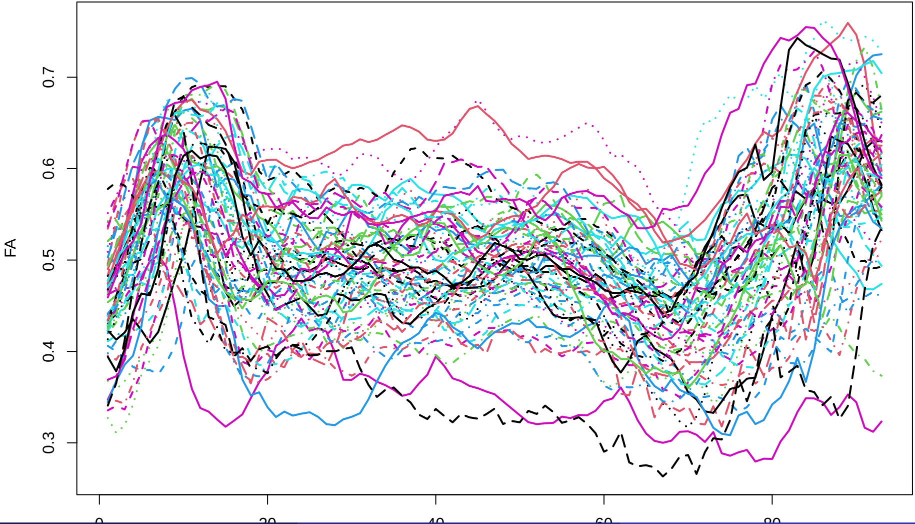
Random Functions

School of Statistics and Data Science

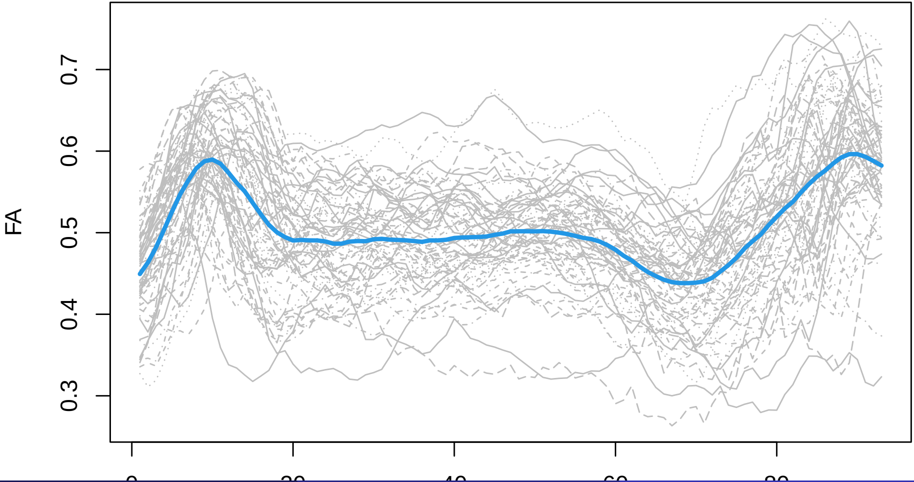
Diffusion Tensor Imaging Example

Fractional anisotropy (FA) is a measure of water diffusion in the brain. We consider FA tract profiles for the corpus callosum (CCA) in MS patients.



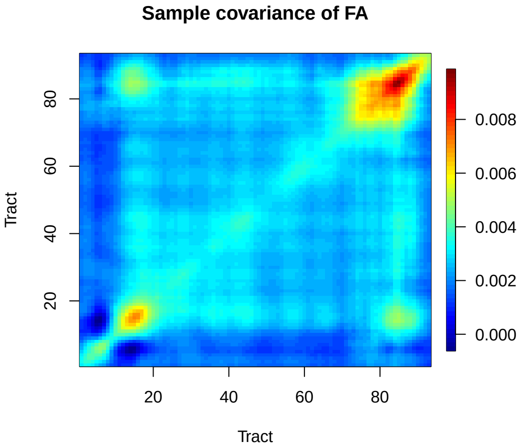
Sample Mean

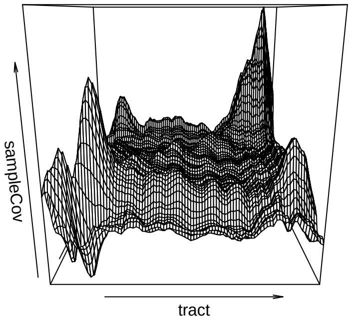
$$\hat{\mu}(t) = \frac{1}{n}\sum_{i=1}^{n} X_i(t)$$



Sample Covariance

$$\hat{c}(s,t) = \frac{1}{n-1}\sum_{i=1}^{n}\{X_i(s)-\bar{X}(s)\}\{X_i(t)-\bar{X}(t)\}$$





Square Integrable Functions

- A function $f(t) : T \to \mathbb{R}$ is said to be *square integrable*, $f \in L^2$ if

$$\int_T f(t)^2 dt < \infty$$

- Let $f, g \in L^2$, the **inner product** $< f, g >$ is defined as

$$\langle f, g \rangle := \int_T f(t)g(t)dt$$

- The **inner product norm** and **inner product distance**

$$||f|| = \left\{ \int_t f(t)^2 dt \right\}^{1/2}; \quad d(f,g) = ||f - g|| = \left[\int_t \{f(t) - g(t)\}^2 \right]^{1/2}$$

- $L^2[T]$ is a Hilbert Space

Random Variables in a General Metric Space S

- Let (Ω, B, P) be a probability space with sample space Ω, σ-algebra B and probability measure P
- Let S be a separable metric space
- We say that the mapping $X : \Omega \to S$ is a random element in S if $X^{-1}(A) \in B$ for all Borel sets A.

Examples

1. $S = \mathbb{R}$, $X : \Omega \to \mathbb{R}$ is a random variable
2. $S = \mathbb{R}^k$, $X : \Omega \to \mathbb{R}^k$ is a random vector
3. $S = L^2[0,1]$, $X : \Omega \to L^2[0,1]$ is a random function

- A random element in $X : \Omega \to S$ induces a probability measure on S defined by

$$\pi(A) = P\{X^{-1}(A)\} = P\{\omega \in \Omega : X(\omega) \in A\} = P(X \in A)$$

- We call π the distribution of X

Example: Random Variable in $L^2[0,1]$

Consider $\Omega = \{0, 1, 2\}$,
$B = \{\phi, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\}$, and P, s.t.
$P(\omega = 0) = P(\omega = 1) = P(\omega = 2) = 1/3$

- Define $X : \Omega \to L^2[0,1]$ as

$$X(t, \omega) = X(t) = X = \begin{cases} 0 & \mathit{if} \quad \omega = 0 \\ \sin(2\pi t) & \mathit{if} \quad \omega = 1 \\ \cos(2\pi t) & \mathit{if} \quad \omega = 2 \end{cases}$$

- $X : \Omega \to L^2[0,1]$ is a random variable in $L^2[0,1]$

Expectation in $L^2[0,1]$

- A random variable $X : \Omega \to H = L^2[0,1]$ is said to be **integrable** if

$$E||X|| = E\left[\left\{\int X(t)^2 dt\right\}^{1/2}\right] < \infty$$

- A random variable $X : \Omega \to H$ is said to be **square integrable** if

$$E||X||^2 = E\left[\int X(t)^2 dt\right] < \infty$$

- *Def.* **Expectation** of X, $E[X] = \mu$. If X is square integrable, then there exist a unique $\mu \in L^2$ s.t.

$$E[\langle f, X\rangle] = \langle f, \mu\rangle$$

for any $f \in L^2$. It follows that $\mu(t) = E[X(t)]$ almost everywhere.

Covariance in $L^2[0,1]$

- Without loss of generality assume $X : \Omega \to L^2[0,1]$ be a square integrable random function with $E(X) = 0$
- *Def.* The **covariance operator** of X, C_X is defined by

$$C_X(f) = E[\langle X, f\rangle X] = E\left[\left\{\int X(s)f(s)ds\right\} X(t)\right] = \int c_X(t,s)f(s)\,ds$$

  for any $f \in L^2[0,1]$
- C_X is completely determined by the **covariance function**

$$c_X(t,s) = E\{X(t)X(s)\}, \quad s,t \in [0,1]$$

- C_X is a covariance operator iff

1. C_X is symmetric, i.e. $\langle C_X(f), y\rangle = \langle f, C_X(y)\rangle$ for any $f, y \in L^2[0,1]$
2. C_X is nnd, i.e. $\langle C_X(f), f\rangle \geq 0$, for any $f \in L^2[0,1]$
3. The eigenvalues of C_X satisfy $\sum_{j=1}^{\infty} \lambda_j < \infty$

Statistical Summaries

Consider $X_1(t), \ldots, X_n(t)$ iid $t \in [0,1]$

- **Sample mean function**

$$\hat{\mu}(t) = \frac{1}{n}\sum_{i=1}^{n} X_i(t)$$

- **Sample covariance surface**

$$\hat{c}_X(s,t) = \frac{1}{n-1}\sum_{i=1}^{n}\{X(s) - \hat{\mu}(s)\}\{X(t) - \hat{\mu}(t)\}$$

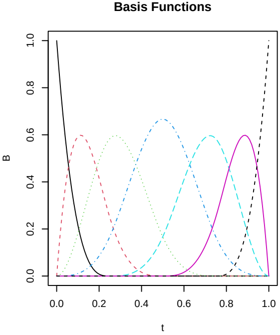
- **Sample covariance operator**

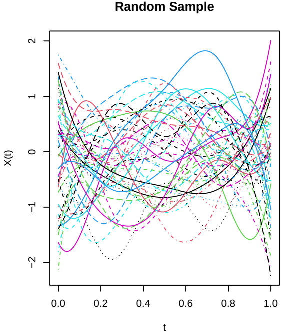
$$\hat{C}_X(t) = \frac{1}{n-1}\sum_{i=1}^{n}\langle X_i(t) - \hat{\mu}(t), x\rangle\{X(t) - \hat{\mu}(t)\}$$

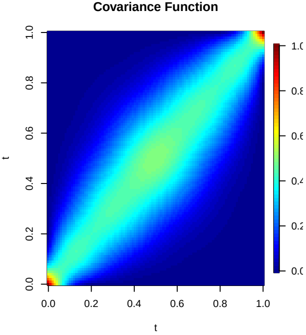
Example: Gaussian Process

- Let $X(t) = B(t)^T\beta$, with $B(t) \in \mathbb{R}^k$ known set of basis functions, and $\beta \sim N(0, I_p)$.
- With $E(X) = 0$, and $c_X\{X(s), X(t)\} = B(s)^T B(t)$

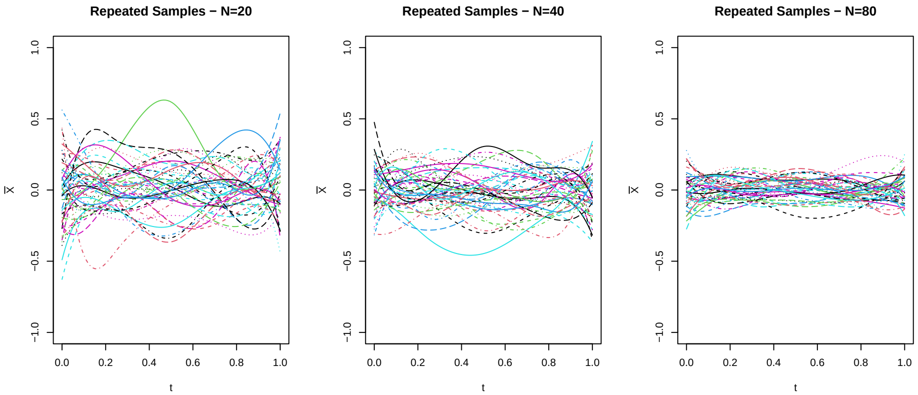
\vskip .5in







Behavior of sample mean function



Convergence of Random Functions

Let $\{X_n\}_{n\geq 1}$ be a sequence of random functions in $L^2[0,1]$, and X be a random function in $L^2[0,1]$.

- *Def.* $\{X_n\}_{n\geq 1}$ **converges in probability** to X, $X_n \to_p X$, if for any $\epsilon > 0$,

  $$P\{\omega \in \Omega : d(X_n(\omega), X(\omega)) > \epsilon\} = P\{d(X_n, X) > \epsilon\} \to 0$$

  Equivalently

  $$\lim_{n\to\infty} P\{d(X_n, X) > \epsilon\} = 0$$

- *Def.* $\{X_n\}_{n\geq 1}$ **converges in distribution** to X, $X_n \to_d X$, if the distribution π_n of X_n converges weakly to the distribution π of X, i.e.

  $$\int g \, d\pi_n \to \int f \, d\pi$$

  for any bounded continuous real function on $[0,1]$.

Consistency of the Sample Mean and Covariance

If $X_1, \dots, X_n$ are iid in L^2 and have the same distribution of X, assumed to be square integrable, with expectation μ and covariance c_X

- *Thm.* **Consistency of the sample mean function $\hat{\mu}$**

$$E(\hat{\mu}) = \mu \quad \text{and} \quad E||\hat{\mu} - \mu||^2 = O(n^{-1})$$

- *Thm.* **Consistency of the sample covariance function $\hat{c}_X$**

$$E\left\{\int [\hat{c}_X(s,t) - c_X(s,t)]\, ds\, dt\right\} \le n^{-1} E||X||^4.$$

  Therefore $\hat{c}_X$ is an L^2-consistent estimator of the covariance function, provided $E||X||^4 < \infty$

Convergence in Distribution: Gaussian Functions

Let $X : \Omega \to H$, with H separable Hilbert space (e.g. $H = L^2$)

- *Def.* The **Characteristic Functional** of a random function X is defined by

$$\varphi_X(f) = E\{\exp(i \langle f, X \rangle)\}$$

  for any $f \in H$

- *Def.* A random function X is said to be **Gaussian** if it characteristic function has the form

$$\varphi_X(f) = \exp\left\{ i \langle \mu, f \rangle - \frac{1}{2} \langle C_X(f), f \rangle \right\}$$

  where μ is the expectation and $C_X(f)$ is the covariance operator.

- **Thm.** A random function is Gaussian with $\mu = 0$ (wlog.) iff $\langle f, X \rangle \sim N(\cdot, \cdot)$, for any $f \in H$

Convergence in Distribution: CLT in Hilbert Space

Consider $X_1, \ldots, X_2$ iid withe the same distribution as $X : \Omega \to H$, square integrable with expected value μ and covariance operator C_X

- **Thm.** The sequence of random functions

$$n^{-1/2} \sum_{i=1}^{n} (X_i - \mu) \to_d Z$$

where Z is a Gaussian random function with $E(Z) = 0$ and covariance operator C_X.

Asymptotic Normality of Sample Summaries

Consider $X_1, \ldots, X_2$ iid withe the same distribution as $X : \Omega \to H$, square integrable ($E||X||^2 < \infty$) with expected value μ and covariance operator C_X

- The sample mean $\hat{\mu}$ is asymptotically normal, i.e.

$$n^{1/2}(\hat{\mu} - \mu) \to_d N(0, C), \text{ in } H$$

  where $C = E[(X - \mu) \otimes (X - \mu)]$

- If $E||X||^4 < \infty$, then

$$n^{1/2}(\hat{C}_X - C_X) \to_d N(0, \Gamma), \text{ in } S$$

  where $\Gamma = E\left[\{(X - \mu) \otimes (X - \mu) - C\} \otimes \{(X - \mu) \otimes (X - \mu) - C\}\right]$

Thank You!

Questions? Comments?