Electron. Commun. Probab. **25** (2020), article no. 54, 1–10.
https://doi.org/10.1214/20-ECP336
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Central limit theorems on compact metric spaces

Steven Rosenberg* Jie Xu†

Abstract

We produce a series of Central Limit Theorems (CLTs) associated to compact metric measure spaces (K, d, η). The main obstacle is the impossibility of averaging K-valued random variables. This is overcome by using isometric images of K inside a Banach space or a Hilbert space, after which we can apply results for CLTs on these spaces.

Keywords: metric spaces; Central Limit Theorems.
MSC2020 subject classifications: 60F05.
Submitted to ECP on March 6, 2020, final version accepted on July 6, 2020.
Supersedes arXiv:2001.02793.

1 Introduction

We produce a series of Central Limit Theorems (CLTs) associated to compact metric measure spaces (K, d, η), with η a reasonable probability measure. The main obstacle is the impossibility of averaging K-valued random variables. This is overcome by using isometric images $\imath_d(K)$ of K inside a Banach space or a Hilbert space, after which we can apply results for CLTs on these spaces. For the first CLT, we can ignore η by isometrically embedding K into $\mathcal{C}(K)$, the space of continuous functions on K with the sup norm, and then applying known CLTs for sample means on Banach spaces (Theorem 3.4). However, the sample mean makes no sense back on K, so using η we develop a CLT for the sample Fréchet mean (Corollary 4.5). This involves working on the closed convex hull of the embedded image of K. To work in the easier Hilbert space setting of $L^2(K, \eta)$, we have to modify the metric d to a related metric d_η. We obtain a CLT for both the sample mean and the sample Fréchet mean (Theorem 5.2) with respect to the modified metric d_η, and we relate the Fréchet sample and population means on the closed convex hull to the Fréchet means on the image of K.

While these CLTs are for random variables taking values in $\imath_d(K)$, there is a bijection between the more natural set of K-valued random variables and the set of $\imath_d(K)$-valued random variables. Thus in the end we produce CLTs for K-valued random variables.

This work is motivated by our previous paper [5], where we investigated the uniqueness of the Fréchet mean on compact subsets K of the space of unlabeled networks with a fixed number of vertices. In some concrete examples, the Fréchet mean is not unique on K as a subset of Euclidean space, but is unique on the closed convex hull of K. This led us to consider when we can find a unique Fréchet mean for general compact metric spaces, and to search for corresponding CLTs.

*Boston University, United States of America. E-mail: sr@math.bu.edu
†Boston University, United States of America. E-mail: xujie@bu.edu

2 Background material

Throughout the paper, (K, d) will be a compact metric space. Recall that a G-valued random variable X is a function $X : \Omega \to G$, where $(\Omega, \mathcal{F}, P)$ is a fixed probability triple. The induced measure/distribution on G is given by $X_*(P)$, with

$$X_*(P)(A) = P(X^{-1}(A)),\ A \subset G.$$

We recall the setup and statement of a Central Limit Theorem on Banach spaces due to Zinn.

Definition 2.1. *(i) Let G be a Banach space. A probability measure γ on G is a Gaussian Radon measure if for every nontrivial linear functional $L : G \to \mathbb{R}$, on G, the pushforward measure $L_*(\gamma)$ is a non-degenerate Gaussian measure on $\mathbb{R}$, i.e., a standard Gaussian measure with non-zero variance.*

(ii) Let $X_1, \dots, X_n, \dots$ be any set of G-valued i.i.d. random variables with common distribution μ. μ satisfies the G-Central Limit Theorem (G-CLT) on G if there exists a Guassian Radon probability measure γ on G such that the distributions, μ_n, of $\frac{X_1+\ldots+X_n}{\sqrt{n}}$ converge, i.e., for every bounded-continuous real-valued function f on G,

$$\int_G f d\mu_n \to \int_G f d\gamma.$$

(iii) The metric d on K implies Gaussian continuity (or d is CGI) if whenever $\{X(s)\}_{s\in K}$ is a separable Gaussian process such that

$$\mathbb{E}\left[|X(t) - X(s)|^2\right] \leqslant d^2(t, s),$$

then X has continuous sample paths a.s.

Let $H_d(K, \epsilon) = \log(N_d(K, \epsilon))$, where $N_d(K, \epsilon)$ is the smallest number of d-balls of diameter at most 2ϵ which cover K.

Proposition 2.2 ([3, Thm. 3.1])**.** *If*

$$\int_0^\infty H_d^{1/2}(K, u) du < \infty, \tag{2.1}$$

then d is GCI.

Let $\mathcal{C}(K)$ be the Banach space of continuous functions on the compact metric space (K, d) equipped with the sup-norm $\|\cdot\|_\infty$. $\mathcal{C}(K)$ becomes a complete metric space with the induced distance function d_∞ by $d_\infty(f, g) = \|f - g\|_\infty, \forall f, g \in \mathcal{C}(K)$.

Definition 2.3. *For the compact metric space (K, d), set*

$$\mathrm{Lip}(d) = \left\{ x \in \mathcal{C}(K) : \sup_{t\neq s} \frac{|x(t) - x(s)|}{d(t, s)} < \infty \right\}.$$

$\mathrm{Lip}(d)$ is nonempty (by letting x be a constant function). We check that $\mathrm{Lip}(d)$ is closed. If $\{x_k\} \in \mathrm{Lip}(d)$ has $\lim_{k\to\infty} x_k = y \in \mathcal{C}(K)$, then for any $\epsilon > 0$ and $t \neq s$,

$$\begin{aligned}\frac{|y(t) - y(s)|}{d(t, s)} &\leq \frac{|y(t) - x_j(t)|}{d(t, s)} + \frac{|x_j(t) - x_j(s)|}{d(t, s)} + \frac{|x_j(s) - y(s)|}{d(t, s)} \\ &\leq 2\epsilon + \frac{|x_j(t) - x_j(s)|}{d(t, s)},\end{aligned}$$

for $j = j(\epsilon) \gg 0$ independent of t, s. This implies that $y \in \mathrm{Lip}(d)$.

Definition 2.4. *A Radon probability measure μ on the Banach space $(G, \|\cdot\|)$ has zero mean and finite variance, respectively, if*

$$\int_G x\ \mu(dx) = 0,\ \int_G \|x\|^2\ \mu(dx) < \infty, \tag{2.2}$$

respectively.

Of course, if a sequence of G-valued random variables X_i has finite expectation, then the new random variable $X_i - \mathbb{E}[X_i]$ has zero mean.

We can now state Zinn's CLT.

Theorem 2.5 ([7])**.** *Let (K, d) be a compact metric space with d CGI. If μ is a Radon probability measure on $Lip(d)$ with zero mean and finite variance, then μ satisfies the Central Limit Theorem on $(\mathcal{C}(K), \|\cdot\|_\infty)$ in the sense of Definition 2.1(ii).*

For our main results, we need to define the Fréchet mean with respect to a probability measure Q on (K, d). This generalizes the notion of centroids from vector spaces to metric spaces.

Definition 2.6. *(i) The Fréchet function $f : K \to \mathbb{R}$ is*

$$f(p) = \int_M d^2(p, z)Q(z)dz, p \in M.$$

If $f(p)$ has a unique minimizer $\mu = \operatorname{argmin}_{p\in K} f(p)$, we call μ the Fréchet mean of Q.

(ii) Given an i.i.d. sequence $X_1, \ldots, X_n \sim Q$ on M, the empirical Fréchet mean is defined to be

$$\mu_n = \operatorname{argmin}_{p\in M} \frac{1}{n}\sum_{i=1}^{n} d^2(p, X_i),$$

provided the argmin is unique.

Unlike centroids in Euclidean space, the uniqueness of Fréchet mean cannot be guaranteed, even in spaces which locally look like Euclidean space.

Example 2.7. We parametrize an open cone (minus a line) $\mathcal{Z} : x^2 + y^2 = z^2$ of height one by

$$F(u, v) = \left(\frac{1}{\sqrt{2}}u\cos v, \frac{1}{\sqrt{2}}u\sin v, \frac{1}{\sqrt{2}}u\right),\ (u, v) \in (0, 1) \times (0, 2\pi).$$

There is a Riemannian isometry from the sector $S = \{(r, \theta) \in (0, \sqrt{2}) \times (0, \sqrt{2}\pi)\}$ to $\mathcal{Z}$ induced by $\alpha : (r, \theta) \mapsto (u = r/\sqrt{2}, v = \sqrt{2}\theta)$, *i.e.,*

$$(r, \theta) \mapsto \left(\frac{r}{2}\cos(\sqrt{2}\theta), \frac{r}{2}\sin(\sqrt{2}\theta), \frac{r}{2}\right).$$

Indeed, the first fundamental form of the sector at (r, θ), respectively the first fundamental form of the cone at (u, v), are

$$\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix},\ \begin{pmatrix} 2 & 0 \\ 0 & \frac{u^2}{2} \end{pmatrix},$$

respectively. It is easy to check that the differential $d\alpha$ preserves these inner products. Thus every point $p \in S$ has a neighborhood U such that the usual Euclidean distance between $q_1, q_2 \in U$ equals the geodesic distance between $\alpha(q_1), \alpha(q_2)$.

It is easy to check that for e.g. the uniform distribution on S, the Fréchet mean/centroid $(\bar{x}, \bar{y})$ is inside S. In contrast, by the rotational symmetry of the geodesic distance function on $\mathcal{Z}$, the minima of the Fréchet function on $\mathcal{Z}$ form a circle containing $\alpha(\bar{x}, \bar{y})$.

For results on CLTs when the Fréchet mean is not unique, see [2].

3 A CLT for compact metric spaces

In this section, we isometrically embed the compact metric (K, d) into the Banach space $(\mathcal{C}(K), d_\infty)$ to obtain a CLT on the image of K.

We define

$$\imath_d : K \to \mathcal{C}(K), x \mapsto f_x := d(x, \cdot).$$

The following proposition is well known.

Proposition 3.1. *$\imath_d : (K, d) \to (\imath_d(K), d_\infty)$ is an isometry.*

Proof. See *e.g.*, [5, Supplement C, §2]. □

It follows that $\imath_d$ is an injection, and $\imath_d(K)$ is a compact subset of $\mathcal{C}(K)$.

We are interested in CLTs associated to K. However, it makes no sense to average K-valued random variables. As a result of Proposition 3.1, we can identify K-valued random variables with $\imath_d(K)$-valued random variables, which can be averaged. To obtain a CLT on $\imath_d(K)$ from Theorem 2.5, we need to verify its hypotheses.

Lemma 3.2. *$\imath_d(K) \subset \mathrm{Lip}(d)$.*

Proof. For $f_x \in \imath_d(K)$, the triangle inequality for $s, t \in K$ gives

$$|f_x(t) - f_x(s)| = |d(x, t) - d(x, s)| \leqslant d(s, t).$$

It follows that

$$\sup_{s \neq t} \frac{|f_x(s) - f_x(t)|}{d(s, t)} \leqslant 1. \qquad \square$$

In the following proof, we strongly use the fact that $\mathcal{C}(K)$ is a "linearization" of K.

Lemma 3.3. *The metric d on K is GCI.*

Proof. We must verify (2.1) in Proposition 2.2. Equivalently, we will show

$$\int_0^\infty H_{d_\infty}^{\frac{1}{2}}(\imath_d(K), u)du < \infty. \tag{3.1}$$

As a compact set, $\imath_d(K)$ can be covered by N balls of radius 1 for some N. Fix any point $x_0 \in \imath_d(K)$, and consider the d_∞-ball $B_\infty(x_0, 1)$ of radius 1 centered at x_0. The closure $\overline{B_\infty(x_0, 1)}$ equals $\imath_d(\overline{B_d(\imath_d^{-1}(x_0), 1)})$ of the corresponding ball in K. It follows that $\overline{B_\infty(x_0, 1)}$ is compact, so we can cover $B(x_0, 1)$ by M balls of radius $\frac{1}{2}$ for some M.

Since d_∞ is translation invariant, M is independent of the choice of x_0. Moreover, d_∞ is homogeneous in the sense that $d_\infty(cf, cg) = |c| d_\infty(f, g)$ for $c \in \mathbb{R}$. Thus for $r > 0$, any d_∞-ball of radius r contained in $\imath_d(K)$ can be covered by M balls of radius $\frac{r}{2}$. Hence

$$N_{d_\infty}(\imath_d(K), 2^{-k}) \leqslant N \cdot M^{k+1}.$$

To estimate (2.1), we integrate over $[0, 1]$ and $[1, \infty)$ separately. Since $\imath_d(K)$ is compact, it is covered by a single d_∞-ball $B_\infty(x_0, R)$ for some $R \gg 0$ and a fixed $x_0 \in \imath_d(K)$. Choose $k_0 \in \mathbb{N}$ such that $2^{k_0} \leqslant R < 2^{k_0+1}$. We have

$$\begin{aligned}
\int_1^\infty H_{d_\infty}^{\frac{1}{2}}(\imath_d(K), u)du &= \int_1^\infty \sqrt{\log(N_{d_\infty}(\imath_d(K), u))}du \\
&\leqslant \sum_{k=1}^{\infty} \sqrt{\log(N_{d_\infty}(\imath_d(K), 2^k))}(2^k - 2^{k-1}) \\
&\leqslant \sum_{k=1}^{k_0+1} \sqrt{\log(N_{d_\infty}(\imath_d(K), 2^k))} 2^{k-1} < \infty.
\end{aligned}$$

For the region $[0,1]$, we have

$$\begin{aligned}&\int_0^1 H^{\frac{1}{2}}_{d_\infty}(\imath_d(K),u)du\\ &=\int_0^1\sqrt{\log(N_{d_\infty}(\imath_d(K),u))}du\leqslant\sum_{k=0}^\infty\sqrt{\log(N_{d_\infty}(\imath_d(K),2^{-k-1}))}(2^{-k}-2^{-k-1})\\ &\leqslant\sum_{k=0}^\infty\sqrt{\log(N\cdot M^{k+2})}\,2^{-k-1}=\sum_{k=0}^\infty\sqrt{(k+2)\log M+\log N}\,2^{-k-1}<\infty.\end{aligned}$$

Adding these estimates gives (3.1). □

This gives our first Central Limit Theorem on K, or really on the isometric space $\imath_d(K)$.

Theorem 3.4 (General CLT)**.** *Let (K,d) be a compact metric space, let μ be a Radon probability measure on K with finite variance and such that $\imath_{d,*}\mu$ has zero mean on $\imath_d(K)$. Then $\imath_{d,*}\mu$ satisfies the G-CLT for $G=(\mathcal{C}(K),\|\cdot\|_\infty)$.*

Proof. By Lemmas 3.2, 3.3, the hypotheses of Theorem 2.1 are satisfied for $\imath_{d,*}\mu$. □

4 A Fréchet CLT associated to a compact metric space

In the previous section, we found a G-CLT on a Banach space associated with the usual sample mean $\sqrt{n}\cdot\frac{1}{n}\sum_{i=1}^n X_i$ on $G=\mathcal{C}(K)$. In this section, we prove a G-CLT on the compact metric space K, endowed with a Radon probability measure η, for the sample Fréchet mean

$$\text{argmin}_{Y\in\imath_d(K)}\frac{1}{n}\sum_{i=1}^n\|X_i-Y\|^2_{2,\eta},\tag{4.1}$$

$X_1,\dots,X_n$ are i.i.d. $\imath_d(K)$-valued random variables, and the L^2 norm is taken with respect to $\imath_{d,*}\eta$.

We emphasize that these CLTs can be applied to the more natural K-valued random variables, since $f\mapsto\imath_d\circ f$ gives a bijection from K-valued random variables to $\imath_d(K)$-valued random variables. Thus we can think of these CLTs as applied to K-valued random variables which are pushed forward to $\imath_d(K)$-valued random variables in order to take averages.

Note that we compute the sample Fréchet mean with respect to the L^2-norm, since we will need a Hilbert space structure below. The minimizer of (4.1) may not be unique, since $\imath_d(K)$ may not be convex in $\mathcal{C}(K)$. Instead, we consider the closed convex hull of $\imath_d(K)$, on which the uniqueness of the Fréchet mean is guaranteed.

Definition 4.1. *Let $\imath_d(K)^c$ be the convex hull of $\imath_d(K)$, i.e., the intersection of all convex subsets of $\mathcal{C}(K)$ containing $\imath_d(K)$, and let*

$$S_d=S_d(K)=\overline{\imath_d(K)^c}$$

be the closure of $\imath_d(K)^c$.

By [1, Thm. 5.35], S_d is a compact subset of $\mathcal{C}(K)$. It is easy to check that $S_d(K)\subset\mathrm{Lip}(d)$. As the minimizer of a convex function on a closed convex space, the Fréchet mean is unique. However, the Fréchet mean in $S_d(K)$ may not lie in the image $\imath_d(K)$ of K, as the next example shows. In this case, we cannot identify the Fréchet mean in $S_d(K)$ with a point in K.

Example 4.2. To continue with Example 2.1, choose a probability measure Q on the cone $\mathcal{Z}$. The sample Fréchet mean for K-valued random variables Y_i lies in the interior of $\mathcal{Z}$ in $\mathbb{R}^3$. It is unclear if the sample Fréchet mean for the $\imath_d(\mathcal{Z})$-valued random variables $X_i = \imath_d \circ Y_i$ lies in $\imath_d(\mathcal{Z})$, but it certainly lies in $S_d(\mathcal{Z})$. (This example doesn't really show the strength of embedding $\mathcal{Z}$ into $\mathcal{C}(\mathcal{Z})$, since $\mathcal{Z}$ lies in a linear space.

Similar remarks apply to the Fréchet minimum. While the Fréchet minimum for the cone $(\mathcal{Z}, \eta)$ is not unique, the Fréchet minimum on $(S_d(\mathcal{Z}), \imath_{d,*}\eta)$, the closed convex hull of the isometric set $(\imath_d(\mathcal{Z}), \imath_{d,*}\eta)$, is unique. (Note that the sector S in Example 2.1 is only locally isometric to $\mathcal{Z}$.) While we have gained uniqueness, there is no reason why the Fréchet minimum need be inside $\imath_d(\mathcal{Z})$, as in Example 2.1. It is shown in [5, Supplement C] that in general the distance from the Fréchet mean in $S_d(K)$ to $\imath_d(K)$ is at most twice the diameter of K.

At this point we have the embeddings $K \hookrightarrow \imath_d(K) \subset S_d(K) \subset \mathcal{C}(K) \subset L^2(K)$, where $L^2(K)$ is taken with respect to a probability measure on K. While $K \hookrightarrow L^2(K)$ is no longer an isometry, there is a known CLT on $S_d(K)$ equipped with the L^2 norm.

Theorem 4.3 (L^2 CLT)**.** *Let μ be a Radon probability measure supported in K such that $\imath_{d,*}\mu$ satisfies (2.2). Then $\imath_{d,*}\mu$ satisfies the G-CLT for $G = (L^2(K), \|\cdot\|_{2,\eta})$. The same result holds if the random variables $\{X_i\}$ in the G-CLT are $\imath_d(K)$-valued and/or μ has support in $S_d(K)$.*

Proof. By [6, Thm. 9.10], the Hilbert space $L^2(K)$ is of type 2 and cotype 2. The existence of a CLT on spaces of type/cotype 2 follows from [4, Thm. 3.5]. □

We also obtain a CLT for the sample Fréchet mean. Here the Hilbert space structure works to our advantage, as the sample Fréchet mean and the usual sample mean coincide.

Proposition 4.4. *For $S_d(K)$-valued random variables $\{X_i\}$, we have*

$$S_n := \frac{1}{n}\sum_{i=1}^n X_i = \mathrm{argmin}_{Y \in S_d} \frac{1}{n}\sum_{i=1}^n \|X_i - Y\|_{2,\eta}^2.$$

Proof. It is well-known that in a finite dimensional Euclidean space, the sample mean coincides with the sample Fréchet mean. For fixed $x \in K$, the real-valued random variables $\{X_i(x)\}$ satisfy

$$\frac{1}{n}\sum_{i=1}^n |X_i(x) - S_n(x)|^2 \leqslant \frac{1}{n}\sum_{i=1}^n |X_i(x) - Y(x)|^2, \ \forall Y \in \mathcal{C}(K).$$

Therefore, for all Y,

$$\frac{1}{n}\sum_{i=1}^n \int_K |X_i(x) - S_n(x)|^2 dQ(x) \leqslant \frac{1}{n}\sum_{i=1}^n \int_K |X_i(x) - Y(x)|^2 dQ(x),$$

so

$$\frac{1}{n}\sum_{i=1}^n \|X_i - S_n\|_{2,\eta}^2 \leqslant \frac{1}{n}\sum_{i=1}^n \|X_i - Y\|_{2,\eta}^2 \Rightarrow S_n = \mathrm{argmin}_{Y \in S_d} \frac{1}{n}\sum_{i=1}^n \|X_i - Y\|_{2,\eta}^2. \qquad \square$$

Combining the Proposition with Theorem 4.3 gives us a CLT for the sample Fréchet mean. We set $\|\cdot\|_{2,\eta}$ be the L^2 norm with respect to a measure μ, and set $\mathcal{C}_0(X)$ to be the set of bounded continuous functions on a topological space X.

Corollary 4.5 (Sample Fréchet mean CLT)**.** *(i) Let $\{X_i\}$ be i.i.d. S_d-valued random variables with distribution μ a Radon probability measure supported in $\imath_d(K)$ satisfying (2.2).*

Then there exists a Gaussian Radon probability measure $\tilde{\gamma}_2$ such that the distributions μ_n of

$$\mathrm{argmin}_{Y\in S_d}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\|X_i-Y\|_{2,\eta}^2$$

converge weakly to $\tilde{\gamma}_2$ in the sense of Defintion 2.1.

(ii) γ_2 in Theorem 4.3 equals $\tilde{\gamma}_2$ in distribution. In particular, for $f\in\mathcal{C}_0(S_d(K))$,

$$\int_{S_d(K)} f d\gamma_2 = \int_{S_d(K)} f d\tilde{\gamma}_2.$$

(iii) Let γ_1 be the limiting measure obtained in Theorem 3.4. Then $\gamma_1=\gamma_2$.

Proof. (i) follows from Proposition 4.4. For (ii), the Proposition implies that the distributions of the sample mean and the sample Fréchet mean are the same a.s. (iii) follows from the uniqueness of γ in Definition 2.1(ii) and comparing Theorems 3.4 and 4.3. □

5 L^2 techniques and G-CLTs

In this section, we embed the compact metric space K, now equipped with a Radon measure η and a modified metric, into $L^2(K,\eta)$ to produce an L^2 version of a G-CLT. In this Hilbert space setting, we are able to relate the Fréchet means of the closed convex hulls to the Fréchet means on the embedded image of K.

We define a seminorm on $\imath_d(K)=\{f_x=d(x,\cdot):x\in K\}$ by

$$\|f_x\|_{2,\eta}^2=\int_{\imath_d(K)}|f_x(y)|^2 d\imath_{d,*}\eta(y)=\int_K d^2(x,y)d\eta(y).$$

Taking the completion modulo the space of norm zero functions gives the Hilbert space $(L^2(\imath_d(K)),\|\cdot\|_{2,\eta})$. More precisely, we will prove $L^2(\imath_d(K))$-CLTs for both the sample mean and the sample Fréchet mean.

The norm $\|\cdot\|_{2,\eta}$ induces a metric $d_{2,\eta}$ on $\imath_d(K)$. Since $\imath_d:(K,d)\to(\imath_d(K),d_{2,\eta})$ is easily continuous, we can pull back $d_{2,\eta}$ to a metric $d_\eta:=\imath_d^* d_{2,\eta}$ on K:

$$d_\eta(x,y)=d_{2,\eta}(f_x,f_y)=\left(\int_K (d(x,z)-d(y,z))^2 d\eta(z)\right)^{1/2}.$$

Thus $\imath_{d_\eta}:(K,d_\eta)\to(\imath_{d_\eta}(K),d_{2,\eta})$, defined by $\imath_{d_\eta}(x)=d_\eta(x,\cdot)$, is an isometry. We interpret (K,d_η) as a modification of (K,d) which keeps track of the L^2 information of η.

We want to relate the various metrics. Let d_∞ be the metric on $\mathcal{C}(K)$ induced by the sup norm $\|\cdot\|_\infty$, and let $[\imath_d(K)]$ be the image of $\imath_d(K)$ in $L^2(K)$. Consider the maps

$$(K,d)\xrightarrow{\imath_d}(\imath_d(K),d_\infty)\xrightarrow{F}([\imath_d(K)],d_{2,\eta})\xrightarrow{G}(K,d_\eta)$$

given by $F(f_x)=[f_x]$, where we take the L^2 equivalence class, and $G([f_x])=\imath_{d_\eta}^{-1}(f_x)=\imath_{d_\eta}^{-1}\imath_d(x)$. (We show that G is well-defined below.) $\imath_d$ is an isometry.

In general, F and G are not injective, since equivalence classes in $L^2(K)$ have many representatives, without a restriction on η.

Lemma 5.1. *Assume that every d-ball $B_\epsilon(x)$ centered at $x\in K$ has $\eta(B_\epsilon(x))>0$. Then F is injective, and G is well-defined and injective.*

Since F and G are trivially surjective, they are bijective under the Lemma's hypothesis. Note that for Lebesgue measure and the standard metric on $\mathbb{R}^N$, the hypothesis is satisfied, while delta functions give rise to Radon measures that do not satisfy the hypothesis.

Proof. For F, it suffices to show that $F \circ \imath_d : x \mapsto [f_x]$ is injective. For $x \neq y$, and $\epsilon < d(x,y)/3$, we have

$$d(y,z) \geq d(x,y) - d(x,z) > 3\epsilon - \epsilon > d(x,z) + \epsilon,$$

for all $z \in B_\epsilon(x)$. Therefore

$$\begin{aligned} d_{2,\eta}([f_x],[f_y])^2 &= \int_K |f_x(z) - f_y(z)|^2 d\eta_z \geq \int_{B_\epsilon(x)} |d(x,z) - d(y,z)|^2 d\eta_z \\ &> \epsilon^2 \eta(B_\epsilon(x)) > 0. \end{aligned}$$

Thus $[f_x] \neq [f_y]$.

To show that G is well-defined, if $f_x, f_y \in [f_x] \in L^2(K)$, then

$$d_{2,\eta}([f_x],[f_y]) = 0 \Rightarrow \int_K |d(x,z) - d(y,z)|^2 d\eta(z) = 0.$$

As above, this implies that $x = y$, so $[f_x]$ has a unique representative of the form f_x. Since $\imath_d, \imath_\eta$ are injective, it follows that G is injective. □

We can now state and prove an L^2 CLT on $S_{d_\eta}(K)$ for both the sample mean and the sample Fréchet mean. As before, let $S_{d_\eta} = S_{d_\eta}(K)$ be the closed convex hall of $\imath_{d_\eta}(K)$ in the sup norm.

Theorem 5.2 (L^2 CLTs for sample means)**.** *Let $\{X_i\}$ be i.i.d. S_{d_η}-valued random variables with distribution μ a Radon probability measure supported in $S_{d_\eta}(K)$ satisfying (2.2). Assume that the hypothesis of Lemma 5.1 holds.*

(i) There exists a Gaussian Radon probability measure γ on S_{d_η} such that the distributions μ_n, of $\frac{X_1+\ldots+X_n}{\sqrt{n}}$ converge to a probability measure γ in the sense of Definition 2.1(ii).

(ii) The distributions $\tilde{\mu}_n$ of $\operatorname{argmin}_{Y\in S_{d_\eta}} \frac{1}{\sqrt{n}} \sum_{i=1}^n \|X_i - Y\|_{2,\eta}^2$ converge in the same sense to the same measure γ.

(iii) Under the hypothesis in Lemma 5.1, the distributions $\tilde{\mu}_n$ of $\operatorname*{argmin}_{Y\in S_{d_\eta}} \frac{1}{\sqrt{n}} \sum_{i=1}^n \|X_i - Y\|_\infty^2$ converge in the same sense to a Gaussian Radon probability measure $\tilde{\gamma}$.

Proof. (i) Replacing the metric d by d_η in Theorem 3.4 gives the CLT for the sample mean.

(ii) Applying Corollary 4.5(i) and (iii) to d_η gives the convergence of $\tilde{\mu}_n$ to the same measure γ.

(iii) The main idea is to use the isometric bijection $\imath_{d_\eta} \circ \imath_d^{-1} : (\imath_d(K), \|\cdot\|_\infty) \to (\imath_{d_\eta}(K), \|\cdot\|_{2,\eta})$. This extends linearly to an isometric bijection

$$\imath_{d_\eta} \circ \imath_d^{-1} : (S_d(K), \|\cdot\|_\infty) \to (S_{d_\eta}(K), \|\cdot\|_{2,\eta}).$$

Set

$$Z_n := \text{argmin}_{Y\in S_{d_\eta}} \frac{1}{n} \sum_{i=1}^n \|X_i - Y\|_\infty^2.$$

By Proposition 4.4,

$$\begin{aligned} Z_n &= \text{argmin}_{Y\in S_{d_{2,\eta}}} \frac{1}{n} \sum_{i=1}^n \|\imath_d \circ \imath_{d_\eta}^{-1} X_i - \imath_d \circ \imath_{d_\eta}^{-1} Y\|_{2,\eta}^2 \\ &= \imath_{d_\eta} \circ \imath_d^{-1} \left(\text{argmin}_{Y\in S_d} \frac{1}{n} \sum_{i=1}^n \|\imath_d \circ \imath_{d_\eta}^{-1} X_i - Y\|_{2,\eta}^2 \right) \\ &= \imath_{d_\eta} \circ \imath_d^{-1} \left(\frac{1}{n} \sum_{i=1}^n \imath_d \circ \imath_{d_\eta}^{-1} X_i \right). \end{aligned}$$

$\{(\imath_d \circ \imath_{d_\eta}^{-1})(X_i)\}$ are S_d-valued i.i.d. random variables with common distribution $(\imath_d \circ \imath_{d_\eta}^{-1})_* \mu$. By Theorem 4.3, we obtain a S_d-CLT with respect to a Gaussian Radon measure γ' on S_d. The isometry $\imath_{d_\eta} \circ \imath_d^{-1}$ then gives the S_{d_η}-CLT with respect to $(\imath_{d_\eta} \circ \imath_d^{-1})_* \gamma'$. □

Because we are in a Hilbert space setting, we can prove that the Fréchet sample and population means on S_{d_η} and on $\imath_{d_\eta}(K)$ have simple relationships.

Let S_{d_η} be the closed convex hull of $\imath_{d_\eta}(K) := K_0$ in $L^2(K, \eta)$, and let $f_y^2 = d_\eta(y, \cdot) \in \mathcal{C}(K)$. Let

$$\begin{aligned} \bar{F}(\bar{x}) &= \int_{S_{d_\eta}} d_{L^2}^2(\bar{x}, \bar{y}) d\imath_{d_\eta,*}\eta(\bar{y}) = \int_{K_0} d_{L^2}^2(\bar{x}, \bar{y}) d\imath_{d_\eta,*}\eta(\bar{y}) = \int_K d_{L^2}^2(\bar{x}, f_y^2) d\eta(y), \\ F(x) &= \int_K d_\eta^2(x, y) d\eta(y) = \int_{K_0} d_{L^2}^2(f_x^2, f_y^2)) d\imath_{d_\eta,*}\eta(f_y^2), \end{aligned}$$

be the L^2 Fréchet functions of S_{d_η} and K, respectively, and let

$$\bar{\mu} = \mathrm{argmin}_{\bar{x} \in S_{d_\eta}} \bar{F}(\bar{x}),\ \mu = \mathrm{argmin}_{x \in K} \int_K d_\eta^2(x, y) d\eta(y)$$

be the population means on S_{d_η} and K, respectively. Set $\mu_0 = \imath_{d_\eta}(\mu)$.

We note that as the minimum of a convex function on a convex set, $\bar{\mu}$ is unique. Also, gradients of differentiable functions exist in Hilbert spaces.

Proposition 5.3 (Relationships between the means)**.** *Assume that (i) μ is unique, (ii) K_0 is the zero set of a Fréchet differentiable function $H : L^2(K, \eta) \to \mathbb{R}$ with $\nabla H_{\mu_0} \neq 0$. Then μ_0 is the closest point in K_0 to $\bar{\mu}$. The same relationship holds for the sample Fréchet means of K_0-valued i.i.d. random variables.*

Proof. Note that $K_0 = \partial K_0$ in $L^2(K, \eta)$, since a compact subset of an infinite dimensional space has no interior. Also, $\bar{\mu} \in K_0$ implies $\bar{\mu} = \mu_0$, so we may assume $\bar{\mu} \notin K_0$.

The method of Lagrangian multipliers is valid in $L^2(K, \eta)$, so there exists $\lambda \in \mathbb{R}$ with $\nabla \bar{F}_{\mu_0} = \lambda \nabla H_{\mu_0}$. The differential $D\bar{F}$ at μ_0 is given by

$$\begin{aligned} D\bar{F}_{\mu_0}(v) &= (d/dt)|_{t=0} \int_{S_{d_\eta}} d_{L^2}^2(\mu_0 + tv, f_y^2) d\eta(y) \\ &= (d/dt)|_{t=0} \int_K \langle \mu_0 + tv - f_y^2, \mu_0 + tv - f_y^2 \rangle d\eta(y) \\ &= 2 \left\langle v, \mu_0 - \int_K f_y^2 d\eta(y) \right\rangle, \end{aligned}$$

where the last term equals the Hilbert space integral

$$\int_{K_0} \bar{y} d\imath_{d_\eta,*}\eta(\bar{y}) = \int_{S_{d_\eta}} \bar{y} d\imath_{d_\eta,*}\eta(\bar{y}).$$

Thus $\nabla \bar{F}_{\mu_0} = 2(\mu_0 - \int_{S_{d_\eta}} \bar{y} d\imath_{d_\eta,*}\eta(\bar{y}))$. Since $\nabla \bar{F}_p = 0$ only at $p = \bar{\mu}$, we see that $\bar{\mu} = \int_{S_{d_\eta}} \bar{y} d\imath_{d_\eta,*}\eta(\bar{y})$. (This is the usual statement that the Fréchet mean is the center of mass of a convex set in $\mathbb{R}^n$.) Thus $\nabla \bar{F}_{\mu_0} = 2(\mu_0 - \bar{\mu})$.

Since $\mu_0 - \bar{\mu}$ is a multiple of ∇H_{μ_0}, which is perpendicular to the level set $\imath_{d_\eta}(K)$, we have $\mu_0 - \bar{\mu} \perp \partial K_0$. We have not used that μ_0 is a minimum, so the same perpendicularity holds at any critical point of $\bar{F}$ on ∂K_0. We can translate in S_{d_η} so that $\bar{\mu} = 0$, in which case $\nabla \bar{F}_p = 2p$ is twice the Euler vector field. The level sets $\bar{F}^{-1}(r)$ are thus spheres centered at the origin in S_{d_η}. Since μ_0 is on the lowest level set of any point in K_0, μ_0 is closer to the origin than any other critical point of $\bar{F}$ on ∂K_0.

If we consider the distance function $D : K_0 \to \mathbb{R}, D(\bar{x}) = d^2(\mu_0, x)$, then a Lagrangian multiplier argument as above shows that at a critical point p of D, we have $\mu_0 - p \perp \partial K_0$. Thus μ_0 is a critical point of D, and by the last paragraph μ_0 must be the closest point in K_0 to μ_0.

The same argument holds for the sample means. $\square$

If a closest point $p(z) \in K_0$ to each $z \in S_{d_\eta}$ can be chosen so that p is continuous, as in the unlikely case that K_0 is convex, then we get a G-CLT on K_0 and hence on K for $p_*\tilde{\mu}_n, p_*\gamma$. This would connect Theorem 5.2 and Proposition 5.3.

References

[1] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis, A Hitchhiker's Guide*. Springer, 3rd edition, 2006. MR-2378491

[2] R. Bhattacharya and L. Lin. Omnibus CLTs for Fréchet means and nonparametric inference on Non-Euclidean spaces. *Proceedings of American Mathematical Society*, 145:413–428, 2017. MR-2378491

[3] R.M. Dudley. The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *Journal of Functional Analysis*, 1:290–330, 1967. MR-0220340

[4] J. Hoffmann-Jørgensen and G. Pisier. The law of large numbers and the Central Limit Theorem in Banach spaces. *The Annals of Probability*, 4(4):587–599, 1976. MR-0423451

[5] Eric D. Kolaczyk, Lizhen Lin, Steven Rosenberg, Jackson Walters, and Jie Xu. Averages of unlabeled networks: Geometric characterization and asymptotic behavior. *Annals of Statistics*, 48(1):514–538, 2020. arXiv:1709.02793v5. MR-0423451

[6] Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces: Isoperimetry and Processes*. Modern Surveys in Mathematics. Springer-Verlag, Berlin, 1991. MR-0423451

[7] Joel Zinn. A note on the Central Limit Theorem in Banach spaces. *The Annals of Probability*, 5(2):283–286, 1977. MR-0443026