TD 01 – Évènements (corrigé)

Exercice 1. *Compléments des indépendants*

1. Montrer qu'une famille d'événements $(A_i)_{1\leqslant i\leqslant n}$ est indépendante si et seulement si la famille des compléments $(\overline{A_i})_{1\leqslant i\leqslant n}$ est indépendante. *Indice : On pourra commencer par montrer que la famille $\{\overline{A_1}, A_2, \dots, A_n\}$ est indépendante ssi la famille $(A_i)_{1\leqslant i\leqslant n}$ est indépendante.*

☞ Comme le complément du complément de A est A, il suffit de montrer une implication pour avoir l'équivalence. De plus, si on montre que les $\{\overline{A_1}, A_2, \dots, A_n\}$ sont indépendants si les A_i le sont, alors en inversant un par un les événements, on aura obtenu que les $\overline{A_i}$ sont mutuellement indépendants. Il suffit donc de prouver l'indice.

Soit $I \subset \{1,\dots,n\}$, si I ne contient pas 1, alors comme les A_i sont indépendants, on a immédiatement $\mathbf{P}\{\cap_{i\in I} A_i\} = \prod_{i\in I}\mathbf{P}\{A_i\}$. On s'intéresse maintenant au cas où $I = \{1\}\cup I'$. On veut montrer que $\mathbf{P}\{\overline{A_1}\cap\cap_{i\in I'} A_i\} = \mathbf{P}\{\overline{A_1}\}\prod_{i\in I'}\mathbf{P}\{A_i\}$. Comme Ω est l'union disjointe de A_1 et $\overline{A_1}$, on sait que $\cup_{i\in I'} A_i$ est l'union disjointe de $\overline{A_1}\cup\cup_{i\in I'} A_i$ et $A_1\cup\cup_{i\in I'} A_i$. On a donc

$$\mathbf{P}\{\cup_{i\in I'} A_i\} = \mathbf{P}\{\overline{A_1}\cup\cup_{i\in I'} A_i\} + \mathbf{P}\{A_1\cup\cup_{i\in I'} A_i\}.$$

Comme les A_i sont indépendants, on en déduit que

$$\begin{aligned}\mathbf{P}\{\overline{A_1}\cup\cup_{i\in I'} A_i\} &= \prod_{i\in I'}\mathbf{P}\{A_i\} - \mathbf{P}\{A_1\}\prod_{i\in I'}\mathbf{P}\{A_i\}\\ &= (1-\mathbf{P}\{A_1\})\prod_{i\in I'}\mathbf{P}\{A_i\}\\ &= \mathbf{P}\{\overline{A_1}\}\prod_{i\in I'}\mathbf{P}\{A_i\}.\end{aligned}$$

C'est bien ce qu'on voulait, on en conclut que les événements $\{\overline{A_1}, A_2, \dots, A_n\}$ sont mutuellement indépendants.

Exercice 2. *Borne inf sur l'univers*

Soit $(\Omega, F, \mathbf{P}\{\})$ un espace de probabilités. On suppose qu'il existe n événements $A_1, \dots, A_n$ mutuellement indépendants tels que $\mathbf{P}\{A_i\} \notin \{0,1\}$ pour tout i.

1. Montrer que $\mathrm{Card}(\Omega) \geq 2^n$. *Indice : On pourra considérer les événements $A_1^{b_1}\cap A_2^{b_2}\cap\dots\cap A_n^{b_n}$ où $A_i^1 = A_i$ et $A_i^0 = \overline{A_i}$.*

☞ On a

$$\Omega = \sqcup_{(b_1,\dots,b_n)\in\{0,1\}^n} A_1^{b_1}\cap A_2^{b_2}\cap\dots\cap A_n^{b_n},$$

où $\sqcup$ désigne une union disjointe. En effet, chaque élément de Ω est soit dans A_1, soit dans $\overline{A_1}$ et soit dans A_2, soit dans $\overline{A_2}$, etc. L'univers Ω est donc une union de 2^n ensembles disjoints. Si on montre que tous ces ensembles sont non vides, alors on aura bien que $\mathrm{Card}(\Omega) \geq 2^n$. Pour montrer que ces ensembles sont non vides, il suffit de montrer que leur probabilité est non nulle. Mais $\mathbf{P}\left\{\cap_{i=1}^n A_i^{b_i}\right\} = \prod_{i=1}^n \mathbf{P}\left\{A_i^{b_i}\right\}$ car les A_i sont indépendants (pour gérer les complémentaires, utiliser l'exercice "complémentaire des indépendants"). Et par hypothèse, $\mathbf{P}\left\{A_i^{b_i}\right\} \neq 0$ car $\mathbf{P}\{A_i\} \notin \{1,0\}$. Donc le produit est non nul, et on obtient le résultat souhaité.

Exercice 3. *Magnets indépendants*

Deux colocataires, Charlie et Diane, achètent indépendamment une boîte de céréales proposant en cadeau un magnet d'un département français. Cette semaine la région mise à l'honneur est la Corse, chaque boîte de céréales contient donc soit le magnet Corse-du-Sud (2A), soit le magnet Haute-Corse (2B). Charlie et Diane espèrent avoir deux départements différents pour compléter la carte de France sur leur frigo.

On va considérer les trois évènements suivants :

A	Charlie obtient le magnet Corse-du-Sud,
B	Diane obtient le magnet Corse-du-Sud,
C	les deux magnets sont identiques.

1. Montrez que ces trois évènements sont indépendants deux à deux, mais pas mutuellement indépendants.

☞ **Indépendance deux à deux.** Indépendance deux à deux : c'est vrai pour A et B par hypothèse. Montrons le par exemple pour A et C, le cas B et C est symétrique. La seule possibilité pour avoir A et C est que les deux magnets soient des 2A. On a donc $\mathbf{P}\{A\cap C\} = 1/4$, ce qui est égal à $\mathbf{P}\{A\}\mathbf{P}\{C\}$.

Indépendance mutuelle. L'évènement $A\cap B\cap C$ est égal à l'événement $A\cap B$. A et B étant indépendants, $\mathbf{P}\{A\cap B\} = \mathbf{P}\{A\}\mathbf{P}\{B\}$. Donc $\mathbf{P}\{A\cap B\cap C\} = \mathbf{P}\{A\}\mathbf{P}\{B\} = 1/4$. Or $\mathbf{P}\{A\}\mathbf{P}\{B\}\mathbf{P}\{C\} = 1/8$. Les trois évènements ne sont donc pas mutuellement indépendants.

2. Jusqu'ici, nous avons supposé que la probabilité d'avoir le département 2A et celle d'avoir le département 2B étaient égales à $1/2$. En notant p la probabilité d'avoir le département 2A (celle d'avoir le département 2B est alors $1-p$), trouvez les valeurs de p pour lesquelles les évènements A et C sont indépendants.

☞ Si la probabilité d'avoir 2A est p tandis que celle d'avoir 2B est $1-p$, on obtient $\mathbf{P}\{A \cap C\} = p^2$ et $\mathbf{P}\{A\} \times \mathbf{P}\{C\} = p \times (p^2 + (1-p)^2)$. Les deux ne sont égaux que si $2p^3 - 3p^2 + p = 0$, i.e. soit il n'y a que des magnets 2A ($p = 1$), soit que des magnets 2B ($p = 0$), soit $p = \frac{1}{2}$.

3. De façon plus générale, proposer un exemple de n événements $A_1, \ldots, A_n$ tels que pour tout $I \subsetneq \{1, \ldots, n\}$, on ait $\mathbf{P}\{\cap_{i \in I} A_i\} = \prod_{i \in I} \mathbf{P}\{A_i\}$ mais qui ne soient pas mutuellement indépendants (i.e. $\mathbf{P}\{\cap_{i=1}^n A_i\} \neq \prod_{i=1}^n \mathbf{P}\{A_i\}$).

☞ La situation précédente revenait essentiellement à faire deux pile ou face, et à regarder si on obtenait deux fois la même chose. Pour généraliser à n événements, on va chercher cette fois-ci à faire $n-1$ lancers, et à définir un n-ième événement A_n tel que, tant qu'on considère une sous-partie stricte des lancers, on ne peut pas déterminer A_n (et on a $\mathbf{P}\{A_n\} = 1/2$), mais dès lors qu'on connaît tous les lancers, on peut déterminer A_n.

Considérons les événements suivants, avec $n-1$ lancers d'une pièce équilibrée :

A_i	Le $i^{\text{ème}}$ lancer donne *pile*	(pour $1 \leq i < n$),
A_n	Il y a un nombre pair de *pile* obtenus au total.	

Alors pour $1 \leq i < n$, les événements A_i ont une probabilité $1/2$ par hypothèse. De plus, il y a autant d'événements fondamentaux avec un nombre pair et avec un nombre impair de *pile*. Comme ces événements sont équiprobables, on obtient $\mathbf{P}\{A_n\} = 1/2$.

Par hypothèse, les événements $A_1, \ldots, A_{n-1}$ sont mutuellement indépendants. Montrons que les événements $A_1, \ldots, A_n$ ne sont pas mutuellement indépendants.

- Supposons que $n-1$ est pair, alors $\cap_{i=1}^n A_i = \cap_{i=1}^{n-1} A_i$ (si tous les lancers donnent *pile*, on obtient un nombre pair de *pile*). Donc $\mathbf{P}\{\cap_{i=1}^n A_i\} = \mathbf{P}\{\cap_{i=1}^{n-1} A_i\} = (1/2)^{n-1}$ car les $(A_i)_{i \leq n-1}$ sont indépendants. Mais $\prod_{i=1}^n \mathbf{P}\{A_i\} = (1/2)^n$, donc les événements A_i ne sont pas mutuellement indépendants.
- Si $n-1$ est impair, on a $\cap_{i=1}^n A_i = \emptyset$, donc $\mathbf{P}\{\cap_{i=1}^n A_i\} = 0 \neq \prod_{i=1}^n \mathbf{P}\{A_i\}$ et les événements ne sont pas non plus indépendants.

Montrons maintenant que pour tout $I \subsetneq \{1, \ldots, n\}$, on a $\mathbf{P}\{\cup_{i \in I} A_i\} = \prod_{i \in I} \mathbf{P}\{A_i\}$. On a déjà dit que c'était vrai si $n \notin I$, car les événements $\{A_1, \ldots, A_{n-1}\}$ sont indépendants. Supposons donc $n \in I$, on montre le résultat par récurrence sur le nombre d'éléments n'appartenant pas à I.

Initialisation. Soit $I = \{1, \ldots, n\} \setminus \{j\}$ pour un certain $j \neq n$. Alors, on a $\cap_{i \in I} A_i = \cap_{i \neq j, n} A_i \cap A_j$ si $n-1$ est pair et $\cap_{i \in I} A_i = \cap_{i \neq j, n} A_i \cap \overline{A_j}$ si $n-1$ est impair. On en déduit que $\mathbf{P}\{\cap_{i \in I} A_i\} = (1/2)^{n-1} = \prod_{i \in I} \mathbf{P}\{A_i\}$ (pour le premier cas, on utilise l'indépendance des A_i quand $i \leq n-1$, et pour le second cas, avec $\overline{A_j}$, on peut utiliser l'exercice "indépendance des compléments", ou simplement dire que $\overline{A_j}$ est indépendant des autres A_i).

Récurrence. Soit $I = \{1, \ldots, n\} \setminus \{j_1, \ldots, j_k\}$ pour un certain $k \geq 2$. On a alors que $\cap_{i \in I} A_i$ est l'union disjointe de $(\cap_{i \in I} A_i) \cap A_{j_1}$ et $(\cap_{i \in I} A_i) \cap \overline{A_{j_1}}$. Par hypothèse de récurrence, $\mathbf{P}\{\cap_{i \in I} A_i\} = \mathbf{P}\left\{(\cap_{i \in I} A_i) \cap A_{j_1}\right\} + \mathbf{P}\left\{(\cap_{i \in I} A_i) \cap \overline{A_{j_1}}\right\} = (1/2)^{n-(k-1)} + (1/2)^{n-(k-1)} = (1/2)^{n-k} = \prod_{i \in I} \mathbf{P}\{A_i\}$ (on utilise encore l'indépendance des compléments).

On conclut que $\mathbf{P}\{\cup_{i \in I} A_i\} = \prod_{i \in I} \mathbf{P}\{A_i\}$ pour tout $I \subsetneq \{1, \ldots, n\}$.

4. Quel lien y a-t-il entre "être indépendants 2 à 2" et "être indépendants mutuellement" ?

☞ Être indépendant mutuellement implique être indépendant 2 à 2, mais pas l'inverse (la question précédente donne un contre exemple). C'est l'inverse de "premiers entre eux 2 à 2" et "premiers entre eux dans leur ensemble".

Exercice 4. *Générer une loi uniforme*

1. Définir une variable aléatoire ayant une loi uniforme sur $\{1, 2, 3\}$ à partir d'une suite infinie de bits aléatoires. Est-ce possible à partir d'une suite finie ?

☞ A partir d'une suite infinie, on peut associer les valeurs 1, 2, 3 à $(00), (01), (10)$ et si on tire le couple (11), le rejeter et tirer à nouveau. A partir d'une suite finie, on ne peut générer que des probabilités en $\frac{1}{2^k}$. En effet, la variable aléatoire X uniforme sur $\{1, 2, 3\}$ est une fonction de B, les différentes suites de bits. Donc $\mathbb{P}(X = 1) = \sum_{b \in X^{-1}(1)} \frac{1}{2^n} = \frac{|X^{-1}(1)|}{2^n}$.

Exercice 5. *Monty Hall*

Monty est le présentateur d'un jeu télévisé qui se déroule de la manière suivante. Il y a trois rideaux : derrière l'un, il y a une voiture à gagner, et derrière chacun des deux autres, il y a une chèvre. Bob, le participant, doit choisir un rideau, en espérant choisir celui derrière lequel se cache la voiture. Appelons le rideau que Bob choisi le rideau 1.

Ensuite, Monty ouvre l'un des deux autres rideaux, où se cache une chèvre (s'il y a une chèvre derrière les 2 autres rideaux, Monty en choisit un uniformément au hasard). Supposons que le rideau que Monty ouvre est le rideau 2. Bob doit ensuite décider s'il garde le rideau qu'il a choisi au début, ou bien s'il change avec l'autre rideau fermé restant. Une fois ce choix effectué, Monty ouvre le rideau choisi par Bob et Bob gagne ce qui se cache derrière.

1. Est-ce que Bob a intérêt de changer de rideau à l'étape intermédiaire, ou bien cela ne fait-il aucune différence ? (on suppose bien évidemment que Bob préfère gagner la voiture plutôt qu'une chèvre.... !)

☞ Oui, Bob a intérêt de changer de rideau.

Stratégie 1 : ne pas changer. Si Bob décide de ne pas changer de porte, alors la probabilité qu'il gagne la voiture est la probilité qu'il ait choisi la bonne porte lors de son choix initial, soit 1/3. Plus formellement, appelons 1 la porte que Bob a choisi, $C \in \{1,2,3\}$ l'emplacement de la voiture et $M \in \{1,2,3\}$ la porte que Monty ouvre. Les données que l'on a sont que $M = 2$ et on veut determiner la probabilité que la voiture se trouve en 1 :

$$\begin{aligned}\mathbf{P}\{C=1|M=2\} &= \frac{\mathbf{P}\{M=2|C=1\}\,\mathbf{P}\{C=1\}}{\mathbf{P}\{M=2|C=1\}\,\mathbf{P}\{C=1\}+\mathbf{P}\{M=2|C=2\}\,\mathbf{P}\{C=2\}+\mathbf{P}\{M=2|C=3\}\,\mathbf{P}\{C=3\}}\\ &= \frac{1/2\cdot 1/3}{1/2\cdot 1/3+0\cdot 1/3+1\cdot 1/3} = \frac{1}{3}\,.\end{aligned}$$

Stratégie 2 : changer. Si Bob décide de changer de porte, alors son deuxième rideau choisi cache la voiture si et seulement si son premier rideau choisi ne la cachait pas. Or à l'état initial, Bob a probabilité 2/3 de choisir un rideau derrière lequel se cache une chèvre. Lorsque Bob change

$$\begin{aligned}\mathbf{P}\{C=3|M=2\} &= \frac{\mathbf{P}\{M=2|C=3\}\,\mathbf{P}\{C=3\}}{\mathbf{P}\{M=2|C=1\}\,\mathbf{P}\{C=1\}+\mathbf{P}\{M=2|C=2\}\,\mathbf{P}\{C=2\}+\mathbf{P}\{M=2|C=3\}\,\mathbf{P}\{C=3\}}\\ &= \frac{1\cdot 1/3}{1/2\cdot 1/3+0\cdot 1/3+1\cdot 1/3} = \frac{2}{3}\,.\end{aligned}$$

On voit donc que la stratégie 2 permet de gagner avec probabilité 2/3 alors que la stratégie 1 ne permet de gagner qu'avec probabilité 1/3.

Exercice 6. *Météo*

Le 14 juillet à Saint Troupaize il fait beau 7 fois sur 10. Le comité des fêtes dispose de deux sources de prévision météorologiques qui se trompent de façon indépendante (la météo nationale n'utilise pas de grenouilles) :

— la météo nationale qui se trompe deux fois sur 100,
— une grenouille verte qui se trompe une fois sur 20.

1. La météo annonce de la pluie alors que le comportement de la grenouille laisse prévoir du beau temps. Déterminez le temps le plus probable.

☞ Notons de la façon suivante les différents évènements :
— B : Il fait beau.
— M : La météo a raison.
— G : La grenouille a raison.
— A : La météo prévoit de la pluie tandis que la grenouille prévoit du beau temps.

La formule de Bayes nous permet d'affirmer :

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B)+P(A|\bar{B})P(\bar{B})}$$

Or $P(A|B) = \frac{2}{100}\times\frac{19}{20} = \frac{38}{2000}$ et $P(A|\bar{B}) = \frac{98}{100}\times\frac{1}{20} = \frac{98}{2000}$. Donc

$$P(B|A) = \frac{38\times 7}{38\times 7+98\times 3} = 0,475.$$

Le plus probable est donc qu'il fasse mauvais.

Autre méthode :
Il y a huit cas : il fait beau ou non, la grenouille a raison ou non, la météo a raison ou non. Mais parmi ces cas il n'y a que deux possibilités car on connaît les prédictions de la météo et de la grenouille. On veut donc comparer $P(B\cap\bar{M}\cap G)$ et $P(\bar{B}\cap M\cap\bar{G})$. Or par indépendance $P(B\cap\bar{M}\cap G) = \frac{7}{10}\frac{2}{100}\frac{19}{20} = \frac{133}{10\,000}$ et $P(\bar{B}\cap M\cap\bar{G}) = \frac{3}{10}\frac{98}{100}\frac{1}{20} = \frac{147}{10\,000}$.
Donc il est plus probable qu'il fasse mauvais.

NB : si $\mathbf{P}\{A \mid B\} = \mathbf{P}\{A \mid \bar{B}\}$, alors A et B sont indépendants :

$$\mathbf{P}\{A\} = \mathbf{P}\{A\cap B\} + \mathbf{P}\{A\cap\bar{B}\} = \mathbf{P}\{A \mid B\}\,(\mathbf{P}\{B\}+\mathbf{P}\{\bar{B}\}) = \mathbf{P}\{A \mid B\},$$

d'où $\mathbf{P}\{A\cap B\} = \mathbf{P}\{A \mid B\}\,\mathbf{P}\{B\} = \mathbf{P}\{A\}\,\mathbf{P}\{B\}$.

Exercice 7. *Formule de Poincaré*

On se donne une suite d'évènements $A_1, A_2, \ldots, A_n$.

1. Démontrez la formule suivante, due à Poincaré :

$$\mathbf{P}\left\{\bigcup_{1\leq i\leq n} A_i\right\} = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leq i_1<\cdots<i_k\leq n} \mathbf{P}\left\{\bigcap_{1\leq j\leq k} A_{i_j}\right\}.$$

☞ Démonstration par récurrence sur $n \geq 2$:

Pour $n = 2$, on retrouve $\mathbf{P}\{A \cup B\} = \mathbf{P}\{A\} + \mathbf{P}\{B\} - \mathbf{P}\{A \cap B\}$. ok.

Pour $n \geq 2$, on a :

$$\begin{aligned}
\mathbf{P}\left\{\bigcup_{1\leq i\leq n+1} A_i\right\} &= \mathbf{P}\left\{A_{n+1} \cup \bigcup_{1\leq i\leq n} A_i\right\} \\
&= \mathbf{P}\{A_{n+1}\} + \mathbf{P}\left\{\bigcup_{1\leq i\leq n} A_i\right\} - \mathbf{P}\left\{A_{n+1} \cap \bigcup_{1\leq i\leq n} A_i\right\} \text{ (cas } n=2) \\
&= \mathbf{P}\{A_{n+1}\} + \mathbf{P}\left\{\bigcup_{1\leq i\leq n} A_i\right\} - \mathbf{P}\left\{\bigcup_{1\leq i\leq n} A_i \cap A_{n+1}\right\} \\
&= \mathbf{P}\{A_{n+1}\} + \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leq i_1<\cdots<i_k\leq n} \mathbf{P}\left\{\bigcap_{1\leq j\leq k} A_{i_j}\right\} \\
&\quad - \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leq i_1<\cdots<i_k\leq n} \mathbf{P}\left\{\bigcap_{1\leq j\leq k} A_{i_j} \cap A_{n+1}\right\} \text{ (HR } \times 2) \\
&= \mathbf{P}\{A_{n+1}\} + \sum_{k=1}^{n} (-1)^{k-1} \sum_{1\leq i_1<\cdots<i_k<n+1} \mathbf{P}\left\{\bigcap_{1\leq j\leq k} A_{i_j}\right\} \\
&\quad + \sum_{k=2}^{n+1} (-1)^{k-1} \sum_{1\leq i_1<\cdots<i_k=n+1} \mathbf{P}\left\{\bigcap_{1\leq j\leq k} A_{i_j}\right\} \\
&= \mathbf{P}\{A_{n+1}\} + \sum_{i_1=1}^{n} \mathbf{P}\{A_{i_1}\} + \sum_{k=2}^{n+1} (-1)^{k-1} \sum_{1\leq i_1<\cdots<i_k\leq n+1} \mathbf{P}\left\{\bigcap_{1\leq j\leq k} A_{i_j}\right\} \\
&= \sum_{k=1}^{n+1} (-1)^{k-1} \sum_{1\leq i_1<\cdots<i_k\leq n+1} \mathbf{P}\left\{\bigcap_{1\leq j\leq k} A_{i_j}\right\}
\end{aligned}$$

Exercice 8. *Collisions*

On considère une fonction de hachage qui, étant donné un fichier, produit une empreinte sous la forme d'une chaîne de 128 bits. On souhaite trouver une collision (c'est-à-dire deux fichiers différents ayant même empreinte) en hachant des fichiers générés aléatoirement.

1. Combien de fichiers faut-il générer pour que la probabilité d'observer une collision soit supérieure à $1/2$?

☞ Notons $N = 2^{128}$ le nombre d'empreintes possibles. Pour $2 \leq k \leq N$, on considère les événements suivants :
C^k : il existe une collision dans les k premiers fichiers,
D^k : les k premiers fichiers sont tous différents.
Il est un peu plus facile de calculer la probabilité des D^k : c'est le nombre d'arrangements de k parmi N, noté A_k^N, divisé par le nombre de suites de k empreintes possibles, soit N^k. On a

$$A_k^N = \frac{N!}{(N-k)!}$$

donc

$$\mathbf{P}\left\{D^k\right\} = \frac{A_k^N}{N^k} = \frac{N!}{(N-k)! \times N^k}.$$

Pour k fichiers, la probabilité d'avoir une collision est donc

$$\mathbf{P}\left\{C^k\right\} = 1 - \mathbf{P}\left\{D^k\right\} = 1 - \frac{(2^{128})!}{(2^{128}-k)! \times (2^{128})^k}.$$

Pour répondre à la question initiale, il "suffit" de faire l'application numérique et de trouver pour quel k cette probabilité dépasse $\frac{1}{2}$. En pratique, c'est un peu plus compliqué, parce que $(2^{128})!$, c'est gros... Une approximation avec

$$\mathbf{P}\left\{D^k\right\} = 1 - \mathrm{e}^{-\frac{k(k-1)}{2N}}$$

nous donne les valeurs suivantes pour $N = 2^{128}$:
- Pour $k = 2^{64}$ fichiers, $\mathbf{P}\left\{C^k\right\} \approx 39,35\%$,
- Pour $k = 2^{65}$ fichiers, $\mathbf{P}\left\{C^k\right\} \approx 86,47\%$.

On peut regarder la formule avec un N avec une taille plus raisonnable : pour un encodage sur 8 bits, soit $N = 2^8$, on obtient

$$\mathbf{P}\left\{C^{19}\right\} \approx 48,73\% \qquad \text{et} \qquad \mathbf{P}\left\{C^{20}\right\} \approx 52,39\%,$$

donc il suffit de 20 fichiers pour générer une collision avec probabilité $\frac{1}{2}$.

Ce problème est aussi connu sous le nom de *paradoxe des anniversaires*.

2. Considérons un groupe de personnes. En supposant que les jours de naissance sont équiprobables (et qu'aucune personne n'est née un 29 février), à partir de combien de personnes la probabilité d'avoir deux personnes avec le même anniversaire dépasse-t-elle 1/2 ?

   ☞ On reprend les calculs de la question précédente, avec $N = 365$. On obtient $\mathbf{P}\left\{D^{22}\right\} \approx 47,57\%$ et $\mathbf{P}\left\{D^{23}\right\} \approx 50,73\%$. Le « paradoxe » est ici qu'il suffit de 23 personnes pour avoir une chance sur deux d'observer une collision, alors qu'on a l'impression que c'est un évènement beaucoup plus rare – et donc nécessitant plus de personnes pour se produire.

3. Quelle est la probabilité pour que deux personnes actuellement dans la pièce partagent la même date d'anniversaire (toujours sous les hypothèses d'équiprobabilité et en négligeant le 29 février) ?

   ☞ Question précédente, avec $N = 365$ et k probablement entre 10 et 25...

4. Quelle est la probabilité pour qu'une personne actuellement dans la pièce partage la même date d'anniversaire que vous ?

   ☞ Cette fois, on cherche la propabilité qu'une personne parmi les $k-1$ restantes aie une date d'anniversaire précise. Ce qui nous donne une probabilité de $1 - (\frac{364}{365})^{k-1}$. Avec 23 personnes, on trouve 5,86% de chance que quelqu'un partage votre date d'anniversaire. Pour avoir une probabilité supérieure à $\frac{1}{2}$, il faudrait 253 personnes !

Exercice 9. *Jouons au poker !*

Comptez les probabilités des combinaisons de cartes au poker (première disposition sans joker).

Le Poker se joue avec un jeu de 52 cartes, formé de 4 couleurs (trèfle, carreau, cœur, pique) contenant chacune 13 figures (2, 3,. . ., 10, valet, dame, roi, as). L'as est la carte la plus forte mais il peut parfois être considéré comme la plus faible (Quinte et quinte flush).
Les combinaisons que l'on peut faire au poker sont :
— La paire : 2 cartes de même valeur qui ne forment pas avec les autres cartes une double paire, un full ou un carré.
— La double paire : 2 séries différentes de deux cartes de même valeur qui ne forment pas avec les autres cartes un full.
— Le brelan : trois cartes de la même valeur mais qui ne forment pas avec les autres cartes un carré ou un full.
— La quinte ou suite : 5 cartes consécutives de différentes couleurs.
— La couleur ou flush : 5 cartes de la même couleur.
— Le full : Trois cartes de la même valeur d'une part et deux cartes de même valeur d'autre part.
— Le carré : 4 cartes de la même valeur.
— La quinte flush : 5 cartes consécutives d'une même couleur.