

Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*
Cotutelle internationale *CERFACS*

Présentée et soutenue le *03/07/2015* par :

ADIL AHIDAR-COUTRIX

**Surfaces quantile:
propriétés, convergences et applications**

JURY

PHILIPPE BERTHET	Directeur	Université Paul Sabatier
BENOÎTE DE SAPORTA	Examineur	Université Montpellier II
JEAN-FRANÇOIS DUPUY	Examineur	INSA Rennes
THIERRY KLEIN	Directeur	Université Paul Sabatier
FABIEN MANGEANT	Examineur	Airbus Group Innovations
DAVIT VARRON	Rapporteur	Université de Franche-Comté

École doctorale et spécialité :

MITT : Domaine Mathématiques : Mathématiques appliquées

Unité de Recherche :

Institut de Mathématiques de Toulouse (UMR 5219)

Directeur(s) de Thèse :

Philippe BERTHET et Thierry KLEIN

Rapporteurs :

Marc HALLIN Université Libre de Bruxelles

Davit VARRON Université de Franche-Comté

*Aux hommes justes disparus,
et à ceux partis dans la grâce
À vous mon grand-père,
Requiescat in pace...*

Avant propos

Ce manuscrit est l'aboutissement du travail de plusieurs années de thèse. En écrivant ces lignes je me souviens du chemin déjà parcouru, des bons moments comme des moins bons, de ces longues nuits de recherche, des moments où la solution apparaît clairement, de la joie que cela procure. Mais aussi de ces instants où l'on croit que rien ne va, ainsi que de ceux où l'on s'endort avec la solution et on se lève avec le doute parfois à raison et d'autres à tort. Des obsessions qui nous accompagnent et qu'on apprend à modérer, et des échanges riches et passionnants entre collègues, tellement indispensables car ils apportent un nouveau regard et permettent de relativiser nos échecs comme nos réussites, et sans lesquels la thèse ressemblerait à une longue traversée du désert. En somme, un pan de vie s'achève et je prends conscience que la thèse ne signe pas la fin, mais marque le début de la vie d'un chercheur, plus qu'un métier, une passion que je souhaite exercer dans les années à venir...

Je tiens à remercier Marc HALLIN et Davit VARRON d'avoir accepté de rapporter cette thèse malgré leurs emplois du temps serrés et les contraintes imposées par les délais. Je remercie également Benoîte DE SAPORTA, Jean-François DUPUY, et Fabien MANGEANT de s'être rendus disponibles afin de faire partie du jury. Je suis très honoré qu'ils aient acceptés.

Mes pensées vont tout d'abord à mes directeurs de thèses. A Philippe BERTHET, qui occupe une place si particulière. Il a su durant ces nombreuses années m'accompagner avec passion et finesse. Il m'a permis de jouir d'une grande liberté tout en se rendant disponible lorsque cela était nécessaire, et ce malgré ses nombreuses responsabilités. J'ai énormément appris à son contact, tant il est une source inépuisable d'inspiration et de stimulation scientifique. En effet, il a toujours su me pousser à approfondir mon raisonnement et encourager mon ouverture à différents domaines théoriques et appliqués, tout en ayant l'expérience nécessaire pour m'empê-

cher de m'égarer. Je souhaite donc ici lui témoigner tout mon respect et ma profonde gratitude. Et à Thierry KLEIN également qui fût à l'origine de ce projet de thèse, et dont je n'oublie pas que rien n'aurait pu se faire sans son concours. Il fût au cours de ces années la voix de la raison, maintenant par sa méthode et son pragmatisme le cap de ce travail vers son but assigné. Il sut encore se rendre disponible et répondre présent toutes les fois où je sollicitai son soutien. Pour tout cela il a mon entière reconnaissance.

Je remercie aussi l'équipe de recherche d'Airbus Group Innovations et tout particulièrement Jayant SEN GUPTA, porteur du projet industriel à l'origine de la thèse, qui s'est montré disponible et enthousiaste et avec qui j'ai eu le plaisir d'échanger à de nombreuses reprises. Je suis conscient du chemin qui reste à parcourir afin de mettre un point final à ce projet, mais j'ai espoir que notre collaboration se poursuive avec la participation de Nabil RACHDI.

Je suis reconnaissant envers le CERFACS et Serge GRATTON, chef d'équipe de recherche ALGO, pour m'avoir accueilli pendant une partie de ma thèse. Merci aussi à Xavier VASSEUR, ainsi qu'à tout le personnel administratif dont Michèle CAMPASSENS, Brigitte YZEL, Chantal NASRI et Lydia OTERO.

Citons encore ces rencontres marquantes qui ponctuent la vie de tout un chacun. C'est ici l'occasion de faire une rétrospective et de rendre hommage aux enseignants, aux professeurs, et à tous ceux qui m'ont transmis leur savoir et leur passion avant tout. Tout d'abord, mon père qui a su éveiller ma curiosité dès le très jeune âge, et m'a permis de faire des expériences dans son laboratoire de physique. Puis ce fût la rencontre avec Monsieur ROCHDI, qui accordait une grande importance à la rigueur mathématique. Il fût certainement durant les années du lycée la personne avec laquelle j'ai le plus échangé. Par sa culture mathématique et plus généralement scientifique il a su me captiver. Ensuite, il y eût Dominique BAKRY, qui à travers son cours « Intégration de Lebesgue » et grâce à son caractère et sa méthode d'enseignement, à su capter mon intérêt et par la même occasion donner à la théorie de la mesure une place bien particulière dans mon esprit. Mais aussi Franck BARTHE, qui m'introduisit aux probabilités, et au-delà il me permit de faire le lien entre probabilités et géométrie. Par sa rigueur et son implication, il a su m'apporter la stimulation dont j'avais besoin. Et encore, Philippe BERTHET - qui deviendra plus tard mon directeur de thèse - pour lequel seul le savoir compte. J'ai encore en mémoire ces cours de 2h s'étalant

sur plus de 5h, les pauses sans cesse repoussées, le programme tout bonnement impressionnant. Ce sont là mes meilleurs souvenirs d'enseignements.

Par ces rencontres, et d'autres encore comme celles avec Patrick CATTIAUX, Thierry KLEIN, Laure COUTIN, j'ai pu apprécier le domaine mathématique dans lequel j'allais finalement poursuivre mes recherches. Il est vrai qu'à cette période-là, alors jeune profane, les proba-stats et plus généralement l'aléatoire ne faisaient pour moi guère partie des mathématiques. Chose que je retrouve aujourd'hui chez certains de mes jeunes étudiants, car au-delà de la recherche, la transmission fait partie de la formation d'un futur docteur. C'est avec grand plaisir que j'ai accompli cette tâche, avec comme source d'inspiration mes propres enseignants.

Parce que nul n'est une île, ces dernières années furent aussi pour moi l'occasion de tisser des liens avec mes collègues, dont certains sont devenus des amis. Je pense à ceux dont le soutien fût précieux, tout particulièrement Fabrice Gamboa, Thibaut Le Gouic, Alexander Frieberg, Laurent Dietrich, Claire Delplancke, Daniel Kiouss, Pierre Monmarché, Mikol Solis, Thibault Espinasse, Paul Rochet, Raphaël Bouyrie (khouya), Nil Venet, Benoît Huou, Claire Boyer. Mais aussi Chloé, Olfa, Malika, Mélanie, Fanny, Selimé, Claire, Julie & Magalie, Hélène, Rim, Salomon, Santiago, Manu, Guillem, Jérémy de Picard, Kévin, Stéphane, Benoît, Xavier, Sofiane, Alhoucine, Youssef, Steven, Fabien, Amar, David, Mikko, Waly, Mawaki, Hung, Guillaume, et bien d'autres que j'oublie sans doute. Je pense également à l'équipe administrative de l'IMT, en particulier Marie-Laure.

Je garde une place toute particulière pour mes parents qui se sont saignés afin que je puisse poursuivre mes études, leur sacrifice m'a permis de réaliser mes rêves. Je suis conscient que malgré la distance qui nous sépare et les nombreuses années d'éloignement, leurs pensées m'ont toujours accompagné, de même que celles de Mamina, de mes frères, de mes tantes et de mes oncles.

Enfin et surtout, à Marine, celle qui m'accompagne depuis toutes ces années, celle qui fût, est et sera toujours mon étoile du berger, celle sur qui je peux toujours compter, et avec qui je peux tout partager. Merci d'être une si bonne épouse pour moi, si sincère et si dévouée, d'avoir été là à chaque instant, dans la joie comme dans la peine, d'avoir fait preuve de tant de patience, ce fut long, je le sais ! Merci pour tout.

Un grand merci à toutes et à tous.

Table des matières

Avant propos	iii
Table des matières	v
Notations et abréviations	ix
Introduction	xvi
0 Origine du problème	1
1 Préliminaires bibliographiques	9
1.1 Processus de quantile empirique réel	9
1.2 Quantile spatial	17
1.2.1 Processus de quantile généralisé et ensembles de vo- lume minimal	17
1.2.2 Ensembles de niveau et excès de masse	21
1.2.3 Fonctions de profondeur	24
1.2.4 Régression quantile	30
1.2.5 Trimming multivarié et ensembles de « Tukey-Nolan »	32
1.2.6 Autres	34
2 Surfaces quantiles via les demi-espaces	35
2.1 Surfaces quantiles via les demi espaces	36
2.1.1 Définition des surfaces quantile	36
2.1.2 Hypothèses de régularité	41
2.1.3 Régularité de la surface quantile	43
2.2 Surfaces quantiles empiriques	52
2.2.1 Estimateurs non-paramétriques	52
2.2.2 Lien avec le quantile réel	57
2.3 Théorèmes de convergence	58
2.3.1 Hypothèses	58

2.3.2	Consistance uniforme	60
2.3.3	Loi du logarithme itéré	61
2.3.4	Principe d'invariance faible	61
2.3.5	Principe d'invariance fort uniforme avec vitesse	63
2.3.6	Représentation de type Bahadur-Kiefer	63
2.4	Corollaires	64
2.5	Preuves	65
2.5.1	Preuve de la consistance	65
2.5.2	Preuve de la loi du logarithme itéré	69
2.5.3	Outils pour les preuves principales	71
2.5.4	Preuve des théorèmes de convergences	79
2.6	Simulations et interprétations	92
2.6.1	La loi gaussienne	92
2.6.2	Lois multimodales : mélanges gaussiens	103
2.6.3	Lois non usuelles	111
3	Surfaces quantiles via des classes générales	119
3.1	Introduction	119
3.2	Surfaces quantiles généralisées indexées par une fonction	120
3.2.1	Définitions	120
3.2.2	Régularité des φ -surfaces quantile	126
3.2.3	Théorèmes de convergences	137
3.2.4	Preuves des théorèmes	141
3.3	Surfaces quantiles généralisées	161
3.3.1	Indexation par des formes générales	161
3.4	Simulations et interprétations	166
3.4.1	La loi gaussienne	166
3.4.2	Lois multimodales : mélange gaussien	170
4	Utilisations des quantiles spatiaux	177
4.1	Introduction	177
4.2	Champs de profondeur	177
4.2.1	Définitions et théorèmes de convergence	178
4.2.2	Champs de profondeur pour les lois symétriques	181
4.2.3	Simulations	185
4.3	Distance quantile	189
4.3.1	Définitions et théorème	189
4.3.2	Test d'adéquation d'un échantillon	194
4.3.3	Test de comparaison d'échantillons spatiaux	197
4.4	Surfaces médianes	203
4.4.1	Définition et théorème	204

4.4.2	Surfaces médianes pour les lois symétriques	206
4.4.3	Simulations	208
4.5	Information auxiliaire	213
4.5.1	Motivation	213
4.5.2	Outil	214
4.5.3	Résultats	217
4.5.4	Simulations	219
	Bibliographie	234

Notations et abréviations

\mathbb{S}_{d-1}	ensemble des vecteurs unité de \mathbb{R}^d
$\mathcal{B}(\mathbb{R}^d)$	ensemble des boréliens de \mathbb{R}^d
$\mathcal{C}(\mathcal{X}, \mathcal{Y})$	ensemble des fonctions continues de \mathcal{X} dans \mathcal{Y}
$\mathcal{D}(\mathcal{X}, \mathcal{Y})$	ensemble des fonctions dérivables de \mathcal{X} dans \mathcal{Y}
$A \setminus B$	complémentaire de B dans A
$A \triangle B$	différence symétrique de A et B
∂A	bord topologique de A
$\text{int}(A)$,	intérieur topologique de A
$d_H(A, B)$	distance de Hausdorff entre les ensembles A et B
d_2	distance euclidienne
$\langle u, v \rangle$	produit scalaire des vecteurs u et v de \mathbb{R}^d
$x \vee y$	maximum de x et y
$x \wedge y$	minimum de x et y
$\lceil x \rceil$	partie entière de x par excès
$\llbracket n, m \rrbracket$	intervalle entier entre n et m
$\mathbb{E}(X)$	espérance de la variable aléatoire X
$\text{var}(X)$	variance de la variable aléatoire X
$\text{cov}(X, Y)$	covariance des variables aléatoires X et Y
$\xrightarrow{\mathcal{L}}$	convergence en loi
p.s.	presque sûrement
LGNU	loi des grands nombres uniforme
LLI	loi du logarithme itéré
TLCU	théorème limite central uniforme
BK	approximation du type Bahadur-Kiefer
PIF	principe d'invariance fort

Introduction

Il existe de nombreuses façons de définir un paramètre θ de localisation d'une loi dans son espace image. Par exemple dans \mathbb{R}^d un paramètre de localisation peut être la translation θ telle que la loi \mathbb{P}_X d'un échantillon X_1, \dots, X_n soit de la forme $\mathbb{P}_X = \mathbb{P}_{X_0 - \theta}$ avec P_{X_0} connue et θ inconnu. Dans ce cas paramétrique simple le paramètre θ s'estime par $\frac{1}{n} \sum_{i=1}^n X_i - \mathbb{E}(X_0)$ et localise entièrement la loi \mathbb{P}_X . Sur \mathbb{R} on dispose d'une autre notion de localisation basée sur les quantiles et leurs estimateurs basés sur la statistique d'ordre de l'échantillon. Malheureusement il n'existe aucune notion d'ordre qui soit pertinente sur \mathbb{R}^d . Une idée naturelle consiste à localiser puis à exclure des zones de « fort volume » et « faible probabilité », ce qui correspond intuitivement à écarter les parties de l'espace se trouvant « loin » de la « zone centrale ». C'est le principe des techniques dites de « trimming ». Pour $\alpha \in (0, 1)$ l'objectif est d'exhiber A_α une région centrale de l'espace telle que $P(A_\alpha) = \alpha$. Cependant il existe une infinité non dénombrable d'ensembles de probabilité α , comment les choisir ?

Une façon consiste à choisir parmi les A_α de probabilité α celles qui minimisent une quantité intrinsèque telle que le volume, le diamètre ou un indice de relation d'ordre arbitraire. Une première restriction au cadre non paramétrique intervient immédiatement : il faut écarter les lois de probabilité admettant des A_α tels que $\mathbb{P}(A_\alpha) = \alpha$ ayant des volumes ou des diamètres ou des indices arbitrairement faibles. Cela exclut typiquement les lois \mathbb{P}_X avec un support de dimension inférieure à d . Une seconde restriction intervient quand les A_α sont sélectionnés par minimisation empirique dans une classe d'ensemble de faible complexité : un grand nombre de lois de probabilité sont exclues de fait. Tout ceci impose de se restreindre à une classe d'ensembles A_α moins riche et de travailler conditionnellement à cette classe, or dans la pratique ce choix n'est ni aisé ni naturel et requiert de l'information supplémentaire sur la mesure de probabilité à trimmer. En outre le choix de telles classes est fortement sensible au nombre de zones modales de la loi inconnue P_X . Supposer connue le nombre de modes est trop exigeant dans un cadre purement non paramétrique. De plus l'idée même de classer

les A_α en fonction du volume, peut être dénuée de tout sens, dans le cas de la dimension infinie par exemple. Dans [LG13] il est introduit une fonction générale de taille et une méthode convergente et bien adaptée au trimming en dimension quelconque (et infinie). Cependant, les résultats de convergence à ce niveau de généralité approprié à la dimension infinie se limitent à la consistance, et les méthodes qui en découlent sur \mathbb{R}^d se heurtent à des calculs de minimisation assez lourds, et difficilement applicables.

Dans la thèse on introduit et on étudie une généralisation spatiale sur \mathbb{R}^d du quantile réel usuel sous la forme d'une *surface quantile* via des formes φ et d'un point d'observation O . En effet, notre point de départ est de simplement admettre la subjectivité due à l'absence de relation d'ordre totale dans \mathbb{R}^d et donc de développer une vision locale et directionnelle des données. Ainsi, les observations seront ordonnées du point de vue d'un observateur se trouvant à un point $O \in \mathbb{R}^d$, via une classe de formes φ et selon un niveau de quantile $\alpha \in \Delta \subset (1/2, 1)$. L'aspect inédit de ce travail est l'obtention de théorème très précis sur ce type d'objets, y compris pour les *demi-espace de Tukey*, largement utilisés en statistiques spatiales depuis plusieurs décennies. L'aspect original est que l'interprétation de ces surfaces fermées plongées dans \mathbb{R}^d est de nature subjective. En effet ces surfaces dépendent du point O qu'il faut ensuite bouger pour capter l'information statistique, et c'est justement ce qui permet des interprétations pertinentes et automatisables, pourvu que des caractérisations précises ressortent et qu'une classification des situations ait été menée au préalable. Dans cette thèse les surfaces elles-même et la convergence de leurs estimateurs non-paramétriques sont étudiés en profondeur, mais la classification s'avère très complexe dans le cas général et on se contente des remarques essentielles de simulations.

Premier chapitre

Dans le chapitre 1, on s'intéresse tout d'abord aux rappels bibliographiques des résultats autour du quantile de lois unidimensionnelles, la quasi totalité de ces résultats, se retrouve généraliser au cas multidimensionnel dans les chapitres 2 et 3. Ensuite, c'est aux différentes généralisations et extensions de la notion du quantile au cas multivarié, déjà présentes dans la littérature, que se consacre la suite du chapitre, à savoir, les ensembles de volume minimal, les ensembles de niveaux et excès de masse, la notion de profondeur (qui est en lien direct avec la notion des champs d'altitudes ou champs de profondeurs introduite dans la chapitre 4) et enfin le trimming multivarié et les ensembles de Tukey-Nolan.

Second chapitre

Dans le chapitre 2, on introduit la notion du quantile vue d'un observateur O dans la direction $u \in \mathbb{S}_{d-1}$ et de niveau α via des formes φ dans le cas où ces formes φ correspondent à des demi-espaces orthogonaux à chaque direction d'observation. Ce choix de classe implique que les résultats de convergence ne dépendent pas du choix de O . Sous des hypothèses minimales de régularité (\mathbf{H}_0^-) et (\mathbf{H}_0^+) (cf. hypothèses 2.11), l'ensemble des points quantile vue de O définit une surface fermée (un ensemble étoilé de points correspondant chacun à un quantile directionnel), cette propriété de régularité des surface du point de vu géométrique est un préalable à l'étude probabiliste et statistique des objets nouvellement introduit.

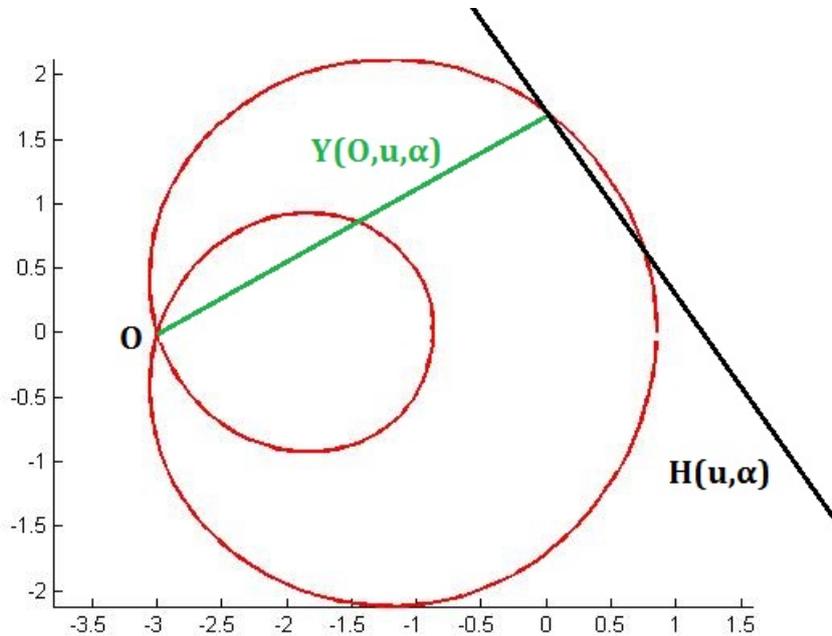


FIGURE 1 – En rouge : la surface quantile d'une gaussienne $N(0_2, I_2)$ pour $O = (-3, 0)$ et $\alpha = 0.8$.
En vert : la distance quantile vue de O dans la direction u de niveau α .
En noir : la frontière du demi-plan de normale u et de masse α .

Dans la suite, l'objectif poursuivit est de dégager des hypothèses minimales (\mathbf{H}_1) – (\mathbf{H}_4) (cf. la section hypothèses 2.3.1) assurant d'un coté la régularité des surfaces et permettant l'estimation uniforme de ces quantiles

spatiaux d'un autre.

Pour les surfaces quantile empiriques associées, on établit les théorèmes limites uniformément en le niveau de quantile et la direction d'observation, avec vitesse asymptotiques et bornes d'approximation non-asymptotiques. Principalement la *loi des grands nombres*, la *loi du logarithme itéré*, le *théorème limite central*, le *principe d'invariance fort uniforme* puis enfin l'*approximation du type Bahadur-Kiefer uniforme*, et avec vitesse d'approximation.

Ces théorèmes limites précisent clairement les corrélations jointes en la direction et l'ordre α du quantile des erreurs d'estimation. Ceci donne accès à des régions de confiance construites explicitement, ou par simulations du processus Gaussiens limites. Le point clef est une approximation de type Bahadur-Kiefer dans \mathbb{R}^d des surfaces quantile par le processus empirique indexé par les demi-espaces. La façon dont elle est obtenue est originale même en dimension $d = 1$, et repose intégralement sur la définition des surfaces quantiles et la compréhension fine du processus empirique, de ses trajectoires et de ses oscillations. Sous une hypothèse de régularité interdirectionnelle « n'exigeant pas de densité dans \mathbb{R}^d », le contrôle joint des surfaces quantile par le processus empirique se fait à une vitesse d'ordre $n^{-1/4}$, comme dans le cas unidimensionnel, et ceci indépendamment de la dimension d .

Troisième chapitre

Dans le chapitre 3, on étend les résultats du chapitre 2 (LGNU, TLCU, LLIU, principe d'invariance fort avec vitesse et approximation du type Bahadur-Kiefer) au cas où les formes φ sont prises dans une classe plus générale (fonctions, surfaces, projections géodésiques, etc) que des demi-espaces qui correspondent à des projections orthogonales par direction.

Dans ce cadre plus général, les résultats dépendent fortement du choix de l'observateur O , et c'est justement cette dépendance qui permet de tirer des interprétations statistiques : détection de modes, multimodalité, localisation de la masse, formes du support, détection de symétrie, etc. Des illustrations par simulation mettent en évidence quelques unes de ces interprétations et permettent de classifier les formes types voire typiques des surfaces quantiles (mélange, uni-modale, support de dimension inférieure,...).

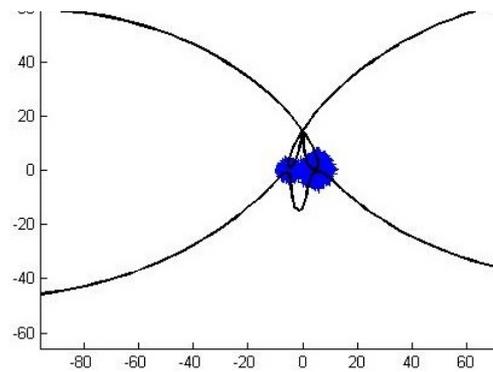


FIGURE 2 – En noir : la surface quantile par des paraboles pour une loi de mélange gaussien

Quatrième chapitre

En cette dernière partie de la thèse des conséquences méthodologiques en statistique inférentielle sont tirées. Tout d'abord on introduit une nouvelle notion de champ de profondeurs directionnelles \vec{D}_α baptisée champ d'altitude. A chaque point de l'espace on associe un vecteur dont la direction (pointant vers la masse) et la norme (quantifiant la distance à la masse) renseignent sur la centralité du point. En faisant varier l'ordre de quantile les déformations du champ aléatoire révèlent beaucoup d'information sur la localisation de la masse de probabilité.

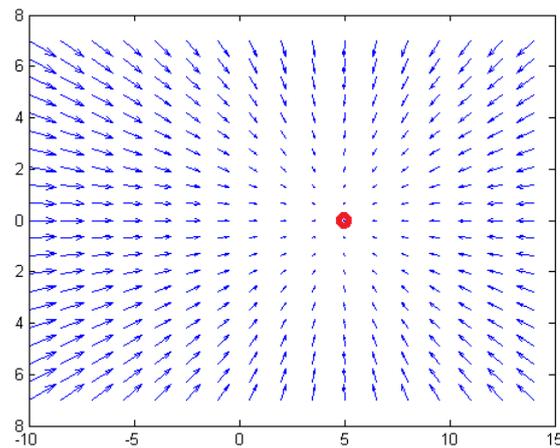


FIGURE 3 – Exemple d'un champ de de profondeurs \vec{D}_α pour $\alpha = 0.6$ d'une loi de mélange gaussien

Ensuite, on définit une notion de distance entre lois de probabilité, basée sur la comparaison des deux collections de surfaces quantile du type Gini-Lorentz. La convergence avec vitesse des mesures empiriques pour cette distance quantile, permet de construire différents tests d'adéquation et de comparaison en contrôlant leurs niveaux et leurs puissances.

Enfin, on donne une version des résultats et des simulations dans le cas où une information auxiliaire est disponible sur une ou plusieurs coordonnées de la variable aléatoire sous la forme de la connaissance exacte de la loi sur une partition finie.

Chapitre 0

Origine du problème

Dans ce chapitre il s'agit principalement de présenter le problème industriel auquel on s'est intéressé lors d'un stage chez EADS et qui a été à l'origine du problème de la thèse, on y présente le cheminement qui a conduit à étudier les quantiles sur \mathbb{R}^d sous la forme développée dans cette thèse.

Dans la première partie, on présente la thématique industrielle avant de donner dans la seconde partie sa formulation mathématique telle qu'elle a été proposée par l'industriel. Ensuite, dans la troisième partie on pointe les problèmes engendrés par ce choix de modélisation pour ensuite proposer une formulation alternative dans l'esprit du problème d'origine, construite grâce aux développements de la thèse et des outils introduits aux chapitres 2, 3 et 4.

1 Présentation de la thématique industrielle

Durant la phase de conception d'un avion par exemple, un ingénieur peut être confronté au problème d'allocation optimale de « marges de conception » aux différents fournisseurs de sous-systèmes. Il voudrait d'un côté laisser le plus de liberté possible à ceux qui contrôlent le moins leur processus de conception (nouvelle technologie, équipes moins expérimentées, etc) ou à ceux dont le coût de fabrication avec grande précision est le plus fort, et d'un autre côté être plus exigeant avec ceux qui ont un bon contrôle sur leur processus de conception (technologie bien rôdée, équipes plus expérimentées, etc) ou ceux dont le coût de fabrication avec une bonne précision est le moindre. En effet plus les marges autorisées sont contraignantes plus le coût de la conception est grand.

Afin de dimensionner un système (par exemple, un train d'atterrissage), outre la minimisation du coût tenant compte des capacités de chaque four-

nisseur, l'ingénieur doit également s'assurer qu'il sera suffisamment résistant pour remplir les différentes missions (manœuvres, scénarios catastrophes, etc) tout en étant le plus performant possible (le plus léger possible). Ainsi, à chaque manœuvre on associe un modèle permettant de calculer les « charges » qui s'exercent sur les différentes composantes (notamment les parties critiques du système). À chaque composante on associe le maximum des charges subies dans les différents scénarios. Ces modèles dépendent des paramètres de conception des différents sous-systèmes (paramètre physiques, géométrique, etc) qui sont généralement estimés donc incertains. Par conséquent, les charges calculées sont incertaines, aléatoires.

Une fois les charges subies par le système calculées, une « fonction de risque » leur est associée afin de déduire à travers un autre modèle la performance du système comme par exemple sa masse (ce nouveau modèle associe à chaque niveau de risque la masse ou l'augmentation de la masse du train d'atterrissage). C'est cette performance qu'on veut avoir plus grande qu'une performance minimale (ou la masse plus petite qu'une masse maximale).

Le problème est donc de répartir les marges de conception à travers une fonction de coût à minimiser sous contrainte que la performance calculée en prenant en compte un risque soit au moins égale à une performance minimale.

2 Formulation mathématique

Dans cette section on propose de traduire le problème ci dessus en termes mathématiques. Ainsi dans toute la suite on pose

- n : le nombre de fournisseurs.
- q : le nombre de scénarios à tester.
- p : le nombre de composantes du système.

Variables de conception. Afin de prendre en compte l'incertitude des paramètres, les variables de conception sont modélisées par un vecteur aléatoire $X = (X_1, \dots, X_n)$. généralement ce vecteur est supposé gaussien d'espérance connue m et de matrice de variance-covariance diagonale

$$\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_n^2).$$

Modèles des q scénarios. À chaque scénario on associe une fonction modélisant la charge subie par chacune des p composantes du système. En d'autres termes, pour $j \in \llbracket 1, q \rrbracket$, le $j^{\text{ème}}$ scénario est modélisé par la fonction

$$\begin{aligned} G^{(j)} : \mathbb{R}^n &\longrightarrow \mathbb{R}^p \\ x &\longmapsto G^j(x). \end{aligned}$$

.

Modèle des charges maximales. Une fois que les modèles des charges dans chaque scénario sont définis, on pose

$$\begin{aligned} G : \mathbb{R}^n &\longrightarrow \mathbb{R}^p \\ x &\longmapsto G(x) = \left(\max_{j \in \llbracket 1, q \rrbracket} |G_i^{(j)}(x)| \right)_{i \in \llbracket 1, p \rrbracket} \end{aligned}$$

comme le maximum des charges pris composante par composante sur les différents scénarios.

Variables des charges. Grâce au modèle des charges maximales, on définit les variables des charges maximales subies par les p composantes du système comme le vecteur aléatoire Y donné par

$$Y = (Y_1, \dots, Y_p) := G(X_1, \dots, X_n).$$

Fonction de risque. Pour un $\alpha \in (0, 1)$ fixé, on considère $y \mapsto q_\alpha(y)$ l'application qui associe à chaque charge y le risque de niveau α associé.

Remarque. Cette fonction reste très complexe à choisir, dans une première étude menée à EADS-IW, il a été considéré pour q_α le quantile composante par composante du vecteur des charges Y , ainsi

$$q_\alpha(Y) = (F_{Y_1}^{-1}(\alpha), \dots, F_{Y_p}^{-1}(\alpha)).$$

Modèle de performance. Comme fonction modélisant la performance du modèle, on considère la fonction $\varphi : \mathbb{R}^p \rightarrow \mathbb{R}^+$ qui associe au risque des charges maximale la masse du système. Cette fonction est généralement supposée connue, continue (ou continue par morceaux) et croissante.

Répartition des marges et la fonction de coût. Dans ce problème, l'attribution des n marges à chaque fournisseur est faite à travers le choix des σ_i . Ainsi plus σ_i est petit plus la loi de X_i est concentrée autour de sa moyenne et plus on est exigeant envers le $i^{\text{ème}}$ fournisseur et vis-versa. Cette allocation de marge est faite travers une fonction de coût f à minimiser

$$\begin{aligned} f : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \sigma &\longmapsto f(\sigma) = \sum_{i=1}^n f_i(\sigma_i) \end{aligned}$$

où le coût $f_i(\sigma_i)$ de la conception avec niveau de précision σ_i tient compte des capacités du $i^{\text{ème}}$ fournisseur. Par exemple on peut prendre $f(\sigma) = \sum_{i=1}^n a_i \sigma_i^2$ une fonction quadratique avec a_i des indices de confiances.

Problème d'optimisation. La formulation mathématique du problème industriel est donc la suivante. Soit $m \in \mathbb{R}^n$, $\sigma = (\sigma_1, \dots, \sigma_n) \in \mathbb{R}_+^n$ et $X^\sigma \sim N(m, \Sigma)$ un vecteur gaussien d'espérance m et de matrice de variance-covariance $\Sigma = \text{diag}(\sigma)$. On cherche

$$\arg \min_{\sigma \in \mathbb{R}_+^n} f(\sigma)$$

sous la contrainte

$$R_\alpha(\sigma) \leq M_{max}$$

avec

- $M_{max} \in \mathbb{R}^+$ la masse maximale.
- f la fonction de coût.
- $Y^\sigma = G(X^\sigma)$ le vecteur de charge maximal.
- $R_\alpha(\sigma) = \varphi(q_\alpha(Y^\sigma))$.

Remarque. La moyenne m est supposée solution, i.e. $\sigma = (0, \dots, 0)$ est solution su problème inverse. En faite cette solution est la plus exigeante qui soit dans le sens où on laisse aucune marge aux différents sous-traitants.

3 Problématique

Comme évoqué plus haut, le choix de la fonction de risque en terme de quantile n'est ni évident ni aisé, d'autant plus que la variable aléatoire en sortie du modèle des charges maximales est multivariée. En plus de devoir choisir ou développer une notion de quantile sur \mathbb{R}^p , il faut que cette dernière

fassent sens d'un point de vue industriel, prenne en compte la loi jointe des variables de charges et satisfasse des propriétés de convergence similaires au quantile réel c'est-à-dire avec de bonnes vitesses de convergence des estimateurs empiriques, si possible avec peu de données. En outre il est souhaitable de se doter d'une notion de quantile spatial permettant des simulations rapides et faciles avec contrôle uniforme des fluctuations stochastiques.

Dans une première étude le choix pour cette fonction s'est porté sur l'application

$$\Psi : \sigma \mapsto (F_{Y_1^\sigma}^{-1}(\alpha), \dots, F_{Y_p^\sigma}^{-1}(\alpha)).$$

Ce choix de quantile conduit très vite à des dépassements de risque global à cause de la non prise en compte des dépendances entre les variables de charges maximales. En effet, en prenant en compte que les quantiles des lois marginales, le problème change complètement de nature puisque par définition des variables de charges maximales

$$Y_i = \max_{j \in \llbracket 1, q \rrbracket} |G_i^{(j)}(X_1, \dots, X_n)| \quad \text{pour } i \in \llbracket 1, p \rrbracket$$

sont fortement corrélées. Ce choix de quantile fait croire aussi que l'on puisse agir de manière indépendante sur chacune des composantes Y_i ce qui conduit irrémédiablement à une perte de contrôle du risque global.

Dans cette thèse, la notion de quantile de niveau α est décrite par une surface ancrée en un point d'observation O que l'on notera $Q(O, \alpha)$. Quel que soit l'observateur O choisi, la connaissance de toutes les surfaces de niveau α caractérise entièrement la loi, ainsi aucune perte d'information (dépendance, corrélation linéaire ou non linéaire, etc) n'est à déplorer. De plus, chaque points de la surface représente un quantile de la loi jointe projetée dans une direction u de la sphère unité S_{d-1} . Dans sa version la plus simple, la surface quantile correspond aux points quantiles des variables aléatoires

$$\langle Y, u \rangle = \sum_{i=1}^p u_i Y_i.$$

Ainsi demander à ce que les quantiles dans toutes les directions soit de niveau α revient à exiger que quelle que soit « l'importance » des variables Y_i les unes par rapport aux autres on soit protégé avec un niveau α de risque dans toutes les directions $u \in S_{d-1}$.

Un tel objet n'a jamais été étudié avant, et s'avère être un très bon candidat pour la généralisation dans \mathbb{R}^p du quantile réel. Dans les chapitres

suiuants on démontre qu'il possède les mêmes propriétés que le quantile usuel.

À cause de la nature du quantile (une surface plongée dans \mathbb{R}^p) et de la fonction de performance, le problème doit être reformulé tout en conservant l'esprit du problème d'origine. Plusieurs formulations sont possibles, parmi ces dernières on cite deux. Néanmoins il manque toute l'étape de discussion avec l'industriel afin de savoir laquelle ou lesquelles semblent les plus raisonnables et pertinentes en terme de capacité à fournir ou à développer les modèles correspondants.

Formulation I. Soit $m \in \mathbb{R}^n$, $\sigma = (\sigma_1, \dots, \sigma_n) \in \mathbb{R}_+^n$ et $X^\sigma \sim N(m, \Sigma)$ un vecteur gaussien d'espérance m et de matrice de variance-covariance $\Sigma = \text{diag}(\sigma)$. On cherche

$$\arg \min_{\sigma \in \mathbb{R}_+^n} f(\sigma)$$

sous la contrainte

$$Q(m, \alpha) \subset \mathcal{R}_{M_{max}}$$

avec

- $Q(m, \alpha)$ est la surface quantile de niveau alpha vue de m
- $\mathcal{R}_{M_{max}}$ un domaine d'admissibilité pour avoir une masse plus petite que la masse maximale M_{max} .
- f la fonction de coût.
- $Y^\sigma = G(X^\sigma)$ le vecteur de charge maximale.

Remarque. *Un candidat dont la consistance reste à vérifier est de tout simplement prendre*

$$\mathcal{R}_{M_{max}} = \varphi^{-1}(]-\infty, M_{max}])$$

avec φ la fonction de performance. Pour le moment aucune étude n'a été menée afin de permettre de valider ou d'invalider ce choix de contrainte.

Dans l'éventualité où fournir un domaine d'admissibilité s'avérerait trop ardue, une autre façon de faire pourrait être en lien avec la notion des champs de profondeurs (champs d'altitudes définis dans le Chapitre 4). Principalement du fait de la lourdeur des notations et de la nécessité d'introduire et de définir les objets des chapitres 2, 3 et 4, on se contente simplement ici de décrire l'idée générale. Ainsi en supposant que l'on ne puissent avoir accès à tout le domaine d'admissibilité, on peut se contenter de quelques points critiques dans l'espace des contraintes, qu'on ne veut

absolument pas voir dans le domaine des admissibles avec une grande probabilité.

Étant donné ces points critiques, on souhaite minimiser la fonction coût sous la contrainte que ces points soient de profondeur plus petite qu'une certaine profondeur tolérable. En effet, plus un point est profond plus sa probabilité d'être dans le domaine des admissibles est grande (dans les cas où une médiane existe, celle-ci est le point le plus profond et plus un point est dans les extrêmes moins il est profond).

4 Conclusion et perspectives

Comme énoncé plus haut, l'objectif poursuivi dans cette thèse est l'introduction puis l'étude d'une généralisation à \mathbb{R}^d du quantile usuel, ainsi que les applications statistiques des objets nouvellement introduits. Un retour au problème industriel est prévu dans la suite de ces travaux, les étapes envisagées afin de mener à bien la suite du projet reste à déterminer avec l'industriel, néanmoins, voici quelques pistes de recherche exploitants au mieux les résultats des travaux de thèse.

1- Une analyse de sensibilité du nouveau modèle associant σ à la surface quantile semble primordiale, cette analyse peut être effectuée soit à travers la distance quantile introduite dans le Chapitre 3, soit grâce à la généralisation des indices de Sobol introduit par F. Gamboa, A. Janon, T. Klein, A. Lagnoux-Renaudie dans [GJKL14].

2- En fonction du choix des contraintes, les champs de profondeur peuvent permettre de simplifier grandement le problème d'optimisation, reste à déterminer la consistance d'un tel choix.

3- À cause de la grande dimension des variables de charges, la prise en compte d'une information auxiliaire pour générer des surfaces quantiles, peut être un véritable atout dans la phase d'exploration dans un algorithme d'optimisation.

Chapitre 1

Préliminaires bibliographiques

Dans ce chapitre, on expose des résultats classiques dans chaque domaine mathématique abordé dans les chapitres de la thèse. Les preuves des théorèmes énoncés, ainsi que les propriétés des objets introduits, peuvent être trouvées dans les références citées tout au long de cette section.

Comme énoncé dans l'introduction, l'un des objectifs principaux de cette thèse, est de proposer une généralisation originale de la notion du quantile usuel sur \mathbb{R} , ainsi que l'étude des comportements asymptotiques et non-asymptotiques des processus empiriques associés. Ainsi, on commence par rappeler les résultats classiques du processus empirique unidimensionnel, ensuite on présente quelques généralisations multidimensionnelles déjà proposées dans la littérature, pouvant servir dans les chapitres suivants à faire les liens avec les notions nouvellement introduites.

1.1 Processus de quantile empirique réel

Dans ce paragraphe, on présente les résultats classiques sur les processus empiriques sur \mathbb{R} , il s'agit principalement de rappeler les théorèmes de convergence faible, le Principe d'invariance fort et l'approximation de Bahadur-Kiefer. Afin de comparer nos résultats par la suite, nous avons fait le choix dans la majorité des cas, de présenter les versions les plus générales des théorèmes existants.

Soit X, X_1, X_2, \dots, X_n des variables aléatoires réelles indépendantes et de même loi $P = \mathbb{P}^X$ définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Supposons que $\mathbb{E}(X) = 0$ et $\text{var}(X) = 1$. Posons $S_n = \sum_{i=1}^n X_i$ et $S_0 = 0$.

Dans les années 50', suite aux travaux novateurs de Erdős et Kac [EK46] introduisant une nouvelle méthodologie de preuve appelée "*Principe d'invariance*", Donsker développa l'outil permettant de construire une suite de

processus stochastiques $\{S_n(t); 0 \leq t \leq 1\}$ sur $C(0, 1)$, à partir des sommes partielles S_0, S_1, \dots, S_n

$$S_n(t) = n^{-1/2} \left(S_{[nt]} + X_{[nt]}(nt - [nt]) \right).$$

En utilisant la version multivariée du théorème de la limite centrale, on a pour tout $0 \leq t_1 < t_2 < \dots < t_k \leq 1$,

$$(S_n(t_1), S_n(t_2), \dots, S_n(t_k)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} (W(t_1), W(t_2), \dots, W(t_k))$$

où \mathcal{L} désigne la convergence faible sur \mathbb{R}^d et W est un processus de Wiener issu de zéro.

En utilisant Kolmogorov [Kol33], [Kol31] et Khinchine [Khi33], il apparaît que la loi limite de $\{S_n(t); 0 \leq t \leq 1\}$ coïncide avec celle de $\{W(t); 0 \leq t \leq 1\}$.

Plus rigoureusement, Donsker dans [Don51], montre que pour toute fonctionnelle $g : C(0, 1) \rightarrow \mathbb{R}$ continue sur $C(0, 1)$ muni de la topologie uniforme, on a

$$g(S_n(t)) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} g(W(t)).$$

Une fois établie la convergence en loi, il est naturel de s'intéresser à la vitesse de convergence. Le premier résultat dans ce sens est dû à Strassen en 1964 [Str64], qui fait une avancée majeure en montrant qu'on peut construire un espace $(\Omega, \mathcal{A}, \mathbb{P})$ sur lequel sont définis simultanément une suite $(X_i)_{i \geq 1}$ et un processus de Wiener $\{W(t); 0 \leq t \leq 1\}$ avec $\mathbb{E}(X) = 0$ et $\text{var}(X) = 1$ tels que

$$\lim_{n \rightarrow \infty} \frac{|S_n - W(n)|}{\sqrt{n \log \log n}} \stackrel{p.s.}{=} 0.$$

Il est évident que ce résultat est plus fort que ceux du type convergence en loi, car d'une part, il s'agit de convergence presque sûre avec vitesse et d'autre part, il implique le résultat de convergence faible de Donsker. Dans ce cas, on parle de « *principe d'invariance fort* » en contraste avec la convergence faible dans le théorème de Donsker.

Définition 1.1. On appelle « *fonction de répartition empirique* » associée à X_1, X_2, \dots, X_n , la fonction aléatoire F_n^X

$$\begin{aligned} F_n^X(x) &= \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{(-\infty, x]}(X_i) \\ &= \frac{1}{n} \# \{i \in \llbracket 1, n \rrbracket : X_i \leq x\}, \quad \forall x \in \mathbb{R}. \end{aligned}$$

Le théorème de Glivenko-Cantelli [SW86, Can17, Gli33], communément appelé le « *théorème fondamental de la statistique* » assure que F_n est un bon estimateur de F , plus précisément on a

Théorème 1.2 (Glivenko-Cantelli [Can17, Gli33]). *Soit $(X_i)_{i=1}^n$ une suite de variables aléatoires i.i.d. alors*

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n^X(x) - F^X(x)| = 0 \text{ p.s.}$$

Ensuite en 1956 Dvoretzky, Kiefer et Wolfowitz [DKW56], donneront une version plus forte impliquant le théorème de Glivenko-Cantelli.

Théorème 1.3 (Dvoretzky, Kiefer et Wolfowitz [DKW56]). *Il existe une constante universelle $C > 0$ telle que pour tout $n \in \mathbb{N}$ et $r > 0$*

$$\mathbb{P} \left(\sqrt{n} \sup_{x \in \mathbb{R}} |F_n^X(x) - F^X(x)| \geq r \right) \leq Ce^{-2r^2}.$$

Le théorème de consistance uniforme de Glivenko-Cantelli garantit qu'avec probabilité 1, observer une infinité de fois des variables aléatoires indépendantes restitue exactement leur loi. La précision d'une telle approximation est liée à la vitesse de convergence, pour cela on s'intéresse au processus normalisé α_n^X , appelé « *processus empirique* »,

$$\alpha_n^X(x) = \sqrt{n} \left(F_n^X(x) - F^X(x) \right), \forall x \in \mathbb{R}.$$

L'étude du processus α_n^X est plus aisée à travers le « processus empirique uniforme », grâce au théorème d'inversion qui stipule que si $U \sim \mathcal{U}(0,1)$ alors $F_X^{-1}(U) \sim X_1$ où F_X^{-1} est l'inverse (à gauche) généralisé de F^X défini par

$$F_X^{-1}(t) = \inf\{x : F^X(x) \geq t\}, \forall t \in (0, 1).$$

Soient U_1, U_2, \dots, U_n des variables aléatoires indépendantes de lois uniformes sur $[0, 1]$, en posant F_n^U la fonction de répartition empirique uniforme et α_n^U le processus empirique uniforme qui s'écrit

$$\alpha_n^U(t) = \sqrt{n} \left(F_n^U(t) - t \right), \forall t \in [0, 1].$$

En posant $X_i = F_X^{-1}(U_i)$ on a pour tout n , $F_n^X = F_n^U \circ F^X$ et $\alpha_n^X = \alpha_n^U \circ F^X$. C'est pourquoi dans la littérature, de nombreux résultats sont établis uniquement pour α_n^U sans perte de généralité.

Le théorème de Donsker assure que α_n^U converge faiblement vers un pont brownien standard B dans $\mathcal{D}(0, 1)$,

Théorème 1.4. Soit $(U_i)_{i \geq 1}$ des variables aléatoires indépendantes de loi uniforme sur $[0, 1]$ alors,

$$\alpha_n^U \xrightarrow{\mathcal{L}} B_P$$

où B_P est un pont brownien, i.e. le processus gaussien centré de structure de covariance pour tout $s, t \in [0, 1]$

$$\text{cov}(B_P(t), B_P(s)) = \min(s, t) - st.$$

Grâce à la transformation des quantiles, on a le théorème de convergence faible dans $\mathcal{D}(0, 1)$ de α_n^X suivant

$$\alpha_n^X \xrightarrow{\mathcal{L}} B_P \circ F^X.$$

De la même façon qu'on a défini F_X^{-1} , on définit la fonction inverse ou quantile de F_n^X

$$F_{X,n}^{-1}(t) = \inf\{x : F_n^X(x) \geq t\}, \quad \forall t \in (0, 1).$$

En suite on définit le processus empirique des quantiles β_n^X

$$\beta_n^X(t) = \sqrt{n} \left(F_{n,X}^{-1}(t) - F_X^{-1}(t) \right), \quad \forall t \in [0, 1].$$

Commençons tout d'abord par faire la même simplification que pour α_n^X . Soit U_1, U_2, \dots, U_n des variables aléatoires indépendantes de loi uniforme sur $[0, 1]$, on notera β_n^X le processus empirique des quantiles associés aux X_1, \dots, X_n . Par le lemme de transformation des quantiles on a

$$F_{n,X}^{-1} = F_X^{-1} \circ F_{n,U}^{-1}.$$

Donc

$$\beta_n^X(t) = \sqrt{n} \left(F_X^{-1} \circ F_{n,U}^{-1}(t) - F_X^{-1}(t) \right).$$

Remarque 1.5. Pour les variables aléatoires uniformes on peut facilement montrer que

$$\sup_{t \in [0,1]} |\alpha_n^U(t)| = \sup_{t \in [0,1]} |\beta_n^U(t)|.$$

Par la remarque précédente on peut immédiatement donner l'analogie pour les processus des quantiles empiriques, du théorème de Glivenko-Cantelli et du théorème de Dvoretzky, Kiefer et Wolfowitz.

Théorème 1.6. *Il existe une constante universelle $C > 0$ telle que pour tout $n \in \mathbb{N}$ et $r > 0$*

$$\mathbb{P} \left(\sup_{t \in [0,1]} |\beta_n^U(t)| \geq r \right) \leq Ce^{-2r^2}.$$

Il en découle facilement le théorème suivant.

Théorème 1.7.

$$\lim_{n \rightarrow \infty} \sup_{t \in [0,1]} |F_{n,U}^{-1}(t) - t| = \lim_{n \rightarrow \infty} \sup_{t \in [0,1]} \frac{1}{\sqrt{n}} |\beta_n^U(t)| = 0 \text{ p.s.}$$

L'analogie de ce théorème pour les variables aléatoires quelconques n'est pas immédiat, de plus il exige des conditions sur le support de la loi qui sont naturellement vérifiées par la loi uniforme. Pour ce faire, dans le cadre des lois à densité par rapport à la mesure de Lebesgue, on introduit la fonction suivante.

Notation 1.8. *Soit X une variable aléatoire de densité f par rapport à la mesure de Lebesgue, on note*

$$h_X = f \circ F_X^{-1}.$$

Cette fonction joue fréquemment un rôle en probabilité, elle apparait de façon naturelle dans les théorèmes limites suivants. Une des façons les plus simples de la faire apparaitre est de dériver F_X^{-1} supposée dérivable

$$\left(F_X^{-1} \right)' = \frac{1}{f \circ F_X^{-1}} = \frac{1}{h_X}.$$

On pose $\varphi = F_X^{-1}$ et on suppose que F_X est dérivable de dérivée f . On sait par la transformation des quantiles, que pour tout $t \in (0,1)$

$$\beta_n^X(t) = \sqrt{n} \left(\varphi \circ F_{n,U}^{-1}(t) - \varphi \circ Id(t) \right).$$

S'il existe $m, M > 0$ tels que

$$0 < m \leq f \circ F_X^{-1} = h_X \leq M < +\infty,$$

on obtient par la delta méthode [VDVW00] que pour tout $t \in (0,1)$

$$\beta_n^X(t) \xrightarrow{\mathcal{L}} \frac{1}{h_X(t)} B_P(t).$$

Théorème 1.9. Soit X une variable aléatoire de fonction de répartition F_X différentiable de dérivée f . Supposons qu'il existe $m, M > 0$ tels que

$$0 < m \leq f \circ F_X^{-1} = h_X \leq M < +\infty$$

alors pour tout $t \in (0, 1)$,

$$\beta_n^X(t) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{t(1-t)}{h_X^2}\right).$$

Par définition de la fonction inverse, si $F_n^X > F^X$ alors $F_{X,n}^{-1} \leq F_X^{-1}$. En d'autres termes,

$$\alpha_n^X > 0 \Rightarrow \beta_n^X \leq 0.$$

On peut ainsi dire que α_n a tendance à contrebalancer/compenser β_n^U . De ce fait on s'intéresse non pas à l'écart entre α_n^U et β_n^U mais à celui entre α_n^U et $-\beta_n^U$, ainsi la quantité d'intérêt est

$$\alpha_n^U + \beta_n^U.$$

Ce processus est souvent appelé « Processus de Bahadur-Kiefer ». En 1966 Bahadur [Bah66] a montré le théorème suivant.

Théorème 1.10. Pour $t \in [0, 1]$, quand n tend vers $+\infty$ on a

$$\frac{n^{1/4}}{(\log n)^{1/2}(\log \log n)^{1/4}} \left(\alpha_n^U(t) + \beta_n^U(t) \right) = O(1) \text{ p.s.}$$

L'ordre de grandeur exact de $(\alpha_n^U(t) + \beta_n^U(t))$ a été donné en 1967 par Kiefer dans [Kie67].

Théorème 1.11. Pour $t \in [0, 1]$

$$\limsup_{n \rightarrow \infty} \pm \frac{n^{1/4}}{(\log \log n)^{3/4}} \left(\alpha_n^U(t) + \beta_n^U(t) \right) = 2^{5/4} 3^{-3/4} (t(1-t))^{1/4} \text{ p.s.}$$

Ensuite, en 1970, Kiefer dans [Kie70] montra deux résultats sur la norme supérieure de ce processus.

Théorème 1.12.

$$\limsup_{n \rightarrow \infty} \frac{n^{1/4}}{(\log n)^{1/2}(\log \log n)^{1/4}} \left\| \alpha_n^U + \beta_n^U \right\|_{\infty} = 2^{-1/4} \text{ p.s.}$$

ainsi que

$$\frac{n^{1/4} \left\| \alpha_n^U + \beta_n^U \right\|}{(\log n)^{1/2} \sqrt{\left\| \alpha_n^U \right\|_{\infty}}} \xrightarrow{P} 1 \text{ quand } n \rightarrow +\infty.$$

Il faudra attendre 1990 pour avoir la version presque sûre de [Kie70], c'est Deheuvels et Mason dans [DM90] qui montrent le théorème suivant.

Théorème 1.13.

$$\lim_{n \rightarrow +\infty} \frac{n^{1/4} \|\alpha_n^U + \beta_n^U\|_\infty}{(\log n)^{1/2} \sqrt{\|\beta_n^U\|_\infty}} = 1 \text{ p.s.}$$

Grâce à la remarque 1.5, on montre immédiatement que

$$\lim_{n \rightarrow +\infty} \frac{n^{1/4} \|\alpha_n^U + \beta_n^U\|_\infty}{(\log n)^{1/2} \sqrt{\|\alpha_n^U\|_\infty}} = 1 \text{ p.s.}$$

Or Par Mogul'skiï [Mog79] on sait que

$$\limsup_{n \rightarrow \infty} (\log \log n)^{-1/2} \|\alpha_n^U\|_\infty = 2^{-1/2} \text{ p.s.}$$

donc

$$\|\alpha_n^U + \beta_n^U\|_\infty = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right).$$

Entre autres, on a

$$\lim_{n \rightarrow +\infty} \|\alpha_n^U + \beta_n^U\|_\infty = 0 \text{ p.s.}$$

or on sait que

$$-\alpha_n^U \xrightarrow{\mathcal{L}} -B_P \stackrel{\mathcal{L}}{\sim} B_P$$

donc

$$-\alpha_n^U + (\alpha_n^U + \beta_n^U) = \beta_n^U \xrightarrow{\mathcal{L}} B_P$$

Ainsi pour la loi uniforme, on obtient

Théorème 1.14.

$$\beta_n^U \xrightarrow{\mathcal{L}} B_P$$

Afin de donner les versions de Bahadur-Kiefer (Théorème 1.12) pour une loi quelconque, on s'intéresse à la quantité suivante.

$$R_n = \left\| \alpha_n^X \circ F^{-1} + h_X \cdot \beta_n^X \right\|_\infty$$

avec $h_X = f \circ F_X^{-1}$. Tout d'abord, voici les hypothèses de régularité garantissant la convergence de R_n .

Hypothèses. Soit X une variable aléatoire de fonction de répartition F dérivable de dérivée f . On note

$$a = \sup \{x : F(x) = 0\} \quad , \quad b = \inf \{x : F(x) = 1\} ,$$

et on suppose

(A0) F est deux fois dérivable.

(A1) $\inf_{t \in (a,b)} f(t) > 0$.

(A2) $\sup_{t \in (a,b)} |f'(t)| < \infty$.

Théorème 1.15. Soit X une variable aléatoire de fonction de répartition F , de dérivée f . Sous les conditions (A0), (A1) et (A2), on a

$$\limsup_{n \rightarrow \infty} \frac{n^{1/4} R_n}{(\log n)^{1/2} (\log \log n)^{1/4}} = 2^{-1/4} \quad p.s.$$

Csörgo et Revez dans [CR78] obtiennent un contrôle presque sûr de l'écart entre β_n^X et β_n^U . Pour ce faire on a besoin des conditions suivantes.

Hypothèses. Soit X une variable aléatoire de fonction de répartition F différentiable de dérivée f . On note

$$a = \sup \{x : F(x) = 0\} \quad , \quad b = \inf \{x : F(x) = 1\}$$

et

$$L^+ = \lim_{x \rightarrow b} f(x) \quad , \quad L^- = \lim_{x \rightarrow a} f(x).$$

(A3) il existe $\gamma > 0$ telle que

$$\sup_{a < x < b} F(x)(1 - F(x)) \frac{|f'(x)|}{f^2(x)} \leq \gamma$$

(A4) $L^+ < \infty$ et $L^- < \infty$

(A5) $\min(L^-, L^+) > 0$

(A'5) si $L^- = 0$ (resp. $L^+ = 0$) alors f est croissante (resp. décroissante) au voisinage de a (resp. au voisinage de b)

Théorème 1.16. Soit X une variable aléatoire de fonction de répartition F différentiable de dérivée f . Sous les conditions (A0)-(A3) on a

$$\sup_{[\delta_n, 1 - \delta_n]} |h_X \beta_n^X - \beta_n^U| \stackrel{p.s.}{=} O(n^{-1/2} \log \log n)$$

avec $\delta_n = 25n^{-1} \log \log n$.

Si de plus les conditions (A4) et (A5) sont vérifiées alors

$$\sup_{[0,1]} |h_X \beta_n^X - \beta_n^U| \stackrel{p.s.}{=} O(n^{-1/2} \log \log n).$$

En remplaçant la condition (A5) par (A'5) on obtient pour $\varepsilon > 0$

$$\sup_{[0,1]} |h_X \beta_n^X - \beta_n^U| \stackrel{p.s.}{=} \begin{cases} O(n^{-1/2} \log \log n) & \text{si } \gamma < 1 \\ O(n^{-1/2} (\log \log n)^2) & \text{si } \gamma = 1 \\ O(n^{-1/2} (\log \log n)^\gamma (\log n)^{(1+\varepsilon)(\gamma-1)}) & \text{si } \gamma > 1 \end{cases}$$

1.2 Quantile spatial

1.2.1 Processus de quantile généralisé et ensembles de volume minimal

La notion de « processus de quantile généralisé » a été introduite par Einmahl et Mason en 1992 [EM92], il s'agit de minimiser une fonction λ sur une classe d'ensemble mesurable de \mathbb{R}^d à valeur réelle, sous contrainte de probabilité au moins α . Dans le cas particulier où λ est la mesure de Lebesgue λ_d sur \mathbb{R}^d , les ensemble réalisant le minimum de λ_d sous contrainte de probabilité au moins α sont appelés « ensembles de volume minimal ».

Définition. Soit X, X_1, X_2, \dots, X_n des variables aléatoires indépendantes à valeurs dans \mathbb{R}^d et de même loi $P = \mathbb{P}_X$. Soit \mathbb{A} une classe d'ensembles P -mesurables de \mathbb{R}^d et λ une fonction de $\mathbb{A} \rightarrow \mathbb{R}$.

On appelle fonction quantile basée sur P , λ et \mathbb{A} , la fonction

$$V(\alpha) = \inf\{\lambda(A) : P(A) \geq \alpha, A \in \mathbb{A}\} \quad 0 < \alpha < 1$$

et la fonction quantile empirique associée

$$V_n(\alpha) = \inf\{\lambda(A) : P_n(A) \geq \alpha, A \in \mathbb{A}\} \quad 0 < \alpha < 1$$

On remarque que cette notion est fortement liée au choix de la fonction λ et de la classe \mathbb{A} .

Exemple 1.17. En dimension $d = 1$, pour $\mathbb{A} = \{(-\infty, t] : t \in \mathbb{R}\}$ et $\lambda((-\infty, t]) = t$, on a

$$\begin{aligned} V(\alpha) &= \inf\{\lambda(A) : \mathbb{P}(X \leq t) \geq \alpha, A \in \mathbb{A}\} \\ &= \inf\{t : \mathbb{P}(X \leq t) \geq \alpha\} \\ &= F^{-1}(\alpha) \end{aligned}$$

de la même manière on peut montrer que

$$V_n(\alpha) = F_n^{-1}(\alpha).$$

Ainsi on retrouve la fonction quantile usuelle dans \mathbb{R} et sa version empirique associée.

Remarque 1.18.

- On peut mentionner aussi les « α -shorth », un autre cas particulier qui a été étudié dans la littérature en rapport aux quantiles ([ABH⁺72], [BEN05]). En effet ces ensembles peuvent s'écrire comme un cas particulier du quantile généralisé pour $d = 1$, $\lambda = \lambda_d$ la mesure de Lebesgue et pour la classe $\mathbb{A} = \{[s, t] : s, t \in \mathbb{R}\}$. Les α -shorth correspondent aux intervalles bornés de plus petite taille et de probabilité au moins α .
- Pour $\lambda = \lambda_d$ la mesure de Lebesgue sur \mathbb{R}^d on parle d'« ensembles de volume minimal » car dans ce cas $V_n(\alpha)$ correspond au volume du plus petit ensemble dans \mathbb{A} contenant au moins une proportion α des points de l'échantillon.

Notation. Pour $f : (\Omega, \mathcal{A}) \rightarrow \mathbb{R}$ on définit la fonction mesurable f^* comme la plus petite fonction mesurable de Ω dans \mathbb{R} tel que pour tout $\omega \in \Omega$ on a $f(\omega) \leq f^*(\omega)$.

Afin de donner le théorème de consistance de $V_n - V$, on introduit les hypothèses suivantes :

Hypothèses 1.19.

- (C'1) \mathbb{A} et λ sont tels que pour tout $0 < \alpha < 1$ on a $V_n^*(\alpha) < \infty$ p.s.
- (C'2) \mathbb{A} est une classe de Glivenko-Cantelli pour P .

Théorème (Polonik 1997). Sous les conditions (C'1) et (C'2), si V est continue sur $[0, 1]$ alors

$$\lim_{n \rightarrow \infty} |V_n^*(\alpha) - V^*(\alpha)| = 0 \text{ p.s.}$$

Ensuite si V est uniformément continue sur un compact $K \subset [0, 1]$, alors on a le théorème suivant

Théorème (Polonik 1997). Sous les conditions (C'1) – (C'2), si V est uniformément continue sur un compact $K \subset [0, 1]$, on a

$$\lim_{n \rightarrow \infty} \sup_{\alpha \in K} |V_n^*(\alpha) - V^*(\alpha)| = 0 \text{ p.s.}$$

Ensuite voici les conditions garantissant un théorème du type Hájek-Bickel [SW86] pour le processus des quantiles empiriques généralisés défini ci-dessous.

Hypothèses 1.20.

- (C1) λ est continue par rapport à la métrique $d(A, B) = \mathbb{P}(A \Delta B)$.
- (C2) λ et \mathbb{A} sont tels que $(V_n)_{n \geq 1}$ sont presque sûrement finis sur $(0, 1)$.
- (C3) Il existe une sous classe dénombrable d'ensemble $\mathbb{D} \subset \mathbb{A}$ telle que pour tout $A \in \mathbb{A}$ il existe $(D_n)_{n \geq 1}$ dans \mathbb{D} avec $\mathbb{1}_{D_n}(x) \rightarrow \mathbb{1}_A(x)$ pour tout $x \in \mathbb{R}$.
- (C4) \mathbb{A} est précompact et, de plus, il existe un espace de probabilité sur lequel on peut construire X_1, \dots, X_n et B_p un pont brownien tel que
- $$\sup_{A \in \mathbb{A}} |\sqrt{n}(P_n(A) - P(A)) - B_p| \rightarrow 0 \text{ a.s.}$$
- (C5) Pour tout $A \in \mathbb{A}$, $0 < P(A) < 1$.
- (C6) $T_t = \{A \in \mathbb{A} : \lambda(A) = V(t), P(A) = t\} \neq \emptyset$ pour tout $0 < t < 1$.
- (C7) Pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour tout $0 \leq t_1, t_2 \leq 1$ avec $|t_1 - t_2| < \delta$ et $A_1 \in T_{t_1}$ il y a un $A_2 \in T_{t_2}$ avec $d(A_1, A_2) < \varepsilon$.
- (C8) V est strictement croissante sur $(0, 1)$ avec $H = V^{-1}$ continûment dérivable avec $H' = h$ sur (L, R) avec $L = \lim_{t \rightarrow 0} V(t)$ et $R = \lim_{t \rightarrow 1} V(t)$.
- (C9) Pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour tout $A \in \mathbb{A}$ vérifiant $0 < t - \delta < P(A) < t < 1$ et $\lambda(A) < V(t)$ il existe un $A' \in T_{H(\lambda(A))}$ tel que $d(A, A') < \varepsilon$.

Voici la définition du processus des quantiles empiriques généralisés.

Définition. Pour $t \in (0, 1)$ on appelle processus des quantiles empiriques généralisés

$$\dot{\beta}_n(t) = n^{1/2} (V_n(t) - V(t))$$

et on note $g = h \circ V$ où $h = (V^{-1})'$.

Notation : Pour tout $t \in [0, 1]$ on note

$$B(t) = \sup_{A \in T_t} B_p(A)$$

Théorème (Einmahl, Mason [EM92]). Sous (C1) – (C9), pour tout $0 < a < b < 1$, quitte à élargir l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ on a

$$\lim_{n \rightarrow \infty} \sup_{a \leq t \leq b} |g(t) \dot{\beta}_n(t) + B(t)|^* \stackrel{p.s.}{=} 0$$

Remarque. Le facteur de renormalisation $g(t)$ est l'analogue au $f \circ F^{-1}(t)$ dans le cadre classique unidimensionnel.

Enfin, voici un théorème de convergence presque sûre uniforme pour la distance de Hausdorff entre les ensembles réalisant la minimisation de $V(\alpha)$ et $V_n(\alpha)$.

Notation. On note \tilde{A}_α et $\tilde{A}_{n,\alpha}$ des ensembles qui réalisent respectivement la minimisation de $V(\alpha)$ et $V_n(\alpha)$. Ainsi on a

$$V(\alpha) = \lambda(\tilde{A}_\alpha) \text{ et } V_n(\alpha) = \lambda(\tilde{A}_{n,\alpha}).$$

On a besoin des hypothèses supplémentaires suivantes.

Hypothèses.

- (C'3) P admet une densité par rapport à la mesure de Lebesgue.
- (C'4) Pour tout $\alpha \in [0, 1]$ il existe un « unique » ensemble \tilde{A}_α tel que $P(\tilde{A}_\alpha) = \alpha$.
- (C'5) Pour tout $\alpha \in [0, 1]$ il existe un ensemble $\tilde{A}_{n,\alpha}$ réalisant la minimisation de $V_n(\alpha)$.
- (C'6) λ est semi continue inférieurement pour d .

Voici un théorème dû à Polonik en 1997 [Pol97a], un résultat de convergence presque sûre uniforme pour la distance d entre \tilde{A}_α et $\tilde{A}_{n,\alpha}$.

Proposition 1.21. Soit $\eta \geq 0$. Sous (C'1) – (C'6), si de plus :

- il existe une loi P telle que (\mathbb{A}, d) est quasi-compact,
- V est continue sur $[\eta, 1 - \eta]$,

alors

$$\lim_{n \rightarrow \infty} \left(\sup_{\alpha \in [\eta, 1 - \eta]} d(\tilde{A}_{n,\alpha}, \tilde{A}_\alpha) \right)^* = 0 \text{ p.s.}$$

En conclusion de ce paragraphe, faisons quelques remarques sur ces généralisations sur \mathbb{R}^d de la notion de quantile. D'une part l'absence de forme intrinsèque des quantiles impose à l'utilisateur de choisir une classe \mathbb{A} , ainsi le quantile d'une même loi pourra être une sphère, un cube, ou une certaine forme si la classe \mathbb{A} est l'ensemble des sphères, des cubes ou autre chose. D'autre part, les conditions garantissant la consistance des quantiles demandent beaucoup de régularité et notamment l'existence d'une densité par rapport à la mesure de Lebesgue. En somme, la géométrie des objets sur \mathbb{R}^d n'est pas exploitée dans la définition même du quantile. Enfin, les résultats

précèdent ne fournissent ni vitesse de convergence ni de représentation du type Bahadur-Kiefer.

Dans cette thèse, l'objectif principal est de définir et d'étudier une notion de quantile sur \mathbb{R}^d qui soit intrinsèque, facile à visualiser et interprétable au point d'en déduire différentes procédures de statistique non paramétrique. Le second objectif est d'obtenir les hypothèses les plus faibles possible garantissant des résultats de convergences aussi précis que ceux cités dans la section 1.1 sur les quantiles réelles.

1.2.2 Ensembles de niveau et excès de masse

Un quantile sur \mathbb{R}^d avec $d \geq 1$ peut être appréhender comme une région ou ensemble de points de \mathbb{R}^d avec une forte concentration de masse. Pour les lois de probabilité absolument continues par rapport à la mesure de Lebesgue, on peut lier les quantiles aux ensembles de niveaux de la fonction densité, pour cela on parle de « level set », le terme anglo-saxon pour « ensemble de niveau ».

Cette notion est très fortement liée à celle d'ensemble de volume minimal présentée dans le paragraphe précédent, ainsi qu'à la notion d'excès de masse définie dans la suite. Pour des lois à densité par rapport à la mesure de Lebesgue λ_d , on s'intéresse au niveau de la fonction densité notée f .

Définition 1.22. Soit $\mu > 0$. on définit l'ensemble μ -niveau de f

$$L_\mu = \{x \in \mathbb{R}^d : f(x) \geq \mu\}.$$

Si μ est assez petit, son volume $\lambda_d(L_\mu)$ est non nul. Sous hypothèses de régularité de f , la frontière de L_μ forme un contour qui localise les zones de fortes probabilité, un des objectifs poursuivis dans cette thèse.

Hypothèses. On fixe $\mu > 0$, on pose $v_\mu = \lambda_d(L_\mu)$ et on suppose que

(C₁^{*}) $v_\mu > 0$ et L_μ est convexe

Remarque. Soit X une variable aléatoire de loi P de densité f , on sait que sous (C₁^{*}) on a

$$\sup_{x \in \mathbb{R}^d} |f(x)| \geq \mu$$

ainsi que

$$p_\mu := P(L_\mu) \geq \mu v_\mu > 0. \quad (1.1)$$

L'ensemble de niveau vérifiant l'inégalité 1.1 est un très bon candidat pour un quantile de niveau p_μ . De plus les contours de L_μ présentent l'avantage d'être

emboîtés, et définis de façon objective et figée, sans avoir besoin de choisir un observateur comme c'est le cas dans les Chapitres 2 et 3. Néanmoins, on exige l'existence d'une densité par rapport à la mesure de Lebesgue, et même plus, et surtout, pour estimer L_μ par plug-in il faut un contrôle du nombre de composantes connexes de L_μ ou bien connaître la régularité de f pour calibrer les choix des paramètres de lissage de l'estimateur, comme une fenêtre et un noyau, ou bien encore, comme on l'évoque ci-dessous, spécifier une petite classe \mathcal{A} (faible entropie métrique par exemple) où chercher L_μ . De notre côté, nos quantiles existent sous hypothèses très faibles sur la loi elle-même, n'exigeant pas l'existence d'une densité, et possédant de bons estimateurs non paramétrique.

Tout d'abord on suppose que X est une variable aléatoire de loi P de densité f et \mathcal{A} une classe d'ensembles P -mesurables, ensuite on définit la fonction d'excès de masse

Définition 1.23 (Fonction excès de masse). *Soit $\mu > 0$. on appelle fonction d'excès de masse, la fonction*

$$\begin{aligned} \mathcal{E}_\mu : \mathcal{A} &\longrightarrow \mathbb{R} \\ A &\longmapsto P(A) - \mu\lambda_d(A) \end{aligned}$$

La fonction \mathcal{E}_μ , nous permet de définir l'ensemble excès de masse et son excès de masse associé

Définition 1.24. *Soit $\mu > 0$. on définit l'ensemble μ -excès de masse*

$$E_\mu = \arg \max_{A \in \mathcal{A}} \{P(A) - \mu\lambda_d(A)\}$$

et le μ -excès de masse associé à E_μ

$$e_\mu = \max_{A \in \mathcal{A}} \{P(A) - \mu\lambda_d(A)\}.$$

Sous la condition (C_1^*) , si $L_\mu \in \mathcal{A}$ on a les égalités suivantes

$$\begin{aligned} L_\mu &= \arg \max_{A \in \mathcal{A}} \{P(A) - \mu\lambda_d(A)\} \\ &= \arg \min_{A \in \mathcal{A}} \{\lambda_d(A) : P(A) \geq p_\mu\}. \end{aligned}$$

De plus, en notant

$$\tilde{A}_\alpha = \arg \min_{A \in \mathcal{A}} \{\lambda_d(A) : \mathbb{P}(X \leq t) \geq \alpha\}$$

l'ensemble qui réalise la minimisation de $V(\alpha)$, et

$$M_v = \arg \max_{A \in \mathcal{A}} \{P(A) : \lambda_d(A) \leq v\}$$

l'ensemble de plus grande probabilité sous contrainte de volume v , on obtient

$$L_\mu = E_\mu = \tilde{A}_{p_\mu} = M_v. \quad (1.2)$$

À la différence avec L_μ , ces ensemble n'exigent pas l'existence d'une densité f et peuvent se contrôler par concentration de la mesure, comme arg min de processus empirique avec dérive.

Ensuite, remarquant qu'on a aussi

$$e_\mu = p_\mu - \mu v_\mu > 0$$

et

$$p_\mu > \mu v_\mu > 0$$

Notation 1.25. on notera L_{μ_α} l'ensemble de niveau de volume minimal et de masse α , i.e.

$$L_{\mu_\alpha} = \arg \min_{A \in \mathcal{A}} \{\lambda_d(A) : P(A) \geq \alpha\}$$

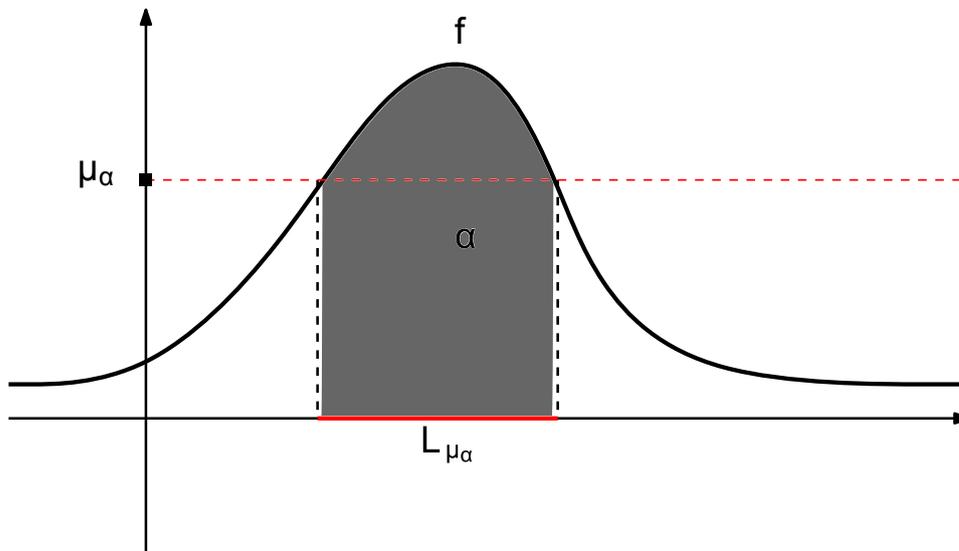


FIGURE 1.1 – μ_α choisie tel que $P(L_{\mu_\alpha}) = \alpha$.

De la même manière qu'on a défini E_μ , M_v et \tilde{A}_α , on définit leurs versions empiriques associées $E_{\mu,n}$, $M_{v,n}$ et $\tilde{A}_{\alpha,n}$. La consistance et la vitesse d'approximation en différence symétrique de ces estimateurs ont été étudiés par Polonik dans [Pol95, Pol97a, Pol97b]. sous diverses hypothèses de régularité de f et sur la classe \mathcal{A} , il est montré, pour un intervalle fermé I

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sup_{\mu \in I} P(L_\mu \Delta L_n) \geq d_n \right) = 0$$

où dans les cas les plus réguliers, $d_n = n^{-1/3}(\log n)^{1/3}$ pour l'estimateur $L_n = E_{\mu,n}$ et $d_n = n^{-1/3+\varepsilon}(\log n)^{1/3+\varepsilon}$ pour l'estimateur $L_n = \tilde{A}_{\alpha,n}$.

Ce phénomène de concentration est décrit plus précisément à l'échelle $n^{1/3}$ dans [BE12] en distance de Hausdorff, à l'aide d'un théorème limite central joint en E_μ , M_v et \tilde{A}_α . Pour se faire, il est supposé que (\mathbf{C}_1^*) est vérifié et que \mathcal{A} est asymptotiquement compatible avec un changement d'échelle au voisinage de la frontière ∂L_μ , de plus, une notion liée à la dérivée de fonctions à valeur ensemble nécessaire à la description limite du processus empirique de surface, autour de ∂L_μ . Le comportement limite en loi jointe des trois ensembles

$$n^{-1/3}(L_\mu \Delta E_{\mu,n}, L_\mu \Delta \tilde{A}_{\alpha,n}, L_\mu \Delta M_{v,n})$$

montre d'un coté que les estimateurs $M_{v,n}$ et $\tilde{A}_{\alpha,n}$ sont équivalents, et d'un autre coté que $E_{\mu,n}$ se comporte différemment, sans pour autant qu'on puisse les comparés, tout dépend de la loi inconnue P . Plus précisément, modulo une transformation bijective de \mathbb{R}^d vers l'espace des cylindres de ∂L_μ , ces ensembles aléatoires convergent en loi vers l'argmax d'un mouvement Brownien indexé par une classe d'ensembles \mathcal{B} dépendant de (\mathcal{A}, L_μ) , avec une dérive dépendant de (P, f, L_μ) . L'argmax est pris sur \mathcal{B} pour $L_\mu \Delta E_{\mu,n}$ et un sous-ensemble de \mathcal{B} pour $L_\mu \Delta \tilde{A}_{\alpha,n}$ et $L_\mu \Delta M_{v,n}$.

Les écarts avec l'ensemble cible L_μ sont décrits dans un repère intrinsèque à celui-ci, à l'aide des normales à sa surface ∂L_μ . C'est encore une fois un point de vue objectif, mais qui exige la convexité, la régularité et connaissance de \mathcal{A} compatible avec L_μ .

Les quantiles définis aux chapitres 2 et 3 mesurent l'écart à l'ensemble cible non pas le long de normales par la distance de Hausdorff mais le long de directions issues d'un observateur extrinsèque, O .

1.2.3 Fonctions de profondeur

Les fonctions de profondeur sont des outils d'inférence en statistique non paramétrique, d'une certaine manière elles permettent de définir un ordre spatial en «partance d'un centre», par rapport à une mesure μ sur \mathbb{R}^d .

Pour une variable aléatoire X de loi $P = \mathbb{P}_X$, la profondeur d'un point $x \in \mathbb{R}^d$ par rapport à la mesure de probabilité P est notée généralement $D(x, P)$ en référence au mot anglais « depth » qui signifie profondeur. L'ordre en partance du centre suggère l'existence d'un point central ainsi que des points dans son voisinage avec une profondeur plus petite que le centre mais plus grande que des points éloignés.

Comme suggéré par R. Liu [Liu88] voici les propriétés désirables pour la fonction de profondeur :

- i. **Invariance affine** : la profondeur d'un point ne doit pas dépendre du système de coordonnées.
- ii. **Maximalité au centre** : pour une mesure ayant un unique point central, la fonction profondeur doit atteindre son maximum en ce point.
- iii. **Décroissance relativement au centre** : en s'éloignant de manière radiale du point le plus profond, la profondeur doit décroître de façon continue.
- iv. **Nullité à l'infini** : la profondeur d'un point doit tendre vers zéro quand sa norme tend vers l'infini.

Voici maintenant une définition formelle ([ZS00a],[ZS00c],[ZS00b]) de la fonction profondeur relativement à une mesure de probabilité P . On note \mathcal{M} l'ensemble des mesures de probabilité sur \mathbb{R}^d et \mathcal{M}_C l'ensemble des mesures de probabilité sur \mathbb{R}^d admettant un centre C :

Définition 1.26. *Une application bornée $D : \mathbb{R}^d \times \mathcal{M} \rightarrow \mathbb{R}^+$ est une fonction de profondeur si :*

- i. **(Invariance affine)** $\forall x \in \mathbb{R}^d, \forall P \in \mathcal{M}, \forall A \in Gl_d(\mathbb{R}), \forall b \in \mathbb{R}^d :$

$$D(AX + b, \mathbb{P}_{AX+b}) = D(x, \mathbb{P}_X) \text{ pour tout vecteur aléatoire } X \text{ dans } \mathbb{R}^d.$$

- ii. **(Maximalité au centre)** $\forall P \in \mathcal{M}_C :$

$$D(C, P) = \sup_{x \in \mathbb{R}^d} D(x, P).$$

- iii. **(Décroissance relativement au centre)** $\forall x \in \mathbb{R}^d, \forall P \in \mathcal{M}_C, \forall \alpha \in [0, 1] :$

$$D(x, P) \geq D(C + \alpha(x - C), P).$$

- iv. **(Nullité à l'infinie)** $\forall x \in \mathbb{R}^d, \forall P \in \mathcal{M}_C :$

$$\lim_{\|x\| \rightarrow +\infty} D(x, P) = 0$$

Étant donné un échantillon X_1, \dots, X_n de loi P , la version empirique de $D(x, P)$ est notée $D_n(x, P_n)$ où P_n est la mesure empirique de P .

Cette définition de la profondeur est généralement utilisée avec un centre C correspondant à une symétrie de la loi. Notons \mathcal{H} l'ensemble des demi-espaces fermés de \mathbb{R}^d .

Définition. Soient $C \in \mathbb{R}^d$ et μ une mesure de probabilité sur \mathbb{R}^d , on dit que μ admet une :

- symétrie centrale autour de C (C -symétrie) si

$$X - C \stackrel{\mathcal{L}}{\sim} C - X$$

- symétrie angulaire autour de C (A -symétrie) si

$$(X - C) / \|X - C\| \stackrel{\mathcal{L}}{\sim} C - X / \|C - X\|$$

- symétrie de demi-espace autour de C (H -symétrie) si

$$\forall H \in \mathcal{H} \text{ avec } C \in H : \mu(X \in H) \geq 1/2.$$

Remarque 1.27.

$$C\text{-symétrie} \Rightarrow A\text{-symétrie} \Rightarrow H\text{-symétrie}.$$

Définition (Profondeur de Tukey). Soit P une mesure de probabilité, l'application

$$D_P^{\mathcal{H}} : \mathbb{R}^d \longrightarrow \mathbb{R}_+ \\ x \longmapsto \inf \{P(H) : x \in H \in \mathcal{H}\}$$

est appelée profondeur de Tukey pour la notion de H -symétrie.

Remarque. La version empirique de $D_P^{\mathcal{H}}$ notée $D_n^{\mathcal{H}}$ correspond au nombre minimal de points de l'échantillon se trouvant d'un côté du demi-plan passant par le point x .

Le théorème suivant montre la consistance uniforme du processus de profondeur de Tukey.

Théorème (Donoho et Gasko, 1992 [DG92]). Soit P une mesure de probabilité à densité bornée, continue et positive et admettant une centre de H -symétrie, on a

$$\lim_{n \rightarrow +\infty} \sup_{x \in \mathbb{R}} |D_n^{\mathcal{H}} - D_P^{\mathcal{H}}| \stackrel{p.s.}{=} 0.$$

Une autre notion de profondeur, principalement étudiée par R. Liu dans [Liu88] est la suivante.

Définition (Profondeur simpliciale). *Soit P une mesure de probabilité et X_1, X_2, \dots, X_{d+1} $d + 1$ vecteurs aléatoires indépendants de même loi P . En notant*

$$S[x_1, x_2, \dots, x_{d+1}] = \left\{ x = \sum_{i=1}^{d+1} \alpha_i x_i : \alpha_i \geq 0, \sum_{i=1}^n \alpha_i = 1 \right\},$$

l'application

$$\begin{aligned} D_P^S : \mathbb{R}^d &\longrightarrow \mathbb{R}_+ \\ x &\longmapsto P(x \in S[X_1, X_2, \dots, X_{d+1}]) \end{aligned}$$

est appelée profondeur simpliciale pour la notion de A -symétrie.

Remarque. $S[x_1, x_2, \dots, x_{d+1}]$ est l'ensemble de toutes les combinaisons convexes de x_1, x_2, \dots, x_{d+1} . Géométriquement $S[x_1, x_2, \dots, x_{d+1}]$ correspond au simplexe en dimension d de sommet x_1, x_2, \dots, x_{d+1} .

Définition (Profondeur simpliciale empirique). *Soit P une mesure de probabilité et X_1, X_2, \dots, X_n n vecteurs aléatoires indépendants de même loi P sur \mathbb{R}^d , l'application*

$$\begin{aligned} D_n^S : \mathbb{R}^d &\longrightarrow \mathbb{R}_+ \\ x &\longmapsto \frac{1}{C_n^{d+1}} \sum_{1 \leq i_1 < \dots < i_{d+1} \leq n} \mathbb{1}_{x \in S[X_{i_1}, \dots, X_{i_{d+1}}]} \end{aligned}$$

est appelée profondeur aléatoire simpliciale empirique.

Maintenant, voici le théorème de consistance uniforme pour la profondeur D_P^S

Théorème (Liu, 1988). *Soit P une mesure de probabilité à densité bornée, on a*

$$\lim_{n \rightarrow \infty} \left\| D_n^S - D_P^S \right\| \stackrel{p.s.}{=} 0.$$

Voici un autre théorème qui montre la consistance de la notion de centre pour D_P^S

Théorème (Liu, 1988). *Soit P une mesure de probabilité à densité bornée et admettant une centre de A -symétrie, on pose*

$$\begin{aligned} \rho &= \arg \max_{x \in \mathbb{R}^d} D_P^S(x) \\ \rho_n &= \arg \max_{x \in \mathbb{R}^d} D_n^S(x) \end{aligned}$$

si la densité de P ne s'annule pas au voisinage de ρ , alors

$$\lim_{n \rightarrow \infty} \|\rho_n - \rho\| \stackrel{p.s.}{=} 0.$$

La littérature statistique sur les fonctions de profondeur est très vaste. Sans donner de liste exhaustive on renvoie à [LS93] pour « Majority depth », [ZS00a] pour « Projection depth », « Mahalanobis depth » et « Oja depth », [KM97] pour « Zonoid depth », [RK11] « Weighted means depth » et « Geometrical depth », [Bar76] pour « Convex-hull peeling depth ». Il ne semble pas exister de résultats de convergence précis de ces fonctions de profondeur à l'exception de quelques travaux comme celui de Arcones et al. dans [ACZ06].

Afin de faire le lien entre la fonction de profondeur et les quantiles présentés au paragraphe précédent, on a besoin d'introduire les objets suivants.

Définition 1.28. Soit $\gamma \geq 0$, P une mesure de probabilité et D une P -profondeur.

– La région intérieure de P -profondeur D au moins γ est définie par

$$I(\gamma, D, P) = \{x \in \mathbb{R}^d, : D(x, P) \geq \gamma\}.$$

– Le γ -contour de P -profondeur D est défini par

$$\partial I(\gamma, D, P) = I(\gamma, D, P) \setminus \text{int}(I(\gamma, D, P))$$

– La plus grande P -profondeur associée aux régions de probabilité au moins $\alpha \in (0, 1)$ est définie par

$$\gamma_\alpha = \sup\{\gamma \geq 0 : P(I(\gamma, D, P)) \geq \alpha\}$$

Remarque 1.29. Tout d'abord remarquons que $I(0, D, P) = \mathbb{R}^d$, ensuite pour $Y = D(X, P)$, et F_Y sa fonction de répartition on a

– Si F_Y est continue alors

$$P(I(\gamma, D, P)) = 1 - F_Y(\gamma).$$

– Si F_Y est continue et strictement croissante alors on a en plus

$$\gamma_\alpha = F_Y^{-1}(1 - \alpha)$$

– Pour $\gamma \uparrow \gamma_\alpha$ l'ensemble $I(\gamma, D, P) \downarrow I(\gamma_\alpha, D, P)$ ce qui implique que

$$\lim_{\gamma \uparrow \gamma_\alpha} P(I(\gamma, D, P)) = P(I(\gamma_\alpha, D, P)) \geq \alpha$$

La définition suivante introduit un objet jouant le rôle d'un « quantile spatial »

Définition 1.30. Soit $\alpha \in (0, 1)$, la plus petite région intérieure de P -profondeur D de probabilité d'au moins α est définie par

$$C(\alpha, D, P) = I(\gamma_\alpha, D, P)$$

Lien avec le quantile généralisé de Einmahl-Mason [Ser02]. Dans cette partie on s'intéresse au lien entre les régions intérieure $I(\gamma, D, P)$ et la fonction quantile généralisée introduite dans la section 1.2.1.

Hypothèses 1.31.

- (A₁') *La mesure de probabilité P admet une densité f sur \mathbb{R}^d avec pour tout $x \in \text{Supp}(P)$, $f(x) > 0$.*
- (A₂') *La profondeur $D(\cdot, P)$ vérifie les deux conditions suivantes*
 - *l'application $D(\cdot, P)$ est continue avec $D(x, P) = 0$ pour tout $x \notin \text{Supp}(P)$ et $\lim_{\|x\| \rightarrow \infty} D(x, P) = 0$.*
 - *l'ensemble $\{x : D(x, P) = \gamma\}$ est non vide pour $0 < \gamma < \gamma^* = \sup_{x \in \mathbb{R}^d} D(x, P)$.*
- (A₃') *Pour \mathbb{A} la classe d'ensemble et $\mu : \mathbb{A} \rightarrow \mathbb{R}$ la fonction de taille, on suppose que*
 - $\mathbb{A} = \{I(\gamma, D, P) : 0 < \gamma < \gamma^*\}$.
 - *Pour tout $0 < \gamma < \gamma^*$ on suppose que $\mu(I(\gamma, D, P))$ est fini, strictement croissante en γ , et continûment dérivable en γ .*

Sous les conditions (A₁'), (A₂') et (A₃') on a

$$\begin{aligned} V(\alpha) &= \inf_{\gamma} \{ \mu(I(\gamma, D, P)) : P(I(\gamma, D, P)) \geq \alpha \} \quad 0 < \alpha < 1 \\ &= \inf_{\gamma} \{ \mu(I(\gamma, D, P)) : 0 < \gamma \leq \gamma_{\alpha} \} \quad 0 < \alpha < 1 \\ &= \mu(I(\gamma_{\alpha}, D, P)) \quad 0 < \alpha < 1 \\ &= \mu(C(\alpha, D, P)) \quad 0 < \alpha < 1. \end{aligned}$$

De la même manière on peut définir V_n la version empirique de V .

Exemple 1.32. *Par exemple si on choisit μ telle que $\mu(I(\gamma, D, P)) = \gamma$ alors $V(\alpha) = I(\gamma_{\alpha}) = F_Y^{-1}(1 - \alpha)$ est le quantile de niveau $1 - \alpha$ de $Y = D(X, P)$.*

Sous la condition (A₃') on sait que V est différentiable donc en posant $g(\alpha) = 1/V'(\alpha)$ et en notant

$$\dot{\beta}_n(\alpha) = n^{1/2}(V_n(\alpha) - V(\alpha)) \quad 0 < \alpha < 1,$$

le processus d'intérêt est $g(\alpha)\dot{\beta}_n(\alpha)$. Dans l'exemple 1.32 $V'(\alpha) = -f_Y \circ F_Y^{-1}(1 - \alpha) =: -h_Y(1 - \alpha)$ où h_Y est communément appelée « densité quantile » de Y et dans ce cas on a donc $g(\alpha) = -1/h_Y(1 - \alpha)$.

Soit B_P un pont brownien i.e. un processus gaussien sur $(\mathcal{C}[0, 1], \|\cdot\|_{\infty})$, centré et de covariance $\text{cov}(B_P(s), B_P(t)) = s \wedge t - st$ pour $s, t \in [0, 1]$.

Théorème 1.33. *Sous les conditions (A'_1) , (A'_2) et (A'_3) , on peut construire sur un même espace probabilisé (Ω, \mathcal{A}, P) une suite de versions de X_n , $n \geq 1$ et un pont brownien B_P tel que pour tout $0 < a < b < 1$ on a*

$$\lim_{n \rightarrow \infty} \sup_{t \in [a, b]} |g(\alpha) \dot{\beta}_n(t) - B_P(t)| = 0 \text{ p.s.}$$

Enfin, parmi les travaux méthodologique gravitant autour de la notion de profondeur [HW97], [Kim00], [Nol99], [Nol91], [VZ00], etc ...

Pour conclure ce bref passage en revue, nous attirons l'attention sur le fait que ce sont les fonctions de profondeur qui ont servi à définir des quantiles. À l'inverse, dans chapitre 4, nous définissons une fonction de profondeur qui découle directement des surfaces quantiles du chapitre 2. Cette nouvelle notion de profondeur est un champs vectoriel qui outre la profondeur de chaque points, indique la direction pointant vers la masse.

De plus, l'exigence de l'existence d'un centre et d'une symétrie et très restrictive dans a définition 1.26 excluant notamment les lois multimodales telles que les mélanges. En outre la question de l'existence d'un centre ou d'une symétrie est intéressante en elle même, ce qui renforce le besoin de disposer d'un outils définis sous hypothèses plus faible, tel que les surfaces quantiles étudiés au chapitre 2.

1.2.4 Régression quantile

Dans cette section on s'intéresse à une notion de régression quantile dans \mathbb{R}^d , celle étudiée dans [HPŠ10] à partir d'une généralisation du quantile réelle en un quantile directionnel. En effet dans certains cas particuliers, il y a un lien avec les surfaces quantiles du chapitre 2, puisque les demi-espaces de Tukey apparaissent explicitement dans les deux cas, mais sous une indexation différente.

Soit X une variable aléatoire réelle de loi \mathbb{P}_X , pour $\alpha \in (0, 1)$ le α -quantile de P_X peut être défini comme la valeur pour laquelle le minimum de la fonction $a \mapsto \mathbb{E} [(X - a)(\alpha - \mathbb{1}_{X < a})]$ est atteint, i.e.

$$F_X^{-1}(\alpha) = \arg \min_{a \in \mathbb{R}} \{ \mathbb{E} [(X - a)(\alpha - \mathbb{1}_{X < a})] \}.$$

Soit Z une variable aléatoire de $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P)$ avec $P = \mathbb{P}_Z$. Supposons que P admette une densité avec un support connexe sur \mathbb{R}^d et que ses moments d'ordre 1 sont finis.

Pour $\alpha \in (0, 1)$ et $u \in \mathbb{S}_{d-1}$ le $(\alpha \cdot u)$ -quantile est défini dans [HPŠ10] comme l'hyperplan $\mathbf{H}(a(u), c(u)) := \{z \in \mathbb{R}^d : \langle c(u), z \rangle = a(u)\}$ où le couple $(a(u), c(u))$ est tel que

$$c(u) = \arg \min_{\substack{c \in \mathbb{R}^d \\ \langle c, u \rangle = 1}} \left\{ \mathbb{E} \left(\left(\langle c, Z \rangle - F_{\langle c, Z \rangle}^{-1}(\alpha) \right) \left(\alpha - \mathbb{1}_{\langle c, Z \rangle < F_{\langle c, Z \rangle}^{-1}(\alpha)} \right) \right) \right\}$$

$$a(u) = F_{\langle c(u), Z \rangle}^{-1}(\alpha).$$

Le vecteur $c(u)$ peut être réécrit sous la forme suivante

$$c(u) = \arg \min_{\substack{c \in \mathbb{R}^d \\ \langle c, u \rangle = 1}} \left\{ (1 - \alpha) \int_{-\infty}^{F_{\langle c, Z \rangle}^{-1}(\alpha)} (-x + F_{\langle c, Z \rangle}^{-1}(\alpha)) P_{\langle c, Z \rangle}(dx) \right. \\ \left. + \alpha \int_{F_{\langle c, Z \rangle}^{-1}(\alpha)}^{+\infty} (x - F_{\langle c, Z \rangle}^{-1}(\alpha)) P_{\langle c, Z \rangle}(dx) \right\}.$$

Afin de faire un lien avec nos travaux, notamment ceux du chapitre 2, on a besoin d'introduire la notation $H(0_d, v, y)$: le demi-espace vue de 0_d dans la direction v à distance y .

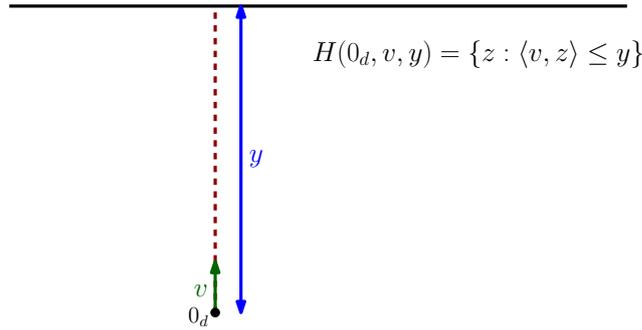
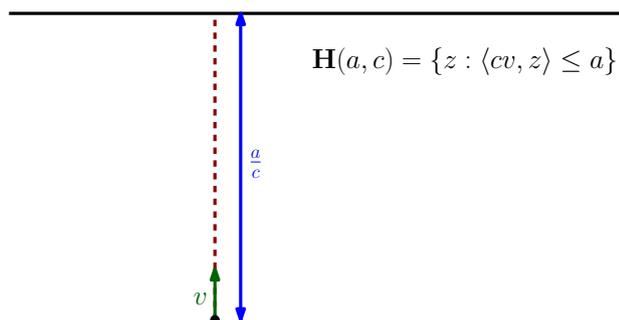


FIGURE 1.2 – $H(0_d, v, y)$

En posant $c(u) = \|c(u)\|_2 v(u)$ avec $v(u) \in \mathbb{S}_{d-1}$, et en remarquant que $F_{\langle c(u), Z \rangle}^{-1}(\alpha) = \|c(u)\|_2 F_{\langle v(u), Z \rangle}^{-1}(\alpha)$, on a

$$\begin{aligned} \mathbf{H}(a(u), c(u)) &= \{z \in \mathbb{R}^d : \langle c(u), z \rangle = a(u)\} \\ &= \{z \in \mathbb{R}^d : \|c(u)\|_2 \langle v(u), z \rangle = \|c(u)\|_2 F_{\langle v(u), Z \rangle}^{-1}(\alpha)\} \\ &= \{z \in \mathbb{R}^d : \langle v(u), z \rangle = F_{\langle v(u), Z \rangle}^{-1}(\alpha)\} \\ &= \partial H(0_d, v(u), F_{\langle v(u), Z \rangle}^{-1}(\alpha)). \end{aligned}$$

FIGURE 1.3 – $\mathbf{H}(a(u), c(u))$

Remarque 1.34. Avec les définitions plus précises du chapitre 2 (les définitions 2.1 et 2.4), on a donc

$$\mathbf{H}(a(u), c(u)) = \partial H(0_d, v(u), F_{\langle v(u), Z \rangle}^{-1}(\alpha))$$

de plus par la Proposition 2.3 on a pour tout $O, O' \in \mathbb{R}^d$ on a

$$H(O, v, F_{\langle v(u), Z - O \rangle}^{-1}(\alpha)) = H(O', v(u), F_{\langle v(u), Z - O' \rangle}^{-1}(\alpha)) = H(v(u), \alpha)$$

alors

$$\mathbf{H}(a(u), c(u)) = \partial H(v(u), \alpha).$$

1.2.5 Trimming multivarié et ensembles de « Tukey-Nolan »

Dans cette partie on s'intéresse à la notion du trimming multivarié d'une mesure de probabilité P introduite dans [Nol92]. Intuitivement cela correspond à une généralisation consistante sur \mathbb{R}^d de l'intervalle quantile $[F^{-1}(\alpha/2), F^{-1}(1 - \alpha/2)]$ pour $\alpha < 1/2$. Afin de généraliser l'intervalle quantile sur \mathbb{R}^d , D. Nolan introduit les ensemble

$$C_\alpha = \bigcap_{\substack{H \text{ demi-espace} \\ P(H) \geq 1 - \alpha}} H$$

ainsi que leur version empirique

$$C_{n,\alpha} = \bigcap_{\substack{H \text{ demi-espace} \\ P_n(H) \geq 1 - \alpha}} H.$$

Afin de quantifier la vitesse de convergence de $C_{n,\alpha}$ vers C_α , on définit pour tout $u \in \mathbb{S}_{d-1}$ les grandeurs suivantes

$$r_\alpha(u) = \inf\{r \geq 0 : ru \notin C_\alpha\}$$

et

$$r_{n,\alpha}(u) = \inf\{r \geq 0 : ru \notin C_{n,\alpha}\}.$$

Sous les hypothèses suivantes, pour α fixé, le théorème principale dans [Nol92], affirme que le processus $\sqrt{n}(r_{n,\alpha} - r_\alpha)$ converge vers un processus gaussien indexé par \mathbb{S}_{d-1} .

Hypothèses 1.35.

- $\text{int}(C_\alpha) \neq \emptyset$.
- la frontière ∂C_α est régulière i.e. pour tout $u \in \mathbb{S}_{d-1}$ il existe un unique plan tangent à C_α en $x := r_\alpha(u)u$, et on note $v_\alpha(u)$ son vecteur normal.
- $P_{v_\alpha(u)}$ la projection de P sur $v_\alpha(u)$ admet une densité par rapport à la mesure de Lebesgue $f_{v_\alpha(u)}$ avec $f_{v_\alpha(u)} > 0$ et continue en u dans \mathbb{S}_{d-1} et en x dans un voisinage de $r_\alpha(u)\langle v_\alpha(u), u \rangle$.
- $\|r_{n,\alpha} - r_\alpha\|_{\mathbb{S}_{d-1}} \rightarrow 0$ p.s.

La condition suffisante de régularité de la frontière proposée dans [Nol92] consiste en la continuité locale en α en distance de Hausdorff pour C_α . Au chapitre 2 sous des hypothèses plus faibles et presque minimales nos résultats impliquent que cette convergence p.s. est uniforme en α .

Théorème 1.36. *Sous les hypothèses 1.35, le processus $\sqrt{n}(r_n - r_\alpha)$ converge vers \mathbb{G} un processus gaussien indexé par \mathbb{S}_{d-1} .*

Ce résultat est impliqué par nos résultats du chapitre 2, sous des hypothèses plus faibles, de plus la vitesse de convergence peut être facilement explicitée. Néanmoins, on tient à rappeler que l'ensemble C_α n'est pas toujours non vide, et n'est pas l'objet central de notre étude.

En lien avec [Nol92] on peut citer [KM12] qui introduit pour $u \in \mathbb{S}_{d-1}$ et $\alpha \in (0, 1/2)$ la réécriture suivante

$$Q_X(u, \alpha) = F_{(X,u)}^{-1}(1 - \alpha)$$

et

$$D(\alpha) = \bigcap_{u \in \mathbb{S}_{d-1}} H(u, Q_X(u, \alpha))$$

où $H(u, y) = \{x \in \mathbb{R}^d : \langle x, u \rangle \leq y\}$. Sous les hypothèse de [KM12] on a les correspondances suivantes avec [Nol92]

$$r_\alpha(u) = Q_X(u, \alpha)$$

et

$$C_\alpha = D(\alpha).$$

1.2.6 Autres

Il est certain que la généralisation du quantile univarié sur \mathbb{R}^d a fait l'objet d'une quantité impressionnante de publications en statistique allant du plus théorique dont certaines ont été rappelées dans ce chapitre, au plus appliquées, dont certaines très spécifiques à la dimension 2 ou 3. Ce chapitre n'a pas la prétention de donner une vision exhaustive des dites généralisations, mais de tout simplement présenter celles en lien avec le quantile spatial introduit dans cette thèse ou celles dont la généralisation semble naturelle et/ou ayant été une piste à creuser durant les années de recherche et d'élaboration de ce manuscrit. Néanmoins, dans cette toute dernière section, il est utile de mentionner d'autres références non traitées dans les sections précédentes, telles que la notion de quantile géométrique dans [Cha96], ou le M -quantile obtenue par M -estimation globale dans [Kol97], [BC88] ou encore le quantile multivarié via les gradients dans [HNO92] ainsi que [Wei08] qui propose une extension du quantile usuel à la dimension 2 avec un quantile conditionnel à la direction bien adapté à des données biologiques. Dans un autre registre on peut aussi référer à la notion des k -means introduite par Cuesta-Albertos, Gordaliza et Matrán dans [CAGM97], [CM88] et [CAF07].

Chapitre 2

Surfaces quantiles via les demi-espaces

Dans ce chapitre on introduit et on étudie une généralisation spatiale sur \mathbb{R}^d du quantile réel usuel sous la forme d'une surface quantile. L'aspect original est que l'interprétation de ces surfaces plongées dans \mathbb{R}^d est de nature subjective. En effet ces surfaces dépendent d'un point d'observation que l'on veut ensuite bouger pour capter l'information statistique. Par contre les théorèmes limites ne dépendent pas de ce point. Cela ne sera plus le cas dans la situation encore plus générale des quantiles à travers des formes φ du Chapitre 3.

Dans cette première généralisation, les formes φ correspondent à des demi-espaces orthogonaux à la direction d'observation et les résultats ne dépendent pas de O . En bougeant O la description des quantiles spatiaux change d'aspect et c'est justement ce qui permet des interprétations statistiques pertinentes et automatisables, pourvu que des caractérisations précises ressortent, et qu'une classification des situations ait été menée au préalable. Dans ce travail cette démarche n'est pas menée au bout, seuls les objets eux-même et leurs convergences sont étudiés pleinement, c'est le préalable à toute utilisation en statistique et constitue l'innovation principale puisque de telles surfaces quantiles n'ont jamais été étudiées. Au chapitre 4 des conséquences sont tirées notamment pour des tests d'adéquation et de comparaison, via des distances entre surfaces quantiles à α fixé ou des champs de profondeur avec O mobile dans un compact. Le fait de bouger O pour observer la déformation des surfaces quantiles constitue une forme d'analyse des données qui ne perd aucune information et nécessite une étude plus approfondie pour en automatiser les interprétations. C'est tout l'esprit de la subjectivité retenue comme paradigme à la description d'une loi depuis un observateur. Pour aboutir à une méthodologie pratique il faut s'intéres-

ser à des questions moins ambitieuses que la description de la loi inconnue P dans \mathbb{R}^d , et accepter des projections en dimension plus faible, ou se servir de l'approche par surface quantile pour une quantification ou la recherche de zones modales, ou encore déterminer des directions principales. Ceci reste la perspective principale, et la motivation première, pour les résultats qui suivent.

L'objectif de ce chapitre est donc de dégager la nature des hypothèses permettant à des quantiles spatiaux d'être bien estimés, et aux surfaces associées d'être régulières, puis d'obtenir les théorèmes limite usuels, à savoir loi des grands nombre, loi du logarithme itéré, théorème limite central, principe d'invariance fort et approximation du type Bahadur-Kiefer, le tout uniformément en le niveau de quantile et la direction d'observation. L'aspect directionnel pose la question des corrélations entre les directions à travers les quantiles. Les réponses qui suivent dans ce chapitre précisent clairement les corrélations directionnelles et entre niveau de quantile, et donnent le moyen de construire explicitement, ou par simulations des processus Gaussiens limites, des régions de confiance autour des surfaces quantile. Le point clef est une approximation de type Bahadur-Kiefer dans \mathbb{R}^d et la façon dont elle est obtenue est originale même en dimension $d = 1$. Cette méthode repose intégralement sur la définition des quantiles et la compréhension fine du processus empirique indexé par des demi-espaces, de ses trajectoires et de ses oscillations. Au final, sous hypothèse de régularité inter-directionnelle, qui demande que les erreurs d'approximation des probabilités par densité directionnelle, de second ordre, soient uniformément contrôlées, le contrôle joint des surfaces quantiles par le processus empirique se fait à la même vitesse d'ordre $n^{-1/4}$ que sur \mathbb{R} .

2.1 Surfaces quantiles via les demi espaces

Dans cette section on donne la définition des surfaces quantiles via les demi espaces, ainsi que les objets qu'il lui sont reliés, ensuite on s'intéresse aux hypothèses garantissant la régularité d'une telle surface.

2.1.1 Définition des surfaces quantile

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $\{X_i\}_{i \in \mathbb{N}}$ une suite de variables aléatoires indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), P)$ et de même loi $P = \mathbb{P}^X$. Dans toute la suite, on munit \mathbb{R}^d du produit scalaire usuel, noté $\langle \cdot, \cdot \rangle$ et on note S_{d-1} la sphère unité de \mathbb{R}^d .

Définition 2.1. Le **demi-espace** dans la direction $u \in \mathbb{S}_{d-1}$ à distance $y \in \mathbb{R}$ du point $O \in \mathbb{R}^d$ est noté

$$H(O, u, y) = \{x \in \mathbb{R}^d : \langle x - O, u \rangle \leq y\}.$$

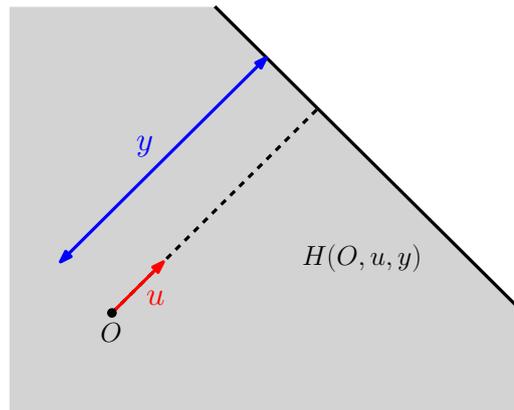


FIGURE 2.1 – Demi-espace $H(O, u, y)$

Définition 2.2. Le **quantile** de la loi P dans la direction $u \in \mathbb{S}_{d-1}$ de niveau $\alpha \in [1/2, 1)$ vu de $O \in \mathbb{R}^d$ est défini par

$$Y(O, u, \alpha) = \inf \{y \in \mathbb{R} : P(H(O, u, y)) \geq \alpha\}.$$

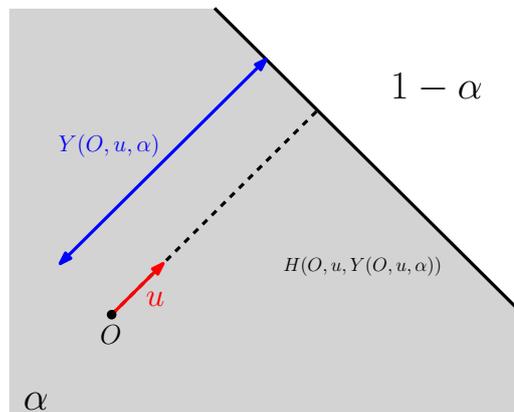


FIGURE 2.2 – Quantile $Y(O, u, \alpha)$

Proposition 2.3. Pour $O \in \mathbb{R}^d, u \in \mathbb{S}_{d-1}$ et $\alpha \in [1/2, 1)$ le demi-espace $H(O, u, Y(O, u, \alpha))$ ne dépend pas du choix de O .

Démonstration. Soit $y \in \mathbb{R}$, pour $O \neq O' \in \mathbb{R}^d$ on a

$$\begin{aligned}
 H(O, u, y) &= \{x \in \mathbb{R}^d : \langle x - O, u \rangle \leq y\} \\
 &= \{x \in \mathbb{R}^d : \langle x - O' + (O' - O), u \rangle \leq y\} \\
 &= \{x \in \mathbb{R}^d : \langle x - O', u \rangle \leq y - \langle O' - O, u \rangle\} \\
 &= H(O', u, y - \langle O' - O, u \rangle)
 \end{aligned} \tag{2.1}$$

et puisque

$$\begin{aligned}
 Y(O, u, \alpha) &= \inf \{y \in \mathbb{R} : P(H(O, u, y)) \geq \alpha\} \\
 &= \inf \{y \in \mathbb{R} : P(H(O', u, y - \langle O' - O, u \rangle)) \geq \alpha\} \\
 &= \inf_{z \in \mathbb{R}} \{z + \langle O' - O, u \rangle : P(H(O', u, z)) \geq \alpha\} \\
 &= \inf_{z \in \mathbb{R}} \{z : P(H(O', u, z)) \geq \alpha\} + \langle O' - O, u \rangle \\
 &= Y(O', u, \alpha) + \langle O' - O, u \rangle
 \end{aligned} \tag{2.2}$$

par les égalités 2.1 et 2.2 on a

$$\begin{aligned}
 H(O, u, Y(O, u, \alpha)) &= H(O', u, Y(O, u, \alpha) - \langle O' - O, u \rangle) \\
 &= H(O', u, Y(O', u, \alpha)).
 \end{aligned}$$

□

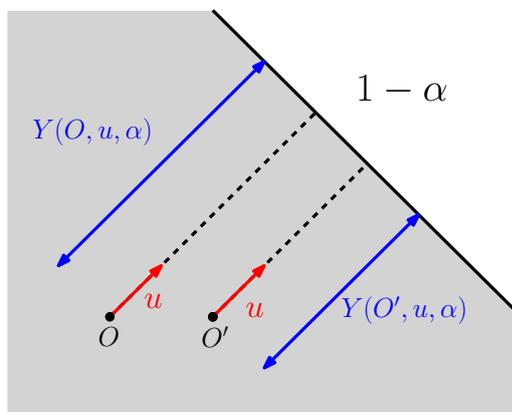
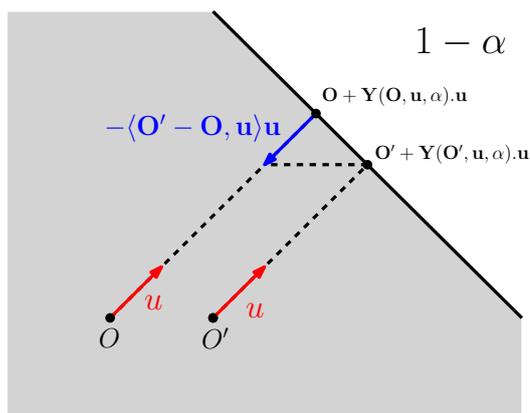


FIGURE 2.3 – Les scalaires $Y(O, u, \alpha)$ et $Y(O', u, \alpha)$

Puisque la dépendance en O disparaît, on simplifie la notation. De même selon les contextes on utilisera les indices O, u, α ou y , quitte à changer le nom d'un même objet, comme le processus empirique par exemple, selon les indices utilisés.

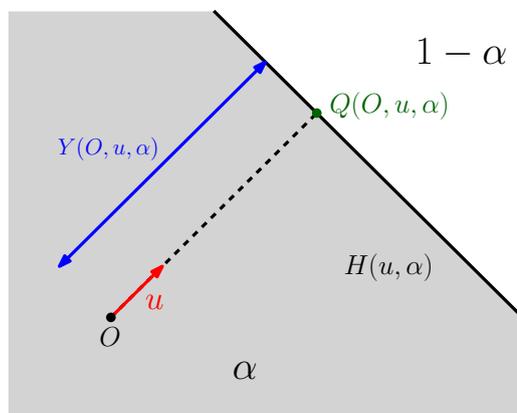
FIGURE 2.4 – Le passage entre $Y(O, u, \alpha)$ et $Y(O', u, \alpha)$

Définition 2.4. Le *demi-espace associé au quantile* $Y(O, u, \alpha)$ pour $u \in \mathbb{S}_{d-1}$, $\alpha \in [1/2, 1)$ et $O \in \mathbb{R}^d$ est noté

$$H(u, \alpha) := H(O, u, Y(O, u, \alpha)).$$

Définition 2.5. Le *point quantile* de la loi P dans la direction $u \in \mathbb{S}_{d-1}$ de niveau $\alpha \in [1/2, 1)$ vu de $O \in \mathbb{R}^d$ est défini par

$$Q(O, u, \alpha) = O + Y(O, u, \alpha) \cdot u.$$

FIGURE 2.5 – Le point $Q(O, u, \alpha)$ et le demi-espace $H(u, \alpha)$

Définition 2.6. L'*ensemble quantile* de la loi P dans la direction $u \in \mathbb{S}_{d-1}$ de niveau $\alpha \in [1/2, 1)$ vu de $O \in \mathbb{R}^d$ est défini par

$$Q_\alpha(O) = \{Q(O, u, \alpha) : u \in \mathbb{S}_{d-1}\}.$$

Définition 2.7. L'ensemble des demi-espaces de niveau $\alpha \in [1/2, 1)$ pour la loi P est défini par

$$\mathcal{H}_\alpha = \{H(u, \alpha) : u \in \mathbb{S}_{d-1}\}.$$

Définition 2.8. L'ensemble de Tukey-Nolan de niveau $\alpha \in (1/2, 1)$ est défini par

$$C_\alpha = \bigcap_{H \in \mathcal{H}_\alpha} H.$$

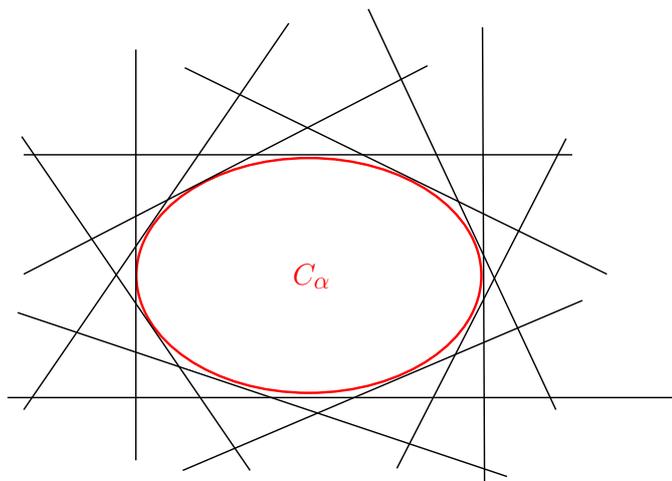


FIGURE 2.6 – L'ensemble de Tukey-Nolan C_α

On a décidé de nommer C_α « ensemble de Tukey-Nolan » en référence à J. Tukey qui a introduit dans [Tuk75] les demi-espaces de Tukey, ainsi qu'à D. Nolan dans [Nol92] qui a étudié cet objet dans le cas des C_α à bord régulier.

Proposition 2.9. Pour $\alpha \in (1/2, 1)$ si l'ensemble C_α est non vide alors il vérifie les propriétés suivantes

- i. C_α ne dépend pas du choix de $O \in \mathbb{R}^d$.
- ii. C_α est un ensemble convexe ou vide.
- iii. C_α est croissant en α (au sens de l'inclusion).

Démonstration. On a

$$C_\alpha = \bigcap_{H \in \mathcal{H}_\alpha} H = \bigcap_{u \in \mathbb{S}_{d-1}} H(u, \alpha)$$

or par la proposition 2.3, $H(u, \alpha)$ ne dépend pas de O , ainsi il en découle immédiatement que C_α ne dépend pas de O . Ensuite, on sait que les $H \in \mathcal{H}_\alpha$

sont convexes, donc C_α est convexe ou vide comme une intersection d'une famille de convexes. Enfin, pour prouver *iii*, il suffit de remarquer que pour tout $\alpha \leq \alpha'$ on a $H(O, u, \alpha) \subseteq H(O, u, \alpha')$. \square

Pour des lois quelconques, cet ensemble C_α est très souvent vide. Cependant, il est non vide si $\alpha > d/(d+1)$. Cette condition suffisante est due à Donoho et Gasko [DG92]. L'ensemble de « Tukey-Nolan » n'est pas l'objet principal de notre étude mais lorsqu'il existe les résultats pour les surfaces quantiles impliquent facilement des résultats similaires pour C_α . En particulier, sous les hypothèses très restrictives de [Nol92], les résultats à venir impliquent ceux de cet article.

2.1.2 Hypothèses de régularité

Avant de donner les hypothèses garantissant que $Q(O, \alpha)$ définit une surface quantile, on introduit quelques objets apparaissant dans toute la suite de ce chapitre.

Définition 2.10. Soit $O \in \mathbb{R}^d$, $u \in S_{d-1}$. Pour $y < z$ on définit la **bande** coincée entre $H(O, u, z)$ et $H(O, u, y)$ par

$$H(O, u, y, z) = H(O, u, z) \setminus H(O, u, y).$$

Voici les hypothèses de base, suffisantes pour garantir de bonnes propriétés géométriques aux surfaces quantiles associées à la loi P . Dans toute la suite, on fixe

$$\Delta = [\alpha^-, \alpha^+] \subset [1/2, 1).$$

Cette restriction du domaine des niveaux des ensembles quantile est motivée par le fait que quand α est proche de 1, il devient plus délicat de formuler les théorèmes limites de ce chapitre. Néanmoins tous les résultats peuvent être adaptés pour α^+ qui tend vers 1 avec n suffisamment lentement dans le même esprit que le théorème 1.16.

Hypothèses 2.11. Soit $O \in \mathbb{R}^d$, $u \in S_{d-1}$ et $\alpha \in \Delta$. On note $\partial H(u, \alpha)$ l'hyperplan frontière de $H(u, \alpha)$.

$$(H_0^-) \quad P(\partial H(u, \alpha)) = 0.$$

$$(H_0^+) \quad P(H(O, u, y, z)) > 0 \quad \text{si } Y(O, u, \alpha^-) \leq y < z \leq Y(O, u, \alpha^+).$$

Il est important de noter qu'en aucun cas on suppose que la loi P est absolument continue par rapport à la mesure de Lebesgue. D'ailleurs dans la section des simulations, on utilise des lois sur \mathbb{R}^2 à support de dimension 1.

- Remarque 2.12.** – La condition (\mathbf{H}_0^-) implique que pour tout $\alpha \in \Delta$ on a $P(H(u, \alpha)) = \alpha$.
- La condition (\mathbf{H}_0^-) exclue les lois dont le support ou une partie du support est un hyperplan.
 - Les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) impliquent que l'application $y \mapsto F_{\langle X, u \rangle}(y)$ est continue et strictement croissante sur $F^{-1}(\Delta)$.
 - Si P admet une densité par rapport à la mesure de Lebesgue sur \mathbb{R}^d avec $f = \frac{dP}{d\lambda_d} > 0$ alors les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) sont satisfaites.

Proposition 2.13. On a, pour tout $O \in \mathbb{R}^d$ fixé,

$$\sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |Y(O, u, \alpha)| < +\infty.$$

Démonstration. Puisque P est tendue alors il existe $r^+ > 0$ tel que

$$P(B(O, r^+)) > \alpha^+$$

où $B(O, r^+)$ désigne la boule de centre O et de rayon r^+ .

Remarquons que $H(u, \alpha^+) \cap B(O, r^+) \neq \emptyset$ car sinon on aurait $P(H(u, \alpha^+) \cup B(O, r^+)) = P(H(u, \alpha^+)) + P(B(O, r^+)) > 2\alpha^+ \geq 1$, ce qui est absurde, ainsi pour tout $u \in \mathbb{S}_{d-1}$ on a

$$|Y(O, u, \alpha^+)| \leq r^+$$

donc pour tout $O \in \mathbb{R}^d$ on a $\sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |Y(O, u, \alpha)| \leq r^+ < +\infty$. □

Définition 2.14. Soit $O \in \mathbb{R}^d$ on définit l'ensemble des distances quantiles autorisées vu de O dans chaque direction par

$$\mathcal{Y}_\Delta(O, u) = \{y : P(H(O, u, y)) \in \Delta\} = F_{\langle X - O, u \rangle}^{-1}(\Delta)$$

D'après la proposition 2.13, $\mathcal{Y}_\Delta(O, u)$ vérifie

$$\sup_{u \in \mathbb{S}_{d-1}} \sup_{y \in \mathcal{Y}_\Delta(O, u)} |y| < \infty.$$

Dans la définition suivante on introduit l'ensemble des bandes admissibles au vu des niveaux Δ .

Définition 2.15. Soit $\varepsilon > 0$ on définit l'ensemble des bandes de largeur ε aux distances autorisées par Δ

$$\mathcal{B}_\varepsilon = \left\{ H(O, u, y, y + \varepsilon) : O \in \mathbb{R}^d, u \in \mathbb{S}_{d-1}, y, y + \varepsilon \in \mathcal{Y}_\Delta(O, u) \right\}.$$

L'ensemble \mathcal{B}_ε des bandes dépend de Δ à travers le domaine des distances admissibles $\mathcal{Y}_\Delta(O, u)$.

2.1.3 Régularité de la surface quantile

Dans cette section on s'intéresse aux propriétés de $Q_\alpha(O)$ pour O et α fixés. Sous les conditions faibles (\mathbf{H}_0^-) et (\mathbf{H}_0^+) on montre qu'il s'agit d'une surface fermée.

De plus, sans conditions supplémentaires, elle présente aussi d'autres bonnes propriétés quant à sa forme géométrique. Dans la plupart des cas elle a au plus deux composantes connexes s'intersectant en O . Une classification des formes possibles lorsque O et α bougent est possible et renseigne très clairement sur la distribution de la masse dans \mathbb{R}^d sous P . On renvoie à la section des simulation pour plus de précisions.

Avant de démontrer la continuité de la surface on commence par établir un lemme technique concernant les demi-espaces indexés par O, u, y .

Lemme 2.16. *Soit $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$, $y \in \mathbb{R}$ et $y_\infty \in \overline{\mathbb{R}}$ avec $y \neq y_\infty$. Pour toute suite $(u_n)_{n \in \mathbb{N}}$ d'éléments de \mathbb{S}_{d-1} avec $u_n \neq u$ et $\lim_{n \rightarrow \infty} u_n = u$ et pour toute suite $(y_n)_{n \in \mathbb{N}}$ de réels avec $y_n \neq y_\infty$ et $\lim_{n \rightarrow \infty} y_n = y_\infty$, il existe une suite croissante d'entiers naturels $(n_k)_{k \in \mathbb{N}}$ avec $\lim_{k \rightarrow \infty} n_k = +\infty$ et il existe une suite d'ensembles $(H_k)_{k \geq 1}$ tels que*

- i. $\forall k \geq 1, H_{k+1} \subset H_k$
- ii. $\bigcap_{k \geq 1} H_k = \emptyset$
- iii. $H(O, u, y) \setminus H(O, u_{n_k}, y_{n_k}) \subset H_k \subset H(O, u, y)$

Démonstration. Soit u une direction dans \mathbb{S}_{d-1} , $y \in \mathbb{R}$ et $y_\infty \in \overline{\mathbb{R}}$ avec $y < y_\infty$. Soit $p_{H(O, u, y)}$ l'application de la projection orthogonale sur $\partial H(O, u, y)$. On note $Q = p_{H(O, u, y)}(O) = O + yu$.

Soit $(u_n)_{n \in \mathbb{N}}$ d'éléments de \mathbb{S}_{d-1} avec $u_n \neq u$ et $\lim_{n \rightarrow \infty} u_n = u$ et $(y_n)_{n \in \mathbb{N}}$ de réels avec $y_n \neq y_\infty$ et $\lim_{n \rightarrow \infty} y_n = y_\infty$, on peut extraire $((u_{m_k}, y_{m_k}))_{k \geq 1}$ suite de $\mathbb{S}_{d-1} \times \mathbb{R}$ telle que la suite $(\langle u_{m_k}, u \rangle)_{k \geq 1}$ est croissante avec $\lim_{k \rightarrow \infty} \langle u_{m_k}, u \rangle = 1$. Le schéma suivant résume les notations.

On pose

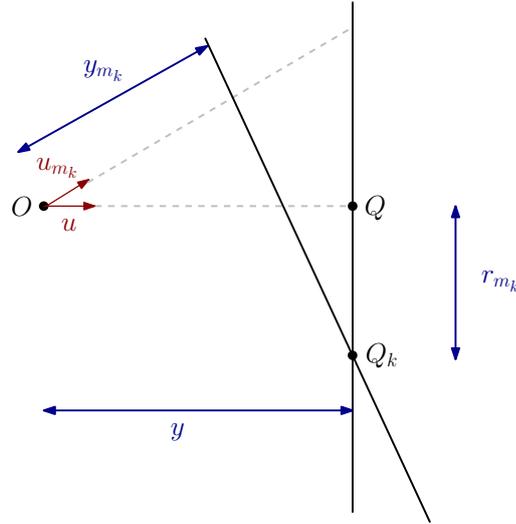
$$D_{m_k} = \partial H(O, u, y) \cap \partial H(O, u_{m_k}, y_{m_k}),$$

l'intersection des bords des demi-espaces $H(O, u, y)$ et $H(O, u_{m_k}, y_{m_k})$. Cette intersection D_{m_k} est non vide car $u_{m_k} \neq u$, c'est un hyperplan de dimension $d - 2$. On définit la distance de Q à D_{m_k} par

$$r_{m_k} = \inf_{Q_k \in D_{m_k}} \|Q_k - Q\|.$$

Étape I : on montre que

$$\lim_{k \rightarrow \infty} r_{m_k} = +\infty$$



Pour cela, on pose $Q' = O + y'u$ avec $y < y' < y_\infty$ un point de la droite (O, u) et $A'_k = H(O, u_{m_k}, y'_{m_k})$ le demi-espace de normale u_{m_k} et qui intersecte (O, u) exactement en Q' , i.e.

$$A'_k \cap (O, u) = Q'.$$

Il est facile de voir que

$$y'_{m_k} = y' \langle u_{m_k}, u \rangle$$

ainsi, la suite $(y_{m_k})_{k \geq 1}$ est croissante avec $\lim_{k \rightarrow \infty} y'_{m_k} = y'$. Pour k suffisamment grand, on a $y'_{m_k} < y_{m_k}$, ainsi on obtient l'inclusion stricte suivante

$$A'_k \subsetneq H(O, u_{m_k}, y_{m_k}).$$

En posant

$$D'_{m_k} = \partial H(O, u, y) \cap \partial A'_k$$

et

$$r'_{m_k} = \inf_{Q'_k \in D'_{m_k}} \|Q'_k - Q\|,$$

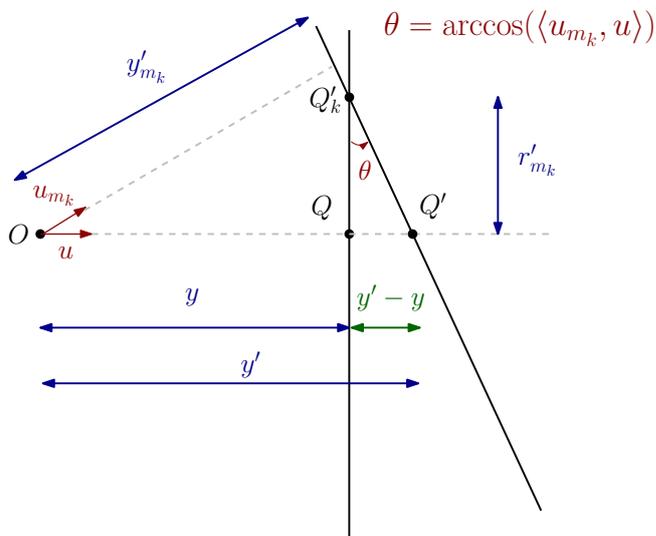
En observant en outre que $A'_k \subsetneq H(O, u_{m_k}, y_{m_k})$, on a $r'_{m_k} < r_{m_k}$. On en déduit que

$$r'_{m_k} = \frac{y' - y}{\tan(\arccos(\langle u_{m_k}, u \rangle))}$$

et de fait $\lim_{k \rightarrow \infty} r'_{m_k} = +\infty$, donc $\lim_{k \rightarrow \infty} r_{m_k} = +\infty$. Ainsi, on peut en extraire une sous suite croissante $(r_{n_k})_{k \geq 1}$ avec

$$\lim_{k \rightarrow \infty} r_{n_k} = +\infty. \quad (2.3)$$

Voici le schéma résumant l'étape I



Étape II : on prouve l'existence d'une suite décroissante $(H_k)_{k \geq 1}$ vérifiant *i. ii. et iii.* du lemme 2.16.

Pour cela on considère pour $k \geq 1$, on l'ensemble des directions

$$\mathbb{V}_k = \{v \in \mathbb{S}_{d-1} : \langle v, u \rangle = \langle u_{n_k}, u \rangle\}$$

et l'ensemble des demi-espaces

$$\mathbb{U}_k = \{H(O, v, y_{n_k}) : v \in \mathbb{V}_k\}$$

obtenus par révolution de $H(O, u_{n_k}, y_{n_k})$ autour de (O, u) . Enfin, on définit

$$\mathbb{T}_k = \bigcap_{v \in \mathbb{V}_k} H(O, v, y_{n_k}) = \bigcap_{\hat{H} \in \mathbb{U}_k} \hat{H}$$

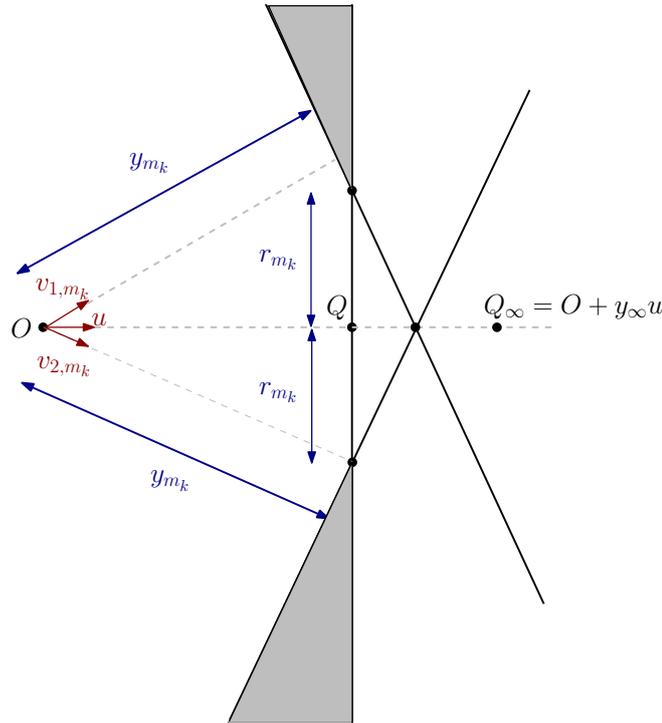
puis

$$H_k = \bigcup_{v \in \mathbb{V}_k} H(O, u, y) \setminus H(O, v, y_{n_k}) = H(O, u, y) \setminus \mathbb{T}_k.$$

Comme $r_{n_k} \uparrow +\infty$ il est clair que $H_{k+1} \subset H_k$. Ensuite, puisque $u_{n_k} \neq u$ alors pour tout $\hat{H} \in \mathbb{U}_k$ on a $H(O, u, y) \cap \hat{H} \neq \emptyset$ donc en particulier,

$$H_k \neq \emptyset.$$

Il est évident que $H_k \subset H(O, u, y)$ et que $u_{n_k} \in \mathbb{V}_k$ on en déduit, par définition de H_k , que $H(O, u, y) \setminus H(O, u_{n_k}, y_{n_k}) \subset H_k$.



En gris $H_k = H(O, u, y) \setminus T_k$

On a $\lim_{k \rightarrow \infty} \langle u_{n_k}, u \rangle = 1$ donc $V_\infty = \{u\}$ et $\lim_{k \rightarrow \infty} y_{n_k} = y$ et de plus $\bigcap_{k \geq 1} H_k = \emptyset$. Ainsi on a montré que si $y < y_\infty$ alors pour toute suite $(u_n)_{n \in \mathbb{N}}$ d'éléments de \mathbb{S}_{d-1} avec $u_n \neq u$ et $\lim_{n \rightarrow \infty} u_n = u$ et pour toute suite $(y_n)_{n \in \mathbb{N}}$ de réels avec $y_n \neq y_\infty$ et $\lim_{n \rightarrow \infty} y_n = y_\infty$, il existe une suite croissante d'entiers naturels $(n_k)_{k \in \mathbb{N}}$ avec $\lim_{k \rightarrow \infty} n_k = +\infty$ et il existe une suite décroissante d'ensembles $(H_k)_{k \geq 1}$ tels que

- i. $\forall k \geq 1, H_{k+1} \subset H_k$
- ii. $\bigcap_{k \geq 1} H_k = \emptyset$
- iii. $H(O, u, y) \setminus H(O, u_{n_k}, y_{n_k}) \subset H_k \subset H(O, u, y)$.

Dans le cas où $y_\infty < y' < y$ on peut montrer, de manière complètement similaire, le même résultat. □

La régularité de la surface quantile de niveau α est établie dans la proposition suivante. les condition (\mathbf{H}_0^-) et (\mathbf{H}_0^+) sont minimales au sens suivant.

Proposition 2.17. *Sous (\mathbf{H}_0^-) on a l'équivalence entre (\mathbf{H}_0^+) et $(u, \alpha) \mapsto Q(O, u, \alpha)$ est bornée et continue sur $\mathbb{S}_{d-1} \times \Delta$ pour $O \in \mathbb{R}^d$.*

Démonstration.

1- Condition nécessaire.

Étape I : on montre que la surface quantile vue de O et de niveau α est bornée.

Par la proposition 2.13, on sait que sous (\mathbf{H}_0^-) et (\mathbf{H}_0^+) on a, pour tout $O \in \mathbb{R}^d$ fixé,

$$\sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |Y(O, u, \alpha)| < +\infty.$$

donc il existe $r > 0$ tel que pour tout $u \in \mathbb{S}^d$, $\alpha \in \Delta$ on a

$$O + Y(O, u, \alpha)u = Q(O, u, \alpha) \in B(O, r)$$

donc pour tout $O \in \mathbb{R}^d$ et $\alpha \in \Delta$ la surface $Q_\alpha(O)$ est bornée.

Étape II : Continuité de la surface quantile.

Si l'application $(u, \alpha) \mapsto Q(O, u, \alpha)$ n'est pas continue, alors il existe une suite $(u_n)_{n \geq 1}$ de \mathbb{S}_{d-1} et $(\alpha_n)_{n \geq 1}$ de Δ avec $u_n \rightarrow u$ et $\alpha_n \rightarrow \alpha$ telles que

$$\lim_{n \rightarrow \infty} Q(O, u_n, \alpha_n) \neq Q(O, u, \alpha).$$

Puisque $Q(O, u_n, \alpha_n)$ est bornée, alors il existe une sous-suite $(u_{n_j})_{j \geq 1}$ telle que $u_{n_j} \rightarrow u$ et une suite $(\alpha_{n_j})_{j \geq 1}$ telle que $\alpha_{n_j} \rightarrow \alpha$ vérifiant de plus

$$\lim_{j \rightarrow \infty} Q(O, u_{n_j}, \alpha_{n_j}) = Q_\infty = O + y_\infty u \in \mathbb{R}^d$$

où $y_\infty = \lim_{j \rightarrow \infty} Y(O, u_{n_j}, \alpha_{n_j}) < +\infty$ et $y_\infty \neq y = Y(O, u, \alpha)$ et ainsi $Q_\infty \neq Q(O, u, \alpha)$.

On suppose que $y < y_\infty < \infty$. En appliquant le lemme précédent, il existe une sous suite croissante $(n_{j(k)})_{k \geq 1}$ avec $n_{j(k)} \rightarrow +\infty$ et une suite décroissante d'ensembles H_k tels que

$$\bigcap_{k \geq 1} H_k = \emptyset, \quad H(O, u, y) \setminus H(u_{n_{j(k)}}, \alpha_{n_{j(k)}}) \subset H_k \subset H(O, u, y)$$

par conséquent la suite $(H(O, u, y) \setminus H_k)_{k \geq 1}$ est croissante au sens de l'inclusion avec

$$\lim_{k \rightarrow \infty} \uparrow (H(O, u, y) \setminus H_k) = \bigcup_{k \geq 1} (H(O, u, y) \setminus H_k) = H(O, u, y) \setminus \bigcap_{k \geq 1} H_k$$

donc

$$\bigcup_{k \geq 1} (H(O, u, y) \setminus H_k) = H(O, u, y)$$

de la propriété de continuité de la mesure et la condition (\mathbf{H}_0^-) , on déduit que

$$\begin{aligned}
\alpha_{n_j(k)} &\leq P(H(O, u, y)) \\
&= \lim_{k \rightarrow \infty} \uparrow P(H(O, u, y) \setminus H_k) \\
&\leq \lim_{k \rightarrow \infty} P(H(O, u, y) \cap H(u_{n_j(k)}, \alpha_{n_j(k)})) \\
&\leq \lim_{k \rightarrow \infty} P(H(u_{n_j(k)}, \alpha_{n_j(k)})) \\
&\leq \alpha
\end{aligned}$$

donc tout $y > Y(O, u, \alpha)$ vérifie

$$\lim_{k \rightarrow \infty} \alpha_{n_j(k)} = \alpha = P(H(O, u, y))$$

et dans ce cas, il vient

$$\begin{aligned}
P(H(O, u, Y(O, u, \alpha), y)) &= P((H(O, u, y) \setminus H(u, \alpha))) \\
&= P((H(O, u, y)) - P(H(u, \alpha))) \\
&= \alpha - \alpha \\
&= 0
\end{aligned}$$

donc $P(H(O, u, Y(O, u, \alpha), y)) = 0$ avec $y > Y(O, u, \alpha)$, ce qui contredit la condition (\mathbf{H}_0^+) . Le cas $y < Y(O, u, \alpha)$ se fait de manière entièrement similaire.

Conclusion 1 : pour tout $O \in \mathbb{R}^d$, sous (\mathbf{H}_0^-) et (\mathbf{H}_0^+) l'application $(u, \alpha) \mapsto Y(O, u, \alpha)$ est bornée et continue sur $S_{d-1} \times \Delta$, donc l'ensemble $Q(O, \alpha)$ étoilé est une surface fermée.

2-Condition suffisante. On veut montrer que si $(u, \alpha) \mapsto Y(O, u, \alpha)$ est bornée et continue sur $S_{d-1} \times \Delta$ alors (\mathbf{H}_0^+) est vérifiée. Pour ce faire, on va montrer que

$$\neg(\mathbf{H}_0^+) \Rightarrow Y(O, \cdot, \cdot) \text{ n'est pas continue sur } S_{d-1} \times \Delta.$$

où $\neg(\mathbf{H}_0^+)$ signifie la contraposée de (\mathbf{H}_0^+) .

Par l'absurde supposons que $\neg(\mathbf{H}_0^+)$ est vrai et que $(u, \alpha) \mapsto Q(O, u, \alpha)$ est bornée et continue sur $S_{d-1} \times \Delta$.

Par $\neg(\mathbf{H}_0^+)$, il existe $\varepsilon_0 > 0$ tel que $B_0 \in \mathcal{B}_{\varepsilon_0}$ avec $P(B_0) = 0$. Soit $u \in S_{d-1}$ tel que $B_0 = H(O, u, y + \varepsilon_0)$ et on pose

$$\alpha_0 = P(H(O, u, y + \varepsilon_0)).$$

On a $P(H(O, u, y + \varepsilon_0)) = P(H(O, u, y)) = \alpha_0$ car

$$\begin{aligned}\alpha_0 &= P(H(O, u, y + \varepsilon_0)) \\ &= P(H(O, u, y) \cup H(O, u, y, y + \varepsilon_0)) \\ &= P(H(O, u, y)) + P(B_0) \\ &= P(H(O, u, y)).\end{aligned}$$

puisque $P(H(O, u, y)) = \alpha_0$ alors $Y(O, u, \alpha_0) \leq y$.

Soit $(\alpha_k^+)_{k \in \mathbb{N}}$ une suite strictement décroissante avec $\alpha_k^+ \downarrow \alpha_0$. Sous (\mathbf{H}_0^-) on a $H(O, u, y + \varepsilon_0) \subsetneq H(u, \alpha_k^+)$ donc

$$Y(O, u, \alpha_k^+) \geq y + \varepsilon_0$$

par continuité de $Y(O, \cdot, \cdot)$ on obtient

$$\lim_{k \rightarrow \infty} Y(O, u, \alpha_k^+) = Y(O, u, \alpha_0) \geq y + \varepsilon_0$$

ainsi on a

$$y + \varepsilon_0 \leq Y(O, u, \alpha_0) \leq y$$

donc $\varepsilon_0 = 0$ ce qui est absurde car $\varepsilon_0 > 0$.

Conclusion 2 : Pour $O \in \mathbb{R}^d$, sous (\mathbf{H}_0^-) si l'application $(u, \alpha) \mapsto Y(O, u, \alpha)$ est bornée et continue sur $S_{d-1} \times \Delta$ alors et (\mathbf{H}_0^+) est vérifiée. \square

L'ensemble $Q_\alpha(O)$ est l'image par une application continue sur S_{d-1} qui est compact, on a immédiatement le corollaire suivant.

Corollaire 2.18. Sous (\mathbf{H}_0^-) et (\mathbf{H}_0^+) , l'ensemble $Q_\alpha(O)$ est une surface fermée, pour tout $O \in \mathbb{R}^d$ et $\alpha \in \Delta$.

Les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) peuvent être réécrites sous une forme fonctionnelle grâce à une fonction Ψ de minoration des probabilités des bandes, pour ce faire, on rappelle que pour $\varepsilon > 0$ l'ensemble des bandes de largeur ε aux distances autorisées par Δ est défini par

$$\mathcal{B}_\varepsilon = \left\{ H(O, u, y, y + \varepsilon) : O \in \mathbb{R}^d, u \in S_{d-1}, y, y + \varepsilon \in \mathcal{Y}_\Delta(O, u) \right\}.$$

Définition 2.19. On définit l'application Ψ de minoration des probabilités des bandes par

$$\begin{aligned}\Psi : \mathbb{R}^+ &\longrightarrow [0, 1] \\ \varepsilon &\longmapsto \inf_{B \in \mathcal{B}_\varepsilon} P(B).\end{aligned}$$

Proposition. Sous (\mathbf{H}_0^-) et (\mathbf{H}_0^+) les deux conditions suivantes sont vérifiées

$$\begin{aligned} (H_{0,\Psi}^-) \quad & \lim_{\varepsilon \rightarrow 0} \Psi(\varepsilon) = 0 \\ (H_{0,\Psi}^+) \quad & \Psi(\varepsilon) > 0 \text{ pour tout } \varepsilon > 0. \end{aligned}$$

Démonstration. En effet, (\mathbf{H}_0^-) implique $(H_{0,\Psi}^-)$, car pour $y = Y(O, u, \alpha)$ avec $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$ on a par continuité de la mesure

$$\lim_{\varepsilon \rightarrow 0} \Psi(\varepsilon) \leq \lim_{\varepsilon \rightarrow 0} P(H(O, u, y, y + \varepsilon)) = P(\partial H(u, \alpha)) = 0.$$

Il reste à montrer que sous (\mathbf{H}_0^-) on a (\mathbf{H}_0^+) implique $(H_{0,\Psi}^+)$. Soit $\varepsilon > 0$ on a

$$\Psi(\varepsilon) = \inf_{u \in \mathbb{S}_{d-1}} \inf_{y, y + \varepsilon \in \mathcal{Y}_\Delta(O, u)} P(H(O, u, y, y + \varepsilon)).$$

Par l'absurde : supposant qu'il existe $\varepsilon_0 > 0$ tel que $\Psi(\varepsilon_0) = 0$, dans ce cas il existe deux suites $u_k \in \mathbb{S}_{d-1}$ et $y_k, y_k + \varepsilon_0 \in \mathcal{Y}_\Delta(O, u_k)$ telles que

$$\lim_{k \rightarrow \infty} P(H(O, u_k, y_k, y_k + \varepsilon_0)) = 0.$$

Puisque $\mathcal{Y}_\Delta(O) = \bigcup_{u \in \mathbb{S}_{d-1}} \mathcal{Y}_\Delta(O, u)$ est compact, alors on peut extraire de la suite $(u_k, y_k) \in \mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O, u_k)$ une sous-suite $(u'_k, y'_k) \in \mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O, u'_k)$ convergente, de limite notée (u_0, y_0) .

On sait que $\mathcal{Y}_\Delta(O, u'_k) = [Y(O, u'_k, \alpha^-), Y(O, u'_k, \alpha^+)]$ donc par continuité de $u \rightarrow Y(O, u, \alpha)$ il vient que $(u_0, y_0) \in \mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O, u_0)$, de même que pour $y_0 + \varepsilon_0$. On pose

$$B'_k = H(O, u'_k, y'_k, y'_k + \varepsilon_0)$$

par hypothèse de l'absurde on a, $\lim_{k \rightarrow \infty} P(B'_k) = 0$. En posant

$$B_0 = H(O, u_0, y_0, y_0 + \varepsilon_0)$$

on a

$$1_{B_0} \geq \lim_{k \rightarrow \infty} 1_{B'_k} \geq 1_{B_0 \setminus \partial B_0} = 1_{B_0} - 1_{\partial B_0}.$$

En effet, si $x \notin B_0$ il existe δ suffisamment petit et un pavé $V_\delta = [u_0 - \delta, u_0 + \delta] \times [y_0 - \delta, y_0 + \delta]$ tel que $x \notin \bigcup_{(u,y) \in V_\delta} H(O, u, y, y + \varepsilon_0)$ donc pour tout k assez grand, $x \notin B'_k$. Si $x \in \partial B_0$ on a toujours $1_{B'_k}(x) \geq 1_{B_0 \setminus \partial B_0}(x) = 0$. Enfin, si $x \in B_0 \setminus \partial B_0$ il existe δ suffisamment petit et un pavé $V_\delta = [u_0 - \delta, u_0 + \delta] \times [y_0 - \delta, y_0 + \delta]$ tel que $x \in \bigcap_{(u,y) \in V_\delta} H(O, u, y, y + \varepsilon_0)$ donc pour tout k assez grand, $x \in B'_k$. Par conséquent,

$$P(B_0) \geq \lim_{k \rightarrow \infty} P(B'_k) \geq P(B_0) - P(\partial B_0)$$

mais d'après (\mathbf{H}_0^-) et (\mathbf{H}_0^+) , on a $P(B_0) > 0$ et $P(\partial B_0) = 0$. On en déduit que $\lim_{k \rightarrow \infty} P(B'_k) = P(B_0) > 0$, ce qui est contradictoire. \square

Ensuite, voici une proposition garantissant la continuité de la fonction Ψ sous les hypothèses (\mathbf{H}_0^-) et (\mathbf{H}_0^-) .

Proposition 2.20. *L'application Ψ est continue à droite en tout $\varepsilon > 0$. Si de plus (\mathbf{H}_0^-) et (\mathbf{H}_0^-) sont vérifiées alors $\Psi(\varepsilon)$ est continue en tout $\varepsilon \geq 0$.*

Démonstration. La fonction monotone Ψ admet des limites à droite en tout $\varepsilon_0 \geq 0$ et à gauche en tout $\varepsilon_0 > 0$. Soit $\varepsilon_k \downarrow \varepsilon_0 > 0$. Pour tout $\theta > 0$ il existe $B_{\theta,0} \in \mathcal{B}_{\varepsilon_0}$ tel que

$$(1 + \theta)\Psi(\varepsilon_0) > P(B_{\theta,0}) \geq \Psi(\varepsilon_0)$$

$$B_{\theta,0} = H(O, u_\theta, y_\theta, y_\theta + \varepsilon_0)$$

et on considère la suite décroissante d'ensembles

$$B_{\theta,k} = H(O, u_\theta, y_\theta, y_\theta + \varepsilon_k)$$

qui sont de limite $\bigcap_k B_{\theta,k} = B_{\theta,0}$ et par continuité de la mesure, vérifient $P(B_{\theta,k}) \downarrow P(B_{\theta,0})$. Ainsi, il existe $k_\theta > 0$ tel que pour tout $k \geq k_\theta$

$$(1 + \theta)\Psi(\varepsilon_0) > P(B_{\theta,k}) \geq P(B_{\theta,0}) \geq \Psi(\varepsilon_0)$$

donc également, par croissance de Ψ ,

$$(1 + \theta)\Psi(\varepsilon_0) > P(B_{\theta,k}) \geq \Psi(\varepsilon_k) \geq \Psi(\varepsilon_0).$$

Comme $\Psi(\varepsilon_k)$ converge vers la limite à droite $\Psi(\varepsilon_0^+)$ de Ψ en ε_0 , on a en fait montré que pour tout $\theta > 0$,

$$(1 + \theta)\Psi(\varepsilon_0) > \lim_{k \rightarrow \infty} \Psi(\varepsilon_k) = \Psi(\varepsilon_0^+) \geq \Psi(\varepsilon_0)$$

c'est-à-dire

$$\lim_{k \rightarrow \infty} \Psi(\varepsilon_k) = \Psi(\varepsilon_0^+) = \Psi(\varepsilon_0).$$

De même, si $\varepsilon_k \uparrow \varepsilon_0 > 0$ alors à tout $\theta > 0$ on associe une suite $B_{\theta,k} \in \mathcal{B}_{\varepsilon_k}$ telle que

$$(1 + \theta)\Psi(\varepsilon_k) > P(B_{\theta,k}) > \Psi(\varepsilon_k)$$

et on extrait, par compacité en u et y , une suite stabilisée

$$B_{\theta,k_n} = H(O, u_{k_n}, y_{k_n}, y_{k_n} + \varepsilon_{k_n})$$

avec $u_{k_n} \rightarrow u_\theta, y_{k_n} \rightarrow y_\theta$.

On pose $B_{\theta,0} = H(O, u_\theta, y_\theta, y_\theta + \varepsilon_0)$. Sous (H_0^-) et (H_0^+) par continuité des surfaces on a $P(B_{\theta,k_n}) \rightarrow P(B_{\theta,0})$ donc pour tout $k \geq k_\theta$ on a

$$\lim_{n \rightarrow \infty} P(B_{\theta,k_n}) \rightarrow \Psi(\varepsilon_0^-) \geq (1 - \theta)P(B_{\theta,0}) \geq (1 - \theta)\Psi(\varepsilon_0)$$

$$\Psi(\varepsilon_0) \geq \Psi(\varepsilon_0^-) \geq (1 - \theta)\Psi(\varepsilon_0)$$

et cela pour tout $\theta > 0$. Par suite,

$$\Psi(\varepsilon_0^-) = \Psi(\varepsilon_0).$$

□

Remarque 2.21. On définit pour tout $t \in [1/2, 1)$ la fonction inverse généralisée de Ψ comme suit

$$\Psi^{-1}(t) = \inf \{ \varepsilon \geq 0 : \Psi(\varepsilon) \geq t \}.$$

Puisque Ψ est continue et croissante alors cette fonction est càdlàg et strictement croissante avec

$$\Psi \circ \Psi^{-1}(t) = t \text{ et } \Psi^{-1} \circ \Psi(t) \geq t.$$

2.2 Surfaces quantiles empiriques

Soit X_1, \dots, X_n une suite de variables aléatoires de \mathbb{R}^d , indépendantes et identiquement distribuées selon la loi P . La mesure empirique dans \mathbb{R}^d associée est définie par

$$P_n^X = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$$

La mesure empirique dans \mathbb{R} associée à la loi projetée le long de la direction (O, u) est définie par

$$P_{n,O,u} = \frac{1}{n} \sum_{i=1}^n \delta_{\langle X_i - O, u \rangle}$$

2.2.1 Estimateurs non-paramétriques

Les objets empiriques suivants, sont définis de manière entièrement analogue à ceux de la section 1.

Définition 2.22. Le **quantile empirique** de la loi P dans la direction $u \in \mathbb{S}_{d-1}$ de niveau $\alpha \in [1/2, 1)$ vu de $O \in \mathbb{R}^d$ est défini par

$$Y_n(O, u, \alpha) = \inf \{ y \in \mathbb{R} : P_n(H(O, u, y)) \geq \alpha \}$$

Proposition 2.23. *Pour $O \in \mathbb{R}^d, u \in \mathbb{S}_{d-1}$ et $\alpha \in [1/2, 1)$ le demi-espace $H(O, u, Y_n(O, u, \alpha))$ ne dépend pas du choix de O .*

Démonstration. Par analogie avec la preuve de la proposition 2.3.

Soit $y \in \mathbb{R}$, pour $O \neq O' \in \mathbb{R}^d$ en remplaçant P par P_n dans l'égalité 2.2 on obtient

$$Y_n(O, u, \alpha) = Y_n(O', u, \alpha) + \langle O' - O, u \rangle \quad (2.4)$$

en utilisant les égalités 2.1 et 2.4 on a

$$\begin{aligned} H(O, u, Y_n(O, u, \alpha)) &= H(O', u, Y_n(O, u, \alpha) - \langle O' - O, u \rangle) \\ &= H(O', u, Y_n(O', u, \alpha)). \end{aligned}$$

□

Proposition 2.24. *Pour $O \in \mathbb{R}^d, u \in \mathbb{S}_{d-1}$ et $\alpha \in [1/2, 1)$ le processus $Y_n(O, u, \alpha) - Y(O, u, \alpha)$ ne dépend pas du choix de O .*

Démonstration. Pour $u \in \mathbb{S}_{d-1}, \alpha \in [1/2, 1)$ et $O \neq O' \in \mathbb{R}^d$ on a

$$\begin{aligned} Y_n(O, u, \alpha) &= Y_n(O', u, \alpha) + \langle O' - O, u \rangle \\ Y(O, u, \alpha) &= Y(O', u, \alpha) + \langle O' - O, u \rangle \end{aligned}$$

donc

$$Y_n(O, u, \alpha) - Y(O, u, \alpha) = Y_n(O', u, \alpha) - Y(O', u, \alpha).$$

□

Définition 2.25. *Le **demi-espace associé au quantile** $Y_n(O, u, \alpha)$ pour $u \in \mathbb{S}_{d-1}, \alpha \in [1/2, 1)$ et $O \in \mathbb{R}^d$ est noté*

$$H_n(u, \alpha) := H(O, u, Y_n(O, u, \alpha)).$$

Définition 2.26. *Le **point quantile empirique** de la loi P dans la direction $u \in \mathbb{S}_{d-1}$ de niveau $\alpha \in [1/2, 1)$ vu de $O \in \mathbb{R}^d$ est défini par*

$$Q_n(O, u, \alpha) = O + Y_n(O, u, \alpha) \cdot u.$$

Par définition de $Q(O, u, \alpha)$ et $Q_n(O, u, \alpha)$ on peut énoncer l'analogie de la proposition 2.24 pour les points quantiles.

Proposition 2.27. *Pour $O \in \mathbb{R}^d, u \in \mathbb{S}_{d-1}$ et $\alpha \in [1/2, 1)$ le processus $Q_n(O, u, \alpha) - Q(O, u, \alpha)$ ne dépend pas du choix de O .*

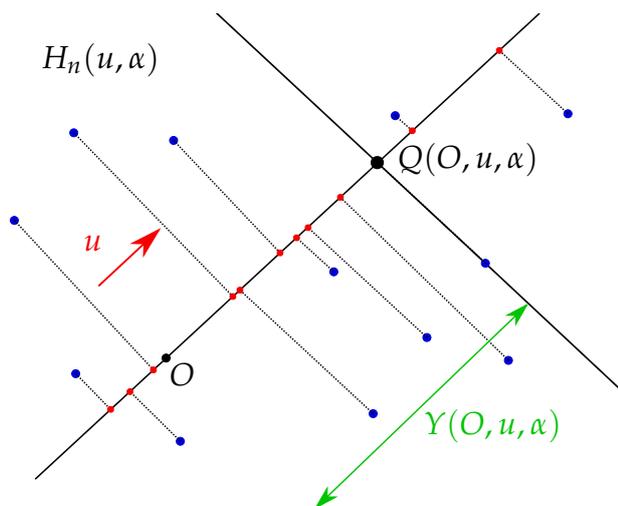


FIGURE 2.7 – $Y_n(O, u, \alpha)$, $Q_n(O, u, \alpha)$ et $H_n(u, \alpha)$

Définition 2.28. *L'ensemble quantile empirique de la loi P dans la direction $u \in \mathbb{S}_{d-1}$ de niveau $\alpha \in [1/2, 1)$ vu de $O \in \mathbb{R}^d$ est défini par*

$$Q_{n,\alpha}(O) = \{Q_n(O, u, \alpha) : u \in \mathbb{S}_{d-1}\}.$$

Définition 2.29. *L'ensemble des demi-espaces empiriques de niveau $\alpha \in [1/2, 1)$ est définie par*

$$\mathcal{H}_{n,\alpha} = \{H_n(u, \alpha) : u \in \mathbb{S}_{d-1}\}.$$

Définition 2.30. *L'intersection de tout les demi-espaces de masse empirique $\alpha \in [1/2, 1)$ est défini par*

$$\mathcal{C}_{n,\alpha} = \bigcap_{H_n \in \mathcal{H}_\alpha} H_n.$$

Proposition 2.31. *Pour $\alpha \in [1/2, 1)$ si l'ensemble $\mathcal{C}_{n,\alpha}$ existe alors il vérifie les propriétés suivantes*

- i. $\mathcal{C}_{n,\alpha}$ ne dépend pas du choix de $O \in \mathbb{R}^d$
- ii. $\mathcal{C}_{n,\alpha}$ est croissant en α (au sens de l'inclusion)

Proposition 2.32. *Pour $O \in \mathbb{R}^d$ et $\alpha \in [1/2, 1)$, l'ensemble $Q_{n,\alpha}(O)$ est une surface sphériques par morceaux.*

Remarque 2.33. Contrairement à la surface $Q_\alpha(O)$, la surface empirique $Q_n(O, \alpha)$ n'est pas continue en α et cela même à u fixé. Cela se comprend rien qu'en considérant $\langle X_i, u \rangle$ et la fonction quantile réelle associée, qui est constante par morceaux avec des sauts en chaque $\alpha = k/N$. De plus, la façon de sauter en même temps que u varie est vraiment complexe.

Proposition 2.34. Sous (\mathbf{H}_0^-) , pour tout $n > d$, l'ensemble des demi-espaces empiriques de masse $\alpha \in \Delta$ peut être réécrit comme suit

$$\mathcal{H}_{n,\alpha} = \left\{ H : H \text{ un demi-espace, } P_n(H) \in \left[\alpha, \alpha + \frac{d}{n} \right] \right\}.$$

Démonstration. On suppose (\mathbf{H}_0^-) et on veut montrer qu'avec probabilité un on a, pour tout O, u, α et tout $n > d$

$$\alpha \leq P_n(H(O, u, Y_n(O, u, \alpha))) = P_n(H_n(u, \alpha)) \leq \alpha + \frac{d}{n}.$$

Par définition de $Y_n(O, u, \alpha)$ on a $P_n(H(O, u, Y_n(O, u, \alpha))) \geq \alpha$. Fixons $n > d$. Sous (\mathbf{H}_0^-) la probabilité que X_1, \dots, X_{d+1} soient dans le même hyperplan est nulle. En effet, en notant $\partial H(x_1, \dots, x_d)$ l'unique hyperplan passant par d points distincts x_1, \dots, x_d , on a, par conditionnement puis par indépendance,

$$\begin{aligned} & \mathbb{P}(X_{d+1} \in \partial H(X_1, \dots, X_d)) \\ &= \int_{x_1 \in \mathbb{R}^d} \dots \int_{x_d \in \mathbb{R}^d} \mathbb{P}(X_{d+1} \in \partial H(x_1, \dots, x_d) \mid X_1 = x_1, \dots, X_d = x_d) dP(x_1) \dots dP(x_d) \\ &= \int_{x_1 \in \mathbb{R}^d} \dots \int_{x_d \in \mathbb{R}^d} \mathbb{P}(X_{d+1} \in \partial H(x_1, \dots, x_d)) dP(x_1) \dots dP(x_d) \\ &= 0 \end{aligned}$$

car $\mathbb{P}(X_{d+1} \in \partial H) = 0$ pour tout hyperplan ∂H , d'après (H_0^-) . Par suite,

$$\begin{aligned} & \mathbb{P}(\{\exists i_1, \dots, i_{d+1} \text{ distincts} : X_{i_{d+1}} \in \partial H(X_{i_1}, \dots, X_{i_d})\}) \\ & \leq \sum_{0 \leq i_1, \dots, i_{d+1} \leq n} \mathbb{P}(X_{i_{d+1}} \in \partial H(X_{i_1}, \dots, X_{i_d})) \\ & = \frac{n!}{(n-d-1)!} \mathbb{P}(X_{d+1} \in \partial H(X_1, \dots, X_d)) \\ & = 0. \end{aligned}$$

Ainsi, presque sûrement, aucun hyperplan ne contient plus de d points de l'échantillon, ce qui s'écrit aussi

$$\mathbb{P}\left(\sup_{H \in \mathcal{H}} P_n(\partial H) \geq \frac{d+1}{n}\right) = 0.$$

avec la notation

$$\text{int}(H(O, u, y)) = \left\{ x \in \mathbb{R}^d : \langle x - O, u \rangle < y \right\}$$

on pose

$$\text{int}(H_n(u, \alpha)) = \text{int}(H(O, u, Y_n(O, u, \alpha)))$$

et

$$\partial H_n(u, \alpha) = H_n(u, \alpha) \setminus \text{int}(H_n(u, \alpha)).$$

Par conséquent, avec probabilité un, pour tout u, α il vient

$$\begin{aligned} P_n(H_n(u, \alpha)) &= P_n(\text{int}(H_n(u, \alpha))) + P_n(\partial H_n(u, \alpha)) \\ &\leq P_n(\text{int}(H_n(u, \alpha))) + \frac{d}{n}. \end{aligned}$$

Enfin, on a aussi

$$P_n(\text{int}(H_n(u, \alpha))) \leq \alpha$$

car si $P_n(\text{int}(H_n(u, \alpha))) > \alpha$ alors il y a au moins $\lceil n\alpha \rceil$ points $X_i \in \text{int}(H_n(u, \alpha))$ donc par définition de $H_n(u, \alpha)$ on a $\langle X_i, u \rangle < Y_n(O, u, \alpha)$ et en notant

$$\overset{\circ}{Y}_n(O, u, \alpha) = \max_{X_i \in \overset{\circ}{H}_n(u, \alpha)} (\langle X_i, u \rangle) < Y_n(O, u, \alpha)$$

il vient

$$P_n(H(O, u, \overset{\circ}{Y}_n(O, u, \alpha))) \geq \frac{\lceil n\alpha \rceil}{n} \geq \alpha$$

ce qui contredit la définition

$$Y_n(O, u, \alpha) = \inf \{ y \in \mathbb{R} : P_n(H(O, u, y)) \geq \alpha \} \leq \overset{\circ}{Y}_n(O, u, \alpha).$$

□

Pour compléter, il est clair qu'un nombre fini d'hyperplans contient exactement d points parmi X_1, \dots, X_n , si $d \leq n$, et il y en a autant que de d -uplets, à savoir $n! / (n-d)!(d)!$. De même, pour tout $k < d \leq n$ un nombre infini d'hyperplans contient k points parmi X_1, \dots, X_n : considérer pour chaque hyperplan contenant k points une légère rotation de l'hyperplan autour de ces points, sans rencontrer d'autre point.

2.2.2 Lien avec le quantile réel

Avant de s'intéresser aux propriétés asymptotiques du quantile multivarié, il serait intéressant d'établir des liens entre cette généralisation multidimensionnelle et la notion du quantile usuel (en dimension un).

Soit $\mathbb{S}^0 = \{-1, +1\}$ la sphère unité sur \mathbb{R} , cet objet représente les deux seules directions possibles dans un espace univarié, une différence majeure avec le cas des dimensions supérieures où il existe une infinité non dénombrable de directions.

Tout d'abord, rappelons encore une fois la fonction quantile en dimension un

$$F_X^{-1}(t) = \inf \{t \in \mathbb{R} : P((-\infty, t]) \geq \alpha\}.$$

ainsi que sa version empirique

$$\begin{aligned} F_{n,X}^{-1}(t) &= \inf \{t \in \mathbb{R} : P_n((-\infty, t]) \geq \alpha\} \\ &= X_{(\lceil nt \rceil)}. \end{aligned}$$

où $X_{(k)}$ la k -ième statistique d'ordre et $\lceil \cdot \rceil$ la partie entière par excès. En dimension un, pour $O \in \mathbb{R}$, $u \in \mathbb{S}^0$ et $y \in \mathbb{R}$ on a :

$$\begin{aligned} H(O, u, y) &= \{x \in \mathbb{R} : \langle x - O, u \rangle \leq y\} \\ &= \{x \in \mathbb{R} : \langle x, u \rangle - \langle O, u \rangle \leq y\} \\ &= \{x \in \mathbb{R} : x \cdot u \leq y + \langle O, u \rangle\}. \end{aligned}$$

Ainsi que

$$\begin{aligned} Y(O, u, \alpha) &= \inf \{y \in \mathbb{R} : P(H(O, u, y)) \geq \alpha\} \\ &= \inf \{y \in \mathbb{R} : \mathbb{P}(X \in H(O, u, y)) \geq \alpha\} \\ &= \inf \{y \in \mathbb{R} : \mathbb{P}(\langle X, u \rangle \leq y + \langle O, u \rangle) \geq \alpha\} \\ &= \inf_{z \in \mathbb{R}} \{z - \langle O, u \rangle : \mathbb{P}(\langle X, u \rangle \leq z) \geq \alpha\} \\ &= \inf \{z \in \mathbb{R} : \mathbb{P}(\langle X, u \rangle \leq z) \geq \alpha\} - \langle O, u \rangle \\ &= F_{\langle X, u \rangle}^{-1}(\alpha) - \langle O, u \rangle. \end{aligned}$$

donc

$$\begin{aligned} Q(O, u, \alpha) &= O + Y(O, u, \alpha) \cdot u \\ &= O + \left(F_{\langle O, u \rangle}^{-1}(\alpha) - O \cdot u \right) \cdot u \\ &= F_{\langle O, u \rangle}^{-1}(\alpha) \cdot u \end{aligned}$$

On remarque que $Q(O, u, \alpha)$ ne dépend pas du choix de O une autre différence majeure avec le cas multidimensionnel.

Du fait que $S^0 = \{-1, +1\}$ et que $Q(O, u, \alpha)$ ne dépend pas de O on peut expliciter complètement Q en fonction seulement de α , ainsi pour $u = +1$ on a

$$Q(O, +1, \alpha) = F_X^{-1}(\alpha)$$

et pour $u = -1$ on a

$$\begin{aligned} Q(O, -1, \alpha) &= -F_{-X}^{-1}(\alpha) \\ &= -\inf \{y \in \mathbb{R} : \mathbb{P}(-X \leq y) \geq \alpha\} \\ &= -\inf \{y \in \mathbb{R} : \mathbb{P}(X \geq -y) \geq \alpha\} \\ &= -\inf \{-z \in \mathbb{R} : \mathbb{P}(X \geq z) \geq \alpha\} \\ &= \sup \{z \in \mathbb{R} : \mathbb{P}(X \leq z) \leq 1 - \alpha\}. \end{aligned}$$

Si F_X^{-1} est continue alors

$$\begin{aligned} Q(O, -1, \alpha) &= \inf \{z \in \mathbb{R} : \mathbb{P}(X \leq z) \geq 1 - \alpha\} \\ &= F_X^{-1}(1 - \alpha) \end{aligned}$$

donc, on a la relation

$$Q(O, -1, \alpha) = Q(O, 1, 1 - \alpha).$$

par conséquent, $Q(O, u, \alpha)$ est bien une généralisation du quantile réel à \mathbb{R}^d . Sous cet angle il est naturel de généraliser les résultats du processus des quantiles empirique du chapitre 1 dans le cadre multidimensionnel.

2.3 Théorèmes de convergence

Dans cette partie, on énonce les principaux théorèmes de convergence du processus des quantiles empirique, mais tout d'abord on regroupe et commente dans la première section les conditions suffisantes aux théorèmes des sections suivantes.

2.3.1 Hypothèses

On s'intéresse aux conditions peu restrictives et parfois minimales garantissant les convergences et les vitesses énoncées dans les théorèmes.

On pose $\Delta = [\alpha^-, \alpha^+] \subset [1/2, 1)$.

(H₀⁻) Pour tout $u \in \mathbb{S}_{d-1}$ et tout $\alpha \in \Delta$ on suppose que

$$P(\partial H(u, \alpha)) = 0.$$

(H₀⁺) Pour tout $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et tout $y, z \in \mathcal{Y}_{Delta}(O, u)$ l'ensemble des distances quantiles autorisées vu de O dans chaque direction dans 2.14, on suppose que

$$P(H(O, u, y, z)) > 0, \quad y < z.$$

Ces deux hypothèses suffisent pour monter le théorème de consistance uniforme et la loi du logarithme itéré. Ensuite, voici une autre hypothèse, qui sera exigée dans la suite et qu'on appelle « Hypothèse de régularité directionnelle ».

(H₁) Pour tout $u \in \mathbb{S}_{d-1}$, la variable aléatoire $\langle X, u \rangle$ admet une densité $f_{\langle X, u \rangle}$ continue sur $F_{\langle X, u \rangle}^{-1}(\Delta)$ avec $f_{\langle X, u \rangle} > 0$, de plus, on suppose qu'il existe m et M tels que $0 < m < M < +\infty$ et

$$m \leq \inf_{\alpha \in \Delta} \inf_{u \in \mathbb{S}_{d-1}} h_u(\alpha) \leq \sup_{\alpha \in \Delta} \sup_{u \in \mathbb{S}_{d-1}} h_u(\alpha) \leq M$$

où $h_u = f_{\langle X, u \rangle} \circ F_{\langle X, u \rangle}^{-1}$.

Notons qu'en aucun cas dans **(H₁)**, on suppose que la loi P est absolument continue par rapport à la mesure de Lebesgue dans \mathbb{R}^d , néanmoins, la condition **(H₁)** implique les deux conditions précédentes **(H₀⁻)** et **(H₀⁺)** avec en plus

$$m(z - y) \leq P(H(0_d, u, y, z)) \leq M(z - y), \quad y < z$$

ce qui se traduit sur la fonction $\Psi : \varepsilon \mapsto \inf_{B \in \mathcal{B}_\varepsilon} P(B)$ par

$$m\varepsilon \leq \Psi(\varepsilon) \leq M\varepsilon \quad \varepsilon \geq 0.$$

Afin de donner la condition **(H₂)** garantissant la convergence faible, pour tout $0 < \gamma < \gamma_0$, on pose

$$\rho(\gamma) = \sup_{|\varepsilon'| < \gamma} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} \left| F_{\langle X, u \rangle}(Y(O, u, \alpha) + \varepsilon') - \alpha - h(u, \alpha)\varepsilon' \right|. \quad (2.5)$$

Dans la preuve du théorème limite central uniforme, cette quantité apparaît naturellement dans le développement limité de $F_{\langle X, u \rangle}$ dans un voisinage

γ de $Y(O, u, \alpha)$. Les quatre hypothèses suivantes sont dites de « régularité inter-directionnelle ».

(H₂) Pour $\gamma > 0$ on suppose que

$$\lim_{\gamma \rightarrow 0} \frac{\sqrt{\log \log(1/\gamma)}}{\gamma} \rho(\gamma) = 0.$$

Ensuite, afin d'assurer l'approximation du type Bahadur-Kiefer, voici une condition plus exigeante que **(H₂)**.

(H₃) Pour $\gamma > 0$ on suppose que

$$\lim_{\gamma \rightarrow 0} \frac{\rho(\gamma) \sqrt{\log \log(1/\gamma)}}{\gamma^{3/2} \sqrt{\log(1/\gamma)}} = 0.$$

Voici deux autres conditions, plus exigeantes mais plus simples que **(H₃)**, assurant aussi l'approximation du type Bahadur-Kiefer,

(H'₃) Il existe $r > 1/2$ et $C^* > 0$ tel que pour tout Pour $0 < \gamma < \gamma_0$ on suppose que on a

$$\rho(\gamma) \leq C^* \gamma^{1+r}.$$

Sous l'hypothèse suivante la condition **(H'₃)** est vraie avec $r = 1$.

(H₄) La fonction $h : (u, \alpha) \mapsto h_u(\alpha)$ est différentiable sur $S_{d-1} \times \Delta$ avec des dérivées uniformément bornées.

Remarque 2.35. On a les implications suivantes

$$(\mathbf{H}_4) \Rightarrow (\mathbf{H}'_3) \Rightarrow (\mathbf{H}_3) \Rightarrow (\mathbf{H}_2).$$

2.3.2 Consistance uniforme

Le théorème de consistance forte suivant se décline sous plusieurs formes au corollaire 2.47

Théorème 2.36. Sous la condition **(H₀⁻)** on l'équivalence entre **(H₀⁺)** et

$$\limsup_{n \rightarrow \infty} \sup_{\alpha \in \Delta} \sup_{u \in S_{d-1}} |Y_n(O, u, \alpha) - Y(O, u, \alpha)| = 0 \quad p.s.$$

Remarque 2.37. Sous **(H₀⁻)** on a aussi l'équivalence entre **(H₀⁺)** et la continuité en $(u, \alpha) \in S_{d-1} \times \Delta$ des surfaces quantiles de P d'après la proposition 2.17. Ainsi, sous **(H₀⁻)** on a l'équivalence entre la continuité des surfaces et la consistance uniforme. Mais rappelons que la surface empirique n'est pas une surface continue en α , alors que sa cible $Q(O, \alpha)$ l'est.

2.3.3 Loi du logarithme itéré

Voici un résultat de convergence presque sûre avec vitesse sur l'écart en norme supérieure entre les quantiles directionnels et leur version empirique, communément appelé « loi du logarithme itéré ». C'est un résultat intermédiaire entre la consistance uniforme et la convergence faible uniforme.

Théorème 2.38. *Sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+)*

$$\limsup_{n \rightarrow \infty} \frac{\|Y_n - Y\|_{\mathbb{S}_{d-1} \times \Delta}}{\Psi_{-1} \left(\sqrt{\frac{\log \log n}{n}} \right)} < \infty \text{ p.s.}$$

Remarque 2.39. *En renforçant légèrement les hypothèses en supposant la condition (\mathbf{H}_1) , la loi du logarithme itéré peut être écrite sous la forme plus classique suivante*

$$\limsup_{n \rightarrow \infty} \frac{\|Y_n - Y\|_{\mathbb{S}_{d-1} \times \Delta}}{\sqrt{\frac{\log \log n}{n}}} < \infty \text{ p.s.}$$

Dans le cas d'une loi C-symétrique, on retrouve exactement la loi du logarithme itéré pour β_n sur \mathbb{R} .

2.3.4 Principe d'invariance faible

Avant d'énoncer le théorème limite central, voici la définition du processus gaussien limite.

Processus gaussiens utiles

1. Soit \mathbb{G} le P -pont brownien indexé par les demi-espaces i.e. le processus gaussien de moyenne 0 sur \mathcal{H} et de covariance

$$\text{cov}(\mathbb{G}(H), \mathbb{G}(H')) = P(H \cap H') - P(H)P(H')$$

pour tout $(H, H') \in \mathcal{H} \times \mathcal{H}$.

2. Soit $O \in \mathbb{R}^d$ et $u \in \mathbb{S}_{d-1}$ on définit $\mathbb{G}_{O,u}$

$$y \mapsto \mathbb{G}_{O,u}(y) := \mathbb{G}(H(O, u, y)).$$

Pour tout $O \in \mathbb{R}^d$, $u, u' \in \mathbb{S}_{d-1}$ et $y, y' \in \mathbb{R}$ on a

$$\mathbb{E}(\mathbb{G}_{O,u}(y)) = \mathbb{E}(\mathbb{G}(H(O, u, y))) = 0$$

et

$$\begin{aligned} & cov(\mathbf{G}_{O,u}(y), \mathbf{G}_{O,u'}(y')) \\ &= P(H(O, u, y) \cap H(O, u', y')) - P(H(O, u, y))P(H(O, u', y')) \\ &= \min(F_{\langle X-O, u \rangle}(y), F_{\langle X-O, u' \rangle}(y')) - F_{\langle X-O, u \rangle}(y)F_{\langle X-O, u' \rangle}(y'). \end{aligned}$$

Définition 2.40 (Processus limite). Soit $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$ on définit $\tilde{\mathbf{G}}(u, \alpha)$

$$\tilde{\mathbf{G}}(u, \alpha) = \frac{\mathbf{G}(H(u, \alpha))}{h(u, \alpha)}.$$

Noter que pour $u, u' \in \mathbb{S}_{d-1}$ et $\alpha, \alpha' \in \Delta$ on a $\mathbb{E}(\tilde{\mathbf{G}}(u, \alpha)) = 0$ et

$$cov(\tilde{\mathbf{G}}(u, \alpha), \tilde{\mathbf{G}}(u', \alpha')) = \frac{P(H(u, \alpha) \cap H(u', \alpha')) - \alpha\alpha'}{h(u, \alpha)h(u', \alpha')}.$$

Théorème limite central uniforme

Afin de construire des bandes de confiance, ou des épaisseurs, au tour des surfaces quantiles empirique il est important de disposer du TLC suivant. Les régions de confiance s'obtiennent soit par Monte-Carlo soit en prenant des grilles points et calculant des bandes jointes issues de leur covariance donnée au Corollaire 2.48.

Théorème 2.41. Sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , le processus $\sqrt{n}(Y_n - Y)$ converge faiblement vers $\tilde{\mathbf{G}}$ sur l'ensemble des fonctions bornées sur $\mathbb{S}_{d-1} \times \Delta$ muni de la norme uniforme.

Les covariances asymptotiques entre les quantiles univariés directionnels sont donc explicites, ce qui fournit une généralisation spatiale inédite au Théorème 1.9 bien connu sur \mathbb{R} . Quelques propriétés des corrélations interquantiles asymptotiques sont évoquées en remarque 2.49.

Corollaire 2.42. Sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , le processus $\sqrt{n}(Q_n - Q)$ indexé $\mathbb{S}_{d-1} \times \Delta$ converge faiblement vers $O + \tilde{\mathbf{G}} \cdot u$ sur $(\mathbb{R}^d, \|\cdot\|)$.

La preuve de la convergence faible n'utilise pas une approche par tension mais exploite la géométrie du problème et procède par couplage avec le processus empirique donné au Théorème 2.45 et l'approximation forte donnée au Théorème 2.43.

2.3.5 Principe d'invariance fort uniforme avec vitesse

Notre technique de contrôle des quantiles directionnels permet fondamentalement de se ramener au processus empirique indexé par les demi-espaces, et d'appliquer le résultat général d'approximation de ce dernier par couplage avec un processus Gaussien de même covariance obtenu dans [BM06a]. C'est pourquoi nous héritons ici des vitesses d'approximations qui dépendent de la dimension d mais pas de la loi P .

Théorème 2.43. *Sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite $\tilde{\mathbb{G}}_n$ de versions de $\tilde{\mathbb{G}}$ de sorte que*

$$Y_n(O, u, \alpha) = Y(O, u, \alpha) + \frac{\tilde{\mathbb{G}}_n(u, \alpha)}{\sqrt{n}} + \frac{Z_n(u, \alpha)}{\sqrt{n}}, \quad O \in \mathbb{R}^d, \quad \alpha \in \Delta, \quad u \in \mathbb{S}_{d-1}, \quad (2.6)$$

où $Z_n = \sqrt{n}(Y_n - Y) - \tilde{\mathbb{G}}_n$ vérifie

$$\lim_{n \rightarrow \infty} \|Z_n\|_{\mathbb{S}_{d-1} \times \Delta} = \lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |Z_n(u, \alpha)| = 0 \quad p.s.$$

Si de plus P satisfait (\mathbf{H}_3) alors on peut construire \mathbb{G}_n tel que pour $v_d = 1/(2 + 10d)$, $w_d = (4 + 10d)/(4 + 20d)$, il existe $n_\omega > 0$ tel que tout pour $n > n_\omega$,

$$\|Z_n\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_d}}{n^{v_d}} \right).$$

Remarque 2.44. *Il ne semble pas possible de s'affranchir de la dépendance en d étant donné que toutes les directions de l'espace entrent en ligne de compte, or dans [BM06b] les vitesses indépendantes de la dimension interviennent sur des classes pauvres, comme les cadrans et non pas les demi-espaces, et pour des lois P absolument continues sur \mathbb{R}^d à densité bornées, ce qui sort de notre cadre non paramétrique au sens assez strict.*

2.3.6 Représentation de type Bahadur-Kiefer

De façon surprenante la vitesse d'approximation de Bahadur-Kiefer sur \mathbb{R}^d est la même que sur \mathbb{R} . Ceci est dû à l'hypothèse de régularité inter-directionnelle, néanmoins ne pas supposer de densité pour P sur \mathbb{R}^d donne toute son ampleur au résultat suivant.

Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$ on note

$$\mathbb{E}_n(u, \alpha) = \sqrt{n} (P_n(H(u, \alpha)) - P(H(u, \alpha))).$$

Théorème 2.45. *Sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , on a*

$$\lim_{n \rightarrow \infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

Sans condition supplémentaire on a,

$$\left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\left(n^{1/2} \rho(\gamma_n) \right) \vee \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right).$$

avec $\gamma_n = \sqrt{\frac{\log \log n}{n}}$ et ρ définit en 2.5 dans la section 2.3.1.

Si de plus P satisfait (\mathbf{H}_3) ou (\mathbf{H}'_3) alors il existe $n_\omega > 0$ tel que tout pour $n > n_\omega$,

$$\left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right).$$

Remarque 2.46.

- *On rappelle que la condition (\mathbf{H}'_3) implique la condition (\mathbf{H}_3) .*
- *D'après les preuve la dimension d apparait dans la constante du $O_{p.s.}$.*
- *L'approche utilisée dans les démonstrations est originale y compris dans le cas $d = 1$, elle exploite directement la définition du quantile par un minimum et le traduit en terme de trajectoire du processus empirique.*

2.4 Corollaires

Dans la suite, on note $\Lambda_n = \sqrt{n}(P_n - P)$ le processus empirique indexé par \mathcal{H} .

Corollaire 2.47. *Sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) , on a les convergences suivantes*

- *Consistance uniforme pour les points quantiles*

$$\lim_{n \rightarrow \infty} \sup_{O \in \mathbb{R}^d} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} \|Q_n(O, u, \alpha) - Q(O, u, \alpha)\|_d = 0 \quad p.s.$$

- *Consistance uniforme pour les surfaces quantiles*

$$\lim_{n \rightarrow \infty} \sup_{O \in \mathbb{R}^d} \sup_{\alpha \in \Delta} d_H(Q_n(O, \alpha), Q(O, \alpha)) = 0 \quad p.s.$$

– Consistance uniforme de l'approximation de Bahadur-Kiefer associée aux points quantiles

$$\lim_{n \rightarrow \infty} \sup_{O \in \mathbb{R}^d} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} \left\| \sqrt{n} (Q_n(O, u, \alpha) - Q(O, u, \alpha)) + \Lambda_n(H(u, \alpha))u \right\|_d = 0 \quad p.s.$$

Corollaire 2.48. Soit $(O_1, u_1, \alpha_1), \dots, (O_k, u_k, \alpha_k)$ dans $\mathbb{R}^d \times \mathbb{S}_{d-1} \times \Delta$. Sous les conditions (H_1) et (H_2) , on a

$$\sqrt{n} \begin{pmatrix} Y_n(O_1, u_1, \alpha_1) - Y(O_1, u_1, \alpha_1) \\ \dots \\ Y_n(O_k, u_k, \alpha_k) - Y(O_k, u_k, \alpha_k) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0_k, \Sigma)$$

avec Σ la matrice de variances-covariances définie par

$$\Sigma_{i,j} = \frac{P(H(O_i, u_i, Y(O_i, u_i, \alpha_i)) \cap H(O_j, u_j, Y(O_j, u_j, \alpha_j))) - \alpha_i \alpha_j}{f_{\langle X-O_i, u_i \rangle} \circ F_{\langle X-O_i, u_i \rangle}^{-1}(\alpha_i) \cdot f_{\langle X-O_j, u_j \rangle} \circ F_{\langle X-O_j, u_j \rangle}^{-1}(\alpha_j)}.$$

Remarque 2.49 (Corrélations). Soit $(u, u') \in \mathbb{S}_{d-1} \times \mathbb{S}_{d-1}$ et $(\alpha, \alpha') \in \Delta \times \Delta$ on a

$$\begin{aligned} c(u, \alpha, u', \alpha') &= \text{corr}(\tilde{G}(u, \alpha), \tilde{G}(u', \alpha')) \\ &= \frac{P(H(u, \alpha) \cap H(u', \alpha')) - \alpha \alpha'}{\sqrt{\alpha(1-\alpha)} \sqrt{\alpha'(1-\alpha')}} \\ &= \text{corr}(\Lambda_n(H(u, \alpha)), \Lambda_n(H(u', \alpha'))) \end{aligned}$$

en fixant $\alpha, \alpha' \in \Delta$ on a

$$\begin{aligned} -\sqrt{\frac{(1-\alpha)(1-\alpha')}{\alpha\alpha'}} &= \frac{\alpha + \alpha' - 1 - \alpha\alpha'}{\sqrt{\alpha(1-\alpha)} \sqrt{\alpha'(1-\alpha')}} \\ &\leq c(u, \alpha, u', \alpha') \\ &\leq \frac{\alpha \vee \alpha' - \alpha\alpha'}{\sqrt{\alpha(1-\alpha)} \sqrt{\alpha'(1-\alpha')}} = \sqrt{\frac{(\alpha \vee \alpha')(1 - \alpha \wedge \alpha')}{(\alpha \wedge \alpha')(1 - \alpha \vee \alpha')}}. \end{aligned}$$

où la borne inférieure est atteinte pour $u = -u'$ et la borne supérieure est atteinte pour $u = u'$.

2.5 Preuves

2.5.1 Preuve de la consistance

Preuve du Théorème 2.36. Pour $\varepsilon > 0$ on considère \mathcal{B}_ε l'ensemble des bandes admissibles de largeur ε , défini dans 2.15.

Preuve par l'absurde : Sous (\mathbf{H}_0^-) et $(H_{0,\Psi}^+)$ supposons qu'il existe $\delta > 0$ et une suite aléatoire croissante $n_k \rightarrow \infty$ telle que

$$|Y_{n_k}(O, u_{n_k}, \alpha_{n_k}) - Y(O, u_{n_k}, \alpha_{n_k})| > \delta.$$

Soit $(u_{n'_k}, \alpha_{n'_k})$ une suite extraite de $\mathbb{S}_{d-1} \times [\alpha^-, \alpha^+] \subset \mathbb{S}_{d-1} \times (0, 1)$ avec $u_{n'_k} \neq u_0$ et $\alpha_{n'_k} \neq \alpha_0$ et $\lim_{k \rightarrow \infty} u_{n'_k} = u_0$ et $\lim_{k \rightarrow \infty} \alpha_{n'_k} = \alpha_0$. Deux configurations sont possible, ou bien il existe $(m_k)_{k \geq 1}$ une suite croissante avec $\lim_{k \rightarrow \infty} m_k = +\infty$ tel que

$$Y(O, u_{m_k}, \alpha_{m_k}) - Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$$

ou bien $Y_{m'_k}(O, u_{m_k}, \alpha_{m_k}) - Y(O, u_{m_k}, \alpha_{m_k}) \geq \delta$.

Ainsi d'un côté, pour $(m_k)_{k \geq 1}$ une suite croissante avec $\lim_{k \rightarrow \infty} m_k = +\infty$ tel que $Y(O, u_{m_k}, \alpha_{m_k}) - Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$, on pose

$$A_k = H_{m_k}(O, u_{m_k}, \alpha_{m_k})$$

$$C_k = H(O, u_{m_k}, \alpha_{m_k})$$

$$B_k = C_k \setminus A_k$$

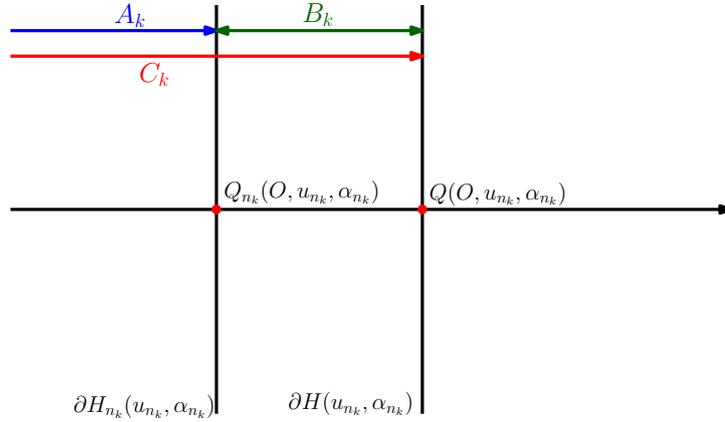


FIGURE 2.8 – Cas 1 : $Y(O, u_{m_k}, \alpha_{m_k}) - Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$

puisque \mathcal{H} est une VC-classe presque sûrement alors

$$\lim_{n \rightarrow \infty} \sup_{H \in \mathcal{H}} |P_n(H) - P(H)| = 0, \text{ p.s.}$$

puisque (\mathbf{H}_0^-) est vérifiée alors par la proposition 2.34, on a

$$\begin{aligned}
\sup_{H \in \mathcal{H}} |P_{m_k}(H) - P(H)| &\geq P_{m_k}(A_k) - P(A_k) \\
&\geq \alpha_{m_k} - \frac{d}{m_k} - (\alpha_{m_k} - P(B_k)) \\
&\geq -\frac{d}{m_k} + P(B_k) \\
&\geq -\frac{d}{m_k} + \Psi(\delta)
\end{aligned}$$

donc

$$\Psi(\delta) \leq \sup_{H \in \mathcal{H}} |P_{m_k}(H) - P(H)| + \frac{d}{m_k} \quad (2.7)$$

lorsque k tend vers l'infini, ceci implique qu'il existe $\delta > 0$ tel que

$$\Psi(\delta) = 0$$

or ceci contredit l'hypothèse $(H_{0,\Psi}^+)$.

D'un autre côté pour $(m'_k)_{k \geq 1}$ une suite croissante avec $\lim_{k \rightarrow \infty} m'_k = +\infty$ telle que $Y(O, u_{m'_k}, \alpha_{m_k}) \leq Y_{m'_k}(O, u_{m'_k}, \alpha_{m'_k})$

on pose

$$A'_k = H_{m'_k}(O, u_{m'_k}, \alpha_{m_k})$$

$$C'_k = H(O, u_{m'_k}, \alpha_{m'_k})$$

$$B'_k = H(O, u_{m'_k}, \alpha_{m'_k}) \setminus H_{m'_k}(O, u_{m'_k}, \alpha_{m'_k})$$

puisque \mathcal{H} est VC-classe on a

$$\lim_{n \rightarrow \infty} \sup_{H \in \mathcal{H}} |P_n(H) - P(H)| = 0, \text{ p.s.}$$

puisque (\mathbf{H}_0^-) est vérifiée alors par la proposition 2.34, on a

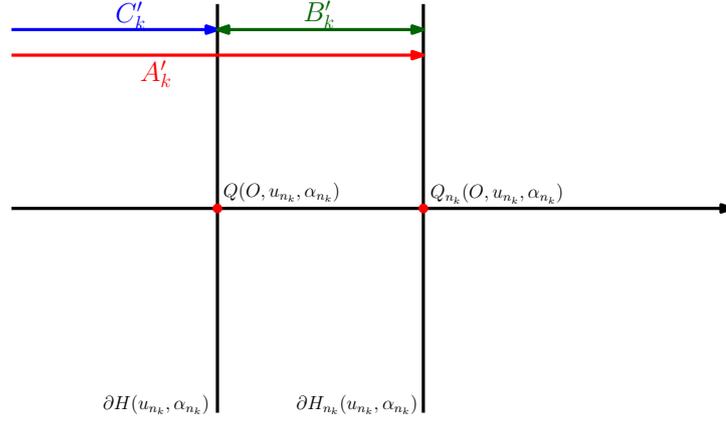


FIGURE 2.9 – Cas 2 : $Y_{m'_k}(O, u_{m_k}, \alpha_{m_k}) - Y(O, u_{m_k}, \alpha_{m_k}) \geq \delta$

$$\begin{aligned}
 \sup_{H \in \mathcal{H}} |P_{m'_k}(H) - P(H)| &\geq P(A'_k) - P_{m'_k}(A'_k) \\
 &\geq P(C'_k \cup B'_k) - P_{m'_k}(A'_k) \\
 &\geq \alpha_{m'_k} + P(B'_k) - (\alpha_{m'_k} + \frac{d}{m'_k}) \\
 &\geq P(B'_k) - \frac{d}{m'_k} \\
 &\geq \Psi(\delta) - \frac{d}{m'_k}
 \end{aligned}$$

donc de la même manière, on a

$$\Psi(\delta) \leq \sup_{H \in \mathcal{H}} |P_{m'_k}(H) - P(H)| + \frac{d}{m_k} \quad (2.8)$$

lorsque k tend vers l'infini, ceci implique qu'il existe $\delta > 0$ tel que

$$\Psi(\delta) = 0$$

or ceci contredit l'hypothèse $(H_{0,\Psi}^+)$.

Conclusion : Sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) on a

$$\lim_{n \rightarrow \infty} \sup_{\alpha \in \Delta} \sup_{u \in \mathbb{S}_{d-1}} |Y_n(O, u, \alpha) - Y(O, u, \alpha)| = 0 \quad p.s.$$

2.5.2 Preuve de la loi du logarithme itéré

Preuve par l'absurde : Sous (H_0^-) et $(H_{0,\Psi}^+)$ supposons qu'il existe une suite aléatoire croissante $n_k \rightarrow \infty$ tel que

$$|Y_{n_k}(O, u_{n_k}, \alpha_{n_k}) - Y(O, u_{n_k}, \alpha_{n_k})| > \delta_{n_k} = \Psi^{-1} \left(\sqrt{\frac{\log \log n_k}{n_k}} \right).$$

Soit $(u_{n'_k}, \alpha_{n'_k})$ une suite de $\mathbb{S}_{d-1} \times [\alpha^-, \alpha^+] \subset \mathbb{S}_{d-1} \times (0, 1)$ avec $u_{n'_k} \neq u_0$ et $\alpha_{n'_k} \neq \alpha_0$ et $\lim_{k \rightarrow \infty} u_{n'_k} = u_0$ et $\lim_{k \rightarrow \infty} \alpha_{n'_k} = \alpha_0$.

Cas I : il existe $(m_k)_{k \geq 1}$ une suite croissante avec $\lim_{k \rightarrow \infty} m_k = +\infty$ telle que

$$Y(O, u_{m_k}, \alpha_{m_k}) - Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$$

on pose

$$A_k = H_{m_k}(O, u_{m_k}, \alpha_{m_k})$$

$$C_k = H(O, u_{m_k}, \alpha_{m_k})$$

$$B_k = C_k \setminus A_k$$

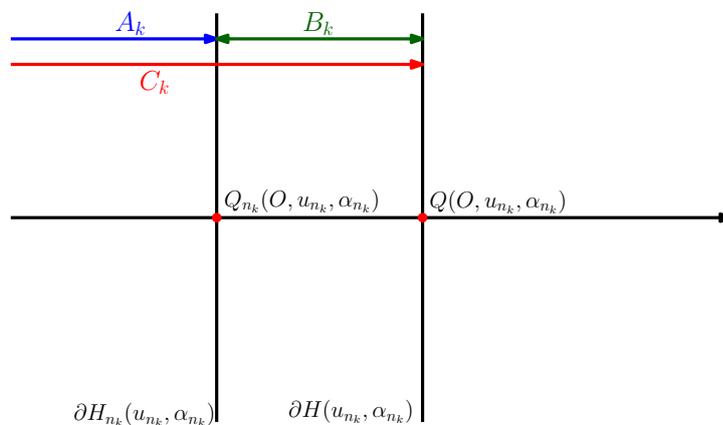


FIGURE 2.10 – Cas 1 : $Y(O, u_{m_k}, \alpha_{m_k}) - Y_{m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$

puisque \mathcal{H} est une VC-classe alors

$$\lim_{n \rightarrow \infty} \sup_{H \in \mathcal{H}} |P_n(H) - P(H)| = 0, \text{ p.s.}$$

et par la loi du logarithme itéré (Alexander 1984 [Ale84]) on a

$$\limsup_{n \rightarrow \infty} \frac{\|P_n - P\|_{\mathcal{H}}}{\sqrt{\frac{\log \log n}{n}}} \leq \sqrt{2}/2 \text{ p.s.}$$

en notant que $4/5 > \sqrt{2}/2$, il existe $k(\omega) > 0$ tel que pour tout $k \geq k(\omega)$

$$\frac{4}{5} \sqrt{\frac{\log \log m_k}{m_k}} \geq \sup_{H \in \mathcal{H}} |P_{m_k}(H) - P(H)|$$

puisque (H_0^-) est vérifiée alors par la Proposition 2.34, la continuité de Ψ de la Proposition 2.20 et la remarque 2.21 on a

$$\begin{aligned} \sup_{H \in \mathcal{H}} |P_{m_k}(H) - P(H)| &\geq P_{m_k}(A_k) - P(A_k) \\ &\geq \alpha_{m_k} - \frac{d}{m_k} - (\alpha_{m_k} - P(B_k)) \\ &\geq -\frac{d}{m_k} + P(B_k) \\ &\geq -\frac{d}{m_k} + \Psi(\delta_n) \\ &\geq -\frac{d}{m_k} + \Psi\left(\Psi^{-1}\left(\sqrt{\frac{\log \log m_k}{m_k}}\right)\right) \\ &\geq \sqrt{\frac{\log \log m_k}{m_k}} - \frac{d}{m_k} \end{aligned}$$

donc

$$\frac{4}{5} \sqrt{\frac{\log \log m_k}{m_k}} \geq \sup_{H \in \mathcal{H}} |P_{m_k}(H) - P(H)| \geq \sqrt{\frac{\log \log m_k}{m_k}} - \frac{d}{m_k} \quad (2.9)$$

ce qui implique que

$$1 \geq \frac{1}{5d} \sqrt{m_k \log \log m_k}$$

ce qui est absurde, donc

$$\limsup_{n \rightarrow \infty} \frac{\|Y_n - Y\|_{S_{d-1} \times \Delta}}{\Psi^{-1}\left(\sqrt{\frac{\log \log n}{n}}\right)} < \infty.$$

Dans le deuxième cas, la preuve est identique.

2.5.3 Outils pour les preuves principales

Dans cette partie on donne les outils nécessaires dans les preuves du théorème limite central, principe d'invariance fort et approximation de Bahadur-Kiefer.

Rappels

- **Suite de Rademacher** $1/2$: une suite de Rademacher $(\tau_i)_{i=1}^n$ est une suite de variables aléatoires indépendantes, de même loi telles que

$$\mathbb{P}(\tau_1 = 1) = \mathbb{P}(\tau_1 = -1) = 1/2.$$

- **Fonction enveloppe** : soit \mathcal{F} une classe de fonctions mesurables de \mathcal{X} dans \mathbb{R} . on appelle fonction enveloppe de \mathcal{F} (lorsqu'elle existe), la fonction mesurable F telle que pour tout $x \in \mathcal{X}$ et toute $f \in \mathcal{F}$ on a

$$|f(x)| \leq F(x).$$

- **Demi-métrique** d_μ : pour μ une mesure de probabilité sur $(\mathcal{X}, \mathcal{A})$, pour $f, h \in \mathcal{F}$, on pose la demi-métrique suivante

$$d_\mu^2(f, h) = \int_{\mathcal{X}} (f - h)^2 d\mu$$

et on note $\mu(f^2) = d_\mu^2(f, 0) = \int_{\mathcal{X}} f^2 d\mu$.

- **Nombre de L^2 -recouvrement** : soit $\varepsilon > 0$ et μ une mesure de probabilité sur $(\mathcal{X}, \mathcal{A})$, la quantité $N(\varepsilon, \mathcal{F}, d_\mu)$ correspond au nombre minimal de boules $\{f \in \mathcal{F} : d_\mu(f, h) < \varepsilon\}$ de d_μ -rayon ε et de centre $h \in \mathcal{F}$ nécessaire pour recouvrir \mathcal{F} .
- **Nombre de L^2 -recouvrement uniforme** :

$$N_F(\varepsilon, \mathcal{F}) = \sup_{\substack{\mu \text{ proba.} \\ 0 < \mu(F^2) < 1}} N(\varepsilon \sqrt{\mu(F^2)}, \mathcal{F}, d_\mu)$$

où F est une fonction enveloppe de \mathcal{F} .

- **Classe de Vapnik-Chervonenkis** : soit \mathcal{F} une classe de fonctions mesurables de \mathcal{X} dans \mathbb{R} , s'il existe $c > 0$ et $v_0 > 0$, et une fonction enveloppe F tels que

$$N_F(\varepsilon, \mathcal{F}) \leq c_0 \varepsilon^{-2(v_0-1)}$$

alors la classe \mathcal{F} est appelé Classe de Vapnik-Chervonenkis (VC-classe). Le plus petit v_0 est appelé dimension de la classe VC \mathcal{F} et il est noté $VC(\mathcal{F}) = v_0$.

Remarque 2.50. *On a fait le choix dans la définition de la classe de Vapnik–Chervonenkis, de ne pas donner la définition usuelle par les propriétés de pulvérisation combinatoire mais plutôt celle par le nombre d'entropie uniforme qui sert pour l'approximation des processus aléatoires.*

Outils

Soit \mathcal{F} une classe de fonctions mesurables de \mathcal{X} à valeurs réelles, on pose les conditions suivantes

(F.i) Pour $S_* > 0$ et pour toute fonction $f \in \mathcal{F}$, $\sup_{x \in \mathcal{X}} |f(x)| \leq S_*/2$.

(F.ii) La classe \mathcal{F} est mesurable point par point, i.e. il existe une sous-classe dénombrable \mathcal{F}_∞ de \mathcal{F} telle que pour toute fonction f il existe une suite $(f_m)_{m \in \mathbb{N}} \subset \mathcal{F}_\infty^{\mathbb{N}}$ avec $\lim_{m \rightarrow \infty} f_m(x) = f(x)$ pour tout $x \in \mathcal{X}$.

La condition **(F.ii)** est imposée afin d'éviter les problèmes de mesurabilité dans les théorèmes suivants et d'avoir recours à la mesure de probabilité extérieure.

Théorème 2.51 (Inégalité de Talagrand [Tal94]). *Soit \mathcal{G} une classe de fonctions satisfaisant les conditions **(F.i)** et **(F.ii)** alors il existe deux constantes $A_0 > 0$ et $A_1 > 0$ telles que pour tout $n \geq 1$ et $t > 0$,*

$$\begin{aligned} \mathbb{P} \left\{ \|\alpha_n\|_{\mathcal{G}} \geq A_0 \left(\mathbb{E} \left(\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i g(X_i) \right\|_{\mathcal{G}} \right) + t \right) \right\} \\ \leq 2 \exp \left(-\frac{A_1 t^2}{\sigma_{\mathcal{G}}^2} \right) + 2 \exp \left(-\frac{A_1 t \sqrt{n}}{S_*} \right) \end{aligned}$$

avec $\sigma_{\mathcal{G}}^2 = \sup_{g \in \mathcal{G}} \text{Var}(g(X))$, et S_* la borne dans **(F.i)**

Dans le théorème précédent les constantes A_0, A_1 sont universelles et donc ne dépendent ni de \mathcal{G} ni de S_* .

Remarque 2.52. *Soit $\varepsilon > 0$ on rappelle que*

$$\mathcal{B}_\varepsilon = \left\{ H(O, u, y, y + \varepsilon) : O \in \mathbb{R}^d, u \in \mathbb{S}_{d-1}, y, y + \varepsilon \in \mathcal{Y}_\Delta(O, u) \right\}.$$

Soit $n \geq 3$, $C > 1$, pour $\varepsilon_n = C \sqrt{\frac{\log \log n}{n}}$, on pose

$$\mathcal{B}_n = \bigcup_{0 < \varepsilon < \varepsilon_n} \mathcal{B}_\varepsilon$$

et

$$\mathcal{F}_n = \{ \mathbb{1}_B : B \in \mathcal{B}_n \}.$$

Il est facile de voir que \mathcal{F}_n vérifie les conditions **(F.i)** et **(F.ii)**, en effet

– pour toute fonction $g \in \mathcal{F}_n$ on a $\sup_{x \in \mathcal{X}} |g(x)| \leq 1 = 2/2$, ainsi avec les notations de **(F.i)** on a

$$S_* = 2$$

– pour tout $\varepsilon > 0$, $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $y \in \mathcal{Y}_{\text{Delta}}(O, u)$ tels que la bande $H(O, u, y, y + \varepsilon) \in \mathbb{B}_n$ il existe une suite de rationnels $\delta_k \rightarrow \varepsilon$, une suite de vecteurs $u_k \rightarrow u$ dans $\mathbb{Q}_{d-1} = \{v \in \mathbb{Q}^2 : \|v\|_2 = 1\}$ et une suite de rationnels $y_k \rightarrow y$ telles que pour tout $x \in \mathbb{R}^d$ on a

$$\lim_{k \rightarrow \infty} g_k(x) = \lim_{k \rightarrow \infty} \mathbb{1}_{H(O, u_k, y_k, y_k + \delta_k)}(x) = g(x) = \mathbb{1}_{H(O, u, y, y + \delta)}(x)$$

Théorème 2.53 (Inégalité des moments [EM00], [GG01]). Soit \mathcal{G} une classe de fonction satisfaisant les conditions **(F.i)** et **(F.ii)** admettant une fonction enveloppe noté G . S'il existe $\beta, v, c > 1$ et $\sigma \leq 1/(8c)$ tel que les conditions suivantes sont vérifiées

$$\mathbb{E}(G^2(X)) \leq \beta^2; \quad N_G(\varepsilon, \mathcal{G}) \leq c\varepsilon^{-v}, 0 < \varepsilon < 1; \quad \sup_{g \in \mathcal{G}} \mathbb{E}(g^2(X)) \leq \sigma^2;$$

et

$$\sup_{g \in \mathcal{G}} \sup_{x \in \mathcal{X}} |g(x)| \leq \frac{\sqrt{n\sigma^2 / \ln(\beta \vee 1/\sigma)}}{2\sqrt{v+1}}$$

alors on a pour une constante universelle A_2 indépendante de β ,

$$\mathbb{E} \left(\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i g(X_i) \right\|_{\mathcal{G}} \right) \leq A_2 \sqrt{v\sigma^2 \ln(\beta \vee 1/\sigma)}.$$

Remarque 2.54. Soit $n \geq 3$, $g \in \mathcal{F}_n$ et $G = 1$ la fonction enveloppe de \mathcal{F}_n , on sait que

$$\mathbb{E}(G(X)^2) = 1 \leq \beta = 2$$

ensuite sous **(H₁)** pour tout $B \in \mathcal{B}_n$ on a

$$E(\mathbb{1}_{X \in B}^2) = P(B) \leq M\varepsilon_n =: \theta_n^2$$

puisque \mathcal{B}_n est une VC classe de dimension $2d + 1$, alors il existe $c > 1$

$$N(\varepsilon, \mathcal{F}_n) \leq c\varepsilon^{-v}, 0 < \varepsilon < 1$$

avec $v = 2((2d + 1) - 1) = 4d$. Enfin, on sait qu'il existe $n_0 > 0$ tel que pour tout $n > n_0$ on a

$$\frac{1}{\theta_n} = \frac{1}{\sqrt{M\varepsilon_n}} = \left(\frac{1}{MC} \sqrt{\frac{n}{\log \log n}} \right)^{1/2} = \frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log \log n)^{1/4}} > 2.$$

Par conséquent on a $\ln(\beta \vee 1/\theta_n) = \log(2 \vee 1/\theta_n) = \log(1/\theta_n)$ et

$$n\theta_n^2 = CM\sqrt{n \log \log n}$$

donc

$$\frac{n\theta_n^2}{\log(\beta \vee \frac{1}{\theta_n})} = \frac{CM\sqrt{n \log \log n}}{\log\left(\frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log \log n)^{1/4}}\right)}$$

d'où

$$\lim_{n \rightarrow \infty} \frac{n\theta_n^2}{\log(\beta \vee \frac{1}{\theta_n})} = +\infty.$$

Puisque $\sup_{g \in \mathcal{F}_n} \sup_{x \in \mathcal{X}} |g(x)| = 1$ on en conclut qu'il existe $n_1 > n_0 > 0$ tel que pour $n \geq n_1$ on a

$$\sup_{g \in \mathcal{F}_n} \sup_{x \in \mathcal{X}} |g(x)| \leq \frac{\sqrt{n\theta_n^2 / \log(\beta \vee 1/\theta_n)}}{2\sqrt{v} + 1}.$$

Théorème 2.55 (Berthet et Mason 2006 [BM06a]). Soit \mathcal{G} une VC-classe de dimension $VC(\mathcal{G})$ satisfaisant les conditions **(F.i)** et **(F.ii)** et admettant une fonction enveloppe $G := S_*/2$. Pour tout $\lambda > 1$ il existe $\rho(\lambda) > 1$ tel que pour tout $n \geq 1$ on peut construire sur un même espace de probabilité, les vecteurs X_1, \dots, X_n et une suite (\mathbf{G}_n) des versions de \mathbf{G} tel que

$$\mathbb{P}\{\|\alpha_n - \mathbf{G}_n\|_{\mathcal{G}} > \rho(\lambda)n^{-v_1}(\log n)^{v_2}\} \leq n^{-\lambda}$$

avec $v_1 = 1/(2 + 5v_0)$ et $v_2 = (4 + 5v_0)/(4 + 10v_0)$ où $v_0 = 2(VC(\mathcal{G}) - 1)$ ainsi que \mathbf{G} un P -pont brownien indexé par \mathcal{G} i.e. le processus gaussien centré de covariance définie, pour tout $f, g \in \mathcal{G}$ par

$$\text{cov}(\mathbf{G}(f), \mathbf{G}(g)) = \mathbb{E}(f(X)g(X)) - \mathbb{E}(f(X))\mathbb{E}(g(X)).$$

Notations Afin de faciliter l'écriture des processus empiriques indexés de différentes façons, on introduit les notations suivantes :

– Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $y \in \mathbb{R}$

$$\alpha_n(O, u, y) = \sqrt{n} (P_n(H(O, u, y)) - P(H(O, u, y))).$$

– Pour $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$

$$\mathbb{E}_n(u, \alpha) = \sqrt{n} (P_n(H(u, \alpha)) - P(H(u, \alpha))).$$

– Pour $H \in \mathcal{H}$

$$\Lambda_n(H) = \sqrt{n}(P_n(H) - P(H))$$

– Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$

$$\mathbb{D}_n(u, \alpha) = \sqrt{n} (Y_n(O, u, \alpha) - Y(O, u, \alpha)).$$

Remarque 2.56. Avec les notations précédentes, pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$

$$\alpha_n(O, u, Y(O, u, \alpha)) = \mathbb{E}_n(u, \alpha) = \Lambda_n(H(u, \alpha)).$$

Remarque 2.57. Pour $H(O, u, y, y + \varepsilon) \in \mathcal{B}_\varepsilon$ on a

$$\begin{aligned} \Lambda_n(H(O, u, y, y + \varepsilon)) &= \sqrt{n} (P_n(H(O, u, y, y + \varepsilon)) - P(H(O, u, y, y + \varepsilon))) \\ &= \sqrt{n} (P_n(H(O, u, y + \varepsilon) \setminus H(O, u, y)) - P(H(O, u, y + \varepsilon) \setminus H(O, u, y))) \\ &= \Lambda_n(H(O, u, y + \varepsilon)) - \Lambda_n(H(O, u, y)). \end{aligned}$$

Notation 2.58. Soit $n \geq 3$, $C > 1$, pour $\varepsilon_n = C \sqrt{\frac{\log \log n}{n}}$, on pose

$$\mathcal{B}_n = \bigcup_{0 < \varepsilon < \varepsilon_n} \mathcal{B}_\varepsilon, \text{ et } \mathcal{F}_n = \{\mathbb{1}_B : B \in \mathcal{B}_n\}.$$

On peut à présent démontrer une proposition cruciale pour la preuve de nos résultats. Il s'agit du contrôle précis du module de continuité du processus empirique Λ_n pour des bandes de largeur plus petite que ε_n .

Proposition 2.59. Sous **(H₁)**, pour des constantes $C_0, C_1 > 0$, $\zeta > 1$ et pour tout $n \geq 3$ on a

$$P \left\{ \|\Lambda_n\|_{\mathcal{B}_n} \geq C_0 \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right\} \leq \frac{C_1}{n^\zeta}.$$

Démonstration. Soit $n \geq 3$, on sait que par la remarque 2.52 que \mathcal{F}_n est classe de fonctions vérifiant les conditions **(F.i)** et **(F.ii)** donc en appliquant l'inégalité de Talagrand [Tal94] il existe deux constantes positives $A_0, A_1 > 0$ telles que

$$\begin{aligned} \mathbb{P} \left\{ \|\Lambda_n\|_{\mathcal{B}_n} \geq A_0 \left(\mathbb{E} \left(\sup_{B \in \mathcal{B}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in B} \right| \right) + t_n \right) \right\} \\ \leq 2 \exp \left(-\frac{A_1 t_n^2}{\sigma_n^2} \right) + 2 \exp \left(-A_1 t_n \sqrt{n} \right) \end{aligned}$$

avec $\sigma_n^2 = \sup_{B \in \mathcal{B}_n} \text{Var}(\mathbb{1}_{X \in B})$ et $t_n = \sqrt{\frac{M}{A_1} \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}}}$.

Étape I. On souhaite montrer qu'il existe $C_1 > 0$ et $\zeta > 1$ telle que

$$2 \exp \left(-\frac{A_1 t_n^2}{\sigma_n^2} \right) + 2 \exp \left(-A_1 t_n \sqrt{n} \right) \leq \frac{C_1}{n^\zeta}.$$

On a par (\mathbf{H}_1)

$$\begin{aligned} \text{Var}(\mathbb{1}_{X \in B}) &= P(B)(1 - P(B)) \\ &\leq M\varepsilon_n(1 - m\varepsilon_n) \\ &\leq M\varepsilon_n =: \theta_n^2. \end{aligned}$$

donc

$$\begin{aligned} \exp\left(-\frac{A_1 t_n^2}{\sigma_n^2}\right) &\leq \exp\left(-\frac{A_1 t_n^2}{M\varepsilon_n}\right) \\ &= \exp\left(C \frac{n^{1/2}}{(\log \log n)^{1/2}} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}}\right)^2\right) \\ &= \exp(-C \log n) \\ &= \frac{1}{n^C}. \end{aligned}$$

De plus, on a,

$$\begin{aligned} \exp(-A_1 t_n \sqrt{n}) &= \exp\left(-A_1 \sqrt{\frac{M}{A_1}} \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \sqrt{n}\right) \\ &= \exp\left(-\sqrt{MA_1} (\log n)^{1/2} (n \log \log n)^{1/4}\right) \end{aligned}$$

car il existe $n(\omega) > 0$ tel que pour $n \geq n(\omega)$ il existe $K_1 > 1$ telle que

$$(\log n)^{1/2} (n \log \log n)^{1/4} \geq \frac{K_1}{\sqrt{MA_1}} \log n$$

donc

$$\begin{aligned} \exp(-A_1 t_n \sqrt{n}) &= \exp\left(-\sqrt{MA_1} (\log n)^{1/2} (n \log \log n)^{1/4}\right) \\ &\leq \exp(-K_1 \log n) \\ &\leq \frac{1}{n^{K_1}} \end{aligned}$$

Ainsi il existe $C_1 > 0$ et $\zeta > 1$ telle que

$$\mathbb{P}\left\{\|\Lambda_n\|_{\mathcal{B}_n} \geq A_0 \left(\mathbb{E}\left(\sup_{B \in \mathcal{B}_n} \left|\frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in B}\right|\right) + t_n\right)\right\} \leq \frac{C_1}{n^\zeta}.$$

Étape II. On s'intéresse à la quantité

$$\mathbb{E} \left(\sup_{B \in \mathcal{B}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in B} \right| \right)$$

On a par la remarque 2.54 on a \mathcal{F}_n vérifie toutes les conditions de l'inégalité des moments [GG01] donc il existe $A_2 > 0$ telle que pour tout $n \geq n_0$ on a

$$\mathbb{E} \left(\sup_{B \in \mathcal{B}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in B} \right| \right) \leq A_2 \sqrt{v\theta_n^2 \log(1 \vee 1/\theta_n)}$$

où $\theta_n =$

$$\begin{aligned} \mathbb{E} \left(\sup_{B \in \mathcal{B}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in B_n} \right| \right) &\leq A_2 \sqrt{4d\theta_n^2 \log(1/\theta_n)} \\ &= A_2 \sqrt{4dCM \frac{(\log \log n)^{1/2}}{n^{1/2}} \cdot \log \left(\frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log \log n)^{1/4}} \right)} \end{aligned}$$

on a

$$\log \left(\frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log \log n)^{1/4}} \right) = \log \left(\frac{1}{\sqrt{MC}} \right) + \log \left(\frac{n^{1/4}}{(\log \log n)^{1/4}} \right)$$

puis que $M > 0$ est un majorant de la borne supérieure de $(u, \alpha) \mapsto h(u, \alpha)$ alors on peut toujours le choisir $M > 1$ ainsi

$$\log \left(\frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log \log n)^{1/4}} \right) \leq \log \left(\frac{n^{1/4}}{(\log \log n)^{1/4}} \right)$$

donc

$$\begin{aligned} \mathbb{E} \left(\sup_{B \in \mathcal{B}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in B_n} \right| \right) &\leq A_2 \sqrt{4dCM \frac{(\log \log n)^{1/2}}{n^{1/2}} \cdot \log \left(\frac{1}{\sqrt{MC}} \cdot \frac{n^{1/4}}{(\log \log n)^{1/4}} \right)} \\ &\leq A_2 \sqrt{4dCM} \frac{(\log \log n)^{1/4}}{n^{1/4}} \cdot \sqrt{\log \left(\frac{n^{1/4}}{(\log \log n)^{1/4}} \right)} \end{aligned}$$

en remarquant que

$$\frac{\log \left(\frac{n^{1/4}}{(\log \log n)^{1/4}} \right)}{\log n} = \frac{\log n}{4 \log n} - \frac{\log \log \log n}{4 \log n} = \frac{1}{4} + o(1)$$

alors il existe $A_4 > 1$ et $\eta \geq 1$ tels que pour tout $n \geq \eta$

$$\sqrt{\log \left(\frac{n^{1/4}}{(\log \log n)^{1/4}} \right)} \leq A_4 (\log n)^{1/2}$$

donc

$$\mathbb{E} \left(\sup_{B \in \mathcal{B}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in B} \right| \right) \leq A_2 A_4 \sqrt{4dCM} \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}}$$

Ainsi il existe $C_0 = A_2 A_4 \sqrt{4dCM}$ et $C_1 > 0$ telles que

$$P \left\{ \|\Lambda_n\|_{\mathcal{B}_n} \geq C_0 \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right\} \leq \frac{C_1}{n^\zeta}.$$

□

Remarque 2.60. Pour $\zeta > 1$ la série de terme général $n^{-\zeta}$ converge donc par le lemme de Borell-Cantelli sous (\mathbf{H}_1) ,

$$\|\Lambda_n\|_{\mathcal{B}_n} = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right).$$

Lemme 2.61. Sous (\mathbf{H}_1) , presque sûrement, il existe $C_\Delta > 0$ et $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$\|Y_n - Y\|_{\mathcal{S}_{d-1} \times \Delta} \leq C_\Delta \sqrt{\frac{\log \log n}{n}}.$$

Démonstration. Comme déjà remarqué dans la section 2.4.1 sous (\mathbf{H}_1) on a

$$m\varepsilon \leq \Psi(\varepsilon) \leq M\varepsilon, \quad \varepsilon \geq 0.$$

En posant

$$\varepsilon = \Psi^{-1} \left(\sqrt{\frac{\log \log n}{n}} \right)$$

on obtient par la proposition 2.20 que pour tout $n > 3$

$$\Psi^{-1} \left(\sqrt{\frac{\log \log n}{n}} \right) \leq \frac{1}{m} \sqrt{\frac{\log \log n}{n}}. \quad (2.10)$$

Ensuite grâce à la loi du logarithme itéré, on sait que presque sûrement, il existe $c_\Delta > 0$ et $n(\omega) > 0$ tels que pour tout $n \geq n(\omega)$ on a

$$\|Y_n - Y\|_{S_{d-1} \times \Delta} \leq c_\Delta \Psi^{-1} \left(\sqrt{\frac{\log \log n}{n}} \right)$$

ainsi par l'inégalité 2.10 et en posant $C_\Delta = c_\Delta/m$, on obtient

$$\|Y_n - Y\|_{S_{d-1} \times \Delta} \leq C_\Delta \sqrt{\frac{\log \log n}{n}}.$$

□

2.5.4 Preuve des théorèmes de convergences

Étape préliminaire. Pour $O \in \mathbb{R}^d$, $u \in S_{d-1}$ et $\alpha \in \Delta$, $\gamma > 0$

$$y_\alpha = Y(O, u, \alpha) \quad \text{et} \quad v_\gamma(y_\alpha) = \{y \in \mathbb{R}, |y - y_\alpha| < \gamma\}.$$

On sait que $\lim_{n \rightarrow \infty} \|Y_n(O, u, \alpha) - y_\alpha\|_{S_{d-1} \times \Delta} = 0$ p.s. donc il existe $\gamma_0 > 0$ tel que pour tout $0 < \gamma < \gamma_0$ on a

$$\begin{aligned} Y_n(O, u, \alpha) &= \inf_{y \in v_\gamma(y_\alpha)} \{P_n(H(O, u, y)) \geq \alpha\} \\ &= \inf_{y \in v_\gamma(y_\alpha)} \{P_n(H(O, u, y)) - P(H(O, u, y)) \geq \alpha - P(H(O, u, y))\} \\ &= \inf_{y \in v_\gamma(y_\alpha)} \left\{ P_n(H(O, u, y)) - P(H(O, u, y)) \geq F_{\langle X-O, u \rangle}(y_\alpha) - F_{\langle X-O, u \rangle}(y) \right\} \\ &= \inf_{y \in v_\gamma(y_\alpha)} \left\{ \alpha_n(O, u, y) \geq n^{1/2} \left(F_{\langle X-O, u \rangle}(y_\alpha) - F_{\langle X-O, u \rangle}(y) \right) \right\} \end{aligned} \quad (2.11)$$

avec

$$\alpha_n(O, u, y) = n^{1/2} (P_n(H(O, u, y)) - P(H(O, u, y))).$$

Sous **(H₁)** et **(H₂)**, la fonction $y \mapsto F_{\langle X-O, u \rangle}(y)$ est continue et dérivable sur \mathbb{R} donc par développement de Taylor à l'ordre 1 au voisinage de y_α , on a pour tout $y \in v_\gamma(y_\alpha)$

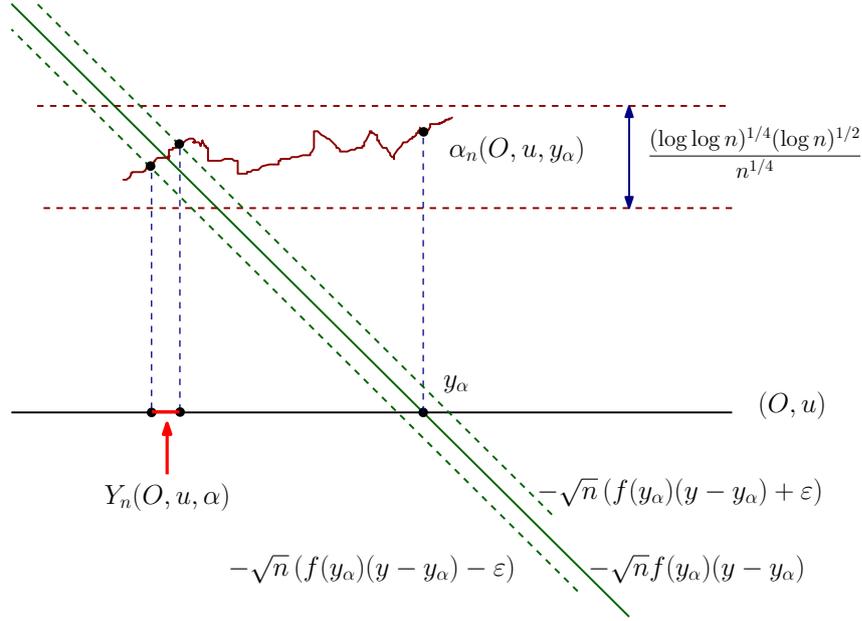
$$F_{\langle X-O, u \rangle}(y_\alpha) - F_{\langle X-O, u \rangle}(y) = f_{\langle X-O, u \rangle}(y_\alpha)(y_\alpha - y) + \varepsilon_\gamma(u, \alpha, y_\alpha - y)$$

avec

$$\lim_{\gamma \rightarrow 0} \sup_{u \in S_{d-1}} \sup_{\alpha \in \Delta} |\varepsilon_\gamma(u, \alpha, y_\alpha - y)| = 0$$

donc

$$Y_n(O, u, \alpha) = \inf_{y \in v_\gamma(y_\alpha)} \left\{ \alpha_n(O, u, y) \geq n^{1/2} \left(f_{\langle X-O, u \rangle}(y_\alpha)(y_\alpha - y) + \varepsilon_\gamma(u, \alpha, y_\alpha - y) \right) \right\}.$$



Preuve du théorème limite central uniforme

Dans cette preuve on commence par montrer dans l'étape I-a sous (\mathbf{H}_1) et (\mathbf{H}_2) l'approximation de Bahadur-Kiefer sans vitesse i.e. $\lim_{n \rightarrow \infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} = 0$ p.s..

Étape I-a : Montrons que sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , on a

$$\lim_{n \rightarrow \infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} = 0 \text{ p.s.}$$

On sait par le lemme 2.61 que, presque sûrement, il existe $C_\Delta > 0$ et $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a pour tout $O \in \mathbb{R}^d$, tout $u \in \mathcal{S}_{d-1}$ et tout $\alpha \in \Delta$, le processus

$$Y_n(O, u, \alpha) \in \left[y_\alpha - C_\Delta \sqrt{\frac{\log \log n}{n}}, y_\alpha + C_\Delta \sqrt{\frac{\log \log n}{n}} \right]$$

on pose

$$\gamma_n = C_\Delta \sqrt{\frac{\log \log n}{n}}.$$

Pour tout $y \in v_{\gamma_n}(y_\alpha)$, on pose

$$z_n(O, u, y, y_\alpha) = \alpha_n(O, u, y_\alpha) - \alpha_n(O, u, y)$$

z_n est un accroissement du processus empirique Λ_n pour des bandes de largeur plus petite γ_n , en effet par la remarque 2.57 on a

$$\alpha_n(O, u, y_\alpha) - \alpha_n(O, u, y) = \Lambda_n(H(O, u, y_\alpha, y))$$

donc par la proposition 2.59,

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right) \quad (2.12)$$

donc

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| = 0 \quad p.s. \quad (2.13)$$

ainsi, pour $n \geq n(\omega)$, on a

$$\begin{aligned} Y_n(O, u, \alpha) &= \inf_{y \in v_{\gamma_n}(y_\alpha)} \left\{ \alpha_n(O, u, y) \geq n^{1/2} \left(f_{\langle X-O, u \rangle}(y_\alpha)(y_\alpha - y) + \varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y) \right) \right\} \\ &= \inf_{y \in v_{\gamma_n}(y_\alpha)} \left\{ y \geq y_\alpha - \frac{\alpha_n(O, u, y)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right\} \\ &= \inf_{y \in v_{\gamma_n}(y_\alpha)} \left\{ y \geq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} - \frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right\} \end{aligned}$$

puisque $|Y_n(O, u, \alpha)| < \infty$, par la dernière égalité, on a d'une part,

$$\begin{aligned} Y_n(O, u, \alpha) &\geq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} \\ &\quad + \inf_{u \in \mathbb{S}^d} \inf_{\alpha \in \Delta} \inf_{y \in v_{\gamma_n}(y_\alpha)} \left(-\frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right) \\ &\geq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} \\ &\quad - \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} \right| - \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \end{aligned}$$

d'autre part, on a

$$\begin{aligned}
& Y_n(O, u, \alpha) \\
& \leq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} \\
& \quad + \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left(-\frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} + \frac{\varepsilon_\gamma(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right) \\
& \leq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} \\
& \quad + \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\langle X-O, u \rangle}(y_\alpha)} \right| + \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right|,
\end{aligned}$$

donc

$$\begin{aligned}
\left| n^{1/2} (Y_n(O, u, \alpha) - y_\alpha) + \frac{\alpha_n(O, u, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| & \leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \quad (2.14) \\
& \quad + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right|.
\end{aligned}$$

Par la condition (\mathbf{H}_1) , on a

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \leq \frac{1}{m} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)|$$

donc, par l'égalité (2.13), on a

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| = 0 \quad p.s.$$

Ensuite, en posant pour tout $0 < \gamma < \gamma_0$

$$\rho(\gamma) = \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_\gamma(y_\alpha)} |\varepsilon_\gamma(u, \alpha, y_\alpha - y)|$$

on obtient que

$$n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \leq \frac{n^{1/2} \rho(\gamma_n)}{m}$$

car $f_{\langle X-O, u \rangle}(y_\alpha) = h(u, \alpha)$ et sous **(H₁)** on a

$$m \leq \inf_{\alpha \in \Delta} \inf_{u \in \mathbb{S}_{d-1}} h_u(\alpha) \leq \sup_{\alpha \in \Delta} \sup_{u \in \mathbb{S}_{d-1}} h_u(\alpha) \leq M.$$

Ensuite, par la condition **(H₂)**, on a

$$\lim_{n \rightarrow \infty} n^{1/2} \rho(\gamma_n) = 0$$

donc

$$\lim_{n \rightarrow \infty} n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| = 0$$

ainsi,

$$\lim_{n \rightarrow \infty} \left| n^{1/2} (Y_n(O, u, \alpha) - y_\alpha) + \frac{\alpha_n(O, u, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| = 0 \quad p.s.$$

or avec les notation précédente, on a

$$\frac{\alpha_n(O, u, y_\alpha)}{f_{\langle X-O, u \rangle}} := \frac{\mathbb{E}_n(u, \alpha)}{h(u, \alpha)}$$

donc

$$\lim_{n \rightarrow \infty} \left\| \sqrt{n} (Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

Étape I-b : Montrons que sous les conditions **(H₁)** et **(H₂)**, on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. (X_n) de loi P et une suite (G_n) de versions de G

$$\lim_{n \rightarrow \infty} \|\mathbb{E}_n - G_n\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

L'ensemble des demi-espaces \mathcal{H} forme une classe de Vapnik–Chervonenkis, donc c'est une classe de Donsker, ainsi on a

$$(\Lambda_n(H))_{H \in \mathcal{H}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} (G(H))_{H \in \mathcal{H}}$$

avec G un pont brownien indexé par \mathcal{H} de covariance donnée par

$$\forall H, H' \in \mathcal{H} \quad \text{cov}(G(H), G(H')) = P(H \cap H') - P(H)P(H').$$

Ensuite, en appliquant le théorème de Berthet et Mason 2006 [BM06a] à Λ_n , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une

suite i.i.d. X_n de loi P et une suite \mathbf{G}_n de versions de \mathbf{G} de sorte que pour tout $\theta > 1$ il existe $K_1 > 0$ tel que

$$\Lambda_n(H(O, u, y)) = \mathbf{G}_n(H(O, u, y)) + \zeta_n(H(O, u, y)) \quad (2.15)$$

avec

$$\lim_{n \rightarrow \infty} \|\zeta_n(H)\|_{\mathcal{H}} = 0 \quad p.s. \quad (2.16)$$

et

$$\mathbb{P} \left(\sup_{u \in \mathcal{S}_{d-1}} \sup_{\alpha \in \Delta} |\zeta_n(u, \alpha)| \geq K_1 \frac{(\log n)^{w_d}}{n^{v_d}} \right) \leq \frac{1}{n^\theta} \quad (2.17)$$

avec la notation $\zeta_n(u, \alpha) = \zeta_n(H(u, \alpha))$ et $v_d = 1/(2 + 10d)$, $w_d = (4 + 10d)/(4 + 20d)$.

Ainsi par les équations (2.15) et (2.16), on a presque sûrement

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathcal{S}_{d-1}} \sup_{\alpha \in \Delta} |\zeta_n(u, \alpha)| = \lim_{n \rightarrow \infty} \sup_{u \in \mathcal{S}_{d-1}} \sup_{\alpha \in \Delta} |\Lambda_n(H(u, \alpha)) - \mathbf{G}_n(H(u, \alpha))| = 0$$

donc avec la notation $\mathbf{G}_n(u, \alpha) = \mathbf{G}_n(H(u, \alpha))$ on a

$$\lim_{n \rightarrow \infty} \|\mathbb{E}_n - \mathbf{G}_n\|_{\mathcal{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

Étape I-c : Montrons que sous les conditions **(H₁)** et **(H₂)**, on a

$$\lim_{n \rightarrow \infty} \left\| \mathbb{D}_n + \frac{\mathbf{G}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

Sous **(H₁)** et **(H₂)**, par l'étape I-a. on a

$$\lim_{n \rightarrow \infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\mathbb{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

et par l'étape I-b. on a

$$\lim_{n \rightarrow \infty} \|\mathbb{E}_n - \mathbf{G}_n\|_{\mathcal{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

donc par l'inégalité triangulaire et la condition **(H₁)**, on a

$$\begin{aligned} \left\| \mathbb{D}_n + \frac{\mathbf{G}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} &\leq \left\| \mathbb{D}_n + \frac{\mathbb{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} + \left\| \frac{\mathbb{E}_n}{h} - \frac{\mathbf{G}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} \\ &\leq \left\| \mathbb{D}_n + \frac{\mathbb{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} + \frac{\|\mathbb{E}_n - \mathbf{G}_n\|_{\mathcal{S}_{d-1} \times \Delta}}{m} \end{aligned}$$

alors sous (\mathbf{H}_1) et (\mathbf{H}_2) , on a

$$\lim_{n \rightarrow \infty} \left\| \mathbb{D}_n + \frac{\mathbf{G}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

Conclusion on a montré que

$$\lim_{n \rightarrow \infty} \left\| \mathbb{D}_n + \frac{\mathbf{G}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

ce qui implique que

$$\lim_{n \rightarrow \infty} d_{PL}(\sqrt{n}(Y_n - Y), -\frac{\mathbf{G}_n}{h}) = \lim_{n \rightarrow \infty} d_{PL}(\sqrt{n}(Y_n - Y), -\frac{\mathbf{G}}{h}) = \lim_{n \rightarrow \infty} d_{PL}(\sqrt{n}(Y_n - Y), \frac{\mathbf{G}}{h}) = 0$$

où d_{PL} est la distance de Prokhorov-Levy. Donc on a

$$\mathbb{D}_n = \sqrt{n}(Y_n - Y) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \tilde{\mathbf{G}}$$

sur l'ensemble des fonctions bornées sur $\mathbb{S}_{d-1} \times \Delta$ muni de la norme uniforme, avec la notation $\tilde{\mathbf{G}} = \frac{\mathbf{G}}{h}$, on a

$$\text{cov}(\tilde{\mathbf{G}}(u, \alpha), \tilde{\mathbf{G}}(u', \alpha')) = \frac{P(H(u, \alpha) \cap H(u', \alpha')) - \alpha \alpha'}{h(u, \alpha)h(u', \alpha')}.$$

Preuve de l'approximation du type Bahadur-Kiefer

On veut montrer que sous les conditions (\mathbf{H}_1) , (\mathbf{H}_2) et $[(\mathbf{H}_3)$ ou $(\mathbf{H}'_3)]$, il existe $n(\omega)$ tel que pour tout $n \geq n(\omega)$

$$\left\| \mathbb{D}_n + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right).$$

La classe des demi-espaces \mathcal{H} est une classe de Vapnik-Červonenkis, d'indice $d + 1$, donc par la loi du logarithme itéré (Alexander 1984 [Ale84]) on a

$$\limsup_{n \rightarrow \infty} \frac{\|\Lambda_n\|_{\mathcal{H}}}{\sqrt{2 \log \log n}} \leq 1/2 \quad p.s.$$

donc avec probabilité 1, il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a pour tout $u \in \mathbb{S}_{d-1}$ le processus

$$\alpha_n(O, u, y) := \Lambda_n(H(O, u, y)) \in \left[-\sqrt{3 \log \log n}, \sqrt{3 \log \log n} \right].$$

Pour tout $n \geq n(\omega)$, on rappelle l'inégalité (2.14)

$$\begin{aligned} \left| n^{1/2} (Y_n(O, u, \alpha) - y_\alpha) + \frac{\alpha_n(O, u, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| &\leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \\ &\quad + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \end{aligned}$$

ainsi que l'égalité (2.12)

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right)$$

donc il existe une constante $C' > 0$, telle que tout $n \geq n(\omega)$

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| \leq C' \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right)$$

alors pour $n \geq n(\omega)$ par **(H₁)** on a

$$\begin{aligned} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| &\leq \frac{1}{m} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| \\ &\leq \frac{C'}{m} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right). \quad (2.18) \end{aligned}$$

Ensuite, en posant pour tout $0 < \gamma < \gamma_0$

$$\rho(\gamma) = \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_\gamma(y_\alpha)} |\varepsilon_\gamma(u, \alpha, y_\alpha - y)|.$$

Sous la condition **(H₃)**, on sait qu'il existe $r > 1/2$ et $C^* > 0$ tel que pour tout $\gamma > 0$ on a

$$\rho(\gamma) \leq C^* \gamma^{1+r}$$

donc pour $n \geq n(\omega)$

$$\begin{aligned}
n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| &\leq \frac{n^{1/2} \rho(\gamma_n)}{m} \\
&\leq \frac{C^* C_\Delta}{m} n^{1/2} \left(\sqrt{\frac{\log \log n}{n}} \right)^{1+r} \\
&\leq \frac{C'_\Delta}{m} \left(\frac{(\log \log n)^{\frac{1+r}{2}}}{n^{\frac{r}{2}}} \right) \quad (2.19)
\end{aligned}$$

où $C'_\Delta = C^* C_\Delta$.

En rappelant les notations suivantes

$$\mathbb{D}_n(u, \alpha) = \sqrt{n}(Y_n(O, u, \alpha) - Y(O, u, \alpha))$$

et

$$\mathbb{E}_n(u, \alpha) = \alpha_n(O, u, y_\alpha)$$

ainsi, on peut écrire

$$\left| n^{1/2}(Y_n(O, u, \alpha) - y_\alpha) + \frac{\alpha_n(O, u, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| := \left| \mathbb{D}_n(u, \alpha) + \frac{\mathbb{E}_n(u, \alpha)}{h(u, \alpha)} \right|$$

on a par les inégalités (2.18) et (2.19), pour tout $n \geq n(\omega)$

$$\begin{aligned}
\left| \mathbb{D}_n(u, \alpha) + \frac{\mathbb{E}_n(u, \alpha)}{h(u, \alpha)} \right| &\leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \\
&\quad + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \\
&\leq \frac{C'_\Delta}{m} \left(\frac{(\log \log n)^{\frac{1+r}{2}}}{n^{\frac{r}{2}}} \right) + \frac{C'}{m} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right) \\
&\leq C''_\Delta \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}. \quad (2.20)
\end{aligned}$$

Ainsi, sous les conditions (\mathbf{H}_1) , (\mathbf{H}_2) et (\mathbf{H}_3) , il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$

$$\left\| \mathbb{D}_n + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right).$$

Sous la condition (\mathbf{H}_3) , on sait que

$$\lim_{\gamma \rightarrow 0} \frac{\rho(\gamma)}{\gamma^{3/2} \sqrt{\log(1/\gamma)}} = 0.$$

Donc pour $n \geq n(\omega)$,

$$\begin{aligned} \left| \mathbb{D}_n(u, \alpha) + \frac{\mathbb{E}_n(u, \alpha)}{h(u, \alpha)} \right| &\leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \\ &\quad + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(u, \alpha, y_\alpha - y)}{f_{\langle X-O, u \rangle}(y_\alpha)} \right| \\ &\leq \frac{n^{1/2} \rho(\gamma_n)}{m} + \frac{C'}{m} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right) \\ &\leq \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4} m} \left(\frac{\rho(\gamma_n) (\log \log n)^{1/2}}{(\gamma_n)^{3/2} (\log n)^{1/2}} + C' \right) \\ &\leq \frac{t_n}{m} \left(\frac{\rho(\gamma_n) (\log \log(1/\gamma_n))^{1/2}}{\gamma_n^{3/2} (\log(1/\gamma_n))^{1/2}} \left(\frac{\log(1/\gamma_n) (\log \log n)}{\log \log(1/\gamma_n) (\log n)} \right)^{1/2} + C' \right) \end{aligned}$$

avec la notation $\log_p = \log \log \cdots \log$, la composée p fois de \log , on a

$$\begin{aligned} \frac{\log(1/\gamma_n) (\log \log n)}{\log \log(1/\gamma_n) (\log n)} &= \frac{\frac{1}{2} (\log n + \log_3 n) \log_2 n}{\left(\log\left(\frac{1}{2}\right) + \log(\log n - \log_3 n) + \log_3 n \right) \log n} \\ &= \frac{\frac{1}{2} \log_2 n}{\log\left(\frac{1}{2}\right) + \log(\log n - \log_3 n) + \log_3 n} \\ &\quad + \frac{\frac{1}{2} \log_3 n \log_2 n}{\left(\log\left(\frac{1}{2}\right) + \log(\log n - \log_3 n) + \log_3 n \right) \log n} \end{aligned}$$

puisque

$$\lim_{n \rightarrow \infty} \frac{\frac{1}{2} \log_2 n}{\log\left(\frac{1}{2}\right) + \log(\log n - \log_3 n) + \log_3 n} = \frac{1}{2}$$

et

$$\lim_{n \rightarrow \infty} \frac{\frac{1}{2} \log_3 n \log_2 n}{\left(\log\left(\frac{1}{2}\right) + \log(\log n - \log_3 n) + \log_3 n \right) \log n} = 0$$

donc

$$\lim_{n \rightarrow \infty} \left(\frac{\log(1/\gamma_n) (\log \log n)}{\log \log(1/\gamma_n) (\log n)} \right)^{1/2} = \frac{1}{2}$$

par **(H₃)**, on a

$$\lim_{n \rightarrow \infty} \frac{\rho(\gamma_n)(\log \log(1/\gamma_n))^{1/2}}{\gamma_n^{3/2}(\log(1/\gamma_n))^{1/2}} = 0$$

donc

$$\lim_{n \rightarrow \infty} \frac{\rho(\gamma_n)(\log \log(1/\gamma_n))^{1/2}}{\gamma_n^{3/2}(\log(1/\gamma_n))^{1/2}} \left(\frac{\log(1/\gamma_n)(\log \log n)}{\log \log(1/\gamma_n)(\log n)} \right)^{1/2} = 0$$

alors, il existe $n_1(\omega) > 0$ tel que pour tout $n > n_1(\omega)$ on a

$$\frac{\rho(\gamma_n)(\log \log n)^{1/2}}{(\gamma_n)^{3/2}(\log n)^{1/2}} \leq C'.$$

Ainsi, sous les conditions **(H₁)**, **(H₂)** et **(H₃)**, il existe $n^*(\omega) = \max(n(\omega), n_1(\omega)) > 0$ tel que pour tout $n \geq n^*(\omega)$

$$\left\| \mathbb{D}_n + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log \log n)^{1/4}(\log n)^{1/2}}{n^{1/4}} \right).$$

Remarque 2.62. Puisque sous **(H₁)** et **(H₂)** on a

$$\lim_{n \rightarrow \infty} n^{1/2} \rho(\gamma_n) = 0$$

alors sans les conditions **(H₃)** et **(H'₃)**, on peut seulement affirmer qu'il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$

$$\left\| \mathbb{D}_n + \frac{\mathbb{E}_n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\left(n^{1/2} \rho(\gamma_n) \right) \vee \frac{(\log \log n)^{1/4}(\log n)^{1/2}}{n^{1/4}} \right).$$

Preuve du principe d'invariance fort

Dans la preuve du théorème limite central uniforme, on a montré que sous **(H₁)** et **(H₂)**, on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. (X_n) de loi P et une suite (G_n) de versions de G tel que pour tout $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$, on a

$$\mathbb{D}_n(u, \alpha) = -\frac{\mathbb{G}_n(u, \alpha)}{h(u, \alpha)} + \mathbb{Z}_n(u, \alpha)$$

où $\mathbb{Z}_n(u, \alpha) = \mathbb{D}_n(u, \alpha) + \frac{\mathbb{G}_n(u, \alpha)}{h(u, \alpha)}$ avec

$$\lim_{n \rightarrow \infty} \|\mathbb{Z}_n\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \text{ p.s.}$$

On veut montrer que sous la condition (\mathbf{H}_3) , il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a

$$\|\mathbf{Z}_n\|_{\mathcal{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_d}}{n^{v_d}} \right)$$

avec $v_d = 1/(2 + 10d)$, $w_d = (4 + 10d)/(4 + 20d)$.

On sait que

$$\begin{aligned} \left\| \mathbf{D}_n + \frac{\mathbf{G}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} &\leq \left\| \mathbf{D}_n + \frac{\mathbf{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} + \left\| \frac{\mathbf{E}_n}{h} - \frac{\mathbf{G}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} \\ &\leq \left\| \mathbf{D}_n + \frac{\mathbf{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} + \frac{1}{m} \|\mathbf{E}_n - \mathbf{G}_n\|_{\mathcal{S}_{d-1} \times \Delta}. \end{aligned}$$

Sous la condition (\mathbf{H}_3) , par l'inégalité (2.20), il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a

$$\left\| \mathbf{D}_n + \frac{\mathbf{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} \leq C''_{\Delta} \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}$$

ensuite par (2.17) et Borel-Cantelli, on a

$$\|\mathbf{E}_n - \mathbf{G}_n\|_{\mathcal{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_d}}{n^{v_d}} \right)$$

donc il existe $K_1 > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$\|\mathbf{E}_n - \mathbf{G}_n\|_{\mathcal{S}_{d-1} \times \Delta} \leq K_1 \left(\frac{(\log n)^{w_d}}{n^{v_d}} \right)$$

$$\begin{aligned} \left\| \mathbf{D}_n + \frac{\mathbf{G}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} &\leq \left\| \mathbf{D}_n + \frac{\mathbf{E}_n}{h} \right\|_{\mathcal{S}_{d-1} \times \Delta} + \frac{1}{m} \|\mathbf{E}_n - \mathbf{G}_n\|_{\mathcal{S}_{d-1} \times \Delta} \\ &\leq C''_{\Delta} \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} + \frac{K_1 (\log n)^{w_d}}{m n^{v_d}} \\ &\leq \frac{D_0 (\log n)^{w_d}}{n^{v_d}} \end{aligned}$$

Conclusion Sous **(H₁)** et **(H₂)** on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. (X_n) de loi P et une suite (G_n) de versions de G de sorte que

$$Y_n(O, u, \alpha) = Y(O, u, \alpha) + \frac{G_n(u, \alpha)}{h(u, \alpha)\sqrt{n}} + \frac{Z_n(u, \alpha)}{\sqrt{n}}, \quad O \in \mathbb{R}^d, \quad \alpha \in \Delta, \quad u \in \mathbb{S}_{d-1},$$

où $Z_n = D_n - \frac{G_n}{h}$ vérifie

$$\lim_{n \rightarrow \infty} \|Z_n\|_{\mathbb{S}_{d-1} \times \Delta} = \lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |Z_n(u, \alpha)| = 0 \quad p.s.$$

si de plus P satisfait **(H₃)** alors on peut construire G_n tel que pour $v_d = 1/(2 + 10d)$, $w_d = (4 + 10d)/(4 + 20d)$, il existe $n_\omega > 0$ tel que tout pour $n > n_\omega$,

$$\|Z_n\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_d}}{n^{v_d}} \right).$$

2.6 Simulations et interprétations

Pour simuler les surfaces quantile, on se place dans \mathbb{R}^2 , à fin de faciliter les interprétations et de permettre des représentations graphiques. On commence par donner les simulations pour des lois gaussiennes ou des mélanges de lois gaussiennes. Cela nous sert en particulier à exhiber les formes génériques des surfaces quantile. Dans différents modèles, on fait dans un premier temps les simulations en fixant le niveau de quantile α et en choisissant différents observateurs ou points repères O , puis en faisant varier le niveau de quantile α à observateur fixé .

2.6.1 La loi gaussienne

2.6.1.a Gaussienne $N(0_2, \Sigma)$ avec $\Sigma = \lambda I_2$

Pour commencer on s'intéresse à la loi gaussienne $N(0_2, I_2)$. Dans un premier temps on fixe α et on s'intéresse à $Q(O, \alpha)$ lorsque O varie. Dans un deuxième temps on s'intéresse à $Q(O, \alpha)$ avec O fixé.

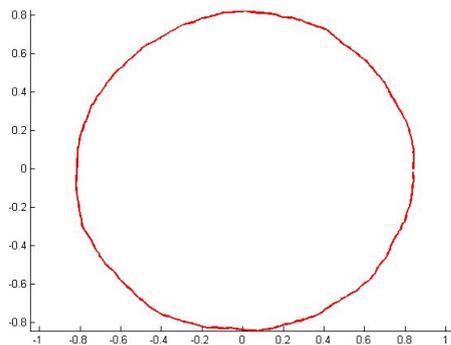


FIGURE 2.11 – $Q(O, \alpha)$ pour $O = \mathbb{E}(X) = (0, 0)$ et $\alpha = 0.8$

On remarque que la surface quantile est un cercle centré en O de rayon ≈ 0.845 , car par exemple dans la direction $u = (1, 0)$ le demi-plan se tenant à distance $y = 0.845$ de O est de masse $\alpha = 0.8$ puisque

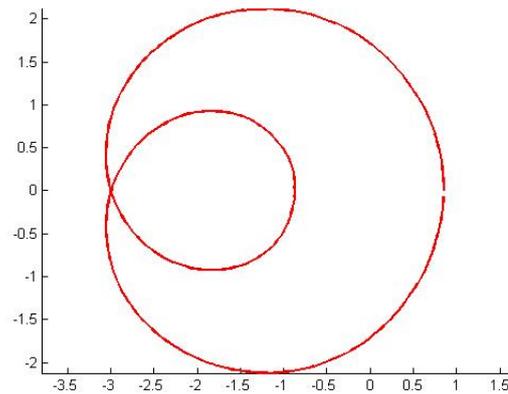
$$\frac{1}{2\pi} \int_{-\infty}^{0.845} \int_{\mathbb{R}} \exp(-(x^2 + y^2)/2) dx dy \approx 0.8.$$

De même le demi-plan de direction $u = (-1, 0)$ se tenant à distance $y = 0.845$ de O est de masse $\alpha = 0.8$ puisque

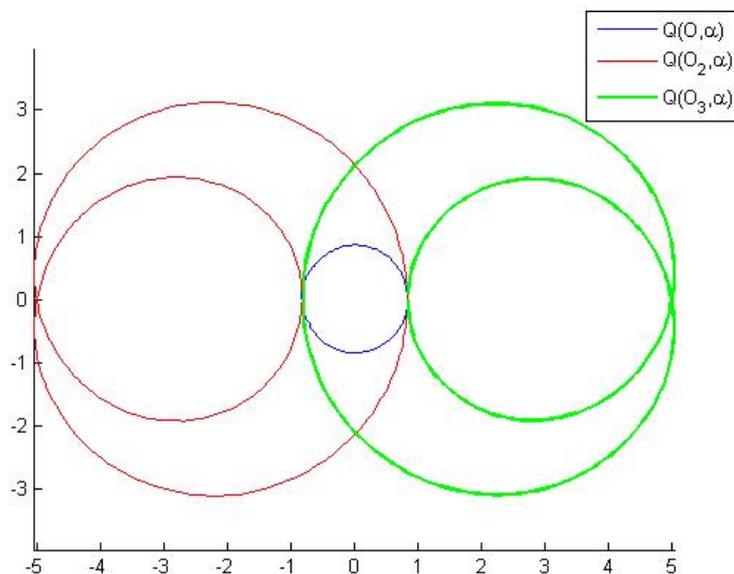
$$\frac{1}{2\pi} \int_{-0.845}^{+\infty} \int_{\mathbb{R}} \exp(-(x^2 + y^2)/2) dx dy \approx 0.8.$$

Dans ce cas particulier, de même que pour toute loi symétrique, les surfaces quantiles $Q(O, \alpha)$ pour $O = \mathbb{E}(X)$ et l'ensemble de Tukey-Nolan \mathcal{C}_α sont confondus. Ceci est dû au fait que dans toutes les directions $u \in \mathbb{S}_{d-1}$, les demi-plan de masse α $H(u, \alpha)$ sont tangents à $Q(O, \alpha)$ ainsi $\mathcal{C}_\alpha = \bigcap_{H \in \mathcal{H}_\alpha} H = Q(O, \alpha)$.

Cas I : α fixe, O mobile. En bougeant l'observateur O , la surface quantile $Q(O, \alpha)$ qui on rappelle, est dépendante du choix de l'observateur, se transforme en une surface d'une certaine forme qu'on appelle « Cardioïde ».



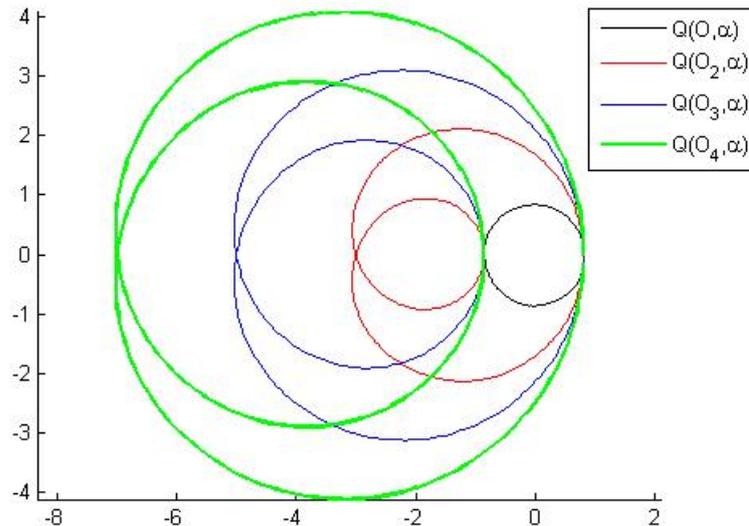
Une première remarque est que la nouvelle surface $Q(O', \alpha)$ pour $O' = (-3, 0)$ est plus « grande » que la surface $Q(O, \alpha)$ pour $O = \mathbb{E}(X) = 0_2$.



Dans le graphique précédent

- En bleu $Q(O, \alpha)$ pour $O = \mathbb{E}(X) = 0_2$.
- En rouge $Q(O_2, \alpha)$ pour $O_2 = (-5, 0)$.
- En vert $Q(O_2, \alpha)$ pour $O_2 = (5, 0)$.

De la même manière on peut tracer les surfaces quantiles de niveau α pour $O \in \{(0, 0), (-3, 0), (-5, 0), (-7, 0)\}$



Dans le graphique précédent

- En noir $Q(O, \alpha)$ pour $O = \mathbb{E}(X) = 0_2$.
- En rouge $Q(O_2, \alpha)$ pour $O_2 = (-3, 0)$.
- En bleu $Q(O_3, \alpha)$ pour $O_3 = (-5, 0)$.
- En vert $Q(O_4, \alpha)$ pour $O_4 = (-7, 0)$.

On remarque principalement deux choses :

- D'un côté plus on s'éloigne de la « médiane » qui correspond dans ce cas spécifique à la moyenne $\mathbb{E}(X) = 0_2$, plus la forme augmente en volume, ce qui en terme d'applications statistiques, peut être un bon indicateur de « **centralité** » de l'observateur.
- D'un autre côté on remarque qu'en déplaçant l'observateur suivant l'axe $(0_2, u_0)$ avec $u_0 = (1, 0)$, toutes les surfaces sont intersectées en deux points M_1, M_2 de cette axe, ici $M_1 \approx (-0.845, 0)$ et $M_2 \approx (0.845, 0)$.

Afin de bien comprendre ces phénomènes, on s'intéresse tout d'abord à la transformation des points de la surface $Q(O, \alpha)$ vers la surface $Q(O', \alpha)$.

En Fixant α et en choisissant une direction quelconque u dans S_{d-1} , on rappelle que par la proposition 2.3, le demi-plan de normale u et de masse α est indépendant du choix de O , de plus, ce demi-plan est unique grâce à l'hypothèse (\mathbf{H}_0^+) . Néanmoins si $\frac{O'-O}{\|O'-O\|} \neq \pm u$ alors $Q(O, u, \alpha) \neq Q(O', u, \alpha)$ car $Q(O, u, \alpha) = p_{H(u, \alpha)}(O)$ et $Q(O', u, \alpha) = p_{H(u, \alpha)}(O')$ (voir graphique ci-dessous), par contre, si $O' = O + \lambda u$ alors $Q(O, u, \alpha) = Q(O', u, \alpha)$.

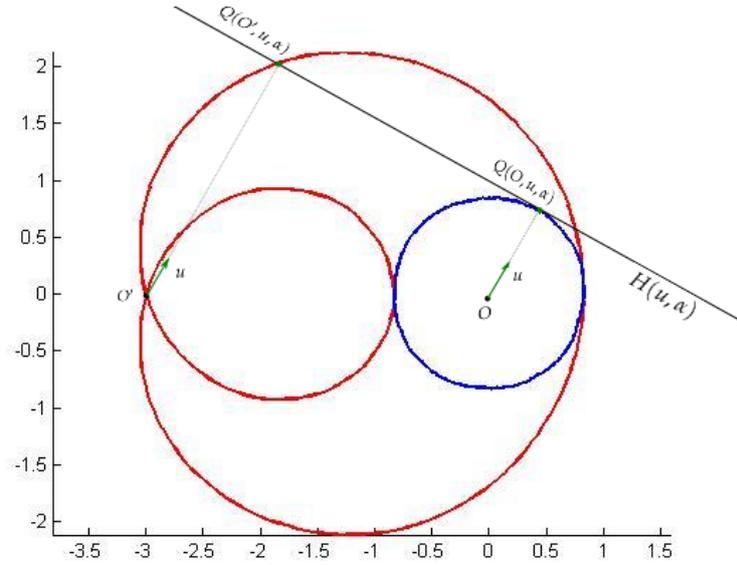


FIGURE 2.12 – En rouge : $Q(O, \alpha)$ pour $O = (-3, 0)$ et $\alpha = 0.8$
En bleu : $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.8$

Ainsi il est aisé de voir sur le graphique que pour tout $O' \in \mathbb{R}^d$ si $u = \frac{O'-O}{\|O'-O\|}$ ou $u = -\frac{O'-O}{\|O'-O\|}$ alors on a $Q(O, u, \alpha) = Q(O', u, \alpha)$, sinon elles sont distinctes partout ailleurs. Ce qui explique que dans le graphique précédent, toutes les surfaces sont « intersectées » en deux points M_1 et M_2 sur l'axe du déplacement $(O, O' - O)$.

Ensuite, pour comprendre la boucle intérieure, on se place dans le cas où par exemple $O' - O = \lambda(Q(O, u_0, \alpha) - O)$ avec $\lambda > 1$, $u_0 = \frac{O'-O}{\|O'-O\|}$ et $O = \mathbb{E}(X)$, comme dans le graphique précédent.

Parmi les demi-plans dont la frontière passe par O' , il existe ici deux demi-plans de masse α , on note u^* et u^{**} les normales respectives à ces demi-plans. Puisque la distance quantile vue de O' dans la direction u^* de masse α est nulle, i.e. $Y(O', u^*, \alpha) = 0$ alors $Q(O', u^*, \alpha) = O'$. Étant donné que $H(u^*, \alpha)$ ne dépend pas du choix de l'observateur, ce demi-plan intersecte

$Q(O, \alpha)$ en $Q(O, u^*, \alpha)$.

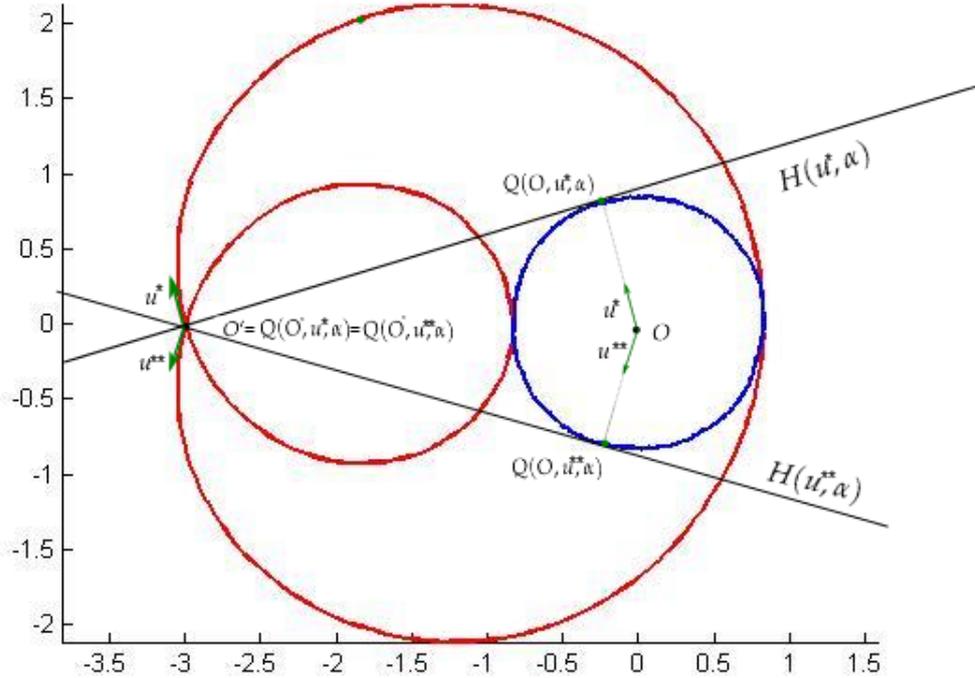


FIGURE 2.13 – En rouge : $Q(O, \alpha)$ pour $O = (-3, 0)$ et $\alpha = 0.8$
 En bleu : $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.8$
 En noir : $\partial H(u^*, \alpha)$ et $\partial H(u^{**}, \alpha)$

En continuant de bouger la direction u , les demi-plans de masse α se retrouvent « derrière » O' pour une partie $S^- \subset \mathbb{S}_{d-1}$, i.e. pour tout $u' \in S^-$ on a $Q(O', u, \alpha) = O' + \lambda' u'$ avec $\lambda' < 0$. Ceci s'explique par le fait qu'il n'y a plus assez de masse à exclure « devant » O' et qu'il faille reculer pour récupérer le bon demi-plan de masse α . Ce phénomène se poursuit en passant par la direction u_0 où $Q(O, -u_0, \alpha) = Q(O', -u_0, \alpha)$ jusqu'au moment où la frontière d'un demi-plan repasse une nouvelle fois par O' avec $\partial H(u^{**}, \alpha)$. Dans cet exemple, la boucle extérieure est indexée par une partie ouverte et connexe de la sphère $S^+ \subset \mathbb{S}_{d-1}$, la boucle intérieure est elle aussi indexée par une partie ouverte et connexe de la sphère $S^- \subset \mathbb{S}_{d-1}$ enfin le point $O' = Q(O', u^*, \alpha) = Q(O', u^{**}, \alpha)$ est indexé par deux vecteurs u^* et u^{**} et $\mathbb{S}_{d-1} = S^+ \cup \{u^*\} \cup S^- \cup \{u^{**}\}$.

Remarque 2.63. *Ce découpage simple n'est en aucun cas une généralité, comme nous le verrons dans un exemple ultérieur, on peut construire des lois pour lesquelles par exemple le point O' est indexé par un continuum de vecteurs de S_{d-1} .*

Remarquons que les surfaces quantiles $Q(O, \alpha)$ pour $O = 0_2$ ont la même forme que les ensembles de niveaux de la loi $N(0_2, I_2)$, néanmoins elles ne suivent pas la même indexation. Par exemple pour $\alpha = 0.8$ la surface quantile $Q(O, \alpha)$ est $C(O, r)$ le cercle centré en O et de rayon $r \approx 0.845$, or dans ce cas on a

$$\frac{1}{2\pi} \int \int_{C(O,r)} \exp(-(x^2 + y^2)/2) dx dy \approx 0.3.$$

Ainsi la surface quantile $Q(O, 0.8)$ correspond à l'ensemble de niveau $L_{0.3}$ de masse 0.3.

Cas II : O fixe, α mobile. On commence par donner les simulations des $Q(O, \alpha)$ avec $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$, pour $O = (0, 0)$ dans un premier temps, puis $O = (-5, 0)$.

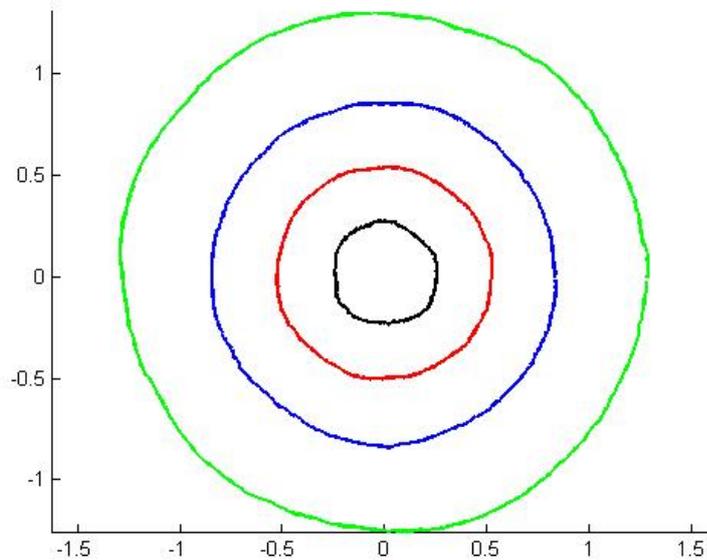


FIGURE 2.14 – En noir : $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.6$
 En rouge : $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.7$
 En bleu : $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.8$
 En vert : $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.9$

Notre première remarque concerne l'emboîtement des $Q(O, \alpha)$, comme déjà énoncé dans la proposition 2.9, ensuite les $Q(O, \alpha)$ pour $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$ sont tous des cercles concentriques en O . Voici maintenant une simulation de $Q(O, \alpha)$ pour $O = (-5, 0)$ et $\alpha \in \{0.6, 0.9\}$.

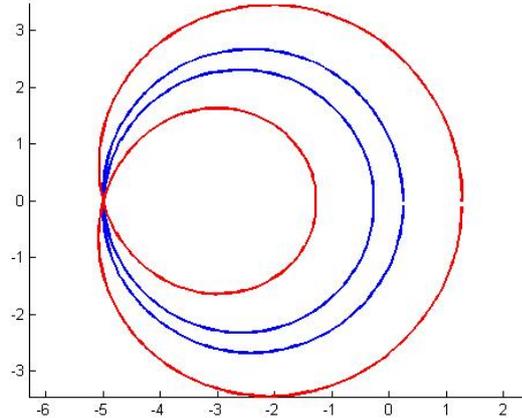
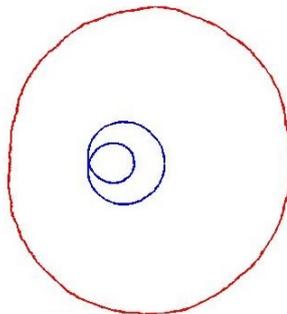


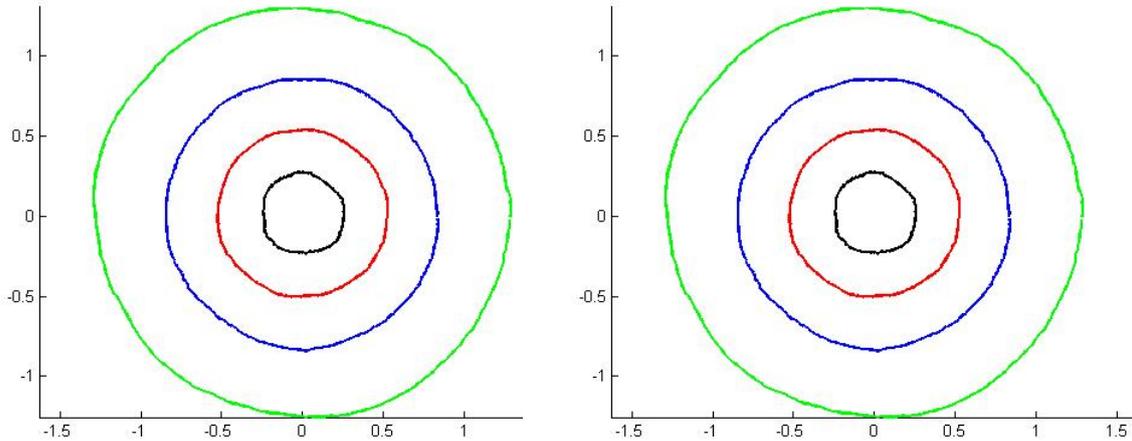
FIGURE 2.15 – En bleu : $Q(O, \alpha)$ pour $O = (-5, 0)$ et $\alpha = 0.6$
En rouge : $Q(O, \alpha)$ pour $O = (-5, 0)$ et $\alpha = 0.9$

Lorsque α croît, on remarque que la boucle extérieure augmente en volume, alors que le volume de la boucle intérieure décroît, cette décroissance de volume est due au fait que lorsque α augmente, $1 - \alpha$ décroît, donc il y a moins de masse à exclure, ce qui implique qu'il ne faille de moins en moins « reculer » pour exclure de la masse. C'est le phénomène inverse qui dirige la boucle extérieure.

Remarque 2.64. *En augmentant α la boucle intérieure peut aussi complètement disparaître, du fait qu'il faille plus du tout aller dans le « dos » de l'observateur O pour exclure de la masse. Voici un exemple, pour $O = (-1, 0)$ et $\alpha = 0.6$ en bleu et $\alpha = 0.95$ en rouge.*



Afin de comprendre l'importance de l'indexation des surfaces par α on commence par donner les graphiques suivants :



(a) Quatre surfaces $Q(O, \alpha)$ de $N(O, I_2)$ pour $O = (0, 0)$ (b) Quatre surfaces $Q(O, \alpha)$ de $N(O, 5I_2)$ pour $O = (0, 0)$

Les deux graphiques précédents sont identiques et représentent quatre cercles concentriques : $C(0, 0.254)$, $C(0, 0.527)$, $C(0, 0.837)$ et $C(0, 0.1.288)$. Pour la loi $N(O, I_2)$ ils représentent les surfaces $Q(O, \alpha)$ pour les niveaux respectifs $\alpha = 0.6$, $\alpha = 0.7$, $\alpha = 0.8$ et $\alpha = 0.9$ alors que pour la loi $N(O, 5I_2)$ ils représentent les surfaces $Q(O, \alpha)$ pour les niveaux respectifs $\alpha = 0.57$, $\alpha = 0.64$, $\alpha = 0.72$ et $\alpha = 0.8$. Ceci montre un fait évident : c'est la connaissance des surfaces quantiles et de leurs « label » correspondants qui permet de décrire la loi.

Avant de s'intéresser à d'autres types de lois, voici une idée sur laquelle nous reviendrons dans un chapitre ultérieur et qui montre une des utilisations possibles de nos quantiles spatiaux en statistiques descriptives. Il s'agit du problème de la localisation de masse. Le principe consiste à bouger à la fois O et α à fin de détecter les zones de forte probabilité.

Dans la simulation suivante, on a pris trois différents observateurs placés à $O_1 = (2 \cos(\pi/6), 2 \sin(\pi/6))$, $O_2 = (-2, 0)$ $O_3 = (2 \cos(-\pi/6), 2 \sin(-\pi/6))$ avec $\alpha = 0.8$.

Nous remarquons que les boucles intérieures « pointent » vers la zone de concentration de la masse, une propriété qui pourrait être mise à profit dans un algorithme de détection de masse automatisé.

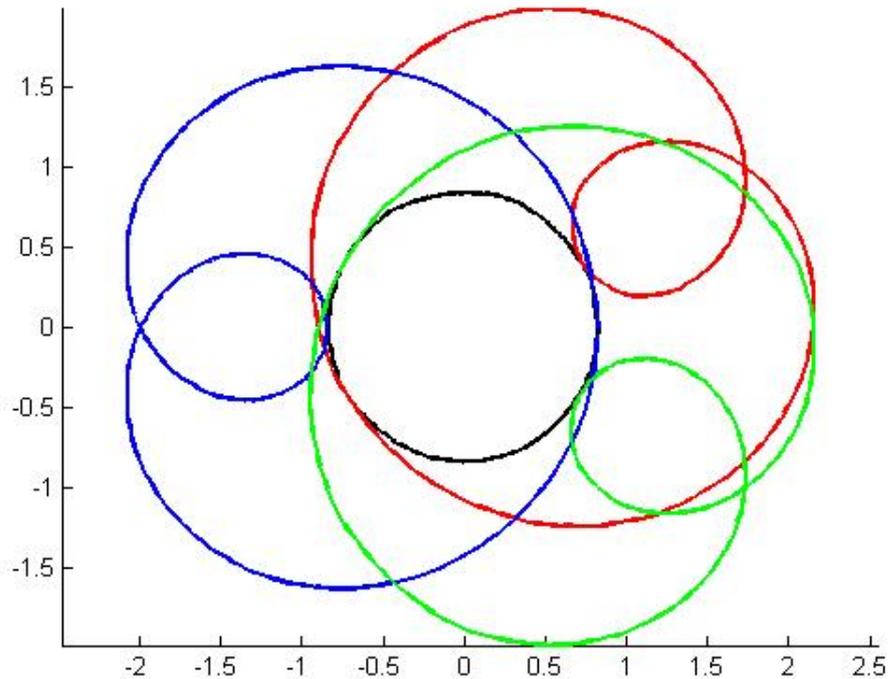


FIGURE 2.17 – En noir : $Q(O, \alpha)$ pour $0 = (0, 0)$ et $\alpha = 0.8$.
 En rouge : $Q(O_1, \alpha)$ pour $O_1 = (2 \cos(\frac{\pi}{6}), 2 \sin(\frac{\pi}{6}))$ et $\alpha = 0.8$.
 En bleu : $Q(O_2, \alpha)$ pour $O_2 = (-2, 0)$ et $\alpha = 0.8$.
 En vert : $Q(O_3, \alpha)$ pour $O_3 = (2 \cos(\frac{-\pi}{6}), 2 \sin(\frac{-\pi}{6}))$ et $\alpha = 0.8$.

2.6.1.b La loi gaussienne $N(0_2, \Sigma)$ avec $\Sigma \neq \lambda I_2$

On pose

$$P_1 = N(0_2, \Sigma) \text{ avec } \Sigma = \begin{pmatrix} 4 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

Cette fois les ensembles de niveaux de la loi ne sont plus des cercles, mais des ellipses.

Voici une première simulation qui rompt avec la forme des surfaces quantiles des lois $N(0, \lambda I_2)$ i.e. avec des ensembles de niveaux circulaires. Nous affichons aussi un échantillon de 10^4 points afin de rendre compte de la répartition des points.

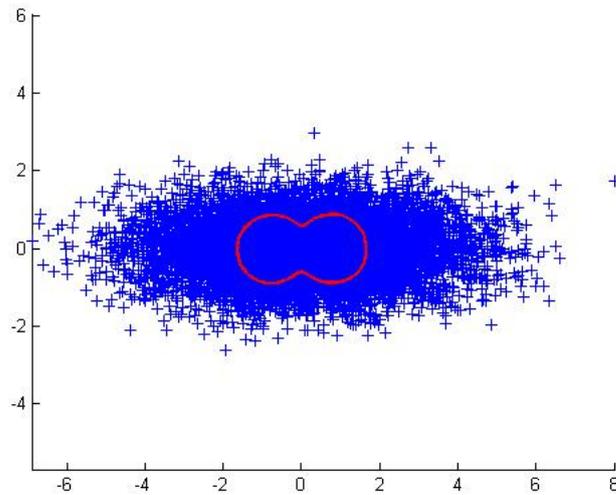


FIGURE 2.18 – $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.8$

On remarque que pour $O = \mathbb{E}X = 0_d$ la surface quantile $Q(O, \alpha)$ n'est plus une surface convexe, donc à la différence du cas précédent, $Q(O, \alpha) \neq \mathcal{C}_\alpha$ car $\mathcal{C}_\alpha = \bigcap_{H \in \mathcal{H}_\alpha} H$ est une intersection de convexe.

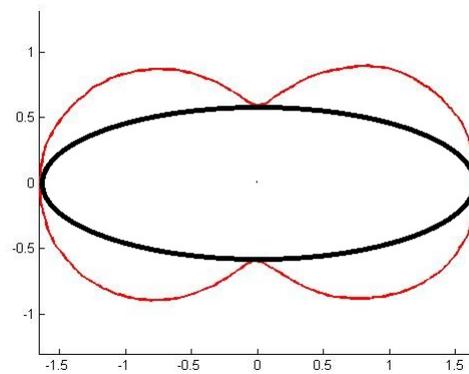


FIGURE 2.19 – En rouge : $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.8$
En noir : \mathcal{C}_α pour $\alpha = 0.8$.

Le « pincement » vertical dans la surface $Q(O, \alpha)$ est dû à la concentration de la masse dans l'axe des ordonnées et son étalement sur l'axe des abscisses. Effectivement avec une simple modification de la matrice de variance-

covariance, par exemple pour $P_2 = N(0_2, \Sigma)$ avec $\Sigma = \begin{pmatrix} 1/2 & 0 \\ 0 & 4 \end{pmatrix}$, on peut observer que les pincements passent au long de l'axe des abscisses.

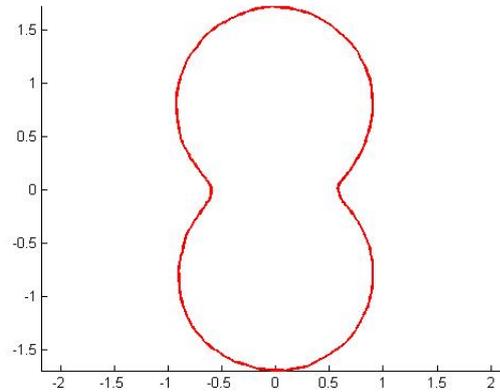


FIGURE 2.20 – $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.8$

Ainsi pour la loi P_1 , si on regarde par exemple dans la direction $u = (0, 1)$ ou $u = (0, -1)$, du fait de la concentration de la masse dans cette direction, il faut très peu s'éloigner de O pour placer un demi-plan de masse α , par contre plus on tourne vers $u = (1, 0)$ ou $u = (-1, 0)$ plus il faut s'éloigner de O pour obtenir une masse égale à α .

Enfin, voici la surface quantile pour $O = (-3, 0)$, on retrouve comme dans la section précédente, une forme en « cardioïde », avec une boucle intérieure pointant dans direction de la masse.

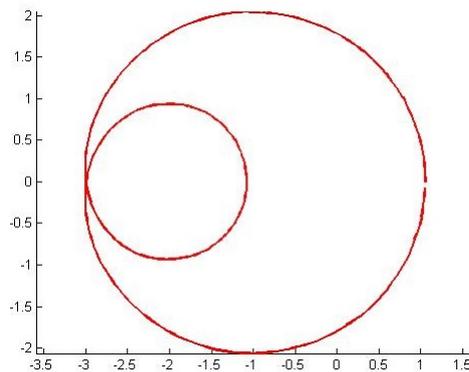


FIGURE 2.21 – $Q(O, \alpha)$ pour $O = (-3, 0)$ et $\alpha = 0.7$

2.6.2 Lois multimodales : mélanges gaussiens

Dans cette partie, on va s'intéresser aux surfaces quantiles de lois avec plusieurs modes. On commence par une loi bimodale, comme mélange de deux gaussiennes

$$P = 1/4N(O_1, \Sigma_1) + 3/4N(O_2, \Sigma_2).$$

2.6.2.a Simulation pour $O_1 = (-5, 0)$, $O_2 = (5, 0)$, $\Sigma_1 = I_2$ et $\Sigma_2 = 3I_2$

Pour commencer, on a volontairement choisi deux modes bien séparés, afin de comprendre l'influence des modes sur la forme de la surface $Q(O, \alpha)$. Dans les simulations nous affichons cette fois encore un échantillon de 10^4 points, afin de rendre compte de la séparation des modes de la loi choisie et du fait qu'avec des échantillons de grandes taille une représentation graphique des point ne suffit pas à rendre compte de la variation de la masse proche des modes. Tracer quelques surfaces quantiles peut servir à représenter les variations des masses autour des modes.

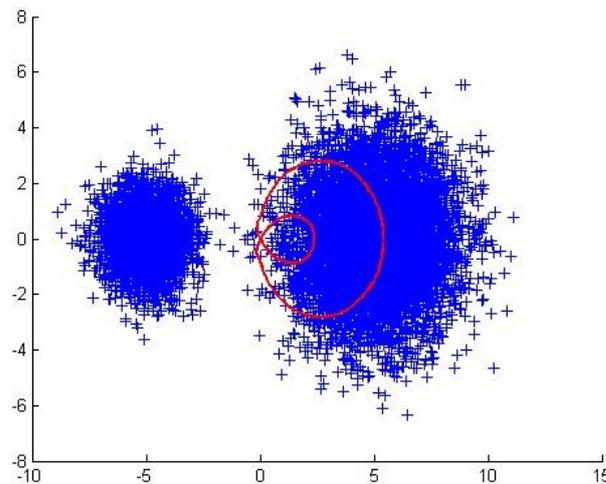


FIGURE 2.22 – $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.7$

De la même manière que dans les sections précédentes, la surface quantile $Q(O, \alpha)$ est en forme de « cardioïde », qui s'explique comme précédemment par l'absence de masse suffisante à exclure dans certaines directions : le mode centré en $(-5, 0)$ est de masse $1/4$ or on a choisi $\alpha = 0.7$ donc on veut exclure 0.3 de masse. En faisant tourner l'observateur le long de la frontière de la grande zone modale, on peut étudier la répartition de la masse autour du mode principale.

Dans la simulation suivante, pour la même loi, on place l'observateur à $O = (5.5, 0)$ à l'intérieur du mode centré en $(5, 0)$ et de masse $3/4$.

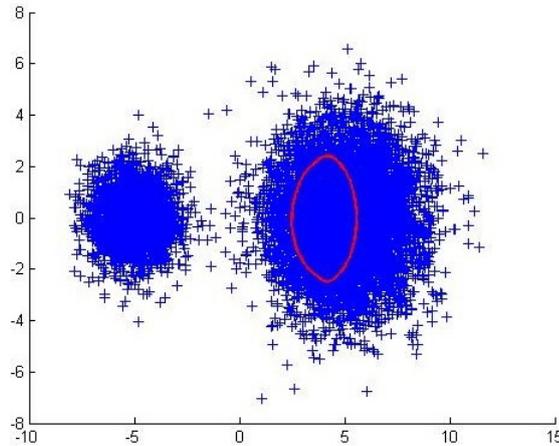


FIGURE 2.23 – $Q(O, \alpha)$ pour $O = (5.5, 0)$ et $\alpha = 0.7$

Dans ce cas on remarque l'absence de la boucle intérieure, grâce au placement dans la zone modale de masse $3/4$, une masse suffisante pour ne pas faire apparaître la boucle intérieure dans la surface quantile.

Par contre il suffit de diminuer α (donc augmenter $1 - \alpha$ la masse à exclure), pour la voir réapparaître.

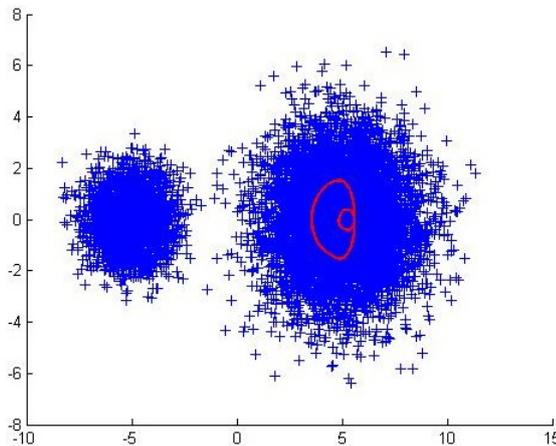


FIGURE 2.24 – $Q(O, \alpha)$ pour $O = (5.5, 0)$ et $\alpha = 0.6$

Si on se place à l'intérieur du mode centré en $(-5, 0)$ et de masse $1/4$,

la surface quantile $Q(O, \alpha)$ avec $O = (-5, 0)$, est composée de deux boucles, une intérieure et une extérieure.

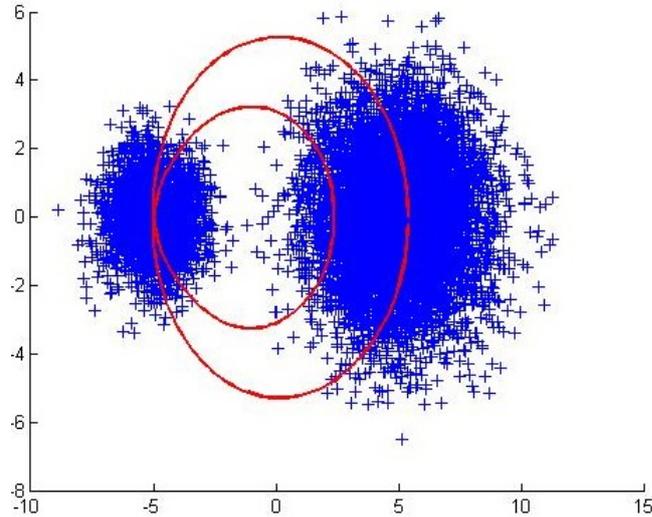


FIGURE 2.25 – $Q(O, \alpha)$ pour $O = (-5, 0)$ et $\alpha = 0.7$

Si on s'intéresse à la surface quantile $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.85$, on voit apparaître les « pincements » comme dans la section précédente, caractéristiques d'une distribution plus étalée sur l'axe des abscisses.

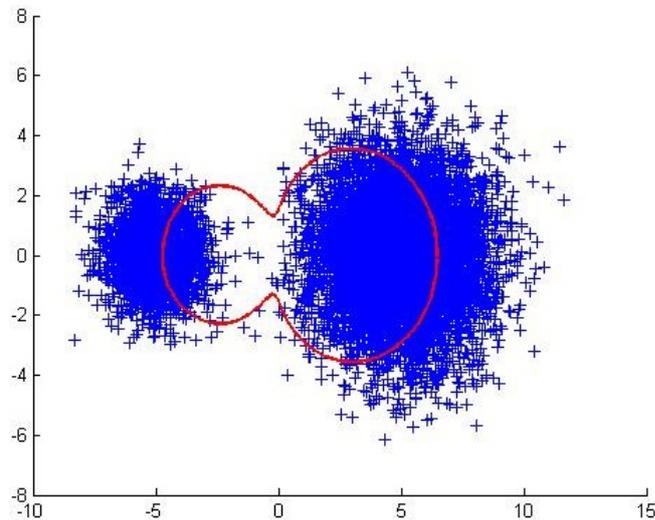


FIGURE 2.26 – $Q(O, \alpha)$ pour $O = (-5, 0)$ et $\alpha = 0.7$

On remarque tout de même une différence avec la figure 2.18, l'asymétrie de la surface étant due à l'asymétrie de la loi. Ensuite si on place l'observateur à $O = (0, 8)$, la surface $Q(O, \alpha)$ avec $\alpha = 0.7$ est une « cardioïde » avec la boucle intérieure qui pointe vers la masse.

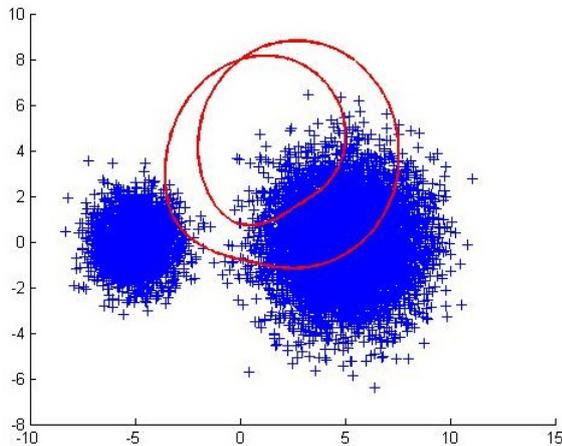


FIGURE 2.27 – $Q(O, \alpha)$ pour $O = (0, 8)$ et $\alpha = 0.7$

Dans la figure suivante, on pose $O = (0, 0)$ et trace les surfaces quantiles $Q(O, \alpha)$ pour $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$.

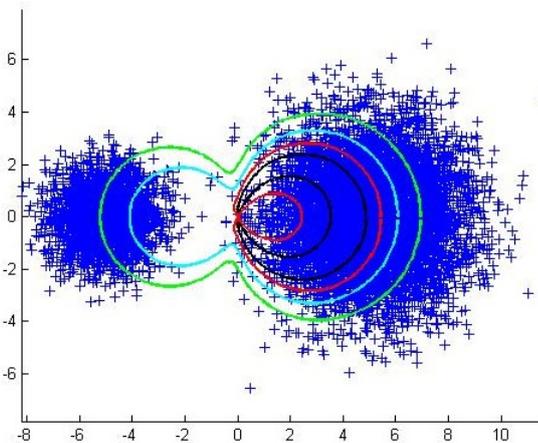


FIGURE 2.28 – $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$

- Dans le graphique précédent, l'observateur est fixé, $O = (0, 0)$, avec
- La surface en noir $Q(O, \alpha)$ pour $\alpha = 0.6$.
 - La surface en rouge $Q(O, \alpha)$ pour $\alpha = 0.7$.

- La surface en cyan $Q(O, \alpha)$ pour $\alpha = 0.8$.
- La surface en vert $Q(O, \alpha)$ pour $\alpha = 0.9$.

On remarque qu'en passant le niveau seuil $\alpha = 0.75$ on passe d'une surface plus concentré autour du mode centré en $(5,0)$ et de masse 0.75 avec une boucle intérieure, à une forme sans boucle intérieure mais plus étalée sur le support de la loi. Ce passage se fait ici de manière discontinue à cause de l'absence de masse entre les deux modes (existence de bandes de probabilité nulle), ceci est dû à la manière avec laquelle sont tracées les surfaces quantile, car elles sont faites à partir d'une mesure empirique et non de la vraie mesure pour laquelle la transition est continue en α .

Voici, l'évolution d'une surface quantile, lorsque O parcourt $\{(-8,0), (-5,0), (0,0), (5,0)\}$

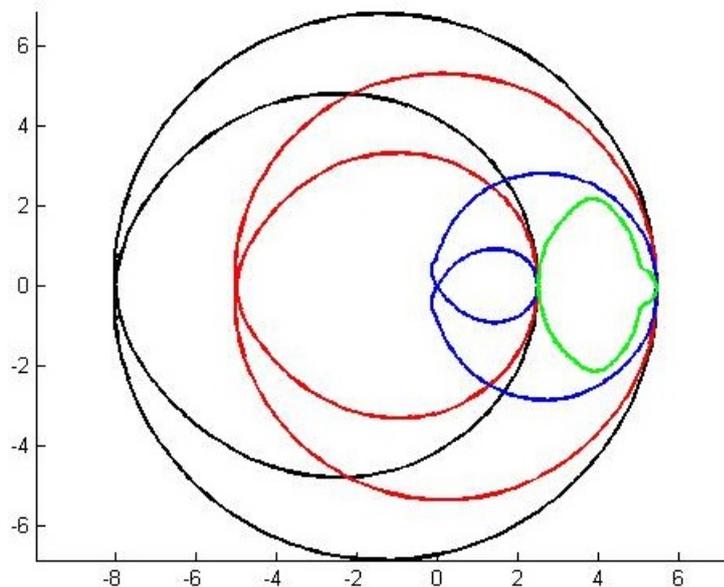


FIGURE 2.29 – $Q(O, \alpha)$ pour $O \in \{(-8,0), (-5,0), (0,0), (5,0)\}$ et $\alpha = 0.7$

Dans le graphique précédent

- En noir $Q(O, \alpha)$ pour $O = (-8,0)$.
- En rouge $Q(O_2, \alpha)$ pour $O_2 = (-5,0)$.
- En bleu $Q(O_3, \alpha)$ pour $O_3 = (0,0)$.
- En vert $Q(O_4, \alpha)$ pour $O_4 = (5,0)$.

Ainsi, on fait les mêmes remarques que dans les cas précédents, plus on s'éloigne de la masse, plus le volume de la boucle intérieure est important,

avec cette dernière qui pointe toujours vers la masse.

Enfin, si on s'intéresse à la « localisation de la masse » voici $Q(O, \alpha)$ pour $O \in \{(-8, -8), (-8, 8), (8, -8), (8, 8)\}$

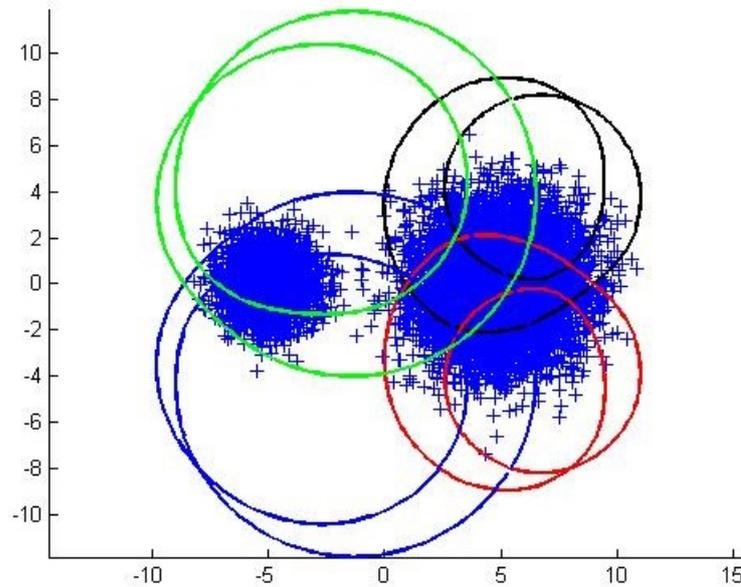


FIGURE 2.30 – $Q(O, \alpha)$ pour $O \in \{(-8, -8), (-8, 8), (8, -8), (8, 8)\}$ et $\alpha = 0.7$

Dans le graphique précédent

- En noir $Q(O_1, \alpha)$ pour $O_1 = (8, 8)$.
- En rouge $Q(O_2, \alpha)$ pour $O_2 = (8, -8)$.
- En bleu $Q(O_3, \alpha)$ pour $O_3 = (-8, -8)$.
- En vert $Q(O_4, \alpha)$ pour $O_4 = (-8, 8)$.

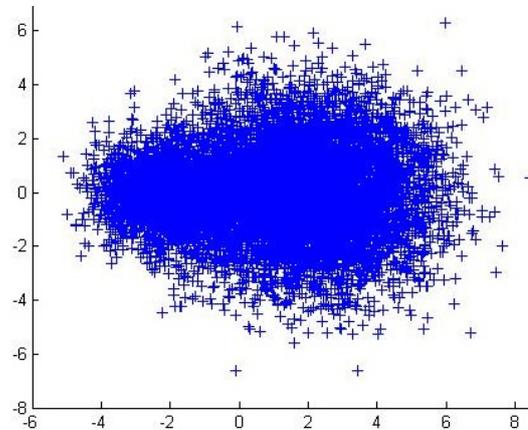
On note que toutes les « boucles intérieures » pointent vers la masse.

2.6.2.b Simulation pour $O_1 = (-2, 0)$, $O_2 = (2, 0)$, $\Sigma_1 = I_2$ et $\Sigma_2 = 3I_2$

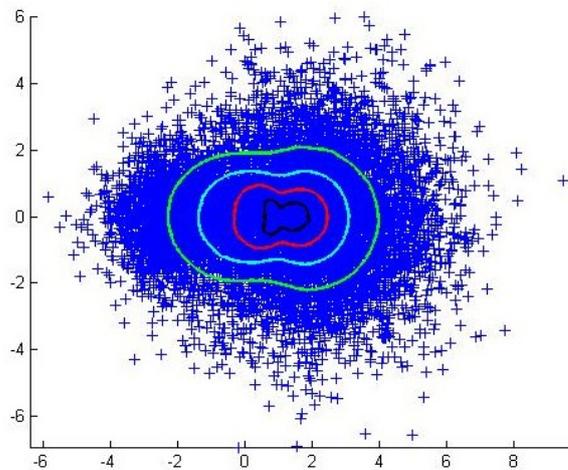
Cette fois, on rapproche les centres des gaussiennes du mélange, en posant $O_1 = (-2, 0)$, $O_2 = (2, 0)$ et

$$P = 1/4N(O_1, I_2) + 3/4N(O_2, 3I_2).$$

Ainsi, on est plus en face d'une loi avec deux modes « éloignés », dont voici un échantillon de 10000 points.

FIGURE 2.31 – Échantillon de 10^4 points de la loi P

En premier lieu, on s'intéresse à $Q(O, \alpha)$ pour $O = E(X) = (1, 0)$ et $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$

FIGURE 2.32 – $Q(O, \alpha)$ pour $O = (1, 0)$ et $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$

Dans le graphique précédent

- En noir $Q(O, \alpha)$ pour $\alpha = 0.6$.
- En rouge $Q(O, \alpha)$ pour $\alpha = 0.7$.
- En cyan $Q(O, \alpha)$ pour $\alpha = 0.8$.

– En vert $Q(O, \alpha)$ pour $\alpha = 0.9$.

Ensuite, voici l'évolution d'une surface quantile, lorsque O parcourt $\{(-4, 0), (-2, 0), (0, 0), (2, 0)\}$

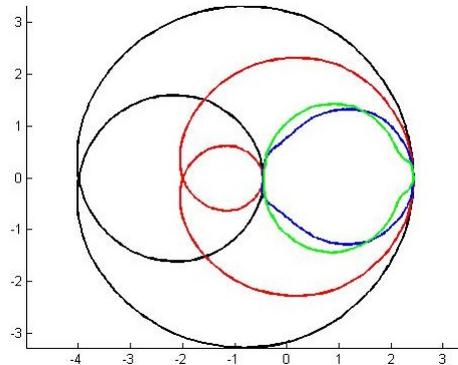


FIGURE 2.33 – $Q(O, \alpha)$ pour $O \in \{(-4, 0), (-2, 0), (0, 0), (2, 0)\}$ et $\alpha = 0.7$

Dans le graphique précédent

- En noir $Q(O, \alpha)$ pour $O = (-4, 0)$.
- En rouge $Q(O_2, \alpha)$ pour $O_2 = (-2, 0)$.
- En bleu $Q(O_3, \alpha)$ pour $O_3 = (0, 0)$.
- En vert $Q(O_4, \alpha)$ pour $O_4 = (2, 0)$.

Pour finir, on s'intéresse à la « localisation de la masse »

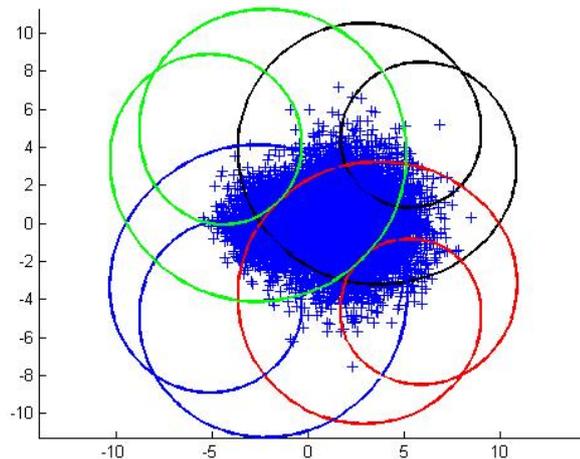


FIGURE 2.34 – $Q(O, \alpha)$ pour $O \in \{(-8, -8), (-8, 8), (8, -8), (8, 8)\}$ et $\alpha = 0.7$

Dans le graphique précédent

- En noir $Q(O_1, \alpha)$ pour $O_1 = (8, 8)$.
- En rouge $Q(O_2, \alpha)$ pour $O_2 = (8, -8)$.
- En bleu $Q(O_3, \alpha)$ pour $O_3 = (-8, -8)$.
- En vert $Q(O_4, \alpha)$ pour $O_4 = (-8, 8)$.

Comme précédemment, les « boucles intérieures » pointent vers la masse.

Remarque 2.65. *Il est très important de noter que les surfaces quantiles fournissent un bon outil de localisation de masse, néanmoins comme on a vu dans les simulations précédentes, ces surfaces ne permettent pas (lorsqu'elles sont issues de demi-espaces) de faire de la détection de modes. Dans les chapitres suivants nous verrons que ceci devient possible lorsque la classe des demi-espaces est remplacée par une classe « plus discriminante », en l'occurrence des formes courbées.*

2.6.3 Lois non usuelles

Dans cette dernière section, on s'intéresse à des lois non usuelles, pas forcément avec une densité par rapport à la mesure de Lebesgue et qui vérifie les conditions (H_0^-) et (H_0^+) .

2.6.3.a Loi à support spiral

On s'intéresse à une loi sur \mathbb{R}^2 avec un support spiral, i.e. supportée sur un sous-espace de dimension inférieure (dimension 1), dont voici un échantillon de 10000 points. dans cet exemple on prends $X = (r \cos(\theta), r \sin(\theta))$ avec $\theta \sim \mathcal{U}([0, 8\pi])/4$ et $r = 1/4 + \theta$.

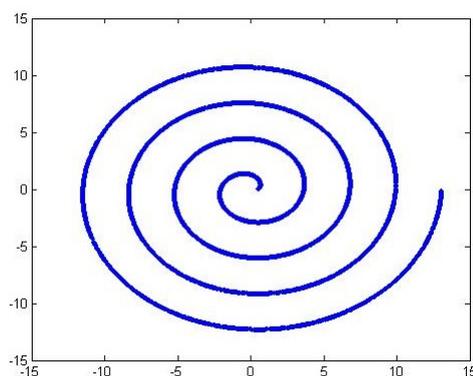


FIGURE 2.35 – Échantillon de la loi P avec $N = 10000$

On commence par remarquer qu'aucune partie du support de la loi n'est linéaire donc pas de projection en masse Dirac. La loi P vérifie les conditions (H_0^-) et (H_0^+) .

Voici maintenant $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$

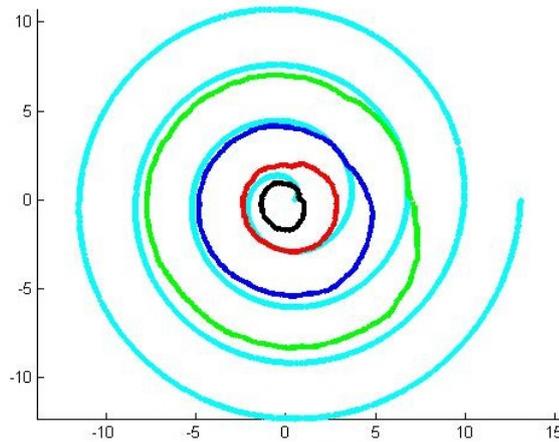


FIGURE 2.36 – $Q(O, \alpha)$ pour $O = (0, 0)$ et $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$

Dans le graphique précédent

- En cyan (X_1, X_2, \dots, X_N) avec $N = 10000$, un échantillon de la loi P .
- En noir $Q(O, \alpha)$ pour $\alpha = 0.6$.
- En rouge $Q(O, \alpha)$ pour $\alpha = 0.7$.
- En bleu $Q(O, \alpha)$ pour $\alpha = 0.8$.
- En vert $Q(O, \alpha)$ pour $\alpha = 0.9$.

Ensuite voici $Q(O, \alpha)$ pour $O = (-8, 0)$ et $\alpha = 0.7$.

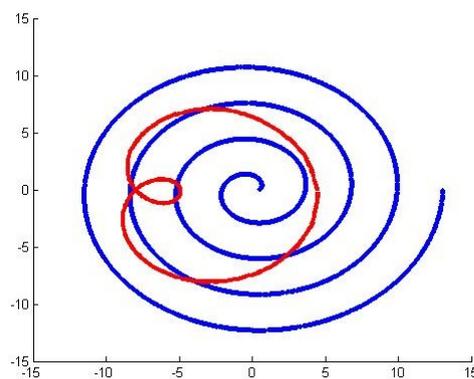


FIGURE 2.37 – $Q(O, \alpha)$ pour $O = (-8, 0)$ et $\alpha = 0.7$

On ne remarque aucune différence de formes possibles de la surface quantile avec les lois à densité par rapport à la mesure de Lebesgue, vue dans la section précédente. La surface quantile est composée de deux boucles : une extérieure et une intérieure qui « pointe » dans la direction de la masse.

2.6.3.b Une loi spéciale

Dans ce paragraphe, on construit une loi qui permet de donner un contre exemple à la supposition suivante : « la surface quantile est toujours composée d'au plus deux boucles. »

Pour ce faire, on fixe un observateur O dans \mathbb{R}^2 et $\alpha = a + b + c$ puis on trace trois demi-espaces de masse α dont la frontière passe par O comme suit :

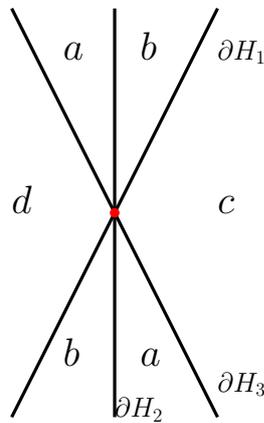


FIGURE 2.38 – En rouge O et en noir la frontière de trois demi-plans avec $a + b + c + d = 1$

On a $P(H_1) = c + a + b = \alpha$, $P(H_2) = b + c + a = \alpha$ et $P(H_3) = a + b + c = \alpha$. Ainsi si on pose $H_1 = H(u, \alpha)$, $H_2 = H(v, \alpha)$ et $H_3 = H(w, \alpha)$ pour $u, v, w \in \mathbb{S}_{d-1}$ on a

$$O = Q(O, u, \alpha) = Q(O, v, \alpha) = Q(O, w, \alpha),$$

donc la surface quantile passe au moins trois fois par O , si on suppose que $a = 0$ alors pour toute direction u' dans le cône de direction entre u et v on a $Q(O, u', \alpha) = 0$. de même pour $b = 0$, de ce fait on suppose que

$$a \neq 0 \text{ et } b \neq 0$$

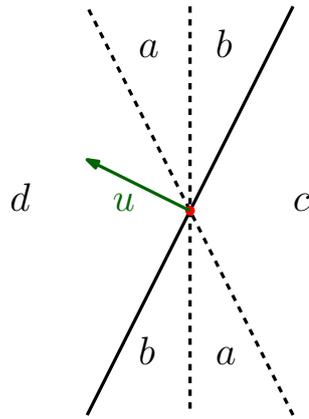


FIGURE 2.39 – En rouge O et en noir $H_1 = H(u, \alpha)$

On commence par s'intéresser au comportement de la surface au voisinage de u .

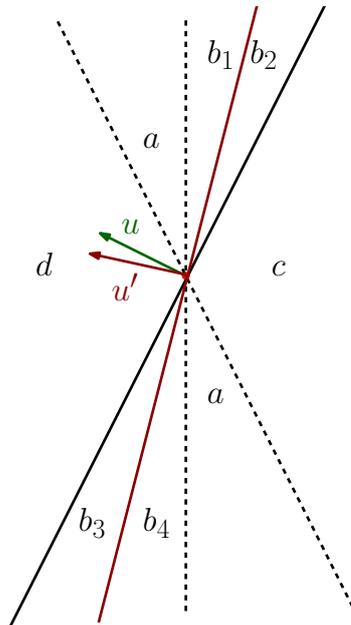


FIGURE 2.40 – En noir $H_1 = H(u, \alpha)$ et en rouge H'_1

Pour u' dans le cône de direction entre u et v , on a $P(H'_1) = a + c + b_2 + b_4$,

en supposant que $b_2 > b_3$ alors on a $P(H(u', \alpha)) > \alpha$ car

$$\begin{aligned}
 P(H'_1) &= a + c + b_2 + b_4 \\
 &= a + b + c - b + b_2 + b_4 \\
 &= \alpha + b_2 + (b_4 - b) \\
 &= \alpha + (b_2 - b_3) \\
 &> \alpha \quad \text{car } b_2 > b_3
 \end{aligned}$$

ainsi il faut « reculer » H'_1 pour avoir α de masse, comme suit

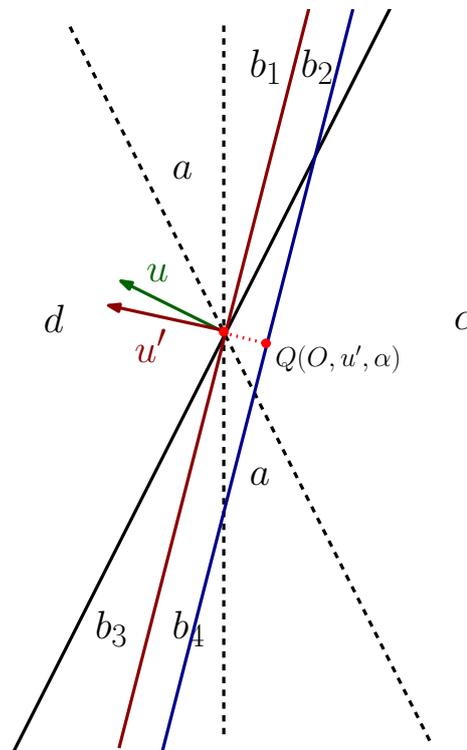
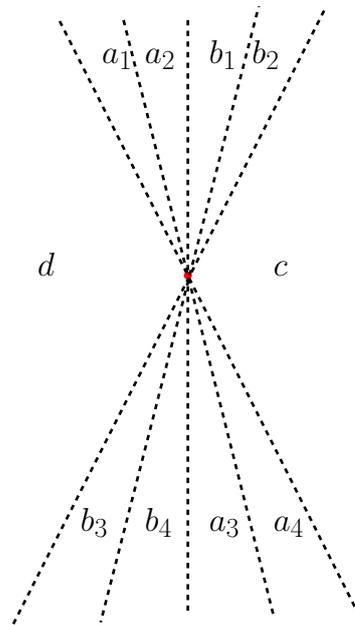
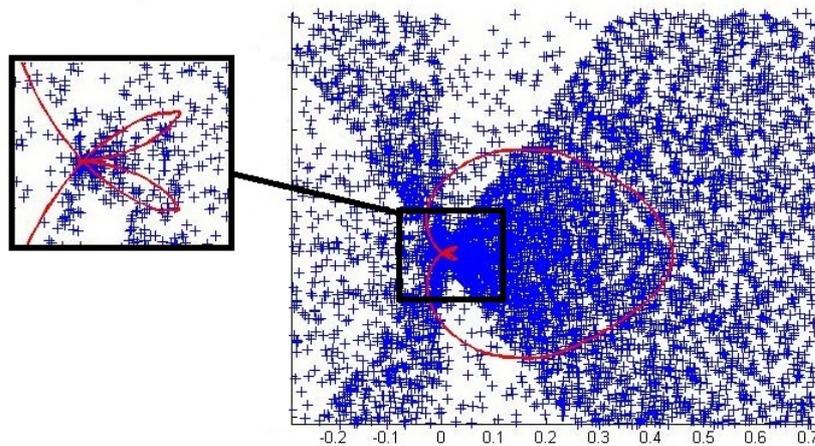


FIGURE 2.41 – En noir $H_1 = H(u, \alpha)$, en rouge H'_1 et en bleu $H(u', \alpha)$

En faisant de manière analogue au voisinage de v et w , on obtient le découpage analogue des zones de masse a , dont la zone supérieure est découpée en $a = a_1 + a_2$ et la zone du bas du plan en $a = a_3 + a_4$ avec $a_2 > a_3$ et $a_1 < a_4$ comme dans la figure suivante.

FIGURE 2.42 – découpage de l'espace par la loi P

Les simulations ci-dessous rendent compte de la forme de la surface quantile pour $\alpha = 0.7$: on voit apparaître deux boucles intérieures et une extérieure.

FIGURE 2.43 – $Q(O, \alpha)$ avec $O = (0, 0)$ et $\alpha = 0.7$

Ce phénomène est local en O et en α , il suffit de faire varier l'observateur ou le niveau pour qu'il ne reste qu'une boucle intérieure.

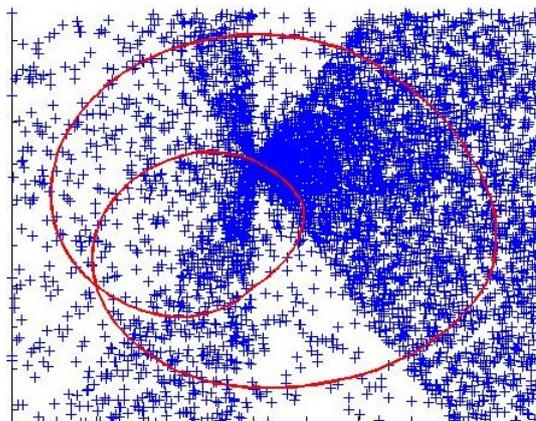


FIGURE 2.44 – $Q(O, \alpha)$ avec $O = (-0.3, -0.3)$ et $\alpha = 0.7$

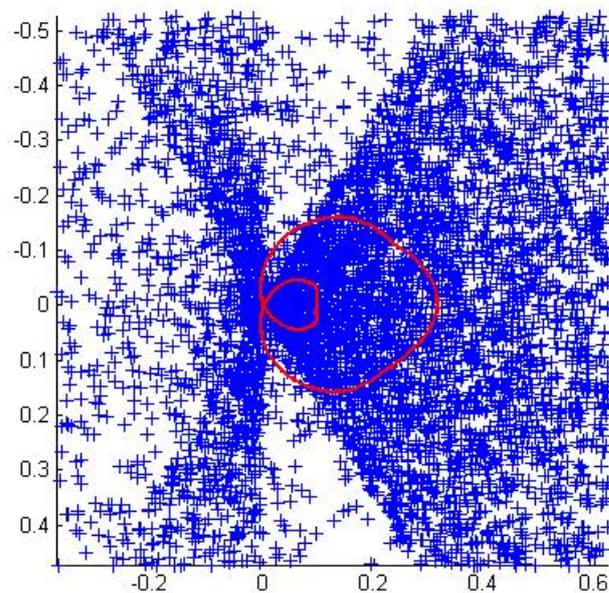


FIGURE 2.45 – $Q(O, \alpha)$ avec $O = (15, 0)$ et $\alpha = 0.8$

Chapitre 3

Surfaces quantiles via des classes générales

Dans ce chapitre on s'intéresse à la généralisation des surfaces quantiles via des demi-espaces aux surfaces quantiles via des classes générales (fonctions, surfaces, géodésiques, ...). On y présente des résultats généraux, admettant pour la plupart comme cas particuliers certains des résultats présentés dans le chapitre précédent.

Dans la première section, on commence par motiver cette nouvelle généralisation, avant de nous intéresser dans la partie suivante aux surfaces quantiles généralisées indexées par une fonction. La troisième partie quant à elle, est consacrée aux énoncés des résultats dont les preuves sont données dans la quatrième partie. Ensuite, on introduit différentes généralisations possibles, sans donner de résultats.

Enfin la dernière partie de ce chapitre est consacrée aux différentes utilisations du quantile indexées par une fonction, sous l'angle de la localisation de masse et de la détection de modes, qui sont une motivation pour le chapitre suivant.

3.1 Introduction

Dans cette partie, il s'agit de généraliser la classe des demi-espaces à des classes plus discriminantes et beaucoup plus riches. Intuitivement la surface quantile $Q_X(O, \alpha)$ définie dans le chapitre précédent via la classe des demi-espaces est la collection de tous les quantiles de niveau α de toutes les combinaisons linéaires $\langle X - O, u \rangle$ pour $u \in \mathbb{S}_{d-1}$. Notre généralisation consiste à quitter ce cadre linéaire en remplaçant le produit scalaire par une autre fonction de la direction u , de l'observateur O et de X .

Il s'agit de discriminer les points de l'espace de manière complètement différente, comme on verra plus tard. Ces surfaces généralisées permettent de décrire la loi de manière différente, certaines classes seront plus adaptées pour la détection des modes d'autres pour la localisation de masse,...

La motivation principale pour cette généralisation est la construction de méthodes automatiques dirigées par les données, afin de localiser la masse d'une loi quelconque et non nécessairement à densité par rapport à la mesure de Lebesgue, de détecter les "modes", ainsi que d'introduire différentes notions de profondeur par niveau de quantile et de comparer des lois, détection de rupture, tester de symétrie, etc...

Les surfaces généralisées étant extrêmement sensibles à la position de l'observateur O , et sans transformation évidente pour passer de la surface vue de O à celle vue de O' elles sont d'autant plus informatives que O est bien choisi ou qu'on compare différents points de vue. En particulier, si les formes servant à construire les surfaces sont bien plus pointues que des demi-espaces, alors positionner O trop en dehors des zones modales peut conduire à des surfaces quantiles avec fortes variations en terme de distance à O , car pour capturer de la masse il faut projeter la pointe de la forme fort loin. Au contraire, avec O dans la zone modale principale les variations de la surface quantile seront très significatives et pointeront bien les directions de faible ou forte masse à proximité de O . Le but d'une méthode automatisée sera alors, après étude approfondie de tels phénomènes, de faire bouger O de façon itérative et guidée par les informations extraites automatiquement de l'échantillon via les O précédents.

3.2 Surfaces quantiles généralisées indexées par une fonction

Dans toute la suite on pose $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et on considère X une variable aléatoire de $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, comme précédemment on note $P := \mathbb{P}_X$.

3.2.1 Définitions

Cette partie est consacrée au cas où la classe discriminante est une fonction : sont exclues les formes fermées, formes en tube et d'autres cas intéressants, qui seront proposées comme seconde généralisation.

Soit $O \in \mathbb{R}^d$, $u_0 = (0, \dots, 0, 1)$, $u \in \mathbb{S}_{d-1}$ et posons $\varphi : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$ une

fonction continue avec $\varphi(0_{d-1}) = 0$ et

$$\varphi^{-1}(\mathbb{R}) = \mathbb{R}^{d-1}.$$

Ces conditions sur la fonction φ , sont appelées dans toute la suite du document les « Conditions (3.2.1) ».

On note p_φ la fonction projection sur $(0_d, u_0)$ suivant les géodésiques parallèles au graphe de φ ,

$$\begin{aligned} p_\varphi : \mathbb{R}^d &\longrightarrow \mathbb{R}^d \\ x &\longmapsto (0, \dots, 0, x_d - \varphi(x_1, \dots, x_{d-1})). \end{aligned}$$

Notation 3.1. Soit $u \in \mathbb{S}_{d-1}$, φ vérifiant les conditions précédentes et R_u la rotation dans $E_u = \text{vect}(u, u_0)$ de centre 0_d et d'angle $\theta_u = \widehat{(u_0, u)}$, on note :

- $r_{\theta_u} = R_u \circ \text{proj}_{E_u} + \text{proj}_{E_u^\perp}$ qui correspond à une rotation dans \mathbb{R}^d laissant invariant E_u^\perp
- $r_{-\theta_u} = R_u^{-1} \circ \text{proj}_{E_u} + \text{proj}_{E_u^\perp}$
- t_O la translation de vecteur $0_d - O$
- t_{-O} la translation de vecteur $O - 0_d$

On pose $p_{O,u,\varphi}$, la fonction projection sur (O, u) suivant les géodésiques parallèles à φ

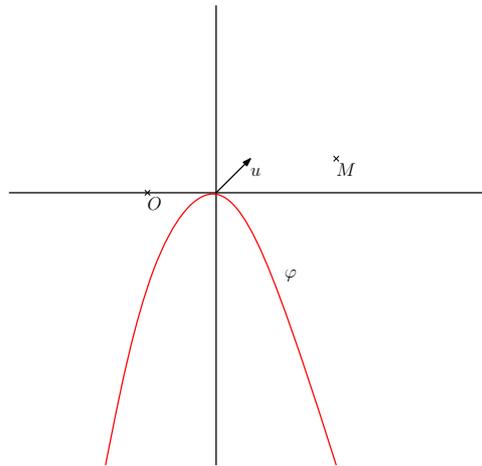
$$p_{O,u,\varphi} = t_{-O} \circ r_{\theta_u} \circ p_\varphi \circ r_{-\theta_u} \circ t_O.$$

Les quantiles généralisés seront définis dans la direction u à l'aide des projetés $p_{O,u,\varphi}$ de l'échantillon, ordonnés selon ces valeurs de projection.

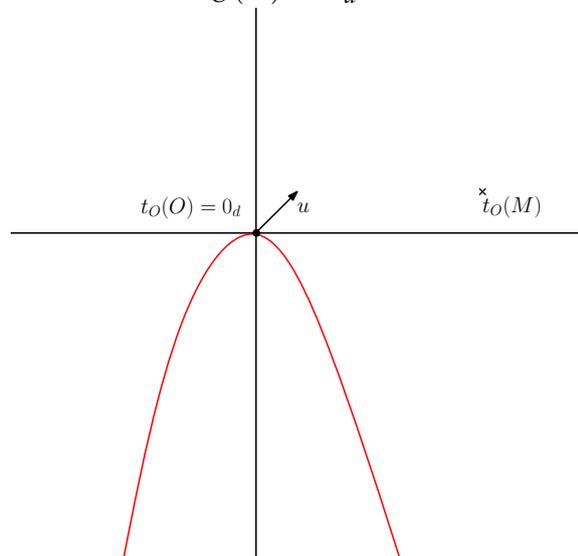
Remarque 3.2. Remarquons que $p_\varphi = p_{0_d, u_0, \varphi}$, car $r_{\theta_{u_0}} = \text{Id}_d$ et $t_{0_d} = t_{-0_d} = \text{Id}_d$. Ensuite puisque φ est supposée continue alors $(O, u, x) \mapsto p_{O,u,\varphi}(x)$ est continue comme composition de fonctions continues.

Remarque 3.3. Voici la méthode de projection expliquée point par point sur \mathbb{R}^2 et $\varphi(x) = -x^2$:

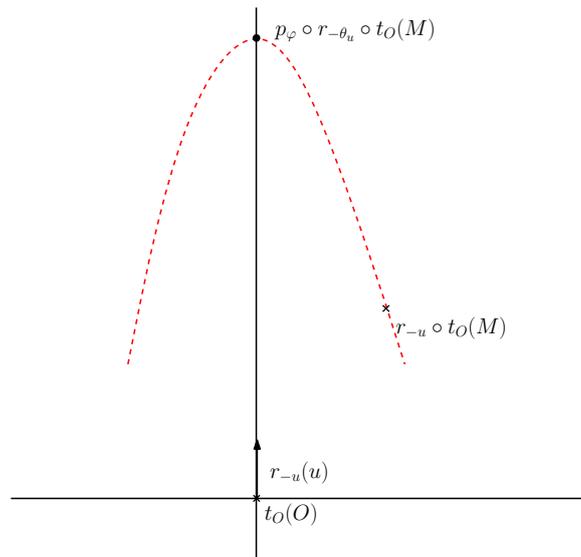
- **Étape I :** on fixe une fonction φ vérifiant les conditions (3.2.1), un point d'observation $O \in \mathbb{R}^d$, la direction de projection $u \in \mathbb{S}_{d-1}$ et le point à projeter $M \in \mathbb{R}^d$.



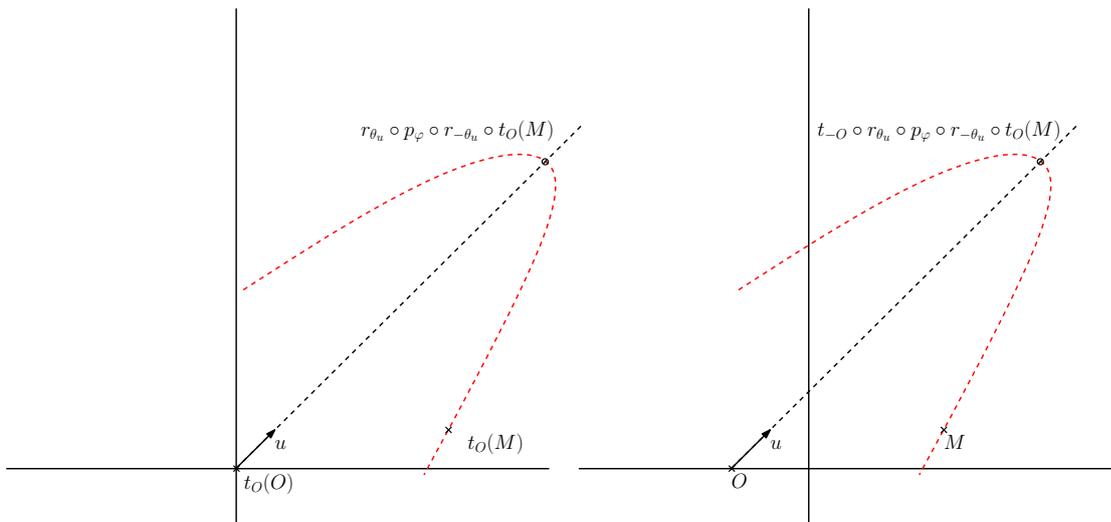
- **Étape II** : on translate le point à projeter M au point $t_O(M) = M - O$ et le point d'observation O à $t_O(O) = 0_d$.



- **Étape III** : on applique la rotation de centre 0_d et d'angle $-\widehat{(u_0, u)}$ au point $t_O(M)$ pour obtenir $M' = r_{-\theta_u}(t_O(M))$ puis on projete M' avec p_φ la projection par φ sur la droite (O, u_0) .



- **Étape IV** : on applique la rotation de centre 0_d et d'angle $\widehat{(u_0, u)}$ puis la translation de vecteur $O - 0_d$ au point $r_{-\theta_u}(t_O(M))$ pour récupérer le point M ainsi que sa projection $p_{\varphi}(r_{-\theta_u}(t_O(M)))$ pour obtenir sa projection dans la direction u vue de O noté $p_{O,u,\varphi}(M) = t_{-O} \circ r_{\theta_u} \circ p_{\varphi} \circ r_{-\theta_u} \circ t_O(M)$.



Exemple 3.4. Soit $O \in \mathbb{R}^2$, $u \in \mathbb{S}_1$ et φ vérifiant 3.2.1. si φ est paire i.e.

$$\forall x \in \mathbb{R}, \varphi(x) = \varphi(-x)$$

alors

$$p_{O,u,\varphi} = p_{O,-u,-\varphi}$$

Montrons tout d'abord que $p_\varphi := p_{0_2,u_0,\varphi} = p_{0_2,-u_0,-\varphi}$.
On sait que

$$p_{0_2,-u_0,-\varphi} = r_{-\pi} \circ p_{-\varphi} \circ r_\pi$$

de plus on a

$$r_{-\pi} = r_\pi = -Id_2$$

donc pour tout $x \in \mathbb{R}^d$ on a

$$\begin{aligned} p_{0_2,-u_0,-\varphi}(x) &= r_{-\pi} \circ p_{-\varphi} \circ r_\pi(x) \\ &= -p_{-\varphi}(-x) \\ &= -(0, -x_2 - (-\varphi(-x_1))) \\ &= (0, x_2 - \varphi(x_1)) \\ &= p_\varphi := p_{0_2,u_0,\varphi}(x) \end{aligned}$$

ensuite on a

$$\theta_{-u} = (\widehat{u_0, -u}) = (\widehat{u_0, u}) + (\widehat{u, -u}) = \theta_u - \pi$$

donc pour tout $x \in \mathbb{R}^d$

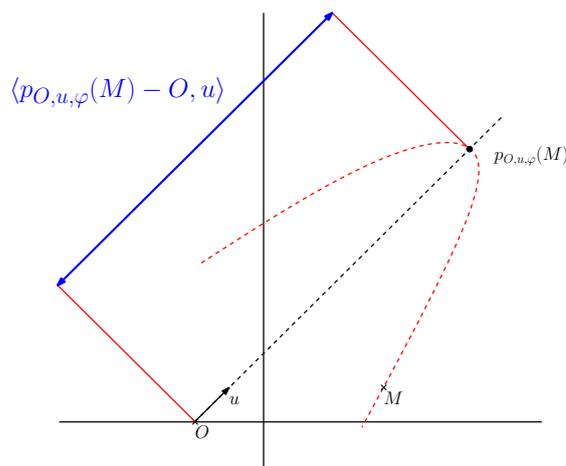
$$\begin{aligned} p_{O,-u,-\varphi}(x) &= t_{-O} \circ r_{\theta_{-u}} \circ p_{-\varphi} \circ r_{-\theta_{-u}} \circ t_O(x) \\ &= t_{-O} \circ r_{\theta_u} \circ r_{-\pi} \circ p_{-\varphi} \circ r_\pi \circ r_{-\theta_u} \circ t_O(x) \\ &= t_{-O} \circ r_{\theta_u} \circ p_{0_d,-u_0,-\varphi} \circ r_{-\theta_u} \circ t_O(x) \\ &= t_{-O} \circ r_{\theta_u} \circ p_\varphi \circ r_{-\theta_u} \circ t_O(x) \\ &= p_{O,u,\varphi}(x). \end{aligned}$$

Voici la définition de la région sous le graphe de φ vue de O dans la direction u et à distance y de O .

Définition 3.5. Soit $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $y \in \mathbb{R}$. on définit la région sous le graphe de φ vue de O dans la direction u et à distance y de O de la façon suivante

$$\dot{G}_\varphi(O, u, y) = \{x \in \mathbb{R}^d : \langle p_{O,u,\varphi}(x) - O, u \rangle \leq y\}.$$

Remarque 3.6. Pour tout $x \in \mathbb{R}^d$ on a $p_{O,u,\varphi}(x) \in (O, u)$ donc $p_{O,u,\varphi}(x) - O$ est colinéaire à u .



Notation 3.7. Dans la suite on notera

$$\pi_{\varphi,O}(x, u) = \langle p_{O,u,\varphi}(x) - O, u \rangle.$$

Par continuité de $(O, u, x) \mapsto p_{O,u,\varphi}(x)$ ainsi que du produit scalaire on obtient la continuité de $(O, u, x) \mapsto \pi_{\varphi,O}(x, u)$.

Voici la définition du α -quantile à travers la fonction φ vu de O dans la direction u .

Définition 3.8. Soit $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in (0, 1)$. Le α -quantile par la fonction φ , de la loi P vu de O dans la direction u et de niveau α est défini par

$$\dot{Y}_{\varphi}(O, u, \alpha) = \inf\{y \in \mathbb{R} : P(\dot{G}_{\varphi}(O, u, y)) \geq \alpha\}.$$

Ensuite, on définit les analogues au point quantile $Q(O, u, \alpha)$, le demi-plan de masse α , ainsi que l'ensemble quantile $Q_{\alpha}(O)$ du Chapitre 2.

Définition 3.9. Soit $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in (0, 1)$.

- Le point α -quantile par la fonction φ , de la loi P vu de O dans la direction u et de niveau α est défini par

$$\dot{Q}_{\varphi}(O, u, \alpha) = O + \dot{Y}_{\varphi}(O, u, \alpha) \cdot u.$$

- La φ -région quantile vue de O de masse α est définie par

$$\dot{G}_{\varphi}(O, u, \alpha) := \dot{G}_{\varphi}(O, u, \dot{Y}_{\varphi}(O, u, \alpha)).$$

- L'ensemble α -quantile par la fonction φ , de la loi P vu de O dans la direction u et de niveau α est défini par

$$\dot{Q}_{\varphi}(O, \alpha) = \{\dot{Q}_{\varphi}(O, u, \alpha) : u \in \mathbb{S}_{d-1}\}.$$

Soit (X_i) une suite de variables aléatoires indépendantes de même loi P . On note $P_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ la mesure empirique associée à l'échantillon (X_1, \dots, X_n) . De manière analogue, les versions empiriques des objets présentés dans cette section sont définis comme suit.

Définition 3.10. Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$, $\alpha \in (0, 1)$, on définit :

- le φ -quantile empirique de la loi P vu de O dans la direction u et de niveau α par

$$\dot{Y}_{\varphi,n}(O, u, \alpha) = \inf\{y \in \mathbb{R} : P_n(\dot{G}_\varphi(O, u, y)) \geq \alpha\}.$$

- le point φ -quantile empirique de la loi P vu de O dans la direction u et de niveau α par

$$\dot{Q}_{\varphi,n}(O, u, \alpha) = O + \dot{Y}_{\varphi,n}(O, u, \alpha) \cdot u.$$

- la φ -région empirique vue de O de masse α est définie par

$$\dot{G}_{\varphi,n}(O, u, \alpha) := \dot{G}_\varphi(O, u, \dot{Y}_{\varphi,n}(O, u, \alpha)).$$

- La φ -surface empirique α -quantile par la fonction φ , de la loi P vu de O dans la direction u et de niveau α par

$$\dot{Q}_{\varphi,n}(O, \alpha) = \{\dot{Q}_{\varphi,n}(O, u, \alpha) : u \in \mathbb{S}_{d-1}\}.$$

Remarque 3.11. Le cas des demi-plans du chapitre précédent correspond au cas particulier où la fonction $\varphi = 0$.

3.2.2 Régularité des φ -surfaces quantile

Dans toute la suite on fixe O ainsi que la fonction φ de plus on suppose qu'elle vérifie les conditions (3.2.1).

3.2.2.a Hypothèses

On pose $\Delta = [\alpha^-, \alpha^+] \subset [1/2, 1)$. Avant d'énoncer les hypothèses et les théorèmes garantissant la régularité des φ -surfaces quantile, on donne la définition d'un ruban de largeur $\varepsilon > 0$, un objet qui joue le même rôle que les bandes de largeur ε dans le cas des demi-plans.

Définition 3.12. Soit $\varepsilon > 0$, $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $y \in \mathbb{R}$. On définit le φ -Ruban de largeur ε par

$$\dot{R}_\varphi(O, u, y, y + \varepsilon) = \dot{G}_\varphi(O, u, y + \varepsilon) \setminus \dot{G}_\varphi(O, u, y)$$

Voici les hypothèses suffisantes à garantir la bonne régularité des φ -surfaces quantile.

Hypothèses. Soit $O \in \mathbb{R}^d$ fixé. Pour tout $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$ on suppose

$$(\dot{H}_0^-) P(\partial \dot{G}_\varphi(O, u, \alpha)) = 0.$$

$$(\dot{H}_0^+) P(\dot{R}_\varphi(O, u, y, z)) > 0 \text{ si } Y_\varphi(O, u, \alpha^+) \geq z > y \geq \dot{Y}_\varphi(O, u, \alpha^-).$$

Remarque 3.13. – La condition (\dot{H}_0^-) permet d'éviter que la loi projetée ait des composantes masses de Dirac. De plus on a $P(\dot{G}_\varphi(O, u, \alpha)) = \alpha$.

- Les conditions (\dot{H}_0^-) et (\dot{H}_0^+) impliquent que l'application $y \mapsto F_{\pi_{\varphi, O}(X, u)}(y)$ est continue et strictement croissante sur $F_{\pi_{\varphi, O}(X, u)}^{-1}(\Delta)$.
- Les conditions (\dot{H}_0^-) et (\dot{H}_0^+) sont satisfaites si P admet une densité $f = \frac{dP}{d\lambda_d} > 0$ par rapport à la mesure de Lebesgue sur \mathbb{R}^d .

Définition 3.14. L'ensemble des régions φ -quantile vue de $O \in \mathbb{R}^d$ est défini par

$$\dot{G}_{O, \Delta} = \{\dot{G}_\varphi(O, u, y) : u \in \mathbb{S}_{d-1}, y \in [\dot{Y}_\varphi(O, u, \alpha^-), \dot{Y}_\varphi(O, u, \alpha^+)]\}$$

Afin d'alléger les notations, on note simplement $\dot{\mathcal{G}} := \dot{G}_{O, \Delta}$ car O et Δ sont toujours fixés.

Une conséquence importante de l'hypothèse (\dot{H}_0^-) est le contrôle suivant des probabilités empiriques en chaque quantile empirique.

Proposition 3.15. Sous la condition (\dot{H}_0^-) , pour tout $n > d$, $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$, on a

$$\alpha \leq P_n(\dot{G}_{\varphi, n}(O, u, \alpha)) \leq \alpha + \frac{d}{n} \text{ p.s.}$$

Démonstration. On suppose (\dot{H}_0^-) et on veut montrer qu'avec probabilité un on a, pour tout O, u, α et tout $n > d$

$$\alpha \leq P_n(\dot{G}_\varphi(O, u, \dot{Y}_{\varphi, n}(O, u, \alpha))) = P_n(\dot{G}_{\varphi, n}(O, u, \alpha)) \leq \alpha + \frac{d}{n}.$$

Par définition de $\dot{Y}_{\varphi, n}(O, u, \alpha)$ on a

$$P_n(\dot{G}_\varphi(O, u, \dot{Y}_{\varphi, n}(O, u, \alpha))) = P_n(\dot{G}_{\varphi, n}(O, u, \alpha)) \geq \alpha.$$

Fixons $n > d$. Soit $\partial G(x_1, \dots, x_d)$ avec $G(x_1, \dots, x_d) \in \dot{\mathcal{G}}$ l'unique hypersurface passant par d points distincts x_1, \dots, x_d , puisque les (X_i) sont indépendantes,

il vient

$$\begin{aligned} & \mathbb{P}(X_{d+1} \in \partial G(X_1, \dots, X_d)) \\ &= \int_{x_1 \in \mathbb{R}^d} \dots \int_{x_d \in \mathbb{R}^d} \mathbb{P}(X_{d+1} \in \partial G(x_1, \dots, x_d) \mid X_1 = x_1, \dots, X_d = x_d) dP(x_1) \dots dP(x_d) \\ &= \int_{x_1 \in \mathbb{R}^d} \dots \int_{x_d \in \mathbb{R}^d} \mathbb{P}(X_{d+1} \in \partial G(x_1, \dots, x_d)) dP(x_1) \dots dP(x_d). \end{aligned}$$

Par (\dot{H}_0^-) on sait que pour tout $G \in \mathcal{G}$ on a $\mathbb{P}(X_{d+1} \in \partial G) = 0$, ce qui implique

$$\mathbb{P}(X_{d+1} \in \partial G(x_1, \dots, x_d)) dP(x_1) \dots dP(x_d) = 0$$

donc

$$\mathbb{P}(X_{d+1} \in \partial G(X_1, \dots, X_d)) = 0.$$

Ensuite, on a

$$\begin{aligned} & \mathbb{P}(\{\exists i_1, \dots, i_{d+1} \text{ deux à deux distincts : } X_{i_{d+1}} \in \partial G(X_{i_1}, \dots, X_{i_d})\}) \\ & \leq \sum_{0 \leq i_1, \dots, i_{d+1} \leq n} \mathbb{P}(X_{i_{d+1}} \in \partial G(X_{i_1}, \dots, X_{i_d})) \\ & = \frac{n!}{(n - (d + 1))!} \mathbb{P}(X_{d+1} \in \partial G(X_1, \dots, X_d)) \\ & = 0. \end{aligned}$$

En d'autres termes, presque sûrement, aucune hypersurface de dimension $d - 1$ ne contient plus de d points de l'échantillon, ce qui s'écrit aussi

$$\mathbb{P}\left(\sup_{G \in \mathcal{G}} P_n(\partial G) \geq \frac{d+1}{n}\right) = 0.$$

avec les notations ∂G et $\text{int}G$ respectivement pour la frontière topologique de G et l'intérieur topologique de G .

on sait que

$$P_n(\text{int}\dot{G}_{\varphi,n}(O, u, \alpha)) \leq \alpha \text{ p.s.}$$

car si

$$P_n(\text{int}\dot{G}_{\varphi,n}(O, u, \alpha)) > \alpha \text{ p.s.}$$

alors il y a au moins $\lceil n\alpha \rceil$ points $X_i \in \text{int}\dot{G}_{\varphi,n}(O, u, \alpha)$ or, par définition de $\dot{G}_{\varphi,n}(O, u, \alpha)$ pour tout $X_i \in \text{int}\dot{G}_{\varphi,n}(O, u, \alpha)$ on a

$$\pi_{\varphi,O}(X_i, u) < \dot{Y}_{\varphi,n}(O, u, \alpha)$$

et en notant

$$\mathring{Y}_{\varphi,n}(O, u, \alpha) := \max_{X_i \in \text{int}\dot{G}_{\varphi,n}(O, u, \alpha)} (\pi_{\varphi, O}(X_i, u))$$

il vient

$$P_n(\dot{G}_{\varphi,n}(O, u, \mathring{Y}_{\varphi,n}(O, u, \alpha))) \geq \frac{\lceil n\alpha \rceil}{n} \geq \alpha$$

ce qui contredit la définition

$$\mathring{Y}_{\varphi,n}(O, u, \alpha) = \inf \{y \in \mathbb{R} : P_n(\dot{G}_{\varphi,n}(O, u, y)) \geq \alpha\}$$

Par conséquent, avec probabilité un, pour tout u, α

$$\begin{aligned} P_n(\dot{G}_{\varphi,n}(O, u, \alpha)) &= P_n(\text{int}\dot{G}_{\varphi,n}(O, u, \alpha)) + P_n(\partial\dot{G}_{\varphi,n}(O, u, \alpha)) \\ &\leq P_n(\text{int}\dot{G}_{\varphi,n}(O, u, \alpha)) + \frac{d}{n} \\ &\leq \alpha + \frac{d}{n}. \end{aligned}$$

□

Une autre conséquence importante des définitions 3.12 et 3.14 est la suivante.

Proposition 3.16. *Soit $O \in \mathbb{R}^d$ fixé, on a*

$$\sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\mathring{Y}_{\varphi}(O, u, \alpha)| < +\infty.$$

Démonstration. Puisque P est tendue alors il existe $r^+ > 0$ tel que

$$P(B(O, r^+)) > \alpha^+$$

où $B(O, r^+)$ désigne la boule de centre O et de rayon r^+ .

Remarquons que $\dot{G}_{\varphi}(O, u, \alpha^+) \cap B(O, r^+) \neq \emptyset$ car sinon on aurait $P(\dot{G}_{\varphi}(O, u, \alpha^+) \cup B(O, r^+)) = P(\dot{G}_{\varphi}(O, u, \alpha^+)) + P(B(O, r^+)) > 2\alpha^+ \geq 1$, ce qui est impossible. Ensuite puisque $\varphi^{-1}(\mathbb{R}) = \mathbb{R}^{d-1}$ alors il existe $y_0 \in \mathbb{R}$ tel que pour tout $u \in \mathbb{S}_{d-1}$ on a $B(O, r^+) \subset \dot{G}_{\varphi}(O, u, y_0)$, donc pour tout $u \in \mathbb{S}_{d-1}$ on a

$$|\mathring{Y}_{\varphi}(O, u, \alpha^+)| \leq y_0$$

par conséquent pour tout $O \in \mathbb{R}^d$ on a

$$\sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\mathring{Y}_{\varphi}(O, u, \alpha)| \leq y_0 < +\infty.$$

□

Notation 3.17. On pose

$$\mathcal{Y}_\Delta(O, u) = \{y : P(\dot{G}_\varphi(O, u, y)) \in \Delta\} = F_{\pi_{\varphi, O}(X, u)}^{-1}(\Delta)$$

l'ensemble des distances quantiles autorisées vues de O dans chaque direction. La proposition 3.16 implique

$$\sup_{u \in \mathbb{S}_{d-1}} \sup_{y \in \mathcal{Y}_\Delta(O, u)} |y| < \infty.$$

Définition 3.18 (L'ensemble des φ -rubans de largeur ε). Soit $\varepsilon > 0$. On définit l'ensemble des φ -Ruban de largeur $\varepsilon > 0$ par

$$\dot{\mathcal{R}}_{\varphi, O, \varepsilon, \Delta} = \{\dot{R}_\varphi(O, u, y, y + \varepsilon) : u \in \mathbb{S}_{d-1}, y, y + \varepsilon \in \mathcal{Y}_\Delta(O, u)\}.$$

Afin d'alléger les notation, on note ici aussi, $\dot{\mathcal{R}}_{\varphi, O, \varepsilon, \Delta} := \dot{\mathcal{R}}_{\varepsilon, \Delta}$.

3.2.2.b Régularité de la φ -surface quantile

Proposition 3.19. Sous (\dot{H}_0^-) et (\dot{H}_0^+) , l'application $(u, \alpha) \rightarrow \dot{Q}_\varphi(O, u, \alpha)$ est bornée et continue sur $\mathbb{S}_{d-1} \times \Delta$.

Démonstration. On fixe $O \in \mathbb{R}^d$, et on pose $u_\infty \in \mathbb{S}_{d-1}$ et $\alpha_\infty \in \Delta$. Comme $\dot{Q}_\varphi(O, u, \alpha) = O + \dot{Y}_\varphi(O, u, \alpha)u$, il suffit de montrer que $\dot{Y}_\varphi(O, \cdot, \cdot)$ est continue sur le compact $\mathbb{S}_{d-1} \times \Delta$.

Preuve par l'absurde : soit $u_n \rightarrow u_\infty$, $\alpha_n \rightarrow \alpha_\infty$ on suppose que $\lim_{n \rightarrow \infty} \dot{Y}_\varphi(O, u_n, \alpha_n) =: Y_\infty > y_\infty := \dot{Y}_\varphi(O, u_\infty, \alpha_\infty)$ donc la limite des $\dot{G}_\varphi(O, u_n, \alpha_n)$ est distincte de $\dot{G}_\varphi(O, u_\infty, \alpha_\infty)$.

On veut montrer que le ruban entre $\dot{G}_\varphi(O, u_\infty, y_\infty)$ et $\dot{G}_\varphi(O, u_\infty, Y_\infty)$ (en gris dans la figure 3.2) est nécessairement de probabilité nulle, ce qui est interdit par la condition (\dot{H}_0^+) . Le même raisonnement vaut pour le cas $Y_\infty < y_\infty$, ce qui montre que $Y_\infty = y_\infty$.

Étape I : on considère une suite de boules $B_k = b(Q_\infty, r_k) \rightarrow \{Q_\infty\}$ de centre fixe $Q_\infty = O + Y_\infty u_\infty$ et de rayon décroissant $r_k = 2^{-k}$, et pour tout k il existe $n_k > 0$ tel que pour tout $n \geq n_k$ le point sommet $Q_n = O + \dot{Y}_\varphi(O, u_n, \alpha_n)u_n$ (le point d'intersection de la frontière de la forme $\dot{G}_\varphi(O, u_n, \alpha_n)$ avec l'axe (O, u_n)) soit dans B_k . Ainsi n_k est tel que

$$\{Q_n : n \geq n_k\} \subset B_k.$$

Étape II : l'ensemble des formes $\dot{G}_\varphi(O, u, y)$ tel que leurs sommets $Q(O, u, y) = O + \dot{Y}_\varphi(O, u, y)u \in B_k$ est défini par

$$G_k = \{\dot{G}_\varphi(O, u, y) : Q(O, u, y) \in B_k\}.$$

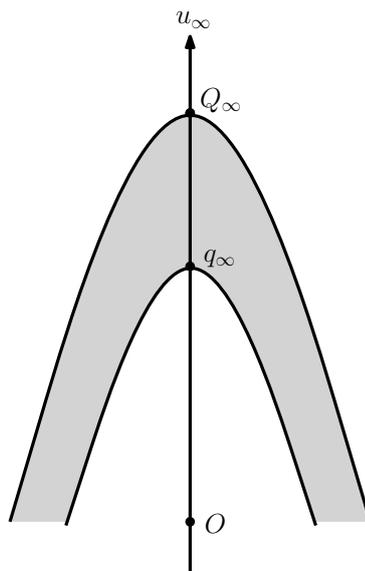


FIGURE 3.2 – $Q_\infty = O + Y_\infty u_\infty$ et $q_\infty = O + y_\infty u_\infty$

On a $\dot{G}_\varphi(O, u_n, \alpha_n) \in G_k$ pour tout $n \geq n_k$, de même, La forme $\dot{G}_\varphi(O, u, Y_\infty)$ de sommet Q_∞ est aussi dans G_k alors que par hypothèse, la vraie forme $\dot{G}_\varphi(O, u_\infty, \alpha_\infty) = \dot{G}_\varphi(O, u_\infty, y_\infty)$ n'est pas dans G_k .

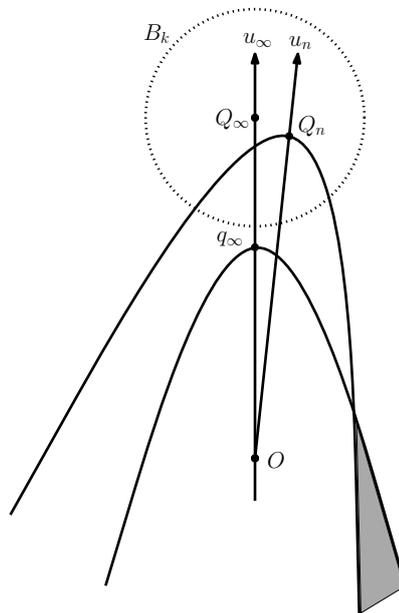


FIGURE 3.3 – En gris $\dot{G}_\varphi(O, u_\infty, \alpha_\infty) \setminus \dot{G}_\varphi(O, u_n, \alpha_n)$

Étape III : on s'intéresse d'abord aux intersections des éléments de G_k avec $\dot{G}_\varphi(O, u_\infty, \alpha_\infty) = \dot{G}_\varphi(O, u_\infty, y_\infty) \notin G_k$ la vraie forme en $(u_\infty, \alpha_\infty)$ qui est coiffée par la suite $\dot{G}_\varphi(O, u_n, \alpha_n)$. On a, pour tout $n \geq n_k$

$$\dot{G}_\varphi(O, u_\infty, \alpha_\infty) \setminus \dot{G}_\varphi(O, u_n, \alpha_n) \subset F_k^- := \bigcup_{\dot{G}_\varphi(O, u, y) \in B_k} \dot{G}_\varphi(O, u_\infty, \alpha_\infty) \setminus \dot{G}_\varphi(O, u, y)$$

avec

$$B_{k+1} \subset B_k, \quad F_{k+1}^- \subset F_k^-, \quad F_k^- \searrow \emptyset$$

donc

$$\lim_{k \rightarrow \infty} P(F_k^-) = P(\emptyset) = 0.$$

Étape IV : par définition on a $\dot{G}_\varphi(O, u_\infty, \alpha_\infty) = \dot{G}_\varphi(O, u_\infty, y_\infty)$,

$$\begin{aligned} \alpha_{n_k} &= P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k})) \\ &\geq P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k}) \cap \dot{G}_\varphi(O, u_\infty, \alpha_\infty)) \\ &= P(\dot{G}_\varphi(O, u_\infty, \alpha_\infty)) - P(\dot{G}_\varphi(O, u_\infty, y_\infty) \setminus \dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k})) \\ &\geq P(\dot{G}_\varphi(O, u_\infty, \alpha_\infty)) - P(F_k^-) \end{aligned}$$

on a $\lim_{k \rightarrow \infty} [P(\dot{G}_\varphi(O, u_\infty, \alpha_\infty)) - P(F_k^-)] = \alpha_\infty$, donc

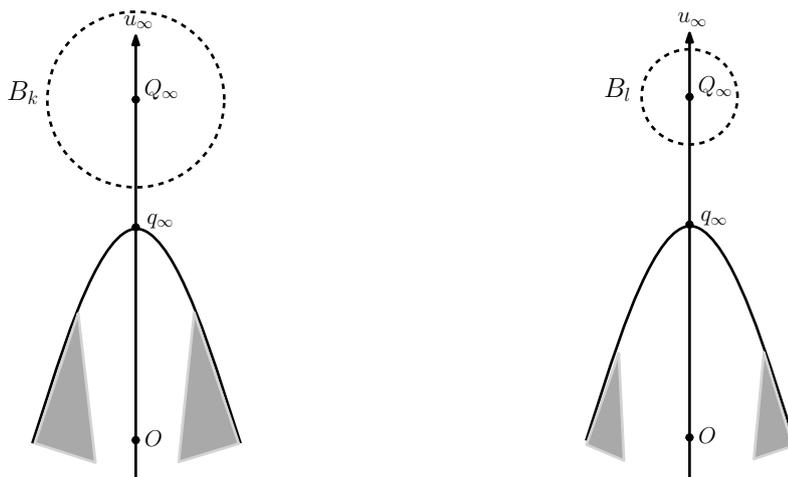
$$\lim_{k \rightarrow \infty} P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k}) \cap \dot{G}_\varphi(O, u_\infty, \alpha_\infty)) = \alpha_\infty$$

et on en déduit

$$\begin{aligned} &\lim_{k \rightarrow \infty} P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k}) \setminus \dot{G}_\varphi(O, u_\infty, \alpha_\infty)) \\ &= \lim_{k \rightarrow \infty} [P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k})) - P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k}) \cap \dot{G}_\varphi(O, u_\infty, \alpha_\infty))] \\ &= \lim_{k \rightarrow \infty} P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k})) - \lim_{k \rightarrow \infty} P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k}) \cap \dot{G}_\varphi(O, u_\infty, \alpha_\infty)) \\ &= \alpha_\infty - \alpha_\infty \\ &= 0 \end{aligned}$$

Étape V : tout est en place pour la contradiction. On considère $y'_\infty \in (y_\infty, Y_\infty)$ et observons que pour tout $n \geq n_k$

$$\dot{G}_\varphi(O, u_\infty, y'_\infty) \setminus G_\varphi(O, u_n, \alpha_n) \subset F_k^+ := \bigcup_{\dot{G}_\varphi(O, u, y) \in B_k} \dot{G}_\varphi(O, u_\infty, y'_\infty) \setminus \dot{G}_\varphi(O, u, y)$$

FIGURE 3.4 – Exemple de F_k^- et F_l^- pour $k < l$

avec de nouveau $F_{k+1}^+ \subset F_k^+, F_k^+ \searrow \emptyset$ donc

$$\lim_{k \rightarrow \infty} P(F_k^+) = P(\emptyset) = 0$$

et

$$\begin{aligned} \alpha_{n_k} &= P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k})) \\ &\geq P(\dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k}) \cap \dot{G}_\varphi(O, u_\infty, y'_\infty)) \\ &= P(\dot{G}_\varphi(O, u_\infty, y'_\infty)) - P(\dot{G}_\varphi(O, u_\infty, y'_\infty) \setminus \dot{G}_\varphi(O, u_{n_k}, \alpha_{n_k})) \\ &\geq P(\dot{G}_\varphi(O, u_\infty, y'_\infty)) - P(F_k^+) \\ &\geq P(\dot{G}_\varphi(O, u_\infty, y_\infty)) - P(F_k^+) \\ &= \alpha_\infty - P(F_k^+) \end{aligned}$$

par passage à la limite on obtient

$$\lim_{k \rightarrow \infty} P(\dot{G}_\varphi(O, u_\infty, y'_\infty)) = \alpha_\infty$$

donc

$$\begin{aligned} P(R_\varphi(O, u_\infty, y_\infty, y'_\infty)) &= P(\dot{G}_\varphi(O, u_\infty, y'_\infty)) - P(\dot{G}_\varphi(O, u_\infty, y_\infty)) \\ &= \alpha_\infty - \alpha_\infty \\ &= 0 \end{aligned}$$

ce qui contredit l'hypothèse (H_0^+) de rubans chargés par P . On peut facilement adapter le même raisonnement au cas où $Y_\infty < y_\infty$ en utilisant de nouveau $y'_\infty \in (Y_\infty, y_\infty)$.

Conclusion : Sous (\dot{H}_0^-) et (\dot{H}_0^+) , l'application $(u, \alpha) \rightarrow \dot{Q}_\varphi(O, u, \alpha)$ est bornée et continue sur $S_{d-1} \times \Delta$. \square

Il en découle immédiatement le corolaire suivant

Corollaire 3.20. *Sous (\dot{H}_0^-) et (\dot{H}_0^+) , l'ensemble $\dot{Q}_\varphi(O, \alpha)$ est une surface fermée, pour tout $\alpha \in \Delta$.*

Afin de donner une réécriture des conditions (\dot{H}_0^-) et (\dot{H}_0^+) sous une forme fonctionnelle on introduit l'application $\dot{\Psi}$.

Définition 3.21. *Soit $\varepsilon > 0$ on définit l'application $\dot{\Psi}$ en fonction de la probabilité des rubans par*

$$\begin{aligned} \dot{\Psi} : \mathbb{R}^+ &\longrightarrow [0, 1] \\ \varepsilon &\longmapsto \inf_{R \in \dot{\mathcal{R}}_{\varepsilon, \Delta}} P(R). \end{aligned}$$

En fait l'application $\dot{\Psi} = \dot{\Psi}_{\varphi, O}$ puisqu'elle dépend de $\dot{\mathcal{R}}_{\varepsilon, \Delta} := \dot{\mathcal{R}}_{\varphi, O, \varepsilon, \Delta} = \{\dot{R}_\varphi(O, u, y, y + \varepsilon) : u \in S_{d-1}, y, y + \varepsilon \in \mathcal{Y}_\Delta(O, u)\}$ avec O et φ fixés.

Remarque 3.22. *Sous (\dot{H}_0^-) et (\dot{H}_0^+) les deux conditions suivantes sont vérifiées*

$$\begin{aligned} (\dot{H}_{0, \dot{\Psi}}^-) \quad &\lim_{\varepsilon \rightarrow 0} \dot{\Psi}(\varepsilon) = 0 \\ (\dot{H}_{0, \dot{\Psi}}^+) \quad &\dot{\Psi}(\varepsilon) > 0 \text{ pour tout } \varepsilon > 0. \end{aligned}$$

- En effet, (\dot{H}_0^-) implique $(\dot{H}_{0, \dot{\Psi}}^-)$ car pour $y = \dot{Y}_\varphi(O, u, \alpha)$ avec $\alpha \in \Delta$ par la continuité de la mesure on a

$$\lim_{\varepsilon \rightarrow 0} \dot{\Psi}(\varepsilon) \leq \lim_{\varepsilon \rightarrow 0} P(\dot{R}_\varphi(O, u, y, y + \varepsilon)) = P(\partial \dot{G}(O, u, \alpha)) = 0.$$

- Ensuite on veut monter que sous (\dot{H}_0^-) on a $(\dot{H}_0^+) \Rightarrow (\dot{H}_{0, \dot{\Psi}}^+)$. Pour ce faire, on considère l'hypothèse de l'absurde i.e. supposons que

$$(\dot{H}_0^-) \text{ et } (\dot{H}_0^+) \text{ et } \neg(\dot{H}_{0, \dot{\Psi}}^+) \text{ est vrai.}$$

Dans ce cas il existe $\varepsilon_0 > 0$ tel que $\dot{\Psi}(\varepsilon_0) = 0$, ainsi il existe deux suites $u_k \in S_{d-1}$ et $y_k \in \mathcal{Y}_\Delta(O, u_k)$ telles que

$$\lim_{k \rightarrow \infty} P(\dot{R}_\varphi(O, u_k, y_k, y_k + \varepsilon_0)) = 0.$$

Par compacité de $\mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O)$ et continuité de $u \mapsto \dot{Y}_\varphi(O, u, \alpha)$ on peut extraire (u'_k, y'_k) une sous suite convergente vers $(u_0, y_0) \in \mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O, u_0)$. On pose

$$R_0 = \dot{R}_\varphi(O, u_0, y_0, y_0 + \varepsilon_0)$$

et

$$R'_k = \dot{R}_\varphi(O, u'_k, y'_k, y'_k + \varepsilon_0).$$

on a

$$\lim_{k \rightarrow \infty} P(R'_k) \geq P(R_0) - P(\partial R_0)$$

car si $x \in R_0 \setminus \partial R_0$ alors il existe un voisinage \mathcal{V}_δ de $(u'_0, y'_0) \in \mathbb{S}_{d-1} \times \mathcal{Y}_\Delta(O, u'_0)$ tel que $x \in \cap_{(u,y) \in \mathcal{V}_\delta} \dot{R}_\varphi(O, u, y, y + \varepsilon_0)$, donc pour k assez grand, $x \in R'_k$. donc

$$P(R_0) \geq \lim_{k \rightarrow \infty} P(R'_k) \geq P(R_0) - P(\partial R_0)$$

or, sous (\dot{H}_0^-) et (\dot{H}_0^+) on a

$$P(R_0) > 0 \text{ et } P(\partial R_0) = 0$$

donc

$$\lim_{k \rightarrow \infty} P(R'_k) > 0$$

ce qui est une contradiction.

Proposition 3.23. *La fonction $\varepsilon \mapsto \dot{\Psi}(\varepsilon)$ est continue à droite en tout $\varepsilon > 0$. Si de plus (\dot{H}_0^-) et (\dot{H}_0^+) sont vérifiées alors $\dot{\Psi}(\varepsilon)$ est continue en tout $\varepsilon \geq 0$.*

Démonstration. L'application $\dot{\Psi}$ est monotone et admet des limites à droite en tout $\varepsilon_0 \geq 0$ et à gauche en tout $\varepsilon_0 > 0$.

Pour tout $\theta > 0$ il existe $R_{\theta,0} \in \dot{\mathcal{R}}_{\varepsilon_0, \Delta}$ tel que

$$(1 + \theta)\dot{\Psi}(\varepsilon_0) > P(R_{\theta,0}) \geq \dot{\Psi}(\varepsilon_0)$$

$$R_{\theta,0} = \dot{R}_\varphi(O, u_\theta, y_\theta, y_\theta + \varepsilon_0)$$

Soit $\varepsilon_k \downarrow \varepsilon_0 > 0$, on considère la suite décroissante d'ensembles

$$R_{\theta,k} = \dot{R}_\varphi(O, u_\theta, y_\theta, y_\theta + \varepsilon_k)$$

qui sont de limite $\cap_k R_{\theta,k} = R_{\theta,0}$ et par continuité de la mesure, vérifient $P(R_{\theta,k}) \downarrow P(R_{\theta,0})$, ainsi, il existe $k_\theta > 0$ tel que pour tout $k \geq k_\theta$

$$(1 + \theta)\dot{\Psi}(\varepsilon_0) > P(R_{\theta,k}) \geq P(R_{\theta,0}) \geq \dot{\Psi}(\varepsilon_0)$$

donc également, par croissance de $\dot{\Psi}$,

$$(1 + \theta)\dot{\Psi}(\varepsilon_0) > P(R_{\theta,k}) \geq \dot{\Psi}(\varepsilon_k) \geq \dot{\Psi}(\varepsilon_0).$$

Comme $\dot{\Psi}(\varepsilon_k)$ converge vers la limite à droite $\dot{\Psi}(\varepsilon_0^+)$ de $\dot{\Psi}$ en ε_0 , on a en fait montré que pour tout $\theta > 0$,

$$(1 + \theta)\dot{\Psi}(\varepsilon_0) > \lim_{k \rightarrow \infty} \dot{\Psi}(\varepsilon_k) = \dot{\Psi}(\varepsilon_0^+) \geq \dot{\Psi}(\varepsilon_0)$$

c'est-à-dire

$$\lim_{k \rightarrow \infty} \dot{\Psi}(\varepsilon_k) = \dot{\Psi}(\varepsilon_0^+) = \dot{\Psi}(\varepsilon_0).$$

De même, si $\varepsilon_k \uparrow \varepsilon_0 > 0$ alors à tout $\theta > 0$ on associe une suite $R_{\theta,k} \in \dot{\mathcal{R}}_{\varepsilon_k, \Delta}$ telle que

$$(1 + \theta)\dot{\Psi}(\varepsilon_k) > P(R_{\theta,k}) > \dot{\Psi}(\varepsilon_k)$$

et on extrait, par compacité en u et y , une suite stabilisée

$$R_{\theta,k_n} = \dot{R}_\varphi(O, u_{k_n}, y_{k_n}, y_{k_n} + \varepsilon_{k_n})$$

avec $u_{k_n} \rightarrow u_\theta, y_{k_n} \rightarrow y_\theta$. On pose

$$R_{\theta,0} = \dot{R}_\varphi(O, u_\theta, y_\theta, y_\theta + \varepsilon_0).$$

Sous (H_0^-) et (H_0^+) par continuité des surfaces on a $P(R_{\theta,k_n}) \rightarrow P(R_{\theta,0})$ donc pour tout $k \geq k_\theta$ on a

$$\begin{aligned} \lim_{k \rightarrow \infty} P(R_{\theta,k}) &= \dot{\Psi}(\varepsilon_0^-) \geq (1 - \theta)P(R_{\theta,0}) \geq (1 - \theta)\dot{\Psi}(\varepsilon_0) \\ \dot{\Psi}(\varepsilon_0) &\geq \dot{\Psi}(\varepsilon_0^-) \geq (1 - \theta)\dot{\Psi}(\varepsilon_0) \end{aligned}$$

et cela pour tout $\theta > 0$. Par suite,

$$\dot{\Psi}(\varepsilon_0^-) = \dot{\Psi}(\varepsilon_0).$$

□

Remarque 3.24. On définit pour tout $t \in \Delta$ la fonction inverse généralisée de $\dot{\Psi}$ comme suit

$$\dot{\Psi}^{-1}(t) = \inf \{ \varepsilon \geq 0 : \dot{\Psi}(\varepsilon) \geq t \}$$

puisque $\dot{\Psi}$ est continue et croissante alors cette fonction est càdlàg et strictement croissante avec

$$\dot{\Psi} \circ \dot{\Psi}^{-1}(t) = t \text{ et } \dot{\Psi}^{-1} \circ \dot{\Psi}(t) \geq t.$$

3.2.3 Théorèmes de convergences

Dans cette partie on donne les hypothèses suffisantes pour garantir les convergences avec ou sans vitesse des théorèmes énoncés dans la partie suivante. Comme précédemment, on fixe O ainsi que la fonction φ vérifiant les conditions (3.2.1).

Sans condition supplémentaire sur la régularité de la fonction φ il est très difficile de garantir que $\dot{\mathcal{G}}$ (cf. définition 3.14) est une VC classe, et encore moins connaître sa dimension VC.

Hypothèses. Soit $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$

(\dot{H}_∞) $\dot{\mathcal{G}}$ est une VC classe de dimension q .

(\dot{H}_0^-) $P(\partial \dot{\mathcal{G}}_\varphi(O, u, \alpha)) = 0$.

(\dot{H}_0^+) $P(\dot{R}_\varphi(O, u, y, z)) > 0$ si $Y_\varphi(O, u, \alpha^+) \geq z > y \geq \dot{Y}_\varphi(O, u, \alpha^-)$.

(\dot{H}_1) La variable aléatoire réelle $\pi_{\varphi, O}(X, u)$ admet une densité $f_{\pi_{\varphi, O}(X, u)}$ continue, avec $f_{\pi_{\varphi, O}(X, u)} > 0$ sur $F_{\pi_{\varphi, O}(X, u)}^{-1}(\Delta)$. De plus, on suppose qu'il existe m et M tels que $0 < m < M < +\infty$ et

$$m \leq \inf_{\alpha \in \Delta} \inf_{u \in \mathbb{S}_{d-1}} h_O(u, \alpha) \leq \sup_{\alpha \in \Delta} \sup_{u \in \mathbb{S}_{d-1}} h_O(u, \alpha) \leq M$$

où $h_O(u, \alpha) = f_{\pi_{\varphi, O}(X, u)} \circ F_{\pi_{\varphi, O}(X, u)}^{-1}(\alpha)$.

Dans les preuves la quantité $h_O(u, \alpha)$ apparait naturellement dans la dérivée de $F_{\pi_{\varphi, O}(X, u)}^{-1}$ en α , et parfois directement sous la forme de $f_{\pi_{\varphi, O}(X, u)}(\dot{Y}_\varphi(O, u, \alpha))$ qui par définition de $\dot{Y}_\varphi(O, u, \alpha)$

$$h_O(u, \alpha) = f_{\pi_{\varphi, O}(X, u)}(\dot{Y}_\varphi(O, u, \alpha)).$$

Comme dans le chapitre précédent, en aucun cas dans (\dot{H}_1), on suppose que la loi P est absolument continue par rapport à la mesure de Lebesgue dans \mathbb{R}^d , néanmoins, la condition (\dot{H}_1) implique les deux conditions précédentes (\dot{H}_0^-) et (\dot{H}_0^+) avec en plus

$$m(z - y) \leq P(\dot{R}_\varphi(O, u, y, z)) \leq M(z - y), \quad \forall y, z \in \mathcal{Y}_\Delta(O, u) \text{ avec } y < z.$$

ce qui se traduit sur la fonction $\dot{\Psi}$ par

$$m\varepsilon \leq \dot{\Psi}(\varepsilon) \leq M\varepsilon \quad \forall \varepsilon \geq 0.$$

Afin de donner la condition (\dot{H}_2), pour tout $\gamma > 0$, on pose

$$\rho(\gamma) = \sup_{|\varepsilon'| < \gamma} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} \left| F_{\pi_{\varphi, O}(X, u)}(\dot{Y}_\varphi(O, u, \alpha) + \varepsilon') - \alpha - h_O(u, \alpha)\varepsilon' \right|.$$

(\dot{H}_2) On suppose que

$$\lim_{\gamma \rightarrow 0} \frac{\sqrt{\log \log(1/\gamma)}}{\gamma} \rho(\gamma) = 0.$$

les conditions (\dot{H}_1) et (\dot{H}_2) suffisent pour garantir le théorème limite central uniforme, l'approximation du type Bahadur-Kiefer sans vitesse, ainsi que le principe d'invariance fort sans vitesse de convergence pour le terme d'erreur de l'approximation.

Afin de pouvoir donner ces vitesses de convergence, on a besoin de l'une des conditions suivantes.

(\dot{H}_3) On suppose que

$$\lim_{\gamma \rightarrow 0} \frac{\rho(\gamma) \sqrt{\log \log(1/\gamma)}}{\gamma^{3/2} \sqrt{\log(1/\gamma)}} = 0.$$

La condition (\dot{H}_3) peut être remplacé par une condition (\dot{H}'_3) plus exigeante, dans le sens que (\dot{H}'_3) \Rightarrow (\dot{H}_3).

(\dot{H}'_3) Il existe $r > 1/2$ et $C^* > 0$ tel que pour tout $\gamma > 0$ on a

$$\rho(\gamma) \leq C^* \gamma^{1+r}.$$

Remarque. En effet, on (\dot{H}'_3) \Rightarrow (\dot{H}_3) car s'il existe $r > 1/2$ et $C^* > 0$ tel que pour tout $\gamma > 0$ on a

$$\rho(\gamma) \leq C^* \gamma^{1+r}$$

alors

$$\frac{\sqrt{\log \log(1/\gamma)}}{\gamma} \rho(\gamma) \leq C^* \gamma^r \sqrt{\log \log(1/\gamma)}$$

or $\lim_{r \rightarrow 0} \gamma^r \sqrt{\log \log(1/\gamma)} = 0$ donc

$$\lim_{\gamma \rightarrow 0} \frac{\rho(\gamma) \sqrt{\log \log(1/\gamma)}}{\gamma^{3/2} \sqrt{\log(1/\gamma)}} = 0$$

3.2.3.a Théorèmes de convergence

Dans toute la suite, on suppose que φ est fixée et vérifie les conditions (3.2.1). On commence par énoncer la loi des grands nombres uniforme.

Théorème 3.25 (Consistance Uniforme). *On fixe $O \in \mathbb{R}^d$, sous les conditions (\dot{H}_∞) , (\dot{H}_0^-) et (\dot{H}_0^+) , on a*

$$\limsup_{n \rightarrow \infty} \sup_{\alpha \in \Delta} \sup_{u \in \mathbb{S}_{d-1}} |\dot{Y}_{\varphi,n}(O, u, \alpha) - \dot{Y}_\varphi(O, u, \alpha)| = 0 \quad p.s.$$

Ensuite, voici la loi du logarithme itéré uniforme pour les distances φ -quantile.

Théorème 3.26 (Loi du logarithme itéré uniforme). *On fixe $O \in \mathbb{R}^d$, sous les conditions (\dot{H}_∞) , (\dot{H}_0^-) et (\dot{H}_0^+) , on a*

$$\limsup_{n \rightarrow \infty} \frac{\|\dot{Y}_\varphi - \dot{Y}_{\varphi,n}\|_{\mathbb{S}_{d-1} \times \Delta}}{\Psi_{-1} \left(\sqrt{\frac{\log \log n}{n}} \right)} < \infty \quad p.s.$$

Notation 3.27. *Comme dans les sections précédentes on fixe $O \in \mathbb{R}^d$ et φ vérifiant 3.2.1. Afin de faciliter l'écriture, on pose les notations suivantes :*

– Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $y \in \mathbb{R}$

$$\alpha_n(O, u, y) = \sqrt{n} (P_n(\dot{G}_\varphi(O, u, y)) - P(\dot{G}_\varphi(O, u, y)))$$

– Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$

$$\mathbb{E}_O^n(u, \alpha) = \sqrt{n} (P_n(\dot{G}_\varphi(O, u, \alpha)) - P(\dot{G}_\varphi(O, u, \alpha)))$$

– Pour $O \in \mathbb{R}^d$ et $G \in \dot{\mathcal{G}}$

$$\Lambda_n(G) = \sqrt{n} (P_n(G) - P(G))$$

– Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$

$$\mathbb{D}_O^n(u, \alpha) = \sqrt{n} (\dot{Y}_{\varphi,n}(O, u, \alpha) - \dot{Y}_\varphi(O, u, \alpha))$$

Remarque 3.28. *Avec les notations précédente, pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$*

$$\alpha_n(O, u, \dot{Y}_\varphi(O, u, \alpha)) = \mathbb{E}_O^n(u, \alpha) = \Lambda_n(\dot{G}_\varphi(O, u, \alpha))$$

Théorème 3.29 (Théorème limite central uniforme). *Soit $O \in \mathbb{R}^d$ fixé, sous les conditions (\dot{H}_∞) et (\dot{H}_1) , sur l'ensemble des fonctions bornées sur $S_{d-1} \times \Delta$ muni de la norme uniforme, le processus \mathbb{D}_O^n converge faiblement vers le processus gaussien \mathbb{B}_O , avec*

$$\mathbb{B}_O(u, \alpha) = \frac{\mathbb{G}(\dot{G}_\varphi(O, u, \alpha))}{h_O(u, \alpha)}$$

où \mathbb{G} est le pont brownien standard indexé par $\dot{\mathcal{G}}$, avec $\mathbb{E}(\mathbb{B}_O(u, \alpha)) = 0$ et

$$\text{cov}(\mathbb{B}_O(u, \alpha), \mathbb{B}_O(u', \alpha')) = \frac{P(\dot{G}_\varphi(O, u, \alpha) \cap \dot{G}_\varphi(O, u', \alpha')) - \alpha\alpha'}{h_O(u, \alpha)h_O(u', \alpha')}.$$

On note q la dimension VC de la classe $\dot{\mathcal{G}}$ et q' la dimension VC de la classe $\mathcal{R}_{\varepsilon, \Delta}$ des rubans admissibles.

Théorème 3.30 (Principe d'invariance fort uniforme avec vitesse). *Soit $O \in \mathbb{R}^d$ fixé, sous les conditions (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite \mathbb{B}_O^n de versions de \mathbb{B}_O de sorte que*

$$\dot{Y}_{\varphi, n}(O, u, \alpha) = \dot{Y}_\varphi(O, u, \alpha) + \frac{\mathbb{B}_O^n(u, \alpha)}{\sqrt{n}} + \frac{\mathbb{Z}_O^n(u, \alpha)}{\sqrt{n}}, \quad O \in \mathbb{R}^d \quad \alpha \in \Delta, \quad u \in S_{d-1}, \quad (3.1)$$

où $\mathbb{Z}_O^n = \mathbb{D}_O^n - \mathbb{B}_O^n$ vérifie

$$\lim_{n \rightarrow \infty} \|\mathbb{Z}_O^n\|_{S_{d-1} \times \Delta} = \lim_{n \rightarrow \infty} \sup_{u \in S_{d-1}} \sup_{\alpha \in \Delta} |\mathbb{Z}_O^n(u, \alpha)| = 0 \quad p.s.$$

si de plus P satisfait la condition (\dot{H}_3) alors on peut construire \mathbb{B}_O^n tel que pour $v_q = 1/(10q - 8)$ et $w_q = (10q - 6)/(20q - 16)$, il existe $n_\omega > 0$ tel que pour tout $n > n_\omega$,

$$\|\mathbb{Z}_O^n\|_{S_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_q}}{n^{v_q}} \right) \quad (3.2)$$

où q est la dimension VC de $\dot{\mathcal{G}}$.

Théorème 3.31 (Représentation de type Bahadur-Kiefer avec vitesse). *Soit $O \in \mathbb{R}^d$ fixé, sous les conditions (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) , on a*

$$\lim_{n \rightarrow \infty} \|\mathbb{D}_O^n h_O + \mathbb{E}_O^n\|_{S_{d-1} \times \Delta} = 0 \quad p.s$$

si de plus P satisfait (\dot{H}_3) ou (\dot{H}'_3) alors il existe $n_\omega > 0$ tel que pour tout $n > n_\omega$,

$$\|\mathbb{D}_O^n h_O + \mathbb{E}_O^n\|_{S_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right)$$

3.2.4 Preuves des théorèmes

Dans cette section on donne les preuves des théorèmes portant sur les quantiles via des formes fonctionnelles.

Dans la suite, on note q la dimension VC de la classe $\dot{\mathcal{G}}$.

3.2.4.a Preuve de la consistance

On fixe $O \in \mathbb{R}^d$, pour $\varepsilon > 0$ on pose $\dot{\mathcal{R}}_{\varepsilon, \Delta}$ comme dans définition 3.18

$$\dot{\mathcal{R}}_{\varepsilon, \Delta} = \{\dot{R}_\varphi(O, u, y, y + \varepsilon) : u \in \mathbb{S}_{d-1}, y, y + \varepsilon \in \mathcal{Y}_\Delta(O, u)\}$$

Ainsi que

$$\Psi(\varepsilon) = \inf_{R \in \dot{\mathcal{R}}_{\varepsilon, \Delta}} P(R).$$

Sous (\dot{H}_0^-) et $(\dot{H}_{0, \Psi}^+)$ supposons qu'il existe $\delta > 0$ et une suite $n_k \rightarrow \infty$ tels que

$$|\dot{Y}_{\varphi, n_k}(O, u_{n_k}, \alpha_{n_k}) - \dot{Y}_\varphi(O, u_{n_k}, \alpha_{n_k})| > \delta$$

ou bien on peut en extraire une sous suite convergente $(m_k)_{k \geq 1}$ avec $\lim_{k \rightarrow \infty} m_k = +\infty$ tel que $\dot{Y}_\varphi(O, u_{m_k}, \alpha_{m_k}) - \dot{Y}_{\varphi, m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$ sinon $\dot{Y}_{\varphi, m_k}(O, u_{m_k}, \alpha_{m_k}) - \dot{Y}_\varphi(O, u_{m_k}, \alpha_{m_k}) \geq \delta$.

Dans le premier cas, i.e. si on peut en extraire une sous suite convergente $(m_k)_{k \geq 1}$ avec $\lim_{k \rightarrow \infty} m_k = +\infty$ tel que $\dot{Y}_\varphi(O, u_{m_k}, \alpha_{m_k}) - \dot{Y}_{\varphi, m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$, on pose

$$A_k = \dot{G}_{\varphi, m_k}(O, u_{m_k}, \alpha_{m_k})$$

$$C_k = \dot{G}_\varphi(O, u_{m_k}, \alpha_{m_k})$$

$$B_k = C_k \setminus A_k$$

puisque par la condition (\dot{H}_∞) on a $\dot{\mathcal{G}} = \{\dot{G}(O, u, y) : u \in \mathbb{S}_{d-1}, y \in \mathcal{Y}_\Delta(O, u)\}$ est une VC-classe alors

$$\lim_{n \rightarrow \infty} \sup_{G \in \dot{\mathcal{G}}} |P_n(G) - P(G)| = 0,$$

ainsi par (\dot{H}_0^-) ,

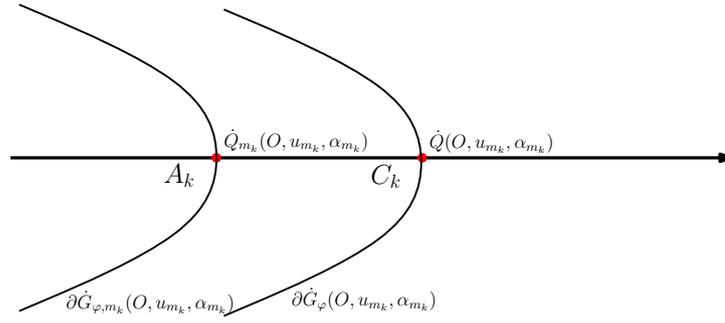


FIGURE 3.5 – Cas 1 : $\dot{Y}_\varphi(O, u_{m_k}, \alpha_{m_k}) - \dot{Y}_{\varphi, m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta$

$$\begin{aligned}
 \sup_{G \in \mathcal{G}} |P_{m_k}(G) - P(G)| &\geq P_{m_k}(A_k) - P(A_k) \\
 &\geq \alpha_{m_k} - \frac{d}{m_k} - (\alpha_{m_k} - P(B_k)) \\
 &\geq -\frac{d}{m_k} + P(B_k) \\
 &\geq -\frac{d}{m_k} + \Psi(\delta)
 \end{aligned}$$

donc

$$\Psi(\delta) \leq \sup_{G \in \mathcal{G}} |P_{m_k}(G) - P(G)| + \frac{d}{m_k} \quad (3.3)$$

lorsque k tends vers l'infinie, ceci implique qu'il existe $\delta > 0$ tel que

$$\Psi(\delta) = 0$$

or ceci contredit l'hypothèse $(\dot{H}_{0, \Psi}^+)$.

Dans le deuxième cas, la preuve est en tout point analogue.

Conclusion : Sous les conditions (\dot{H}_0^-) et (\dot{H}_0^+) on a pour tout $O \in \mathbb{R}^d$:

$$\lim_{n \rightarrow \infty} \sup_{\alpha \in \Delta_0} \sup_{u \in \mathbb{S}_{d-1}} |\dot{Y}_{\varphi, n}(O, u, \alpha) - \dot{Y}_\varphi(O, u, \alpha)| = 0 \quad p.s.$$

3.2.4.b Preuve de la loi du logarithme itéré

Démonstration. Preuve par l'absurde : On fixe $O \in \mathbb{R}^d$, sous les conditions $(\dot{H}_{0,\Psi}^-)$ et $(\dot{H}_{0,\Psi}^+)$, supposons qu'il existe une suite aléatoire croissante $n_k \rightarrow \infty$ tels que

$$|\dot{Y}_{\varphi,n_k}(O, u_{n_k}, \alpha_{n_k}) - \dot{Y}_{\varphi}(O, u_{n_k}, \alpha_{n_k})| > \delta_{n_k} = \Psi^{-1} \left(\sqrt{\frac{\log \log n_k}{n_k}} \right)$$

Deux cas sont possibles : ou bien on peut en extraire une sous suite convergente $(m_k)_{k \geq 1}$ avec $\lim_{k \rightarrow \infty} m_k = +\infty$ tel que $\dot{Y}_{\varphi}(O, u_{m_k}, \alpha_{m_k}) - \dot{Y}_{\varphi,m_k}(O, u_{m_k}, \alpha_{m_k}) \geq \delta_{m_k}$ sinon $\dot{Y}_{\varphi,m_k}(O, u_{m_k}, \alpha_{m_k}) - \dot{Y}_{\varphi}(O, u_{m_k}, \alpha_{m_k}) \geq \delta_{m_k}$.

Dans le premier cas on pose $\dot{Y}_{\varphi}(O, u_{m_k}, \alpha_{m_k}) \geq \dot{Y}_{\varphi,m_k}(O, u_{m_k}, \alpha_{m_k})$, on pose

$$A_k = \dot{G}_{\varphi,m_k}(O, u_{m_k}, \alpha_{m_k})$$

$$C_k = \dot{G}_{\varphi}(O, u_{m_k}, \alpha_{m_k})$$

$$B_k = C_k \setminus A_k.$$

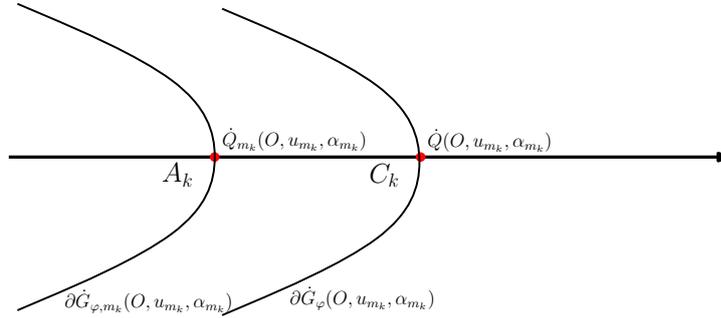


FIGURE 3.6 – Cas 1 : $\dot{Y}_{\varphi}(O, u_{m_k}, \alpha_{m_k}) \geq \dot{Y}_{\varphi,m_k}(O, u_{m_k}, \alpha_{m_k})$

Par la condition (\dot{H}_{∞}) la classe $\dot{\mathcal{G}}$ est une VC-classe, donc on a

$$\limsup_{n \rightarrow \infty} \sup_{G \in \dot{\mathcal{G}}} |P_n(G) - P(G)| = 0,$$

et par la loi du logarithme itéré (Alexander 1984 [Ale84]) on a

$$\limsup_{n \rightarrow \infty} \frac{\|P_n - P\|_{\dot{\mathcal{G}}}}{\sqrt{\frac{\log \log n}{n}}} \leq \sqrt{2}/2 \text{ p.s.}$$

comme $4/5 \geq \sqrt{2}/2$, il existe $k(\omega) > 0$ tel que pour tout $k \geq k(\omega)$

$$\frac{4}{5} \sqrt{\frac{\log \log m_k}{m_k}} \geq \sup_{G \in \mathcal{G}} |P_{m_k}(G) - P(G)|.$$

Enfin, puisque (H_0^-) est vérifiée alors par la Proposition 3.15 et la continuité de Ψ de la proposition 3.23, on a

$$\begin{aligned} \sup_{G \in \mathcal{G}} |P_n(G) - P(G)| &\geq P_{m_k}(A_k) - P(A_k) \\ &\geq \alpha_{m_k} - \frac{d}{m_k} - (\alpha_{m_k} - P(B_k)) \\ &\geq -\frac{d}{m_k} + P(B_k) \\ &\geq -\frac{d}{m_k} + \Psi(\delta_n) \\ &\geq -\frac{d}{m_k} + \Psi\left(\Psi^{-1}\left(\sqrt{\frac{\log \log m_k}{m_k}}\right)\right) \\ &\geq \sqrt{\frac{\log \log m_k}{m_k}} - \frac{d}{m_k} \end{aligned}$$

donc

$$\frac{4}{5} \sqrt{\frac{\log \log m_k}{m_k}} \geq \sup_{G \in \mathcal{G}} |P_n(G) - P(G)| \geq \sqrt{\frac{\log \log m_k}{m_k}} - \frac{d}{m_k} \quad (3.4)$$

ce qui implique que pour $k \geq k(\omega)$

$$\frac{4}{5} \sqrt{\frac{\log \log m_k}{m_k}} \geq \sqrt{\frac{\log \log m_k}{m_k}} - \frac{d}{m_k}$$

donc

$$1 \geq \frac{1}{5d} \sqrt{m_k \log \log m_k}$$

ce qui est absurde, donc

$$\limsup_{n \rightarrow \infty} \frac{\|\dot{Y}_\varphi - \dot{Y}_{\varphi,n}\|_{S_{d-1} \times \Delta}}{\Psi^{-1}\left(\sqrt{\frac{\log \log n}{n}}\right)} < \infty$$

Dans le deuxième cas, la preuve est identique. □

3.2.4.c Outils pour les preuve de convergence faible et approximation forte

On fixe $O \in \mathbb{R}^d$ et soit $\varepsilon > 0$ nous rappelons que

$$\dot{\mathcal{R}}_{\varepsilon, \Delta} = \{\dot{R}_\varphi(O, u, y, y + \varepsilon) : u \in \mathbb{S}_{d-1}, y \in \mathcal{Y}_\Delta(O, u)\}$$

Lemme 3.32. *Soit O fixé dans \mathbb{R}^d . Pour tout $\varepsilon > 0$, $u \in \mathbb{S}_{d-1}$ et $y, y + \varepsilon \in \mathcal{Y}(O, \Delta)$ on a*

$$\Lambda_n(\dot{R}_\varphi(O, u, y, y + \varepsilon)) = \Lambda_n(\dot{G}_\varphi(O, u, y + \varepsilon)) - \Lambda_n(\dot{G}_\varphi(O, u, y)).$$

Démonstration. pour $\dot{R}_\varphi(O, u, y, y + \varepsilon) \in \dot{\mathcal{R}}_{\varepsilon, \Delta}$ on a

$$\begin{aligned} \Lambda_n(\dot{R}_\varphi(O, u, y, y + \varepsilon)) &= \sqrt{n} (P_n(\dot{R}_\varphi(O, u, y, y + \varepsilon)) - P(\dot{R}_\varphi(O, u, y, y + \varepsilon))) \\ &= \sqrt{n} (P_n(\dot{G}_\varphi(O, u, y + \varepsilon) \setminus \dot{G}_\varphi(O, u, y)) \\ &\quad - P(\dot{G}_\varphi(O, u, y + \varepsilon) \setminus \dot{G}_\varphi(O, u, y))) \\ &= \sqrt{n} (P_n(\dot{G}_\varphi(O, u, y + \varepsilon)) - P_n(\dot{G}_\varphi(O, u, y)) \\ &\quad - P(\dot{G}_\varphi(O, u, y + \varepsilon)) + P(\dot{G}_\varphi(O, u, y))) \\ &= \Lambda_n(\dot{G}_\varphi(O, u, y + \varepsilon)) - \Lambda_n(\dot{G}_\varphi(O, u, y)). \end{aligned}$$

□

Soit $\varepsilon_n = C \sqrt{\frac{\log \log n}{n}}$ pour $C > 1$, on pose

$$\dot{\mathcal{R}}_n = \bigcup_{\varepsilon < \varepsilon_n} \dot{\mathcal{R}}_{\varepsilon, \Delta}$$

et la classe de fonction

$$\mathcal{I}_n = \{\mathbb{1}_R, R \in \mathcal{R}_n\}$$

Lemme 3.33. *On fixe $n \in \mathbb{N}$, pour toute fonction $f \in \mathcal{I}_n$ on a*

- i. *il existe $S_* > 0$ tel que $\sup_{x \in \mathbb{R}^d} |f(x)| \leq S_*/2$*
- ii. *il existe une sous classe dénombrable \mathcal{I}_∞ de \mathcal{I}_n telle que pour toute fonction $f \in \mathcal{I}_n$ il existe une suite de fonction (f_k) dans \mathcal{I}_∞ telle que pour tout $x \in \mathbb{R}^d$ on a $\lim_{k \rightarrow \infty} f_k(x) = f(x)$*

Démonstration. On fixe $n \in \mathbb{N}$,

- i. Pour toute fonction $f \in \mathcal{I}_n$ il existe $R \in \mathcal{R}_n$ tel que $f = \mathbb{1}_R$ donc

$$\sup_{x \in \mathbb{R}^d} |f(x)| = \sup_{x \in \mathbb{R}^d} \mathbb{1}_R(x) \leq 1$$

par conséquent pour $S_* = 2$ on a $\sup_{x \in \mathbb{R}^d} |f(x)| \leq S_*/2$

- ii. Pour tout $R \in \mathcal{R}_n$ il existe $\varepsilon \leq \varepsilon_n$, $u \in \mathbb{S}_{d-1}$ et $y \in \mathcal{Y}(O, \Delta)$ tel que tels que le rubans $R = \dot{R}_\varphi(O, u, y, y + \varepsilon)$.

Soit (δ_k) une suite de rationnelles avec $\delta_k \rightarrow \varepsilon$, (u_k) une suite de $\mathbb{Q}_{d-1} = \{v \in \mathbb{Q}^2 : \|v\|_2 = 1\}$ et (y_k) une suite de rationnelle avec $y_k \rightarrow y$, par définition $\dot{R}_\varphi(O, u, y, y + \varepsilon) = \dot{G}_\varphi(O, u, y + \varepsilon) \setminus \dot{G}_\varphi(O, u, y)$ et $\dot{G}_\varphi(O, u, y) \subset \dot{G}_\varphi(O, u, y + \varepsilon)$ donc $\mathbb{1}_{\dot{R}_\varphi(O, u, y, y + \varepsilon)} = \mathbb{1}_{\dot{G}_\varphi(O, u, y + \varepsilon)} - \mathbb{1}_{\dot{G}_\varphi(O, u, y)}$ ensuite par continuité de $\pi_{\varphi, O}$ on a pour tout $x \in \mathbb{R}^d$

$$\begin{aligned} \lim_{k \rightarrow \infty} \mathbb{1}_{\dot{R}_\varphi(O, u_k, y_k, y_k + \varepsilon_k)}(x) &= \lim_{k \rightarrow \infty} (\mathbb{1}_{\dot{G}_\varphi(O, u_k, y_k + \varepsilon_k)}(x) \times \mathbb{1}_{\dot{G}_\varphi(O, u_k, y_k)^c}(x)) \\ &= \mathbb{1}_{\dot{G}_\varphi(O, u, y + \varepsilon)}(x) \times \mathbb{1}_{\dot{G}_\varphi(O, u, y)^c}(x) \\ &= \mathbb{1}_{\dot{R}_\varphi(O, u, y, y + \varepsilon)}(x). \end{aligned}$$

□

Remarque 3.34. Le lemme 3.33 permet de montrer que la classe de fonctions \mathcal{I}_n associée à la classe d'ensembles \mathcal{R}_n vérifie les deux condition (F.i) et (F.ii) de l'inégalité de Talagrand 2.51, l'inégalité des moments 2.53 et le théorème 2.55 de Berthet-Mason.

Dans le lemme suivant, on montre que sous (\dot{H}_∞) , i.e. $\dot{\mathcal{G}}$ est une classe VC de dimension q , on vérifie bien les conditions de l'inégalité des moments 2.53.

Lemme 3.35. Sous (\dot{H}_∞) et (\dot{H}_1) , Il existe $n_0 \geq 3$ telle que pour tout $n \geq n_0$, il existe G une fonction enveloppe de \mathcal{I}_n et $\beta, v, c > 1$, $\theta_n \leq 1/(8c)$ tels que pour toute fonction $g = \mathbb{1}_R \in \mathcal{I}_n$ on a

1. $\mathbb{E}(G^2(X)) \leq \beta$
2. $N(\varepsilon, \mathcal{I}_n) \leq c\varepsilon^{-v}, 0 < \varepsilon < 1$
3. $\mathbb{E}(g^2(X)) \leq \theta_n^2$
4. $\sup_{g \in \mathcal{I}_n} \sup_{x \in \mathcal{X}} |g(x)| \leq \frac{\sqrt{n\theta_n^2 / \log(\beta \vee 1/\theta_n)}}{2\sqrt{v} + 1}$.

Démonstration.

1. Puisque

$$\mathcal{I}_n = \{\mathbb{1}_R, R \in \mathcal{R}_n\}$$

alors $G = 1$ est une fonction enveloppe de \mathcal{I}_n , et dans ce cas $\mathbb{E}(G^2(X)) = 1 < 2$ ainsi on peut poser

$$\beta = 2.$$

2. Par (\dot{H}_∞) on a $\dot{\mathcal{G}}$ est une VC classe de dimension q alors \mathcal{I}_n est une VC classe de dimension q' avec $q' \leq 2q - 1$ donc \mathcal{I}_n est une VC classe de dimension q' . Par définition des VC classes on sait qu'il existe $c > 1$ et $v = 2q' - 1$

tel que

$$N(\varepsilon, \mathcal{I}_n) \leq c\varepsilon^{-2(q'-1)}, 0 < \varepsilon < 1.$$

3. On a $\mathbb{E}(g^2(X)) = \mathbb{E}(1_R(X)) = P(R)$ sous (\dot{H}_1) on $P(R) \leq M\varepsilon_n$ donc pour $\theta_n^2 := M\varepsilon_n = MC\sqrt{\frac{\log \log n}{n}}$ on a

$$\mathbb{E}(g^2(X)) \leq \theta_n^2$$

4. D'une part on a

$$\sup_{g \in \mathcal{I}_n} \sup_{x \in \mathcal{X}} |g(x)| \leq 1$$

D'autre part, $\lim_{n \rightarrow \infty} 1/\theta_n = +\infty$ donc il existe $n_1 > 3$ tel que pour tout $n \geq n_0$ on a $1/\theta_n > 2$ ce qui implique que

$$\log(\beta \vee 1/\theta_n) = \log(1/\theta_n) = \log\left(n^{1/4} / \left(\sqrt{MC}(\log \log n)^{1/4}\right)\right)$$

puisque $n\theta_n^2 = MC\sqrt{n \log \log n}$ alors

$$\lim_{n \rightarrow \infty} \frac{n\theta_n^2}{\log(\beta \vee 1/\theta_n)} = +\infty$$

par conséquent il existe $n_0 > n_1$ tel que pour tout $n > n_0$

$$\frac{\sqrt{n\theta_n^2 / \log(\beta \vee 1/\theta_n)}}{2\sqrt{v+1}} \geq 1$$

donc

$$\sup_{g \in \mathcal{I}_n} \sup_{x \in \mathcal{X}} |g(x)| \leq \frac{\sqrt{n\theta_n^2 / \log(\beta \vee 1/\theta_n)}}{2\sqrt{v+1}}.$$

□

De cette façon, sous les hypothèses (\dot{H}_∞) et (\dot{H}_1) , la classe \mathcal{I}_n vérifie, toutes les conditions de l'inégalité de Talagrand 2.51, l'inégalité des moments 2.53 et le théorème 2.55 de Berthet-Mason. Dans le lemme suivant on applique l'inégalité des moments à X_1, \dots, X_n suite de variables aléatoires i.i.d. de loi P .

Lemme 3.36. Soit $n \geq n_0$, Sous (\dot{H}_∞) et (\dot{H}_1) il existe $C_1 > 0$ tel que

$$\mathbb{E} \left(\sup_{R \in \mathcal{R}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in R} \right| \right) \leq C_1 \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}$$

où (τ_i) est une suite de Rademacher de paramètre 1/2.

Démonstration. Par le lemme 3.35 la classe \mathcal{I}_n vérifie toutes les conditions de l'inégalité des moments, donc il existe $A_2 > 0$ tel que

$$\mathbb{E} \left(\sup_{R \in \mathcal{R}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in R} \right| \right) \leq A_2 \sqrt{v\theta_n^2 \log(1/\theta_n)}$$

avec $\theta_n^2 = MC \sqrt{\frac{\log \log n}{n}}$, donc

$$\mathbb{E} \left(\sup_{R \in \mathcal{R}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in R} \right| \right) \leq A_2 \sqrt{vMC} \frac{(\log \log n)^{1/4}}{n^{1/4}} \left[\log \left(\frac{n^{1/4}}{\sqrt{MC}(\log \log n)^{1/4}} \right) \right]^{1/2}$$

on a

$$\log \left(\frac{n^{1/4}}{\sqrt{MC}(\log \log n)^{1/4}} \right) = \frac{1}{4} \log n - \frac{1}{4} \log \log \log n + \log \left(\frac{1}{\sqrt{MC}} \right)$$

donc

$$\lim_{n \rightarrow \infty} \frac{\log \left(\frac{n^{1/4}}{\sqrt{MC}(\log \log n)^{1/4}} \right)}{\log n} = 1/4$$

ainsi, il existe A_3 et $n_0 > 3$ tel que pour tout $n \geq n_0$ on a

$$\log \left(\frac{n^{1/4}}{\sqrt{MC}(\log \log n)^{1/4}} \right) \leq A_3 \log n$$

par conséquent, il existe $C_1 = A_2 \sqrt{A_3 v MC} > 0$ tel que

$$\mathbb{E} \left(\sup_{R \in \mathcal{R}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in R} \right| \right) \leq C_1 \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}$$

□

Proposition 3.37. *Sous (\dot{H}_∞) , (\dot{H}_1) , il existe $n_0 > 3$ tel que pour tout $n \geq n_0$ on a, pour $C_0, C_1 > 0$, $\zeta > 1$,*

$$P \left\{ \|\Lambda_n\|_{\mathcal{R}_n} \geq C_0 \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right\} \leq \frac{C_1}{n^\zeta}$$

Démonstration. Par le lemme 3.33, la classe de fonctions \mathcal{I}_n associé à la classe d'ensembles \mathcal{R}_n vérifie les hypothèses de l'inégalité de Talagrand 2.51 donc il existe deux constantes positives $A_0, A_1 > 0$ telle que

$$\begin{aligned} \mathbb{P} \left\{ \|\Lambda_n\|_{\mathcal{R}_n} \geq A_0 \left(\mathbb{E} \left(\sup_{R \in \mathcal{R}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in R} \right| \right) + t_n \right) \right\} \\ \leq 2 \exp \left(-\frac{A_1 t_n^2}{\sigma_n^2} \right) + 2 \exp \left(-A_1 t_n \sqrt{n} \right) \end{aligned}$$

avec $\sigma_n^2 = \sup_{R \in \mathcal{R}_n} \text{Var}(\mathbb{1}_{X \in R})$ et

$$t_n = \sqrt{\frac{M}{A_1} \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}}}$$

on sait par lemme 3.36 que sous (\dot{H}_∞) , (\dot{H}_1) il existe $K_1 > 0$ tel que

$$\mathbb{E} \left(\sup_{R \in \mathcal{R}_n} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \tau_i \mathbb{1}_{X_i \in R} \right| \right) \leq K_1 \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}$$

en posant $C_0 = A_0 K_1 + \sqrt{\frac{M}{A_1}}$ on obtient

$$\mathbb{P} \left\{ \|\Lambda_n\|_{\mathcal{R}_n} \geq C_0 \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right\} \leq 2 \exp \left(-\frac{A_1 t_n^2}{\sigma_n^2} \right) + 2 \exp \left(-A_1 t_n \sqrt{n} \right).$$

Sous (\dot{H}_1) on a $m\varepsilon_n P(R) \leq M\varepsilon_n$ donc pour il existe $n_0 > 3$ tel que pour tout $n \geq n_0$

$$\begin{aligned} \text{Var}(\mathbb{1}_{X \in R}) &= P(R) (1 - P(R)) \\ &\leq M\varepsilon_n (1 - m\varepsilon_n) \\ &\leq M\varepsilon_n = MC \sqrt{\frac{\log \log n}{n}} \quad \text{avec } C > 1 \end{aligned}$$

d'une part, on a

$$\begin{aligned} \exp \left(-\frac{A_1 t_n^2}{\sigma_n^2} \right) &\leq \exp \left(-\frac{A_1 t_n^2}{M\varepsilon_n} \right) \\ &= \exp \left(-C \sqrt{\frac{n}{\log \log n}} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right)^2 \right) \\ &= \exp(-C \log n) \\ &= \frac{1}{n^C} \end{aligned}$$

d'autre part,

$$\begin{aligned} \exp(-A_1 t_n \sqrt{n}) &= \exp\left(-A_1 \sqrt{\frac{M}{A_1}} \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \sqrt{n}\right) \\ &= \exp\left(-\sqrt{MA_1} (\log n)^{1/2} (n \log \log n)^{1/4}\right) \end{aligned}$$

or, il existe $A_4 > 0$ et $n_1 > 3$ tel que pour tout $n \geq n_1$

$$(\log n)^{1/2} (n \log \log n)^{1/4} \geq \frac{A_4}{A_1} \log n.$$

donc

$$\begin{aligned} \exp(-A_1 t_n \sqrt{n}) &\leq \exp(-A_4 \log n) \\ &\leq \frac{1}{n^{A_4}}. \end{aligned}$$

En conclusion, il existe $n_* > 3$, $C_0 > 0$, $C_1 > 0$ et $\zeta > 1$ telle que

$$P \left\{ \|\Lambda_n\|_{\mathcal{R}_n} \geq C_0 \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right\} \leq \frac{C_1}{n^\zeta}$$

□

Lemme 3.38. Soit $O \in \mathbb{R}^d$, sous (\dot{H}_∞) , (\dot{H}_1) , presque sûrement, il existe $C_\Delta > 0$ et $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$\|\mathbb{D}_O^n\|_{\mathcal{S}_{d-1} \times \Delta} \leq C_\Delta \sqrt{\log \log n}$$

Démonstration. Sous (\dot{H}_1) on a

$$m\varepsilon \leq \dot{\Psi}(\varepsilon) \leq M\varepsilon, \quad \varepsilon \geq 0$$

ainsi par la proposition 3.23 et la remarque, on a

$$\dot{\Psi}^{-1}(\varepsilon) \leq \frac{\varepsilon}{m}, \quad \varepsilon \geq 0$$

donc pour tout $n > 3$

$$\dot{\Psi}^{-1} \left(\sqrt{\frac{\log \log n}{n}} \right) \leq \frac{1}{m} \sqrt{\frac{\log \log n}{n}}. \quad (3.5)$$

Ensuite, grâce à la loi du logarithme itéré, on sait que presque sûrement, il existe $c_\Delta > 0$ et $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$\|\dot{Y}_{\varphi,n} - \dot{Y}_\varphi\|_{\mathbb{S}_{d-1} \times \Delta} \leq c_\Delta \dot{\psi}^{-1} \left(\sqrt{\frac{\log \log n}{n}} \right)$$

ainsi par l'inégalité 3.5 et en posant $C_\Delta = c_\Delta/m$, on obtient

$$\|\dot{Y}_{\varphi,n} - \dot{Y}_\varphi\|_{\mathbb{S}_{d-1} \times \Delta} \leq C_\Delta \sqrt{\frac{\log \log n}{n}}$$

□

3.2.4.d Preuves de la convergence faible, la représentation de type Bahadur-Kiefer et l'approximation forte

Préliminaire

On fixe $O \in \mathbb{R}^d$, par le théorème de consistance uniforme on sait que

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\dot{Y}_{\varphi,n}(O, u, \alpha) - \dot{Y}_\varphi(O, u, \alpha)| = 0 \text{ p.s.}$$

donc il existe $\gamma_0 > 0$ tel que pour tout $0 < \gamma < \gamma_0$, il existe $n_\gamma > 0$ tel que pour tout $n > n_\gamma$

$$\dot{Y}_{\varphi,n}(O, u, \alpha) \in \{y \in \mathbb{R}, |y - \dot{Y}_\varphi(O, u, \alpha)| < \gamma\}.$$

Pour $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$, on pose

$$y_\alpha = \dot{Y}_\varphi(O, u, \alpha) \text{ et } v_\gamma(y_\alpha) = \{y \in \mathbb{R}, |y - y_\alpha| < \gamma\}.$$

On a

$$\begin{aligned} \dot{Y}_{\varphi,n}(O, u, \alpha) &= \inf_{y \in v_\gamma(y_\alpha)} \{P_n(\dot{G}_\varphi(O, u, y)) \geq \alpha\} \\ &= \inf_{y \in v_\gamma(y_\alpha)} \{P_n(\dot{G}_\varphi(O, u, y)) - P(\dot{G}_\varphi(O, u, y)) \geq \alpha - P(\dot{G}_\varphi(O, u, y))\} \\ &= \inf_{y \in v_\gamma(y_\alpha)} \left\{ (P_n - P)(\dot{G}_\varphi(O, u, y)) \geq F_{\pi_{\varphi,O}(X,u)}(y_\alpha) - F_{\pi_{\varphi,O}(X,u)}(y) \right\} \\ &= \inf_{y \in v_\gamma(y_\alpha)} \left\{ \alpha_n(O, u, y) \geq n^{1/2} \left(F_{\pi_{\varphi,O}(X,u)}(y_\alpha) - F_{\pi_{\varphi,O}(X,u)}(y) \right) \right\} \end{aligned}$$

avec

$$\alpha_n(O, u, y) = n^{1/2} (P_n(\dot{G}_\varphi(O, u, y)) - P(\dot{G}_\varphi(O, u, y)))$$

Sous (\dot{H}_1) la fonction $y \mapsto F_{\pi_{\varphi,O}(X,u)}(y)$ est dérivable sur \mathbb{R} donc par développement de Taylor à l'ordre 1 au voisinage de y_α , on a pour tout $y \in v_\gamma(y_\alpha)$

$$F_{\pi_{\varphi,O}(X,u)}(y_\alpha) - F_{\pi_{\varphi,O}(X,u)}(y) = f_{\pi_{\varphi,O}(X,u)}(y_\alpha)(y_\alpha - y) + \varepsilon_\gamma(O, u, \alpha, y_\alpha - y)$$

avec

$$\lim_{\gamma \rightarrow 0} \varepsilon_\gamma(O, u, \alpha, y_\alpha - y) = 0$$

par conséquent,

$$\dot{Y}_{\varphi,n}(O, u, \alpha) = \inf_{y \in v_\gamma(y_\alpha)} \left\{ \alpha_n(O, u, y) \geq n^{1/2} \left(f_{\pi_{\varphi,O}(X,u)}(y_\alpha)(y_\alpha - y) + \varepsilon_\gamma(O, u, \alpha, y_\alpha - y) \right) \right\}$$

Preuve de l'approximation de Bahadur-Kiefer sans vitesse.

On veut montrer que sous les conditions (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) , on a

$$\lim_{n \rightarrow \infty} \left\| \sqrt{n}(\dot{Y}_{\varphi,n} - Y) + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

avec la notation $\mathbb{E}_O^n(u, \alpha) = \sqrt{n} (P_n(\dot{G}_\varphi(O, u, \alpha)) - P(\dot{G}_\varphi(O, u, \alpha)))$.

Par le lemme 3.38 on sait que, presque sûrement, il existe $C_\Delta > 0$ et $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a pour tout $O \in \mathbb{R}^d$, tout $u \in \mathbb{S}_{d-1}$ et tout $\alpha \in \Delta$, le processus

$$\dot{Y}_{\varphi,n}(O, u, \alpha) \in \left[y_\alpha - C_\Delta \sqrt{\frac{\log \log n}{n}}, y_\alpha + C_\Delta \sqrt{\frac{\log \log n}{n}} \right]$$

donc en posant $\gamma_n = C_\Delta \sqrt{\frac{\log \log n}{n}}$, on a $\dot{Y}_{\varphi,n}(O, u, \alpha) \in v_{\gamma_n}(y_\alpha)$.

Puisque, $\Lambda_O^n(\dot{R}_\varphi(O, u, y, y_\alpha)) = \alpha_n(O, u, y) - \alpha_n(O, u, y_\alpha)$ alors pour tout $y \in v_{\gamma_n}(y_\alpha)$

$$z_n(O, u, y, y_\alpha) := \alpha_n(O, u, y) - \alpha_n(O, u, y_\alpha)$$

est un accroissement du processus empirique sur les rubans de largeur plus petite que $C_\Delta \sqrt{\frac{\log \log n}{n}}$, donc sous (\dot{H}_∞) , (\dot{H}_1) on a par la proposition 3.37,

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right) \quad (3.6)$$

donc

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| = 0 \quad p.s. \quad (3.7)$$

ainsi, pour $n \geq n(\omega)$, on a

$$\begin{aligned}
& \dot{Y}_{\varphi,n}(O, u, \alpha) \\
&= \inf_{y \in v_{\gamma_n}(y_\alpha)} \left\{ \alpha_n(O, u, y) \geq n^{1/2} \left(f_{\pi_{\varphi,O}(X,u)}(y_\alpha)(y_\alpha - y) + \varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y) \right) \right\} \\
&= \inf_{y \in v_{\gamma_n}(y_\alpha)} \left\{ y \geq y_\alpha - \frac{\alpha_n(O, u, y)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right\} \\
&= \inf_{y \in v_{\gamma_n}(y_\alpha)} \left\{ y \geq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} - \frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right\}
\end{aligned}$$

On sait que $|\dot{Y}_{\varphi,n}(O, u, \alpha)| < \infty$, donc par la dernière égalité, on a d'une part,

$$\begin{aligned}
& \dot{Y}_{\varphi,n}(O, u, \alpha) \\
&\geq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} + \inf_{u \in \mathbb{S}^d} \inf_{\alpha \in \Delta} \inf_{y \in v_{\gamma_n}(y_\alpha)} \left(-\frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right) \\
&\geq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \\
&\quad - \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right| - \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right|.
\end{aligned}$$

d'autre part, on a

$$\begin{aligned}
& \dot{Y}_{\varphi,n}(O, u, \alpha) \\
&\leq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} + \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left(-\frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} + \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right) \\
&\leq y_\alpha - \frac{\alpha_n(O, u, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \\
&\quad + \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{n^{1/2} f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right| + \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right|
\end{aligned}$$

donc

$$\begin{aligned}
& \left| n^{1/2}(\dot{Y}_{\varphi,n}(O, u, \alpha) - y_\alpha) + \frac{\alpha_n(O, u, y_\alpha)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right| \tag{3.8} \\
&\leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right| + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi,O}(X,u)}(y_\alpha)} \right|
\end{aligned}$$

Par la condition (\dot{H}_1), on sait que

$$m \leq \inf_{\alpha \in \Delta} \inf_{u \in \mathbb{S}^{d-1}} f_{\pi_{\varphi, O}(X, u)}(y_\alpha) \leq \sup_{\alpha \in \Delta} \sup_{u \in \mathbb{S}^{d-1}} f_{\pi_{\varphi, O}(X, u)}(y_\alpha) \leq M$$

donc

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| \leq \frac{1}{m} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)|$$

ainsi, par l'égalité (3.7), on a

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| = 0 \quad p.s.$$

Ensuite, en posant pour tout $0 < \gamma < \gamma_0$

$$\rho(\gamma) = \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_\gamma(y_\alpha)} |\varepsilon_\gamma(O, u, \alpha, y_\alpha - y)|$$

donc

$$n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| \leq \frac{n^{1/2} \rho(\gamma_n)}{m}$$

par la condition (\dot{H}_2), on a

$$\lim_{n \rightarrow \infty} n^{1/2} \rho(\gamma_n) = 0$$

donc

$$\lim_{n \rightarrow \infty} n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| = 0$$

ainsi,

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \left| n^{1/2} (\dot{Y}_{\varphi, n}(O, u, \alpha) - y_\alpha) + \frac{\alpha_n(O, u, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| = 0 \quad p.s.$$

or avec les notations précédentes, on a

$$\frac{\alpha_n(O, u, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} = \frac{\mathbb{E}_O^n(u, \alpha)}{h_O(u, \alpha)}$$

donc

$$\lim_{n \rightarrow \infty} \left\| \sqrt{n} (\dot{Y}_{\varphi, n} - Y) + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}^{d-1} \times \Delta} = 0 \quad p.s. \quad (3.9)$$

Dans la suite, on note

$$\mathbb{D}_O^n(u, \alpha) = \sqrt{n} (\dot{Y}_{\varphi, n}(O, u, \alpha) - \dot{Y}_\varphi(O, u, \alpha))$$

Preuve du principe d'invariance fort sans vitesse.

On veut montrer que sous les conditions (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite \mathbb{B}_O^n de versions de \mathbb{B}_O de sorte que

$$\dot{Y}_{\varphi,n}(O, u, \alpha) = \dot{Y}_\varphi(O, u, \alpha) + \frac{\mathbb{B}_O^n(u, \alpha)}{h_O(u, \alpha)\sqrt{n}} + \frac{\mathbb{Z}_O^n(u, \alpha)}{\sqrt{n}}, \quad O \in \mathbb{R}^d, \quad \alpha \in \Delta, \quad u \in \mathbb{S}_{d-1},$$

$$\text{où } \mathbb{Z}_O^n = \mathbb{D}_O^n - \frac{\mathbb{B}_O^n}{h_O} \text{ vérifie}$$

$$\lim_{n \rightarrow \infty} \|\mathbb{Z}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = \lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\mathbb{Z}_O^n(u, \alpha)| = 0 \quad p.s.$$

Pour ce faire, on pose \mathbb{G} le pont brownien indexé par $\dot{\mathcal{G}}$ centré et de covariance donnée par

$$\begin{aligned} C_O(u, u', \alpha, \alpha') &= \text{cov}(\mathbb{G}(\dot{\mathcal{G}}_\varphi(O, u, \alpha)), \mathbb{G}(\dot{\mathcal{G}}_\varphi(O, u', \alpha'))) \\ &= P(\dot{\mathcal{G}}_\varphi(O, u, \alpha) \cap \dot{\mathcal{G}}_\varphi(O, u', \alpha')) - P(\dot{\mathcal{G}}_\varphi(O, u, \alpha))P(\dot{\mathcal{G}}_\varphi(O, u', \alpha')) \\ &= P(\dot{\mathcal{G}}_\varphi(O, u, \alpha) \cap \dot{\mathcal{G}}_\varphi(O, u', \alpha')) - \alpha\alpha' \end{aligned}$$

Par (\dot{H}_∞) l'ensemble $\dot{\mathcal{G}}$ forme une classe de Vapnik–Chervonenkis de dimension q , donc en appliquant le théorème de Berthet et Mason 2006 [BM06a] à Λ_n , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite \mathbb{G}_n de versions de \mathbb{G} de sorte que pour tout $\theta > 1$ il existe $K_1 > 0$ tel que

$$\Lambda_n(\dot{\mathcal{G}}_\varphi(O, u, \alpha)) = \mathbb{G}_n(\dot{\mathcal{G}}_\varphi(O, u, \alpha)) + \zeta_n(\dot{\mathcal{G}}_\varphi(O, u, \alpha)) \quad (3.10)$$

avec

$$\lim_{n \rightarrow \infty} \|\zeta_n\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s. \quad (3.11)$$

et

$$\mathbb{P} \left(\sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\zeta_n^n(u, \alpha)| \geq K_1 \frac{(\log n)^{w_d}}{n^{v_d}} \right) \leq \frac{1}{n^\theta} \quad (3.12)$$

avec la notation $\zeta_n^n(u, \alpha) = \zeta_n(\dot{\mathcal{G}}_\varphi(O, u, \alpha))$ et $v_q = 1/(10q - 8)$ et $w_q = (10q - 6)/(20q - 16)$ où q est la dimension VC de $\dot{\mathcal{G}}$.

Ainsi par les équations (3.10) et (3.11), on a presque sûrement

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\zeta_n^n(u, \alpha)| = \lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\Lambda_n(\dot{\mathcal{G}}_\varphi(O, u, \alpha)) - \mathbb{G}_n(\dot{\mathcal{G}}_\varphi(O, u, \alpha))| = 0$$

donc avec les notations

$$\mathbb{E}_O^n(u, \alpha) = \Lambda_n(\dot{G}_\varphi(O, u, \alpha)), \quad \mathbb{G}_O^n(u, \alpha) = \mathbb{G}_n(\dot{G}_\varphi(O, u, \alpha))$$

on conclut que

$$\lim_{n \rightarrow \infty} \|\mathbb{E}_O^n - \mathbb{G}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s. \quad (3.13)$$

Il reste à montrer que sous les conditions (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) ,

$$\lim_{n \rightarrow \infty} \left\| \mathbb{D}_O^n + \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

On sait que

$$\begin{aligned} \left\| \mathbb{D}_O^n + \frac{\mathbb{G}_O^n}{h} \right\|_{\mathbb{S}_{d-1} \times \Delta} &\leq \left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} + \left\| \frac{\mathbb{E}_O^n}{h_O} - \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} \\ &\leq \left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} + \frac{1}{m} \|\mathbb{E}_O^n - \mathbb{G}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} \end{aligned}$$

alors sous (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) , par les limite (3.9) et (3.13) on a

$$\lim_{n \rightarrow \infty} \left\| \mathbb{D}_O^n + \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s. \quad (3.14)$$

Conclusion On a montrer que sous les conditions (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite $\mathbb{B}_O^n := -\mathbb{G}_O^n$ de versions de $\mathbb{B}_O := -\mathbb{G}_O$ de sorte que

$$\dot{Y}_{\varphi, n}(O, u, \alpha) = \dot{Y}_\varphi(O, u, \alpha) + \frac{\mathbb{B}_O^n(u, \alpha)}{h_O(u, \alpha)\sqrt{n}} + \frac{\mathbb{Z}_O^n(u, \alpha)}{\sqrt{n}}, \quad \alpha \in \Delta, \quad u \in \mathbb{S}_{d-1},$$

où $\mathbb{Z}_O^n = \mathbb{D}_O^n - \frac{\mathbb{B}_O^n}{h_O}$ vérifie

$$\lim_{n \rightarrow \infty} \|\mathbb{Z}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = \lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\mathbb{Z}_O^n(u, \alpha)| = 0 \quad p.s.$$

Preuve du théorème limite central uniforme.

On note d_{PL} la distance Prokhorov-levy. Sous (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) on a par la limite (3.14)

$$\lim_{n \rightarrow \infty} \left\| \mathbb{D}_O^n + \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

par conséquent

$$\begin{aligned} \lim_{n \rightarrow \infty} d_{PL}(\sqrt{n}(\dot{Y}_{\varphi,n} - \dot{Y}_{\varphi}), -\frac{\mathbf{G}_O^n}{h_O}) &= \lim_{n \rightarrow \infty} d_{PL}(\sqrt{n}(\dot{Y}_{\varphi,n} - \dot{Y}_{\varphi}), -\frac{\mathbf{G}_O}{h_O}) \\ &= \lim_{n \rightarrow \infty} d_{PL}(\sqrt{n}(\dot{Y}_{\varphi,n} - \dot{Y}_{\varphi}), \frac{\mathbf{G}_O}{h_O}) \\ &= 0 \end{aligned}$$

donc sous les conditions (\dot{H}_{∞}) , (\dot{H}_1) et (\dot{H}_2) on a

$$\mathbb{D}_O^n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathbb{B}_O$$

sur l'ensemble des fonctions bornées sur $\mathbb{S}_{d-1} \times \Delta$ muni de la norme uniforme et avec la notation $\mathbb{B}_O = \frac{\mathbf{G}_O}{h_O}$ et

$$\text{cov}(\mathbb{B}_O(u, \alpha), \mathbb{B}_O(u', \alpha')) = \frac{P(\dot{G}_{\varphi}(O, u, \alpha) \cap \dot{G}_{\varphi}(O, u', \alpha')) - \alpha\alpha'}{h_O(u, \alpha)h_O(u', \alpha')}$$

Preuve de représentation de type Bahadur-Kiefer avec vitesse

On fixe O et on veut montrer que sous les conditions (\dot{H}_{∞}) , (\dot{H}_1) , (\dot{H}_2) et $[(\dot{H}_3)$ ou $(\dot{H}'_3)]$, il existe $n(\omega)$ tel que pour tout $n \geq n(\omega)$

$$\left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right).$$

Par (\dot{H}_{∞}) la classe des demi-espaces $\dot{\mathcal{G}}$ est une classe de Vapnik-Červonenkis, de dimension q , donc par la loi du logarithme itéré (Alexander 1984 [Ale84]) on a

$$\limsup_{n \rightarrow \infty} \frac{\|\Lambda_n\|_{\dot{\mathcal{G}}}}{\sqrt{2 \log \log n}} \leq \frac{1}{2} \quad p.s.$$

donc avec probabilité 1, il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a pour tout $u \in \mathbb{S}_{d-1}$ le processus

$$\alpha_n(O, u, y) := \Lambda_n(\dot{G}_{\varphi}(O, u, y)) \in \left[-\sqrt{3 \log \log n}, \sqrt{3 \log \log n} \right]$$

Pour tout $n \geq n(\omega)$, on rappelle l'inégalité (3.8)

$$\begin{aligned} & \left| n^{1/2}(\dot{Y}_{\varphi,n}(O, u, \alpha) - y_{\alpha}) + \frac{\alpha_n(O, u, y_{\alpha})}{f_{\pi_{\varphi, O}(X, u)}(y_{\alpha})} \right| \\ & \leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_{\alpha})} \left| \frac{z_n(O, u, y, y_{\alpha})}{f_{\pi_{\varphi, O}(X, u)}(y_{\alpha})} \right| + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_{\alpha})} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_{\alpha} - y)}{f_{\pi_{\varphi, O}(X, u)}(y_{\alpha})} \right| \end{aligned}$$

ainsi que l'égalité (3.6)

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right)$$

donc il existe une constante $C' > 0$, telle que tout $n \geq n(\omega)$

$$\sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| \leq C' \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right)$$

alors pour $n \geq n(\omega)$

$$\begin{aligned} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| &\leq \frac{1}{m} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} |z_n(O, u, y, y_\alpha)| \\ &\leq \frac{C'}{m} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right) \end{aligned} \quad (3.15)$$

Ensuite, en posant pour tout $\gamma > 0$

$$\rho(\gamma) = \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_\gamma(y_\alpha)} |\varepsilon_\gamma(O, u, \alpha, y_\alpha - y)|$$

Sous la condition (H'_3) , on sait qu'il existe $r > 1/2$ et $C^* > 0$ tel que pour tout $\gamma > 0$ on a

$$\rho(\gamma) \leq C^* \gamma^{1+r}$$

donc pour $n \geq n(\omega)$

$$\begin{aligned} n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| &\leq \frac{\rho(\gamma_n)}{m} \\ &\leq \frac{C^* C_\Delta}{m} n^{1/2} \left(\sqrt{\frac{\log \log n}{n}} \right)^{1+r} \\ &\leq \frac{C'_\Delta}{m} \left(\frac{(\log \log n)^{\frac{1+r}{2}}}{n^{\frac{r}{2}}} \right) \end{aligned} \quad (3.16)$$

où $C'_\Delta = C^* C_\Delta$.

avec les notations précédentes on a

$$\left| n^{1/2} (\dot{Y}_{\varphi, n}(O, u, \alpha) - y_\alpha) + \frac{\alpha_n(O, u, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| = \left| \mathbb{D}_O^n(u, \alpha) + \frac{\mathbb{E}_O^n(u, \alpha)}{h_O(u, \alpha)} \right|$$

on a par les inégalités (3.15) et (3.16), pour tout $n \geq n(\omega)$

$$\begin{aligned}
& \left| \mathbb{D}_O^n(u, \alpha) + \frac{\mathbb{E}_O^n(u, \alpha)}{h_O(u, \alpha)} \right| \\
& \leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| \\
& \leq \frac{C'_\Delta}{m} \left(\frac{(\log \log n)^{\frac{1+r}{2}}}{n^{\frac{r}{2}}} \right) + \frac{C'}{m} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right) \\
& \leq C''_\Delta \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}. \tag{3.17}
\end{aligned}$$

Ainsi, sous les conditions (\dot{H}_∞) , (\dot{H}_1) , (\dot{H}_2) et (\dot{H}_3) , il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$

$$\left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}^{d-1} \times \Delta} = O_{p.s} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right).$$

Sous la condition (\dot{H}_3) , on sait que

$$\lim_{\gamma \rightarrow 0} \frac{\rho(\gamma)}{\gamma^{3/2} \sqrt{\log(1/\gamma)}} = 0$$

donc pour $n \geq n_1(\omega)$, avec la notation $t_n = n^{-1/4} (\log n)^{1/2} (\log \log n)^{1/4}$

$$\begin{aligned}
\left| \mathbb{D}_O^n(u, \alpha) + \frac{\mathbb{E}_O^n(u, \alpha)}{h_O(u, \alpha)} \right| & \leq \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{z_n(O, u, y, y_\alpha)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| \\
& \quad + n^{1/2} \sup_{u \in \mathbb{S}^d} \sup_{\alpha \in \Delta} \sup_{y \in v_{\gamma_n}(y_\alpha)} \left| \frac{\varepsilon_{\gamma_n}(O, u, \alpha, y_\alpha - y)}{f_{\pi_{\varphi, O}(X, u)}(y_\alpha)} \right| \\
& \leq \frac{n^{1/2} \rho(\gamma_n)}{m} + \frac{C'}{m} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right) \\
& \leq \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4} m} \left(\frac{\rho(\gamma_n) (\log \log n)^{1/2}}{(\gamma_n)^{3/2} (\log n)^{1/2}} + C' \right) \\
& \leq C'' \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}
\end{aligned}$$

car sous (\dot{H}_3) ,

$$\lim_{n \rightarrow \infty} \frac{\rho(\gamma_n) (\log \log n)^{1/2}}{(\gamma_n)^{3/2} (\log n)^{1/2}} = 0.$$

Par conséquent, sous les conditions (\dot{H}_∞) , (\dot{H}_1) , (\dot{H}_2) et (\dot{H}_3) , il existe un entier $n^*(\omega) = \max(n(\omega), n_1(\omega)) > 0$ tel que pour tout $n \geq n^*(\omega)$,

$$\left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} \right).$$

Preuve de l'approximation forte avec vitesse

Dans la preuve du théorème limite central uniforme, on a montré que sous (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite \mathbb{G}_O^n de versions de \mathbb{G}_O tel que pour tout $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$, on a

$$\mathbb{D}_O^n(u, \alpha) = -\frac{\mathbb{G}_O^n(u, \alpha)}{h_O(u, \alpha)} + \mathbb{Z}_O^n(u, \alpha)$$

où $\mathbb{Z}_O^n(u, \alpha) = \mathbb{D}_O^n(u, \alpha) + \frac{\mathbb{G}_O^n(u, \alpha)}{h_O(u, \alpha)}$ avec

$$\lim_{n \rightarrow \infty} \|\mathbb{Z}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = 0 \quad p.s.$$

On veut montrer que sous la condition (\dot{H}_3) , il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a

$$\|\mathbb{Z}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_q}}{n^{v_q}} \right)$$

avec $v_q = 1/(10q - 8)$ et $w_q = (10q - 6)/(20q - 16)$.

On sait que

$$\begin{aligned} \left\| \mathbb{D}_O^n + \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} &\leq \left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} + \left\| \frac{\mathbb{E}_O^n}{h_O} - \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} \\ &\leq \left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} + \frac{1}{m} \|\mathbb{E}_O^n - \mathbb{G}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} \end{aligned}$$

Sous la condition (\dot{H}_3) , par l'inégalité (3.17), il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$, on a

$$\left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} \leq C_\Delta'' \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}}$$

ensuite par (3.12) et Borel-Cantelli, on a

$$\|\mathbb{E}_O^n - \mathbb{G}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_q}}{n^{v_q}} \right)$$

donc il existe $K_1 > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$\left\| \mathbb{E}_O^n - \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} \leq K_1 \left(\frac{(\log n)^{w_q}}{n^{v_q}} \right)$$

$$\begin{aligned} \left\| \mathbb{D}_O^n + \frac{\mathbb{G}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} &\leq \left\| \mathbb{D}_O^n + \frac{\mathbb{E}_O^n}{h_O} \right\|_{\mathbb{S}_{d-1} \times \Delta} + \frac{1}{m} \|\mathbb{E}_O^n - \mathbb{G}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} \\ &\leq C'_\Delta \frac{(\log \log n)^{1/4} (\log n)^{1/2}}{n^{1/4}} + K_1 \frac{(\log n)^{w_q}}{n^{v_q}} \\ &\leq \frac{D_0 (\log n)^{w_q}}{n^{v_q}} \end{aligned}$$

Conclusion Sous (\dot{H}_∞) , (\dot{H}_1) et (\dot{H}_2) on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite $\mathbb{B}_O^n = -\mathbb{G}_O^n$ de versions de $\mathbb{B}_O = -\mathbb{G}_O$ de sorte que

$$\dot{Y}_{\varphi,n}(O, u, \alpha) = \dot{Y}_\varphi(O, u, \alpha) + \frac{\mathbb{B}_O^n(u, \alpha)}{h_O(u, \alpha)\sqrt{n}} + \frac{\mathbb{Z}_O^n(u, \alpha)}{\sqrt{n}}, \quad O \in \mathbb{R}^d, \quad \alpha \in \Delta, \quad u \in \mathbb{S}_{d-1},$$

où $\mathbb{Z}_O^n = \mathbb{D}_O^n - \frac{\mathbb{B}_O^n}{h_O}$ vérifie

$$\lim_{n \rightarrow \infty} \|\mathbb{Z}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = \lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\mathbb{Z}_O^n(u, \alpha)| = 0 \quad p.s.$$

si de plus P satisfait (\dot{H}_3) alors on peut construire \mathbb{G}_O^n tel que pour $v_q = 1/(10q - 8)$ et $w_q = (10q - 6)/(20q - 16)$, il existe $n_\omega > 0$ tel que tout pour $n > n_\omega$,

$$\|\mathbb{Z}_O^n\|_{\mathbb{S}_{d-1} \times \Delta} = O_{p.s.} \left(\frac{(\log n)^{w_q}}{n^{v_q}} \right).$$

3.3 Surfaces quantiles généralisées

3.3.1 Indexation par des formes générales

Soit O un point de \mathbb{R}^d , u_0 un vecteur de \mathbb{S}_{d-1} et g une fonction continue de \mathbb{R}^d dans \mathbb{R} telle que $\lim_{\|x\| \rightarrow \infty} g(x) = +\infty$ vérifiant

- la propriété de croissance : pour tout $y_1 \leq y_2$

$$g^{-1}((-\infty, y_1]) := G_{y_1} \subset G_{y_2} \quad (3.18)$$

- la propriété de non constance : pour tout $y \in \mathbb{R}$

$$\lambda_d(g^{-1}(y)) = 0 \quad (3.19)$$

Remarque 3.39. Alors que dans le cas du quantile à travers une forme fonctionnelle, c'est par la même fonction φ fixé (modulo une translation et rotation) qu'on réalisait la projection dans toute les directions, ici, dans chaque direction la forme peut « évoluer ». Ci-dessous deux exemples de géodésique de projection dans le cas général.

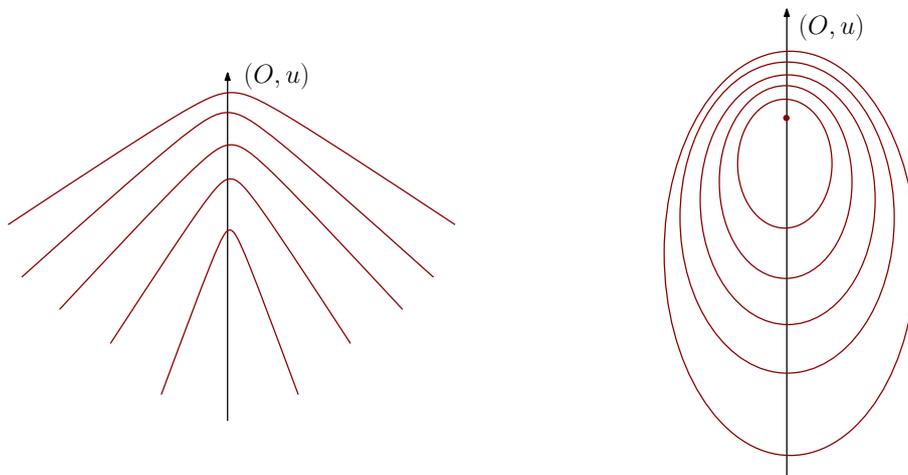


FIGURE 3.7 – Deux exemples de géodésique de projection dans le cas général.

Notation 3.40. On reprend les notations 3.1 de la section précédente

- r_{θ_u} la rotation de centre 0_d et d'angle $\widehat{(u_0, u)}$ laissant invariant $\text{vect}(u, u_0)^\perp$
- t_O la translation de vecteur $0_d - O$
- t_{-O} la translation de vecteur $O - 0_d$
- $G_g(O, u, y) = t_O \circ r_{\theta_u}(G_y)$

Remarque 3.41. Pour $u = u_0$ et $O = 0_d$ si on pose $g = \langle p_\varphi, u_0 \rangle$ alors

$$G_y = \dot{G}_g(0_d, u_0, y)$$

Comme dit précédemment, le cas des quantiles par la fonction φ n'est qu'un cas particulier de la généralisation suivante.

Le g -quantile vu de O dans la direction u et de niveau α correspond au $y \in \mathbb{R}$ tel que $G_g(O, u, y)$ est de probabilité α .

Définition 3.42. Soit $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$, $\alpha \in (0, 1)$ et g vérifiant (3.18), (3.19). Le g -quantile vu de O dans la direction u et de niveau α est défini par

$$\check{Y}_g(O, u, \alpha) = \inf\{y \in \mathbb{R} : P(G_g(O, u, y)) \geq \alpha\}.$$

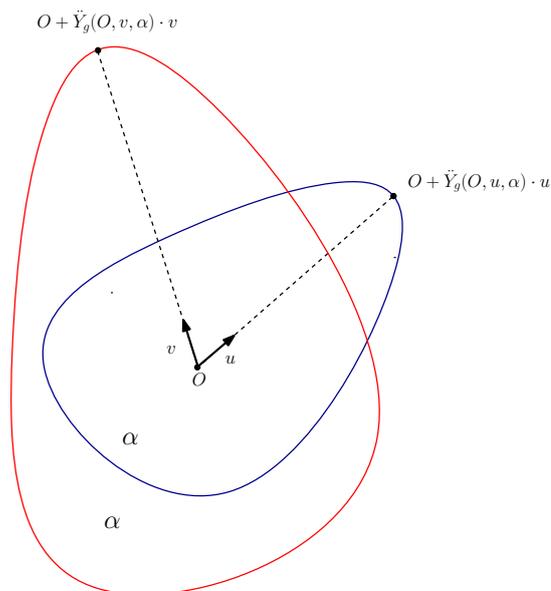


FIGURE 3.8 – Exemple de $\check{Y}_g(O, u, \alpha)$ et $\check{Y}_g(O, v, \alpha)$ pour G_y fermé.

Ensuite on définit la g -région vue de O de masse α

Définition 3.43. Soit $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in (0, 1)$.

La g -région vue de O de masse α est définie par

$$G_g(O, u, \alpha) := G(O, u, \check{Y}_g(O, u, \alpha))$$

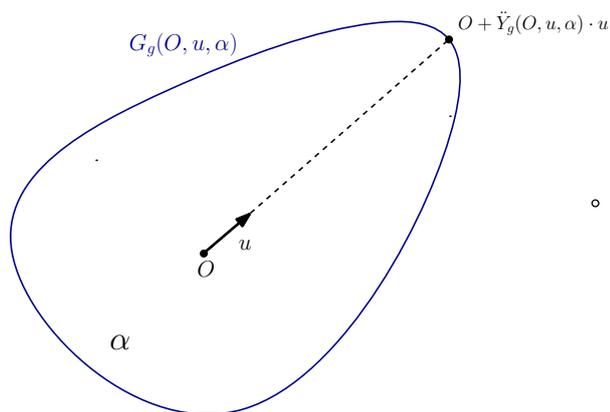
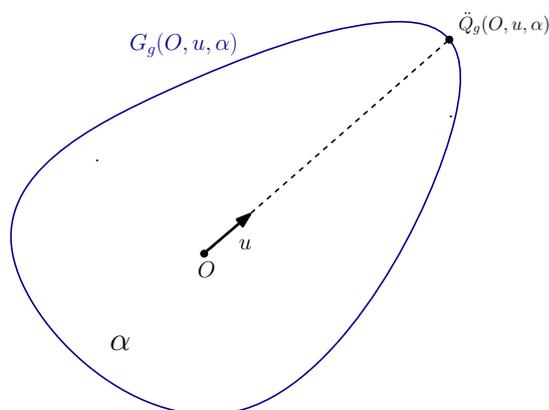
Voici un exemple (fig. 3.9) d'une g -région vue de O de masse α , avec g choisie de sorte que $g^{-1}((-\infty, y])$ soit un fermé.

Remarque 3.44. Pour des fonctions g quelconques, $G_g(O, u, \alpha)$ dépend fortement de O , ainsi on perd la propriété d'indépendance qu'on avait dans le cadre des hyperplans.

Définition 3.45. Soit $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in (0, 1)$.

Le point g -quantile vu de O dans la direction u et de niveau α est défini par

$$\check{Q}_g(O, u, \alpha) = O + \check{Y}_g(O, u, \alpha) \cdot u$$

FIGURE 3.9 – Exemple de $G_g(O, u, \alpha)$ pour G_Y fermé.FIGURE 3.10 – Exemple de $G_g(O, u, \alpha)$ et $\check{Q}_g(O, u, \alpha)$ pour G_Y fermé.

Les points g -quantiles vu de O dans la direction u et de niveau α étant définis dans toutes les directions u de S_{d-1} , on peut donner la définition de la « g -surface » quantile vue de O de niveau α .

Définition 3.46. Soit $O \in \mathbb{R}^d$, $u \in S_{d-1}$ et $\alpha \in (0, 1)$.

La g -surface quantile vue de O de niveau α est définie par

$$\check{Q}_g(O, \alpha) = \{\check{Q}_g(O, u, \alpha) : u \in S_{d-1}\}$$

Sans condition sur la loi P et la classe de fonction g , rien ne peut garantir ni la régularité ni la continuité de cette surface, néanmoins, le fait que $\check{Q}_g(O, u, \alpha)$ soit bien défini pour tout $O \in \mathbb{R}^d$, $u \in S_{d-1}$ et $\alpha \in (0, 1)$ permet de définir l'ensemble $\check{Q}_g(O, \alpha)$.

Contrairement au chapitre sur les quantiles via les hyperplans, le fait que les $G_g(O, u, \alpha)$ soient fortement liés à l'observateur ne permet pas de

définir un objet intrinsèque, qui joue le même rôle que l'ensemble de Tukey-Nolan.

Même si cette dépendance en O peut sembler rédhibitoire, elle peut être un atout car le fait de changer de point d'observation permet d'avoir une information différente, en croisant les informations obtenues non seulement en changeant de point de vue, mais aussi la forme φ et le niveau de quantile α , permet inexorablement d'enrichir la connaissance (modes, masses, dispersion, etc ...) sur la loi inconnue P .

Soit X_1, \dots, X_n une suite de variables aléatoires indépendantes de même loi P . on note $P_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ la mesure empirique associé à P . Voici, de manière complètement analogue, les versions empiriques des objets présentés dans cette section.

Définition 3.47. *Pour $O \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in (0, 1)$, on définit*

- le g -quantile vu de O dans la direction u et de niveau α par

$$\check{Y}_{g,n}(O, u, \alpha) = \inf\{y \in \mathbb{R} : P_n(G(O, u, y)) \geq \alpha\}.$$

- le point g -quantile vu de O dans la direction u et de niveau α par

$$\check{Q}_{g,n}(O, u, \alpha) = O + \check{Y}_{g,n}(O, u, \alpha) \cdot u.$$

- la g -région empirique vue de O de masse α est définie par

$$G_{g,n}(O, u, \alpha) := G(O, u, \check{Y}_{g,n}(O, u, \alpha)).$$

- la g -ensemble de quantile empirique vue de O de niveau α par

$$\check{Q}_{g,n}(O, \alpha) = \{\check{Q}_{g,n}(O, u, \alpha) : u \in \mathbb{S}_{d-1}\}$$

Remarque 3.48. *Les mêmes théorèmes que dans le cas précédent peuvent être obtenus pour ces quantiles généralisés, néanmoins les hypothèses doivent être renforcées afin de garantir que $\check{Q}_g(O, \alpha)$ est bien une surface.*

3.4 Simulations et interprétations

Dans cette section, on se place dans \mathbb{R}^2 et on s'intéresse aux simulations pour différents types de lois (symétrique/asymétrique, unimodale/multimodale), à travers des exemples on pointerá les différences sur la surface quantile avec le cas des surfaces quantiles via demi-espaces. Grâce à ces exemples bien choisis, on montre les différentes utilisations possibles des surfaces quantile, pour la localisation de masse ainsi que la détection de modes.

3.4.1 La loi gaussienne

3.4.1.a Gaussienne $N(0_2, \Sigma)$ avec $\Sigma = \lambda I_2$

Ici on pose

$$P = N(0_2, I_2)$$

Les surfaces sont générées avec un échantillon de taille $N = 10000$, pour commencer on choisit comme forme de référence $\varphi(x) = ax^2$ où $a \in \mathbb{R}$ dont la valeur sera précisée dans chaque exemple (le but ici n'est pas de jouer sur la courbure, car ceci fait partie d'un autre type de problèmes auxquels on ne s'est pas intéressé dans cette thèse).

Généralement on prend $a = 1$ une valeur qui permet dans la plupart des cas étudiés, de faire des interprétations en termes de détection de modes et de localisation de masse.

Voici une première simulation pour $a = 1$, $O = \mathbb{E}(X) = (0, 0)$ et $\alpha = 0.8$

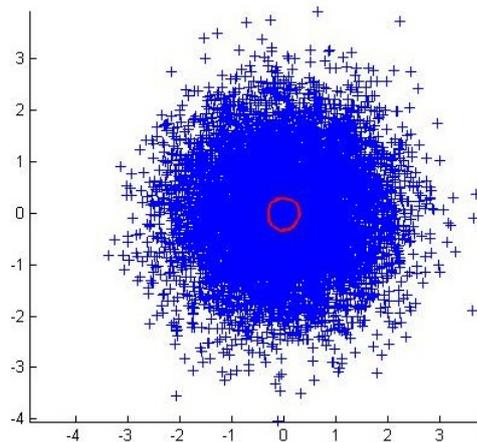
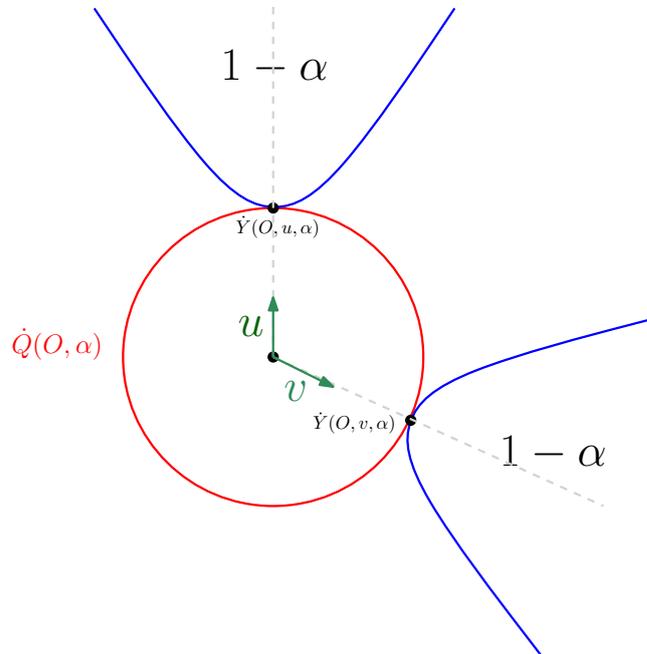


FIGURE 3.11 – $\hat{Q}(O, \alpha)$ pour $O = \mathbb{E}(X) = (0, 0)$ et $\alpha = 0.8$

La surface $\dot{Q}(O, \alpha)$ est un cercle de centre 0_2 , ceci est dû à l'invariance de la loi par rotation. On rappelle que chaque point est obtenu comme suit



En faisant décroître $\alpha = 0.714$ on atteint la « valeur seuil » pour la courbure $a = 1$, i.e. la surface $\dot{Q}(O, \alpha)$ est réduite à un point.

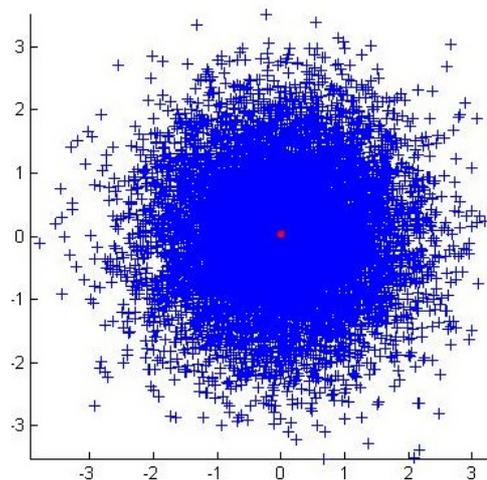
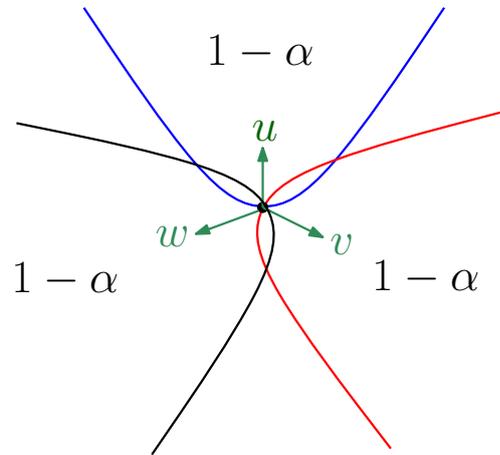


FIGURE 3.12 – $\dot{Q}(O, \alpha)$ pour $O = \mathbb{E}(X) = (0, 0)$ et $\alpha = 0.714$

Ceci s'explique par le fait que dans n'importe quelle direction u , la forme $\varphi_{u,O}$ exclue $1 - \alpha \approx 0.286$



Voici, $\dot{Q}(O, \alpha)$ lorsque $\alpha \in \{0.7, 0.8, 0.9\}$

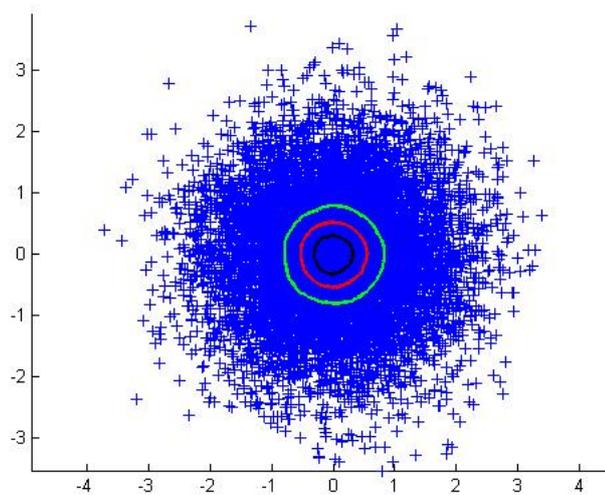


FIGURE 3.13 – $\dot{Q}(O, \alpha)$ pour $O = \mathbb{E}(X) = (0, 0)$ et $\alpha \in \{0.7, 0.8, 0.9\}$

Ensuite si pose $O = (-2, 0)$ alors voici la surface $\dot{Q}(O, 0.6)$ pour $\varphi(x) = x^2$.

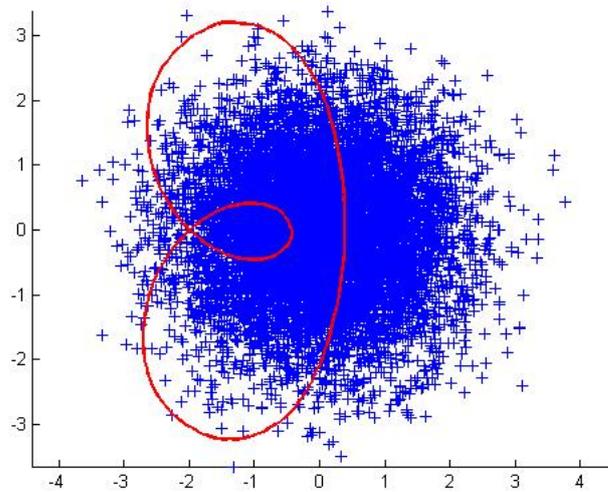


FIGURE 3.14 – $\dot{Q}(O, \alpha)$ pour $O = (-2, 0)$ et $\alpha = 0.6$

Voici, une collection de $\dot{Q}(O, \alpha)$ pour $O = (-2, 0)$ et $\alpha \in \{0.7, 0.8, 0.9\}$

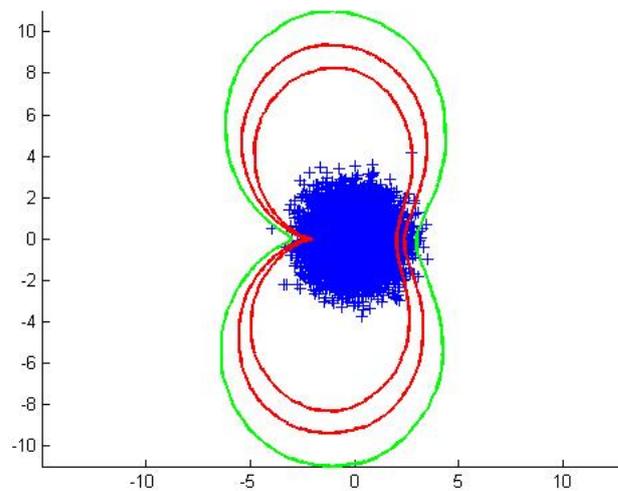


FIGURE 3.15 – $\dot{Q}(O, \alpha)$ pour $O = (-2, 0)$ et $\alpha \in \{0.7, 0.8, 0.9\}$

Enfin, pour une collection de $O \in \{(-5, -5), (-5, 5), (5, -5), (5, 5)\}$ voici $\dot{Q}(O, \alpha)$ avec $\alpha = 0.8$
Avec toujours les boucles intérieures qui pointent vers la masse.

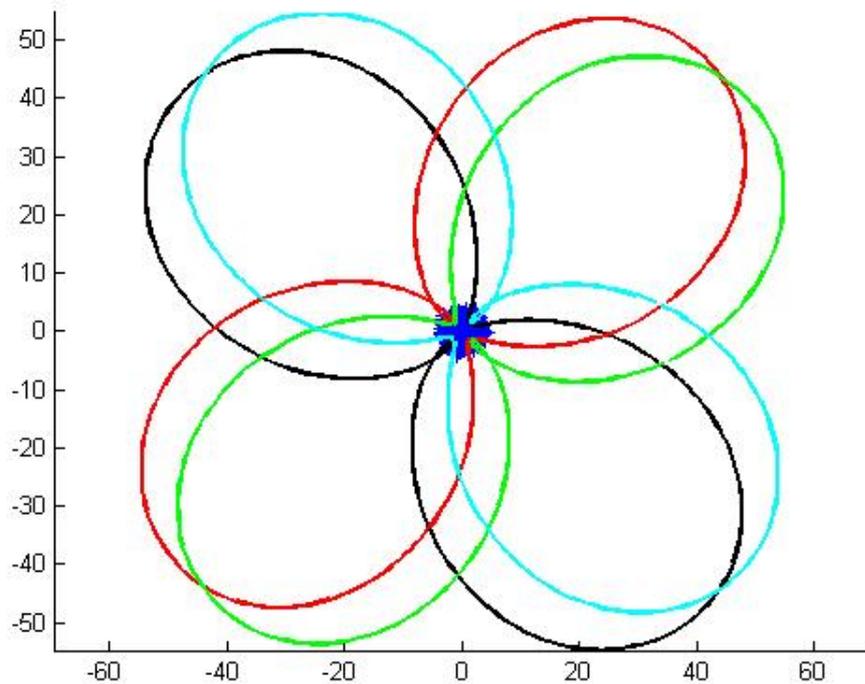


FIGURE 3.16 – $\dot{Q}(O, \alpha)$ pour $O \in \{(-5, -5), (-5, 5), (5, -5), (5, 5)\}$ et $\alpha = 0.8$

3.4.2 Lois multimodales : mélange gaussien

Dans cette partie, on s'intéresse aux surfaces quantiles de lois avec plusieurs modes. On commence par une loi bimodale, comme mélange de deux gaussiennes,

$$P_1 = 1/4N(O_1, \Sigma_1) + 3/4N(O_2, \Sigma_2).$$

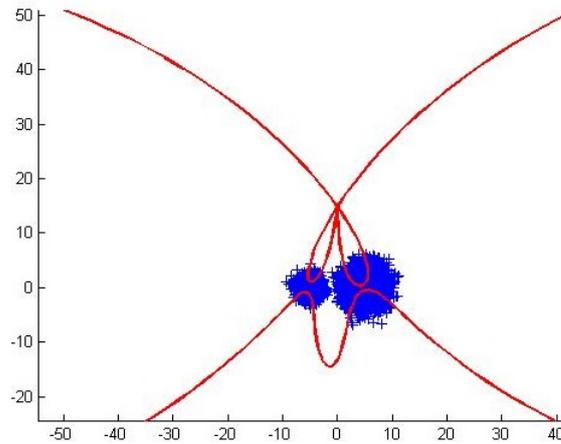
Ensuite on s'intéresse à une loi avec trois modes,

$$P_2 = 1/4N(O_1, \Sigma_1) + 5/12N(O_2, \Sigma_2) + 5/12N(O_2, \Sigma_2) + 1/3N(O_3, \Sigma_3).$$

3.4.2.a Simulation avec P_1 pour $O_1 = (-5, 0)$, $O_2 = (5, 0)$, $\Sigma_1 = I_2$ et $\Sigma_2 = 3I_2$

Dans cette section également, on commence par choisir deux modes bien séparés afin de comprendre l'influence des modes sur la forme de la surface $\dot{Q}(O, \alpha)$ avec $\varphi(x) = x^2$.

Voici un grossissement de $\dot{Q}(O, \alpha)$ pour $O = (0, 15)$ et $\alpha = 0.8$

FIGURE 3.17 – $\dot{Q}(O, \alpha)$ pour $O = (0, 15)$ et $\alpha = 0.8$

Ici on remarque que $\dot{Q}(O, \alpha)$ est une surface avec deux boucles intérieures dont chacune pointe vers un des modes.

En choisissant différents points d'observations on peut faire plusieurs remarques, pour ce faire on pose la grille suivante :

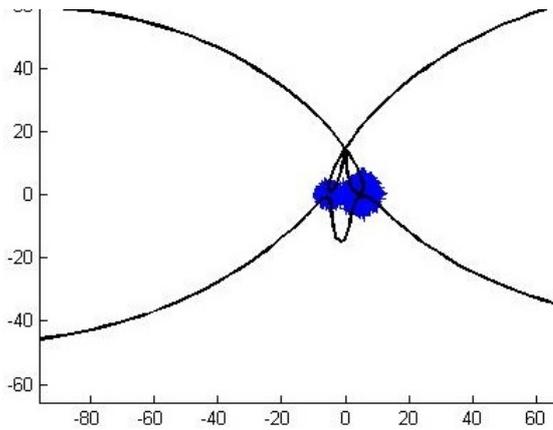
$$O \in \{(0, 15), (0, -15), (-10, 15), (10, -15), (-15, 0), (15, 0), (-5, 0), (5, 0), (0, 0)\}$$

avec $\alpha = 0.8$. Pour des raisons de lisibilité des graphiques, on donne des graphiques séparés pour chaque point d'observation choisi.

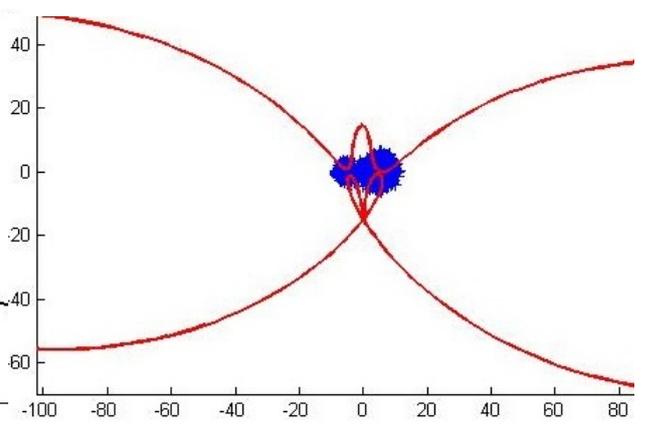
Les graphiques suivants 3.18a, 3.18b, 3.18c, 3.18d, représentent des surfaces $\dot{Q}(O, \alpha)$ avec un observateur O permettant uniquement à partir de la forme de la surface, d'en déduire que la loi P est bimodale.

Ensuite, à partir des graphiques 3.18e, 3.18f, la surface $\dot{Q}(O, \alpha)$ est composée d'une seule boucle intérieure, ainsi il nous est impossible de savoir uniquement à partir de ce point d'observation, que la loi P est bimodale. Ceci est dû à l'alignement de l'observateur avec deux des modes, ainsi à partir de son point d'observation, il ne peut « voir » les deux modes de manière séparée. La même observation peut être faite pour des observateurs choisis à l'intérieur des zones modales (voir les graphiques 3.19a et 3.19b).

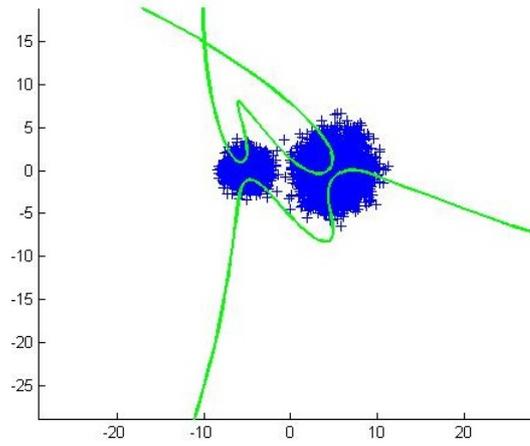
Enfin, le graphique 3.20 représente la surface $\dot{Q}(O, \alpha)$ pour $O = (0, 0)$, on observe une surface avec deux boucles dont chacune pointe vers un des modes, une observation qui indique de manière visuelle, l'existence de deux modes.



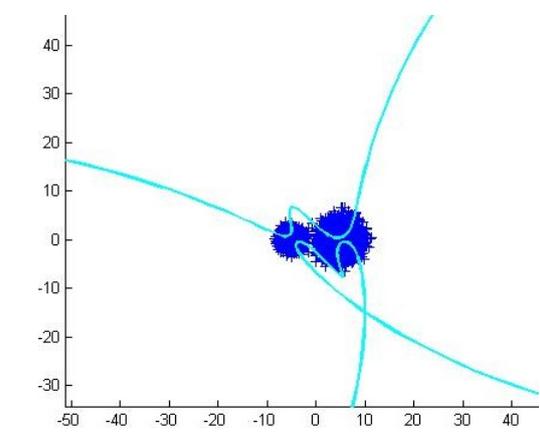
(a) $\dot{Q}(O, \alpha)$ pour $O = (0, -15)$ et $\alpha = 0.8$



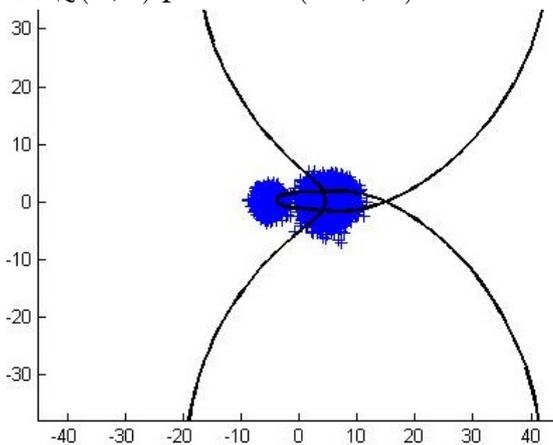
(b) $\dot{Q}(O, \alpha)$ pour $O = (0, -15)$ et $\alpha = 0.8$



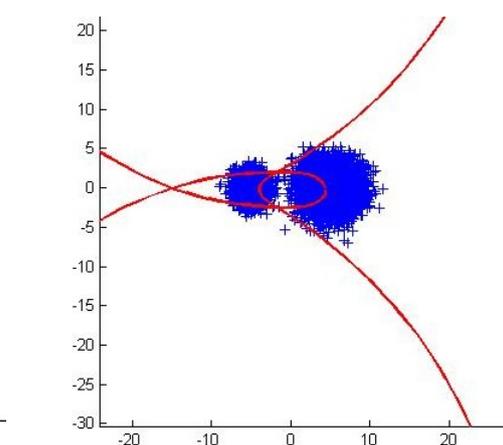
(c) $\dot{Q}(O, \alpha)$ pour $O = (-10, 15)$ et $\alpha = 0.8$



(d) $\dot{Q}(O, \alpha)$ pour $O = (10, -15)$ et $\alpha = 0.8$



(e) $\dot{Q}(O, \alpha)$ pour $O = (-15, 0)$ et $\alpha = 0.8$



(f) $\dot{Q}(O, \alpha)$ pour $O = (-15, 0)$ et $\alpha = 0.8$

FIGURE 3.18

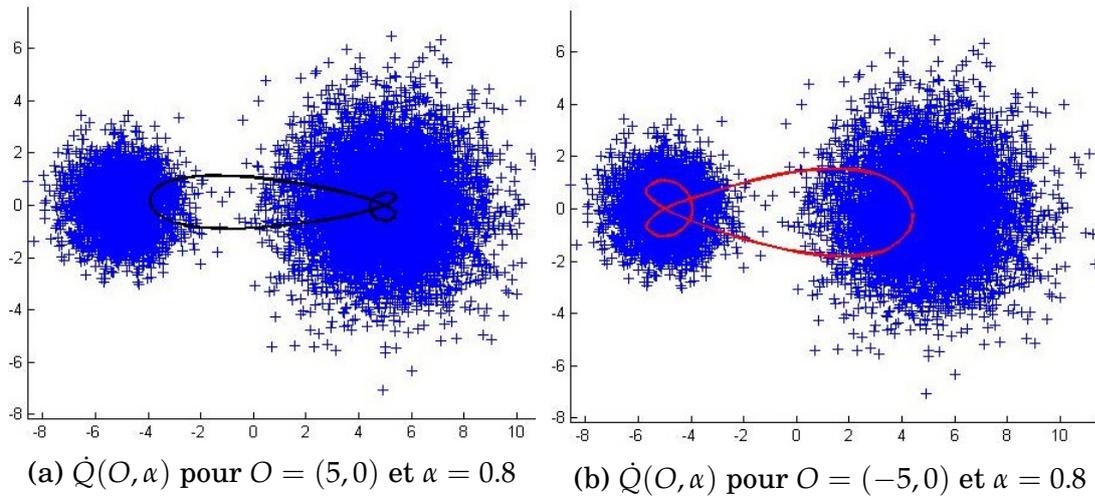
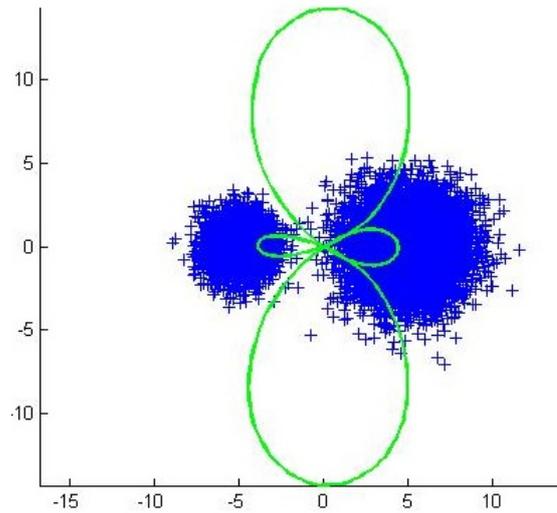


FIGURE 3.19

FIGURE 3.20 – $\dot{Q}(O, \alpha)$ pour $O = (0, 0)$ et $\alpha = 0.8$

3.4.2.b Simulation avec P_2

Cas I : trois modes « alignés » On pose

$$O_1 = (-10, 0), O_2 = (0, 0), O_3 = (10, 0)$$

et

$$\Sigma_1 = I_2, \Sigma_2 = 3/2I_2, \Sigma_3 = 1/2I_2$$

On choisit différents points d'observations, pour ce faire on pose la grille suivante :

$$O \in \{(0, 20), (-10, 10), (10, 10), (6, -15), (-6, 0), (0, 0), (-20, 0), (20, 0)\}$$

avec $\alpha = 0.8$. Pour des raisons de lisibilité des graphiques, on donne des graphiques séparés pour chaque point d'observation choisi.

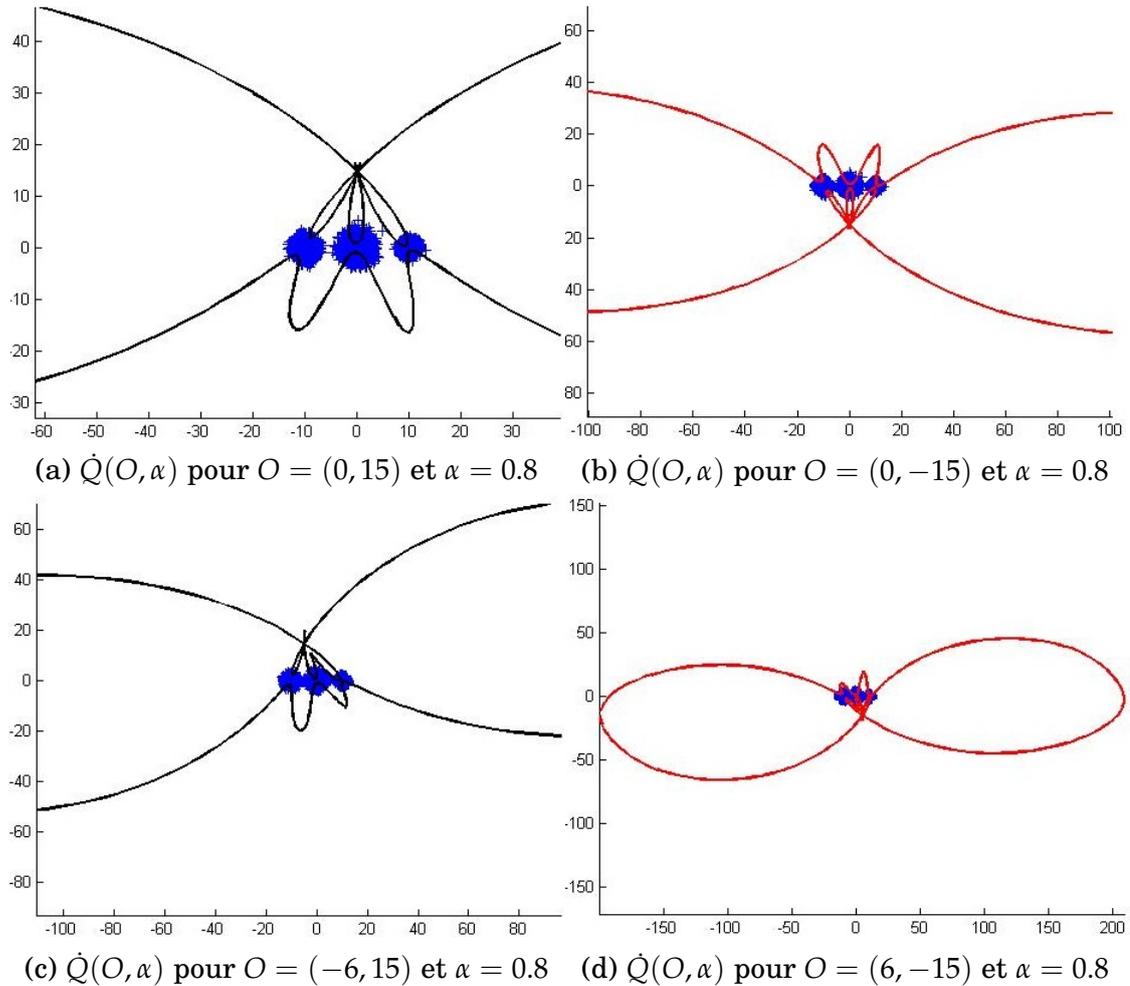


FIGURE 3.21 – Trois modes sur trois détectés.

Les graphiques précédents 3.21a, 3.21b, 3.21c, 3.21d, représentent des surfaces $\dot{Q}(O, \alpha)$ avec un observateur O permettant, uniquement à partir de la forme de la surface quantile (trois boucles), d'en déduire que la loi P_2 est tri modale.

Par contre ce cas-ci (3.22a, 3.22b), on est face au cas où le point O est aligné avec les centres des modes de P_2 avec au moins un mode de chaque côté, de ce fait la surface quantile $\hat{Q}(O, \alpha)$ est une surface avec deux boucles, ainsi avec seulement ces points d'observations, on peut uniquement affirmer que la loi P_2 est une loi au moins bimodale.

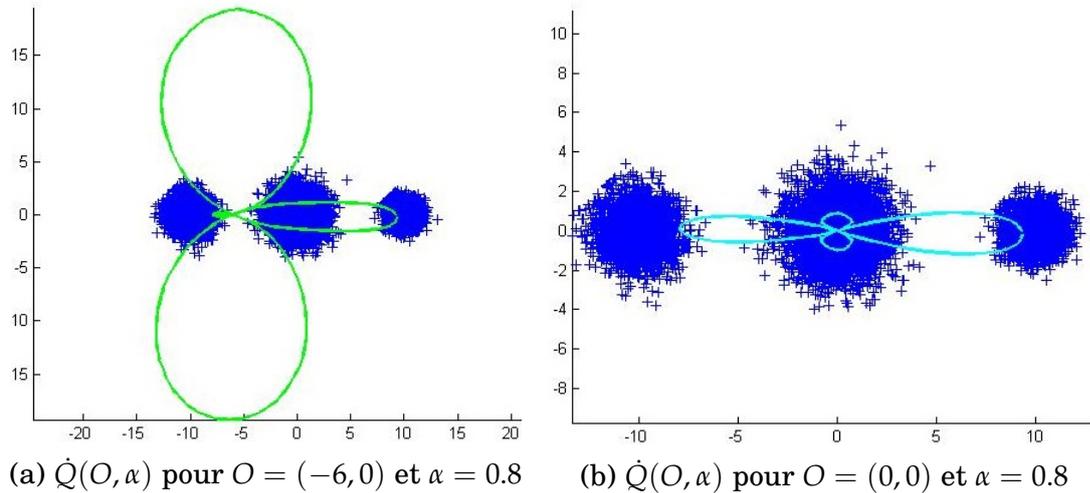


FIGURE 3.22 – Deux modes sur trois détectés.

Enfin les deux derniers graphiques (3.23a, 3.23b) représentent le cas où le point O est aligné avec les centres des modes de P_2 avec les trois modes d'un côté ou de l'autre.

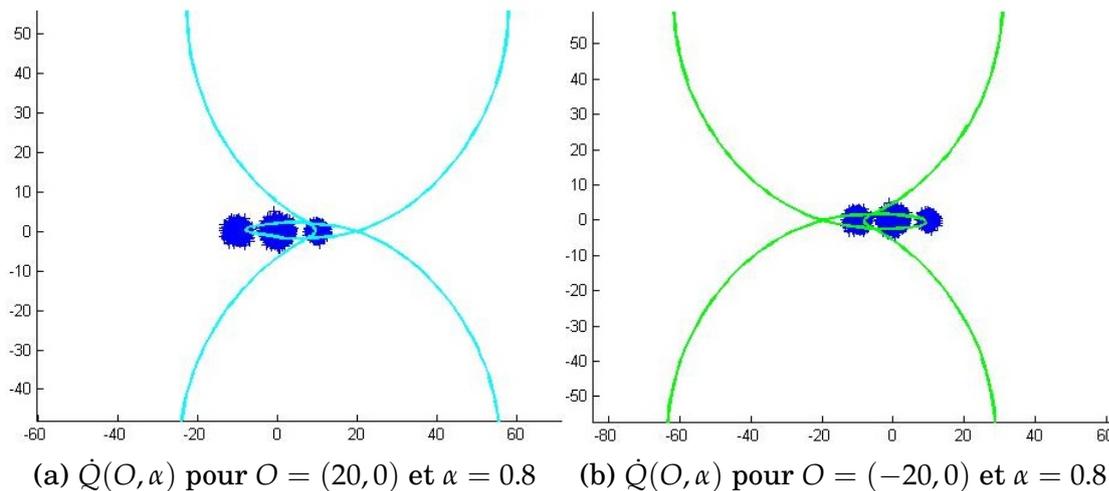


FIGURE 3.23 – Un seul mode sur trois détecté.

Cas II : trois modes « non-alignés » On pose

$$O_1 = (-10, 0), O_2 = (0, 10), O_3 = (10, 0)$$

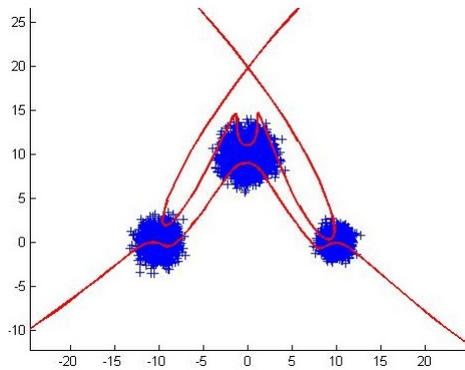
et

$$\Sigma_1 = I_2, \Sigma_2 = 3/2I_2, \Sigma_3 = 1/2I_2$$

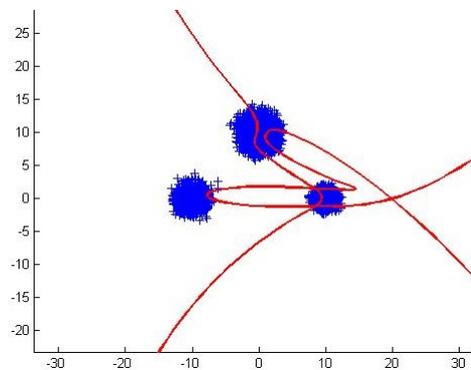
Dans ce cas, la forme de la surface quantile peut indiquer de façon « visuelle » deux situations seulement :

- détecter deux modes : dus à l'alignement de l'observateur O avec deux des trois modes (figure 3.24b).
- détecter trois modes (figure 3.24a).

Voici deux graphiques représentant chacune des situations citées



(a) Trois modes détectés :
 $\hat{Q}(O, \alpha)$ pour $O = (0, 20)$ et $\alpha = 0.8$



(b) Deux modes détectés :
 $\hat{Q}(O, \alpha)$ pour $O = (20, 0)$ et $\alpha = 0.8$

Chapitre 4

Utilisations des quantiles spatiaux

4.1 Introduction

Ce chapitre regroupe quelques applications statistiques des quantiles spatiaux et tout particulièrement les quantiles via les demi-espaces.

Dans la première section il est introduit une nouvelle notion appelée « champs de profondeur », un objet qui permet d'exploiter la dépendance des surfaces quantiles en l'observateur pour construire des champs de vecteurs permettant d'exhiber des propriétés sur la loi de probabilité. La seconde partie quant à elle, est consacrée d'abord à l'introduction et à l'étude d'une distance quantile entre deux lois de probabilité, puis à la construction des tests d'adéquation et de comparaison via la distance entre les surfaces quantile. Ensuite la troisième section s'intéresse aux surfaces quantiles médianes basées sur les quantiles via les demi-plans, dans cette partie il s'agit principalement de montrer par des simulations les formes des surfaces quantiles pour des lois aussi bien symétriques que non symétriques (dans le cas symétrique, ces surfaces correspondent bien au centre de symétrie). Enfin la dernière section est dédiée à l'utilisation de « l'information auxiliaire » (connaissance de la probabilité d'une partition finie) pour améliorer l'estimation des surfaces quantiles.

4.2 Champs de profondeur

Dans cette section on introduit une notion de « champs de profondeurs » ou « champs d'altitudes », à partir des surfaces quantile, en associant à

chaque point O de \mathbb{R}^d un vecteur $\vec{D}_\alpha(O)$ dont la direction pointe vers la masse principale et la norme quantifie la distance à la masse. On se place dans le cadre des surfaces quantiles par des demi-espaces. Les conditions (\mathbf{H}_0^+) , (\mathbf{H}_0^-) , (\mathbf{H}_1) , (\mathbf{H}_2) , (\mathbf{H}_3) et (\mathbf{H}'_3) font références aux hypothèses 2.3.1 du chapitre 2.

4.2.1 Définitions et théorèmes de convergence

Dans cette partie $d \geq 2$, le cas dimension 1, ne présente aucun intérêt. Sauf mention du contraire, dans toute la suite X_1, \dots, X_n désigne un échantillon sur \mathbb{R}^d de loi $P = \mathbb{P}_X$ et de mesure empirique associée $P_n = 1/n \sum_{i=1}^n \delta_{X_i}$

Définition 4.1. Soit $O \in \mathbb{R}^d$, $\alpha \in \Delta$, le vecteur de profondeur de O de niveau α est défini par

$$\vec{D}_\alpha(O) = \int_{S_{d-1}} (Q(O, u, \alpha) - O) du$$

Définition 4.2. Soit $O \in \mathbb{R}^d$, $\alpha \in \Delta$, le vecteur de profondeur empirique de O de niveau α est défini par

$$\vec{D}_\alpha^n(O) = \int_{S_{d-1}} (Q_n(O, u, \alpha) - O) du$$

Remarque 4.3. On a

$$\vec{D}_\alpha(O) = \int_{S_{d-1}} (Q(O, u, \alpha) - O) du = \int_{S_{d-1}} Q(O, u, \alpha) du - \text{Vol}(S_{d-1}) \cdot O$$

$$\vec{D}_\alpha^n(O) = \int_{S_{d-1}} (Q_n(O, u, \alpha) - O) du = \int_{S_{d-1}} Q_n(O, u, \alpha) du - \text{Vol}(S_{d-1}) \cdot O$$

Proposition 4.4. Soit $O \in \mathbb{R}^d$ et $\alpha \in \Delta$. Le processus défini par $\vec{D}_\alpha^n(O) - \vec{D}_\alpha(O)$ ne dépend pas de O .

Démonstration. Soit $O \in \mathbb{R}^d$ et $\alpha \in \Delta$. On a

$$\begin{aligned} \vec{D}_\alpha^n(O) - \vec{D}_\alpha(O) &= \int_{S_{d-1}} (Q_n(O, u, \alpha) - O) du - \int_{S_{d-1}} (Q(O, u, \alpha) - O) du \\ &= \int_{S_{d-1}} (Q_n(O, u, \alpha) - Q(O, u, \alpha)) du \end{aligned} \quad (4.1)$$

par la proposition 2.27 on sait que $Q(O, u, \alpha) - Q_n(O, u, \alpha)$ est indépendant du choix de O , donc $\vec{D}_\alpha^n(O) - \vec{D}_\alpha(O)$ ne dépend pas de O . \square

Théorème 4.5. *Sous les conditions (H_0^-) et (H_0^+) on a*

$$\lim_{n \rightarrow \infty} \sup_{\alpha \in \Delta} \left| \vec{D}_\alpha^n(O) - \vec{D}_\alpha(O) \right| = 0 \text{ p.s.}$$

Ce théorème est une conséquence immédiate du théorème de consistance uniforme pour les quantiles via des demi-plans.

Voici la définition du processus de profondeur empirique, qui par la proposition 4.4, ne dépend que du niveau $\alpha \in \Delta$.

Définition 4.6. *On définit le processus de profondeur empirique par*

$$\Theta_n^\alpha = \sqrt{n} \left(\vec{D}_\alpha^n(O) - \vec{D}_\alpha(O) \right), \quad \alpha \in \Delta.$$

Par l'égalité (4.1), on remarque que

$$\Theta_n^\alpha = \int_{\mathbb{S}_{d-1}} \sqrt{n} (Q_n(O, u, \alpha) - Q(O, u, \alpha)) du$$

Soit $\alpha \in \Delta$, on note $(\mathbf{G}_\alpha(u))_{u \in \mathbb{S}_{d-1}}$ le processus gaussien centré et de covariance

$$u, v \in \mathbb{S}_{d-1} \quad \text{cov}(\mathbf{G}_\alpha(u), \mathbf{G}_\alpha(v)) = \frac{P(H(u, \alpha) \cap H(v, \alpha)) - \alpha^2}{h_\alpha(u)h_\alpha(v)}$$

avec $h_\alpha(u) = f_{\langle X, u \rangle} \circ F_{\langle X, u \rangle}^{-1}(\alpha)$. Ensuite, on pose

$$\mathbb{B}_\alpha = \int_{\mathbb{S}_{d-1}} \mathbf{G}_\alpha(u) \cdot u du$$

et pour tout $i \in \llbracket 1, d \rrbracket$, on note \mathbb{B}_α^i la $i^{\text{ème}}$ coordonnée de \mathbb{B}_α définie comme suit

$$\mathbb{B}_\alpha^i = \int_{(u_1, u_2, \dots, u_d) \in \mathbb{S}_{d-1}} \mathbf{G}_\alpha(u_1, u_2, \dots, u_d) \cdot u_i du_1 du_2 \dots du_d$$

\mathbb{B}_α est centré de covariance pour tout $i, j \in \llbracket 1, d \rrbracket$

$$\text{cov}(\mathbb{B}_{\alpha'}^i, \mathbb{B}_\alpha^j) = \int_{\mathbb{S}_{d-1}} \int_{\mathbb{S}_{d-1}} \mathbb{E}(\mathbf{G}_\alpha(u) \mathbf{G}_{\alpha'}(v)) \cdot u_i v_j du_1 du_2 \dots du_d dv_1 dv_2 \dots dv_d.$$

Théorème 4.7. *Sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) le processus $(\Theta_n^\alpha)_{\alpha \in \Delta}$ converge faiblement vers le processus gaussien centré $(\mathbb{B}_\alpha)_{\alpha \in \Delta}$ de matrice de covariance pour $\alpha, \alpha' \in \Delta$*

$$\left(\text{cov}(\mathbb{B}_{\alpha'}^i, \mathbb{B}_{\alpha'}^j) \right)_{i, j \in \llbracket 1, d \rrbracket} = \int_{\mathbb{S}_{d-1}} \int_{\mathbb{S}_{d-1}} \mathbb{E}(\mathbf{G}_\alpha(u) \mathbf{G}_{\alpha'}(v)) \cdot u_i v_j dudv.$$

Démonstration. Le principe d'invariance fort uniforme (cf. théorème 2.43) implique que sous (\mathbf{H}_1) et (\mathbf{H}_2) , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite \mathbf{G}^n de versions de \mathbf{G} de sorte que

$$\sqrt{n}(Q_n(O, u, \alpha) - Q(O, u, \alpha)) = \mathbf{G}_\alpha^n(u) \cdot u + \mathbf{Z}_n(u, \alpha) \cdot u, \quad O \in \mathbb{R}^d, \alpha \in \Delta, u \in \mathbb{S}_{d-1}$$

où $\mathbf{Z}_n = \sqrt{n}(Y_n - Y) - \mathbf{G}^n$ vérifie

$$\lim_{n \rightarrow \infty} \|\mathbf{Z}_n\|_{\mathbb{S}_{d-1} \times \Delta} = \lim_{n \rightarrow \infty} \sup_{u \in \mathbb{S}_{d-1}} \sup_{\alpha \in \Delta} |\mathbf{Z}_n(u, \alpha)| = 0 \quad p.s.$$

Par suite

$$\begin{aligned} \Theta_n^\alpha &= \int_{\mathbb{S}_{d-1}} \sqrt{n} (Q_n(O, u, \alpha) - Q(O, u, \alpha)) du \\ &= \int_{\mathbb{S}_{d-1}} \mathbf{G}_\alpha^n(u) \cdot u du + \int_{\mathbb{S}_{d-1}} \mathbf{Z}_n(u, \alpha) \cdot u du. \end{aligned}$$

Notons que $\int_{\mathbb{S}_{d-1}} \mathbf{G}_\alpha^n(u) \cdot u du \stackrel{\mathcal{L}}{\sim} \mathbb{B}_\alpha$, donc en posant

$$\mathbb{B}_{n, \alpha} := \int_{\mathbb{S}_{d-1}} \mathbf{G}_\alpha^n(u) \cdot u du$$

et

$$\kappa_\alpha^n := \int_{\mathbb{S}_{d-1}} \mathbf{Z}_n(u, \alpha) \cdot u du$$

on obtient

$$\Theta_n^\alpha = \mathbb{B}_{n, \alpha} + \kappa_\alpha^n, \quad \alpha \in \Delta, u \in \mathbb{S}_{d-1}$$

avec $\lim_{n \rightarrow \infty} \|\kappa_n\|_\Delta = 0 \quad p.s.$ donc le processus $(\Theta_n^\alpha)_{\alpha \in \Delta}$ converge faiblement vers le processus gaussien centré $(\mathbb{B}_\alpha)_{\alpha \in \Delta}$ de covariance, pour tout $i, j \in \llbracket 1, d \rrbracket, \alpha, \alpha' \in \Delta$

$$cov(\mathbb{B}_\alpha^i, \mathbb{B}_{\alpha'}^j) = \int_{\mathbb{S}_{d-1}} \int_{\mathbb{S}_{d-1}} \mathbb{E}(\mathbf{G}_\alpha(u) \mathbf{G}_{\alpha'}(v)) \cdot u_i v_j \, dudv.$$

□

La covariance limite du champ de profondeur est donnée par

$$\begin{aligned} cov(\mathbb{B}_\alpha^i, \mathbb{B}_{\alpha'}^j) &= \int_{u \in \mathbb{S}_{d-1}} \int_{v \in \mathbb{S}_{d-1}} \mathbb{E}(\mathbf{G}_\alpha(u) \mathbf{G}_{\alpha'}(v)) \cdot u_i v_j \, dudv \\ &= \int_{u \in \mathbb{S}_{d-1}} \int_{v \in \mathbb{S}_{d-1}} \frac{P(H(u, \alpha) \cap H(v, \alpha')) - \alpha \alpha'}{h_\alpha(u) h_{\alpha'}(v)} \cdot u_i v_j \, dudv \\ &= \int_{u \in \mathbb{S}_{d-1}} \int_{v \in \mathbb{S}_{d-1}} \frac{P(H(u, \alpha) \cap H(v, \alpha')) - \alpha \alpha'}{h_\alpha h_{\alpha'}} \cdot u_i v_j \, dudv. \end{aligned}$$

Remarque 4.8. *Avant de s'intéresser au cas des lois symétriques, voici quelques remarques en liens avec les fonctions de profondeur.*

- *En contraste avec les fonctions de profondeurs rappelées dans le Chapitre 1, aucune notion de centre n'est exigée par avance.*
- *Sous (\mathbf{H}_0) et (\mathbf{H}_0^\dagger) le champs de profondeur est bien défini, aucune hypothèse forte n'est exigée.*
- *Les propriétés désirables pour les champs de profondeurs similaires à celles de la définition 1.26 peuvent être adaptées, en effet les champs de profondeurs présentent la propriété « d'invariance affine », le vecteur d'altitude est « minimal autour du centre de la masse » et croît en norme lorsque O s'éloigne de la masse, puis enfin il est de norme infinie quand $\|O\|$ tends vers l'infini.*
- *On est convaincu que les champs d'altitudes par des quantiles généralisés du Chapitre 3, sont beaucoup plus informatifs que ceux obtenus par des demi-espaces car les surfaces quantiles très dépendantes de chaque point O du champs. Néanmoins pour le moment on n'a pas fourni de méthodologie statistique exploitant l'information contenue dans les champs d'altitudes généralisés.*

4.2.2 Champs de profondeur pour les lois symétriques

Soit $P = \mathbb{P}_X$ une loi de probabilité admettant une symétrie autour d'un point $M \in \mathbb{R}^d$. En d'autres termes, pour tout $u, v \in \mathbb{S}_{d-1}$ on a

$$F_{\langle X-M, u \rangle} = F_{\langle X-M, v \rangle}.$$

On en déduit que pour tout $\alpha \in \Delta$ que

$$Y(M, u, \alpha) = Y(M, v, \alpha)$$

De plus comme pour tout $O, O' \in \mathbb{R}^d$ on a

$$Y(O, u, \alpha) = Y(O', u, \alpha) + \langle O' - O, u \rangle$$

alors pour tout $O \in \mathbb{R}^d$ on obtient que

$$Y(M, u, \alpha) = Y(O, v, \alpha) + \langle O - M, v - u \rangle \quad \forall u, v \in \mathbb{S}_{d-1}.$$

Ainsi que $(u, \alpha) \rightarrow h(u, \alpha)$ ne dépend pas de u . On note $h_\alpha := h(u, \alpha)$. Soit \mathbb{B}_α le processus centré défini dans la section précédente. La covariance de \mathbb{B}_α devient pour tout $i, j \in \llbracket 1, d \rrbracket$

$$\begin{aligned}
\text{cov}(\mathbb{B}_\alpha^i, \mathbb{B}_\alpha^j) &= \int_{u \in \mathbb{S}_{d-1}} \int_{v \in \mathbb{S}_{d-1}} \frac{P(H(u, \alpha) \cap H(v, \alpha)) - \alpha^2}{h_\alpha h_\alpha} \cdot u_i v_j \, dudv \\
&= \frac{1}{h_\alpha^2} \int_{u \in \mathbb{S}_{d-1}} \int_{v \in \mathbb{S}_{d-1}} P(H(u, \alpha) \cap H(v, \alpha)) \cdot u_i v_j \, dudv \\
&\quad - \frac{\alpha^2}{h_\alpha^2} \int_{u \in \mathbb{S}_{d-1}} u_i \, du \times \int_{v \in \mathbb{S}_{d-1}} v_j \, dv
\end{aligned}$$

et, en remarquant que $\int_{u \in \mathbb{S}_{d-1}} u_i \, du_1 du_2 \cdots du_d = 0$ comme intégrale d'une fonction impaire sur un domaine symétrique, la covariance se simplifie en

$$\forall i, j \in \llbracket 1, d \rrbracket, \quad \text{cov}(\mathbb{B}_\alpha^i, \mathbb{B}_\alpha^j) = \frac{1}{h_\alpha^2} \int_{u \in \mathbb{S}_{d-1}} \int_{v \in \mathbb{S}_{d-1}} P(H(u, \alpha) \cap H(v, \alpha)) \cdot u_i v_j \, dudv$$

Corollaire 4.9. Soit $P = \mathbb{P}_X$ une loi de probabilité admettant une symétrie autour d'un point M et $\alpha \in \Delta$. Sous la condition (\mathbf{H}_1) le vecteur Θ_n^α converge faiblement vers le vecteur gaussien centré \mathbb{B}_α de covariance

$$\forall i, j \in \llbracket 1, d \rrbracket, \quad \text{cov}(\mathbb{B}_\alpha^i, \mathbb{B}_\alpha^j) = \frac{1}{h_\alpha^2} \int_{u \in \mathbb{S}_{d-1}} \int_{v \in \mathbb{S}_{d-1}} P(H(u, \alpha) \cap H(v, \alpha)) \cdot u_i v_j \, dudv$$

Dans ce cas particulier, le vecteur de profondeur présente la particularité de toujours pointer vers le point de symétrie M

Proposition 4.10. Soit $P = \mathbb{P}_X$ une loi de probabilité admettant une symétrie autour d'un point M et $\alpha \in \Delta$. Il existe $k_d > 0$ tel que pour tout $O \in \mathbb{R}^d$ on a

$$\vec{D}_\alpha(O) = k_d(M - O)$$

Démonstration. On rappelle que $d \geq 2$ et on note $M = (x_1^M, \dots, x_d^M)$.

Comme remarqué précédemment pour tout $O = (x_1^O, \dots, x_d^O) \in \mathbb{R}^d$, $u = (u_1, \dots, u_d) \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$, on a

$$Y(O, u, \alpha) = Y(M, u, \alpha) + \langle M - O, u \rangle$$

donc

$$\begin{aligned}
\vec{D}_\alpha(O) &= \int_{\mathbb{S}_{d-1}} (Q(O, u, \alpha) - O) \, du \\
&= \int_{\mathbb{S}_{d-1}} Y(O, u, \alpha) u \, du \\
&= \int_{\mathbb{S}_{d-1}} Y(M, u, \alpha) u \, du + \int_{\mathbb{S}_{d-1}} \langle M - O, u \rangle u \, du
\end{aligned}$$

or pour tout $u, v \in \mathbb{S}_{d-1}$ on a $Y(M, u, \alpha) = Y(M, v, \alpha) := y_\alpha^M$ donc

$$\int_{\mathbb{S}_{d-1}} Y(M, u, \alpha) u \, du = y_\alpha^M \int_{\mathbb{S}_{d-1}} u \, du = 0.$$

Par suite

$$\begin{aligned} \vec{D}_\alpha(O) &= \int_{\mathbb{S}_{d-1}} \langle M - O, u \rangle u \, du \\ &= \left(\int_{\mathbb{S}_{d-1}} \langle M - O, u \rangle u_1 \, du, \dots, \int_{\mathbb{S}_{d-1}} \langle M - O, u \rangle u_d \, du \right) \end{aligned} \quad (4.2)$$

notons que

$$\int_{\mathbb{S}_{d-1}} \langle M - O, u \rangle u_1 \, du = \int_{\mathbb{S}_{d-1}} \langle M, u \rangle u_1 \, du - \int_{\mathbb{S}_{d-1}} \langle O, u \rangle u_1 \, du$$

avec

$$\begin{aligned} \int_{\mathbb{S}_{d-1}} \langle M, u \rangle u_1 \, du &= \int_{\mathbb{S}_{d-1}} x_1^M u_1^2 + \dots + x_d^M u_d u_1 \, du \\ &= \int_{\mathbb{S}_{d-1}} x_1^M u_1^2 \, du + \int_{\mathbb{S}_{d-1}} x_2^M u_2 u_1 + \dots + x_d^M u_d u_1 \, du \end{aligned}$$

d'autre part, on a

$$\begin{aligned} \int_{\mathbb{S}_{d-1}} x_1^M u_1^2 \, du &= x_1^M \int_{-1}^1 u_1^2 \left(\int_{-1}^1 \dots \int_{-1}^1 \mathbb{1}_{\{u_2^2 + \dots + u_d^2 \leq 1 - u_1^2\}} \, du_2 \dots du_d \right) du_1 \\ &= x_1^M \int_{-1}^1 u_1^2 (\text{vol}(\mathbb{S}^{d-2}) (1 - u_1^2)^{(d-2)/2}) du_1 \\ &= \text{vol}(\mathbb{S}^{d-2}) x_1^M \int_{-1}^1 u_1^2 (1 - u_1^2)^{(d-2)/2} du_1 \\ &= \text{vol}(\mathbb{S}^{d-2}) C_d x_1^M \end{aligned} \quad (4.3)$$

avec $C_d = \int_{-1}^1 s^2 (1 - s^2)^{(d-2)/2} ds$ une constante réelle et strictement positive.

D'un autre côté, on a

$$\int_{\mathbb{S}_{d-1}} x_2^M u_2 u_1 + \dots + x_d^M u_d u_1 \, du = \int_{-1}^1 u_1 \left(\int_{-1}^1 \dots \int_{-1}^1 \sum_{i=2}^d x_i^M u_i \mathbb{1}_{\{u \in \mathbb{S}_{d-1}\}} \, du_2 \dots du_d \right) du_1$$

or,

$$\begin{aligned} \int_{-1}^1 \cdots \int_{-1}^1 \sum_{i=2}^d x_i^M u_i \mathbb{1}_{\{u \in \mathbb{S}_{d-1}\}} du_2 \cdots du_d &= \sum_{i=2}^d x_i^M \int_{-1}^1 \cdots \int_{-1}^1 u_i \mathbb{1}_{\{u \in \mathbb{S}_{d-1}\}} du_2 \cdots du_d \\ &= \sum_{i=2}^d x_i^M \times 0 \\ &= 0 \end{aligned}$$

donc

$$\int_{\mathbb{S}_{d-1}} x_2^M u_2 u_1 + \cdots + x_d^M u_d u_1 du = 0. \quad (4.4)$$

Les égalités (4.3) et (4.4) implique que

$$\int_{\mathbb{S}_{d-1}} \langle M, u \rangle u_1 du = \text{vol}(\mathbb{S}^{d-2}) C_d x_1^M.$$

Par analogie, on obtient que pour tout $i \in \llbracket 1, d \rrbracket$

$$\int_{\mathbb{S}_{d-1}} \langle M - O, u \rangle u_i du = \text{vol}(\mathbb{S}^{d-2}) C_d (x_i^M - x_i^O)$$

donc

$$\int_{\mathbb{S}_{d-1}} \langle M - O, u \rangle u du = \text{vol}(\mathbb{S}^{d-2}) C_d (M - O)$$

Conclusion : Par l'égalité (4.2) on a

$$\vec{D}_\alpha(O) = C_d \times \text{vol}(\mathbb{S}^{d-2})(M - O)$$

avec $C_d = \int_{-1}^1 s^2(1-s^2)ds$ une constante réelle et strictement positive. \square

Remarque 4.11. Grâce à la proposition précédente, on peut immédiatement affirmer les points suivants :

- on a $\vec{D}_\alpha(M) = \vec{0}$
- pour $O, O' \in \mathbb{R}^d$, si $\|O' - M\| > \|O - M\|$ alors $\|\vec{D}_\alpha(O')\| > \|\vec{D}_\alpha(O)\|$.

4.2.3 Simulations

On se place dans le cadre de $d = 2$. Dans cette section, on présente les champs de profondeur de trois types de loi : les lois symétriques, les lois multimodales et les lois avec support de dimension inférieure.

Cas symétrique loi gaussienne $N(0_2, \Sigma)$

Pour commencer on s'intéresse à la loi gaussienne $P = N(0_2, I_2)$, les simulations sont faites avec un échantillon de $N = 10^3$. Dans ce cas la loi est symétrique autour de son espérance. Ainsi, par la proposition 4.10, on sait que tous les vecteurs « pointent » vers $\mathbb{E}(X) = 0_2$. Dans cet exemple, on considère une grille E avec 441 observateurs $O = (x, y)$ suivante

- $x = -2 + k \cdot 0.2$ pour $0 \leq k \leq 20$
- $y = -2 + k \cdot 0.2$ pour $0 \leq k \leq 20$

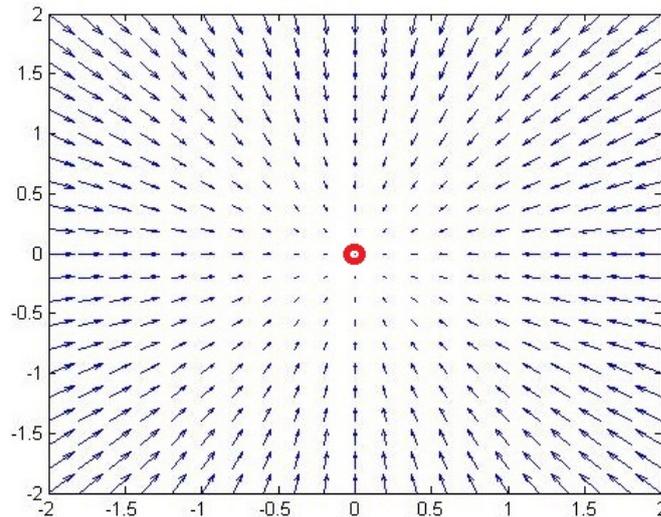


FIGURE 4.1 – Champ de vecteurs $\vec{D}_\alpha(O)$ de $N(0_2, I_2)$ pour $\alpha = 0.6$ et $O \in E$

Avec la modification suivante de la matrice de variance-covariance, l'espérance reste un bon candidat pour la médiane, ainsi comme précédemment tous les vecteurs de profondeur « pointent » vers $\mathbb{E}(X) = 0_2$

$$P = N(0_2, \Sigma) \text{ avec } \Sigma = \begin{pmatrix} 4 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

Dans cet exemple, on se fixe la grille E avec 441 observateurs $O = (x, y)$ suivante

- $x = -7 + k \cdot 0.7$ pour $0 \leq k \leq 20$
- $y = -3 + k \cdot 0.3$ pour $0 \leq k \leq 20$

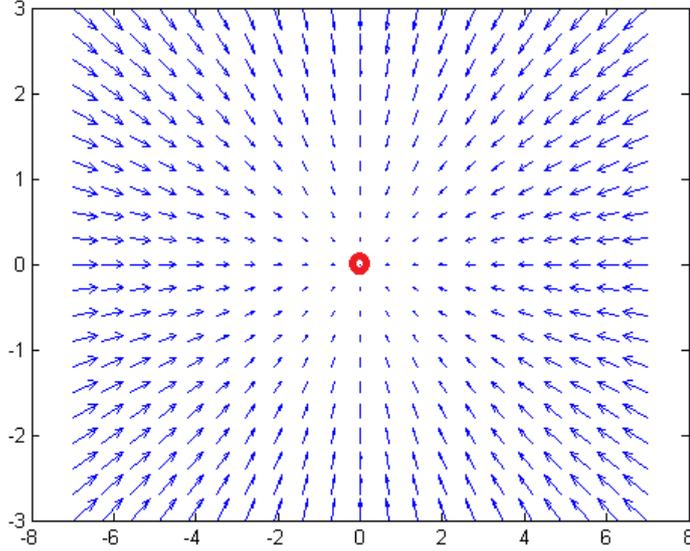


FIGURE 4.2 – Champ de vecteurs $\vec{D}_\alpha(O)$ de $N(O_2, \Sigma)$ pour $\alpha = 0.7$ et $O \in E$

De manière générale, dans le cas symétrique le niveau $\alpha \in \Delta$ n'a pas d'influence sur la direction des vecteurs, cette propriété ne sera plus vraie pour des lois asymétriques.

Lois Multimodales : Mélange Gaussien

Dans cette partie, on va s'intéresser aux champs de vecteurs de profondeur pour des lois à plusieurs modes. On commence par une loi bimodale, comme mélange de deux gaussiennes

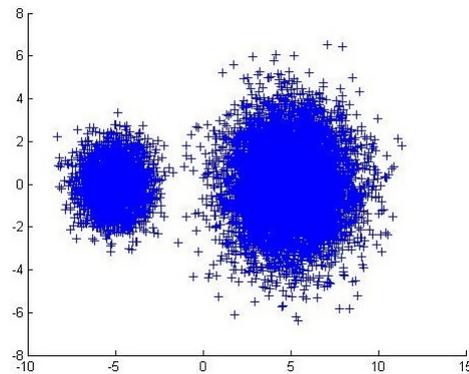
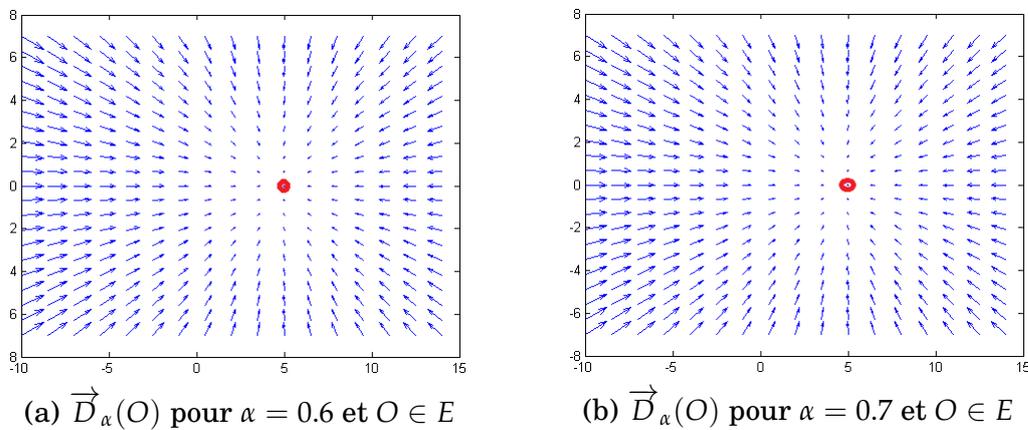
$$P = \frac{1}{4}N(O_1, \Sigma_1) + \frac{3}{4}N(O_2, \Sigma_2).$$

On pose $O_1 = (-5, 0)$, $O_2 = (5, 0)$, $\Sigma_1 = I_2$ et $\Sigma_2 = 3I_2$.

Dans cet exemple, on se fixe la grille E avec 357 observateurs $O = (x, y)$ suivante

- $x = -10 + k \cdot 1.5$ pour $0 \leq k \leq 20$
- $y = -7 + k \cdot 0.7$ pour $0 \leq k \leq 20$

On commence par s'intéresser à $\vec{D}_\alpha(O)$ pour $O \in E$ et $\alpha \in (1/2, 3/4)$. On remarque que tous les vecteurs $\vec{D}_\alpha(O)$ pointent vers le mode $O_2 = (5, 0)$ centre de la zone modale de masse $3/4$ (figure 4.4a, 4.4b).

FIGURE 4.3 – $1/4N(O_1, \Sigma_1) + 3/4N(O_2, \Sigma_2)$ FIGURE 4.4 – Champ de vecteurs $\vec{D}_\alpha(O)$ pour $\alpha \in \{0.6, 0.7\}$ et $O \in E$

Pour $\vec{D}_\alpha(O)$ avec $O \in E$ et $\alpha > 3/4$ un changement de direction survient, les vecteurs $\vec{D}_\alpha(O)$ pointent vers une zone autour de $(1, 0)$, ceci est dû au fait que la zone modale de masse $3/4$ ne suffit plus et qu'il faille chercher de la masse dans la zone modale de probabilité $1/4$ (figure 4.5a, 4.5b).

Lois à support spiral

Dans cette partie, on s'intéresse aux champ de vecteurs de profondeur pour une loi sans densité par rapport à la mesure de Lebesgue de \mathbb{R}^2 . on pose

$$X = (r \cos(\theta), r \sin(\theta))$$

avec $\theta \sim \mathcal{U}([0, 8\pi])/4$ et $r = 1/4 + \theta$.

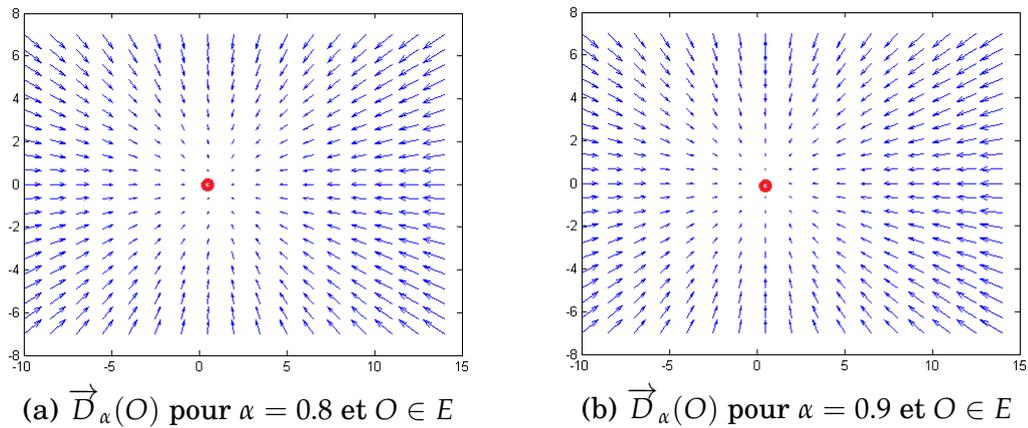


FIGURE 4.5 – Champ de vecteurs $\vec{D}_\alpha(O)$ pour $\alpha \in \{0.8, 0.9\}$ et $O \in E$

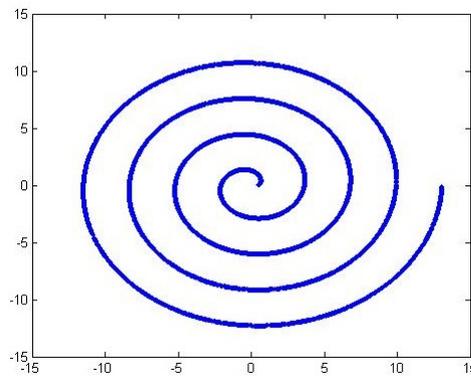
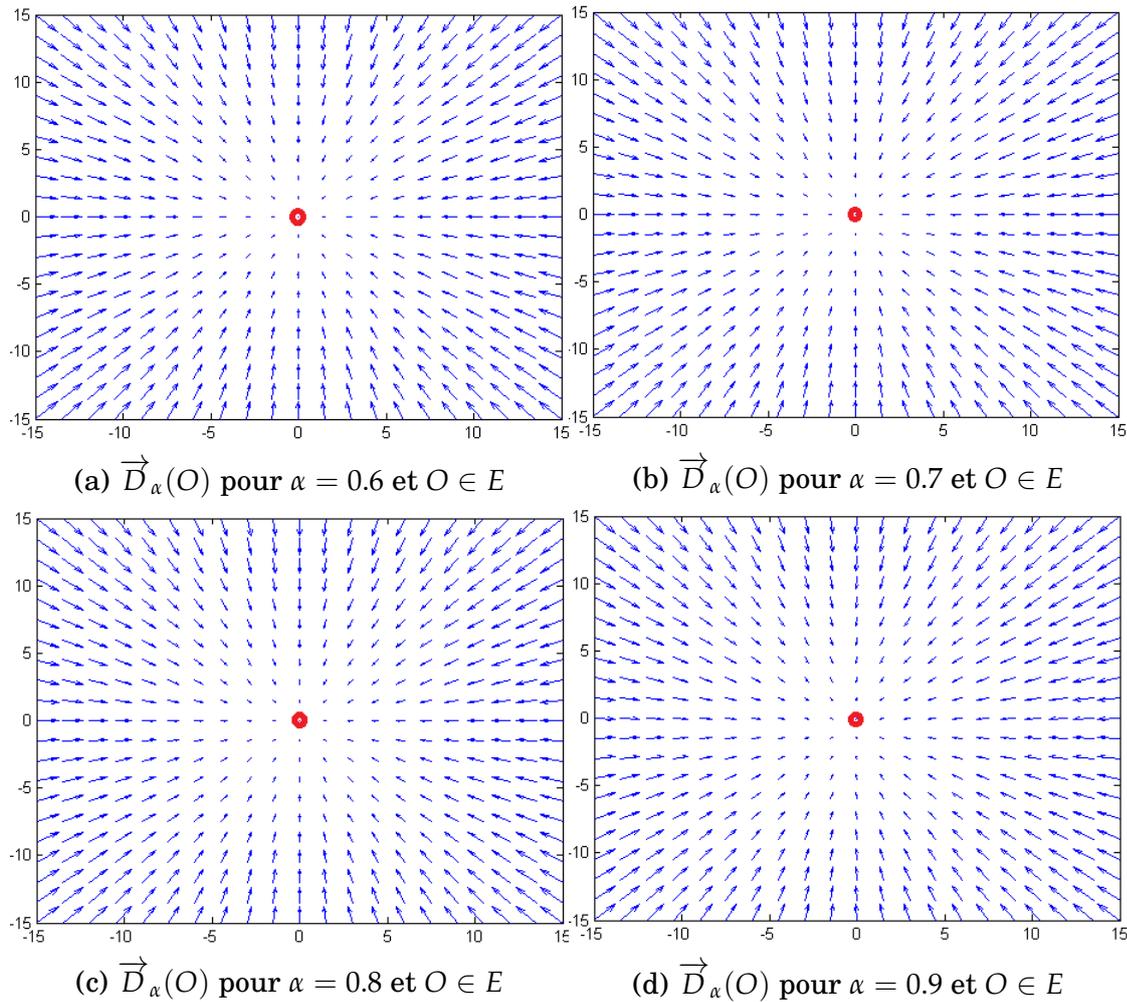


FIGURE 4.6 – Échantillon de la loi P avec $N = 10000$

Dans cet exemple, on se fixe la grille E avec 441 observateurs $O = (x, y)$ suivante

- $x = -15 + k \cdot 1.5$ pour $0 \leq k \leq 20$
- $y = -15 + k \cdot 1.5$ pour $0 \leq k \leq 20$

Bien que la loi ne soit pas symétrique, tous les vecteurs $\vec{D}_\alpha(O)$ pointent « autour » de $(0,0)$ quel que soit $\alpha \in \Delta$. voici les champs de vecteurs $\vec{D}_\alpha(O)$ pour $O \in E$ et $\alpha \in \{0.6, 0.7, 0.8, 0.9\}$.



4.3 Distance quantile

Dans cette section il s'agit d'introduire puis d'étudier une distance entre deux lois de probabilité P et Q à travers leurs surfaces de quantile associées à niveau fixé α . Comme première application de cette distance, on propose un test d'adéquation puis un test de comparaison.

4.3.1 Définitions et théorème

On note \mathcal{P}_d l'ensemble des mesure de probabilité sur \mathbb{R}^d .

Définition 4.12. Soit P et Q deux lois de probabilité et X une variable aléatoire de loi P et Z une variable aléatoire de loi Q . Pour tout $O \in \mathbb{R}^d$ et tout

$\alpha \in [1/2, 1)$, on définit la semi-distance quantile $d_\alpha : \mathcal{P}_d \times \mathcal{P}_d \rightarrow \mathbb{R}^+$ par

$$d_\alpha(P, Q) = \int_{\mathbb{S}_{d-1}} \left| F_{\langle X-O, u \rangle} \circ F_{\langle Z-O, u \rangle}^{-1}(\alpha) - \alpha \right| du + \int_{\mathbb{S}_{d-1}} \left| F_{\langle Z-O, u \rangle} \circ F_{\langle X-O, u \rangle}^{-1}(\alpha) - \alpha \right| du$$

Remarque 4.13. Avec la notation $Y^P(O, u, \alpha)$ le quantile de la loi P vu de $O \in \mathbb{R}^d$ dans la direction $u \in \mathbb{S}_{d-1}$ et de niveau $\alpha \in \Delta$, la semi-distance quantile peut être écrite comme suit

$$d_\alpha(P, Q) = \int_{\mathbb{S}_{d-1}} \left| F_{\langle X-O, u \rangle}(Y^Q(O, u, \alpha)) - \alpha \right| du + \int_{\mathbb{S}_{d-1}} \left| F_{\langle Y-O, u \rangle}(Y^P(O, u, \alpha)) - \alpha \right| du$$

puis l'écriture en terme de demi-espace suivante

$$d_\alpha(P, Q) = \int_{\mathbb{S}_{d-1}} \left| P(H(O, u, Y^Q(O, u, \alpha))) - \alpha \right| du + \int_{\mathbb{S}_{d-1}} \left| Q(H(O, u, Y^P(O, u, \alpha))) - \alpha \right| du$$

ainsi par la définition 2.4, on a

$$H(O, u, Y^P(O, u, \alpha)) := H^P(u, \alpha)$$

et

$$H(O, u, Y^Q(O, u, \alpha)) := H^Q(u, \alpha)$$

donc

$$d_\alpha(P, Q) = \int_{\mathbb{S}_{d-1}} \left| P(H^Q(u, \alpha)) - \alpha \right| du + \int_{\mathbb{S}_{d-1}} \left| Q(H^P(u, \alpha)) - \alpha \right| du.$$

Grâce à cette dernière écriture on remarque que d_α ne dépend pas du choix de $O \in \mathbb{R}^d$.

- Ensuite, pour tout $\alpha \in \Delta$, l'application d_α est bien une semi-distance car
- pour tout $P, Q \in \mathcal{P}_d$ on a $d_\alpha(P, Q) = d_\alpha(Q, P) \geq 0$
 - si $P = Q$ alors

$$\begin{aligned} d_\alpha(P, Q) &= \int_{\mathbb{S}_{d-1}} \left| F_{\langle X, u \rangle} \circ F_{\langle X, u \rangle}^{-1}(\alpha) - \alpha \right| du + \int_{\mathbb{S}_{d-1}} \left| F_{\langle X, u \rangle} \circ F_{\langle X, u \rangle}^{-1}(\alpha) - \alpha \right| du \\ &= \int_{\mathbb{S}_{d-1}} |\alpha - \alpha| du + \int_{\mathbb{S}_{d-1}} |\alpha - \alpha| du \\ &= 0 \end{aligned}$$

Lemme 4.14. Soient P, Q deux mesures de probabilité vérifiant les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) . Pour $\alpha \in \Delta$ on a

$$\lim_{n \rightarrow \infty} d_\alpha(P_n, Q) = d(P, Q) \quad p.s.$$

Démonstration.

$$d_\alpha(P_n, Q) = \int_{\mathcal{S}_{d-1}} |P_n(H^Q(u, \alpha)) - \alpha| du + \int_{\mathcal{S}_{d-1}} |Q(H_n^P(u, \alpha)) - \alpha| du$$

l'ensemble des demi-espaces forme une classe VC donc

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathcal{S}_{d-1}} |P_n(H^Q(u, \alpha)) - P(H^Q(u, \alpha))| = 0.$$

ainsi

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathcal{S}_{d-1}} \left| |P_n(H^Q(u, \alpha)) - \alpha| - |P(H^Q(u, \alpha)) - \alpha| \right| = 0$$

ensuite sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) on a

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathcal{S}_{d-1}} |Q(H_n^P(u, \alpha)) - Q(H^P(u, \alpha))| = 0$$

et

$$\lim_{n \rightarrow \infty} \sup_{u \in \mathcal{S}_{d-1}} \left| |Q(H_n^P(u, \alpha)) - \alpha| - |Q(H^P(u, \alpha)) - \alpha| \right| = 0.$$

Finalement

$$\lim_{n \rightarrow \infty} d_\alpha(P_n, Q) = d(P, Q) \text{ p.s.}$$

□

Dans la proposition suivante on s'intéresse à la convergence en demi-distance quantile de la mesure empirique P_n vers la mesure P .

Corollaire 4.15. *Sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) , pour tout $\alpha \in \Delta$ on a*

$$\lim_{n \rightarrow \infty} d_\alpha(P_n, P) = 0 \text{ p.s}$$

Démonstration. Sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) on a

$$\lim_{n \rightarrow \infty} d_\alpha(P_n, P) = d(P, P) \text{ p.s}$$

donc

$$\lim_{n \rightarrow \infty} d_\alpha(P_n, P) = 0 \text{ p.s.}$$

□

Voici, le théorème limite central pour le semi-distance d_α

Théorème 4.16. *Sous les conditions (\mathbf{H}_1) , (\mathbf{H}_2) et (\mathbf{H}'_3) pour tout $\alpha \in \Delta$ on a*

$$\sqrt{n} d_\alpha(P_n, P) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} 2 \int_{\mathbb{S}_{d-1}} |\mathbf{G}(u, \alpha)| du$$

où \mathbf{G} est un processus gaussien centré de covariance donnée par

$$\forall H, H' \in \mathcal{H}, \quad \text{cov}(\mathbf{G}(H), \mathbf{G}(H')) = P(H \cap H') - P(H)P(H')$$

Démonstration. Soit $\alpha \in \Delta$, on a

$$d_\alpha(P_n, P) = \int_{\mathbb{S}_{d-1}} |P_n(H(u, \alpha)) - \alpha| du + \int_{\mathbb{S}_{d-1}} |P(H_n(u, \alpha)) - \alpha| du$$

Étape 1. Sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , par approximation forte du processus empirique indexé par les hyperplan, on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite \mathbf{G}_n de versions de \mathbf{G} de sorte que

$$P_n(H(u, \alpha)) = P(H(u, \alpha)) + \frac{\mathbf{G}_n(u, \alpha)}{\sqrt{n}} + \frac{\mathbf{Z}_n(u, \alpha)}{\sqrt{n}}$$

avec $\lim_{n \rightarrow \infty} \|\mathbf{Z}_n\|_{\mathbb{S}_{d-1} \times \Delta} = 0$ p.s. donc

$$\sqrt{n} \int_{\mathbb{S}_{d-1}} |P_n(H(u, \alpha)) - \alpha| du = \int_{\mathbb{S}_{d-1}} |\mathbf{G}_n(u, \alpha)| du + \text{vol}(\mathbb{S}_{d-1}) \times o_{p.s.}(1)$$

Étape 2. On sait que

$$\sqrt{n} \int_{\mathbb{S}_{d-1}} |P(H_n(u, \alpha)) - \alpha| du = \sqrt{n} \int_{\mathbb{S}_{d-1}} \left| F_{\langle X, u \rangle}(Y_n(O, u, \alpha)) - F_{\langle X, u \rangle}(Y(O, u, \alpha)) \right| du$$

sous la condition (\mathbf{H}_1) par la loi du logarithme itéré du Théorème 2.38, il existe $C > 0$ et $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$\sup_{u \in \mathbb{S}_{d-1}} |Y_n(O, u, \alpha) - Y(O, u, \alpha)| \leq C \sqrt{\frac{\log \log n}{n}},$$

donc

$$\begin{aligned} \sqrt{n} \left(F_{\langle X-O, u \rangle}(Y_n(O, u, \alpha)) - F_{\langle X-O, u \rangle}(Y(O, u, \alpha)) \right) &= \sqrt{n} h_\alpha(u) (Y_n(O, u, \alpha) - Y(O, u, \alpha)) \\ &\quad + \sqrt{n} \rho \left(C \sqrt{\frac{\log \log n}{n}} \right) \end{aligned}$$

Sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) on a par le Théorème 2.43 et la Remarque 2.62,

$$Y_n(O, u, \alpha) = Y(O, u, \alpha) - \frac{\mathbf{G}_n(H(u, \alpha))}{\sqrt{n} h_\alpha(u)} + \frac{\mathbf{Z}'_n(u, \alpha)}{\sqrt{n} h_\alpha(u)}$$

avec $\lim_{n \rightarrow \infty} \|\mathbf{Z}'_n\|_{\mathcal{S}_{d-1} \times \Delta} = 0$ p.s. De plus, sous (\mathbf{H}_2) on a

$$\lim_{n \rightarrow \infty} \sqrt{n} \rho \left(C \sqrt{\frac{\log \log n}{n}} \right) = 0$$

donc

$$\sqrt{n} \int_{\mathcal{S}_{d-1}} |P(H_n(u, \alpha)) - \alpha| du = \int_{\mathcal{S}_{d-1}} |\mathbf{G}_n(u, \alpha)| du + \text{vol}(\mathcal{S}_{d-1}) \times o_{p.s.}(1)$$

Conclusion : sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , il existe $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$d_\alpha(P_n, P) = 2 \int_{\mathcal{S}_{d-1}} |\mathbf{G}_n(u, \alpha)| du + o_{p.s.}(1)$$

En rappelant que $\mathbf{G}_n \stackrel{\mathcal{L}oi}{\sim} \mathbf{G}$, on conclut que

$$\sqrt{n} d_\alpha(P_n, P) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} 2 \int_{\mathcal{S}_{d-1}} |\mathbf{G}(u, \alpha)| du$$

□

Remarque 4.17. Soit $P = \mathbb{P}_X$ et $Q = \mathbb{P}_{X'}$ deux mesures de probabilité de fonctions de répartition respectives F et G vérifiant la condition (\mathbf{H}_1) . Si $d_\alpha(P, Q) = 0$ alors pour tout $u \in \mathcal{S}_{d-1}$ on a

$$\left| P(H^Q(u, \alpha)) - \alpha \right| = 0$$

et

$$\left| Q(H^P(u, \alpha)) - \alpha \right| = 0$$

donc

$$Q(H^P(u, \alpha)) = P(H^Q(u, \alpha)) = \alpha$$

puisque

$$Q(H^P(u, \alpha)) = G_{\langle X', u \rangle} \circ F_{\langle X, u \rangle}^{-1}(\alpha)$$

donc sous (\mathbf{H}_1) on a $P - p.s.$ et $Q - p.s.$

$$Y^Q(0, u, \alpha) = Y^P(0, u, \alpha) \tag{4.5}$$

ainsi

$$H(0, u, Y^P(0, u, \alpha)) = H(0, u, Y^Q(0, u, \alpha)) \text{ p.s.}$$

donc $d_\alpha(P, Q) = 0$ implique

$$H^P(u, \alpha) = H^Q(u, \alpha) := H(u, \alpha).$$

4.3.2 Test d'adéquation d'un échantillon

Dans cette section, on souhaite construire le test d'adéquation suivant Soient P et P_0 deux mesures de probabilité vérifiant les conditions (\mathbf{H}_1) , (\mathbf{H}_2) et (\mathbf{H}'_3) . On souhaite tester pour P_0 fixée

$$(\mathbf{Hyp}_0) : d_\alpha(P, P_0) = 0 \text{ contre } (\mathbf{Hyp}_1) : d_\alpha(P, P_0) > 0.$$

Remarque 4.18. (\mathbf{Hyp}_0) n'implique évidemment pas que $P = P_0$. De même l'hypothèse

$$(\mathbf{Hyp}'_0) : d_\alpha(P, P_0) = 0 \quad \alpha \in \Delta$$

n'implique pas que $P = P_0$ car $\Delta \subset [1/2, 1)$. Cette hypothèse se test en remplaçant $d_\alpha(P_n, P_0)$ par $\sup_{\alpha \in \Delta} d_\alpha(P_n, P_0)$ ou $\int_{\alpha \in \Delta} d_\alpha(P_n, P_0) d\alpha$, dans le théorème suivant.

Théorème 4.19. Soit X_1, \dots, X_n un échantillon de loi P et P_n sa mesure empirique associée. On a

– sous (\mathbf{Hyp}_0)

$$\sqrt{n}d_\alpha(P_n, P_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \int_{\mathcal{S}_{d-1}} \left(1 + \frac{h^{P_0}(u, \alpha)}{h^P(u, \alpha)}\right) |\mathbf{G}_\alpha(u)| du$$

avec \mathbf{G}_α^P un processus gaussien de covariance

$$\forall u, u' \in \mathcal{S}_{d-1} \quad \text{cov}(\mathbf{G}_\alpha(u), \mathbf{G}_\alpha(u')) = P(H(u, \alpha) \cap H(u', \alpha)) - \alpha^2$$

– sous (\mathbf{Hyp}_1)

$$\lim_{n \rightarrow \infty} \sqrt{n}d_\alpha(P_n, P_0) = +\infty.$$

Démonstration. Soit $X \sim P$ et $X_0 \sim P_0$

Étape I : sous (\mathbf{Hyp}_1) , par le lemme 4.14 on sait que sous la condition (\mathbf{H}_1)

$$\lim_{n \rightarrow \infty} d_\alpha(P_n, P_0) = d_\alpha(P, P_0)$$

donc

$$\lim_{n \rightarrow \infty} \sqrt{n}d_\alpha(P_n, P_0) = +\infty$$

Étape II : sous **(Hyp₀)** on a par la remarque 4.17 on sait que sous la condition **(H₁)**

$$Q(H^P(u, \alpha)) = P(H^Q(u, \alpha)) = \alpha$$

$$Y^Q(O, u, \alpha) = Y^P(O, u, \alpha)$$

et

$$H^P(u, \alpha) = H^Q(u, \alpha) := H(u, \alpha)$$

donc

$$d_\alpha(P_n, P_0) = \int_{\mathbb{S}_{d-1}} |P_n(H(u, \alpha)) - \alpha| du + \int_{\mathbb{S}_{d-1}} |P_0(H_n(u, \alpha)) - \alpha| du$$

sous les condition **(H₁)**, **(H₂)** et **(H₃)**, par approximation forte du processus empirique indexé par les hyperplan du Théorème 2.43 on sait qu'on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite G_n^P de versions de G^P de sorte que

$$P_n(H(u, \alpha)) = P(H(u, \alpha)) + \frac{G_n^P(H(u, \alpha))}{\sqrt{n}} + \frac{Z_n(u, \alpha)}{\sqrt{n}}$$

avec $\lim_{n \rightarrow \infty} \|Z_n\|_{\mathbb{S}_{d-1} \times \Delta} = 0$ p.s. donc

$$\sqrt{n} \int_{\mathbb{S}_{d-1}} |P_n(H(u, \alpha)) - \alpha| du = \int_{\mathbb{S}_{d-1}} |G_n^P(H(u, \alpha))| du + o_{p.s.}(1).$$

Ensuite, on a

$$P_0(H_n(u, \alpha)) - P_0(H(u, \alpha)) = F_{\langle X_0, u \rangle}(Y_n(0_d, u, \alpha)) - F_{\langle X_0, u \rangle}(Y(0_d, u, \alpha)).$$

Sous **(H₁)** par la loi du logarithme itéré on sait qu'il existe $C > 0$ et $n(\omega) > 0$ tel que pour tout $n \geq n(\omega)$ on a

$$\sup_{u \in \mathbb{S}_{d-1}} |Y_n(O, u, \alpha) - Y^P(O, u, \alpha)| \leq C \sqrt{\frac{\log \log n}{n}}.$$

Ainsi, sous les condition **(H₁)**, **(H₂)** pour tout $n \geq n(\omega)$ et tout $u \in \mathbb{S}_{d-1}$

$$\begin{aligned} & F_{\langle X-O, u \rangle}(Y_m^Q(O, u, \alpha)) - F_{\langle X-O, u \rangle}(Y^Q(O, u, \alpha)) \\ & \leq f_{\langle X_0, u \rangle}^{P_0}(Y^Q(0_d, u, \alpha))(Y_n(0_d, u, \alpha) - Y(0_d, u, \alpha)) + \rho \left(C \sqrt{\frac{\log \log n}{n}} \right) \end{aligned}$$

sous (\mathbf{H}'_3) il existe $r > 1/2$ et $C^* > 0$ tel que pour tout $\gamma \leq \gamma_0$ on a

$$\rho(\gamma) < C^* \gamma^{1+r}$$

avec la notation

$$h^{P_0}(u, \alpha) = f_{\langle X, u \rangle}^{P_0}(Y(0_d, u, \alpha))$$

on a

$$\begin{aligned} F_{\langle X_0, u \rangle}(Y_n(0_d, u, \alpha)) - F_{\langle X_0, u \rangle}(Y(0_d, u, \alpha)) &= h^{P_0}(u, \alpha)(Y_n(0_d, u, \alpha) - Y(0_d, u, \alpha)) \\ &\quad + o_{p.s.} \left(\frac{1}{n^{(r+1)/2}} \right). \end{aligned}$$

Ensuite, par l'approximation forte du processus des quantiles empirique du Théorème 2.43 on a

$$Y_n(0_d, u, \alpha) = Y(0_d, u, \alpha) - \frac{\mathbf{G}_n^P(H(u, \alpha))}{h^P(u, \alpha)\sqrt{n}} + \frac{\mathbf{Z}'_n(u, \alpha)}{h^P(u, \alpha)\sqrt{n}}$$

avec $\lim_{n \rightarrow \infty} \|\mathbf{Z}'_n\|_{\mathcal{S}_{d-1} \times \Delta} = 0$ p.s. donc

$$\sqrt{n}(Y_n(0_d, u, \alpha) - Y(0_d, u, \alpha)) = -\frac{\mathbf{G}^P n(H(u, \alpha))}{h^P(u, \alpha)} + \frac{\mathbf{Z}'_n(u, \alpha)}{h^P(u, \alpha)}$$

ainsi pour $n \geq n(\omega)$

$$\begin{aligned} \sqrt{n}(P(H_n(u, \alpha)) - P(H(u, \alpha))) &= \sqrt{n} \left(F_{\langle X_0, u \rangle}(Y_n(0_d, u, \alpha)) - F_{\langle X_0, u \rangle}(Y(0_d, u, \alpha)) \right) \\ &= -\frac{h^{P_0}(u, \alpha)}{h^P(u, \alpha)} \mathbf{G}_n^P(H(u, \alpha)) + \frac{h^{P_0}(u, \alpha)}{h^P(u, \alpha)} \mathbf{Z}'_n(u, \alpha) + o_{p.s.}(1) \end{aligned}$$

ainsi, on a

$$\sqrt{n}d_\alpha(P_n, P_0) = \int_{\mathcal{S}_{d-1}} \left(1 + \frac{h^{P_0}(u, \alpha)}{h^P(u, \alpha)}\right) |\mathbf{G}_n^P(H(u, \alpha))| du + o_{p.s.}(1)$$

donc, puisque la loi de $\mathbf{G}_n^P(H(\cdot, \alpha))$ ne dépend pas de n , on a pour $\alpha \in \Delta$

$$\sqrt{n}d_\alpha(P_n, P_0) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \int_{\mathcal{S}_{d-1}} \left(1 + \frac{h^{P_0}(u, \alpha)}{h^P(u, \alpha)}\right) |\mathbf{G}_\alpha(u)| du$$

avec \mathbf{G}_α un processus gaussien centré de covariance

$$\forall u, u' \in \mathcal{S}_{d-1} \quad \text{cov}(\mathbf{G}_\alpha(u), \mathbf{G}_\alpha(u')) = P(H(u, \alpha) \cap H(u', \alpha)) - \alpha^2$$

□

4.3.3 Test de comparaison d'échantillons spatiaux

Soit P_n et Q_n deux mesures empiriques associées à deux échantillons de \mathbb{R}^d indépendants entre eux, X_1, \dots, X_n et X'_1, \dots, X'_m de lois respectives P et Q et de fonctions de répartition F et G .

On veut tester l'hypothèse **(Hyp₀)** : $d_\alpha(P, Q) = 0$ contre l'hypothèse **(Hyp₁)** : $d_\alpha(P, Q) > 0$.

On considère une statistique de test de type $v_{n,m}d_\alpha(P_n, Q_m)$ car sous l'hypothèse **(Hyp₁)**, si $n \rightarrow \infty$ et $m \rightarrow \infty$ on a

$$d_\alpha(P_n, Q_m) \rightarrow d_\alpha(P, Q) > 0$$

donc, si $v_{n,m} \rightarrow \infty$ la statistique de test diverge. reste à choisir $v_{n,m}$ pour que sous **(Hyp₀)** la statistique converge vers une loi limite à déterminer.

Théorème 4.20. Soit P_n et Q_m deux mesure empiriques associées à deux échantillons de \mathbb{R}^d indépendants entre eux, X_1, \dots, X_n et X'_1, \dots, X'_m de lois respectives P et Q vérifiant les conditions **(H₁)**, **(H₂)** et **(H₃)**.

On fixe $\alpha \in \Delta$ et on suppose que $n/(n+m) \rightarrow \delta \in]0, 1[$, on a

– sous **(Hyp₀)**

$$\sqrt{\frac{nm}{n+m}} d_\alpha(P_n, Q_m) \xrightarrow[n, m \rightarrow \infty]{\mathcal{L}} \int_{S_{d-1}} (h^P(u, \alpha) + h^Q(u, \alpha)) \left| \tilde{\tilde{G}}_\alpha(u) \right| du$$

avec

$$\tilde{\tilde{G}}_\alpha = \sqrt{1-\delta} \tilde{G}_\alpha^P - \sqrt{\delta} \tilde{G}_\alpha^Q$$

avec \tilde{G}^P et \tilde{G}^Q deux processus gaussiens centrés indépendants de fonctions covariance respectivement

$$\forall u, u' \in S_{d-1} \quad \text{cov}(\tilde{G}_\alpha^P(u), \tilde{G}_\alpha^P(u')) = \frac{P(H(u, \alpha) \cap H(u', \alpha)) - \alpha^2}{h^P(u, \alpha)h^P(u', \alpha)}$$

et

$$\forall u, u' \in S_{d-1} \quad \text{cov}(\tilde{G}_\alpha^Q(u), \tilde{G}_\alpha^Q(u')) = \frac{Q(H(u, \alpha) \cap H(u', \alpha)) - \alpha^2}{h^Q(u, \alpha)h^Q(u', \alpha)}$$

– sous **(Hyp₁)**

$$\lim_{n, m \rightarrow \infty} \sqrt{\frac{nm}{n+m}} d_\alpha(P_n, Q_m) = +\infty \text{ p.s.}$$

Démonstration. Tout d'abord on sait par la remarque 4.17 que sous **(Hyp₁)** pour tout $u \in \mathbb{S}_{d-1}$ on a

$$Q(H^P(u, \alpha)) = P(H^Q(u, \alpha)) = \alpha$$

$$Y^Q(O, u, \alpha) = Y^P(O, u, \alpha)$$

et

$$H^P(u, \alpha) = H^Q(u, \alpha) := H(u, \alpha)$$

Dans la suite pour une loi de probabilité \mathcal{L} sur \mathbb{R}^d on note $(\mathbf{G}^{\mathcal{L}})_{H \in \mathcal{H}}$ le processus gaussien centré de covariance

$$\forall H, H' \in \mathcal{H} \quad \text{cov}(\mathbf{G}^{\mathcal{L}}(H), \mathbf{G}^{\mathcal{L}}(H')) = \mathcal{L}(H \cap H') - \mathcal{L}(H)\mathcal{L}(H')$$

- Sous **(Hyp₀)** on suppose que $n/(n+m) \rightarrow \delta \in]0, 1[$, on a

$$\begin{aligned} & \sqrt{\frac{nm}{n+m}} d_\alpha(P_m, Q_m) \\ &= \sqrt{\frac{nm}{n+m}} \int_{u \in \mathbb{S}_{d-1}} |P_n(H_m^Q(u, \alpha)) - \alpha| du + \sqrt{\frac{nm}{n+m}} \int_{u \in \mathbb{S}_{d-1}} |Q_m(H_n^P(u, \alpha)) - \alpha| du \\ &= \sqrt{\frac{nm}{n+m}} \int_{u \in \mathbb{S}_{d-1}} |P_n(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha))| du \\ &+ \sqrt{\frac{nm}{n+m}} \int_{u \in \mathbb{S}_{d-1}} |Q_m(H_n^P(u, \alpha)) - Q(H^P(u, \alpha))| du \end{aligned} \quad (4.6)$$

dans la suite, on s'intéresse uniquement à la quantité

$$\int_{u \in \mathbb{S}_{d-1}} |P_n(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha))| du$$

l'autre quantité jouent un rôle symétrique. On pose

$$\Lambda_n^P = \sqrt{n}(P_n - P)$$

On sait que

$$\begin{aligned} & P_n(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha)) \\ &= P_n(H_m^Q(u, \alpha)) - P(H_m^Q(u, \alpha)) + P(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha)) \\ &= \left(P_n(H_m^Q(u, \alpha)) - P(H_m^Q(u, \alpha)) \right) - \left(P_n(H^Q(u, \alpha)) - P(H^Q(u, \alpha)) \right) \\ &+ \left(P_n(H^Q(u, \alpha)) - P(H^Q(u, \alpha)) \right) + \left(P(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha)) \right) \end{aligned}$$

donc

$$P_n(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha)) = \frac{1}{\sqrt{n}} \left(\Lambda_n^P(H_m^Q(u, \alpha)) - \Lambda_n^P(H^Q(u, \alpha)) \right) \quad (4.7)$$

$$+ (P_n(H(u, \alpha)) - P(H(u, \alpha))) \quad (4.8)$$

$$+ \left(P(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha)) \right) \quad (4.9)$$

Étape 1. On s'intéresse à la quantité $\Lambda_n^P(H_m^Q(u, \alpha)) - \Lambda_n^P(H^Q(u, \alpha))$. Par la remarque 2.57, on a

$$\begin{aligned} & \Lambda_n^P(H_m^Q(u, \alpha)) - \Lambda_n^P(H^Q(u, \alpha)) \\ &= \Lambda_n^P(H^Q(O, u, Y_m^Q(O, u, \alpha))) - \Lambda_n^P(H^Q(O, u, Y^Q(O, u, \alpha))) \\ &= \Lambda_n^P(H^Q(O, u, Y_m^Q(O, u, \alpha), Y_m(O, u, \alpha))) \end{aligned}$$

où $H^Q(O, u, Y_m^Q(O, u, \alpha), Y^Q(O, u, \alpha))$ est la bande coincée entre $H^Q(u, \alpha)$ et $H_m^Q(u, \alpha)$.

Sous **(H₁)** par la loi du logarithme itéré on sait qu'il existe $C > 0$ et $m_1(\omega) > 0$ tel que pour tout $m \geq m_1(\omega)$ on a

$$|Y_m^Q(O, u, \alpha) - Y^Q(O, u, \alpha)| \leq C \sqrt{\frac{\log \log m}{m}}.$$

puisque $n/(n+m) \rightarrow \delta \in]0, 1[$ donc $n/m = \delta/(1-\delta) + o(1)$ alors il existe $C_\delta > 0$ telle que

$$\sup_{u \in \mathbb{S}_{d-1}} |Y_m^Q(O, u, \alpha) - Y^Q(O, u, \alpha)| \leq \varepsilon_n = C_\delta \sqrt{\frac{\log \log n}{n}}.$$

On pose $\mathcal{B}_n = \bigcup_{\varepsilon \leq \varepsilon_n} \mathcal{B}_\varepsilon$ on a

$$\left| \Lambda_n^P(H_m^Q(u, \alpha)) - \Lambda_n^P(H^Q(u, \alpha)) \right| \leq \sup_{B \in \mathcal{B}_n} \left| \Lambda_n^P(B) \right|$$

par le lemme 2.59 on sait qu'il existe $K > 0$ telle que

$$\sup_{B \in \mathcal{B}_n} \left| \Lambda_n^P(B) \right| \leq K \frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}}$$

donc

$$\Lambda_n^P(H_m^Q(u, \alpha)) - \Lambda_n^P(H^Q(u, \alpha)) = o_{p.s.}(1).$$

De la même façon, on peut montrer qu'il existe $n_1(\omega) > 0$ tel que pour tout $n \geq n_1(\omega)$

$$\Lambda_m^Q(H_n^P(u, \alpha)) - \Lambda_m^Q(H^P(u, \alpha)) = o_{p.s.}(1).$$

Étape 2. Pour étudier $P_n(H(u, \alpha)) - P(H(u, \alpha))$ dans (4.8), on sait que par approximation forte du processus empirique indexé par les hyperplan, on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite G_n^P de versions de G^P de sorte que

$$P_n(H(u, \alpha)) = P(H(u, \alpha)) + \frac{G_n^P(H(u, \alpha))}{\sqrt{n}} + \frac{Z_n(u, \alpha)}{\sqrt{n}}$$

avec $\lim_{n \rightarrow \infty} \|Z_n\|_{S_{d-1} \times \Delta} = 0$ p.s.

Concernant $Q_n(H(u, \alpha)) - Q(H(u, \alpha))$ de manière analogue on montre qu'on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_m de loi Q et une suite G_m^Q de versions de G^Q de sorte que

$$Q_n(H(u, \alpha)) = Q(H(u, \alpha)) + \frac{G_m^Q(H(u, \alpha))}{\sqrt{m}} + \frac{\bar{Z}_m(u, \alpha)}{\sqrt{m}}$$

avec $\lim_{m \rightarrow \infty} \|\bar{Z}_m\|_{S_{d-1} \times \Delta} = 0$ p.s.

Étape 3. Enfin, on s'intéresse à la quantité

$$P(H_m^Q(u, \alpha)) - P(H^Q(u, \alpha)) = F_{\langle X-O, u \rangle}(Y_m^Q(O, u, \alpha)) - F_{\langle X-O, u \rangle}(Y^Q(O, u, \alpha))$$

Sous les condition **(H₁)**, **(H₂)** il existe $m_2(\omega) > 0$ tel que pour tout $m \geq m_2(\omega)$

$$\begin{aligned} & F_{\langle X-O, u \rangle}(Y_m^Q(O, u, \alpha)) - F_{\langle X-O, u \rangle}(Y^Q(O, u, \alpha)) \\ &= f_{\langle X-O, u \rangle}^P(Y^Q(O, u, \alpha))(Y_m^Q(O, u, \alpha) - Y(O, u, \alpha)) + \rho \left(C \sqrt{\frac{\log \log m}{m}} \right) \end{aligned}$$

sous **(H₃)** il existe $r > 1/2$ et $C^* > 0$ tel que pour tout $\gamma \leq \gamma_0$ on a

$$\rho(\gamma) < C^* \gamma^{1+r}$$

de plus, sous l'hypothèse nulle par l'égalité (4.5) on a

$$f_{\langle X-O, u \rangle}^P(Y^Q(O, u, \alpha)) = f_{\langle X-O, u \rangle}^P(Y^P(O, u, \alpha)) = h^P(u, \alpha)$$

donc

$$F_{\langle X-O, u \rangle}(Y_m^Q(O, u, \alpha)) - F_{\langle X-O, u \rangle}(Y^Q(O, u, \alpha)) = h^P(u, \alpha)(Y_m^Q(O, u, \alpha) - Y(O, u, \alpha)) \\ + o_{p.s.}\left(\frac{1}{m^{(r+1)/2}}\right)$$

Ensuite, par l'approximation forte du processus des quantiles empirique dans le théorème 2.43 on a montré que

$$Y_m^Q(O, u, \alpha) = Y(O, u, \alpha) - \frac{\mathbf{G}_m^Q(H(u, \alpha))}{h^Q(u, \alpha)\sqrt{m}} + \frac{\mathbf{Z}'_m(u, \alpha)}{h^Q(u, \alpha)\sqrt{m}}$$

avec $\lim_{m \rightarrow \infty} \|\mathbf{Z}'_m\|_{\mathcal{S}_{d-1} \times \Delta} = 0$ p.s. donc

$$\sqrt{m} \left(Y_m^Q(O, u, \alpha) - Y(O, u, \alpha) \right) = -\frac{\mathbf{G}_m^Q(H(u, \alpha))}{h^Q(u, \alpha)} + \frac{\mathbf{Z}'_m(u, \alpha)}{h^Q(u, \alpha)}$$

ainsi pour $m \geq m_2(\omega)$

$$\sqrt{m} \left(P(H_m^Q(u, \alpha)) - P(H(u, \alpha)) \right) = \sqrt{m} \left(F_{\langle X-O, u \rangle}(Y_m^Q(O, u, \alpha)) - F_{\langle X-O, u \rangle}(Y^Q(O, u, \alpha)) \right) \\ = -\frac{h^P(u, \alpha)}{h^Q(u, \alpha)} \mathbf{G}_m^Q(H(u, \alpha)) + \frac{h^P(u, \alpha)}{h^Q(u, \alpha)} \mathbf{Z}'_m(u, \alpha) + o_{p.s.}(1)$$

De la même manière, on montre qu'il existe $n_2(\omega)$ tel que pour $n \geq n_2(\omega)$ on a

$$\sqrt{n} \left(Q(H_n^P(u, \alpha)) - Q(H(u, \alpha)) \right) = \sqrt{n} \left(G_{\langle X'-O, u \rangle}(Y_n^P(O, u, \alpha)) - F_{\langle X'-O, u \rangle}(Y^P(O, u, \alpha)) \right) \\ = -\frac{h^Q(u, \alpha)}{h^P(u, \alpha)} \mathbf{G}_n^P(H(u, \alpha)) + \frac{h^Q(u, \alpha)}{h^P(u, \alpha)} \bar{\mathbf{Z}}'_n(u, \alpha) + o_{p.s.}(1)$$

avec $\lim_{n \rightarrow \infty} \|\bar{\mathbf{Z}}'_n\|_{\mathcal{S}_{d-1} \times \Delta} = 0$ p.s.

Conclusion : on pose $n(\omega) = n_1(\omega) \vee n_2(\omega)$ et $m(\omega) = m_1(\omega) \vee m_2(\omega)$, par les **Étapes 1, 2 et 3**, on a pour $n \geq n(\omega)$ et $m \geq m(\omega)$

$$\sqrt{\frac{nm}{n+m}} (P_n(H_m^Q(u, \alpha))) - \alpha = \sqrt{\frac{m}{n+m}} \mathbf{G}_n^P(H(u, \alpha)) - \sqrt{\frac{n}{n+m}} \frac{h^P(u, \alpha)}{h^Q(u, \alpha)} \mathbf{G}_m^Q(H(u, \alpha)) \\ + \sqrt{\frac{m}{n+m}} \bar{\mathbf{Z}}_n(u, \alpha) + \sqrt{\frac{n}{n+m}} \frac{h^P(u, \alpha)}{h^Q(u, \alpha)} \mathbf{Z}'_m(u, \alpha) + o_{p.s.}(1).$$

avec $\lim_{m \rightarrow \infty} \|\mathbf{Z}_n\|_{\mathcal{S}_{d-1} \times \Delta} = \lim_{m \rightarrow \infty} \|\mathbf{Z}'_m\|_{\mathcal{S}_{d-1} \times \Delta} = 0$ p.s.

ainsi que

$$\begin{aligned} \sqrt{\frac{nm}{n+m}}(Q_n(H_n^P(u, \alpha)) - \alpha) &= \sqrt{\frac{n}{n+m}} \mathbf{G}_m^Q(H(u, \alpha)) - \sqrt{\frac{m}{n+m}} \frac{h^Q(u, \alpha)}{h^P(u, \alpha)} \mathbf{G}_n^P(H(u, \alpha)) \\ &+ \sqrt{\frac{n}{n+m}} \bar{\mathbf{Z}}_m(u, \alpha) + \sqrt{\frac{m}{n+m}} \frac{h^Q(u, \alpha)}{h^P(u, \alpha)} \bar{\mathbf{Z}}'_n(u, \alpha) + o_{p.s.}(1). \end{aligned}$$

avec $\lim_{m \rightarrow \infty} \|\bar{\mathbf{Z}}_n\|_{\mathcal{S}_{d-1} \times \Delta} = \lim_{m \rightarrow \infty} \|\bar{\mathbf{Z}}'_m\|_{\mathcal{S}_{d-1} \times \Delta} = 0$ p.s.

Afin de mener les calculs on pose les notations suivantes

$$\mathbb{K}_{m,n}(u) = \sqrt{\frac{m}{n+m}} \frac{\mathbf{G}_n^P(H(u, \alpha))}{h^P(u, \alpha)} - \sqrt{\frac{n}{n+m}} \frac{\mathbf{G}_m^Q(H(u, \alpha))}{h^Q(u, \alpha)}.$$

$$R_{m,n}(u) = \sqrt{\frac{m}{n+m}} \mathbf{Z}_n(u, \alpha) + \sqrt{\frac{n}{n+m}} \frac{h^P(u, \alpha)}{h^Q(u, \alpha)} \mathbf{Z}'_m(u, \alpha) + o_{p.s.}(1).$$

$$\bar{R}_{m,n}(u) = \sqrt{\frac{n}{n+m}} \bar{\mathbf{Z}}_m(u, \alpha) + \sqrt{\frac{m}{n+m}} \frac{h^Q(u, \alpha)}{h^P(u, \alpha)} \bar{\mathbf{Z}}'_n(u, \alpha) + o_{p.s.}(1).$$

puisque sous la condition **(H₁)** il existe deux constantes $0 < M_* \leq M^*$ telles que

$$M_* \leq \inf_{u \in \mathcal{S}_{d-1}} |h^P(u, \alpha)| < \sup_{u \in \mathcal{S}_{d-1}} |h^P(u, \alpha)| \leq M^*$$

alors

$$\lim_{n \rightarrow \infty} \|R_{m,n}\|_{\mathcal{S}_{d-1}} = 0 \text{ p.s.}$$

de même on a

$$\lim_{n \rightarrow \infty} \|\bar{R}_{m,n}\|_{\mathcal{S}_{d-1}} = 0 \text{ p.s.}$$

$$\begin{aligned} \sqrt{\frac{nm}{n+m}}(P_n(H_m^Q(u, \alpha)) - \alpha) &= h^P(u, \alpha) \mathbb{K}_{m,n}(u) + R_{m,n}(u) \\ \sqrt{\frac{nm}{n+m}}(Q_n(H_n^P(u, \alpha)) - \alpha) &= -h^Q(u, \alpha) \mathbb{K}_{m,n}(u) + \bar{R}_{m,n}(u). \end{aligned}$$

pour tout $a, b \in \mathbb{R}$ on a

$$-|b| \leq |a + b| - |a| \leq |b|$$

donc pour tout $u \in \mathbb{S}_{d-1}$, on a

$$-|R_{m,n}(u)| \leq |h^P(u, \alpha)\mathbb{K}_{m,n}(u) + R_{m,n}(u)| - |h^P(u, \alpha)\mathbb{K}_{m,n}(u)| \leq |R_{m,n}(u)|$$

$$- \|R_{m,n}\|_{\mathbb{S}_{d-1}} \leq |h^P(u, \alpha)\mathbb{K}_{m,n}(u) + R_{m,n}(u)| - |h^P(u, \alpha)\mathbb{K}_{m,n}(u)| \leq \|R_{m,n}(u)\|_{\mathbb{S}_{d-1}}$$

puisque $\lim_{n \rightarrow \infty} \|R_{m,n}\|_{\mathbb{S}_{d-1}} = 0$ p.s. alors

$$\sqrt{\frac{nm}{n+m}} \int_{\mathbb{S}_{d-1}} |P_n(H_m^Q(u, \alpha)) - \alpha| du = \int_{\mathbb{S}_{d-1}} |h^P(u, \alpha)\mathbb{K}_{m,n}(u)| du + o_{p.s.}(1)$$

donc

$$d_\alpha(P_n, Q_m) = \int_{\mathbb{S}_{d-1}} (h^P(u, \alpha) + h^Q(u, \alpha)) |\mathbb{K}_{m,n}(u)| du + o_{p.s.}(1)$$

Puisque G_m^P et G_m^Q sont indépendants par construction et de lois respectives $G_\alpha^P := G^P(H(\cdot, \alpha))$ et $G_\alpha^Q := G^Q(H(\cdot, \alpha))$ alors avec la notation

$$\tilde{G}_\alpha^P = \frac{G^P(H(u, \alpha))}{h^P(u, \alpha)} \quad \text{et} \quad \tilde{G}_\alpha^Q = \frac{G^Q(H(u, \alpha))}{h^Q(u, \alpha)}$$

on a

$$\mathbb{K}_{m,n} \stackrel{\mathcal{L}oi}{\sim} \sqrt{\frac{m}{n+m}} \tilde{G}_\alpha^P - \sqrt{\frac{n}{n+m}} \tilde{G}_\alpha^Q.$$

On pose

$$\tilde{\tilde{G}}_\alpha = \sqrt{1-\delta} \tilde{G}_\alpha^P - \sqrt{\delta} \tilde{G}_\alpha^Q$$

Sous (**Hyp₀**)

$$\sqrt{\frac{nm}{n+m}} d_\alpha(P_n, Q_m) \xrightarrow[n, m \rightarrow \infty]{\mathcal{L}} \int_{\mathbb{S}_{d-1}} (h^P(u, \alpha) + h^Q(u, \alpha)) |\tilde{\tilde{G}}_\alpha(u)| du$$

□

4.4 Surfaces médianes

La question d'une médiane spatiale est très naturelle et a évidemment un grand nombre de candidats. Pour la caractériser il faut disposer d'une notion d'ordre ou de quantile dans \mathbb{R}^d , ou d'une fonction classant les points. Divers auteurs ont fait des propositions, à chaque fois sous hypothèses assez fortes sur la loi afin de garantir l'existence des objets étudiés par exemple, injectant \mathbb{R}^d dans \mathbb{R} , ce qui est une contrainte forte comme dans le cas de

"bonnes" fonctions de profondeur $D(x, P)$ de x pour P (cf. Chapitre 1 section 1.3).

Dans notre cas, la notion de quantile est subjective et donne lieu à des surfaces. Pourtant on cherche un point ou une région faisant office de mode. Prendre l'intersection des demi-espaces H de probabilité $1/2$ n'est pas une solution car $\mathcal{C}_{1/2} = \bigcap_{H \in \mathcal{H}_{1/2}} H$ est le plus souvent vide, l'autre possibilité sous nos hypothèses étant qu'elle soit réduite à un point. C'est le cas notamment pour des lois symétriques par rapport à un point M qui est de fait candidat naturel à toute notion de localisation et notamment la médiane (qui coïncide donc avec l'espérance), et cela quelle que soit l'approche générale retenue pour la notion de quantile multivarié. La question qui se pose est donc : que faire si $\mathcal{C}_{1/2} = \emptyset$?

L'idée dans ce court chapitre consiste donc à considérer les surfaces $Q(O, \alpha)$ pour $\alpha = 1/2$ dans les définitions de la surface quantile dans le chapitre 2. L'étude complète du cas $\alpha = 1/2$ dans les surfaces quantiles est un sujet à part entière car comme on verra dans les simulations, elles ont potentiellement des formes très irrégulières avec un grand nombre de franchissements et de boucles si la loi n'est pas symétrique et même si O est une presque médiane. Par conséquent, leur version empirique risque d'être elle aussi très instable, et le processus gaussien limite de variance très faible.

4.4.1 Définition et théorème

Dans cette section il s'agit tout simplement de rappeler les théorèmes et définitions du chapitre 2 pour le cas particulier de $\alpha = 1/2$. En effet, ces théorèmes restent valables quand $\alpha = 1/2$ car la restriction pour les $\alpha \in \Delta = [\alpha^-, \alpha^+] \subset [1/2, 1)$ ne concerne que les α proches de 1. Ensuite on montre dans le cas simple des lois symétriques, la bonne propriété des surfaces. Enfin, la dernière partie de cette section est consacrée aux simulations.

Définition 4.21. Soit $O \in \mathbb{R}^d$ et $u \in \mathbb{S}_{d-1}$.

- Le **quantile médian** de la loi P vu de O dans la direction u est défini par

$$\hat{Y}(O, u) = \inf \{y \in \mathbb{R} : P(H(O, u, y)) \geq 1/2\}.$$

- Le **point quantile médian** de la loi P vu de et $O \in \mathbb{R}^d$ dans la direction $u \in \mathbb{S}_{d-1}$ est défini par

$$\hat{Q}(O, u) = O + \hat{Y}(O, u) \cdot u.$$

- Le **demi-espace médian** de la loi P vu de et $\mathbf{O} \in \mathbb{R}^d$ dans la direction $\mathbf{u} \in \mathbb{S}_{d-1}$ est défini par

$$\hat{H}(\mathbf{u}) := H(\mathbf{O}, \mathbf{u}, \hat{Y}(\mathbf{O}, \mathbf{u})).$$

- La **surface quantile médiane** de la loi P vu de et $\mathbf{O} \in \mathbb{R}^d$ dans la direction $\mathbf{u} \in \mathbb{S}_{d-1}$ est définie par

$$\hat{Q}(\mathbf{O}) = \{\hat{Q}(\mathbf{O}, \mathbf{u}) : \mathbf{u} \in \mathbb{S}_{d-1}\}.$$

Ensuite en fixant $\alpha = 1/2$, on a immédiatement les propositions et théorèmes suivants comme corollaires de ceux du chapitre 1

Corollaire 4.22.

- Soit $\mathbf{O} \in \mathbb{R}^d$, sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) , l'ensemble $\hat{Q}(\mathbf{O})$ est une surface fermée.
- **Consistance uniforme** : sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) on a

$$\lim_{n \rightarrow \infty} \sup_{\mathbf{u} \in \mathbb{S}_{d-1}} |\hat{Y}_n(\mathbf{O}, \mathbf{u}) - \hat{Y}(\mathbf{O}, \mathbf{u})| = 0 \quad p.s.$$

- **Loi du logarithme itéré** : sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+)

$$\limsup_{n \rightarrow \infty} \frac{\|\hat{Y}_n - \hat{Y}\|_{\mathbb{S}_{d-1}}}{\Psi^{-1}\left(\sqrt{\frac{\log \log n}{n}}\right)} < \infty \quad p.s.$$

- **Théorème limite central uniforme** : sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , sur l'ensemble des fonctions bornées sur \mathbb{S}_{d-1} muni de la norme uniforme, le processus $\sqrt{n}(\hat{Y}_n - \hat{Y})$ converge faiblement vers $\hat{\mathbf{G}}$ un processus centré et covariance

$$\text{cov}(\hat{\mathbf{G}}(\mathbf{u}), \hat{\mathbf{G}}(\mathbf{u}')) = \frac{P(\hat{H}(\mathbf{u}) \cap \hat{H}(\mathbf{u}')) - 1/4}{\hat{h}(\mathbf{u})\hat{h}(\mathbf{u}')}$$

où $\hat{h}(\mathbf{u}) = f_{\langle X, \mathbf{u} \rangle}(F_{\langle X, \mathbf{u} \rangle}^{-1}(1/2))$.

- **Principe d'invariance forte uniforme avec vitesse** : sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , on peut construire sur un même espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ une suite i.i.d. X_n de loi P et une suite $\hat{\mathbf{G}}_n$ de versions de $\hat{\mathbf{G}}$ de sorte que

$$\hat{Y}_n(\mathbf{O}, \mathbf{u}) = \hat{Y}(\mathbf{O}, \mathbf{u}) + \frac{\hat{\mathbf{G}}_n(\mathbf{u})}{\sqrt{n}} + \frac{\hat{\mathbf{Z}}_n(\mathbf{u})}{\sqrt{n}}, \quad \mathbf{O} \in \mathbb{R}^d, \quad \mathbf{u} \in \mathbb{S}_{d-1}, \quad (4.10)$$

où $\hat{Z}_n = \sqrt{n}(\hat{Y}_n - \hat{Y}) - \hat{G}_n$ vérifie

$$\lim_{n \rightarrow \infty} \|\hat{Z}_n\|_{S_{d-1}} = \lim_{n \rightarrow \infty} \sup_{u \in S_{d-1}} |\hat{Z}_n(u)| = 0 \quad p.s.$$

si de plus P satisfait (\mathbf{H}_3) alors on peut construire \hat{G}_n tel que pour $v_d = 1/(2 + 10d)$, $w_d = (4 + 10d)/(4 + 20d)$, il existe $n_\omega > 0$ tel que tout pour $n > n_\omega$,

$$\|\hat{Z}_n\|_{S_{d-1}} = O_{p.s.} \left(\frac{(\log n)^{w_d}}{n^{v_d}} \right).$$

– **Représentation de type Bahadur-Kiefer** : sous les conditions (\mathbf{H}_1) et (\mathbf{H}_2) , on a

$$\lim_{n \rightarrow \infty} \left\| \sqrt{n}(Y_n - Y) + \frac{\hat{\mathbb{E}}_n}{\hat{h}} \right\|_{S_{d-1}} = 0 \quad p.s.$$

Si de plus P satisfait (\mathbf{H}_3) ou (\mathbf{H}'_3) alors il existe $n_\omega > 0$ tel que tout pour $n > n_\omega$,

$$\left\| \sqrt{n}(\hat{Y}_n - \hat{Y}) + \frac{\hat{\mathbb{E}}_n}{\hat{h}} \right\|_{S_{d-1}} = O_{p.s.} \left(\frac{(\log n)^{1/2} (\log \log n)^{1/4}}{n^{1/4}} \right)$$

où $\hat{\mathbb{E}}_n(u) = \sqrt{n} (P_n(\hat{H}(u)) - P(\hat{H}(u)))$

Remarque 4.23. Pour tout $O \in \mathbb{R}^d$, sous les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) , les surfaces quantiles $\hat{Q}(O)$ sont des surfaces « double » i.e. chaque point de la surface correspond à un point quantile $Q(O, u, 1/2)$ et son dual $Q(O, -u, 1/2)$ pour tout $u \in S_{d-1}$.

4.4.2 Surfaces médianes pour les lois symétriques

Soit $P = \mathbb{P}_X$ une loi de probabilité admettant une H-symétrie autour d'un point M , i.e. pour tout demi-espace H contenant M on a $P(H) \geq 1/2$. Sous la condition (\mathbf{H}_0^-) ceci implique que pour tout $u \in S_{d-1}$ on a

$$P(H(M, u, 0)) = 1/2$$

où $H(M, u, 0)$ est le demi-espace fermé de normale u avec $M \in \partial H(M, u, 0)$. Dans ce cas particulier, sous (\mathbf{H}_0^-) et (\mathbf{H}_0^+) on a pour tout $u \in S_{d-1}$

$$\hat{Y}(M, u) = Y(O, u, 1/2) = 0$$

Proposition 4.24. Soit $P = \mathbb{P}_X$ une loi de probabilité admettant une H -symétrie autour d'un point M et vérifiant les conditions (\mathbf{H}_0^-) et (\mathbf{H}_0^+) .

Pour tout $O \in \mathbb{S}_{d-1}$ la surface quantile de P vu de O de niveau $1/2$ est une sphère de centre $(O + M)/2$ et de diamètre $d(O, M)$, i.e.

$$\hat{Q}(O) = \mathbb{S}(A, r)$$

avec $A = (O + M)/2$ et $r = d(O, M)/2$.

Démonstration. Soient $O \in \mathbb{R}^d \setminus \{M\}$ et $u \in \mathbb{S}_{d-1}$ on sait que

$$\hat{Y}(O, u) = \hat{Y}(M, u) + \langle M - O, u \rangle$$

or $\hat{Y}(M, u) = 0$ donc

$$\hat{Q}(O, u) = O + \langle M - O, u \rangle u$$

on note $\lambda = \langle O - M, u \rangle$, ainsi $O - \hat{Q}(O, u) = \lambda u$. On a

$$\begin{aligned} \langle O - \hat{Q}(O, u), M - \hat{Q}(O, u) \rangle &= \langle \lambda u, M - (O - \lambda u) \rangle \\ &= \langle \lambda u, M - O + \lambda u \rangle \\ &= \lambda (\langle u, M - O \rangle + \lambda \langle u, u \rangle) \\ &= \lambda (-\lambda + \lambda) \\ &= 0 \end{aligned}$$

donc pour tout $u \in \mathbb{S}_{d-1}$ on a $O - \hat{Q}(O, u)$ est orthogonale à $M - \hat{Q}(O, u)$ donc la surface quantile de P vu de O de niveau $1/2$ est une sphère de centre $(O + M)/2$ et de diamètre $d(O, M)$, i.e.

$$\hat{Q}(O) = \mathbb{S}(A, r)$$

avec $A = (O + M)/2$ et $r = d(O, M)/2$. □

4.4.3 Simulations

On se place dans \mathbb{R}^2 . Dans cette section, on présente les simulation des surfaces médianes pour trois type de loi : des lois symétriques, des lois multimodales et des lois avec support de dimension inférieure.

4.4.3.a Gaussienne $N(m, \Sigma)$

on pose

$$P = N(m, \Sigma)$$

une loi gaussienne d'espérance $m = (0,0)$ et de matrice de variance-covariance $\Sigma = I_2$. Les surfaces sont générées avec un échantillon de taille $N = 10000$. Cette loi étant symétrique (C-symétrie) la vrai surface médiane en $O = m$ est un cercle de rayon $\|O - m\|_2$ si $O \neq \mathbb{E}(X)$ (figures 4.8).

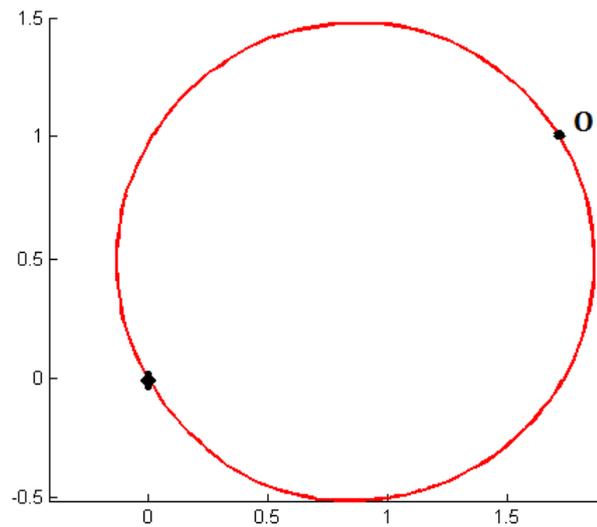
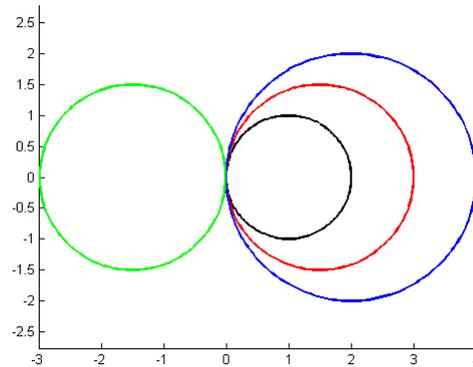


FIGURE 4.8 – Surface médiane de $N(m, \Sigma)$ vu de $O = (2 \cos(\frac{\pi}{6}), 2 \sin(\frac{\pi}{6}))$

Ensuite dans la figure 4.9 on représente quatre surfaces médianes pour $O \in \mathcal{O} = \{(0, -3), (0, 2), (0, 3), (0, 4)\}$, avec les correspondances suivantes

- en vert $\hat{Q}(O)$ pour $O = (-3, 0)$
- en noir $\hat{Q}(O)$ pour $O = (2, 0)$
- en rouge $\hat{Q}(O)$ pour $O = (3, 0)$
- en bleu $\hat{Q}(O)$ pour $O = (4, 0)$.

FIGURE 4.9 – Surface médiane de $N(m, \Sigma)$ vu de $O \in \mathcal{O}$

4.4.3.b Lois multimodales : mélanges gaussiens

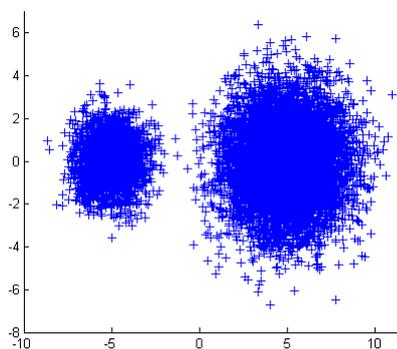
Dans cette partie, on va s'intéresser aux surfaces quantiles de lois avec plusieurs modes. On commence par une loi bimodale, comme mélange de deux gaussiennes

$$P = 1/4N(O_1, \Sigma_1) + 3/4N(O_2, \Sigma_2).$$

Simulation pour $O_1 = (-5, 0)$, $O_2 = (5, 0)$, $\Sigma_1 = I_2$ **et** $\Sigma_2 = 3I_2$

On commence par s'intéresser au surface de la loi

$$P = 1/4N(O_1, I_2) + 3/4N(O_2, 3I_2).$$

FIGURE 4.10 – Échantillon de P de taille 10^4

On commence par tracer la surface médiane de P pour $O \in \{(-5,0), (0,0)\}$ (figure 4.11a et 4.11b).

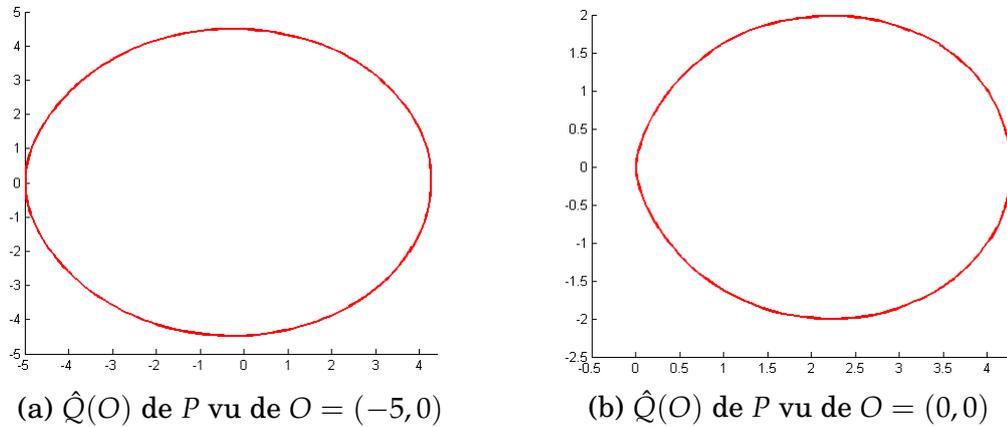


FIGURE 4.11 – Surfaces médiane de P pour $O \in \{(-5,0), (0,0)\}$

On fait remarquer que toutes les surfaces médiane sont en réalité des surface « double », dans le sens où chaque point de la surface correspond à la fois à $\hat{Q}(O, u)$ et $\hat{Q}(O, -u)$ pour tout $u \in \mathbb{S}_{d-1}$, ceci peut aussi être expliquer de manière plus visuelle et intuitive de la façon suivante : lorsque $\alpha \rightarrow 1/2$, la « boucle intérieure » observée dans le chapitre 2, croît à l'intérieur de la « boucle extérieure » jusqu'à ce que les boucles soient confondues. Dans la figure suivante (4.12a) on trace quatre surfaces médiane pour $O \in \mathcal{O} = \{(-15,6), (-5,6), (-5,-6), (15,6)\}$.

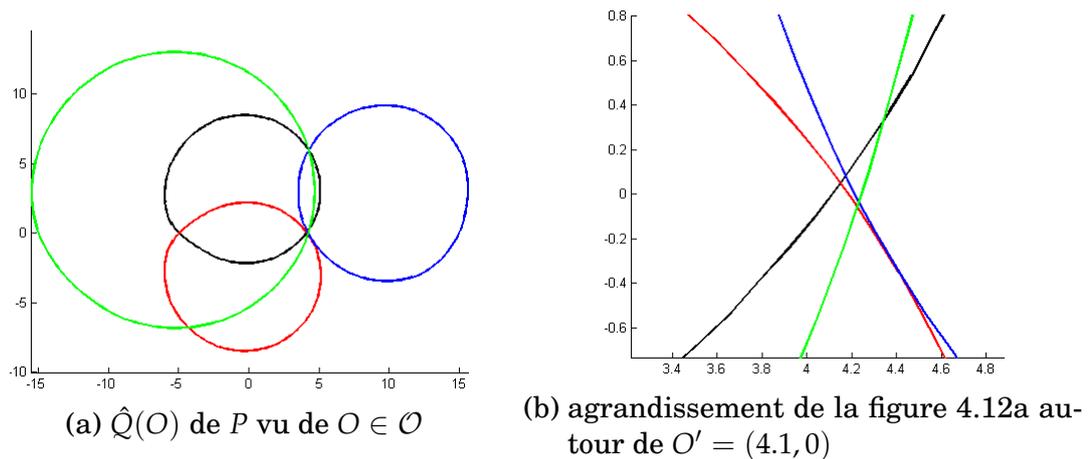


FIGURE 4.12 – Surfaces médiane de P pour $O \in \mathcal{O}$

Avec les correspondances suivantes

- en vert $\hat{Q}(O)$ pour $O = (-15, 6)$.
- en noir $\hat{Q}(O)$ pour $O = (-5, 6)$
- en rouge $\hat{Q}(O)$ pour $O = (-5, -6)$
- en bleu $\hat{Q}(O)$ pour $O = (15, 6)$

Remarquons que les deux figures précédentes butent dans l'axe des abscisses sur une zone autour de $O' = (4.1, 0)$, dont voici la surface $\hat{Q}(O')$ (figure 4.13)

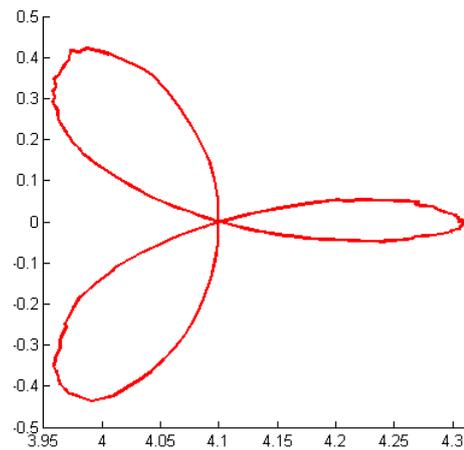


FIGURE 4.13 – Surfaces médiane de P pour $O' = (4.1, 0)$

Dans la figure 4.13 la surface s'intersecte avec elle même, ceci est dû à l'inexistence d'un centre de symétrie ce qui implique des passages répétés par O' en parcourant S_{d-1} .

4.4.3.c Lois à support spiral

Comme précédemment, on s'intéresse aux surfaces médiane pour une loi sans densité par rapport à la mesure de Lebesgue de \mathbb{R}^2 . On pose

$$X = (r \cos(\theta), r \sin(\theta))$$

avec $\theta \sim \mathcal{U}([0, 8\pi])/4$ et $r = 1/4 + \theta$.

On commence par tracer quatre surfaces médiane (figure 4.15) pour $O \in \mathcal{O} = \{(-15, -15), (-15, 15), (15, -15), (15, 15)\}$, avec les correspondances suivantes

- en vert $\hat{Q}(O)$ pour $O = (15, -15)$

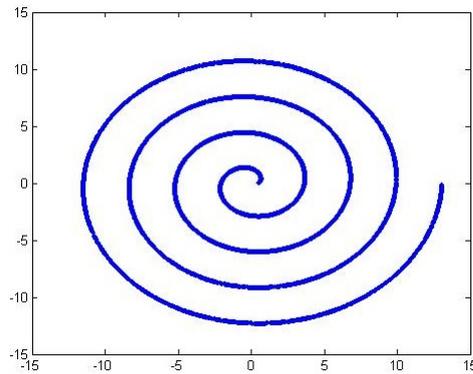


FIGURE 4.14 – Échantillon de la loi P avec $N = 10000$

- en noir $\hat{Q}(O)$ pour $O = (15, 15)$
- en rouge $\hat{Q}(O)$ pour $O = (-15, 15)$
- en bleu $\hat{Q}(O)$ pour $O = (-15, -15)$

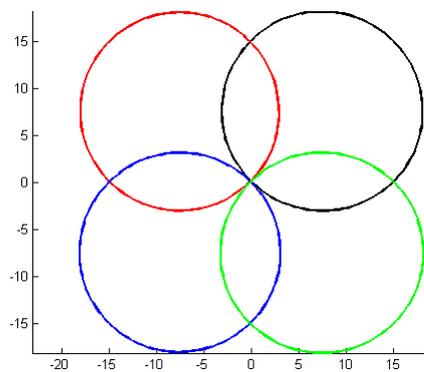
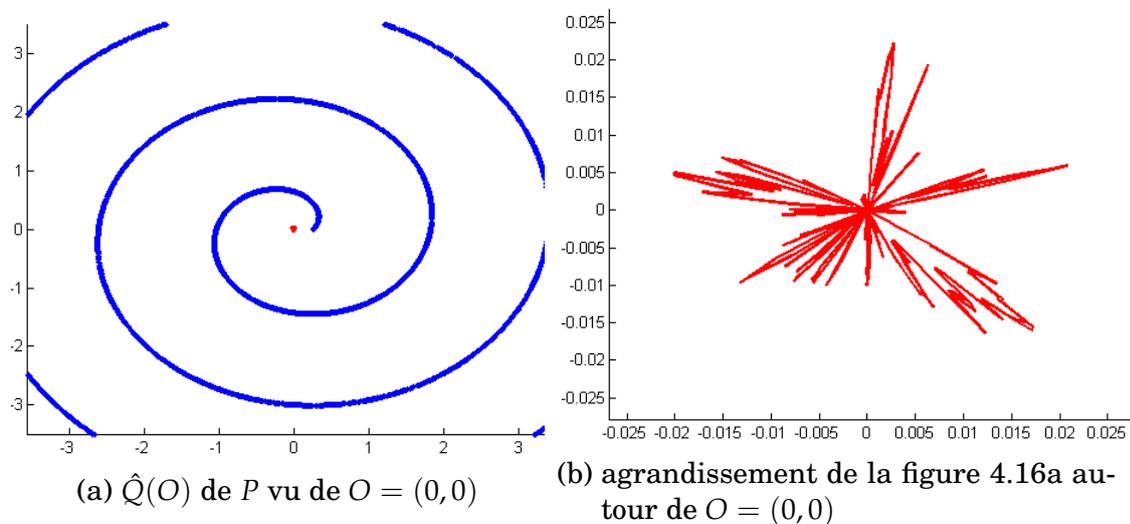


FIGURE 4.15 – Surfaces médiane de P pour $O \in \mathcal{O}$

Dans ce cas, même si la loi P n'est pas symétrique, la zone autour de $O = (0, 0)$ est un bon candidat pour le choix du centre, néanmoins la surface $\hat{Q}(O)$ s'auto-intersecte un grand nombre de fois avec elle-même (figure 4.16a, 4.16b). De plus, sous (\mathbf{H}_0^-) et (\mathbf{H}_0^+) chaque point de cette surface est un point double car $\hat{Q}(O, u) = \hat{Q}(O, -u)$ pour tout $u \in \mathbb{S}_{d-1}$.

FIGURE 4.16 – Surfaces médiane de P pour $O \in \mathcal{O}$

4.5 Information auxiliaire

4.5.1 Motivation

Dans [Ber14] P. Berthet explore les deux paradigmes suivants.

Tout d'abord, en supposant que la mesure inconnue P soit en partie connue, quelles vitesses d'estimation non paramétrique en découlent ? Si en effet on acquiert une conviction forte, comme celle d'un expert, au sujet d'une classe d'événements, d'une variable, ou d'une combinaison de variables φ , il est clair que l'estimation de la probabilité d'événements en lien avec ceux bien connus est améliorée si la conviction correspond à la réalité. La question est de savoir en quoi celle des autres événements est améliorable aussi, ou non. Il s'agit donc de trouver une modification de la mesure empirique \mathbb{P}_n qui permet d'améliorer l'estimation de la loi P lorsqu'elle est connue ou mieux approximée sur une classe d'ensembles mesurables, une sous-tribu, notamment engendrée par une ou plusieurs partitions. La modification de \mathbb{P}_n dépend de la nature de P , φ et de la classe \mathcal{C} des éléments sur lesquels il reste à estimer P .

Ce paradigme se prolonge de façon concrète dans le second : supposons que l'on puisse payer moins cher les observations partielles, comme par exemple l'observation d'une seule coordonnée de X , de son appartenance à tel ensemble d'une partition, ou la valeur d'une statistique subordonnée $Y = \varphi(X)$, moins complexe. Ainsi observer X_1 coûte 1 et observer à la place $\varphi(X_1)$ coûte $c < 1$, pour un budget de n qui représente à la fois le temps

et le coût expérimental, acquérir $\varphi(X_1)$ est plus facile et plus court que X_1 complet. La question est alors d'équilibrer $n = n_1 + cn_2$ où n_2 observations partielles permettent d'acquérir une bonne conviction sur $P^Y = P \circ \varphi^{-1}$ et d'utiliser la mesure empirique modifiée comme par le premier paradigme par cette meilleure connaissance, pour estimer P sur \mathcal{C} via n_1 observations.

Il est intéressant d'appliquer ces paradigmes d'estimation au cas des surfaces quantiles. En effet, dans de nombreux contextes il est réaliste de supposer certaines coordonnées de X mieux appréhendées que d'autres par des connaissances expertes, ou un historique d'estimations antérieures. Ou alors moins coûteuses que l'expérience complète $X(\omega)$. Concernant les surfaces quantiles, l'information supposée connue peut être diverse. Deux cas se dégagent et ne semblent pas relever des mêmes techniques de traitement :

(i) la loi d'une coordonnée de X est connue, ou de façon équivalente la loi d'une combinaison linéaire des coordonnées, ou la loi de X discrétisée selon une partition,

(ii) une ou plusieurs surfaces quantiles sont connues et il faut interpoler entre elles grâce à l'échantillon.

L'information donnée par (ii) est assez complexe mais encore très incomplète, comme en (i). Dans chaque situation (i) et (ii) la connaissance exacte peut plus réalistement être remplacée par une approximation meilleure que celle possible avec un n -échantillon, donc meilleure que $n^{-1/2}$. En appliquant [Ber14] on obtient facilement des variantes des résultats des chapitre 2 et chapitre 3.

4.5.2 Outil

On observe X_1, \dots, X_n i.i.d. sur $(\Omega, \mathcal{T}, \mathbb{P})$ et à valeurs dans $(\mathcal{X}, \mathcal{A})$. Supposons que P soit connue sur une partition finie. Ainsi $P(A_j)$ sont connus, $j \leq m$ avec $\mathcal{A}_m = \{A_j : j \leq m\} \subset \mathcal{A}$ une partition mesurable de \mathcal{X} . La mesure empirique modifiée par cette information étudiée dans [Ber14] est

$$\widehat{P}_n(A) = \sum_{j \leq m} \mathbb{P}_n(A | A_j) P(A_j) = \sum_{j \leq m} \mathbb{P}_n(A \cap A_j) \frac{P(A_j)}{\mathbb{P}_n(A_j)}$$

qui est l'analogie de

$$P(A) = \sum_{j \leq m} P(A | A_j) P(A_j) = \sum_{j \leq m} P(A \cap A_j) \frac{P(A_j)}{P(A_j)}.$$

Ainsi $\widehat{P}_n(A_j) = P(A_j)$ et de plus \widehat{P}_n préserve les probabilités conditionnelles empiriques, qui ne peuvent évidemment pas être améliorées par la connais-

sance des seuls $P(A_j)$, en effet on a bien $\widehat{P}_n(A | A_j) = \mathbb{P}_n(A | A_j)$.

La mesure \widehat{P}_n est discrète et possède les mêmes points de saut que \mathbb{P}_n . Le poids associé à X_i par \widehat{P}_n est aléatoire, c'est $P(A_j)/n\mathbb{P}_n(A_j)$ si $X_i \in A_j$. Il est facile de montrer que \widehat{P}_n minimise la distance de Kullback-Leibler $d_{KL}(Q, \mathbb{P}_n)$ entre \mathbb{P}_n et l'ensemble des probabilités discrètes Q sur X_1, \dots, X_n qui coïncident avec P sur \mathcal{A}_m . Les processus empiriques associés sont

$$\alpha_n(A) = \sqrt{n} (\mathbb{P}_n(A) - P(A)), \quad \widehat{\alpha}_n(A) = \sqrt{n} (\widehat{P}_n(A) - P(A))$$

et on a

$$\begin{aligned} \widehat{\alpha}_n(A) &= \sqrt{n} \sum_{j \leq m} \left(\mathbb{P}_n(A \cap A_j) \frac{P(A_j)}{\mathbb{P}_n(A_j)} - P(A \cap A_j) \right) \\ &= \sqrt{n} \sum_{j \leq m} \frac{P(A_j)}{\mathbb{P}_n(A_j)} (\mathbb{P}_n(A \cap A_j) - \mathbb{P}_n(A_j)P(A | A_j)) \\ &= \sqrt{n} \sum_{j \leq m} \frac{P(A_j)}{\mathbb{P}_n(A_j)} (\mathbb{P}_n(A \cap A_j) - P(A \cap A_j) - (\mathbb{P}_n(A_j) - P(A_j))P(A | A_j)) \\ &= \sum_{j \leq m} \frac{P(A_j)}{\mathbb{P}_n(A_j)} (\alpha_n(A \cap A_j) - \alpha_n(A_j)P(A | A_j)). \end{aligned}$$

En s'intéressant par exemple à une VC-classe $\mathcal{C} \subset \mathcal{A}$ et dans le cas le plus simple d'une partition finie \mathcal{A}_m comme ci-dessus, le comportement de $\widehat{\alpha}_n$ indexé par \mathcal{C} décrit dans le théorème ci-dessous. L'indice de \mathcal{C} est le plus petit entier VC tel que \mathcal{C} ne pulvérise aucun ensemble $E \subset \mathcal{X}$ de cardinal VC , au sens où il est impossible d'écrire chacun des 2^{VC} sous-ensembles de E sous la forme $E \cap C$ avec $C \in \mathcal{C}$.

Théorème 4.25. (Berthet, 2014) *Pour une partition finie \mathcal{A}_m , le processus $\widehat{\alpha}_n$ indexé par une VC-classe \mathcal{C} d'indice VC vérifie $\|\widehat{\alpha}_n\|_{\mathcal{C}} = O_{p.s.}(\sqrt{\log \log n})$ et converge en loi vers le processus $\widehat{G} = G - G^\nabla$ où G est le pont Brownien indexé par \mathcal{C} et $G^\nabla(A) = \sum_{j \leq m} G(A_j)P(A | A_j)$, $A \in \mathcal{C}$. Pour tout $\beta > 0$ il existe c_β telle qu'il est possible de construire sur le même espace de probabilité, pour tout n , une version de X_1, \dots, X_n et une version \widehat{G}_n de \widehat{G} telles que*

$$\mathbb{P} \left(\|\widehat{\alpha}_n - \widehat{G}_n\|_{\mathcal{C}} > c_\beta \left(\frac{\log n}{n} \right)^{1/(10VC-8)} (\log n)^{1/2} \right) < \frac{1}{n^\beta}.$$

Si la partition est de taille $m = m_n \rightarrow \infty$ alors un théorème similaire est valable, en tenant compte de m_n . Pour la VC-classe $\mathcal{C} = \mathcal{H}$ des demi-espaces de \mathbb{R}^d , utilisée au chapitre 2, on a $VC = d + 1$. Remarquer que si $\mathcal{A}_m \subset \mathcal{C}$ alors $\widehat{G} = 0$ sur \mathcal{A}_m . Puisque G est un processus additif, on a aussi

$$\widehat{G}(A) = \sum_{j \leq m} (G(A_j \cap A) - G(A_j)P(A | A_j))$$

donc, en posant

$$p_j = P(A_j), \quad q_j = P(A | A_j), \quad a_j = P(A \cap A_j) = p_j q_j, \quad G_j(A) = G(A_j \cap A) - q_j G(A_j)$$

il vient

$$\text{Var}(\widehat{G}(A)) = \text{Var}(\sum_{j \leq m} G_j(A)) = \sum_{j \leq m} \text{Var}(G_j(A)) + 2 \sum_{1 \leq i < j \leq m} \text{cov}(G_i(A), G_j(A)).$$

Puisque $\text{cov}(G(A), G(B)) = P(A \cap B) - P(A)P(B)$, on a d'une part

$$\begin{aligned} \text{Var}(G_j(A)) &= a_j(1 - a_j) + q_j^2 p_j(1 - p_j) - 2q_j(a_j - a_j p_j) \\ &= a_j(1 - a_j) + q_j a_j(1 - p_j) - 2q_j a_j(1 - p_j) \\ &= a_j - a_j^2 - a_j q_j + q_j a_j p_j \\ &= a_j(1 - q_j) \\ &= a_j - p_j q_j^2 \end{aligned}$$

et d'autre part, si $1 \leq i < j \leq m$,

$$\begin{aligned} \text{cov}(G_i(A), G_j(A)) &= \text{cov}(G(A_i \cap A) - q_i G(A_i), G(A_j \cap A) - q_j G(A_j)) \\ &= -a_i a_j - q_j(-a_i p_j) - q_i(-a_j p_i) + q_i q_j(-p_i p_j) \\ &= 0 \end{aligned}$$

d'où

$$\begin{aligned} \text{Var}(\widehat{G}(A)) &= \sum_{j \leq m} a_j - \sum_{j \leq m} p_j q_j^2 \\ &= P(A) - \sum_{j \leq m} P(A_j) P^2(A | A_j) \\ &\leq P(A) - \left(\sum_{j \leq m} P(A_j) P(A | A_j) \right)^2 \\ &= P(A) - P^2(A) = \text{Var}(G(A)) = \text{Var}(\alpha_n(A)). \end{aligned}$$

On a aussi $Var(\alpha_n^*(A)) = Var(\widehat{G}(A))$ avec le processus dominant dans $\widehat{\alpha}_n(A)$ donné par

$$\begin{aligned}\alpha_n^*(A) &= \sum_{j \leq m} (\alpha_n(A \cap A_j) - \alpha_n(A_j)P(A | A_j)) \\ &= \alpha_n(A) - \sum_{j \leq m} \alpha_n(A_j)P(A | A_j).\end{aligned}$$

La réduction de variance est donc de l'ordre de

$$\begin{aligned}Var(\alpha_n(A)) - Var(\alpha_n^*(A)) &= Var(G(A)) - Var(\widehat{G}(A)) \\ &= \sum_{j \leq m} P(A_j)P^2(A | A_j) - P^2(A) \\ &= \sum_{j \leq m} \frac{P^2(A \cap A_j)}{P(A_j)} - P(A) \sum_{j \leq m} P(A \cap A_j) \\ &= \sum_{j \leq m} P(A \cap A_j) (P(A | A_j) - P(A)).\end{aligned}$$

En particulier, pour chaque $A_j \subset A$ on a $P(A | A_j) = 1$ et on économise la variance

$$\begin{aligned}P(A_j) (1 - P(A)) &= P(A_j) (1 - P(A_j)) - P(A_j)P(A \setminus A_j) \\ &= Var(\mathbb{P}_n(A_j)) + cov(\mathbb{P}_n(A_j), \mathbb{P}_n(A \setminus A_j))\end{aligned}$$

du fait de la disparition de $\mathbb{P}_n(A_j)$ dans $\mathbb{P}_n(A) = \mathbb{P}_n(A_j) + \mathbb{P}_n(A \setminus A_j)$ quand on se sert de $\widehat{\mathbb{P}}_n(A) = P(A_{j,n}) + \widehat{\mathbb{P}}_n(A \setminus A_{j,n})$.

On obtient de même

$$\begin{aligned}Cov(\alpha_n^*(A), \alpha_n^*(B)) &= Cov(\widehat{G}(A), \widehat{G}(B)) \\ &= P(A \cap B) - \sum_{j \leq m} P(A_j)P(A | A_j)P(B | A_j)\end{aligned}$$

et on montre que, pour tout B_1, \dots, B_k il vient, en notant Σ la matrice de covariance de $(G(B_1), \dots, G(B_k))$ et $\widehat{\Sigma}$ celle de $(\widehat{G}(B_1), \dots, \widehat{G}(B_k))$,

$$\Sigma - \widehat{\Sigma} \text{ est semi-définie positive.}$$

4.5.3 Résultats

Supposons que $P(A_j)$ est connue pour une partition mesurable A_1, \dots, A_m de \mathbb{R}^d . Le cas générique est une partition selon une coordonnée de X

très bien connue, disons X^1 , ainsi $A_j = B_j \times \mathbb{R}^{d-1}$ avec $B_1, \dots, B_m =]-\infty, x_1], \dots,]x_1, \infty[$ une partition de \mathbb{R} en intervalles. Associons à chaque X_i le poids aléatoire

$$C_i = \sum_{j \leq m} \frac{P(A_j)}{n\mathbb{P}_n(A_j)} 1_{X_i \in A_j} = \frac{P(A_j)}{n\mathbb{P}_n(A_j)} \quad \text{pour } X_i \in A_j$$

ce qui donne bien $\sum_{i \leq n} C_i = 1$, et repondérons la fonction de répartition des projections directionnelles sous la forme

$$\widehat{F}_n(O, u)(y) = \sum_{i \leq n} C_i \langle X_i - O, u \rangle.$$

Ainsi \widehat{F}_n est une fonction de répartition associée à une mesure discrète du type évoquée dans la section précédente et la surface quantile modifiée est

$$\widehat{Q}_n(O, \alpha) = \left\{ \widehat{Q}_n(O, u, \alpha) : u \in \mathbb{S}_{d-1} \right\}$$

où

$$\widehat{Q}_n(O, u, \alpha) = O + \widehat{Y}_n(O, u, \alpha)u, \quad \widehat{Y}_n(O, u, \alpha) = \widehat{F}_n^{-1}(O, u)(\alpha).$$

Considérons le processus d'écart des surfaces quantiles théoriques et empiriques,

$$\widehat{Y}_n(u, \alpha) = \widehat{Y}_n(O, u, \alpha) - Y(O, u, \alpha), \quad u \in \mathbb{S}_{d-1}, \quad \alpha \in \Delta,$$

qui de nouveau ne dépend pas de $O \in \mathbb{R}^d$.

Théorème 4.26. *Pour P vérifiant la condition (\mathbf{H}_1) il vient*

$$\lim_{n \rightarrow \infty} \sup_{\alpha \in \Delta} \sup_{u \in \mathbb{S}_{d-1}} |\widehat{Y}_n(u, \alpha)| = 0 \quad p.s.$$

Si de plus, P vérifie (\mathbf{H}_2) alors $\sqrt{n}|\widehat{Y}_n|$ converge en loi vers le processus \widehat{G} indexé par $\mathbb{S}_{d-1} \times \Delta$, gaussien, centré, et de covariance

$$\begin{aligned} & \text{Cov}(\widehat{G}(u_1, \alpha_1), \widehat{G}(u_2, \alpha_2)) \\ &= \frac{1}{g(u_1, \alpha_1)g(u_2, \alpha_2)} \left(P(H(O, u_1, \alpha_1) \cap H(O, u_2, \alpha_2)) - \right. \\ & \left. \sum_{j \leq m} P(A_j)P(H(O, u_1, \alpha_1) | A_j)P(H(O, u_2, \alpha_2) | A_j) \right). \end{aligned}$$

4.5.4 Simulations

Dans cette section on se place dans le cas où une information auxiliaire sur la loi P est disponible sous la forme de la connaissance de P sur une partition en peu d'intervalles. On se propose de visualiser, à travers des simulations dans \mathbb{R}^2 , le gain en terme de précision dans l'estimation avec très peu de points. La « vrai » surface sera visualisée à l'aide de la surface empirique avec un très grand nombre de points N .

Dans les graphiques qui suivent on choisit comme convention de dessiner :

- en vert $Q(O, \alpha)$: la surface quantile « empirique » obtenue avec $N = 10^5$ points.
- en rouge $Q_1(O, \alpha)$: la surface quantile empirique obtenue avec petit nombre de points $N_1 \ll N$.
- en noire $Q_{aux}(O, \alpha)$: la surface quantile empirique obtenue avec les mêmes $N_1 \ll N$ points, en tenant compte de l'information auxiliaire.

Partition 1 : comme première information auxiliaire, on considère que P est connue sur $A = \{I_k, 1 \leq k \leq n\}$ une partition de taille n de l'axe des abscisses en n intervalles, i.e. on suppose connues les probabilités $P(I_k \times \mathbb{R})$ pour $k = 1, \dots, n$.

-Loi uniforme sur $[0, 1] \times [0, 1]$ On pose

$$A = \{I_k \times \mathbb{R}, 1 \leq k \leq 4\}$$

avec $I_1 = [0, 1/4]$, $I_2 = [1/4, 1/2]$, $I_3 = [1/2, 3/4]$, $I_4 = [3/4, 1]$, et

$$P(I_k \times \mathbb{R}) = \frac{1}{4}, \quad 1 \leq k \leq 4$$

- La figure 4.17 est réalisée avec $N_1 = 40$ points.
- La figure 4.18 est réalisée avec $N_1 = 75$ points.
- La figure 4.19 est réalisée avec $N_1 = 150$ points.

Dans chaque figure la surface en noire $Q_{aux}(O, \alpha)$ est plus proche (en moyenne) de $Q(O, \alpha)$ en vert que la surface rouge $Q_1(O, \alpha)$, ainsi l'erreur d'approximation de la surface quantile est moindre en tenant compte de l'information auxiliaire. Néanmoins, à cause de la nature de cette information qui concerne uniquement la première coordonnée, l'amélioration ne se fait que dans les directions proches de $u = (\pm 1, 0)$, et aucune amélioration n'est à noter dans les directions $u = (0, \pm 1)$.

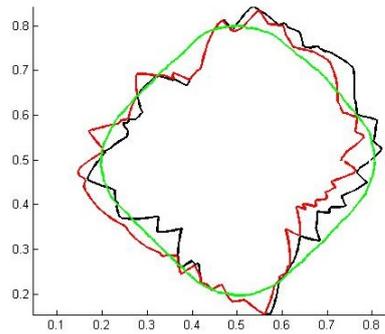


FIGURE 4.17 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0.5,0.5)$ avec $N_1 = 40$

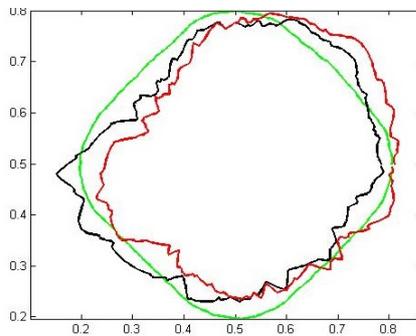


FIGURE 4.18 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0.5,0.5)$ avec $N_1 = 75$

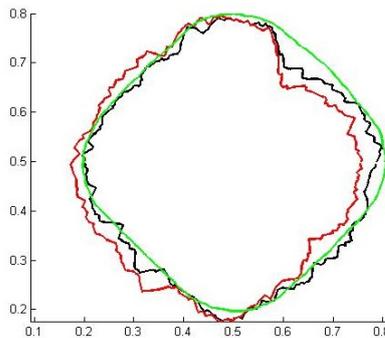


FIGURE 4.19 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0.5,0.5)$ avec $N_1 = 150$

-Loi normale $N(0_2, I_2)$ on pose $\psi = F_{N(0,1)}^{-1}(3/4) \simeq 0.6743$ et

$$A = \{I_k \times \mathbb{R}, 1 \leq k \leq 4\}$$

avec $I_1 = [-\infty, -\psi]$, $I_2 = [-\psi, 0]$, $I_3 = [0, \psi]$, $I_4 = [\psi, +\infty]$, et

$$P(I_k \times \mathbb{R}) = \frac{1}{4}, \quad 1 \leq k \leq 4$$

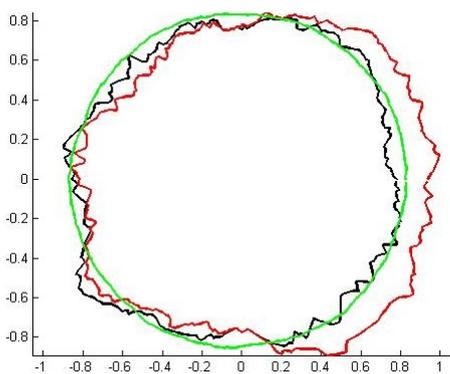
- La figure 4.21 est réalisée avec $N_1 = 40$ points.
- La figure 4.22 est réalisée avec $N_1 = 75$ points.
- La figure 4.23 est réalisée avec $N_1 = 150$ points.

Comme dans le cas des uniformes, on remarque que la surface en noire $Q_{aux}(O, \alpha)$ est plus proche (en moyenne) de $Q(O, \alpha)$ en vert que la surface rouge $Q_1(O, \alpha)$, avec une amélioration concentrée dans les directions proches de $u = (\pm 1, 0)$, et aucune amélioration n'est à noter dans les directions $u = (0, \pm 1)$, à cause de la non information supplémentaire dans ces direction.

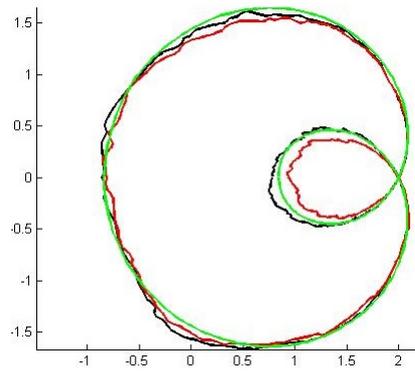
Remarque 4.27. Dans toutes les simulations, on a choisi de montrer uniquement les simulation pour un seul choix de l'observateur O , ceci est du au fait que l'erreur de l'approximation ne dépend pas du choix de l'observateur. En effet, pour $O, O' \in \mathbb{R}^d$, $u \in \mathbb{S}_{d-1}$ et $\alpha \in \Delta$ on a

$$\begin{aligned} Y_n(O, u, \alpha) - \widehat{Y}_n(O, u, \alpha) &= Y_n(O', u, \alpha) + \langle O' - O, u \rangle - \left(\widehat{Y}_n(O', u, \alpha) + \langle O' - O, u \rangle \right) \\ &= Y_n(O', u, \alpha) - \widehat{Y}_n(O', u, \alpha). \end{aligned}$$

Voici une simulation pour une loi normale $N(0_2, I_2)$ avec $N_1 = 150$ points et deux observateurs $O = (0, 0)$ et $O' = (2, 0)$.



(a) $Q(O, 0.8)$, $Q_1(O, 0.8)$ et $Q_{aux}(O, 0.8)$ avec $O = (0, 0)$



(b) $Q(O, 0.8)$, $Q_1(O, 0.8)$ et $Q_{aux}(O, 0.8)$ avec $O = (2, 0)$

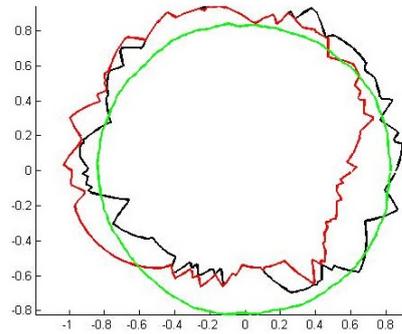


FIGURE 4.21 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0,0)$ avec $N_1 = 40$

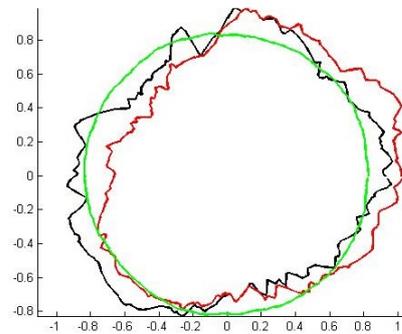


FIGURE 4.22 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0,0)$ avec $N_1 = 75$

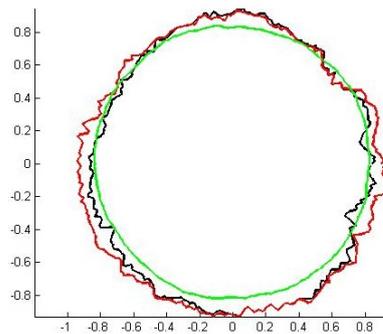


FIGURE 4.23 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0,0)$ avec $N_1 = 150$

Partition 2 : cette fois on considère que P est connue sur une partition de taille $n \times m$ avec $A = \{I_k \times J_l, 1 \leq k \leq n, 1 \leq l \leq m\}$ où $\{I_k, 1 \leq k \leq n\}$ est une partition de l'axe des abscisses en n intervalles et $\{J_l, 1 \leq l \leq m\}$ est une partition de l'axe des ordonnées en m intervalles, i.e. on suppose connue les probabilités $P(I_k \times J_l)$ pour $1 \leq k \leq n, 1 \leq l \leq m$.

-Loi uniforme sur $[0, 1] \times [0, 1]$

Dans ce cas, on pose

$$A = \{I_k \times I_l, 1 \leq k \leq 4, 1 \leq l \leq 4\}$$

avec $I_1 = [0, 1/4]$, $I_2 = [1/4, 1/2]$, $I_3 = [1/2, 3/4]$, $I_4 = [3/4, 1]$, avec l'information auxiliaire suivante :

$$P(I_k \times I_l) = \frac{1}{16}, \quad 1 \leq k \leq 4, 1 \leq l \leq 4$$

- La figure 4.24 est réalisée avec $N_1 = 40$ points.
- La figure 4.25 est réalisée avec $N_1 = 75$ points.
- La figure 4.26 est réalisée avec $N_1 = 150$ points.

-Loi normale $N(0, I_2)$

Dans ce cas, on pose $\psi = F_{N(0,1)}^{-1}(3/4) \simeq 0.6743$ et

$$A = \{I_k \times I_l, 1 \leq k \leq 4, 1 \leq l \leq 4\}$$

avec $I_1 = [-\infty, -\psi]$, $I_2 = [-\psi, 0]$, $I_3 = [0, \psi]$, $I_4 = [\psi, +\infty]$, avec

$$P(I_k \times I_l) = \frac{1}{16}, \quad 1 \leq k \leq 4, 1 \leq l \leq 4$$

- La figure 4.27 est réalisée avec $N_1 = 40$ points.
- La figure 4.28 est réalisée avec $N_1 = 75$ points.
- La figure 4.29 est réalisée avec $N_1 = 150$ points.

Commençons par remarquer que comme dans les cas précédents les surfaces quantiles $Q_{aux}(O, \alpha)$ sont clairement plus proches de $Q(O, u, \alpha)$ que celles sans prise en compte de l'information auxiliaire, Ainsi, l'erreur d'estimation est moindre.

À la différence avec le cas précédent, cette partition étant plus précise et couvrant bien l'espace, l'estimation est améliorée quelle que soit la direction, cela a permis dans certains cas de replacer complètement la surface quantile, en lui faisant épouser la forme de $Q(O, \alpha)$.

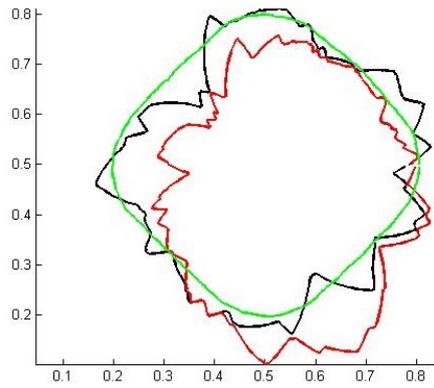


FIGURE 4.24 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0.5,0.5)$ avec $N_1 = 40$

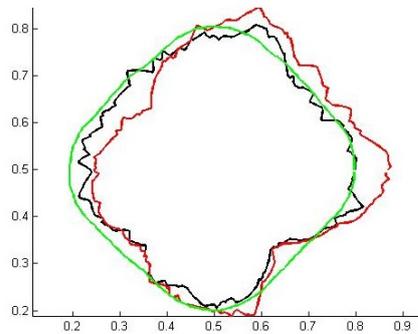


FIGURE 4.25 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0.5,0.5)$ avec $N_1 = 75$

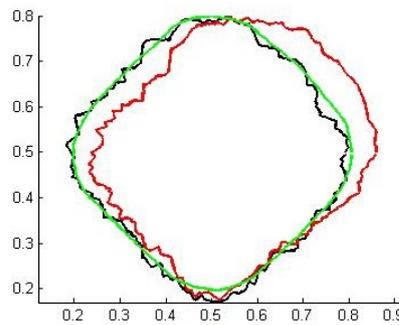


FIGURE 4.26 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0.5,0.5)$ avec $N_1 = 150$

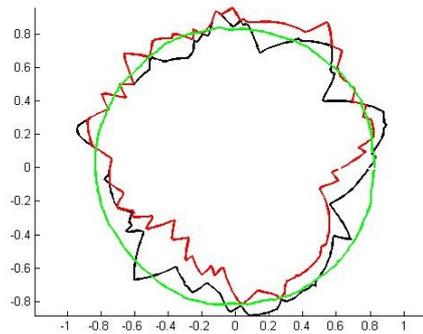


FIGURE 4.27 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0.5,0.5)$ avec $N_1 = 40$

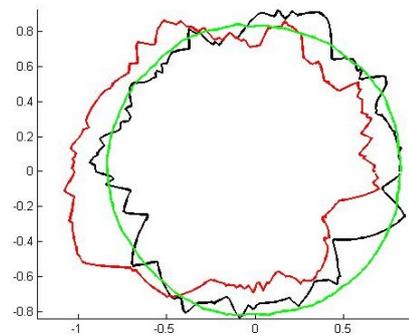


FIGURE 4.28 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0,0)$ avec $N_1 = 75$

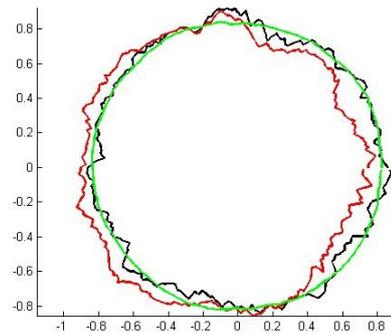


FIGURE 4.29 – Surfaces Quantiles $Q(O,0.8)$, $Q_1(O,0.8)$ et $Q_{aux}(O,0.8)$ avec $O = (0,0)$ avec $N_1 = 150$

-Loi normale $N(0_2, \Sigma)$

Pour la loi normale centré et de matrice de variance-covariance

$$\Sigma = \begin{pmatrix} 1 & -1 \\ -1 & 3/2 \end{pmatrix}$$

on se place dans le cas où l'information auxiliaire est donnée uniquement sur l'axe des abscisse, ainsi, comme précédemment on pose $\psi = F_{N(0,1)}^{-1}(3/4) \simeq 0.6743$ et

$$A = \{I_k \times \mathbb{R}, 1 \leq k \leq 4\}$$

avec $I_1 = [-\infty, -\psi]$, $I_2 = [-\psi, 0]$, $I_3 = [0, \psi]$, $I_4 = [\psi, +\infty]$, et

$$P(I_k \times \mathbb{R}) = \frac{1}{4}, \quad 1 \leq k \leq 4$$

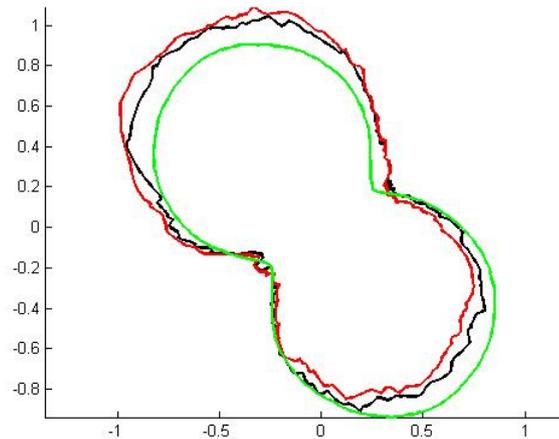


FIGURE 4.30 – Surfaces Quantiles $Q(O, 0.8)$, $Q_1(O, 0.8)$ et $Q_{aux}(O, 0.8)$ avec $O = (0, 0)$ avec $N_1 = 150$

Dans ce cas aussi, les surfaces quantiles $Q_{aux}(O, \alpha)$ sont clairement plus proches de $Q(O, u, \alpha)$ que celles construites uniquement avec l'échantillon de taille N_1 . Néanmoins l'amélioration de l'estimation de $Q(O, \alpha)$ n'est notable que dans les directions proche de $(\pm 1, 0)$ et elle est quasi nulle ailleurs.

Combinaison de plusieurs partitions : dans cette partie, on s'intéresse au cas où plusieurs informations sont disponibles, dans cet exemple on dispose de deux informations apportées par deux partitions A et B de direction et de taille distinctes. Pour la loi normale $N(0_2, I_2)$ on pose

$$A = \{I_k \times \mathbb{R}, 1 \leq k \leq 8\}$$

avec

$$P(I_k \times \mathbb{R}) = \frac{1}{8}, \quad 1 \leq k \leq 8$$

et

$$B = \{\mathbb{R} \times J_l, 1 \leq l \leq 4\}$$

avec

$$P(\mathbb{R} \times J_l) = \frac{1}{4}, \quad 1 \leq l \leq 4.$$

On note

- $Q_{Aux}^A(O, \alpha)$ la surface quantile empirique construite avec l'échantillon (X_i) de taille N_1 avec les poids aléatoires associés (C_i^A) .
- $Q_{Aux}^B(O, \alpha)$ la surface quantile empirique construite avec l'échantillon (X_i) de taille N_1 avec les poids aléatoires associés (C_i^B) .
- $Q_{Aux}^\delta(O, \alpha)$ pour $\delta \in [0, 1]$ la surface quantile empirique construite avec l'échantillon (X_i) de taille N_1 avec les poids aléatoires associés $(\delta C_i^A + (1 - \delta)C_i^B)$.

Tout d'abord, comme dans toutes les simulation précédentes, la diminution de l'erreur est fonction de la nature de la partition, ainsi dans la figure 4.31a qui représente en noire la surface quantile $Q_{Aux}^A(O, \alpha)$ l'amélioration de l'estimation est principalement le long de l'axe des abscisses, puis dans la figure 4.31b en noire la surface quantile $Q_{Aux}^B(O, \alpha)$ qui améliore l'estimation le long de l'axe des ordonnées.

Les deux informations étant sur deux directions distinctes, on souhaite tirer profit des deux partitions, en associant un « indice de confiance » δ à une des partition et $1 - \delta$ à l'autre, de la façon suivante : on associe à l'échantillon de taille N_1 un nouveau poids C_i^δ combinaison linéaire de C_i^A et C_i^B , ainsi on pose

$$C_i^\delta = \delta C_i^A + (1 - \delta)C_i^B.$$

Ce nouveau paramètre peut être choisi en fonction de la richesse ou la finition de l'information dans chaque direction , par exemple pour $\delta = 1$ on ne tient pas compte de l'information apporter par la partition B et inversement on ne prend en compte que de l'information apportée par B lorsque $\delta = 0$. Le choix le plus intéressant serait de prendre δ dans $(0, 1)$ à calibrer, permettant ainsi de tenir compte des deux informations.

Dans les figures 4.31c et 4.31d

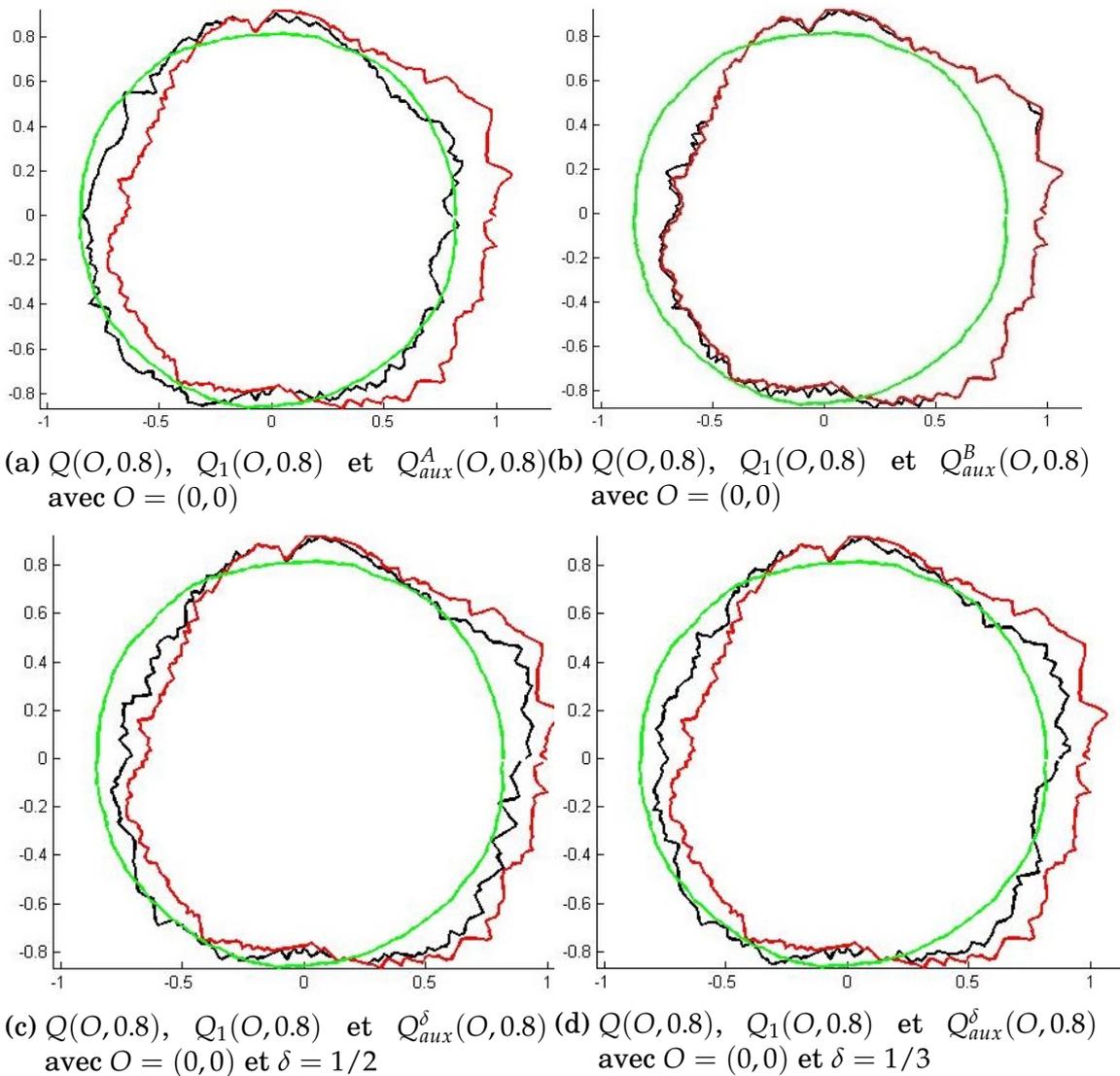


FIGURE 4.31 – $Q(O, 0.8)$, $Q_1(O, 0.8)$, $Q_{aux}^A(O, 0.8)$, $Q_{aux}^B(O, 0.8)$ et $Q_{aux}^\delta(O, 0.8)$

Remarque 4.28. *Comme on vient de l'observer, l'information auxiliaire a permis d'améliorer l'estimation des surfaces quantile, cette amélioration est naturellement liée au type d'information disponible. Un autre type d'information auxiliaire pourrait être la connaissance de certaines surfaces quantiles pour des niveaux α_1, α_2 , dans le but d'améliorer l'approximation des surfaces de niveau $\alpha \neq \alpha_1, \alpha_2$, néanmoins, ce type d'information auxiliaire ne rentre pas dans le cadre d'une partition finie, car une surface quantile est décrite par une infinité non dénombrable de demi-espaces indexés par S_{d-1} .*

Bibliographie

- [Ale84] Alexander, K. S. (1984); “Probability inequalities for empirical processes and a law of the iterated logarithm”. The Annals of Probability, 12(4) :pp. 1041–1067.
- [ABH⁺72] Andrews, D. F., Bickel, P. J., Hampel, F. R., Huber, P., H. Rogers, J. W. and Tukey, J. W. (1972); “Robust estimates of location. survey and advances.” Princeton University Press, Princeton, N.J., 179(4072) :pp. 469–470.
- [ACZ06] Arcones, M., Cui, H. and Zuo, Y. (2006); “Empirical depth processes”. TEST, 15(1) :pp. 151–177. ISSN 1133-0686.
- [Bah66] Bahadur, R. R. (1966); “A note on quantiles in large samples”. The Annals of Mathematical Statistics, 37(3) :pp. 577–580.
- [Bar76] Barnett, V. (1976); “The ordering of multivariate data”. Journal of the Royal Statistical Society. Series A (General), 139(3) :pp. 318–355.
- [BMV99] Beirlant, J., Mason, D. M. and Vynckier, C. (1999); “Goodness-of-fit analysis for multivariate normality based on generalized quantiles”. Computational Statistics and Data Analysis, 30(2) :pp. 119–142.
- [Ber14] Berthet, P. (2014); “The strong belief empirical process”. preprint.
- [BE12] Berthet, P. and Einmahl, J. H. J. (2012); “Central limit theorems for level set estimators”. preprint.
- [BEN05] Berthet, P. and El-Nouty, C. (2005); “Almost sure asymptotic behaviour of the shorth estimators”. Mathematical Methods of Statistics, 14(4) :pp. 379–403.
- [BM06a] Berthet, P. and Mason, D. M. (2006); “Revisiting two strong approximation results of dudley and philipp”. High Dimensional Probability, 51 :pp. 155–172.

- [BM06b] Berthet, P. and Mason, D. M. (2006); “Strong invariance principles for empirical processes indexed by functions”. preprint.
- [BLG14] Boissard, E. and Le Gouic, T. (2014); “On the mean speed of convergence of empirical and occupation measures in Wasserstein distance”. Annales de l’Institut Henri Poincaré, Probabilités et Statistiques, 50(2) :pp. 539–563.
- [BC88] Breckling, J. and Chambers, R. (1988); “ M -quantiles”. Biometrika, 75(4) :pp. 761–771.
- [Can17] Cantelli, F. P. (1917); “Sulla probabilità come limite della frequenza”. Atti della Accademia Nazionale dei Lincei, 26(1) :pp. 39–45.
- [Cha96] Chaudhuri, P. (1996); “On a geometric notion of quantiles for multivariate data”. Journal of the American Statistical Association, 91(434) :pp. 862–872.
- [CCHM86] Csörgő, M., Csörgő, S., Horváth, L. and Mason, D. M. (1986); “Weighted empirical and quantile processes”. The Annals of Probability, 14(1) :pp. 31–85.
- [CR78] Csorgo, M. and Revesz, P. (1978); “Strong approximations of the quantile process”. The Annals of Statistics, 6(4) :pp. 882–894.
- [CR81] Csörgő, M. and Révész, P. (1981); Strong approximations in probability and statistics. Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London.
- [CM88] Cuesta, J. A. and Matrán, C. (1988); “The strong law of large numbers for k -means and best possible nets of Banach valued random variables”. Probability Theory and Related Fields, 78(4) :pp. 523–534.
- [CAF07] Cuesta-Albertos, J. A. and Fraiman, R. (2007); “Impartial trimmed k -means for functional data”. Computational Statistics & Data Analysis, 51(10) :pp. 4864–4877.
- [CAGM97] Cuesta-Albertos, J. A., Gordaliza, A. and Matrán, C. (1997); “Trimmed k -means : an attempt to robustify quantizers”. The Annals of Statistics, 25(2) :pp. 553–576.
- [Cue90] Cuevas, A. (1990); “On pattern analysis in the nonconvex case”. Kybernetes, 19(6) :pp. 26–33.
- [Deh92] Deheuvels, P. (1992); “Pointwise Bahadur-Kiefer-type theorems. I”. In “Probability theory and applications”, Math. Appl., volume 80, pp. 235–255. Kluwer Acad. Publ., Dordrecht.

- [Deh00] Deheuvels, P. (2000); “Strong approximation of quantile processes by iterated Kiefer processes”. The Annals of Probability, 28(2) :pp. 909–945.
- [DM90] Deheuvels, P. and Mason, D. M. (1990); “Bahadur-kiefer-type processes”. The Annals of Probability, 18(2) :pp. 669–697.
- [DG92] Donoho, D. L. and Gasko, M. (1992); “Breakdown properties of location estimates based on halfspace depth and projected outlyingness”. The Annals of Statistics, 20(4) :pp. 1803–1827.
- [Don51] Donsker, M. D. (1951); “An invariance principle for certain probability limit theorems”. Memoirs of the American Mathematical Society, 1951(6) :p. 12.
- [DKW56] Dvoretzky, A., K., J. and Wolfowitz, J. (1956); “Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator”. The Annals of Mathematical Statistics, 27 :pp. 642–669.
- [Ein96] Einmahl, J. H. J. (1996); “A short and elementary proof of the main Bahadur-Kiefer theorem”. The Annals of Probability, 24(1) :pp. 526–531.
- [EGS10] Einmahl, J. H. J., Gantner, M. and Sawitzki, G. (2010); “Asymptotics of the shorth plot”. Journal of Statistical Planning and Inference, 140(11) :pp. 3003–3012. ISSN 0378-3758.
- [EM92] Einmahl, J. H. J. and Mason, D. M. (1992); “Generalized quantile processes”. The Annals of Statistics, 20(2) :pp. pp. 1062–1078.
- [EM00] Einmahl, U. and Mason, D. M. (2000); “An empirical process approach to the uniform consistency of kernel-type function estimators”. Journal of Theoretical Probability, 13(1) :pp. 1–37.
- [EK46] Erdős, P. and Kac, M. (1946); “On certain limit theorems of the theory of probability”. Bulletin of the American Mathematical Society, 52 :pp. 292–302.
- [GJKL14] Gamboa, F., Janon, A., Klein, T. and Lagnoux, A. (2014); “Sensitivity analysis for multidimensional and functional outputs”. Electronic Journal of Statistics, 8(1) :pp. 575 – 603.
- [GG01] Ginè, E. and Guillou, A. (2001); “On consistency of kernel density estimators for randomly censored data : rates holding uniformly over adaptive intervals”. Annales de l’Institut Henri Poincaré (B) Probability and Statistics, 37(4) :pp. 503 – 522.

- [Gli33] Glivenko, V. (1933); “Sulla determinazione empirica della legge di probabilita”. Giornale dell’Istituto Italiano degli Attuari, 4 :pp. 92–99.
- [HPŠ10] Hallin, M., Paindaveine, D. and Šiman, M. (2010); “Multivariate quantiles and multiple-output regression quantiles : From L_1 -optimization to halfspace depth”. The Annals of Statistics, 38(2) :pp. 635–669.
- [HW97] He, X. and Wang, G. (1997); “Convergence of depth contours for multivariate datasets”. The Annals of Statistics, 25(2) :pp. 495–504.
- [HNO92] Hettmansperger, T. P., Nyblom, J. and Oja, H. (1992); “On multivariate notions of sign and rank”. In “ L_1 -statistical analysis and related methods (Neuchâtel, 1992)”, pp. 267–278. North-Holland, Amsterdam.
- [Khi33] Khintchine, A. (1933); Asymptotische Gesetze der Wahrscheinlichkeitsrechnung, volume 4. Springer-Verlag.
- [Kie67] Kiefer, J. (1967); “On bahadur’s representation of sample quantiles”. The Annals of Mathematical Statistics, 38(5) :pp. 1323–1342.
- [Kie70] Kiefer, J. (1970); “Deviations between the sample quantile process and the sample df”. In “Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)”, pp. 299–319. Cambridge Univ. Press, London.
- [Kim00] Kim, J. (2000); “Rate of convergence of depth contours : with application to a multivariate metrically trimmed mean”. Statistics & Probability Letters, 49(4) :pp. 393–400.
- [Kol31] Kolmogorov, A. N. (1931); “Eine verallgemeinerung des laplace-liapounoffschen satzes.” Bulletin of the Academy of Sciences of the USSR, 1931(7) :pp. 959–962.
- [Kol33] Kolmogorov, A. N. (1933); “Sulla determinazione empirica di una legge di distribuzione”. Giornale dell’Istituto Italiano degli Attuari, 4 :pp. 83–91.
- [Kol97] Koltchinskii, V. I. (1997); “ M -estimation, convexity and quantiles”. The Annals of Statistics, 25(2) :pp. 435–477.
- [KM12] Kong, L. and Mizera, I. (2012); “Quantile tomography : using quantiles with multivariate data”. Statistica Sinica, 22(4) :pp. 1589–1610.

- [KM97] Koshevoy, G. and Mosler, K. (1997); “Zonoid trimming for multivariate distributions”. The Annals of Statistics, 25(5) :pp. 1998–2017.
- [LG13] Le Gouic, T. (2013); “Localisation de masse et espaces de Wasserstein”. PhD Thesis.
- [Liu88] Liu, R. Y. (1988); “On a notion of simplicial depth”. Proceedings of the National Academy of Science USA, 85(6) :pp. 1732–1734.
- [LS93] Liu, R. Y. and Singh, K. (1993); “A quality index based on data depth and multivariate rank tests”. Journal of the American Statistical Association, 88(421) :pp. 252–260.
- [Mas02] Massé, J.-C. (2002); “Asymptotics for the Tukey median”. Journal of Multivariate Analysis, 81(2) :pp. 286–300.
- [Mog79] Mogul’skii, A. A. (1979); “The law of the iterated logarithm in Chung’s form for function spaces”. Teoriya Veroyatnostei i ee Primeneniya, 24(2) :pp. 399–407.
- [Nol91] Nolan, D. (1991); “The excess-mass ellipsoid”. Journal of Multivariate Analysis, 39(2) :pp. 348–371.
- [Nol92] Nolan, D. (1992); “Asymptotics for multivariate trimming”. Stochastic Processes and their Applications, 42(1) :pp. 157–169.
- [Nol99] Nolan, D. (1999); “On min-max majority and deepest points”. Statistics & Probability Letters, 43(4) :pp. 325–333.
- [Pol95] Polonik, W. (1995); “Measuring mass concentrations and estimating density contour clusters-an excess mass approach”. Ann. Statist., 23(3) :pp. 855–881.
- [Pol97a] Polonik, W. (1997); “Minimum volume sets and generalized quantile processes”. Stochastic Processes and their Applications, 69(1) :pp. 1 – 24.
- [Pol97b] Polonik, W. (1997); “Minimum volume sets in statistics : Recent developments”. In R. Klar and O. Opitz, editors, “Classification and Knowledge Organization”, Studies in Classification, Data Analysis, and Knowledge Organization, pp. 187–194. Springer Berlin Heidelberg. ISBN 978-3-540-62981-8.
- [RK11] Rainer, D. and Karl, M. (2011); “Weighted-mean trimming of multivariate data”. Journal of Multivariate Analysis, 102(3) :pp. 405 – 421.
- [Ser02] Serfling, R. (2002); “Generalized quantile processes based on multivariate depth functions, with applications in nonparametric

- tric multivariate analysis”. Journal of Multivariate Analysis, 83(1) :pp. 232 – 247.
- [SW86] Shorack, G. R. and Wellner, J. A. (1986); Empirical processes with applications to statistics. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.
- [Str64] Strassen, V. (1964); “An invariance principle for the law of the iterated logarithm”. zeitschrift für wahrscheinlichkeitstheorie und verwandte gebiete, 3 :pp. 211–226.
- [Tal94] Talagrand, M. (1994); “Sharper bounds for gaussian and empirical processes”. The Annals of Probability, 22(1) :pp. 28–76.
- [Tuk75] Tukey, J. W. (1975); “Mathematics and the picturing of data”. In “Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2”, pp. 523–531. Canad. Math. Congress, Montreal, Que.
- [VDVW00] Van Der Vaart, A. and Wellner, J. (2000); Weak Convergence and Empirical Processes : With Applications to Statistics. Springer Series in Statistics. Springer. ISBN 0387946403.
- [VZ00] Vardi, Y. and Zhang, C. (2000); “The multivariate L_1 -median and associated data depth”. Proceedings of the National Academy of Sciences of the United States of America, 97(4) :pp. 1423–1426.
- [Wei08] Wei, Y. (2008); “An approach to multivariate covariate-dependent quantile contours with application to bivariate conditional growth charts”. Journal of the American Statistical Association, 103(481) :pp. 397–409.
- [ZS00a] Zuo, Y. and Serfling, R. (2000); “General notions of statistical depth function”. The Annals of Statistics, 28(2) :pp. 461–482.
- [ZS00b] Zuo, Y. and Serfling, R. (2000); “On the performance of some robust nonparametric location measures relative to a general notion of multivariate symmetry”. Journal of Statistical Planning and Inference, 84(1-2) :pp. 55–79.
- [ZS00c] Zuo, Y. and Serfling, R. (2000); “Structural properties and convergence results for contours of sample statistical depth functions”. The Annals of Statistics, 28(2) :pp. 483–499.

Résumé

Dans la thèse on introduit et on étudie une généralisation spatiale sur \mathbb{R}^d du quantile réel usuel sous la forme d'une surface quantile via des formes φ et d'un point d'observation O . Notre point de départ est de simplement admettre la subjectivité due à l'absence de relation d'ordre totale dans \mathbb{R}^d et donc de développer une vision locale et directionnelle des données. Ainsi, les observations seront ordonnées du point de vue d'un observateur se trouvant à un point $O \in \mathbb{R}^d$.

Dans le chapitre 2, on introduit la notion du quantile vue d'un observateur O dans la direction $u \in \mathbb{S}_{d-1}$ et de niveau α via des demi-espaces orthogonaux à chaque direction d'observation. Ce choix de classe implique que les résultats de convergence ne dépendent pas du choix de O . Sous des hypothèses minimales de régularité, l'ensemble des points quantile vue de O définit une surface fermée. Sous hypothèses minimales, on établit pour les surfaces quantile empiriques associées les théorèmes limites uniformément en le niveau de quantile et la direction d'observation, avec vitesses asymptotiques et bornes d'approximation non-asymptotiques. Principalement la LGNU, la LLI, le TCLU, le principe d'invariance fort uniforme puis enfin l'approximation du type Bahadur-Kiefer uniforme, et avec vitesse d'approximation.

Dans le chapitre 3, on étend les résultats du chapitre précédent au cas où les formes φ sont prises dans une classe plus générale (fonctions, surfaces, projections géodésiques, etc) que des demi-espaces qui correspondent à des projections orthogonales par direction. Dans ce cadre plus général, les résultats dépendent fortement du choix de O , et c'est ce qui permet de tirer des interprétations statistiques.

Dans le chapitre 4, des conséquences méthodologiques en statistique inférentielle sont tirées. Tout d'abord on introduit une nouvelle notion de champ de profondeurs directionnelles baptisée champ d'altitude. Ensuite, on définit une notion de distance entre lois de probabilité, basée sur la comparaison des deux collections de surfaces quantile du type Gini-Lorentz. La convergence avec vitesse des mesures empiriques pour cette distance quantile, permet de construire différents tests en contrôlant leurs niveaux et leurs puissances. Enfin, on donne une version des résultats dans le cas où une information auxiliaire est disponible sur une ou plusieurs coordonnées sous la forme de la connaissance exacte de la loi sur une partition finie.

Abstract

The main issue of the thesis is the development of spatial generalizations on \mathbb{R}^d of the usual real quantile. Facing the usual fact that \mathbb{R}^d is not naturally ordered, our idea is to simply admit subjectivity and thus to define a local viewpoint rather than a global one, anchored at some point of reference O and arbitrary shape φ with the motivation of crossing information gathered by changing viewpoint O , shape φ and α -th order of quantile.

In Chapter 2, we study the spatial quantile points seen from an observer O in a direction $u \in \mathbb{S}_{d-1}$ of level α through the class of the half-spaces orthogonal to the direction u . This choice implies that the convergence theorems do not depend on the choice of O . Under minimal regularity assumptions, the set of all quantile points seen from O is a closed surface. Under minimal assumptions, we establish for the associated empirical quantile surfaces the convergence theorems uniformly on the quantile level and the observation direction with the asymptotic speed and non-asymptotic bounds of approximation. Mainly, we establish the ULLN, the ULIL, the UCLT, the uniform strong invariance principle and finally the Bahadur-Kiefer type embedding, with the approximation speed rate.

In Chapter 3, all the results of the previous chapter are extended to the case where the shapes φ are taken in a class more general (functions, surfaces, geodesic projections, etc) than orthogonal projections (half-spaces). In this general setting, the results depend strongly on the choice of O . It is this dependence which permit to draw statistical interpretations : modes detection, mass localization, etc.

In Chapter 4, some methodological consequences in inferential statistics are drawn. First we introduce a new concept of directional depth fields called altitude fields. In a second application is defined a new distances between probability distributions, based on the comparison of two collections of quantile surfaces, which are indexes of the type Gini-Lorentz. The convergence with speed of the empirical quantile measures for these distances, can build different tests with control of their level and their power. A third use of the quantile surfaces is for the case where $\alpha = 1/2$. Finally, we give a version of our theorems in the case where auxiliary information is available on one or more coordinates of the random variable. By assuming known the probability of the elements of a finite partition, the asymptotic variance of the limiting process decreases and the simulations with few points clearly shows the reframe of the estimated surfaces to the real ones.