J. P. Rasson

Estimation de formes convexes du plan

Statistique et analyse des données, tome 4, n° 1 (1979), p. 31-46.
<http://www.numdam.org/item?id=SAD_1979__4_1_31_0>

Statistiques et Analyse des Données
1 - 1979 pp. 31, 46.

ESTIMATION DE FORMES CONVEXES DU PLAN

Dr. J.P. RASSON
Facultés Universitaires de Namur, Belgique.

Nous suggérons une solution au problème suivant proposé par le Prof. D.G. KENDALL : "Supposons qu'une réalisation d'un processus de Poisson dans le plan ait été *mutilée* de telle façon que nous ne puissions observer que les points intérieurs à un domaine convexe compact D , inconnu. *Trouvez* D en utilisant des méthodes d'inférence statistique".

Ceci est présenté comme l'analogue, dans le plan, du célèbre *Problème du Taxi* dont la généralisation unidimensionnelle est l'estimation des deux bornes inconnues d'un intervalle alors que nous observons un nombre fixé de points à l'intérieur de celui-ci. Nous justifions le choix des estimations classiques pour ce problème par les principes de l'estimation équivariante appliqués à la recherche d'estimateurs optimaux (U.M.V.).

Nous attirons alors l'attention sur la différence essentielle entre ce problème et son analogue dans le plan ; ici, il y a une nouvelle inconnue appelée la *forme*. Ce paramètre, qui n'est pas de dimension finie, apparaît comme un paramètre de nuisance pour l'estimation de la position et de la surface. Nous procédons à son élimination en le remplaçant par la forme qui maximise la vraisemblance marginale de ce paramètre et de la surface. Nous donnons alors les estimateurs de l'aire et de la position du domaine D qui correspondent au choix que nous avions fait dans le cas unidimensionnel, correspondance que nous détaillons. Notre solution finale sera une homothétie de l'enveloppe convexe de l'échantillon à partir du centre de gravité de celle-ci.

Finalement, nous signalons comment cette solution se généralise naturellement à $\mathbb{R}^n$ et nous donnons les possibilités d'application dans différents domaines.

1 - ESTIMATEURS EQUIVARIANTS

Ce rappel de notions et de résultats classiques de l'estimation équivariante est destiné essentiellement à faciliter la lecture de l'article. Le lecteur pourra trouver plus de détails dans Zacks [7] et Ferguson [2] .

1.1 - Structure des estimateurs équivariants

Supposons qu'il existe un groupe $\underline{\underline{G}}$ de transformations bijectives sur l'espace des échantillons $(\underline{\underline{X}} , \underline{\underline{B}})$ et un groupe $\overline{\underline{\underline{G}}}$ de transformations sur l'espace des paramètres Θ tels que

$$\forall\ g \in \underline{\underline{G}}\ , \quad \exists !\ \overline{g} \in \overline{\underline{\underline{G}}} \quad \text{tel que} \quad P_\theta(B) = P_{\overline{g}\theta}(g\ B) \quad \forall\ B \in \underline{\underline{B}}\ .$$

Dans ce cadre, nous dirons que :

1. une statistique T est *invariante* pour $\underline{\underline{G}}$ si
$$T(X) = T(g\ X) \quad \forall\ X \in \underline{\underline{X}}\ , \quad g \in \underline{\underline{G}} \qquad ;$$
2. l'*orbite* d'un point $X_0 \in \underline{\underline{X}}$ pour l'action du groupe $\underline{\underline{G}}$ est l'ensemble
$$G(X_0) = \{X\ ;\ \exists g \in \underline{\underline{G}} \mid g(X_0) = X\} \qquad ;$$
3. une statistique invariante est *maximale* si elle prend une valeur différente sur chaque orbite;
4. un estimateur est *équivariant* si
$$d(g\ X) = \overline{g}\ d(X) \qquad \forall\ g \in \underline{\underline{G}} \qquad ;$$
5. le problème de la recherche d'estimateurs rendant minimum le risque quadratique moyen est *invariant* pour $\underline{\underline{G}}$ si
$$(\theta - d)^2 = (\overline{g}\ \theta - \overline{g}\ d)^2 \qquad \forall\ g \in \overline{\underline{\underline{G}}}\ .$$

1.2 - Les paramètres de nuisance pour l'estimation équivariante

Soient $(X_1 , \dots , X_N)$ un échantillon de N variables aléatoires indépendantes de distribution commune, F_θ ,et un groupe $\underline{\underline{G}}$ de transformations sur $\underline{\underline{X}}$.

Supposons qu'il existe une partition du paramètre θ en (θ_1 , θ_2) telle que θ_2 est invariante pour $\underline{\underline{G}}$.

Si $U(X_1 , \dots , X_N)$ est une statistique invariante maximale pour $\underline{\underline{G}}$, la distribution de U ne dépend pas de θ_1 , mais, puisque θ_2 est invariante pour $\overline{\underline{\underline{G}}}$, cette distribution dépendra généralement de θ_2 . Nous la désignerons par $K_{\theta_2}(u)$. Si nous notons

$$(X_1 , \dots , X_N) = (T(X_1 , \dots , X_N) \ , \ U(X_1 , \dots , X_N))$$

la décomposition de la statistique, alors la distribution de T , étant donné $U = u$, dépendra généralement de θ_1 et θ_2 . Nous la noterons $P_{\theta_1,\theta_2}(t \mid U)$.

On peut montrer que, si $d(X)$ est un estimateur équivariant de θ_1 , son risque quadratique moyen ne dépend que de θ_2 et de l'orbite particulière à laquelle appartient θ_1 .

Supposons, de plus, que, si θ_2 est connu, il existe un estimateur équivariant qui minimise uniformément le risque quadratique moyen. Puisque ce risque varie en fonction de θ_2 , ce résultat ne sera plus vrai en général si θ_2 est inconnu. Dans ce cas, θ_2 est appelé un paramètre de *nuisance*.

1.3 - Estimation d'un paramètre de position

Considérons un problème statistique dont le *seul* paramètre inconnu est un paramètre de position, ceci signifiant que la densité $f_\theta(x)$ de chacune des variables aléatoires de l'échantillon s'écrit $f(x-\theta)$. $\underline{\underline{G}}$ est le groupe des translations sur $\underline{\underline{X}}$. Nous rechercons un estimateur équivariant de θ qui minimise le risque quadratique moyen $E_\theta(\theta - d)^2$.

Si $X = (T(X), U(x))$ est la décomposition adoptée au point 1.2, la solution du problème conditionnel, à savoir la recherche de la constante $\phi(u)$ qui minimise

$$E_\theta [(\theta - (T(X) + \phi(u)))^2 \mid U = u]$$

sera donnée par $\phi(u) = -E_0 [T \mid U = u]$ et la solution optimale du problème sera donc

$$d(X) = T(X) - E_0 [T(X) \mid U(X)]$$

De plus, $d(X)$ sera également la solution du problème original non conditionnel et aura la propriété supplémentaire d'être un estimateur sans biais de θ .

2 - LE PROBLEME UNIDIMENSIONNEL

Soient $X_1, \ldots, X_N$ des variables aléatoires indépendantes et identiquement distribuées, de densité uniforme sur l'intervalle $[a, b]$, $-\infty \leq a < b \leq +\infty$. Nous désirons estimer les paramètres inconnus a et b .

Nous exposerons uniquement le point de vue de l'estimation équivariante qui devra nous servir de point de comparaison avec le travail original. Ceci fera l'objet du paragraphe suivant. C'est évidemment le seul intérêt de ce développement qui n'est qu'un application triviale des théories de l'estimation équivariante

Considérons $(X_{(1)}, X_{(N)}) \equiv (\text{Min } X_i, \text{Max } X_i)$ qui est la statistique exhaustive minimale et complète pour le problème et décomposons-la en $((X_{(1)} + X_{(N)}) / 2 , X_{(N)} - X_{(1)})$.

Nous constatons que, pour l'action du groupe des translations, $X_{(N)} - X_{(1)} \equiv U(X_{(1)}, X_{(N)})$ est une statistique invariante maximale. La décomposition

$$(X_{(1)}, X_{(N)}) = ((X_{(1)} + X_{(N)}) / 2 , X_{(N)} - X_{(1)})$$

correspond donc au type de décomposition $X = (T(X), U(X))$ adoptée au § 1.

Nous pouvons aussi décomposer (a, b) en $(\theta_1, \theta_2) \equiv ((a+b)/2 , b-a)$ et constater que

$b-a$ est également invariant pour le groupe des translations $\underline{\underline{\bar{G}}} \equiv \underline{\underline{G}}$.

2.1 - Estimation équivariante de θ_1

Suivant le point 1.2, notre premier soin doit être de vérifier si, θ_2 étant connu, il existe un estimateur équivariant qui minimise uniformément le risque quadratique moyen. Il nous faudra ensuite vérifier si cela reste vrai lorsque θ_2 est inconnu.

Si θ_2 est connu, le seul paramètre inconnu étant un paramètre de position, la solution nous sera donnée par le point 1.3 :

$$d(X) = T(X) - E_0\,[\,T(X) \mid U(X)\,]$$

Or, la densité de $X = (X_1, \dots, X_N)$ s'écrit

$$f_{\theta_1,\theta_2}(x) = \frac{1}{\theta_2^N} \prod_{i=1}^{N} 1_{[\theta_1-\theta_2/2,\theta_1+\theta_2/2]}(x_i)$$

$$= \frac{1}{\theta_2^N}\; 1_{[\theta_1-(\theta_2-u(x))/2,\theta_1+(\theta_2-u(x))/2]}(t(x))$$

Pour la densité jointe de $t(x)$ et $u(x)$, nous aurons

$$f_{\theta_1,\theta_2}(t(x)\,,u(x)) = f_{\theta_1,\theta_2}(x) \quad N \quad (N-1) \quad u(x)^{N-2}$$

La densité marginale de u sera donc

$$K_{\theta_2}(u(x)) = \frac{(\theta_2 - u(x))}{\theta_2^N} \quad N \quad (N-1) \quad u(x)^{N-2}$$

et la densité conditionnelle de t :

$$P_{\theta_1,\theta_2}(t(x) \mid u(x)) = 1_{[\theta_1-(\theta_2-u(x))/2,\theta_1+(\theta_2-u(x))/2]}(t(x))\;(\theta_2 - u(x))^{-1}$$

de sorte que

$$P_{0,\theta_2}(t(x) \mid u(x)) = 1_{[-(\theta_2-u(x))/2,(\theta_2-u(x))/2]}(t(x))\;(\theta_2 - u(x))^{-1}$$

et

$$E_0(T(X) \mid U(X)) = 0$$

En conclusion, $T(X) = (X_{(1)} + X_{(N)})/2$ est uniformément optimal si θ_2 est connu, et, puisque il ne dépend pas de θ_2 , il le restera si θ_2 est inconnu. De plus, cet estimateur sera sans biais.

2.2 - Estimation équivariante de θ_2

Nous désirons trouver un estimateur équivariant de $\theta_2 = b-a$. La condition d'équivariance pour un estimateur $d(X)$ tel que $d(g\,X) = \bar{g}\,d(X)$ nous indique que, puisque θ_2 est invariant pour $\underline{\underline{\bar{G}}}$, nous devons avoir

$$d(X) \equiv \psi(T(X)\,,U(X)) = d(g(X)) = \psi(T(X)+g\,,U(X)) \quad \forall\, g \text{ tel que } -\infty < g < +\infty \quad .$$

Si nous prenons $g = -T(X)$, nous en déduisons que la forme générale d'un estimateur équivariant de θ_2 sera $d(X) = \psi(U(X))$.

Cependant, comme chaque fonction mesurable de $U(X)$ sera un estimateur équivariant de θ_2 , nous n'avons aucun guide qui nous permette de trouver l'estimateur optimal. Si nous ajoutons la condition supplémentaire pour les estimateurs d'être sans biais, nous pouvons trouver l'estimateur *U.M.V.U.* qui sera unique. En effet, il est facile de constater que $E(T(X))/\theta_2 = 1-2/(N+1)$ et donc que $(N+1)/(N-1)\;T(X)$ sera l'estimateur que nous cherchons.

3 - LE PROBLEME DANS LE PLAN

Soient X_1 , ... , X_N des variables aléatoires indépendantes et identiquement distribuées sur le domaine convexe compact D du plan.

Nous désirons estimer D .

3.1 - Introduction

Soient m la mesure de Lebesgue sur $\mathbf{R}^2$ muni de sa tribu borélienne, $\underline{\underline{D}}$ la classe des parties convexes compactes du plan de mesure strictement positive et $g(E)$ le centre de gravité d'un borélien E , par rapport à m pour un membre de $\underline{\underline{D}}$.

Pour de tels ensembles, notons

$$s(E) \equiv E - g(E) \qquad \text{(soustraction vectorielle)} \quad ,$$

sa *forme* . Signalons que cette notion de forme inclut la notion d'orientation.

Pour des ensembles A et B de $\mathbf{R}^2$, notons

$H(A) \equiv$ enveloppe convexe de A ;

$A + B$ = somme vectorielle de A et B ;

$A \ominus B = \{y \mid y+B \subseteq A\}$ (*Minkowski substraction* [3]) .

Nous définissons alors

$$A_B \underset{\text{déf.}}{\equiv} (A \ominus B) + B \quad ,$$

l'union de tous les translatés de B contenus dans A . Suivant cette définition, $A_B \subset A$ et $A_B \ominus B = A \ominus B$.

Supposons que nous observions $X = (X_1\,, \dots, X_N)$ où les X_i sont des variables aléatoires i.i.d. de densité uniforme sur D . La densité de X s'écrira

$$f_D(x) = \frac{1}{m(D)^N} \prod_{i=1}^{N} 1_D(x_i)$$

$$= \frac{1}{m(D)^N} 1_D(H(X))$$

Ceci implique que $H(X)$ est une statistique exhaustive minimale.

3.2 - Estimation équivariante de la position

Nous pouvons décomposer la statistique $H(X)$ en $(g(H(X)), s(H(X)))$ et, si nous considérons l'action du groupe des translations planes, $s(H(X)) \equiv U(X)$ est une statistique invariante maximale.

Conformément aux notations adoptées au point 1.2, nous pourrons donc écrire

$$(g(H(X)), s(H(X))) \equiv (T(X), U(X))$$

Nous pouvons également décomposer D en $(\theta_1, \theta_2) \equiv (g(D), s(D))$ et constater que $s(D)$ est invariante pour $\bar{\underline{\underline{G}}} = \underline{\underline{G}}$.

Suivant la méthode exposée au point 1.2, nous devons d'abord supposer θ_2 connu et caractériser l'estimateur optimal que nous obtenons dans ce cas. Suivant le point 1.3, celui-ci sera donné par

$$d(X) = T(X) - E_0 [T(X) \mid U(X)]$$

Or, la densité jointe de $T(X)$ et $U(X)$ s'écrit

$$f_{\theta_1,\theta_2}(t(x), u(x)) = 1_{[\theta_1+(\theta_2 \ominus u(x)]}(t(x)) \quad (1/m(\theta_2)^N N (N-1) \dots (N-(V_N+1)) m(u(x))^{N-V_N}$$

où V_N est le nombre de points appartenant à l'enveloppe convexe $H(X)$, de sorte que la densité marginale de $u(x)$ sera donnée par

$$K_{\theta_2}(u(x)) = m(\theta_2 \ominus u(x)) \quad m(\theta_2)^{-N} \quad m(u(x))^{n-V_N} \quad N \quad (N-1) \dots (N-(V_N+1))$$

et la densité conditionnelle de $T(X)$ par

$$P_{\theta_1,\theta_2}(t(x) \mid u(x)) = 1_{[\theta_1+(\theta_2 \ominus u(x))]}(t(x)) \quad m(\theta_2 \ominus u(x))^{-1}$$

Si $\theta_1 = 0$ (origine des translations),

$$P_{o,\theta_2}(t(x) \mid u(x)) = 1_{[\theta_2 \ominus u(x)]}(t(x)) \quad m(\theta_2 \ominus u(x))^{-1}$$

Ceci nous donne

$$E_0(T(X) \mid U(X)) = g(\theta_2 \ominus u(x))$$

L'estimateur optimal sera donc, dans notre cas,

$$d(X) = g(H(X)) - g(s(D) \ominus s(H(X)))$$

Nous constatons que, si $s(D)$ est connu, cet estimateur est uniformément optimal, mais, lorsque $s(D)$ est inconnu, cette propriété ne sera évidemment plus vraie. Il s'agit bien, dans ce cas, d'un paramètre de *nuisance* (point 1.2).

Nous ne désirons pas, à ce niveau, prendre position dans la querelle à laquelle se livrent les différentes écoles de statistique à propos de l'élimination des paramètres de nuisance. Nous nous rangerons plutôt à l'avis de A.W.F. Edwards [1] :

> *"I see no reason to suppose that it is always possible to eliminate a nuisance parameter. All our likelihood arguments are conditional on particular probability models : in a sence, THE MODEL ITSELF IS A NUISANCE PARAMETER. WE WOULD LIKE TO ARGUE WITHOUT IT, BUT CANNOT. IT SHOULD THEREFORE COME AS NO SURPRISE THAT IN SOME PROBLEMS, THE CONDITIONS UNDER WHICH WE ARGUE MAY HAVE TO INCLUDE SPECIFIC VALUES FOR NUISANCE PARAMETER".*

Or, si nous comparons soigneusement les sens du paramètre θ_2 dans $\mathbb{R}$ et dans $\mathbb{R}_2$, nous pouvons découvrir une différence essentielle. En effet, dans le cas unidimensionnel, $\theta_2 = b - a > 0$ peut être considéré comme un coefficient d'échelle de dimension 1 et, dans ce cas, la forme s peut se décomposer en $(b - a , [-1/2 , 1/2])$. Dans $\mathbb{R}_2$, si nous décomposons $s(D)$ en $(m(D) , s(D) / m(D))$, nous pouvons isoler un coefficient d'échelle de dimension 1 , $m(D)$, qui est bien l'équivalent de la longueur $b - a$. La différence essentielle se situera donc au niveau de la *forme unitaire* inconnue, $s(D) / m(D)$, dont force nous est de constater qu'elle est une partie essentielle de la distribution elle-même. Le paramètre de position $g(D)$ est d'ailleurs inextricablement lié à cette inconnue, de par la définition-même du centre de gravité qui n'a de sens que si nous nous référons à une forme unitaire fixée. C'est pourquoi, suivant le point de vue d'Edwards, nous ne pourrons traiter ce problème qu'après avoir attribué à $s(D)$ ou, en tout cas à $s(D) / m(D)$, une valeur spécifique. Cette dernière distinction est fondamentale car, si nous nous reportons au cas unidimensionnel, seule la forme unitaire était connue et nous n'avons pas eu besoin de donner à $b - a$ une valeur bien précise pour trouver une solution uniformément optimale. De toute façon, que nous arrivions à ne devoir donner de valeur spécifique qu'à $s(D) / m(D)$ ou qu'il nous faille le faire pour $s(D)$ tout entier, cette valeur ne peut pas raisonnablement être tout à fait quelconque. En effet, nous nous proposons d'estimer $s(D)$ et il serait donc assez illogique d'estimer $g(D)$ en se fixant une valeur $s(D)$ différente de celle que nous estimons. D'autre part, la structure même du problème et le type d'estimation que nous envisageons ne peuvent être qu'un encouragement à adopter cette attitude. En effet, nous avons pu décomposer la densité jointe de $g(H(X))$ et $s(H(X))$ en

$$P_{g(D),s(D)}(g(H(X)) , s(H(X))) = P_{g(D),s(D)}(g(H(X)) \mid s(H(X)) \quad K_{s(D)}(s(H(X)) \quad ;$$

la densité marginale de $s(H(X))$ ne dépend pas de $g(D)$ et donc, si nous désirons utiliser une valeur estimée de $s(D)$ afin de pouvoir estimer $g(D)$, il serait préférable de se restreindre à cette densité marginale.

Seulement, comme la densité conditionnelle de $g(H(X))$ dépend de $s(D)$, nous nous priverions, en faisant cette restriction, d'une partie de l'information donnée par la statistique exhaustive concernant $s(D)$. Cependant, si nous adoptons le point de vue de l'estimation équivariante de $s(D)$, tout estimateur sera de la forme

$$d(X) \equiv \psi(g(H(X)), s(H(X))) = d(g(X)) = \psi(g(H(X)) + k, s(H(X)) \quad \forall k \in \mathbb{R}_2 .$$

Il nous suffit de prendre de prendre $k = -g(H(X))$ pour découvrir que la forme générale d'un tel estimateur sera $d(X) = \psi(s(H(X)))$ et que donc nous n'utilisons de toute façon pas l'information concernant $s(D)$ qui nous est fournie par la densité conditionnelle de $g(H(X))$.

3.3 - Estimation équivariante de la forme unitaire

Comme nous venons de l'établir, toute estimation équivariante de $s(D)$ se fera sur base de la densité marginale de $s(H(X))$. Pour ce paramètre $s(D)$ de dimension ∞ , il semblerait logique, à défaut d'autres critères qui s'imposeraient naturellement, de s'intéresser à l'estimateur de vraisemblance maximale. Cependant, si nous considérons de nouveau la partition $s(D) = (m(D), s(D)/m(D))$, nous nous apercevons que ce type d'estimation est, tout au moins en ce qui concerne la partie $m(D)$, beaucoup moins cohérente avec le cas unidimensionnel où nous avions pu estimer $b-a$ suivant certains critères d'optimalité bien précis.

Notre premier réflexe doit donc être de tenter de séparer ces deux composantes du paramètre $s(D)$ en essayant de factoriser la densité de $s(H(X))$ en une densité marginale pour l'une des composantes et une densité conditionnelle pour l'autre. Cela s'est malheureusement révélé impossible. Et, dans ces conditions, nous avons adopté le point de vue de Sprott et Kalbleisch [6] dont la logique est appréciée par Edwards [1] , p. 118, qui sera d'attribuer à l'un des paramètres, la valeur qui rend maximale la vraisemblance lorsque l'autre paramètre reste fixé. Ce maximum relatif peut évidemment dépendre de la valeur fixée de l'autre paramètre. Dans notre cas, supposons $m(D)$ fixé. Nous aurons alors le

Théorème

Soit P un polygone convexe et $0 < m(P) < A$.
Alors $m(D \ominus P)$ est maximisé sur l'ensemble $\{D \mid D \in \underline{\underline{D}}, m(D) = A\}$ par $D = [A/m(D)]^{1/2} P$.

Démonstration

Voir Annexe I.

En conséquence, la vraisemblance relative de $s(D)/m(D)$ sera maximale pour la forme unitaire $s(D)/m(D)$ qui est exactement la forme unitaire de l'enveloppe convexe, et ce, *quelle que soit* $m(D)$. Ceci signifie que notre estimateur de la forme unitaire est, fort heureusement, invariant pour le groupe des transformations affines.

3.4 - Retour à l'estimation de $g(D)$

Si nous remplaçons, dans la formule de l'estimateur optimal de $g(D)$, la forme *unitaire* inconnue $s(D)/m(D)$ par sa valeur estimée $s(H(X))/m(H(X))$, nous obtenons

$$d(X) = g(H(X)) - g(s(D) \ominus s(H(X)))$$

$$d(X) = g(H(X)) - g(m(D)\,[\,s(H(X))\,/\,m(H(X))\,] \ominus m(H(X))\,[\,s(H(X))\,/\,m(H(X))\,])$$
$$= g(H(X)) - 0$$
$$= g(H(X))$$

C'est ainsi que, comme dans le cas du problème unidimensionnel, cet estimateur est optimal lorsque $m(D)$ est connu et le reste lorsque $m(D)$ est inconnu !

3.5 - Estimation de la surface du domaine

Nous recherchons un estimateur sans biais de $m(D)$, qui nous fournira, conjointement avec la solution apportée à la section 3.3, l'estimation complète de $s(D)$, à savoir une expansion homothétique de $H(X)$ dont le coefficient sera fixé par l'estimation de $m(D)$.

Soient A_N et V_N respectivement la surface et le nombre de sommets de l'enveloppe convexe $H(X)$. Supposons qu'un autre point x_{N+1} soit distribué uniformément dans D, indépendamment de X. Alors,

$$P(x_{N+1} \text{ est un sommet de } H(\{x_1, \dots, x_{N+1}\}) \mid X) = 1 - A_N / m(D)$$

$$P(x_{N+1} \text{ est un sommet de } H(\{x_1, \dots, x_{N+1}\})) = 1 - E(A_N) / m(D) \quad .$$

Nous servant alors de la symétrie des indices,

$$E(V_{N+1}) = \sum_{i=1}^{N+1} P(x_i \text{ est un sommet de } H(\{x_1, \dots, x_{N+1}\}))$$
$$= (N+1)\;[\,1 - E(A_N) / m(D)\,]$$

et donc

$$E(m(A_N)) / m(A) = 1 - E(V_{N+1}) / (N+1) \quad ,$$

ce qui est parfaitement l'analogue de la formule unidimensionnelle

$$E(X_{(N)} - X_{(1)}) / (b-a) = 1 - 2 / (N+1)$$

puisque, évidemment, dans ce cas-là, $E(V_{N+1}) = 2$.

Dans notre cas, cette formule est inutilisable brutalement car, bien sûr, notre échantillon ne comporte que N points et non pas $N+1$.

Si maintenant, nous remplaçons la forme unitaire par la valeur estimée au point 3.3, nous sommes à même d'utiliser directement cette relation. En effet, la valeur $E(V_{N+1})$ ne dépend que de la forme unitaire et est indépendante de la valeur $m(D)$. De plus, nous avons obtenu une forme unitaire estimée qui ne dépendrait pas de $m(D)$.

Dans ce cas (voir annexe II), nous pouvons calculer explicitement $E(V_{N+1})$ et donc, pour ce problème où nous remplaçons la forme unitaire inconnue par sa valeur estimée, nous obtenons un estimateur sans biais de la surface

$$(N+1) / ((N+1) - E_{s(H(X))/m(H(X))}(V_{N+1}))\quad m(H(X))$$

qui est exactement la valeur correspondante du cas unidimensionnel, à savoir

$$\frac{(N+1)}{(N+1)-2} \cdot (X_{(N)} - X_{(1)})$$

3.6 - Conclusions et remarques

1. Notre estimateur final sera donc

$$g(H(X)) + c \quad s(H(X)) \qquad \text{avec} \qquad c = (N+1) / ((N+1) - E_{s(H(X))/m(H(X))}(V_{N+1})$$

une homothétie de l'enveloppe convexe à partir de son centre de gravité.

Dans les figures 1 et 2, nous trouvons les résultats de simulation pour des domaines elliptiques et carrés, au moyen d'un échantillon de 200 points.

2. Si, dans un sens, le passage du cas linéaire au problème dans le plan n'était pas immédiat, notre solution pour $\mathbf{R}^2$, qui, comme nous l'avons vu, se particularise facilement à $\mathbf{R}$, se généralise presque trivialement à n'importe quel espace $\mathbf{R}^n$. Les sections traitant de la surface et de la position restent valables intégralement moyennant la convention que m représente la mesure de Lebesgue sur $\mathbf{R}^n$. Seule la démonstration du théorème concernant la forme nécessite une adaptation suivant le cas, bien qu'il ne s'agisse, en fait, que d'une simple transposition pour l'espace concerné.

ANNEXE I

Théorème 1

Soit P un polygone convexe et $0 < m(P) < A$. Alors, $m(D \ominus P)$ est maximisée sur l'ensemble $\{D \mid D \in \underline{\underline{D}}\ , m(D) = A\}$ par $D = [A / m(P)]^{1/2}\ P$. Cette solution est unique aux translations près.

Démonstration

Remarquons tout d'abord que, pour la solution $D = c\ p$, nous avons $m(D \ominus P) = (c-1)^2\ m(P) > 0$ de sorte qu'il nous suffit de nous limiter aux $D \in \underline{\underline{D}}$ tels que $m(D) = A$ et $m(D \ominus P) > 0$. Nous allons montrer que nous pouvons transformer un tel ensemble D en trois étapes, augmentant $m(D \ominus P)$ à chacune d'entre elles, pour obtenir finalement la forme de D proposée dans la solution et l'accroissement total sera strictement positif, sauf si D possède déjà la forme précédente.

(i) Soit $c = [m(D) / m(D_p)]^{1/2} \geq 1$. Si $c = 1$, D contient l'ensemble convexe compact D_p de même surface, ce qui entraîne évidemment que $D = D_p$. Supposons donc que $c > 1$. Alors, $m(D \ominus P) = m(D_p \ominus P) < c^2\ m(D_p \ominus P) = m(c(D_p \ominus P)) = m(c(D_p) \ominus c\ P) \leq m(c(D_p) \ominus P)$, cette dernière inégalité résultant de ce que $A \subset B$ implique $c \ominus B \subset c \ominus A$. donc, si D n'est pas

identique à D_P , nous obtenons une solution strictement meilleure en remplaçant D par $D' = c(D_P)$.
Mais, d'après le lemme suivant, nous aurons

$$(D')_P = [c(D_P)]_P = [(c\ D)_{cP}]_P = [(c\ D)_{P+(c-1)P}]_P = (c\ D)_{P+(c-1)P} = (c\ D)_{cP} = D' ,$$

de sorte que nous n'aurons à considérer que les $D \in \underline{\underline{D}}$ tels que $D = D_P$.

Lemme
$(D_{S+T})_S = D_{S+T}$ pour tout $D , S , T \in \underline{\underline{D}}$.

Démonstration
Evidemment, nous avons $(D_{S+T})_S \subset D_{S+T}$. Supposons que $y \in D_{S+T}$, c'est-à-dire qu'il existe $s \in S$, $t \in T$ tels que $(y - s - t) + S + T \subset D$. Cependant, $y \in (D_{S+T})_S$ si et seulement si il existe $s' \in S$ tel que $(y - s') + S \subset D_{S+T}$, c'est-à-dire $(y + s'' + s') \subset D_{S+T}$ pour tout $s'' \in S$.
Pour cela, il est nécessaire et suffisant qu'il existe $s''' \in S$ et $t' \in T$ tels que $(y + s'' - s' - s''' - t') + S + T \subset D$. Nous pouvons choisir $s' = s$, $s''' = s''$ et $t' = t$, ce qui ramène exactement cette condition à l'explicitation de $y \in D_{S+T}$. C.Q.F.D.

(ii) Nous supposons donc maintenant que $D = D_P$, c'est-à-dire que $D = \cup \{y + P \mid y + P \in D\}$. Fixons un ordre, par exemple celui des aiguilles d'une montre, et appelons a_1 , a_2 , ... , a_r les côtés de P dans cet ordre. Soit a_i' le vecteur extérieur perpendiculaire à a_i , et soit y_i un point de $D \ominus P$ dont la projection sur a_i' est maximale. Alors, $A_i \underset{\text{déf.}}{\equiv} y_i + a_i \subset \delta D$ puisqu'aucun point de D ne peut avoir sur a_i' une projection plus grande que celle d'un point de A_i . Puisque D est convexe, les translatés A_i des côtés de P doivent se présenter dans le même ordre dans δD . Désignons par B_i l'arc de δD joignant A_i et A_{i+1} (où $A_{r+1} = A_1$). Considérons le sommet de P qui est le point commun à a_1 et a_2 . L'ensemble des points que celui-ci peut atteindre lorsque l'on translate P à l'intérieur de D est borné par B_1 et la courbe fermée complétée par les translatés des arcs B_2 , ... , B_r (éventuellement dégénérés) dans l'ordre (cf. figure 3). Donc, $D \ominus P$ est bornée par un translaté de cette courbe et $m(D \ominus P)$ est la surface délimitée par cette courbe. Supposons que D n'est pas un polygone de r côtés parallèles à ceux de P . Alors, au moins un arc B_s n'est pas l'union de deux segments de droites prolongeant A_s et A_{s+1} . Si nous modifions uniquement cet arc, en gardant D convexe, on peut augmenter $m(D)$ et $m(D \ominus P)$ de deux quantités égales (cf. figure 3). Définissons un groupe de (A_i) comme étant un ensemble de segments de droites de cet ensemble, ayant chacun un point extrême commun avec le segment voisin si et seulement si celui-ci est inclus. Un groupe sera donc une suite maximale d'exemplaires des côtés consécutifs de P reliés par des arcs B_i dégénérés. On peut au moins déplacer un de ces groupes vers l'intérieur, en ajustant les arcs aux extrémités afin que la figure reste convexe. Une telle modification diminue $m(D) - m(D \ominus P)$ d'une valeur égale à la surface comprise entre le groupe initial et le groupe déplacé et les segments de droites joignant les points extrêmes (cf. figure 4). En combinant ces deux sortes de modifications, on peut donc augmenter strictement $m(D \ominus P)$ sans changer $m(D)$. Un nombre fini de telles modifications nous permet d'obtenir pour D la forme souhaitée.

(iii) Nous considérons donc maintenant uniquement les polygones convexes avec r côtés parallèles à ceux de P et de longueurs au moins égale.
Notons ces côtés par

$$a''_i = (p_i + 1)\, a_i \qquad (i = 1, \dots, r) \quad .$$

Soit t le plus grand indice parmi ceux des côtés a_i qui fait un angle inférieur à π avec a_1. Notons $[a, b]$ la surface de ce triangle de côtés a et b. Nous pouvons diviser D en triangles dont les sommets sont le premier point de a''_1 (suivant le sens indiqué) et les derniers points de $a''_2, \dots, a''_t$, et ensuite en prenant le premier point de a''_{t+1} et les derniers points de $a''_{t+2}, \dots, a''_r$, de sorte que la surface de D sera

$$m(D) = \sum_{i=1}^{t-1} [a''_1 + a''_2 + \dots + a''_i, a''_{i+1}] + \sum_{j=t+1}^{r-1} [a''_{t+1} + \dots + a''_j, a''_{j+1}]$$

(cf. figure 5).
Puisque

$$[\lambda\, a + u\, b, c] = \lambda\, [a, c] + u\, [b, c]$$

si $\lambda, u > 0$ et $0 \leqslant$ angle entre c et $b \leqslant$ angle entre c et $a \leqslant \pi$, nous pouvons exprimer $m(D)$ par

$$\Sigma\, (p_i + 1)\, (p_j + 1)\, [a_i, a_j] \qquad ,$$

la sommation s'effectuant sur tous les couples (i, j) tels que $1 \leqslant i < j \leqslant t$ ou $t+1 \leqslant i < j \leqslant 1$.
Soit $I(f(i, j)) = \Sigma\, f(i, j)\, [a_i, a_j]$ pour ces mêmes couples (i, j). Alors $m(D) = I((p_i + 1)\,(p_j + 1))$ et $m(D \ominus P) = I(p_i, p_j)$ puisque $D \ominus P$ est un polygone de côtés $p_1 a_1, \dots, p_r a_r$. Considérons une perturbation qui garde fermés les polygones D et $D \ominus P$, soit $p_i \to c(a)\,(p_i + a)$ où $c(a)$ est choisi de telle sorte que l'on ait toujours $m(D) = A$. Alors, puisque $c(0) = 1$,

$$m(D) = c(a)^2\, I((p_i + a)\,(p_j + a)) + c(a)\, I(p_i + p_j + 2\,a) + I(1)$$

$$0 = \left.\frac{d\, m(D)}{d\, a}\right|_0 = 2 \left.\frac{d\, c}{d\, a}\right|_0 I(p_i\, p_j) + I(p_i + p_j) + \left.\frac{d\, c}{d\, a}\right|_0 I(p_i + p_j) + 2\, I(1)$$

$$\left.\frac{d\, m(D \ominus P)}{d\, a}\right|_0 = 2 \left.\frac{d\, c}{d\, a}\right|_0 I(p_i\, p_j) + I(p_i + p_j)$$

$$= I(p_i + p_j) - 2\, I(p_i + p_j + 2)\, I(p_i\, p_j) / I(p_i + p_j + 2\, p_i\, p_j)$$

$$= I((p_i - p_j)^2) / I(p_i + p_j + 2\, p_i\, p_j)$$

Ce terme est positif et égale 0 si et seulement si tous les p_i sont égaux (puisque chaque $[a_i, a_j]$ est positif).

Nous avons donc montré que nous pouvons accroître strictement la valeur de $m(D \ominus P)$ pour tout $D \in \underline{\underline{D}}$ $(m(D) = A)$ qui ne serait pas une expansion homothétique de P, et ce, par une suite de modifications qui conduisent exactement à cette expansion. Celle-ci sera donc la solution optimale, ce qui était l'énoncé du théorème. C.Q.F.D.

ANNEXE II

Supposons donc que P soit un polynôme convexe de r côtés $a_1, \ldots, a_r$. Nous supposerons qu'aucun a_i n'est parallèle à un autre a_j. Soit K_{ij} la probabilité que $x_1 x_2$ soit un côté de $H(X)$ et qu'il coupe les côtés a_i et a_j de P. Nous aurons alors

$$E(V_{N+1}) = \sum_{1 \leq s < t \leq N+1} P(x_s x_t \text{ est un côté de } H(X))$$

$$= \binom{N+1}{2} P(x_1 x_2 \text{ est un côté de } H(X))$$

$$= \binom{N+1}{2} \sum_{i<j} K_{ij}$$

Si nous supposons que P est d'aire unitaire,

$$K_{ij} = \int F^{N-1} + (1-F)^{N-1} \, d x_1 \, d x_2$$

où nous intégrons sur les couples (x_1, x_2) tels que la droite joignant ces deux points coupe a_i et a_j et F est définie sur la figure 6 .

Exactement comme dans ([8] , équation 18), nous obtenons

$$K_{ij} = \int_0^{|a_i|} \int_0^{|a_j|} [F^{N-1} - (1-F)^{N-1}] \frac{1}{3} \sin^2 \theta \, (a+a')(b+b') \, d a \, d b$$

où $F = \frac{1}{2} (a\, b + a'\, b + a\, b') \sin \theta$ et où a, b, b' et θ sont définis sur la figure 6.

Nous effectuons les substitutions suivantes :

$$z = \sqrt{\sin \theta / 2} \quad , \quad x = z\,(a+a') \quad , \quad x' = z\, a' \quad , \quad x'' = z\,(a' + |a_i|)$$

$$y = z\,(b+b') \quad , \quad y' = z\, b' \quad \text{et} \quad y'' = z\,(b' + |a_j|) \quad .$$

Nous obtenons alors

$$K_{ij} = \frac{4}{3} \int_{x'}^{x''} \int_{y'}^{y''} [(1 + c_1 - x\,y)^{N-1} + (x\,y - c_1)^{N-1}] \; x\, y \; d x \; d y$$

où c_1 = surface (B D E) .

Effectuant alors la double intégration et posant c_2 = surface (A B C D) , c_3 = surface (A B D) et c_4 = surface (B C D) (cf. figure 6),

$$f(m) = c_2^m - c_3^m - c_4^m \quad , \quad g(m) = (1-c_2)^m - (1-c_3)^m - (1-c_4)^m$$

nous obtenons

$$K_{ij} = \frac{4}{3} \left[\frac{f(N+1) + g(N+1)}{(N+1)^2} - \frac{1}{(N+1)^N} \sum_{m=1}^{N} \{(-c_1)^{N+1-m} f(m) + (1+c_1)^{N+1-m} g(m)\} \right] .$$

Si la surface de P est $A \neq 1$, nous aurons simplement à diviser chaque c_i par A .

REFERENCES

[1] A.W.F. EDWARDS : "Likelihood", Cambridge University Press, 1972.

[2] T.S. FERGUSON : "Mathematical Statistics", Academic Press, 1967.

[3] G. MATHERON : "Random Sets and Integral Geometry", Wiley, 1975.

[4] D.G. KENDALL : "Window-Framed Poisson Point Processes", Communication personnelle.

[5] B.D. RIPLEY et J.P. RASSON : "Finding the Edge of a Poisson Forest", Journal of Applied Probability 14, pp. 483-491, 1977.

[6] J.D. KALFLEISCH et A. SPROOT : "Application of Likelihood Methods ...", J.R.S.S., B, 32, pp. 175-208, 1970.

[7] S. ZACKS : "The Theory of Statistical Inference", Wiley, New York, 1971.

[8] A. RENYI et R. SULANKE : "Uber die Konvexe Hülle von n Zufällig Gewählten Punkten", Z. Wahrscheinlichkeitstheorie 2, pp. 75-84, 1973 et 3, pp. 138-147, 1964.

Note .

Certains des résultats présentés ici ont été obtenus en collaboration avec le Professeur B.D. RIPLEY, Imperial College (cf. [5]).

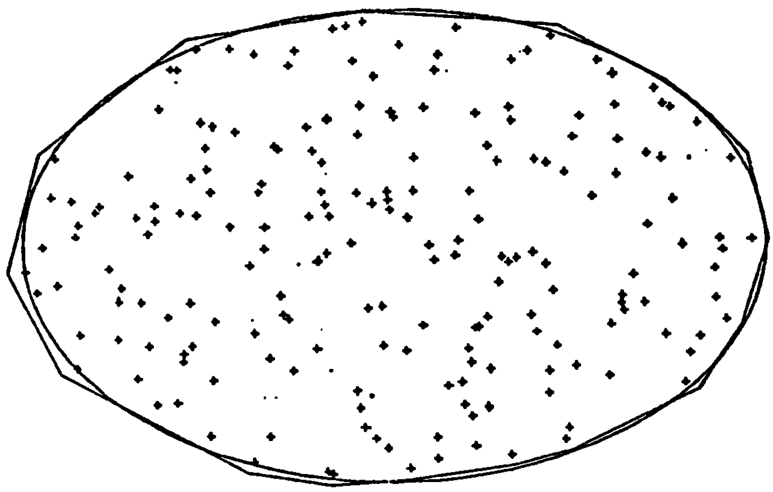

Figure 1

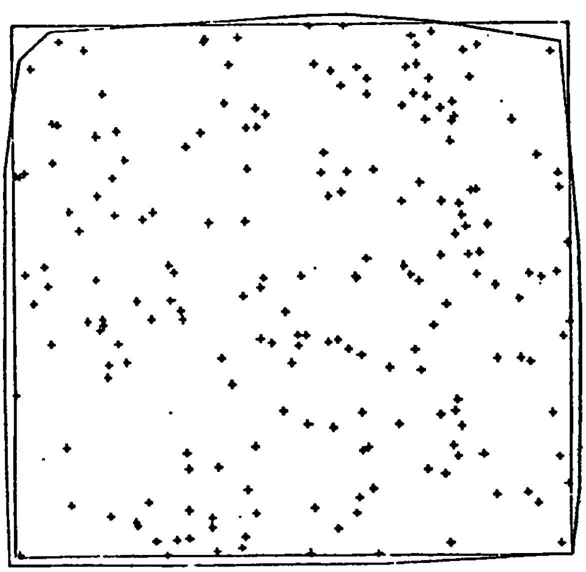

Figure 2

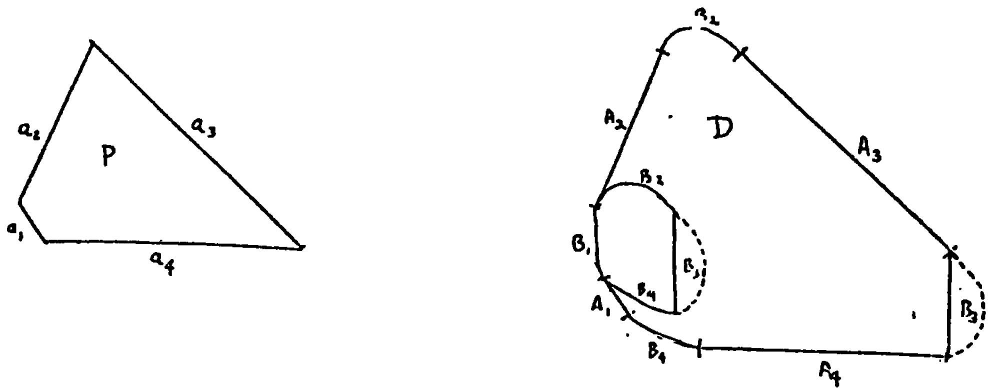

Figure 3

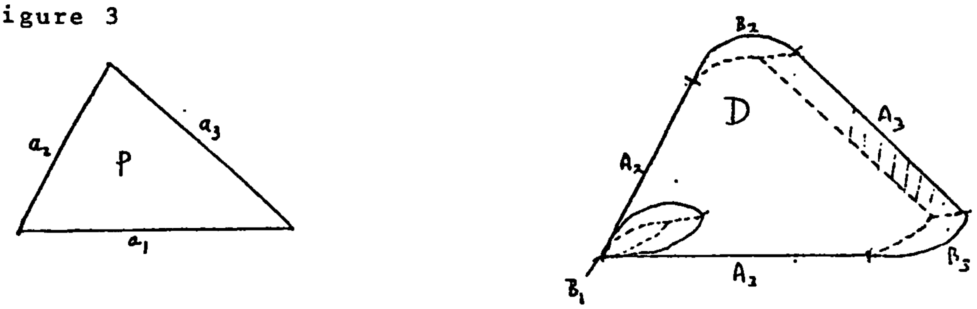

Figure 4

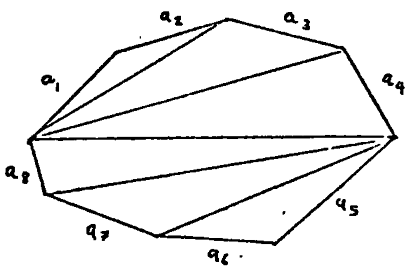

Figure 5

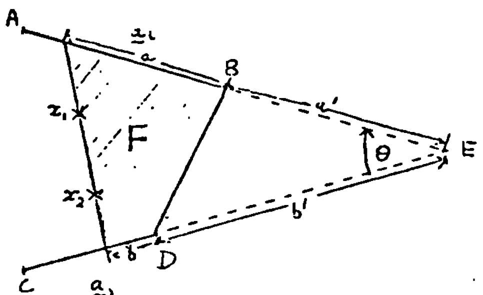

Figure 6