



**PRÉPASIMPLE**

[prepasimple.fr](http://prepasimple.fr)

---

**HEC — 2012**

*Corrigé de mathématiques approfondies*

---

**Mathématiques approfondies — Prépa ECG**

*Annales de concours · Corrigé*

## Corrigé

---

Rappels et conventions utilisés dans tout le corrigé.

On identifie  $\mathbb{R}^n$  à  $\mathcal{M}_{n,1}(\mathbb{R})$ ; le produit scalaire canonique s'écrit  $\langle X, X' \rangle = {}^t X X'$  et la norme associée  $\|X\|$ . On note  $\mathbf{1}$  la colonne dont tous les coefficients valent 1 et  $\mathbf{x}$  la colonne de coefficients  $x_1, \dots, x_n$ : ce sont les deux colonnes de  $M$ , et  $\mathcal{F} = \text{Vect}(\mathbf{x}, \mathbf{1})$ .

Une matrice  $Q$  de  $\mathcal{M}_n(\mathbb{R})$  est orthogonale si  ${}^t Q Q = I$ ; elle est alors inversible d'inverse  ${}^t Q$ , donc  $Q {}^t Q = I$  aussi. C'est le cas d'une matrice de passage entre deux bases orthonormales.

Règles de calcul sur les variances et covariances (admisses par l'énoncé pour les variables à densité): bilinéarité et symétrie de la covariance,  $\text{Cov}(J, J) = V(J)$ ,  $\text{Cov}(J, c) = 0$  pour une constante  $c$ , et  $\text{Cov}(J_1, J_2) = 0$  si  $J_1$  et  $J_2$  sont indépendantes. En particulier, si  $J_1, \dots, J_n$  sont mutuellement indépendantes et  $c_1, \dots, c_n$  réels,

$$V\left(\sum c_i J_i\right) = \sum c_i^2 V(J_i)$$

et

$$\text{Cov}\left(\sum c_i J_i, \sum d_i J_i\right) = \sum c_i d_i V(J_i).$$

Lois normales: si  $N \hookrightarrow \mathcal{N}(m, s^2)$  et si  $c \neq 0, d$  réels, alors  $cN + d \hookrightarrow \mathcal{N}(cm + d, c^2 s^2)$ ; une somme de variables normales mutuellement indépendantes suit une loi normale (d'espérance et de variance les sommes des espérances et des variances).

Lois gamma: la loi  $\gamma(\nu)$  ( $\nu > 0$ ) a pour espérance et variance  $\nu$ ; si deux variables indépendantes suivent les lois  $\gamma(\nu_1)$  et  $\gamma(\nu_2)$ , leur somme suit la loi  $\gamma(\nu_1 + \nu_2)$ .

## Partie I. Quelques résultats statistiques et algébriques

1. a)  $s_x^2$  est une somme de carrés divisée par  $n$ , donc positive ; on montre qu'elle n'est pas nulle.

Si  $s_x^2 = 0$ , alors

$$\sum_{i=1}^n (x_i - \bar{x})^2 = 0,$$

somme de termes positifs, donc  $x_i = \bar{x}$  pour tout  $i$  : les  $x_i$  seraient tous égaux, ce qui est exclu.

### Conclusion

$$s_x^2 > 0.$$

- b) On développe et on utilise

$$\sum_{i=1}^n x_i = n\bar{x}$$

et

$$\sum_{i=1}^n y_i = n\bar{y}.$$

$$\sum_{i=1}^n (x_i - \bar{x})y_i = \sum_{i=1}^n x_i y_i - \bar{x} \sum_{i=1}^n y_i = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}$$

En appliquant cette formule à la série  $(x_i)$  à la place de  $(y_i)$ , et comme

$$\sum_{i=1}^n (x_i - \bar{x}) = n\bar{x} - n\bar{x} = 0 :$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_i - \bar{x})x_i - \bar{x} \sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i^2 - n\bar{x}^2$$

### Conclusion

les deux formules sont établies.

c) Les  $\alpha_i$  sont bien définis car  $s_x^2 > 0$ . On utilise

$$\sum_i (x_i - \bar{x}) = 0$$

et b).

$$\sum_{i=1}^n \alpha_i = \frac{1}{ns_x^2} \sum_{i=1}^n (x_i - \bar{x}) = 0$$

$$\sum_{i=1}^n \alpha_i x_i = \frac{1}{ns_x^2} \sum_{i=1}^n (x_i - \bar{x})x_i = \frac{1}{ns_x^2} \left( \sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = \frac{ns_x^2}{ns_x^2} = 1$$

$$\sum_{i=1}^n \alpha_i^2 = \frac{1}{n^2 s_x^4} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{ns_x^2}{n^2 s_x^4} = \frac{1}{ns_x^2}$$

## Conclusion

$$\sum \alpha_i = 0,$$

$$\sum \alpha_i x_i = 1$$

et

$$\sum \alpha_i^2 = \frac{1}{ns_x^2}.$$

2. a) Le rang de  $M$  est le rang de la famille de ses deux colonnes  $(\mathbf{x}, \mathbf{1})$ ; on montre que cette famille est libre.

Soient  $\lambda, \mu$  réels tels que  $\lambda \mathbf{x} + \mu \mathbf{1} = 0$ , c'est-à-dire  $\lambda x_i + \mu = 0$  pour tout  $i$ . Si  $\lambda \neq 0$ , alors  $x_i = -\frac{\mu}{\lambda}$  pour tout  $i$  : les  $x_i$  seraient tous égaux, ce qui est exclu. Donc  $\lambda = 0$ , puis  $\mu = 0$ . La famille est libre; comme  $M$  n'a que deux colonnes, son rang vaut 2.

## Conclusion

$\text{rg}(M) = 2$ . En particulier  $\dim \mathcal{F} = 2$  et l'application  $\theta \mapsto M\theta$  de  $\mathcal{M}_{2,1}(\mathbb{R})$  dans  $\mathcal{M}_{n,1}(\mathbb{R})$  est injective (son noyau est nul par le théorème du rang), d'image  $\mathcal{F}$ .

- b) On calcule le produit, puis on justifie l'inversibilité de deux façons.

$${}^t M M = \begin{pmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{pmatrix} = n \begin{pmatrix} s_x^2 + \bar{x}^2 & \bar{x} \\ \bar{x} & 1 \end{pmatrix}$$

en utilisant 1. b) :

$$\sum x_i^2 = ns_x^2 + n\bar{x}^2.$$

**Méthode 1 — Noyau.** Soit  $X \in \mathcal{M}_{2,1}(\mathbb{R})$  tel que  ${}^tMMX = 0$ . Alors  ${}^tX {}^tMMX = {}^t(MX)(MX) = \|MX\|^2 = 0$ , donc  $MX = 0$ , puis  $X = 0$  par injectivité (question a)). La matrice carrée  ${}^tMM$  a un noyau nul, elle est donc inversible.

**Méthode 2 — Critère pour les matrices d'ordre deux.** Une matrice  $\begin{pmatrix} p & q \\ r & t \end{pmatrix}$  est inversible si et seulement si  $pt - qr \neq 0$ , et son inverse est alors

$$\frac{1}{pt - qr} \begin{pmatrix} t & -q \\ -r & p \end{pmatrix}.$$

Ici  $pt - qr = n^2 (s_x^2 + \bar{x}^2 - \bar{x}^2) = n^2 s_x^2 > 0$ , donc  ${}^tMM$  est inversible et :

$$({}^tMM)^{-1} = \frac{1}{ns_x^2} \begin{pmatrix} 1 & -\bar{x} \\ -\bar{x} & s_x^2 + \bar{x}^2 \end{pmatrix}$$

La première méthode est générale (elle vaut pour toute matrice  $M$  de rang égal à son nombre de colonnes) ; la seconde est propre à l'ordre deux, mais fournit l'inverse explicitement.

## Conclusion

$${}^tMM = \begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix}$$

est inversible.

3. a) On cherche à minimiser  $\|y - M\theta\|^2$ ; on présente une méthode géométrique et une méthode par l'étude d'une fonction de deux variables.

**Méthode 1 — Projection orthogonale.** Pour tout  $\theta$ ,

$$\sum_{i=1}^n u_i^2 = \|y - M\theta\|^2,$$

et  $M\theta$  décrit  $\mathcal{F}$  lorsque  $\theta$  décrit  $\mathcal{M}_{2,1}(\mathbb{R})$ . D'après la caractérisation de la projection orthogonale par minimisation de la distance,  $z \mapsto \|y - z\|$  atteint sur  $\mathcal{F}$  son minimum en un unique point,  $p_{\mathcal{F}}(y)$ , projeté orthogonal de  $y$  sur  $\mathcal{F}$ . Comme  $\theta \mapsto M\theta$  est une bijection de  $\mathcal{M}_{2,1}(\mathbb{R})$  sur  $\mathcal{F}$  (question 2. a)), il existe un unique  $\hat{\theta}$  réalisant le minimum, caractérisé par  $M\hat{\theta} = p_{\mathcal{F}}(y)$ .

---

Relation.  $M\hat{\theta} = p_{\mathcal{F}}(y)$  équivaut à  $y - M\hat{\theta} \in \mathcal{F}^\perp$ , c'est-à-dire à l'orthogonalité de  $y - M\hat{\theta}$  avec les deux colonnes de  $M$ , qui engendrent  $\mathcal{F}$  :  ${}^tM(y - M\hat{\theta}) = 0$ , soit  ${}^tMM\hat{\theta} = {}^tMy$ .

**Méthode 2 — Fonction de deux variables.** Posons

$$f(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2,$$

fonction polynomiale, donc de classe  $C^2$  sur l'ouvert convexe  $\mathbb{R}^2$ .

Points critiques.

$$\frac{\partial f}{\partial a}(a, b) = -2 \sum_{i=1}^n x_i (y_i - ax_i - b)$$

et

$$\frac{\partial f}{\partial b}(a, b) = -2 \sum_{i=1}^n (y_i - ax_i - b).$$

Matriciellement,  $\nabla f(a, b) = -2 {}^t M(y - M\theta)$ . Le gradient s'annule si et seulement si  ${}^t M M \theta = {}^t M y$ , soit, d'après 2. b),  $\theta = ({}^t M M)^{-1} {}^t M y$  : il y a un unique point critique  $\hat{\theta}$ .

Hessienne. Les dérivées secondes sont constantes :

$$\frac{\partial^2 f}{\partial a^2} = 2 \sum x_i^2,$$

$$\frac{\partial^2 f}{\partial a \partial b} = 2 \sum x_i,$$

$\frac{\partial^2 f}{\partial b^2} = 2n$ , donc la hessienne en tout point est  $H = 2 {}^t M M$ , matrice symétrique. Si  $\lambda$  est une valeur propre de  $H$  et  $X \neq 0$  un vecteur propre associé, alors  $\lambda \|X\|^2 = {}^t X H X = 2 \|M X\|^2 > 0$  car  $M X \neq 0$  (injectivité). Donc toutes les valeurs propres de  $H$  sont strictement positives, en tout point de l'ouvert convexe  $\mathbb{R}^2$ .

Conclusion de la méthode 2. D'après le critère d'extremum global sur un ouvert convexe,  $f$  admet en son point critique  $\hat{\theta}$  un minimum global. Tout point où  $f$  atteint son minimum sur l'ouvert  $\mathbb{R}^2$  est un point critique, donc le minimum n'est atteint qu'en  $\hat{\theta}$ .

La méthode géométrique explique pourquoi la solution est la projection de  $y$  sur  $\mathcal{F}$  (ce qui sert dans la suite) ; la méthode analytique est la méthode standard d'optimisation, et produit directement l'équation  ${}^t M M \hat{\theta} = {}^t M y$  (« équations normales »).

## Conclusion

le problème admet une unique solution  $\hat{\theta}$ , et  ${}^tMM\hat{\theta} = {}^tMy$ .

**b)** On résout le système  ${}^tMM\hat{\theta} = {}^tMy$ .

On a

$${}^tMy = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}.$$

Le système s'écrit :

$$\left( \sum_{i=1}^n x_i^2 \right) \hat{a} + n\bar{x}\hat{b} = \sum_{i=1}^n x_i y_i \quad \text{et} \quad n\bar{x}\hat{a} + n\hat{b} = n\bar{y}$$

La seconde équation donne  $\hat{b} = \bar{y} - \hat{a}\bar{x}$ . En reportant dans la première :

$$\hat{a} \left( \sum x_i^2 - n\bar{x}^2 \right) = \sum x_i y_i - n\bar{x}\bar{y},$$

soit, d'après 1. b),

$$\hat{a} ns_x^2 = \sum_{i=1}^n (x_i - \bar{x}) y_i,$$

puis

$$\hat{a} = \sum_{i=1}^n \frac{x_i - \bar{x}}{ns_x^2} y_i.$$

### Conclusion

$$\hat{a} = \sum_{i=1}^n \alpha_i y_i$$

$$\text{et } \hat{b} = \bar{y} - \hat{a} \bar{x}.$$

c) On exprime le projeté  $Ky = p_{\mathcal{F}}(y)$  grâce à a), pour tout  $y$ .

Pour tout  $y \in \mathcal{M}_{n,1}(\mathbb{R})$ , d'après a),  $Ky = p_{\mathcal{F}}(y) = M\hat{\theta}$  avec  $\hat{\theta} = ({}^tMM)^{-1} {}^tMy$ . Donc  $Ky = M({}^tMM)^{-1} {}^tMy$  pour toute colonne  $y$ ; en prenant pour  $y$  les vecteurs de la base canonique, on obtient l'égalité des colonnes des deux matrices.

### Conclusion

$$K = M({}^tMM)^{-1} {}^tM.$$

d) On utilise  $Ky = M\hat{\theta}$  et  $GM = 0$ .

$\hat{u} = y - M\hat{\theta} = y - Ky = (I - K)y = Gy$ . Par ailleurs,  $KM = M({}^tMM)^{-1}({}^tMM) = M$ , donc  $GM = M - KM = 0$ . Comme  $y = M\theta + u$ ,  $Gy = GM\theta + Gu = Gu$ .

### Conclusion

$$\hat{u} = Gy = Gu.$$

e) On utilise la symétrie de  $G$  et la relation  $G^2 = G$ .

Propriétés de  $G$ . D'après c),  ${}^tK = M({}^t({}^tMM)^{-1}) {}^tM = M({}^tMM)^{-1} {}^tM = K$ , car  ${}^tMM$  est symétrique et la transposée de l'inverse est l'inverse de la transposée. De plus  $K^2 = M({}^tMM)^{-1}({}^tMM)({}^tMM)^{-1} {}^tM = K$ . Donc  ${}^tG = I - {}^tK = G$  et  $G^2 = I - 2K + K^2 = I - K = G$ .

Par définition du produit matriciel,

$${}^t\hat{u}\hat{u} = \sum_{i=1}^n \hat{u}_i^2.$$

Avec d) :

$${}^t\hat{u}\hat{u} = {}^t(Gy)(Gy) = {}^ty{}^tGGy = {}^tyG^2y = {}^tyGy$$

et de même, avec  $\hat{u} = Gu$ ,  ${}^t\hat{u}\hat{u} = {}^tuG^2u = {}^tuGu$ .

### Conclusion

$${}^t\hat{u}\hat{u} = \sum_{i=1}^n \hat{u}_i^2 = {}^tyGy = {}^tuGu.$$

## Partie II. Le modèle de régression linéaire

4. a) On utilise la linéarité de l'espérance,  $E(Y_i) = ax_i + b$ , et 1. c).

Pour tout  $i$ ,  $E(Y_i) = ax_i + b + E(U_i) = ax_i + b$ . Donc :

$$E(A_n) = \sum_{i=1}^n \alpha_i(ax_i + b) = a \sum_{i=1}^n \alpha_i x_i + b \sum_{i=1}^n \alpha_i = a$$

$$E(\bar{Y}_n) = a\bar{x} + b, \text{ puis } E(B_n) = E(\bar{Y}_n) - \bar{x}E(A_n) = a\bar{x} + b - a\bar{x} = b.$$

## Conclusion

$A_n$  et  $B_n$  sont des estimateurs sans biais de  $a$  et  $b$ .

b) On écrit  $A_n$  et  $B_n$  comme combinaisons linéaires des variables indépendantes  $Y_i$ .

Les  $Y_i = ax_i + b + U_i$  sont mutuellement indépendantes (chacune est fonction de la seule  $U_i$ ) et  $V(Y_i) = V(U_i) = \sigma^2$ .

$$V(A_n) = \sum_{i=1}^n \alpha_i^2 \sigma^2 = \frac{\sigma^2}{ns_x^2}$$

d'après 1. c).

Posons  $\beta_i = \frac{1}{n} - \bar{x}\alpha_i$ , de sorte que

$$B_n = \sum_{i=1}^n \beta_i Y_i.$$

Alors :

$$\sum_{i=1}^n \beta_i^2 = \sum_{i=1}^n \left( \frac{1}{n^2} - \frac{2\bar{x}}{n} \alpha_i + \bar{x}^2 \alpha_i^2 \right) = \frac{1}{n} - 0 + \frac{\bar{x}^2}{ns_x^2}$$

donc

$$V(B_n) = \sigma^2 \sum \beta_i^2 = \left( 1 + \frac{\bar{x}^2}{s_x^2} \right) \frac{\sigma^2}{n}.$$

### Conclusion

$$V(A_n) = \frac{\sigma^2}{ns_x^2} \text{ et } V(B_n) = \left(1 + \frac{\bar{x}^2}{s_x^2}\right) \frac{\sigma^2}{n}.$$

c) On utilise la bilinéarité de la covariance et l'indépendance des  $Y_i$ .

$$\begin{aligned} \text{Cov}(A_n, B_n) &= \text{Cov}\left(\sum \alpha_i Y_i, \sum \beta_i Y_i\right) = \sigma^2 \sum_{i=1}^n \alpha_i \beta_i = \sigma^2 \left(\frac{1}{n} \sum \alpha_i - \bar{x} \sum \alpha_i^2\right) \\ &= -\frac{\bar{x} \sigma^2}{ns_x^2} \end{aligned}$$

### Conclusion

$$\text{Cov}(A_n, B_n) = -\frac{\bar{x} \sigma^2}{ns_x^2}. \text{ En particulier, } A_n \text{ et } B_n \text{ ne sont pas corrélées si et seulement si } \bar{x} = 0.$$

5. On applique l'inégalité de Bienaymé-Tchebychev, les estimateurs étant sans biais, en suivant la dépendance en  $n$  des quantités de la question 4.

On suppose donnée une suite  $(x_i)_{i \geq 1}$  telle que, pour tout  $n \geq 3$ ,  $x_1, \dots, x_n$  ne soient pas tous égaux, et une suite  $(U_i)_{i \geq 1}$  de variables indépendantes vérifiant les hypothèses du modèle. On note  $\bar{x}_n$  et  $s_n^2$  la moyenne et la variance empiriques de  $x_1, \dots, x_n$ ; par hypothèse,  $\bar{x}_n \rightarrow \lambda$  et  $s_n^2 \rightarrow \mu^2 > 0$ .

Soit  $\varepsilon > 0$ . Comme  $E(A_n) = a$ , l'inégalité de Bienaymé-Tchebychev et 4. b) donnent :

$$0 \leq P(|A_n - a| \geq \varepsilon) \leq \frac{V(A_n)}{\varepsilon^2} = \frac{\sigma^2}{ns_n^2 \varepsilon^2}$$

Comme  $ns_n^2 \rightarrow +\infty$  (car  $s_n^2 \rightarrow \mu^2 > 0$ ), le majorant tend vers 0. De même :

$$0 \leq P(|B_n - b| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \left(1 + \frac{\bar{x}_n^2}{s_n^2}\right)$$

et  $1 + \frac{\bar{x}_n^2}{s_n^2} \rightarrow 1 + \frac{\lambda^2}{\mu^2}$ , donc le majorant tend vers 0. Par encadrement, les deux probabilités tendent vers 0.

### Conclusion

$(A_n)_{n \geq 3}$  converge en probabilité vers  $a$  et  $(B_n)_{n \geq 3}$  converge en probabilité vers  $b$ .

6. a) Par linéarité, avec 4. a) :  $E(\widehat{U}_i) = E(Y_i) - x_i E(A_n) - E(B_n) = ax_i + b - ax_i - b$ .

### Conclusion

pour tout  $i$ ,  $E(\widehat{U}_i) = 0$ .

b) On exprime  $\widehat{U}_i$  en fonction des  $U_j$  et de  $A_n - a$ .

Comme  $\bar{Y}_n = a\bar{x} + b + \bar{U}_n$  et  $B_n = \bar{Y}_n - A_n\bar{x}$  :

$$\begin{aligned}\widehat{U}_i &= (Y_i - \bar{Y}_n) - A_n(x_i - \bar{x}) = a(x_i - \bar{x}) + (U_i - \bar{U}_n) - A_n(x_i - \bar{x}) = (U_i - \bar{U}_n) \\ &\quad - (A_n - a)(x_i - \bar{x})\end{aligned}$$

Par ailleurs, d'après 1. c),

$$A_n - a = \sum_j \alpha_j (ax_j + b + U_j) - a = \sum_j \alpha_j U_j,$$

et

$$\sum_i (x_i - \bar{x})(U_i - \bar{U}_n) = \sum_i (x_i - \bar{x})U_i = ns_x^2 \sum_i \alpha_i U_i = ns_x^2 (A_n - a).$$

En développant le carré et en sommant :

$$\sum_{i=1}^n \widehat{U}_i^2 = \sum_{i=1}^n (U_i - \bar{U}_n)^2 - 2(A_n - a) \cdot ns_x^2(A_n - a) + (A_n - a)^2 \sum_{i=1}^n (x_i - \bar{x})^2$$

et le dernier terme vaut  $ns_x^2(A_n - a)^2$ .

### Conclusion

$$\sum_{i=1}^n \widehat{U}_i^2 = \sum_{i=1}^n (U_i - \bar{U}_n)^2 - ns_x^2(A_n - a)^2.$$

c) On calcule l'espérance de chaque terme de b).

D'après 1. b) appliqué à la série  $(U_i)$ ,

$$\sum_i (U_i - \bar{U}_n)^2 = \sum_i U_i^2 - n\bar{U}_n^2.$$

Or  $E(U_i^2) = V(U_i) = \sigma^2$  et  $E(\bar{U}_n^2) = V(\bar{U}_n) = \frac{\sigma^2}{n}$  (car  $E(\bar{U}_n) = 0$  et les  $U_i$  sont indépendantes). Donc

$$E\left(\sum_i (U_i - \bar{U}_n)^2\right) = n\sigma^2 - \sigma^2 = (n-1)\sigma^2.$$

Comme  $E(A_n) = a$ ,  $E((A_n - a)^2) = V(A_n) = \frac{\sigma^2}{ns_x^2}$ , donc  $E(ns_x^2(A_n - a)^2) = \sigma^2$ .

Par linéarité,

$$E\left(\sum_i \widehat{U}_i^2\right) = (n-1)\sigma^2 - \sigma^2 = (n-2)\sigma^2.$$

Comme  $n \geq 3$ ,  $n-2 > 0$ .

## Conclusion

$$E \left( \sum_{i=1}^n \widehat{U}_i^2 \right) = (n-2)\sigma^2,$$

et

$$\frac{1}{n-2} \sum_{i=1}^n \widehat{U}_i^2$$

est un estimateur sans biais de  $\sigma^2$ .

## Partie III. Hypothèse de normalité et prévision

Lemme L (combinaisons linéaires des  $U_i$ ). Si  $c_1, \dots, c_q$  sont des réels non tous nuls et si  $V_1, \dots, V_q$  sont mutuellement indépendantes de loi  $\mathcal{N}(0, \sigma^2)$ , alors pour tout réel  $d$ ,

$$d + \sum_{j=1}^q c_j V_j$$

suit la loi

$$\mathcal{N} \left( d, \sigma^2 \sum_j c_j^2 \right),$$

de variance non nulle.

Démonstration. Soit  $J = \{j : c_j \neq 0\}$ , non vide. Pour  $j \in J$ ,  $c_j V_j \hookrightarrow \mathcal{N}(0, c_j^2 \sigma^2)$ ; pour  $j \notin J$ ,  $c_j V_j = 0$ . Les variables  $c_j V_j$ ,  $j \in J$ , sont mutuellement indépendantes (coalitions), donc leur somme suit une loi normale d'espérance 0 et de variance

$$\sigma^2 \sum_{j \in J} c_j^2 = \sigma^2 \sum_j c_j^2 > 0;$$

l'ajout de la constante  $d$  translate l'espérance.

7. a) On applique le lemme L aux combinaisons linéaires des  $Y_i$ , puis on exhibe une combinaison constante des  $Y_i - \bar{Y}_n$ .

Le vecteur  $(Y_1, \dots, Y_n)$  est normal. Soit  $(\rho_1, \dots, \rho_n) \neq (0, \dots, 0)$ . Alors

$$\sum_i \rho_i Y_i = \sum_i \rho_i (ax_i + b) + \sum_i \rho_i U_i.$$

D'après le lemme L (avec  $c_i = \rho_i$ , non tous nuls, et

$$d = \sum \rho_i (ax_i + b)$$

), cette variable suit une loi normale de variance

$$\sigma^2 \sum \rho_i^2 \neq 0.$$

Le second vecteur n'est pas normal. Avec  $\rho = (1, 1, \dots, 1) \neq 0$  :

$$\sum_{i=1}^n (Y_i - \bar{Y}_n) = n\bar{Y}_n - n\bar{Y}_n = 0.$$

Cette variable est constante, de variance nulle ; elle ne suit pas une loi normale de variance non nulle.

### Conclusion

$(Y_1, \dots, Y_n)$  est normal,  $(Y_1 - \bar{Y}_n, \dots, Y_n - \bar{Y}_n)$  ne l'est pas.

- b) On écrit  $A_n, B_n$  et leurs combinaisons comme combinaisons linéaires non triviales des  $Y_i$ .

Lois.

$$A_n = \sum \alpha_i Y_i$$

avec  $(\alpha_i)$  non tous nuls (car

$$\sum \alpha_i^2 = \frac{1}{ns_x^2} > 0$$

), et

$$B_n = \sum \beta_i Y_i$$

avec  $\beta_i = \frac{1}{n} - \bar{x}\alpha_i$ , non tous nuls (car

$$\sum \beta_i = 1$$

). Comme  $(Y_1, \dots, Y_n)$  est normal,  $A_n$  et  $B_n$  suivent des lois normales, dont l'espérance et la variance ont été calculées en 4. :

$$A_n \hookrightarrow \mathcal{N}\left(a, \frac{\sigma^2}{ns_x^2}\right) \quad B_n \hookrightarrow \mathcal{N}\left(b, \left(1 + \frac{\bar{x}^2}{s_x^2}\right) \frac{\sigma^2}{n}\right)$$

Le vecteur  $(A_n, B_n)$ . Soit  $(\rho_1, \rho_2) \neq (0, 0)$ . Alors

$$\rho_1 A_n + \rho_2 B_n = \sum_i (\rho_1 \alpha_i + \rho_2 \beta_i) Y_i.$$

Si tous les coefficients  $\rho_1 \alpha_i + \rho_2 \beta_i$  étaient nuls, en sommant sur  $i$  on obtiendrait  $\rho_1 \cdot 0 + \rho_2 \cdot 1 = 0$ , donc  $\rho_2 = 0$ , puis  $\rho_1 \alpha_i = 0$  pour tout  $i$ , d'où  $\rho_1 = 0$  (les  $\alpha_i$  n'étant pas tous nuls) : contradiction. Les coefficients ne sont donc pas tous nuls, et,  $(Y_1, \dots, Y_n)$  étant normal,  $\rho_1 A_n + \rho_2 B_n$  suit une loi normale de variance non nulle.

## Conclusion

$A_n$  et  $B_n$  suivent les lois normales ci-dessus, et le vecteur  $(A_n, B_n)$  est normal.

8. a) On écrit

$$\sum \rho_i T_i$$

comme combinaison linéaire des  $U_j$  dont les coefficients ne sont pas tous nuls.

Soit  $\rho \in \mathcal{M}_{n,1}(\mathbb{R})$  non nul, de coefficients  $\rho_1, \dots, \rho_n$ . Alors

$$\sum_i \rho_i T_i = {}^t \rho T = {}^t \rho S U = {}^t ({}^t S \rho) U = \sum_j c_j U_j,$$

où  $c = {}^t S \rho$ . Comme  $S$  est inversible,  ${}^t S$  l'est aussi, donc  $c \neq 0$ . D'après le lemme L,

$$\sum_i \rho_i T_i$$

suit une loi normale de variance  $\sigma^2 \|c\|^2 \neq 0$ .

## Conclusion

le vecteur  $(T_1, \dots, T_n)$  est normal.

b) On calcule les covariances des  $T_i$ , puis on applique le résultat admis.

Pour tous  $i, j$ , par bilinéarité et indépendance des  $U_k$  :

$$\text{Cov}(T_i, T_j) = \text{Cov} \left( \sum_k S_{ik} U_k, \sum_\ell S_{j\ell} U_\ell \right) = \sigma^2 \sum_k S_{ik} S_{jk} = \sigma^2 (S^t S)_{ij}$$

(Ces covariances existent au sens de l'énoncé :  $T_i, T_j$  et  $T_i + T_j$  sont, d'après a), des

variables normales de variance non nulle, donc à densité.) Si  $S$  est orthogonale,  $S^t S = I$ , donc  $\text{Cov}(T_i, T_j) = 0$  pour  $i \neq j$ . Le vecteur  $(T_1, \dots, T_n)$  étant normal d'après a), le résultat admis s'applique.

### Conclusion

si  $S$  est orthogonale,  $T_1, \dots, T_n$  sont mutuellement indépendantes; de plus chaque  $T_i$  suit la loi

$$\mathcal{N}\left(0, \sigma^2 \sum_k S_{ik}^2\right) = \mathcal{N}(0, \sigma^2),$$

car  $(S^t S)_{ii} = 1$ .

9. a) On applique, pour chaque  $\omega$ , les résultats de la question 3 avec  $y = Y(\omega)$ .

Pour tout  $\omega \in \Omega$ ,

$$A_n(\omega) = \sum \alpha_i Y_i(\omega)$$

et  $B_n(\omega) = \bar{Y}_n(\omega) - A_n(\omega)\bar{x}$  sont, d'après 3. b), les réels  $\hat{a}$  et  $\hat{b}$  associés à la colonne  $y = Y(\omega)$ . Donc, d'après 3. d),  $\hat{U}(\omega) = Y(\omega) - M\hat{\theta} = GY(\omega)$ . Ainsi  $\hat{U} = GY = G(M\theta + U) = GU$ , car  $GM = 0$  (question 3. d)).

### Conclusion

$$\hat{U} = GU.$$

b)  $G$  est la matrice, dans la base canonique, de  $\text{id} - p_{\mathcal{F}}$ , qui est le projecteur orthogonal sur  $\mathcal{F}^\perp$ , avec  $\dim \mathcal{F} = 2$  et  $\dim \mathcal{F}^\perp = n - 2$ .

**Méthode 1 — Base orthonormale adaptée.** Soit  $(e_1, \dots, e_{n-2})$  une base orthonormale de  $\mathcal{F}^\perp$  et  $(e_{n-1}, e_n)$  une base orthonormale de  $\mathcal{F}$ . Comme  $\mathbb{R}^n = \mathcal{F}^\perp \oplus \mathcal{F}$  et que ces sous-espaces sont orthogonaux,  $(e_1, \dots, e_n)$  est une base orthonormale de  $\mathbb{R}^n$ . Le projecteur orthogonal sur  $\mathcal{F}^\perp$  envoie  $e_j$  sur  $e_j$  pour  $j \leq n-2$  et sur 0 pour  $j \geq n-1$  : sa matrice dans cette base est  $D = \text{diag}(1, \dots, 1, 0, 0)$ . Soit  $R$  la matrice de passage de la base canonique à  $(e_1, \dots, e_n)$  : c'est une matrice de passage entre bases orthonormales, donc orthogonale, et la formule de changement de base donne  $G = RDR^{-1} = RD^tR$ .

**Méthode 2 — Théorème spectral.** D'après 3. e),  $G$  est symétrique réelle; par le théorème spectral, il existe  $R$  orthogonale et  $D$  diagonale telles que  $G = RD^tR$ , les coefficients diagonaux de  $D$  étant les valeurs propres de  $G$ . Si  $GX = \lambda X$  avec  $X \neq 0$ , alors  $\lambda X = GX = G^2X = \lambda^2 X$ , donc  $\lambda \in \{0, 1\}$ . Enfin  $D = {}^tRGR$  a le même rang que  $G$  (multiplication par des matrices inversibles), soit  $\dim \text{Im } G = \dim \mathcal{F}^\perp = n-2$ ; le rang d'une matrice diagonale à coefficients dans  $\{0, 1\}$  est le nombre de ses coefficients égaux à 1. Quitte à réordonner la base orthonormale de vecteurs propres associée (ce qui permute les colonnes de  $R$  et les coefficients de  $D$ ), on peut supposer que les 1 figurent en premier.

La première méthode construit directement  $R$  et  $D$  à partir de la géométrie du problème; la seconde est la démarche générale de réduction d'une matrice symétrique et montre que, quelle que soit la décomposition choisie,  $D$  contient exactement  $n-2$  fois 1 et deux fois 0.

### Conclusion

il existe  $R$  orthogonale et  $D$  diagonale telles que  $G = RD^tR$ ; les éléments diagonaux de  $D$  sont 1 (répété  $n-2$  fois) et 0 (répété deux fois). Dans la suite, on choisit  $D = \text{diag}(1, \dots, 1, 0, 0)$ .

- c) Question adaptée au programme actuel (la loi  $\Gamma(\beta, \nu)$  est définie dans l'énoncé). On applique 8. à  $S = {}^tR$ , puis on détermine la loi d'une somme de carrés de normales indépendantes à l'aide des lois  $\gamma$ .

Indépendance et lois des  $Z_i$ . La matrice  ${}^tR$  est inversible (d'inverse  $R$ ) et orthogonale ( ${}^t({}^tR) {}^tR = R {}^tR = I$ ). D'après 8.,  $Z_1, \dots, Z_n$  sont mutuellement indépendantes et suivent toutes la loi  $\mathcal{N}(0, \sigma^2)$ .

Loi du carré d'une normale centrée réduite. Soit  $V \hookrightarrow \mathcal{N}(0, 1)$ , de densité  $\varphi$ , et  $\Phi$  sa fonction de répartition. Pour  $t < 0$ ,  $P\left(\frac{V^2}{2} \leq t\right) = 0$ , et pour  $t = 0$  cette probabilité vaut  $P(V = 0) = 0$ ; pour  $t > 0$ ,  $P\left(\frac{V^2}{2} \leq t\right) = P(|V| \leq \sqrt{2t}) = 2\Phi(\sqrt{2t}) - 1$ . Cette fonction de répartition est continue sur  $\mathbb{R}$ , de classe  $C^1$  sur  $\mathbb{R}^*$ , de dérivée sur  $]0, +\infty[$  :

$$t \mapsto 2\varphi(\sqrt{2t}) \cdot \frac{1}{\sqrt{2t}} = \frac{2}{\sqrt{2\pi}} \cdot \frac{e^{-t}}{\sqrt{2t}} = \frac{1}{\sqrt{\pi}} t^{-1/2} e^{-t}$$

Ainsi  $\frac{V^2}{2}$  a une densité nulle sur  $] -\infty, 0]$  et proportionnelle à  $t \mapsto t^{-1/2} e^{-t}$  sur  $]0, +\infty[$ , comme la densité de  $\gamma\left(\frac{1}{2}\right)$ . Deux densités proportionnelles sont égales (leurs intégrales valent 1), donc  $\frac{V^2}{2} \hookrightarrow \gamma\left(\frac{1}{2}\right)$  (et, au passage,  $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ ).

Conclusion pour la somme. Pour  $i \leq n-2$ ,  $\frac{Z_i}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$ , donc  $\frac{Z_i^2}{2\sigma^2} \hookrightarrow \gamma\left(\frac{1}{2}\right)$ . Ces  $n-2$  variables sont mutuellement indépendantes (coalitions). Par récurrence et stabilité des lois  $\gamma$ ,

$$\sum_{i=1}^{n-2} \frac{Z_i^2}{2\sigma^2} \hookrightarrow \gamma\left(\frac{n-2}{2}\right)$$

(avec  $n-2 \geq 1$ ).

## Conclusion

$$\sum_{i=1}^{n-2} Z_i^2$$

suit la loi  $\Gamma\left(2\sigma^2, \frac{n-2}{2}\right)$ .

- d) Question adaptée au programme actuel (la loi  $\Gamma(\beta, \nu)$  est définie dans l'énoncé en 9. c)). On exprime

$$\sum \widehat{U}_i^2$$

à l'aide de  $Z$ .

D'après 9. a) et les propriétés  ${}^tG = G, G^2 = G$  (question 3. e)) :

$$\sum_{i=1}^n \widehat{U}_i^2 = {}^t\widehat{U}\widehat{U} = {}^tU {}^tGGU = {}^tUGU = {}^tURD {}^tRU = {}^t({}^tRU) D ({}^tRU) = {}^tZDZ = \sum_{i=1}^{n-2} Z_i^2$$

car  $D = \text{diag}(1, \dots, 1, 0, 0)$ . D'après c) :

### Conclusion

$$\sum_{i=1}^n \widehat{U}_i^2$$

suit la loi  $\Gamma\left(2\sigma^2, \frac{n-2}{2}\right)$ . On retrouve

$$E\left(\sum \widehat{U}_i^2\right) = 2\sigma^2 \cdot \frac{n-2}{2} = (n-2)\sigma^2,$$

conformément à 6. c).

- e) On se ramène à une variable dont la loi ne dépend d'aucun paramètre, puis on applique le théorème de la bijection.

Posons

$$Q = \frac{1}{\sigma^2} \sum_{i=1}^n \widehat{U}_i^2.$$

D'après d),  $\frac{Q}{2} \hookrightarrow \gamma\left(\frac{n-2}{2}\right)$  : la loi de  $Q$  ne dépend que de  $n$ , pas de  $a, b, \sigma^2$ . Sa fonction de répartition  $F_Q$  est continue sur  $\mathbb{R}$  et, sa densité

$$x \mapsto \frac{1}{2} \cdot \frac{(x/2)^{\frac{n-2}{2}-1} e^{-x/2}}{\Gamma\left(\frac{n-2}{2}\right)}$$

étant strictement positive sur  $]0, +\infty[$ ,  $F_Q$  est strictement croissante sur  $]0, +\infty[$ , avec  $F_Q(0) = 0$  et  $\lim_{+\infty} F_Q = 1$ . La fonction  $c \mapsto P([Q \geq c]) = 1 - F_Q(c)$  (car  $P(Q = c) = 0$ ) réalise donc une bijection continue strictement décroissante de  $]0, +\infty[$  sur  $]0, 1[$ . Comme  $p \in ]0, 1[$ , il existe un unique  $c_n > 0$  tel que  $P([Q \geq c_n]) = p$ , c'est-à-dire

$$P\left(\left[\sum \widehat{U}_i^2 \geq c_n \sigma^2\right]\right) = p.$$

### Conclusion

un tel réel  $c_n$  existe ; il ne dépend que de  $n$  et de  $p$ .

10. a) On minimise  $g(r) = \|r\|^2$  sur l'ensemble  $\mathcal{C}$  des  $r$  vérifiant les deux contraintes linéaires  $\langle r, \mathbf{1} \rangle = 1$  et  $\langle r, \mathbf{x} \rangle = x_{n+1}$ .

Vérification préalable.  $r^* \in \mathcal{C}$  : d'après 1. c),

$$\sum_i r_i^* = 1 + (x_{n+1} - \bar{x}) \sum_i \alpha_i = 1$$

et

$$\sum_i x_i r_i^* = \bar{x} + (x_{n+1} - \bar{x}) \sum_i \alpha_i x_i = x_{n+1}.$$

De plus  $r^* = \frac{1}{n} \mathbf{1} + \frac{x_{n+1} - \bar{x}}{n s_x^2} (\mathbf{x} - \bar{x} \mathbf{1})$  appartient à  $\mathcal{F} = \text{Vect}(\mathbf{x}, \mathbf{1})$ .

**Méthode 1 — Théorème de Pythagore.** Soit  $r \in \mathcal{C}$  et  $h = r - r^*$ . Alors  $\langle h, \mathbf{1} \rangle = 1 - 1 = 0$  et  $\langle h, \mathbf{x} \rangle = x_{n+1} - x_{n+1} = 0$ , donc  $h \in \mathcal{F}^\perp$ , et en particulier  $h \perp r^*$ . D'après le théorème de Pythagore :

$$g(r) = \|r^* + h\|^2 = \|r^*\|^2 + \|h\|^2 \geq g(r^*)$$

avec égalité si et seulement si  $h = 0$ , c'est-à-dire  $r = r^*$ .

**Méthode 2 — Condition nécessaire sous contraintes linéaires.** La fonction  $g$  est de classe  $C^1$  sur  $\mathbb{R}^n$ , de gradient  $\nabla g(r) = 2r$ ; les gradients des contraintes sont  $\mathbf{1}$  et  $\mathbf{x}$ .

Existence d'un minimum.  $\mathcal{C}$  est fermé (défini par des égalités de fonctions continues) et contient  $r^*$ . L'ensemble  $\mathcal{C}' = \{r \in \mathcal{C} : g(r) \leq g(r^*)\}$  est fermé, borné (contenu dans la boule fermée de centre 0 et de rayon  $\|r^*\|$ ) et non vide; la fonction continue  $g$  y atteint un minimum, qui est aussi son minimum sur  $\mathcal{C}$  (hors de  $\mathcal{C}'$ ,  $g > g(r^*)$ ).

Identification. En un point  $r_0$  où ce minimum est atteint, la condition nécessaire d'extremum sous contraintes d'égalités linéaires impose  $\nabla g(r_0) \in \text{Vect}(\mathbf{1}, \mathbf{x})$  : il existe  $\lambda', \mu'$  réels tels que  $r_0 = \lambda' \mathbf{1} + \mu' \mathbf{x}$ . Les contraintes donnent  $n\lambda' + n\bar{x}\mu' = 1$  et

$$n\bar{x}\lambda' + \mu' \sum x_i^2 = x_{n+1}.$$

La première donne  $\lambda' = \frac{1}{n} - \bar{x}\mu'$ ; en reportant,

$$\mu' \left( \sum x_i^2 - n\bar{x}^2 \right) = x_{n+1} - \bar{x},$$

soit  $\mu' = \frac{x_{n+1} - \bar{x}}{ns_x^2}$ . Donc  $r_0$  a pour coordonnées  $\lambda' + \mu' x_i = \frac{1}{n} + \mu'(x_i - \bar{x}) = r_i^*$  :

$r_0 = r^*$ , ce qui prouve à la fois que le minimum est atteint en  $r^*$  et qu'il ne l'est qu'en ce point.

La méthode géométrique est la plus rapide une fois le candidat  $r^*$  connu, et donne l'unicité directement ; la méthode analytique permet de trouver le candidat sans le connaître à l'avance.

### Conclusion

$g$  admet sous les deux contraintes un minimum absolu, atteint en l'unique point  $r^*$  de coordonnées  $r_i^* = \frac{1}{n} + (x_{n+1} - \bar{x})\alpha_i$ .

**b)** On traduit la condition d'espérance en les deux contraintes de a), et la variance en  $\sigma^2 g(r)$ .

Condition d'espérance.

$$E(\widehat{Y}_{n+1}^{(r)}) = \sum_i r_i(ax_i + b) = a \sum_i r_i x_i + b \sum_i r_i$$

et  $E(Y_{n+1}) = ax_{n+1} + b$ . L'égalité pour tout  $(a, b)$  équivaut, en prenant  $(a, b) = (0, 1)$  puis  $(1, 0)$  pour le sens direct, à

$$\sum_i r_i = 1$$

et

$$\sum_i x_i r_i = x_{n+1},$$

c'est-à-dire  $r \in \mathcal{C}$ .

Variance. Les  $Y_i$  étant indépendantes de variance  $\sigma^2$ ,

$$V(\widehat{Y}_{n+1}^{(r)}) = \sigma^2 \sum_i r_i^2 = \sigma^2 g(r).$$

Comme  $\sigma^2 > 0$ , minimiser la variance sur  $\mathcal{C}$  revient à minimiser  $g$  sur  $\mathcal{C}$  : d'après a), le minimum est atteint pour  $r = r^*$  et seulement pour lui.

Expression.

$$\begin{aligned}\widehat{Y}_{n+1}^{(r^*)} &= \frac{1}{n} \sum_i Y_i + (x_{n+1} - \bar{x}) \sum_i \alpha_i Y_i = \bar{Y}_n + (x_{n+1} - \bar{x}) A_n = (B_n + A_n \bar{x}) \\ &+ (x_{n+1} - \bar{x}) A_n.\end{aligned}$$

### Conclusion

parmi les prédicteurs linéaires sans biais (pour tout  $(a, b)$ ),  $\widehat{Y}_{n+1}^{(r^*)}$  est l'unique prédicteur de variance minimale, et  $\widehat{Y}_{n+1}^{(r^*)} = A_n x_{n+1} + B_n$ .

11. a) On écrit la variable comme combinaison linéaire non triviale des  $U_1, \dots, U_{n+1}$  et on applique le lemme L.

D'après 10. b),

$$A_n x_{n+1} + B_n = \sum_i r_i^* Y_i = \sum_i r_i^* (a x_i + b) + \sum_i r_i^* U_i = a x_{n+1} + b + \sum_i r_i^* U_i,$$

car  $r^* \in \mathcal{C}$ . Donc :

$$Y_{n+1} - (A_n x_{n+1} + B_n) = U_{n+1} - \sum_{i=1}^n r_i^* U_i$$

combinaison linéaire des variables indépendantes  $U_1, \dots, U_{n+1}$ , de loi  $\mathcal{N}(0, \sigma^2)$ , de coefficients non tous nuls (celui de  $U_{n+1}$  vaut 1). D'après le lemme L, elle suit la loi  $\mathcal{N}(0, \sigma^2 (1 + g(r^*)))$ . Enfin, avec 1. c) :

$$g(r^*) = \sum_{i=1}^n \left( \frac{1}{n^2} + \frac{2(x_{n+1} - \bar{x})}{n} \alpha_i + (x_{n+1} - \bar{x})^2 \alpha_i^2 \right) = \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{n s_x^2}$$

## Conclusion

$$Y_{n+1} - (A_n x_{n+1} + B_n) \hookrightarrow \mathcal{N}(0, \sigma^2 \tau_n^2) \text{ avec } \tau_n^2 = 1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{n s_x^2}.$$

b) On centre et on réduit, puis on utilise la bijectivité de  $\Phi$ .

Notons  $\Delta = Y_{n+1} - (A_n x_{n+1} + B_n)$  et  $\tau_n > 0$ . D'après a),  $N = \frac{\Delta}{\sigma \tau_n} \hookrightarrow \mathcal{N}(0, 1)$ . Pour  $d > 0$ , par symétrie de la loi normale centrée réduite :

$$P(|\Delta| \leq d\sigma) = P\left(|N| \leq \frac{d}{\tau_n}\right) = 2\Phi\left(\frac{d}{\tau_n}\right) - 1$$

Cette quantité vaut  $p$  si et seulement si  $\Phi\left(\frac{d}{\tau_n}\right) = \frac{1+p}{2}$ . La fonction  $\Phi$  est continue et strictement croissante sur  $\mathbb{R}$ , de limites 0 et 1 : c'est une bijection de  $\mathbb{R}$  sur  $]0, 1[$ . Comme  $\frac{1+p}{2} \in \left]\frac{3}{4}, 1\right[$ ,  $\Phi^{-1}\left(\frac{1+p}{2}\right)$  existe et est strictement positif (car  $\Phi(0) = \frac{1}{2}$ ).

## Conclusion

$$d_n = \tau_n \Phi^{-1}\left(\frac{1+p}{2}\right) = \sqrt{1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{n s_x^2}} \Phi^{-1}\left(\frac{1+p}{2}\right)$$

convient ; il ne dépend que de  $p$ , de  $n$  et des  $x_i$ , et non de  $a, b, \sigma^2$ .

c) On remplace  $\sigma$  par la majoration fournie par 9. e) et on minore la probabilité d'une intersection par la formule du crible.

Posons  $\mathcal{E}_1 = [|\Delta| \leq d_n \sigma]$  et

$$\mathcal{E}_2 = \left[\sum \widehat{U}_i^2 \geq c_n \sigma^2\right],$$

où  $c_n$  est le réel de 9. e) associé au même  $p$ . On a  $P(\mathcal{E}_1) = P(\mathcal{E}_2) = p$ . Comme  $c_n > 0$  et

$\sigma > 0$ ,

$$\mathcal{E}_2 = \left[ \sigma \leq \sqrt{\frac{1}{c_n} \sum \widehat{U}_i^2} \right].$$

Sur  $\mathcal{E}_1 \cap \mathcal{E}_2$  :

$$|\Delta| \leq d_n \sigma \leq d_n \sqrt{\frac{1}{c_n} \sum_{i=1}^n \widehat{U}_i^2}$$

Par la formule du crible,  $P(\mathcal{E}_1 \cap \mathcal{E}_2) = P(\mathcal{E}_1) + P(\mathcal{E}_2) - P(\mathcal{E}_1 \cup \mathcal{E}_2) \geq 2p - 1$ , qui est strictement positif car  $p > \frac{1}{2}$ . Posons

$$L_n = d_n \sqrt{\frac{1}{c_n} \sum_{i=1}^n \widehat{U}_i^2};$$

les  $\widehat{U}_i = Y_i - A_n x_i - B_n$  ne dépendent que des  $Y_i$  ( $i \leq n$ ) et des  $x_i$ .

## Conclusion

l'intervalle  $[A_n x_{n+1} + B_n - L_n, A_n x_{n+1} + B_n + L_n]$ , dont les bornes ne dépendent que des  $Y_i$  ( $1 \leq i \leq n$ ), des  $x_i$  ( $1 \leq i \leq n+1$ ), de  $c_n$  et de  $d_n$ , contient  $Y_{n+1}$  avec une probabilité supérieure ou égale à  $2p - 1$ . Ce n'est pas un intervalle de confiance au sens usuel : un intervalle de confiance encadre un paramètre inconnu mais déterministe, alors qu'ici on encadre la variable aléatoire  $Y_{n+1}$  (il s'agit d'un intervalle de prévision) ; de plus, seule une minoration de la probabilité est obtenue.

## Relevé des adaptations

9. c) et 9. d) L'original utilise la notation  $\Gamma\left(2\sigma^2, \frac{n-2}{2}\right)$  sans la définir ; le référentiel ne contient que la loi gamma de taux 1. L'énoncé adapté définit, au début de 9. c), la loi  $\Gamma(\beta, \nu)$  comme la loi d'une variable  $X$  telle que  $\frac{X}{\beta}$  suive la loi  $\gamma(\nu)$ . Justification :

rendre la question conforme au référentiel sans changer son contenu (c'est la convention de l'original : densité  $x \mapsto \frac{x^{\nu-1}e^{-x/\beta}}{\beta^\nu \Gamma(\nu)}$  sur  $]0, +\infty[$ ). Les autres questions sont conformes ; les méthodes ont seulement été choisies dans le référentiel (invertibilité de  ${}^tMM$  sans déterminant général, étude de la hessienne par ses valeurs propres en 3. a), extremum sous contraintes linéaires en 10. a)).