MATHÉMATIQUES I - HEC-ESSEC E 2020

Proposition de corrigé par David Meneu

Lycée Champollion - Grenoble, pour Major-Prépa

Ce sujet s'intéressait au problème de la *double dépense de bitcoins* par un groupe d'individus mal intentionnés.

Partie I - Deux résultats généraux

Calcul d'une probabilité

Soient 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 $x \in \mathbb{R}^+$, on pose $H(x) = \mathbb{P}([X \leqslant Y] \cap [Y \leqslant x])$.

1. a) Pour tous réels x et y de $\mathbb{R}^+$ tels que $x \leqslant y$:

   $[Y \leqslant x]$ implique $[Y \leqslant y]$, ce qui en termes d'événements s'écrit $[Y \leqslant x] \subset [Y \leqslant y]$,

   et donc $[X \leqslant Y] \cap [Y \leqslant x] \subset [X \leqslant Y] \cap [Y \leqslant y]$.

   La propriété de croissance de la probabilité assure alors que :

   $\mathbb{P}([X \leqslant Y] \cap [Y \leqslant x]) \leqslant \mathbb{P}([X \leqslant Y] \cap [Y \leqslant y])$, et prouve ainsi que :

$$\forall (x,y) \in (\mathbb{R}_+)^2, \quad x \leqslant y \Longrightarrow H(x) \leqslant H(y)$$

   et donc par définition, que H est une fonction croissante sur $\mathbb{R}^+$.

   Comme de plus $H(x)$ est par définition une probabilité, on a : $\quad \forall x \in \mathbb{R}_+,\ H(x) \leqslant 1$.

   La fonction H étant croissante et majorée par 1 sur $\mathbb{R}_+$, elle admet donc une limite finie en $+\infty$, d'après le théorème de limite monotone pour les fonctions.

   b) La fonction H étant croissante, on a : $\lim\limits_{x\to+\infty} H(x) = \lim\limits_{n\to+\infty} H(n)$ où n est une variable entière.

   Comme vu précédemment : $\quad \forall n \in \mathbb{N},\ [X \leqslant Y] \cap [Y \leqslant n] \subset [X \leqslant Y] \cap [Y \leqslant n+1]$, ce qui signifie que $\big([X \leqslant Y] \cap [Y \leqslant n]\big)_{n\in\mathbb{N}}$ est une suite croissante d'événements.

   La propriété de limite monotone pour les probabilités donne donc :

$$\lim_{n\to+\infty} H(n) = \lim_{n\to+\infty} \mathbb{P}([X \leqslant Y] \cap [Y \leqslant n]) = \mathbb{P}\Big(\bigcup_{n=0}^{+\infty} [X \leqslant Y] \cap [Y \leqslant n]\Big) = \mathbb{P}\Big([X \leqslant Y] \cap \bigcup_{n=0}^{+\infty} [Y \leqslant n]\Big) = \mathbb{P}(X \leqslant Y)$$

   puisque $\bigcup\limits_{n=0}^{+\infty} [Y \leqslant n]$ est un événement presque-certain, étant donné que Y est une variable aléatoire réelle.

   On a donc bien démontré que $\lim\limits_{x\to+\infty} H(x) = \mathbb{P}(X \leqslant Y)$, et d'autre part :

   $[X \leqslant Y] \cap [Y \leqslant 0] \subset [Y \leqslant 0]$, donc $0 \leqslant \mathbb{P}([X \leqslant Y] \cap [Y \leqslant 0]) \leqslant \mathbb{P}(Y \leqslant 0) = 0$ toujours par croissance de la probabilité, et puisque Y est annoncée à valeurs positives ; ainsi $H(0) = 0$.

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

a) On peut écrire :

$[X \leqslant Y] \cap [Y \leqslant v] = [X \leqslant Y] \cap \big([Y \leqslant u] \cup [u < Y \leqslant v]\big) = \big([X \leqslant Y] \cap [Y \leqslant u]\big) \cup \big([X \leqslant Y] \cap [u < Y \leqslant v]\big)$

où l'union est disjointe puisque $[Y \leqslant u] \cap [u < Y \leqslant v] = \varnothing$, donc :

$$\mathbb{P}\big([X \leqslant Y] \cap [Y \leqslant v]\big) = \mathbb{P}([X \leqslant Y] \cap [Y \leqslant u]) + \mathbb{P}([X \leqslant Y] \cap [u < Y \leqslant v])$$
$$\iff \mathbb{P}([X \leqslant Y] \cap [u < Y \leqslant v]) = H(v) - H(u)$$

Par ailleurs, on dispose des implications :

- Si $[X \leqslant u]$ **et** $[u < Y \leqslant v]$ sont réalisés, **alors** par transitivité de l'inégalité, $[X \leqslant Y]$ et $[u < Y \leqslant v]$ le sont aussi, donc :
  $[X \leqslant u] \cap [u < Y \leqslant v] \Longrightarrow [X \leqslant Y] \cap [u < Y \leqslant v]$, donc par croissance de la probabilité :

$$\mathbb{P}([X \leqslant u] \cap [u < Y \leqslant v]) \leqslant \mathbb{P}([X \leqslant Y] \cap [u < Y \leqslant v]) \iff \mathbb{P}(X \leqslant u) \times \mathbb{P}(u < Y \leqslant v) \leqslant \mathbb{P}([X \leqslant Y] \cap [u < Y \leqslant v])$$

puisque X et Y sont indépendantes, ce qui s'écrit :

$$F_X(u) \times \big(F_Y(v) - F_Y(u)\big) \leqslant H(v) - H(u)$$

- De même si $[X \leqslant Y]$ et $[u < Y \leqslant v]$ sont réalisés, alors par transitivité de l'inégalité, $[X \leqslant v]$ et $[u < Y \leqslant v]$ le sont aussi, donc :
  $[X \leqslant Y] \cap [u < Y \leqslant v] \subset [X \leqslant v] \cap [u < Y \leqslant v]$ et par croissance de la probabilité :

$$\mathbb{P}([X \leqslant Y] \cap [u < Y \leqslant v]) \leqslant P(X \leqslant v) \times \mathbb{P}(u < Y \leqslant v) \iff H(v) - H(u) \leqslant F_X(v).\big(F_Y(v) - F_Y(u)\big)$$

par transitivité de l'inégalité, on peut rassembler les deux inégalités en une seule, et en divisant par $v - u > 0$ on obtient bien :

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

b) L'expression $\dfrac{H(v) - H(u)}{v - u}$ doit ici être vue comme le taux d'accroissement de la fonction H en v, ou en u (puisque c'est aussi égal à $\dfrac{H(u) - H(v)}{u - v}$).

La restriction de la fonction densité f_Y étant continue sur $[0; +\infty[$, la restriction de la fonction de répartition F_Y à $\mathbb{R}^+$ est de classe $\mathcal{C}^1$ sur cet intervalle, et :

$$\lim_{u \to v^-} \frac{F_Y(v) - F_Y(u)}{v - u} = F_Y'(v) = f_Y(v) \quad \text{et} \quad \lim_{v \to u^+} \frac{F_Y(v) - F_Y(u)}{v - u} = F_Y'(u) = f_Y(u)$$

Comme par ailleurs F_X est continue sur $\mathbb{R}^+$ (puisque X est à densité) : $\lim\limits_{u \to v^-} F_X(u) = F_X(v)$ et $\lim\limits_{v \to u^+} F_X(v) = F_X(u)$.

Le théorème d'encadrement s'applique donc deux fois pour dire que :

- Pour tout $v \in \mathbb{R}_+$, puisque $\lim\limits_{u \to v-} F_X(u).\dfrac{F_Y(v) - F_Y(u)}{v - u} = F_X(v).f_Y(v) = \lim\limits_{u \to v^-} F_X(v).\dfrac{F_Y(v) - F_Y(u)}{v - u}$, alors
$$\lim_{u \to v^-} \frac{H(v) - H(u)}{v - u} = F_X(v).f_Y(v)$$
- Pour tout $u \in \mathbb{R}_+$, puisque $\lim\limits_{v \to u^+} F_X(u).\dfrac{F_Y(v) - F_Y(u)}{v - u} = F_X(u).f_Y(u) = \lim\limits_{v \to u^+} F_X(v).\dfrac{F_Y(v) - F_Y(u)}{v - u}$, alors
$$\lim_{v \to u^+} \frac{H(v) - H(u)}{v - u} = F_X(u).f_Y(u)$$

On en déduit que pour tout x de $\mathbb{R}^+$, on peut écrire :

$$\lim_{y\to x^+}\frac{H(x)-H(y)}{x-y}=F_X(x).f_Y(x)=\lim_{y\to x^-}\frac{H(x)-H(y)}{x-y},$$

ce qui prouve que la fonction H est dérivable en tout point x de $\mathbb{R}^+$, avec :

$$\forall x\in\mathbb{R}^+,\quad H'(x)=F_X(x).f_Y(x)$$

c) De tout ce qui précède, on déduit que :

$$\forall x\in\mathbb{R}^+,\quad \int_0^x F_X(t).f_Y(t)\ \mathrm{d}t=\Big[H(t)\Big]_0^x=H(x)-H(0)=H(x)$$

puisqu'on a vu que $H(0)=0$.

3. Le fait de savoir que $\lim\limits_{x\to+\infty}H(x)=\lim\limits_{x\to+\infty}\int_0^x F_X(t).f_Y(t)\mathrm{d}t$ existe et est finie (combinaison des résultats des questions 1.a) et 2.c)), assure que l'intégrale impropre $\int_0^{+\infty}F_X(t).f_Y(t)\mathrm{d}t$ est convergente, et le résultat de 1.b) donne bien :

$$\int_0^{+\infty}F_X(t).f_Y(t)\mathrm{d}t=\lim_{x\to+\infty}H(x)=\mathbb{P}(X\leqslant Y).$$

4. L'énoncé admettait qu'en utilisant la fonction $K:\ x\mapsto\mathbb{P}([X<Y]\cap[Y\leqslant x])$, on obtenait de même (et c'était admis) que :

$$\mathbb{P}(X<Y)=\int_0^{+\infty}F_X(t).f_Y(t)\mathrm{d}t=\mathbb{P}(X\leqslant Y)$$

On en déduit évidemment que : $\quad\mathbb{P}(X=Y)=\mathbb{P}(X\leqslant Y)-\mathbb{P}(X<Y)=0.$

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 non nul.

a)

b) Pour tout réel x, d'après le cours sur la loi exponentielle :

$$F_X(x)=\mathbb{P}(U-\theta\leqslant x)=\mathbb{P}(U\leqslant x+\theta)=F_U(x+\theta)=\begin{cases}0 & \text{si } x+\theta\leqslant 0\iff x\leqslant-\theta\\ 1-e^{-\lambda(x+\theta)} & \text{si } x+\theta>0\iff x>-\theta\end{cases}$$

c) La variable aléatoire V vérifie bien, puisqu'elle suit la loi exponentielle, les mêmes propriétés que la variable aléatoire Y générale qui intervient dans tout ce qui précède : elle est à valeurs positives et possède une densité $f_V:\ x\mapsto\begin{cases}0 & \text{si } x<0\\ \mu.e^{-\mu x} & \text{si } x\geqslant 0\end{cases}$ dont la restriction à $\mathbb{R}_+$ est continue sur cet intervalle. $X=U-\theta$ est bien indépendante de V puisque c'est le cas de U et V, en application du lemme des coalitions.

Les formules précédentes s'appliquent donc à X et V et on en déduit que pour tout $\theta\geqslant 0$:

$$\mathbb{P}(U-\theta\leqslant V)=\mathbb{P}(X\leqslant V)=\int_0^{+\infty}F_X(t).f_V(t)\mathrm{d}t=\int_0^{+\infty}\left(1-e^{-\lambda(t+\theta)}\right).\mu.e^{-\mu t}\mathrm{d}t$$

$$= \int_0^{+\infty} \mu.e^{-\mu t}\mathrm{d}t - \mu.e^{-\lambda.\theta}\int_0^{+\infty} e^{-(\lambda+\mu)t}\mathrm{d}t$$

La première intégrale vaut 1 puisque c'est en fait la valeur de $\int_{-\infty}^{+\infty} f_V(t)\mathrm{d}t$. Selon le même principe avec une loi exponentielle de paramètre $\lambda+\mu$, on a :

$\int_0^{+\infty}(\lambda+\mu).e^{-(\lambda+\mu)t}\mathrm{d}t = 1 \iff \int_0^{+\infty} e^{-(\lambda+\mu)t}\mathrm{d}t = \dfrac{1}{\lambda+\mu}$, donc :

$$\mathbb{P}(U-\theta \leqslant V) = 1 - \frac{\mu}{\lambda+\mu}e^{-\lambda.\theta} \qquad CQFD$$

Inégalité de Boole

6. On considère $(B_k)_{k\in\mathbb{N}^*}$ une famille d'événements d'un même espace probabilisé.

a) Montrons par récurrence sur $\mathbb{N}^*$ que la propriété $\mathcal{P}(n)$: "$\mathbb{P}\Big(\bigcup_{k=1}^{n} B_k\Big) \leqslant \sum_{k=1}^{n}\mathbb{P}(B_k)$", est vraie pour tout $n \in \mathbb{N}^*$.

I. Pour $n=1$, en présence d'un seul événement donc : $P\Big(\bigcup_{k=1}^{1} B_k\Big) = \mathbb{P}(B_1) = \sum_{k=1}^{1}\mathbb{P}(B_k)$, donc $\mathcal{P}(1)$ est évidemment vraie.

Il est en fait intéressant ici de s'intéresser au cas de $n=2$ événements : selon la formule du crible,

$\mathbb{P}(B_1\cup B_2) = \mathbb{P}(B_1)+\mathbb{P}(B_2)-\mathbb{P}(B_1\cap B_2) \leqslant \mathbb{P}(B_1)+\mathbb{P}(B_2)$ puisque $\mathbb{P}(B_1\cap B_2)\geqslant 0$ (c'est une probabilité).

On va d'ailleurs utiliser cela dans l'hérédité :

H. Supposons $\mathcal{P}(n)$ vraie pour un certain $n\in\mathbb{N}^*$, et sous cette hypothèse, montrons que $\mathcal{P}(n+1)$ est encore vraie.

On considère donc l'union $\bigcup_{k=1}^{n+1} B_k$, qu'on peut écrire $\Big(\bigcup_{k=1}^{n} B_k\Big)\cup B_{n+1}$.

L'inégalité de Boole démontrée pour deux éléments donne alors :

$$P\Big(\bigcup_{k=1}^{n+1} B_k\Big) \leqslant \mathbb{P}\Big(\bigcup_{k=1}^{n} B_k\Big) + \mathbb{P}(B_{n+1}) \overset{H.R.}{\leqslant} \sum_{k=1}^{n}\mathbb{P}(B_k) + \mathbb{P}(B_{n+1})$$

ce qui donne bien, par transitivité de l'inégalité :

$$\mathbb{P}\Big(\bigcup_{k=1}^{n+1} B_k\Big) \leqslant \sum_{k=1}^{n+1}\mathbb{P}(B_k)$$

donc $\mathcal{P}(n+1)$ est vraie si $\mathcal{P}(n)$ l'est.

C. La propriété est initialisée et héréditaire : elle est donc vraie pour tout $n\in\mathbb{N}^*$, d'après le principe de récurrence.

b) On suppose que la série $\sum_{k\geqslant 1}\mathbb{P}(B_k)$ converge.

La propriété de limite monotone donne toujours, dans le cas le plus général :

$\mathbb{P}\Big(\bigcup_{k\geqslant 1} B_k\Big) = \lim_{n\to+\infty}\mathbb{P}\Big(\bigcup_{k=1}^{n} B_k\Big)$, et l'hypothèse faite sur la série permet alors de passer à la limite dans l'inégalité précédente lorsque n tend vers $+\infty$, ce qui assure en effet que :

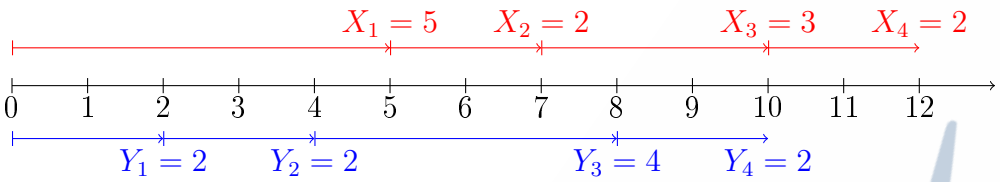
$$\mathbb{P}\Big(\bigcup_{k\geqslant 1} B_k\Big) \leqslant \sum_{k=1}^{+\infty}\mathbb{P}(B_k)$$

Partie II - Une compétition entre deux groupes

7. Avec les notations fournies par la (longue) introduction de cette partie :

   a) La variable aléatoire $\sum_{k=1}^{n} X_k$ représente le temps total mis par le groupe A pour résoudre les n premiers problèmes qui lui sont posés.

   b) On suppose que $X_1 = 5,\ X_2 = 2,\ X_3 = 3,\ X_4 = 2,\ Y_1 = 2,\ Y_2 = 2,\ Y_3 = 4,\ Y_4 = 2$.
   Un dessin aide toujours à bien visualiser le problème :



   Les problèmes résolus dans cet exemple sont dans l'ordre : $(Q_1, Q_2, P_1, P_2, Q_3, P_3, Q_4, P_4)$, et donc :

$$U_1 = 0 = U_2,\ U_3 = 1 = U_4,\ U_5 = 0,\ U_6 = 1,\ U_7 = 0$$

   et on n'a pas assez d'information pour déterminer U_8, au cas où le groupe B s'attèlerait à résoudre un cinquième problème qui peut encore être fini avant P_4, ou pas.

   c) Le script `Scilab` complété ci-dessous simule le jeu et pour $n,\ p$ donnés, affiche la liste des valeurs $U_1, U_2, \ldots, U_n$:

```
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)   // calcul du premier temps minimal de résolution

for k = 1:n
    if sommeX == mini then //si le dernier problème résolu l'a été par le gpe A
        U(k) =           //alors la nouvelle variable U_k prend la valeur 1
        sommeX = sommeX + grand(1, 1, "exp", 1/p)  //et le groupe A résout un
nouveau problème
    else                          //sinon le dernier problème résolu l'a été par le
groupe B, et U_k reste nul
        sommeY = sommeY + grand(1, 1, "exp", 1/q) // et le groupe B résout un
nouveau problème
    end
    mini = min(sommeX, sommeY)  // calcul du nouveau temps minimal de résolution
end
disp(U)     // affichage du vecteur complet contenant les valeurs U_1,...,U_n
```

   d) Le nombre total S_n de problèmes résolus par A parmi les n premiers problèmes résolus, correspond au nombre de 1 présents dans le vecteur `U` : il est donc clair que $S_n = \sum_{k=1}^{n} U_k$, et donc qu'on obtient très simplement la valeur de S_n en rajoutant à la fin du script précédent, la commande :

```
disp(sum(U))
```

8. *Loi de U_n.*

Dans cette question, on démontre par récurrence sur $n \geqslant 1$ que $\mathbb{P}(U_n = 1) = p$.

a) Il est clair que l'événement $[U_1 = 1]$ est réalisé si et seulement si le groupe A a résolu son premier problème avant que le groupe B ait fait de même, donc $[U_1 = 1] = [X_1 \leqslant Y_1]$, et $\mathbb{P}(U_1 = 1) = \mathbb{P}(X_1 \leqslant Y_1)$ se calcule avec la formule obtenue en 5.b) avec $\theta = 0$, et en adaptant les paramètres des lois exponentielles utilisées ici :

$$\mathbb{P}(U_1 = 1) = \mathbb{P}(X_1 - 0 \leqslant Y_1) = 1 - \frac{q}{p+q}.e^0 = 1 - q = p \quad \text{puisque } q = 1 - p$$

b) i. Sachant $[U_1 = 1]$, donc $[X_1 \leqslant Y_1]$ réalisé : $Y_1 - X_1$ prend alors une valeur positive, et il est donc évident que $\mathbb{P}_{[U_1=1]}(Y_1 - X_1 \leqslant x) = 0$ si x est un réel négatif.

ii. Soit x un réel positif ou nul :

$$\mathbb{P}_{[U_1=1]}(Y_1 - X_1 \leqslant x) = \frac{\mathbb{P}([U_1 = 1] \cap [Y_1 - X_1 \leqslant x])}{\mathbb{P}(U_1 = 1)} = \frac{\mathbb{P}([X_1 \leqslant Y_1] \cap [Y_1 \leqslant X_1 + x])}{p} = \frac{1}{p}\mathbb{P}([X_1 \leqslant Y_1 \leqslant X_1 + x])$$

Or : $\mathbb{P}(X_1 \leqslant Y_1 \leqslant X_1 + x) = \mathbb{P}(Y_1 \leqslant X_1 + x) - \mathbb{P}(Y_1 < X_1)$ (par une argumentation semblable à celle utilisée en 2.a)), donc :

$$\mathbb{P}(X_1 \leqslant Y_1 \leqslant X_1 + x) = \mathbb{P}(Y_1 - x \leqslant X_1) - \mathbb{P}(Y_1 - 0 \leqslant X_1) = 1 - \frac{p}{q+p}e^{-qx} - 1 + \frac{p}{q+p}e^0 = p - pe^{-qx}$$

donc en divisant par p : $\mathbb{P}_{[U_1=1]}(Y_1 - X_1 \leqslant x) = 1 - e^{-qx}$.

c) On a ainsi obtenu : $\mathbb{P}_{[U_1=1]}(Y_1 - X_1 \leqslant x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 - e^{-qx} & \text{si } x \geqslant 0 \end{cases}$, ce qui correspond à la fonction de répartition d'une loi exponentielle de paramètre q, qu'on interprète comme la *loi conditionnelle de $Y_1 - X_1$ sachant $[U_1 = 1]$*.

Par analogie, en échangeant les rôles de p et q, on peut considérer que la loi conditionnelle de $X_1 - Y_1$ sachant $[U_1 = 0]$ est la loi exponentielle de paramètre p.

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

Sachant que $[U_1 = 1]$ est réalisé, et donc que le groupe A a terminé son premier problème avant le groupe B : le temps supplémentaire que met le groupe B à terminer son premier problème est égal à $Y_1 - X_1$, et suit sous cette condition la même loi $\mathcal{E}(q)$ que Y_1 sans conditionnement : tout se passe donc comme si, après que le groupe A a résolu son premier problème, on "remettait les compteurs à zéro", X_2 devenant le temps de résolution du premier problème par le groupe A, $Y_1 - X_1$ le temps de résolution du premier problème par le groupe B, X_3 le temps de résolution du deuxième problème par le groupe A, Y_2 le temps de résolution du deuxième problème par le groupe B, etc... Illustration supplémentaire de la propriété d'absence de mémoire caractéristique des lois exponentielles !

Selon cette logique, et toujours sachant $[U_1 = 1]$, le $n+1$-ième problème est résolu par le groupe A si le n-ième problème résolu *après* le premier, l'est par le groupe A : en termes de probabilités et selon cette interprétation, on peut bien dire que :

$$\mathbb{P}_{[U_1=1]}(U_{n+1} = 1) = \mathbb{P}(U_n = 1) \overset{H.R.}{=} p$$

Une argumentation en tout point semblable pour le cas où $[U_1 = 0]$ est supposé réalisé, voit $X_1 - Y_1$ prendre la place de X_1 et suivre la même loi exponentielle de paramètre p, pour conclure que de même :

$$\mathbb{P}_{[U_1=1]}(U_{n+1} = 1) = \mathbb{P}(U_n = 1) = p$$

e) La formule des probabilités totales avec le s.c.e. $\big([U_1=0],[U_1=1]\big)$ permet de conclure :

$$\mathbb{P}(U_{n+1}=1)=\mathbb{P}(U_1=0).\mathbb{P}_{[U_1=0]}(U_{n+1}=1)+\mathbb{P}(U_1=1).\mathbb{P}_{[U_1=1]}(U_{n+1}=1)=p.\big(\mathbb{P}(U_1=0)+\mathbb{P}(U_1=1)\big)=p.1=p$$

Donc $\mathbb{P}(U_n=1)=p$ implique $\mathbb{P}(U_{n+1}=1)$ et la propriété est héréditaire : comme elle est aussi initialisée à $n=1$, elle est donc vraie pour tout $n\in\mathbb{N}^*$ d'après le principe de récurrence, et :

$$\forall n\in\mathbb{N}^*,\quad \mathbb{P}(U_n=1)=p$$

9. L'énoncé admettait ici que les variables aléatoires $U_1,\dots,U_n$ sont mutuellement indépendantes : $S_n=\sum_{k=1}^{n}U_k$ apparaît dans ce cas comme la somme de n variables de Bernoulli mutuellement indépendantes et de même paramètre p : S_n suit donc la loi binomiale de paramètres (n,p).

Soit $r\in\mathbb{N}$, on s'intéresse dans les questions qui suivent, à la probabilité a_r de l'événement

A_r : « il existe un $n\geqslant r$ 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) Il est certain que lorsqu'aucun problème n'a été résolu, le groupe A en a résolu 0 de plus que le groupe B : lorsque $r=0$, l'entier $n=0\geqslant r$ assure que l'événement A_0 est certainement réalisé, donc $a_0=\mathbb{P}(A_0)=1$.

b) Si $[U_1=1]$ est réalisé, alors le groupe A a terminé son premier problème avant que le groupe B ait fait de même : sachant ce fait, le groupe A a pris un problème d'avance, et la probabilité que le groupe A finisse par avoir r problèmes d'avance sur le groupe B (reformulation de l'événement A_r) est égale à la probabilité qu'à partir de la fin de la résolution en avance de son premier problème, le groupe A finisse par prendre $(r-1)$ problèmes d'avance supplémentaires, ce qui justifie la relation : $\mathbb{P}_{[U_1=1]}(A_r)=\mathbb{P}(A_{r-1})$.

De même, sachant $[U_1=0]$: le groupe A finit par avoir r problèmes d'avance si et seulement si à partir de la fin de la résolution de son premier problème par le groupe B, où tout est remis à zéro, le groupe A a refait son retard et finit par avoir $r+1$ problèmes d'avance sur le groupe B (sans compter donc son premier problème résolu), ce qui justifie également : $\mathbb{P}_{[U_1=0]}(A_r)=\mathbb{P}(A_{r+1})$.

c) La formule des probabilités totales et le s.c.e. $\big([U_1=0],[U_1=1]\big)$ permettent alors d'écrire, pour tout $r\geqslant 1$:

$$\mathbb{P}(A_r)=\mathbb{P}(U_1=0).\mathbb{P}_{[U_1=0]}(A_r)+\mathbb{P}(U_1=1).\mathbb{P}_{[U_1=1]}(A_r)\iff a_r=q.a_{r+1}+q.a_{r-1}\iff a_{r+1}=\frac{1}{q}a_r-\frac{p}{q}a_{r-1}.$$

d) Au vu de cette relation, la suite $(a_r)_{r\in\mathbb{N}}$ est donc récurrente linéaire d'ordre 2, d'équation caractéristique :

$$x^2=\frac{1}{q}x-\frac{p}{q}=0\iff qx^2-x+p=0$$

Cette équation du second degré a pour discriminant $\Delta=(-1)^2-4.pq=1-4pq$.

Comme $1-4pq=1-4p(1-p)=1-4p+4p^2=(1-2p)^2$ d'après une identité remarquable (!), on en déduit que $\Delta\geqslant 0$, et donc que l'équation admet deux solution réelles, distinctes ou confondues suivant que $p=1/2$ ou pas :

- Si $p\neq 1/2$, il y a alors deux racines :

$$x_1=\frac{1-(1-2p)}{2q}=\frac{p}{q}\quad\text{et}\quad x_2=\frac{1+(1-2p)}{2q}=\frac{2(1-p)}{2q}=1\quad\text{puisque } 1-p=q$$

Il existe donc deux réels α et β tels que : $\forall r\in\mathbb{N},\ a_r=\alpha.\left(\frac{p}{q}\right)^r+\beta$.

- Si $p=\dfrac{1}{2}$, alors l'équation caractéristique a pour unique solution $x_0=-\dfrac{-1}{2q}=1$.

Il existe donc deux réels α et β tels que : $\forall r\in\mathbb{N},\ a_r=\alpha+\beta.n$.

Si α et β sont les deux réels auxquels l'énoncé faisait référence, alors on n'est pas censé aller plus loin dans le calcul ici.

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

- Si $p = 1/2$: alors $a_r = \alpha + \beta.r$: sachant qu'il s'agit d'une probabilité, a_r est toujours compris entre 0 et 1, la suite est bornée. Or si $\beta \neq 0$, alors $\lim\limits_{r\to+\infty} \alpha + \beta.r$ vaut $+\infty$ où $-\infty$ selon le signe de β, ce qui est incompatible avec le caractère borné de la suite.

  La seule solution est donc que $\beta = 0$, et donc que $a_r = \alpha$ pour tout $r \in \mathbb{N}$: la suite $(a_r)_{r\in\mathbb{N}}$ est bien constante, toujours égale à son premier terme $a_0 = 1$.

- Par un argument similaire : si $p > 1/2$, alors $q = 1-p < 1/2 < p$, donc $\dfrac{p}{q} > 0$ et $\lim\limits_{r\to+\infty} \left(\dfrac{p}{q}\right)^r = +\infty$ dans ce cas.

  Toujours à cause des mêmes propriétés de la suite (a_r), on en déduit que dans l'expression de a_r obtenue ci-dessus, il faut que α soit nul, et donc que $a_r = \beta$ pour tout $r \in \mathbb{N}$: la suite (a_r) est bien constante, toujours égale à $a_0 = 1$.

12. Le cas $p < 1/2$.

a) Soit k un entier naturel.

i. L'événement A_{2k} est réalisé si et seulement si le groupe A finit par avoir $2k$ problèmes résolus d'avance sur le groupe B : il existe donc un entier j tel que le groupe B ayant résolu j problèmes, le groupe A en a résolu $j + 2k$; le nombre total de problèmes résolus est alors $2j + 2k = 2(j + k)$, et on peut écrire :

$$A_{2k} = \bigcup_{j\geqslant 0} [S_{2(j+k)} = j + 2k] \overset{[i=j+k]}{=} \bigcup_{i\geqslant k} [S_{2i} = i + k]$$

ii. La loi de S_n a été obtenue à la question 9 : chaque variable S_{2i} suit la loi binomiale de paramètres $(2i, p)$, donc :

Pour tout $i \geqslant k$, $\mathbb{P}(S_{2i} = i + k) = \binom{2i}{i+k} p^{i+k} q^{2i-(i+k)} = \binom{2i}{i+k} p^{i+k} q^{i-k}$.

sachant d'ailleurs que si $i \geqslant k$, alors $2i \geqslant i + k$ et $i + k \in S_{2i}(\Omega)$.

iii. On sait que $\sum\limits_{j=0}^{2i} \binom{2i}{j} = 2^{2i} = 4^i$: c'est une propriété de la somme de tous les coefficients binomiaux de la même ligne (ici $2i$) du triangle de Pascal, qu'on peut aussi voir comme un cas particulier de la formule du binôme de Newton :

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

S'agissant d'une somme de termes tous positifs, chaque terme est inférieur à la somme de tous les termes ; et comme on l'a déjà remarqué : pour tout entier $i \geqslant k$, $0 \leqslant i+k \leqslant 2i$ donc on a bien $\binom{2i}{i+k} \leqslant 4^i$.

iv. De tout ce qui précède, on déduit que : pour tout entier $i \geqslant k$,

$$\mathbb{P}(S_{2i} = i + k) \leqslant 4^i.p^{i+k}.q^{i-k} \iff \mathbb{P}(S_{2i} = i + k) \leqslant (4pq)^i.\left(\frac{p}{q}\right)^k.$$

Une étude classique de la fonction trinôme $p \mapsto p(1 - p) = -p^2 + p$ nous rappelle qu'elle a pour racines évidentes 0 et 1,que sa parabole représentative est "tournée vers le bas" et a pour sommet le point d'abscisse $\alpha = -\dfrac{b}{2a} = \dfrac{1}{2}$ et pour ordonnée $\dfrac{1}{2}(1-\dfrac{1}{2}) = \dfrac{1}{4}$, et qu'on peut donc écrire :

$\forall p \in]0;1/2[,\ 0 < p(1-p) < \dfrac{1}{4} \Longrightarrow 0 < 4pq < 1$. La série $\displaystyle\sum_{i\geqslant k}(4pq)^i.\left(\frac{p}{q}\right)^k$ est donc convergente, comme série géométrique de raison $4pq \in]0;1[$: par comparaison de séries à termes positifs, la série $\displaystyle\sum_{i\geqslant k}\mathbb{P}(S_{2i}=k+i)$ converge aussi et :

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

d'après la formule générale pour la somme d'une série géométrique à partir d'un certain rang (ici k).

b) D'après l'inégalité de Boole obtenue à la question 6 :

$$a_{2k} = \mathbb{P}(A_{2k}) \overset{q.12.a)i.}{=} \mathbb{P}\Big(\bigcup_{i\geqslant k}[S_{2i}=k+i]\Big) \leqslant \sum_{i=k}^{+\infty}\mathbb{P}(S_{2i}=k+i) \overset{12.a)iv.}{\leqslant} \left(\frac{p}{q}\right)^k.\frac{(4pq)^k}{1-4pq}$$

où comme on a déjà eu l'occasion de le dire, pour $p<1/2$ on a $0<\dfrac{p}{q}<1$ et $0<4pq<1$, donc $\lim\limits_{k\to+\infty}\left(\dfrac{p}{q}\right)^k = \lim\limits_{k\to+\infty}(4pq)^k = 0$, donc $\lim\limits_{k\to+\infty}\left(\dfrac{p}{q}\right)^k.\dfrac{(4pq)^k}{1-4pq} = 0$.

Comme une probabilité est toujours positive : $\forall k\in\mathbb{N},\ 0 \leqslant a_{2k} \leqslant \left(\dfrac{p}{q}\right)^k.\dfrac{(4pq)^k}{1-4pq} = 0$ et le théorème d'encadrement permet de conclure que

$$\text{si } p<1/2 \text{ alors} \quad \lim_{k\to+\infty} a_{2k} = 0$$

c) En reprenant le résultat de 10.d), on peut écrire que : $\forall k\in\mathbb{N},\ a_{2k} = \alpha.\left(\dfrac{p}{q}\right)^{2k} + \beta$; on a toujours, pour $p<1/2,\ 0<\dfrac{p}{q}<1$, donc $\lim\limits_{k\to+\infty}\left(\dfrac{p}{q}\right)^{2k} = 0$ et $\lim\limits_{k\to+\infty} a_{2k} = \beta$ avec cette expression.
Mais on vient aussi de démontrer que $\lim\limits_{k\to+\infty} a_{2k} = 0$, donc par unicité de la limite :
$\beta = 0$ et $\forall r\in\mathbb{N},\ a_r = \alpha.\left(\dfrac{p}{q}\right)^r$.

La suite (a_r) est donc dans ce cas, géométrique de raison $\dfrac{p}{q}$, et comme son premier terme a_0 est égal à 1, alors $\alpha = 1$ et on a bien :

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

On a ainsi établi dans les questions 11 et 12 :

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

Partie III - La blockchain et la stratégie de la double dépense

13. Avec les notations de l'énoncé, on cherche la loi de la variable aléatoire T_n égale au nombre de problèmes résolus par le groupe A lorsqu'on place Q_n dans la liste des problèmes résolus.

a) $T_n(\Omega) = \mathbb{N}$ car a priori le temps nécessaire au groupe B pour résoudre n problèmes peut être arbitrairement grand, et pour tout $k \in \mathbb{N}$, l'événement $[T_n = k]$ signifie :

- que le problème Q_n est le dernier problème résolu par l'un des deux groupes, en l'occurrence le groupe B
- et qu'auparavant, le groupe A a résolu k problèmes et le groupe B $n-1$ problèmes, pour un total de $n+k-1$ problèmes : c'est l'événement $[S_{n+k-1} = k]$. Le problème Q_n est donc le $n+k$-ième problème résolu, ce qui correspond à l'événement $[U_{n+k} = 0]$.

D'où en effet, l'égalité d'événements : $[T_n = k] = [S_{n+k-1} = k] \cap [U_{n+k} = 0]$.

b) La variable aléatoire S_{n+k-1} est la somme des variables aléatoires $U_1, \ldots, U_{n+k-1}$, toutes indépendantes de U_{n+k} (admis question 9), donc par le lemme des coalitions, S_{n+k-1} est indépendante de U_{n+k}, et :

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(T_n = k) = \mathbb{P}(S_{n+k-1} = k) \times \mathbb{P}(U_{n+k} = 0) = \binom{n+k-1}{k} p^k q^{n-1} \times q = \binom{n+k-1}{k} p^k q^n.$$

14. a) La formule des probabilités totales appliquée avec le système complet d'événements $\left([T_n = k]\right)_{k \in \mathbb{N}}$ pour l'événement G_n, s'écrit :

$$\mathbb{P}(G_n) = \sum_{k=0}^{+\infty} \mathbb{P}(T_n = k).\mathbb{P}_{[T_n=k]}(G_n)$$

où :

- pour tout entier $k \geqslant n+1$: si l'événement $[T_n = k]$ est réalisé, alors au moment où le problème Q_n est mis dans la liste des problèmes résolus, le groupe A a déjà résolu un nombre de problèmes strictement supérieur à n, et par conséquent dès l'instant t la condition de réalisation de l'événement G_n est réalisée, et $\mathbb{P}_{[T_n=k]}(G_n) = 1$.
- pour tout entier k compris entre 0 et n : pour que G_n soit réalisé il faut qu'à partir du moment où le problème Q_n a été mis dans la liste des problèmes résolus, il existe un instant auquel le groupe A a rattrapé son retard de $n-k$ problèmes résolus sur le groupe B et en a résolu un de plus encore par rapport au nombre de ceux que le groupe B a résolus après Q_n ; bref, il faut que lors de la deuxième étape le groupe A finisse par avoir $n-k+1$ problèmes d'avances sur le groupe B. Par indépendance des temps de résolution des problèmes, on peut donc dire que $\mathbb{P}_{[T_n=k]}(G_n) = \mathbb{P}(A_{n-k+1}) = a_{n-k+1}$

Il reste donc à séparer la somme précédente en deux à l'indice n, ce qui permet d'écrire :

$$\mathbb{P}(G_n) = \sum_{k=0}^{n} \mathbb{P}(T_n = k).a_{n-k+1} + \sum_{k=n+1}^{+\infty} \mathbb{P}(T_n = k).1 = \sum_{k=0}^{n} \mathbb{P}(T_n = k).a_{n-k+1} + \mathbb{P}(T_n \geqslant n+1)$$

b) Lorsque $p \geqslant \dfrac{1}{2}$, on a vu que pour tout entier r, $a_r = 1$, donc dans ce cas :

$$\mathbb{P}(G_n) = \sum_{k=0}^{n} \mathbb{P}(T_n = k) + \mathbb{P}(T_n \geqslant n+1) = \mathbb{P}(T_n \leqslant n) + \mathbb{P}(T_n \geqslant n+1) = 1$$

puisque les événements $[T_n \leqslant n]$ et $[T_n \geqslant n+1]$ sont contraires l'un de l'autre.

c) Lorsque $p < \dfrac{1}{2}$, d'après le résultat de la question 12., on peut écrire :

$$\mathbb{P}(G_n) = \mathbb{P}(T_n \geqslant n+1) + \sum_{k=0}^{n} \mathbb{P}(T_n = k).\left(\frac{p}{q}\right)^{n-k+1} = 1 - \sum_{k=0}^{n} \mathbb{P}(T_n = k) + \sum_{k=0}^{n} \mathbb{P}(T_n = k).\left(\frac{p}{q}\right)^{n-k+1}$$

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

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

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

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

a) Pour tout $n \in \mathbb{N}^*$:

$$1 - u_n(p) + \frac{p}{q} u_n(q) = 1 - (1-p)^n \sum_{k=0}^{n} \binom{n+k-1}{k} p^k + \frac{p}{q}(1-q)^n \sum_{k=0}^{n} \binom{n+k-1}{k} q^k$$

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

puisque $p = 1-q$ et $q = 1-p$, donc $(1-p)^n p^k = p^k q^n$ et $\dfrac{p}{q}(1-q)^n q^k = p^{n+1} q^{k-1}$.

b) Pour tout $x \in [0;1]$ et $n \in \mathbb{N}^*$:

$$u_{n+1}(x) = (1-x)^{n+1} \sum_{k=0}^{n+1} \binom{n+k}{k} x^k = (1-x)^{n+1} \Big[\sum_{k=0}^{n} \binom{n+k}{k} x^k + \binom{2n+1}{n+1} x^{n+1} \Big]$$

$$= (1-x)^n \sum_{k=0}^{n+1} \binom{n+k}{k} (x^k - x^{k+1}) + (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1}$$

$$= (1-x)^n \Big[\sum_{k=0}^{n} \binom{n+k}{k} x^k - \sum_{k=0}^{n} \binom{n+k}{k} x^{k+1} \Big] + (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1}$$

$$= (1-x)^n \Big[\binom{n}{0} x^0 + \sum_{k=1}^{n} \Big(\binom{n+k-1}{k-1} + \binom{n+k-1}{k} \Big) x^k - \sum_{j=1}^{n+1} \binom{n+j-1}{j-1} x^j \Big]$$
$$+ (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1}$$

Formule de Pascal à gauche, changement d'indice $j = k+1$ à droite

$$= (1-x)^n \Big[1 + \sum_{k=1}^{n} \binom{n+k-1}{k} x^k + \underbrace{\sum_{k=1}^{n} \binom{n+k-1}{k-1} x^k - \sum_{k=1}^{n+1} \binom{n+k-1}{k-1} x^k}_{\text{télescopage}} \Big]$$
$$+ (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1}$$

$$= (1-x)^n \sum_{k=0}^{n} \binom{n+k-1}{k} x^k - (1-x)^n \binom{2n}{n} x^{n+1} + (1-x)^{n+1} \binom{2n+1}{n+1} x^{n+1}$$

$$= u_n(x) + (1-x)^n.x^{n+1}\Big[-\binom{2n}{n} + \binom{2n+1}{n+1}(1-x)\Big]$$

$$= u_n(x) + (1-x)^n.x^{n+1}\Big[\binom{2n+1}{n+1} - \binom{2n}{n} - \binom{2n+1}{n+1}x\Big]$$

$$u_{n+1}(x) = u_n(x) + (1-x)^n.x^{n+1}\Big[\binom{2n}{n+1} - \binom{2n+1}{n+1}x\Big] \quad \text{formule de Pascal à nouveau}$$

c) De la relation obtenue en 15.a), on déduit que :

$$\mathbb{P}(G_{n+1}) = 1 - u_{n+1}(p) + \frac{p}{q}u_{n+1}(q)$$

$$= 1 - \Big[u_n(p) + q^n p^{n+1}\Big(\binom{2n}{n+1} - \binom{2n+1}{n+1}p\Big)\Big] + \frac{p}{q}\Big[u_n(q) + p^n q^{n+1}\Big(\binom{2n}{n+1} - \binom{2n+1}{n+1}q\Big)\Big]$$

$$= 1 - u_n(p) + \frac{p}{q}u_n(q) + \binom{2n}{n+1}.\big(p^{n+1}q^n - p^{n+1}q^n\big) + \binom{2n+1}{n+1}.\big(q^n p^{n+2} - p^{n+1}q^{n+1}\big)$$

$$\mathbb{P}(G_{n+1}) = \mathbb{P}(G_n) - \Big(1 - \frac{p}{q}\Big)(pq)^{n+1}\binom{2n+1}{n+1}$$

d) On peut alors montrer par récurrence que $\mathcal{P}(n)$: "$\mathbb{P}(G_n) = \frac{p}{q} - \Big(1-\frac{p}{q}\Big)\sum_{k=1}^{n}\binom{2k-1}{k}(pq)^k$", est vraie pour tout $n \in \mathbb{N}^*$.

I. Pour $n = 1$: d'après 14.c), $\mathbb{P}(G_1) = 1 - \sum_{k=0}^{1}\binom{k}{k}(p^k q - p^2 q^{k-1}) = 1 - (q - \frac{p^2}{q} + pq - p^2)$

$= p + \frac{p^2}{q} + pq - p^2 = \frac{p(q+p)}{q} + pq - p^2 = \frac{p}{q} + pq - p^2$, et d'autre part :

$\frac{p}{q} - \Big(1-\frac{p}{q}\Big)\sum_{k=1}^{1}\binom{2k-1}{k}(pq)^k = \frac{p}{q} - \Big(1-\frac{p}{q}\Big)\binom{1}{1}(pq)^1 = \frac{p}{q} - pq + p^2$, donc $\mathcal{P}(1)$ est vraie.

H. Supposons $\mathcal{P}(n)$ vraie pour un certain $n \in \mathbb{N}^*$, et montrons qu'alors $\mathcal{P}(n+1)$ est encore vraie :

$$\mathbb{P}(G_{n+1}) = \mathbb{P}(G_n) - \Big(1-\frac{p}{q}\Big)(pq)^{n+1}\binom{2n+1}{n+1}$$

$$\overset{H.R.}{=} \frac{p}{q} - \Big(1-\frac{p}{q}\Big)\sum_{k=1}^{n}\binom{2k-1}{k}(pq)^k - \Big(1-\frac{p}{q}\Big)\binom{2n+1}{n+1}(pq)^{n+1}$$

$$= \frac{p}{q} - \Big(1-\frac{p}{q}\Big)\sum_{k=1}^{n+1}\binom{2k-1}{k}(pq)^k$$

puisque le dernier terme correspond bien à celui de la somme lorsque $k = n+1$; ainsi, $\mathcal{P}(n+1)$ est vraie si $\mathcal{P}(n)$ l'est.

C. La propriété est initialisée et héréditaire : elle est donc vraie pour tout $n \in \mathbb{N}^*$, d'après le principe de récurrence.

16. *Application à la sécurisation des transactions*

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

a) La formule demandée vient du cours, elle est souvent connue comme la "formule sans nom", et se démontre avec l'expression explicite des coefficients binomiaux : pour $k \in [\![1;n]\!]$,

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

La notion de fonction récursive n'étant plus explictement au programme, on peut se servir de cette relation en l'itérant :

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1} = \frac{n}{k}\times\frac{n-1}{k-1}\binom{n-2}{k-2} = \ldots = \frac{n}{k}\times\frac{n-1}{k-1}\times\ldots\times\frac{n-(k-1)}{1}\binom{n-k}{0} = \prod_{i=0}^{k-1}\frac{n-i}{k-i}$$

D'où le script `Scilab`, où le résultat voulu est la valeur finale de la variable `p` :

```
function p = binom(n,k)
    p = 1
    for i = 0:(k-1)
        p = p * (n-i)/(k-i)
    end
endfunction
```

b) On utilise alors la valeur de $\mathbb{P}(G_1)$ et la relation de récurrence entre $\mathbb{P}(G_{n+1})$ et $\mathbb{P}(G_n)$ obtenues précédemment pour calculer le plus petit entier n tel que $\mathbb{P}(G_n) \leqslant \varepsilon$ pour $p < \dfrac{1}{2}$ et $\varepsilon > 0$ saisis au clavier par l'utilisateur :

```
p = input('p = ')
eps = input('epsilon = ')

q = 1-p
g = p/q-p*q+p^2
n = 1

while g > eps
    g = g - (1-p/q)*(p*q)^(n+1)*binom(2*n+1,n+1)
    n = n+1
end
disp(n)
```

Pour le plaisir (ce n'était pas demandé), on donne ci-dessous le script qui permet effectivement d'afficher le graphique fourni par l'énoncé :

```
T = []
eps = 1e-4

for p = 0.1:0.01:0.32
    q = 1-p
    g = p/q-p*q+p^2
    n = 1

    while g > eps
        g = g - (1-p/q)*(p*q)^(n+1)*binom(2*n+1,n+1)
        n = n+1
    end
    T = [T,n]
end

plot(0.1:0.01:0.32,T)
```