Un Contraste de Hipótesis para datos funcionales (Sección Introductoria)

Cristhian Leonardo Urbano Leon
Javier Olaya Ochoa.

Universidad del Valle

Escuela de Estadística
Santiago de Cali

Mayo 2018

Motivación

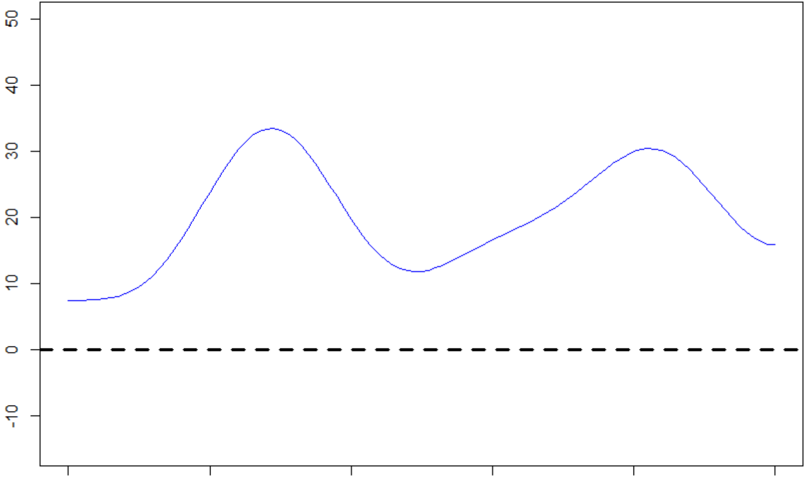
Cuando trabajamos con datos escalares, la comparación de medias de dos poblaciones, es un tema ampliamente estudiado. En la literatura se encuentran gran cantidad de variantes, a tal grado, que podríamos considerar el problema ya resuelto.

Motivación

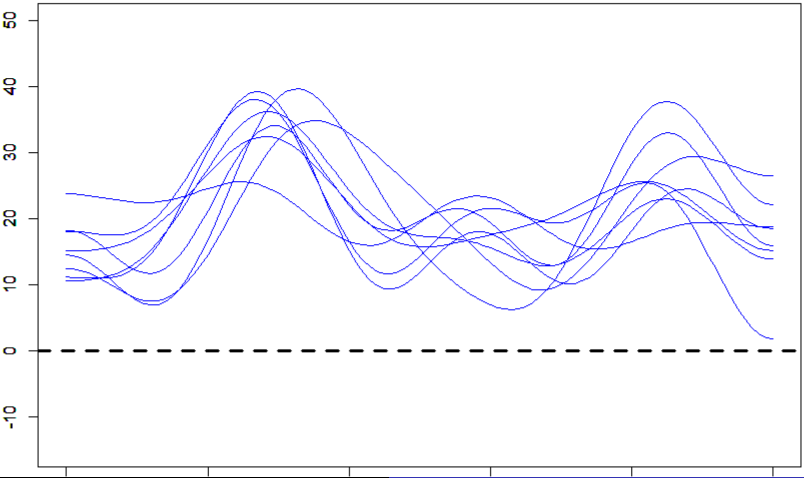
Cuando trabajamos con datos escalares, la comparación de medias de dos poblaciones, es un tema ampliamente estudiado. En la literatura se encuentran gran cantidad de variantes, a tal grado, que podríamos considerar el problema ya resuelto.

Sin embargo, en los últimos años se ha abierto camino una nueva alternativa de representación y tratamiento de datos llamada Análisis de Datos Funcionales, en donde los objetos de estudio son curvas

Motivación



Motivación



Motivación

En este nuevo enfoque:

Motivación

En este nuevo enfoque:

- Se parte de los datos escalares para construir curvas suaves valiéndose de técnicas modificadas del análisis numérico

Motivación

En este nuevo enfoque:

- Se parte de los datos escalares para construir curvas suaves valiéndose de técnicas modificadas del análisis numérico
- Cada curva se llama dato Funcional

Motivación

En este nuevo enfoque:

- Se parte de los datos escalares para construir curvas suaves valiéndose de técnicas modificadas del análisis numérico
- Cada curva se llama dato Funcional
- El dato funcional en una función continua

Motivación

En este nuevo enfoque:

- Se parte de los datos escalares para construir curvas suaves valiéndose de técnicas modificadas del análisis numérico
- Cada curva se llama dato Funcional
- El dato funcional en una función continua
- Los objetos de estudio son curvas de funciones, por lo cual es posible trabajar dentro de un espacio funcional

Motivación

Supongamos que se quiere, constatar la utilidad de dos estaciones encargadas de monitorear la calidad del aire bajo las siguientes consideraciones:

Motivación

Supongamos que se quiere, constatar la utilidad de dos estaciones encargadas de monitorear la calidad del aire bajo las siguientes consideraciones:

- Las estaciones están ubicadas dentro de una misma ciudad.
- Cada estación toma mediciones sobre la presencia de material particulado en el aire que tienen un diámetro menor a 2.5 micrómetros cada 10 segundos y reporta la media cada hora.

Motivación

Se tiene que:

- Si se considera el subconjunto de datos escalares correspondiente a un día, para la construcción de una curva, se obtiene el dato funcional de ese día.

Motivación

Se tiene que:

- Si se considera el subconjunto de datos escalares correspondiente a un día, para la construcción de una curva, se obtiene el dato funcional de ese día.
- Los datos funcionales diarios construidos para una misma estación pueden no ser independientes ya que provienen de un proceso estocástico.

Motivación

Se tiene que:

- Si se considera el subconjunto de datos escalares correspondiente a un día, para la construcción de una curva, se obtiene el dato funcional de ese día.
- Los datos funcionales diarios construidos para una misma estación pueden no ser independientes ya que provienen de un proceso estocástico.
- Si las estaciones se encuentran muy cerca entre si, al tratarse de mediciones de un contaminante presente en el aire, se espera que los datos de las dos estaciones presenten alguna estructura de dependencia espacial.

El Problema

Si se toma un conjunto de datos funcionales con una unidad de medida que provea una estructura de dependencia entre ellos, como por ejemplo, datos de dos estaciones de vigilancia de la calidad del aire para contaminación por partículas con un diámetro menor a 2.5 micrometros, que se encuentren lo suficientemente cerca y que reporten las mediciones a la misma hora y fecha.

El Problema

Si se toma un conjunto de datos funcionales con una unidad de medida que provea una estructura de dependencia entre ellos, como por ejemplo, datos de dos estaciones de vigilancia de la calidad del aire para contaminación por partículas con un diámetro menor a 2.5 micrometros, que se encuentren lo suficientemente cerca y que reporten las mediciones a la misma hora y fecha.

Entonces los datos funcionales diarios de cada estación pueden ser tomados como dos muestras que se encuentran emparejadas.

Antecedentes

El problema de comparar estadisticamente curvas provenientes de datos funcionales, es un problema que se ha estado atacando recientemente.

Febrero et al(2004)

Un primer antecedente teórico para comparar estadisticamente dos curvas se debe a **[Febrero et al. (2004)]**. Aquí se propone una prueba para la igualdad de medias de dos muestras consideradas independientes, tomadas cada una de dos poblaciones.

Febrero et al(2004)

- Este método se es similar a una prueba tipo F en datos escalares, extendida para la comparación de los dos medias funcionales

Febrero et al(2004)

- Este método se es similar a una prueba tipo F en datos escalares, extendida para la comparación de los dos medias funcionales
- Los autores proponen un procedimiento numérico de Monte Carlo para observar la distribución asintótica de la estadística de prueba.

Febrero et al(2004)

- Este método se es similar a una prueba tipo F en datos escalares, extendida para la comparación de los dos medias funcionales
- Los autores proponen un procedimiento numérico de Monte Carlo para observar la distribución asintótica de la estadística de prueba.
- Se generan una serie de muestras independientes de cada familia.

Febrero et al(2004)

- Este método se es similar a una prueba tipo F en datos escalares, extendida para la comparación de los dos medias funcionales
- Los autores proponen un procedimiento numérico de Monte Carlo para observar la distribución asintótica de la estadística de prueba.
- Se generan una serie de muestras independientes de cada familia.
- El valor calculado del estadístico de prueba bootstrap se compara con la distribución simulada y se informa un valor p

Febrero et al(2004)

La implementación practica para el programa estadístico R de este test se describe en un articulo posterior de Febrero-Bande y Oviedo de la Fuente en el año 2012 **[Febrero. B and De la Fuente. O.]**.

Degras (2011)

Un segundo acercamiento se debe a Degras en el año 2011 **[Degras, D. A.(2011)]**. El cual se implementa tres años más tarde por el mismo autor **[Degras, D.(2014)]**. En este método se utilizan bandas de confianza con las siguientes consideraciones:

Degras (2011)

- Como primera medida se calculan las debidas medias funcionales de las dos muestras extraídas de las dos poblaciones.

Degras (2011)

- Como primera medida se calculan las debidas medias funcionales de las dos muestras extraídas de las dos poblaciones.
- Se construyen bandas de confianza para una de los dos medias funcionales.

Degras (2011)

- Como primera medida se calculan las debidas medias funcionales de las dos muestras extraídas de las dos poblaciones.
- Se construyen bandas de confianza para una de los dos medias funcionales.
- se verifica si la otra media funcional está o no incluida por esas bandas.

Degras (2011)

- Utilizan dos tipos de bandas de confianza, una basada en normalidad asintótica y la otra fundamentada en un procedimiento de bootstrap.

Degras (2011)

- Utilizan dos tipos de bandas de confianza, una basada en normalidad asintótica y la otra fundamentada en un procedimiento de bootstrap.
- Finalmente también se informan un valor p

Degras (2011)

- Utilizan dos tipos de bandas de confianza, una basada en normalidad asintótica y la otra fundamentada en un procedimiento de bootstrap.
- Finalmente también se informan un valor p

La implementacion de este procedimiento para el programa estadístico R se describe en Degras (2014)

Cox (2015)

Uno de los ultimos avances se debe a Cox D, en el año 2015. Aquí se realizan las siguientes consideraciones.

Cox (2015)

Uno de los ultimos avances se debe a Cox D, en el año 2015. Aquí se realizan las siguientes consideraciones.

- La prueba propuesta se basa en una adaptación de la metodología de Neyman aplicada en el análisis de perfiles de dos grupos en el análisis multivariante.

Cox (2015)

Uno de los ultimos avances se debe a Cox D, en el año 2015. Aquí se realizan las siguientes consideraciones.

- La prueba propuesta se basa en una adaptación de la metodología de Neyman aplicada en el análisis de perfiles de dos grupos en el análisis multivariante.
- Se utiliza el estadístico T^2 de Hotelling como una generalizacion del estadistico T de student.

Cox (2015)

Uno de los ultimos avances se debe a Cox D, en el año 2015. Aquí se realizan las siguientes consideraciones.

- La prueba propuesta se basa en una adaptación de la metodología de Neyman aplicada en el análisis de perfiles de dos grupos en el análisis multivariante.
- Se utiliza el estadístico T^2 de Hotelling como una generalizacion del estadistico T de student.

Cox (2015)

- Se supone que las distribuciones de las dos poblaciones son las mismas.

Cox (2015)

- Se supone que las distribuciones de las dos poblaciones son las mismas.
- Proponen métodos de aleatorización para encontrar una distribución nula.

Cox (2015)

- Se supone que las distribuciones de las dos poblaciones son las mismas.
- Proponen métodos de aleatorización para encontrar una distribución nula.

Vale decir que a la fecha aun no se encuentra un procedimiento para la implementación de esta metodología en algún programa estadístico

Generalidades

Vale decir, que debido a la naturaleza de estos nuevos objetos de estudio, es necesario hacer uso de algunas herramientas propias de áreas especificas de la matemática tanto discreta y continua, como computacional.

- Álgebra Moderna: Estructuras algebraicas.
- Análisis Matemático: Espacios funcionales.
- Análisis Numérico: Métodos de interpolación (Modificados)

Grupo

La estructura básica y que sirve como soporte para todas las demás, se conoce como **Grupo** y se define como sigue [FRALEIGH, J. (1988)]

Grupo

La estructura básica y que sirve como soporte para todas las demás, se conoce como **Grupo** y se define como sigue [FRALEIGH, J. (1988)]:

Grupo

Sea G un conjunto no vació, y sea $\oplus$ una operación definida en G. Se dice que $(G, \oplus)$ es un **Grupo** si, y solo si :

- la operación $\oplus$ es cerrada en G
- $\oplus$ es Asociativa en G
- Existe un elemento neutro bajo $\oplus$ en G
- Existe un elemento simétrico para cada elemento de G
- Si la operación es conmutativa, el grupo se dice **Grupo Abeliano**.

Anillo

Con base a la estructura anterior definimos la siguiente [FRALEIGH, J. (1988)], [Landin, J. (2012)]:

Anillo:

Sea A un conjunto no vacio, y sean $\oplus$ y $\cdot$ dos operación definidas en A. Se dice que A es un **Anillo** si, y solo si:

- $(A, \oplus)$ es un grupo abeliano.
- $\cdot$ distribuye sobre $\oplus$.
- $\cdot$ es Asociativa en A.
- Si existe un elemento neutro bajo $\cdot$ en A, se dice **Anillo Unitario**.
- Si $\cdot$ es conmutativa, se dice **Anillo Conmutativo**.

Cuerpo

Sea $\mathbb{K}$ un conjunto no vacío, y sean $\oplus$ y $\cdot$ dos operación definidas en $\mathbb{K}$. Se dice que $(\mathbb{K}, \oplus, \cdot)$ es un **Cuerpo** si, y solo si : [Landin, J. (2012)]

Cuerpo:

- $(\mathbb{K}, \oplus, \cdot)$ es un anillo conmutativo.
- $(\mathbb{K} - \{0\}, \cdot)$ es un Grupo Abeliano.

Espacio Vectorial

Espacio Vectorial:

Sean $\mathcal{V}$ y $\mathbb{K}$ conjuntos no vacíos, y sean dos operaciones $\oplus$ Interna sobre $\mathcal{V}$ y $\cdot$ externa en $\mathcal{V}$. Se dice que $\mathcal{V}$ es un **Espacio Vectorial Sobre** $\mathbb{K}$ si, y solo si [Landin, J. (2012)]:

- $(\mathcal{V}, \oplus)$ es un grupo aveliano.
- $\mathbb{K}$ es un cuerpo.
- y se verifican propiedades de distribución y asociatividad entre $\mathcal{V}$ y $\mathbb{K}$

Observaciones

- Los espacios vectoriales no son únicos, existen gran variedad de estos, sobre los cuales puede definirse otras operaciones o funciones a fin de dotar de propiedades deseables a estas estructuras.

Observaciones

- Los espacios vectoriales no son únicos, existen gran variedad de estos, sobre los cuales puede definirse otras operaciones o funciones a fin de dotar de propiedades deseables a estas estructuras.
- Son necesarios para nuestro trabajo aquellos dotados de una estructura de medición. Se realizan las definiciones pertinentes tomando como base a Kolmogorov, W.(1957), [Rudin, W.(1979)] y [Landin, J. (2012)].

Definición

Espacio Normado: *Sea $\mathcal{V}$ un espacio vectorial sobre un cuerpo $\mathbb{K}$. Una norma en $\mathcal{V}$ es una función $\|\cdots\|$ de $\mathcal{V}$ en $\mathbb{R}_0^+$, tal que si $x, y \in \mathcal{V}$ y $\lambda \in \mathbb{K}$:*

- $\|x\| = 0 \Rightarrow x = 0$
- $\|\lambda x\| = |\lambda| \, \|x\|$
- $\|x + y\| \leq \|x\| + \|y\|$
- *Los espacios sobre los cuales se define una norma, se dicen* **Espacios Normados**

Espacio Métrico:

Una métrica (distancia), sobre un conjunto A es una función $m :\longrightarrow \mathbb{R}_0^+$ tal que para $x, y \in A$:

- $m(x) \geq 0$
- $m(x, y) = 0 \Leftrightarrow x = y$
- $m(x, y) = m(y, x)$
- $m(x + y) \leq m(x) + m(y)$
- Si sobre el conjunto A se ha definido dicha función, entonces A se denomina **Espacios Metrico**

Definición

Espacio Interior: *Sea $\mathcal{V}$ un espacio vectorial sobre un cuerpo $\mathbb{K}$. Un producto interno sobre $\mathcal{V}$ es una función $\phi : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ tal que:*

- $\phi(v + w, z) = \phi(v, z) + \phi(w, z)$ *y* $\phi(\alpha \cdot v, z) = \alpha \cdot \phi(v, z)$
- $\phi(v, w) = \phi(w, v)$
- $\phi(v, v) \geq 0$
- *Un espacio Vectorial sobre el cual se ha definido un producto interno se llama* **Espacio Interior**

Definición

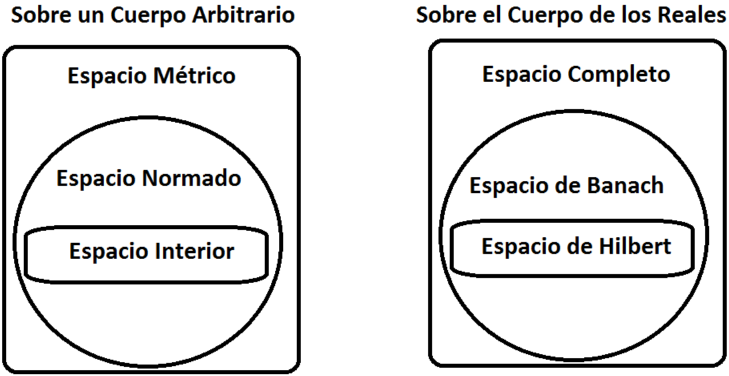
Espacios de Hilbert y Banach

- *Un espacio de Hilbert, es un espacio vectorial métrico, normado con producto interno sobre el cuerpo de los numeros reales* $\mathbb{R}$.
- *Un espacio de Banach es un espacio metrico normado igualmente sobre* $\mathbb{R}$.

Observación

Todo producto interno induce una norma dentro de un espacio vectorial y esta a su vez define una métrica sobre el mismo espacio [Kolmogorov et al., 1957]. Esta relación entre funciones permite relacionar también los diferentes espacios que las contienen

Relación existente entre los diferentes espacios



Ya con la teoría definida anteriormente, nos adentramos ahora en un espacio mas puntual para nuestro estudio.

Ya con la teoría definida anteriormente, nos adentramos ahora en un espacio mas puntual para nuestro estudio.

Definición

Espacio Funcional
Un espacio funcional es una estructura algebraica en la cual sus elementos son funciones. **Rudin, W., 1979** *[3]*

Definición

El espacio de Hilbert L_2: [Rudin, W., 1979] *El conjunto de funciones cuadrado integrables L_2 en el intervalo cerrado $[a, b]$, es un espacio vectorial normado sobre el cuerpo $\mathbb{R}$ de los números reales y que posee producto interno. A saber:*

- *Producto interno:* $\langle f_1, f_2 \rangle = \int_b^a f_1(x) f_2(x) dx$
- *Norma:* $\|f\| = \langle f, f \rangle^{1/2} = (\int_b^a f^2(x) dx)^{1/2}$
- *La norma definida induce una métrica en* L_2
- *Con lo anterior,* L_2 *es un espacio de Hilbert*

Observaciones

- Si bien un dato funcional es en si una función continua, la forma de obtener este objeto es mediante suavización de mediciones discretas, por tal motivo se hace necesario tener alguna forma de generar una representación de la función a partir de dichas mediciones.[Ramsay, J.O and Silverman, B (1997)]
- Como los datos funcionales están soportado sobre la teoría de espacios funcionales que como ya se vio, no son mas que simples espacios vectoriales, estos poseen una forma de representación.

Definición

Sistema de Generadores:

- *Sean $\mathcal{V}$ un espacio vectorial sobre $\mathbb{K}$ y sea $S \subseteq \mathcal{V}$. Se dice que S es un conjunto generador de $\mathcal{V}$ si todo elemento de $\mathcal{V}$ se puede escribir como combinación lineal de los elementos de S.*
- *Un espacio vectorial se denomina Finito, si puede ser generado por un conjunto finito de elementos.*
- *Un conjunto $B \subseteq \mathcal{V}$ es una base para $\mathcal{V}$ si B es un conjunto generador y además es linealmente independiente.*
- *Si un espacio se puede representar por medio de un numero finito de elementos se dice que tiene dimensión finita, de no ser así, sera de*

Observaciones

- El espacio funcional de Hilbert L_2 es un espacio de dimensión infinita.

Observaciones

- El espacio funcional de Hilbert L_2 es un espacio de dimensión infinita.
- En la practica, se trabaja siempre, en un subespacio de dimensión finita de este.

Observaciones

- El espacio funcional de Hilbert L_2 es un espacio de dimensión infinita.
- En la practica, se trabaja siempre, en un subespacio de dimensión finita de este.
- Basta ahora escoger una base para dicha representación

Observaciones

- El espacio funcional de Hilbert L_2 es un espacio de dimensión infinita.
- En la practica, se trabaja siempre, en un subespacio de dimensión finita de este.
- Basta ahora escoger una base para dicha representación
- Dependiendo de la base escogida, se trabaja en uno u otro subespacio de dimensión finita

Observaciones

- El espacio funcional de Hilbert L_2 es un espacio de dimensión infinita.
- En la practica, se trabaja siempre, en un subespacio de dimensión finita de este.
- Basta ahora escoger una base para dicha representación
- Dependiendo de la base escogida, se trabaja en uno u otro subespacio de dimensión finita
- Existen en general muchas bases para el espacio L_2. Sin embargo, para su aplicación en datos funcionales es común utilizar solo un puñado de ellas

Base de Fourier

Un subespacio vectorial del espacio de Hilbert $L_2(I)$ puede ser generado por el conjunto de elementos:

$$\sqrt{\frac{1}{T}}, \sqrt{\frac{2}{T}} Sen\left(\frac{2N\pi}{T}t\right), \sqrt{\frac{2}{T}} Cos\left(\frac{2N\pi}{T}t\right)$$

Para $N = 1, 2, ..., k$, con$T$ la medida de I
Dicho conjunto constituye una base finita ortonormal para un subespacio de $L_2(I)$ **[Ferraty and Vieu, 2006]** [4]

Base B-Splines

Una base B-spline, es una base de estructura polinomial fundamentada en polinomios de menor grado llamados **polinomios de Berstein**, definidos como sigue para grado n [Ferraty and Vieu, 2006] [4]:

$$B_n = [(1-x)+x]^n$$

De esta manera, los elementos de la base tienen la forma:

$$S(t) = \sum_{m+L-1}^{k=1} c_k B_k(t,\tau)$$

Donde τ son los nodos donde se unirán las funciones, $(L-1)$ el numero de nodos interiores y m el orden del polinomio.

Base Spline de Grado 2

Si se quiere generar una función spline de grado 2 utilizamos los términos de la expansión binomial:

Base Spline de Grado 2

Si se quiere generar una función spline de grado 2 utilizamos los términos de la expansión binomial:

$$(x + (1 - x))^2 = x^2 + 2x(1 - x) + (1 - x)^2$$

Base Spline de Grado 2

Si se quiere generar una función spline de grado 2 utilizamos los términos de la expansión binomial:

$$(x + (1 - x))^2 = x^2 + 2x(1 - x) + (1 - x)^2$$

De esta manera, se puede representar polinomios de grado menor o igual que 2, utilizando 3 polinomios. A saber:

Base Spline de Grado 2

Si se quiere generar una función spline de grado 2 utilizamos los términos de la expansión binomial:

$$(x + (1 - x))^2 = x^2 + 2x(1 - x) + (1 - x)^2$$

De esta manera, se puede representar polinomios de grado menor o igual que 2, utilizando 3 polinomios. A saber:

- $B_2^2 = x^2$

Base Spline de Grado 2

Si se quiere generar una función spline de grado 2 utilizamos los términos de la expansión binomial:

$$(x + (1 - x))^2 = x^2 + 2x(1 - x) + (1 - x)^2$$

De esta manera, se puede representar polinomios de grado menor o igual que 2, utilizando 3 polinomios. A saber:

- $B_2^2 = x^2$
- $B_1^2 = 2x(1 - x) = -2x^2 + 2x$

Base Spline de Grado 2

Si se quiere generar una función spline de grado 2 utilizamos los términos de la expansión binomial:

$$(x+(1-x))^2 = x^2 + 2x(1-x) + (1-x)^2$$

De esta manera, se puede representar polinomios de grado menor o igual que 2, utilizando 3 polinomios. A saber:

- $B_2^2 = x^2$
- $B_1^2 = 2x(1-x) = -2x^2 + 2x$
- $B_0^2 = (1-x)^2 = x^2 - 2x + 1$

Base Spline

Note que: $x^2 + \frac{1}{2}(-2x^2 + 2x) = x$. Es decir:

Base Spline

Note que: $x^2 + \frac{1}{2}(-2x^2 + 2x) = x$. Es decir:

$$x = 1B_2^2 + \frac{1}{2}B_1^2 + 0B_0^2$$

Base Spline

Note que: $x^2 + \frac{1}{2}(-2x^2 + 2x) = x$. Es decir:

$$x = 1B_2^2 + \frac{1}{2}B_1^2 + 0B_0^2$$

Los coeficientes de representación del polinomio $p(x) = x$ son

$$(1, \frac{1}{2}, 0)$$

Base de Fourier

Una de las bases de Fourier más utilizada tiene como componente principal la función Coseno. Sin embargo esta función puede ser representada por medio de polinomios de Chebyshev. los cuales son definidos de manera recurrente como:

Base de Fourier

Una de las bases de Fourier más utilizada tiene como componente principal la función Coseno. Sin embargo esta función puede ser representada por medio de polinomios de Chebyshev. los cuales son definidos de manera recurrente como:

- $T_0(x) = 1$

Base de Fourier

Una de las bases de Fourier más utilizada tiene como componente principal la función Coseno. Sin embargo esta función puede ser representada por medio de polinomios de Chebyshev. los cuales son definidos de manera recurrente como:

- $T_0(x) = 1$
- $T_1(x) = x$

Base de Fourier

Una de las bases de Fourier más utilizada tiene como componente principal la función Coseno. Sin embargo esta función puede ser representada por medio de polinomios de Chebyshev. los cuales son definidos de manera recurrente como:

- $T_0(x) = 1$
- $T_1(x) = x$
- $T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$

Base de Fourier

De esta manera. los cosenos podrán ser descritos como sigue:

Base de Fourier

De esta manera. los cosenos podrán ser descritos como sigue:

- $Cos(2\theta) = T_2(x) = 2x^2 - 1$

Base de Fourier

De esta manera. los cosenos podrán ser descritos como sigue:

- $Cos(2\theta) = T_2(x) = 2x^2 - 1$
- $Cos(3\theta) = T_3(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x$

Base de Fourier

De esta manera. los cosenos podrán ser descritos como sigue:

- $Cos(2\theta) = T_2(x) = 2x^2 - 1$
- $Cos(3\theta) = T_3(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x$
- $Cos(4\theta) = T_4(x) = 8x^4 - 8x^2 + 1$

Base de Fourier

De esta manera. los cosenos podrán ser descritos como sigue:

- $Cos(2\theta) = T_2(x) = 2x^2 - 1$
- $Cos(3\theta) = T_3(x) = 2x(2x^2 - 1) - x = 4x^3 - 3x$
- $Cos(4\theta) = T_4(x) = 8x^4 - 8x^2 + 1$
- $Cos(5\theta) = T_5(x) = 16x^5 - 20x^3 + 5x$

Medidas en datos funcionales

Las herramientas utilizadas para el análisis descriptivo de datos escalares son validas para datos funcionales gracias a una extensión de los conceptos como sigue según [Ramsay, Silverman (1997)]. Dado un conjunto de datos funcionales $\mathcal{X}_1, \mathcal{X}_2, ..., \mathcal{X}_n$ definidos en $t \in \mathbb{R}$, las funciones descriptivas se definen como:

Medidas en datos funcionales

- **Media:** $\bar{\mathcal{X}}(t) = n^{-1} \sum_{i=1}^{n} \mathcal{X}_i(t)$
- **Varianza:** $Var(\mathcal{X})(t) = (N-1)^{-1} \sum_{i=1}^{n} (\mathcal{X}_i(t) - \bar{\mathcal{X}}(t))$
- **Desviación Estándar:** $Dev(\mathcal{X}(t)) = \sqrt{Var(\mathcal{X}(t))}$
- **Covarianza:** $Cov(\mathcal{X}_i(t), \mathcal{X}_j(t)) = (N-1)^{-1} \sum_{i=1}^{n} (\mathcal{X}_i(t) - \bar{\mathcal{X}}(t))(\mathcal{X}_i(t) - \bar{\mathcal{X}}(t))$
- **Correlación:** $Corr(\mathcal{X}_i(t), \mathcal{X}_j(t)) = \frac{Cov(\mathcal{X}_i(t), \mathcal{X}_j(t))}{\sqrt{Var(\mathcal{X}(t)) Var(\mathcal{X}(t))}}$

Gracias!

- Cox, D. D., Lee, J. S. and Follen, M. *A two sample test for functional data.*, Communications for Statistical Applications and Methods. 22:121-135, 2015.
- Degras, D. *SCBmeanfd: Simultaneous Confidence Bands for the Mean of Functional Data.*, R package version 1.1, 2014.
- Febrero. B and De la Fuente. O. *Statistical computing in functional data analysis: The R package fda.usc*, Journal of Statistical Software. 51(4):1-28, 2012.
- Landin. *JAn Introduction to Algebraic Structures*, Dover Books on Mathematics Series. 2012.
- Degras, D. A. *Simultaneous confidence bands for nonparametric regression with functional data*, Statistica Sinica. 21:1735-1765, 2011. (2010).

- Ramsay, J. O., Graves, S. and Hooker, G. *Functional Data Analysis with R and MATLAB*, Springer. 2010.
- Zhang, J and Chen, J. *Statical Inferences For Functional Data*, The Annals of Statistics, Vol. 35, No. 3, 2007.
- Cuesta-Albertos, J and Febrero-Bande, M. *A simple multiway anova for funcional data*. Business and Economic,19:537-557, 2007.
- Ferraty, F. and Vieu, P. *Nonparametric Functional Data Analysis Theory and Practice*, Springer. 2006.
- Ramsay, J. O. and Silverman, B. W. *Functional Data Analysis, 2nd. ed*, Springer. 2005.

- Cuevas, A. Febrero, M. and Fraiman, R. *A anova test for functional data*, Computational Statistics and Data Analysis, 47:111-122, 2004.
- Shen, Q. and Faraway, J. *An f test for linear models whit functional responses*, Statistica Sinica, 14:1239-1257, 2004.
- Febrero, M., Cuevas, A. and Fraiman, R. *An anova test for functional data*, Computational Statistics and Data Analysis. 47:111-122, 2002.
- Fan, J. and Lin, S. *Test of significance when data are curves*, J. Amer Statistics. Assoc. 93:111-122, 1998.
- Ramsay, J.O and Silverman, B.W. *Funcional Data Analysis*, Springer, New York, 1997.

- Ramsay, J. O. and Dalzell, C. J. *Some tools for functional data analysis*, Journal of the Royal Statistical Society, 3:593-572, 1991.
- FRALEIGH, J.*Álgebra Abstracta*, Addison-Wesley Iberoamericana, S.A. Wilmington, Delaware, E.U.A.1988.
- Rudin, W.*Análisis Funcional*, Reverte. 1979.
- HERSTEIN, I.N*Álgebra Moderna.* Trillas. México. 1970.
- Kolmogorov, A. N., Fomin, S. V., and Fomin, S. *Elements of the theory of functions and functional analysis: Volumen I METRIC AND NORMED SPACES.*, Graylock Press. 1957.