



UNIVERSIDAD NACIONAL “PEDRO RUIZ GALLO”
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA



**“ La Ecuación Funcional Aditiva de Cauchy de
 \mathbb{R} en \mathbb{R} ”**

TESIS

**“PRESENTADA PARA OBTENER EL TÍTULO
PROFESIONAL DE LICENCIADO EN MATEMÁTICAS”.**

Presentado por:

Bach. Mat. Cabanillas Banda, Wilson Alberto

Bach. Mat. Vera Rubio, José Gilmer

Asesor:

Mg. Malca Villalobos, Amado

LAMBAYEQUE – PERÚ

2017

UNIVERSIDAD NACIONAL“ PEDRO RUIZ GALLO”

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

ESCUELA PROFESIONAL DE MATEMÁTICA

Los firmantes por la presente certifican que han leído y recomiendan a la Facultad de Ciencias Físicas y Matemáticas la aceptación de la tesis titulada “ **La Ecuación Funcional Aditiva de Cauchy de \mathbb{R} en \mathbb{R}** ”, presentada por los Bachilleres en Matemáticas, en el cumplimiento parcial de los requisitos necesarios para la obtención del título profesional de Licenciado en Matemáticas.

Dra. Ortiz Basauri, Gloria María

Presidente Jurado de Tesis

M.Sc. Abramonte Ato, Carlos Arturo

Secretario Jurado de Tesis

Mg. Pérez Herrera, Adelmo

Vocal Jurado de Tesis

Fecha de Defensa: 15 de Febrero del 2017

UNIVERSIDAD NACIONAL“ PEDRO RUIZ GALLO”

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

ESCUELA PROFESIONAL DE MATEMÁTICA

**“ La Ecuación Funcional Aditiva de Cauchy de
 \mathbb{R} en \mathbb{R} ”**

Bach. Mat. Cabanillas Banda, Wilson Alberto

Autor

Bach. Mat. Vera Rubio, José Gilmer

Autor

Mg. Malca Villalobos, Amado

Asesor

Lambayeque – Perú

Febrero - 2017

Agradecimiento

Agradecemos a Dios por la vida y la salud y sobre todo por su Hijo **Jesucristo**, pues por medio de Él es que alcanzamos la salvación de nuestras almas, la felicidad y gozo verdadero.

Agradecemos a nuestro profesor Mg. Amado Malca por ser nuestro asesor en esta tesis.

Resumen

El estudio de las funciones aditivas se remonta a A.M Legendre, quien primero intentó determinar la solución de la ecuación funcional aditiva de Cauchy

$$f(x+y) = f(x) + f(y), \quad (1)$$

para todo $x, y \in \mathbb{R}$, donde $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función. Sin embargo el estudio sistemático de la ecuación funcional aditiva de Cauchy (1) fue iniciado por A.L. Cauchy en su libro *Cours d'Analyse* en 1821. Además estudió también otras tres ecuaciones funcionales :

$$f(xy) = f(x) + f(y), \quad (2)$$

$$f(x+y) = f(x)f(y), \quad (3)$$

$$f(xy) = f(x)f(y). \quad (4)$$

A lo largo de la presente tesis se realiza un estudio de la ecuación funcional aditiva de Cauchy (1), donde se responde a la siguiente interrogante :

¿Bajo qué condiciones una función $f : \mathbb{R} \rightarrow \mathbb{R}$ que satisface la ecuación funcional aditiva de Cauchy (1) es lineal, de la forma $f(x) = c \cdot x$, donde $c \in \mathbb{R}$?

Además se presenta la solución de las tres ecuaciones funcionales restantes (2), (3), (4).

Comenzamos mostrando que una condición es la continuidad, y que otras más débiles son integrabilidad local, continuidad en un punto, estar acotada superior o inferiormente. También se explora el comportamiento de funciones discontinuas que satisfacen la ecuación funcional (1) y se muestra que ellas manifiestan un comportamiento muy extraño: sus gráficas son subconjuntos densos del plano \mathbb{R}^2 . Además se discute brevemente las bases de Hamel con el fin

de usarlas para construir funciones aditivas discontinuas y se finaliza el presente trabajo exhibiendo dos aplicaciones : la caracterización de distribución geométrica y la caracterización de la distribución normal.

Abstract

The study of additive functions dates back to A.M. Legendre who first attempted to determine the solution of the additive Cauchy functional equation

$$f(x+y) = f(x) + f(y), \quad (5)$$

for all $x, y \in \mathbb{R}$, where $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function. However systematic study of the additive Cauchy functional equation (5) was initiated by A.L. Cauchy in his book *Cours d'Analyse* in 1821. He also studied another three functional equations, which are :

$$f(xy) = f(x) + f(y), \quad (6)$$

$$f(x+y) = f(x)f(y), \quad (7)$$

$$f(xy) = f(x)f(y). \quad (8)$$

Throughout this thesis, a study of the additive Cauchy functional equation (5) is carried out, where the following question is answered :

Which are the conditions that should be imposed upon f , so that f may be linear?

where $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function satisfying the equation (5). Also the solution of the three remaining functional equations (6), (7) and (8) is shown.

We begin exhibiting that a condition is continuity, and that another are local integrability, continuity at only one point, bounded from the above or from below. Also, we explore the behavior of nonlinear discontinuous functions that satisfies (5) and show that they display a very strange behavior: their graphs are dense subsets in the plane \mathbb{R}^2 . Furthermore, we briefly discuss the Hamel basis in order to use them to build discontinuous additive functions.

Lastly, we finish up this thesis presenting two applications : Characterization of Geometric Distribution and Characterization of Normal Distribution.

Introducción

Las ecuaciones funcionales son una parte fundamental del Análisis Matemático, que tienen estrechos vínculos con otras áreas de la matemática, como el Álgebra o la Geometría. Históricamente podemos afirmar que las primeras ecuaciones funcionales aparecieron relacionadas con ciertos problemas geométricos y/o físicos. Sólo por mostrar un ejemplo, citemos que en 1352 Nicolás Oresme (1320 - 1382) se interesó por la ecuación funcional

$$\frac{f(x_1) - f(x_2)}{f(x_2) - f(x_3)} = \frac{x_1 - x_2}{x_2 - x_3} \quad \text{para todo } x_1, x_2, x_3 \in \mathbb{R} \text{ tal que } x_1 > x_2 > x_3.$$

El matemático N. Oresme se refería a las funciones que resuelven esta ecuación como *cantidades uniformemente deformadas* y se planteó su caracterización. Posteriormente, en 1638 Galileo Galilei (1564-1642) utilizó para la verificación de su tesis sobre la caída libre de los cuerpos, la ecuación funcional

$$\frac{f((n+1)t) - f(nt)}{f(nt) - f((n-1)t)} = \frac{2n+1}{2n-1} \quad \text{para todo } t \in \mathbb{R} \text{ y todo } n \in \mathbb{N},$$

que es satisfecha por todo monomio cuadrático $f(x) = ax^2$. Posteriormente, en 1769, en conexión con el problema físico de describir el comportamiento de la superposición (o suma) de dos fuerzas a partir de sus propiedades axiomáticas, problema que sirve para motivar (y, de hecho, justificar) la ley del paralelogramo, D'Alembert estudiaría la ecuación funcional

$$f(x+y) + f(x-y) = 2f(x)f(y).$$

Con el tiempo muchas otras ecuaciones funcionales fueron apareciendo progresivamente en la literatura. Especialmente importantes fueron las aportaciones de Cauchy (1789-1857), quien

introdujo y estudió en su famoso texto de 1821 ([7]), las siguientes cuatro ecuaciones:

$$f(x+y) = f(x) + f(y),$$

$$f(xy) = f(x) + f(y),$$

$$f(x+y) = f(x)f(y),$$

$$f(xy) = f(x)f(y),$$

(todas ellas planteadas originalmente para funciones reales de variable real). Cauchy se interesó por estas ecuaciones porque ellas resultan de gran utilidad si queremos definir cuidadosamente algunas de las funciones elementales, como el logaritmo o las funciones trigonométricas.

El motivo de este trabajo fue la necesidad de tener una referencia de carácter elemental y actualizada, que sirva como guía introductoria a esta materia para los estudiantes de matemáticas. El contenido de este trabajo está dirigido a los estudiantes de la escuela profesional de Matemáticas que tengan un conocimiento del curso de Análisis Real en una variable y conocimientos básicos de Probabilidad.

Para el desarrollo de la tesis se ha tomado como guía el libro de P. Sahoo y P. Kannappan: *Introduction to Functional Equations* ([13]).

En el Capítulo I, se introducen las nociones básicas relacionadas a Bases Hamel y su existencia, Variables Aleatorias y su distribución, en particular la distribución geométrica y normal y el Método de Estimación de Máxima Verosimilitud. Lo mencionado anteriormente constituyen el fundamento teórico matemático que usaremos a lo largo de la tesis.

En el Capítulo II, se estudia la ecuación funcional aditiva de Cauchy y se muestra que funciones aditivas continuas, localmente integrables, continuas en un punto, acotadas superiormente o superiormente son lineales. Se explora además el comportamiento de funciones aditivas discontinuas y se prueba que ellas presentan un comportamiento muy extraño: sus gráficas son densas en el plano. Finalmente, se discute brevemente sobre las bases de Hamel y su uso para construir funciones aditivas discontinuas. Una discusión sobre funciones aditivas complejas es también presentada.

En el Capítulo III, se presenta las tres ecuaciones funcionales de Cauchy restantes, y se halla su solución general.

En el Capítulo IV, se presenta dos aplicaciones : usando dos ecuaciones funcionales de Cauchy se caracteriza la distribución geométrica y la distribución normal.

Los Autores.

Índice general

Resumen	I
Abstract	III
Introducción	V
4 CAPÍTULO 1	
Preliminares	
1.1. Conjuntos Ordenados	4
1.2. El Axioma de Elección y El Lema de Zorn	6
1.3. Bases de Hamel	8
1.4. Variables Aleatorias y Distribuciones de Probabilidad	11
1.5. Muestras, Estadísticas y Estimadores	22
1.6. El Método de Máxima Verosimilitud	24
1.7. Parámetros y Familias de Ubicación y de Escala	26
29 CAPÍTULO 2	
La Ecuación Funcional Aditiva de Cauchy	
2.1. Ecuaciones Funcionales y Definición	29
2.2. Solución Lineal de la Ecuación Funcional Aditiva de Cauchy	30
2.3. Otros Criterios para la Linealidad	36
2.4. Construcción de una Función Aditiva Discontinua	41
2.5. Funciones Aditivas sobre el Plano Complejo	43

49 | CAPÍTULO 3

Otras Ecuaciones Funcionales de Cauchy

3.1. Solución de la Ecuación Exponencial de Cauchy	49
3.2. Solución de la Ecuación Logarítmica de Cauchy	51
3.3. Solución de la Ecuación Multiplicativa de Cauchy	53

58 | CAPÍTULO 4

Algunas Aplicaciones de la Ecuación Funcional de Cauchy

4.1. Caracterización de la Distribución Geométrica	58
4.2. Caracterización de la Distribución Normal	66

Conclusiones	70
---------------------	-----------

Bibliografía	73
---------------------	-----------

Capítulo 1:

Preliminares

En este capítulo se presenta brevemente las herramientas matemáticas necesarias para comprender el resto de la tesis.

1.1 Conjuntos Ordenados

Definición 1.1. Una relación binaria sobre un conjunto X es un subconjunto no vacío R de $X \times X$.

Definición 1.2. Dado un conjunto X , una relación binaria $R \subset X \times X$ es llamada un orden parcial sobre X si ella es

1. reflexiva: $(x, x) \in R$ para todo $x \in X$,
2. antisimétrica: si (x, y) y (y, x) están en R , entonces $x = y$,
3. transitiva: si (x, y) y (y, z) están en R , entonces $(x, z) \in R$.

Se escribe \leq para denotar un orden parcial sobre X . Un conjunto X equipado con un orden parcial \leq es llamado un conjunto parcialmente ordenado y es denotado por (X, \leq) . A continuación presentamos dos ejemplos de orden parcial.

Ejemplo 1.1. Considere $X = \mathbb{N}$ y defina una relación \leq sobre $\mathbb{N} \times \mathbb{N}$ como sigue: dados $a, b \in \mathbb{N}$

$$a \leq b \Leftrightarrow a \text{ divide } b$$

Es fácil ver que \leq satisface todas las condiciones de la definición anterior y por lo tanto es una orden parcial sobre \mathbb{N} .

Otro ejemplo clásico, es el siguiente:

Ejemplo 1.2. Considere A un conjunto no vacío y sea $\mathbf{P}(A)$ el conjunto potencia de A . Luego defina una relación \leq sobre $\mathbf{P}(A) \times \mathbf{P}(A)$ como sigue : dados $A, B \in \mathbf{P}(A)$

$$A \leq B \Leftrightarrow A \subset B$$

Es fácil ver que \leq satisface todas las condiciones de la definición anterior y por lo tanto es un orden parcial sobre $\mathbf{P}(A)$.

Definición 1.3. Sea X un conjunto parcialmente ordenado con orden parcial \leq . Dos elementos x, y de X se dice que son comparables si $x \leq y$ o $y \leq x$. Además, el conjunto X es llamado totalmente ordenado si para cualquier $x, y \in X$, se tiene que x e y son comparables y a dicho orden se le denomina un orden total.

Definición 1.4. Sea X un conjunto parcialmente ordenado. Una cadena \mathcal{C} es un subconjunto de X que está totalmente ordenado cuando se le dota del orden que hereda de X .

Definición 1.5. Sea (X, \leq) un conjunto parcialmente ordenado y sea $A \subseteq X$.

- Un elemento $x \in X$ es una cota superior de A si $a \leq x$ para todo $a \in A$.
- Si x_0 es una cota superior de A y si cualquier otra cota superior x de A satisface $x_0 \leq x$; entonces se dice que x_0 es el supremo de A .
- Un elemento $x_0 \in X$ se dice que es el máximo en X si $x \leq x_0$ para todo $x \in X$.

Definición 1.6. Sea (X, \leq) un conjunto parcialmente ordenado y sea $A \subseteq X$. Un elemento $m \in A$ se dice que es un elemento maximal en A si

$$\forall p \in A, m \leq p \Rightarrow m = p.$$

Las definiciones de ínfimo, mínimo y elemento minimal se introducen de modo completamente similar.

A continuación presentamos algunos ejemplos para aclarar un poco las definiciones dadas anteriormente.

- Ejemplo 1.3.**
1. El conjunto \mathbb{R} de números reales, con el orden usual \leq es un conjunto totalmente ordenado, en particular, parcialmente ordenado, que **no** tiene elementos maximales.
 2. Sea $S \subset \mathbb{N}$ consistiendo de todas las potencias de 2, con el orden parcial del Ejemplo 1.1. Entonces S es un conjunto totalmente ordenado, es decir, es una cadena en \mathbb{N} .
 3. El conjunto $P = \{2, 4, 8, 3, 9, 27\}$ es un conjunto parcialmente ordenado con el orden del Ejemplo 1.1. Note que P tiene dos elementos maximales, a saber, 8 y 27. Mientras que el conjunto $Q = \{2, 3, 5, 7, 11\}$ con el mismo orden, es parcialmente ordenado pero todo elemento de Q es maximal y minimal a la vez!
 4. Sea $U = \{1, 2\}$ y su conjunto potencia $\mathbf{P}(U) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ con el orden parcial \subset . Entonces

$$\begin{aligned} \text{máximo de } \mathbf{P}(U) &= \bigcup_{A \in \mathbf{P}(U)} A = \{1, 2\} \\ \text{mínimo de } \mathbf{P}(U) &= \bigcap_{A \in \mathbf{P}(U)} A = \emptyset \end{aligned}$$

1.2 El Axioma de Elección y El Lema de Zorn

El Axioma de Elección es un axioma de la teoría de conjuntos que postula la existencia de ciertos objetos sin dar ninguna indicación de cómo obtenerlos. Desde su aparición ha resultado ser un principio muy controvertido. Su aceptación, en términos generales, se sustenta sobre la creencia de que nuestra percepción sobre los conjuntos finitos se puede ampliar a los conjuntos infinitos, pero más allá de eso, el principal argumento para su aceptación es que dicho axioma

es tremendamente útil (ver, por ejemplo [11] para una discusión sobre este tema). Muchos resultados importantes y fundamentales en Análisis Real, Análisis Funcional, Álgebra, etc, se pueden demostrar si se acepta, sin limitaciones, el Axioma de Elección.

Definición 1.7. Una función de elección es una aplicación f , definida en una familia \mathcal{C} de conjuntos no vacíos tal que para todo conjunto $A \in \mathcal{C}$, $f(A) \in A$, o dicho de otra forma, la función de elección f elige exactamente un elemento de cada conjunto en \mathcal{C} .

$$f : \mathcal{C} \longrightarrow \bigcup_{A \in \mathcal{C}} A$$

Entre las numerosas formas equivalentes del Axioma de Elección que existen, tal vez una de las más populares es la siguiente :

Axioma 1. (Axioma de Elección). Si $(X_\alpha)_{\alpha \in J}$ es una familia de conjuntos tal que X_α es no vacío para todo $\alpha \in J$, entonces existe al menos una función de elección para la familia $(X_\alpha)_{\alpha \in J}$.

También otra forma equivalente del Axioma de Elección, es el así llamado Lema de Zorn, un resultado formulado por M. Zorn en 1935 y que resulta ser extremadamente útil en varias ramas del quehacer matemático. Por ejemplo, el Lema de Zorn es fundamental para demostrar resultados importantes tales como el Teorema de Hahn-Banach y en especial **la prueba de la existencia de una base de Hamel en cualquier espacio vectorial no trivial**.

Lema 1.1. (Lema de Zorn) Sea (X, \leq) un conjunto no vacío parcialmente ordenado. Si cualquier cadena \mathcal{C} en X posee una cota superior, entonces X posee un elemento maximal.

Con mucha frecuencia, el Lema de Zorn se utiliza para familias \mathcal{F} de subconjuntos de un conjunto X , ordenados por la relación de inclusión \subset , con la propiedad de que cualquier cadena $\mathcal{C} \subseteq X$, la unión $\bigcup_{C \in \mathcal{C}} C$, también esté en \mathcal{F} .

En la siguiente sección hablaremos acerca de lo que es una base de Hamel, demostrando luego su existencia para todo espacio vectorial, este hecho será de crucial importancia para poder construir una función aditiva discontinua en la Sección 2.4.

1.3 Bases de Hamel

Consideremos X un \mathbb{K} -espacio vectorial, donde \mathbb{K} representa algún cuerpo, ya sea \mathbb{Q} , \mathbb{R} o \mathbb{C} .

Definición 1.8. Sea $\mathcal{A} \subset X$ un conjunto finito, $\mathcal{A} \neq \emptyset$, digamos que $\mathcal{A} = \{x_1, \dots, x_n\}$. Diremos que \mathcal{A} es linealmente independiente si

$$\sum_{i=1}^n \lambda_i x_i = 0, (\lambda_i \in \mathbb{K}, i = 1, \dots, n)$$

implica que $\lambda_i = 0, i = 1, \dots, n$.

Definición 1.9. Sea $\mathcal{B} \subset X$, $\mathcal{B} \neq \emptyset$. Diremos que \mathcal{B} es linealmente independiente si todo subconjunto finito no vacío $\mathcal{A} \subset \mathcal{B}$ es linealmente independiente.

Definición 1.10. Sea $\mathcal{C} \subset X$ no vacío. Diremos que \mathcal{C} es maximal linealmente independiente si \mathcal{C} es linealmente independiente y además, si \mathcal{D} es tal que $\mathcal{D} \supsetneq \mathcal{C}$ entonces \mathcal{D} no es linealmente independiente.

Definición 1.11. Un subconjunto maximal linealmente independiente de X se llama base de Hamel de X .

El teorema sobre existencia de bases de Hamel es importante, sin embargo, no provee una construcción de la base de Hamel.

Teorema 1.1. *Todo espacio vectorial no trivial $X \neq \{0\}$ posee una base de Hamel.*

Demostración. La demostración de este teorema hará uso del Lema de Zorn. Como $X \neq \{0\}$ existe $x \in X$ tal que $x \neq 0$. Evidentemente $\{x\}$ es un conjunto linealmente independiente.

Sea ahora $\mathcal{S} := \{L \subset X, L \text{ es linealmente independiente}\}$. Se sigue fácilmente que $\mathcal{S} \neq \emptyset$, pues $\{x\} \in \mathcal{S}$. Ahora definimos una relación \leq sobre $\mathcal{S} \times \mathcal{S}$ como sigue: dados $A, B \in \mathcal{S}$

$$A \leq B \Leftrightarrow A \subset B.$$

Se verifica fácilmente que \leq es un orden parcial en \mathcal{S} por lo cual (\mathcal{S}, \leq) es un conjunto parcialmente ordenado.

Antes de continuar, recuerde que una cadena en \mathcal{S} es un subconjunto \mathcal{C} de \mathcal{S} tal que si

$$A, B \in \mathcal{C} \Rightarrow A \subset B \text{ o bien } B \subset A.$$

Procedemos entonces a tomar una cadena \mathcal{C} en \mathcal{S} y definimos

$$M_{\mathcal{C}} := \bigcup_{L \in \mathcal{C}} L.$$

Es evidente que $M_{\mathcal{C}}$ es un subconjunto de X y que

$$L \subset M_{\mathcal{C}}, \quad \forall L \in \mathcal{C}. \quad (1.1)$$

La ecuación (1.1) nos dice claramente que $M_{\mathcal{C}}$ es una cota superior para \mathcal{C} . Ahora bien, nos faltaría probar que $M_{\mathcal{C}} \in \mathcal{S}$, es decir, probar que $M_{\mathcal{C}}$ es linealmente independiente, para luego aplicar el Lema de Zorn.

En efecto, sea N un subconjunto finito no vacío de $M_{\mathcal{C}}$, digamos, $N = \{x_1, x_2, \dots, x_n\}$. Como $N \subset M_{\mathcal{C}}$, entonces para cada x_i existe un conjunto $L_i \in \mathcal{C}$ tal que $x_i \in L_i$ para todo $i = 1, \dots, n$. Pero como \mathcal{C} es una cadena, existe un conjunto, digamos $L_k \in \mathcal{C}$, que contiene a todos los demás conjuntos L_i con $i \neq k$ y por lo tanto $L_k \supset \{x_1, x_2, \dots, x_n\}$ y como L_k es linealmente independiente entonces $N = \{x_1, x_2, \dots, x_n\}$ también lo es y debido a la arbitrariedad de N se sigue que $M_{\mathcal{C}}$ es linealmente independiente, es decir que, $M_{\mathcal{C}} \in \mathcal{S}$.

En resumen, se ha demostrado que toda cadena \mathcal{C} en \mathcal{S} tiene una cota superior $M_{\mathcal{C}}$ que está en \mathcal{S} . Por el lema de Zorn, existe un elemento maximal $B \in \mathcal{S}$, el cual constituye una base para X . \square

En particular el teorema anterior muestra la existencia de una base de Hamel para \mathbb{R} visto como \mathbb{Q} -espacio vectorial.

Teorema 1.2. *Sea \mathcal{B} una base de Hamel para X . Entonces, para cada $x \in X$, $x \neq 0$, existen: $n \in \mathbb{N}$, elementos únicos x_1, \dots, x_n de \mathcal{B} y escalares $\lambda_1, \dots, \lambda_n$ únicos, tales que $|\lambda_i| > 0$ y $x = \sum_{i=1}^n \lambda_i x_i$.*

Demostración. Existencia: Sea $x \in X$, $x \neq 0$. Si $x \in \mathcal{B}$, una tal representación es clara con $n = 1$, $x_1 = x$ y $\lambda_1 = 1$. Supongamos ahora que $x \notin \mathcal{B}$. Entonces, el subconjunto $\mathcal{B} \cup \{x\}$

de X contiene propiamente a \mathcal{B} , y ya que \mathcal{B} es maximal linealmente independiente, se tiene que $\mathcal{B} \cup \{x\}$ no es linealmente independiente, luego existe $\mathcal{A} \subset \mathcal{B} \cup \{x\}$ finito y no vacío tal que \mathcal{A} no es linealmente independiente. Forzosamente $x \in \mathcal{A}$, pues de otro modo tendríamos $\mathcal{A} \subset \mathcal{B}$ y así \mathcal{A} debería ser linealmente independiente, lo cual es imposible.

Supongamos que $\mathcal{A} = \{x, x_1, \dots, x_n\}$. Como \mathcal{A} no es linealmente independiente existen escalares $\lambda, \lambda_1, \dots, \lambda_n$ no todos nulos tales que

$$\lambda x + \sum_{i=1}^n \lambda_i x_i = 0.$$

Afirmamos que $\lambda \neq 0$, porque en otro caso tendríamos que $\sum_{i=1}^n \lambda_i x_i = 0$ con $x_i \in \mathcal{B}$, para todo $i = 1, \dots, n$ y siendo \mathcal{B} linealmente independiente tendríamos que $\lambda_i = 0$ para todo $i = 1, \dots, n$, lo cual es una contradicción.

Así, $\lambda \neq 0$ y por consiguiente

$$x = \sum_{i=1}^n -\frac{\lambda_i}{\lambda} x_i.$$

Ahora, demostraremos la unicidad. Supongamos que existen $m, n \in \mathbb{N}$, $m \neq n$ y elementos $x_1, \dots, x_m, y_1, \dots, y_n$ tales que

$$x_i \in \mathcal{B} \quad \forall i = 1, \dots, m, \quad y_j \in \mathcal{B} \quad \forall j = 1, \dots, n$$

y sean $\lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_n$ escalares tales que

$$|\lambda_i| > 0 \quad \forall i = 1, \dots, m, \quad |\mu_j| > 0 \quad \forall j = 1, \dots, n$$

y

$$x = \sum_{i=1}^m \lambda_i x_i = \sum_{j=1}^n \mu_j y_j.$$

Si $m > n$ entonces existe $\exists i_0 \in \{1, \dots, m\}$ tal que x_{i_0} se expresa como combinación lineal de x_i , $i \in \{1, \dots, m\} \setminus \{i_0\}$ y y_j , $j \in \{1, \dots, n\}$, lo cual es imposible, puesto que los x_i, y_j constituyen un subconjunto linealmente independiente de \mathcal{B} . De igual forma se prueba que $n \leq m$, por lo tanto $m = n$.

Ahora, sea i tal que $1 \leq i \leq m$. Si $x_i \neq y_j$ para cada $1 \leq j \leq n$, entonces $\lambda_i = 0$, ya que de otro modo quedaría expresado x_i en términos de los otros elementos y esto es imposible. De nuevo contradecimos el hecho de que $|\lambda_i| > 0$. Así, para cada $1 \leq i \leq m$ existe i_j , $1 \leq i_j \leq n$ tal que $x_i = y_{i_j}$. De nuevo, la independencia lineal asegura que $\lambda_i = \mu_{i_j}$. \square

1.4 Variables Aleatorias y Distribuciones de Probabilidad

En esta sección introduciremos brevemente los conceptos de σ -álgebra, medida de probabilidad o simplemente probabilidad, variable aleatoria, función de distribución, función de probabilidad y de densidad de una variable aleatoria. También presentaremos lo que es una distribución geométrica y normal.

Definición 1.12. Sea Ω un conjunto no vacío. Una colección no vacía \mathcal{F} de subconjuntos de Ω es llamada una σ -álgebra si cumple las siguientes condiciones:

1. si $A \in \mathcal{F}$ entonces $A^c \in \mathcal{F}$,
2. si $A_j \in \mathcal{F}$ para $j \geq 1$, entonces $\bigcup_{j=1}^{\infty} A_j \in \mathcal{F}$.

A continuación definimos lo que es un evento, visto como un miembro de una σ -álgebra.

Definición 1.13. Sea Ω un conjunto no vacío. Todo subconjunto $A \subseteq \Omega$ tal que $A \in \mathcal{F}$ será llamado evento. Ω es el evento cierto, \emptyset es el evento imposible. Si $\omega \in \Omega$ es tal que $\{\omega\} \in \mathcal{F}$, el evento $\{\omega\}$ será llamado elemental o simple.

Definición 1.14. Una medida de probabilidad o probabilidad es una función $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ definida sobre la σ -álgebra $\mathcal{F} \subseteq 2^{\Omega}$ que satisface las siguientes condiciones:

1. $\mathbb{P}(\emptyset) = 0$ y $\mathbb{P}(\Omega) = 1$,
2. si los conjuntos $A_j \in \mathcal{F}$ para $j \geq 1$ son disjuntos dos a dos entonces:

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Definición 1.15. Un espacio de probabilidad es una terna $(\Omega, \mathcal{F}, \mathbb{P})$, donde

- (a) Ω es un conjunto no vacío,
- (b) \mathcal{F} es una σ -álgebra de subconjuntos de Ω y

(c) \mathbb{P} es una probabilidad definida en \mathcal{F} .

En el siguiente ejemplo presentamos un espacio de probabilidad discreto.

Ejemplo 1.4. Considere el experimento de lanzar una moneda, entonces existen dos posibles resultados, cara (C) y sello (S). El espacio muestral es $\Omega = \{C, S\}$, la σ -álgebra a considerar es

$$\mathcal{F} := \{\{C, S\}, \{C\}, \{S\}, \emptyset\}$$

y los eventos son

$$\{C, S\}, \{C\}, \{S\}, \emptyset.$$

Si la moneda es justa, es decir, si una cara o sello son igualmente probables, entonces podemos asignar igual probabilidades a los dos posibles resultados y definir \mathbb{P} como sigue :

$$\mathbb{P}(\{C\}) := \mathbb{P}(\{S\}) := 0,5$$

Además, $\mathbb{P}(\{C, S\}) := 1$ y $\mathbb{P}(\emptyset) := 0$, de esta manera tenemos una medida de probabilidad

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1],$$

definida sobre toda la σ -álgebra \mathcal{F} . La terna $(\Omega, \mathcal{F}, \mathbb{P})$ se le llama un espacio de probabilidad discreto.

A continuación mostramos dos propiedades muy útiles y conocidas de la medida de probabilidad, que nos hablan acerca de su continuidad inferior y superior.

Proposición 1.1. (*Continuidad superior*) Sea $\{A_n : n \in \mathbb{N}\}$ una sucesión no decreciente de eventos, esto es, $A_1 \subseteq A_2 \subseteq \dots$. Entonces

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n). \quad (1.2)$$

Demostración. Como $A_n \subseteq A_{n+1}$ tenemos que $\mathbb{P}(A_n) \leq \mathbb{P}(A_{n+1})$. Por lo tanto la sucesión numérica $\{\mathbb{P}(A_n) : n \in \mathbb{N}\}$ es no decreciente y acotada superiormente. Entonces el límite de esta sucesión existe y el lado derecho de la igualdad (1.2) tiene sentido. Defina los eventos

$$\begin{aligned}
B_1 &= A_1, \\
B_n &= A_n - A_{n-1} \quad \text{para } n \geq 2.
\end{aligned}$$

La sucesión $\{B_n : n \in \mathbb{N}\}$ es una colección de eventos disjuntos dos a dos y tal que

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n.$$

Por lo tanto

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(B_n) \\
&= \mathbb{P}(B_1) + \sum_{n=2}^{\infty} \mathbb{P}(B_n) \\
&= \mathbb{P}(A_1) + \sum_{n=2}^{\infty} \mathbb{P}(A_n - A_{n-1}) \\
&= \mathbb{P}(A_1) + \sum_{n=2}^{\infty} (\mathbb{P}(A_n) - \mathbb{P}(A_{n-1})) \\
&= \mathbb{P}(A_1) + \lim_{m \rightarrow \infty} \sum_{n=2}^m (\mathbb{P}(A_n) - \mathbb{P}(A_{n-1})) \\
&= \mathbb{P}(A_1) + \lim_{m \rightarrow \infty} \mathbb{P}(A_m) - \mathbb{P}(A_1) \\
&= \lim_{m \rightarrow \infty} \mathbb{P}(A_m).
\end{aligned}$$

□

Las medidas de probabilidad también son continuas respecto de sucesiones no crecientes de eventos. Esta afirmación es el contenido del siguiente resultado que se demuestra a partir de la proposición anterior.

Proposición 1.2. (*Continuidad inferior*) Sea $\{A_n : n \in \mathbb{N}\}$ una sucesión no creciente de eventos, esto es, $A_1 \supseteq A_2 \supseteq \dots$. Entonces

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

Demostración. Observe que si $A_n \supseteq A_{n+1}$ entonces $A_n^c \subseteq A_{n+1}^c$. Por la proposición anterior,

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n^c).$$

Aplicando las leyes de De Morgan,

$$1 - \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) = \lim_{n \rightarrow \infty} (1 - \mathbb{P}(A_n^c)),$$

de donde se sigue fácilmente el resultado. \square

Definición 1.16. (Probabilidad Condicional) Sean A y B dos eventos. La probabilidad de A dado el evento B , denotada por $\mathbb{P}(A|B)$, es definida de la siguiente manera:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)},$$

cuando $\mathbb{P}(B) > 0$.

Definición 1.17. Dado $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Una variable aleatoria es una función

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

tal que para cada $x \in \mathbb{R}$, se tiene que

$$X^{-1}\left(]-\infty, x]\right) = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}.$$

Notación: De aquí en adelante, escribiremos $(X \leq x)$ en vez de $\{\omega \in \Omega : X(\omega) \leq x\}$ y $\mathbb{P}(X \leq x)$ en vez de $\mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\})$.

A continuación presentamos un ejemplo de una variable aleatoria asociada con el espacio de probabilidad dado en el Ejemplo 1.4.

Ejemplo 1.5. Dado el espacio de probabilidad del Ejemplo 1.4, definimos la función $X : \Omega \rightarrow \mathbb{R}$, de la siguiente manera :

$$X(\omega) := \begin{cases} 1, & \omega = C \\ 0, & \omega = S \end{cases}$$

Probemos ahora que X es una variable aleatoria. En efecto, no es difícil verificar que:

$$X^{-1}\left(]-\infty, x]\right) = \begin{cases} \Omega, & x \geq 1 \\ \{S\}, & 0 \leq x < 1 \\ \emptyset, & x < 0 \end{cases} \quad (1.3)$$

En conclusión, $X^{-1}\left(]-\infty, x]\right) \in \mathcal{F}$ para toda $x \in \mathbb{R}$, lo cual prueba que efectivamente X es una variable aleatoria.

Toda variable aleatoria tiene asociada una función llamada **función de distribución**. En lo que sigue se define esta importante función.

Definición 1.18. La función de distribución de una variable aleatoria X es la función $F : \mathbb{R} \rightarrow [0, 1]$, definida como sigue

$$F(x) := \mathbb{P}(X \leq x),$$

para todo $x \in \mathbb{R}$.

Cuando sea necesario especificar la variable aleatoria en cuestión, se escribe $F_X(x)$, pero en general se omite el subíndice X cuando no haya posibilidad de confusión. El argumento de la función es la letra minúscula x que puede tomar cualquier valor real. A esta función se le conoce también con el nombre de función de acumulación de probabilidad, o función de probabilidad acumulada.

Observe que la función de distribución de una variable aleatoria está definida sobre la totalidad del conjunto de números reales, y toma valores en el intervalo $[0, 1]$. La función de distribución es importante pues ella contiene toda la información de la variable aleatoria y la correspondiente medida de probabilidad.

En el siguiente ejemplo presentamos la función de distribución asociada a la variable aleatoria del Ejemplo 1.5.

Ejemplo 1.6. La función de distribución F asociada a la variable aleatoria X del Ejemplo 1.5,

viene dada por $F(x) = \mathbb{P}(X \leq x) = \mathbb{P} \circ X^{-1}(\,] - \infty, x])$, entonces de (1.3) obtenemos

$$F(x) = \begin{cases} \mathbb{P}(\Omega), & x \geq 1 \\ \mathbb{P}(\{S\}), & 0 \leq x < 1 \\ \mathbb{P}(\emptyset), & x < 0 \end{cases}$$

y así,

$$F(x) = \begin{cases} 1, & x \geq 1 \\ 0,5, & 0 \leq x < 1 \\ 0, & x < 0 \end{cases}$$

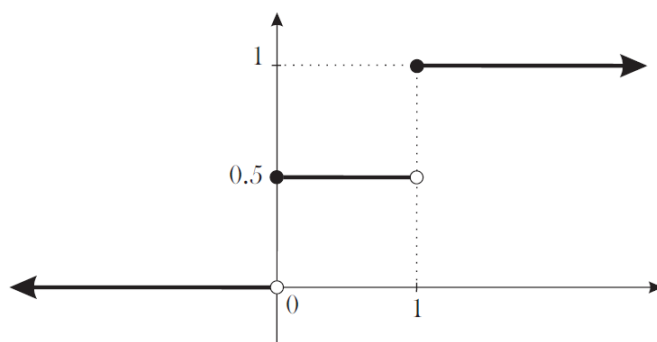


Figura 1.1: Gráfica de la función de distribución F de X . Esta función presenta dos puntos de discontinuidad o saltos en los puntos $x = 0$ y $x = 1$. Note además que el tamaño de estos saltos, no es mas que la probabilidad de que X tome los valores $x = 0$ y $x = 1$

A continuación damos propiedades importantes de la función de distribución F .

Proposición 1.3. *Sea $F(x)$ la función de distribución de una variable aleatoria. Entonces*

1. $\lim_{x \rightarrow +\infty} F(x) = 1$.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$.
3. Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2)$.
4. $F(x)$ es continua por la derecha.

Demostración. 1. Sea $(x_n)_{n \in \mathbb{N}}$ una sucesión cualquiera de números reales creciente que tiende a $+\infty$ y sean los eventos $A_n = (X \leq x_n)$. Entonces $\{A_n\}_{n \in \mathbb{N}}$ es una sucesión de eventos creciente cuya unión es Ω . Luego,

$$\begin{aligned} \lim_{n \rightarrow +\infty} F(x_n) &= \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq x_n) \\ &= \lim_{n \rightarrow +\infty} \mathbb{P}(A_n) \\ &= \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \\ &= \mathbb{P}(\Omega) = 1, \end{aligned}$$

donde en la tercera igualdad se ha usado el hecho que la medida de probabilidad \mathbb{P} es continua superiormente y como \mathbb{R} es un espacio métrico lo anterior implica que $F(x)$ converge a uno cuando x tiende a más infinito.

2. Sea $(x_n)_{n \in \mathbb{N}}$ una sucesión cualquiera de números reales decreciente que tiende a $-\infty$ y sean los eventos $A_n = (X \leq x_n)$. Entonces $\{A_n\}_{n \in \mathbb{N}}$ es una sucesión de eventos decreciente cuya intersección es el conjunto vacío \emptyset . Luego,

$$\begin{aligned} \lim_{n \rightarrow +\infty} F(x_n) &= \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq x_n) \\ &= \lim_{n \rightarrow +\infty} \mathbb{P}(A_n) \\ &= \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right) \\ &= \mathbb{P}(\emptyset) = 0, \end{aligned}$$

donde en la tercera igualdad se ha usado el hecho que la medida de probabilidad \mathbb{P} es continua inferiormente y como \mathbb{R} es un espacio métrico lo anterior implica que $F(x)$ converge a cero cuando x tiende a menos infinito.

3. Si $x_1 \leq x_2$ entonces el evento $(X \leq x_1)$ implica el evento $(X \leq x_2)$, es decir, $(X \leq x_1) \subset (X \leq x_2)$ y así

$$F(x_1) = \mathbb{P}(X \leq x_1) \leq \mathbb{P}(X \leq x_2) = F(x_2),$$

donde en la desigualdad anterior se ha usado el hecho que \mathbb{P} es monótona.

4. Sea $(x_n)_{n \in \mathbb{N}}$ una sucesión cualquiera de números reales no negativos decreciente que tiende a cero. Entonces, si $A_n := (x < X \leq x + x_n)$, tenemos

$$\begin{aligned} F(x + x_n) &= \mathbb{P}(X \leq x + x_n) \\ &= \mathbb{P}\left((X \leq x) \cup (x < X \leq x + x_n)\right) \\ &= \mathbb{P}(X \leq x) + \mathbb{P}(x < X \leq x + x_n) \\ &= F(x) + \mathbb{P}(A_n), \end{aligned} \tag{1.4}$$

donde en la tercera igualdad se debe a que los eventos $(X \leq x)$ y $(x < X \leq x + x_n)$ son disjuntos. Luego como $(A_n)_{n \in \mathbb{N}}$ es una sucesión decreciente cuya intersección es el vacío, se sigue por la continuidad inferior de la medida \mathbb{P} que

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(\emptyset) = 0. \tag{1.5}$$

Finalmente de (1.4) y (1.5) obtenemos que $\lim_{n \rightarrow \infty} F(x + x_n) = F(x)$, es decir, F es continua por la derecha. \square

Definición 1.19. (Variable aleatoria discreta) La variable aleatoria X se llama discreta si su correspondiente función de distribución $F(x)$ es una función constante por tramos.

Sean x_1, x_2, \dots los puntos de discontinuidad de $F(x)$. En cada uno de estos puntos el tamaño de la discontinuidad es $\mathbb{P}(X = x_i) = F(x_i) - F(x_i-) > 0$, donde $F(x_i-) = \lim_{x \rightarrow x_i^-} F(x)$. A la función $f(x)$ que indica estos incrementos se le llama **función de probabilidad** de X , y se define como sigue

$$f(x) = \begin{cases} \mathbb{P}(X = x), & \text{si } x = x_0, x_1, \dots, \\ 0, & \text{otro caso.} \end{cases}$$

En este caso se dice también que la función de distribución es discreta, además la función de probabilidad $f(x)$ siempre existe, y se le llama también función de masa de probabilidad. También se acostumbra usar el término función de densidad, como una analogía con el caso de variables aleatorias absolutamente continuas definidas más adelante.

Definición 1.20. (Variable aleatoria continua) La variable aleatoria X se llama continua si su correspondiente función de distribución $F(x)$ es continua. Cuando existe una función integrable $\varphi \geq 0$ tal que

$$F(x) = \int_{-\infty}^x \varphi(u) du,$$

para cualquier valor de x , entonces se dice que X es “absolutamente continua”. En tal caso a la función φ se le llama “función de densidad” de X .

Comentario 1. Se sigue fácilmente de la definición anterior que

$$\int_{\mathbb{R}} \varphi(u) du = 1,$$

pues,

$$\begin{aligned} 1 = \lim_{x \rightarrow +\infty} F(x) &= \lim_{x \rightarrow +\infty} \int_{-\infty}^x \varphi(u) du \\ &= \int_{-\infty}^{+\infty} \varphi(u) du. \end{aligned}$$

Definición 1.21. (Independencia de variables aleatorias) Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Dos variables aleatorias $X, Y : \Omega \rightarrow \mathbb{R}$ son independientes si

$$\mathbb{P}\left((X \leq x) \cap (Y \leq y)\right) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y),$$

para todo $x, y \in \mathbb{R}$.

Definición 1.22. Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Dos variables aleatorias $X, Y : \Omega \rightarrow \mathbb{R}$ son idénticamente distribuidas si tienen la misma función de probabilidad (si son discretas) ó la misma función de densidad (si son continuas).

Definición 1.23. (Distribución Bernoulli.) Un “ensayo Bernoulli” es un experimento aleatorio con únicamente dos posibles resultados, llamados “éxito” y “fracaso”, y con probabilidades

respectivas p y $1 - p$. Se define la variable aleatoria X como aquella función que lleva el resultado “éxito” al número 1 y el resultado “fracaso” al número 0. Entonces se dice que X tiene una distribución Bernoulli con parámetro $p \in]0, 1[$. Se escribe $X \sim Ber(p)$ y su correspondiente función de probabilidad es

$$f(x) = \begin{cases} 1 - p, & \text{si } x = 0 \\ p, & \text{si } x = 1 \\ 0, & \text{otro caso.} \end{cases}$$

Definición 1.24. (Distribución geométrica.) Suponga que se tiene una sucesión infinita de ensayos independientes Bernoulli. Se define X como el número ensayos que son realizados hasta que el primer éxito E ocurra. Se dice entonces que X tiene una distribución geométrica con parámetro $p \in]0, 1[$. Se escribe $X \sim geo(p)$ y su correspondiente función de probabilidad es

$$f(x) = \begin{cases} p(1 - p)^{x-1}, & \text{si } x = 1, 2, \dots \\ 0, & \text{otro caso.} \end{cases}$$

La justificación de esta fórmula es inmediata porque, si el primer éxito E aparece en la posición x , el resultado total es $\underbrace{FF\dots F}_{x-1}E$ cuya probabilidad (por independencia) es $(1 - p)^{x-1}p$. Es claro de la definición que X toma los valores $1, 2, 3, \dots$

En la siguiente proposición veremos algunas propiedades relacionadas con la distribución geométrica que nos serán útiles más adelante.

Proposición 1.4. Sean $X : \Omega \rightarrow \mathbb{N}$ una variable aleatoria con distribución geométrica, F su función de distribución y $n \in \mathbb{N}$. Entonces

1. $\mathbb{P}(X > n) = (1 - p)^n$.
2. $F(n) = 1 - \mathbb{P}(X > n)$.

Demostración. Probemos 1. En efecto, aprovechando que X toma valores en \mathbb{N} , se sigue fácilmente que para todo $k \in \mathbb{N}$

$$(X > k) = (X \geq k + 1)$$

y que

$$(X \geq k+1) = \bigcup_{j=k+1}^{\infty} (X = j).$$

Por lo tanto

$$\begin{aligned} \mathbb{P}(X > n) &= \mathbb{P}(X \geq n+1) \\ &= \mathbb{P}\left(\bigcup_{x=n+1}^{\infty} (X = x)\right) \\ &= \sum_{x=n+1}^{\infty} \mathbb{P}(X = x) \\ &= \sum_{x=n+1}^{\infty} (1-p)^{x-1} p \\ &= (1-p)^n, \end{aligned}$$

donde la tercera igualdad es por que los eventos $(X = i)$ y $(X = j)$ son disjuntos dos a dos para todo $i \neq j$. Finalmente probemos 2. En efecto, dado que $\Omega = (X > n) \cup (X \leq n)$, se sigue que

$$1 = \mathbb{P}(\Omega) = \mathbb{P}((X > n) \cup (X \leq n)) = \mathbb{P}(X > n) + \mathbb{P}(X \leq n), \quad (1.6)$$

de lo cual, $F(n) = \mathbb{P}(X \leq n) = 1 - \mathbb{P}(X > n)$. Note que la segunda desigualdad de (1.6) se debe nuevamente al hecho de que los eventos $(X > n)$ y $(X \leq n)$ son disjuntos dos a dos. \square

La siguiente distribución de probabilidad es posiblemente la de mayor importancia.

Definición 1.25. (Distribución normal) Se dice que X tiene una distribución normal o Gaussiana con media μ y varianza σ^2 si su función de densidad es

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2},$$

en donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$ son dos parámetros. Se escribe $X \sim N(\mu, \sigma^2)$.

El parámetro μ recibe el nombre de parámetro de ubicación y σ^2 recibe el nombre de parámetro de escala. Estos conceptos serán explicados más adelante en la Definición 1.33

1.5 Muestras, Estadísticas y Estimadores

En lo que sigue, los conceptos de estadística y estimador son introducidos.

Definición 1.26. El conjunto de valores de una característica (observable) asociada a una colección de individuos u objetos de interés es dicho ser una **población**.

Por ejemplo, si estamos estudiando el peso de hombres con edad entre 30 y 60 años, la población es el conjunto de los pesos de todos los hombres del mundo con edades entre 30 y 60 años.

Podemos considerar a una población como una variable aleatoria X definida sobre cierto espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, con distribución F .

Cualquier parte (o subconjunto) de una población es denominada una muestra. De manera más formal tenemos: Dada una variable aleatoria X con función de densidad (caso continuo) o función de probabilidad (caso discreto) $f(x|\theta)$, entonces

Definición 1.27. Una sucesión X_1, \dots, X_n de n variables aleatorias independientes idénticamente distribuidas con función de densidad o función de probabilidad $f(x|\theta)$ recibe el nombre de **muestra** aleatoria de tamaño n de la distribución de X . En este caso tenemos que la función de probabilidad o densidad conjunta de (X_1, \dots, X_n) es

$$f(x_1, \dots, x_n|\theta) = \prod_{i=1}^n f(x_i|\theta) = f(x_1|\theta) \times \dots \times f(x_n|\theta) \quad (1.7)$$

Cabe resaltar que en la definición anterior, X y $X_i, i = 1, \dots, n$ tienen la misma función de probabilidad (o de densidad) y por lo tanto la misma distribución.

La función de densidad (o de probabilidad) conjunta dada en (1.7) es denominada **función de verosimilitud** de θ , correspondiente a la muestra observada $\mathbf{x} = (x_1, \dots, x_n)'$ y será denotada por

$$L(\theta, \mathbf{x}) = \prod_{i=1}^n f(x_i|\theta).$$

Definición 1.28. Cualquier función de la muestra que no depende de parámetros desconocidos es denominada una **estadística**.

En el ejemplo que presentamos a continuación, consideramos varias estadísticas.

Ejemplo 1.7. Sean X_1, \dots, X_n una muestra aleatoria de la variable aleatoria X , con f.d. o f.p. $f(x|\theta)$. Ejemplos de estadísticas son

$$(i) X_{(1)} = \min\{X_1, \dots, X_n\},$$

$$(ii) X_{(n)} = \max\{X_1, \dots, X_n\},$$

$$(iii) \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$(iv) \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

En (i) y (ii), mín y máx denotan, respectivamente el mínimo y el máximo muestral observado. Por otro lado, \bar{X} y $\hat{\sigma}^2$ denotan, respectivamente, la media y varianza muestrales.

Definición 1.29. El conjunto Θ en el cual θ toma valores es denominado **espacio paramétrico**.

Ejemplo 1.8. Sean X_1, \dots, X_n una muestra aleatoria de la variable aleatoria $X \sim N(\mu, \sigma^2)$.

(i) Si $\sigma^2 = 1$ y $\theta = \mu$ es el parámetro desconocido entonces

$$\Theta = \{\mu, -\infty < \mu < \infty\}.$$

(ii) Si $\mu = 0$ y $\theta = \sigma^2$ es el parámetro desconocido entonces

$$\Theta = \{\sigma^2, \sigma^2 > 0\}.$$

(iii) Si μ y σ^2 son desconocidos entonces $\theta = (\mu, \sigma^2)$ y

$$\Theta = \{(\mu, \sigma^2), -\infty < \mu < \infty \text{ y } \sigma^2 > 0\}.$$

Definición 1.30. Cualquier estadística que asuma valores en Θ es un **estimador** para θ .

1.6 El Método de Máxima Verosimilitud

En esta sección vamos a considerar un método de estimación que posibilita la obtención de estimadores en situaciones específicas. El método que consideramos es el método de máxima verosimilitud en el que los estimadores son obtenidos a partir de la maximización de la función de verosimilitud.

Definición 1.31. Sean X_1, \dots, X_n una muestra aleatoria de tamaño n de la variable aleatoria X , con función de densidad (o de probabilidad) $f(x|\theta)$, con $\theta \in \Theta$, donde Θ es el espacio paramétrico. La función de verosimilitud de θ correspondiente a la muestra aleatoria observada es dada por

$$L(\theta, \mathbf{x}) = \prod_{i=1}^n f(x_i|\theta). \quad (1.8)$$

En muchos casos importantes existe un único valor θ que maximiza la función de verosimilitud, el cual llamamos el estimador de máxima verosimilitud (EMV) y el cual es a menudo obtenido por diferenciación.

Definición 1.32. El estimador de máxima verosimilitud de θ es el valor $\hat{\theta} \in \Theta$ que maximiza la función de verosimilitud $L(\theta, \mathbf{x})$.

El logaritmo natural de la función de verosimilitud de θ es denotado por

$$l(\theta, \mathbf{x}) = \ln L(\theta, \mathbf{x}). \quad (1.9)$$

En el caso uniparamétrico donde Θ es un intervalo de la recta y $l(\theta, \mathbf{x})$ es derivable, el estimador de máxima verosimilitud puede ser encontrado como la raíz de la ecuación de verosimilitud

$$l'(\theta, \mathbf{x}) = \frac{\partial l(\theta, \mathbf{x})}{\partial \theta} = 0. \quad (1.10)$$

En algunas situaciones simples, la solución de la ecuación de verosimilitud puede ser obtenida explícitamente. En situaciones más complicadas, la solución de la ecuación (1.10) es obtenida

por procedimientos numéricos. Para concluir que la solución $\hat{\theta}$ de la ecuación (1.10) es un máximo, es necesario verificar que

$$l''(\hat{\theta}, \mathbf{x}) = \frac{\partial^2 \ln L(\theta, \mathbf{x})}{\partial \theta^2} \Big|_{\theta=\hat{\theta}} < 0. \quad (1.11)$$

Ejemplo 1.9. Sean X_1, \dots, X_n una muestra aleatoria de la distribución de la variable aleatoria $X \sim N(\mu, \sigma^2)$ donde solamente uno de los parámetros es conocido. Determine el EMV del otro parámetro (desconocido).

Discusión: Con x_1, \dots, x_n siendo los valores observados de X_1, \dots, X_n , tenemos:

(i) Sea μ el parámetro desconocido. Entonces

$$\begin{aligned} L(\mu, \mathbf{x}) &= \prod_{i=1}^n f(x_i | \mu) \\ &= \prod_{i=1}^n \left\{ \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{1}{2\sigma^2} (x_i - \mu)^2 \right] \right\} \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \times \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right], \end{aligned}$$

luego,

$$\begin{aligned} \ln L(\mu, \mathbf{x}) &= \ln \left\{ \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \times \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right] \right\} \\ &= -n \ln(\sqrt{2\pi}\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2, \end{aligned}$$

de manera que $\frac{\partial}{\partial \mu} \ln L(\mu, \mathbf{x}) = \frac{n(\bar{x} - \mu)}{\sigma^2} = 0$, y de aquí que $\mu = \bar{x}$. Además, $\frac{\partial^2}{\partial \mu^2} \ln L(\mu, \mathbf{x}) = -\frac{n}{\sigma^2} < 0$, para todo μ . Por lo tanto, se sigue que el EMV de μ es $\hat{\mu} = \bar{x}$.

(ii) Sea σ^2 el parámetro desconocido. Entonces

$$\begin{aligned} \ln L(\sigma^2, \mathbf{x}) &= \ln \left\{ \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \times \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right] \right\} \\ &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2, \end{aligned}$$

de manera que $\frac{\partial}{\partial \sigma^2} \ln L(\sigma^2, \mathbf{x}) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0$, de aquí que $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$; sea $\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 = s^2$. Entonces

$$\begin{aligned} \frac{\partial^2}{\partial (\sigma^2)^2} \ln L(\sigma^2, \mathbf{x}) &= \frac{n}{2\sigma^4} - \frac{2n}{2\sigma^6} \times \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \\ &= \frac{n}{2(\sigma^2)^2} - \frac{2n}{2(\sigma^2)^3} s^2, \end{aligned}$$

de manera que $\frac{\partial^2}{\partial (\sigma^2)^2} \ln L(\sigma^2, \mathbf{x})|_{\sigma^2=s^2} = \frac{n}{2(s^2)^2} - \frac{2ns^2}{2(s^2)^3} = -\frac{n}{2s^4} < 0$. Se sigue que el EMV de σ^2 es $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$.

1.7 Parámetros y Familias de Ubicación y de Escala

Una distribución de probabilidad es caracterizada por parámetros de ubicación y de escala. Estos parámetros son típicamente usados en aplicaciones de modelado.

Suponga que Z es una variable aleatoria fija tomando valores en \mathbb{R} . Para $a \in \mathbb{R}$ y $b \in \mathbb{R}_+$, definamos una nueva variable aleatoria X , como

$$X := a + bZ.$$

Cada una de éstas variables aleatorias X , tiene asociada una función de distribución $F_{a,b}$ que depende de dos parámetros a y b .

Definición 1.33. La familia bi-paramétrica $\mathcal{F} = \{F_{a,b}, a \in \mathbb{R}, b \in \mathbb{R}_+\}$ de funciones de distribución asociadas a X es llamada una **familia de ubicación-escala** asociada con la distribución de Z .

Además a es llamado el parámetro de ubicación y b es llamado el parámetro de escala.

En el caso especial en que $b = 1$, la familia \mathcal{F} recibe el nombre de **familia de ubicación**. Y en el caso especial en el que $a = 0$ la familia \mathcal{F} recibe el nombre de **familia de escala**.

A continuación aclararemos lo dicho anteriormente usando funciones de distribución y de densidad y considerando a Z como una variable absolutamente continua. Veamos, sean G y g

las funciones de distribución y de densidad de Z respectivamente, entonces $X = a + bZ$ tiene función de distribución F y de densidad f dadas por :

$$(1) F(x) = G\left(\frac{x-a}{b}\right), x \in \mathbb{R}.$$

$$(2) f(x) = \frac{1}{b}g\left(\frac{x-a}{b}\right), x \in \mathbb{R}$$

En lo que sigue damos la prueba, la cual es sencilla. Para $x \in \mathbb{R}$,

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(a + bZ \leq x) = \mathbb{P}\left(Z \leq \frac{x-a}{b}\right) = G\left(\frac{x-a}{b}\right),$$

lo cual prueba el primer ítem. Ahora probemos el segundo ítem. Primero note que la función de distribución F de X es continua, puesto que $F(x) = G\left(\frac{x-a}{b}\right)$ y G es continua. Por otro lado, para hallar la densidad de f usamos el siguiente hecho (que lo utilizaremos sin dar una prueba, pues se escapa de nuestros objetivos) : que una función de densidad es la derivada de una función de distribución, más precisamente $g = G'$ y $f = F'$, de lo cual se sigue que,

$$f(x) = F'(x) = \frac{1}{b}G'\left(\frac{x-a}{b}\right) = \frac{1}{b}g\left(\frac{x-a}{b}\right),$$

lo cual prueba el ítem (2). Ahora bien, note que en el ítem (2), que si $b = 1$ y $a > 0$ entonces la gráfica de f es obtenida moviendo la gráfica a unidades a la derecha y si $a < 0$ entonces es obtenida moviendo $-a$ unidades a la izquierda, lo cual nos da una idea de cómo influye el parámetro de ubicación a .

Por otro lado, si $a = 0$ y $b > 1$ entonces la gráfica de f es obtenida de la gráfica de g estirándola horizontalmente y comprimiéndola verticalmente por un factor de b y si $0 < b < 1$ sucede lo contrario.

Por ejemplo, si X tiene distribución normal con media 0 y varianza 1, es decir su función de densidad es

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2},$$

entonces la gráfica de $g(x - \mu) := \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-\mu)^2}$, $\mu \in \mathbb{R}$ no es más que una traslación a la derecha o izquierda de f dependiendo del parámetro de ubicación μ . Y la colección

$$\{g(x - \mu), \mu \in \mathbb{R}\} \tag{1.12}$$

forma una familia uniparamétrica de densidades. A esta familia también se le suele llamar **familia de ubicación uniparamétrica** asociada a la variable aleatoria X .

En la siguiente figura vemos la gráfica de $f(x)$ y $g(x - 4)$.

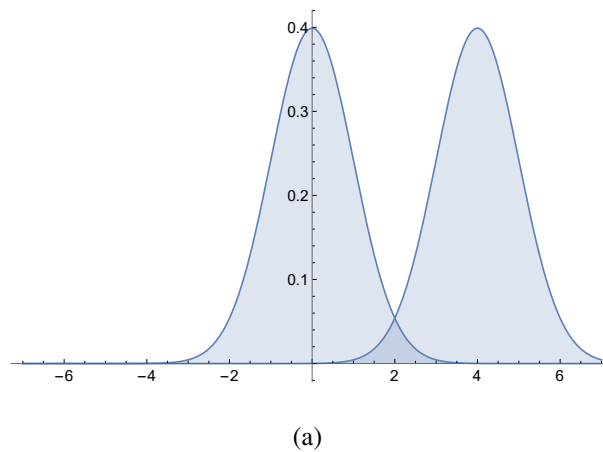


Figura 1.2: La gráfica de la izquierda es la gráfica de $f(x)$, mientras que la de la derecha es la gráfica de $g(x - 4)$, que se ha obtenido moviendo $\mu = 4$ unidades hacia la derecha de f . Tanto $f(x)$ como $g(x - 4)$ son miembros de la familia de ubicación (1.12)

Capítulo 2:

La Ecuación Funcional Aditiva de Cauchy

A lo largo de todo el capítulo se presenta la ecuación funcional aditiva de Cauchy y se halla soluciones lineales bajo ciertas condiciones de regularidad. Con la ayuda de una base de Hamel para \mathbb{R} , se construye una función aditiva discontinua y finalmente se discute de manera breve funciones aditivas sobre el plano complejo.

2.1 Ecuaciones Funcionales y Definición

Una ecuación involucrando una función desconocida y una o más de sus derivadas es llamada una ecuación diferencial. Ejemplos de ecuaciones diferenciales son

$$f'(x) + mx = 15$$

y

$$f''(x) + f'(x) + \sin(x) = 2.$$

Ecuaciones involucrando integrales de una función desconocida son llamadas ecuaciones integrales. Algunos ejemplos de ecuaciones integrales son

$$f(x) = e^x - \int_0^x e^{x-t} f(t) dt$$

y

$$f(x) = \int_0^x [1 - \cos(xt)] f(t) dt.$$

Definición 2.1. Las ecuaciones funcionales son ecuaciones en las cuales las incógnitas son funciones.

Algunas de éstas ecuaciones son

$$\begin{aligned} f(x+y) &= f(x) + f(y) \\ f(xy) &= f(x) + f(y) \\ f(x+y) &= f(x)f(y) \\ f(xy) &= f(x)f(y) \\ f(x+y) + f(x-y) &= 2f(x)f(y), \text{ etc.} \end{aligned}$$

El campo de las ecuaciones funcionales incluye ecuaciones diferenciales, ecuaciones en diferencia e iteraciones y ecuaciones integrales. Las ecuaciones funcionales es un campo de las matemáticas el cual tiene más de 260 años de antigüedad. Más de 5000 papers han sido publicados en esta área. La teoría de las ecuaciones funcionales forma una disciplina matemática moderna la cual ha sido desarrollada muy rápidamente en las últimas 6 décadas.

Resolver una ecuación funcional significa encontrar todas las funciones que satisfacen la ecuación funcional. Con el fin de obtener una solución, las funciones a menudo son restringidas a una naturaleza específica, tal como el ser acotadas, continuas, diferenciables, convexas, monótonas, medibles, holomorfas, etc.

2.2 Solución Lineal de la Ecuación Funcional Aditiva de Cauchy

En esta sección se introduce la ecuación funcional aditiva de Cauchy y se halla su solución bajo ciertas hipótesis. La ecuación funcional

$$f(x+y) = f(x) + f(y), \quad x, y \in \mathbb{R} \tag{2.1}$$

donde es una función $f : \mathbb{R} \rightarrow \mathbb{R}$ recibe el nombre de **ecuación funcional aditiva de Cauchy**.

Definición 2.2. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ se dice ser una función aditiva si satisface la ecuación funcional aditiva de Cauchy (2.1).

Definición 2.3. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es llamada una función lineal si y solo si es de la forma

$$f(x) = cx, \quad \forall x \in \mathbb{R},$$

donde c es una constante real arbitraria.

En el siguiente teorema veremos que toda solución continua de la ecuación funcional aditiva de Cauchy (2.1) es una función lineal.

Teorema 2.1. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ una función continua que satisface la ecuación funcional aditiva de Cauchy (2.1) entonces f es una función lineal, i.e, $f(x) = cx$, donde c es una constante real arbitraria.

Demostración. Primero fijemos $x \in \mathbb{R}$, luego expresamos $f(x)$ como

$$\int_0^1 f(x) dy,$$

pero $f(x) = f(x+y) - f(y)$, entonces

$$f(x) = \int_0^1 [f(x+y) - f(y)] dy$$

$$f(x) = \int_0^1 f(x+y) dy - \int_0^1 f(y) dy.$$

Haciendo el cambio de variable $u = x+y$ se tiene

$$f(x) = \int_x^{1+x} f(u) du - \int_0^1 f(y) dy.$$

Dado que f es continua, usando el Teorema Fundamental del Cálculo, obtenemos

$$f'(x) = f(1+x) - f(x).$$

Como f es aditiva se tiene

$$f'(x) = f(1+x) - f(x) = f(1) + f(x) - f(x) = f(1),$$

y así $f'(x) = c$, donde $c = f(1)$, luego $f(x) = cx + d$, donde d es una constante arbitraria. Como f satisface la ecuación funcional (2.1), entonces de $f(x+y) = f(x) + f(y)$, se tiene

$$c(x+y) + d = cx + d + cy + d,$$

y así $d = 2d$ y por lo cual $d = 0$. Por lo tanto, $f(x) = cx$, $c \in \mathbb{R}$. □

A continuación presentamos una segunda prueba para el teorema anterior, pero antes damos una definición y un teorema que nos facilitarán el trabajo.

Definición 2.4. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ se dice ser racionalmente homogénea si y solo si

$$f(rx) = rf(x) \tag{2.2}$$

para todo $x \in \mathbb{R}$ y todo $r \in \mathbb{Q}$. Además diremos que f es lineal sobre \mathbb{Q} si existe alguna constante $\lambda \in \mathbb{R}$ tal que $f(q) = \lambda q$, $\forall q \in \mathbb{Q}$

El siguiente teorema muestra que cualquier solución de la ecuación funcional aditiva de Cauchy es racionalmente homogénea y lineal sobre \mathbb{Q} .

Teorema 2.2. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una solución de la ecuación funcional aditiva de Cauchy (2.1). Entonces f es racionalmente homogénea. Además f es lineal sobre el conjunto de números racionales \mathbb{Q} .

Demostración. Como f satisface la ecuación funcional aditiva de Cauchy (2.1) entonces si $x = y = 0$, vemos que $f(0) = f(0) + f(0)$ y de aquí

$$f(0) = 0.$$

Ahora, sustituyendo $y = -x$ en (2.1) y del hecho que $f(0) = 0$, vemos que

$$f(-x) = -f(x),$$

para todo $x \in \mathbb{R}$, es decir f es una función impar en \mathbb{R} . Por lo tanto, hasta ahora hemos mostrado que una solución de la ecuación funcional aditiva de Cauchy es cero en el origen y es una función impar.

A continuación mostraremos que una solución de la ecuación funcional aditiva de Cauchy es racionalmente homogénea. Afirmamos que $\forall x \in \mathbb{R}$ y $\forall n \in \mathbb{N}$ se cumple que

$$f(nx) = nf(x).$$

Pues si definimos $S := \{n \in \mathbb{N} \mid f(nx) = nf(x)\}$ se tiene que $1 \in S$ y si $n \in S$ entonces

$$f((n+1)x) = f(nx+x) = f(nx) + f(x) = nf(x) + f(x) = (n+1)f(x),$$

es decir, $n+1 \in S$ y así $n+1 \in S$. Por lo tanto $S = \mathbb{N}$.

Ahora bien, si n es un entero negativo, entonces $-n$ es un número entero positivo, luego como $f(mx) = mx$ para todo $m \in \mathbb{N}$, se sigue que

$$\begin{aligned} f(nx) &= f(-(-nx)) \\ &= -f(-nx) \\ &= -(-n)f(x) \\ &= nf(x). \end{aligned}$$

Por lo tanto, hemos mostrado que $f(nx) = nf(x)$ para todos los enteros n y para todo $x \in \mathbb{R}$.

Ahora sea r un número racional arbitrario. De aquí tenemos

$$r = \frac{k}{l},$$

donde k es un entero y l es un número natural. Además, $kx = l(rx)$. Usando la homogeneidad entera de f , obtenemos

$$kf(x) = f(kx) = f(l(rx)) = lf(rx),$$

luego $\frac{k}{l}f(x) = f(rx)$ y así $rf(x) = f(rx)$. Por lo tanto, f es racionalmente homogénea. Además, si $x = 1$ en la igualdad anterior, obtenemos $rf(1) = f(r)$ y definiendo $c = f(1)$ se tiene que

$$f(r) = cr,$$

para todo los números racionales $r \in \mathbb{Q}$. De aquí que f es lineal sobre el conjunto de los números racionales y la prueba está completa. \square

Ahora presentamos la segunda prueba del Teorema 2.1 usando el hecho de que toda función f que satisface la ecuación (2.1) es lineal sobre \mathbb{Q} .

Demostración. Sea f una solución continua de la ecuación funcional aditiva de Cauchy. Para cualquier número real x existe una sucesión $\{r_n\}$ de números racionales con $r_n \rightarrow x$. Dado que f satisface la ecuación funcional aditiva de Cauchy, por el Teorema 2.2, f es lineal sobre el conjunto de números racionales. Es decir,

$$f(r_n) = cr_n,$$

para todo $n \in \mathbb{N}$. Ahora usando la continuidad de f , obtenemos

$$\begin{aligned} f(x) &= f\left(\lim_{n \rightarrow \infty} r_n\right) \\ &= \lim_{n \rightarrow \infty} f(r_n) \\ &= \lim_{n \rightarrow \infty} cr_n \\ &= c \lim_{n \rightarrow \infty} r_n \\ &= cx \end{aligned}$$

y la prueba está ahora completa. □

Definición 2.5. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ se dice ser localmente integrable si y solo si es integrable sobre todo intervalo acotado.

En lo que sigue demostramos que toda solución localmente integrable de la ecuación funcional aditiva de Cauchy es también lineal. Esta prueba está basada en un argumento dado por Shapiro en [14].

Teorema 2.3. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es localmente integrable y satisface la ecuación funcional aditiva de Cauchy (2.1) entonces f es una función lineal, i.e, $f(x) = cx$, donde $c \in \mathbb{R}$.

Demostración. Asumiendo que f es una solución localmente integrable de la ecuación funcional aditiva de Cauchy, se tiene que $f(x+y) = f(x) + f(y)$ para todo $x, y \in \mathbb{R}$. Usando este hecho y usando la integrabilidad local de f obtenemos

$$\begin{aligned} yf(x) &= \int_0^y f(x) dz \\ &= \int_0^y [f(x+z) - f(z)] dz \\ &= \int_0^y f(x+z) dz - \int_0^y f(z) dz. \end{aligned}$$

Haciendo un cambio de variable $u = x + z$ en la primera integral, tenemos

$$yf(x) = \int_x^{x+y} f(u)du - \int_0^y f(z)dz,$$

lo cual se puede expresar también como

$$\begin{aligned} yf(x) &= \int_x^{x+y} f(u)du - \int_0^y f(u)du \\ &= \int_0^{x+y} f(u)du - \int_0^x f(u)du - \int_0^y f(u)du. \end{aligned}$$

El lado de la derecha de la igualdad anterior es invariante bajo el intercambio de x e y , así que

$$yf(x) = xf(y),$$

para todo $x, y \in \mathbb{R}$. Ahora, fijando y en $\mathbb{R} \setminus \{0\}$ y haciendo variar x en $\mathbb{R} \setminus \{0\}$, tenemos

$$\frac{f(x)}{x} = c,$$

donde $c = \frac{f(y)}{y}$. Esto implica que $f(x) = cx$, para todo $x \in \mathbb{R} \setminus \{0\}$, pero f satisface la ecuación funcional aditiva de Cauchy, entonces

$$f(0+0) = f(0) + f(0),$$

de lo cual se tiene que $f(0) = 0$. Por lo tanto $f(x) = cx$ para todo $x \in \mathbb{R}$. □

El siguiente teorema es debido a Darboux ([8]).

Teorema 2.4. *Sea f una solución de la ecuación funcional aditiva de Cauchy (2.1). Si f es continua en un punto, entonces es continua en todo punto.*

Demostración. Sea f continua en $t \in \mathbb{R}$ y sea $x \in \mathbb{R}$ un punto arbitrario. De aquí tenemos que $\lim_{y \rightarrow t} f(y) = f(t)$, (pues f es continua en t). Luego mostraremos que f es continua en x . En efecto,

$$\begin{aligned} \lim_{y \rightarrow x} f(y) &= \lim_{y \rightarrow x} f(y - x + x - t + t) \\ &= \lim_{y \rightarrow x} [f(y - x + t) + f(x - t)] \\ &= \lim_{y \rightarrow x} f(y - x + t) + \lim_{y \rightarrow x} f(x - t) \\ &= f(t) + f(x - t) \\ &= f(t) + f(x) - f(t) \\ &= f(x). \end{aligned}$$

Esto prueba que f es continua en x y como x es un punto arbitrario de la recta, f es continua en todo punto. Y así la prueba está completa. \square

El siguiente teorema es obvio usando el Teorema 2.1 y el Teorema 2.4.

Teorema 2.5. *Sea f una solución de la ecuación funcional aditiva de Cauchy (2.1). Si f es continua en un punto, entonces f es una función lineal, i.e, $f(x) = cx$, donde $c \in \mathbb{R}$.*

Demostración. Si f es continua en un punto $x_0 \in \mathbb{R}$, entonces por el Teorema 2.4, f es continua en \mathbb{R} , luego por el Teorema 2.1 se tiene que $f(x) = cx$, donde $c \in \mathbb{R}$. \square

2.3 Otros Criterios para la Linealidad

En esta sección presentamos algunas otras condiciones de regularidad que fuerzan a una función que satisface (2.1) ser lineal. Pero primero damos un teorema que nos será de gran ayuda.

Definición 2.6. La gráfica de una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es el conjunto

$$G = \{(x, y) \in \mathbb{R}^2 \mid x \in \mathbb{R}, y = f(x)\}.$$

Es fácil notar que la gráfica G de una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es un subconjunto del plano \mathbb{R}^2 .

Teorema 2.6. *La gráfica de toda solución no lineal $f : \mathbb{R} \rightarrow \mathbb{R}$ de la ecuación (2.1) es un conjunto denso en el plano \mathbb{R}^2 .*

Demostración. Sea f solución de la ecuación funcional aditiva de Cauchy. La gráfica G de f es dada por

$$G = \{(x, y) \in \mathbb{R}^2 \mid x \in \mathbb{R}, y = f(x)\}.$$

Elija un $x_1 \in \mathbb{R} \setminus \{0\}$ fijo. Afirmamos que si f es no lineal entonces existe un número real diferente de cero x_2 tal que

$$\frac{f(x_1)}{x_1} \neq \frac{f(x_2)}{x_2} \quad (2.3)$$

En efecto, pues si $\frac{f(x_1)}{x_1} = \frac{f(x_2)}{x_2}$ para todo $x_2 \in \mathbb{R} \setminus \{0\}$, entonces $f(x_2) = cx_2 \forall x_2 \in \mathbb{R} \setminus \{0\}$, donde $c = \frac{f(x_1)}{x_1}$. Ahora bien, como f es solución de la ecuación funcional aditiva de Cauchy entonces $f(0) = 0$ y así $f(x) = cx$ para todo $x \in \mathbb{R}$. La cual es una contradicción, pues f es no lineal.

Luego (2.3) implica que

$$\begin{vmatrix} x_1 & f(x_1) \\ x_2 & f(x_2) \end{vmatrix} \neq 0,$$

de manera que los vectores $v_1 = (x_1, f(x_1))$ y $v_2 = (x_2, f(x_2))$ son linealmente independientes y por lo tanto generan todo el plano \mathbb{R}^2 . Esto significa que para cualquier vector $v = (x, f(x))$ existen números reales r_1 y r_2 tal que

$$v = r_1 v_1 + r_2 v_2.$$

Definamos

$$\tilde{G} := \{(x, y) \in \mathbb{R}^2 \mid x = \rho_1 x_1 + \rho_2 x_2, y = f(\rho_1 x_1 + \rho_2 x_2), \rho_1, \rho_2 \in \mathbb{Q}\}.$$

Afirmamos que \tilde{G} es denso en \mathbb{R}^2 . Para ver esto, tome $v \in \mathbb{R}^2$, luego $v = r_1 v_1 + r_2 v_2$ donde $r_1, r_2 \in \mathbb{R}$. Luego sean $(\rho_1^n)_{n \in \mathbb{N}}$ y $(\rho_2^n)_{n \in \mathbb{N}}$ sucesiones en \mathbb{Q} tal que $\rho_1^n \rightarrow r_1$ y $\rho_2^n \rightarrow r_2$. Ahora defina

$$v_n := \rho_1^n v_1 + \rho_2^n v_2$$

para todo $n \in \mathbb{N}$, luego

$$\begin{aligned} \lim_{n \rightarrow \infty} v_n &= \lim_{n \rightarrow \infty} \rho_1^n v_1 + \lim_{n \rightarrow \infty} \rho_2^n v_2 \\ &= r_1 v_1 + r_2 v_2 = v \end{aligned}$$

y además $(v_n)_{n \in \mathbb{N}} \subset \tilde{G}$, pues para $n \in \mathbb{N}$

$$\begin{aligned} v_n &= \rho_1^n v_1 + \rho_2^n v_2 \\ &= \rho_1^n (x_1, f(x_1)) + \rho_2^n (x_2, f(x_2)) \\ &= (\rho_1^n x_1 + \rho_2^n x_2, \rho_1^n f(x_1) + \rho_2^n f(x_2)) \\ &= (\rho_1^n x_1 + \rho_2^n x_2, f(\rho_1^n x_1 + \rho_2^n x_2)) \end{aligned}$$

y así $v_n \in \tilde{G}$ para todo $n \in \mathbb{N}$. Por lo tanto para $v \in \mathbb{R}^2$, existe una sucesión $(v_n)_{n \in \mathbb{N}} \subset \tilde{G}$ tal que $\lim_{n \rightarrow \infty} v_n = v$ y así \tilde{G} es denso en \mathbb{R}^2 y como $\tilde{G} \subset G$, concluimos que G es denso en \mathbb{R}^2 . \square

Observación 2.1. En lo que sigue utilizaremos la Definición 2.2, o sea, en vez de decir que f satisface la ecuación funcional aditiva de Cauchy, diremos simplemente que f es aditiva.

Comentario 2. La gráfica de una función continua aditiva es una línea recta que pasa por el origen, mientras que la gráfica de una función aditiva no lineal f es densa en el plano, es decir, todo disco en \mathbb{R}^2 contiene un punto (x, y) tal que $y = f(x)$. También hemos visto que una función aditiva f se convierte en lineal cuando se impone continuidad sobre f . Incluso se puede debilitar esta condición de continuidad a continuidad en un punto y aún conseguir que f sea lineal.

A continuación presentamos dos condiciones más de regularidad que fuerzan a una función aditiva para que sea lineal.

Teorema 2.7. *Si una función aditiva real f es acotada superiormente o inferiormente, entonces f es lineal, i.e., $f(x) = c \cdot x$, donde $c \in \mathbb{R}$.*

Demostración. Suponga que f no es lineal, luego por Teorema 2.6, la gráfica de f es un conjunto denso en el plano. Ahora bien, supongamos sin pérdida de generalidad que f es acotada superiormente, luego existe una constante $M \in \mathbb{R}$ tal que la función aditiva f satisface

$$f(x) \leq M, \quad \forall x \in \mathbb{R}.$$

Luego la gráfica de f no intersecta el conjunto $A = \{(x, y) \in \mathbb{R}^2 \mid y = f(x) > M\}$. Por lo tanto, no puede ser densa en el plano, lo cual es una contradicción. Por lo tanto f es lineal. Cuando f es acotada inferiormente, se procede de manera similar. \square

Proposición 2.1. *Si f es aditiva y acotada sobre \mathbb{R} entonces $f(x) = 0$ para todo $x \in \mathbb{R}$.*

Demostración. Para ver esto, suponga que existe $x_0 \in \mathbb{R}$ tal que $f(x_0) \neq 0$. Dado que f es aditiva, se tiene que $f(nx_0) = nf(x_0)$ para todo $n \in \mathbb{N}$.

Luego afirmamos que f no está acotada, pues de lo contrario existiría una constante $M > 0$ tal que $|f(x)| < M$ para todo $x \in \mathbb{R}$.

Luego $\forall n \in \mathbb{N}$,

$$n|f(x_0)| = |f(nx_0)| < M. \quad (2.4)$$

Ahora tomando $n \in \mathbb{N}$ en (2.4) tal que $n > \frac{M}{|f(x_0)|}$ se llega a una contradicción. Por lo tanto f no está acotada. Y así $f(x) = 0$ para todo $x \in \mathbb{R}$. \square

En el siguiente teorema no asumimos que f está acotada sobre \mathbb{R} , sino asumimos que f está acotada sobre un intervalo cerrado $[a, b]$ no degenerado para $a, b \in \mathbb{R}$.

Teorema 2.8. *Si una función aditiva real f es acotada sobre un intervalo $[a, b]$ no degenerado, entonces existe una constante c tal que $f(x) = cx$ para todo $x \in \mathbb{R}$.*

Demostración. Suponga que $f : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva y acotada sobre un intervalo $[a, b]$. Primero mostramos que $f(x)$ está acotada sobre el intervalo $[0, b - a]$. Dado que $f(x)$ está acotada sobre $[a, b]$, existe un número positivo M tal que

$$|f(y)| < M,$$

para todo $y \in [a, b]$. Si $x \in [0, b - a]$, entonces $x + a \in [a, b]$, de manera que de la siguiente ecuación

$$f(x) = f(x + a) - f(a)$$

se tiene

$$\begin{aligned} |f(x)| &= |f(x + a) - f(a)| \\ &\leq |f(x + a)| + |f(a)| \\ &< M + |f(a)|, \end{aligned}$$

y así $|f(x)| < M + |f(a)|$. Si llamamos $\alpha = b - a$, entonces $f(x)$ está acotada sobre $[0, \alpha]$. Sea $c = \frac{f(\alpha)}{\alpha}$ y defina $\phi(x) := f(x) - cx$. Entonces ϕ satisface:

$$\begin{aligned} \phi(x + y) &= f(x + y) - c(x + y) \\ &= f(x) + f(y) - cx - cy \\ &= f(x) - cx + f(y) - cy \\ &= \phi(x) + \phi(y) \end{aligned}$$

y también tenemos $\phi(\alpha) = f(\alpha) - c\alpha = 0$. Se sigue que $\phi(x)$ es periódica de periodo α , pues

$$\phi(x + \alpha) = \phi(x) + \phi(\alpha) = \phi(x)$$

para todo $x \in \mathbb{R}$. Además, como $\phi(x)$ es la diferencia de funciones acotadas sobre $[0, \alpha]$, la función $\phi(x)$ está acotada sobre $[0, \alpha]$.

Ahora probaremos que $\phi(x)$ está acotada sobre todo \mathbb{R} . Primero probaremos que ϕ está acotada sobre \mathbb{R}_0^+ .

Sea

$$S := \left\{ n \in \mathbb{N} \mid \phi \text{ está acotada en } [(n-1)\alpha, n\alpha] \right\}.$$

Como ϕ está acotada sobre $[0, \alpha]$, entonces $1 \in S$. Ahora suponga que $n \in S$, entonces existe una constante $M > 0$ tal que

$$\forall x \in [(n-1)\alpha, n\alpha], \quad |\phi(x)| \leq M.$$

Ahora veamos si $n+1 \in S$. Para esto, tome $x \in [n\alpha, (n+1)\alpha]$, entonces

$$x - \alpha \in [(n-1)\alpha, n\alpha].$$

Luego por la hipótesis inductiva $|\phi(x - \alpha)| \leq M$. Ahora bien,

$$\begin{aligned} |\phi(x)| &= |\phi(x - \alpha) + \phi(\alpha)| \\ &\leq |\phi(x - \alpha)| + |\phi(\alpha)| \\ &\leq M + 0 \\ &= M. \end{aligned}$$

Por lo tanto ϕ está acotada en $[n\alpha, (n+1)\alpha]$ y así $n+1 \in S$. Como $\mathbb{R}_0^+ = \bigcup_{n=1}^{\infty} [(n-1)\alpha, n\alpha]$, concluimos que ϕ está acotada sobre \mathbb{R}_0^+ . Ahora probaremos que ϕ está acotada en \mathbb{R}_0^- . Sea $x \in \mathbb{R}_0^-$, entonces existe $n \in \mathbb{Z}_0^-$ tal que $x \in [(n-1)\alpha, n\alpha]$, luego $-x \in [-n\alpha, (1-n)\alpha] \subset \mathbb{R}_0^+$. Como ϕ es impar, entonces $|\phi(x)| = |-\phi(-x)| = |\phi(-x)|$. Pero $-x \in [-n\alpha, (1-n)\alpha] \subset \mathbb{R}_0^+$, entonces existe $C > 0$ tal que $|\phi(-x)| \leq C$ y así $|\phi(x)| \leq C$. Por lo tanto ϕ está acotada sobre \mathbb{R}_0^- .

De aquí por la Proposición 2.1, dado que $\phi(x)$ es una función aditiva y acotada sobre \mathbb{R} , se tiene que, $\phi(x) = 0$ para todo $x \in \mathbb{R}$, es decir $f(x) = cx$. Esto completa la prueba del teorema. \square

2.4 Construcción de una Función Aditiva Discontinua

Primero veremos que existe una conexión cercana entre funciones aditivas y bases de Hamel. Para exhibir una función aditiva es suficiente dar sus valores sobre una base de Hamel, estos valores pueden ser asignados arbitrariamente. Este es el contenido de los siguientes dos teoremas.

Teorema 2.9. *Sea B una base de Hamel para \mathbb{R} , considerado como \mathbb{Q} -espacio vectorial. Si dos funciones aditivas tienen el mismo valor en cada elemento de B , entonces ellas son iguales.*

Demostración. Sean f_1 y f_2 dos funciones aditivas tomando el mismo valor en cada miembro de B . Luego $f_1 - f_2$ es aditiva. Escribamos $f = f_1 - f_2$ y sea x cualquier número real. Luego existen números b_1, b_2, \dots, b_n en B y números racionales r_1, r_2, \dots, r_n tal que

$$x = r_1 b_1 + r_2 b_2 + \dots + r_n b_n,$$

de aquí

$$\begin{aligned} f_1(x) - f_2(x) &= f(x) \\ &= f(r_1 b_1 + r_2 b_2 + \dots + r_n b_n) \\ &= f(r_1 b_1) + f(r_2 b_2) + \dots + f(r_n b_n) \quad (\text{pues } f \text{ es aditiva}) \\ &= r_1 f(b_1) + r_2 f(b_2) + \dots + r_n f(b_n) \quad (\text{gracias al Teorema 2.2}) \\ &= r_1 [f_1(b_1) - f_2(b_1)] + r_2 [f_1(b_2) - f_2(b_2)] + r_n [f_1(b_n) - f_2(b_n)]. \\ &= 0 \end{aligned}$$

Por lo tanto, tenemos $f_1 = f_2$ y la prueba está completa. □

Teorema 2.10. *Sea B una base de Hamel para \mathbb{R} , considerado como \mathbb{Q} -espacio vectorial. Sea $g : B \rightarrow \mathbb{R}$ una función arbitraria definida sobre B . Entonces existe una función aditiva $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $f(b) = g(b)$ para cada $b \in B$.*

Demostración. Para cada número real x , existen $b_1, b_2, \dots, b_n \in B$ y existen números racionales r_1, r_2, \dots, r_n tal que

$$x = r_1 b_1 + r_2 b_2 + \dots + r_n b_n,$$

pues B es una base de Hamel para \mathbb{R} .

Ahora definamos $f : \mathbb{R} \rightarrow \mathbb{R}$ de la siguiente forma:

$$f(x) := r_1 g(b_1) + r_2 g(b_2) + \dots + r_n g(b_n),$$

esto define $f(x)$ para todo x . Esta definición no es ambigua, puesto que, para cada x , la elección de $b_1, b_2, \dots, b_n, r_1, r_2, \dots, r_n$ es única, excepto por el orden en el cual b_i y r_i son seleccionados. Para cada b en B , tenemos $f(b) = g(b)$, por definición de f . A continuación mostraremos que f es aditiva sobre los reales. Sean x e y dos números reales arbitrarios. Entonces

$$x = r_1 a_1 + r_2 a_2 + \dots + r_n a_n,$$

$$y = s_1 b_1 + s_2 b_2 + \dots + s_n b_n$$

donde $r_1, r_2, \dots, r_n, s_1, s_2, \dots, s_n$ son números racionales y $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ son miembros de la base de Hamel. Los dos conjuntos $\{a_1, a_2, \dots, a_n\}$ y $\{b_1, b_2, \dots, b_m\}$ puede tener algunos elementos en común, así que sea $\{c_1, c_2, \dots, c_l\}$ la unión de estos dos conjuntos, donde $l \leq m + n$, luego

$$x = u_1 c_1 + u_2 c_2 + \dots + u_n c_l,$$

$$y = v_1 c_1 + v_2 c_2 + \dots + v_n c_l$$

donde $\{u_1, u_2, \dots, u_l\}, \{v_1, v_2, \dots, v_l\}$ son números racionales y varios de los cuales pueden ser cero. Luego

$$x + y = (u_1 + v_1)c_1 + (u_2 + v_2)c_2 + \dots + (u_l + v_l)c_l$$

y

$$\begin{aligned} f(x+y) &= f\left((u_1 + v_1)c_1 + (u_2 + v_2)c_2 + \dots + (u_l + v_l)c_l\right) \\ &= (u_1 + v_1)g(c_1) + (u_2 + v_2)g(c_2) + \dots + (u_l + v_l)g(c_l) \\ &= [u_1 g(c_1) + u_2 g(c_2) + \dots + u_l g(c_l)] + [v_1 g(c_1) + v_2 g(c_2) + \dots + v_l g(c_l)] \\ &= f(x) + f(y). \end{aligned}$$

Por lo tanto f es aditiva sobre el conjunto de los números reales \mathbb{R} y la prueba del teorema está ahora completa. □

En lo que sigue, con la ayuda de una base de Hamel, se construirá una **función aditiva discontinua**. Sea B una base de Hamel para el conjunto de números reales \mathbb{R} . Sea $b \in B$ cualquier elemento no nulo de B . Defina $g : B \rightarrow \mathbb{R}$ de la siguiente manera:

$$g(x) = \begin{cases} 0 & \text{si } x \in B \setminus \{b\} \\ 1 & \text{si } x = b. \end{cases}$$

Por el teorema anterior, existe una función aditiva $f : \mathbb{R} \rightarrow \mathbb{R}$ con $f(x) = g(x)$ para cada $x \in B$. Afirmamos que f no puede ser lineal. Pues si f es lineal entonces

$$f(x) = kx \quad \forall x \in \mathbb{R}, \quad (2.5)$$

donde $k \in \mathbb{R}$ es una constante. Luego

$$\frac{f(x)}{x} = k, \quad \forall x \neq 0.$$

Ahora tome $x \in B \setminus \{0, b\}$, entonces

$$k = \frac{f(x)}{x} = \frac{g(x)}{x} = \frac{0}{x} = 0, \quad (2.6)$$

Y así de (2.6) en (2.5) se tiene que $f(x) = 0$. Luego $0 = f(b) = g(b) = 1$, lo cual es una contradicción. Por lo tanto f es no lineal. Luego por el Teorema 2.1 concluimos que f es **discontinua**.

Ningún ejemplo concreto de una base de Hamel para \mathbb{R} es conocida; solamente sabemos que existe. Por otro lado, la gráfica de una función aditiva discontinua sobre \mathbb{R} no es fácil de dibujar, debido a que es un subconjunto denso del plano \mathbb{R}^2 .

2.5 Funciones Aditivas sobre el Plano Complejo

En esta sección, se presenta algunos resultados concernientes a funciones aditivas de variable compleja.

Definición 2.7. El sistema de números complejos \mathbb{C} es el conjunto de pares ordenados de números reales (x, y) con las operaciones de adición y multiplicación definidas por

$$\begin{aligned}(x, y) + (u, v) &= (x + u, y + v) \\ (x, y)(u, v) &= (xu - yv, xv + yu)\end{aligned}$$

para todo $x, y, u, v \in \mathbb{R}$.

Considerando a un número real como x o $(x, 0)$ y denotando como i al número puramente imaginario $(0, 1)$, podemos escribir la siguiente expresión

$$(x, y) = (x, 0) + (0, 1)(y, 0)$$

como

$$(x, y) = x + iy.$$

Si denotamos el lado izquierdo de la representación anterior por z entonces tenemos $z = x + iy$. El número real x es llamado la parte real de z y es denotado por $\operatorname{Re} z$. Similarmente, el número real y es llamado la parte imaginaria de z y es denotada por $\operatorname{Im} z$. Si z es un número complejo de la forma $x + iy$, entonces el número complejo $x - iy$ es llamado el conjugado de z y es denotado por \bar{z} .

Una función arbitraria $f : \mathbb{C} \rightarrow \mathbb{C}$ puede ser escrita como

$$f(z) = f_1(z) + if_2(z), \quad (2.7)$$

donde $f_1 : \mathbb{C} \rightarrow \mathbb{R}$ y $f_2 : \mathbb{C} \rightarrow \mathbb{R}$ son dadas por

$$f_1(z) = \operatorname{Re} f(z) \text{ y } f_2(z) = \operatorname{Im} f(z). \quad (2.8)$$

Si f es aditiva, entonces por (2.7) y (2.8) tenemos

$$\begin{aligned}f_1(z_1 + z_2) &= \operatorname{Re} f(z_1 + z_2) \\ &= \operatorname{Re}[f(z_1) + f(z_2)] \\ &= \operatorname{Re} f(z_1) + \operatorname{Re} f(z_2) \\ &= f_1(z_1) + f_1(z_2)\end{aligned}$$

y

$$\begin{aligned}f_2(z_1 + z_2) &= \operatorname{Im} f(z_1 + z_2) \\ &= \operatorname{Im}[f(z_1) + f(z_2)] \\ &= \operatorname{Im} f(z_1) + \operatorname{Im} f(z_2) \\ &= f_2(z_1) + f_2(z_2)\end{aligned}$$

En resumen si $f : \mathbb{C} \rightarrow \mathbb{C}$ es aditiva, entonces $\operatorname{Re} f$, $\operatorname{Im} f$ son también aditivas.

Ahora veremos dos teoremas relacionados con la ecuación funcional aditiva de Cauchy en dos variables, los cuales nos serán útiles para continuar con nuestro estudio de funciones aditivas $f : \mathbb{C} \rightarrow \mathbb{C}$. Considere $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ que satisface la siguiente ecuación

$$f(x_1 + y_1, x_2 + y_2) = f(x_1, x_2) + f(y_1, y_2) \quad (2.9)$$

para todo $(x_1, x_2) \in \mathbb{R}^2$, $(y_1, y_2) \in \mathbb{R}^2$.

Teorema 2.11. *La solución general $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de la ecuación funcional (2.9) es dada por :*

$$f(x_1, x_2) = A_1(x_1) + A_2(x_2)$$

donde $A_1, A_2 : \mathbb{R} \rightarrow \mathbb{R}$ son aditivas.

Demostración. Sea $x_2 = y_2 = 0$ en (2.9), luego

$$f(x_1 + y_1, 0) = f(x_1, 0) + f(y_1, 0).$$

Luego defina $A_1 : \mathbb{R} \rightarrow \mathbb{R}$ por $A_1(x) = f(x, 0)$. Luego de la ecuación anterior se tiene

$$A_1(x_1 + y_1) = A_1(x_1) + A_1(y_1).$$

De aquí que $A_1 : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva. Similarmente, haciendo $x_1 = y_1 = 0$ en (2.9), obtenemos

$$f(0, x_2 + y_2) = f(0, x_2) + f(0, y_2).$$

Definiendo $A_2 : \mathbb{R} \rightarrow \mathbb{R}$ por $A_2(x) = f(0, x)$, obtenemos al reemplazar en la ecuación anterior

$$A_2(x_2 + y_2) = A_2(x_2) + A_2(y_2).$$

Por lo tanto $A_2 : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva. Luego sustituimos en (2.9)

$$y_1 = 0 = x_2,$$

para obtener

$$\begin{aligned} f(x_1, y_2) &= f(x_1, 0) + f(0, y_2) \\ &= A_1(x_1) + A_2(y_2). \end{aligned}$$

Por lo tanto

$$f(x, y) = A_1(x) + A_2(y),$$

para $x, y \in \mathbb{R}$ donde $A_1, A_2 : \mathbb{R} \rightarrow \mathbb{R}$ son funciones aditivas sobre \mathbb{R} . \square

Este teorema dice que cualquier función aditiva de dos variables sobre \mathbb{R}^2 puede ser descompuesta como la suma de dos funciones aditivas en una variable. Es decir

$$f(x, y) = A_1(x) + A_2(y),$$

donde $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ y $A_1, A_2 : \mathbb{R} \rightarrow \mathbb{R}$.

El teorema anterior puede también ser escrito como el siguiente teorema.

Teorema 2.12. *Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es aditiva sobre el plano \mathbb{R}^2 , entonces existen funciones aditivas $A_1, A_2 : \mathbb{R} \rightarrow \mathbb{R}$ tal que*

$$f(x_1, x_2) = A_1(x_1) + A_2(x_2)$$

para todo $x_1, x_2 \in \mathbb{R}$.

El siguiente teorema se sigue del anterior y del Teorema 2.1.

Teorema 2.13. *Si $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es una función aditiva continua sobre el plano \mathbb{R}^2 , entonces existen constantes c_1, c_2 reales tal que*

$$f(x_1, x_2) = c_1 x_1 + c_2 x_2,$$

para todo $x_1, x_2 \in \mathbb{R}$.

Ahora regresemos al plano complejo \mathbb{C} . A continuación un teorema referente a funciones aditivas $f : \mathbb{C} \rightarrow \mathbb{C}$.

Teorema 2.14. *Si $f : \mathbb{C} \rightarrow \mathbb{C}$ es aditiva, entonces existen funciones aditivas $f_{kj} : \mathbb{R} \rightarrow \mathbb{R}$ ($k, j = 1, 2$) tal que*

$$f(z) = f_{11}(\operatorname{Re} z) + f_{12}(\operatorname{Im} z) + i f_{21}(\operatorname{Re} z) + i f_{22}(\operatorname{Im} z).$$

Demostración. Para $z \in \mathbb{C}$, $f(z) = f_1(z) + i f_2(z)$ donde $f_1 : \mathbb{C} \rightarrow \mathbb{R}$ y $f_2 : \mathbb{C} \rightarrow \mathbb{R}$. Dado que f es una función aditiva, f_1 y f_2 también son aditivas. Y debido a que f_1 y f_2 pueden ser consideradas como funciones de \mathbb{R}^2 en \mathbb{R} , entonces aplicando el Teorema 2.12, tenemos que

$$f_1(x, y) = A_{11}(x) + A_{12}(y),$$

$$f_2(x, y) = A_{21}(x) + A_{22}(y),$$

donde $A_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ son funciones aditivas para $i, j = 1, 2$. Luego considerando $z = (x, y) = x + iy$ tenemos

$$\begin{aligned} f(z) &= f_1(z) + if_2(z) \\ &= A_{11}(\operatorname{Re} z) + A_{12}(\operatorname{Im} z) + iA_{21}(\operatorname{Re} z) + iA_{22}(\operatorname{Im} z). \end{aligned}$$

□

El siguiente teorema trata sobre la forma de funciones aditivas continuas sobre el plano complejo.

Teorema 2.15. Si $f : \mathbb{C} \rightarrow \mathbb{C}$ es una función aditiva continua, entonces existen constantes complejas c_1 y c_2 tal que

$$f(z) = c_1 z + c_2 \bar{z}.$$

Demostración. Como f es aditiva, por el teorema anterior, obtenemos

$$f(z) = f_{11}(\operatorname{Re} z) + f_{12}(\operatorname{Im} z) + if_{21}(\operatorname{Re} z) + if_{22}(\operatorname{Im} z)$$

donde $f_{kj} : \mathbb{R} \rightarrow \mathbb{R}$ ($k, j = 1, 2$) son funciones aditivas. La continuidad de f implica la continuidad de cada función f_{kj} y de aquí

$$f_{kj}(x) = c_{kj}x,$$

donde $c_{k,j}$ ($k, j = 1, 2$) son constantes reales. Por lo tanto, usando la forma de $f(z)$ y la forma de f_{kj} , obtenemos

$$\begin{aligned} f(z) &= c_{11}\operatorname{Re} z + c_{12}\operatorname{Im} z + ic_{21}\operatorname{Re} z + ic_{22}\operatorname{Im} z \\ &= (c_{11} + ic_{21})\operatorname{Re} z + (c_{12} + ic_{22})\operatorname{Im} z \\ &= a\operatorname{Re} z + b\operatorname{Im} z, \text{ donde } a = c_{11} + ic_{21}, \quad b = c_{12} + ic_{22} \\ &= a\operatorname{Re} z - i(bi)\operatorname{Im} z \\ &= \frac{a+bi}{2}\operatorname{Re} z + \frac{a-bi}{2}\operatorname{Re} z - \frac{a+bi}{2}i\operatorname{Im} z + \frac{a-bi}{2}i\operatorname{Im} z \\ &= \frac{a-bi}{2}\operatorname{Re} z + \frac{a-bi}{2}i\operatorname{Im} z + \frac{a+bi}{2}\operatorname{Re} z - \frac{a+bi}{2}i\operatorname{Im} z \\ &= \frac{a-bi}{2}(\operatorname{Re} z + i\operatorname{Im} z) + \frac{a+bi}{2}(\operatorname{Re} z - i\operatorname{Im} z) \\ &= \frac{a-bi}{2}z + \frac{a+bi}{2}\bar{z}, \end{aligned}$$

donde $c_1 = \frac{a-bi}{2}$ y $c_2 = \frac{a+bi}{2}$ son constantes complejas. Esto completa la prueba del teorema. \square

Note que a diferencia de las funciones aditivas continuas $f : \mathbb{R} \rightarrow \mathbb{R}$, las funciones aditivas continuas $f : \mathbb{C} \rightarrow \mathbb{C}$ no son lineales.

La linealidad puede ser restaurada si uno asume una condición mas fuerte tal como holomorfía en vez de continuidad.

Definición 2.8. Una función $f : \mathbb{C} \rightarrow \mathbb{C}$ se dice ser holomorfa en \mathbb{C} si para todo $z_0 \in \mathbb{C}$, existe

$$\lim_{h \rightarrow 0} \frac{f(z_0 + h) - f(z_0)}{h}, \quad (h \in \mathbb{C})$$

Teorema 2.16. Si $f : \mathbb{C} \rightarrow \mathbb{C}$ es una función aditiva holomorfa, entonces existe una constante $c \in \mathbb{C}$ tal que

$$f(z) = cz,$$

es decir, f es lineal.

Demostración. Dado que f es aditiva, entonces para $z_1, z_2 \in \mathbb{C}$, tenemos

$$f(z_1 + z_2) = f(z_1) + f(z_2)$$

Luego como f es holomorfa, podemos derivar la expresión anterior con respecto a z_1 y así obtenemos

$$f'(z_1 + z_2) = f'(z_1),$$

para todo $z_1, z_2 \in \mathbb{C}$. Luego, haciendo $z_1 = 0$ y $z_2 = z$ obtenemos

$$f'(z) = f'(0).$$

De lo anterior vemos que $f(z) = cz + b$, donde $c = f'(0)$ y $b \in \mathbb{C}$. Pero como f es aditiva, entonces de

$$f(z_1 + z_2) = f(z_1) + f(z_2),$$

se obtiene

$$c(z_1 + z_2) + b = cz_1 + b + cz_2 + b,$$

de lo cual $b = 0$. Por lo tanto $f(z) = cz$, $c \in \mathbb{C}$, lo cual completa la prueba del teorema. \square

Capítulo 3:

Otras Ecuaciones Funcionales de Cauchy

En el capítulo 5 de su libro Cours d'Analyse ([7]) , A.L. Cauchy (1821) también estudió otras tres ecuaciones funcionales, a saber,

$$f(x+y) = f(x)f(y) \tag{3.1}$$

$$f(xy) = f(x) + f(y) \tag{3.2}$$

$$f(xy) = f(x)f(y). \tag{3.3}$$

para todo $x, y \in \mathbb{R}$. Este capítulo está dedicado a resolver estas tres ecuaciones funcionales. La solución general de cada una de estas ecuaciones funcionales está determinada en términos de una función aditiva.

3.1 Solución de la Ecuación Exponencial de Cauchy

En esta sección se determina la solución general de la ecuación funcional exponencial de Cauchy (3.1) sin asumir alguna condición de regularidad tal como continuidad, acotación o diferenciabilidad de la función desconocida f .

Teorema 3.1. *Si la ecuación funcional (3.1), es decir,*

$$f(x+y) = f(x)f(y)$$

se cumple para todo número real x e y , entonces la solución general de (3.1) es dado por

$$f(x) = e^{A(x)} \text{ y } f(x) = 0, \forall x \in \mathbb{R}$$

donde $A : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva y e es la base Neperiana del logaritmo.

Demostración. Es fácil ver que $f(x) = 0$ para todo $x \in \mathbb{R}$ es una solución de (3.1). De aquí en adelante suponemos que $f(x)$ no es idénticamente cero, i.e., $\exists x \in \mathbb{R}$ tal que $f(x) \neq 0$ y bajo esta suposición podemos probar que $f(x)$ nunca toma el valor cero.

Para esto, supongamos lo contrario. Luego existiría un y_0 tal que $f(y_0) = 0$. De (3.1) obtenemos

$$\begin{aligned} f(y) &= f((y - y_0) + y_0) \\ &= f(y - y_0)f(y_0) \end{aligned}$$

para todo $y \in \mathbb{R}$, luego $f(x)$ es idénticamente cero, lo cual es una contradicción a nuestra suposición que $f(x)$ no es idénticamente cero. De aquí que $f(x)$ nunca es cero.

Ahora sea $x = \frac{t}{2} = y$ en (3.1), luego vemos que

$$f(t) = f\left(\frac{t}{2}\right)^2$$

para todo $t \in \mathbb{R}$. De aquí que $f(x)$ es estrictamente positivo. Ahora tomando logaritmo natural en ambos lados de (3.1), obtenemos

$$\ln f(x+y) = \ln f(x) + \ln f(y).$$

Definiendo $A : \mathbb{R} \rightarrow \mathbb{R}$ por $A(x) = \ln f(x)$, se tiene

$$A(x+y) = A(x) + A(y),$$

es decir $A(x)$ es aditiva. Luego despejando, se obtiene que $f(x) = e^{A(x)}$ y la prueba está completa. \square

El siguiente corolario es obvio del teorema anterior.

Corolario 3.1. Si la ecuación funcional (3.1), es decir, $f(x+y) = f(x)f(y)$, se cumple para todo número real x e y , entonces la solución general continua de (3.1) es dada por

$$f(x) = e^{cx} \text{ y } f(x) = 0, \forall x \in \mathbb{R}$$

donde c es una constante real arbitraria.

Demostración. En el Teorema 3.1 se probó que la solución general de (3.1) es dada por $f(x) = e^{A(x)}$ y $f(x) = 0, \forall x \in \mathbb{R}$ donde $A : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva.

Pero si $A(x)$ es además continua, entonces por teorema 2.2.1 $A(x)$ es lineal, es decir, $A(x) = cx$ donde c es una constante arbitraria. Por lo tanto la solución general continua de (3.1) es dada por

$$f(x) = e^{cx} \text{ y } f(x) = 0, \forall x \in \mathbb{R}.$$

□

Definición 3.1. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es llamada una función exponencial real si satisface $f(x+y) = f(x)f(y)$ para todo $x, y \in \mathbb{R}$.

3.2 Solución de la Ecuación Logarítmica de Cauchy

Ahora nosotros consideramos la segunda ecuación funcional de Cauchy (3.2). Esta ecuación es conocida como la ecuación logarítmica de Cauchy.

Teorema 3.2. Si la ecuación funcional (3.2),

$$f(xy) = f(x) + f(y)$$

se cumple para todo $x, y \in \mathbb{R} \setminus \{0\}$, entonces la solución general de (3.2) es dada por

$$f(x) = A(\ln |x|), \quad \forall x \in \mathbb{R} \setminus \{0\}$$

donde A es una función aditiva.

Demostración. Primero sustituimos $x = t$ e $y = t$ en (3.2) para obtener

$$f(t^2) = 2f(t).$$

Similarmente, haciendo $x = -t$ e $y = -t$ en (3.2), tenemos

$$f(t^2) = 2f(-t).$$

Luego vemos que

$$f(t) = f(-t), \quad \forall t \in \mathbb{R} \setminus \{0\}. \quad (3.4)$$

Ahora, suponga que la ecuación funcional (3.2) se cumple para $x > 0$ e $y > 0$. Sea

$$x = e^s \text{ e } y = e^t \quad (3.5)$$

de manera que

$$s = \ln x \text{ y } t = \ln y. \quad (3.6)$$

Note que $s, t \in \mathbb{R}$, dado que $x, y \in \mathbb{R}^+$. Sustituyendo (3.5) en (3.2), obtenemos

$$f(e^{s+t}) = f(e^s) + f(e^t).$$

Definiendo

$$A(s) = f(e^s) \quad (3.7)$$

y reemplazando en la penúltima ecuación tenemos

$$A(s+t) = A(s) + A(t)$$

para todo $s, t \in \mathbb{R}$. Luego utilizando la primera ecuación de (3.6) en (3.7) se tiene

$$f(x) = A(\ln x), \quad \forall x \in \mathbb{R}^+, \quad (3.8)$$

donde A es una función aditiva y dado que $f(t) = f(-t), \forall t \in \mathbb{R} \setminus \{0\}$, vemos que la solución general de (3.2) es $f(x) = A(\ln |x|), \forall x \in \mathbb{R} \setminus \{0\}$ y la prueba ahora está completa. \square

Corolario 3.2. *La solución general de la ecuación funcional (3.2), es decir, $f(xy) = f(x) + f(y)$, cumpliéndose para todo $x, y \in \mathbb{R}^+$ es dada por*

$$f(x) = A(\ln x), \quad (3.9)$$

donde $A : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva.

Corolario 3.3. *La solución general de la ecuación funcional (3.2) cumpliéndose para todo $x, y \in \mathbb{R}$ es dada por*

$$f(x) = 0, \quad \forall x \in \mathbb{R} \quad (3.10)$$

Demostración. Sea $x \in \mathbb{R}$, luego sustituyendo $y = 0$ en (3.2) obtenemos $f(0) = f(x) + f(0)$ y de aquí tenemos $f(x) = 0$. \square

Corolario 3.4. *La solución general continua de la ecuación funcional (3.2), es decir, $f(xy) = f(x) + f(y)$, cumpliéndose para todo $x, y \in \mathbb{R} \setminus \{0\}$ es dado por*

$$f(x) = c \ln |x|, \quad \forall x \in \mathbb{R} \setminus \{0\} \quad (3.11)$$

donde c es una constante real arbitraria.

Demostración. Sabemos por Teorema 3.2 que la solución general de (3.2) es dada por $f(x) = A(\ln |x|)$, $\forall x \in \mathbb{R} \setminus \{0\}$ donde A es una función aditiva. Ahora bien, si $f(x)$ es continua entonces $A(x)$ es continua, luego como $A(x)$ es aditiva entonces $A(x)$ es lineal, es decir, $A(x) = cx$ donde $c \in \mathbb{R}$ es una constante arbitraria. Así $f(x) = A(\ln |x|) = c \ln |x|$, $\forall x \in \mathbb{R} \setminus \{0\}$. \square

Definición 3.2. Una función $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ es llamada una función logarítmica si satisface $f(xy) = f(x) + f(y)$ para todo $x, y \in \mathbb{R}^+$.

3.3 Solución de la Ecuación Multiplicativa de Cauchy

Ahora tratamos la última ecuación de Cauchy (3.3). Esta ecuación es la más complicada de las tres ecuaciones consideradas en este capítulo. En el siguiente teorema necesitamos la noción de la función signo. La función signo es denotada por $\text{sgn}(x)$ y definida como

$$\text{sgn}(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -1 & \text{si } x < 0 \end{cases} \quad (3.12)$$

Teorema 3.3. *La solución general de la ecuación funcional multiplicativa (3.3), es decir,*

$$f(xy) = f(x)f(y)$$

cumpliendo para todo $x, y \in \mathbb{R}$ es dada por

$$f(x) = 0 \quad (3.13)$$

$$f(x) = 1 \quad (3.14)$$

$$f(x) = e^{A(\ln|x|)} |sgn(x)| \quad (3.15)$$

y

$$f(x) = e^{A(\ln|x|)} sgn(x) \quad (3.16)$$

donde $A : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva y e es la base Napieriana del logaritmo.

Demostración. Sea $x = y = 0$ en (3.3) para obtener $f(0) = f(0)f(0)$, luego $f(0)[1 - f(0)] = 0$ y de aquí

$$f(0) = 0 \quad \text{o} \quad f(0) = 1. \quad (3.17)$$

Similarmente, sustituyendo $x = 1 = y$ en (3.3), tenemos $f(1) = f(1)f(1)$, luego $f(1)[1 - f(1)] = 0$ y de aquí

$$f(1) = 0 \quad \text{o} \quad f(1) = 1. \quad (3.18)$$

Sea x un número real positivo, es decir $x > 0$. Luego en (3.3)

$$f(x) = f(\sqrt{x} \cdot \sqrt{x}) = f(\sqrt{x})f(\sqrt{x}) = f(\sqrt{x})^2 \geq 0. \quad (3.19)$$

Suponga que existe algún $x_0 \in \mathbb{R}, x_0 \neq 0$ tal que $f(x_0) = 0$. Sea $x \in \mathbb{R}$ un número real arbitrario. Luego de (3.3) tenemos $f(x) = f(x_0 \cdot \frac{x}{x_0}) = f(x_0)f(\frac{x}{x_0}) = 0$ para todo $x \in \mathbb{R}$ y obtenemos la solución (3.13).

De ahora en adelante suponemos que $f(x) \neq 0$ para todo $x \in \mathbb{R} \setminus \{0\}$. De (3.17) tenemos $f(0) = 0$ o $f(0) = 1$. Si $f(0) = 1$, entonces haciendo $y = 0$ en (3.3), obtenemos

$$f(0) = f(x)f(0)$$

y de aquí $f(x) = 1$ para todo $x \in \mathbb{R}$. Por lo tanto tenemos la solución deseada (3.14).

Luego consideramos el caso $f(0) = 0$. En este caso afirmamos que f nunca es cero en $\mathbb{R} \setminus \{0\}$.

Suponga que no. Entonces existe un $y_0 \in \mathbb{R} \setminus \{0\}$ tal que $f(y_0) = 0$. Sea $y = y_0$ en (3.3), entonces tenemos

$$f(xy_0) = f(x)f(y_0) = 0.$$

De aquí $f(x) = 0, \forall x \in \mathbb{R} \setminus \{0\}$, lo cual es una contradicción a nuestra suposición que f no es idénticamente cero en $\mathbb{R} \setminus \{0\}$. Por lo tanto f nunca es cero en $\mathbb{R} \setminus \{0\}$. Del hecho que f es nunca cero en $\mathbb{R} \setminus \{0\}$ y de (3.19), tenemos

$$f(x) > 0 \quad \text{para } x > 0. \quad (3.20)$$

Sea

$$x = e^s \quad \text{e} \quad y = e^t \quad (3.21)$$

de manera que

$$s = \ln x \quad \text{y} \quad t = \ln y. \quad (3.22)$$

Note que $s, t \in \mathbb{R}$ dado que $x, y \in \mathbb{R}_+$. Sustituyendo (3.21) en (3.3), obtenemos

$$f(e^{s+t}) = f(e^s)f(e^t).$$

Dado que $f(t) > 0$ para todo $t > 0$, tomando el logaritmo natural en ambos lados de la última ecuación, tenemos $\ln f(e^{s+t}) = \ln f(e^s) + \ln f(e^t)$, luego definiendo

$$A(s) = \ln f(e^s), \quad \forall s \in \mathbb{R} \quad (3.23)$$

tenemos $A(s+t) = A(s) + A(t)$. Por lo tanto A es una función aditiva. De (3.23) y (3.22), obtenemos

$$f(x) = e^{A(\ln|x|)}, \quad \forall x \in \mathbb{R}^+. \quad (3.24)$$

De (3.18) vemos que $f(1) = 0$ o $f(1) = 1$. Si $f(1) = 0$, haciendo $y = 1$ en (3.3), obtenemos $f(x) = f(x \cdot 1) = f(x)f(1) = 0$, luego

$$f(x) = 0, \quad \forall x \in \mathbb{R} \setminus \{0\},$$

lo cual es contrario a nuestra suposición de que f no es idénticamente cero sobre $\mathbb{R} \setminus \{0\}$.

De aquí $f(1) = 1$. Ahora haciendo $x = -1 = y$ en (3.3), obtenemos $f(1) = f(-1)^2$ y de aquí

$$f(-1) = 1 \quad \text{o} \quad f(-1) = -1. \quad (3.25)$$

Si $f(-1) = 1$, entonces haciendo $y = -1$ en (3.3), tenemos $f(-x) = f(x)f(-1) = f(x)$ para todo $x \in \mathbb{R} \setminus \{0\}$. Por lo tanto (3.24) nos resulta

$$f(x) = e^{A(\ln|x|)},$$

para todo $x \in \mathbb{R} \setminus \{0\}$. Dado que $f(0) = 0$, tenemos

$$f(x) = \begin{cases} e^{A(\ln|x|)} & \text{si } x \in \mathbb{R} \setminus \{0\} \\ 0 & \text{si } x = 0, \end{cases}$$

lo cual se puede expresar como $f(x) = e^{A(\ln|x|)}|\text{sgn}(x)|$, la cual es la solución deseada (3.15).

Si $f(-1) = -1$, entonces haciendo $y = -1$ en (3.3), tenemos $f(-x) = f(x)f(-1) = -f(x)$ para todo $x \in \mathbb{R} \setminus \{0\}$. De aquí (3.24) nos lleva a que

$$f(x) = \begin{cases} e^{A(\ln|x|)} & \text{si } x > 0 \\ -e^{A(\ln|x|)} & \text{si } x < 0 \end{cases}$$

para todo $x \in \mathbb{R} \setminus \{0\}$. Luego junto con el hecho que $f(0) = 0$, tenemos

$$f(x) = \begin{cases} e^{A(\ln|x|)} & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -e^{A(\ln|x|)} & \text{si } x < 0 \end{cases}$$

lo cual puede ser expresada así :

$$f(x) = e^{A(\ln|x|)}\text{sgn}(x),$$

lo cual es la solución deseada (3.16). Ahora la prueba del teorema está completa. \square

En virtud del teorema anterior tenemos el siguiente corolario.

Corolario 3.5. *La solución general continua de la ecuación funcional (3.3), es decir, $f(xy) = f(x)f(y)$, cumpliéndose para todo $x, y \in \mathbb{R}$ es dada por*

$$f(x) = 0 \tag{3.26}$$

$$f(x) = 1 \tag{3.27}$$

$$f(x) = |x|^\alpha \tag{3.28}$$

y

$$f(x) = |x|^\alpha \text{sgn}(x) \tag{3.29}$$

donde α es una constante real positiva arbitraria.

Demostración. Por el Teorema 3.3 $f = 0$, o $f = 1$, o f tiene la forma (3.15) o (3.16), donde $A : \mathbb{R} \rightarrow \mathbb{R}$ es una función aditiva. Dado que f es continua y $A(t) = \ln f(e^t)$, $A(t)$ es también continua sobre \mathbb{R} . Por lo tanto

$$A(t) = \alpha t,$$

donde $\alpha \in \mathbb{R}$ es una constante arbitraria. De aquí de (3.15) y (3.16), obtenemos

$$f(x) = |x|^\alpha$$

y

$$f(x) = |x|^\alpha \operatorname{sgn}(x)$$

respectivamente. Lo único que falta mostrar es que $\alpha > 0$. Si tuviéramos $\alpha = 0$, entonces de (3.28) resulta que $f(x) = 1$ para $x \neq 0$, y por continuidad de f se tiene $f(0) = 1$. De aquí tendremos que $f = 1$, es decir, cuando $\alpha = 0$, nos lleva a la solución (3.27). La fórmula (3.29) con $\alpha = 0$ nos lleva a lo siguiente

$$f(x) = 1 \quad \text{para } x > 0$$

y

$$f(x) = -1 \quad \text{para } x < 0$$

y por lo tanto f no puede ser continua. Similarmente si $\alpha < 0$, entonces f dada por (3.28) y (3.29) satisface

$$\lim_{x \rightarrow 0^+} f(x) = \infty$$

y de aquí que f no puede ser continua en 0. Ahora la prueba del corolario está completa. \square

Definición 3.3. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ es llamada una función multiplicativa si satisface $f(xy) = f(x)f(y)$ para todo $x, y \in \mathbb{R}$.

Capítulo 4:

Algunas Aplicaciones de la Ecuación Funcional de Cauchy

En este capítulo se presenta dos aplicaciones de las ecuaciones funcionales de Cauchy en la teoría de la probabilidad: se caracteriza la distribución de probabilidad geométrica por medio de la propiedad de Pérdida (o Falta) de Memoria y se finaliza presentando la caracterización de la distribución de probabilidad normal.

4.1 Caracterización de la Distribución Geométrica

En esta sección se da la caracterización de la distribución geométrica en términos de la propiedad de Pérdida de Memoria usando la ecuación funcional exponencial de Cauchy. Se comienza dando la siguiente definición.

Definición 4.1. Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Una variable aleatoria $X : \Omega \rightarrow \mathbb{N}$ se dice que tiene la propiedad de Pérdida de Memoria si satisface lo siguiente

$$\mathbb{P}(X > m + n \mid X > n) = \mathbb{P}(X > m)$$

para todo $m, n \in \mathbb{N}$.

Proposición 4.1. *La variable aleatoria $X : \Omega \rightarrow \mathbb{N}$ tiene la propiedad de Pérdida de Memoria si y solo si satisface la siguiente condición*

$$\mathbb{P}(X > m+n) = \mathbb{P}(X > m)\mathbb{P}(X > n).$$

Demostración. Usando la definición de probabilidad condicional se tiene que :

$$\mathbb{P}(X > m+n \mid X > n) = \frac{\mathbb{P}\left((X > m+n) \cap (X > n)\right)}{\mathbb{P}(X > n)}. \quad (4.1)$$

Pero como X satisface la propiedad de Pérdida de Memoria, se tiene que

$$\mathbb{P}(X > m+n \mid X > n) = \mathbb{P}(X > m). \quad (4.2)$$

Luego de las ecuaciones (4.1) y (4.2) se obtiene que

$$\mathbb{P}(X > m)\mathbb{P}(X > n) = \mathbb{P}\left((X > m+n) \cap (X > n)\right),$$

de lo cual se sigue que

$$\mathbb{P}(X > m+n) = \mathbb{P}(X > m)\mathbb{P}(X > n)$$

para todo $m, n \in \mathbb{N}$.

Ahora sea X una variable aleatoria que satisface

$$\mathbb{P}(X > m+n) = \mathbb{P}(X > m)\mathbb{P}(X > n).$$

Sabemos que $(X > m+n) \cap (X > n) = (X > m+n)$, entonces

$$\mathbb{P}\left((X > m+n) \cap (X > n)\right) = \mathbb{P}(X > m+n) = \mathbb{P}(X > m)\mathbb{P}(X > n),$$

luego, bajo el supuesto de que $\mathbb{P}(X > n) > 0$ se tiene

$$\frac{\mathbb{P}\left((X > m+n) \cap (X > n)\right)}{\mathbb{P}(X > n)} = \mathbb{P}(X > m)$$

y así

$$\mathbb{P}(X > m+n \mid X > n) = \mathbb{P}(X > m).$$

□

Ahora se presenta la caracterización de la distribución geométrica en términos de la propiedad de Pérdida de Memoria usando la ecuación funcional exponencial de Cauchy:

Teorema 4.1. *Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Una variable aleatoria discreta $X : \Omega \rightarrow \mathbb{N}$ es geométrica (es decir, tiene distribución geométrica) si y sólo si satisface la propiedad de Pérdida de Memoria.*

Demostración. Si X es geométrica, entonces usando el ítem 1 de la Proposición 1.4 se sigue que

$$\begin{aligned} \mathbb{P}(X > m+n) &= (1-p)^{m+n} \\ &= (1-p)^m (1-p)^n \\ &= \mathbb{P}(X > m) \mathbb{P}(X > n). \end{aligned}$$

Luego por la proposición anterior podemos concluir que X satisface la propiedad de Pérdida de Memoria.

Ahora supongamos que X es una variable aleatoria que satisface la propiedad de Pérdida de Memoria,

$$\mathbb{P}(X > m+n \mid X > n) = \mathbb{P}(X > m) \quad (4.3)$$

para todo $m, n \in \mathbb{N}$. Queremos mostrar que X es geométrica. Definamos

$$g : \mathbb{N} \rightarrow \mathbb{R}$$

de la siguiente manera:

$$g(n) = \mathbb{P}(X > n). \quad (4.4)$$

Luego gracias a la proposición anterior, la ecuación dada en (4.3) es equivalente a la ecuación

$$\mathbb{P}(X > m+n) = \mathbb{P}(X > m) \mathbb{P}(X > n)$$

la cual expresada en términos de g equivale a

$$g(m+n) = g(n)g(m),$$

para todo $m, n \in \mathbb{N}$.

Vemos que g es no negativa (pues está definida en términos de una probabilidad) y no creciente, pues si $m < n$ entonces $\{X > n\} \subset \{X > m\}$ y así $g(n) = \mathbb{P}(X > n) \leq \mathbb{P}(X > m) = g(m)$. La solución general (aun sin continuidad) es dada por

$$g(n) = a^n,$$

donde a es una constante, a saber, $a = g(1)$. Luego

$$\mathbb{P}(X > n) = a^n.$$

Gracias al ítem 2 de la Proposición 1.4 lo anterior se puede expresar como:

$$1 - F(n) = a^n,$$

donde $F(n)$ es la función de distribución de X . Luego usando el ítem 1 de la Proposición 1.3 tenemos que

$$\begin{aligned} 1 &= \lim_{n \rightarrow \infty} F(n) \\ &= \lim_{n \rightarrow \infty} (1 - a^n). \end{aligned}$$

De lo anterior, se sigue que $\lim_{n \rightarrow \infty} a^n = 0$. De lo cual $a \in [0, 1[$. Pero como la función de distribución de una variable aleatoria no puede ser constante, se tiene entonces que $a \neq 0$ y así concluimos que $0 < a < 1$. Por conveniencia escribamos $a = 1 - p$, donde $0 < p < 1$. Luego obtenemos

$$F(n) = 1 - (1 - p)^n.$$

La función de de probabilidad de la variable aleatoria X es entonces dado por

$$\begin{aligned} f(1) &= F(1) = p \\ f(2) &= F(2) - F(1) \\ &= 1 - (1 - p)^2 - p \\ &= (1 - p)p \\ f(3) &= F(3) - F(2) \\ &= 1 - (1 - p)^3 - 1 + (1 - p)^2 \\ &= (1 - p)^2 p. \end{aligned}$$

Por lo tanto por inducción, obtenemos

$$f(x) = (1 - p)^{x-1}p$$

para $x = 1, 2, 3, \dots$. Por lo tanto

$$X \sim \text{Geo}(p)$$

lo cual completa la prueba. □

A continuación presentamos un problema original relacionado con la propiedad de pérdida o falta de memoria:

Sourav, estudiante hábil de Matemáticas de la *FACFYM* hace un momento dio un examen de Probabilidad, digamos que ese momento es $t = 0$. Sourav quiere saber cual es la probabilidad de que su profesor entregue los exámenes al cabo de “ n ” días. Obviamente para tal objetivo necesita más datos, así que se propone realizar algunas indagaciones.

Profesor, ¿ cuál es la probabilidad de que entregue los exámenes la próxima clase, que es dentro de $t = 3$ días?, pregunta Sourav. El profesor contesta al salón que en un plazo de $t = 7$ días recibirán sus exámenes.

A lo que Sourav, pregunta si existe alguna probabilidad positiva de que la entrega se retrase algunos días más. Entonces el profesor responde :

si me pasaría del plazo ofrecido, la probabilidad de que me retrase n días más es la misma probabilidad de que entregue los exámenes después de n días de haberlo tomado...

Y para retirarse ya, agrega unas cuantas cosas más:

No me interesa si me paso del plazo de 7 días ofrecido, lo único que es de relevancia para mí son los días extra que me demore después del plazo ofrecido. Es más, lo mismo podría decir para cualquier plazo de tiempo en días.

Pasados los días, 15 días para ser exactos, ¡ el profesor aún no entrega los exámenes ! y los alumnos incómodos y preocupados le reclaman. El profesor enojado le contesta que nunca les entregará sus exámenes y que están todos jalados. En ese momento Sourav le responde de una manera segura y vehemente: ¡ eso es imposible profesor, pues la probabilidad de que nunca entregue los exámenes es cero !. Inmediatamente el profesor, un tanto absorto de su respuesta,

se dirige a todos alumnos y les dice Sourav está del todo correcto y que es el único que merece aprobar el curso de probabilidad!!!

Explique usted, en que se basa la afirmación contundente de Sourav y la decisión inesperado del profesor.

Solución

Definamos la variable aleatoria $X : \Omega \rightarrow \mathbb{N}$ como sigue :

$X :=$ días que esperan los alumnos hasta que se les den sus exámenes

Entendamos mejor a X :

1. el evento $(X = 3)$ significa que los alumnos esperan 3 días para recibir sus exámenes.
2. el evento $(X \leq 15)$ significa que los alumnos esperan a lo mucho 15 días para recibir sus exámenes.
3. el evento $(X > 10)$ significa que los alumnos esperan más de 10 días para recibir sus exámenes o que se les entrega los exámenes después de 10 días

Ahora examinemos lo que dice el profesor por segunda vez:

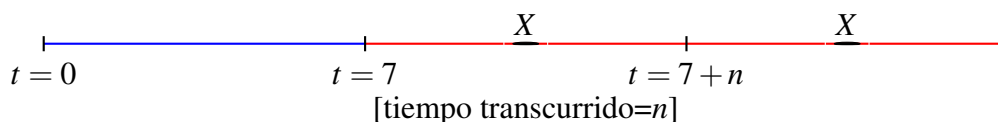
si ya me pasé del plazo ofrecido, la probabilidad de que me retrase n días más es la misma probabilidad de que entregue los exámenes después de n días de haberlo tomado...

Bien, lo que haremos es expresar la expresión anterior en una ecuación en términos de X y la probabilidad \mathbb{P} . Note que en esa expresión tenemos una igualdad en la palabra es. Cuando dice : ya me pasé del plazo ofrecido lo podemos escribir como el evento $(X > 7)$. Cuando dice : que me retrase más de n días, lo podemos expresar como el evento $(X > 7 + n)$. Cuando dice: que entregue los exámenes después de n días de haberlo tomado, lo podemos expresar como $(X > n)$. Luego expresando ese párrafo en terminos de X y \mathbb{P} tendríamos la siguiente ecuación

:

$$\mathbb{P}(X > 7 + n | X > 7) = \mathbb{P}(X > n). \quad (4.5)$$

Note que el segundo miembro de la ecuación depende sólo de n que es el tiempo transcurrido desde $t = 7$, mientras que el primer miembro de la ecuación depende de 7 y n .

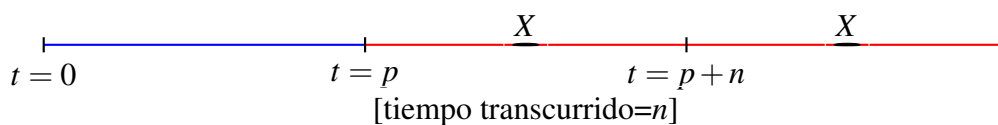


Ahora analicemos la otra afirmación que hace el profesor

No me interesa si me paso del plazo de 7 días ofrecido, lo único que es de relevancia para mí son los días extra que me demore después del plazo ofrecido. Es más, lo mismo podría decir para cualquier plazo de tiempo en días...

Lo anterior se interpreta que hay una pérdida de memoria del tiempo transcurrido $t = 7$ días, que es el plazo ofrecido. Por otro lado, vemos que el plazo puede ser cualquiera, eso nos lleva a generalizar la ecuación (4.5) a la siguiente:

$$\mathbb{P}(X > p + n | X > p) = \mathbb{P}(X > n), \quad (4.6)$$



donde p representa un plazo arbitrario en días, es decir $p \in \mathbb{N}$. Además se tiene también que $n \in \mathbb{N}$. En resumen hemos llegado a la conclusión que X satisface la propiedad de pérdida o falta de memoria. Luego gracias al Teorema 4.1 acerca de la caracterización de la distribución geométrica, tenemos que $X \sim \text{geo}(p)$ para algún $p \in]0, 1[$.

Ahora bien, lo que nos resta justificar es que la probabilidad que el profesor nunca entregue los exámenes en cero. Lo que haremos es expresar el evento que “el profesor nunca

entregue los exámenes” en términos de X . Como en ese evento el tiempo de entrega es indeterminado, lo expresaremos así

$$(X = +\infty)$$

Luego, procederemos a demostrar que la probabilidad del tal evento es cero, es decir que

$$\mathbb{P}(X = +\infty) = 0$$

En efecto, podemos expresar $(X = +\infty)$ como una intersección de los siguientes eventos $(X > n)$, para ser más precisos

$$(X = \infty) = \bigcap_{n=1}^{\infty} (X > n),$$

luego,

$$\mathbb{P}(X = \infty) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} (X > n)\right). \quad (4.7)$$

Como los eventos $(X > n)$ forman una sucesión decreciente, entonces debido a la continuidad inferior de la medida \mathbb{P} , tenemos que

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} (X > n)\right) = \lim_{n \rightarrow \infty} \mathbb{P}(X > n). \quad (4.8)$$

Luego de las ecuaciones (4.7) y (4.8) se sigue que

$$\mathbb{P}(X = \infty) = \lim_{n \rightarrow \infty} \mathbb{P}(X > n), \quad (4.9)$$

pero por el ítem 1 de la Proposición 1.4

$$\mathbb{P}(X > n) = (1 - p)^n, \quad (4.10)$$

Finalmente de (4.9) y (4.10) obtenemos

$$\mathbb{P}(X = \infty) = \lim_{n \rightarrow \infty} (1 - p)^n = 0,$$

pues $p \in]0, 1[$. Lo cual justifica que la probabilidad que el profesor nunca entregue los exámenes es cero y por lo tanto Sourav toda tiene la razón y al parecer el profesor hizo todas afirmaciones con el fin de ver si alguien entendía el problema aleatorio que ocurría...

4.2 Caracterización de la Distribución Normal

Es conocido en la estadística inferencial que, si x_1, x_2, \dots, x_n es una muestra aleatoria de una distribución normal con media μ y varianza σ^2 , entonces la estimación de máxima verosimilitud (EMV) del parámetro μ es dada por la media muestral $\bar{x} = \sum_{i=1}^n x_i/n$. (Lo dicho anteriormente se puede ver en detalle en el ítem (i) del Ejemplo 1.9, Sección 1.6.)

Por otro lado, si la estimación de máxima verosimilitud de un parámetro de una población es dada por la media muestral, ¿es verdad que la distribución de la población es normal? La respuesta a esta cuestión es afirmativa y fue respondida por Gauss en 1809.

En esta sección, usando la ecuación funcional aditiva de Cauchy, presentamos la caracterización más antigua de la distribución normal. La demostración que se da a continuación es una versión adaptada de la prueba reciente dada por Azzalini y Genton en [4] para una muestra de tamaño $n \geq 3$.

Teorema 4.2. *Considere una familia de ubicación uniparamétrica para una variable aleatoria $Z : \Omega \rightarrow \mathbb{R}$ absolutamente continua, tal que para alguna elección de $\mu \in \mathbb{R}$ la correspondiente función de densidad en el punto $x \in \mathbb{R}$ es $f(x - \mu)$. Asuma que una muestra aleatoria de tamaño $n \geq 3$ es extraída de un miembro de esta familia uniparamétrica, y que las siguientes condiciones se cumplen:*

1. *$f(x)$ es una función positiva diferenciable de x y su derivada $f'(x)$ es continua en al menos un punto $\tilde{x} \in \mathbb{R}$;*
2. *para cada conjunto de valores muestrales x_1, x_2, \dots, x_n la media muestral $\bar{x} = \sum_{i=1}^n x_i/n$ es una solución de la ecuación de máxima verosimilitud para el parámetro μ .*

Entonces la función de densidad $f(x - \mu)$ es la función de densidad normal dada por

$$f(x - \mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

para algún σ^2 positivo.

Demostración. Sea la función de verosimilitud asociada con la muestra x_1, \dots, x_n

$$L(\mu, \mathbf{x}) = f(x_1 - \mu) \times \cdots \times f(x_n - \mu).$$

Luego la función log-verosimilitud es entonces dada por

$$\ln L(\mu, \mathbf{x}) = \sum_{i=1}^n \ln f(x_i - \mu). \quad (4.11)$$

Y el estimador de máxima verosimilitud para μ se obtiene derivando la ecuación anterior con respecto a μ , luego igualando a cero y resolviendo la ecuación (llamada ecuación de verosimilitud)

$$\frac{\partial}{\partial \mu} \ln L(\mu, \mathbf{x}) = \sum_{i=1}^n \frac{d}{d\mu} \ln f(x_i - \mu) = 0. \quad (4.12)$$

Ahora lo que haremos es expresar la ecuación anterior en términos de una función $g : \mathbb{R} \rightarrow \mathbb{R}$ definida de la siguiente manera :

$$g(x) = \frac{d}{dx} \ln f(x), \quad (4.13)$$

y así la ecuación de máxima verosimilitud (4.12) se escribiría como

$$\frac{\partial}{\partial \mu} \ln L(\mu, \mathbf{x}) = \sum_{i=1}^n g(x_i - \mu) = 0 \quad (4.14)$$

Por la condición (2) del teorema, tenemos que $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ es solución de (4.14) y por lo cual $\frac{\partial}{\partial \mu} \ln L(\mu, \mathbf{x})|_{\mu=\bar{x}} = 0$, es decir,

$$\sum_{i=1}^n g(x_i - \bar{x}) = 0. \quad (4.15)$$

Recuerde que x_1, x_2, \dots, x_n es una muestra arbitraria. Así que haciendo $x_1 = x_2 = \cdots = x_n = 0$, se sigue que $\bar{x} = 0$, y luego en (4.15) obtenemos que $\sum_{i=1}^n g(0 - 0) = 0$, de lo cual se desprende que

$$g(0) = 0. \quad (4.16)$$

Considere ahora la muestra $x_1 = 2u, x_2 = 0, x_3 = \dots = x_n = u, u \in \mathbb{R}$, donde de nuevo $\bar{x} = u$, luego en (4.15) obtenemos

$$g(2u - u) + g(0 - u) + \sum_{i=3}^n g(u - u) = 0,$$

usando (4.16) en la expresión anterior y despejando $g(-u)$ vemos que

$$g(-u) = -g(u), \quad (4.17)$$

lo cual nos dice que g es una función impar.

Ahora bien, para dos puntos u y v en \mathbb{R} cualesquiera, considere la muestra $x_1 = u$, $x_2 = v$, $x_3 = -u - v$, $x_4 = x_5 = \dots = x_n = 0$, con lo cual $\bar{x} = 0$ y reemplazando en (4.15) obtenemos

$$g(u - 0) + g(v - 0) + g(-u - v - 0) + (n - 3)g(0 - 0) = 0, \quad (4.18)$$

luego usando (4.16) y (4.17), la ecuación anterior se reduce a

$$g(u + v) = g(u) + g(v), \quad (4.19)$$

para todo $u, v \in \mathbb{R}$, es decir, g es una función aditiva. Por la condición (1) del teorema, f' es continua en un punto \tilde{x} , lo cual nos conduce a que g sea también continua en \tilde{x} . Luego usando el Teorema 2.5 se sigue que g es lineal de la forma

$$g(x) = -cx, \quad (4.20)$$

para alguna constante $c \in \mathbb{R}$ que no depende de x , donde el signo menos de c es insertado por mera conveniencia de notación. Luego de (4.20) en (4.13) obtenemos la siguiente ecuación diferencial

$$-cx = \frac{d}{dx} \ln f(x),$$

cuya solución es $f(x) = e^{d - \frac{1}{2}cx^2}$ y por lo cual $f(x - \mu) = e^{d - \frac{1}{2}c(x - \mu)^2}$, donde d es alguna constante real.

Dado que $f(x - \mu)$ es una función de densidad, c debe ser mayor que cero, de lo contrario f no sería integrable sobre \mathbb{R} . Haciendo $c = 1/\sigma^2$, donde $\sigma \in \mathbb{R}$, tenemos

$$f(x - \mu) = e^{d - \frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2}.$$

De nuevo, dado que $f(x - \mu)$ es una función de densidad, su integral sobre \mathbb{R} debe ser igual a 1, es decir,

$$\int_{\mathbb{R}} e^{d - \frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2} = 1. \quad (4.21)$$

Es conocido que $\int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi}$, entonces con un cambio de variable adecuado se puede probar que

$$\int_{\mathbb{R}} e^{-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2} dx = \sqrt{2\pi\sigma^2}. \quad (4.22)$$

Así que de las ecuaciones (4.21) y (4.22) concluimos que $e^d = \frac{1}{\sqrt{2\pi\sigma^2}}$. Por lo tanto

$$f(x - \mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2}.$$

Lo cual completa la prueba del teorema. □

Conclusiones

- En la presente tesis únicamente se mostraron resultados y ejemplos ya existentes de una manera más detallada. La mayoría de estos resultados están dados para funciones en la recta real \mathbb{R} .
- Hemos visto que las condiciones de regularidad que debe cumplir una función f que satisface (2.1) para ser lineal, son continuidad, integrabilidad local, continuidad en un punto y acotada superior o inferiormente. No obstante, no descartamos la existencia de otras condiciones de regularidad.
- Por otro lado, así como existen funciones continuas que satisfacen la ecuación (2.1), a saber, las funciones lineales, también existen funciones discontinuas que satisfacen dicha ecuación. Esta última clase de funciones presenta un comportamiento extraño: sus gráficas son subconjuntos denso del plano \mathbb{R}^2 . Cabe resaltar aquí la importancia de las bases de Hamel, pues sin ellas sería imposible la construcción de dichas funciones discontinuas.
- Finalmente, en la aplicaciones vemos que es posible expresar muchas veces un problema matemático en términos de una ecuación funcional.

Sugerencias

Hasta ahora se ha estudiado solamente las ecuaciones funcionales de Cauchy en la recta. Sin embargo el conocimiento no tiene límites, por tal motivo nos permitimos sugerirles lo siguiente:

- Para trabajos posteriores se sugiere estudiar las ecuaciones funcionales de Cauchy en el \mathbb{R}^n .
- También el estudio de la extensión de funciones aditivas es una tema muy interesante que surge de la siguiente interrogante:
Dado un intervalo $[a, b]$ de la recta y $f : [a, b] \rightarrow \mathbb{R}$ una función aditiva solamente sobre el intervalo $[a, b]$. Existirá alguna función aditiva $A : \mathbb{R} \rightarrow \mathbb{R}$ tal que $A|_{[a, b]} = f$?
- Existen muchas ecuaciones funcionales más complejas y bellas que podrían ser estudiadas, tales como la ecuación funcional de Jensen, Pexider, Pompeiu, Hosszú, Abel, tomando como guía el libro de P. Sahoo y P. Kannappan.

Otras sugerencias

- A nuestros compañeros de estudios, seguir desarrollando trabajos relacionados a este tema y que tomen como base lo que nosotros estamos presentando en nuestra tesis.
- A nuestros profesores que promuevan el desarrollo de estos temas que tal vez desconocidos por muchos alumnos de matemáticas de la Facfym, pero muy populares en el exterior. Solo para nombrar, en la Olimpiada Internacional de Matemáticas (IMO) es muy común que se presente una ecuación funcional en las pruebas. (Puede ver [9])

- A las autoridades de la Facultad de Ciencias Físicas y Matemáticas que impulsen trabajos relacionados a este tema y sus aplicaciones.

Bibliografía

- [1] **Aczél, J.** *Functional equations: History, Applications and Theory*. Reidel D. Publishing Company, New York, 1984.
- [2] **Aczél, J. y Dhombres, J.** *Functional equations in several variables*. Cambridge University Press, 1989.
- [3] **Aczél, J.** *Lectures on functional equations and their applications*. Academic Press, 1966.
- [4] **Azzalini, Adelchi y Genton, Marc G.** *On Gauss's characterization of the normal distribution*. Bernoulli Society for Mathematical Statistics and Probability, Vol. 13, No. 1, pp. 169-174, Feb., 2007.
- [5] **Bartle, Robert G.** *The Elements of Integration and Lebesgue Measure*. Wiley Classics Library Edition Published, 1995.
- [6] **Bolfarine, Heleno y Carneiro, Mônica.** *Introducao à Inferência Estatística*. Universidade de Sao Paolo, 2010.
- [7] **Cauchy, A. L.** *Cours d'analyse de l'Ecole Polytechnique, Vol. 1. Analyse Algebrique. Cap V*. Paris, 1821.
- [8] **Darboux, G.** *Sur la composition des forces en statique*. Bull. Sci.Math., 9:281-288, 1875.
- [9] **Djukić, D; Janković, V; Matić, I; Petrović, N.** *The IMO Compendium. A Collection of Problems Suggested for The International Mathematical Olympiads: 1959-2009*. Springer, Second Edition, 2011.

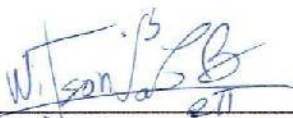
-
- [10] **Grossman, Stanley y Flores, José.** *Álgebra Lineal*.- Séptima Edición, McGraw-Hill, 2012.
- [11] **Jech, T.** *Set Theory*. Springer, 2003.
- [12] **Lages, Elon.** *Curso de Analise Matemático. Volumen 1*. IMPA, 1991.
- [13] **Sahoo, Prasanna y Kannappan, Palaniappan.** *Introduction to Functional Equations*. Chapman & Hall/CRC, 2011.
- [14] **Shapiro, H. N.** *A micronote on a functional equation*. Amer. Math.Monthly, 80:1041, 1973.
- [15] **Rincón, Luis.** *Curso intermedio de Probabilidad*. Facultad de Ciencias Unam, 2005.
- [16] **Roussas, George G.** *An Introduction to Probability and Statistical Inference*. Second Edition, Elsevier, 2015.
- [17] **Christopher G. Small.** *Functional Equations and How to Solve Them*. Springer, 2007.
- [18] **Sierpinski, W.** *Sur l'equation fonctionnelle $f(x+y) = f(x) + f(y)$* . Fund. Math., pp. 116-122, 1920
-

UNIVERSIDAD NACIONAL "PEDRO RUIZ GALLO"

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

ESCUELA PROFESIONAL DE MATEMÁTICA

**"La Ecuación Funcional Aditiva de Cauchy de
 \mathbb{R} en \mathbb{R} "**



Bach. Mat. Cabanillas Banda, Wilson Alberto

Autor



Bach. Mat. Vera Rubio, José Gilmer

Autor



Mg. Malca Villalobos, Amado

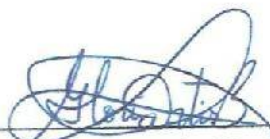
Asesor

Lambayeque – Perú

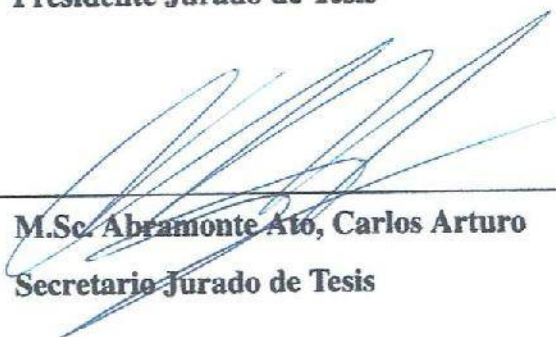
Febrero - 2017

UNIVERSIDAD NACIONAL "PEDRO RUIZ GALLO"
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA

Los firmantes por la presente certifican que han leído y recomiendan a la Facultad de Ciencias Físicas y Matemáticas la aceptación de la tesis titulada "**La Ecuación Funcional Aditiva de Cauchy de \mathbb{R} en \mathbb{R}** ", presentada por los Bachilleres en Matemáticas, en el cumplimiento parcial de los requisitos necesarios para la obtención del título profesional de Licenciado en Matemáticas.



Dra. Ortiz Basauri, Gloria María
Presidente Jurado de Tesis



M.Sc. Abramonte Ato, Carlos Arturo
Secretario Jurado de Tesis



Mg. Pérez Herrera, Adelmo
Vocal Jurado de Tesis

Fecha de Defensa: 15 de Febrero del 2017