



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Instituto de Ciencias Matemáticas

**“UN MODELO ESTRUCTURAL DE SERIES DE TIEMPO PARA LA
PREDICCIÓN DE LA DEMANDA DE ATENCIÓN MÉDICA EN EL
SISTEMA MUNICIPAL DE SALUD”**

Caso: M.I. Municipalidad

T E S I S D E G R A D O

Previo a la obtención del Título de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

PRESENTADO POR:

Richard John Parra Suárez

GUAYAQUIL – ECUADOR

2012

AGRADECIMIENTO

Primeramente agradezco a Dios por haberme dado la fortaleza suficiente de estar con vida, y por permitirme alcanzar mi meta propuesta con salud.

A mi madre y padre quienes me apoyaron desde el principio con mi preparación académica dándome buenos consejos cada día.

A mi esposa e hijas que me han dado fuerza para seguir adelante en mis estudios.

Al Matemático John Ramírez, director de mi tesis de grado, por haberme dado la oportunidad de realizarla, con su ayuda y guía durante la elaboración de la misma.

DEDICATORIA

Para mi padre Sr. Vicente Parra que siempre me apoyo en mis estudios; en lo económico, moral y con sus palabras de aliento constante que me impulsaban a seguir luchando cada día para poder realizarme profesionalmente como ingeniero.

A mis hijas quienes me inspiraron a seguir luchando fervientemente para darles una mejor vida y ser su ejemplo a seguir.

TRIBUNAL DE GRADUACIÓN

Ing. Vanessa Salazar

PRESIDENTA

Mat. John Ramírez Figueroa

DIRECTOR DE TESIS

Ing. Sandra García

VOCAL

DECLARACIÓN EXPRESA

"La responsabilidad del contenido de este trabajo final de graduación de grado, nos corresponde exclusivamente; y el patrimonio intelectual de la misma a la Escuela Superior Politécnica del Litoral".

(Reglamento de Graduación de la ESPOL)

RICHARD JOHN PARRA SUAREZ

RESUMEN

La presente tesis consiste en buscar un modelo estructural de series de tiempo para la predicción de la demanda de pacientes que pueden asistir a un centro de salud ubicado en la ciudad de Guayaquil, utilizando el filtro de Kalman. La tesis está conformada por cuatro capítulos. En el primer capítulo se describe en breves rasgos el tipo de atención médica que debe de tener un paciente, y la importancia de ser atendido a tiempo. En el capítulo dos se describen los antecedentes e importancia de tener una buena atención en el centro de salud, dado las dificultades que existen para tener un cupo y ser atendido, además se describen las metodologías a utilizarse por medio de estadísticos de pruebas para determinar un modelo adecuado. En el tercer capítulo se determinara el modelo adecuado, mediante pruebas estadísticas y análisis de predicción, para medir la demanda de pacientes que asisten al centro de salud en el área de consulta general y odontología, para lo cual se usara el software ox metric versión 4.10 y Excel. En el capítulo cuatro se describen los resultados del proyecto mediante las conclusiones obtenidas.

SIMBOLOGÍA

\bar{x} : Media muestral

Med: mediana

S^2 : Varianza Muestral

μ_t : Tendencia

Ψ_t : Ciclo

γ_t : Componente estacional

ε_t : Componente irregular

σ_ε^2 : Varianza

ω_t : Perturbación

m_t : Momento de los Errores

C : Kurtosis

N_{BS} : Estadístico de Bowman Shenton

ÍNDICE GENERAL

AGRADECIMIENTO	II
DEDICATORIA	III
TRIBUNAL DE GRADUACIÓN	IV
DECLARACIÓN EXPRESA	V
RESUMEN	VI
SIMBOLOGÍA	VII
CAPÍTULO I	
SALUD E HIGIENE.....	
1.1.INTRODUCCIÓN	1
1.2.ANTECEDENTES DE LA SALUD.....	2
1.3.POLÍTICA DE LA SALUD	2
1.4.SALUD BASADA EN LA ATENCIÓN PIMARIA	3

CAPÍTULO II

2.	DESCRIPCIÓN DEL PROBLEMA Y MARCO TEÓRICO.....
2.1.	ANTECEDENTES DEL PROBLEMA.....	5
2.2.	DESCRIPCIÓN DEL PROBLEMA.....	6
2.3.	OBJETIVO GENERAL Y DESCRIPCIÓN DE LOS DATOS EN ESTUDIO.....	6
2.4.	DETERMINACIÓN Y DESCRIPCIÓN DE LOS MÉTODOS DE SOLUCIÓN.....	7
2.4.1.	LEVANTAMIENTO DE INFORMACIÓN.....	7
2.5.	MARCO TEÓRICO.....	8
2.5.1.	CONCEPTOS BÁSICOS.....	8
2.5.1.1.	VARIABLE.....	8
2.5.1.1.1.	VARIABLE CUANTITATIVA	8
2.5.1.1.2.	VARIABLE CUALITATIVA.....	9
2.5.1.1.3.	VARIABLE DEPENDIENTE.....	9
2.5.1.1.4.	VARIABLE INDEPENDIENTE.....	9
2.5.2.	MÉTODOS USADO EN EL ANÁLISIS ESTADÍSTICO.....	10
2.5.2.1.	ESTADÍSTICA DESCRIPTIVA.....	10
2.6.	MODELOS ESTRUCTURALES DE SERIES TEMPORALES.....	12
2.6.1.	MODELOS DE SERIES TEMPORALES DE UNA VARIABLE.....	13
2.7.	PRUEBAS PARA DETERMINAR EL MODELO ADECUADO.....	31
2.7.1.	NORMALIDAD.....	31
2.7.2.	HETEROSCEDASTICIDAD.....	33

2.7.3. AUTOCORRELACIÓN.....	34
2.7.4. PEV (PREDICTION ERROR VARIANCE).....	36
2.7.5. DETECCIÓN DE DATOS ATÍPICOS Y RUPTURAS ESTRUCTURALES.....	36

CAPÍTULO III

3. APLICACIÓN ESTADÍSTICA.....
3.1. INTRODUCCIÓN.....	39
3.2. ANÁLISIS ESTADÍSTICO	40
3.2.1. VARIABLE CONSULTA GENERAL.....	41
3.2.1.1. PRIMERA	
ETAPA.....	42
MODELO 1.....	42
MODELO 2.....	45
3.2.1.2. SEGUNDA ETAPA.....	48
MODELO 3.....	48
MODELO 4.....	53
3.2.1.3. TERCERA ETAPA.....	59
ANÁLISIS DE PREDICCIÓN DEL MODELO.....	59
3.2.2. VARIABLE ODONTOLOGÍA.....	62
3.2.2.1. PRIMERA ETAPA.....	63
MODELO 1.....	63
MODELO 2.....	66

3.2.2.2. SEGUNDA ETAPA.....	70
MODELO 3.....	70
3.2.2.3. TERCERA ETAPA.....	75
ANÁLISIS DE PREDICCIÓN DEL MODELO.....	75

CAPÍTULO IV

4. CONCLUSIONES Y RECOMENDACIONES.....	77
4.1. VARIABLE DE CONSULTA GENERAL.....	77
4.2. VARIABLE DE ODONTOLOGÍA.....	80
4.3. RECOMENDACIONES.....	82

BIBLIOGRAFÍA

REFERENCIAS

BIBLIOGRÁFICAS.....	
----------------------------	--

CAPÍTULO I

1. SALUD E HIGIENE

1.1. INTRODUCCIÓN

La Salud es muy importante para todas las personas, ya que en ella influyen muchos factores, esto es hablar de todos los beneficios que necesita el cuerpo humano para una buena calidad de vida.

Es evidente la estrecha relación que existe entre salud y calidad de vida. No obstante, con frecuencia no se es consciente de la relación, directa, entre la construcción de un hábito alimenticio, el mantenimiento de un cuerpo saludable, salud e higiene.

Para garantizar la atención a personas en un centro médico; es necesario la creación de un sistema que permite mejorar la atención al paciente. En este sentido, la implantación de infraestructuras y procesos de control hacen posible los impresionantes avances en la mejora de la calidad y la esperanza de vida.

1.2. ANTECEDENTES DE LA SALUD

En algunas ciudades del País, la mayoría de los Centros Médicos o dispensarios de Salud, efectúan campañas de emergencia, brindando atención primaria a las personas mas vulnerable como son los niños, jóvenes, mujeres, adultos mayores, personas con discapacidad, personas en riesgo, pobres y excluidas, personas de escasos recursos, brindando atención médica y medicina gratuita.

Otras Instituciones que también prestan servicios de salud en una ciudad son: Dirección Provincial de Salud del Guayas, Instituto Ecuatoriano de Seguridad Social, Fuerzas Armadas, Policía Nacional, Junta de Beneficencia, SOLCA, Universidad de Guayaquil a través del Hospital Universitario y diversas Organizaciones.

1.3. POLÍTICA DE LA SALUD

Una política de Salud resulta del esfuerzo organizado y disciplinario de un conjunto de personas que, con experiencia y conocimiento le han dado buenos resultados el programa de ATENCION PRIMARIA ⁽¹⁾. Primero, porque permite involucrarse en todas las enfermedades que se dan en desventaja a la

1 “**ATENCIÓN PRIMARIA**, es la asistencia sanitaria esencial puesta al alcance de todos los individuos y familias de la comunidad por medios que le resulten aceptables.

salud del cuerpo y segundo, porque regulariza los recursos hacia las prioridades de salud locales como son; Nutrición apropiada, Atención de salud materno infantil, Inmunización, Tratamiento básico de los problemas de salud y Suministro de medicamentos esenciales.

1.4. SALUD BASADA EN LA ATENCION PRIMARIA

Como atención primaria a los pacientes podemos encontrar algunos departamentos de suma importancia como el departamento médico, que tiene como objetivo preservar la salud del ser humano, elevando el nivel de complejidad de la atención médica, desarrollando cada uno de sus componentes esenciales como son la Curación, rehabilitación, prevención de enfermedades, promoción y fomento de la Salud.

Otro departamento es la atención Gineco-Obstétrica, la cual comprende la atención a la mujer que se encuentra en estado de gestación o fuera de esta, en forma general en la que se realiza el control prenatal con todos los medios necesarios básicos de detención oportuna del embarazo, control de inmunizaciones, exámenes de laboratorio, control de peso, ecosonografía, entrega de medicinas, etc.

Los médicos especializados tratan de brindar a todo el público una buena atención, acorde a las necesidades del paciente, así lo demuestran todas las personas que acuden a dicho programa ⁽²⁾.

Además de las atenciones inmediatas; se ofrecen las atenciones de Pediatría, Cardiología, Quirúrgica, laboratorio de patología, odontología y medicina veterinaria.

CAPÍTULO II

2. DESCRIPCIÓN DEL PROBLEMA Y MARCO TEÓRICO

2.1. ANTECEDENTES DEL PROBLEMA

Numerosos pacientes podrían estar siendo atendidos en estos momentos en cualquier centro médico, haciendo innumerables filas con personas que también quieren ser atendidos en el mismo horario, sin saber si serán revisados o ingresados a consulta o no.

Como resultado de estas filas indeterminadas de pacientes, surgen personas que no son atendidas, ya que en todo centro médico o dispensario de salud existe un horario de atención y no logran abastecerse para la cantidad de personas o pacientes que llegan a este lugar.

2.2. DESCRIPCIÓN DEL PROBLEMA

La mayoría de los centros de salud y dispensarios médicos, tienen un sistema informático para el manejo de historias clínicas e ingresos de atención a pacientes; las cuales no ayudan completamente a satisfacer las necesidades de los mismos, ya sea porque no asisten a la cita, porque no hay turno o porque se colapsó el sistema.

Por esta razón o simplemente por tener un mejor sistema de manejo y atención a personas o pacientes; surge la idea de implementar un sistema que ayude a saber cuántos pacientes pueden ser atendidos en un día, mes o año específicamente; y así de esta manera se logrará dar mejores resultados y satisfacer la demanda de pacientes en el Centro Municipal de Salud.

2.3. OBJETIVO GENERAL Y DESCRIPCIÓN DE LOS DATOS EN ESTUDIO

El objetivo general de este estudio es realizar un modelo estructural basado en series de tiempo que ayuden a predecir la demanda de atención médica en el sistema Municipal de Salud.

Se definió como objeto de estudio a la variable que indica la cantidad de pacientes que han sido atendidos en consulta general y odontología en los meses de enero a diciembre de 3 años en particular.

El estudio se lo realizó en el Municipio de una ciudad de la Provincia del Guayas, los datos fueron recopilados en el área de estadística de la Dirección de Salud e Higiene. Se consideró a los pacientes que acudieron a consulta en Medicina General y Odontología desde el mes de Enero del 2007 hasta diciembre del 2009.

2.4. DETERMINACIÓN Y DESCRIPCIÓN DE LOS MÉTODOS DE SOLUCIÓN.

2.4.1. LEVANTAMIENTO DE INFORMACIÓN

Para realizar el levantamiento de la información se realizaron varias visitas a la Dirección de Salud e Higiene con la finalidad de obtener los datos específicos de los meses y años en particular como objeto de nuestro estudio.

Además se recopilaron datos del año 2010, como objeto de comparación; para determinar el modelo que mejor se ajuste a los datos reales.

2.5. MARCO TEÓRICO

En el presente apartado se dará a conocer las diferentes técnicas estadísticas a utilizarse en el estudio, definiciones y descripciones que ayudarán a la obtención de los resultados y la realización del análisis en los capítulos posteriores.

2.5.1. CONCEPTOS BÁSICOS

2.5.1.1. VARIABLE

Una variable no es sino un conjunto de las distintas modalidades o valores que toma un carácter. Las variables se clasifican en:

2.5.1.1.1 Variable Cuantitativa (numérica o Continua): Son las que pueden expresarse numéricamente. Se pueden cuantificar los resultados experimentales por medio de instrumentos adoptando unidades de medida para valorar los diferentes resultados. Discretas son aquellas cuyas categorías sólo pueden tomar valores enteros. Continuas son aquellas cuyas

categorías pueden fraccionarse según cualquier criterio.

2.5.1.1.2. Variable Cualitativa (Categóricas): Son aquellas que no aparecen en forma numérica, así como categorías o atributos (sexo, profesión, color de ojos). Las variables cualitativas sólo pueden ser nominales u ordinales ⁽³⁾.

2.5.1.1.3. Variable Dependiente: La variable dependiente puede ser definida como los cambios sufridos por los sujetos como consecuencia de la manipulación de la variable independiente por parte del experimentador.

2.5.1.1.4 Variable Independiente: Es la que modifica de una u otra manera a la variable dependiente, llamándose también según el caso factor de riesgo, factor predictivo, etc.

³ **Variable Categórica Ordinal:** recogen la idea de orden, pero no medirse distancia entre una categoría y otra. **Variable Categórica nominal:** lo único que puede hacerse es establecer frecuencias en cada atributo y la igualdad o desigualdad entre los diferentes casos, ver cuál es el grupo que tiene mayor frecuencia alcanzando el concepto de "moda" (valor que se repite con mayor frecuencia)

2.5.2. MÉTODOS USADO EN EL ANÁLISIS ESTADÍSTICO

Para el análisis estadístico de los datos, se usarán diversas técnicas que han sido aprendidas a lo largo de la carrera profesional, se usarán modelos estructurales de series temporales.

2.5.2.1. ESTADÍSTICA DESCRIPTIVA

La estadística descriptiva implica la abstracción de varias propiedades de conjunto de observaciones, mediante el empleo de métodos gráficos, tabulares o numéricos.

a) Medidas de Tendencia Central.

Al describir grupos de observaciones, con frecuencia se desea describir un grupo con un solo número. Para tal fin, desde luego, no se usará el valor más elevado ni el valor más pequeño como único representante, ya que sólo representan los extremos en lugar de valores típicos. Por consiguiente, sería más conveniente buscar el valor central. De estas medidas las que se imponen son la media, mediana y moda.

➤ **MEDIA ARITMÉTICA**

La idea de media o promedio formaliza el concepto intuitivo de punto de equilibrio de las observaciones. Es decir, es el punto medio del recorrido de la variable según la cantidad de valores obtenidos. Su fórmula matemática es.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{N} = \frac{\sum_{i=1}^N x_i}{N} \quad (4)$$

➤ **MEDIANA**

La mediana, a diferencia de la media no busca el valor central del recorrido de la variable según la cantidad de observaciones, sino que busca determinar el valor que tiene aquella observación que divide la cantidad de observaciones en dos mitades iguales. Por lo tanto es necesario ordenar los datos, y debido a ello, este cálculo depende de la posición relativa a los valores obtenidos. Es necesario, antes que nada, ordenar los datos de menor a mayor (o viceversa).

$$med = x_{\left(\frac{N+1}{2}\right)} \text{ o } med = \frac{N}{2} \text{ en caso de que } N \text{ es impar; y}$$

$$med = \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2} \text{ en caso de que } N \text{ sea par} \quad (5)$$

4 Fórmula para obtener la Media de una población.

5 Fórmula para obtener la Mediana de un conjunto de observaciones

b) Medidas de Dispersión

Se llaman medidas de dispersión aquellas que permiten retratar la distancia de los valores de la variable a un cierto valor central, o que permiten identificar la concentración de los datos en un cierto sector de recorrido de la variable. Se trata de coeficiente para variables cuantitativas.

➤ VARIANZA

Esta es la medida de dispersión más utilizada para notar la variabilidad de los datos con respecto a la media aritmética. Es un estimador insesgado de la varianza poblacional. Es expresada por S^2 y se la obtiene a través de la siguiente fórmula.

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (6)$$

2.6. MODELOS ESTRUCTURALES DE SERIES TEMPORALES ⁽⁷⁾

Debido a la facilidad que existe hoy en día para obtener datos y, sobre todo, la posibilidad de construir series más largas y con mayor frecuencia de observación, se ha estimulado el uso, e

6 Fórmula para obtener la Varianza de una población.

7 Tomado de Modelos estructurales en el contexto de las series temporales económicas – Gloria Martín Rodríguez

incluso el desarrollo, de técnicas estadístico-económicas para su tratamiento. Así mismo, para el investigador resulta cada vez más difícil mantener el supuesto de un patrón de comportamiento estable a lo largo del tiempo; de hecho, las propiedades estadísticas con respecto a la salud poseen un carácter evolutivo y, en este sentido, los modelos estructurales de series temporales constituyen una técnica apropiada, puesto que admiten la posibilidad de que cada uno de los componentes típicos de una serie temporal posea una naturaleza estocástica.

2.6.1. MODELOS DE SERIES TEMPORALES DE UNA VARIABLE

Un modelo univariado de series temporales se puede definir como

$$y_t = \mu_t + \Psi_t + \gamma_t + \varepsilon_t$$

Donde μ_t representa la tendencia, Ψ_t el ciclo, γ_t el componente estacional y ε_t el componente irregular, que recoge los movimientos no sistemáticos. Lo interesante de estos modelos estructurales es que en su flexibilidad para recoger los cambios en el comportamiento de la serie mediante la consideración de sus diferentes componentes como procesos estocásticos dirigidos por

perturbaciones aleatorias. Así se comprueban las distintas especificaciones, teniendo que, un modelo estructural no necesariamente debe ser definido en términos de todos los componentes que no se observan.

El modelo más simple para el componente irregular es un proceso de ruido blanco, es decir, una sucesión de variables aleatorias no correlacionadas en series de tiempo, con media constante igual a cero y varianza constante e igual a σ_ε^2 .

Existen modelos más generales en los que el proceso estocástico ε_t es un proceso de autorregresión estacionario de orden p .

La tendencia es el componente que indica la dirección en que se mueve la serie en el largo plazo; quiere decir que no se especifica como una función determinista del tiempo, sino que, por el contrario, se asume que el componente de tendencia es un proceso estocástico cuyo valor en un instante dado del tiempo (nivel estocástico) es el resultado de añadir al valor del

componente en el instante anterior un determinado incremento de tipo aleatorio (pendiente estocástica) y un término de perturbación aleatoria. Por supuesto, cabe la posibilidad de que el nivel o la pendiente no posean un carácter estocástico y, más aún, la pendiente puede estar o no presente.

Así la formulación se puede flexibilizar para permitir que el nivel de la serie cambie en el tiempo, dando lugar al *modelo de nivel local*, en el que el nivel en cada momento del tiempo es el nivel del período anterior más un elemento aleatorio. En este modelo, definido como:

$$y_t = \mu_t + \varepsilon_t$$

con
$$\mu_t = \mu_{t-1} + \eta_t \text{ para } t = 1, \dots, n$$

De donde $\eta_t \sim N(0, \sigma_\eta^2)$ ⁽⁸⁾ son independientes; existe un término de perturbación aleatoria alrededor de un nivel subyacente que oscila sin una dirección particular. De ahí que también se conoce por el nombre de caminata aleatoria con ruido. Si la varianza del nivel es cero, se obtiene el modelo de nivel

8 La expresión $N(0, \sigma^2)$ denota que las variables aleatorias son serialmente independientes y normalmente distribuidas con media nula y varianza σ^2 .

determinista antes especificado. Si la varianza del componente irregular es cero, pero la varianza del nivel es distinta de cero, la serie sólo posee componente de nivel cuyo proceso generador es una caminata aleatoria ⁽⁹⁾.

Si a los elementos descritos hasta el momento para el componente de nivel se añade una pendiente, es posible considerar que el incremento medio por unidad de tiempo sea constante, pero también se puede recurrir a formulaciones más flexibles. Una de ellas es el *modelo de tendencia lineal local* ⁽¹⁰⁾, definido como:

$$y_t = \mu_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \zeta_t \text{ para } t=1, \dots, n$$

Así tenemos que $\zeta_t \sim N(0, \sigma_\zeta^2)$ ⁽¹¹⁾. El término de perturbación ζ_t otorga a la pendiente un carácter estocástico,

9 La inclusión de un componente de caminata aleatoria significa que el proceso es no estacionario y, además, que es equivalente a un proceso ARIMA(0,1,1), donde el parámetro MA sólo cubre la mitad del espacio usual del parámetro, es decir, $-1 \leq \theta \leq 0$; además, si $0.2 = \sigma_\eta$, se obtiene un proceso ARIMA en el que el parámetro MA toma el valor -1 .

10 Aproximación local de una tendencia lineal.

11 El modelo de tendencia lineal local es equivalente a un proceso ARIMA (0,2,2); además, será un proceso no invertible cuando $\sigma_\zeta^2 = 0$ y, por el contrario, si $\sigma_\zeta^2 > 0$ el proceso es invertible independientemente de cual sea el valor de $\sigma_{\zeta\eta}^2$.

mientras que, dados el nivel y la pendiente del período anterior, la naturaleza estocástica del nivel en el instante actual se deriva de la presencia del término de perturbación η_t ⁽¹²⁾. Así pues, si la varianza de ζ_t es nula, de forma que la pendiente es constante, el modelo resultante conocido como *modelo de nivel local con derivada* considera un nivel estocástico y una pendiente fija, es decir,

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \quad \text{para } t=1, \dots, n \quad (13)$$

Si la pendiente no es sólo constante sino nula, el modelo anterior se transforma en el que se ha denominado modelo de nivel local. Finalmente, es posible mantener el carácter estocástico de la pendiente y, al mismo tiempo, suponer que, dados los valores del nivel y la pendiente en el instante $t - 1$, el valor del nivel en el instante t es no estocástico, es decir,

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} \\ \beta_t &= \beta_{t-1} + \zeta_t \quad \text{para } t=1, \dots, n \end{aligned} \quad (14)$$

12 Se supone que estos dos términos de perturbación aleatoria son mutua y seriamente independientes.

13 Si se supone, además, que $\sigma_\zeta^2 = 0$, se obtiene una tendencia lineal determinística.

14 El modelo de nivel fijo y pendiente estocástica se denomina también *modelo de tendencia* dado que el componente tendencial es una caminata aleatoria integrado cuya estimación fluctúa suavemente, sin incrementos o decrementos abruptos.

Por otra parte, el modelo de tendencia lineal local se puede modificar introduciendo un factor de amortiguación en el componente de pendiente de la tendencia, es decir,

$$\beta_t = \rho\beta_{t-1} + \zeta_t \text{ para } t=1, \dots, n$$

Donde $0 \leq \rho \leq 1$. Se obtiene entonces el *modelo de tendencia lineal local amortiguado* ⁽¹⁵⁾.

Así tenemos que, a la hora de elegir una de estas especificaciones, es importante, antes, examinar el comportamiento de la serie a lo largo de toda la muestra y, después, comprobar cuál produce un mejor ajuste; aunque tampoco se pueden ignorar las implicaciones de cada especificación y los objetivos que se persiguen con el modelo.

En muchas series temporales es importante distinguir entre una tendencia de largo plazo y los movimientos asociados a un componente cíclico, que admite también tanto una

15 Estos modelos pueden ser utilizados para recoger tendencias globales no lineales como, por ejemplo, las presentes en aquellos casos en lo que existe un nivel de saturación. Esta solución, aunque es viable en algunos momentos del tiempo, no capta ciertas características de comportamiento de manera tan adecuada como otras formulaciones alternativas, tales como curvas de crecimiento.

especificación determinística como estocástica. Un ciclo determinista describe un comportamiento perfectamente periódico con una frecuencia dada, que puede ser recogido mediante una función de senos y cosenos del siguiente tipo,

$$\Psi_t = \alpha \cos(\lambda t) + \beta \operatorname{sen}(\lambda t),$$

Donde λ es la frecuencia angular del ciclo, en radianes ⁽¹⁶⁾. Si se admite que los parámetros α y β evolucionen aleatoriamente en el tiempo, el ciclo adquiere naturaleza estocástica. Pero antes de incorporar el carácter estocástico, conviene garantizar la continuidad del componente cíclico expresándolo de la forma recursiva siguiente:

$$\begin{pmatrix} \Psi_t \\ \Psi'_t \end{pmatrix} = \begin{pmatrix} \cos \lambda & \operatorname{sen} \lambda \\ -\operatorname{sen} \lambda & \cos \lambda \end{pmatrix} \begin{pmatrix} \Psi_{t-1} \\ \Psi'_{t-1} \end{pmatrix}, \quad t = 1 \dots, T$$

Con $\Psi_t = \alpha$ y $\Psi'_t = \beta$. Los nuevos parámetros son Ψ_{t-1} , el valor del ciclo en $t-1$, y Ψ'_{t-1} , que aparece por construcción en la formación de Ψ_t . Entonces, el ciclo es estocástico al

16 El parámetro λ representa el número de veces que el ciclo se repite en un período de tiempo de longitud 2π . De este modo, el cociente $\lambda/2\pi$ expresa el número de veces que se repite el ciclo por unidad de tiempo.

introducir dos términos de perturbación ruido blanco normal k_t y k'_t de forma que:

$$\begin{pmatrix} \Psi_t \\ \Psi'_t \end{pmatrix} = \begin{pmatrix} \cos \lambda & \text{sen } \lambda \\ -\text{sen } \lambda & \cos \lambda \end{pmatrix} \begin{pmatrix} \Psi_{t-1} \\ \Psi'_{t-1} \end{pmatrix} + \begin{pmatrix} k_t \\ k'_t \end{pmatrix}$$

Se supone que k_t y k'_t están mutuamente intercorrelacionadas y poseen la misma varianza σ_k^2 . Por otra parte, para recoger los comportamientos pseudocíclicos que caracterizan a muchas series temporales, conviene incluir un factor de amortiguación (*damping*) ρ que aporta mayor flexibilidad al ciclo estocástico, es decir,

$$\begin{pmatrix} \Psi_t \\ \Psi'_t \end{pmatrix} = \rho \begin{pmatrix} \cos \lambda & \text{sen } \lambda \\ -\text{sen } \lambda & \cos \lambda \end{pmatrix} \begin{pmatrix} \Psi_{t-1} \\ \Psi'_{t-1} \end{pmatrix} + \begin{pmatrix} k_t \\ k'_t \end{pmatrix}$$

Dónde $0 \leq \rho \leq 1$. Si $\lambda = 0$ o $\lambda = \pi$, el ciclo estocástico se convierte en un proceso de autorregresión de primer orden. Si $\lambda = 0$ y $\rho = 1$, el componente cíclico se transforma en el modelo de nivel local especificado para la tendencia.

El componente cíclico puede ser combinado con el de tendencia de muchas maneras, pero las formulas más habituales son: *modelos cíclicos*,

$$y_t = \mu + \Psi_t + \varepsilon_t$$

Con $0 \leq \rho \leq 1$; *modelo tendencia-ciclo*,

$$y_t = \mu_t + \Psi_t + \varepsilon_t$$

Donde μ_t recoge una tendencia lineal local; *modelo de tendencia cíclica*,

$$y_t = \mu_t + \varepsilon_t$$

Donde

$$\mu_t = \mu_{t-1} + \Psi_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \zeta_t \quad \text{para } t=1, \dots, n$$

Otras especificaciones consisten en añadir un componente AR(2) a

$$y_t = \mu_t + \varepsilon_t \quad \text{o} \quad \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$

Si las observaciones de la serie en cuestión se registran con un periodo inferior a la anual, suelen aparecer efectos estacionales y, por ello, conviene introducir este componente en el modelo. Si las fluctuaciones estacionales son estables, el componente estacional se puede considerar determinista o fijo. Si el número de estaciones es s , y dado que la suma de los efectos estacionales a lo largo del año debe ser nula, el componente estacional en el instante t se puede especificar como:

$$\gamma_t = \sum_{j=1}^{s-1} z_{j,t} \gamma_j$$

Donde $z_{j,t}$, $j=1, \dots, s-1$, es igual a 1 si t pertenece a la estación j , -1 si t pertenece a la estación s y 0 en otro caso.

Formalmente, se tiene que:

$$\gamma_t = \begin{cases} \gamma_j, & \text{si } t \text{ pertenece a la estación } j \\ -\sum_{j=1}^{s-1} \gamma_j, & \text{si } t \text{ pertenece a la estación } s \end{cases} \quad j=1, \dots, s-1$$

Por otro lado, si el patrón estacional evoluciona a lo largo del tiempo, es más adecuado incorporar un componente estacional estocástico, que admite, al menos, dos formulaciones posibles. Una forma de conseguir que el patrón estacional evolucione en el tiempo es permitir que la suma de los efectos estacionales a lo largo del año no sea estrictamente nula, sino igual a un término de perturbación aleatoria. Es decir, los valores del componente estacional en los instantes del tiempo $t, t-1, \dots, t-s+1$ están ligados por la relación $\sum_{j=0}^{s-1} \gamma_{t-j} = \omega$, o bien, $\gamma_t = \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t$, siendo ω_t un término de perturbación aleatoria tal que $\omega_t \sim N(0, \sigma_\omega^2)$.

Otra posibilidad es suponer que el efecto de cada estación evolucione de acuerdo a una caminata aleatoria. Es decir, si t pertenece a la estación j , $\gamma_t = \gamma_{j,t}$; donde

$$\gamma_{j,t} = \gamma_{j,t-1} + \omega_{j,t}, j = 1, \dots, s$$

y ω_t sigue siendo un término de perturbación aleatorio tal que

$$\omega_t \sim N(0, \sigma_\omega^2).$$

Otra alternativa es la formulación trigonométrica, en la que

$$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{j,t}$$

Donde $[s/2]$ denota la parte entera de $s/2$ y cada $\gamma_{j,t}$ es generada por

$$\begin{pmatrix} \gamma_{j,t} \\ \gamma'_{j,t} \end{pmatrix} = \begin{pmatrix} \cos \lambda_j & \text{sen } \lambda_j \\ -\text{sen } \lambda_j & \cos \lambda_j \end{pmatrix} \begin{pmatrix} \gamma_{j,t-1} \\ \gamma'_{j,t-1} \end{pmatrix} + \begin{pmatrix} \omega_{j,t} \\ \omega'_{j,t} \end{pmatrix}$$

Para $j = 1, \dots, [s/2]$ y $t = 1, \dots, T$; $\lambda_j = 2\pi j/s$ es la frecuencia, en radianes, y $\omega_{j,t}$ y $\omega'_{j,t}$ son dos términos de perturbación aleatoria mutuamente intercorrelacionadas distribuidos $N(0, \sigma_\omega^2)$ en ambos casos. Si s es par, $\gamma_{s/2,t} = -\gamma_{s/2, t-1} + \omega_{s/2,t}$. Nótese que $\gamma'_{j,t}$ interviene por razones de construcción con el mismo fin que en el componente cíclico.

Observe también que en cualquiera de las dos especificaciones el número de parámetros que intervienen en el componente estacional es el mismo y, si $\sigma_{\omega}^2 = 0$, el patrón estacional es determinístico.

Dado que el patrón estacional cambia de forma relativamente suave a lo largo del tiempo, algunas veces podría ser razonable eliminar algunos de los términos trigonométricos correspondientes a las frecuencias más altas. Aunque, cuando dichos términos son responsables de una parte importante de la variación estacional, ésta no sería una forma adecuada de proceder.

Así ocurre cuando se dispone de observaciones semanales en las que existe un importante efecto intermensual, que se refleja en ciclos de períodos dos y cuatro semanas; en este caso, el componente estacional podría formularse usando los términos trigonométricos asociados a las frecuencias bajas así como a las frecuencias $\pi/2$ y π , es decir,

$$y_t = \mu_t + \gamma_t^{bf} + \gamma_t^{im} + \varepsilon_t,$$

Donde γ_t^{bf} contiene, por ejemplo, las variaciones correspondientes a las frecuencias $2\pi j/52$, $j = 1, \dots, 6$, y γ_t^{im} asociadas a las frecuencias $\pi/2$ y π .

También en el caso de datos semanales, el número de semanas por año varía entre 52 y 53 y la misma semana no recoge siempre el mismo período del año. Además, determinadas fiestas como la Semana Santa tampoco caen siempre en el mismo período del año. Una forma de abordar este problema consiste en considerar que el componente estacional es el resultado de combinar dos componentes, es decir,

$$\gamma_t = \gamma_t^p + \gamma_t^f,$$

Donde γ_t^p representa el componente periódico, que es una función del número de días del año que han transcurrido hasta que se registra la observación, y γ_t^f el componente calendario, que recoge los efectos asociados a días festivos que siempre caen el mismo día de la semana pero en diferentes fechas en los distintos años.

Cuando se dispone de observaciones diarias se puede incorporar al modelo un componente θ_t para recoger el comportamiento atribuible al día de la semana al que pertenece la observación registrada en el momento t . Si se denota por w el número de tipos diferentes de días en una semana y por k_j el número de días del tipo j -ésimo en la semana, para $j=1,\dots,w$, el efecto asociado con el j -ésimo tipo de día es $\theta_{j,t}$ donde

$$\theta_{j,t} = \theta_{j,t-1} + \chi_{j,t}, j=1,\dots,w-1,$$

Siendo $\chi_{j,t}$ un término de perturbación de media cero, varianza $\sigma_\chi^2(1 - k_j^2/k)$, $j=1,\dots, w-1$, con $K=\sum_{j=1}^w k_j^2$, y covarianzas entre los términos de perturbación $\chi_{i,t}$ y $\chi_{h,t}$ igual a $-\sigma_\chi^2 k_j k_h / K$, $h, j = 1, \dots, w - 1, j \neq h$. Aunque cada uno de los w efectos diarios, $\theta_{j,t}$, evoluciona en el tiempo, el valor del componente diario en un instante dado del tiempo, θ_t , viene dado por el valor del efecto asociado al día de la semana al que pertenece la observación registrada en el instante t , es decir, $\theta_t = \theta_{j,t}$, $t = 1, \dots, T$, donde j es el tipo de día al que pertenece la observación.

Como el modelo se formula para los primeros $w - 1$ efectos diarios, el efecto para el tipo de día w se obtiene a partir de la condición de que, en cualquier momento del tiempo, la suma de los efectos diarios correspondientes a los diferentes días de una semana es igual a cero, es decir,

$$\theta_{w,t} = -k_w^{-1} \sum_{j=1}^{w-1} k_j \theta_{j,t}$$

El ámbito de aplicación de este último modelo no está limitado a los efectos diarios, sino que puede ser interpretado como una generalización del modelo formulado para el componente estacional en la que se admite que no todas las estaciones sean igualmente ponderadas.

Por último, si aún siendo el periodo de las observaciones inferior a la anual, el análisis revela la ausencia de variaciones estacionales o se da el caso de que los datos han sido ajustados estacionalmente, se puede plantear un modelo estructural sin componente estacional.

Combinando la tendencia bajo la formulación del modelo de tendencia lineal local y el componente estacional bajo cualquiera de las dos formulaciones estocásticas básicas, es decir, sin considerar posibles efectos intermensuales ni de calendario, se obtiene el *modelo estructural básico (MEB)*. Si se elige la formulación trigonométrica de la estacionalidad, el *MEB* queda especificado, por tanto, del modo siguiente:

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2)$$

$$\beta_t = \beta_{t-1} + \zeta_t, \quad \zeta_t \sim N(0, \sigma_\zeta^2)$$

$$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{j,t}$$

$$\begin{pmatrix} \gamma_{j,t} \\ \gamma'_{j,t} \end{pmatrix} = \begin{pmatrix} \cos \lambda_j & \text{sen } \lambda_j \\ -\text{sen } \lambda_j & \cos \lambda_j \end{pmatrix} \begin{pmatrix} \gamma_{j,t-1} \\ \gamma'_{j,t-1} \end{pmatrix} + \begin{pmatrix} \omega_{j,t} \\ \omega'_{j,t} \end{pmatrix}, \quad j=1, \dots, [s/2], \quad \lambda_j = 2\pi j/s$$

$$\omega_{j,t} \sim N(0, \sigma_\omega^2), \quad \omega'_{j,t} \sim N(0, \sigma_\omega^2), \quad \text{Cov}\{\omega_{j,t}, \omega'_{j,t}\} = 0$$

La especificación del modelo se completa suponiendo que no existe correlación entre los términos de perturbación de los

distintos componentes debido a razones de identidad de datos. Y, dado que las propiedades estadísticas del modelo dependen del conjunto completo de perturbaciones, la función de autocovarianzas de la serie observada es igual a la suma de las autocovarianzas de cada uno de los componentes.

Por último, como una extensión natural, el modelo puede incluir valores retardados de la variable dependiente así como variables explicativas exógenas, es decir,

$$y_t = \mu_t + \gamma_t + \Psi_t + \sum_{\tau=1}^p \varphi_{\tau} y_{t-\tau} + \sum_{i=1}^k \sum_{\tau=0}^{qk} \delta_{i\tau} x_{i,t-\tau} + \varepsilon_t$$

Donde $x_{i,t}$ son variables exógenas y φ_{τ} y $\delta_{i,\tau}$ son parámetros desconocidos. En el caso de que el modelo tenga incorporados retardos de la variable dependiente, por ejemplo

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \mu_t + \varepsilon_t ,$$

La tendencia puede ser redefinida como

$$\mu_t^+ = \Phi_p^{-1}(L)\mu_t ,$$

De forma que

$$y_t = \mu_t^+ + \Phi_p^{-1}(L)\varepsilon_t ,$$

Siendo $\Phi_p(L) = (1 - \phi_1 L - \dots - \phi_p L^p)$. En este caso, tanto el componente de tendencia como el irregular están sometidos a efectos autorregresivos similares.

Cuando se dispone de datos de variabilidad, las especificaciones habituales del componente estacional requieren considerar un número elevado de parámetros que pueden traer apareados problemas de estimación.

Un caso de particular interés es aquél en el que las variables explicativas son variables de intervención que permiten recoger observaciones anómalas y/o cambios estructurales. En el caso más sencillo, es decir, una observación inválida que tiene lugar en un momento determinado del tiempo, se toma mediante una variable impulso. Sin embargo, un cambio estructural caracterizado por un aumento o disminución permanente en el componente de nivel de la serie a partir de un determinado momento del tiempo puede ser capturado mediante una variable intervención tipo escalón o, alternativamente, añadiendo una variable intervención tipo impulso en la ecuación del nivel. Si, por el

contrario, se produce un cambio permanente en la pendiente del componente tendencial a partir de un determinado momento, dicha alteración puede ser tomada mediante una variable intervención que toma los valores 1, 2, 3,... a partir del momento del tiempo en el que es observado el cambio o, alternativamente, añadiendo una variable intervención tipo impulso en la ecuación de la pendiente.

2.7. PRUEBAS PARA DETERMINAR EL MODELO ADECUADO

Las perturbaciones de los modelo son variables aleatorias normales e independientes en series, con varianzas constantes, si el modelo está bien especificado, los errores de predicción o innovaciones estandarizadas $e_t = \frac{v_t}{\sqrt{F_t}}$, para $t=1, \dots, n$ (o para $t=2, \dots, n$ en el caso difuso) también deben ser normales con media 0, varianza 1 y no correlacionados. Se pueden comprobar si estas propiedades se cumplen mediante las pruebas citadas a continuación.

2.7.1. NORMALIDAD.

Los primeros cuatro momentos de los errores de predicción estandarizados están considerados como:

$$m_1 = \frac{1}{n} \sum_{t=1}^n e_t$$

$$m_i = \frac{1}{n} \sum_{t=1}^n (e_t - m_1)^i, \quad i = 2, 3, 4$$

Donde el sesgo y la curtosis están dados respectivamente por:

$$S = \frac{m_3}{\sqrt{m_2^3}}, \quad C = \frac{m_4}{m_2^2}$$

Bajo las suposiciones fundamentales del modelo se tiene que, asintóticamente:

$$S \sim N\left(0, \frac{6}{N}\right), \quad C \sim N\left(3, \frac{24}{N}\right)$$

Mediante pruebas estadísticas se pueden analizar los valores obtenidos en los diferentes modelos de los dos estadísticos para ver si son consistentes con sus densidades asintóticas. También se puede utilizar en una prueba combinada tanto S como C . El estadístico de Bowman Shenton N_{BS} dado por:

$$N_{BS} = n \left(\frac{S^2}{6} + \frac{(C-3)^2}{24} \right)$$

La cual es una distribución χ_2^2 asintótica bajo la hipótesis de normalidad en las perturbaciones.

El cuantil $\chi^2_{2(0.05)}$ es igual a 5.991. Así, si el valor del estadístico N_{BS} es mayor a 5.991 se rechaza la hipótesis nula de normalidad en los residuos.

Ho: Se acepta normalidad en los residuos del estadístico Bowman Shenton

H1: Se rechaza la hipótesis nula bajo la presencia de normalidad en los residuos

Si existen valores altos de N_{BS} , evidenciaremos la presencia de datos atípicos ya sea en el componente irregular, en el nivel o la pendiente del modelo.

2.7.2. HETEROSCEDASTICIDAD.

La Heteroscedasticidad se presenta en un modelo de regresión lineal cuando la varianza de las perturbaciones no es una constante a lo largo de las observaciones, esta prueba consiste en comparar la suma de los cuadrados de dos subconjuntos no traslapados de la muestra, por medio del siguiente estadístico:

$$H(m) = \frac{\sum_{t=n-m+1}^n e_t^2}{\sum_{t=1}^m e_t^2}$$

el cual presenta una distribución $F_{m,m}$ para un entero positivo m , bajo la hipótesis nula de homoscedasticidad (Ho: Se acepta la hipótesis nula bajo la presencia de Homoscedasticidad). Un valor elevado de H indica que la varianza de los residuos se incrementa a lo largo del tiempo.

2.7.3. AUTOCORRELACIÓN.

Un modelo es adecuado cuando los errores de predicción son seriamente no correlacionados, de aquí que el correlograma de los residuos presentará una evidencia de ello, es decir mostrará correlaciones insignificantes. El estadístico apropiado para probar la existencia o ausencia de correlación serial es el estadístico de Ljung-Box:

$$Q(k, p) = n(n+2) \sum_{j=1}^k \frac{c_j^2}{n-j}$$

Donde k es un entero positivo y c_j es la j -ésima correlación de los residuos:

$$c_j = \frac{1}{n} \sum_{t=j+1}^n \frac{(e_t - m_1)(e_{t-j} - m_1)}{m_2}$$

El estadístico de Ljung-Box permite probar si las primeras k autocorrelaciones son iguales a cero. Bajo la hipótesis nula de no correlación de los residuos (H_0 : Se acepta la hipótesis nula de no correlación de los residuos) Q se distribuye asintóticamente como una variable aleatoria χ_p^2 donde $p=k-q+1$, y q es el número de hiperparámetros.

También tenemos el estadístico para detectar la presencia de auto correlación, Durbin Watson, este permite probar si los residuos presentan auto correlación de primer orden y está dado por:

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \cong 2(1 - r(1))$$

donde:

$r(1)$ es la auto correlación muestral de primer orden.

Si el modelo está bien especificado, es decir, si los residuos no presentan auto correlación de primer orden, el estadístico DW se distribuye aproximadamente como una variable aleatoria $N\left(2, \frac{4}{n}\right)$

2.7.4. PEV (Prediction Error Variance).

PEV (σ_p^2): Es la varianza del error de predicción. Se define como la varianza de los errores de predicción o innovaciones del modelo en el período muestral. La PEV indica que tan buena es la bondad de ajuste de un modelo comparado con otro. Un modelo será preferido a otro alternativo si su PEV es menor que la del modelo alterno.

2.7.5. DETECCIÓN DE DATOS ATÍPICOS Y RUPTURAS ESTRUCTURALES.

Los residuos alisados estandarizados están dados por:

$$u_t^* = \frac{\hat{\varepsilon}_t}{\sqrt{V[\hat{\varepsilon}_t]}}, \quad r_t^* = \frac{\hat{\eta}_t}{\sqrt{V[\hat{\eta}_t]}} = \frac{r_t}{\sqrt{N_t}}, \quad t = 1, \dots, n$$

Estos residuos que son estimadores de las perturbaciones, sirven para detectar la presencia de datos atípicos. La existencia de una

observación anómala (es decir bastante grande en valor absoluto) generará un valor grande de $\hat{\varepsilon}_t$, y un cambio brusco en el nivel generará; así mismo, un valor grande del residuo $\hat{\eta}_t$, si se establece un nivel de confianza, por ejemplo un 95%, en presencia de un dato atípico se tiene:

$$|u_{t_0}^*| = \left| \frac{\hat{\varepsilon}_{t_0}}{\sqrt{V[\hat{\varepsilon}_{t_0}]}} \right| > 1.96$$

y en presencia de una ruptura estructural:

$$|r_{t_0}^*| = \left| \frac{\hat{\eta}_{t_0}}{\sqrt{V[\hat{\eta}_{t_0}]}} \right| = \left| \frac{r_{t_0}}{\sqrt{N_{t_0}}} \right| > 1.96$$

Los gráficos de los residuos estandarizados para las perturbaciones de nivel y pendiente permiten detectar irregularidades estadísticas y eventos particulares como rupturas de nivel o pendiente.

Al construir un modelo de componentes no observables es importante tomar en cuenta lo siguiente:

1. La estimación de los hiperparámetros permite conocer si los componentes evolucionan de forma aleatoria. Si el valor de un determinado hiperparámetro es cero, significa que el correspondiente comportamiento es determinista.

2. Al aplicar el filtro de Kalman se obtiene dos series: una filtrada y otra alisada, según Durban, Koopman y otros autores, es la serie alisada la que se debe utilizar para estimar los componentes no observables, puesto que el alisamiento toma en cuenta todos los datos de la muestra.
3. La tendencia de la serie es el componente que al extrapolarse indica el movimiento futuro a largo plazo, por esta razón es importante que no contenga ningún comportamiento estacional o cíclico.

CAPÍTULO III

3. APLICACIÓN ESTADÍSTICA

3.1. INTRODUCCIÓN.

A continuación se detalla una estimación de los componentes no observables de la serie de tiempo correspondiente a la medición de la variación del número de personas en ser atendidos, en Consulta General y Odontología⁽¹⁷⁾; los datos a ser analizados son desde enero del 2007 a diciembre del 2009; así mismo los datos del 2010 son importantes para realizar pruebas en los modelos establecidos.

Estas series serán designadas con sus variables respectivas siendo estas **Y**, **X**. Se utilizará el enfoque de los modelos estructurales, descomponiendo dicha serie en elementos tales como tendencia, estacionalidad, ciclo y componente irregular, elementos que permitirán hacer un análisis de la situación actual de la cantidad de pacientes atendidos mes a mes.

17 Especialidades tomadas del hospital del Día ubicado en el sur de Guayaquil

Los resultados obtenidos pueden ser utilizados por los diferentes Hospitales del Día ⁽¹⁸⁾. Además; se podrá determinar la cantidad de pacientes que se acercan a atender mes a mes en los siguientes años para poder brindar una mejor atención contratando o disminuyendo el personal de acuerdo a la demanda de pacientes que asisten a estos Hospitales.

Los análisis son realizados a través de paquetes estadísticos STAMP 4.10 DE OXMETRIC VERSION 4.10 y la aplicación Microsoft Excel 2007, además se usarán las herramientas descritas en el capítulo anterior

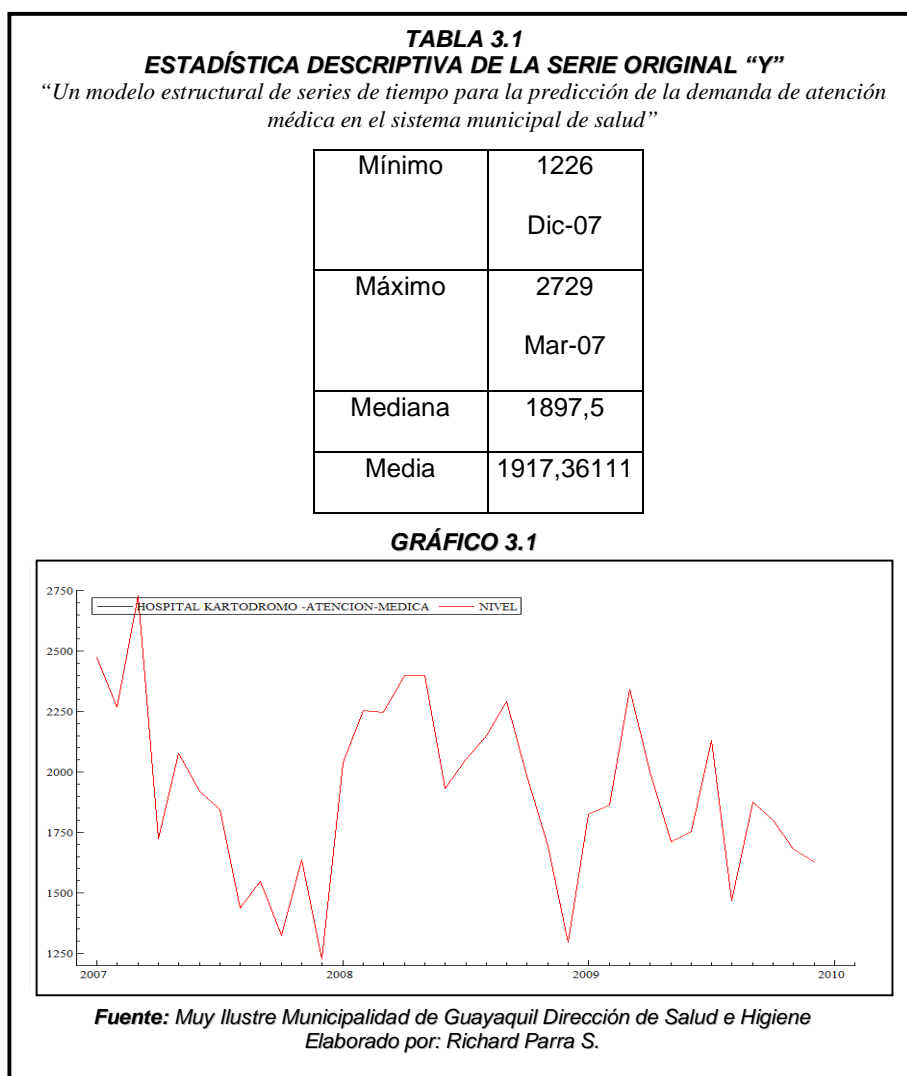
3.2. ANÁLISIS ESTADÍSTICO

En esta investigación se consideraran tres etapas, en la primera etapa se realizara un análisis exploratorio para identificar los componentes no observables con varianza cero, y los que no sean significativos para el modelo. En la segunda etapa, se realizan pruebas estadísticas con el objeto de depurar el modelo obtenido en la primera etapa. Dichas pruebas se realizan en base a los residuos estandarizados del modelo, en busca de desviaciones de la normalidad, de la presencia de auto correlación y datos atípicos. En la última etapa, se utiliza el modelo obtenido en las etapas anteriores para generar proyecciones a corto y mediano plazo.

18 Hospitales del Día y Clínicas Móviles que maneja la Municipalidad

3.2.1. VARIABLE CONSULTA GENERAL

La serie original, objeto del análisis es la correspondiente a la variable “Y” que son las personas que se atienden en Consulta General. El período de muestra abarca desde enero de 2007 hasta diciembre de 2009, tal como se observa en el gráfico 3.1.



Tal como observamos en la tabla 3.1 el número de personas que asiste al Hospital del día para ser atendidos mediante consulta general, es bajo en el mes de diciembre de 2007, tomando en consideración que la etapa invernal empieza; en ese mes los insectos aparecen más de lo normal y la afectación del cambio de clima ocasiona enfermedades en algunas personas aumentando el índice de pacientes con diferentes síntomas; observaremos que el índice se incrementa mes a mes hasta llegar a su punto máximo en el mes de marzo de 2007 donde llega a 2.729 personas en ser atendidos.

3.2.1.1. PRIMERA ETAPA:

❖ MODELO 1

Se considera un modelo estructural básico; al que denominaremos modelo 1, sin estacionalidad. En este proceso analizaremos los hiperparámetros estimados en las varianzas de las perturbaciones del componente irregular del nivel. Los resultados obtenidos en este modelo por medio de las estimaciones nos denotaran las propiedades del mismo para lo que utilizaremos la serie dada de enero del 2007 hasta diciembre de 2009. Las estimaciones estadísticas obtenidas en el modelo están resumidas en la tabla 3.2.

TABLA 3.2		
ESTIMACIONES ESTADÍSTICAS DEL MODELO 1		
<i>“Un modelo estructural de series de tiempo para la predicción de la demanda de Atención médica en el sistema municipal de salud”</i>		
UC(1) MODELIZANDO HOSPITAL -ATENCION-MEDICA POR MAXIMA VEROSIMILITUD.		
The selection sample is: 2007(1) - 2009(12)		
The model is: $Y = \text{Level} + \text{Irregular}$		
Log-Likelihood is -206.761 (-2 LogL = 413.521).		
Prediction error variance is 131494		
Summary statistics		
std.error	362.62	
Normality	1.4915	
H(11)	0.62792	(0.7737)
r(1)	-0.42734	
r(3)	-0.20635	
DW	2.8351	
Q(3,2)	9.9311	(0.006974)
R^2	-0.060873	
Variances of disturbances.		
Component	Value	(q-ratio)
Level	1.3525e+005	(1.0000)
Irregular	0.00000	(0.0000)
State vector analysis at period 2009(12)		
- level is 1628 with stand.err .NaN.		
Fuente: <i>Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene</i> <i>Elaborado por: Richard Parra S.</i>		

En el modelo 1, los valores de la varianzas de las perturbaciones consideran que este modelo contiene componente irregular igual a cero además tiene un nivel determinista, es decir:

$$y_t = \mu_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \eta_t \text{ para } t = 1, \dots, n$$

Observando el valor del estadístico de Bowman-Shenton detallado en el modelo 1 descrito por la normalidad, el cual es **1.4915**; siendo este menor que **5.991** se acepta la hipótesis nula, mientras que el estadístico $H(11) = 0.62792$ con un valor p asociado igual a 0.7737 permite aceptar la hipótesis nula bajo efecto de homoscedasticidad el cual indica que la varianza de cada perturbación μ es una constante igual a σ^2 .

Luego analizando el estadístico de Ljung-Box el cual tiene un valor de 9.9311 con un valor $p=0.006974$ dado que cae fuera de la región de aceptación se rechaza la hipótesis nula de no correlación en los residuos, por lo que este modelo indica la presencia de auto correlación.

Dado los resultados en el estudio del modelo 1 por medio del estadístico de Ljung-Box se puede determinar que este modelo no es adecuado para poder realizar predicciones.

❖ MODELO 2

En el siguiente modelo estructural básico, denominado modelo 2 analizaremos los hiperparámetros estimados en las varianzas de las perturbaciones del componente irregular del nivel y la pendiente. Donde los resultados que se obtengan en este modelo por medio de las estimaciones nos ayudaran a definir si este es el adecuado. Las estimaciones estadísticas obtenidas en el modelo están resumidas en la tabla 3.3.

En el modelo 2, los valores de la varianzas de las perturbaciones consideran que este modelo contiene componente irregular igual a 44787 y nivel igual 42428 mientras que la pendiente es igual a cero el cual lo podemos denotar en el gráfico 3.3, así tenemos la ecuación:

$$y_t = \mu_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \zeta_t \text{ para } t=1, \dots, n$$

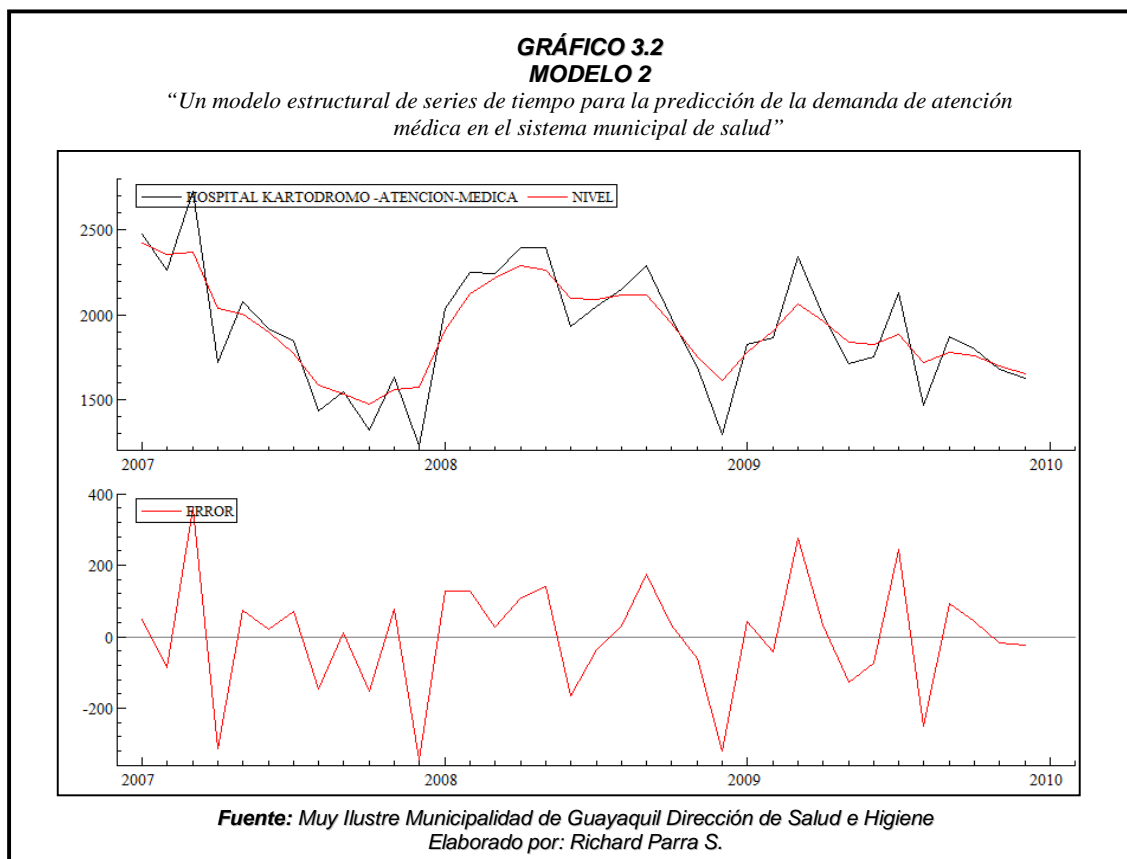
TABLA 3.3
ESTIMACIONES ESTADÍSTICAS DEL MODELO 2

“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

UC(2) MODELIZANDO HOSPITAL KARTODROMO -ATENCION-MEDICA POR MAXIMA VEROSIMILITUD.	
The selection sample is: 2007(1) - 2009(12)	
The model is: $Y = \text{Trend} + \text{Irregular}$	
Log-Likelihood is -200.357 (-2 LogL = 400.714).	
Prediction error variance is 108258	
Summary statistics	
std.error	329.03
Normality	0.70315
H(11)	0.53409 (0.8434)
r(1)	-0.069239
r(3)	-0.16088
DW	2.0700
Q(3,1)	1.3684 (0.2421)
Variances of disturbances.	
Component	Value (q-ratio)
Level	42428 (0.9473)
Slope	0.00000 (0.0000)
Irregular	44787 (1.0000)
State vector analysis at period 2009(12)	
- level is 1653.03 with stand.err 166.689.	
- slope is -22.0195 with stand.err 35.4744.	

Fuente: *Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene*
Elaborado por: Richard Parra S.

En cuanto al valor del estadístico de Bowman-Shenton detallado en el modelo 2 de la tabla 3.3 descrito por la normalidad, se observa claramente un valor de **0.70315** siendo menor a **5.991** se considera que existe normalidad en los residuos. Observando el estadístico $H(11) = 0.53409$ el cual tiene un valor p asociado igual a 0.8434 se acepta la hipótesis nula bajo efecto de homoscedasticidad el cual me indica que la varianza es constante mientras que el estadístico de Ljung-Box tiene un valor de 1.3684 con un valor $p=0.2421$ dado que cae en la región de aceptación se acepta la hipótesis nula de no correlación en los residuos con un 95% de confianza. También se puede determinar la presencia de auto correlación mediante el estadístico de Durbin Watson (2.070) el cual es cercano a 2 donde se puede concluir que no existe auto correlación de orden 1 en los residuos. Dado los resultados obtenidos en este modelo; podemos decir que es un modelo aceptable para realizar predicciones. Obsérvese el grafico 3.2, de donde se muestra un mejor ajuste en los datos..



3.2.1.2. SEGUNDA ETAPA

❖ MODELO 3

Con este modelo denominado modelo 3 se van a ajustar los datos de la serie estudiada mediante la componente estacional, para

poder determinar si este modelo es el adecuado. Los resultados estadísticos de la serie se detallan en la tabla 3.4.

TABLA 3.4		
ESTIMACIONES ESTADÍSTICAS DEL MODELO 3		
<i>“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”</i>		
UC(3) MODELIZANDO HOSPITAL KARTODROMO - ATENCION-MEDICA POR MAXIMA VEROSIMILITUD.		
The selection sample is: 2007(1) - 2009(12)		
The model is: $Y = \text{Trend} + \text{Seasonal} + \text{Irregular}$		
Log-Likelihood is -149.357 (-2 LogL = 298.714).		
Prediction error variance is 47969.1		
Summary statistics		
std.error	219.02	
Normality	0.97728	
H(7)	0.41963	(0.8627)
r(1)	-0.052497	
r(2)	-0.10052	
DW	1.9806	
Q(2,-1)	0.34872	(0.5548)
Variances of disturbances.		
Component	Value	(q-ratio)
Level	22547	(0.6809)
Slope	0.00000	(0.0000)
Seasonal	0.00000	(0.0000)
Irregular	33113	(1.0000)
State vector analysis at period 2009(12)		
- level is 2035.36 with stand.err 168.019.		
- slope is -4.46415 with stand.err 26.3576.		
Fuente: Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene		

Tal como se observa; este modelo muestra un ajuste en la estacionalidad, mientras que el estadístico de Bowman-Shenton sigue siendo el adecuado manifestando normalidad en los residuos con un valor de 0.97728, indicando que en la serie no existen rupturas estructurales que afecten su nivel ni su pendiente, el estadístico $H(7)$ sigue siendo bajo en este modelo con (0.41963) y un valor $p = 0.8627$ con una estacionalidad constante tal como se muestra en la tabla 3.4, bajo estos argumentos se prueba la homoscedasticidad de los residuos es decir la varianza no se incrementa se mantiene constante cada año. Observando el gráfico 3.3 y la tabla 3.5 observaremos el comportamiento del componente estacional determinista en este modelo el cual es constante cada año además los coeficientes de las variables ficticias asociado a cada mes del año.

El estadístico de Ljung-Box tiene un valor de 0.3487 con un valor $p=0.5548$ lo que me indica que no existe autocorrelación en los datos. En este modelo mediante el estadístico de Durbin Watson igual a (1.9806) el cual es cercano a 2 nos indica que no existe autocorrelación de orden 1 en los residuos.

GRÁFICO 3.3
ESTACIONALIDAD DE LOS DATOS MODELO 3

“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

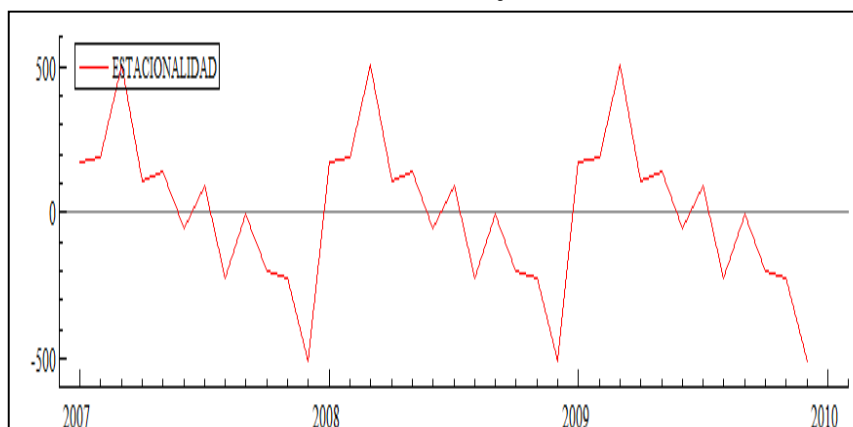


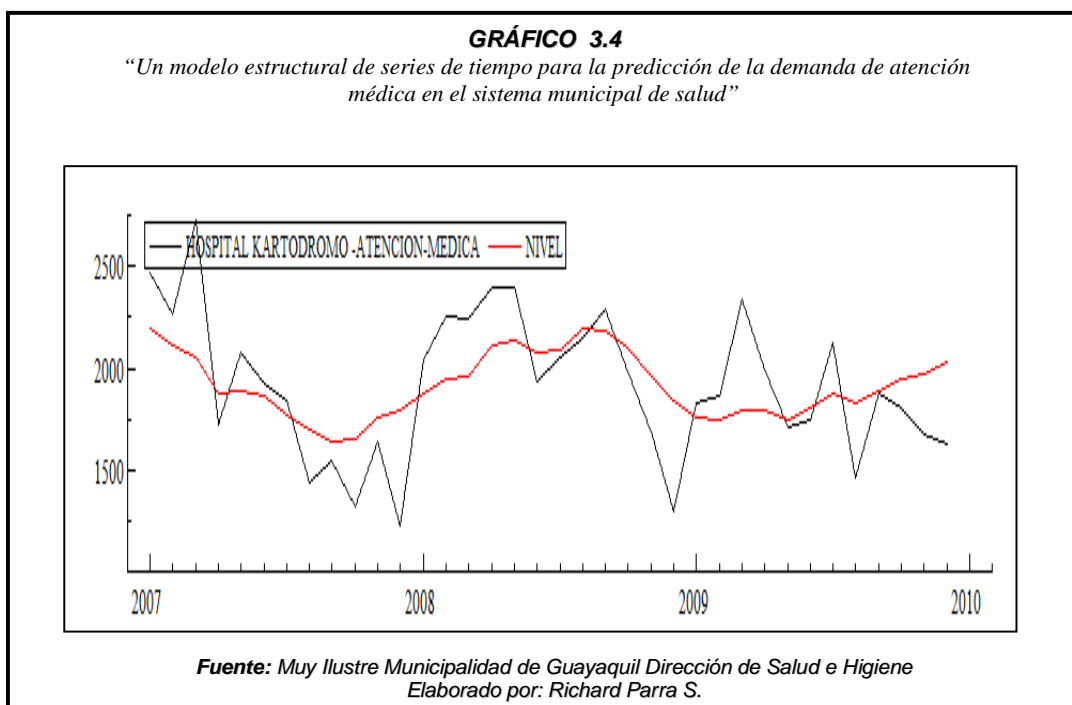
TABLA 3.5

COEFICIENTE DEL COMPONENTE ESTACIONAL CON UNA DISTRIBUCIÓN
CHI CUADRADO CON 11GL Y SU ESTADISTICO DE 29.4598

Period	Value	stand.err
Enero	170.419415	140.182509
Febrero	189.883562	137.720286
Marzo	506.347708	135.718349
Abril	112.478522	134.197299
Mayo	138.942669	133.173614
Junio	-51.593185	132.658810
Julio	95.204296	132.658810
Agosto	-225.331558	133.173614
Septiembre	-1.534077	134.197299
Octubre	-198.403264	135.718349
Noviembre	-226.605784	137.720286
Diciembre	-509.808304	140.182509

Fuente: *Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene*
Elaborado por: Richard Parra S.

En el gráfico 3.4 notamos como el nivel de la serie original empieza a suavizarse dando a notar la pendiente del modelo que oscila en forma constante.



Observemos en la tabla 3.5 el componente estacional, el cual muestra como el número de pacientes asiste al Hospital del Día, con un crecimiento desde el mes Enero, llegando a su máximo valor en el mes de Marzo, este fenómeno ocurre por la etapa invernal dado que en este periodo de tiempo las enfermedades aumentan por el cambio de clima, mientras que desde ese mes comienza a decrecer paulatinamente con unos pequeños repuntes

en los meses de Mayo y Julio hasta llegar a su valor más bajo en el mes de Diciembre.

❖ **MODELO 4**

En este modelo utilizaremos todos los componentes para su análisis: **nivel, pendiente, tendencia, estacionalidad, ciclo y componente irregular:**

$$y_t = \mu_t + \beta_t + \gamma_t + \Psi_t + \varepsilon_t$$

Donde

$$\mu_t = \mu_{t-1} + \Psi_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \zeta_t \quad \text{para } t=1, \dots, n$$

Las estimaciones para el modelo se muestran en la siguiente tabla:

TABLA 3.6
MODELO 4

“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

**UC(4) MODELIZANDO HOSPITAL KARTODROMO -ATENCION-MEDICA
POR MAXIMA VEROSIMILITUD.**

The selection sample is: 2007(1) - 2009(12)

The model is: $Y = \text{Trend} + \text{Seasonal} + \text{Cycle} + \text{Irregular}$

Log-Likelihood is -146.48 (-2 LogL = 292.959).

Prediction error variance is 25895.6

Summary statistics

std.error	160.92
Normality	1.4895
H(7)	0.35868 (0.9002)
r(1)	-0.073880
r(2)	-0.20983
DW	2.0662
Q(2,-2)	1.3482 (0.5096)

Variances of disturbances.

Component	Value	(q-ratio)
Level	0.00000	(0.0000)
Slope	0.00000	(0.0000)
Seasonal	0.00000	(0.0000)
Cycle	0.061282	(0.0000)
Irregular	39655	(1.0000)

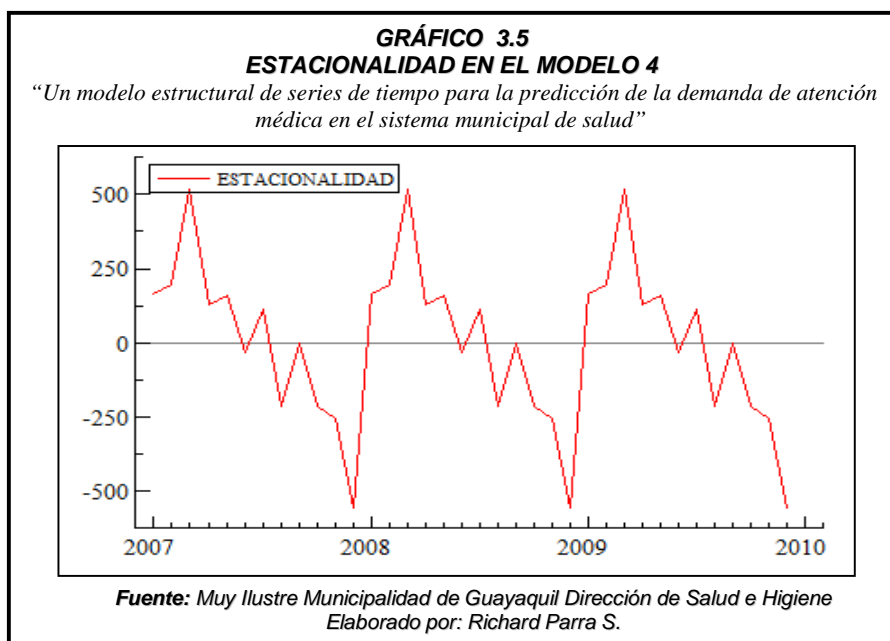
Parameters in Cycle

Variance	30641.
Period	19.398
Period in years	1.6165
Frequency	0.32392
Damping factor	1.0000
State vector analysis at period 2009(12)	
level is	1939.48 with stand.err 71.7177.
slope is	0.243858 with stand.err 3.62717.

Fuente: *Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene*
Elaborado por: Richard Parra S.

Tal como se observa en la tabla 3.6, el valor del estadísticos de Bowman-Shenton con un valor de (1.4895) y con 95 % de confianza, indica que los datos son normales mientras que el estadístico H(7) sigue manteniéndose bajo con (0.35868) y un valor $p = 0.9002$, por lo que se prueba que no existe heteroscedasticidad en los residuos es decir la varianza no se incrementa se mantiene constante.

También tenemos el estadístico de Ljung-Box con un valor de 1.3482 asociado a $p=0.5096$ dado que este cae en la región de aceptación se acepta la hipótesis nula de no correlación en los residuos.



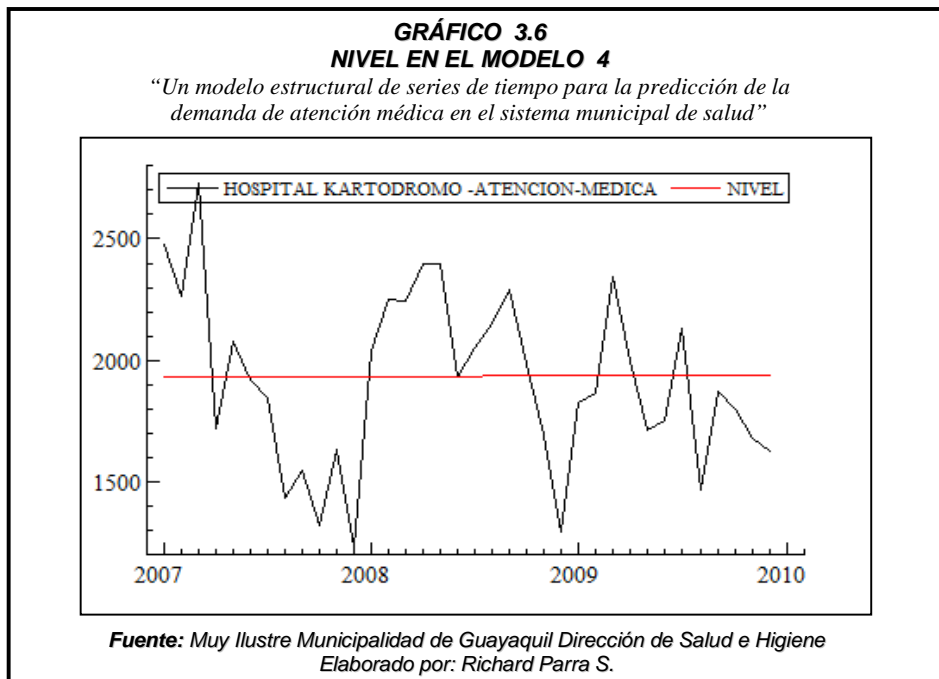
Obsérvese en la tabla 3.7 y el gráfico 3.5 del componente estacional del modelo 4 el cual muestra algo similar que el componente estacional en el modelo 3 donde tenemos como el número de pacientes asiste al Hospital del Día en una forma constante cada año, de donde se muestra un crecimiento desde el mes Enero, llegando a su máximo valor en el mes de Marzo, mientras que desde ese mes comienza a decrecer paulatinamente con unos pequeños repuntes en los meses de Mayo y Julio hasta llegar a su valor más bajo en el mes de Diciembre.

TABLA 3.7
“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

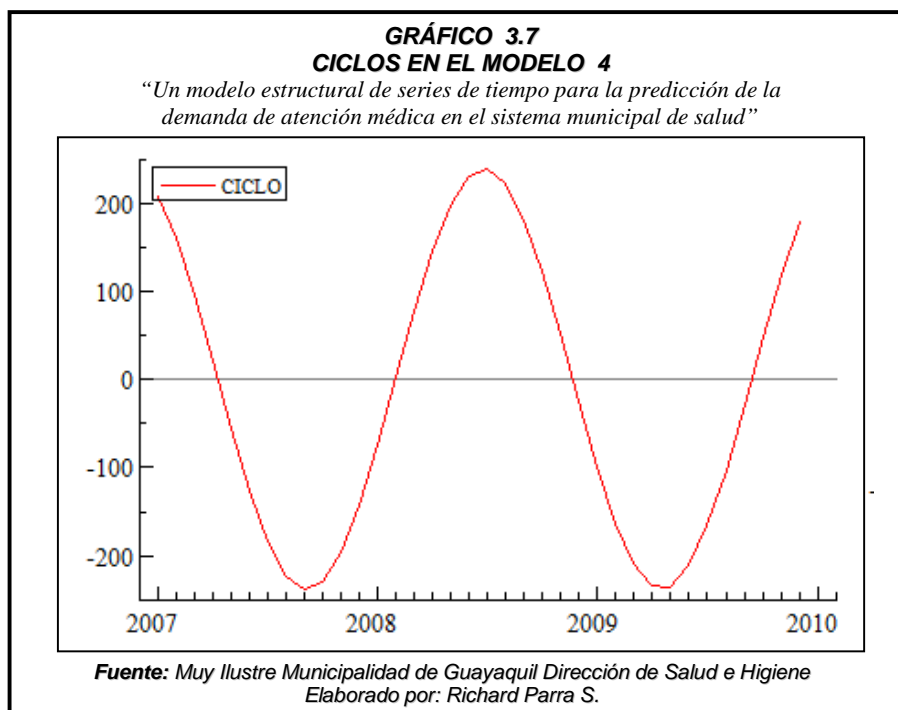
COEFICIENTE DEL COMPONENTE ESTACIONAL CON UNA DISTRIBUCION CHI CUADRADO CON 11GL Y SU ESTADISTICO DE 59.6174		
period	value	stand.err
Enero	166.811880	111.760203
Febrero	193.451219	111.172414
Marzo	517.066466	110.736900
Abril	129.090873	110.434274
Mayo	159.088479	110.245396
Junio	-31.131207	110.155121
Julio	112.242604	110.155121
Agosto	-215.590332	110.245396
Septiembre	-2.694070	110.434274
Octubre	-213.424480	110.736900
Noviembre	-257.496242	111.172414
Diciembre	-557.415190	111.760203
- amplitude of Cycle 1 is 238.255		

*Fuente: Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene
Elaborado por: Richard Parra S.*

Se observa en el gráfico 3.6, que con una desviación estándar de 71.7177 aproximándose al valor inicial calculado en la media de la tabla 3.1 que es 1917.3611 indica que éste es un buen modelo. También tenemos una pendiente igual a 0.243858 con una desviación estándar 3.62717.



En el gráfico 3.7 se observan los ciclos del modelo 4 donde en el mes de enero de 2007 existe una alta demanda de pacientes para “consulta general” disminuyendo gradualmente hasta llegar a un número mínimo de pacientes en ser atendidos en el mes de Octubre donde comienza a aumentar el número de pacientes hasta obtener nuevamente un valor máximo en los meses de Abril y Mayo de 2008. Así el componente cíclico indica una fase de oscilaciones constante el cual varía de acuerdo a la época estacional con respecto al cambio de clima.



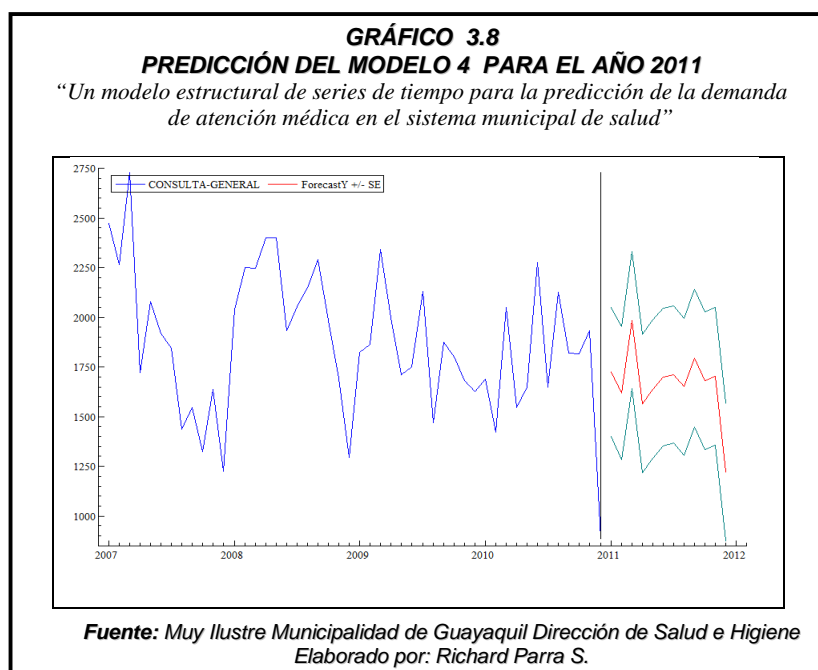
3.2.1.3. TERCERA ETAPA

❖ ANÁLISIS DE PREDICCIÓN DEL MODELO

Dado el análisis realizado en los modelos 1, 2, 3 y 4 se determina que el modelo adecuado para realizar predicciones en el Hospital del día para atención en consulta General es el modelo 4, el cual muestra una mejor aproximación a los datos reales, además de

tener un valor PEV adecuado dado que el modelo 4 tiene mejor bondad de ajuste siendo este el menor valor entre los 4 modelos.

Tal como observamos en el gráfico 3.8, las predicciones del año 2011 se ajustan a las predicciones del modelo 4 del año 2010, manteniendo la tendencia de pacientes en ser atendidos en años anteriores.



Denotemos los datos de las predicciones del año 2011 del modelo 4, detallada en la tabla 3.8.

TABLA 3.9
PREDICCIÓN DEL MODELO 4 PARA EL AÑO 2010

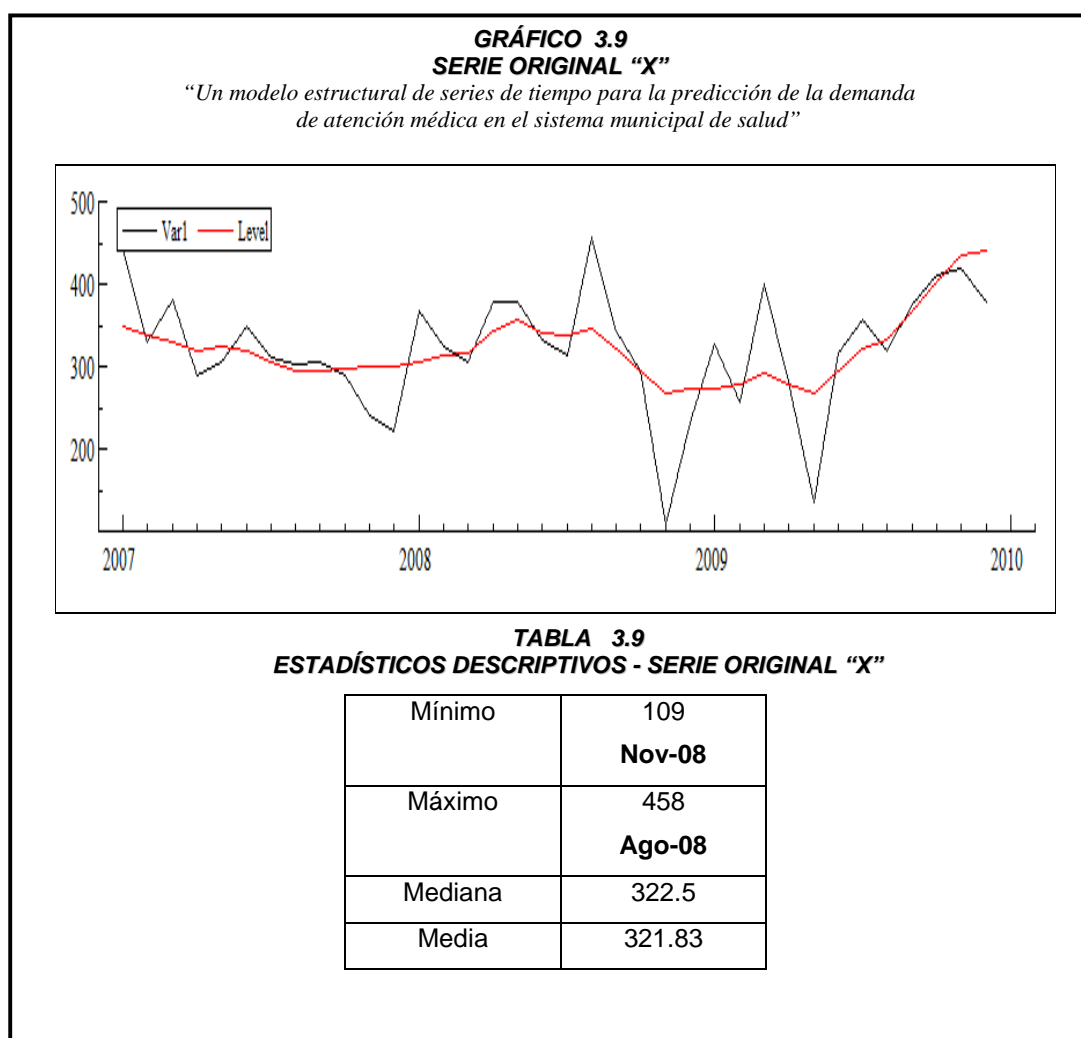
“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

PERIODO	DATO- PREDICCIÓN
Ene-11	1725
Feb-11	1619
Mar-11	1984
Abr-11	1560
May-11	1635
Jun-11	1700
Jul-11	1710
Ago-11	1652
Sep-11	1790
Oct-11	1680
Nov-11	1700
Dic-11	1220

Fuente: *Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene*
Elaborado por: Richard Parra S.

3.2.2. VARIABLE ODONTOLÓGÍA

La serie original, objeto del análisis es la correspondiente a la variable “X”, personas que se atienden en el área Odontología del Hospital Municipal. El período muestral abarca desde enero de 2007 hasta diciembre de 2009, tal como muestra en el gráfico 3.9.



Tal como observamos en la tabla 3.9 y el gráfico 3.9; el número de pacientes que asiste al Hospital para ser atendidos en el área de Odontología, es bajo en el mes de Noviembre de 2008 con 109 y alto en el mes de agosto del mismo año con un valor de 458; además, se tiene una media de alrededor de 321 pacientes atendidos mensualmente en esos años.

3.2.2.1. PRIMERA ETAPA

❖ MODELO 1

Tomando el modelo estructural básico, tal como se describe en la ecuación 2.2.2 (al que denominaremos modelo 1), donde se analiza los hiperparámetros estimados en las varianzas de las perturbaciones del componente irregular del nivel. De donde los resultados obtenidos en este modelo por medio de las estimaciones, nos denotaran las propiedades del mismo, para lo que utilizaremos la serie dada de enero del 2007 hasta diciembre de 2009. Las estimaciones estadísticas obtenidas en el modelo están resumidas en la tabla 3.10.

TABLA 3.10
ESTIMACIONES ESTADÍSTICAS DEL MODELO 1

“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

UC(1) MODELIZANDO HOSPITAL KARTODROMO-ODONTOLOGIA POR MAXIMA VEROSIMILITUD.

The selection sample is: 2007(1) - 2009(12)

The model is: $Y = \text{Trend} + \text{Irregular}$

Log-Likelihood is -150.992 (-2 LogL = 301.984).

Prediction error variance is 5904.51

Summary statistics

std.error 76.841

Normality 5.9737

H(11) 2.1112 (0.1155)

r(1) 0.041706

r(3) 0.0045305

DW 1.7652

Q(3,1) 2.9382 (0.4013)

Variances of disturbances.

Component	Value (q-ratio)
Level	1089.2 (0.3101)
Slope	0.80902 (0.0002)
Irregular	3512.5 (1.0000)

State vector analysis at period 2009(12)

- level is 386.444 with stand.err 39.5226.

- slope is 0.945421 with stand.err 6.59345.

Fuente: *Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene*
Elaborado por: Richard Parra S.

En el modelo 1, los valores de la varianzas de las perturbaciones consideran que este modelo contiene componente irregular además tiene un nivel determinista, es decir:

$$y_t = \mu_t + \varepsilon_t \quad (3.3.4)$$

$$\mu_t = \mu_{t-1} + \eta_t \text{ para } t = 1, \dots, n$$

Observando el valor del estadístico de Bowman-Shenton detallado en el modelo 1 descrito por la normalidad, el cual es **5.9737** siendo este menor que **5.991** se acepta la hipótesis nula bajo efecto de normalidad en los residuos, mientras que el estadístico $H(11) = 2.1112$ con un valor $p=0.1155$ lo cual permite aceptar la hipótesis nula indicándome la presencia de homoscedasticidad en los datos. Analizando el estadístico de Ljung-Box el cual tiene un valor de 2.9382 con un valor $p=0.4013$ dado que cae dentro de la región de aceptación se acepta la hipótesis nula de no correlación en los residuos. Dado los resultados del análisis de los estadísticos del modelo 1, se puede tomar en cuenta este modelo para predecir datos en el futuro.

❖ MODELO 2

En el siguiente modelo estructural básico al que denominaremos modelo 2, analizaremos los hiperparámetros estimados en las varianzas de las perturbaciones del componente irregular del nivel, la pendiente y estacionalidad.

Donde los resultados que se obtengan en este modelo por medio de las estimaciones nos ayudaran a definir si este es el adecuado. Las estimaciones estadísticas obtenidas en el modelo están resumidas en la tabla 3.11.

TABLA 3.11-1	
ESTIMACIONES ESTADÍSTICAS DEL MODELO 2	
<i>“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”</i>	
UC(2) MODELIZANDO HOSPITAL KARTODROMO-ODONTOLOGIA POR MAXIMA VEROSIMILITUD.	
The selection sample is: 2007(1) - 2009(12)	
The model is: $Y = \text{Trend} + \text{Seasonal} + \text{Irregular}$	
Log-Likelihood is -119.827 (-2 LogL = 239.654).	
Prediction error variance is 3802.41	
Summary statistics	
std.error	61.664
Normality	0.35619
Fuente: <i>Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene</i> <i>Elaborado por: Richard Parra S.</i>	

TABLA 3.11-2	
ESTIMACIONES ESTADÍSTICAS DEL MODELO 2	
<i>“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”</i>	
H(7)	1.8980 (0.2085)
r(1)	-0.010367
r(2)	-0.26420
DW	1.9190
Q(2,-1)	1.9140 (0.3840)
Variances of disturbances.	
Component	Value (q-ratio)
Level	1234.4 (0.4242)
Slope	5.0488 (0.0017)
Seasonal	0.00000 (0.0000)
Irregular	2909.9 (1.0000)
State vector analysis at period 2009(12)	
- level is 441.215 with stand.err 45.3405.	
- slope is 7.03021 with stand.err 9.57534.	
- joint seasonal chi2 test is 14.4831 with 11 df.	
Fuente: <i>Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene</i> <i>Elaborado por: Richard Parra S.</i>	

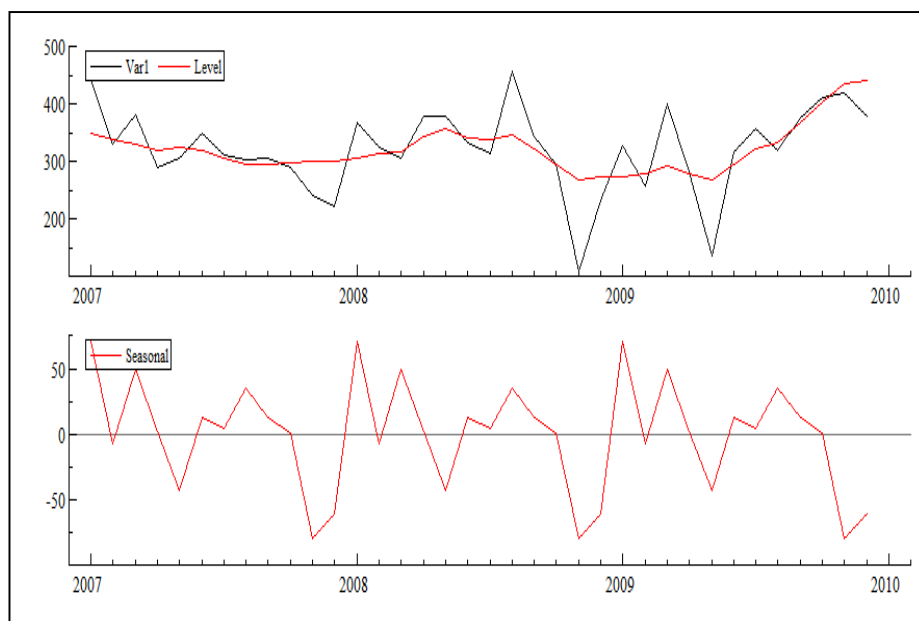
En el modelo 2, los valores de la varianzas de las perturbaciones consideran que este modelo contiene componente irregular igual a 2909.9 y nivel igual 1234.4 mientras que la pendiente es igual a 5.0488 el cual lo podemos denotar en el gráfico 3.10, así tenemos la ecuación:

$$y_t = \mu_t + \varepsilon_t$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t \quad \text{para } t=1, \dots, n$$

En cuanto al valor del estadístico de Bowman-Shenton detallado en el modelo 2 de la tabla 3.11 descrito por la normalidad, el cual es 0.35619 siendo menor a **5.991** se considera que existe normalidad en los residuos. Observando el estadístico $H(7) = 1.8980$ el cual tiene un valor p asociado igual a 0.2085 este permite aceptar la hipótesis nula bajo efecto de homoscedasticidad el cual me indica que la varianza es constante mientras que el estadístico de Ljung-Box tiene un valor de 1.9140 con un valor $p=0.3840$ dado que cae en la región de aceptación se acepta la hipótesis nula de no correlación en los residuos.

Las pruebas estadísticas de este análisis se lo han realizado con una confianza del 95%.

GRÁFICO 3.10
ESTIMACIONES ESTADÍSTICAS DEL MODELO 2
“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”



Fuente: Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene
Elaborado por: Richard Parra S.

Observando el gráfico 3.10 podemos notar como el componente estacional se ajusta cada año en forma similar, haciendo que cada año los pacientes asistan en forma constante.

3.2.2.2. SEGUNDA ETAPA

❖ MODELO 3

En este modelo utilizaremos todos los componentes para su análisis: **nivel, pendiente, tendencia, estacionalidad, ciclo y componente irregular:**

$$y_t = \mu_t + \beta_t + \gamma_t + \Psi_t + \varepsilon_t$$

Donde

$$\mu_t = \mu_{t-1} + \Psi_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \zeta_t \quad \text{para } t=1, \dots, n$$

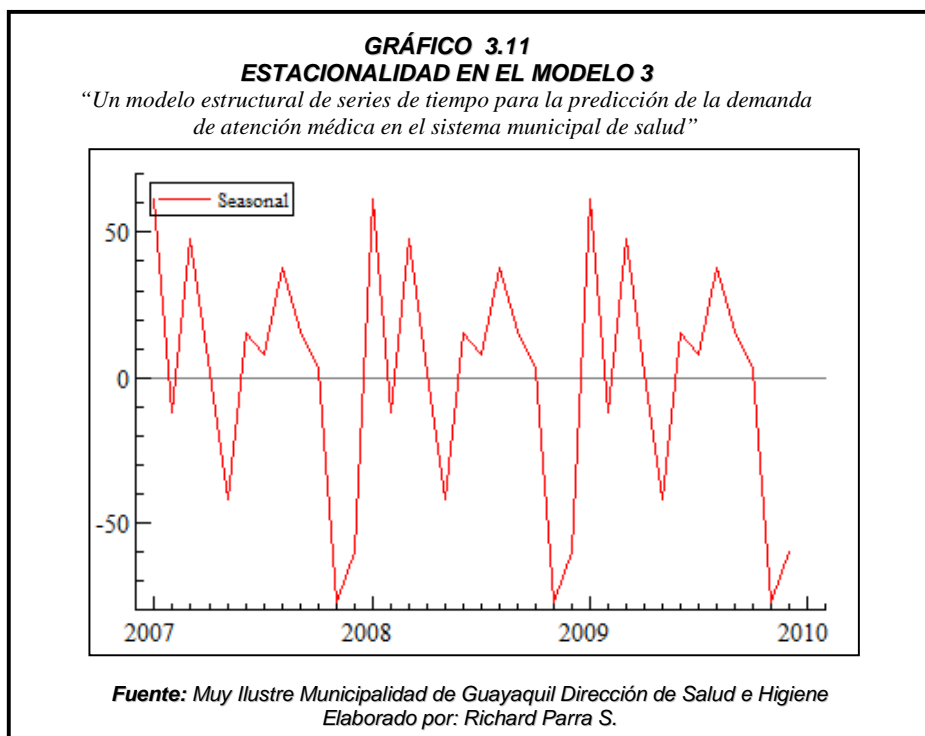
Tal como se observa en la tabla 3.12, el valor del estadísticos de Bowman-Shenton el cual con un valor de (1.0476) con 95 % de confianza indica que los residuos son normales mientras que el estadístico H(7) es igual a 1.8391 con un valor p =0.22001, por lo que se prueba que existe homoscedasticidad en los residuos es decir la varianza no se incrementa, se mantiene constante.

También tenemos el estadístico de Ljung-Box con un valor de 2.197 asociado a p=0.3333 dado que este cae en la región de

aceptación se acepta la hipótesis nula de no correlación en los residuos.

Las estimaciones para el modelo se muestran en la siguiente tabla:

TABLA 3.12	
ESTIMACIONES ESTADÍSTICAS DEL MODELO 3	
<i>“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”</i>	
UC(4) MODELIZANDO HOSPITAL KARTODROMO	
ODONTOLOGIA POR MAXIMA VEROSIMILITUD.	
The selection sample is: 2007(1) - 2009(12)	
The model is: $Y = \text{Trend} + \text{Seasonal} + \text{Irregular} + \text{Cycle 1}$	
Log-Likelihood is -119.271 (-2 LogL = 238.541).	
Prediction error variance is 3349.79	
Summary statistics	
std.error	57.877
Normality	1.0476
H(7)	1.8391 (0.22001)
r(1)	0.029683
r(2)	-0.28178
DW	1.8351
Q(2,-2)	2.1970 (0.3333)
Variances of disturbances.	
Component	Value (q-ratio)
Level	2.2821 (0.0006)
Slope	6.1103 (0.0017)
Seasonal	0.00000 (0.0000)
Cycle	22.903 (0.0064)
Irregular	3555 (1.0000)
Parameters in Cycle	
Variance	1293.7
Period	19.189
Period in years	1.5991
Frequency	0.32743
Damping factor	0.99111
State vector analysis at period 2009(12)	
- level is 382.402 with stand.err 36.7004.	
- slope is 8.75828 with stand.err 7.10283.	
Fuente: <i>Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene</i>	



Obsérvese en la tabla 3.13 y el gráfico 3.11. El componente estacional del modelo 3 muestra algo similar que el componente estacional en el modelo 2 así tenemos como el número de pacientes del área de Odontología del Hospital del Día asiste en forma constante cada año, además podemos notar como en los meses de febrero, mayo, noviembre y diciembre baja el número de pacientes en ser atendidos, para luego en los meses de enero, marzo y agosto aumentar el número de pacientes.

TABLA 3.13
COMPONENTE ESTACIONAL

“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

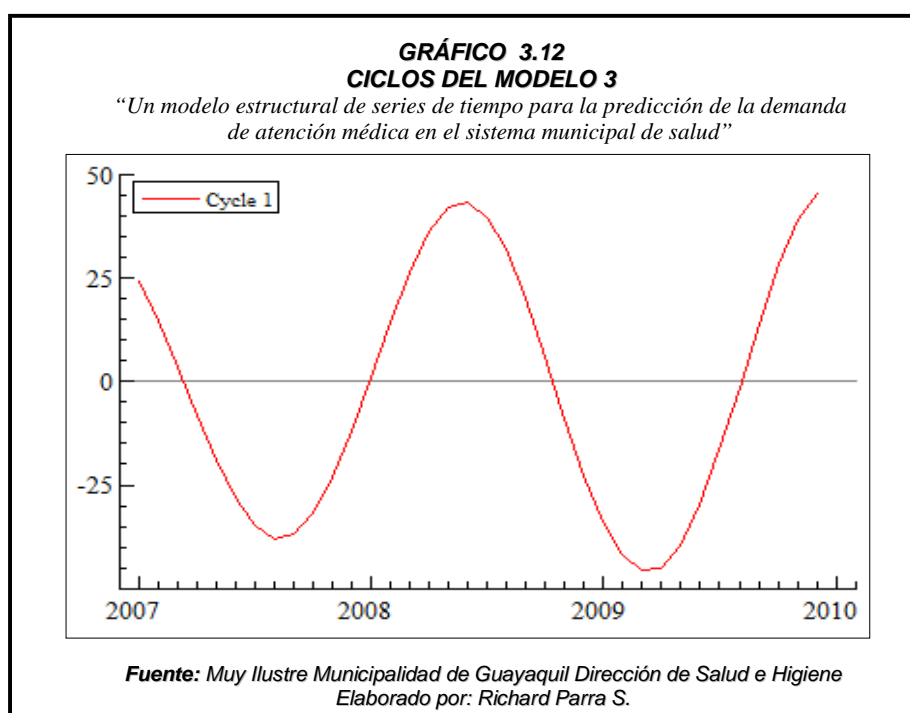
COEFICIENTE DEL COMPONENTE ESTACIONAL CON UNA DISTRIBUCION CHI CUADRADO CON 11gl Y SU ESTADISTICO DE 14.8601		
period	value	stand.err
Enero	61.697355	34.864984
Febrero	-11.895756	34.329148
Marzo	47.435511	33.959850
Abril	1.947962	33.723478
Mayo	-41.720001	33.587640
Junio	15.452661	33.526768
Julio	7.866209	33.526768
Agosto	37.626068	33.587640
Septiembre	15.865609	33.723478
Octubre	2.733495	33.959850
Noviembre	-77.286579	34.329148
Diciembre	-59.722535	34.864984

- amplitude of Cycle 1 is 47.0702

*Fuente: Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene
Elaborado por: Richard Parra S.*

En el gráfico 3.12 se observan los ciclos del modelo 3 donde en el mes de enero de 2007 existe una alta demanda de pacientes

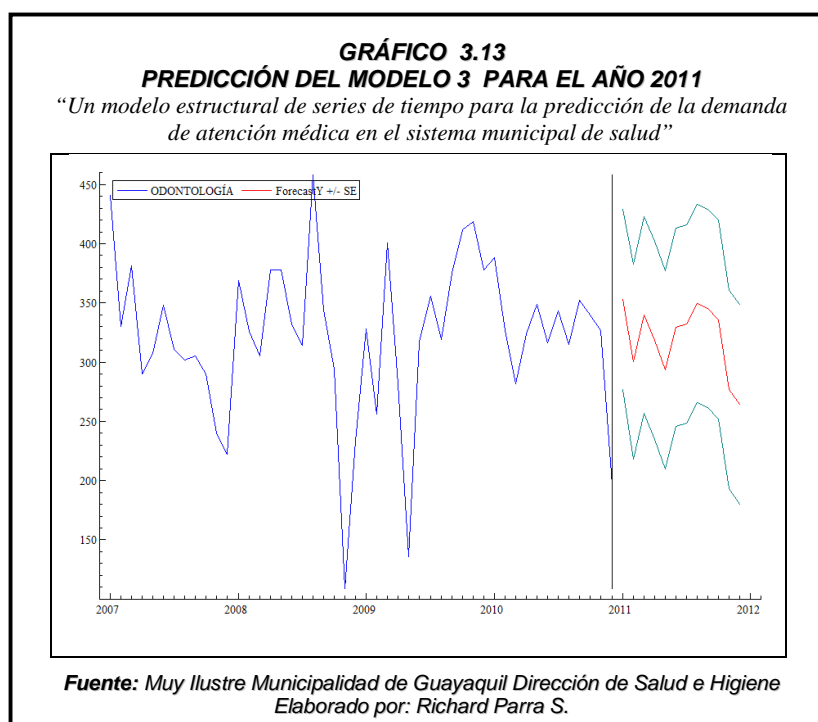
en el área de “Odontología” disminuyendo gradualmente hasta llegar a un número mínimo de pacientes en ser atendidos en el mes de Mayo y Noviembre donde comienza a aumentar el número de pacientes hasta obtener nuevamente un valor máximo en los meses de Enero y Marzo de 2008. Así el componente cíclico indica una fase de oscilaciones constante cada año.



3.2.2.3. TERCERA ETAPA

❖ ANÁLISIS DE PREDICCIÓN DEL MODELO

Según los análisis realizados en las predicciones de los modelos 1, 2 y 3 se ha logrado establecer que el modelo adecuado para realizar predicciones en el Hospital Municipal para atención en el área de Odontología es el del modelo 3 el cual muestra una mejor aproximación a los datos reales, además el PEV es el adecuado en este modelo dado que tiene su bondad de ajuste idóneo.



Tal como observamos en el gráfico 3.13, las predicciones del año 2011 se ajustan a las predicciones del modelo 4 del año 2010, manteniendo la tendencia de pacientes en ser atendidos en años anteriores.

Denotemos los datos de las predicciones del año 2011 del modelo 4, detallada en la tabla 3.14.

TABLA 3.14
PREDICCIÓN DEL MODELO 3 PARA EL AÑO 2011
“Un modelo estructural de series de tiempo para la predicción de la demanda de atención médica en el sistema municipal de salud”

PERIODO	DATO-PREDICCIÓN
Ene-11	352
Feb-11	294
Mar-11	332
Abr-11	310
May-11	285
Jun-11	321
Jul-11	325
Ago-11	341
Sep-11	332
Oct-11	327
Nov-11	270
Dic-11	256

Fuente: *Muy Ilustre Municipalidad de Guayaquil Dirección de Salud e Higiene*
Elaborado por: Richard Parra S.

CAPÍTULO IV

4. CONCLUSIONES Y RECOMENDACIONES

En el presente capítulo se dará a conocer las conclusiones del estudio de las variables (consulta general y odontología) que se usaron para poder determinar el modelo estructural que mejor se ajustó para predecir la demanda de atención médica en el sistema municipal de salud.

4.1. Variable de Consulta General.

Para determinar el modelo adecuado se analizaron los siguientes parámetros; estadísticos de Bowman - Shenton con un valor de (0.97728) y con 5 % de significancia, el cual indica que los datos son normales mientras que el estadístico H(7) tiene un valor bajo con (0.41963) y un valor $p = 0.8627$, por lo que se prueba que no

existe heteroscedasticidad en los residuos es decir la varianza no se incrementa se mantiene constante.

También tenemos el estadístico de Ljung-Box con un valor de 0.3487 asociado a $p=0.5548$ dado que este cae en la región de aceptación con un nivel de confianza del 95% se acepta la hipótesis nula de no correlación en los residuos.

En el mes de enero de 2007 existe una alta demanda de pacientes para “consulta general” disminuyendo gradualmente hasta llegar a un número mínimo de pacientes en ser atendidos en el mes de Octubre donde comienza a aumentar el número de pacientes hasta obtener nuevamente un valor máximo en los meses de Abril y Mayo de 2008. Así el componente cíclico indica una fase de oscilaciones constante el cual varía de acuerdo a la época estacional con respecto al cambio de clima en la ciudad de Guayaquil.

Nótese que las predicciones del modelo, con proyecciones del 2011 se ajustan a los datos de años anteriores manteniendo la tendencia.

Dado al análisis realizado se determina que el modelo adecuado para realizar predicciones en el Hospital del día para atención en consulta General es este modelo, el cual muestra una mejor aproximación a los datos reales, además de tener un valor PEV adecuado dado que el modelo 4 tiene mejor bondad de ajuste siendo este, el menor valor entre los 4 modelos.

4.2. Variable de Odontología

Para determinar que este modelo es el adecuado se analizan los valores de los estadísticos de Bowman-Shenton el cual muestra un valor de (1.0476) con 95 % de confianza indicando que los residuos son normales, mientras que el estadístico $H(7)$ es igual a 1.8391 con un valor $p = 0.22001$, por lo que se prueba que existe homoscedasticidad en los residuos es decir la varianza no se incrementa, se mantiene constante.

También tenemos el estadístico de Ljung-Box con un valor de 2.197 asociado a $p = 0.3333$ dado que este cae en la región de aceptación se acepta la hipótesis nula de no correlación en los residuos.

En el mes de enero de 2007 existe una alta demanda de pacientes para el área de "Odontología" disminuyendo gradualmente hasta llegar a un número mínimo de pacientes en ser atendidos en el mes de Mayo y Noviembre para lo cual desde este mes empieza a aumentar el número de pacientes

hasta obtener nuevamente un valor máximo en los meses de Enero y Marzo de 2008. Así el componente cíclico indica una fase de oscilaciones constante cada año.

Las predicciones del modelo, con proyecciones del año 2011 se ajustan en su totalidad a los datos de los años anteriores manteniendo la tendencia en sus datos.

Según los análisis realizados en las predicciones de los modelos 1, 2 y 3 se ha logrado establecer que el modelo adecuado para realizar predicciones en el Hospital Municipal para atención en el área de Odontología es el del modelo 3 el cual muestra una mejor aproximación a los datos reales, además el PEV del modelo es bajo en comparación a los otras predicciones.

4.3. RECOMENDACIONES

Para poder llevar a cabo un estudio adecuado en el que se pueda predecir la demanda de pacientes que se hacen atender en un centro de Salud; se recomienda los siguientes puntos:

- Determinar la Variable a estudiar con datos cuantitativos (numéricos).
- Determinar el software adecuado para el análisis, en este caso STAMP 4.10 DE OXMETRIC VERSION 4.10, el cual nos ayuda a determinar los estadísticos de las pruebas para tomar un buen modelo.
- Tomar datos de por lo menos 4 años consecutivos, mensuales para un mejor análisis.
- De los 4 años que se toman, realizar el estudio con los 3 primeros años, para determinar un modelo de series temporales, el cual se ajustara a los datos reales del 4to año. El modelo que mejor se aproxime a los datos reales se lo tomara para predecir datos futuros.

- Actualizar el modelo de predicción cada año, dado que los datos evolucionan en el tiempo y este puede variar dependiendo de la demanda de pacientes que se hacen atender en años posteriores al actual.
- Finalmente, determinar el modelo adecuado, considerando todas las pruebas estadísticas que se dan en el análisis.

REFERENCIAS BIBLIOGRÁFICAS

[1]. Dirección de Salud e Higiene Municipal - Revista Más Salud 2000 -2006.

[2]. Hamilton, J. D. Time series analysis. Princeton, New Jersey: Princeton University Press. 1994

[3]. Harvey A. Forecasting, structural time series models and the Kalman filter. Cambridge University Press. Reino Unido. 2003.
Harvey A. y Shephard N. Structural Time Series Models. Handbook of Statistics, Vol. 11. Pág. 261-302. 1993 .

[4]. Harvey A. y Koopman S. J. "Diagnostic Checking of Unobserved-Components Time Series Models," Journal of Business & Economic Statistics, American Statistical Association, vol. 10(4), pages 377-89, October 1992.

- [5].Koopman S.J. y J. Time series analysis by state space models. Oxford University Press. Reino Unido. 2001
- [6].Martín, G. “Modelos estructurales en el contexto de las series temporales económicas” www.fceye.ull.es/invest/docum/ull-ulpgc/2002-04.pdf - (Documento de trabajo 2002 – 2004),
- [7].Mendenhall, W., Scheaffe, R., Wackerly, Dennis. “Estadística Matemática con Aplicaciones”, Grupo Editorial Iberoamérica, México D.F., México.
- [8].Ramírez, J. Análisis Estructural del IPC mediante el Filtro de Kalman, Revista Ecuatoriana de Estadística del Instituto Nacional de Estadísticas y Censos No 2 de Junio de 2010 Pág. 5-22.
- [9].Tovar J. Structural Time Series Models and the Kalman Filter: a concise review. Faculty of Economics and Politics, University of Cambridge. Reino Unido. 2009 .