MATHÉMATIQUES III - HEC E 2008

Proposition de corrigé par David Meneu

Lycée Champollion - Grenoble, pour Major-Prépa

EXERCICE

Étant donné un entier n supérieur ou égal à 2, on considère un nuage de n points du plan, c'est-à-dire un n-uplet $((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n))$ d'éléments de $\mathbb{R}^2$. On suppose que les réels $x_1, x_2, \ldots, x_n$ (resp. $(y_1, y_2, \ldots, y_n)$ ne sont pas tous égaux.

On appelle *moyenne arithmétique* $\overline{x}$ et *écart-type* σ_x du n-uplet $x = (x_1, \ldots, x_n)$, les réels suivants :

$$\overline{x} = \frac{1}{n}\sum_{k=1}^{n} x_k \quad \text{ et } \sigma_x = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2}$$

On définit de même la moyenne arithmétique $\overline{y}$ et l'écart-type σ_y du n-uplet $y = (y_1, y_2, \ldots, y_n)$.
La covariance $\operatorname{cov}(x, y)$ et le coefficient de corrélation linéaire $r(x, y)$ du couple (x, y) sont donnés par :

$$\operatorname{cov}(x, y) = \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})(y_k - \overline{y}) \quad \text{ et } \quad r(x, y) = \frac{\operatorname{cov}(x, y)}{\sigma_x \times \sigma_y}$$

Soit f la fonction définie sur $\mathbb{R}^2$ à valeurs réelles qui, à tout couple (a, b) de $\mathbb{R}^2$, associe le réel $f(a, b)$ tel que :

$$f(a, b) = \sum_{k=1}^{n}(ax_k + b - y_k)^2$$

1. La fonction f est de classe C^2 sur $\mathbb{R}^2$ car c'est une fonction polynômiale en les deux variables a et b.
2. a) Les points critiques de la fonction f de classe C^2 sur $\mathbb{R}^2$, sont les solutions du système :

$$(S) : \begin{cases} \partial_1(f)(a, b) & = 0 \\ \partial_2(f)(a, b) & = 0 \end{cases} \iff \begin{cases} \displaystyle\sum_{k=1}^{n} 2x_k.(ax_k + b - y_k) & = 0 \\ \displaystyle\sum_{k=1}^{n} 2(ax_k + b - y_k) & = 0 \end{cases} \iff \begin{cases} \displaystyle\sum_{k=1}^{n} x_k.(ax_k + b - y_k) & = 0 \\ \displaystyle\sum_{k=1}^{n}(ax_k + b - y_k) & = 0 \end{cases}$$

b) On résout le système (S) d'inconnues a, b en commençant par séparer a et b dans les deux équations, par linéarité de la somme :

$$(S) \iff \begin{cases} a.\displaystyle\sum_{k=1}^{n} x_k^2 + b\sum_{k=1}^{n} x_k & = \displaystyle\sum_{k=1}^{n} x_k.y_k \\ a.\displaystyle\sum_{k=1}^{n} x_k + b.n & = \displaystyle\sum_{k=1}^{n} y_k \end{cases} \quad \text{puis on réalise : } L_1 \leftarrow n.L_1 - \left(\sum_{k=1}^{n} x_k\right).L_2$$

$$\iff \begin{cases} a.\left[n.\displaystyle\sum_{k=1}^{n} x_k^2 - \left(\sum_{k=1}^{n} x_k\right)^2\right] & = n\displaystyle\sum_{k=1}^{n} x_k.y_k - \left(\sum_{k=1}^{n} x_k\right).\left(\sum_{k=1}^{n} y_k\right) \\ a.\displaystyle\sum_{k=1}^{n} x_k + b.n & = \displaystyle\sum_{k=1}^{n} y_k \end{cases}$$

Il faut maintenant introduire les notations de l'énoncé pour simplifier l'écriture du système. On remarque que :

$\sum_{k=1}^{n} x_k = n.\overline{x}$, puis $n.\sigma_x^2 = \sum_{k=1}^{n} x_k^2 - 2x_k.\overline{x} + \overline{x}^2 = \sum_{k=1}^{n} x_k^2 - 2\overline{x}.\sum_{k=1}^{n} + n\overline{x}^2 = \sum_{k=1}^{n} x_k^2 - n.\overline{x}^2$, et enfin

$n.\text{cov}(x,y) = \sum_{k=1}^{n} x_k.y_k - \overline{x}.y_k - x_k.\overline{y_k} + \overline{x}.\overline{y} = \sum_{k=1}^{n} x_k.y_k - \overline{x}\sum_{k=1}^{n} y_k - \overline{y}.\sum_{k=1}^{n} x_k + n\overline{x}.\overline{y} = \sum_{k=1}^{n} x_k.y_k - n.\overline{x}.\overline{y}$,

donc : $(S) \iff \begin{cases} a.n.\sigma_x^2 & = n.\text{cov}(x,y) \\ a.n.\overline{x} + b.n & = n.\overline{y} \end{cases} \iff \begin{cases} a & = \dfrac{\text{cov}(x,y)}{\sigma_x^2} \\ b & = \overline{y} - \overline{x}.a = \overline{y} - \overline{x}.\dfrac{\text{cov}(x,y)}{\sigma_x^2} \end{cases}$

On a bien obtenu un unique couple solution $(\widehat{a}, \widehat{b})$.

c) La hessienne de f au point critique $(\widehat{a}, \widehat{b})$ est :

$$H_{(\widehat{a},\widehat{b})} = \begin{pmatrix} \partial^2_{(1,1)}(f)(\widehat{a},\widehat{b}) & \partial^2_{(1,2)}(f)(\widehat{a},\widehat{b}) \\ \partial^2_{(2,1)}(f)(\widehat{a},\widehat{b}) & \partial^2_{(2,2)}(f)(\widehat{a},\widehat{b}) \end{pmatrix} = \begin{pmatrix} 2\sum_{k=1}^{n} x_k^2 & 2\sum_{k=1}^{n} x_k \\ 2\sum_{k=1}^{n} x_k & 2n \end{pmatrix} = 2n.\begin{pmatrix} (\sigma_x^2 + \overline{x}^2) & \overline{x} \\ \overline{x} & 1 \end{pmatrix}.$$

$2n > 0$ donc les valeurs propres de $H_{(\widehat{a},\widehat{b})}$ sont du même signe que celles de $M = \begin{pmatrix} (\sigma_x^2 + \overline{x}^2) & \overline{x} \\ \overline{x} & 1 \end{pmatrix}$, où :

λ est valeur propre de $M \iff M - \lambda.I_2 = \begin{pmatrix} \sigma_x^2 + \overline{x}^2 - \lambda & \overline{x} \\ \overline{x} & 1-\lambda \end{pmatrix}$ est non-inversible

$\iff (\sigma_x^2 + \overline{x}^2 - \lambda)(1-\lambda) - \overline{x}^2 = 0$ d'après le critère d'inversibilité des matrices 2×2.
Les valeurs propres de M sont donc les solutions de l'équation : $\lambda^2 - \lambda.(\sigma_x^2 + \overline{x}^2 + 1) + \sigma_x^2 = 0$.
M étant symétrique réelle, on sait qu'elle est diagonalisable, et admet donc deux valeurs propres distinctes ou confondues λ_1 et λ_2.

L'équation précédente est donc équivalente à : $(\lambda - \lambda_1)(\lambda - \lambda_2) = 0 \iff \lambda^2 - (\lambda_1 + \lambda_2).\lambda + \lambda_1.\lambda_2 = 0$, donc par identification des coefficients :
$\lambda_1.\lambda_2 = \sigma_x^2 > 0$: les deux valeurs propres sont de même signe, et : $\lambda_1 + \lambda_2 = \sigma_x^2 + \overline{x}^2 + 1 > 0$: les deux valeurs propres de M, donc de $H_{(\widehat{a},\widehat{b})}$ sont donc toutes deux strictement positives.

On en conclut que f admet un minimum local en $(\widehat{a}, \widehat{b})$.

d) Calcul de $f(\widehat{a}, \widehat{b})$ à l'aide des relations établies en 2.b) :

$$\begin{aligned} f(\widehat{a},\widehat{b}) &= \sum_{k=1}^{n} (\widehat{a}.x_k + \overline{y} - \overline{x}.\widehat{a} - y_k)^2 = \sum_{k=1}^{n} (\widehat{a}(x_k - \overline{x}) - (y_k - \overline{y}))^2 \\ &= \widehat{a}^2 \sum_{k=1}^{n} (x_k - \overline{x})^2 - 2\widehat{a}\sum_{k=1}^{n} (x_k - \overline{x})(y_k - \overline{y}) + \sum_{k=1}^{n} (y_k - \overline{y})^2 \\ &= \frac{[\text{cov}(x,y)]^2}{\sigma_x^4}.n.\sigma_x^2 - 2\frac{\text{cov}(x,y)}{\sigma_x^2}.n.\text{cov}(x,y) + n.\sigma_y^2 \\ &= n\sigma_y^2 - n.\frac{[\text{cov}(x,y)]^2}{\sigma_x^2} = n.\sigma_y^2.\left[1 - \frac{[\text{cov}(x,y)]^2}{\sigma_x^2 \times \sigma_y^2}\right] = n.\sigma_y^2.[1 - r^2(x,y)] \quad \text{CQFD} \end{aligned}$$

a) En remarquant que, pour tout $(a,b) \in \mathbb{R}^2$, $f(a,b)$ est une somme de carrés de nombres réels, on a bien sûr : $\forall (a,b) \in \mathbb{R}^2,\ f(a,b) \geqslant 0$, et en particulier :
$f(\widehat{a},\widehat{b}) \geqslant 0 \iff n.\sigma_y^2.(1 - r^2(x,y)) \geqslant 0 \overset{n.\sigma_y^2>0}{\iff} 1 - r^2(x,y) \geqslant 0 \iff r^2(x,y) \leqslant 1 \iff |r(x,y)| \leqslant 1$ (par stricte croissance de la fonction racine carrée sur $\mathbb{R}^+$).

b) Si $|r(x,y)| = 1$, alors $f(\widehat{a}, \widehat{(b)}) = 0 \iff \sum_{k=1}^{n} (\widehat{a}w_k + b - y_k)^2 = 0 \iff \forall k \in [\![1,n]\!],\ \widehat{a}x_k + \widehat{b} - y_k = 0$ car une somme de réels positifs est nulle, si et seulement si chacun des termes de la somme est nul.
Mais alors : $\forall k \in [\![1,n]\!],\ y_k = \widehat{a}.x_k + \widehat{b}$, c'est-à-dire que tous les points du nuage appartiennent à la droite D d'équation : $y = \widehat{a}.x + \widehat{b}$, et sont donc alignés !

Remarque : dans le cas général, $f(a,b) = \sum_{k=1}^{n} (ax_k + b - y_k)^2$ représente la somme des carrés des distances (verticales) de chacun des points du nuage, à la droite d'équation $y = ax + b$. Le résultat obtenu est que cette valeur est minimale lorsqu'on choisit les paramères $(\widehat{a}, \widehat{b})$: la droite D s'appelle *droite de régression linéaire au sens des moindres carrés*, c'est la droite qui "approche au mieux" le nuage de points.

PROBLÈME

Dans tout le problème, N désigne un entier naturel fixé supérieur ou égal à 2, et p un réel fixé de l'intervalle $]0,1[$. On pose : $q = 1 - p$. Soit n un entier naturel quelconque.

Dans une population de N individus, on s'intéresse à la propagation d'un certain virus. Chaque jour, on distingue dans cette population trois catégories d'individus : en premier lieu, les individus sains, c'est-à-dire ceux qui ne sont pas porteurs du virus, ensuite les individus qui viennent d'être contaminés et qui sont inoffensifs pour les autres, et enfin, les individus contaminés par le virus et qui sont contagieux.

Ces trois catégories évoluent jour après jour selon le modèle suivant :

- chaque jour n, chaque individu sain peut être contaminé par n'importe lequel des individus contagieux ce jour avec la même probabilité p, ces contaminations éventuelles étant indépendantes les unes des autres ;
- un individu contaminé le jour n devient contagieux le jour $n+1$;
- chaque individu contagieux le jour n redevient sain le jour $n+1$.

On note alors X_n le nombre aléatoire d'individus contagieux le jour n.

On remarquera que si, pour un certain entier naturel i, on a $X_i = 0$, alors on a aussi $X_{i+1} = 0$.

Les variables aléatoires $X_0, X_1, \ldots, X_n, \ldots$ sont supposées définies sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$ et $E(X_n)$ désigne, pour tout n de $\mathbb{N}$, l'espérance de X_n.

Partie I. Un cas particulier

Dans cette partie uniquement, on suppose que l'on a : $N = 3$ et $p = 1/3$.

On considère les matrices S et R suivantes :

$$S = \begin{pmatrix} 9 & 4 & 4 & 9 \\ 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 0 & 1 & -6 & 1 \\ 1 & 0 & 5 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

L'ensemble $\mathcal{M}_{4,1}(\mathbb{R})$ des matrices colonnes à quatre lignes est confondu avec l'espace vectoriel $\mathbb{R}^4$.

1. La mise en oeuvre de la méthode de Gauss permet de prouver que la matrice R est bien inversible, d'inverse
$$R^{-1} = \begin{pmatrix} 0 & 1/6 & -5/6 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1/6 & 1/6 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

2. a) Un réel λ est valeur propre de S si et seulement si $S - \lambda.I_4$ est non-inversible.

$$S - \lambda.I_4 = \begin{pmatrix} 9-\lambda & 4 & 4 & 9 \\ 0 & 4-\lambda & 5 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix} \xrightarrow{L_2 \leftrightarrow L_3} \begin{pmatrix} 9-\lambda & 4 & 4 & 9 \\ 0 & 1 & -\lambda & 0 \\ 0 & 4-\lambda & 5 & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix}$$

$$\xrightarrow{L_3 \leftarrow L_3 - (4-\lambda)L_2} \begin{pmatrix} 9-\lambda & 4 & 4 & 9 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 5+4\lambda-\lambda^2 & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix}$$; la matrice $S - \lambda.I_4$ est maintenant échelonnée, donc non-inversible si et seulement si l'un des coefficients diagonaux de la réduite de Gauss est nul, soit :

$$\begin{cases} & 9-\lambda = 0 \\ \text{ou} & 5+4\lambda-\lambda^2 = 0 \\ \text{ou} & -\lambda = 0 \end{cases}$$. Le trinôme $-\lambda^2+4\lambda+5$ a bien pour racines -1 et 5, donc : $\mathrm{Sp}(S) = \{-1, 0, 5, 9\}$.

b) Notons D le produit matriciel $R^{-1}SR$:

$$D = \begin{pmatrix} 0 & 1/6 & -5/6 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1/6 & 1/6 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 9 & 4 & 4 & 9 \\ 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & -6 & 1 \\ 1 & 0 & 5 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1/6 & -5/6 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1/6 & 1/6 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -30 & 9 \\ -1 & 0 & 25 & 0 \\ 1 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

On remarque donc que D est une matrice diagonale, dont les éléments diagonaux sont les valeurs propres de S. L'énoncé nous a donc donné une matrice de passage R vers une base de vecteurs propres de S, qui est bien diagonalisable puisque c'est une matrice de $\mathcal{M}_4(\mathbb{R})$ admettant 4 valeurs propres distinctes.

c) $D = R^{-1}SR \iff S = RDR^{-1}$. Si $n = 0$, évidemment $S^0 = I$, sinon on sait qu'on est dans le cas où l'on démontre par récurrence que :
$\forall n \in \mathbb{N}^*,\ S^n = RD^nR^{-1}$.

$\boxed{\text{I.}}$ C'est vrai pour $n = 1$: $S = RDR^{-1}$.

$\boxed{\text{H.}}$ supposons la propriété vraie pour un certain $n \in \mathbb{N}^*$: au rang suivant,
$S^{n+1} = S^n.S \overset{H.R.}{=} (RD^nR^{-1})(RDR^{-1}) = RD^n(RR^{-1})DR^{-1} = RD^nIRD^{-1} = RD^{n+1}R^{-1}$.

$\boxed{\text{C.}}$ La propriété est donc héréditaire ; comme elle est initialisée pour $n = 1$, le théorème de récurrence assure donc qu'elle est vraie pour tout $n \in \mathbb{N}^*$.

Pour tout $n \in \mathbb{N}^*$, puisque D est une matrice diagonale :

$D^n = \begin{pmatrix} (-1)^n & 0 & 0 & 0 \\ 0 & 0^n & 0 & 0 \\ 0 & 0 & 5^n & 0 \\ 0 & 0 & 0 & 9^n \end{pmatrix}$ (attention, on pourra écrire $0^n = 0$ pour $n \in \mathbb{N}^*$; par contre $0^0 = 1$).

Le produit matriciel $S^n = RD^nR^{-1}$ donne, tous calculs faits :

$$\boxed{\forall n \in \mathbb{N}^*,\ S^n = \begin{pmatrix} 9^n & 9^n - 5^n & 9^n - 5^n & 9^n \\ 0 & \dfrac{5^{n+1} + (-1)^n}{6} & \dfrac{5(-1)^{n+1} + 5^{n+1}}{6} & 0 \\ 0 & \dfrac{5^n + (-1)^{n+1}}{6} & \dfrac{5(-1)^n + 5^n}{6} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}.$$

Remarque : la formule n'est donc pas valide pour $n = 0$.

3. Soit $n \in \mathbb{N}$ fixé.

a) Sachant qu'à l'instant n il n'y a aucune personne contagieuse : personne ne peut donc être contaminé à l'instant n, et il n'y aura donc aucun contaminé à l'instant $n+1$, soit :
$\boxed{P_{X_n=0}(X_{n+1} = 0) = 1 \text{ et } P_{X_n=0}(X_{n+1} = 1) = P_{X_n=0}(X_{n+1} = 2) = P_{X_n=0}(X_{n+1} = 3) = 0}$.

b) Sachant qu'à l'instant n les trois personnes sont toutes contagieuses : elles ne peuvent pas être contaminées (n'étant pas saines), et redeviennent toutes saines à l'instant $n+1$, donc :
$\boxed{P_{X_n=3}(X_{n+1} = 0) = 1 \text{ et } P_{X_n=3}(X_{n+1} = 1) = P_{X_n=3}(X_{n+1} = 2) = P_{X_n=3}(X_{n+1} = 3) = 0}$.

c) Sachant qu'à l'instant n il y a deux personnes contagieuses (qui redeviendront saines à l'instant $n+1$) : la personne restante est soit contaminée à l'instant n, donc contagieuses à l'instant $n+1$, soit restée saine à l'instant n, donc n'est pas contagieuse à l'instant $n+1$.
Il y a donc au plus une personne contagieuse à l'instant $n+1$, et en particulier :
$\boxed{P_{X_n=2}(X_{n+1} = 2) = P_{X_n=2}(X_{n+1} = 3) = 0}$
Sachant $(X_n = 2)$: $(X_{n+1} = 0)$ est réalisé si et seulement si les deux personnes contagieuses n'ont pas contaminé la troisième personne à l'instant n, ce qui arrive avec la probabilité $\dfrac{2}{3}$ à chaque fois, et indépendamment l'une de l'autre, d'où : $\boxed{P_{X_n=2}(X_{n+1} = 0) = \left(\dfrac{2}{3}\right)^2 = \dfrac{4}{9}}$. Enfin, comme $((X_n = i))_{i \in [\![0,3]\!]}$ forment un système complet d'événements : $\boxed{P_{X_n=2}(X_{n+1} = 1) = 1 - \dfrac{4}{9} - 0 - 0 = \dfrac{5}{9}}$.

La loi conditionnelle de X_{n+1} sachant $[X_n = 2]$ est la loi de Bernoulli de paramètre $\frac{5}{9}$, donc la loi binomiale $\mathcal{B}(1, \frac{5}{9})$.

Sachant qu'à l'instant n il y a une personne contagieuses, qui redeviendra saine à l'instant $n+1$:
Il y aura donc , au maximum, 2 personne contagieuses à l'instant $n+1$ (contaminées à l'instant n).
Donc : $\boxed{P_{X_n=1}(X_{n+1} = 3) = 0}$.
Sachant $(X_n = 1)$: $(X_{n+1} = 2)$ est réalisé si et seulement si les deux autres personnes ont été contaminées à l'instant n par la seule personne contagieuse, avec la probabilité $\frac{1}{3}$ à chaque fois.

Les contaminations étant indépendantes : $\boxed{P_{X_n=1}(X_{n+1} = 2) = \left(\frac{1}{3}\right)^2 = \frac{1}{9} = \binom{2}{2}(1/3)^2(2/3)^{2-2}}$.

Toujours sachant $(X_n = 1)$: $(X_{n+1} = 0)$ est réalisé si et seulement si la seule personne contagieuse a "échoué" à contaminer les deux autres, et ce avec la probabilité $\frac{2}{3}$ à chaque fois.

Ainsi : $\boxed{P_{X_n=1}(X_{n+1} = 0) = \left(\frac{2}{3}\right)^2 = \frac{4}{9} = \binom{2}{0}(1/3)^0(2/3)^{2-0}}$.

Enfin : $\sum_{i=0}^{3} P_{X_n=1}(X_{n+1} = i) = 1$, donc : $\boxed{P_{X_n=1}(X_{n+1} = 1) = 1 - \frac{4}{9} - \frac{1}{9} = \frac{4}{9} = \binom{2}{1}(1/3)^1(2/3)^{2-1}}$.

La loi conditionnelle de X_{n+1} sachant $[X_n = 1]$ est donc bien la loi binomiale de paramètres $(2, \frac{1}{3})$.

d) On note $E(X_{n+1}/X_n = i)$ l'espérance de la loi de probabilité conditionnelle de X_{n+1} sachant l'événement $[X_n = i]$.
$E(X_{n+1}/X_n = 1)$ est donc l'espérance d'une loi de binomiale de paramètres $(2, \frac{1}{3})$, donc :

$$E(X_{n+1}/X_n = 1) = 2.\frac{1}{3} = \frac{2}{3}\ .$$

De même, $E(X_{n+1}/X_n = 2)$ est l'espérance d'une loi de Bernoulli de paramètre $\frac{5}{9}$, donc :

$$E(X_{n+1}/X_n = 2) = \frac{5}{9}.$$

Remarque : selon ce principe, on a aussi : $E(X_{n+1}/X_n = 0) = 0 = E(X_{n+1}/X_n = 3)$ comme espérance d'une v.a.r. certaine égale à 0 sachant que $[X_n = 0]$ ou $[X_n = 3]$ est réalisé.

4. On suppose, *uniquement dans cette question*, que X_0 suit la loi binomiale de paramètres $(3, \frac{1}{3})$:
$\forall i \in [\![0, 3]\!],\ P(X_0 = i) = \binom{3}{i}\left(\frac{1}{3}\right)^i\left(\frac{2}{3}\right)^{3-i}$.

a) D'après la formule des probabilités totales, appliquée avec le système complet d'événements $((X_0 = i))_{i\in[\![0,3]\!]}$:

$P(X_1 = 0) = \sum_{i=0}^{3} P_{X_0=i}(X_1 = 0).P(X_0 = i)$
$= 1.P(X_0 = 0) + \frac{4}{9}P(X_0 = 1) + \frac{4}{9}P(X_0 = 2) + 1.P(X_0 = 3)$
$= \left(\frac{2}{3}\right)^3 + \frac{4}{9}.3.\frac{1}{3}.\left(\frac{2}{3}\right)^2 + \frac{4}{9}.3.\left(\frac{1}{3}\right)^2.\frac{2}{3} + \left(\frac{1}{3}\right)^3 = \frac{3.8 + 16 + 8 + 3}{81} = \boxed{\frac{51}{81}}$.

$P(X_1 = 1) = \sum_{i=0}^{3} P_{X_0=i}(X_1 = 1)P(X_0 = i)$
$= 0.P(X_0 = 0) + \frac{4}{9}P(X_0 = 1) + \frac{5}{9}P(X_0 = 2) + 0.P(X_0 = 3)$
$= \frac{4}{9}.3.\frac{1}{3}.\left(\frac{2}{3}\right)^2 + \frac{5}{9}.3.\left(\frac{1}{3}\right)^2.\frac{2}{3} = \frac{16 + 10}{81} = \boxed{\frac{26}{81}}$

$P(X_1 = 2) = \sum_{i=0}^{3} P_{X_0=i}(X_1 = 2)P(X_0 = i)$
$= 0.P(X_0 = 0) + \frac{1}{9}.P(X_0 = 1) + 0.P(X_0 = 2) + 0.P(X_0 = 3) = \frac{1}{9}.3.\frac{1}{3}.\left(\frac{2}{3}\right)^2 = \boxed{\frac{4}{81}}$.

$P(X_1=3)=\sum_{i=0}^{3} P_{X_0=i}(X_1=3)P(X_0=i)=0$: en fait, $X_1(\Omega)=\{0,1,2\}$.

$E(X_1)=0.P(X_1=0)+1.P(X_1=1)+2.P(X_1=2)=\frac{26}{81}+2.\frac{4}{81}=\boxed{\frac{34}{81}}$.

b) Le calcul de somme donne, en utilisant les résultats de 3.d) :

$$\begin{aligned}\sum_{i=0}^{3} E(X_1/X_0=i).P(X_0=i)&=0.P(X_0=0)+\frac{2}{3}.P(X_0=1)+\frac{5}{9}.P(X_0=2)+0.P(X_0=3)\\&=\frac{2}{3}.3.\frac{1}{3}.\left(\frac{2}{3}\right)^2+\frac{5}{9}.3.\left(\frac{1}{3}\right)^2.\frac{2}{3}=\frac{24}{81}+\frac{10}{81}\\&=\frac{34}{81}=E(X_1)\end{aligned}$$

On a vérifié sur cet exemple la *formule de l'espérance totale* (hors-programme en ECE2).

5. Pour tout entier naturel, on considère le vecteur U_n de $\mathbb{R}^4$ défini par : $U_n=\begin{pmatrix}u_n\\v_n\\w_n\\t_n\end{pmatrix}=\begin{pmatrix}P([X_n=0])\\P([X_n=1])\\P([X_n=2])\\P([X_n=3])\end{pmatrix}$.

a) Pour tout $n\in\mathbb{N}$, $((X_n=i))_{i\in[\![0,3]\!]}$ est un système complet d'événements, et par conséquent : $\sum_{i=0}^{3} P(X_n=i)=1 \iff u_n+v_n+w_n+t_n=1$.

b) Pour tout $n\in\mathbb{N}$, avec le système complet d'événements ci-dessus, on écrit la formule des probabilités totales pour exprimer la loi de X_{n+1} :

$P(X_{n+1}=0)=\sum_{i=0}^{3} P_{X_n=i}(X_{n+1}=0)P(X_n=i)$
$=1.P(X_n=0)+\frac{4}{9}.P(X_n=1)+\frac{4}{9}.P(X_n=2)+1.P(X_n=3)$

$P(X_{n+1}=1)=\sum_{i=0}^{3} P_{X_n=i}(X_{n+1}=1)P(X_n=i)$
$=0.P(X_n=0)+\frac{4}{9}.P(X_n=1)+\frac{5}{9}.P(X_n=2)+0.P(X_n=3)$

$P(X_{n+1}=2)=\sum_{i=0}^{3} P_{X_n=i}(X_{n+1}=3)P(X_n=i)$
$=0.P(X_n=0)+\frac{1}{9}.P(X_n=1)+0.P(X_n=2)+0.P(X_n=3)$

$P(X_{n+1}=3)=\sum_{i=0}^{3} P_{X_n=i}(X_{n+1}=3)P(X_n=i)$
$=0.P(X_n=0)+0.P(X_n=1)+0.P(X_n=2)+0.P(X_n=3)$,
soit :

$$\forall n\in\mathbb{N},\ U_{n+1}=MU_n \text{ avec } M=\begin{pmatrix}1&4/9&4/9&1\\0&4/9&5/9&0\\0&1/9&0&0\\0&0&0&0\end{pmatrix}.$$

La matrice M est bien indépendante de n.

c) Clairement : $\boxed{M=\frac{1}{9}.S}$. On peut donc déduire les valeurs propres de M à partir de celles de S calculées précédemment, en effet :
soit λ une valeur propre de M, $X\neq 0$ un vecteur propre associé. L'égalité $MX=\lambda.X$ est équivalente à : $\frac{1}{9}SX=\lambda.X \iff SX=9\lambda.X$, c'est-à-dire que : λ est valeur propre de M si et seulement si 9λ est valeur propre de S. On en déduit que les valeurs propres de M sont : $\boxed{-\frac{1}{9},0,\frac{5}{9},1}$.

d) On sait alors qu'on doit prouver, à l'aide d'une simple récurrence, que :

$\forall n \in \mathbb{N}^*,\ U_n = M^n.U_0$, où : $M^n = \left(\frac{1}{9}.S\right)^n = \frac{1}{9^n}.S^n$, soit :

$\forall n \in \mathbb{N}^*$,

$$U_n = \begin{pmatrix} u_n \\ v_n \\ w_n \\ t_n \end{pmatrix} = \frac{1}{9^n}\begin{pmatrix} 9^n & 9^n - 5^n & 9^n - 5^n & 9^n \\ 0 & \dfrac{5^{n+1} + (-1)^n}{6} & \dfrac{5(-1)^{n+1} + 5^{n+1}}{6} & 0 \\ 0 & \dfrac{5^n + (-1)^{n+1}}{6} & \dfrac{5(-1)^n + 5^n}{6} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} u_0 \\ v_0 \\ w_0 \\ t_0 \end{pmatrix},$$

donc en particulier : $\forall n \in \mathbb{N}^*$,

$u_n = \frac{1}{9^n}.(9^n.u_0 + (9^n - 5^n).v_0 + (9^n - 5^n).w_0 + 9^n.t_0)$

$= u_0 + \left(1 - \left(\frac{5}{9}\right)^n\right).(v_0 + w_0) + t_0$, soit

$\boxed{\forall n \in \mathbb{N}^*,\ u_n = 1 - \left(\frac{5}{9}\right)^n (v_0 + w_0)}$

puisque $u_0 + t_0 = 1 - (v_0 + w_0)$. On a aussi : $\forall n \in \mathbb{N}^*$,

$\boxed{v_n = \frac{1}{9^n}.\left(\frac{5^{n+1} + (-1)^n}{6}.v_0 + \frac{5.(-1)^{n+1} + 5^{n+1}}{6}.w_0\right)}$.

6. On pose $F = \bigcup_{n=0}^{+\infty} [X_n = 0]$

a) L'événement F est réalisé si et seuement si l'un des événements $[X_n = 0]$ est réalisé, donc si et seulement si il existe un jour n auquel le virus a disparu.

b) Comme on l'a vu, et comme l'énoncé le précise (!) : si à un jour donné n, le virus a disparu ($[X_n = 0]$ réalisé), alors le virus est encore absent au jour suivant.
C'est-à-dire que : $\forall n \in \mathbb{N},\ [X_n = 0] \subset [X_{n+1} = 0]$, et la suite $([X_n = 0])_{n\in\mathbb{N}}$ est une suite croissante d'événements.
La propriété de limite monotone permet d'écrire : $P\left(\bigcup_{n=0}^{+\infty} [X_n = 0]\right) = \lim_{n\to+\infty} P(X_n = 0) = \lim_{n\to+\infty} u_n = 1$ puisque $0 < \frac{5}{9} < 1$. On peut donc dire que le virus finit *presque sûrement* par disparaître, et ce quelle que soit la loi initiale de la v.a.r. X_0.

Partie II. Le cas général

On suppose que pour tout entier naturel n et pour tout entier i de $[\![0, N]\!]$, on a : $P([X_n = i]) > 0$. On suppose également que pour tout couple (i, j) de $[\![0, N]\!]^2$, le réel $q_{i,j}$ défini par : $q_{i,j} = P_{[X_n=i]}([X_{n+1} = j])$, est indépendant de n.
Soit Q la matrice de $\mathcal{M}_{N+1}(\mathbb{R})$ définie par : $Q = (q_{i,j})_{0\leqslant i,j\leqslant N}$.

1. a) Sachant qu'il n'y a plus de personnes contaminées à l'instant n, le virus ne peut plus se propager, donc : $\boxed{\forall j \in [\![0, N]\!],\ q_{0,j} = 0}$.
De même : sachant que les N personnes sont contaminées à l'instant n, elles guérissent toutes à l'instant suivant et personne ne peut être à nouveau contaminé à l'instant suivant : $\boxed{\forall j \in [\![0, N]\!],\ q_{N,j} = 0}$.
Enfin : qu'il n'y ait plus aucun contaminé à l'instant n, ou qu'il y en ait $i \in [\![1, N]\!]$, il est impossible que la population toute entière soit contaminée à l'instant suivant : les contaminés du jour précédent sont guéris ! Par conséquent : $\boxed{\forall i \in [\![0, N]\!],\ q_{i,N} = 0}$, c'est-à-dire que la dernière colonne de la matrice Q est forcément nulle.

b) Sachant qu'il y a i contaminés à l'instant n : ces personnes sont toutes guéries à l'instant suivant, ayant pu contaminer entre temps tout ou partie des $N - i$ personnes restantes, et pas davantage, ce qui implique bien : $\boxed{\text{si } j > N - i, \quad \text{alors } q_{i,j} = 0}$.

c) Soit $i \in [\![1, N - 1]\!]$; sachant que $[X_n = i]$ est réalisé, d'après les données de l'expérience aléatoire :
Au jour suivant, chacune des $N - i$ personnes restantes a été ou non contaminée, ce qui correspond à

la réalisation de $N-i$ expérience de Bernoulli identiques et indépendantes, où le succès est ici "être contaminé par l'une des i personnes du jour n". Les contaminations éventuelles étant indépendantes : la probabilité de n'être contaminé par *aucune* des i personnes du jour n, est donc q^i, et la probabilité d'être contaminé (par *au moins une personne*), est donc $1-q^i$.
La variable aléatoire X_{n+1} compte alors le nombre de succès en $N-i$ épreuves ; on peut donc conclure que

$$\boxed{\text{La loi conditionnelle de } X_{n+1} \text{ sachant } [X_n = i] \text{ est la loi binomiale de paramètres}(N-i, 1-q^i)}$$

2. a) Un résultat très classique sur les matrices *stochastiques* : comme la famille $([X_{n+1}=j])_{0\leqslant j\leqslant N}$ est un système complet d'événements, quelle que soit la probabilité utilisée, y compris la probabilité conditionnelle sachant $[X_n = i]$:

$$\forall i \in [\![0, N]\!],\ \sum_{j=0}^{N} P_{[X_n=i]}([X_{n+1}=j]) = 1 \iff \sum_{j=0}^{N} q_{i,j} = 1$$

ce qui revient à dire que la somme des éléments de chaque ligne de la matrice Q vaut 1.

Or, matriciellement : le produit de Q par $J = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, matrice colonne de $\mathcal{M}_{N+1,1}(\mathbb{R})$ qui ne contient que des 1, donne bien une matrice colonne dont les coefficients sont les coefficients de chaque ligne de Q !
En clair, avec ce vecteur J non nul : $QJ = J$, ce qui prouve bien que 1 est valeur propre de Q, admettant J pour vecteur propre associé.

b) Soit λ une valeur propre de Q, et $V = \begin{pmatrix} V(0) \\ V(1) \\ \vdots \\ V(N) \end{pmatrix}$ un vecteur propre associé à λ.

On pose : $|V(i)| = \max\limits_{0\leqslant j\leqslant N} |V(j)|$, et ainsi : $\forall j \in [\![0, N]\!],\ 0 \leqslant |V(j)| \leqslant |V(i)|$.
Par conséquent, si $V(i) = 0$, alors $\forall j \in [\![0, N]\!],\ 0 \leqslant |V(j)| \leqslant 0 \iff |V(j)| = 0 \iff V(j) = 0$, mais alors : V est un vecteur nul, ce qui est contradictoire avec le fait que ce soit un vecteur propre ! Donc $V(i) \neq 0$.

Par identification des coefficients à la ligne i de la relation matricielle : $QV = \lambda V$, on obtient :

$$\sum_{j=0}^{N} q_{i,j}.V(j) = \lambda.V(i) \Longrightarrow |\lambda.V(i)| = \left|\sum_{j=0}^{N} q_{i,j}.V(j)\right|$$

$$\Longrightarrow |\lambda|.|V(i)| \leqslant \sum_{j=0}^{N} q_{i,j}|V(j)| \qquad \text{d'après l'inégalité triangulaire}$$

sachant que $q_{i,j} \geqslant 0$ puisque ces réels sont des probabilités. Enfin :

$$|\lambda|.|V(i)| \leqslant \sum_{j=0}^{N} q_{i,j}|V(j)| \leqslant \underbrace{\sum_{j=0}^{N} q_{i,j}}_{=1} |V(i)|$$

$$\Longrightarrow |\lambda|.|V(i)| \leqslant |V(i)| \Longrightarrow |\lambda| \leqslant 1 \text{ puisque } |V(i)| > 0$$

On vient de refaire la démonstration classique du fait que les valeurs propres d'une matrice stochastique appartiennent toutes à l'intervalle $[-1, 1]$.

On pose pour tout entier naturel n : $U_n = \begin{pmatrix} P([X_n=0]) \\ P([X_n=1]) \\ \dots \\ P([X_n=N]) \end{pmatrix}$.

La formule des probabilités totales utilisée avec le système complet d'événements $([X_n = j])_{0\leqslant j\leqslant N}$ donne, pour tout entier $n \in \mathbb{N}$:

$$\forall i \in [\![0, N]\!],\ P([X_{n+1} = i]) = \sum_{j=0}^{N} P_{[X_n=j]}([X_{n+1} = i]) \times P([X_n = j]) \iff P([X_{n+1} = i])$$

$$= \sum_{j=0}^{N} q_{j,i}.P([X_n = j]) \iff U_{n+1}(i) = \sum_{j=0}^{N} m_{i,j}.U_n(j)$$

Ce qui correspond bien à la définition du produit matriciel : $U_{n+1} = MU_n$,
où $M = (m_{i,j})_{0 \leqslant i,j \leqslant N} = (q_{j,i})_{0 \leqslant i,j \leqslant N}$ est en fait la **transposée de** Q.

On suppose jusqu'à la fin de la partie II que la matrice M est diagonalisable, et que $\mathcal{B} = (V_0, \ldots, V_N)$ est une base de vecteurs propres de M telle que, pour tout k de $[\![0, N]\!]$, le vecteur propre V_k est associé à une valeur propre λ_k.

De plus, on suppose que : $\lambda_0 = 1,\ V_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, et que pour tout k de $[\![1, N]\!]$, on a : $|\lambda_k| < 1$ (ce qui revient à supposer, à la sute du résultat de 2.b), que -1 n'est pas valeur propre).

3. On décompose alors le vecteur U_0 sur la base $\mathcal{B}$: $U_0 = \sum_{k=0}^{N} \alpha_k.V_k$.

a) Une récurrence immédiate donne : $\forall n \in \mathbb{N},\ U_n = M^n U_0$ à partir de la relation : $\forall n \in \mathbb{N},\ U_{n+1} = MU_n$. D'autre part :

Par définition, chacun des vecteurs propres V_k $(0 \leqslant k \leqslant N)$ vérifie la relation : $MV_k = \lambda_k.V_k$. Une récurrence immédiate là encore, donne : $\forall n \in \mathbb{N},\ M^n V_k = \lambda_k^n.V_k$.

Des deux relations précédentes, on déduit :

$$\forall n \in \mathbb{N},\ U_n = M^n U_0 = M^n \sum_{k=0}^{N} \alpha_k.V_k = \sum_{k=0}^{N} \alpha_k.M^n V_k = \sum_{k=0}^{N} \alpha_k.\lambda_k^n.V_k$$

b) On note, pour tout couple (k, i) de $[\![0, N]\!]^2$, $V_k(i)$ la $(i+1)$-ième composante du vecteur V_k. Pour tout entier $n \in \mathbb{N}$, et pour tout $i \in [\![0, N]\!]$, $P([X_n = i])$ est la i-ième composante du vecteur U_n, donc par indentification des coefficients dans la relation obtenue à la question précédente :

$$P([X_n = i]) = \sum_{k=0}^{N} \alpha_k.\lambda_k^n.V_k(i)$$

c) Pour tout entier $k \in [\![1, N]\!]$, l'énoncé donne $|\lambda_k| < 1$, donc : $\lim_{n \to +\infty} \lambda_k^n = 0$, et par conséquent :

$$\lim_{n \to +\infty} \alpha_k.\lambda_k^n.V_k(i) = 0.$$

Dans la somme précédente, tous les termes convergent donc vers 0, sauf le premier puisque $\lambda_0^n = \lambda_0 = 1$ reste constant :
$\forall i \in [\![1, N]\!],\ \lim_{n \to +\infty} P([X_n = i]) = \alpha_0.V_0(i) = 0$ vu que toutes les composantes du vecteur V_0 sont nulles à part celle d'indice 0.

d) Du résultat précédent on déduit : $\lim_{n \to +\infty} \sum_{i=1}^{N} P([X_n = i]) = 0$ (somme finie de suites de limite nulle).

Et comme, vu que $X_n(\Omega) \subset [\![0, N]\!]$ pour tout $n \in \mathbb{N}$, on a toujours : $P([X_n = 0]) = 1 - \sum_{i=1}^{N} P([X_n = i])$,

alors : $\lim_{n \to +\infty} P([X_n = 0]) = 1 = P\left(\bigcup_{n=0}^{+\infty} [X_n = 0]\right)$ (même argument qu'à la dernière question de la partie I), ce qui redit bien que le virus finit par disparaître presque sûrement, quelle que soit la loi de la variable aléatoire initiale X_0.

Partie III. Estimations ponctuelle et par intervalle de confiance de p

On suppose qu'un paramètre p, qui exprime la probabilité qu'un individu contagieux transmette le virus à un individu sain, est inconnu, et on cherche à l'estimer. On note comme d'habitude $q = 1 - p$.
Pour m entier supérieur ou égal à 1, on considère un m-échantillon $(Y_1, Y_2, \dots, Y_m)$ de variables aléatoires indépendantes, de même loi de Bernoulli de paramètre p, définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$.
On pose : $\overline{Y_m} = \dfrac{1}{m}\displaystyle\sum_{i=1}^{m} Y_i$.
Dans toute la suite de cette partie, on note ε un réel strictement positif quelconque.

1. a) S'il y a bien un résultat de cours à avoir retenu sur l'estimation, c'est l'utilisation de la *moyenne empirique* comme estimateur sans biais de l'espérance de la loi de l'échantillon !
   Par linéarité de l'espérance : $E(\overline{Y_n}) = \dfrac{1}{m}\sum\limits_{i=1}^{m} E(Y_i) = \dfrac{1}{m} \times m.p = p$ puisque $Y_i \hookrightarrow \mathcal{B}(p)$, donc $\overline{Y_n}$ est bien un estimateur sans biais de p.
   Comme les v.a.r. Y_i sont mutuellement indépendantes :
   $V(\overline{Y_n}) = \dfrac{1}{m^2}\sum\limits_{i=1}^{m} V(Y_i) = \dfrac{1}{m^2} \times m.p.q = \dfrac{pq}{m}$. C'est aussi le risque quadratique de l'estimateur $\overline{Y_n}$, vu qu'il est sans biais.

   b) L'inégalité de Bienaymé-Tchebicheff s'écrit, pour l'estimateur sans biais $\overline{Y_n}$:
$$\begin{aligned}\forall \varepsilon > 0, P(|\overline{Y_n} - E(\overline{Y_n})| \geqslant \varepsilon) &\leqslant \frac{V(\overline{Y_n})}{\varepsilon^2}\\ \iff 1 - P(|\overline{Y_n} - p| < \varepsilon) &\leqslant \frac{pq}{m\varepsilon^2}\\ \iff P(\overline{Y_n} - \varepsilon < p < \overline{Y_n} + \varepsilon) &\geqslant 1 - \frac{pq}{m\varepsilon^2}\end{aligned}$$
   La dernière relation obtenue exprime que $\left[\overline{Y_n} - \varepsilon ; \overline{Y_n} + \varepsilon\right]$ est un intervalle de confiance au niveau 0.95 dès que :
   $1 - \dfrac{pq}{m\varepsilon^2} \geqslant 0.95 \iff \dfrac{pq}{m\varepsilon^2} \leqslant 0.05 \iff \varepsilon^2 \geqslant \dfrac{pq}{m \times 0.05} \iff \varepsilon^2 \geqslant \dfrac{20pq}{m}$ puisque $0.05 = \dfrac{5}{100} = \dfrac{1}{20}$.
   La majoration extrêmement classique : $\forall p \in [0;1],\ pq = p(1-p) \leqslant \dfrac{1}{4}$ (étudiez le trinôme $x \mapsto x(1-x)$) assure que l'inégalité précédente est vraie dès que : $\varepsilon^2 \geqslant \dfrac{5}{m} \iff \varepsilon \geqslant \sqrt{\dfrac{5}{m}}$.
   Ainsi : $P\left(\overline{Y_n} - \sqrt{\dfrac{5}{m}} \leqslant p \leqslant \overline{Y_n} + \sqrt{\dfrac{5}{m}}\right) \geqslant 0.95$, ce qui exprime bien que l'intervalle $\left[\overline{Y_n} - \sqrt{\dfrac{5}{m}} ; \overline{Y_n} + \sqrt{\dfrac{5}{m}}\right]$ est un intervalle de confiance de p au niveau 95%.

2. Soit θ un réel positif.
   a) Certaines inégalités se démontrent par simple manipulation des événements, comme ici :
   $P(\overline{Y_n} - p \geqslant \varepsilon) = P(\overline{Y_n} \geqslant p + \varepsilon) = P(m\theta\overline{Y_n} \geqslant m\theta(p+\varepsilon)) = P(e^{m\theta\overline{Y_n}} \geqslant e^{m\theta(p+\varepsilon)})$ car m et θ sont positifs, et par stricte croissance de l'exponentielle sur $\mathbb{R}$.

   b) Encore une question de cours : on demande ici de redémontrer l'*inégalité de Markov* pour une v.a.r. T discrète finie, à valeurs positives et d'espérance $E(T)$.
   Pour tout réel a positif, l'univers-image $T(\Omega)$ se divise en deux catégories : les valeurs qui sont strictement inférieures à a, et celles qui sont supérieures ou égales à a. On écrit :
$$\begin{aligned}E(T) &= \sum_{t \in T(\Omega)} t.P(T = t)\\ &= \sum_{t \in T(\Omega)|\ t<a} t.P(T = t) + \sum_{t \in T(\Omega)|\ t \geqslant a} t.P(T = t)\end{aligned}$$
   Comme T est à valeurs positives : on peut simplement minorer la première somme par 0.
   Dans la deuxième, on considère des valeurs $t \in T(\Omega)$ telles que :
   $t \geqslant a \Longrightarrow t.P(T = t) \geqslant a.P(T = t)$ car une probabilité est toujours positive. Par sommation de ces

inégalités, on obtient :
$\sum_{t\in T(\Omega)|\ t\geqslant a} t.P(T=t) \geqslant \sum_{t\in T(\Omega)|\ t\geqslant a} a.P(T=t) = a.\sum_{t\in T(\Omega)|\ t\geqslant a} P(T=t) = a.P(T\geqslant a).$

La synthèse des inégalités obtenues donne ($a>0$) :

$$E(T) \geqslant a.P(T\geqslant a) \iff P(T\geqslant a) \leqslant \frac{E(T)}{a}$$

c) Soit g la fonction définie sur $\mathbb{R}^+$ par : $g(x)=\ln(pe^x+q)$.
On repart de : $P(\overline{Y_n}-p\geqslant\varepsilon) = P(e^{m\theta\overline{Y_n}} \geqslant e^{m\theta(p+\varepsilon)})$, où l'on va poser $T=e^{m\theta\overline{Y_n}}$ et $a=e^{m\theta(p+\varepsilon)}$.
L'inégalité de Markov donne alors :

$$P(\overline{Y_m}-p\geqslant\varepsilon) \leqslant \frac{E(e^{m\theta\overline{Y_n}})}{e^{m\theta(p+\varepsilon)}}$$

où : $T=e^{m\theta\overline{Y_n}} = e^{\theta.\sum_{i=1}^m Y_i} = \prod_{i=1}^m e^{\theta.Y_i}$. Comme les Y_i sont indépendantes, on peut écrire :

$E(T)=\prod_{i=1}^m E(e^{\theta.Y_i})$, où d'après le théorème de transfert :
$E(e^{\theta.Y_i}) = e^{m\theta.0}.P(Y_i=0)+e^{m\theta.1}.P(Y_i=1) = p.e^\theta+q$.
D'où : $E(T)=(p.e^\theta+q)^m = e^{m\ln(p.e^\theta+q)} = e^{mg(\theta)}$ (forme exponentielle de la puissance.
De tous les calculs précédents, vient bien :

$$P(\overline{Y_n}-p\geqslant\varepsilon) \leqslant \frac{e^{mg(\theta)}}{e^{m\theta(p+\varepsilon)}} = e^{m(g(\theta)-\theta(p+\varepsilon))}$$

d) La fonction g est bien définie sur $\mathbb{R}^+$ puisque $pe^x+q\geqslant p+q=1>0$ pour tout x positif, et de classe $\mathcal{C}^2$ sur cet intervalle comme composée de fonctions continues.
$\forall x\in\mathbb{R}^+ :\ g'(x)=\dfrac{pe^x}{pe^x+q}$.
$\forall x\in\mathbb{R}^+ :\ g''(x)=\dfrac{pe^x(pe^x+q)-pe^x.pe^x}{(pe^x+q)^2} = \dfrac{pqe^x}{(pe^x+q)^2}$.
Sur $\mathbb{R}^+$, $g''(x)$ est donc positive, et :
$\forall x\in\mathbb{R}^+,\ \dfrac{1}{4}-|g''(x)| = \dfrac{1}{4}-\dfrac{pqe^x}{p^2e^{2x}+2pqe^x+q^2} = \dfrac{p^2e^{2x}+2pqe^x+q^2-4pqe^x}{(pe^x+q)^2}$
$= \dfrac{p^2e^{2x}-2pqe^x+q^2}{(pe^x+q)^2} = \dfrac{(pe^x-q)^2}{(pe^x+q)^2} \geqslant 0$, donc : $\forall x\in\mathbb{R}^+,\ 0\leqslant g''(x)\leqslant\dfrac{1}{4}$.

e) On pose la fonction $f:\ x\mapsto g(x)-x.p-\dfrac{x^2}{8}$, qui est bien de classe $\mathcal{C}^2$ sur $\mathbb{R}^+$ avec :
$\forall x\geqslant0,\ g'(x)=f'(x)-p-\dfrac{x}{4}$, et $\forall x\geqslant0,\ f''(x)=g''(x)-\dfrac{1}{4}\leqslant0$.
La fonction f' est donc décroissante sur $\mathbb{R}^+$, et : $\forall x\geqslant0,\ f'(x)\leqslant h'(0)=g'(0)-p=\dfrac{p}{p+q}-p=0$.
La fonction f est donc elle-même décroissante sur $\mathbb{R}^+$, avec :
$\forall x\geqslant0,\ f(x)\leqslant f(0)=g(0)=\ln(p+q)=0$.
On a bien : $\forall\theta\in\mathbb{R}^+,\ f(\theta)\leqslant0 \iff g(\theta)\leqslant\theta p+\dfrac{\theta^2}{8}$.

f) La fonction $h:\ x\mapsto\dfrac{x^2}{8}-\varepsilon.x$ est un trinôme du second degré, de coefficient dominant $a=\dfrac{1}{8}>0$, et qui admet donc un minimum en $x=-\dfrac{b}{2a}=4\varepsilon$ (sommet de la parabole...)
Ainsi : h est décroissante sur $[0;4\varepsilon]$, et croissante sur $[4\varepsilon;+\infty[$. Le minimum vaut : $h(4\varepsilon)=-2\varepsilon^2$.
En repartant de l'inégalité : $P(\overline{Y_m}-p\geqslant\varepsilon)\leqslant e^{m(g(\theta)-\theta(p+\varepsilon))}$, et au vu des résultats précédents :
$P(\overline{Y_m}-p\geqslant\varepsilon)\leqslant e^{m(\theta.p+\theta^2/8-\theta.p-\theta.\varepsilon)} = e^{m.h(\theta)}$.
Ici l'inégalité : $h(\theta)\geqslant-2\varepsilon^2$ n'est pas dans le bon sens pour continuer à travailler par majorations successives. Mais l'inégalité précédente est en fait vraie pour tout réel $\theta\geqslant0$, donc en particulier pour $\theta=4\varepsilon$, ce qui donne bien :

$$P(\overline{Y_m}-p\geqslant\varepsilon)\leqslant e^{-2m\varepsilon^2}$$

3. On pose $\overline{W_m} = \dfrac{1}{m}\sum_{i=1}^{m}(1-Y_i)$. Comme les Y_i suivent une loi de Bernoulli de paramètre p, les v.a.r. $W_i = (1-Y_i)$ suivent toutes une loi de Bernoulli de paramètre $q = 1-p$

(puisque $Y_i = 1 \iff 1-Y_i = 0$, et vice-versa) ; elles sont toujours indépendantes, $\overline{W_m} = \dfrac{1}{m}\sum_{i=1}^{m} W_i$ donc

on peut donc réutiliser toute l'étude précédente en remplaçant partout Y_i par W_i, $\overline{Y_m}$ par $\overline{W_m}$ et p par q, ce qui donne bien :

$$P(\overline{W_m} - q \geqslant \varepsilon) \leqslant e^{-2m\varepsilon^2}$$

4. a) L'événement $[|\overline{Y_m} - p| \geqslant \varepsilon]$ peut toujours s'écrire sous la forme :
$[|\overline{Y_m} - p| \geqslant \varepsilon] = [\overline{Y_m} - p \geqslant \varepsilon] \cup [\overline{Y_m} - p \leqslant -\varepsilon]$, réunion de deux événements incompatibles.
On fait le lien avec la question 3. en écrivant que :
$\overline{W_m} = \dfrac{1}{m}\sum_{i=1}^{m}(1-Y_i) = \dfrac{m}{m} - \dfrac{1}{m}\sum_{i=1}^{m} Y_i = 1 - \overline{Y_m}$, et donc :
$P(\overline{Y_m} - p \leqslant -\varepsilon) = P(1 - \overline{W_m} - 1 + q \leqslant -\varepsilon) = P(\overline{W_m} - q \geqslant \varepsilon)$.
Ainsi : $P(|\overline{Y_m} - p| \geqslant \varepsilon) = P(\overline{Y_m} - p \geqslant \varepsilon) + P(\overline{W_m} - q \geqslant \varepsilon) \leqslant 2e^{-2m\varepsilon^2}$ au vu des majorations obtenues en 2.f) et 3.

   b) Pour $\varepsilon = \sqrt{\dfrac{1.844}{m}}$: $2e^{-2m\varepsilon^2} = 2e^{-2\times 1.844} = 2e^{-3.688} \approx 2 \times 0.025 = 0.05$.
Avec cette valeur de ε, on peut donc écrire :

$$\begin{aligned} P\left(|\overline{Y_m} - p| \geqslant \sqrt{\frac{1.844}{m}}\right) = 0.05 &\iff 1 - P\left(|\overline{Y_m} - p| < \sqrt{\frac{1.844}{m}}\right) = 0.05 \\ &\iff P\left(\overline{Y_m} - \sqrt{\frac{1.844}{m}} < p < \overline{Y_m} + \sqrt{\frac{1.844}{m}}\right) = 0.95 \end{aligned}$$

Une relation qui exprime cette fois que $\left[\overline{Y_m} - \sqrt{\dfrac{1.844}{m}} ; \overline{Y_m} + \sqrt{\dfrac{1.844}{m}}\right]$ est un intervalle de confiance de p au niveau de confiance 0.95.
Cet intervalle de confiance est évidemment plus restreint que celui obtenu à la question 1.b), puisque $1.844 < 5$, alors qu'on a le même niveau de confiance : il est donc plus intéressant ! Mais a coûté plus d'efforts pour l'obtenir...