



UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES
ACATLAN



APLICACIÓN DE LAS SERIES DE TIEMPO A LA CREACIÓN
DE UN MODELO DE PRONOSTICO, EL CUAL PODRÁ SER
UTILIZADO PARA MONITOREAR EL CONTROL DEL
PORCENTAJE DE POBLACIÓN INFECTADA DE S.I.D.A.
EN MÉXICO

T E S I S

QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN MATEMÁTICAS APLICADAS
Y COMPUTACIÓN

P R E S E N T A N

CARLOS ALBERTO ÁLVAREZ GARCÍA
HÉCTOR GABRIEL RIVERA VARGAS

ASESORA:

MTRA. MARÍA DEL CARMEN GONZÁLEZ VIDEGARAY

JUNIO, 2004





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ESTA TESIS NO SALE
DE LA BIBLIOTECA

OBJETIVO GENERAL

Crear un modelo de pronóstico para las estadísticas obtenidas mediante el uso de series de tiempo, el cuál podrá ser utilizado para monitorear el control del porcentaje de población infectada de S. I. D. A. en México.

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: Rivera Vargas
Héctor Gabriel

FECHA: 29-junio-04

FIRMA: H. Gabriel Rivera V.

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional.

NOMBRE: Alvarez García
Carlos Alberto

FECHA: 29 de junio de 2004

FIRMA: Alvarez García Carlos A.

A DIOS:

Por haberme dado la oportunidad de existir.

A MIS PADRES, REYNALDA Y RAFAEL:

Por el amor, el apoyo y la confianza que siempre me han manifestado y por ser un ejemplo a seguir. Gracias, los quiero mucho.

**A MIS HERMANOS, JOSÉ EUGENIO,
LAURA ANGÉLICA, LUIS ALFREDO Y
FERNANDO:**

Por su amor, cariño, apoyo y por todos los momentos que hemos compartido.

A GUADALUPE CHEPETLA:

Por su apoyo y por sus valiosos consejos que me ha dado de manera incondicional.

A MANUEL RAMÍREZ:

Por ser uno de mis mejores amigos.

**A LA MAESTRA
MARICARMEN:**

Por ser un ejemplo a seguir como profesionista y como ser humano.

Sinceramente
HÉCTOR GABRIEL

AL TODOPODEROSO

A MIS PADRES, CLEMENTINA y EMILIO:

Por su gran paciencia, comprensión y apoyo. Les debo todo a ustedes y gracias por haber sido lo que son; porque siempre hemos estado unidos en los momentos buenos y malos apoyándonos. Los quiero mucho.

**A MI HERMANO,
JOSÉ LUIS,** quien siempre me ha demostrado ser eso, mi hermano. Gracias.

A ERÉNDIRA (Topita):
Porque hemos empezado a correr juntos las colinas; muy pronto correremos juntos también por la vida. ¡Gracias gran mujer por tu apoyo!

**A LA MAESTRA
MARICARMEN:**
Porque transmite demasiada confianza, que yo necesité; ella ha hecho posible este trabajo y seguirá siendo una guía para mi. MUCHAS GRACIAS MAESTRA.

Sinceramente
CARLOS ALBERTO

ÍNDICE

INTRODUCCIÓN.....	7
CAPÍTULO I SERIES DE TIEMPO.....	9
1. ¿Qué es una serie de tiempo?.....	9
1.1 Series de tiempo vistas como procesos estocásticos.....	10
2. Objetivos de análisis de una serie de tiempo.....	11
2.1 Elementos Estadísticos en el análisis de series de tiempo.....	12
3. Componentes de una serie de tiempo.....	12
3.1 Variación cíclica.....	13
3.2 Tendencia.....	14
3.3 Variación estacional.....	15
3.4 Fluctuación irregular.....	18
4. Metodología de BOX-JENKINS.....	19
4.1 Identificación del modelo.....	22
4.2 Estimación de parámetros.....	26
4.3 Diagnóstico de la adecuación del modelo.....	28
4.4 Pronósticos.....	30
CAPÍTULO II RECOPIACIÓN Y ANÁLISIS DE DATOS.....	33
1. Datos muestrales.....	33
1.2 Casos de S. I. D. A. por mes en México notificados del año 1990 a 1999.....	34
1.3 Casos de S. I. D. A. por mes en México notificados en el año 2001.....	39
2. Fuente de Información.....	40
3. Gráfica de la serie.....	40
4. Identificación de los posibles modelos.....	42
CAPÍTULO III EL SIDA.....	45
1. ¿Qué es el SIDA?.....	45
2. Historia del SIDA.....	46
2.1 Teoría basada en que su origen proviene de los chimpancés.....	46
2.2 Teoría del “mono verde”.....	47
2.3 Teoría de la “siempre existencia”.....	47
2.4 Teoría de la creación del virus en laboratorio.....	48
3. Vías de transmisión.....	48

3.1	Contacto sexual.....	48
3.2	Vía sanguínea.....	48
3.3	Vía perinatal.....	49
3.4	Otros.....	49
4	Focos de Infección.....	49
4.1	Homosexualidad (en hombres).....	49
4.2	Prostitución (en mujeres).....	50
4.3	Drogadicción (via inyección).....	50
5	Formas de prevención de la transmisión del VIH.....	50
5.1	Uso del condón.....	50
5.2	Abstinencia.....	53
5.3	Fidelidad Mutua.....	53
6	El SIDA en México.....	54
6.1	Introducción.....	54
6.2	CONASIDA.....	54
6.3	Mitos del SIDA en nuestro país.....	55

CAPÍTULO IV PROPUESTA Y VALIDACIÓN DEL MODELO.....57

1.	Diagnóstico de la adecuación del modelo.....	57
2	Gráfica de los residuales.....	58
3	Histograma de los residuales.....	60
4.	La ACF y la PACF de los residuales de cada uno de los posibles modelos.....	61
5.	Periodograma de los residuales.....	64
6.	Periodograma integral de los residuales.....	65
7.	Modelo de pronóstico elegido.....	66
8.	Pronósticos.....	67

CAPÍTULO V DECISIONES PARA CONTROL.....71

1.	Obtención de datos muestrales.....	72
2.	Estadísticas generadas con los datos muestrales.....	77
3.	La información como propuesta para controlar el porcentaje de S. I. D. A. en México.....	83
4.	El uso del condón como propuesta para controlar el porcentaje de S. I. D. A. en México.....	85

CONCLUSIONES.....89

GLOSARIO.....91

BIBLIOGRAFÍA.....95

INTRODUCCIÓN

Las preguntas que nos hacemos muy a menudo como parte de una sociedad a la que pertenecemos son: ¿Cómo erradicar la enfermedad que no respeta sexo, religión, edad, condición social y ningún otro aspecto que fuese importante destacar? ¿Qué hacer para controlar de manera adecuada este problema? ¿Qué información debo de conocer para prevenirme de la enfermedad? ¿Qué información debo de conocer para sentirme colaborador de una sociedad que poco a poco hace conciencia de la gravidez del problema? Nosotros como parte de esa sociedad y pensando en que podemos lograr más que simplemente preguntarnos y esperar a que el problema lo resuelvan y ya, hemos decidido informarnos más al respecto y hacer del tema nuestro trabajo de tesis.

La poca o mucha experiencia con la que podamos contar nos ha hecho saber que la información es el arma más importante que tenemos para erradicar este problema y la mayoría de los problemas. Por esta razón, creemos que es de suma importancia que el trabajo contenga aspectos relevantes y concretos del SIDA y no con información que confundiera al lector, ya que al respecto suele haber demasiada información que se presta para ello.

Las matemáticas nos darán una idea rápida y concreta de la situación en nuestro país; gracias a ellas y a la computadora podremos representarla en gráficas que nos permitan visualizar aspectos relevantes como por ejemplo: una supuesta tendencia creciente, esta suposición de acuerdo a la información que hasta ahora las campañas informativas de gobierno nos han brindado.

El objetivo de este trabajo será el de crear un modelo de pronóstico para las estadísticas que se recopilarán mediante un muestreo que se realizará con escasos recursos, únicamente con lo que dos estudiantes puedan aportar para la realización de ello, es importante mencionar que se trabajará de esa manera, ya que describiremos a una población grande que aproximadamente la componen 100 millones de elementos, que necesita sin lugar a duda, inmensos recursos para poder describirla.

La técnica que se utilizará para crear dicho modelo son las series de tiempo, la razón es la de contar con el número de observaciones suficientes para utilizarlas.

Una vez que podamos tener una estimación de nuestra población en relación a la información que tiene del S. I. D. A., a sus tendencias sexuales, hábitos sexuales, etc. y al tener un modelo de pronóstico que nos prediga la tendencia de los casos de S. I. D. A. en México, podemos monitorear el control de la situación mediante propuestas que traigan consigo tendencias decrecientes de la gráfica que representa los casos de S. I. D. A. en México.

Sabemos que tendremos excesivos problemas debido a los recursos con los que contamos y a los objetivos fijados, pero no nos detendremos en la realización de ello, ya que junto con nosotros hay gente versada en el tema (como nuestro asesor: Maricarmen González Videgaray y algunos profesores con los que hemos platicado del caso), que confían en nosotros, en el trabajo y en los resultados que este arrojará pese a esos problemas de recursos.

Esperamos que este sencillo pero esmerado trabajo, sea un producto que sirva para hacer pensar a la gente que todos podemos contribuir en solucionar el problema de una manera fácil y eficiente: informándonos.

CAPÍTULO I

SERIES DE TIEMPO

Objetivo: Conocer la terminología necesaria para poder utilizar la metodología de Box-Jenkins y poder crear nuestro modelo.

1. ¿QUÉ ES UNA SERIE DE TIEMPO?

Como el objetivo de este trabajo es el de crear un modelo de pronóstico mediante el uso de series de tiempo, creemos conveniente indicar algo acerca de los pronósticos.

Para la elaboración de pronósticos contamos básicamente con dos enfoques: el cualitativo y el cuantitativo. El primero es importante cuando no se cuenta con información histórica y se dice que es altamente subjetivo y de criterio. Nuestro estudio no hace uso de este enfoque puesto que contamos con información que se ha ido generando a partir de un determinado momento hasta contar con información numérica lo suficientemente útil para nuestro propósito. Por lo tanto, hacemos uso de los métodos cuantitativos que para la elaboración de pronósticos se pueden subdividir en dos tipos: *series de tiempo y causales*. Con la información numérica que disponemos **NO** intentaremos predecir los movimientos futuros de una variable relacionándola con un conjunto de otras variables en una estructura causal (o sea, la determinación de factores que se relacionan con la variable a predecir); en cambio, se hace una predicción con base sólo en el comportamiento pasado de esa variable. Por esa razón hacemos uso de series de tiempo.

De esta forma, podemos decir que una Serie de Tiempo¹ es una colección de observaciones numéricas generadas en forma secuencial a través del tiempo. Los datos se ordenan con respecto al tiempo (por lo general con espacimientos iguales) y las observaciones sucesivas son generalmente dependientes entre sí.

La Serie de Tiempo observada es una realización de un cierto proceso. Una “realización” es toda una secuencia de observaciones. Se le llama de esta manera porque si se pudiera realizar de nuevo el proceso (lo cual en la mayoría de los casos reales es imposible), se tendrían resultados diferentes, es decir, una realización distinta. El objeto del análisis es describir el proceso teórico en forma de un modelo que tenga propiedades similares al proceso real.

Como ejemplos de Series de Tiempo tenemos:

1. Número de casos registrados de infección por VIH/SIDA mensualmente.
2. El número de mexicanos que salen anualmente al extranjero (con visa y pasaporte).
3. La temperatura promedio diaria.
4. El índice semanal de contaminación.
5. La cantidad de personas que usan a diario el metro.
6. El índice mensual de inflación.
7. Las ventas mensuales de cierta empresa.
8. Número de asaltos denunciados en el D. F. diariamente.
9. Número de turistas que arriban a nuestro país, por mes, en el aeropuerto.
10. La población mundial analizada mensualmente.

1.1 Series de Tiempo vistas como procesos estocásticos.

Recordemos que un proceso o sucesión de eventos que se desarrolla en el tiempo, en el cual el resultado en cualquier etapa contiene algún elemento que depende del azar se denomina un *proceso aleatorio* o *proceso estocástico*. El marco teórico del análisis de Series de Tiempo es el de los procesos estocásticos. Más específicamente, si definimos un proceso estocástico como una sucesión de variables aleatorias $\{X_t\}$, con $t = 1, 2, \dots, n$, ordenadas en el tiempo, una Serie de Tiempo puede ser considerada como una realización o trayectoria de un proceso estocástico. Esto es, el valor observado de la serie en el instante t puede ser considerado como una muestra aleatoria de tamaño uno de la variable X_t del proceso estocástico definida en dicho instante. Nótese que X_t y X_{t-1} están separadas por k retardos si $|t_k - t_{k-1}| = k$.

Con base en lo anterior, se concibe entonces a una Serie de Tiempo como a la sucesión de observaciones generadas por un proceso estocástico cuyo conjunto de variables

¹ Nosotros creemos que la mejor forma de nombrar a una Serie de Tiempo es “sucesión cronológica”, debido a que el término “serie” en Matemáticas, recordemos, se refiere más bien a una suma infinita de valores de una variable; sin embargo, continuaremos haciendo mención a Series de Tiempo, debido a que ésta es la terminología que más se usa y, además, es la que se conoce en mayor grado.

aleatorias se toma en relación al tiempo; es decir, una realización de un proceso estocástico, como se ha mencionado en el párrafo anterior. Por lo tanto, la inferencia que se realice será acerca de las características del proceso estocástico generador de la serie observada. Además, así como existen procesos estocásticos discretos y continuos, existirán también Series de Tiempo discretas y continuas.

2. OBJETIVOS DE ANÁLISIS DE UNA SERIE DE TIEMPO

Básicamente, el análisis de Series de Tiempo puede tener cuatro objetivos²:

1. *Descripción (¿qué es?).* Se refiere a enunciar el comportamiento de un proceso, permitiendo una visión clara.
2. *Explicación (¿por qué?).* Permite un análisis de tipo causa-efecto, al relacionar un fenómeno con el medio circundante.
3. *Pronóstico (¿qué sucederá?).* Se refiere a la estimación de valores futuros con el fin de tomar decisiones o prever comportamientos.
4. *Control (¿cómo mantener el proceso de cierta forma?).* Esto sucede cuando no sólo se desea observar sino mantener el proceso dentro de ciertos límites.
5. *Optimización (¿cómo mejorar?).* Se pretende dar propuestas que modifiquen el comportamiento del proceso para obtener alguna mejora.

Hay que recordar que los análisis de Series de Tiempo no dan la respuesta a lo que nos reserva el futuro, pero resulta valioso en el proceso de pronóstico y ayuda a reducir errores en ésta.

Aunque los métodos de regresión pueden ser útiles al analizar datos de Series de Tiempo, las observaciones de Y no pueden considerarse como representativas de una muestra aleatoria. De hecho, pueden encontrarse correlacionadas entre sí, con una correlación que aumenta a medida que el intervalo de tiempo entre un par de observaciones decrece. Por ejemplo, es probable que el cambio en la tasa de desempleo para un mes en específico se encuentre relacionada con la que se observará para el siguiente mes. De esta forma, algunas de las suposiciones que son necesarias para el desarrollo de procedimientos inferenciales posiblemente no se verifiquen para los datos de una Serie de Tiempo.

² Notas de Clase.

2.1 ELEMENTOS ESTADÍSTICOS EN EL ANÁLISIS DE SERIES DE TIEMPO.

Las Series de Tiempo consisten de datos numéricos y por esta razón se hace uso de herramientas estadísticas para describirlas y analizarlas, de la misma forma como ocurre con cualquier otro conjunto de información numérica.

Recordemos que la Estadística utiliza dos enfoques básicos: 1) *el enfoque descriptivo*, que se ocupa esencialmente de resumir y describir en forma concisa, ya sea mediante gráficas o a través de unas cuantas medidas descriptivas (media, mediana, moda, desviación estándar, varianza, principalmente), la información con que se cuente; y 2) *el enfoque inferencial*, cuyo objetivo fundamental es el de utilizar muestras representativas para realizar inferencias (deducciones) que sean válidas para toda la población de donde se obtuvo la muestra.

Dentro de los *elementos descriptivos* de una Serie de Tiempo se encuentran pues, las gráficas y las medidas descriptivas, y posiblemente el orden en que se mencionó a estos elementos represente su orden de importancia, ya que es primordial construir gráficas antes de llevar a cabo cualquier tipo de cálculo, aunque sea tan solo para verificar visualmente la congruencia de los datos.

Por su parte, en los *elementos de inferencia* de una Serie de Tiempo, la población sobre la cual se desea inferir, depende fundamentalmente del tipo de análisis y/o modelo que se emplee. Por consiguiente, conviene mencionar que el análisis de una Serie de Tiempo dada puede ser realizado de maneras distintas; por ejemplo, uno de los métodos de análisis considerado como clásico es el conocido como *descomposición de Series*, del cual se hablará más adelante.

3. COMPONENTES DE UNA SERIE DE TIEMPO³

DESCOMPOSICIÓN DE SERIES. Este método presupone que la Serie de Tiempo está formada por una componente de *tendencia-ciclo*, que representa el movimiento de largo plazo de la serie; otra componente de *estacionalidad* cuya utilidad es la de representar básicamente a los efectos producidos por fenómenos que se repiten cada año con cierta constancia; y una componente más de *irregularidad* que sirve para caracterizar los movimientos imprevisibles y considerados como aleatorios, es decir, identificar los factores que ejercen influencia sobre cada uno de los valores periódicos. Sin embargo, podemos hablar en general de cuatro componentes:

- Variación cíclica
- Tendencia
- Variación estacional
- Fluctuación irregular

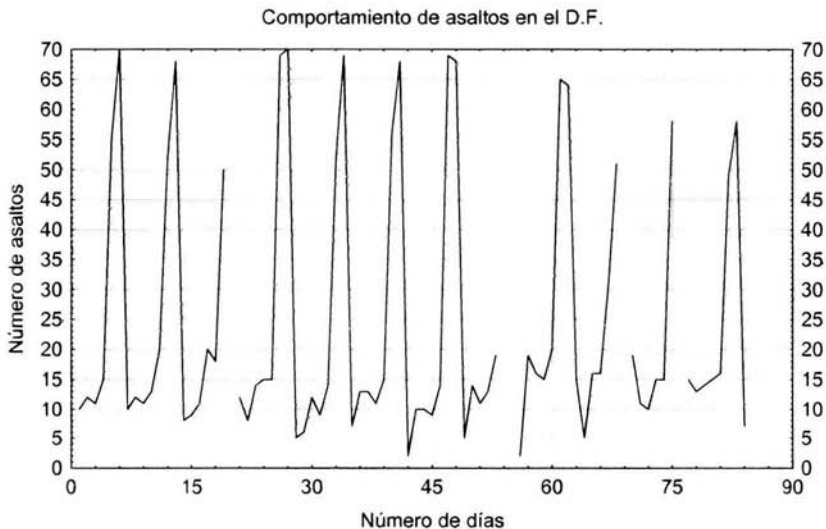
Con la metodología de descomposición de series se pretende identificar y estimar cada una de las componentes por separado, de tal manera que la serie pueda proyectarse a

³ Obtenido de "Pronósticos en los negocios", de John E. Hanke y Arthur G. Reitsch, editorial Prentice-Hall, 1996.

futuro y utilizarse en pronósticos tanto de corto como de largo plazos. Además, de acuerdo con este enfoque, podría pensarse que las componentes de tendencia, variación cíclica y estacionalidad constituyen la parte determinista o semideterminista de la serie, mientras que la componente irregular vendría a ser su parte no-determinista o estocástica. Por esta razón, se desea inferir sobre el conjunto de series de tiempo formadas por la parte determinista o semideterminista, combinada con todas las posibles realizaciones imaginables de la parte estocástica. Veamos concretamente a lo que se refiere cada una de las componentes de una Serie de Tiempo.

3.1 VARIACIÓN CÍCLICA

Ligado a la tendencia, el componente cíclico es la fluctuación en forma de onda o ciclo, de más de un año de duración, producida por cambios en las condiciones económicas cambiantes. El componente cíclico se identifica eliminando o promediando los efectos de la tendencia. Ya que esta componente constituye lo que queda después de dichos ajustes, se le refiere como el *método residual*. A continuación se presenta una gráfica que muestra la variación cíclica. Ésta se refiere al número de asaltos denunciados en el D. F.



Fuente: Delegación Gustavo A. Madero.

Observe en la gráfica que existe un ciclo que se repite cada 15 días y aún dentro del mismo podemos observar que cada 7 u 8 casos se repite uno más; podemos suponer que la causa de esto es debido a que en el D. F. la mayoría de las empresas pagan cada 15 días y otras tantas cada semana.

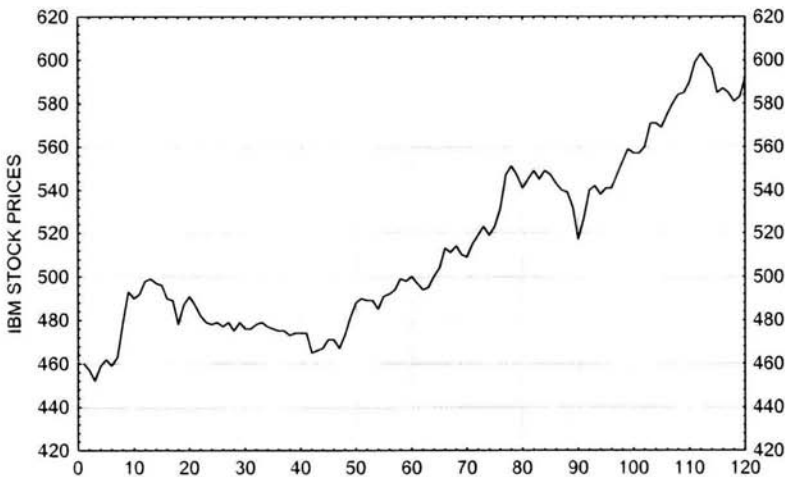
3.2 TENDENCIA

La tendencia es el componente de largo plazo que constituye la base del crecimiento o declinación de una serie. Las fuerzas básicas que producen o afectan la tendencia pueden ser por ejemplo: cambios en la población, cambios de precios, cambios tecnológicos, incrementos en la productividad, ciclos de vida de los productos.

Antes de medir la tendencia se debe primero conocer el propósito de su medición. Este conocimiento nos guía en la elección del método y dimensión de la Serie de Tiempo a utilizar en la medición. Existen dos propósitos fundamentales: proyectar la tendencia y eliminarla de los datos originales.

En el análisis de la tendencia la variable independiente es el tiempo. Se grafican los datos y en este sentido uno debe notar que no siempre una gráfica de los datos y una declaración de la finalidad de la medición nos permite efectuar la elección final. En ocasiones es necesario calcular y graficar dos o más tendencias junto con los datos originales para ver cuáles de ellas se ajustan a la serie. También lo que parece un cambio de nivel en una serie pequeña puede no serlo al tener más datos, convirtiéndose en un movimiento cíclico. Aquí será indispensable el conocimiento del fenómeno para establecer el modelo adecuado.

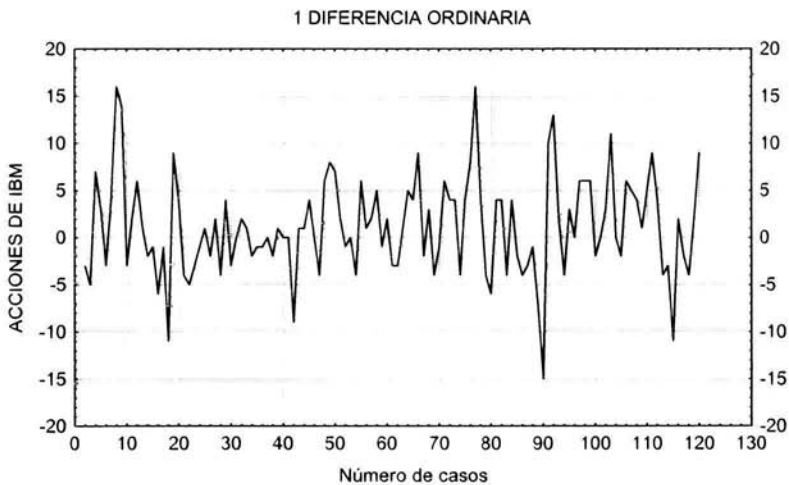
Obsérvese el siguiente ejemplo. La gráfica mostrada a continuación presenta el precio de la acción por mes de IBM, y es un claro ejemplo para observar la tendencia creciente que muestran sus acciones.



Fuente: Ejemplo del paquete estadístico STATGRAPHICS

Para eliminar la tendencia podríamos usar varios métodos como por ejemplo el de un modelo de regresión el cual se usa cuando la tendencia es fija y determinística. Aunque es muy importante diferenciar entre ajustar una serie y modelarla para efectuar un pronóstico. El problema real radica en que es muy difícil saber cuándo un cambio en el nivel de la serie se debe a que la tendencia sea determinística o estocástica. Por esta y otras razones se usan las *diferencias*. No siempre es lo más efectivo, aunque generalmente funciona para series con un número grande de observaciones.

Veamos lo que sucedió en el ejemplo del precio de las acciones de IBM, al aplicar 1 diferencia ordinaria a las observaciones. Se observa la eliminación de la tendencia.



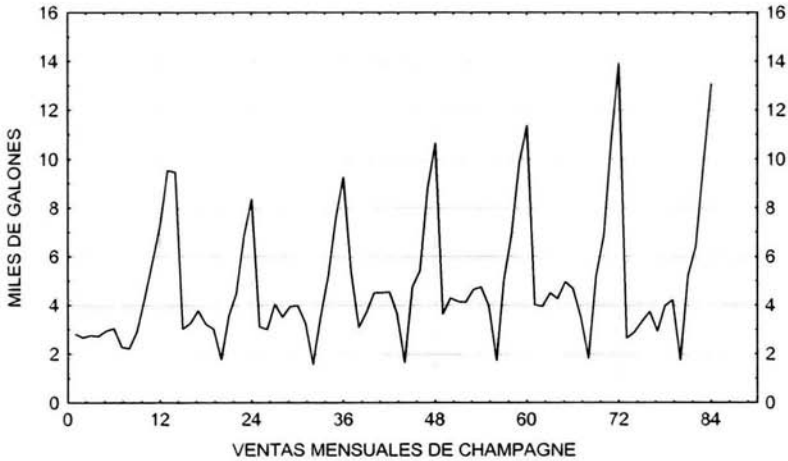
Esto hace claramente una serie estacionaria con media constante y varianza estabilizada, por lo que en la serie es posible obtener los pronósticos.

3.3 VARIACIÓN ESTACIONAL

Las fluctuaciones estacionales se encuentran típicamente en los datos clasificados por trimestre, mes o semana. La variación estacional se refiere a un patrón de cambio regularmente recurrente a través del tiempo. El movimiento se completa dentro de la duración de un año, y se repite a sí mismo año tras año. La estacionalidad es generalmente clara después de eliminar la tendencia, y en algunos casos es evidente aun con la tendencia, pero es importante observar que al analizar la estacionalidad es imposible determinar sus efectos sin poner atención a la tendencia, eliminándola primero, si es que existe. Además, la identificación de la estacionalidad es importante, debido a que proporciona información

acerca de la “regularidad” en la serie que puede ayudarnos a hacer un pronóstico y para eso es importante la función de autocorrelación.

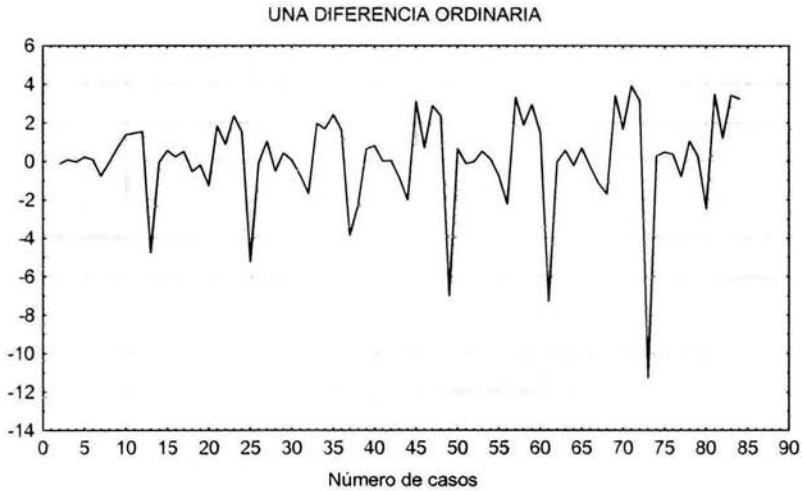
La siguiente gráfica muestra la variación estacional que se registra cada año en las ventas de champagne. Obsérvese cómo la gráfica muestra claramente heteroscedasticidad (varianza no estable) y tendencia creciente.



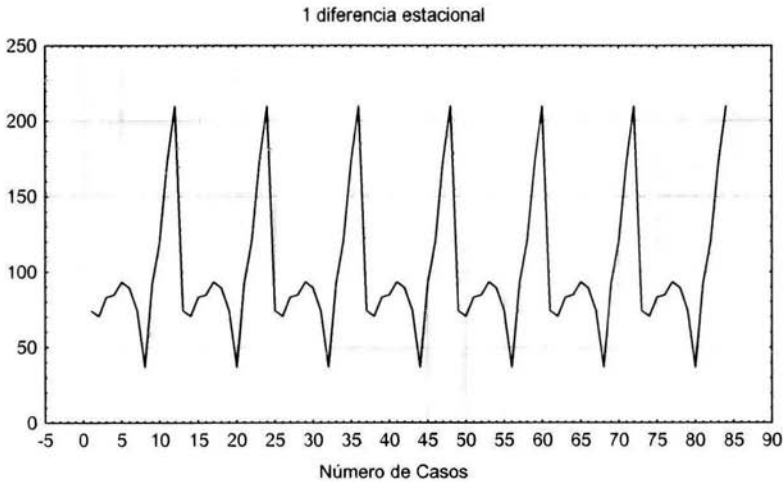
La gráfica requiere de una transformación precisamente para hacer estable la serie. Cabe recordar que se escogerá la gráfica menos dispersa y, por esa razón, más estable.

Hablando de la estacionalidad, existen varios métodos para manejarla. La mayoría son métodos de “ajuste”, basados únicamente en la información contenida en la serie. Como procedimiento de ajuste pueden hacerse diferencias estacionales, de manera similar a las diferencias consecutivas (ordinarias).

Al aplicar una diferencia ordinaria a la gráfica observamos la eliminación de la tendencia. Obsérvese:

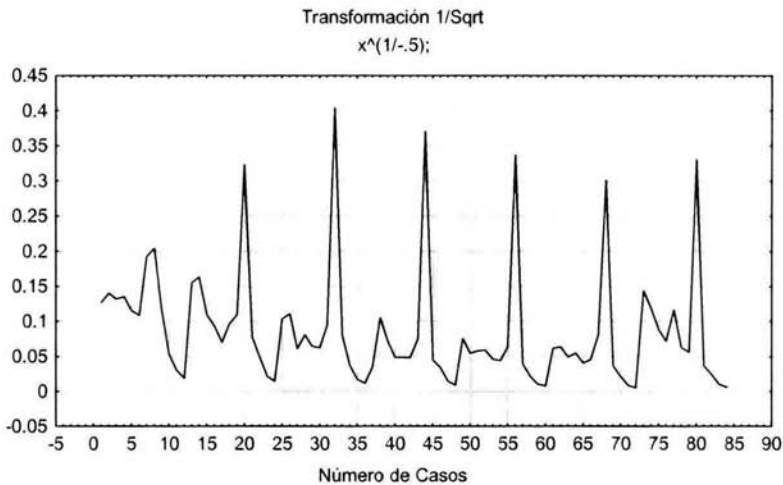


Aun así, la varianza no es estable. La variación estacional, se eliminará aplicando diferencias estacionales; posteriormente haremos una transformación a la serie original. Retomemos el ejemplo de las ventas mensuales de champagne y su gráfica, observando lo que sucede:



Claramente se observa aun el componente estacional, aunque la tendencia no se observa ya.

Como decíamos anteriormente, la serie original de la venta de champagne requiere de una transformación. Al parecer la que mejor se presta para nuestros propósitos es la $1/\sqrt{x}$.

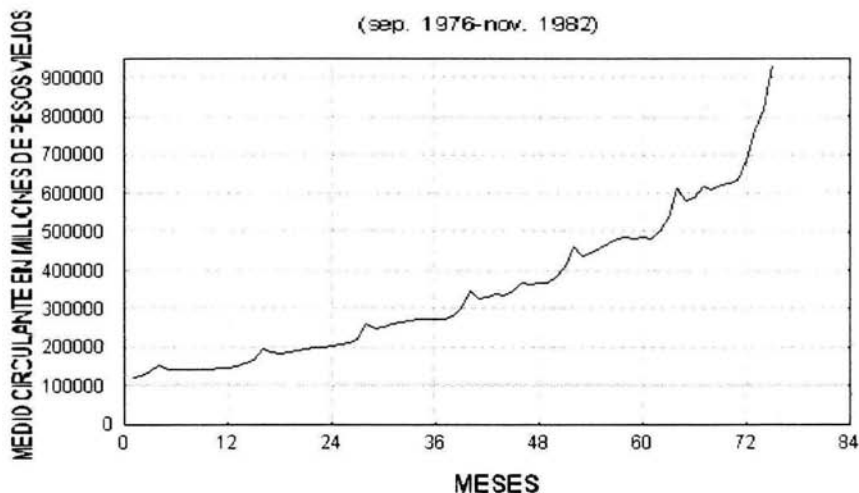


Nótese que la gráfica tiene una varianza más estable comparada con las anteriores gráficas por lo que la transformación fue de mucha ayuda.

3.4 FLUCTUACIÓN IRREGULAR

El comportamiento irregular está compuesto por fluctuaciones causadas por sucesos impredecibles o no periódicos, como un estado del tiempo poco usual, huelgas, guerras, rumores de guerras, elecciones y cambios en las leyes. Es importante el análisis de este componente debido principalmente a que se sospeche que estos sucesos o fenómenos anormales puedan afectar la estimación de los parámetros y los resultados del modelo en general. Se observa esta componente cuando, por ejemplo, uno o más de los residuales son mayores de lo esperado. Sin embargo, si dichos residuales pueden explicarse por la ocurrencia de fenómenos ajenos al comportamiento de la serie, convendría reducir la influencia que estos fenómenos ejercen sobre los resultados del modelo. Básicamente esto se realiza con algo que se conoce como la técnica de *análisis de intervención*, que se atribuye al artículo de Box y Tiao (1975)⁴. Veamos como ejemplo la siguiente gráfica que muestra la serie del medio circulante dentro de la economía mexicana durante el periodo septiembre 1976 a noviembre de 1982 y que abarca la devaluación de la moneda y nacionalización de la banca durante el sexenio de José López Portillo:

⁴ Para un análisis más detallado de esta técnica consúltense "*Análisis estadístico de series de tiempo económicas*", de Víctor M. Guerrero, editado por la UAM unidad Iztapalapa, 1991.



Esta gráfica muestra estacionalidad, pero obsérvese el comportamiento “extraño” después de la devaluación de la moneda (septiembre 1976 y febrero y agosto 1982) y del inicio del IVA (Impuesto al Valor Agregado) a partir de enero de 1980.

4. METODOLOGÍA DE BOX-JENKINS (ARIMA)

Aun cuando la metodología de descomposición de Series de Tiempo es bastante útil, actualmente su empleo se ve limitado en esencia a la estimación de la componente estacional a fin de obtener series desestacionales. Esto se debe, en parte, a la aparición y unificación de ciertos métodos estadísticos relacionados con procesos estocásticos, que han probado su eficacia en lo que toca a la construcción de modelos para Series de Tiempo. Una de las principales ventajas del enfoque de procesos estocásticos aplicada a éstas es la gran flexibilidad que se logra para representar a un buen número de fenómenos reales mediante una sola clase general de modelos; otra ventaja es la facilidad y gran precisión que se obtiene al realizar pronósticos y, finalmente, una ventaja más es la posible generalización de manera natural de modelos para series individuales a modelos para varias series consideradas simultáneamente.

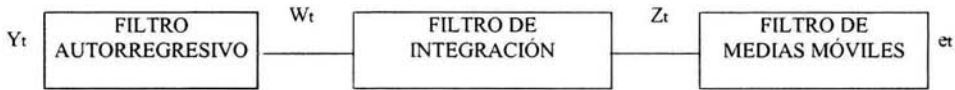
La idea fundamental de George E. P. Box y Gwilym M. Jenkins en su enfoque propuesto se halla en la estrategia que ellos proponen para construir modelos, los cuales no sólo deben ser adecuados para representar el comportamiento de los datos observados, sino que su elección debe ser sugerida por los datos mismos. Recordemos que el enfoque tradicional busca simplemente lograr el mejor ajuste de modelos preconcebidos, a los datos con que se cuenta.

El análisis de Series de Tiempo estudia cómo construir un modelo para explicar y prever la evolución de una Serie. La metodología Box-Jenkins es un procedimiento de

análisis estadístico para ajustar a una Serie un tipo especial de modelos, denominados ARIMA (*Autorregresive Integrated Moving Average*).

Esta técnica no asume ningún patrón particular en los datos históricos de la serie a pronosticar. Utiliza un enfoque iterativo de identificación de un modelo útil a partir de modelos de tipo general. El modelo elegido se verifica con los datos históricos para ver si describe la serie con precisión. El modelo se ajusta bien si los residuos entre el modelo de pronóstico y los puntos de datos históricos son reducidos, distribuidos de manera aleatoria e independiente. Si el modelo especificado no es satisfactorio, se repite el proceso utilizando otro modelo diseñado para mejorar el original. Este proceso se repite hasta encontrar un modelo satisfactorio.

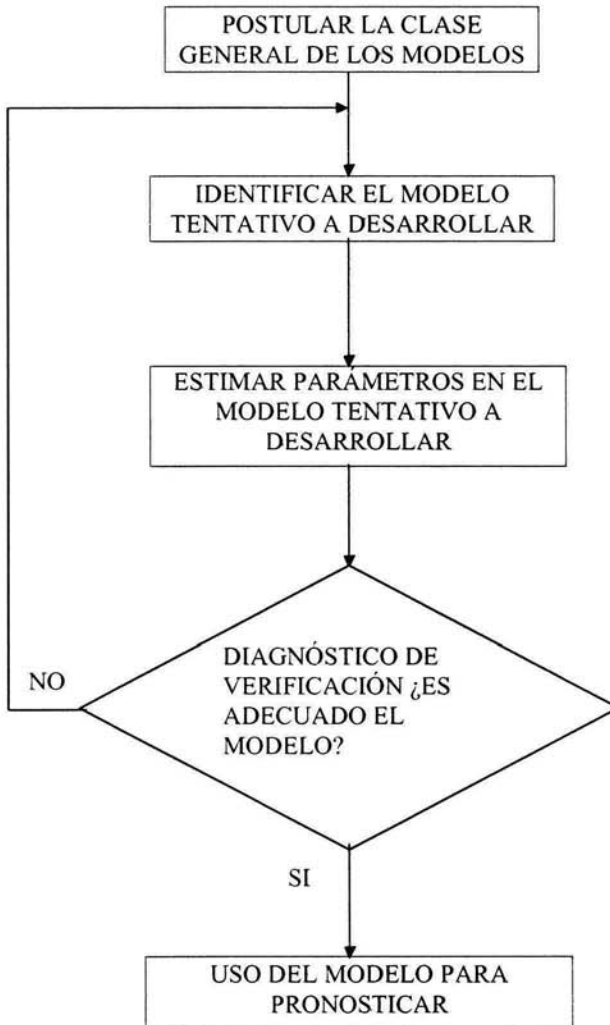
El método de Box-Jenkins consiste en extraer los movimientos predecibles de los datos observados. La Serie de Tiempo se descompone en varios componentes algunas veces llamados “filtros” precisamente porque la filosofía del método consiste en detectar las distintas componentes usando los filtros correspondientes, hasta obtener residuales no predecibles cuyo comportamiento tiene poca influencia en el resultado final. El enfoque de Box-Jenkins hace principalmente uso de tres filtros finales: el autorregresivo, el de integración y el de medias móviles.



El proceso iterativo de Box-Jenkins para construir modelos lineales de Series de Tiempo consiste de cuatro pasos:

- Identificación de las especificaciones preliminares del modelo.
- Estimación de los parámetros.
- Diagnóstico de la adecuación del modelo.
- Pronóstico de realizaciones futuras.

La siguiente figura muestra el proceso iterativo (diagrama de flujo):



En el análisis de Series de Tiempo y pronóstico se ve que los modelos matemáticos que se necesiten contienen algunas constantes o parámetros cuyos valores pueden ser estimados a partir de los datos. Entonces, en la práctica será necesario emplear el menor número posible de parámetros para una adecuada representación de nuestro modelo. El papel central que juega este **principio de parsimonia**⁵ (ahorro) en el uso de parámetros es de trascendental importancia por lo que es el que deberíamos de seguir.

El objetivo es, entonces, obtener modelos parsimoniosos (buscar el equilibrio) aunque adecuados. Los procesos de pronósticos y control pueden ser seriamente deficientes si son inadecuados en el uso de los parámetros. Por lo que se necesita cuidado, precisión, atención, esmero y esfuerzo para la selección del modelo. El proceso de selección es necesariamente iterativo, es decir, un proceso de evolución, adaptación y error.

EXPLICACIÓN DEL PROCESO ITERATIVO DE BOX-JENKINS

Por medio de la gráfica de las observaciones muestrales se examina si existen varianzas no constante (heteroscedasticidad), tendencia y variación estacional; en algunos casos todo esto se puede realizar a simple vista.

La varianzas se puede estabilizar por medio de una transformación, por ejemplo la de BOX-COX (incluida en el paquete estadístico STATGRAPHICS y STATISTICA). Si se hizo alguna transformación debemos examinar brevemente sus funciones de autocorrelación (ACF) y de autocorrelación parcial (PACF). Si existe tendencia, podremos aplicar alguna diferencia ordinaria ya que la n -ésima diferencia de un polinomio de grado " n " lo convierte en constante (cabe recordar que en cada diferencia se pierde un dato de la serie original). Si llegáramos a encontrar variaciones estacionales podremos eliminarlas mediante diferencias estacionales de periodo " s " (recordando que con cada diferencia se pierden " s " datos).

4.1 IDENTIFICACIÓN DEL MODELO

El primer paso consiste en determinar si la serie es estacionaria; o sea, si el valor de la media no varía a través del tiempo; esto es lo que significa la estacionaridad. Si la serie no lo es, en general se puede convertir a una serie estacionaria mediante diferenciaciones. Una vez obtenida una serie estacionaria se debe identificar la forma del modelo a utilizar. Esto se logra mediante la comparación de los ACF y PACF de nuestra serie para determinar aproximadamente qué modelo(s) se ajusta(n) a nuestros datos. Entre los modelos de tipo ARIMA más comunes están:

⁵ Fue mencionado por primera vez por Durand de Saint-Pourçain (1270-1334), quien fue teólogo dominicano francés. William de Ockham, escolástico medieval, lo usó con tal frecuencia y de forma tan aguda que en inglés se conoce también con el nombre de <<Ockham's razor>>. Parsimonia, en latín significa frugalidad.

a) Modelos Autorregresivos de Orden p : $AR(p)$

$$Z_t = F_1 Z_{t-1} + F_2 Z_{t-2} + \dots + F_p Z_{t-p} + e_t$$

donde e_t es un proceso de ruido blanco independiente de Z_{t-h} , para todo $h = 1$, para todo t .

Si B es el operador retardo:

$$BX_t = X_{t-1}, \dots, B^k X_t = X_{t-k}$$

la ecuación que resulta es:

$$F_p(B)X_t = e_t$$

Donde

$$F_p(B) = 1 - F_1 B - \dots - F_p B^p$$

Un proceso estacionario y autorregresivo $AR(p)$ se caracteriza porque los primeros p coeficientes de la PACF son no nulos y el resto cero. Estos modelos poseen media constante y varianza constante (con restricciones de estacionaridad) y en general la ACF presenta muchos coeficientes no nulos que decrecen con el retardo con mezcla de exponenciales y sinusoidales; el comportamiento de la ACF será decreciente infinita.

b) Modelos de Medias Móviles de orden q : $MA(q)$

$$Z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

donde e_t es un proceso de ruido blanco. Considerando el operador retardo B , la ecuación que resulta es:

$$X_t = \theta_q(B)e_t$$

Donde:

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q.$$

Todo proceso de media móvil es estacionario. Un proceso de media móvil $MA(q)$ se caracteriza porque los q primeros coeficientes de la ACF son no nulos y el resto cero; y la PACF, en general, presenta muchos coeficientes no nulos que decrecen con el retardo con mezcla de exponenciales y sinusoidales. Estos modelos poseen media constante, al igual que la varianza y ACF que se trunca en " q ". Tienen restricciones de invertibilidad.

c) Modelos Mezclados Autorregresivos y de Medias Móviles: ARMA(p,q)

$$Z_t = F_1 Z_{t-1} + F_2 Z_{t-2} + \dots + F_p Z_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

En estos modelos e_t es un proceso de ruido blanco y al igual que los modelos anteriores, considerando el operador de retardo B, la ecuación que resulta es:

$$F_p(B)X_t = \theta_q(B)e_t$$

donde

$$F_p(B) = 1 - F_1 B - \dots - F_p B^p$$

$$\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$$

Dado que los procesos ARMA son una mezcla de procesos AR y procesos MA, la ACF y PACF de un proceso estacionario ARMA(p,q) será una superposición de las PACF y ACF de los procesos AR(p) y MA(q) correspondientes. Tendrán por supuesto condiciones tanto de estacionaridad como de invertibilidad y su función de autocorrelación será decreciente infinita (ya que incluye una parte de tipo autorregresivo), en la cual el efecto de la parte de medias móviles desaparece después de "q" periodos.

d) Procesos autorregresivos integrados de medias móviles: ARIMA(p,d,q)

$$(1 - F_1 B - \dots - F_p B^p) \nabla^d X_t = (1 - \theta_1 B - \dots - \theta_q B^q) e_t$$

donde e_t es un proceso de ruido blanco y ∇ es el operador de diferencia:

$$\nabla X_t = (1 - B)X_t = X_t - X_{t-1}, \dots, \nabla^d X_t = (1 - B)^d X_t = (1 - B)^{d-1} (X_t - X_{t-1})$$

Equivalentemente:

$$F_p(B)W_t = \theta_q(B)e_t$$

donde

$$W_t = \nabla^d X_t$$

De esta manera, el proceso ARIMA(p,d,q) para X_t es equivalente al proceso ARMA(p,q) para W_t . Si el proceso W_t es estacionario, el proceso ARIMA para X_t quedará caracterizado por las ACF y PACF del proceso W_t y por d , el orden de diferenciación del proceso original X_t .

Aun cuando en general no será posible hacer coincidir exactamente los datos con las distribuciones teóricas, se pueden efectuar pruebas durante el siguiente paso o etapa para

determinar si el modelo es adecuado. Entonces, si el modelo no es satisfactorio, se puede intentar un modelo alternativo, recordando el principio de parsimonia mencionado anteriormente. Aquí son de vital importancia las ACF y PACF. Esta última es aquella que mide el efecto o la contribución de una variable de un modelo de orden superior con respecto a un modelo de orden inferior (de tipo autorregresivo). Por ejemplo, para un modelo puramente autorregresivo, en general la PACF se trunca en “p”.

En esta etapa debemos saber identificar las autocorrelaciones que caen exponencialmente a cero. Si éstas descienden de esta forma, el proceso indicado es el AR; si son las autocorrelaciones parciales las que descienden a cero, entonces el proceso indicado es el MA; y, si tanto los coeficientes de autocorrelación como los de autocorrelación parcial descienden a cero, el indicado es un proceso mixto ARIMA.

Por su parte para medir el efecto parcial de cada variable autorregresiva en un modelo de medias móviles, debemos “invertir” este modelo a uno que posea exclusivamente partes autorregresivas.

En las siguientes figuras se muestran las distribuciones teóricas para los modelos ARIMA más comunes y que nos ayudarán en la selección de un modelo adecuado.

Podemos resumir mediante los siguientes cuadros lo expuesto anteriormente acerca de los modelos teóricos y de modelos *no estacionales*:

Tabla 1.

MODELO	ACF	PACF
AR(p)	Decae infinitamente	Se trunca después de p
MA(q)	Se trunca después de q	Decae infinitamente
ARMA(p,q)	Decae infinitamente	Decae infinitamente

Tabla 2.

MODELO	EQUIVALE A:	CONDICIONES
AR(p)	MA(∞)	Estacionaridad
MA(q)	AR(∞)	Invertibilidad
ARMA(p,q)	MA(∞) AR(∞)	Estacionaridad Invertibilidad

Es conveniente mencionar algo acerca de las series de tiempo con características estacionales, debido a que se presentan frecuentemente en la práctica.

e) *Modelos estacionales*. Se entiende por una serie estacional la que, aparte de contener una tendencia (y/o ciclos) de larga duración, muestre fluctuaciones que se repiten anualmente, tal vez con cambios notables a través de los años. Aunque es importante

señalar que aun cuando la estacionalidad en general se considera como un fenómeno repetitivo anual, esto no implica que no pueda existir un cierto patrón de comportamiento periódico con duración menor a un año.

Al analizar los residuales es posible observar el comportamiento estacional en la serie de tiempo. De esta forma, podremos encontrar modelos *puramente estacionales* (de tipo autorregresivo, de medias móviles y mezclados), así como modelos generales multiplicativos estacionales. El comportamiento de estos modelos se asemeja al de los modelos *no estacionales*, de tal forma que podemos identificar los modelos multiplicativos estacionales por:

$$\text{ARIMA}(p, d, q) * (p, d, q)_s$$

donde:

- p = orden del proceso AR estacional
- q = orden del proceso MA estacional
- d = orden de diferencias estacionales
- s = longitud del periodo.

También podemos basarnos de la siguiente tabla para identificar los modelos de este tipo:

Tabla 3.

MODELO	ACF	PACF
AR(p) * SAR(P)	Decae	Se trunca después de p+sP
MA(q) * SMA(Q)	Se trunca después de q+sQ	Decae
Mezclados	Decae	Decae

4.2 ESTIMACIÓN DE LOS PARÁMETROS

A no ser que el fenómeno se comporte como ruido blanco, se entiende que el modelo que, en un momento dado nosotros escogimos como tentativo a desarrollar, deberá contener al menos un parámetro. La obtención de parámetros se realiza por medio de los siguientes métodos:

- a. Método de los momentos.
- b. Método de máxima verosimilitud.
- c. Método de los mínimos cuadrados.

El primero método busca igualar los momentos muestrales a los momentos poblacionales; sin embargo, al tratarse de modelos con varios parámetros, se tendría un poco más de dificultad al utilizar este método.

El segundo tiene como objetivo el MAXIMIZAR la función de distribución conjunta de los datos muestrales. Para esto se debe suponer que se tiene una muestra aleatoria independiente de una misma población.

El tercer método tiene como meta principal encontrar un modelo que minimice la suma de los cuadrados de las distancias de los puntos muestrales al modelo teórico.

Tanto el método de los mínimos cuadrados como el de máxima verosimilitud nos dan los mismos estimadores. Así mismo, podría hacerse mención de que en ambos métodos, si lleváramos a cabo un análisis riguroso, se llegaría a la conclusión de que se debe de minimizar la expresión:

$$S(\Phi, \theta) = \sum_{t=1}^n e_t^2$$

donde:

$F = (F_1, F_2, \dots, F_p)$ y $\theta = (\theta_1, \theta_2, \dots, \theta_q)$; y por esta razón los estimadores obtenidos mediante los métodos de mínimos cuadrados y máxima verosimilitud son exactamente los mismos.

La expresión anterior se minimiza obteniendo las derivadas parciales con respecto a cada uno de los parámetros, teniéndose así $p + q$ ecuaciones iguales a cero, de las cuales deberán despejarse los valores de los parámetros. Aquí se pueden presentar dos problemas, a saber:

1) los errores e_t están en función de variables no observadas. Para resolver esta problema se procede a realizar lo siguiente:

- sustituir los valores no observados por sus valores esperados incondicionales, es decir, sus valores esperados independientes del comportamiento del modelo. Para Z_t , el valor esperado incondicional es la media de Z , y para e_t es cero.
- Utilizando los valores esperados incondicionales; calcular los valores esperados condicionales de Z_t utilizando el pronóstico hacia atrás (Backforecasting).

Si el número de observaciones es grande comparado con "p" y "q", y si el fenómeno es estacionario será suficiente con los valores esperados incondicionales.

2) los modelos que incluyen parte de medias móviles generan un sistema de ecuaciones no lineales. Para resolver este problema, debemos recordar que se trata de linealizar las ecuaciones utilizando una serie de Taylor de orden 1 multivariada, alrededor de los valores iniciales de los parámetros. Los valores obtenidos al resolver ecuaciones linealizadas pueden usarse como nuevos valores iniciales hasta encontrar la convergencia (de ser posible).

Una vez que se ha solucionado el primer problema calculando los valores esperados incondicionales, entonces se genera un sistema de ecuaciones minimizando la expresión. Todo esto es posible obtenerlo gracias a los paquetes estadísticos con los que nos podremos auxiliar para la obtención de los parámetros, tales como el statgraphics o el paquete STATISTICA.

4.3. DIAGNÓSTICO DE VERIFICACIÓN. EXAMEN DEL MODELO

Antes de usar el modelo para pronóstico, se debe verificar si es adecuado. Debemos analizar el ajuste y los pronósticos obtenidos a través de este modelo mediante el diagnóstico de los residuales, lo cual constituye el aspecto de mayor relevancia dentro de la metodología de BOX-JENKINS.

Como primera instancia para comprobar su adecuación se pueden revisar los términos de error e_t , para asegurarse de que son aleatorios. Esto se puede realizar observando que las autocorrelaciones de los residuales NO SEAN diferentes de cero en forma significativa. Si algunos retrasos de orden menor o estacionales son diferentes de cero en forma significativa, entonces el modelo resulta inadecuado y debemos regresar a la primera etapa.

También podemos verificar lo adecuado del modelo haciendo una prueba ji-cuadrada (χ^2), conocida como la estadística Q de BOX-PIERCE, sobre las autocorrelaciones de los residuales. La estadística de prueba es:

$$Q = (N - d) \sum_{k=1}^m r_k^2$$

la cual está distribuida aproximadamente como una distribución ji-cuadrada con k-p-q grados de libertad.

En la expresión anterior:

N = longitud de la serie de Tiempo

k = primeras k autocorrelaciones que se verifican

m = número máximo de retrasos verificados

r_k = función de autocorrelación de la muestra del k -ésimo término de residuo

d = grado de diferenciación para obtener una serie estacionaria

si el valor calculado de Q es mayor que la χ^2 para $k-p-q$ grados de libertad, entonces se debe considerar que el modelo es inadecuado.

Ninguna de estas dos pruebas de adecuación debe considerarse como la última palabra, aunque en general deben utilizarse juntas, además de un juicio considerable por parte de nosotros. Es posible que se puedan juzgar dos o más modelos como aproximadamente iguales, aunque ninguno de ellos se ajuste exactamente a los datos. En este caso, debe prevalecer el principio de parsimonia y elegir el modelo más sencillo.

En la elección de un modelo suele presentarse el caso de haber incluido algún(os) parámetro(s) que no es (son) significativo(s). Este es el caso de que el modelo esté *sobreespecificado*. Para determinar si hay que omitir algún parámetro G se debe realizar una prueba de hipótesis, tal que:

$$H_0: G = 0$$

$$H_a: G > 0$$

Se construye el estadístico $T = G / s_G$ que se distribuye como una t de student con $(n-p-q)$ ó $(n-p-q-1)$ grados de libertad. La regla de decisión será rechazar H_0 si $|T| > 2$. Si se cumple H_0 , entonces el parámetro es significativo y por lo tanto debe omitirse tomando en cuenta las siguientes observaciones:

- Si el parámetro significativo está en cualquier orden distinto del final, entonces, si hay correlación entre el parámetro no significativo y el último, se debe eliminar el último; en caso contrario, se requerirá de un modelo multiplicativo.
- Si el parámetro no significativo está al final del modelo, simplemente se elimina reduciendo el orden del modelo (de uno en uno).

Si los residuales no se comportan como ruido blanco, sino que siguen algún modelo del tipo ARMA(p,q), este modelo debe incorporarse al modelo tentativo original, de manera que los residuales se conviertan en ruido blanco. Este proceso se refiere a aquellos modelos que están *subespecificados*.

Convendría mostrar a continuación en forma resumida las recomendaciones que hay que considerar para elegir de un modelo sencillo al más complejo:

- Análisis de estacionaridad.
- Análisis de los residuales.
- Modelo sobreespecificado.
- modelo subespecificado.

Por último, es importante indicar cuáles son los diversos estadísticos y medidas de bondad de ajuste que se requieren para esta etapa:

- valor del estimador
- estimador del error estándar de cada estimador
- el cociente del estimador entre su error estándar (estadístico t)
- límites del 95% de confianza
- gráfica de los residuales normalizada.
- gráfica de los residuales y periodograma de los mismos.
- estimador de la desviación estándar de los residuales.
- suma de los errores al cuadrado ($R^2 = 1 - (\sum e_i^2 / \sum Z_i^2)$).
- periodograma integral de los residuales.
- estadístico BOX-PIERCE de los residuales.

4.4 PRONÓSTICOS

La última etapa de la metodología mencionada consiste en pronosticar. Una vez que se encontró un modelo adecuado, se pueden realizar pronósticos para uno o varios periodos a futuro. También se pueden formular intervalos de confianza sobre estas estimaciones. En general, entre más a futuro se pronostica, mayor será el intervalo de confianza. Estos pronósticos e intervalos de confianza se pueden obtener mediante algún paquete estadístico computacional, por ejemplo los que hemos mencionado, y otros: STATGRAPHICS, STATISTICA, SPSS para windows, SAS, Hojas electrónicas, etc.

Creemos importante hacer las siguientes indicaciones:

- Al haber más datos disponibles, se puede utilizar el mismo modelo para revisar los pronósticos, escogiendo otro periodo de origen.
- Si la serie parece cambiar a través del tiempo, pudiera ser necesario recalcular los parámetros, o incluso desarrollar un modelo nuevo por completo.

Si se aprecian pequeñas diferencias en los errores de pronóstico, pudieran indicar que es necesario recalcular los parámetros, en este caso regresaríamos a la etapa 2 a estimar parámetros. Cuando se aprecian grandes diferencias en la dimensión de los errores, pudieran indicar que se requiere un modelo completamente nuevo, regresando así hasta el comienzo de nuestro análisis.

CAPÍTULO II

RECOPIACIÓN Y ANÁLISIS DE DATOS

Objetivo: Recopilar los datos muestrales necesarios para realizar el análisis mediante la metodología de Box-Jenkins.

1. DATOS MUESTRALES

Los datos muestrales recopilados fueron proporcionados en la Dirección General de Epidemiología de la Secretaría de Salud; siendo ésta, una institución confiable que proporcionó la información de una manera fácil y amable, ya que cuenta con una biblioteca y personal capacitado para brindar este tipo de datos. En un principio tratamos de conseguir tal información a través del CONASIDA, pero dicha institución no tiene como objetivos principales brindar ésta, ya que entre ellos se encuentran el brindar información relacionadas con el S. I. D. A. y no con las estadísticas que se tengan de la enfermedad.

La información utilizada no está estratificada en ningún aspecto, como lo pueden ser: género, edad, vía de transmisión, grupo sexual, etc. y representan el número de casos notificados de personas infectadas por el VIH en México por mes. Los datos muestrales son recopilados por este período de tiempo (por mes) por las siguientes razones:

- a. Contamos con más de 80 datos, que es la cantidad mínima para realizar un análisis por medio de la metodología de Box-Jenkins.
- b. Los datos varían de manera significativa utilizando dicho periodo de tiempo.

Los datos muestrales que se utilizarán para el análisis son los siguientes:

I.2 CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS DEL AÑO 1990 A 1999.

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO 1990

No. de observación	MES	Numero de casos
1	ENERO	139
2	FEBRERO	297
3	MARZO	244
4	ABRIL	182
5	MAYO	177
6	JUNIO	129
7	JULIO	141
8	AGOSTO	241
9	SEPTIEMBRE	200
10	OCTUBRE	291
11	NOVIEMBRE	291
12	DICIEMBRE	242
	TOTAL	2574

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO 1991

No. de observación	MES	Numero de casos
13	ENERO	200
14	FEBRERO	218
15	MARZO	185
16	ABRIL	280
17	MAYO	192
18	JUNIO	188
19	JULIO	661
20	AGOSTO	446
21	SEPTIEMBRE	287
22	OCTUBRE	156
23	NOVIEMBRE	169
24	DICIEMBRE	184
	TOTAL	3166

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1992

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
25	ENERO	165
26	FEBRERO	117
27	MARZO	203
28	ABRIL	238
29	MAYO	191
30	JUNIO	359
31	JULIO	249
32	AGOSTO	201
33	SEPTIEMBRE	218
34	OCTUBRE	308
35	NOVIEMBRE	593
36	DICIEMBRE	352
	TOTAL	3194

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1993

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
37	ENERO	248
38	FEBRERO	357
39	MARZO	354
40	ABRIL	330
41	MAYO	202
42	JUNIO	487
43	JULIO	669
44	AGOSTO	438
45	SEPTIEMBRE	679
46	OCTUBRE	566
47	NOVIEMBRE	322
48	DICIEMBRE	394
	TOTAL	5046

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1994

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
49	ENERO	331
50	FEBRERO	238
51	MARZO	432
52	ABRIL	205
53	MAYO	290
54	JUNIO	232
55	JULIO	259
56	AGOSTO	372
57	SEPTIEMBRE	347
58	OCTUBRE	370
59	NOVIEMBRE	341
60	DICIEMBRE	637
	TOTAL	4054

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1995

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
61	ENERO	355
62	FEBRERO	261
63	MARZO	254
64	ABRIL	112
65	MAYO	249
66	JUNIO	391
67	JULIO	553
68	AGOSTO	524
69	SEPTIEMBRE	700
70	OCTUBRE	344
71	NOVIEMBRE	215
72	DICIEMBRE	338
	TOTAL	4296

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1996

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
73	ENERO	201
74	FEBRERO	514
75	MARZO	190
76	ABRIL	324
77	MAYO	476
78	JUNIO	492
79	JULIO	262
80	AGOSTO	568
81	SEPTIEMBRE	407
82	OCTUBRE	359
83	NOVIEMBRE	243
84	DICIEMBRE	158
	TOTAL	4194

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1997

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
85	ENERO	307
86	FEBRERO	381
87	MARZO	315
88	ABRIL	310
89	MAYO	210
90	JUNIO	311
91	JULIO	316
92	AGOSTO	212
93	SEPTIEMBRE	576
94	OCTUBRE	405
95	NOVIEMBRE	126
96	DICIEMBRE	183
	TOTAL	3652

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1998

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
97	ENERO	595
98	FEBRERO	556
99	MARZO	414
100	ABRIL	172
101	MAYO	316
102	JUNIO	500
103	JULIO	362
104	AGOSTO	329
105	SEPTIEMBRE	494
106	OCTUBRE	388
107	NOVIEMBRE	279
108	DICIEMBRE	340
	TOTAL	4745

CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
1999

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
109	ENERO	1
110	FEBRERO	11
111	MARZO	42
112	ABRIL	26
113	MAYO	36
114	JUNIO	142
115	JULIO	88
116	AGOSTO	79
117	SEPTIEMBRE	199
118	OCTUBRE	160
119	NOVIEMBRE	157
120	DICIEMBRE	151
	TOTAL	1092

**CASOS DE S. I. D. A. POR MES EN MÉXICO NOTIFICADOS EN EL AÑO
2000**

No. DE OBSERVACIÓN	MES	NÚMERO DE CASOS
121	ENERO	2
122	FEBRERO	8
123	MARZO	15
124	ABRIL	43
125	MAYO	90
126	JUNIO	137
127	JULIO	60
128	AGOSTO	122
129	SEPTIEMBRE	131
130	OCTUBRE	129
131	NOVIEMBRE	165
132	DICIEMBRE	208
	TOTAL	1100

1.3 CASOS DE S. I. D. A. POR MES DEL AÑO 2001 EN MÉXICO

Estos casos son de gran importancia para nosotros, ya que no serán utilizados como datos muestrales pero sí para compararlos con los pronósticos que genere el modelo elegido.

Los casos se muestran a continuación:

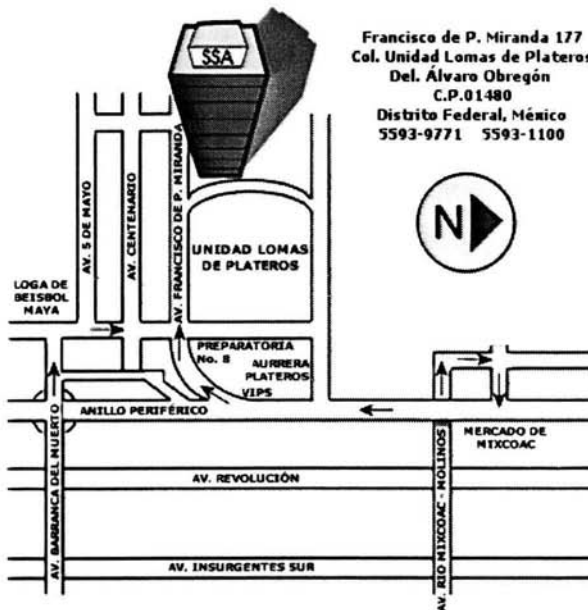
MES	NÚMERO DE CASOS	MES	NÚMERO DE CASOS
ENERO	0	JULIO	59
FEBRERO	6	AGOSTO	121
MARZO	20	SEPTIEMBRE	164
ABRIL	49	OCTUBRE	202
MAYO	88	NOVIEMBRE	104
JUNIO	151	DICIEMBRE	206

Fuente: Instituto Nacional de Epidemiología.

2. FUENTE DE INFORMACIÓN

La información fue proporcionada por el Instituto Nacional de Epidemiología, dicha Institución es bastante confiable y tiene como principal objetivo el de proveer información y conocimientos epidemiológicos oportunos y de calidad, para orientar las acciones de salud. Confiable al ser una institución estratégica de experiencia y dependiente de la Secretaría de Salud.

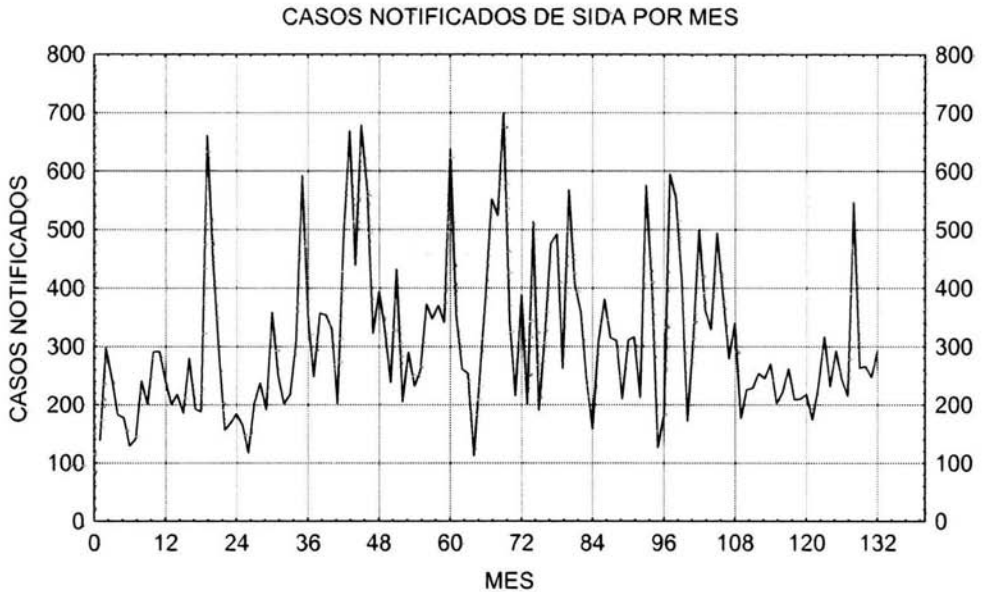
Dicha institución se ubica en Francisco de P. Miranda 177. Col. Unidad Lomas de Plateros, delegación Álvaro Obregón en el Distrito Federal. Se proporciona un croquis para su mejor ubicación:



3. GRÁFICA DE LA SERIE

El paquete estadístico utilizado para generar las gráficas y utilizar la metodología de Box-Jenkins es el STATISTICA, este paquete estadístico brinda todas las estimaciones y gráficas necesarias para utilizar dicha metodología.

Las gráficas que generan los datos muestrales es la siguiente:



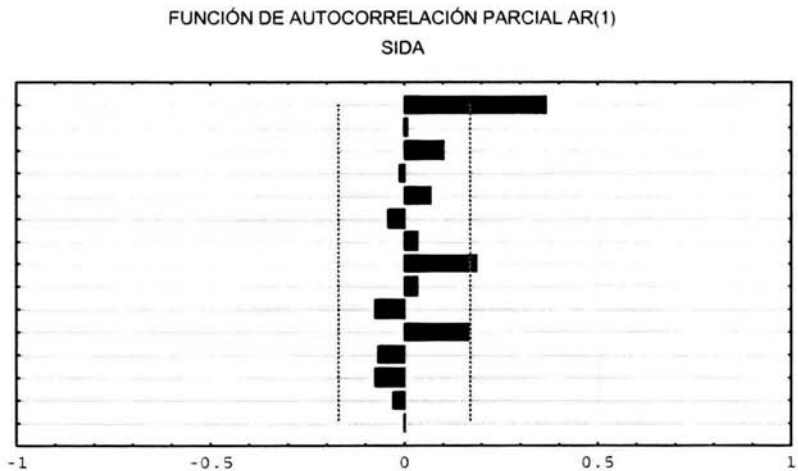
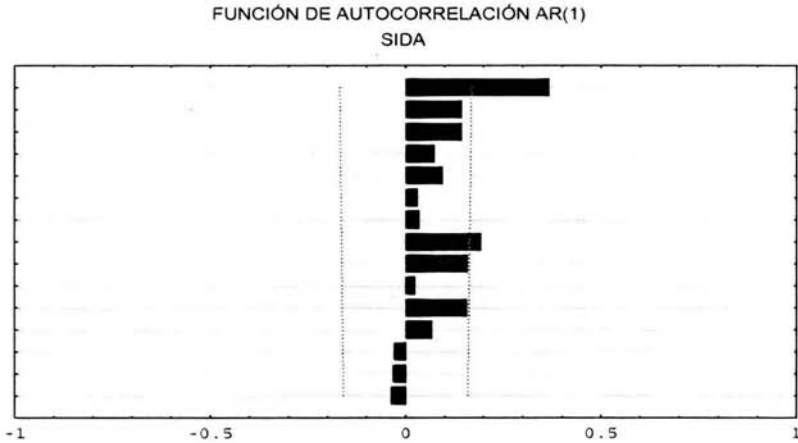
Se puede observar en la gráfica que no existe tendencia (nosotros esperábamos encontrar una tendencia creciente, de acuerdo a la información que se nos ha estado brindando en forma masiva), no existe heteroscedasticidad (varianza no constante) y no existen fluctuaciones estacionales que nos lleven a pensar que cada cierto periodo el comportamiento se repite; esto último pareciera ser evidente ya que no hay razón para pensar que cada año, por ejemplo, el número de casos notificados de S. I. D. A. se repiten.

Por lo tanto, no será necesario realizar transformaciones a la serie, ya que la varianza es constante, y por otro lado no será necesario realizar diferencias, dado que no existe una tendencia (creciente por lo regular y supuesta en forma muy probable en nuestro caso). Esto es bueno ya que no existen patrones que traigan cambios significativos en la gráfica, los cuáles tendrían que ser eliminados para analizar de manera más fácil a la serie. Otra ventaja de no haber encontrado esa supuesta tendencia creciente es que en ocasiones la tendencia pareciera ser determinística y en otras tantas pareciera ser producto de un conjunto de fenómenos aleatorios.

4. IDENTIFICACIÓN DE LOS POSIBLES MODELOS

Este proceso se realiza con la gráfica de la función de autocorrelación (ACF) y la gráfica de la función de autocorrelación parcial (PACF) de la serie.

Veamos las gráficas de la ACF y las PACF:



Observando las gráficas podemos mencionar que posiblemente los datos muestrales se ajusten muy bien a un AR (1), MA (1) o ARMA (1,1); pero tendremos que realizar algunas pruebas estadísticas y análisis estadísticos para seleccionar el modelo óptimo para el caso. Y claro, seleccionándolo siguiendo el principio de parsimonia.

CAPÍTULO III

EL SIDA

Objetivo: Conocer aspectos relevantes del tema que nos permitan realizar las estadísticas necesarias para el monitoreo destinado para el control del porcentaje de población infectada de SIDA.

1. ¿QUÉ ES EL SIDA?

El Síndrome de Inmunodeficiencia Adquirida (S. I. D. A) es un conjunto de síntomas y signos caracterizado por un grave trastorno del sistema inmunológico, o sea, una falta de defensas contra las enfermedades; y adquirida, dado que es una enfermedad contagiosa.

El S. I. D. A. se produce por un virus llamado VIH (Virus de Inmunodeficiencia Humana) que, infecta las células del sistema inmunológico es decir, destruye las defensas del organismo.

El S. I. D. A. es una enfermedad, pero hay que mencionar también que la gente no muere de S. I. D. A.; la gente con este mal muere a causa de las enfermedades a la queda expuesta, debido a que el virus deja a las personas infectadas, sin protección. Es decir, el sistema inmunológico ya no puede rechazar estas enfermedades.

Se puede estar infectado por el VIH durante ocho años o más sin presentar síntomas ni desarrollar la enfermedad. Durante este tiempo, las personas infectadas pueden transmitir el virus a otras sin saberlo. Por esta razón, nos parece importante conocer por lo menos cuáles son las vías de transmisión del virus y algo de la historia de este mal, ya que conociendo algo de ello podemos pensar en soluciones que nos ayuden a disminuir el porcentaje de población infectada de S. I. D. A en nuestro país.

2. HISTORIA DEL SIDA

El SIDA es una enfermedad que se manifestó hasta el inicio de la década de los 80's.

Es muy difícil precisar dónde se dieron los primeros casos de esta enfermedad. Sin embargo, podemos afirmar que las primeras descripciones datan de 1981. En esta fecha se dieron a conocer en Estados Unidos los casos de cinco adultos varones -jóvenes homosexuales- los cuáles presentaban una neumonía y otras infecciones. Dado de que presentaban síntomas no habituales, esto llamó la atención de las autoridades sanitarias de tal país.

En ese mismo año, se observaron varios casos donde la gente presentaba un cuadro de inmunodeficiencia adquirida, que les causaba infecciones en sus vías respiratorias; por otro lado, lo que había que hacer notar aquí es que todos presentaban un mismo cuadro hasta ese entonces no conocido. Por otra parte, todos los sujetos referidos en el informe inicial, eran varones homosexuales, lo que hizo creer, en un principio, que esta enfermedad era propia de este grupo de personas.

Se siguieron presentando nuevos casos y se observó que las infecciones no eran exclusivamente respiratorias. Poco después empezaban a observarse casos de SIDA en mujeres, niños, personas que no habían recibido transfusiones, hemofílicos y adictos a las drogas. Ello vino a demostrar que el SIDA no era una enfermedad exclusivamente de los homosexuales hombres, aunque sean ellos un grupo especialmente afectado.

Las características de la enfermedad, su propagación, el tipo de sujetos que la padecían, etc., hicieron que el SIDA se le considerara como una enfermedad de causas infecciosas y la situación de inmunodeficiencia que lo caracterizaba como inducida por un agente vírico. Estas suposiciones fueron posteriormente confirmadas.

2.1 TEORÍA BASADA QUE SU ORIGEN PROVIENE DE LOS CHIMPANCÉS

Las pruebas contundentes de que el virus del SIDA tuvo su origen en los chimpancés pueden acelerar la búsqueda de una vacuna o el mejoramiento de los tratamientos. Estos han dado a conocer las pruebas, que al parecer eliminan las dudas acerca del origen del virus que ha infectado a 35 millones de personas en todo el mundo. Gente de la Universidad de Alabama afirman que el virus tiene su origen en un subgrupo de la especie llamada Pan Troglodita. Las pruebas genéticas de la sangre y muestras de tejidos de un chimpancé hembra llamada Marilyn permitieron detectar un virus de inmunodeficiencia simio, VIS, que se asemeja mucho al VIH, el virus en humanos. El VIS infecta a los chimpancés pero no los enferma.

Los científicos creen que estudiando a los animales infectados en la vida silvestre descubrirán por qué no se enferman, información que será muy valiosa en la lucha contra el VIH y el SIDA.

Se supone que los chimpancés pueden ofrecer nueva información sobre cómo el virus causa la infección y la respuesta inmunológica en su contra. Se trabaja con los investigadores para desarrollar dos vacunas diseñadas para actuar juntas y así estimular las células T en el sistema inmunológico para que combatan el virus VIH. El descubrimiento podría ayudar a una mejor comprensión de cómo se originó la enfermedad "y quizás a (considerar) otros ángulos inesperados de la terapia".

Los científicos creen que el virus simio podría haberse trasladado a humanos a través de monos infectados o carne de mono. Comer carne de mono es común en algunas partes de África y el virus podría propagarse a través de la sangre salpicada durante su sacrificio.

2.2 TEORÍA DEL "MONO VERDE"

Se cree que el VIH es la consecuencia de una mutación. Es probable que el virus que se presenta en el mono verde africano, haya sufrido una mutación, se haya convertido en el VIH-1 y en el VIH-2 y se conoce que estas mutaciones se producen. De hecho los virus de la gripe mutan con gran frecuencia., de ahí que haya que modificar las vacunas para prevenir dicha enfermedad cada vez que tiene lugar la mutación del virus. Por lo tanto, algo parecido debió ocurrir con el virus productor del S. I. D. A.

Se piensa que, dada la promiscuidad de la convivencia que existe entre el mono verde africano y el hombre, el virus pasó a este último una vez sufrida la mutación. Y es por esta razón que los primeros casos de S. I. D. A. se localizaran en África Central. Dado que su proceso de incubación es largo pasan años entre el momento en que se produce la infección y la aparición de los síntomas del S. I. D. A. y que la situación sanitaria de África es muy deficiente, es posible que fuese extendiéndose entre la población africana sin que nadie se percatara de ello, todo esto aproximadamente en la década de los 60's.

En los años 70's las relaciones entre África y Haití fueron demasiado fluidas. Se cree que el flujo de africanos al Caribe, especialmente a Haití, permitió que la infección pasara del África a las islas caribeñas. De hecho, un buen número de los primeros casos de SIDA se detectaron en jóvenes norteamericanos varones homosexuales, que habían pasado vacaciones en Haití, donde habían mantenido relaciones sexuales con la población nativa. Probablemente, jóvenes homosexuales norteamericanos se infectaron con el SIDA durante su estancia en el Caribe y después importaron y extendieron la infección entre las comunidades gay de las grandes urbes norteamericana. Es de esta forma, como el virus ha ido extendiéndose a todo el mundo.

2.3 TEORÍA DE LA "SIEMPRE EXISTENCIA"

Según esta teoría, el virus siempre ha existido en ciertas poblaciones africanas, que viven apartadas de toda civilización. No les provocaba ningún daño, dado que se habían adaptado genéticamente al virus que lo transmitían por generaciones. Por razones que aun

no se conocen, el virus pasó a otras comunidades, por supuesto, mucho más sensibles al virus, y así es como apareció la enfermedad.

2.4 TEORÍA DE LA CREACIÓN DEL VIRUS

Según esta teoría, el virus es fruto de una manipulación realizada en los laboratorios, cuya creación era bélica y destructiva. Nosotros descartamos esta teoría, ya que de haber sido así, no se hubiese salido de control, extendiéndose por todo el mundo y sin contar con una vacuna para contrarrestarlo.

3. VÍAS DE TRANSMISIÓN

Las formas de transmisión del virus, más comunes son:

- contacto sexual
- vía sanguínea
- vía perinatal

3.1 CONTACTO SEXUAL

La transmisión del virus del S. I. D. A. se puede realizar a través de las relaciones sexuales, cuando hay intercambio de fluidos corporales tales como: líquido preeyaculatorio, semen, secreción vaginal y sangrado menstrual.

Esta es la forma de transmisión más importante, ya que es la causante de aproximadamente el 90% de los casos de S. I. D. A. en México. Siendo los homosexuales hombres, un grupo de riesgo importante y afectado, ya que este conforma más del 50% de los casos acumulados en México.

3.2 VÍA SANGUÍNEA

La transmisión se puede realizar por una transfusión de sangre infectada o por compartir objetos punzo cortantes que no han sido esterilizados y que tienen restos de sangre fresca contaminada por el virus, como lo pueden ser las agujas y jeringas.

En nuestro país existen leyes que no permiten la comercialización de sangre y obligan a que la sangre transfundida sea previamente analizada. Sin embargo, esta forma de transmisión, tiene otro problema, los drogadictos que usan drogas inyectables, ya que es común entre este grupo de personas, el compartir jeringas, agujas o cualquier otro tipo de instrumento para la administración de las drogas inyectables.

3.3 VÍA PERINATAL

Un bebé puede contraer el virus del S. I. D. A. al momento de nacer o bien durante el embarazo o la lactancia, siempre y cuando su madre esté infectada.

Esta forma de transmisión, constituye el principal mecanismo de contagio en los niños, ya que 2 de 3 casos de S. I. D. A. en ellos se deben a esta forma de contagio. México cuenta con una política nacional de tratamiento gratuito, la cual intenta proteger al 100% de las mujeres embarazadas y con esto mantener un nivel bajo de infección en mujeres embarazadas.

3.4 OTROS

Por trasplante de órganos o por inseminación artificial cuando esos órganos con semen provienen de sujetos VIH (+). A través de material médico (instrumentos quirúrgicos, material de endoscopia, material de dentistas, agujas de acupuntura, etc.) u otro tipo de material (depilación, peluquería, cepillos dentales, instrumentos utilizados en tatuajes, etc.) que contengan sangre infectada por el VIH.

Por supuesto, estos últimos mecanismos, aunque posibles, tienen un mínimo interés con los tres fundamentales, que son: sangre, sexual (semen, secreciones vaginales) y materno fetal.

También es importante señalar que la infección VIH no se transmite por ninguno de los siguientes mecanismos:

- contacto con una persona VIH + asintomática o enfermo de SIDA, ya sea, en la escuela, el trabajo, las relaciones sociales, familiares etc.
- a través de los estornudos, al hablar, al besar o al platicar a una persona infectada por el VIH o enferma de SIDA.
- a través del agua, alimentos, aire, insectos, objetos, etc.
- a través de la administración de vacunas con jeringas no infectadas, etc.

4. FOCOS DE INFECCIÓN

4.1 HOMOSEXUALIDAD (EN HOMBRES)

Los homosexuales hombres constituyen un grupo de riesgo importante, ya que la práctica del sexo anal, es predominante en este grupo.

La práctica del sexo anal, es un mecanismo fundamental de la infección, ya que, la mucosa rectal es especialmente rica en tejido linfático, mucho mas que la mucosa vaginal, lo que hace que el recto sea una zona muy vulnerable al VIH. El VIH presente en el semen,

líquido que contiene linfocitos, se deposita en el recto y encuentra en él un terreno más apropiado que en la vagina.

4.2 PROSTITUCIÓN (EN MUJERES)

La prostitución constituye un grupo de riesgo, ya que las mujeres, que se dedican a esta actividad, mantienen una intensa actividad sexual con varios hombres y corren el riesgo de infectarse en forma heterosexual. Es importante tomar en cuenta este grupo, ya que sabemos que la enfermedad poco a poco está abarcando el terreno heterosexual.

4.3. DROGADICCIÓN (VÍA INYECCIÓN)

Este grupo constituye, dentro de este tipo de transmisión el más importante, ya que corren el riesgo de quedar infectados al utilizar jeringas, objetos punzo cortantes, agujas, etc., contaminadas. En nuestro país, este grupo constituye casi el 100% de infección por vía sanguínea.

5. FORMAS DE PREVENCIÓN DE LA TRANSMISIÓN DEL VIH

La mejor manera de prevenir la infección por el VIH es el conocer los mecanismos de transmisión del virus y evitarlos.

Hay que evitar, fundamentalmente, los comportamientos de riesgo relacionados con el sexo y con las drogas. A este respecto, distintos organismos internacionales de reconocido prestigio han propuesto las siguientes recomendaciones:

- Practicar la abstinencia o mantener relaciones sexuales con el compañero(a) estable no infectado (a), siendo ambos monógamos.
- No mantener relaciones sexuales con personas desconocidas.
- Evitar el coito anal.
- Evitar el consumo de drogas por vía intravenosa, y no compartir agujas, jeringuillas u otros utensilios para su administración.
- Utilizar siempre preservativos de látex con espermicida (nonxynol-a) cuando se mantengan relaciones con una persona sobre la que no tengamos suficiente información. Los preservativos disminuyen el riesgo de infección por el VIH y otras enfermedades de transmisión sexual, aunque esta protección no es total.

5.1 USO DEL CONDÓN

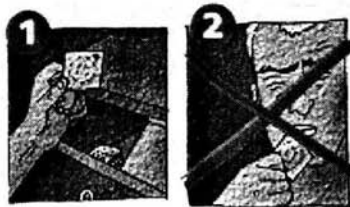
El uso del condón ha desencadenado respuestas polémicas entre los diversos sectores de la sociedad, como en la Iglesia y algunas instituciones educativas. Su uso,

nosotros creemos, es un asunto fundamental de salud pública que depende de su adecuada y constante información de su benéfica utilidad.

El uso correcto del condón podría reducir de manera radical el porcentaje de infecciones por el VIH y otras enfermedades de transmisión sexual, dado que se sabe que el 86.7%¹ de las infecciones por el VIH que se han notificado hasta ahora en México han sido por contacto sexual. Dicha reducción no llegaría a efectuarse sin la información que las instancias gubernamentales deben brindar en relación a su uso.

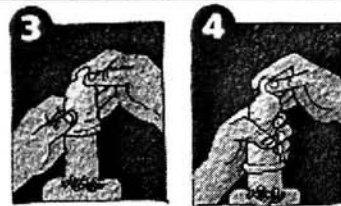
El uso correcto y constante protege al usuario en 95% de contraer la infección por el Virus de la Inmunodeficiencia Humana (VIH). El uso incorrecto y/o la falta de constancia son las principales causas de fallas.

El uso constante y correcto del condón es una de las estrategias recomendadas por la Organización Mundial de la Salud para prevenir el VIH/SIDA, otras Enfermedades de Transmisión Sexual y el embarazo. Esto significa usarlo cada vez que se tengan relaciones sexuales, desde el inicio hasta el final de la misma con penetración, ya sea anal, vaginal u oral. A continuación se muestra gráficamente los pasos que deben seguirse para la correcta colocación del condón:



1 Guarde el condón en lugares frescos y secos. Verifique la fecha de caducidad o que la fecha de fabricación no pase de tres años

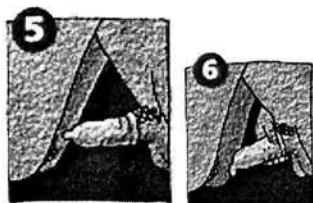
2 Fijese que el empaque no esté roto y que al oprimirlo forme una bolsa de aire. No lo abra con los dientes



3 Presione la punta para sacar el aire mientras lo coloca en el pene erecto.

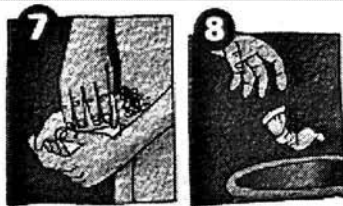
4 Sin soltar la punta, desenróllelo hasta llegar a la base del pene.

¹ Datos obtenidos del Reporte Especial de la Gaceta CCH, de la Universidad Nacional Autónoma de México, del día 3 de diciembre de 2001.



Siempre colóquese el condón antes de cualquier contacto sexual. No use cremas cosméticas, ni vaselina como lubricantes, porque deterioran el condón.

Después de eyacular, sujete el condón por su base, retire el pene mientras aún está erecto.



Quítese el condón cuidando que no se goteé.

Tire el condón a la basura anudándolo

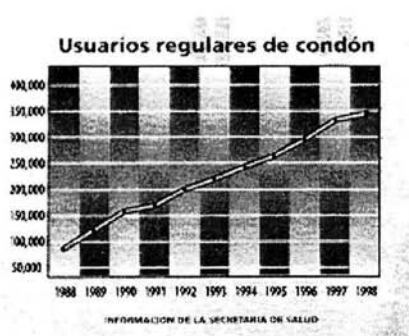
El condón puede fallar por tres razones, y éstas son, en orden de frecuencia:

- uso inconstante (no se usa en todas las relaciones sexuales);
- uso incorrecto (no se usa desde el principio hasta el final de la relación sexual ni se cuidan los pasos que ayudan a evitar fallas); y
- por deslizamiento (que el condón se salga completamente del pene) y/o ruptura durante la relación sexual o al retirar el pene.

La falta de constancia en el uso del condón se debe a circunstancias de tipo económico, cultural y/o psicológico. El uso incorrecto es atribuible a la falta de habilidad de la persona para manejar adecuadamente el condón. El deslizamiento y/o ruptura puede ocurrir por cometer algún error al colocarlo y, en menor medida, por el grosor del pene o por el tamaño del condón en relación con éste.

El condón es una excelente opción para prevenir Enfermedades de Transmisión Sexual/VIH y embarazo, tanto para parejas casadas o unidas, hombres y mujeres adultos no casados, personas con múltiples parejas, personas con parejas sexuales ocasionales, adolescentes, hombres que tienen relaciones con hombres, hombres que se relacionan sexualmente con mujeres, hombres y mujeres dedicados a la prostitución y sus clientes, y para prácticas sexuales anales, orales o vaginales. Lo importante es protegernos de consecuencias no deseadas.

Por otra parte, en México, el uso del condón se ha incrementado sustancialmente año con año -según datos de Secretaría de Salubridad y Asistencia- sin embargo es importante que el personal de salud se involucre en acciones que coloquen al condón como una medida de salud pública, más que nada para su buen uso. Obsérvese la gráfica siguiente:



5.2 ABSTINENCIA

La abstinencia es uno de los métodos de prevención menos usados por las personas sexualmente activas aunque su eficacia sea del 100%. Muchas de las personas que recurren a este tipo de prevención practican lo que se conoce “como el sexo seguro”, que consiste en mantener toda práctica sexual sin recurrir a la penetración y al intercambio de fluidos biológicos.

5.3 FIDELIDAD MUTUA²

Consiste en mantener todo tipo de relaciones sexuales con una sola persona. Nosotros pensamos que esta forma de prevención es la más insegura e inadecuada, ya que uno puede practicarla con una persona que previamente ha sido infectada. Debemos de mencionarlo puesto que se sabe que una persona puede vivir normalmente durante aproximadamente ocho años sin darse cuenta de su infección. Esto debido a que el virus tarda en incubarse hasta ocho años, en la mayoría de los casos.

² Nosotros creemos que “Fidelidad Mutua” pudiera desatar confusión en nuestra sociedad ya que el nombre parece garantizar la no infección.

6. EL SIDA EN MÉXICO

6.1. INTRODUCCIÓN

En México, se han implantado acciones en lucha contra el SIDA, tales como: la información a toda la comunidad respecto a los medios y mecanismos de la infección; la búsqueda de la eliminación de estereotipos y de falsas ideas del sentido común del mexicano; buscar una educación la cuál pueda despertar una conciencia social positiva del problema y para promover solidaridad social con el enfermo, con el infectado y con sus familiares.

En realidad en nuestro país, el problema con respecto a lo que se refiere al porcentaje de población infectada no es tan grave comparado con otros países, como lo son por ejemplo: Estados Unidos y Haití, los cuáles tienen un gran número de gente infectada de SIDA dentro de su población. No por esto tenemos que pensar que en nuestro país no se deban tomar las medidas pertinentes con respecto a su propagación. Por tal motivo, se ha creado el CONASIDA, y a su vez ésta fue creada para brindar apoyo y asesoría a todo aquella persona que lo solicite.

6.2. CONASIDA

El Consejo Nacional para la Prevención y Control del SIDA (CONASIDA) fue creado mediante decreto presidencial publicado en el Diario Oficial de la Federación el día 24 de agosto de 1988. A partir de ese momento las actividades del CONASIDA se financiaron con recursos provenientes de fondos internacionales, principalmente del Programa Global de SIDA de la Organización Mundial de la Salud (OMS) y en 1991 le fue asignado presupuesto federal ante la progresiva e importante disminución del financiamiento internacional y el crecimiento de la epidemia. Este presupuesto específico quedaba dentro del presupuesto operativo de la Dirección General de Epidemiología. En 1992 a CONASIDA le fue asignado presupuesto federal propio (clave presupuestal KOO).

A partir del 22 de diciembre de 1994 el CONASIDA quedó incluido en la estructura de la Dirección General de Epidemiología, dependiente de la Subsecretaría de Servicios de Salud.

Actualmente el CONASIDA es un órgano desconcentrado de la Secretaría de Salud con nivel de Dirección General y cuenta con plazas federales propias y depende directamente de la Subsecretaría de Prevención y Control de Enfermedades.

El programa de prevención y control del SIDA se plantea cuatro objetivos principales:

1. Prevención de la transmisión sanguínea, sexual y perinatal, a través de diversas estrategias como la promoción de medidas preventivas, la detección oportuna y el tratamiento adecuado de otras enfermedades de transmisión sexual, el control y la

detección del VIH en todo paquete de sangre y hemoderivados, y la prestación de asesoría y de servicios a la población con prácticas de riesgo.

2. Reducción del impacto individual y social: fortalecer y promover las acciones que favorezcan atención médica y apoyo psicológico y social adecuado a las personas afectadas y a sus familiares; promover acciones que disminuyan el impacto social y económico del VIH/SIDA y que favorezcan el respeto a los derechos humanos, así como el rechazo a la discriminación.
3. Movilización y unificación de esfuerzos nacionales: movilizar al Sistema Nacional de Salud, al resto del sector público y a la sociedad en su conjunto en la lucha contra el VIH/SIDA.
4. Investigación: promover y realizar actividades de investigación para conocer las características particulares de la enfermedad en nuestro país desde el punto de vista epidemiológico, biomédico y psicosocial.

6.3 MITOS DEL S. I. D. A. EN NUESTRO PAÍS

Desde la aparición del VIH en nuestro país la sociedad ha creado un conjunto de mitos que son originados por la ignorancia del tema; esa ignorancia aún existe y ha sido aminorada en gran medida por las campañas informativas de gobierno, que de alguna manera se ha preocupado por mantener un cierto nivel de conciencia social.

El creer que el virus del S. I. D. A. es exclusivo de los homosexuales hombres es un mito que apareció en el mismo momento que se dieron a conocer los primeros casos de S. I. D. A., ya que estos se presentaron en personas que pertenecen a dicho grupo sexual. Esta idea no ha sido erradicada por completo y ha sobrevivido por la gran aberración que existe hacia ese grupo sexual en nuestra sociedad.

Muchos mitos se tienen en relación a las vías de transmisión del virus, haciendo aún más difícil la vida de las personas que son portadoras, pensando que cualquier contacto con estas personas pudiera ser un mecanismo de contagio, ejemplos de ellos son:

- Un simple beso
- Utilizar su cepillo de dientes
- Utilizar el sanitario que ocupa
- Hacer contacto con sus prendas u objetos personales
- Mantener un contacto físico, como podría ser un abrazo o un saludo de mano.

Todas las ideas anteriores son tajantemente inválidas, ya que afortunadamente contamos con la información suficiente para saber que solamente existen tres formas principales de contagio: la vía sexual, la vía intravenosa y la vía perinatal.

El uso del condón es una medida preventiva contra el virus del VIH, su uso ha sido promovido también como un método anticonceptivo, esto ha originado en algunas personas una gran confusión: pensar que los métodos anticonceptivos en general pueden actuar como una medida de prevención en contra del virus del VIH.

La fidelidad mutua entre la pareja sexual ha sido también promovida como una medida de prevención contra el virus del VIH. Sin embargo, debemos tomar en cuenta que una pareja sexual puede mantenerse completamente fiel sin saber que alguno de los miembros que la conforman es portadora de dicho virus; dicha persona puede contagiar a su pareja e iniciar después de cierto periodo de tiempo cada una de ellas otra relación íntima utilizando la misma medida preventiva. La fidelidad mutua no garantiza la no infección, ya que una persona que desconozca que esté infectada del virus puede mantener completamente fiel sus relaciones de pareja sexual que pudiera realiza durante un considerable número de años, y sin que se presente algún síntoma del virus en su organismo.

CAPÍTULO IV

PROPUESTA Y VALIDACIÓN DEL MODELO

Objetivo: Proponer un modelo de pronóstico obtenido del análisis de datos realizado mediante la Metodología de Box-Jenkins; realizar la optimización y validación de dicho modelo, aplicando la misma metodología.

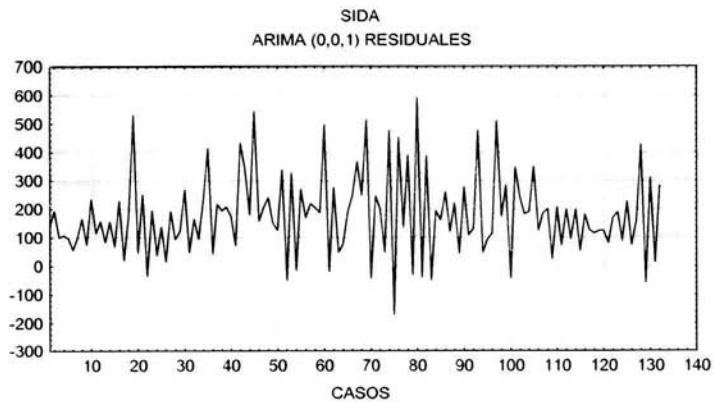
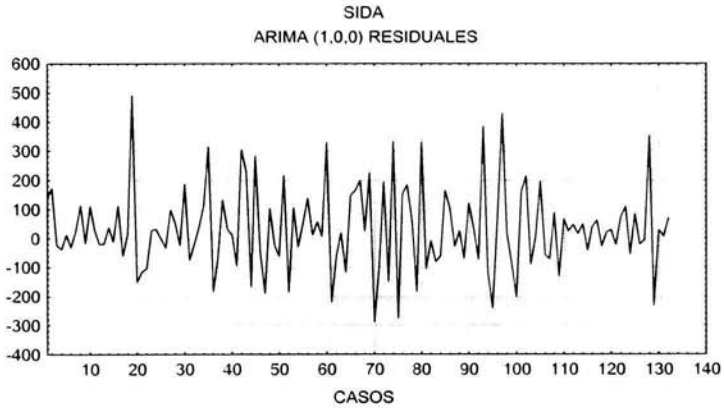
1. DIAGNÓSTICO DE LA ADECUACIÓN DEL MODELO

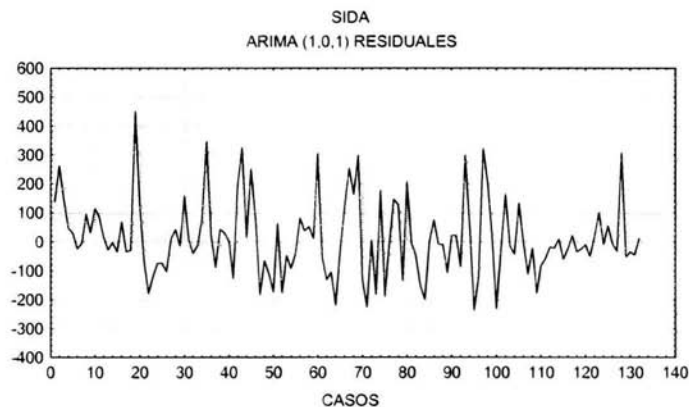
Una vez que se han analizado y observado las gráficas de la ACF y las gráficas de la PACF de los posibles modelos de pronóstico que se adecuan a los datos muestrales utilizados, proponemos como modelos de pronóstico más adecuados a AR (1), MA (1) y a un ARMA (1,1); esto también siguiendo el principio de parsimonia.

Más adelante, en este mismo capítulo, se analizarán los residuales de dichos modelos propuestos para elegir el modelo más adecuado y posteriormente se hará la validación de dichas propuestas de los modelos de pronóstico. Dicha validación se hará retomando la información que se mencionó en el capítulo II referente a los casos de S. I. D. A. notificados en el año 2001 y compararlos con los pronósticos proporcionados por el modelo elegido.

2. GRÁFICA DE LOS RESIDUALES

La gráfica de los residuales para cada uno de los modelos propuestos es:





Al analizar las gráficas de los residuales se observa que:

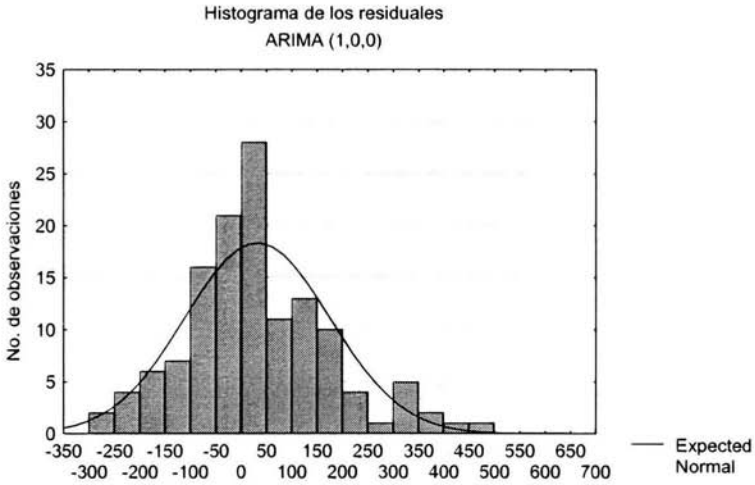
- Existe una media constante.
- Los residuales se centran alrededor de cero, suponiendo, por lo tanto, media igual a cero, esto únicamente para los residuales de los modelos AR(1) y ARMA(1,1); en el caso de los residuales del modelo MA(1) se observa una media que pareciera ser distinta de cero.
- La varianza es constante, ya que la gráfica no muestra una tendencia creciente o decreciente; es a lo que nosotros le llamamos que no existe heteroscedasticidad.
- Es importante realizar este tipo de análisis ya que al observar una media significativamente distinta de cero en los residuales (caso de los residuales del modelo MA (1)), implicaría que existe una parte determinista que no ha sido tomada en cuenta en el modelo propuesto.

Por los puntos que conforman el análisis de las gráficas de los residuales de los modelos sugeridos, tendremos en tela de juicio al modelo MA (1).

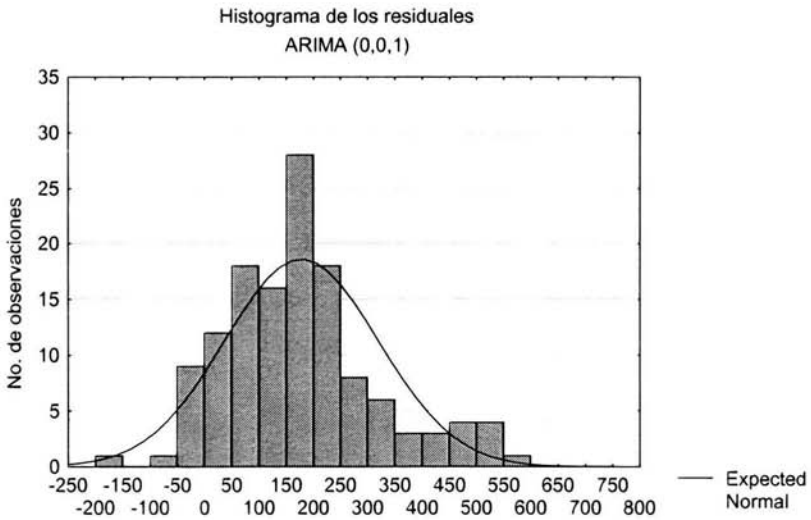
Ahora veamos el histograma de los residuales para cada uno de los modelos sugeridos:

Ahora veamos, el histograma de los residuales para cada uno de los modelos

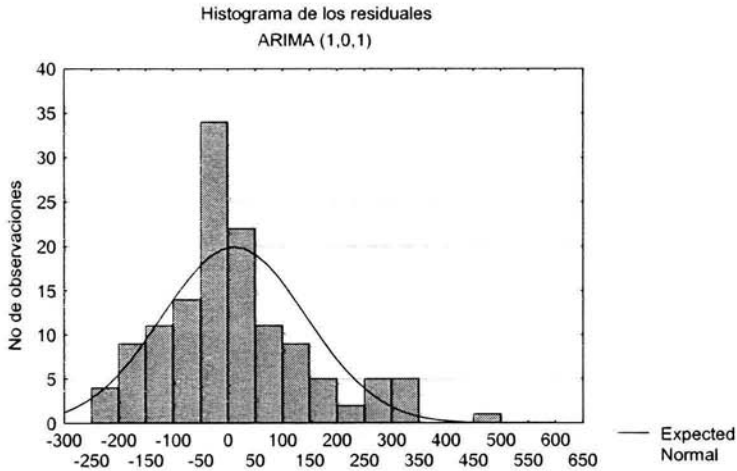
3. HISTOGRAMA DE LOS RESIDUALES



Se puede observar, que no existen datos anómalos, ya que cualquiera de ellos, se encuentran en (-3, 3)



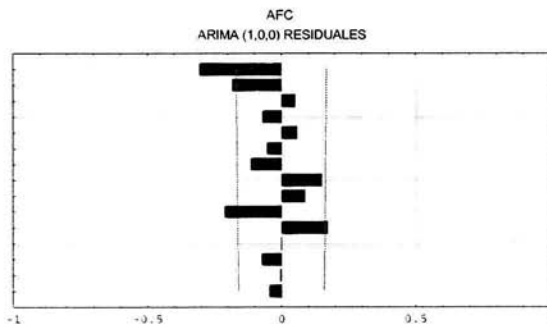
Es evidente que el modelo MA (1), no se adapta a nuestro datos muestrales, por tal motivo hemos decidido descartarlo. Cheque como la gráfica muestra datos que existen fuera de (-3,3).

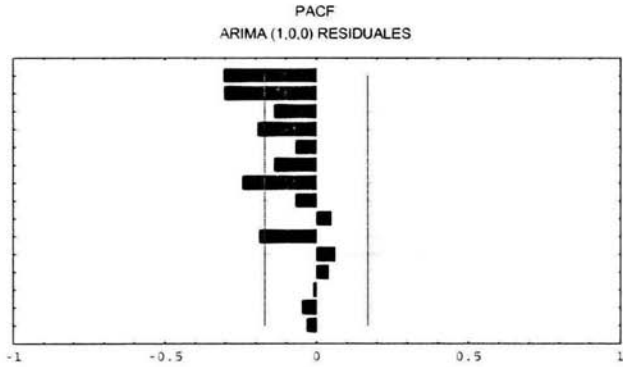


El mismo caso que el histograma de los residuales de AR (1), ya que tampoco existen observaciones “sospechosas” que estén fuera del intervalo (-3,3).

4. LA ACF Y LA PACF DE LOS RESIDUALES DE CADA UNO DE LOS POSIBLES MODELOS

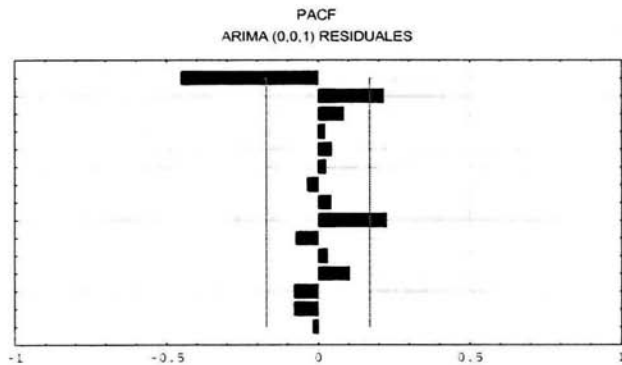
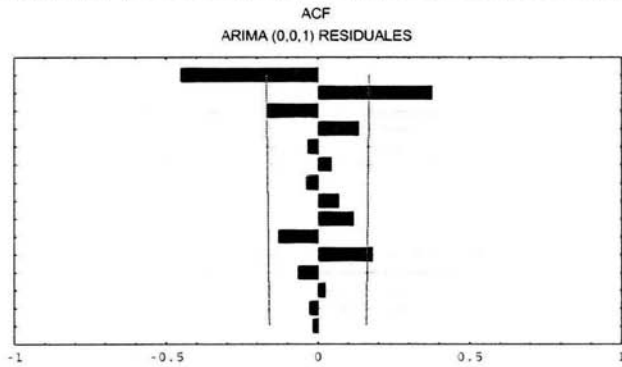
Ahora observemos las gráficas de la ACF y PACF de los residuales de cada uno de los modelos de nuestra propuesta:





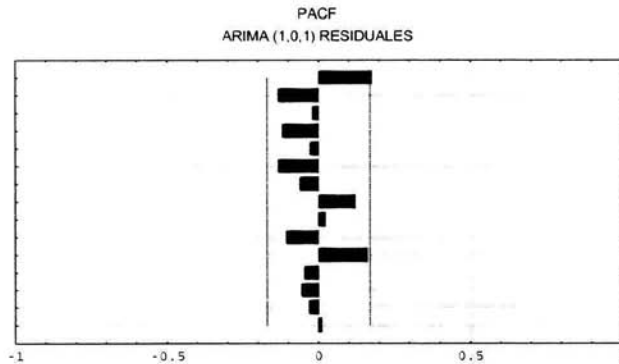
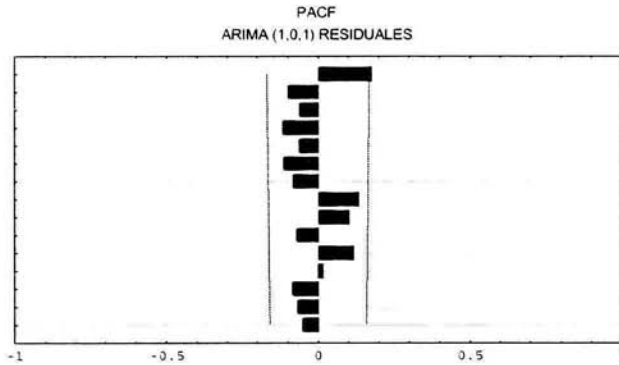
Tanto en la ACF y PACF de los residuales de este modelo podemos observar que no precisamente se están comportando como ruido blanco.

Observemos ahora la ACF y la PACF de los residuales del modelo MA (1):



Las gráficas de la ACF y de la PACF de los residuales del modelo MA (1), parecen no comportarse como ruido blanco, por lo tanto, hay suficiente evidencia para descartar dicho modelo de nuestra propuesta, pero seguiremos tomándolo en cuenta para sus comparación con los otros.

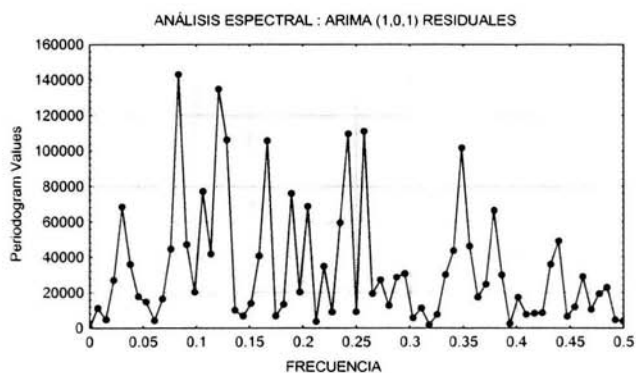
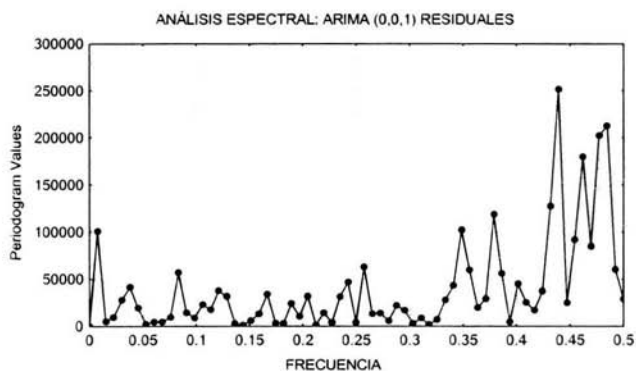
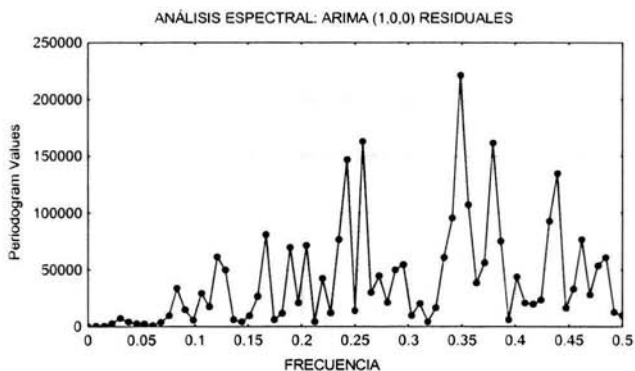
Analicemos ahora la ACF y la PACF de los residuales del modelo ARMA (1,1):



Las gráficas nos muestran un clásico comportamiento de ruido blanco en los residuales, por lo que concluimos hasta ahora, que el modelo que más se adecua a nuestros datos es ARMA (1,1).

5. PERIODOGRAMA DE LOS RESIDUALES

Veamos a continuación el periodograma de los residuales

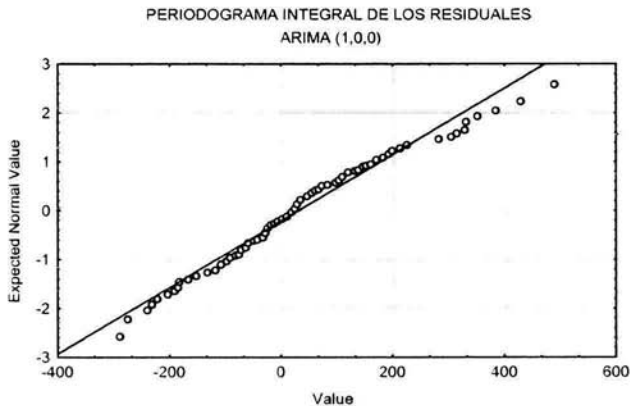


Lo que observamos en estas gráficas es que en ninguna de ellas se observan picos que pudieran considerarse que el modelo esté subespecificado. Esto significa que cualquiera de los tres modelos pudiera ser utilizado para pronosticar; sin embargo, analizaremos todo el examen y la validación de los tres para decidimos por uno de ellos.

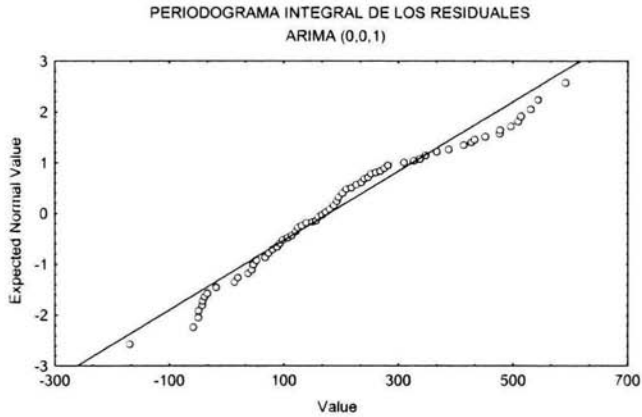
También observemos que de los tres el que menos eleva sus picos es el periodograma integral de los residuales del modelo ARMA (1,1).

6. PERIODOGRAMA INTEGRAL DE LOS RESIDUALES

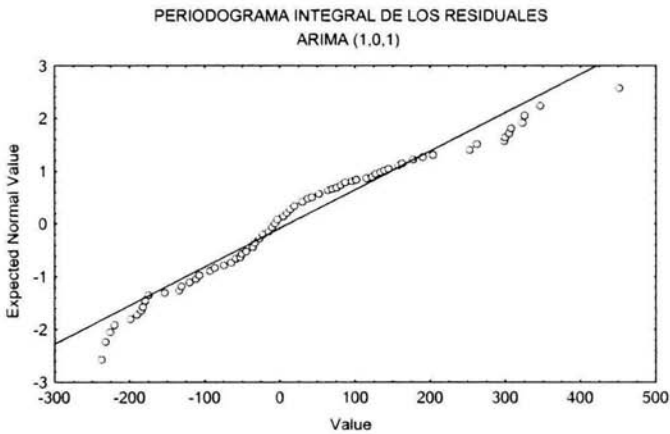
A continuación se analizan los periodogramas integrales de los residuales



El comportamiento del periodograma integral se ajusta al problema o al fenómeno bajo estudio, pero el modelo se comparará con los resultados obtenidos con los demás modelos que funcionan como alternativas.



El periodograma nos muestra un menor ajuste al modelo que el caso de alternativa anterior.



Un buen ajuste es lo que podemos observar en el periodograma integral.

7. MODELO DE PRONÓSTICO ELEGIDO

Recordando y comparando todos los análisis y observaciones que se hicieron con los modelos tentativos, seleccionamos al modelo **ARMA (1,1)** como el modelo que mejor se ajusta a nuestros datos.

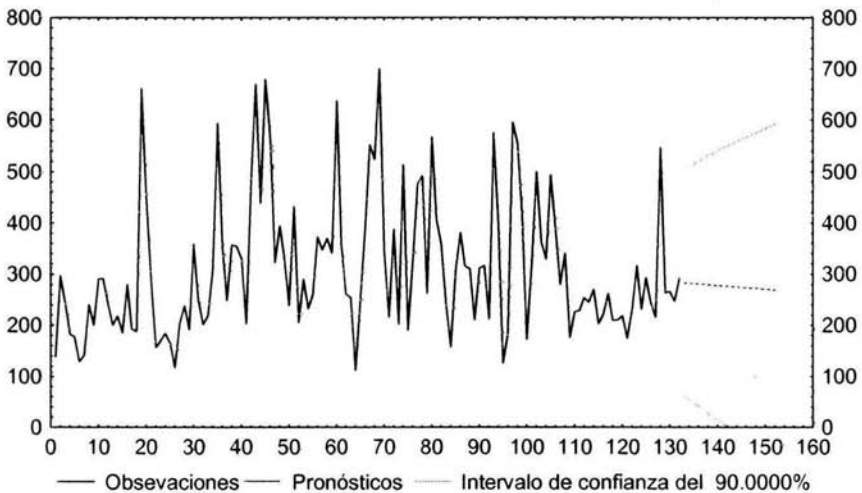
La estimación de los valores de los parámetros se muestra a continuación:

PARÁMETRO	VALOR
$p(1)$	0.997500
$q(1)$	0.747550

8. PRONÓSTICOS

Se generarán 20 pronósticos que corresponderán a los meses de ENERO DEL 2001 a AGOSTO DEL 2002. Será importante mencionar que los pronósticos generados por el modelo elegido se ajuste a las observaciones reales que ya conocemos, lo anterior debido a que ya conocemos el número de casos por mes de los meses que corresponden al año 2001, esto implicaría generar pseudo-pronósticos, más adelante mencionaremos su comparación y su conclusión.

CASOS NOTIFICADOS DE SIDA POR MES EN MÉXICO
Y 20 PRONÓSTICOS GENERADOS POR EL MODELO ARMA (1,1)



Los pronósticos son los siguientes, se muestra la estimación puntual y los límites del intervalo de confianza del 90 % de confiabilidad:

No. DE OBSERVACIÓN	PRONÓSTICO	LÍMITE INFERIOR	LÍMITE SUPERIOR
133	282	62	503
134	282	55	509
135	281	47	515
136	280	40	520
137	280	33	526
138	279	27	531
139	278	20	536
140	277	14	541
141	277	8	546
142	276	2	550
143	275	-4	555
144	275	-10	559
145	274	-16	564
146	273	-21	568
147	273	-27	572
149	272	-32	576
150	271	-37	580
151	271	-43	584
152	270	-48	588
153	269	-53	591

Se observa que hay una ligera tendencia decreciente en los pronósticos, esto significa que de alguna manera, las medidas que hasta ahora se han tomado en relación al problema han sido efectivas

Los pronósticos que presentamos corresponden a los casos notificados desde el mes de ENERO DEL 2001 hasta AGOSTO DEL 2002. Meses que ya han transcurrido y donde se sabemos lo que ocurrió, la información correspondiente a los meses ENERO DEL 2001 A DICIEMBRE DEL 2001 aunque se cuenta con la información no fue utilizada para obtener el modelo de pronóstico y se utilizará para compararla con los pronósticos obtenidos:

MES	NÚMERO DE CASOS (DATOS REALES)	PRONÓSTICO	INTERVALO DE CONFIANZA DEL 90%
ENERO 2001	0	282	62 - 503
FEBRERO 2001	6	282	55 - 509
MARZO 2001	20	281	47 - 515
ABRIL 2001	49	280	40 - 520
MAYO 2001	88	280	33 - 526
JUNIO 2001	151	279	27 - 531
JULIO 2001	59	278	20 - 536
AGOSTO 2001	121	277	14 - 541
SEPTIEMBRE 2001	164	277	8 - 546
OCTUBRE 2001	202	276	2 - 550
NOVIEMBRE 2001	104	275	0 - 555
DICIEMBRE 2001	206	275	0 - 559

Los datos reales a excepción de los casos de enero del 2001 a marzo del mismo año caen dentro del intervalo de confianza, podemos concluir que el modelo ARMA (1, 1) nos genera pronósticos aceptables.

CAPÍTULO V

DECISIONES PARA CONTROL

Objetivo: Realizar propuestas para controlar el porcentaje de población infectada de S. I. D. A. en México utilizando el modelo de pronóstico y la información obtenida de las estadísticas realizadas.

En el presente capítulo se darán decisiones para monitorear el control de porcentaje de personas infectadas de S. I. D. A. en México.

Esta parte es muy importante ya que nuestro trabajo no tiene como único objetivo el de proporcionar los pronósticos, sino también monitorear las posibles situaciones en que la gráfica de casos relacionados con personas infectadas se reduzca.

El procedimiento que se realizará para tal efecto se menciona a continuación:

1. Realizar un muestreo para la obtención de datos muestrales.
2. Capturar en un paquete estadístico la información recopilada.
3. Calcular porcentajes y estadísticas.
4. Generar pronósticos utilizando el modelo de pronóstico elegido mediante la metodología de Box-Jenkins.
5. Retomar los pronósticos obtenidos y modificarlos de acuerdo a los porcentajes obtenidos en las encuestas.
6. Analizar cada aspecto (formas de prevención, grupos de riesgo y formas de contagio), para monitorear y sugerir que medidas son las más apropiadas, no para estabilizar (porque de hecho ya lo está), sino para observar una tendencia decreciente en la gráfica de la serie de tiempo con respecto a la población infectada de S. I. D. A en México.

I. OBTENCIÓN DE DATOS MUESTRALES

La técnica que se empleó en la investigación de campo para la obtención de la información es el cuestionario. Se acordó que fuese esta técnica, ya que la encuesta, por ejemplo, traería consigo la incomodidad de la gente seleccionada al tratar de contestar preguntas como: ¿A qué grupo sexual perteneces? ¿Cómo te proteges contra el virus del S. I. D. A.? etc. Se pensó también en esta técnica (el cuestionario) ya que con esto se evitaría la presión de nosotros referente al tiempo. La persona sería libre de pensar y analizar sus posibles respuestas, así como la de confiar en nosotros para poder aclarar cualquier duda referente al conjunto de preguntas que se harían. Y lo más importante, pensamos en lo confidencial de la información proporcionada al utilizar dicha técnica.

El cuestionario que se aplicó para obtener la información se muestra a continuación; en el se trató de encontrar cuáles eran los conocimientos que tiene la sociedad en general con respecto al virus del S. I. D. A., tratando de detallar esos conocimientos en formas de prevención y formas de contagio que conoce; se trató de averiguar cuáles eran los mecanismos que utilizan para protegerse de dicho virus y por último cuáles son los grupos de alto riesgo que frecuentan en su vida sexual.

Las preguntas que se hicieron fueron formuladas utilizando palabras sencillas, esto para no causar confusión a las personas que lo contestaran. Sin embargo, no lo logramos por completo, ya que aproximadamente el 70% de las personas que contestaron el cuestionario no comprendió el término heterosexual, preguntándonos por el significado del mismo.

CUESTIONARIO

El siguiente cuestionario se desarrolló para ser utilizado en un estudio estadístico que se realiza en la Universidad Nacional Autónoma de México Campus Acatlán. La información que nos proporcione será estrictamente anónima y confidencial, por lo que le rogamos conteste de la manera más sincera posible. Gracias.

1. Localidad: _____ Estado: _____
2. Sexo: Femenino() Masculino()
3. Edad: _____
4. Grupo sexual al que pertenece:
a. heterosexual b. homosexual c. bisexual
5. ¿Usted ha oído hablar del VIH y las formas de prevenir y contraer su infección?
i. Sí ii. No
6. Si su respuesta anterior fue afirmativa, mencione las formas de prevenir y contraer la infección del virus del VIH:

7. ¿Qué medidas de prevención utiliza contra el virus del SIDA?
a. Abstinencia b. Condón (preservativo) c. fidelidad mutua
d. Otros, menciónelos: _____
8. ¿Usted tiene algún contacto sexual con alguien que forme parte de alguno o algunos de los siguientes grupos?
a. Homosexuales hombres b. Prostitutas
c. Drogadictos que utilizan drogas inyectables d. Ninguno de los anteriores

El muestreo se realizó de manera aleatoria entre personas que considerábamos capaces de contestar el cuestionario; decidiendo que como edad mínima para ser tomado en consideración la de 12 años.

Como ya se mencionó el muestreo que se utilizó para recabar la información fue de manera aleatoria, enseguida explicaremos en que consiste dicho tipo de muestreo.

Muestreo aleatorio. Este muestreo consiste en recabar la información al considerar a un subconjunto de elementos de la población (muestra) y considerando que todos los elementos de la población tienen la misma probabilidad de ser seleccionados (esto se logra al no hacer preferencias o discriminaciones al realizar la investigación de campo). Con esto hacemos que el subconjunto sea heterogéneo y considere todas las características de la población, logrando con esto que nuestro subconjunto sea representativo de ella.

En específico, nosotros realizamos un muestreo de 907 personas (tamaño de la muestra) considerando únicamente el conocimiento que tenemos de nuestra población; así como de los pocos recursos con los que contábamos para efectuarlo. Esto último fue nuestro principal problema para pensar que estos podrían considerarse de alguna manera representativos de la población, para efectuar nuestras inferencias. Para tratar de resolver el problema (de los recursos con los que contábamos) recurrimos al muestreo de juicio, el cual es un tipo de muestreo que justifica de alguna manera el problema de los recursos. Este muestreo consiste en hacer una selección de los elementos que conformarán a la muestra (muestreo no aleatorio) mediante el conocimiento que se tiene de la población, sin embargo, esto trae consigo un problema mayor: resulta ser peligroso tratar de realizar deducciones o inferencias a partir de un estudio realizado de esta manera ya que al seleccionar de una manera no aleatoria los elementos que conformarán la muestra estaríamos forzando de alguna manera los resultados que se obtendrán.

Con este problema tratamos de realizar una investigación de campo la cual traerá consigo una recopilación de información lo más representable de la población y confiable posible; para ello se buscaron lugares estratégicos; de tal forma, que encuestáramos a personas de diferente estrato social o económico y que vivieran o radicaran en cualquier comunidad de los diferentes estados de la República; por supuesto, sin importar el tipo de comunidad que fuese (rural o urbana), género, edad, grado máximo de estudios, etc.; ya que las inferencias se realizaron no a un grupo específico de personas, sino a toda la población, en este caso, la población se definiría como: mexicanos radicados en el territorio nacional con edades entre 12 y 80 años de edad; este intervalo de edades pensado en que las personas que se encuentren en él, de una forma u otra han oído hablar del tema.

Los lugares donde se realizaron las encuestas son:

- *Zócalo capitalino.* Este lugar fue pensado como lugar estratégico para realizar nuestras encuestas, ya que en él podemos encontrar cualquier estrato social y económico de mexicanos que radiquen no solo en la zona metropolitana sino de cualquier lugar de la república mexicana.

- *TAPO, localizada en el D F.* Algunas de las encuestas fueron realizadas aquí, ya que nosotros pensamos que en dicho lugar podemos encontrar cualquier estrato social y económico de diferentes estados de la república mexicana localizados en la zona oriente, como por ejemplo: Oaxaca, Guerrero, Chiapas, Veracruz, etc.
- *Autobuses del Norte, localizada en el D F.* Este lugar fue de gran ayuda para nuestro muestreo, ya que en el encuestamos a personas que radicaban en distintos estados de la zona norte, como por ejemplo: Guanajuato, Tamaulipas, Nuevo León, etc.
- *San Lucas Tecopilco, Tlaxcala.* Necesitábamos una comunidad rural que representará de alguna forma a las demás comunidades que existen en nuestro país. Nosotros acordamos como lugar para realizar a nuestras encuestas esta comunidad, porque pensamos que nos podría aportar una información muy interesante debido a que en Tlaxcala se encuentra el mayor porcentaje de casos de S I D A en México.
- *Celaya, Guanajuato.* Ciudad considerada como industrial; donde en ella se pueden observar las grandes diferencias socio-económicas en la gente, por esta razón y aunado a que se tenían los recursos para visitarla, la seleccionamos para que en ella se realizaran algunas encuestas.
- *Naucalpan, Estado de México.* Este lugar fue considerado debido a que en esta zona la realización de las encuestas se nos iba a facilitar y no necesitábamos muchos recursos para ello.
- *Puebla, Puebla.* Esta zona fue escogida por su cercanía al D F y por pensar que esta ciudad puede caracterizar a las demás en ciertos aspectos importantes por el contenido de nuestras encuestas como lo son: la información que se ha venido generando en relación al tema, la infraestructura con la que se cuenta, etc.

Nomenclatura utilizada en la captura de información.

Hemos utilizado la siguiente notación para representar las diferentes categorías empleadas en el cuestionario aplicado

Entidad Se refiere a la entidad federativa, o sea, al Estado de la República Mexicana y D F en el cual se aplicó el cuestionario

Tipo de comunidad Representado como **TIPO_COM**, es el lugar donde se aplicó el cuestionario:

- U Urbana
- R Rural

Género Es el sexo al que representa la persona encuestada:

- Masculino**
- Femenino**

Edad Se muestra en años, destacando el intervalo de 12 a 81 años

Grupo sexual A qué clase de personas pertenece o frecuenta, sexualmente hablando En el cuadro de datos se marca como **A, B, C**, respectivamente

- a. Heterosexual**
- b. Homosexual**
- c. Bisexual**

Información básica sobre SIDA Esto hace mención al hecho de que la persona conozca o esté enterado acerca del SIDA y lo que es Se representa como INF_SIDA en el cuadro La mayor cantidad de respuestas relacionadas con las siguientes situaciones hacen creer que la persona sí ha oído hablar del VIH y las formas de contraer y prevenir Si este es el caso, se escribe S En caso contrario, una N

- ABSTINEN** Abstinencia total
- FIDELI** Fidelidad
- OTROS**
- REL_SEX** Relaciones sexuales
- TRAN_SAN** Transfusiones sanguíneas
- PERINATA** El contagio de la madre embarazada a su hijo

Medidas utilizadas Trata acerca de las precauciones que toman las personas encuestadas para prevenir el SIDA Se representa como MED_UTI y en el cuadro de datos se usa como **A, B, C, D**, respectivamente:

- a. Abstinencia**
- b. Condón**
- c. Fidelidad Mutua**
- d. Si practica otros, que los mencione**

Grupo de riesgo Denotado como G_RIESGO, hace mención a las personas que tienen que tratar sexualmente los encuestados, utilizando en el cuadro ya mencionado las respuestas **A, B, C, D**, respectivamente:

- a. Homosexuales hombres**
- b. Prostitutas**
- c. Drogadictos que utilizan drogas inyectables**
- d. Ninguno de los anteriores**

2. ESTADÍSTICAS GENERADAS CON LOS DATOS MUESTRALES

Anteriormente se mencionó cuales habían sido las características del muestreo realizado para la obtención de datos muestrales; es por esta razón que nosotros podemos inferir hacia la población con estos resultados muestrales.

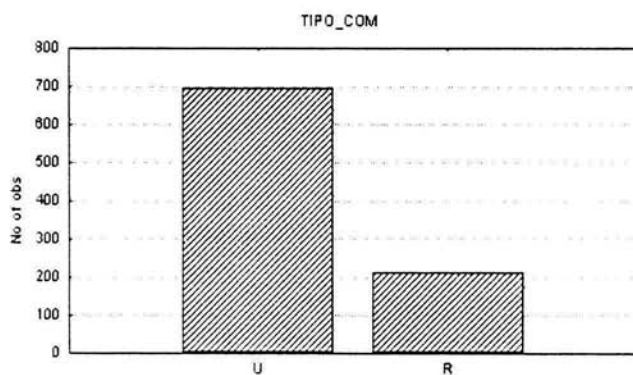
Las observaciones se hicieron en lugares estratégicos considerando los recursos que se tenían para realizar el muestreo, trayendo consigo los siguientes porcentajes con respecto al número de personas encuestadas que radican en las diferentes entidades del país:

ENTIDAD	NÚMERO DE PERSONAS	PORCENTAJE
MICHOACAN	3	0.33
TLAXCALA	66	7.28
VERACRUZ	11	1.21
GUERRERO	4	0.44
GUANAJUATO	111	12.24
TAMAULIPAS	1	0.11
COAHUILA	1	0.11
NUEVO LEON	1	0.11
BAJA CAL.	2	0.22
ZACATECAS	1	0.11
AGUASCALIENTES	1	0.11
COLIMA	1	0.11
CHIAPAS	5	0.55
D. F.	308	33.96
DURANGO	1	0.11
EDOMEX	331	36.49
HIDALGO	4	0.44
JALISCO	2	0.22
NAYARIT	1	0.11
OAXACA	2	0.22
PUEBLA	36	3.97
QUERETARO	4	0.44
S. L. P.	3	0.33
SONORA	1	0.11
TABASCO	5	0.55
YUCATAN	1	0.11
NULOS	0	0.00
TOTAL	907	100.00

Con respecto al tipo de comunidad tenemos las siguientes estadísticas:

TIPO_COM	NO. DE PERSONAS	PORCENTAJE
U	694	76.52
R	213	23.48
SIN CONTAR	0	0.00
TOTAL	907	100.00

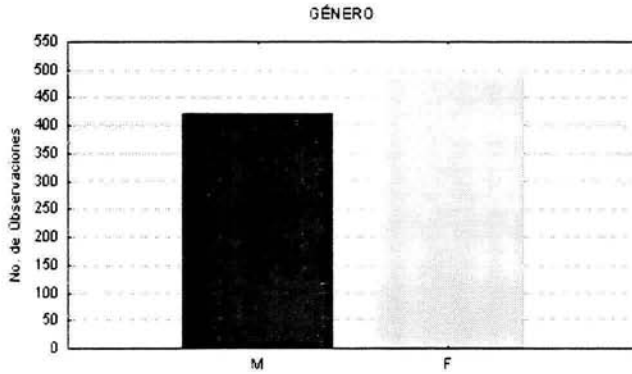
La gráfica de barras que se genera es la siguiente:



El género es una variable booleana bastante interesante, veamos qué estadísticas arrojaron nuestros datos muestrales:

GÉNERO	NO. DE PERSONAS	PORCENTAJE
M	422	46.53
F	485	53.47
SIN CONTAR	0	0.00
Total	907	100

La gráfica de barras que representa la situación es la siguiente:



Podemos mencionar como conclusión estadística que se estima que el 46.53% de las personas con edades entre 12 y 81 años pertenecen al sexo masculino y el 53.47% al femenino.

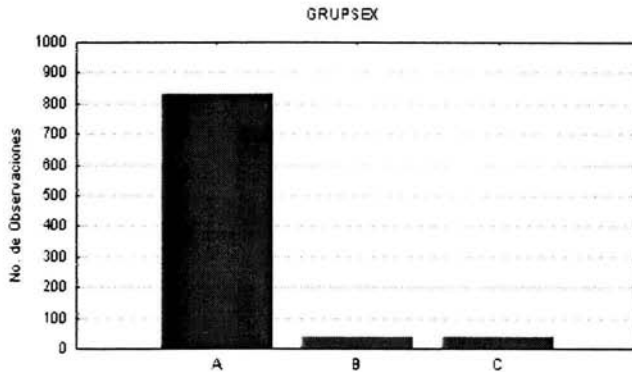
Las estadísticas generadas con respecto a la variable EDAD se mencionan a continuación:

	MEDIA MUESTRAL	DESVIACIÓN ESTÁNDAR	MÍNIMO VALOR CONSIDERADO	MÁXIMO VALOR CONSIDERADO	TAMAÑO DE LA MUESTRA	SIN CONTAR
EDAD	27.0396913	12.048127	12	81	907	0

Las estadísticas generadas con respecto a la variable GRUPO SEXUAL es la siguiente:

GRUPO SEXUAL	No. DE PERSONAS	PORCENTAJE
A	832	91.73
B	39	4.30
C	36	3.97
SIN CONTAR	0	0.00
TOTAL	907	100

La gráfica de barras que representa la información es la siguiente:



Como conclusión estadística y tomando como base la estimación puntual podemos establecer los siguientes intervalos de confianza del 95%:

- Se estima que entre 89.94% y el 93.52% de las personas que viven en el territorio nacional son heterosexuales.
- Se estima que entre 2.98% y el 5.62% de las personas que viven en el territorio nacional son homosexuales.
- Se estima que entre el 2.7% y el 5.24% de las personas que viven en el territorio nacional son bisexuales.

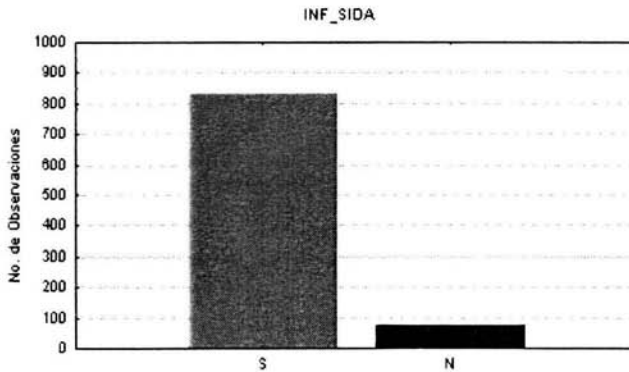
La interpretación de dicha conclusión estadística es en relación al porcentaje de homosexuales que existen, sabemos que es un grupo sexual que es considerado foco de infección y por lo tanto es benéfico que el porcentaje obtenido sea relativamente reducido. Aunque hay que hacer énfasis de que la enfermedad no es exclusiva de algún grupo sexual, social o de cualquier otro índole.

La variable que inicia es de acuerdo con el conocimiento de la gente referente con la palabra SIDA, en general se preguntó lo siguiente: “¿USTED HA OÍDO HABLAR DEL VIH Y LAS FORMAS DE PREVENIR Y CONTRAER SU INFECCIÓN?”

La información que se generó es la siguiente:

INF_SIDA	NO. DE PERSONAS	PORCENTAJE
S	830	91.51
N	77	8.49
SIN CONTAR	0	0.00
TOTAL	907	100

La gráfica de barras que representa la información es la siguiente:



Con esto, nosotros hemos inferido que la población mexicana realmente está consciente de que existe un fenómeno social llamado S. I. D. A., pero no podemos concretar diciendo que realmente la gente conoce de él, por eso nosotros recurrimos al análisis de las respuestas más comunes que nos dieron referente a lo que conoce de los mecanismos de defensa y contagio que conoce, el análisis es el siguiente:

En el capítulo II se mencionó cuales eran las formas o vías de contagio y las medidas de prevención más comunes y/o eficaces, éstas son:

Formas o vías de contagio

- Relaciones sexuales
- Transfusión sanguínea
- Perinatal

Medidas de Prevención

- Condón
- Abstinencia
- Fidelidad mutua

Para tratar de juzgar si realmente una persona está verdaderamente bien informada sobre las formas de prevención y contagio realizamos la siguiente valoración con respecto a las respuestas más comunes que nos arrojaron.

Pregunta no. 4

¿Usted ha oído hablar del VIH y las formas de prevenir y contraer su infección?
Si la respuesta fue sí entonces la valoramos por 2 puntos.

Pregunta no. 5

Si su respuesta anterior fue afirmativa, mencione las formas de prevenir y contraer la infección del virus del VIH

Si mencionó como formas de prevenir:

Uso del condón, 2 puntos
Abstinencia 1 punto
Fidelidad mutua 1 punto

Si mencionó como formas de contraer la infección:

Relaciones sexuales (se sobre entiende con una persona infectada del virus) 2 puntos
Transfusiones sanguíneas 1 punto
Vía perinatal 1 punto

En total son 10 puntos los que consideramos para juzgar de una manera muy particular y como parte de un equipo que conformamos, si realmente una persona tiene los conocimientos básicos del tema. También cabe mencionar que en la muestra algunas personas mencionaron algunas respuestas no muy acertadas tomando en consideración los elementos básicos sobre el tema expuesto en el capítulo II; tal fue el caso de algunas personas que mencionaban como medidas de prevención del VIH a los métodos anticonceptivos en general, esta cuestión fue base para no considerar otras respuestas

Los resultados sobre esta variable son los siguientes:

Según nuestro criterio ya antes mencionado consideramos que 596 de 907 personas conoce realmente los conocimientos básicos del VIH, lo que corresponde a un porcentaje

del 65.71%. Y por lo tanto, el 34.29% de la población es el que no tiene esos conocimientos básicos.

Esta variable es considerada importante y se utilizará como propuesta para controlar el porcentaje de personas infectadas de S. I. D. A. en México.

3. LA INFORMACIÓN COMO PROPUESTA PARA CONTROLAR EL PORCENTAJE DE S. I. D. A. EN MÉXICO

La tendencia de los pronósticos es a la erradicación de la infección; consideramos que ésta es ocasionada por la información que hay sobre el tema; la tendencia creciente que hubo en los primeros años después de presentarse el virus, en forma muy probable fue ocasionada por la desinformación que había en ese entonces.

Anteriormente inferimos que el 34.29% de la población es la que no tiene la información básica, este porcentaje es el que se utilizará para monitorear el control para proponer a LA INFORMACIÓN COMO PROPUESTA PARA EL CONTROL.

PRONÓSTICO	INTERVALO DE CONFIANZA DEL PRONÓSTICO	PRONÓSTICO ALTERADO (ÚNICAMENTE EL 34.29% NO TIENE LA INFORMACIÓN SUFICIENTE DEL S. I. D. A.)	INTERVALO DE PRONÓSTICO ALTERADO
282	62 – 503	97	21 – 172
282	55 – 509	97	19 – 175
281	47 – 515	96	16 – 177
280	40 – 520	96	14 – 178
280	33 – 526	96	11 – 180
279	27 – 531	96	9 – 182
278	20 – 536	95	7 – 184
277	14 – 541	95	5 – 186
277	8 – 546	95	3 – 187
276	2 – 550	95	1 – 189
275	0 – 555	94	0 – 190

4. EL USO DEL CÓNDON COMO PROPUESTA PARA CONTROLAR EL PORCENTAJE DE S. I. D. A. EN MÉXICO

En el capítulo II se mencionó que una de las formas más eficaces para prevenir la infección es el uso correcto del condón, sabemos que es eficaz, pero no en un 100%; esto debe ser tomado en consideración ya que no garantiza que al utilizarlo la persona no pueda transmitir el virus a otra persona o que sea ella la que quede infectada. En nuestro estudio, tenemos que:

MED_UTI	NO. DE PERSONAS	PORCENTAJE
B	364	40.13
C	184	20.29
A	169	18.63
N	37	4.08
ABC	19	2.09
AO	1	0.11
BC	81	8.93
CO	1	0.11
AB	21	2.32
AC	10	1.10
D	6	0.66
BD	11	1.21
BCD	2	0.22
CD	1	0.11
SIN CONTAR	0	0
TOTAL	907	100

El 40.13% usa condón, el 8.93% usa condón y dice mantener una fidelidad mutua con su pareja, el 2.3% mencionó que usa condón y se abstiene de tener relaciones sexuales, el 1.2% mencionó que usa el condón y utiliza otros métodos para prevenirse de la enfermedad y el 0.22% mencionó que usa el condón, mantiene una fidelidad mutua con su pareja y utiliza otros métodos de prevención con su pareja.

El 2.3% mencionado en el párrafo anterior no lo tomaremos en cuenta, ya que de forma muy probable no es una información fidedigna, de acuerdo a la respuesta que se proporcionó.

El 0.22% tampoco se considerará, ya que nunca mencionaron cuáles otros métodos de prevención de la infección utilizan.

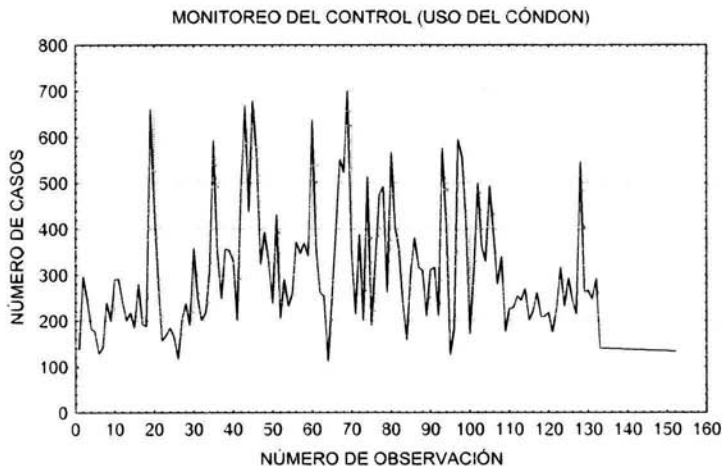
Por lo tanto, podemos concluir o inferir que el 50.24% de los mexicanos en edad sexualmente activa usa el condón como parte de un método en contra de la infección del virus del VIH. Y el 49.74% de esa población es la que no utiliza esa forma de prevención.

Si realmente esto tuviera una efectividad del 100% y se le diera una mayor promoción al uso del condón los resultados se verían reflejados en el porcentaje de personas infectadas de S. I. D. A en México, vea la tabla comparativa de las 2 situaciones:

PRONÓSTICO	INTERVALO DE CONFIANZA DEL PRONÓSTICO	PRONÓSTICO ALTERADO (EL 49.74% NO USA EL CONDÓN)	INTERVALO DE PRONÓSTICO ALTERADO
282	62 – 503	140	30 – 250
282	55 – 509	140	27 – 253
281	47 – 515	140	23 – 256
280	40 – 520	139	20 – 259
280	33 – 526	139	16 – 262
279	27 – 531	139	13 – 264
278	20 – 536	138	10 – 267
277	14 – 541	138	7 – 269
277	8 – 546	138	4 – 272
276	2 – 550	137	1 – 274
275	0 – 555	137	0 – 276
275	0 – 559	137	0 – 278
274	0 – 564	136	0 – 281
273	0 – 568	136	0 – 283
273	0 – 572	136	0 – 285

272	0 – 576	135	0 – 287
271	0 – 580	134	0 – 289
271	0 – 584	134	0 – 291
270	0 – 588	134	0 – 293
269	0 – 591	134	0 – 294

El monitoreo registró la siguiente gráfica:



Observe el decaimiento radical de la gráfica al representar los pronósticos obtenidos mediante el modelo ARMA (1,1) y considerando únicamente el 49.76% de los casos de cada pronóstico, ya que el porcentaje de población que dice utilizar el condón para protegerse del virus es del 50.24%.

CONCLUSIONES

Es increíble la forma en que este trabajo nos ha ayudado a comprender el problema del S. I. D. A. El supuesto que anteriormente teníamos sobre la tendencia creciente ha sido eliminado por completo de una forma rápida con la ayuda de las matemáticas y la computadora, esto al observar la gráfica de la serie de tiempo que generan los casos de S. I. D. A en México. Supuesto no exclusivo de nosotros sino de un gran porcentaje de personas, ya que mientras elaborábamos el trabajo cuestionamos a un buen número de personas sobre el caso percatándonos de ello. En el presente trabajo se aplicaron diversos temas vistos en la carrera, siendo de los principales la Estadística, ya que realizamos una investigación de campo que requería de conocimientos sobre muestreo, población, estimación, etc. Sin lugar a duda, podemos concluir que esta investigación nos ha ayudado a madurar de alguna manera como matemáticos aplicando las técnicas mediante el uso de la computadora y tomando decisiones al respecto. Debemos mencionarlo así, ya que es este el quehacer de los egresados de la carrera.

El S. I. D. A. es un problema mundial y por lo tanto un gran problema, debemos de respetarlo y no ignorarlo, porque de ser así ocasionaría grandes pérdidas humanas, se nos saldría de control y traería consigo la destrucción de sociedades enteras. El problema debe de ser considerado por: gobernantes, investigadores, doctores, matemáticos, etc.; pero hemos llegado a concluir que el problema debe de tomar conciencia aún con mayor importancia a las sociedades.

Creemos que el tema que abordamos en nuestro trabajo guarda una cierta importancia ante nuestra sociedad. Nuestras ideas de cómo lograr un cierto control nunca carecieron de un cierto fundamento o razón, haciéndonos pensar con esto que las instituciones correspondientes puedan lanzar campañas de concienciación con el objetivo de erradicar la infección del VIH en México. La manera en que se puede lograr esa concienciación es haciéndole ver a la sociedad de que ella forma parte de esa solución.

Una investigación de campo requiere de bastantes recursos para ser realizado, sin embargo, la carencia y/o en algunos casos la insuficiencia de los mismos no logró que nosotros nos rindiéramos, ya que los resultados obtenidos caracterizan de alguna manera a nuestra sociedad. Sociedad conciente del problema, pero sociedad ignorante aún ante el mismo.

La información que existe en relación al S. I. D. A. es bastante, pero no se tiene acceso con la suficiente en el aspecto de estadísticas, el gobierno o las instituciones gubernamentales, deben de ser un poco más abiertas para brindar ese tipo de información, ya que eso traerá consigo menos problemas a personas que requieran ese tipo de información. Lo anterior debido a que nosotros no pudimos obtener la información en CONASIDA, institución que se encarga de apoyar y brindar información a la sociedad en relación al S. I. D. A.

La gente en su mayoría tuvo la amabilidad y la disposición para contestar el cuestionario, con esto podemos concluir que nuestra sociedad ya esta abierta, ya podemos hablar de temas que en la época del pasado siglo pudieran ser considerados como tabúes. Esto nos beneficia en gran medida ya que entenderemos como sociedad de que se trata de un problema de esa índole y no exclusivo de algún sector de ella misma.

La serie, en un principio, no causa conflicto. Con esto decimos que no hubo la necesidad de realizar diferencias o transformaciones. Sin embargo, el problema es grande y por esa razón insistimos en que alguien lleve a cabo un análisis mucho más profundo de lo que nosotros hicimos, entendiéndose con esta afirmación que se destinen recursos, como suficiente dinero, pero también tiempo.

Este trabajo proporciona una idea de los cambios significativos que habría en la gráfica, la cual representa al número de casos notificados de S. I. D. A. en México; a partir de esta idea, la iniciativa privada o las estancias gubernamentales correspondientes pueden realizar investigaciones para encontrar la medida preventiva que fuese más conveniente incentivar para reducir el porcentaje de población infectada de S. I. D. A. en México.

Mientras el mal siga acabando con la gente, pensamos que este trabajo contribuye a fortalecer el ánimo, encaminado a la erradicación de la enfermedad. Lo mejor que pudimos observar es:

1. Estacionaridad de la serie.
2. El pronóstico indica constancia. Esto es, que si no se hace algo más al respecto, las cosas seguirán igual o peor. Se espera que en poco tiempo se llegue a una cifra récord porque hay demasiados casos que no se habían notificado desde un principio, lo que hace elevar el número de infectados. Esto está fuera de nuestro estudio.
3. El mejor instrumento con el que cuenta la población sigue siendo la información.
4. La aplicación de las series de tiempo no sólo se debe buscar en la economía. Actualmente se observa que se puede aplicar en muchísimas ramas de las ciencias, sean físicas o sociales; este trabajo es un ejemplo.

GLOSARIO

ASINTOMÁTICO. Que carece de síntomas.

AUTORREGRESIVO. Expresa un pronóstico como una función de los valores previos de la serie de tiempo.

ESTACIONALIDAD. Es una característica propia de un proceso estocástico y que consiste en una pauta regular de comportamiento periódico.

ESTACIONARIEDAD. Cuando las características del proceso estocástico no cambian con el tiempo, tales como que la media y la varianza sean constantes, estamos trabajando con un proceso estacionario.

ESTADÍSTICO. Es el valor numérico de una característica de la muestra y que se utiliza para hacer inferencias acerca de un parámetro poblacional.

ESTIMACIÓN. Es el resultado de valorar un cierto parámetro mediante un estimador, así como la valoración que determina el estimador a partir de una muestra. Pero también existe la posibilidad de obtenerlo por medio de un intervalo, que se denomina, *de confianza*, en lugar de un valor único.

ESTIMADOR. Es un estadístico que se utiliza para estimar el valor de un cierto parámetro.

ESTIMADOR DE MÁXIMA VEROSIMILITUD. Estadístico que se aproxima al valor del parámetro que maximiza el valor de la probabilidad de la muestra aleatoria observada. En otras palabras, encontrar el valor del parámetro que maximiza la función de verosimilitud. Se centra en el hecho de que poblaciones diferentes generan muestras diferentes.

HETEROSEXUAL. Persona que siente atracción por el sexo opuesto y por supuesto mantiene relación erótica.

HOMOSEXUAL. En oposición al ser heterosexual, inclinación sexual o erótica hacia individuos del mismo sexo.

INFERENCIA. Proceso mediante el cual a partir de una muestra se sacan conclusiones acerca de una población.

INVERTIBILIDAD. Condición análoga a la de estacionariedad. Consiste en que la parte de promedio móvil del proceso ARMA puede invertirse en un proceso autorregresivo puro.

MEDIA MÓVIL. Método de pronóstico que se obtiene encontrando la media de un conjunto específico de valores para que después se utilice para pronosticar el siguiente periodo.

MODELO. Es la reproducción ideal y concreta de un objeto o de un fenómeno con fines de estudio y experimentación. El modelo es la simplificación de la realidad, con las características más importantes de la misma.

MODELO SOBRESPECIFICADO. Se refiere a que en el modelo examinado de pronóstico se puedan omitir parámetros redundantes o excesivos. Se observa que los coeficientes sean significativamente distintos de cero con varias técnicas.

MODELO SUBESPECIFICADO. Es el resultado de agregar un parámetro adicional al modelo tentativo de pronóstico para verificar que contiene el número apropiado de parámetros y con esto observar si se obtiene un mejor modelo. Se recomienda no agregar jamás un parámetro AR (autorregresivo) y uno MA (Media Móvil) simultáneamente.

MOMENTOS. Son los valores esperados de ciertas funciones de una variable aleatoria X y forman una colección de medidas descriptivas que pueden emplearse para caracterizar la distribución de probabilidad de X (por esta razón también es correcto emplear la frase *momentos de la distribución de probabilidad X*) y especificarla si todos los momentos de X son conocidos. Su uso es muy útil principalmente porque lo más probable es que experimentador tenga un desconocimiento total de la distribución de probabilidad

MUESTRA. Es un subconjunto *representativo* seleccionado de una población. Con esto se quiere decir que la muestra refleja las características esenciales de la población de la cual se obtuvo.

OPERADOR DE RETARDO. Se denota por lo regular con una letra mayúscula B e indica que en una serie de tiempo los valores se desfasan o "retardan" una cierta cantidad de veces. Dicho operador se define mediante la relación

$$BZ_t = Z_{t-1} \quad \text{para toda } t$$

Por aplicación sucesiva del operador B se obtiene:

$$\begin{aligned} B^2 Z_t &= B(BZ_t) = Z_{t-2} \\ B^3 Z_t &= B(B^2 Z_t) = Z_{t-3} && \text{para } k=0, 1, 2, \dots \text{ y toda } t \\ \dots & \dots \\ B^k Z_t &= B(B^{k-1} Z_t) = Z_{t-k} \end{aligned}$$

PARÁMETRO. Es un valor que caracteriza de alguna manera a la población. Por extensión, una caracterización numérica de la distribución de la población, de manera que describe, parcial o completamente, la función de densidad de probabilidad de la característica de interés.

PERINATAL. Se refiere básicamente a lo relacionado con la gestación, nacimiento y primeros meses del niño, pero en especial, a la forma de transmisión del VIH de la madre a su hijo durante el embarazo.

POBLACIÓN. Colección de toda la posible información que caracteriza a un fenómeno y que en Estadística se refiere a cualquier colección ya sea de un número finito de mediciones o una colección grande, virtualmente infinita, de datos acerca de algo de interés.

PROCESO ESTOCÁSTICO. Es un proceso o sucesión de eventos que se desarrolla en el tiempo, en el cual, el resultado en cualquier etapa contiene algún elemento que depende del azar.

RESIDUAL. Básicamente un residual es la diferencia entre un valor real y su valor de pronóstico. Esto motiva la creación de métodos para resumir los errores generados por una técnica particular de pronóstico donde la mayoría de estas mediciones implican promediar alguna función de la diferencia entre el valor real y su valor de pronóstico.

RUIDO BLANCO. En una serie de tiempo, cuyos valores sucesivos pueden ser altamente dependientes, puede considerarse generada a partir de una serie de choques aleatorios independientes. Estos se suponen que son realizaciones independientes de una variable aleatoria con media y varianzas constantes. A esta sucesión de variables aleatorias se le conoce como ruido blanco.

SERIE DE TAYLOR. Es una forma de representar funciones particulares como sumas de series de potencias y este nombre se da en honor del matemático inglés Brook Taylor (1685-1731). Además, si una función f tiene un desarrollo de serie de potencias en un valor a , ha de ser de la forma:

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n \\ &= f(a) + \frac{f'(a)}{1!} (x-a) + \frac{f''(a)}{2!} (x-a)^2 + \frac{f^{(3)}(a)}{3!} (x-a)^3 + \dots \end{aligned}$$

y es precisamente a esta ecuación la que lleva el nombre de *serie de Taylor de la función f en a* (alrededor de a , o centrada en a).

SÍNDROME. Conjunto de signos y síntomas que constituyen un estado patológico y caracterizan el cuadro clínico de una enfermedad.

VARIABLE ALEATORIA. Es el resultado de un experimento aleatorio, que varía de unas pruebas a otras. Este resultado es un número real que es asociado a cada elemento de todos los posibles resultados del experimento (que se conoce como espacio muestral), por lo que la variable aleatoria es una función.

VARIABLE DEPENDIENTE. Es aquella variable que se encuentra bajo la influencia o que depende de otra variable que se denomina explicativa o independiente.

VARIABLE INDEPENDIENTE. Representa cualquier valor que se puede usar en la función, o sea, del dominio en el que está definida ésta.

VARIABLE BOOLEANA. Es un símbolo que se utiliza para representar magnitudes lógicas. Cualquier variable puede tener un valor 1 ó 0.

BIBLIOGRAFÍA

1. **BOX**, George E. P. and **JENKINS**, Gwilym M., Time series analysis. Forecasting and control, Holden-Day, San Francisco, 1976.
2. **GONZÁLEZ VIDEGARAY**, MariCarmen, Modelos de decisión con procesos estocásticos II (Metodología de Box and Jenkins), Ediciones Acatlán, México, 1990.
3. **MERCADO H.**, Salvador, ¿Cómo hacer una tesis?, Editorial LIMUSA, México, 1994.
4. **DELGADO RUBIO**, Alfonso, SIDA. Todo lo que la sociedad necesita saber, Ediciones mensuales, Barcelona, 1994.
5. **GALVÁN DÍAZ**, Francisco, El SIDA en México: los efectos sociales, Ediciones de cultura popular, México, 1988.
6. **RIVETT**, Patrick, Construcción de modelos para análisis de decisiones, Editorial LIMUSA, México, 1983.
7. **HANKE E.**, John y **REITSH G.**, Arthur, Pronósticos en los negocios, Editorial Prentice-Hall, México, 1996.
8. **GUERRERO**, Víctor M., Análisis estadístico de series de tiempo económicas, Universidad Autónoma Metropolitana, México, 1991.
9. **PINDYCK**, Robert S. y **RUBINFELD**, Daniel L., Econometría: modelos y pronósticos, Editorial McGraw-Hill, México, 2001.
10. **CUÉLLAR AGUAYO**, Ada Ruth, La simulación como herramienta de pronóstico en el área epidemiológica, Tesis para obtener el título de Licenciado en Matemáticas Aplicadas y Computación, UNAM Campus Acatlán, México, 1994.
11. **OTERO**, José María, Econometría. Series Temporales y predicción, Editorial AC (Alfa Centauro), Madrid, 1993.
12. **NAVA**, F. Alejandro, Procesamiento de Series de Tiempo, Fondo de Cultura Económica, México, 1994.
13. **NELSON**, C. R. , Applied Time Series Analysis for managerial forecasting, Ed. Holden-Day, San Francisco, 1977.

-
14. **BOWERMAN, B. L. y O'CONNELL, R. T.** , Forecasting and Times-Series, North Scituate, Mass., 1986.
 15. **WEI, William**, Time Series Analysis. Univariate and Multivariate Methods, Addison-Wesley, Massachusetts, 1990.
 16. **CHATFIELD, C.**, The Analysis of Time Series: An introduction, Chapman and Hall, Londres, 1980.
 17. **MAKRIDAKIS, S.; WHEELWRIGHT, S. C., y McGEE, V. E.**, Forecasting: Methods and applications, John Wiley & Sons, Nueva York, 1983.
 18. **CANAVOS, George C.**, Probabilidad y Estadística: Aplicaciones y Métodos, McGraw-Hill/Interamericana, México, 1988.
 19. **URIEL, E.**, Análisis de series temporales. Modelos ARIMA, Paraninfo, Madrid, 1985.
 20. **McCLEARY, R., y HAY, R. A.**, Applied Time Series Analysis for the Social Science, Sage Publications, Beverly Hills (California), 1980.
 21. **MENDENHALL, William**, et. al., Estadística Matemática con Aplicaciones, Grupo Editorial Iberoamérica, México, 1994.
 22. **ESPINOSA de los Monteros, Julián** (coordinador general), Diccionario de Matemáticas, Cultural, S.A., Madrid, 2000.