Strong Consistency of Set-Valued Frechet Sample Mean in Metric Spaces

Cedric E. Ginestet

Department of Mathematics and Statistics
Boston University

JSM – 2013
The Frechet Mean

Barycentre as Average

- Given a set of points \( \mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^k \);
- And a set of masses \( w_1, \ldots, w_n \in \mathbb{R} \);
- The barycentre is

\[
\bar{\mathbf{x}} := \frac{1}{n} \sum_{i=1}^{n} w_i \mathbf{x}_i.
\]  

(1)

- When the distribution of mass is uniform, this is the centroid.
The Frechet Mean

Barycentre as Average

- Given a set of points \( x_1, \ldots, x_n \in \mathbb{R}^k \);
- And a set of masses \( w_1, \ldots, w_n \in \mathbb{R} \);
- The barycentre is
  \[
  \bar{x} := \frac{1}{n} \sum_{i=1}^{n} w_i x_i.
  \]  

- When the distribution of mass is uniform, this is the centroid.

Barycentre as Minimizer

- Take the \( L^2 \)-metric on \( \mathbb{R}^k \);
- The barycentre is a minimizer,
  \[
  \bar{x} = \arg\min_{x' \in \mathbb{R}^k} \frac{1}{n} \sum_{i=1}^{n} w_i \| x_i - x' \|_2^2.
  \]
The Frechet Mean

The Most ‘Central’ Element

- Given a metric space \((\mathcal{X}, d)\),
- Let an abstract-valued r.v. \(X : (\Omega, \mathcal{F}, \mathbb{P}) \mapsto (\mathcal{X}, \mathcal{B})\),
- Fréchet (1948) defined the ‘mean’ value, for every \(r \geq 1\),

\[
\Theta^r := \arg\inf_{x' \in \mathcal{X}} \int_{\mathcal{X}} d(x, x')^r d\mu(x).
\]
The Frechet Mean

The Most ‘Central’ Element

- Given a metric space \((\mathcal{X}, d)\),
- Let an abstract-valued r.v. \(X : (\Omega, \mathcal{F}, \mathbb{P}) \mapsto (\mathcal{X}, \mathcal{B})\),
- Fréchet (1948) defined the ‘mean’ value, for every \(r \geq 1\),

\[ \Theta^r := \arg\inf_{x' \in \mathcal{X}} \int_{\mathcal{X}} d(x, x')^r d\mu(x). \]

Frechet Sample Mean

- Given a family of abstract-valued r.v.s \(X_i\)’s, \(i = 1, \ldots, n\),

\[ \hat{\Theta}^r_n := \arg\inf_{x' \in \mathcal{X}} \frac{1}{n} \sum_{i=1}^{n} d(X_i, x')^r. \]

- In general, \(\Theta^r, \hat{\Theta}^r_n\) will not be unique: i.e. \(\Theta^r, \hat{\Theta}^r_n \subseteq \mathcal{X}\).
Motivating Example

Figure: A Sample of Two Simple Graphs, $G$ with $|V(G)| = 4$. 
Motivating Example

Figure: A Sample of Two Simple Graphs, $G$ with $|V(G)| = 4$.

Hamming Distance

$$d(G, G') := \sum_{i<j} I\{e_{ij} \neq e'_{ij}\}.$$
Non-unique Means

Figure: A Sample of Two Simple Graphs, $G$ with $|V(G)| = 4$.

Figure: Set of Frechet means ($\Theta$) is a superset of the sampled graphs.
Questions and Assumptions

Convergence of Frechet Sample Mean

\[ \hat{\Theta}_n^r := \arg\inf_{x' \in \mathcal{X}} \frac{1}{n} \sum_{i=1}^n d(X_i, x')^r \to \arg\inf_{x' \in \mathcal{X}} \int_{\mathcal{X}} d(x, x')^r \, d\mu(x) =: \Theta^r, \]

- Ziezold (1977) in separable finite metric spaces, for \( r = 2; \)
- Sverdrup-Thygeson (1981) in compact metric spaces, for \( r \geq 1. \)
Questions and Assumptions

Convergence of Frechet Sample Mean

\[ \hat{\Theta}_n^r := \operatorname{arginf}_{x' \in \mathcal{X}} \frac{1}{n} \sum_{i=1}^{n} d(X_i, x')^r \rightarrow \operatorname{arginf}_{x' \in \mathcal{X}} \int_{\mathcal{X}} d(x, x')^r d\mu(x) =: \Theta^r, \]

- Ziezold (1977) in separable finite metric spaces, for \( r = 2 \);
- Sverdrup-Thygeson (1981) in compact metric spaces, for \( r \geq 1 \).

Topological and Measure-theoretic Assumptions

- Separable bounded \((\mathcal{X}, d)\).
- Possibly empty \(\hat{\Theta}_n^r, \Theta^r\).
- Possibly non-unique \(\hat{\Theta}_n^r, \Theta^r\).
Set-valued Outer and Inner Limits

Figure: Outer and inner limits, defined wrt set inclusion on $\mathcal{X}$.

\[ \limsup_{n \to \infty} S_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} S_n, \quad \text{and} \quad \liminf_{n \to \infty} S_n := \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} S_n. \]

- **Outer Limit:** The $x$’s that belong to *infinitely many* $S_n$.
- **Inner Limit:** The $x$’s that belong to *all but finitely many* $S_n$. 
Kuratowski Upper Limit

Equivalent Definitions

Given a sequence of subsets $A_n \subseteq \mathcal{X}$:

- Set of cluster points of sequences $a_n \in A_n$;
- The Kuratowski upper limit is

$$\text{Limsup } A_n := \left\{ x \in \mathcal{X} : \liminf_{n \to \infty} d(x, A_n) = 0 \right\}.$$ 

where liminf and Limsup are taken with respect to real numbers and subsets of $\mathcal{X}$, respectively.

Lemma

Given a metric space $(\mathcal{X}, d)$, for any sequence of sets $A_n \subseteq \mathcal{X}$,

$$\text{Limsup } A_n = \bigcap_{n \to \infty} \bigcup_{n=1}^{\infty} A_m.$$
Almost Sure Consistency of Frechet Sample Mean

Main Result

- In a separable bounded metric space \((X, d)\),
- For any \(X_1, \ldots, X_n\) be an iid sequence,
- Convergence of the infima:

\[
\hat{\sigma}_n^r := \inf_{x' \in X} \frac{1}{n} \sum_{i=1}^{n} d(X_i, x')^r \rightarrow \inf_{x' \in X} \int_X d(x, x')^r \, dx =: \sigma^r \quad \text{a.s.},
\]

- Convergence of the arginf's:

\[
\limsup_{n \to \infty} \hat{\Theta}_n^r \subseteq \Theta^r \quad \text{a.s.}
\]

- For every \(r > 0\).
Glivenko-Cantelli Lemma (Rao, 1962)

i. Let $\mathcal{F}(\mathcal{X})$ be a class of real-valued functions on separable $\mathcal{X}$,

ii. Let a sequence of finite measures $\mu_n$, and $\mu$,

iii. $\mathcal{F}(\mathcal{X})$ is dominated by a continuous integrable function on $\mathcal{X}$,

iv. $\mathcal{F}(\mathcal{X})$ is equicontinuous, and

v. $\mu_n \Rightarrow \mu$, a.s.;

then we obtain uniform a.s. weak convergence,

$$\lim_{n \to \infty} \sup_{f \in \mathcal{F}} \left| \int f \, d\mu_n - \int f \, d\mu \right| = 0, \quad \text{a.s..}$$
Definition
For some \( z \in \mathcal{X} \), the \( z \)-point function is

\[ d_z(x) := d(z, x), \]

for every \( x \in \mathcal{X} \). The class of point functions on \((\mathcal{X}, d)\) is then denoted by

\[ \mathcal{D}^r(\mathcal{X}) := \{ d^r_z : \forall z \in \mathcal{X} \}. \]

for every \( r \geq 1 \).

Lemma
If \((\mathcal{X}, d)\) is a bounded metric space, then for every \( r \geq 1 \), \( \mathcal{D}^r(\mathcal{X}) \) is

i. Uniformly bounded;

ii. Uniformly equicontinuous.
Almost Sure Consistency of Frechet Sample Mean

Strengthening of a.s. Weak Convergence

Using the Glivenko-Cantelli lemma, we obtain for every \( r \geq 1, \)

\[
\lim_{n \to \infty} \sup_{z \in D_r} \left| \int_X d_x^r \, d\mu_n - \int_X d_x^r \, d\mu \right| = 0 \quad \text{a.s.}
\]

where \( \mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{X_i}. \)
Almost Sure Consistency of Frechet Sample Mean

Strengthening of a.s. Weak Convergence

Using the Glivenko-Cantelli lemma, we obtain for every \( r \geq 1 \),

\[
\lim_{n \to \infty} \sup_{z \in \mathcal{D}^r} \left| \int_X d_z \, d\mu_n - \int_X d_z \, d\mu \right| = 0 \quad \text{a.s.}
\]

where \( \mu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i} \).

Almost Sure Convergence of \( \hat{\sigma}_n^r \)

- In order to show that \( \hat{\sigma}_n^r \to \sigma^r \) a.s.,
- We also need the following ‘sandwich’ relationship,

\[
T_n^r \leq \hat{\sigma}_n^r - \sigma^r \leq T_n^r.
\]
Almost Sure Consistency of Frechet Sample Mean

Sandwich Argument I

By the minimality of $\theta \in \Theta^2$,

$$T_n(\hat{\theta}_n) := \frac{1}{n} \sum_{i=1}^{n} d^2_{\hat{\theta}_n}(X_i) - \int_{X} d^2_{\hat{\theta}_n}(x) d\mu(x)$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} d^2_{\hat{\theta}_n}(X_i) - \int_{X} d^2_{\theta}(x) d\mu(x) =: T^*_n(\hat{\theta}_n).$$
Almost Sure Consistency of Frechet Sample Mean

Sandwich Argument I
By the minimality of $\theta \in \Theta^2$,

$$T_n(\hat{\theta}_n) := \frac{1}{n} \sum_{i=1}^{n} d_{\hat{\theta}_n}^2(X_i) - \int_{\mathcal{X}} d_{\hat{\theta}_n}^2(x) d\mu(x)$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} d_{\hat{\theta}_n}^2(X_i) - \int_{\mathcal{X}} d_{\theta}^2(x) d\mu(x) =: T_n^*(\hat{\theta}_n).$$

Sandwich Argument II
By the minimality of $\hat{\theta}_n \in \widehat{\Theta}_n^2$,

$$T_n^*(\hat{\theta}_n) := \frac{1}{n} \sum_{i=1}^{n} d_{\hat{\theta}_n}^2(X_i) - \int_{\mathcal{X}} d_{\theta}^2(x) d\mu(x)$$

$$\leq \frac{1}{n} \sum_{i=1}^{n} d_{\theta}^2(X_i) - \int_{\mathcal{X}} d_{\theta}^2(x) d\mu(x) =: T_n(\theta).$$
Almost Sure Consistency of Frechet Sample Mean

Main Result

- In a separable bounded metric space \((X, d)\),
- For any \(X_1, \ldots, X_n\) be an iid sequence,
- Convergence of the infima:

\[
\hat{\sigma}_n^r := \inf_{x' \in X} \frac{1}{n} \sum_{i=1}^{n} d(X_i, x')^r \to \inf_{x' \in X} \int_X d(x, x')^r \, dx =: \sigma^r \ a.s.,
\]

- Convergence of the arginf’s:

\[
\limsup_{n \to \infty} \hat{\Theta}_n^r \subseteq \Theta^r \ a.s.,
\]

- For every \(r > 0\).
Conclusion and Extensions

Consistency of Frechet Sample Mean

- Under separable $\mathcal{X}$ and bounded $d$.
- For non-unique means.
- For every $r \geq 1$.

Possible Extensions

- Distance functions without the triangle inequality.
- Non-iid random variables (See Kendall et al., 2012).
- Rate of convergence.
- Statistical inference.
Publications & Funding Agencies

Papers

- Ginestet, C.E. (Submitted). Strong Consistency of Fréchet Sample Mean Sets for Graph-Valued Random Variables. *ArXiv*.

Acknowledgements

- Eric D. Kolaczyk (Boston University)
- Tom Nichols and Wilfrid S. Kendall (Warwick University)

Funding Agency

- Air Force Office for Scientific Research (AFOSR).