

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA.
ESCUELA DE MATEMÁTICA



TRABAJO DE GRADO TITULADO:

ESTUDIO DE LA GEOMETRÍA FRACTAL CON
APLICACIONES A FINANZAS Y VULCANOLOGÍA

PRESENTADO POR:

BR. MILTON ARNOLDO RIVAS MORALES

PARA OPTAR AL TÍTULO DE:

LICENCIADO EN MATEMÁTICA

CIUDAD UNIVERSITARIA, NOVIEMBRE 2,017

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA.
ESCUELA DE MATEMÁTICA



TRABAJO DE GRADO TITULADO:

**ESTUDIO DE LA GEOMETRÍA FRACTAL CON
APLICACIONES A FINANZAS Y VULCANOLOGÍA**

PRESENTADO POR:

BR. MILTON ARNOLDO RIVAS MORALES

PARA OPTAR AL TÍTULO DE:

LICENCIADO EN MATEMÁTICA

DOCENTE ASESOR:

MSC. WALTER OTONIEL CAMPOS GRANADOS

DOCENTE ESCUELA DE MATEMÁTICA

CIUDAD UNIVERSITARIA, NOVIEMBRE 2,017

AUTORIDADES

UNIVERSIDAD DE EL SALVADOR

RECTOR:

MSc. ROGER ARMANDO ARIAS ALVARADO

SECRETARIO GENERAL:

LIC. CRISTÓBAL HERNÁN RÍOS BENÍTEZ

**FACULTAD DE CIENCIAS NATURALES Y
MATEMÁTICA**

DECANO:

LIC. MAURICIO HERNÁN LOVO

SECRETARIA:

LICDA. DAMARIS MELANY HERRERA TURCIOS

ESCUELA DE MATEMÁTICA

DIRECTOR:

DR. JOSÉ NERYS FUNES TORRES

SECRETARIA:

MSc. ALBA IDALIA CÓRDOVA CUÉLLAR

Índice general

PRELIMINARES	5
0.1. NOCIONES DE PROBABILIDAD	7
0.2. AXIOMAS DE LA TEORÍA DE PROBABILIDAD.	7
0.3. FUNCIÓN DE PROBABILIDAD.	8
0.4. ESPACIO FRACTAL Y GENERACIÓN DE FRACTALES	8
1. FRACTALES Y FINANZAS: DESARROLLO DEL MÉTODO	13
1.1. CONTRACCIONES EN EL ESPACIO DE FUNCIÓN DE DISTRIBUCIÓN	14
1.2. OPERADORES IFS	16
1.3. ESTIMACIÓN Y APLICACIONES DE LA FUNCIÓN DE DISTRIBUCIÓN	24
1.4. Propiedades asintóticas	26
2. APLICACIÓN DE FRACTALES A FINANZAS	27
3. FRACTALES Y VULCANOLOGÍA	52
3.1. El concepto de dimensión de recuento de cajas.	52
3.2. ENFOQUE DE LA TRANSFORMADA DE FOURIER.	64
3.3. APLICACIÓN A VULCANOLOGÍA	70
A. BASE DE DATOS PARA FINANZAS	78
B. BASE DE DATOS PARA VULCANOLOGÍA	79

INTRODUCCIÓN

En el presente trabajo se dará conocer la vinculación que existe entre la geometría fractal con la finanzas y vulcanología poniendo así en evidencia la forma en como la matemática está íntimamente relacionada con la realidad.

Históricamente los orígenes de la geometría fractal se remontan a finales del siglo *XIX* y comienzos del siglo *XX* con la aparición en el campo de la matemática de conjuntos geométricos de propiedades aparentemente paradójicas.

En dichos conjuntos (curvas de Peano y Koch, conjunto de Cantor, ...) parecía existir una discordancia entre su tamaño real y su configuración espacial como conjunto de puntos (curvas con área o con longitud infinita entre dos de sus puntos, etc...).

Los ejemplos plantearon la necesidad de establecer una medida adecuada de su tamaño, por una parte, y por otra el estudio de su forma o propiedades geométricas. De esta forma se inició una nueva parcela del análisis matemático que hoy denominamos **Teoría Geométrica de la Medida**, que tuvo su punto de arranque con la definición del concepto de **dimensión Hausdorff**, concepto que establecía la distinción del tamaño de los conjuntos paradójicos y que sentó sus bases con los trabajos de Besicovitch durante los años 20 y 30, en los que con particular ingenio estudio las propiedades geométricas de los conjuntos planos que él llamo **irregulares**: eran el arquetipo de lo que hoy llamamos fractales, y sus trabajos fueron el sólido cimiento de lo que hoy llamamos Geometría Fractal.

A partir de entonces, matemáticos como H.Federer, J.M.Marstrand, R.O. Davies, K.J Falconer, P.Matila, etc continúan el estudio de este tipo de conjuntos configurándose la teoría iniciada por Besicovitch.

En 1977, B.Mandelbrot, en su famosa obra **The Fractal Geometry of Nature**, además de acuñar el termino fractal defiende una idea que se convertiría con el tiempo en la razón del crecimiento exponencial de las aplicaciones de estos y de la actual popularización del término: las formas la naturaleza son fractales y múltiples procesos de la misma se rigen por comportamientos fractales.

En 1981 el australiano Hutchinson, publica un trabajo que a la larga se convertiría en uno de los mas citados en la literatura especializada. En el se desarrolla un concepto básico en geometría fractal: el de conjunto autosemejante, el cual ha tenido una gran importancia en el desarrollo posterior de la misma.

El presente trabajo estará estructurado en cuatro capítulos, desarrollados como se detalla a continuación:

En el capítulo uno se desarrolla el artículo: APPROXIMATING DISTRIBUTION FUNCTIONS BY ITERATED FUNCTION SYSTEMS.

En el capítulo dos aplicamos la teoría desarrollada en el capítulo uno para analizar una base de datos de una entidad financiera y encontrar el valor operacional de la pérdidas totales y hacer predicciones sobre la entidad.

En el capítulo tres se desarrolla los fundamentos matemáticos de la geometría fractal que nos permitirán desarrollar una aplicación en vulcanología, a una situación ideal, herramienta que nos sirve para comprender un los datos en el caso real. En este trabajo, realizamos a una aproximación de la dimensión de caja de la serie temporal de emisiones de CO_2 en el año 2005 en el Volcán de San Salvador.

En el desarrollo de los dos capítulos de aplicaciones es necesario utilizar software, que nos permita hacer cálculos de una manera más eficiente, en este sentido nos auxiliaremos de el software R.

PRELIMINARES.

A lo largo de la historia, la industria bancaria ha sufrido importantes pérdidas por fallos operacionales. Consciente de ello, el Comité de Basilea publicó, en 2004, un Nuevo Acuerdo de capital en el que instaba a las entidades financieras medir, controlar y gestionar su riesgo operacional. En este contexto el valor en Riesgo (VaR) se convierte en un elemento crucial para la medición del riesgo operacional.

¿Qué es el valor en riesgo-VAR?. Se trata de un método para cuantificar la exposición al riesgo de mercado, utilizando técnicas de estadísticas tradicionales. Partamos de la base que los agentes económicos de hoy enfrentan riesgos de diferente naturaleza, como por ejemplo de crédito, de mercado, de liquidez, operacional, legal, etc. El Valor en Riesgo vendría a medir la pérdida que se podría sufrir en condiciones normales de mercado en un intervalo de tiempo y con un cierto nivel de probabilidad o de confianza.

Un ejemplo práctico puede ser que una entidad financiera podría considerar que el VaR diario de una cartera operativa es de 50 millones de euros, con un nivel de confianza del 90%. Esto quiere decir que solamente hay 1 posibilidad en 10, en condiciones normales de mercado, de que haya una pérdida superior a los 50 millones de euros.

La medición, en términos de capital económico, se convierte en el aspecto más complejo y , a la vez, más trascendental en el tratamiento operacional. El comité de Basilea propone tres enfoques para calcular los requerimientos de capital por dicho riesgo que, de menor a mayor grado de sofisticación y sensibilidad al riesgo, son:

- (1) Método del Indicador Básico (Basic Indicator Approach, BIA).
- (2) Método Estándar (Standardised Approach, SA).
- (3) Metodologías de Medición Avanzada (Advanced Measurement Approach, AMA)

A su vez, dentro de los modelos AMA, se describen tres metodologías: el Modelo de Medición Interna (Internal Measurement Approach, IMA); los Cuadros de Mando (Scorecards); y el Modelo de Distribución de pérdidas (Loss Distribution Approach, LDA).

Los modelos de Enfoque de Medición Avanzada

Los modelos AMA se basan en el cálculo interno de una entidad financiera de la frecuencia y la severidad (monto de pérdida) por un evento de riesgo operativo. Los modelos LDA

combinan las distribuciones de frecuencia y severidad para construir una distribución de pérdidas totales y, con base en esta distribución, calcular la medida de riesgo requerida como un cuantil de dicha distribución.

Aunque diversos modelos han sido utilizados para cuantificar el riesgo operacional, aún no ha surgido el que se pueda considerar "mejor". La principal dificultad con estos métodos es la escasez de datos de buena calidad que reflejan las operaciones reales de la compañía.

Tipos de Riesgo¹

De acuerdo con el comité de supervisión Bancaria de Basilea (2004), los tipos de riesgo se dividen en:

1) **Riesgo de Crédito.**

Esta asociado a la pérdida resultante de los incumplimientos o de las contrapartes vinculados a transacciones financieras y crediticias. Este género de fallas pueden derivarse de las condiciones y calidad propias del cliente o contraparte, así como por el Riesgo soberano, si se tratará de transacciones internacionales. Este último se deriva de los controles impuestos por un país, que imposibilitan a la contraparte a cumplir con sus obligaciones.

2) **Riesgo de mercado.**

Es la pérdida, a la que está expuesta la institución, producto de movimientos adversos en los mercados financieros y que afectan las tasas de interés, el tipo de cambio, o la cotización de los activos en el portafolio de bienes, servicios e inversiones.

3) **Riesgo de seguro.**

Es la pérdida derivada de la suscripción de malos riesgos, de precios inadecuados, de la exposición contractual no anticipada en el diseño o precio del producto, de la desviación del número o cuantía de los reclamos, de la constitución inapropiada de las reservas, de eventos propios del ciclo económico, y de un inadecuado programa de reaseguro.

4) **Riesgo operacional.**

Es el riesgo de pérdida debido a la inadecuación o a fallos de los procesos, el personal, los sistemas internos o bien a causa de acontecimientos externos. Incluye el riesgo legal pero excluye el riesgo estratégico y el de reputación.

5) **Riesgo estratégico.**

Es la posibilidad (probabilidad) e impacto asociado a que los objetivos y metas propuestos por la alta directiva se vean afectados por amenazas tanto internas como externas a la organización.

6) **Riesgo de reputación.**

Es todo aquello capaz de afectar negativamente la imagen que tienen los clientes, trabajadores, accionistas y la colectividad en general en relación a la empresa.

¹Tomado del libro: Diseño e implantación de Indicadores Claves de Riesgo Operacional para la infraestructura Tecnológica de Mercantil Seguros.

7) **Riesgo legal.**

Es la posibilidad de ser sancionado, multado u obligado a pagar daños punitivos como resultado de acciones supervisoras o de acuerdos privados entre las partes.

Factores de Riesgo Operacional

Entendiendo como factor cualquier fuente capaz de originar un riesgo, se pueden nombrar cinco factores de riesgo operacional:

Procesos: Agrupa las pérdidas que son causadas por deficiencias o ausencia de procedimientos, recursos dispuestos o políticas internas. Incumplimiento de leyes y regulaciones.

Gente: Incluye las pérdidas asociadas a violaciones, intencionales o no, de leyes, normas o políticas internas por parte de los empleados.

Tecnológico: Pérdidas causadas por problemas en los sistemas, en las telecomunicaciones, datos y elementos informáticos.

Físico: Pérdidas causadas por interrupciones del negocio producto de fuerzas naturales, vandalismo, terrorismo, daños en la infraestructura.

Externo: Pérdidas causadas por cambios regulatorios, en contratos a terceros, o como consecuencia de hechos de política, fraude, robo y atraco.

0.1. NOCIONES DE PROBABILIDAD

Con base al hecho de que podamos predecir con exactitud o no la ocurrencia de un evento o experimento podemos definir dos tipos de experimentos, éstos son:

Definición 1 (Experimento determinístico). *Es aquel en el cual, bajo las mismas condiciones experimentales, las repeticiones del experimento absolutamente todas, siempre producen el mismo resultado.*

Definición 2 (Experimento Aleatorio). *Conservando las mismas condiciones experimentales, los resultados no se pueden predecir, con exactitud, para ninguna repetición.*

Definición 3. *Todos los posibles resultados de un experimento aleatorio, conforman el espacio muestral que representaremos por S , a cualquier subconjunto del espacio muestral se le denomina suceso o evento aleatorio y lo denotaremos con " E ", $E \subset S$. Cada uno de los elementos del espacio muestral se denomina evento elemental " e ". $e \in E \subset S$.*

0.2. AXIOMAS DE LA TEORÍA DE PROBABILIDAD.

Para todo experimento, la probabilidad de ocurrencia de un evento A , $p(A)$, es una función que cumple con los siguientes axiomas:

- 1) $p(A) \geq 0$ Toda probabilidad es no negativa.
- 2) $p(S) = 1$ La probabilidad del espacio muestral es 1.
- 3) Si dos o mas sucesos son incompatibles entre sí, entonces la probabilidad de la unión de ellos, es igual a la suma de sus probabilidades respectivas.

$$p(A \cap B) = 0 \rightarrow p(A \cup B) = p(A) + p(B)$$

Definición 4. Una variable aleatoria es una función definida sobre un espacio muestral a los números reales. Si ese espacio muestral especificado como dominio es numerable, decimos que la variable es de tipo **discreta**, en caso contrario es de tipo **continua**.

0.3. FUNCIÓN DE PROBABILIDAD.

Las variables aleatorias, transforman los eventos del espacio muestral en eventos numéricos, los cuales desde luego, tienen asociada una probabilidad de ocurrencia.

Definición 5 (FUNCIÓN DE PROBABILIDAD $f(x)=p(X=x)$). Es una función definida sobre una variable aleatoria a los reales en el intervalo $[0, 1]$ que cumple con los axiomas de la teoría de la probabilidad.

Definición 6. *FUNCIÓN DE DISTRIBUCIÓN* $F(x) = p(X \leq x)$

Una función $f : [0, 1] \rightarrow [0, 1]$ es llamada función de distribución si satisface las siguientes propiedades:

- a) $F(0) = 0, F(1) = 1$
- b) $\lim_{x \rightarrow a^+} F(x) = F(a), \forall a \in [0, 1]$
- c) $x_1 > x_2$ implica que $F(x_1) > F(x_2)$

0.4. ESPACIO FRACTAL Y GENERACIÓN DE FRACTALES

Cuando se va a trabajar ciertas clases de objetos conviene precisar en primer lugar de qué tipo de objetos se trata, dónde van a vivir y como se pueden distinguir unos de otros.

Consideremos el espacio \mathbb{R}^n dotado de la clásica distancia euclídea que, para cada $x, y \in \mathbb{R}^n$ con $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ viene definida por:

$$d(x, y) = \|x - y\| = \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{\frac{1}{2}}$$

Definición 7 (Espacio Hausdorff). Sea (X, d) un espacio métrico completo, se define $\mathbb{H}(X)$ como el conjunto cuyos puntos son los subconjuntos compactos no vacíos de X .

Como vamos a tratar el problema de aproximar objetos (fractales) mediante fractales que nosotros podamos generar, se hace necesario disponer de una métrica que nos permita medir la distancia entre los distintos fractales, es decir la distancia entre los elementos del espacio \mathbb{R}^n . Esta métrica nos va a contestar a la cuestión: ¿cómo se distinguen, unos de otros, los conjuntos fractales?, ya que por la definición de métrica o distancia, la condición necesaria y suficiente para que dos fractales sean iguales es que la distancia entre ellos sea cero.

Dado un punto $x \in X$ y $B \in \mathbb{H}(X)$, se define:

$$d(x, B) = \text{mín}\{d(x, y), y \in B\}$$

Se observa que $d(x, B) = 0$ si y solo si $x \in B$. Además en esta definición se refiere al "mínimo" y no al ínfimo, esto se debe a que dicho mínimo se alcanza (ya que B al ser compacto es cerrado). En efecto, consideremos la aplicación:

$$f : X \rightarrow \mathbb{R}$$

$$f(y) = \|d(x, y) - T\|$$

donde T es el ínfimo del conjunto $\{d(x, y), y \in B\}$

Para cada n natural, tómesese un punto $y_n \in B$ de forma que $\|d(x, y_n) - T\| < \frac{1}{n}$. Por ser B compacto la sucesión $\{y_n\}$ (convergente a cero) ha de contener una subsucesión $\{y_{n_k}\}$ convergente a algún punto $y_0 \in B$. Por la continuidad de f se deduce que $\{f(y_{n_k})\}$ converge a $f(y_0) = \|d(x, y_0) - T\|$ que tiene que ser cero. De esta forma $d(x, y_0) = T$ y el mínimo se alcanza.

Ahora definimos la distancia entre dos compactos $A, B \in \mathbb{H}(\mathbb{R}^n)$ como sigue:

$$\bar{d}(A, B) = \text{máx}\{d(x, B), x \in A\}$$

Observese que el máximo se alcanza, además que $\bar{d}(A, B) = 0$ si y solo si $A \subset B$, lo que significa que $(\mathbb{H}(\mathbb{R}^n), \bar{d})$ no es un espacio métrico.

Considérese dos discos concéntricos cerrados de radio r, R , ver 1

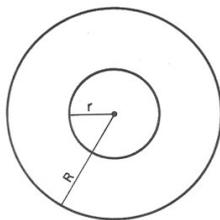


Figura 1: Discos concéntricos

Claramente $\bar{d}(A, B) = 0$ mientras que $\bar{d}(B, A) = R - r$ por lo que \bar{d} no es una métrica y por tanto no es una distancia.

La métrica que nosotros vamos a considerar en el espacio de los conjuntos fractales va a ser

la métrica de Hausdorff, que viene definida por:

$$d_H(A, B) = \max\{\bar{d}(A, B), \bar{d}(B, A)\} \quad A, B \in \mathbb{H}(\mathbb{R}^n)$$

donde:

$$\bar{d}(A, B) = \max\{d(x, B), x \in B\}$$

$$\bar{d}(B, A) = \max\{d(y, A), y \in A\}$$

Teorema 1. *La aplicación d_H es una distancia en $\mathbb{H}(\mathbb{R}^n)$*

Demostración ver [2, pag 15-16]

Una de las formas mas populares para generar fractales es utilizar el concepto de los sistemas iterados de funciones, que introdujo Michael Barnsley en el año 1988 [8], dados que estos conjuntos presentan semejanza, es decir, el conjunto se puede descomponer en un número finito de partes de modo que cada una de ellas es idéntica, salvo escala.

Teorema 2. *El espacio métrico $(\mathbb{H}(\mathbb{R}^n), d_H)$ es completo.*

Demostración ver [2, pag 17-20]

Definición 8. *Un Sistema de Funciones Iteradas (IFS, por sus siglas en inglés) en un espacio métrico completo X , consiste en un número finito de transformaciones contractivas de X , w_1, \dots, w_N , denotado por $\{X; w_1, \dots, w_N\}$.*

Lema 1. *Sea (X, d) un espacio métrico y sea $w : X \rightarrow X$ una aplicación continua y $A \in \mathbb{H}(X)$, entonces $w(A) \in \mathbb{H}(X)$.*

Demostración. Simplemente se usará la definición de compacidad. Sea $\{w(x_n)\}$ una sucesión de puntos de $w(A)$. Entonces $\{x_n\}$ es una sucesión de puntos de A que, por ser este conjunto compacto, admite una subsucesión $\{x_{n_k}\}$ convergente ($\lim\{x_{n_k} = x_0\}$). Por ser w continua, $\lim\{w(x_{n_k})\} = w(x_0)$ por lo que se ha encontrado una subsucesión convergente de $\{w(x_n)\}$ y el lema queda probado. \square

Teorema 3. *Si $w : X \rightarrow X$ es una transformación contractiva de factor de contracción k , entonces $w : \mathbb{H}(X) \rightarrow \mathbb{H}(X)$ es también contractiva con el mismo factor de contracción.*

Demostración. Por ser w contractiva se tiene que:

$$\begin{aligned} \bar{d}(w(B), w(C)) &= \max_{x \in B} \{ \min_{y \in C} \{ d(w(x), w(y)) \} \} \\ &\leq k \max_{x \in B} \{ \min_{y \in C} \{ d(x, y) \} \} \\ &= k \bar{d}(B, C) \end{aligned}$$

Igualmente se deduce que:

$$\begin{aligned}\bar{d}(w(C), w(B)) &= \underbrace{\mbox{máx}}_{y \in C} \{ \underbrace{\mbox{mín}}_{x \in B} \{ d(w(y), w(x)) \} \} \\ &\leq k \underbrace{\mbox{máx}}_{y \in C} \{ \underbrace{\mbox{mín}}_{x \in B} \{ d(x, y) \} \} \\ &= k\bar{d}(C, B)\end{aligned}$$

Por lo que:

$$\begin{aligned}d_H(w(B), w(C)) &= \mbox{máx} \{ \bar{d}(w(B), w(C)), \bar{d}(w(C), w(B)) \} \\ &\leq k \mbox{máx} \{ \bar{d}(B, C), \bar{d}(C, B) \} \\ &= kd_H(B, C)\end{aligned}$$

□

Sean w_1, \dots, w_N transformaciones contractivas de X con factores de contracción k_1, \dots, k_N respectivamente. Defínase entonces una transformación:

$$W : \mathbb{H}(X) \rightarrow \mathbb{H}(X)$$

$$W(A) = w_1(A) \cup w_2(A) \cup \dots \cup w_N(A)$$

Esto es, W transforma una determinada "imagen" en la unión de las "imágenes" producidas por cada una de las $w_i, i = 1, \dots, N$.

Teorema 4. *Sea la transformación $W : \mathbb{H}(X) \rightarrow \mathbb{H}(X)$ asociada al IFS $\{X; w_1, \dots, w_N\}$ con factor de contracción k_i para cada w_i respectivamente, con $i = 1, \dots, N$, es decir:*

$$W : \mathbb{H}(X) \rightarrow \mathbb{H}(X)$$

definida por

$$W(B) = \bigcup_{i=1}^n w_i(B)$$

para cada $B \in \mathbb{H}(X)$.

entonces W es una transformación contractiva con factor de contracción $k = \mbox{máx} \{ k_i, i = 1, \dots, N \}$

Demostración ver [8, pag 81-82]

Teorema 5. *Sea X un espacio métrico completo y sea $f : X \rightarrow X$ una contracción. Entonces f admite un único punto fijo, ese punto puede ser encontrado como límite de la sucesión $(x_0, f(x_0), f(f(x_0)), \dots)$, para cualquier punto $x_0 \in X$*

Demostración ver [2, pag 26-28]

En este momento surge la pregunta natural ¿Qué es un fractal?.

Bajo el contexto que hemos venido trabajando tenemos la siguiente definición(ver [2, pag:65]):

Definición 9. *Al compacto A_∞ , punto fijo de la transformación W se le llama **Fractal** asociado al IFS $\{X; w_1, \dots, w_N\}$, se le conoce también como el "atractor" del IFS.*

Ahora, debido al amplio campo donde se encuentran y aplican los fractales y debido a que es un tema que esta en investigación se pueden encontrar otras definiciones, las más usuales son²:

- 1) Los Fractales son los objetos matemáticos que conforman la Teoría del Caos.
- 2) Los Fractales son objetos cuya dimensión es un número racional.
- 3) Un fractal es un objeto que su dimensión de **Hausdorff-Besicovich** es mayor que su dimensión topológica.
- 4) Un objeto fractal es aquel que tiene autosemejanza y su dimensión es racional.

Lo que nos lleva a la necesidad de definir el siguiente concepto³:

Definición 10. Dimensión Fractal:*Sentido genérico: Número que sirve para cuantificar el grado de irregularidad y fragmentación de un conjunto geométrico o de un conjunto natural. La dimensión fractal no es necesariamente entera.*

² Tomado de:<http://ucua.ujaen.es/jnavas/mayores/fractales.pdf>

³Tomado de:Una Introducción a la geometría fractal, pág:5

Capítulo 1

FRACTALES Y FINANZAS: DESARROLLO DEL MÉTODO

En los últimos años, fundamentalmente desde el surgimiento del Nuevo Acuerdo de Basilea, que planteó el requerimiento de incorporar el riesgo operacional en la relación de solvencia, han surgido diversas alternativas metodológicas para su cuantificación.

En la gama de alternativas de medición avanzada, el método de distribución de pérdidas (LDA, por sus siglas en inglés), que involucra herramientas del cálculo actuarial, ha sido el más generalmente utilizado en este campo.

A partir de la distribución de pérdidas agregadas, obtenida generalmente a través de métodos numéricos, se calcula la carga de capital, o el llamado Valor en Riesgo Operacional, mediante el cuantil correspondiente al nivel de confianza determinado, que para el caso de riesgo operacional ha sido establecido por Basilea en el 99.9%.

En el presente trabajo se presenta la llamada estimación fractal, basada en sistemas de funciones iteradas, como posible alternativa para la medición del riesgo. No se hace ninguna comparación específica con otras metodologías, ya que esta es una introducción al método para exponer como se aplica la geometría fractal en el área de finanzas.

En aplicaciones prácticas un problema crucial es el llamado problema inverso, que se formula de la siguiente manera: dado f en un espacio métrico (S, d) , encontrar una contracción $T : S \rightarrow S$ que admita un único punto fijo $\hat{f} \in S$ tal que $d(f, \hat{f})$ sea lo suficientemente pequeño. Si se resuelve el problema inverso con precisión adecuada, se puede identificar f con el operador T que tiene a \hat{f} como punto fijo.

Una solución aproximada del problema inverso es la solución del siguiente problema de optimización restringida, que llamaremos el problema P :

$$\min_{p \in C} d_{\infty}(T_p F, F) \quad (P)$$

donde $T_p F$ es un operador que aproxima a F y que depende de un vector de probabilidades P que es tomado de un conjunto C y $d_{\infty}(x, y)$ es la métrica del supremo.

1.1. CONTRACCIONES EN EL ESPACIO DE FUNCIONES DE DISTRIBUCIÓN

Definición 11. Denotamos por $F([0, 1])$ el espacio de funciones de distribución de probabilidad en el intervalo $[0, 1]$.

Definición 12. Denotamos por $B([0, 1])$ el espacio de funciones reales acotadas en $[0, 1]$.

Definición 13 (Métrica del Supremo). Sea X un conjunto y sea $f : X \rightarrow \mathbb{R}$ y $g : X \rightarrow \mathbb{R}$, dos funciones definidas en el conjunto X , entonces la distancia entre f, g se define por:

$$d_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|$$

Teorema 6. Toda sucesión de Cauchy en un espacio métrico (X, d) está acotada.

Demostración. Sea x_n una sucesión de Cauchy y sea $\epsilon = 1$.

Por ser de Cauchy existe n_0 tal que si $m, n > n_0$, se tiene que $d(x_m, x_n) < \epsilon = 1$.

de modo si $n > n_0$ y tomando $m = n_0 + 1$, se tiene que $d(x_{n_0+1}, x_n) < \epsilon = 1$, es decir que $x_n \in B(x_{n_0+1}, 1), \forall n = n_0 + 1 > n_0$.

Quedan un número finito de términos fuera de esta bola, x_1, x_2, \dots, x_{n_0} , así definimos

$$r = \text{máx}\{d(x_1, x_{n_0+1}), d(x_2, x_{n_0+1}), \dots, d(x_{n_0}, x_{n_0+1})\}$$

esto significa que $d(x_n, x_{n_0+1}) \leq r, \forall n$.

Así:

$$\{x_n : n : 1, 2, \dots, \infty\} \subset B(x_{n_0+1}, r + 1)$$

□

Teorema 7. Sea X un conjunto cualquiera. Entonces el espacio métrico $(B(X, \mathbb{R}), d_\infty)$, donde $B(X, \mathbb{R}) = \{f : X \rightarrow \mathbb{R} : f \text{ acotada}\}$, es completo.

Demostración. Sea $(f_n)_{n=1}^\infty$ una sucesión de funciones de Cauchy en $B(X, \mathbb{R})$. Entonces para cada $x \in X$, la sucesión de números $(f_n(x))_{n=1}^\infty$ es de Cauchy en $(\mathbb{R}, | \cdot |)$.

Como $(\mathbb{R}, | \cdot |)$ es completo, se tiene que para cada $x \in X$ la sucesión $(f_n(x))_{n=1}^\infty$ converge a un punto en \mathbb{R} que llamamos $f(x)$.

Es decir

$$\lim_{n \rightarrow \infty} (f_n(x))_{n=1}^\infty = f(x), x \in X$$

A partir de este límite, definimos $f : X \rightarrow \mathbb{R}$ tal que a cada $x \in X$ le hace corresponder el límite de la sucesión $(f_n(x))_{n=1}^\infty$, que llamamos $f(x)$.

$$\begin{aligned} f & : X \rightarrow \mathbb{R} \\ x & \rightarrow \lim_{n \rightarrow \infty} (f_n(x))_{n=1}^\infty = f(x), x \in X \end{aligned}$$

Veamos que $(f_n)_n$ converge a f .

Como la sucesión es de Cauchy, tendremos que $\forall \epsilon > 0$, existe n_0 tal que si $m, n > n_0$, se tiene

$$d_\infty(f_n, f_m) = \sup\{|f_n(x) - f_m(x)| : x \in X\} < \epsilon$$

En particular, si tomamos $n > n_0$ (hijo) y $p \in \mathbb{N}$, tendremos que

$$d_\infty(f_n, f_{n+p}) = \sup\{|f_n(x) - f_{n+p}(x)| : x \in X\} < \epsilon$$

Así $\forall x \in X$ se cumple

$$|f_n(x) - f_{n+p}(x)| < \epsilon$$

si tomamos $p \rightarrow \infty$, tendremos que $\forall x \in X$

$$|f_n(x) - f_{n+p}(x)| \rightarrow |f_n(x) - f(x)|$$

esto significa que

$$|f_n(x) - f(x)| < \epsilon$$

y por tanto la conclusión es que si $n > n_0$ entonces

$$d_\infty(f_n, f) = \sup\{|f_n(x) - f(x)| : x \in X\} < \epsilon$$

lo que implica que $(f_n)_n$ converge a f .

Esto también implica que $f \in B(X, \mathbb{R})$. Como $(f_n)_n$ es de Cauchy por teorema 6 está acotada. Así existe $M > 0$ tal que $(f_n)_n \subset B(0, M)$. Como 0 es la función idénticamente nula, es decir $d_\infty(0, f_n) < M, \forall n \in \mathbb{N}$.

Entonces, dado $\epsilon = 1 > 0$, como $(f_n)_n$ converge a f , existe n_1 tal que si $n > n_1$ entonces $d_\infty(f_n, f) < \epsilon = 1$.

Por tanto podemos poner (utilizando la desigualdad triangular):

$$\begin{aligned} d_\infty(0, f) &\leq d_\infty(0, f_n) + d_\infty(f_n, f) \\ d_\infty(0, f) &\leq M + 1 \end{aligned}$$

□

Corolario 1. *El espacio métrico $(B([0, 1]), d_\infty)$ es completo*

Demostración. Es inmediato del teorema 7

□

Teorema 8. *El espacio $(F([0, 1]), d_\infty)$ es un espacio métrico completo.*

Demostración. Sea $(f_n)_n$ una sucesión de Cauchy en $f([0, 1])$, además observar que $F([0, 1]) \subset B([0, 1])$.

Si definimos $f(x) = \lim_{n \rightarrow \infty} f_n(x)$, entonces f_n converge a f en $(B([0, 1]), d_\infty)$

Además es verdad que:

$$f(0) = \lim_{n \rightarrow \infty} f_n(0) = 0$$

$$f(1) = \lim_{n \rightarrow \infty} f_n(1) = 1$$

y si $x_1 \geq x_2$ entonces

$$f(x_1) = \lim_{n \rightarrow \infty} f_n(x_1) \geq \lim_{n \rightarrow \infty} f_n(x_2) = f(x_2)$$

Para probar la continuidad por la derecha usamos la convergencia uniforme de f_n a f .

$$\begin{aligned} \lim_{x \rightarrow a^+} f(x) &= \lim_{x \rightarrow a^+} \lim_{n \rightarrow \infty} f_n(x) \\ &= \lim_{n \rightarrow \infty} \lim_{x \rightarrow a^+} f_n(x) \\ &= f(a) \end{aligned}$$

□

Ahora introducimos los sistemas de funciones iteradas de una manera mas general, lo que nos permitirá aproximar funciones de distribución, que es lo que perseguimos.

1.2. OPERADORES IFS

Sea K un espacio métrico compacto y

$$\omega = \{\omega_i : i = 1, 2, \dots, d\}$$

una colección finita de funciones Borel medibles. $\omega_i : K \rightarrow K$. Sea $C(K)$ el espacio de Banach de funciones reales de valor real sobre si, con la norma

$$\|f\|_\infty = \max\{|f(x)| : x \in K\}, \quad f \in C(K)$$

Usamos la notación $P = \{p_i : i = 1, 2, \dots, d\}$ para un conjunto de probabilidades $p_i \geq 0$, con

$$\sum_{i=1}^d p_i = 1$$

Escribiremos $P > 0$ para indicar que $p_i > 0$ para cada i .

Definición 14. $\{K, \omega\}$ es un sistema de funciones iteradas (IFS) si y solo si existe un conjunto asociado de probabilidades P tales que el operador T en $C(K)$, dado por

$$(Tf)(x) = \sum_{i=1}^d p_i (f \circ \omega_i)(x), \quad f \in C(K)$$

tiene la propiedad

$$T : C(K) \rightarrow C(K)$$

Teorema 9 (Banach). Sea (F, d) un espacio métrico completo y $T : F \rightarrow F$ un operador contractivo, tal que existe una constante $c \in [0, 1)$ (factor de contracción) para el cual:

$$d(Tf_1, Tf_2) \leq cd(f_1, f_2)$$

Entonces existe un único elemento $\hat{f} \in F$ tal que:

$$T\hat{f} = \hat{f}, \quad y \quad \lim_{n \rightarrow \infty} d(T^n f, \hat{f}) = 0, \forall f \in F$$

donde T^n es la n -ésima iteración del operador T . \hat{f} es llamado el punto fijo del operador T

El teorema de Banach establece que para n "suficientemente grande" $T^n f$ es una "buena" aproximación del elemento $\hat{f} \in F$. Además da una solución al siguiente problema directo: "dado un operador T , encontrar, si existe, su punto fijo f ".

Consideremos el espacio métrico completo $\mathbb{H}([0, 1] \times [0, 1], d_H)$, donde $\mathbb{H}([0, 1] \times [0, 1])$ denota la colección de subconjuntos compactos de $[0, 1] \times [0, 1]$ y d_H denota la distancia de Hausdorff entre los conjuntos compactos $A, B \subset [0, 1] \times [0, 1]$.

Con esta idea se ajusta una imagen en blanco y negro como un elemento C de $\mathbb{H}([0, 1] \times [0, 1])$ y para construir una aproximación de C uno tiene que construir un operador $T : \mathbb{H}([0, 1] \times [0, 1]) \rightarrow \mathbb{H}([0, 1] \times [0, 1])$ que tiene a C como único punto fijo. Este tipo de operadores son llamados operadores IFS.

En la práctica es difícil encontrar un operador contractivo que tenga la imagen como único punto fijo. Por lo tanto, se puede relajar esta solicitud, buscando un operador de contracción que tenga el punto fijo "cercano" a la imagen dada. Es evidente que, en aplicaciones prácticas, el siguiente problema inverso es más importante que el directo:

"dado $f \in F$, encontrar un operador contractivo $T : F \rightarrow F$ tal que admite un único punto fijo $\hat{f} \in F$, tal que $d(f, \hat{f})$ es suficientemente pequeña."

En este trabajo una imagen será una función de distribución. Es decir, el trabajo será encontrar un operador contractivo del espacio de funciones en sí mismo que tenga un punto fijo "cercano" a una función de distribución dada. De una manera aproximada, este problema puede resolverse utilizando el resultado, conocido como Teorema del Collage.

Sea $N \in \mathbb{N}$ fijo y consideremos los siguientes puntos:

- (i) $w_i : [a_i, b_i] \rightarrow [c_i, d_i], i = 1, \dots, N - 1$
 $w_N : [a_N, b_N] \rightarrow [c_N, d_N]$
con
 $a_1 = c_1 = 0$ y $b_N = d_N = 1$;
 $c_{i+1} = d_i, i = 1, \dots, N - 1; [a_i, b_i] \subset [0, 1]$
- (ii) $w_i, i = 1, \dots, N$ son crecientes y continuas;
- (iii) $\bigcup_{i=1}^{N-1} [c_i, d_i] \cup [c_N, d_N] = [0, 1]$
- (iv) Si $i \neq j$, entonces $[c_i, d_i] \cap [c_j, d_j] = \emptyset$
- (v) $p_i \geq 0, i = 1, \dots, N, \delta_i \geq 0, i = 1, \dots, N, \sum_{i=1}^N p_i + \sum_{i=1}^{N-1} \delta_i = 1$

En $(F([0, 1]), d_\infty)$, definimos un operador de la siguiente manera:

$$T(F(x)) = \begin{cases} p_i f(w_i^{-1}(x)) & \text{si } x \in [c_1, d_1] \\ p_i f(w_i^{-1}(x)) + \sum_{j=1}^{i-1} p_j + \sum_{j=1}^{i-1} \delta_j & \text{si } x \in [c_i, d_i], i = 2, \dots, N-1 \\ P_N f(w_N^{-1}(x)) + \sum_{j=1}^{N-1} p_j + \sum_{j=1}^{N-1} \delta_j & \text{si } x \in [c_N, d_N]. \end{cases}$$

Donde $f \in F([0, 1])$.

Suponemos que w_i y los parámetros δ_i son fijos, y los p_i deben ser seleccionados. Entonces como, T depende del vector $P = (p_1, \dots, p_N)$ se denota por T_P .

Teorema 10. T_P es un operador de $F([0, 1])$ en si mismo.

Demostración. Se observa que $T_P(f(0)) = 0$ y $T_P(f(1)) = 1$

Sabemos que:

$$\omega_1 : [0, b_1] \rightarrow [0, d_1]$$

$$\omega_N : [a_N, 1] \rightarrow [c_N, 1]$$

De esto, tenemos:

$$\omega_1^{-1}(0) = 0, \omega_N(1) = 1.$$

Así tenemos:

$$T_P f(0) = p_1 f(\omega_1^{-1}(0)) \quad (1.1)$$

$$= p_1 f(0) \quad (1.2)$$

$$= 0 \quad (1.3)$$

$$T_P f(1) = p_N f(\omega_N^{-1}(1)) + \sum_{j=1}^{N-1} p_j + \sum_{j=1}^{N-1} \delta_j \quad (1.4)$$

$$= p_N + \sum_{j=1}^{N-1} p_j + \sum_{j=1}^{N-1} \delta_j \quad (1.5)$$

$$= \sum_{j=1}^N p_j + \sum_{j=1}^{N-1} \delta_j \quad \text{hipótesis} \quad (1.6)$$

$$= 1 \quad (1.7)$$

Además, si $x_1 > x_2$, sin pérdida de generalidad podemos considerar los siguientes dos casos:

(i) $x_1, x_2 \in \omega_i([a_i, b_i])$

(ii) $x_1 \in \omega_{i+1}([a_{i+1}, b_{i+1}])$ y $x_2 \in \omega_i([a_i, b_i])$

Caso (i): Sabemos que ω_i son mapeos crecientes, así tenemos que:

$$\omega_i^{-1}(x_1) > \omega_i^{-1}(x_2), (x_1 > x_2)$$

Esto significa que:

$$p_i f(\omega_i^{-1}(x_1)) \geq p_i f(\omega_i^{-1}(x_2))$$

Ya que f es una función de distribución.

Por lo tanto : $T_P f(x_1) \geq T_P f(x_2)$.

Caso (ii): $x_1 \in \omega_{i+1}([a_{i+1}, b_{i+1}))$

$x_2 \in \omega_i([a_i, b_i))$

Así tenemos que:

$$T_P f(x_1) = p_{i+1} f(\omega_{i+1}^{-1}(x_1)) + \sum_{j=1}^i p_j + \sum_{j=1}^i \delta_j$$

$$T_P f(x_2) = p_i f(\omega_i^{-1}(x_2)) + \sum_{j=1}^i p_j + \sum_{j=1}^i \delta_j$$

Esto significa que:

$$\begin{aligned} T_P f(x_1) - T_P f(x_2) &= p_i + \delta_i + p_{i+1} f(\omega_{i+1}^{-1}(x_1)) - p_i f(\omega_i^{-1}(x_2)) \\ &= p_i(1 - f(\omega_i^{-1}(x_2))) + p_{i+1} f(\omega_{i+1}^{-1}(x_1)) + \delta_i \\ &\geq 0 \end{aligned}$$

Esto significa que $T_P f(x)$ es creciente.

Veamos la continuidad por la derecha.

Esto equivale a ver que:

$$\lim_{n \rightarrow \infty} T_P f\left(x + \frac{1}{n}\right) = T_P f(x)$$

En efecto, analicemos el siguiente límite:

$$J = \lim_{n \rightarrow \infty} f\left(\omega_i^{-1}\left(x + \frac{1}{n}\right)\right)$$

Tenemos que:

$$J = f\left(\lim_{n \rightarrow \infty}\left(\omega_i^{-1}\left(x + \frac{1}{n}\right)\right)\right)$$

Esto porque f es continua por la derecha. De aquí, usando el hecho que ω_i^{-1} es continua, se tiene que:

$$J = f(\omega_i^{-1}(x))$$

Por lo anterior se tiene que:

$$\lim_{n \rightarrow \infty} T_P f\left(x + \frac{1}{n}\right) = T_P f(x)$$

Como se deseaba. □

Teorema 11. Si $c = \underbrace{\mbox{máx}}_{i=1,\dots,N} p_i < 1$.

Entonces T_P es una contracción en $(F([0, 1]), d_\infty)$ con constante de contracción c .

Demostración. Sea $f, g \in (F([0, 1]), d_\infty)$ y supóngase que $x \in w_i([a_1, b_i])$
Tenemos:

$$T_P(f(x)) = p_i f(w_i^{-1}(x)) + \sum_{j=1}^{i-1} p_j + \sum_{j=1}^{i-1} \delta_j$$

$$T_P(g(x)) = p_i g(w_i^{-1}(x)) + \sum_{j=1}^{i-1} p_j + \sum_{j=1}^{i-1} \delta_j$$

De donde:

$$\begin{aligned} |T_P(f(x)) - T_P(g(x))| &= p_i |f(w_i^{-1}(x)) - g(w_i^{-1}(x))| \\ &\leq p_i d_\infty(f, g), \forall x \end{aligned}$$

Sabemos que

$$d_\infty(f, g) = \sup_{x \in [0,1]} |T_P(f(x)) - T_P(g(x))|$$

De donde

$$d_\infty(T_P(f(x)), T_P(g(x))) \leq p_i d_\infty(f(x), g(x)) \leq c d_\infty(f(x), g(x))$$

□

El siguiente teorema establece que pequeñas variaciones en los parámetros P_i produce pequeñas variaciones en el punto fijo del operador.

Teorema 12. Sean $P, P^* \in \mathbb{R}^n$ tal que $T_P(f_1) = F_1$ y $T_{P^*}(f_2) = F_2$ con $f_1, f_2 \in F([0, 1])$.
Entonces:

$$d_\infty(f_1, f_2) \leq \frac{1}{1-c} \sum_{j=1}^N |p_j - p_j^*|$$

donde c es la constante de contracción de T_P .

Demostración.

$$\begin{aligned} |p_i(f_1(w_i^{-1}(x))) + \sum_{j=1}^{i-1} p_j - p_i^*(f_2(w_i^{-1}(x))) - \sum_{j=1}^{i-1} p_j^*| &\leq \left| \sum_{j=1}^{i-1} p_j - p_j^* \right| + |p_i f_1(w_i^{-1}(x)) \\ &\quad - p_i^* f_2(w_i^{-1}(x)) - p_i f_2(w_i^{-1}(x)) \\ &\quad + p_i f_2(w_i^{-1}(x))| \\ &\leq \sum_{j=1}^{i-1} |p_j - p_j^*| + |p_i f_1(w_i^{-1}(x)) \\ &\quad - p_i f_2(w_i^{-1}(x))| + |p_i f_2(w_i^{-1}(x)) \\ &\quad - p_i^* f_2(w_i^{-1}(x))| \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{j=1}^{i-1} |p_j - p_j^*| + p_i |f_1(w_i^{-1}(x)) - f_2(w_i^{-1}(x))| \\
&\quad + \underbrace{f_2(w_i^{-1}(x))}_{\leq 1} |p_i - p_i^*| \\
&\leq \sum_{j=1}^{i-1} |p_j - p_j^*| + p_i d_\infty(f_1, f_2) + |p_i - p_i^*| \\
d_\infty(f_1, f_2) &\leq cd_\infty(f_1, f_2) + \sum_{j=1}^N |p_j - p_j^*|
\end{aligned}$$

□

Tomamos $f \in F([0, 1], d_\infty)$. Recordemos que el objetivo consiste en encontrar un mapeo contractivo:

$$T : F([0, 1]) \rightarrow F([0, 1])$$

que tiene un punto fijo "cercano" a f . De hecho si es posible resolver el problema inverso exactamente, se puede identificar el operador T con su punto fijo.

En el contexto de riesgo operacional esa f que deseamos estimar será la función de distribución de las pérdidas agregadas en una línea específica de negocio, o en toda la entidad.

Definición 15. Una función Ψ definida en un intervalo abierto (a, b) es **convexa** si para cualquier números no negativos λ, μ tal que $\lambda + \mu = 1$, y x, y tal que $a < x < y < b$, tenemos:

$$\Psi(\lambda x + \mu y) \leq \lambda \Psi(x) + \mu \Psi(y)$$

Los puntos extremos a, b pueden tomar los valores $\infty, -\infty$ respectivamente. Si, para todos los números positivos λ, μ tal que $\lambda + \mu = 1$ tenemos

$$\Psi(\lambda x + \mu y) < \lambda \Psi(x) + \mu \Psi(y)$$

Ψ se dice que es **estrictamente convexa**.

Definición 16. Un conjunto C es **convexo** si y solo si para todo $a, b \in C$, $[a, b] \subset C$, es decir:

$$\begin{aligned}
&\forall t \in [0, 1]: \\
&(1-t)a + tb \in C
\end{aligned}$$

Con un sistema fijo de mapeos w_i y parámetros δ_i , el problema inverso puede ser resuelto, si es posible, usando los parámetros p_i . Estos deben ser elegidos en el siguiente conjunto convexo:

$$C = \left\{ P \in \mathbb{R}^N : p_i \geq 0, i = 1, \dots, N, \sum_{i=1}^N p_i = 1 - \sum_{i=1}^{N-1} \delta_i \right\}$$

En efecto, a continuación probamos que efectivamente el conjunto C es convexo.

Teorema 13. *El conjunto C , definido por:*

$$C = \left\{ P \in \mathbb{R}^N : p_i \geq 0, i = 1, \dots, N, \sum_{i=1}^N p_i = 1 - \sum_{i=1}^{N-1} \delta_i \right\}$$

es convexo

Demostración. Sean $A, B \in C$, definimos la recta $l : A + (B - A)t, t \in \mathbb{R}$. Consideremos el punto $x \in \mathbb{R}^N$ que pertenece a la recta l , así tenemos que:

$$x = A + (B - A)t$$

$$\begin{aligned} \sum_{i=1}^N x_i &= \sum_{i=1}^N (A_i + (B_i - A_i)t) = \sum_{i=1}^N A_i + \left(\sum_{i=1}^N B_i - \sum_{i=1}^N A_i \right) t \\ &= \sum_{i=1}^N A_i(1 - t) + t \sum_{i=1}^N B_i \\ &= \left(1 - \sum_{i=1}^{N-1} \delta_i \right) (1 - t) + t \left(1 - \sum_{i=1}^{N-1} \delta_i \right) \\ &= \left(1 - \sum_{i=1}^{N-1} \delta_i \right) (1 - t + t) \\ &= \left(1 - \sum_{i=1}^{N-1} \delta_i \right) \end{aligned}$$

Es decir:

$$\sum_{i=1}^N x_i = \left(1 - \sum_{i=1}^{N-1} \delta_i \right)$$

Se ha hecho uso del hecho que:

$$\begin{aligned} \sum_{i=1}^N A_i &= \left(1 - \sum_{i=1}^{N-1} \delta_i \right) \\ \sum_{i=1}^N B_i &= \left(1 - \sum_{i=1}^{N-1} \delta_i \right) \end{aligned}$$

Por lo tanto C es convexo. □

Teorema 14. *La función*

$$\begin{aligned} D(P) &: \mathbb{R}^N \rightarrow \mathbb{R} \\ &= d_\infty(T_P(f), f) \end{aligned}$$

es convexa.

Demostración. Si escogemos $P_1, P_2 \in \mathbb{R}^N$ y $\lambda \in [0, 1]$, entonces:

$$D(\lambda P_1 + (1 - \lambda)P_2) = \sup_{x \in [0, 1]} |T_{\lambda P_1 + (1 - \lambda)P_2} f(x) - f(x)|$$

Notar que: $-f(x) = -\lambda f(x) - (1 - \lambda)f(x)$.

Tenemos que:

$$\begin{aligned}
|T_{\lambda P_1 + (1-\lambda)P_2}(f(x)) - f(x)| &\leq |\lambda(P_1 f(w_i^{-1}(x))) + \sum_{j=1}^{i-1} P_{1j} + \sum_{j=1}^{i-1} \delta_j + \sum_{j=1}^{i-1} \delta_j \\
&\quad + (1 - \lambda)(P_2 f(w_i^{-1}(x))) + \sum_{j=1}^{i-1} P_{2j} - f(x)| \\
&\leq \lambda |T_{P_1}(f(x)) - f(x)| + (1 - \lambda) |T_{P_2}(f(x)) - f(x)| \\
&\leq \lambda \sup_{x \in [0,1]} |T_{P_1}(f(x)) - f(x)| + (1 - \lambda) \sup_{x \in [0,1]} |T_{P_2}(f(x)) - f(x)| \\
&= \lambda D(P_1) + (1 - \lambda) D(P_2)
\end{aligned}$$

De donde:

$$D(\lambda P_1 + (1 - \lambda)P_2) \leq \lambda D(P_1) + (1 - \lambda) D(P_2)$$

Por lo tanto la función D es convexa. □

Estos dos últimos teoremas nos dicen que lo que hemos estado trabajando es un problema de optimización convexa, pues la función objetivo a minimizar ($D(P)$) es convexa y el conjunto solución (C), de dicho problema, es convexo.

Teorema 15 (Teorema de Collage). *Tomando $\epsilon > 0$ y $P \in C$ tal que $\forall P_i > 0$ para algún i , si*

$$d_\infty(T_P(f), f) \leq \epsilon$$

entonces

$$d_\infty(f, \tilde{f}) \leq \frac{\epsilon}{1 - c}$$

donde \tilde{f} es el punto fijo de T_P en $F([0, 1])$ y $c = \max_{i=1, \dots, N} P_i$ es la constante de contracción.

Demostración. Haciendo uso de la desigualdad triangular y del hecho que T_P es una contracción, se tiene:

$$\begin{aligned}
d_\infty(f, \tilde{f}) &\leq d_\infty(f, T_P(f)) + d_\infty(T_P(f), \tilde{f}) \\
&\leq \epsilon + d_\infty(T_P(f), \tilde{f}) \\
&\leq \epsilon + d_\infty(T_P(f), T_P(\tilde{f})) \\
&\leq \epsilon + c d_\infty(f, \tilde{f})
\end{aligned}$$

De aquí, es evidente que:

$$d_\infty(f, \tilde{f}) \leq \frac{\epsilon}{1 - c}$$

□

1.3. ESTIMACIÓN Y APLICACIONES DE LA FUNCIÓN DE DISTRIBUCIÓN

En lugar de tratar de resolver exactamente el problema P utilizamos las propiedades de las funciones de distribución para obtener un buen aproximador de F que se pueda utilizar directamente en la estimación de la función de distribución.

Mostramos que bajo condiciones adecuadas en los mapeos w_i , el estimador de la función de distribución IFS propuesto es asintóticamente eficiente para muestras grandes.

Es usual en aplicaciones estadísticas, dada una muestra de n observaciones independientes e idénticamente distribuidas (x_1, \dots, x_n) , extraídas de una función de distribución desconocida $f \in F([0, 1])$ uno puede fácilmente construir la (EDF) \hat{F}_n dada por:

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \chi_{(-\infty, x]}(x_i), x \in \mathbb{R}$$

donde χ_A es la función característica de un conjunto A .

Esta función tiene una representación IFS que es exacta y puede ser encontrado sin resolver ningún problema de optimización.

Asumimos que todos los x_i de la muestra son diferentes (esta suposición es habitual ya que F es una función de distribución continua).

Sea

$$w_i(x) : [x_{i-1}, x_i) \rightarrow [x_{i-1}, x_i), i = 1, \dots, n$$

y

$$w_1 : [0, x_1) \rightarrow [0, x_1)$$

$$w_{n+1}(x) : [x_n, x_{n+1}] \rightarrow [x_n, x_{n+1}]$$

con $x_0 = 0$ y $x_{n+1} = 1$.

Suponga además que cada mapeo es de la forma:

$$w_i(x) = x$$

Si escojemos $p_i = \frac{1}{n}, i = 2, \dots, n + 1, p_1 = 0$ y $\delta_1 = \frac{n-1}{n^2}, \delta_i = \frac{-1}{n^2}$.

Entonces se cumple la siguiente representación:

$$T(F\hat{x}) = \begin{cases} 0 & i = 1 \\ \frac{1}{n}\hat{F}_n(x) + \frac{n-1}{n^2} & i = 2 \\ \frac{1}{n}\hat{F}_n(x) + \frac{i-1}{n} + \frac{n-i+1}{n^2} & i = 3, \dots, n + 1 \end{cases}$$

Para $x \in [x_{i-1}, x_i)$.

Para producir un estimador, primero deberemos proporcionar una buena aproximación de F .

De ahora en adelante asumimos que $\delta_i = 0$ para todo i .

Así que se fija un $F \in F([0, 1])$ y escojemos $N + 1$ -eneada (x_1, \dots, x_{N+1}) tal que $x_1 = 0$ y $x_{N+1} = 1$.

Definimos los mapeos w_i y coeficientes p_i como sigue: para $i = 1, \dots, N$

$$p_i(F) = F(x_{i+1}) - F(x_i), \quad w_i(x) : [0, 1] \rightarrow [x_i, x_{i+1}] = (x_{i+1} - x_i) \cdot x + x_i \quad (1.8)$$

El funcional T_P podemos denotarlo como T_N con este conjunto dado de mapeos y coeficientes ya que dependen del número de puntos y de sus valores.

Para algún $\mu \in F([0, 1])$, T_N podemos escribirla para $x \in \mathbb{R}$ como:

$$T_N(\mu(x)) = \sum_{i=1}^N p_i \mu(w_i^{-1}(x)) = \sum_{i=1}^N (F(x_{i+1}) - F(x_i)) \cdot \mu \left(\frac{x - x_i}{x_{i+1} - x_i} \right) \quad (1.9)$$

T_N es una contracción en $(F([0, 1]), d_\infty)$ y $T_N(\mu(x_i)) = F(x_i), \forall i$.

Este funcional es de hecho una función de F y no se puede utilizar directamente en aplicaciones estadísticas ya que F es desconocido.

Para este fin, tomamos los puntos x_i para ser los cuantiles q_i de F , es decir, escojase $N + 1$ puntos $\mu_1 = 0 < \mu_2 < \dots < \mu_N < \mu_{N+1} = 1$ igualmente espaciados en $[0, 1]$ y el conjunto $q_i = F^{-1}(\mu_i)$. Así, la función T_N se convierte en:

$$T_N(\mu(x)) = \sum_{i=1}^N \frac{1}{N} \mu \left(\frac{x - q_i}{q_{i+1} - q_i} \right), x \in \mathbb{R} \quad (1.10)$$

y T_N depende solo de los cuantiles, q_i , de F .

De hecho, si dos cuantiles q_i y q_{i+1} son relativamente distintos cada uno del otro, entonces F esta creciendo lentamente en el intervalo (q_i, q_{i+1}) y viceversa.

Como los cuantiles pueden ser fácilmente estimados de una muestra, ahora tenemos un posible candidato par un *IFS* del estimador de la función de distribución.

Así, sea x_1, \dots, x_n una muestra de F y sea $\hat{q}_i, i = 1, \dots, N + 1$ los cuantiles empíricos de orden $\frac{1}{N}$ tal que $\hat{q}_1 = 0$ y $\hat{q}_{N+1} = 1$.

Entonces proponemos como *IFS* estimador de la función de distribución el punto fijo del siguiente *IFS*:

$$\hat{T}_N \mu(x) = \sum_{i=1}^N \frac{1}{N} \mu \left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i} \right) \quad (1.11)$$

1.4. Propiedades asintóticas

Sea $N = N_n$ una sucesión dependiendo de el número de la muestra n . Denotamos el punto fijo de \hat{T}_{N_n} por $\hat{T}_{N_n}^*$. Entonces $\hat{T}_{N_n}^*$ satisface

$$\hat{T}_{N_n}^*(x) = \sum_{i=1}^{N_n} \frac{1}{N_n} \hat{T}_{N_n}^* \left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i} \right), \quad x \in \mathbb{R} \quad (1.12)$$

Teorema 16. *Glivenko-Cantelli* Sea $N_n \rightarrow \infty$ cuando $n \rightarrow \infty$. Entonces para cualquier punto fijo F

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |\hat{T}_{N_n}^*(x) - F(x)| = 0 \quad (1.13)$$

Para teorema de Glivenko-Cantelli ver [3]

Capítulo 2

APLICACIÓN DE FRACTALES A FINANZAS

En este momento estamos preparados para desarrollar una aplicación de la teoría estudiada en el capítulo anterior. Lo que pretendemos es encontrar una función de distribución que mejor modele un conjunto de datos y con ella poder encontrar el VAR-Operacional. Esto significa encontrar el percentil 99.9 de los datos que estudiaremos, a través de la función de distribución que estimaremos, para cuantificar el grado de riesgo al que la entidad financiera está expuesta. Es decir encontrar el valor en pérdidas al que está expuesta a sufrir la entidad con una probabilidad de 0.001.

Los datos que contiene la base de datos expresan las pérdidas sufridas por la entidad en los áreas de Fraude interno, fraude externo, y sistemas; dichas pérdidas están cuantificadas en el periodo comprendido de Enero del año 2000 a Diciembre del año 2008.

Para trabajar con los datos y aplicar el análisis fractal reunimos todas las pérdidas en una sola tabla, es decir construimos una tabla de pérdidas totales con las tres áreas de interés, Fraude interno, fraude externo, y sistemas; y trabajamos con ella para encontrar en VAR-operacional de la entidad financiera.

Al conocer la base de datos podemos encontrar la función de distribución empírica de éstos, pero esta es una aproximación, digamos discreta, de los datos; ahora a través de el estimador que hemos desarrollado en el capítulo anterior y aplicaremos en este, la aproximación que haremos nos permitirá conocer el valor de la función de distribución en valores que desconocíamos al inicio, es decir hacemos una estimación continua de los datos.

Aplicaremos la teoría del capítulo anterior a una base de datos de perdidas mensuales en las áreas de:

- 1) Fraude interno: Número de inconsistencia en caja menor.
- 2) Fraude externo: número de cheques con firmas falsificadas, hacking de sistemas de información.

3) Sistema: Número de fallas en el sistema.

Por ejemplo, a continuación presentamos una tabla que muestra las pérdidas asociadas en Fraude Externo: hacking de sistemas de información.

Tabla de fraude externo 2000-2008

4a. Fraude Externo		
Mes	Hacking de sistemas de información	Pérdida Asociada
feb-00	1	\$ 7,462,610
ene-01	1	\$ 5,654,110
feb-01	1	\$ 3,883,620
abr-02	1	\$ 8,065,620
sep-02	1	\$ 2,032,930
oct-02	1	\$ 1,707,430
ago-03	1	\$ 5,000,240
mar-04	1	\$ 4,681,870
nov-04	1	\$ 2,517,920
ago-05	1	\$ 5,602,570
oct-05	1	\$ 25,534,400
may-06	1	\$ 19,340,100
jul-06	1	\$ 9,621,800
oct-07	1	\$ 11,993,500
ene-08	1	\$ 53,953,700
may-08	1	\$ 2,690,690
jun-08	1	\$ 6,916,870
nov-08	1	\$ 3,855,270
dic-08	1	\$ 6,977,950

Fuente: Elaboración propia

Para desarrollar el procedimiento, *IFS*, se transforman los datos de la muestra a el intervalo $[0, 1]$; esto debido a que el método ha sido desarrollado para trabajar con funciones de soporte finito, en este caso funciones de distribución, entre las múltiples alternativas usamos la transformación $f(x) = \left(\frac{x-\text{mín}}{\text{máx}-\text{mín}}\right)^{\frac{1}{2}}$ donde mín y máx, son respectivamente, el mínimo y el máximo de la muestra.

La función f así definida transforma los datos de $[\text{mín}, \text{máx}]$ a datos en $[0, 1]$, esto debido a que es una función continua y $f(\text{mín}) = 0$, $f(\text{máx}) = 1$

Para el desarrollo de la aplicación de esta teoría a un caso concreto nos valdremos de el software R-Studio, programa informático desarrollado para aplicaciones estadísticas, para hacer estimaciones numéricas de los datos y hacer un análisis estadístico de los mismos.

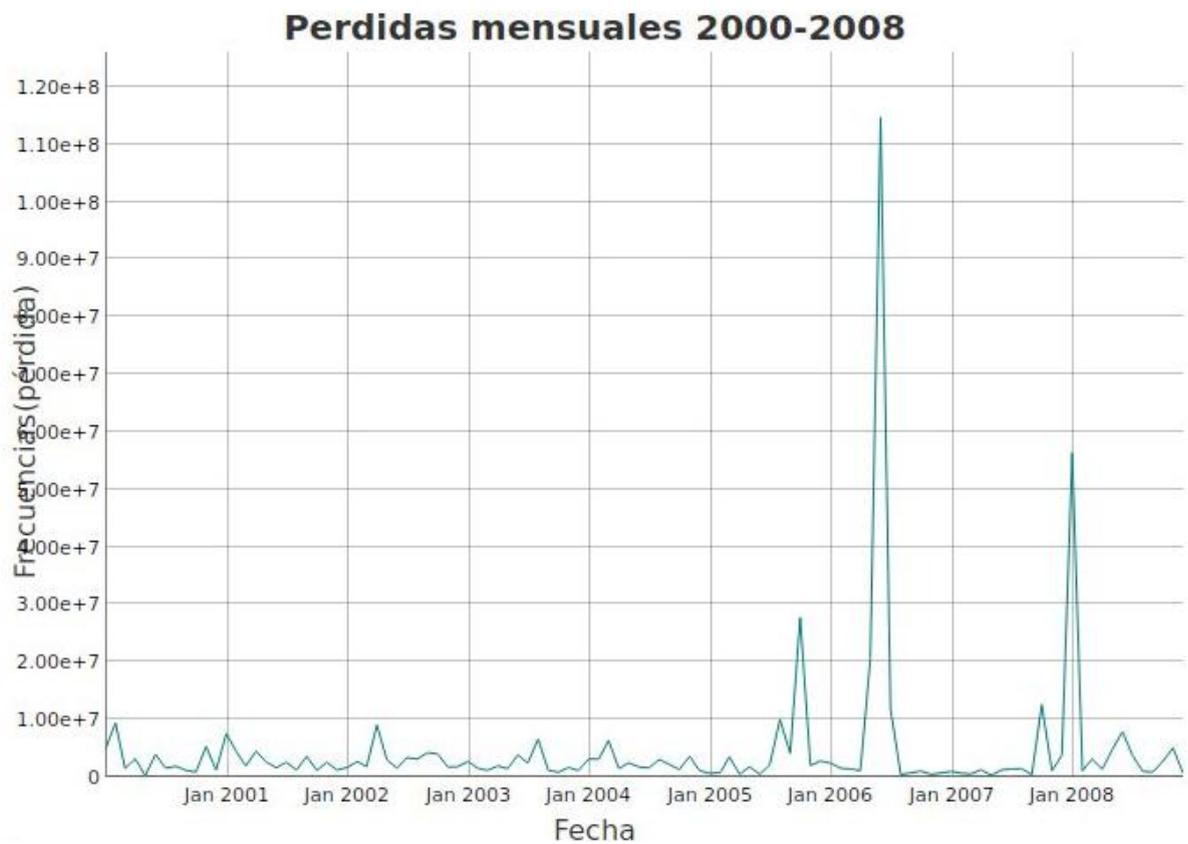
Como tenemos cuatro áreas donde se han tenido pérdidas en los 8 años a analizar, entonces construimos una tabla resumen de las pérdidas totales mensuales tenidas en la empresa, que presentamos a continuación.

Pérdidas totales mensuales 2000-2008

##	Jan	Feb	Mar	Apr	May
## 2000	4720566.9	9389342.8	1488557.3	3163696.2	240462.2
## 2001	7528091.7	4316079.3	1914818.5	4486207.3	2609332.4
## 2002	1634311.3	2688666.1	1779702.4	9041768.2	2998347.9
## 2003	2690431.0	1486728.1	1159835.4	1905096.8	1451992.3
## 2004	3088024.5	3148085.0	6353706.5	1472401.5	2428068.6
## 2005	636173.8	706059.6	3493612.1	462902.1	1758834.9
## 2006	2387470.6	1509856.2	1386588.6	1100408.5	20537989.8
## 2007	959501.1	653178.8	523717.6	1248877.8	298033.3
## 2008	56357645.3	1075864.4	3123679.6	1364401.5	4728136.2
##	Jun	Jul	Aug	Sep	Oct
## 2000	3943336.1	1544034.0	1865173.9	1144671.5	850278.7
## 2001	1613915.5	2540327.9	1221633.4	3586310.5	1135851.2
## 2002	1535143.8	3368980.5	3110483.9	4227353.5	4041043.2
## 2003	3830353.2	2409861.8	6577461.5	1147296.3	825461.1
## 2004	1749400.2	1578612.1	3050330.4	2165384.7	1302521.4
## 2005	467399.2	2095285.8	10012942.8	4104886.3	27725665.5
## 2006	114625709.3	11849265.0	436775.6	711546.0	1014926.9
## 2007	1221112.1	1387284.0	1434429.9	390666.5	12571170.3
## 2008	7899175.6	3782123.8	966685.5	853580.4	2735991.6
##	Nov	Dec			
## 2000	5296861.4	1181022.6			
## 2001	2537418.6	1216094.2			
## 2002	1751863.7	1782043.5			
## 2003	1645904.0	1109802.0			
## 2004	3591160.0	1077254.9			
## 2005	1982035.2	2778488.0			
## 2006	458560.1	703889.5			
## 2007	1082589.4	3798755.4			
## 2008	5043038.7	623933.4			

Fuente:Elaboración propia

A partir de la tabla anterior elaboramos el siguiente gráfico:

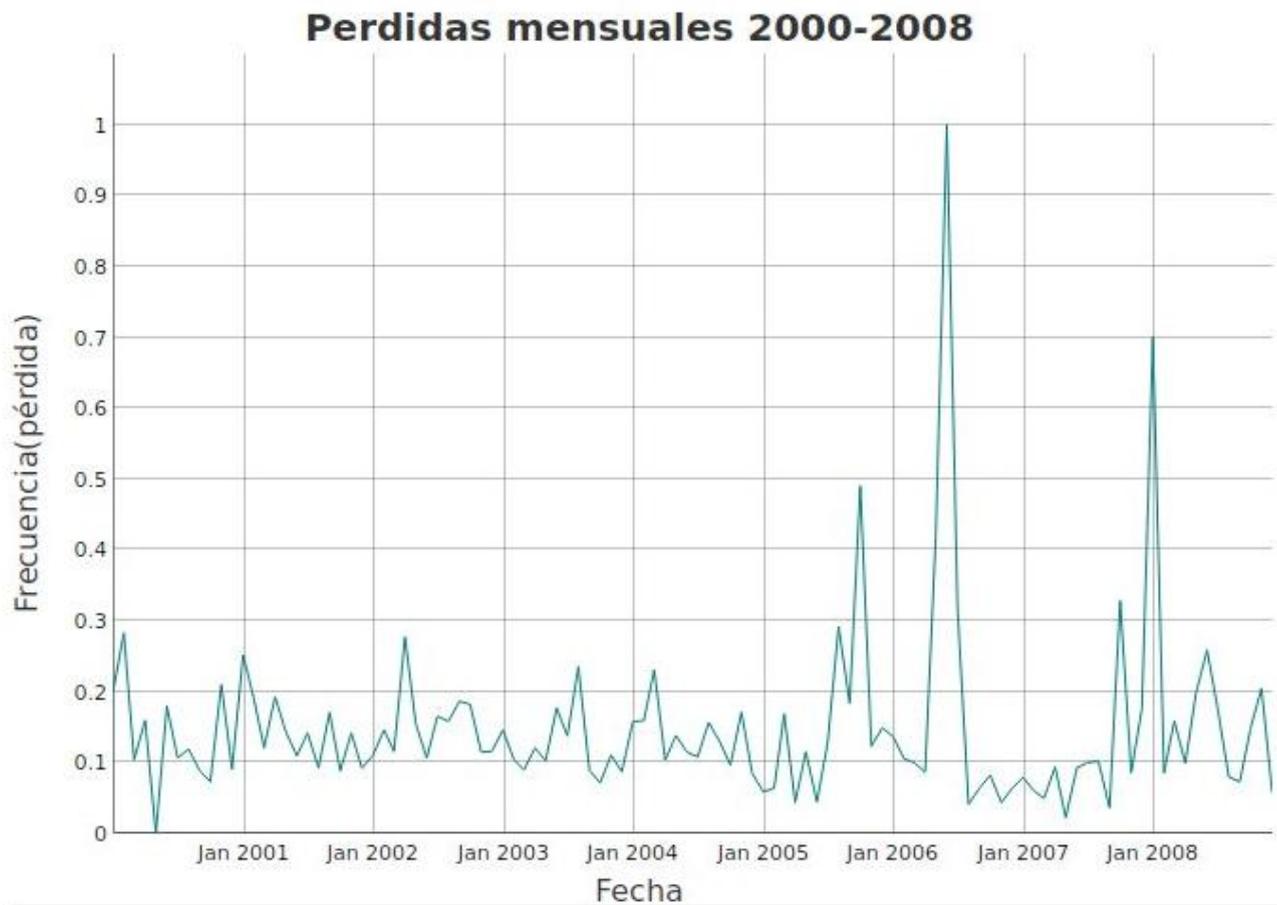


Fuente:Elaboración propia

Para poder trabajar los datos necesitamos tenerlos en el intervalo $[0, 1]$ por lo que les aplicamos la transformación

$$f(x) = \left(\frac{x - \text{mín}}{\text{máx} - \text{mín}} \right)^{\frac{1}{2}}$$

y al graficarlos luego de la transformación tenemos el gráfico siguiente:



Fuente:Elaboración propia

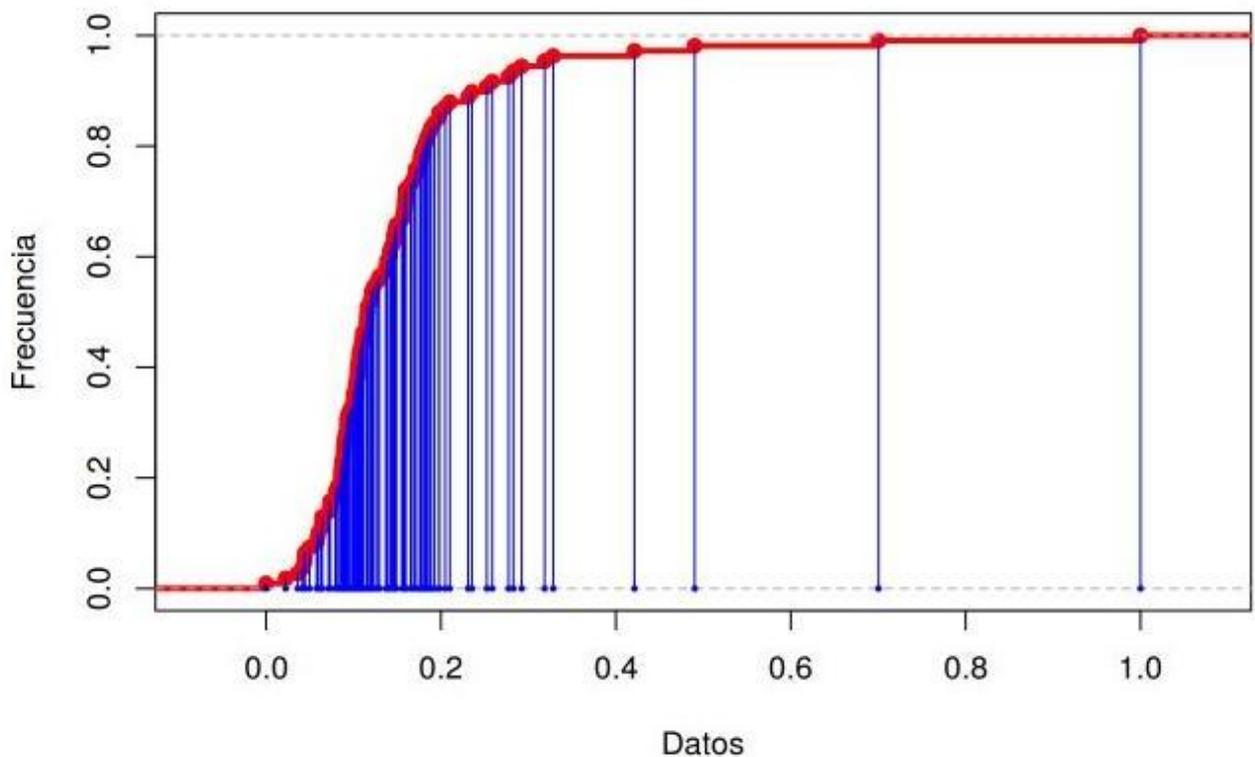
A partir de la muestra de los datos la función de distribución que nos refleja el comportamiento de éstos y que se nos hace fácil conocer es la Distribución acumulativa empírica. hacer notar que la variable que estamos estudiando es:

X : Pérdida total en la entidad financiera en el período comprendido de enero de 2000 a diciembre de 2008.

Esta variable es considerada continua, pues las pérdidas que puede estar sufriendo la entidad pueden suceder en cualquier momento.

Al graficar la función de distribución empírica para esta variable tenemos:

Función de distribución acumulativa empírica para la variable pérdida



Fuente:Elaboración propia

Con este gráfico tenemos una aproximación a la función de distribución que siguen los datos de la muestra. Ahora lo que haremos será obtener la función de distribución por medio del operador:

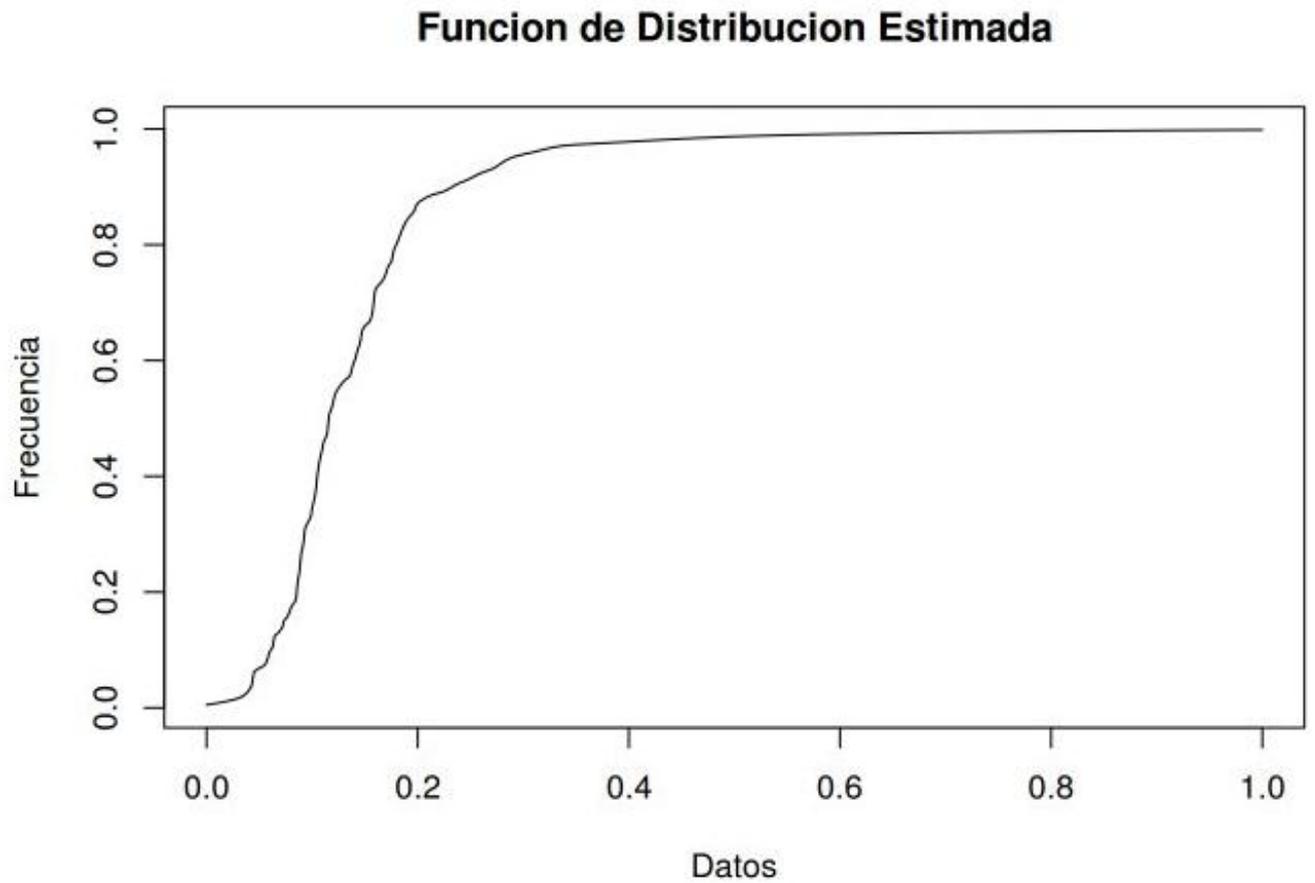
$$\hat{T}_N \mu(x) = \sum_{i=1}^N \frac{1}{N} \cdot \mu \left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i} \right) \quad (2.1)$$

Notar que la función $\mu \in F([0, 1])$ que aparece en este operador es arbitraria, esto significa que podemos utilizar "cualquiera" para hacer la estimación de la función de distribución. El principio en el que se fundamenta la funcionalidad de este operador es el Teorema de Banach, lo que significa que la función de distribución que buscamos es el punto fijo de éste operador.

Sabemos que el punto fijo es el límite cuando iteramos infinitamente el operador, pero este es un proceso (iterar indefinidamente) que en la realidad no es posible, por lo que fijamos un número finito de iteraciones que haremos con en el operador y hacemos variar la función de distribución μ para conocer la mejor estimación de la función de distribución de las pérdidas en la entidad financiera.

Para decidir con cual estimación nos quedamos para hacer el análisis de las pérdidas en la entidad, debemos decir cuál es la mejor estimación hecha por el operador. Hasta este momento lo que conocemos es la función de distribución empírica de los datos y ésta será nuestro parámetro para decidir con cual estimación haremos el análisis. Lo que haremos será calcular la estimación con diversas funciones de distribución μ en el operador y calcularemos el Error cuadrático medio (ECM) entre dicha estimación y la función de distribución empírica de los datos, esto debido a que la función de distribución estimada debe seguir el comportamiento de la función de distribución empírica de los datos.

Al desarrollar y graficar el operador que nos proporciona la estimación de la función de distribución desconocida, usando la distribución normal, tenemos el siguiente gráfico:



Fuente:Elaboración propia

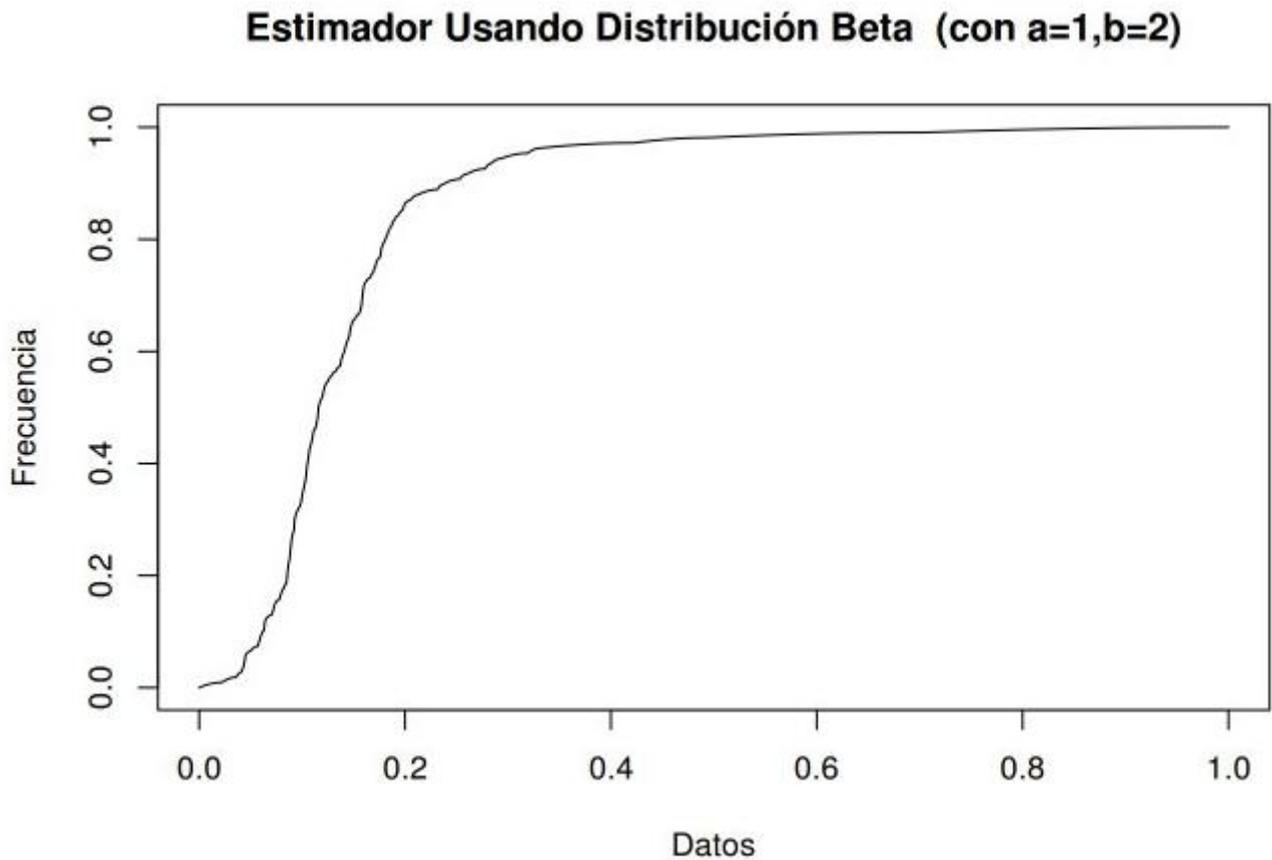
La función de distribución beta es una familia de distribuciones de probabilidad continua con dos parámetros a, b , cuya función de densidad para valores $0 \leq x \leq 1$ es:

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}$$

donde

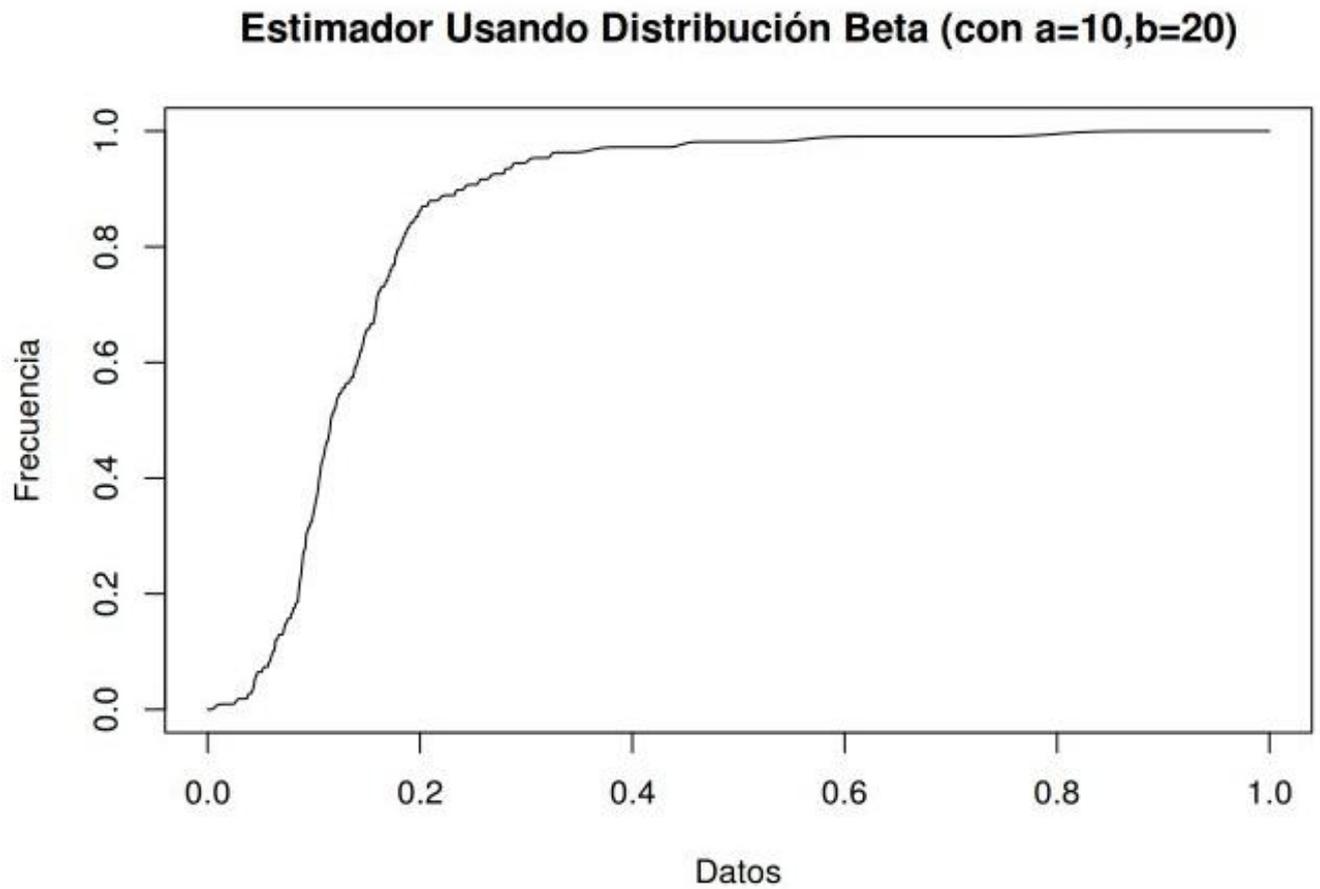
$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

es la función gamma.



Fuente:Elaboración propia

Hacemos una estimación de la distribución de pérdidas, usando $\text{Beta}(a = 10, b = 20)$.



Fuente:Elaboración propia

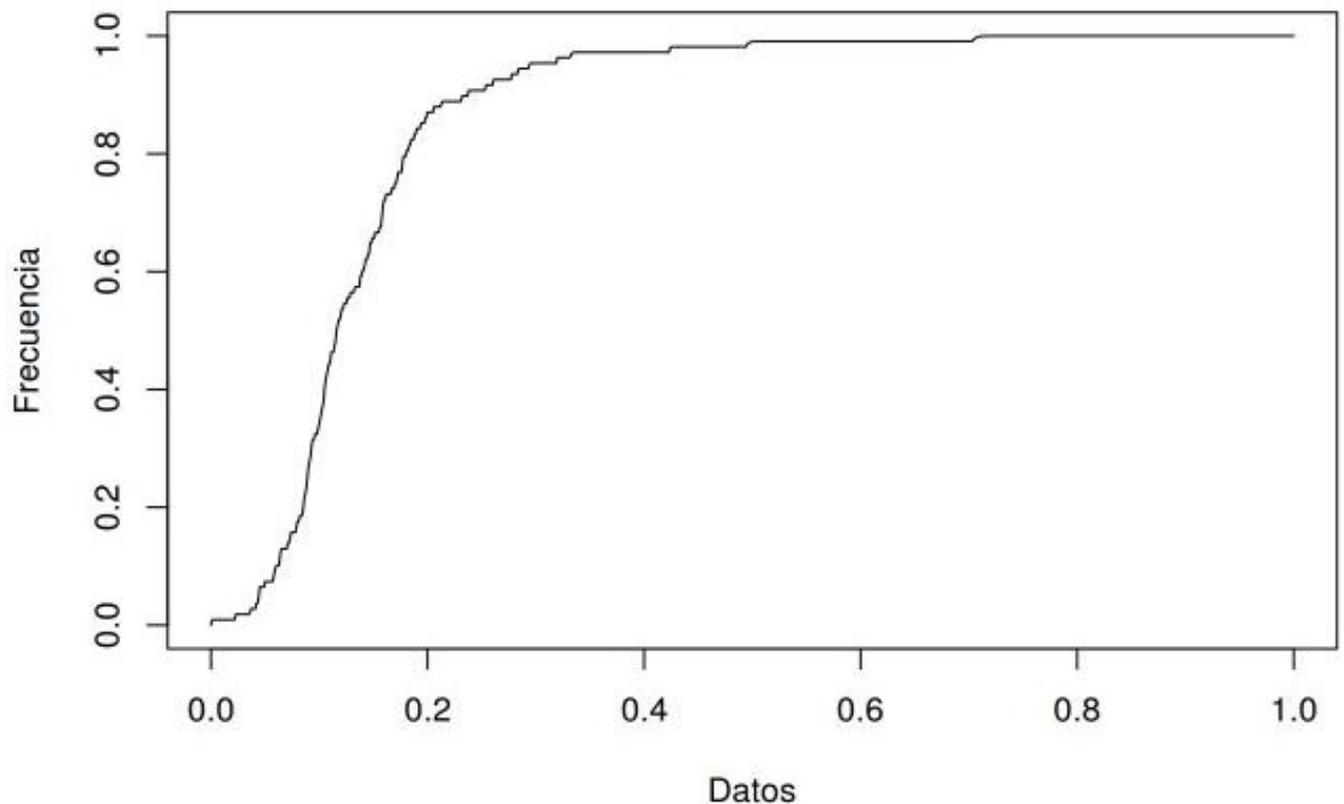
La distribución exponencial es una distribución de probabilidad continua con un parámetro $\lambda \geq 0$ cuya función de densidad es:

$$f(x) = P(x) = \begin{cases} \lambda e^{-\lambda x} & \text{para } x \geq 0 \\ 0 & \text{otro caso} \end{cases} \quad (2.2)$$

Su función de distribución acumulada es:

$$f(x) = P(X \leq x) = \begin{cases} 0 & \text{para } x < 0 \\ 1 - \lambda e^{-\lambda x} & \text{para } x \geq 0 \end{cases} \quad (2.3)$$

Estimador Usando Distribución Exponencial (lambda=100)



Fuente:Elaboración propia

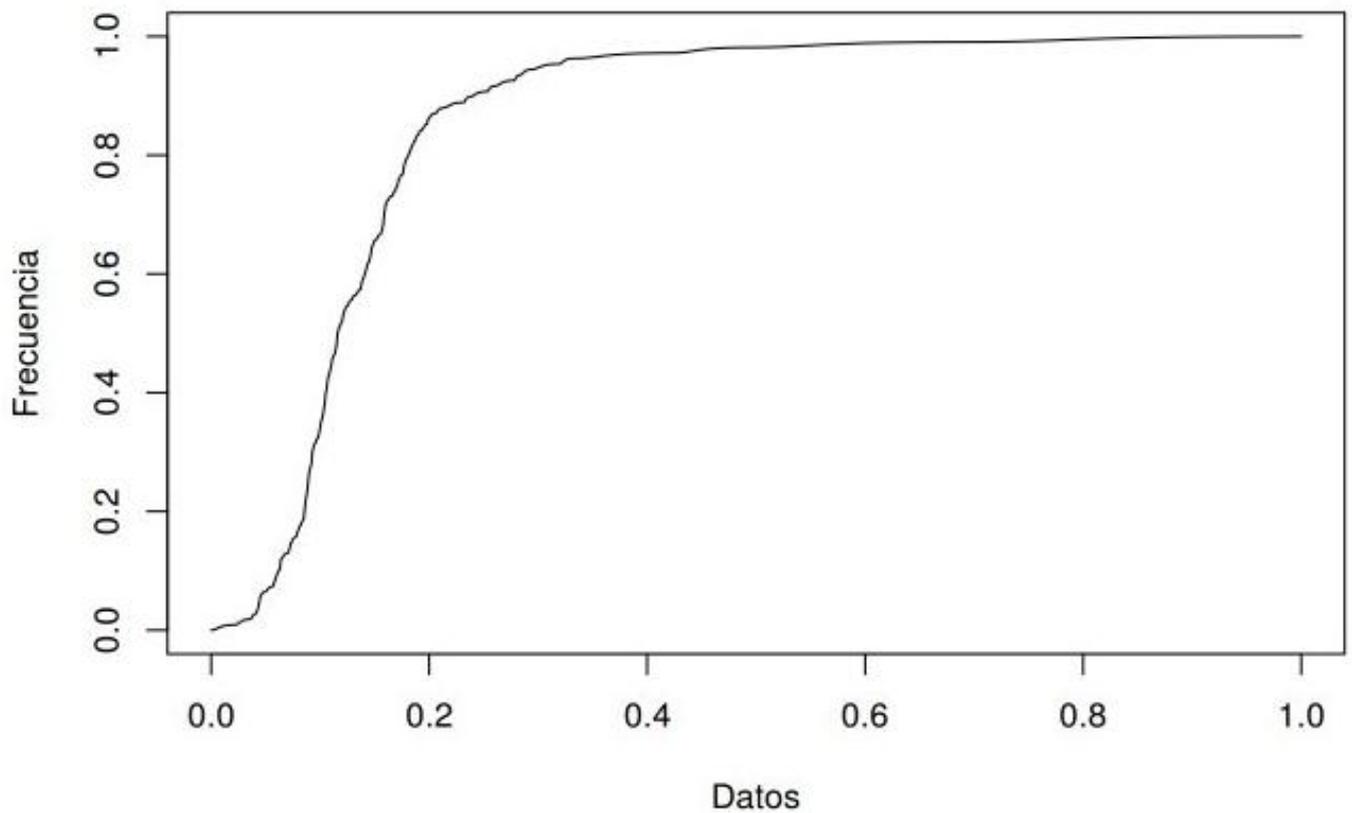
La función de distribución Gamma es una distribución de probabilidad continua con dos parámetros k y λ , cuya función de densidad es dada por:

$$f(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{k-1}}{\Gamma(k)}$$

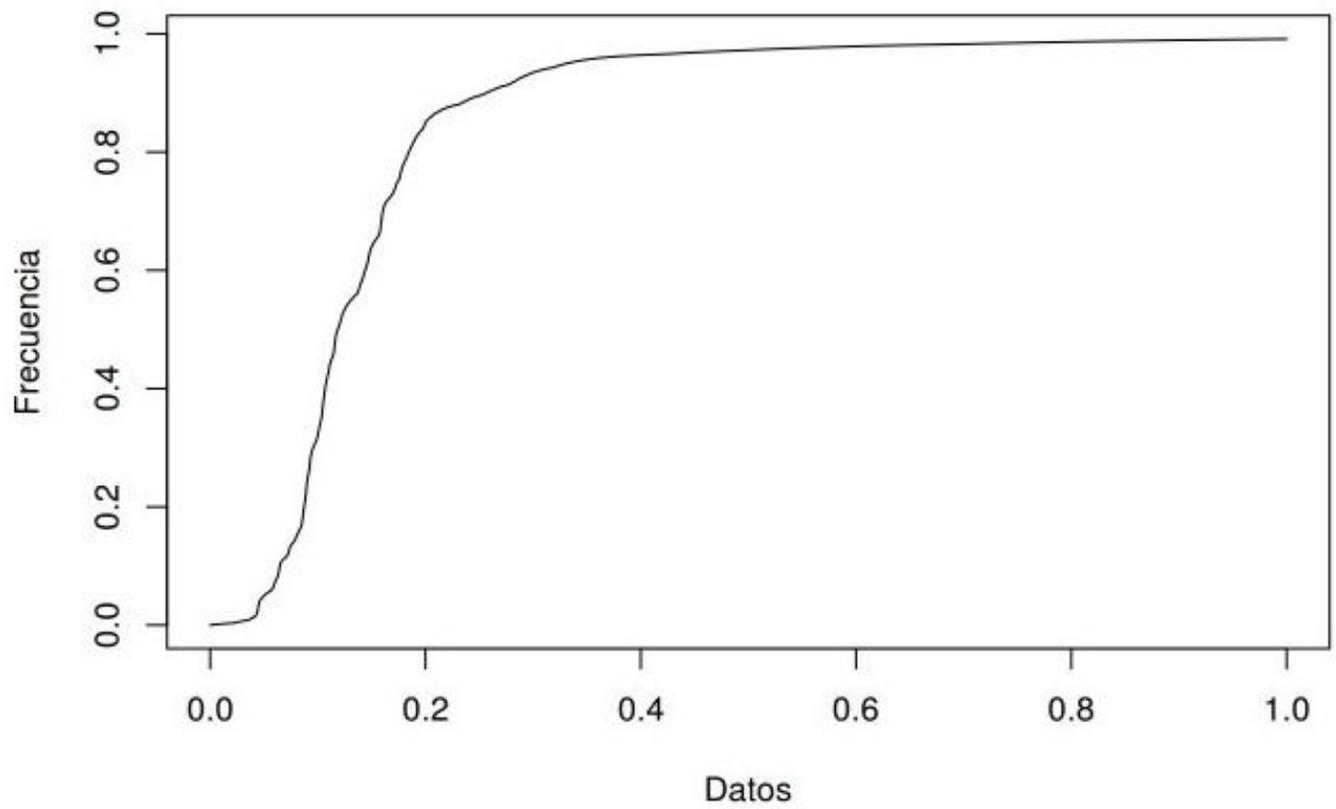
donde Γ es la función Gamma dada por:

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

Estimador Usando Distribución Gamma (k=3,lambda=8)

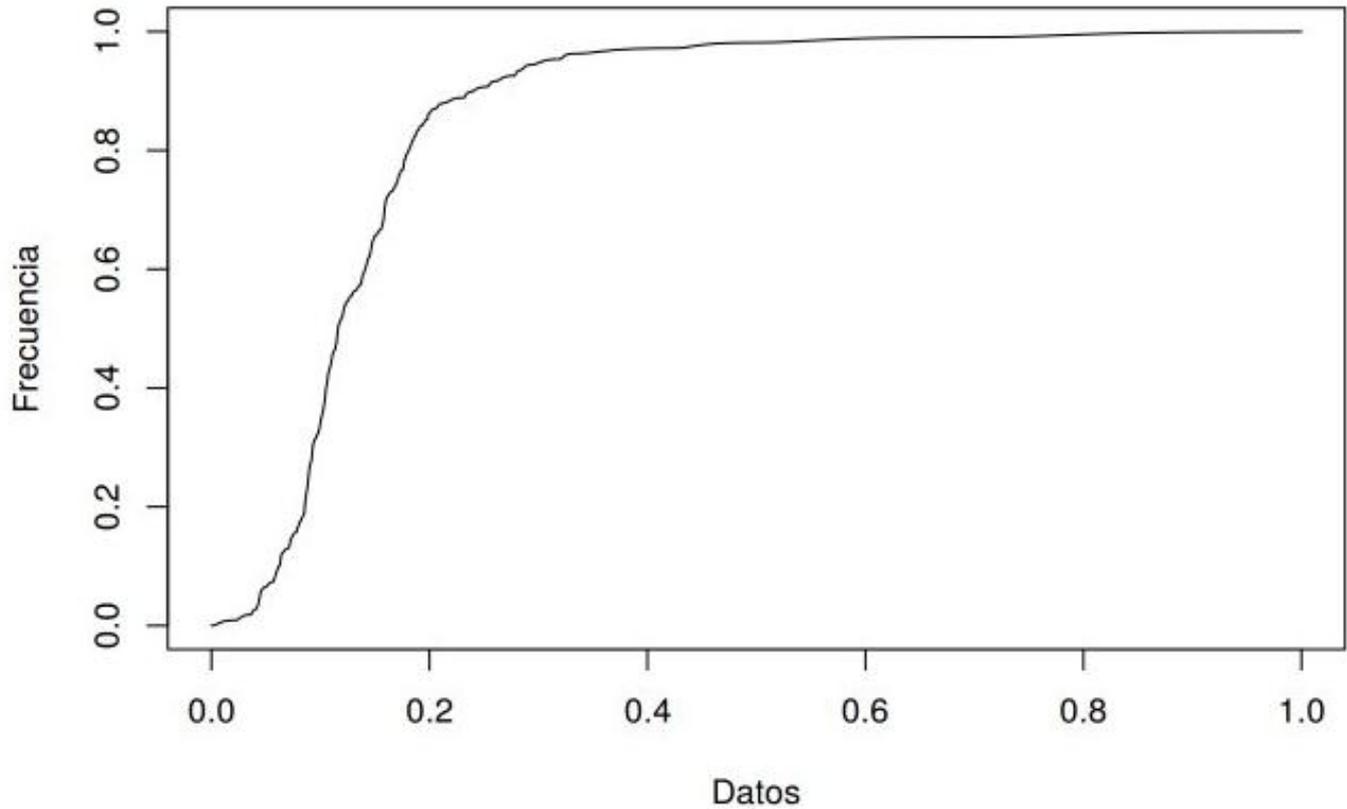


Fuente:Elaboración propia

Estimador Usando Distribución Gamma ($k=0.5, \lambda=1$)

Fuente:Elaboración propia

Estimador Usando Distribución Poisson (lambda=8)



Fuente:Elaboración propia

Hemos hecho varias estimaciones de la función de distribución de pérdidas por medio del estimador:

$$\hat{T}_N \mu(x) = \sum_{i=1}^N \frac{1}{N} \cdot \mu \left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i} \right) \quad (2.4)$$

variando μ y haciendo una iteración en el operador.

Podemos notar como se aproximan éstas estimaciones de la función de distribución de pérdidas a la distribución empírica de los datos. Lo que falta es definir cual es la mejor estimación, que lo haremos utilizando el Error cuadrático medio (ECM).

Antes de encontrar la mejor estimación de la función de distribución de pérdidas, haremos una simulación de un conjunto de datos que a priori sabemos que función de distribución los modela. Esto nos ayudará a visualizar de una mejor manera la efectividad del operador.

La función de distribución beta(20,10) tiene el siguiente grafico:

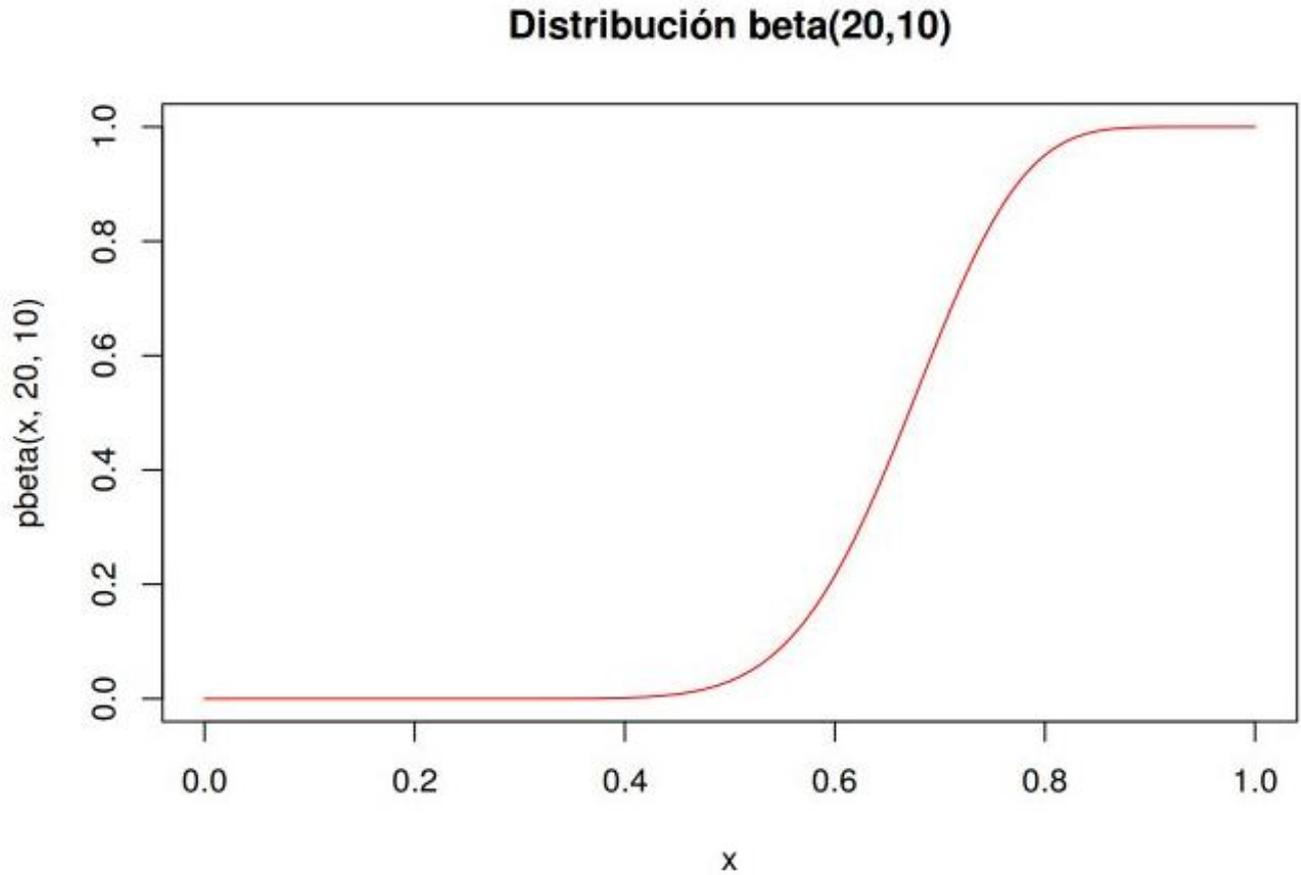


Figura 2.1: Fuente:Elaboración propia

Lo que pretendemos es hacer una aproximación de esta función usando el operador

$$\hat{T}_N \mu(x) = \sum_{i=1}^N \frac{1}{N} \cdot \mu \left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i} \right) \quad (2.5)$$

Para hacer el estudio, como ya expusimos, usamos el Software R. La idea es la siguiente: crear un vector de números de manera aleatoria que sigan la distribución beta(20,10), luego aplicarles el procedimiento de los *IFS* por medio del operador construido y graficar dicha estimación, y finalmente comparar con la función de distribución original (beta(20,10)).

El operador $\hat{T}_N \mu(x)$ necesita que le definamos una función de distribución μ , en este caso le definimos la función de distribución exponencial con parámetro 1 (exp(1)).

Generamos dos vectores de números, de longitud 100 y 500, que siguen la distribución $\text{beta}(20, 10)$.

El siguiente código elaborado en R, crea los vectores de datos de longitud 100 y 500 que siguen una distribución $\text{beta}(20, 10)$, grafica la estimación y compara con la gráfica original.

```
#muestreo de beta(20,10)
h100<-rbeta(100,20,10) # generamos length(h)=100 numeros que siguen la
  distribucion beta(20,10)
d100<-quantile(h100,probs = seq(0,1,(1/length(h100)))) # calculamos los
  cuantiles de orden (1/length(h)) h

# un muestreo aleatorio de beta ,
h500<-rbeta(500,20,10) # generamos length(h)=500 numeros que siguen la
  distribucion beta
d500<-quantile(h500,probs = seq(0,1,(1/length(h500)))) # calculamos los
  cuantiles de orden (1/length(h)) de h

#Operador para d100
T100<-function(x,y=d100){
  #y<-cuantiles
  g1<-rep(0,(length(y)-1))
  for(i in 1:(length(y)-1)){
    g1[i]<-(1/(length(y)-1))*pexp(((x-y[i])/(y[i+1]-y[i])),1)
  }
  return(sum(g1))
}

ps<-seq(0,1,0.0001)
sp<-rep(0,length(ps))
for(i in 1:length(ps)){
  sp[i]<-T100(ps[i],y=d100)
}

plot(ps,sp,type = "l",col="green",main="FDA (Aproximacion de beta a traves de
  la exponencial(1))",xlab = "Datos",ylab = "Frecuencia")
#superponer graficos
plot(ps,sp,type = "l",col="green",xlab = "Datos", ylab = "Frecuencia",main="
  Superponiendo graficos")
curve(pbeta(x,20,10),col="red",add = T)

#Operador para d500
T33<-function(x,y=d500){
  #y<-cuantiles
  g1<-rep(0,(length(y)-1))
  for(i in 1:(length(y)-1)){
    g1[i]<-(1/(length(y)-1))*pexp(((x-y[i])/(y[i+1]-y[i])),1)
  }
  return(sum(g1))
}

sss<-seq(0,1,0.0001)
e<-rep(0,length(sss))
for(i in 1:length(sss)){
  e[i]<-T33(sss[i],y=d500)
}
```

```
}  
plot(sss,e,type = "l",col="blue",main="FDA (Aproximacion de beta a traves de la  
    exponencial(1))",xlab = "Datos",ylab = "Frecuencia")  
#superponer graficos  
plot(sss,e,type = "l",col="blue",xlab = "Datos", ylab = "Frecuencia",main="  
    Superponiendo graficos")  
curve(pbeta(x,20,10),col="red",add = T)
```

Al hacer el proceso fractal para el vector de 100 valores y graficarlo tenemos:

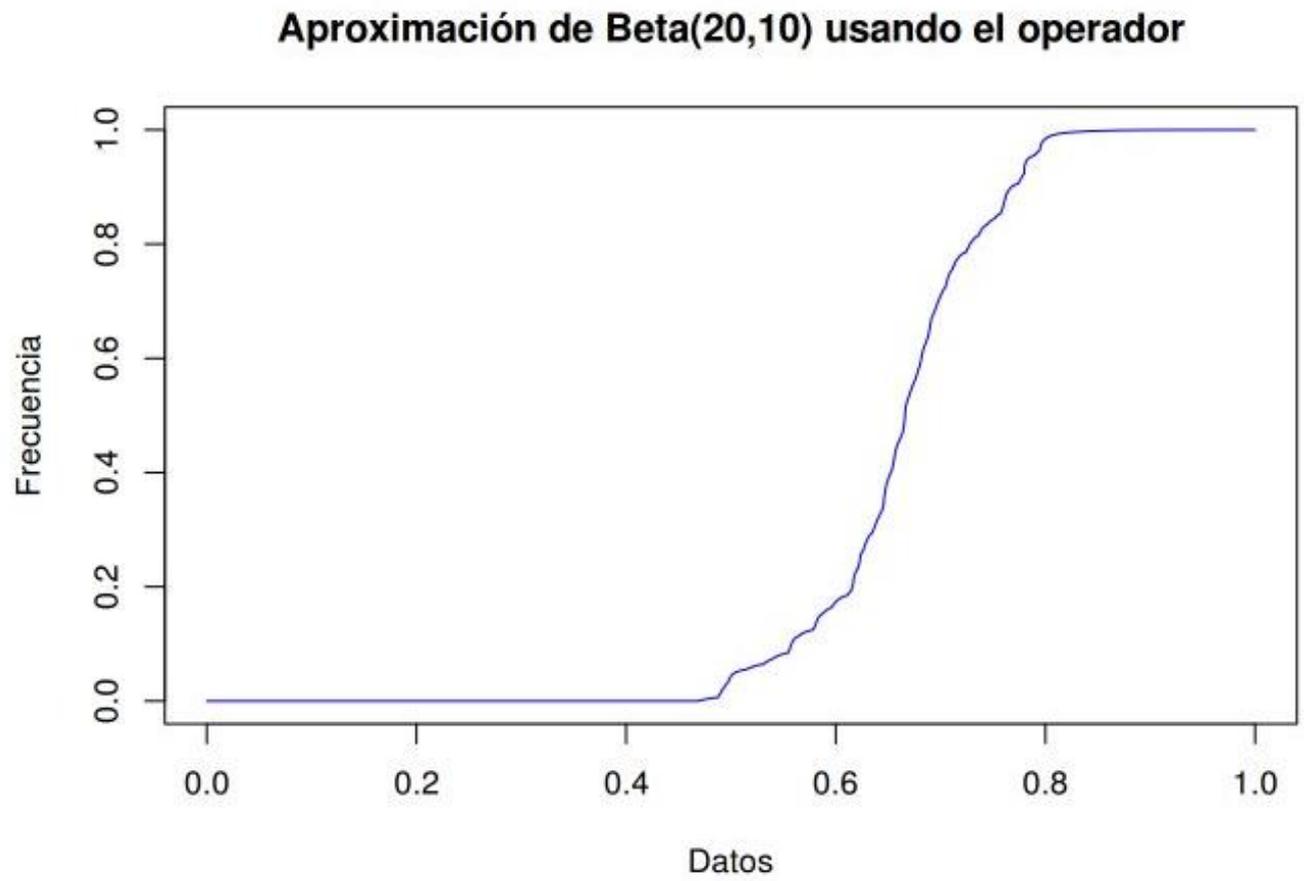


Figura 2.2: Fuente:Elaboración propia

Veamos los dos gráficos, el de $\beta(20, 10)$ y la estimación hecha por el operador con el vector de 100 valores, en un mismo sistema de coordenadas.

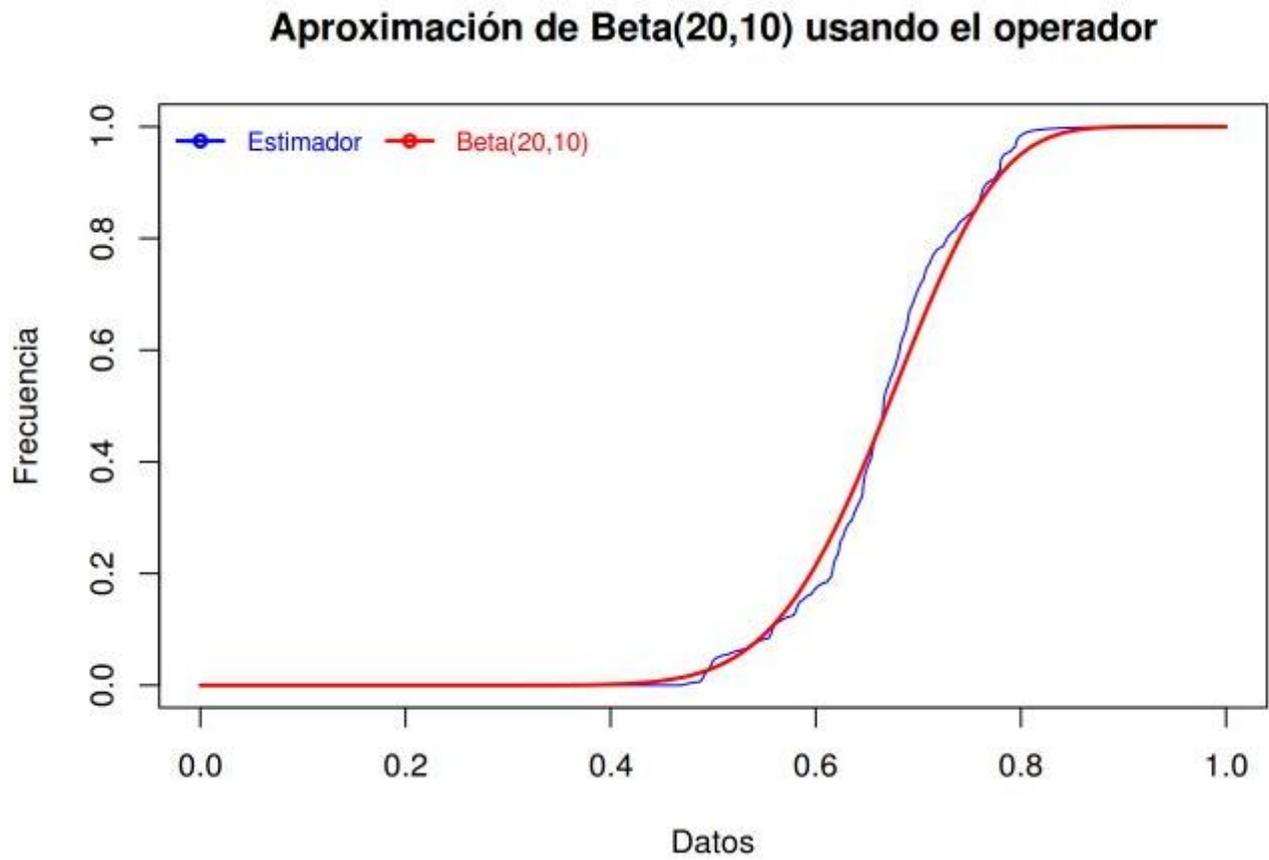


Figura 2.3: Fuente:Elaboración propia

Ahora al hacer el proceso fractal para el vector de 500 valores y graficarlo tenemos:

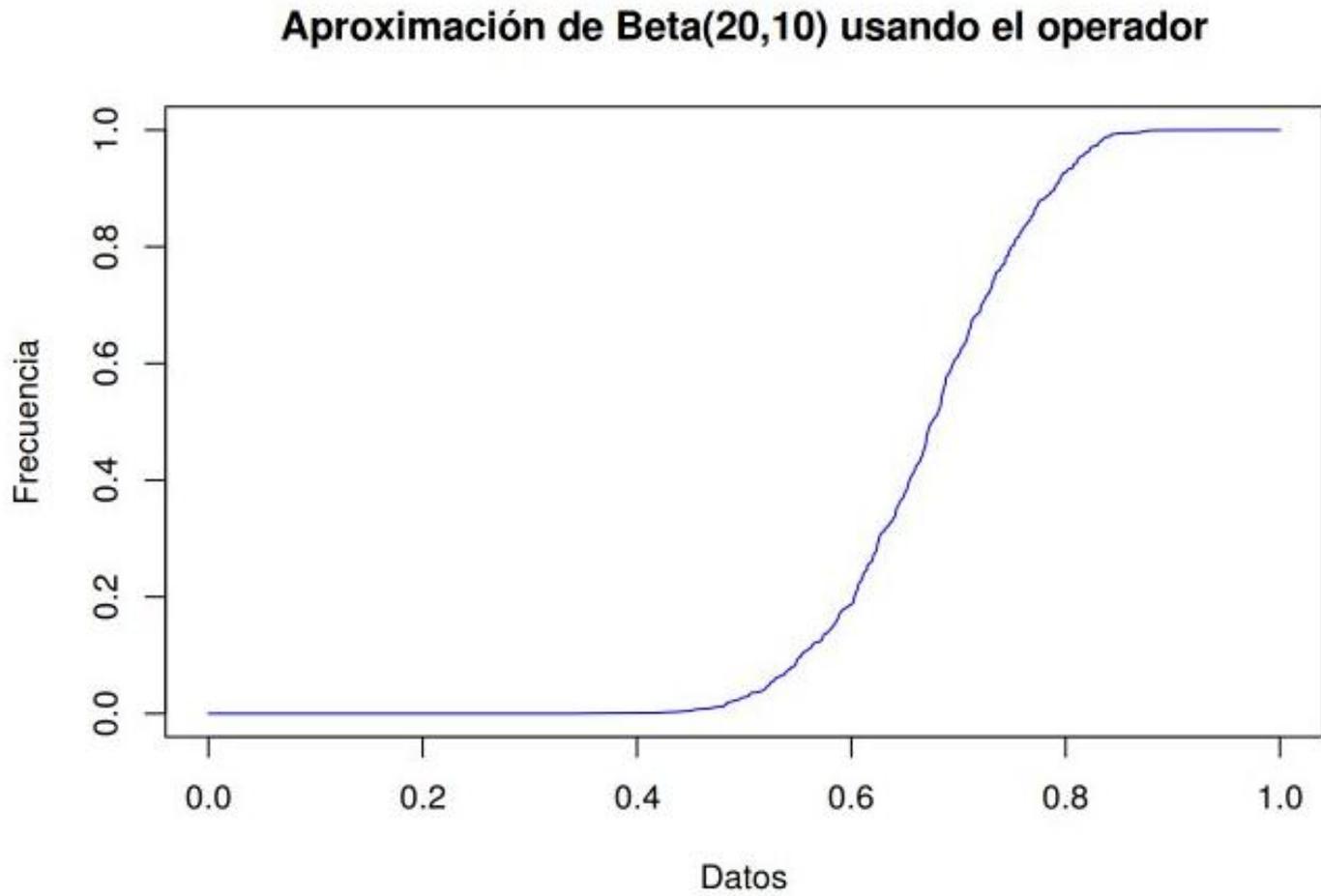


Figura 2.4: Fuente:Elaboración propia

Veamos los dos gráficos, el de $beta(20, 10)$ y la estimación hecha por el operador con el vector de 500 valores, en un mismo sistema de coordenadas.

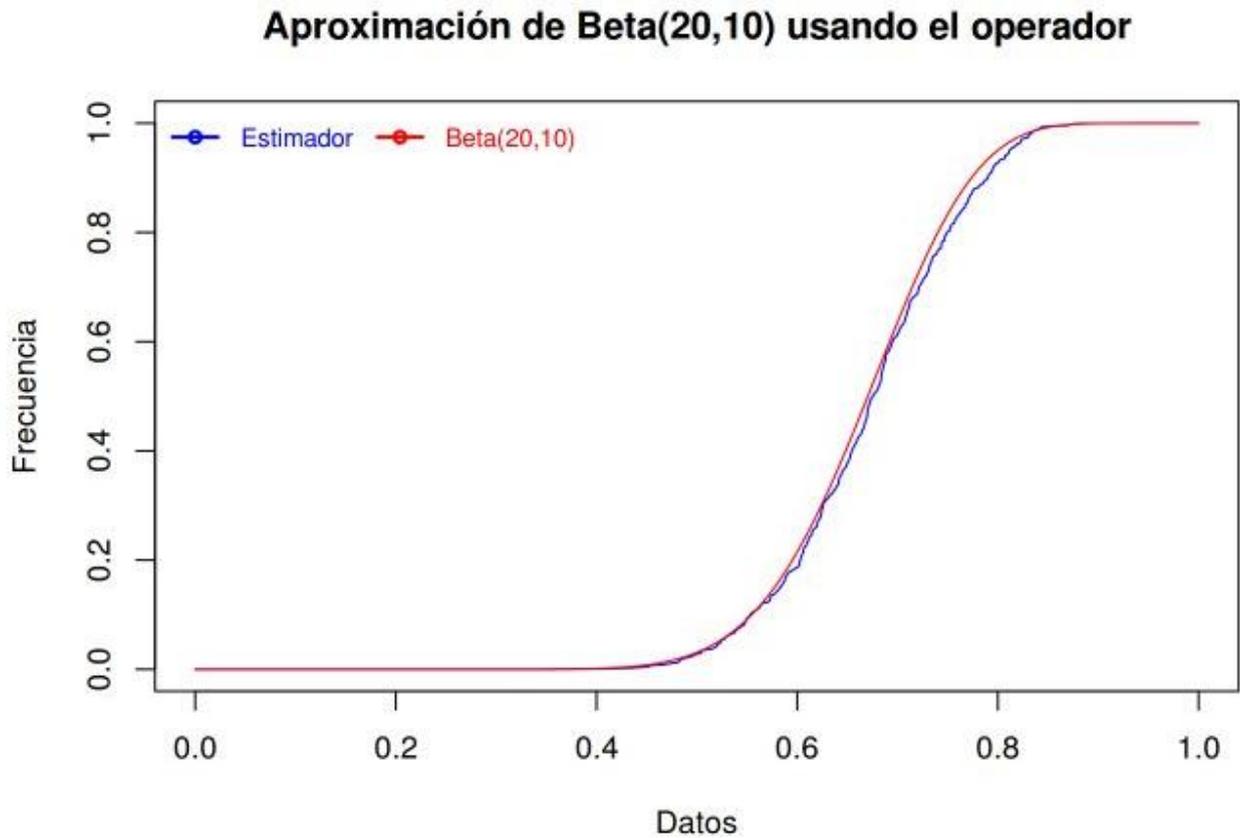


Figura 2.5: Fuente:Elaboración propia

Lo anterior nos muestra la efectividad de el método y como la aproximación va mejorando a medida que aumentamos la cantidad de los datos muestrados. Cabe resaltar que en este proceso estamos trabajando con una sola iteración del operador.

Luego de haber hecho la simulación para $beta(20, 10)$ y obtener una aproximación por medio de el operador, a través del proceso fractal, pasamos a encontrar cual es la mejor estimación hecha por el operador para la función de distribución de pérdidas de la entidad financiera analizada.

Si nos detenemos a analizar un momento nos damos cuenta que hay estimadores que son muchos mejores que otros, es decir; en la ecuación:

$$\hat{T}_N \mu(x) = \sum_{i=1}^N \frac{1}{N} \cdot \mu \left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i} \right). \quad (2.6)$$

Se utiliza la función $\mu \in F([0, 1])$ sin ninguna restricción. Como el estimador funciona para cualquier $\mu \in F([0, 1])$ para muestras grandes, según el teorema 16 (Glivenko-Cantelli), entonces para muestras pequeñas, como en este caso, los resultados varían según la función μ utilizada. Además recordar que hemos hecho solo una iteración en el operador y con ello hemos encontrado la estimación, nos sería irrelevante cual función μ utilizáramos si el proceso de iteración lo desarrolláramos infinitamente, lo que sabemos que humanamente es imposible. Entonces para elegir la mejor aproximación comparamos el estimador (\hat{T}_N) con la función de distribución empírica (\hat{F}_N), a través de el Error Cuadrático Medio (ECM), donde:

$$ECM = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.7)$$

donde

Y : variable a predecir.

\hat{Y} : variable predictora.

A continuación estudiamos el ECM de una lista de funciones de distribución:

Distribución	ECM
$Unif[0, 1]$	$2.56093e - 05$
$poisson(\lambda = 8)$	0.001692901
$Gamma(k = 0.5, \lambda = 1)$	$7.304158e - 05$
$Beta(a = 10, b = 20)$	$4.53853e - 05$
$Beta(a = 1, b = 2)$	$3.846762e - 05$

Donde el ECM fue estimado con el siguiente código:

```

yu # aproximacion hecha por la distribucion uniforme
w1<-ecdf(a)(xu)
w1 # aproximacion de la distribucion empirica
mean((w1-yu)^2) # ECM
w2<-ecdf(a)(lm1)
w2 # aproximacion de la distribucion empirica
lm2 # aproximacin hecha con la poisson (lanmda=8)
mean((w2-lm2)^2) #ECM

w3<-ecdf(a)(m1)
w3 # aproximacion de la distribucion empirica

```

```

m2 # aproximacion hecha por Distribucion Gamma (k=0.5,lambda=1)
mean((w3-m2)^2) #ECM

w4<-ecdf(a)(x12)
w4 # aproximacion de la distribucion empirica
hh12 #aproximacion hecha por Distribucion Beta (con a=10,b=20)
mean((w4-hh12)^2) # ECM

w5<-ecdf(a)(x1) # aproximacion de la distribucion empirica
hh1 # aproximacion hecha por Distribucion Beta (a=1,b=2)
mean((w5-hh1)^2) #ECM

```

Observamos que la distribución que mejor funciona es la Uniforme en $[0, 1]$, esto porque es la que menor ECM tiene respecto a la distribución empírica. Así que esta será la distribución que utilizaremos para hacer la estimación de el valor operacional.

Para ello necesitamos calcular el percentil 99.9 de la distribución, es decir resolver la ecuación:

$$0.999 = \hat{T}_N \mu(x) = \sum_{i=1}^N \frac{1}{N} \cdot \mu \left(\frac{x - \hat{q}_i}{\hat{q}_{i+1} - \hat{q}_i} \right) \quad (2.8)$$

donde μ es la distribución Uniforme en $[0, 1]$.

Resolviendo esta ecuación en R, tenemos que el percentil 99.9 es 0.9679457. Para ello basta ejecutar el siguiente código:

```

vo<-function(x,y=cuantiles){
  j<-(0.999-Tu(x,y=cuantiles))
  return(j)
}
uniroot(vo,interval=c(0,1),tol=10^(-100),maxiter=1000)

```

Tenemos $x = 0.9679457$, este valor lo transformamos al intervalo [máx, mín], donde máx y mín son el máximo y mínimo, respectivamente, de los datos que estamos analizando, y este valor es el valor operacional buscado.

Así el resultado conduce, según las transformaciones utilizadas, a que el percentil 99.9 estimado para la distribución de pérdidas totales mensuales es de 107, 410, 160.

Esto indicaría que para la entidad analizada el *VAR* mensual estimado por esta metodología, bajo el supuesto de que no se provisionen las pérdidas esperadas, estaría dado por: \$107, 410, 160.

Esto significa que solamente hay 1 posibilidad de 1000, en condiciones normales de mercado, de que haya una pérdida superior a los \$107, 410, 160.

Como conclusión podemos decir que este método es bastante efectivo en modelar la función de distribución de probabilidad, prueba de ello es el análisis que se le hizo a la función *beta*(20, 10).

Como recomendaciones sugerimos tratar el operador IFS en funciones de soporte mas general, es decir en un intervalo $[a, b]$, para tratar de explicar funciones mas generales.

También proponemos exponer a mayor detalle el desarrollo teórico del operador en conjuntos abstractos, ya que este esta relacionado con teoría de la medida.

Otra recomendación es hacer un estudio comparativo entre el método aquí presentado, el operador fractal, con otras metodologías con la intención de ver la metodología que mejor funcione.

Capítulo 3

FRACTALES Y VULCANOLOGÍA

3.1. El concepto de dimensión de recuento de cajas.

La noción de dimensión es fundamental en geometría fractal. Las dimensiones fractales extienden, a las clases anchas de sistemas, de que una línea recta lisa es 1-dimensional, una superficie es bidimensional y así sucesivamente.

Aproximadamente, una dimensión indica, de alguna manera, cuanto espacio un conjunto ocupa cerca de cada uno de sus puntos.

Fundamental para la mayoría de las definiciones de dimensión fractal es la idea de "medición de un conjunto a escala δ ". Para cada δ , medimos un conjunto de una manera que detectara irregularidades de tamaño δ , y vemos como estas medidas se comportan cuando $\delta \rightarrow 0$.

En este trabajo se estudia principalmente la dimensión de conteo de cajas, que tiene una definición intuitiva y sencilla y es una de las dimensiones más utilizadas. Esta definición se remonta al menos a la década de 1930.

Sea F cualquier subconjunto no vacío de \mathbb{R}^n y sea $N_\delta(F)$ el menor número de conjuntos de diámetro menor o igual a δ , los cuales cubren F . Sea $\{\delta_k\}$ una sucesión de números positivos que cumple dos condiciones:

1. $\lim_{k \rightarrow \infty} \delta_k = 0$
2. $\lim_{k \rightarrow \infty} \frac{\log \delta_k}{\log \delta_{k-1}} = 1$

Definición 17. *Las dimensiones de recuento de cajas inferior y superior son respectivamente:*

$$\underline{\dim}_B F = \underline{\lim}_{k \rightarrow \infty} \frac{\log(N_{\delta_k} F)}{-\log \delta_k}$$
$$\overline{\dim}_B F = \overline{\lim}_{k \rightarrow \infty} \frac{\log(N_{\delta_k} F)}{-\log \delta_k}$$

*Si estos son iguales, se conoce al valor común como la **dimensión de recuento de cajas** o **dimensión de caja de F***

$$\dim_B F = \lim_{k \rightarrow \infty} \frac{\log(N_{\delta_k} F)}{-\log \delta_k}$$

Malla de cuadrados de lado δ

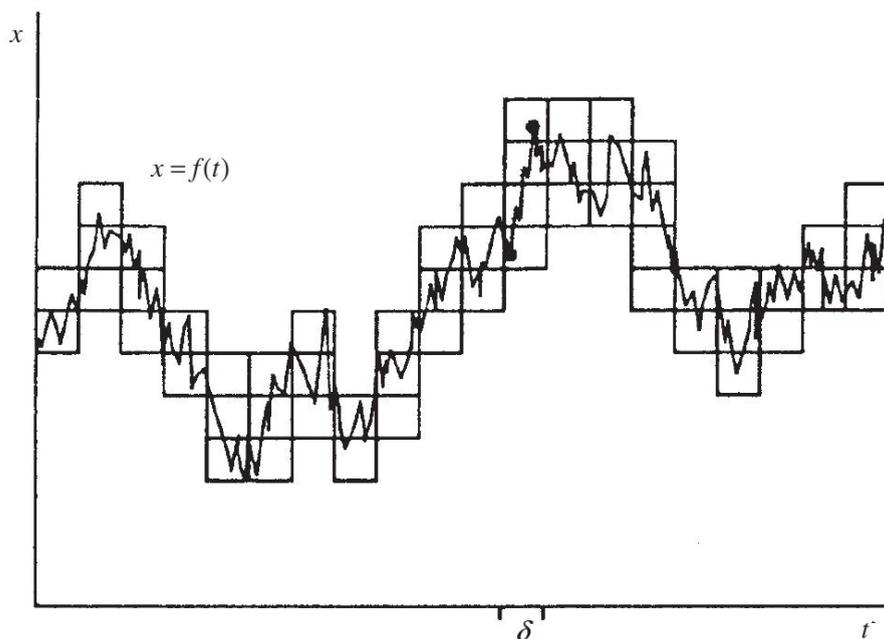


Figura 3.1: Fuente: [11]

Las medidas de dimensión fractal se derivan de los fractales que están bien definidos (matemáticamente), como por ejemplo conjunto de Cantor, triángulo de Sierpinski, etc. Sin embargo, muchos fenómenos del mundo real exhiben propiedades fractales. Entonces suele ser útil calcular la dimensión fractal de un conjunto de datos muestreados. La medida de la dimensión fractal de una serie de tiempo no puede ser calculada exactamente, por lo que tiene que ser estimada, procurando hacer una buena estimación de ella. Las estimaciones de la dimensión son muy sensibles a la cantidad de datos muestreados.

Una propiedad fractal puede ser espacial, puede ser temporal, como en una serie de datos tomados en un sistema a lo largo de un intervalo de tiempo definido.

Series

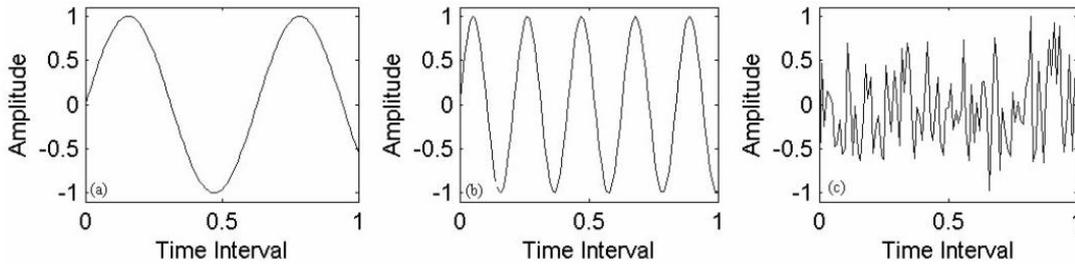


Figura 3.2: Fuente: [11]

Consideremos la gráfica de funciones como se muestra en la figura 3.2, una curva sinusoidal y una curva aleatoria irregular. Al superponer estas gráficas en un mismo sistema de ejes coordinado es evidente que la curva irregular llena más el espacio que las dos primeras series. Una definición utilizada en la práctica de la dimensión fractal es una medida de esta propiedad de llenado del espacio. Observamos que hasta este momento no hemos hablado de la auto-similitud de la curva irregular. El análisis de la dimensión fractal se realiza porque da una medida de la conveniencia de describir la complejidad estructural de los objetos.

La complejidad fractal de las señales en el dominio del tiempo se calcula utilizando los métodos de Katz y Sevcik. En el dominio del tiempo el método parece ser simple y puede ser utilizado en muchas aplicaciones.

La dimensión fractal así calculada es una medida de la complejidad de la curva que representa la señal en un plano. La complejidad se refiere al grado de llenado de espacio en el plano 2D. La complejidad de una señal puede caracterizarse por su FD (dimensión fractal) directamente en el espacio del tiempo. Generalmente, la complejidad de la señal puede analizarse en el dominio del tiempo, en el dominio de la frecuencia o en el espacio de fase del sistema que generó la señal. El análisis en el dominio de la frecuencia requiere la transformada de Fourier de la señal, mientras que el análisis en el espacio de fase requiere la incorporación de los datos en un espacio dimensional más alto. La estimación de la FD que se adopta aquí se deriva de una operación directa sobre la señal y no en cualquier espacio de fase. Esto significa que la serie de datos no tiene que ser incrustada en un espacio de mayor dimensión para la estimación de FD.

Bajo ciertas circunstancias, el gráfico de una función $f : [a, b] \rightarrow \mathbb{R}$

$$\text{graph}(f) = \{(t, f(t)) \mid a \leq t \leq b\}$$

considerado como un conjunto del plano (x, t) , puede ser un fractal. Primero derivamos algunas estimaciones simples pero ampliamente aplicables para la definición de caja de los gráficos. Dada una función f y un intervalo $[t_1, t_2]$, escribimos R_f para el rango máximo de f sobre un intervalo

$$R_f[t_1, t_2] = \sup_{t_1 \leq t, u \leq t_2} |f(t) - f(u)|$$

Teorema 17. Sea $f : [0, 1] \rightarrow \mathbb{R}$ continua. Sea $0 < \delta < 1$, y sea m el menor entero mayor o igual a $\frac{1}{\delta}$. Entonces, si N_δ es el número de cuadrados de δ -malla que intersecta el gráfico de f ,

$$\delta^{-1} \sum_{i=0}^{m-1} R_f[i\delta, (i+1)\delta] \leq N_\delta \leq 2m + \delta^{-1} \sum_{i=0}^{m-1} R_f[i\delta, (i+1)\delta] \quad (3.1)$$

La desigualdad de la derecha sigue siendo válida incluso si f no es continua.

Demostración. El número de cuadrados de la malla de lado δ , en la columna, por encima del intervalo $[i\delta, (i+1)\delta]$, que intercepta el gráfico de f es como mucho $2 + R_f[i\delta, (i+1)\delta]/\delta$, y si f es continua es por lo menos $R_f[i\delta, (i+\delta)\delta]/\delta$

N_δ^* : es el número de cuadrados de δ -malla que interceptan el gráfico de f en el intervalo $[i\delta, (i+1)\delta]$

de esto se tiene:

$$\sum_{i=0}^{m-1} N_\delta^* = N_\delta \quad (3.2)$$

En cada intervalo tenemos la siguiente relación:

$$\frac{R_f[i\delta, (i+1)\delta]}{\delta} \leq N_\delta^* \leq 2 + \frac{R_f[i\delta, (i+1)\delta]}{\delta} \quad (3.3)$$

Sumando a ambos lados de la desigualdad tenemos:

$$\sum_{i=0}^{m-1} \frac{R_f[i\delta, (i+1)\delta]}{\delta} \leq \sum_{i=0}^{m-1} N_\delta^* \leq \sum_{i=0}^{m-1} 2 + \frac{R_f[i\delta, (i+1)\delta]}{\delta} \quad (3.4)$$

$$\delta^{-1} \sum_{i=0}^{m-1} R_f[i\delta, (i+1)\delta] \leq N_\delta \leq 2m + \delta^{-1} \sum_{i=0}^{m-1} R_f[i\delta, (i+1)\delta] \quad (3.5)$$

que es lo que se deseaba demostrar. \square

Corolario 2. Sea $f : [0, 1] \rightarrow \mathbb{R}$ una función continua

(a) Supóngase

$$|f(t) - f(u)| \leq c|t - u|^{2-s}, \quad (0 \leq t, u \leq 1) \quad (3.6)$$

donde $c > 0$ y $1 \leq s < 2$. Entonces

$$\underline{\dim}_B \text{graph}(f) \leq \overline{\dim}_B \text{graph}(f) \leq s$$

Esto sigue siendo cierto, si para algunos $\delta > 0$, la desigualdad anterior se sostiene cuando $|t - u| < \delta$

(b) Supóngase que hay números $c > 0$, $\delta_0 > 0$ y $1 \leq s < 2$ con la siguiente propiedad: para cada $t \in [0, 1]$ y $0 < \delta \leq \delta_0$, existe u tal que $|t - u| \leq \delta$ y

$$|f(t) - f(u)| \geq c\delta^{2-s} \quad (3.7)$$

Entonces, $s \leq \underline{\dim}_B \text{graph}(f)$

Demostración. 1) Es inmediato de 3.6 que:

$$R_f[t_1, t_2] \leq c|t_1 - t_2|^{2-s} \quad 0 \leq t_1, t_2 \leq 1$$

Con la notación de 3.1, tenemos $m < 1 + \frac{1}{\delta} = 1 + \delta^{-1}$.

Igualmente de la ecuación 3.1, tenemos:

$$\begin{aligned} N_\delta &\leq 2m + \delta^{-1}c\delta^{2-s}m \\ &\leq m(2 + \delta^{-1}c\delta^{2-s}) \\ &\leq (1 + \delta^{-1})(2 + \delta^{-1}c\delta^{2-s}) \\ &\leq c_1\delta^{-s} \end{aligned}$$

Así tenemos que:

$$\frac{\log(N_\delta)}{-\log(\delta)} \leq \frac{-c_1}{\log(\delta)} + s \quad (3.8)$$

- 2) Del mismo modo de 3.7 implica que $c|t_1 - t_2|^{2-s} \leq R_f[t_1, t_2]$.
Ya que $\delta^{-1} \leq m$, obtenemos de 3.1 que:

$$N_\delta \geq \delta^{-1} \sum_{i=0}^{m-1} R_f[i\delta, (i+1)\delta] \quad (3.9)$$

$$\geq \delta^{-1} \sum_{i=1}^{m-1} c|t_1 - t_2|^{2-s} \quad (3.10)$$

$$\geq \delta^{-1}cm\delta^{2-s} \quad (3.11)$$

$$\geq \delta^{-1}\delta^{-1}c\delta^{2-s} \quad (3.12)$$

$$\geq c\delta^{-s} \quad (3.13)$$

Así que equivalente a la definición, se tiene:

$$s \leq \underline{\dim}_B \text{graph}(f)$$

□

Ejemplo 1 (La función de Weierstrass). Fijar $\lambda > 0$ y $1 < s < 2$. Definimos $f : [0, 1] \rightarrow \mathbb{R}$ por

$$f(t) = \sum_{k=1}^{\infty} \lambda^{(s-2)k} \sin(\lambda^k t) \quad (3.14)$$

Entonces, siempre que λ sea lo suficientemente grande, $\dim_B \text{graph} f = s$.

En efecto, dado $0 < h < \lambda^{-1}$, sea N el entero tal que

$$\lambda^{-(N+1)} \leq h < \lambda^{-N} \quad (3.15)$$

Entonces

$$\begin{aligned}
|f(t+h) - f(t)| &\leq \sum_{k=1}^N \lambda^{(s-2)k} |\sin(\lambda^k(t+h)) - \sin(\lambda^k t)| \\
&\quad + \sum_{k=N+1}^{\infty} \lambda^{(s-2)k} |\sin(\lambda^k(t+h)) - \sin(\lambda^k t)| \\
&\leq \sum_{k=1}^N \lambda^{(s-2)k} \lambda^k h + \sum_{k=N+1}^{\infty} 2\lambda^{(s-2)k}
\end{aligned}$$

Usamos el hecho $|\sin(u) - \sin(v)| \leq |u - v|$ (esto debido al teorema del valor medio), y que $|\sin(u)| \leq 1$ en el resto, y sumamos las series geométricas.

$$\begin{aligned}
h \sum_{k=1}^N (\lambda^{s-1})^k &= h \left(\frac{\lambda^{s-1} - (\lambda^{s-1})^{N+1}}{1 - \lambda^{s-1}} \right) \\
&= h \frac{\lambda^{s-1}(1 - (\lambda^{s-1})^N)}{\lambda^{s-1}(\lambda^{1-s} - 1)} \\
&= h \frac{((\lambda^{s-1})^N - 1)}{(1 - \lambda^{1-s})} \\
&\leq \frac{h(\lambda^{s-1})^N}{1 - \lambda^{1-s}}
\end{aligned}$$

$$\begin{aligned}
2 \sum_{k=N+1}^b \lambda^{(s-2)k} &= 2 \left(\frac{(\lambda^{s-2})^{N+1} - (\lambda^{s-2})^{b+1}}{1 - \lambda^{s-2}} \right) \\
&\leq 2 \frac{\lambda^{(s-2)N+1}}{1 - \lambda^{s-2}}
\end{aligned}$$

Aquí tomamos $b \rightarrow \infty$.

Así tenemos que

$$\begin{aligned}
|f(t+h) - f(t)| &\leq \frac{h\lambda^{(s-1)N}}{1 - \lambda^{s-1}} + \frac{2\lambda^{(s-2)(N+1)}}{1 - \lambda^{s-2}} \\
&\leq \frac{h\left(\frac{1}{h}\right)^{s-1}}{1 - \lambda^{1-s}} + \frac{2\left(\frac{1}{h^2}\right)^{s-2}}{1 - \lambda^{s-2}} \\
&= \frac{h^{2-s}}{1 - \lambda^{1-s}} + \frac{2\left(\frac{\lambda}{h}\right)^{s-2}}{1 - \lambda^{s-2}} \\
&= h^{2-s} \left(\frac{1}{1 - \lambda^{1-s}} + \frac{2\lambda^{s-2}}{1 - \lambda^{s-2}} \right) \\
&\leq ch^{2-s}
\end{aligned}$$

donde c es independiente de h

Por corolario 2.(a), tenemos que $\overline{\dim}_B \text{graph} f \leq s$.

De la misma manera, dividiendo la suma en tres partes: los primeros $N-1$ términos, término N y el resto tenemos que:

$$\begin{aligned}
& |f(t+h) - f(t) - \lambda^{(s-2)N}(\sin(\lambda^N(t+h)) - \sin(\lambda^N t))| \\
\leq & |f(t+h) - f(t)| + \lambda^{(s-2)N} |\lambda^N(t+h) - \lambda^N t| \\
\leq & \frac{h\lambda^{(s-1)N}}{1 - \lambda^{1-s}} + \frac{2\lambda^{(s-2)(N+1)}}{1 - \lambda^{s-2}} + h\lambda^{(s-2)N+N} \\
\leq & \frac{h\lambda^{(s-1)N}}{1 - \lambda^{1-s}} + \frac{2\lambda^{(s-2)(N+1)}}{1 - \lambda^{s-2}} + \lambda^{(s-2)N} \\
\leq & \frac{(\lambda^{N(s-2)} + \lambda^{N(s-2)}) - \lambda^{(s-2)N-s+1}}{1 - \lambda^{1-s}} + \frac{2\lambda^{(s-2)(N+1)}}{1 - \lambda^{s-2}} \\
\leq & \frac{\lambda^{N(s-2)} - \lambda^{(s-2)N-s+1}}{1 - \lambda^{1-s}} + \frac{2\lambda^{(s-2)(N+1)}}{1 - \lambda^{s-2}}
\end{aligned}$$

Veamos lo siguiente:

$$\begin{aligned}
2\lambda^{N(s-2)} - \lambda^{(s-2)N-s+1} &= \lambda^{N(s-2)}(2 - \lambda^{1-s}) \\
&\leq \lambda^{N(s-2)}\lambda^{1-s}
\end{aligned}$$

Así tenemos que:

$$|f(t+h) - f(t) - \lambda^{(s-2)N}(\sin(\lambda^N(t+h)) - \sin(\lambda^N t))| \quad (3.16)$$

$$\leq \frac{\lambda^{(s-2)N-s+1}}{1 - \lambda^{1-s}} + \frac{2\lambda^{(s-2)(N+1)}}{1 - \lambda^{s-2}} \quad (3.17)$$

Suponga que $\lambda > 2$ es lo suficientemente grande para que el lado derecho de (3.17) sea menor que $\frac{1}{20}\lambda^{(s-2)N}$ para todo N .

Para $\delta < \lambda^{-1}$ tomar N tal que

$$\lambda^{-1} \leq \delta < \lambda^{-(N-1)}$$

Para cada t , podemos elegir h con $\frac{1}{\lambda^{N+1}} \leq h < \frac{1}{\lambda^N} < \delta$ tal que

$$|\sin(\lambda^N(t+h)) - \sin(\lambda^N t)| > \frac{1}{10}$$

Recordemos lo siguiente.

$$|a| - |b| < |a - b| < |a| + |b|$$

y además $|a - b| = |b - a|$.

y así tomando

$$a = f(t+h) - f(t)$$

$$b = \lambda^{(s-2)N} (\sin(\lambda^N(t+h)) - \sin(\lambda^N(t)))$$

tenemos, usando (3.17):

$$\begin{aligned} |f(t+h) - f(t)| &\geq \frac{1}{10} \lambda^{(s-2)N} - \frac{1}{20} \lambda^{(s-2)N} \\ &= \frac{1}{20} \lambda^{(s-2)N} \\ &\geq \frac{1}{20} \lambda^{s-2} \delta^{2-s} \end{aligned}$$

Esta última desigualdad es debido a que:

$$\frac{1}{\lambda^N} \leq \delta < \frac{1}{\lambda^{N-1}}$$

Así $\delta < \lambda^{N-1}$.

o lo que es lo mismo $\delta \leq \frac{\lambda}{\lambda^N}$
de donde tenemos

$$\begin{aligned} \lambda^N &\leq \frac{\lambda}{\delta} \\ (\lambda^N)^{2-s} &\leq \left(\frac{\lambda}{\delta}\right)^{s-2} \\ \lambda^{s-2} \delta^{2-s} &\leq \lambda^{N(s-2)} \end{aligned}$$

Veamos cuatro gráficas de la función Weierstrass $f(t) = \sum_{k=1}^{\infty} \lambda^{(s-2)k} \sin(\lambda^k t)$ con $\lambda = 1.5$ y en: (a) $s = 1.1$, (b) $s = 1.3$, (c) $s = 1.5$, (d) $s = 1.7$

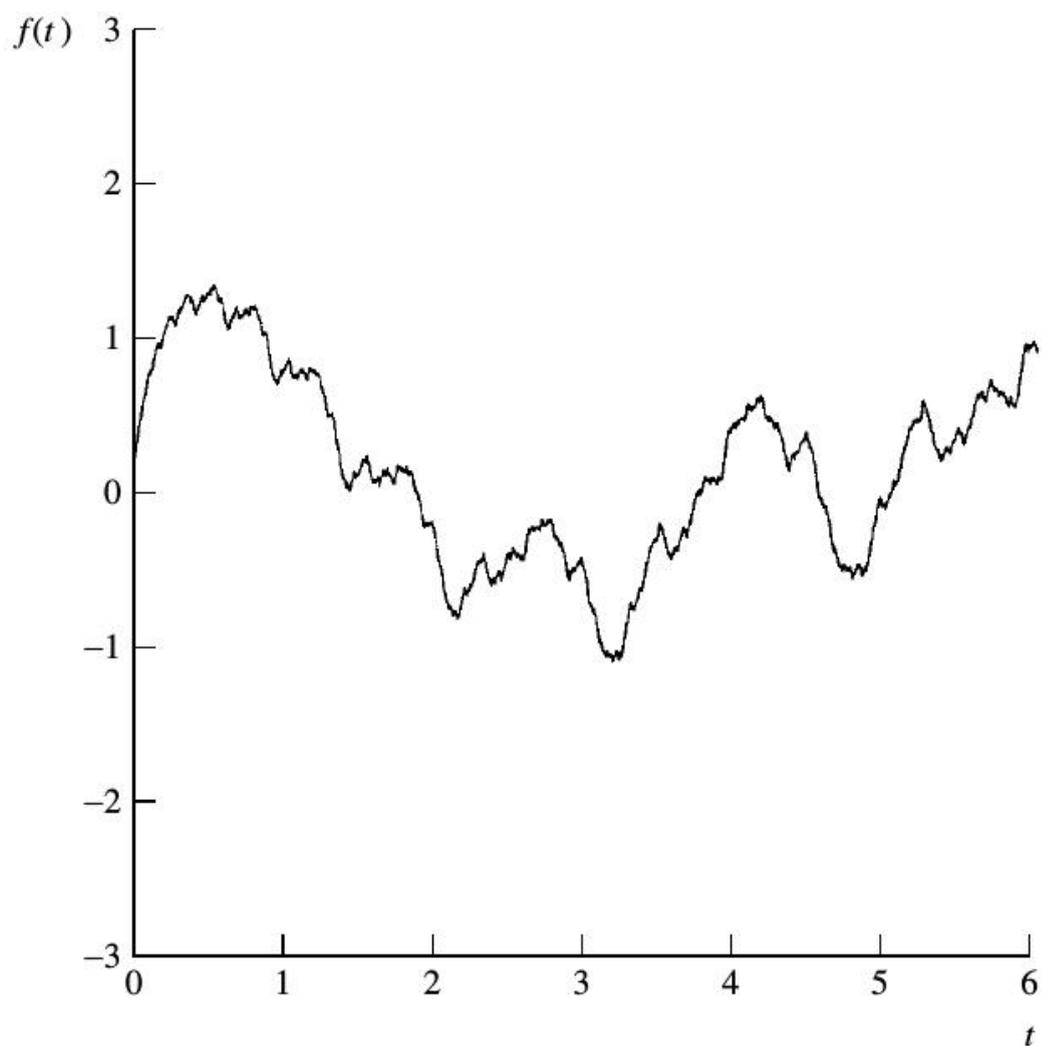
Función Weierstrass con $s = 1.1$ 

Figura 3.3: Fuente:[11]

Función Weierstrass con $s = 1.3$

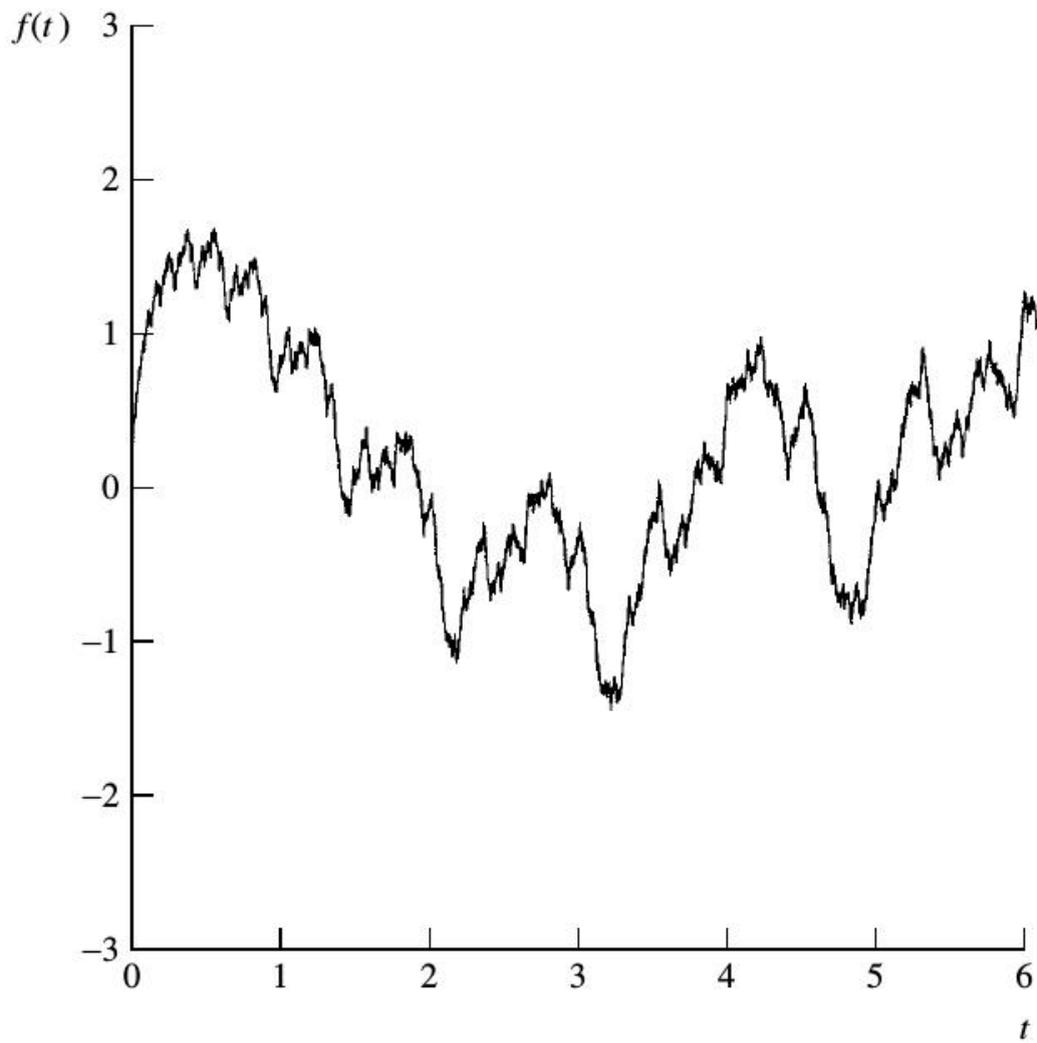


Figura 3.4: Fuente:[11]

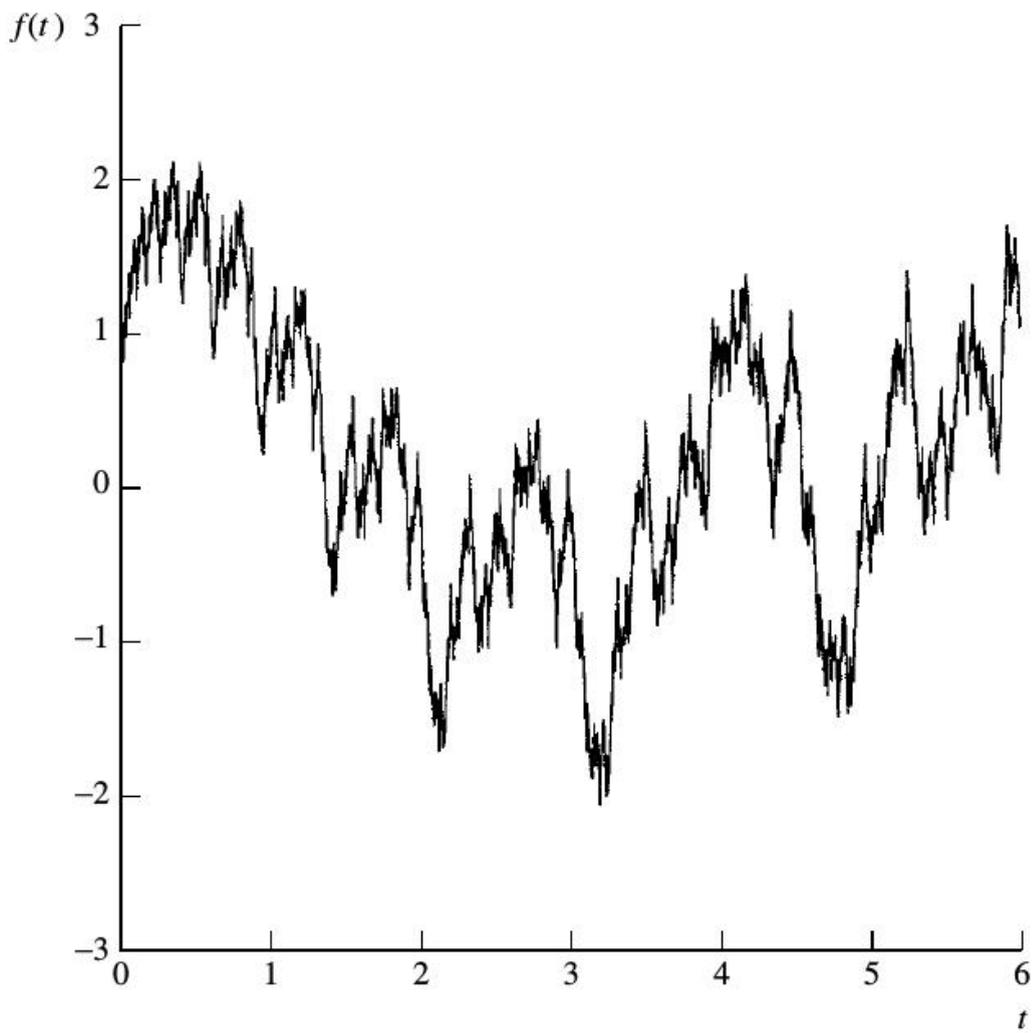
Función Weierstrass con $s = 1.5$ 

Figura 3.5: Fuente:[11]

Función Weierstrass con $s = 1.7$

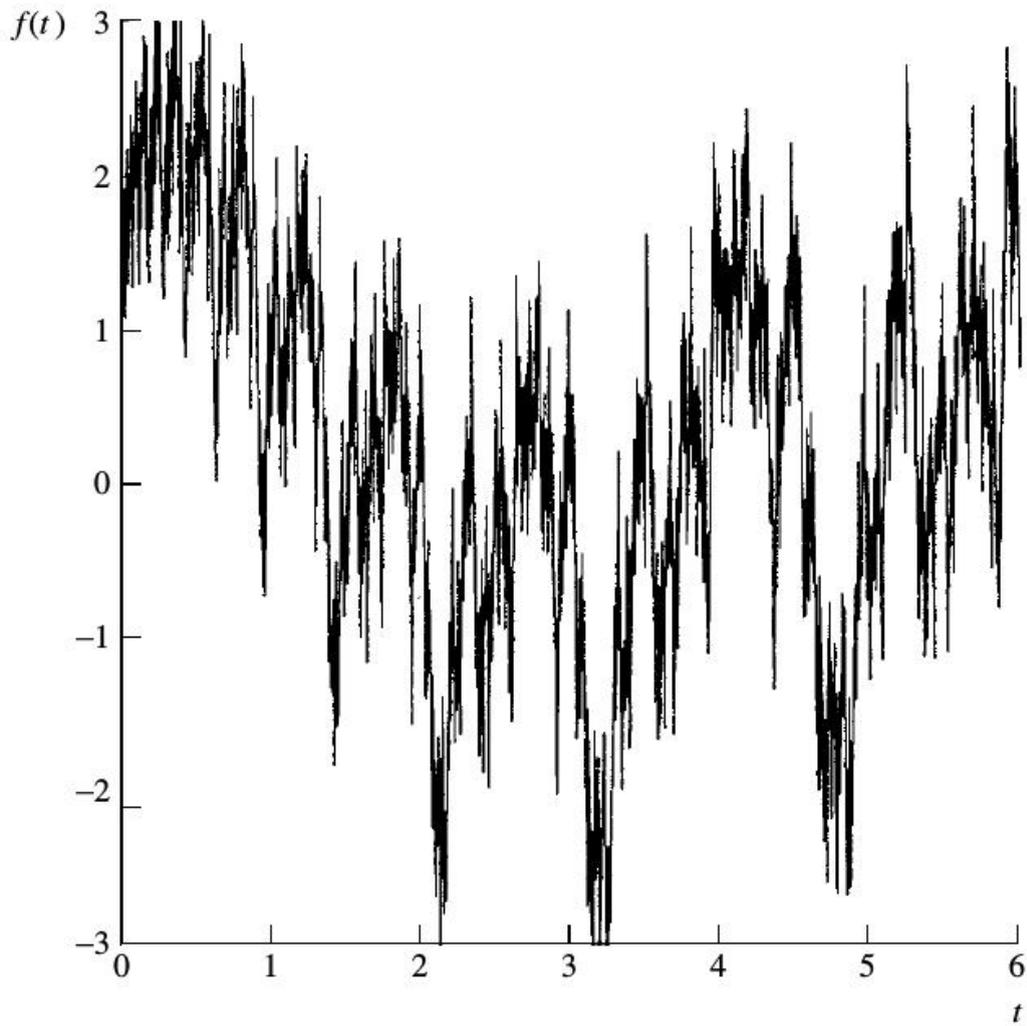


Figura 3.6: Fuente:[11]

3.2. ENFOQUE DE LA TRANSFORMADA DE FOURIER.

En esta sección expondremos una manera de encontrar la dimensión fractal utilizando la transformada de Fourier y la correlación de la señal en estudio.

Definición 18. Sea $f, g \in L^1(\mathbb{R}^n)$. Definimos la convolución $f * g$ por

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x - y)dy$$

Definición 19. La n -dimensional Transformada de Fourier de una función Lebesgue integrable f se define por:

$$\hat{f}(u) = \int_{\mathbb{R}^n} f(x)e^{-ixu} dx, u \in \mathbb{R}^n$$

Definición 20. Sea f una función Lebesgue integrable y \hat{f} su transformada de Fourier, entonces se tiene la siguiente relación:

$$f(y) = \frac{1}{2\pi} \int_{\mathbb{R}^n} \hat{f}(y)e^{iyt} dy$$

Teorema 18. La transformada de Fourier de la convolución de dos funciones es el producto de las transformadas, es decir:

$$(\widehat{f_1 * f_2})(x) = \hat{f}_1(x)\hat{f}_2(x)$$

Autocorrelacion de funciones fractales

Las series temporales evolucionan normalmente con cierta inercia, que se manifiesta en dependencia entre sus valores presentes y sus valores pasados. Para medir la dependencia lineal entre el presente y el pasado se utilizan las funciones de autocovarianzas y autocorrelaciones y que generalizan la idea de covarianza y correlación entre dos variables.

Las cantidades que varían con el tiempo a menudo resultan tener gráficos fractales. Una forma en que su naturaleza fractal se manifiesta a menudo es por un comportamiento de ley de potencia de la correlación entre las mediciones separadas por el tiempo h .

Para la conveniencia del análisis asumimos que $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función continua y acotada, y consideremos el comportamiento promedio de f en el período $[-T, T]$.

Escribimos por \bar{f} por el valor medio de f , esto es:

$$\bar{f} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t)dt$$

Definición 21. La medida de correlación entre f separada en un tiempo h la proporciona la función de auto-correlación, dada por:

$$C(h) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t+h) - \bar{f})(f(t) - \bar{f})dt \quad (3.18)$$

De donde, tenemos la siguiente relación:

$$C(h) = C(0) - \frac{1}{2} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t+h) - f(t))^2 dt \quad (3.19)$$

En efecto:

$$\begin{aligned} C(h) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t+h) - \bar{f})(f(t) - \bar{f}) dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t+h)f(t) - \bar{f}f(t+h) - \bar{f}f(t) + (\bar{f})^2) dt \\ &= \lim_{T \rightarrow \infty} \int_{-T}^T f(t+h)f(t) dt - \bar{f} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t+h) dt - \bar{f} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t) dt + \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (\bar{f})^2 dt \\ &= \lim_{T \rightarrow \infty} \int_{-T}^T f(t+h)f(t) dt - (\bar{f})^2 - (\bar{f})^2 + (\bar{f})^2 \\ &= \lim_{T \rightarrow \infty} \int_{-T}^T f(t+h)f(t) dt - (\bar{f})^2 \end{aligned}$$

De 3.18 vemos que $c(h)$ es positiva si $f(t+h) - \bar{f}$ y $f(t) - \bar{f}$ tiene el mismo signo, y es negativa si tienen signo opuesto. Si no hay correlación $c(h) = 0$.

Sabemos que:

$$\int (f(t+h) - f(t))^2 dt = \int f(t+h)^2 dt + \int f(t)^2 dt - 2 \int f(t+h)f(t) dt$$

Lo que significa que:

$$\int f(t+h)f(t) dt = \frac{1}{2} \left(\int f(t+h)^2 dt + \int f(t)^2 dt - \int (f(t+h) - f(t))^2 dt \right)$$

Así:

$$\begin{aligned} C(h) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t+h)f(t) dt - (\bar{f})^2 \\ &= \frac{1}{2} \left(\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t+h)^2 dt + \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t)^2 dt - \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t+h) - f(t))^2 dt \right) - (\bar{f})^2 \\ &= \bar{f}^2 - (\bar{f})^2 - \frac{1}{2} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t+h) - f(t))^2 dt \end{aligned}$$

Observemos que $C(0) = \bar{f}^2 - (\bar{f})^2$.

Por lo tanto tenemos:

$$C(h) = C(0) - \frac{1}{2} \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (f(t+h) - f(t))^2 dt \quad (3.20)$$

Donde

$$\bar{f}^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t)^2 dt$$

es la media cuadrática de f , asumiendo que es positivo y finito.

Con $C(h)$ en la forma 3.20, podemos inferir la relación entre la función de autocorrelación y el dimensión del gráfico de f . La pista es el corolario 2. Supóngase que f es una función satisfaciendo 2 y también 3.20. Entonces existen constantes c_1 y c_2 tal que:

$$c_1 h^{2-s} \leq f(t+h) - f(t) \leq c_2 h^{2-s}$$

De donde tenemos:

$$c_1 h^{4-2s} \leq \frac{1}{2T} \int_{-T}^T (f(t+h) - f(t))^2 dt \leq c_2 h^{4-2s}$$

para h pequeño.

Así, la función de autocorrelación f satisface:

$$C(0) - C(h) \simeq ch^{4-2s} \quad (3.21)$$

Para h pequeño, es razonable esperar que la dimensión de caja del gráfico de f sea s .

Un enfoque estándar para el análisis de series de tiempo es llevar a cabo una transformada de Fourier de la serie, lo utilizaremos y llegaremos a la misma ecuación 3.21, bajo la hipótesis de que f tiene media cero.

Una serie de tiempo puede prescribirse en el dominio físico como $f(t)$ o en el dominio de la frecuencia en términos del espectro $y(\omega, T)$, donde ω es la frecuencia.

$$y(\omega, T) = \int_{-T}^T f(t) e^{i\omega t} dt$$

La densidad espectral de potencia (espectro de potencia) de $f(t)$ es definido por:

$$s(\omega) = \frac{1}{T} |y(\omega, T)|^2$$

en el límite $T \rightarrow \infty$.

La función de auto-correlación está estrechamente relacionada con el espectro de potencia de f .

Se demuestra que el espectro de potencia es la Transformada de Fourier de la función de auto correlación.

$$\hat{C}(\omega) = S(\omega)$$

En efecto, definimos:

$$f_T(t) = \begin{cases} f(t), & |t| \leq T \\ 0, & \text{otro caso} \end{cases} \quad (3.22)$$

y

$$\begin{aligned} C_T(h) &= \frac{1}{2T} \int_{-\infty}^{\infty} f_T(t+h) f_T(t) dt \\ &= \frac{1}{2T} f_T^- * f_T(-h) \end{aligned}$$

donde $f_T^-(t) = f_T(-t)$ y $*$ denota la convolución.

Por teorema 18 tenemos:

$$\begin{aligned} \hat{C}_T(w) &= \frac{1}{2T} \hat{f}_T^-(w) \hat{f}_T(w) \\ &= \frac{1}{2T} \hat{f}_T(-w) \hat{f}_T(w) \\ &= \frac{1}{2T} \left(\int_{-\infty}^{\infty} f(u) e^{i w u} du \int_{-\infty}^{\infty} f(u) e^{-i w u} du \right) \\ &= \frac{1}{2T} |\hat{f}_T(w)|^2 \end{aligned}$$

Notar que:

$$\lim_{T \rightarrow \infty} C_T(h) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-\infty}^{\infty} f_T(t+h) f_T(t) dt = C(h)$$

Así, es inmediato que:

$$\hat{C}_T(w) = \hat{C}(w)$$

Lo que significa que:

$$\hat{C}(w) = S(w)$$

Como las funciones C, S son reales, la transformación es transformación coseno. Así:

$$S(\omega) = \int_{-\infty}^{\infty} C(t) e^{i t \omega} dt = \int_{-\infty}^{\infty} C(t) \cos(\omega t) dt \quad (3.23)$$

Y, mediante, la fórmula para inversión de transformadas de Fourier

$$C(h) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{-i h \omega} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) \cos(\omega h) d\omega \quad (3.24)$$

Las correlaciones nos proporcionan métodos para para estimar la dimensión del gráfico de f . Podemos calcular la función de correlación $C(h)$ o equivalentemente el cambio de la media cuadrática en la señal (f) en el tiempo h durante un periodo largo, de modo que

$$2[C(0) - C(h)] \simeq \frac{1}{2T} \int_{-T}^T (f(t+h) - f(t))^2 dt$$

Si el comportamiento de ley de potencia

$$C(0) - C(h) \simeq ch^{4-2s} \quad (3.25)$$

se observa para h pequeño, se esperaría que la dimensión de caja del gráfico sea s , es decir:

$$\dim_{Bgraph}(f) = 2 - \lim_{h \rightarrow 0} \frac{\log C(0) - C(h)}{2 \log h}$$

Alternativamente podemos trabajar desde el espectro de potencia $S(\omega)$ para encontrar la función de auto-correlación.

La situación de mayor interés es cuando el espectro de potencia obedece una ley de potencia. Para series temporales que son fractales auto-afines, la densidad espectral de potencia tiene una dependencia de ley de potencia de la frecuencia:

$$s(\omega) \sim \omega^{-\alpha}$$

para w grande, en ese caso, se tiene

$$C(0) - C(h) \sim bh^{\alpha-1} \quad (3.26)$$

para h pequeño y alguna constante b .

Según Mandelbrot, se cumple la siguiente relación:

$$\alpha = 5 - 2s$$

En efecto, de 3.24, se tiene:

$$\begin{aligned} 2\pi(C(0) - C(h)) &= \int_{-\infty}^{\infty} S(\omega) \cos(0) d\omega - \int_{-\infty}^{\infty} S(\omega) \cos(\omega h) d\omega \\ &= 2 \int_0^{\infty} S(\omega) (1 - \cos(\omega h)) d\omega \end{aligned}$$

Utilizando la identidad

$$\text{sen}^2(t) = \frac{1 - \cos(2t)}{2}$$

y

$$t = \omega h$$

tenemos

$$\pi(C(0) - C(h)) = \int_0^{\infty} S(\omega) 2 \text{sen}^2\left(\frac{\omega h}{2}\right) d\omega \quad (3.27)$$

$$= 2 \int_0^{\infty} S(\omega) \text{sen}^2\left(\frac{\omega h}{2}\right) d\omega \quad (3.28)$$

$$= 2 \int_0^{\infty} \omega^{-\alpha} \text{sen}^2\left(\frac{\omega h}{2}\right) d\omega \quad (3.29)$$

donde

$$S(\omega) = \omega^{-\alpha}$$

Hacemos $u = h\omega$. Así, tenemos que:

$$\omega^{-\alpha} = \left(\frac{u}{h}\right)^{-\alpha}, \quad d\omega = \frac{du}{h}$$

Así tenemos

$$\begin{aligned}\pi(C(0) - C(h)) &= 2 \int_{-\infty}^{\infty} \left(\frac{u}{h}\right)^{-\alpha} \operatorname{sen}^2\left(\frac{1}{2}u\right) \frac{du}{u} \\ &= 2h^{\alpha-1} \int_{-\infty}^{\infty} u^{-\alpha} \operatorname{sen}^2\left(\frac{1}{2}u\right) du \\ \frac{1}{2}\pi(C(0) - C(h)) &= h^{\alpha-1} \int_{-\infty}^{\infty} u^{-\alpha} \operatorname{sen}^2\left(\frac{1}{2}u\right) du\end{aligned}$$

Comparando 3.25 y 3.26, tenemos

$$4 - 2s = \alpha - 1$$

de donde

$$s = \frac{5 - \alpha}{2} \tag{3.30}$$

Esto nos permite encontrar la dimensión fractal s en términos de α , para una función f que cumple las hipótesis de los teoremas.

3.3. APLICACIÓN A VULCANOLOGÍA

Se cuenta con una base de datos, de las mediciones de diferentes factores físicos, como el CO_2 , temperatura del aire, la presión atmosférica, entre otros factores; tomados en el volcán de San Salvador en el período comprendido desde el primero de enero al 31 de diciembre del año 2005. Las mediciones se hicieron en intervalos de una hora de separación. Dentro de la variedad de parámetros de interés para los especialistas, en este trabajo estudiamos el factor Dióxido de carbono (CO_2).

Se hará un estudio de la estimación de la dimensión fractal s . Esto nos permitirá determinar el grado de complejidad de la serie temporal. La dimensión fractal es una herramienta originada para analizar fractales definidos matemáticamente como por ejemplo el conjunto de cantor entre otros, en este caso, la estamos utilizando para estudiar la serie temporal de emisiones de CO_2 en el año 2005 en el volcán de San Salvador.

Debe observarse que no exigimos de antemano que la serie tenga un comportamiento fractal, como autosimilitud, pues sabemos que los fenómenos de la naturaleza no se comportan de manera ideal, sin embargo si podemos calcular su dimensión fractal para analizar su grado de llenado en el espacio. En este proceso no se entra en conflicto con la teoría, sabemos que el algoritmo para calcular la dimensión de caja siempre nos devolverá un número sin importar el objeto de interés.

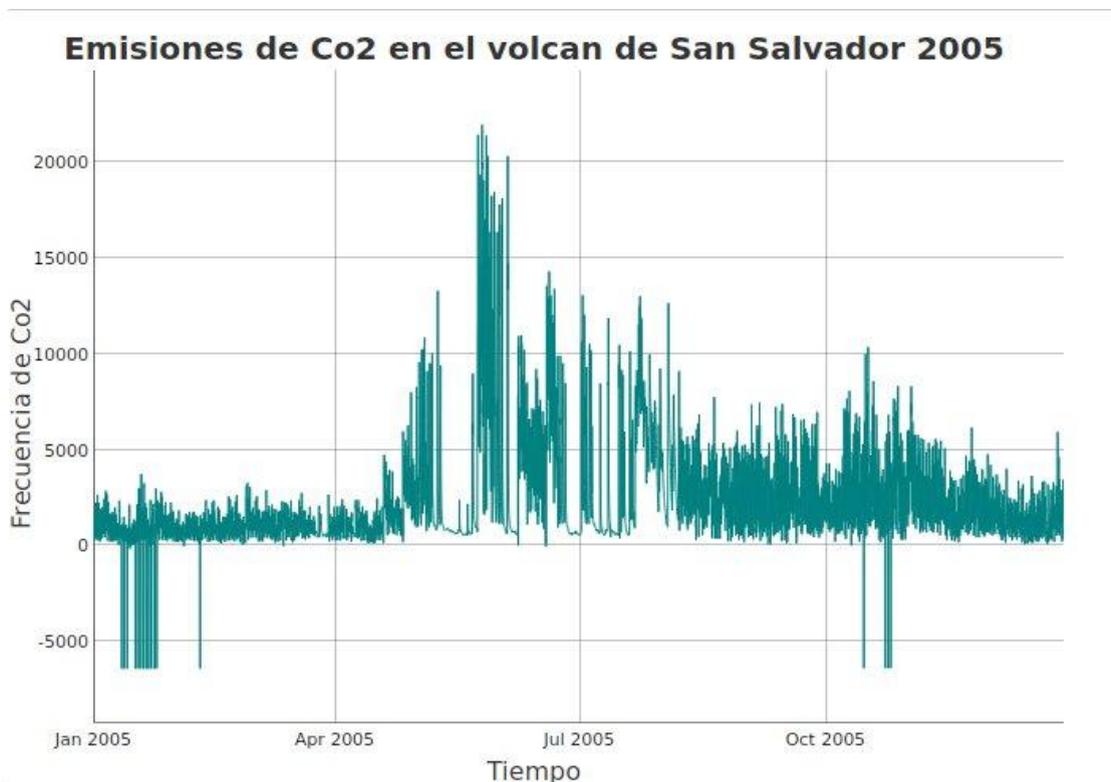


Figura 3.7: Fuente: Elaboración propia

Los datos originales de emisiones de Co2 presentan valores negativos, esto es imposible, pues las emisiones son valores mayores o iguales a cero. Esta anomalía se debe a fallos en el sistema que estaba tomando los datos.

Por esta razón, los valores negativos se aproximan por la media aritmética de los 5 datos anteriores. Para ello se crea una función llamada "corrige" en el Software R que corrige los datos.

Veamos la función:

```
#Creamos una funcion que corrige los datos: un dato negativo lo sustituye por
# el promedio de los cinco valores anteriores
corrige<-function(v){
  w<-rep(0,length(v))
  for(i in 1:5){
    w[i]<-v[i]
  }
  for(i in 6:length(v)){
    if(v[i] > 0){
      w[i]<-v[i]
    }
    else {
      w[i]<-sum(w[i-1]+w[i-2]+w[i-3]+w[i-4]+w[i-5])/5
    }
  }
  return(w)
}
```

Ahora si se hacen pasar los datos originales a través de esta función y se grafican tenemos lo siguiente:

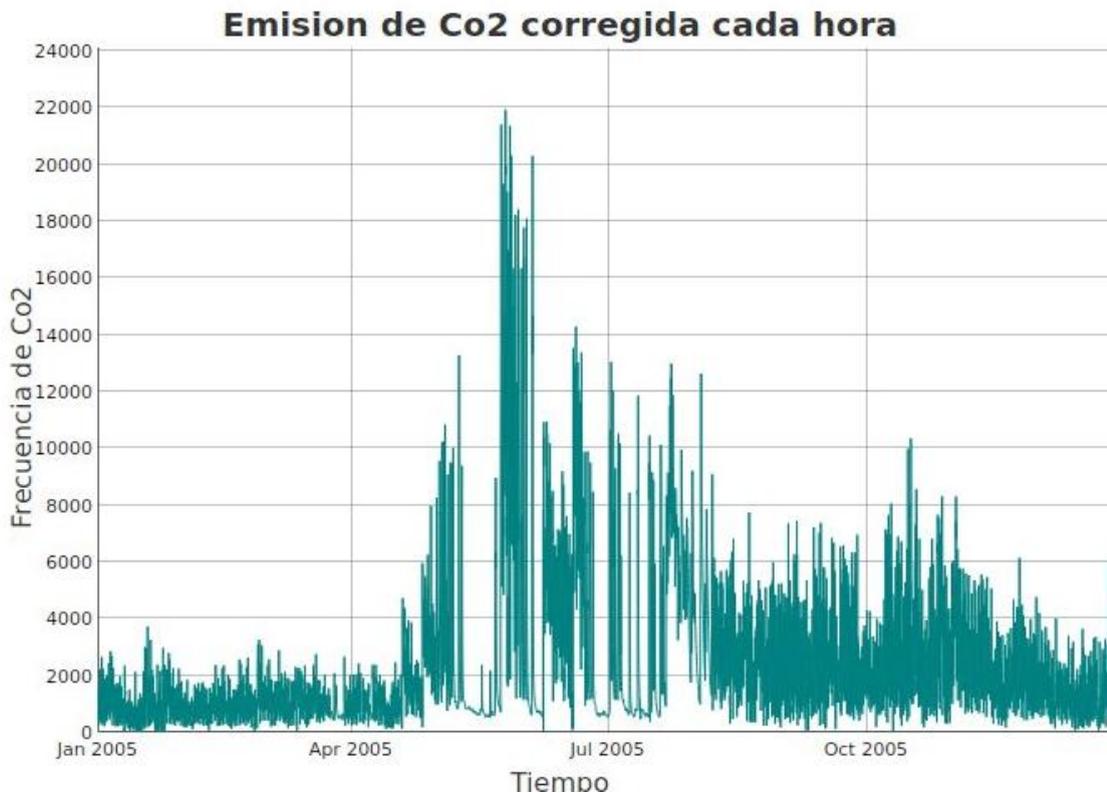


Figura 3.8: Fuente:Elaboración propia

Como podemos observar, no podemos aplicar los métodos y técnicas descritas en el capítulo anterior para determinar la dimensión fractal, esto debido a que no se cumplen las hipótesis de los teoremas a aplicar así y el hecho de desconocer la función (señal) F , es por ello que desarrollamos un algoritmo computacional que nos da una aproximación de la dimensión fractal.

El algoritmo para el cálculo de la dimensión fractal por medio de el conteo de cajas propone obtener la dimensión fractal para varios tamaños de caja y hacer un ajuste lineal a una gráfica $\log - \log$ de $N(r)$ sobre (r) . La pendiente de la recta de mínimos cuadrados se toma como una estimación de la dimensión fractal de la curva (serie).

Iniciamos fijando un radio dt , que será la distancia entre dos puntos de la serie, es decir, si $y(i)$ pertenece a la serie también lo hace $y(i + dt)$. Luego encontramos la altura h de la serie en el bloque $y(i), y(i + dt)$, es decir; $h = \max y(j) - \min y(j), j \in [i, i + dt]$.

Así en cada bloque de la serie son necesarias $b = \lceil \frac{h}{dt} \rceil$ cajas de lado dt para cubrir el bloque de la serie en análisis, donde $\lceil a \rceil$ denota el menor entero mayor que a .

El total de cajas de resolución dt necesarias para cubrir la serie es $N(r) = \sum_{k=1}^s b_s$, donde s es el número de bloques en que dt dividió la serie. El procedimiento se repite haciendo variar

el radio dt .

Creamos una función llamada "bloques" que calcula el número de cajas cuadradas, de lado dado, necesarias para cubrir toda la serie.

```

bloques<-function(u=datos , dt){
  xx<-rep(0 , length(u))
  pos<-rep(0 , length(u))
  xx[1]<-u[1]
  pos[1]<-1
  for(i in 2:length(u)){
    if ((i*dt)<length(u)){
      xx[i]<-u[((i-1)*dt)]
      pos[i]<-(i-1)*dt
    }
    x<-xx[xx>0]
    f<-pos[pos>0]
  }
  n<-0 # numero de cajas
  for(i in 1:(length(f)-1)){
    w<-u[c(f[i]:f[i+1])]# genera el bloque de elementos de u de longitud dt, f
      tiene indices
    ha<-(max(w)-min(w))
    n<-ceiling(ha/dt)+n
    if (i==(length(f)-1)){ # analizando la cola de u de longitud menor que dt
      cola<-u[c(length(f):length(u))] # contiene los ultimos elementos de u
      ha2<-(max(cola)-min(cola))
      n<-ceiling(ha2/dt)+n
    }
  }

  return(list(x,f,n))
# x: vector con frecuencia dt
#f: vector con las posiciones(indices) en u de los valores de x
#n: numero de cajas de lado dt para cubrir u
}

```

Veamos el código para la función "creadatos", que lo que hace es crear un vector con el número de cajas, dado un radio con el que se crea una lista de radios:

```
#recibe la serie y el radio l(limite superior)
#crea un vector de radios desde 1 hasta l
#regresa el vector de cajas
creadatos<-function(u=datos,l){# l: limite superior de las frecuencias
  multiples<-1:l
  cajas<-rep(0,l) # contiene numero de cajas de ancho l
  for(i in 1:l){
    cajas[i]<-bloques(u=datos,multiples[i])[3]
  }
  boxx<-as.numeric(cajas)
  return(boxx)
}
```

De igual manera se crean las funciones REGRESION1,REGRESION2 y REGRESION3, que regresan el summary de la función lm, el logaritmo de un vector de radios y el logaritmo de un vector del número de cajas respectivamente. Estos son sus códigos:

```
REGRESION1<-function(u=datos ,j){ # regresa el summary
  q<-1:j
  box<-as.numeric(creadatos(u=datos ,j)) # boxx
  rad<-(q) # multiplos
  radi<-1/rad
  box.log<-log10(box)
  radi.log<-log10(radi)
  f<-lm(box.log~radi.log)
  A<- list(f , radi.log , box.log)
  return( summary(f) )
}
```

```
REGRESION2<-function(u=datos ,j){ # regresa el log del vector radio
  q<-1:j
  box<-as.numeric(creadatos(u=datos ,j)) # boxx
  rad<-(q) # multiplos
  radi<-1/rad
  box.log<-log10(box)
  radi.log<-log10(radi)
  f<-lm(box.log~radi.log)
  A<- list(f , radi.log , box.log)
  return(radi.log)
}
```

```
REGRESION3<-function(u=datos ,j){ # regresa el log del vector numero de caja
  q<-1:j
  box<-as.numeric(creadatos(u=datos ,j)) # boxx
  rad<-(q) # multiplos
  radi<-1/rad
  box.log<-log10(box)
  radi.log<-log10(radi)
  f<-lm(box.log~radi.log)
  A<- list(f , radi.log , box.log)
  return(box.log)
}
```

Ahora, apoyandonos de las funciones antes descritas, estimaremos la dimensión de caja para la serie de tiempo de las emisiones de Co2 en el volcán de San Salvador.

Recordemos que debemos fijar un radio para hacer la estimación de la dimensión, ya que las funciones así están diseñadas, aquí fijamos el radio en 50, es decir, para los valores desde uno hasta cincuenta, se calcula el número de cajas necesarias para cubrir la serie para cada radio.

Con estos valores, tanto de los radios como de el número de cajas, calculamos la recta de regresión lineal entre ambas variables y la pendiente de la recta de regresión es la dimensión fractal de la serie.

Veamos como se implementa en R:

```
p<-lm(REGRESION3(u=datos,50)~REGRESION2(u=datos,50))
plot(REGRESION2(u=datos,50),REGRESION3(u=datos,50), main=" log(radio) vs log(N
  cajas)", xlab = "log(radio)", ylab = "log(N cajas)")
abline(p)
```

Esto nos genera lo siguiente:

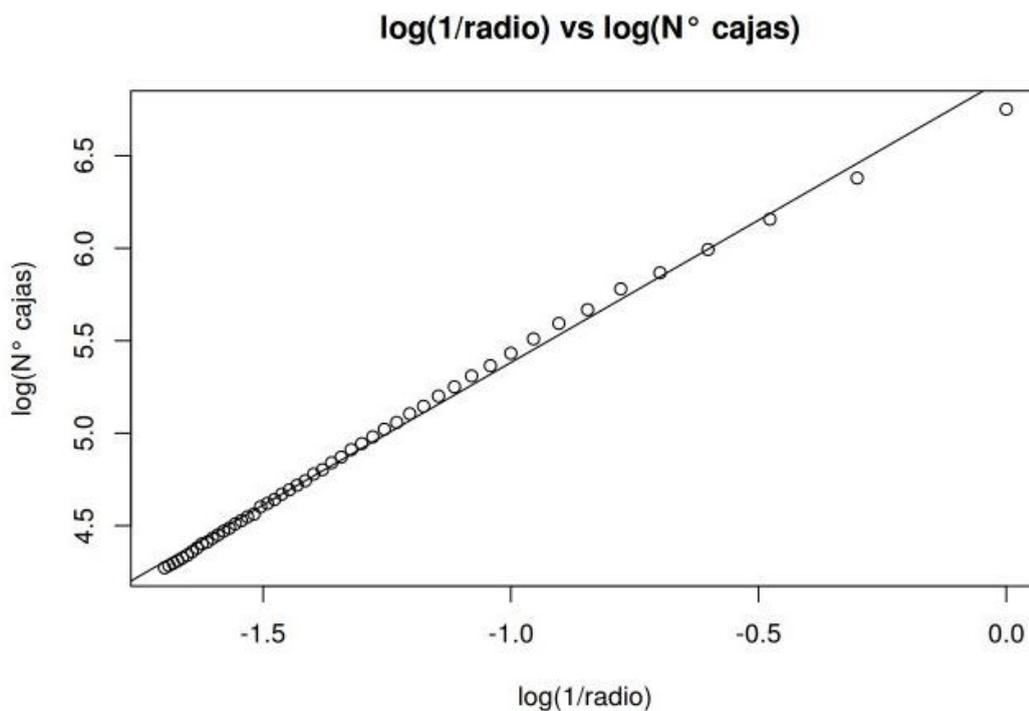


Figura 3.9: Fuente:Elaboración propia

Cuando ejecutamos la función REGRESION1, tenemos lo siguiente:

Resumen de la regresión

```
##
## Call:
## lm(formula = box.log ~ radi.log)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.16998 -0.02193 -0.00060  0.03057  0.06291
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.92195    0.02045   338.4  <2e-16 ***
## radi.log     1.54030    0.01521   101.3  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0411 on 48 degrees of freedom
## Multiple R-squared:  0.9953, Adjusted R-squared:  0.9952
## F-statistic: 1.026e+04 on 1 and 48 DF,  p-value: < 2.2e-16
```

Figura 3.10: Fuente:Elaboración propia

De la tabla anterior deducimos que la recta de regresión que modela los datos es: $y = 1.54030x + 6.92195$.

De esto concluimos que la dimensión de caja de la series es: 1.54030. Este número seguirá aumentado a medida que ejecutemos el código con un número más grande de radios. Sin embargo es una muy buena aproximación de la dimensión fractal de la serie temporal de las emisiones de CO_2 .

Los algoritmos de estimación de la dimensión fractal dan un número independientemente de si el objeto es o no un fractal. También es posible tener dos conjuntos fractales diferentes que tengan la misma dimensión fractal.

Observemos también que a medida incrementamos el número de radios la aproximación de la dimensión fractal será mejor.

Este número nos ayuda, como ya expusimos, a determinar el grado de llenado que la serie tiene en el espacio.

Sin embargo ningún objeto físico es un fractal, porque no tiene propiedades similares a todas las escalas. Esto lleva al hecho de que el análisis de la dimensión fractal no distingue entre objetos fractales y no fractales, sino que de alguna manera da la medida de la conveniencia para describir el objeto utilizando modelos fractales.

Recomendamos trabajar con las series temporales de las emisiones de CO_2 en un período más grande de tiempo, por ejemplo 10 años, y conocer la dimensión fractal de cada año; lo que nos permitirá conocer la forma en que las series temporales han ido llenado el espacio en dicho periodo.

También sugerimos estudiar series temporales que tiene un comportamiento que cumple las condiciones impuestas por la teoría y así poder comparar su dimensión fractal obtenida por el algoritmo y la que dicta la teoría.

Apéndice A

BASE DE DATOS PARA FINANZAS

Pérdidas totales mensuales 2000-2008

##	Jan	Feb	Mar	Apr	May
## 2000	4720566.9	9389342.8	1488557.3	3163696.2	240462.2
## 2001	7528091.7	4316079.3	1914818.5	4486207.3	2609332.4
## 2002	1634311.3	2688666.1	1779702.4	9041768.2	2998347.9
## 2003	2690431.0	1486728.1	1159835.4	1905096.8	1451992.3
## 2004	3088024.5	3148085.0	6353706.5	1472401.5	2428068.6
## 2005	636173.8	706059.6	3493612.1	462902.1	1758834.9
## 2006	2387470.6	1509856.2	1386588.6	1100408.5	20537989.8
## 2007	959501.1	653178.8	523717.6	1248877.8	298033.3
## 2008	56357645.3	1075864.4	3123679.6	1364401.5	4728136.2
##	Jun	JuI	Aug	Sep	Oct
## 2000	3943336.1	1544034.0	1865173.9	1144671.5	850278.7
## 2001	1613915.5	2540327.9	1221633.4	3586310.5	1135851.2
## 2002	1535143.8	3368980.5	3110483.9	4227353.5	4041043.2
## 2003	3830353.2	2409861.8	6577461.5	1147296.3	825461.1
## 2004	1749400.2	1578612.1	3050330.4	2165384.7	1302521.4
## 2005	467399.2	2095285.8	10012942.8	4104886.3	27725665.5
## 2006	114625709.3	11849265.0	436775.6	711546.0	1014926.9
## 2007	1221112.1	1387284.0	1434429.9	390666.5	12571170.3
## 2008	7899175.6	3782123.8	966685.5	853580.4	2735991.6
##	Nov	Dec			
## 2000	5296861.4	1181022.6			
## 2001	2537418.6	1216094.2			
## 2002	1751863.7	1782043.5			
## 2003	1645904.0	1109802.0			
## 2004	3591160.0	1077254.9			
## 2005	1982035.2	2778488.0			
## 2006	458560.1	703889.5			
## 2007	1082589.4	3798755.4			
## 2008	5043038.7	623933.4			

Fuente:Elaboración propia

Apéndice B

BASE DE DATOS PARA VULCANOLOGÍA

Emisiones de CO_2 año 2005 volcán de San Salvador

##	[1]	575.202881	903.454651	1203.251343	1498.982666	1359.755249
##	[6]	1335.364990	833.332703	541.666260	223.577148	326.219360
##	[11]	662.601013	384.146149	662.601013	476.625824	1582.316040
##	[16]	398.373718	727.641418	1042.682129	717.479126	570.121582
##	[21]	1013.210571	453.251709	860.771728	1422.763184	1344.510986
##	[26]	2207.315186	1861.787354	1442.072510	2007.112671	2236.786865
##	[31]	1325.202148	321.138061	273.373840	349.593292	446.137940
##	[36]	490.853363	475.609375	2651.421387	356.707062	786.584778
##	[41]	1245.934082	1905.486694	838.413940	802.844849	897.357056
##	[46]	453.251801	975.609131	1228.657715	718.495239	1596.543701
##	[51]	2234.754639	1356.706421	1946.136841	843.495178	1689.022949
##	[56]	1082.316162	893.291992	271.341370	512.194824	214.430801
##	[61]	589.430481	868.901672	1876.015259	1782.518799	398.373718
##	[66]	333.333130	393.292419	347.560761	501.015991	512.194824
##	[71]	935.974670	624.999512	898.373413	1165.649658	1217.478516
##	[76]	1318.088623	916.666016	1997.965698	952.235107	493.902191
##	[81]	946.137390	397.357452	696.137756	1058.942261	1084.348877
##	[86]	902.438232	1572.153076	1124.999023	1967.478149	913.617248
##	[91]	1076.218750	619.918213	721.544190	1245.934326	857.722839
##	[96]	1206.299927	1575.202515	1961.380615	2443.087646	1881.096191
##	[101]	1088.413818	1712.397339	1301.828369	845.527649	455.284210
##	[106]	558.942749	784.552307	1058.942139	1216.462158	1452.234619
##	[111]	2549.795166	2853.656494	1159.552002	1646.340210	1433.942261
##	[116]	777.438293	861.788086	907.519409	741.869446	1003.047974
##	[121]	1769.307861	1452.234375	2287.599609	2688.006348	1973.575684
##	[126]	1553.860718	1135.161621	588.414307	523.373596	283.536438
##	[131]	246.951126	982.722839	745.934326	1056.909546	768.292114
##	[136]	702.235107	2272.356445	1753.047607	564.024048	776.422119
##	[141]	923.779724	732.722900	740.852966	871.950500	886.178162
##	[146]	1116.869263	1673.779419	1233.738892	1545.730591	1154.470459
##	[151]	773.373291	429.877777	285.568970	246.951126	213.414551
##	[156]	213.414566	914.633789	774.389526	1129.063965	1029.470703
##	[161]	1425.812134	1581.299561	933.942383	457.316894	235.772278
##	[166]	690.040161	682.926209	1138.210449	1158.535767	1737.803589
##	[171]	1067.072021	1403.454346	910.568481	758.129456	891.259460
##	[176]	666.666199	213.414551	296.747833	523.373657	623.983215
##	[181]	1600.608643	1393.291992	1140.243164	1722.559570	449.186676
##	[186]	341.463226	402.438782	332.316925	670.731140	685.975159

Fuente:Elaboración propia

Bibliografía

- [1] STEFANO MARIA IACUS y DAVIDE LA TORRE: *APPROXIMATING DISTRIBUTION FUNCTIONS BY ITERATED FUNCTION SYSTEMS*, Department of Economics, Business and Statistics, Faculty of Political Science, University of Milan: 2004
- [2] W.O CAMPOS y G.V CHAVEZ: *Introducción a la geometría fractal*, tesis, San Salvador: 2004
- [3] GALEN R. SHORACK y JON A. WELLNER: *Empirical Processes with Applications to Statistics*, John Wiley y Sons, New York, 1986
- [4] D.LA TORRE F.MENDIVIL y E.R VRSCAY: *Iterated function systems on functions of bounded variation*, 2015
- [5] BRUNO FORTE y EDWARD R. VRSCAY: *Theory of Generalized Fractal Transforms*, 1995
- [6] SABOGAL SONIA y ARENAS GILBERTO *Introducción a la Geometría Fractal*, Bocaranga: 2011
- [7] LUIS CEFERINO FRANCO ARBALÁEZ y HERMILSON VELÁSQUEZ CEBALLOS: *CUANTIFICACIÓN DEL RIESGO OPERACIONAL UTILIZANDO SISTEMAS DE FUNCIONES ITERADAS*, Universidad de Medellín: 2011
- [8] MICHAEL BARNSLEY: *FRACTAL EVERYWHERE*, School of Mathematics Georgia Institute of Technology Atlanta, Georgia and Iterated Systems, Inc. Atlanta, Georgia
- [9] RICARDO UMBERTO SILVIO MÉNDEZ: *Diseño e Implementación de Indicadores Clave de Riesgo Operacional para la Infraestructura Tecnológica de Mercantil de Seguros*, Universidad Simón Bolívar: 2010
- [10] DONALD L. TURCOTTE: *FRACTALS AND CHAOS IN GEOLOGY AND GEOPHYSICS*, Cornell University
- [11] KENNETH FALCONER : *FRACTAL GEOMETRY, Mathematical Foundations and Applications*, University of St Andrews, UK
- [12] <http://ucua.ujaen.es/jnavas/mayores/fractales.pdf>