

CORRECTION DS5 - HEC ECE MATHS I 2020 (ADAPTÉ)

Partie I - Deux résultats généraux

Calcul d'une probabilité

1. (a) Soient $(x, y) \in \mathbb{R}^2$ avec $0 \leq x \leq y$. Comme $Y \leq x \Rightarrow Y \leq y$, on a l'inclusion entre événements :

$$[Y \leq x] \subset [Y \leq y]$$

et donc

$$([X \leq Y] \cap [Y \leq x]) \subset ([X \leq Y] \cap [Y \leq y])$$

d'où en passant aux probabilités :

$$H(x) \leq H(y)$$

$$\boxed{H \text{ est donc croissante sur } \mathbb{R}_+}.$$

- (b) Pour tout $x \in \mathbb{R}$, on a l'inclusion d'événements :

$$[X \leq Y] \cap [Y \leq x] \subset [X \leq Y]$$

Donc en passant aux probabilités :

$$H(x) \leq P(X \leq Y).$$

Par passage à la limite (qui existe d'après la question précédente), on obtient :

$$\boxed{\lim_{x \rightarrow +\infty} H(x) \leq P([X \leq Y]).}$$

- (c) On a :

$$([X \leq Y] \cap [Y \leq 0]) \subset [Y \leq 0]$$

donc $H(0) \leq P(Y \leq 0) = P(Y = 0) = 0$ (Y est à valeurs positives, et est à densité) ; et finalement

$$\boxed{H(0) = 0.}$$

2. (a) On a l'union disjointe :

$$([X \leq Y] \cap [Y \leq v]) = ([X \leq Y] \cap [Y \leq u]) \sqcup ([X \leq Y] \cap [u < Y \leq v])$$

donc en prenant les probabilités :

$$\boxed{H(v) = H(u) + P([X \leq Y] \cap [u < Y \leq v])}$$

On souhaiterait alors séparer en utilisant l'indépendance de X et Y , mais c'est impossible avec l'événement $[X \leq Y]$...

On passe encore par une inclusion. Si $X \leq u$ et $u < Y \leq v$ on a clairement $X \leq Y$; d'où :

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

et de même :

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

En prenant les probabilités :

$$P([X \leq u] \cap [u < Y \leq v]) \leq H(v) - H(u) \leq P([X \leq v] \cap [u < Y \leq v])$$

et on peut maintenant utiliser l'indépendance de X et Y pour écrire :

$$P([X \leq u])P([u < Y \leq v]) \leq H(v) - H(u) \leq P[X \leq v]P([u < Y \leq v])$$

ce qui donne

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

et enfin en divisant par $v - u > 0$:

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

(b) On utilise un taux de variation.

Soit $v = x \geq 0$, et $v = x + h$ avec $h > 0$: l'encadrement précédent s'écrit :

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

et pour $h \rightarrow 0$, $h > 0$:

- $F_X(x+h) \rightarrow F_X(x)$ par continuité de F_X (X à densité) ;
- $\frac{F_Y(x+h) - F_Y(x)}{h} \rightarrow F'_Y(x) = f_Y(x)$ car Y est à densité f_Y continue sur $\mathbb{R}_+$ donc F_Y est $\mathcal{C}^1$ sur $\mathbb{R}_+$ avec $F'_Y = f_Y$.
(NB : en 0 on n'a qu'une dérivée à droite...)

Par théorème des gendarmes on a donc

$$\forall x \geq 0, \lim_{h \rightarrow 0, h > 0} \frac{H(x+h) - H(x)}{h} = F_X(x) f_Y(x)$$

Pour $x > 0$, on trouve de même avec $u = x + h \geq 0$ et $v = x > 0$:

$$\forall x > 0, \lim_{h \rightarrow 0, h < 0} \frac{H(x+h) - H(x)}{h} = F_X(x) f_Y(x)$$

donc finalement $\boxed{H \text{ est dérivable sur } \mathbb{R}_+ \text{ et } H' = F_X f_Y}$ (avec seulement une dérivée à droite en $x = 0$).

(c) On en déduit directement :

$$\forall x \geq 0, H(x) - H(0) = \int_0^x H'(t) dt = \boxed{\int_0^x F_X(t) f_Y(t) dt}$$

d'où le résultat avec $H(0) = 0$.

3. On a encore :

$$P([X \leq Y]) = \lim_{x \rightarrow +\infty} H(x) = \lim_{x \rightarrow +\infty} \left(\int_0^x F_X(t) f_Y(t) dt \right)$$

d'où

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

(avec au passage la convergence de cette intégrale).

4. Avec $P([X \leq Y]) = P([X < Y])$ on a immédiatement :

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

5. *Application aux lois exponentielles*

(a) Pour tout réel x : $F_X(x) = P(X \leq x) = P(U \leq x + \theta) = F_U(x + \theta)$; et avec l'expression de la fonction de répartition de $\mathcal{E}(\lambda)$:

$$F_X(x) = \begin{cases} 0 & \text{si } x + \theta \leq 0 \\ 1 - e^{-\theta(x+\theta)} & \text{si } x + \theta > 0 \end{cases} = \boxed{\begin{cases} 0 & \text{si } x \leq -\theta \\ 1 - e^{-\theta(x+\theta)} & \text{si } x > -\theta \end{cases}}$$

- (b) On applique la formule de la question 3 avec F_X ci-dessus, f_V la densité usuelle de $\mathcal{E}(\mu)$ ($X = U - \theta$ et $Y = V$ vérifient bien les hypothèses de travail de la question 3 :

$$P(U - \theta \leq V) = \int_0^{+\infty} F_X(t) f_V(t) dt$$

Sur $[0, +\infty[$, $t \geq 0 \geq -\theta$, donc l'intégrale s'écrit :

$$P(U - \theta \leq V) = \int_0^{+\infty} (1 - e^{-\lambda(t+\theta)}) \mu e^{-\mu t} dt = \int_0^{+\infty} \mu e^{-\mu t} dt - e^{-\lambda\theta} \mu \int_0^{+\infty} e^{-(\lambda+\mu)t} dt$$

(où les deux intégrales convergent séparément car $\mu > 0$, $\lambda + \mu > 0$) et avec $\int_0^{+\infty} \lambda e^{-\lambda t} dt = 1$ (intégrale de la densité de $\mathcal{E}(\lambda)$) on a :

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

Inégalité de Boole

6. (a) Procédons par récurrence.

- **Initialisation** : Pour $n = 1$, on a à démontrer $P(B_1) \leq P(B_1)$ ce qui est trivialement vrai.
- **Hérédité** : Soit $n \in \mathbb{N}^*$ fixé ; on suppose

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

Comme pour deux événements A et B on a $P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B)$, on peut écrire, en utilisant l'hypothèse au rang n :

$$P\left(\left(\bigcup_{k=1}^n B_k\right) \cup B_{n+1}\right) \leq P\left(\bigcup_{k=1}^n B_k\right) + P(B_{n+1}) \leq \sum_{k=1}^n P(B_k) + P(B_{n+1}) = \sum_{k=1}^{n+1} P(B_k)$$

ce qui donne l'hérédité.

On a bien :

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

- (b) La série $\sum P(B_k)$ étant convergente et à termes positifs, on peut écrire :

$$\forall n \in \mathbb{N}^*, P\left(\bigcup_{k=1}^n B_k\right) \leq \sum_{k=1}^n P(B_k) \leq \sum_{k=1}^{+\infty} P(B_k)$$

La formule admise étant dans l'ancien programme d'ECE, c'est une conséquence du théorème de la limite monotone (pour les probabilités) :

$$P\left(\bigcup_{k=1}^{+\infty} B_k\right) = \lim_{n \rightarrow +\infty} P\left(\bigcup_{k=1}^n B_k\right)$$

et donc en faisant tendre $n \rightarrow +\infty$ dans l'inégalité précédente :

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

Partie II - Une compétition entre deux groupes

7. (a) $\sum_{k=1}^n X_k$ est le temps passé par le groupe A pour résoudre les n premiers problèmes $P_1, \dots, P_n$.

(b) Avec ces valeurs :

- P_1 est résolu au temps $X_1 = 5$;
- P_2 est résolu au temps $X_1 + X_2 = 7$;
- P_3 est résolu au temps $X_1 + X_2 + X_3 = 10$;
- P_4 est résolu au temps 12;
- Q_1 est résolu au temps $Y_1 = 2$;
- Q_2 est résolu au temps $Y_1 + Y_2 = 4$;
- Q_3 est résolu au temps 8;
- Q_4 est résolu au temps 10.

Donc les problèmes rangés dans l'ordre chronologique de leur résolution sont

$$Q_1, Q_2, P_1, P_2, Q_3, P_3, Q_4, P_4$$

en rangeant les « ex-æquo » P_3 et Q_4 (temps 10) selon la convention de l'énoncé.

Comme $U_i = 1$ ssi le i -ème problème de la liste ci-dessus est un P , et 0 sinon, on obtient

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

On ne peut pas connaître U_8 sans avoir la valeur de Y_5 .

En effet si $Y_5 = 1$, Q_5 sera résolu au temps 11, donc avant P_4 , et on aura $U_8 = 0$; tandis que si $Y_5 = 4$, Q_5 est résolu après P_4 et on a $U_8 = 1$.

(c)

```

1  import numpy as np
   import numpy.random as rd

   p = float(input("p = "))
5  n = int(input("n = "))
   q = 1 - p

   U = np.zeros(1,n)
   sommeX = rd.exponential(1/p)
10  sommeY = rd.exponential(1/q)

   mini = np.min(sommeX, sommeY)

   for k in range(n):
15     if sommeX == mini :
         U[k] = 1
         sommeX = sommeX + rd.exponential(1/p)
       else:
         sommeY = sommeY + rd.exponential(1/q)
20     mini = np.min(sommeX, sommeY)

   print(U)
```

En effet, `sommeX` et `sommeY` comptent les temps cumulés par les équipes A et B (respectivement) pour résoudre leurs problèmes successifs. En initialisation on compare X_1 et Y_1 et la première composante de U vaudra 1 ssi $X_1 \leq Y_1$.

À l'intérieur de la boucle **for**, dès qu'une équipe résout un problème on fixe la valeur de $U(k)$ correspondante, puis on ajoute à la **somme** le temps de la résolution suivante ; on compare à nouveau les temps cumulés de résolution pour avoir la valeur de la prochaine composante de U .

- (d) S_n est par définition le nombre de « 1 » de la liste U construite ci-dessus ; comme U ne contient que des 0 et des 1, il suffit de sommer ses composantes.

On ajoute donc :

```
1 print(np.sum(U))
```

8. Loi de U_n

- (a) Par définition, $U_1 = 1$ ssi le premier problème résolu l'est par l'équipe A ; donc ssi P_1 est résolu avant (au sens large) Q_1 ; ce qui équivaut à $X_1 \leq Y_1$.

D'où $P([U_1 = 1]) = P([X_1 \leq Y_1])$.

$X_1 \hookrightarrow \mathcal{E}(p)$ et $Y_1 \hookrightarrow \mathcal{E}(q)$ vérifient les hypothèses de la partie I ; on peut donc appliquer la formule de 5b avec $\theta = 0$, $\lambda = p > 0$ et $\mu = q > 0$.

On a alors :

$$P(X_1 \leq Y_1) = 1 - \frac{q}{p+q}e^{-0} = 1 - q = \boxed{p}$$

- (b) i. Pour $x < 0$:

$$P_{[U_1=1]}(Y_1 - X_1 \leq x) = \frac{P([U_1 = 1] \cap [Y_1 - X_1 \leq x < 0])}{P([U_1 = 1])} = \frac{P([X_1 \leq Y_1] \cap [Y_1 - X_1 \leq x < 0])}{P([U_1 = 1])} = \boxed{0}$$

car l'événement au numérateur est impossible.

- ii. Cette fois, avec $x \geq 0$:

$$\begin{aligned} P_{[U_1=1]}(Y_1 - X_1 \leq x) &= \frac{P([U_1 = 1] \cap [Y_1 - X_1 \leq x])}{P([U_1 = 1])} \\ &= \frac{P([X_1 \leq Y_1 \leq X_1 + x])}{P([U_1 = 1])} \\ &= \boxed{\frac{1}{p}P([X_1 \leq Y_1 \leq X_1 + x])} \quad \text{d'après 8a.} \end{aligned}$$

Il nous faut donc calculer $P([X_1 \leq Y_1 \leq X_1 + x])$.

On cherche à utiliser 5b. On a l'union disjointe :

$$[X_1 \leq Y_1] = [X_1 \leq Y_1 \leq X_1 + x] \sqcup [Y_1 > X_1 + x]$$

d'où :

$$\begin{aligned} P([X_1 \leq Y_1 \leq X_1 + x]) &= P([X_1 \leq Y_1]) - P([Y_1 > X_1 + x]) \\ &= p - (1 - P([X_1 + x \geq Y_1])) \\ &= p - 1 + P([Y_1 - x \leq X_1]) \end{aligned}$$

et on peut appliquer 5b avec $\theta = x \geq 0$, $\lambda = q > 0$, $\mu = p > 0$.

On trouve, avec $p + q = 1$:

$$P([X_1 \leq Y_1 \leq X_1 + x]) = p - 1 + \left(1 - \frac{p}{p+q}e^{-qx}\right) = p(1 - e^{-qx})$$

et donc finalement

$$\boxed{P_{[U_1=1]}(Y_1 - X_1 \leq x) = 1 - e^{-qx}}$$

- (c) En regroupant les résultats de 8b :

$$\forall x \in \mathbb{R}, P_{[U_1=1]}(Y_1 - X_1 \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-qx} & \text{si } x \geq 0 \end{cases}$$

et on peut conclure que

$$\boxed{\text{Conditionnellement à } [U_1 = 1], Y_1 - X_1 \hookrightarrow \mathcal{E}(q)}$$

Pour examiner la loi de $X_1 - Y_1$ conditionnellement à $[U_1 = 0]$, il suffit d'échanger les rôles des X_i et des Y_i ; donc de p et q . Donc :

$$\boxed{\text{Conditionnellement à } [U_1 = 0], X_1 - Y_1 \hookrightarrow \mathcal{E}(q)}$$

(NB : il y a quand même le cas d'égalité $X_1 = Y_1$ qui donne $[U_1 = 1]$ et rompt cette symétrie. On aura en fait à examiner la probabilité $P([Y_1 < X_1 < Y_1 + x])$, mais on obtiendra les mêmes expressions en vertu (d'une petite généralisation) de la question 4.

(d) Sachant $[U_1 = 1]$, le premier problème résolu est P_1 .

Si on prend la résolution de P_1 comme nouvelle origine des temps, le problème en (X_i) et (Y_j) se ramène à un problème similaire en les variables $(X_2, X_3, \dots)$ et $(Y_1 - X_1, Y_2, Y_3, \dots)$.

On effectue alors des calculs avec la mesure de probabilité donnée par la loi conditionnelle $P_{[U_1=1]}$.

U_1 ne dépend que de X_1 et Y_1 , donc par lemme des coalitions, $U_1, X_2, X_3, \dots, Y_2, Y_3, \dots$ sont mutuellement indépendantes : donc **pour la mesure de probabilité** $P_{[U_1=1]}$, les $X_{i \geq 2}$ suivent toujours $\mathcal{E}(p)$ et les $Y_{i \geq 2}$ suivent $\mathcal{E}(q)$; enfin $Y_1 - X_1 \hookrightarrow \mathcal{E}(q)$ d'après 8c.

Considérer U_{n+1} se ramène à considérer la résolution de $(n+1)$ problèmes ; avec notre décalage temporel on se ramène à la résolution de n problèmes et on peut appliquer l'hypothèse de récurrence (pour la probabilité $P_{[U_1=1]}$) :

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

Sachant $[U_1 = 0]$ c'est la même chose. On décale l'origine des temps à la résolution de Q_0 , et conditionnellement à $[U_1 = 0]$:

- $X_1 - Y_1, X_2, X_3, \dots \hookrightarrow \mathcal{E}(p)$;
- $Y_2, Y_3, \dots \hookrightarrow \mathcal{E}(q)$

et on conclut de même :

$$\boxed{P_{[U_1=0]}([U_{n+1} = 1]) = P([U_n = 1]) = p}$$

(e) La formule des probabilités totales avec le SCE ($[U_1 = 0], [U_1 = 1]$) donne alors

$$P([U_{n+1} = 1]) = pP([U_1 = 0]) + pP([U_1 = 1]) = p$$

donc on a l'hérédité.

Par initialisation (8a) et hérédité on a donc :

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

9. Les U_i sont donc mutuellement indépendantes, et suivent toutes $\mathcal{B}(p)$ d'après 8 ; donc

$$\boxed{S_n = \sum_{k=1}^n U_k \hookrightarrow \mathcal{B}(n, p)}$$

10. (a) A_0 s'écrit :

« Il existe un $n \geq 0$ tel que, lorsque n problèmes en tout ont été résolus, le groupe A en a résolu autant que le groupe B »

$n = 0$ convient clairement (A et B ont alors résolu 0 problème chacun), donc A_0 est certain :

$$\boxed{a_0 = P(A_0) = 1}$$

- (b) Sachant $[U_1 = 1]$ (A résout son problème en premier), des arguments identiques à ceux de 8d permettent de fixer l'origine des temps à la résolution de P_1 .

A ayant maintenant un problème d'avance, il aura r problèmes d'avance sur B à un instant donné ssi il a $r - 1$ problèmes d'avance sur l'intervalle de temps observé.

Donc

$$P_{[U_1=1]}(A_r) = P(A_{r-1})$$

De même, si $[U_1 = 0]$, A part maintenant avec un problème de retard, et doit en remonter $r + 1$:

$$P_{[U_1=0]}(A_r) = P(A_{r+1})$$

- (c) Par probas totales avec le SCE ($[U_1 = 0], [U_1 = 1]$) :

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

ce qui donne bien :

$$\forall r \geq 1, \quad a_{r+1} = \frac{1}{q}a_r - \frac{p}{q}a_{r-1}$$

soit encore :

$$\forall r \in \mathbb{N}, \quad a_{r+2} = \frac{1}{q}a_{r+1} - \frac{p}{q}a_r$$

- (d) On a une suite récurrente linéaire d'ordre 2 ; le discriminant de son équation caractéristique vaut $\Delta = \frac{1-4pq}{q^2} = \left(\frac{1-2p}{q}\right)^2$ avec l'indication et on trouve les deux solutions $x_1 = 1$ et $x_2 = \frac{p}{q}$ (avec $x_1 = x_2$ dans le cas $\Delta = 0$ ($\Leftrightarrow p = q = \frac{1}{2}$)).

- Pour $p \neq \frac{1}{2}$, il existe $(\alpha, \beta) \in \mathbb{R}^2$ telles que :

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

- Pour $p = \frac{1}{2}$, il existe $(\alpha', \beta') \in \mathbb{R}^2$ telles que :

$$\forall r \in \mathbb{N}, \quad a_r = (\alpha' + \beta'r)$$

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

- Supposons $p > \frac{1}{2}$. Alors d'après la question 10d il existe $(\alpha, \beta) \in \mathbb{R}^2$ tel que

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

Montrons par l'absurde que $\beta = 0$: supposons $\beta \neq 0$. On a $p > \frac{1}{2}$ donc $\frac{p}{q} > 1$, d'où $\lim_{r \rightarrow +\infty} a_r = +\infty$ (si $\beta > 0$) ou $\lim_{r \rightarrow +\infty} a_r = -\infty$ (si $\beta < 0$), ce qui contredit le fait que $a_r \in [0, 1]$ pour tout $r \in \mathbb{N}$ (c'est la probabilité d'un événement). Ainsi $\beta = 0$ et $a_r = \alpha$ pour tout $r \in \mathbb{N}$. Par ailleurs pour $r = 0$ la question 10a donne $\alpha = 1$.

- Supposons $p = \frac{1}{2}$. Alors d'après la question 10d il existe $(\alpha', \beta') \in \mathbb{R}^2$ tel que

$$\forall r \in \mathbb{N}, \quad a_r = \alpha' + \beta' r$$

On vérifie de même par l'absurde que $\beta' = 0$: si $\beta' \neq 0$, alors $\lim_{t \rightarrow +\infty} a_r = \pm\infty$, contredisant $a_r \in [0, 1]$ pour tout $r \in \mathbb{N}$. Ainsi $a_r = \alpha'$ pour tout $r \in \mathbb{N}$. De plus pour $r = 0$ la question 10a donne $\alpha' = 1$.

Finalement on a montré que $\boxed{\text{si } p \geq \frac{1}{2}, \text{ alors } (a_r)_{n \in \mathbb{N}} \text{ est constante égale à } 1.}$

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

(a) i. Justifions cette égalité par double inclusion.

- Pour tout $i \geq k$, $[S_{2i} = i + k]$ signifie qu'au moment où $2i$ problèmes ont été résolus, $i + k$ d'entre eux ont été résolus par A , donc $2i - (i + k) = i - k$ par B . Ainsi à cet instant A a bien réalisé $2k$ problèmes de plus que B , d'où $[S_{2i} = i + k] \subset A_{2k}$. Cette inclusion étant valable pour tout $i \geq k$, on a bien $\bigcup_{i \geq k} [S_{2i} = i + k] \subset A_{2k}$.
- Supposons A_{2k} réalisé et notons $n \geq 2k$ le nombre total de problèmes résolus à l'instant où A en a résolu $2k$ de plus que B . Alors à l'issue du n -ème problème, A en a résolu S_n et B en a résolu $S_n - 2k$, d'où $S_n + (S_n - 2k) = n$, ou encore $n = 2(S_n - k)$. Ainsi n est pair et de la forme $n = 2i$ avec $i = S_n - k$. Il existe alors bien un entier naturel $i \geq k$ tel que l'événement $[S_{2i} = i + k]$ est réalisé. On a bien montré $A_{2k} \subset \bigcup_{i \geq k} [S_{2i} = i + k]$.

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

ii. Soit $i \geq k$ un entier. D'après la question 9, $S_{2i} \hookrightarrow \mathcal{B}(2i, p)$, donc

$$P(S_{2i} = i + k) = \binom{2i}{i+k} p^{i+k} q^{2i-(i+k)} = \boxed{\binom{2i}{i+k} p^{i+k} q^{i-k}}$$

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

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

Soit $i \geq k$ un entier. Alors $i + k \in \llbracket 0, 2i \rrbracket$ donc on peut décomposer la somme ci-dessus :

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

or $\sum_{j=0}^{i+k-1} \binom{2i}{j} \geq 0$ et $\sum_{j=i+k+1}^{2i} \binom{2i}{j} \geq 0$ comme sommes de termes positifs, ainsi

$$\boxed{4^i \geq \binom{2i}{i+k}}$$

iv. D'après les deux questions précédentes, pour tout entier $i \geq k$, puisque $p^{i+k} q^{i-k} \geq 0$:

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

Or $p < \frac{1}{2}$ donc $1 - 4pq = (1 - 2p)^2 > 0$, donc $0 < 4pq < 1$, par conséquent la série géométrique de raison $4pq$ converge, et en sommant l'inégalité précédente pour $i \geq k$:

$$\begin{aligned} \sum_{i=k}^{+\infty} P(S_{2i} = i + k) &\leq \left(\frac{p}{q}\right)^k \sum_{i=k}^{+\infty} (4pq)^i \\ &\leq \left(\frac{p}{q}\right)^k \sum_{j=0}^{+\infty} (4pq)^{k+j} \\ &\leq \boxed{\left(\frac{p}{q}\right)^k \frac{(4pq)^k}{1 - 4pq}} \end{aligned}$$

(b) En utilisant l'inégalité de Boole, puis le résultat de la question précédente :

$$\begin{aligned} 0 \leq P(A_{2k}) &= P\left(\bigcup_{i \geq k} [S_{2i} = i + k]\right) \\ &\leq \sum_{i \geq k} P(S_{2i} = i + k) \\ &\leq \left(\frac{p}{q}\right)^k \frac{(4pq)^k}{1 - 4pq} \end{aligned}$$

Or $p < \frac{1}{2}$ donc $0 < \frac{p}{q} < 1$, et par ailleurs on a vu que $0 < 4pq < 1$, donc

$$\lim_{k \rightarrow +\infty} \left(\frac{p}{q}\right)^k \frac{(4pq)^k}{1 - 4pq} = 0$$

ainsi d'après le théorème des gendarmes, $\boxed{\lim_{k \rightarrow +\infty} a_{2k} = 0}$.

(c) Si $p < \frac{1}{2}$, alors d'après la question 10d,

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

On en déduit en particulier, puisque $\frac{p}{q} < 1$, $\lim_{k \rightarrow +\infty} a_{2k} = \alpha$, donc d'après la question précédente et par unicité de la limite, $\boxed{\alpha = 0}$.

Par ailleurs d'après la question 10a, $a_0 = 1$, ce qui donne $\beta \left(\frac{p}{q}\right)^0 = 1$, d'où $\boxed{\beta = 1}$.

Finalement,

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

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

13. (a) Soit $k \in \mathbb{N}$. $[T_n = k]$ correspond à l'événement "A a résolu k problèmes à l'instant où B résout son n -ème problème", c'est-à-dire "le $(n + k)$ -ème problème dans l'ordre de tous les problèmes résolus est résolu par B, et A en avait résolu k parmi les $n + k - 1$ qui précèdent", autrement dit " U_{n+k} vaut 0 et S_{n+k-1} vaut k ".

Ainsi on a bien $\boxed{[T_n = k] = [S_{n+k-1} = k] \cap [U_{n+k} = 0]}$.

- (b) Soit $k \in \mathbb{N}$. $(U_i)_{i \in \mathbb{N}^*}$ est une suite de variables aléatoires mutuellement indépendantes et $S_{n+k-1} = \sum_{i=1}^{n+k-1} U_i$, donc d'après le lemme des coalitions S_{n+k-1} et U_{n+k} sont indépendantes, ainsi

$$\begin{aligned} P(T_n = k) &= P(S_{n+k-1} = k)P(U_{n+k} = 0) = \binom{n+k-1}{k} p^k q^{n+k-1-k} \times q \\ &= \boxed{\binom{n+k-1}{k} p^k q^n} \end{aligned}$$

14. (a) $T_n(\Omega) = \mathbb{N}$ donc d'après la formule des probabilités totales appliquée au système complet d'événements $\{[T_n = k] : k \in \mathbb{N}\}$,

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

Pour tout entier $k \geq n+1$, $[T_n = k] \subset G_n$ (si à l'instant où B résout son n -ème problème, A en a déjà résolu $k \geq n+1$, alors A a gagné), donc $[T_n = k] \cap G_n = [T_n = k]$. Ainsi

$$\begin{aligned} P(G_n) &= \sum_{k=0}^n P([T_n = k] \cap G_n) + \sum_{k=n+1}^{+\infty} P(T_n = k) \\ &= \sum_{k=0}^n P_{[T_n=k]}(G_n)P(T_n = k) + P(T_n \geq n+1) \end{aligned}$$

Pour tout $k \in \llbracket 0, n \rrbracket$, l'événement $[T_n = k]$ signifie qu'à l'instant t le joueur A a un retard de $n-k$ problèmes par rapport à B. Ainsi, en raisonnant à partir de l'instant t , et sachant l'événement $[T_n = k]$, A gagnera si à un instant donné, il a résolu $n-k+1$ problèmes de plus que B. Ainsi $P_{[T_n=k]}(G_n) = P(A_{n-k+1})$.

Finalement on a :

$$P(G_n) = P(T_n \geq n+1) + \sum_{k=0}^n P(T_n = k)a_{n-k+1}$$

- (b) D'après les résultats de la partie II, si $p \geq \frac{1}{2}$ alors $a_r = 1$ pour tout $r \in \mathbb{N}$. Alors,

$$P(G_n) = P(T_n \geq n+1) + \sum_{k=0}^n P(T_n = k) = P(T_n \geq n+1) + P(T_n \leq n) = 1$$

car $\{[T_n \geq n+1], [T_n \leq n]\}$ forme un système complet d'événements. Ainsi on a bien $P(G_n) = 1$ si $p \geq \frac{1}{2}$.

- (c) Supposons $p < \frac{1}{2}$. Remarquons d'abord qu'on peut écrire :

$$P(T_n \geq n+1) = 1 - P(T_n \leq n) = 1 - \sum_{k=0}^n P(T_n = k)$$

donc

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

et en utilisant les résultats des questions 12c et 13b,

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

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

(a) Pour tout $n \in \mathbb{N}^*$, d'après la question précédente

$$\begin{aligned}
 P(G_n) &= 1 - q^n \sum_{k=0}^n \binom{n-k+1}{k} p^k + \frac{p}{q} p^n \sum_{k=0}^n \binom{n-k+1}{k} q^k \\
 &= 1 - \underbrace{(1-p)^n \sum_{k=0}^n \binom{n-k+1}{k} p^k}_{=u_n(p)} + \frac{p}{q} \times \underbrace{(1-q)^n \sum_{k=0}^n \binom{n-k+1}{k} q^k}_{=u_n(q)} \\
 &= \boxed{1 - u_n(p) + \frac{p}{q} u_n(q)}
 \end{aligned}$$

(b) Soient $x \in [0, 1]$ et $x \in \mathbb{N}^*$. Dans les calculs qui suivent on utilise la convention $\binom{j}{i} = 0$ si $i \leq 0$ ou $i > j$.

En écrivant $(1-x)^{n+1} = (1-x)^n - x(1-x)^n$,

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

D'après la formule de Pascal, pour tout $k \in \llbracket 0, n+1 \rrbracket$, on a

$$\binom{n+k}{k} = \binom{n+k-1}{k-1} + \binom{n+k-1}{k},$$

donc

$$\begin{aligned}
 u_{n+1}(x) &= (1-x)^n \sum_{k=0}^{n+1} \binom{n+k-1}{k-1} x^k + (1-x)^n \sum_{k=0}^{n+1} \binom{n+k-1}{k} x^k - x(1-x)^n \sum_{k=0}^n \binom{n+k}{k} x^k \\
 &\quad - (1-x)^n x^{n+2} \binom{2n+1}{n+1} \\
 &= (1-x)^n \sum_{k=0}^{n+1} \binom{n+k-1}{k-1} x^k + \underbrace{(1-x)^n \sum_{k=0}^n \binom{n+k-1}{k} x^k}_{=u_n(x)} + (1-x)^n x^{n+1} \binom{2n}{n+1} \\
 &\quad - (1-x)^n \sum_{k=0}^n \binom{n+k}{k} x^{k+1} - (1-x)^n x^{n+2} \binom{2n+1}{n+1}
 \end{aligned}$$

On remarque que par changement d'indice :

$$\begin{aligned}
 &(1-x)^n \sum_{k=0}^{n+1} \binom{n+k-1}{k-1} x^k - (1-x)^n \sum_{k=0}^n \binom{n+k}{k} x^{k+1} \\
 &= (1-x)^n \sum_{k=1}^{n+1} \binom{n+k-1}{k-1} x^k - (1-x)^n \sum_{k=1}^{n+1} \binom{n+k-1}{k-1} x^k \\
 &= 0
 \end{aligned}$$

Ainsi

$$\begin{aligned}
 u_{n+1}(x) &= u_n(x) + (1-x)^n x^{n+1} \binom{2n}{n+1} - (1-x)^n x^{n+2} \binom{2n+1}{n+1} \\
 &= \boxed{u_n(x) + (1-x)^n x^{n+1} \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} x \right)}
 \end{aligned}$$

(c) Pour tout $n \in \mathbb{N}^*$, d'après la question 15a puis la question précédente

$$\begin{aligned}
 P(G_{n+1}) &= 1 - u_{n+1}(p) + \frac{p}{q}u_{n+1}(q) \\
 &= 1 - u_n(p) - (1-p)^n p^{n+1} \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} p \right) \\
 &\quad + \frac{p}{q} \left(u_n(q) + (1-q)^n q^{n+1} \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} q \right) \right) \\
 &= 1 - u_n(p) + \underbrace{\frac{p}{q}u_n(q)}_{=P(G_n)} - q^n p^{n+1} \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} p \right) \\
 &\quad + p^{n+1} q^n \left(\binom{2n}{n+1} - \binom{2n+1}{n+1} q \right) \\
 &= P(G_n) - p^{n+1} q^n \left(\binom{2n+1}{n+1} q - \binom{2n+1}{n+1} p \right) \\
 &= \boxed{P(G_n) - (pq)^{n+1} \left(1 - \frac{p}{q} \right) \binom{2n+1}{n+1}}
 \end{aligned}$$

(d) Pour tout $n \in \mathbb{N}^*$, on considère l'identité

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

Pour $n = 1$ on a d'après la question 14c,

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

ainsi (E_1) est vérifiée.

Soit $n \in \mathbb{N}^*$. Supposons (E_n) vérifiée. Alors d'après la question précédente et l'hypothèse de récurrence,

$$\begin{aligned}
 P(G_{n+1}) &= P(G_n) - (pq)^{n+1} \left(1 - \frac{p}{q} \right) \binom{2n+1}{n+1} \\
 &= \frac{p}{q} - \left(1 - \frac{p}{q} \right) \sum_{k=1}^n \binom{2k-1}{k} (pq)^k - \left(1 - \frac{p}{q} \right) \binom{2n+1}{n+1} (pq)^{n+1} \\
 &= \boxed{\frac{p}{q} - \left(1 - \frac{p}{q} \right) \sum_{k=1}^{n+1} \binom{2k-1}{k} (pq)^k}
 \end{aligned}$$

et l'identité (E_{n+1}) est vérifiée.

Ainsi on a montré par récurrence que

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

16. Application à la sécurisation des transactions

(a) Soit $n \in \mathbb{N}$ et $k \in \llbracket 0, n \rrbracket$.

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

En itérant cette relation :

$$\binom{n}{k} = \frac{n}{k} \binom{n-1}{k-1} = \frac{n}{k} \frac{n-1}{k-1} \binom{n-2}{k-2} = \dots = \frac{n}{k} \frac{n-1}{k-1} \dots \frac{n-k+1}{1} \binom{n-k}{0} = \frac{n}{k} \frac{n-1}{k-1} \dots \frac{n-k+1}{1}$$

On peut ainsi définir une fonction à l'aide d'une boucle **for** :

```
1 def binom(k,n):
    c = 1
    for i in range(0,k):
        c = (n-i)/(i+1) * c
5     return c
```

- (b) On utilise la relation de la question 15c pour calculer les termes successifs de la suite $(P(G_n))_{n \in \mathbb{N}^*}$.

```
1 p = float(input("p = ?"))
  eps = float(input("epsilon = ?"))
  q = 1 - p
  n = 1
5 PGn = p/q - (1-p/q)*p*q

  while PGn > eps:
      PGn = PGn - (p*q)**(n+1)*(1-p/q)*binom(n+1,2*n+1)
      n = n+1
10 print(n)
```