ÉCOLES NORMALES SUPÉRIEURES

ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES

ÉCOLE DES MINES DE PARIS

ÉCOLES CENTRALE

CONCOURS D'ADMISSION SESSION 2025

FILIÈRE BCPST

COMPOSITION DE MATHÉMATIQUES

Épreuve commune aux ENS de Lyon, Paris, Paris-Saclay, à l'ENPC, aux Écoles centrale et aux Mines Paris

Durée : 4 heures

L'utilisation des calculatrices n'est pas autorisée pour cette épreuve.
Il est recommandé de lire les énoncés attentivement et patiemment. Il est demandé de veiller au soin de la présentation, ainsi qu'à la rigueur et à la concision des raisonnements.

***

Début de l'épreuve

Le sujet porte sur l'estimation de l'espérance de données structurées. Il comprend 4 pages numérotées de 1 à 4, et comporte 4 parties.
La Partie I porte sur les variables aléatoires gaussiennes.
La Partie II établit un résultat d'algèbre d'égalité matricielle.
La Partie III s'intéresse à l'estimation de l'espérance d'une suite de variables aléatoires gaussiennes.
La Partie IV propose un modèle d'évolution structurée sur un arbre.
Les résultats de la Partie I (qui pourront être admis) pourront être réutilisés dans tout le reste du sujet. Les Parties II, III et IV sont indépendantes entre elles, sauf mention explicite du contraire.
Il est recommandé de lire attentivement le sujet. Il est demandé de veiller au soin de la présentation, ainsi qu'à la rigueur et à la concision des raisonnements.

Notations et rappels

— On note $\mathbb{R}$ l'ensemble des réels, et $\mathbb{R}_+$, $\mathbb{R}^*$, $\mathbb{R}_+^*$, respectivement, les réels positifs, les réels non nuls, et les réels strictement positifs. $\mathbb{N}$ et $\mathbb{N}^*$ désignent respectivement l'ensemble des entiers naturels et l'ensemble des entiers naturels non nuls.

— Pour tout $(i,j) \in \mathbb{N}^2$, $i \leq j$, on note $[\![i\,,j]\!] = \{k \in \mathbb{N} : i \leq k \leq j\}$ l'ensemble des entiers dans l'intervalle entre i et j (inclus).

— Pour tout $k, n \in \mathbb{N}^*$, on note $\mathcal{M}_{n,k}(\mathbb{R})$ l'ensemble des matrices à coefficients dans $\mathbb{R}$ avec n lignes et k colonnes, et $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées de taille n à coefficients réels. $I_n \in \mathcal{M}_n(\mathbb{R})$ désigne la matrice identité de $\mathcal{M}_n(\mathbb{R})$, $\mathbb{1}_n \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne de taille n ne comportant que des 1, et $0_n \in \mathcal{M}_{n,1}(\mathbb{R})$ le vecteur colonne ne comportant que des 0. Pour toute matrice $M \in \mathcal{M}_{n,k}(\mathbb{R})$, on note $\ker(M)$ son noyau, et $M^T \in \mathcal{M}_{k,n}(\mathbb{R})$ sa transposée. Dans toute la suite, on identifie $\mathbb{R}^n$ avec $\mathcal{M}_{n,1}(\mathbb{R})$ l'ensemble des vecteurs colonnes de $\mathbb{R}$.

— On note $X \sim \mathcal{N}(\mu, \sigma^2)$ une variable aléatoire gaussienne X à valeurs dans l'ensemble des réels $\mathbb{R}$, d'espérance $\mathbb{E}[X] = \mu$, $\mu \in \mathbb{R}$, et de variance $\mathbb{V}[X] = \sigma^2$, avec $\sigma \in \mathbb{R}_+^*$ l'écart type. On rappelle la fonction de densité associée :

$$\phi_{\mu,\sigma} : \begin{cases} \mathbb{R} \to \mathbb{R}_+ \\ x \mapsto \phi_{\mu,\sigma}(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right), \end{cases}$$

avec $\int_{-\infty}^{+\infty} \phi_{\mu,\sigma}(x)dx$ une intégrale convergente, et telle que $\int_{-\infty}^{+\infty} \phi_{\mu,\sigma}(x)dx = 1$.

Partie I : Quelques propriétés des variables aléatoires gaussiennes

Tous les résultats de la **Partie I** pourront être utilisés dans la suite du problème.

1. Soit $X \sim \mathcal{N}(\mu, \sigma^2)$ une variable aléatoire gaussienne d'espérance $\mu \in \mathbb{R}$ et de variance $\sigma^2 \in \mathbb{R}_+^*$.

(a) Soient $a \in \mathbb{R}^*$ et $b \in \mathbb{R}$. Montrez que la variable aléatoire $Y = aX + b$ suit une loi gaussienne dont on précisera l'espérance et la variance.

(b) Montrez que la variable $Z = \frac{X-\mu}{\sigma}$ suit une loi normale centrée réduite : $Z \sim \mathcal{N}(0,1)$.

2. Soit $\alpha \in \,]0,1[$. On note $\phi_{0,1}$ la densité de la loi normale centrée réduite, définie au début du sujet.

(a) Montrez qu'il existe un unique réel $u(\alpha)$ tel que $\int_{-\infty}^{u(\alpha)} \phi_{0,1}(x)dx = \alpha$. Montrez que l'application $u :]0,1[\to \mathbb{R}$ qui à $\alpha \in \,]0,1[$ associe $u(\alpha)$ est continue et strictement croissante.

(b) Montrez : $u(1-\alpha) = -u(\alpha)$.

3. Soient $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ deux variables aléatoires gaussiennes indépendantes d'espérances $\mu_1, \mu_2 \in \mathbb{R}$, et de variances $\sigma_1^2, \sigma_2^2 \in \mathbb{R}_+^*$, avec $\sigma_1 \leq \sigma_2$. Soit $\delta > 0$.
Montrez : $\mathbb{P}[|X_1 - \mu_1| \geq \delta] \leq \mathbb{P}[|X_2 - \mu_2| \geq \delta]$. Interprétez (vous pourrez vous aider d'un schéma représentant les deux densités).

4. Dans cette question, on cherche à démontrer le résultat suivant :

Lemme 1. *Soit $n \in \mathbb{N}^*$ et $a \in \mathbb{R}^n$ un vecteur de réels non identiquement nul ($a \neq 0_n$). Pour tout $i \in [\![1\,,n]\!]$, soit $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ une variable aléatoire gaussienne d'espérance $\mu_i \in \mathbb{R}$ et de variance $\sigma_i^2 \in \mathbb{R}_+^*$. On suppose que ces variables sont mutuellement indépendantes. Alors la variable aléatoire $Y = \sum_{i=1}^n a_i X_i$ est une variable gaussienne, d'espérance $\mu = \sum_{i=1}^n a_i \mu_i$ et de variance $\sigma^2 = \sum_{i=1}^n a_i^2 \sigma_i^2$, i.e. $Y \sim \mathcal{N}(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$.*

(a) Soient $X_1 \sim \mathcal{N}(0, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(0, \sigma_2^2)$ deux variables aléatoires gaussiennes centrées indépendantes d'espérance 0, de variances σ_1^2 et σ_2^2 dans $\mathbb{R}_+^*$, et de fonctions de densité f_1 et f_2. On définit la variable aléatoire $Y = X_1 + X_2$.

i. Donnez l'espérance et la variance de Y.

ii. Montrez que, pour tout couple de réels (x, y), on a :

$$f_1(y-x)f_2(x) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{y^2}{2(\sigma_1^2+\sigma_2^2)} - \frac{\left(x - \frac{y\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2}{2\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}} \right).$$

iii. En déduire que :

$$\int_{-\infty}^{+\infty} f_1(y-x)f_2(x)dx = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} \exp\left(-\frac{y^2}{2(\sigma_1^2+\sigma_2^2)} \right).$$

iv. Conclure la démonstration du lemme 1 dans le cas de deux variables ($n = 2$).

(b) Démontrez le lemme 1 dans le cas général.

(c) Démontrez le corollaire suivant :

Corollaire 1. *Soit $n, k \in \mathbb{N}^*$, $M \in \mathcal{M}_{n,k}(\mathbb{R})$, et $a \in \mathbb{R}^n$ tel que $a \notin \ker(M^T)$. Pour tout $i \in [\![1\,,k]\!]$, soit $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ une variable aléatoire gaussienne d'espérance $\mu_i \in \mathbb{R}$ et de variance $\sigma_i^2 \in \mathbb{R}_+^*$. On suppose que ces variables sont mutuellement indépendantes. On note $X = (X_1, \ldots, X_k) \in \mathbb{R}^k$ le vecteur colonne contenant toutes les variables $(X_i)_{i \in [\![1,k]\!]}$, $\mu = (\mu_1, \ldots, \mu_k) \in \mathbb{R}^k$ le vecteur colonne de toutes les espérances, et $\Sigma \in \mathcal{M}_k(\mathbb{R})$ la matrice diagonale telle que $\Sigma_{ii} = \sigma_i^2$ pour tout $i \in [\![1\,,k]\!]$. Alors $a^T M X \sim \mathcal{N}(a^T M \mu, a^T M \Sigma M^T a)$.*

Partie II : Identité matricielle de Woodbury

Dans cette partie, on établit le résultat suivant :

Lemme 2 (identité de Woodbury)**.** *Soient $k, n \in \mathbb{N}^*$, $A \in \mathcal{M}_n(\mathbb{R})$ et $C \in \mathcal{M}_k(\mathbb{R})$ deux matrices inversibles, et soient $U_1 \in \mathcal{M}_{n,k}(\mathbb{R})$ et $U_2 \in \mathcal{M}_{k,n}(\mathbb{R})$ deux matrices quelconques. Alors la matrice $A + U_1 C U_2$ est inversible si et seulement si la matrice $C^{-1} + U_2 A^{-1} U_1$ est inversible, et alors :*

$$(A + U_1 C U_2)^{-1} = A^{-1} - A^{-1} U_1 (C^{-1} + U_2 A^{-1} U_1)^{-1} U_2 A^{-1}.$$

1. Soient $f : \mathbb{R}^n \to \mathbb{R}^k$ et $g : \mathbb{R}^k \to \mathbb{R}^n$ deux applications linéaires admettant pour représentation matricielle dans les bases canoniques, respectivement, CU_2 et $-A^{-1}U_1$.

(a) Soit $x \in \ker(A + U_1 C U_2)$. Montrez que $f(x) \in \ker(C^{-1} + U_2 A^{-1} U_1)$, et que $g \circ f(x) = x$.

(b) Soit $x \in \ker(C^{-1} + U_2 A^{-1} U_1)$. Montrez que $g(x) \in \ker(A + U_1 C U_2)$, et que $f \circ g(x) = x$.

(c) Montrez que f est un isomorphisme de $\ker(A + U_1 C U_2)$ dans $\ker(C^{-1} + U_2 A^{-1} U_1)$.

2. Démontrez le lemme 2.

3. Soit $t \in \mathbb{R}_+$ et $A \in \mathcal{M}_n(\mathbb{R})$ symétrique inversible, telle que $\forall x \in \mathbb{R}^n, x \neq 0 \implies x^T A x > 0$.

(a) Montrez que $V = A + t\mathbb{1}_n \mathbb{1}_n^T$ est inversible, et donnez son inverse.

(b) Soit $y \in \mathbb{R}^n$ un vecteur colonne. On note $p_A = \mathbb{1}_n^T A^{-1} \mathbb{1}_n$ et $\mu_A = \mathbb{1}_n^T A^{-1} y$. Montrez que :

$$\mathbb{1}_n^T V^{-1} \mathbb{1}_n = \frac{p_A}{1 + t p_A} \quad \text{et} \quad \mathbb{1}_n^T V^{-1} y = \frac{\mu_A}{1 + t p_A}.$$

(c) Soit $y \in \mathbb{R}^n$ un vecteur colonne. On suppose de plus que A est diagonale, avec $A_{ii} = \frac{1}{r_i}$, $r_i > 0$ pour $1 \leq i \leq n$. On note $r = (r_1, \ldots, r_n) \in \mathbb{R}^n$ le vecteur colonne contenant tous les r_i, $1 \leq i \leq n$. Montrez :

$$(\mathbb{1}_n^T V^{-1} \mathbb{1}_n)^{-1} V^{-1} \mathbb{1}_n = \left(\mathbb{1}_n^T r\right)^{-1} r.$$

Partie III : Estimation d'espérance

Dans cette partie, on cherche à estimer l'espérance d'une suite de variables aléatoires dans différentes conditions. La question 1 établit des résultats d'analyse et ne fait pas appel à des notions de probabilité.

1. Soit $n \in \mathbb{N}^*$, soit $a \in \mathbb{R}^n$ tel que $a^T \mathbb{1}_n = 1$, et soit $r \in (\mathbb{R}_+^*)^n$.

(a) Montrez que $\sum_{i=1}^n a_i^2 \geq \frac{1}{n}$, et que l'égalité est réalisée pour le vecteur $b = \frac{1}{n}\mathbb{1}_n$.
[Indication : on pourra remarquer que $\sum_{i=1}^n a_i^2 = \sum_{i=1}^n (a_i - b_i + b_i)^2$, et développer.]

(b) Montrez que $\sum_{i=1}^n r_i^{-1} a_i^2 \geq \left(\sum_{i=1}^n r_i\right)^{-1}$, et que l'égalité est réalisée pour $b = \left(\sum_{i=1}^n r_i\right)^{-1} r$.

(c) Soit m un entier tel que $m \geq n$, et $M \in \mathcal{M}_{m,n}(\mathbb{R})$ une matrice de rang égal à n.

i. Montrez que $V = M^T M$ est symétrique, et telle que : $\forall x \in \mathbb{R}^n, x \neq 0_n \implies x^T V x > 0$.

ii. Montrez que V est diagonalisable, et que toutes ses valeurs propres sont strictement positives.

iii. Montrez que V est inversible, et que : $\forall x \in \mathbb{R}^n, x \neq 0_n \implies x^T V^{-1} x > 0$.

iv. Montrez que $a^T V a \geq (\mathbb{1}_n^T V^{-1} \mathbb{1}_n)^{-1}$, et que l'égalité est réalisée pour le vecteur $b = (\mathbb{1}_n^T V^{-1} \mathbb{1}_n)^{-1} V^{-1} \mathbb{1}_n$.
[Indication : on pourra remarquer que $a^T V a = (a - b + b)^T V (a - b + b)$, et développer.]

v. Montrez que les deux questions précédentes (1-a, 1-b) sont des cas particuliers de ce résultat.

2. Dans cette question, on considère une suite de variables aléatoires mutuellement indépendantes $(\epsilon_i)_{i\geq 1}$, avec, pour tout $i \geq 1$, $\epsilon_i \sim \mathcal{N}(0,1)$ et $Y_i = \mu + \sigma\epsilon_i$, où $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+^*$. Pour tout $n \geq 1$, on note $M_n = \frac{1}{n}\sum_{i=1}^n Y_i$.

(a) Soit $n \geq 1$. Donnez la loi de M_n.

(b) On dit qu'une quantité E_n est un *estimateur linéaire sans biais* de μ si il existe $a \in \mathbb{R}^n$ un vecteur déterministe tel que $E_n = \sum_{i=1}^n a_i Y_i$ (linéaire) et que $\mathbb{E}[E_n] = \mu$ (sans biais). Montrez que M_n est un estimateur linéaire sans biais de μ.

(c) Soit $n \geq 1$, $F_n = \{x \in \mathbb{R}^n : x^T \mathbb{1}_n = 1\}$, et $Q_n : F_n \to \mathbb{R}$ l'application qui à $a \in F_n$ associe $Q_n(a) = \sum_{i=1}^n a_i Y_i$. Soit $a \in F_n$. Quelle est la loi de $Q_n(a)$?

(d) Montrez que, pour tout $a \in F_n$, $Q_n(a)$ est un estimateur linéaire sans biais de μ.

(e) Montrez que, pour tout $a \in F_n$, $\mathbb{V}[Q_n(a)] \geq \mathbb{V}[M_n]$. Interprétez.

(f) Soit $a \in F_n$ et $\delta \in \mathbb{R}_+^*$. Montrez : $\mathbb{P}[|M_n - \mu| \geq \delta] \leq \mathbb{P}[|Q_n(a) - \mu| \geq \delta]$. Interprétez.

(g) Soit $\delta \in \mathbb{R}_+^*$ un réel strictement positif. Montrez que $\lim_{n\to\infty} \mathbb{P}[|M_n - \mu| \geq \delta] = 0$. Que peut-on en déduire sur l'estimateur M_n lorsque la quantité de données devient importante ?

3. Dans cette question, on considère de même une suite de variables aléatoires mutuellement indépendantes $(\epsilon_i)_{i\geq 1}$, avec, pour tout $i \geq 1$, $\epsilon_i \sim \mathcal{N}(0,1)$ et $Y_i = \mu + \frac{\sigma}{\sqrt{r_i}}\epsilon_i$, avec $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+^*$ et $r_i \in \mathbb{R}_+^*$ pour tout i, tels que pour tout $n \geq 1$, $\sum_{i=1}^n \frac{1}{r_i} \leq n$. Comme précédemment, pour tout $n \geq 1$, $M_n = \frac{1}{n}\sum_{i=1}^n Y_i$, et on note $Q_n(r) = \left(\sum_{i=1}^n r_i\right)^{-1} \sum_{i=1}^n r_i Y_i$.

(a) Donnez la loi de $Q_n(r)$. Soit $\delta \in \mathbb{R}_+^*$. A-t-on que $\lim_{n\to\infty} \mathbb{P}[|Q_n(r) - \mu| \geq \delta] = 0$?

(b) En supposant que les poids r_i sont connus pour tout $1 \leq i \leq n$, entre M_n et $Q_n(r)$, quel estimateur de μ préférez-vous utiliser ? Interprétez la forme de l'estimateur $Q_n(r)$.

4. Dans cette question, on considère toujours une suite de variables aléatoires mutuellement indépendantes $(\epsilon_i)_{i\geq 0}$, avec, pour tout $i \geq 0$, $\epsilon_i \sim \mathcal{N}(0,1)$ (où i part bien de 0 ici). Pour tout $i \geq 1$, on définit : $Y_i = \mu + \sigma\left(\frac{1}{\sqrt{r_0}}\epsilon_0 + \frac{1}{\sqrt{r_i}}\epsilon_i\right)$, avec $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+^*$ et $r_i \in \mathbb{R}_+^*$ pour tout $i \geq 0$, avec $\sum_{i=1}^n \frac{1}{r_i} \leq n$ pour tout $n \geq 1$.
Soit $n \geq 1$. Donnez la loi de $M_n = \frac{1}{n}\sum_{i=1}^n Y_i$. Soit $\delta \in \mathbb{R}_+^*$. A-t-on que $\lim_{n\to\infty} \mathbb{P}[|M_n - \mu| \geq \delta] = 0$? Interprétez.

5. Dans cette question, on fixe $m \geq n \geq 1$, et on considère toujours une suite de variables aléatoires mutuellement indépendantes $(\epsilon_i)_{1\leq i\leq m}$, avec, pour tout $i \in [\![1, m]\!]$, $\epsilon_i \sim \mathcal{N}(0,1)$ et on note $E = (\epsilon_1, \ldots, \epsilon_m) \in \mathbb{R}^m$ le vecteur colonne de composé de toutes ces variables. Soit $M \in \mathcal{M}_{n,m}(\mathbb{R})$ une matrice réelle de rang égal à n. On définit $Y = \mu\mathbb{1}_n + \sigma M E$, avec $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$.

(a) Soit $a \in \mathbb{R}^n$ un vecteur colonne tel que $a^T \mathbb{1}_n = 1$. Donnez la loi de $Q_n(a) = a^T Y$.

(b) Proposez un estimateur pour l'espérance μ, qui ne dépend que de Y et de M.
(c) Montrez que les questions 2, 3, et 4 peuvent être exprimées dans le cadre général présenté ici.

6. On se replace dans le cadre de la **question 4** de cette partie. En utilisant les résultats de la **Partie II**, proposez un estimateur de μ dans ce cas, et donnez-en l'expression en fonction de r. Donnez sa loi. Que dire de sa convergence ? Interprétez.

Partie IV : Évolution structurée sur un arbre

Dans cette partie, on considère un système simple avec deux espèces notées A et B, qui ont un ancêtre commun noté AB. On suppose que l'ancêtre dérive lui-même d'une espèce ancestrale notée R (voir la figure 1). On suppose qu'un temps t_A s'est écoulé entre l'époque où vivait l'espèce AB et l'époque où vit l'espèce A, et de même qu'un temps t_B s'est écoulé entre l'époque où vivait l'espèce AB et l'époque où vit l'espèce B ($t_A = t_B$ si les espèces A et B sont contemporaines). Dans chaque espèce, on mesure un trait quantitatif représentant une caractéristique continue de l'espèce, par exemple le logarithme de la masse corporelle moyenne au sein de l'espèce. On note X_{AB}, X_A, X_B les variables aléatoires réelles représentant la valeur du trait dans, respectivement, les espèces AB, A et B. Soient ϵ_{AB}, ϵ_A et ϵ_B trois variables aléatoires gaussiennes centrées et indépendantes deux à deux, telles que $\epsilon_i \sim \mathcal{N}(0, \sigma^2 t_i)$ pour $i \in \{AB, A, B\}$, avec $\sigma \in \mathbb{R}_+^*$ un réel strictement positif. Le modèle est décrit par :

$$X_{AB} = \mu + \epsilon_{AB}; \quad X_A = X_{AB} + \epsilon_A; \quad X_B = X_{AB} + \epsilon_B,$$

avec $\mu \in \mathbb{R}$ représentant le trait dans l'espèce ancestrale R.

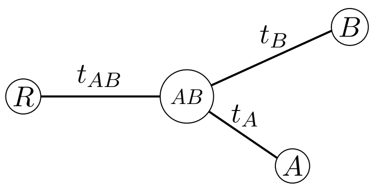

Figure 1 – Un arbre phylogénétique simple. L'espèce AB est l'ancêtre des espèces A et B. Un temps t_A (respectivement, t_B) s'est écoulé entre l'ancêtre AB et l'espèce A (resp., B). Un temps t_{AB} s'est écoulé entre la racine R et l'espèce AB.

1. **(a)** Donnez les lois de X_{AB}, X_A et X_B.
(b) On note $Y = (X_A, X_B) \in \mathbb{R}^2$ le vecteur colonne contenant les valeurs du trait pour les deux espèces A et B, et $E = (\epsilon_{AB}, \epsilon_A, \epsilon_B) \in \mathbb{R}^3$ le vecteur colonne des incréments. Montrez que l'on peut écrire $Y = \mu \mathbb{1}_2 + M_2 E$, avec

$$M_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

(c) Soit $a \in \mathbb{R}^2$ un vecteur colonne déterministe tel que $a^T \mathbb{1}_2 = 1$. Montrez que $Q_2(a) = a^T Y$ est une variable gaussienne, d'espérance μ, et de variance $\sigma^2 a^T M_2 D_{AB} M_2^T a$, où $D_{AB} \in \mathcal{M}_3(\mathbb{R})$ est la matrice diagonale telle que $[D_{AB}]_{11} = t_{AB}$, $[D_{AB}]_{22} = t_A$ et $[D_{AB}]_{33} = t_B$.
(d) Montrez l'égalité :

$$M_2 D_{AB} M_2^T = t_{AB} \mathbb{1}_2 \mathbb{1}_2^T + \begin{pmatrix} t_A & 0 \\ 0 & t_B \end{pmatrix}.$$

2. Justifiez que ce modèle peut être utilisé pour modéliser l'évolution du trait quantitatif pour un groupe d'espèces liées par un arbre phylogénétique. En particulier, on dit que les ϵ_i pour $i \in \{AB, A, B\}$ représentent les "incréments" du trait sur une branche. Justifiez ce terme. Quelle est la signification de μ dans ce modèle ? Pourquoi est-il intéressant d'en trouver un estimateur ?

3. Montrez que ce modèle est un cas particulier d'un des modèles étudiés dans la **Partie III**. En déduire un estimateur de μ, et donnez son expression. Interprétez. Si l'on avait accès à la mesure du trait X_C, X_D, ... de plusieurs autres espèces C, D, etc. descendant directement de AB, pourrait-on mieux estimer μ ?

Fin de l'épreuve