TEXTE DE L'EPREUVE
DE PROBABILITES ET STATISTIQUES

Conventions et notations.

a. Dans ce problème, on aura lieu de considérer, sur l'ensemble $\mathbf{R}^n$,

— le *produit scalaire* usuel (noté .) et la *norme* qui lui est associée (norme euclidienne, notée $|\ \ |$);

— la *topologie usuelle,* définie par la norme $|\ \ |$ (pour toute partie A de $\mathbf{R}^n$, on note $\mathring{A}$ son intérieur et $\bar{A}$ sa fermeture) et la *tribu borélienne* (notée $\mathcal{B}^n$) qui lui est associée;

— l'ensemble des parties *convexes* : étant donné une partie A de $\mathbf{R}^n$, on appelle enveloppe convexe (resp enveloppe convexe fermée) de A la plus petite (au sens de l'inclusion) partie convexe (resp convexe fermée) contenant A.

b. Tous les espaces mesurables envisagés dans ce problème sont *euclidiens :* $(\Omega, \mathcal{A})$ est dit euclidien si, et seulement si, il existe n, entier strictement positif, tel que $\Omega \in \mathcal{B}^n$, et que $\mathcal{A}$ soit la tribu induite par $\mathcal{B}^n$ sur Ω (c'est-à-dire $\mathcal{A} = \{ A \in \mathcal{B}^n \ ; \ A \subset \Omega \}$).

c. Pour toute probabilité P sur $(\mathbf{R}^n, \mathcal{B}^n)$, on appelle *support* de P, et on note Supp (P) la plus petite (au sens de l'inclusion) partie fermée F de $\mathbf{R}^n$ vérifiant $P(F) = 1$ (on admettra son existence).

Définitions.

Soient deux entiers strictement positifs, n et m; soit $\Omega \in \mathcal{B}^n$ et soit $\Theta \subset \mathbf{R}^m$; soit (P_θ) une famille de probabilités sur l'espace mesurable euclidien $(\Omega, \mathcal{A})$, indicée par Θ; Ω est appelé l'ensemble des *résultats,* et Θ l'ensemble des *paramètres de* la famille (P_θ).

On dit que la famille (P_θ) est *exponentielle* si et seulement si il existe :

une mesure σ-finie μ sur $(\Omega, \mathcal{A})$,
un entier strictement positif k,
une application mesurable T de $(\Omega, \mathcal{A})$ dans $(\mathbf{R}^k, \mathcal{B}^k)$,
une application U de Θ dans $\mathbf{R}^k$,
une application mesurable a de $(\Omega, \mathcal{A})$ dans $(\mathbf{R}, \mathcal{B})$,
une application b de Θ dans $\mathbf{R}$,

tels que, pour tout θ, P_θ soit absolument continue par rapport à μ et admette, pour densité par rapport à μ, la fonction

$$\omega \longmapsto \exp [T(\omega) \, . \, U(\theta) - a(\omega) - b(\theta)] \ .$$

Le quadruplet (T, U, a, b) est appelé une *représentation* de *référence* μ (ou μ-représentation), de *dimension* k, de la famille (P_θ); cette représentation est dite :

résultats-identique si $n = k$ et, pour tout $\omega \in \Omega$, $T(\omega) = \omega$,
paramètres-identique si $m = k$ et, pour tout $\theta \in \Theta$, $U(\theta) = \theta$,
de *type nul* si, pour tout $\omega \in \Omega$, $a(\omega) = 0$;

la fonction b est appelée *fonction de cumul.*

I. *Généralités sur les familles exponentielles.*

1° Démontrer que les familles suivantes sont exponentielles :

a. Étant donné un entier strictement positif h, $\Omega = \{0, 1, \ldots, h\}$; $\Theta =]0, 1[$; pour tout θ, P_θ est la loi binomiale d'ordre h et paramètre θ.

b. $\Omega = \mathbf{N}$; $\Theta = \mathbf{R}_+^*$ $(=]0, +\infty[)$; pour tout θ, P_θ est la loi de Poisson de paramètre θ.

c. $\Omega = \mathbf{R}$; $\Theta = \mathbf{R}$; pour tout θ, P_θ est la loi de Laplace-Gauss (dite aussi loi normale) réduite (c'est-à-dire de variance égale à 1), de moyenne θ.

d. $\Omega = \mathbf{R}$; $\Theta = \mathbf{R}_+^*$; pour tout θ, P_θ est la loi de Laplace-Gauss centrée (c'est-à-dire de moyenne égale à 0), de variance θ.

e. $\Omega = \mathbf{R}$; $\Theta = \mathbf{R} \times \mathbf{R}_+^*$; pour tout $\theta = (\theta_1, \theta_2)$, P_θ est la loi de Laplace-Gauss de moyenne θ_1 et variance θ_2.

2° Soit (P_θ) une famille exponentielle, d'ensemble des résultats Ω et ensemble des paramètres Θ ; soit (T, U, *a*, *b*) une μ-représentation de dimension k de (P_θ).

a. Deux mesures σ-finies sur $(\Omega, \mathcal{A})$ sont dites équivalentes si chacune est absolument continue par rapport à l'autre.

Démontrer que, pour toute mesure σ-finie ν équivalente à μ, la famille (P_θ) admet une représentation de référence ν et de dimension k.

Démontrer que, pour tout $\theta^0 \in \Theta$, la famille (P_θ) admet une représentation de référence P_{θ^0} et de dimension k, qui est de type nul.

b. Une application surjective φ, de Θ sur un ensemble Θ', est appelée un codage compatible avec la famille (P_θ) si, et seulement si, pour tout couple $(\theta^1, \theta^2) \in \Theta^2$, tel que $\varphi(\theta^1) = \varphi(\theta^2)$, on a $P_{\theta^1} = P_{\theta^2}$; l'unique famille, notée $(({}_\varphi P)_{\theta'})$, admettant Ω pour ensemble des résultats et Θ' pour ensemble des paramètres, définie par

$$(\forall \theta \in \Theta) \qquad P_\theta = ({}_\varphi P)_{\varphi(\theta)},$$

est dite *codée* de (P_θ) par φ.

Soit $\Theta' = U(\Theta)$.

Démontrer que U est un codage compatible avec la famille (P_θ).

Démontrer que la famille $(({}_U P)_{\theta'})$, codée de (P_θ) par U, est exponentielle et admet une μ-représentation de dimension k, paramètres-identique.

c. f étant une application mesurable de $(\Omega, \mathcal{A})$ dans un espace mesurable $(\Omega', \mathcal{A}')$, on appelle *transformée* de (P_θ) par f la famille, notée $(({}^f P)_\theta)$, admettant Ω' pour ensemble des résultats et Θ pour ensemble des paramètres, et où, pour tout $\theta \in \Theta$, $({}^f P)_\theta$ est l'image de P_θ par f (c'est-à-dire que, pour tout $A' \in \mathcal{A}'$, on a

$$({}^f P)_\theta(A') = P_\theta\,[f^{-1}(A')]) \quad .$$

Soit $\Omega' \in \mathcal{B}^k$, tel que $T(\Omega) \subset \Omega'$; T peut être considéré comme une application mesurable de $(\Omega, \mathcal{A})$ dans l'espace mesurable euclidien $(\Omega', \mathcal{A}')$.

Soit ${}^T\mu$ la mesure, sur $(\Omega', \mathcal{A}')$, image de μ par T; démontrer que la famille $(({}^T P)_\theta)$, transformée de (P_θ) par T, est exponentielle et admet une ${}^T\mu$-représentation de dimension k, résultats-identique.

II. *Représentation canonique d'une famille exponentielle.*

A partir de toute famille exponentielle admettant une représentation de dimension k on peut obtenir, par les opérations détaillées en I 2° (changement de référence, codage, transformation), une famille exponentielle (P_θ), admettant une représentation de référence P_{θ^0} (où θ^0 appartient à l'ensemble des paramètres), de dimension k, qui est de type nul, paramètres-identique et résultats-identique; si de plus $\theta^0 = 0$ (ce qui est toujours réalisable par un codage défini par une application bijective de $\mathbf{R}^k$ sur lui-même), la P_{θ^0}-représentation est dite *canonique*.

1° Soit P_0 une probabilité sur $(\mathbf{R}^k, \mathcal{B}^k)$.

a. On note b l'application de $\mathbf{R}^k$ dans $\mathbf{R} \cup \{+\infty\}$ définie par

$$(\forall y \in \mathbf{R}^k) \qquad b(y) = \log \int_{\mathbf{R}^k} \exp(x \cdot y)\, P_0(dx) \quad ;$$

soit $\Theta = b^{-1}(\mathbf{R})$; démontrer que Θ est une partie convexe non vide de $\mathbf{R}^k$.

b. On note Ω l'enveloppe convexe fermée de Supp (P_0); démontrer qu'il existe une famille exponentielle, dont l'ensemble des résultats est Ω et l'ensemble des paramètres Θ, admettant une P_0-représentation canonique de fonction de cumul b (on s'autorise les « abus de langage » consistant à confondre :

1° P_0, probabilité sur $(\mathbf{R}^k, \mathcal{B}^k)$, avec la probabilité qu'elle définit, par restriction, sur $(\Omega, \mathcal{A})$ (car $P_0(\Omega) = 1$),

2° b, application de $\mathbf{R}^k$ dans $\mathbf{R} \cup \{+\infty\}$, avec l'application, à valeurs dans $\mathbf{R}$, qui s'en déduit par restriction à Θ).

La famille exponentielle ainsi définie est dite *engendrée* par P_0.

2° Dans chacun des cas ci-dessous (de *a.* à *e.*), caractériser (en donnant l'ensemble des résultats (Ω'), l'ensemble des paramètres (Θ') et la fonction de cumul de la représentation P_0'-canonique (b')) la famille exponentielle engendrée par P_0', et démontrer que cette famille peut être obtenue, par codage et transformation, à partir de la famille exponentielle étudiée dans le cas correspondant de la question I 1°.

a. Étant donné un entier strictement positif h, P_0' est l'unique probabilité sur $(\mathbf{R}, \mathcal{B})$ qui vérifie

$$(\forall\, i \in \{0, 1, \dots, h\}) \qquad P_0'(\{i\}) = C^i_h \left(\frac{1}{2}\right)^h .$$

b. P_0' est l'unique probabilité sur $(\mathbf{R}, \mathcal{B})$ qui vérifie

$$(\forall\, i \in \mathbf{N}) \qquad P_0'(\{i\}) = \frac{1}{e}\,\frac{1}{i\,!} .$$

c. P_0' est la loi de Laplace-Gauss de dimension 1, centrée et réduite.

d. P_0' est la probabilité sur $(\mathbf{R}, \mathcal{B})$ qui admet, pour densité par rapport à la mesure de Lebesgue, la fonction f_0 définie par :

$$\left|\begin{array}{lll} \text{si} \quad x \leqslant 0 \;, & & f_0(x) = 0 \\ \text{si} \quad x > 0 \;, & & f_0(x) = \dfrac{1}{\sqrt{\pi}}\, e^{-x} x^{-\frac{1}{2}} . \end{array}\right.$$

e. P_0' est la probabilité définie sur $(\mathbf{R}^2, \mathcal{B}^2)$ par les conditions suivantes :

$$P_0'(\{(x_1, x_2) \in \mathbf{R}^2 ; \quad x_2 = \tfrac{1}{2}(x_1)^2\}) = 1 ,$$

la première marge de P_0' (c'est-à-dire l'image de P_0' par la projection $\pi_1 : (x_1, x_2) \longmapsto x_1$) est la loi de Laplace-Gauss centrée et réduite.

3° Soit P_0 une probabilité sur $(\mathbf{R}^k, \mathcal{B}^k)$; soit (P_θ) la famille exponentielle (d'ensemble des résultats Ω et ensemble des paramètres Θ) engendrée par P_0, et soit b (application de $\mathbf{R}^k$ dans $\mathbf{R} \cup \{+\infty\}$) la « fonction de cumul » de sa P_0-représentation canonique.

a. Démontrer que, en tout élément θ de l'intérieur $\mathring{\Theta}$ de Θ, la fonction b est dérivable à tous les ordres.

b. Pour tout $i \in \{1, \dots, k\}$, on note π_i la projection de Ω $(\subset \mathbf{R}^k)$ dans $\mathbf{R}$ définie par :

$$(\forall\, (x_j)_{1 \leqslant j \leqslant k} \in \Omega) \qquad \pi_i((x_j)_{1 \leqslant j \leqslant k}) = x_i .$$

Démontrer que, pour tout $\theta \in \mathring{\Theta}$ $(\theta = (\theta_j)_{1 \leqslant j \leqslant k})$, on a les résultats suivants :

— pour tout $i \in \{1, \dots, k\}$, l'*espérance mathématique* de π_i par rapport à P_θ est égale à $\dfrac{\partial b}{\partial \theta_i}(\theta)$;

— pour tout $i \in \{1, \dots, k\}$, la *variance* de π_i par rapport à P_θ est égale à $\dfrac{\partial^2 b}{\partial \theta_i^2}(\theta)$;

— pour tout couple $(i, j) \in \{1, \ldots, k\}^2$, tel que $i \neq j$, la *covariance* de π_i et π_j, par rapport à P_θ, est égale à $\dfrac{\partial^2 b}{\partial\theta_i\,\partial\theta_j}(\theta)$.

III. *Estimation par maximum de vraisemblance.*

Soit P_0 une probabilité sur $(\mathbf{R}^k, \mathcal{B}^k)$; soit (P_θ) la famille exponentielle (d'ensemble des résultats Ω et ensemble des paramètres Θ) engendrée par P_0, et soit b (application de $\mathbf{R}^k$ dans $\mathbf{R} \cup \{+\infty\}$) la « fonction de cumul » de sa P_0-représentation canonique (voir II 1°). On suppose que Supp (P_0) n'est contenu dans aucun hyperplan.

On va étudier le problème de l'estimation, en un résultat ω, du paramètre θ par la méthode du maximum de vraisemblance.

1° On appelle fonction convexe à k dimensions toute application f de $\mathbf{R}^k$ dans $\mathbf{R} \cup \{+\infty\}$ vérifiant :

$$\left|\begin{array}{l} (\forall (x^1, x^2) \in (\mathbf{R}^k)^2) \quad (\forall \lambda \in]0,1[) \\ \qquad\qquad f(\lambda x^1 + (1-\lambda)x^2) \leqslant \lambda f(x^1) + (1-\lambda) f(x^2), \\ f^{-1}(\mathbf{R}) \neq \varnothing \quad (f^{-1}(\mathbf{R}) \text{ est appelé le } \textit{domaine} \text{ de } f, \text{ et noté } D_f). \end{array}\right.$$

Si de plus la fonction convexe f vérifie :

$$(\forall (x^1, x^2) \in (\mathring{D}_f)^2) \quad (\forall \lambda \in]0,1[) \quad :$$

$$x^1 \neq x^2 \quad \Rightarrow \quad f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(x^1) + (1-\lambda) f(x^2) \ ,$$

elle est dite *stricte.*

Démontrer que b est une fonction convexe stricte, de domaine Θ.

2° On rappelle qu'une fonction convexe f est continue sur $\mathring{D}_f \cup \complement \overline{D}_f$; elle est dite *fermée* si elle est semi-continue inférieurement sur $\mathbf{R}^k$, c'est-à-dire vérifie : $(\forall x \in \mathbf{R}^k) \quad f(x) = \liminf\limits_{x' \to x} f(x')$.

(On rappelle que, pour que $l = \liminf\limits_{x' \to x} f(x')$, il faut que :

$$(\forall l' < l) \quad (\exists \eta > 0) \quad (\forall x' \in \mathbf{R}^k) \quad [\,|x' - x| < \eta \ \Rightarrow \ f(x') \geqslant l'\,])$$

Cette condition équivaut aussi au fait, pour f, d'être enveloppe supérieure des fonctions affines qu'elle majore, c'est-à-dire que, pour tout $x \in \mathbf{R}^k$,

$$f(x) = \sup \{\alpha \in \mathbf{R} \ ; (\exists y \in \mathbf{R}^k)(\forall x' \in \mathbf{R}^k)\ y.(x' - x) + \alpha \leqslant f(x')\} \ .$$

a. Démontrer que, pour qu'une fonction convexe f soit fermée, il suffit qu'elle vérifie :

$$(\forall x \in \mathbf{R}^k) \qquad f(x) \leqslant \liminf_{x' \to x} f(x') \ .$$

b. Démontrer que b est une fonction convexe fermée.

3° Deux fonctions convexes fermées à k dimensions, f et g, sont dites *compatibles* si, et seulement si, elles vérifient :

$$(\forall (x, y) \in (\mathbf{R}^k)^2) \qquad f(x) + g(y) \geqslant x.y \ .$$

a. Démontrer que, pour toute fonction convexe fermée f, l'ensemble des fonctions convexes fermées compatibles avec f est non vide, et admet un plus petit élément, noté f^*.

b. Démontrer que, pour toute fonction convexe fermée f, on a $(f^*)^* \leqslant f$.

Soit l une fonction affine; démontrer, en calculant l^* et $(l^*)^*$, que $l = (l^*)^*$.

En déduire que, pour toute fonction convexe fermée f, on a $(f^*)^* = f$.

On dit que deux fonctions convexes fermées f et g sont *conjuguées* si, et seulement si, elles vérifient $f^* = g$ (ce qui équivaut à $f = g^*$).

c. Démontrer que l'estimation du paramètre par maximum de vraisemblance, en l'observation $\omega \in \Omega$, est, *s'il existe et est unique*, l'élément θ de Θ tel que $b(\theta) + b^*(\omega) = \theta \cdot \omega$.

4° On sait que l'ensemble des paramètres Θ est égal à D_b, domaine de la fonction convexe b; le but de cette question est de comparer Ω, ensemble des résultats, et D_{b^*}, domaine de la fonction convexe b^*.

On note S la boule unité fermée de $\mathbf{R}^k$; pour tout $x \in \mathbf{R}^k$ et tout $\delta > 0$, $x + \delta S$ désigne la boule fermée de centre x et rayon δ

$$x + \delta S = \{ x' \in \mathbf{R}^k \quad ; \quad |x' - x| \leqslant \delta \}$$

a. Nous allons démontrer que $D_{b^*} \subset \Omega$.

Soit $x_0 \in \complement_{\mathbf{R}^k} \Omega$.

Démontrer qu'il existe $y_0 \in S$, et $\alpha < 0$, tels que, pour tout $\omega \in \Omega$, on ait $y_0 \cdot (\omega - x_0) \leqslant \alpha$.

Démontrer qu'on a :

$$\lim_{r \to +\infty} [(ry_0) \cdot x_0 - b(ry_0)] = +\infty.$$

En déduire

$$x_0 \notin D_{b^*} .$$

b. Nous allons démontrer que $\mathring{\Omega} \subset D_{b^*}$.

Soit $\omega_0 \in \mathring{\Omega}$.

On rappelle que Ω est l'enveloppe convexe fermée de Supp (P_0); on admet qu'il en résulte qu'il existe Ω', partie *finie* de Supp (P_0), telle que ω_0 appartienne à l'intérieur de l'enveloppe convexe de Ω'.

Soit y un élément de $\mathbf{R}^k$ de norme 1 ; on note :

$H(y)$	l'hyperplan	$\{ x \in \mathbf{R}^k \;;\; y \cdot x = y \cdot \omega_0 \}$	,
$K^+(y)$	le demi-espace	$\{ x \in \mathbf{R}^k \;;\; y \cdot x \geqslant y \cdot \omega_0 \}$	,
$K^-(y)$	le demi-espace	$\{ x \in \mathbf{R}^k \;;\; y \cdot x \leqslant y \cdot \omega_0 \}$	;

pour tout $\omega' \in \Omega'$, on note $d(\omega', y)$ la distance de ω' à $H(y)$

$$\left(d(\omega', y) = \inf_{x \in H(y)} |\omega' - x|\right)$$

et on pose

$$d(y) = \min\left[\max_{\omega' \in K^+(y)} d(\omega', y) \quad , \quad \max_{\omega' \in K^-(y)} d(\omega', y) \right] ;$$

Démontrer que dans ces conditions on a : $d(y) > 0$.

Soit $d_0 = \inf_{|y|=1} d(y)$; démontrer qu'on a : $d_0 > 0$.

Démontrer que, pour tout élément ω de Supp (P_0) et tout $\delta > 0$, on a : $P_0(\omega + \delta S) > 0$.

On pose : $$\alpha = \inf_{\omega' \in \Omega'} P_0(\omega' + d_0 S) .$$

Démontrer qu'on a : $\alpha > 0$.

Démontrer que, pour tout y de norme 1, on a : $P_0[K^+(y)] \geqslant \alpha$.

En déduire : $\omega_0 \in D_{b^*}$.

c. Démontrer l'égalité $\mathring{D}_{b^*} = \mathring{\bar{\Omega}}$.

5° On appelle *sous-gradient*, en un point $x \in \mathbf{R}^k$, de la fonction convexe à k dimensions f, la partie de $\mathbf{R}^k$.

$$\partial f(x) = \{ y \in \mathbf{R}^k \quad ; \quad (\forall x' \in \mathbf{R}^k) \qquad f(x') \geqslant f(x) + y \cdot (x' - x) \} \; .$$

a. Soit $(x, y) \in (\mathbf{R}^k)^2$, et soit (f, g) un couple de fonctions convexes fermées à k dimensions, conjuguées. Démontrer l'équivalence des 3 propositions suivantes :

$$\left| \begin{array}{l} y \in \partial f(x) \quad , \\ f(x) + g(y) = x \cdot y \quad , \\ x \in \partial g(y) \quad . \end{array} \right.$$

En déduire une condition nécessaire et suffisante d'existence, en $\omega \in \Omega$, d'une estimation du paramètre par maximum de vraisemblance.

b. Une fonction convexe fermée à k dimensions est dite *douce* si, et seulement si, en tout $x \in \mathbf{R}^k$, $\partial f(x)$ a au plus un élément; on admettra que cette condition équivaut au couple de conditions suivantes :

pour tout $x \notin \mathring{\mathrm{D}}_f$, $\partial f(x) = \varnothing$,

pour tout $x \in \mathring{\mathrm{D}}_f$, f est dérivable en x et , $df(x)$ dénotant le gradient de f en x, on a $\partial f(x) = \{df(x)\}$.

Démontrer que, si la fonction convexe fermée f est stricte et douce, il en est de même de f^* (on pourra démontrer, à chaque fois par l'absurde, que f^* est douce, puis que f^* est stricte).

c. On suppose que la fonction b est douce.

Démontrer qu'alors l'estimation par maximum de vraisemblance

n'est pas définie aux points ω appartenant à la frontière de Ω,

est définie en tout point ω appartenant à $\mathring{\Omega}$, et vaut $db^*(\omega)$ (dire quelle est alors l'espérance mathématique, par rapport à $\mathrm{P}_{db^*(\omega)}$, de chacune des projections π_i $(1 \leqslant i \leqslant k)$ (voir II 3°)) .

6° Pour chacun des exemples étudiés en II 2°, on demande de :

a. s'assurer que la fonction de cumul est douce,

b. calculer la fonction $(b')^*$,

c. préciser $\mathring{\Omega}'$ et donner la valeur, pour tout $\theta' \in \Theta'$, de $\mathrm{P}'_{\theta'}(\mathring{\Omega}')$,

d. calculer, en tout $\omega' \in \mathring{\Omega}'$, la valeur de l'estimation par maximum de vraisemblance.

RAPPORT SUR L'EPREUVE DE PROBABILITES ET STATISTIQUES

GENERALITES

Ce problème est consacré aux familles exponentielles de Probabilités sur un espace mesurable euclidien. Ces familles exponentielles jouent un rôle essentiel dans toute la Statistique Mathématique, car les structures statistiques où elles interviennent fournissent l'existence d'outils (statistiques exhaustives et complètes, rapports de vraisemblance monotone) qui permettent de construire des procédures statistiques possédant de nombreuses qualités (estimateurs sans biais de variance minimum, tests uniformément plus puissants pour le problème de test unilatéral relatif au paramètre d'une structure exponentielle à une dimension,...) ; nous renvoyons le lecteur intéressé par ces propriétés aux chapitres X et XI de [2].

La définition d'une famille exponentielle est donnée en introdution au problème ; il s'agit d'une famille de probabilités absolument continues par rapport à une mesure σ- finie donnée (ces notions figurent en [3], II, 2° (sous la forme "théorème de Radon-Nikodym") et dont les densités ont une forme remarquable donnée.

Dans la première partie, on s'assure tout d'abord qu'un certain nombre de familles classiques de lois de probabilités sont exponentielles ; les seuls outils nécessaires sont l'absolue continuité et, bien sûr, la connaissance des lois étudiées : binomiale, Poisson, Laplace-Gauss ([4], I, 6e alinéa). On étudie ensuite le comportement des familles exponentielles relativement aux opérations suivantes : changement de mesure de référence, changement de l'espace des paramètres, changement de l'espace des observations ; toutes ces études ne font intervenir que le théorème de Radon-Nikodym et le changement de variable dans les intégrales.

La deuxième partie est consacrée à la génération d'une famille exponentielle par une probabilité (la famille est alors obtenue sous forme d'une représentation dite canonique) ; il s'agit là d'une procédure fort classique, qu'on peut trouver par exemple dans [I], 5, p. 5.4 ; et [2], chap. X, § 2 (où l'on considère, plus généralement, la famille exponentielle engendrée par une "mesure modérée" sur $\mathbb{R}^n$). On donne, pour chacune des familles étudiées en I, une probabilité qui l'engendre ; vérifier, dans chaque cas, qu'il en est bien ainsi, ne fait intervenir aucune notion nouvelle, mais exige une certaine habileté de calcul. On précise ensuite le lien entre les dérivées successives de la "fonction de cumul" de la famille exponentielle engendrée par une probabilité P_0, et les moments des différentes probabilités de cette famille ; il s'agit là d'un résultat bien classique (voir par exemple [2], X, §2, th.2) fondé sur des calculs formellement simples, (dérivations sous le signe somme), mais qui doivent être justifiés ([3], II,3).

La troisième partie, la plus proprement statistique, est consacrée à l'estimation par maximum de vraisemblance ([4], IV, 2e alinéa) du paramètre d'une famille exponentielle ; cette étude est extraite de [I], 6. La recherche de maxima que cette étude nécessite utilise la convexité de la fonction de cumul ; pour démontrer que la fonction de cumul est strictement convexe fermée on utilise des propriétés des espaces L ([3], II, I.), (inégalité de Hölder), et la convergence sous le signe somme ([3], II,3), (théorème de Fatou-Lebesgue). On applique ensuite à la fontion de cumul la théorie des fonctions convexes conjuguées ; on donne ou on fait redémontrer les éléments ici nécessaires de cette théorie (que le lecteur intéressé pourra trouver dans [5], 12). Cette dernière étude nécessite, pour être menée à bien, une réelle aisance dans le maniement de l'analyse.

Partie I

I 1°) 99 candidats à l'agrégation masculine (sur 451 ayant composé en Probabilités et Statistiques) ont obtenu la note 0 à cette épreuve ; ils ont donc arrêtés dès cette première question ; les raisons essentielles semblent être, outre chez certains l'incapacité de lire un énoncé - ici d'assimiler la notion même de famille exponentielle -, l'ignorance des lois de probabilités élémentaires considérées, et l'incompréhension de la notion d'absolue continuité ; en particulier, si la plupart des candidats savent que la probabilité de Laplace-Gauss est "absolument continue" (en oubliant souvent de préciser que c'est par rapport à la mesure de Lebesgue), nombreux sont ceux qui ne savent pas que les probabilités définies sur un ensemble fini ou dénombrable Ω sont toutes absolument continues par rapport à la mesure de dénombrement (cardinalité), la densité p d'une probabilité P par rapport à cette mesure étant donnée par : $(\forall\, \omega \in \Omega)\ p(\omega) = P(\{\omega\})$.

I 2°) a. Tout repose ici sur le caractère strictement positif des densités utilisées ; de nombreux candidats ont été pénalisés pour ne pas l'avoir remarqué.

b. Question facile, en général correctement traitée ; il faut cependant démontrer que b se factorise à travers U, ce qui résulte de l'égalité (due au fait que, pour tout θ, P_θ est une probabilité): $b(\theta) = \log \int_{\mathbb{R}^k} \exp\,[T(\omega).U(\theta) - a(\omega)]\ p(d\omega)$; ceci entraîne que U est un codage compatible avec la famille (P_θ), et permet d'introduire b', fonction de cumul de la famille codée (application de Θ' dans $\mathbb{R}$, telle que $b = b' \circ U$) ; un certain manque de contrôle dans la rédaction (confusion entre b et b', absence de précision sur les espaces de départ et d'arrivée des applications en jeu) rend, dans d'assez nombreuses copies, ce passage très confus.

c. La plupart des candidats ne savent pas, étant données deux probabilités P et μ, définies sur un même espace (Ω, α), (P absolument continue par rapport à μ), et une application T, mesurable de (Ω, α) dans (Ω', α'), exprimer f', densité de ${}^{T}P$ par rapport à ${}^{T}\mu$(${}^{T}P$ et ${}^{T}\mu$ sont les images de P et μ par l'application T) en fonction de f, densité de P par rapport à μ : cette densité f' est l'espérance conditionnelle de f, relativement à la mesure μ, et par rapport à l'application T. L'ignorance de ce résultat est sans gravité, si on se ramène, grâce à a), au cas où la famille exponentielle est de type nul ; en effet alors toutes les densités f_θ des probabilités P_θ par rapport à la mesure de référence μ se factorisent à travers T ; on a $f_\theta = f'_\theta \circ T$, et on peut vérifier directement (simple calcul de changement de variable dans les intégrales) que f'_θ est la densité ${}^{T}P_\theta$ par rapport à ${}^{T}\mu$; plusieurs candidats, ne songeant pas à se ramener à une famille de type nul, mais voyant que cela les arrangeait, ont affirmé que (ou fait "comme si"), dans le cas général, la densité f_θ se factorise à travers T, ce qui est évidemment faux.

Partie II

II 1°) Pour que $b(y)$ soit fini, il faut et il suffit que $\exp b(y)$ le soit ; la convexité de l'ensemble des y tels que $\exp b(y)$ est fini est alors une conséquence immédiate de la convexité de l'exponentielle ; des candidats ayant voulu démontrer ici la convexité de l'application b qui n'est demandée qu'en III 1°) ont commis des erreurs que nous analysons plus loin.

II 2°) Pour être lisible par les correcteurs, cette question exige une rédaction et une présentation claire des résultats (avec respect des notations de l'énoncé !) qui font souvent défaut ; certains ont semblé trouver difficiles les calculs suivants :

en d. retrouver la loi du κ^2 à un degré de liberté à partir de la loi de Laplace-Gauss de dimension 1 (transformation de la densité par l'application $x \rightsquigarrow x^2$).

en e. intégrer par rapport à la probabilité P_b , définie sur $(\mathbb{R}^2, \mathscr{B}^2)$ mais concentrée sur la parabole d'équation $x_2 = \frac{1}{2}(x_1)^2$ [on a évidemment, P_b^1 désignant ici la première projection de P_b : $\int_{\mathbb{R}^2} f(x_1,x_2)\, P_b(dx_1,dx_2) = \int_{\mathbb{R}} f(x_1, \frac{1}{2}(x_1)^2)\, P_b^1(dx_1)$, égalité que les candidats qui en font usage omettent de justifier ; la justification la plus simple est bien sûr la P_b - $p.s.$ égalité des applications f et $(x_1,x_2) \rightsquigarrow f(x_1, \frac{1}{2}(x_1)^2)$].

II 3°) a. Cette question n'a été que très exceptionnellement correctement traitée ; les correcteurs se satisfaisaient cependant d'une démonstration correcte de la dérivabilité à l'ordre 1, par rapport à la variable θ_1 (le cas d'un ordre supérieur ne faisant pas appel à des arguments nouveaux) ; mais il fallait au moins :

- énoncer correctement qu'on devait trouver, pour tout $\theta^o \in \overset{o}{\Theta}$, un voisinage V_{θ^o} de θ^o, contenu dans $\overset{o}{\Theta}$, et une application M_{θ^o}, P_o-intégrable, de $\mathbb{R}^k$ dans $\mathbb{R}_+$, de telle sorte que, pour tout $\theta \in V_{\theta^o}$, M_{θ^o} domine $\frac{\partial b}{\partial \theta_1}(\theta)$;

- présenter les principes de la construction de M_{θ^o} (on peut, étant donné $\varepsilon > 0$ tel que $\prod_{i=1}^{k} [\theta_i^o - \varepsilon, \theta_i^o + \varepsilon] \subset \overset{o}{\Theta}$, prendre $V_{\theta^o} = \prod_{i=1}^{k}[\theta_i^o - \frac{\varepsilon}{2}, \theta_i^o + \frac{\varepsilon}{2}]$ et choisir $L \geqslant 1$ tel que, pour tout $x \geqslant L$, on ait $x \leqslant e^{x\frac{\varepsilon}{2}}$; on peut alors définir M_{θ^o} par :

pour $|x_1| > L$, $\quad M_{\theta^o}(x_1, \ldots, x_k) = \prod_{i=1}^{k} \exp\, [x_i\, (\theta_i^o + sgn\, (x_i)\, \frac{\varepsilon}{2})]$

pour $|x_1| \leqslant L$, $\quad M_{\theta^o}(x_1, \ldots, x_k) = |x_1| \prod_{i=1}^{k} \exp\, [x_i\, (\theta_i^o + sgn\, (x_i)\, \frac{\varepsilon}{2})]$

Les correcteurs ont souvent regretté que les candidats n'aient pas préféré ne pas traiter cette question plutôt que de fournir de fausses justifications de dérivabilité sous le signe somme ; ils tiennent à insister sur la gravité des lacunes que l'on rencontre ici chez presque tous les candidats (même fort bons par ailleurs).

Partie III

III 1°) La convexité de la fonction b est une conséquence directe de l'inégalité de Hölder (ou de l'inégalité de Schwartz si on se contente de démontrer que

$$[\ \forall (x^1,x^2) \in (\overset{o}{D} f)^2\ ] \qquad b\left(\frac{x^1 + x^2}{2}\right) \leqslant \frac{b(x^1) + b(x^2)}{2}$$

ce qui est licite à condition de s'assurer que b est continue) ; de nombreux candidats ont cru démontrer cette convexité de proche en proche, en utilisant successivement la convexité de la fonction exponentielle, la linéarité et positivité de l'intégrale, puis enfin la "convexité" de la fonction logarithme (qui est évidemment concave !).

Peu de candidats ont vu l'importance, pour la **stricte** convexité de b, de l'hypothèse "Supp (P_o) n'est contenu dans aucun hyperplan" ; or, étant données 2 fonctions f et g, P_o - intégrables, et deux nombres réels p et q, strictement supérieurs à 1, tels que $\frac{1}{p} + \frac{1}{q} = 1$, l'"égalité de Hölder" $\int_{\mathbb{R}^k} |f(x)\, g(x)|\, P_o(dx) = \left(\int_{\mathbb{R}^k} |f(x)|^p\, P_o(dx)\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^k} |g(x)|^q\, P_o(dx)\right)^{\frac{1}{q}}$

équivaut à la P_o- presque sûre égalité de f et g, autrement dit, si f et g sont continues, à leur coïncidence sur Supp (P_o) ; ici où on applique l'inégalité de Hölder à $f(x) = e^{\lambda(x.y^1)}$ et $g(x) = e^{\lambda(x.y^2)}$, cette condition équivaut au fait pour Supp (P_o) d'être contenu dans l'hyperplan $\{x \in \mathbb{R}^k\ ;\ x.(y^1 - y^2) = 0\}$ ce qui est contraire à l'hypothèse faite.

III 2°) Question rarement abordée, mais les candidats qui y sont arrivés ont en général su voir que l'inégalité $b(y) \leqslant \liminf_{y' \to y} b(y')$ est vérifiée en vertu du théorème de Fatou-Lebesgue, appliqué à une suite de fonctions positives $x \rightsquigarrow e^{y_n \cdot x}$ (où $\lim_{n \to \infty} y_n = y$) (on a en effet alors $\int_{\mathbb{R}^k} (\liminf_{n \to \infty} e^{y_n \cdot x})\, P_0\,(dx) < \liminf_{n \to \infty} \int_{\mathbb{R}^k} e^{y_n \cdot x}\, P_0(dx)$).

III 3°) Ceux des candidats qui ont essayé de traiter cette question ont souvent échoué faute d'avoir d'avoir su effectuer le calcul, pourtant facile, qui permet d'obtenir l^* (si $y \neq y_0$, $l^*(y) = +\infty$; $l(y_0) = \alpha$).

III 4°) à 6°) Ces questions n'ayant été que très rarement abordées, nous renvoyons le lecteur intéressé à l'ouvrage [1], 6, d'où elles ont été extraites.

CONCLUSION

Mise à part l'ignorance de nombre d'entre eux sur certaines notions de base du calcul des probabilités (lois fondamentales en particulier) il semble que les candidats aient surtout éprouvé des difficultés dues à leur manque de maîtrise des méthodes de l'analyse (dérivation sous le signe somme, par exemple) et à leur maladresse dans la conduite de calculs pourtant élémentaires.

BIBLIOGRAPHIE

[1] BARNDORFF NIELSEN O. *Exact exponential families* Varions Publications Series, 19, Mathematisk Institut, Aarhus Universitet, 1971.

[2] BARRA J.R. *Notions fondamentales de Statistique Mathématique* Dunod, Paris, 1971.

[3] PROGRAMMES des épreuves écrites et orales des Agrégations de Mathématiques (candidats et candidates) pour la session de 1970, *Composition d'Analyse*, Note du 24 Septembre 1969 ; *Bull. Off. Ed. Nat.* n° 37 (2-10-1969) (reconduits pour 1973).

[4] PROGRAMMES des épreuves écrites et orales des Agrégations de Mathématiques (candidats et candidates) pour la session de 1970, *Probabilités et Statistiques*, note du 24 Septembre 1969, *Bull. Off. Ed. Nat.* n° 37 (2-10-1969) (reconduits pour 1973).

[5] ROCKAFELLAR R.T. *Convex Analysis*, Princeton University Press 1970.

RENSEIGNEMENTS STATISTIQUES

Agrégation masculine :
Nombre de copies corrigées : 451
Moyenne : 10,9 sur 40
Quantiles supérieurs :
- d'ordre 2 (médiane): 10 (223 notes ⩾ 10)
- d'ordre 3 : 14 (156 notes ⩾ 14)
- d'ordre 4 : 17 (119 notes ⩾ 17)
- d'ordre 5 : 19 (92 notes > 19)

Agrégation féminine :
Nombre de copies corrigées : 264
Moyenne : 11,56 sur 40
Quantiles supérieurs :
- d'ordre 2 (médiane) : 10 (134 notes > 10)
- d'ordre 3 : 15 (92 notes > 15)
- d'ordre 4 : 18 (71 notes > 18)
- d'ordre 5 : 20 (59 notes > 20)

Répartition des notes par classes d'amplitude 5 :

	0	1 à 5	6 à 10	11 à 15	16 à 20	21 à 25	26 à 30	31 à 35	36 à 40
Candidats	99	77	61	88	52	41	10	17	6
Candidates	51	39	42	44	36	25	17	7	3

(Candidats : 2 copies ont obtenu la note 40)
(Candidates : une note 40, une note 38, une note 37).