Université Paris 5 - René Descartes
UFR de Mathématiques et Informatique
45, rue des Saints-Pères 75270 Paris cedex 06

Maîtrises Math-Mass 2001-2002

Corrigé du partiel du 13 novembre 2001

1)a) Le n-échantillon $(X_1, \dots, X_n)$ est à valeurs dans $\mathbb{R}^n$; on choisit donc $\Omega^{(n)} = \mathbb{R}^n$ comme espace des observations, muni de la tribu des boréliens $\mathcal{A}^{(n)} = \mathcal{B}(\mathbb{R}^n)$. La loi $\mathbb{P}_\theta^{(n)}$ est celle de n gaussiennes indépendantes, et sa densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$ vaut

$$\forall (x_1, \dots, x_n) \in \Omega^{(n)} \quad \forall \theta \in \Theta = \mathbb{R}$$

$$\begin{aligned}
\frac{d\mathbb{P}_\theta^{(n)}}{d\lambda_{\mathbb{R}^n}}(x_1, \dots, x_n) &= f_\theta(x_1, \dots, x_n) \\
&= \prod_{i=1}^n \frac{\exp\left(-\frac{(x_i-\theta)^2}{2}\right)}{\sqrt{2\pi}} = \frac{\exp\left(-\frac{\sum_{i=1}^n (x_i-\theta)^2}{2}\right)}{(2\pi)^{\frac{n}{2}}}.
\end{aligned}$$

b) Le modèle est exponentiel : en effet, la densité de la loi $\mathbb{P}_\theta^{(n)}$ peut se mettre sous la forme

$$\forall (x_1, \dots, x_n) \in \Omega^{(n)}$$

$$\begin{aligned}
f_\theta(x_1, \dots, x_n) &= \exp\left(n\theta \bar{X}(n) - n\frac{\theta^2}{2}\right) \frac{\exp\left(-\frac{\sum_{i=1}^n x_i^2}{2}\right)}{(2\pi)^{\frac{n}{2}}} \\
&= \exp\big(Q(\theta)\bar{X}(n) - \phi(\theta)\big)\, h(x_1, \dots, x_n)
\end{aligned}$$

avec

$$\begin{aligned}
Q(\theta) &= n\theta \\
\phi(\theta) &= n\frac{\theta^2}{2} \\
h(x_1, \dots, x_n) &= \frac{\exp\left(-\frac{\sum_{i=1}^n x_i^2}{2}\right)}{(2\pi)^{\frac{n}{2}}}.
\end{aligned}$$

L'ensemble image $Q(\Theta) = \mathbb{R}$ est d'intérieur non vide ; il n'y a pas de relation linéaire non triviale entre les différentes statistiques présentes sous l'exponentielle car il n'y en a qu'une seule ; on en déduit que $\bar{X}(n)$ est une statistique exhaustive et complète, donc minimale.

c) Le biais de $\bar{X}(n)$ vaut

$$\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[\bar{X}(n)] - \theta = \frac{1}{n}\sum_{i=1}^n \mathbb{E}_{\mathbb{P}_\theta^{(n)}}[X_i] - \theta = \frac{1}{n} n\theta - \theta = 0.$$

L'estimateur $\bar{X}(n)$ est donc sans biais. Son risque est alors égal à sa variance et vaut

$$R(\theta, \bar{X}(n)) = \operatorname{var}_{\mathbb{P}_\theta^{(n)}}\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^n \operatorname{var}_{\mathbb{P}_\theta^{(n)}}(X_i) = \frac{1}{n^2} n = \frac{1}{n}$$

où l'on s'est servi de ce que la variance d'une somme de v.a. indépendantes est égale à la somme des variances respectives. L'estimateur $\bar{X}(n)$ est donc de risque constant, car son risque ne dépend pas de θ.

d) Le modèle statistique est exponentiel. L'ensemble $\Theta = \mathbb{R}$ des paramètres est un ouvert de $\mathbb{R}$. La fonction Q est évidemment de dérivée continue et ne s'annulant pas. On en déduit que le modèle est régulier, et que $\bar{X}(n)$ est un estimateur UVMB, régulier et efficace de son espérance qui est $\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[\bar{X}(n)] = \theta$.
Il s'ensuit que la borne de Fréchet-Cramer-Rao est atteinte par cet estimateur :

$$\operatorname{var}_{\mathbb{P}_\theta^{(n)}}\left(\bar{X}(n)\right) = \left(\frac{d\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[\bar{X}(n)]}{d\theta}\right)^2 \frac{1}{I_n(\theta)} = \frac{1}{I_n(\theta)}$$

Il en résulte que $I_n(\theta) = \frac{1}{\operatorname{var}_{\mathbb{P}_\theta^{(n)}}(\bar{X}(n))} = n$.

2)a) Soient S et T deux estimateurs de θ, ayant $R(\theta, S)$ et $R(\theta, T)$ comme fonctions de risque respectives. L'estimateur S est dit meilleur que T si et seulement si

$$\forall \theta \in \Theta \quad R(\theta, S) \leq R(\theta, T)$$
$$\exists \theta \in \Theta \quad R(\theta, S) < R(\theta, T)$$

Un estimateur est dit admissible s'il n'existe pas d'estimateur meilleur que lui.
Soit T un estimateur de θ. Supposons le meilleur que $\bar{X}(n)$. Il existe donc θ_0 tel que

$$R(\theta_0, T) < R(\theta_0, \bar{X}(n)).$$

S'il est sans biais comme $\bar{X}(n)$, alors le risque est égal à la variance, et cette inégalité se récrit comme

$$\operatorname{var}_{\mathbb{P}_{\theta_0}^{(n)}}(T) < \operatorname{var}_{\mathbb{P}_{\theta_0}^{(n)}}\left(\bar{X}(n)\right);$$

mais comme $\bar{X}(n)$ est de variance minimale parmi les estimateurs sans biais, on a aussi que, pour tout $\theta \in \Theta$,

$$\operatorname{var}_{\mathbb{P}_\theta^{(n)}}(T) \geq \operatorname{var}_{\mathbb{P}_\theta^{(n)}}\left(\bar{X}(n)\right)$$

Ces deux dernières inégalités se contredisent. L'estimateur T est nécessairement biaisé.
b) Soit $T = T(X_1, \ldots, X_n)$ un estimateur de θ, tel que $\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[T^2] < +\infty$. Alors

$$\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[T] = \int_{\mathbb{R}^n} g(\theta, x_1, \ldots, x_n) dx_1 \ldots dx_n,$$

avec $g(\theta, x_1, \ldots, x_n) = T(x_1, \ldots, x_n) f_\theta(x_1, \ldots, x_n)$. L'intégrale

$$\int_{\mathbb{R}^n} |g(\theta, x_1, \ldots, x_n)| dx_1 \ldots dx_n = \mathbb{E}_{\mathbb{P}_\theta^{(n)}}[|T|] \leq \sqrt{\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[T^2]}$$

est convergente pour toute valeur de θ car T est de carré intégrable. Soit $A > 0$ quelconque. Montrons que $\left|\frac{\partial g}{\partial \theta}\right|$ est dominée par une fonction intégrable sur $\mathbb{R}^n$, uniformément sur $\theta \in]-A, A[$.

En effet, si $\theta \in]-A, A[$, alors

$$\begin{aligned}
\left|\frac{\partial g}{\partial \theta}\right| &= \frac{|T|}{(2\pi)^{n/2}} n|\bar{x}-\theta| \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2\right) \text{ avec } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \\
&= \frac{|T|}{(2\pi)^{n/2}} n(|\bar{x}|+A) \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + nA|\bar{x}|\right) \\
&= \frac{|T|}{(2\pi)^{n/2}} n(|\bar{x}|+A) \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \sqrt{n}A\sqrt{\sum_{i=1}^{n} x_i^2}\right) \text{ car } |\bar{x}| \leq \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \\
&= \frac{|T|}{(2\pi)^{n/2}} n(|\bar{x}|+A) \exp\left(-\frac{1}{3}\sum_{i=1}^{n} x_i^2\right) C
\end{aligned}$$

avec

$$C = \sup_{(x_1,\dots x_n)\in\mathbb{R}^n} \exp\left(-\frac{\sum_{i=1}^{n} x_i^2}{6} + \sqrt{n}A\sqrt{\sum_{i=1}^{n} x_i^2}\right) = \sup_{u\geq 0} \exp\left(-\frac{u}{6} + \sqrt{n}A\sqrt{u}\right)$$

La constante C est finie, car la fonction $u \mapsto \exp\left(-\frac{u}{6} + \sqrt{n}A\sqrt{u}\right)$, qui est continue sur $[0, +\infty[$ et qui tend vers 0 en $+\infty$, est nécessairement bornée. On a ainsi majoré $\left|\frac{\partial g}{\partial \theta}(\theta, x_1, \dots, x_n)\right|$ par une fonction qui ne dépend pas de θ, et dont on vérifie facilement qu'elle est intégrable grâce à l'inégalité de Cauchy-Schwarz :

$$\begin{aligned}
&\int_{\mathbb{R}^n} \frac{|T|}{(2\pi)^{n/2}} n(|\bar{x}|+A) \exp\left(-\frac{1}{3}\sum_{i=1}^{n} x_i^2\right) C dx_1 \dots dx_n \\
&\leq C\sqrt{\int_{\mathbb{R}^n} \frac{T^2}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) dx_1 \dots dx_n} \\
&\times\sqrt{\int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2}} n^2(|\bar{x}|+A)^2 \exp\left(-\frac{1}{6}\sum_{i=1}^{n} x_i^2\right) dx_1 \dots dx_n}
\end{aligned}$$

Or la première intégrale du membre de droite de l'inégalité est égale à $\mathbb{E}_{\mathbb{P}_0^{(n)}}[T^2] < +\infty$ et n'apparaissent dans la deuxième intégrale que les deux premiers moments - finis - d'une gaussienne centrée. L'intégrale située à gauche de l'inégalité est donc convergente.
On déduit de ce résultat de domination et de la convergence de $\int_{\mathbb{R}^n} |g(\theta, x_1, \dots, x_n)| dx_1 \dots dx_n$ que l'intégrale $\int_{\mathbb{R}^n} g(\theta, x_1, \dots, x_n) dx_1 \dots dx_n$ est dérivable par rapport à θ, et que sa dérivée vaut $\int_{\mathbb{R}^n} \frac{\partial f}{\partial \theta}(\theta, x_1, \dots, x_n) dx_1 \dots dx_n$, sur l'intervalle $]-A, A[$ quel que soit $A > 0$, donc sur tout $\mathbb{R}$. En d'autres termes, ceci signifie que, pour tout $\theta \in \mathbb{R}$,

$$\begin{aligned}
\frac{d}{d\theta}\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[T] &= = \frac{d}{d\theta}\int_{\mathbb{R}^n} g(\theta, x_1, \dots, x_n) dx_1 \dots dx_n = \int_{\mathbb{R}^n} \frac{\partial g}{\partial \theta}(\theta, x_1, \dots, x_n) dx_1 \dots dx_n \\
&= \int_{\mathbb{R}^n} T(x_1, \dots, x_n) \frac{\partial f_\theta}{\partial \theta}(x_1, \dots, x_n) dx_1 \dots dx_n \\
&= \int_{\mathbb{R}^n} T(x_1, \dots, x_n) \frac{\partial \ln f_\theta}{\partial \theta}(x_1, \dots, x_n) f_\theta(x_1, \dots, x_n) dx_1 \dots dx_n \\
&= \mathbb{E}_{\mathbb{P}_\theta^{(n)}}\left[T\frac{\partial \ln f_\theta(x_1, \dots, x_n))}{\partial \theta}\right]
\end{aligned}$$

c'est-à-dire que T est un estimateur régulier. On en déduit aussi que le biais $\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[T] - \theta$ est bien dérivable.
Comme T est régulier, il résulte de l'inégalité de Fréchet-Cramer-Rao que

$$\operatorname{var}_{\mathbb{P}_\theta^{(n)}}(T) \geq \left(\frac{d\mathbb{E}_{\mathbb{P}_\theta^{(n)}}[T]}{d\theta}\right)^2 \frac{1}{I_n(\theta)} = \left(\frac{d(\theta + b(\theta)}{d\theta}\right)^2 \frac{1}{n} = \frac{(1 + b'(\theta))^2}{n}$$

D'autre part, on sait que le risque (quadratique) est la somme de la variance et du biais au carré ; d'où

$$R(\theta, T) = b^2(\theta) + \operatorname{var}_{\mathbb{P}_\theta^{(n)}}(T) \geq b^2(\theta) + \frac{(1 + b'(\theta))^2}{n}$$

c) Si T est un estimateur meilleur que $\bar{X}$, son risque est plus faible que celui de $\bar{X}$. Or $R(\theta, \bar{X}) = 1/n$ quel que soit θ. Il en ressort que

$$\frac{1}{n} \geq R(\theta, T) \geq b^2(\theta) + \frac{(1 + b'(\theta))^2}{n},$$

ce qui est l'inégalité demandée. On en déduit quel que soit θ

$$\begin{aligned} b^2(\theta) &\leq \frac{1}{n} \\ (1 + b'(\theta)^2 &\leq 1 \end{aligned}$$

ce qui implique que b soit une fonction bornée et que $1 + b'(\theta)$ soit compris entre -1 et 1, autrement dit que $b'(\theta) \in [-2, 0]$ quel que soit θ. Puisque b' est une fonction négative, il s'ensuit que b est décroissante. On en déduit que b, fonction décroissante et minorée, admet une limite en $+\infty$, que l'on notera l_+, et qui vaut $\inf_{\theta \in \mathbb{R}} b(\theta)$; on en déduit de même que b, fonction décroissante et majorée, admet une limite en $-\infty$, que l'on notera l_-, et qui vaut $\sup_{\theta \in \mathbb{R}} b(\theta)$. Supposons $l_+ \neq 0$. Il existe alors $A \in \mathbb{R}$ tel que

$$\forall \theta > A \quad |b(\theta)| \geq \frac{1}{2}|l_+|$$

(b ne tendant pas vers 0, il existe un voisinage de l'infini $]A, +\infty[$ où b est hors du voisinage $]-\frac{1}{2}|l_+|, \frac{1}{2}|l_+|[$ de 0). Cela implique que pour tout $\theta > A$, on ait

$$\frac{(1 + b'(\theta))^2}{n} \leq \frac{1}{n} - b^2(\theta) \leq \frac{1}{n} - \frac{1}{4}l_+^2$$

D'où l'on déduit

$$\forall \theta > A \quad b'(\theta) \leq \sqrt{1 - \frac{n}{4}l_+^2} - 1 < 0,$$

puis

$$\forall \theta > A \quad b(\theta) = b(A) + \int_A^\theta b'(u)du \leq b(A) + (\theta - u)\left(\sqrt{1 - \frac{n}{4}l_+^2} - 1\right)$$

Cette inégalité a pour conséquence $\lim_{\theta \to +\infty} b(\theta) = -\infty$. On montre exactement de même que si on a $l_- \neq 0$, alors $\lim_{\theta \to -\infty} b(\theta) = +\infty$. Donc, si les limites l_+ et l_- ne sont pas toutes deux nulles, alors b n'est pas bornée.

d) Si T est un estimateur de θ meilleur que $\bar{X}$, alors, d'après la question précédente, son biais b est borné, décroissant, et ses limites en l'infini sont alors nécessairement nulles. Il s'ensuit que b elle-même est la fonction nulle. Donc T est sans biais. Or, d'après la question a), un estimateur meilleur que $\bar{X}$ est forcément biaisé. Il y a là une contradiction. Il n'existe donc pas d'estimateur de θ meilleur que $\bar{X}$, qui est par conséquent un estimateur admissible.

3) On sait qu'une règle de décision de risque constant et admissible est minimax. D'après les questions 1c) et 2d), $\bar{X}$ est un estimateur de risque constant et admissible. Il est donc minimax.

4) a) Soit $x \in \mathbb{R}$ quelconque. Sous $\mathbb{P}_\theta^{(\delta_x,t)}$, $(X_i, i \geq 0)$ est une suite de v.a. indépendantes de même loi, telles que

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}_\theta^{(\delta_x,t)}}[|X_1|] &= (1-t)\int_{-\infty}^{+\infty} |u| \frac{e^{-\frac{(u-\theta)^2}{2}}}{\sqrt{2\pi}} du + t|x| \\
&= (1-t)\mathbb{E}_{\mathbb{P}_\theta}[|X_i|] + t|x| < +\infty
\end{aligned}$$

On déduit alors de la loi forte des grands nombres que la moyenne empirique $\bar{X}(n)$ converge $\mathbb{P}_\theta^{(\delta_x,t)}$-p.s. vers $\mathbb{E}_{\mathbb{P}_\theta^{(\delta_x,t)}}[X_1]$. Autrement dit, $\bar{X}_\infty(\theta, \delta_x, t)$ existe et vaut

$$\bar{X}_\infty(\theta, \delta_x, t) = \mathbb{E}_{\mathbb{P}_\theta^{(\delta_x,t)}}[X_1] = (1-t)\mathbb{E}_{\mathbb{P}_\theta}[X_1] + tx = (1-t)\theta + tx.$$

Il s'ensuit que $\frac{\partial T_\infty}{\partial t}(\theta, \delta_x, t) = x - \theta$, et donc que

$$\sup_{x \in \mathbb{R}} \left| \frac{\partial T_\infty}{\partial t}(\theta, \delta_x, 0) \right| = +\infty.$$

La famille d'estimateurs $(\bar{X}(n))_{n \geq 1}$ n'est donc pas robuste.
b) Soit ψ telle que $\psi(X_1) = \mathbb{E}_{\mathbb{P}_\theta}[1_{Y \leq X_1} | X_1]$ $\mathbb{P}_\theta$-p.s. Alors

$$\begin{aligned}
\psi(x) &= \mathbb{E}_{\mathbb{P}_\theta}[1_{Y \leq X_1} | X_1 = x] \\
&= \mathbb{E}_{\mathbb{P}_\theta}[1_{Y \leq x} | X_1 = x] \\
&= \mathbb{E}_{\mathbb{P}_\theta}[1_{Y \leq x}] \text{ car } Y \text{ est indépendante de } X_1 \\
&= \mathbb{P}_\theta(Y \leq x) = \phi(x)
\end{aligned}$$

D'où il ressort qu'effectivement $\phi(X_1) = \psi(X_1) = \mathbb{E}_{\mathbb{P}_\theta}[1_{Y \leq X_1} | X_1]$ $\mathbb{P}_\theta$-p.s. On en déduit

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}\theta}[\phi(X_1)] &= \mathbb{E}_{\mathbb{P}_\theta}\left[\mathbb{E}_{\mathbb{P}_\theta}[1_{Y \leq X_1} | X_1]\right] = \mathbb{E}_{\mathbb{P}_\theta}\left[1_{Y \leq X_1}\right] \\
&= \mathbb{P}_\theta(Y \leq X_1) = \mathbb{P}_\theta\left(\frac{Y - X_1 + \theta}{\sqrt{2}} \leq \frac{\theta}{\sqrt{2}}\right) = \phi\left(\frac{\theta}{\sqrt{2}}\right)
\end{aligned}$$

car, sous $\mathbb{P}_\theta$, $\frac{Y - X_1 + \theta}{\sqrt{2}}$ est une gaussienne centrée réduite. Comme ϕ est bijective, on en déduit la formule demandée, à savoir $\theta = \sqrt{2}\phi^{-1}\left(\mathbb{E}_{\mathbb{P}_\theta}\left[\phi(X_1)\right]\right)$.
c) La méthode de substitution consiste à remplacer l'espérance $\mathbb{E}_{\mathbb{P}_\theta}\left[\phi(X_1)\right]$ par l'espérance empirique $\frac{1}{n}\sum_{i=1}^n \phi(X_i)$. On en déduit la famille d'estimateurs

$$\left(U_n = \sqrt{2}\phi^{-1}\left(\frac{1}{n}\sum_{i=1}^n \phi(X_i)\right), n \geq 1 \right).$$

Sous $\mathbb{P}_\theta (= \mathbb{P}_\theta^{(\nu,t=0)})$ comme sous $\mathbb{P}_\theta^{(\nu,t)}$, la famille $(\phi(X_n)n \geq 1)$ est une suite de v.a. indépendantes, de même loi et bornées donc intégrables. Il résulte alors de la loi forte des grands nombres que

$$\lim_{n \to +\infty} \frac{1}{n}\sum_{i=1}^n \phi(X_i) = \mathbb{E}_{\mathbb{P}_\theta^{(\nu,t)}}[\phi(X_1)] \quad \mathbb{P}_\theta^{(\nu,t)}\text{-p.s.}$$

Comme ϕ^{-1} est continue (c'est la réciproque d'une fonction strictement monotone et continue), on en déduit que $\mathbb{P}_\theta^{(\nu,t)}$-p.s. U_n converge vers $U_\infty(\theta,\nu,t) = \sqrt{2}\phi^{-1}(\mathbb{E}_{\mathbb{P}_\theta^{(\nu,t)}}[\phi(X_1)])$.

d) Soit $x \in \mathbb{R}$ quelconque. Comme $\mathbb{E}_{\mathbb{P}_\theta^{(\delta_x,t)}}[\phi(X_1)] = (1-t)\mathbb{E}_{\mathbb{P}_\theta}[\phi(X_1)] + t\phi(x) = (1-t)\phi(\theta/\sqrt{2}) + t\phi(x)$, il s'ensuit le calcul :

$$\begin{aligned} \frac{\partial U_\infty}{\partial t}(\theta,\nu,t) &= \sqrt{2}(\phi(x) - \phi(\theta/\sqrt{2}))\frac{1}{\phi' \circ \phi^{-1}((1-t)\phi(\theta/\sqrt{2}) + t\phi(x))} \\ &= 2\sqrt{\pi}(\phi(x) - \phi(\theta/\sqrt{2}))\exp\left(\frac{1}{2}\phi^{-1}\left((1-t)\phi(\theta/\sqrt{2}) + t\phi(x)\right)^2\right) \end{aligned}$$

Quand t tend vers 0, $\frac{\partial U_\infty}{\partial t}(\theta,\nu,t)$ tend vers $2\sqrt{\pi}(\phi(x) - \phi(\theta/\sqrt{2}))\exp\left(\frac{1}{2}\phi^{-1}(\phi(\theta/\sqrt{2}))^2\right)$, qui est borné uniformément en x car $\phi(x)$ appartient à $]0,1[$. On en conclut que $(U_n, n \geq 1)$ est une suite robuste d'estimateurs de θ.