



Benemérita Universidad Autónoma
de Puebla



Facultad de Ciencias Físico-Matemáticas

Estudio multiparamétrico de series de tiempo biológicas

tesis presentada para obtener el grado de

LICENCIATURA EN CIENCIAS FÍSICA

por

Karina González León

asesorado por

Dr. Eduardo Moreno Barbosa

Dr. Benito de Celis Alonso

Puebla, Pue.
Diciembre 2018

Índice general

Agradecimientos	v
Resumen	vii
1. Introducción	1
2. Teoría	3
2.1. Series de Tiempo	3
2.1.1. Componentes de una Serie de tiempo	3
2.1.2. Clasificación descriptiva de las series temporales	4
2.1.3. Estimación de la tendencia	4
2.1.4. Diferenciación de la serie	7
2.1.5. Estimación de la estacionalidad	8
2.2. Transformada rápida de Fourier	11
2.3. Prueba de Corridas por arriba y abajo de la media	12
2.4. Estadística	12
2.4.1. Correlación Lineal	12
2.4.2. Prueba de Hipotesis	13
2.5. El Corazón y el Electrocardiograma	13
2.6. Frecuencia Cardíaca	15
2.7. Síndrome de Apnea-Hipopnea obstructiva del sueño	15
3. Metodología	17
3.1. Aplicación de la prueba de corridas por arriba y abajo de la media	17
3.2. Aplicando la Transformada discreta de Fourier	19
3.3. Obteniendo parametrización de las señales de los electrocardiogramas	20
4. Resultados	23
4.1. Prueba de corridas por arriba y abajo de la media	23
4.2. Transformada Rápida de Fourier	27
4.3. Parametrizaciones en las señales de los electrocardiogramas	30
4.4. Correlación de Datos	34
5. Conclusión	43
A. Tablas	45
B. Prueba de corridas por arriba y abajo de la media	51
C. Transformada Rapida de Fourier	57

D. Parametrización de las señales de los Electroencefalogramas	61
Bibliografía	67

Agradecimientos

Agradezco a la vida por permitirme conocer sus maravillosas y complicaciones de ella.

Le agradezco al Dr. Eduardo Moreno Barbosa por ser mi maestro y asesor de tesis, estoy muy agradecida por sus enseñanzas, apoyo y paciencia.

También quiero agradecerle a mi asesor el Dr. Benito de Celis Alonso por las revisiones y aportaciones en el desarrollo de este trabajo de tesis.

Estoy muy agradecida con mis asesores porque sin su ayuda este trabajo no hubiese sido posible, ustedes son mi ejemplo a seguir.

Al M. C. Moisés Zeleny Mora por sus consejos y enseñanzas. Gracias por ser mi compañero, amigo y novio. Te agradezco por apoyarme y acompañarme en todo momento. Siempre estaré muy agradecida con tigo por confiar y creer en mí. Gracias, I Fluvín You.

Las personas a las que estaré agradecida toda mi vida por motivarme a seguir son:

A mi papá Alberto González Hernández que siempre me motivo a seguir adelante y deseaba con todo su corazón que llegaré este día, se que si estuvieras aquí entre nosotros estarías orgulloso de mí.

A mi mamá Araceli León Rosas por darme la vida, sus enseñanzas y todo su apoyo.

A mis hermanas Lesly y Karol Naomi que siempre me apoyan en los momentos más difíciles.

A mi tía Josefina León Rosas por quererme, cuidarme y apoyarme en todo momento de mi vida. Gracias, te quiero mucho.

A mis amigos y compañeros que me acompañaron en este paso de mi vida.

Resumen

Una serie de tiempo es el conjunto de observaciones de una variable que cambia con el tiempo. El análisis de series de tiempo permite entender el comportamiento de esta variable que a su vez representa alguna característica de una población.

En este trabajo de tesis se propuso trabajar con las series de tiempo de electrocardiogramas de pacientes sanos y pacientes que padecen Síndrome de la Apnea Obstructiva del Sueño (SAOS). El SAOS es un trastorno en el cual una persona deja de respirar parcial o totalmente de forma repetida durante el sueño [1]. El factor más importante en la aparición de este síndrome es el sobrepeso. Además, los hombres son más propensos a desarrollar este síndrome. Las personas que lo padecen suelen tener problemas de la presión arterial, fallos cardíacos, latidos irregulares, diabetes y/o depresión. En la actualidad existen tratamientos para el SAOS como la ventilación nasal, dispositivos bucales, aparatos para la respiración o si es necesario una cirugía [1]. Existen diferentes estudios clínicos para la detección de este síndrome como: el análisis de las señales respiratorias de tórax o la variabilidad de la frecuencia cardíaca.

En este trabajo se estudió y analizó el comportamiento del corazón cuando se presenta y no el SAOS. Se utilizaron los datos de 57 ECG de la base de datos de Physionet, donde 12 personas eran sanas y el resto presentaba SAOS. Para analizar la información obtenida, se propusieron tres métodos: Prueba de corridas por arriba y abajo de la media, transformada de Fourier y parametrización de la señal. Inicialmente se aplicó la prueba de corridas por arriba y abajo de la media. Esto, para detectar si existió un patrón de comportamiento que incrementase o decreciese las señales. En la prueba de corridas por arriba y abajo de la media se calculó la frecuencia de 0's y 1's en L . Este valor depende de si el elemento $L[i]$ está por arriba o abajo de la media la señal. Posteriormente se analizaron los histogramas de 0's y 1's. En un segundo análisis se aplicó la transformada rápida de Fourier en la señal, la cual fue empleada para transformar las señales a un dominio de frecuencias. Este análisis permitió encontrar la frecuencia cardíaca (f_c) y el valor de la amplitud de la frecuencia cardíaca a_f . Durante la parametrización de la señal se calculó la duración de la actividad eléctrica en la onda P, grupo QRS y onda T para conocer la variación del tiempo y la amplitud de la onda del ECG en los dos tipos de voluntarios. El lenguaje de programación Python, fue el elegido para desarrollar los programas necesarios de análisis de señales ECG.

Finalmente, y tras completar los tres análisis descritos anteriormente, se buscaron correlaciones entre los datos de la frecuencia cardíaca (f_c), índice de masa corporal (IMC), edad (E), número de apneas por hora (AI), número de Hipopneas por hora (HI) y parámetros temporales. Todas estas variables permitieron conocer los cambios que hay en las señales del ECG en personas que presentan y no SAOS.

Mediante pruebas de hipótesis se llegaron a una serie de conclusiones, de entre las que destacan:

1. Se establecieron unos valores umbral de los distintos parámetros estudiados aquí, bajo los cuales una persona pudo ser identificada como con SAOS.
2. Las personas con sobrepeso y/o obesidad fueron más propensas a presentar SAOS. Al presentar

esta enfermedad la (f_{CR}) no se encontró dentro de lo normal. Esto quiere decir que las amplitudes y duraciones de las ondas ECG en sus fases P, T y QRS fueron afectadas.

Palabras claves: Series de tiempo, Análisis de series de tiempo, correlación, Electrocardiograma, Síndrome de Apnea-Hipopnea obstructiva del sueño.

Capítulo 1

Introducción

En este trabajo se estudiaron diferentes métodos para poder analizar series de tiempo de la señal dada por un electrocardiograma, y así poder interpretar y entender el comportamiento del corazón. Los datos de los electrocardiogramas fueron utilizados con el propósito de el cual es un trastorno en donde la respiración se interrumpe o se hace muy superficial durante el sueño. Utilizamos los datos obtenidos del electrocardiograma (ECG) el cual mide el ritmo y la fuerza de cada latido, de 57 personas de edad entre 27 hasta 63 años, donde 45 presentan el síndrome del sueño y 12 no. A estos grupos de personas las vamos a catalogar como personas “enfermas” y “sanas”, respectivamente.

La finalidad de este trabajo es encontrar correlaciones de las Series de Tiempo y datos médicos como edad, índice de masa corporal, AI, HI y frecuencia cardíaca, los cuales permitan identificar y comprender el comportamiento de los electrocardiogramas de las personas que presentan el SAOS.

En esta tesis se describirán algunos métodos importantes para el estudio de series de tiempo, de los cuales se escogieron tres y se utilizaron para analizar y estudiar los ECG de las personas que presentan o no SAOS. Inicialmente se aplicó la prueba de corridas por arriba y abajo de la media en las señales de los ECG, para detectar entre los valores de la duración de las ondas del ECG si existe un patrón que incremente o decrezca. Luego se aplicó la Transformada rápida de Fourier en la señal, la cual fue empleada para transformar las señales a un dominio de las frecuencias. De esta forma se pudo encontrar la frecuencia cardíaca de cada persona y el valor de la norma en la frecuencia de cada persona.

Posteriormente utilizando las señales de los electrocardiogramas, se calculó la duración de la actividad eléctrica en la onda **P**, grupo **QRS** y onda **T** para encontrar valores característicos de las señales del ECG en personas sanas y enfermas.

Utilizando los datos de la frecuencia cardíaca (f_c), índice de masa corporal (IMC), edad (E), número de apneas por hora (AI), número de Hipopneas por hora (HI) y parámetros temporales, se buscaron correlaciones entre estas, para poder encontrar valores los cuales describan a una persona que presenta o no el SAOS. En este trabajo se usó el lenguaje de programación Python para hacer los programas necesarios los cuales permitirán analizar y estudiar las señales de los ECG.

Capítulo 2

Teoría

2.1. Series de Tiempo

Una serie temporal es una sucesión de observaciones de una variable particular con respecto al tiempo, estas observaciones se registran en tiempos iguales.

Estudiar los datos de las series de tiempo tiene como objetivo encontrar patrones los cuales nos darán medidas descriptivas básicas y predecir el comportamiento de futuras observaciones [2].

2.1.1. Componentes de una Serie de tiempo

El estudio descriptivo de series temporales se basa en la idea de descomponer la variación de una serie en varias componentes básicas. Para poder encontrar el patrón de las series de tiempo, es necesario conocer sus componentes o fuentes de variación que se consideran son las siguientes:

- **Tendencia** La tendencia de una serie de tiempo caracteriza el movimiento hacia arriba o hacia abajo con respecto a un periodo de tiempo. Esta tendencia refleja el crecimiento o declinación de larga duración en la serie de tiempo.
- **Ciclo** El ciclo de una serie de tiempo caracteriza el movimiento hacia arriba o hacia abajo alrededor de los niveles de tendencia.
- **Variación estacional** Las variaciones estacionales son patrones periódicos los cuales ocurren año tras año en los mismos meses (o en los mismos trimestres) del año poco más o menos con la misma intensidad.
- **Fluctuaciones irregulares** Las fluctuaciones irregulares son movimientos erráticos que siguen un patrón indefinido o irregular.

Dadas estas componentes, se puede representar la señal y_t como:

$$y_t = T_t + E_t + I_t \quad (2.1)$$

donde T_t es la tendencia, E_t es la componente estacional e I_t es el ruido o parte aleatoria que constituyen la señal.

2.1.2. Clasificación descriptiva de las series temporales

Las series temporales se pueden clasificar en:

- **Estacionarias:** Una serie es estacionaria cuando la media μ y la varianza s^2 no cambian con el tiempo es decir son constantes. Este tipo de series son mucho mas fáciles de predecir su comportamiento en el futuro.

Como ejemplo podemos graficar el número de reprobados de la clase Cálculo Diferencial por año como se muestra en la figura 2.1, esta serie es estable alrededor de un valor central. Si representamos un histograma de esta serie, se podrá describir adecuadamente la información: en promedio hay 8 alumnos reprobados por año, por lo tanto la mejor predicción para el próximo año sería la media de la serie.

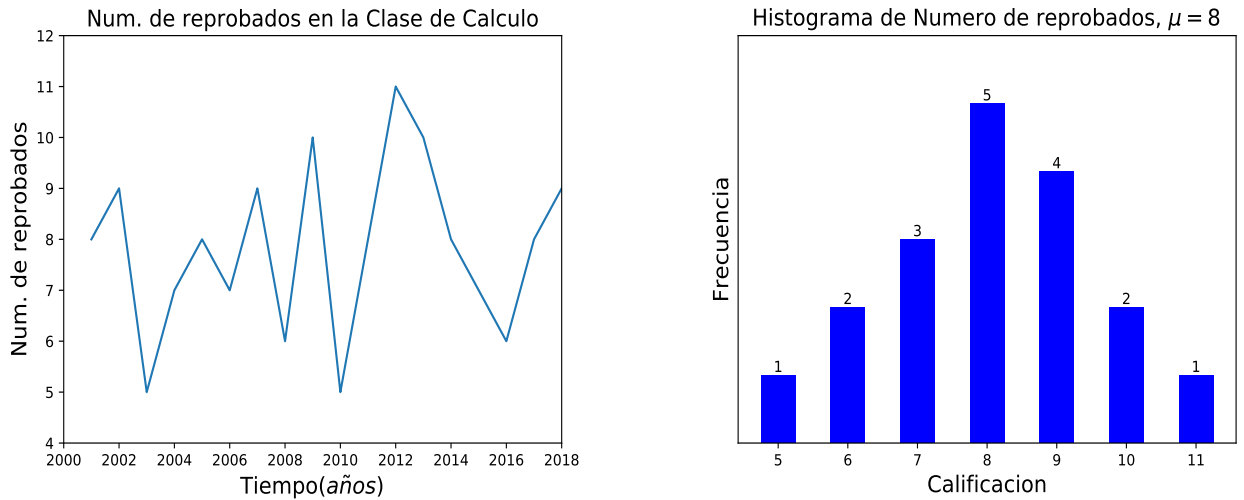


Figura 2.1: Número de reprobados en la clase de Calculo Diferencial y su histograma

- **No Estacionarias:** Son series en las cuales la media y/o variabilidad cambian en el tiempo, es decir los cambios en la media determinan una tendencia a crecer o decrecer a largo plazo, por lo que la serie no oscila alrededor de un valor constante.

2.1.3. Estimación de la tendencia

Para estimar la tendencia se supondrá que y_t es una serie no estacionaria sin componente estacional, es decir que la serie se puede descomponer en $y_t = T_t + I_t$.

- **Tendencia determinista :** Una serie tiene una tendencia determinista si no hay incertidumbre alguna sobre la evolución futura de la tendencia, es decir es posible prever su futuro conociendo su pasado. La función mas sencilla posible que describa la tendencia determinista es una recta, es decir,

$$T_t = a + bt \tag{2.2}$$

donde a y b se estiman mediante el método de mínimos cuadrados, la forma de estimar estas constantes es mediante un modelo de regresión lineal entre las variables y_t y el tiempo $t = 1, 2, 3, \dots$ donde

$$a = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{n(\sum x_i^2) - (\sum x_i)^2}, \quad b = \frac{(\sum y_i) - a(\sum x_i)}{n}$$

De esta forma, si estimamos los parámetros a y b , entonces la componente irregular será $I_t = X_t - a - bt$.

Como ejemplo se presenta la ventas por semana de un nuevo producto para el crecimiento del cabello durante 1 mes, al calcular su aproximación por mínimos cuadrados, se encuentra una aproximación lineal como se muestra en la figura 2.2,

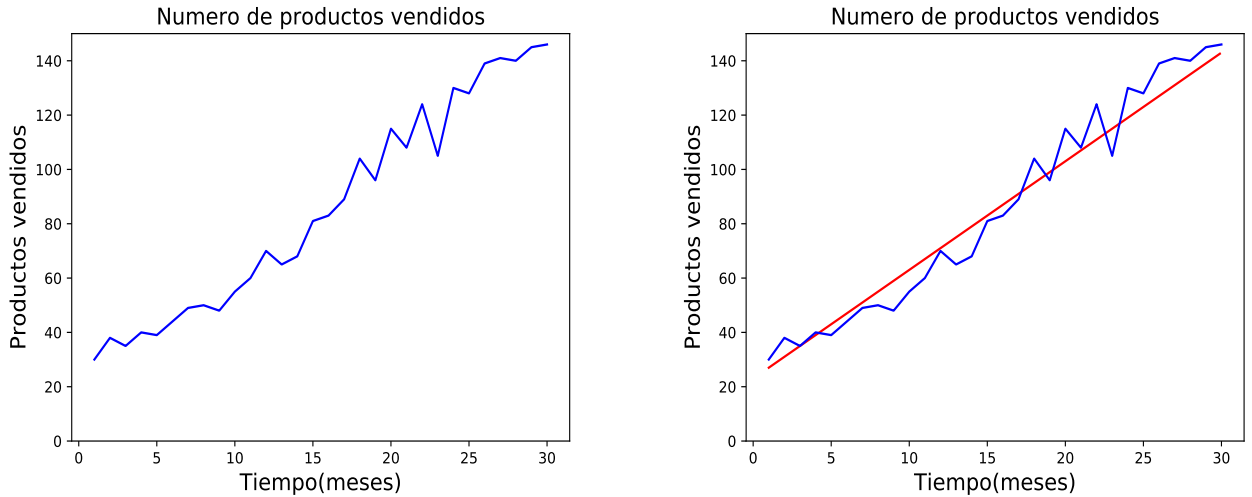


Figura 2.2: Número de reprobados en la clase de Calculo Diferencial y su tendencia.

Usando la ecuación 2.2 obtenemos que la recta que se aproxima es $T_t = 4 * x + 23$. Al graficar la componente irregular I_t , se obtiene la siguiente gráfica (figura 2.3):

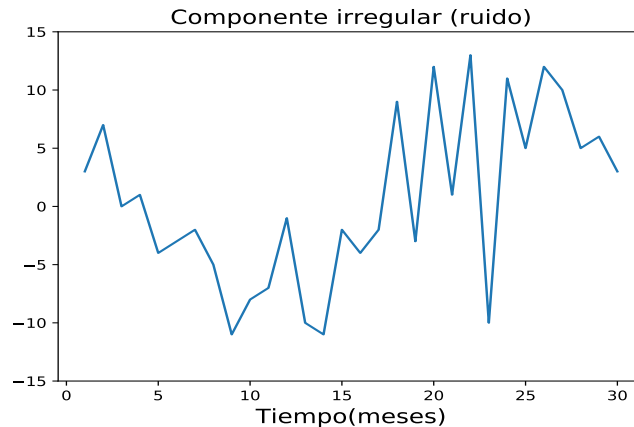


Figura 2.3: Componente irregular del número de productos vendidos.

Lo que queda es una serie estacionaria. En algunos casos, no es posible ajustar la tendencia mediante una recta, en estos casos lo mejor es ajustar la tendencia a un polinomio o a la curva que mejor se pueda ajustar.

- **Tendencia evolutiva** La tendencia evolutiva es una función que evoluciona lentamente y que puede aproximarse en intervalos muy cortos, por una función simple del tiempo, donde la representación de la tendencia por una recta es válida para tres períodos consecutivos de tiempo $t - 1$, t y $t + 1$, se representan las tendencias en los tres periodos consecutivos de la siguiente manera:

$$\begin{aligned} T_{t-1} &= T_t - \text{crecimiento} \\ T_t &= T_t \\ T_{t+1} &= T_t + \text{crecimiento} \end{aligned} \tag{2.3}$$

Si calculamos la media (m) de tres observaciones consecutivas,

$$m_t = \frac{y_{t-1} + y_t + y_{t+1}}{3} \tag{2.4}$$

y como la serie se descompone en $y_t = T_t + I_t$, entonces

$$m_t = T_t + \frac{I_{t-1} + I_t + I_{t+1}}{3} \tag{2.5}$$

y como la componente irregular tiene media cero, la media de los tres valores de la componente irregular se puede suponer que es despreciable frente a la tendencia, y m_t representa la tendencia en ese instante, esta operación se denomina **media móvil** de orden tres. Realizando esta operación se pierde la primera observación y la última.

Como ejemplo se muestra el reporte de ventas de una marisquería correspondiente a este año durante 48 días, estos datos se muestran en la siguiente figura 2.4:

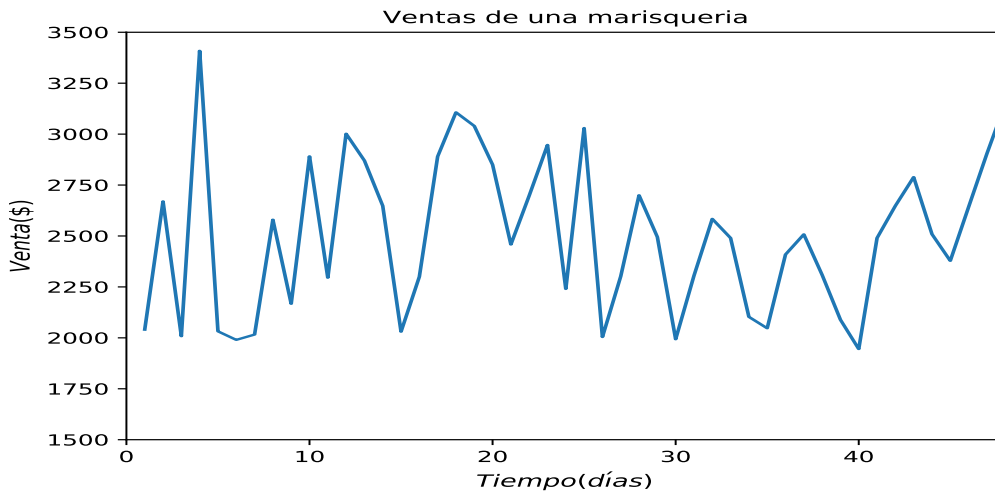


Figura 2.4: Datos de la venta de un negocio en 48 días

Usando la media móvil de orden tres, se obtiene una lista de números en la cual cada uno es el promedio de un subconjunto de los datos originales, esta es representada en la figura 2.5. Y al graficar la componente irregular se obtiene la figura 2.6, donde la tendencia esta dada por $T_t = -0.801 * x + 24.91$.

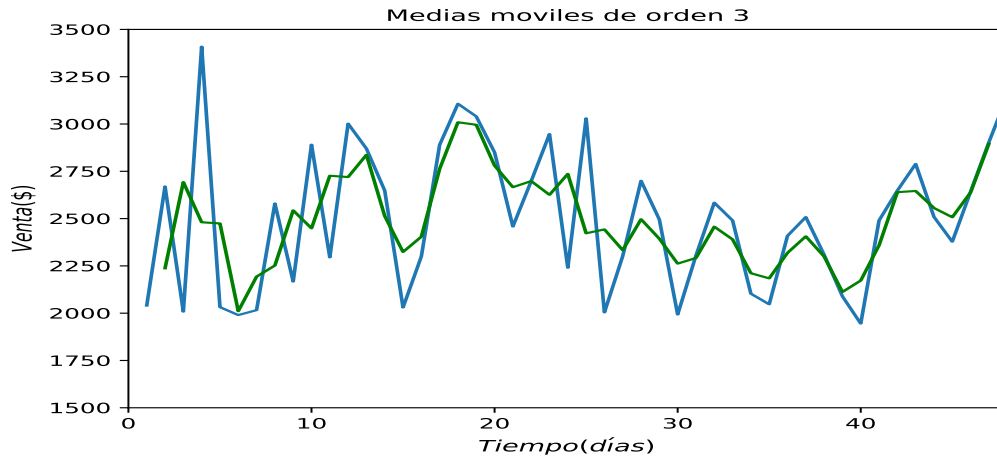


Figura 2.5: En color verde se representa la media móvil de orden 3

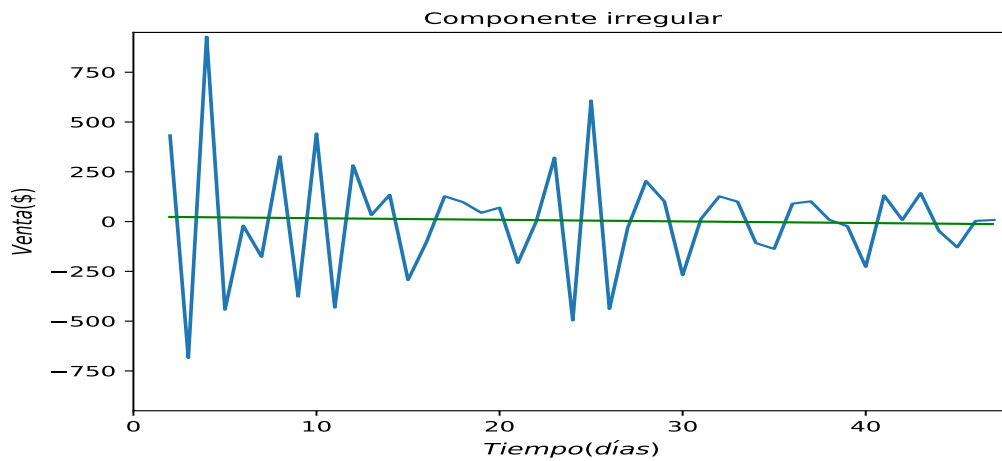


Figura 2.6: En color verde se representa la componente irregular

Como se a visto en la estimación de la tendencia, al encontrar la tendencia de una serie de datos y_t , esta es eliminada para poder encontrar la componente irregular, a continuación se da un método mas.

2.1.4. Diferenciación de la serie

Este método permite eliminar la tendencia el cual consiste en suponer que la tendencia evoluciona lentamente en el tiempo, de manera que en el instante t la tendencia debe estar próxima a la tendencia en el instante $t - 1$. De esta forma, si restamos a cada valor de la serie el valor anterior, la serie resultante estará aproximadamente libre de tendencia, este método consiste en pasar de la serie original y_t a la serie x_t mediante:

$$x_t = y_t - y_{t-1} \quad (2.6)$$

De este modo, la serie diferenciada resulta **ser estacionaria**.

Como ejemplo mostramos el valor promedio del dolar (\$) en los meses de agosto, septiembre y octubre del año en curso como se muestra a continuación en la figura 2.7:

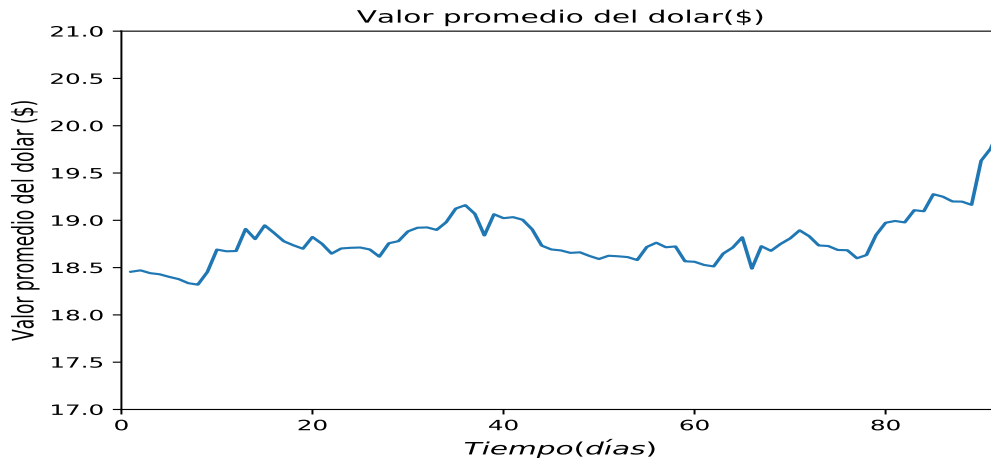


Figura 2.7: Valor promedio del dolar(\$), de los últimos tres meses en curso

Al diferenciar estos valores se obtiene la siguiente figura 2.8:

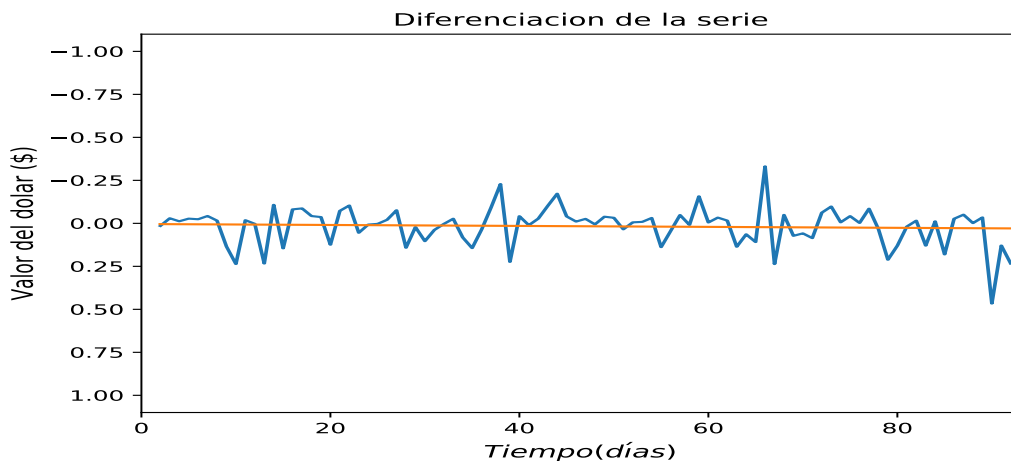


Figura 2.8: Diferenciación del valor del dolar en los últimos tres meses

2.1.5. Estimación de la estacionalidad

Un método para estimar el efecto estacional (de cada mes) es considerar cómo varía la media del período (mes) respecto de la media global.

Meses	1	2	...	n	Medias	S
Enero	x_{11}	x_{12}	...	x_{1n}	\bar{x}_1	S_1
Febrero	x_{21}	x_{22}	...	x_{2n}	\bar{x}_2	S_2
⋮	⋮	⋮	...	⋮	⋮	⋮
Noviembre	x_{111}	x_{112}	...	x_{11n}	\bar{x}_{11}	S_{11}
Diciembre	x_{121}	x_{122}	...	x_{12n}	\bar{x}_{12}	S_{12}

Existe un efecto estacional mensual, por tanto, existen 12 coeficientes estacionales, uno para cada mes del año, para encontrar estos valores primero se calcula la media de las observaciones para cada mes, $\bar{x}_1, \dots, \bar{x}_{12}$, y el coeficiente estacional se calcula como :

$$S_i = \bar{x}_i - M, \quad i = 1, \dots, 12 \quad (2.7)$$

donde M es la media total de las observaciones.

Como ejemplo se muestran los valores del precio (\$) de un producto en una dulcería durante los últimos 7 años en la siguiente tabla, donde se calculó la media y los coeficientes estacionales, estos datos se muestran a continuación:

Meses	2012	2013	2014	2015	2016	2017	2018	Media	Coficiente estacional
Enero	5.49	5.65	5.75	5.83	5.89	6.03	6.49	5.8757	-0.1123
Febrero	5.45	5.62	5.71	5.76	5.83	5.97	6.20	5.7914	-0.1966
Marzo	5.59	5.76	5.87	5.89	6.01	6.04	6.35	5.93	-0.0580
Abril	5.59	5.75	5.85	5.85	5.98	6.13	6.25	5.9142	-0.0738
Mayo	5.60	5.76	5.87	5.89	6.04	6.16	6.21	5.9328	-0.0552
Junio	5.75	5.92	6.05	6.08	6.16	6.28	6.40	6.0914	0.1033
Julio	5.90	6.02	6.14	6.20	6.31	6.43	6.75	6.25	0.2619
Agosto	5.85	6.00	6.15	6.22	6.33	6.41	6.70	6.2371	0.2490
Septiembre	5.74	5.87	6.00	6.05	6.14	6.23	6.50	6.0757	0.0876
Octubre	5.61	5.72	5.85	5.88	6.01	6.13	6.30	5.9285	-0.0595
Noviembre	5.47	5.60	5.72	5.74	5.89	5.97	6.35	5.82	-0.1680
Diciembre	5.63	5.72	5.82	5.85	6.00	6.30	6.75	6.01	0.0219

En la figura 2.9 se graficaron los coeficientes estacionales, y de igual manera se muestra la tendencia en verde.

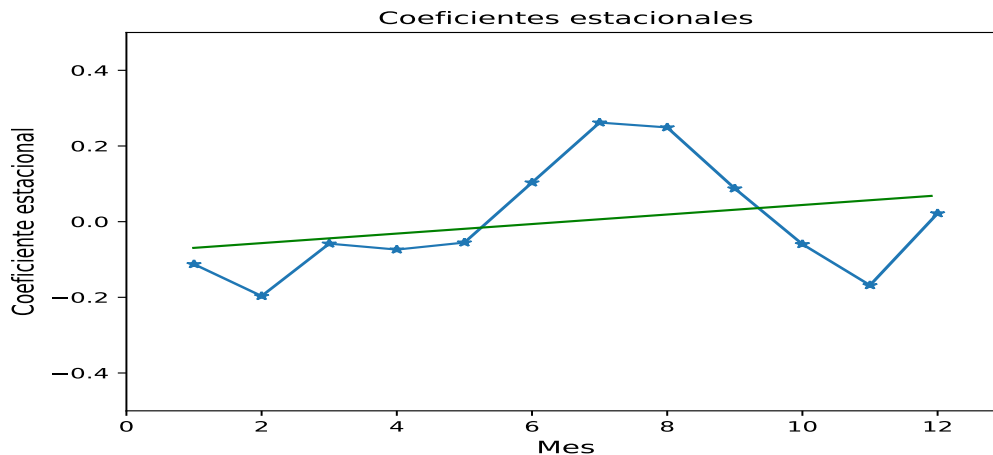


Figura 2.9: Valores de los coeficientes estacionales de cada mes

Se denomina *serie desestacionalizada* a una serie donde se ha eliminado el efecto de cada mes y el cual se obtiene restando al valor de cada mes el coeficiente estacional de dicho mes. En este ejemplo, la serie desestacionalizada se muestra en la figura 2.10:

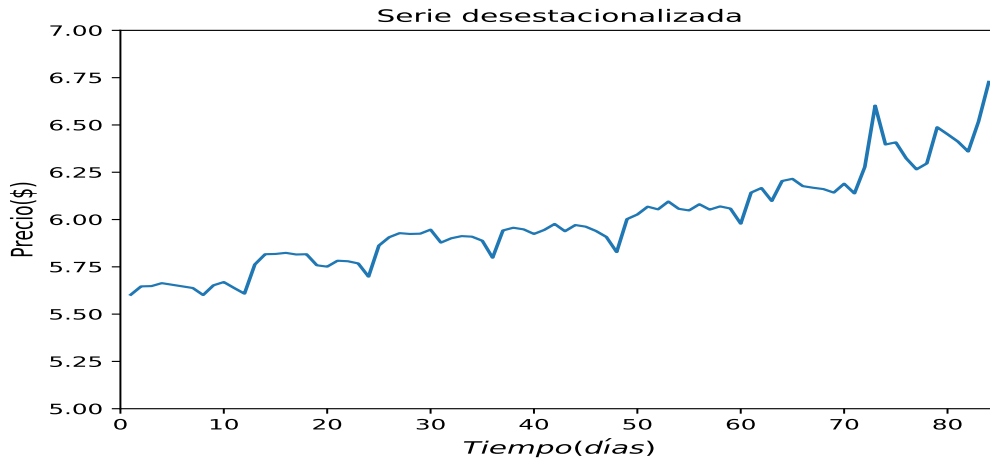


Figura 2.10: Serie desestacionalizada de los precios del producto

es posible observar que esta serie ya no tiene estacionalidad, pero todavía tiene tendencia. Para eliminar la tendencia se procedió a aplicar *Diferenciación de la serie* a la serie desestacionalizada, al diferenciar esta serie se obtiene la siguiente figura 2.11:

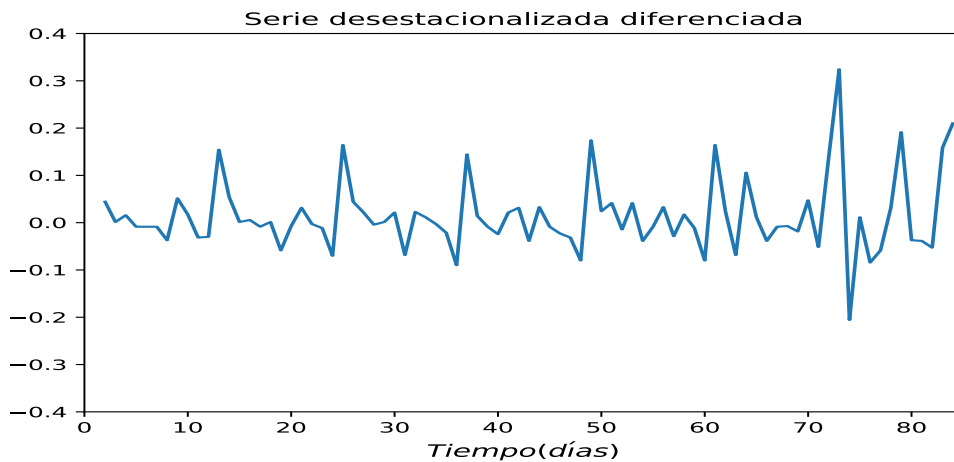


Figura 2.11: Diferenciación de la serie desestacionalizada.

esta serie vuelve a ser estacional, esto es, porque al diferenciar ha aparecido otra vez un efecto estacional. Para solucionar este problema, lo más apropiado, es primero eliminar la tendencia de la serie y en segundo lugar desestacionalizar la serie.

Por último se muestra en la figura 2.12 la serie año a año. La variación de esta serie es muy parecida todos los años, lo cual indica que hay *estacionalidad*. Las gráficas de los siete años están en orden creciente, es decir la del primer año está por debajo de la del segundo año, la del segundo año por

debajo de la del tercer año y así sucesivamente. Esto indica que la serie tiene una tendencia creciente.

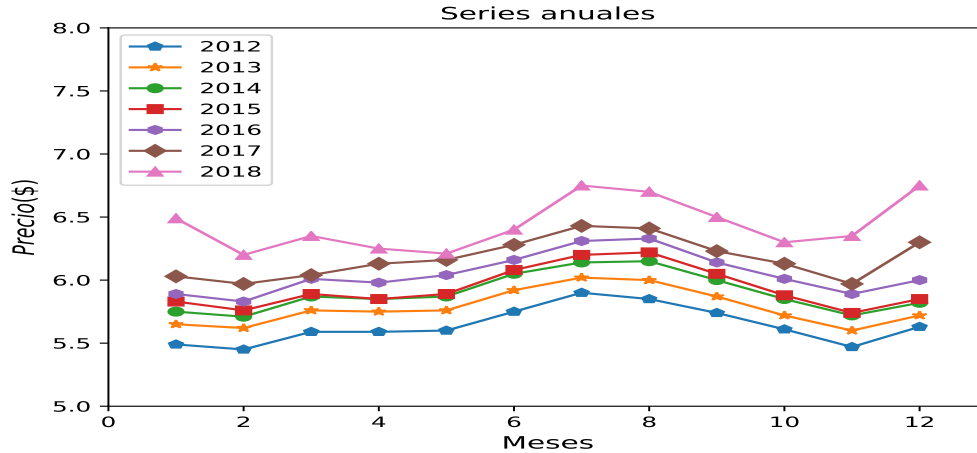


Figura 2.12: Precio del dulce en los últimos 7 años.

2.2. Transformada rápida de Fourier

Al considerar la transformada de Fourier en un espacio discreto de puntos esta será una aproximación de la transformada continua, la cual se llama Transformada Discreta de Fourier (DFT) [3]. Se considera una DFT de N puntos $F_N(k)$, donde $k = 0, \dots, N - 1$:

$$F_N(k) = \sum_{n=0}^{N-1} f(n)e^{i2\pi kn/N} \quad (2.8)$$

La Transformada Rápida de Fourier (FFT) es un algoritmo para el cálculo de la DFT. La FFT elimina una gran parte de los cálculos de la DFT, por lo tanto el cálculo es más rápido. Es útil en el procesamiento digital de señales ya que nos permite analizar la señal en el dominio del tiempo.

Para calcular la transformada de Fourier discreta, en Python se usó `numpy.fft.fft`, esta función calcula la DFT usando la Transformada Rápida de Fourier.

Como ejemplo se creó una gráfica con valores aleatorios la cual tenía una distribución gaussiana con una media $\mu = 0$ y desviación estándar $\sigma = 0.25$. Usando la Transformada Rápida de Fourier se obtuvo la siguiente figura 2.14, la cual nos muestra las frecuencias de esta serie.

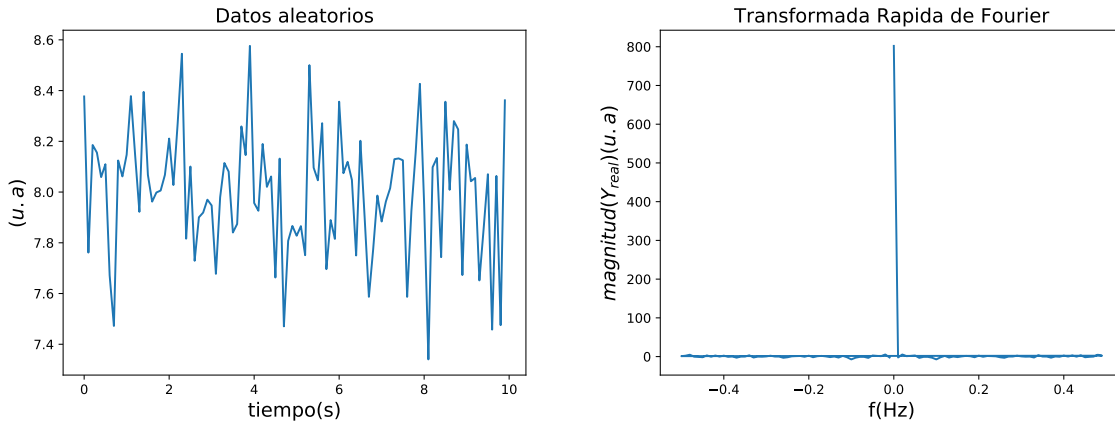


Figura 2.13: Valores aleatorios con una media $\mu=0$ y Figura 2.14: Transformada de Fourier de los valores desviación estandar $\sigma=0.25$.

2.3. Prueba de Corridas por arriba y abajo de la media

La prueba del corridas sirve para probar la aleatoriedad de algún conjunto de datos. Esta prueba detecta entre los valores de la señal si existe un patrón que incremente o decrece.

El procedimiento de esta prueba consiste en determinar una secuencia de unos y ceros, de acuerdo con una comparación entre los números del conjunto y_i y μ . Posteriormente se determina el número de corridas observadas, y los valores de n_0 y n_1 . C_0 es el número de corridas en la secuencia, determinado de la misma manera que en la prueba de corridas arriba y abajo; n_0 es igual a la cantidad de ceros en la secuencia, y n_1 es igual a la cantidad de unos en la secuencia, cumpliéndose que $n_0 + n_1 = n$. Luego se calcula el valor esperado, la varianza del número de corrida y el estadístico

El cual consiste en lo siguiente:

Denotamos con el valor 0 aquel número que se encuentre por debajo de la media.

Denotamos con el valor 1 aquel número que se encuentre por encima de la media.

2.4. Estadística

2.4.1. Correlación Lineal

El coeficiente de correlación lineal r que es una medida numérica de la fuerza de la relación entre dos variables que representa datos cuantitativos. La fuerza de la relación se determina mediante la magnitud del efecto que cualquier cambio en una variable ejerce sobre la otra. Sean x y y las variable, se tendrá el siguiente criterio para r : si al crecer x no se observan modificaciones definidas en los valores de y entonces diremos que no se encuentra correlación y si cuando aumentamos x observamos cambios definidos en los valores de y , diremos que la correlación es positiva cuando y aumenta y negativa cuando y decrece [4].

El coeficiente de correlación lineal r , tendrá siempre un valor entre -1 y +1. Los posibles valores de r depende de su diagrama de dispersión como se observa en la figura 2.15. La correlación lineal perfecta aparece en la situación en que todos los puntos quedan exactamente situados sobre una recta, por lo tanto $r = +1$ o $r = -1$. Si los datos quedan en una recta horizontal o vertical, hablamos de una correlación cero ($r = 0$), entonces los cambios en una variable no tienen efecto sobre la otra por lo tanto se dice que no hay correlación lineal [5]. Se considerará que la correlación es alta cuando se aproxima a +1 o -1

y baja cuando su valor se acerca a cero. El valor del coeficiente lineal para una muestra se calcula por medio de la siguiente formula:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}} \quad (2.9)$$

donde:

n representa el número de pares de datos presentes.

Cuando se ha calculado el valor de un coeficiente de correlación a partir de un conjunto de datos, lo denotamos por r^* .

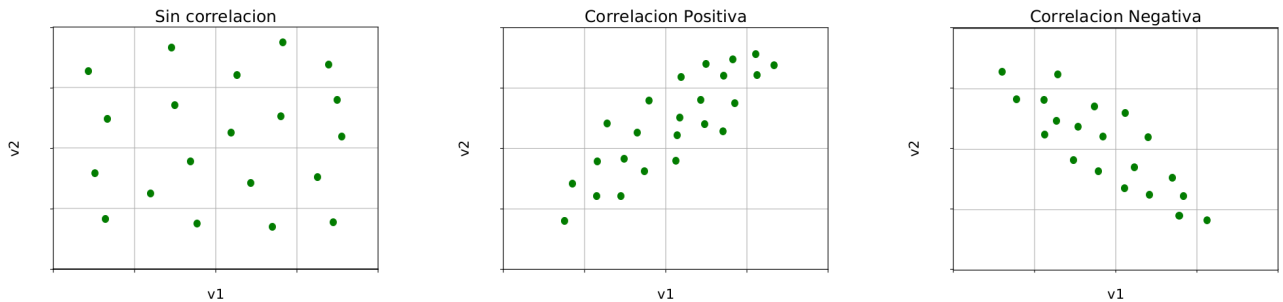


Figura 2.15: Diagramas de dispersión

2.4.2. Prueba de Hipotesis

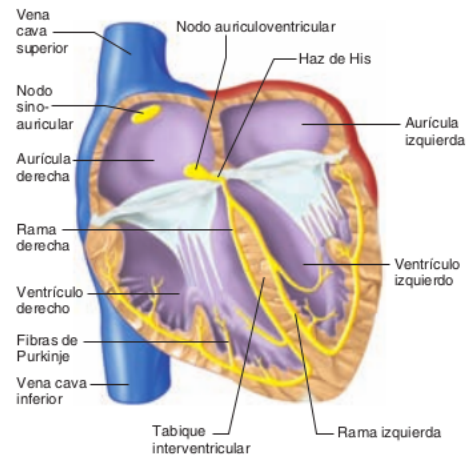
La prueba de hipótesis consiste en formular una hipótesis nula (H_o) y una hipótesis alterna (H_a), es decir determinar el parámetro de la población en interés y proponer la validez de un valor para las dos hipótesis. Al rechazar H_o se acepta H_a . Si los valores del estadístico de prueba se encuentran en el intervalo de aceptación se decide no rechazar H_o y si se encuentran en la región crítica se rechaza. La región crítica se determina de manera que los valores del estadístico de prueba se encuentren en la región de rechazo de H_o sea igual al nivel de significancia (α). Al tomar una decisión de rechazar o no rechazar H_o permite llegar a una conclusión.

2.5. El Corazón y el Electrocardiograma

El corazón es el órgano principal del sistema circulatorio. Con ayuda de los vasos sanguíneos hace circular la sangre a través del cuerpo, permitiendo llegar el oxígeno, nutrientes a los tejidos del cuerpo.

El corazón está formado de 3 capas: Pericardio (capa que lo envuelve por fuera), miocardio (músculo cardíaco, encargado de impulsar la sangre por el cuerpo mediante su contracción) y endocardio (capa fina que lo recubre por dentro). Este órgano muscular hueco está dividido en dos mitades (izquierda y derecha) por un músculo-membranoso, donde cada una constituye a su vez dos cavidades: una superior, la aurícula y una inferior, el ventrículo [6]. Estas partes se muestran en la figura 2.16.

Figura 2.16: Partes del corazón, imagen tomada de [6].



El latido cardíaco se origina en un sistema de conducción cardíaca y se extiende por este sistema a todas las partes del miocardio. El sistema de conducción está formado por músculo cardíaco modificado, el cual posee menos estriaciones y límites indistintos. Las diversas partes del sistema de conducción y las partes del miocardio, son capaces de emitir una descarga espontánea [6].

Las partes del corazón laten en una secuencia ordenada: la contracción de las aurículas (sístole auricular) que va seguida de la contracción de los ventrículos (sístole ventricular) y durante la diástole las cuatro cavidades se relajan. Esta secuencia puede ser observada en la figura 2.17, donde la onda P se produce por la despolarización auricular, el complejo QRS por la despolarización ventricular y la onda T por la repolarización ventricular. En el la figura 3.4 se muestran estas las ondas.

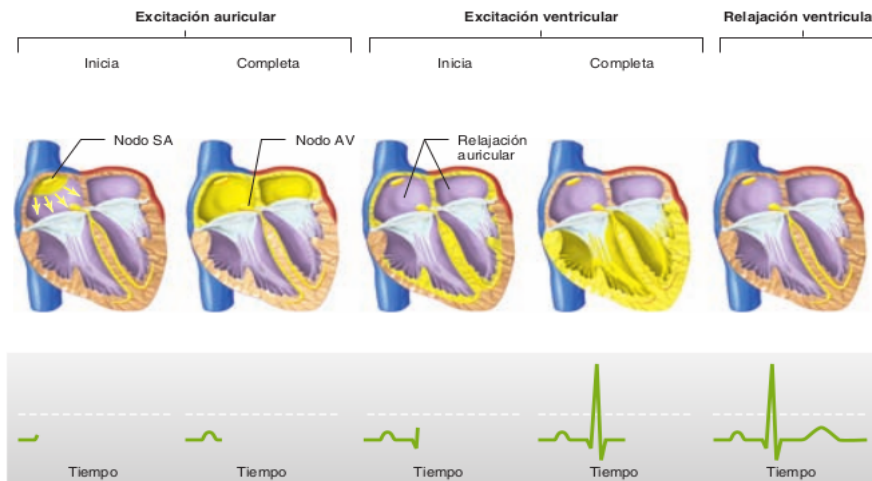


Figura 2.17: Electrocardiograma, imagen tomada de [6].

La buena conducción eléctrica de los líquidos corporales hacen que el cuerpo humano sea un conductor de volumen. Si sobre la piel se colocan electrodos en posiciones adecuadas, es posible registrar las variaciones del potencial, el registro gráfico se le conoce como **electrocardiograma** [7].

El electrocardiograma puede ser medido, empleando dos electrodos activos o mediante un electrodo activo y uno indiferente que se conecta a potencial cero, en dos puntos distintos. La disposición de las conexiones de cada par de electrodos se conoce como derivación, esta es clasificada en tres grupos: derivaciones estándar de las extremidades, derivaciones aumentadas de las extremidades y derivaciones precordiales.

Derivaciones estándar de las extremidades (también conocidas como derivaciones de Einthoven), se registran las diferencias de potenciales entre dos extremidades, donde los electrodos pueden ser conectados en el brazo izquierdo, brazo derecho y pierna izquierda.

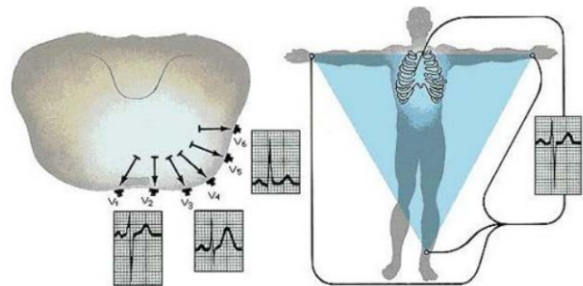


Figura 2.18: Derivaciones precordiales, imagen tomada de [7].

Las derivaciones aumentadas de las extremidades consiste en tomar la diferencia de potencial entre una extremidad y un electrodo virtual cuyo potencial es el promedio de los tres electrodos utilizados en las derivaciones de Einthoven (denominado electrodo central de Wilson).

Las derivaciones precordiales miden la diferencia de potencial entre el electrodo central de Wilson y distintos puntos del tórax, a este conjunto de derivaciones unipolares son llamadas V1,V2,...,V6. La ubicación del electrodo activo son las siguientes: V1 (Cuarto espacio intercostal, dos centímetros a la derecha de la línea medio-esternal), V2 (Cuarto espacio intercostal, dos centímetros a la izquierda de la línea medio-esternal), V3 (Punto medio entre V2 y V4), V4 (Quinto espacio intercostal izquierdo, en la línea media clavicular),V5 (Quinto espacio intercostal izquierdo, en la línea axilar anterior) y V6 (Quinto espacio intercostal izquierdo, en la línea axilar media), estas son señaladas en la figura 2.18.

Estas derivaciones precordiales, permitirán registrar la actividad eléctrica del corazón en el plano horizontal, a diferencia de todas las anteriores que lo hacen en un plano coronal.

2.6. Frecuencia Cardíaca

La frecuencia cardíaca (f_C) representa el número de veces que el corazón se contrae durante un minuto. Es importante que los deportistas conozcan su frecuencia cardíaca para realizar un entrenamiento deportivo adecuado y de esta manera cuidar su integridad física [8]. Frecuencia Cardíaca en Reposo (f_{CR}) es el valor de la frecuencia cardíaca cuando se tiene menor actividad física, es decir en reposo . El valor de f_{CR} depende de los hábitos alimenticios, el entrenamiento, el sueño, el nivel de estrés mental. Los valores medios de la F_{CR} en función de la edad y del sexo se muestran en la siguiente tabla:

Tabla 2.1: Valores de la frecuencia cardíaca en reposo en hombres y mujeres

HOMBRES					MUJERES				
Edad	Mal	Normal	Bien	Excelente	Edad	Mal	Normal	Bien	Excelente
20 - 29	86 +	70-84	62-68	60 o menos	20 - 29	96 +	78-94	72-76	70 o menos
30 - 39	86 +	72-84	64-70	62 o menos	30 - 39	98 +	80-96	72-78	70 o menos
40 - 49	90 +	74-88	66-72	64 o menos	40 - 49	100 +	80-98	74-78	72 o menos
50 +	90 +	76-88	68-74	66 o menos	50 +	104 +	84-102	76-86	74 o menos

La frecuencia cardiaca refleja la actividad del sistema nervioso y su disminución se asocia a la fibrilación auricular, infarto de miocardio, obesidad, resistencia de la insulina y aterosclerosis [9].

2.7. Síndrome de Apnea-Hipopnea obstructiva del sueño

El síndrome de apnea-hipopnea obstructiva del sueño (*SAHOS*) es un trastorno en el cual una persona deja de respirar parcial o totalmente de forma repetida durante el sueño [10]. Es caracterizado por episodios recurrentes de un colapso parcial o completo de las vías respiratorias, estos episodios provocan descensos en la saturación de oxígeno que es la causa probable de las alteraciones cardiovasculares a medio-largo plazo [11] .

Una apnea se define como la ausencia o reducción superior al 90 % en la amplitud de la señal de flujo respiratorio de más de 10 segundos de duración. La apnea es obstructiva si se acompaña de esfuerzo respiratorio, es central en ausencia de dicho esfuerzo respiratorio y es mixta si comienza como central y termina con esfuerzo respiratorio. La hipopnea es una reducción de la señal de flujo superior a 30 % e inferior a 90 % de más de 10 segundos de duración [11] .

Este síndrome suele interrumpir el sueño durante la noche, por lo que puede provocar estar somnoliento durante el día, ya que el sueño es necesario para mantener el equilibrio metabólico calórico, el equilibrio térmico y la función inmunitaria. La obstrucción de las vías respiratorias provoca que las personas ronquen durante la noche. Las personas con apnea del sueño tienen mayor riesgo de sufrir accidentes de tránsito, accidentes de trabajo y otros problemas médicos. La apnea del sueño puede tratarse

mediante ventilación nasal [1], para mejorar la calidad de vida de una persona con este síndrome puede acudir a dispositivos bucales, aparatos para la respiración o si es necesario una cirugía.

La apnea puede afectar a cualquier edad, aunque los hombres son mas propensos a desarrollar este síndrome. El riesgo es también mayor para aquellos: que son mayores de 40 años, con sobrepeso, con amígdalas grandes, una lengua larga o mandíbulas pequeñas, con historia familiar de apnea del sueño, con una obstrucción nasal debido a un tabique desviado, alergias o un problema sinusal. La apnea obstructiva del sueño (OSA) afecta a cerca de 24 % de varones y 9% de las mujeres en edad madura en Estados Unidos [6].

Los problemas que puede causar si no se trata el problema de la apnea son: Alta presión arterial, derrame, fallos cardíacos, latidos irregulares, ataque cardíaco, diabetes y/o depresión [1].

Capítulo 3

Metodología

Las señales del ECG fueron analizadas con: Transformada rápida de Fourier, Prueba de corridas por arriba y abajo de la media y parametrización de la señal. Todo esto en personas que presentan el SAOS (enfermas) y otras que no (sanas).

En *physionet.org* se encontró bases de datos de señales obtenidas de electrocardiogramas en pacientes que presentaban y no SAOS. También en esta pagina se proporciono datos personales de cada persona como la edad, altura, peso, AI, HI y AHI(ver tabla A.1). Se decidió trabajar con estas señales porque son señales periódicas de las que es muy fácil obtener parámetros temporales.

Se obtuvo los datos de ECG de 70 personas, 57 hombres y 13 mujeres. Estas señales fueron tomadas usando derivaciones precordiales. Como la población de sexo masculino fue mayor que del femenino, solo se tomaron en cuenta las bases de datos en hombres de edad entre 27 y 63 años. En esta muestra hay 12 personas que no presentan valores de AI, HI y AHI y todas las demás si presentan.

Se comenzó por la instalación del paquete de Software WFDB para leer y convertir los registros de formato EDF en un archivo de texto (txt), escribiendo en la terminal:

```
wget dirección_del_archivo
rdsamp -r nombre_archivo.edf-p >nombre_archivo.txt
```

El registro de los datos fué tomados desde el inicio hasta el final del sueño. Todas las bases de datos tienen mediciones en un intervalo de tiempo de 0.01 s. En esta tesis se utilizaron los datos registrados a partir de 20 minutos para aplicar los métodos seleccionados y de esta manera poder estudiar y analizar el comportamiento de la señal.

3.1. Aplicación de la prueba de corridas por arriba y abajo de la media

Sea la señal (x, y) , donde x es el tiempo, y y la señal del electrocardiograma, dada en milivolts (mv). Utilizando estos datos se buscó la tendencia determinista (ver sección 2.1.3), utilizando el método de mínimos cuadrados, se encuentran los valores de \mathbf{a} y \mathbf{b} , para poder conocer la función promedio, como se muestra en la figura 3.1

$$P(x) = a * x + b$$

Se creo una nueva lista \mathbf{l} con 0's y 1's, los cuales fueron calculados de la siguiente manera. Para algún elemento en la posición i de la lista \mathbf{y} , $\mathbf{l}[i]$ depende del siguiente criterio:

3.3. Obteniendo parametrización de las señales de los electrocardiogramas

Al graficar la actividad eléctrica del corazón dada por la base de datos, se pueden observar las ondas de electrocardiogramas, es decir, la onda "P", "Q", "R", "S" y "T".

Se hizo un programa en Python mostrado en el apéndice D, el cual calculó la duración temporal (s), así como los valores alcanzados en milivoltios (mv), en las que duró la onda P, el grupo QRS y la onda T, para encontrar estos valores se encontraron las posiciones de inicio y final de las ondas P, QRS y T de las señales de los electrocardiogramas, estos puntos serán nombrados (x_1, y_1) , (x_2, y_2) , ..., (x_9, y_9) , las posiciones de estos puntos se muestran en la siguiente figura 3.4, la medición de estos puntos se hizo para 10 s , como ejemplo en la figura 3.5. Se pueden observar los puntos dados por el programa en la señal. Los valores de estos puntos fueron encontrados para cada persona.

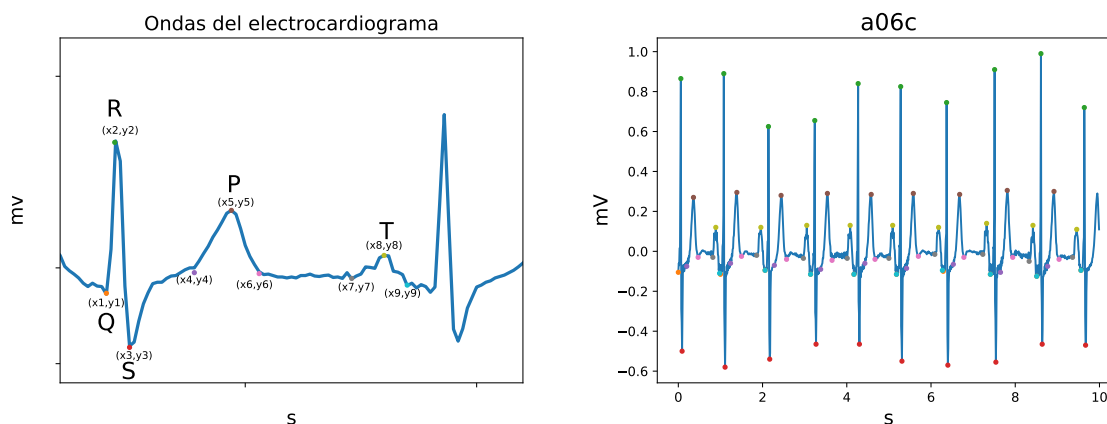


Figura 3.4: Puntos que identifican el Grupo QRS, la onda P y onda T. Figura 3.5: Puntos encontrados en la señal utilizando el programa

Dado a que se obtienen tres puntos descriptivos de la onda, se calcularon las diferencias descritas en la tabla 3.1, como ejemplo se muestra en la figura 3.6 las diferencias de $x_3 - x_1$, $y_2 - y_1$ y $y_3 - y_2$, para la onda QRS, de igual manera se calculó $x_6 - x_4$, $y_5 - y_4$ y $y_6 - y_5$ para la onda P y finalmente $x_9 - x_7$, $y_8 - y_7$ y $y_9 - y_8$ para la onda T.

Tabla 3.1

	Diferencia 1		Diferencia 2
Onda QRS	$x_3 - x_1$	$y_2 - y_1$	$y_3 - y_2$
Onda P	$x_6 - x_4$	$y_5 - y_4$	$y_6 - y_5$
Onda T	$x_9 - x_7$	$y_8 - y_7$	$y_9 - y_8$

CAPÍTULO 3. METODOLOGÍA
3.3. OBTENIENDO PARAMETRIZACIÓN DE LAS SEÑALES DE LOS ELECTROCARDIOGRAMAS

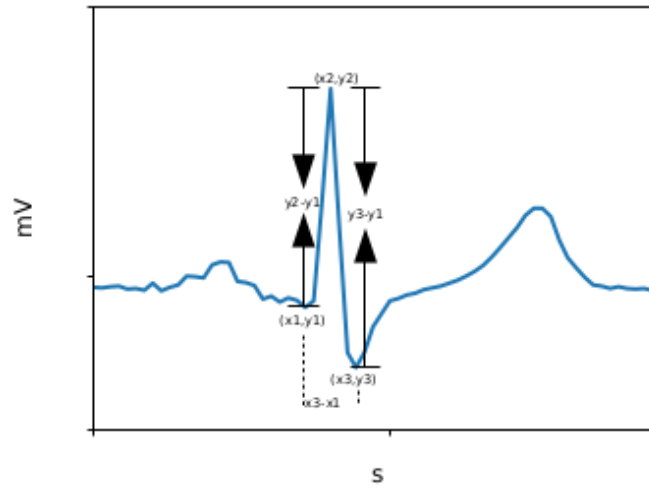


Figura 3.6: Amplitud y duración de la onda QRS.

A este método de análisis se le llamó **Método de parametrización** de las señales. Al obtener los valores de la amplitud (mv) y la duración (s) de las ondas P y T y grupo QRS, se estudio y analizo para poder encontrar valores que hagan describir personas con un indice de masa corporal normal, con sobrepeso y con obesidad.

Capítulo 4

Resultados

Considerando el Índice de Masa Corporal(IMC), como se muestra a continuación:

	IMC
Bajo peso	15 - 19.99
Peso normal	20 - 24.99
Sobrepeso	25 - 29.99
Obesidad I	30 - 34.99
Obesidad II	35 - 39.99
Obesidad III	40 -

Como no se contó con suficientes datos de personas con Bajo peso y Obesidad I,II y III, en este trabajo se considero el peso normal con $IMC < 25$, sobrepeso con $25 < IMC < 30$ y obesidad con $IMC > 30$.

El índice de Masa Corporal se calcula de la siguiente manera:

$$IMC = \frac{masa}{estatura^2} = \frac{(kg)}{(m^2)} \quad (4.1)$$

En necesario mencionar que a todas las graficas se calculo el valor de correlación (r) y el valor de P, para plantear prueba de hipótesis, en caso de ser posible.

4.1. Prueba de corridas por arriba y abajo de la media

En la señal del ECG, durante 10s, fue usado el método mencionado se encontró el número de repeticiones consecutivas (N_r) de 1's y 0's en cada L_i donde $i= 1, , 57$, es decir se calculó la frecuencia (f_{N_r}) de encontrar n veces consecutivas de ceros o unos. A continuación se grafican las listas **cero** _{i} y **uno** _{i} donde $i= 1, , 57$.

En la figuras 4.1,4.2 y 4.3 se muestra un ejemplo de los histogramas 1's y 0's de personas con un IMC normal, sobrepeso y obesidad respectivamente, de igual manera se grafico el promedio de los histogramas 1's y 0's de las personas con los diferentes IMC.

Se puede observar de los histogramas obtenidos que tienen una distribución *asimétrica positiva*, es decir que la "cola" a la derecha de la media es más larga que la de la izquierda, por lo tanto hay valores más separados de la media a la derecha. En todo los histogramas mostrados a continuación se calculó la Moda (m), el valor de la frecuencia donde se encuentre la moda de (f_m), la media (\bar{x}), la mediana (x), la varianza (S^2) y la desviación estándar (σ) tanto para el histograma de 1's y 0's, los datos se muestran en las tablas A.2 y A.3 del apéndice A.

CAPÍTULO 4. RESULTADOS

4.1. PRUEBA DE CORRIDAS POR ARRIBA Y ABAJO DE LA MEDIA

Figura 4.1: Ejemplo del histogramas 1's y 0's en una persona con un IMC normal y el promedio de los histogramas de todas las personas con un IMC normal

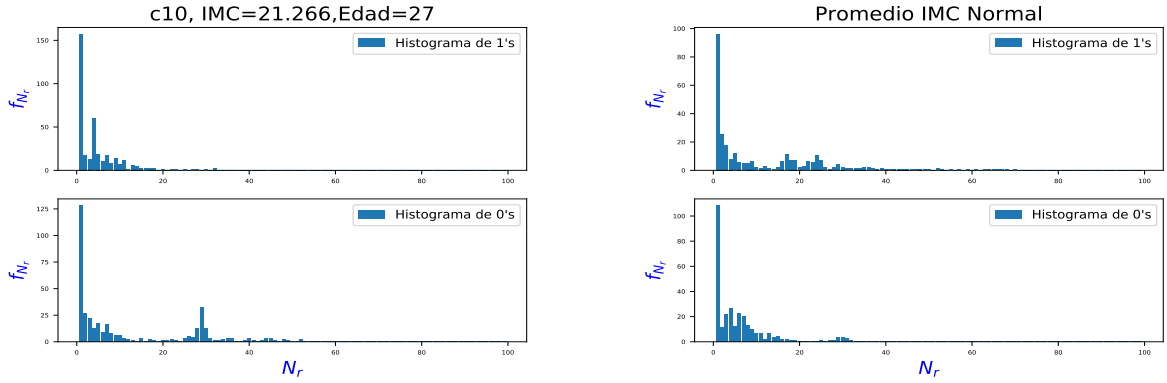


Figura 4.2: Ejemplo del histogramas 1's y 0's en una persona con sobrepeso y el promedio de los histogramas de todas las personas con sobrepeso

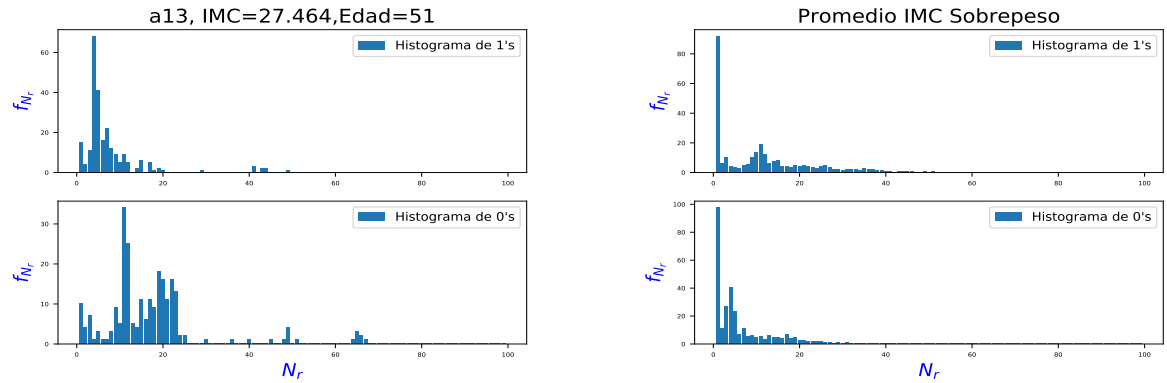
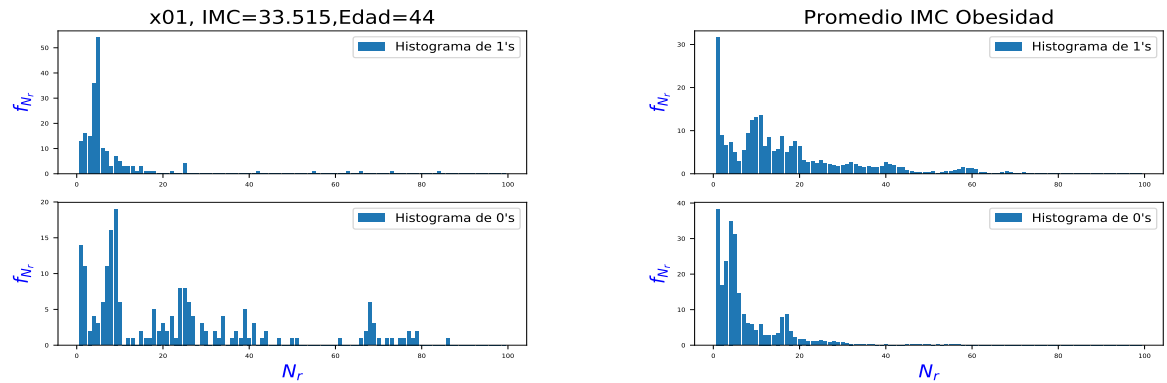


Figura 4.3: Ejemplo del histogramas 1's y 0's en una persona con obesidad y el promedio de los histogramas de todas las personas con obesidad



CAPÍTULO 4. RESULTADOS

4.1. PRUEBA DE CORRIDAS POR ARRIBA Y ABAJO DE LA MEDIA

Se calculó el promedio de los histogramas 1's y 0's del conjunto de personas con un IMC; normal, sobrepeso y obesidad. Y de los nuevos histogramas promedio se calculó la media y desviación estándar.

Tabla 4.1: Valores de la media (\bar{x}) y la desviación estándar (σ) de los histogramas 1's y 0's promedio.

Grafica	Normal	Sobrepeso	obesidad
Histograma de 1's	$(\bar{x})=3.088, \sigma=11.74$	$(\bar{x})=3.027, \sigma=11.06$	$(\bar{x})=2.44, \sigma=6.75$
Histograma de 0's	$(\bar{x})=3.09, \sigma=10.14$	$(\bar{x})=3.03, \sigma=9.55$	$(\bar{x})=2.44, \sigma=4.23$

Se observa de los histogramas que al aumentar el valor del índice de masa corporal la distribución aumenta hacia la derecha. Es decir para las personas con un índice de masa corporal normal la duración de las ondas QRS, P y T es casi constante, mientras que para personas que presentan sobrepeso u obesidad es posible observar que el valor de la duración es las ondas QRS, P y T van cambiando con el tiempo, es decir su duración no es estable.

Se encontró una correlación considerable entre la media (\bar{x}) de los histogramas de 0's y 1's y el índice de masa corporal, dado a que mientras aumenta el IMC de igual manera aumenta la media como se observa en la figura 4.4. De igual manera se graficó la desviación estándar (σ) y la media (\bar{x}), dando como resultado que las personas con un IMC-Normal se encuentran en una media menor a 50, y una σ entre 2.5 y 7.5, mientras que las personas con un IMC mayor de 25, los valores de la media pueden alcanzar hasta 400 con una σ hasta 20. De esta manera se pueden identificar a las personas sanas y enfermas.

Figura 4.4: Graficas de la media y el índice de masa corporal, de los datos obtenidos de los histograma de 1's y 0's

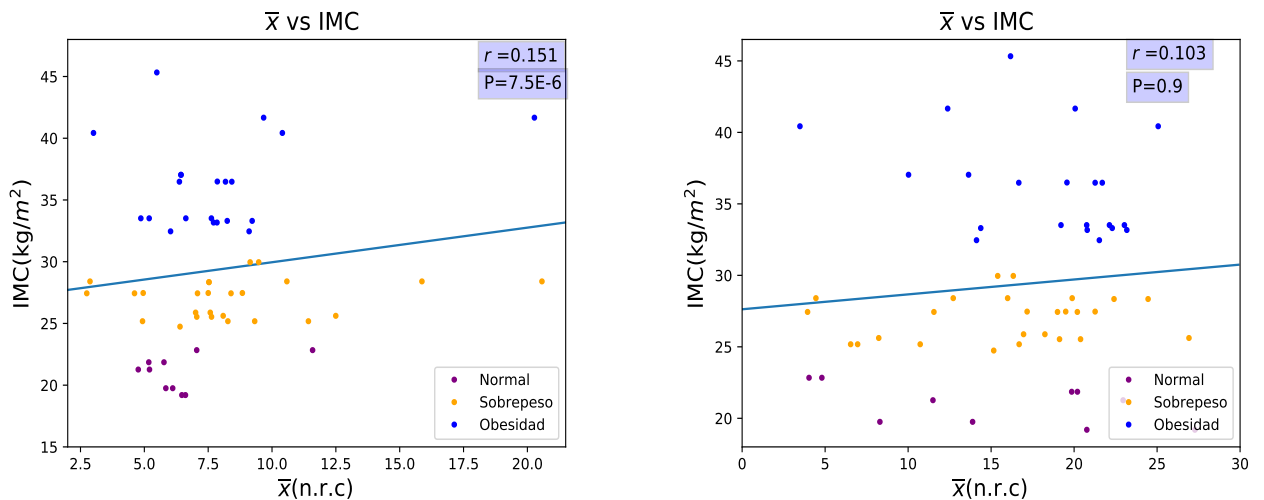
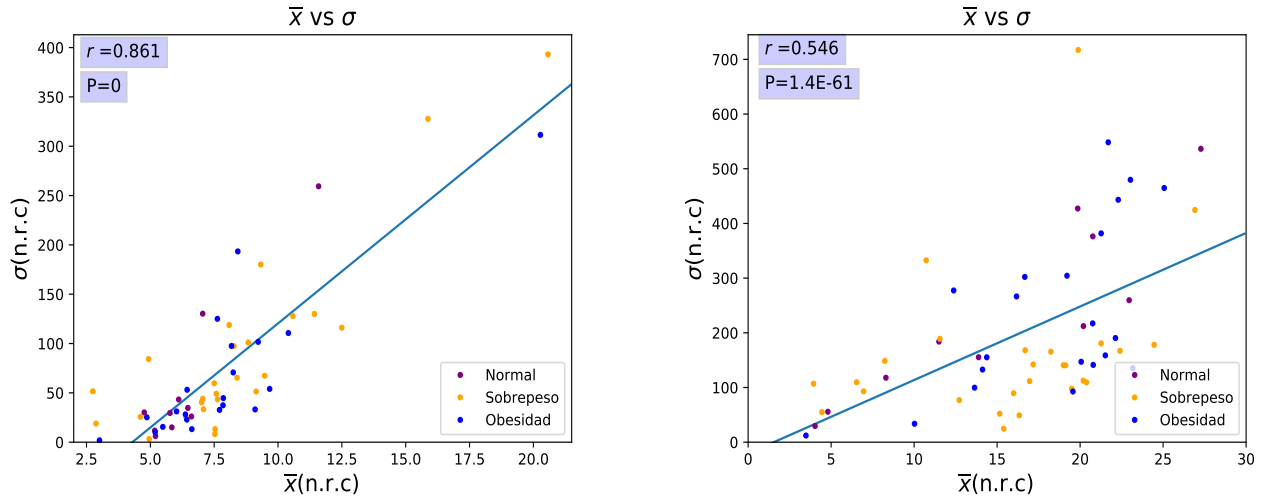


Figura 4.5: Graficas de la media y la desviación estándar, de los datos obtenidos de los histograma de 1's y 0's



Se puede concluir que una persona con un IMC normal, y una media $4 \leq \bar{x} \leq 7.5$ en su histograma de 1's, no presenta el síndrome de la Apnea Obstructora del sueño, por lo tanto se planteó la siguiente prueba de hipótesis:

$$H_o : \bar{x} = 7.5 (\leq)$$

$$H_a : \bar{x} > 7.5$$

Al rechazar o no rechazar la hipótesis se llegara a la conclusión de saber si el paciente es sano o enfermo.

4.2. Transformada Rápida de Fourier

Por medio del programa mostrado en el Apéndice C, se calculó la Transformada Rápida de Fourier de la señal del electrocardiograma de cada persona, durante 10 minutos.

Inicialmente se contó el número de periodos que hay en la señal durante 10 s, luego se calculó la frecuencia, al cual se llamó f_{aprox} . Se sabe que solo se graficó la mitad positiva de la serie dada por la FFT, así que se buscó el valor máximo de la magnitud en un intervalo de $0.066 < f_{aprox} < 0.058$ y el valor correspondiente a esa magnitud máxima se le nombra (f). Los valores de f_{aprox} y f se muestran en la figura 4.6, en ella se observa que la frecuencia se mantiene constante por un periodo de 10 minutos.

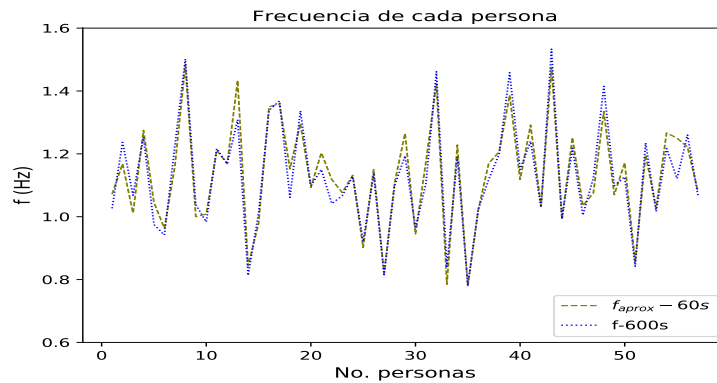
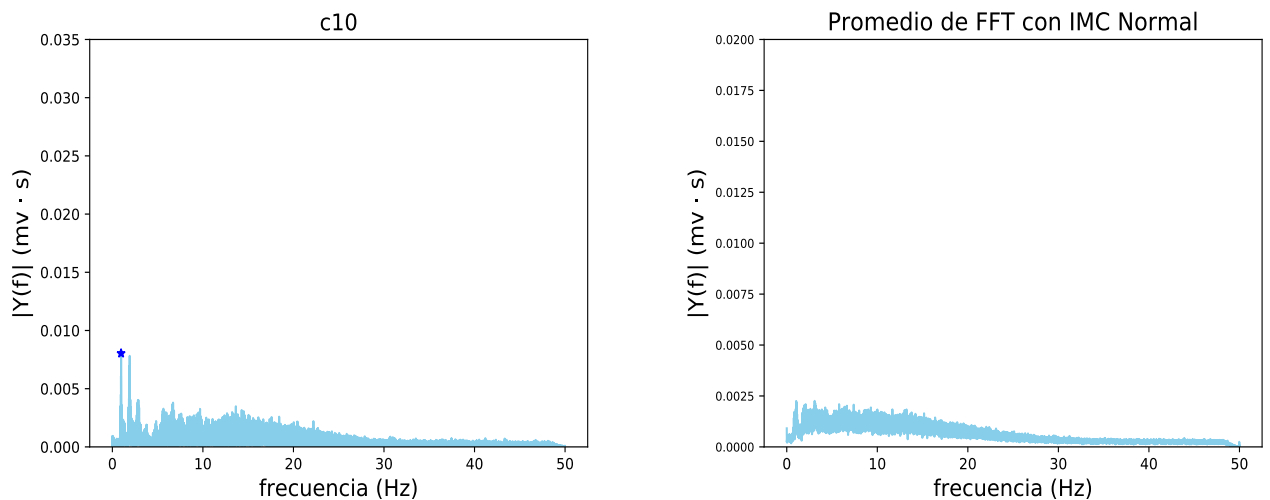


Figura 4.6: Frecuencia (f) encontrada en un intervalo cercano a f_{aprox} .

Utilizando este programa fue posible obtener la frecuencia (f), los pulsos por minuto (PPM) y el valor de la amplitud de la frecuencia (a_f) de cada persona, estos datos se muestran en la tabla A.4 del apéndice A. Utilizando el programa ya antes mencionado se graficó el espectro de las frecuencias, en la figura 4.7 se muestran las gráficas para personas con un índice de masa corporal normal, en la figura 4.8 para personas con Sobrepeso y para personas con obesidad en la figura 4.9.

Figura 4.7: Ejemplo de la FFT en una persona con IMC normal y el promedio de la FFT de todas las personas con IMC normal.



CAPÍTULO 4. RESULTADOS
4.2. TRANSFORMADA RÁPIDA DE FOURIER

Figura 4.8: Ejemplo de la FFT en una persona con sobrepeso y el promedio de la FFT de todas las personas con sobrepeso.

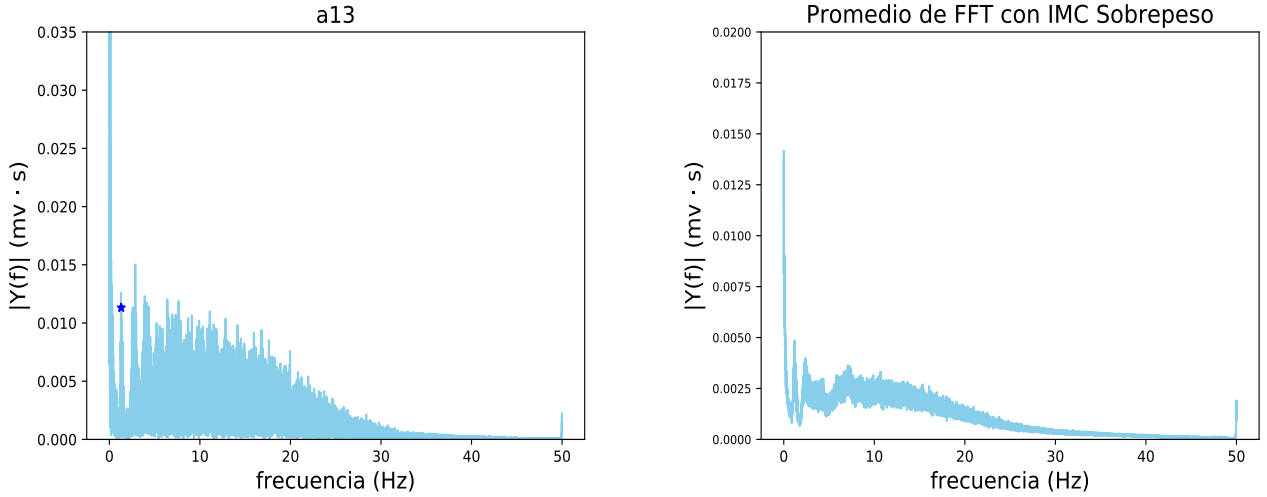
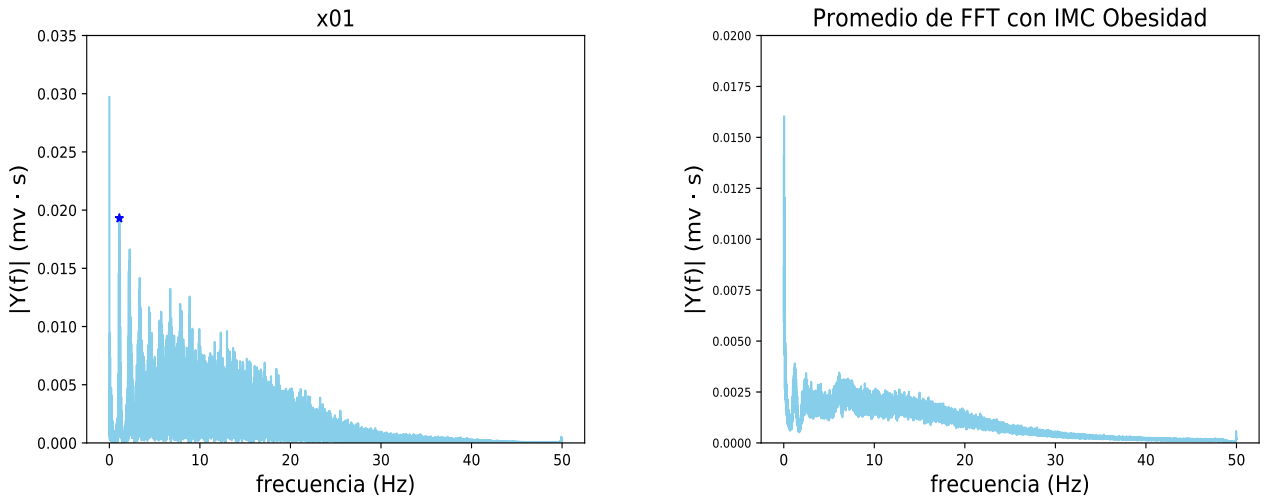


Figura 4.9: Ejemplo de la FFT en una persona con obesidad y el promedio de la FFT de todas las personas con obesidad.



En la figura 4.10 se puede observar que las personas con un IMC normal los valores de a_f van desde 0.005 hasta 0.010 aproximadamente y el valor de la frecuencia va desde 0.7 hz hasta 1.1 hz, para personas con sobrepeso y obesidad los valores de a_f van desde 0.0025 hasta 0.030 ($\text{mv} \cdot \text{s}$) mientras que los valores de la frecuencia van desde 0.95 hz hasta 1.6 hz. En la figura 4.11, las personas con un índice de masa corporal mayor a 25 los valores de a_f son muy variados, pero en el caso de las personas con un IMC normal los valores de a_f van de 0.0025 hasta 0.009 ($\text{mv} \cdot \text{s}$).

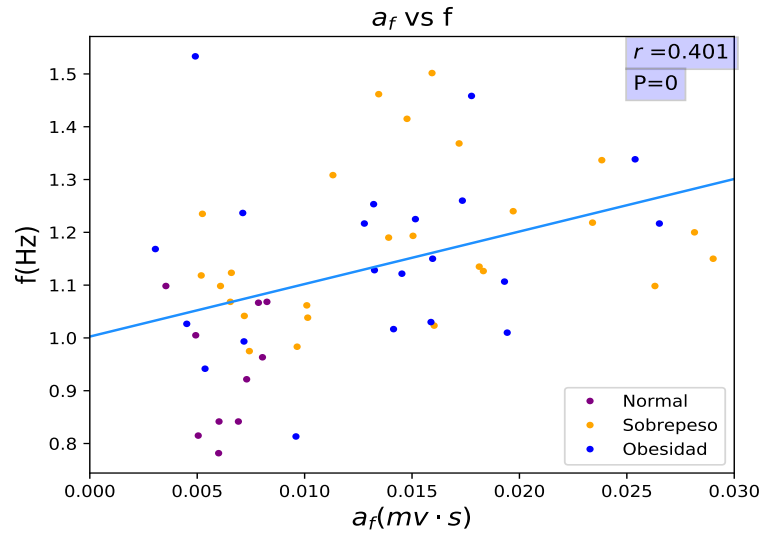


Figura 4.10: Grafica de la magnitud onde se encuentra la frecuencia de la señal (a_f) y la frecuencia (f).

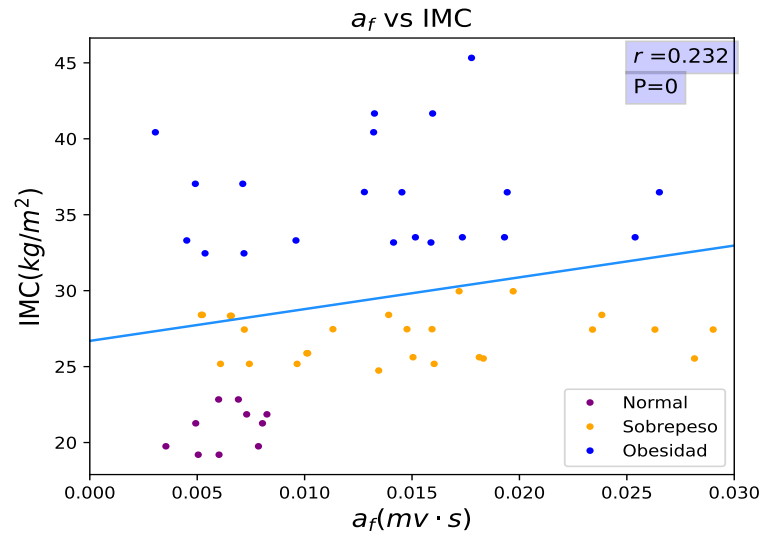


Figura 4.11: Grafica de la magnitud onde se encuentra la frecuencia de la señal (a_f) y el Indice de Masa Corporal.

Se puede concluir que una persona con un IMC normal, el valor de la amplitud en su frecuencia cardiaca se encontrará en un intervalo de $0.0025 \leq a_f \leq 0.009$, por lo tanto la paciente no presenta el SAOS, así que se planteó la siguiente prueba de hipótesis:

$$H_o : \bar{x} = 0.009 (\leq)$$

$$H_a : \bar{x} > 0.009$$

Al rechazar o no rechazar la hipótesis se llegara a la conclusión de saber si el paciente es sano o enfermo.

4.3. Parametrizaciones en las señales de los electrocardiogramas

Utilizando las diferencias presentadas en la tabla 3.1, se graficó la segunda columna y la tercera columna, también se graficó la segunda columna y la cuarta columna. El objetivo de graficar estas dos diferencias para poder observar si existe cambios notables en personas con un IMC normal, con sobrepeso u obesidad. A continuación se muestran los datos obtenidos de ambas diferencias en el grupo QRS, ondas P y T, se graficaron estos datos por edades, dando:

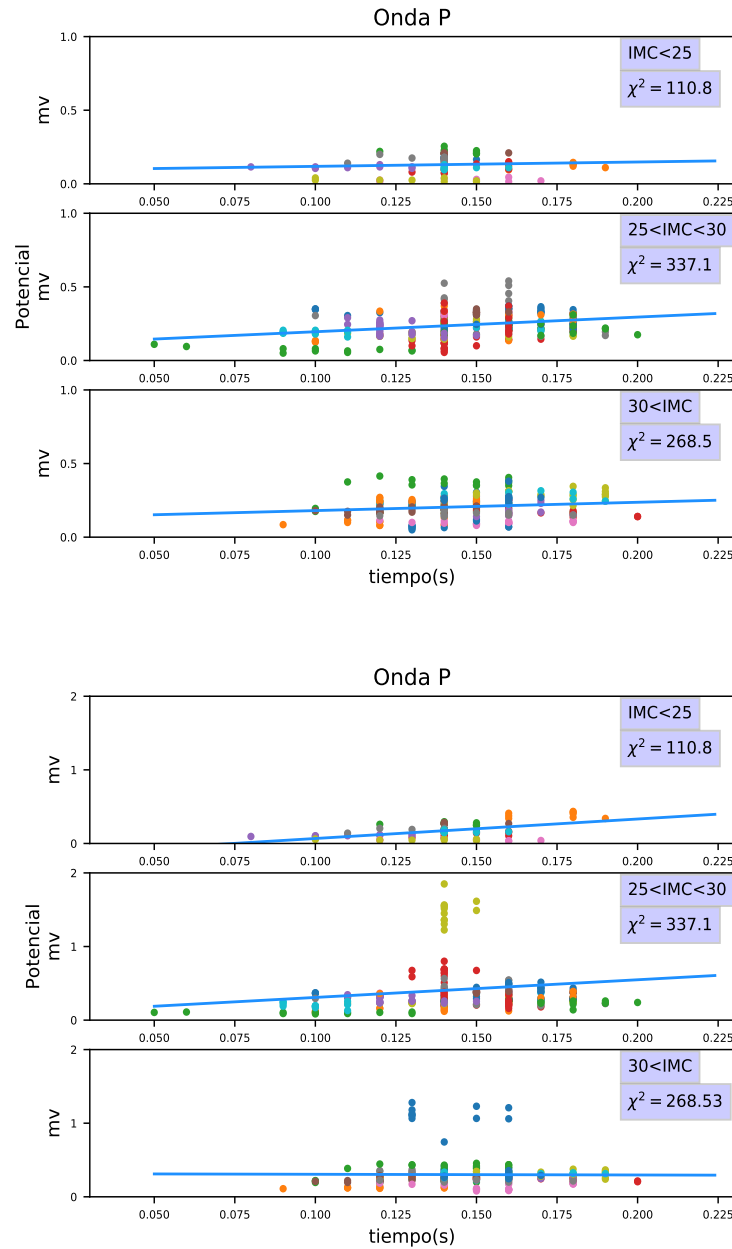


Figura 4.12: Onda P, utilizando la diferencia 1 y diferencia 2 respectivamente

4.3. PARAMETRIZACIONES EN LAS SEÑALES DE LOS ELECTROCARDIOGRAMAS

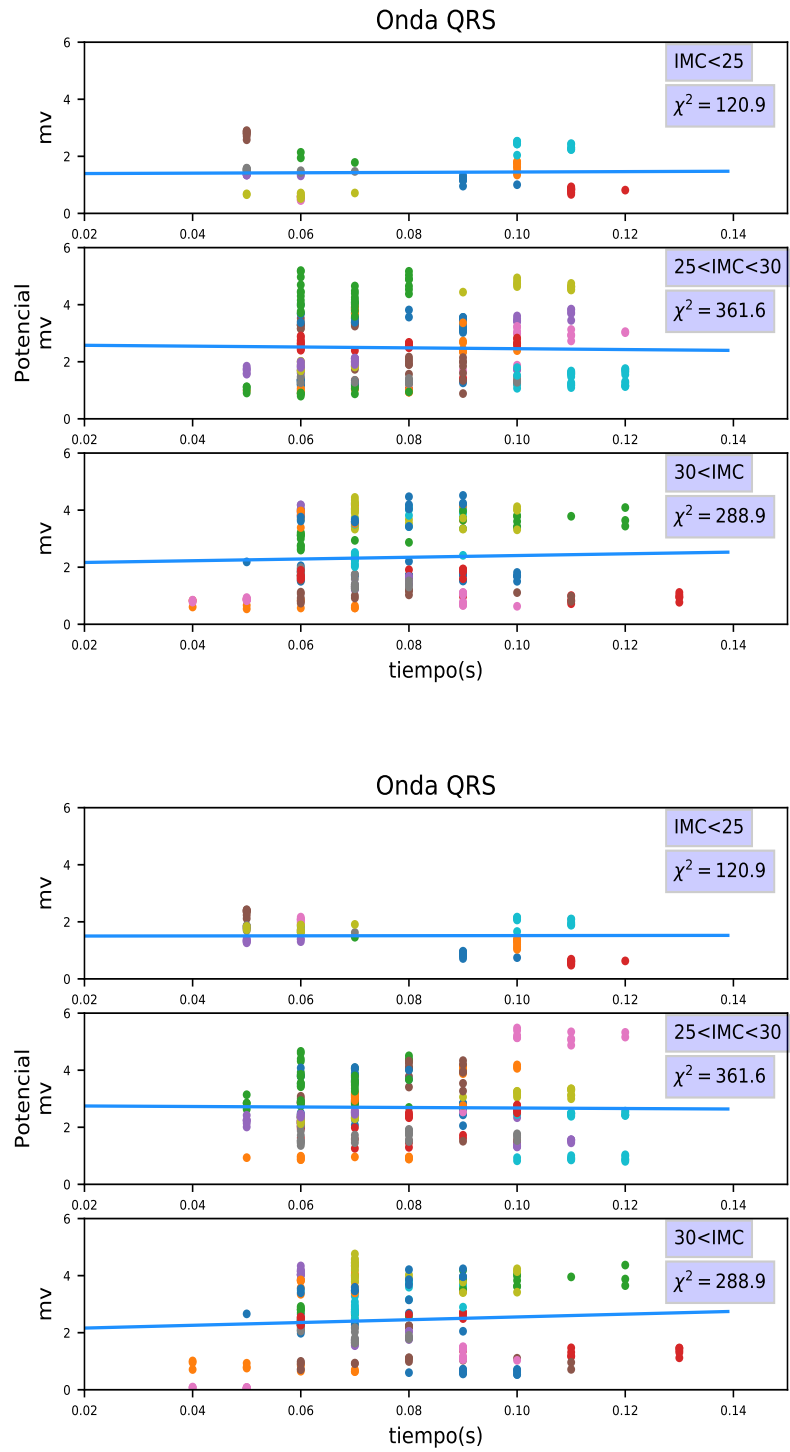


Figura 4.13: Onda QRS, utilizando la diferencia 1 y diferencia 2 respectivamente

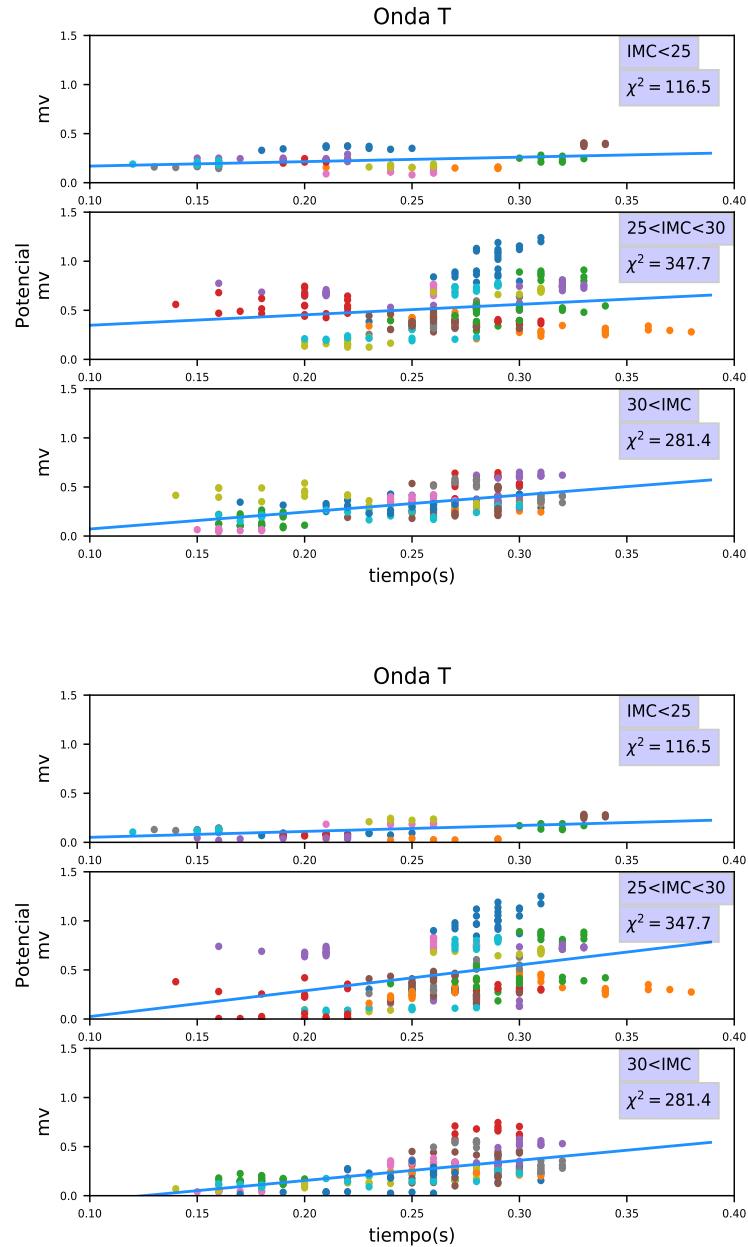


Figura 4.14: Onda T, utilizando la diferencia 1 y diferencia 2 respectivamente

Se puede observar de esta gráfica que la onda QRS se encuentra en un intervalo de tiempo de 0.05 s hasta 0.12 s, la onda P tiene una duración de 0.10 s hasta 0.2 s, y la onda T se encuentra alrededor de 0.15 s y 0.35 s. De igual manera se puede observar que para la onda QRS el valor de la amplitud puede alcanzar valores de 0.5 mV hasta 5.5 mV, mientras que en la onda P alcanza valores de 0.1 mV hasta 0.5 mV y la onda T puede tener valores de 0.1 mV hasta 1 mV aproximadamente.

Al analizar los valores de χ^2 en la onda P, se puede concluir que en pacientes con un IMC normal se obtendrá $\chi^2 \leq 110.8$, es decir la paciente no presenta el SAOS. Entonces se planteó la siguiente prueba

de hipótesis:

$$\begin{aligned}H_o &: \chi_2 = 116.5(\leq) \\H_a &: \chi_2 > 0.009\end{aligned}$$

Al rechazar o no rechazar la hipótesis se llegara a la conclusión de saber si el paciente esta enfermo o sano, si la hipótesis no es rechazada por lo tanto se concluye que el paciente es sano y tiene un IMC normal, en caso de que sea rechazada la hipótesis nula se podrá plantear otra prueba de hipótesis:

$$\begin{aligned}H_o &: \chi_2 = 281.4(\leq) \\H_a &: \chi_2 > 281.4\end{aligned}$$

ahora al rechazar o no rechazar la hipótesis se llegara a la conclusión de saber si el paciente tiene sobrepeso u obesidad.

4.4. Correlación de Datos

En las siguientes graficas que se muestran a continuación se utilizaron los datos mostrados en la tabla A.1.

En la figura 4.15 se graficó la edad contra el número de apneas por hora (AI), el número de Hipopneas por hora (HI) y también con AHI señalando con diferente color a las personas que tienen índice de masa corporal normal, con sobrepeso y obesidad, de igual manera se calculo la correlación utilizando la ecuación 2.9 y el valor de P.

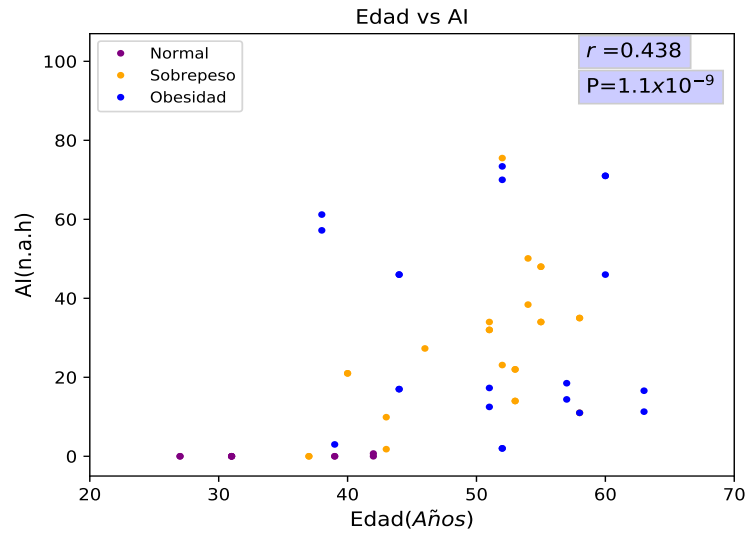


Figura 4.15: Datos de la edad (años) y el número de apneas por hora (n.a.m) de cada persona.

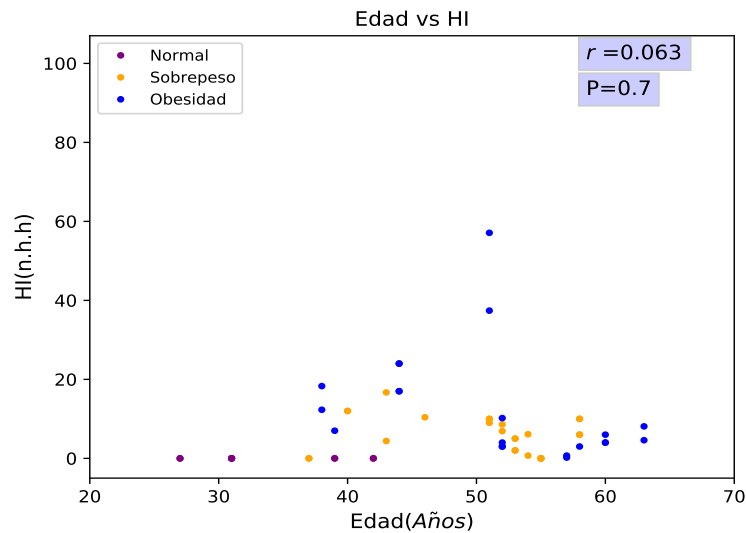


Figura 4.16: Datos de la edad y el número de hipopneas por hora (n.h.m) de cada persona.

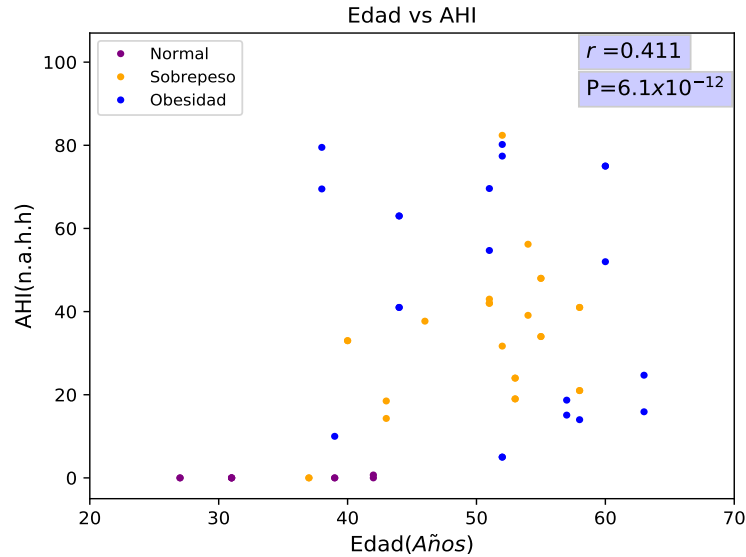


Figura 4.17: Datos de la edad y el número de la suma de la apnea e hipopnea por hora (n.a.m) de cada persona.

En el caso de los datos de la figura 4.15 y 4.17 el valor de P es menor a $\alpha = 0.05$ por lo tanto la hipótesis que se proponga será rechazada. Por lo tanto se llega a la conclusión; los pacientes menores de 45 años y valores de HI menor a 2 es un paciente sano, es así que se propuso la siguiente prueba de hipótesis,

$$H_o : P = 0.05(\geq)$$

$$H_a : P < 0.05$$

y puesto que el valor de $P=0.7$ entonces no se rechaza la hipótesis nula. De las figuras 4.15, 4.16 y 4.17 se pudo concluir que personas menores a 35 años no presentan valores de AI o HI, es decir son personas sanas. Por lo tanto se puede proponer la siguiente prueba de hipótesis:

$$H_o : E = 35(\leq)$$

$$H_a : E > 35$$

al tomar la decisión de rechazar la hipótesis se llegara a la conclusión de que el paciente es candidato a presentar el SAOS y al no rechazar la hipótesis se podrá decir que el paciente es sano.

En las próximas cuatro figuras se graficó la edad y los pulsos por minuto, considerando la clasificaciones de la tabla 2.1, se grafico por separado a las personas con indice de masa corporal normal, con sobrepeso y obesidad, figuras 4.18, 4.19 y 4.20 respectivamente. En la figura 4.19 y 4.20 se muestran con diferente color los valores que son $AI < HI$ (gris), $AI > HI$ (amarillo) y $AI = HI$ en morado.

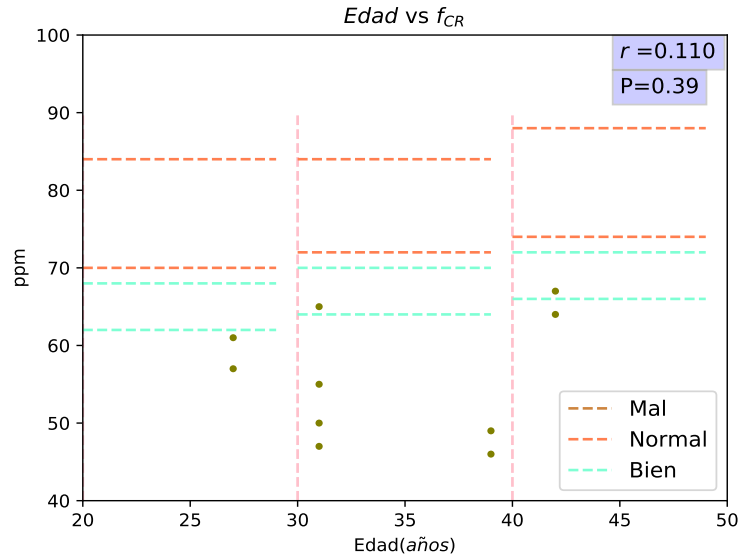


Figura 4.18: Representación de la edad y la frecuencia cardiaca en reposo de personas con un IMC normal.

En las siguientes figuras 4.19, 4.20 se puede notar que los valores de pulsos por minuto (ppm) en personas con sobrepeso y obesidad se encuentran por debajo de lo normal, es decir conforme aumenta la edad disminuye la frecuencia cardiaca, lo cual puede ser engañoso porque alguien puede creer que solo por tener su frecuencia cardiaca baja es una persona sana. En personas con sobrepeso se puede observar que el valor de AI es mayor de HI y en el caso de personas con obesidad se encuentran dos casos $AI < HI$.

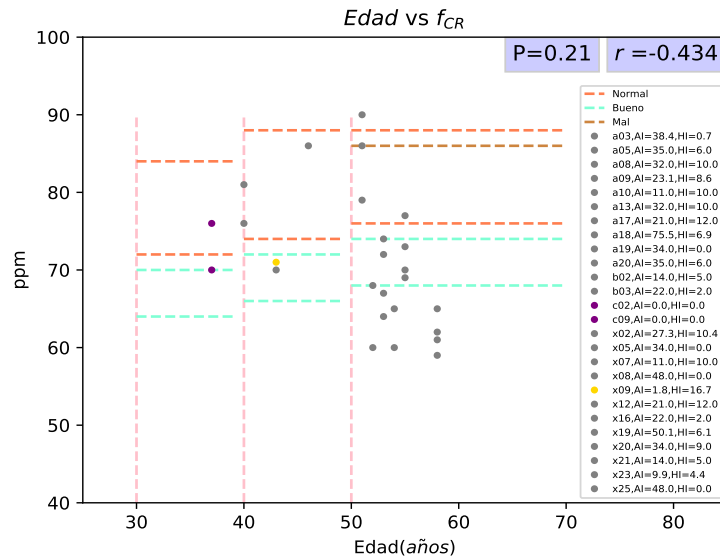


Figura 4.19: Representación de la edad y la frecuencia cardiaca en reposo de personas con sobrepeso.

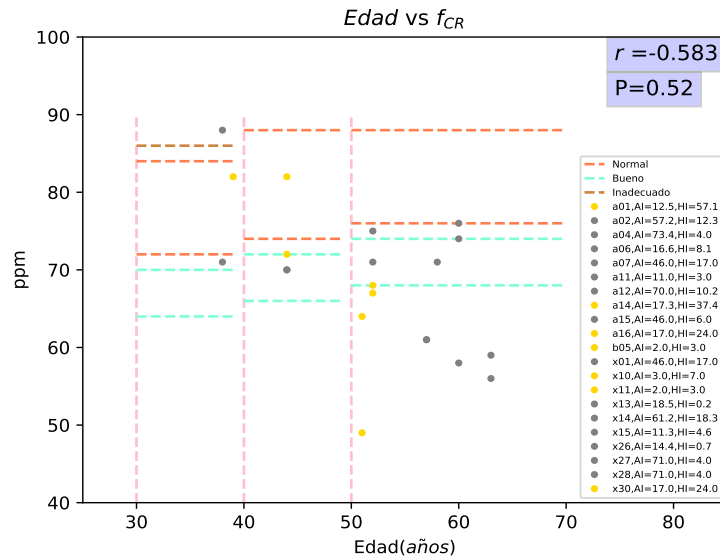


Figura 4.20: Representación de la edad y la frecuencia cardiaca en reposo de personas con obesidad.

En la figura 4.21, se grafico la Edad vs f_{CR} de las 57 personas, donde se colocó una recta la cual está descrita por $y(x) = m(x - x_1) + y_1$, donde $m = -1.4$ y $y_1 = 85$, los datos que se encuentran por debajo de esta recta se encuentra la mayoría de las personas con un IMC normal, mientras que por arriba de esta recta se encuentran los datos de personas con sobrepeso y obesidad, entonces se calculo la correlación de personas con sobrepeso y obesidad.

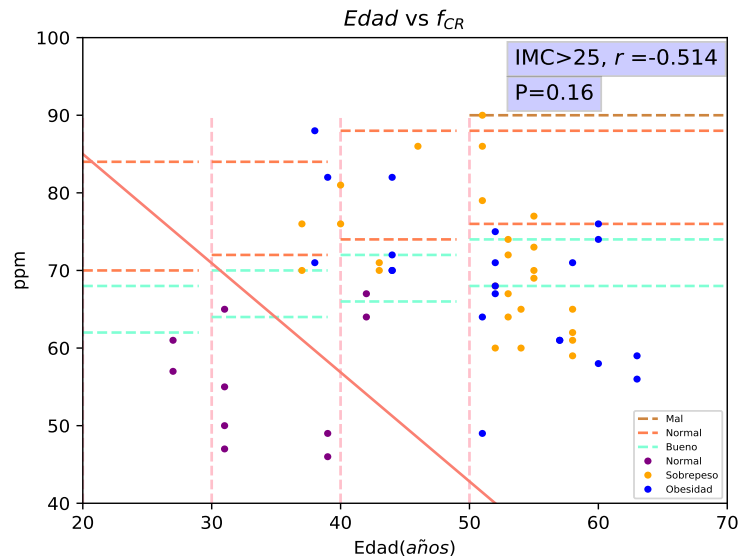


Figura 4.21: Representación de la edad y la frecuencia cardiaca en reposo.

Se llega a la conclusión de que valores por debajo de la recta se encontraran los valores de pacientes

sanos y por arriba los pacientes enfermos. Por lo tanto se propone la siguiente prueba de hipótesis,

$$H_o : y(e_p) = p(e_p, f_{CRp})(\leq)$$

$$H_a : y(e_p) > p(e_p, f_{CRp})$$

donde e_p es la edad del paciente y f_{CRp} . En caso de que el punto $p(edad, f_{CR})$ sea menor que $y(edad)$ la hipótesis nula no se rechaza por lo tanto el paciente esta sano. En caso de rechazar la hipótesis el paciente esta enfermo.

Posteriormente en la figura 4.23 se graficó la edad con el índice de masa corporal, haciendo notar con diferentes colores el IMC; normal, sobrepeso y obesidad, donde se calculo para cada caso el valor de r y P. Es posible notar que la personas después de los 35 años su IMC es mayor a 25 (kg/m^2).

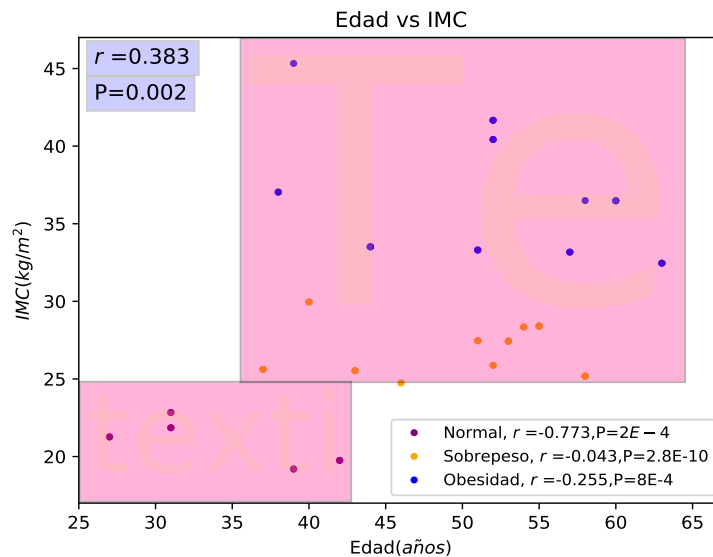


Figura 4.22: Datos de la edad y el IMC de los 57 hombres.

En la figura 4.23 se grafica f_{CR} y el índice de masa corporal haciendo notar con diferente color a las personas con un IMC; Normal, Sobrepeso y Obesidad, esto con el objetivo de encontrar alguna característica del comportamiento de f_{CR} . En la grafica 4.24 de igual manera se grafica PPM con el IMC, pero ahora con la diferencia de hacer notar con diferente color las edades, para hacer esta clasificación fueron usados los datos de la tabla 2.1, esto con el objetivo de encontrar algún comportamiento f_{CR} dependiendo de la edad.

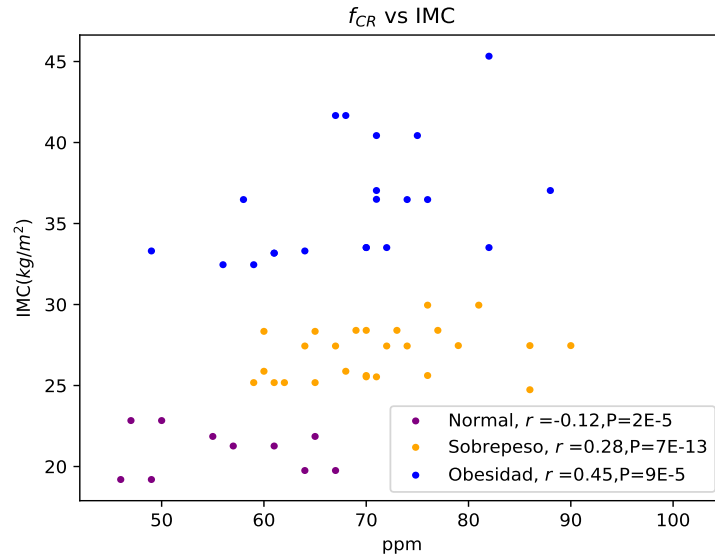


Figura 4.23: Datos de los pulsos por minuto y el índice de masa corporal.

De esta ultima figura se puede concluir que en personas con $IMC < 25$ los pulsos por minutos se encuentran en un rango de $40 < ppm < 70$, para personas con sobrepeso $60 < ppm < 90$ y en personas con sobrepeso $50 < ppm < 90$.

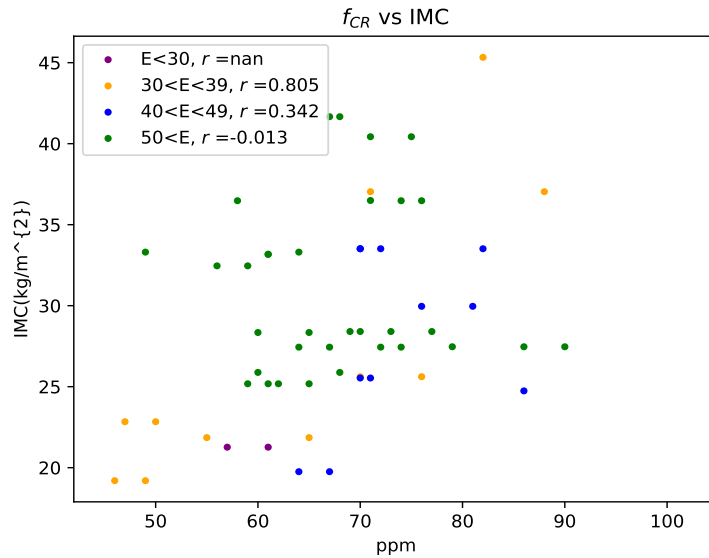


Figura 4.24: Diferencia de los datos del pulsos por minuto y el índice de masa corporal en diferentes edades.

Los valores de P para $30 < E < 39$ es de 0.09, $40 < E < 49$ es de 0.45 y $50 < E$ es de 0.005. De esta figura solo es posible concluir que las personas mayores de 50 años su IMC es mayor de 25.

En las graficas que se muestran a continuación, se graficó el IMC con el número de Apneas presentadas por hora (AI), el número de Hipopneas por hora (HI) y AHI(es decir, la suma de AI y HI):

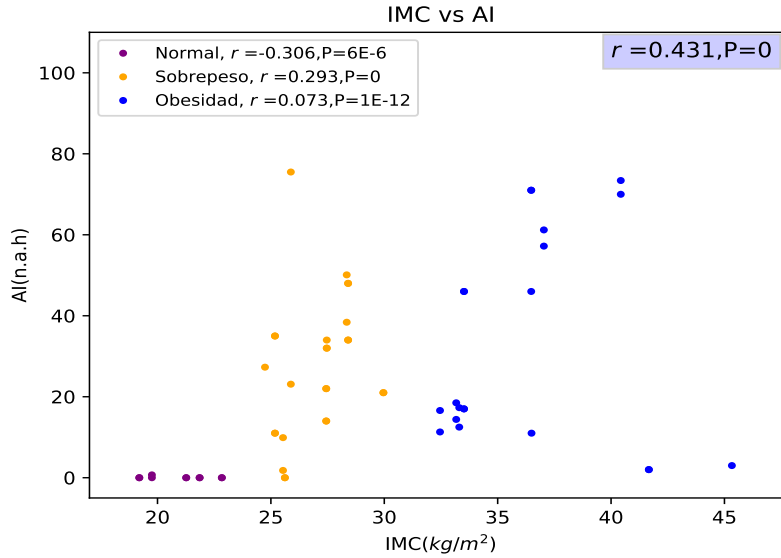


Figura 4.25: Datos del IMC y el número de apneas por hora (n.a.h).

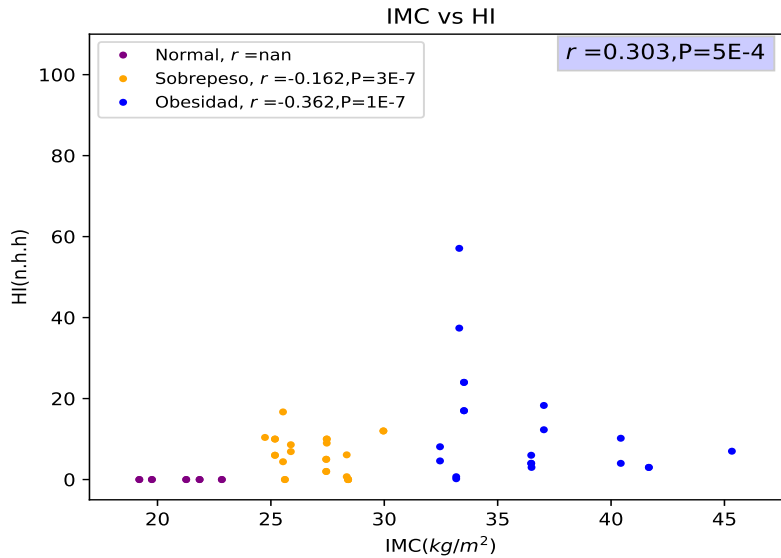


Figura 4.26: Datos del IMC y el número de hipopneas por hora(n.h.h).

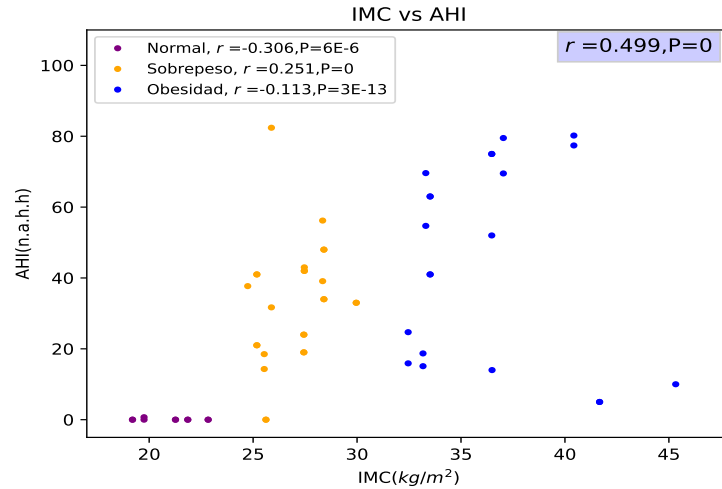


Figura 4.27: Datos del IMC y el número de la suma de apneas e hipopneas por hora (n.a.h).

De estas últimas tres gráficas, las personas con un IMC menor a 25 no presentan valores mayores a 5 de AI, HI y AHI lo cual nos permite concluir que las personas con un IMC menor a 25 es una **persona sana**. De igual manera se puede concluir que las personas con un IMC mayor a 25 son **personas enfermas**.

Las gráficas 4.28, 4.29 y 4.30 describen el comportamiento al graficar los pulsos por minuto (PPM) con AI, HI y AHI, mostrando con diferentes colores los posibles índices de masa corporal, esto con el objetivo de poder observar si existe algún comportamiento dependiendo del IMC.

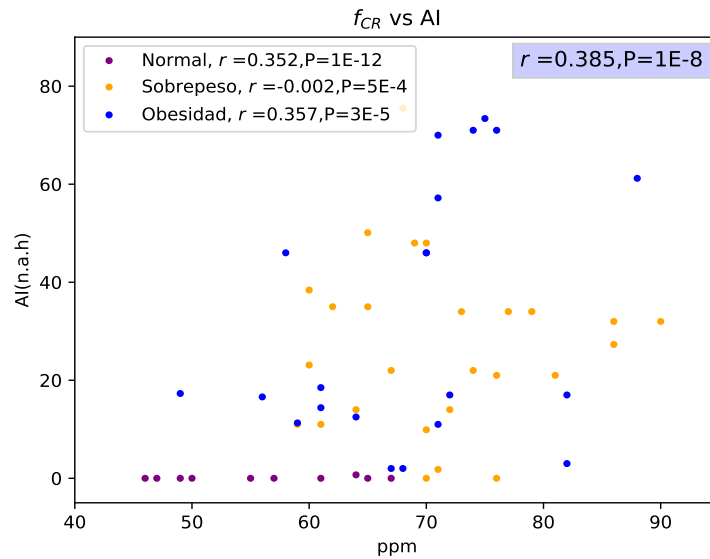


Figura 4.28: Datos del f_{CR} y el número de apneas por hora (n.a.h).

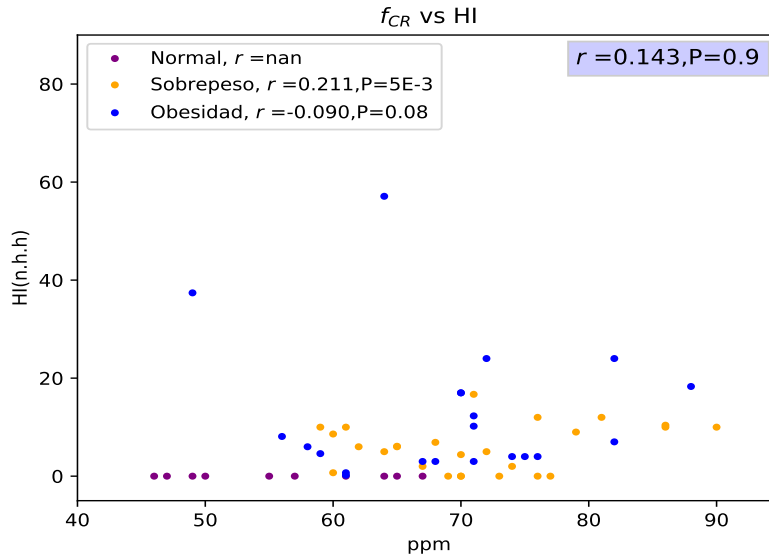


Figura 4.29: Datos del f_{CR} y el número de hipopneas por hora (n.h.h).

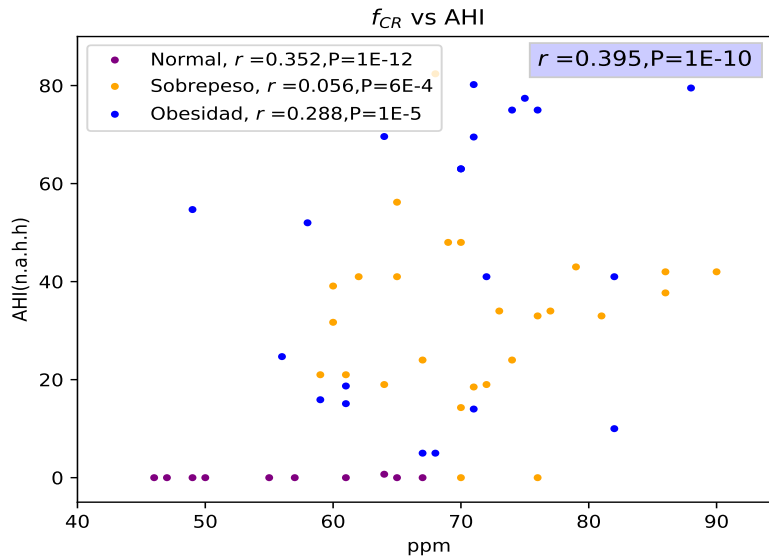


Figura 4.30: Datos del f_{CR} y el número de la suma de apneas e hipopneas por hora (n.a.h.h).

Estas graficas nos permiten entender que no una relación clara con el valor de los pulsos por minuto y los valores de la AI, HI Y AHI. Lo unico que se puede concluir es que a partir de 80 ppm el valor de AHI es mayores de 5.

Capítulo 5

Conclusión

Al aplicar los métodos propuestos en las señales de los electroencefalogramas, se encontraron diferentes variables y comportamientos que variaron entre pacientes con y sin SAOS y que también dependieron de su IMC.

Una persona con un IMC normal con valor de su media $4 \leq \bar{x} \leq 7.5$ en su histograma de 1's, no presenta el síndrome de la Apnea Obstructora del sueño, es decir es una persona sana. En el caso de que no se cumpla esta condición se está hablando de una persona con el SAOS.

Si el valor de una persona del IMC es menor a $25(kg/m^2)$ y el valor de la amplitud en su frecuencia cardiaca se encontrará en un intervalo de $0.0025 \leq a_f \leq 0.009$, por lo tanto es sano y en caso contrario es una persona enferma.

La onda QRS se encuentra en un intervalo de tiempo de 0.05 s hasta 0.12 s. La onda P tiene una duración de 0.10 s hasta 0.2 s. La onda T se encuentra alrededor de 0.15 s y 0.35 s.

Valores de $\chi^2 \leq 110.8$ en la onda P indican que el paciente tiene un IMC normal y no presenta SAOS. En caso no cumplir con esta condición se concluye que el paciente está enfermo. Si $\chi^2 \leq 281.4$ entonces el paciente tiene sobrepeso y de ser $\chi^2 > 281.4$ el paciente tiene sobrepeso.

Se pudo observar en los datos que las personas con una edad menor a 35 años no se presentaron valores de AI y HI, por lo tanto son sanos. Se observó de las figuras 4.19 y 4.20 que el valor de los pulsos por minuto se encuentran en la región de lo normal y bueno, véase la tabla 2.1, por lo tanto no es suficiente saber el valor de la frecuencia cardiaca en reposo f_{CR} para saber si una persona es sana o no.

Dada la $y(x)$ en el capítulo 5.4 es posible decir que una persona es sana si $y(e_p) \leq p(e_p, f_{CRP})$, donde e_p es la edad del paciente y f_{CRP} la frecuencia cardiaca en reposo del paciente.

Por lo tanto se puede concluir que las personas con sobrepeso u obesidad son muy propensas a presentar SAOS y que al presentar esta enfermedad la f_{CR} no se encuentran dentro de lo normal. Esto quiere decir que el tiempo y la amplitud de duración de las ondas P, T y QRS son afectadas.

Al aplicar el método de prueba de corridas por arriba y abajo fue posible comprender el comportamiento de las ondas del electrocardiograma por medio de histogramas en personas sanas y enfermas. Al aplicar la transformada discreta de Fourier permitió encontrar la frecuencia de cada persona y comprender el comportamiento de a_f en personas sanas y enfermas. Y al calcular la parametrización de las

señales fue posible observar la variación del tiempo y la amplitud que existe en las ondas del electrocardiograma.

La aplicación de los métodos propuestos permitió analizar y estudiar los datos de las señales del ECG ampliamente, lo cual nos permitió encontrar comportamientos diferentes en personas sanas y enfermas.

Apéndice A

Tablas

En la pagina <https://physionet.org/cgi-bin/atm/ATM>, se obtuvo los datos de ECG de 70 personas, 57 hombres y 13 mujeres. Como la población de sexo masculino fue mayor que del femenino, por lo tanto solo se utilizó los datos de hombres, donde hay 12 personas que no presentan valores de AI, HI y AHI, y todas las demás personas si presentan valores.

Tabla A.1: Datos del número de apneas por hora(AI), número de hipopneas (HI), suma de AI e HI (AHI), edad, sexo, altura, peso e índice de masa corporal, de cada persona.

	AI	HI	AHI	Edad	sexo	Altura(cm)	Peso(kg)	IMC
a01	12.5	57.1	69.6	51	M	175	102	33.306
a02	57.2	12.3	69.5	38	M	180	120	37.037
a03	38.4	0.7	39.1	54	M	168	80	28.344
a04	73.4	4	77.4	52	M	173	121	40.429
a05	35	6	41	58	M	176	78	25.180
a06	16.6	8.1	24.7	63	M	179	104	32.458
a07	46	17	63	44	M	177	105	33.515
a08	32	10	42	51	M	179	88	27.464
a09	23.1	8.6	31.7	52	M	178	82	25.880
a10	11	10	21	58	M	176	78	25.180
a11	11	3	14	58	M	168	103	36.493
a12	70	10.2	80.2	52	M	173	121	40.429
a13	32	10	42	51	M	179	88	27.464
a14	17.3	37.4	54.7	51	M	175	102	33.306
a15	46	6	52	60	M	176	113	36.479
a16	17	24	41	44	M	177	105	33.515
a17	21	12	33	40	M	179	96	29.961
a18	75.5	6.9	82.4	52	M	178	82	25.880
a19	34	0	34	55	M	178	90	28.405
a20	35	6	41	58	M	176	78	25.180
b02	14	5	19	53	M	176	85	27.440
b03	22	2	24	53	M	176	85	27.440
b04	0.7	0	0.7	42	M	180	64	19.753
b05	2	3	5	52	M	180	135	41.666
c01	0	0	0	31	M	184	74	21.857

APÉNDICE A. TABLAS

	AI	HI	AHI	Edad	sexo	Altura(cm)	Peso(kg)	IMC
c02	0	0	0	37	M	180	83	25.617
c03	0	0	0	39	M	184	65	19.198
c08	0	0	0	42	M	180	64	19.753
c09	0	0	0	37	M	180	83	25.617
c10	0	0	0	27	M	184	72	21.266
x01	46	17	63	44	M	177	105	33.515
x02	27.3	10.4	37.7	46	M	167	69	24.740
x04	0	0	0	39	M	184	65	19.198
x05	34	0	34	55	M	178	90	28.405
x06	0	0	0	31	M	170	66	22.837
x07	11	10	21	58	M	176	78	25.180
x08	48	0	48	55	M	178	90	28.405
x09	1.8	16.7	18.5	43	M	177	80	25.535
x10	3	7	10	39	M	170	131	45.328
x11	2	3	5	52	M	180	135	41.666
x12	21	12	33	40	M	179	96	29.961
x13	18.5	0.2	18.7	57	M	171	97	33.172
x14	61.2	18.3	79.5	38	M	180	120	37.037
x15	11.3	4.6	15.9	63	M	179	104	32.458
x16	22	2	24	53	M	176	85	27.440
x18	0	0	0	27	M	184	72	21.266
x19	50.1	6.1	56.2	54	M	168	80	28.344
x20	34	9	43	51	M	179	88	27.464
x21	14	5	19	53	M	176	85	27.440
x23	9.9	4.4	14.3	43	M	177	80	25.535
x24	0	0	0	31	M	170	66	22.837
x25	48	0	48	55	M	178	90	28.405
x26	14.4	0.7	15.1	57	M	171	97	33.172
x27	71	4	75	60	M	176	113	36.479
x28	71	4	75	60	M	176	113	36.479
x30	17	24	41	44	M	177	105	33.515
x35	0	0	0	31	M	184	74	21.857

APÉNDICE A. TABLAS

Tabla A.2: Utilizando los valores de los histogramas 1's se calculo la moda (m), el valor de la frecuencia donde se encuentre la moda de (f_m), la media (\bar{x}), la mediana, la varianza (S^2) y la desviación estándar (σ).

	m	f_m	\bar{x}	x	S^2	σ		m	f_m	\bar{x}	x	S^2	σ
a01	1	51	8.246	4.5	70.708	8.409	c10	1	157	4.762	3.0	30.160	5.492
a02	1	85	6.440	2.0	53.119	7.288	x01	5	54	7.626	5.0	125.044	11.182
a03	5	53	7.527	7.0	8.206	2.865	x02	3	128	6.397	3.0	25.734	5.073
a04	4	68	3.009	3.0	1.878	1.370	x04	4	37	6.610	5.0	26.046	5.104
a05	4	54	11.425	6.5	130.018	11.403	x05	4	42	15.871	9.0	327.672	18.102
a06	4	35	9.103	10.0	33.200	5.762	x06	1	120	11.589	3.0	259.388	16.106
a07	5	58	6.625	5.0	13.234	3.638	x07	1	348	4.929	1.0	84.285	9.181
a08	4	81	4.955	5.0	3.227	1.796	x08	1	607	2.872	1.0	18.866	4.344
a09	4	65	7.582	4.0	49.079	7.006	x09	4	64	7.631	4.0	43.336	6.583
a10	1	167	9.322	3.0	180.066	13.419	x10	5	48	5.489	5.0	15.532	3.941
a11	3	72	7.858	4.0	44.700	6.686	x11	4	68	9.672	6.0	54.041	7.351
a12	1	163	10.405	7.0	110.653	10.519	x12	4	54	9.478	5.0	67.344	8.206
a13	4	68	7.502	5.0	59.731	7.729	x13	4	56	7.845	4.0	37.416	6.117
a14	4	51	9.222	4.0	101.675	10.083	x14	1	78	6.429	5.0	22.915	4.787
a15	5	46	8.426	5.0	193.361	13.905	x15	1	79	6.027	4.0	31.191	5.585
a16	5	79	5.190	5.0	10.206	3.195	x16	4	53	7.082	4.0	33.354	5.775
a17	4	55	9.143	5.0	51.352	7.166	x18	4	60	5.202	5.0	6.161	2.482
a18	4	86	7.008	4.0	40.416	6.357	x19	5	55	7.535	7.0	13.299	3.647
a19	10	14	20.571	11.0	393.102	19.827	x20	5	50	8.839	5.0	101.100	10.055
a20	1	100	8.272	4.0	97.386	9.868	x21	1	747	2.746	1.0	51.633	7.186
b02	3	67	8.394	3.0	65.358	8.084	x23	4	65	7.053	4.0	44.015	6.634
b03	1	174	4.615	3.0	25.687	5.068	x24	1	268	7.051	1.0	130.282	11.414
b04	1	200	6.111	3.0	43.271	6.578	x25	4	41	10.582	8.0	127.861	11.308
b05	1	13	20.275	13.0	311.518	17.650	x26	3	58	7.710	5.0	32.749	5.723
c01	1	87	5.769	5.0	29.579	5.439	x27	6	52	6.373	6.0	28.206	5.311
c02	3	62	12.497	3.0	116.091	10.775	x28	5	62	8.178	5.0	97.551	9.877
c03	1	55	6.470	5.0	34.742	5.894	x30	5	79	4.863	5.0	25.215	5.021
c08	1	99	5.842	6.0	15.051	3.880	x35	7	65	5.174	6.0	11.718	3.423
c09	1	111	8.087	4.0	118.821	10.900							

APÉNDICE A. TABLAS

Tabla A.3: Utilizando los valores de los histogramas O 's se calculo la moda (m), el valor de la frecuencia donde se encuentre la moda de (f_m), la media (\bar{x}), la mediana, la varianza (S^2) y la desviación estándar (σ).

	m	f_m	\bar{x}	x	S^2	σ		m	f_m	\bar{x}	x	S^2	σ
a01	8	48	14.377	9.0	155.254	12.460	c10	1	127	11.496	4.0	183.887	13.560
a02	1	50	13.642	12.0	99.688	9.984	x01	9	19	23.031	17.0	479.753	21.903
a03	10	40	24.457	26.0	178.044	13.343	x02	10	64	15.158	14.0	52.148	7.221
a04	11	62	25.066	12.0	464.775	21.559	x04	17	26	27.277	19.0	536.471	23.162
a05	10	23	16.689	12.0	168.205	12.969	x05	7	24	12.718	9.0	76.978	8.774
a06	20	37	21.520	20.0	158.750	12.600	x06	1	196	4.031	1.0	29.546	5.436
a07	8	31	22.130	20.0	190.430	13.800	x07	1	328	6.538	1.0	109.658	10.472
a08	11	41	19.494	19.0	97.711	9.885	x08	1	614	4.445	1.0	55.023	7.418
a09	20	31	18.241	19.0	165.485	12.864	x09	12	50	20.388	15.0	109.368	10.458
a10	1	185	6.959	1.0	92.996	9.643	x10	1	32	16.170	8.0	266.511	16.325
a11	16	48	19.569	16.0	92.631	9.624	x11	10	40	20.064	14.0	147.080	12.128
a12	1	202	3.483	2.0	12.157	3.487	x12	13	44	16.336	14.0	49.094	7.007
a13	11	34	17.169	15.0	142.206	11.925	x13	19	46	23.176	19.0	135.399	11.636
a14	9	30	22.296	17.0	443.246	21.053	x14	9	54	10.025	10.0	33.701	5.805
a15	1	22	21.697	13.0	548.343	23.417	x15	1	62	14.121	12.0	132.844	11.526
a16	9	29	20.758	22.0	217.205	14.738	x16	11	37	20.195	16.0	112.439	10.604
a17	14	39	15.398	14.0	24.617	4.962	x18	9	48	22.953	23.0	259.632	16.113
a18	11	41	16.956	17.0	111.610	10.565	x19	11	30	22.405	25.0	167.031	12.924
a19	3	37	19.884	4.0	717.177	26.780	x20	21	18	21.266	19.0	180.708	13.443
a20	1	83	10.728	3.0	332.511	18.235	x21	1	740	3.943	1.0	106.955	10.342
b02	11	77	19.000	11.0	140.731	11.863	x23	12	52	19.122	14.0	140.527	11.854
b03	1	160	11.555	9.0	189.002	13.748	x24	1	256	4.804	1.0	55.651	7.460
b04	1	190	8.300	2.0	117.816	10.854	x25	8	23	15.996	17.0	89.456	9.458
b05	4	42	12.383	4.0	277.427	16.656	x26	19	39	20.790	18.0	141.328	11.888
c01	1	74	19.855	18.0	427.269	20.670	x27	10	33	21.267	11.0	381.846	19.541
c02	9	25	26.921	11.0	424.563	20.605	x28	1	35	16.668	11.0	302.205	17.384
c03	1	35	20.767	17.0	376.279	19.398	x30	1	40	19.210	11.0	304.367	17.446
c08	24	57	13.885	7.0	155.332	12.463	x35	17	31	20.200	18.0	212.262	14.569
c09	1	139	8.231	2.0	148.672	12.193							

APÉNDICE A. TABLAS

Tabla A.4: Al calcular la Transformada Rápida de Fourier se encontró la frecuencia (f), los pulsos por minuto (PPM) y el valor de la amplitud de la frecuencia de cada persona, la cual llamaremos como $|Y|_f$, de cada persona.

Nombre	a_f	PPM	frecuencia(f)	Nombre	a_f	PPM	frecuencia(f)
a01	0.004513	616	1.026667	c10	0.008038	578	0.963333
a02	0.007121	742	1.236667	x01	0.019305	664	1.106667
a03	0.006542	641	1.068333	x02	0.013448	877	1.461667
a04	0.013213	752	1.253333	x04	0.006016	505	0.841667
a05	0.007428	585	0.975000	x05	0.013912	714	1.190000
a06	0.005365	565	0.941667	x06	0.005997	469	0.781667
a07	0.015157	735	1.225000	x07	0.016030	614	1.023333
a08	0.015931	901	1.501667	x08	0.005188	671	1.118333
a09	0.010144	623	1.038333	x09	0.028149	720	1.200000
a10	0.009650	590	0.983333	x10	0.017770	875	1.458333
a11	0.012781	730	1.216667	x11	0.015959	690	1.150000
a12	0.003052	701	1.168333	x12	0.019709	744	1.240000
a13	0.011318	785	1.308333	x13	0.015882	618	1.030000
a14	0.009597	488	0.813333	x14	0.004912	920	1.533333
a15	0.019432	606	1.010000	x15	0.007177	596	0.993333
a16	0.025385	803	1.338333	x16	0.023403	731	1.218333
a17	0.017191	821	1.368333	x18	0.004931	603	1.005000
a18	0.010101	637	1.061667	x19	0.006590	674	1.123333
a19	0.023835	802	1.336667	x20	0.014767	849	1.415000
a20	0.006084	659	1.098333	x21	0.026308	659	1.098333
b02	0.029019	690	1.150000	x23	0.018320	676	1.126667
b03	0.007194	625	1.041667	x24	0.006917	505	0.841667
b04	0.007851	640	1.066667	x25	0.005242	741	1.235000
b05	0.013251	677	1.128333	x26	0.014140	610	1.016667
c01	0.007302	553	0.921667	x27	0.026515	730	1.216667
c02	0.018129	681	1.135000	x28	0.014529	673	1.121667
c03	0.005048	489	0.815000	x30	0.017344	756	1.260000
c08	0.003542	659	1.098333	x35	0.008246	641	1.068333
c09	0.015046	716	1.193333				

Apéndice B

Prueba de corridas por arriba y abajo de la media

Se hizo en python el siguiente programa a continuación para guardar solo cierta cantidad de datos en unos nuevos archivos, esos nuevos archivos fueron utilizados para analizar y estudiar cada método deseado.

```
import numpy as np
import pylab as plt
import os
import pandas as pd
from scipy import signal

w1=0
w2=57

os.mkdir("datos")
direccion="/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/archivos/"
archivos=os.listdir(direccion) # se obtiene la lista de nombres en la direccion dada
direccion2="/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/0-1/datos/"

n=6000
for i in archivos:
    archivo=open(direccion+'{b}'.format(b=i), 'r')
    archivo1=open(direccion2+'datos%','%','w')
    k=-1
    for l in archivo:
        k +=1
        if k < n:
            archivo1.writelines(str(l))
    archivo1.close()
    archivo.close()
```

Utilizando la prueba descrita anteriormente se calculó la frecuencia de ceros y unos en cada base de datos, utilizando el programa que a continuación se describe:

```
import random
import numpy as np
import pandas as pd
import pylab as plt
import os
import numpy as num
from decimal import Decimal
from scipy import stats # importando scipy.stats

direccion="/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/0-1/datos/"
archivos=os.listdir(direccion) # se obtiene la lista de nombres en la direccion dada
archivos.sort()
```

APÉNDICE B. PRUEBA DE CORRIDAS POR ARRIBA Y ABAJO DE LA MEDIA

```
#crea un array con cada fila correspondiente a cada archivo de la carpeta datos

w1=0
w2=len(archivos)
y=[]
for k in range(w2):
    y.append(np.loadtxt(direccion+archivos[k], delimiter='\\t', skiprows=0, usecols=[0]))

y = np.array(y)

n=len(y[0])
#-----

x = {i:np.arange(0, int(n*0.010), 0.010) for i in range(w1, w2)}

#-----
#Minimos cuadrados para encontrar el promedio
Lista0=[]
Lista1=[]
A=[]
B=[]
fcero=[]
funo=[]

for i in range(w1, w2):
    x1=x[i]
    y1=y[i]

    sumax=0
    for j in x1:
        sumax=sumax+j

    sumax22=sumax**2

    sumay=0
    for j in y1:
        sumay=sumay+j

    sumax2=0
    for j in x1:
        sumax2=sumax2+j**2

    sumaxy=0
    for j in range(n):
        sumaxy=sumaxy+(x1[j]*y1[j])

    a=(n*sumaxy-(sumax*sumay))/((n*sumax2)- sumax22)
    b=((sumay-(a*sumax)))/(n)

    A.append(a)
    B.append(b)

#Funcion por minimos cuadrados
def y2(x2):
    return a*x2+b

#-----
#Calculamos el valor del promedio en el punto dado
def promedio(n):
    lisp=[0]*n
    for i in range(n):
        lisp[i]=y2(x1[i])
    return lisp
prom=promedio(n)
#-----
#Calculamos los 1's y 0's
def lista01(n):
    lis=[0]*n
    for i in range(n):
        if y1[i]<prom[i]:
            lis[i]=0
        else:
            lis[i]=1
    return lis

l=lista01(n)
#-----
```

APÉNDICE B. PRUEBA DE CORRIDAS POR ARRIBA Y ABAJO DE LA MEDIA

```
#Conteo de 1's y 0's
list=1
lista0=[]
lista1=[]
ceros=[0]*n
unos=[0]*n

while len(list)!=1:
    if 1 and 0 in list:
        k0 = list.index(0)
        k1 = list.index(1)
        if k0<k1:
            resta = k1-k0
            lista0.append(resta)
            ceros[resta]=ceros[resta]+1
            list= list[k1:]
        else:
            resta = k0-k1
            lista1.append(resta)
            unos[resta]=unos[resta]+1
            list= list[k0:]
    if 0 not in list:
        break
    if 1 not in list:
        break

#-----
lis=list
val=len(lis)
if lis[0]=0:
    ceros[val]=ceros[val]+1
else:
    unos[val]=unos[val]+1

m=100
nceros=ceros[1:m]
nunos=unos[1:m]

fcero.append(nceros)
funo.append(nunos)
Lista0.append(lista0)
Lista1.append(lista1)

fcero=np.array(fcero)
funo=np.array(funo)
Lista1=np.array(Lista1)
Lista0=np.array(Lista0)
#####

medidas0=[]
nmi0=[]
nma0=[]
def funcion0(v):
    minimo=min(v)
    maximo=max(v)
    moda=stats.mode(v)[0][0]
    frecuencia=stats.mode(v)[1][0]
    media=np.mean(v)
    mediana=np.median(v)
    varianza=np.var(v)
    desviacion=np.std(v)
    elemento=[moda, frecuencia, "{med:.3f}".format(media=media), mediana, "{var:.3f}".format(var=varianza), "{d"]
    ###
    medidas0.append(elemento)
    nmi0.append(minimo)
    nma0.append(maximo)

#####
medidas1=[]
nmi1=[]
nma1=[]
def funcion1(v):
    minimo=min(v)
    maximo=max(v)
    moda=stats.mode(v)[0][0]
    frecuencia=stats.mode(v)[1][0]
    media=np.mean(v)
```

APÉNDICE B. PRUEBA DE CORRIDAS POR ARRIBA Y ABAJO DE LA MEDIA

```
mediana=np.median(v)
varianza=np.var(v)
desviacion=np.std(v)
elemento=[moda,frecuencia,"{med:.3f}".format(med=media),mediana,"{var:.3f}".format(var=varianza),"{d}"]
###
medidas1.append(elemento)
nmil.append(minimo)
nma1.append(maximo)

#####
#GRaficas

#Archivo donde estan los datos de cada persona
a=pd.read_csv("dat.txt",delimiter='\t',skiprows=0,usecols=[0,1,3,4,5,6,7,8,11],names=['nombres','tiempo','minimo'])
#Se leen los datos del IMC en el archivo
nomb=np.array(a['nombres'])
direccion2="/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/0-1/"
b=np.loadtxt('dat.txt',delimiter='\t',skiprows=0,usecols=[11])
f=15
f1=16
for i in range(w1,w2):
    v1=np.arange(1,len(funo[i])+1)
    v2=np.arange(1,len(fcer0[i])+1)
    funcion0(Lista0[i])
    funcion1(Lista1[i])
    plt.figure()
    plt.subplot(2,1,1)
    plt.title("{n},{_IMC}={c},Edad={d}".format(n=a['nombres'][i],c=a['IMC'][i],d=a['edad'][i]),fontsize=f)
    plt.bar(v1,funo[i],label="Histograma_de_1's")
    plt.legend(loc="upper_right")
    #plt.xlabel(r"$N_{r}$",color='blue')
    plt.ylabel(r"$f_{N_{r}}$",color='blue',fontsize=f)
    plt.xticks(fontsize=5)
    plt.yticks(fontsize=5)
    plt.subplot(2,1,2)
    plt.bar(v2,fcer0[i],label="Histograma_de_0's")
    plt.xlabel(r"$N_{r}$",color='blue',fontsize=f)
    plt.ylabel(r"$f_{N_{r}}$",color='blue',fontsize=f)
    plt.xticks(fontsize=5)
    plt.yticks(fontsize=5)
    plt.legend(loc="upper_right")
    plt.savefig("/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/0-1/graficas/graficas-todas/")
    if b[i]<24.7:
        plt.savefig("/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/0-1/graficas/IMC-Nor/")
    if 24.7<=b[i]<30:
        plt.savefig("/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/0-1/graficas/IMC-Sol/")
    if 30<=b[i]:
        plt.savefig("/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/0-1/graficas/IMC-Ob/")

print nmi1, len(nmi1)
print nma1, len(nma1)
print nmi0, len(nmi0)
print nma0, len(nma0)
#####
ji1=[]
ji0=[]
for i in range(w1,w2):
    n1=nma1[i]-1
    n2=nmi1[i]-1
    rango=n1-n2
    sumaf=0
    for j in funo[i]:
        sumaf=sumaf+j
    valoresperado=sumaf/rango
    ###
    suma=0
    for k in range(n2,n1+1):
        if funo[i][k]>0:
            suma=suma+((funo[i][k]-valoresperado)**2)/(valoresperado)
    ji1.append(suma)
#####
n1=nma0[i]-1
n2=nmi0[i]-1
rango=n1-n2
sumaf=0
```

APÉNDICE B. PRUEBA DE CORRIDAS POR ARRIBA Y ABAJO DE LA MEDIA

```
for j in fzero[i]:
    sumaf=sumaf+j
valoresperado=sumaf/rango
###
suma=0
for k in range(n2,n1+1):
    if fzero[i][k]>0:
        suma=suma+ ((fzero[i][k]-valoresperado)**2)/(valoresperado)
    ji0.append(suma)

print ji1
print ji0
#####

DAT0 = num.column_stack((nomb, medidas0, ji0))
num.savetxt('medidastendencia0.txt', DAT0, delimiter=" ", fmt="% ")

DAT1 = num.column_stack((nomb, medidas1, ji1))
num.savetxt('medidastendencia1.txt', DAT1, delimiter=" ", fmt="% ")

plt.show()
```


Apéndice C

Transformada Rápida de Fourier

El programa que se hizo para calcular la transformada de Fourier a cada base de datos, se muestra a continuación:

```
import random
import numpy as np
import pylab as plt
import os
import pandas as pd
from pylab import plot, show, title, xlabel, ylabel, subplot
from scipy import fft, arange
from decimal import Decimal
import numpy as num

direccion="/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/nuevo-programa/datos-10m/"
archivos=os.listdir(direccion)
archivos.sort()

w1=0
w2=len(archivos)
y1=[]
for k in range(w2):
    y1.append(np.loadtxt(direccion+archivos[k], delimiter='\\t', skiprows=0, usecols=[0]))

n=len(y1[0])
y1 = np.array(y1)

Fs = len(y1[0]) # rata de muestreo
Ts = 0.010; # intervalo de muestreo
t = np.arange(0, int(n*Ts), Ts) # vector tiempo
tiempo=int(Fs*Ts)
intervalo=Decimal(50)/Decimal(n/2) #intervalo de la FFT
print intervalo

os.mkdir("graficas/Todas-{z}x".format(z=tiempo))
os.mkdir("graficas/IMC-N-{z}x".format(z=tiempo))
os.mkdir("graficas/IMC-S-{z}x".format(z=tiempo))
os.mkdir("graficas/IMC-O-{z}x".format(z=tiempo))

l3=0
l4=0.035
frecuencias=[]
maximos=[]
def plotSpectrum(y, Fs):
    n = len(y) # longitud de la senal
    k = arange(n)
    T = n/Fs
    frq = (k/T) # 2 lados del rango de PPM
    frq = frq[range(n/2)] # Un lado del rango de PPM
    m=Decimal(n/2)/Decimal(50)
    frq1=frq/m #Frecuencia
```

APÉNDICE C. TRANSFORMADA RAPIDA DE FOURIER

```

print m
Y = fft(y)/n # fft calcula la normalizacion
Y = Y[range(n/2)]
absolt=abs(Y)
limite1=int(Decimal(c[i])/Decimal(intervalo)) - 80
limite2=int(Decimal(c[i])/Decimal(intervalo)) + 70
print limite1, limite2
lista=list(absolt[limite1 : limite2])
maximo=max(lista)
maximos.append(maximo)
posicion=lista.index(maximo) + limite1
ppm=frq[posicion]
frec=Decimal(ppm)/Decimal((n/2)/(50))
print frec
elemento=[absolt[posicion],ppm,frec]
frecuencias.append(elemento)
j=i+1
print j, absolt[posicion],ppm,frec
#print '-----'
plot(frq1,absolt,'skyblue') # grafica el espectro de frecuencia
plt.ylim(13,14)
plt.xlabel(r"frecuencia_(Hz)",fontsize = 15)
plt.ylabel(r"|Y(f)|_(mv_.$\cdot\$_s)",fontsize = 15)

def plotSpectrum1(y,Fs):
n = len(y) # longitud de la senal
k = arange(n)
T = n/Fs
frq = k/T # 2 lados del rango de frecuencia
frq = frq[range(n/2)] # Un lado del rango de frecuencia
Y = fft(y)/n # fft calcula la normalizacion
Y = Y[range(n/2)]
absolt=abs(Y)
#print '-----'
#grafica del espectro de frecuencia
plot(frq[(n/2)/(75) :(n/2)/25],absolt[(n/2)/(75) :(n/2)/25],'skyblue')
xlabel('Frecuencia_(Hz)')
ylabel('|Y(f)|')

#-----

#Archivo donde estan los datos de cada persona
a=pd.read_csv("dat.txt",delimiter='\t',skiprows=0,usecols=[0],names=['nombres'])
#Se leen los datos del IMC en el archivo
nomb=np.array(a['nombres'])
b=np.loadtxt('dat.txt',delimiter='\t',skiprows=0,usecols=[11])
c=np.loadtxt('/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/FFT/frecs.txt',
delimiter='\t',skiprows=0,usecols=[2])

for i in range(w1,w2):
#Proceso de graficar la senal
plt.figure()
#subplot(2,1,1)
plt.title("{d}".format(d=a['nombres'][i]),fontsize = 15)
plotSpectrum(y1[i],Fs)
plot(frecuencias[i][2],frecuencias[i][0], '*', color='b')
#plt.xticks(fontsize = 5)
#plt.yticks(fontsize = 5)
#subplot(2,1,2)
#Se llama a la funcion con la senal y la rata de muestreo
#plotSpectrum1(y1[i],Fs)
#plot(frecuencias[i][1],frecuencias[i][0], '*', color='b')
#plt.xticks(fontsize = 5)
#plt.yticks(fontsize = 5)
plt.savefig("graficas/Todas-{z}x/{d}-{z}s.pdf".format(d=a['nombres'][i],z=tiempo))
if b[i]<24.7:
plt.savefig("graficas/IMC-N-{z}x/{d}-{z}s.pdf".format(d=a['nombres'][i],z=tiempo))
if 24.7<=b[i]<30:
plt.savefig("graficas/IMC-S-{z}x/{d}-{z}s.pdf".format(d=a['nombres'][i],z=tiempo))
if 30<=b[i]:
plt.savefig("graficas/IMC-O-{z}x/{d}-{z}s.pdf".format(d=a['nombres'][i],z=tiempo))

DAT = num.column_stack((nomb, frecuencias))
num.savetxt('frecuencias-{z}.txt'.format(z=tiempo), DAT, delimiter="_", fmt="%")

```

```
#plt.show()
```


Apéndice D

Parametrización de las señales de los Electroencefalogramas

```
import random
import numpy as np
from scipy import signal
import pylab as plt

#####Se importan los datos de txt y se convierte en una matriz#####
nombre='{0: 'x02c' }
file='/home/karina/Documentos/python/tesis/Apnea-ECG-Database/fft/parametrizacion/Graficas/datos/
datos{a}.txt'.format(a=nombre[0])
data1=np.loadtxt(file, delimiter='\t', skiprows=0, usecols=[0])

###Se convierte la matriz en una lista##

lista = data1.tolist()
y=list(np.array(lista))
n=len(y)
x = np.arange(0, int(n*0.010), 0.010)

n=len(lista)
I=0.010
m=n*I
l=n/100

l1=6 #num de valores antes del maximo donde se encuentran los minimos
l2=1 #limite donde se encuentran los maximos
l3=6 #num de valores despues del maximo para encontrar el punto 3
l4=12 #num de valores antes del punto 5 para encontrar el punto 4
l6=10 #num de valores despues del punto 5 para encontrar el punto 6
l7=7 #numeros de valores antes del punto 8, para encontrar el punto 7

list1=list(y[0:50])
min1=min(list1)
max1=max(list1)
p1=list1.index(min1)
p21=list1.index(max1)
x1= p1 * I
x21= p21 * I

list11=list(y[500:550])
min11=np.min(list11)
max11=np.max(list11)
p11=list11.index(min11) +500
p31=list11.index(max11) +500
x11= p11 * I
x31= p31 * I

list2=list(y[50:100])
min2=np.min(list2)
max2=np.max(list2)
p2=list2.index(min2) +50
p22=list2.index(max2) +50
x2= p2 * I
x22= p22 * I

list2=list(y[550:600])
min12=np.min(list12)
max12=np.max(list12)
p12=list12.index(min12) +550
p32=list12.index(max12) +550
x12= p12 * I
x32= p32 * I

list3=list(y[100:150])
list13=list(y[600:650])
```

APÉNDICE D. PARAMETRIZACIÓN DE LAS SEÑALES DE LOS ELECTROENCEFALOGRAMAS

```

min3=np.min(list3)
max3=np.max(list3)
p3=list3.index(min3) +100
p23=list3.index(max3) +100
x3= p3 * I
x23= p23 * I

list4=list(y[150:200])
min4=np.min(list4)
max4=np.max(list4)
p4=list4.index(min4) +150
p24=list4.index(max4) +150
x4= p4 * I
x24= p24 * I

list5=list(y[200:250])
min5=np.min(list5)
max5=np.max(list5)
p5=list5.index(min5) +200
p25=list5.index(max5) +200
x5= p5 * I
x25= p25 * I

list6=list(y[250:300])
min6=np.min(list6)
max6=np.max(list6)
p6=list6.index(min6) +250
p26=list6.index(max6) +250
x6= p6 * I
x26= p26 * I

list7=list(y[300:350])
min7=np.min(list7)
max7=np.max(list7)
p7=list7.index(min7) +300
p27=list7.index(max7) +300
x7= p7 * I
x27= p27 * I

list8=list(y[350:400])
min8=np.min(list8)
max8=np.max(list8)
p8=list8.index(min8) +350
p28=list8.index(max8) +350
x8= p8 * I
x28= p28 * I

list9=list(y[400:450])
min9=np.min(list9)
max9=np.max(list9)
p9=list9.index(min9) +400
p29=list9.index(max9) +400
x9= p9 * I
x29= p29 * I

list10=list(y[450:500])
min10=np.min(list10)
max10=np.max(list10)
p10=list10.index(min10) +450
p30=list10.index(max10) +450
x10= p10 * I
x30= p30 * I

minimos=[min1, min2, min3, min4, min5, min6, min7, min8, min9, min10, min11, min12, min13,
min14, min15, min16, min17, min18, min19, min20]
posicion1=[p1, p2, p3, p4, p5, p6, p7, p8, p9, p10, p11, p12, p13, p14, p15, p16, p17, p18, p19, p20]
maximos=[max1, max2, max3, max4, max5, max6, max7, max8, max9, max10, max11, max12, max13, max14, max15,
max16, max17, max18, max19, max20]
posicion2=[p21, p22, p23, p24, p25, p26, p27, p28, p29, p30, p31, p32, p33, p34, p35, p36, p37, p38, p39, p40]
valorx1=[x1, x2, x3, x4, x5, x6, x7, x8, x9, x10, x11, x12, x13, x14, x15, x16, x17, x18, x19, x20]
valorx2=[x21, x22, x23, x24, x25, x26, x27, x28, x29, x30, x31, x32, x33, x34, x35, x36, x37, x38, x39, x40]

q2= []
P2= []
pos2=[]
for i in range(len(maximos)):

```

APÉNDICE D. PARAMETRIZACIÓN DE LAS SEÑALES DE LOS ELECTROENCEFALOGRAMAS

```
    if maximos[i] > 12:
        #print i
        q2.append(maximos[i])
        P2.append(valorx2[i])
        pos2.append(posicion2[i])

q1= [y[0]]
P1= [0]
pos1= [0]
for i in pos2:
    a=i-11
    if a>0:
        lista=list(y[a:i])
        mi=min(lista)
        t=lista.index(mi) + a
        w=t * I
        pos1.append(t)
        P1.append(w)
        q1.append(mi)

print 'Punto_uno'
print q1
print P1
print pos1

print 'Punto_dos'
print 'Numeros_de_maximos: ' + repr(len(q2))
print q2
print P2
print pos2

q3= []
P3= []
pos3= []
for i in pos2:
    a=i+13
    if a<=n:
        lista=list(y[i:a])
        mi=min(lista)
        t=lista.index(mi) + i
        w=t * I
        pos3.append(t)
        P3.append(w)
        q3.append(mi)

print 'Punto_tres'
print 'Numero_de_minimos: ' + repr(len(q3))
print q3
print P3
print pos3

l5=(pos1[1]-pos3[0])/2

q5= []
P5= []
pos5=[]
for i in pos3:
    a=i+15
    if a<=n:
        lista=list(y[i:a])
        ma=max(lista)
        t=lista.index(ma) + i
        z=t * I
        P5.append(z)
        q5.append(ma)
        pos5.append(t)

q4= []
P4= []
pos4= []
for i in pos5:
    a=i-14
    if a<=n:
        lista=list(y[a:i])
        mi=min(lista)
```

APÉNDICE D. PARAMETRIZACIÓN DE LAS SEÑALES DE LOS ELECTROENCEFALOGRAMAS

```
t=lista.index(mi) + a
w=t * I
pos4.append(t)
P4.append(w)
q4.append(mi)

print 'Punto_cuatro'
print P4
print q4
print pos4

print 'Punto_cinco'
print P5
print q5
print pos5

q6= []
P6= []
pos6= []
for i in pos5:
    a=i+16
    if a<=n-1:
        lista=list(y[i:a])
        mi=min(lista)
        t=lista.index(mi) + i
        w=t * I
        pos6.append(t)
        P6.append(w)
        q6.append(mi)
print 'Punto_seis'
print q6
print P6
print pos6

q8= []
P8= []
pos8=[]
for i in pos1:
    a=i-15+10
    if a >0:
        lista=list(y[a : i] )
        ma=max(lista)
        t=lista.index(ma) + a
        w=t * I
        P8.append(w)
        q8.append(ma)
        pos8.append(t)

q7= []
P7= []
pos7=[]
for i in pos8:
    a=i-17
    if a >0:
        lista=list(y[a : i] )
        mi=min(lista)
        t=lista.index(mi) + a
        w=t * I
        P7.append(w)
        q7.append(mi)
        pos7.append(t)
print 'Punto_siete'
print P7
print q7
print pos7

print 'punto_ocho'
print P8
print q8
print pos8

q9= []
P9= []
pos9= []
for i in pos8:
    j=(pos1[1]-pos8[0])-4
```

APÉNDICE D. PARAMETRIZACIÓN DE LAS SEÑALES DE LOS ELECTROENCEFALOGRAMAS

```
a=i+j
if a<=n-1:
    lista=list( y[i : a] )
    mi=min( lista )
    t=lista.index(mi) + i
    w=t * I
    P9.append(w)
    q9.append(mi)
    pos9.append(t)
print 'Punto_nueve'
print q9
print P9
print pos9

punto= [q1,P1,q2,P2,q3,P3,q4,P4,q5,P5,q6,P6,q7,P7,q8,P8,q9,P9]
puntos= {0:'q1',1:'P1',2:'q2',3:'P2',4:'q3',5:'P3',6:'q4',7:'P4',8:'q5',9:'P5',10:'q6',11:'P6',
12:'q7',13:'P7',14:'q8',15:'P8',16:'q9',17:'P9'}

for i in range(0,len(punto)):
    archivo = open('/home/karina/Documentos/python/tesis/
    ~~~~~Apnea-ECG-Database/fft/parametrizacion/Graficas/coordenadas/
    ~~~~~coordenadas{a}{b}.txt'.format(a=puntos[i],b=nombre[0]),'w')
    for j in punto[i]:
        k=str(j)
        archivo.write('%s\n'%k)
    archivo.close()

X=[P1,P2,P3,P4,P5,P6,P7,P8,P9]
Y=[q1,q2,q3,q4,q5,q6,q7,q8,q9]

plt.figure()
plt.title("{b}".format(b=nombre[0]))
plt.plot(x,y)
for i in range(0,9):
    plt.plot(X[i],Y[i],'.')

plt.xlabel("s")
plt.ylabel("mV")
plt.savefig("graficas/{b}-10s.pdf".format(b=nombre[0]))
plt.savefig("/home/karina/Documentos/python/tesis/Apnea-ECG-Database/
fft/parametrizacion/Graficas/imagenes-latex/prog-x02.pdf")

plt.show()
```


Bibliografía

- [1] RG Mark AL Goldberger JH Peter T. Penzel, GB Moody. The apnea-ecg database. *Computers in Cardiology 2000. (Cat. 00CH37163)*, 27:53 – 74, 2000.
- [2] Richard T.O'connell y Anne B. Koehler Bowerman Bruce L. *Pronosticos, series de tiempo y regresion*. CENGAGE learning, cuarta edición edition, 2006.
- [3] Gabriel Silva. *Una formalización de la Transformada Rapida de Fourier*.
- [4] Robert Johnson. *Estadística Elemental*. Trillas, 1976.
- [5] Mario F. Triola. *Estadística*. Greg Tobin, 2009.
- [6] Kim E. Barrett. *Fisiología médica*. MCGRAW-HILL INTERAMERICANA EDITORES, 23 edición edition, 2010.
- [7] Marcos Matías Álvarez. *Procesamiento de señales de Electrocardiografía en Equipos Portables mediante Estrategias de Inteligencia Artificial*. PhD thesis, Universidad Nacional de Córdoba, Facultad de Ciencias Exactas, Físicas y Naturales, Facultad de Ciencias Médicas e Ingeniería Biomédica, Octubre 2016.
- [8] Juan M. García Manso. Aplicación de la variabilidad de la frecuencia cardiaca al control del entrenamiento deportivo: análisis en modo frecuencia. *Arch. Med. Deporte.*, 30:43, 2013.
- [9] Buenaventura Brito Díaz, José Juan Alemán Sánchez, and Antonio Cabrera de León. Frecuencia cardiaca en reposo y enfermedad cardiovascular. *Medicina Clínica*, 143(1):34 – 38, 2014.
- [10] Karem Josefina Parejo-Gallardo. Definition of obstructive sleep apnea-hypopnea syndrome (OSAHS). *Rev. Fac. Med. 2017 Vol. 65 Sup: S9-10*, 2017.
- [11] V. M. Eguia and J. A. Cascante. Síndrome de apnea-hipopnea del sueño: Concepto, diagnóstico y tratamiento médico. *Anales del Sistema Sanitario de Navarra*, 30:53 – 74, 2007.