Lycée Pierre-Gilles de Gennes 2024-2025

BCPST2 – Mathématiques

Modélisation mathématique et informatique- 3h

L'usage d'une calculatrice est autorisée pour cette épreuve (mais pas nécessaire).

Chaque candidat est responsable de la vérification de son sujet d'épreuve : pagination et impression de chaque page. Ce contrôle doit être fait en début d'épreuve. En cas de doute, le candidat doit alerter au plus tôt le surveillant qui vérifiera et, éventuellement, remplacera le sujet.

Ce sujet comporte 7 pages numérotées de 1 à 7.
Une annexe des commandes Python utiles se trouve en page 7.

Si, au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il le signale sur sa copie et poursuit sa composition en expliquant les raisons des initiatives qu'il a été amené à prendre.

Dans ce problème on étudie un modèle pour décrire la transmission d'allèles et l'évolution des fréquences alléliques au cours du temps dans une population diploïde.
Pour simplifier, on se place dans le cadre suivant :

— on étudie un seul gène sur un locus précis et qui se présente sous deux allèles distincts `A` et `a`,

— la taille de la population est supposée constante au cours du temps égale à $N \in \mathbb{N}^*$; il y a donc $2N$ locus dans la population totale,

— les générations ne se chevauchent pas : à chaque instant k, la k-ième génération meurt et donne naissance aux N individus de la génération suivante,

— la reproduction à l'instant k ne dépend pas des reproductions précédentes.

Chaque individu est de l'un des trois types suivants :

`AA` (type 1) ; `aa` (type 2) ; `Aa` (type 3).

1 Modèle de Wright-Fisher

Dans ce modèle, on néglige les phénomènes de mutation et de sélection.
On s'intéresse aux nombres d'allèles `A` présents sur le locus étudié et pour tout $n \in \mathbb{N}$, on note X_n la variable aléatoire à valeurs dans $[\![0, 2N]\!]$ donnant le nombre d'allèles de type `A` à la génération n dans une population de taille N.
On considère que pour tout entier n :

$$\forall (i,j) \in [\![0, 2N]\!]^2, \quad \mathbb{P}(X_{n+1} = j \mid X_n = i) = \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j}$$

où $\mathbb{P}(A|B)$ désigne la probabilité conditionnelle de A sachant B.

1.1 Étude d'un cas particulier

On suppose dans cette partie que $N = 1$ et on note pour tout $n \in \mathbb{N}$:

$$V_n = \begin{pmatrix} \mathbb{P}(X_n = 0) \\ \mathbb{P}(X_n = 1) \\ \mathbb{P}(X_n = 2) \end{pmatrix}.$$

1. Soit $n \in \mathbb{N}$. On utilise la formule des probabilités totales avec le système complet d'événements $([X_n = 0], [X_n = 1], [X_n = 2])$:

$$\forall j \in \{0, 1, 2\}, \quad \mathbb{P}(X_{n+1} = j) = \sum_{i=0}^{2} \mathbb{P}(X_{n+1} = j \mid X_n = i)\mathbb{P}(X_n = i).$$

Ainsi :

$$\mathbb{P}(X_{n+1} = 0) = \mathbb{P}(X_n = 0) + \frac{1}{4}\mathbb{P}(X_n = 1) + 0\mathbb{P}(X_n = 2)$$
$$\mathbb{P}(X_{n+1} = 1) = 0\mathbb{P}(X_n = 0) + \frac{1}{2}\mathbb{P}(X_n = 1) + 0\mathbb{P}(X_n = 2)$$
$$\mathbb{P}(X_{n+1} = 2) = 0\mathbb{P}(X_n = 0) + \frac{1}{4}\mathbb{P}(X_n = 1) + \mathbb{P}(X_n = 2).$$

Ainsi, on a :

$$V_{n+1} = \begin{pmatrix} 1 & \frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & 1 \end{pmatrix} V_n.$$

2. (a) Soit $\lambda \in \mathbb{R}$.

$$rg(M - \lambda I_3) = rg \begin{pmatrix} 1-\lambda & 1/4 & 0 \\ 0 & 1/2-\lambda & 0 \\ 0 & 1/4 & 1-\lambda \end{pmatrix} = rg \begin{pmatrix} 1-\lambda & 0 & 1/4 \\ 0 & 0 & 1/2-\lambda \\ 0 & 1-\lambda & 1/4 \end{pmatrix}$$
$$= rg \begin{pmatrix} 1-\lambda & 0 & 1/4 \\ 0 & 1-\lambda & 1/4 \\ 0 & 0 & 1/2-\lambda \end{pmatrix}$$

en inversant les deux dernières colonnes puis les deux dernières lignes (opérations qui ne changent pas le rang).

On en déduit : $rg(M - \lambda I_3) < 3 \iff \lambda = 1 \quad \text{ou} \quad \lambda = \frac{1}{2}$.

Ainsi $\mathrm{Spec}(M) = \{\frac{1}{2}, 1\}$.

— Sous-espace propre associé à 1 : soit $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R})$.

$$MX = X \iff \begin{cases} x + \frac{1}{4}y &= x \\ \frac{1}{2}y &= y \\ \frac{1}{4}y + z &= z \end{cases} \iff y = 0.$$

Ainsi $E_1(M) = \text{Vect}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right)$ et on vérifie facilement que $\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right)$ est une base de $E_1(M)$.

— Sous-espace propre associé à 1/2 : soit $X = \begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathcal{M}_{3,1}(\mathbb{R})$.

$$MX = \frac{1}{2}X \iff \begin{cases} x + \frac{1}{4}y &= \frac{1}{2}x \\ \frac{1}{2}y &= \frac{1}{2}y \\ \frac{1}{4}y + z &= \frac{1}{2}z \end{cases} \iff \begin{cases} \frac{1}{2}x + \frac{1}{4}y &= 0 \\ \frac{1}{4}y + \frac{1}{2}z &= 0 \end{cases}$$
$$\iff \begin{cases} 2x + y &= 0 \\ y + 2z &= 0 \end{cases}$$
$$\iff \begin{cases} y &= -2x \\ z &= x \end{cases}$$

Ainsi $E_{\frac{1}{2}}(M) = \text{Vect}\left(\begin{pmatrix}1\\-2\\1\end{pmatrix}\right)$.

(b) La somme des dimensions des sous-espaces propres est 3, égale à la dimension de la matrice M. Ainsi M est diagonalisable.
D'après ce qui précède, en posant :

$$P = \begin{pmatrix}1 & 1 & 0\\0 & -2 & 0\\0 & 1 & 1\end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix}1 & 0 & 0\\0 & \frac{1}{2} & 0\\0 & 0 & 1\end{pmatrix}$$

les formules de changement de bases donnent :

$$M = PDP^{-1}.$$

3. Deux méthodes pour le calcul de M^n. Une seule des deux sous-questions peut être traitée.

(a) On montre facilement par récurrence que :

$$\forall n \in \mathbb{N}, \quad M^n = PD^nP^{-1}.$$

— **Initialisation :** les cas $n = 0$ est immédiat.

— **Hérédité :** soit $n \in \mathbb{N}$ et supposons que $M^n = PD^nP^{-1}$. Alors avec la question précédente :

$$M^{n+1} = M^nM = PD^nP^{-1}PDP^{-1} = PD^nI_3DP^{-1} = PD^nDP^{-1} = PD^{n+1}P^{-1}.$$

— **Conclusion :** pour tout $n \in \mathbb{N}$: $M^n = PD^nP^{-1}$.

Ainsi :

$$M^n = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1}.$$

Il suffit de déterminer P^{-1} puis faire le produit matriciel ci-dessus. On trouve alors :

$$P^{-1} = \frac{1}{2} \begin{pmatrix} 2 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \quad \text{puis} \quad M^n = \begin{pmatrix} 1 & \frac{2^n-1}{2^{n+1}} & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & \frac{2^n-1}{2^{n+1}} & 1 \end{pmatrix}.$$

(b) — **Initialisation :** évident.

— **Hérédité :** soit $n \in \mathbb{N}$. On suppose le résultat acquis au rang n. Alors :

$$M^{n+1} = MM^n = M \begin{pmatrix} 1 & \frac{2^n-1}{2^{n+1}} & 0 \\ 0 & \frac{1}{2^n} & 0 \\ 0 & \frac{2^n-1}{2^{n+1}} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{2^{n+1}-1}{2^{n+2}} & 0 \\ 0 & \frac{1}{2^{n+1}} & 0 \\ 0 & \frac{2^{n+1}-1}{2^{n+2}} & 1 \end{pmatrix}.$$

La propriété est donc vérifiée au rang $n+1$.

— **Conclusion :** par récurrence le résultat voulu est démontré.

4. (a) Soit $n \in \mathbb{N}$. Les variables aléatoires X_n et X_0 sont de support fini donc elles possèdent une espérance.

Notons $V_0 = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Une récurrence immédiate montre que :

$$V_n = M^n V_0 = \begin{pmatrix} a + b\frac{2^n-1}{2^{n+1}} \\ \frac{b}{2^n} \\ b\frac{2^n-1}{2^{n+1}} + c \end{pmatrix}.$$

Ainsi :

$$\begin{aligned} E(X_n) &= 0 \times \mathbb{P}(X_n = 0) + 1 \times \mathbb{P}(X_n = 1) + 2 \times \mathbb{P}(X_n = 2) \\ &= 0 \times (a + b\frac{2^n-1}{2^{n+1}}) + \frac{b}{2^n} + 2 \times (b\frac{2^n-1}{2^{n+1}} + c) \\ &= 0 \times a + b + 2c \\ &= E(X_0). \end{aligned}$$

(b) Soit $n \in \mathbb{N}$. En concervant les notations de la question précédente, on a :

$$\begin{aligned} \mathbb{P}(X_n\{0,2\}) = \mathbb{P}(X_n = 0) + \mathbb{P}(X_n = 2) &= a + b\frac{2^n-1}{2^{n+1}} + +b\frac{2^n-1}{2^{n+1}} + c \\ &\xrightarrow[n\to+\infty]{} a + b + c = 1. \end{aligned}$$

1.2 Cas général.

On suppose désormais que $N \geq 1$.

5. Python. Dans cette question et uniquement dans cette question, on suppose que X_0 suit une loi certaine de paramètre `x0`$\in [\![0, 2N]\!]$.

(a) On considère la fonction suivante :

```
def X(N,n,x0):
    X = x0
    for k in range(n):
        X =rd.binomial(2*N,X/(2*N))
    return X
```

La variable `x0` désigne le nombre d'allèles `A` dans la génération initiale.
La commande `X =rd.binomial(2*N,X/(2*N))` remplace la valeur de `X` par une réalisation d'une loi binomiale de paramètres 2N, $\dfrac{\mathtt{X}}{2\mathtt{N}}$: cela simule le nombre d'allèle `A` à la génération suivante.
La boucle `for` permet de répéter ce procédé n fois donc la fonction simule la variable X_n.

(b) On suppose que `x0`, `n` et `N` ont déjà été déclarés dans le script. Compléter les lignes suivantes pour simuler 1000 fois la variable $X_{\mathtt{n}}$ et afficher une valeur approchée de $\mathbb{P}(X_{\mathtt{n}} \in \{0, 2\mathtt{N}\})$ dans la variable `s` :

```
s = 0
for k in range(1000):
        Xn = X(N,n,x0)
        if Xn == 0 or Xn == 2*N :
                s + = 1/1000
print(s)
```

(c) Avec la commande de la question précédente, on réalise 1000 fois de manière indépendante l'expérience de Bernoulli consistant à simuler X_n et dont le succès est $X_n \in \{0, 1\}$. La variable `s`, qui est la fréquence de succès, est donc une valeur approchée de $\mathbb{P}(X_n \in \{0, 1\})$.

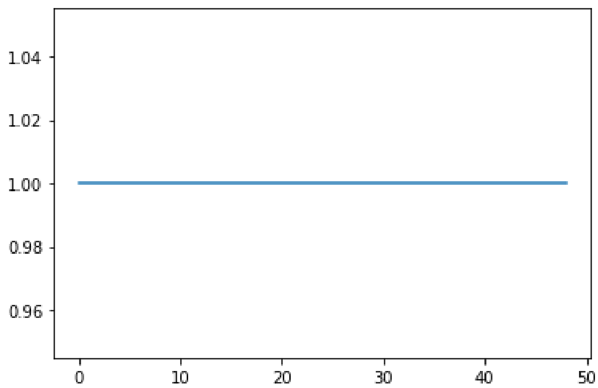

FIGURE 1 – $\mathbb{P}(X_{1000} \in \{0, 2N\})$ pour $N = 1, \ldots 50$.

Comme $n = 1000$ est « grand », on peut conjecturer que $\lim\limits_{n \to +\infty} \mathbb{P}(X_n \in \{0, 2N\}) \simeq$

$\mathbb{P}(X_{1000} \in \{0, 2N\}) \simeq 1$ en cohérence avec le résultat trouvé en **4.(b)** dans le cas $N = 1$.

Cela signifie qu'après un grand nombre de générations l'un des deux allèles disparaissent presque-sûrement.

6. **(a)** Soit $i \in [\![0, 2N]\!]$. On reconnaît dans la somme

$$S_i = \sum_{j=0}^{2N} j \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j}$$

l'espérance d'une variable aléatoire de loi $\mathcal{B}(2N, \frac{i}{2N})$.

Ainsi : $S_i = 2N \times \dfrac{i}{2N} = i$.

(b) Soit $n \in \mathbb{N}$. Les variables sont à valeurs dans $[\![0, 2N]\!]$ donc possèdent une espérance :

$$E(X_{n+1}) = \sum_{j=0}^{2N} j\mathbb{P}(X_{n+1} = j).$$

En appliquant la formule des probabilités totales avec le système complet d'événements $([X_n = i])_{i \in [\![0,2N]\!]}$ on obtient pour tout $j \in [\![0, 2N]\!]$:

$$\begin{aligned}\mathbb{P}(X_{n+1} = j) &= \sum_{i=0}^{2N} \mathbb{P}(X_{n+1} = j \mid X_n = i)\mathbb{P}(X_n = i) \\ &= \sum_{i=0}^{2N} \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j} \mathbb{P}(X_n = i).\end{aligned}$$

Ainsi :

$$E(X_{n+1}) = \sum_{j=0}^{2N} j \sum_{i=0}^{2N} \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j} \mathbb{P}(X_n = i).$$

et en intervertissant les sommes :

$$\begin{aligned}E(X_{n+1}) &= \sum_{i=0}^{2N} \mathbb{P}(X_n = i) \sum_{j=0}^{2N} j \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j} \\ &= \sum_{i=0}^{2N} \mathbb{P}(X_n = i) S_i \\ &= \sum_{i=0}^{2N} \mathbb{P}(X_n = i) i \quad \text{par } \mathbf{6.(a)} \\ &= E(X_n).\end{aligned}$$

(c) Le nombre moyen d'allèle de type `A` est constant au sein d'une population.

7. On considère la suite u de terme général $u_n = \mathbb{P}(X_n \in \{0, 2N\})$.

(a) Soit $n \in \mathbb{N}$ et $k \in [\![1, 2N-1]\!]$. On sait que :

$$\mathbb{P}(X_{n+1} = 0 \mid X_n = k) = \left(1 - \frac{k}{2N}\right)^{2N}.$$

Or $k \leq 2N - 1$ donc $\dfrac{k}{2N} \leq 1 - \dfrac{1}{2N}$ et ainsi :

$$1 - \frac{k}{2N} \geq \frac{1}{2N}.$$

Par croissance de $x \mapsto x^{2N}$ sur $\mathbb{R}_+$:

$$\mathbb{P}(X_{n+1} = 0 \mid X_n = k) = \left(1 - \frac{k}{2N}\right)^{2N} \geq \left(\frac{1}{2N}\right)^{2N}.$$

De même, on a :

$$\mathbb{P}(X_{n+1} = 2N \mid X_n = k) = \left(\frac{k}{2N}\right)^{2N} \geq \left(\frac{1}{2N}\right)^{2N}$$

car $k \geq 1$. Finalement, on a donc

$$\mathbb{P}(X_{n+1} \in \{0, 2N\} \mid X_n = k) \geq 2\left(\frac{1}{2N}\right)^{2N}.$$

(b) Soit $n \in \mathbb{N}$. On note $A_n = [X_n \in \{0, 2N\}]$. Par la formule des probabilités totales toujours avec le même système complet d'événements :

$$\begin{aligned}
u_{n+1} &= \sum_{k=0}^{2N} \mathbb{P}(A_{n+1} \mid X_n = k)\mathbb{P}(X_n = k) \\
&= \sum_{k=1}^{2N-1} \mathbb{P}(A_{n+1} \mid X_n = k)\mathbb{P}(X_n = k) \\
&\quad + \mathbb{P}(A_{n+1} \mid X_n = 0)\mathbb{P}(X_n = 0) + \mathbb{P}(A_{n+1} \mid X_n = 2N)\mathbb{P}(X_n = 2N) \\
&\geq 2\left(\frac{1}{2N}\right)^{2N} \sum_{k=1}^{2N-1} \mathbb{P}(X_n = k) \\
&\quad + \mathbb{P}(A_{n+1} \mid X_n = 0)\mathbb{P}(X_n = 0) + \mathbb{P}(A_{n+1} \mid X_n = 2N)\mathbb{P}(X_n = 2N) \\
&\geq 2\left(\frac{1}{2N}\right)^{2N} \mathbb{P}(X_n \notin \{0, 2N\}) + \mathbb{P}(X_n = 0) + \mathbb{P}(X_n = 2N) \\
&\geq 2\left(\frac{1}{2N}\right)^{2N} (1 - u_n) + \mathbb{P}(X_n \in \{0, 2N\}) \\
&\geq 2\left(\frac{1}{2N}\right)^{2N} (1 - u_n) + u_n.
\end{aligned}$$

(c) Soit $\alpha \in]0,1[$. On considère la suite w définie par :

$$w_0 = u_0 \quad \text{et} \quad \forall n \in \mathbb{N},\ w_{n+1} = w_n + \alpha(1 - w_n).$$

On a pour tout $n \in \mathbb{N}$:

$$w_{n+1} = (1-\alpha)w_n + \alpha.$$

Il s'agit donc d'une suite arithmético-géométrique. On en déduit, *via* la méthode d'étude de telles suites, que :

$$\forall n \in \mathbb{N}, \quad w_n = (1-\alpha)^n(w_0 - 1) + 1.$$

Comme $0 < \alpha < 1$, alors $\lim\limits_{n\to+\infty}(1-\alpha)^n = 0$ et on obtient que w est convergente avec 1 pour limite.

(d) On pose $\alpha = 2\left(\dfrac{1}{2N}\right)^{2N} \in]0,1[$ (car $N \geq 1$) et w la suite définie à la question précédente.

Montrons par récurrence que pour tout $n \in \mathbb{N}$, $u_n \geq w_n$.

— **Initialisation :** $u_0 = w_0$ donc c'est évident.

— **Hérédité :** soit $n \in \mathbb{N}$ et supposons que $u_n \geq w_n$. Par la question **7.(b)** et l'hypothèse de récurrence on a :

$$u_{n+1} \geq \alpha(1-u_n) + u_n = (1-\alpha)u_n + \alpha \geq (1-\alpha)w_n + \alpha = w_{n+1}.$$

Ainsi le résultat est vérifié au rang $n+1$.

— **Conclusion :** pour tout $n \in \mathbb{N}$, $u_n \geq w_n$.

Comme u_n est une probabilité, on a donc pour tout $n \in \mathbb{N}$:

$$w_n \leq u_n \leq 1.$$

D'après la question précédente, w converge vers 1 donc le théorème des gendarmes permet de conclure que u converge vers 1.

Cela prouve la conjecture de la question **5.(c)**.

8. Hétérozygotie. On s'intéresse à la probabilité que deux locus tirés aux hasard (sans remise) à la génération n portent des allèles différents. On note $h(n)$ cette probabilité.

(a) Soit B_n l'événement « choisir deux locus au hasard uniformément sans remise » à la génération n. La formule des probabilités totales (encore...) donne :

$$h(n) = \sum_{j=0}^{2N} \mathbb{P}(A_n \mid X_n = j)\mathbb{P}(X_n = j).$$

Sachant l'événement $[X_n = j]$ réalisé, on tire deux allèles parmi $2N$ allèles dont j sont des `A` et $2N - j$ des `a`. En tirer un de chaque revient à en tirer un parmi les j `A` et 1 parmi les $2N - j$ `a`. D'où :

$$\mathbb{P}(A_n \mid X_n = j) = \frac{\binom{j}{1}\binom{2N-j}{1}}{\binom{2N}{2}} = \frac{j(2N-j)}{\frac{2N(2N-1)}{2}} = \frac{j(2N-j)}{N(2N-1)}.$$

On trouve bien :

$$h(n) = \sum_{j=0}^{2N} \frac{j(2N-j)}{N(2N-1)}\mathbb{P}(X_n = j).$$

(b) Soit $n \in \mathbb{N}$. La question précédente et la formule des probabilités totales (...et toujours) donnent :

$$\begin{aligned}
h(n+1) &= \sum_{j=0}^{2N} \frac{j(2N-j)}{N(2N-1)} \mathbb{P}(X_{n+1}=j) \\
&= \sum_{j=0}^{2N} \frac{j(2N-j)}{N(2N-1)} \sum_{i=0}^{2N} \mathbb{P}(X_{n+1}=j \mid X_n=i)\mathbb{P}(X_n=i) \\
&= \sum_{i=0}^{2N} \mathbb{P}(X_n=i) \sum_{j=0}^{2N} \frac{j(2N-j)}{N(2N-1)} \mathbb{P}(X_{n+1}=j \mid X_n=i) \\
&= \sum_{i=0}^{2N} \mathbb{P}(X_n=i) \frac{1}{N(2N-1)} \sum_{j=0}^{2N} j(2N-j)\binom{2N}{j}\left(\frac{i}{2N}\right)^j\left(1-\frac{i}{2N}\right)^{2N-j}.
\end{aligned}$$

La somme $\displaystyle\sum_{j=0}^{2N} j(2N-j)\binom{2N}{j}\left(\frac{i}{2N}\right)^j\left(1-\frac{i}{2N}\right)^{2N-j}$ n'est autre que l'espérance de $X(2N-X)$ où X suit la loi $\mathcal{B}(2N, i/2N)$ et vaut donc avec Koenig-Huygens :

$$\begin{aligned}
\mathbb{E}(X(2N-X)) = 2N\mathbb{E}(X) - \mathbb{E}(X^2) &= 2Ni - \mathbb{V}(X) - \mathbb{E}(X)^2 \\
&= 2Ni - i\left(1-\frac{i}{2N}\right) - i^2 \\
&= i\left(2N - i - \frac{2N-i}{2N}\right) \\
&= i(2N-i)\left(1-\frac{1}{2N}\right).
\end{aligned}$$

On obtient donc finalement

$$\begin{aligned}
h(n+1) &= \sum_{i=0}^{2N} \mathbb{P}(X_n=i) \frac{1}{N(2N-1)} \sum_{j=0}^{2N} j(2N-j)\binom{2N}{j}\left(\frac{i}{2N}\right)^j\left(1-\frac{i}{2N}\right)^{2N-j} \\
&= \sum_{i=0}^{2N} \mathbb{P}(X_n=i) \frac{1}{N(2N-1)} i(2N-i)\left(1-\frac{1}{2N}\right) \\
&= \left(1-\frac{1}{2N}\right) h(n).
\end{aligned}$$

(c) La suite $(h(n))$ est géométrique d'après la question précédente donc :

$$\forall n \in \mathbb{N}, \quad h(n) = \left(1-\frac{1}{2N}\right)^n h(0).$$

2 Équilibre de Hardy-Weinberg

Soient p_1, p_2 et p_3 trois réels strictement positifs tels que $p_1+p_2+p_3=1$.
On suppose que pour tout $i = 1, 2, 3$, p_i représente la fréquence du génotype de type i au locus étudié et on note N_i la variable aléatoire de loi $\mathcal{B}(N, p_i)$ donnant le nombre d'individus de type i.

9. Cf cours.

10. On considère la matrice $W = \begin{pmatrix} \mathbb{V}(N_1) & \mathrm{Cov}(N_1, N_2) \\ \mathrm{Cov}(N_2, N_1) & \mathbb{V}(N_2) \end{pmatrix}$.

(a) Par symétrie de la covariance, W est symétrique à coefficients réels donc diagonalisable.

(b) Soit $(a, b) \in \mathbb{R}^2$. D'une part :

$$\begin{aligned} \mathbb{V}(aN_1 + bN_2) &= \mathbb{V}(aN_1) + 2\mathrm{Cov}(aN_1, bN_2) + \mathbb{V}(bN_2) \\ &= a^2\mathbb{V}(N_1) + 2ab\mathrm{Cov}(N_1, N_2) + b^2\mathbb{V}(N_2). \end{aligned}$$

D'autre part :

$$\begin{aligned} \begin{pmatrix} a & b \end{pmatrix} W \begin{pmatrix} a \\ b \end{pmatrix} &= \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} a\mathbb{V}(N_1) + b\mathrm{Cov}(N_1, N_2) \\ a\mathrm{Cov}(N_2, N_1) + b\mathbb{V}(N_2) \end{pmatrix} \\ &= a^2\mathbb{V}(N_1) + ab\mathrm{Cov}(N_1, N_2) + ab\mathrm{Cov}(N_2, N_1) + b^2\mathbb{V}(N_2) \\ &= a^2\mathbb{V}(N_1) + 2ab\mathrm{Cov}(N_1, N_2) + b^2\mathbb{V}(N_2) \end{aligned}$$

par symétrie de la covariance.

On a donc bien :

$$\mathbb{V}(aN_1 + bN_2) = \begin{pmatrix} a & b \end{pmatrix} W \begin{pmatrix} a \\ b \end{pmatrix}$$

(c) Soit λ une valeur propre et $\begin{pmatrix} a \\ b \end{pmatrix}$ un vecteur propre associé.

D'une part :

$$\begin{pmatrix} a & b \end{pmatrix} W \begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \lambda(a^2 + b^2).$$

D'autre part : $\mathbb{V}(aN_1 + bN_2) = \begin{pmatrix} a & b \end{pmatrix} W \begin{pmatrix} a \\ b \end{pmatrix}$.

Ainsi :

$$\mathbb{V}(aN_1 + bN_2) = \lambda(a^2 + b^2).$$

Comme $\begin{pmatrix} a \\ b \end{pmatrix}$ est un vecteur propre, c'est un vecteur non nul donc a et b ne sont pas simultanément nuls et $a^2 + b^2 > 0$. Par ailleurs, une variance est toujours positives donc :

$$\lambda = \frac{\mathbb{V}(aN_1 + bN_2)}{a^2 + b^2} \geq 0.$$

(d) Supposons que 0 soit valeur propre de W et soit $\begin{pmatrix} a \\ b \end{pmatrix}$ un vecteur propre associé.

Alors d'après ce qui précède on a :

$$\mathbb{V}(aN_1 + bN_2) = 0.$$

Par conséquent la variable $aN_1 + bN_2$ est de variance nulle donc suit une loi certaine.

Or il est clair que $aN_1 + bN_2$ ne suit pas une loi certaine : si $N_1 = 0$ et $N_2 = 0$ alors $aN_1 + bN_2 = 0$, si $N_1 = 1$ et $N_2 = 0$ alors $aN_1 + bN_2 = a$ et si $N_1 = 0$ et

$N_2 = 1$ alors $aN_1 + bN_2 = b$. Or a et b ne peuvent pas être simultanément nuls car $\begin{pmatrix} a \\ b \end{pmatrix}$ est un vecteur propre donc un vecteur non nul.

Par conséquent, 0 ne peut pas être valeur propre et avec la question précédente cela montre que les valeurs propres de W sont strictement positives.

(e) D'après la question **10.(a)**, W est diagonalisable. Il existe donc P inversible et Δ diagonale dont les coefficients diagonaux sont les valeurs propres de W telles que :

$$W = P\Delta P^{-1}.$$

Soit D la matrice dont les coefficients diagonaux sont les racines carrées des coefficients diagonaux de Δ : c'est possible d'après **10.(c)**. Alors :

$$D^2 = \Delta.$$

Ainsi on a

$$W = PD^2P^{-1}.$$

(f) Comme les valeurs propres de W sont strictement positives, il en va de même des coefficients diagonaux de Δ donc de D. Ainsi D est diagonale à coefficients diagonaux non nuls donc est inversible.

On note $A = D^{-1}P^{-1}$ et on considère les variables aléatoires Y_1 et Y_2 telles que :

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = A \begin{pmatrix} N_1 - Np_1 \\ N_2 - Np_2 \end{pmatrix}.$$

On admet que :

— $AWA^t = \begin{pmatrix} \mathbb{V}(Y_1) & \mathrm{Cov}(Y_1, Y_2) \\ \mathrm{Cov}(Y_2, Y_1) & \mathbb{V}(Y_2) \end{pmatrix}$

— $P^{-1} = P^t$

où M^t désigne la transposée de la matrice M.

11. Il suffit de calculer AWA^t. Compte tenu de la question **10.(e)** et des résultats admis ci-dessus :

$$\begin{aligned} AWA^t = D^{-1}P^{-1}PD^2P^{-1}(D^{-1}P^{-1})^t = D^{-1}D^2P^{-1}(D^{-1}P^{-1})^t &= DP^{-1}(D^{-1}P^{-1})^t \\ &= DP^{-1}(P^{-1})^t(D^{-1})^t \\ &= DP^{-1}(P^t)^tD^{-1} \\ &= DP^{-1}PD^{-1} \\ &= I_2. \end{aligned}$$

Ainsi, avec le résultat admis on a :

$$\mathrm{Cov}(Y_1, Y_2) = 0 \quad ; \quad \mathbb{V}(Y_1) = \mathbb{V}(Y_2) = 1.$$

12. **(a)** La variable $N_1 + N_2$ compte le nombre d'individu de type 1 ou 2 dans une population de N individus où la proportion d'individus de type 1 ou 2 est $p_1 + p_2$.

Ainsi $N_1 + N_2$ suit la loi $\mathcal{B}(N, p_1 + p_2)$ donc :

$$\mathbb{V}(N_1 + N_2) = N(p_1 + p_2)(1 - p_1 - p_2) = Np_3(1 - p_3).$$

(b) On sait que :

$$\mathbb{V}(N_1+N_2) = \mathbb{V}(N_1)+2\mathrm{Cov}(N_1,N_2)+\mathbb{V}(N_2) = Np_1(1-p_1)+2\mathrm{Cov}(N_1,N_2)+Np_2(1-p_2).$$

Avec la question précédente, on en déduit en se rappelant que $p_1+p_2+p_3=1$:

$$\begin{aligned}\mathrm{Cov}(N_1,N_2) &= \frac{Np_3(1-p_3)-Np_1(1-p_1)-Np_2(1-p_2)}{2}\\ &= N\frac{(1-p_1-p_2)(p_2+p_1)-Np_1(1-p_1)-Np_2(1-p_2)}{2}\\ &= N\frac{p_1(1-p_1-p_2-(1-p_1))+p_2(1-p_1-p_2-(1-p_2))}{2}\\ &= -Np_1p_2.\end{aligned}$$

(c) On a :

$$\begin{aligned}\det(W) = \mathbb{V}(N_1)\mathbb{V}(N_2)-\mathrm{Cov}(N_1,N_2)^2 &= N^2p_1(1-p_1)p_2(1-p_2)-N^2p_1^2p_2^2\\ &= N^2p_1p_2((1-p_1)(1-p_2)-p_1p_2)\\ &= N^2p_1p_2(1-p_1-p_2)\\ &= N^2p_1p_2p_3\end{aligned}$$

car $p_3 = 1-p_1-p_2$.
On en déduit que

$$W^{-1} = \frac{1}{N^2p_1p_2p_3}\begin{pmatrix}\mathbb{V}(N_2) & -\mathrm{Cov}(N_1,N_2)\\ -\mathrm{Cov}(N_1,N_2) & \mathbb{V}(N_1)\end{pmatrix} = \frac{1}{Np_1p_2p_3}\begin{pmatrix}p_2(1-p_2) & p_1p_2\\ p_1p_2 & p_1(1-p_1)\end{pmatrix}.$$

(d) Compte tenu de la définition de A et des propriétés admises avant la question 11, on a :

$$\begin{aligned}A^tA = (D^{-1}P^{-1})^tD^{-1}P^{-1} &= (D^{-1}P^t)^tD^{-1}P^{-1}\\ &= (P^t)^t(D^{-1})^tD^{-1}P^{-1}\\ &= PD^{-1}D^{-1}P^{-1}\\ &= PD^{-2}P^{-1}\\ &= (PD^2P^{-1})^{-1}\\ &= W^{-1}.\end{aligned}$$

Avec l'indication et vu la définition de Y_1 et Y_2 on trouve alors :

$$\begin{aligned}Y_1^2+Y_2^2 = \begin{pmatrix}Y_1 & Y_2\end{pmatrix}\begin{pmatrix}Y_1\\ Y_2\end{pmatrix} &= \left(A\begin{pmatrix}N_1-Np_1\\ N_2-Np_2\end{pmatrix}\right)^t A\begin{pmatrix}N_1-Np_1\\ N_2-Np_2\end{pmatrix}\\ &= \begin{pmatrix}N_1-Np_1 & N_2-Np_2\end{pmatrix}A^tA\begin{pmatrix}N_1-Np_1\\ N_2-Np_2\end{pmatrix}\\ &= \begin{pmatrix}N_1-Np_1 & N_2-Np_2\end{pmatrix}W^{-1}\begin{pmatrix}N_1-Np_1\\ N_2-Np_2\end{pmatrix}\end{aligned}$$

Il ne reste plus qu'à faire le calcul matriciel :

$$\begin{aligned}
&\begin{pmatrix} N_1 - Np_1 & N_2 - Np_2 \end{pmatrix} W^{-1} \begin{pmatrix} N_1 - Np_1 \\ N_2 - Np_2 \end{pmatrix} \\
&= \frac{1}{Np_1p_2p_3} \begin{pmatrix} N_1 - Np_1 & N_2 - Np_2 \end{pmatrix} \begin{pmatrix} (N_1 - Np_1)p_2(1-p_2) + p_1p_2(N_2 - Np_2) \\ (N_1 - Np_1)p_1p_2 + p_1(1-p_1)(N_2 - Np_2) \end{pmatrix} \\
&= \frac{((N_1 - Np_1)^2 p_2(1-p_2) + 2p_1p_2(N_2 - Np_2)(N_1 - Np_1) + (N_2 - Np_2)^2 p_1(1-p_1))}{Np_1p_2p_3} \\
&= \frac{((N_1 - Np_1)^2 p_2(p_1+p_3) + 2p_1p_2(N_2 - Np_2)(N_1 - Np_1) + (N_2 - Np_2)^2 p_1(p_2+p_3))}{Np_1p_2p_3} \\
&= \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} + \frac{(N_1 - Np_1)^2 + 2(N_2 - Np_2)(N_1 - Np_1) + (N_2 - Np_2)^2}{Np_3} \\
&= \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} + \frac{(N_1 + N_2 - N(p_1+p_2))^2}{Np_3} \\
&= \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} + \frac{(N - N_3 - N(1-p_3))^2}{Np_3} \\
&= \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} + \frac{(N_3 - Np_3)^2}{Np_3}.
\end{aligned}$$

Ainsi

$$Y_1^2 + Y_2^2 = \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} + \frac{(N_3 - Np_3)^2}{Np_3}.$$

3 Étude de la loi limite

Soient Z_1 et Z_2 deux variables aléatoires indépendantes de loi normale centrée et réduite. On note $T = Z_1^2 + Z_2^2$.

13. On a tracé sur la figure 2, la fonction de répartition F_T de T.

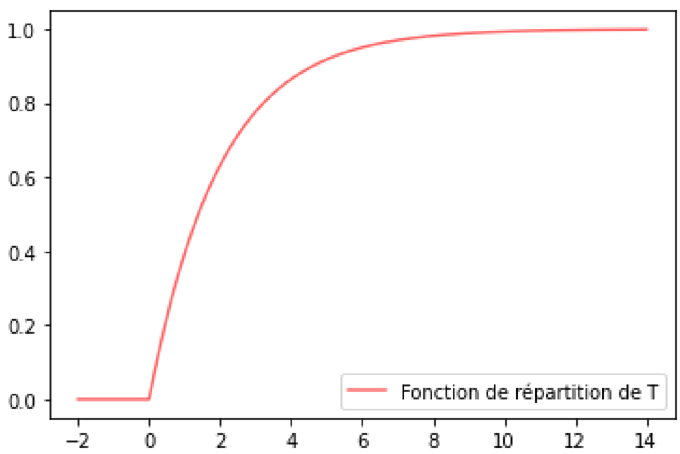

FIGURE 2 – Fonction de répartition de T

(a) La courbe représente une fonction qui semble croissante et avoir pour limite 0 en $+\infty$ et 1 en $+\infty$: c'est normal car se sont des propriétés des fonctions de répartition.

On peut aussi voir que F_T est nulle sur $]-\infty, 0]$: cela vient du fait que T est presque-sûrement positive.

(b) La courbe n'a pas de « cassure » donc F_T semble continue sur $\mathbb{R}$.

La courbe n'a pas de « d'angle » sauf en 0 donc on peut conjecturer que F_T est de classe C^1 sur $\mathbb{R}^*$.

On peut donc conjecturer que T est à densité.

14. **(a)** Écrire une fonction `question_12d(N,p1,p2,p3)` qui prend en argument le nombre d'individus N de la population, les fréquences p_1, p_2, p_3 des trois génotypes et qui simule la variable $Y_1^2 + Y_2^2$ de la question **12.(d)** :

```
def question12_d(N,p1=0.3,p2=0.5,p3=0.2):
        a = (rd.binomial(N, p1)-N*p1)**2/(N*p1)
        b = (rd.binomial(N, p2)-N*p2)**2/(N*p2)
        c = (rd.binomial(N, p3)-N*p3)**2/(N*p3)
    return a+b+c
```

(b) On écrit, à la suite de la fonction précédente, la fonction suivante où `x` est un réel :

```
def mystere(x,N,p1,p2,p3):
    s = 0
    for k in range(1000):
        if question_12d(N,p1,p2,p3)<=x:
            s += 1/1000
    return s
```

On réalise 1000 simulation de $Y_1^2 + Y_2^2$ et on regard la proportion de fois où l'événement $[Y_1^2 + Y_2^2 \leq x]$ est réalisé. Ainsi, la valeur renvoyée est une valeur approchée de :

$$\mathbb{P}(Y_1^2 + Y_2^2 \leq x).$$

Cette fonction est une approximation de la fonction de répartition de $Y_1^2 + Y_2^2$.

(c) On suppose N, p_1, p_2, p_3 déjà déclarés dans Python et on souhaite tracer la fonction `x` $\mapsto$ `mystere(x,N,p1,p2,p3)` entre -2 et 14 avec 1000 points.

Compléter les lignes de codes :

```
x = np.linspace(-2 ,14 ,1000)
L = [mystere(t,N,p1,p2,p3) for t in x]
plt.plot(x,L)
plt.show()
```

(d) On a effectué les tracés pour différentes valeurs de N (voir figures 3 suivantes) :

La courbe représentant l'approximation de $x \mapsto \mathbb{P}(Y_1^2 + Y_2^2 \leq x)$ (en bleu) semble s'approcher de plus en plus de F_T quand N tend vers $+\infty$. On peut donc conjecturer que pour tout $x \in \mathbb{R}$:

$$\lim_{N \to +\infty} \mathbb{P}(Y_1^2 + Y_2^2 \leq x) = F_T(x).$$

15. Soit $x \in \mathbb{R}$:

$$F_{Z_1^2}(x) = \mathbb{P}(Z_1^2 \leq x).$$

— Si $x < 0$ alors $F_{Z_1^2}(x) = \mathbb{P}(Z_1^2 \leq x) = 0.$.

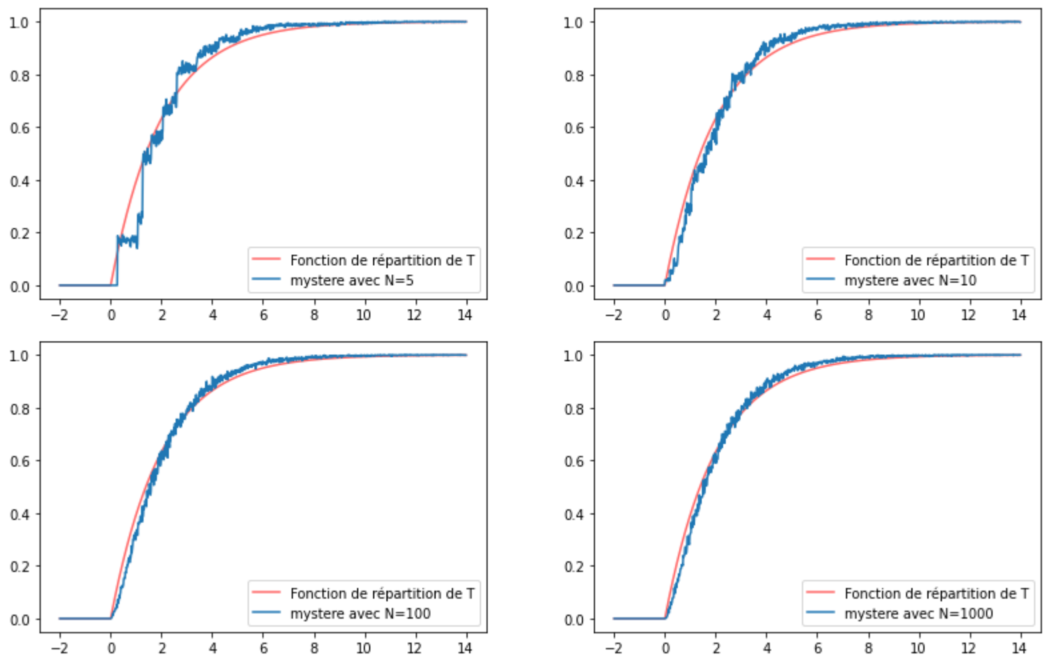

FIGURE 3 – Fonction mystère pour $N \in \{5, 10, 100, 1000\}$.

- Si $x = 0$ alors $F_{Z_1^2}(x) = \mathbb{P}(Z_1 = 0) = 0$ car Z_1 est à densité.
- Si $x > 0$ alors par croissance stricte de la fonction racine carrée :

$$F_{Z_1^2}(x) = \mathbb{P}(Z_1^2 \leq x) = \mathbb{P}(-\sqrt{x} \leq Z_1 \leq \sqrt{x}) = \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{x}}^{\sqrt{x}} e^{-\frac{u^2}{2}} du = \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$$

  où Φ est la fonction de répartition d'une loi normale centrée réduite.

Ainsi :

$$F_{Z_1^2} : x \mapsto (\Phi(\sqrt{x}) - \Phi(-\sqrt{x}))1_{\mathbb{R}^{+*}}(x).$$

Comme Φ est de classe C^1, $F_{Z_1^2}$ est de classe C^1 sur $\mathbb{R}_+^*$ et continue sur $\mathbb{R}$.
Ainsi Z_1^2 est à densité et on a pour tout $t \neq 0$:

$$\begin{aligned} F'_{Z_1^2}(t) &= \frac{1}{2\sqrt{t}}(\Phi'(\sqrt{t}) + \Phi'(-\sqrt{t}))1_{\mathbb{R}^{+*}}(t) \\ &= \frac{1}{\sqrt{2\pi t}} e^{-t/2} 1_{\mathbb{R}^{+*}}(t) \end{aligned}$$

Ainsi $f_1 : t \mapsto \dfrac{1}{\sqrt{2\pi t}} e^{-t/2} 1_{\mathbb{R}^{+*}}(t)$ est bien une densité de Z_1^2.

16. L'espérance de Z_1^2 n'est autre que le moment d'ordre 2 de Z_1 (qui existe bien d'après le cours). D'après la formule de Koenig-Huygens :

$$\mathbb{E}(Z_1^2) = \mathbb{V}(Z_1) + \mathbb{E}(Z_1)^2 = 1.$$

Par le théorème de transfert, Z_1^2 possède une variance si et seulement si $\displaystyle\int_{-\infty}^{+\infty} t^4 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt$ converge absolument.

Comme l'intégrande est positive et paire, il suffit de montrer la convergence de $\displaystyle\int_0^{+\infty} t^4 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt$.

Les $u : t \mapsto t^3$ et $v : t \mapsto -e^{-\frac{t^2}{2}}$ sont de classe C^1 et $\lim\limits_{t\to+\infty} u(t)v(t) = 0$ par croissance comparée. Donc par intégration par partie, les intégrales :

$$\int_0^{+\infty} t^4 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt \quad \text{et} \quad \int_0^{+\infty} -3t^2 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt$$

sont de même nature et en cas de convergence on a :

$$\int_0^{+\infty} t^4 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt = \lim_{t\to+\infty} u(t)v(t) - u(0)v(0) - \int_0^{+\infty} -3t^2 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt.$$

La dernière intégrale converge (moment d'ordre 2 d'un loi normale) donc l'autre aussi et on a :

$$\begin{aligned}\int_0^{+\infty} t^4 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt &= \lim_{t\to+\infty} u(t)v(t) - u(0)v(0) - \int_0^{+\infty} -3t^2 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt \\ &= 3\int_0^{+\infty} t^2 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt.\end{aligned}$$

Ainsi Z_1^2 possède un moment d'ordre 2 et par parité :

$$\begin{aligned}\mathbb{E}(Z_1^4) = 2\int_0^{+\infty} t^4 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt &= 3\times 2\int_0^{+\infty} t^2 \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}dt \\ &= 3E(Z_1^2) \\ &= 3.\end{aligned}$$

Par Koenig-Huygens :

$$\mathbb{V}(Z_1^2) = \mathbb{E}(Z_1^4) - E(Z_1^2)^2 = 3 - 1^2 = 2.$$

17. On considère la fonction $h : x \mapsto \displaystyle\int_0^x \frac{1}{\sqrt{t(x-t)}}dt$ et on admet que h est correctement définie sur $\mathbb{R}^{+*}$.

Soit $x \in \mathbb{R}_+^*$. On effectue le changement de variable $t = xu$ qui est licite car $u \mapsto xu$ est de classe C^1 et strictement monotone de $[0,1]$ dans $[0,x]$. On a donc :

$$h(x) = \int_0^1 \frac{1}{\sqrt{xu(x-xu)}}xdu = \int_0^1 \frac{1}{\sqrt{u(1-u)}}du = h(1)$$

car $x > 0$.

Ainsi h est constante sur $\mathbb{R}^{+*}$ et on note C cette constante (on ne cherchera pas à la déterminer).

18. On rappelle que si deux variables aléatoires U_1 et U_2 sont indépendantes et de densités respectives f_{U_1} et f_{U_2} alors $U_1 + U_2$ est une variable aléatoire à densité de densité :

$$x \mapsto \int_{-\infty}^{+\infty} f_{U_1}(x-t) f_{U_2}(t) dt.$$

(a) On sait que Z_1^2 et Z^2 sont indépendantes (lemme des coalitions) et de densité $f_1 : t \mapsto \dfrac{1}{\sqrt{2\pi t}} e^{-t/2} 1_{\mathbb{R}^{+*}}(t)$.

D'après la formule de convolution, T est donc à densité de densité :

$$f_T : x \mapsto \int_{-\infty}^{+\infty} f_1(x-t) f_1(t) dt.$$

Or, pour tout $x \in \mathbb{R}$ on a :

$$\begin{aligned}
\int_{-\infty}^{+\infty} f_1(x-t) f_1(t) dt &= \frac{1}{2\pi} \int_0^{+\infty} e^{-t/2} e^{-(x-t)/2} \frac{1}{\sqrt{t(x-t)}} 1_{\mathbb{R}_+^*}(x-t) dt \\
&= \frac{e^{-x/2}}{2\pi} \int_0^{+\infty} \frac{1}{\sqrt{t(x-t)}} 1_{\mathbb{R}_+^*}(x-t) dt.
\end{aligned}$$

— Si $x \leq 0$ alors pour tout $t \in [0, +\infty[$, $x - t \leq 0$ donc $1_{\mathbb{R}_+^*}(x-t) = 0$: l'intégrande est alors la fonction nulle. Ainsi $f_T(x) = 0$.

— Si $x > 0$ alors pour tout $t \in [0, +\infty[$, $x - t > 0 \iff t < x$ donc :

$$f_T(x) = \frac{e^{-x/2}}{2\pi} \int_0^{x} \frac{1}{\sqrt{t(x-t)}} 1_{\mathbb{R}_+^*}(x-t) dt = \frac{Ce^{-x/2}}{2\pi}.$$

Finalement :

$$f_T : x \longmapsto \frac{Ce^{-x/2}}{2\pi} 1_{R_+^*}(x) = \begin{cases} \dfrac{Ce^{-x/2}}{2\pi} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

(b) Comme f_T est une densité, on doit avoir :

$$\int_{-\infty}^{+\infty} f_T(x) dx = 1.$$

Or

$$\begin{aligned}
1 = \int_{-\infty}^{+\infty} f_T(x) dx &= \frac{C}{2\pi} \int_0^{+\infty} e^{-x/2} dx \\
&= \frac{C}{2\pi} \lim_{A \to +\infty} \int_0^{A} e^{-x/2} dx \\
&= \frac{C}{2\pi} \lim_{A \to +\infty} \left[-2e^{-x/2}\right]_0^A \\
&= \frac{C}{\pi}.
\end{aligned}$$

Ainsi $C = \pi$.

D'après la question **16** et par linéarité de l'espérance :

$$\mathbb{E}(T) = \mathbb{E}(Z_1^2) + \mathbb{E}(Z_2^2) = 2.$$

D'après la question **16** et par indépendance :

$$\mathbb{V}(T) = \mathbb{V}(Z_1^2) + \mathbb{V}(Z_2^2) = 4.$$

4 Annexe des fonctions Python utiles

Dans le module `matplotlib.pyplot` importé sous l'alias `plt` :

```
plt.plot(X,Y)
```

prend en entrée deux vecteurs ou deux listes de même taille, et réalise le tracé des points d'abscisses prises dans X et d'ordonnées prises dans Y. On utilise `plt.show()` pour afficher le tracé.

Dans le module `numpy` importé sous l'alias `np`

```
np.linspace(a,b,n)
```

crée une matrice unidimensionnelle de n coefficients régulièrement espacés dans l'intervalle $[a, b]$.

Dans le module `numpy.random` importé sous l'alias `rd` :

```
rd.binomial(n,p)
```

simule une variable aléatoire suivant une loi de binomiale de paramètres n et p.