les X_k sont les mêmes que dans un jeu de pile ou face biaisé. Or on sait dans ce cas qu'il finira par sortir un pile, ce qui correspond ici à écrire le texte de la pièce d'une traite.

La probabilité que le chimpanzé écrive *Hamlet* au bout d'un certain nombre (aléatoire, mais fini) de tentatives vaut donc 1. Néanmoins, en pratique, il y a fort à parier que le chimpanzé (ou la machine à écrire) arrivera à épuisement bien avant.