PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 3, Fasc. 2 (1984), p. 259-282

SUR UN THÉORÈME DE MINIMAX ET SON APPLICATION AUX TESTS

PAR

LUCIEN BIRGÉ (PARIS)

Abstract. In this paper we try to extend in some sense the results of Huber and Strassen concerning the tests between two sets which are under two-alternating capacities. We just assume here that those sets are convex and weakly compact (this last assumption may be weakened in good cases). We get non-asymptotic bounds for the errors of our tests and prove that these bounds are asymptotically optimal in the sense of an exponential rate of decrease. These results are applied to the particular case of Hellinger balls: we find very simple tests with good properties which are proved to be useful in estimation theory.

1. Introduction. Soient $(\Omega, \mathscr{A})$ un espace mesuré, $\mathscr{M}_1$ l'ensemble des mesures de probabilité sur $(\Omega, \mathscr{A})$ et $\mathscr{P}$, $\mathscr{Q}$ deux sous-ensembles disjoints de $\mathscr{M}_1$. Le problème que nous allons considérer, à savoir l'obtention de „bons" tests de $\mathscr{P}$ contre $\mathscr{Q}$, a été abordé par de nombreux auteurs et n'est que très partiellement résolu. Si l'on met à part certains cas paramétriques très particuliers pour lesquels il existe des tests uniformément plus puissants, il n'y a pas de moyen canonique pour comparer les tests et l'on peut distinguer deux points de vue largement utilisés et adaptés en fait aux structures des ensembles à tester.

1. *Le point de vue minimax.* Supposons que l'on dispose de n observations; une *fonction de test* $\varphi_n(x_1, \ldots, x_n)$ est une application mesurable de Ω^n dans $[0, 1]$. Posons

$$\mathscr{P}^n = \{P_1 \otimes \ldots \otimes P_n \mid P_i \in \mathscr{P}\}.$$

Le point de vue minimax, utilisé le plus souvent lorsque $\mathscr{P}$ et $\mathscr{Q}$ sont des voisinages de P_0 et Q_0 respectivement, pour une certaine topologie, définit le niveau α_n et l'erreur β_n des tests de $\mathscr{P}^n$ contre $\mathscr{Q}^n$ par

$$(1.1)\quad \alpha_n = \sup_{P \in \mathscr{P}^n} \int \varphi_n(x_1, \ldots, x_n)\, dP, \qquad \beta_n = \sup_{Q \in \mathscr{Q}^n} \int [1 - \varphi_n(x_1, \ldots, x_n)]\, dQ,$$

c'est-à-dire que l'on suppose les variables indépendantes mais non nécessairement équidistribuées. Un test sera dit *minimax* s'il minimise l'erreur pour une valeur donnée du niveau. La recherche de tels tests est difficile et on ne les obtient qu'en imposant des hypothèses assez restrictives sur $\mathscr{P}$ et $\mathscr{Q}$.

Les résultats les plus importants ont été obtenus par Huber et Strassen [10] (voir aussi Birgé [3], Bednarski [2], Tusnády [19]) et concernent le cas où $\mathscr{P}$ et $\mathscr{Q}$ sont deux ensembles convexes faiblement compacts et dominés par des capacités bi-alternatives. Ceci englobe le cas des boules pour les distances en variation totale et de Prokhorov. Pour tout n et tout niveau α_n, Huber et Strassen ont prouvé l'existence de tests minimax qui sont simplement les tests de rapport de vraisemblance, au niveau α_n, d'un couple de probabilités $(\bar{P}, \bar{Q})$ dans $\mathscr{P} \times \mathscr{Q}$, indépendant de n et α_n. Ce couple „le plus défavorable" a été construit explicitement dans divers cas par Huber [9], Rieder [16] et Österreicher [15].

Par contre, en l'absence de domination par des capacités bi-alternatives, de tels couples n'existent plus (cf. [2]). Cette hypothèse de domination est assez restrictive et exclut le cas des boules pour la distance de Hellinger définie par

$$(1.2)\qquad h(P, Q) = \tfrac{1}{2} \int \left(\sqrt{dP} - \sqrt{dQ}\right)^2$$

(cf. [11]), ainsi que d'autres convexes plus généraux.

Le Cam [11] a obtenu d'importants résultats sur les tests de deux ensembles convexes en majorant la valeur minimale de $\alpha_n + \beta_n$, mais il ne détermine pas cette valeur ni ne donne l'expression explicite des tests qui permettraient d'obtenir les majorations indiquées.

2. *Les vitesses de décroissance exponentielle optimales.* Lorsque l'on veut tester deux ensembles $\mathscr{P}$ et $\mathscr{Q}$ plus généraux (en particulier $\mathscr{P}$ contre $\mathscr{P}^c$) on est amené àdéfinir le niveau α_n et l'erreur $\beta_n(Q)$ en Q par

$$\alpha_n = \sup_{P \in \mathscr{P}} \int \varphi_n(x_1, \ldots, x_n)\, dP \otimes \ldots \otimes dP$$

et

$$\beta_n(Q) = \int [1 - \varphi_n(x_1, \ldots, x_n)]\, dQ \otimes \ldots \otimes dQ,$$

en supposant que les variables sont indépendantes équidistribuées.

Si l'on fixe α_n, on ne sait généralement pas minimiser $\beta_n(Q)$ pour tout Q dans $\mathscr{Q}$ et l'on adopte le point de vue asymptotique suivant fondé sur le fait que α_n et $\beta_n(Q)$ se comportent le plus souvent comme des fonctions exponentielles de n. Pour une suite de fonctions de tests $\{\varphi_n\}_{n\in N}$, on définit un log-niveau α et une log-erreur $\beta(Q)$ en Q par

$$\alpha = \overline{\lim}\, n^{-1}\log \alpha_n \quad \text{et} \quad \beta(Q) = \overline{\lim}\, n^{-1}\log \beta_n(Q).$$

$\alpha < 0$ étant fixé, on cherche des suites de tests qui minimisent $\beta(Q)$ pour tout Q dans $\mathcal{Q}$. Les premiers résultats de ce type ont été obtenus par Brown [6] et généralisés par Tusnády [18] puis Birgé [4]. Cependant, l'optimalité de tels tests est essentiellement asymptotique et il est clair, d'après leur construction même, que leurs performances pour des valeurs moyennes ou même assez grandes de n risquent d'être très médiocres. En plus, leur construction paraît difficilement utilisable pratiquement et leur champ d'application sera donc réduit.

Pour pallier ces inconvénients, nous allons adopter ici un point de vue intermédiaire qui permet d'affaiblir les hypothèses de Huber et Strassen en s'affranchissant de la domination par des capacités bi-alternatives et, tout en conservant l'optimalité asymptotique au sens des erreurs exponentielles, de garantir de „bonnes performances" pour n fini. Le niveau α_n et l'erreur β_n des tests φ_n sont définis par (1.1) et seront tels que $\beta = \overline{\lim}\, n^{-1}\log\beta_n$ est minimal pour une valeur de $\alpha = \overline{\lim}\, n^{-1}\log\alpha_n$ donnée (optimalité asymptotique). De plus, pour tout $\varepsilon > 0$, on pourra choisir φ_n de manière que l'on ait

$$\alpha_n \leqslant (1+\varepsilon)\alpha^n \quad \text{et} \quad \beta_n \leqslant (1+\varepsilon)\beta^n. \tag{1.3}$$

L'application principale de ces résultats sera la possibilité d'obtenir des tests explicites et très simples entre deux boules pour la distance de Hellinger, ce qui pourra être utilisé dans la construction d'estimateurs (cf. [11], [12] et [5]). De plus, dans ce cas, les inégalités (1.3) sont réalisées avec $\varepsilon = 0$.

Le second chapitre sera consacré à l'étude des transformées de Hellinger entre les éléments de $\mathcal{P}$ et $\mathcal{Q}$ et de leur enveloppe supérieure. Dans le troisième nous obtiendrons un certain nombre de théorèmes de minimax que nous appliquerons aux tests dans un quatrième chapitre afin d'obtenir les résultats annoncés. Enfin le dernier chapitre sera consacré à l'application aux boules en distance de Hellinger. On y montrera en particulier que ces boules ne sont pas dominées par des capacités bi-alternatives, et qu'on ne peut donc pas leur appliquer la théorie de Huber et Strassen.

2. Les propriétés de l'enveloppe des transformées de Hellinger. Si P et Q sont deux probabilités sur $(\Omega, \mathcal{A})$ et μ une mesure qui domine $P+Q$, on peut définir pour tout t dans $[0, 1]$

$$\Psi(t, P, Q) = \int (dQ/d\mu)^t (dP/d\mu)^{1-t}\, d\mu = \int dQ^t\, dP^{1-t}, \tag{2.1}$$

en adoptant la convention $0^0 = 0$. Cette quantité est indépendante de μ et définit une fonction continue sur [0, 1] à valeurs dans [0, 1] dite *transformée de Hellinger* du couple (P, Q) et telle que

$$\Psi(0, P, Q) = P\{dQ/dP > 0\}, \quad \Psi(1, P, Q) = Q\{dP/dQ > 0\}.$$

Nous utiliserons aussi $\psi(t, P, Q) = \log \Psi(t, P, Q)$. La fonction $\psi(t, P, Q)$ est convexe et de classe $\mathscr{C}^1$ sur [0, 1], $\mathscr{C}^\infty$ sur]0, 1[(sauf le cas trivial où P et Q sont étrangères, d'où $\psi(t, P, Q) \equiv -\infty$).

On sait (cf. [8], [19] et [4]) que les log-erreurs et les log-niveaux des tests de rapports de vraisemblance entre P et Q sont déterminés par la fonction ψ et plus précisément par ses tangentes. On peut s'attendre à ce que ceux des tests de $\mathscr{P}$ contre $\mathscr{Q}$ soient liés à la fonction $\bar{\psi}$ définie sur [0, 1] par

$$\bar{\psi}(t) = \sup_{\substack{P \in \mathscr{P} \\ Q \in \mathscr{Q}}} \psi(t, P, Q). \tag{2.2}$$

Pour le montrer, nous aurons besoins des hypothèses suivantes:

Hypothèses (T). *L'ensemble $\mathscr{M}$ des mesures finies sur $(\Omega, \mathscr{A})$ est muni d'une topologie $\mathscr{T}$ d'espace localement convexe séparé telle que*:

(T1) *$\mathscr{P}$ et $\mathscr{Q}$ sont convexes et compacts.*

(T2) *Il existe un ensemble F de fonctions mesurables positives telles que pour tout f dans F et t dans]0, 1[les intégrales $\int f^t dP$ (resp. $\int f^{t-1} dQ$) définissent des fonctions de $\mathscr{P}$ (resp. $\mathscr{Q}$) dans $\mathbf{R}$ qui sont semi-continues supérieurement (s. c. s.).*

(T3) *Pour tous t dans]0, 1[, P dans $\mathscr{P}$, Q dans $\mathscr{Q}$ et $\varepsilon > 0$, on peut trouver f dans F tel que*

$$(1-t)\int f^t dP + t\int f^{t-1} dQ < \Psi(t, P, Q) + \varepsilon. \tag{2.3}$$

Nous ferons aussi usage de la condition suivante:

Hypothèse (AC). *Tous les éléments de $\mathscr{P}$ (resp. $\mathscr{Q}$) sont mutuellement absolument continus.*

Il est bien clair que les hypothèses (T) subsistent si l'on remplace F par

$$F' = \{\lambda f \mid \lambda \in \mathbf{R}^+, f \in F\},$$

ce qui nous permettra de supposer toujours que F est un cône.

Exemples. Ces hypothèses sont à rapprocher de celles utilisées dans [4]. Il est facile de voir que (T2) et (T3) seront satisfaites si F est un ensemble de fonctions mesurables, positives, log-bornées (dont le logarithme est borné), dense dans tous les espaces $L^1_+(P+Q)$ (c'est-à-dire la partie de $L^1(P+Q)$ formée de fonctions positives p.s.) tel que si f est dans F, f^t l'est aussi pour t dans $[-1, 1]$ et l'ensemble des mesures est muni de la topologie définie par $\mu_n \to \mu$ si et seulement si $\int f d\mu_n \to \int f d\mu$ pour tout $f \in F$.

On peut en particulier considérer les cas suivants:

(i) F est l'ensemble de toutes les fonctions mesurables log-bornées et la topologie est celle de la norme.

(ii) F est formé des combinaisons linéaires de fonctions indicatrices d'ensembles appartenant à une classe qui engendre la σ-algèbre $\mathscr{A}$.

(iii) F est l'ensemble des fonctions continues log-bornées sur l'espace métrique σ-compact Ω muni de la σ-algèbre borélienne $\mathscr{A}$, et la topologie est celle de la convergence étroite.

Dans tous ces cas, (2.3) est obtenu en choisissant pour f une fonction „proche" du rapport de vraisemblance dQ/dP comme le montre la preuve du théorème 3.2.

Nous noterons désormais s.c.i. (resp. s.c.s.) pour semi-continu inférieurement (resp. supérieurement).

Propriétés de $\bar{\psi}$. Il est clair que la fonction $\bar{\psi}$ est convexe et s.c.i., donc continue sur $[0, 1]$; elle admet des dérivées à droite $\bar{\psi}'_d$ et à gauche $\bar{\psi}'_g$ qui sont égales sauf en un nombre dénombrable de points. Nous supposerons désormais sans toujours le rappeler que les hypothèses (T) sont satisfaites.

PROPOSITION 2.1. *Soient t dans $[0, 1]$ et c un nombre tel que*

$$\bar{\psi}'_g(t) \leqslant c \leqslant \bar{\psi}'_d(t) \quad si\ t \in]0, 1[,$$

$$\bar{\psi}'_d(t) = c \quad si\ t = 0,$$

$$\bar{\psi}'_g(t) = c \quad si\ t = 1.$$

Alors il existe un couple $(\bar{P}, \bar{Q})$ dans $\mathscr{P} \times \mathscr{Q}$, dépendant de t, tel que

$$\bar{\psi}(t) = \psi(t, \bar{P}, \bar{Q}), \tag{2.4}$$

$$c = \psi'(t, \bar{P}, \bar{Q}). \tag{2.5}$$

De plus, si (P, Q) est un autre couple qui vérifie (2.4) avec t dans $]0, 1[$,

$$d\bar{P}/d\bar{Q} = dP/dQ \quad \bar{P}+P+\bar{Q}+Q \ p.s. \tag{2.6}$$

La démonstration repose sur les deux lemmes suivants:

LEMME 2.1. *Pour toute fonction mesurable positive g, on a*

$$(1-t)\int g^t\, dP + t\int g^{t-1}\, dQ \geqslant \Psi(t, P, Q) \quad pour\ tout\ t \in]0, 1[. \tag{2.7}$$

Preuve. Si μ domine $P+Q$, posons

$$p = \frac{dP}{d\mu}, \quad q = \frac{dQ}{d\mu}, \quad h = g\frac{p}{q}1_{\{pq>0\}};$$

alors

$$\begin{aligned}(1-t)\int g^t\, dP + t\int g^{t-1}\, dQ \\ \geqslant (1-t)\int_{\{pq>0\}} h^t q^t p^{1-t}\, d\mu + t\int_{\{pq>0\}} h^{t-1} q^t p^{1-t}\, d\mu.\end{aligned} \tag{2.8}$$

Le minimum de la fonction $x \to (1-t)x^t + tx^{t-1}$ étant obtenu pour $x = 1$ et égal à 1, la conclusion s'ensuit.

Lemme 2.2. *L'application* $(P, Q) \to \Psi(t, P, Q)$ *est concave pour tout* t *dans* $[0, 1]$. *Pour tous* u *et* t *dans* $]0, 1[$, *on a*

(2.9) $\Psi(t, uP+(1-u)P', uQ+(1-u)Q') \geqslant u\Psi(t, P, Q)+(1-u)\Psi(t, P', Q')$

et l'inégalité est stricte sauf si

$$dP/dQ = dP'/dQ' \quad P+P'+Q+Q' \ p.s.$$

Preuve. On vérifie facilement que l'application $(x, y) \to x^t y^{1-t}$ de $\boldsymbol{R}^+ \times \boldsymbol{R}^+$ dans $\boldsymbol{R}^+$ est concave (y compris pour $t = 0$ ou $t = 1$ si l'on convient que $xy^0 = x1_{\{y>0\}}$) et l'on a si $t \in]0, 1[$

(2.10) $\quad [ux+(1-u)x']^t[uy+(1-u)y']^{1-t} \geqslant ux^t y^{1-t}+(1-u)x'^t y'^{1-t}$

avec égalité seulement si $xy' = yx'$; d'où le résultat.

Preuve de la proposition 2.1. Le lemme 2.1. et (T3) entraînent

(2.11) $\quad \Psi(t, P, Q)$

$$= \inf_{f \in F} \left[(1-t)\int f^t\, dP+t\int f^{t-1}\, dQ\right] \quad \text{pour tout } t \in]0, 1[.$$

Il s'ensuit que $\Psi(t, P, Q)$ est une fonction s.c.s. du couple (P, Q) et qu'elle atteint son maximum sur le compact $\mathscr{P} \times \mathscr{Q}$ en un point (P', Q') qui satisfait à (2.4). Si (P, Q) est un couple analogue,

$$\Psi\left(t, \frac{P+P'}{2}, \frac{Q+Q'}{2}\right) \geqslant \tfrac{1}{2}\Psi(t, P, Q)+\tfrac{1}{2}\Psi(t, P', Q') = \exp[\bar{\psi}(t)].$$

Comme on a nécessairement l'égalité, le lemme 2.2 entraîne (2.6). De plus, tout couple (P, Q) qui vérifie (2.4) vérifie également

(2.12) $$\bar{\psi}'_g(t) \leqslant \psi'(t, P, Q) \leqslant \bar{\psi}'_d(t).$$

Sinon, il existerait un ε tel que $\psi(t+\varepsilon, P, Q) > \bar{\psi}(t+\varepsilon)$, ce qui est impossible.

Soient alors t dans $[0, 1[$ et $\{\varepsilon_n\}_{n\in N}$ une suite de réels > 0, tendant vers 0 telle que $\bar{\psi}'(t+\varepsilon_n)$ existe et tende vers $\bar{\psi}'_d(t)$. D'après ce qui précède et (2.12), il existe des couples (P_n, Q_n) tels que

$$\psi(t+\varepsilon_n, P_n, Q_n) = \bar{\psi}(t+\varepsilon_n) \quad \text{et} \quad \psi'(t+\varepsilon_n, P_n, Q_n) = \bar{\psi}'(t+\varepsilon_n).$$

Par convexité, ceci entraîne

$$\psi(u, P_n, Q_n) \geqslant \bar{\psi}(t+\varepsilon_n)+(u-t-\varepsilon_n)\bar{\psi}'(t+\varepsilon_n) \quad \text{pour tout } u \in [0, 1].$$

Supposons, quitte à extraire une sous-suite que $P_n \to \hat{P}$ et $Q_n \to \hat{Q}$. Comme ψ est s.c.s., on aura

$$\psi(u, \hat{P}, \hat{Q}) \geqslant \lim_n \psi(u, P_n, Q_n) \geqslant \bar{\psi}(t)+(u-t)\bar{\psi}'_d(t) \quad \text{pour tout } u \in [0, 1],$$

ce qui implique que $(\hat{P}, \hat{Q})$ satisfait (2.4), donc (2.12), d'oò $\psi'(t, \hat{P}, \hat{Q}) = \bar{\psi}'_d(t)$. Si $t = 0$, on a seulement la seconde inégalité de (2.12) et on obtient (2.4) et (2.5). Le même raisonnement prouve que si t est dans $]0, 1]$, il existe $(\hat{P}; \hat{Q}')$ tels que

$$\psi(t, \hat{P}', \hat{Q}') = \bar{\psi}(t) \quad \text{et} \quad \psi'(t, \hat{P}', \hat{Q}') = \psi'_g(t).$$

Si $t = 1$, on a obtenu (2.4) et (2.5). Si t est dans $]0, 1[$, on pose

$$\bar{P} = \lambda\hat{P} + (1-\lambda)\hat{P}' \quad \text{et} \quad \bar{Q} = \lambda\hat{Q} + (1-\lambda)\hat{Q}'.$$

D'après (2.6) il vient $d\bar{P}/d\bar{Q} = d\hat{P}'/d\hat{Q}' = d\hat{P}/d\hat{Q}$ pour $\lambda \in [0, 1]$ et d'après le lemme 2.2 on a $\psi(t, \bar{P}, \bar{Q}) = \bar{\psi}(t)$. Comme, en utilisant les notations du lemme 2.1,

$$\Psi'(t, P, Q) = \int p \log\frac{q}{p} \exp\left[t \log\frac{q}{p}\right] d\mu,$$

il est immédiat que

$$\Psi'(t, \bar{P}, \bar{Q}) = \lambda\Psi'(t, \hat{P}, \hat{Q}) + (1-\lambda)\Psi'(t, \hat{P}', \hat{Q}')$$

et donc que

$$\psi'(t, \bar{P}, \bar{Q}) = \frac{\Psi'(t, \bar{P}, \bar{Q})}{\exp[\bar{\psi}(t)]} = \lambda\bar{\psi}'_d(t) + (1-\lambda)\bar{\psi}'_g(t).$$

Il s'ensuit que, pour un choix convenable de λ dans $[0, 1]$, le couple $(\bar{P}, \bar{Q})$ vérifiera à la fois (2.4) et (2.5).

3. Quelques résultats de minimax. Soient c un nombre réel et I un intervalle de $[0, 1]$. On sait (cf. [4]) que le log-niveau des tests de rapports de vraisemblance de P contre Q au seuil c est donné par $\inf_{t \in [0,1]} [\psi(t, P, Q) - tc]$. Posons par analogie

$$\delta_I(c) = \inf_{t \in I} [\bar{\psi}(t) - tc]; \tag{3.1}$$

on aura donc

$$\delta_I(c) \geqslant \sup_{\mathscr{P} \times \mathscr{Q}} \inf_{t \in I} [\psi(t, P, Q) - tc]. \tag{3.2}$$

Si l'on note $\bar{I}$ l'adhérence de I dans $[0, 1]$, on peut démontrer le résultat de minimax suivant:

Proposition 3.1. *On a*

$$\delta_I(c) = \sup_{\mathscr{P} \times \mathscr{Q}} \inf_{t \in I} [\psi(t, P, Q) - tc]. \tag{3.3}$$

De plus, il existe un point u dans $\bar{I}$ et un couple $(\bar{P}, \bar{Q})$ dans $\mathscr{P} \times \mathscr{Q}$ tels que

$$\delta_I(c) = \bar{\psi}(u) - uc = \psi(u, \bar{P}, \bar{Q}) - uc = \inf_{t \in I} [\psi(t, \bar{P}, \bar{Q}) - tc]. \tag{3.4}$$

Preuve. Le résultat est immédiat si I est réduit à un point et il est clair que l'on peut supposer I fermé par continuité. Dans ce cas, il existe u dans I tel que $\delta_I(c) = \bar{\psi}(u) - uc$. Si u est intérieur à I, on a nécessairement $\bar{\psi}'_g(u) \leqslant c \leqslant \bar{\psi}'_d(u)$ et d'après la proposition 2.1 il existe $(\bar{P}, \bar{Q})$ avec $\psi(u, \bar{P}, \bar{Q}) = \bar{\psi}(u)$ et $\psi'(u, \bar{P}, \bar{Q}) = c$. Par convexité on obtient (3.4). Si u est une extrémité de I, *la* gauche par exemple, $c \leqslant \bar{\psi}'_d(u)$. On peut encore trouver $(\bar{P}, \bar{Q})$ tels que $\psi(u, \bar{P}, \bar{Q}) = \bar{\psi}(u)$ et $\bar{\psi}(u, \bar{P}, \bar{Q}) = \bar{\psi}_d(u) \geqslant c$, ce qui prouve (3.4). L'égalité (3.3) en est une conséquence immédiate compte-tenu de (3.2).

Il s'ensuit en utilisant (2.11) que, si I ne contient ni 0 ni 1,

$$\delta_I(c) = \sup_{\mathscr{P}\times\mathscr{Q}} \inf_{t\in I} \inf_{f\in F} \{\log[(1-t)\int f^t\,dP + t\int f^{t-1}\,dQ] - tc\},$$

ou encore, en remplaçant f par fe^c puisque F est un cône,

$$\exp[\delta_I(c)] = \sup_{\mathscr{P}\times\mathscr{Q}} \inf_{t\in I} \inf_{f\in F} [(1-t)\int f^t\,dP + te^{-c}\int f^{t-1}\,dQ]. \tag{3.5}$$

Nous allons appliquer à (3.5) le théorème de minimax classique suivant que l'on trouvera par exemple dans [14]:

THÉORÈME 3.1. *Soient K un espace compact et $\mathscr{F}$ un ensemble convexe de fonctions de K dans $\mathbb{R}\cup\{-\infty\}$ tel que tout f dans $\mathscr{F}$ soit s.c.s. et prenne une valeur $\geqslant 0$. Il existe une probabilité de Radon μ sur K telle que l'on ait*

$$\int_K f(x)\,d\mu(x) \geqslant 0 \quad \textit{pour tout } f\in\mathscr{F}.$$

Si de plus K est un compact convexe d'un espace localement convexe séparé et si tous les éléments f de $\mathscr{F}$ sont concaves, alors on peut supposer que μ est une mesure de Dirac sur K portée par un point x_0.

Soient I et c comme précédemment, à tout élément (t, f) de $I\times F$ on associe une fonction H sur $\mathscr{P}\times\mathscr{Q}$, linéaire et s.c.s. d'après (T2), par

$$H(P, Q) = (1-t)\int f^t\,dP + te^{-c}\int f^{t-1}\,dQ. \tag{3.6}$$

On notera $\mathscr{H}(c, I)$ l'ensemble de ces fonctions lorsque (t, f) parcourt $I\times F$ et $\tilde{\mathscr{H}}(c, I)$ son enveloppe convexe. On obtient, comme conséquence immédiate du théorème 3.1, de la proposition 3.1 et de (3.5),

COROLLAIRE 3.1. *Sous les hypothèses* (T), *il existe un couple $(\bar{P}, \bar{Q})$ dans $\mathscr{P}\times\mathscr{Q}$ tel que, si I ne contient ni 0 ni 1, l'on ait*

$$\begin{aligned}\delta_I(c) &= \inf_{t\in I}[\bar{\psi}(t) - tc] = \inf_{t\in I}[\psi(t, \bar{P}, \bar{Q}) - tc] \\ &= \log\sup_{\mathscr{P}\times\mathscr{Q}} \inf_{H\in\mathscr{H}(c,I)} H(P, Q) = \log \inf_{\tilde{H}\in\tilde{\mathscr{H}}(c,I)} \sup_{\mathscr{P}\times\mathscr{Q}} \tilde{H}(P, Q).\end{aligned} \tag{3.7}$$

Le couple $(\bar{P}, \bar{Q})$ n'est pas unique a priori, seul son rapport de vraisemblance l'est d'après la proposition 2.1. D'autre part, il est très important de noter qu'ici, le ou les couples qui vérifient $\bar{\psi}(t) = \psi(t, \bar{P}, \bar{Q})$ dépendent de t comme le montre l'exemple suivant:

$\Omega = \{1, 2, 3, 4, 5\}$ est muni de la σ-algèbre de ses parties et l'on considère 4 mesures définies par les masses qu'elles donnent à ces 5 points:

$$P_0 = \{1/4, 3/4, 0, 0, 0\}, \qquad Q_0 = \{0, 3/4, 1/4, 0, 0\},$$
$$P_1 = \{0, 0, 0, 1/10, 9/10\}, \qquad Q_1 = \{0, 0, 0, 9/10, 1/10\}.$$

On pose $P_x = xP_1 + (1-x)P_0$, $Q_x = xQ_1 + (1-x)Q_0$ avec $0 \leqslant x \leqslant 1$ et $\mathscr{P} = \{P_x\}_{x\in[0,1]}$, $\mathscr{Q} = \{Q_x\}_{x\in[0,1]}$. Dans ce cas pour $x, y \in [0, 1]$ on a

$$\Psi(t, P_x, Q_y) = \tfrac{3}{4}(1-y)^t(1-x)^{1-t} + \frac{9^t + 9^{1-t}}{10} y^t x^{1-t}.$$

Il s'ensuit par un calcul facile de dérivées que

$$\bar{\psi}(t) = \log\max\left\{\frac{3}{4}; \frac{9^t+9^{1-t}}{10}\right\}$$

et donc que

$$\bar{\psi}(t) = \begin{cases} \psi(t, P_0, Q_0) & \text{si } t \in [t_0, 1-t_0], \\ \psi(t, P_1, Q_1) & \text{si } t \leqslant t_0 \text{ ou } t \geqslant 1-t_0, \end{cases}$$

où t_0 ($\approx 0{,}1845$) et $1-t_0$ sont les racines de l'équation

$$\frac{3}{4} = \frac{9^t + 9^{1-t}}{10}.$$

Il est alors bien clair que le couple $(\bar{P}, \bar{Q})$ du corollaire 3.1 dépendra à la fois de I et de c.

Par contre, s'il existe un couple $(\bar{P}, \bar{Q})$ tel que $\psi(t, \bar{P}, \bar{Q}) = \bar{\psi}(t)$ pour tout $t \in [0, 1]$, ce couple vérifiera (3.7) indépendamment de I et de c. C'est en particulier ce qui se produit pour les couples „les plus défavorables" construits par Huber et Strassen [10], Huber [9] ou Rieder [16] lorsque $\mathscr{P}$ et $\mathscr{Q}$ sont dominés par des capacités bi-alternatives. Mais la réciproque est fausse, c'est-à-dire que l'égalité $\psi(t, \bar{P}, \bar{Q}) = \bar{\psi}(t)$ pour tout $t \in [0, 1]$ n'entraîne pas nécessairement que le couple $(\bar{P}, \bar{Q})$ est „le plus défavorable", c'est-à-dire que le test de rapport de vraisemblance de $\bar{P}$ contre $\bar{Q}$, de niveau α et erreur $\beta(\alpha)$, est un test de $\mathscr{P}$ contre $\mathscr{Q}$, aux mêmes niveau et erreur, pour toute valeur de α. En effet, ceci entraînerait que pour (P, Q) dans $\mathscr{P} \times \mathscr{Q}$, le test de niveau α de P contre Q a une erreur inférieure à $\beta(\alpha)$. Ceci n'est pas vrai en général comme le montre le contre-exemple suivant dû à Torgersen ([17], p. 233).

On considère sur $\boldsymbol{R}$ les probabilités P_0 et Q_0 portées par trois points A, B, C et définies par $P_0(B) = Q_0(B) = 1-\xi$ et $P_0(A) = \xi = Q_0(C)$, ainsi que $P_1 = \mathscr{N}(0, \sigma^2)$ et $Q_1 = \mathscr{N}(1, \sigma^2)$ deux lois normales de variance σ^2. On pose $P_x = xP_1 + (1-x)P_0$ et $Q_x = xQ_1 + (1-x)Q_0$ avec x dans $[0, 1]$. Alors $\mathscr{P}$ est le segment $\{P_x \mid 0 \leqslant x \leqslant 1\}$ et $\mathscr{Q}$ l'analogue. Il s'ensuit que $\psi(t, P_0, Q_0) \leqslant \psi(t, P_1, Q_1)$, $t \in [0, 1]$, dès que $1-\xi \leqslant e^{-1/8\sigma^2}$ et dans ce cas on a $\bar{\psi}(t) = \psi(t, P_1, Q_1)$. Par ailleurs, Torgersen montre que si

$$1-\xi > 2\Phi\left(-\frac{1}{2\sigma}\right) \quad \text{avec } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du,$$

il existe un niveau α pour lequel la puissance du test de rapport de vraisemblance de P_1 contre Q_1 est plus grande que celle du test de même niveau de P_0 contre Q_0. On obtient le contre-exemple cherché en choisissant

$$e^{-1/8\sigma^2} \geqslant 1-\xi > 2\Phi\left(-\frac{1}{2\sigma}\right),$$

ce qui est possible, puisque $\Phi(-t) < e^{-t^2/2}$ pour $t > 0$.

L'application aux tests du corollaire 3.1 ne donne qu'une majoration de la somme des deux erreurs. Pour majorer chacune d'elles séparément, il est nécessaire d'améliorer l'égalité (3.7) comme suit:

THÉORÈME 3.2. *Soient I un intervalle de $[0, 1]$ et $(\bar{P}, \bar{Q})$ un couple tel que*

$$\delta_I(c) = \inf_{t \in I} [\psi(t, \bar{P}, \bar{Q}) - tc].$$

Alors pour tout $p \geqslant 1$ il existe une fonction positive f_p et un réel t_p, $0 < t_p < 1$, tels que

$$\sup_{P \in \mathscr{P}} \int f_p^{t_p}\, dP \leqslant \exp[\delta_I(c)] + 1/p, \tag{3.8}$$

$$\sup_{Q \in \mathscr{Q}} \int f_p^{t_p - 1}\, dQ \leqslant \exp[\delta_I(c) + c] + 1/p. \tag{3.9}$$

De plus, si $\delta_I(c) = \bar{\psi}(t) - tc$ avec t différent de 0 et 1, on peut choisir $t_p = t$ pour $p \geqslant 1$ et

$$f_p \to e^{-c} \frac{d\bar{Q}}{d\bar{P}} \quad \bar{P} + \bar{Q} \text{ p.s.} \quad \text{lorsque } p \to \infty. \tag{3.10}$$

Preuve. Supposons d'abord que $\delta_I(c) = \bar{\psi}(t) - tc = \delta_{\{t\}}(c)$, où t appartient à $]0, 1[$. D'après (3.7) appliqué avec $I = \{t\}$, pour tout $\varepsilon > 0$ il existe des réels positifs $a_1, \ldots, a_{k(\varepsilon)}$ avec $\sum_{j=1}^{k(\varepsilon)} a_j = 1$ et des fonctions $\bar{g}_j$, $j = 1, \ldots, k(\varepsilon)$, telles que pour tout (P, Q) dans $\mathscr{P} \times \mathscr{Q}$

$$\bar{\psi}(t) - tc + \varepsilon > \log\left[\sum_{j=1}^{k(\varepsilon)} a_j \bar{H}_j(P, Q)\right] \tag{3.11}$$

avec

$$\bar{H}_j(P, Q) = (1-t) \int \bar{g}_j^t dP + te^{-c} \int \bar{g}_j^{t-1} dQ.$$

Si l'on pose $g^t = e^{tc} \sum_{j=1}^{k(\varepsilon)} a_j \bar{g}_j^t$, il est clair par convexité de la fonction $x \to x^{(t-1)/t}$ que

$$g^{t-1} \leqslant \left(\sum_{j=1}^{k(\varepsilon)} a_j \bar{g}_j^{t-1}\right) e^{(t-1)c}, \tag{3.12}$$

d'où finalement pour tous les (P, Q) dans $\mathscr{P} \times \mathscr{Q}$

$$\exp[\bar{\psi}(t)+\varepsilon] > (1-t) \int g^t dP + t \int g^{t-1} dQ. \tag{3.13}$$

Définissons μ, p, q, h, ν par

$$\mu = \bar{P}+\bar{Q}, \quad p = \frac{d\bar{P}}{d\mu}, \quad q = \frac{d\bar{Q}}{d\mu}, \tag{3.14}$$

$$g = h\frac{q}{p} 1_{\{pq>0\}} + g 1_{\{pq=0\}}, \quad \frac{d\nu}{d\mu} = q^t p^{1-t},$$

où h est nulle sur l'ensemble $\{pq = 0\}$. On obtient alors un analogue de (2.8). Si l'on remarque que

$$(1-t)x^t + tx^{t-1} \geqslant 1 + (1-x^t)^2 \frac{1-t}{2t} 1_{\{0<x<1\}} + (1-x^{t-1})^2 \frac{t}{2(1-t)} 1_{\{x>1\}}$$

(en développant en série $(1-u)^{(t-1)/t}$ et $(1-u)^{1/(t-1)}$ pour $u > 0$), il vient, d'après (3.13) et le fait que $\Psi(t, \bar{P}, \bar{Q}) = \nu(\Omega)$,

$$\begin{aligned} \exp[\bar{\psi}(t)+\varepsilon] > \nu(\Omega) &+ \\ + \int_{\{0<h<1\}} (1-h^t)^2 \frac{1-t}{2t} d\nu &+ \int_{\{h>1\}} (1-h^{t-1})^2 \frac{t}{2(1-t)} d\nu + \\ &+ (1-t) \int_{\{pq=0\}} g^t d\bar{P} + t \int_{\{pq=0\}} g^{t-1} d\bar{Q} \geqslant \Psi(t, \bar{P}, \bar{Q}). \end{aligned} \tag{3.15}$$

Soient alors une suite positive $\{\eta_n\}_{n\geqslant 1}$ tendant vers 0 et des fonctions g_n et h_n définies comme précédemment de façon à satisfaire (3.15) avec $\varepsilon = \eta_n$. Comme $\nu(\Omega) = \Psi(t, \bar{P}, \bar{Q}) = \exp[\bar{\psi}(t)]$, les 4 dernières intégrales de (3.15) vont tendre vers 0 quand n tend vers ∞. Donc

$$h_n^t 1_{\{h_n<1\}} \xrightarrow{L^2(\nu)} 1 \quad \text{et} \quad h_n^{t-1} 1_{\{h_n>1\}} \xrightarrow{L^2(\nu)} 1,$$

$$g_n^t 1_{\{pq=0\}} \xrightarrow{L^1(\bar{P})} 0 \quad \text{et} \quad g_n^{t-1} 1_{\{pq=0\}} \xrightarrow{L^1(\bar{Q})} 0.$$

Quitte à extraire des sous-suites on peut donc supposer que

$$h_n \to 1 \ \nu \text{ p.s.,} \quad g_n 1_{\{pq=0\}} \to 0 \ \bar{P} \ p.s. \quad et \quad g_n^{-1} 1_{\{pq=0\}} \to 0 \ \bar{Q} \text{ p.s.,}$$

ce qui entraîne, d'après la définition de ν, que $g_n \to q/p \ \mu$ p.s. D'autre part, on a

$$g_n^t \geqslant h_n^t 1_{\{h_n<1\}}(q/p)^t + (q/p)^t 1_{\{h_n \geqslant 1\}},$$

ce qui par intégration entraîne

$$\int g_n^t d\bar{P} \geqslant \Psi(t, \bar{P}, \bar{Q}) - \int_{\{h_n<1\}} (1-h_n^t) d\nu. \tag{3.16}$$

Si l'on pose $\varepsilon_n = \exp[\bar{\psi}(t)+\eta_n] - \Psi(t, \bar{P}, \bar{Q})$, on obtient, d'après (3.15),

$$\int_{\{h_n<1\}} (1-h_n^t)^2 d\nu \leqslant \frac{2t}{1-t}\varepsilon_n$$

et par l'inégalité de Schwartz

$$\int_{\{h_n<1\}} (1-h_n^t) d\nu \leqslant \sqrt{\frac{2t\varepsilon_n \nu(\Omega)}{1-t}}.$$

Alors (3.16) implique

$$\int g_n^t d\bar{P} \geqslant \Psi(t, \bar{P}, \bar{Q}) - \sqrt{\frac{2t\varepsilon_n \nu(\Omega)}{1-t}}$$

et on obtient de même

$$\int g_n^{t-1} d\bar{Q} \geqslant \Psi(t, \bar{P}, \bar{Q}) - \sqrt{\frac{2(1-t)\varepsilon_n \nu(\Omega)}{t}}.$$

Comme d'après (3.13) pour tout couple (P, Q)

$$\Psi(t, \bar{P}, \bar{Q}) + \varepsilon_n > (1-t)\int g_n^t dP + t\int g_n^{t-1} dQ,$$

il s'ensuit que, pour tout (P, Q) dans $\mathscr{P} \times \mathscr{Q}$,

$$\int g_n^t dP < \frac{\varepsilon_n}{1-t} + \sqrt{\frac{2t\varepsilon_n \nu(\Omega)}{1-t}} + \Psi(t, \bar{P}, \bar{Q}), \tag{3.17}$$

$$\int g_n^{t-1} dQ < \frac{\varepsilon_n}{t} + \sqrt{\frac{2(1-t)\varepsilon_n \nu(\Omega)}{t}} + \Psi(t, \bar{P}, \bar{Q}). \tag{3.18}$$

Comme $\delta_t(c) = \psi(t, \bar{P}, \bar{Q}) - tc$ et $\varepsilon_n \to 0$ lorsque $n \to \infty$, on obtiendra (3.8) et (3.9) en extrayant une sous-suite $\{f_p\}_{p\geqslant 1}$ de $\{e^{-c}g_n\}_{n\geqslant 1}$. Comme $g_n \to q/p \ \bar{P}+\bar{Q}$ p.s., (3.10) est immédiat.

Finalement, si $t = 0$ ou 1, il suffit de faire le même raisonnement à partir de $\bar{\psi}(t_p)$, où $\bar{\psi}(t_p)$ est suffisamment proche de $\bar{\psi}(t)$, pour obtenir (3.8) et (3.9) puisque

$$\delta_I(c) = \lim_{t_p \to t} [\bar{\psi}(t_p) - t_p c].$$

COROLLAIRE 3.2. *Sous la condition* (AC), *si* $\delta_I(c) = \bar{\psi}(t) - tc$ *avec* t *dans* $]0, 1[$ *et* $(\bar{P}, \bar{Q})$ *satisfait à* (3.7), *alors*

$$\sup_{P \in \mathscr{P}} e^{-tc} \int (d\bar{Q}/d\bar{P})^t \, dP \leqslant \exp[\delta_I(c)], \tag{3.19}$$

$$\sup_{Q \in \mathscr{Q}} e^{-tc} \int (d\bar{Q}/d\bar{P})^{t-1} \, dQ \leqslant \exp[\delta_I(c)], \tag{3.20}$$

où $d\bar{Q}/d\bar{P} = q/p$.

Preuve. Supposons qu'on ait (3.8), (3.9) et (3.10); comme P est absolument continu par rapport à μ, le lemme de Fatou implique que

$$\int e^{-tc} (d\bar{Q}/d\bar{P})^t \, dP \leqslant \varliminf_{p} \int f_p^t \, dP,$$

où f_p est donné par le théorème 3.2 et satisfait (3.10), puisque les fonctions f_p sont toutes positives. (3.19) s'ensuit et (3.20) de même après simplification par e^c grâce à (3.8) et (3.9).

4. Application aux tests. Nous allons considérer tout particulièrement le cas où $I = [0, 1]$ et nous noterons pour simplifier

$$\delta(c) = \delta_{[0,1]}(c) \quad \text{et} \quad \gamma(c) = \delta(c) + c. \tag{4.1}$$

PROPOSITION 4.1. *Soient* $\varepsilon > 0$, n *et* c *fixés. Il existe une région de rejet* $W_n \subset \Omega^n$ *telle que le test de* $\mathscr{P}^n$ *contre* $\mathscr{Q}^n$ *défini par* $\varphi_n = 1_{W_n}$ *ait des erreurs* α_n *et* β_n *majorées respectivement par* $(1+\varepsilon) \exp[n\delta(c)]$ *et* $(1+\varepsilon) \exp[n\gamma(c)]$, *c'est-à-dire*

$$\sup_{\substack{P_i \in \mathscr{P} \\ 1 \leqslant i \leqslant n}} P_1 \otimes \ldots \otimes P_n [W_n] \leqslant (1+\varepsilon) \exp[n\delta(c)], \tag{4.2}$$

$$\sup_{\substack{Q_i \in \mathscr{Q} \\ 1 \leqslant i \leqslant n}} Q_1 \otimes \ldots \otimes Q_n [W_n^c] \leqslant (1+\varepsilon) \exp[n\gamma(c)]. \tag{4.3}$$

Preuve. Soit p tel que

$$[\exp[\delta(c)] + 1/p]^n \leqslant (1+\varepsilon) \exp[n\delta(c)], \quad [\exp[\gamma(c)] + 1/p]^n \leqslant (1+\varepsilon) \exp[n\gamma(c)]. \tag{4.4}$$

Considérons alors f_p qui satisfait à (3.8) et (3.9) avec $I = [0, 1]$ et posons

$$W_n = \{x_i, \ 1 \leqslant i \leqslant n \mid \sum_{i=1}^{n} \log f_p(x_i) > 0\}.$$

Alors d'après l'inégalité exponentielle, puisque $t_p > 0$,

$$P_1 \otimes \ldots \otimes P_n[W_n] \leqslant \int \exp\Big[t_p\big[\sum_{i=1}^{n} \log f_p(x_i)\big]\Big]\, dP_1(x_1)\ldots dP_n(x_n)$$

$$\leqslant \prod_{i=1}^{n} \int f_p^{t_p}(x_i)\, dP(x_i) \leqslant \big[\exp[\delta(c)] + 1/p\big]^n$$

et l'on obtient (4.2) grâce à (4.4). Le calcul concernant l'autre erreur est identique et donne (4.3).

De façon analogue, à partir du corollaire 3.2, on obtient

COROLLAIRE 4.1. *Sous la condition* (AC), *si* $\delta(c) = \bar{\psi}(t) - tc$ *avec* t *dans* $]0, 1[$ *et* $(\bar{P}, \bar{Q})$ satisfait *à* (3.7), *on peut obtenir* (4.2) *et* (4.3) *avec* $\varepsilon = 0$ *si l'on choisit*

$$W_n = \{x_i,\ 1 \leqslant i \leqslant n \mid 1/n \sum_{i=1}^{n} \log d\bar{Q}/d\bar{P}(x_i) > c\}.$$

Nous pouvons aussi retrouver des résultats analogues à ceux de Le Cam [11]. Choisissons $I = \{\frac{1}{2}\}$ et $c = 0$. Alors $\Psi(\frac{1}{2}, P, Q) = \int \sqrt{dPdQ}$ est dite *affinité de Hellinger* et notée $\varrho(P, Q)$ et

$$\varrho(P, Q) = 1 - h^2(P, Q).$$

Donnons-nous alors des couples $(\mathscr{P}_i, \mathscr{Q}_i)$, $1 \leqslant i \leqslant n$, vérifiant les hypothèses (T) et non nécessairement égaux et posons

$$\varrho_i = \varrho(\mathscr{P}_i, \mathscr{Q}_i) = \sup_{\substack{P \in \mathscr{P}_i \\ Q \in \mathscr{Q}_i}} \varrho(P, Q) = \varrho(\bar{P}_i, \bar{Q}_i),$$

l'existence de $\bar{P}_i$, $\bar{Q}_i$ étant assurée par la proposition 3.1. On obtient alors le résultat suivant dont la démonstration est identique à celle de la proposition 4.1:

PROPOSITION 4.2. *Pour tout* $\varepsilon > 0$, *il existe un test de région de rejet* W_n *tel que pour tout* P_i *dans* $\mathscr{P}_i$ *et* Q_i *dans* $\mathscr{Q}_i$, $1 \leqslant i \leqslant n$, *on ait*

$$(4.5)\quad P_1 \otimes \ldots \otimes P_n[W_n] \leqslant (1+\varepsilon) \prod_{i=1}^{n} \varrho_i, \qquad Q_1 \otimes \ldots \otimes Q_n[W_n^c] \leqslant (1+\varepsilon) \prod_{i=1}^{n} \varrho_i.$$

De plus, si tous les couples $(\mathscr{P}_i, \mathscr{Q}_i)$ *vérifient la condition* (AC), *en posant*

$$W_n = \{\sum_{i=1}^{n} \log d\bar{Q}_i/d\bar{P}_i > 0\},$$

on a $\varepsilon = 0$ *dans* (4.5).

Remarque. La proposition 4.2 est l'analogue de 4.1 avec $I = \{\frac{1}{2}\}$ et $c = 0$. Il est clair que ce choix est arbitraire et destiné à l'application aux affinités de Hellinger; on pourrait obtenir des résultats identiques avec $I = \{t\}$, $t \in]0, 1[$, et $c \in \boldsymbol{R}$.

Il est clair que les majorations des erreurs obtenues dans les précédentes propositions n'ont aucun caractère d'optimalité, ce sont simplement des bornes exponentielles utilisables pratiquement. Cependant asymptotiquement, les coefficients des exponentielles qui interviennent dans (4.2) et (4.3) sont les meilleurs possibles. Pour le montrer, nous rappellerons tout d'abord quelques résultats concernant les erreurs asymptotiques des tests et qui sont contenus dans [4], [8] et [19].

Soit $\Phi = \{\varphi_n\}_{n \geq 1}$ une suite de tests de $\mathscr{P}$ contre $\mathscr{Q}$ fondés sur n observations. Notons $\alpha(\Phi)$ le log-niveau de la suite Φ et $\beta(\Phi)$ sa log-erreur définis par

$$\alpha(\Phi) = \overline{\lim}\, n^{-1} \log \alpha_n \quad \text{et} \quad \beta(\Phi) = \overline{\lim}\, n^{-1} \log \beta_n,$$

α_n et β_n étant données par (1.1). Soit $A > 0$ fixé; on définit

$$B(A, \mathscr{P}, \mathscr{Q}) = \sup_{\{\Phi \mid \alpha(\Phi) \leqslant -A\}} \{-\beta(\Phi)\}.$$

Lorsque $\mathscr{P}$ et $\mathscr{Q}$ sont réduits chacun à un point, on a le résultat suivant (cf. [4]):

PROPOSITION 4.3. *$B(A, P, Q) = +\infty$ si $A \leqslant -\psi(0, P, Q)$. Pour $A > -\psi(0, P, Q)$, $B(A, P, Q)$ est une fonction continue décroissante de A telle qu'il existe un unique c pour lequel*

$$-A = \inf_{t \in [0,1]} [\psi(t, P, Q) - tc], \qquad B(A, P, Q) = A - c.$$

Si Φ est une suite de tests telle que $\alpha(\Phi) = -A$ et $\beta(\Phi) = B(A, \mathscr{P}, \mathscr{Q})$, elle sera dite *logarithmiquement asymptotiquement optimale* (L.A.O.).

THÉORÈME 4.1. (i) *Si $A > -\bar{\psi}(0)$, il existe c tel que*

$$\delta(c) = -A, \quad \gamma(c) = -B(A, \mathscr{P}, \mathscr{Q}) \tag{4.6}$$

et la suite de tests donnés par la proposition 4.1 *ou le corollaire* 4.1 *avec le même c est L.A.O.*

(ii) *Si $\bar{\psi}(0) < 0$, $B(-\bar{\psi}(0), \mathscr{P}, \mathscr{Q}) = +\infty$ et pour toute suite c_n tendant vers $-\infty$ on a*

$$\delta(c_n) \to \bar{\psi}(0) \text{ lorsque } n \to \infty, \qquad \gamma(c_n) \to -\infty \text{ lorsque } n \to \infty,$$

la suite de tests donnés par la proposition 4.1 *ou le corollaire* 4.1 *avec $c = c_n$ étant L.A.O.*

Preuve. D'après (3.1) et la convexité de $\bar{\psi}$, la fonction $\delta(c)$ est décroissante et continue, avec

$$\lim_{c \to -\infty} \delta(c) = \bar{\psi}(0) \quad \text{et} \quad \lim_{c \to +\infty} \delta(c) = -\infty.$$

Donc si $A > -\bar{\psi}(0)$, il existe c tel que $\delta(c) = -A$ et $\gamma(c) = -A+c$. Soient $\bar{P}$, $\bar{Q}$ satisfaisant à (3.4), d'après la proposition 4.3 on aura $B(A, \bar{P}, \bar{Q}) = A-c = -\gamma(c)$. Comme il est clair que $B(A, \mathscr{P}, \mathscr{Q}) \leqslant B(A, \bar{P}, \bar{Q})$, on a $B(A, \mathscr{P}, \mathscr{Q}) \leqslant -\gamma(c)$. Or la suite des tests donnés par la proposition 4.1 a un log-niveau $\delta(c) = -A$ et une log-erreur $\gamma(c)$, ce qui entraîne $B(A, \mathscr{P}, \mathscr{Q}) \geqslant -\gamma(c)$ et donc (4.6).

Soit $\{c_n\}_{n\geqslant 1}$ une suite tendant vers $-\infty$. Alors $\delta(c_n) \to \bar{\psi}(0)$, $\gamma(c_n) \to -\infty$ et les tests donnés par la proposition 4.1 ont un log-niveau $\bar{\psi}(0)$ et une log-erreur $-\infty$, ce qui entraîne que $B(-\bar{\psi}(0), \mathscr{P}, \mathscr{Q}) = +\infty$ et achève la démonstration.

Remarque. D'après ce qui précède, on a

$$B(A, \mathscr{P}, \mathscr{Q}) = A-c \quad \text{si } -A = \inf_{t\in[0,1]} [\bar{\psi}(t)-tc] < \bar{\psi}(0),$$

c'est-à-dire que la fonction $B(A, \mathscr{P}, \mathscr{Q})$ est entièrement déterminée par $\bar{\psi}$. Il s'ensuit que toutes les majorations d'erreurs précédemment obtenues subsistent si l'on remplace $\mathscr{P}$ et $\mathscr{Q}$ par des ensembles plus gros $\mathscr{P}'$ et $\mathscr{Q}'$ pourvu que $\bar{\psi}$ reste inchangé. Cela sera souvent utile lorsque $\mathscr{P}$ et $\mathscr{Q}$ ne sont pas compacts; on pourra utiliser des compactifications $\mathscr{P}'$ et $\mathscr{Q}'$, par exemple en grossissant l'espace support Ω comme on le verra dans le cas des boules pour la distance de Hellinger.

5. Tests entre des boules en distance de Hellinger. Nous allons montrer ici comment la théorie précédente s'applique à deux boules pour la distance de Hellinger définie par (1.2): $\mathscr{P} = \mathscr{B}(P_0, \varepsilon)$ et $\mathscr{Q} = \mathscr{B}(Q_0, \eta)$, où $\mathscr{B}(P, r)$ désigne la boule fermée de centre P et de rayon r. On supposera que $\varepsilon+\eta < h(P_0, Q_0)$ de façon que ces boules soient disjointes. Comme $h^2(P, Q) = 1 - \Psi(\frac{1}{2}, P, Q)$, le lemme 2.2 entraîne qu'elles sont aussi convexes.

Si l'on suppose que Ω est un espace métrique localement compact, F l'ensemble des fonctions continues, positives log-bornées, et $\mathscr{T}$ la topologie de la convergence étroite, (T2) et (T3) seront vérifiées; de plus, la fonction $\Psi(\frac{1}{2}, P, Q)$ du couple (P, Q) est s.c.s., donc $h(P, Q)$ est s.c.i. et les boules de Hellinger sont fermées, mais généralement non compactes.

Si l'espace Ω est compact, la topologie $\mathscr{T}$ est compacte et les boules $\mathscr{P}$ et $\mathscr{Q}$ également. Sinon, il suffit de considérer le compactifié d'Alexandroff $\bar{\Omega} = \Omega \cup \Delta$. Les boules $\bar{\mathscr{P}}$ et $\bar{\mathscr{Q}}$ ont les mêmes centres et les mêmes rayons que $\mathscr{P}$ et $\mathscr{Q}$, mais pour les probabilités portées par $\bar{\Omega}$. Les boules $\bar{\mathscr{P}}$ et $\bar{\mathscr{Q}}$ vérifient alors les hypothèses (T) et on peut leur appliquer les résultats précédents. La proposition suivante montre qu'on n'a rien perdu en remplaçant $\mathscr{P}$ et $\mathscr{Q}$ par $\bar{\mathscr{P}}$ et $\bar{\mathscr{Q}}$ en ce sens qu'on n'a pas changé la fonction $\bar{\psi}(t)$. Ce résultat est d'ailleurs vrai si Δ est un ensemble quelconque, ce qui autorise des compactifications autres, pourvu que (T) soit vérifiée.

Proposition 5.1. *Soient $\bar{P}$ et $\bar{Q}$ dans $\bar{\mathscr{P}}$ et $\bar{\mathscr{Q}}$ respectivement; alors on peut trouver P dans $\mathscr{P}$ et Q dans $\mathscr{Q}$ avec*

$$\varrho(P, P_0) \geqslant \varrho(\bar{P}, P_0), \qquad \varrho(Q, Q_0) \geqslant \varrho(\bar{Q}, Q_0),$$

$$\int_\Omega d\bar{P}^t d\bar{Q}^{1-t} \leqslant \int_\Omega dP^t dQ^{1-t} \qquad \textit{pour } 0 \leqslant t \leqslant 1.$$

Preuve. Soient $a = \bar{P}(\Delta)$ et $P = 1_\Omega \bar{P} + R$ avec $R(\Omega) = a$ et $R(\Delta) = 0$. On a alors

$$\varrho(P, P_0) = \int_\Omega \sqrt{d\bar{P}+dR}\, dP_0 \geqslant \int_\Omega \sqrt{d\bar{P}}\, dP_0 = \varrho(\bar{P}, P_0)$$

puisque P_0 est porté par Ω. On a de la même façon $\varrho(Q, Q_0) \geqslant \varrho(\bar{Q}, Q_0)$ si $Q = 1_\Omega \bar{Q} + S$ avec $S(\Omega) = \bar{Q}(\Delta) = b$ et $S(\Delta) = 0$. On aura alors, d'après (2.10) avec $u = \frac{1}{2}$,

$$\int_\Omega dP^t dQ^{1-t} = \int_\Omega (d\bar{P}+dR)^t (d\bar{Q}+dS)^{1-t} \geqslant \int_\Omega d\bar{P}^t d\bar{Q}^{1-t} + \int_\Omega dR^t dS^{1-t}.$$

On peut supposer que a, b ne sont pas tous deux nuls, sinon le problème est résolu; si par exemple $a \neq 0$, on choisira $S = ba^{-1} R$, et on a finalement

$$\int_\Omega dP^t dQ^{1-t} \geqslant \int_\Omega d\bar{P}^t d\bar{Q}^{1-t} + a^t b^{1-t} \geqslant \int_\Omega d\bar{P}^t d\bar{Q}^{1-t}$$

puisque

$$\int_\Delta d\bar{P}^t d\bar{Q}^{1-t} \leqslant [\bar{P}(\Delta)]^t [\bar{Q}(\Delta)]^{1-t}.$$

Cependant, pour $t = \frac{1}{2}$, les résultats du chapitre 3 prennent un aspect particulièrement simple parce qu'il est très facile de construire des couples $(\bar{P}, \bar{Q})$ tels que $\psi(\frac{1}{2}, \bar{P}, \bar{Q}) = \bar{\psi}(\frac{1}{2})$ comme le montre le lemme suivant, pour lequel on supposera que $\mathscr{P}$ et $\mathscr{Q}$ ne sont pas toutes deux réduites à un point, le problème étant alors tout résolu:

Lemme 5.1. *Si $\varepsilon + \eta > 0$, il existe un et un seul couple $(\bar{P}, \bar{Q})$, $\bar{P} \in \mathscr{P}$, $\bar{Q} \in \mathscr{Q}$, tel que*

$$\varrho(\bar{P}, \bar{Q}) = \sup_{\substack{P \in \mathscr{P} \\ Q \in \mathscr{Q}}} \varrho(P, Q) < 1 - [h(P_0, Q_0) - \varepsilon - \eta]^2, \tag{5.1}$$

et si $P_0 = p_0\, d\mu$, $Q_0 = q_0\, d\mu$, $\bar{P}$ et $\bar{Q}$ sont absolument continus par rapport à μ de densités $\bar{p}$, $\bar{q}$ avec

$$\sqrt{\bar{p}} = \alpha \sqrt{p_0} + \alpha' \sqrt{q_0} \quad \textit{et} \quad \sqrt{\bar{q}} = \beta' \sqrt{p_0} + \beta \sqrt{q_0}, \tag{5.2}$$

α, α', β, β' étant tels que $\varrho(P_0, \bar{P}) = 1 - \varepsilon^2$ et $\varrho(Q_0, \bar{Q}) = 1 - \eta^2$.

Preuve. On va aisément ramener le problème à un problème euclidien en petite dimension. Soient R dans $\mathscr{P}$, S dans $\mathscr{Q}$. On peut supposer que R

$= r\,d\mu$ et $S = s\,d\mu$. Il est clair alors que $2h$ est la distance $\mathbb{L}^2$ entre les racines des densités par rapport à μ et ϱ le produit scalaire associé. Donc il existe f et g, p' et q' dans $\mathbb{L}^2(\mu)$, de norme 1 (mais non nécessairement $\geqslant 0$), tels que f et g sont orthogonaux à $\sqrt{p_0}$ et $\sqrt{q_0}$, p' et q' sont dans l'espace engendré par $\sqrt{p_0}$ et $\sqrt{q_0}$ et

$$\sqrt{r} = \sqrt{(1-\alpha^2)}\,p' + \alpha f, \qquad \sqrt{s} = \sqrt{(1-\beta^2)}\,q' + \beta g.$$

On a alors

$$\langle \sqrt{r}, \sqrt{s} \rangle = \sqrt{(1-\alpha^2)(1-\beta^2)}\,\langle p' \cdot q' \rangle + \alpha\beta \langle f \cdot g \rangle.$$

Il est clair que pour maximiser $\langle \sqrt{r}, \sqrt{s} \rangle$, on peut supposer que $f = \pm g$. Mais on est alors ramené à un problème tri-dimensionnel: trouver le minimum de la distance entre deux calottes sphériques. Ce minimum est bien sûr atteint par un couple unique qui se trouve dans le plan engendré par $\sqrt{p_0}$ et $\sqrt{q_0}$, d'où (5.2). De plus, on a

$$\|\sqrt{p_0} - \sqrt{\bar{p}}\|_2^2 = 4\varepsilon^2, \qquad \|\sqrt{q_0} - \sqrt{\bar{q}}\|_2^2 = 4\eta^2, \qquad \|\sqrt{p_0} - \sqrt{q_0}\|_2^2 = 4h^2(P_0, Q_0);$$

comme ces 4 points sont sur un même cercle et non pas alignés,

$$\|\sqrt{\bar{p}} - \sqrt{\bar{q}}\|_2 > 2[h(P_0, Q_0) - \varepsilon - \eta],$$

ce qui entraîne (5.1).

Il est facile de voir que $\alpha' > 0$ si $\varepsilon > 0$ et $\beta > 0$ si $\eta > 0$, ce qui implique que si $\varepsilon > 0$, $\sqrt{\bar{q}/\bar{p}}$ est borné supérieurement, et que si $\eta > 0$, $\sqrt{\bar{q}/\bar{p}}$ admet un minorant > 0. Dans tous les cas, $P_0 + Q_0$ et $\bar{P} + \bar{Q}$ sont équivalentes.

Nous pouvons alors améliorer le théorème 3.2 de la manière suivante: ayant choisi $\mu = P_0 + Q_0$, nous noterons Λ le support de μ. Considérons des versions $\bar{p}$ et $\bar{q}$ des densités $d\bar{P}/d\mu$ et $d\bar{Q}/d\mu$ telles que si nous posons

(5.3) $\qquad f = \bar{q}/\bar{p}\,1_{\{\Lambda^c\}} + 1_{\{\Lambda\}}$ (avec la convention $0/0 = 1$),

l'inf de f coïncide avec son inf essentiel et de même le sup avec le sup essentiel. Si $\varepsilon\eta > 0$, f est une fonction log-bornée et $\bar{P}$, $\bar{Q}$ sont mutuellement absolument continues. Nous allons démontrer

THÉORÈME 5.1. *Pour tout P dans $\mathscr{P}$ et tout Q dans $\mathscr{Q}$, on a*

(5.4) $\qquad \int \sqrt{f}\,dP \leqslant \varrho(\bar{P}, \bar{Q})$ *et* $\int \sqrt{f^{-1}}\,dQ \leqslant \varrho(\bar{P}, \bar{Q})$.

Preuve. La démonstration du théorème 3.2 avec $I = \{\frac{1}{2}\}$ et $c = 0$ montre l'existence d'une suite $\{f_n\}$ de fonctions $\geqslant 0$ telles que $f_n \to \bar{q}/\bar{p}$ $\bar{P} + \bar{Q}$ p.s. lorsque $n \to \infty$ et (d'après (3.17) et (3.18))

$$\overline{\lim} \int \sqrt{f_n}\,dP \leqslant \varrho(\bar{P}, \bar{Q}), \qquad \overline{\lim} \int \sqrt{f_n^{-1}}\,dQ \leqslant \varrho(\bar{P}, \bar{Q})$$

pour tout P dans $\mathscr{P}$ et tout Q dans $\mathscr{Q}$. Comme $\bar{P} + \bar{Q}$ et μ sont équivalentes, il s'ensuit par le lemme de Fatou que (5.4) est satisfaite si P et Q sont absolument continues par rapport à μ. Supposons que cette condition ne soit pas réalisée, par exemple par P, le cas de Q se traitant de manière analogue. Ceci implique que P est différent de P_0, donc $\varepsilon > 0$ et $\bar{q}/\bar{p}$ est bornée supérieurement. Puisque P a une partie étrangère à μ, on peut écrire $P = h\mu + gv$ (v étrangère à μ). Nous poserons

$$M = \text{ess sup}\,\frac{\bar{q}}{\bar{p}} = \sup\frac{\bar{q}}{\bar{p}}.$$

Il s'ensuit que $f \leqslant M$. Soit $\delta > 0$; alors $\bar{Q}\{f \geqslant M-\delta\} = \delta' > 0$. Supposons que $\int g\,dv = \gamma$ et posons

$$P' = h\mu + \frac{\gamma}{\delta'} 1_{\{f \geqslant M-\delta\}} \bar{Q}.$$

P' est alors une probabilité telle que

$$\varrho(P_0, P') \geqslant \int \sqrt{hp_0}\,d\mu = \varrho(P_0, P).$$

Donc P' est un point de $\mathscr{P}$, ayant une densité par rapport à μ. Il s'ensuit que

$$\varrho(\bar{P}, \bar{Q}) \geqslant \int \sqrt{f}\,dP' \geqslant \int \sqrt{f}\,h\,d\mu + \gamma/\delta' \sqrt{M-\delta}\,\delta',$$

soit $\varrho(\bar{P}, \bar{Q}) \geqslant \gamma\sqrt{M-\delta} + \int \sqrt{f}\,h\,d\mu$. D'autre part,

$$\int \sqrt{f}\,dP \leqslant \int \sqrt{f}\,h\,d\mu + \sqrt{M}\int g\,dv = \int \sqrt{f}\,h\,d\mu + \gamma\sqrt{M}.$$

Comme δ est arbitraire, on a finalement

$$\int \sqrt{f}\,dP \leqslant \varrho(\bar{P}, \bar{Q}) \qquad \text{pour tout } P \in \mathscr{P}.$$

On en déduit, par une démonstration semblable à celle de la proposition 4.1,

Corollaire 5.1. *Soit le test de $\mathscr{P}$ contre $\mathscr{Q}$ fondé sur n observations $x_1, \ldots, x_n$ qui rejette $\mathscr{P}$ lorsque*

$$\sum_{i=1}^{n} \log f(x_i) > 0,$$

f étant défini par (5.3). *Les deux erreurs de ce test sont majorées par*

$$\varrho^n(\bar{P}, \bar{Q}) < \exp[-nh^2(\mathscr{P}, \mathscr{Q})].$$

Ce corollaire permet d'obtenir très aisément des tests de $\mathscr{P}$ contre $\mathscr{Q}$, qui pourront être utilisés pour construire certains estimateurs comme dans [12] et [5]. Pour mettre en oeuvre ces tests, il suffit de calculer $\bar{p}$ et $\bar{q}$, lesquels sont donnés par (5.2). Il s'ensuit aisément que pour que $\bar{P}$ et $\bar{Q}$ soient des probabilités, il faut et il suffit que α, α', β, β' soient les racines positives des équations

$$\alpha^2+\alpha'^2+2\alpha\alpha'(1-a)=1, \quad \beta^2+\beta'^2+2\beta\beta'(1-a)=1,$$
$$\alpha+\alpha'(1-a)=1-\varepsilon^2, \quad \beta+\beta'(1-a)=1-\eta^2,$$

où $a=h^2(P_0, Q_0)$.

En particulier, lorsque $\varepsilon=\eta$, on obtient en convenant de poser $p_0/q_0=1$ lorsque $p_0=q_0=0$ ainsi que sur A^c:

$$\sqrt{f}=(\alpha\sqrt{q_0/p_0}+\alpha')(\alpha'\sqrt{q_0/p_0}+\alpha)^{-1}.$$

Si la théorie précédente s'applique très bien, comme on vient de le voir, aux boules en distance de Hellinger sur un espace métrique localement compact, il serait plus intéressant d'obtenir des test optimaux pour n fixé en cherchant un couple „le plus défavorable" comme dans [10] ou [9]. Malheureusement, si les boules de Hellinger engendrent des capacités bi-alternatives, elles ne sont pas dominées par ces capacités comme on va le voir, ce qui fait qu'il est impossible de leur appliquer les résultats de [10].

Soit $\mathscr{B}$ la boule fermée de centre P et de rayon $r>0$. Définissons la fonction d'ensemble v par

$$v(A)=\sup_{Q\in\mathscr{B}} Q(A) \quad \text{pour tout } A\in\mathscr{A}. \tag{5.5}$$

On rappelle que v est bi-alternative si et seulement si

$$v(A)+v(B)\geqslant v(A\cup B)+v(A\cap B).$$

Nous démontrerons d'abord un lemme, qui nous a été indiqué par Patrice Assouad:

LEMME 5.2. *Soit la fonction d'ensemble w définie par $w(A)=g[P(A)]$, où P est une probabilité et g une fonction concave. Alors w est bi-alternative.*

Preuve. Il suffit de voir que pour $t\geqslant 0$ et $0\leqslant a\leqslant b$ on a

$$g(a)+g(b)\geqslant g(a-t)+g(b+t),$$

ce qui est une conséquence de la concavité de g.

PROPOSITION 5.2. *Supposons que $(\Omega, \mathscr{A})$ est un espace métrique compact muni de sa tribu borélienne. Alors la fonction d'ensemble définie par (5.5) est une capacité bi-alternative.*

Preuve. $\mathscr{B}$ est convexe et fermée pour la topologie de la convergence étroite, donc compacte puisque Ω l'est, ce qui d'après [8] entraîne que v est une capacité. Pour vérifier qu'elle est bi-alternative, nous utiliserons le lemme 5.2. Soit Q dans $\mathscr{B}$ telle que $1\geqslant Q(A)\geqslant P(A)$. Pour déterminer $v(A)$, il suffit de trouver le maximum de $Q(A)$ lorsque Q parcourt $\mathscr{B}$. Pour cela nous chercherons d'abord le maximum de $\varrho(P, Q)$ en fonction de $Q(A)$. Si μ domine $P+Q$ et $p = dP/d\mu$, $q = dQ/d\mu$, on peut écrire

$$\varrho(P, Q) = \int_A \sqrt{p/q}\, dQ + \int_{A^c} \sqrt{q/p}\, dP.$$

D'après l'inégalité de Jensen

$$1/Q(A) \int_A \sqrt{p/q}\, dQ \leqslant 1/\sqrt{Q(A)}\sqrt{\int_A p/q\, dQ} \leqslant \sqrt{P(A)/Q(A)},$$

l'égalité des termes extrêmes étant réalisée si et seulement si p/q est constant Q p.s. sur A. Il s'ensuit que pour $P(A)$ et $Q(A)$ fixés, le maximum de $\int_A \sqrt{p/q}\, dQ$ est $\sqrt{P(A)Q(A)}$. En raisonnant de même avec A^c, on obtient la valeur maximale de $\varrho(P, Q)$ en fonction de $Q(A)$ donnée par

$$\sqrt{P(A)Q(A)} + \sqrt{P(A^c)Q(A^c)}, \qquad P(A) \leqslant Q(A) \leqslant 1.$$

Posons $k(s, u) = \sqrt{su} + \sqrt{(1-s)(1-u)}$ pour $0 \leqslant s \leqslant u \leqslant 1$. La fonction $k(s, u)$ est décroissante en u sur $[s, 1]$ de 1 à $\sqrt{s}$. Le maximum de $\varrho(P, Q)$ en fonction de $P(A)$ et $Q(A)$ étant $k(P(A), Q(A))$, il s'ensuit que le maximum de $Q(A)$ sous les contraintes $P(A) = s$ et $\varrho(P, Q) \geqslant 1-r^2$ (soit $h(P, Q) \leqslant r$) est donné par la solution $g(s) \geqslant s$ de $k(s, g(s)) = 1-r^2$ qui existe si $1-r^2 \geqslant \sqrt{s}$. Sinon, il suffit de prendre $Q(A) = 1$ puisque $k(s, 1) = \sqrt{s} > 1-r^2$. Donc $v(A)$ est de la forme $g(P(A))$, où g est définie par

$$\begin{array}{lll} k(s, g(s)) = 1-r^2, & s \leqslant g(s) \leqslant 1 & \text{si } \sqrt{s} \leqslant 1-r^2, \\ g(s) = 1 & & \text{si } 1-r^2 \leqslant \sqrt{s} \leqslant 1. \end{array}$$

On vérifie que g est croissante et concave sur $[0, 1]$, ce qui, grâce au lemme 5.2, achève la démonstration.

A v on peut associer l'ensemble $\mathscr{V}$ dominé par la capacité, c'est-à-dire:

$$\mathscr{V} = \{Q \mid Q(A) \leqslant v(A) \quad \text{pour tout } A \in \mathscr{A}\}.$$

Il est bien clair que $\mathscr{B} \subset \mathscr{V}$. Si on avait égalité, on pourrait utiliser les résultats de [10], mais on va voir que non seulement $\mathscr{V}$ n'est pas inclus dans $\mathscr{B}$, mais qu'il n'est pas possible d'inclure $\mathscr{V}$ dans une boule de centre P et de rayon λr pour un λ indépendant de r.

On prendra $\Omega = [0, 1]$, P la mesure de Lebesgue, δ_0 la masse de Dirac en 0 et Q définie par

$$Q = \frac{r^2}{2}\delta_0 + \left(1 + \frac{r}{\sqrt{2x}}\right) 1_{[0,1/2]} \cdot P + (1-2r-r^2)\, 1_{[1/2,1]} \cdot P,$$

ce qui fait que l'on a

$$Q[0, s] = \begin{cases} (\sqrt{s}+r/\sqrt{2})^2 & \text{si } 0 \leqslant s \leqslant \frac{1}{2}, \\ s+(2r+r^2)(1-s) & \text{si } s \geqslant \frac{1}{2}. \end{cases}$$

Proposition 5.3. *Q est un élément de $\mathscr{V}$ si $r^2+r \leqslant \frac{1}{2}$ ($r \geqslant 0$, $r^2+2r \leqslant 1$).*

Preuve. Il suffit de voir que l'on a $Q(A) \leqslant v(A)$ pour tout borélien de $[0, 1]$ et donc que $Q(A) \leqslant g[P(A)]$. Mais dQ/dP étant décroissante sur $[0, 1]$, on a nécessairement $Q(A) \leqslant Q[0, P(A)]$, et il suffit de vérifier

$$Q[0, s] \leqslant g(s), \quad 0 \leqslant s \leqslant 1. \tag{5.6}$$

Si $\sqrt{s} \geqslant 1-r^2$, $g(s) = 1$ et (5.6) est vérifié. Si $\sqrt{s} \leqslant (1-r^2)$, $k(s, g(s)) = 1-r^2$. Comme $k(s, y)$ est une fonction décroissante de y pour $y \in [s, 1]$ et que $Q[0, s] \geqslant s$, il suffit de vérifier que

$$k(s, Q[0, s]) \geqslant 1-r^2 \quad \text{si } \sqrt{s} \leqslant (1-r^2).$$

Si $s \leqslant \frac{1}{2}$, on se ramène à vérifier en posant $\sqrt{s} = u$

$$\sqrt{(1-u^2)[1-(u+r/\sqrt{2})^2]} \geqslant 1-r^2-u(u+r/\sqrt{2}) \tag{5.7}$$

$$\text{avec } u \leqslant 1/\sqrt{2} \text{ et } u \leqslant 1-r^2.$$

On peut montrer que ceci équivaut si le second membre de (5.7) est $\geqslant 0$ à $r^2-3/2+2u^2+ur\sqrt{2} \leqslant 0$. Ces conditions sont remplies si $s \leqslant \frac{1}{2}$ et $r^2+r \leqslant \frac{1}{2}$. Lorsque $s \geqslant \frac{1}{2}$, on a $Q[0, s] = R[0, s]$, où R est défini par

$$dR/dP = (1+r)^2\, 1_{[0,1/2]} + (1-2r-r^2)\, 1_{[1/2,1]}.$$

Or, on a $\varrho(R, P) = \frac{1}{2}[\sqrt{(1+r)^2}+\sqrt{1-2r+r^2}] \geqslant 1-r^2$ dès que $r^2+r \leqslant \frac{1}{2}$. Donc R est dans $\mathscr{B}$ et $g(s) \geqslant R[0, s]$, ce qui achève la démonstration.

Proposition 5.4. *Q n'appartient pas à $\mathscr{B}$ et l'on a*

$$h^2(P, Q) = -Cr^2 \log r + O(r^2) \quad \textit{lorsque } r \to 0,$$

où C est une constante > 0.

Preuve. On calcule

$$\varrho(P, Q) = \int_0^{1/2} \sqrt{1+r/\sqrt{2x}}\, dx + \tfrac{1}{2}\sqrt{1-2r-r^2}.$$

Si l'on pose dans l'intégrale

$$\sqrt{x} = \frac{r}{\sqrt{2}} \operatorname{sh}^2 t,$$

on obtient

$$\varrho(P, Q) = \frac{r^2}{4}\left[\frac{\operatorname{sh} 4a}{4} - a\right] + \frac{1}{2}\sqrt{1-2r-r^2} \quad \text{avec } \operatorname{sh}^2 a = r^{-1},$$

donc

$$\frac{\operatorname{sh} 4a}{4} = r^{-2}(r+2)\sqrt{r+1}.$$

$\operatorname{sh}^2 a = r^{-1}$ implique que

$$-\tfrac{1}{2}\operatorname{Log} r = -\operatorname{Log} 2 + a + \operatorname{Log}[1-e^{-2a}],$$

d'où il vient

$$a = -\tfrac{1}{2}\operatorname{Log} r + \operatorname{Log} 2 + O(r).$$

La conclusion s'en déduit par un calcul facile.

Je tiens à remercier tout particulièrement un des rapporteurs qui a découvert une grossière erreur dans une première version de ce travail et a fait de nombreuses remarques concernant en particulier les exemples.

Travaux cités

[1] R. R. Bahadur, *Some limit theorems in statistics*, Regional Conference Series in Applied Mathematics 4, SIAM, Philadelphia 1971.

[2] T. Bednarski, *Binary experiments, minimax tests and 2-alternating capacities*, Ann. Statist. 10 (1982), p. 226-232.

[3] L. Birgé, *Tests minimax robustes dans „Théorie de la robustesse et estimation d'un paramètre"*, Astérisque 43-44 (1976).

[4] – *Vitesses maximales de décroissance des erreurs et tests optimaux associés*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 55 (1981), p. 261-273.

[5] – *Approximation dans les espaces métriques et théorie de l'estimation* (1980) (à paraître).

[6] L. D. Brown, *Non-local asymptotic optimality of appropriate likelihood ratio tests*, Ann. Math. Statist. 42 (1971), p. 1206-1240.

[7] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, ibidem 23 (1952) p. 493-507.

[8] D. Dacunha-Castelle, Ecole d'Eté de Probabilités de Saint-Flour VII – 1977, Lectures Notes in Math. 678, Springer-Verlag, Berlin 1978.

[9] P. J. Huber, *A robust version of the probability ratio test*, Ann. Math. Statist. 36 (1965), p. 1753-1758.

[10] – and V. Strassen, *Minimax tests and the Neyman-Pearson lemma for capacities*, Ann. Statist. 1 (1973), p. 251-263.

[11] L. Le Cam, *Convergence of estimates under dimensionality restrictions*, ibidem 1 (1973), p. 38-53.

[12] – *On local and global properties in the theory of asymptotic normality of experiments*, Stochastic Processes and Related Topics 1 (1975), p. 13-54.

[13] – *An inequality concerning Bayes estimates*, manuscrit, 1978 (à paraître).

[14] B. Maurey, Séminaire Maurey-Schwartz, Exposé 10, *Théorèmes de Nikishin*, Publications du Centre de Mathématiques de l'Ecole Polytechnique, Paris 1973.

[15] F. Österreicher, *On the constructions of least favorable distributions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 43 (1978), p. 49-55.

[16] H. Rieder, *Least favorable pairs for special capacities*, Ann. Statist. 5 (1977), p. 909-921.
[17] E. N. Torgersen, *Comparison of experiments when the parameter space is finite*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 16 (1970), p. 219-249.
[18] G. Tusnády, *On asymptotically optimal tests*, Ann. Statist. 5 (1977), p. 385-393.
[19] — *Testing statistical hypotheses (an information theoretic approach)*, preprint, 1979.

U. E. R. d'Economie
Université Paris X Nanterre
200 Avenue de la République
92001 Nanterre Cedex, France

Received on 28. 9. 1979;
revised version on 9. 10. 1981