



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

1 de 2

Neiva, JULIO 2019

Señores

CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN  
UNIVERSIDAD SURCOLOMBIANA  
Ciudad

El (Los) suscrito(s):

JESSICA FRINETH CULMA PERDOMO , con C.C. No. 1.075.278.836 ,

LAURA JULIANA BOHORQUEZ VERGARA , con C.C. No. 1.118.562.091 ,

Autor(es) de la tesis y/o trabajo de grado titulado: EL ESTUDIO ALELICO DE LA ENFERMEDAD "LA FIBROSIS QUISTICA". Presentado y aprobado en el año 2018 como requisito para optar al título de: Matemático.

Autorizo (amos) al CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN de la Universidad Surcolombiana para que, con fines académicos, muestre al país y el exterior la producción intelectual de la Universidad Surcolombiana, a través de la visibilidad de su contenido de la siguiente manera:

Los usuarios puedan consultar el contenido de este trabajo de grado en los sitios web que administra la Universidad, en bases de datos, repositorio digital, catálogos y en otros sitios web, redes y sistemas de información nacionales e internacionales "open Access" y en las redes de información con las cuales tenga convenio la Institución.

Permita la consulta, la reproducción y préstamo a los usuarios interesados en el contenido de este trabajo, para todos los usos que tengan finalidad académica, ya sea en formato Cd-Rom o digital desde internet, intranet, etc., y en general para cualquier formato conocido o por conocer, dentro de los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia.

Continúo conservando los correspondientes derechos sin modificación o restricción alguna; puesto que, de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación del derecho de autor y sus conexos.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

EL AUTOR/ESTUDIANTE:

Firma:

EL AUTOR/ESTUDIANTE:

Firma:

Vigilada Mineducación



**TÍTULO COMPLETO DEL TRABAJO: EL ESTUDIO ALELICO DE LA ENFERMEDAD “LA FIBROSIS QUISTICA”**

**AUTOR O AUTORES:**

Primero y Segundo Apellido	Primero y Segundo Nombre
CULMA PERDOMO	JESSICA FRINETH
BOHORQUEZ VERGARA	LAURA JULIANA

**DIRECTOR Y CODIRECTOR TESIS:**

Primero y Segundo Apellido	Primero y Segundo Nombre
MEDINA ARCE	YINETH

**ASESOR (ES):**

Primero y Segundo Apellido	Primero y Segundo Nombre
MEDINA ARCE	YINETH

**PARA OPTAR AL TÍTULO DE: MATEMATICO**

**FACULTAD:** CIENCIAS EXACTAS Y NATURALES.

**PROGRAMA O POSGRADO:** MATEMATICA APLICADA

**CIUDAD:** NEIVA      **AÑO DE PRESENTACIÓN:** 2018      **NÚMERO DE PÁGINAS:** 122

**TIPO DE ILUSTRACIONES** (Marcar con una X):

Diagramas\_\_X\_ Fotografías\_X\_ Grabaciones en discos\_\_\_ Ilustraciones en general\_\_X\_ Grabados\_\_\_  
Láminas\_\_\_ Litografías\_\_\_ Mapas\_X\_ Música impresa\_\_\_ Planos\_\_\_ Retratos\_\_\_ Sin ilustraciones\_\_\_  
Tablas o Cuadros\_X\_

**SOFTWARE** requerido y/o especializado para la lectura del documento: EXCEL, POPULUS, R-PROJECT, TEXT-MAKER.

**MATERIAL ANEXO: TABLAS DE DATOS SUMINISTRADOS POR LA SECRETARIA DE SALUD DEPARTAMENTAL.**

Vigilada mieducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional [www.usco.edu.co](http://www.usco.edu.co), link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



<b>CÓDIGO</b>	<b>AP-BIB-FO-07</b>	<b>VERSIÓN</b>	<b>1</b>	<b>VIGENCIA</b>	<b>2014</b>	<b>PÁGINA</b>	<b>2 de 3</b>
---------------	---------------------	----------------	----------	-----------------	-------------	---------------	---------------

**PREMIO O DISTINCIÓN** (En caso de ser LAUREADAS o Meritoria):

**PALABRAS CLAVES EN ESPAÑOL E INGLÉS:**

<u>Español</u>	<u>Inglés</u>	<u>Español</u>	<u>Inglés</u>
1. Alelos	Alleles	6. Frecuencias	Frequencies
2. Población	Population	7. Modelo Hardy - Weinberg	Hardy Model - Weinberg
3. Fibrosis Quística	Cystic Fibrosis	8. Cadenas de Markov	Markov Chains
4. Leyes de Mendel	Mendel's Laws	9. CFTR	CFTR
5. Probabilidades	Probabilities	10. neumonía	pneumonia

**RESUMEN DEL CONTENIDO:** (Máximo 250 palabras)

Este trabajo de grado se centra en la problemática presente en el estudio alélico de la enfermedad "fibrosis quística", la cual es una enfermedad hereditaria que afecta a la raza caucásica y de cual se ve afectada gran parte de la población mundial.

Se realizó un estudio alélico en torno a la problemática, de la enfermedad fibrosis quística, caracterizada como una enfermedad multisistémica y catalogada como enfermedad huérfana en Colombia, que no afecta la inteligencia ni la apariencia; generalmente, suele diagnosticarse en la etapa de la infancia debido a la presencia de neumonía o retrasos en el crecimiento.

A nivel departamental, el objetivo fue, realizar un estudio estadístico de los casos de la enfermedad "fibrosis quística" presentados en el Huila, para identificar sus parámetros poblacionales y verificarlos mediante el modelo matemático de Hardy - Weinberg.

**ABSTRACT:** (Máximo 250 palabras)

This work of degree focuses on the problems present in the allelic study of the disease "cystic fibrosis", which is a hereditary disease that affects the Caucasian race and of which a large part of the world population is affected.

An allelic study was conducted around the problem of cystic fibrosis disease, characterized as a



DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO

CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	3 de 3
--------	--------------	---------	---	----------	------	--------	--------

multisystemic disease and classified as an orphan disease in Colombia, which does not affect intelligence or appearance; Generally, it is usually diagnosed at the childhood stage due to the presence of pneumonia or growth retardation.

At the departmental level, the objective was to carry out a statistical study of the cases of the disease "cystic fibrosis" presented in Huila, to identify their population parameters and verify them using the mathematical model of Hardy – Weinberg.

### APROBACION DE LA TESIS

Nombre Presidente Jurado: Mauro Montealegre Cárdenas.

Firma:

Nombre Jurado: Mauro Montealegre Cárdenas.

Firma:

Nombre Jurado: Luis Arturo Polania Quizá.

Firma:



# **EL ESTUDIO ÁLELICO DE LA ENFERMEDAD "LA FIBROSIS QUÍSTICA"**

**JESSICA FRINETH CULMA PERDOMO**

**Código:2011198830**

**LAURA JULIANA BOHORQUEZ VERGARA**

**Código: 20122112825**

UNIVERSIDAD SURCOLOMBIANA  
FACULTAD DE CIENCIAS EXACTAS Y NATURALES  
MATEMÁTICA APLICADA  
NEIVA-HUILA, COLOMBIA  
2018

# **EL ESTUDIO ÁLELICO DE LA ENFERMEDAD "LA FIBROSIS QUÍSTICA"**

**JESSICA FRINETH CULMA PERDOMO**

**Código:2011198830**

**LAURA JULIANA BOHORQUEZ VERGARA**

**Código:20122112825**

Trabajo de grado presentado como requisito parcial para optar al título de:  
**MATEMÁTICO**

Supervisor: MG. YINETH MEDINA ARCE

Línea de Investigación:  
Biomatemática  
Universidad Surcolombiana  
Facultad de Ciencias Exactas  
Matemática Aplicada  
Neiva - Huila, Colombia  
2018

# DEDICATORIA

Este trabajo de grado se lo queremos dedicar a nuestros padres por el apoyo inconmensurable en nuestras vidas tanto personal como profesional, al formarnos mujeres integras para la sociedad; cabe destacar a nuestras hijas por el inmenso amor que nos brindan con cada gesto y sonrisa, otorgandonos la fuerza y valentia para lograr cada una de nuestras metas propuestas.

Un matemático, como un pintor o un poeta, es un creador de patrones. Si sus patrones son más permanentes que los de ellos, es porque están hechos con ideas.

G.H. Hardy

# AGRADECIMIENTOS

Primeramente agradecemos a Dios, a nuestros padres e hijas por ser los principales promotores de nuestros sueños, gracias a ellos por cada día confiar y creer en nosotras y nuestras expectativas, gracias a nuestras madres por estar dispuestas a acompañarnos en nuestras largas y agotadoras noches de estudio, en las que su compañía y las llegadas de sus cafés eran para nosotras como agua en el desierto; gracias a nuestros padres por siempre desearnos lo mejor para nuestras vidas, gracias por los consejos y por cada una de sus palabras que nos guiaron durante este proceso y por apoyarnos siempre; gracias a nuestras hijas por ser la fuente de nuestra motivación e inspiración para poder superarnos cada día más y así poder luchar para que la vida nos depare un futuro mejor.

Gracias a la Universidad Surcolombiana por habernos aceptado ser parte de ella y abriendo las puertas de su seno para poder estudiar nuestra carrera, así como también a los diferentes docentes que nos brindaron sus conocimientos y su apoyo para seguir adelante día a día.

Agradecemos a nuestra asesora del Trabajo de Grado la Mg. Yineth Medina Arce, por habernos brindado la oportunidad de recurrir a su capacidad y conocimiento, desde luego habernos tenido toda la paciencia del mundo para guiarnos durante este proceso, creer en nosotras y apoyarnos en todo.

Finalizando, también agradecemos a nuestros compañeros de clase durante los semestres de la carrera, ya que gracias al compañerismo, amistad y apoyo moral han aportado nuestras ganas de seguir adelante con nuestra carrera profesional.

# RESUMEN

Este trabajo de grado se centra en la problemática presente en el estudio alélico de la enfermedad "fibrosis quística", la cual es una enfermedad hereditaria que afecta a la raza caucásica y de cual se ve afectada gran parte de la población mundial.

Se realizó un estudio alélico en torno a la problemática, de la enfermedad fibrosis quística, caracterizada como una enfermedad multisistémica y catalogada como enfermedad huérfana en Colombia, que no afecta la inteligencia ni la apariencia; generalmente, suele diagnosticarse en la etapa de la infancia debido a la presencia de neumonía o retrasos en el crecimiento.

A nivel departamental, el objetivo fue, realizar un estudio estadístico de los casos de la enfermedad "fibrosis quística" presentados en el Huila, para identificar sus parámetros poblacionales y verificarlos mediante el modelo matemático de Hardy - Weinberg.

**Palabras clave:** Alelos, Población, Fibrosis Quística, CFTR, Leyes de Mendel, Probabilidades, Frecuencias, Modelo Hardy - Weinberg, Cadenas de Markov.

# SUMMARY

This work of degree focuses on the problems present in the allelic study of the disease cystic fibrosis", which is a hereditary disease that affects the Caucasian race and of which a large part of the world population is affected.

An allelic study was conducted around the problem of cystic fibrosis disease, characterized as a multisystemic disease and classified as an orphan disease in Colombia, which does not affect intelligence or appearance; Generally, it is usually diagnosed at the childhood stage due to the presence of pneumonia or growth retardation.

At the departmental level, the objective was to carry out a statistical study of the cases of the disease cystic fibrosis" presented in Huila, to identify their population parameters and verify them using the mathematical model of Hardy - Weinberg.

**Keywords:** Alleles, Population, Cystic Fibrosis, CFTR, Mendel's Laws, Probabilities, Frequencies, Hardy Model - Weinberg, Markov Chains.

# Índice general

---

<b>Agradecimientos</b>	<b>4</b>
<b>Resumen</b>	<b>5</b>
<b>Índice general</b>	<b>6</b>
<b>1 INTRODUCCIÓN</b>	<b>8</b>
<b>2 PLANTEAMIENTO DEL PROBLEMA</b>	<b>10</b>
<b>3 OBJETIVOS</b>	<b>11</b>
3.1. OBJETIVO GENERAL . . . . .	11
3.2. OBJETIVOS ESPECIFICOS . . . . .	11
<b>4 JUSTIFICACIÓN</b>	<b>12</b>
4.1. EL GEN CFTR . . . . .	16
4.1.1. LA MUTACIÓN DF508 . . . . .	18
4.1.2. EL ESPECTRO DE MUTACIONES DEL GEN CFTR . . . . .	19
4.1.3. CORRELACIÓN GENOTIPO - FENOTIPO . . . . .	19
4.1.4. LA PROTEÍNA CFTR Y SUS FUNCIONES . . . . .	20
4.1.5. FUNCIONES DE LA PROTEÍNA . . . . .	21
4.2. CLASIFICACIÓN DE LA FIBROSIS QUISTICA Y SUS MANIFESTACIONES	24
<b>5 MARCO TEÓRICO</b>	<b>26</b>
5.1. BINOMIO DE NEWTON . . . . .	26
5.2. TRIÁNGULO DE PASCAL . . . . .	28
5.3. PROBABILIDAD Y ESTADÍSTICA . . . . .	29
5.4. PROCESOS DE MARKOV O CADENAS DE MARKOV . . . . .	31
5.5. LEYES MENDELIANAS . . . . .	32
5.5.1. 1ª LEY DE MENDEL: Ley de la uniformidad de los híbridos de la primera generación filial. . . . .	32
5.5.2. 2ª LEY DE MENDEL:Ley de la segregación de los caracteres en la segunda generación filial. . . . .	33
5.5.3. 3ª LEY DE MENDEL: Ley de la independencia de los caracteres heredi- tarios. . . . .	35
5.6. FIBROSIS QUISTICA (FQ) . . . . .	36

5.6.1.	ANTECEDENTES HISTÓRICOS . . . . .	37
5.7.	MODELO MATEMÁTICO DE HARDY - WEINBERG . . . . .	39
5.7.1.	EL MODELO DENOMINADO LEY HARDY - WEINBERG (EQUILIBRIO) . . . . .	39
5.7.2.	POBLACIÓN INFINITA . . . . .	42
5.8.	METODOLOGÍAS ESTADÍSTICA . . . . .	45
5.8.1.	MEDIDAS DE DISPERSIÓN . . . . .	46
5.8.2.	MEDICIÓN DE LA MORTALIDAD . . . . .	47
<b>6</b>	<b>METODOLOGÍA</b>	<b>48</b>
<b>7</b>	<b>ANÁLISIS ESTADÍSTICO DE LA FQ CON RESPECTO AL HUILA</b>	<b>49</b>
<b>8</b>	<b>RESULTADOS</b>	<b>58</b>
8.0.1.	E840 . . . . .	58
8.0.2.	E841 . . . . .	61
8.0.3.	E848 . . . . .	63
8.0.4.	E849 . . . . .	66
8.1.	RESULTADOS DEL ANÁLISIS DEL CUADRO (7-2) . . . . .	73
8.1.1.	E840 . . . . .	73
8.1.2.	E841 . . . . .	76
8.1.3.	E848 . . . . .	78
8.1.4.	E849 . . . . .	81
8.2.	RESULTADOS DEL ANÁLISIS DEL CUADRO (7-3) . . . . .	87
8.2.1.	E840 . . . . .	87
8.2.2.	E841 . . . . .	89
8.2.3.	E849 . . . . .	92
8.2.4.	TASA ESPECIFICA DE MORTALIDAD . . . . .	98
8.3.	ESTUDIO 1: UN GRUPO DE FAMILIAS PORTADORAS DE LA ENFERMEDAD FQ . . . . .	99
8.4.	ESTUDIO 2: UNA FAMILIA PORTADORA DE LA ENFERMEDAD FQ . . . . .	102
8.5.	SIMULACIÓN EN SOFTWARE EXCEL Y POPULUS . . . . .	107
8.5.1.	SIMULACIÓN EJEMPLO 1 . . . . .	107
8.5.2.	SIMULACIÓN EJEMPLO 2: . . . . .	113
<b>9</b>	<b>CONCLUSIONES</b>	<b>115</b>
<b>A</b>	<b>TÉRMINOS, CONCEPTOS Y DEFINICIONES</b>	<b>116</b>
	<b>Bibliografía</b>	<b>120</b>

## *Capítulo 1*

---

# INTRODUCCIÓN

---

La fibrosis quística (FQ); es una enfermedad multisistémica con un patrón de herencia autosómica recesiva, que afecta preferentemente a la población de origen caucásico y con una incidencia global de 1 : 2500 recién nacidos. Se produce, a partir de una mutación del gen, que codifica la proteína reguladora de la conductancia transmembrana o CFTR, ubicado en el brazo largo del cromosoma 7. La relación entre el genotipo y el fenotipo pulmonar es menos patente, probablemente debido a las modificaciones genéticas y a factores ambientales.

Para el estudio alélico de esta enfermedad se llevó acabo unas series de formulas y conceptos matemáticos, estadísticos y genéticos, para lograr una mejor proyección en la afectación y equilibrio de esta enfermedad a la población huilense.

Los avances en el estudio de la FQ, han contribuido favorablemente en la supervivencia de estos pacientes, permitiéndoles alcanzar en la década de la 90 una edad media de 25 - 30 años, según países y pasando a integrarse en su asistencia los médicos y especialistas de adultos. La creación de la Unidades de FQ, ha demostrado ser un gran avance en las mejoras del control y seguimiento de los pacientes con FQ, pues esta enfermedad multisistémica precisa de la existencia de un equipo de profesionales conducidos por un experto en FQ que aglutine las diferentes especialidades.

Los acontecimientos más significativos han sido los siguientes:

- Avances en la terapia con los nuevos enzimas pancreáticos (microesférulas y minimicroesférulas) y las mejoras en la nutrición.
- Identificación del gen que produce la FQ, Collins, Tsui y Riordan; la detección de las mutaciones responsables. Diagnóstico prenatal y neonatal. Terapia génica.
- Ampliación de los conocimientos sobre funcionamiento y mecanismo en el ámbito de las membranas del canal INTRODUCCIÓN CFTR que regula y afecta a los canales de los iones, cloro y sodio.

- Nuevos antibióticos y otras terapias (antiinflamatorios, mucolíticos, tratamiento en "el domicilio", etc.) para mejorar la lucha contra la infección pulmonar.
- Desarrollo de los trasplantes, pulmonar (1990) y hepático.
- Colaboración internacional en la investigación y sobre todo en la divulgación de los conocimientos sobre la FQ y la difusión por todos los medios informativos de la enfermedad, permitiendo que la sociedad sea más participativa en la lucha contra la FQ.

## *Capítulo 2*

---

# **PLANTEAMIENTO DEL PROBLEMA**

---

Aportar una herramienta científica, que sirva de apoyo especialmente a las entidades de salud pública del departamento del Huila, a la hora de tomar decisiones en la investigación, vigilancia, control y lucha contra la enfermedad de la fibrosis quística. El componente principal de la investigación es científico; se empleó el modelo matemático de Hardy - Weinberg para determinar la dinámica de la transmisión genética de los alelos involucrados en la fibrosis quística, el análisis de los equilibrios y la estimación de parámetros de esta enfermedad tanto para la zona rural como la urbana del departamento del Huila.

## Capítulo 3

---

# OBJETIVOS

---

### 3.1. OBJETIVO GENERAL

Aplicar el modelo matemático de Hardy- Weinberg, para simular la dinámica respecto a la enfermedad de la fibrosis quística, y determinar las poblaciones de equilibrio en la población opita ; con lleve a la construcción de un sistema dinámico que modele la probabilidad de transmisión hereditaria, tanto en las familias y personas portadoras de la fibrosis quística, con el cual se apoyen las entidades de salud pública departamental puedan crear estrategias orientadas a tratar, prevenir y disminuir su incidencia de la enfermedad en el Huila.

### 3.2. OBJETIVOS ESPECIFICOS

- Recolectar datos e información pertinente sobre los pacientes que poseen la enfermedad de la fibrosis quística, los posibles portadores de esta enfermedad, incluyendo sobre los recursos, los factores ambientales y económicos.
- Desarrollar un modelo matemático que indique la probabilidad hereditaria de la fibrosis quística, teniendo en cuenta el núcleo familiar.
- Mediante los métodos matemáticos clásicos con ayuda de modeladores computacionales estimar los parámetros de transmisión hereditaria de la fibrosis quística.
- Dar a conocer más esta enfermedad a las personas del Huila y despertar conciencia ciudadana en cuanto a su prevención, control y consecuencias.

## Capítulo 4

---

# JUSTIFICACIÓN

---

Considerando los informes de la Organización Mundial de la Salud (OMS), Sobre el estudio de los **Defectos congénitos** a nivel mundial, **Control de Enfermedades Genéticas** y el estudio de las **Enfermedades Raras**; en el Capítulo XVII de la décima revisión de la Clasificación Estadística Internacional de Enfermedades y Problemas de Salud Conexos (**CIE-10**), dedicado a las «**Malformaciones congénitas, deformaciones y anomalías cromosómicas**», figuran ciertos defectos congénitos que forman un grupo heterogéneo de trastornos de origen prenatal que pueden obedecer a la presencia de un solo gen defectuoso, a alteraciones cromosómicas, a una combinación de factores hereditarios, agentes teratógenos presentes en el medio ambiente o a carencias de micronutrientes.

Además, los avances de las últimas dos décadas en el terreno de la genómica han puesto de relieve que la categoría tradicional de enfermedades genéticas, abarca sólo las dolencias en la contribución de los genes es particularmente importante como la interacción de los genes entre sí y con factores ambientales determina muchos de los aspectos de la salud humana y las enfermedades. Las enfermedades genéticas suelen clasificarse en trastornos monogénicos (hemoglobinopatías, fibrosis quística y hemofilia) y trastornos cromosómicos (por ejemplo, el síndrome de Down).

En los trastornos monogénicos encontramos la enfermedad de la Fibrosis Quística; la cual vamos abarcar en el trabajo de grado y se debe a la presencia de un gen anormal en cualquiera de los autosomas; donde este gen es recesivo y la enfermedad sólo se expresa cuando el gen defectuoso se hereda de los dos progenitores (en este caso se habla de trastorno autosómico recesivo); en donde ambos padres son portadores sanos de la enfermedad y tienen un riesgo del 25 % de heredarla a sus hijos) más común en la población caucásica (raza blanca) y se presenta en 1/2,000 a 1/2,500 en el mundo; con un riesgo de recurrencia en cada gestación del 25 % (homocigota), los portadores (heterocigotos) suman un 50 % y el otro 25 % son sanos. Los pacientes con fibrosis quística (FQ) tienen alterado un conducto en las células que es el que

permite regular, a través de la entrada y la salida del Cloro, la cantidad de agua que sale al exterior.

Cuando éste se encuentra dañado las secreciones son espesas en las vías aéreas, páncreas e intestinos, por la falta de agua. En su forma clásica, las manifestaciones clínicas de la FQ afectan a las glándulas sudoríparas y a los tejidos pulmonar, digestivo y reproductor.

El número de casos de FQ en cada país varía según el origen étnico de las poblaciones. Las razas humanas se distinguen unas de otras por posesión de agregados de diferentes genes, o de los mismos genes en distintas frecuencias o de ambas cosas. Todas las razas son poblaciones mendelianas que cambian en el tiempo por la influencia de los factores ambientales. En algunas zonas de Gran Bretaña se encuentran 1/367 recién nacidos afectados, en la población asiática 1/90,000 y en Europa los datos se encuentran (**la Tabla 1**). Aunque la FQ es más frecuente en la población caucásica, una de cada 25 personas es portadora de la enfermedad. Los estudios genéticos en familias con FQ se iniciaron en 1985; utilizando marcadores genéticos que se localizaban en la región del gen **CFTR(cystic fibrosis transmembrane conductance regulator)**. Esto supuso un avance importante para las familias, ya que se pudo abordar la detección de portadores a partir del núcleo familiar y ofrecer el diagnóstico prenatal a las parejas de alto riesgo.

**Tabla 1. Incidencia de la Fibrosis Quística en Europa**

PAÍS	INCIDENCIA
Noruega	1/6500
Suecia	1/4000
Dinamarca	1/4700
Finlandia	1/25000
Francia	1/4000
Bretaña (Francia)	1/2913
Italia (Verona)	1/2730
Suiza	1/2000
Holanda	1/3600
Irlanda	1/1461
Reino Unido	1/2500

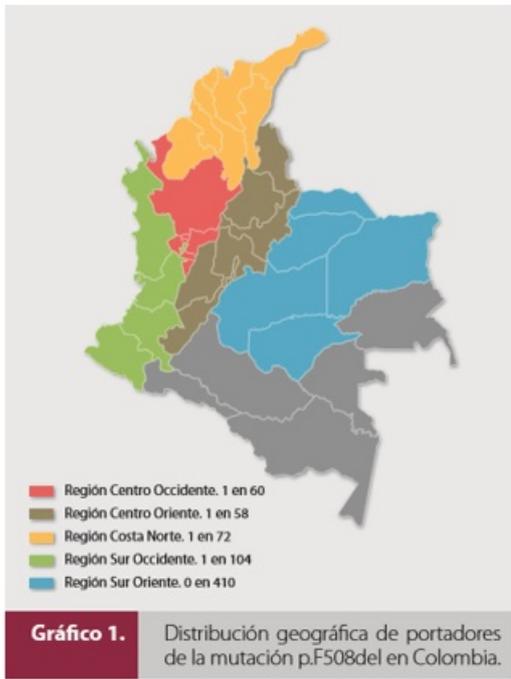
Entre los hijos de padre y madre heterocigóticos, se puede esperar que una cuarta parte de machos y hembras sean homocigóticos y se encuentren afectados. Sin embargo, en la genética humana, la heterocigosis del padre y la madre se descubre solamente cuando entre los hijos hay por lo menos uno de estos afectados. Si una persona afectada por un rasgo recesivo se casa con otra que sea homocigótica normal, ninguno de los hijos resultará afectados; pero todos ellos serán portadores heterocigóticos. Si un individuo afectado por un rasgo recesivo se casa con un heterocigótico portador del mismo gen recesivo, apróximadamente la mitad de los hijos se verán afectados y resultara un patrón genealógico semejante en apariencia al rasgo dominante. Para mantener el equilibrio entre los nuevos genes añadidos por nuevas mutaciones y los que se han

perdido por la muerte prematura de los homocigotos, se requiere un índice de mutación de 1 a 1500 gametos por generación.

En Canadá concretamente en Quebec, se reseña una incidencia de 1 por cada 891 individuos recién nacidos vivos en el caso de en Norte América 1/3.000 y su frecuencia en Latinoamérica oscila entre 1/3.862 y 2.000 recién nacidos en Cuba y de 1/9.000 en México; los datos de Uruguay donde la prevalencia es de 6.9 por cada 100.000 habitantes y en Argentina, según datos de los Registros Nacional y Provincial, se registran más de 1000 pacientes con FQ. A nivel provincial, los hospitales de niños y adultos tratan aproximadamente 250 pacientes (Facultad de Ciencias Exactas, UNLP Argentina); los primeros estudios realizados en la provincia de Buenos Aires informan datos preocupantes: en 52.024 muestras analizadas con el método de la tripsina inmunoreactiva, se detectaron 14 pacientes, con una incidencia en esta población de 1:3.468; de estos datos se infiere que la prevalencia de portadores sanos en la población general sería de 1:30.

En Colombia no se conoce el dato exacto de los niños que nacen afectados por la enfermedad; sin embargo, gracias a los resultados de esta y otras investigaciones, podemos decir que cerca de uno de cada cinco mil niños colombianos tendrán FQ (fibrosis quística). No obstante, resulta preocupante que en el país la FQ se diagnostica tardíamente, cuando ya se ha presentado el daño pulmonar y pancreático, razón por la cual causa gran mortalidad en los afectados. Al realizar comparaciones de los datos de la frecuencia de la mutación  $\Delta F508$  del gen en Colombia con otros países, los investigadores de la Universidad del Rosario no encuentran diferencias con Argentina, Chile o Uruguay, hecho sorprendente si se tiene en cuenta que esas son poblaciones con una fuerte influencia de ancestros europeos y deberían mostrar muchos más casos de personas portadoras de la mutación en uno de sus dos genes, razón por la cual se esperaban menos portadores en Colombia.

Dado que en Colombia es un país con bastante diversidad étnica en las diferentes regiones del país (**Grafico 1**), un estudio realizado por grupo de científicos de la Escuela de Medicina y Ciencias de la Salud de la Universidad del Rosario, luego de encontrar que la frecuencia de portadores de la mutación  $\Delta F508$  del (causante del 34,5 % de los casos de FQ) para el país es de un portador en cada ochenta y cuatro personas, (treinta y uno de los dos mil sesenta y ocho analizados)(**Tabla 1. Incidencia de la FQ en Colombia**).



Región	Número de personas	Portadores	Frecuencia de portadores	Incidencia estimada
<b>Centro-occidente</b> Antioquia, Caldas y Quindío	422	7	1/60	1/7.711
<b>Centro-oriental</b> Bogotá, Cundinamarca, Boyacá, Tolima y Santander	641	11	1/58	1/7.225
<b>Costa Norte</b> Atlántico y Magdalena	509	7	1/72	1/5.208
<b>Sur-occidente</b> Valle, Huila y Nariño	626	6	1/104	1/19.685
<b>Sur-oriental</b> Meta, Casanare y Guaviare	410	0	0	0
<b>Total</b>	2.608	31	1/84	1/5.025

**Tabla 1.** Incidencia estimada de FQ en la población colombiana. Frecuencia de portadores de la mutación F508del. Fuente: Grupo de Investigación GeniURos, Escuela de Medicina y Ciencias de la Salud, Universidad del Rosario.

Figura 4.1: Tabla 1. Incidencia de la FQ en Colombia

De acuerdo a un informe del Registro Colombiano de fibrosis quística del año 2009 de la Asociación Colombiana de Neumología Pediátrica, se analizaron un total de 178 pacientes de los cuales 97 han sido diagnosticados en la última década (55,2 %): 11 en el 2.009 (6,2 %), 14 en el 2.008 (7,9 %), 13 en el 2.007 (7,3 %), 8 en el 2.006 (4,5 %), 12 en el 2.005 (6,7 %), 10 en el 2.004 (5,6 %), 13 en el 2.003 (7,3 %), 11 en el 2.002 (6,2 %) y 5 en el 2.000 (2,8 %) y 5 en el 2.000 (2,8 %) (**Gráfico 2**).

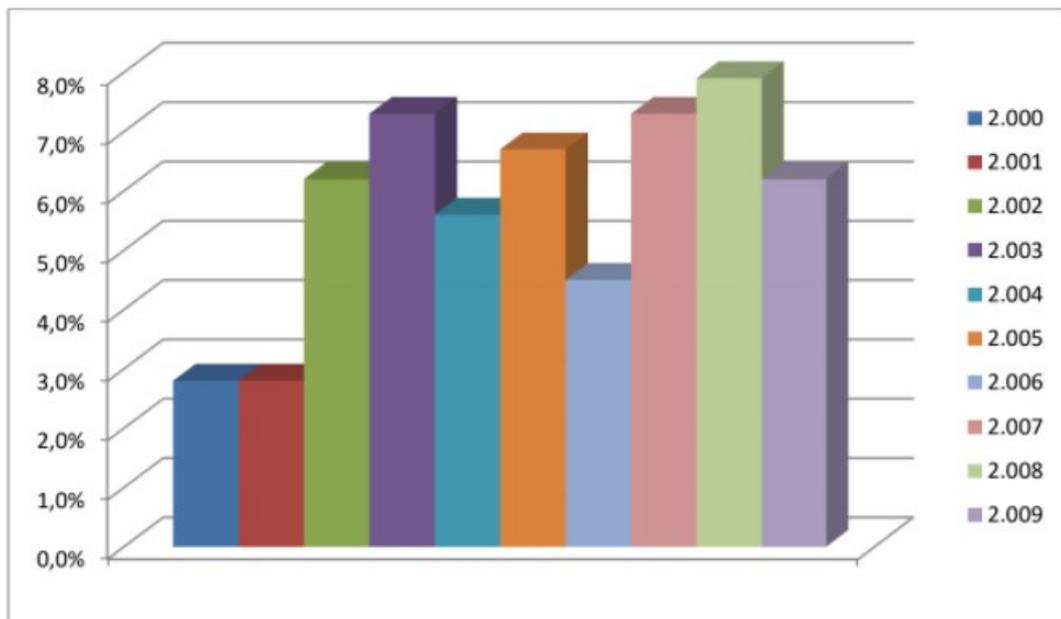


Figura 4.2: Gráfico 2

Teniendo en cuenta las proyecciones nacionales del DANE, es decir que en el año 2.009 habría una población total de 44.978.832 habitantes en el territorio colombiano, se estimaría una prevalencia de la enfermedad de 3,95 casos por cada 1.000.000 de habitantes (0,000395 %). Por otra parte, la tasa de incidencia para esta población sería de 2,44 casos nuevos por cada 10 millones de habitantes (que equivale a 11 casos nuevos diagnosticados en el 2.009 en la población analizada).

Con respecto al Huila y como en Neiva no se tiene unos estudios establecidos durante los últimos 9 años sobre la fibrosis quística, ya que es catalogada en nuestro país, departamento y municipio como una de las enfermedades huérfanas de las cuales son muy pocos son los casos que se presentan con respecto a los años; por lo cual, el sector educativo superior tratara técnicas y métodos que permitan estimar la evaluación de la enfermedad para así tener una eficacia contra la **FQ**.

La intención, por lo tanto, es adaptar el modelo matemático de Hardy - Weinberg para analizar el equilibrio de la población del Huila que padece la enfermedad de la fibrosis quística y utilizando los métodos estadísticos y probabilidad; que con lleve a la construcción de un sistema dinámico que modele la probabilidad de transmisión hereditaria, tanto en las familias y personas portadoras de la fibrosis quística, con el cual se apoyen las entidades de salud pública departamental puedan crear estrategias orientadas a tratar, prevenir y disminuir su incidencia de la enfermedad en el Huila.

## **4.1. EL GEN CFTR**

El gen **CFTR** (**cystic fibrosis transmembrane conductance regulator**) es una glucoproteína transportadora de membrana dependiente del AMPc, responsable de una de las vías de transporte de iones cloro en las células epiteliales, que también puede controlar la función de otras proteínas de membrana como los canales auxiliares de cloro y los canales de sodio. Desde principios de 1980, varios grupos de investigadores iniciaron estudios de ligamiento genético entre la FQ y distintos marcadores de DNA, con objeto de identificar el gen de la FQ.

Esta proteína se localiza en la membrana apical de muchos tipos de células epiteliales: células epiteliales de las vías aéreas, glándulas submucosas del tracto gastrointestinal, hígado, vesícula biliar y páncreas. La alteración de la proteína CFTR impide que pueda realizar su acción de transporte y el resultado final de todas las mutaciones detectadas que alteran la función de CFTR es el mismo: la imposibilidad de transportar cloruro. Esto explica la historia natural de la enfermedad en las glándulas sudoríparas, aparato respiratorio, páncreas, aparato genital masculino y sistema hepatobiliar.

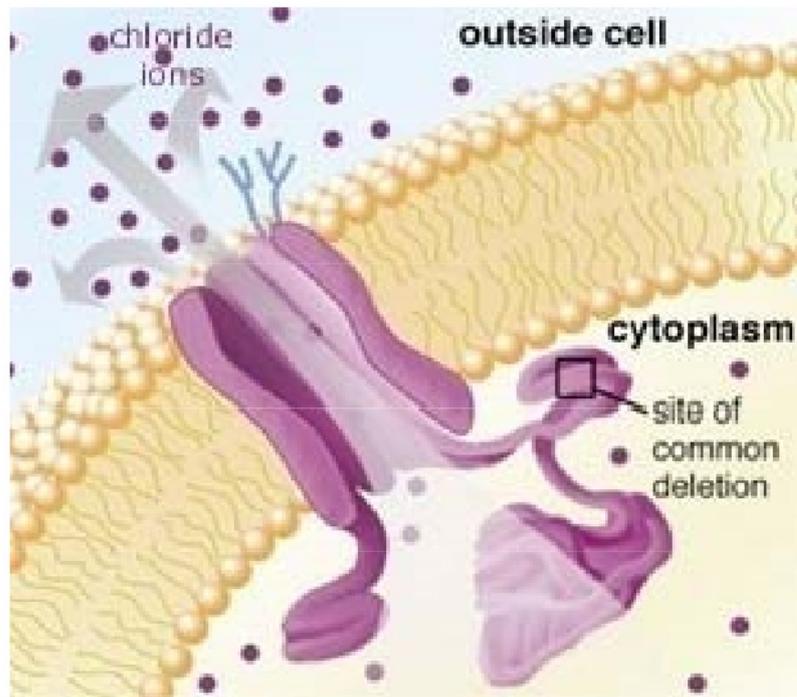


Figura 4.3: Función CFTR



Figura 4.4: Mapa

A finales de 1985, Tsui describieron el ligamiento genético entre la FQ y el marcador ADN DOCRI917, asignado luego al cromosoma 7, estudios que se ampliaron en años posteriores, descubriéndose nuevos marcadores (MET y D7S8,...) ligados al gen FQ, que permitieron la construcción de un mapa de ligamiento genético entre estos marcadores y la enfermedad por otros grupos de investigadores en distintas partes del mundo entre los años 1985 y 1989, de forma que el análisis de fragmentos de restricción polimórficos (RFLP) durante esos cuatro

años permitió determinar la existencia de un único locus, el gen de la FQ, en el brazo largo del cromosoma 7 humano, concretamente en la región 7q31 (Figura 4.5). Abarca 250kb y contiene 27 exones que codifican una proteína de 1480 aminoácidos y 168kDa.

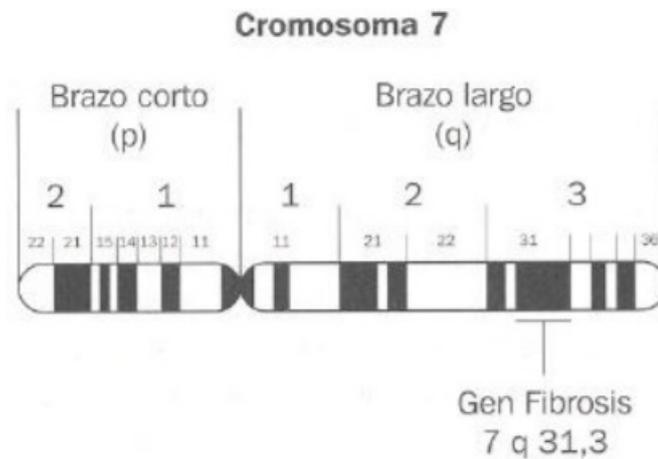


Figura 4.5: Localización cromosómica del gen CFTR (Modificado de Morales-Pérez P, Fernández-Soria VM18).

En agosto de 1989, en un estudio coordinado por Tsui en Toronto y por Collins en Michigan, se clonó el ADN comprendido en esta región, con estrategias de walking y jumping cromosómico para llegar a secuenciar el ADN del gen de la FQ, al que se denominó CFTR, al mismo tiempo que se logró caracterizar la mutación más frecuente de este gen, la mutación DF508.

La proteína CFTR está formada por cinco dominios: 2 dominios transmembrana (MSD), 2 dominios de unión al ATP(NBD) y 1 dominio regulador (R) con múltiples puntos de fosforilación.

El gen de la FQ fue el primer gen humano identificado sin conocer la proteína para la que codificaba, ni disponer de claves citogenéticas que permitiesen un avance rápido en su identificación, mediante la tecnología denominada «genética reversa», que consiste en la identificación y aislamiento de un gen determinado, sin tener datos sobre la función y características bioquímicas de la proteína para la que éste codifica.

#### 4.1.1. LA MUTACIÓN DF508

La mutación más frecuente es la deleción de tres nucleótidos que tiene como resultado la pérdida de una fenilalanina en la posición 508 (DF508). Esta mutación está presente en 70 % de los alelos FQ y en las poblaciones de origen europeo.

La extensión de esta mutación en la población europea se ha explicado por una ventaja selectiva de los individuos heterocigotos portadores de ella. Según esta hipótesis, los heterocigotos estarían más protegidos frente a la deshidratación provocada por las diarreas causadas por las enterotoxinas bacterianas. La proteína DF508, junto a una menor expresión de CFTR en las

superficies epiteliales, disminuirían la entrada de patógenos en el epitelio intestinal, protegiendo frente a estas infecciones.

#### **4.1.2. EL ESPECTRO DE MUTACIONES DEL GEN CFTR**

Una de las particularidades de CFTR es la presencia de transcritos que muestran la delección de uno o más exones entre los individuos normales. Estos transcritos se deben a alteraciones en el proceso de ajuste o splicing alternativo de la que la más frecuente es la delección del exón 9. La presencia o ausencia de este exón se correlaciona con la presencia de un polimorfismo de la secuencia del intrón 8 cercano al punto aceptor del splicing (5T, 7T ó 9T). Si se encuentra la forma 7T ó 9T se produciría un procesamiento normal en 90 % de las moléculas de RNAm, sin embargo la presencia de la forma 5T sólo garantiza el procesamiento normal de entre 10 y 40 % de las moléculas de RNAm.

#### **4.1.3. CORRELACIÓN GENOTIPO - FENOTIPO**

- La DF508 por sí sola suma el 66 % de las mutaciones, 40 % de los pacientes con FQ son heterocigotos mostrando la DF508 en uno de sus alelos y otra mutación en el otro alelo de CFTR.
- Generalmente, tanto la función pulmonar como la edad de aparición de la enfermedad y la cantidad de sales de cloruro se relacionan con el genotipo.
- Por otro lado, la diferente severidad y variedad de los síntomas dentro de la misma familia indican que el genotipo por sí sólo, a nivel de alteraciones en CFTR, no puede explicar completamente el fenotipo clínico.
- Únicamente la mutación A455E se asocia fuertemente con la función pulmonar.
- En cuanto a la función pancreática los pacientes con una o dos mutaciones con cambio de sentido conservan la función exocrina (SP), mientras que los pacientes que muestran los dos alelos con mutaciones de splicing, sin sentido o con cambio de marco de lectura y algunas con cambio de sentido muestran insuficiencia de esta función (IP). Los pacientes que muestran en un alelo una mutación causante de IP y en el otro una mutación causante de SP muestran un fenotipo SP. Con sólo una mutación SP la función CFTR es suficiente para mantener la función pancreática.
- El efecto de una mutación puede ser modificado por una segunda mutación en si sobre el mismo alelo.
- El polimorfismo de la zona de polipirimidinas del intrón 8 modifica la penetrancia de algunas mutaciones.

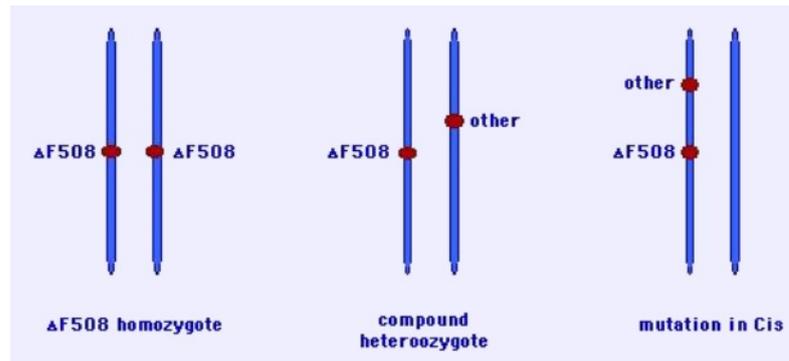


Figura 4.6: Mutaciones

#### 4.1.4. LA PROTEÍNA CFTR Y SUS FUNCIONES

##### Estructura de la proteína CFTR

- La proteína CFTR está formada por 1480 aminoácidos.
- Esta proteína está formada por dos motivos repetidos compuestos cada uno por un dominio hidrofílico que atraviesa la membrana (MSD, membrane-spanning domain) que contiene seis hélices y una región hidrofílica importante de unión a ATP (NBF, nucleotide binding fold). Ambos motivos se encuentran unidos por un dominio citoplasmático (dominio R) codificado por el exón 13 que contiene varios residuos cargados y la mayor parte de los lugares de fosforilación (probablemente sustratos de las proteínquinasas A y/o C).

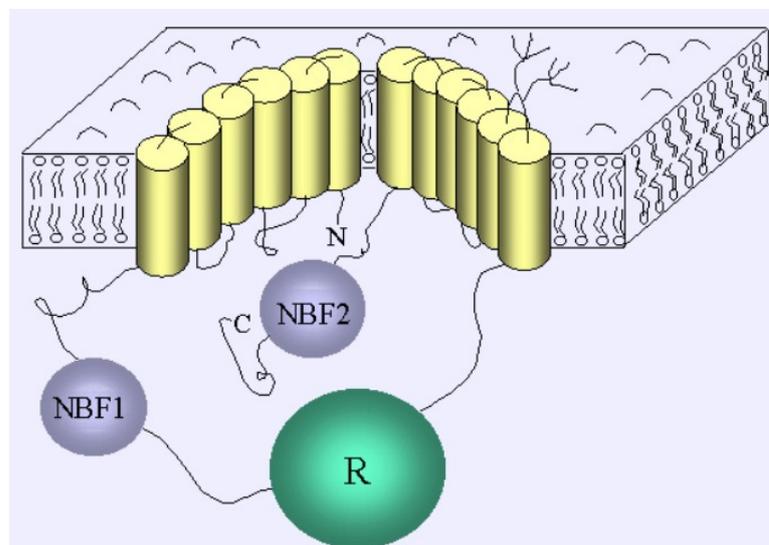


Figura 4.7: Estructura propuesta para la proteína CFTR, por Riordan et al., 1989 - Pascale Fanen.

- Existe homología entre la estructura primaria de la proteína CFTR y miembros de familias de proteínas de membrana, como la familia de los transportadores ABC (ATP-binding cassette), que son responsables del transporte activo de sustratos a través de la membrana celular, en el que la hidrólisis del ATP es la fuente de energía.

## 4.1.5. FUNCIONES DE LA PROTEÍNA

### Función del canal de Cl

Las primeras hipótesis sobre la posible función de la proteína CFTR tuvieron en cuenta dos posibilidades.

- La primera hipótesis postulaba que la proteína CFTR era un canal de Cl. Esta hipótesis era compatible con los defectos que se observaban en la permeabilidad de los iones Cl en las membranas apicales epiteliales de los pacientes con FQ.
- La segunda hipótesis proponía que la proteína CFTR no fuera un canal de iones sino que jugara un papel en la regulación de los canales de Cl ya sea por asociación con éstos o por medio del transporte de algún factor regulador de los canales de Cl dentro o fuera de la célula.

Finalmente quedó demostrada la primera de las hipótesis, que la función de CFTR era la de canal de Cl.

### La proteína CFTR, una proteína multifuncional

Los descubridores del gen CFTR lo denominaron regulador de la conductancia transmembrana" (Cystic Fibrosis Transmembrane conductance Regulator). De hecho, la proteína CFTR regula también otros canales, como el canal de Cl<sup>-</sup> ORCC (Outwardly Rectifying Chloride Channel), el canal de Na<sup>+</sup> epithelial (ENaC, Epithelial Na<sup>+</sup> Channel) y al menos los dos canales de K<sup>+</sup> ROMK1 y ROMK2. Además de un regulador de canales, también juega un papel en el transporte de ATP, modificando la exocitosis y endocitosis, y en la regulación del pH de los orgánulos intracelulares.

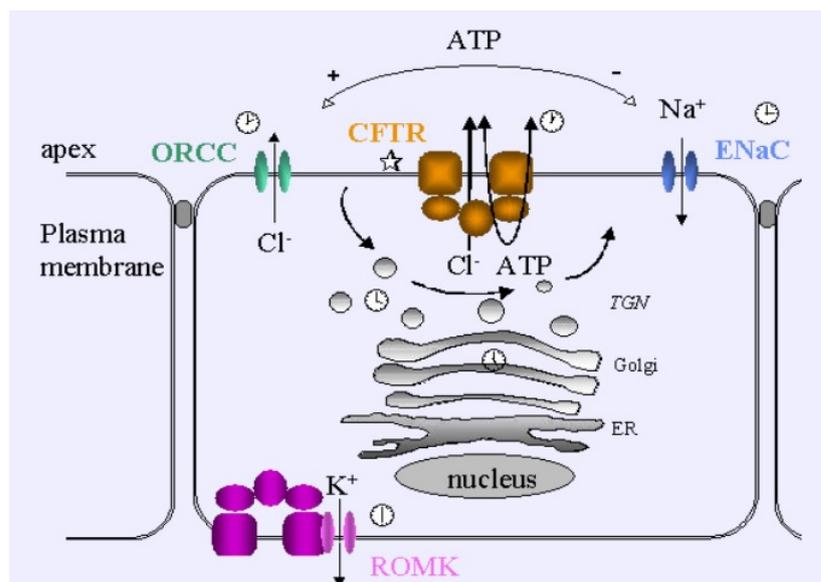


Figura 4.8: CFTR, una proteína multifuncional, por Schwiebert et, 1999 - Pascale Fanen.

## Correlación de las mutaciones en el gen CFTR con la función del canal de Cl

Las alteraciones moleculares tiene diversos efectos sobre la proteína CFTR y sus funciones. Welsh y Smith han propuesto la clasificación de estas alteraciones en relación a la función del canal de Cl (Welsh & Smith, 1993) (Figura 4.9).

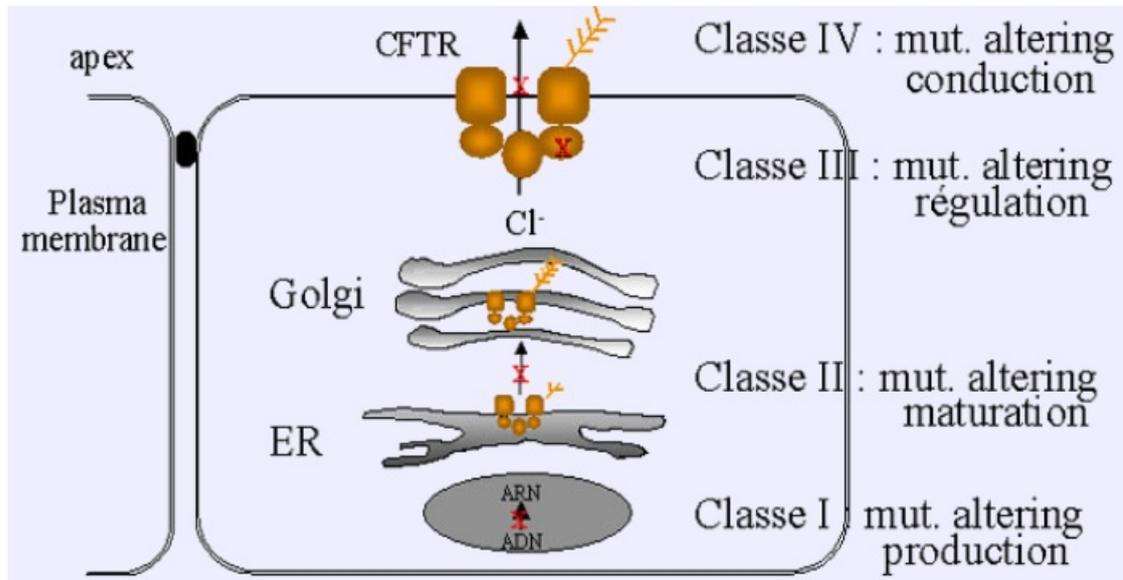


Figura 4.9: Clasificación de las mutaciones del gen CFTR, por Welsh & Smith, 1993 - Pascale Fanen

- **Clase 1: mutaciones que alteran la producción de la proteína** Estas mutaciones tienen como consecuencia la ausencia total o parcial de la proteína. En esta clase se incluyen las mutaciones sin sentido y aquellas que conducen a la formación prematura de un codón de parada de la traducción (algunas de las cuales producen alteraciones del splicing mientras que otras producen un cambio en el marco de lectura). En algunos casos la no producción de la proteína se debe a que el RNAm formado es inestable. En otros casos, aunque se produce la proteína, ésta es inestable y se degrada rápidamente. Funcionalmente todas estas mutaciones se caracterizan por una pérdida de la conductancia del canal de Cl en el epitelio afectado.
- **Clase 2: mutaciones que alteran el proceso de maduración celular de la proteína** Ciertas mutaciones alteran la maduración de la proteína y su transporte a la membrana plasmática. Ello conduce a que esta proteína no esté presente en la membrana o que lo esté en muy baja cantidad. Las mutaciones de esta clase son las más frecuentes (DF508).
- **Clase 3: mutaciones que alteran la regulación del canal de Cl** Estas mutaciones se encuentran frecuentemente en el dominio de unión al ATP.
- **Clase 4: mutaciones que alteran la conducción a través del canal de Cl** Ciertos segmentos de los dominios que atraviesan la membrana participan en la formación de un poro

de iones. Las mutaciones con cambio de sentido situadas en estas regiones generan una proteína correctamente posicionada con actividad de canal de  $Cl^-$  cAMP dependiente. Sin embargo sus características son distintas a las del canal CFTR nativo ya que muestran una disminución del flujo de iones y una selectividad modificada.

- **Clase 5: mutaciones que alteran la estabilidad del RNAm**
- **Clase 6: mutaciones que alteran la estabilidad de la proteína madura**

### Secreción defectuosa de $Cl^-$ y absorción de $Na^+$ aumentada en FQ

La forma de interacción de CFTR con ENaC y otros canales iónicos aún es desconocida y continúa bajo investigación. Dado que los defectos en CFTR llevan a una regulación defectuosa del canal ENaC, la FQ se caracteriza por alteraciones en la secreción y absorción de electrolitos. En relación con la enfermedad pulmonar de riesgo vital en pacientes con FQ, se han descrito dos mecanismos fisiopatológicos importantes: 1) la secreción alterada de  $Cl^-$ , la cual se localiza de preferencia en las glándulas submucosas y 2) la absorción aumentada de Na y la consecuente hiperabsorción de electrolitos, la cual ocurre en el epitelio superficial figura (4.10).

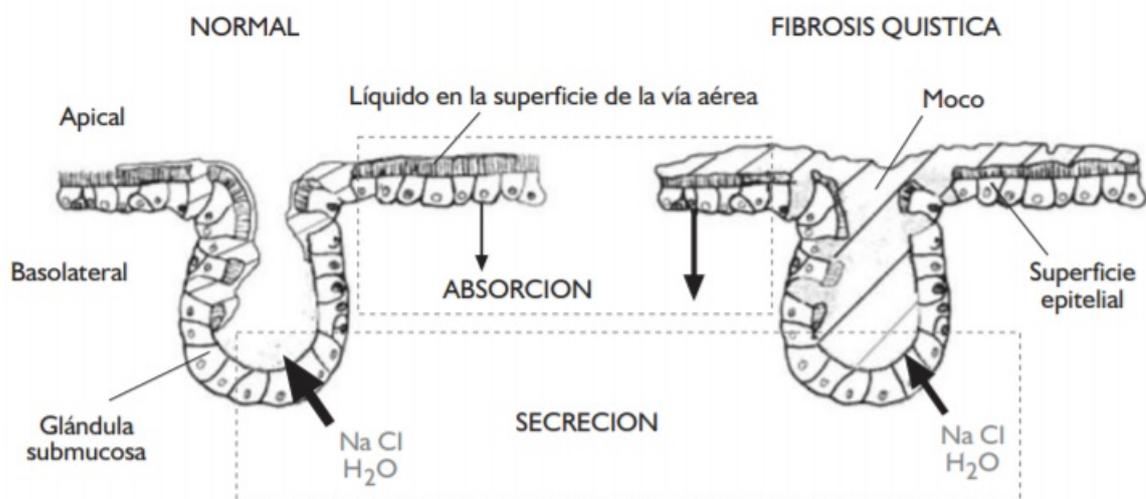


Figura 4.10: Actividad de la proteína CFTR en la membrana celular del humano

En las vías aéreas normales, el transporte iónico cambia desde una absorción neta bajo condiciones de reposo a una secreción cuando se expone a secretagogos. Sin embargo, las vías aéreas en pacientes con FQ son incapaces de aumentar su transporte secretor con estimulación. Por el contrario, la conductancia de  $Na^+$  epitelial aumentada en membranas apicales, en forma paralela a una permeabilidad paracelular para  $Cl^-$  y la elevada permeabilidad al agua no permite gradientes osmóticos transepiteliales mayores, lo que lleva a hiperabsorción en los epitelios FQ. Por lo tanto, la capa de líquido superficial de la vía aérea (ASL) disminuye, las glándulas mucosas ya no son despejadas del mucus y el clearance mucociliar se ve fuertemente alterado en las vías aéreas con FQ.

## 4.2. CLASIFICACIÓN DE LA FIBROSIS QUISTICA Y SUS MANIFESTACIONES

### FIBROSIS QUISTICA CON MANIFESTACIONES PULMONARES

Las manifestaciones respiratorias de la fibrosis quística, las infecciones respiratorias con sus reagudizaciones, así como las complicaciones pulmonares de las mismas, son la principal causa de la morbi-mortalidad en la FQ. La presencia de una exacerbación pulmonar viene definida por los siguientes criterios:

- Aumento de la tos y/o dificultad respiratoria, menos tolerancia al ejercicio.
- Cambios en el aspecto o en el volumen de la expectoración.
- Nuevos signos a la auscultación.
- Nuevos hallazgos radiológicos.
- Pérdida del apetito.
- Fiebre y disminución de la función pulmonar.

**La bronquiolitis** es una infección viral común durante los dos primeros años de la vida, siendo, en ocasiones, en los países donde no se realiza el cribaje neonatal de la FQ, el momento donde, por su evolución más tórpida de lo habitual, se hace el diagnóstico clínico de la enfermedad.

Las sobreinfecciones, sobre todo por enterobacteriáceas o por *Stafilococcus aureus*, se observan a menudo en el curso de una bronquiolitis VRS positiva, siendo difícil de establecer si la sobreinfección bacteriana que padecen estos niños es primaria o secundaria. Además, las infecciones virales parecen tener un papel en la reactividad de gérmenes como la *Pseudomonas aeruginosa*, colonizadora habitual de las vías aéreas de estos pacientes.

**Insuficiencia respiratoria** Los pacientes afectados de FQ desarrollan un fallo respiratorio hipoxémico, lo que conduce a una hipertensión pulmonar y a un cor pulmonale. Al mismo tiempo, la obstrucción de la vía aérea disminuye la capacidad vital, el volumen tidal y la ventilación alveolar, con un progresivo aumento del CO<sub>2</sub> y un fallo respiratorio.

## FIBROSIS QUISTICA CON MANIFESTACIONES INTESTINALES

Las manifestaciones gastrointestinales de la FQ acorde a la edad son:

- Niños menores de 24 meses: Recién nacido con antecedente de Ileo meconial, colestasis neonatal, deshidratación sin causa aparente, hiponatremia, hipoproteinemia, anemia, diátesis hemorrágica, desnutrición, esteatorrea, neumatosis intestinal y mal incremento ponderal.
- 2-12 años: diarrea crónica, malabsorción, reflujo gastroesofágico, invaginación recurrente, vólvulos, prolapso rectal, enfermedad hepática de causa desconocida, hipocratismo digital y enlentecimiento del crecimiento.
- Mayor de 12 años: hipocratismo digital, hipertensión portal, diarrea crónica, obstrucción intestinal distal, pancreatitis recurrente, hepatomegalia, cirrosis biliar primaria, colelitiasis y desnutrición.

La gran mayoría de los pacientes presentan insuficiencia pancreática (85 %), es decir, requieren suplemento de enzimas pancreáticas. Para evaluar la necesidad de este tratamiento se requiere de una historia clínica adecuada consignando los síntomas pertinentes como: detención del crecimiento, diarreas, esteatorrea, flatulencia excesiva, distensión abdominal, dolor abdominal, etc. Además, nos podemos ayudar de algunas pruebas diagnósticas como determinación de grasa fecal (Van der kammer, esteatocrito, sudán III) y pruebas de malabsorción (carotinemias, D-xilosemia a la hora).

## Capítulo 5

---

# MARCO TEÓRICO

---

Hay ciertas relaciones de las matemáticas con otras ciencias exactas que sirven de colaboración para el desarrollo de grandes temas de investigación, entre esas ciencias la más destacada es la Estadística, la cual se encarga del estudio de la recolección de datos (muestras) y el análisis e interpretación de la información, ya sea para ayudar a la toma de mejores decisiones.

De ellas abarcaremos temas de gran relevancia que abarcan estas ramas la relación que existe entre el binomio de Newton, el triángulo de pascal, el análisis combinatorio, la aleatoriedad las probabilidades, diseños de experimentos y los procesos estocásticos.

Además, este trabajo para el desarrollo de la investigación de la enfermedad la Fibrosis quística, otra área de las que utilizaremos será el estudio de la genética abarcando como objeto principal las leyes de Mendel, la mutación genética y el equilibrio de HARDY - WEINBERG.

### 5.1. BINOMIO DE NEWTON

El teorema del binomio fue descubierto en el año 1665, fue notificado por primera vez en dos cartas que fueron enviadas por el funcionario y administrativa de la Royal Society, Henry Oldenburg en el año 1676. La primera carta fue fechada el 13 de junio de 1676, en respuesta a un pedido del filósofo, jurista y matemático alemán Gottfried Wilhelm von Leibniz, quien quería tener conocimiento de las labores e investigaciones de matemáticos británicos sobre series infinitas. Por lo cual Newton envía el enunciado de su teorema y un ejemplo ilustrativo. Leibniz responde, en una carta fechada el 17 de agosto de 1676, que se encuentra ante una técnica general que le permite obtener distintos resultados sobre las cuadraturas, las series, etc; donde denomina algunas de sus ramificaciones por las investigaciones de Leibniz; Newton responde también con una carta en la que detalla cómo ha descubierto la serie de binomios.

A partir de este hallazgo Newton intuyó que era posible operar con series infinitas del mismo modo que con expresiones polinómicas finitas. Newton no se encargó de publicar jamás el teorema del binomio. Lo hizo el matemático británico, John Wallis en el año 1685 en su Algebra, en la cual atribuyó a Newton el gran hallazgo; Newton desarrolló la fórmula para así proceder al cálculo de las potencias de un binomio usando para esto números combinatorios..

El teorema del binomio, también llamado binomio de Newton, corresponde a un polinomio que se encuentra formado por dos términos y expresa la enésima potencia de un binomio como un polinomio. El desarrollo del binomio  $(a + b)^n$  posee singular importancia, ya que aparece con mucha frecuencia en Matemáticas y posee diversas aplicaciones en otras áreas del conocimiento.

### FÓRMULA GENERAL DEL BINOMIO

Sea el binomio de la forma  $(a + b)$ , si a este binomio se le multiplica sucesivamente por si mismo se obtienen las siguientes potencias:

$$\begin{aligned} (a + b)^1 &= a + b \\ (a + b)^2 &= \underbrace{(a + b)(a + b)}_{2 \text{ veces}} = a^2 + 2ab + b^2 \\ (a + b)^3 &= \underbrace{(a + b)(a + b)(a + b)}_{3 \text{ veces}} = a^3 + 3a^2b + 3ab^2 + b^3 \\ (a + b)^4 &= \underbrace{(a + b) \cdots (a + b)}_{4 \text{ veces}} = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \\ (a + b)^5 &= \underbrace{(a + b) \cdots (a + b)}_{5 \text{ veces}} = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5 \end{aligned}$$

para n- veces

$$(a + b)^n = \binom{n}{0} a^n + \binom{n}{1} a^{n-1}b + \binom{n}{2} a^{n-2}b^2 + \cdots + \binom{n}{n-1} ab^{n-1} + \binom{n}{n} b^n$$

(En el caso en que en el binomio figure un signo menos, los signos del desarrollo deben irse alternando de la forma  $+ - + - + \cdots$ )

La fórmula generalizada del binomio de Newton:

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$$

Siendo  $\binom{n}{k}$  los coeficientes binomiales.

## 5.2. TRIÁNGULO DE PASCAL

El triángulo de Pascal ha jugado un papel relevante en diferentes momentos del desarrollo cultural; fue símbolo místico y vitalista para los antiguos chinos y griegos, un motivo para la recreación matemática. Pascal ideó una manera sencilla de calcular números combinatorios (aunque en algunos textos esta idea se atribuye a Tartaglia):

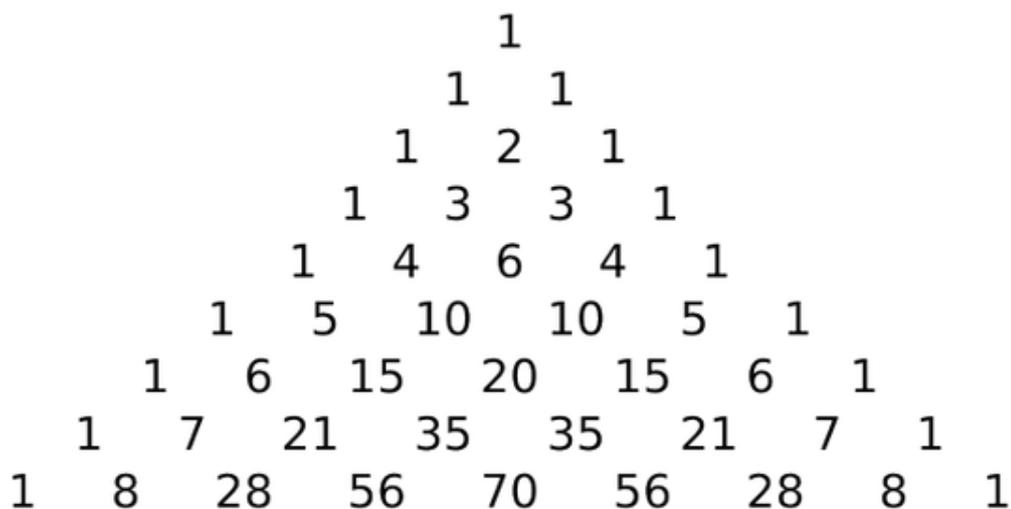
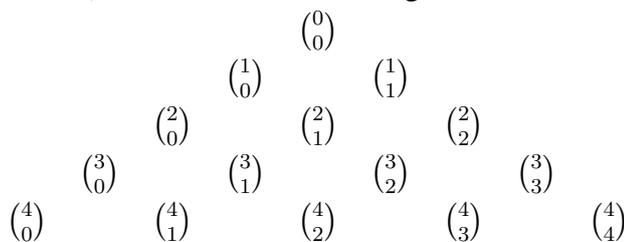


Figura 5.1: Triángulo de Pascal

Los números del triángulo de Pascal coinciden con los números combinatorios. El número combinatorio  $C_n^m$  ( $n$  sobre  $m$ ) se encuentra en el triángulo en la fila  $n + 1$ , en el lugar  $m + 1$ .



Podemos saber que el número de parejas posibles que decíamos antes es 6 si miramos el tercer número de la quinta fila. Esto hace que el triángulo sea útil como representación de estos números, y proporciona una buena forma de intuir sus propiedades.

Por el contrario, a la fórmula de los números combinatorios se le puede dar el carácter de fórmula general del triángulo para saber, sin necesidad de construir todas las filas anteriores, cuál es el número que ocupa un lugar determinado:

$$C_n^m = \binom{m}{n} \frac{m!}{n!(m-n)!}$$

## 5.3. PROBABILIDAD Y ESTADÍSTICA

**La Teoría de las probabilidades:** es la parte de las matemáticas que trata de los modelos adecuados para las situaciones en las que intervienen el Azar; ante un fenómeno aleatorio esta teoría plantea un modelo cuantitativo (aleatorio), a partir del cual se podrán hacer predicciones en la medida en que se concuerde o se contradigan la realidad observada y esto servirán para *validarlo*. Si el modelo depende de uno o varios parámetros, se pueden plantear una estimación y el contraste de hipótesis para estos parámetros, utilizando *métodos estadísticos*.

**Definición de Probabilidades:** Dado un *suceso*  $A$  de un espacio muestral  $\Omega$  asociado a un fenómeno aleatorio, podemos estar interesados en algún indicador del "grado de confianza" que tenemos en que realizar una observación del fenómeno observemos precisamente un resultado que este en  $A$  (diremos que el suceso  $A$  se ha realizado").

**Definición de Probabilidad Clásica:** Supongamos que  $\Omega$  es finito y hay simetría entre los sucesos elementales (esto es, que no hay ninguno que podamos pensar que esté especialmente favorecido o perjudicado respecto a los demás, en cuanto al hecho de poder ser observado). entonces, dado un suceso  $A$  de  $\Omega$ , se define su probabilidad como

$$P(A) = \frac{\#A}{\#\Omega} = \frac{\text{"casos favorables"}}{\text{"casos posibles"}}$$

El grado de confianza que tenemos en que se realice  $A$  habrá de ser proporcional al número de sucesos elementales que contiene, que son *los casos favorables*, bajo la hipótesis de finitud y simetría en  $\Omega$ . Para normalizar se divide por  $\#\Omega$ , que son *los casos posibles* y esta confianza se define como la  $P(A)$ .

Algunas propiedades de las probabilidades:

1.  $0 \leq P(A) \leq 1$  (ya que  $0 \leq \#A \leq \#\Omega$ )
2.  $P(\Omega) = 1$
3.  $P(\emptyset) = 0$
4.  $P(A \cup B) = P(A) + P(B)$  si  $A$  y  $B$  son conjuntos disyuntos.

En general para un número finito de sucesos,  $P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)$  si  $A_1, \dots, A_n$  son disyuntos dos a dos, (es decir,  $A_i \cap A_j = \emptyset$  para todo  $i \neq j$ ).

**Definición de Probabilidad Total:**

Si  $B_1, \dots, B_n$  forman una partición de  $\Omega$  (su unión es todo  $\Omega$  y son disyuntos dos a dos) y sus probabilidades no son ceros (para poder condicionar) entonces

$$P(A) = P(A/B_1)P(B_1) + \dots + P(A/B_n)P(B_n)$$

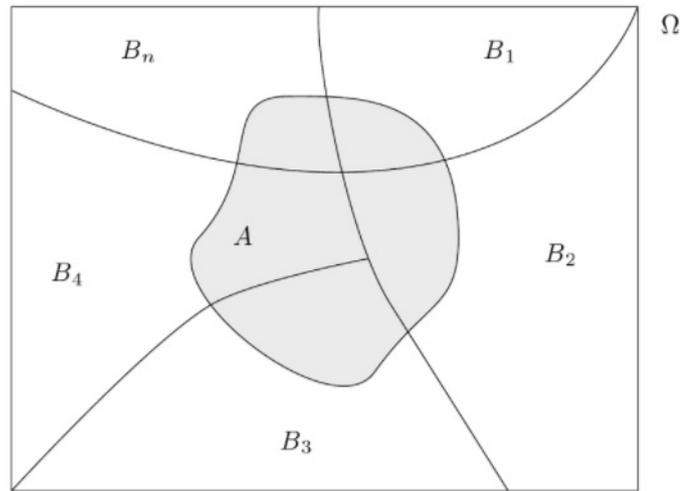


Figura 5.2: Probabilidad Total

### PROBABILIDAD CONDICIONADA Ó FORMULA DE BAYES

Si  $B_1, \dots, B_n$  forman una partición de  $\Omega$  esto es su unión es todo  $\Omega$  y son disyuntos dos a dos) y sus probabilidades no son cero(para poder condicionar), entonces si  $A \subset \Omega$  con  $P(A) > 0$ , se tiene que para cualquier valor de  $i = 1, \dots, n$ .

$$P(B_i/A) = \frac{P(A/B_i) P(B_i)}{P(A)}$$

Luego se utiliza la probabilidad total para expresar la probabilidad del denominador,  $P(A)$  de la ecuación anterior y se entra a reemplazar el valor,

$$P(B_i/A) = \frac{P(A/B_i) P(B_i)}{P(A/B_1) P(B_1), \dots, P(A/B_n) P(B_n)}$$

Este teorema juega un papel fundamental en una rama de la Estadística moderna llamada la Estadística Bayesiana.

### ANÁLISIS COMBINATORIO

En muchos casos el número de puntos muestrales en un espacio muestral no es muy grande y así la enumeración o cuenta directa de los puntos del muestreo necesarios para obtener las probabilidades no es difícil. Sin embargo, surgen problemas cuando la cuenta directa se convierte en una imposibilidad práctica. En tales casos se emplea el análisis combinatorio, que podría llamarse una forma sofisticada de contar.

## 5.4. PROCESOS DE MARKOV O CADENAS DE MARKOV

**Definición 5.1** Una *cadena de Markov* es una sucesión de ensayos similares u observaciones en la cual cada ensayo tiene el mismo número finito de resultados posibles, y en donde también la probabilidad de cada resultado para un ensayo dado depende sólo del resultado del ensayo inmediatamente precedente y no de cualesquier resultado previo.

**Definición 5.2** Consideremos un proceso de Markov en que el sistema posee  $n$  estados posibles, dados por los números  $1, 2, 3, \dots, n$ . Denotemos con  $p_{ij}$  la probabilidad de que el sistema pase al estado  $j$  después de cualquier ensayo en donde su estado era  $i$  antes del ensayo. Los números  $p_{ij}$  se denominan las probabilidades de transición y la matriz  $n \times n P = [p_{ij}]$  se conoce por **matriz de transición** del sistema.

**Teorema 5.4.1** Si  $P$  denota la matriz de transición de una cadena de markov y  $A_k$  es la matriz del estado despues de  $K$  ensayos, entonces la matriz  $A_{k+1}$  después del ensayo siguiente está dada por

$$A_{k+1} = A_k P$$

**Teorema 5.4.2** Se dice que una matriz de trasiación  $P$  es **regular** si para algun entero positivo  $k$  la matriz  $P^k$  no tiene elementos iguales a cero. Si  $P$  es una matriz de transición regular, entonces sin importar la matriz de estado inicial, las matrices de estados sucesivas se aproximan a alguna matriz de estado fija  $B$  en donde  $BP = B$ . La matriz  $B$  se denomina matriz estacionaria (**o de estado estable**) del sistema.  $A_{k+1}$

Es un proceso aleatorio  $x(t, \omega)$  esta definido como un proceso markoviano si para cualquier secuencia ordenada de tiempos  $\{t_1, t_2, \dots, t_k\}$ ,  $t_1 < t_2 < \dots < t_k$  y cualquier  $K$  entero se da la siguiente propiedad:

$$P_{ij} = \text{Prob}\{X_{t+1} = j \mid X_t = i, X_{t-1} = i_{t-1}, X_{t-2} = i_{t-2}, \dots, X_1 = i_1\}$$

Se tiene que  $X_t = i$ , significa que el proceso se encuentra en el estado  $i$  en un instante  $t$  y de la misma manera para  $X_{t+1} = j$ . Ahora, para  $P_{ij}$  como la probabilidad de pasar del estado  $i$  al estado  $j$  en  $n$  etapa o en  $n$  transiciones.

$$P_{ij}(n) = P_{ij}^{(n)} = P_{ij} = \text{Prob}\{X_{t+1} = j \mid X_t = i\}$$

Con la matriz de transición de estados

$$\text{Prob} = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1S} \\ P_{21} & P_{22} & \dots & P_{2S} \\ \vdots & \vdots & \vdots & \vdots \\ P_{S1} & P_{S2} & \dots & P_{SS} \end{bmatrix}$$

Se dan las siguientes condiciones

1.  $P_{ij} \geq 0$  donde para todo  $i, j = 1, 2, \dots, S$
2. Para todo  $i, j$ ;  $\sum_{j=1}^S P_{ij} = 1$

## 5.5. LEYES MENDELIANAS

La **genética mendeliana (leyes de Mendel)** son un conjunto de reglas básicas sobre la transmisión hereditaria de los caracteres de cada especie, que se realiza exclusivamente mediante gametos. Estas reglas básicas de herencia constituyen el fundamento de la genética. Las leyes se derivan del trabajo realizado por **Gregor Mendel** publicado en el año 1865 y en 1866, aunque fue ignorado por mucho tiempo hasta su redescubrimiento en 1900.

En primer lugar, determino que proporción de la descendencia de las plantas originales heredaban los distintos caracteres, y luego siguió el desarrollo de dicha proporción a lo largo de generaciones sucesivas. Además, estableció el análisis de solamente un par de caracteres homólogos.

Las tres leyes de Mendel explican y predicen cómo van a ser los caracteres físicos (fenotipo) de un nuevo individuo. Frecuentemente se han descrito como «leyes para explicar la transmisión de caracteres» (herencia genética) a la descendencia. Desde este punto de vista, de transmisión de caracteres, estrictamente hablando no correspondería considerar la primera ley de Mendel (Ley de la uniformidad).

### 5.5.1. 1ª LEY DE MENDEL: Ley de la uniformidad de los híbridos de la primera generación filial.

Establece que si se cruzan dos razas puras (una con genotipo dominante y otra con genotipo recesivo) para un determinado carácter, los descendientes de la primera generación serán todos iguales entre sí fenotípicamente y genotípicamente, e iguales fenotípicamente a uno de los progenitores (de genotipo dominante), independientemente de la dirección del cruzamiento. Expresado con letras mayúsculas las dominantes (A = amarillo) y minúsculas las recesivas (a = verde), se representaría así: AA + aa = Aa, Aa, Aa, Aa. En pocas palabras, existen factores para cada carácter los cuales se separan cuando se forman los gametos y se vuelven a unir cuando ocurre la fecundación.



dos durante la producción de gametos mediante una división celular meiótica. Esto significa que cada gameto va a contener un solo alelo para cada gen. Lo cual permite que los alelos materno y paterno se combinen en el descendiente, asegurando la variación.

Para cada característica, un organismo hereda dos alelos, uno de cada pariente. Esto significa que en las células somáticas, un alelo proviene de la madre y otro del padre. Éstos pueden ser homocigotos o heterocigotos.

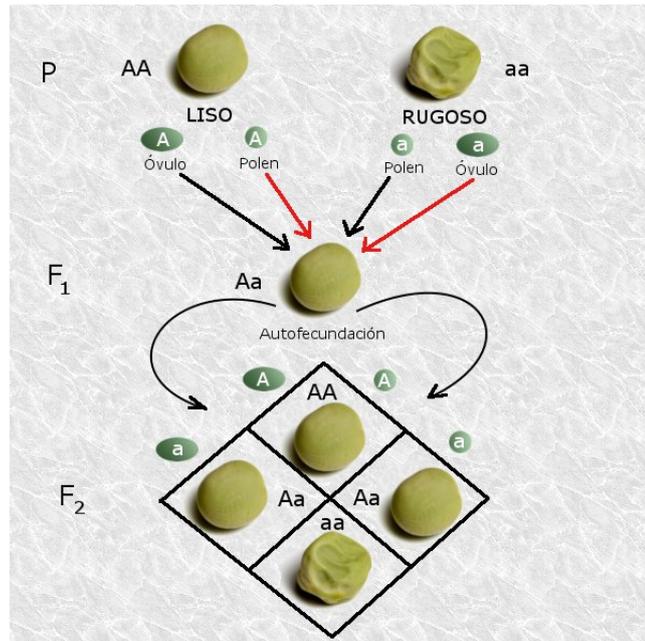


Figura 5.5: 2ª Ley de Mendel

Así, los caracteres recesivos que, como consecuencia de un cruzamiento entre dos líneas puras resultan enmascarados en la  $F_1$ , reaparecen en la  $F_2$  (la segunda generación filial) una proporción del 25 %.

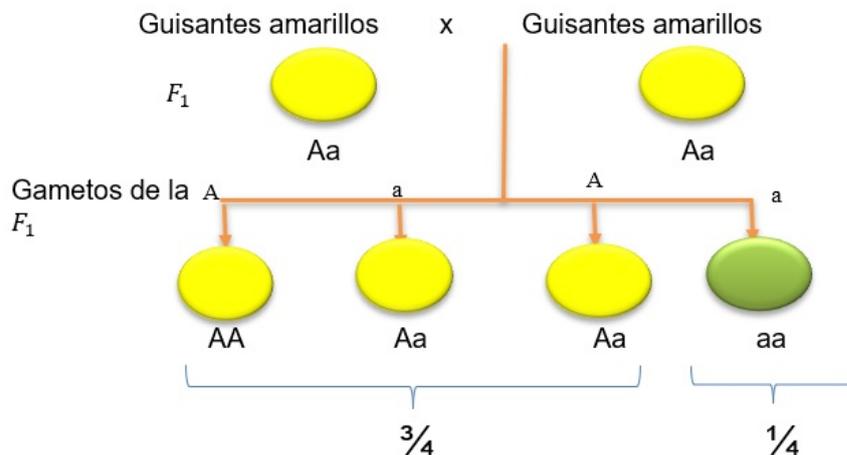


Figura 5.6: La segunda generación filial

### 5.5.3. 3ª LEY DE MENDEL: Ley de la independencia de los caracteres hereditarios.

En ocasiones es descrita como la 2ª Ley, en caso de considerar solo dos leyes (criterio basado en que Mendel solo estudió la transmisión de factores hereditarios y no su dominancia/expresividad). Mendel concluyó que diferentes rasgos son heredados independientemente unos de otros, no existe relación entre ellos, por lo tanto el patrón de herencia de un rasgo no afectará al patrón de herencia de otro. Sólo se cumple en aquellos genes que no están ligados (es decir, que están en diferentes cromosomas) o que están en regiones muy separadas del mismo cromosoma. En este caso la descendencia sigue las proporciones. Representándolo con letras, de padres con dos características AALL y aall (donde cada letra representa una característica y la dominancia por la mayúscula o minúscula), por entrecruzamiento de razas puras (1era Ley), aplicada a dos rasgos, resultarían los siguientes gametos: AL + al =AL,Al, aL, al. Al intercambiar entre estos cuatro gametos, se obtiene la proporción AALL, AALl, AaLL, AaLl, AAIL, AAIL, AaLL, AaLl, AaLL, AaLl, aALL, aALL, aAIL, aAIL, aaLL, aaLl, aall. Como conclusión tenemos: 9 con (A) y (L) dominantes,

ZIGOTO (16)	GENOTIPO (9)	FENOTIPO (4)	PROPORCIÓN	TOTAL	
AABB	AABB	verde-liso	1/16	9/16	
AABb1	AABb	verde-liso	2/16		
AABb2					
AaBB1	AaBB	verde-liso	2/16		
AaBB2					
AaBb1	AaBb	verde-liso	4/16	3/16	
AaBb2					
AaBb3					
AaBb4					
AAbb	AAbb	verde-rugoso	1/16		
Aabb1	Aabb	verde-rugoso	2/16		
Aabb2					
aaBB	aaBB	amarillo-liso	1/16		3/16
aaBb1	aaBb	amarillo-liso	2/16		
aaBb2					
aabb	aabb	amarillo-liso	1/16		1/16

Figura 5.7: 3ª Ley de Mendel

3 con (a) y (L), 3 con (A) y (l) y 1 con genes recesivos (aall); en palabras del propio Mendel: por tanto, no hay duda de que a todos los caracteres que intervinieron en los experimentos se aplica el principio de que la descendencia de los híbridos en que se combinan varios caracteres esenciales diferentes, presenta los términos de una serie de combinaciones, que resulta de la reunión de las series de desarrollo de cada pareja de caracteres diferenciales.

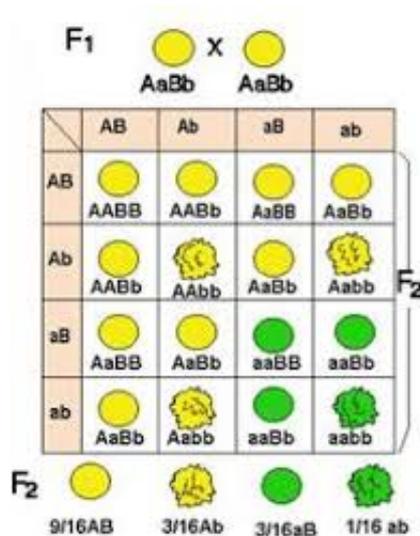


Figura 5.8: tabla 3<sup>a</sup> Ley de Mendel

## 5.6. FIBROSIS QUÍSTICA (FQ)

La fibrosis quística (FQ) es una enfermedad multisistémica y de herencia Mendeliana autosómica recesiva con un riesgo de recurrencia en cada gestación del 25 % (Homocigotas). Los portadores (Heterocigotos) suman un 50 % y 25 % son sanos. Constituye una enfermedad hereditaria letal más frecuente en la raza blanca y más común entre la raza caucásica. Se produce a partir de una mutación del gen, que codifica la proteína reguladora de la conductancia transmembrana o **CFTR**, ubicado en el brazo largo del cromosoma 7 (Cr7). La neumopatía crónica no sólo es un condicionante en localidad de vida de los pacientes sino también de la supervivencia de los mismos, de tal forma que se acepta que entre el 80-95 % de los pacientes con FQ fallecen por fracaso respiratorio condicionado por el número de exacerbaciones y la carga bacteriana en las secreciones respiratorias.

En su forma clásica, las manifestaciones clínicas de la FQ afectan a las glándulas sudoríparas y a los tejidos pulmonar, digestivo y reproductor. Mientras que aquellos tipos de FQ, que determinan una disminución severa en la función del canal (menor al 1 % de la actividad del mismo), conducen a un cuadro clínico grave con marcada insuficiencia pancreática. Es por lo tanto primordial, en estos pacientes la realización de un diagnóstico temprano, puesto que la sintomatología aparece en los primeros meses / años de vida. La anomalía de la CFTR, proteína normalmente presente en las membranas de diferentes tejidos, reguladora de los transportes de sodio y cloro a través de las células epiteliales, lleva a una alteración en los canales de Cloruro, con una mayor concentración del Cloruro de sodio (ClNa) en el sudor y la elevación de los niveles sanguíneos de sustancias como la tripsina pancreática, que se valora mediante la prueba sanguínea de **Tripsina Inmunorreactiva (TIR)**.

En los recién nacidos con FQ se encuentran cifras más elevadas de TIR que en los no afectados, y estas cifras se mantienen altas al repetir la determinación dentro de los 20 días de vida. El diagnóstico se debe confirmar con la determinación posterior del Test del Sudor (TS).

En niños con más de 30 días de vida, los resultados dudosos o negativos de TIR no permiten descartar la enfermedad, porque los niveles de TIR disminuyen en forma progresiva a valores normales, debido a que el tejido pancreático normal es reemplazado paulatinamente por tejido graso.

### 5.6.1. ANTECEDENTES HISTÓRICOS

Las primeras observaciones que se hicieron sobre la entidad patológica de la enfermedad de la FQ se recoge desde el folclore, en el Norte de Europa, en las que se aseguraban que los niños al besarlos tenían un sabor salado era porque estaban embrujados y morirían prematuramente; el sabor salado se recoge de un manuscrito alemán del siglo XV y les llamaban "**los niños encantados**".

Las primeras descripciones histopatológicas se realizan en 1595, por Peter Paw en Holanda; este realizó la autopsia de una niña de 11 años de edad, supuestamente hechizada que había padecido síntomas extraños durante ocho años de su vida, la paciente se encontraba muy delgada, y su páncreas estaba abultado, cirroso y de color blanco y brillante después de cortarlo determino que la causa de la muerte de la paciente era de una enfermedad en el páncreas. Además Gerardus Blasius, en 1677, observo las mismas alteraciones en un niño de 9 años.

Un pediatra sueco Nils Rosen Von Rosenstein (1706 - 1773) detallo en su libro sobre las enfermedades infantiles a la enfermedad **fluxus coeliacus**, consistía en diarrea, distrofia, falta de crecimiento, debilidad, edemas en las manos y pies, vientre distendido y páncreas endurecido, probablemente poseía la descripción de la FQ.

En 1936, Fanconi estableció una relación entre la afectación del páncreas y la parte respiratoria, publicada en una revista europea las características clínicas de una enfermedad que presentaban algunos miembros de una familia, denominandola "Fibromatosis congénita familiar del páncreas con bronquiectasias".

En el año 1938, la Doctora **Dorothy Hansine Andersen** , estadounidense especializada en patología, tras realizar estudios exhaustivos en los órganos obtenidos de las autopsias dio el nombre por primera vez de "FIBROSIS QUÍSTICA DEL PANCREAS" en 1954 Andersen con Hodges concluyeron que la enfermedad concuerda con una herencia autosómica recesiva. En aquella época el diagnóstico se realizaba por la familiaridad y la afectación respiratoria y digestiva.



Figura 5.9: Dr. Dorothy Hansine Andersen

La década de los 1950 se caracterizó por el desarrollo de centros dedicados al estudio y manejo de la FQ, en Europa y Norteamérica. En 1955 se crea la US National CF Research Foundation (posteriormente CF Foundation) y en 1959 la Fundación Canadiense de FQ, y en 1965 se forma en Paris la CF International Association.

La determinación del cloro y el sodio en el sudor comenzó a construir el método diagnóstico de la FQ. Inicialmente se exponían a los pacientes a altas temperaturas para hacerlos sudar, la prueba de estimulación del sudor mediante iontoforesis con pilocarpina, la cual permitía de forma segura analizar las concentraciones electrolíticas de cloro en el sudor.

En 1983, Paul M. Quinton en la Universidad de California, en Riverside, descubrió en la capa epitelial de los conductos sudoríparos existía un defecto específico en la reabsorción del cloro y por ello algunos pacientes con FQ mostraban un sabor salado. El esfuerzo culminó, con el trabajo de los investigadores La chep Tsui y John R. Riordan del Hospital for sick children, en Toronto, con el descubrimiento del gen que genera la fibrosis quística, y la proteína codificada por el gen y relacionada con el gradiente del cloro la denominaron Cystic Fibrosis Transmembrane.

También se identificó la anomalía genética más frecuente que afectaba al 70 % de los enfermos, conocida como el DF508, consiste en una deleción de tres nucleótidos que originaba la alteración de los aminoácido fenilalanina en la posición 508. Por técnicas de inmunohistoquímica se ha identificado el ARN mensajero del CFTR en las glándulas sudoríparas, páncreas, criptas intestinales, conductos biliares y en grandes cantidades, en los túbulos renales, donde no se expresa la enfermedad, quizás debido a la existencia de un canal alternativo de cloro.

## 5.7. MODELO MATEMÁTICO DE HARDY - WEINBERG

La genética de poblaciones estudia la distribución de los genes en las poblaciones, así como los factores que se mantienen o cambian con la constitución genética de las mismas, de una generación a otra. Una de sus leyes más importante es la **ley de Hardy - Weinberg** es el equilibrio genético, piedra angular de la genética de población, que proporciona un modelo matemático para el estudio de los cambios evolutivos en la frecuencia de los alelos dentro de una población que permanecen constantes de una generación a otra. En 1908, el matemático británico **Godfrey Harold Hardy** y el médico alemán **Wilhelm Weinberg** creían que la frecuencias de los genes es la variable pertinente en la genética de una población. Los alelos que se encuentran en un locus determinado forman un "consorcio de genes".

El modelo denominado **Ley Hardy - Weinberg** utiliza

- Los principios de las leyes mendelianas de segregación.
- Las teorías simples de las probabilidades
- Para explicar las relaciones de las frecuencias alélicas y genotípicas de una población.

Este modelo describe con lo que pasa con los alelos y los genotipos en una población "ideal" que sea

- Infinitamente grande
- Que los emparejamientos que se produzcan sean aleatorios.
- Las generaciones son discretas (no hay solapamiento de generaciones)
- La Migración es despreciable

### 5.7.1. EL MODELO DENOMINADO LEY HARDY - WEINBERG (EQUILIBRIO)

De acuerdo a esto la **ley de equilibrio** hace unas predicciones:

- Las frecuencias de los alelos en el acervo genético no varían con el tiempo.
- Si se examinan dos alelos de los cuales son A y a, después de una generación de emparejamiento aleatorio, las frecuencias de los genotipos AA, Aa, aa dentro de la población puede calcularse mediante la formula

$$p^2 + 2pq + q^2 = (p + q)^2 = 1 \quad (1)$$

Donde p es la frecuencia del alelo A, q es la frecuencia del alelo a.

- $p^2$ : AA; frecuencia genotípica del homocigoto dominante.
- $2qp$ : Aa; frecuencia genotípica del heterocigoto.
- $q^2$ : aa; frecuencia genotípica del homocigoto recesivo.

	A (p)	a (q)
A (p)	AA	Aa
A (q)	Aa	aa

Cuadro 5-1: Cuadro de Punnett para el equilibrio de Hardy - Weinberg

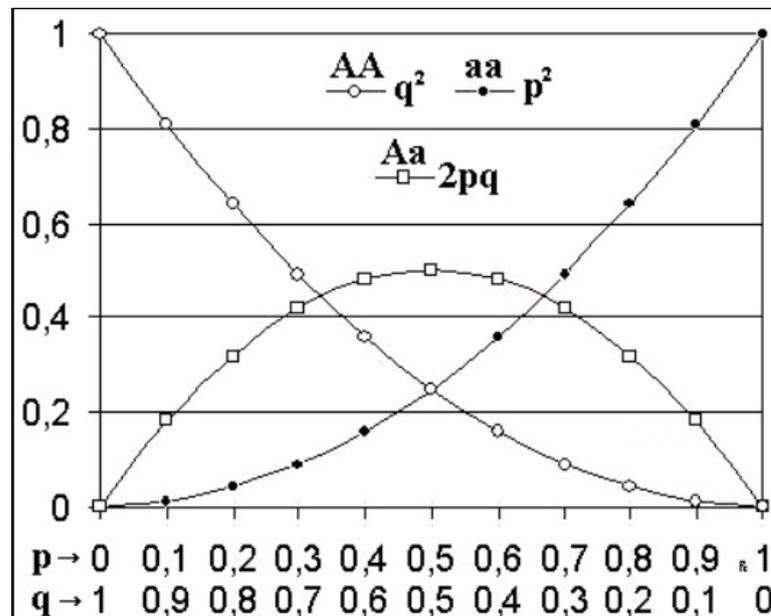


Figura 5.10: Grafica H-W, donde el eje horizontal muestra las dos frecuencias alélicas p y q, el eje vertical muestra la frecuencia de los genotipos y los tres posibles genotipos se representan por los distintos glifos.

La utilidad analítica de las razones fundamentales de los genotipos ( $p^2 : 2pq : q^2$ ) se muestra el hecho de que conociendo la frecuencia del fenotipo recesivo ( $q^2$ ) se puede calcular la proporción de los heterocigotos ( $2pq$ ), que es relativamente grande aunque el fenotipo recesivo sea raro.

Ahora el cálculo para la frecuencia de los heterocigotos

$$\begin{aligned}
 (p + q)^2 &= 1 \\
 \sqrt{(p + q)^2} &= \sqrt{1} \\
 p + q &= 1
 \end{aligned}$$

Donde encontramos para hallar la frecuencia para p

$$p = 1 - q \quad (2)$$

Donde encontramos para hallar la frecuencia para q

$$q = 1 - p \quad (3)$$

Ahora para la frecuencia de los heterocigotos

$$2pq = 2(1 - q)q \quad (4)$$

## Herencia de la fibrosis quística:

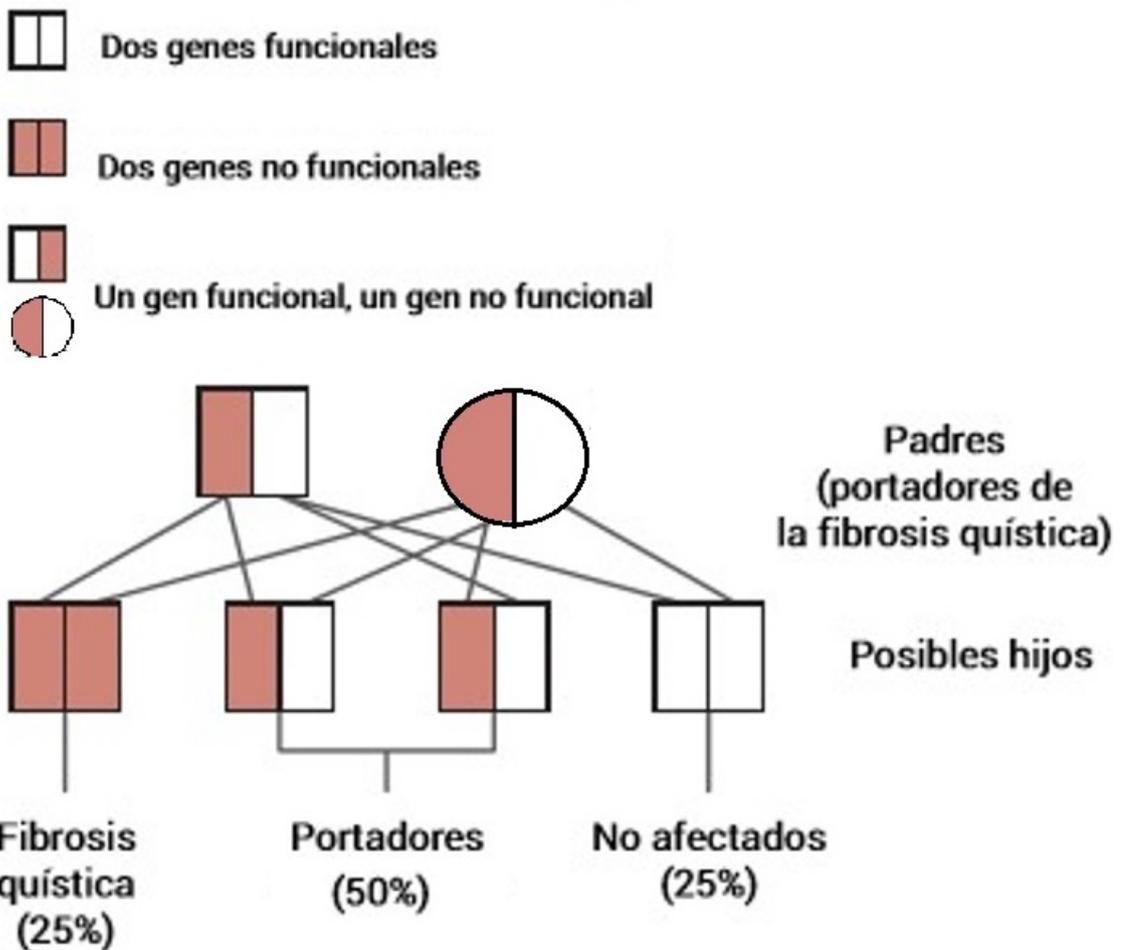


Figura 5.11: Cruce Genético FQ

## 5.7.2. POBLACIÓN INFINITA

Consideramos una **población infinita** (muy grande en cualquiera aplicación práctica) de individuos diploides respecto a una determinada localización de gen. Supóngase que en esta localización existen dos pares de alelos, A y a. existirán tres genotipos distintos, AA, Aa y aa, con las respectivas frecuencias relativas.

$$\alpha, 2\beta, \gamma \quad (1)$$

De aquí se determina la frecuencia de los genes A y a, que son

$$p = \alpha + \beta \quad , \quad q = \beta + \gamma \quad (2)$$

Suponemos que la generación siguiente se origina mediante un apareamiento aleatorio (la unión de los gametos se produce en forma aleatoria con independencia del genotipo de los padres y cada apareamiento da lugar a un mismo número de descendientes). En el cuadro 5.12. Podemos ver el apareamiento, sus frecuencias relativas y frecuencias genotipos de la generación hija. Las frecuencias genotípicas  $\langle \tilde{\alpha}, 2\tilde{\beta}, \tilde{\gamma} \rangle$  en la generación hija viene entonces dadas por la regla de

$$\tilde{\alpha} = (\alpha + \beta)^2 \quad , \quad \tilde{\beta} = (\alpha + \beta)(\beta + \gamma) \quad , \quad \tilde{\gamma} = (\beta + \gamma)^2 \quad (3)$$

Con la ecuación (3) define una aplicación en el conjunto

$$M = \{(\alpha, \beta, \gamma) : \alpha, \beta, \gamma \geq 0, \alpha + 2\beta + \gamma = 1\} \subset \mathbb{R}^3$$

Posibilidades de apareamiento de los genotipos	Frecuencia	Desdoblamiento en la generación hija		
		AA	Aa	aa
AA x AA	$\alpha^2$	$\alpha^2$	0	0
AA x Aa	$4\alpha\beta$	$2\alpha\beta$	$2\alpha\beta$	0
Aa x Aa	$4\beta^2$	$\beta^2$	$2\beta^2$	$\beta^2$
AA x aa	$2\alpha\gamma$	0	$2\alpha\gamma$	0
Aa x aa	$4\beta\gamma$	0	$2\beta\gamma$	$2\beta\gamma$
aa x aa	$\gamma^2$	0	0	$\gamma^2$
Frecuencias de genotipos en la generación hija		$(\alpha + \beta)^2$	$2[(\alpha + \beta)(\beta + \gamma)]$	$(\beta + \gamma)^2$

Figura 5.12: Tabla de Uniones al Azar

En sí mismo. Si designamos las frecuencias de genotipos en las generaciones sucesivas por  $\alpha_v, 2\beta_v, \gamma_v$  ( $v = 0$  para la población inicial), se obtiene a partir de (3) la ecuación en diferencias

$$\alpha_{v+1} = (\alpha_v + \beta_v)^2 \quad , \quad \beta_{v+1} = (\alpha_v + \beta_v)(\beta_v + \gamma_v) \quad , \quad \gamma_{v+1} = (\beta_v + \gamma_v)^2 \quad (4)$$

De (3) y (2) se deduce

$$\tilde{\alpha} = p^2 \quad , \quad 2\tilde{\beta} = 2pq \quad , \quad \tilde{\gamma} = q^2 \quad (5)$$

Las frecuencias de genotipos las generaciones hijas depende pues solamente de las frecuencias de genes en la generación de los padres. En la generación hija las frecuencias de genes son ahora  $p^2 + pq = p$  y  $pq + q^2 = q$ . Las frecuencias de genes son por lo tanto constantes y las frecuencias de genotipos son constantes a partir de la generación hija.

$$\alpha_v = p^2 \quad , \quad \beta_v = pq \quad , \quad \gamma_v = q^2 \quad v = 1, 2, \dots \quad (6)$$

Toda solución de la ecuación (4) que empiece en M alcanza con un paso el estado de equilibrio. El estado descrito en (2) recibe el nombre de equilibrio de Hardy Weinberg.

Se puede añadir todavía algunas consideraciones. Se observa que es siempre

$$p^2 + q^2 \geq 2pq \quad (7)$$

Y, en consecuencia, la mitad a lo sumo de todos los individuos son heterocigotos. La condición:

$$\alpha\gamma = \beta^2 \quad (8)$$

Es necesaria y suficiente para que las poblaciones se encuentren en el equilibrio de Hardy-Weinberg. Se sigue en efecto de (4).

$$\begin{aligned} \tilde{\alpha} &= (\alpha + \beta)^2 = \alpha^2 + 2\alpha\beta + \beta^2 = \alpha^2 + 2\alpha\beta + \alpha\gamma = \alpha(\alpha + 2\beta + \gamma) = \alpha(1) = \alpha \\ \tilde{\beta} &= (\alpha + \beta)(\beta + \gamma) = \alpha\beta + \beta^2 + \beta\gamma + \alpha\gamma = \alpha\beta + \beta\gamma + 2\beta^2 = \beta(\alpha + 2\beta + \gamma) \\ &= \beta(1) = \beta \\ \tilde{\gamma} &= \gamma \end{aligned}$$

Si en el lugar considerando se encuentran no dos, si no  $n \geq 2$  alelos  $a_1, a_2, \dots, a_n$ , existirán entonces  $n^2$  genotipos  $a_1a_1, a_1a_2, \dots, a_na_n$ , (identificándose siempre  $a_ja_k$  y  $a_ka_j$ ). Sean

$$a_{jk} = a_{kj} \geq 0 \quad j, k = 1, \dots, n \quad (9)$$

Las frecuencias de los genotipos en la población inicial con,

$$\sum_{j,k} a_{jk} = 1 \quad (10)$$

La matriz  $A = (a_{jk})$  es simétrica y de elemento no negativos. Bajo la hipótesis de apareamiento aleatorio, las frecuencias de genotipos de la generación siguiente se obtienen mediante la fórmula.

$$\tilde{a}_{jk} = \left( \sum_i a_{ji} \right) \left( \sum_l a_{kl} \right) \quad j, k = 1, \dots, n. \quad (11)$$

La matriz  $A = (\tilde{a}_{jk})$  tiene por lo tanto una representación  $\tilde{a}_{jk} = P_j P_k$ ,  $\sum_j P_j = 1$  y es por consiguiente una matriz de rango 1 (una diada). Si por otra parte la matriz inicial es ya una diada,  $a_{jk} = P_j P_k$ , se sigue entonces

$$\tilde{a}_{jk} = \left( \sum_i P_j P_i \right) \left( \sum_l P_k P_l \right) = P_j P_k \quad (12)$$

También aquí se alcanza con un paso el estado de equilibrio, el cual viene descrito mediante las ecuaciones

$$a_{jk} = P_j P_k \quad j, k = 1, \dots, n \quad (13)$$

La relación (13) es la ley de Hardy- Weinberg para n- alelo.

### Ejemplo 1

La expresión algebraica de la tabla de las frecuencias genotípicas se conservan constantes de una generación a la siguiente.

Genes alélicos	Frecuencia Asignada	Ejemplo numerico	Genotipos	Frecuencia en la Población	Ejemplo numérico	Ejemplo numérico %
A	p	0.90	AA	$p^2$	0.81	81
a	q	0.10	Aa	$2pq$	0.18	18
			aa	$q^2$	0.01	1
Juntos		1.00	Todos los tipos		1	100

Cuadro 5-2: Caracteres de la generación parental

Padres ♂ ♀	Frecuencia del tipo de unión	Frecuencia de la prole*			Ejemplo numérico %		
		AA	Aa	aa	AA	Aa	aa
AA x AA	$p^4$	$p^4$			6561	0	0
AA x Aa	$4p^3q$	$2p^3q$	$2p^3q$		1458	1458	0
Aa x AA							
AA x aa	$2p^2q^2$		$2p^2q^2$		0	162	0
aa x AA							
Aa x Aa	$4p^2q^2$	$p^2q^2$	$2p^2q^2$	$p^2q^2$	81	162	81
Aa x aa							
aa x Aa	$4pq^3$		$2pq^3$	$2pq^3$	0	18	18
aa x aa							
	$q^4$			$q^4$	0	0	1
Todos los tipos		$p^2$	$2pq$	$q^2$	8100	1800	100

Cuadro 5-3: Prole de Uniones al Azar

Para los resultados de las columnas que corresponde a las frecuencias de la prole \* se compone de la suma de la columna AA, la columna Aa y la columna aa; de acuerdo a esto se da lo siguiente:

a) La suma de la columna AA es

$$p^4 + 2p^3q + p^2q^2 = p^2(p^2 + 2pq + q^2) = p^2(p + q)^2$$

b) La suma de la columna Aa es

$$2p^3q + 4p^2q^2 + 2pq^3 = 2pq(p^2 + 2pq + q^2) = 2pq(p + q)^2$$

c) La suma de la columna aa

$$p^2q^2 + 2pq^3 + q^4 = q^2(p^2 + 2pq + q^2) = q^2(p + q)^2$$

### Frecuencias de apareamiento

En el apareamiento al azar, supuesto principio de Hardy-weinberg, la frecuencia de los apareamientos de diferentes tipos es el producto de las frecuencias de los individuos apareados. Esto se observa en el cuadro 5-3; si  $p^2$  es la frecuencia de las personas AA, entonces  $p^2 \times p^2 = p^4$ , es la frecuencia de los apareamientos  $AA \times AA$ . Análogamente para las frecuencias Aa y aa.

## 5.8. METODOLOGÍAS ESTADÍSTICA

Las medidas de Posición o medidas de Tendencia central, nos permiten determinar la posición de un valor con respecto a un conjunto de datos, el cual lo consideramos como representativo o típico, para el total de las observaciones.

Un **PROMEDIO** es valor en la escala de las X correspondiente a una distribución de frecuencias, cuyo valor se calcula para representarla en calidad de medida típica del grupo. La expresión corriente de promedio, suele en la mayoría de los casos referirse a la **media aritmética**.

**MEDIA ARITMÉTICA**  $\bar{X} = \mu$ .

La mas conocida y utilizada por todos, por su gran estabilidad es la preferida en el muestreo, sus fórmulas admiten tratamiento algebraico. Se define como " **Suma de todos los valores observados, divididos por el mismo número total de observaciones**", de esta forma se aplica para datos sin agrupar o denominados datos originales.

$$\mu = \bar{X} = \frac{\sum X_i}{N}, \text{ siendo } \mu = \frac{X_1 + X_2 + \dots + X_N}{N}$$

**MEDIANA**  $M_e$  La mediana es el valor de la variable que ocupa la posición central, cuando los datos se disponen en orden de magnitud.

- a) Si n es impar, la mediana es el valor que ocupa la posición  $\frac{n+1}{2}$  una vez que los datos han sido ordenados (en orden creciente o decreciente), porque éste es el valor central. Es decir:  $M_e = x_{\frac{n+1}{2}}$ .

- b) Si  $n$  es par, la mediana es la media aritmética de los dos valores centrales. Cuando  $n$  es impar, los dos datos que están en el centro de la muestra ocupan las posiciones  $\frac{n}{2}$  y  $(\frac{n}{2}) + 1$ . Es decir:

$$M_e = \frac{(x_{\frac{n}{2}} + x_{\frac{n}{2}+1})}{2}$$

La **MODA** de una distribución se define como el valor de la variable que más se repite. En un polígono de frecuencia la moda corresponde al valor de la variable que está bajo el punto más alto del gráfico. Una muestra puede tener más de una moda.

### 5.8.1. MEDIDAS DE DISPERSIÓN

Las medidas de dispersión entregan información sobre la variación de la variable. Pretenden resumir en un solo valor la dispersión que tiene un conjunto de datos. Las medidas de dispersión más utilizadas son: Rango, Varianza, Desviación estándar, Coeficiente de variación.

**RANGO U OSCILACIÓN** Se define como la diferencia entre el mayor valor de la variable y el menor valor de la variable.

**VARIANZA**  $\sigma^2$  (sigma cuadrado) para el universo o población y con el símbolo  $S^2$  (s cuadrado), se define como la media aritmética de los cuadrados de las desviaciones respecto a la media aritmética; su importancia radica especialmente en que da origen a otra medida de dispersión mucho más significativa denominada desviación típica o estandar ( $S$ ). En muchas situaciones es preciso estimar la varianza de una población a partir de una muestra. Si se toma una muestra con reemplazo ( $x_1, \dots, x_n$ ) de  $n$  valores de ella, de entre todos los estimadores posibles de la varianza de la población de partida, existen dos de uso corriente:

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

**DESVIACIÓN ESTANDAR O TÍPICA** ( $S$  u  $\sigma$ ) se define como la raíz cuadrada de la varianza, tomada siempre con el signo positivo, también se puede definir como la raíz cuadrada de las desviaciones respecto a la media. Es la medida de dispersión más conocida y por tal motivo la más utilizada; desempeña un papel muy importante en el análisis de los datos estadísticos. En datos no agrupados la desviación típica será:

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

En demografía, se emplea el concepto de mortalidad cuando se produce la defunción o la acción de muerte sobre los integrantes de una población. La muerte es un riesgo al que está expuesta

una persona durante toda la vida. Obviamente, es un hecho que ocurre una sola vez, por ende, toda la población está expuesta al riesgo de morir.

## 5.8.2. MEDICIÓN DE LA MORTALIDAD

El estudio de la mortalidad se realiza a través de indicadores que permiten medir su incidencia y comportamiento. De un lado, es posible su estudio con datos absolutos, es decir, de los hechos ocurridos, en este caso, defunciones y la población expuesta al riesgo de morir. De otro lado, su estudio se basa en medidas relativas, que pueden ser expresadas en Tasas.

### TASA ESPECIFICA DE MORTALIDAD

La tasa específica de mortalidad es el indicador más utilizado en la medición de la mortalidad. Se obtiene de la relación entre el número de defunciones ocurridas en un período de tiempo determinado (generalmente un año) y una estimación de la población expuesta al riesgo de morir en el mismo período.

La estimación de la población supone calcular el tiempo vivido por aquella durante dicho período. Dadas las dificultades que presenta su cálculo, se estima la población a mitad de periodo. Así:

$$d_z = \frac{D_z}{N_{30-VI-z}} \times 1000$$

Donde:

- $d_z$  es la Tasa específica de Mortalidad del año abarcado.
- $D_z$  son las Defunciones ocurridas en el año  $z$ .
- la población estimada al 30 de Junio del año  $z$ .

La tasa multiplicada por mil, representa la frecuencia relativa con la que ocurren las defunciones en una población durante un año.

## *Capítulo 6*

---

# **METODOLOGÍA**

---

Para este proyecto se utilizó los datos estadísticos, con cuenta salud pública departamental sobre las poblaciones que padecen la enfermedad de fibrosis quística en Neiva - Huila.

Se sistematizó la información, para luego de aplicar el modelo matemático, se analizó las características alélicas y determinísticas que se identifican y establecen la población adecuada en el ambiente.

El presente proyecto culminará con el diseño de una propuesta científico - académica consistente en la publicación de un artículo, sobre el estudio de la información de la fibrosis quística con base en la parte genética desde una familia hasta una población infinita y multialélica ( polimorfismo genético ) sobre el modelo matemático propuestos que mejor se ajusta para encontrar soluciones del problema a nivel del departamento del Huila.

## Capítulo 7

---

# ANÁLISIS ESTADÍSTICO DE LA FQ CON RESPECTO AL HUILA

---

La información de los casos presentados en el departamento del Huila del periodo comprendido del 2009 al 2017, en los eventos de morbilidad y mortalidad reportados por todas las IPS y ESE del departamento; teniendo en cuenta los códigos CIE 10 en la cual se clasifica las características de dicha enfermedad:

- E840 - FIBROSIS QUÍSTICA CON MANIFESTACIONES PULMONARES
- E841 - FIBROSIS QUÍSTICA CON MANIFESTACIONES INTESTINALES
- E848 - FIBROSIS QUÍSTICA CON OTRAS MANIFESTACIONES
- E849 - FIBROSIS QUÍSTICA, SIN OTRA ESPECIFICACIÓN

### ANÁLISIS ESTADÍSTICO DEL CUADRO (7-1) MORBILIDAD (FQ): PERSONAS ATENDIDAS EN URGENCIAS O CONSULTA EXTERNA

Diagnostico	2009	2010	2011	2012	2013	2014	2015	2016	2017	TOTAL
E840	1	0	5	3	8	5	4	10	8	44
E841	1	0	0	0	1	3	8	1	2	16
E848	0	4	2	2	5	4	3	1	1	22
E849	13	19	24	26	34	23	11	6	9	175
TOTAL	15	23	31	31	48	35	26	28	20	257

Cuadro 7-1: Morbilidad: Número de personas atendidas en urgencias y por consulta externa con diagnósticos de fibrosis quística. Huila 2009 a 2017.

Con respecto al cuadro (7-1) se hace el estudio estadístico de morbilidad del número de personas atendidas en urgencias y por consulta externa de la enfermedad la fibrosis quística en

el Huila; del cual se extiende con varias manifestaciones que son consecuencia de esta enfermedad, donde clasifican por medio de unos códigos que se obtienen del libro (CIE 10); se va a trabajar con un periodo de tiempo de 9 años y utilizando el **software R proyect**, se basó por medio de unos algoritmos para obtener los resultados necesarios para el desarrollo del trabajo.

El diagnóstico **E840 - Fibrosis quística con manifestaciones pulmonares**, la cual es uno de los diagnósticos de la enfermedad principal de la FQ. Abarcando el periodo de los años 2009 hasta el 2017 y tomando como referencia la primera fila, un promedio de 4,9; por consiguiente se presentaron 5 casos de esta causa de la FQ. De todas las patologías de FQ corresponde el 17 % a la E840.

Se organizó los datos de menor a mayor y se obtuvo la mediana con un resultado de 5; además de eso se obtuvo el rango con un valor de 10, donde es el valor máximo de los datos de la primera fila de diagnóstico, se presentó en el año 2016, donde hubieron 10 casos de esta enfermedad y encontramos el valor mínimo 0 del año 2010 en donde no se presentaron casos con fibrosis quística. Se logra evidenciar que los números que más se repiten son 5 y 8, por ende, se dice que es una distribución bimodal.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 11 de los datos de los usuarios con este padecimiento de la enfermedad de la fibrosis quística con manifestación pulmonar, durante los 9 años y una desviación estándar o típica de 3 casos con la E840.

El diagnóstico **E841 - Fibrosis Quística con manifestaciones intestinales** y tomando como referencia la segunda fila, se obtiene un promedio de 1,8; por consiguiente se presentaron 2 casos de esta causa de la FQ. De todas las patologías de FQ corresponde al 6 % a la E841.

Se organizó los datos de menor a mayor y se obtuvo la mediana con un resultado de 1; además de eso se obtuvo el rango con un valor de 8; el valor máximo de los datos de la primera fila de diagnóstico, se presentó en el año 2015, donde hubieron 8 casos de esta enfermedad y encontramos el valor mínimo 0 en el año 2010, 2011 y 2012 en donde no se presentaron casos con fibrosis quística. Se logra evidenciar que los números que más se repiten son 0 y 1, por ende, se dice que es una distribución bimodal.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 6 de los datos de los usuarios con este padecimiento de la enfermedad fibrosis quística con manifestación intestinales durante los 9 años y una desviación estándar o típica de 3 casos con la E841.

El diagnóstico **E848- Fibrosis Quística con otras manifestaciones**, tomando como referencia la tercera fila, se obtuvo un promedio de 2.4 % por consiguiente se presentaron 2 casos de esta causa de la FQ. De todas las patologías de FQ corresponde al 9 % a la E848.

Se organizó los datos de menor a mayor y se obtuvo la mediana con un valor de 2; además de eso se obtuvo el rango con un valor de 5, el valor máximo de los datos de la tercera fila del diagnóstico que se presentó en el año 2013, donde hubieron 5 casos de esta enfermedad y encontramos el valor mínimo 0 en el año 2009 en donde no se presentaron casos con fibrosis quística.

Se logra evidenciar que los números que más se repiten son 1, 2 y 4, por ende, se dice que es una distribución trimodal. Luego abarcando las medidas de dispersión obtuvo el valor esperado fue de 3 de los datos de los usuarios con este padecimiento de la enfermedad fibrosis quística con otras manifestaciones durante los 9 años y una desviación estándar o típica de 2 casos con la E848.

El diagnóstico **E849 - Fibrosis Quística sin otra especificación**, y tomando como referencia la cuarta fila, se obtuvo un promedio de 19,4 por consiguiente se presentaron 19 casos de esta causa de la FQ. De todas las patologías de FQ corresponde al 68 % a la E849.

Se organizó los datos de menor a mayor y se obtuvo la mediana con un resultado de 19; además de eso se obtuvo el rango con un valor de 25, el valor máximo de los datos de la primera fila de diagnóstico se presentó en el año 2013, lo cual hubieron 34 casos de esta enfermedad y encontramos el valor mínimo 9 en el año 2017 en donde solo se presentaron 9 casos con fibrosis quística.

Se logró evidenciar que ningún valor se repite por lo tanto no tiene moda. Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 65 de los datos de los usuarios con este padecimiento de la enfermedad fibrosis quística sin otra especificación durante los 9 años y una desviación estándar o típica de 8 casos con la E849.

Concluyendo el **análisis total** del cuadro (7-1) se obtuvo un promedio de 7,1; un rango de 34 y una suma total 257 datos en el recorrido de esos 9 años. Del cual se puede evidenciar una desviación típica o estándar con un valor de 9.

Se explora las relaciones entre todas las parejas de variables, en particular la relación de años con cada una de las variables independientes en este caso (E840, E841, E848, E849). Esto lo detectaremos a través de las correlaciones y la función pairs de R-proyect, la cual produce un gráfico matricial para las variables dadas (vease 8.29). Además, se realizó un análisis exploratorio

a los datos que permitan visualizar la correlación existente entre las variables (vease 8.30). De lo anterior se logró apreciar que existen las siguientes correlaciones:

1. 79.4 % entre la variable E840 y el AÑO.
2. 63.3 % entre la variable E848 y E849.

Por ende, se hace una regresión lineal entre estas variables  $E840 \sim \text{AÑO}$ . Donde se da la recta de regresión lineal que mejor se ajusta es  $y = 0,96x - 1941,01$ .

Estos dos parámetros pueden interpretarse del siguiente modo: -1941,01 es la cantidad de casos E840 dados en el año 0, lo cual no tiene sentido. De hecho, en multitud de ocasiones la interpretación de este parámetro no es relevante y todo el interés recae sobre la interpretación del resto de parámetros.

El parámetro 0,96 indica la cantidad que incrementan los casos de E840, por el aumento en 1 año a partir del año 2009.

Abarcando el segundo punto del análisis de las otras dos variables de las correlaciones, surge la recta de regresión lineal que mejor se ajusta es  $y = 0,1306x - 0,09574$ ; lo cual se puede interpretar como un aumento de 0,1306 casos E848 cuando se da 1 caso de E849.

## ANÁLISIS ESTADÍSTICO DEL CUADRO (7-2) MORBILIDAD (FQ): PERSONAS ATENDIDAS EN HOSPITALIZACIÓN

Diagnostico	2010	2011	2012	2013	2014	2015	2016	2017	TOTAL
E840	1	0	1	0	0	1	3	1	7
E841	0	0	0	1	0	0	0	0	1
E848	0	0	0	0	1	0	0	0	1
E849	1	1	1	0	1	1	1	0	6
TOTAL	2	1	2	1	2	2	4	1	15

Cuadro 7-2: Morbilidad: Número de personas atendidas en hospitalización con diagnósticos de fibrosis quística. Huila 2010 a 2017.

Con respecto al cuadro (7-2) se hace el estudio estadístico de morbilidad del número de personas atendidas en hospitalización con diagnóstico de la enfermedad la fibrosis quística en el Huila; del cual se extiende con varios diagnósticos que son consecuencia de cuya enfermedad, donde clasifican por medio de unos códigos que se obtienen del libro (CIE 10); se va a trabajar con un periodo de tiempo de 8 años de los cuales se abarcan los años 2010 hasta el 2017, ya que en el año 2009 no se presentaron casos de esta enfermedad. Utilizando el **software R proyect**, se

basa por medio de unos algoritmos para obtener los resultados necesarios para el desarrollo del trabajo.

El diagnóstico **E840 - Fibrosis quística con manifestaciones pulmonares**, la cual es uno de los diagnósticos de la enfermedad principal de la FQ, y tomando como referencia la primera fila, se obtiene un promedio de 0,9 por consiguiente se presentaron 1 casos de esta causa de la FQ. De todas las patologías de FQ corresponde al 47 % a la E840.

Se organizan los datos de menor a mayor y se obtiene la mediana con un valor de 1; además el rango con un valor de 3; el valor máximo de los datos de la primera fila de diagnóstico se presenta en el año 2016 donde hubieron 3 casos de esta enfermedad, encontramos el valor mínimo 0 en los años 2011, 2013 y 2014 en donde no se presentaron casos con fibrosis quística. Se logra evidenciar que los números que más se repiten son 0 y 1, por ende, se dice que es una distribución bimodal.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 1 de los datos de los usuarios con este padecimiento de la enfermedad fibrosis quística con manifestación pulmonar durante los 8 años y una desviación estándar o típica de un solo caso con respecto a este diagnóstico.

El diagnóstico **E841 - Fibrosis Quística con manifestaciones intestinales** y tomando como referencia la segunda fila, se obtiene un promedio de 0,13 por consiguiente no se presentaron casos de esta causa de la FQ. De todas las patologías de FQ corresponde al 7 % a la E841.

Se organizan los datos de menor a mayor y se obtiene la mediana con un resultado de 0; además de eso se obtiene el rango con un valor de 1; el valor máximo de los datos de la primera fila de diagnóstico se presenta en el año 2013 donde hubo solo 1 casos de esta enfermedad, encontramos el valor mínimo 0 en los años 2010, 2011, 2012, 2014, 2015, 2016 y 2017 en donde no se presentaron casos con fibrosis quística. Se logra evidenciar que los números que más se repiten son 0, por ende, se dice que es la moda es 0.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 0,13 de los datos de los usuarios con este padecimiento de la enfermedad fibrosis quística con manifestación intestinales durante los 8 años y una desviación estándar o típica de 0,35 casos con la E841.

El diagnóstico **E848- Fibrosis Quística con otras manifestaciones**, tomando como referencia la tercera fila, se obtiene un promedio de 0,13 por consiguiente no se presentaron casos de esta causa de la FQ. De todas las patologías de FQ corresponde al 7 % a la E848.

Se organizan los datos de menor a mayor y se obtiene la mediana con un valor de 0; además obtiene el rango con un valor de 1; el valor máximo de los datos de la primera fila de diagnóstico se presenta en el año 2014 donde hubo solamente 1 caso de esta enfermedad, encontramos el valor mínimo 0 en los años 2010, 2011, 2012, 2013, 2015, 2016 y 2017 en donde no se presentaron casos con fibrosis quística. Se logra evidenciar que los números que más se repiten son 0, por ende, se dice que la moda es 0.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 0,13 de los datos de los usuarios con este padecimiento de la enfermedad fibrosis quística con otras manifestaciones durante los 8 años y una desviación estándar o típica de 0,35 casos con la E848.

La causa **E849 - Fibrosis Quística sin otra especificación**, y tomando como referencia la cuarta fila, se obtiene un promedio de 0,8 por consiguiente se presentó 1 caso de esta causa de la FQ. De todas las patologías de FQ corresponde al 40 % a la E849.

Se organizan los datos de menor a mayor y se obtiene la mediana con un valor de 1; además de eso se obtiene el rango con un valor de 1; el valor máximo de los datos de la primera fila de diagnóstico se presenta en los años 2010, 2011, 2012, 2014, 2015 y 2016 lo cual hubo solamente 1 caso de esta enfermedad y encontramos el valor mínimo 0 en los años 2013 y 2017 en donde no se presentaron casos con fibrosis quística. Se logra evidenciar que el valor que más repite es 1, por ende es la moda.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 0,21 de los datos de los usuarios con este padecimiento de la enfermedad fibrosis quística sin otra especificación durante los 8 años y una desviación estándar o típica de 0,46 casos con la E849.

Concluyendo el **análisis total** del cuadro (7-2) se obtiene un promedio de 0,5; con un rango de 3 y una suma total 15 datos en el recorrido de esos 9 años. Del cual se puede evidenciar una desviación típica o estándar con un valor de 1.

Se explora las relaciones entre todas las parejas de variables, en particular la relación de años con cada una de las variables independientes en este caso (E840, E841, E848, E849). Esto lo detectaremos a través de las correlaciones y la función pairs de R-proyect, la cual produce un gráfico matricial para las variables dadas (vease 8.63). Además, se realiza un análisis exploratorio a los datos que permitan visualizar la correlación existente entre las variables (vease 8.64). De lo anterior se logra apreciar que existen las siguientes correlaciones:

1. 66.1 % entre la variable E841 y E849

Por ende, se hace una regresión lineal entre estas variables  $E841 \sim E849$ . Donde se da la recta de regresión lineal que mejor se ajusta es  $y = 0,5 - 0,5x$ .

Es inversamente proporcional, con pendiente negativa, lo cual se puede interpretar como una disminución de 0,5 casos E841 cuando se da un caso de E849.

### **ANÁLISIS ESTADÍSTICO DEL CUADRO (7-3) MORTALIDAD: PERSONAS FALLECIDAS CON DIAGNOSTICO DE LA (FQ)**

Diagnostico	2010	2011	2013	2016	2017	TOTAL
E840	0	0	1	0	1	2
E841	0	0	0	1	0	1
E849	1	1	0	0	0	2
TOTAL	1	1	1	1	1	5

**Cuadro 7-3:** Mortalidad: Número de personas fallecidas con diagnósticos de fibrosis quística. Huila 2010 a 2017

Se analizó las tendencias de mortalidad por Fibrosis quística en el departamento del Huila - Colombia para el periodo 2010-2017, a partir de datos que se disponen de la fuente DANE (Departamento Administrativo Nacional de Estadística); con sus bases de datos de la FTP ( Protocolo de transferencia de archivos (File Transfer Protocol o FTP)) y con acceso exclusivo para el departamento. En el cuadro (7-3) del documento se puede observar la mortalidad de Fibrosis Quística (FQ), debida a las causas que se clasifica mediante la CIE 10 de esta enfermedad como **E840**, **E841** y **E849**; como se puede observar en ese mismo cuadro no se encuentra la causa **E848 - FIBROSIS QUISTICA CON OTRAS MANIFESTACIONES**, lo que se deduce que no hubieron casos de mortalidad durante los años tomados para el estudio de esta enfermedad. Además, abarcando los años 2009, 2012, 2014 y 2015, no se presentaron casos de personas fallecidas con causas de la enfermedad (FQ).

La causa **E840 - Fibrosis quística con manifestaciones pulmonares**, la cual es una de las principales fuentes de fallecimiento de la FQ, y tomando como referencia la primera fila, un promedio de 40 % de los casos presentados de esta causa de la FQ.

Se organizaron los datos de menor a mayor y se obtiene la mediana con un valor de 0; el rango con un valor de 1; el valor máximo de los datos de la primera fila de diagnóstico se presenta en los años 2013 y 2017 donde hubo solamente 1 caso de esta enfermedad y encontramos el valor mínimo 0 en los años 2010, 2011, 2012, 2014, 2015 y 2016 en donde no se presentaron casos con fibrosis quística. Se logró evidenciar que los números que más se repiten son 0, por ende, se dice que la moda es 0.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 0,21 de los datos de los pacientes con este padecimiento de la enfermedad fibrosis quística con manifestación pulmonar durante los años trabajados y una desviación estándar o típica de 0,46 casos.

La causa **E841 - Fibrosis Quística con manifestaciones intestinales** y tomando como referencia la segunda fila, se obtiene un promedio de 20 % de los casos presentados de esta causa de la FQ.

Se organizaron los datos de menor a mayor y se obtiene la mediana con un valor de 0; el rango con un valor de 1; el valor máximo de los datos de la primera fila de diagnóstico se presenta en el año 2016 donde hubo solamente 1 caso de esta enfermedad y encontramos el valor mínimo 0 en los años 2010, 2011, 2012, 2013, 2014, 2015 y 2017 en donde no se presentaron casos con fibrosis quística. Se logró evidenciar que los números que más se repiten son 0, por ende, se dice que la moda es 0.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 0,13 de los datos de los pacientes con este padecimiento de la enfermedad fibrosis quística con manifestación intestinal y una desviación estándar o típica de 0,35 casos.

El diagnóstico **E849 - Fibrosis Quística sin otra especificación**, y tomando como referencia la tercera fila, se obtiene un promedio de 40 % de los casos presentados de esta causa de la FQ.

Se organizaron los datos de menor a mayor y se obtiene la mediana con un valor de 0; además de eso se obtiene el rango con un valor de 1; el valor máximo de los datos de la primera fila de diagnóstico se presenta en los años 2010 y 2011 donde hubo solamente 1 caso de esta enfermedad y encontramos el valor mínimo 0 en los años 2012, 2013, 2014, 2015, 2016 y 2017 en donde no se presentaron casos con fibrosis quística. Se logra evidenciar que los números que más se repiten son 0, por ende, se dice que la moda es 0.

Luego abarcando las medidas de dispersión obtenemos el valor esperado fue de 0,21 de los datos de los pacientes con este padecimiento de la enfermedad fibrosis quística sin otra especificación y una desviación estándar o típica de 0,46 casos.

Concluyendo el **análisis total** del cuadro (7-3) se obtiene un promedio de 0,3; un rango de 1 y una suma total 5 datos en el recorrido de esos 5 años. Del cual se evidenció una desviación típica o estándar con un valor de 0.

Se exploró las relaciones entre todas las parejas de variables, en particular la relación de años con cada una de las variables independientes en este caso (E840, E841, E849). Esto lo detectaremos a través de las correlaciones y la función pairs de R-project, la cual produce un gráfico matricial

para las variables dadas (vease 8.89). Además, se realizó un análisis exploratorio a los datos que permitan visualizar la correlación existente entre las variables (vease 8.90). De lo anterior se logró apreciar que existen las siguientes correlaciones:

1. 75.6 % entre la variable E849 y el AÑO.

Por ende, se hizo una regresión lineal entre estas variables  $E849 \sim \text{AÑO}$ . Donde se da la recta de regresión lineal que mejor se ajusta es  $y = 287,8929 - 0,1429x$ .

Es inversamente proporcional, con pendiente negativa, estos dos parámetros pueden interpretarse del siguiente modo: 287,8929 es la cantidad de casos E849 dados en el año 0, lo cual no tiene sentido. De hecho, en multitud de ocasiones la interpretación de este parámetro no es relevante y todo el interés recae sobre la interpretación del resto de parámetros. El parámetro -0,1429 indica la cantidad que disminuyen los casos de E849 por el aumento en 1 año a partir del año 2010.

Utilizando intervalos de confianza pueden ser calculados para tasas específicas, en este caso abarcamos la **Tasa específica de mortalidad** la cuál fue del 9 ‰ de las defunciones de la FQ en la población huilense.

Con respecto a la causa **E840 - FIBROSIS QUÍSTICA CON MANIFESTACIONES PULMONARES** se presentó un caso en el año 2013 y un caso en el año 2017 de resto no hubieron casos en el transcurso del periodo de tiempo tomado al estudio. Luego utilizando la tasa específica de mortalidad abarcando el año 2013 fue del 13 ‰ y la del año 2017 fue 23 ‰ la diferencia de estas tasas específicas fue la población estimada durante el año tomado, incluso en el año 2017 se optó por tomar base de datos parciales.

La causa **E841 - FIBROSIS QUÍSTICA CON MANIFESTACIONES INTESTINALES** sólo se presentó un caso en el año 2016. En los otros años tomados no se evidenció casos de mortalidad de esta patología. Tomando la tasa específica de mortalidad abarcando el año 2016 fue 12 ‰ de la población FQ estimada durante ese año.

La causa **E849 - FIBROSIS QUÍSTICA, SIN OTRA ESPECIFICACIÓN** se presentó un caso en el año 2010 y un caso en el año 2011; no se presentó casos de mortalidad, en el transcurso de el periodo abarcado al estudio de la FQ. Luego utilizando la tasa específica de mortalidad abarcando el año 2010 fue del 15 ‰ y la del año 2011 fue 14 ‰ la variación de estas tasas específicas fue la población de defunciones estimada durante este periodo.

## Capítulo 8

---

# RESULTADOS

---

### 8.0.1. E840

Utilizando el software **R PROJECT** obtamos por el primer paso la creación del cuadro (7-1) tomando la primera fila en donde se encuentra el código E840, se obtiene:

```
> datos<-c(1,0,5,3,8,5,4,10,8)
> año<-c(2009,2010,2011,2012,2013,2014,2015,2016,2017)
> z<-cbind(año,datos)
> z
```

	año	datos
[1,]	2009	1
[2,]	2010	0
[3,]	2011	5
[4,]	2012	3
[5,]	2013	8
[6,]	2014	5
[7,]	2015	4
[8,]	2016	10
[9,]	2017	8

Figura 8.1: Fila 1 E840

Se organizan los datos para poder obtener los resultados:

```
> sort(datos)
[1] 0 1 3 4 5 5 8 8 10
```

Figura 8.2: Datos Organizados

En dónde encontramos los siguientes resultados, utilizando las medidas de tendencia central de estadísticas:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  3.000   5.000   4.889  8.000  10.000

> var(datos)
[1] 11.11111

> sd(datos)
[1] 3.333333
```

Figura 8.3: Resultados E840

- La Media(Mean):4,889  $\approx$  4,9
- La Mediana(Median): 5,00  $\approx$  5
- Rango: 10
- Valor mínimo(Min):0,000  $\approx$  0
- Valor Máximo(Max):10,000  $\approx$  10
- Primer Cuartil(1st Qu):3,000  $\approx$  3
- Tercer Cuartil(3rd Qu):8,000  $\approx$  8
- Moda: 5 y 8.

Aplicando las medidas de dispersión en estadística:

- Varianza(var): 11,11111  $\approx$  11
- Desviación Estandar(sd):3,33333  $\approx$  3

Ahora,veamos el código y la gráfica de la frecuencias con la causa E840 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,
legend.text = NULL, beside = FALSE, horiz = FALSE,border = par("fg"),
main = "MORBILIDAD E840", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')
```

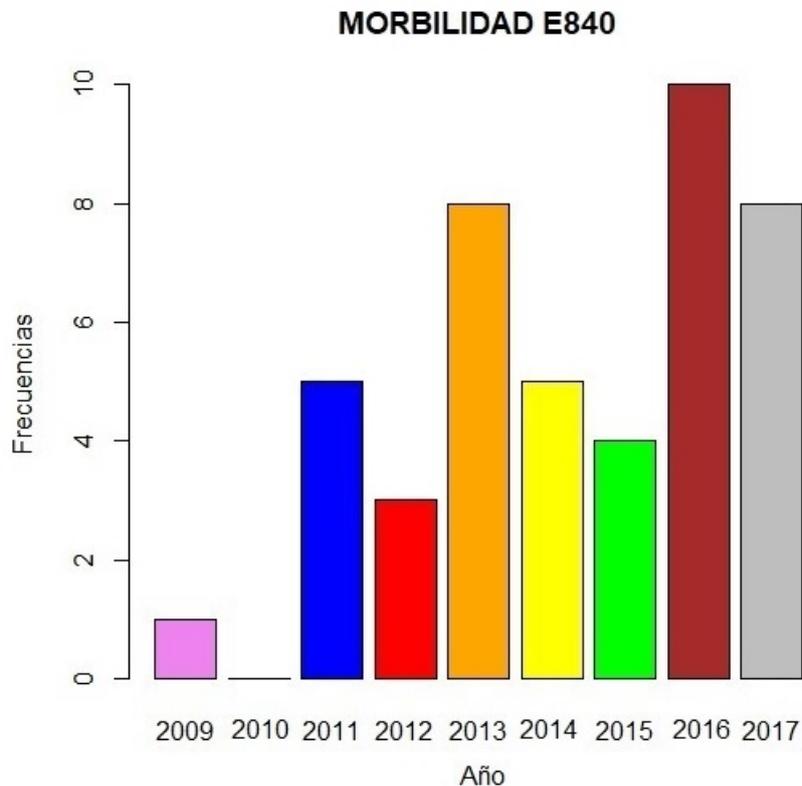


Figura 8.4: Histograma de frecuencias de la E840

Luego, aplicamos método de caja y bigotes (Box-plot) al estudio de la E840.

```
boxplot(datos,main='Morbilidad E840', xlab='años',ylab='numero de pacientes')
```

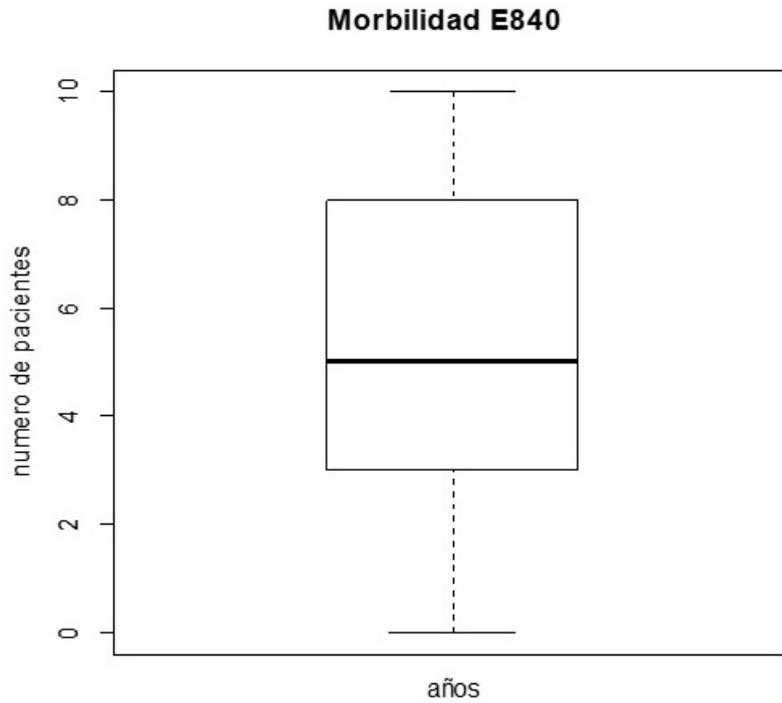


Figura 8.5: Diagrama de Box-plot

Aplicando el diagrama de dispersión:

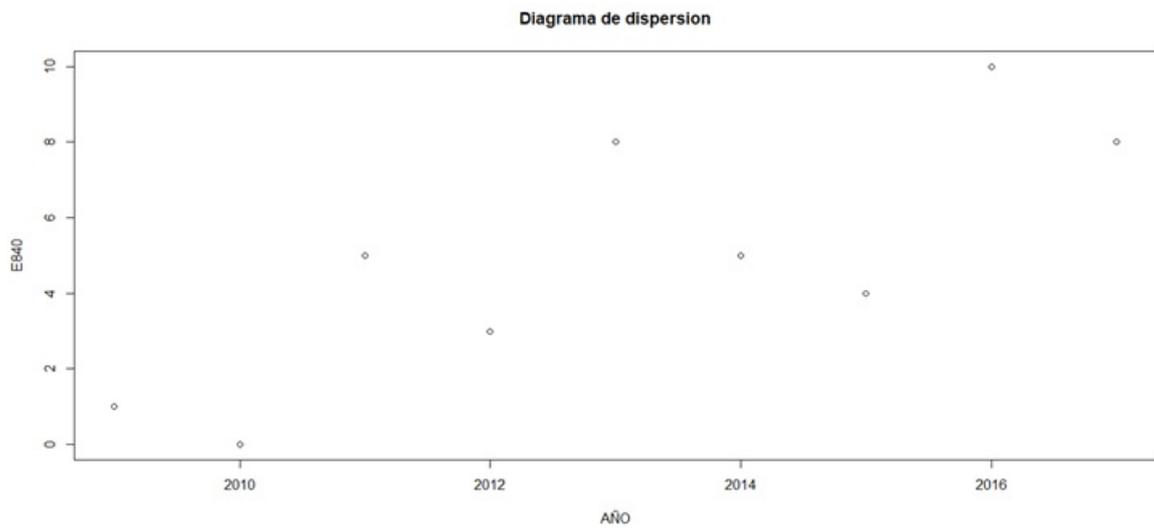


Figura 8.6: Diagrama de dispersión E840

## 8.0.2. E841

El estudio de la **E841-Fibrosis Quística con manifestaciones Intestinales**.

Del cuadro (7-1) se toma la segunda fila, en donde se encuentra el código E841 y se obtiene:

```
> datos<-c(1,0,0,0,1,3,8,1,2)
> #####creamos vector con los años
> año<-c(2009,2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
      año datos
[1,] 2009     1
[2,] 2010     0
[3,] 2011     0
[4,] 2012     0
[5,] 2013     1
[6,] 2014     3
[7,] 2015     8
[8,] 2016     1
[9,] 2017     2
```

Figura 8.7: Fila 2 E841

Se organizan los datos para obtener los resultados:

```
> sort(datos)
[1] 0 0 0 1 1 1 2 3 8
```

Figura 8.8: Datos Organizados

En donde encontramos los siguientes resultados, utilizando las medidas de tendencia central de estadísticas:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   1.000   1.778   2.000   8.000

> var(datos)
[1] 6.444444

> sd(datos)
[1] 2.538591
```

Figura 8.9: Resultados E841

- La Media(Mean): $1,778 \approx 1,8$
- La Mediana(Median):  $1,000 \approx 1$
- Rango: 8
- Valor mínimo(Min): $0,000 \approx 0$
- Valor Máximo(Max): $8,000 \approx 8$
- Primer Cuartil(1st Qu): $0,000 \approx 0$

- Tercer Cuartil(3rd Qu):2,000  $\approx$  2
- Moda: 0 y 1.

Aplicando las medidas de dispersión en estadística:

- Varianza(var): 6,44444  $\approx$  6,4
- Desviación Estandar(sd):2,538591  $\approx$  2,5

Ahora,veamos el código y la gráfica de la frecuencias con la causa E841 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,
legend.text = NULL, beside = FALSE, horiz = FALSE,border = par("fg"),
main ="MORBILIDAD E841", sub = NULL, xlab ='Año', ylab ='Frecuencias')
```

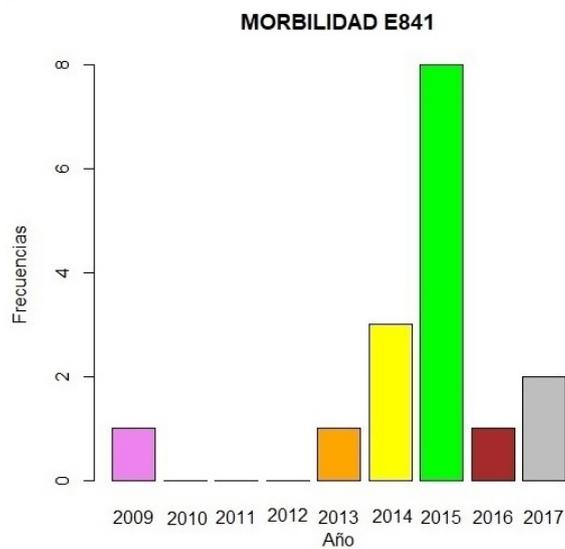


Figura 8.10: Histograma de frecuencias de la E841

Luego, aplicamos método de caja y bigotes (Box-plot) al estudio de la E841.

```
boxplot(datos,main='Morbilidad E841', xlab='años',ylab='numero de pacientes')
```

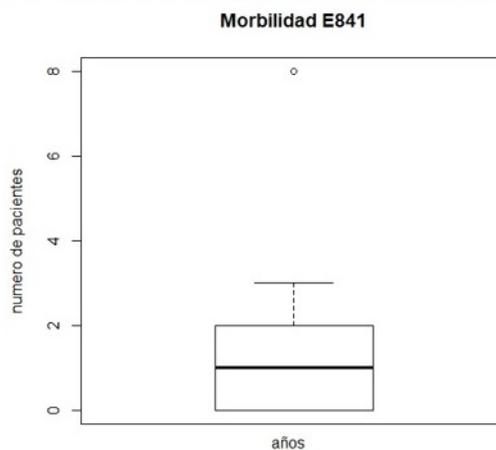


Figura 8.11: Diagrama de Box-plot

Aplicando el diagrama de dispersión para esta patología:

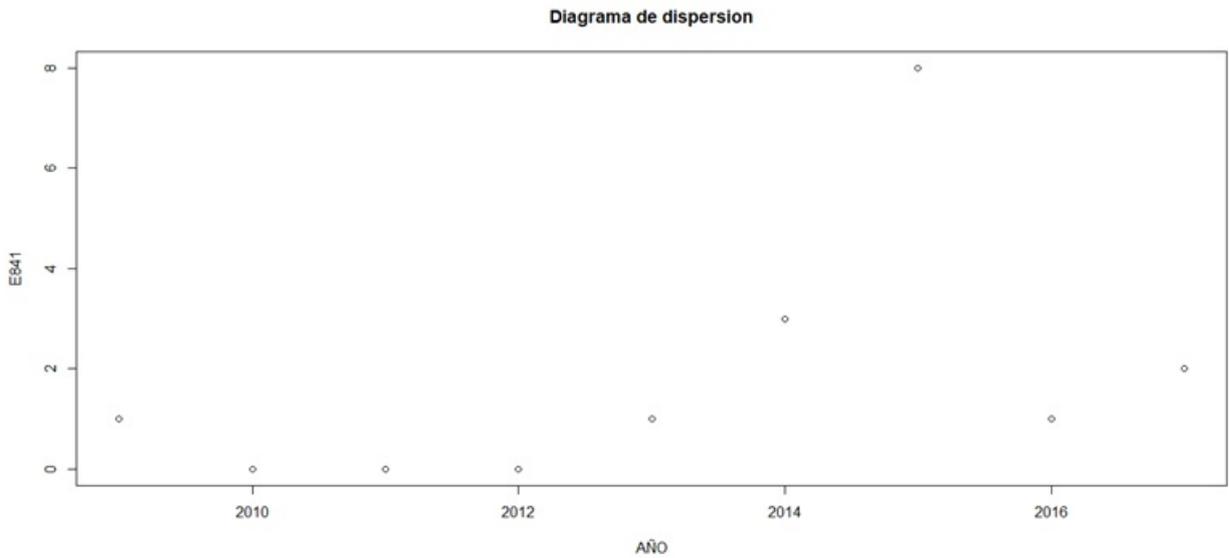


Figura 8.12: Diagrama de dispersion E841

### 8.0.3. E848

El estudio de la **E848-Fibrosis Quística con otras manifestaciones.**

Del cuadro (7-1) se toma la segunda fila, en donde se encuentra el código E848 y se obtiene:

```
> datos<-c(0,4,2,2,5,4,3,1,1)
> #####creamos vector con los años
> año<-c(2009,2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
```

	año	datos
[1,]	2009	0
[2,]	2010	4
[3,]	2011	2
[4,]	2012	2
[5,]	2013	5
[6,]	2014	4
[7,]	2015	3
[8,]	2016	1
[9,]	2017	1

Figura 8.13: Fila 3 E848

Se organizan los datos para poder obtener los resultados.

```
> sort(datos)
[1] 0 1 1 2 2 3 4 4 5
```

Figura 8.14: Datos organizados E848

Luego encontramos los siguientes resultados, utilizando las medidas de tendencias central:

```

> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  1.000   2.000   2.444  4.000   5.000

> var(datos)
[1] 2.777778

> sd(datos)
[1] 1.666667

```

Figura 8.15: Resultados de E848

- La Media(Mean):2,444  $\approx$  2,4
- La Mediana(Median): 2,000  $\approx$  2
- Rango: 5
- Valor mínimo(Min):0,000  $\approx$  0
- Valor Máximo(Max):5,000  $\approx$  5
- Primer Cuartil(1st Qu):1,000  $\approx$  1
- Tercer Cuartil(3rd Qu):4,000  $\approx$  4
- Moda: 1,2 y 4.

Aplicando las medidas de dispersión en estadística:

- Varianza(var): 2,77778  $\approx$  2,8
- Desviación Estandar(sd):1,66667  $\approx$  1,7

Ahora, veamos el código y la gráfica de las frecuencias con la causa E848 con el histograma:

```

barplot(datos, width = 1, space = NULL, names.arg = año,
legend.text = NULL, beside = FALSE, horiz = FALSE, border = par("fg"),
main = "MORBILIDAD E848", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')

```

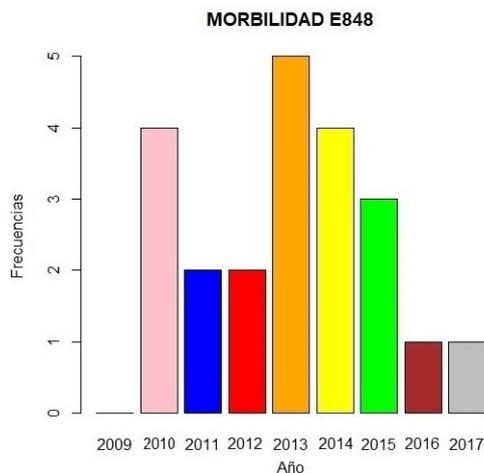


Figura 8.16: Histograma de frecuencias de la E848

Luego, aplicamos el método de caja bigotes(Box-plot)al estudio de la E848.

```
boxplot(datos,main='Morbilidad E848', xlab='años',ylab='numero de pacientes')
```



Figura 8.17: Diagrama de Box-plot

Aplicando el diagrama de dispersión para esta patología:

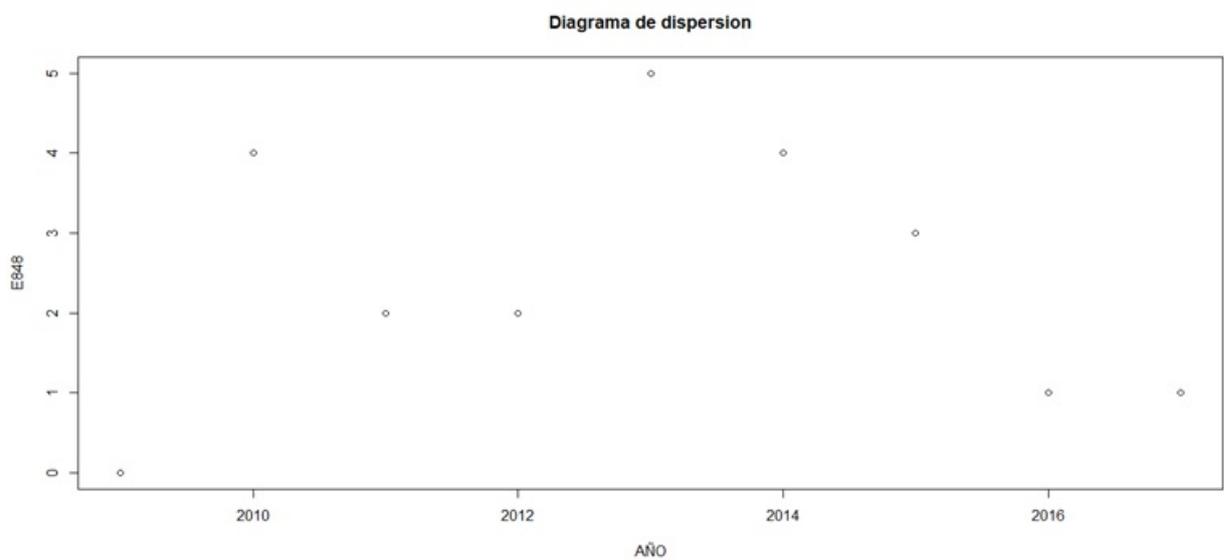


Figura 8.18: Diagrama de dispersión de la E848

## 8.0.4. E849

El estudio de la **E849- Fibrosis Quística sin otras especificaciones** del cuadro (7-1) se toma la cuarta fila, en donde se encuentra el código E849 y se obtiene:

```
> datos<-c(13,19,24,26,34,23,11,16,9)
> #####creamos vector con los años
> año<-c(2009,2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
```

	año	datos
[1,]	2009	13
[2,]	2010	19
[3,]	2011	24
[4,]	2012	26
[5,]	2013	34
[6,]	2014	23
[7,]	2015	11
[8,]	2016	16
[9,]	2017	9

Figura 8.19: Fila 4 E849

Se organizan los datos para obtener los resultados:

```
> sort(datos)
[1] 9 11 13 16 19 23 24 26 34
```

Figura 8.20: Datos organizados

Luego, encontramos los siguientes resultados, utilizando las medidas de tendencias central:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  9.00  13.00   19.00   19.44  24.00   34.00

> var(datos)
[1] 65.27778

> sd(datos)
[1] 8.079466
```

Figura 8.21: Resultados de la E849

- La Media(Mean): $19,44 \approx 19,4$
- La Mediana(Median):  $19,00 \approx 19$
- Rango: 25
- Valor mínimo(Min): $9,00 \approx 9$
- Valor Máximo(Max): $34,00 \approx 34$
- Primer Cuartil(1st Qu): $13,00 \approx 13$
- Tercer Cuartil(3rd Qu): $24,00 \approx 24$
- Moda: no se repiten valores.

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $65,27778 \approx 65,30$
- Desviación Estandar(sd):  $8,079466 \approx 8,0$

Ahora, veamos el código y la gráfica de las frecuencias con la causa E849 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,  
legend.text = NULL, beside = FALSE, horiz = FALSE, border = par("fg"),  
main = "MORBILIDAD E849", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')
```

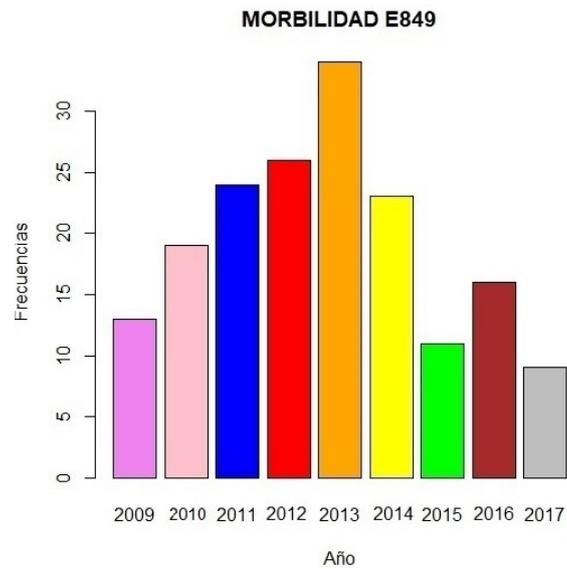


Figura 8.22: Histograma de frecuencias absolutas de morbilidad E849

Luego, aplicamos el método de caja y bigotes (Box-plot) al estudio de la E849.

```
boxplot(datos, main = 'Morbilidad E849', xlab = 'años', ylab = 'numero de pacientes')
```

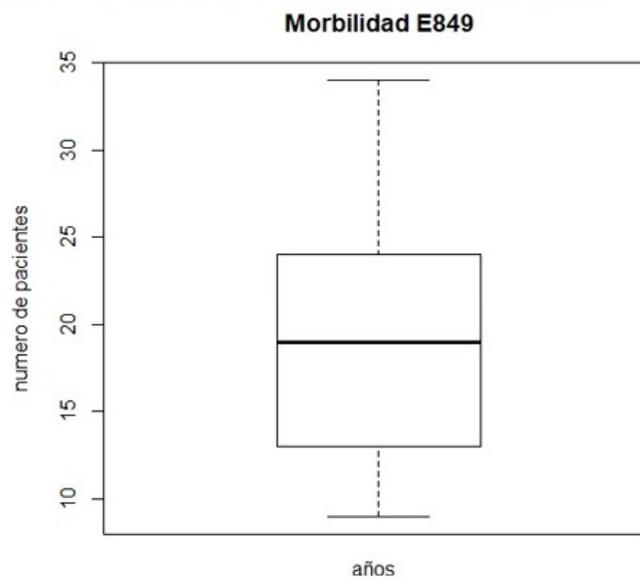


Figura 8.23: Diagrama de Box-plot E849

Aplicando el diagrama de dispersión para esta patología:

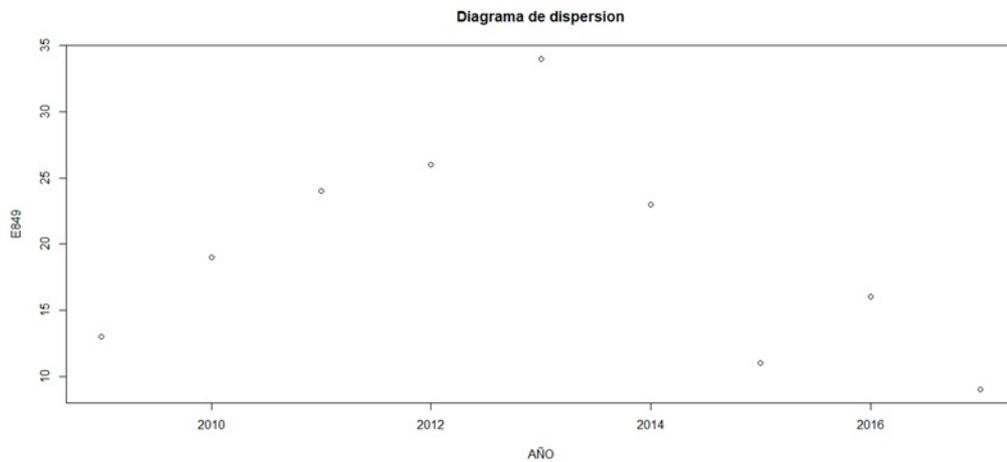


Figura 8.24: Diagrama de dispersión de la E849

Por último, abarcamos todo el análisis de los resultados del primer cuadro de las manifestaciones en personas atendidas por urgencias y por consulta externas, se obtiene lo siguiente :

```
E840<-c(1,0,5,3,8,5,4,10,8)
E841<-c(1,0,0,0,1,3,8,1,2)
E848<-c(0,4,2,2,5,4,3,1,1)
E849<-c(13,19,24,26,34,23,11,16,9)
m1<-cbind(E840,E841,E848,E849)
rownames(z)<-c("2009","2010","2011","2012","2013","2014","2015","2016","2017")
```

	E840	E841	E848	E849
2009	1	1	0	13
2010	0	0	4	19
2011	5	0	2	24
2012	3	0	2	26
2013	8	1	5	34
2014	5	3	4	23
2015	4	8	3	11
2016	10	1	1	16
2017	8	2	1	9

Figura 8.25: Cuadro 1 total de datos

- Media: 7.1
- Mediana: 4
- Moda: 1
- Rango: 34
- Valor mínimo: 0
- Valor Máximo: 34
- Suma total: 257

Aplicando las medidas de dispersión en estadística:

- Desviación Estandar: 8.53
- Varianza: 72.8

Veamos la gráfica total de los histogramas del cuadro (7-1)

```
barplot(m1,beside=T,
legend.text=rownames(m1), ylim=c(-0.5,35),
main='Distribucion de manifestaciones por año',
xlab='Manifestacion',ylab='frecuencia por cantidad de pacientes')
```

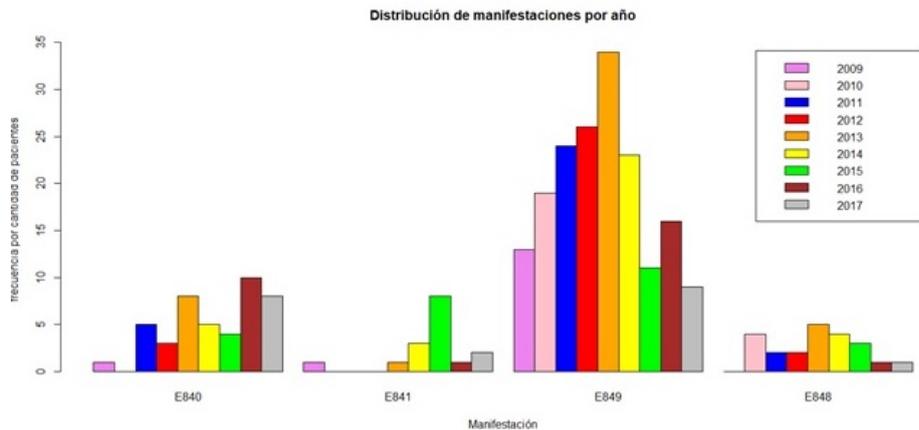


Figura 8.26: Histograma del cuadro 7-1

Veamos la gráfica total del Box-plot:

```
boxplot(m1,main='Morbilidad Personas atendidas \n por urgencias y consulta externa ', xlab='Manifestaciones',ylab='numero de pacientes')
```

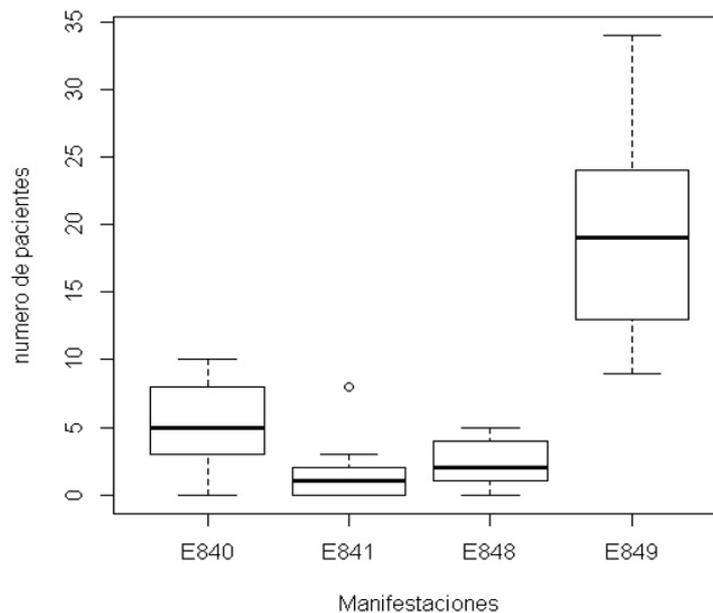


Figura 8.27: Diagrama del Box-Plot Total

Primero se explorará las relaciones entre todas las parejas de variables, en particular la relación de años con cada una de las variables independientes en este caso (E840, E841, E848, E849). Esto lo detectaremos a través de las correlaciones y la función pairs de R-proyect, la cual produce un gráfico matricial para las variables dadas.

```
#####diagrama de dispersion
AÑO<-c(2009,2010,2011,2012,2013,2014,2015,2016,2017)
#Graficas individuales
x11()
plot(E840~AÑO,main="Diagrama de dispersion")
x11()
plot(E841~AÑO,main="Diagrama de dispersion")
x11()
plot(E848~AÑO,main="Diagrama de dispersion")
x11()
plot(E849~AÑO,main="Diagrama de dispersion")
x11()
### todos contra todos
pairs(~E840+E841+AÑO+E848+E849)
```

Figura 8.28: Código para graficar los diagramas de dispersión

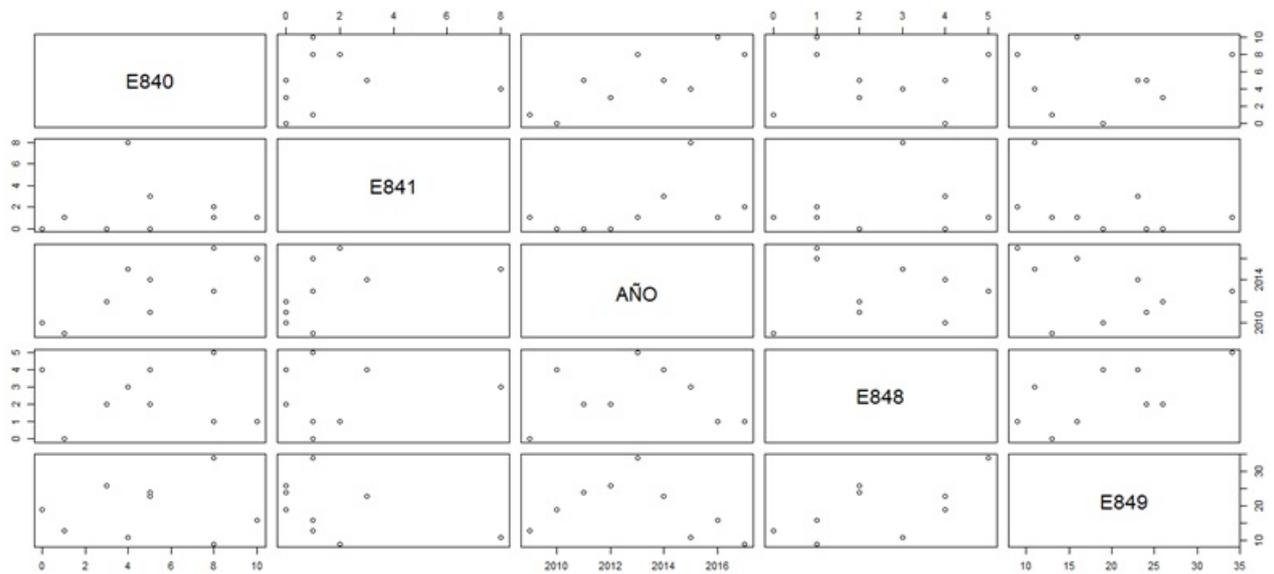


Figura 8.29: Matriz de diagramas de dispersión total

Se realiza un análisis exploratorio a los datos que permitan visualizar la correlación existente entre las variables.

```
> round(cor(mt), 3)
      AÑO  E840  E841  E849  E848
AÑO    1.000  0.794  0.467 -0.305 -0.027
E840   0.794  1.000  0.056  0.086 -0.035
E841   0.467  0.056  1.000 -0.446  0.144
E849  -0.305  0.086 -0.446  1.000  0.633
E848  -0.027 -0.035  0.144  0.633  1.000
```

Figura 8.30: Datos para la correlación

Se hace una regresión lineal entre estas variables.  $E840 \sim AÑO$

```
> regresion1<-lm(E840~AÑO)
> regresion1

Call:
lm(formula = E840 ~ AÑO)

Coefficients:
(Intercept)      AÑO
-1941.0111      0.9667
```

Figura 8.31: Regresión lineal

La recta de regresión lineal que mejor se ajusta es  $y = 0,96x - 1941,01$ .

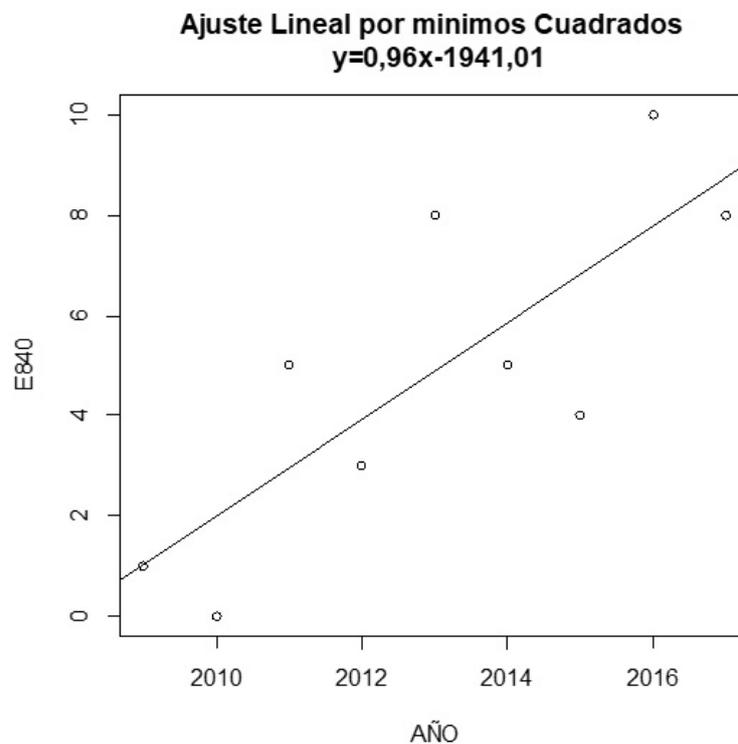


Figura 8.32: Regresión lineal E840 - AÑO

Se hace una regresión lineal entre estas variables. E848~ E849.

```
> regresion2<-lm(E848~E849)
> regresion2

Call:
lm(formula = E848 ~ E849)

Coefficients:
(Intercept)      E849
  -0.09574      0.13064
```

Figura 8.33: Regresión lineal

La recta de regresión lineal que mejor se ajusta es  $y = 0,1306x - 0,09574$

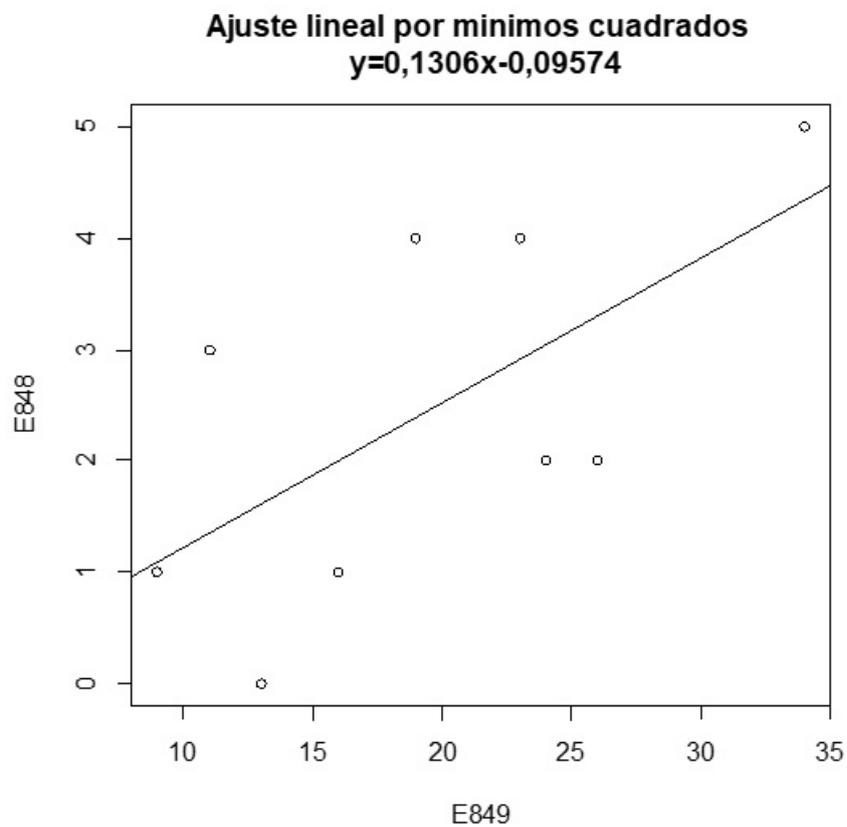


Figura 8.34: Regresión lineal E848 - E849

## 8.1. RESULTADOS DEL ANÁLISIS DEL CUADRO (7-2)

### 8.1.1. E840

Continuando con la utilización del software R PROYECT; se crea el segundo cuadro (7-2) Morbilidad: Numero de personas atendidas en hospitalización con diagnosticos de fibrosis quística; en el Huila 2010 al 2017. Tomando la primera fila en donde se encuentra el código E840, se obtiene:

```
> datos<-c(1,0,1,0,0,1,3,1)
> #####creamos vector con los años
> año<-c(2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
```

	año	datos
[1,]	2010	1
[2,]	2011	0
[3,]	2012	1
[4,]	2013	0
[5,]	2014	0
[6,]	2015	1
[7,]	2016	3
[8,]	2017	1

Figura 8.35: Fila 1 E840

Se organizan los datos para poder obtener los resultados:

```
> sort(datos)
[1] 0 0 0 1 1 1 1 3
```

Figura 8.36: Datos Organizados

En donde, encontramos los siguientes resultados, utilizando las medidas de tendencia central de estadística:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   1.000   0.875   1.000   3.000

> var(datos)
[1] 0.9821429

> sd(datos)
[1] 0.9910312
```

Figura 8.37: Resultados E840

- La Media(Mean): $0,875 \approx 0,88$
- La Mediana(Median):  $1,000 \approx 1$
- Rango: 3
- Valor mínimo(Min): $0,00 \approx 0$
- Valor Máximo(Max): $3,00 \approx 3$
- Primer Cuartil(1st Qu): $0,00 \approx 0$
- Tercer Cuartil(3rd Qu): $1,00 \approx 1$
- Moda: valores 0 y 1.

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $0,9821429 \approx 0,98$
- Desviación Estandar(sd): $0,9910312 \approx 0,99$

Ahora, veamos el código y la gráfica de las frecuencias, con la causa E840 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,
legend.text = NULL, beside = FALSE, horiz = FALSE, border = par("fg"),
main = "MORBILIDAD E840", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')
```

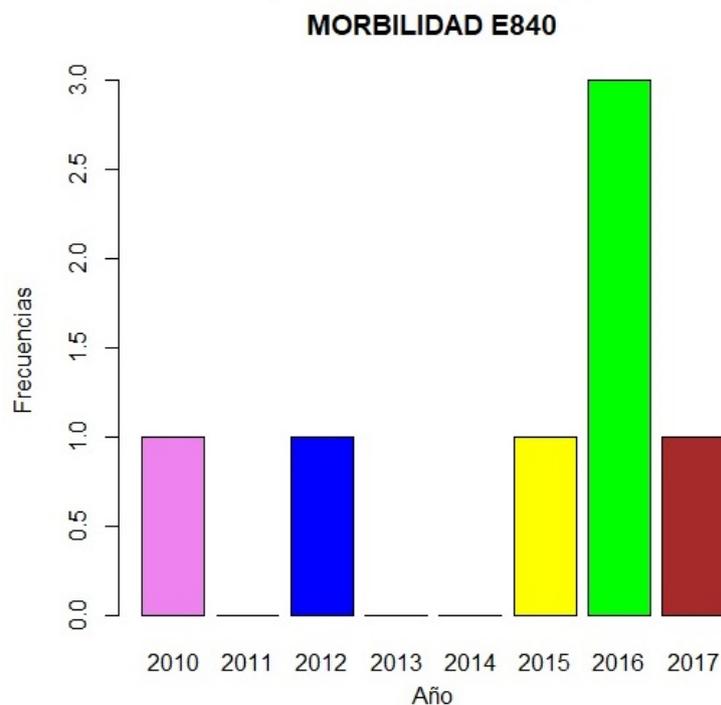


Figura 8.38: Histograma de frecuencias absolutas de Morbilidad E840

Luego, aplicamos el método de caja y bigotes(Box-plot) al estudio de la E840.

```
boxplot(datos,main='Morbilidad E840', xlab='años',ylab='número de pacientes')
```

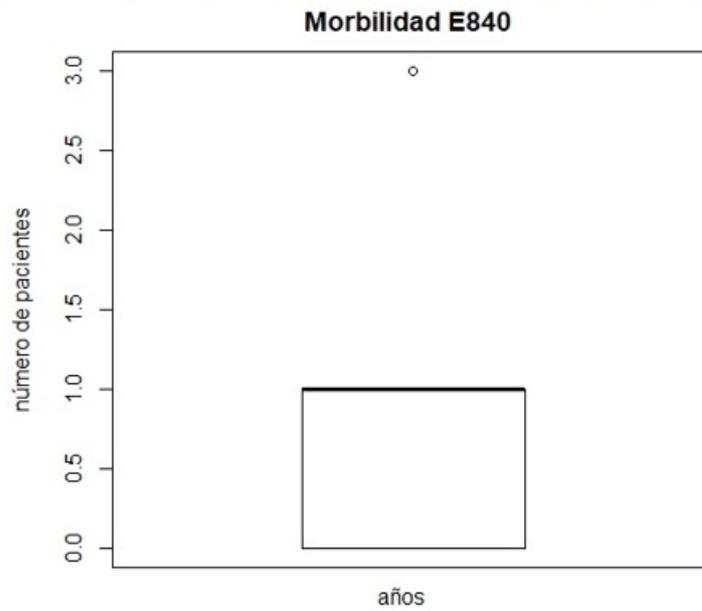


Figura 8.39: Diagrama de Box-plot E840

Aplicando el diagrama de dispersión para el análisis de esta patología:

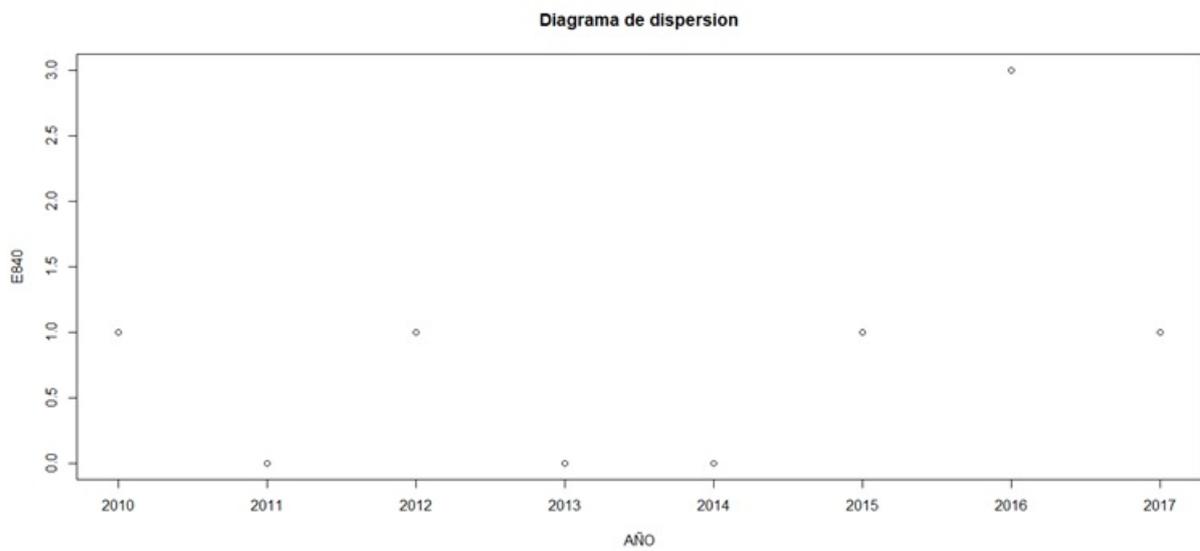


Figura 8.40: Diagrama de dispersión de la E840

## 8.1.2. E841

El estudio de la **E841- Fibrosis Quística con manifestaciones Intestinales** del cuadro(7-2) se toma la segunda fila, en donde se encuentra el código E841 y se obtiene:

```
> datos<-c(0,0,0,1,0,0,0,0)
> #####creamos vector con los años
> año<-c(2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
      año datos
[1,] 2010     0
[2,] 2011     0
[3,] 2012     0
[4,] 2013     1
[5,] 2014     0
[6,] 2015     0
[7,] 2016     0
[8,] 2017     0
```

Figura 8.41: Fila 2 E841

Se organizan los datos para obtener los resultados:

```
> sort(datos)
[1] 0 0 0 0 0 0 0 1
```

Figura 8.42: Datos Organizados

En donde, encontramos los siguientes resultados, utilizando las medidas de tendencia central de estadística:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.000  0.000   0.000   0.125  0.000   1.000

> var(datos)
[1] 0.125

> sd(datos)
[1] 0.3535534
```

Figura 8.43: Resultados de la E841

- La Media(Mean): $0,125 \approx 0,13$
- La Mediana(Median):  $0,000 \approx 0$
- Rango: 1
- Valor mínimo(Min): $0,000 \approx 0$
- Valor Máximo(Max): $1,000 \approx 1$
- Primer Cuartil(1st Qu): $0,000 \approx 0$
- Tercer Cuartil(3rd Qu): $0,00 \approx 0$

- Moda: valores 0 y 1.

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $0,125 \approx 0,13$
- Desviación Estandar(sd):  $0,3535534 \approx 0,35$

Ahora, veamos el código y la gráfica de las frecuencias, con la causa E841 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,
legend.text = NULL, beside = FALSE, horiz = FALSE, border = par("fg"),
main = "MORBILIDAD E841", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')
```

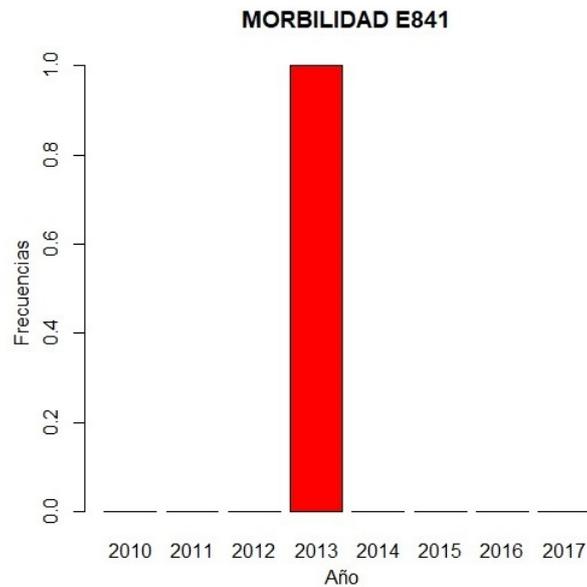


Figura 8.44: Histograma de frecuencias absolutas de Morbilidad E841

Luego, aplicamos el método de caja y bigotes (Box-Plot) al estudio de la E841.

```
boxplot(datos, main = 'Morbilidad E841', xlab = 'años', ylab = 'numero de pacientes')
```

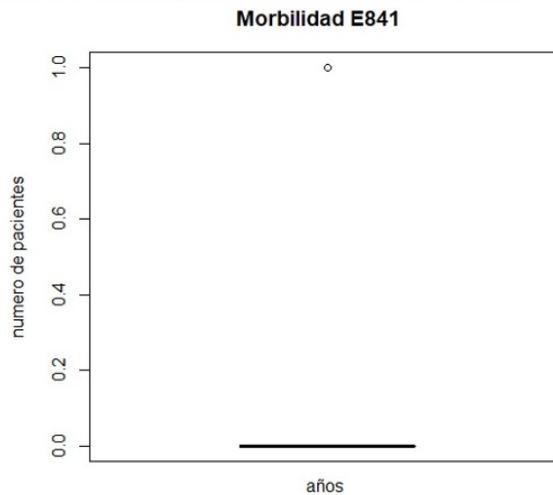


Figura 8.45: Diagrama de Box-plot E841

Aplicando el diagrama de dispersión para esta patología:

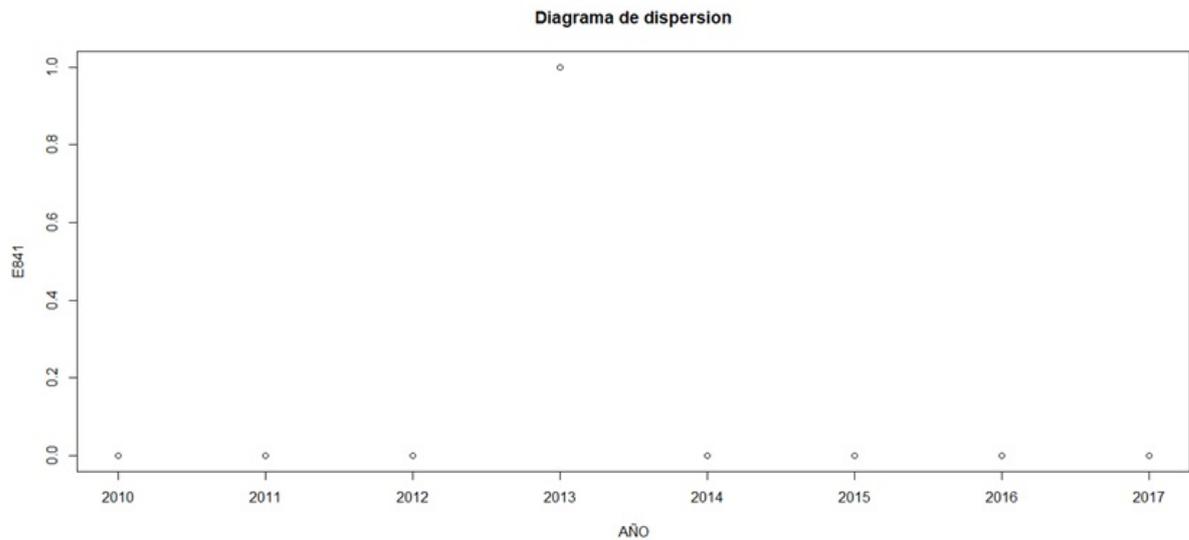


Figura 8.46: Diagrama de dispersión de la E841

### 8.1.3. E848

Del cuadro (7-2) se toma la tercera fila, en donde se encuentra el **Código E848** y se obtiene:

```
> datos<-c(0,0,0,0,1,0,0,0)
> #####creamos vector con los años
> año<-c(2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
```

	año	datos
[1,]	2010	0
[2,]	2011	0
[3,]	2012	0
[4,]	2013	0
[5,]	2014	1
[6,]	2015	0
[7,]	2016	0
[8,]	2017	0

Figura 8.47: Fila 3 E848

Se organizan los datos para poder obtener los resultados.

```
> sort(datos)
[1] 0 0 0 0 0 0 0 1
```

Figura 8.48: Datos organizados E848

Luego, encontramos los siguientes resultados, utilizando las medidas de tendencia central:

```

> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   0.000   0.125  0.000   1.000

> var(datos)
[1] 0.125

> sd(datos)
[1] 0.3535534

```

Figura 8.49: Resultados de la E848

- La Media(Mean): $0,125 \approx 0,13$
- La Mediana(Median):  $0,000 \approx 0$
- Rango: 1
- Valor mínimo(Min): $0,000 \approx 0$
- Valor Máximo(Max): $1,000 \approx 1$
- Primer Cuartil(1st Qu): $0,000 \approx 0$
- Tercer Cuartil(3rd Qu): $0,00 \approx 0$
- Moda: valores 0 .

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $0,125 \approx 0,13$
- Desviación Estandar(sd): $0,3535534 \approx 0,35$

Ahora, veamos el código y la gráfica de las frecuencias, con la causa E848 con el histograma:

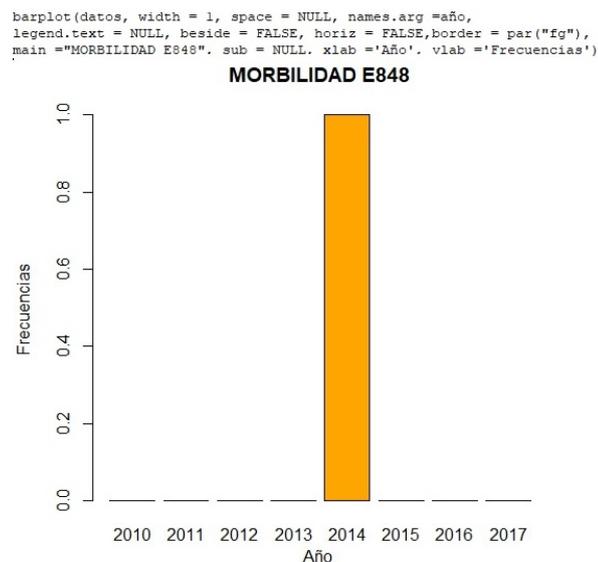


Figura 8.50: Histograma de frecuencias absolutas de Morbilidad E848

Luego, aplicamos el método de caja y bigotes (Box-plot) al estudio de la E848.

```
boxplot(datos,main='Morbilidad E848', xlab='años',ylab='numero de pacientes')
```

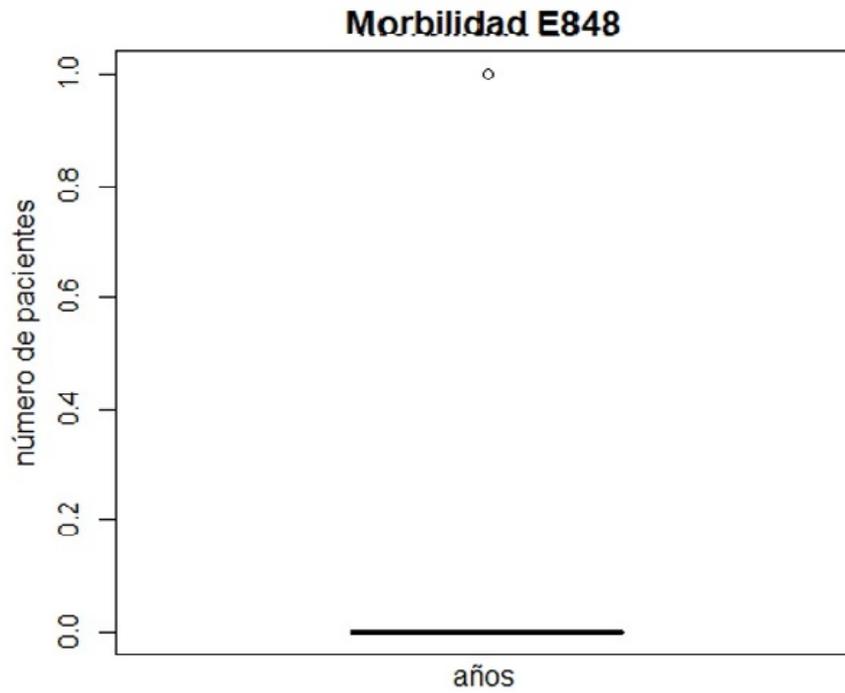


Figura 8.51: Diagrama de Box-plot E848

Aplicando el diagrama de dispersión para esta patología:

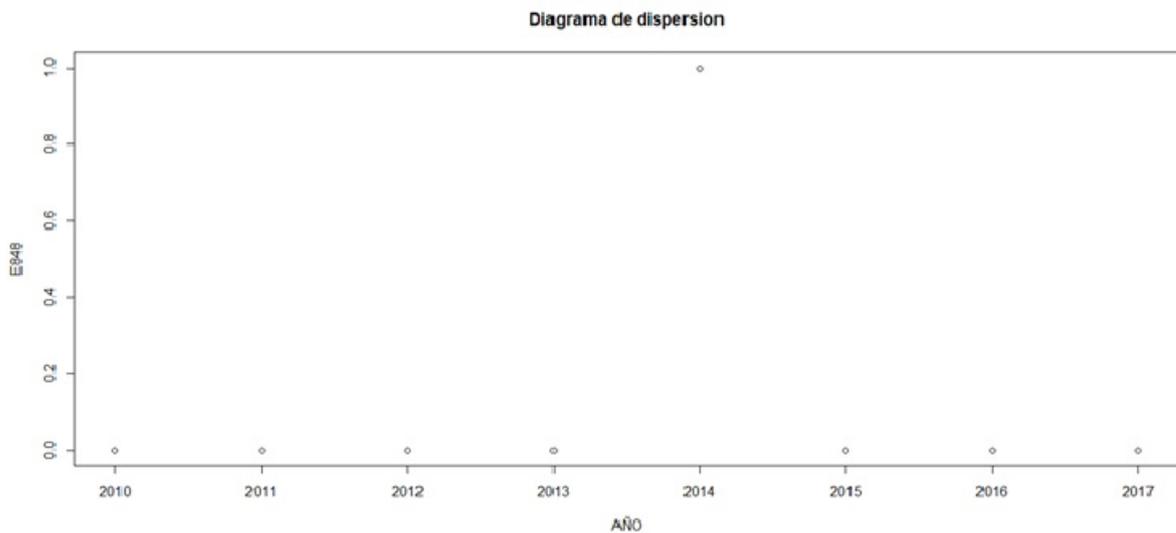


Figura 8.52: Diagrama de Dispersión de la E848

### 8.1.4. E849

El estudio de la **E849- Fibrosis quística sin otras especificaciones** del cuadro (7-2) se toma la cuarta fila, en donde se encuentra el código E849 y se obtiene:

```
> datos<-c(1,1,1,0,1,1,1,0)
> #####creamos vector con los años
> año<-c(2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
```

	año	datos
[1,]	2010	1
[2,]	2011	1
[3,]	2012	1
[4,]	2013	0
[5,]	2014	1
[6,]	2015	1
[7,]	2016	1
[8,]	2017	0

Figura 8.53: Fila 4 E849

Se organizan los datos para poder obtener los resultados:

```
> sort(datos)
[1] 0 0 1 1 1 1 1 1
```

Figura 8.54: Datos organizados E849

Luego, encontramos los siguientes resultados, utilizando las medidas de tendencias central:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00   0.75   1.00   0.75   1.00   1.00

> var(datos)
[1] 0.2142857

> sd(datos)
[1] 0.46291
```

Figura 8.55: Resultados de la E849

- La Media(Mean): $0,75 \approx 0,8$
- La Mediana(Median):  $1,00 \approx 1$
- Rango: 1
- Valor mínimo(Min): $0,00 \approx 0$
- Valor Máximo(Max): $1,00 \approx 1$
- Primer Cuartil(1st Qu): $0,75 \approx 0,8$
- Tercer Cuartil(3rd Qu): $1,00 \approx 1$
- Moda: valores 1.

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $0,2142857 \approx 0,21$
- Desviación Estandar(sd):  $0,46291 \approx 0,46$

Ahora, veamos el código y la gráfica de las frecuencias, con la causa E849 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,  
legend.text = NULL, beside = FALSE, horiz = FALSE, border = par("fg"),  
main = "MORBILIDAD E849", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')
```

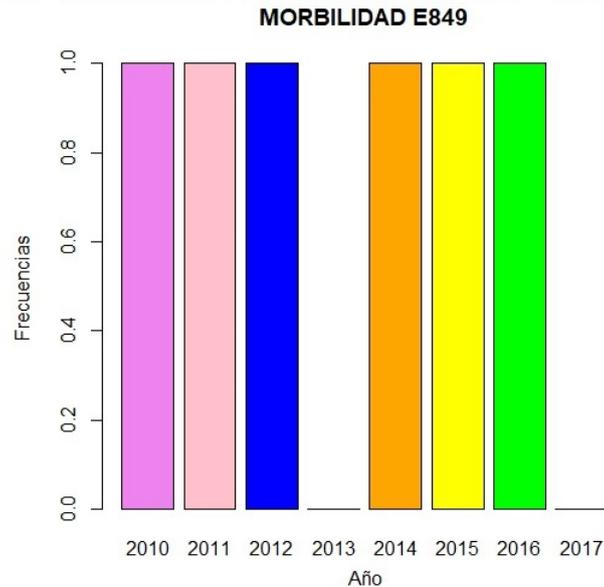


Figura 8.56: Histograma de frecuencias absolutas de morbilidad E849

Luego, aplicamos el método de caja y bigotes (Box-plot) al estudio de la E849.

```
boxplot(datos, main = 'Morbilidad E849', xlab = 'años', ylab = 'numero de pacientes')
```

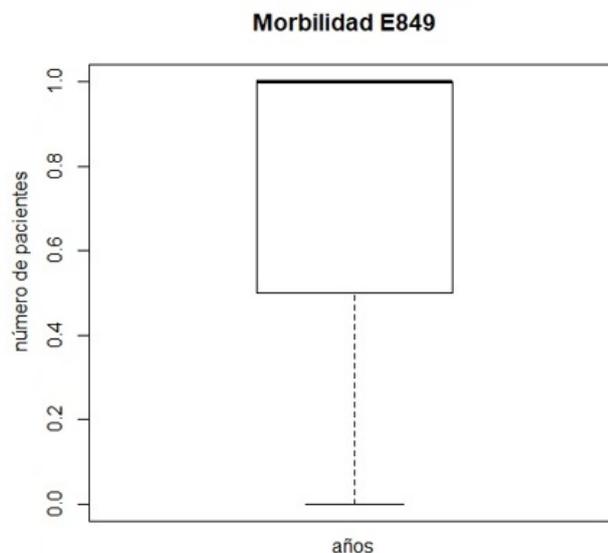


Figura 8.57: Diagrama de Box-plot E849

Aplicando el diagrama de dispersión para esta patología:

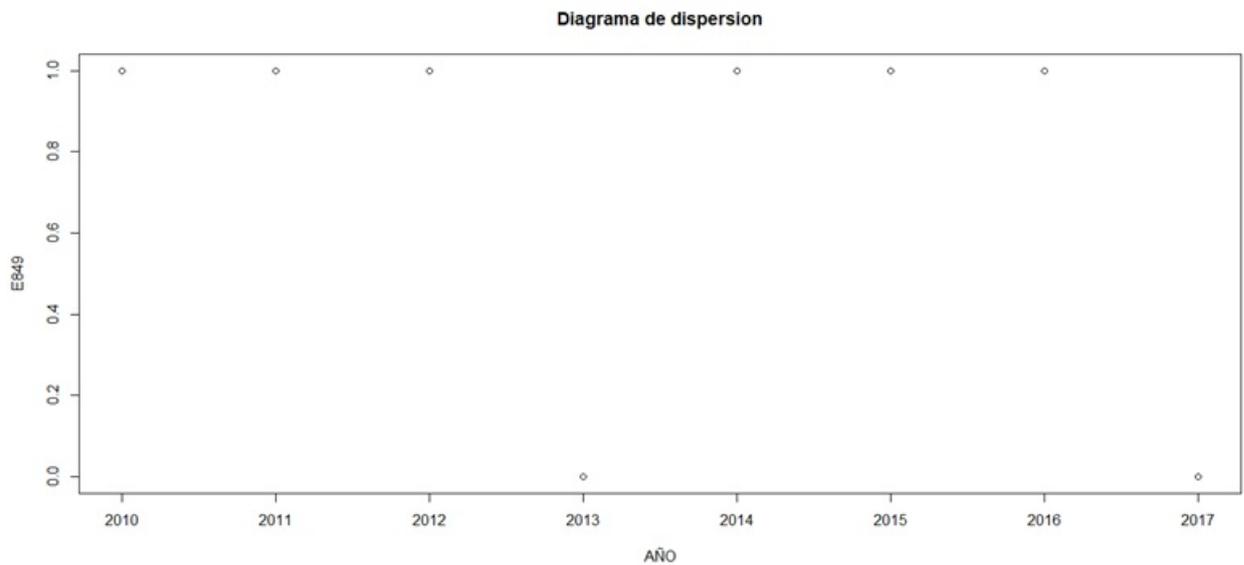


Figura 8.58: Diagrama de dispersión de la E849

Por último, abarcamos todo el análisis de los resultados del segundo cuadro de las manifestaciones en el número de personas atendidas en hospitalización con diagnóstico de fibrosis quística en el Huila durante los años 2010-2017.

```
> ml<-cbind(E840,E841,E848,E849)
> rownames(ml)<-c("2010","2011","2012","2013","2014","2015","2016","2017")
> ml
```

	E840	E841	E848	E849
2010	1	0	0	1
2011	0	0	0	1
2012	1	0	0	1
2013	0	1	0	0
2014	0	0	1	1
2015	1	0	0	1
2016	3	0	0	1
2017	1	0	0	0

Figura 8.59: Cuadro 2 total de datos

- La Media:0,5
- Rango: 3
- Valor mínimo(Min):0
- Valor Máximo(Max):3
- Suma total: 15

Aplicando las medidas de dispersión en estadística:

- Desviación Estandar(sd):1

Veamos la gráfica total de los histogramas del cuadro (7-2).

```
x11()
barplot(m1,beside=T,
legend.text=rownames(m1), ylim=c(-0.5,4),
main='Distribucion de manifestaciones en personas atendidas \n en Hospitalización por año',
xlab='Manifestacion'.vlab='frecuencia por cantidad de pacientes')
```

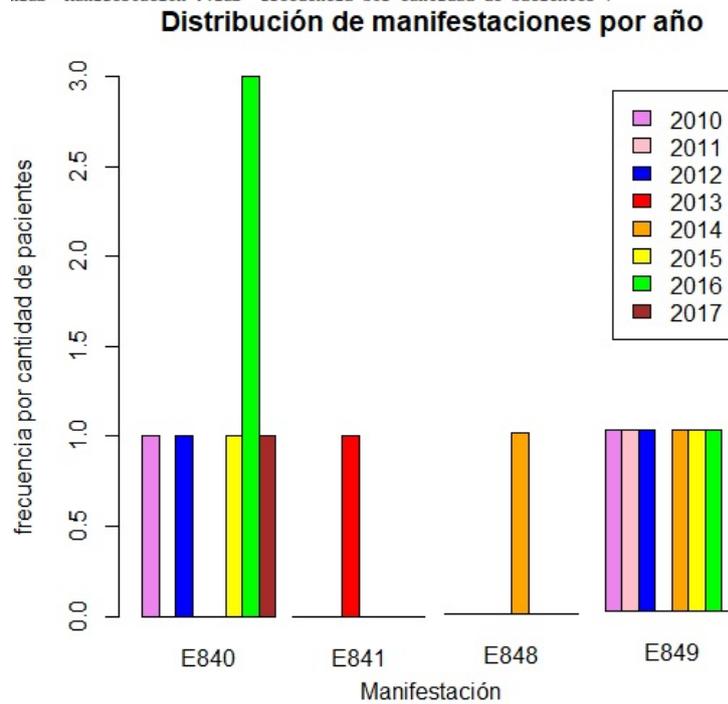


Figura 8.60: Histograma de frecuencias absolutas del cuadro (7-2)

Aplicamos el método de caja y bigotes (Box-plot)al estudio total del cuadro (7-2).

```
boxplot(m1,main='Morbilidad Personas atendidas \n en Hospitalización ', xlab='Manifestaciones',ylab='número de pacientes')
```

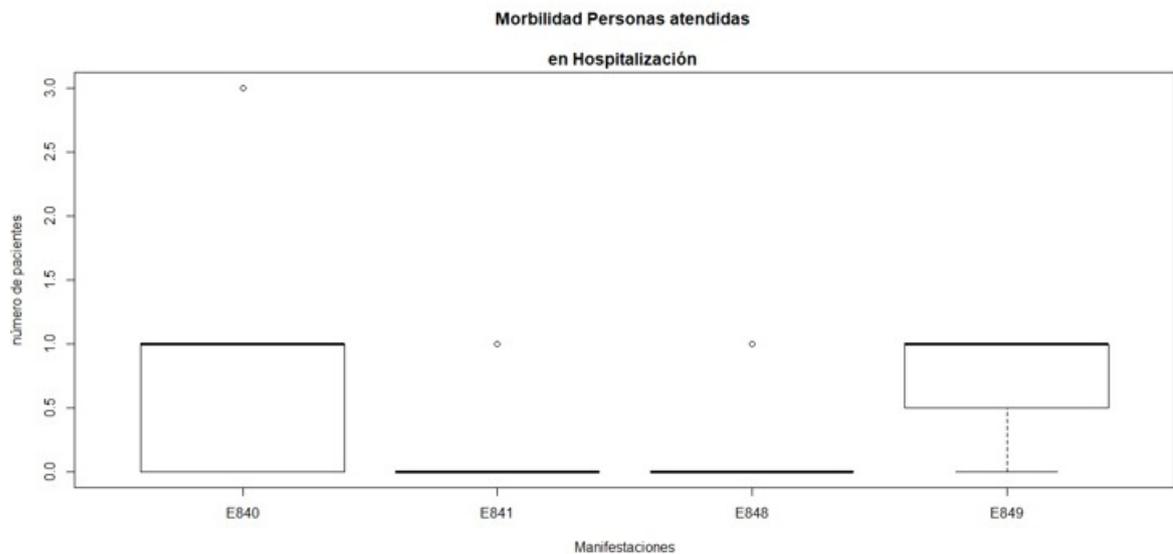


Figura 8.61: Diagrama total del Box-plot

Se explorará las relaciones entre todas las parejas de variables, en particular la relación de años con cada una de las variables independientes en este caso (E840, E841, E848, E849). Esto lo detectaremos a través de las correlaciones y la función pairs de R-proyect, la cual produce un gráfico matricial para las variables dadas.

```
#####diagrama de dispersion
AÑO<-c(2010,2011,2012,2013,2014,2015,2016,2017)
#Graficas individuales
x11()
plot(E840~AÑO,main="Diagrama de dispersion")
x11()
plot(E841~AÑO,main="Diagrama de dispersion")
x11()
plot(E848~AÑO,main="Diagrama de dispersion")
x11()
plot(E849~AÑO,main="Diagrama de dispersion")
x11()
#### todos contra todos
pairs(~E840+E841+AÑO+E848+E849) |
```

Figura 8.62: Código para los diagramas de dispersiones

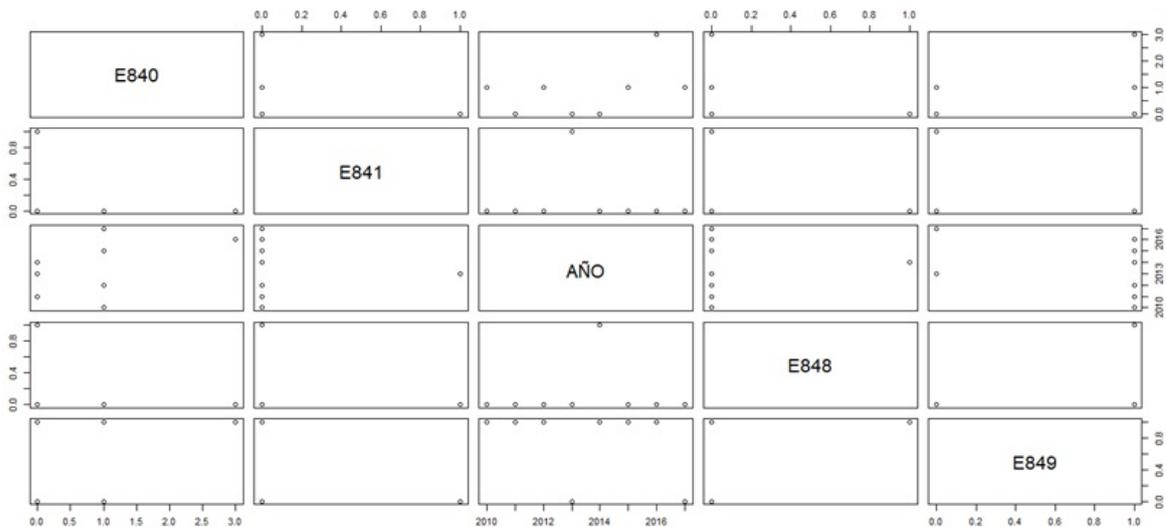


Figura 8.63: Matriz de diagramas de dispersiones

Se realiza un análisis exploratorio a los datos que permitan visualizar la correlación existente entre las variables.

Se hace una regresión lineal entre estas variables. E841~ E849.

La recta de regresión lineal que mejor se ajusta es  $y = 0,5 - 0,5x$ .

```
> round(cor(mt), 3)
      AÑO  E840  E841  E849  E848
AÑO    1.000  0.393 -0.137 -0.207  0.000
E840   0.393  1.000 -0.359  0.238 -0.359
E841  -0.137 -0.359  1.000 -0.661 -0.125
E849  -0.207  0.238 -0.661  1.000  0.189
E848   0.000 -0.359 -0.125  0.189  1.000
```

Figura 8.64: Datos para la correlación

```
> regresion1<-lm(E841~E849)
> regresion1

Call:
lm(formula = E841 ~ E849)

Coefficients:
(Intercept)      E849
          0.5         -0.5
```

Figura 8.65: Regresión Lineal

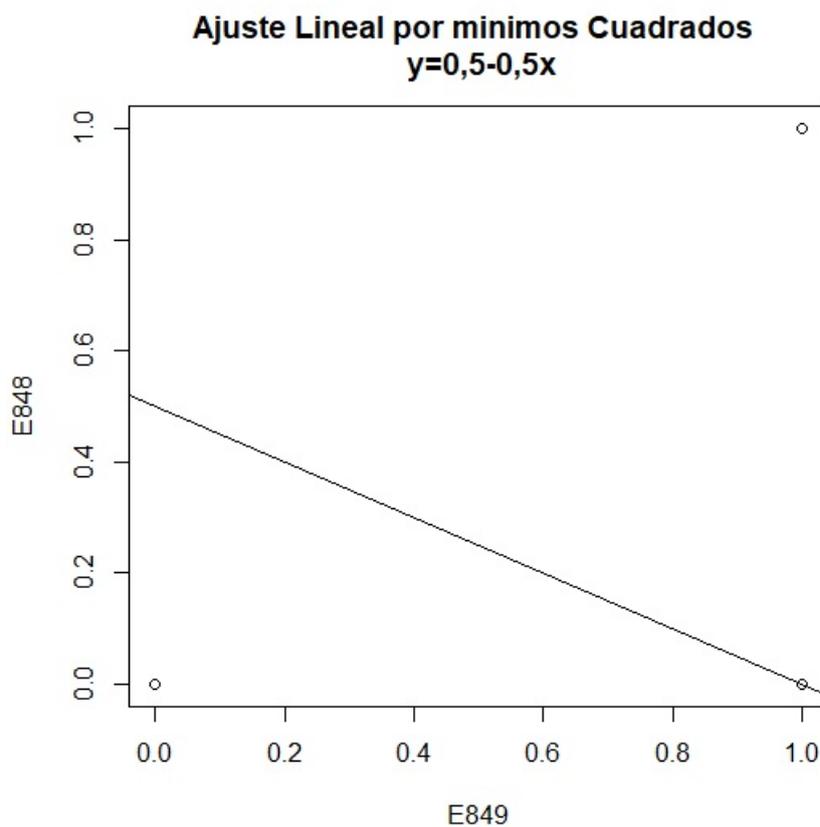


Figura 8.66: Regresión lineal E841 - E849

## 8.2. RESULTADOS DEL ANÁLISIS DEL CUADRO (7-3)

### 8.2.1. E840

Utilizando el software R PROYECT; continuamos creando el tercer cuadro (7-3) **Mortalidad: Número de personas fallecidas con diagnóstico de Fibrosis Quística; Huila 2010-2017.**

```
> datos<-c(0,0,0,1,0,0,0,1)
> #####creamos vector con los años
> año<-c(2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
      año datos
[1,] 2010     0
[2,] 2011     0
[3,] 2012     0
[4,] 2013     1
[5,] 2014     0
[6,] 2015     0
[7,] 2016     0
[8,] 2017     1
```

Figura 8.67: Fila 1 E840

Se organizan los datos para un mejor resultado:

```
> sort(datos)
[1] 0 0 0 0 0 0 1 1
```

Figura 8.68: Datos Organizados E840

Luego, encontramos los siguientes resultados, utilizando las medidas de tendencias central:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.00  0.00   0.00   0.25  0.25   1.00

> var(datos)
[1] 0.2142857

> sd(datos)
[1] 0.46291
```

Figura 8.69: Resultados de la E840

- La Media(Mean): $0,25 \approx 0,3$
- La Mediana(Median):  $0,00 \approx 0$
- Rango: 1
- Valor mínimo(Min): $0,00 \approx 0$
- Valor Máximo(Max): $1,00 \approx 1$
- Primer Cuartil(1st Qu): $0,00 \approx 0$
- Tercer Cuartil(3rd Qu): $0,25 \approx 0,3$

- Moda: valores 0 y 1.

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $0,2142857 \approx 0,21$
- Desviación Estandar(sd):  $0,46291 \approx 0,46$

Ahora, veamos el código y la gráfica de las frecuencias, con la causa E840 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,
legend.text = NULL, beside = FALSE, horiz = FALSE, border = par("fg"),
main = "MORTALIDAD E840", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')
```

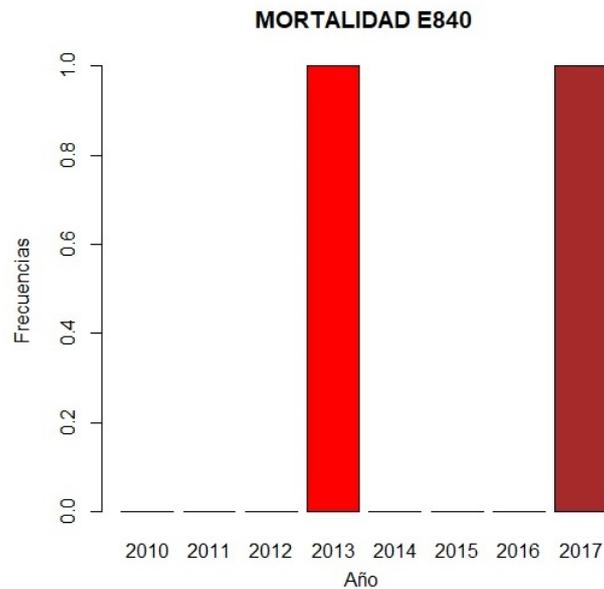


Figura 8.70: Histograma de frecuencias absolutas de Mortalidad E840

Aplicamos el método de caja y bigotes (Box-plot) al E840.

```
boxplot(datos, main = 'Mortalidad E840', xlab = 'años', ylab = 'número de personas fallecidas')
```

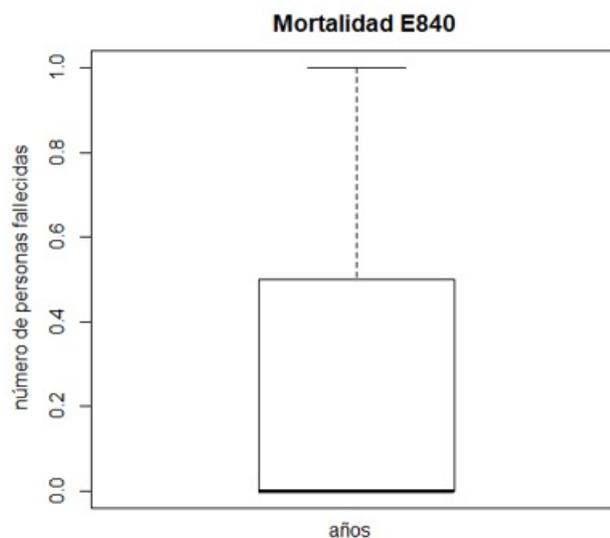


Figura 8.71: Diagrama de Box-plot E840

Aplicamos el diagrama de dispersión para el estudio de esta patología:

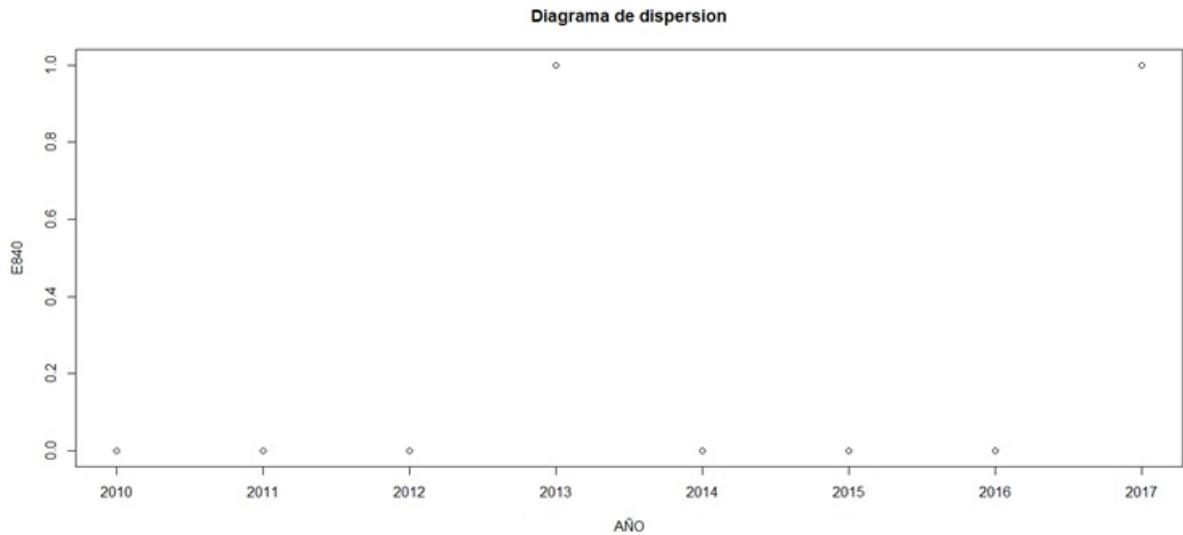


Figura 8.72: Diagrama de dispersion de la E840

## 8.2.2. E841

El estudio de la **E841 - Fibrosis quística con manifestaciones Intestinales** del cuadro(7-3)Se toma la segunda fila, en donde se encuentra el codigo E841 y se obtiene:

```
> datos<-c(0,0,0,0,0,0,1,0)
> #####creamos vector con los años
> año<-c(2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
```

	año	datos
[1,]	2010	0
[2,]	2011	0
[3,]	2012	0
[4,]	2013	0
[5,]	2014	0
[6,]	2015	0
[7,]	2016	1
[8,]	2017	0

Figura 8.73: Fila 2 E841

Se organizan los datos para obtener los resultados:

```
> sort(datos)
[1] 0 0 0 0 0 0 0 1
```

Figura 8.74: Datos organizados

Luego, encontramos los siguientes resultados, utilizando las medidas de tendencias central:

```

> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.000  0.000   0.000   0.125  0.000   1.000

> var(datos)
[1] 0.125

> sd(datos)
[1] 0.3535534

```

Figura 8.75: Resultados de la E841

- La Media(Mean): $0,125 \approx 0,13$
- La Mediana(Median):  $0,00 \approx 0$
- Rango: 1
- Valor mínimo(Min): $0,00 \approx 0$
- Valor Máximo(Max): $1,00 \approx 1$
- Primer Cuartil(1st Qu): $0,00 \approx 0$
- Tercer Cuartil(3rd Qu): $0,00 \approx 0,0$
- Moda: valor 0 .

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $0,125 \approx 0,13$
- Desviación Estandar(sd): $0,3535534 \approx 0,35$

Ahora, veamos el código y la gráfica de las frecuencias, con la causa E841 con el histograma:

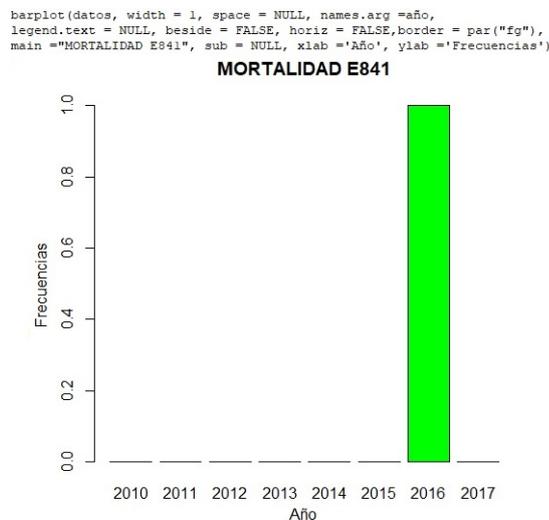


Figura 8.76: Histograma de frecuencias absolutas de Mortalidad E841

Aplicamos el método de caja y bigotes (Box-plot) al E841.

```
boxplot(datos,main='Mortalidad E841', xlab='años',ylab='número de personas fallecidas')
```

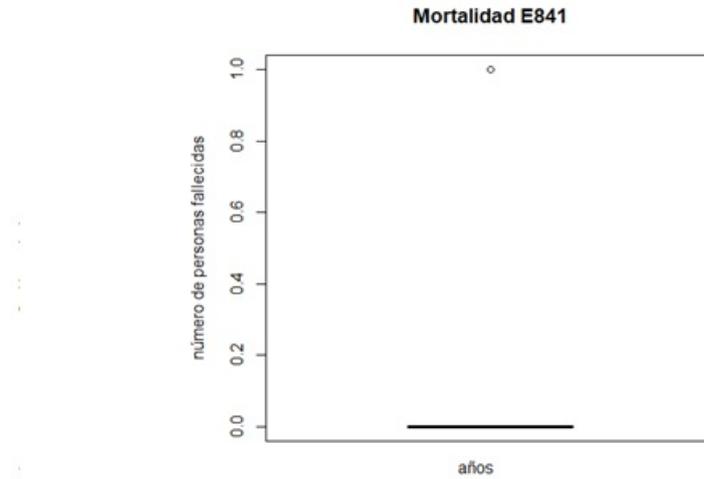


Figura 8.77: Diagrama de Box-plot E841

Aplicamos el diagrama de dispersión para el estudio de esta causa:

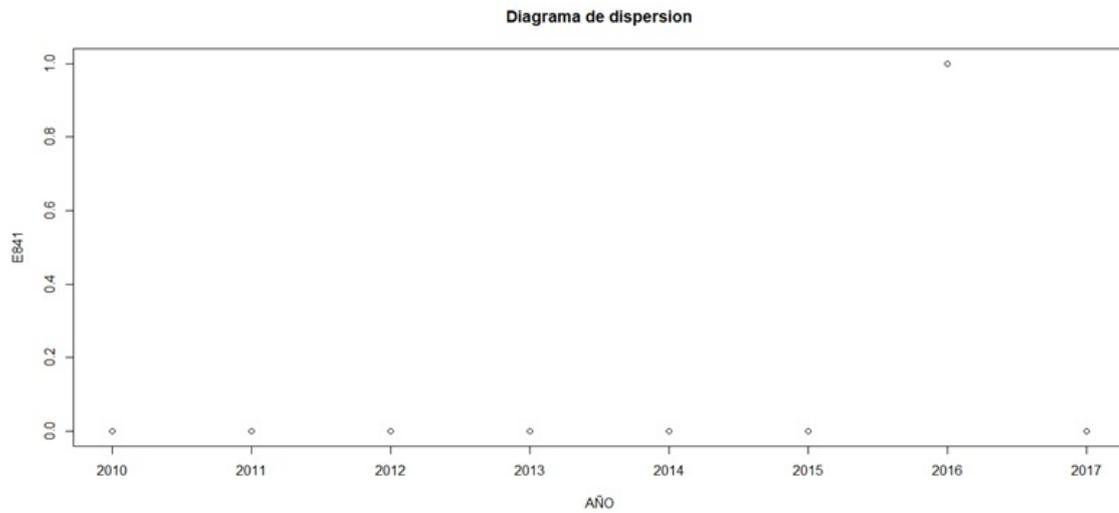


Figura 8.78: Diagrama de dispersión

### 8.2.3. E849

Del cuadro (7-3) se toma la tercera fila, en donde se encuentran el Código E849 y se obtiene:

```
> datos<-c(1,1,0,0,0,0,0,0)
> #####creamos vector con los años
> año<-c(2010,2011,2012,2013,2014,2015,2016,2017)
> ##### creamos tabla vertical con año y datos
> z<-cbind(año,datos)
> z
      año datos
[1,] 2010     1
[2,] 2011     1
[3,] 2012     0
[4,] 2013     0
[5,] 2014     0
[6,] 2015     0
[7,] 2016     0
[8,] 2017     0
```

Figura 8.79: Fila 3 E849

Se organizan los datos para obtener los resultados:

```
> sort(datos)
[1] 0 0 0 0 0 0 1 1
```

Figura 8.80: Datos organizados

Luego, encontramos los siguientes resultados, utilizando las medidas de tendencias central:

```
> summary(datos)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.00  0.00    0.00   0.25   0.25   1.00
> var(datos)
[1] 0.2142857

> sd(datos)
[1] 0.46291
```

Figura 8.81: Resultados de la E849

- La Media(Mean):0,25
- La Mediana(Median): 0,00  $\approx$  0
- Rango: 1
- Valor mínimo(Min):0,00  $\approx$  0
- Valor Máximo(Max):1,00  $\approx$  1
- Primer Cuartil(1st Qu):0,00  $\approx$  0
- Tercer Cuartil(3rd Qu):0,25
- Moda: valores 0 y 1 .

Aplicando las medidas de dispersión en estadística:

- Varianza(var):  $0,2142857 \approx 0,21$
- Desviación Estandar(sd):  $0,46291 \approx 0,46$

Ahora, veamos el código y la gráfica de las frecuencias, con la causa E849 con el histograma:

```
barplot(datos, width = 1, space = NULL, names.arg = año,
legend.text = NULL, beside = FALSE, horiz = FALSE, border = par("fg"),
main = "MORTALIDAD E849", sub = NULL, xlab = 'Año', ylab = 'Frecuencias')
```

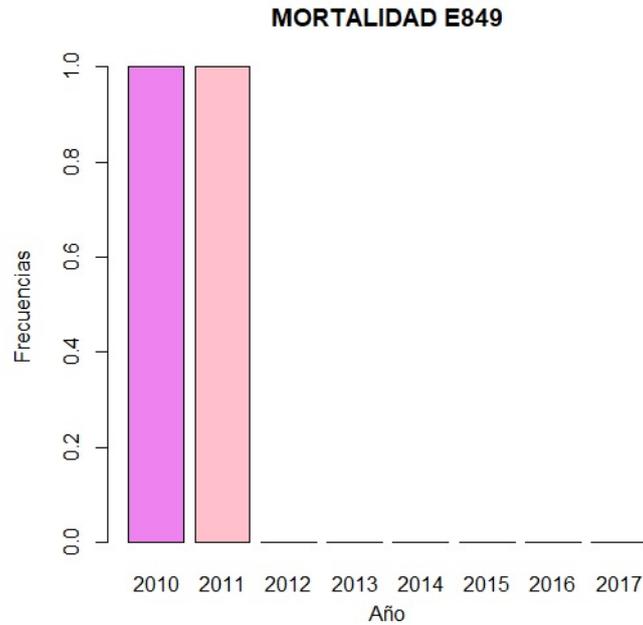


Figura 8.82: Histograma de frecuencias absolutas de Mortalidad E849

Aplicamos el método de caja y bigotes (Box-plot) al E849.

```
boxplot(datos, main = 'Mortalidad E849', xlab = 'años', ylab = 'número de personas fallecidas')
```

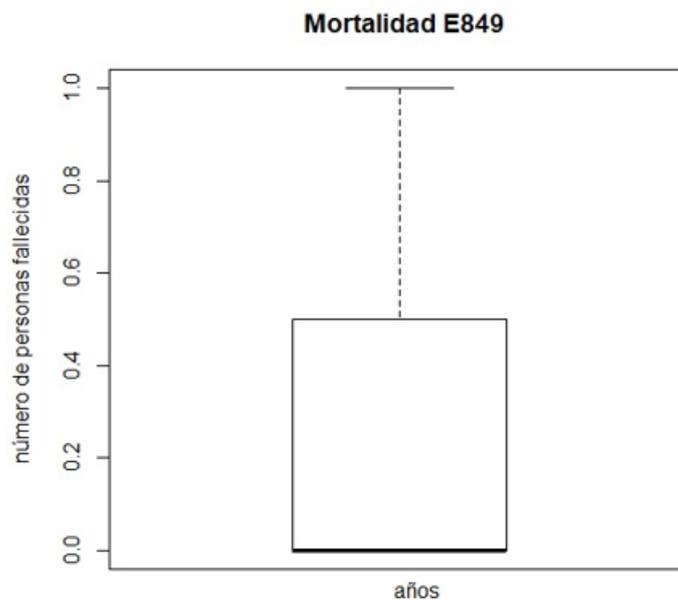


Figura 8.83: Diagrama de Box-plot E849

Aplicando el diagrama de dispersión para el estudio de esta causa:

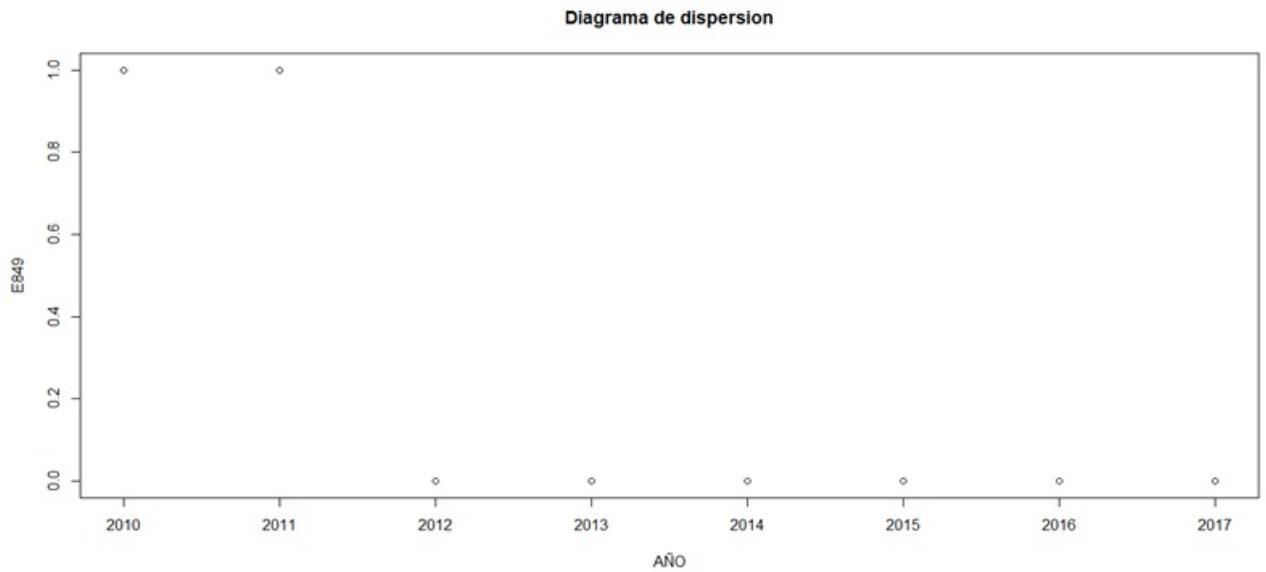


Figura 8.84: Diagrama de dispersión de la E849

Por último, abarcamos todo el análisis de los resultados del tercer cuadro Mortalidad: Número de personas fallecidas con diagnósticos de fibrosis quística Huila 2010-2017.

```
> ml<-cbind(E840,E841,E849)
> rownames(ml)<-c("2010","2011","2012","2013","2014","2015","2016","2017")
> ml
```

	E840	E841	E849
2010	0	0	1
2011	0	0	1
2012	0	0	0
2013	1	0	0
2014	0	0	0
2015	0	0	0
2016	0	1	0
2017	1	0	0

Figura 8.85: Cuadro 3 total de datos

- La Media: 0.33
- La Mediana: 0
- Rango: 1
- Valor mínimo: 0
- Valor Máximo: 1
- Moda: 0

Aplicando las medidas de dispersión en estadística:

- Varianza: 0.23
- Desviación Estandar: 0.48

Ahora, veamos el código y la gráfica de las frecuencias, con las causas totales de mortalidad de la fibrosis quística con el histograma:

```
barplot(ml,beside=T,
legend.text=rownames(ml), ylim=c(-0.5,4),
main='Distribución de Mortalidad',
xlab='Manifestación',ylab='frecuencia por cantidad de personas fallecidas')
```

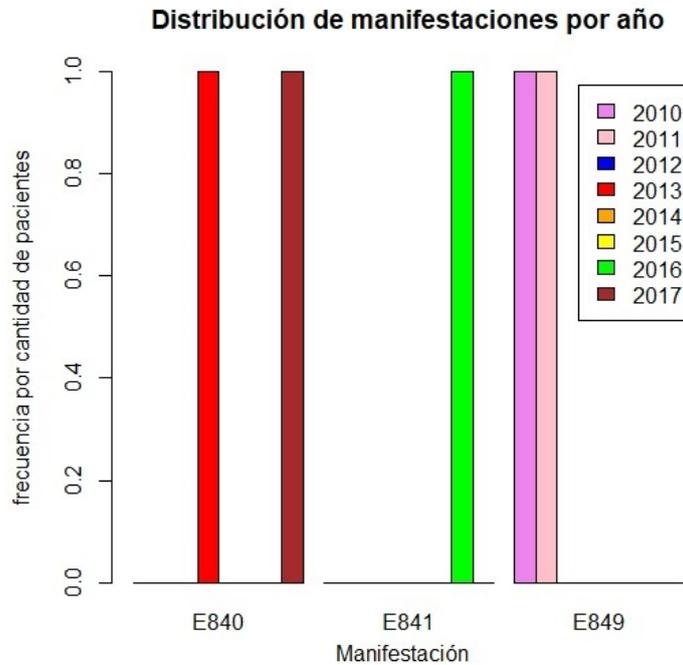


Figura 8.86: Histograma de frecuencias absolutas de Mortalidad

Aplicamos el método de caja y bigotes (Box-plot) a las causas totales de mortalidad de la FQ.

```
boxplot(ml,main='Mortalidad', xlab='Manifestaciones',
ylab='número de personas fallecidas')
```

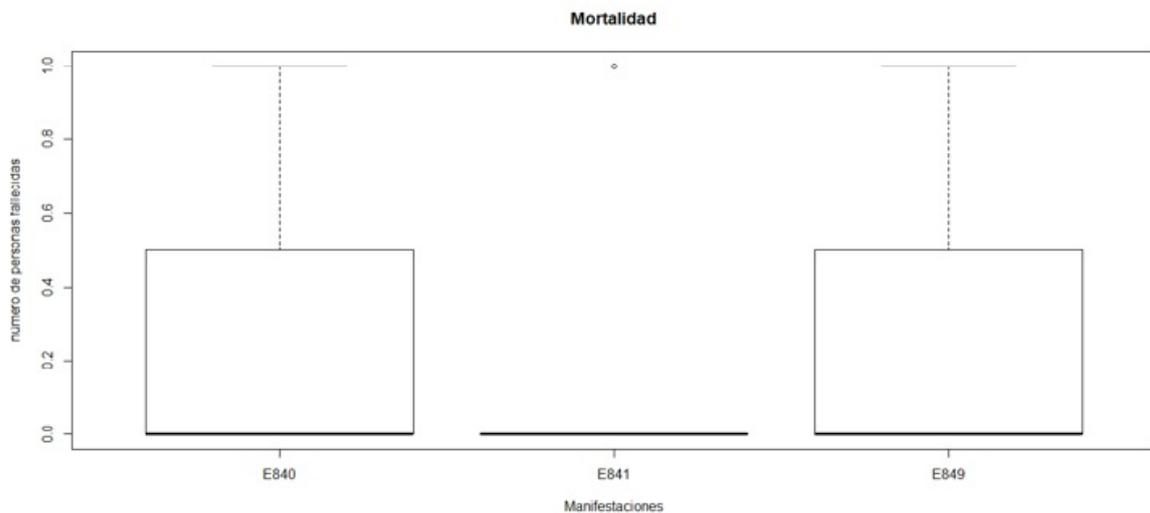


Figura 8.87: Diagrama de Box-plot

Primero se explorará las relaciones entre todas las parejas de variables, en particular la relación de años con cada una de las variables independientes en este caso (E840, E841, E849). Esto lo detectaremos a través de las correlaciones y la función pairs de R-proyect, la cual produce un gráfico matricial para las variables dadas.

```
#####diagrama de dispersion
AÑO<-c(2010,2011,2012,2013,2014,2015,2016,2017)
#Graficas individuales
x11()
plot(E840~AÑO,main="Diagrama de dispersion")
x11()
plot(E841~AÑO,main="Diagrama de dispersion")
x11()
plot(E849~AÑO,main="Diagrama de dispersion")
x11()
#### todos contra todos
pairs(~E840+E841+AÑO+E849)
```

Figura 8.88: Código para los diagramas de dispersiones

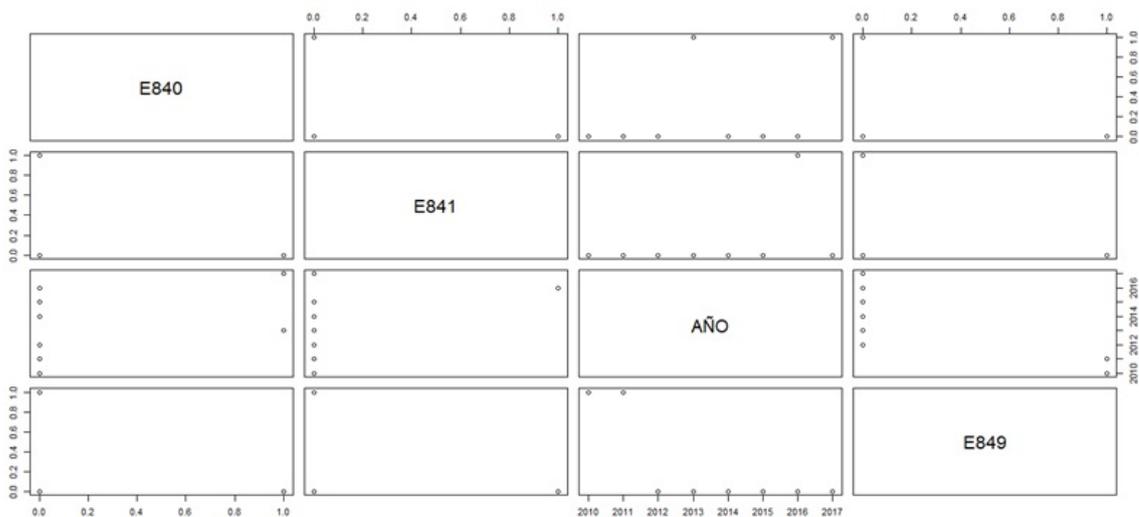


Figura 8.89: Matriz de los diagramas de dispersiones

Se realiza un análisis exploratorio a los datos que permitan visualizar la correlación existente entre las variables.

```
> round(cor(mt), 3)
      AÑO  E840  E841  E849
AÑO  1.000  0.378  0.412 -0.756
E840  0.378  1.000 -0.218 -0.333
E841  0.412 -0.218  1.000 -0.218
E849 -0.756 -0.333 -0.218  1.000
```

Figura 8.90: Datos para la correlación

Se hace una regresión lineal entre la variable E849~ AÑO.

```
> regresion1  
  
Call:  
lm(formula = E849 ~ AÑO)  
  
Coefficients:  
(Intercept)      AÑO  
  287.8929      -0.1429
```

Figura 8.91: Regresión Lineal

La recta de regresión lineal que mejor se ajusta es  $y = 287,8929 - 0,1429x$ .

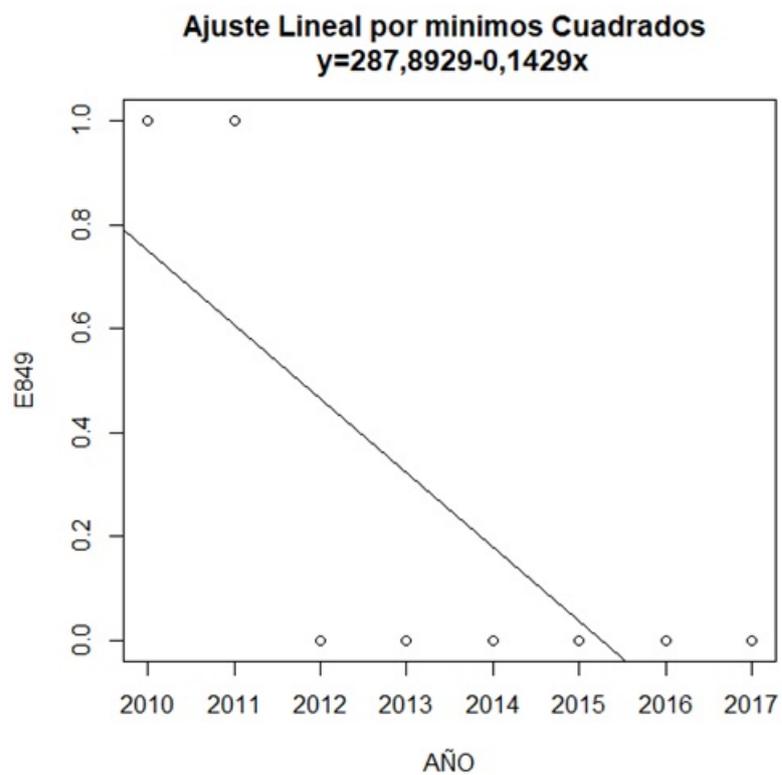


Figura 8.92: Regresión lineal E849 - AÑO

## 8.2.4. TASA ESPECIFICA DE MORTALIDAD

Para el desarrollo de la tasa específica de mortalidad se tuvieron en cuenta el total de la población de defunción de cada año:

Años	No fetales	Fetales	Total defunciones
2010	4779	2009	6788
2011	4736	2577	7313
2012	4885	2411	7296
2013	4938	2770	7708
2014	5092	2648	7740
2015	5688	2497	8185
2016	5795	2664	8459
2017	3141	1239	4380

Figura 8.93: Defunciones durante el periodo abarcado para el estudio de la FQ

### ■ E840-FIBROSIS QUÍSTICA CON MANIFESTACIONES PULMONARES

$$d_{2013} = \frac{1}{7708} = 0,12973534 \Rightarrow 13\text{‰}$$

$$d_{2017} = \frac{1}{4380} = 0,2283105 \Rightarrow 23\text{‰}$$

### ■ E841-FIBROSIS QUÍSTICA CON MANIFESTACIONES INTESTINALES

$$d_{2016} = \frac{1}{8459} = 0,11821728 \Rightarrow 12\text{‰}$$

### ■ E849-FIBROSIS QUÍSTICA SIN OTRA ESPECIFICACIÓN

$$d_{2010} = \frac{1}{6788} = 0,1473188 \Rightarrow 15\text{‰}$$

$$d_{2011} = \frac{1}{7313} = 0,13674279 \Rightarrow 14\text{‰}$$

### ■ Tasa específica total de mortalidad de la FQ.

El total de las defunciones de los años 2010-2017 fue 57869 de la población de mortalidad en el Huila.

$$d_T = \frac{5}{57869} = 0,08640205 \Rightarrow 9\text{‰}$$

### 8.3. ESTUDIO 1: UN GRUPO DE FAMILIAS PORTADORAS DE LA ENFERMEDAD FQ

Un grupo aleatorio de 16 familias, con dos hijos cada una cuyos padres son heterocigóticos para un rasgo recesivo (figura 8.94). En esas 16 familias las probabilidades son de 4 de los 16 primogénitos resulten afectados; igualmente las probabilidades son de que 4 de los segundogénitos hereden el rasgo; pero como esto es un acontecimiento independiente del primero, una sola de las 16 familias tendrá afectados el *primero* y el *segundo* hijos ( $\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$ ).

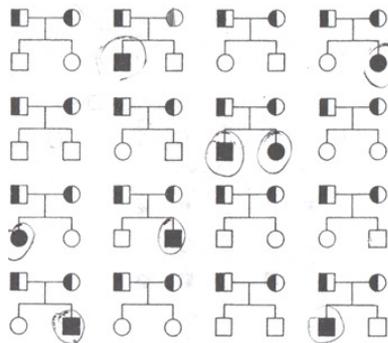


Figura 8.94: En un grupo de dieciséis familias, cada una con ambos progenitores heterocigóticos para el mismo gen recesivo y cada una con dos hijos, los hallazgos serán como aquí indica. Los símbolos en negro indican los hijos afectados por un rasgo recesivo simple.

Como la investigación se realiza cuando se encuentran niños afectados, solamente 7 de las 16 familias son identificables (figura 8.95).

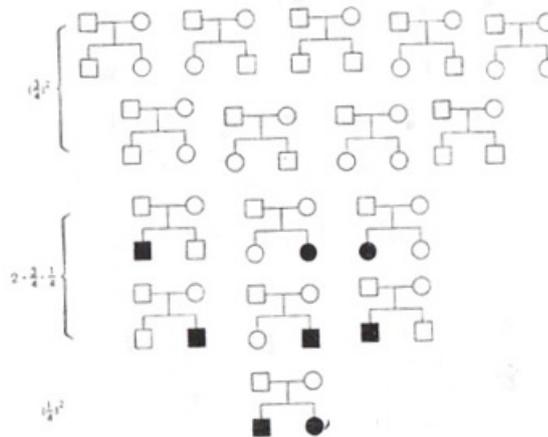


Figura 8.95: Solamente siete de las dieciséis familias podrán evaluarse, porque sólo siete tienen un hijo afectado.

La fórmula general para los hermanos no investigados es de  $\left(\frac{3}{4}\right)^n$ , en la que  $n$  es el número de los hermanos y  $\frac{3}{4}$  la probabilidad del fenotipo dominante.

Obsérvese que los hermanos no investigados han sido suprimidos del binomio desarrollado  $(p + q)^n$ , en el que  $p$  es la probabilidad del fenotipo dominante (en este caso  $\frac{3}{4}$ ),  $q$  la probabilidad del fenotipo recesivo ( $\frac{1}{4}$ ) y  $n$  el número de hermanos.

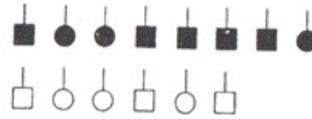


Figura 8.96: La proporción de niños afectados en las familias investigadas no es de 1 en 4, si no de 8 en 14 (57%)

La proporción de niños afectados en esas 7 no es de  $\frac{1}{4}$ , sino de  $\frac{8}{16}$  (o sea, 57%) (figura 8.96).

La figura 8.97 da las fórmulas y la proporción de familias con 2, 3 y 4 hijos de los cuales algunos se encuentran afectados. Fórmula general

$$(p + q)^n$$

Familias de 2 hijos:

$$\left(\frac{3}{4}\right)^2 + 2\left(\frac{3}{4}\right)\left(\frac{1}{4}\right) + \left(\frac{1}{4}\right)^2$$

↓                      ↓                      ↓

$$\frac{9}{16} \qquad \frac{6}{16} \qquad \frac{1}{16}$$

Los 2 hijos normales      1 hijo normal, 1 hijo afectado      Los 2 hijos afectados

Familias de 3 hijos:

$$\left(\frac{3}{4}\right)^3 + 3\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right) + 3\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3$$

↓                      ↓                      ↓                      ↓

$$\frac{27}{64} \qquad \frac{27}{64} \qquad \frac{9}{64} \qquad \frac{1}{64}$$

Los 3 hijos normales      2 hijos normales, 1 hijo afectado      1 hijo normal, 2 hijos afectados      Los 3 hijos afectados

Familias de 4 hijos:

$$\left(\frac{3}{4}\right)^4 + 4\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right) + 6\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 + 4\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^3 + \left(\frac{1}{4}\right)^4$$

↓                      ↓                      ↓                      ↓                      ↓

$$\frac{81}{256} \qquad \frac{108}{256} \qquad \frac{54}{256} \qquad \frac{12}{256} \qquad \frac{1}{256}$$

Los 4 hijos normales      3 hijos normales, 1 hijo afectado      2 hijos normales, 2 hijos afectados      1 hijo normal, 3 hijos afectados      Los 4 hijos afectados

Figura 8.97: Diversas proporciones de hermanos afectados por un rasgo autosómico recesivo, en familias de dos tres y cuatro hijos, en la que ambos progenitores son heterocigóticos.

Para una familia con 5 hijos la formula es la siguiente

$$\left(\frac{3}{4}\right)^5 + 5\left(\frac{3}{4}\right)^4\left(\frac{1}{4}\right) + 10\left(\frac{3}{4}\right)^3\left(\frac{1}{4}\right)^2 + 10\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^3 + 5\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^4 + \left(\frac{1}{4}\right)^5$$

y así sucesivamente.

Por consiguiente la generalización ,

$$\left[\left(\frac{3}{4}\right) + \left(\frac{1}{4}\right)\right]^n$$

Los coeficientes en el desarrollo del binomio pueden obtenerse fácilmente del triángulo de Pascal:

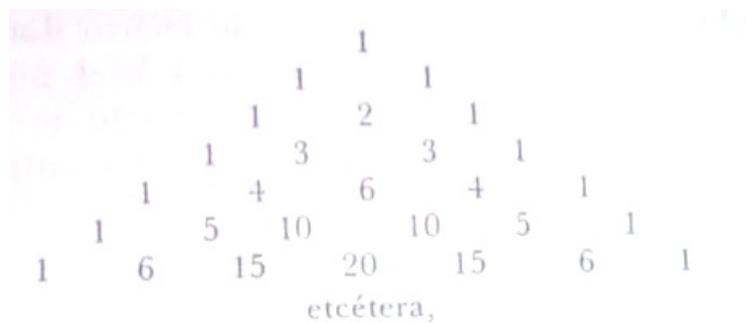


Figura 8.98: Triangulo de Pascal

o pueden obtenerse derivandolos de ls fórmula

$$\frac{n!}{a!b!}$$

En la figura 8.99 se ilustra la segregación de las familias con tres hijos (con ambos padres heterocigóticos).

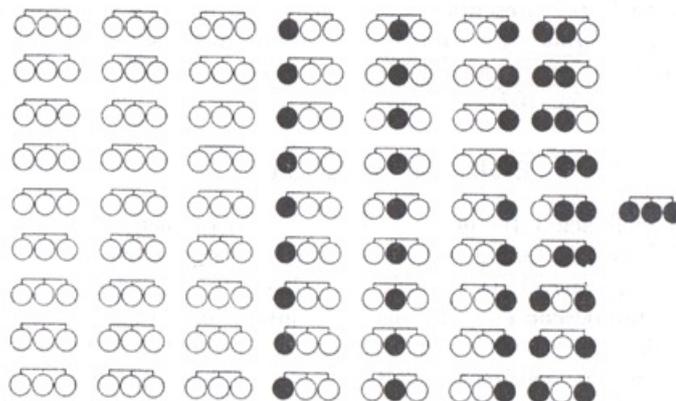


Figura 8.99: Segregación de un rasgo recesivo simple(símbolos negros).

## 8.4. ESTUDIO 2: UNA FAMILIA PORTADORA DE LA ENFERMEDAD FQ

Un matrimonio de heterocigotos para el gen de la fibrosis quística (c) desea tener 4 hijos y desean saber la probabilidad de que ninguno, algunos o todos sus hijos estén afectados.

$$Cc \times Cc$$

- 4 hijos (cada uno es un evento independiente)

NO AFECTADOS	AFECTADOS	PROBABILIDAD
4	0	NNNN
3	1	NNNA
2	2	NNAA
1	3	NAAA
0	4	AAAA

Cuadro 8-1: Tabla experimento de una familia portadora FQ

- N: normales
- A: afectados

Aplicando probabilidades:

NO AFECTADOS	AFECTADOS	PROBABILIDAD
4	0	NNNN

$$1 \left( \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \right) = 1 \left( \frac{81}{256} \right) \approx 0,31$$

Donde 3/4

- 3: Es la frecuencia absoluta de los hijos no afectados.
- 4: Cantidad de individuos.

NO AFECTADOS	AFECTADOS	PROBABILIDAD
3	1	NNNA

$$4 \left( \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} \right) = 4 \left( \frac{27}{256} \right) \approx 0,42$$

Donde 1/4

- 1: Es la frecuencia absoluta de los hijos afectados.
- 4: Cantidad de individuos.

NO AFECTADOS	AFECTADOS	PROBABILIDAD
2	2	NNAA

$$6 \left( \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} \right) = 6 \left( \frac{9}{256} \right) \approx 0,21$$

- 1: Es la frecuencia absoluta de los hijos afectados.
- 4: Cantidad de individuos.

NO AFECTADOS	AFECTADOS	PROBABILIDAD
1	3	NAAA

$$4 \left( \frac{3}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \right) = 4 \left( \frac{3}{256} \right) \approx 0,04$$

- 1: Es la frecuencia absoluta de los hijos afectados.
- 4: Cantidad de individuos.

NO AFECTADOS	AFECTADOS	PROBABILIDAD
0	4	AAAA

$$1 \left( \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \right) = 1 \left( \frac{1}{256} \right) \approx 3,90625 \times 10^{-3}$$

La sumatoria de todas las probabilidades:

$$\sum_{j=1}^n P_j = 1$$

Entonces las frecuencias genóticas de los individuos enfermos es  $q^2$  ósea que de 4 hijos solo tiene la enfermedad 1.

$$P + Q = 1$$

$$q^2 = \frac{1}{4} \approx 0,25$$

$$q = \frac{1}{2} \approx 0,5$$

Aplicando ecuación (4) surgen los portadores de la enfermedad; los heterocigóticos Cc

$$2pq = 2(0,5 \times 0,5) = 0,5$$

Aplicando ecuación (2), surgen los sanos CC.

$$p = 1 - q$$

$$p = 1 - (0,5) = 0,5$$

$$p^2 = (0,5)^2 = 0,25$$

La Estimación de cuantos individuos enfermos nacieran aproximadamente en cada 4 nacimientos; si los progenitores ambos son portadores y no manifiesten la enfermedad.

$$\frac{(2pq)^2}{(1 - q)^2} = \frac{(0,5)^2}{(1 - 0,5)^2} = 1$$

Individuos enfermos si nacieran de padres que no manifiesten la enfermedad en un grupo familiar será

$$\frac{1}{2} \times \frac{(0,5)^2}{(1 - 0,5)^2} = 0,5 \approx 50 \%$$

Se toma  $\frac{1}{2}$ , porque ahora en día las familias se reducen a tener máximo dos hijos por cuestiones económicas; dado el caso que si una pareja es portadora y allegara a tener dos hijos la probabilidad es el 50 %, de que alguno de los dos puede ser el primero o el segundo que posea esa enfermedad.

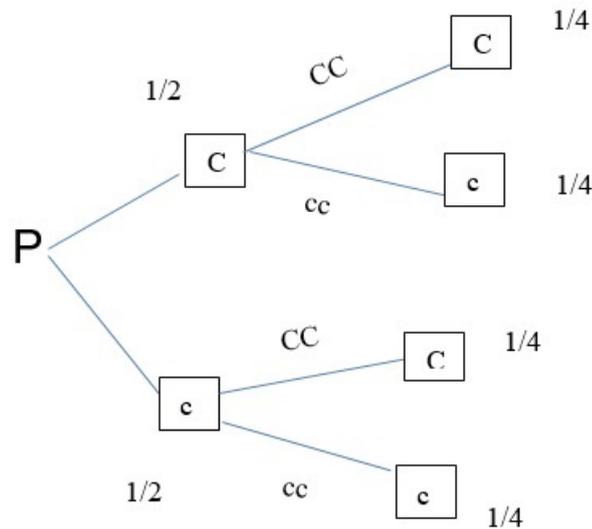


Figura 8.100: Diagrama de Arbol

Considere solo dos estados: normal o afectado, la probabilidad de tener un niño normal es de 25 %; si el anterior fue normal, pero si el último niño que nació fue afectado la probabilidad de tener uno normal es de 25 %.

## Matriz de transición

Definimos la variable:

X: Estado de la enfermedad

Donde,

$$X = \begin{cases} 0 & \text{Normal} \\ 1 & \text{Afectado} \end{cases}$$

$$P \{X_{t+1} = 0 \mid X_t = 0\} = 0,25$$

$$P \{X_{t+1} = 0 \mid X_t = 1\} = 0,25$$

Luego, realizamos la matriz transición.

ESTADO ACTUAL $x_t$	ESTADO SIGUIENTE $x_{t+1}$	
	0	1
0	0.25	0.75
1	0.25	0.75

Cuadro 8-2: Tabla de Estados de la matriz

Diagrama de transición.

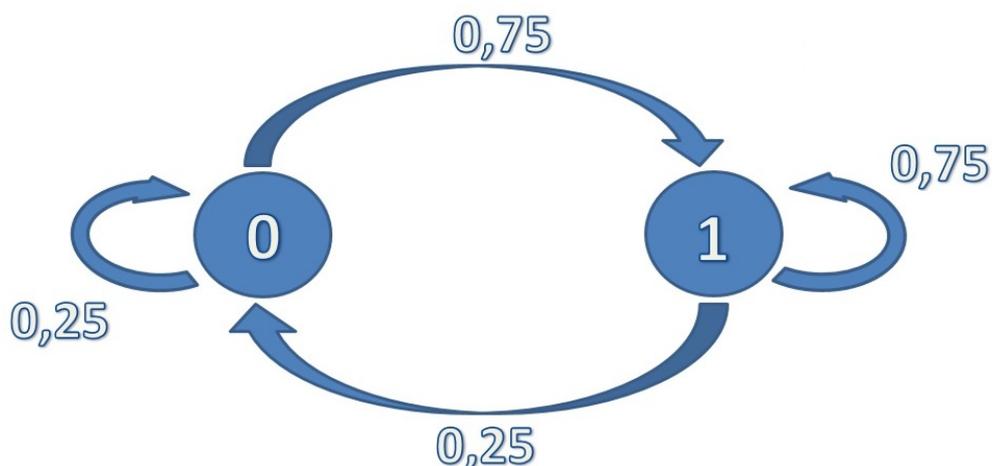


Figura 8.101: Diagrama de transición de Estados

Para la matriz transición de 2x2 se plantean de la siguiente manera:

$$\pi_0 = P_{00}\pi_0 + P_{10}\pi_1$$

$$\pi_1 = P_{01}\pi_0 + P_{11}\pi_1$$

$$1 = \pi_0 + \pi_1$$

Luego, calcular:

- $\pi_0$ : probabilidad de llegar a uno normal.
- $\pi_1$ : probabilidad de llegar a uno afectado.

Tenemos que:

$$\pi_0 = 0,25\pi_0 + 0,25\pi_1 \quad (1)$$

$$\pi_0 = 0,75\pi_0 + 0,75\pi_1 \quad (2)$$

$$1 = \pi_0 + \pi_1 \quad (3)$$

Despejamos la ecuación (3)

$$\pi_0 = 1 - \pi_1 \quad (4)$$

Tomamos la ecuación (4) y reemplazamos en la ecuación (1).

$$1 - \pi_1 = 0,25(1 - \pi_1) + 0,25\pi_1$$

$$1 - \pi_1 = 0,25 - 0,25\pi_1 + 0,25\pi_1$$

$$1 - \pi_1 = 0,25$$

$$\pi_1 = 1 - 0,25$$

$$\pi_1 = 0,75$$

Y finalmente reemplazamos el valor de  $\pi_1$  en la ecuación (4) ya que la tenemos despejada

$$\pi_0 = 1 - \pi_1$$

$$\pi_0 = 1 - 0,75$$

$$\pi_0 = 0,25$$

Se concluye, el estudio de los rasgos autosómicos recesivos, se examinan de ordinario familia con ambos progenitores heterocigóticos y por ello tienen un 25 % de probabilidad de pro-crear hijos afectados, solamente cuando por lo menos uno de los hijos lo está.

## 8.5. SIMULACIÓN EN SOFTWARE EXCEL Y POPULUS

### 8.5.1. SIMULACIÓN EJEMPLO 1

Con respecto al enunciado se aplicó a las 16 familias y se desarrollo en Excel una tabla la cual ejerce la aleatoriedad de los gametos y los cigotos con respecto a la problemática del trabajo.

gametos Y	gametos X	cigotas
A	r	Ar
r	A	Ar
r	r	rr
A	A	AA
r	A	Ar
A	r	Ar
r	r	rr
A	A	AA
A	A	AA
A	r	Ar
A	A	AA
A	A	AA
r	A	Ar
A	A	AA
A	A	AA
r	r	rr

Figura 8.102: Tabla aleatoria de los gametos y cigotos

Para desarrollar la tabla (8.102)se utiliza la siguiente formula de aleatoriedad para formar las columnas de los gametos:

$$= SI((ALEATORIO()) > 0,5;"A";"r")$$

Para formar la columna de los cigotos se utiliza la siguiente formula:

$$= SI(A_n = B_n; A_n \& B_n;"Ar")$$

con n:número de fila con la que esta cruzando los datos de los gametos, en este caso sería  $n = 1, 2, \dots, 16$ .

E	F	G	H	I	J	K
N° DE GAMETOS M.		N° DE GAMETOS H		Genotipos	Frecuencia absoluta	Frecuencias de genotipos
A	12	A	12	AA	9	0,5625
r	4	r	4	Ar	6	0,375
				rr	1	0,0625
				Total	16	

Figura 8.103: Cuadros de Gametos y Genotipos

Se realiza un cuadro para clasificar el número de gametos simbolizados con la letra A para los dominantes y la letra r para los recesivos; utilizando las siguientes formulas:

$$= \text{CONTAR.SI}(A2 : A17; "A")$$

$$= \text{CONTAR.SI}(A2 : A17; "a")$$

Luego, se realiza el cuadro que da lugar a los Genotipos de los cuales son AA,Ar,rr; la clasificación de la Frecuencia relativa y las Frecuencias (absolutas) de los genotipos o frecuencias genotípicas.

La formula para clasificar la Frecuencia relativa (cantidad de individuos) que posee los genotipos:

$$AA = \text{CONTAR.SI}(C2 : C17; "AA")$$

$$Ar = \text{CONTAR.SI}(C2 : C17; "Ar")$$

$$rr = \text{CONTAR.SI}(C2 : C17; "rr")$$

Verificando la Frecuencia relativa (cantidad total de individuos):

$$= \text{SUMA}(J2; J3; J4)$$

Luego, la formula para realizar las **Frecuencias de los genotipos** es la siguiente:

$$f(\text{genotipica}) = \frac{\# \text{Genotipos}}{\# \text{Total muestra}}$$

Se realiza el cuadro del Número de alelos y las Frecuencias alélicas

	M	N	O	P
	N° de Alelos		Frecuencias de alelos	
1				
2	A	24	A(p)	0,75
3	r	8	r(q)	0,25
4	Total	32	Total	1

Figura 8.104: Cuadro de las Frecuencias Alelicas

La formula para hallar **Las Frecuencias Alélicas** es la siguiente:

$$f(\text{Alelos}) = \frac{\# \text{Alelos observados}}{\# \text{Total de alelos}}$$

Además abarcamos el cuadro del Ejemplo 1; donde se presenta Caracteres de la generación parental y se encuentra el modelo de Hardy-Weinberg.

Q	R	S	T	U	V	W
Genes Alélicos	Frecuencia asignada	Ejemplo	Genotipos	Frecuencia en la población	Ejemplo numérico	Ejemplo numérico %
A	p	0,75	AA	$p^2$	0,5625	56,25
r	q	0,25	Ar	$2pq$	0,375	37,5
			rr	$q^2$	0,0625	6,25
Juntos		1	Todos los tipos		1	100

Figura 8.105: Caracteres de la generación parental

Utilizando las frecuencias asignadas de las cuales surgen de las frecuencias alélicas, se hallan los valores con respecto a la frecuencia en la población, donde se observa la fórmula del equilibrio que corresponde al modelo de Hardy-Weinberg, del cual se aplica las siguientes fórmulas en Excel.

Para **Las frecuencias en la población** se tiene:

$$AA = p^2 = (S2)^2$$

$$Ar = 2pq = 2 * (S2 * S3)$$

$$rr = r^2 = (S3)^2$$

El valor total de las frecuencias en la población se halla de la siguiente manera:

$$p^2 + 2pq + r^2 = SUMA(V2; V3; V4) = 1$$

En la columna Ejemplo numérico en % se toman los valores de la columna anterior la cual es Ejemplo numérico, cada valor se multiplica por 100 y el total se halla con la suma de cada uno de los valores:

$$= SUMA(W2; W3; W4)$$

Por otra parte, se realiza el cuadro del **Prole de uniones al azar**, para comenzar a detallar las generaciones de esas familias a medida que surgen el cruce entre los alelos aleatoriamente.

X	Y	Z	AA	AB	AC	AD	AE
Padres	Frecuencia del tipo de unión	Frecuencia de la prole*			Ejemplo numérico %		
		AA	Ar	rr	AA	Ar	rr
AA x AA	$p^4$	$p^4$			0,31640625	0	0
AA x Ar	$4p^3q$	$2p^3q$	$2p^3q$		0,2109375	0,2109375	0
Ar x AA							
AA x rr	$2p^2q^2$		$2p^2q^2$		0	0,0703125	0
rr x AA							
Ar x Ar	$4p^2q^2$	$p^2q^2$	$2p^2q^2$	$p^2q^2$	0,03515625	0,0703125	0,03515625
Ar x rr	$4pq^3$		$2pq^3$	$2pq^3$	0	0,0234375	0,0234375
rr x Ar							
rr x rr	$q^4$			$q^4$	0	0	0,00390625
Todos los tipos		$p^2$	$2pq$	$q^2$	0,5625	0,375	0,0625

Figura 8.106: Prole de las uniones al azar

Para hallar la frecuencia de la Prole se utilizan las siguientes formulas:

$$\begin{aligned}
 AA \times AA &= (P2)^4 \\
 AA \times Ar &= 2 * (((P2)^3) * (P3)) \\
 Ar \times AA &= 2 * (((P2)^3) * (P3)) \\
 AA \times rr &= 2 * (V2 * V4) \\
 rr \times AA &= 2 * (V2 * V4) \\
 Ar \times Ar &= V2 * V4 \\
 &= 2 * (V2 * V4) \\
 &= V2 * V4 \\
 Ar \times rr &= 2 * (P2 * ((P3)^3)) \\
 rr \times Ar &= 2 * (P2 * ((P3)^3)) \\
 rr \times rr &= (P3)^4
 \end{aligned}$$

La suma total de cada una de las casillas del ejemplo númerico, volvemos y obtenemos el resultado de las frecuencias en la población:

$$\begin{aligned}
 p^2 &= SUMA(AC3; AC4; AC5; AC6; AC7; AC8; AC9; AC10; AC11) \\
 2pq &= SUMA(AD3; AD4; AD5; AD6; AD7; AD8; AD9; AD10; AD11) \\
 r^2 &= SUMA(AE3; AE4; AE5; AE6; AE7; AE8; AE9; AE10; AE11)
 \end{aligned}$$

Después de haber realizado la parte algebraica con los calculos desarrollados; se opta por utilizar el programa Populus para graficar y observar el equilibrio con respecto a las generaciones.

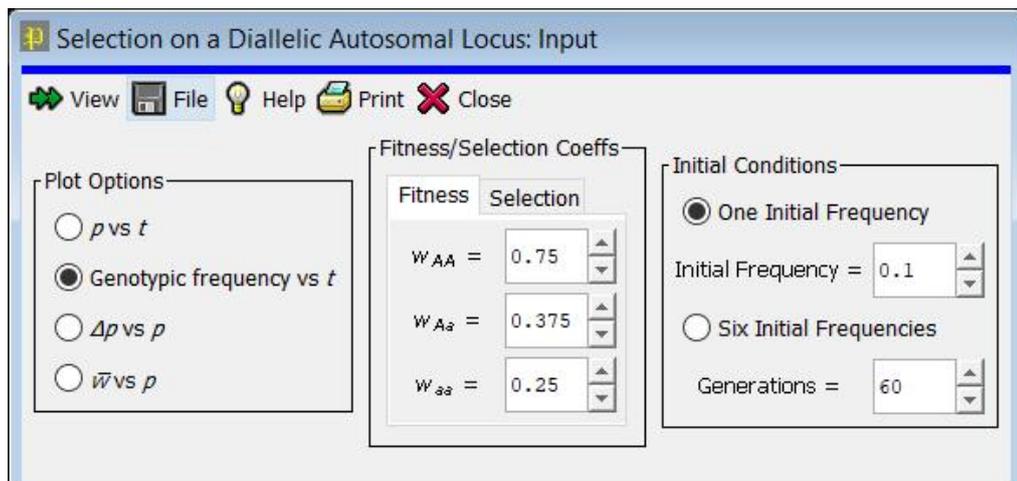


Figura 8.107: Ventana del simulador

En la figura (8.107) se ingresan los valores de las Frecuencias de los genotipos y el simulador aproxima los resultados en la última cifra de algunos valores; luego el valor hasta cuantas generaciones se quiere observar el proceso.

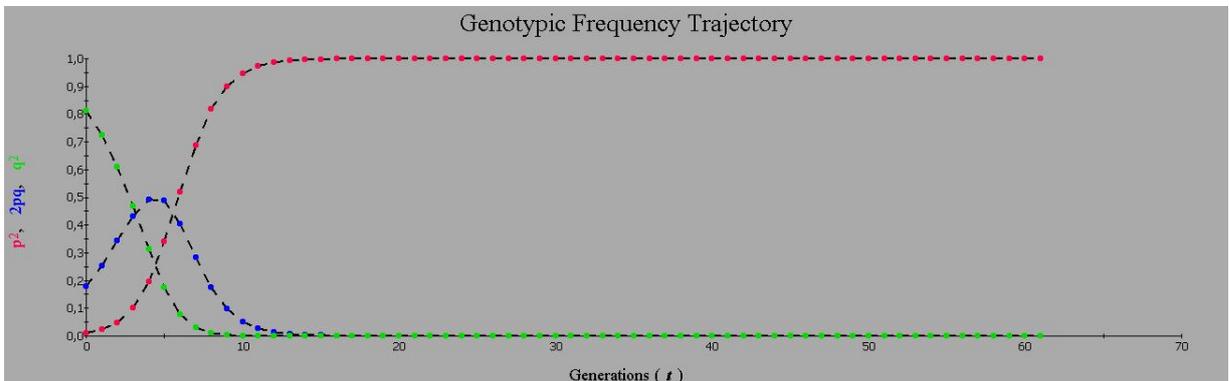


Figura 8.108: Gráfica del Equilibrio H-W

### Análisis cualitativo de la figura 8.108

La curva roja es llamada  $p^2$  nos indica los valores de los individuos(familiares) sanos o que no poseen la enfermedad, la cual es una curva creciente hasta la generación 12, después de la generación 12 comienza a tomar un valor constante (1.0) de las frecuencias genotípicas y el comportamiento es estable.

Ahora, la curva azul es llamada  $2pq$  nos indica los valores de los individuos(familiares) portadores del gen, la cual posee un crecimiento hasta la cuarta generación, donde alcanza su máximo valor de 0,5 con respecto a las frecuencias genotípicas y surge un decrecimiento de portadores desde la quinta (5) generación hasta la generación 13; abarcando los valores de la generación 13 hasta la generación 16 el comportamiento de la curva es constante, optando a desaparecer los portadores después de la generación 16.

La curva verde es llamada  $q^2$  nos indica los valores de los individuos (familiares) enfermos poseedores del gen y la enfermedad FQ. Además, es una curva exponencial y su comportamiento es decreciente con respecto a las frecuencias genotípicas, en la generación 10 (decima generación) surge un comportamiento estable y tiende a desaparecer el gen de la enfermedad con respecto a las generaciones.

Se concluye al tomar un valor de las frecuencias genotípicas de la aleatoriedad de los alelos, con respecto a las familias se ajusta al modelo de Hardy-Weinberg, obteniendo el equilibrio alélico en la octava generación de acuerdo a la enfermedad la Fibrosis Quística; optando por resultar más descendientes sanos a medida que transcurre las generaciones y tendiendo a desaparecer la

enfermedad dentro de las familias.

Ahora, tomando un valor arbitrario de la aleatoriedad de las frecuencias genotípicas, se detalla que se pierde el equilibrio de la enfermedad en las generaciones de la familias.

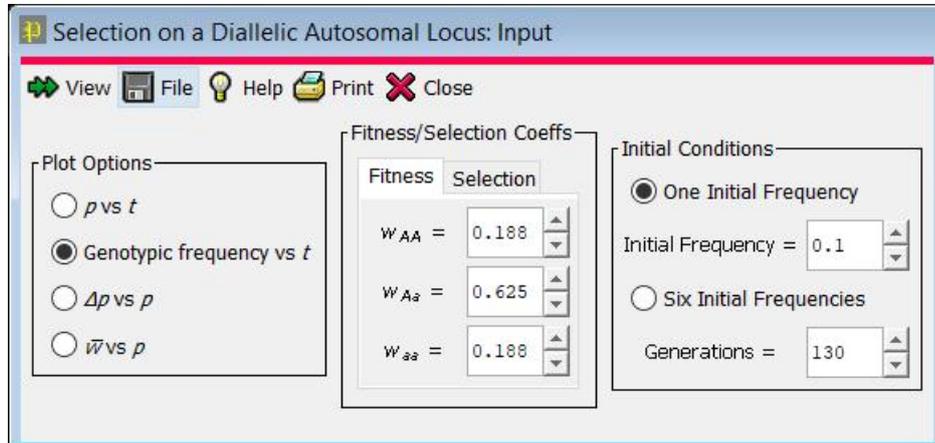


Figura 8.109: Ventana del simulador

Se observa la gráfica de este proceso.

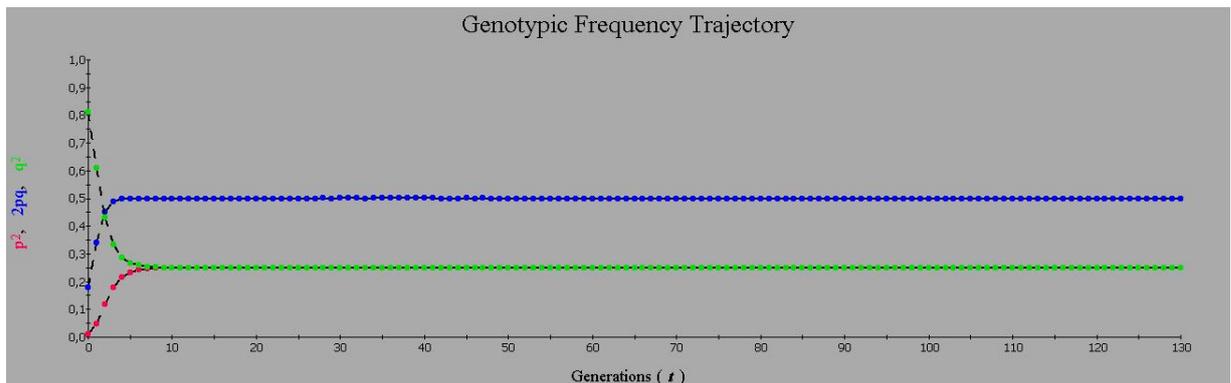


Figura 8.110: Perdida del Equilibrio

### Análisis cualitativo 8.110

La curva roja es llamada  $p^2$  nos indica los valores de los individuos(familiares) sanos o que no poseen la enfermedad FQ. Además, es una curva exponencial y su comportamiento es creciente hasta la decimo tercera generación (13), después de la generación 13 continúan los valores de los enfermos a lo que se refiere que esos descendientes poseen la enfermedad.

La curva azul es llamada  $2pq$  nos indica los valores de los individuos(familiares) portadores del gen, la cual posee un crecimiento hasta la cuarta generación, donde alcanza su máximo valor de 0,5 con respecto a las frecuencias genotípicas y surge un cambio desde la quinta generación

donde comienza la curva a volverse estable a medida que transcurre las generaciones.

Se intersecan las dos curvas las cuales son  $2pq$  portadores del gen y  $q^2$  los enfermos de FQ, en la segunda generación (2) con respecto al valor 0.44 de la frecuencia genotípica.

La curva verde es llamada  $q^2$  nos indica los valores de los individuos (familiares) enfermos poseedores del gen y la enfermedad FQ. Además, es una curva exponencial y su comportamiento es decreciente hasta el valor 0.245 de las frecuencias genotípicas, surgiendo un cambio de la curva volviendose constante (estable) durante la novena generación (9) y continuando su trayectoria de la enfermedad a los descendientes.

Se concluye que al tomar un valor arbitrario de las frecuencias genotípicas de la aleatoriedad de los alelos, con respecto a las familias no se obtuvo el equilibrio de Hardy-Weinberg de la enfermedad FQ de acuerdo a sus descendientes, surgiendo más portadores del gen y más enfermos de FQ a medida que transcurre las generaciones.

### 8.5.2. SIMULACIÓN EJEMPLO 2:

Análogamente se desarrolla el proceso en excel y abarcando la simulación en populus se tiene:

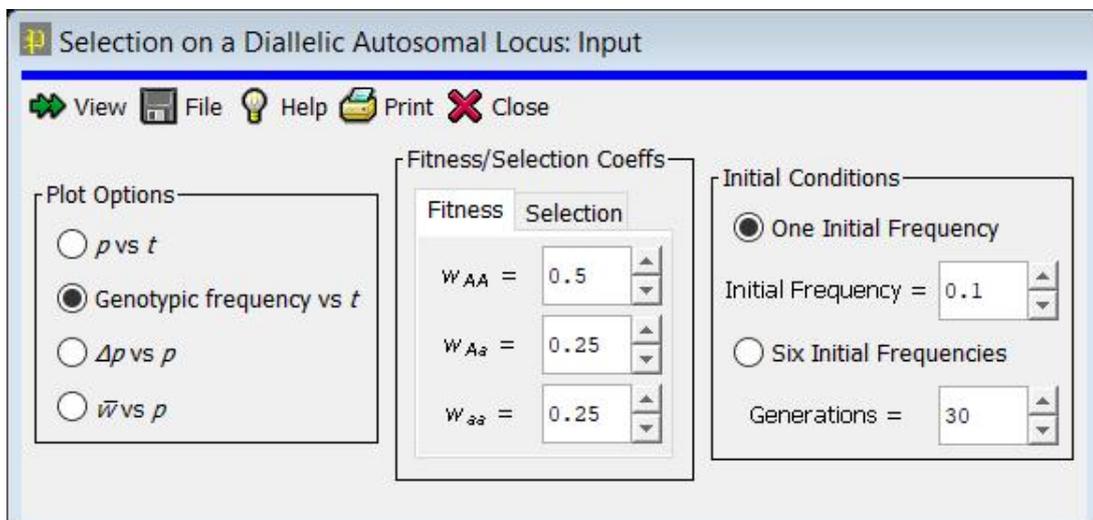


Figura 8.111: Ventana de la simulación

En la figura(8.111) se ingresan los valores de las Frecuencias genotípicas y se obtiene la gráfica que nos muestra el proceso del equilibrio alélico de la familia con respecto al modelo de Hardy-Weinberg.

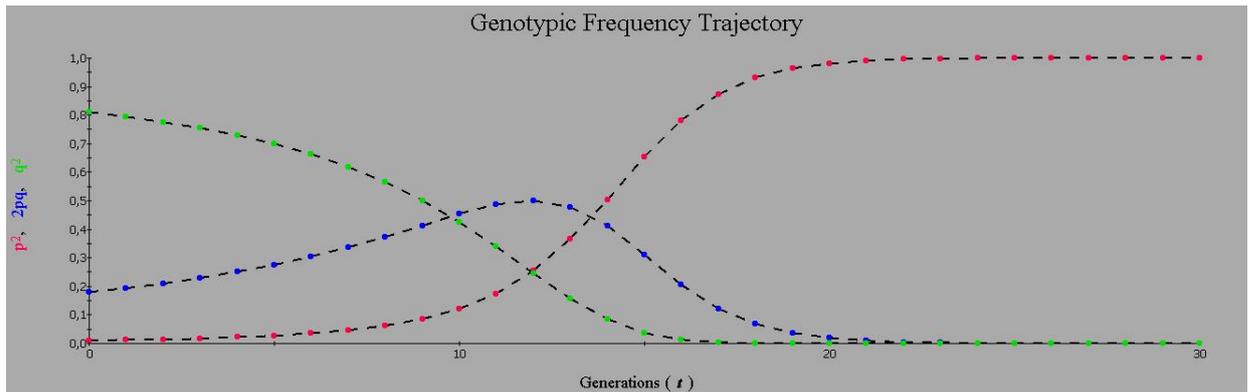


Figura 8.112: Grafica del modelo H-W para una familia portadora de FQ

### Análisis cualitativo de la figura 8.112

La curva roja es llamada  $p^2$  nos indica los valores de los individuos(familiares) sanos o que no poseen la enfermedad, la cual es una curva inicia con valores constantes hasta la quinta(5) generación; luego surge un cambio de crecimiento hasta la generación 20, después de la generación 20 comienza a tomar un valor constante (1.0) de las frecuencias genotípicas y el comportamiento de la curva se vuelve estable.

Ahora, la curva azul es llamada  $2pq$  nos indica los valores de los individuos(familiares) portadores del gen, la cual posee un crecimiento moderado hasta la generación 12, donde alcanza su máximo valor de 0,5 con respecto a las frecuencias genotípicas y surge un decrecimiento de portadores desde la generación 13 hasta la generación 20; abarcando los valores de la generación 20 hasta la generación 30 y los portadores tienden a desaparecer o a convertirse en enfermos.

La curva verde es llamada  $q^2$  nos indica los valores de los individuos (familiares) enfermos poseedores del gen y la enfermedad FQ. Además, es una curva exponencial y su comportamiento es decreciente con respecto a las frecuencias genotípicas hasta la generación 16, tiende a desaparecer los descendientes enfermos.

Se concluye al tomar unos valores de las frecuencias genotípicas de la aleatoriedad de los alelos, con respecto a la familia se ajusta al modelo de Hardy-weinberg, obteniendo el equilibrio alélico entre la generación 9 y la generación 13 de acuerdo a la enfermedad la Fibrosis Quística; optando por resultar más descendientes sanos a medida que transcurre las generaciones y tendiendo a desaparecer la enfermedad dentro del pedigrí familiar.

## Capítulo 9

---

# CONCLUSIONES

---

- A partir de la selección de una sola familia se pudo proyectar a futuro la afectación de la enfermedad en sus descendientes, logrando un equilibrio genético.
- Se puede determinar mediante la simulación que a partir de la generación 20 la enfermedad se extingue dentro del grupo familiar.
- De acuerdo a el análisis realizado sobre la fibrosis quística en el departamento del Huila; se obtuvo una tasa específica de mortalidad 9 ‰ ; es decir, que por cada 1000 muertos durante el periodo 2010 al 2017, nueve individuos fallecieron con causa antecedente originaria de FQ. Al estudio alélico se presentó un aumento progresivo del 1 ‰ con respecto a cada año.
- Se aplicó el modelo de Hardy-Weinberg para población infinita ; se puede desarrollar para enfermedades multialélicas (polimorfismo genético) y así lograr un equilibrio, ya cual sea la enfermedad que este relacionada y avanzar nuevas investigaciones con respecto a este modelo.
- Para el análisis la proyección a futuro, permite evidenciar como ha disminuido la afectación de la enfermedad FQ, con respecto a los tratamientos para mejorar la calidad de vida de los huilenses y colombianos.

## Apéndice A

---

# TÉRMINOS, CONCEPTOS Y DEFINICIONES

---

A lo largo del trabajo de grado es necesario tener en claro algunos conceptos básicos que permitirán facilitar el desarrollo del mismo.

**Probabilidades:** Dado un suceso A de un espacio muestral  $\omega$  (omega) asociado a un fenómeno aleatorio, podemos estar interesados en algún indicador del "grado de confianza" que tenemos en que realizar una observación del fenómeno observemos precisamente un resultado que este en A (diremos que el suceso A se ha realizado").

**Malformación congénita:** Que se presenta durante el período de gestación y puede ser hereditaria o no.

**Cromosomas:** Cromo (color) y soma (cuerpo o elemento) a cada una de las estructuras altamente organizadas, formadas por ADN y proteínas, que contiene la mayor parte de la información genética de un individuo.

**Prenatal:** Hace referencia a las primeras etapas del desarrollo de los seres vivos en especial en los seres humanos, abarca todo el proceso en que transcurre desde la concepción o la fecundación del nuevo ser.

**Teratógenos:** Es todo aquello capaz de alterar el desarrollo embrionario normal del ser vivo, causando malformaciones.

**Genómica:** Es un campo interdisciplinario de la ciencia dentro del campo de la biología molecular. Un genoma es un conjunto completo de ADN dentro de una sola célula de un organismo, y como tal, la genómica se enfoca en la estructura, función, evolución y mapeo de los genomas.

**Trastornos monogénicos:** Las enfermedades hereditarias monogénicas son aquellas producidas por alteraciones en la secuencia de ADN de un solo gen.

**Hemoglobinopatías:** Es un grupo de trastornos en los cuales hay una estructura y producción anormal de la molécula de la hemoglobina. Se transmite de padres a hijos (hereditario).

**Hemofilia:** Enfermedad hereditaria que se caracteriza por un defecto de la coagulación de la sangre debido a la falta de uno de los factores que intervienen en ella y que se manifiesta por una persistencia de las hemorragias.

**Autosomas:** El conjunto de cromosomas no sexuales del organismo. Según su definición, intervienen en la transmisión de todas las fisonomías y circunstancias genéticas, excepto las vinculadas con el sexo.

**CFTR:** Nombre de la proteína y del gen que se encuentran alterados en la FQ, viene de las siglas en inglés de Regulador de la Conductancia Transmembrana de la FQ.

**Frecuencias:** Es una magnitud que mide el número de repeticiones por unidad de tiempo de cualquier fenómeno o suceso periódico. Para calcular la frecuencia de un suceso, se contabilizan un número de ocurrencias de éste, teniendo en cuenta un intervalo temporal, y luego estas repeticiones se dividen por el tiempo transcurrido.

**Alelos:** Es cada una de las dos o más versiones de un gen. Un individuo hereda dos alelos para cada gen, uno del padre y otro de la madre.

**Población:** Es un conjunto de seres vivos de una especie que habita en un determinado lugar. Se utiliza también para referirse al conjunto de viviendas, de forma similar al término 'localidad'. Procede del latín *populat*o, -*nis*.

**Acervo genético:** Conjunto de posibles combinaciones de genes para una determinada especie o población. Un mayor acervo genético implica mejores posibilidades de la especie para subsistir y desarrollarse en distintos entornos.

**Gametos:** Un gameto es una célula que tiene una función reproductora. En el ser humano podemos distinguir los gametos masculinos (espermatozoides) y los gametos femeninos (óvulos). Cada gameto tiene 23 cromosomas. Los espermatozoides se producen en los testículos del hombre y los óvulos en los ovarios de la mujer. Los dos tipos de gametos se ponen en contacto durante una relación sexual o también durante una fecundación in vitro. La unión de ambos

gametos (fecundación) puede dar lugar a la formación de un embrión (futuro bebé).

**Homocigoto:** Respecto a un gen cuando los dos alelos codifican la misma información para un carácter, por ejemplo, color de la flor en la arvejilla. Para nombrarlos se utilizan letras mayúsculas y minúsculas; así se dice que AA es Homocigota Dominante y aa es Homocigota Recesivo. Homocigoto dominante es para una característica particular que posee dos copias idénticas. Los alelos denominados alelos dominantes, se representan con una letra mayúscula. Cuando un organismo es homocigota dominante para una característica particular, el genotipo está representado por una duplicación del símbolo de ese rasgo. Un individuo que es homocigoto recesivo para un rasgo particular lleva dos copias idénticas y recesivas del alelo que codifica para el rasgo recesivo. Los alelos denominados alelos recesivos, se representan generalmente por la forma minúscula de la letra utilizada para el rasgo dominante correspondiente. El genotipo de un organismo que es homocigótico recesivo para un rasgo particular se representa por una duplicación de la letra apropiada.

**Heterocigoto:** Un heterocigoto es un organismo que posee diferentes alelos en un gen. Este organismo lleva formas diferentes de un gen, donde esas formas producen resultados ligeramente diferentes (esto es, cabello castaño o negro, ojos azules o verdes, etc.) En cada caso, el mismo gen tiene ligeras variaciones, por lo tanto, estas son formas "hetero" diferentes de la misma cosa.

**Mortalidad:** La mortalidad en el ámbito demográfico, es la relación que existe entre el número de defunciones ocurridas durante un tiempo determinado. El fenómeno de la mortalidad se expresa a través de la tasa o índice de mortalidad, la cual puede definirse como el número de muertos por cada mil habitantes en relación con la población total a lo largo de un periodo establecido.

**Morbilidad:** Es un término de uso médico y científico, sirve para señalar la cantidad de personas o individuos considerados enfermos o víctimas de una enfermedad en un espacio y tiempo determinados. La morbilidad es, entonces un dato estadístico de altísima importancia para poder comprender la evolución y avance o retroceso de una enfermedad, así también como las razones de su surgimiento y las posibles soluciones.

**AMPc:** Siglas del adenosinmonofosfato cíclico, enzima de la clase hidrolasa: adenosinmonofosfato + agua = monofosfato de inosina + amoniaco. Se encuentra en proporción muy alta en el tejido muscular y es la principal fuente de iones amonio durante la contracción muscular. El AMPc es un segundo mensajero, empleado en las rutas de transducción de la señal en las células como respuesta a un estímulo externo o interno, como puede ser una hormona como el glucagón o la adrenalina, o una respuesta de regulación postraduccional. Suele estar relacionado con la activación de proteína kinasas variadas. En bacterias, es un regulador catabólico que controla la

expresión de genes relacionados con la degradación de azúcares en función de la concentración de glucosa.

**Porcentaje:** Proporción expresada en porcentaje respecto del total. Ejemplo, el porcentaje de defunciones de menores de un año, respecto del total de defunciones.

**Tasa o coeficiente:** Se refiere a la frecuencia relativa con la que ocurren ciertos hechos en la población durante un tiempo determinado, generalmente un año. Sin embargo, la palabra tasa, ha ido adquiriendo un significado más amplio y es usada para designar a indicadores obtenidos mediante operaciones complejas o incluso, como sinónimo de relación, proporción o porcentaje. Frecuentemente, se emplean ponderadas por una constante, 100 o 1000, a fin de que adquieran valores significativos.

---

# Bibliografía

---

- [1] ROSARIO DELGADO DE LA TORRE, *Probabilidad y Estadística para ciencias e Ingenierías*, Universidad Autónoma de Barcelona, Facultad de Ciencias, España, 2008.
- [2] ALBERTO LUCEÑO VAZQUEZ, FRANCISCO JAVIER GONZALEZ ORTIZ, *Métodos Estadísticos para medir, Describir y Controlar la Variabilidad*, Universidad De Cantabria, España, 2004.
- [3] FRANK SPITZER, *Pinciples of Random Walk*, New york , 1964.
- [4] Z.L.KOVÁCS, *TEORIA DA PROBABILIDADE E PROCESSOS ESTOCÁSTICOS*, professor da Escola Politécnica da Univeridade de Sao Paulo, Brasil , 1996.
- [5] PAUL C. SHIELDS, *THE ERGODIC THEORY OF DISCRETE SAMPLE PATHS*, United States of America , 1996.
- [6] R DAWKINS, *LEYES DE MENDEL*, Bases científicas, academia.edu.
- [7] R PLOMIN, JC DEFRIES, GE MACCLEARN, L PEZZI, *GENÉTICA DE LA CONDUCTA*, 1984.
- [8] J.M. VAQUERO BARRIOS, *TERAPIA RESPIRATORIA DE LA FIBROSIS QUÍSTICA; TRATAMIENTO ANTIBIÓTICO E INHALADO EN FIBROSIS QUÍSTICA*, Servicio de Neumología, Hospital Reina Sofía, Córdoba, España, 2005. (PDF Download Available).
- [9] ROSA MARIA GIRÓN MORENO, ANTONIO SALSEDO POSADA, *FIBROSIS QUÍSTICA* , Monografías NeumoMadrid, volumen VIII, España, 2005.
- [10] FEDERACION ESPAÑOLA CONTRA LA FIBROSIS QUISTICA, *LIBRO BLANCO DE ATENCION A LA FIBROSIS QUISTICA* ,Asociación cántabra de fibrosis quística, España, 2017.
- [11] FACULTAD DE SALUD, UNIVERSIDAD DEL ROSARIO, *Fibrosis quística, 11 Fascículo Interactivo de 16 afectaría a uno de cada cinco mil recién nacidos*, ESCUELA DE MEDICINA Y CIENCIAS DE LA SALUD; Universidad del Rosario; Tomo V.

- [12] ASOCIACIÓN BIOQUÍMICA ARGENTINA. M. M. TIBOLLA, N. G. MÁRQUEZ, M. GUASTAVINO, C. ZACHAR-ZEWSKI, G. DUSSE, C. INSAURRALDE, A. ERRASTI CÁTEDRA DE BIOQUÍMICA CLÍNICA III, *MÉTODO SECUENCIAL PARA DETECCIÓN TEMPRANA DE FIBROSIS QUÍSTICA EN LA CIUDAD DE POSADAS* Departamento de Bioquímica Clínica, Carrera de Bioquímica, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones. REVISTA BIOQUIMICA Y PATOLOGIA CLINICA VOL 72 N° 3 2008;
- [13] HÉCTOR ESCOBAR Y AMAYA SOJO, *Artículo FIBROSIS QUÍSTICA*, Protocolos diagnósticos y terapéuticos en pediatría.
- [14] MCKUSICK., *GENETICA HUMANA*
- [15] K.P. HADELER., *MATEMÁTICA PARA BIÓLOGOS*, Versión española, Dr.Bartolomé Frontera Marquez, España, 1982.
- [16] DR. EDGARDO SEGAL., *Documento CONSENSO DE FIBROSIS QUÍSTICA*
- [17] ARUPLABORATORIES, *Folleto ESTUDIO DE PORTADORES DE FIBROSIS QUÍSTICA*
- [18] ORLANDO MESA BETANCURT, *Triangulo de pascal en el planteamiento de un problema*
- [19] DR. LUIS E. VEGA-BRICEÑO(PEDIATRA BRONCOPULMONAR), *CFTR: Más que un canal de cloro, Hospital Padre Hurtado, junio 2017.* <http://www.neumologia-pediatria.cl>
- [20] ROSA MARIA GIRON MORENO, ANTONIO SALSEDO POSADAS, *MONOGRAFÍAS FIBROSIS QUÍSTICA* ,Monografías NEUMOMADRID, Volumen VIII, España, 2005.
- [21] DR. PEDRO ASTUDILLO(NEUMÓLOGO PEDIATRA, *HISTORIA DE LA FIBROSIS QUÍSTICA* ,Clinica INDISA, 2017.<http://www.neumologia-pediatria.cl>
- [22] CIRO MARTINEZ BENCARDINO, *ESTADÍSTICA Y MUESTREO*,decima tercera edición, 2012.
- [23] CIRO MARTINEZ BENCARDINO, *ESTADÍSTICA BÁSICA APLICADA* ,cuarta edición, 2011.
- [24] ORGANIZACIÓN PANAMERICANA DE LA SALUD; ORGANIZACIÓN MUNDIAL DE LA SALUD, *LINEAMIENTOS BÁSICOS PARA EL ANÁLISIS DE LA MORTALIDAD*,Washington D.C , 2017.
- @articleColombia Médica, author = LUIS ORTIGOSA, M.D.\*, title = Fibrosis quística. Aspectos diagnósticos, year = 2007, OPTvolume = Vol. 38, OPTnumber = N° 1, OPTmonth = (Enero-Marzo),

@articleColombia,COLECCION SUS METODOLOGIAS ESTADISTICAS author = Centro de Documentación del INEI , title = Metodología para el Cálculo de los Indicadores de Mortalidad , year = 2000, OPTnumber = N° 08, OPTmonth = (Enero-Marzo),

@articleMexico,ARCHIVOS DE MEDICINA author = Alejandra Wendoly Hernández-Rodríguez1, Flor de María TrejoMedinilla, title = Estudio Genético Poblacional de Frecuencias Alélicas para 15 marcadores STR presentes en la Población del Estado de Zacatecas Aplicado a la Práctica Forense , year = 2014, OPTnumber = N° 10, <http://www.imedpub.com>,

## **WEBGRAFIA**

<http://atlasgeneticsoncology.org/Educ/CistFibID30032SS.html>,

[http://apps.who.int/gb/ebwha/pdf\\_files/EB126/B126\\_10-sp.pdf](http://apps.who.int/gb/ebwha/pdf_files/EB126/B126_10-sp.pdf),

[http://apps.who.int/gb/archive/pdf\\_files/EB116/B116\\_3-sp.pdf](http://apps.who.int/gb/archive/pdf_files/EB116/B116_3-sp.pdf)

<https://es.slideshare.net/zushu/fibrosis-quistica-pag-29>

[http://www.gastronutriped.com/files/publicaciones/publicacion\\_146.pdf](http://www.gastronutriped.com/files/publicaciones/publicacion_146.pdf)

[http://www.exactas.unlp.edu.ar/articulo/2017/9/3/fibrosis-quistica\\_\\_acuedo\\_con\\_hopitales\\_para\\_investigar](http://www.exactas.unlp.edu.ar/articulo/2017/9/3/fibrosis-quistica__acuedo_con_hopitales_para_investigar)

[http://portal.neumopediatricacolombia.com/?page\\_id=5743](http://portal.neumopediatricacolombia.com/?page_id=5743)

<https://es.slideshare.net/elmerserrano338/unidad-3-gentica-de-poblaciones>

<https://www.sefh.es/54congresoInfo/documentos/ponencias/1094.pdf>

<http://www.aruplab.com/topics/cf/resources>

[http://132.248.164.227/publicaciones/docs/apuntes\\_matematicas/38](http://132.248.164.227/publicaciones/docs/apuntes_matematicas/38).

<https://www.sangakoo.com/es/temas/binomio-de-newton\y\triangulo-de-pascal%20Teorema%20del%20Binomio.pdf>

<https://www.cun.es/diccionario-medico/terminos/ampc>

<https://www.pediatriaintegral.es/numeros-antiguos/publicacion-2012-03/fibrosis-quistica-y-sus-manifestaciones-respiratorias/>

<https://encolombia.com/medicina/revistas-medicas/neumologia/vns-124/neumol2400con-alteraciones/>