On Resampling Schemes for Uniform Polytopes

Weinan Qi

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the degree of Master of Science in Mathematics[1]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

[1]The M.Sc. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Abstract

The convex hull of a sample is used to approximate the support of the underlying distribution. This approximation has many practical implications in real life. For example, approximating the boundary of a finite set is used by many authors in environmental studies and medical research. To approximate the functionals of convex hulls, asymptotic theory plays a crucial role. Unfortunately, the asymptotic results are mostly very complicated. To address this complication, we suggest a consistent bootstrapping scheme for certain cases. Our resampling technique is used for both semi-parametric and non-parametric cases. Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random points uniformly distributed on an unknown convex set. Our bootstrapping scheme relies on resampling uniformly from the convex hull of $X_1, X_2, \ldots, X_n$. In this thesis, we study the asymptotic consistency of certain functionals of convex hulls. In particular, we apply our bootstrapping technique to the Hausdorff distance between the actual convex set and its estimator. We also provide a conjecture for the application of our bootstrapping scheme to Gaussian polytopes. Moreover, some other relevant consistency results for the regular bootstrap are developed.

Dedications

This research is dedicated to my great parents, who have given me the support and love in countless ways.

Acknowledgement

First, I would like to thank my parents, for encouraging me to pursue my education. Without their unconditional support and faith in me, I would not be studying at the University of Ottawa.

Next, I would like to express my sincere gratitude to my supervisor, Dr. Mahmoud Zarepour, for his expertise, assistance, encouragement and patience throughout the process of this thesis. His guidance helped me during research and completing this thesis. It is a great honour to work under his supervision.

I would also like to show my appreciation to Dr. Imre Bárány, from Alfréd Rényi Mathematical Institute, for his kind help.

I would like to show my deepest thanks to my girlfriend, Yan Su, for her love, support and understanding.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

In plug-in principle, the regular bootstrap method always resamples from the sample. This is because the empirical distribution approximates the underlying distribution. When dealing with high-dimensional sample points, the convex hull of the sample, i.e. the random polytopes, is used as a good estimator for the support of the underlying distributions. Because of the complexity of the asymptotic distribution for different functionals of convex hulls, we develop a new bootstrapping scheme based on stochastic properties of random polytopes and study its asymptotic consistency. Since our main results are based on the existing theory of random polytopes, it is worth developing this asymptotic theory. These investigations enable us to know when and how the suggested resampling technique can be used.

1.2 Origin and application of random polytopes

In this thesis, let $\mathcal{E}(\mathbb{R}^d)$ be the collection of all non-empty compact subsets of d-dimensional Euclidean space $\mathbb{R}^d$ ($d \in \mathbb{Z}^+$). Also let $\mathcal{K}(\mathbb{R}^d)$ be the collection of all non-empty convex compact subsets of $\mathbb{R}^d$. Throughout this thesis, let $(\Omega, \mathcal{F}, P)$ be our probability space.

Definition 1.2.1. *For any $S \in \mathcal{E}(\mathbb{R}^d)$, we say C is the convex hull of S (denoted by $C = \operatorname{Conv}(S)$), if*

$$C = \bigcap_{A \in \mathcal{K}(\mathbb{R}^d), A \supset S} A.$$

Remark 1.2.2. The set C can also be written as

$$C = \bigcup_{n=1}^{\infty} \left\{ \sum_{i=1}^{n} a_i x_i : x_i \in S,\ a_i \geq 0,\ i = 1, 2, \ldots, n,\ a_1 + a_2 + \cdots + a_n = 1 \right\}.$$

Also notice that C is the smallest convex set containing S.

From the definition of convex hulls, it is easy to see the following property.

Proposition 1.2.3. *If S is a finite set, then its convex hull, $\operatorname{Conv}(S)$, is the smallest convex polytope containing S. Obviously, all the vertices of $\operatorname{Conv}(S)$ belong to S.*

Definition 1.2.4. *We say C is the random convex hull of S if S is a random element taking values in $\mathcal{E}(\mathbb{R}^d)$ and for almost all $\omega \in \Omega$, $C(\omega)$ is the convex hull of $S(\omega)$. When S is a finite point set, we call C a random polytope.*

The research on random polytopes originates from Sylvester's Four-Point Problem which studies the probability that the convex hull of four points is quadrilateral if four points are randomly (e.g. uniformly) selected from a planar region K. (See Figure 1.1.)

Sylvester's Four-Point Problem can be easily extended to n points as follows. Let X_i, $1 \leq i \leq n$ be a sequence of random points in $\mathbb{R}^d$ with identical underlying distribution F (i.i.d). Define $K_n = \operatorname{Conv}\{X_1, X_2, \ldots, X_n\}$. What is the probability that the convex hull of n points is exactly a polygon with n edges? Furthermore, we can study some asymptotic properties of K_n, or more specifically, the functionals of

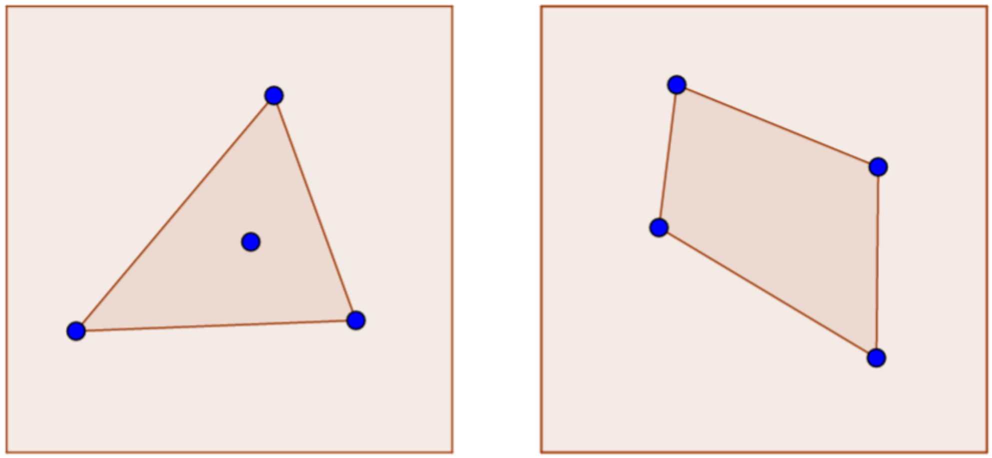

Figure 1.1: Two situations of Four-point problem when K is a square.

K_n. To our knowledge, this has been rarely studied before two classical papers [35] and [34] by Rényi and Sulanke. They assume the underlying distribution F is the uniform distribution on a convex set $K \subset \mathbb{R}^2$ (see Figure 1.2). They derive many interesting results about the expected number of vertices, the area (or volume) and the perimeter (or superficial area) of K_n.

The random polytope, K_n, has been studied for several reasons. For example, MacDonald et al. [29] apply K_n to the estimation for the territorial range of a wildlife species by tagging an individual of this species with a radio transmitter and recording the position as X_i after releasing. Ripley and Rasson [37] apply K_n to the estimation for the support of the homogeneous planar Poisson process. Moreover in order statistics, the method of convex hull peeling constructs a nested sequence of random polytopes based on the sample points. See Tukey [39] and Barnett [5] for an account of this application.

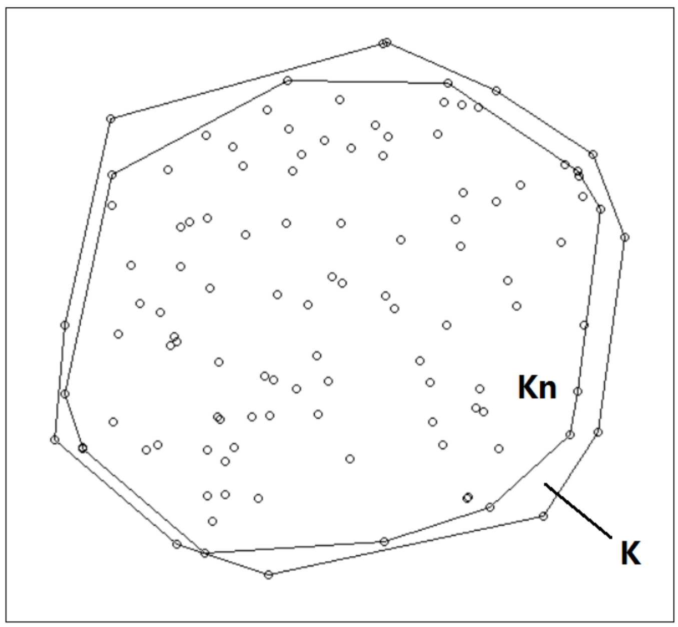

Figure 1.2: K is a convex polygon. F is the uniform distribution on K. The sample size $n = 100$. K_n is the convex hull of the sample points.

1.3 Different aspects of random polytopes

The research on random polytopes is developed in many directions such as the types of polytopes, the functionals and their asymptotic behaviors. Before giving a short list of these results, let us stipulate some notations and introduce some new concepts. Throughout this thesis, we use the following notations. For any set A, its volume is denoted by $|A|$. The sample points, $X_1, X_2, \ldots, X_n$, are random elements taking values in $\mathbb{R}^d$ and drawn from the underlying distribution F. If F is the uniform distribution on some convex set K, we call K the underlying set to draw sample from. For an independent and identically distributed (i.i.d.) sample $X_1, X_2, \ldots, X_n$, we denote the convex hull of the sample (i.e. the random polytope) by K_n, i.e. $K_n = \text{Conv}\{X_1, X_2, \ldots, X_n\}$. For functionals of K_n, we denote the number of vertices by N_n, the volume by $|K_n|$, the perimeter by L_n, the probability content by $F(K_n)$ and the s-dimensional faces by $f_s(K_n)$ ($f_0(K_n)$ is equivalent to N_n). Moreover if K is the underlying set, we denote the set difference $K \setminus K_n$ by D_n and the Hausdorff distance between K and K_n by $H(K, K_n)$.

Definition 1.3.1. *Let d_e be the Euclidean metric on $\mathbb{R}^d$. For any two sets $X,\ Y \subset \mathbb{R}^d$,*

$$H(X, Y) = \inf\{\delta > 0 : X \subset Y_\delta,\ Y \subset X_\delta\} \tag{1.3.1}$$

where

$$X_\delta = \{\boldsymbol{x} : d_e(\boldsymbol{x}, \boldsymbol{y}) < \delta \text{ for some } \boldsymbol{y} \in X\}.$$

The following observation about Hausdorff distance is useful in this thesis. Let $Y \subset \mathbb{R}^2$ be a polygon with vertices $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_r}$. Suppose X is a convex subset of Y. Then

$$H(X, Y) = \sup_{1 \le k \le r} d_e(\boldsymbol{c_k}, X) \tag{1.3.2}$$

where $d_e(\boldsymbol{x}_k, X) = \inf\{d_e(\boldsymbol{c_k}, \boldsymbol{x}) : \boldsymbol{x} \in X\}$.

As for the types of random polytopes, we always name the random polytopes by their underlying distributions. For example, the most common underlying distributions are uniform distribution and Gaussian distribution (with identity covariance matrix). In these cases, we call K_n a uniform polytope and a Gaussian polytope respectively. Besides, there are other types of random polytopes with extensive interest which are called Poisson polytopes and regularly varying polytopes. (See below for the definitions.)

Definition 1.3.2. *Let F be the uniform distribution on $K \subset \mathbb{R}^d$ ($d \geq 2$) and $\{X_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables with the common distribution F. Suppose $X(n) = \sum_{i=1}^{N} \varepsilon_{X_i}$, where N is a Poisson random variable with mean n ($\mathrm{E}\,(N) = n$) and independent from X_i, and $\varepsilon_{\boldsymbol{x}}(\cdot)$ is the point measure at $\boldsymbol{x}$ (see Appendix A.2). We draw a sample, $X_1, X_2, \ldots X_N$, uniformly from K. Then the Poisson polytope Π_n is defined by*

$$\Pi_n = \mathrm{Conv}\,\{X_1, \ldots, X_N\}.$$

Remark 1.3.3. We can prove that $X(n) = \sum_{i=1}^{N} \varepsilon_{X_i}$ is a Poisson point process with intensity $nF(\cdot)$ (see the definitions in Appendix A.3 and A.5). That is why we call Π_n a Poisson polytope. (See Resnick [36].)

Definition 1.3.4. *Suppose $\mathbb{H} = \mathbb{R}^d \setminus \{\mathbf{0}\}$. Let $X_1, X_2 \ldots, X_n$ be a sequence of i.i.d. sample points drawn from the underlying distribution F. We say F is regularly varying (or satisfies the regularly varying condition) if there exist positive constants a_n, such that $\lim_{n\to\infty} a_n = \infty$ and*

$$nF(a_n \cdot) = nP[a_n^{-1} X_1 \in \cdot\ ] \xrightarrow{V} \mu,$$

where "$\xrightarrow{V}$" denotes vague convergence (see Appendix A.6) and μ is a non-degenerate Radon measure (see Appendix A.1) on $\mathbb{H}$.

Example 1.3.5. A bivariate Cauchy distribution satisfies the regularly varying condition. The probability density function for a centered bivariate Cauchy distribution with independent components can be written as follows:

$$f(x, y) = \frac{1}{\pi^2 \gamma_1 \gamma_2 (1 + (x/\gamma_1)^2)(1 + (y/\gamma_2)^2)},$$

for $(x, y) \in \mathbb{R}^2$. Figure 1.3 plots the convex hull of 1000 sample points drawn from the bivariate Cauchy distribution with independent components which $\gamma_1 = \gamma_2 = 1$. The probability density function of a common bivariate Cauchy distribution can be given by:

$$f(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2\pi |\boldsymbol{\Sigma}|^{1/2} [1 + (\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})]^{1.5}},$$

for $\boldsymbol{x} \in \mathbb{R}^2$. Figure 1.4 plots the convex hull of 1000 sample points drawn from the bivariate Cauchy distribution which $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma}$ is the identity matrix.

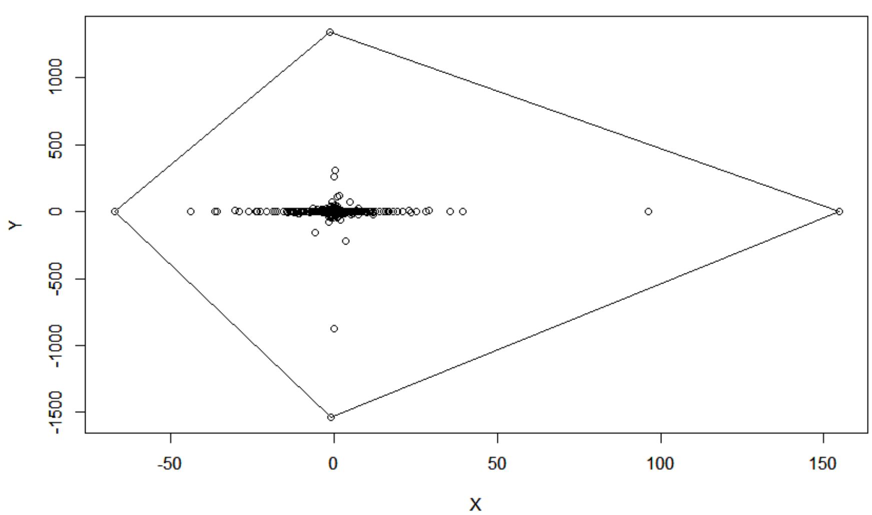

Figure 1.3: The convex hull of 1000 sample points drawn from a bivariate Cauchy distribution with independent components.

As for the asymptotic behaviors of random polytopes, researchers are mainly interest-

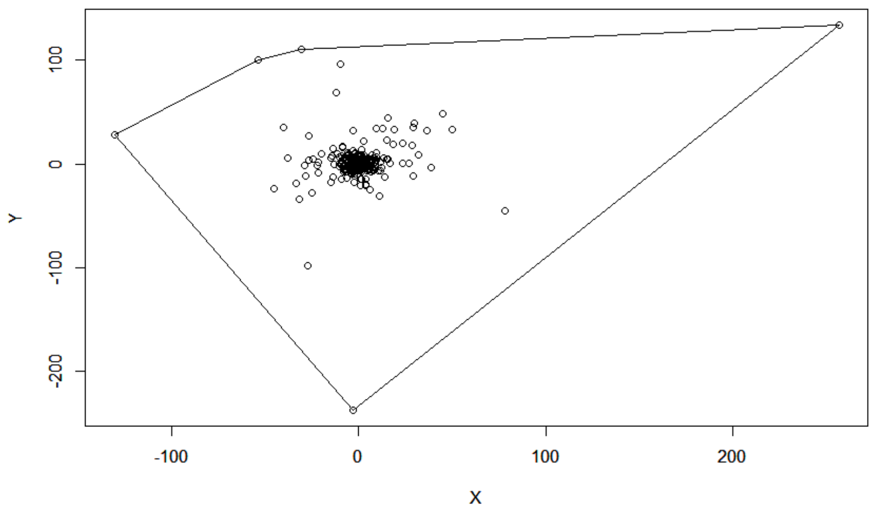

Figure 1.4: The convex hull of 1000 sample points drawn from a bivariate cauchy distribution.

ed in three important characterizations: expected moments, limit theorems and their relationships. For the moments, problems include giving a bound on expectations or variances as a function of the sample size. For the limit theorem, the interest is mainly concentrated on the limiting distributions of the functionals after appropriate normalizations and centering. For the relationships, some authors study the recursion formulas for expected moments or the mathematical relationships among the moments for different functionals.

1.4 A short list of relevant results

Here we give a short list of some relevant results which directly or indirectly serve as the theoretical basis of our main topics. The results directly cited in our derivation are denoted in the form of theorems. As we state in the previous section, we classify these results according to the types of polytope, the functionals and the asymptotic behaviors. Again, here and later, the sample points $X_1, X_2, \ldots, X_n$ are always drawn from the underlying distribution F, and $K_n = \text{Conv}\{X_1, \ldots, X_n\}$.

1.4.1 Limit theorems of the number of vertices on uniform polytopes

Theorem 1.4.1. *(Groeneboom [20]) Let F be the uniform distribution on a convex set $K \subset \mathbb{R}^2$. If K is a convex polygon with r (≥ 3) vertices, and N_n is the number of vertices of K_n, then*

$$\left\{N_n - \frac{2}{3} r \log n\right\} / \sqrt{\frac{10}{27} r \log n} \xrightarrow{D} \mathcal{N}(0,1), \tag{1.4.1}$$

where $\mathcal{N}(0,1)$ denotes the standard normal distribution. If K is a unit disk, then

$$(N_n - 2\pi c_1 n^{1/3}) / \sqrt{2\pi c_2 n^{1/3}} \xrightarrow{D} \mathcal{N}(0,1), \tag{1.4.2}$$

where c_1 and c_2 are two constants.

1.4.2 Limit theorems of the set difference and the perimeter on uniform polytopes

As an extension for the technique in Groeneboom [20], Cabo and Groeneboom [13] prove the central limit theorem for the set difference $D_n := K \setminus K_n$.

Theorem 1.4.2. *(Cabo and Groeneboom [13]) Let F be the uniform distribution on a convex set $K \subset \mathbb{R}^2$. If K is a convex polygon with r (≥ 3) vertices, then as $n \to \infty$,*

$$\frac{|D_n|/|K| - \beta_n}{\alpha_n} \xrightarrow{D} \mathcal{N}(0,1), \tag{1.4.3}$$

where $\alpha_n = \frac{1}{n}\sqrt{\frac{28}{27} r \log n}$ and $\beta_n = \frac{2}{3n} r \log n$.

Here α_n and β_n are the correct version. (The original result in Cabo and Groeneboom [13] is $\alpha = \sqrt{\frac{100}{189} r n^{-1} \log n}$ and $\beta = \frac{2}{3} r n^{-1} \log n$, which is proved to be incorrect by Buchta [12] (see Section 1.4.10)). To derive the correct version, Groeneboom [21] introduces a theorem in Nagaev and Khamdamov [31] (not published):

$$\left(\frac{10}{27} r \log n\right)^{-1/2} \left(N_n - \frac{2}{3} r \log n,\ n|D_n|/|K| - \frac{2}{3} r \log n\right) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), \tag{1.4.4}$$

where the covariance matrix $\mathbf{\Sigma}$ is given by

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 1 \\ 1 & \frac{14}{5} \end{bmatrix}. \tag{1.4.5}$$

It is easy to check (1.4.4) and (1.4.3) give the same marginal for D_n

1.4.3 Expected moments of the set difference on uniform polytopes

Let F be the uniform distribution on a convex set $K \subset \mathbb{R}^d$ with smooth boundary ∂K. The definition of "smooth" here is $\partial K \in \mathcal{C}^k(\mathbb{R}^{d-1})$ (i.e. with k continuous derivatives) for some $k \geq 2$ and the Gaussian curvature κ (see Appendix A.7) are bounded away from 0 and ∞. Also let $D_n = K \setminus K_n$. Bárány [3] proves that for any $d \geq 2$,

$$\mathrm{E}\,|D_n| = c(d) \left(\frac{n}{|K|}\right)^{-2/(d+1)} \int_{\partial K} (\kappa(z))^{1/(d+1)} dz + O\left(n^{-3/(d+1)} \log^2 n\right), \tag{1.4.6}$$

where $c(d)$ is a constant only depending on d. This result is improved by Schütt [38] as follows.

Theorem 1.4.3. *(Schütt [38]) For any convex set $K \subset \mathbb{R}^d$,*

$$\lim_{n\to\infty} \mathrm{E}\,|D_n| = c(d) \left(\frac{n}{|K|}\right)^{-2/(d+1)} \int\limits_{\partial K} (\kappa(z))^{1/(d+1)} dz, \tag{1.4.7}$$

where κ is the generalized Gaussian curvature (see Schütt [38]) and $c(d)$ is a constant only depending on d.

1.4.4 Limit theorems of the volume, the perimeter and the number of s-dimensional faces on uniform polytopes

Let F be the uniform distribution on a unit disk K. Hsing [23] proves the central limit theorem for the $|K_n|$ (the volume of K_n) as follows.

$$n^{5/6}(|K_n| - \mathrm{E}\,|K_n|) \xrightarrow{D} \mathcal{N}(0, \sigma^2), \tag{1.4.8}$$

where σ is an unknown constant. (See the forthcoming sections.)
Bräker and Hsing [9] develop the technique in Hsing [23] for general convex sets with certain kind of smooth boundary and derives the central limit theorem for the joint distribution of the volume $|K_n|$ and the perimeter L_n.

Theorem 1.4.4. *(Bräker and Hsing [9]) Let F be the uniform distribution on a convex set $K \subset \mathbb{R}^2$, where the boundary of K, i.e. ∂K, has a curvature (see Appendix A.7) bounded away from 0 and ∞, and $\partial K \in \mathcal{C}^2(\mathbb{R})$. Let L and L_n be the perimeters of K and K_n respectively. Then as $n \to \infty$, we have*

$$n^{5/6}(|K_n| - \mathrm{E}\,|K_n|, L_n - \mathrm{E}\,L_n) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), \tag{1.4.9}$$

where $\mathbf{\Sigma}$ *is a constant matrix. The convergences of expectations are given by*

$$\lim_{n\to\infty} n^{2/3}\mathrm{E}\left(L - L_n\right) = c_1 \tag{1.4.10}$$

and

$$\lim_{n\to\infty} n^{2/3}\mathrm{E}\left(|D_n|\right)) = c_2, \tag{1.4.11}$$

where c_1 *and* c_2 *are two constants.*

When $K \subset \mathbb{R}^d$ ($d \geq 2$), Reitzner [33] proves the following central limit theorem of $|K_n|$.

Theorem 1.4.5. *(Reitzner [33]) let* d (≥ 2) *and* s $(\leq d-1)$ *be two nonnegative integers. Let* $K \subset \mathbb{R}^d$ *be a convex set with* $\partial K \in \mathcal{C}^2(\mathbb{R}^{d-1})$ *and positive Gaussian curvature (see Appendix A.7) bounded away from 0 and* ∞*. Let* Φ *be the cumulative distribution function of the standard normal distribution. Then there exist constants* C_n *and* $C_n^{'}$*, bounded between two constants, depending only on* K*, and another two constants* c *and* $c^{'}$ *such that*

$$\left| P\left(\frac{|K_n| - \mathrm{E}\,|K_n|}{\sqrt{C_n n^{-1-2/(d+1)}}} \leq t\right) - \Phi(t)\right| \leq c\varepsilon(n) \tag{1.4.12}$$

and

$$\left| P\left(\frac{f_s(K_n) - \mathrm{E}\left(f_s(K_n)\right)}{\sqrt{C_n^{'} n^{1-2/(d+1)}}} \leq t\right) - \Phi(t)\right| \leq c^{'}\varepsilon_s(n), \tag{1.4.13}$$

where

$$\varepsilon(n) = n^{-1/(2(d+1))}(\log n)^{2+2/(d+1)}$$

and

$$\varepsilon_s(n) = n^{-1/(2(d+1))}(\log n)^{2+3s+2/(d+1)}.$$

Let $D_n = K \setminus K_n$. Then the inequality in (1.4.12) can be also written as

$$n^{\frac{1}{2}+\frac{1}{d+1}} \left(|D_n| - \mathrm{E}\,|D_n|\right) \xrightarrow{D} \mathcal{N}(0,\sigma^2), \quad \text{as } n \to \infty,$$

where σ is a constant.

1.4.5 The Hausdorff distance on uniform polytopes

Let F be the uniform distribution on a convex set $K \subset \mathbb{R}^d$ with non-empty interior. Suppose $X_1, X_2, \ldots, X_n$ is a sequence of i.i.d. random points drawn from F. Let $K = \mathrm{Conv}\,\{X_1, \ldots, X_n\}$. The Hausdorff distance between the underlying set K and the random polytopes K_n, i.e. $H(K, K_n)$, is studied by many authors. Dümbgen and Walther [17] initially derive the rates of the almost sure convergence for $H(K, K_n)$ as follows:

$$H(K, K_n) = O((\log n/n)^{1/d}) \text{ almost surely.} \tag{1.4.14}$$

By adding smoothness condition, a stronger version is also derived by Dümbgen and Walther [17]. Let d_e be the Euclidean metric on $\mathbb{R}^d$ which equipped with standard inner product $\langle \cdot, \cdot \rangle$. Let $B(\mathbf{0}, 1) = \{\boldsymbol{y} \in \mathbb{R}^d : d_e(\boldsymbol{x}, \boldsymbol{y}) \le 1\}$. For each $\boldsymbol{x} \in \partial K$, there is a unique $\theta(\boldsymbol{x}) \in \partial B(\mathbf{0}, 1)$ such that $\langle \boldsymbol{y}, \theta(\boldsymbol{x}) \rangle \le \langle \boldsymbol{x}, \theta(\boldsymbol{x}) \rangle$ for all $\boldsymbol{y} \in K$. Let $|\theta(\boldsymbol{x}) - \theta(\boldsymbol{y})| \le l|\boldsymbol{x} - \boldsymbol{y}|$ for some constant $l \in \mathbb{R}$ and any $\boldsymbol{x},\ \boldsymbol{y} \in \partial K$. Then

$$H(K, K_n) = O((\log n/n)^{2/(d+1)}) \text{ almost surely.} \tag{1.4.15}$$

Bräker et al. [10] develop the technique in Hsing [23] and derive the weak convergence of $H(K, K_n)$ if $K \subset \mathbb{R}^2$ is a convex polygon or has certain smooth boundary.

Theorem 1.4.6. *(Bräker et al. [10]) Let F be the uniform distribution on a convex* *set $K \subset \mathbb{R}^2$. Let K be a convex polygon with angles $\theta_1, \theta_2, \ldots, \theta_r$. Then*

$$\lim_{n\to\infty} P\{\sqrt{n}H(K, K_n) \le x\} = G_1(x), \tag{1.4.16}$$

where

$$G_1(x) = \prod_{i=1}^{r}(1 - p_i(x))$$

with

$$p_i(x) = \begin{cases} \displaystyle\int_0^{\theta_i} h_i(x,\theta)d\theta + \exp\left\{-\frac{x^2}{2|K|}\tan\theta_i\right\}, & 0 < \theta_i < \dfrac{\pi}{2} \\ \displaystyle\int_{\theta_i-\pi/2}^{\pi/2} h_i(x,\theta)d\theta, & \dfrac{\pi}{2} \le \theta_i < \pi \end{cases}$$

and

$$h_i(x,\theta) = \exp\left\{-\frac{x^2}{2|K|}\big(\tan\theta_i + \tan(\theta_i - \theta)\big)\right\}\frac{x^2}{2|K|}\tan^2\theta.$$

Theorem 1.4.7. *(Bräker et al. [10]) Let F be the uniform distribution on a bounded convex set $K \subset \mathbb{R}^2$. Suppose the boundary of K has length L and we parameterize it (positively oriented) as $t \to c(t)$, where $t =$ the arc length between $c(0)$ and $c(t)$. Suppose the curvature $\mathcal{K}(t) = |\ddot{c}(t)|$ is well-defined and bounded away from 0 and ∞ and has a bounded derivative. Define the function*

$$\lambda(t) = |K|\sqrt{\mathcal{K}(t)}, \quad 0 \le t < L,$$

and let $\lambda_0 := \max_{t\in[0,L)} \lambda(t)$. Suppose that there exit some bounded sequence of non-negative constants ν_n and positive constant μ, such that

$$\frac{(\log n)^{\nu_n}}{|K|}\int_0^L \exp\left\{-\gamma_n\left(\frac{\lambda_0}{\lambda(t)} - 1\right)\log n\right\}dt \longrightarrow \mu \in (0,\infty), \quad \text{as } n\to\infty, \tag{1.4.17}$$

where

$$\gamma_n = \frac{1}{3} + \left(\frac{2}{3} - \nu_n\right)\frac{\log\log n}{\log n}.$$

Denote

$$c_n = \frac{3\sqrt{2}}{8}\lambda_0\gamma_n$$

and

$$a_n = n^{-2/3}(\log n)^{-1/3}, \quad b_n = n^{-2/3}(c_n \log n)^{2/3}.$$

Then

$$\lim_{n\to\infty} P\left(\frac{H(K, K_n) - b_n}{a_n} \le x\right) = G_2(x), \quad x \in \mathbb{R}, \tag{1.4.18}$$

where

$$G_2(x) = \exp\left\{-d_1 e^{-d_2 x}\right\}$$

with

$$d_1 = c_0^{2/3}\mu, \quad d_2 = (2c_0)^{-2/3}.$$

1.4.6 Limit theorems of the diameter on uniform polytopes

Let F be the uniform distribution on a d-dimensional unit ball ($d \ge 2$). Suppose $X_1, X_2, \dots, X_n$ is a sequence of i.i.d. random points drawn from F. Also Let $K_n = \mathrm{Conv}\{X_1, \dots, X_n\}$. Denote the diameter of K_n, the longest distance between two points in the interior of K_n, by $\mathrm{diam}(K_n)$. Mayer and Molchanov [30] prove for $t > 0$,

$$\lim_{n\to\infty} P\left(n^{\frac{4}{d+3}}(2 - \mathrm{diam}(K_n)) \le t\right) = G(t; d), \tag{1.4.19}$$

where

$$G(t; d) = 1 - \exp\left\{-\frac{2^d d\Gamma(\frac{d}{2}+1)}{\sqrt{\pi}(d+1)(d+3)\Gamma(\frac{d+1}{2})} t^{\frac{d+3}{2}}\right\}$$

with

$$\Gamma(x) = \int_0^\infty s^{x-1}e^{-s}ds.$$

1.4.7 Expected moments of the number of s-dimensional faces on Gaussian polytopes

Let F be the Gaussian distribution on $\mathbb{R}^d$ $(d \geq 1)$ with identity covariance matrix. Suppose $X_1, X_2, \ldots, X_n$ is a sequence of i.i.d. random points drawn from F. Let $K_n = \operatorname{Conv}\{X_1, \ldots, X_n\}$. Recall that $f_s(K_n)$ means the number of s-dimensional faces of random polytopes. Raynaud [32] computes $\mathrm{E}(f_s(K_n))$ for Gaussian polytopes.

Theorem 1.4.8. *(Raynaud [32]) Let F be the Gaussian distribution on $\mathbb{R}^d$ $(d \geq 1)$ with identity covariance matrix. For $s = 0, 1, \ldots, d-1$, we have*

$$\mathrm{E}(f_s(K_n)) = \frac{2^d}{\sqrt{d}} \binom{d}{s+1} \beta_{s,d-1} (\pi \log n)^{(d-1)/2} (1 + o(1)), \tag{1.4.20}$$

where $\beta_{s,d-1}$ is the internal angle of the regular $(d-1)$-simplex at one of its s-dimensional faces. When $s = 0$, (1.4.20) is equivalent to

$$\mathrm{E}\, N_n = \Theta\left((\log n)^{(d-1)/2}\right). \tag{1.4.21}$$

Hug and Reitzner [25] give an upper bound of variance of f_s for d-dimensional Gaussian polytopes.

Theorem 1.4.9. *Let d (≥ 2) and s $(\leq d-1)$ be positive integers. Let F be the Gaussian distribution on $\mathbb{R}^d$ with identity covariance matrix. Then there exists a positive constant c_d, only depending on d, such that*

$$\mathrm{Var}(f_s(K_n)) \leq c_d (\log n)^{(d-1)/2}. \tag{1.4.22}$$

When $s = 0$, (1.4.22) is equivalent to

$$\mathrm{Var}(N_n) = O((\log n)^{(d-1)/2}). \tag{1.4.23}$$

1.4.8 Limit theorems of the number of s-dimensional faces, the volume and the perimeter on Gaussian polytopes

Let F be the Gaussian distribution on $\mathbb{R}^2$ with identity covariance matrix. For the Gaussian polytope K_n, Hueter [24] proves the central limit theorems for the number of vertices N_n, the perimeter L_n and the volume $|K_n|$ as follow:

$$\frac{N_n - 2\sqrt{2\pi \log n}}{(2\sqrt{2\pi \log n}(1+\pi c_0))^{1/2}} \xrightarrow{D} \mathcal{N}(0,1) \tag{1.4.24}$$

where c_0 is a constant,

$$(L_n - 2\pi\sqrt{2\log n})/\sqrt{4\pi^{3/2}\log n} \xrightarrow{D} \mathcal{N}(0,1) \tag{1.4.25}$$

and

$$(|K_n| - 2\pi \log n)/\sqrt{2\pi^{3/2}(\log n)^2} \xrightarrow{D} \mathcal{N}(0,1). \tag{1.4.26}$$

For the Gaussian polytope K_n on $\mathbb{R}^d$ ($d \geq 2$), Bárány and Vu [4] derive the central limit theorem for the volume $|K_n|$ and the number of s-dimensional faces $f_s(K_n)$ as follow. Denote the cumulative distribution function of the standard normal distribution by Φ. Then

$$\left| P\left(\frac{|K_n| - E(|K_n|)}{\sqrt{\mathrm{Var}\,(|K_n|)}} \leq t \right) - \Phi(t) \right| \leq \varepsilon(n) \tag{1.4.27}$$

and

$$\left| P\left(\frac{f_s(K_n) - \mathrm{E}\,(f_s(K_n))}{\sqrt{\mathrm{Var}\,(K_n)}} \leq t \right) - \Phi(t) \right| \leq \varepsilon(n) \tag{1.4.28}$$

for any value of t, where $\varepsilon(n)$ can be taken as $(\log n)^{-(d-1)/4+o(1)}$.

1.4.9 Limit theorems of the volume and the number of s-dimensional faces on Poisson polytopes

Let F be the uniform distribution on a non-empty convex set $K \subset \mathbb{R}^d$. Suppose the sample points $X_1, X_2, \ldots, X_n$ is a sequence of i.i.d. random points drawn from F and $K_n = \mathrm{Conv}\{X_1, \ldots, X_n\}$. Reitzner [33] proves the central limit theorem for Poisson polytopes (see Definition 1.3.2), i.e. $\Pi_n = \mathrm{Conv}\{X_1, \ldots, X_N\}$ where N is a Poisson random variable with mean n. Denote the boundary of K by ∂K. Suppose $\partial K \in \mathcal{C}^2(\mathbb{R}^{d-1})$ and has a positive Gaussian curvature (see Appendix A.7) bounded away from 0 and ∞. Let d (≥ 2) and s $(\leq d-1)$ be positive integers. Then there exist constants c_1 and c_2 such that

$$\left| P\left(\frac{|\Pi_n| - E(|\Pi_n|)}{\sqrt{\mathrm{Var}\,(|\Pi_n|)}} \leq t \right) - \Phi(t) \right| \leq c_1 \varepsilon_1(n) \tag{1.4.29}$$

and

$$\left| P\left(\frac{f_s(\Pi_n) - \mathrm{E}\,(f_s(\Pi_n))}{\sqrt{\mathrm{Var}\,(\Pi_n)}} \leq t \right) - \Phi(t) \right| \leq c_2 \varepsilon_2(n) \tag{1.4.30}$$

for any value of t, where

$$\varepsilon_1(n) = n^{-1/2+1/(d+1)} (\log n)^{2+2/(d+1)}$$

and

$$\varepsilon_2(n) = n^{-1/2+1/(d+1)} (\log n)^{2+3s+2/(d+1)}.$$

With the same assumption as above, Reitzner also builds the relationships between Π_n and K_n as follow. There are constants c_3 and c_4 depending on K such that for any t,

$$|P(|K_n| \leq t) - P(|\Pi_n| \leq t)| \leq c_3 (\log n/n)^{2/(d+1)}$$

and

$$|P(f_s(K_n) \le t) - P(f_s(\Pi_n) \le t)| \le c_4(\log n/n)^{2/(d+1)}.$$

1.4.10 Weak convergence of regularly varying polytopes

Suppose $\mathcal{E}(\mathbb{R}^2)$, the collection of all non-empty compact subsets of $\mathbb{R}^2$, is metrized by the Hausdorff distance. Suppose F is a distribution on $\mathbb{R}^2$ which satisfies the regularly varying condition $nF(a_n\cdot) \xrightarrow{V} \mu$ (see Definition 1.3.4). Let $X_1, X_2, \dots, X_n$ be a sequence of i.i.d. random points drawn from F. Let $\eta := \sum_{i=1}^{\infty} \varepsilon_{\boldsymbol{j_k}}$ be a Poisson point process (see Appendix A.5) with intensity μ. Davis et al. [15] prove

$$\sum_{i=1}^{n} \varepsilon_{a_n^{-1}X_i} \xrightarrow{D} \eta$$

and

$$\text{Conv}\,\{a_n^{-1}X_1, \dots, a_n^{-1}X_n\} \xrightarrow{D} \text{Conv}\,\{\boldsymbol{j_k} : k \ge 1\}$$

in $\mathcal{E}(\mathbb{R}^2)$.

1.4.11 Almost sure limit sets of random sample

Davis et al. [16] study the almost sure convergence of random sample in the first quadrant. For all $\boldsymbol{x},\ \boldsymbol{y} \in \mathbb{R}^d$, $\boldsymbol{x} = (x_i, \dots, x_d)$ and $\boldsymbol{y} = (y_1, \dots, y_d)$, denote $\boldsymbol{x} \le \boldsymbol{y}$ if $x_i \le y_i$ for every i. Let $X_1, X_2, \dots, X_n$ be a sequence of i.i.d. random points in $[0,\infty)^d$. Suppose $h : [0,\infty)^d \to [0,\infty)$ such that $h(\boldsymbol{x}) = -\log P(X_i \ge \boldsymbol{x})$, and for all $\boldsymbol{x} \in [0,\infty)^d \setminus \{\boldsymbol{0}\}$,

$$\lim_{t\to\infty} \frac{h(t\boldsymbol{x})}{h(t\boldsymbol{1})} = v(\boldsymbol{x}) \tag{1.4.31}$$

where v satisfies that for some $\alpha > 0$ and any $\boldsymbol{x}, \boldsymbol{y} \in [0,\infty)^d \setminus \{\boldsymbol{0}\}$,

$$v(t\boldsymbol{x}) = t^{\alpha} v(\boldsymbol{x}) \text{ for any } t > 0, \tag{1.4.32}$$

and

$$v(\boldsymbol{x}) < v(\boldsymbol{y}),\ \text{if } \boldsymbol{x} \leq \boldsymbol{y} \text{ and } \boldsymbol{x} \neq \boldsymbol{y}. \tag{1.4.33}$$

Let b_n satisfy $h(b_n\mathbf{1}) \sim \log n$ and define

$$S_n = \{X_i/b_n,\ 1 \leq i \leq n\},$$
$$S_v = \{\boldsymbol{x} \geq \mathbf{0} :\ v(\boldsymbol{x}) \leq 1\}.$$

Then Davis et al. [16] prove that as $n \to \infty$,

$$S_n \xrightarrow{a.s.} S_v, \tag{1.4.34}$$

and

$$\text{Conv}\,(S_n) \xrightarrow{a.s.} \text{Conv}\,(S_v), \tag{1.4.35}$$

in $\mathcal{E}(\mathbb{R}^d)$.

When $d = 2$, suppose v is continuous and satisfies (1.4.31), (1.4.32) and that for all $\boldsymbol{x} \in [0,\infty)^2 \setminus \{\mathbf{0}\}$ there exists $\beta > 0$ such that

$$\liminf_{t\to\infty}(v(x_1, tx_2) \wedge v(tx_1, x_2))/t^\beta = \infty. \tag{1.4.36}$$

Let h satisfy

$$\lim_{t\to\infty} \sup_{\|\boldsymbol{x}\|=1} \left| \frac{h(t\boldsymbol{x})}{h(t\mathbf{1})} - v(\boldsymbol{x}) \right| = 0, \tag{1.4.37}$$

where $\|\cdot\|$ is the Euclidean norm. Then for all $\boldsymbol{x} \in [0,\infty)^2 \setminus \{\mathbf{0}\}$,

$$\lim_{t\to\infty} \frac{-\log P(X \geq t\boldsymbol{x})}{h(t\mathbf{1})} = \inf_{s>1}(v(x_1, sx_2) \wedge v(sx_1, x_2)). \tag{1.4.38}$$

Both results above can also be extended to the whole plane. (See the extension technique in Davis et al. [16].) Using (1.4.38) and the extension technique, Davis et al. [16] derive that if $X_1, X_2, \ldots, X_n$ is a sequence of i.i.d. bivariate normal random points with the density

$$f(x_1, x_2) = c \exp\{-(x_1^2 - 2\rho x_1 x_2 + x_2^2)/(2(1-\rho^2))\}, \quad (x_1, x_2) \in \mathbb{R}^2 \tag{1.4.39}$$

and $|\rho| < 1$, then

$$\{a_n^{-1} X_i\}_{i=1}^n \xrightarrow{a.s} \{(x_1, x_2) : x_1^2 - 2\rho x_1 x_2 + x_2^2 \leq 2(1-\rho^2)\}, \tag{1.4.40}$$

where $a_n = \sqrt{\log n}$.

1.4.12 Relationships among expected moments of different functionals

The relationships among the moments of different functionals play an important role in the limit theory of random polytopes. For instance, Efron [18] relates probability content $F(K_n)$ with the number of vertices N_n and derives:

$$\mathrm{E}\,(F(K_n)) = 1 - \frac{\mathrm{E}\,N_{n+1}}{n+1} \tag{1.4.41}$$

where F can be any distribution on $\mathbb{R}^2$ or $\mathbb{R}^3$. This identity appears in many articles, like Reitzner [33] and Bárány [3], and bridges the first moment problems between the number of vertices and the area (or the volume).

Based on purely geometric methods, Buchta [12] extends the identity (1.4.41) to higher dimensions and higher moments as follows.

Theorem 1.4.10. *(Buchta [12]) For any probability distribution F on $\mathbb{R}^d$ such that*

$\forall x \in \mathbb{R}^d,\ F(x) = 0$, *we have*

$$\mathrm{E}\left(F^k(K_n)\right) = \mathrm{E}\left[\prod_{i=1}^{k}\left(1 - \frac{N_{n+k}}{n+i}\right)\right]. \tag{1.4.42}$$

If F is the uniform distribution on a convex set $K \subset \mathbb{R}^d$, Theorem 1.4.10 makes it possible to get the variance of the probability content from the variance of the number of vertices. With the help of Theorem 1.4.1 and 1.4.10, Buchta not only points out the mistake in Cabo and Groeneboom [13] but also provides the correct result (see Theorem 1.4.2). Buchta also improves Theorem 1.4.8 as follows:

$$\frac{D_n/|K| - 2\pi c_1 n^{-2/3}}{\sqrt{2\pi\left(\frac{1}{3}(c_1+c_2)\right)n^{-5/3}}} \xrightarrow{D} \mathcal{N}(0,1), \tag{1.4.43}$$

where c_1 and c_2 are the same constants as in (1.4.2).

Let F be the uniform distribution on a convex polygon $K \subset \mathbb{R}^2$. Buchta [11] introduces recursion formulas for the m-th moment of N_n (the number of vertices of K_n). For any $m,\ n \in \mathbb{N}$, the m-th moment of N_n is the unique solution of the second order difference equation

$$\begin{aligned}&\frac{n(n+2)}{2}\mathrm{E}\left(N_n^m\right) - (n^2-n+1)\mathrm{E}\left(N_{n-1}^m\right) + \frac{(n-2)(n-1)}{2}\mathrm{E}\left(N_{n-2}^m\right)\\ &= \mathrm{E}\left((N_{n-1}+1)^m\right) - \mathrm{E}\left(N_{n-1}^m\right)\end{aligned} \tag{1.4.44}$$

with initial values $\mathrm{E}\left(N_1^m\right) = 1$ and $\mathrm{E}\left(N_2^m\right) = (2^m+2)/3$. And

$$\mathrm{E}\left(N_n^m\right) = 1 + \sum_{k=1}^{n-1}\omega_{k,n}(\mathrm{E}\left((N_k+1)^m\right) - \mathrm{E}\left(N_k^m\right)) \tag{1.4.45}$$

with weights

$$\omega_{k,n} = 2k(k+1)\sum_{i=k+1}^{n}\frac{1}{(i^2-1)i^2}.$$

Furthermore Buchta proves:

$$\mathrm{E}\left(N_n^m\right) = \left(\frac{2}{3}\log n\right)^m + O\left((\log n)^{m-1}\right), \quad \text{as } n \to \infty. \tag{1.4.46}$$

1.5 Organization of this thesis

This thesis is organized as follows. In Chapter 2, we first review some basic background on different types of convergence in probability theory. The properties introduced in this chapter will be very useful throughout this thesis. The proofs of some lemmas are provided to have a more concise proofs for our future results.

In Chapter 3, our first main result is summarized in Theorem 3.2.1. This result is a simple application of the regular bootstrap method on uniform convex polytopes. Since Theorem 3.2.1 is mainly based on the works of Bräker et al. [10] and Zarepour [40], several parts of proof which are only the mimics of the proofs in this two papers will be omitted.

In Chapter 4, we discuss our new bootstrapping scheme starting with defining the concept of the scope of the sample. Two more examples are also given to fully explain our bootstrapping scheme.

Chapter 5 contains our main results in this thesis - the asymptotic consistency of our new bootstrapping scheme applied on uniform polytopes with respect for the Hausdorff distance. We mainly study two types of convergence. For convex polygon cases, the proof of Theorem 5.1.1 and Theorem 5.1.2 fully discribes our technique. For smooth boundary cases, we prove Theorem 5.2.1 and Theorem 5.2.2 following a similar technique. Moreover, we give two examples to demonstrate that our bootstrapping scheme can also work on other functionals.

In Chapter 6, we present a conjecture about the generalization of our theory to Gaussian polytopes. In Chapter 7, we include simulation studies for our main results in Chapter 5 to verify our proofs using descriptive tools. In Chapter 8, we briefly discuss some future works to be considered later.

1.6 Notations by default

The following notations will be used by default.

$f(n) = O(g(n))$: there exists a constant C such that $f(n) \leq Cg(n)$;

$f(n) = \Omega(g(n))$: there exists a constant C such that $f(n) \geq Cg(n)$;

$f(n) = \Theta(g(n))$: $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$;

$m = o(n)$: $m/n \to 0$, as $n, m \to \infty$;

$m \sim n$: $m/n \to 1$, as $n, m \to \infty$;

$\mathcal{C}^k(S)$: the set of all functions with k continuous derivatives with domain S;

∂A: the boundary of the set A;

$x \wedge y := \min(x, y)$;

$A_n \uparrow A$: $A = \bigcup_{n=1}^{\infty} A_n$, where $A_n \subset A_m$ if $n < m$;

$A_n \downarrow A$: $A = \bigcap_{n=1}^{\infty} A_n$, where $A_n \supset A_m$ if $n < m$.

Chapter 2

Preliminary

2.1 Basic convergence of random variables

2.1.1 Almost sure convergence

Let $\{X_n\}_{n=0}^{\infty}$ be a sequence of real-valued random variables. We say $\{X_n\}_{n=1}^{\infty}$ converges almost surely (a.s.) to X_0 if

$$P(\lim_{n\to\infty} X_n = X_0) = 1,$$

i.e. for almost all $\omega \in \Omega$,

$$X_n(\omega) \to X_0(\omega), \quad \text{as } n \to \infty.$$

2.1.2 Convergence in probability

Let $\{X_n\}_{n=0}^{\infty}$ be a sequence of real-valued random variables. We say $\{X_n\}_{n=1}^{\infty}$ converges to X_0 in probability (written $X_n \xrightarrow{P} X_0$) if for any given $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X_0| > \varepsilon) = 0.$$

Almost sure convergence implies convergence in probability but the converse is not true.

2.1.3 Convergence in distribution

Suppose S is a complete, separable metric space and $C(S)$ is the collection of all bounded, continuous real-valued functions on S. Let $\{X_n\}_{n=0}^{\infty}$ be a sequence of random elements taking values in S. Denote $P_n = P \circ X_n^{-1}$ for all integer $n \geq 0$. We say $\{X_n\}_{n=1}^{\infty}$ converges to X_0 in distribution, or weakly (written $X_n \xrightarrow{D} X_0$ or $P_n \xrightarrow{D} P_0$) if for any $f \in C(S)$,

$$\lim_{n\to\infty} \mathrm{E}\,(f(X_n)) = \mathrm{E}\,(f(X_0)).$$

Convergence in probability implies convergence in distribution and also the converse is not true. But if X_0 is a constant $c \in S$, i.e. $X_0 \equiv c$ almost surely, then

$$X_n \xrightarrow{P} c \text{ iff } X_n \xrightarrow{D} c,$$

as $n \to \infty$.

2.2 Continuous mapping theorem

(Billingsley [7]) Let $(S_1, \mathfrak{S}_1, d_1)$ and $(S_2, \mathfrak{S}_2, d_2)$ be two metric spaces. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random elements of $(S_1, \mathfrak{S}_1, d_1)$ such that

$$X_n \xrightarrow{D} X_0, \quad \text{as } n \to \infty.$$

For any $f : S_1 \to S_2$, denote by D_h the collection of all points $x \in S_1$ such that f is not continuous at x. Then if

$$P(X_0 \in D_h) = 0,$$

we have

$$f(X_n) \xrightarrow{D} f(X_0), \quad \text{as } n \to \infty,$$

in S_2.

2.3 Skorokhod's representation theorem

Let P_n, $n \geq 0$ be a sequence of distributions on a metric space S. Suppose $P_n \xrightarrow{D} P_0$ where P_0 has the separable support. Then there exist random elements $\{X_n\}_{n=0}^{\infty}$ defined on a common probability space $(\Omega, \mathcal{F}, P)$ such that the law of X_n is P_n for all $n \geq 0$ and $X_n \xrightarrow{a.s.} X_0$.

2.4 Weak convergence for the random probability measure

In this section, we refer to the concept of weak convergence in probability. This convergence plays a very important role in our main results.

Let $\{X_n\}_{n=1}^{\infty}$ and $\{Y_n\}_{n=1}^{\infty}$ be two sequences of real-valued random variables. It is easy to see that

$$g(Y_n) = P(X_n \leq x | \ Y_n) \text{ is a random variable.}$$

Let G be a distribution function. Denote the law of $X_n | \ Y_n$ by $\mathcal{L}(X_n | \ Y_n)$. If

$$P(X_n \leq x | \ Y_n) \xrightarrow{P} G(x), \quad \text{as } n \to \infty,$$

then

$$\mathcal{L}(X_n | \ Y_n) \to G \ \text{ in probability.}$$

2.5 Some lemmas

Lemma 2.5.1. *For any events A, B and C, we have*

$$P(A|\ B\cap C)P(C|\ B) = P(A\cap C|\ B).$$

Proof:

$$P(A|\ B\cap C) = \frac{P(A\cap B\cap C)}{P(B\cap C)} = \frac{P(A\cap C|\ B)P(B)}{P(C|\ B)P(B)} = \frac{P(A\cap C|\ B)}{P(C|\ B)}.$$

∎

Lemma 2.5.2. *If $P \subset K \subset \mathbb{R}^2$ such that P is a convex polygon and K is a convex set, we have*

$$L(P) \leq L(K)$$

where $L(P)$ and $L(K)$ are the perimeters of P and K respectively.

Proof: Suppose P has r vertices. let $e_1, e_2, \dots, e_r$ be the edges of P. It is enough to show that there exist disjoint segments $s_1, s_2, \dots, s_r$ on the boundary of K such that

$$e_i \leq s_i, \quad i = 1, 2, \dots, r.$$

Consider the segments, $s_1, s_2, \dots, s_r$, in the Figure 2.1. Since the length of e_i, for any $1 \leq i \leq r$, is the distance between the two parallel lines, the length of s_i is not shorter than the length of e_i. Hence

$$L(P) = \sum_{i=1}^{r} e_i \leq \sum_{i=1}^{r} s_i \leq L(K).$$

∎

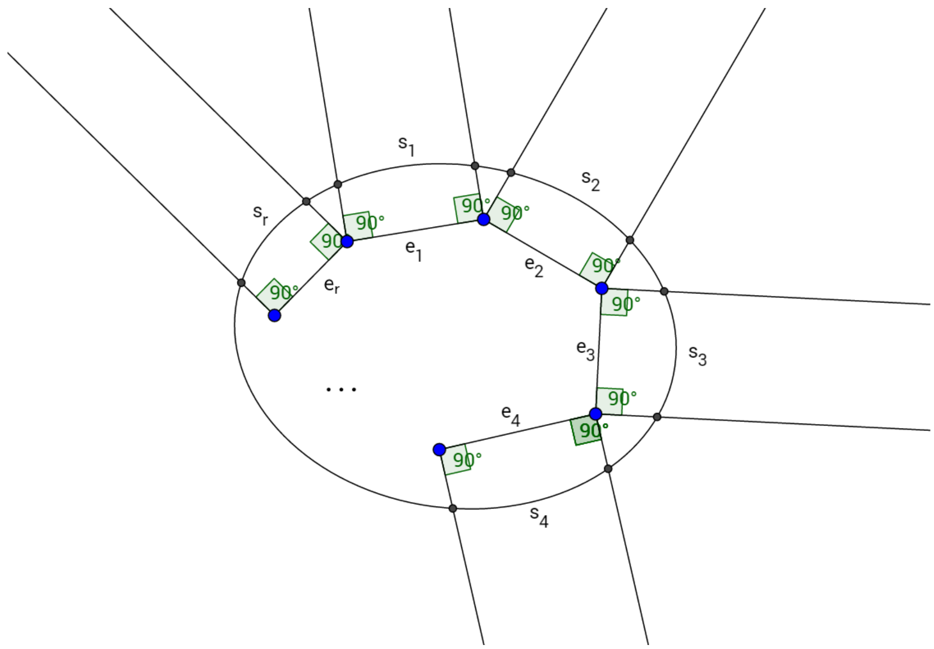

Figure 2.1: The edges $e_1, e_2, \ldots, e_r$ and the segments $s_1, s_2, \ldots, s_r$.

Lemma 2.5.3. *(Billingsley [8]) Suppose d is the metric on S. If (X_n, Y_n) are random elements of $S \times S$ such that $X_n \xrightarrow{D} X_0$ and $d(X_n, Y_n) \xrightarrow{D} 0$, then $Y_n \xrightarrow{D} X_0$.*

Proof: Define $A_\delta := \{x : d(x, A) \leq \delta\}$ where $d(x, A) := \inf\{d(x, y) : y \in A\}$, then

$$P(Y_n \in A) \leq P(X_n \in A_\delta) + P(d(X_n, Y_n) \geq \delta) \tag{2.5.1}$$

and

$$P(X_n \in A) \leq P(Y_n \in A_\delta) + P(d(X_n, Y_n) \geq \delta) \tag{2.5.2}$$

Using our assumptions, we have

$$\limsup_n P(Y_n \in A) \leq \limsup_n P(X_n \in A_\delta) \leq P(X_0 \in A_\delta) \tag{2.5.3}$$

and

$$P(X_0 \in A) \leq \liminf_n P(X_n \in A) \leq \liminf_n P(Y_n \in A_\delta) \tag{2.5.4}$$

If A is closed, then $A_\delta \downarrow A$ as $\delta \downarrow 0$. ∎

Lemma 2.5.4. *If $|A_n - B_n| \xrightarrow{P} 0$, $A_n \geq C_n$, $C_n \xrightarrow{P} Z$, $B_n \leq D_n$ and $D_n \xrightarrow{P} Z$. Then we have $A_n \xrightarrow{P} Z$.*

Proof: For any given $\varepsilon > 0$, as $n \to \infty$,

$$\begin{aligned} P(A_n - Z > \varepsilon) &= P(A_n - B_n + B_n - D_n + D_n - Z > \varepsilon) \\ &\leq P(A_n - B_n > \varepsilon/3) + P(B_n - D_n > \varepsilon/3) + P(D_n - Z > \varepsilon/3) \to 0. \\ P(Z - A_n > \varepsilon) &= P(Z - C_n + C_n - A_n) \\ &\leq P(Z - C_n > \varepsilon/2) + P(C_n - A_n > \varepsilon/2) \to 0. \end{aligned}$$

∎

Lemma 2.5.5. *Let $\boldsymbol{x_i}$, $1 \leq i \leq n$, be a sequence of points in $\mathbb{R}^d$. Suppose d_e is the Euclidean metric. Define the distance between a point $\boldsymbol{x} \in \mathbb{R}^d$ and a set $A \subset \mathbb{R}^d$ by $d_e(\boldsymbol{x}, A) := \inf\{d_e(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{y} \in A\}$. Then the function $d_e(\boldsymbol{x}, \mathrm{Conv}\,\{\boldsymbol{x_1}, \ldots, \boldsymbol{x_n}\})$ is continuous with respect to the topology induced by Euclidean metric on $\mathbb{R}^{dn}$.*

Proof: We only give the proof when $d = 2$, which can be easily generalized to higher dimensions. Suppose $\widetilde{x} = (\boldsymbol{x_1}, \ldots, \boldsymbol{x_n}) \in \mathbb{R}^{2n}$, where $\boldsymbol{x_i} \in \mathbb{R}^2$, $1 \leq i \leq n$. Each $\widetilde{\boldsymbol{x}}$ gives a value of K_n, i.e. $K_n(\widetilde{\boldsymbol{x}}) = \mathrm{Conv}\,\{\boldsymbol{x_1}, \ldots, \boldsymbol{x_n}\}$. We take $\widetilde{\boldsymbol{y}} = (\boldsymbol{y_1}, \ldots, \boldsymbol{y_n}) \in \mathbb{R}^{2n}$, where $\boldsymbol{y_i} \in \mathbb{R}^2$, $1 \leq i \leq n$, such that $d_e(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{y}}) \leq \varepsilon$, where $\varepsilon > 0$ is small enough. Therefore for any $1 \leq i \leq n$, we have $d_{e,2}(\boldsymbol{x_i}, \boldsymbol{y_i}) \leq \varepsilon$, where $d_{e,2}$ is the Euclidean metric on $\mathbb{R}^2$. Then as in Figure 2.2, we make a circle which centers at each vertex

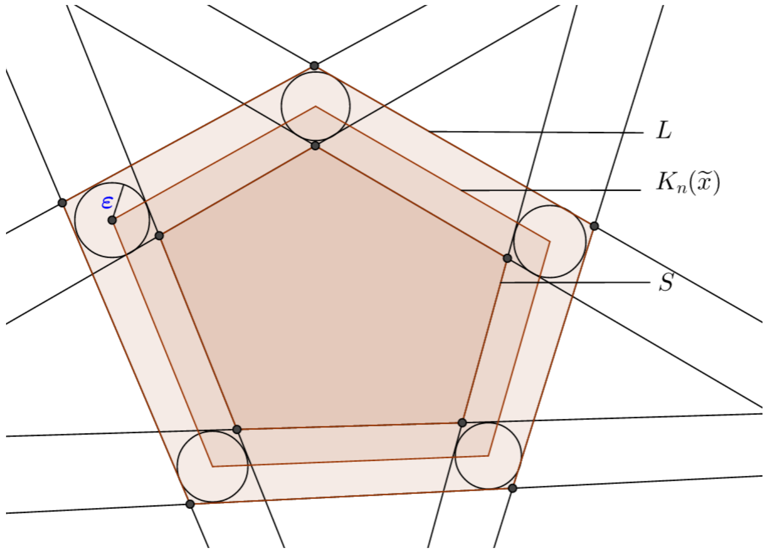

Figure 2.2: The convex polygons L and S.

of $K_n(\widetilde{\boldsymbol{x}})$. By drawing the tangent lines, we get two other convex polygons L and S. It is easy to show that there exist a constant c, which is determined by $\widetilde{\boldsymbol{x}}$ (by all the interior angles of $K_n(\widetilde{\boldsymbol{x}})$), such that $H(L, S) \leq c\varepsilon$. ∎

Now we are ready to start our proof that $d_e(\boldsymbol{x}, \mathrm{Conv}\{\boldsymbol{x_1}, \ldots, \boldsymbol{x_n}\})$ is continuous. When $\boldsymbol{x} \in A$, the proof is trivial. When $\boldsymbol{x} \notin A$, we can take ε small enough such that $\boldsymbol{x} \notin L$. It is enough to show

$$|d_e(\boldsymbol{x}, K_n(\widetilde{\boldsymbol{y}})) - d_e(\boldsymbol{x}, K_n(\widetilde{\boldsymbol{x}}))| \leq c\varepsilon.$$

Since $S \subset K_n(\widetilde{\boldsymbol{y}}), K_n(\widetilde{\boldsymbol{x}}) \subset L$, we have

$$d_e(\boldsymbol{x}, L) \leq d_e(\boldsymbol{x}, K_n(\widetilde{\boldsymbol{x}})), d_e(\boldsymbol{x}, K_n(\widetilde{\boldsymbol{y}})) \leq d_e(\boldsymbol{x}, S),$$

which is followed by

$$|d_e(\boldsymbol{x}, K_n(\widetilde{\boldsymbol{y}})) - d_e(\boldsymbol{x}, K_n(\widetilde{\boldsymbol{x}}))| \leq d_e(\boldsymbol{x}, S) - d_e(\boldsymbol{x}, L).$$

Therefore it is enough to show

$$d_e(\boldsymbol{x}, S) - d_e(\boldsymbol{x}, L) \leq c\varepsilon,$$

which is given by the facts that

$$d_e(\boldsymbol{x}, L) + H(S, L) \geq d_e(\boldsymbol{x}, S)$$

and

$$H(S, L) \leq c\varepsilon.$$

Chapter 3

Main result; the bootstrap

3.1 Introduction

Let $\{X_i\}_{i=1}^n$ be a sequence of i.i.d. random points drawn from the distribution F. Let $\phi_n(X_1, \ldots, X_n; F)$ be the functional of interest. To draw a precise inference for ϕ_n, we desire to know the exact form of the underlying distribution F. Yet, in practice, this is usually impossible. To address this problem, we may use other known approximate distribution to replace F. Efron [19] recommends to replace F with the empirical distribution F_n, where

$$F_n(-\infty, x] = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x) \tag{3.1.1}$$

and I is the indicator function. This method is called the plug-in principle. The plug-in principle applied on any statistics can be used to estimate its distribution. This method is called the bootstrap by Efron [19]. The consistency of this method is coming from the Glivenko-Cantelli theorem:

$$\sup_{x \in \mathbb{R}} |F_n(-\infty, x] - F(-\infty, x]| \to 0, \text{ almost surely,} \tag{3.1.2}$$

which tells us F_n approximates F almost surely. However, the replacement still needs to be inspected for consistency for the target functionals. Specifically, suppose

$$P\left(\frac{\phi_n(X_1,\ldots,X_n;F)-b_n}{a_n}\leq x\right)\to G(x),\ \text{as}\ n\to\infty, \tag{3.1.3}$$

for any continuity point $x\in\mathbb{R}$. Conditionally on $X_1, X_2, \ldots, X_n$, draw points $X^*_{n,1}$, $X^*_{n,2}$, $\ldots$, $X^*_{n,m}$ from F_n. For convenience, it is usually assumed $m = n$. However, in the forthcoming sections, we do not necessarily assume $m = n$. We would like to have

$$P\left(\frac{\phi_n(X^*_{n,1},\ldots,X^*_{n,n};F_n)-b_n}{a_n}\leq x\,\middle|\, X_1,X_2,\ldots,X_n\right)\xrightarrow{a.s.} G(x). \tag{3.1.4}$$

We call this "the bootstrap consistency". As the sample mean is the most common and useful functional, most of articles focus on the bootstrap consistency of the mean.

Example 3.1.1. Let ϕ_n be the difference between the sample mean $\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$ and expectation $\mathrm{E}\,(X_1) = \int xdF$, i.e.

$$\phi_n(X_1,\ldots,X_n;F)=\bar{X}_n-\mathrm{E}\,(X_1).$$

Conditionally on $X_1, X_2, \ldots, X_n$, let

$$\bar{X}^*_{n,m}=\frac{1}{m}(X^*_{n,1}+\cdots+X^*_{n,m}),\quad \mathrm{E}^*(X^*_{n,1}):=\mathrm{E}\,(X^*_{n,1}|\ X_1,\ldots,X_n)=\int xdF_n$$

and

$$\phi_n(X^*_{n,1},\ldots,X^*_{n,m};F_n)=\bar{X}^*_{n,m}-\mathrm{E}^*(X^*_{n,1})=\bar{X}^*_{n,m}-\bar{X}_n.$$

Suppose $\sigma^2 = \mathrm{Var}\,(X_1) < \infty$. Then Bickel and Freedman [6] successfully prove the convergence (3.1.4) when G is $\mathcal{N}(0,\sigma^2)$, $b_n = 0$ and $a_n = n^{-1/2}$.

A noticeable question about Example 3.1.1 is whether we can discard the condition

$\mathrm{Var}(X_1) < \infty$ or more specifically what happens if the distribution is heavy-tailed. In the convergence (3.1.4), there are two important questions to be answered. The first question is whether m must be equal to n, and the other question is whether the convergence is in "a.s." sense. Indeed the discussion of the above greatly enriches the results of the bootstrap consistency of the mean. See for example, Athreya [2], Arcones and Giné [1], Knight [27], and Hall [22].

3.2 The regular bootstrap method on uniform polytopes on $\mathbb{R}^2$

Suppose $X_1, X_2, \ldots, X_n$ is a sequence of i.i.d. random points uniformly drawn from a polygon K with vertices $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_r}$ and angles $\theta_1, \theta_2, \ldots, \theta_r$. Let d_e be the Euclidean metric. Define the distance between a point $\boldsymbol{x} \in \mathbb{R}^2$ and a set $A \subset \mathbb{R}^2$ by $d_e(\boldsymbol{x}, A) := \inf\{d_e(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{y} \in A\}$. Let $K_n = \mathrm{Conv}\{X_1, \ldots, X_n\}$. Given the vertex $\boldsymbol{c_k}$, Bräker et al. [10] prove

$$\lim_{n\to\infty} P\left(\sqrt{n} d_e(\boldsymbol{c_k}, K_n) \le x\right) = 1 - p_k(x) \tag{3.2.1}$$

where

$$p_k(x) = \begin{cases} \displaystyle\int_0^{\theta_k} h_k(x,\theta)d\theta + \exp\left\{-\frac{x^2}{2|K|}\tan\theta_k\right\}, & 0 \le \theta_k < \pi/2 \\ \displaystyle\int_{\theta_k - \pi/2}^{\pi/2} h_k(x,\theta)d\theta, \quad \theta \ge \pi/2 & \end{cases}$$

with

$$h_k(x,\theta) = \exp\left\{-\frac{x^2}{2|K|}\big(\tan\theta_k + \tan(\theta_k - \theta)\big)\right\}\frac{x^2}{2|K|}\tan^2\theta.$$

See Figure 3.1 for $d_e(\boldsymbol{c_k}, K_n)$. Now conditionally on $X_1, X_2, \ldots, X_n$, the bootstrapping points $X^*_{n,1}, X^*_{n,2}, \ldots, X^*_{n,m}$ are drawn from the empirical distribution $F_n(-\infty, x] =$

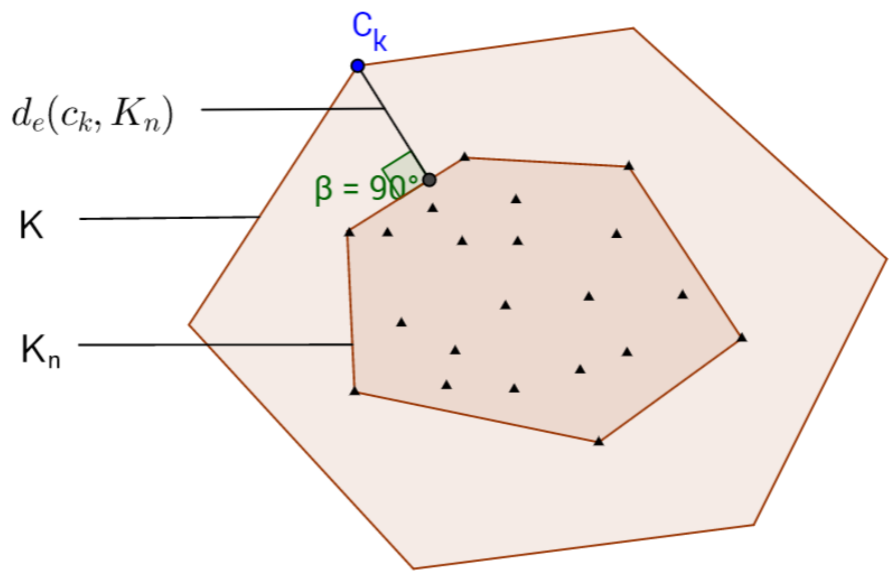

Figure 3.1: A situation of $d_e(\boldsymbol{c_k}, K_n)$, where the sample points are drawn uniformly from K, and K_n is the convex hull of the sample.

$\frac{1}{n}\sum_{i=1}^{n} I(X_i \leq x)$. Define $K^*_{n,m} := \text{Conv}\{X^*_{n,1}, X^*_{n,2}\ldots, X^*_{n,m}\}$. We would like to know whether the bootstrapping approximation is valid for $d_e(\boldsymbol{c_k}, K^*_{n,m})$. To show consistency, we will combine the techniques in Zarepour [40] and Bräker et al. [10]. Using the notations above, our theorem is stated as follows.

Theorem 3.2.1. *Use the same assumption above. Suppose $m = o(n)$, i.e. $m/n \to 0$, as $n \to \infty$. Then for any k, $0 \leq k \leq r$, we have*

$$\sqrt{m}d_e(\boldsymbol{c_k}, K^*_{n,m}) \xrightarrow{D} Z \quad \textit{in probability} \tag{3.2.2}$$

as $n, m \to \infty$, where Z has the distribution function $1 - p_k(x)$ defined in (3.2.1).

Proof: Our proof is divided into three steps.

Step 1: the construction of the original point process

Here and later, the additions and subtractions between a set and a point are Minkowski sum and difference (see Appendix A.8). Denote λ as Lebesgue measure. For any measurable subset $A \subset \mathbb{R}^2$, we have

$$\begin{aligned} nP(\sqrt{n}(X_1 - \boldsymbol{c_k}) \in A) &= nP\left(X_1 \in \left(\frac{A}{\sqrt{n}} + \boldsymbol{c_k}\right)\right) \\ &= \frac{n}{|K|}\lambda\left(\frac{A}{\sqrt{n}} \cap (K - \boldsymbol{c_k})\right) \\ &= \frac{1}{|K|}\lambda(A \cap \sqrt{n}(K - \boldsymbol{c_k})) \\ &\to \frac{1}{|K|}\lambda(A), \qquad \text{as } n \to \infty. \end{aligned} \tag{3.2.3}$$

This is due to the fact that $\sqrt{n}(K - \boldsymbol{c_k}) \uparrow \mathbb{R}^2$. Therefore according to Proposition A.6.1, we have

$$nP(\sqrt{n}(X_1 - \boldsymbol{c_k}) \in \cdot\ ) \xrightarrow{V} \frac{1}{|K|}\lambda(\cdot). \tag{3.2.4}$$

Define the point process $\xi_n = \sum_{i=1}^{n} \delta_{\sqrt{n}(X_i - \boldsymbol{c_k})}$. Here ξ_n can be regarded as a random element in $M_p(\text{Cone } \boldsymbol{c_k})$ (see Appendix A.2) endowed with the vague topology (see

Appendix A.6). The Laplace functional (see Appendix A.4) of ξ_n will be:

$$\begin{aligned}
\Phi_{\xi_n}(f) &= \left(\int_{\mathbb{R}^2} \exp\{-f(\sqrt{n}(\boldsymbol{x}-\boldsymbol{c_k}))\}\frac{1}{|K|}\mu(d\boldsymbol{x})\right)^n \quad (\mu = \lambda(\,\cdot \cap K)) \\
&= \left(1 - \frac{\int_{\mathbb{R}^2}(1-e^{f(\boldsymbol{y})})\mu^{'}(d\boldsymbol{y})}{n}\right)^n \quad \left(\boldsymbol{y} = \sqrt{n}(\boldsymbol{x}-\boldsymbol{c_k})\right) \\
&\to \exp\left\{-\int_{\mathbb{R}^2}(1-e^{f(\boldsymbol{y})})\mu^{'}(d\boldsymbol{y})\right\},
\end{aligned}$$

where as $n \to \infty$,

$$\mu^{'}(A) = \frac{n}{|K|}\lambda\left(\left(\frac{A}{\sqrt{n}} + \boldsymbol{c_k}\right)\cap K\right) = \frac{1}{|K|}\lambda\left(A \cap \sqrt{n}(K-\boldsymbol{c_k})\right) \to \frac{1}{|K|}\lambda(A)$$

for any measurable set $A \in \mathcal{R}^2$ (the Borel σ-field of $\mathbb{R}^2$). Hence $\xi_n \xrightarrow{D} \xi$, where ξ is a Poisson point process with intensity measure $\frac{1}{|K|}\lambda$ (see Appendix A.3 and A.5).

Step 2: the construction of the bootstrapping point process

Conditionally on $X_1, X_2, \dots, X_n$, let $\xi^*_{m,k} = \sum_{i=1}^{m} \delta_{\sqrt{m}(X^*_{n,i}-\boldsymbol{c_k})}$, which is also a random element in $M_p(\text{Cone } \boldsymbol{c_k})$. The Laplace functional of $\xi^*_{m,k}$ is

$$\begin{aligned}
\Phi_{\xi^*_{m,k}}(f) &= \left(\int_{\mathbb{R}^2} e^{-f\left(\sqrt{m}(\boldsymbol{x}-\boldsymbol{c_k})\right)} F_n(d\boldsymbol{x})\right)^m \\
&= \left(\int_{\mathbb{R}^2} e^{-f\left(\sqrt{m}(\boldsymbol{x}-\boldsymbol{c_k})\right)} \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}(d\boldsymbol{x})\right)^m \\
&= \left(\frac{1}{n}\sum_{i=1}^{n} e^{f\left(\sqrt{m}(X_i-\boldsymbol{c_k})\right)}\right)^m \\
&= \left(\int_{\mathbb{R}^2} e^{f(\boldsymbol{y})}\frac{1}{n}\sum_{i=1}^{n}\delta_{\sqrt{m}(X_i-\boldsymbol{c_k})}(d\boldsymbol{y})\right)^m \\
&= \left(1 - \frac{\int_{\mathbb{R}^2}(1-e^{-f(\boldsymbol{y})})\frac{m}{n}\sum_{i=1}^{n}\delta_{\sqrt{m}(X_i-\boldsymbol{c_k})}(d\boldsymbol{y})}{m}\right)^m.
\end{aligned} \tag{3.2.5}$$

If

$$\eta_{n,m,k} = \frac{m}{n} \sum_{i=1}^{n} \delta_{\sqrt{m}(X_i - \boldsymbol{c_k})}, \tag{3.2.6}$$

then the equation (3.2.5) becomes

$$\Phi_{\xi^*_{m,k}}(f) = \left(1 - \frac{\int_{\mathbb{R}^2}(1 - e^{-f(\boldsymbol{y})})\eta_{n,m,k}(d\boldsymbol{y})}{m}\right)^m. \tag{3.2.7}$$

Since $mP(\sqrt{m}(X_1 - \boldsymbol{c_k}) \in \cdot\ ) \xrightarrow{V} \frac{1}{|K|}\lambda(\cdot)$ and $(e^x - 1)/x \to 1$ as $x \to 0$, we have

$$\begin{aligned}
&\mathrm{E}\left[\exp\{t\eta_{n,m,k}(A)\}\right] \\
&= \left(e^{tm/n}P(\sqrt{m}(X_1 - \boldsymbol{c_k}) \in A) + 1 - P(\sqrt{m}(X_1 - \boldsymbol{c_k}) \in A)\right)^n \\
&= \left(\frac{mP(\sqrt{m}(X_1 - \boldsymbol{c_k}) \in A)\left(\frac{n}{tm}(e^{tm/n} - 1)\right)t}{n} + 1\right)^n \\
&\to \exp\left\{\frac{t}{|K|}\lambda(A)\right\}, \quad \text{as } n, m \to \infty.
\end{aligned} \tag{3.2.8}$$

Therefore with the help of relevant propositions of DC-semiring (Kallenberg [26]), used in Zarepour [40], we have

$$\eta_{n,m,k} \xrightarrow{P} \frac{1}{|K|}\lambda, \quad \text{as } n, m \to \infty. \tag{3.2.9}$$

Combining (3.2.9) with (3.2.7) implies that as $n, m \to \infty$,

$$\xi^*_{m,k} \xrightarrow{D} \xi \ \text{ in probability.} \tag{3.2.10}$$

Step 3: the continuous mapping theorem

For any measure $\eta \in M_p(\text{Cone } \boldsymbol{c_k})$, suppose f_k maps η to the smallest distance between the origin and the convex hull of the points of η. It is easy to find

$$f_{\boldsymbol{k}}(\xi^*_{m,k}) = \sqrt{m} d_e(\boldsymbol{c_k}, K^*_{n,m}).$$

Now apply the Skorokhod's representation theorem, Lemma 2.5.5 and the continuous mapping theorem (see Section 2.3 and 2.2) on (3.2.10) to get

$$\sqrt{m} d_e(\boldsymbol{c_k}, K^*_{n,m}) \xrightarrow{D} f_{\boldsymbol{k}}(\xi) \quad \text{in probability,} \tag{3.2.11}$$

as $n, m \to \infty$. From Bräker [10], we have

$$P(f_{\boldsymbol{k}}(\xi) \leq x) = 1 - p_k(x),$$

therefore the proof is complete. ∎

The following remark will be used in next chapter.

Remark 3.2.2. (cf. Zarepour [40]) Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random variables uniformly distributed on the interval $[a, b]$ with unknown parameters a and b. Also suppose $m = o(n)$. Like (3.2.3), for any $A \subset \mathbb{R}$, we can show

$$nP(nX_1 \in A) \to \frac{1}{b-a}\lambda(A) \tag{3.2.12}$$

where λ is the Lebesgue measure on $\mathbb{R}$ and therefore

$$nP(nX_1 \in \cdot\ ) \xrightarrow{V} \frac{1}{b-a}\lambda(\cdot). \tag{3.2.13}$$

Moreover, similar to (3.2.9), we can prove

$$\eta_{n,m} := \frac{m}{n}\sum_{i=1}^{n} \delta_{mX_i} \xrightarrow{P} \frac{1}{b-a}\lambda, \quad \text{as } n, m \to \infty. \tag{3.2.14}$$

Chapter 4

New bootstrapping scheme on uniform samples

4.1 Scope of the sample

Let X_i, $1 \leq i \leq n$ be a sequence of i.i.d. random variables uniformly distributed on the interval $[a, b]$ where both a and b are unknown parameters. The statistic $(\min_{1\leq i\leq n} X_i, \max_{1\leq i\leq n} X_i)$ is the maximum likelihood estimator for (a, b). Indeed,

$$[\min_{1\leq i\leq n} X_i, \max_{1\leq i\leq n} X_i]$$

is the smallest interval containing all the sample points. It is easy to show that

$$\min_{1\leq i\leq n} X_i \xrightarrow{a.s.} a, \quad \max_{1\leq i\leq n} X_i \xrightarrow{a.s.} b \tag{4.1.1}$$

which means the boundary of the random interval $[\min_{1\leq i\leq n} X_i, \max_{1\leq i\leq n} X_i]$ converges to the boundary of the support of the underlying distribution almost surely. Note that both $\min_{1\leq i\leq n} X_i$ and $\max_{1\leq i\leq n} X_i$ are monotone random variables in n. To understand our suggested bootstrapping scheme, notice that when the sample size n is very large, we can deem the interval $[\min_{1\le i\le n} X_i, \max_{1\le i\le n} X_i]$ as the true support of the underlying distribution and replace the unknown underlying distribution with the uniform distribution on $[\min_{1\le i\le n} X_i, \max_{1\le i\le n} X_i]$ in our simulations. As it can be seen this bootstrapping scheme is different from the regular bootstrapping scheme. Now given $X_1, X_2, \ldots, X_n$, let $X^*_{n,1}, X^*_{n,2}, \ldots, X^*_{n,m}$ be i.i.d. random points uniformly drawn from $[\min_{1\le i\le n} X_i, \max_{1\le i\le n} X_i]$.It is easy to see as $n \to \infty$,

$$\begin{aligned} P\left(n\left(\min_{1\le i\le n} X_i - a\right) > x,\ n\left(b - \max_{1\le i\le n} X_i\right) > y\right) &= \left(F\left(b - \frac{y}{n}\right) - F\left(a + \frac{x}{n}\right)\right)^n \\ &= \left(1 - \frac{y+x}{n(b-a)}\right)^n \\ &\to \exp\left\{-\frac{y+x}{b-a}\right\}, \qquad (4.1.2) \end{aligned}$$

and as $n, m \to \infty$, we have

$$\begin{aligned} &P\left(m\left(\min_{1\le i\le m} X^*_{n,i} - \min_{1\le i\le n} X_i\right) > x,\ m\left(\max_{1\le i\le n} X_i - \max_{1\le i\le m} X^*_{n,i}\right) > y \,\middle|\, X_1, \ldots, X_n\right) \\ &= \left(1 - \frac{y+x}{m\left(\max_{1\le i\le n} X_i - \min_{1\le i\le n} X_i\right)}\right)^m \\ &\xrightarrow{a.s.} \exp\left\{-\frac{y+x}{b-a}\right\}. \qquad (4.1.3) \end{aligned}$$

From (4.1.2) and (4.1.3), we can conclude that our bootstrap approximation is valid for almost all samples. This bootstrapping technique relies on the knowledge of uniformity of the distribution on $[a, b]$.

To give a comparison of our bootstrapping scheme with the regular bootstrapping scheme, conditionally on $X_1, X_2, \ldots, X_n$, let $X^{**}_{n,1}, X^{**}_{n,2}, \ldots, X^{**}_{n,m}$ be a sequence of i.i.d. random variables drawn from the empirical distribution F_n. Suppose $m = o(n)$, then using (3.2.14), as $n, m \to \infty$, we have

$$
\begin{aligned}
&P\left(m\left(\min_{1\le i\le m} X^{**}_{n,i} - \min_{1\le i\le n} X_i\right) > x,\ m\left(\max_{1\le i\le n} X_i - \max_{1\le i\le m} X^{**}_{n,i}\right) > y \,\middle|\, X_1,\dots,X_n\right)\\
&= F_n^m\left(\min_{1\le i\le n} X_i + \frac{x}{m}, \max_{1\le i\le n} X_i - \frac{y}{m}\right)\\
&= \left(1 - \frac{mF_n\left(\left(\min_{1\le i\le n} X_i, \min_{1\le i\le n} X_i + \frac{x}{m}\right) \cup \left(\max_{1\le i\le n} X_i - \frac{y}{m}, \max_{1\le i\le n} X_i\right)\right)}{m}\right)^m\\
&= \left(1 - \frac{\eta_{n,m}\left(\left(m\min_{1\le i\le n} X_i, m\min_{1\le i\le n} X_i + x\right) \cup \left(m\max_{1\le i\le n} X_i - y, m\max_{1\le i\le n} X_i\right)\right)}{m}\right)^m\\
&\xrightarrow{P} \exp\left\{-\frac{y+x}{b-a}\right\}. \qquad (4.1.4)
\end{aligned}
$$

From (4.1.4), we find that the asymptotic consistency of the regular bootstrap holds only in the sense of weak convergence in probability (not the weak convergence almost surely like in (4.1.3)). Moreover, to attain the consistency, we still need the additional assumption of $m = o(n)$ for the regular bootstrap. Therefore, our bootstrapping scheme is better than the regular bootstrap for this example. It should be emphasized that the uniformity knowledge was imposed on our case.

We call $SS(X_1, \dots, X_n) = [\min_{1\le i\le n} X_i, \max_{1\le i\le n} X_i]$ the scope of the sample $X_1, X_2, \dots, X_n$. $SS(X_1, \dots, X_n)$ plays an important role in our bootstrapping scheme where it suggests the replacement of the support of the underlying distribution with $SS(X_1, \dots, X_n)$. The above simple example was provided only for clarification of our bootstrapping scheme. Now we generalize the scope of the sample to spaces with higher dimensions.

Definition 4.1.1. *Suppose $X_1, X_2, \dots, X_n$ is a sequence of i.i.d. random points uniformly drawn from the distribution F with non-empty convex support $K \subset \mathbb{R}^d$. We say $SS(X_1, \dots, X_n)$ is the scope of the sample if $SS(X_1, \dots, X_n)$ is the maximum likelihood estimator of K.*

Like (4.1.1), it seems reasonable to add the following assumption,

$$H(K, SS(X_1, \ldots, X_n)) \xrightarrow{a.s.} 0 \tag{4.1.5}$$

where $H(\cdot, \cdot)$ is the Hausdorff distance (see Definition 1.3.1). So far, a formal definition for our bootstrapping scheme relies on resampling uniformly from $SS(X_1, \ldots, X_n)$ when (4.1.5) holds. Bearing this definition in mind, consider the following examples.

4.2 Two examples for our bootstrapping scheme on disks and rectangles

Example 4.2.1. Let d_e be the Euclidean metric on $\mathbb{R}^2$. Define the Euclidean norm $\|\boldsymbol{x}\| = d_e(\mathbf{0}, \boldsymbol{x}), \forall \boldsymbol{x} \in \mathbb{R}^2$. Denote $B(\mathbf{0}, r) := \{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| \leq r\}$. Let $X_1, X_2, \ldots, X_n$ be the i.i.d. random points uniformly drawn from the disk $C := B(\mathbf{0}, R)$. Let $C_n := B(\mathbf{0}, R_n)$ where

$$R_n = \max_{1 \leq i \leq n} \|X_i\|,$$

i.e. C_n is the smallest disk with center at origin and containing all the sample points. (See Figure 4.1.) It is easy to check that C_n is also the maximum likelihood estimator for C, i.e. $SS(X_1, \ldots, X_n) = C_n$. Since the sample points are drawn independently and uniformly from the C, we have

$$\begin{aligned} P(|R_n - R| > \varepsilon) &= P(R_n < R - \varepsilon) \\ &= \left(\frac{\pi(R-\varepsilon)^2}{\pi R^2}\right)^n \\ &\to 0 \quad \text{as } n \to \infty, \end{aligned}$$

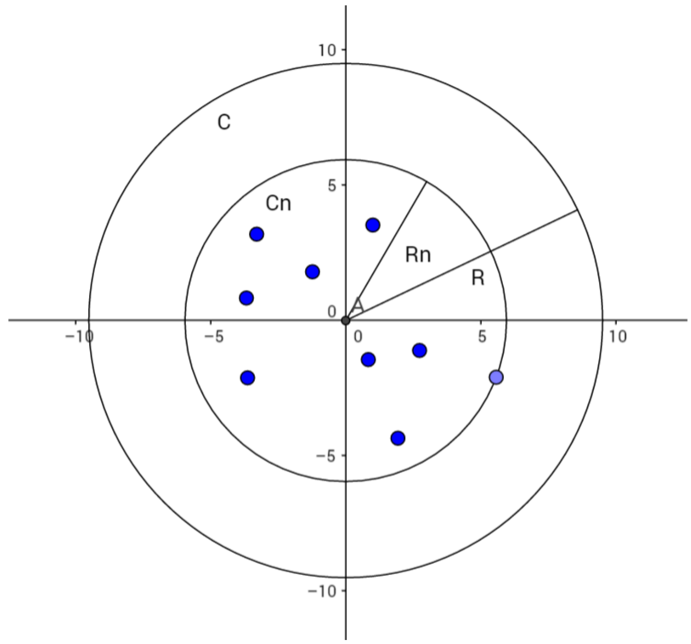

Figure 4.1: Draw sample points uniformly from the unit disk. C_n is the smallest disk with center at origin and containing all the sample points.

which implies $R_n \xrightarrow{P} R$. Indeed,

$$\operatorname{Var} R_n = \left(\frac{n}{n+1} - \frac{4n^2}{(2n+1)^2} \right) R^2 = \Theta(n^{-2}). \tag{4.2.1}$$

Therefore,

$$\sum_{n=1}^{\infty} P(|R_n - R| > \varepsilon) \leq \varepsilon^{-2} \sum_{n=1}^{\infty} \operatorname{Var} R_n < \infty, \tag{4.2.2}$$

which implies R_n converges to R completely (i.e. $R_n \xrightarrow{C} R$, which is stronger than almost sure convergence). Hence

$$H(C, SS(X_1, \ldots, X_n)) = H(C, C_n) = R - R_n \xrightarrow{a.s.} 0. \tag{4.2.3}$$

Therefore this case satisfies the conditions of our bootstrapping scheme. Now we can evaluate the asymptotic validity of other functionals using our bootstrapping scheme on this sample. Given $X_1, X_2, \ldots, X_n$, define

$$R^*_{n,m} = \max_{1\leq i\leq m} \|X^*_{n,i}\|$$

where $X^*_{n,i}$, $1 \leq i \leq m$, are drawn uniformly from C_n. Notice that as $n \to \infty$, we have

$$\begin{aligned} P(n(R - R_n) > x) &= P(R_n < R - x/n) \\ &= \left(\frac{\pi(R - x/n)^2}{\pi R^2}\right)^n \\ &= \left(1 - \frac{x}{nR}\right)^{2n} \\ &\to \exp\left\{-2x/R\right\}, \end{aligned} \tag{4.2.4}$$

and as $n, m \to \infty$, we have

$$\begin{aligned} P(m(R_n - R^*_{n,m}) > x|\ C_n) &= P(R^*_{n,m} < R_n - x/m|\ C_n) \\ &= \left(\frac{\pi(R_n - x/m)^2}{\pi R_n^2}\right)^m \\ &= \left(1 - \frac{x}{mR_n}\right)^{2m} \\ &\xrightarrow{a.s.} \exp\left\{-2x/R\right\}. \end{aligned} \tag{4.2.5}$$

From (4.2.4) and (4.2.5), it can be seen the bootstrap approximation is valid in this example. Notice that the standard condition of $m = o(n)$ (m out of n bootstrapping) is not required.

Example 4.2.2. Let (X_i, Y_i), $1 \leq i \leq n$ be a sequence of i.i.d. random points uniformly distributed on a rectangle $S = [0, a] \times [0, b]$, where $a,\ b > 0$ are two unknown parameters. Let $S_n := [\min_{1\leq i\leq n} X_i, \max_{1\leq i\leq n} X_i] \times [\min_{1\leq i\leq n} Y_i, \max_{1\leq i\leq n} Y_i]$, i.e. smallest rectangle which circumscribes all the sample points, and all the edges are parallel to the coordinate axises. It is easy to show S_n is the maximum likelihood estimator for S. Moreover, we can verify $H(S,S_n) \xrightarrow{a.s.} 0$. Indeed, it is easy to show

$$\min_{1\le i\le n} X_i \xrightarrow{a.s.} 0, \quad \max_{1\le i\le n} X_i \xrightarrow{a.s.} a, \quad \min_{1\le i\le n} Y_i \xrightarrow{a.s.} 0, \quad \max_{1\le i\le n} Y_i \xrightarrow{a.s.} b,$$

and we have

$$H(S,S_n) = \max\{R_{1,n}, R_{2,n}, R_{3,n}, R_{4,n}\} \xrightarrow{a.s.} 0 \tag{4.2.6}$$

where

$$R_{1,n} = \sqrt{\left(\min_{1\le i\le n} X_i\right)^2 + \left(\min_{1\le i\le n} Y_i\right)^2},$$

$$R_{2,n} = \sqrt{\left(\min_{1\le i\le n} X_i\right)^2 + \left(b - \max_{1\le i\le n} Y_i\right)^2},$$

$$R_{3,n} = \sqrt{\left(a - \max_{1\le i\le n} X_i\right)^2 + \left(b - \max_{1\le i\le n} Y_i\right)^2},$$

and

$$R_{4,n} = \sqrt{\left(a - \max_{1\le i\le n} X_i\right)^2 + \left(\min_{1\le i\le n} Y_i\right)^2}.$$

(See Figure 4.2.)

We use "∂" to represent boundary, thus ∂S is the four edges of S and ∂S_n is the four edges of S_n. For any $A, B \subset \mathbb{R}^2$, let $h(A,B)$ be the shortest horizontal or vertical distance between A and B, i.e.

$$h(A,B) = \min\{x_0, y_0\}$$

where

$$x_0 = \inf\{|x_1 - x_2| : (x_1, y) \in A,\ (x_2, y) \in B,\ y \in \mathbb{R}\}$$

and

$$y_0 = \inf\{|y_1 - y_2| : (x, y_1) \in A,\ (x, y_2) \in B,\ x \in \mathbb{R}\}.$$

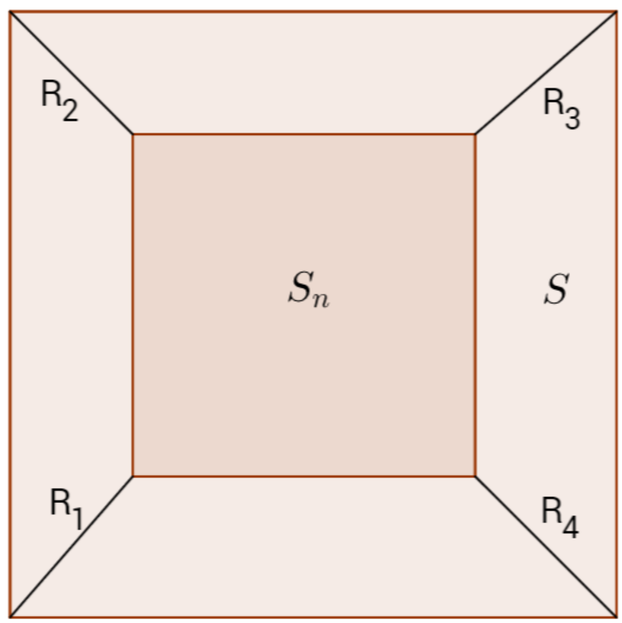

Figure 4.2: Draw sample points uniformly from S. S_n is the smallest rectangle which circumscribes all the sample points, and all the edges are parallel to the coordinate axises.

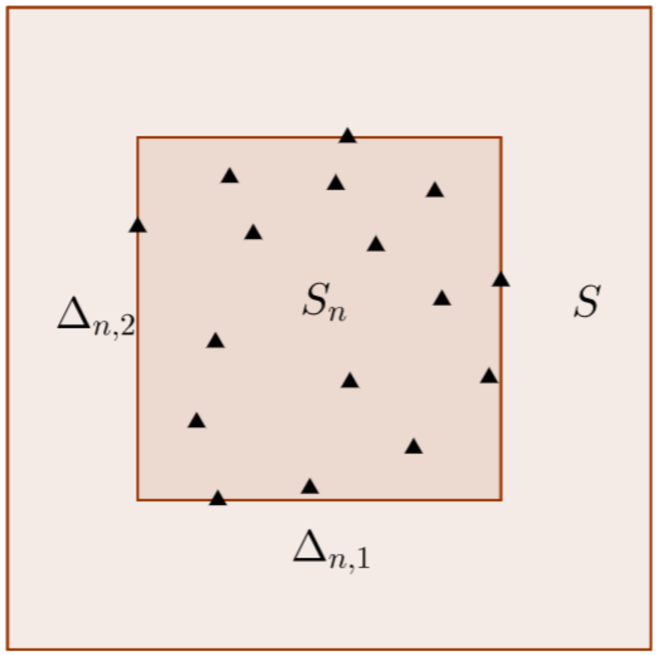

Figure 4.3: The rectangle S_n with $\Delta_{n,1}$ and $\Delta_{n,2}$.

Therefore as $n \to \infty$, we have

$$\begin{aligned} P(nh(\partial S, \partial S_n) > x) &= P(h(\partial S, \partial S_n) > x/n) \\ &= P^n(h(\partial S, X_i) > x/n) \\ &= a^{-1}b^{-1}(a - 2x/n)^n(b - 2x/n)^n \\ &\to e^{-2x(a^{-1}+b^{-1})}. \end{aligned} \tag{4.2.7}$$

Given S_n, let $(X^*_{n,i}, Y^*_{n,i})$, $1 \le i \le m$, be a sequence of i.i.d. random points uniformly drawn from S_n. Denote

$$\Delta_{n,1} := \max_{1\le i\le n} X_i - \min_{1\le i\le n} X_i, \quad \Delta_{n,2} := \max_{1\le i\le n} Y_i - \min_{1\le i\le n} Y_i.$$

(See Figure 4.3.) It is easy to show

$$\Delta_{n,1} \xrightarrow{a.s.} a, \quad \Delta_{n,2} \xrightarrow{a.s.} b.$$

Therefore, as $n, m \to \infty$, we have

$$\begin{aligned} P(mh(\partial S_n, \partial S^*_{n,m}) > x|\ S_n) &= P(h(\partial S_n, \partial S^*_{n,m}) > x/m|\ S_n) \\ &= P^m(h(\partial S_n, X^*_{n,i}) > x/m|\ S_n) \\ &= \Delta_{n,1}^{-m}\Delta_{n,2}^{-m}\left(\Delta_{n,1} - \frac{2x}{m}\right)^m \left(\Delta_{n,2} - \frac{2x}{m}\right)^m \\ &\xrightarrow{a.s.} e^{-2x(a^{-1}+b^{-1})}. \end{aligned} \tag{4.2.8}$$

From (4.2.7) and (4.2.8), we can conclude that our bootstrap approximation is valid for the functional h in this example. Again, the condition $m = o(n)$ is not necessary.

Remark 4.2.3. Notice that the convergence (4.2.5) and the convergence (4.2.8) are both in almost sure sense. Indeed, when the underlying distribution F and the target functional are more complicated, it is hard to derive such strong results directly by calculating the conditional probability like in the convergence (4.2.5) and the convergence (4.2.8). Therefore, we choose to control the resample size and construct the convergence in probability. Now we seek a more general resampling scheme for more complex functionals and we let resample size be flexible as long as asymptotic consistency holds in probability.

4.3 New bootstrapping scheme on general convex support

In Example 4.2.1 and Example 4.2.2, the support of the underlying distribution is either a disk or a rectangle. Now we move toward a more general set. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. random points uniformly drawn from an unknown non-empty convex set K. We would like to examine whether we can apply our new bootstrap scheme to any unknown convex support. We also follow the same procedures as in Example 4.2.1 and Example 4.2.2. We use the following steps:

1) Find the scope of the sample, $SS(X_1, \ldots, X_n)$, by calculating the maximum likelihood estimator of the convex support K.

2) Prove

$$H(K, SS(X_1, \ldots, X_n)) \xrightarrow{a.s.} 0.$$

Suppose $\lambda(\cdot)$ is Lebesgue measure. Let $K_n = \text{Conv}\{X_1, \ldots, X_n\}$. We get the join probability distribution

$$\begin{aligned} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) &= \lambda(K)^{-n} I\left(x_i \in K, 1 \leq i \leq n\right) \\ &= \lambda(K)^{-n} I\left(\text{Conv}\{x_1, \ldots, x_n\} \subset K\right) \end{aligned}$$

Therefore, K_n is the maximum likelihood estimator for K and $SS(X_1, \ldots, X_n) = K_n$. Remember that in Chapter 1, we discussed that Dümbgen and Walther [17] show $H(K, K_n) = O((\log n/n)^{1/k}))$ for any convex set $K \subset \mathbb{R}^d$. Indeed, the condition can be more general. Suppose the underlying distribution F has the convex support $K \subset \mathbb{R}^d$. Denote the boundary of K by ∂K. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. random points drawn from F. If there exist constants $c, k, h > 0$ such that for any $\delta \in [0, h]$,

$$\inf_{\boldsymbol{z} \in \partial K} F(B(\boldsymbol{z}, \delta)) \geq c\delta^k,$$

then

$$H(K, K_n) = O((\log n/n)^{1/k}),$$

where $B(\boldsymbol{z}, \delta)$ denotes the d-dimensional ball centering at z with the radius of δ. (See Dümbgen and Walther [17] for details.) It seems that our bootstrapping scheme can also be applied on any non-empty convex support.

The maximum likelihood estimator for the convex support is exactly the uniform polytopes. As we introduced in Chapter 1, the uniform polytopes have been heavily studied in the last few decades. Many of the existing work can help us to construct the weak convergence results like in (4.2.4) and (4.2.7). Moreover, the uniform polytopes inherit many geometric advantages from polytopes. For example, using the same assumptions in Theorem 3.2.1, the Hausdorff distance $H(K, K_n)$ can be easily calculated:

$$H(K, K_n) = \sup_{1 \leq k \leq r} d_e(\boldsymbol{c_k}, K_n)$$

where $\boldsymbol{c_k}$, $1 \leq k \leq r$ are all the vertices of the convex polygon K. To the best of our knowledge, we are the first to propose this new bootstrapping scheme and apply it on the functionals of random polytopes. More complex cases with more details are derived in the next chapter.

Chapter 5

The bootstrap consistency of the Hausdorff distance on uniform polytopes

Suppose F is the uniform distribution on an unknown convex set $K \subset \mathbb{R}^d$ ($d \geq 2$). Let F_A be the restriction of F on a set $A \subset \mathbb{R}^d$, i.e.

$$F_A(B) = \frac{F(B \cap A)}{F(A)}, \quad \forall B \subset \mathbb{R}^d. \tag{5.0.1}$$

Let $X_1, X_2, \ldots, X_n$ be a sequence of i.i.d. random points drawn from F. Let $K_n = \operatorname{Conv}\{X_1, \ldots, X_n\}$. In Chapter 4, we proved that the scope of the sample

$$SS(X_1, \ldots, X_n) = K_n,$$

i.e. K_n is the maximum likelihood estimator for K. Also we discussed that

$$H(K, K_n) \xrightarrow{a.s.} 0$$

where $H(K, K_n)$ is the Hausdorff distance between K and K_n (see Section 1.3.1). This demonstrates that our bootstrapping scheme can be applied in this case.

Suppose

$$\frac{\phi(X_1, \ldots, X_n; F) - \beta_n}{\alpha_n} \xrightarrow{D} Z. \tag{5.0.2}$$

where α_n and β_n are real numbers and Z is a random variable. Since F is the uniform distribution on K, F_{K_n} is the uniform distribution on K_n. Let $X^*_{n,1}, X^*_{n,2}, \ldots, X^*_{n,m}$ be a sequence of i.i.d. random points drawn from F_{K_n}. We would like to have the following two types of convergence,

$$(I) \qquad \frac{\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F) - \beta_m}{\alpha_m} \xrightarrow{D} Z, \quad \text{as } n, m \to \infty, \text{ in probability.}$$

$$(II) \qquad \frac{\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F_{K_n}) - \beta_m}{\alpha_m} \xrightarrow{D} Z, \quad \text{as } n, m \to \infty, \text{ in probability.}$$

From Theorem 1.4.6 and Theorem 1.4.7, the convergence in (5.0.2) holds if we let

$$\phi(X_1, \ldots, X_n; F) = H(K, K_n).$$

Hence in Section 5.1 and Section 5.2, we will prove type I convergence where

$$\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F) = H(K, \text{Conv}\{X^*_{n,1}, \ldots, X^*_{n,m}\}),$$

and type II convergence where

$$\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F_{K_n}) = H(K_n, \text{Conv}\{X^*_{n,1}, \ldots, X^*_{n,m}\}).$$

In Section 5.3, we also derive results for the symmetric difference and the perimeter difference (see the forthcoming definitions). The methods, used to derive the convergence in (5.0.2), are derived by using Theorem 1.4.5 and Theorem 1.4.4 respectively. We only give brief proofs for these two examples since the technique is similar to this section.

Throughout this chapter, the following notations are used. Let F be the underlying distribution. For any point set A, denote the convex hull of A by $\text{Conv}\,(A)$. Let $S_n := \{X_i\}_{i=1}^n$ be a sequence of i.i.d. random points drawn from F. Given S_n, let F_{K_n} be the uniform distribution on $K_n = \text{Conv}\,(S_n)$. Let $S^*_{n,m} := \{X^*_{n,i}\}_{i=1}^m$ be a sequence of i.i.d. random points drawn from F_{K_n}. Let $S'_m := \{X'_i\}_{i=1}^m$ be another sequence of i.i.d. random points drawn from F and independent from S_n. For any set A, we denote $A^n = \underbrace{A \times \cdots \times A}_{n \text{ times}}$. For simplicity, we also use the following notations:

$$K'_m = \text{Conv}\,(S'_m),$$
$$K^*_{n,m} = \text{Conv}\,(S^*_{n,m}).$$

The following convergences are equivalent:

$$(1)\ m(1 - F(K_n)) \xrightarrow{P} 0, \text{ as } n, m \to \infty;$$
$$(2)\ m \log\,(F(K_n)) \xrightarrow{P} 0, \text{ as } n, m \to \infty. \tag{5.0.3}$$

(Note that $P(m(1 - F(K_n)) > m(1 - \exp\{-\varepsilon/m\})) \to 0$ iff
$P(|m \log\,(F(K_n))| > \varepsilon) \to 0$ where $\varepsilon > 0$ and $m(1 - \exp\,(-\varepsilon/m)) \to \varepsilon$, as $m \to \infty$.)

5.1 Uniform distribution on convex polygons

Suppose F is the uniform distribution on $K \subset \mathbb{R}^2$ where K is a convex polygon with r vertices $\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_r}$ and angles $\theta_1, \theta_2, \ldots, \theta_r$. From Theorem 1.4.6, we have

$$\sqrt{n} H(K, K_n) \xrightarrow{D} Z_1, \quad \text{as } n \to \infty, \tag{5.1.1}$$

where Z_1 has the distribution function

$$G_1(x) = \prod_{i=1}^{r}(1 - p_i(x)) \tag{5.1.2}$$

with

$$p_i(x) = \begin{cases} \displaystyle\int_0^{\theta_i} h_i(x,\theta)d\theta + \exp\left\{-\frac{x^2}{2|K|}\tan\theta_i\right\}, & 0 < \theta_i < \frac{\pi}{2} \\ \displaystyle\int_{\theta_i-\pi/2}^{\pi/2} h_i(x,\theta)d\theta, & \frac{\pi}{2} \le \theta_i < \pi \end{cases}$$

and

$$h_i(x,\theta) = \exp\left\{-\frac{x^2}{2|K|}\big(\tan\theta_i + \tan(\theta_i - \theta)\big)\right\}\frac{x^2}{2|K|}\tan^2\theta.$$

The following theorem shows that

$$\sqrt{m}H(K, K^*_{n,m}) \xrightarrow{D} Z_1 \text{ in probability,}$$

if $n, m \to \infty$ satisfy certain relationship. This means for any continuity point of G_1, say $x \in \mathbb{R}$, as $n, m \to \infty$, we have

$$P(\sqrt{m}H(K, K^*_{n,m}) \le x|\ S_n) \xrightarrow{P} G_1(x).$$

Theorem 5.1.1. *Let $m \log n/n \to 0$, as $n \to \infty$. Then as $n, m \to \infty$, we have*

$$\sqrt{m}H(K, K^*_{n,m}) \xrightarrow{D} Z_1 \quad \textit{in probability,}$$

where Z_1 has the distribution function $G_1(x)$ as in Theorem 1.4.6.

Proof: Define the event:

$$E_{n,m} \triangleq \{S'_m \subset K_n\}.$$

First, we show that as $n, m \to \infty$,

$$P(E_{n,m}|\ S_n) \xrightarrow{P} 1. \tag{5.1.3}$$

This is equivalent to

$$\log\left(P(E_{n,m}|\ S_n)\right) \to 0. \tag{5.1.4}$$

Notice that $P(E_{n,m}|\ S_n) = (|K_n|/|K|)^m$. Then we have

$$\log\left(P(E_{n,m}|\ S_n)\right) = m \log\left(|K_n|/|K|\right). \tag{5.1.5}$$

Let $|D_n| := |K \setminus K_n| = |K| - |K_n|$. Using the equivalence in (5.0.3), we only need to show

$$m|D_n|/|K| \xrightarrow{P} 0. \tag{5.1.6}$$

Notice that Theorem 1.4.2 shows $(|D_n|/|K| - \beta_n)/\alpha_n$ converges weakly where $\alpha_n = \frac{1}{n}\sqrt{\frac{28}{27} r \log n}$ and $\beta_n = \frac{2}{3n} r \log n$, and also notice that as $n, m \to \infty$, $m\alpha_n \to 0$ and $m\beta_n \to 0$ according to our assumption about m. As $n, m \to \infty$, we have

$$\begin{aligned} m|D_n|/|K| &= m\left(\alpha_n \frac{|D_n|/|K| - \beta_n}{\alpha_n} + \beta_n\right) \\ &= m\alpha_n \left(\frac{|D_n|/|K| - \beta_n}{\alpha_n}\right) + m\beta_n \\ &\xrightarrow{P} 0 \end{aligned} \tag{5.1.7}$$

Now we develop our proof for Theorem 5.1.1. Define the function $f_n : \mathbb{R}^n \to \mathbb{R}$, such that

$$f_n(x_1, \ldots, x_n) = \sqrt{n} H(K, \operatorname{Conv}\{x_1, \ldots, x_n\}). \tag{5.1.8}$$

From Theorem 1.4.6, we get

$$f_n(X_1, \dots, X_n) \xrightarrow{D} Z_1, \quad \text{as } n \to \infty, \tag{5.1.9}$$

i.e. for any continuity point of G_1, say $x \in \mathbb{R}$,

$$P(f_n(X_1, \dots, X_n) \leq x) \to G_1(x), \quad \text{as } n \to \infty. \tag{5.1.10}$$

Let $A_n := (f_n^{-1}(-\infty, x]) \cap K^n$ where $f_n^{-1} : \mathbb{R} \to \mathbb{R}^n$ denotes the inverse mapping of f_n. Since, given S_n, the random vector $(X^*_{n,1}, \dots, X^*_{n,m})$ is independent from $E_{n,m}$, we have

$$\begin{aligned} P((X^*_{n,1}, \dots, X^*_{n,m}) \in A_m | \; S_n) & \\ &= P((X^*_{n,1}, \dots, X^*_{n,m}) \in A_m | \; S_n, E_{n,m}) \\ &= P((X^*_{n,1}, \dots, X^*_{n,m}) \in A_m \cap K_n^m | \; S_n, E_{n,m}) \\ &= P((X_1^{'}, \dots, X_m^{'}) \in A_m \cap K_n^m | \; S_n, E_{n,m}). \end{aligned} \tag{5.1.11}$$

The last equation holds because the joint distribution of $(X_1^{'}, \dots, X_m^{'})$ is the same as $(X^*_{n,1}, \dots, X^*_{n,m})$ when its distribution is restricted on K_n. According to lemma 2.5.1, we have

$$\begin{aligned} P((X_1^{'}, \dots, X_m^{'}) \in A_m \cap K_n^m | \; S_n, E_{n,m}) P(E_{n,m} | \; S_n) & \\ &= P(\{(X_1^{'}, \dots, X_m^{'}) \in A_m \cap K_n^m\} \cap E_{n,m} | \; S_n) \\ &= P((X_1^{'}, \dots, X_m^{'}) \in A_m \cap K_n^m | \; S_n). \end{aligned} \tag{5.1.12}$$

From (5.1.3), we have $P(E_{n,m} | \; S_n) \xrightarrow{P} 1$, as $n \to \infty$. Therefore (5.1.12) implies

$$|P((X_1^{'}, \dots, X_m^{'}) \in A_m \cap K_n^m | \; S_n, E_{n,m}) - P((X_1^{'}, \dots, X_m^{'}) \in A_m \cap K_n^m | \; S_n)|$$

$$\xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \tag{5.1.13}$$

Since $(A_m \cap K_n^m) \subset A_m$ and $A_m \setminus (A_m \cap K_n^m) \subset D_n^m$, we have

$$\begin{aligned} |P((X_1', \dots, X_m') \in A_m \cap K_n^m | \, S_n) - P((X_1', \dots, X_m') \in A_m | \, S_n)| \\ &\le P((X_1', \dots, X_m') \in D_n^m | \, S_n) \\ &= (|D_n|/|K|)^m \\ &\xrightarrow{a.s.} 0, \quad \text{as } n, m \to \infty. \end{aligned} \tag{5.1.14}$$

(Note that for almost all $\omega \in \Omega$, $|D_n(\omega)|/|K| < |D_3(\omega)|/|K| < 1$ implies $(|D_n(\omega)|/|K|)^m \to 0$, as $n, m \to \infty$.) Thus combining (5.1.11), (5.1.13) and (5.1.14) implies that

$$\begin{aligned} |P((X_{n,1}^*, \dots, X_{n,m}^*) \in A_m | \, S_n) - P((X_1', \dots, X_m') \in A_m | \, S_n)| \\ \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \end{aligned}$$

This is equivalent to:

$$\begin{aligned} |P(f_m(X_{n,1}^*, \dots, X_{n,m}^*) \le x | \, S_n) - P(f_m(X_1', \dots, X_m') \le x | \, S_n)| \\ \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \end{aligned} \tag{5.1.15}$$

Since $f_m(X_1', \dots, X_m')$ is independent from S_n, using (5.1.10), we have

$$\begin{aligned} P(f_m(X_1', \dots, X_m') \le x | \, S_n) &= P(f_m(X_1', \dots, X_m') \le x) \\ &\to G_1(x), \quad \text{as } n, m \to \infty. \end{aligned} \tag{5.1.16}$$

Finally, (5.1.15) and (5.1.16) imply

$$P(f_m(X_{n,1}^*, \dots, X_{n,m}^*) \le x | \, S_n) \xrightarrow{P} G_1(x), \quad \text{as } n, m \to \infty. \tag{5.1.17}$$

∎

Theorem 5.1.2. *Under the same conditions of the Theorem 5.1.1, we have*

$$\sqrt{m}H(K_n, K^*_{n,m})) \xrightarrow{D} Z_1, \quad \text{as } n, m \to \infty,$$

in probability.

Proof: In the proof of Theorem 5.1.1, given $S_n = \{X_i\}_{i=1}^n$, we reset $A_m = (g_m^{-1}(-\infty, x]) \cap K^m$ where

$$g_m(x_1, \ldots, x_m) = \sqrt{m}H(K_n, \text{Conv}\{x_1, \ldots, x_m\}).$$

Let x be a continuity point of G_1. Then follow the steps of the proof of Theorem 5.1.1 to get

$$|P((X^*_{n,1}, \ldots, X^*_{n,m}) \in A_m | \ S_n) - P((X'_1, \ldots, X'_m) \in A_m | \ S_n)| \xrightarrow{P} 0, \quad \text{as } n, m \to \infty.$$

This is equivalent to:

$$|P(g_m(X^*_{n,1}, \ldots, X^*_{n,m}) \le x | \ S_n) - P(g_m(X'_1, \ldots, X'_m) \le x | \ S_n)| \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \tag{5.1.18}$$

Therefore, we have

$$\left| P\left(\sqrt{m}H(K_n, K^*_{n,m}) \le x | \ S_n\right) - P\left(\sqrt{m}H(K_n, K'_m) \le x | \ S_n\right)\right| \xrightarrow{P} 0, \tag{5.1.19}$$

as $n, m \to \infty$. From Theorem 5.1.1, we have

$$P\left(\sqrt{m}H(K_n, K^*_{n,m}) \le x|\ S_n\right) \ge P\left(\sqrt{m}H(K, K^*_{n,m}) \le x|\ S_n\right) \tag{5.1.20}$$

and

$$P\left(\sqrt{m}H(K, K^*_{n,m}) \le x|\ S_n\right) \xrightarrow{P} G_1(x), \tag{5.1.21}$$

as $n, m \to \infty$. Notice that $H(K, K^*_{n,m}) \ge H(K_n, K^*_{n,m})$. Suppose $G_{1,m}$ is the distribution function of $f_m(X'_1, \ldots, X'_m)$, where f_m is the same as in the proof of Theorem 5.1.1. Then $\lim_{m\to\infty} G_{1,m}(x) = G_1(x)$. Since the Hausdorff distance satisfies triangle inequality:

$$H(K_n, K'_m) \ge H(K, K'_m) - H(K, K_n), \tag{5.1.22}$$

we have

$$\begin{aligned} P\left(\sqrt{m}H(K_n, K'_m) \le x|\ S_n\right) &\le P\left(\sqrt{m}(H(K, K'_m) - H(K, K_n)) \le x|\ S_n\right) \\ &= P\left(\sqrt{m}H(K, K'_m) \le \sqrt{m}H(K, K_n) + x|\ S_n\right) \\ &= G_{1,m}\left(\sqrt{\frac{m}{n}}(\sqrt{n}H(K, K_n)) + x\right) \\ &\xrightarrow{P} G_1(x), \quad \text{as } n, m \to \infty, \end{aligned} \tag{5.1.23}$$

where the last convergence is given by Theorem 1.4.6 and a simple application of the Skorokhod's representation theorem. Then use lemma 2.5.4, (5.1.19), (5.1.20), (5.1.21) and (5.1.23) to get

$$P\left(\sqrt{m}H(K_n, K^*_{n,m}) \le x|\ S_n\right) \xrightarrow{P} G_1(x), \quad \text{as } n, m \to \infty.$$

∎

5.2 Uniform distribution on convex sets with smooth boundary

The similar technique can also be applied to the smooth case. Let $K \subset \mathbb{R}^2$ be a convex set with $\partial K \subset \mathcal{C}^2(\mathbb{R})$ (see Chapter 1). Suppose the curvature of ∂K is bounded away from 0 and ∞ and has a bounded derivative. Suppose our underlying distribution F is the uniform distribution on K. Also suppose F satisfies the equation (1.4.17). From Theorem 1.4.7, we get

$$\frac{H(K, K_n) - b_n}{a_n} \xrightarrow{D} Z_2, \quad \text{as } n \to \infty. \tag{5.2.1}$$

The following theorem proves

$$\frac{H(K, K^*_{n,m}) - b_m}{a_m} \xrightarrow{D} Z_2, \quad \text{as } n, m \to \infty, \quad \text{in probability}, \tag{5.2.2}$$

where $P(Z_2 \leq x) = G_2(x) = \exp(-d_1 e^{-d_2 x})$ with the same d_1 and d_2 in Theorem 1.4.7. This means for any continuity point of G_2, say $x \in \mathbb{R}$, as $n, m \to \infty$, we have

$$P\left(\frac{H(K, K^*_{n,m}) - b_m}{a_m} \leq x \,\middle|\, S_n \right) \xrightarrow{P} G_2(x). \tag{5.2.3}$$

Theorem 5.2.1. *Let $m/n^{2/3} \to 0$, as $n \to \infty$. Then we have*

$$\frac{H(K, K^*_{n,m}) - b_m}{a_m} \xrightarrow{D} Z_2, \quad \text{as } n, m \to \infty,$$

in probability, where Z_2 has the distribution $G_2(x) = \exp\{-d_1 e^{-d_2 x}\}$, and a_m, b_m and G_2 are all defined in Theorem 1.4.7.

Proof: Our proof is a mimic of the proof for Theorem 5.1.1. Define the events

$E_{n,m} \triangleq \{S'_m \subset K_n\}$. First we need to show as $n, m \to \infty$

$$P(E_{n,m}|\ S_n) \xrightarrow{P} 1. \tag{5.2.4}$$

Suppose $D_n = K \setminus K_n$. Since $P(E_{n,m}|\ S_n) = (|K_n|/|K|)^m$, taking log from both sides and using the equivalence in (5.0.3), then (5.2.4) becomes

$$m|D_n| \to 0. \tag{5.2.5}$$

Notice that

$$\begin{aligned} m|D_n| &= m(|K| - |K_n|) \\ &= \left(mn^{-2/3}\right)\left(n^{2/3}\mathrm{E}\left(|D_n|\right)\right) - \left(mn^{-5/6}\right)\left(n^{5/6}\left(|K_n| - \mathrm{E}\left(|K_n|\right)\right)\right) \end{aligned}$$

where $mn^{-2/3} \to 0$ and $mn^{-5/6} \to 0$ accroding to our assumption about m. Finally using the results in Theorem 1.4.4, we complete the proof for $P(E_{n,m}|\ S_n) \xrightarrow{P} 1$. Moreover, define the function $f_n : \mathbb{R}^n \to \mathbb{R}$ such that

$$f_n(x_1, \ldots, x_n) = \frac{H(K, \mathrm{Conv}\{x_1, \ldots, x_n\}) - b_n}{a_n}.$$

We only need to follow the steps of the proof of Theorem 5.1.1 and use the result of Theorem 1.4.7. Therefore for any x which is the continuity point of G_2, we get

$$P(f_m(X^*_{n,1}, \ldots, X^*_{n,m}) \le x|\ S_n) \xrightarrow{P} G_2(x), \quad \text{as } n, m \to \infty. \tag{5.2.6}$$

∎

Theorem 5.2.2. *Under the same conditions of the Theorem 5.2.1, we have*

$$\frac{H\left(K_n, K^*_{n,m}\right) - b_m}{a_m} \xrightarrow{D} Z_2, \quad \text{as } n, m \to \infty,$$

in probability.

Proof: Here we use the same technique as in Theorem 5.1.1. Given $S_n = \{X_i\}_{i=1}^n$, we reset $A_m = (g_m^{-1}(-\infty, x]) \cap K^m$ where

$$g_m(x_1, \ldots, x_m) = \frac{H(K_n, \operatorname{Conv}\{x_1, \ldots, x_m\}) - b_m}{a_m}.$$

Let x be a continuity point of G_2. Then follow the steps of the proof of Theorem 5.1.1 to get

$$|P((X^*_{n,1}, \ldots, X^*_{n,m}) \in A_m |\ S_n) - P((X'_1, \ldots, X'_m) \in A_m |\ S_n)| \\ \xrightarrow{P} 0, \quad \text{as } n, m \to \infty.$$

This is equivalent to:

$$|P(g_m(X^*_{n,1}, \ldots, X^*_{n,m}) \leq x |\ S_n) - P(g_m(X'_1, \ldots, X'_m) \leq x |\ S_n)| \\ \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \tag{5.2.7}$$

Therefore, we have

$$\left| P\left(\frac{H(K_n, K^*_{n,m}) - b_m}{a_m} \leq x \middle|\ S_n \right) - P\left(\frac{H(K_n, K'_m) - b_m}{a_m} \leq x \middle|\ S_n \right) \right| \xrightarrow{P} 0, \tag{5.2.8}$$

as $n, m \to \infty$. From Theorem 5.2.1, we have

$$P\left(\frac{H(K_n, K^*_{n,m}) - b_m}{a_m} \leq x \middle|\ S_n \right) \geq P\left(\frac{H(K, K^*_{n,m}) - b_m}{a_m} \leq x \middle|\ S_n \right) \tag{5.2.9}$$

and

$$P\left(\frac{H(K,K^*_{n,m})-b_m}{a_m}\leq x\,\middle|\, S_n\right)\xrightarrow{P} G_2(x), \tag{5.2.10}$$

as $n,m\to\infty$. Suppose $G_{2,m}$ is the distribution function of $f_m(X_1^{'},\ldots,X_m^{'})$, where f_m is the same as in the proof of Theorem 5.2.1. Then $\lim_{m\to\infty}G_{2,m}(x)=G_2(x)$. Since the Hausdorff distance satisfies triangle inequality:

$$H(K_n,K_m^{'})\geq H(K,K_m^{'})-H(K,K_n), \tag{5.2.11}$$

we have

$$\begin{aligned}
P\left(\frac{H(K_n,K_m^{'})-b_m}{a_m}\leq x\,\middle|\, S_n\right) &\leq P\left(\frac{H(K,K_m^{'})-H(K,K_n)-b_m}{a_m}\leq x\,\middle|\, S_n\right)\\
&= P\left(\frac{H(K,K_m^{'})-b_m}{a_m}\leq \frac{H(K,K_n)}{a_m}+x\,\middle|\, S_n\right)\\
&= G_{2,m}\left(\frac{H(K,K_n)}{a_m}+x\right).
\end{aligned} \tag{5.2.12}$$

Since $a_n/a_m\to 0$ and $b_n/a_m\to 0$, we get

$$\frac{H\left(K,K_n\right)}{a_m}=\frac{a_n}{a_m}\frac{H\left(K,K_n\right)-b_n}{a_n}+\frac{b_n}{a_m}\xrightarrow{D}0,\quad \text{as } n,m\to\infty. \tag{5.2.13}$$

Therefore, according to (5.2.13) and a simple application of the Skorokhod's representation theorem, we have

$$G_{2,m}\left(\frac{H(K,K_n)}{a_m}+x\right)\xrightarrow{P}G_2(x),\quad \text{as } n,m\to\infty. \tag{5.2.14}$$

Finally use lemma 2.5.4, (5.2.8), (5.2.9), (5.2.10), (5.2.12) and (5.2.14) to get

$$P\left(\frac{H(K_n,K^*_{n,m})-b_m}{a_m}\leq x|\ S_n\right)\xrightarrow{P}G_2(x),\quad \text{as } m\to\infty. \tag{5.2.15}$$

∎

Remark 5.2.3. We can also get another bonus convergence result:

$$\frac{H(K_n, K_m^{'}) - b_m}{a_m} \xrightarrow{D} Z_2, \quad \text{as } m \to \infty.$$

(This result is only a byproduct of our proof and has nothing to do with our bootstrapping scheme.) By lemma 2.5.3 and Theorem 1.4.7, it is sufficient to prove

$$\frac{H(K, K_n)}{a_m} \xrightarrow{D} 0, \quad \text{as } m \to \infty,$$

which is similar to (5.2.13).

5.3 Two more examples

Besides the Hausdorff distance, the same technique can also be applied to other functionals. Two noticeable examples are the symmetric difference and the perimeter difference.

The first example is the symmetric difference on $\mathbb{R}^d$. Suppose d (≥ 2) is a fixed integer and K is a convex set in $\mathbb{R}^d$ such that $\partial K \in \mathcal{C}^2(\mathbb{R}^{d-1})$ with positive Gaussian curvature (see Appendix A.7) bounded away from 0 and ∞. With these assumptions, we can use the results in Theorem 1.4.5 and Theorem 1.4.3.

Example 5.3.1. Let $\phi_1(A, B)$ be the area of symmetric difference between A and B, i.e. $\phi_1(A, B) = |(A \setminus B) \cup (B \setminus A)|$. Since $K_n \subset K$, we have $\phi_1(K, K_n) = |D_n|$. Theorem 1.4.5 proves that

$$n^{\frac{1}{2}+\frac{1}{d+1}} \left(|D_n| - \mathrm{E}\left(|D_n|\right)\right) \xrightarrow{D} \mathcal{N}(0, \sigma^2), \quad \text{as } n \to \infty, \tag{5.3.1}$$

and Theorem 1.4.3 gives

$$\lim_{n\to\infty} n^{\frac{2}{d+1}} E|D_n| = c, \tag{5.3.2}$$

where σ and c are constants. Rewrite (5.3.1) using ϕ_1, to get

$$\frac{\phi_1(K, K_n) - b_n}{a_n} \xrightarrow{D} \mathcal{N}(0,1), \quad \text{as } n \to \infty, \tag{5.3.3}$$

where $a_n = \sigma n^{-\frac{1}{2}-\frac{1}{d+1}}$ and $b_n = \mathrm{E}\,|D_n| = \Theta(n^{-\frac{2}{d+1}})$. Our goal is to find a proper assumption about m such that the type I convergence holds for ϕ_1, i.e. as $n, m \to \infty$, we have

$$\frac{\phi_1(K, K^*_{n,m}) - b_m}{a_m} \xrightarrow{D} \mathcal{N}(0,1), \quad \text{in probability.} \tag{5.3.4}$$

As before, define the events $E_{n,m} \triangleq \{S'_m \subset K_n\}$ for the positive integers m and n. To impose a proper assumption about m, we must ensure

$$P(E_{n,m}|\ S_n) = P^m\left(|K_n|/|K|\right) \xrightarrow{P} 1, \quad \text{as } n, m \to \infty. \tag{5.3.5}$$

Taking log from the both sides and using the equivalence in (5.0.3), we have

$$m|D_n| \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \tag{5.3.6}$$

Since

$$\begin{aligned} m|D_n| &= m\phi_1(K, K_n) &(5.3.7)\\ &= m\left[a_n \frac{\phi_1(K, K_n) - b_n}{a_n} + b_n\right] \\ &= ma_n\left[\frac{\phi_1(K, K_n) - b_n}{a_n}\right] + mb_n, &(5.3.8) \end{aligned}$$

the assumption required for m is

$$ma_n \to 0, \quad mb_n \to 0,$$

as $n, m \to \infty$. Therefore the resample size m should satisfy $m = o(n^{\frac{2}{d+1}})$. To get

$$\frac{\phi_1(K, K^*_{n,m}) - b_m}{a_m} \xrightarrow{D} \mathcal{N}(0,1), \quad \text{as } n, m \to \infty, \quad \text{in probability,} \tag{5.3.9}$$

the details are omitted and the proofs are only the mimics of the proofs for Theorem 5.1.1 and Theorem 5.2.1.

Moreover, to prove type II convergence for ϕ_1, i.e.

$$\frac{\phi_1(K_n, K^*_{n,m}) - b_m}{a_m} \xrightarrow{D} \mathcal{N}(0,1), \quad \text{as } n, m \to \infty, \quad \text{in probability,} \tag{5.3.10}$$

notice that

$$\phi_1(K, K^*_{n,m}) \geq \phi_1(K_n, K^*_{n,m})$$

and

$$\phi_1(K_n, K^{'}_m) \geq \phi_1(K, K^{'}_m) - \phi_1(K, K_n),$$

which are easy to verify. We still need to show

$$a_n/a_m \to 0, \quad b_n/a_m \to 0,$$

as $n, m \to \infty$, which are obvious.

Now we can consider the perimeter difference (defined below) on $\mathbb{R}^2$. To use the result in Theorem 1.4.4, suppose K satisfies $\partial K \in \mathcal{C}^2(\mathbb{R})$ with positive curvature bounded away from 0 and ∞.

Example 5.3.2. Define $\phi_2(A, B) = |\text{perimeter of } A - \text{perimeter of } B|$. Theorem

1.4.4 proves

$$n^{5/6}(L_n - \mathrm{E}\, L_n) \xrightarrow{D} \mathcal{N}(0, (\sigma^{'})^2), \quad \text{as } n \to \infty, \tag{5.3.11}$$

and

$$\lim_{n\to\infty} n^{2/3}\mathrm{E}\,(L - L_n) = c^{'}, \tag{5.3.12}$$

where $\sigma^{'}$ and $c^{'}$ are two constants. Rewriting (5.3.11) in terms of ϕ_2, we have

$$\frac{\phi_2(K, K_n) - b_n^{'}}{a_n^{'}} \xrightarrow{D} \mathcal{N}(0, 1), \quad \text{as } n \to \infty, \tag{5.3.13}$$

where $a_n^{'} = \sigma^{'} n^{-5/6}$ and $b_n^{'} = \mathrm{E}\,(L - L_n) = \Theta(n^{-2/3})$. To prove the type I convergence, i.e.

$$\frac{\phi_2(K, K_{n,m}^*) - b_n^{'}}{a_n^{'}} \xrightarrow{D} \mathcal{N}(0, 1), \quad \text{as } n, m \to \infty \ \text{ in probability}, \tag{5.3.14}$$

we need to show

$$P(E_{n,m}|\ S_n) = P^m\left(|K_n|/|K|\right) \xrightarrow{P} 1, \quad \text{as } n, m \to \infty, \tag{5.3.15}$$

Similar to Example 5.3.1, m should satisfy the conditions:

$$m|D_n| \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \tag{5.3.16}$$

as $n, m \to \infty$. Thus, our assumption for m should follow $m = o(n^{2/3})$. (Note that $d = 2$.)

To prove the type II convergence, i.e.

$$\frac{\phi_2(K_n, K_{n,m}^*) - b_n^{'}}{a_n^{'}} \xrightarrow{D} \mathcal{N}(0, 1), \quad \text{as } n, m \to \infty \ \text{ in probability}, \tag{5.3.17}$$

notice that

$$\phi_2(K, K^*_{n,m}) \geq \phi_2(K_n, K^*_{n,m})$$

and

$$\phi_2(K_n, K^{'}_m) \geq \phi_2(K, K^{'}_m) - \phi_2(K, K_n),$$

which are easy to verify using Lemma 2.5.2. We also need to show

$$a^{'}_n/a^{'}_m \to 0, \quad b^{'}_n/a^{'}_m \to 0,$$

as $n, m \to \infty$, which are obvious.

Chapter 6

New bootstrapping scheme on Gaussian polytopes

6.1 The general case

Suppose $S_n := \{X_i\}_{i=1}^n$ is a sequence of i.i.d. random points drawn from the distribution F. Let $K_n = \text{Conv}\{X_1, \ldots, X_n\}$. Suppose $S_m' := \{X_i'\}_{i=1}^m$ is another sequence of i.i.d. random points drawn from F and independent from S_n. If F is not the uniform distribution, say a Gaussian distribution with unknown parameters, the scope of the sample $SS(X_1, \ldots, X_n)$ (see Definition 4.1.1) is not meaningful as the support of F is $\mathbb{R}^d$. Therefore we can not apply the suggested bootstrapping scheme to this case. However, a careful review of our proof in Chapter 5 would suggest if the convergence (5.0.2) and (5.1.3) hold, i.e. the functional ϕ satisfies: there exist real numbers α_n, β_n and a random variable Z such that as $n \to \infty$,

$$\frac{\phi(X_1, \ldots, X_n; F) - \beta_n}{\alpha_n} \xrightarrow{D} Z, \tag{6.1.1}$$

and

$$P(E_{n,m} |\ S_n) \xrightarrow{P} 1, \tag{6.1.2}$$

where $E_{n,m} \triangleq \{S_m^{'} \subset K_n\}$, we can still construct the type I convergence, i.e.

$$\frac{\phi(X_{n,1}^*, \ldots, X_{n,m}^*; F) - \beta_m}{\alpha_m} \xrightarrow{D} Z, \quad \text{as } n, m \to \infty, \text{ in probability,}$$

by drawing the sample $X_{n,1}^*, X_{n,2}^*, \ldots, X_{n,m}^*$ from F_{K_n} (the restriction of F on K_n, see the definition in Chapter 5). Moreover if

$$\alpha_n/\alpha_m \to 0, \quad \beta_n/\alpha_m \to 0, \quad \text{as } n, m \to \infty,$$

and ϕ satisfies:

$$\phi(K, K_{n,m}^*) \geq \phi(K_n, K_{n,m}^*) \text{ and } \phi(K_n, K_m^{'}) \geq \phi(K, K_m^{'}) - \phi(K, K_n),$$

where K is the support of F,

$$\phi(K, K_n) = \phi(X_1, \ldots, X_n; F) \text{ and } \phi(K_n, K_{n,m}^*) = \phi(X_{n,1}^*, \ldots, X_{n,m}^*; F_{K_n}),$$

then the type II convergence can also be concluded, i.e.

$$\frac{\phi(X_{n,1}^*, \ldots, X_{n,m}^*; F_{K_n}) - \beta_m}{\alpha_m} \xrightarrow{D} Z, \quad \text{as } n, m \to \infty, \text{ in probability.}$$

Our bootstrapping scheme for general situation relies on drawing $X_{n,i}^*$, $1 \leq i \leq m$, from F_{K_n}. Suppose the convergence (6.1.1) holds. The two types of convergence we are concerned with are:

$$(I) \qquad \frac{\phi(X_{n,1}^*, \ldots, X_{n,m}^*; F) - \beta_m}{\alpha_m} \xrightarrow{D} Z, \quad \text{as } n, m \to \infty, \text{ in probability.}$$

and

$$(II) \qquad \frac{\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F_{K_n}) - \beta_m}{\alpha_m} \xrightarrow{D} Z, \quad \text{as } n, m \to \infty, \text{ in probability.}$$

From the discussion above, the critical points of our technique require to show the convergence (6.1.1) and the convergence (6.1.2). The construction of the convergence (6.1.1) for certain functional is beyond the scope of this thesis. However, the convergence (6.1.2) is worth discussing since it has nothing to do with the functional. For simplicity, let $S_n = \{X_i\}_{i=1}^n$. Suppose $S_m^{'} = \{X_i^{'}\}_{i=1}^m$ is a sequence of i.i.d. random points drawn from F. According to the convergence (6.1.2), we need to find a proper assumption about m such that

$$P(S_m^{'} \subset K_n | S_n) \xrightarrow{P} 1, \quad \text{as } n, m \to \infty. \tag{6.1.3}$$

Since

$$P(S_m^{'} \subset K_n | S_n) = F^m(K_n),$$

take log from the both sides and use the equivalent relation in (5.0.3), and then get

$$m(1 - F(K_n)) \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \tag{6.1.4}$$

The convergence (6.1.4) is not trivial. From the convergence (6.1.4), we can ensure the proper assumption about m. We have seen the cases for uniform distribution where $1 - F(K_n) = |D_n|$, i.e. $K \setminus K_n$. Since the weak convergence of $|D_n|$ can be transformed to the weak convergence of $F(K_n)$, we can proved (6.1.4) easily by using the existing results about $|D_n|$. However, the convergence (6.1.4) is very difficult to prove in general. In the next section, we will discuss the assumption about m when K_n is a Gaussian polytope.

6.2 Assumption about m with respect to Gaussian polytopes

For Gaussian polytopes in $\mathbb{R}^d$ ($d \geq 2$), To our knowledge, no weak convergence for $F(K_n)$ is available. However, we find a paragraph at the end of Bárány and Vu [4] which explains how to apply their technique on the probability content to construct the central limit theorem. In author's recent contact with Bárány, he acknowledged that "the result has never been published. It was clear that the proof method of the paper can be used for further results (including the central limit theorem for the probability content), but the paper was long enough anyway and we decided not to include it there."

Here we give a reasonable conjecture.

$$\frac{F(K_n) - \beta_n}{\alpha_n} \xrightarrow{D} \mathcal{N}(0,1), \quad \text{as } n \to \infty, \tag{6.2.1}$$

where $\alpha_n = \sqrt{\mathrm{Var}\,(F(K_n))}$, $\beta_n = \mathrm{E}\,(F(K_n))$. According to Theorem 1.4.10, we have

$$\mathrm{E}\,((F(K_n))^k) = \mathrm{E}\left(\prod_{i=1}^{k}\left(1 - \frac{N_{n+k}}{n+i}\right)\right). \tag{6.2.2}$$

From Theorem 1.4.8, we get

$$\mathrm{E}\,(N_n) = \Theta\left((\log n)^{(d-1)/2}\right), \tag{6.2.3}$$

and from Theorem 1.4.9, we get

$$\mathrm{Var}\,(N_n) = O\left((\log n)^{(d-1)/2}\right). \tag{6.2.4}$$

(6.2.3) and (6.2.4) imply

$$\mathrm{E}\left(N_n^2\right) = \Theta\left((\log n)^{d-1}\right). \tag{6.2.5}$$

Combining (6.2.2), (6.2.3) and (6.2.5) implies

$$\mathrm{E}\left(F(K_n)\right) = 1 + \Theta\left(\frac{(\log n)^{(d-1)/2}}{n}\right) \tag{6.2.6}$$

and

$$\mathrm{E}\left[(F(K_n))^2\right] = 1 + \Theta\left(\frac{(\log n)^{(d-1)/2}}{n}\right). \tag{6.2.7}$$

Then we get

$$\mathrm{Var}\left(F(K_n)\right) = O\left(\frac{(\log n)^{(d-1)/2}}{n}\right). \tag{6.2.8}$$

Finally, we get

$$\alpha_n = \sqrt{\mathrm{Var}\left(F(K_n)\right)} = O\left(\frac{(\log n)^{(d-1)/4}}{\sqrt{n}}\right)$$

and

$$1 - \beta_n = 1 - \mathrm{E}\left(F(K_n)\right) = \Theta\left(\frac{(\log n)^{(d-1)/2}}{n}\right).$$

According to our conjecture (6.2.1), if as $n, m \to \infty$,

$$m\alpha_n \to 0, \quad m(1-\beta_n) \to 0,$$

we get

$$m(1 - F(K_n)) = m\left(-\alpha_n \frac{F(K_n) - \beta_n}{\alpha_n} + 1 - \beta_n\right) \xrightarrow{P} 0, \quad \text{as } n, m \to \infty. \tag{6.2.9}$$

Therefore we can take

$$m = o\left(\sqrt{n}(\log n)^{-(d-1)/4}\right).$$

Chapter 7

Bootstrapping; simulation results

7.1 Algorithm for general cases

For general cases, our assumptions and notations are the same as in Chapter 6. For the functional ϕ and the underlying distribution F, suppose we have the following convergence result:

$$\frac{\phi(X_1, \ldots, X_n; F) - \beta_n}{\alpha_n} \xrightarrow{D} Z,$$

where Z has the distribution function G. This simulation procedure is designed to describe our results in Chapter 5. The simulation algorithm is as follows.

1. Draw n points $X_1, X_2, \ldots, X_n$ from the underlying distribution F. Find the convex polytope

$$K_n = \text{Conv}\{X_1, \ldots, X_n\}.$$

2. Draw m points $X^*_{n,1}, X^*_{n,2} \ldots, X^*_{n,m}$ from the distribution F_{K_n} (the restriction of F on K_n, see the definition in Chapter 5) and find convex polytope as

$$K^*_{n,m} = \text{Conv}\{X^*_{n,1}, \ldots, X^*_{n,m}\}.$$

3. Calculate the normalized target functional $T_{n,m}$ or $T^{(1)}_{n,m}$, where

$$T_{n,m} = \frac{\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F) - \beta_m}{\alpha_m}, \quad T^{(1)}_{n,m} = \frac{\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F_{K_n}) - \beta_m}{\alpha_m}.$$

4. Repeat the Step 1-3, N times and get a group of data $t_1, t_2, \ldots, t_N$ for $T_{n,m}$ or $T^{(1)}_{n,m}$.

5. Depict the empirical distribution F_{T_N} for $t_1, t_2, \ldots, t_N$.

6. Compare F_{T_N} with G (the limiting distribution of Z).

7.2 Bootstrapping for the Hausdorff distance on uniform polygons

7.2.1 The bootstrap simulations to verify Theorem 5.1.1

In Theorem 5.1.1, the limiting distribution function G in Step 6 of the algorithm is G_1 (see G_1 in Theorem 5.1.1). Here F is the uniform distribution on K, therefore F_{K_n} (in Step 2) is the uniform distribution on K_n. Let us consider the normalized target functional

$$T_{n,m} = \sqrt{m} H(K, K^*_{n,m}).$$

In this section, the empirical distribution F_{T_N} (in Step 5) is denoted by $F_{T_N,1}$. To construct an example of a uniform polygon, we draw 50 points uniformly from a unit disk and we take our underlying uniform polygon K as the convex hull of these 50 points (see Figure 7.1). Note that the size 50 is not the sample size and this step is only for constructing a uniform distribution on a given polygon.

To verify our assumption: $m \log n/n \to 0$ as $n \to \infty$, we take the sample size $n = 400$ and the resample size $m = 20$. Figure 7.2 provides the comparison of $F_{T_N,1}$ with G_1 for $N = 1000$ (the number of times to repeat in Step 4).

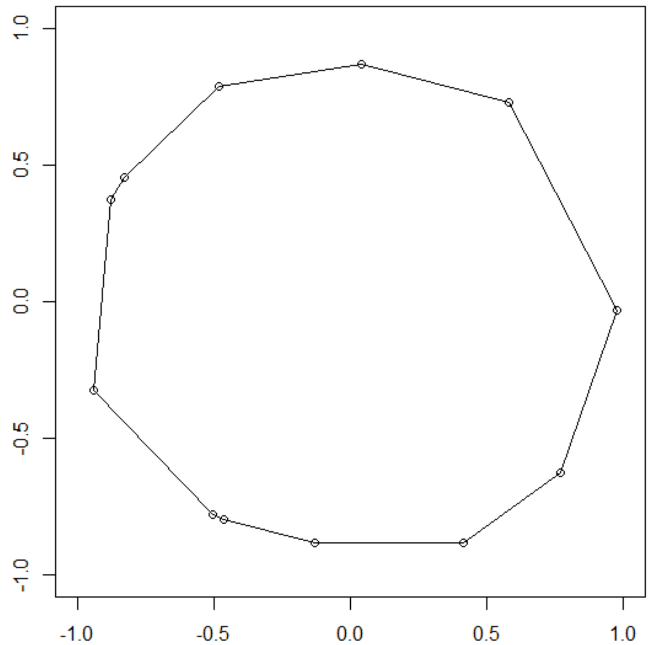

Figure 7.1: The underlying convex polygon K.

Since there is a gap between $F_{T_N,1}$ and G_1 in Figure 7.2, the simulation result does not seem to be precise. However, it is also possible for the gap to be narrowed with increasing n and m. To show this point, notice that in the proof of Theorem 5.1.1, we have

$$|P(f_m(X^*_{n,1},\ldots,X^*_{n,m}) \leq x|\ S_n) - P(f_m(X'_1,\ldots,X'_m) \leq x)| \xrightarrow{P} 0, \quad \text{as } m \to \infty, \tag{7.2.1}$$

where

$$P(f_m(X^*_{n,1},\ldots,X^*_{n,m}) \leq x|\ S_n) = P(\sqrt{m}H(K, K^*_{n,m}) \leq x|\ S_n) \tag{7.2.2}$$

and

$$P(f_m(X'_1,\ldots,X'_m) \leq x) = G_{1,m}(x) \tag{7.2.3}$$

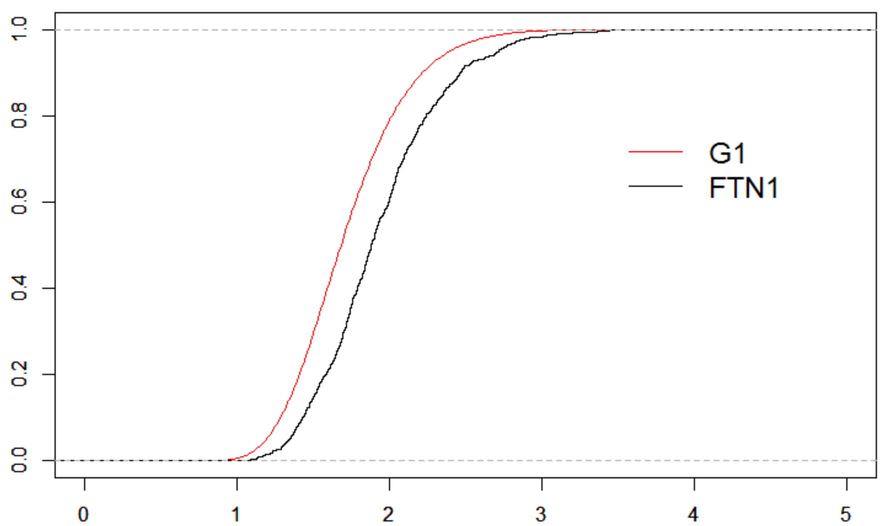

Figure 7.2: The comparison of $F_{T_N,1}$ with G_1 for $n = 400$, $m = 20$ and $N = 1000$.

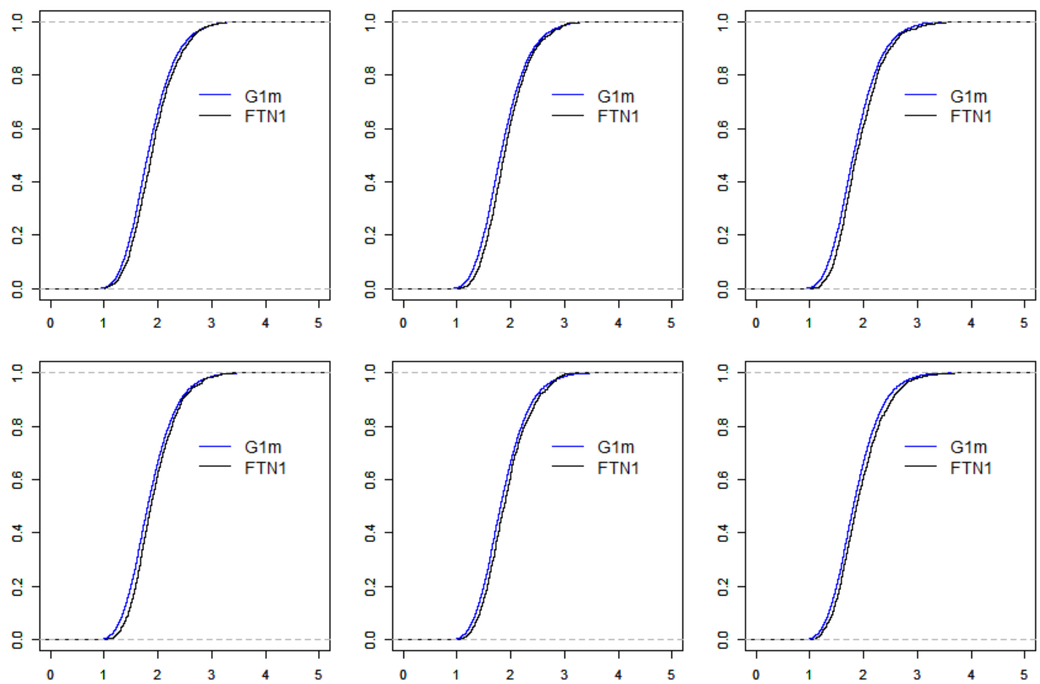

Figure 7.3: The comparisons of $F_{T_N,1}$ with $G_{1,20}$ for $n = 400$, $m = 20$ and $N = 1000$.

with

$$\lim_{m\to\infty} G_{1,m}(x) = G_1(x), \tag{7.2.4}$$

where x is any continuity point of G. Since the empirical distribution $F_{T_N,1}$ is exactly a simulation for $P(\sqrt{m}H(K, K^*_{n,m}) \leq x|S_n)$, (7.2.4) shows that $G_{1,m}$ is a reasonable approximation (therefore a good replacement) for G_1 in our simulation. Thus we make comparisons of $G_{1,20}$ and $F_{T_N,1}$ in Figure 7.3. To verify the stability of accuracy, we repeat the simulations 6 times.

In Figure 7.3, the gap between distributions disappears by replacing G_1 with $G_{1,20}$ in our simulations. Moreover, we depict a comparison of G_1 with $G_{1,20}$ (see Figure 7.4). As it can be seen there is a noticeable gap between G_1 and $G_{1,20}$ which is caused by inadequate sizes of n and m. (This gap is rooted in Theorem 1.4.6 and has nothing to do with our bootstrapping scheme.) To this point, it is reasonable to conclude that

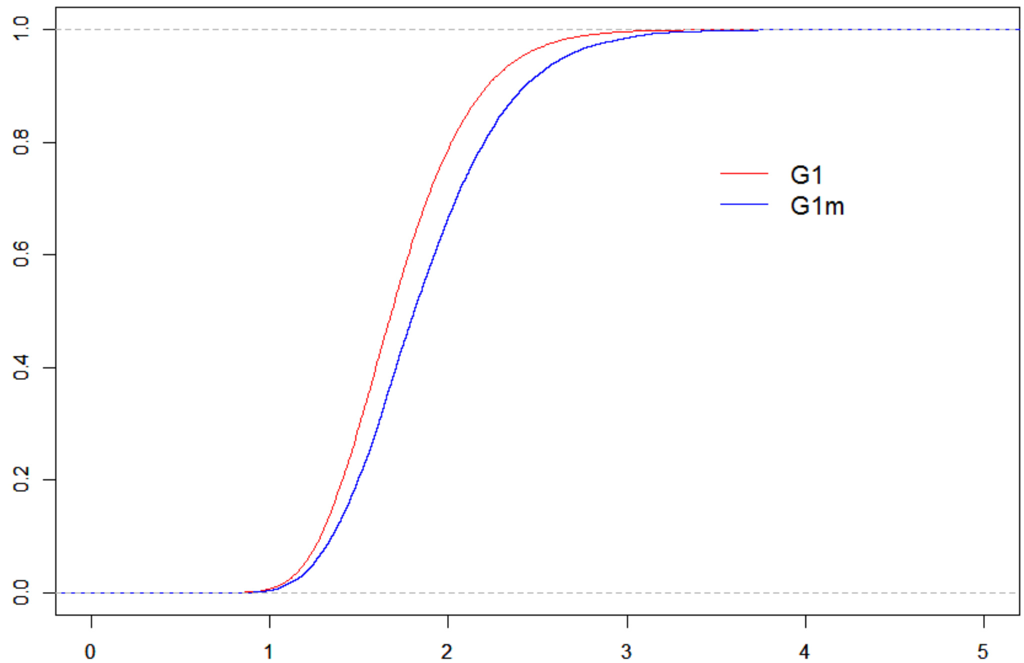

Figure 7.4: The comparison of G_1 with $G_{1,20}$ for $m = 20$ and $N = 1000$.

the gap between $F_{T_N,1}$ and G_1 in Figure 7.2 also disappears with the increases of n and m.

To verify whether our assumption about m (i.e. $m \log n/n \to 0$, as $n, m \to \infty$) is acceptable, we set the resample size $m = m_i$, $i = 1, 2, 3$, where

$$m_1 = \sqrt{n}, \quad m_2 = n^{2/3}, \quad m_3 = n. \tag{7.2.5}$$

It is obvious that m_1 and m_2 satisfy our assumption but m_3 does not. We compare $F_{T_N,1}$ for three different values of m in (7.2.5) with $G_{1,20}$ (see Figure 7.5). As we can see in Figure 7.5, when $m = m_1, m_2$, $F_{T_N,1}$ is close to $G_{1,20}$. However when $m = m_3$, $F_{T_N,1}$ is not a good approximation to $G_{1,20}$. This result also meets our expectations.

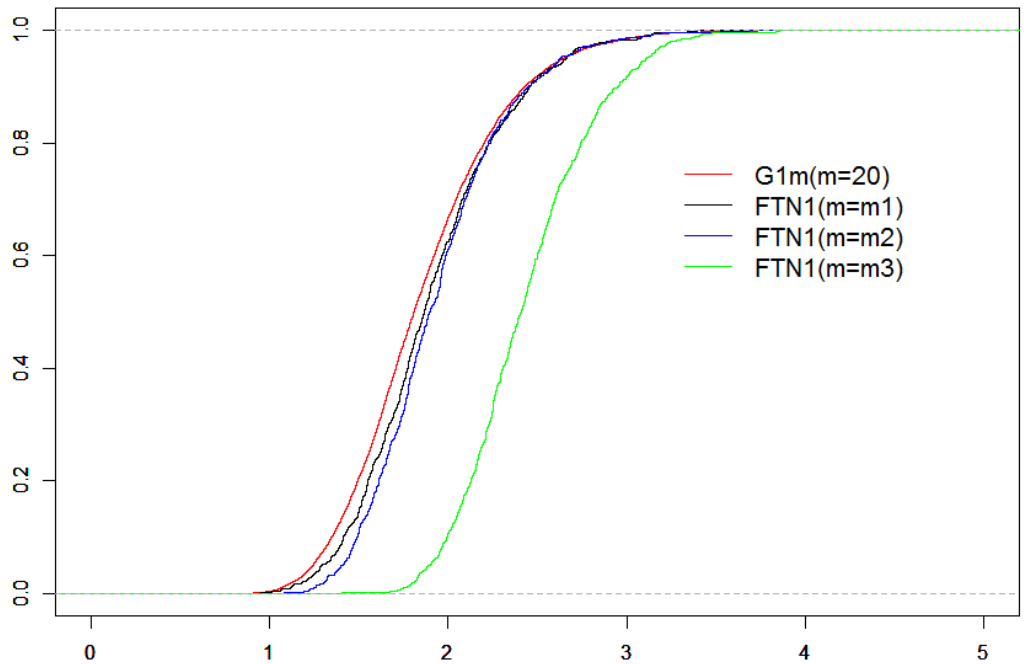

Figure 7.5: The comparison of $F_{T_N,1}$ with $G_{1,20}$ for $n = 400$, $m = m_1, m_2, m_3$ and $N = 1000$.

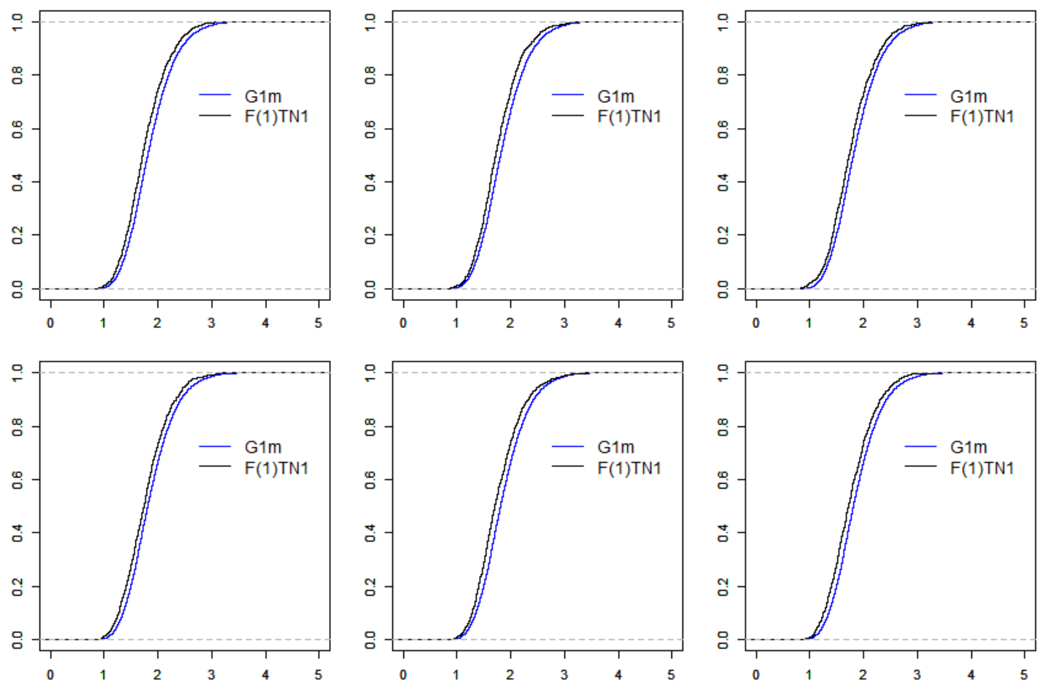

Figure 7.6: The comparisons of $F^{(1)}_{T_N,1}$ with $G_{1,20}$ for $n = 400$, $m = 20$ and $N = 1000$.

7.2.2 The bootstrap simulations to verify Theorem 5.1.2

The same situation happens for Theorem 5.1.2. Set $T^{(1)}_{n,m} = \sqrt{m}H(K_n, K_{n,m})$ (see Step 3 of our algorithm). In this section, the empirical distribution F_{T_N} (in Step 5) is denoted by $F^{(1)}_{T_N,1}$. Replacing G_1 with $G_{1,20}$, we repeat the simulations 6 times (see Figure 7.6). As it can be seen $G_{1,20}$ and $F^{(1)}_{T_N,1}$ are close, which also implies $F^{(1)}_{T_N,1}$ and G_1 are also close to each other if n and m are large enough.

Also to verify whether our assumption about m (i.e. $m \log n/n \to 0$, as $n, m \to \infty$) is acceptable, we set the resample size $m = m_i$, $i = 1, 2, 3$, where $m_1 = \sqrt{n}$, $m_2 = n^{2/3}$ and $m_3 = n$. From Figure 7.7, we notice that the choices $m = m_1, m_2$ works while $m = m_3$ fails to work. In conclusion, the assumption $m \log n/n \to 0$, as $n \to \infty$, is required and our simulation result also confirms this necessary condition.

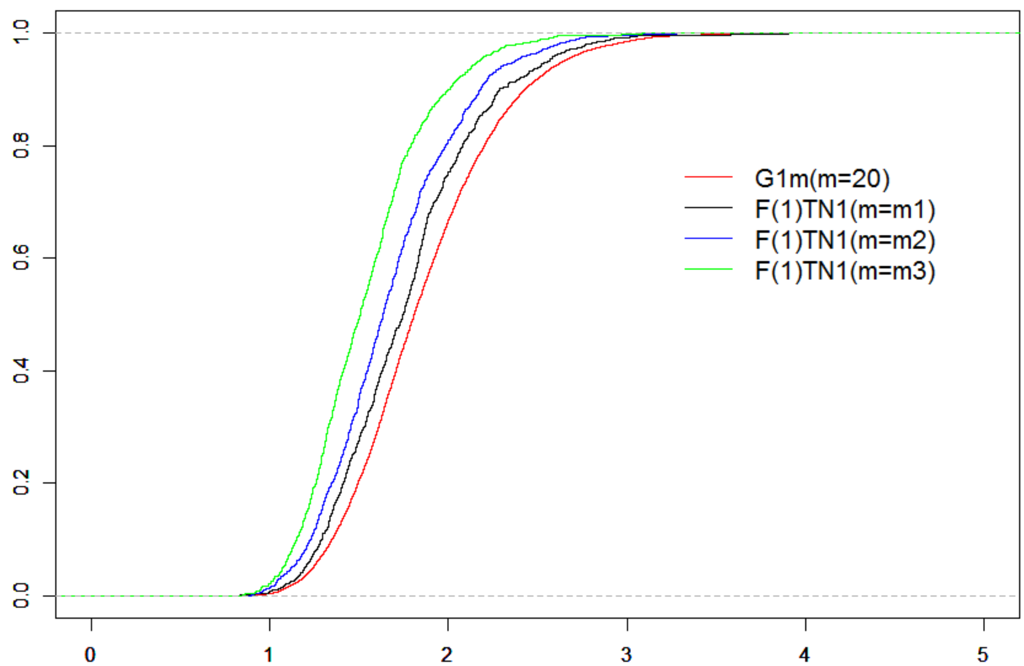

Figure 7.7: The comparison of $F^{(1)}_{T_N,1}$ with $G_{1,20}$ for $n = 400$, $m = m_1, m_2, m_3$ and $N = 1000$.

7.3 Bootstrapping the Hausdorff distance on uniform polytopes with smooth boundary

For the smooth cases in Theorem 5.2.1 and Theorem 5.2.2, we consider the uniform distribution F on a unit disk. The limiting distribution G (in Step 6) is G_2 (see Theorem 1.4.7), where

$$G_2(x) = \exp\{-d_1 e^{-d_2 x}\}$$

with $d_1 = (\pi/2)^{2/3}$ and $d_2 = (2/\pi)^{2/3}$. The normalized target functionals are

$$T_{n,m} = \frac{H(K, K^*_{n,m}) - b_m}{a_m}, \quad T^{(1)}_{n,m} = \frac{H(K_n, K^*_{n,m}) - b_m}{a_m}$$

as in Theorem 5.2.1 and Theorem 5.2.2 respectively. In our case, a_m and b_m should be

$$a_m = m^{-2/3}(\log m)^{-1/3}, \quad b_m = m^{-2/3}(c_m \log m)^{2/3},$$

with

$$c_m = 2^{-5/2}\pi\left(1 + \frac{2\log\log m}{m}\right).$$

As it was discussed, we need to have

$$m/n^{2/3} \to 0, \text{ as } n \to \infty,$$

Notice that in the proof of Theorem 5.2.1, we have

$$|P(f_m(X^*_{n,1}, \dots, X^*_{n,m}) \le x |\, S_n) - P(f_m(X'_1, \dots, X'_m) \le x)| \xrightarrow{P} 0, \quad \text{as } m \to \infty, \tag{7.3.1}$$

where

$$P(f_m(X^*_{n,1},\ldots,X^*_{n,m})\leq x|\ S_n) = P\left(\frac{H(K,K^*_{n,m})-b_m}{a_m}\leq x\middle|\ S_n\right) \tag{7.3.2}$$

and

$$P(f_m(X_1^{'},\ldots,X_m^{'})\leq x) = G_{2,m}(x) \tag{7.3.3}$$

with

$$\lim_{m\to\infty} G_{2,m}(x) = G_2(x), \tag{7.3.4}$$

where x is any continuity point of G_2. For a similar reason as it was discussed in the previous simulations, we can replace G_2 with $G_{2,m}$ in our simulations.

7.3.1 The bootstrap simulations to verify Theorem 5.2.1

In Theorem 5.2.1, the normalized target functional

$$T_{n,m} = \frac{H(K,K^*_{n,m})-b_m}{a_m}.$$

Take $N = 1000$ in Step 4 of our algorithm. In this section, the empirical distribution F_{T_N} (in Step 5) is denoted by $F_{T_N,2}$. We still take $n = 400$ and $m = 20$. We replace G_2 with $G_{2,20}$ and repeat our simulations 6 times (see Figure 7.8). As it can be seen that $F_{T_N,2}$ and $G_{2,20}$ are very close which meets our expectations. To verify our assumption about m, we examine $m_1 = n^{1/3}$, $m_2 = \sqrt{n}$ and $m_3 = n$. From Figure 7.9, it can be seen that $F_{T_N,2}$ and $G_{2,20}$ are close when $m = m_1, m_2$, and very far when $m = m_3$. Since only m_3 does not satisfy our assumption, this simulation result also supports our theory.

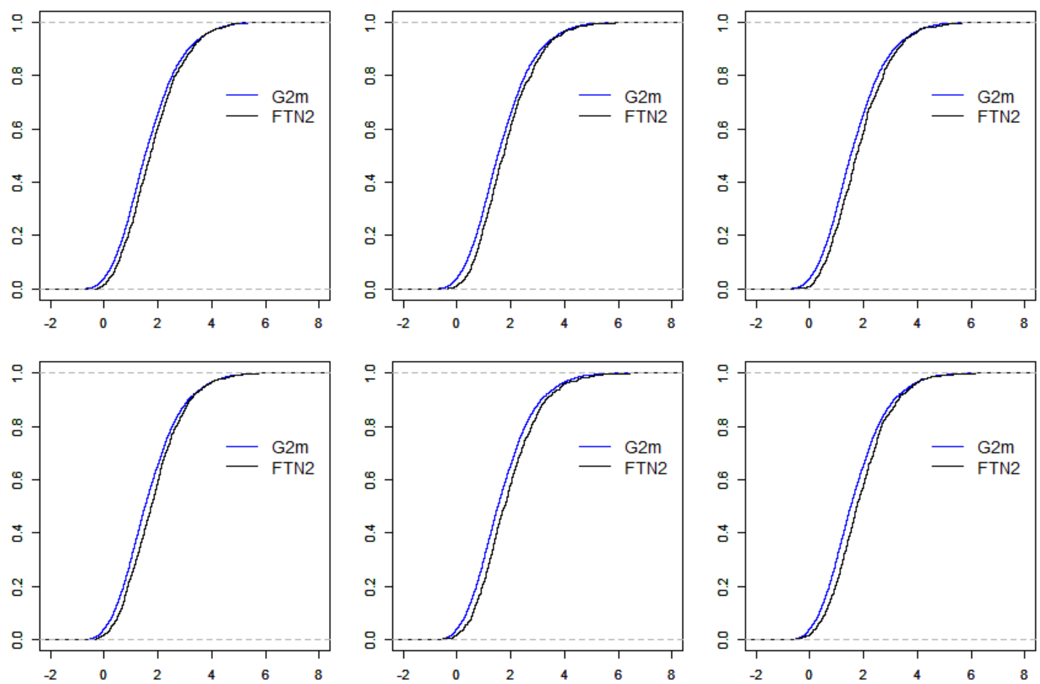

Figure 7.8: The comparisons of $G_{2,20}$ with $F_{T_N,2}$ for $n = 400$, $m = 20$ and $N = 1000$.

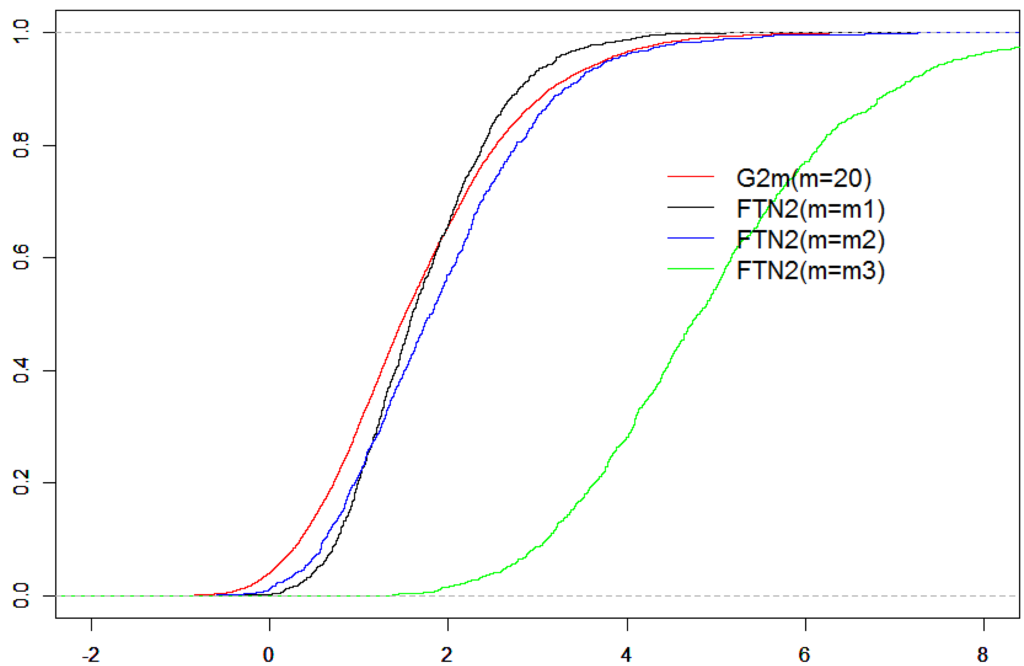

Figure 7.9: The comparison of $G_{2,20}$ with $F_{T_N,2}$ for $n = 400$, $m = m_1, m_2, m_3$ and $N = 1000$.

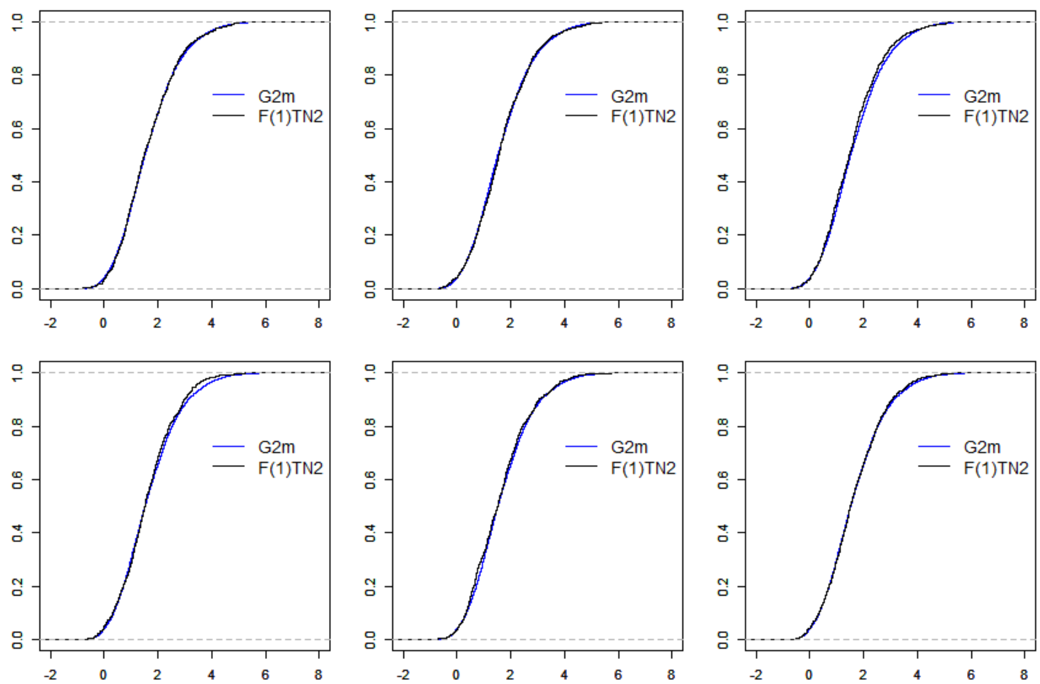

Figure 7.10: The comparisons of $G_{2,20}$ with $F^{(1)}_{T_N,2}$ for $n = 400$, $m = 20$ and $N = 1000$.

7.3.2 The bootstrap simulations to verify Theorem 5.2.2

In Theorem 5.2.2, let the normalized target functional be

$$T^{(1)}_{n,m} = \frac{H(K_n, K^*_{n,m}) - b_m}{a_m}.$$

Take $N = 1000$ in Step 4. In this section, the empirical distribution F_{T_N} (in Step 5) is denoted by $F^{(1)}_{T_N,2}$. We also replace G_2 with $G_{2,20}$ and repeat our simulation 6 times (see Figure 7.10). $F^{(1)}_{T_N,2}$ and $G_{2,20}$ are very close. Again, the simulation results support our theory. To verify our assumption about m, we still examine $m_1 = n^{1/3}$, $m_2 = \sqrt{n}$ and $m_3 = n$. From Figure 7.11, it can be seen that $F^{(1)}_{T_N,2}$ and $G_{2,20}$ are relatively close when $m = m_1, m_2$ and very far when $m = m_3$. This result confirms the fact that $m = m_3$ does not satisfy our assumption.

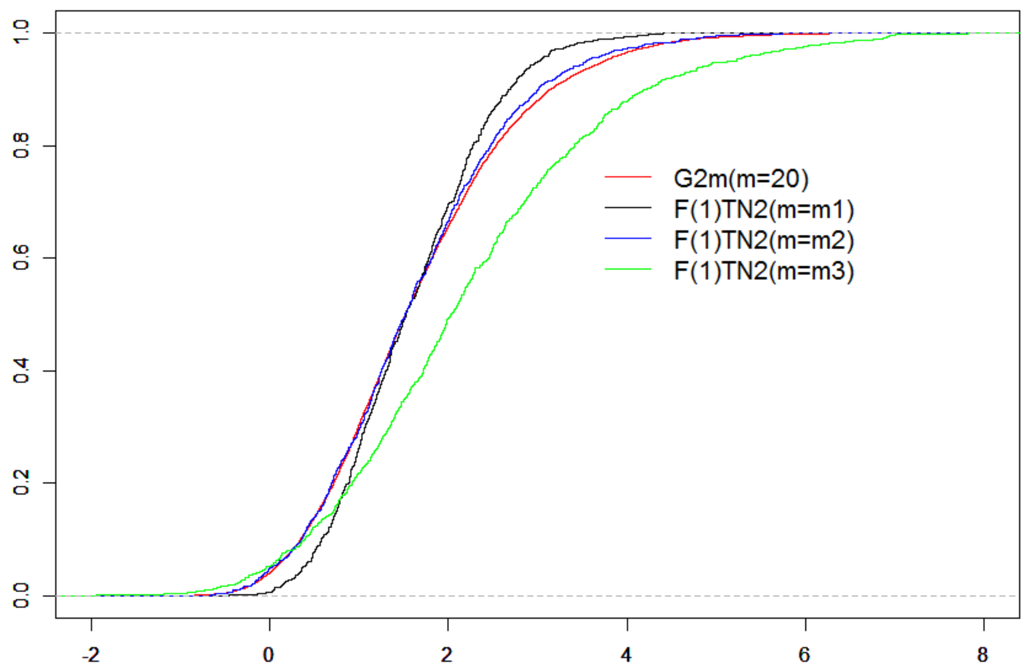

Figure 7.11: The comparison of $G_{2,20}$ with $F^{(1)}_{T_N,2}$ for $n = 400$, $m = m_1, m_2, m_3$ and $N = 1000$.

Chapter 8

Future work

8.1 Assumption about the resample size

As it was discussed in Chapter 6, our assumption about the resample size m is from the following convergence result,

$$m(1 - F(K_n)) \xrightarrow{P} 0, \quad \text{as } n, m \to \infty, \tag{8.1.1}$$

This assumption has nothing to do with the type of functionals. For example, in Chapter 6, we directly give $m(n) = o\left(\sqrt{n}(\log n)^{-(d-1)/4}\right)$ from our conjecture for Gaussian polytopes regardless of the functional. (See Conjecture (6.2.1).) However we are not always certain whether the convergence (8.1.1) is a necessary and sufficient condition for the success of our bootstrapping scheme (see Section 6.1). This means that the condition (8.1.1) may be too strong. Indeed, considering relevant work in the regular bootstrap, the assumption $m = o(n)$ is sufficient mostly. This will be addressed in our future research.

8.2 Generalization for the underlying distribution

Canal [14] introduces three different types of distribution with exponential tails, algebraic tails and truncated tails as follow.

Definition 8.2.1. *A function $S(x)$ is slowly varying as $x \to \infty$ if for all positive λ,*

$$\lim_{x\to\infty} \frac{S(\lambda x)}{S(x)} = 1.$$

Definition 8.2.2. *Suppose S is slowly varying. For $\boldsymbol{x} \in \mathbb{R}^d$ ($d \in \mathbb{Z}^+$), A distribution F has a truncated tail with order $k > 0$ if $F(1 - \boldsymbol{x}) = \Theta(\boldsymbol{x}^k)$. F has an exponential tail if $\boldsymbol{x} = S(1/F(\boldsymbol{x}))$. F has an algebraic tail with order $k > 0$ if $F(\boldsymbol{x}) = \boldsymbol{x}^{-k}S(x)$.*

All the distributions appeared in this thesis (like uniform distribution, Gaussian distribution and regularly varying distribution) belong to these types of distribution. We would like to generalize our bootstrapping scheme to these distributions.

8.3 Type *II* problem

For type *II* problem (see Section 6.1), i.e.

$$\frac{\phi(X^*_{n,1}, \ldots, X^*_{n,m}; F_{K_n}) - \beta_m}{\alpha_m} \xrightarrow{D} Z, \quad \text{as } n, m \to \infty, \text{ in probability,}$$

we have discussed the uniform polytopes when the functionals are the Hausdorff distance (Theorem 5.1.2 and Theorem 5.2.2), the symmetric difference (Example 5.3.1) and the perimeter difference (Example 5.3.2). A common point of these functionals, say ϕ, is that all of them satisfy the following inequalities

$$\phi(K, K^*_{n,m}) \geq \phi(K_n, K^*_{n,m})$$

and

$$\phi(K_n, K_m^{'}) \geq \phi(K, K_m^{'}) - \phi(K, K_n),$$

which are indispensable properties for our techniques. However sometimes, these two conditions are not necessary for our bootstrapping scheme. For example, the functional h defined in Example 4.2.2 does not satisfy the condition:

$$h(K_n, K_m^{'}) \geq h(K, K_m^{'}) - h(K, K_n)$$

(see Figure 8.1), where $h(K, K_m^{'}) > 0$ and $h(K, K_n) = h(K_n, K_m^{'}) = 0$. We can still prove the type *II* convergence for h in Example 4.2.2. Therefore, we would like to improve our techniques to deal with more functionals about the type *II* convergence in our future research.

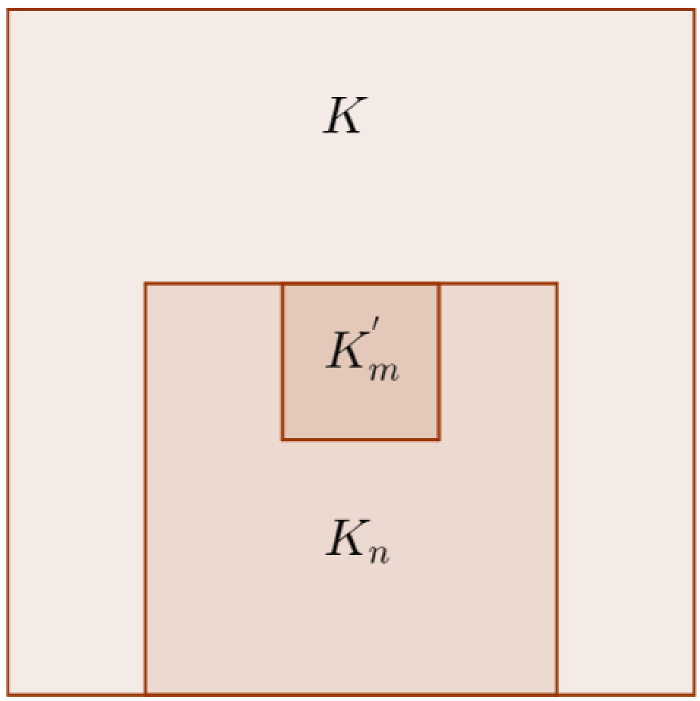

Figure 8.1: Rectangles K, K_n and $K_m^{'}$.

Appendix A

More concepts

Let $(\Omega, \mathcal{F}, P)$ be our probability space. Let $\mathbb{R}^d$ $(d \in \mathbb{Z}^+)$ be the d-dimensional Euclidean space with the Borel σ-field $\mathcal{R}^d$.

A.1 Radon measure

The measure m is called a Radon measure if m is a measure on the Borel σ-field of a separate space, which is inner regular and locally finite.

A.2 Point measure and point process

Suppose the space S is locally compact with countable basis. Denote the Borel σ-field of S by $\mathfrak{S}$. Define the measure ε_x on S by

$$\varepsilon_x(A) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

for any $A \in \mathfrak{S}$. let $\{x_i\}_{i=1}^{\infty}$ be a countable collection of points in S. A measure m is call a point measure on S if

$$m = \sum_{i=1}^{\infty} \varepsilon_{x_i}$$

and $m(K)$ is finite if $K \in \mathfrak{S}$ is compact.

Let $M_p(S)$ be the collection of all non-negative point measures on S. Let $\mathcal{M}_p(S)$ be the smallest σ-field making maps $m \to m(A)$ measurable for any $m \in M_p(S)$, $A \in \mathfrak{S}$. A point process on S is a measurable map from $(\Omega, \mathcal{F})$ to $(M_p(S), \mathcal{M}_p(S))$. In other words, A point process on S is a random element taking values in $M_p(S)$.

A.3 Intensity measure

The measure μ is called the intensity measure of the point process ξ if

$$\mu(A) = \mathrm{E}\,\xi(A), \quad \text{for any } A \in \mathfrak{S}.$$

A.4 Laplace functional

Designate by $M_p(S)$ the collection of all point processes on S. Suppose $\xi \in M_p(S)$ and $f : (S, \mathfrak{S}) \to (\mathbb{R}, \mathcal{R})$ is measurable. Define $\xi(f)$ by

$$\xi(f) := \int_S f d\xi.$$

For $\xi \in M_p(S)$ the Laplace functional is given by:

$$\Psi_\xi(f) = \mathrm{E} \exp\{-\xi(f)\}.$$

Proposition A.4.1. *The Laplace functional Ψ_ξ uniquely determines the law of the point process ξ.*

Proposition A.4.2. (cf. Resnick [36]) *Let $C_k^+(S)$ be the collection of all continuous non-negative functions on S with compact support. Suppose $\xi_n,\ \xi \in M_p(S)$. The following are equivalent:*

(i) $\xi_n \xrightarrow{D} \xi$, as $n \to \infty$.

(ii) $\Psi_{\xi_n}(f) \to \Psi_\xi(f)$, as $n \to \infty$, for all $f \in C_k^+(S)$.

A.5 Poisson point process

$\xi \in M_p(S)$ is called a Poisson point process with intensity μ if for any $A \in \mathfrak{S}$ and $k \geq 0$,

$$P(\xi(A) = k) = \begin{cases} \frac{\exp\{\mu(A)\}(\mu(A))^k}{k!}, & \text{if } \mu(A) < \infty \\ 0, & \text{if } \mu(A) = \infty \end{cases}$$

and for any $n \geq 1$, $\{\xi(A_i)\}_{i=1}^n$ are independent random variables when $A_1, A_2, \dots, A_n \in \mathfrak{S}$ are mutually disjoint.

Proposition A.5.1. *For any $f \in C_k^+(S)$, the Laplace functional of a Poisson point process ξ is given by:*

$$\Phi_\xi(f) = \exp\left\{ - \int_S (1 - \exp\{-f(x)\})\mu(dx) \right\}. \tag{A.5.1}$$

Conversely if the Laplace functional of ξ has the form of (A.5.1), then ξ is a Poisson process with intensity μ.

A.6 Vague convergence

Resnick [36] defines the vague convergence as follows. For μ_n, $\mu \in M_p(S)$, we say μ_n converges vaguely to μ (written as $\mu_n \xrightarrow{V} \mu$) if

$$\mu_n(f) \to \mu(f), \quad \text{as } n \to \infty,$$

for any $f \in C_k^+(S)$.

Proposition A.6.1. *Suppose $\mu_n,\ \mu \in M_p(S)$. The following are equivalent:*

(i) $\mu_n \xrightarrow{V} \mu$;

(ii) $\mu_n(A) \to \mu(A)$ *for all relatively compact* $A \in \mathfrak{S}$ *for which* $\mu(\partial A) = 0$.

A.7 Curvature and Gaussian curvature

Let the planar curve $c(t) = (x(t), y(t))$ where $x,\ y \in \mathcal{C}^2(\mathbb{R})$ and t is the arc length between $c(0)$ and $c(t)$. Then the curvature $\kappa(t) = |\ddot{c}(t)|$.

Let S be a surface in Euclidean space with second fundamental form II (see Kobayashi and Nomizu [28]). Given a point z on S, suppose $\{v_1, v_2\}$ is an orthonormal basis of tangent vectors at z. Suppose κ_1 and κ_2 are the eigenvalues of the following matrix

$$\begin{bmatrix} II(v_1, v_1) & II(v_1, v_2) \\ II(v_2, v_1) & II(v_2, v_2) \end{bmatrix}.$$

Then the Gaussian curvature at the point z is given by $\kappa = \kappa_1 \kappa_2$.

A.8 Minkowski sum and difference

Suppose $A,\ B \subset \mathbb{R}^d$. Then the Minkowski sum and difference of A and B is defined as:

$$A + B = \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in A,\ \mathbf{b} \in B\}$$

and

$$A - B = \{\mathbf{c} : \mathbf{c} + B \subset A\}.$$

Note that $A - B$ is not generally equal to $A + (-B)$. If $B = \{\mathbf{b}\}$, then $A - B = A + (-B)$.

Appendix B

Programs

```
library(LaplacesDemon)

#==================== Functions ====================
#the function for calculating the convex hull of a sample
ConvexHull <- function(XX){
  #order by x axis for delete
  XX=XX[,order(XX[1,])]

  #delete repeated points
  Y=XX[,1]
  NN=dim(XX)[2]
  for(i in 1:(NN-1)){
    if(XX[1,i+1]!=XX[1,i]|XX[2,i+1]!=XX[2,i]) {Y=c(Y,XX[,i+1])}
  }
  Y=matrix(Y,nrow=2)
  Nd=dim(Y)[2]
```

```
#find p0, the lower left point
p0=Y[,1]
j=1
Nr=Nd-1  #number of the rest points
for(i in 1:Nr){
  if(Y[2,i+1]<p0[2]) {
    p0=Y[,i+1]
    j=i+1
  }
  if(Y[2,i+1]==p0[2]) {
    if(Y[1,i+1]<p0[1]){
      p0=Y[,i+1]
      j=i+1
    }
  }
}
Yr=Y[,-j] #the rest points in Y except p0


#add one more row as angle vector, sorted by angle
a=matrix(0,1,Nr)
Yr=rbind(Yr,a)


#the ordered angles of the rest points Yr
for(i in 1:Nr){
```

```
  Yr[3,i]=acos((Yr[1,i]-p0[1])/sqrt((Yr[2,i]-p0[2])^2+
                                    (Yr[1,i]-p0[1])^2))
}
Oang=order(Yr[3,])
Yro=matrix(0,3,Nr)
for(i in 1:Nr){
  Yro[,i]=Yr[,Oang[Nd-i]]
}


#delete the repeated angles
#only leave the one with the longest distance to p0
Nrr=Nr
Yrro=Yro
i=2
while(i<=Nrr){
  if(Yrro[3,i]==Yrro[3,i-1]){
    if(((Yrro[1,i]-p0[1])^2+(Yrro[2,i]-p0[2])^2)<=
       ((Yrro[1,i-1]-p0[1])^2+(Yrro[2,i-1]-p0[2])^2)){
      Yrro=Yrro[,-i]
    }else{
      Yrro=Yrro[,-(i-1)]
    }
    Nrr=Nrr-1
  }else{
    i=i+1
  }
}
```

```
#output
if(Nrr<3)
{print("cannot find the convex hull!")
}else{
  S=cbind(c(p0,0),Yrro[,1])  #output group
  P=Yrro #possible vertex
  i=2
  L=2
  while(i<=Nrr){
    S=cbind(S,P[,i])
    L=L+1
    if(((S[1,L-1]-S[1,L-2])*(S[2,L]-S[2,L-1])-
        (S[2,L-1]-S[2,L-2])*(S[1,L]-S[1,L-1]))<0)
    {
      i=i+1
    }else{
      P=P[,-(i-1)]
      S=S[,-L]
      S=S[,-(L-1)]
      L=L-2
      Nrr=Nrr-1
      i=i-1
    }
  }
}
return(S)
```

```
}


#the function for calculating the area of a polygon
#the data of vertex should be clockwise
Area <- function(XX){
  n=dim(XX)[2]
  S=0
  for(i in 3:n){
    a=sqrt((XX[1,i-1]-XX[1,1])^2+(XX[2,i-1]-XX[2,1])^2)
    b=sqrt((XX[1,i]-XX[1,1])^2+(XX[2,i]-XX[2,1])^2)
    c=sqrt((XX[1,i]-XX[1,i-1])^2+(XX[2,i]-XX[2,i-1])^2)
    r=(a+b+c)/2
    A=sqrt(r*(r-a)*(r-b)*(r-c))
    S=S+A
  }
  return(S)
}


#the function for calculating the Hausdorff distance H(K,Kn)
#when K is a convex polygon
HD <- function(O,I){
  n=dim(O)[2]
  m=dim(I)[2]
  sd=matrix(0,1,n)
  for(i in 1:n){
    aa=matrix(0,1,m)
```

```
    for(j in 1:m){
      aa[j]<-sqrt((O[1,i]-I[1,j])^2+(O[2,i]-I[2,j])^2)
    }
    os<-order(aa)[1]
    sd[i]<-aa[os]
    if(sd[i]!=0){
      if(os>=2){osm<-os-1
      }else{osm<-m}
      if(os<=m-1){osp<-os+1
      }else{osp<-1}
      a1<-c(I[1,osm]-I[1,os],I[2,osm]-I[2,os])
      a2<-c(O[1,i]-I[1,os],O[2,i]-I[2,os])
      a3<-c(I[1,osp]-I[1,os],I[2,osp]-I[2,os])
      if(a1%*%a2>0){sd[i]<-sqrt(a2%*%a2-((a1%*%a2)/sqrt(a1%*%a1))^2)}
      if(a3%*%a2>0){sd[i]<-sqrt(a2%*%a2-((a3%*%a2)/sqrt(a3%*%a3))^2)}
    }
  }
  return(sd[order(sd)[n]])
}


#the function of approximate Hausdorff distance H(K,Kn)
#when K is a unit disk
HDC <- function(XX){
  X=matrix(0,2,1000)
  for(i in 1:1000){
    X[1,i]=cos(-i*pi/500)
    X[2,i]=sin(-i*pi/500)
```

```
  }
  HD(X,XX)
}
```

```
#the approximate distribution function for Hausdorff distance
#when K is a convex polygon
ADHP <- function(x,XX){
  k=Area(XX)
  r=dim(XX)[2]
  thelta=XX[1,]
  YY=cbind(XX,XX[,1])
  for(i in 2:r){
    b1=YY[1:2,(i-1)]
    b2=YY[1:2,i]
    b3=YY[1:2,(i+1)]
    thelta[i]=acos((b3-b2)%*%(b1-b2)/sqrt((b3-b2)%*%(b3-b2))/
                     sqrt((b1-b2)%*%(b1-b2)))
  }
  b1=XX[1:2,r]
  b2=XX[1:2,1]
  b3=XX[1:2,2]
  thelta[1]=acos((b3-b2)%*%(b1-b2)/sqrt((b3-b2)%*%(b3-b2))/
                   sqrt((b1-b2)%*%(b1-b2)))
  p=matrix(0,1,r)
  for(i in 1:r){
    hf <- function(th){
```

```
      y=exp(-x^2/(2*k)*(tan(th)+tan(thelta[i]-th)))*x^2/(2*k)*
        (tan(th))^2
      return(y)
    }
    if(thelta[i]>0&&thelta[i]<pi/2){
      p[i]=integrate(hf,lower=0,upper=thelta[i])[[1]]+
        exp(-x^2/(2*k)*tan(thelta[i]))
    }else{
      p[i]=integrate(hf,lower=thelta[i]-pi/2,upper=pi/2)[[1]]
    }
  }
  y=1
  for(j in 1:r){
    y=y*(1-p[j])
  }
  return(y)
}


#the function for drawing n points uniformly from the unit disk
sample <- function(T){
  X=matrix(0,2,1)
  a=NULL
  while(dim(X)[2]<=T){
    x1=runif(1,-1,1)
    x2=runif(1,-1,1)
    if(x1^2+x2^2<=1){
      a=c(x1,x2)
```

```
      X=cbind(X,a)
    }
  }
  X=X[,-1]
  return(X)
}


#the function for drawing n points uniformly from a polygon
#the polygon is a subset of the unit disk
samplepolygon <- function(XX,T){
  r=dim(XX)[2]
  X=matrix(0,2,1)
  while(dim(X)[2]<=T){
    j=0
    a0=sample(1)
    a1=XX[,r]
    a2=XX[,1]
    if((a2[1]-a0[1])*(a1[2]-a0[2])>(a1[1]-a0[1])*(a2[2]-a0[2])){
      for(i in 1:(r-1)){
        a1=XX[,i]
        a2=XX[,(i+1)]
        if((a2[1]-a0[1])*(a1[2]-a0[2])<=(a1[1]-a0[1])*(a2[2]-a0[2])){
          break
        }else{j=j+1}
      }
    }
    if(j==r-1){X=cbind(X,a0)}
```

```
  }
  X=X[,-1]
  return(X)
}


#==================== Figures ====================
#Figure 1.2 K and Kn, n=100
par(mfrow=c(1,1))
K=ConvexHull(sample(50))
KK=cbind(K,K[,1])
n=100
plot(KK[1,],KK[2,],xlim=c(-1,1),ylim=c(-1,1),xlab="",ylab="")
lines(KK[1,],KK[2,])
lines(KK[1,dim(KK)[2]],KK[1,1])
par(new=T)
ss=samplepolygon(K,n)
Kn=ConvexHull(ss)
Kn=cbind(Kn,Kn[,1])
plot(ss[1,],ss[2,],xlim=c(-1,1),ylim=c(-1,1),xlab="",ylab="")
lines(Kn[1,],Kn[2,])
lines(Kn[1,dim(Kn)[2]],Kn[1,1])


#Figure 1.3 the convex hull of 1000 points
#from bivariate cauchy with independent components
s=1000
b=matrix(0,2,s)
```

```
b[1,]=rcauchy(s)
b[2,]=rcauchy(s)
plot(b[1,],b[2,],xlab="X",ylab="Y")
par(new=T)
a=ConvexHull(b)
aa=cbind(a,a[,1])
plot(aa[1,],aa[2,],xlab="X",ylab="Y")
lines(aa[1,],aa[2,])
lines(aa[1,dim(aa)[2]],aa[1,1])


#Figure 1.4 the convex hull of 1000 points
#from a common bivariate cauchy
#need the package "LaplacesDemon"
b <- rmvc(1000, rep(0,2), diag(2))
b=t(b)
plot(b[1,],b[2,],xlab="X",ylab="Y")
par(new=T)
a=ConvexHull(b)
aa=cbind(a,a[,1])
plot(aa[1,],aa[2,],xlab="X",ylab="Y")
lines(aa[1,],aa[2,])
lines(aa[1,dim(aa)[2]],aa[1,1])


#Figure 8.1 underlying convex polygon K
par(pin=c(5,5))
par(mfrow=c(1,1))
```

```
K=ConvexHull(sample(50))
KK=cbind(K,K[,1])
#Figure 8.1 K
plot(KK[1,],KK[2,],xlim=c(-1,1),ylim=c(-1,1),xlab="",ylab="")
lines(KK[1,],KK[2,])
lines(KK[1,dim(KK)[2]],KK[1,1])


#sample size n, resample size m,
n=400
m=floor(sqrt(n))
N=1000


#G1
#cdf
cdf=matrix(0,2,500)
for(i in 1:500){
  cdf[2,i]=ADHP(0.001+i*0.01,K)
  cdf[1,i]=0.001+i*0.01
}


#G1m
prec=10000
G1m=matrix(0,1,prec)
for(i in 1:prec){
  Km=ConvexHull(samplepolygon(K,m))
```

```
  G1m[i]=sqrt(m)*HD(K,Km)
}


#FTN1
frequ=matrix(0,1,N)
for(i in 1:N){
  Kn=ConvexHull(samplepolygon(K,n))
  Knm=ConvexHull(samplepolygon(Kn,m))
  frequ[i]=sqrt(m)*HD(K,Knm)
}


#Figure 8.2  G1 FTN1
par(mar=c(5,5,5,5))
par(mfrow=c(1,1))
plot(cdf[1,],cdf[2,],xlim=c(0,5.0),ylim=c(0,1),type="l",col="red",
     xlab="",ylab="",main="")
par(new=T)
plot(ecdf(frequ),col="black",xlim=c(0,5.0),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="")
legend(3.4,0.8,cex=1.5,lty=1,col=c("red","black"),
       legend=c("G1","FTN1"),bty="n")


#Figure 8.3 G1m FTN1
par(mar=c(2,2,2,2))
par(mfrow=c(2,3))
```

```
for(j in 1:6){
  frequ=matrix(0,1,N)
  for(i in 1:N){
    Kn=ConvexHull(samplepolygon(K,n))
    Knm=ConvexHull(samplepolygon(Kn,m))
    frequ[i]=sqrt(m)*HD(K,Knm)
  }
  plot(ecdf(G1m),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,col="blue",
       xlab="",ylab="",main="")
  par(new=T)
  plot(ecdf(frequ),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,col="black",
       xlab="",ylab="",main="")
  legend(2.5,0.8,cex=1.3,lty=1,col=c("blue","black"),
         legend=c("G1m","FTN1"),bty="n")
}


#Figure 8.4 G1 G1m
par(mfrow=c(1,1))
plot(cdf[1,],cdf[2,],xlim=c(0,5.0),ylim=c(0,1),type="l",col="red",
     xlab="",ylab="",main="")
par(new=T)
plot(ecdf(G1m),xlim=c(0,5.0),ylim=c(0,1),col="blue",cex=0.1,
     xlab="",ylab="",main="")
legend(3.4,0.8,cex=1.3,lty=1,col=c("red","blue"),
       legend=c("G1","G1m"),bty="n")
```

```
#Figure 8.5 G1m FTN1 m=m1,m2,m3
m1=floor(sqrt(n))
m2=floor(n^(2/3))
m3=n
frequ1=matrix(0,1,N)
frequ2=matrix(0,1,N)
frequ3=matrix(0,1,N)
par(mfrow=c(1,1))
for(i in 1:N){
  Kn=ConvexHull(samplepolygon(K,n))
  Knm=ConvexHull(samplepolygon(Kn,m1))
  frequ1[i]=sqrt(m1)*HD(K,Knm)
}
for(i in 1:N){
  Kn=ConvexHull(samplepolygon(K,n))
  Knm=ConvexHull(samplepolygon(Kn,m2))
  frequ2[i]=sqrt(m2)*HD(K,Knm)
}
for(i in 1:N){
  Kn=ConvexHull(samplepolygon(K,n))
  Knm=ConvexHull(samplepolygon(Kn,m3))
  frequ3[i]=sqrt(m3)*HD(K,Knm)
}
plot(ecdf(G1m),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,col="red",
     xlab="",ylab="",main="")
par(new=T)
plot(ecdf(frequ1),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="black")
```

```
par(new=T)
plot(ecdf(frequ2),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="blue")
par(new=T)
plot(ecdf(frequ3),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="green")
legend(3.2,0.8,cex=1.3,lty=1,col=c("red","black","blue","green"),
       legend=c("G1m(m=20)","FTN1(m=m1)","FTN1(m=m2)","FTN1(m=m3)"),
       bty="n")


#Figure 8.6 G1m F(1)TN1
par(mar=c(2,2,2,2))
par(mfrow=c(2,3))
for(j in 1:6){
  frequ=matrix(0,1,N)
  for(i in 1:N){
    Kn=ConvexHull(samplepolygon(K,n))
    Knm=ConvexHull(samplepolygon(Kn,m))
    frequ[i]=sqrt(m)*HD(Kn,Knm)
  }
  plot(ecdf(G1m),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,col="blue",
       xlab="",ylab="",main="")
  par(new=T)
  plot(ecdf(frequ),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,col="black",
       xlab="",ylab="",main="")
  legend(2.5,0.8,cex=1.3,lty=1,col=c("blue","black"),
         legend=c("G1m","F(1)TN1"),bty="n")
```

```
}

#Figure 8.7 G1m F(1)TN1 m=m1,m2,m3
m1=floor(sqrt(n))
m2=floor(n^(2/3))
m3=n
frequ1=matrix(0,1,N)
frequ2=matrix(0,1,N)
frequ3=matrix(0,1,N)
par(mfrow=c(1,1))
for(i in 1:N){
  Kn=ConvexHull(samplepolygon(K,n))
  Knm=ConvexHull(samplepolygon(Kn,m1))
  frequ1[i]=sqrt(m1)*HD(Kn,Knm)
}
for(i in 1:N){
  Kn=ConvexHull(samplepolygon(K,n))
  Knm=ConvexHull(samplepolygon(Kn,m2))
  frequ2[i]=sqrt(m2)*HD(Kn,Knm)
}
for(i in 1:N){
  Kn=ConvexHull(samplepolygon(K,n))
  Knm=ConvexHull(samplepolygon(Kn,m3))
  frequ3[i]=sqrt(m3)*HD(Kn,Knm)
}
plot(ecdf(G1m),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,col="red",
     xlab="",ylab="",main="")
```

```
par(new=T)
plot(ecdf(frequ1),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="black")
par(new=T)
plot(ecdf(frequ2),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="blue")
par(new=T)
plot(ecdf(frequ3),xlim=c(0,5.0),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="green")
legend(3.2,0.8,cex=1.3,lty=1,col=c("red","black","blue","green"),
       legend=c("G1m(m=20)","F(1)TN1(m=m1)","F(1)TN1(m=m2)",
                "F(1)TN1(m=m3)"),bty="n")


#normalization coefficients, an and bn
a <- function(n){
  an=((log(n))^(-1/3))*(n^(-2/3))
  return(an)
}


b <- function(n){
  gamma=1/3
  gamman=gamma+(1-gamma)*log(log(n))/log(n)
  bn=0.5*((pi/2)^(2/3))*((gamman/gamma)^(2/3))*((log(n))^(2/3))*
    ((n)^(-2/3))
  return(bn)
}
```

```
#G2m
prec=10000
G2m=matrix(0,1,prec)
for(i in 1:prec){
  Km=ConvexHull(sample(m))
  G2m[i]=(HDC(Km)-b(m))/a(m)
}


#Figure 8.8 G2m FTN2
par(mfrow=c(2,3))
for(i in 1:6){
  frequ=matrix(0,1,N)
  for(i in 1:N){
    Kn=ConvexHull(sample(n))
    Knm=ConvexHull(samplepolygon(Kn,m))
    frequ[i]=(HDC(Knm)-b(m))/a(m)
  }
  plot(ecdf(G2m),xlim=c(-2,8),ylim=c(0,1),cex=0.1,col="blue",
       xlab="",ylab="",main="")
  par(new=T)
  plot(ecdf(frequ),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
       xlab="",ylab="",main="",col="black")
  legend(4.0,0.8,cex=1.3,lty=1,col=c("blue","black"),
         legend=c("G2m","FTN2"),bty="n")
}
```

```
#Figure 8.9 G2m FTN2 m=m1,m2,m3
m1=floor(n^(1/3))
m2=floor(sqrt(n))
m3=n
par(mfrow=c(1,1))
for(i in 1:N){
  Kn=ConvexHull(sample(n))
  Knm=ConvexHull(samplepolygon(Kn,m1))
  frequ1[i]=(HDC(Knm)-b(m1))/a(m1)
}
for(i in 1:N){
  Kn=ConvexHull(sample(n))
  Knm=ConvexHull(samplepolygon(Kn,m2))
  frequ2[i]=(HDC(Knm)-b(m2))/a(m2)
}
for(i in 1:N){
  Kn=ConvexHull(sample(n))
  Knm=ConvexHull(samplepolygon(Kn,m3))
  frequ3[i]=(HDC(Knm)-b(m3))/a(m3)
}
plot(ecdf(G2m),xlim=c(-2,8),ylim=c(0,1),cex=0.1,col="red",
     xlab="",ylab="",main="")
par(new=T)
plot(ecdf(frequ1),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="black")
par(new=T)
```

```
plot(ecdf(frequ2),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="blue")
par(new=T)
plot(ecdf(frequ3),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="green")
legend(4.2,0.8,cex=1.3,lty=1,col=c("red","black","blue","green"),
       legend=c("G2m(m=20)","FTN2(m=m1)","FTN2(m=m2)","FTN2(m=m3)"),
       bty="n")


#Figure 8.10 G2m  F(1)TN2
par(mfrow=c(2,3))
for(i in 1:6){
  frequ=matrix(0,1,N)
  for(i in 1:N){
    Kn=ConvexHull(sample(n))
    Knm=ConvexHull(samplepolygon(Kn,m))
    frequ[i]=(HD(Kn,Knm)-b(m))/a(m)
  }
  plot(ecdf(G2m),xlim=c(-2,8),ylim=c(0,1),cex=0.1,col="blue",
       xlab="",ylab="",main="")
  par(new=T)
  plot(ecdf(frequ),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
       xlab="",ylab="",main="",col="black")
  legend(3.5,0.8,cex=1.3,lty=1,col=c("blue","black"),
         legend=c("G2m","F(1)TN2"),bty="n")
}
```

```
#Figure 8.11 G2m F(1)TN2 m=m1,m2,m3
m1=floor(n^(1/3))
m2=floor(sqrt(n))
m3=n
par(mfrow=c(1,1))
for(i in 1:N){
  Kn=ConvexHull(sample(n))
  Knm=ConvexHull(samplepolygon(Kn,m1))
  frequ1[i]=(HD(Kn,Knm)-b(m1))/a(m1)
}
for(i in 1:N){
  Kn=ConvexHull(sample(n))
  Knm=ConvexHull(samplepolygon(Kn,m2))
  frequ2[i]=(HD(Kn,Knm)-b(m2))/a(m2)
}
for(i in 1:N){
  Kn=ConvexHull(sample(n))
  Knm=ConvexHull(samplepolygon(Kn,m3))
  frequ3[i]=(HD(Kn,Knm)-b(m3))/a(m3)
}
plot(ecdf(G2m),xlim=c(-2,8),ylim=c(0,1),cex=0.1,col="red",
     xlab="",ylab="",main="")
par(new=T)
plot(ecdf(frequ1),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="black")
par(new=T)
plot(ecdf(frequ2),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
```

```
     xlab="",ylab="",main="",col="blue")
par(new=T)
plot(ecdf(frequ3),xlim=c(-2,8),ylim=c(0,1),cex=0.1,
     xlab="",ylab="",main="",col="green")
legend(3.5,0.8,cex=1.3,lty=1,col=c("red","black","blue","green"),
       legend=c("G2m(m=20)","F(1)TN2(m=m1)","F(1)TN2(m=m2)",
                "F(1)TN2(m=m3)"),bty="n")
```

Bibliography

[1] M. A. Arcones and E. Giné. The bootstrap of the mean with arbitrary bootstrap sample size. *Annales de l'Institut Henri Poincaré*, 25:457–481, 1989.

[2] K. Athreya. Bootstrap of the mean in the infinite variance case. *The Annals of Statistics*, 15:724–731, 1987.

[3] I. Bárány. Random polytopes in smooth convex bodies. *Mathematika*, 39:81–92, 1992.

[4] I. Bárány and V. Vu. Central limit theorems for Gaussian polytopes. *The Annals of Probability*, 35:1593–1621, 2007.

[5] V. Barnett. The ordering of multivariate data. *Journal of the Royal Statistical Society. Series A (General)*, 139:318–355, 1976.

[6] P. J. Bickel and D. Freedman. Some asymptotic theory for the bootstrap. *The Annals of Statistics*, 9:1196–1217, 1981.

[7] P. Billingsley. *Probability and Measure.* JOHN WILEY & SONS, third edition, 1995.

[8] P. Billingsley. *Convergence of Probability measures.* JOHN WILEY & SONS, second edition, 1999.

[9] H. Bräker and T. Hsing. On the area and perimeter of a random convex hull in a bounded convex set. *Probability Theory and Related Fields*, 111:517–550, 1998.

[10] H. Bräker, T. Hsing, and N. H. Bingham. On the Hausdorff distance between a convex set and an interior random convex hull. *Advances in Applied Probability*, 30:295–316, 1998.

[11] C. Buchata. On the boundary structure of the convex hull of random points. *Advances in Geometry*, 12:179–190, 2012.

[12] C. Buchta. An identity relating moments of functionals of convex hulls. *Discrete and Computational Geometry*, 33:125–142, 2005.

[13] A. J. Cabo and P. Groeneboom. Limit theorems for functionals of convex hulls. *Probability Theory and Related Fields*, 100:31–55, 1994.

[14] H. Carnal. Die konvexe Hülle von n rotationssymmetrisch verteilten Punkten. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 15:168–176, 1970.

[15] R. Davis, E. Mulrow, and S. Resnick. The convex hull of a random sample in $\mathbb{R}^2$. *Communications in Statistics. Stochastic Models*, 3:1–27, 1987.

[16] R. Davis, E. Mulrow, and S. Resnick. Almost sure limit sets of random samples in $\mathbb{R}^d$. *Advances in Applied Probability*, 20:573–599, 1988.

[17] L. Dümbgen and G. Walther. Rates of convergence for random approximations of convex sets. *Advances in Applied Probability*, 28:384–393, 1996.

[18] B. Efron. The convex hull of a random set of points. *Biometrika*, 52:331–343, 1965.

[19] B. Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7:1–26, 1979.

[20] P. Groeneboom. Limit theorems for convex hulls. *Probability Theory and Related Fields*, 79:327–368, 1988.

[21] P. Groeneboom. Convex hull of uniform samples from a convex polygon. *Advances in Applied Probability*, 44:330–342, 2012.

[22] P. Hall. Asymptotic properties of the bootstrap for heavy-tailed distributions. *The Annals of Probability*, 18:1342–1360, 1990.

[23] T. Hsing. On the asymptotic distribution of the area outside a random convex hull in a disk. *The Annals of Applied Probability*, 4:478–493, 1994.

[24] I. Hueter. The convex hull of a normal sample. *Advances in Applied Probability*, 26:855–875, 1994.

[25] D. Hug and M. Reitzner. Gaussian polytopes: variances and limit theorems. *Advances in Applied Probability*, 37:297–320, 2005.

[26] O. Kallenberg. *Random Measure*. Akademie-Verlag, Berlin, 1983.

[27] K. Knight. On the bootstrap of the sample mean in the infinite variance case. *The Annals of Statistics*, 17:1168–1175, 1989.

[28] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. Wiley-Interscience, New edition, 1996.

[29] D. W. MacDonald, F. G. Ball, and N. G. Hough. *The evaluation of home range size and configuration using radio tracking data*. In: A Handbook on Biotelemetry and Radio Tracking; ed: D.W. MacDonald and C.J. Amlaner. Pergamon Press, Oxford, 1980.

[30] M. Mayer and I. Molchanov. Limit theorems for the diameter of a random sample in the unit ball. *Extremes*, 10:129–150, 2007.

[31] A. V. Nagaev and I. M. Khamdamov. Limit theorems for functionals of random convex hulls. *Preprint of Institute of Mathematics, Academy of Sciences of Uzbekistan. Tashkent.*, 1991. In Russian, 50 pages.

[32] H. Raynaud. Sur l'enveloppe convexe des nuages de points aléatoires dans $\mathbb{R}^n$. *Journal of Applied Probability*, 7:35–48, 1970.

[33] M. Reitzner. Central limit theorems for random polytopes. *Probability Theory and Related Fields*, 133:483–507, 2005.

[34] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 2:75–84, 1963.

[35] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten II. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 3:138–147, 1964.

[36] S. I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer, New York, 1987.

[37] B. D. Ripley and J. P. Rasson. Finding the edge of a Poisson forest. *Journal of Applied Probability*, 14:483–491, 1977.

[38] C. Schütt. Random polytopes and affine surface area. *Mathematische Nachrichten*, 170:227–249, 1994.

[39] J. W. Tukey. Mathematics and the picturing of data. *In International Congress of Mathematicians*, 2:523–532, 1974.

[40] M. Zarepour. Bootstapping convex hulls. *Statistics and Probability Letters*, 45:55–63, 1999.