HEC-ESSEC 2020 - VOIE E

Correction proposée par G. Dupont pour http://maths-concours.fr.

On s'intéresse dans ce sujet au problème de la *double dépense* de *bitcoins* par un groupe d'individus mal intentionnés.
On rappelle que le *bitcoin* est une monnaie virtuelle dont l'utilisation pour des transactions est associée à une structure unique appelée *blockchain*, partagée sur le réseau des usagers de cette monnaie et ayant pour but de sécuriser ces transactions.
La modélisation étudiée ne nécessite pas de connaissances particulières sur le *bitcoin* et la *blockchain*.

Partie I – Deux résultats généraux

On démontre dans cette partie deux résultats préliminaires, aux questions 5 et 6. Ces résultats seront utilisés dans la suite du sujet et pourront être admis.

Calcul d'une probabilité

Soit X et Y deux variables aléatoires sur un espace probabilisé, à densité et indépendantes.
On note F_X et F_Y les fonctions de répartition de X et Y.
On suppose que Y est à valeurs positives et possède une densité f_Y dont la restriction à $[0, +\infty[$ est continue sur cet intervalle.
Pour tout $x \in \mathbf{R}_+$, on pose $H(x) = \mathbf{P}([X \leq Y] \cap [Y \leq x])$.

1. (a) Montrer que H est une fonction croissante sur $\mathbf{R}_+$ qui admet une limite finie en $+\infty$.
   (b) En utilisant la suite $(H(n))_{n\in\mathbf{N}}$, montrer que $\lim\limits_{x\to+\infty} H(x) = \mathbf{P}([X \leq Y])$.

   Que vaut $H(0)$?
2. Soit (u, v) un couple de réels positifs tels que $u < v$.
   (a) Montrer que $H(v) - H(u) = \mathbf{P}([X \leq Y] \cap [u < Y \leq v])$ puis que :

$$F_X(u)\frac{F_Y(v) - F_Y(u)}{v-u} \leq \frac{H(v) - H(u)}{v-u} \leq F_X(v)\frac{F_Y(v) - F_Y(u)}{v-u}.$$

   (b) En déduire que pour tout $x \in \mathbf{R}_+$, H est dérivable en x et $H'(x) = F_X(x)f_Y(x)$.
   (c) En conclure que pour tout x réel positif, $H(x) = \displaystyle\int_0^x F_X(t)f_Y(t)\ dt$.
3. Montrer que $\mathbf{P}([X \leq Y]) = \displaystyle\int_0^{+\infty} F_X(t)f_Y(t)\ dt$.
4. En utilisant la fonction $K : x \mapsto \mathbf{P}([X < Y] \cap [Y \leq x])$, on montrerait de même et nous l'admettrons que :

$$\mathbf{P}([X < Y]) = \int_0^{+\infty} F_X(t)f_Y(t)\ dt = \mathbf{P}([X \leq Y]).$$

   Que peut-on en déduire pour $\mathbf{P}([X = Y])$?
5. *Application aux lois exponentielles*
   On suppose que U et V sont deux variables aléatoires indépendantes suivant des lois exponentielles de paramètres respectifs λ et μ, réels strictement positifs.
   Soit θ un réel positif ou nul.
   (a) Déterminer la fonction de répartition de la variable aléatoire $X = U - \theta$.
   (b) En déduire que pour tout $\theta \geq 0$,

$$\mathbf{P}([U - \theta \leq V]) = 1 - \frac{\mu}{\lambda+\mu}e^{-\lambda\theta}.$$

Inégalité de Boole

6. On considère $(B_k)_{k\in\mathbf{N}^*}$ une famille d'événements d'un espace probabilisé.
   (a) Montrer par récurrence sur $n \in \mathbf{N}^*$ que $\mathbf{P}\left(\bigcup_{k=1}^{n} B_k\right) \leq \sum_{k=1}^{n} \mathbf{P}(B_k)$.
   (b) On suppose que la série $\sum_{k\geq 1} \mathbf{P}(B_k)$ converge. Montrer que :

$$\mathbf{P}\left(\bigcup_{k\geq 1} B_k\right) \leq \sum_{k=1}^{+\infty} \mathbf{P}(B_k).$$

Partie II – Une compétition entre deux groupes

Dans toute la suite du sujet, on désigne par p un réel de l'intervalle $]0,1[$ et on pose $q = 1-p$.
On modélise une compétition entre deux groupes d'individus A et B avec les règles suivantes :

— Le groupe A doit résoudre une suite de problèmes $(P_k)_{k\geq 1}$ dans l'ordre des indices. Au temps $t = 0$, le groupe commence la résolution du problème P_1, ce qui lui prend un temps représenté par la variable aléatoire X_1. Une fois P_1 résolu, le groupe aborde immédiatement le problème P_2, et on note X_2 le temps consacré à la résolution de P_2 par le groupe A, et ainsi de suite.
Pour tout $k \in \mathbf{N}^*$, on note X_k la variable aléatoire donnant le temps consacré à la résolution du problème P_k par le groupe A.

— De même, le groupe B doit résoudre dans l'ordre une suite de problèmes $(Q_k)_{k\geq 1}$; la résolution du premier problème Q_1 commence au temps $t = 0$ et on note, pour tout $k \in \mathbf{N}^*$, Y_k la variable aléatoire donnant le temps consacré par le groupe B à la résolution du problème Q_k.

— À ce jeu est associé à un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ sur lequel sont définies les suites de variables aléatoire $(X_k)_{k\geq 1}$ et $(Y_k)_{k\geq 1}$ et on fait les hypothèses suivantes :
   — pour tout $k \in \mathbf{N}^*$, X_k suit la loi exponentielle de paramètre p, notée $\mathcal{E}(p)$, et Y_k suit la loi exponentielle $\mathcal{E}(q)$;
   — pour tout $k \in \mathbf{N}^*$, les variables aléatoires $X_1, \ldots, X_k, Y_1, \ldots, Y_k$ sont indépendantes.

— On établit alors la liste de tous les problèmes résolus *dans l'ordre où ils le sont par les deux groupes*. En cas de simultanéité temporelle de la résolution par les deux groupes d'un de leurs problèmes, on placera d'abord le problème résolu par A dans la liste puis celui résolu par B.
Pour tout $n \in \mathbf{N}^*$, on note U_n la variable aléatoire de Bernoulli associée à l'événement « le n-ième problème placé dans la liste est un problème résolu par le groupe A ».
Par exemple, si la liste des cinq premiers problèmes résolus est $(P_1, P_2, Q_1, P_3, Q_2)$, alors $U_1 = 1$, $U_2 = 1$, $U_3 = 0$, $U_4 = 1$ et $U_5 = 0$.

— Pour tout $n \geq 0$, on note aussi S_n la variable aléatoire donnant le nombre de problèmes qui ont été résolus par A présents dans la liste des n premiers problèmes résolus. En particulier, S_0 vaut toujours 0.

7. (a) Que représente la variable aléatoire $\sum_{k=1}^{n} X_k$?
   (b) On suppose que $X_1 = 5$, $X_2 = 2$, $X_3 = 3$, $X_4 = 2$, $Y_1 = 2$, $Y_2 = 2$, $Y_3 = 4$, $Y_4 = 2$.
   Déterminer $U_1, \ldots, U_7$.
   Peut-on aussi en déduire la valeur de U_8 ?
   (c) Compléter le script Scilab suivant pour qu'il simule le jeu et, pour n, p donnés, affiche la liste des valeurs $U_1, \ldots, U_n$:

```
p = input('p=')
n = input('n=')
q = 1-p

U = zeros(1,n)

sommeX = grand(1,1,'exp',1/p)
sommeY = grand(1,1,'exp',1/q)
```

```
mini = min(sommeX,sommeY)

for k=1:n
    if sommeX == ...
        U(k) = ...
        sommeX = sommeX + grand(1,1,'exp',1/p)
    else
        sommeY = ...
    end
    mini = min(sommeX,sommeY)
end
...
```

(d) Quelle(s) instruction(s) faut-il ajouter pour afficher la valeur de S_n ?

8. *Loi de U_n*

Dans cette question, on démontre par récurrence sur $n \geq 1$ que $\mathbf{P}([U_n = 1]) = p$.

(a) Montrer que $\mathbf{P}([U_1 = 1]) = \mathbf{P}([X_1 \leq Y_1]) = p$.

(b) i. Montrer que pour tout réel $x < 0$, $\mathbf{P}_{[U_1=1]}([Y_1 - X_1 \leq x]) = 0$.

ii. Soit x un réel positif ou nul.

Établir : $\mathbf{P}_{[U_1=1]}([Y_1 - X_1 \leq x]) = \dfrac{1}{p}\mathbf{P}([X_1 \leq Y_1 \leq X_1 + x])$, puis calculer $\mathbf{P}_{[U_1=1]}([Y_1 - X_1 \leq x])$.

(c) On peut interpréter ce résultat en disant que la *loi conditionnelle de $Y_1 - X_1$ sachant* $[U_1 = 1]$ est une loi exponentielle. Quel est son paramètre ?

Par analogie, quelle est la loi conditionnelle de $X_1 - Y_1$ sachant $[U_1 = 0]$? (on n'attend pas une démonstration précise mais un argument de bon sens pour justifier le résultat proposé).

(d) On suppose que $n \in \mathbf{N}^*$ et $\mathbf{P}([U_n = 1]) = p$.

Déduire de cette hypothèse et de la question précédente que $\mathbf{P}_{[U_1=1]}([U_{n+1} = 1]) = p$ et $\mathbf{P}_{[U_1=0]}([U_{n+1} = 1]) = p$.

(e) Conclure

9. On montrerait aussi par récurrence, et nous l'admettrons, que pour tout $n \in \mathbf{N}^*$, les variables aléatoires $U_1, \ldots, U_n$ sont mutuellement indépendantes.

En déduire la loi de S_n.

Soit $r \in \mathbf{N}$. On s'intéresse, dans les questions qui suivent à la probabilité a_r de l'événement :
A_r : « il existe un $n \geq r$ tel que, lorsque n problèmes en tout ont été résolus, le groupe A en a résolu r de plus que le groupe B ».

10. (a) Justifier que $a_0 = 1$.

(b) Montrer que pour tout $r \geq 1$, $\mathbf{P}_{[U_1=1]}(A_r) = \mathbf{P}(A_{r-1})$ et $\mathbf{P}_{[U_1=0]}(A_r) = \mathbf{P}(A_{r+1})$.

(c) En déduire que pour tout $r \geq 1$, $a_{r+1} = \dfrac{1}{q}a_r - \dfrac{p}{q}a_{r-1}$.

(d) En remarquant que $1 - 4pq = (1-2p)^2$, donner une expression de a_r en fonction de p, q, r et de deux constantes que l'on introduira.

11. Le cas $p \geq \dfrac{1}{2}$.

Montrer que, dans les cas $p = \dfrac{1}{2}$ et $p > \dfrac{1}{2}$, la suite $(a_r)_{r\in\mathbf{N}}$ est constante égale à 1.

12. Le cas $p < \dfrac{1}{2}$.

(a) Soit k un entier naturel.

i. Établir : $A_{2k} = \bigcup_{i\geq k} [S_{2i} = i + k]$.

ii. Montrer que pour tout $i \geq k$, on a $\mathbf{P}([S_{2i} = i + k]) = \binom{2i}{i+k} p^{i+k} q^{i-k}$.

iii. Après avoir donné la valeur de la somme $\sum_{j=0}^{2i} \binom{2i}{j}$, montrer que pour tout entier $i \geq k$, $\binom{2i}{i+k} \leq 4^i$.

iv. En déduire l'inégalité :

$$\sum_{i=k}^{+\infty} \mathbf{P}\left([S_{2i} = i + k]\right) \leq \left(\frac{p}{q}\right)^k \frac{(4pq)^k}{1 - 4pq}.$$

(b) Montrer en utilisant l'inégalité de Boole (voir question 6) que si $p < \frac{1}{2}$, alors $\lim\limits_{k \to +\infty} a_{2k} = 0$.

(c) Conclure en utilisant la question 10.d que si $p < \frac{1}{2}$, alors pour tout entier naturel r, $a_r = \left(\frac{p}{q}\right)^r$.

On a ainsi établi dans les questions 11 et 12 :

$$\forall r \in \mathbf{N}, \quad a_r = \begin{cases} \left(\frac{p}{q}\right)^r & \text{si} \quad p < \frac{1}{2} \\ 1 & \text{si} \quad p \geq \frac{1}{2}. \end{cases}$$

Ce résultat pourra être admis et utilisé dans la suite du sujet.

Partie III – La *blockchain* et la stratégie de la *double dépense*

On utilise, dans cette partie, les notations et résultats de la partie II.
Soit n un entier supérieur ou égal à 1.
La *blockchain* est formée d'une suite de blocs, chacun associé à plusieurs transactions. Elle contient l'historique de toutes les transactions effectuées depuis la création du *bitcoin*.
Avant d'être placé dans la *blockchain*, un nouveau bloc doit être validé. Cette validation nécessite la mise en œuvre d'une grande puissance de calcul pour résoudre un problème dépendant fortement du contenu du bloc et des blocs qui le précédent.
Les individus qui valident les blocs sont appelés mineurs.
Il est possible qu'à un instant donné, coexistent sur le réseaux deux *blockchains*, valides et différentes. Dans ce cas, le réseau choisira celle qui comporte le plus de blocs et l'autre sera abandonnée.
Par prudence, lorsqu'un bloc est validé, il est recommandé d'attendre que $n - 1$ blocs le suivant soient aussi validés pour considérer que les transactions incluses dans le bloc soient honnêtes.
Un groupe de mineurs mal intentionnés, noté A, peut essayer de dépenser deux fois les mêmes *bitcoins* en procédant ainsi :

- Le groupe A demande la validation de l'achat d'un bien d'un montant de s *bitcoins* qu'il a en sa possession.
- Lorsqu'un bloc K incluant cette transaction est proposé à la validation sur le réseau, A modifie ce bloc en K', qu'il ne diffuse pas, en remplaçant l'achat par une vente de s *bitcoins* en euros à son profit par exemple. Il se met alors à la validation de ce nouveau bloc et crée ainsi une deuxième instance de la *blockchain* qu'il continue à développer sans la diffuser.
- Lorsque le groupe B, représentant l'ensemble des autres mineurs du réseau, a validé K ainsi que les $n - 1$ blocs suivants, le vendeur du bien considère que la transaction est valide et fournit le bien.
- Le groupe A attend alors d'avoir une *blockchain* plus longue que celle de B, qui est publique, pour la diffuser donc invalider la *blockchain* publique et l'achat du bien. Le crédit en *bitcoins* du vendeur du bien est alors annulé.

On reprend et on complète la modélisation de la partie précédente pour déterminer la probabilité que la stratégie de la *double dépense* réussisse et le choix de n pour que cette probabilité soit faible.
Une première phase du jeu, décrit dans la partie II, s'achève à l'instant aléatoire t où le problème Q_n est ajouté à la liste des problèmes résolus.
Le groupe de mineurs A est ensuite déclaré vainqueur s'il se trouve un instant $t' \geq t$ où le nombre de problèmes résolus par A dans la liste des problèmes résolus depuis le début du jeu, est strictement supérieur au nombre de ceux résolus par B dans cette même liste. On note G_n cet événement.
On détermine, dans cette partie, la probabilité de G_n en fonction de n et de p.

13. On s'intéresse tout d'abord à la loi de la variable aléatoire T_n égale au nombre de problèmes résolus par le groupe A lorsque l'on place Q_n dans la liste des problèmes résolus.

(a) Montrer que pour tout $k \in \mathbf{N}$, $[T_n = k] = [S_{n+k-1} = k] \cap [U_{n+k} = 0]$.

(b) En déduire que $\mathbf{P}\left([T_n = k]\right) = \binom{n+k-1}{k} p^k q^n$.

14. (a) En utilisant la formule des probabilités totales, établir :

$$\mathbf{P}\left(G_n\right) = \mathbf{P}\left([T_n \geq n+1]\right) + \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right) a_{n+1-k}.$$

(b) Dans le cas où $p \geq \dfrac{1}{2}$, en déduire que $\mathbf{P}\left(G_n\right) = 1$.

(c) De même lorsque $p < \dfrac{1}{2}$, montrer que :

$$\mathbf{P}\left(G_n\right) = 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} (p^k q^n - p^{n+1} q^{k-1}).$$

15. *Une meilleure expression de* $\mathbf{P}\left(G_n\right)$ *lorsque* $p < \dfrac{1}{2}$.

Pour tout $x \in [0,1]$ et $n \in \mathbf{N}^*$, on pose :

$$u_n(x) = (1-x)^n \sum_{k=0}^{n} \binom{n+k-1}{k} x^k.$$

(a) Vérifier que pour tout $n \in \mathbf{N}^*$: $\mathbf{P}\left(G_n\right) = 1 - u_n(p) + \dfrac{p}{q} u_n(q)$.

(b) Pour tout $x \in [0,1]$ et $n \in \mathbf{N}^*$, établir la relation :

$$u_{n+1}(x) = u_n(x) + (1-x)^n x^{n+1} \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} x \right).$$

(c) En déduire que pour tout $n \in \mathbf{N}^*$:

$$\mathbf{P}\left(G_{n+1}\right) = \mathbf{P}\left(G_n\right) - \left(1 - \frac{p}{q}\right) (pq)^{n+1} \binom{2n+1}{n+1}.$$

(d) Montrer par récurrence, que pour tout $n \in \mathbf{N}^*$:

$$\mathbf{P}\left(G_n\right) = \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k.$$

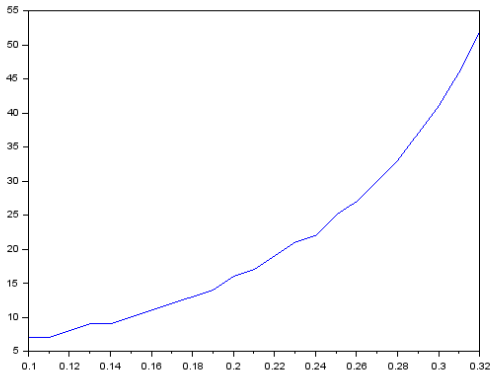
16. *Application à la sécurisation des transactions*

Connaissant $p < \dfrac{1}{2}$, on cherche à limiter le risque que la stratégie mise en place par le groupe de mineurs A réussisse.

(a) Après avoir établi la formule $\binom{n}{k} = \dfrac{n}{k}\binom{n-1}{k-1}$ lorsque $k \in [\![1, n]\!]$, écrire une fonction Scilab qui calcule les coefficients binomiaux.

(b) Écrire un script Scilab qui détermine n_p, le plus petit entier n tel que $\mathbf{P}\left(G_n\right) \leq \epsilon$ pour $p < \dfrac{1}{2}$ et $\epsilon > 0$ saisis au clavier par l'utilisateur.

NB : Pour $\epsilon = 10^{-4} = 0,1‰$ et p variant entre 10% et 32%, on obtient pour la représentation de n_p en fonction de p :



Solution :

Partie I – Deux résultats généraux

Calcul d'une probabilité

1. (a) Soient $x, y \in \mathbf{R}_+$. On a :

$$\begin{aligned} x \leq y &\Rightarrow [Y \leq x] \subset [Y \leq y] \\ &\Rightarrow [X \leq Y] \cap [Y \leq x] \subset [X \leq Y] \cap [Y \leq y] \\ &\Rightarrow \mathbf{P}\left([X \leq Y] \cap [Y \leq x]\right) \leq \mathbf{P}\left([X \leq Y] \cap [Y \leq y]\right) \\ &\Rightarrow H(x) \leq H(y). \end{aligned}$$

Ainsi :

$$\boxed{H \text{ est une fonction croissante sur } \mathbf{R}_+.}$$

(b) La suite $([Y \leq n])_{n\in\mathbf{N}}$ est une suite croissante d'événements tels que $\bigcup_{n\in\mathbf{N}} [Y \leq n] = \Omega$. Ainsi, $[X \leq Y] = \bigcup_{n\in\mathbf{N}} ([X \leq y] \cap [Y \leq n])$ et donc

$$\mathbf{P}\left([X \leq Y]\right) = \mathbf{P}\left(\bigcup_{n\in\mathbf{N}} \left([X \leq Y] \cap [Y \leq n]\right)\right)$$

mais puisque la suite d'événements est croissante, il suit du théorème de la limite croissante que :

$$\mathbf{P}\left(\bigcup_{n\in\mathbf{N}} \left([X \leq Y] \cap [Y \leq n]\right)\right) = \lim_{n\to+\infty} \mathbf{P}\left([X \leq Y] \cap [Y \leq n]\right) = \lim_{n\to+\infty} H(n).$$

La fonction H étant croissante et majorée par 1, elle admet une limite finie en $+\infty$ et alors $\lim_{n\to+\infty} H(n) = \lim_{x\to+\infty} H(x)$ de sorte que :

$$\boxed{\mathbf{P}\left([X \leq Y]\right) = \lim_{x\to+\infty} H(x).}$$

En outre, $H(0) = \mathbf{P}\left(Y \leq 0\right) = 0$ car Y est une variable aléatoire à densité à valeurs dans $\mathbf{R}_+$.

2. Soit (u, v) un couple de réels positifs tels que $u < v$.

(a) On a $[X \leq Y] \cap [X \leq u] \subset [X \leq Y] \cap [X \leq v]$ donc

$$\mathbf{P}\left([X \leq Y] \cap [u < X \leq v]\right) = \mathbf{P}\left(([X \leq Y] \cap [X \leq v]) \setminus ([X \leq Y] \cap [X \leq u])\right) = H(v) - H(u).$$

Par ailleurs, par indépendance de X et Y, on a :

$$F_X(u)(F_Y(v) - F_Y(u)) = \mathbf{P}\left(X \leq u\right)\mathbf{P}\left(u < Y \leq v\right) = \mathbf{P}\left([X \leq u] \cap [u < Y \leq v]\right)$$

$$F_X(v)(F_Y(v) - F_Y(u)) = \mathbf{P}\left(X \leq v\right)\mathbf{P}\left(u < Y \leq v\right) = \mathbf{P}\left([X \leq v] \cap [u < Y \leq v]\right)$$

mais

$$[X \leq u] \cap [u < Y \leq v] \subset [X \leq Y] \cap [u < Y \leq v] \subset [X \leq v] \cap [u < Y \leq v]$$

donc

$$F_X(u)(F_Y(v) - F_Y(u)) \leq H(v) - H(u) \leq F_X(v)(F_Y(v) - F_Y(u))$$

et, en divisant par $v - u > 0$, il vient

$$\boxed{F_X(u)\frac{F_Y(v)-F_Y(u)}{v-u} \leq \frac{H(v)-H(u)}{v-u} \leq F_X(v)\frac{F_Y(v)-F_Y(u)}{v-u}.}$$

(b) Pour tout $x \geq 0$ et h tel que $x+h \in \mathbf{R}_+$, il suit de la question précédente que :

$$F_X(x)\frac{F_Y(x+h)-F_Y(x)}{h} \leq \frac{H(x+h)-H(x)}{h} \leq F_X(x+h)\frac{F_Y(x+h)-F_Y(x)}{h}$$

mais, puisque f_Y est continue, F_Y est $\mathcal{C}^1$ et donc

$$\lim_{h\to 0}\frac{F_Y(x+h)-F_Y(x)}{h} = F_Y'(x) = f_Y(x).$$

Par ailleurs, X est à densité donc F_X est continue et :

$$\lim_{h\to 0} F_X(x+h) = F_X(x)$$

de sorte que, d'après le théorème d'encadrement :

$$\lim_{h\to 0}\frac{H(x+h)-H(x)}{h} = F_X(x)f_Y(x).$$

En conclusion :

$$\boxed{H \text{ est dérivable sur } \mathbf{R}_+ \text{ et } H' = F_X f_Y.}$$

(c) $H' = F_X f_Y$ sur $\mathbf{R}_+$ donc H est une primitive de $F_X f_Y$ sur cet intervalle. Par ailleurs, on a observé en 1.b que $H(0) = 0$ de sorte que H est l'unique primitive de $F_X f_Y$ sur $\mathbf{R}_+$ qui s'annule en 0. D'après le théorème fondamental de l'analyse, il vient :

$$\boxed{\forall x \geq 0, \quad H(x) = \int_0^x F_X(t)f_Y(t)\ dt.}$$

3. D'après la question 1.b, on a $\mathbf{P}([X \leq Y]) = \lim\limits_{x\to+\infty} H(x)$ mais alors :

$$\lim_{x\to+\infty} H(x) = \lim_{x\to+\infty}\int_0^x F_X(t)f_Y(t)\ dt = \int_0^{+\infty} F_X(t)f_Y(t)\ dt.$$

On a donc :

$$\boxed{\mathbf{P}([X \leq Y]) = \int_0^{+\infty} F_X(t)f_Y(t)\ dt.}$$

4. On a $[X \leq Y] = [X < Y] \sqcup [X = Y]$ et, par incompatibilité, il vient $\mathbf{P}([X \leq Y]) = \mathbf{P}([X < Y]) + \mathbf{P}([X = Y])$. Or, d'après ce que nous avons établi en 3 et ce qui a été admis dans l'énoncé de cette question, on a $\mathbf{P}([X \leq Y]) = \mathbf{P}([X < Y])$ de sorte que

$$\boxed{\mathbf{P}([X = Y]) = 0.}$$

5. *Application aux lois exponentielles*

(a) Pour tout $x \in \mathbf{R}$, on a

$$F_X(x) = \mathbf{P}\left(U - \theta \leq x\right) = \mathbf{P}\left(U \leq x + \theta\right) = F_U(x+\theta) = \left\{ \begin{array}{ccc} 0 & \text{si} & x+\theta < 0 \\ 1 - e^{-\lambda(x+\theta)} & \text{si} & x + \theta \geq 0. \end{array} \right.$$

Ainsi,

$$\boxed{\forall x \in \mathbf{R}, \quad F_X(x) = \left\{ \begin{array}{ccc} 0 & \text{si} & x < -\theta \\ 1 - e^{-\lambda(x+\theta)} & \text{si} & x \geq -\theta. \end{array} \right.}$$

(b) D'après la question 3, on a

$$\begin{aligned} \mathbf{P}\left([U - \theta \leq V]\right) &= \int_0^{+\infty} F_X(t) f_Y(t)\ dt \\ &= \int_0^{+\infty} (1 - e^{-\lambda(\theta+t)}) \mu e^{-\mu t}\ dt \\ &= \int_0^{+\infty} \mu e^{-\mu t} - \mu e^{-\lambda\theta} e^{-(\lambda+\mu)t}\ dt \\ &= \int_0^{+\infty} \mu e^{-\mu t}\ dt - \mu e^{-\lambda\theta} \int_0^{+\infty} e^{-(\lambda+\mu)t}\ dt \\ &= 1 - \mu e^{-\lambda\theta} \times \frac{1}{\lambda+\mu}, \end{aligned}$$

la séparation des intégrales et la dernière égalité étant dues au fait que l'on a reconnu des densités de lois exponentielles de paramètres respectifs μ et $\lambda + \mu$. En conclusion :

$$\boxed{\mathbf{P}\left([U - \theta \leq V]\right) = 1 - \frac{\mu}{\lambda+\mu} e^{-\lambda\theta}.}$$

Inégalité de Boole

6. (a) Montrons par récurrence sur $n \in \mathbf{N}^*$ que $\mathbf{P}\left(\bigcup_{k=1}^{n} B_k\right) \leq \sum_{k=1}^{n} \mathbf{P}\left(B_k\right)$.

Initialisation : Pour $n = 1$, on a bien $\mathbf{P}\left(B_1\right) \leq \mathbf{P}\left(B_1\right)$.

Hérédité : Soit $n \in \mathbf{N}^*$ tel que $\mathbf{P}\left(\bigcup_{k=1}^{n} B_k\right) \leq \sum_{k=1}^{n} \mathbf{P}\left(B_k\right)$. Alors

$$\begin{aligned} \mathbf{P}\left(\bigcup_{k=1}^{n+1} B_k\right) &= \mathbf{P}\left(\left(\bigcup_{k=1}^{n} B_k\right) \cup B_{n+1}\right) \\ &= \mathbf{P}\left(\bigcup_{k=1}^{n} B_k\right) + \mathbf{P}\left(B_{n+1}\right) - \underbrace{\mathbf{P}\left(\left(\bigcup_{k=1}^{n} B_k\right) \cap B_{n+1}\right)}_{\geq 0} \\ &\leq \mathbf{P}\left(\bigcup_{k=1}^{n} B_k\right) + \mathbf{P}\left(B_{n+1}\right) \\ &\underset{\text{H.R.}}{\leq} \left(\sum_{k=1}^{n} \mathbf{P}\left(B_k\right)\right) + \mathbf{P}\left(B_{n+1}\right) \\ &= \sum_{k=1}^{n+1} \mathbf{P}\left(B_k\right). \end{aligned}$$

Ainsi :

$$\forall n \in \mathbf{N}^*, \quad \mathbf{P}\left(\bigcup_{k=1}^{n} B_k\right) \leq \sum_{k=1}^{n} \mathbf{P}\left(B_k\right).$$

(b) Notons, pour tout $n \in \mathbf{N}^*$, $A_n = \bigcup_{k=1}^{n} B_k$, alors $(A_n)_{n\in\mathbf{N}^*}$ est une suite croissante d'événements donc, d'après le théorème de la limite croissante :

$$\mathbf{P}\left(\bigcup_{k\geq 1} B_k\right) = \mathbf{P}\left(\bigcup_{n\in\mathbf{N}^*} A_n\right) = \lim_{n\to+\infty} \mathbf{P}\left(A_n\right).$$

Or, d'après 6.a,

$$\mathbf{P}\left(A_n\right) = \mathbf{P}\left(\bigcup_{k=1}^{n} B_k\right) \leq \sum_{k=1}^{n} \mathbf{P}\left(B_k\right)$$

et donc, en passant à la limite,

$$\lim_{n\to+\infty} \mathbf{P}\left(A_n\right) \leq \sum_{k=1}^{+\infty} \mathbf{P}\left(B_k\right).$$

On a donc établi l'*inégalité de Boole* :

$$\mathbf{P}\left(\bigcup_{k\geq 1} B_k\right) \leq \sum_{k=1}^{+\infty} \mathbf{P}\left(B_k\right).$$

Intermède culturel : Georges Boole (1815-1864) est un mathématicien et logicien britannique qui a participé à la fondation de la logique moderne. On lui doit notamment la découverte de l'*algèbre de Boole*, particulièrement utile en informatique pour formaliser les tests conditionnels. C'est à lui que fait référence le qualificatif *booléen* utilisé pour les variables `%T` et `%F` en Scilab.

Partie II – Une compétition entre deux groupes

7. (a) La variable aléatoire $\sum_{k=1}^{n} X_k$ représente le temps mis par le groupe A pour résoudre n problèmes.

(b) Avec $X_1 = 5$, $X_2 = 2$, $X_3 = 3$, $X_4 = 2$, $Y_1 = 2$, $Y_2 = 2$, $Y_3 = 4$, $Y_4 = 2$, la liste des problèmes résolus par les deux groupes avec leur temps de résolution est :

Groupe A			**Groupe B**		
Problème	**Temps**	**Total**	**Problème**	**Temps**	**Total**
P_1	5	5	Q_1	2	2
P_2	2	7	Q_2	2	4
P_3	3	10	Q_3	4	8
P_4	2	12	Q_4	2	10

Ainsi, la liste des problèmes dans l'ordre de résolution par les deux groupes est :

$$(Q_1, Q_2, P_1, P_2, Q_3, P_3, Q_4, \ldots)$$

On obtient donc les valeurs suivantes :

U_1	U_2	U_3	U_4	U_5	U_6	U_7
0	0	1	1	0	1	0

Pour U_8, on ne peut pas conclure car on ne sait pas quel problème a été résolu en huitième position. En effet, si $Y_5 = 1$, alors ce huitième problème sera Q_5 et aura été résolu par le groupe B, mais si $Y_5 \geq 2$, alors ce huitième problème sera P_4 et aura été résolu par le groupe A.

(c) On complète le script Scilab de la manière suivante :

```
p = input('p=')
n = input('n=')
q = 1-p

U = zeros(1,n)

sommeX = grand(1,1,'exp',1/p)
sommeY = grand(1,1,'exp',1/q)

mini = min(sommeX,sommeY)

for k=1:n
    if sommeX == mini
        U(k) = 1
        sommeX = sommeX + grand(1,1,'exp',1/p)
    else
        sommeY = sommeY + grand(1,1,'exp',1/q)
    end
    mini = min(sommeX,sommeY)
end
disp(U,'U=')
```

(d) Pour afficher la valeur de S_n, il suffit d'ajouter en fin de script les lignes suivantes :

```
S = sum(U)
disp(S,'S=')
```

8. *Loi de U_n*

(a) On a $[U_1 = 1] = [X_1 \leq Y_1]$ car cela signifie que le premier problème a été résolu par le groupe A, et donc que celui-ci l'a résolu plus vite que le groupe B. En particulier, on a :

$$\boxed{\mathbf{P}([U_1 = 1]) = \mathbf{P}([X_1 \leq Y_1]).}$$

(b) i. Soit $x < 0$. Alors $[Y_1 - X_1 \leq x] = [Y_1 \leq X_1 + x] \subset [Y_1 < X_1]$ et donc

$$[Y_1 - X_1 \leq x] \cap [U_1 = 1] \subset [Y_1 < X_1] \cap [X_1 \leq Y_1] = \emptyset$$

et alors

$$\mathbf{P}_{[U_1=1]}([Y_1 - X_1 \leq x]) = \frac{\mathbf{P}([Y_1 - X_1 \leq x] \cap [U_1 = 1])}{\mathbf{P}([U_1 = 1])} = 0.$$

Ainsi,

$$\boxed{\forall x < 0, \quad \mathbf{P}_{[U_1=1]}([Y_1 - X_1 \leq x]) = 0.}$$

ii. Soit $x \geq 0$. Alors :

$$\begin{aligned}\mathbf{P}_{[U_1=1]}([Y_1 - X_1 \leq x]) &= \frac{\mathbf{P}([Y_1 - X_1 \leq x] \cap [U_1 = 1])}{\mathbf{P}([U_1 = 1])} \\ &= \frac{1}{p}\mathbf{P}([Y_1 \leq X_1 + x] \cap [X_1 \leq Y_1])\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{p}\mathbf{P}\left(X_1 \leq Y_1 \leq X_1 + x\right) \\
&= \frac{1}{p}\left(\mathbf{P}\left(Y_1 \leq X_1 + x\right) - \mathbf{P}\left(Y_1 \leq X_1\right)\right) \\
&\underset{\text{5.b}}{=} \frac{1}{p}\left[\left(1 - \frac{p}{q+p}e^{-qx}\right) - \left(1 - \frac{p}{q+p}\right)\right] \\
&= \frac{1}{p}\left[-pe^{-qx} + p\right] \\
&= 1 - e^{-qx}.
\end{aligned}
$$

(c) On reconnaît la fonction de répartition d'une loi $\mathcal{E}(q)$ donc :

> la loi conditionnelle de $Y_1 - X_1$ sachant $[U_1 = 1]$ est une loi $\mathcal{E}(q)$.

En intervertissant les rôles des groupes A et B, on obtiendrait de même :

> la loi conditionnelle de $X_1 - Y_1$ sachant $[U_1 = 0]$ est une loi $\mathcal{E}(p)$.

(d) Supposons que $[U_1 = 1]$ est réalisé. Alors $Y_1' = Y_1 - X_1$ est le temps de résolution mis par le groupe B pour son premier problème à partir du moment où le premier problème a été réalisé par A. Puisque $Y_1' \rightsquigarrow \mathcal{E}(q)$ d'après la question précédente, et puisque Y_1' est indépendante de $Y_2, Y_3, \ldots$, et de $X_2, X_3, \ldots$, d'après le lemme des coalitions, on se retrouve dans le même modèle que précédemment, avec un problème en moins à résoudre. Ainsi, si on note U_i' la variable aléatoire définie à partir de ce processus de la même manière que U_i dans le modèle précédent, U_{i+1} suit la même loi que U_i'. Or il suit de l'hypothèse de récurrence que $\mathbf{P}\left(U_n' = 1\right) = p$ et donc, $\mathbf{P}_{[U_1=1]}(U_{n+1} = 1) = p$.
De même, si $[U_1 = 0]$ est réalisé, on pose $X_1' = X_1 - Y_1$ qui suit une loi $\mathcal{E}(p)$ d'après la question précédente. Et en étudiant le processus associé à $X_1', X_2, \ldots$ et $Y_2, Y_3, \ldots$, on obtient de même $\mathbf{P}_{[U_1=0]}(U_{n+1} = 1) = p$.
En conclusion :

> $\forall n \in \mathbf{N}^*, \quad \left(\mathbf{P}\left(U_n = 1\right) = p \quad \Rightarrow \quad \mathbf{P}_{[U_1=0]}([U_{n+1} = 1]) = \mathbf{P}_{[U_1=1]}([U_{n+1} = 1]) = p\right)$

Remarque : Dit autrement, ce que nous avons observé dans cette question est une *absence de mémoire* du processus étudié, laquelle est caractéristique des lois exponentielles suivies par les X_i et les Y_i. Cette observation sera d'un grand secours dans les questions suivantes.

(e) On sait déjà que $\mathbf{P}\left([U_1 = 1]\right) = p$. Par ailleurs, si $n \in \mathbf{N}^*$ est tel que $\mathbf{P}\left([U_n = 1]\right) = p$, il suit de la question précédente et de la formule des probabilités totales que :

$$
\begin{aligned}
\mathbf{P}\left([U_{n+1} = 1]\right) &= \mathbf{P}_{[U_n=0]}([U_{n+1} = 1])\mathbf{P}\left([U_n = 0]\right) + \mathbf{P}_{[U_n=1]}(U_{n+1} = 1)\mathbf{P}\left([U_n = 1]\right) \\
&\underset{\text{8.d}}{=} p\mathbf{P}\left([U_n = 0]\right) + p\mathbf{P}\left([U_n = 1]\right) \\
&= p.
\end{aligned}
$$

et donc la propriété est héréditaire.
Ainsi, par récurrence sur n, on a montré que :

> $\forall n \in \mathbf{N}^*, \quad \mathbf{P}\left([U_n = 1]\right) = p.$

9. Soit $n \in \mathbf{N}^*$. On observe que $S_n = \sum_{k=1}^{n} U_k$. Mais alors, les variables aléatoires $U_1, \ldots, U_n$ étant de loi $\mathcal{B}(p)$ d'après la question 8 et mutuellement indépendantes d'après l'énoncé, on sait que :

> $S_n \rightsquigarrow \mathcal{B}(n, p).$

10. (a) Pour $n = 0$, 0 problème sont résolus et A en a donc résolu 0 de plus que B. Autrement dit, l'événement A_0 est systématiquement réalisé. Il s'ensuit que

$$\boxed{a_0 = 1.}$$

(b) Si $[U_1 = 1]$ est réalisé, A a déjà résolu un problème et a donc un problème d'avance sur B. Ainsi, du fait de l'absence de mémoire du processus observée à la question 8.d, si $[U_1 = 1]$ est réalisé, la probabilité que A_r se réalise est la même que celle que A_{r-1} se réalise dans le processus général et donc $\mathbf{P}_{[U_1=1]}(A_r) = A_{r-1}$. De la même manière, si $[U_1 = 0]$ est réalisé, A a un problème de retard et, du fait de l'absence de mémoire, la probabilité que A_r se réalise est la même que celle que A_{r+1} se réalise dans le processus général. Ainsi, $\mathbf{P}_{[U_1=0]}(A_r) = A_{r+1}$.
On a donc montré :

$$\boxed{\forall r \geq 1, \quad \mathbf{P}_{[U_1=1]}(A_r) = A_{r-1} \quad \text{et} \quad \mathbf{P}_{[U_1=0]}(A_r) = A_{r+1}.}$$

(c) Soit $r \geq 1$. Alors, d'après la formule des probabilités totales, on a

$$\begin{aligned} a_r &= \mathbf{P}(A_r) \\ &= \mathbf{P}_{[U_1=0]}(A_r)\mathbf{P}([U_1 = 0]) + \mathbf{P}_{[U_1=1]}(A_r)\mathbf{P}([U_1 = 1]) \\ &= \mathbf{P}(A_{r+1})\, q + \mathbf{P}(A_{r-1})\, p \end{aligned}$$

et donc, en divisant par q :

$$\boxed{a_{r+1} = \frac{1}{q}a_r - \frac{p}{q}a_{r-1}.}$$

(d) Posons $P = X^2 - \frac{1}{q}X + \frac{p}{q}$ le polynôme caractéristique associé à la suite $(a_r)_{r\in\mathbf{N}}$. Il admet les mêmes racines que $Q = qX^2 - X + p$ donc le discriminant est $\Delta = 1 - 4pq = (1-2p)^2 \geq 0$. Ces racines sont donc

$$r_1 = \frac{1-(1-2p)}{2q} = \frac{p}{q} \quad \text{et} \quad r_2 = \frac{1+(1-2p)}{2q} = \frac{2(1-p)}{q} = 1.$$

Ainsi, il existe des réels λ et μ tels que :

$$\boxed{\forall r \in \mathbf{N}, \quad a_r = \lambda + \mu\left(\frac{p}{q}\right)^r.}$$

11. Si $p = \dfrac{1}{2}$, alors $q = \dfrac{1}{2}$ et donc $\dfrac{p}{q} = 1$. Il s'ensuit que, pour tout $r \in \mathbf{N}$, on a $a_r = \lambda + \mu$ de sorte que la suite (a_r) est constante. Puisque $a_0 = 1$, il s'ensuit que, pour tout $r \in \mathbf{N}$, $a_r = 1$.
Supposons $p > \dfrac{1}{2}$ de sorte que $\dfrac{p}{q} > 1$. Si $\mu \neq 0$, on a

$$|a_r| \underset{r\to+\infty}{\sim} |\mu|\left(\frac{p}{q}\right)^r \xrightarrow[r\to+\infty]{} +\infty,$$

une contradiction puisque chaque a_r est une probabilité. Ainsi, $\mu = 0$ et donc, pour tout $r \in \mathbf{N}$, $a_r = \lambda$. La suite (a_r) est donc constante égale à $a_0 = 1$.
Ainsi,

$$\boxed{\text{Si } p \geq \frac{1}{2}, \text{ alors } (a_r)_{r\in\mathbf{N}} \text{ est constante égale à 1.}}$$

12. Le cas $p < \dfrac{1}{2}$.

(a) Soit k un entier naturel.

i. Si A_{2k} est réalisé, alors il existe un entier $n \geq 2k$ et des entiers a, b tels que A a résolu a problèmes, B a résolu b problèmes avec $a + b = n$ et $a = b + 2k$. Mais alors $n = a + b = 2b + 2k$ est nécessairement pair et peut donc s'écrire $2i$, avec $i \geq k$.

On a alors :

$$A_{2k} = \bigcup_{2i \geq 2k} [S_{2i} = (2i - S_{2_i}) + 2k] = \bigcup_{i \geq k} [2S_{2i} = 2i + 2k] = \bigcup_{i \geq k} [S_{2i} = i + k].$$

Ainsi :

$$\boxed{A_{2k} = \bigcup_{i \geq k} [S_{2i} = i + k].}$$

ii. On a observé à la question 9 que $S_{2i} \rightsquigarrow \mathcal{B}(2i, p)$ donc, pour tout $k \leq i$, on a $i + k \in [\![0, 2i]\!]$ et donc :

$$\mathbf{P}([S_{2i} = i + k]) = \binom{2i}{i+k} p^{i+k} q^{2i-(i+k)} = \binom{2i}{i+k} p^{i+k} q^{i-k}.$$

On a donc bien :

$$\boxed{\forall i \geq k, \quad \mathbf{P}([S_{2i} = i + k]) = \binom{2i}{i+k} p^{i+k} q^{i-k}.}$$

iii. D'après la formule du binôme de Newton, on a :

$$\sum_{j=0}^{2i} \binom{2i}{j} = \sum_{j=0}^{2i} \binom{2i}{j} 1^j 1^{2i-j} = (1+1)^{2i} = 2^{2i} = 4^i.$$

Soit alors $i \geq k$. Puisque $i + k \in [\![0, 2i]\!]$, la positivité des termes dans la somme suivante implique :

$$\boxed{\binom{2i}{i+k} \leq \sum_{j=0}^{2i} \binom{2i}{j} \leq 4^i.}$$

iv. Soit $i \geq k$. On a

$$\mathbf{P}([S_{2i} = k + i]) = \binom{2i}{i+k} p^{i+k} q^{i-k} \leq 4^i p^{i+k} q^{i-k} = (4pq)^i \left(\frac{p}{q}\right)^k.$$

Or $p < \dfrac{1}{2}$ donc $|pq| \leq \dfrac{1}{4} < 1$ et alors :

$$\sum_{i \geq k} (4pq)^i \left(\frac{p}{q}\right)^k = \left(\frac{p}{q}\right)^k \sum_{i \geq k} (4pq)^i$$

est une série convergente et

$$\begin{aligned} \sum_{i=k}^{+\infty} (4pq)^i \left(\frac{p}{q}\right)^k &= \left(\frac{p}{q}\right)^k \sum_{i=k}^{+\infty} (4pq)^i \\ &= \left(\frac{p}{q}\right)^k (4pq)^k \sum_{i=0}^{+\infty} (4pq)^i \end{aligned}$$

$$= \left(\frac{p}{q}\right)^k (4pq)^k \frac{1}{1-4pq}$$

et, par comparaison :

$$\boxed{\sum_{i=k}^{+\infty} \mathbf{P}\left([S_{2i} = k+i]\right) \leq \left(\frac{p}{q}\right)^k (4pq)^k \frac{1}{1-4pq}.}$$

(b) Soit $k \in \mathbf{N}$. Alors d'après l'inégalité de Boole :

$$0 \leq a_{2k} = \mathbf{P}\left(\bigcup_{i \geq k} [S_{2i} = i+k]\right) \leq \sum_{i=k}^{+\infty} \mathbf{P}\left([S_{2i} = k+i]\right) \leq \left(\frac{p}{q}\right)^k (4pq)^k \frac{1}{1-4pq}$$

mais $p < \dfrac{1}{2}$ donc $|4pq[< 1$ et $\left|\dfrac{p}{q}\right| < 1$ et ainsi $\left(\dfrac{p}{q}\right)^k (4pq)^k \xrightarrow[k\to+\infty]{} 0$. Par encadrement, il vient

$$\boxed{\lim_{k\to+\infty} a_{2k} = 0.}$$

(c) On a établi à la question 10.d qu'il existe des réels λ et μ tels que, pour tout $r \in \mathbf{N}$,

$$a_r = \lambda + \mu \left(\frac{p}{q}\right)^r.$$

En particulier, pour tout $k \in \mathbf{N}$,

$$a_{2k} = \lambda + \mu \left(\frac{p}{q}\right)^{2k}.$$

Si $p < \dfrac{1}{2}$, on a $\left(\dfrac{p}{q}\right)^{2k} \xrightarrow[k\to+\infty]{} 0$ et donc $\lim\limits_{k\to+\infty} a_{2k} = \lambda$. Par unicité de la limite, il suit de la question 12.b que $\lambda = 0$. On a donc :

$$\forall r \in \mathbf{N}, \quad a_r = \lambda \left(\frac{p}{q}\right)^r$$

et puisque $a_0 = 1$, il vient $\lambda = 1$. On a donc

$$\boxed{\forall r \in \mathbf{N}, \quad a_r = \left(\frac{p}{q}\right)^r.}$$

Partie III – La *blockchain* et la stratégie de la *double dépense*

13. (a) Soit $k \in \mathbf{N}$. $[T_n = k]$ signifie que quand Q_n est ajouté à la liste des problèmes résolus, alors A a résolu k problèmes. Ceci signifie donc que A avait déjà résolu k problèmes à l'étape précédente (qui est la $(n+k-1)$-ième) et que le dernier problème résolu l'a été par B. Autrement dit :

$$\boxed{[T_n = k] = (S_{n+k-1} = k) \cap (U_{n+k} = 0).}$$

(b) D'après le lemme des coalitions, $S_{n+k-1} = \sum\limits_{i=1}^{n+k-1} U_i$ est indépendante de U_{n+k} et donc :

$$\begin{aligned}\mathbf{P}\left([T_n = k]\right) &= \mathbf{P}\left([S_{n+k-1} = k] \cap [U_{n+k} = 0]\right) \\ &= \mathbf{P}\left([S_{n+k-1} = k]\right) \mathbf{P}\left([U_{n+k} = 0]\right)\end{aligned}$$

$$= \binom{n+k-1}{k} p^k q^{n-1} \times q \quad \text{(d'après 8 et 12.a.ii)}$$
$$= \binom{n+k-1}{k} p^k q^n.$$

Ainsi :

$$\boxed{\forall k \in \mathbf{N}, \quad \mathbf{P}\left([T_n = k]\right) = \binom{n+k-1}{k} p^k q^n.}$$

14. (a) En appliquant la formule des probabilités totales avec le système complet d'événements $\{[T_n = k]\}_{k\in\mathbf{N}}$, on a :

$$\mathbf{P}\left(G_n\right) = \sum_{k=0}^{+\infty} \mathbf{P}\left(G_n \cap [T_n = k]\right).$$

Pour tout $k \geq n+1$, si $[T_n = k]$ est réalisé, alors G_n est automatiquement réalisé (il suffit de prendre $t' = t$), ce qui signifie que $[T_n = k] \subset G_n$ et donc $\mathbf{P}\left(G_n \cap [T_n = k]\right) = \mathbf{P}\left([T_n = k]\right)$. On a donc

$$\mathbf{P}\left(G_n\right) = \sum_{k=0}^{n} \mathbf{P}\left(G_n \cap [T_n = k]\right) + \sum_{k=n+1}^{+\infty} \mathbf{P}\left([T_n = k]\right) = \sum_{k=0}^{n} \mathbf{P}\left(G_n \cap [T_n = k]\right) + \mathbf{P}\left([T_n \geq n+1]\right).$$

Par ailleurs, si $k \in [\![0, n]\!]$, on a

$$\mathbf{P}\left(G_n \cap [T_n = k]\right) = \mathbf{P}_{[T_n=k]}(G_n)\mathbf{P}\left([T_n = k]\right)$$

mais $\mathbf{P}_{[T_n=k]}(G_n) = \mathbf{P}\left(A_{n+1-k}\right)$ car, sachant qu'au temps t, le groupe A a résolu k problèmes et B en a résolu n, il faut encore que A résolve $n+1-k$ problèmes de plus que B par la suite, et l'absence de mémoire du processus (voir 8.d) indique cette probabilité est égale à $\mathbf{P}\left(A_{n+1-k}\right)$. Alors :

$$\boxed{\mathbf{P}\left(G_n\right) = \mathbf{P}\left([T_n \geq n+1]\right) + \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right)\mathbf{P}\left(A_{n+1-k}\right).}$$

(b) Si $p \geq \dfrac{1}{2}$, la suite $(a_r)_{r\in\mathbf{N}}$ étant constante égale à 1, on a

$$\mathbf{P}\left(G_n\right) = \mathbf{P}\left([T_n \geq n+1]\right) + \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right) = \sum_{k=0}^{+\infty} \mathbf{P}\left([T_n = k]\right) = 1$$

de sorte que

$$\boxed{p \geq \frac{1}{2} \quad \Rightarrow \quad \mathbf{P}\left(G_n\right) = 1.}$$

(c) Si $p < \dfrac{1}{2}$, on a :

$$\begin{aligned}
\mathbf{P}\left(G_n\right) &= \mathbf{P}\left(T_n \geq n+1\right) + \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right)\mathbf{P}\left(A_{n+1-k}\right) \\
&\underset{12}{=} \mathbf{P}\left(T_n \geq n+1\right) + \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right)\left(\frac{p}{q}\right)^{n+1-k} \\
&= \left[1 - \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right)\right] + \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right)\left(\frac{p}{q}\right)^{n+1-k}
\end{aligned}$$

$$
\begin{aligned}
&= 1 - \sum_{k=0}^{n} \mathbf{P}\left([T_n = k]\right) \left[1 - \left(\frac{p}{q}\right)^{n+1-k}\right] \\
&\underset{\text{13.b}}{=} 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} p^k q^n \frac{q^{n+1-k} - p^{n+1-k}}{q^{n+1-k}} \\
&= 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} \left(p^k q^n - \frac{p^{n+1}}{q^{1-k}}\right) \\
&= 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} (p^k q^n - p^{n+1} q^{k-1})
\end{aligned}
$$

On a donc :

$$\boxed{p < \frac{1}{2} \quad \Rightarrow \quad \mathbf{P}\left(G_n\right) = 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} (p^k q^n - p^{n+1} q^{k-1}).}$$

15. *Une meilleure expression de* $\mathbf{P}\left(G_n\right)$ *lorsque* $p < \frac{1}{2}$.

(a) Soit $n \in \mathbf{N}^*$. On a :

$$
\begin{aligned}
1 - u_n(p) + \frac{p}{q} u_n(q) &= 1 - (1-p)^n \sum_{k=0}^{n} \binom{n+k-1}{k} p^k + \frac{p}{q}(1-q)^n \sum_{k=0}^{n} \binom{n+k-1}{k} q^k \\
&= 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} \left[q^n p^k - \frac{p}{q} p^n q^k\right] \\
&= 1 - \sum_{k=0}^{n} \binom{n+k-1}{k} [p^k q^n - p^{n+1} q^{k-1}] = \mathbf{P}\left(G_n\right).
\end{aligned}
$$

On a donc bien :

$$\boxed{\forall n \in \mathbf{N}^*, \quad \mathbf{P}\left(G_n\right) = 1 - u_n(p) + \frac{p}{q} u_n(q).}$$

(b) En un mot, la réponse à cette question est un long calcul utilisant à plusieurs reprises le triangle de Pascal, la symétrie du coefficient binomial, des changements d'indices, quelques factorisations et développements bien choisis et un télescopage. Si on s'y prend méthodiquement, il est plus long que difficile. Mais la relative longueur laisse à penser que l'on a manqué un raccourci quelque part...

$$
\begin{aligned}
u_{n+1}(x) &= (1-x)^{n+1} \sum_{k=0}^{n+1} \binom{n+k}{k} x^k \\
&= \left[(1-x)^{n+1} \sum_{k=0}^{n} \binom{n+k}{k} x^k\right] + (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1} \\
&= (1-x) \left[(1-x)^n \sum_{k=0}^{n} \left(\binom{n+k-1}{k} + \binom{n+k-1}{k-1}\right) x^k\right] + (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1} \\
&= (1-x) \left[(1-x)^n \left(\sum_{k=0}^{n} \binom{n+k-1}{k} x^k + \sum_{k=0}^{n} \binom{n+k-1}{k-1} x^k\right)\right] + (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1} \\
&= (1-x) \left[u_n(x) + (1-x)^n \sum_{k=0}^{n} \binom{n+k-1}{k-1} x^k\right] + (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1}
\end{aligned}
$$

$$
\begin{aligned}
&= (1-x)\left[u_n(x) + (1-x)^n \sum_{k=0}^{n-1} \binom{n+k}{k} x^{k+1}\right] + (1-x)^{n+1}\binom{2n+1}{n+1}x^{n+1} \\
&= u_n(x) - x u_n(x) + (1-x)^n \sum_{k=0}^{n-1} \binom{n+k}{k} x^{k+1} - \sum_{k=0}^{n-1} \binom{n+k}{k} x^{k+2} + (1-x)^{n+1}\binom{2n+1}{n+1}x^{n+1} \\
&= u_n(x) + (1-x)^n \left[-\sum_{k=0}^{n} \binom{n+k-1}{k} x^{k+1} + \sum_{k=0}^{n-1} \binom{n+k}{k} x^{k+1} - \sum_{k=0}^{n-1} \binom{n+k}{k} x^{k+2}\right. \\
&\qquad \left. + \binom{2n+1}{n+1}x^{n+1} - \binom{2n+1}{n+1}x^{n+2}\right] \\
&= u_n(x) + (1-x)^n \left[\sum_{k=0}^{n-1} \left(\binom{n+k}{k} - \binom{n+k-1}{k}\right) x^{k+1} - \binom{2n-1}{n} x^{n+1} - \sum_{k=0}^{n-1} \binom{n+k}{k} x^{k+2}\right. \\
&\qquad \left. + \binom{2n+1}{n+1}x^{n+1} - \binom{2n+1}{n+1}x^{n+2}\right] \\
&= u_n(x) + (1-x)^n \left[-\binom{2n-1}{n-1}x^{n+1} + \binom{2n+1}{n+1}x^{n+1} - \binom{2n-1}{n}x^{n+1} - \binom{2n+1}{n+1}x^{n+2}\right] \\
&= u_n(x) + (1-x)^n x^{n+1}\left[\binom{2n+1}{n+1} - \left(\binom{2n-1}{n-1} + \binom{2n-1}{n}\right) - \binom{2n+1}{n+1}x\right] \\
&= u_n(x) + (1-x)^n x^{n+1}\left[\binom{2n+1}{n} - \binom{2n}{n} - \binom{2n+1}{n+1}x\right] \\
&= u_n(x) + (1-x)^n x^{n+1}\left[\binom{2n}{n-1} - \binom{2n+1}{n+1}x\right] \\
&= u_n(x) + (1-x)^n x^{n+1}\left[\binom{2n}{n+1} - \binom{2n+1}{n+1}x\right].
\end{aligned}
$$

On a donc bien :

$$\boxed{u_{n+1}(x) = u_n(x) + (1-x)^n x^{n+1}\left[\binom{2n}{n+1} - \binom{2n+1}{n+1}x\right].}$$

(c) Soit $n \in \mathbf{N}^*$. Alors :

$$
\begin{aligned}
\mathbf{P}(G_{n+1}) &= 1 - u_{n+1}(p) + \frac{p}{q}u_{n+1}(q) \\
&= 1 - \left[u_n(p) + (1-p)^n p^{n+1}\left(\binom{2n}{n+1} - \binom{2n+1}{n+1}p\right)\right] + \frac{p}{q}(1-q)^n q^{n+1}\left(\binom{2n}{n+1} - \binom{2n+1}{n+1}q\right) \\
&= 1 - u_n(p) - \frac{p}{q}u_n(q) - q^n p^{n+1}\left(\binom{2n}{n+1} - \binom{2n+1}{n+1}p\right) + \frac{p}{q}p^n q^{n+1}\left(\binom{2n}{n+1} - \binom{2n+1}{n+1}q\right) \\
&= \mathbf{P}(G_n) - \binom{2n}{n+1}q^n p^{n+1} + \binom{2n+1}{n+1}q^n p^{n+2} + \binom{2n}{n+1}p^{n+1}q^n - \binom{2n+1}{n+1}p^{n+1}q^{n+1} \\
&= \mathbf{P}(G_n) + p^{n+1}q^{n+1}\left(\frac{p}{q} - 1\right)\binom{2n+1}{n+1}
\end{aligned}
$$

et donc

$$\boxed{\forall n \in \mathbf{N}^*, \quad \mathbf{P}(G_{n+1}) = \mathbf{P}(G_n) - (pq)^{n+1}\left(1 - \frac{p}{q}\right)\binom{2n+1}{n+1}}$$

(d) **Initialisation :** Pour $n = 1$, on a :

$$\frac{p}{q} - \left(1 - \frac{p}{q}\right)\binom{1}{1}(pq) = \frac{p}{q} - pq + p^2 = \frac{p - pq^2 + p^2 q}{q}$$

et, par ailleurs, il suit de 14.c que

$$\begin{aligned}\mathbf{P}(G_1) &= 1 - \left(q - \frac{p^2}{q}\right) - (pq - p^2) \\ &= \frac{q - q^2 + p^2 - pq^2 + p^2 q}{q} \\ &= \frac{q(1-q) + p^2 + p^2 q - pq^2}{q} \\ &= \frac{p(q+p) + p^2 q - pq^2}{q} \\ &= \frac{p + p^2 q - pq^2}{q}\end{aligned}$$

et donc la propriété est vraie pour $n = 1$.

Hérédité : Soit $n \in \mathbf{N}^*$ tel que $\mathbf{P}(G_n) = \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k$. Alors

$$\begin{aligned}\mathbf{P}(G_{n+1}) &= \mathbf{P}(G_n) - (pq)^{n+1} \left(1 - \frac{p}{q}\right) \binom{2n+1}{n+1} \\ &\underset{\text{H.R.}}{=} \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k - (pq)^{n+1} \left(1 - \frac{p}{q}\right) \binom{2n+1}{n+1} \\ &= \frac{p}{q} - \left(1 - \frac{p}{q}\right) \left(\left[\sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k \right] + (pq)^{n+1} \binom{2n+1}{n+1} \right) \\ &= \frac{p}{q} - \left(1 - \frac{p}{q}\right) \left(\sum_{k=1}^{n+1} \binom{2k-1}{k} (pq)^k \right).\end{aligned}$$

Ainsi, on a montré que :

$$\forall n \in \mathbf{N}^*, \quad \mathbf{P}(G_n) = \frac{p}{q} - \left(1 - \frac{p}{q}\right) \sum_{k=1}^{n} \binom{2k-1}{k} (pq)^k.$$

16. *Application à la sécurisation des transactions*

(a) Soit $n \in \mathbf{N}^*$ et $k \in [\![1, n]\!]$, alors :

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)!}{k(k-1)!((n-1)-(k-1))!} = \frac{n}{k} \binom{n-1}{k-1}$$

donc nous proposons le script récursif suivant :

```
function c=binom(n,k)
    if k==0 then
        c = 1
    else
        c = n/k*binom(n-1,k-1)
    end
endfunction
```

(b) Pour déterminer n_p, le plus petit entier n tel que $\mathbf{P}(G_n) \leq \epsilon$ pour $p < \frac{1}{2}$ et $\epsilon > 0$ saisis au clavier par l'utilisateur, nous proposons d'écrire le script suivant à la suite de la fonction précédente :

```
function r = PG(n,p)
    q = 1-p
    S = 0
    for k=1:n
        S = S+binom(2*k-1,k)*(p*q)^k
    end
    r = p/q-(1-p/q)*S
endfunction

function r = np(p,epsilon)
    if (p>=1/2) then
        r = 'Erreur : Proba p >= 1/2'
    else
        n = 1
        while (PG(n,p)>epsilon)
            n=n+1
        end
        r = n
    end
endfunction
```

Remarque : Pour obtenir la figure proposée dans l'énoncé, nous avons complété notre script avec les lignes suivantes :

```
P = .1:.01:.32
Y = zeros(1,23)
for k=1:23
    Y(k) = np(P(k),.0001)
end
plot(P,Y)
```

□