FES

ACATLÁN

"Diseño de modelos matemáticos para obtener pronósticos confiables de ventas de una empresa del sector juguetero"

Asesor: Mtra. María del Carmen González Videgaray

Presenta: Sergio Esqueda Miyamoto

Acatlán Estado de México, Diciembre 2007

Agradecimientos

El presente trabajo esta dedicado con mucho amor y cariño a mi esposa Alejandra y a mis hijos Sergio y Alejandro, como dice el Sr. Alberto Cortés “querer es poder”

A mi madre, mi primera maestra, quien me enseño a ser competitivo en cualquier área, para aprender y destacar.

A mi mamy Angelita, donde quiera que tú me estés observando, es para y por ti

A la FES Acatlán, donde siempre he tenido el apoyo de los miembros de la Jefatura de la carrera de Actuaría, en especial a mi Asesora la Maestra María del Carmen González Videgaray, con infinita gratitud por todas sus enseñanzas.

Sergio Esqueda Miyamoto.

Índice de la tesis

Introducción

Diseño de Modelos Matemáticos, para obtener pronósticos confiables de Ventas de una Empresa del Sector Juguetero

Introducción

Los modelos para la predicción de ventas tienen una importancia vital para la planeación de toda empresa, en el mercado Juguetero, la efectividad de los cálculos había sido relativamente fácil en los años noventa, ya que el mercado estaba en un inicio empezando y era controlado por sólo dos compañías importadoras americanas, Hasbro y Mattel, las cuales distribuían casi el 85% del juguete del mercado, y no existía casi importación directa de parte de los clientes, lo que hacía muy predecible el pronóstico de las ventas estando en una o en otra compañía.

Sin embargo con la apertura de las fronteras, la madurez que ya esta teniendo el mercado del juguete al tener más ventas los video juegos y sumándole las necesidades de los clientes de tener más volúmenes de ganancias, los motivó a recurrir cada vez más a la Importación directa, y a comprar a nuevas empresas importadoras y productoras que se incorporaron al mercado nacional, lo cual implicó una nueva repartición del mismo y provocó que cada vez sea más difícil alcanzar los volúmenes de ventas como se hacía antes, donde lo que se traía era lo que se vendía.

Ahora se tiene la necesidad de saber responder: ¿hasta dónde una compañía puede llegar a incrementar sus ventas contra los años pasados?, ¿cuánto se colocará con cada cliente?, ¿cómo puedo ser más eficiente para mantener controlados mis gastos de promoción vs distribución?, ¿cuánto más puede crecer el mercado, cuando la tendencia mundial de los juguetes tradicionales tienen a disminuir sobre todo por el crecimiento de los video juegos?

Estas preguntas son hechas por todas las compañías jugueteras y partiendo de la constante actual de que casi todo el juguete que llega a México es de importación, y la mayoría de las compañías deben solicitar a sus plantas de Oriente (China principalmente) el producto que van a vender con 6 meses de anticipación nos lleva a la realidad de que la planeación es vital para continuar vivo en este mercado.

A esto sumado el hecho de que el juguete es en un gran porcentaje un artículo de moda por eventos determinados y marcados en el tiempo como son: estreno de películas, series de televisión, día del Niño, navidad, día de Reyes etc., por lo que un alto porcentaje de productos son conocidos como productos de entrada por salida, y cabría hacerse la pregunta: ¿cuánto pueden venderse?, ¿porqué generalmente cuando son un hit, se tienen problemas de des-abasto?, y como la mercancía se pide con 6 meses antes de su lanzamiento al mercado, sin saber como va a funcionar ya no es posible reaccionar para tener más producto y se presenta un fenómeno de “no venta”.

O pueden ser llegar a ser un fracaso, y causan un sobre inventario en los clientes lo que provoca la necesidad de tomar acciones como son: aceptar devoluciones, autorizar aportaciones para rebajas, El tener producto que no se vende y ocupe un espacio físico en la tienda ó almacén, del cliente causa una mala imagen de la compañía importadora, problema que, con el pasar del tiempo menos se vende (el juguete que ya no esta de moda se dice que apesta en el anaquel), y el empaque se maltrata y no queda más que recogerlo en devolución y tenerlo en el almacén de la compañía, y rematarlo por abajo de su costo a un Saldero. Este fenómeno es llamado **"sobreventa"**.

Por esto que se pone mucho cuidado en realizar una adecuada planeación de las ventas la cual es responsabilidad del área comercial (ventas y mercadotecnia), quienes se deben de coordinar entre sí para obtener cada vez mejores estimaciones de lo que se puede colocar y digerir por el mercado en la fecha adecuada y con esto garantizar un negocio para los clientes y la compañía.

Tradicionalmente esto se hace por medio de juntas con los gerentes de marca que tienen un número global de cuanto piensan por producto que el mercado puede consumir y ventas dice cuanto fue el histórico desplazado por ese producto o uno similar si es nuevo, en el pasado, y da su recomendación de ¿Cómo?, ¿Cuánto? ¿Dónde?

El presente trabajo pretende desarrollar **modelos matemáticos** que pronostiquen las ventas de una empresa juguetera utilizando principalmente herramientas estadísticas para ayudar a planear las estrategias funcionales que lleven a un mejor resultado financiero, comercial y operativo en los años venideros.

Una estimación confiable de las ventas, que no sólo sea por el estómago de los gerentes de ventas/mercadotecnia ayudará a dar mejores resultados.

El trabajo esta dividido en 5 capítulos:

- Antecedentes
- Elementos requeridos para la proyección
- Justificación del Modelo
- Modelo Matemático
- Análisis de resultados y estrategias recomendadas en base a los resultados del modelo matemático.

Resumen por capítulo:

Capítulo 1 Antecedentes

Este capítulo hablará sobre el mercado de los juguetes, de los principales retos y logros alcanzados por las distribuidoras en México, y lo que se espera en un futuro cercano. Además se detallará sobre la necesidad de contar con una adecuada planeación de las futuras ventas para establecer cuales son los elementos esenciales para la competitividad de una compañía como son optimizar las campañas de publicidad, tener las llegadas de producto a tiempo, costos de almacenaje cero, etc.

Se definirán algunos detalles sobre los elementos que influyen en la facturación por mes y por año para que estos elementos sirvan de información adicional en los filtros adaptativos y en la elección del mejor modelo.

Capítulo 2 Elementos requeridos para el cálculo

Esta parte del trabajo se presentará toda la información disponible para la estimación del modelo, como son los históricos de ventas por pesos y piezas, de una compañía juguetera, además se detallarán todos los elementos económicos y financieros que se deben de tomar en cuenta para realizar un pronóstico y hacerlo dinámico y apegado a la realidad de las condiciones del mercado que son muy cambiantes.

Además se dará un breve panorama de las condiciones económicas que se requieren para continuar con un crecimiento de la compañía en el mercado juguetero.

Se darán algunos ejemplos del uso que se le da a la inferencia estadística en algunas compañías para la toma de decisiones, y algunas propuestas para las jugueteras que históricamente sólo usan la parte que corresponde a la estadística descriptiva.

Capítulo 3 Justificación del modelo

Primeramente se dará una introducción a la metodología de Box-Jenkins para pruebas de tendencia, ciclos y variantes de estación. Se extenderá en el desarrollo de las fórmulas y criterios que se requieren para lograr una mejor decisión en el mejor modelo que ajuste a la serie.

Se calcularán parámetros descriptivos que den una primera estimación de la futura facturación anual por mes y cliente de la compañía para tres años.

Después se determinará los patrones que tienen los históricos de ventas de la compañía, como son tendencia, ciclos, estacionalidad y se hará un estimado para 2007-2009, considerando medias móviles y regresión lineal de un histórico de ventas contra inversión publicitaria.

Posteriormente se analizará la disminución del ruido blanco de los datos utilizando las técnicas de filtros adaptativos que consiste en asignación de los pesos apropiados a las observaciones y el abatimiento del ruido asociado, y se generan las gráficas de las autocorrelaciones muestrales (ACF) y autocorrelaciones muestrales parciales (PACF).

Capítulo 4 Modelo matemático

En este capítulo se utilizarán los elementos teóricos que se expusieron en el capítulo anterior, y de la serie del histórico de ventas por mes se harán pruebas de diferentes modelos que mejor se ajusten a los datos y que den los mejores resultados futuros, con la ayuda del programa SPSS.

Se darán pronósticos de ventas de algunos clientes importantes utilizando algunos datos financieros cuya correlación es alta con las ventas alcanzadas

Se harán proyecciones de algunos productos o líneas continuas, cuya proyección ayudará a predecir el desempeño de productos nuevos.

Además se calcularán los márgenes de variación que se obtienen, la correlación que muestran y determinación de pruebas de consistencia con los datos reales.

Capítulo 5 Análisis de resultados y estrategias

Este capítulo se dará una interpretación de los resultados obtenidos, además de realizarse una comparación, hasta donde sea posible con datos reales hasta la fecha de la terminación de la tesis.

Se proyectará considerando la serie de tiempo 1993-2006, el año 2007, y se comparará con los datos reales disponibles a la fecha.

Se realizará un resumen de las proyecciones de ventas 200-2009, de acuerdo a las hipótesis económicas y financieras, correlacionadas con las posibles estrategias comerciales de la empresa.

Se hará un reporte resumen de las expectativas de ventas de los principales clientes, de acuerdo a factores económicos y comerciales, resaltando fortalezas y debilidades de cada un de los analizados.

Se darán algunas recomendaciones de los pasos que a criterio de un servidor, ayudarán para alcanzar los objetivos de venta estimados con el modelo.

Conclusiones

Se darán algunos elementos adicionales mezclando por una parte lo que indica la parte matemática, y por otra parte la experiencia de 10 años en el mercado juguetero del expositor.

Anexo 1
Se darán un listado con las principales fórmulas utilizadas a lo largo del texto.

Anexo 2
Se darán diversos ejemplos de pronósticos de ventas con diversos métodos utilizados en la práctica.

Capítulo 1
Antecedentes

Capítulo 1

Antecedentes

1.1 Situación actual del mercado juguetero.

1.1.1 Compañías actuales del mercado juguetero.

En los años noventa, el mercado juguetero mexicano estaba dominado por dos grandes compañías norteamericanas, Hasbro y Mattel, cada una de las cuales tenían un crecimiento tan rápido que existían años donde crecían casi al doble en ventas. Las dos compañías tenían segmentos de mercado diferentes en la mayoría de los casos y generalmente había pocos segmentos donde competían de frente. Esto cambió a partir de 1999, cuando al mercado se incorporaron al mercado diversas empresas, así como la Importación directa de los clientes, modificó el panorama y la repartición del mercado se hizo más particionada.

El mercado juguetero se clasifica principalmente en los siguientes segmentos:

- Niños: figuras de acción, vehículos, modelismo, accesorios de aventura y acción, artículos deportivos, al aire libre

- Niñas: fashion dolls (muñecas de moda, tipo Barbie, Bratz), muñecas de mecanismo, muñecos con accesorios, muñecas de trapo, peluches (interactivos o tradicionales), juegos de cocina (muebles o vajillas, accesorios de cocina), estuches de maquillaje

- Mixtos: carros montables y eléctricos, bicicletas, inflables, balones ó pelotas, juguetes tradicionales (canicas, trompos y yoyos), juegos de mesa y rompecabezas, juegos electrónicos, resbaladillas o columpios de metal o plástico (role play), preescolares, didácticos y científicos.

Hasbro tradicionalmente domina en el segmento de juegos de mesa (marca Milton Bradley), Parker Brothers) , y algunos segmentos de niños como son las marcas de figuras de Action Man, Tranformers, Star Wars y recientemente incorporó las figuras de Marvel, en especial Spiderman. En niñas tiene el segmento de muñecas tipo Nenuco y Cabbage Pach con accesorios.

Mattel es líder en el mercado, domina el segmento de fashion dolls, con Barbie, preescolar con Fisher Price, y de Niños con figuras como Max Steel, Power Rangers, Figuras de D.C Comics como Batman, Superman entre otras, y en el de vehículos con los sumamente conocidos Hot Weels, donde el carrito es el producto más vendido en unidades en cualquier juguetería a lo largo del año, sumándole a esto que sus autopistas son las más publicitadas y buscadas por los niños. Además cuentan desde 2003 con la marca Match Box

En didácticos están las compañías de Lego y Mega Blocks, donde en base a licencias, y al interés de los padres en que sus hijos jueguen con cosas que les desarrollen habilidades, Play Doh es una marca de Hasbro en este segmento que esta teniendo cada vez más importancia, sustituyendo a la plastilina tradicional.

Los Juegos Electrónicos son dominados por Sony y Nintendo, que con sus consolas tienen un mercado que cada vez es más amplio ya que los video-juegos no sólo son jugados por niños y adolescentes, ya que cada vez son más los adultos que se interesan por tener lo último de estos productos.

Debemos además de considerar a empresas como MGA con Bratz y a Learning Leap Frog de Leap Frog, en los segmentos de fashion dolls y preescolar como serios adversarios, ya que en E.U. en el año pasado en México, tuvieron un excelente desempeño dándoles serios dolores de cabeza a los líderes de las categorías que son Barbie y Fisher Price

Además de existir marcas propias de juguete (importación directa del cliente) que cada vez son más conocidas como Kid Conection, Britanny en Wal Mart, Kiddi Toys en Juguetrón., Lilly Ledy de Juguetibici, etc., quienes traen productos similares a los que son anunciados en televisión por parte de Mattel, Hasbro, Leap Frog, MGA etc.

Además existen los llamados Brokers, que son Importadores que representan a compañías sobre todo Orientales que ofrecen una infinita variedad de productos hechos en China cuyo costo les permite ganar un gran margen de utilidad, existen importadores de la talla de Lion Fung, International Sources, Smoby International, Innovaciones de América, Famosa, por mencionar algunos.

Finalmente habría que mencionar a los productos nacionales los cuales con tanta competencia se han visto obligados a mejorar su calidad y algunos lo que están haciendo es producir en China, importar como piezas sueltas y poner algo hecho en México y armar paquetes de juguetes.

Consideramos que los productos de montables y bicicletas, y los Juegos científicos y de armar son los únicos que sobrevivirán a esta desigual competencia. Juguetes Impala, Inflables Kay, Montables Pricel, Juguetes Mi Alegría, Bicicletas Maggistroni, Muñecas Elizabeth y Juanita Pérez, por mencionar algunos de los más conocidos.

Como mencionamos es cada vez más competitivo y maduro el mercado juguetero, y al considerar muchas variables financieras, económicas, más competencia, hacen resaltar la importancia de que el factor estocástico este presente cada vez más en el desarrollo del negocio.

Adicional a esto el mercado se ha visto disminuido sus ventas por factores económicos y sociales (las niñas cada vez juegan menos años con juguetes, y los niños tienen a su alcance productos de tecnología que cuestan los que un juguete, ejemplo simple: el celular, Ipod, Game boys, MP3, Laptops, etc.

1.1.2 Resumen General de los Principales Clientes

Como se verá en los datos, el 80% de las Ventas esta concentrada en los Autoservicios, quienes tienen como principal cualidad el poder crecer sus espacios de venta en la temporada, con lo que concentran mayor cantidad de mercancía.

El principal de todos ellos con un estimado del 47% del mercado Juguetero, y con 500 tiendas, y con planes de seguir creciendo, Wal Mart de México sigue a la cabeza con su política de precios bajos todos los días, a la cual le sumo el año pasado la de además rebajar cada trimestre 8,000 artículos, y una agresivo plan de aperturas que en el 2006 fueron cerca de 100 en todos los formatos que manejan (Tiendas Wal Mart Bodega Aurrera, SAM´S, Suburbia y Vip´s), es decir van todo, y siguen siendo el mejor cliente en cuestión de pagos oportunos, e incluso anticipados, un centro de distribución de primer mundo (la mercancía en algunas tiendas llega incuso en horas después de haber sido entregada), una planeación muy anticipada, gran volumen. Todo esto tiene un precio, "El precio más bajo que el proveedor pueda ofrecer", o como ocurre frecuentemente "el Precio que Wal Mart pida"

El segundo retail, para juguetería es Comercial Mexicana, (con un 14% de participación en el mercado), con 180 tiendas distribuidas a lo largo de todo el territorio nacional, excepto zona Norte, han luchado fuertemente con Wal Mart, con un concepto diferenciador que es en la atención de su personal, la variedad de marcas que ofrece que es más extensa que Wal Mart, la facilidad del apartado o más meses sin intereses, son elementos que le ayudan a mantenerse con una clientela que le es muy fiel, sobre todo por la oferta de frutas y verduras donde son hasta ahora líderes del mercado, y con una política de estar competitivos, "no tienen el precio más barato", pero si están muy cerca de este, y con los elementos diferenciadores que mencionamos arriba cautiva a un segmento importante del mercado (se estima en un 18%). Este año tiene pensado abrir unas 25 tiendas (Megas, Bodegas, Al Super, California), y llevar este ritmo de aperturas durante 5 años hasta casi duplicar las unidades que actualmente tienen.

Soriana, aunque es el segundo retail por volumen de ventas en general (15% del mercado) en juguetería es el tercero, pero cada vez más cerca de Comercial Mexicana, con una presencia de tiendas en 31 estados de la república (excepto D.F.), y con una lealtad de la población de la zona norte del país y una fuerte presencia en el centro del país. Actualmente Tiendas Soriana buscan atacar un segmento de población de medio a bajo poder adquisitivo con un formato de tipo Bodega, los cuales se ven cada vez más en las ciudades de más alta población de la zona metropolitana (el año pasado abrieron 40 tiendas nuevas en diverso formatos), y piensan continuar en ese ritmo. Además con programa de crédito y puntos con beneficios tipo compra-con compra, y tomando como gancho algunos artículos con precios más bajos que los Wal Mart ó Comercial Mexicana más cercanos, buscan estar con mejor posición en el mercado que lo percibe como una tienda que tiene mucho juguete, pero no de marca.

Gigante es el cuarto cliente en importancia (6.7% de participación en el mercado), esta cadena con más de 150 tiendas, pero con muchos problemas financieros que les ha llevado a la necesidad de cerrar sucursales, lejos de lo que están haciendo sus competidores. Gigante no cumple con los plazos de pago establecidos con sus proveedores, lo que le ha causado que muchos de ellos no le vendan o le vendan menos, lo que al final lo lleva a desventajas como falta de surtido o la no venta. De acuerdo a los datos financieros que publica por Internet para sus accionistas la pérdida continua, pero cada año disminuye, esta cadena se rumora puede ser comprada de un momento a otro por Soriana. En nuestra proyección este Cliente deberá tender a la baja a menos que sea comprado por un grupo de la talla del arriba mencionado.

Las cadenas departamentales, Liverpool, El Palacio de Hierro, Sears y Sanborn´s, participan con cerca del 12% del mercado, y siguen un ritmo de crecimiento de un promedio de 13% por año, influenciado sobre todo por el dinamismo de Liverpool que ha crecido su espacio de ventas en un 50% en 4 años, pero se ha detenido el dinamismo de su crecimiento, sobre todo por que los plazos de meses sin intereses, ya se ofrecen también en los autoservicios, además que en el caso del mercado juguetero, los video juegos, que se vendían exclusivamente en estas tiendas, ahora también se ofrece en los autoservicios.

Los especializados o mayoristas se están enfocando en tener tiendas pequeñas más rentables, cerca de centros comerciales, y son cada vez menor su venta al mayoreo y, su crecimiento de venta se mantiene con un incremento de un dígito, sobre todo por Juguetrón, que con nuevos negocios, como catálogos navideños y exposiciones navideñas, y algunas aperturas buscan seguir como los más experimentados del ramo, y con un surtido completo durante cualquier época del año.

1.2 Importancia de realizar pronósticos de ventas en las compañías (Caso especial Industria Juguetera)

El compromiso de cualquier empresa es mejorar cada vez dar mayor valor al servicio que se otorga a sus clientes. Pero para lograr esto se debe de realizar una mejor planeación y los pronósticos de ventas juegan un papel determinante, ya que son la base de los objetivos y decisiones de las direcciones de cualquier empresa. Por lo que un pronóstico acertado puede determinar el éxito o fracaso de todo un año de trabajo.

El objetivo que se plantea en el presente trabajo es precisamente presentar algunos modelos estadísticos que estimen las futuras ventas para una empresa juguetera, para el periodo 2006-2009, estipulando ventas por cadena, categoría, mes y de algunos artículos nuevos y continuos, herramienta que en la practica ayudaría a establecer un plan de llegadas, cuotas y objetivos por mes.

La empresa que presentamos sus datos, por ser una de las principales del ramo nos da una muestra significativa de cómo se comportó el mercado y que se espera sucederá en el futuro.

Por esto se dará una breve explicación de cómo se alcanzaron los resultados de años anteriores, y como se planeó el siguiente para identificar las variables que están en juego cuando se realizaba un pronóstico de manera tradicional.

Los modelos expuestos contemplarán diversas hipótesis, como son la inflación, cotización del dólar, estrenos de series y películas, promoción y publicidad, fecha de arribo de mercancía al país, etc., factores que darán al modelo una aproximación más apegada a la realidad.

Los modelos deberán contemplar todas estas posibles variantes ya que el mercado es uno de los más cambiantes día a día, y siempre genera una expectativa hacia todos aquellos compradores y vendedores involucrados en el día a día.

Hasta ahora no se sabe si el realizar pronósticos de ventas son una ciencia o un arte; ahora con la disponibilidad de una gran cantidad de computadoras que realizan numerosos cálculos en velocidades impresionantes, han desarrollado nuevas técnicas diseñadas por todo tipo de profesionistas las cuales en su mayoría requieren de una gran cantidad de términos y teorías complejas y difíciles de aplicar a problemas reales.

Por esto que además de aplicar técnicas clásicas como son las series de tiempo, regresión, promedios de históricos, etc. se considera el juicio ejecutivo de gerentes de ventas y hasta compradores.

1.3 Políticas básicas ocupadas por las empresas comercializadoras de juguetes para realizar estimaciones.

Para realizar un pronóstico de ventas se pueden utilizar métodos como: [1]

- Juicio o conocimiento de una o varias personas.
- Obtener estimados de la fuerza de ventas.
- Determinar expectativas de los clientes por encuestas o investigación de mercado.
- Realizar proyectos en conjunto con los Clientes sobre las expectativas de crecimiento, en base a Apertura de Nuevas Tiendas.

[1]**Práctica común en Hasbro en el área de Ventas en conjunto con Dirección**

Existen otros métodos más técnicos como: [2]

- Usar análisis de series de tiempo, usando cifras de ventas pasadas y promediándolas para obtener una línea de dirección que pueda extenderse al futuro.
- Utilizar técnicas de econometría tratando de encontrar una relación predictiva entre sus ventas y una serie estadística externa como el P.I.B. (producto interno bruto), P.N.B. (producto nacional bruto), o el índice de inflación, etc.

Muchas compañías usan otras técnicas, sin embargo para que éstas sean exitosas se han definido ocho: [3]

1.- Apéguese a formatos familiares

Esto consiste básicamente en ocupar las formas que la mayoría conoce, esto con motivo de ayudar a la comprensión de la gente que ocupe los datos.

2.- Hacer pronósticos fundamentales en un esfuerzo de equipo

Realizar pronósticos consultando a personas importantes como son Directores, Ejecutivos o Gerentes de Cuenta Clave.

3.- Enfatice la precisión

Evitar problemas como exceso de existencias o faltantes son elementos importantes para obtener precisión en los cálculos.

4.- Conozca su pasado para conocer su futuro

Esta marca la importancia de conocer las ventas pasadas y que factores estacionales se presentan, así como cambios de estilo

5.- Anticipe cambios

La empresa del juguete esta siempre sujeta a cambios como son precio, publicidad, tendencia del mercado, etc.

6.- Experimentar

Utilizar elementos nuevos como indicadores económicos específicos, revisión del pasado, suprimir elementos que propiciaron errores, entre otros.

---

[2] **John E. Hanke/Dean W. Wicher Pronósticos en los negocios**

[3]**White R. Harry Sales Executives Club of New York**

7.- Mantenerse Actualizado

Los buenos sistemas de pronóstico siempre necesitan un seguimiento constante de ser posible realizar revisiones periódicas.

8.- Conozca a su competencia

Cada es más notoria la percepción de precio por parte del Consumidor, y es necesario dar seguimiento a la tendencia del mercado, por que el Juguete es de Moda y actualmente muy Sustituible, ya que esta dirigido a un Consumidor muy volátil, el niño(a), y es comprado en un 70% por la mamá (por Decisión de compra, ya que las mujeres compran generalmente los juguetes de su hijos, sobre todo en edades menores a los 3 años).

1.4 Los elementos de cambio de decisiones

Cuando se estudian las cifras de ventas de los últimos 10 años, de seguro encontrará un comportamiento de altas y bajas que surgen de procesos de cambio. Sin este elemento las gráficas serían una línea horizontal y podría suponerse que el volumen de ventas sería el mismo para periodos futuros así como fue en el pasado.

Existen muchas cosas que producen el cambio que hacen que las cifras varíen hacia arriba o hacia abajo. Básicamente hay 5 tipos de cambio que él pronosticador trata de identificar y estimar su efecto en futuras ventas:

- **Tendencia:**

  Es el movimiento a largo plazo en una serie de tiempo. Son las recesiones y depresiones (Ciclos), la empresa ha mostrado una tendencia creciente en toda su historia. Sin embargo hay factores como comerciales, películas y caricaturas, pueden ser la proporción de crecimiento o declinación.

- **Los Ciclos:**

  Son fluctuaciones parecidas a olas que flotan sobre y bajo la línea de tendencia a largo plazo. Los ciclos de ventas se incorporaron usualmente a un ciclo un ciclo general. Casi todos los negocios experimentan estos ciclos, son recurrentes, pero no periódicos, lo cual les dificulta el hacer. Se mueven en Juguetería en el rango entre uno a diez años en longitud, con un promedio de cuatro años.

- **Las fluctuaciones estacionales:**

  Que son más fáciles de predecir en pronósticos a corto plazo, porque son tanto recurrentes como periódicos, usualmente ocurren en los mismos meses, como es la Temporada de Navidad o Reyes en juguetería, y donde la venta a los Clientes o Distribuidores se desarrolla desde los meses de septiembre y octubre.

- **Las variaciones en el calendario:**

  Estas afectan las estimaciones mensuales. Años bisiestos.

- **Las fluctuaciones irregulares:**

  Son aquellas que pueden atribuirse a cualquiera de las 4 fuerzas anteriores, también llamadas variables aleatorias, son acontecimientos como ventas de remate, falta de series de T.V (caricaturas), que estaban programadas a salir y que se quedan guardadas, llegada tardía de producto, problemas de aduana.

  Como se observa, el cambio es el elemento básico para el pronóstico.

1.5 Ciclo de vida de los juguetes

La mayoría de los productos y en general casi cualquier cosa, pasan por un ciclo de vida, que consiste en general en un período de crecimiento, de estabilidad o prosperidad y un intervalo de tiempo de declinación. Este concepto es fundamental para la empresa ya que algunos de los juguetes dejan salir al mercado varios años, para evitar saturar su imagen al consumidor.

Para el pronosticador el ciclo de vida le proporciona una referencia de las posibles fuerzas subyacentes que afecta al producto o a la categoría de productos y poder dar los pasos necesarios para ajustarse a ellos.

Por lo tanto, al pronosticar la demanda de su producto, debería de ver su estimado de acuerdo con el ciclo de vida del producto. El primer paso deberá contemplar si se han dado cambios en la curva y velocidad de los ciclos de vida de la línea o categoría del juguete.

Cuestionarse si la declinación de ventas se deba a efectos estacionales a corto plazo, o factores cíclicos a mediano plazo en el ciclo de marca de la compañía a las acciones de las empresas competidoras o a sucesos externos que van más allá del control de la compañía.

Para evitar que algún producto sufra una depresión, normalmente se deben de realizar nuevas estrategias o cambio de ajuste de precio, un nuevo empaque o un relanzamiento con publicidad, todas estás medidas son tomadas en el transcurso del año.

Al mantener el concepto de la prosperidad, o declinación, se debe considerar el factor llamado “ciclo de vida de un juguete”, esto no solo proporcionará pronósticos más exactos sino también se interpretara mejor los datos sobre los cuales se basará el pronóstico.

En Juguetes se piensa que el ciclo de vida de un producto de Licencia, de una película de moda, es de unos 6 meses (ejemplos: Superman, Harry Potter, Spider Man), en una Licencia Clásica, después de una película o serie (Ejemplo: Winnie the Pooh, Mickey Mouse, Star Wars) puede ser de 3 á 5 años.

1.6 El Liderazgo (juicio ejecutivo)

Existen muchos métodos para estimar las ventas y aun ahora él más extensamente utilizado, es el pronóstico con juicio ejecutivo, sin importar el tamaño de la empresa. Métodos sofisticados, como el análisis de series de tiempo y regresiones econométricas cada vez más utilizadas, sin embargo en la práctica las corazonadas del mercado siguen desempeñando un papel vital en casi todos los pronósticos de ventas sin importar cuantas técnicas matemáticas se utilicen.

Hasta ahora no existe para el negocio sustituto para la intuición de un ejecutivo de ventas experto, la ventaja de una relación estrecha con los distribuidores y comerciantes.

El pronosticar reflexivamente se hace de dos métodos, mejor dicho modos. Por un individuo experimentado (por lo general una compañía pequeña), o por un grupo de individuos, algunas veces llamados **"Jurado de Opinión"** [4]. El enfoque del grupo a su vez usa dos métodos: (1) Los Ejecutivos de Cuentas Clave se turnan los estimados de forma independiente sin discutirlos y estos son promediados con un pronóstico por el Director. (2) El grupo se reúne, cada persona presenta su estimado, las dificultades se resuelven y se llega a un consenso. Una variante interesante del método de grupo resulta ser cuando un Ejecutivo asume el papel de la unidad compradora de un Cliente que toma decisiones, lo que permite evaluar a partir de los datos disponibles de lo que los Clientes esperan recibir en el siguiente período.

Esta forma de pronosticar ventas se utiliza generalmente cuando:

- El Ejecutivo o ejecutivos que pronostican tiene la experiencia y poseen un buen sentido del mercado y comprenden a la perfección las necesidades del cliente
- Cuando el presupuesto de pronosticar ventas sea de corto tiempo y dinero
- Cuando el volumen de ventas es estable y el mercado esta bien definido
- Cuando el riesgo o consecuencias de un error importante en el pronóstico es relativamente bajo.

---

[4] **White R. Harry Sales Executives of New York**

Las desventajas de este método son básicamente:

- Descansa únicamente en puntos de vista personales y pueden ser en gran parte un trabajo de adivinar. La Compañía trata de minimizar esa dificultad dotando a sus ejecutivos de antecedentes en la historia de sus ventas y supuestos que tengan que ver con el panorama general del mercado.
- No proporciona una manera de pensar y evaluar las opiniones individuales, puesto que no existe un procedimiento estándar establecido.
- En los pronósticos de grupo, algunos individuos dominantes suelen determinar las decisiones del grupo y tienden a desalentar a los miembros más tímidos, este problema se puede resolver mediante la técnica de Delphi, cuya explicación se dará en el siguiente subtema..
- Tiende a dar un peso igual a las predicciones de los pronosticadores pasados sin importar que tan malos hayan sido sus resultados.
- Ya que no existe fórmula establecida, los nuevos ejecutivos tienen dificultades para llegar a pronósticos razonables.

1.6.1 La técnica de Delphi [5]

Esta fue desarrollada por la Compañía Rand, y es un método de minimización de algunas de las desventajas del grupo de pronosticadores mencionados anteriormente, consiste en llenar cuestionarios cuyas respuestas no requieren de firmarse. Los resultados se resumen después y se distribuyen a los miembros del grupo, quienes luego son interrogados sobre si ellos mismos desean sus predicciones sobre la base de otros estimados. Este proceso puede ser repetido entre los miembros del grupo si surge alguna divergencia de opinión muy señalada. Finalmente la opinión de grupo es considerada ser equivalente al promedio de las opiniones individuales finales.

Esta técnica se puede ver mucho en la práctica y actualmente se llama inteligencia de las masas, como ha podido determinar en base a los grandes números, sucede como ejemplo que para el cálculo de un número pelotas dentro de un carro, el número más exacto se obtuvo del promedio de miles de personas, y muchas personas con intelectos superiores estuvieron lejos de ese promedio, esto también es conocido como “la inteligencia de las masas”.

Cuestiones así pasan con restaurantes, donde la mejor comida es donde están más llenos, y esto lo aplicamos como un axioma, sobre todo cuando no conocemos el lugar, y lo clásico es ir a donde hay más personas.

Estos ejemplos en los negocios, ocurre mucho en los productos de consumo, por eso las compañías frecuentemente invierten grandes cantidades de dinero en muestreos, donde al lanzar un producto, le entregan al consumidor final una muestra gratis de su producto, y con esto les sirve para medir estadísticamente si el producto será o no del agrado de la mayoría de la población.

---

[5] **Hurwood D.L., Grossman E.S & Bailey L Sales Forecasting**

Existen Clientes, como Wal Mart, que tiene un sistema de ventas tan avanzado, que pueden saber con que producto se vendió en más ocasiones con el nuestro, por ejemplo Chocolate Hersey´s con Carritos, Danonino con Preescolares, Bey Blade con Sofoul, etc.

Este permite conocer al consumidor final, que rango de edad y monto de compra tiene, además de permitir hacer ofertas cruzadas, o cual puede ser un detonador de venta incremental.

1.6.2 Ejemplo en la compañía juguetera

En Hasbro, se confía extensamente en el juicio ejecutivo para la evaluación de los datos estadísticos y la efectividad de planes promocionales tanto como grado de los niveles en las ventas pronosticadas. Sus datos de productos continuos como el Adivina Quién?, Figuras Star Wars, Sr. Cara de Papa, entre otros son conocidos por los niños, son vendidos en casi todas las cadenas de autoservicio, departamentales y especializadas a lo largo del territorio nacional, por lo que gastos frecuentes en los gastos de promoción y pronósticos de ventas permiten que sus líneas puedan enfrentar y darle lucha a la competencia.

Trimestralmente se reúnen las Direcciones de Ventas, Mercadotecnia, Operaciones y Finanzas para sostener una discusión en grupo sobre el curso del producto y las tenencias del negocio. Después de haber estudiado los reportes de ventas y desplazamientos recopilados por el departamento de administración de ventas de la Compañía y de otras fuentes externas como son agencias de publicidad, determinan los volúmenes de unidades (pronósticos de piezas a vender), que se perseguirán para cada ramo, y así mismo para el total de ventas por mes, trimestre y total año, así como el calendario de llegadas de China, ya que todo el producto es importado, y se tiene que considerar contar con el suficiente inventario para cumplir la meta establecida.

La información siguiente es compilada y usada por el grupo para su uso posterior, a favor de un pronóstico, lo más preciso posible:

- Calendario promocional al comercio
- Historial de ventas
- Situación y posición en el mercado
- Tendencia de las marcas
- Calendario de llegadas (sourcing)

Capítulo 2

Elementos requeridos para la proyección

Capítulo 2

Elementos requeridos para la proyección

2.1 Histórico de ventas de los años 1993-2006

2.1.1 Antecedentes importantes por año

Los datos de facturación por año reflejan una tendencia creciente, que son el resultado de la incorporación de nuevos productos, gran difusión en los medios masivos de comunicación y una buena distribución a lo largo y ancho del país. Los datos de facturación bruta se les descontarán las devoluciones (excedente de mercancía que es devuelta por los clientes) y aportaciones (rebajas de producto que son descontadas de las facturas pagadas por los clientes. Los datos no son los datos exactos por protección a la empresa que nos los proporciono, sin embargo conservan su tendencia, estacionalidad y ciclaje.

Durante el año de 1993, inicia la operación de la empresa en cuestión, los clientes desconocían a la compañía. En 1994 se presenta después de buenos resultados un importante incremento del 68% de las ventas y cada vez más clientes se interesan por los artículos nuevos.

Sin embargo 1995 representó un año difícil, con la crisis financiera ocurrida en diciembre de 1994, el dólar se disparo, con lo que la empresa, que importa la gran mayoría de sus artículos sé vio obligada a incrementar sus precios y la colocación de los mismos era concentrada en algunos clientes, ocasionando grandes cantidades de devoluciones.

Llegó el año de 1996, con una economía más estable, se avanzó facturando cantidades adecuadas por los históricos de ventas, y a un mayor número de clientes cuyas necesidades de producto eran más altas, ya que los consumidores contaban con mejores posibilidades económicas, y esto logró que el nivel de devoluciones se disminuyera notablemente.

En 1997, se tomo como base a la experiencia obtenida en el año previo, y se planeó una venta en base muy agresiva, adelantando la facturación de temporada a meses como agosto y septiembre, cuando tradicionalmente empezaba hasta octubre. Pero esto originó un descontrol y algunos clientes recibieron más mercancía de la debida, algunos se quedaron sin surtir algunos pedidos, y esto tuvo su costo, grandes devoluciones de producto. Sin embargo el incremento de venta fue de un 60% de incremento, lo cual en un país subdesarrollado, con una crisis económica reciente y con grandes conflictos internos, debe de destacarse.

En 1998, se presentó un nuevo modelo de negocio, donde se trabajó en crecer los meses de no-temporada (llamados en el medio resurtido), que son de enero-agosto. Este año se logro colocar con la mayoría de los clientes, y se consiguió además una planeación de venta por mes y por cliente, para analizar el potencial de cada uno.

En 1999, ha sido uno de los mejores años en colocación de producto, se alcanzó un crecimiento impresionante, si embargo ese año la compañía que estamos estudiando su comportamiento de ventas, perdió mucho dinero ya que gran parte de la mercancía enviada sobre todo en diciembre no estaba sustentada con un folio real de recibo, con lo cual los clientes no pagaron esa mercancía y los únicos grandes ganadores fueron los transportistas. Sin embargo este año fue fabuloso, ya que dos conceptos como lo fueron Star Wars, Episodio 1, Pokémon y Furby fueron todo un fenómeno además del lanzamiento de Action Man. Si embargo Mattel este año también empezó a distribuir Max Steel, con las figuras de acción como Marvel y Spider Man, las cuales a la postre harían que ellos controlarán el mercado de las figuras de acción en el siguiente año.

El 2000 fue un año de supervivencia, ya que representó un año de reconstrucción por que el fenómeno de Pokémon y Furby no se pudo compensar y muchos de los clientes obtuvieron de Mattel una serie de beneficios que hicieron crecer enormemente la brecha y el líder Mattel duplicó con muchos clientes su facturación y por lo tanto dejaron a otras empresa con un presupuesto de compra inferior al año pasado. Sin embargo el 2000 se tuvieron nuevas novedades desde inicio de año con lo que lo que se empezó a trabajar una nueva estrategia por parte de la Dirección General que dieron apoyo a las nuevas negociaciones del año siguiente.

Este año, 2001 fue el año de los grandes cambios, se rentó un nuevo centro de distribución, se cambió la forma de entrega, (primero se surtía la mercancía, después se facturaba y ese mismo día se entregaba, o en un lapso máximo de 5 días en plazas foráneas.), lo que dio certidumbre y credibilidad a la empresa con los clientes. Además este año se modificaron las condiciones de compra de muchos clientes permitiendo una mayor flexibilidad de negociación y se mostró reflejada en las cifras, donde con por ejemplo los departamentales pesaron más que los mayoristas, y los autoservicios, continuaron su crecimiento acelerado, en especial el grupo Cifra, que empezaba a convertirse en Wal Mart de México y comenzó a abrir nuevas sucursales por todo el país y de manera más acelerada nuevos puntos de venta.

En este año, 2002 se empezó a tener una fuerte dependencia del Grupo Wal Mart, ya que este año se convierten todas la tiendas Aurrera en Wal Mart, lo que hace que este cliente tenga un control del mercado cercano al 35% y con esto empiece a relegar a otros como Comercial Mexicana, y Gigante, este último empieza a disminuir su facturación por problemas financieros y además de ser un factor de riesgo desde el punto de vista financiero.

Este año se trabaja en desarrollo de productos que son propiedad de la compañía (denominados core brands) y las licencias, se vuelven un segundo plano. Se tuvo el empuje de la película de Star Wars episodio 2, la cual a diferencia de la estrenada en 1999, contó con una línea figuras a un precio accesible, y fueron importantes en la facturación de todo el año..

En este año 2003, se tuvo un crecimiento exorbitante por parte de los Clientes que tenían un mejor sistema de resurtido automático, tal es el caso de Wal Mart, Comercial Mexicana, Liverpool y Carrefour, ya que se presentó un producto que da un claro ejemplo de porque este mercado es tan cambiante Beyblade, el cual por si sólo fue más del 30% de las ventas de todo el año de la compañía, lo cual lo hizo un básico tan importante que llegó a tener ventas en algunas cadenas, en pesos cercanas a la leche Alpura la milanesa, es decir productos de la canasta básicos, y fuera de temporada fuerte (diciembre-enero). Si embargo no todo fue bueno ya que este fenómeno, sumado a novedades como Bratz, y Nenuco, hicieron que la compañía descuidara algunas otras líneas que siempre tenían buen desempeño (como es el caso de juegos de mesa), y se tuvo un retroceso en las ventas de Beyblade, a partir de noviembre, lo que orilló a tener sobrexcedentes que frenaría la colocación de producto al siguiente año.

En este año 2004, se presentó dificultad de mover el excedente de mercancía y se tuvo que realizar promociones del 2x1, y eventos que funcionaron a medias, y la colocación de novedades a principio de año ayudó a mover producto. Este año se caracterizó por no tener una línea ganadora, ni tampoco lo tuvo la competencia, y el mercado estuvo poco entusiasta para comprar y a esto sumado la gran cantidad de productos que se ofrecieron al consumidor, hicieron que los buenos resultados de 2003, no se vieran este año y se tuvo que recoger una gran cantidad de inventario sobrante. Con muchos clientes se alcanzaron niveles históricos de devoluciones. Las marcas más representativas fueron los “Fur Real Friends” (Peluches Interactivos) y las “Tortugas Ninja” y los personajes de “Toy Story” se acabaron en todas partes.

En 2005, la empresa empieza poco a poco a cambiar su estrategia y ha replanteado su lista de precios, considerando novedades para todos los clientes desde enero, bajo precios en muchos artículos continuos, con muchos años en el mercado como lo son los juegos de mesa, sin embargo no fue fácil el inicio del primer trimestre ya que muchos clientes quedaron con mucha mercancía remanente, la gran ayuda para colocar en el segundo trimestre y en verano fue el estreno de la película “Star Wars Episodio 3”, la cual tuvo un éxito impresionante, sin embargo a final de temporada lo no vendido a principio de año, y falta de más mercancía de Star Wars. La competencia cerró muy fuerte con “Batman” y “Harry Potter”, en el caso de Niñas, fue bueno para “Nenuco” y “Cabagge Pach”, y para Mattel el boom de “Bratz”, le empieza a causar dolores de cabeza, ya que “Barbie” debe de ser rebajada en algunos modelos. Preescolar se agota en muchas tiendas, sin importar la marca, y Leap Frog hace un excelente año, poco producto en el mercado de construcción, que sigue sin ser tan importante como lo es en Europa, por ejemplo.

Para 2006, las cosas no fueron fáciles para Hasbro, Mattel tenía todos los juguetes de las películas de Moda como son “Cars” y “Superman”, por lo que tuvo que recurrir a sus líneas clásicas, por lo que relanzó algunos productos (refresh en el empaque), sobre todo el área preescolar, y bajo nuevamente precios en lo continuo en un 8%, lo que le orillo a tener un buen volumen durante todo el año con ciertos Clientes, sobre todo Wal Mart.

A fin de año se vendieron los productos caros como son los radio controles, Nenucos, y figuras de acción, aunque se tuvieron resultados similares a los del 2005, es decir sin ganar ningún punto de participación de mercado.

Lo importante es que sin película de “Star Wars”, y sin muchas novedades, Hasbro, tuvo un crecimiento de 2 dígitos y se mantuvo bien, y esta soportado sobre todo por Wal Mart, quien ya representa un 45% de todo el mercado.

Como se observa en este resumen de estos 14 años, el mercado juguetero es una variable muy caprichosa y difícil de predecir, por lo que debemos de ser muy cautelosos al considerar todas las variables que están dentro del cálculo de las estimaciones.

Para este 2007, se tienen una gran cantidad de películas para primavera-verano, y están repartidas, “Spider Man 3” se estrena en mayo, y “Los 4 Fantásticos”, los venderá Hasbro, “Shrek 3” lo Distribuirá MGA, “Harry Potter”, “Piratas del Caribe, Tortugas Ninja, “Ratatouille” de Disney-Pixar, por Mattel Lo que pensamos que para Temporada, no habrá concepto que domine, y será de muchos conceptos.

2.1.2 Datos del Histórico de Ventas 1993-2006

El histórico de ventas en pesos es el siguiente:

Histórico de Ventas

En millones

Año	*Millones $*	*% Inc.*
1993	$ 171.67	
1994	$ 283.25	65%
1995	$ 304.66	8%
1996	$ 397.91	31%
1997	$ 639.95	61%
1998	$ 825.20	29%
1999	$ 994.59	21%
2000	$ 954.27	-4%
2001	$ 1,182.83	24%
2002	$ 1,316.08	11%
2003	$ 1,499.02	14%
2004	$ 1,733.93	16%
2005	$ 1,801.70	4%
2006	$ 2,013.72	12%

El histórico de ventas en miles de piezas es el siguiente:

Histórico de Ventas

En miles de piezas

Año	*Piezas*	*% Inc.*
1993	4,539	
1994	6,926	53%
1995	5,311	-23%
1996	6,124	15%
1997	7,062	15%
1998	7,782	10%
1999	8,007	3%
2000	7,792	-3%
2001	8,891	14%
2002	10,319	16%
2003	11,801	14%
2004	11,507	-2%
2005	12,222	6%
2006	13,543	11%

2.2 Facturado por mes 1993-2006

2.2.1 Antecedentes por mes

En la facturación por mes observamos datos y tendencias interesantes de destacar, como son las siguientes:

Enero:

La facturación del mes está compuesta por el envió de pedidos de emergencia de los de ciertas Tiendas muy importantes en los días de Reyes, y actualmente empiezan a ser los llenados de planograma en el caso de Wal Mart

Febrero:

La facturación de este mes poco a poco empieza a ser más interesente, por la llegada de novedades y es el mes de altas y cambios de precio con los clientes.

Este mes empieza la reactivación de compra con la gran mayoría del mercado, y se rellena con productos continuos agotados en la temporada (50%) y novedades (50%)

Marzo:

Es el mes que se factura la partida de fuerte de la mini temporada del “Día del Niño”, donde por ser una época de reajustes en la economía de los Papás se factura producto a precio de oportunidad, el cual es producto de la temporada, con un precio de venta rebajado hasta un 50%.

Abril:

La gran mayoría de los pedidos del "Día del Niño" son colocados en marzo, y se factura en este mes a clientes de mayoreo. También la facturación de las películas de verano son facturadas desde este mes, para tener un suficiente tiempo de exposición el en anaquel.

Mayo:

Este es mes de películas importantes generalmente y su facturación se concentra en estas películas, así como la apertura de la entrega de temporada a todos los mayoristas, y algunos departamentales que tengan tiendas sobre todo en el norte y pacífico.

Junio:

Este mes se le empieza a convertir en un mes de alta facturación de mayoristas con temporada abierta de meses pasados, además de lanzamientos por las películas de verano, así como eventos como mundiales ocurren en este mes que son motivos de lanzamientos y productos nuevos de lato precio que son enviados a todos los clientes que se les ha surtido temporada, También se libera a alg unos autoservicios con tiendas lejanas del centro del país como Chedraui (en Veracruz y Sureste), y Casa Ley (en pacífico),

Julio:

Mes de películas, por lo que se acaba de surtir órdenes de productos faltantes, llegada de mercancía de alto precio que es surtida a la mitad de los clientes, que ya están en temporada, ofertas y ventas especiales a Comercial Mexicana, que es el cliente con mayor atención de parte del consumidor ya que es una campaña muy fuerte en medios.

Agosto:

Se libera temporada de autoservicios en general, destacando el fuerte surtido que se hace en Comercial Mexicana, Soriana, y Chedraui. Las Tiendas Departamentales son consideradas dentro de las prioridades, aunque la mayoría ya habían recibido algo de Temporada en los meses pasados.

Septiembre:

Es un mes de surtido a todo el mercado (incluido el de mayor volumen que es Wal Mart, y es el mes de mayor trabajo en tiendas, ya que es el mes de inicio de armado de anexos, y espacios adicionales, y esto implica mucha planeación para que las llegadas de las novedades sean surtidas a todo el mercado

Octubre:

Es el mes que históricamente se ha llevado la mayor facturación desde la fundación de la empresa, y son los primeros 5 días del mes donde se entrega casi el 20% de la facturación de todo el año en el centro de distribución de Wal Mart, lo que muestra la importancia de este cliente y la planeación que se debe tener para atender a otros.

Noviembre:

Es el segundo mes de facturación fuerte del año después de octubre, este mes se termina de entregar un 95%, de todos los pedidos de temporada, los cuales casi todos cancelan en este mes, es un mes difícil ya que la estrategia de asignación y pronóstico al principio de año para todos los clientes fue satisfactoria o no.

Diciembre:

Este mes es el de entrega a clientes especiales (sobre todo del norte y pacifico), donde la temporada llega hasta el 24 de diciembre, ya que la gente no tiene la tradición de Reyes, y es donde se observan los primeros indicios de que tan buena se espera la temporada. Este mes se tenía planeado desaparecer su facturación para correr menos riesgos, sin embargo casi siempre se embarca lo faltante por colocar o para llegar al objetivo anual comprometido, y esto se hace con ciertos clientes especiales, con condiciones igualmente especiales. La desventaja de esto es que cada vez más el consumidor se espera a comprar en los días cercanos, en busca de ofertas o descuentos que les ayuden a no mermar más su economía.

2.2.2 Datos facturado por mes

Los pesos facturados por año y por mes de la compañía son los siguientes:

Facturado por Mes (miles de pesos)

Año	Ene	feb	Mar	abr	may	jun	Jul
1993	$ 163	$ 280	$ 967	$ 1,121	$ 6,343	$ 4,763	$ 10,230
1994	$ 2,346	$ 1,331	$ 3,954	$ 9,737	$ 15,786	$ 20,341	$ 16,299
1995	$ 7,631	$ 2,461	$ 4,084	$ 12,345	$ 4,171	$ 1,512	$ 12,046
1996	$ 5,149	$ 4,406	$ 7,048	$ 17,493	$ 12,196	$ 12,453	$ 18,158
1997	$ 8,775	$ 16,016	$ 28,110	$ 24,795	$ 21,613	$ 23,473	$ 32,862
1998	$ 10,317	$ 19,976	$ 34,103	$ 25,915	$ 32,779	$ 27,992	$ 42,044
1999	$ 13,033	$ 21,370	$ 45,590	$ 38,610	$ 34,094	$ 51,674	$ 49,908
2000	$ 5,400	$ 7,896	$ 54,838	$ 25,949	$ 42,237	$ 39,367	$ 37,463
2001	$ 1,205	$ 12,331	$ 41,477	$ 24,827	$ 35,967	$ 84,075	$ 66,620
2002	$ 8,149	$ 36,670	$ 54,912	$ 59,365	$ 78,301	$ 101,984	$ 62,940
2003	$ 11,013	$ 28,060	$ 72,045	$ 77,440	$ 121,501	$ 85,074	$ 75,003
2004	$ 13,120	$ 45,327	$ 84,721	$ 53,906	$ 88,487	$ 111,218	$ 93,638
2005	$ 6,843	$ 51,934	$ 88,821	$ 74,488	$ 104,259	$ 122,652	$ 121,439
2006	$ 15,229	$ 52,614	$ 98,340	$ 62,572	$ 102,712	$ 129,098	$ 108,692

Facturado por Mes (miles de pesos)

Año	ago	sep	oct	Nov	Dic	Total
1993	$ 14,346	$ 24,911	$ 30,571	$ 60,512	$ 17,464	$ 171,673
1994	$ 16,449	$ 31,765	$ 67,909	$ 92,109	$ 5,226	$ 283,252
1995	$ 9,817	$ 26,309	$ 108,318	$ 85,075	$ 30,895	$ 304,663
1996	$ 21,312	$ 85,020	$ 104,749	$ 104,529	$ 5,392	$ 397,906
1997	$ 66,714	$ 116,436	$ 166,727	$ 129,542	$ 4,887	$ 639,949
1998	$ 76,861	$ 141,514	$ 233,969	$ 173,259	$ 6,479	$ 825,207
1999	$ 98,655	$ 145,868	$ 248,656	$ 216,525	$ 30,608	$ 994,591
2000	$ 109,775	$ 126,963	$ 215,877	$ 298,466	$ (9,961)	$ 954,269
2001	$ 117,523	$ 216,185	$ 256,090	$ 275,054	$ 51,478	$ 1,182,832
2002	$ 126,987	$ 269,873	$ 282,609	$ 188,247	$ 46,044	$ 1,316,080
2003	$ 163,031	$ 216,955	$ 324,937	$ 262,051	$ 61,914	$ 1,499,023
2004	$ 209,856	$ 190,343	$ 411,637	$ 395,934	$ 35,746	$ 1,733,933
2005	$ 156,100	$ 151,630	$ 476,580	$ 417,618	$ 29,336	$ 1,801,700
2006	$ 243,591	$ 220,942	$ 477,811	$ 460,630	$ 41,492	$ 2,013,722

2.2.3 Comentarios relevantes de los importes facturados por mes

Como se podrá observar los meses presentan movimientos característicos de series de tiempo como son:

- Los meses de septiembre, octubre y noviembre presentan movimientos **estacionales** que son las grandes entregas de Pedidos de temporada ya que el producto debe de estar para su venta al consumidor que es en los meses de diciembre y enero (los cuales son los meses de facturación y ventas más pequeños de todos los meses del año).

- Los meses de marzo y abril presentan movimientos **cíclicos** de mejoría en comparación con los dos meses anteriores y los dos meses posteriores, que salvo la facturación de temporada no tienen tanta representatividad en el año.

- Diciembre presenta un comportamiento **irregular**, en algunos años su facturación es poco representativa, más sin embargo en 93 y 95 representó el 10% de la facturación anual, lo cual por experiencia es malo por tener poco tiempo de exposición en el punto de venta, y por lo tanto la oportunidad de desplazarse se vuelve más baja. En el 2000 se tuvo un retroceso tan fuerte producto del mal desempeño de 1999 que se tuvo ventas negativas que fueron Pedidos que llegaron en diciembre pero que habían cancelado desde noviembre. Del 2001 al 2004, la facturación de diciembre queda a criterio de la Dirección de Ventas que generalmente es lo faltante para alcanzar el objetivo. La intención en un futuro es entregar mercancía de novedad del siguiente año, para con esto empezar a dar a conocer novedades del año entrante, lo que ayuda a mantener un espacio permanente.

- Los meses de agosto y septiembre tienden a largo plazo a **incrementar** su participación en la facturación de temporada, por lo que se debe planear la pronta llegada del producto al almacén.

2.3 Facturado por Cliente 1993-2006

La facturación por cadena incluirá algunos clientes en los cuales sólo se tendrán datos a partir de 1994, año que se empezó a trabajar con todo el mercado, los datos los presentaremos por canal de distribución:

FACTURACION POR CADENA

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Wal Mart/Aurrera	38	67	87	100	155	160	148	165	173	238	340	367	385	445
Bod Aurrera	14	25	50	52	69	77	156	115	175	214	294	334	342	394
Com Mex	14	26	35	47	77	113	142	132	139	113	146	158	154	168
Gigante	16	33	38	48	54	94	128	50	98	109	24	32	48	51
Soriana	4	5	6	17	35	58	59	45	67	75	107	128	144	169
Carrefour	-	4	5	10	22	24	21	27	32	39	47	60	-	-
Chedraui	6	6	10	9	16	32	41	46	45	45	52	89	126	145
Bog Comex	1	4	4	7	11	10	13	8	20	22	35	34	38	41
Bga Gigante/Blanco	3	7	4	5	10	6	14	13	19	31	3	1	2	5
Casa Ley	-	2	4	5	11	13	38	41	30	24	24	29	39	47
Auchan	-	-	-	1	1	2	2	2	3	5	-	-		-
Otros	1	6	8	11	28	4	4	43	37	50	39	24	2	4
Total Autoservicios	**97**	**186**	**249**	**312**	**489**	**594**	**765**	**685**	**838**	**965**	**1,112**	**1,255**	**1,280**	**1,467**

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Liverpool	7	6	11	11	28	32	32	41	48	69	102	103	112	113
Sanborn´s	5	5	6	7	7	18	20	19	20	24	33	38	40	42
Palacio de Hierro	4	4	4	6	9	19	24	12	14	19	24	27	27	27
Sears	-	-	-	-	6	12	12	16	14	18	21	24	27	28
Colonial de Mex	2	2	2	2	4	4	3	3	4	3	2	-	-	-
Sybele	-	-	-	1	2	2	2	1	1	-	-	-	-	-
Woolworth	1	1	0	1	2	6	10	12	18	-	-	-	-	5
Otras	1	3	0	0	2	3	2	4	4	3	-	3	3	4
Total Departamentales	**19**	**21**	**24**	**30**	**60**	**95**	**105**	**108**	**122**	**136**	**180**	**194**	**209**	**218**

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Juguetibici	10	15	8	10	21	22	19	22	15	21	25	27	15	18
Julio Cepeda	9	12	4	10	21	37	32	41	47	48	34	48	57	63
Merc del Refugio	11	11	4	9	15	13	9	10	15	-	-	-		-
Volcanes	3	6	-	6	12	12	12	0	-	-	-	-		-
Juguetrón	3	6	-	5	9	12	11	9	19	22	30	35	35	41
Tony	3	5	0	4	5	6	5	5	3	3	10	5	3	4
Centro Juguetero	-	-	-	-	-	-	8	8	6	14	10	20	17	19
Toy´s Land	-	-	-	1	1	2	1	2	-	-	-	-		-
Otros	1	5	6	3	(2)	2	2	0	0	-	5	14	55	30
Total Mayoristas	**40**	**61**	**22**	**47**	**81**	**106**	**98**	**97**	**105**	**109**	**114**	**149**	**182**	**175**
Otros/Frontera	15	16	10	9	10	30	27	65	117	107	93	136	131	154
Total Cia	**172**	**283**	**305**	**398**	**640**	**825**	**995**	**954**	**1,183**	**1,316**	**1,499**	**1,734**	**1,802**	**2,014**

2.4 Hipótesis Econométricas, Demográficas y Financieras

2.4.1 Inflación e Incremento de Lista de Precios

Se consideran las siguientes tasas promedio de incremento de lista de precios de 1993 a 2006 y las reales de inflación publicadas en internet por el Banco de México y nuestra proyección para el 2007-2009, es importante destacar que la lista de precios esta generalmente un poco arriba e la inflación pero cada vez en menor proporción.

Estos parámetros servirán para proyectar las ventas en pesos que se espera por efecto directo de este factor. Se estudiará en el Capítulo 5 la relación entre el histórico de ventas y los indicadores económicos del país, a fin de detectar similitudes, oportunidades y proponer soluciones.

El incremento real de ventas será medido como el porcentaje adicional de ventas por mayor volumen de unidades con respecto al año pasado.

2.4.2 Datos de inflación e incremento de lista de precios

Los datos de inflación publicados por el Banco de México, y los incrementos anuales que se registraron en la compañía juguetera fueron los siguientes:

AÑO	INFLACIÓN	INC. DE PRECIOS PROM
1993	8.01%	11.0%
1994	7.05%	9.00%
1995	51.97%	52.00%
1996	27.70%	29.00%
1997	15.72%	16.00%
1998	18.61%	25.00%
1999	12.32%	13.0%
2000	8.96%	9.0%
2001	4.40%	5.0%
2002	5.70%	6.0%
2003	3.98%	4.5%
2004	5.19%	6.2%
2005	3.33%	2.0%
2006	4.05%	4.09%
2007	**4.02%**	**4.05%**
2008	**3.90%**	**4.00%**
2009	**3.85%**	**4.00%**
2010	**3.80%**	**3.09%**

Fuente: Banxico.org.mx y Listas de Precio HASMEX Promedio en el año
Negrillas: Proyectado

2.4.3 Crecimiento del PIB y del Salario Mínimo del D.F.

AÑO	PIB	% Variación PIB	Salario Mínimo del D.F. ($)
1993	100.00		14.27
1994	106.40	6.40%	15.27
1995	131.20	23.31%	16.34
1996	175.60	33.84%	20.65
1997	212.10	20.79%	26.47
1998	248.00	16.93%	30.20
1999	290.80	17.26%	34.45
2000	332.70	14.41%	37.90
2001	357.70	7.51%	40.35
2002	384.30	7.44%	42.15
2003	419.20	9.08%	43.65
2004	440.10	4.99%	45.24
2005	457.30	3.91%	46.80
2006	474.80	3.83%	48.67
2007	**491.30**	**3.48%**	**50.57**
2008	**510.46**	**3.90%**	**52.85**
2009	**530.11**	**3.85%**	**55.00**
2010	**550.26**	**3.78%**	**57.15**

1993=100% Fuente Banxico
Negrillas: Estimado

2.4.4 Grupos Quinquenales de Edad (Posibles Compradores)

Edad Quinquenal	De 0 a 4 años	De 5 a 9 años	De 10 a 14 años	Total Población
Hombres:	**5,175,913**	**5,339,913**	**5,545,910**	**50,249,955**
Mujeres:	**5,010,330**	**5,172,611**	**5,406,213**	**53,013,433**
Total:	**10,186,243**	**10,511,738**	**10,952,123**	**100,263,388**

Fuente INEGI: Censo de Población y vivienda 2000

Se observa que en los grupos quinquenales de 0 a 4 y el de 5 a 9 existen mas varones, y ya a partir de 10 en adelante son más mujeres, esto es una de las razones del por que, los Padres gastan más en juguetes de niños que de niñas.

Esto sumado a otros factores como que los niños coleccionan más juguetes que las niñas (las niñas coleccionan generalmente un solo personaje, y no importa si es juguete, toalla, camiseta, etc.), además los niños en promedio juegan más tiempo de su vida, con juguetes que en promedio lo hacen las niñas, ya a que las niñas les gusta antes la ropa, la moda, pintarse, etc.

2.5 Breve resumen de las condiciones macroeconómicas de México de los años 1993-2006

México ha sufrido crisis financieras durante toda su historia, en enero de 1992 a noviembre de 1994[6], el país tenía una economía estable, los principales indicadores macroeconómicos como son el P.I.B. y el P.N.B, presentaban incrementos entre un 4 y 5%, la inflación andaba cerca de un dígito, y aparentemente con la firma del Tratado de Libre Comercio las condiciones serían muy favorables y se estaba encaminando a una supuesta prosperidad.

El comercio y la industria constituyen en este país la primera fuente de riqueza en cuanto a valor de producción.

El comercio exterior todavía tenía una importancia relativa en el P.I.B. el petróleo, café algodón y azúcar constituyen los principales rubros de exportación hacia los E.U. que es el principal cliente y proveedor del país.

El tratado de libre comercio firmado en agosto de 1992, creó fuertes expectativas económicas, pero a corto plazo las tensiones especulativas provocaron una gran devaluación del peso que llevó a la economía mexicana a una profunda crisis.

La devaluación del peso frente al dólar americano fue de más de un 50% y se tuvo que implementar un plan económico de emergencia para reducir el gasto público y el déficit de la balanza comercial.

Esta devaluación golpeó en todos los mercados, sobre todo en Empresas que importan sus productos en dólares y que se vio en el incremento de precios que se

[6] **Sistema Financiero de México, Eduardo Villegas; Rosa Ma. Ortega Mc Graw Hill, México 2005**

Determinó, para ese año, y las ventas de 1995 que tuvo la compañía con respecto a 1994, fue con un crecimiento real por debajo de la inflación.

En 1996 con los planes de ajuste, implantados por el Gobierno, que pretenden seguir con el control de la paridad peso-dólar y conservar la estabilidad financiera, para mantener y captar nuevos capitales extranjeros y la especulación financiera.

Sin embargo la dependencia económica de México, por Capitales extranjeros sigue siendo hasta nuestros días un dolor de cabeza ya que cualquier especulación política genera una especulación financiera, y las fugas de capitales se pueden dar en cualquier momento, como sucedió e el famoso error de diciembre de 1994. [7]

Una protección que tiene la Compañía que estamos estudiando es poner como cláusula en sus condiciones de venta, que pueden incrementar sus precios en cualquier momento basados al tipo de cambio del dólar frente al peso en el momento de facturar pedidos de temporada, los cuales como hemos comentado se establecen con meses antes de la Temporada y de una posible devaluación.

En 1998 nuestro país pierde con China el ser el primer país exportador de los E.U, además este año, la Empresas Transnacionales cierran maquiladoras en México y se las llevan a China, por tener una mano de obra más barata. La Compañía que estamos analizando cerro plantas en Tijuana, y dejó de maquilar algunos Juegos de Mesa y Componentes en México, y traer todo o casi todo de China.

En 1999 el 95% del producto llega de China, lo que provoca que se tenga que pedir altas cantidades de producto ya que China establece cantidades mínimas para producir y le da preferencia a E.U, Japón, Europa, y hasta el último América Latina. Esto supone dos peligros, traer mercancía de **más** y que no se venda y se tenga que rematar o recoger en devolución o que la cantidad que se traiga o sea **insuficiente** y dejar de surtir por no tener más producto por lo menos 6 meses, ya que es el tiempo de producción requerido por China, y como sabemos el tiempo de vida de un juguete es muy corto y el tiempo es vital en este negocio. Un producto que sea muy demandado en septiembre, es seguro que no llegara a Navidad y Reyes, pero es imposible reaccionar y la fecha de llegada de este sería hasta febrero del siguiente año.

En el 2000, la Empresa tuvo un desajuste ya que las líneas fuertes de año precio se cayeron y no se tenían las novedades del 99’ como lo fueron Pokémon, Furby y Satr Wars, (que representaban el 50% del negocio), además se tenían problemas de credibilidad con los Clientes en el renglón de las entregas, y su tuvo que hacer una reingeniería de los procesos. Sin embargo el efecto electoral de este año hace que se reactive la economía y esto provocó un excedente de

---

[7] **Evolución del Sistema Financiero Mexicano Capítulo 1 Sistema Financiero de México, Eduardo Villegas; Rosa Ma. Ortega Mc Graw Hill, México 2005**

Efectivo, que fue utilizado para la compra de productos no básicos y parte de esto, se destinó a los juguetes, sobre todo los de precios accesibles, de la euforia democrática provocó que mucha gente gastara más de la cuenta y como las tasas de interés empezaron a bajar, muchos clientes como Liverpool, Palacio de Hierro, Sears, etc. ofrecían los planes de meses sin interés. Además empezaba la estrategia de "Precios Bajos todos los días" de Wal Mart [8] , que revoluciono el mercado al quitar una costumbre comercial muy arraigada, el descuento.

En el 2001, ya se habían consolidado muchas marcas de la Compañía, habían pasado los años de mala imagen con Clientes, se establecieron nuevos Convenio de Condiciones Comerciales con Clientes Claves, lo que permitió contar con más volumen, y medir el potencial que si explotaban otras Compañías. Adicional a esto la compañía empezó a manejar líneas de terceros que le dieron mejor penetración en el mercado, aunque todavía muy lejos del líder Mattel.. El mercado empieza a estabilizarse económicamente el Gobierno estabiliza la inflación en un dígito, lo cual hace que los factores macroeconómicos tengan menor peso en desempeño de las ventas.

En el 2002, se dio el año del ajuste económico, es decir se redujo la línea, se quedaron productos continuos que habían demostrado en años anteriores que seguían en la mente del consumidor, además se manejan licencias importantes de terceros entre las que destacan marcas muy conocidas como son las Bratz, Nenuco, Muñecos y Figuras de Películas de Disney, con lo que las ventas del segmento de niñas y una parte de niños pequeños crecieron para esta empresa un 100%.. Mattel ese año consolidó más su mercado Niños con Hot wheels y Figuras Spider Man, las cuales fueron un hit por estreno de la película.

En 2003 fue desde el punto de vista de rentabilidad el mejor año de la Compañía, ya que se alcanzó superar la cifra de los cien millones de dólares, una línea líder en todo el mercado fue Beyblade, la cual como hemos comentado anteriormente se vendía con niveles de un básico en los autoservicios, y fue tan fuerte su impacto que llego a ser un problema surtir y después el mercado lo empezó a rechazar, la teoría es que alcanzó su techo y en temporada un producto de bajo precio no es atractivo aunque sea de moda Sin embargo el resultado anual de desplazamiento promedio estuvo arriba del 85%, lo que permitió negociar y reactivar con la mayoría de los clientes, desde enero del 2004

En el 2004, se tenían muchas expectativas de crecimiento, pero un panorama difícil ya que se esperaba que Beyblade continuara más tiempo lo cual no ocurrió y lo que se tuvo que hacer es lanzar diversas marcas de productos del segmento de niños para compensar esa baja del mercado, además se trabajó en un lanzamiento de productos electrónicos de lato precio como lo fueron Video Now y Fur Real Friend productos netos de temporada

[8] **Excélsior, mayo 2000 sección Financiera y Negocios**

El resultado fue un año regular ya que las ventas estuvieron contraídas económicamente, y fue una temporada atípica ya que no se mencionó ningún juguete o línea como clave o dominante en ese año y quedo mucho inventario excedente el cual todavía en muchos casos se tuvo que aceptar niveles muy altos de Devoluciones.

El 2005, fue un año de Película con Star Wars, lo que hizo que mejorará las ventas sobre todo en los meses de abril y mayo, donde un 70% de las ventas en estos meses eran de productos de esta Licencia, en la temporada, estuvo apoyada en productos de alto precio como los Nenucos, Radio Controles, y la Línea de Action Man que cada vez gana más terreno a Mattel, estimamos que esta en este año 2 a 1, es decir por cada dos Max Steel de Mattel, se vende un Action Man de Hasbro.

El resultado del año fue bueno sin ser espectacular, se tuvo que trabajar en rematar muchos productos y sobre todo de Juegos de Mesa, líneas como los Preescolares se agotaron.

El 2006, representó la llegada de nuevas marcas para Hasbro como figuras de Spider Man y Marvel Hérores, lo que le dio nuevos bríos a la categoría de Figuras de Action, Mattel, sin embargo Mattel manejo todas las películas de Verano como fueron Superman, Cars, Piratas del Caribe (entre otras), lo que dificultó el desplazamiento de las figuras, ya que le ofrecimos al Consumidor demasiadas propuestas.

Finalmente, 2007 representa un año lleno de retos ya que se tienen muchas marcas en el mercado y la venta esta muy diluida entre todas las marcas de las distintas Compañías, el mercado esta muy contraído afectado por la dependencia de nuestra economía a la de los E.U. que esta teniendo una desaceleración de la economía y un aumento mayor de inflación de lo esperado, reduciendo márgenes de utilidad que limitan nuevas plazas de trabajo, y por ende aI consumo .y si consideramos que el segundo monto de Divisas de nuestro país lo recibimos de los residentes en E.U. el problema se ve gris, más sin embargo las cadenas establecidas en México le están invirtiendo fuerte en Aperturas con lo que abrirá el abanico de posibilidades de venta, pero la mayoría de las cadenas a Tiendas iguales se estima que tendrán un incremento de ventas promedio de cero, durante el año.

Capítulo 3

Justificación del modelo

Capítulo 3

JUSTIFICACIÓN DEL MODELO

3.1 Cálculo de parámetros estadísticos del histórico de ventas

Obtención de la media, varianza, y desviación estándar de ventas 1993-2006.

3.1.1 Definiciones

Media: La media $\overline{X}$, de la muestra es el primer momento, cuyo estimador es el sumar todos los datos y y dividir el resultado por el número total de ellos es decir:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

En donde Xi para el presente trabajo es el importe de venta de cada uno de los (i) años, que se están promediando. El promedio o media mensual de ventas del período 1993-2006

Varianza:

La Varianza es una media de la Dispersión de los datos respecto a la media, se obtiene restando la media a cada valor, elevando al cuadrado el resultado, sumando todos los valores ahí obtenidos y dividiendo la suma entre el total número de datos, es decir::

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Se utilizará en realidad la que se conoce como la cursivarianza que se obtiene dividiendo entre n-1 y no entre n.

Desviación Estándar

El cálculo de la Desviación Estándar que es la raíz cuadrada de la Varianza.

El cálculo de estos tres parámetros para los 14 años de estudio es el siguiente:

Facturación en Pesos

Mes	Media a 14 años	Varianza a 14 años	Desviación Estándar
ene	7,740.89	21,064,414	4,589.60
feb	21,476.37	349,888,042	18,705.29
mar	44,215.10	1,108,917,542	33,300.41
abr	36,326.04	618,306,350	24,865.77
may	50,031.84	1,641,196,931	40,511.69
jun	58,262.47	2,108,537,334	45,918.81
jul	53,381.48	1,311,331,907	36,212.32
ago	102,215.47	5,473,563,784	73,983.54
sep	140,336.73	6,114,107,649	78,192.76
oct	243,317.13	20,483,748,478	143,121.45
nov	225,682.25	17,163,944,856	131,011.24
dic	25,499.93	444,437,548	21,081.69
total	1,008,485.70	369,118,081,168	607,550.89

Facturación en Piezas

Mes	Media a 14 años	Varianza a 14 años	Desviación Estándar
ene	53,583.36	1,958,757,241	44,257.85
feb	370,426.04	24,065,488,265	155,130.55
mar	708,865.54	62,193,860,053	249,386.97
abr	516,086.19	31,887,059,955	178,569.48
may	651,241.56	82,771,603,795	287,700.55
jun	804,716.49	65,401,287,942	255,736.76
jul	648,164.88	42,037,814,328	205,031.25
ago	824,255.30	127,153,991,335	356,586.58
sep	803,369.93	214,707,615,090	463,365.53
oct	1,632,369.01	338,843,582,834	582,102.73
nov	1,358,458.44	175,556,828,564	418,995.02
dic	330,432.77	141,085,043,494	375,612.89
total	8,701,969.50	7,647,373,119,832	2,765,388.42

Facturación en Pesos/Piezas Promedio de Venta

Mes	Media a 14 años	Varianza a 14 años	Desviación Estándar
ene	168.90	10,536	102.64
feb	51.41	1,026	32.03
mar	54.11	1,103	33.22
abr	63.04	840	28.98
may	66.70	1,277	35.73
jun	64.16	1,724	41.52
jul	74.54	1,158	34.03
ago	119.59	7,222	84.98
sep	177.70	5,516	74.27
oct	143.50	3,973	63.03
nov	156.54	3,827	61.86
dic	121.08	8,677	93.15
total	105.63	1,611	40.14

3.1.2 Comentarios hacia la media

- Se observa una gran discrepancia entre los resultados, los meses de enero y febrero se tiene en promedio un bajo porcentaje de participación en las ventas del total año, aunque cada vez empiezan a ser un poco más importantes por la llegada de producto nuevo a fines de febrero.
- El siguiente bimestre, marzo y abril tienen un promedio un poco mejor que enero y febrero por la mini temporada del Día del Niño, el crecimiento se observa sobre todo en marzo que es le mes de colocación de los Pedidos de dicho evento, y en abril algunas llegadas de novedades de alguna película de verano se han vuelto indispensables para atraer la atención de lo pequeños en una fecha cada vez menos recordada por los papás, sobre todo cuando coincide con semana santa, ya que en lugar de gastar en juguetes, los papás deciden un viaje de vacaciones.

- El siguiente bimestre empieza la facturación de pedidos de temporada de clientes denominados mayoristas, los cuales por su reventa que hacen a otros clientes, requieren mercancía para ofrecer mercancía, ya sea por catálogo a través de su fuerza de ventas, los cuales se concentran en provincia principalmente, por lo tanto requieren de tener producto antes, por que su venta se hace por medio del apartado. Esta facturación incrementa mucho el número, y representa más del 60% de la facturación de estos dos meses.

- Los meses siguientes que son julio y agosto son generalmente más fuertes que los anteriores, ya que se surten por una parte lo pendiente de las órdenes abiertas de los Mayoristas, y por otra generalmente hay un incremento natural de venta por alguna película, además de ser la época de promociones de Clientes Importantes como son Comercial Mexicana con su famoso “Julio Regalado”,y Liverpool en la denominada “Gran Barata”, además se libera la Facturación con plazo de temporada de Clientes Departamentales como Palacio de Hierro, Sears, Sanborn´s.

- Septiembre y Octubre son meses de mucho trabajo ya que continua, la Facturación de los Pedidos abiertos de los Departamentales y Mayoristas, además de que en septiembre se empieza los surtidos de Autoservicios fuertes como Comercial Mexicana, Chedraui, Casa Ley, Gigante. Y Octubre es el mes más fuerte de facturación por iniciar la Facturación del Grupo más poderoso actualmente de los Autoservicios, Wal-Mart, que en sus dos formatos Super Center y Bodega Aurrera los cuales reciben hasta el 70% de su orden de Temporada en estos dos meses.

- Los últimos meses del año son contrastantes, mientras que en noviembre se factura los faltantes de todos los Clientes, y se factura casi un 70% de lo facturado en octubre, lo que lo convierte en el segundo mes en facturación de acuerdo al histórico de ventas, diciembre se vuelve un mes de casi nada de facturación, si acaso de algún producto muy importante que se haya agotado y que se tengan algún remanente, y generalmente existen casos aislados de Tiendas que requieran resurtidos, sobre todo en Tiendas de Navidad, donde el desplazamiento es únicamente en diciembre.

3.1.2 Comentarios hacia la varianza

- Como se observa en las tabla, la varianza es dramáticamente elevada (incluso por piezas es mayor a la de por pesos por lo que sólo se trabajara en pesos para las transformaciones).

- La varianza creciente muestra que año con año, los históricos de venta cambian drásticamente, además como cada vez se adelanta la facturación de Temporada más Clientes, antes del término del primer semestre del año, como sabemos la varianza alta indica que la media esta muy lejos ser un buen estimador del comportamiento de las ventas, muy por debajo de la facturación de los últimos años y muy elevada para los primeros años. Esto sucede por que en los primeros años, las Compañía no tenía ventas al todos los Clientes, y además que la venta estaba muy concentrada a sólo finales de año y en los años recientes, la Facturación fuerte que son los Pedidos de Temporada, se facturan cada vez antes, fenómeno que se concentra en Clientes Departamentales y de Mayoreo, además de algunos autoservicios locales que anticipan también su entrega como son Casa Ley, Comercial Mexicana, Chedraui, Soriana.

- Octubre es el mes de mayor facturación y por lo tanto tiene la mayor desviación con respecto a la media, esto se debe a que como la Demanda del producto es creciente y la variación de pesos año con año es fuerte, en octubre por su volumen se presenta más fuerte cada día, más adelante se trabajará el concepto de Tendencia para determinar si existe este fenómeno.

3.1.4 Frecuencia y Moda

- Estos dos conceptos estadísticos no se detectan con facilidad, ya que año con año se modifican los importes de ventas y es difícil encontrar valores semejantes, incluso meses con una facturación similar en dos o más años, de los 12 estudiados en el presente trabajo.

- Como observación encontramos que una manera de estabilizar la varianza es utilizar algunas transformaciones como puede ser la logarítmica o alguna raíz cuadrada o una potencia, y hacer diferencias estacionales.

3.1.5 Transformaciones para los Datos

Las transformaciones que se operan frecuentemente son:

1. Raíz cuadrada: $W_t = \sqrt{Z_t}$

2. Cambio porcentual: $Z_t\% = 100*(Z_t - Z_{t-1}) / Z_{t-1}$

3. Cambio porcentual en Logs: $Z_t\% = 100* \text{Log}(Z_t / Z_{t-1})$

   Note que: $100* \text{Log}(Z_t / Z_{t-1}) = 100*[\text{Log}(Z_t) - \text{Log}(Z_{t-1})]$

   Si el crecimiento de Z es chico los dos caminos dan resultados muy similares ya que al desarrollar hasta orden dos en la serie de Taylor se tiene:

   $\text{Log}(Z_t / Z_{t-1}) \sim [Z_t - Z_{t-1}] / Z_{t-1} - \{ [Z_t / Z_{t-1}] -1\}^2 /2$

   El termino cuadrático $[\{ [Z_t / Z_{t-1}] -1\}^2 /2]$ es chico si el crecimiento es moderado.

4. Logaritmos $Z_t = \text{Log}(Z_t)$ este requiere que $Z_t > 0$.

5. Diferencias de logaritmos:

   $W_t = \text{Log}(X_t) - \text{Log}(Y_t) = \text{Log}(X_t / Y_t)$

6. Primeras diferencias $DZ_t = (Z_t - Z_{t-1})$

7. Segundas diferencias $D^2Z_t = DZ_t - DZ_{t-1}$

   o sea: $D^2Z_t = Z_t - 2 Z_{t-1} + Z_{t-2}$

8. D-ésimas diferencias $D^dZ_t = (1-B)^dZ_t$

9. Diferencia estacional $(1 - B^4) Z_t = Z_t - Z_{t-4}$

10. Diferencia estacional $(1 - B^{12}) Z_t = Z_t - Z_{t-12}$

Las dos últimas son usadas con datos trimestrales y mensuales respectivamente, lo que hacen es filtrar la componente estacional, es decir la eliminan.
La transformación Logarítmica sería de la siguiente forma:

Aplicando primero Logaritmos (4) a la serie y diferencias a estos logaritmos (6) obtenemos los siguientes datos:

Facturación en Pesos usando Logaritmo Natural

Mes	Media a 14 años	Varianza a 14 años	Desviación Estándar
Ene	168.90	10,536	102.64
Feb	51.41	1,026	32.03
Mar	54.11	1,103	33.22
Abr	63.04	840	28.98
May	66.70	1,277	35.73
Jun	64.16	1,724	41.52
Jul	74.54	1,158	34.03
Ago	119.59	7,222	84.98
Sep	177.70	5,516	74.27
Oct	143.50	3,973	63.03
Nov	156.54	3,827	61.86
Dic	121.08	8,677	93.15
total	105.63	1,611	40.14

Facturación en Pesos usando Diferencias de Logaritmo Natural

Mes	Media a 14 años	Varianza a 14 años	Desviación Estándar
Ene	168.90	10,536	102.64
Feb	51.41	1,026	32.03
Mar	54.11	1,103	33.22
abr	63.04	840	28.98
may	66.70	1,277	35.73
jun	64.16	1,724	41.52
jul	74.54	1,158	34.03
ago	119.59	7,222	84.98
sep	177.70	5,516	74.27
oct	143.50	3,973	63.03
nov	156.54	3,827	61.86
dic	121.08	8,677	93.15
total	105.63	1,611	40.14

3.2 Cálculo de promedios variables simples

Dado un conjunto de números:
$X_1, X_2, X_3 ...$

Definimos un promedio móvil de orden N como la sucesión de medias aritméticas:

$$\frac{X_1 + X_2 + X_3 ... + X_N}{N}, \frac{X_2 + X_3 + X_4 ... + X_{N+1}}{N}, \frac{X_3 + X_4 + X_5 ... + X_{N+2}}{N}, ...$$

Las sumas en el numerador de la sucesión se llaman totales móviles de orden N, el recorrido de los valores será desde el año de inicio de la proyección hasta el año de promedio que se requiera (dos o tres años), a continuación se darán para cada mes del período 1993-2004 el promedio para dos años variable por mes y su pronóstico considerando los promedios móviles de dos años para 2007-2009.

Los cálculos de los promedios móviles para 3 años considerando los 12 años del histórico de ventas quedarían como sigue:

PESOS Promedio de Medias Móviles	**ene**	**feb**	**mar**	**abr**	**may**	**jun**	**jul**
Promedio últimos 10 años	9,308	29,219	60,296	46,787	66,195	77,660	69,061
Promedio últimos 7 años	8,708	33,547	70,736	54,078	81,923	96,210	80,828
Promedio últimos 6 años	9,260	37,822	73,386	58,766	88,538	105,683	88,055
Promedio últimos 5 años	10,871	42,921	79,768	65,554	99,052	110,005	92,342
Promedio últimos 3 años	11,730	49,958	90,627	63,656	98,486	120,989	107,923

PESOS Promedio de Medias Móviles	**ago**	**sep**	**oct**	**nov**	**dic**	**total**
Promedio últimos 10 años	136,909	179,671	309,489	281,733	29,802	1,296,131
Promedio últimos 7 años	160,980	198,984	349,363	328,286	36,578	1,500,223
Promedio últimos 6 años	169,515	210,988	371,611	333,256	44,335	1,591,215
Promedio últimos 5 años	179,913	209,948	394,715	344,896	42,906	1,672,891
Promedio últimos 3 años	203,182	187,638	455,342	424,727	35,525	1,849,785

3.3. Estimación de ventas 2006-2009, a partir de histórico de datos históricos de ventas, con promedios móviles y regresión lineal.

3.3.1 Medias Móviles

3.3.2

Tomando como base a los datos reales de las ventas obtenidas los años 1993-2006, se calculó un incremento promedio de ventas que se utilizó sobre las medias móviles para obtener los siguientes resultados para ser una primera estimación de las ventas del Período 2006-2009 considerando además la inflación acumulada en esos períodos.

Las Tablas estarán considerando períodos móviles de 5 y 3 años.

Pronóstico 2007 **Considera Inflación acumulada de 3 y 5 años**

PESOS

Promedio de Medias Móviles	**ene**	**feb**	**mar**	**abr**	**may**	**jun**	**jul**
+ Inc de Ventas Lista de Precios							
Promedio últimos 3 años	13,205	56,238	102,019	71,657	110,866	136,198	121,489
Promedio últimos 5 años	13,289	52,471	97,516	80,140	121,091	134,481	112,888

PESOS

Promedio de Medias Móviles	**ago**	**sep**	**oct**	**nov**	**dic**	**total**
Promedio últimos 3 años	228,722	211,224	512,579	478,116	39,990	2,082,303
Promedio últimos 5 años	219,943	256,662	482,539	421,636	52,453	2,045,110

Pronóstico 2008

Promedio de Medias Móviles	**ene**	**feb**	**mar**	**abr**	**may**	**jun**	**jul**
+ Inc de Ventas Lista de Precios							
Promedio últimos 3 años	13,889	59,151	107,304	75,369	116,609	143,253	127,782
Promedio últimos 5 años	14,079	55,589	103,312	84,903	128,288	142,474	119,598

PESOS

Promedio de Medias Móviles	**ago**	**sep**	**oct**	**nov**	**dic**	**total**
Promedio últimos 3 años	240,441	222,047	538,842	502,613	42,039	2,189,340
Promedio últimos 5 años	233,016	271,917	511,219	446,696	55,571	2,166,665

Pronóstico 2009

Promedio de Medias Móviles	**ene**	**feb**	**mar**	**abr**	**may**	**jun**	**jul**
+ Inc de Ventas Lista de Precios							
Promedio últimos 3 años	14,439	61,493	111,552	78,353	121,225	148,924	132,840
Promedio últimos 5 años	14,541	58,767	108,392	85,053	129,371	147,996	126,598

PESOS

Promedio de Medias Móviles	**ago**	**sep**	**oct**	**nov**	**dic**	**total**
Promedio últimos 3 años	250,045	230,916	560,365	522,689	43,718	2,276,559
Promedio últimos 5 años	235,365	274,658	516,373	451,200	56,131	2,204,446

3.3.2. Estimación por Mínimos Cuadrados

El método consiste en ajustar los datos a una recta cuya ecuación es la siguiente:

$$Y = \bar{Y} + \left(\frac{XY}{X^2}\right) X * (1 + inc\_vta)$$

Donde X representa los años de cálculo e Y los montos de venta de cada uno esos años. Además se incorporo una variable llamada incremento de Venta que es un factor de inflación que incrementa los precios año con año.

Para estimar 2007, se consideraron los años 93-2006, para 2008 se considera desde el 93, hasta el estimado 2007., y así consecutivamente. Utilizando cada uno de los años de proyección se obtuvieron los siguientes valores:

	Pronóstico contra Real 2006			**Inc de Prec**
Pronóstico	**Regresión Lineal Simple**			3.50%
X	**y**	**y *(1+inc_prec)**	**Real 2006**	**% Var**
1	11,830	11,830	15,229	-22.32%
2	46,960	46,960	52,614	-10.75%
3	92,917	92,917	98,340	-5.52%
4	70,141	70,141	62,572	12.10%
5	105,679	105,679	102,712	2.89%
6	124,004	124,004	129,098	-3.95%
7	105,311	108,997	108,692	0.28%
8	208,578	215,878	243,591	-11.38%
9	239,774	248,166	220,942	12.32%
10	452,022	467,842	477,811	-2.09%
11	409,130	423,449	460,630	-8.07%
12	43,632	45,159	41,492	8.84%
Total 06	**1,909,977**	**1,961,023**	**2,013,723**	-2.62%

Donde X=1, es el mes de Enero y X=12 es diciembre por ejemplo.

Conclusiones:

El modelo de Regresión Lineal Simple, en total, tiene da una buena aproximación por abajo de los valores observados con los reales 2006, esto se debe principalmente a que los crecimientos de la Compañía han sido de un dígito en los último 3 años, sin embargo, no es un buen estimador de meses con mucha inestabilidad en las ventas como es enero y febrero.

Además en los meses de películas (caso julio/agosto), se presenta un fenómeno que no contemplado que es la estacionalidad, por películas o series que estos meses siempre presentan estos cambios bruscos. Además se ve septiembre real

Por debajo del estimado por que como se coloco más de lo estimado en agosto, septiembre era natural que disminuyera.

Considero que como variable independiente del modelo como los meses, no nos parecen lo más adecuado para estimar por este método, ya que se comporta más como una variable cualitativa, y por experiencia es una variable más cuantitativa

Los pronósticos de los años 2007-2009 que se obtienen utilizando la misma técnica estadística de Regresión Lineal Simple son los siguientes:

Pronóstico 2007-2009

Pronóstico	**Regresión Lineal Simple**		
x	**Y=2007**	**Y=2008**	**Y=2009**
ene	13,027	13,939	14,734
feb	53,619	58,813	63,535
mar	105,737	115,721	124,773
abr	79,211	86,247	92,580
may	120,313	131,712	142,051
jun	141,267	154,719	166,925
jul	119,168	129,929	139,634
ago	236,746	258,656	278,474
sep	267,452	288,755	307,670
oct	508,575	552,349	591,606
nov	459,261	498,010	532,673
dic	48,697	52,596	56,061
total	**2,153,071**	**2,341,445**	**2,510,715**

Conclusiones: El Modelo estima los valores de la misma forma, toman en cuenta las hipótesis de Inflación como incremento de precios, y en el total creemos que obtenemos una buena estimación, si la comparamos con la regresión de Costo de Publicidad vs. Facturación, sin embargo en la información mes a mes, presenta muchas deficiencias, lo que lo haría un modelo poco aplicable en la asignación de la cuota de Ventas de la Compañía.

3.3.3 Regresión Lineal Costo de Publicidad y Facturación total año.

En base a la experiencia, se tiene una variable independiente que puede ser más cercana al importe de ventas, que es el Costo de Publicidad, que son un importe de las ventas que se destina para su desplazamiento, o visto de otro modo, la cantidad necesaria para gastar en los medios para desplazar lo facturado, es decir existe una fuerte correlación entre ellas, y es mucho más lineal, como se observa en las siguiente tabla, y en una salida en SPSS.

Como los gastos son registrados hasta el final del año, la estimación, sólo podrá ser por año:

Regresión Lineal con Costo de Publicidad vs Facturación

x	x	Y	% Inc.
1993	$ 22,332.43	$ 171,655.87	
1994	$ 36,992.72	$ 282,602.87	65%
1995	$ 42,935.63	$ 304,683.06	8%
1996	$ 55,457.10	$ 398,455.92	31%
1997	$ 82,031.52	$ 634,918.91	59%
1998	$ 98,308.83	$ 825,292.39	30%
1999	$ 108,402.20	$ 994,378.74	20%
2000	$ 105,823.87	$ 954,219.22	-4%
2001	$ 130,188.69	$ 1,182,297.47	24%
2002	$ 147,587.71	$ 1,316,219.68	11%
2003	$ 151,816.08	$ 1,498,942.54	14%
2004	$ 196,298.00	$ 1,733,767.34	16%
2005	**$ 230,000.00**	**$ 1,803,717.98**	4%
2006	**$ 260,000.00**	**$ 2,000,795.39**	11%
2007	**$ 285,000.00**	**$ 2,168,983.74**	8%
2008	**$ 305,000.00**	**$ 2,375,186.41**	10%
2009		**$ 2,540,148.54**	7%

Negritas: Estimados con Excel

Conclusiones: En base a esta técnica, se obtiene los siguientes coeficientes de autocorrelación muy cercanos a la unidad.

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0.994207669
Coeficiente de determinación R^2	0.988448889
R^2 ajustado	0.987293778
Error típico	57788.69215
Observaciones	12

El análisis de Varianza que determina Excel, es el siguiente:

ANÁLISIS DE VARIANZA

	Grados de libertad	*Suma de cuadrados*	*Promedio de los cuadrados*	*F*	*Valor crítico de F*
Regresión	1	2.8577E+12	2.8577E+12	855.71762	5.08534E-11
Residuos	10	3.3395E+10	3339532941		
Total	11	2.8911E+12			

Las constantes, el error típico y la probabilidad del modelo, son las siguientes:

	Coeficientes	*Error típico*	*Estadístico t*	*Probabilidad*
Intercepción	-88100.18174	36394.887	-2.42067469	0.03602003
Variable X 1	9.637480348	0.32945664	29.2526515	5.0853E-11

3.4 Metodología de Box-Jenkins (modelos ARMA, ARIMA, SARIMA)

3.4.1 Modelos de Decisión con Procesos Estocásticos

Una Serie de Tiempo es una colección de observaciones generadas en forma secuencial a través del tiempo. Los datos se ordenan con respecto al tiempo y las observaciones sucesivas son generalmente dependientes entre sí.

En nuestro estudio, serán las unidades vendidas por mes, el número de tiendas que abren los Clientes por año, ventas por mes, cadena, por año, el índice de inflación, los montos de publicidad por mes y año, etc.

La serie de tiempo observada es una realización de un cierto proceso. Una "realización" es una secuencia de observaciones. Se le llama de esta forma por que si se pudiera realizar de nuevo, se obtendrían resultados diferentes, es decir, una realización distinta.

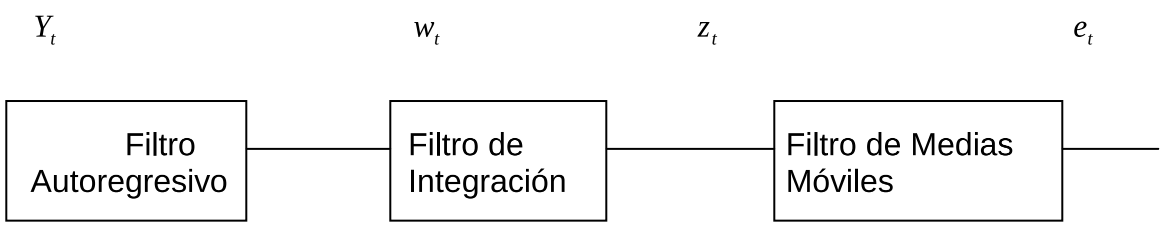
Objetivos del Análisis de Series de Tiempo:

- Descripción: enunciar el comportamiento de un proceso
- Explicación: permiten dar un análisis de de tipo causa-efecto, al relacionar por ejemplo como ayuda las promociones a la facturación.
- Pronóstico: que es el objetivo del presente trabajo, estimar los valores de futuras ventas, apegados a la estadística.
- Control: las variables que se estudian y que tienen correlación ayudarán a desarrollar planes que ayudarán para alcanzar las metas de manera ordenada entre otras áreas como Operaciones y Mercadotecnia.

3.4.2. Metodología de Box-Jenkins

En 19976, George E.P. Box y Gwilym M. Jenkins, publican su libro
"Time Series análisis. Forecastin & Control, en el cual mencionan 4 aplicaciones prácticas del series de tiempo.

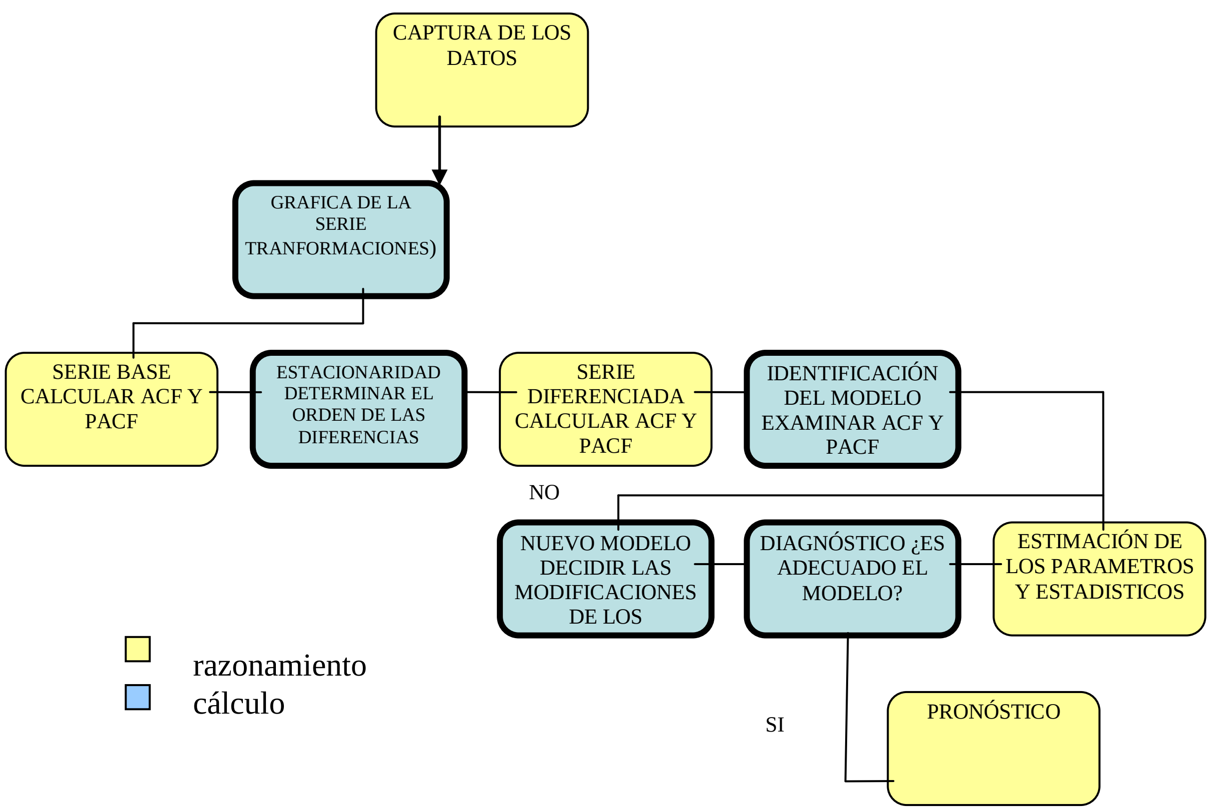
El método de Box-Jenkins consiste en extraer los movimientos predecibles de los datos observados. La Serie de Tiempo se descompone en varios componentes llamados “filtros”, precisamente por que la filosofía consiste en detectar las distintas componentes usando los filtros correspondientes hasta obtener residuales no predecibles, cuyo comportamiento tiene poca influencia en el resultado final. El método utiliza tres filtros lineales principalmente:
El autoregresivo, el de integración y el de medias móviles.



El proceso interactivo para construir modelos lineales consiste en 4 pasos:

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

A continuación se presenta un diagrama funcional del método de Box-Jenkins:



Los procedimientos empleados en la identificación son inexactos y requieren de mucha experiencia del fenómeno en cuestión por lo que se hará uso de la experiencia o juicio ejecutivo para determinar el comportamiento del modelo.

3.4.3. Estacionalidad

Si la Serie muestra una tendencia creciente (Las ventas por ejemplo lo muestran), el problema se complica, si se ve a la serie como una realización de todas las series de tiempo que podrían ser generadas por el mismo mecanismo, se tiene una muestra de tamaño 1.Con esta muestra deberá de estimarse la media, que puede ser en función del tiempo. Si no hay tendencia puede asumirse que la media es constante y el valor observado para cada periodo puede representarse por esta media:

$$E(Y_t) = E(Y_{t+h}) = \mu$$

$$\bar{Y} = E(Y_t)(estimada) = \frac{1}{n}\sum_{t=1}^{n} Y_t$$

Esta estimación es sólo una condición para la estacionalidad. La segunda condición es que la varianza del proceso también sea constante:

$$Var Y_t = Var Y_{t+h} = \sigma_Y^{\ 2}$$

Por sencillez, puede usarse la variable transformada que representa las desviaciones con respecto a la media:

$$Y_t = Y_t - \bar{Y}$$

Finalmente, se impone una condición más sobre la correlación de los datos. La autocorrelación mide la correlación entre un valor observado Y_t y otro Y_{t+h}, que están separados por un intervalo de longitud k:

$$\rho_k = corr(Y_t, Y_{t+k})$$

$$= \frac{E(Y_t, Y_{t+h})}{\left(\operatorname{var} Y_t \operatorname{var} Y_{t+h}\right)^{\frac{1}{2}}}$$

$$= \frac{E(Y_t, Y_{t+h})}{\sigma_y^{\ 2}}$$

$$-1 \leq \rho_k \leq 1$$

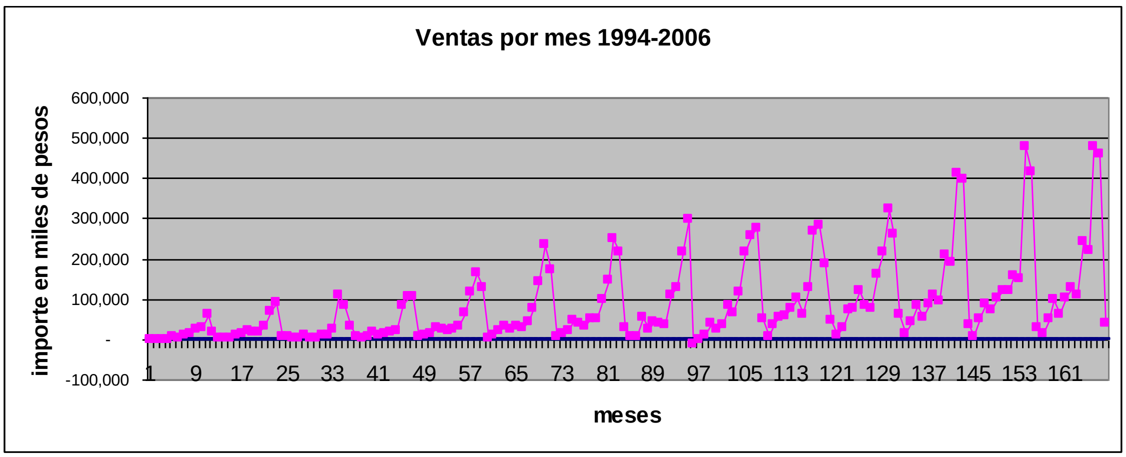
Este valor sólo dependerá de k, en el caso de las series estacionarias, es decir la relación entre dos variables sólo depende del intervalo que las separa.

Si el proceso es estacionario estos valores pueden estimarse tomando la serie de tiempo como una muestra de tamaño *n*, cuyos parámetros son constantes, donde *n* es el número de observaciones.

3.4.4. No Estacionarios

Lo más recomendable cuando se trabaje con series de tiempo, es graficar los datos disponibles contra el tiempo, para visualizar la tendencia, si la varianza es no constante, estacionalidad, discontinuidades y datos discrepante o influyentes.

Tomando los datos de las ventas por mes de 1994 a 2006, observamos lo siguiente:



Gráfica de las ventas de una Empresa Juguetera del mes de enero de 1993 al mes de diciembre del 2006.

En la gráfica se nota lo siguiente:

1.- Varianza no constante
2.- Posible tendencia creciente
3.- Fluctuación estacional (anual)
4.- Variación aleatoria (por películas/series de T.V que incrementan las ventas y son en diferentes meses cada año)

En los datos se observa que no hay estacionalidad, hay un patrón recurrente de incremento de ventas al final de año, principalmente en octubre y noviembre, e incluso un diciembre muy inestable, donde hay un dato negativo en la facturación.

Por lo que será necesario **estabilizar la varianza, tratar de eliminar la tendencia, y finalmente la variación estacional.**

3.4.5 Estabilización de la Varianza

Existen transformaciones de la serie, con el objeto de inducir a un varianza constante. La idea es convertir curvas a rectas y hacer a la varianza constante.

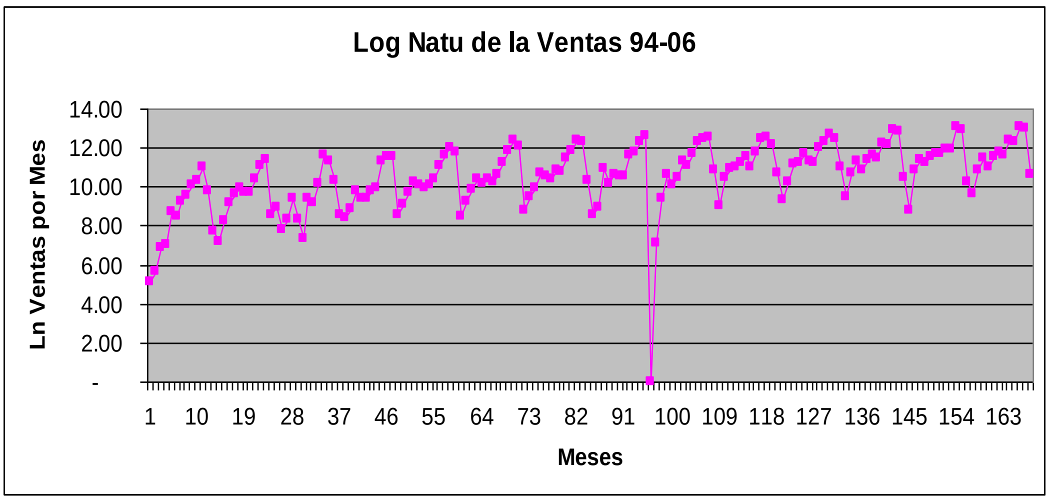
Las transformaciones más comunes son la logarítmica y las raíces cuadradas.

Estas transformaciones son útiles cuando:

- La varianza es proporcional al crecimiento de la serie lo cual ocurre en las ventas, ya que los incrementos de precios se deben en parte a la inflación, la cual aumenta con forme pasa el tiempo.

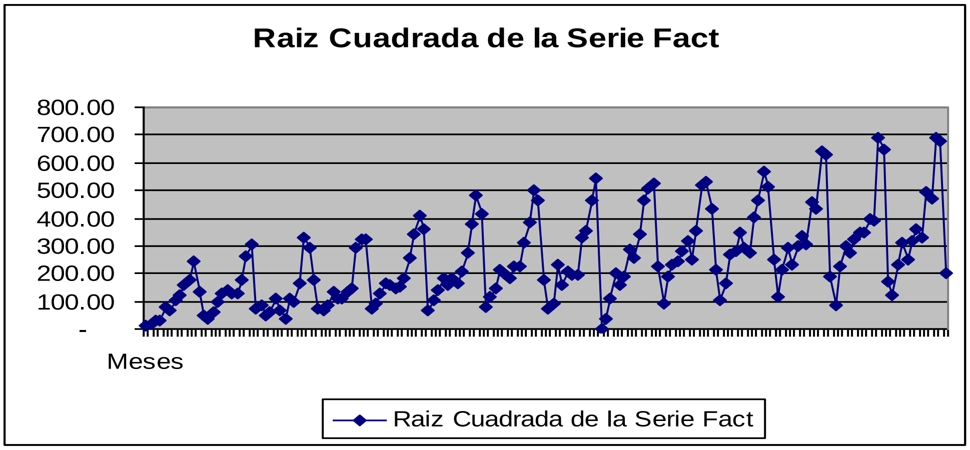
- El nivel medio de la serie aumenta a una tasa constante. Esta hipótesis generalmente ocurre a largo/mediano plazo.

Presentamos a la serie original como la función $Z_t = LN(Y_t)$, la cual tiene la siguiente gráfica:



Esta gráfica nos permite ver que el Logaritmo Natural de la serie original es una buena transformación, existe un valor atípico en diciembre del 2000 el cual es una facturación negativa, el cual se descartará su valor (SPSS lo considerará como un Missing Value), por no ser un dato influyente ni recurrente.

Presentamos también otra transformación que es la raíz cuadrada, SPSS la ocupa frecuentemente y que también se analizara para estabilizar la varianza.



3.4.6 Eliminación de la tendencia.

Se define tendencia como el cambio sistemático en el nivel de una serie de tiempo.

Es importante notar que lo que parece un cambio de nivel en una serie pequeña puede no serlo al tener más datos, convirtiéndolo en un movimiento cíclico. Es por esto la importancia de conocer el proceso que se esta estudiando para establecer el modelo adecuado.

Existen varios métodos para eliminar la tendencia, uno de ellos es usar el método de regresión, este se usa cuando la tendencia es fija y determinística.

El problema radica en que es sumamente difícil saber cuando el cambio de un nivel de la serie se debe a que la tendencia sea determinística o estocástica. Por esto y otras razones se usan diferencias, esto no es lo más efectivo pero para series con un gran número de observaciones, generalmente lo es.

El método consiste en restar los valores de las observaciones uno de otro en un orden preestablecido. Tomando las primeras diferencias de una serie con tendencia lineal, por ejemplo la tendencia desaparece. En general un polinomio de grado 1 se vuelve constante al aplicar una diferencia, uno de grado 2 se vuelve constante al aplicar dos diferencias así sucesivamente. En procesos que se presentan en la Ciencias Sociales **difícilmente aparecen series que requieran de más de dos diferencias para eliminar la tendencia.**

Nótose que las segundas diferencias no equivalen a restar los períodos “t” “t-2”
Es decir la diferencia seria: $\Delta Y_t = Y_t - Y_{t-1}$

$$\Delta Y^2{}_t = \Delta Y_t - \Delta Y_{t-1}$$
$$= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$$
$$= Y_t - 2Y_{t-1} + Y_{t-2}$$

Cada vez que se diferencia se pierde una observación

Una vez eliminada la tendencia, seguir haciendo diferencias producirá series sin tendencia, sin embargo se pierden observaciones y, como se ve, el modelo empieza a complicarse:

$$Y_t = e_t$$
$$ARIMA(0,0,0)$$
$$z_t = y_t - y_{t-1} = e_t - e_{t-1}$$
$$ARIMA(0,1,1)$$

La idea básica de eliminar la tendencia no es olvidar que existe, sino obtener una nueva serie que pueda ser analizada más fácilmente, y después volver a introducir la tendencia, por ejemplo, si se obtuvieron las primeras diferencias:

$$z_t = Y_t - Y_{t-1}$$

Puede reconstruirse la serie original Y_t, en forma recursiva, por la ecuación:

$$Y_t = z_t + Y_{t-1}$$

$$= z_t + z_{t-1} + Y_{t-2}$$

.

.

.

$$= z_t + z_{t-1} + z_{t-2} + z_{t-3} + \ldots$$

3.4.6 Fluctuaciones Estacionales

Al analizar la estacionalidad es imposible determinar sus efectos sin poner atención a la tendencia eliminándola primero, si es que existe en el proceso. La estacionalidad aparece después de eliminar la tendencia.

Existen varios métodos para manejar la estacionalidad. La mayoría son métodos de “autoajuste” basados únicamente en la información de la serie, esto utilizando la intuición, originada por el conocimiento del fenómeno.

Como procedimiento de ajuste pueden hacerse diferencias estacionales, de manera similar a las diferencias consecutivas ordinarias. Sea s (span), la longitud del período de la fluctuación estacional. En las ventas por mes, s tiene longitud 12, es decir se completa un período cada año, por lo tanto:

$$\Delta_s Y_t = Y_t - \Delta Y_{t-s}$$

Se pierden “s” observaciones cada vez que se efectúan estas diferencias

El cálculo de la varianza se define como:

$$Var(X) = E[X - E(X)]^2$$
$$si$$
$$E(X) = 0 \Rightarrow Var(X) = E[X]^2$$

3.4.7 Tipos de Modelos

El proyecto consistirá en analizar una serie de tiempo real sobre la cuál será importante realizar un modelo de pronóstico para ayudarnos en la toma de decisiones en la compañía.

La serie de tiempo debe de ser de tener al menos 100 datos para dar confiabilidad al modelo, y es de interés ya que como lo hemos mencionado este modelo apoyará a tomar mejores decisiones en los pronósticos de ventas.

Los pasos como lo habíamos mencionado que seguirán son los siguientes:

- Identificación del modelo
- Estimación de Parámetros
- Diagnóstico Adecuado
- Pronóstico

3.4.7.1 Identificación del modelo

La metodología que se utilizará para la construcción de un modelo de pronóstico es el de Box-Jenkins con el cual se pretende ajustar una serie de tiempo a un modelo de tipo ARIMA (Autorregresive Integreted Moved Average), Modelos Autorregresivos Integrales y de Médias Móviles.

La forma general del modelo ARIMA *(p,d,q)* es:

$$Y = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varphi_3 Y_{t-3} + \ldots + \varphi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q}$$

Donde: $Y_t = \Delta^d Y_t$ si se hicieran diferencias ordinarias (eliminación de tendencia)

$Y_t = \Delta_s^D Y_t$, si se hicieran diferencias estacionales (eliminación de la variación estacional, el período es de longitud *s*).

$e_t = v.a.i.i.d \sim N(0,\sigma^2)$, variable aleatoria independiente, que se conoce como error aleatorio, o ruido blanco y que se distribuye como una normal estándar.

Los parámetros: $\varphi_1, \varphi_2, \varphi_3, ..., \varphi_p, \theta_1, \theta_2, ..., \theta_q$ indican la contribución de cada variable al modelo.

Para identificar el modelo es necesario determinar las variables *p* y *q.*

El Modelo ARIMA*(p,d,q)* puede expresarse en forma reducida como:

$$\varphi(B)Z_t = \theta(B)e_t$$

Donde: $$\varphi(B) = 1 - B\varphi_1 - B^2\varphi_2 - ... - B^p\varphi_p$$
$$\theta(B) = 1 - B\theta_1 - B^2\theta_2 - ... - B^q\theta_q$$

B es el llamado operador de salto hacia atrás, definido de la siguiente forma:

$$BZ_t = Z_{t-1}$$
$$B^2Z_t = Z_{t-2} = B(Z_{t-1}) = Z_{t-2}$$

.

.

.

$$Z_t = \Delta^d Y_t (diferencias\_de\_la\_variable\_original\_o\_transformada)$$

3.4.7.2 Modelos Autorregresivos AR

El Modelo AR(1)= $AR(1) = \varphi_1 Z_{t-1} + e_t \_con\_e_t \sim N(0,\sigma^2)$

Este modelo debe cumplir con lo siguiente:

1) **Media Constante**

$$E(Z_t) = E(\varphi_1 Z_{t-1}) + E(e_t)$$
$$= \varphi_1 E(Z_{t-1}) + 0$$

Si suponemos que: $E(Z_t) = E(Z_{t-1})$ entonces tenemos:

$$(1-\varphi_1)E(Z_t) = 0 \Rightarrow E(Z_t) = \frac{0}{(1-\varphi_1)} = 0$$

$$para : \varphi_1 \neq 0$$

2) **Varianza Constante**:

$$Var(Z_t) = Var(\varphi_1 Z_{t-1} + e_t) = \varphi_1^2 Var(Z_{t-1}) + \sigma^2$$
$$(1-\varphi_1^2)Var(Z_t) = \sigma^2$$

Si suponemos que $Var(Z_t) = Var(Z_{t-1})$, entonces:

$$Var(Z_t) = \frac{\sigma^2}{(1-\varphi_1^2)} > 0$$

$para\_\left|1-\varphi_1^2\right| > 0; \varphi_1^2 < 1; \left|\varphi_1\right| < 1$ *Condición de estacionalidad*

$$-1 < \varphi_1 < 1$$

3) La función de autocorrelación debe depender sólo de k, es decir, del intervalo entre dos variables.

La función de autocorrelación se define como (ACF)

$$\rho_k = \frac{Cov(Z_{t-1}, Z_{t-k})}{\sqrt{Var(Z_t)Var(Z_{t-k})}}$$

Si se cumple la condición 2.

$$\rho_k = \frac{Cov(Z_{t-1}, Z_{t-k})}{Var(Z_t)} = \frac{\gamma_k}{\gamma_0}$$

Para k=1

$$\begin{aligned}
\gamma_1 &= Cov(Z_t, Z_{t-1}) = E\left[(Z_t - E(Z_t))(Z_{t-1} - E(Z_{t-1}))\right] \\
&= E(Z_t Z_{t-1}) \\
&= E\left[(\varphi_1 Z_{t-1} + e_t)(Z_{t-1})\right] \\
&= \varphi_1 E\left[(Z_{t-1})^2\right] + E[e_t Z_{t-1}] \\
&= \varphi_1 E[(Z_{t-1})^2] \\
&= \varphi_1 Var(Z_{t-1}) = \varphi_1 \gamma_0
\end{aligned}$$

Como la $E[Z_{t-1}]^2 = E[Z_{t-1} - 0]^2 = Var(Z_{t-1}) = \gamma_0$

$$\gamma_1 = \varphi_1 \gamma_0$$

Para k=2

$$
\begin{aligned}
\gamma_1 &= Cov(Z_t, Z_{t-2}) = \\
&= E[(Z_t - E(Z_{t-2})(EZ_t - Z_{t-2})] \\
&= E(Z_t Z_{t-2}] \\
&= E\left[(\varphi_1 Z_{t-1} + e_t)(Z_{t-2})\right] \\
&= \varphi_1 E\left[(Z_{t-1} Z_{t-2})\ \right] + E[e_t Z_{t-2}] \\
&= \varphi\ {}_1 E(\varphi_1 Z_{t-2} + e_t)(Z_{t-2}) \\
&= \varphi^2{}_1 E(Z_{t-2})^2 \\
&= \varphi_1^2 Var(Z_{t-2}) \\
&= \varphi_1^2 \gamma_0
\end{aligned}
$$

Ello implica que:

$$\gamma_2 = \varphi^2{}_1 \gamma_0$$

$$\gamma_3 = \varphi^3{}_1 \gamma_0$$

. Por Inducción Matemática

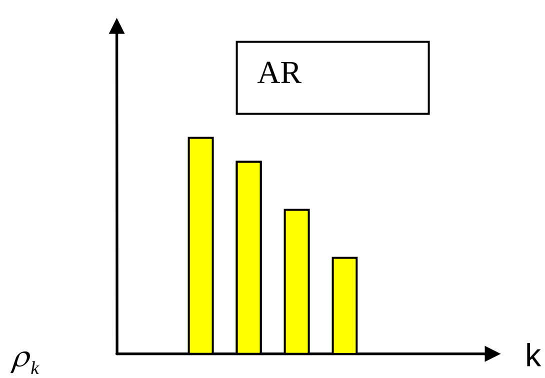
.

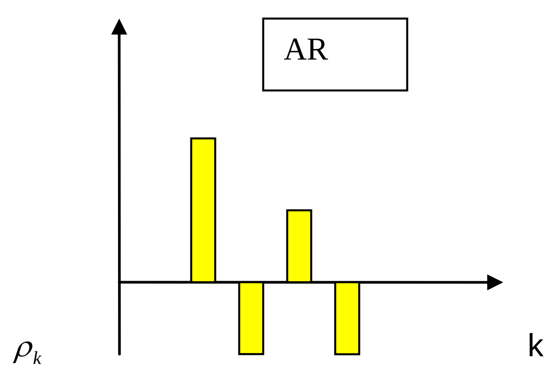
.

$$\gamma_k = \varphi^k{}_1 \gamma_0$$

Entonces $\rho_k = \dfrac{\gamma_k}{\gamma_0} = \varphi_1^k \dfrac{\gamma_0}{\gamma_0} = \rho_k = \varphi_1^k$; esto es la función de ACF con $|\varphi_1| < 1$

Si analizamos la función obtenida observamos que ρ_k decrece conforme k aumenta:





Otra herramienta de identificación es la ***función de memoria*** que indica el efecto de cada variable de error aleatorio en el tiempo.

$$Z_t = \varphi_1 Z_{t-1} + e_t ; \_ Z_{t-1} = \varphi_1 Z_{t-2} + e_{t-1}$$
$$Z_t = \varphi_1^2 Z_{t-2} + \varphi_1 e_{t-1} + e_t ;$$
$$Z_{t-2} = \varphi_1 Z_{t-3} + e_{t-2}$$
$$= \varphi_1^2 (\varphi_1 Z_{t-3} + e_{t-2}) + \varphi_1 e_{t-1} + e_t$$
$$= \varphi_1^2 (\varphi_1 Z_{t-3} + e_{t-2}) + \varphi_1 e_{t-1} + e_t$$
$$= \varphi_1^3 Z_{t-3} + \varphi^2{}_1 e_{t-2} + \varphi_1 e_{t-1} + e_t$$
$$= \varphi^2{}_1 e_{t-2} + \varphi_1 e_{t-1} + e_t + \varphi_1^3 Z_{t-3}$$

El último término es un modelo medias móviles *MA*(∞)

Hasta aquí analizamos un modelo *AR(1)*. En general para un modelo *AR(p)*

$$Z_t = \varphi_1 Z_{t-1} + \varphi_1 Z_{t-2} + ... + \varphi_p Z_{t-p} \quad *$$

Se tienen las condiciones de estacionalidad para los parámetros $\varphi_1, \varphi_1, ...\varphi_p$

* La función de correlación será decreciente infinita y puede obtenerse a partir de las ecuaciones de Yule-Walker:

$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + ... + \varphi_p \rho_{k-p}$, para k>0

Esta expresión se obtiene multiplicando el modelo por Z_{k-p} tomando la esperanza *E[Z]* y dividiendo entre γ_0

Es decir:

$$Z_t Z_{t-k} = \varphi_1 Z_{t-1} Z_{t-k} + \varphi_1 Z_{t-2} Z_{t-k} + ... + \varphi_p Z_{t-p} Z_{t-k}$$
$$= E(\varphi_1 Z_{t-1} Z_{t-k} + \varphi_1 Z_{t-2} Z_{t-k} + ... + \varphi_p Z_{t-p} Z_{t-k})$$
$$= \varphi_1 E(Z_{t-1} Z_{t-k}) + \varphi_1 E(Z_{t-2} Z_{t-k}) + ... + \varphi_p E(Z_{t-p} Z_{t-k})$$

El primer término quedará expresado en función de k-1 (t-1-(t-k)=k-1), el segundo en términos de k-2, y el p-ésimo en función de k-p, lo que implica que:

$$E(Z_t Z_{t-k}) = \rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + ... + \varphi_p \rho_{k-p} \quad \text{ACF}$$

3.4.7.3 Procesos de Medias Móviles

El Proceso *MA(1)* $Z_t = e_t - \theta_1 e_{t-1}$, debe de ser estacionario entonces debe cumplir:

1.- Media Constante

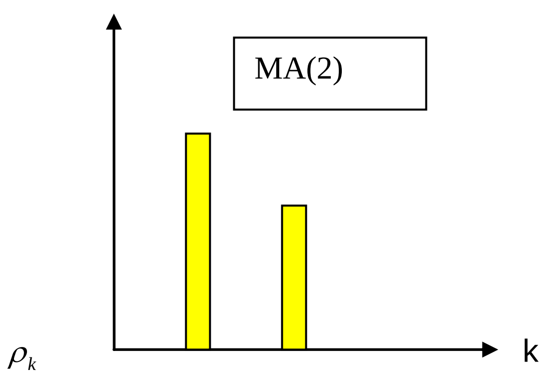
$$E(Z_t) = E(e_t - \theta_1 e_{t-1})$$
$$E(Z_t) = E(e_t) - \theta E(e_{t-1}) = 0$$

2.- Varianza Constante

$$Var(Z_t) = Var(e_t - \theta_1 e_{t-1})$$
$$\gamma_0 = \sigma^2 - \theta_1^2 \sigma^2 = \sigma^2 (1 - \theta_1^2)$$

3.- La ACF depende sólo de ***k***

$$\gamma_1 = Cov(Z_t Z_{t-1}) = E[(e_t - \theta_1 e_{t-1})(e_{t-1} - \theta_1 e_{t-2})]$$
$$= E(e_t e_{t-1} - \theta_1 e_t e_{t-2} - \theta_1 e^2{}_{t-2} + \theta_1 e_{t-1} e_{t-2}) = E(-\theta_1 (e_{t-1})^2)$$
$$= -\theta_1 Var(e_{t-1}) = -\theta_1 \sigma^2 \Rightarrow$$
$$\gamma_1 = \frac{-\theta_1 \sigma^2}{\sigma^2 (1 - \theta^2)} = \frac{-\theta_1}{(1 - \theta_1^2)}$$
$$\gamma_2 = Cov(Z_{t-1} Z_{t-2}) = E[(e_t - \theta_1 e_{t-1})(e_{t-2} - \theta_1 e_{t-3})] = 0$$
$$\gamma_3 = Cov(Z_{t-2} Z_{t-3}) = 0$$
$$\gamma_k = 0$$
$$\therefore \rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} \dfrac{-\theta_1}{(1-\theta_1}; k = 1 \\ 0; \_\_\_\_\_\ k > 1 \end{cases}$$



En forma inversa el modelo $MA(1)$ equivale a un $AR(\infty)$

$$Z_t = e_t - \theta_1 e_{t-1} \Rightarrow e_t = Z_t + \theta_1 e_{t-1}$$
$$e_{t-1} = Z_{t-1} + \theta_1 e_{t-2}$$

$$Z_t = e_t - \theta_1(Z_{t-1} + \theta_1 e_{t-2}) = e_t - \theta_1 Z_{t-1} - \theta_1^2 e_{t-2}$$
$$= e_t - \theta_1 Z_{t-1} - \theta_1^2(Z_{t-2} + \theta_1 e_{t-3})$$
$$= e_t - \theta_1 Z_{t-1} - \theta_1^2 Z_{t-2} - \theta_1^3 e_{t-3}$$

.

.

.

$$Z_t = e_t - \theta_1 Z_{t-1} - \theta_1^2 Z_{t-2} - \theta_1^3 e_{t-3} - ....AR(\infty)$$

A esta ecuación se le llama forma invertida del proceso de medias móviles y representa un proceso con un número infinito de términos autoregresivos.

Para que esta ecuación tenga sentido se requiere que el parámetro cumpla con:

$-1 < \theta_1 < 1\_ó\_|\theta_1| < 1$ Condicional de Invertibilidad

3.4.7.4 Modelos de Medias Móviles de Orden Mayor

El proceso de MA(q) será:

$$Z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - .... - \theta_q e_{t-q}$$

La función de memoria será la gráfica de los valores: 1, $-\theta_1, -\theta_2, -\theta_3 .... -\theta_q$

El efecto del choque e_t, persistirá por “q” períodos. La condición de invertibilidad para MA(q) no es sencilla de derivar, ya que involucra funciones no lineales de los q parámetros.

Si analizamos la función ACF de un MA(q) observaremos que este se trunca en q.

Por ejemplo el proceso de un MA(2) será:

Conclusiones en este momento:

- Un proceso finito MA( ) puede representarse como proceso infinito AR( )
- Un proceso finito AR( ) puede representarse como proceso infinito MA( )
- Un proceso AR tiene condición de estacionalidad no de invertibilidad
- (ya es invertible por si sólo)

- Un proceso MA no tiene condición de estacionalidad, pero si es invertible (siempre es estacionario)
- La ACF de un proceso MA( ) se trunca, en tanto que la de un proceso AR es monótona decreciente infinita.
- Un proceso ARMA( ) tiene condición de *estacionalidad* y de *invertibilidad* , ya que incluye una parte de MA( ) y una de AR( ).

3.4.7.5 Modelos con término constante.

Cuando la media de una serie de tiempo no es cero, se requiere de un término constante que ajuste a la media. El modelo general se expresa:

$$\varphi(B)Z_t = \theta(B)e_t + \delta$$
$$donde: \varphi(B) = 1 - B\varphi_1 - B^2\varphi_2 - ... - B^p\varphi_p$$
$$\theta(B) = 1 - B\theta_1 - B^2\theta_2 - ... - B^q\theta_q$$

Las cuales se obtiene aplicando el operador de salta hacia atrás *B definido como:*

$$BZ_t = Z_{t-1}$$
$$B^2Z_t = B(Z_{t-1}) = Z_{t-2}$$
$$B^3Z_t = B^2(Z_{t-1}) = B(Z_{t-2}) = Z_{t-3}$$

$$\varphi(B)Z_t = Z_t - Z_{t-1}\varphi_1 - Z_{t-2}\varphi_2 - ... - Z_{t-p}\varphi_p$$
$$[e_t]\theta(B) = e_t - e_tB\theta_1 - e_tB^2\theta_2 - ... - e_tB^q\theta_q$$

Aplicando el operador:

$$-e_t\theta(B) = e_t - e_{t-1}\theta_1 - e_{t-2}\theta_2 - ... - e_{t-q}\theta_q$$

De ahí obtenemos la forma general:

$$\varphi(B)Z_t = \theta(B)e_t + \delta$$

En este caso, con un término constante, la media de la serie será:

$$E(Z_t - Z_{t-1}\varphi_1 - Z_{t-2}\varphi_2 - ... - Z_{t-p}\varphi_p) =$$
$$= E(e_t - e_tB\theta_1 - e_tB^2\theta_2 - ... - e_tB^q\theta_q) + E(\delta)$$

Si suponemos que la media es constante:

$$E(Z_t) = E(Z_{t-1})$$
$$E(Z_t) = \varphi_1 E(Z_t) - \varphi_2 E(Z_t) - ... - \varphi_p E(Z_t) =$$
$$= E(e_t) - \theta_1 E(e_t) - \theta_2 E(e_{t-1}) - ... - \theta_q E(e_{t-q}) + E(\delta)$$
$$E(Z_t) = (1 - \varphi_1 - \varphi_2 - ... - \varphi_p) = 0 + E(\delta)$$
$$E(Z_t) = \frac{\delta}{(1 - \varphi_1 - \varphi_2 - ... - \varphi_p)}$$

Esta sería la fórmula de la media de la serie, con término constante.

3.4.7.6 Modelos Mezclados ARMA(1,1)

El Proceso ARMA(1,1)

$$Z_t = \varphi_1 Z_{t-1} - \theta_1 e_{t-1} + e_t$$

1.- Media Constante

$$E(Z_t) = E(\varphi_1 Z_{t-1}) - \theta_1 e_{t-1} + e_t$$
$$\varphi_1 E(Z_{t-1}) - \theta_1 E(e_{t-1}) + E(e_t)$$
$$E(Z_t) = \varphi_1 E(Z_{t-1})$$
$$E(Z_t) = \varphi_1 E(Z_t)$$
$$E(Z_t)(1 - \varphi_1) = 0$$
$$E(Z_t) = \frac{0}{1 - \varphi_1} = 0$$

Siempre que $\varphi_1 \neq 0$

2.- Varianza Constante

$$Var(Z_t) = (E(Z_t - E(Z_t))^2$$
$$= E(Z_t)^2 = E(\varphi_1 Z_{t-1} - \theta_1 e_{t-1} + e_t)^2$$
$$= E[(\varphi_1 Z_{t-1})^2 - (\varphi_1 Z_{t-1})(\theta_1 e_{t-1}) + (\varphi_1 Z_{t-1})(e_t) - (\varphi_1 Z_{t-1})(\theta_1 e_{t-1}) + (\theta_1 e_{t-1})^2$$
$$- (\theta_1 e_{t-1})(e_t) + (\varphi_1 Z_{t-1} e_t) - (\theta_1 e_{t-1})(e_t) + (e_t)^2]$$
$$= E[(\varphi_1 Z_{t-1})^2 - 2(\varphi_1 Z_{t-1})(\theta_1 e_{t-1}) + 2(\varphi_1 Z_{t-1})(e_t) - 2(\theta_1 e_{t-1})(e_t) + (\theta_1 e_{t-1})^2 + (e_t)^2]$$
$$\Rightarrow Var(Z_t) = \varphi_1^2 Var(Z_{t-1}) - 2E[(\varphi_1 Z_{t-1})(\theta_1 e_{t-1})] + \theta_1 \sigma^2 + \sigma^2$$
$$E[-2(\varphi_1 Z_{t-1})(\theta_1 e_{t-1})\ ] = -2\varphi_1\theta_1 E(Z_{t-1}e_{t-1})$$
$$Z_t = \varphi_1 Z_{t-1} - \theta_1 e_{t-1} + e_t$$

Ahora $Z_{t-1} = \varphi_1 Z_{t-2} - \theta_1 e_{t-2} + e_{t-1} \Rightarrow$

$$E[Z_{t-1}\theta_1 e_{t-1}] = E[\varphi_1 Z_{t-2} - \theta_1 e_{t-2} + e_{t-1}(e_{t-1})] = E[\varphi_1 Z_{t-2} e_{t-1}] - E[\theta_1 e_{t-2}(e_t)] + E[e_t]^2$$
$$= \varphi_1(0) - \theta_1(0) + \sigma^2 \Rightarrow E[-2(\varphi_1 Z_{t-1})(\theta_1 e_{t-1}) = 2\varphi_1\theta_1\sigma^2$$

Con lo que obtenemos:

$$Var(Z_t) = \varphi_1^2 Var(Z_{t-1}) - 2\varphi_1\theta_1\sigma^2 + \theta_1\sigma^2 + \sigma^2$$

Si suponemos la varianza constante, $Var(Z_t) = Var(Z_{t-1})$

$$Var(Z_t) = \varphi_1^2 Var(Z_t) - 2\varphi_1\theta_1\sigma^2 + \theta_1\sigma^2 + \sigma^2 \Rightarrow$$
$$Var(Z_t)(1 - \varphi_1^2) = -2\varphi_1\theta_1\sigma^2 + \theta_1\sigma^2 + \sigma^2 \Rightarrow$$
$$Var(Z_t) = \frac{-2\varphi_1\theta_1\sigma^2 + \theta_1\sigma^2 + \sigma^2}{(1 - \varphi_1^2)} = \frac{\sigma^2[-2\varphi_1\theta_1 + \theta_1 + 1]}{1 - \varphi_1^2}$$

3.- Las ACF depende sólo de K

$$\rho_k = \frac{Cov(Z_t, Z_{t-1})}{Var(Z_t)} = \begin{cases} \dfrac{(1 - \varphi_1\theta_1)(\varphi_1 - \theta_1)}{(1 + \theta^2{}_1 + 2\varphi_1\theta)}; k = 1 \\ \varphi_1\rho_k \ \_\_\_\_\_\_ k > 1 \end{cases}$$

3.4. Graficas de Estacionalidad y ciclaje de las series históricas (ACF y PACF) y periodograma integral.

Función de autocorrelación parcial (PACF)

Al construir un modelo es conveniente investigar si la inclusión de una nueva variable es no significativa, esto es, si esta variable contribuye a mejorar el modelo.

Modelos Autoregresivos

En estos modelos si se agrega una nueva Z_{t-k} se deseará comprobar si el modelo AR(k) es más adecuado que el modelo AR(k-1), lo cual se determina con base en el valor del coeficiente φ_k.

Este coeficiente mide el exceso de la correlación no tomado en cuenta por el modelo AR(k), esto es mide el efecto parcial de Z_{t-k} para explicar el comportamiento de Z_t en un modelo que ya incluye a:

$Z_{t-1}, Z_{t-2}, ..Z_{t-(k-1)}$

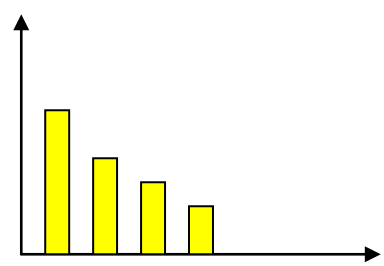
Entonces el coeficiente autorregresivo de mayor orden φ_k, se define como la autocorrelación parcial del intervalo k y se denota como: ρ_{kk}

Estamos analizando: $AR(1) = Z_t = \varphi_1 Z_{t-1} + e_t \Rightarrow AR(K-1)$

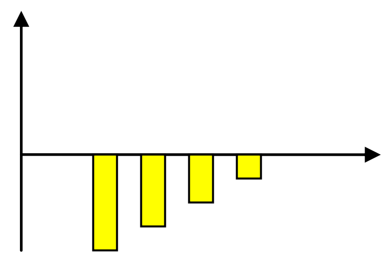
Si le sumamos una nueva variable: $AR(2) = Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + e_t \Rightarrow AR(K) \Rightarrow \varphi_k = \varphi_2$

Se llama autocorrelación parcial por que nada más estamos analizando que significativa va a ser nuestro parámetro φ_k

Las gráficas en el caso positivo serían:



Las gráficas en el caso negativo serían:



Suponga un modelo $AR(1) = Z_t = \varphi_1 Z_{t-1} + e_t \Rightarrow \rho_{kk} = \varphi_1 = \rho_{11}$

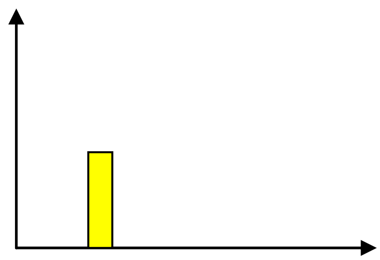
Para calcularse ρ_{22} deberá probarse un modelo AR(2)

$$Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + e_t \Rightarrow \rho_{22} = \varphi_2$$

Si el modelo adecuado es el AR(1) el valor de ρ_{22} deberá de cero, (porque su contribución es nula). En forma similar: ρ_{33}= ρ_{44}=….=0

En general la PACF para un modelo AR(1)

$$\rho_{kk} \begin{cases} \varphi_{1,} k = 1 \\ 0, k > 1 \end{cases}$$



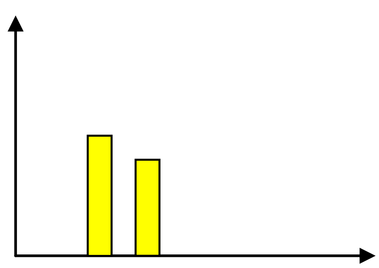
Se trunca en 1

La ACF de un AR(1) es decreciente infinita (grado de correlación entre dos variables)

La PACF de un AR(1) se trunca en k=1 (contribución al modelo del último coeficiente o parámetro)
Si el modelo adecuado fuera un AR(2) entonces ρ_{22} sería distinto de cero y ρ_{33}= ρ_{44}=….=0

En general para un modelo AR(2) la PACF será:

$$\rho_{kk} \begin{cases} \varphi_{1,} k = 1 \\ \varphi_2, k = 2 \\ 0, k > 2 \end{cases}$$



Se trunca en 2

En general para un modelo AR(K) la P.A.C.F será:

$$\rho_{kk} \begin{cases} \varphi_{i,} i \le k \\ 0, i > k \end{cases}$$ La gráfica de PACF se trunca en i=k

3.6 ACF y PACF Muestrales

El estimador de la autocorrelación con las mejores propiedades:

$$\hat{\rho}_k = r_k = \frac{C_k}{C_0}$$

$$donde:$$

$$C_k = \frac{1}{n}\sum_{t=1}^{n-k} Z_t Z_{t-k}$$

Ahora bien para obtener el estimador de ρ_{kk} es conveniente utilizar las ecuaciones de Yule-Walker

Para un proceso *AR(p)* definimos la ACF como:

$$\rho_k = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2} + \ldots + \varphi_p\rho_{k-p}$$

Esta expresión la obtenemos de la matriz de covarianzas y varianzas para un proceso estacionario.

Covarianza: Relación de una variable con respecto a ella misma en diferentes unidades de tiempo.

Definimos a la matriz Γ_N, como la matriz simétrica, cuadrada, cuya diagonal es la varianza:

$$\begin{array}{cc} & \begin{array}{cccc} Y_1 & Y_2 & Y_3 & Y_N \end{array} \\ \Gamma_N = & \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots\cdots\gamma_{N-1} \\ \gamma_1 & \gamma_0 & \gamma_2 & \cdots\cdots\gamma_{N-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots\cdots\gamma_{N-3} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \gamma_{N-1} & \gamma_{N-2} & \gamma_{N-3} & \gamma_0 \end{pmatrix}, \\ & \begin{array}{cccc} Y_1 & Y_2 & Y_3 & Y_N \end{array} \end{array}$$

$$P_N = \frac{1}{\gamma_0}\Gamma_N = \begin{pmatrix} 1 & \rho_1 & \rho_2 & .....\rho_{N-1} \\ \rho_1 & 1 & \rho_2 & ......\rho_{N-2} \\ \rho_2 & \rho_1 & 1 & ....\rho_{N-3} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \rho_{N-1} & \rho_{N-2} & \rho_{N-3} & 1 \end{pmatrix}$$

Donde $\gamma_0 = VAR(Y_t) > 0$

A partir de esto se genera la siguiente ecuación:

$$P_N \Phi = \rho_k$$

Es decir: $\begin{pmatrix} 1 & \rho_1 & \rho_2 & .....\rho_{N-1} \\ \rho_1 & 1 & \rho_2 & .....\rho_{N-2} \\ \rho_2 & \rho_1 & 1 & ......\rho_{N-3} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ \rho_{N-1} & \rho_{N-2} & \rho_{N-3} & 1 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_{.3} \\ . \\ . \\ . \\ \varphi_k \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_{.3} \\ . \\ . \\ . \\ \rho_k \end{pmatrix}$

Para un AR(1)

$\rho_{11} = \varphi_1 = \rho_1$

Para un AR(2)

$\rho_1 = \varphi_1 + \rho_1\varphi_2$

$\rho_2 = \varphi_1\rho_1 + \varphi_2$

Utilizando determinantes para resolver la ecuación obtenemos por ejemplo para φ_2 , el siguiente resultado:

$\varphi_2 = \dfrac{\rho_2 - \rho_1^{\ 2}}{1 - \rho_1^2}$ Sustituyendo los valores de ρ_k por su estimador r_k

$$\hat{\rho}_{11} = \varphi_1 = r_1$$

$$\hat{\rho}_{22} = \varphi_2 = \frac{r_2 - r_1^2}{1 - r_1^2} = r_{22}$$

En general se obtiene:

$$\hat{\rho}_{kk} = r_{kk} = \begin{cases} r_1 \_ k = 1 \\ \dfrac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \_ k = 2,3 \end{cases}$$

Varianza de los Estimadores.

Algunos expertos en la materia definen a la varianza como:

$$VAR(\rho_n) \approx \frac{1}{n}\left[1 + 2(\rho_1^2 + \rho_2^2 + \ldots + \rho_q^2\right] \text{ para } k>q$$

Donde *q* es el orden de la parte *MA( )*

Utilizando su estimador:

$$V\hat{A}R(r_n) \approx \frac{1}{n}\left[1 + 2(r_1^2 + r_2^2 + \ldots + r_q^2\right]$$

$$k>q$$

$$V\hat{A}R(r_n) \approx \frac{1}{n}\left[1 + 2\sum_{i=1}^{q} r_i^2\right]$$

Otros la define como:

$$V\hat{A}R(r_n) \approx \frac{1}{n} \_ para \_ k > p$$ donde **p** es el orden de la parte *AR( )*

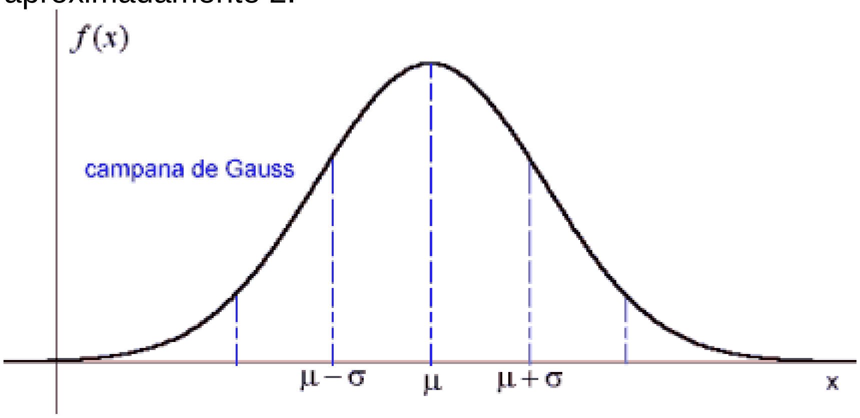
3.7 Estadísticos e Intervalos de Confianza

No interesa probar la hipótesis de que los coeficientes de Autocorrelación, son significativos, es decir diferente de cero.

$H_0 = \hat{\rho}_k = 0$
$H_a = \hat{\rho}_k \neq 0$ Se construye el estadístico $t = \frac{r_k}{Sr_k}$ Sr_k = desviación estándar de r_k

Que se distribuye asintóticamente como una t-student, donde los grados de libertad corresponden al número de variables independientes menos el número de parámetros estimados.

Como se supone que se tendrán 100 o más datos (observaciones de las ventas), puede emplearse un nivel de confianza del 95%, un valor de t-student de aproximadamente 2.



Entonces la regla de aceptación sería:

$$Aceptar\_H_0\_si\_|t| < 2$$

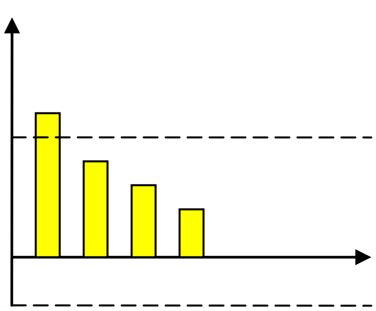
$$\hat{\rho}_k = 0$$

Esto quiere decir que $\hat{\rho}_k$ es estadísticamente no significativo

Por otro lado, un intervalo de confianza del 95% para $\hat{\rho}_k$ será:

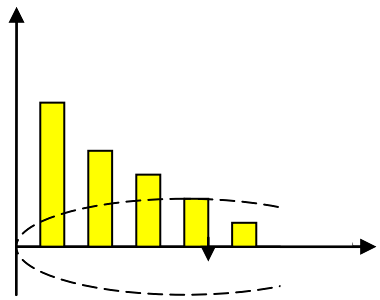
$$r_k - 2Sr_k \leq \hat{\rho}_k \leq 2Sr_k + r_k$$

ACF



No significativas

PACF



No significativas

Para la PACF, nos interesa probar la hipótesis:

$$H_0 = \rho_{kk} = 0$$

$$H_a = \rho_{kk} \neq 0$$

Se construye el estadístico $t = \dfrac{r_{kk}}{\sqrt{\dfrac{1}{n}}}$; $\sqrt{\dfrac{1}{n}}$ =desviación estándar r_{kk}

Que se distribuye como una t-student.

Regla de decisión:

Aceptar H_0 si $|t| < 2$ lo cual indica que $\hat{\rho}_{kk}$ es no significativa.

Para un nivel de confianza del 95% el intervalo para $\hat{\rho}_{kk}$ será:

$$r_{kk} - 2Sr_{kk} \leq \hat{\rho}_{kk} \leq 2Sr_{kk} + r_{kk}$$

Además de las pruebas vistas anteriormente existe la prueba PORTMANTEAU que se utiliza para ver si un grupo de autocorrelaciones puede clasificarse como no significativas.

Las pruebas de hipótesis quedarían de la siguiente manera:

$H_0 = \rho_1 = \rho_2 = ... = \rho_k = 0$ ruido blanco

$H_a = \rho_i \neq 0$
$i = 1,2,..k$ al menos una diferente de cero

Se utiliza el estadístico de Box_Pierce

$Q(k) = n(n+2)\sum_{k=1}^{k} \frac{1}{n-k} r_k^2$, que se distribuye como una χ^2 con k grados de libertad.

La regla de decisión será: Aceptar H_0 si $Q(k) < \chi^2(k)$

Nota: Al efectuarse estas pruebas, debe tomarse en cuenta que involucra muchas aproximaciones por lo que la decisión no debe ser muy estricta. Por ejemplo, si la autocorrelación ρ_{12} parece significativa pero existe razón para pensar que el período 12 sea influyente, es mejor no tomarla en cuenta.

Capitulo 4

Modelo matemático

Capítulo 4

4.1 Identificación del Modelo (Teoría)

4.1.1 Estimación de los Modelos

Una vez que se ha hecho la especificación del modelo es decir, una vez que se ha elegido los valores de *p,q,d* para el modelo

$$\Phi(B)W_t = \Theta(B)e_t \ldots\ldots\ldots(1)$$

Deben obtenerse los estimadores de los parámetros autorregresivos AR $\varphi_1, \varphi_2 \ldots \varphi_p$, y los parámetros $\theta_1, \theta_2 \ldots \theta_q$.
Como en el caso de regresión lineal, se eligen los parámetros que minimizen la suma de cuadrados de las diferencias entre la serie real observada W_t y la serie estimada $\hat{W}_t$.

En máxima verosimilitud se deseará encontrar los valores de los parámetros que maximicen la probabilidad de ocurrencia de W_t

De la ecuación (1) se obtiene: $e_t = \Theta^{-1}(B)\Phi(B)W_t \ldots\ldots\ldots(2)$

Se desea encontrar un vector de parámetros AR: $\Phi = (\varphi_1, \varphi_2 \ldots \varphi_p)$ y un vector de parámetros MA: $\Theta = (\theta_1, \theta_2 \ldots \theta_q)$ que minimizen la suma de los errores al cuadrado:

$$S(\Phi, \Theta) = \sum_t e_t^2 \ldots\ldots\ldots(3)$$

Los valores que minimizan la ecuación (3) se denotan como:

$$\hat{\Phi} = (\hat{\varphi}_1, \hat{\varphi}_2 \ldots \hat{\varphi}_p)$$
$$\hat{\Theta} = (\hat{\theta}_1, \hat{\theta}_2 \ldots \hat{\theta}_q)$$

Y los residuales asociados se denotan por e_t. De modo que la ecuación (2) quedará expresada como:

$$\hat{e}_t = \hat{\Theta}^{-1}(B)\hat{\Phi}(B)W_t.$$

$$\Rightarrow S(\hat{\Phi}, \hat{\Theta}) = \sum_t \hat{e}_t^{\,2}$$

La estimación puede resultar muy compleja si existen términos de medias móviles, ya que la ecuación (3) será "no lineal". Además por ejemplo e_1 se requerirá de valores no observados $e_0, e_{-1}, \ldots, e_{q+1}$ y $W_0, W_{-1}, \ldots, W_{p+1}$ que deberán estimarse utilizando un procedimiento que permita inicializar una serie.

Después de estimar los parámetros del modelo se usará un proceso de "diagnóstico" si la especificación inicial del modelo fue correcta. Es decir esperaríamos que los errores residuales $\hat{e}_t\_t = 1,2,..n$ se parezcan lo más posible a los errores e_t, los cuales están supuestamente intercorrelacionados. El chequeo de diagnóstico se usa para probar si los residuales en realizad no están correlacionados, ya que si lo están se deberá de replantear los valores para *p,q,d* para determinar otro modelo y volver a realizar el diagnóstico.

4.1.2 Procedimiento de Estimación.

Si se asume que el total de *T+d* observaciones están disponibles para una serie de tiempo homogénea no estacionaria de orden d (es decir requiere "d" diferencias ordinarias para ser estacionaria), Y_t, y se denotan estás observaciones como $Y_{-d+1}, .. Y_0, Y_1, \ldots, Y_T$; después de diferenciar esta serie "d" veces se obtiene una serie estacionaria z_t con *T* observaciones $z_1, \ldots, z_T$. El problema es estimar los vectores de parámetros Φ y Θ para el modelo ARMA(p,q) especificado para la serie z. Para hacer esto se utiliza el hecho de que los términos de los errores e_t son distribuidos normalmente e independientes, cono media cero y varianza σ^2. Entonces la función de verosimilitud condicional asociada con los parámetros (Φ, Θ, σ^2) esta dada por:

$$F(\Phi, \Theta, \sigma^2) = [\frac{1}{\sqrt{(2\pi\sigma^2)}}]^T e^{\frac{(-S(\Phi,\Theta)}{2\sigma^2}}$$

$$= [2\pi\sigma^2]^{-\frac{T}{2}} e^{\frac{(-S(\Phi,\Theta)}{2\sigma^2}}$$

Donde T es el número de observaciones disponibles

Obteniendo el Logaritmo a esta función, obtenemos:

$$L(\Phi,\Theta,\sigma^2) = -\frac{T}{2}\ln(2\pi\sigma^2) - \frac{S(\Phi,\Theta)}{2\sigma^2}.....(4)$$

El segundo término de la igualdad se debe de minimizar $\frac{S(\Phi,\Theta)}{2\sigma^2}$, para que la maximizar la función logarímica $L(\Phi,\Theta,\sigma^2)$, que se llama función de verosimilitud condicional por que la suma de cuadrados $S(\Phi,\Theta)$ depende de valores no observables: $e_0, e_{-1},...,e_{q+1}$ y $W_0, W_{-1},...,W_{p+1}$.

Esto se puede observar escribiendo la ecuación para el primer error observable e_1 en la forma expandida del modelo ARMA.

$$e_1 = W_1 - \varphi_1 W_0 - \varphi_2 W_{-1} - ....- \varphi_p W_{-p+1}$$

Los estimadores de máxima verosimilitud de Φ y Θ estarán dados por la minimización de la suma de residuales al cuadrado $S(\Phi,\Theta)$. Entonces, bajo la suposición de normalidad, el estimador de máxima verosimilitud es igual al de mínimos cuadrados.

4.1.3 Inicialización de la serie

Ya que la función suma de cuadrado y la función de verosimilitud son ambas condicionales con respecto a valores no observables de $z_t\_y\_e_t$, los estimadores de mínimos cuadrados que se obtengan dependen de la elección de los valores $W_0, W_{-1},...,W_{p+1}$, etc.

Una posible eyección es tomar $W_0, W_{-1},...,W_{p+1}, e_0, e_{-1},...,e_{q+1}$ iguales a sus valores esperados *incondicionales* de $e_0, e_{-1},...,e_{q+1}$ son todos cero, y si $\delta = 0$ (suponiendo que se manejan las desviaciones con respecto a la media), los valores esperados incondicionales de $W_0, W_{-1},...,W_{p+1}$ serán también ceros.

En caso de que $\delta \neq 0$, se usa la siguiente expresión:

$$E(W_t) = \frac{\delta}{1 - \varphi_1 - \varphi_2 - ... - \varphi_p}$$

Se usa como la media muestral, $\bar{W}$ como el valor esperado incondicional.

Generalmente esta operación es suficiente si los valores de $\varphi_1,\varphi_2,..\varphi_p,\theta_1,\theta_2,..\theta_q$ no son cercanos a 1 ó -1, y si el número de observaciones T es muy “grande” en comparación con p y q.

También existe un procedimiento para encontrar los valores esperados **condicionales** para $W_0,W_{-1},...,W_{p+1}$. Inicialmente se utilizan los valores esperados incondicionales para estimar el modelo; el modelo se estima minimizando $S(\Phi,\Theta)$, condicionando a los valores a cero. Una vez estimada se pronostica hacia atrás con el modelo para generar nuevos valores de $W_0,W_{-1},...,W_{p+1}$. Esto es válido ya que la serie es estacionaria.

Operador salto hacia delante F

$$\Phi(F)W_t = \Theta(F)e_t$$
$$F(W_t) = W_{t+1}$$

$$F^2(W_t) = W_{t+2}$$
$$\Rightarrow W_t = \Phi^{-1}(F)\Theta(F)e_t$$

Por ejemplo, para un A(1)

$$W_t = \varphi_1 W_{t-1} + e_t$$

$$W_0 = \varphi_1 W_1 + e_0$$
$$W_1 = \varphi_1 W_2 + e_1$$

Esta última ecuación se utiliza para encontrar los valores de $\hat{W}_0,\hat{W}_{-1},...,\hat{W}_{-p+1}$ en función de los valores observados $W_0,W_{-1},...,W_{p+1}$ y de los valores estimados de $e_1,e_2,...,e_T$, que serán los residuales obtenidos de restar los valores observados W_t menos los estimados inicialmente $\hat{W}_t$, con estos valores pronosticados se obtienen nuevos estimadores de mínimos cuadrados para Φ y Θ, minimizando la suma de los errores al cuadrado, condicionando a los nuevos valores $\hat{W}_0,\hat{W}_{-1},...,\hat{W}_{-p+1}$.

Este proceso se repite hasta que $\hat{\Phi}$ y $\hat{\Theta}$ dejan de cambiar en forma significativa, es decir se obtenga convergencia.

No hay garantía que este proceso converja, si no la hay, deberá de usarse los estimadores incondicionales. Por lo general en los ARMA en aplicaciones económicas y Financieras son de orden bajo (p y q≤ 3). En esos casos la inicialización no será problema.

4.1.4 Estimación No Lineal de los Parámetros del Modelo

El problema es encontrar los valores de $\hat{\Phi}$ y $\hat{\Theta}$ que minimicen la suma de cuadrados: $S(\Phi,\Theta) = \sum_t e_t^2 = \sum_t^T [e_t | \Phi,\Theta,W]^2 \ldots.(4)$

Donde W es el vector de las observaciones W_t

Suponiendo que se toman los valores incondicionales iguales a cero:

$$S(\Phi,\Theta) = \sum_t^T [e_t | \Phi,\Theta,W]^2 \ldots.(5)$$

Supóngase que el modelo es puramente autoregresivo de la forma:

$$\Phi(B)W_t = e_t \ldots..(6)\_ó$$

$$W_t - \varphi_1 W_{t-1} - \varphi_2 W_{t-2} - \ldots - \varphi_p W_{t-p} = e_t$$

$$W_t = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} + \ldots + \varphi_p W_{t-p} + e_t \ldots..(7)$$

Obsérvese que la ecuación (7) sigue la forma de un modelo de regresión múltiple.

$$Y = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + e_t \ldots..(8)$$

Este modelo se estima utilizando la siguiente forma matricial:

$$\hat{\varphi}_1 = (X^T X)^{-1} X^T Y$$

$$\Phi = (\varphi_1, \varphi_2, \ldots, \varphi_p)$$

Donde: $X = \begin{pmatrix} W_p & W_{p-1} & \ldots & W_1 \\ W_{p+1} & W_p & \ldots & W_2 \\ . & & & \\ . & & & \\ . & & & \\ W_{T-1} & W_{T-2} & \ldots & W_{T-p} \end{pmatrix}$ $Y = \begin{pmatrix} W_{p+1} \\ W_{p+2} \\ . \\ . \\ . \\ W_T \end{pmatrix}$ $e = \begin{pmatrix} e_{p+1} \\ e_{p+2} \\ . \\ . \\ . \\ e_T \end{pmatrix}$

$$Y = \Phi X + = e$$

El modelo original: $W_t = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} + \ldots + \varphi_p W_{t-p} + e_t \_\_ y$

$$W_{p+1} = \varphi_1 W_p + \varphi_2 W_{p-1} + \varphi_3 W_{p-2} + \ldots + \varphi_p W_1 + e_{p+1}$$

Será para este ejemplo para t=1 hasta T

Se desea minimizar

$$\sum_t e_t^2 = S(\Phi,\Theta) = \sum_t^T [e_t \mid \Phi,\Theta,W]^2 =$$

$$\sum (\Phi B W_t)^2 = \sum_t^T [W_t - \varphi_1 W_{t-1} - \varphi_2 W_2 - \ldots - \varphi_p W_{t-p}]^2$$

Obteniendo las parciales con respecto a $\varphi_1, \varphi_2, \ldots \varphi_p$:

$$\frac{\partial(\Phi,\Theta)}{\partial \varphi_1} = -2\sum_t^T [W_t - \varphi_1 W_{t-1} - \varphi_2 W_{t-2} - \ldots - \varphi_p W_{t-p}]\ W_{t-1}$$

.

.

.

$$\frac{\partial(\Phi,\Theta)}{\partial \varphi_p} = -2\sum_t^T [W_t - \varphi_1 W_{t-1} - \varphi_2 W_{t-2} - \ldots - \varphi_p W_{t-p}]\ W_{t-p}$$

Igualando a cero las parciales se obtiene las siguientes ecuaciones:

$$\sum_t^T W_t W_{t-1} = \varphi_1 \sum_t^T W_{t-1}{}^2 + \varphi_2 \sum_t^T W_{t-1} W_{t-2} + \ldots + \varphi_p \sum_t^T W_{t-1} W_{t-p}$$

$$\sum_t^T W_t W_{t-2} = \varphi_1 \sum_t^T W_{t-2} W_{t-1} + \varphi_2 \sum_t^T W^2{}_{t-2} + \ldots + \varphi_p \sum_t^T W_{t-2} W_{t-p}$$

.

.

.

$$\sum_t^T W_t W_{t-p} = \varphi_1 \sum_t^T W_{t-p} W_{t-1} + \varphi_2 \sum_t^T W_{t-p} W_{t-2} + \ldots + \varphi_p \sum_t^T W^2{}_{t-p}$$

Este sistema de ecuaciones puede escribirse matricialmente como:

$$(X^T Y) = (X^T X)\varphi \ \_ó$$

$$\varphi = (X^T X)^{-1}(X^T Y)$$

que seria una forma matricial de un modelo de regresión lineal múltiple.

Por lo que concluimos que para un modelo puramente autorregresivo el proceso de estimación es esencialmente una regresión lineal, el problema es más difícil si el modelo contiene un componente de medias móviles, como por ejemplo:

$$W_t = e_t - \theta_1 e_{t-1} \Rightarrow$$

$e_t = W_t + \theta_1 e_{t-1} \Rightarrow$ minimizando los cuadrados

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

$$\sum_{t=1}^{T} (W_t + \theta_1 e_{t-1})^2$$

de aquí que:

$$S(\Theta) = \sum_{t=1}^{T} e_t^2 = \sum_{t=1}^{T} [(1-\theta_1 B)^{-1} W_t \ ]^2$$

Aplicando el operador salto hacia atrás:

$$W_t = e_t - \theta_1 B e_t$$

$$W_t = (1-\theta_1 B) e_t \Rightarrow$$

$$e_t = W_t (1-\theta_1 B)^{-1}$$

Para Minimizar se obtiene:

$$\frac{\partial S(\Theta)}{\partial \theta_1} = 2\sum_{t=1}^{T} [(1-\theta_1 B)^{-1} W_t \ ][-1(1-\theta_1 B)^{-2} (-BW_t)] =$$

$$= 2\sum_{t=1}^{T} [(1-\theta_1 B)^{-1} W_t \ ][W_{t-1} (1-\theta_1 B)^{-2}]$$

igualando a cero

$$= 2\sum_{t=1}^{T} [(1-\theta_1 B)^{-3} W_t \ W_{t-1}]$$

$$\sum_{t=1}^{T} [(1-\theta_1 B)^{-3} W_t \ W_{t-1}] = 0 \Rightarrow$$

$$\frac{1}{(1-\theta_1 B)^3} \sum_{t=1}^{T} W_t \ W_{t-1} = 0$$

Como se obtiene una ecuación no lineal, no puede aplicarse un método analítico o numérico, y la solución consistirá en aproximar la función e_t a una función lineal.

Se usarán los dos primeros términos de la expansión por serie de Taylor, alrededor de un valor inicial de Φ y Θ, se realiza una regresión lineal en la nueva ecuación linealizada, se obtienen estimadores de mínimos cuadrados de Φ y Θ estos estimadores se usan como variables iniciales para una nueva inicialización. Este proceso se repite hasta que los valores converjan. Del Libro "Modelos de Decisión con Modelos Estocásticos, de la Act. M. Carmen González Videgaray, que es de donde se obtienen casi toda la base teórica del presente trabajo, tenemos:

Para observar este proceso iterativo sea β el vector que represente los *p+q* parámetros de Φ y Θ que se desean estimar. Se desea elegir valores para β que minimicen:

$$S(\beta) = \sum_{t=1}^{T} [e_t | W_t, \beta]^2 = \sum_{t=1}^{T} [e_t]^2$$

Si extendemos a e_t en una serie de Taylor alrededor de un valor inicial β_0 para los valores de los parámetros.

$$[e_t] = [e | W_t \beta_0] + \sum_{t=1}^{p+q} (\beta_i - \beta_{i,0}) \frac{\partial [e_t]}{\partial \beta_i}\bigg|_{\beta=\beta_0} + \frac{1}{2} \sum_{t=1}^{p+q} (\beta_i - \beta_{i,0})^2 \frac{\partial^2 [e_t]}{\partial \beta_i^{\,2}}\bigg|_{\beta=\beta_0} + \ldots$$

Aquí $\beta_{i,0}$ es el valor inicial de β_i, es decir, el í-ésimo componente del vector β_0.
Se aproximará $[e_t]$ con los primeros términos de de la expansión de la Serie de Taylor, entonces sea:

$$x_{i,t} = -\frac{\partial [e_t]}{\partial \beta_i}\bigg|_{\beta=\beta_0} \; y$$

$$[e_{t,0}] = [e_t | W_t, \beta_0]$$

Se sigue que:

$$[e_t] \approx [e_{t,0}] - \sum_{i=1}^{p+q} (\beta_i - \beta_0) x_{i,t}$$

que puede rescribirse como:

$$[e_{t,0}] = \sum_{i=1}^{p+q} \beta_{i,0} x_{i,t} = \sum_{i=1}^{p+q} \beta_0 x_{i,t} + [e_{t,0}] \ldots\ldots (\xi)$$

El lado izquierdo de la ecuación (ξ) puede verse como una variable dependiente compuesta que tendrá diferentes valores para t=1,2,….T (Notar que $[e_{t,0}]$ es el valor del error al tiempo t, dado el valor inicial β_0. Del lado derecho de la ecuación hay *p+q* variables independientes (multiplicadas por *p+q* parámetros desconocidos β_i así como un error aditivo.

Queda claro que los parámetros β_i pueden estimarse de la ecuación (ξ) por regresión lineal:

$$Y = \begin{bmatrix} [e_{t,0}] + \sum_{i=1}^{p+q} \beta_{i,0} x_{i,1} \\ . \\ . \\ . \\ [e_{t,0}] + \sum_{i=1}^{p+q} \beta_{i,0} x_{i,T} \end{bmatrix} \quad \text{X}= \begin{bmatrix} x_{1,1} & x_{2,1} & \dots & x_{p+q,1} \\ & & . & \\ & & . & \\ x_{1,T} & x_{2,T} & \dots & x_{p+q,T} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ . \\ . \\ . \\ \beta_{p+q} \end{bmatrix}$$

Que puede ser expresado como:

$$Y = X\beta + [e_t]$$

Se realiza entonces la regresión ordinaria por mínimos cuadrados para producir un nuevo estimador de β llamado β_i

En seguida, usando una nueva expresión de Taylor de $[e_t]$ alrededor de β_1, se obtiene una nueva versión de (ξ) que también se puede estimar por mínimos cuadradros ordinarios para generar un nuevo estimador β_2. Este proceso se repite una y otra vez hasta que:

$$\begin{matrix} |\beta_k - \beta_{k-1}| \approx 0 \\ |\beta_k - \beta_{k-1}| \leq \varepsilon \end{matrix} \ \text{ó}$$

el valor “k” se le llama valor de convergencia, así como el número de iteraciones requeridas para que ocurra la convergencia. β_k, será nuestro estimador final de $\varphi_1, \varphi_2, ..\varphi_p, \theta_1, \theta_2, ..\theta_q$

Algunos paquetes estadísticos no sólo toman esta diferencia como criterio de convergencia, sino también la suma de los cuadrados de los residuales (analizando las variaciones sucesivas de ésta), detenido el proceso cuando se cumpla alguna de las dos condiciones.

Los errores estándar y los estadísticos *t* de nuestros estimadores se calcula a partir de la última linealización, de la misma forma que en cualquier procedimiento de estimación lineal. De modo que modo que los estimadores "t" tienen poca relevancia, indicando sólo la significancia de los estimadores de mínimos cuadrados obtenidos en la linealización final de la ecuación no lineal. De igual manera se puede calcular R^2 (Coeficiente de determinación), que tendrá una importancia similar. Por esta razón no necesariamente deberá rechazarse una ecuación con una R^2 pequeña:

$$R^2 = 1 - \frac{\sum e_t^2}{\sum w_t^2} = \frac{\sum w_t^2 - \sum e_t^2}{\sum w_t^2}$$

Aunque se obtenga un valor pequeño para R^2 en la última iteración es posible que el modelo no-lineal tenga considerable valor predictivo.

Es importante señalar que no hay garantía que el proceso de estimación descrito arriba converja siempre a un estimador final de los parámetros. Es muy posible que el proceso diverja y que los estimadores sucesivos $\beta_{1,,}\beta_2 \ldots \beta_k$ se alejen cada vez más. Aun más, también es posible que existan soluciones múltiples, en este caso ocurre la convergencia pero con diferentes valores iniciales lleva a diferentes estimadores finales de uno o más parámetros. El hecho que exista divergencia o múltiples soluciones depende de la especificación del modelo ARIMA como de los datos.

4.1.5 Obtención de un valor inicial de los parámetros.

Antes de poder realizar la estimación no lineal, debe hacerse la inicialización tentativa de los parámetros

$$\beta = (\Phi, \Theta) = (\hat{\varphi}_{1,0}, \hat{\varphi}_{2,0}, ..., \hat{\varphi}_{p,0}, \hat{\theta}_{1,0}, \hat{\theta}_{2,0}, ..., \hat{\theta}_{q,0})$$

Estos valores deben ser lo más cercanos posibles a los valores verdaderos para que el modelo converja.

Existen varias opciones para obtener estos valores:

Utilizar la ACF muestral,

Para un AR(1)

PACF= $\hat{\rho}_{11} = \varphi_1 = r_1$

Para un *AR(2)*

PACF= $$\hat{\rho}_{11} = \varphi_1 = r_1$$
$$\hat{\rho}_{22} = r_{22} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

Si es un *MA(1)*

$$\hat{\rho}_1 = \frac{-\theta_1}{1-\theta_1^2} \Rightarrow \hat{\theta}_1 = r_1\theta_1^2 - r_1 \Rightarrow \hat{\theta}_1 - r_1\theta_1^2 + r_1 = 0$$

La cual puede resolverse como una ecuación de segundo grado, con lo que obtenemos los siguientes valores:

$$\theta_1^2 = \frac{1 \pm \sqrt{1-4r_1}}{-2r_1}$$

Y se toma el valor que satisfaga $|\theta_1| < 1$ (condición de invertibilidad)

Algunos autores sugieren que los valores iniciales de los parámetros pueden ser cero.

Si es un *ARMA(1,1)*

Si es un modelo mezclado de orden (1,1), es un modelo no-lineal

$$(1-\varphi_1 B)W_t = (1-\theta_1 B)e_t \Rightarrow$$

$$e_t = \frac{(1-\varphi_1 B)W_t}{1-\theta_1 B}$$

Esta ecuación es no-lineal con respecto a θ_1, pero puede aproximarse con una expansión en serie de Taylor. Esto implica calcular las primeras derivadas de e_t con respecto a $\theta_{1,0}$ y $\varphi_{1,0}$.

$$x_{1,t} = -\frac{\partial[e_t]}{\partial\varphi_1}\bigg|_{\substack{\varphi_{1,0}\\ \theta_{1,0}}} = \frac{B}{1-\theta_{1,0}B}W_t ......(\psi)$$

$$x_{2,t} = -\frac{\partial[e_t]}{\partial\theta_1}\bigg|_{\substack{\varphi_{1,0}\\ \theta_{1,0}}} = \frac{B-\varphi_{1,0}B^2}{(1-\theta_{1,0}B)^2}W_t ......(\zeta)$$

Pueden calcularse series de tiempo numéricas para $x_{1,t}$ y $x_{2,t}$, (para hacer una regresión lineal) para t=1,2,…,T. Por ejemplo de (ψ):

$$(1-\theta_{1,0}B)x_{1,t} = BW_t$$

$$x_{1,t} - \theta_{1,0}x_{1,t-1} = W_{t-1}$$

$$x_{1,t} = \theta_{1,0}x_{1,t-1} + W_{t-1} .....(\varsigma)$$

Haciendo $W_0 = x_{1,0} = 0$ puede resolverse la ecuación (ς) en forma sucesiva para generar la serie $x_{1,t}$:

$$x_{1,1} = 0$$
$$x_{1,2} = W_1$$
$$x_{1,3} = \theta_{1,0} W_1 + W_2$$

.

.

.

El mismo proceso se hace para (ζ)

$$x_{2,t} = 2\theta_{1,0} x_{2,t-1} - \theta^2{}_{1,0} x_{2,t-2} - W_{t-1} + \varphi_{1,0} W_{t-2}$$
$$con\_x_{2,1} = x_{2,0} = W_0 = W_{-1} = 0$$
$$x_{2,1} = 0$$
$$x_{2,2} - W_1$$
$$x_{2,3} = 2\theta_{1,0} x_{2,2} - z_{2,} + \varphi_{1,0} W_1$$

.

.

.

Finalmente se debe calcularse una serie para $[e_{t,0}]$, que son los residuales basados en la primera aproximación de $\theta_{1,0}$ y $\varphi_{1,0}$

Ahora se puede regresar la regresión lineal con:

$$e_{t,0} = \varphi_{1,0} x_{1,t} + \theta_{1,0} x_{2,t} = \varphi_1 x_{1,t} + \theta_1 x_{2,t} + e_t$$

El proceso se repite de forma iterativa.

4.1.6 Examen del modelo

Una vez que el modelo se ha identificado y sus parámetros se han estimado, es necesario verificar la adecuación del modelo y analizar si se puede mejorarse.

Para el estimador del parámetro pueden calcularse varias estadísticas:

1.- Media
2.- Estadístico t
3.- χ^2
4.- Valor de la constante
5.- Intervalo de confianza

Además se tiene una serie de herramientas para realizar el análisis de los residuales:
1.- ACF y PACF de los residuales
2.- Gráfica de los residuales
3.- Suma de cuadrados de los residuales
4.- Intervalos de confianza
5.- Graficar residuales contra la normal
6.- Periodograma
7.- Periodograma integral

Se realizaron 4 pruebas de diagnóstico y con base en todas ellas se tomará la decisión de modificar o no el modelo inicial, éstas son:

a) Análisis de Estacionalidad
b) Análisis de residuales
c) Modelo sub-especificado (agregar parámetros)
d) Modelo sobre-especificado (quitar parámetros)

4.1.7 Análisis de Estacionaridad

Se toma la serie de residuales como una nueva serie de tiempo para analizar su estacionalidad, si se detectan problemas deberán de analizarse de nuevo la estacionaridad del modelo; debe tenerse cuidado en no sobrediferenciar el modelo; ya que se pierden dos observaciones y el modelo se complica de manera innecesaria.

4.1.8 Análisis de los residuales

Los residuales se analizarán bajo los siguientes aspectos:

1.- Gráfica de residuales, se detectan posibles datos discrepantes y puede aparecer variación estacional.
2.- La ACF y la PACF teóricamente deben contener únicamente ceros, esto puede determinarse en forma usual, por el hecho de que todos los coeficientes deben quedar dentro del intervalo de confianza.

En forma de prueba de hipótesis se realiza la prueba de χ^2 con el estadístico de Box_Pierce:

Para probar: $\begin{matrix} H_0 := \rho_1 = \rho_2 = \ldots = \rho_k \\ H_a := \rho_k \neq 0 \end{matrix}$ alguna diferente de cero

$$Q(k) = n(n+2)\sum_{k=1}^{k} \frac{r_k^{\,2}(e)}{n-k}$$

$$Q(k) < \chi_{(k)}^{\;2}$$

Donde $r_k^2(e)$ es la k-ésima autocorrelación de los residuales

$Q(k) \sim \chi^2$

con .grados de libertad

$k-(p+q)$

Regla de decisión: Aceptar H_0 si $Q(k) < \chi^2$

Ejemplo 1: Supóngase que se decidió estimar un modelo MA(1)

$$w_t = a_t - \theta_1 a_{t-1} = a_t - \theta_1(B)a_t = (1-\theta_1(B))a_t$$

Y se encuentra evidencia de que los residuales a_t siguen también un proceso MA(1) (la ACF se trunca en 1 y la PACF es decreciente infinita)

$a_t = e_t - \lambda e_{t-1}$, donde supuestamente e_t es ruido blanco

Sustituyendo
:

$$(e_t - \lambda e_{t-1}) - \theta(B)(e_t - \lambda e_{t-1}) = e_t - \lambda e_{t-1} - \theta(B)e_t + \theta(B)\lambda e_{t-1}$$
$$= e_t - (\lambda + \theta)e_{t-1} + \theta\lambda e_{t-2} \ldots MA(2)$$

La influencia del error aleatorio nos indica que el modelo a analizar deberá ser un MA(2).

Ejemplo 2: Supongamos ahora que los residuales tienen un comportamiento AR(1).¿Cuál es el siguiente modelo a probar?

$$w_t = a_t - \theta_1 a_{t-1} = a_t - \theta_1(B)a_t = (1-\theta_1(B))a_t \ldots Mod\_Original$$
$$a_t = \theta_1 a_{t-1} + e_t \ldots residuales$$
$$sustituyendo:$$
$$w_t = (1-\theta_1(B))(\varphi_1 a_{t-1} + e_t) = \varphi_1 a_{t-1} + e_t - \theta_1(B)\varphi_1 a_{t-1} - \theta_1(B)e_t =$$
$$w_t = \varphi_1 a_{t-1} + e_t - \theta_1\varphi_1 a_{t-2} - \theta_1 e_{t-1} \ldots\ldots ARMA(2,1)$$

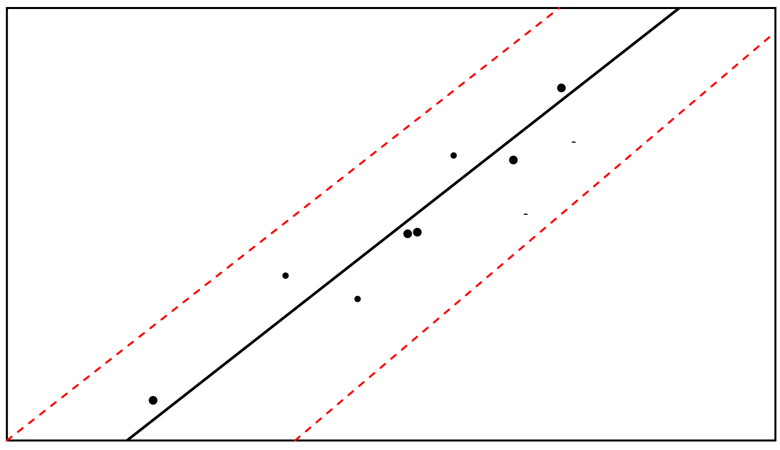
Ejemplo 3: Calcular una diferencia a los residuales:

$$\Delta e_t - e_t - e_{t-1} \ldots . MA(1)$$

Los residuales son “ruido blanco”, entonces sus primeras diferencias deben seguir un modelo MA(1) con θ_1 cercano a 1.

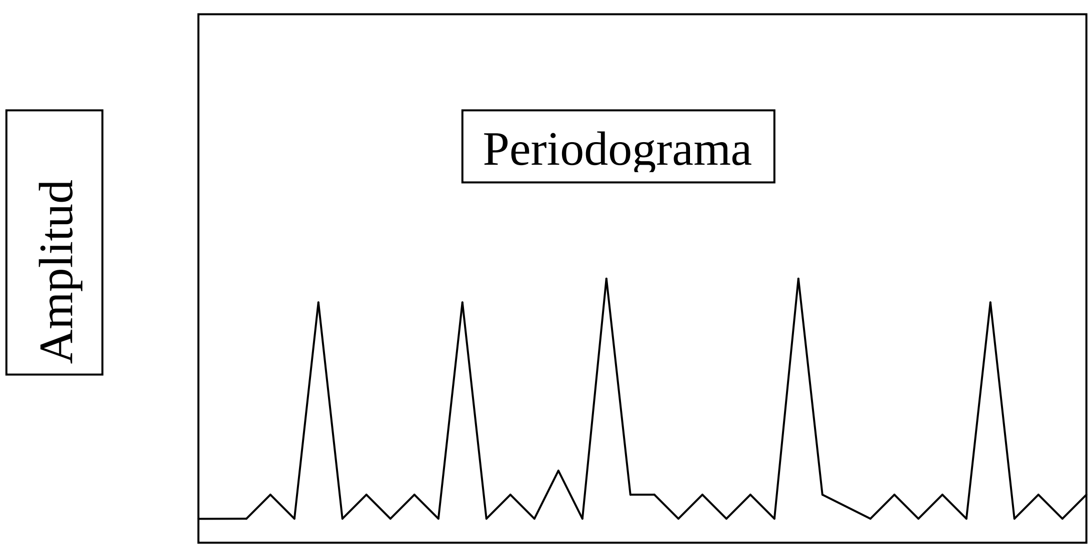
Ejemplo 4:

Se prueba el ajuste de residuales a la normal. Se compara la gráfica de frecuencias relativas acumuladas de la serie contra una distribución normal.

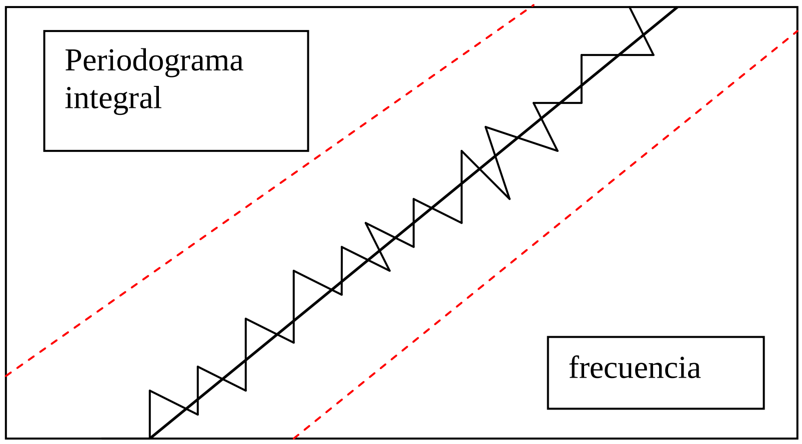


Los datos deben ajustarse a la línea recta y no salir del intervalo.

Ejemplo 5: El periodograma y el periodograma integral se usan para detectar posibles variaciones estacionales.



Si los datos presentan variación estacional, entonces el periodograma mostrará picos en cada frecuencia.



El periodograma integral también se conoce como periodograma normalizado acumulativo.

Si los residuales son ruido blanco, entonces su periodograma se aproximará a la diagonal.

Si existe periodicidad de orden S, se encontrarán escalones o saltos en los períodos S, 2S, 3S, etc. Es decir en las frecuencias $\frac{1}{S}, \frac{2}{S}, \frac{3}{S}, ...$

4.1.9 Modelos sobre-especificados

Los parámetros redundantes pueden localizarse utilizando el estadístico “t” de cada estimador y la matriz de correlación entre los estimadores.

Para probar la hipótesis de que un parámetro β no es significativo para el modelo, se realiza la prueba de hipótesis:

$$H_0 : \hat{\beta}_i = 0$$

$H_a : \hat{\beta} \neq 0;$ Regla de decisión: $Aceptar\_H_0\_\_si\_|T| < 2$

$$T = \frac{\hat{\beta}}{\sigma_\beta}$$

Observaciones:

1.- si el parámetro no significativo **NO** es el de mayor orden (ya sea de la parte AR o Ma, se requiere de mayor análisis
2.- Si el parámetro no significativo es el de mayor orden, entonces se elimina.
3.- Si el parámetro no significativo esta correlacionado con el de mayor orden, entonces se elimina (el de mayor orden)
4.- Si no están correlacionados entonces, implica un modelo multiplicativo.

4.1.10 Modelo sub-especificados (agregar parámetros)

Para verificar que el modelo es adecuado, puede incluirse un parámetro adicional (conforme al comportamiento de los residuales y verificar si se mejora o no el modelo).

Recomendación: Para no sobre-especificar el modelo se deben agregar sólo un parámetro a la vez. (por que si no, no sabremos cual es el que podría estar afectando a nuestro modelo).
Finalmente después de efectuar al diagnóstico de todos los modelos tentativos, es conveniente resumir los resultados.

4.1.10.1 Modelos Estacionales

En algunas series de tiempo existen similitudes en ciertos periodos, por ejemplo cada 7 días, 12 meses, cada navidad o verano, etc.

Para analizar estas series de tiempo se utilizarán modelos estacionales.

4.1.10.2 Modelos Estacionales Autoregresivos

Estos modelos se aplican si el valor actual de la serie Z_t puede expresarse como una función lineal del valor de la serie “S” periodos atrás, “12 periodos atrás, etc.; donde “S“ representa el orden de la variación estacional y además el error aleatorio e_t.
el modelo autorregresivo de orden 1 es:

SAR(1) ó $AR(1)_s$ (en este modelo se usan unidades de periodo, ya no de tiempo)

$$AR(1)_s = SAR(1) = \varphi_1 Z_{t-s} + e_t = Z_t$$

El modelo SAR(2) o $AR(2)_s$

$$AR(2)_s = SAR(2) = \varphi_1 Z_{t-s} + \varphi_2 Z_{t-2s} + e_t = Z_t$$

Si se aplican diferencias estacionales, entonces:

$$Z_t = \Delta_s^D W_t$$

En general SAR(p): $Z_t = \varphi_1 Z_{t-s} + \varphi_2 Z_{t-2s} + \ldots + \varphi_p Z_{t-ps} + e_t$

De otra forma:

$$e_t = Z_t - \varphi_1 Z_{t-s} - \varphi_2 Z_{t-2s} - \ldots - \varphi_p Z_{t-ps}$$
$$e_t = Z_t - \Phi(B^S) Z_t$$
$$donde:$$
$$\Phi(B^S) = 1 - \varphi_1 B^S - \varphi_2 B^{2S} - \ldots - \varphi_P B^{PS}$$
$$y$$
$$B^S Z_t = Z_{t-s}$$
$$B^{2S} Z_t = Z_{t-2s}$$

P = grado del modelo

.

.

.

$$B^{PS} Z_t = Z_{t-PS}$$

Si quisiéramos calcular las ACF de un modelo SAR(1) ó $AR(1)_s$. Sería la siguiente fórmula:

ACF: $\rho_k = \dfrac{\operatorname{cov}(Z_t, Z_{t-1})}{\operatorname{var}(Z_t)} = E(\dfrac{(Z_t, Z_{t-k})}{\operatorname{var}(Z_t)} = \dfrac{\gamma_k}{\gamma_0}$

$$Z_t = \varphi_1 Z_{t-s} + e_t$$
$$e_t \sim N(0, \sigma^2)$$

$$\gamma_0 = Var(\varphi_1 Z_{t-s} + e_t) = \varphi_1^{\,2} Var(Z_{t-s}) + \sigma^2$$
$$Var(Z_t) = Var(Z_{t-s}) \Rightarrow \varphi_1^{\,2}\gamma_0 + \sigma^2 \Rightarrow \gamma_0(1 - \varphi_1^{\,2}) = \sigma^2$$
$$\therefore \gamma_0 = \frac{\sigma^2}{1 - \varphi_1^{\,2}}$$

Ahora para γ_1

$$\frac{\gamma_1}{\gamma_0} = \frac{E(Z_{t,} Z_{t-1})}{Var(Z_t)} \Rightarrow \gamma_1 = E[(\varphi_1 Z_{t-s} + e_t)(Z_{t-1})] = E[\overbrace{\varphi_1 Z_{t-s} Z_{t-1}}^{0} + \underbrace{e_t Z_{t-1}}_{0}] \Rightarrow \gamma_1 = 0$$

Para γ_2

$$\gamma_2 = E(Z_{t,}Z_{t-2}) = E[(\varphi_1 Z_{t-s} + e_t)(Z_{t-2})] = E[\overbrace{\varphi_1 Z_{t-s} Z_{t-2}}^{0}] + \underbrace{E[e_t Z_{t-2}]}_{0} = 0$$

$$\Rightarrow \gamma_2 = 0$$

Para γ_S

$$\gamma_S = E(Z_t, Z_{t-s}) = E(\varphi_1 Z_{t-s} + e_t)(Z_{t-s})) = E[\varphi_1 Z_{t-s}{}^2 + e_t Z_{t-s}] = E[\varphi_1 Z_{t-s}{}^2] + \underbrace{E[e_t Z_{t-s}]}_{0} \Rightarrow$$

$$\gamma_S = \varphi_1 E[Z_{t-s}{}^2] = \varphi_1 Var(Z_{t-s}) = \varphi_1 \gamma_0$$

.

.

.

$$\gamma_{S+1} = 0$$

$$\gamma_{S+2} = 0$$

.

.

.

$$\gamma_{2s} = E(Z_t, Z_{t-2s}) = \varphi_1 Z_{t-s} + e_t)(Z_{t-2s})) = E[(\varphi_1 Z_{t-s} Z_{t-2s}) + (e_t Z_{t-2s})] = E[\varphi_1 Z_{t-s} Z_{t-2s}] + \underbrace{E[e_t Z_{t-2s}]}_{0}$$

$$= E[\varphi_1(\varphi_1 Z_{t-2s} + e_{t-s})(Z_{t-2s})]\_\_ya\_que\_Z_{t-s} = \varphi_1 Z_{t-2s} + e_{t-s}$$

$$E[\varphi_1{}^2 Z_{t-2s} + \varphi_1 e_{t-s})(Z_{t-2s})] = E[(\varphi_1{}^2 Z_{t-2s}{}^2)] + \underbrace{E[(\varphi_1 e_{t-s})(Z_{t-2s})]}_{0} = \varphi_1{}^2 E[Z_{t-2s}{}^2] \Rightarrow \gamma_{2s} = \varphi_1{}^2 Var(Z_{t-2s}) =$$

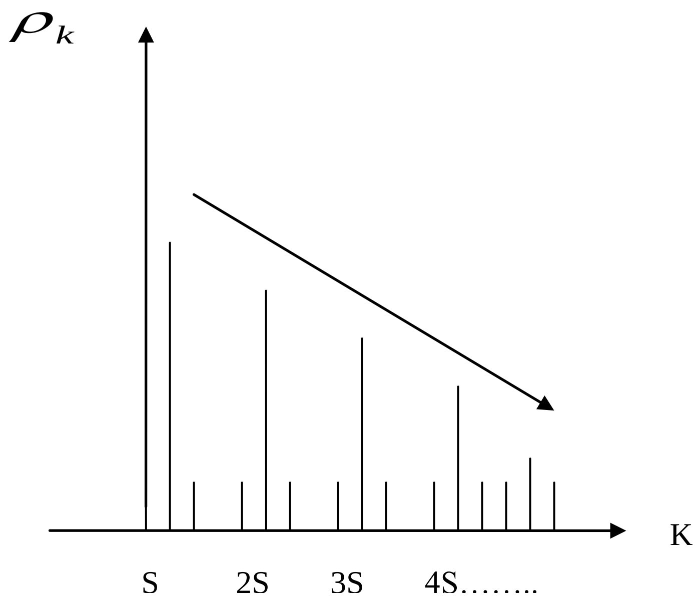
$$\varphi_1{}^2 \gamma_0$$

En general: $\gamma_k = \begin{cases} \varphi^{\frac{k}{s}} \gamma_0 & k = s, 2s, 3s, \ldots, \\ 0 & cualquier\_otro \end{cases}$

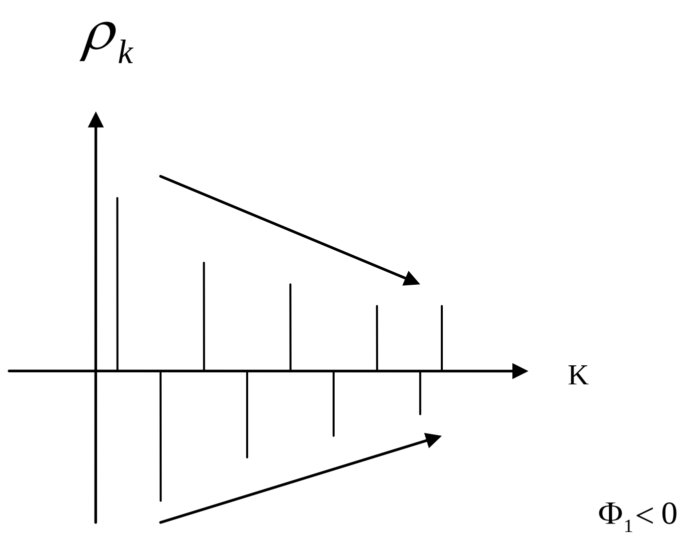
Por lo tanto la ACF:

$$\rho_k = \begin{cases} \varphi^{\frac{k}{s}} \gamma_0 & k = s, 2s, 3s, \ldots, \\ 0 & cualquier\_otro \end{cases}$$

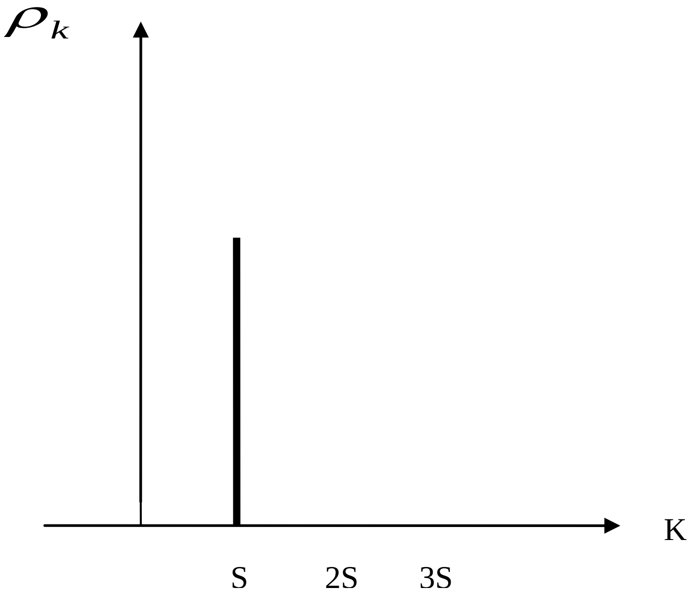
Gráficamente:



La ACF es decreciente infinita $\Phi > 0$



Como observamos la ACF muestra la misma forma que la de los modelos ordinarios excepto por los valores significativos en S, 2S,3S… etc. De igual forma para la PACF tendremos que debe de truncarse en S (para un SAR(1)).



Por lo tanto, para un modelo autorregresivo general AR(p):

ACF es decreciente infinita con valores significativos en los múltiplos de S.

PACF, se trunca en Incorporados estos datos en P_S (que tiene valores significativos en los múltiplos de S, anteriores al valor en que se trunca).

4.1.11 Modelos de Medias Móviles Estacionales

Se usan si el valor actual de Z_t puede representarse como una función lineal del error actual e_t, del error S periodos atrás e_{t-s}, el de 2S periodos atrás e_{t-2s}, etc.
El modelo SMA(1) ó $MA(1)_s$:

$$Z_t = e_t - \theta_1 e_{t-S} - \theta_2 e_{t-2S} - ... - \theta_Q e_{t-QS}$$

De otra manera:

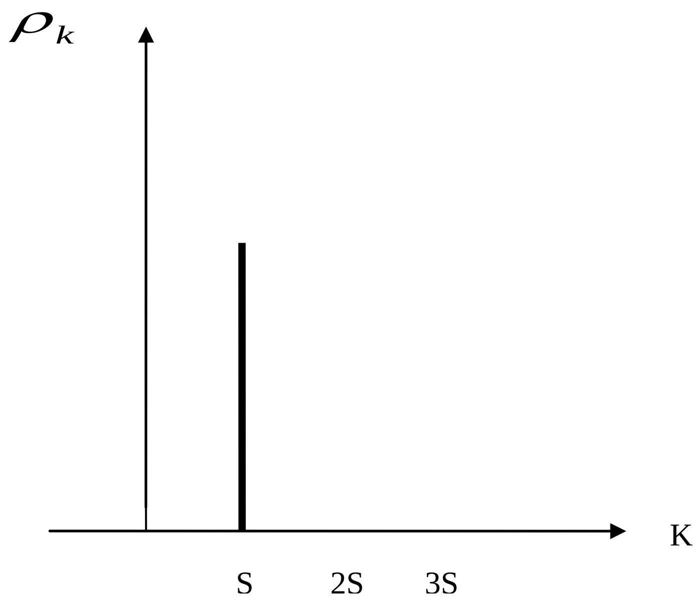
$$Z_t = \Theta B^s (e_t)$$

Donde:

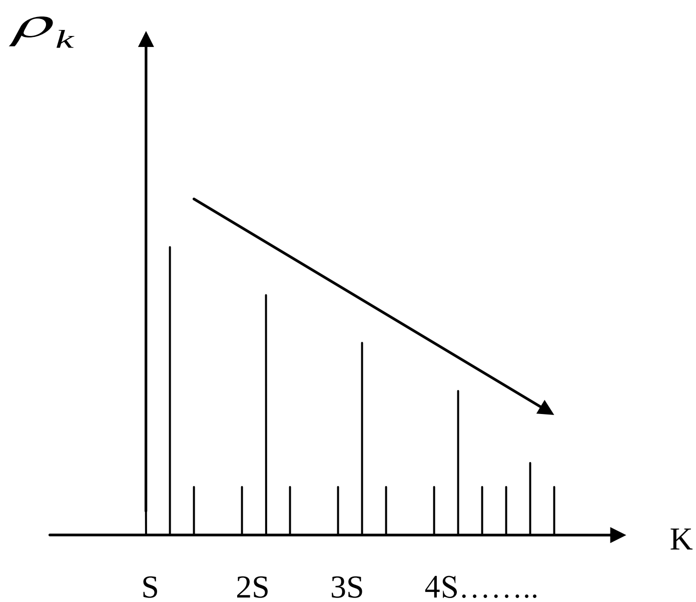
$$\Theta B^s = (1 - \theta_1 B^s - \theta_2 B^{2s} - ... - \theta_Q B^{Qs})$$

Si hubo diferencias estacionales: $Z_t = \Delta_s^D W_t$

Para un SMA(1) $ACF = \upsilon_k = \begin{cases} \dfrac{-\theta_1}{1+\theta_1^2} & k = s \\ 0 & cualquier\_otro \end{cases}$



ACF: Se trunca



PACF: Decreciente infinita

En General:

ACF: Se trunca en Q_S con valores significativos en múltiplos de S (que sean menores al valor que se trunca)

PACF: Decreciente infinita con valores significativos en múltiplos de S.

4.1.12 Modelos de Medias Móviles Estacionales

Se componen de una parte SMA y una SAR, tiene forma: ARIMA(P,D,$Q)_S$ ó SARIMA(P,D.Q):

$$Z_t = \varphi_1 Z_{t-s} + \varphi_2 Z_{t-2s} + ... + \varphi_p Z_{t-ps} - \theta_1 e_{t-s} - \theta_2 e_{t-2s} - ... - \theta_Q e_{t-Qs} + e_t$$

Para estos modelos se tiene que:

ACF: es decreciente infinita, con valores significativos en múltiplos de S.
PACF: es decreciente infinita con valores significativos en múltiplos de S.

4.1.13 Modelos Generales multiplicativos Estacionales

Finalmente pueden combinarse los modelos en una base general que para la mayoría de las series de tiempo proporciona muy buen ajuste y pronóstico adecuados. Esta clase se llama “Modelos ARIMA multiplicativos”, y se expresa como:

$$\Phi(B)\Phi(B^s)Z_t = \Theta(B)\Theta(B^s)e_t$$

Es decir:
ARIMA(p,d,q)X(P,D,$Q)_S$

Por ejemplo:
ARIMA(1,0,0)X(1,0,0$Q)_4$

El primer modelo es un AR(1), el segundo factor es un SAR(1), con S=4

$$Ar(1) = (1-\varphi_1)BZ_t$$

$$SAR(1)\_con\_s = 4 = (1-\varphi_1)B^4 Z_t = e_t$$

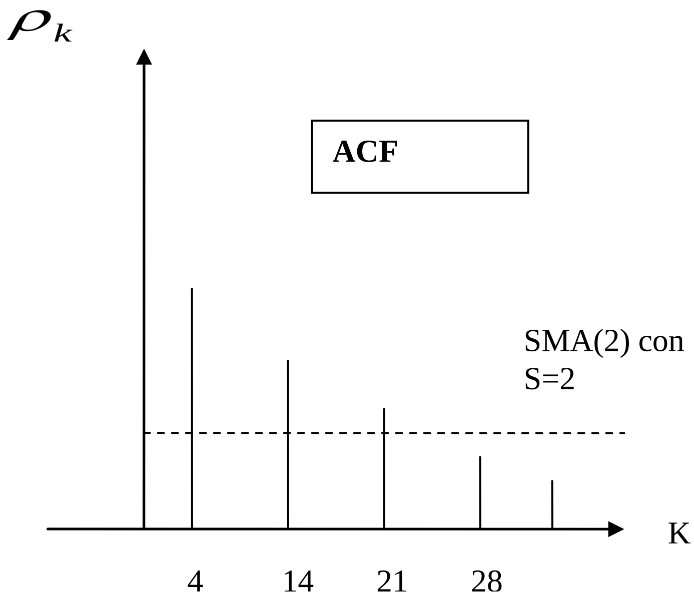
$$ARIMA\_X\_SAR(1) = Z_t = (1-\Theta_1)BZ_t(1-\Phi_1)B^4 Z_t = e_t \Rightarrow$$

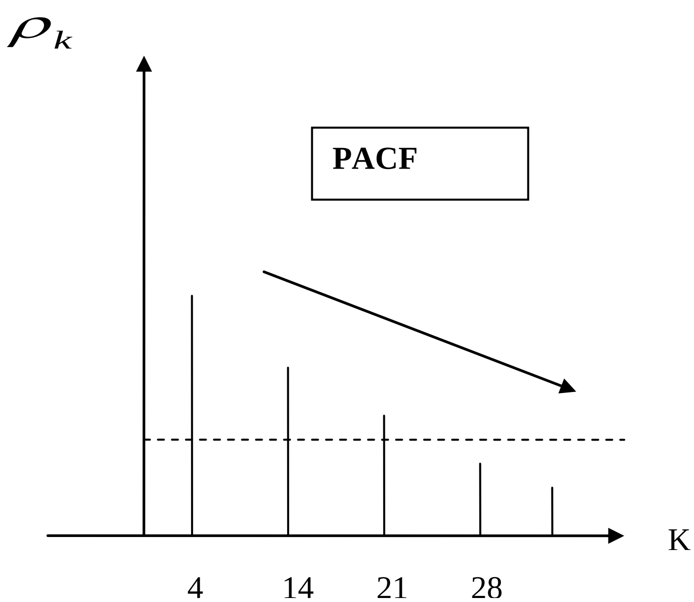
$$Z_t = \frac{e_t}{(1-\varphi_1)B(1-\Theta_1)B^4} \Rightarrow Z_t = \varphi_1 Z_{t-1} + \Theta Z_{t-4} + e_t$$

$$AR(1)XMA(1)_3 = ARIMA(1,0,0)X(0,0,1)_3$$

$$Z_t = \varphi_1 Z_{t-1} - \theta_1 e_{t-3} + e_t(1-\theta)BZ_t = (1-\Theta)B^3 e_t$$

$$Z_t = \frac{(1-\Theta)B^3 e_t}{(1-\theta)B}$$





Multiplicativos

$$\varphi(B)\Phi(B^s)W_t = \theta(B)\Theta(B)e_t$$

ARIMA(p,d,q)XSARIMA(P,D,Q)

Para identificar un modelo multiplicativo es recomendable realizar el proceso de identificación por etapas y analizar los residuales.

Ejemplo: Suponga que los datos W_t se modelan con un SAR(1) o un AR$(1)_4$:

$$(1-\Phi(B^4))W_t = e_t$$

$$ó$$

$$W_t = \Phi W_{t-4} + e_t$$

Ahora suponga que los residuales tienen un comportamiento semejante a un modelo AR(1):

$$(1-\Phi_1 B)W_t = e_t$$

$$ó$$

$$W_t = \Phi W_{t-1} + e_t$$

El modelo general es de la forma:

$$(\varphi_1(B)\Phi(B^S)W_t = e_t$$

$$sustituyendo:$$

$$(1-\varphi_1 B)(1-\Phi B^4)W_t = e_t$$

$$desarrollando\_la\_parte\_ordinaria:$$

$$(W_t - \varphi_1 W_{t-1})(1-\Phi B^4) = e_t$$

$$W_t - \underbrace{\Phi_1 W_{t-1}}_{\alpha} - \underbrace{\varphi_1 W_{t-1}}_{\beta} + \underbrace{\varphi_1 \Phi W_{t-S}}_{\alpha\beta} = e_t$$

(β) …ARIMA(1,0,0) X SARIMA(1,0,0)… (α) con s=4

Ejemplo: Sea

$$W_t = SMA(1)\_con\_s = 4 \Rightarrow$$

$$SMA(1) = (1-\theta B)e_t = W_t$$

$$\Rightarrow W_t = e_t + \Theta e_{t-4}$$

Residuales de MA(1)

$$MA(1) = W_t = (1-\theta B)e_t$$

$$\Rightarrow W_t = e_t + \theta e_{t-1}$$

El modelo general es de la forma:

$$W_t = \theta(B)\Theta(B^S)e_t$$

$$Sustituyendo:$$

$$W_t = (1-\theta(B))e_t(1-\Theta(B^4))$$

$$W_t = (e_t - \theta e_{t-1})(1-\Theta(B^4)$$

$$W_t = e_t - \Theta e_{t-4} - \theta e_{t-1} + \theta\Theta e_{t-5}$$

Modelo a probar: ARIMA(0,0,1) X SARIMA(0,0,1)

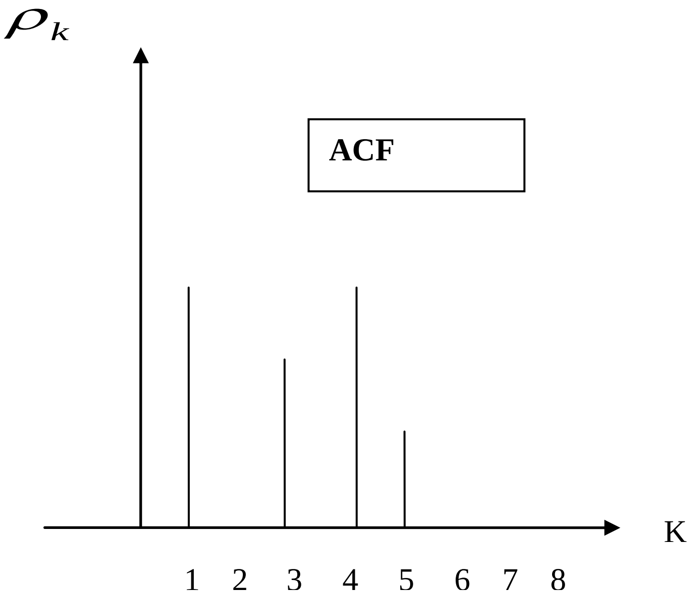
Para este ejemplo la ACF sería:

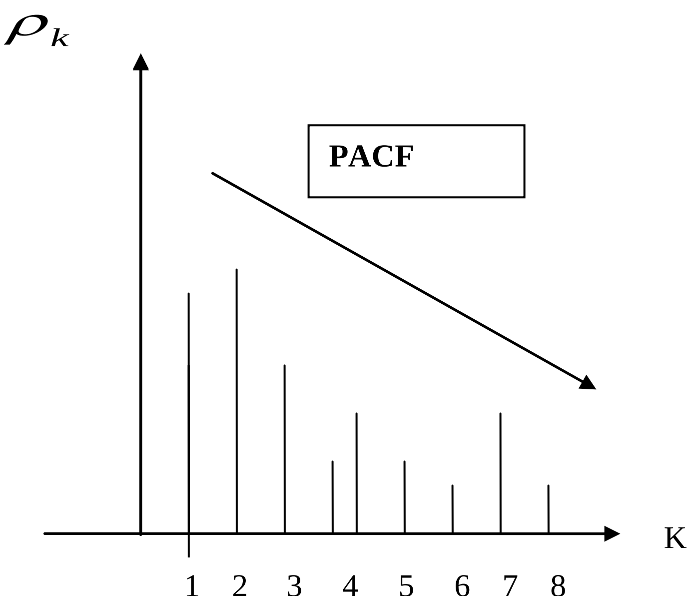
$$\gamma_0 = Var(W_t) = Var(e_t - \Theta e_{t-4} - \theta e_{t-1} + \theta\Theta e_{t-5}) =$$
$$\sigma^2 + \Theta^2\sigma^2 + \theta^2\sigma^2 + \theta^2\Theta^2\sigma^2) = \sigma^2(1 + \Theta^2 + \theta^2 + \theta^2\Theta^2)$$
$$\gamma_1 = Cov(W_t, W_{t-1}) = E[W_t, W_{t-1}] =$$
$$E[(\underbrace{e_t}_{0} - \Theta\underbrace{e_{t-4}}_{0} - \theta e_{t-1} + \theta\Theta e_{t-5})(e_{t-1} - \Theta e_{t-5} - \theta\underbrace{e_{t-2}}_{0} + \theta\Theta\underbrace{e_{t-6}}_{0}) =$$
$$E[-\theta e^2{}_{t-1} - \theta\Theta^2 e^2{}_{t-5}] = Var(-\theta e^2{}_{t-1}) + Var(-\theta\Theta^2 e^2{}_{t-5}) = -\theta\sigma^2 - (\theta\Theta^2)\sigma^2 =$$
$$\Rightarrow \gamma_1 = -\theta\sigma^2(1 + \Theta^2)$$
$$\gamma_2 = Cov(W_t, W_{t-1}) = E[W_t, W_{t-2}] = E[(\underbrace{e_t}_{0} - \Theta\underbrace{e_{t-4}}_{0} - \theta\underbrace{e_{t-1}}_{0} + \theta\Theta\underbrace{e_{t-5}}_{0})(\underbrace{e_{t-2}}_{0} - \Theta\underbrace{e_{t-6}}_{0} - \theta\underbrace{e_{t-3}}_{0} + \theta\Theta\underbrace{e_{t-7}}_{0})$$
$$\Rightarrow \gamma_2 = 0$$
$$\gamma_3 = Cov(W_t, W_{t-3}) = E[W_t, W_{t-3}] =$$
$$E[(\underbrace{e_t}_{0} - \Theta e_{t-4} - \theta\underbrace{e_{t-1}}_{0} + \theta\Theta\underbrace{e_{t-5}}_{0})(\underbrace{e_{t-3}}_{0} - \Theta\underbrace{e_{t-7}}_{0} - \theta e_{t-4} + \theta\Theta\underbrace{e_{t-8}}_{0}) =$$
$$E[\Theta\theta e^2{}_{t-4}] = Var(\theta\Theta e^2{}_{t-4}) \Rightarrow \gamma_3 = \theta\Theta\sigma^2$$
$$\gamma_4 = Cov(W_t, W_{t-4}) = E[W_t, W_{t-4}] = E[(\underbrace{e_t}_{0} - \Theta e_{t-4} - \theta\underbrace{e_{t-1}}_{0} + \theta\Theta e_{t-5})(e_{t-4} - \Theta\underbrace{e_{t-8}}_{0} - \theta e_{t-5} + \theta\Theta\underbrace{e_{t-9}}_{0})$$
$$E[\Theta e^2{}_{t-4} - \theta^2\Theta e^2{}_{t-5}] = Var(\Theta e^2{}_{t-4}) + Var(\theta^2\Theta e^2{}_{t-5}) \Rightarrow \gamma_4 = -\Theta\sigma^2(1 + \theta^2)$$

$$\gamma_5 = Cov(W_t, W_{t-5}) = E[W_t, W_{t-5}] = E[(\underbrace{e_t}_{0} - \Theta\underbrace{e_{t-4}}_{0} - \theta\underbrace{e_{t-1}}_{0} + \theta\Theta e_{t-5})(e_{t-5} - \Theta\underbrace{e_{t-9}}_{0} - \theta e_{t-6} + \theta\Theta\underbrace{e_{t-10}}_{0})$$
$$E[\theta\Theta e^2{}_{t-5}] = Var(\theta\Theta e^2{}_{t-5}) \Rightarrow \gamma_5 = \Theta\theta\sigma^2$$
$$\gamma_6 = 0$$
$$\gamma_7 = 0$$
$$\gamma_8 = 0,\_etc$$

$$\rho_k = \begin{cases} \dfrac{\gamma_1}{\gamma_0} = \dfrac{-\theta\sigma^2(1+\Theta^2)}{\sigma^2(1-\Theta^2+\theta^{2\prime}+\Theta^2\theta^2)} \\ \dfrac{\gamma_2}{\gamma_0} = \dfrac{0}{\sigma^2(1-\Theta^2+\theta^{2\prime}+\Theta^2\theta^2)} = 0 \\ \dfrac{\gamma_3}{\gamma_0} = \dfrac{\theta\Theta\sigma^2}{\sigma^2(1-\Theta^2+\theta^{2\prime}+\Theta^2\theta^2)} \\ \dfrac{\gamma_4}{\gamma_0} = \dfrac{-\Theta\sigma^2(1+\theta^2)}{\sigma^2(1-\Theta^2+\theta^{2\prime}+\Theta^2\theta^2)} \\ \dfrac{\gamma_5}{\gamma_0} = \dfrac{\Theta\theta\sigma^2}{\sigma^2(1-\Theta^2+\theta^{2\prime}+\Theta^2\theta^2)} \\ \dfrac{\gamma_5}{\gamma_0} = 0 \end{cases}$$

La ACF se trunca en S+1=5, no es significativa en k=2, es simétrica alrededor de S=4.





Continuando con los modelos multiplicativos se tienen las siguientes observaciones:

ACF de un modelo MA(q) X SMA(Q) s=4:

- Se trunca en S+1; esto es s+q
- La función es simétrica alrededor de S
- Las PACF del mismo serán decreciente infinita, debida a que un proceso $AR(\infty) \sim MA(q)$

Otro ejemplo para un modelo AR(2) X SAR(1) será:

Si la forma general de un modelo multiplicativo es:

$$\underbrace{\varphi(B)}_{AR(2)}\underbrace{\Phi(B)W_t}_{SAR(1)} = e_t$$

$$W_t = \varphi_1 W_{t-1} + \varphi_2 W_{t-2} \Rightarrow W_t - \varphi_1 W_{t-1} + \varphi_2 W_{t-2} =$$

$$W_t(1-\varphi_1 B + \varphi_2 B^2)\_\_\ y$$

$$W = \Phi W_{t-S} = e_t \Rightarrow (1-\Phi(B^S))W_t = e_t$$

$$Ahora\_sustituyendo:$$

$$(1-\varphi_1 B + \varphi_2 B^2)(1-\Phi(B^S))W_t = e_t \Rightarrow$$

$$W_t - \Phi W_{t-4} - \varphi_1 W_{t-1} + \varphi_1 \Phi W_{t-5} - \varphi_2 W_{t-2} + \varphi_2 \Phi W_{t-6} \approx AR(6)$$

Es decir se aproxima a un AR(6)

Las ACF debe de ser decreciente infinita, valores sobresalientes en múltiplos de S.
PACF se trunca en 6, es decir S+q

Para otro tipo de modelo multiplicativo con una parte AR(SAR) y uno MA(SMA) tenemos:

ACF: Decreciente infinita con valores sobresalientes en múltiplos de S
PACF: Decreciente infinita con valores sobresalientes en múltiplos de S

Si los modelos fueran:

AR(1) SAR(1) la PACF se trunca en: S+1
AR(2) SAR(1) la PACF se trunca en: S+2
AR(1) SAR(2) la PACF se trunca en: 2S+1

4.2 Identificación del modelo propuesto que se ajuste a los datos y que sirva para predecir valores futuros

4.2.1 Estimación de la Serie de Tiempo de las Ventas de la Empresa, período 1993-2006.

Regresando a la serie de tiempo que forman las ventas mensuales del Período 1993-2006, la analizaremos mediante la metodología de Box-Jenkins. Como se tiene como objetivo predecir la facturación del año en curso y 2 años más, se tratará de construir un Modelo que no sólo ajuste bien a los datos observados sino que, además garantice de alguna manera las predicciones también se ajustara a las observaciones futuras. Por esta razón la estimación del modelo se realizará desde 1993 hasta el primer semestre de 2006, comparando los valores observados con las predicciones del modelo. Si los resultados obtenidos en el periodo de validación se consideran satisfactorios, es de esperar que también los sean los obtenidos al predecir la facturación correspondiente al período 2007-2009.

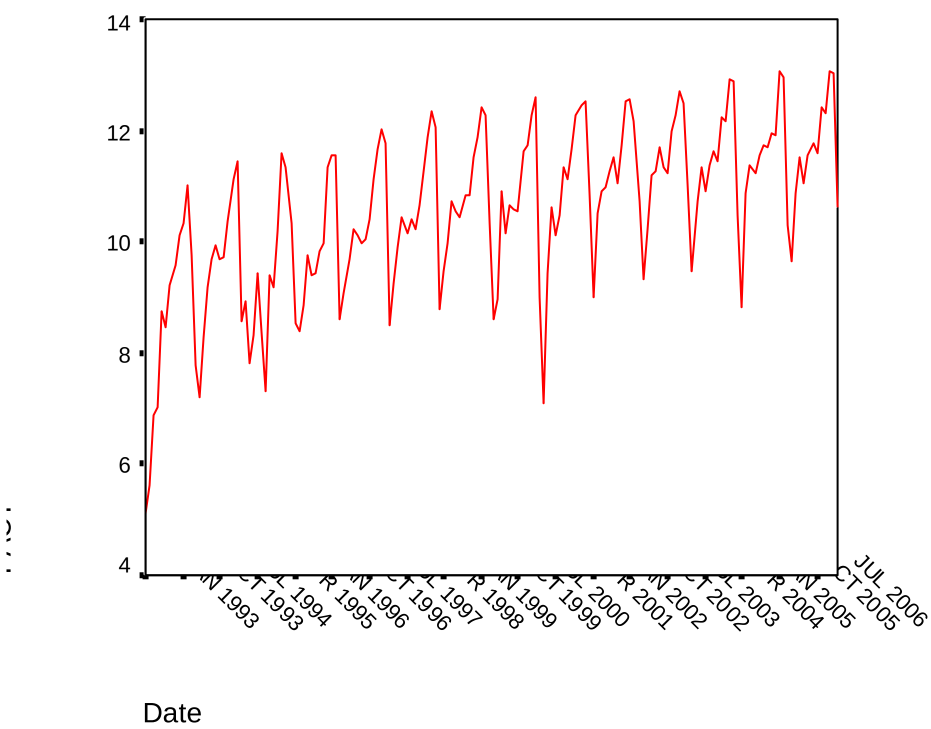
4.2.2.- Identificación del modelo adecuado a la serie de Tiempo del histórico de ventas.

Como se ha mostrado en el capítulo 3, la serie tiene una varianza muy alta y nos es constante, por eso la serie que se esta corriendo se transformo en el

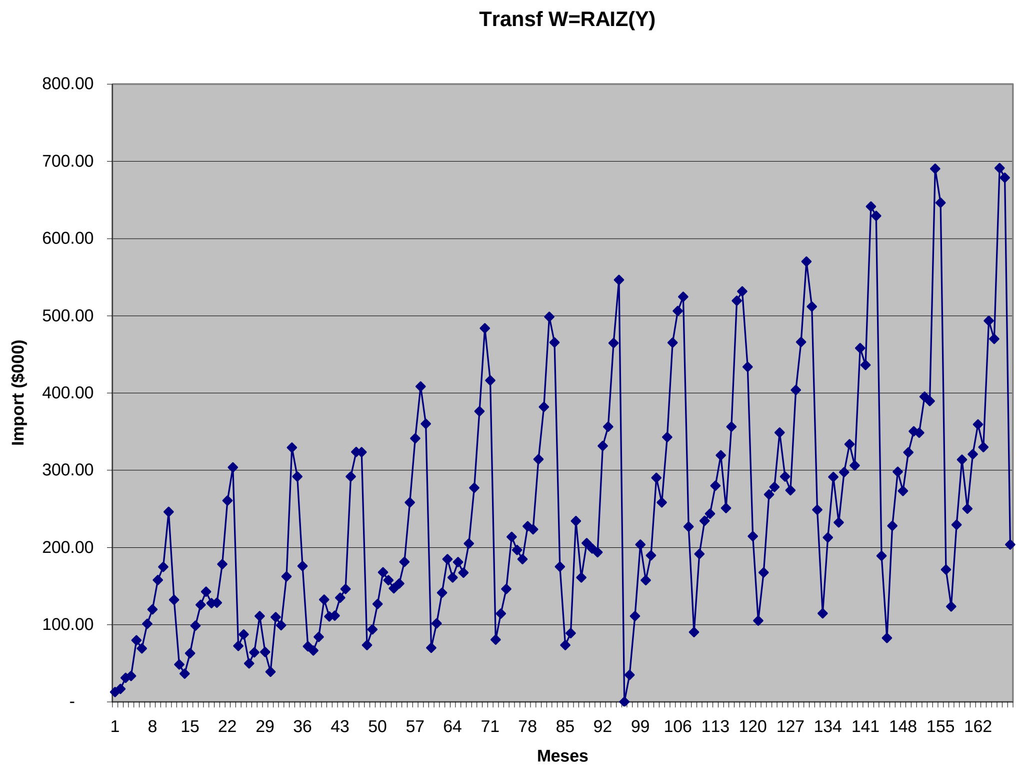
$W_t = LN(Y_t)$ ó $W_t = \sqrt{Y_t}$

Esta transformación puede estabilizar la varianza, y proporcionar normalidad.

Presentamos nuevamente la gráfica que representa la serie de Facturación Mensual con la serie. $W_t = LN(Y_t)$



Presentamos ahora la grafica de la transformación $W_t = \sqrt{Y_t}$



Esta transformación al correr el modelo, es mucho mejor que la logarítmica.

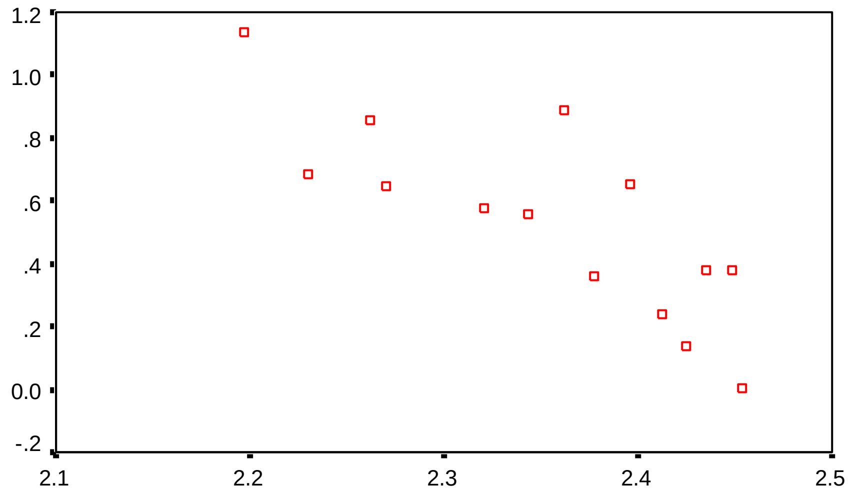
Para detectar si la varianza de la serie ***Fact*** es constante, si no cambia con el tiempo, aplicaremos la prueba de Levene para la homogenidad de varianzas.

4.3 Análisis de la estabilidad en varianzas y transformaciones para conseguir homocedasticidad.

Dado que el periodo de ajuste se tienen años completos desde 1993-2005 tenemos 13 grupos con 12 observaciones cada uno de ellos. La prueba permitirá contrastar la hipótesis nula de que no existen diferencias significativas entre las varianzas de la serie ***Fact*** en los 13 grupos definidos.

El cuadro siguiente proporciona el gráfico de nivel y dispersión y la Prueba de Levene

Spread vs. Level Plot of FACT By YEAR_



Level

* Plot of LN of Spread vs LN of Level

Slope = -2.959 Power for transformation = 3.959

El gráfico de nivel y dispersión representa a cada uno de los 13 años, el Logaritmo neperiano del rango intercuartílico de la serie ***Fac.*** (como medida de la dispersión de las observaciones), frente al logaritmo natural de la mediana (como medida de tendencia central).

Test of Homogeneity of Variance

		Levene Statistic	df1	Df2	Sig.
FACT	**Based on Mean**	**1.154**	**13**	**154**	**.318**
	Based on Median	**.958**	**13**	**154**	**.496**
	Based on Median and with adjusted df	**.958**	**13**	**125.374**	**.497**
	Based on trimmed mean	**1.118**	**13**	**154**	**.348**

La pendiente de la recta de regresión para la nube de puntos representada es muy negativa, lo que implica que la tendencia de la varianza esta disminuyendo, lo cual no es aspecto suficiente para afirmar que es constante. La confirmación de la estabilidad en varianzas nos la proporciona el (p-valor) asociado al estadístico de Levene esta muy arriba de 0.05 de nivel de significación o contraste (Ver en Anexo 1, definición de Estadístico Levene). En consecuencia seguiremos con la serie.

4.4 Análisis de la estabilidad en medias y de la estacionalidad.

Puede suceder que la serie observada presente un patrón de comportamiento creciente o decreciente. En dicho caso, la serie no sería estable en medias. Si el gráfico de la serie la presencia de tendencia es creciente (decreciente), no fuera evidente se recomienda como alternativa representar las ACF estimadas. La Función de Autocorrelación Simple de orden k, como lo hemos definido se obtiene como la correlación entre la serie observada $X_t = (x_1, x_2, \ldots, x_n)$ y la serie retardada *k* veces, $Y_t = (y_1, y_2, \ldots, y_{n-k})$ donde $y_1 = x_{k+1}, \ldots, y_{n-k} = x_n$.

Si la serie observada presentara tendencia creciente o decreciente, su ACF presentaría estructura positiva con decrecimiento lento hacia cero. La tendencia puede estabilizarse aplicando sucesivas diferencias regulares:

Es decir:

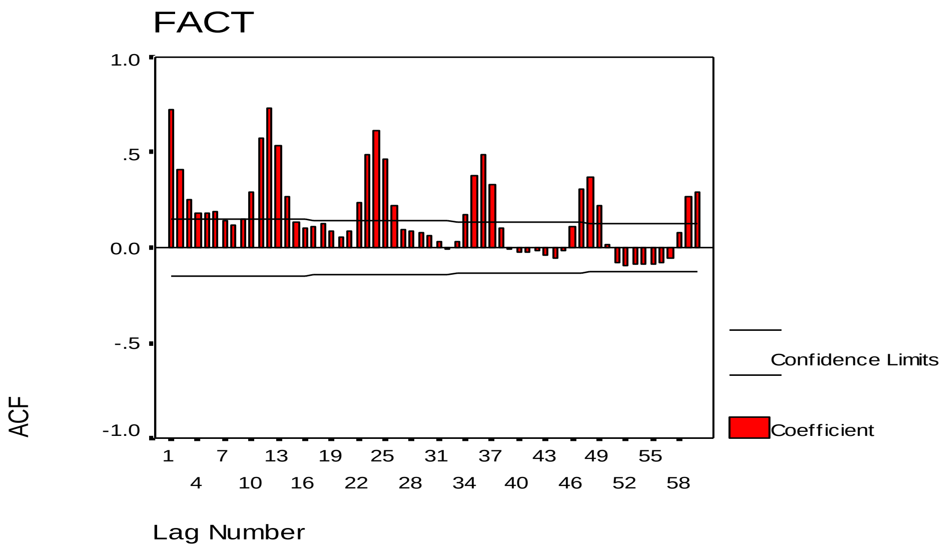
$W_t = \nabla^d Y_t = (1 - B)^d Y_t$, hasta que la serie obtenida W_t presente media constante.

Un caso particular de no estacionariedad presente en muchas series es la estacionalidad, que se entiende como una pauta regular de comportamiento periódico en la serie. Si en el gráfico de la serie la presencia de estacionalidad no es evidente, la alternativa es presentar las ACF estimadas. Si la serie presentara estacionalidad de periodo *s* , su ACF presentaría coeficientes altos con decremento en los retardos s, 2s, 3s…La estacionalidad puede eliminarse aplicando sucesivamente diferencias estacionales de período s:

$$W_t = \nabla_s^D Y_t = (1 - B^s)^D Y_t$$

Hasta, que la serie obtenida, W_t , no presente estacionalidad.

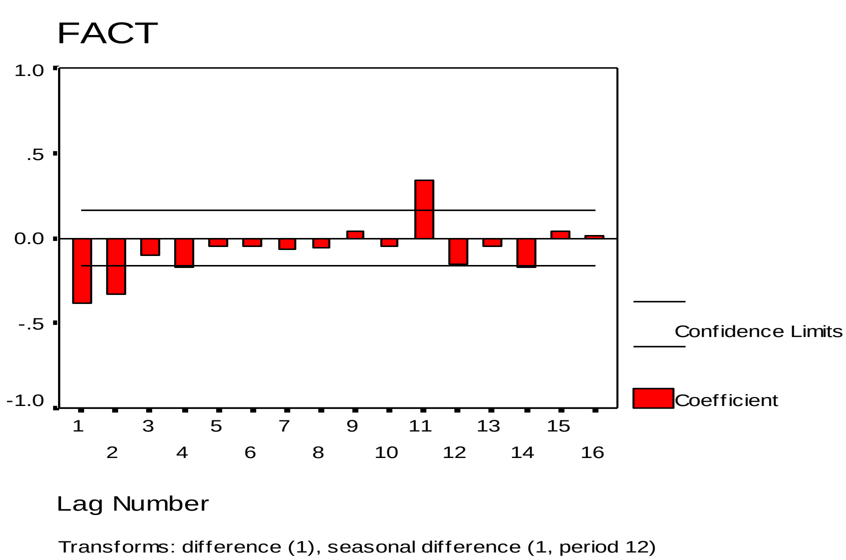
Para detectar que la variable ***Fact*** presenta tenencia creciente o decreciente, calcularemos sus ACF, sobre los primeros 60 retardos.

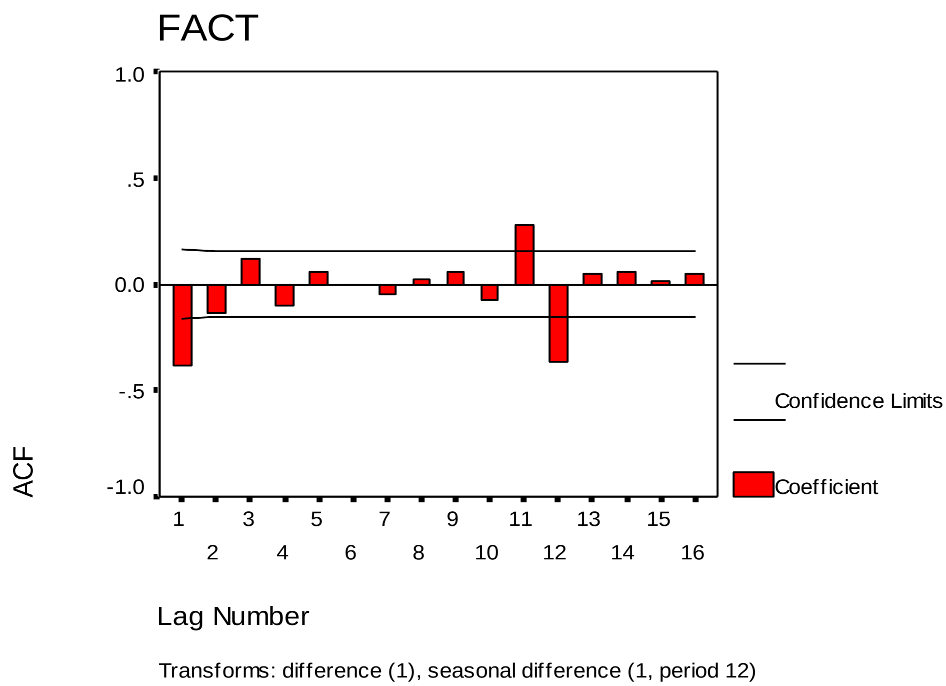


Obsérvese que la representación gráfica incluye una banda centrada en el cero. Los límites de dicha banda corresponden a los límites de los intervalos de confianza al 95% para el cero. Como la serie presenta tendencia creciente o decreciente , la ACF presenta estructura positiva con decrecimiento lento hacia cero que se detecta en los retardos de los múltiplos de 12 (12,24,36,48, etc), lo que corrobora la estacionalidad del periodo 12 que mostraba el gráfico de la serie.

Como la serie presenta tendencia será necesario diferenciarla regularmente, y por lo tanto el parámetro de *d* del modelo ARIMA, deberá también de ser al menos 1, además de al menos una diferencia estacional para poderla eliminar la estacionalidad.

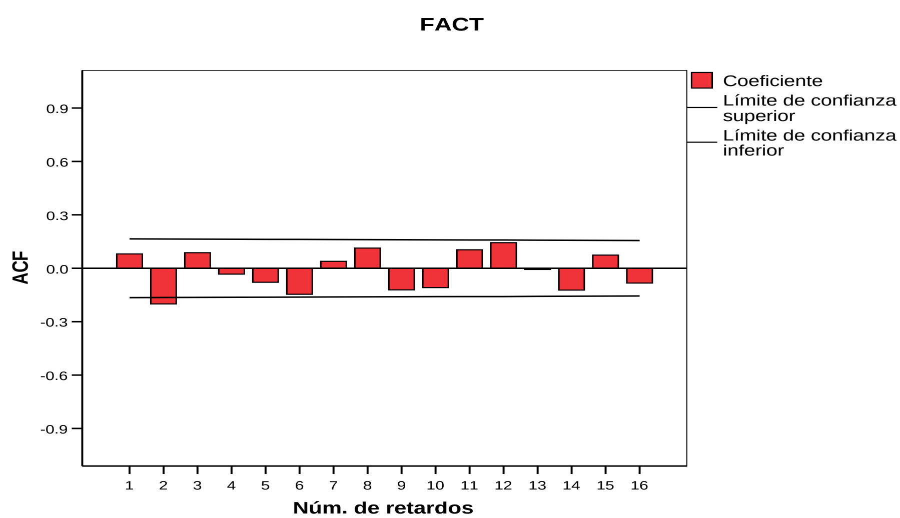
Ahora presentaremos las ACF y las PACF sobre los primeros 16 retardos, después de tomar una diferencia regular y una diferencia estacional sobre la serie ***Fact***

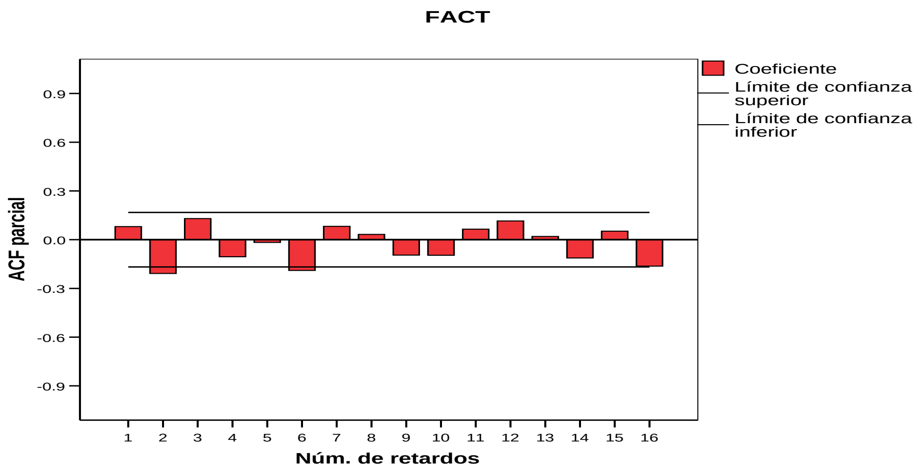




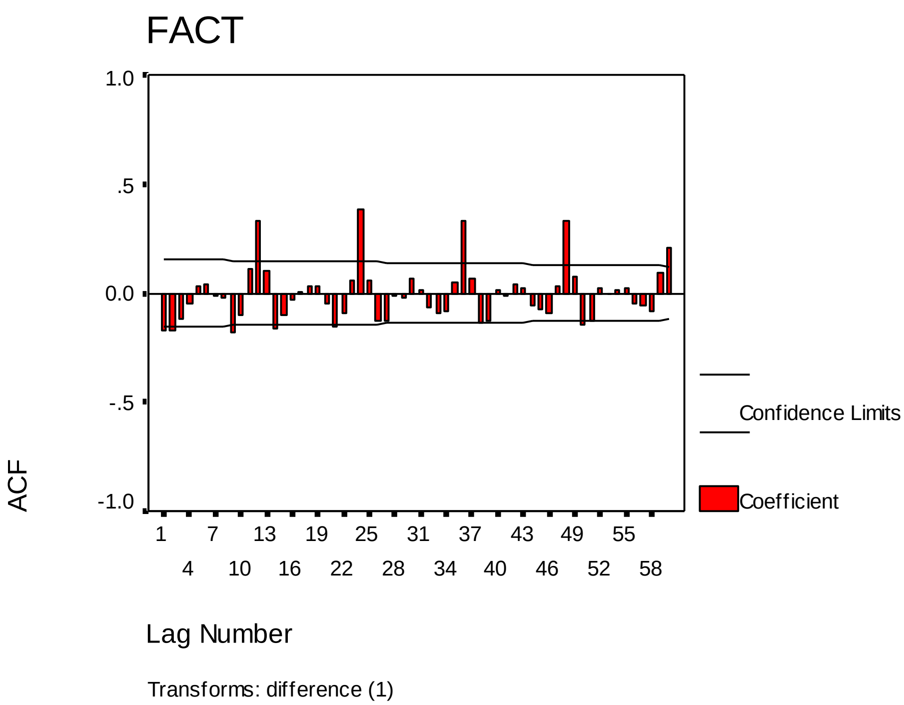
Como se observa la serie con una diferencia regular presenta valores muy negativos, sólo hay un valor posito hasta el retraso 12.

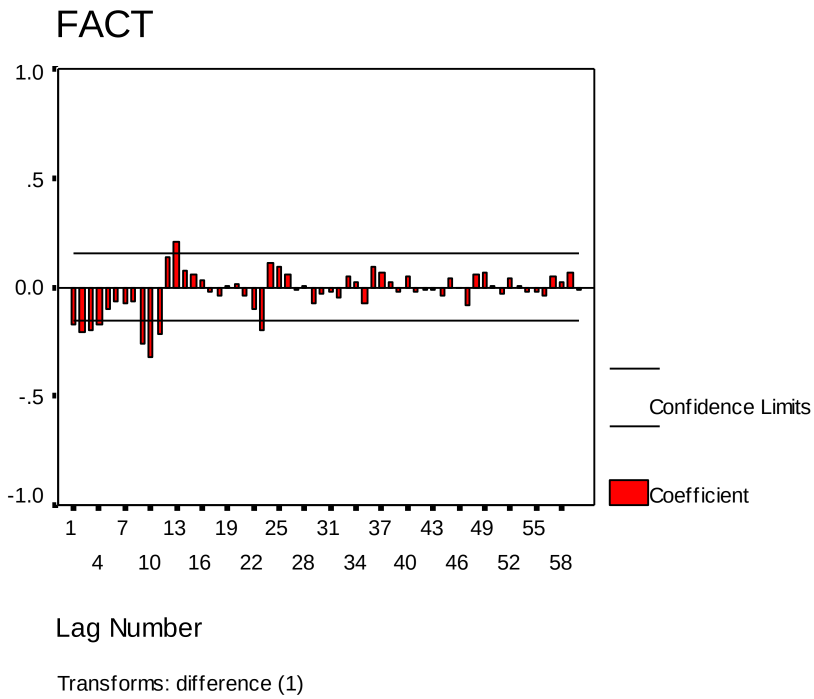
Se obtienen nuevamente las ACF y la PACF, con el parámetro *d=0*, y D=1 y se obtiene:





Ahora si consideramos sólo hacer una diferencia regular las ACF y las PACF, serán las siguientes:





Conclusiones hasta ahora:

Las observaciones que tengo al momento, son las siguientes:

- La serie muestra que necesita una diferencia ordinaria, ya que presenta muchas variaciones año contra año.

- La serie se muestra como estacionaria, con s=12, que es la variación estacional, por lo que necesita una diferencia.

- Según la teoría hasta ahora expuesta se requiere probar un modelo, donde se aplique una diferencia estacional s=12 para quitar picos y suavizar la curva.

- Se presume se requiere un modelo SARIMA(P,D,Q), el fue definido en la anterior sección, ya que por la naturaleza del fenómeno del juguete, existen elementos autoregresivos y estacionales importantes.

- Si el modelo no presentara esta estacionalidad, un posible candidato sería un AR(1), pero según considero adecuado será un modelo multiplicativo con d=1, y D=1, ahora determinaremos los otros parámetros.

- Por experiencia el parámetro de media móvil estacional Q=1 será quizá necesario por los cambios de facturación que se tienen en los meses, por eventos de películas o lanzamientos.

4.5 Estimación de los parámetros del histórico de ventas por mes

4.5.1 Determinación de los órdenes p, d, q, P, D Y Q del modelo ARIMA estacional (del histórico de ventas por mes)

- Como lo establece la metodología los órdenes de p, d, q, P, D Y Q se determinan a partir de la estructura de las ACF y PACF, de la serie w_t . Las consideraciones que se toman en la práctica son las siguientes:

- Si la serie presenta media constante y no es estacional, los parámetros d, D, P y Q son iguales a 0 y, en consecuencia, el modelo se ajusta a un ARIMA(p,q). Los parámetros p y q se estiman a partir de las ACF y las PACF de la propia serie observada.

- Como hemos analizado que la serie ***Fact*** es estable en varianzas y que con una diferencia (d=1), ya no presenta tendencia creciente ni decreciente. Además al presentar estacionalidad en el período s=12, fue necesario aplicar un diferencia estacional para eliminarla (D=1). En consecuencia el modelo que se ajustará será un

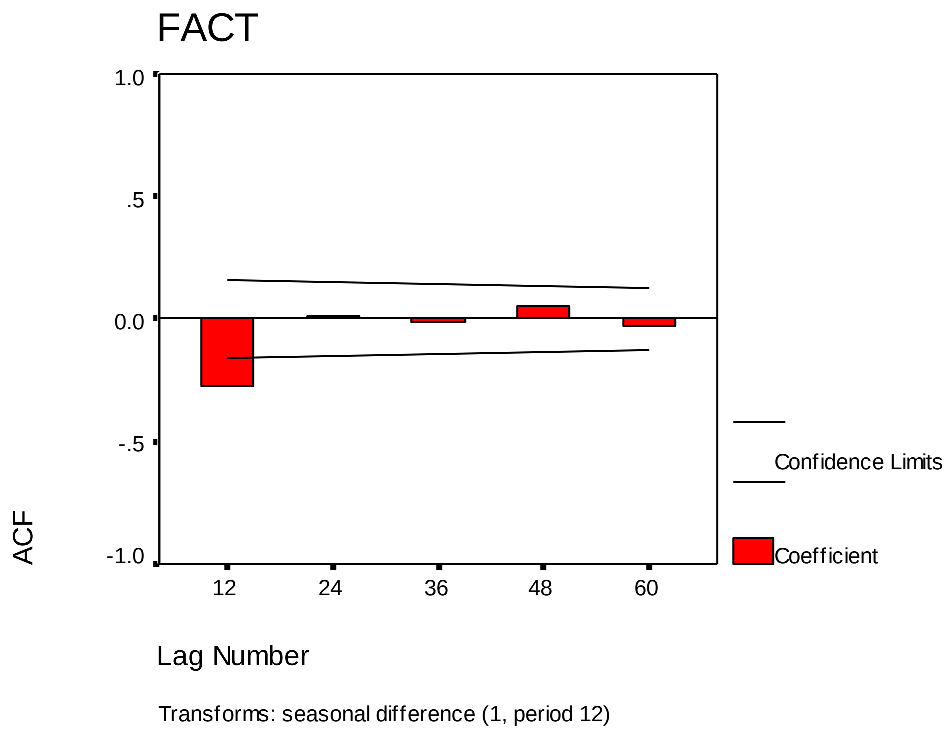
$$ARIMA(p,1,q)(P,1,Q)_{12}$$

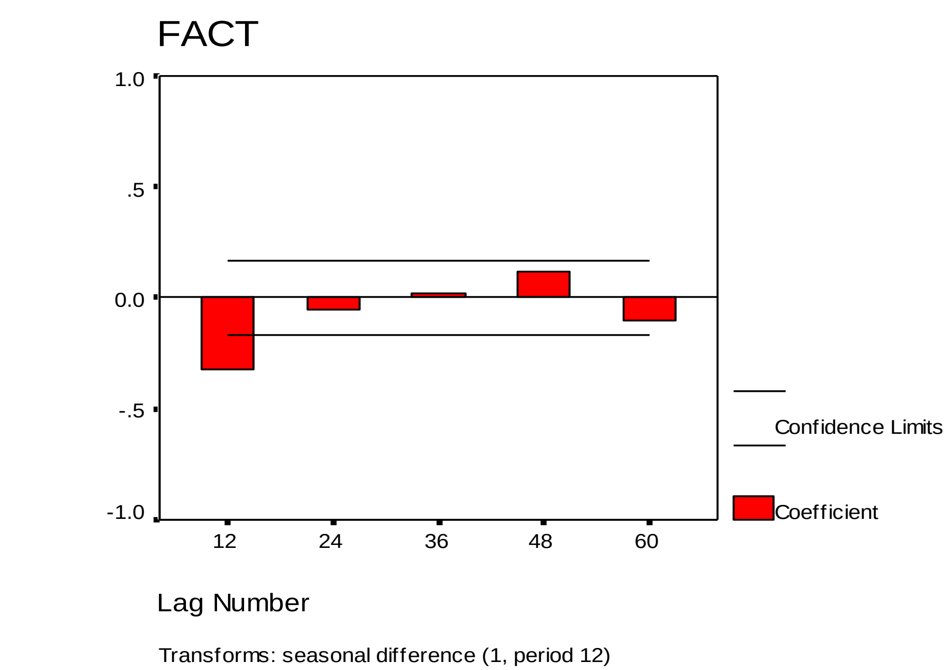
- La identificación de los parámetros p y q, órdenes de los polinomios autoregresivos y de medias móviles de la parte regular del modelo, se realizará a partir de las ACF y PACF para la serie diferenciada estacionalmente, sobre los retardos.

- Si sólo hacemos una diferencia estacional, se observa que en las ACF, los primeros 3 coeficientes son no nulos y van decreciendo con el retardo. Las PACF el único significativamente distinto de cero es el primero. Esto indicaría que como primera tentativa será entonces suponer que a serie presenta, en la parte regular, una estructura de un modelo autorregresivo de orden *p*=1.

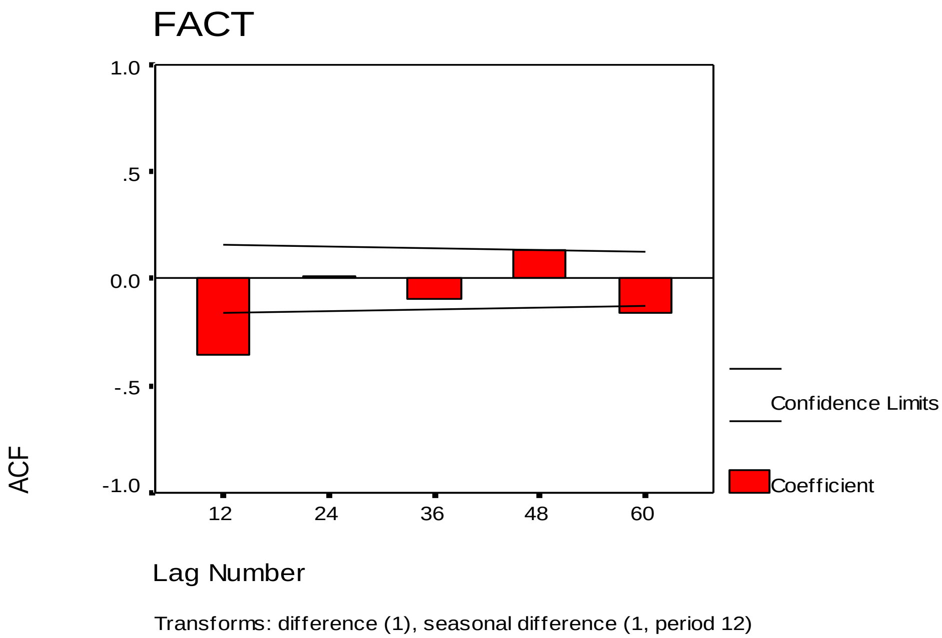
- Si sólo hacemos una diferencia estacional y una regular, se observa que en las ACF y las PACF, los coeficientes no son significativos salvo el primero y van decreciendo con el retardo. Existe en ambas, un pico positivo en el final del año 2000, y principios del 2001. Entonces esto hace suponer que a serie presenta, en la parte regular, una estructura de un modelo autorregresivo de orden *p*=2.

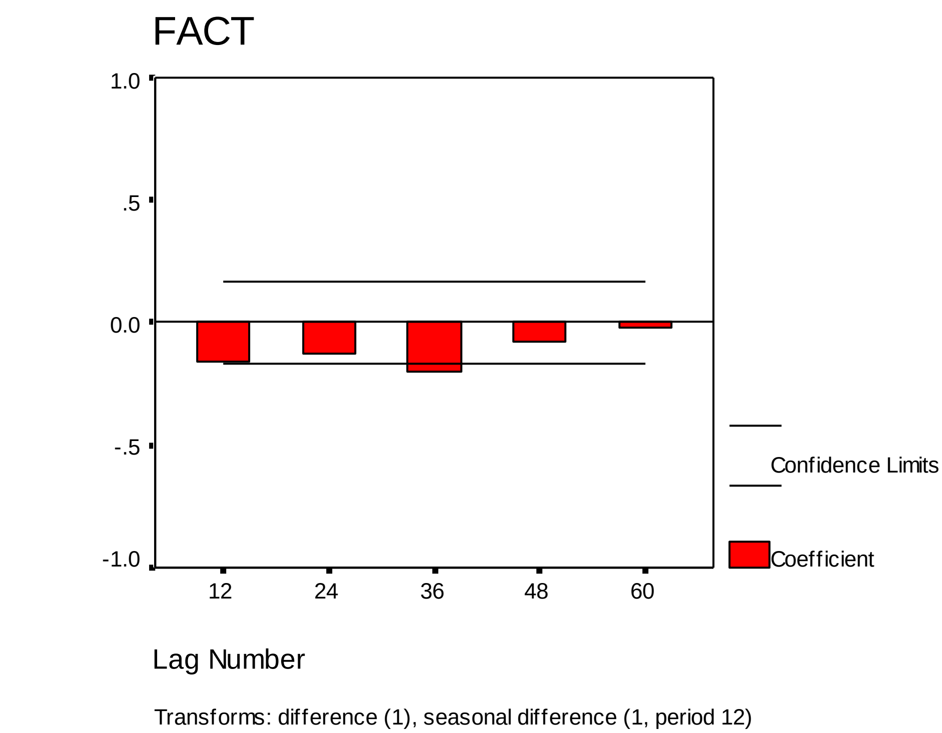
- La identificación de los parámetros *P* y *Q*, órdenes de autoregresión y de medias móviles de la parte estacional del modelo, se realizará a partir de las ACF y PACF para la serie diferenciada estacionalmente, considerando exclusivamente los retardos estacionales. El siguiente cuadro muestra los primeros 5 retardos estacionales de la serie ***Fact.***





Tanto en ACF como en las PACF, los coeficientes están dentro de lo límites de la banda de confianza para el cero (excepto el primero), por lo que considero que no hay ninguna evidencia fuerte para pensar en que persista una estructura estacional. Sin embargo se presenta la duda sobre si es necesaria una diferencia regular, ya que si calculamos los 5 primeros retardos con una diferencia estacional y una regular obtenemos casi los mismos valores de las ACF y las PACF, como se muestran es las siguientes gráficas:





Por esto consideraremos de inicio dos modelos:

El $ARIMA(1,0,0)(0,1,0)_{12}$ y el $ARIMA(1,1,0)(0,1,0)_{12}$

4.6 Examen y Ajuste del modelo

ARIMA: estimación de los parámetros

Una vez determinados los órdenes de p, d, q, P, D Y Q de un proceso ARIMA(p, d, q)(P,D,Q $)_s$, con periodo conocido *s*, se trata de , a partir de una realización del proceso u observación de la serie $\{w_1, w_2, \ldots, w_n\}$, estimar los parámetros del modelo

$$\varphi(B)\Phi(B^s)W_t = \theta(B)\Theta(B)e_t$$

Es decir, se trata de estimar los parámetros:

$$\beta_i = \varphi_i; i = 1,..,p$$
$$\beta_i = \Phi_{i-p}; i = p+1,..,P+p$$
$$\beta_i = \theta_{i-(p+P)}; i = P+p,,..,p+P+q$$
$$\beta_i = \Theta_{i-(p+P+q)}; i = P+p+q+1,..,p+P+q+Q$$
$$\beta_i = c; i = 0$$

Si B_i es la estimación β_i, la primera etapa en la validación del modelo consistirá en comprobar si los coeficientes B_i son significativamente distintos de cero. Para ello, sobre cada parámetro, se planteará la hipótesis nula:

$$H_0 = \beta_i = 0$$

Dicha hipótesis puede ser interpretada como que la variable asociada al parámetro β_i no mejora el ajuste con respecto al obtenido con las restantes variables incluidas en el modelo. Si el *p*-valor asociado al valor des estadístico de contraste *t* es menor que α, se rechazará la hipótesis nula al nivel de significación α. Si denotamos F_t a la observación de la serie FACT en el instante ***t***, la expresión general del modelo $ARIMA(1,0,0)(0,1,0)_{12}$ identificado es:

$$\varphi_1(B)(1 - B^{12})F_t = e_t$$

ó

$$(1 - \varphi_1 B)(1 - B^{12})F_t = e_t$$

$$BF_t = F_{t-1}, \ldots, B^{12}F_t = F_{t-12}$$

y $\{e_t\}$ es un proceso de ruido blanco. Identificado el modelo, el siguiente paso será ajustarlo sobre la serie observada, procediendo a la estimación de los parámetros φ_1 y e_t.

El siguiente cuadro proporciona las estimaciones de los parámetros del modelo

$$ARIMA(1,0,0)(0,1,0)_{12}$$

Modelo 1

Model Description

			Model Type
Model ID	fact	Model_1	ARIMA(1,0,0)(0,1,0)

Model Summary

Parámetros del modelo ARIMA

					Estimación	ET	t	Sig.
FACT-Modelo_1	FACT	Log natural	Constante		.254	.084	3.022	.003
			AR	Retardo 1	.366	.079	4.654	.000
			Diferenciación estacional		1			

Fit statistic	**Mean 5**
Stationary R-squared	0.097
R-squared	0.678
RMSE	54,016.49
MAPE	72.16
MAXAPE	1,090.87
MAE	31,902.90
MAXAE	222,919.66
Normalizad BIC	21.863

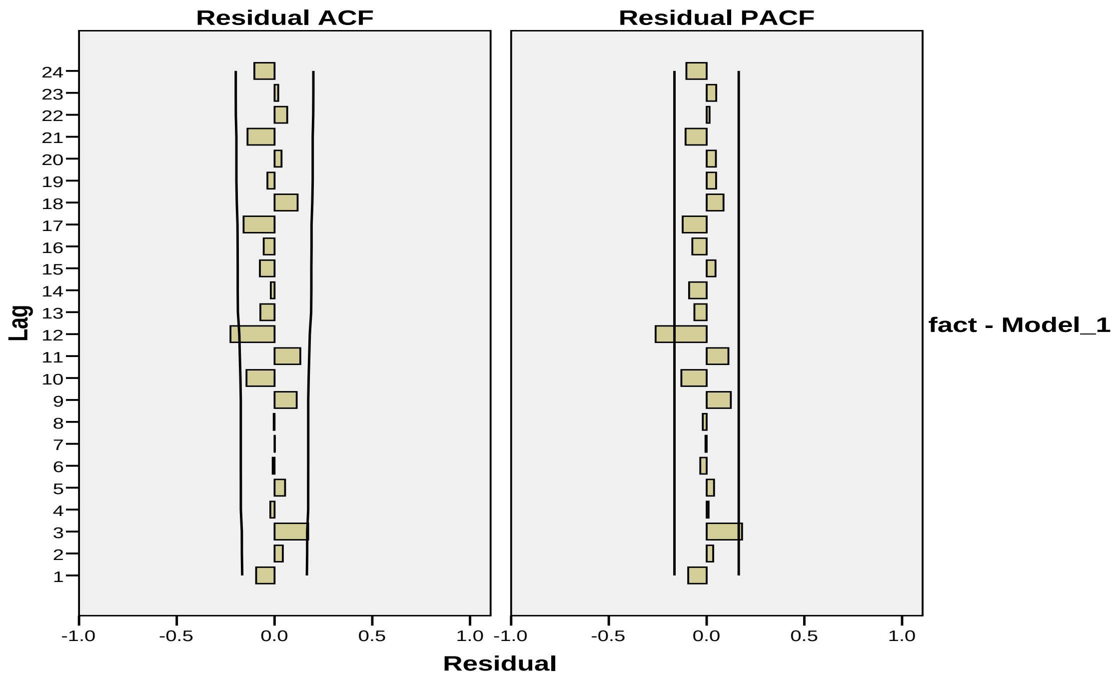
Ljung-Box Q(18)

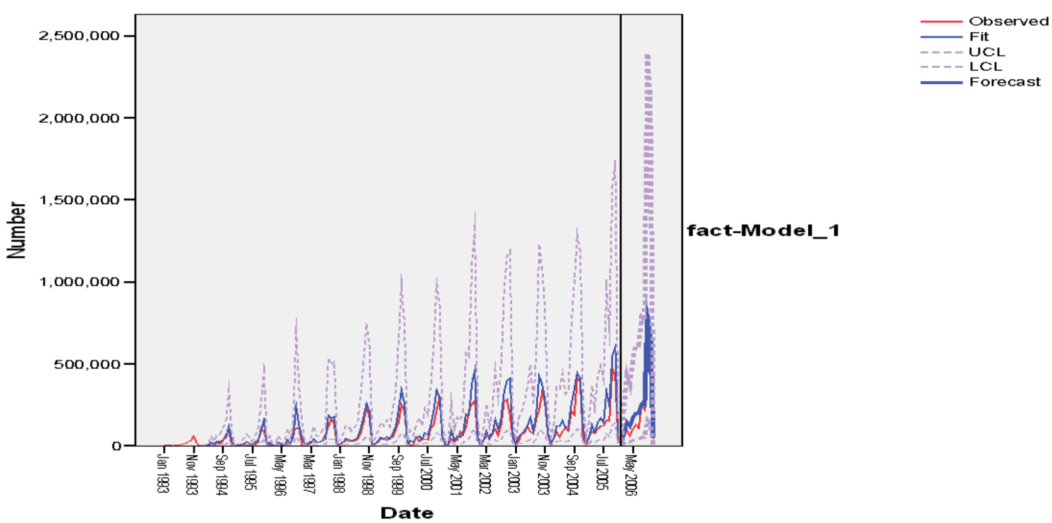
Statistic	DF	Sig
30.966	17.00	0.020

Forecast

Model		Jan 2006	Feb 2006	Mar 2006	Apr 2006	May 2006	Jun 2006	Jul 2006	Aug 2006	Sep 2006	Oct 2006	Nov 2006	Dec 2006
Fac.-Model_1	Forecast	9176.88	79508.70	141831.7	120689.8	169812.3	200148.7	198306.4	254971.7	247693.8	778539.3	682229.0	47924.76
	UCL	26509.81	242614.1	435816.7	371197.4	522345.4	615670.7	610005.2	784312.0	761924.7	2394845	2098587	147420.1
	LCL	2108.46	16368.02	28781.76	24444.43	34384.80	40526.12	40152.91	51626.41	50152.79	157637.8	138137.0	9703.76

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.





Al final del bloque de resultados encontramos otros parámetros cuyos valores, son las estimaciones o predicciones mediante el modelo, los residuos, los límites inferior y superior del intervalo de confianza al 95% para las predicciones y los errores típicos de las predicciones.

Observamos que el Coeficiente de Autocorrelación explica el 68% del comportamiento del modelo, lo cual si se hace la transformación por medio de la raíz cuadrada se obtiene un mejor valor, como lo muestra la siguiente corrida:

Modelo 1.2

Model Description

			Model Type
Model ID	fact	Model_1	ARIMA(1,0,0)(0,1,0)

Parámetros del modelo ARIMA

					Estimación	ET	t	Sig.
FACT-Modelo_1	FACT	Raíz cuadrada	Constante		20.871	4.307	4.846	.000
			AR	Retardo 1	.150	.085	1.775	.078
			Diferenciación estacional		1			

Fit statistic	Mean 5
Stationary R-squared	**0.023**
R-squared	**0.918**
RMSE	**27,213.168**
MAPE	**61.267**
MAXAPE	**1,538.143**
MAE	**17,486.840**
MAXAE	**111,817.055**
Normalizad BIC	**20.492**

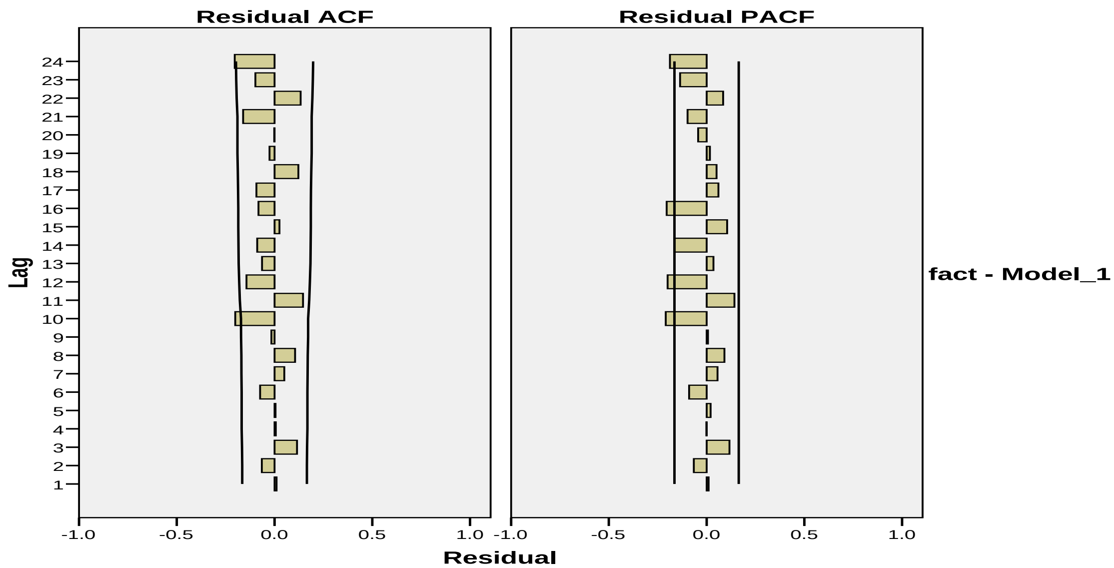
Ljung-Box Q(18)

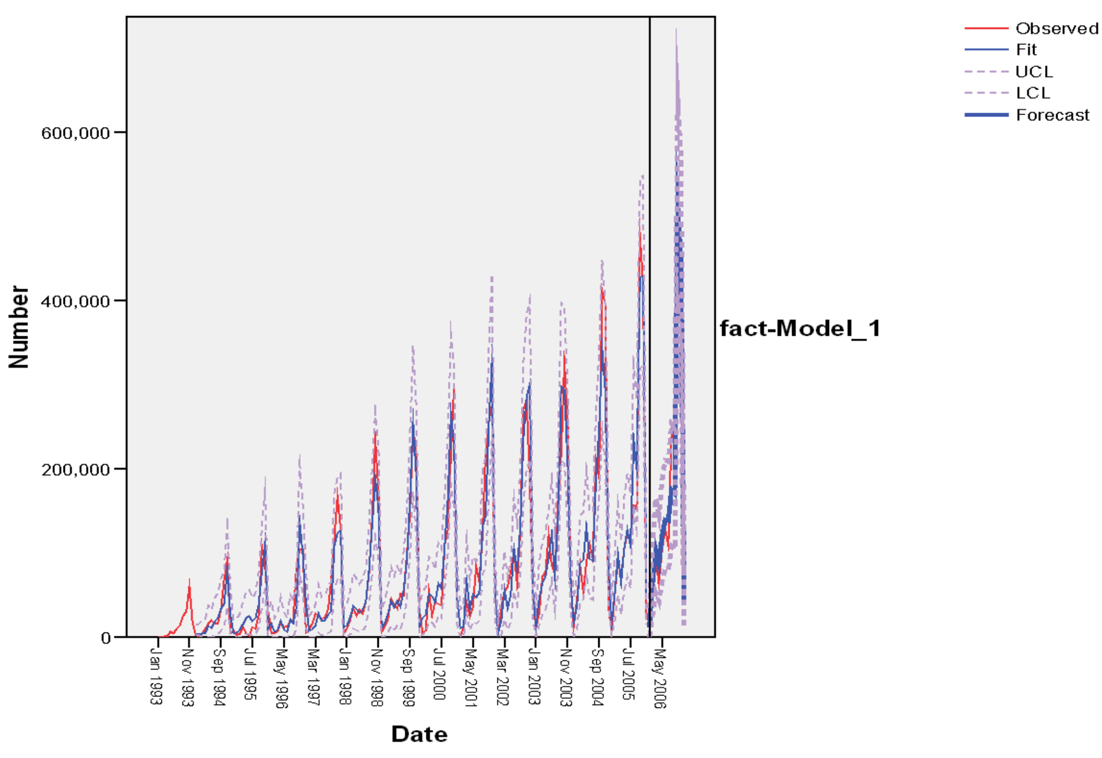
Statistic	**DF**	**Sig**
25.272	17	.089

Forecast

Model		Jan 2006	Feb 2006	Mar 2006	Apr 2006	May 2006	Jun 2006	Jul 2006	Aug 2006	Sep 2006	Oct 2006	Nov 2006	Dec 2006
Fac.-Model_1	Forecast	11477.52	63407.56	103573.4	88263.88	120130.7	139665.5	138380.0	174986.9	170279.2	507791.3	446988.8	38880.97
	UCL	33969.30	112479.3	165064.4	145380.8	185996.1	210313.8	208724.5	253489.9	247785.0	637970.7	569447.2	78210.89
	LCL	126.95	25728.14	53480.42	42545.05	65663.44	80415.29	79433.71	107882.0	104171.6	389010.1	335928.5	10949.16

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.


Residual ACF
Residual PACF
Lag
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
-1.0 -0.5 0.0 0.5 1.0
-1.0 -0.5 0.0 0.5 1.0
Residual
fact - Model_1


Observed
Fit
UCL
LCL
Forecast
Number
600,000
400,000
200,000
0
fact-Model_1
Jan 1993
Nov 1993
Sep 1994
Jul 1995
May 1996
Mar 1997
Jan 1998
Nov 1998
Sep 1999
Jul 2000
May 2001
Mar 2002
Jan 2003
Nov 2003
Sep 2004
Jul 2005
May 2006
Date

El modelo expuesto anteriormente es un pronóstico que le falta quizá una diferencia ordinaria, y tiene un bajo coeficiente de correlación estacionario, por lo que correremos un modelo como el anterior, pero adicionándole la diferencia ordinaria.

El cuadro que sigue muestra las estimaciones de los parámetros del modelo 2

$$ARIMA(1,1,0)(0,1,0)_{12}$$

Modelo 2

Model Description

			Model Type
Model ID	Fac.	Model_1	ARIMA(1,1,0)(0,1,0)

Parámetros del modelo ARIMA

					Estimación	ET	t	Sig.
FACT-Modelo_1	FACT	Raíz cuadrada	Constante		-.232	3.224	-.072	.943
			AR	Retardo 1	-.396	.080	-4.965	.000
			Diferencia		1			
			Diferenciación estacional		1			

Fit statistic	Mean 5
Stationary R-squared	0.157
R-squared	0.867
RMSE	34,751.636
MAPE	74.959
MAXAPE	1,541.493
MAE	21,997.684
MAXAE	129,459.430
Normalizad BIC	20.982

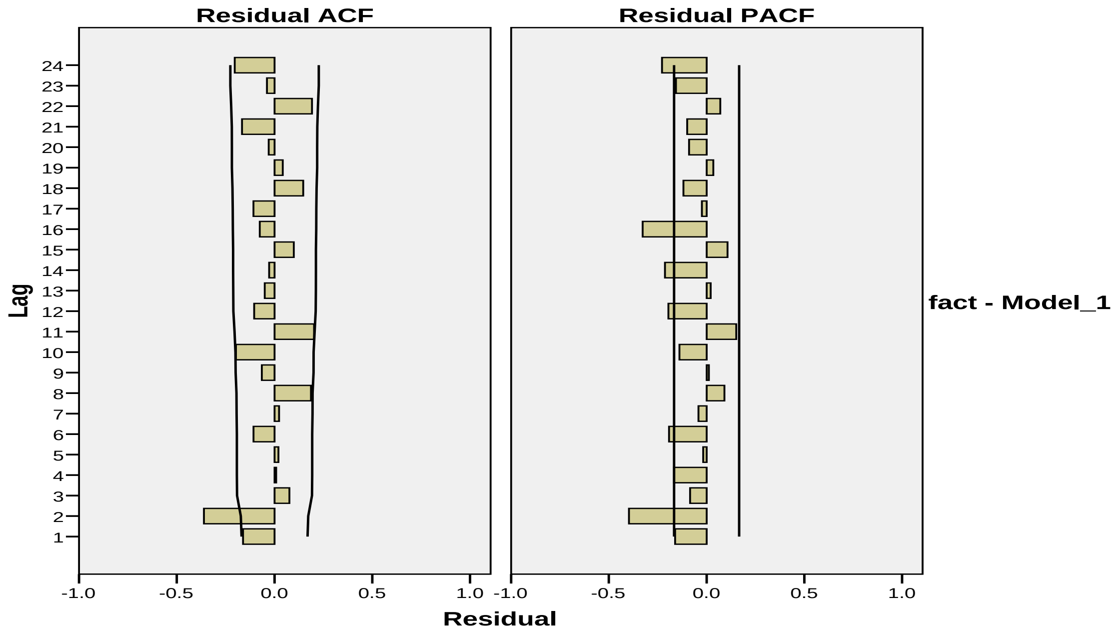
Ljung-Box Q(18)

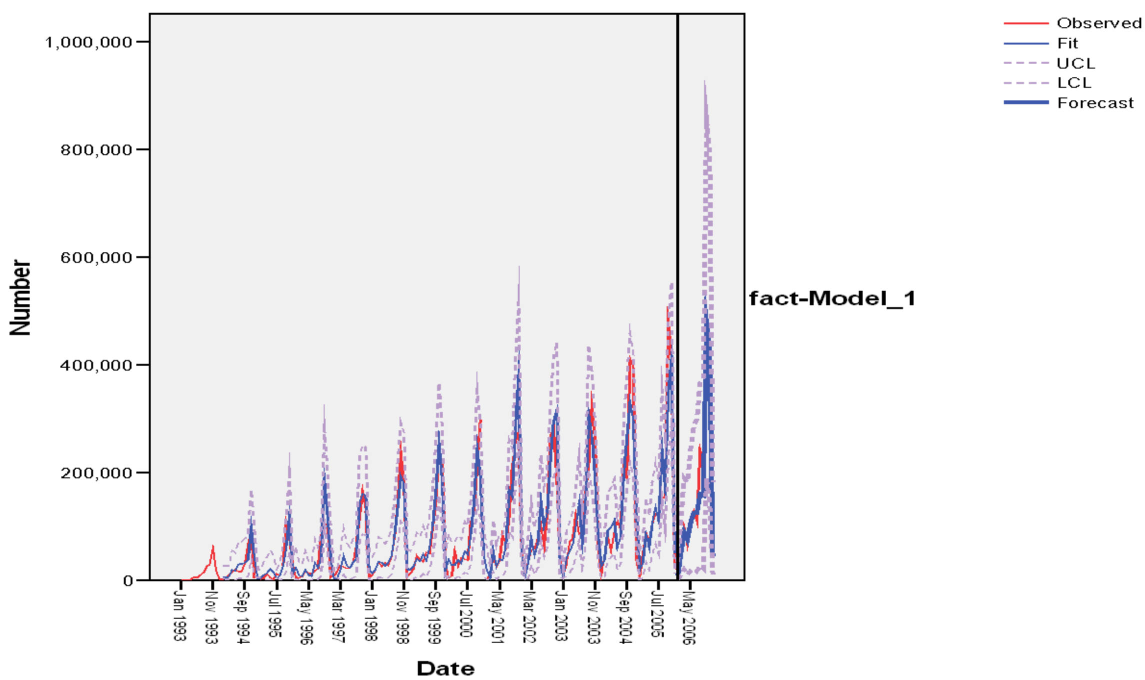
Statistic	DF	Sig
53.056	17	.000

Forecast

Model		Jan 2006	Feb 2006	Mar 2006	Apr 2006	May 2006	Jun 2006	Jul 2006	Aug 2006	Sep 2006	Oct 2006	Nov 2006	Dec 2006
Fac.-Model_1	Forecast	8932.51	51284.42	89495.07	76350.99	106732.3	125798.7	125930.0	160929.6	157831.5	477836.1	420918.7	43986.04
	UCL	33404.18	115469.9	189612.2	181896.6	243263.1	286237.4	299233.4	367696.3	375465.9	855881.7	795552.3	182693.4
	LCL	674.08	9217.74	20879.19	10214.70	18088.39	21496.31	17102.33	26942.27	21294.43	189200.1	144009.2	11316.35

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.


Residual ACF
Residual PACF
Lag
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
-1.0
-0.5
0.0
0.5
1.0
Residual
fact - Model_1


Observed
Fit
UCL
LCL
Forecast
1,000,000
800,000
600,000
400,000
200,000
0
Number
fact-Model_1
Jan 1993
Nov 1993
Sep 1994
Jul 1995
May 1996
Mar 1997
Jan 1998
Nov 1998
Sep 1999
Jul 2000
May 2001
Mar 2002
Jan 2003
Nov 2003
Sep 2004
Jul 2005
May 2006
Date

El Paquete estadístico tiene una opción de model expecter el cual nos puede ayudar a entender mejor el comportamiento de la Serie de Tiempo del histórico de ventas. SPSS (versión 15.0), nos recomienda un Winters´Additive, cuyos valores son los siguientes:

Modelo 3

Model Description

	Model Type
ID del modelo FACT Modelo_1	Winters´ Additive

Exponential Smoothing Model Parameters

Model			Estimate	SE	t	Sig.
fact-Model_1	No Transformation	Alpha (Level)	5.30E-007	.030	1.79E-005	1.000
		Gamma (Trend)	.015	853.389	1.71E-005	1.000
		Delta (Season)	1.000	.085	11.768	.000

Fit statistic	Mean 5
Stationary R-squared	0.438
R-squared	0.919
RMSE	26,678.246
MAPE	83.908
MAXAPE	2,026.971
MAE	17,444.297
MAXAE	122,057.595
Normalizad BIC	20.481

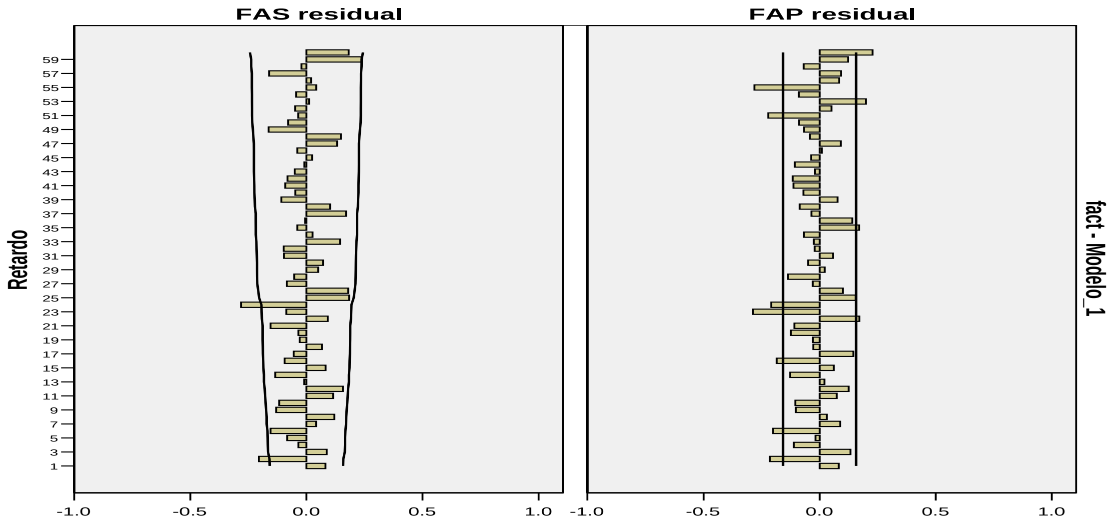
Ljung-Box Q(18)

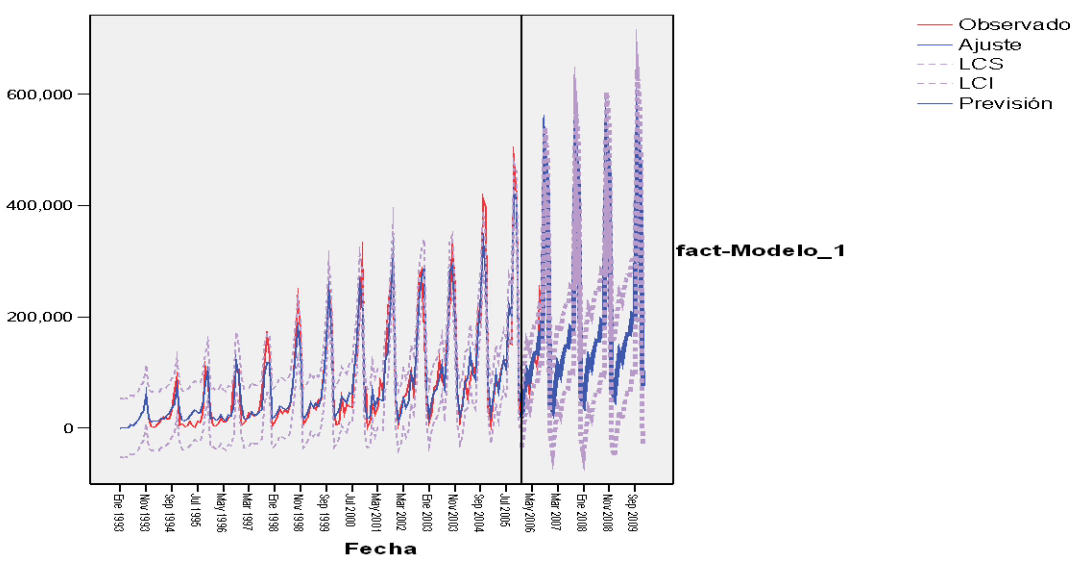
Statistic	DF	Sig
35.646	15	.002

Forecast

Model		Jan 2006	Feb 2006	Mar 2006	Apr 2006	May 2006	Jun 2006	Jul 2006	Aug 2006	Sep 2006	Oct 2006	Nov 2006	Dec 2006
Fac.-Model_1	Forecast	18669.96	63760.68	100648.5	86314.50	116086.2	134478.8	133265.0	167928.6	163458.4	488404.7	429445.0	41163.90
	UCL	71378.01	116468.7	153356.6	139022.5	168794.2	187186.8	185973.1	220636.7	216166.4	541112.8	482153.0	93871.94
	LCL	-34038.1	11052.63	47940.49	33606.45	63378.15	81770.74	80556.99	115220.6	110750.4	435696.7	376736.9	-11544.2

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.





Este modelo proporciona un excelente coeficiente de determinación y un buen coeficiente de correlación estacional, sin embargo, la estimación para los meses más significativos del 2006 que son octubre y noviembre están muy arriba de los reales, por esto se calcularán otros modelos que sean más exactos contra lo real, para esto se volverán a revisar otros modelos ARIMA, el que sigue no considera parámetro autoregresivo esto por aprendizaje de este modelo, por es decir el siguiente será:

Modelo 4

Model Description

			Model Type
ID del modelo	FACT	Modelo_1	ARIMA(0,0,0)(0,1,1)

ARIMA Model Parameters

					Estimate	SE	T	Sig.
fact-Model_1	fact	Square Root	Constant		21.015	2.906	7.231	.000
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	.227	.089	2.551	.012

Fit statistic	Mean 5
Stationary R-squared	0.030
R-squared	0.916
RMSE	27,569.161
MAPE	62.963
MAXAPE	1,559.416
MAE	17,763.462
MAXAE	117,994.549
Normalizad BIC	20.518

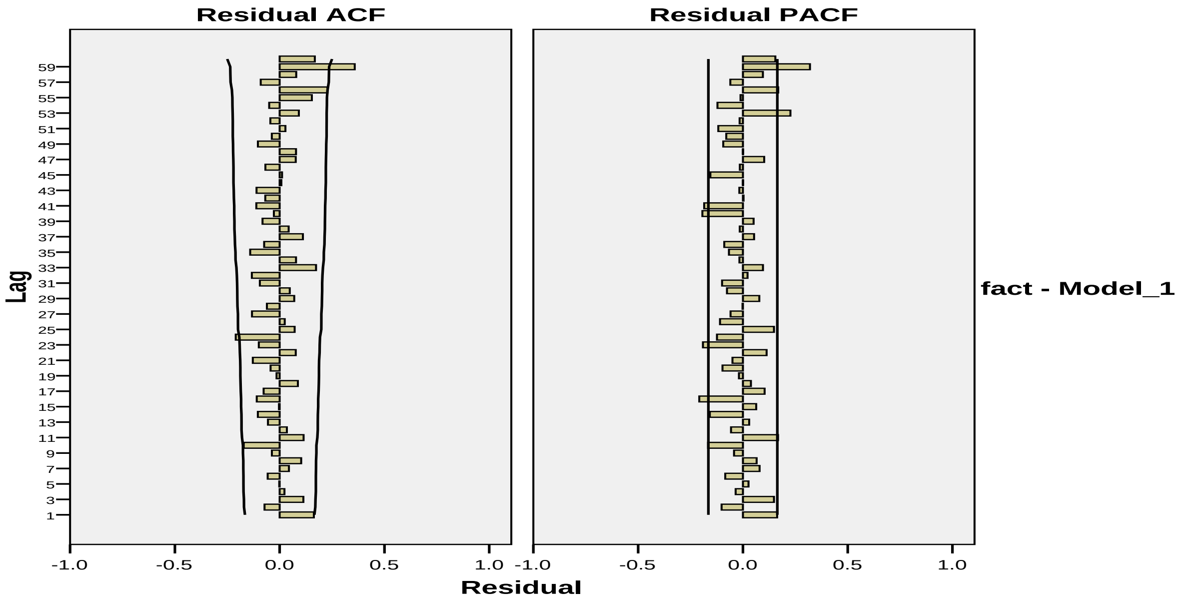
Ljung-Box Q(18)

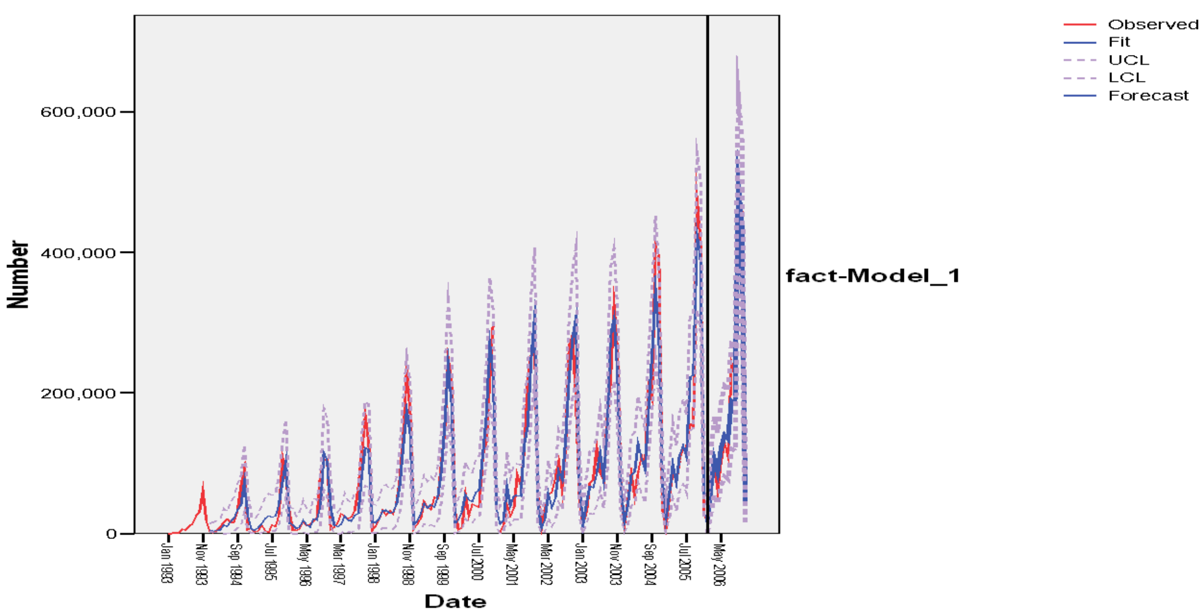
Statistic	DF	Sig
22.421	17	0.169

Forecast

Model		Jan 2006	Feb 2006	Mar 2006	Apr 2006	May 2006	Jun 2006	Jul 2006	Aug 2006	Sep 2006	Oct 2006	Nov 2006	Dec 2006
fact-Model_1	Forecast	15424.42	64093.14	105576.6	87513.15	121537.6	140013.9	134460.3	189514.4	186038.3	495070.8	441282.7	43984.31
	UCL	40958.01	112533.7	166514.0	143391.5	186588.4	209497.1	202644.1	269584.7	265414.7	621463.7	560893.4	84810.07
	LCL	915.12	26676.86	55663.58	42659.09	67511.05	81555.12	77300.77	120468.4	117686.2	379702.1	332696.2	14182.85

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.


Residual ACF
Residual PACF
Lag
Residual
fact - Model_1


Observed
Fit
UCL
LCL
Forecast
Number
Date
fact-Model_1

Este modelo tiene casi los mismos valores para la Correlación Cuadrada (alta 0.90), y la Correlación Estacionaria (0.030).

La estimación es muy acertada para los valores correspondientes al 2006, sobre todo para los meses más cargados de facturación (octubre y noviembre), sin embargo, buscaremos un modelo con una mejor correlación estacional, aumentado en algún otro parámetro, por lo que no podemos descartar de manera tajante un parámetro autoregresivo.

Modelo 5

Model Description

			Model Type
ID del modelo	FACT	Modelo_1	ARIMA(1,0,0)(0,1,1)

ARIMA Model Parameters

					Estimate	SE	t	Sig.
fact-Model_1	fact	Square Root	Constant		20.938	3.359	6.233	.000
			AR	Lag 1	.160	.086	1.872	.063
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	.240	.090	2.678	.008

Fit statistic	Mean 5
Stationary R-squared	0.055
R-squared	0.915
RMSE	27,860.908
MAPE	60.859
MAXAPE	1,307.940
MAE	17,644.704
MAXAE	119,120.076
Normalizad BIC	20.574

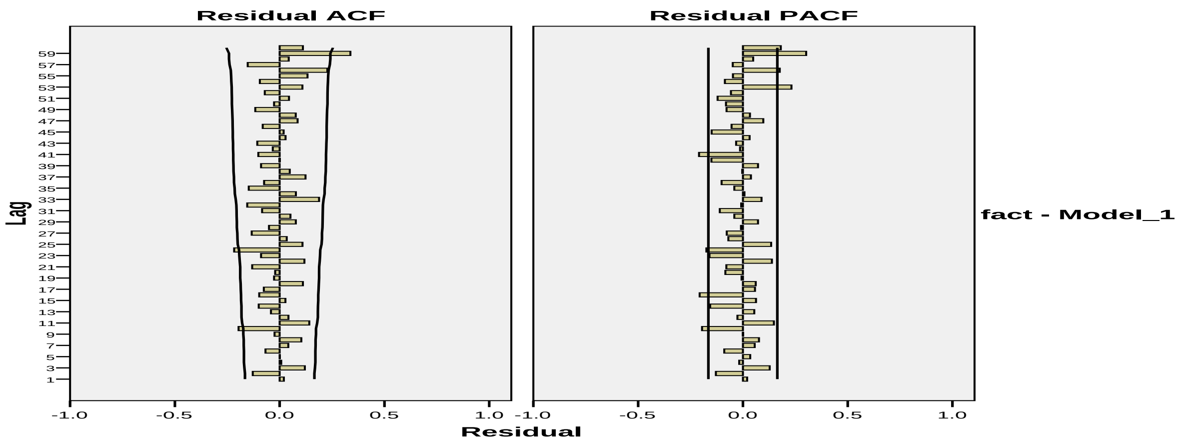
Ljung-Box Q(18)

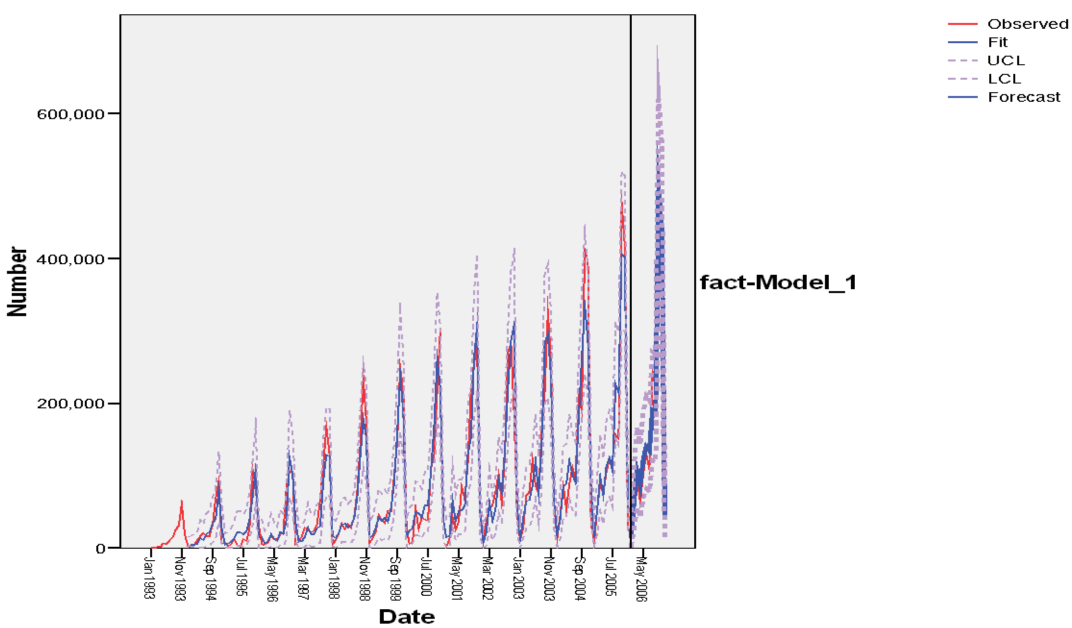
Statistic	DF	Sig
23.131	16	0.110

Forecast

Model		Jan 2006	Feb 2006	Mar 2006	Apr 2006	May 2006	Jun 2006	Jul 2006	Aug 2006	Sep 2006	Oct 2006	Nov 2006	Dec 2006
fact-Model_1	Forecast	13465.25	63311.42	105467.6	87494.68	121644.6	139961.7	134147.3	190182.5	187113.2	493925.5	440383.0	44369.98
	UCL	37231.90	111627.1	166578.8	143555.2	186938.8	209663.2	202483.3	270649.7	266967.2	620586.2	560264.0	85497.07
	LCL	507.22	26081.82	55449.64	42527.57	67443.85	81353.48	76904.61	120808.7	118352.6	378358.2	331595.4	14336.28

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.





Este modelo obtiene una estimación muy similar se observa que se requiere una diferencia ordinaria, ya que el modelo presenta en las ACF y las PACF una serie de picos, que una diferencia ordinaria ayudaría a disminuir.

Modelo 6

Model Description			
			Model Type
Model ID	**fact**	**Model_1**	ARIMA(1,1,0)(0,1,1)

ARIMA Model Parameters

					Estimate	SE	t	Sig.
fact-Model_1	fact	Square Root	Constant		-.169	2.763	-.061	.951
			AR	Lag 1	-.366	.083	-4.402	.000
			Difference		1			
			Seasonal Difference		1			
			MA, Seasonal	Lag 1	.177	.095	1.862	.065

Fit statistic	Mean 5
Stationary R-squared	0.169
R-squared	0.863
RMSE	35,340.196
MAPE	76.526
MAXAPE	1,930.888
MAE	21,989.540
MAXAE	143,840.292
Normalizad BIC	21.050

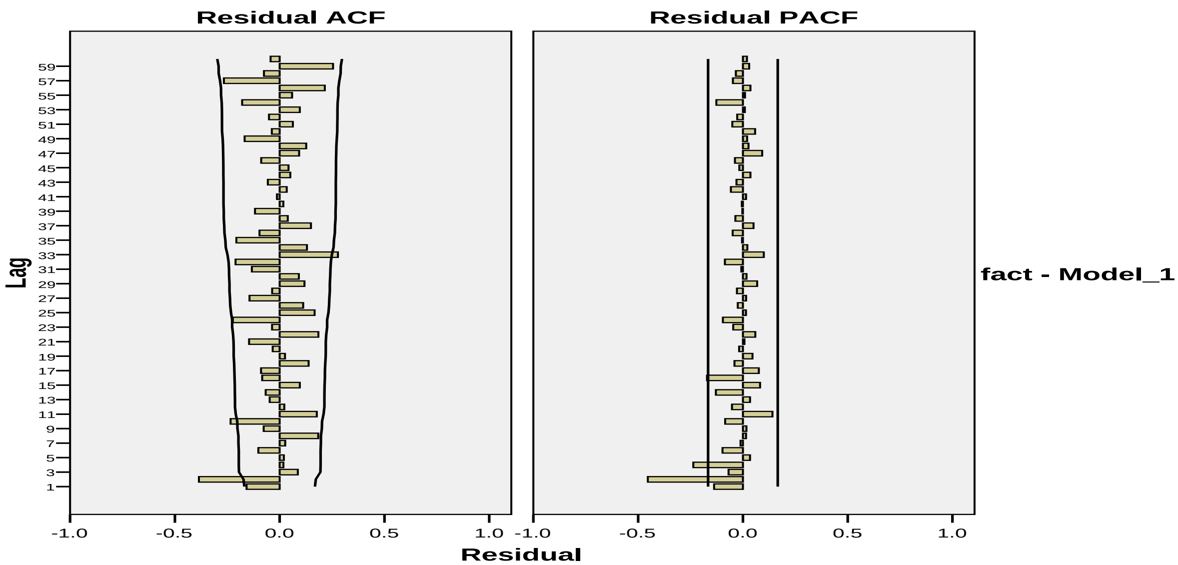
Ljung-Box Q(18)

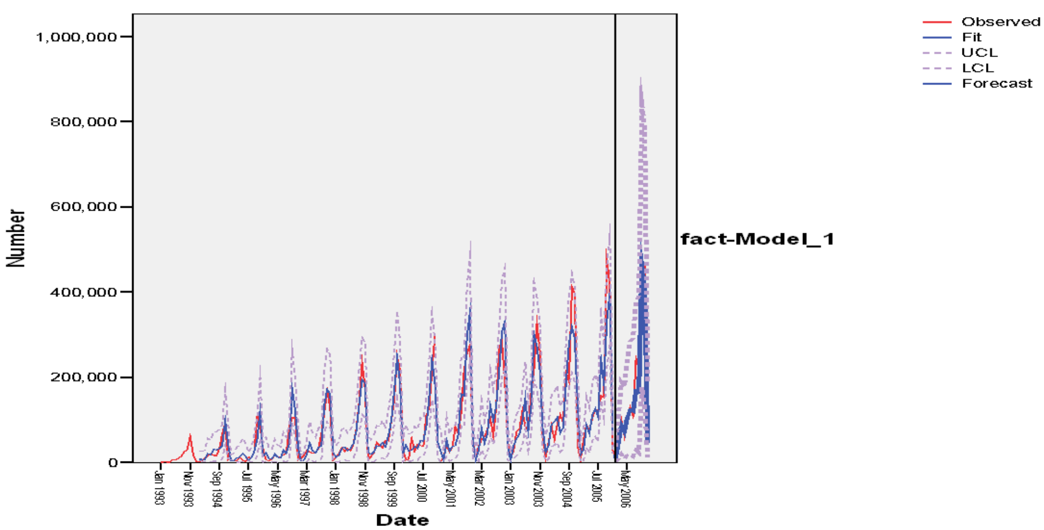
Statistic	DF	Sig
54.774	16	.000

Forecast

Model		Jan 2006	Feb 2006	Mar 2006	Apr 2006	May 2006	Jun 2006	Jul 2006	Aug 2006	Sep 2006	Oct 2006	Nov 2006	Dec 2006
fact-Model_1	Forecast	9966.70	48672.96	87722.07	72639.29	104045.6	122382.6	119623.7	167491.7	164587.5	461747.5	411119.1	45754.25
	UCL	35634.22	111999.2	187789.8	177136.7	240776.5	283100.6	291495.7	381202.1	389815.6	839826.8	787769.0	189381.2
	LCL	359.95	7864.30	19644.16	8433.59	16327.79	19244.30	13954.40	28586.33	22774.52	175690.4	135099.5	11365.52

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.





Se observa un buen ajuste, ligeramente abajo del real con un mejor ajuste estacional, todavía pequeño, esto podría indicar que es necesario que el parámetro de medias móviles es decir ***q*** sea al menos igual a 1, además observamos que la probabilidad significativa de la constante asociado al estadístico "T" para contrastar la hipótesis nula de que el parámetro correspondiente es igual a cero a un nivel de significación del 0.05 no puede ser rechazada, por lo que como el término independiente (constante) es significativamente diferente de cero se puede prescindir de él. Teniendo en cuenta estas dos observaciones proponemos el siguiente modelo.

4.7 Predicción futura por mes y cliente de los años 2007-2009

Empezaremos a incrementar algunos parámetros, y el SPSS nos ayudará a estimar pronósticos para los años 2007-2009.

Modelo 7

Model Description

			Model Type
Model ID	**Fac.**	**Model_1**	ARIMA(1,1,1)(0,1,1)

Parámetros del modelo ARIMA

					Estimación	ET	t	Sig.
FACT-Modelo_1	FACT	Raíz cuadrada	AR	Retardo 1	.193	.094	2.055	.042
			Diferencia		1			
			MA	Retardo 1	.993	.204	4.873	.000
			Diferenciación estacional		1			
			MA, estacional	Retardo 1	.228	.093	2.453	.015

Fit statistic	Mean 5
Stationary R-squared	.391
R-squared	.906
RMSE	29,254.918
MAPE	66.886
MAXAPE	1316.848
MAE	18,235.908
MAXAE	140,706.553
Normalizad BIC	20.672

Ljung-Box Q(18)

Statistic	DF	Sig
26.186	15	.036

FORECAST 2006

Modelo		Ene 2006	Feb 2006	Mar 2006	Abr 2006	May 2006	Jun 2006	Jul 2006	Ago 2006	Sep 2006	Oct 2006	Nov 2006	Dic 2006
FACT-Modelo_1	FCST	12500.63	61561.26	103315.2	85642.30	119363.6	137656.0	132118.1	186853.0	183457.2	490385.9	436892.6	42788.74
	LCS	36052.45	110324.5	165221.1	142427.1	185562.6	208395.1	201520.5	268466.2	264373.9	619430.4	558996.1	84205.11
	LCI	244.95	24342.69	52977.14	40428.12	64736.37	78489.54	74289.24	116814.1	114115.7	372917.4	326366.0	12950.15

Para cada modelo, las predicciones comienzan después del último valor no perdido del rango del período de estimación solicitado y finalizan en el último período para el que hay disponibles valores no perdidos de todos los predictores o en la fecha de finalización del período de predicción solicitado, lo que ocurra antes.

Forecast 2007

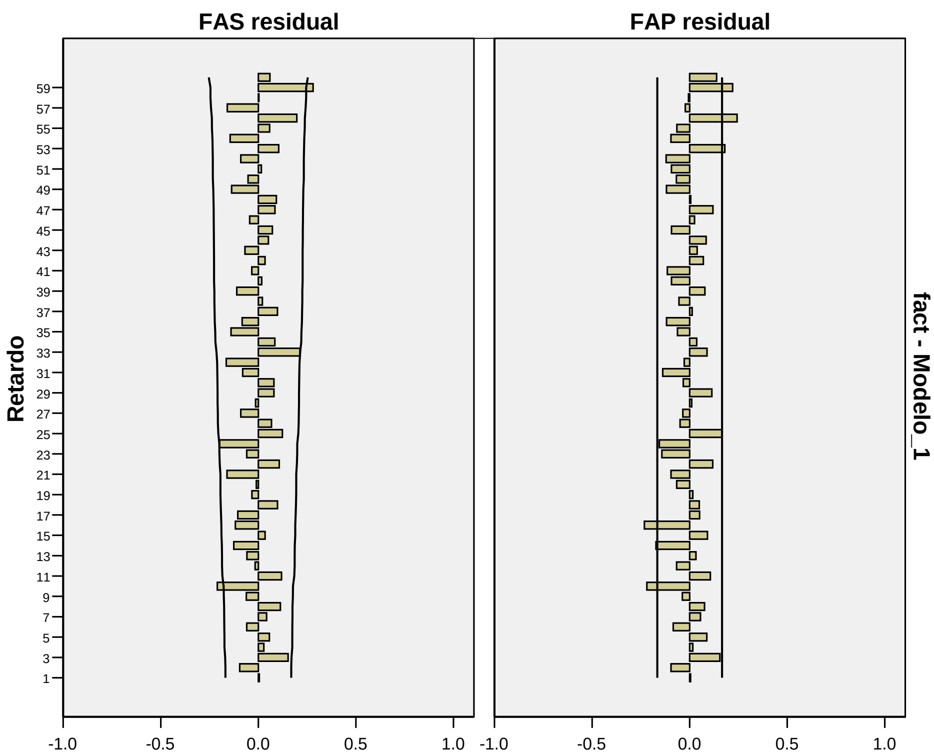
Modelo		Ene 2007	Feb 2007	Mar 2007	Abr 2007	May 2007	Jun 2007	Jul 2007	Ago 2007	Sep 2007	Oct 2007	Nov 2007	Dic 2007
Fac.-Modelo_1	Forecast	18423.88	72371.48	116685.6	97890.24	133574.3	152815.1	146996.7	204287.6	200745.7	517749.2	462802.2	51807.68
	LCS	55027.61	140539.4	201436.5	176123.7	223763.8	248773.8	241261.5	314065.3	309643.2	687782.8	624025.6	110541.9
	LCI	158.62	22837.37	50590.04	38316.91	62048.09	75522.46	71400.58	113181.3	110522.2	366392.3	320258.2	11755.56

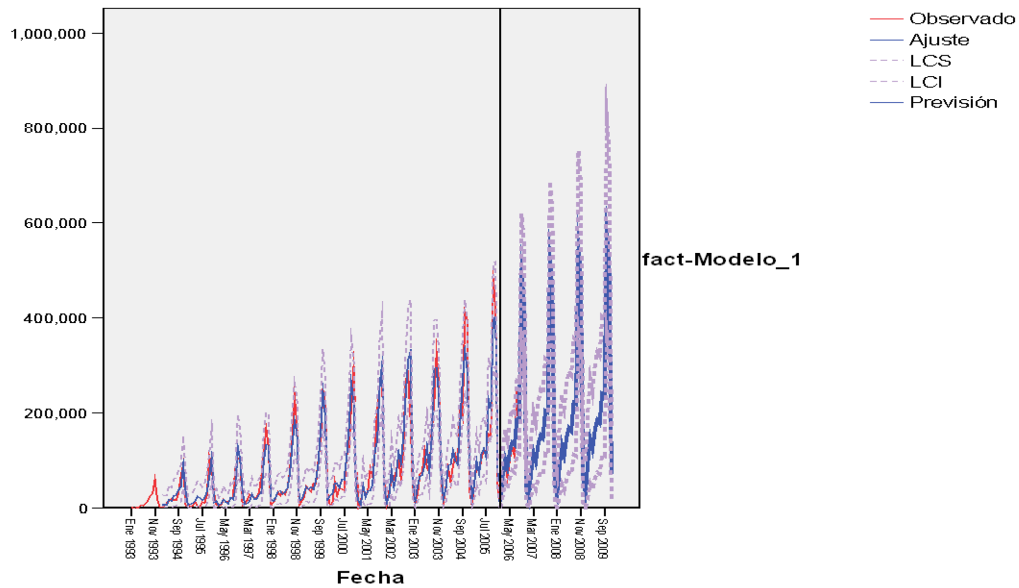
Forecast 2008

Modelo		Ene 2008	Feb 2008	Mar 2008	Abr 2008	May 2008	Jun 2008	Jul 2008	Ago 2008	Sep 2008	Oct 2008	Nov 2008	Dic 2008
Fac.-Modelo_1	Forecast	24568.29	83667.65	130710.3	110838.9	148493.5	168684.7	162586.3	222433.5	218745.6	545824.1	489423.5	61538.54
	LCS	74575.94	170833.9	237422.2	209875.0	261625.4	288621.3	280531.7	358669.1	353949.5	753100.0	686322.0	137624.0
	LCI	121.21	22382.31	49907.88	37720.78	61285.75	74677.86	70576.21	112138.7	109488.0	364500.1	318482.4	11416.05

Forecast 2009

Modelo		Ene 2009	Feb 2009	Mar 2009	Abr 2009	May 2009	Jun 2009	Jul 2009	Ago 2009	Sep 2009	Oct 2009	Nov 2009	Dic 2009
Fac.-Modelo_1	Forecast	31424.81	95676.09	145447.4	124500.1	164125.5	185267.2	178889.0	241292.7	237459.0	574612.6	516758.7	71983.43
	LCS	95672.79	202055.7	274005.1	244356.5	299979.2	328847.3	320216.4	403378.9	398381.7	817321.9	747702.2	165841.4
	LCI	137.05	22603.08	50232.97	38000.38	61638.43	75063.26	70947.16	112601.5	109940.7	365317.1	319238.3	11558.02





Este modelo presenta los parámetros con un valor significativo menor al 5%, lo cual implica que los p-valores son lo suficientemente pequeños para rechazar la hipótesis nula de que el parámetro correspondiente es igual a cero. Podemos suponer que el modelo 7 es adecuado en el sentido de que todos sus coeficientes son significativamente diferentes de cero.

4.8 Sobre-ajuste y validación del modelo

Se analizará diferentes valores de los parámetros, para detectar si al modificar algún parámetro, el modelo no obtiene mejores estimaciones de los valores de ventas.

4.8.1 Modelo 8 incrementado un parámetro autoregresivo estacional:

Descripción del modelo

			Tipo de modelo
ID del modelo	fact	Modelo_1	ARIMA(1,1,1)(1,1,1)

Parámetros del modelo ARIMA

					Estimación	ET	T	Sig.
fact-Modelo_1	Fac.	Raíz cuadrada	AR	Retardo 1	.197	.095	2.062	.041
			Diferencia		1			
			MA	Retardo 1	.993	.274	3.620	.000
			AR, estacional	Retardo 1	.410	.326	1.259	.210
			Diferenciación estacional		1			
			MA, estacional	Retardo 1	.651	.294	2.213	.029

Fit statistic	Mean 5
Stationary R-squared	.404
R-squared	.913
RMSE	28,267.510
MAPE	68.215
MAXAPE	1,327.350
MAE	17,860.410
MAXAE	125,876.224
Normalizad BIC	20.639

Ljung-Box Q(18)

Statistic	DF	Sig
25.817	14	.027

FORECAST 2006

Modelo		Ene 2006	Feb 2006	Mar 2006	Abr 2006	May 2006	Jun 2006	Jul 2006	Ago 2006	Sep 2006	Oct 2006	Nov 2006	Dic 2006
Fac.-Modelo _1	Previsión	13052.01	60573.25	101968.3	86737.22	118339.6	135743.1	130456.3	181123.1	188434.9	476645.2	422616.3	45703.04
	LCS	36852.84	108399.7	162777.5	143181.9	183503.7	205203.7	198646.6	260620.9	269427.3	602497.6	541426.9	87860.17
	LCI	366.70	24039.60	52476.08	41612.36	64496.48	77604.34	73588.86	112948.9	118766.9	362118.1	315131.8	14872.90

Para cada modelo, las predicciones comienzan después del último valor no perdido del rango del período de estimación solicitado y finalizan en el último período para el que hay disponibles valores no perdidos de todos los predictores o en la fecha de finalización del período de predicción solicitado, lo que ocurra antes.

FORECAST 2007

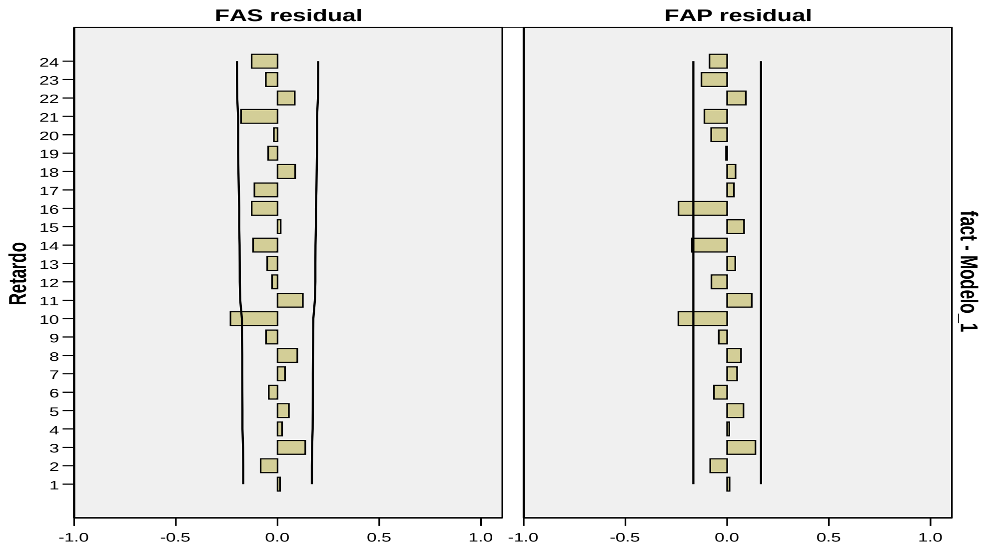
Modelo		Ene 2007	Feb 2007	Mar 2007	Abr 2007	May 2007	Jun 2007	Jul 2007	Ago 2007	Sep 2007	Oct 2007	Nov 2007	Dic 2007
Fac.-Modelo _1	Forecast	20519.47	70569.61	114943.9	98677.93	132048.4	149478.1	142238.2	201375.9	214400.2	491494.3	438679.5	58397.06
	LCS	58229.01	136724.7	197582.5	175764.2	220137.0	242748.4	233405.8	308466.7	324669.2	654454.3	593090.0	119230.1
	LCI	532.86	22427.86	50340.38	39631.53	62002.75	74253.31	69118.77	112335.9	122184.7	346590.4	302327.7	15625.28

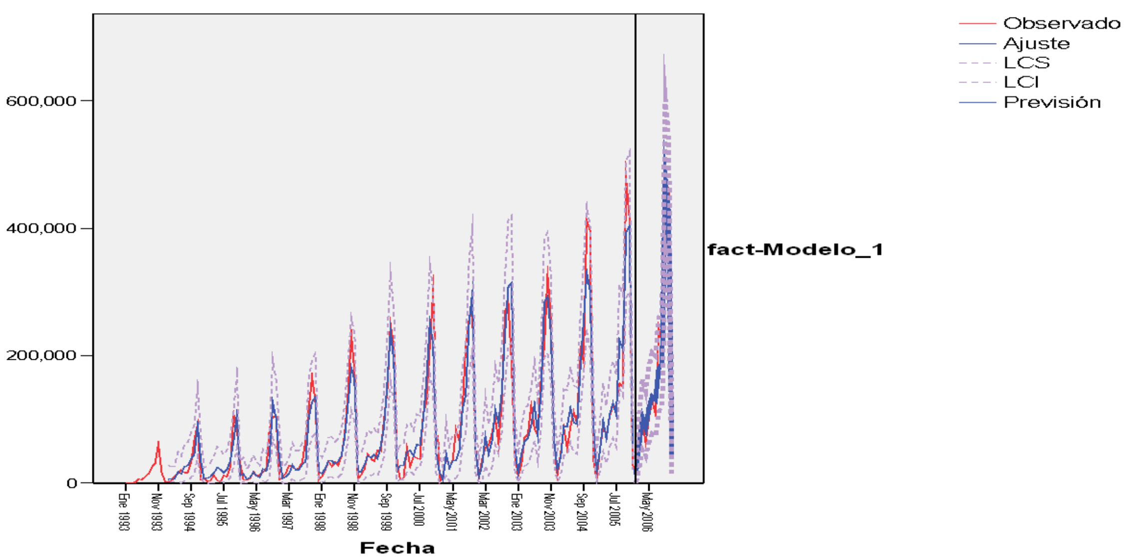
FORECAST 2008

Modelo		Ene 2008	Feb 2008	Mar 2008	Abr 2008	May 2008	Jun 2008	Jul 2008	Ago 2008	Sep 2008	Oct 2008	Nov 2008	Dic 2008
Fac.-Modelo _1	Forecast	27527.79	80950.95	128120.3	110897.4	146033.2	163931.8	155630.8	219969.8	235854.8	512833.6	459738.2	69656.00
	LCS	77219.69	161935.4	227886.6	204338.0	251976.0	275652.7	264732.6	347937.5	368039.1	703064.2	640415.7	145476.6
	LCI	852.48	23227.76	51638.99	40749.35	63388.24	75513.69	69836.69	115314.8	126988.1	345925.6	302388.3	17167.85

FORECAST 2009

Modelo		Ene 2009	Feb 2009	Mar 2009	Abr 2009	May 2009	Jun 2009	Jul 2009	Ago 2009	Sep 2009	Oct 2009	Nov 2009	Dic 2009
Fac.-Modelo _1	Forecast	34604.66	91880.36	141764.5	123618.6	160518.8	179077.1	170095.9	238227.8	255713.6	537261.3	483269.9	80612.85
	LCS	95700.53	186873.4	257402.2	232308.1	282936.1	307791.4	295834.9	385029.3	407420.5	751877.0	687475.0	170363.3
	LCI	1301.75	24917.11	54184.31	42997.31	66178.09	78447.37	72449.20	119526.5	132114.6	350761.3	307188.3	18993.66





Este modelo nos indica que la probabilidad en los parámetros son poco significativos, excepto el autoregresivo, correremos un pronóstico con ese modelo sin tomar en cuenta ese parámetro, con eso quizá disminuiremos la suma de los errores cuadrados. Tiene un buen coeficiente de correlación estacional, y la suma de sus errores al cuadrado son menores al modelo 7 y 6.

4.8.2 Modelo 9 sin considerar parámetros autoregresivos

En este modelo disminuiremos los parámetros autoregresivos, es decir subvaluaremos el modelo.

Descripción del modelo

			Tipo de modelo
ID del modelo	fact	Modelo_1	ARIMA(0,1,1)(0,1,1)

Parámetros del modelo ARIMA

					Estimación	ET	t	Sig.
fact-Modelo_1	fact	Raíz cuadrada	Diferencia		1			
			MA	Retardo 1	.901	.050	17.981	.000
			Diferenciación estacional		1			
			MA, estacional	Retardo 1	.211	.092	2.292	.023

Fit statistic	Mean 5
Stationary R-squared	.377
R-squared	.907
RMSE	29,134.095
MAPE	68.526
MAXAPE	1,526.000
MAE	18,355.721
MAXAE	143,574.170
Normalizad BIC	20.629

FORECAST 2006

Modelo		Ene 2006	Feb 2006	Mar 2006	Abr 2006	May 2006	Jun 2006	Jul 2006	Ago 2006	Sep 2006	Oct 2006	Nov 2006	Dic 2006
FACT-Modelo_1	FCST	12611.14	58170.49	97754.47	80505.25	113089.2	131110.9	126050.5	178295.5	174541.2	479122.1	426051.5	38931.28
	LCS	36677.72	106216.0	159008.4	136819.7	179284.6	202362.1	196362.0	261347.1	257171.3	613028.2	553247.5	81101.07
	LCI	209.29	21903.58	48393.73	36198.60	59016.30	72096.66	68090.59	107710.1	104491.9	357911.5	311665.5	9686.17

Para cada modelo, las predicciones comienzan después del último valor no perdido del rango del período de estimación solicitado y finalizan en el último período para el que hay disponibles valores no perdidos de todos los predictores o en la fecha de finalización del período de predicción solicitado, lo que ocurra antes.

Forecast 2007

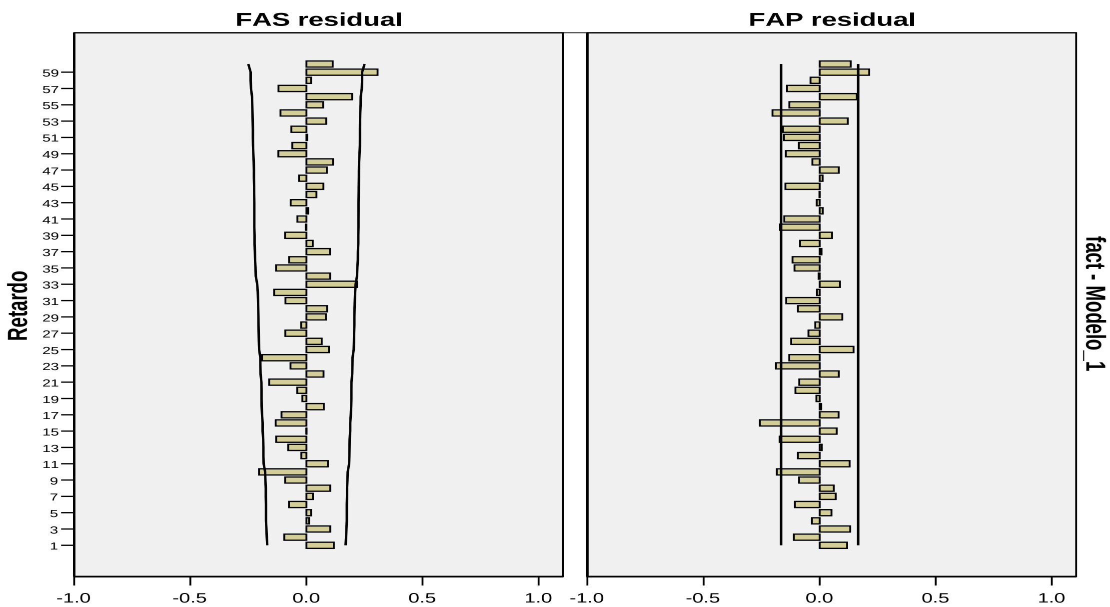
Modelo		Ene 2007	Feb 2007	Mar 2007	Abr 2007	May 2007	Jun 2007	Jul 2007	Ago 2007	Sep 2007	Oct 2007	Nov 2007	Dic 2007
Fac.-Modelo_1	Forecast	16839.77	65444.13	106694.0	88831.68	122648.7	141296.3	136119.1	189923.7	186111.4	496906.8	442992.8	45621.72
	LCS	55735.66	137680.3	197553.4	173191.7	221131.6	247229.8	241126.6	313239.9	309259.4	693022.4	630194.2	112454.6
	LCI	59.24	15690.00	38683.26	27686.89	47747.73	59311.24	55426.78	91289.17	88011.79	326206.3	281573.0	4937.09

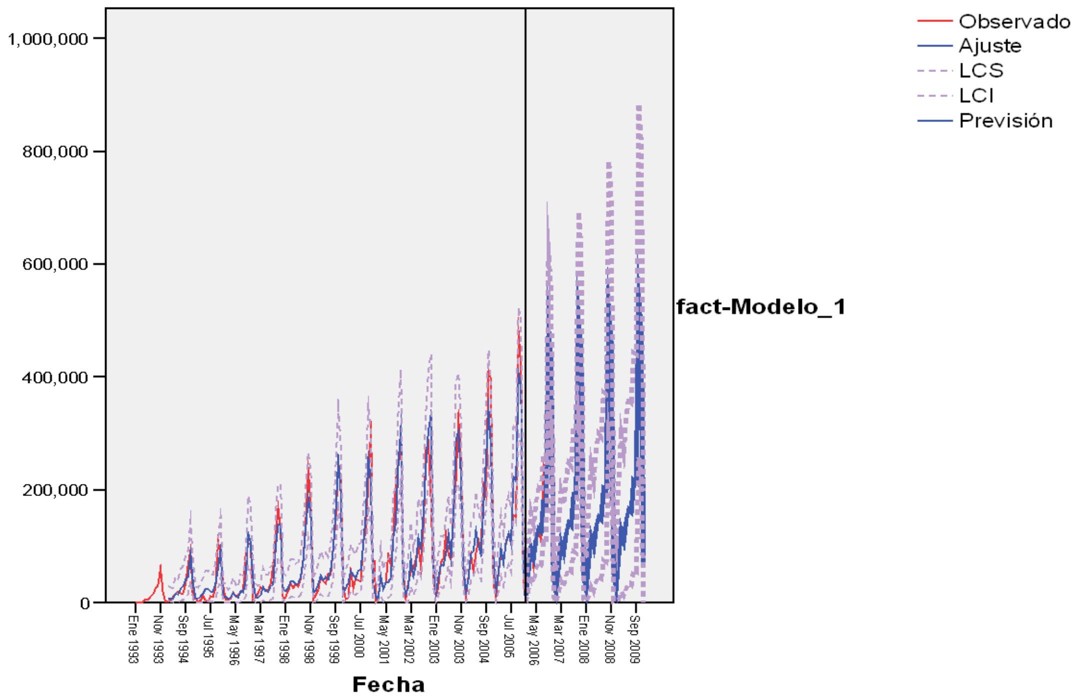
Forecast 2008

Modelo		Ene 2008	Feb 2008	Mar 2008	Abr 2008	May 2008	Jun 2008	Jul 2008	Ago 2008	Sep 2008	Oct 2008	Nov 2008	Dic 2008
Fac.-Modelo_1	Forecast	22086.15	73759.88	116700.2	98249.20	133323.8	152621.7	147352.4	202740.9	198895.2	515929.5	461196.7	53599.15
	LCS	80156.18	175304.8	242762.9	216219.4	270047.3	299441.9	293337.6	373078.0	369408.3	782708.6	716765.3	151050.7
	LCI	1044.93	10004.91	29188.61	19591.32	36673.88	46636.30	42962.92	74760.91	71500.17	293029.8	250268.5	1549.26

Forecast 2009

odelo		Ene 2009	Feb 2009	Mar 2009	Abr 2009	May 2009	Jun 2009	Jul 2009	Ago 2009	Sep 2009	Oct 2009	Nov 2009	Dic 2009
ac.- o_1	orecast	28644.01	83411.60	128066.8	109051.7	145408.4	165381.1	160044.0	217041.1	213186.3	536484.1	480956.9	63157.42
	CS	111682.9	221226.7	296956.6	268104.8	328346.1	361345.1	355287.3	443281.9	439988.7	885034.1	815763.1	198688.3
	CI	3719.83	5009.45	19888.16	12007.93	25778.30	34023.09	30704.79	58002.67	54884.57	257733.0	217247.8	21.85





En este modelo se observa una disminución de la probabilidad de ocurrencia de los parámetros, sin embargo la suma de los errores cuadrados aumenta, lo cual no nos garantiza que mejore el modelo de manera real, por lo que definiremos que hasta ahora el modelo 8 es el más adecuado para estimar los datos de la serie de ventas en el tiempo.

4.8.3 Aumento de los parámetros (sobrevaluación del modelo)

Al incrementar los valores de lo parámetros, tampoco encontramos una mejoría sustancial, como se observa en la suma de los cuadrados, y en las probabilidades de los parámetros que cada vez se vuelven más significativas:

a) Si aumentamos el parámetro de Medias Móviles se obtiene lo siguiente:

Descripción del modelo

			Tipo de modelo
ID del modelo	fact	Modelo_1	ARIMA(0,1,2)(0,1,1)

Parámetros del modelo ARIMA

					Estimación	ET	t	Sig.
fact-Modelo_1	fact	Raíz cuadrada	Diferencia		1			
			MA	Retardo 1	.750	1.148	.654	.514
				Retardo 2	.248	.321	.773	.441
			Diferenciación estacional		1			
			MA, estacional	Retardo 1	.253	.093	2.709	.008

Fit statistic	Mean 5
Stationary R-squared	.396
R-squared	.908
RMSE	29074.134
MAPE	66.039
MAXAPE	1215.445
MAE	18133.106
MAXAE	136477.091
Normalizad BIC	20.660

Se observa que los errores al cuadrado crecen, y la probabilidad de cada parámetro crece mucho, por lo que por el principio de parsimonia, este modelo no aporta un mejor pronóstico.

b) Si aumentamos el parámetro de medias móviles estacional los resultados son los siguientes:

			Tipo de modelo
ID del modelo	fact	Modelo_1	ARIMA(0,1,1)(0,1,2)

Parámetros del modelo ARIMA

					Estimación	ET	T	Sig.
Fac.-Modelo_1	fact	Raíz cuadrada	Diferencia		1			
			MA	Retardo 1	.937	.053	17.815	.000
			Diferenciación estacional		1			
			MA, estacional	Retardo 1	.181	.095	1.918	.057
				Retardo 2	.192	.101	1.904	.059

Fit statistic	Mean 5
Stationary R-squared	.394
R-squared	.919
RMSE	27156.660
MAPE	71.641
MAXAPE	1665.737

MAE	17827.855
MAXAE	118968.265
Normalizad BIC	20.523

Este modelo disminuye la suma de los errores cuadrados, sin embargo la probabilidad del parámetro es un 6% significativamente diferente de cero, por el principio mencionado, no se queda este modelo óptimo para pronosticar.

c) Finalmente analizaremos un modelo con dos diferencias estacionales, para ver si aumentando el coeficiente de correlación estacional, se obtiene mejores resultados en nuestro pronóstico.

Descripción del modelo

			Tipo de modelo
ID del modelo	fact	Modelo_1	ARIMA(0,1,1)(0,2,1)

Parámetros del modelo ARIMA

					Estimación	ET	T	Sig.
Fac.-Modelo_1	fact	Raíz cuadrada	Diferencia		1			
			MA	Retardo 1	.727	.076	9.529	.000
			Diferenciación estacional		2			
			MA, estacional	Retardo 1	.999	19.072	.052	.958

Fit statistic	Mean 5
Stationary R-squared	.616
R-squared	.877
RMSE	34234.134
MAPE	70.350
MAXAPE	2019.108
MAE	22520.246
MAXAE	133205.727
Normalizad BIC	20.957

Como se ve en la corrida, el coeficiente de autocorrelación estacionario, es mejor, sin embargo se desvía del comportamiento real de la serie ya que aumentan en un buen porcentaje los errores al cuadrado.

Por esto nuestra recomendación para estimar la serie de Ventas 2007-2009, es el modelo 8, en el capítulo 5, parte dos, expondremos un estudio de Incorrelación, Normalidad y Predicción.

Capítulo 5

Análisis de resultados y estrategias

Capítulo 5

Análisis de resultados y estrategias recomendadas en base a resultados del modelo matemático.

5.1 Análisis de la correlación entre el entorno del mercado, considerando estados financieros de los principales clientes con los datos obtenidos en el pronóstico en los últimos 3 años

5.1.1 Correlación entre las ventas de la distribuidora de juguetes considerando variables macroeconómicas

De acuerdo con lo revisado en el capítulo 3, existe una fuerte correlación de los gastos de publicidad con las ventas de la empresa juguetera, esta parte se analizará que además los factores macroeconómicos tienen un peso importante en las proyecciones de ventas, primero daremos algunas variables que los principales clientes analizan y utilizan para estudiar su inversión, establecer sus planes de crecimiento, delimitan a donde quieren llegar. Algunos de estos parámetros son los que mencionamos en el capítulo 2, en la sección de hipótesis económicas y financieras.

5.1.2 Regresión lineal múltiple de correlación entre las ventas de la distribuidora de juguetes considerando Costos de Publicidad e Inflación.

Regresión lineal múltiple de ventas con costos publicidad e inflación del período 1993-2006, Ventas en pesos (variable dependiente), costos de publicidad en pesos e inflación en porcentaje (variables independientes).

Regresión Lineal Vtas con Ctos Pub y PIB

Resumen

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0.99692931
Coeficiente de determinación R^2	0.99386806
R^2 ajustado	0.99275316
Error típico	51616.95
Observaciones	14

ANÁLISIS DE VARIANZA

	Grados de libertad	*Suma de cuadrados*	*Promedio de los cuadrados*	*F*	*Valor crítico de F*
Regresión	2	4.7502E+12	2.3751E+12	891.442531	6.7887E-13
Residuos	11	2.9307E+10	2664309523		
Total	13	4.7795E+12			

	Coeficientes	*Error típico*	*Estadístico t*	*Probabilidad*	*Inferior 95%*	*Superior 95%*	*Inferior 95.0%*	*Superior 95.0%*
Intercepción	-171051.594	40631.2812	-4.20984986	0.00146113	-260480.441	-81622.7466	-260480.441	-81622.7466
x1	4.2275018	0.64964331	6.50741992	4.383E-05	2.79764652	5.65735708	2.79764652	5.65735708
x3	2286.68293	351.955172	6.49708575	4.4454E-05	1512.03482	3061.33104	1512.03482	3061.33104

Análisis de los residuales

Observación	*Pronóstico y*	*Residuos*	*Residuos estándares*
1	152027.08	19628.7859	0.41340567
2	228638.243	53964.6278	1.13655951
3	310471.667	-5788.61092	-0.12191506
4	464934.898	-66478.9776	-1.40012666
5	660742.267	-25823.3586	-0.54387077
6	811646.53	13645.8652	0.28739822
7	952186.29	42192.4468	0.88862332
8	1037098.4	-82879.1806	-1.74553453
9	1197267.8	-14970.3229	-0.31529288
10	1331647.98	-15428.2996	-0.32493842
11	1429328.63	69613.9038	1.46615196
12	1665167.73	68599.604	1.44478959
13	1846973.93	-43255.943	-0.91102182
14	2013815.93	-13020.5404	-0.27422813

Resultados de datos de probabilidad

Percentil	*y*
3.57142857	171655.865
10.7142857	282602.871
17.8571429	304683.056
25	398455.921
32.1428571	634918.909
39.2857143	825292.395
46.4285714	954219.22
53.5714286	994378.737
60.7142857	1182297.47
67.8571429	1316219.68
75	1498942.54
82.1428571	1733767.34
89.2857143	1803717.98
96.4285714	2000795.39

Los resultados nos arrojan un coeficiente de correlación arriba del 98%, que se había obtenido con los Costos de Publicidad con lo que podemos afirmar que observamos que existe muy importante correlación entre el costo de publicidad y el crecimiento de PIB, es decir la economía en global esta muy relaciona con el consumo de productos, ya que cuando las personas tienen más recursos en promedio gasta en productos y el juguete se vuelve un producto de consumo masivo como un perecedero, es un hecho importante que los mayores consumidores de juguetes sean las clases con menos poder adquisitivo, hecho muy notable en temporada.

Las estimaciones para 2007-2009 son las siguientes:

Pronóstico para 2007-09

Año	**Costo de Publicidad**	**PIB**	**Pronostico Vtas**	**% inc Costo de Publicidad**	**% inc Pronostico de vtas**
2007	$ 260,000.00	3.48%	$ 2,153,920		
2008	$ 285,000.00	3.90%	$ 2,350,633	9.62%	9.13%
2009	$ 305,000.00	3.85%	$ 2,509,289	7.02%	6.75%

Un hecho importante es que si no existe la inflación la estimación es bastante lineal y el crecimiento sería plano, pero como la economía real tiene inflación, nuestra siguiente estimación la incluirá.

5.1.3 Regresión lineal múltiple de correlación entre las ventas de la distribuidora de juguetes considerando más variables independientes Costos de Publicidad e Inflación, PIB, Salario Mínimo.

Ahora haremos una Regresión lineal múltiple de ventas con más variables independientes, para saber si incrementamos el costos publicidad con una inflación, estable y de acuerdo a los pronósticos del PIB del período 1993-2006, Las ventas están en pesos (variable dependiente), los costos de publicidad, salario mínimo e inflación en pesos , PIB en porcentaje (todas ellas como variables independientes)

Regresión Lineal Múltiple con Costo de Publicidad, PIB, Inflación y Sal Mínimo

Resumen

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0.9972649
Coeficiente de determinación R^2	0.9945372
R^2 ajustado	0.9921093
Error típico	53861.243
Observaciones	14

ANÁLISIS DE VARIANZA

	Grados de libertad	*Suma de cuadrados*	*F*	*Valor crítico de F*
Regresión	4	4.7534E+12	409.626	3.60404E-10
Residuos	9	2.6109E+10		
Total	13	4.7795E+12		

	Coeficientes	*Error típico*	*Probabilidad*	*Inferior 95%*	*Superior 95%*	*Inferior 95.0%*	*Superior 95.0%*
Intercepción	-83873.15	99029.1892	0.41897	-307892.7376	140146.4409	-307892.7376	140146.4409
Variable X 1	4.0307296	0.76822869	0.00053	2.292875528	5.768583581	2.292875528	5.768583581
Variable X 2	-138187.9	143705.078	0.36137	-463271.3458	186895.5971	-463271.3458	186895.5971
Variable X 3	3161.2691	1482.81825	0.06183	-193.0987969	6515.637058	-193.0987969	6515.637058
Variable X 4	-9212.355	13738.2955	0.51934	-40290.53825	21865.82884	-40290.53825	21865.82884

Análisis de los residuales

Observación	*Pronóstico para Y*	*Residuos*
1	179740.59	-8084.7269
2	251178.62	31424.2522
3	281601.17	23081.8893
4	466265.1	-67809.176
5	651704.76	-16785.847
6	792448.03	32844.3673
7	937973.49	56405.2444
8	1032898.6	-78679.378
9	1193869.4	-11571.953
10	1329711.3	-13491.596
11	1445641.3	53301.2747
12	1674686.7	59080.6493
13	1853103.2	-49385.184
14	2011125.2	-10329.817

Resultados de datos de probabilidad

Percentil	*Y*
3.571428571	171655.8655
10.71428571	282602.8707
17.85714286	304683.0557
25	398455.9208
32.14285714	634918.9088
39.28571429	825292.3948
46.42857143	954219.2199
53.57142857	994378.7369
60.71428571	1182297.474
67.85714286	1316219.677
75	1498942.538
82.14285714	1733767.338
89.28571429	1803717.982
96.42857143	2000795.39

Al incluir inflación, se observa que su coeficiente es inversamente proporcional al estimado de ventas, por lo que su importancia esta demostrada en la siguiente estimación, donde se observa que cada vez debemos invertir más dinero en publicidad en proporción para poder incrementar las ventas, lo cual sabemos, no es posible hacerlo, sobre todo por que si no se tiene un incremento que sea por participación nueva en el mercado, los incrementos están muy cercanos a la inflación, lo que indica que el crecimiento real de la compañía podría ser cero o negativo, lo cual indica que otro competidor esta creciendo y esta quitando a la Distribuidora participación de mercado.

Las estimaciones para el año 2007-2009 son las siguientes:

Pronóstico para 2007-09

Año	Costo de Publicidad	Inflación	PIB	SMDF	Pronostico Vtas	% inc Costo de Publicidad	% inc Pronostico de vtas
2007	$260,000.00	4.02%	$ 491.30	$ 50.57	$ 2,045,824		
2008	$285,000.00	3.90%	$ 510.46	$ 52.85	$ 2,186,324	9.62%	6.87%
2009	$305,000.00	3.85%	$ 530.11	$ 55.00	$ 2,309,320	7.02%	5.63%

Otro aprendizaje que nos da es que no necesariamente al incrementar la publicidad, las ventas crecerán en la misma proporción, el mercado tiene un techo, por lo que se hace necesario tener en cuenta hasta donde es una mala inversión invertir en publicidad, ya que no ayudará igualmente a mejorar ventas. Como se observa la correlación es altísima, por lo que se debe de tener planes de expansión de línea, o empezar a desarrollar nuevos Clientes, ya que no se tiene mucho margen de maniobra para crecer. Por último concluimos que el costo por publicidad es una variable que esta muy correlacionada con el importe de ventas y es necesario para hacer estimaciones utilizando regresión múltiple, conocer su importe.

5.1.4 Desempeño de Ventas 2006 del Sector de Autoservicios, Departamentales y Especializados

Se superan las expectativas de ventas anuales en el 2006 (**11.3% arriba de 2005**), por factores como: incremento en el número de tiendas, el calendario, elecciones federales que incrementan el circulante por la campañas políticas, y un incremento de un 32% en los vales de D.F. pero a nuestra consideración fue por el entorno económico favorable e incremento en precios de alimentos y no perecederos.[11]

Factores adicionales como el Empleo, que según el IMSS creció a una tasa del 6.8% al cierre de 2006, alcanzando un total de $14´080,367. De esta forma se crearon 895 mil nuevos empleos, que representaron el 66% de las nuevas fuentes de trabajo formal de los últimos 6 años. Con respecto a los salarios, de acuerdo al INEGI, las remuneraciones medias en el sector manufacturero crecieron en promedio 0.5% durante ene-dic del 2006. También se incremento el crédito al consumo, de acuerdo con el Banco de México, al mes de diciembre del 2006 el crédito al consumo bancario registraba un tasa de crecimiento de 37.9% en términos reales. Finalmente las remesas, acumularon US$ 23,054 millones de dólares en el año de 2006, que representan un incremento del 16.5% contra el 2005.

[11] **Infobasic, S.A de C.V. Información, Análisis y Consultoria Marzo 2007**

5.1.5 Indicadores de Desempeño de Ventas 2006 del Sector de Autoservicios

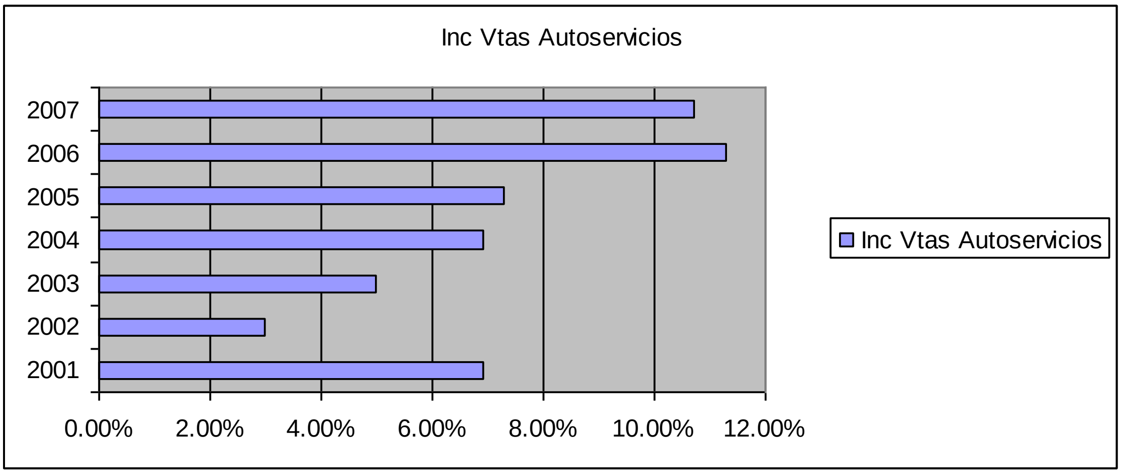
En este apartado se darán algunos datos importantes de los principales clientes de la compañía juguetera: los autoservicios, quienes han tenido un crecimiento enorme, sobre todo por su histórico número de aperturas, concentradas en dos cadenas principalmente: Wal Mart y Soriana que tienen el 70% de las inversiones. Además se ve incremento en la productividad o ventas a Tiendas iguales, que crecieron un 3.6%.

Las otras Cadenas han mejorado y modernizado su piso de ventas, e incluso cerrado Tiendas que no sean rentables, lo que da un total de 169 unidades nuevas que representan un histórico en la industria del autoservicio, en el transcurso del año.

Los incrementos por año de las ventas de los Autoservicios son los siguientes:

Año	% inc. Por año
2001	6.90%
2002	3.00%
2003	5.00%
2004	6.90%
2005	7.30%
2006	11.30%
2007	10.70%

Sector Tiendas de Autoservicio- Ventas a Tiendas Totales



5.1.6 Perspectivas de los autoservicios para el 2007

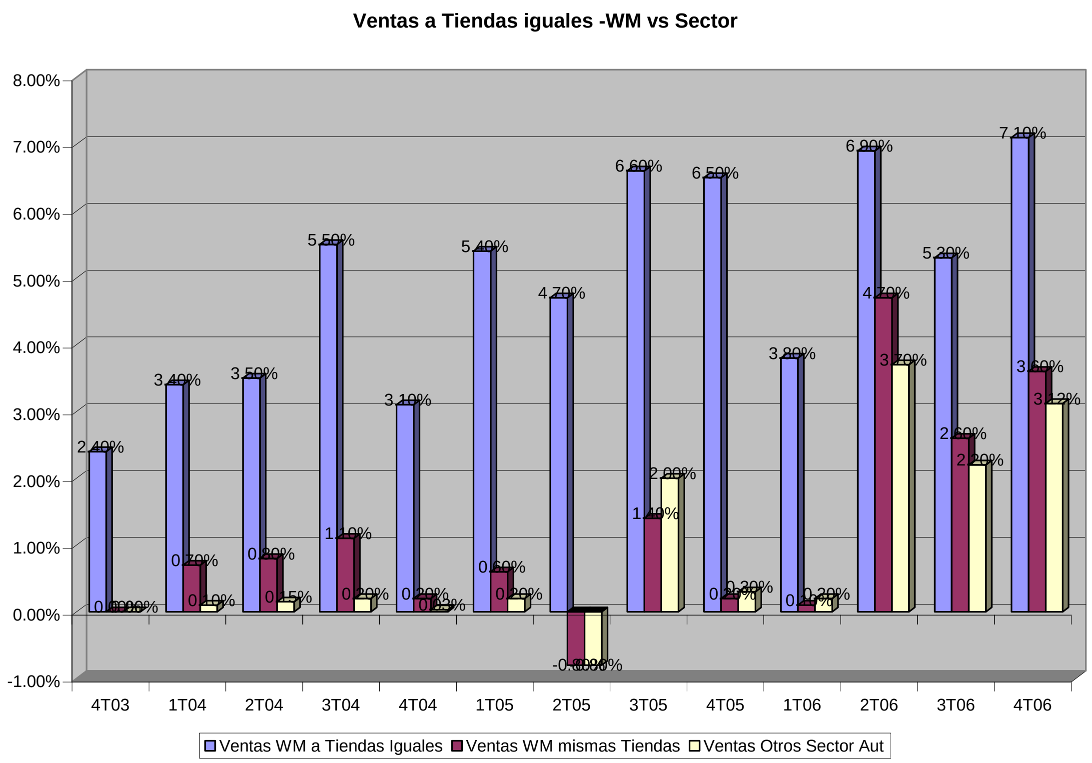
Los autoservicios apuntan a un desarrollo sostenido, seguirán las inversiones en punto de venta nuevos, pero a un menor dinamismo, se pronostica una disminución del PIB, sobre todo influenciado por la desaceleración de la economía de los E.U. y su influencia para México.

Se espera que el consumo privado crecerá a un ritmo de un 4.5% en el 2007 [12], los empleos se estiman en 700,000 nuevos empleos. El crédito al consumo seguirá su tendencia ascendente aunque se pronostica una desaceleración, y se estima un crecimiento de un 2.3% a tiendas iguales y a ventas a tiendas totales (es decir con nuevas tiendas), un crecimiento de un 10.7%. Los meses claves al igual que en el mercado juguetero, los meses claves son octubre, noviembre y diciembre, donde se venden los vales del D.F. que en el 2006 son redimidos en dichas sucursales.

5.1.7 Ejemplo Cliente: Wal Mart

Como lo hemos mencionado anteriormente es el principal cadena del País, con más de 893 tiendas a lo largo de toda la república.

Las Ventas de esta cadena tienen los mayores incrementos de todas las cadenas comerciales, las cuales si las agrupamos en un sector, los resultados de los últimos 12 trimestres, se ven reflejado en la siguiente gráfica:



Donde lo azul son las venta de Wal Mart, y es donde se nota la diferencia, con respecto a sus competidores, ya que cada año crece más el diferencia y el porcentaje de mercado que Wal Mart contra el resto del autoservicio.

[12] **Infobasic, S.A de C.V. Información, Análisis y Consultoria Marzo 2007**

Los Estados Financieros publicados en Internet por el mismo Wal Mart son los siguientes:

WAL MEX

Balance General

En millones de Pesos	2005	2006	(%)	Inc Real %
Activo Total	**88,231**	**98,862**	**100%**	**12%**
Activo Creciente	32,654	37,412	37.8%	15%
Efectivo	14,735	14,985	15.2%	2%
Clientes y Doctos x Cob	2,735	3,815	3.9%	39%
Inventarios	14,669	18,058	18.3%	23%
Otros	514	555	0.6%	8%
Activo Fijo Neto	55,559	61,450	62.2%	11%
Otros Activos	-	-		
Pasivo Total	**34,314**	**39,366**	39.8%	**15%**
Pasivo a Corto Plazo	28,018	31,360	31.7%	12%
Con Costo	-	-		
Sin Costo	28,018	31,360	31.7%	12%
Pasivo a Largo Plazo				
Con Costo				
Sin Costo				
Otros Pasivos	6,296	8,006	8.1%	27%
Capital Contable Mayoritario	**53,898**	**59,496**	**60.2%**	**10%**
Interés Minoritario	-	-		
Cap.Contable Mayoritario	53,898	59,496	60.2%	10%

e: Estimado por WalMart
Datos proporcionados en Internet por Wal Mart, Febrero 1007

El Estado de Resultados es el siguiente:

Wal Mart Estado de Resultados

En millones de Pesos	Acumulado 2005	2006	(%)	Inc Real %	Trimestral 2005	2006	(%)	Inc Real %
Ventas	**171,705**	**198,971**	**100.00%**	**16%**	**52,243**	**60,814**	**100.00%**	**16%**
Costo de Ventas	135,116	156,020	78.41%	15%	40,867	47,473	78.06%	16%
Utilidad Bruta	36,589	42,952	21.59%	17%	11,376	13,341	21.94%	17%
Gastos de Operación	23,835	26,834	13.49%	13%	6,619	7,565	12.44%	14%
Utilidad de Operación	**12,755**	**16,117**	**8.10%**	**26%**	**4,757**	**5,777**	**9.50%**	**21%**
Ut Oper+Depreciación	15,991	19,418	9.76%	21%	5,519	6,632	10.91%	20%
Ut Oper+Dep+Int Gan.	17,218	20,364	10.23%	18%	5,800	6,865	11.29%	18%
Ut Oper+Dep+Int Netos	17,169	20,263	10.18%	18%	5,751	6,837	11.24%	19%
Resultado Cambiario	- 8	- 42	-0.02%	425%	16	7	0.01%	-56%
Repomo	- 265	- 519	-0.26%	96%	- 180	- 261	-0.43%	45%
Resultado Neto	**9,851**	**12,425**	**6.24%**	**26%**	**3,682**	**4,384**	**7.21%**	**19%**
Result Neto -Part Extraord	9,851	12,425	6.24%	26%	3,682	4,384	7.21%	19%
R. Neto+Dep+Repomo+R. Camb.	12,814	15,163	7.62%	18%	4,280	4,984	8.20%	16%

Wal Mart Balance General	**Información Nominal (millones) 2006**	**2007 e**	**2008 e**	**Márgenes 2006**	**2007 e**	**2008 e**	**Variación Real ´07/06**	**´08/07**
Activo Total	**98,862**	**110,798**	**125,214**	**100%**	**0%**	**100.00%**	**8.77%**	**9.91%**
Activo Circulante	37,412	40,206	45,679	**37.84%**	**100.00%**	**36.48%**	**4.17%**	**10.51%**
Efectivo e Inversiones Temporales	14,985	15,175	16,350	**15.16%**	**36.29%**	**13.06%**	**-2.03%**	**4.64%**
Clientes y Docs por cobrar	1,770	2,113	2,482	**1.79%**	**13.70%**	**1.98%**	**16.08%**	**14.36%**
Inventarios	18,058	19,938	23,428	**18.27%**	**1.91%**	**18.71%**	**7.11%**	**14.40%**
Otros Activos Circulantes	2,599	2,980	3,419	**2.63%**	**17.99%**	**2.73%**	**11.36%**	**11.63%**
Activo a Largo Plazo	**-**	-	-	**0.00%**	**2.69%**	**0.00%**		
Inmuebles, Planta y Equipo Neto	61,450	70,592	79,535	**62.16%**	**0.00%**	**63.52%**	**11.58%**	**9.57%**
Activo Diferido Neto	-	-	-	**0.00%**	**63.71%**	**0.00%**		
Otros Activos	-	-	-	**0.00%**	**0.00%**	**0.00%**		
Pasivo Total	**39,366**	**43,360**	**48,359**	**39.82%**	**0.00%**	**38.62%**	**6.85%**	**8.43%**
Pasivo Circulante	31,360	35,088	39,827	**31.72%**	**39.13%**	**31.81%**	**8.59%**	**10.41%**
Pasivo sin Costo	31,360	35,088	39,827	**31.72%**	**31.67%**	**31.81%**	**8.59%**	**10.41%**
Proveedores	25,865	29,409	33,970	**26.16%**	**31.67%**	**27.13%**	**10.40%**	**12.41%**
Impuestos por pagar	474	489	505	**0.48%**	**26.54%**	**0.40%**	**-0.14%**	**0.17%**
Otros	5,022	5,189	5,352	**5.08%**	**0.44%**	**4.27%**	**0.03%**	**0.04%**
Pasivos con Costo	-	-	-	**0.00%**	**4.68%**	**0.00%**		
Pasivo a Largo Plazo		-		**0.00%**	**0.00%**	**0.00%**		

	-		-					
Pasivos sin Costo	-	-	-	**0.00%**	**0.00%**	**0.00%**		
Pasivos con Costo	-	-	-	**0.00%**	**0.00%**	**0.00%**		
Créditos Diferidos	-	-	-	**0.00%**	**0.00%**	**0.00%**		
Otros Pasivos	8,006	8,272	8,532	**8.10%**	**0.00%**	**6.81%**	**0.02%**	**0.04%**
Capital Contable Consolidado	**59,496**	**67,438**	**76,855**	**60.18%**	**0.00%**	**61.38%**	**10.05%**	**10.86%**
Participación Minoritaria	-	-	-					
Capital Contable Mayoritario	59,496	67,438	76,855	**60.18%**	**0.00%**	**61.38%**	10.05%	10.86%

En cuanto a Estados Financieros, Wal Mart ha tenido crecimientos de 2 dígitos, porcentaje, muy similar al porcentaje de crecimiento de expansión de nuevas unidades de negocio. El reporte del cuarto trimestre de 2006, las ventas totales de este cliente suman $60,814 millones de pesos, un incremento del 16.4% en términos reales con respecto al mismo periodo de 2005. En ventas totales durante el 2006 ascendieron a $198,971 millones de pesos, contra los 178,705 millones de pesos del 2005, lo que significa un 16% de incremento de ventas.

Pero para un inversionista, le interesa la Utilidad de operación creció un 21% en términos reales, al sumar $5,777 millones de pesos. La utilidad neta trimestral totalizó $4,384 millones de pesos, un incremento de 19% en la comparación anual. En el periodo octubre-diciembre de 2006, el margen bruto se ubicó en 21.9%, lo que represento una ampliación de 10 puntos base con respecto al mismo periodo del año anterior. Por último los gastos de operación, disminuyeron 30 puntos base (0.3 puntos porcentuales) como proporción de los ingresos totales.

Proyecciones WM

Se tiene proyectado un crecimiento en ventas más moderado. Durante el 2007, prevemos que las ventas totales de Wal Mart alcanzarán $232,080 millones de pesos lo que representa un incremento de 12.9% en términos reales con respecto al año anterior. Este desempeño es menor al avance del 15.9% en los ingresos totales registrado en el 2006. En principio, creemos que el ascenso en la productividad o ventas en tiendas iguales será de entre 4.0% y 4.5% en el año en curso, en tanto que en el 2006 fue de 5.9%. Asimismo, el número de aperturas previsto para el 2007 es de 125 unidades, muy similar a las 120 del 2006. En el rubro del margen de comercialización o bruto, creemos que el intenso entorno competitivo contrarresterá las eficiencias que se pudieran alcanzar, como por ejemplo en el centro de distribución de congelados y refrigerados. Por lo anterior, esperamos que el margen bruto se mantenga en los mismos niveles del 2006 es decir 21.6%. De acuerdo a las expectativas del Cliente la utilidad de operación sumará $19,105 millones de pesos al crecer 14.7% en términos reales. La utilidad neta alcanzará $14,480 millones de pesos, un incremento estimado de 12.8%.

5.1.8 Estimación de Ventas Wal Mart en la comercializadora juguetera

Como las Ventas en Wal Mart mantienen de acuerdo a lo visto anteriormente un crecimiento exponencial y este esta muy ligado a años recientes, se aplicaran dos métodos sencillos de estimación de las ventas, conocidos como suavizamiento exponencial simple y suavizamiento lineal de Holt, cuyas expresiones requieren del cálculo de una serie de parámetros, y cuyas expresiones son las siguientes:

Suavizamiento exponencial simple:

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$$
$$\alpha = parámetro\_de\_suavizamiento < 1$$

Expresión alterna equivalente:

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + ...$$
$$\alpha = parámetro\_de\_suavizamiento < 1$$

Suavizamiento lineal de Holt

Estimado del nivel actual de la serie suavizada exponencialmente:
$$L_t = \alpha Y_t + (1-\alpha)(L_t + Y_{t-1})$$

Estimado de tendencia:
$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1})$$
$$\beta = Parámetro\_de\_tendencia$$

Pronóstico del periodo ***p*** en el futuro:
$$\hat{Y}_{t+p} = L_t + pT_t$$
$$p = periodo$$

Estos métodos son utilizados frecuentemente para los pronósticos de corto plazo. Sus mayores ventajas son bajo costo y simplicidad. Este método se basa sus pronósticos en promedios ponderados de mediciones pasadas. La lógica es que los valores pasados contienen información de lo que va a ocurrir en el futuro. Dado que los valores pasados incluyen fluctuaciones aleatorias, así como información acerca del patrón subyacente de la variable, se hace un intento por suavizar los valores. Este enfoque supone que las fluctuaciones extremas representan lo aleatorio en una serie de observaciones históricas.

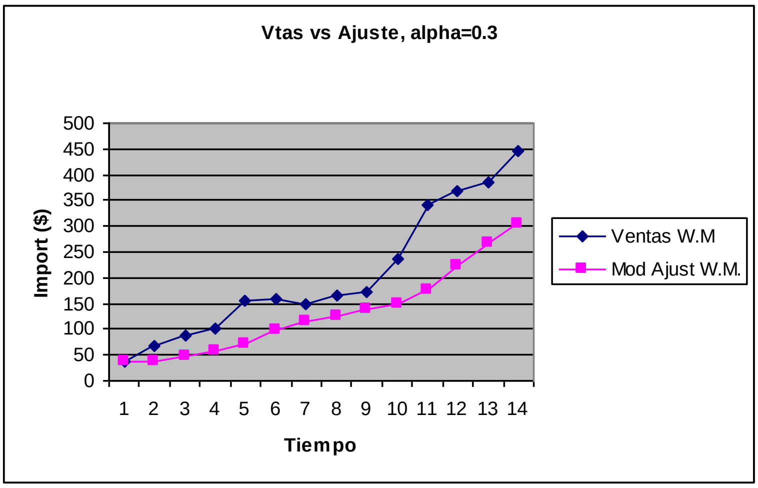
Por lo que el método del suavizamiento exponencial genera pronósticos que asignan pesos que disminuyen en forma exponencial conforme las observaciones se vuelven más antiguas.

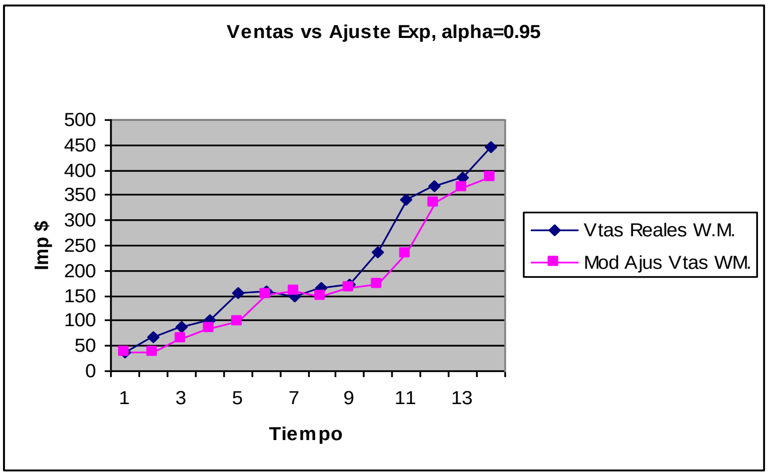
Los cálculos son los siguientes:

Suavizamiento Exponencial
Importe de Venta de un Autoservicio Wal*Mart

Año	Yt	Val Pronos SuaExp =0.1	Error Pronóst et	Val Pronos SuaExp =0.6	Error Pronóst et	Error SuaExp =0.1	Error SuaExp =0.95
1993	38	38.00	-	38.00	-	-	-
1994	67	38.00	29.00	38.00	29.00	841.00	841.00
1995	87	46.70	40.30	65.55	21.45	1,624.09	460.10
1996	100	58.79	41.21	85.93	14.07	1,698.26	198.04
1997	155	71.15	83.85	99.30	55.70	7,030.32	3,102.89
1998	160	96.31	63.69	152.21	7.79	4,056.79	60.61
1999	148	115.41	32.59	159.61	- 11.61	1,061.78	134.81
2000	165	125.19	39.81	148.58	16.42	1,584.80	269.60
2001	173	137.13	35.87	164.18	8.82	1,286.42	77.81
2002	238	147.89	90.11	172.56	65.44	8,119.21	4,282.53
2003	340	174.93	165.07	234.73	105.27	27,249.65	11,082.21
2004	367	224.45	142.55	334.74	32.26	20,321.15	1,040.94
2005	385	267.21	117.79	365.39	19.61	13,873.68	384.68
2006	445	302.55	142.45	384.02	60.98	20,292.18	3,718.64
					MSE:=	**8,387.64**	**1,973.37**

Las gráficas que describe las ventas reales contra el pronóstico, con diferentes parámetros son las siguientes:





El parámetro alpha=0.95, describe bastante bien el comportamiento de las ventas, sin embargo nos interesa optimizar el modelo, por lo que haremos el pronóstico con el segundo método:

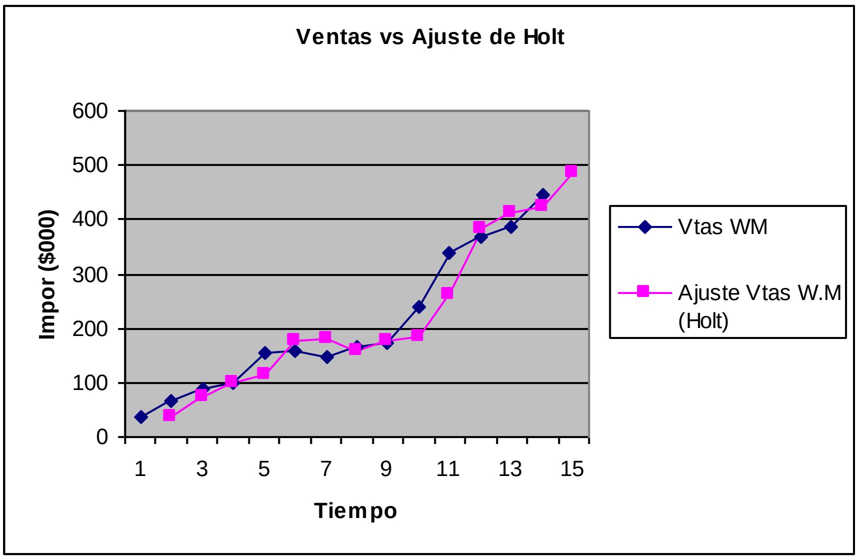
Suavizamiento Exponencial ajustado a la Tendencia (Holt o Exponencial doble)
Importe de Venta

Alpha= 0.30 Beta= 0.1
P= 1

Año	Yt	Lt	Tt	Yt+p	et	et(2)
1993	38	38.00	-			-
1994	67	65.55	8.27	38.00	29.0	841.00
1995	87	86.34	12.02	73.82	13.2	173.84
1996	100	99.92	12.49	98.36	1.6	2.68
1997	155	152.87	24.63	112.41	42.6	1,814.14
1998	160	160.87	19.64	177.50	- 17.5	306.19
1999	148	149.63	10.37	180.52	- 32.5	1,057.29
2000	165	164.75	11.80	160.00	5.0	25.00
2001	173	173.18	10.79	176.55	- 3.5	12.60
2002	238	235.30	26.19	183.97	54.0	2,919.78
2003	340	336.07	48.56	261.49	78.5	6,164.48
2004	367	367.88	43.54	384.64	- 17.6	311.11
2005	385	386.32	36.01	411.42	- 26.4	697.97
2006	445	443.87	42.47	422.33	22.7	513.99
2007	**486**	**486.34**	**42.47**	**486.34**	-	-
2008	**529**	**528.80**	**42.47**	**528.80**	-	-

MSE= 1,060.00

Con este método de Holt, se disminuye más de un 80% del error al cuadrado, y en la gráfica se observa que es mejor estimación que el suavizamiento exponencial, y un pronóstico para Wal Mart de corto plazo.



5.2 Comparación de las cifras reales del 2006 por mes, con las obtenidas en la estimación.

La comparación entre las ventas reales y el pronóstico esta basado en la diferencia de estos dos, conocido como error, en esta parte analizaremos como al modelo se le pueden hacer una serie de pruebas que nos validen que este es el mejor estimador del evento.

5.2.1 Validación del modelo

La validación del modelo consistirá en comprobar si se satisface la hipótesis del modelo ARIMA relativa a que $\{e_t\}$ es un proceso de ruido blanco. El inconveniente radica en que, dado los que los coeficientes del modelo son estimaciones, a partir de la muestra observada, de los verdaderos parámetros, los verdaderos errores e_t son desconocidos. Por consiguiente, la comprobación de las hipótesis de proceso de ruido blanco se realizará sobre una estimación de los mismos. Si $\hat{y}_t$ es la estimación de y_t, mediante el modelo ajustado, una estimación de la realización del proceso de ruido blanco $\{e_t\}$ será el residuo: $e_t = y_t - \hat{y}_t$

Recordemos que el error (residual de modelo_1), es el valor de las ventas menos el pronóstico del modelo ARIMA(1,1,1)X(1,1,1), su validación consistirá en comprobar que el error tiene media cero, que es estable en varianzas, que sus observaciones están incorreladas y que procede de una distribución Normal.

5.2.2 Media Cero y varianza constante

Respecto a la nulidad de la media y a la estabilidad en varianzas de la serie residual de la Fact., es consecuencia del método de estimación de los coeficientes del modelo, mientras que la segunda es consecuencia de la estabilidad en varianzas de la serie Fact., comprobada al inicio del análisis.

5.2.3 Incorrelación

Un proceso habitual para verificar que los residuos son incorrelados es incluir en el gráfico de la función de autocorrelación simple los límites del intervalo de confianza al 95% para el 0 y comprobar si todos los coeficientes de la función están en la banda. En promedio, cinco de cada 100 coeficientes de correlación estimados saldrán fuera. Por ello la aparición de algún coeficiente significativo es esperable. Sin embargo, un valor fuera de los límites de confianza en los retardos iniciales debería ser considerado como un indicio claro de que el modelo no es correcto. Un estadístico adecuado para contrastar la hipótesis de independencia de los residuos es el Box-Ljung que, para cada coeficiente de la función de autocorrelación simple, permite contrastar la hipótesis nula de que todos los coeficientes anteriores, hasta el correspondiente, son nulos.
Si el **p-valor** asociado es menor que α , se rechazará la hipótesis nula al nivel de significación α

El siguiente cuadro proporciona la ACF de la serie residual Fact.

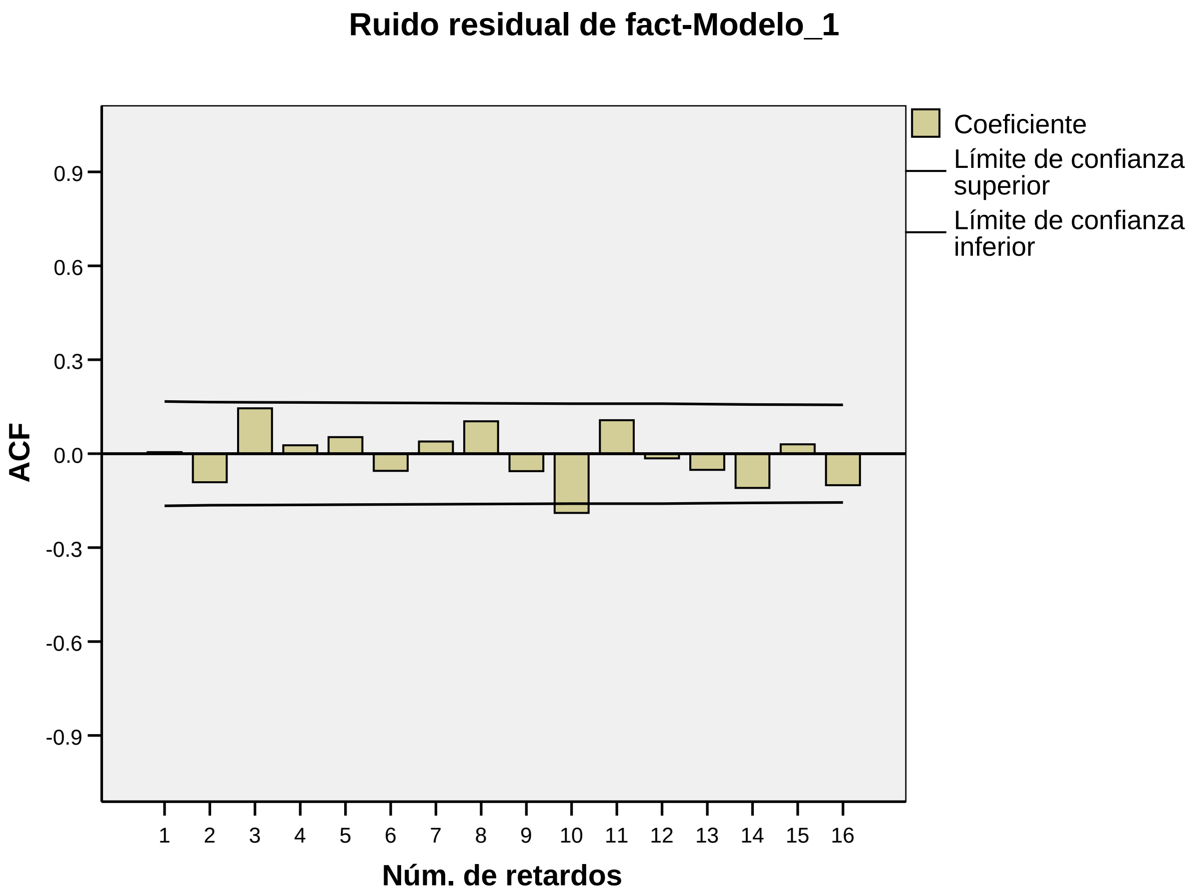
Ruido residual de fact-Modelo_1

Autocorrelaciones

Serie: Ruido residual de fact-Modelo_1

Retardo	Autocorrelación	Error típico(a)	Estadístico de Box-Ljung Valor	gl	Sig.(b)
1	.005	.083	.003	1	.955
2	-.091	.082	1.232	2	.540
3	.145	.082	4.349	3	.226
4	.027	.082	4.456	4	.348
5	.053	.081	4.875	5	.431
6	-.055	.081	5.333	6	.502
7	.039	.081	5.566	7	.591
8	.103	.080	7.215	8	.514
9	-.056	.080	7.699	9	.565
10	-.189	.080	13.309	10	.207
11	.107	.080	15.103	11	.178
12	-.015	.080	15.138	12	.234
13	-.052	.079	15.565	13	.273
14	-.110	.079	17.508	14	.230
15	.030	.078	17.654	15	.281
16	-.101	.078	19.320	16	.252

a El proceso subyacente asumido es la independencia (ruido blanco).
b Basado en la aproximación chi cuadrado asintótica.



Excepto por un valor en 11, todos los coeficientes de caen dentro de la banda de confianza de un 95%. Otro criterio es que todos los p-valores asociados al estadístico de Box-Ljung para cada retardo la probabilidad ***Sig(b)*** son lo suficiente grandes como para no rechazar la hipótesis nula de que, en cada caso, todos los coeficientes anteriores, hasta el correspondiente son nulos. En consecuencia, podemos concluir en general que los residuos están incorrelados.

5.2.4 Normalidad

Los residuos observados y los esperados bajo la hipótesis de normalidad deben ser parecidos, por lo que comprobar la hipótesis de normalidad se reduce a comparar la distribución de ambas variables residuales. Una alternativa para el estudio de la normalidad es el gráfico de probabilidad normal, a partir del que se puede observar la magnitud de las diferencias entre las distribuciones de las series residuales. Sin embargo, este tipo de gráfico no permite, salvo situaciones extremas de coincidencia entre las dos distribuciones, concluir que las diferencias no son estadísticamente significativas. Una prueba que permite contrastar la hipótesis nula de distribución Normal es la prueba de bondad de ajuste de Kolmogorov-Smirnov. Si el p-valor asociado al valor del estadístico de contraste es menor que α, se rechazará la hipótesis nula al nivel de significación α.

El siguiente cuadro proporciona la prueba de Kolmogorov-Smirnov para contrastar la hipótesis nula de normalidad de la serie residual Fact (resultado del cuadro anterior). El p-valor asociado al estadístico de contraste (Sig. Asintót bilateral=0.649) es lo suficientemente grande como para concluir que la diferencia observada entre la distribución Normal no es estadísticamente significativa. En consecuencia, no se puede rechazar la hipótesis de distribución Normal.

Pruebas no paramétricas

Prueba de Kolmogorov-Smirnov para una muestra

		Ruido residual de fact-Modelo_1
N		139
Parámetros normales(a,b)	Media	-2.2334
	Desviación típica	44.69467
Diferencias más extremas	Absoluta	.063
	Positiva	.058
	Negativa	-.063
Z de Kolmogorov-Smirnov		.737
Sig. asintót. (bilateral)		.649

a La distribución de contraste es la Normal.
b Se han calculado a partir de los datos.

5.3 Reporte estadístico de las cifras obtenidas en la proyección 2007-2009, comparado con la proyección que la compañía tiene proyectada para ser competitiva en los próximos 3 años. (PREDICCION)

Si la serie observada es una realización $(y_1,..y_T)$ de un proceso ARIMA(p,d,q)(P,D,Q)$_s$, el problema de la predicción consiste en obtener estimaciones de las observaciones futuras $y_{T+h} > 0$

Recordemos que la expresión de un proceso autorregresivo integrado de medias móviles de órdenes p,d,q,P,D yQ es:

$$\Phi_p(B^s)\varphi_p(B)(1-B)^d(1-B^s)^D Y_t = \Theta_Q(B^s)\theta_q(B)e_t + c$$

O, equivalentemente:

$$(1-\Phi_1 B^s - \ldots - \Phi_p B^{Ps})(1-\varphi_1 B - \ldots - \varphi_p B^p)(1-B)^d (1-B^s)^D Y_t =$$
$$= (1-\Theta_1 B^s - \ldots - \Theta_Q B^{Qs})(1-\theta_1 B - \ldots - \theta_q B^q) e_t + c$$

Desarrollando la ecuación anterior la observación $y_t, t \le T$, puede expresarse como:

$$y_t = \alpha_1 y_{t-1} + .. + \alpha_k y_{t-k} + e_t - \beta_1 e_{t-1} - \ldots - \beta_m e_{t-m} + c$$

Donde *k=p+d+s(P+D)* y *m=q+sQ*

Si $\hat{\alpha}_1$ y $\hat{\beta}_1$ son las estimaciones de α_1 y β_1, respectivamente, i=1,…k y j=1,…m, y $\hat{c}$ es la estimación de c, la estimación de $y_t, k < t \le T$, mediante el modelo ajustado será:

$$\hat{y}_t = \hat{\alpha}_1 y_{t-1} + .. + \hat{\alpha}_k y_{t-k} - \hat{\beta}_1 e_{t-1} - \ldots - \hat{\beta}_m e_{t-m} + \hat{c}$$

Donde:

$$e_{t-1} = \begin{cases} y_{t-1} - y_{t-1}\_donde\_t - i > k \\ 0\_donde\_t - i \le k \end{cases}$$

A partir de esta expresión, la predicción $\hat{y}_{T+h}$ para y_{T+h} puede calcularse como:

$$\hat{y}_{T+h} = \hat{\alpha}_1 \hat{y}_{T+h-1} + .. + \hat{\alpha}_{h-1} \hat{y}_{T+1} + \hat{\alpha}_h y_T + \ldots + \hat{\alpha}_k y_{T+h-k} - \hat{\beta}_h e_T - \ldots - \hat{\beta}_n e_{T+h-n} + \hat{c}$$

Esta ecuación indica que después de unos valores iniciales, la predicción queda determinada por la parte autorregresiva del modelo. Es decir, para *h > q*, las predicciones satisfacen la ecuación:

$$(1 - \hat{\alpha}_1 B - \ldots - \hat{\alpha}_k B^k) \hat{y}_{t-k} = 0$$
$$donde: \hat{y}_{T+h-i} = y_{T+h-i}; i = h, h+1, .., k,$$
$$si\_h \le k$$

Al construir un modelo con fines predictivos es conveniente dividir la serie en un período histórico o de estimación y un período de validación. El método consiste en desarrollar un modelo basado en las observaciones en el período histórico y generar a partir de él, predicciones para el período de validación. La comparación entre las predicciones y las observaciones en el período de validación permitirá evaluar la capacidad predictiva del modelo. Esta misma metodología permitirá comparar distintos ajustes para una misma serie, eligiéndose aquel modelo cuya capacidad predictiva sea mayor. Dado que el objetivo es comprobar la capacidad del modelo para predecir observaciones en un período futuro, en nuestro modelo el período de validación es un año y el de pronóstico es 3.

Hemos comprobado que el modelo ARIMA de la forma:

$$(1-\varphi_1 B)(1-B^{12})F_t = (1-\theta_1 B)(1-\Theta_1 B^{12})e_t$$

Es adecuado para predecir observaciones futuras de la serie Fact Las estimaciones de los parámetros del modelo se obtuvieron en el capítulo anterior, considerando únicamente las observaciones de enero de 1994 hasta diciembre del 2005 (período histórico) y, posteriormente se comprobó la capacidad predictiva sobre el año 2006 (período de validación). Para predecir los años 2007-2009 ajustaremos el modelo sobre todo el período observado (incluido el 2006), de forma que los datos más recientes también sean considerados.

La corrida arroja que los valores de los estimadores de los parámetros del modelo, ajustando sobre todo el período de observación, difieren muy poco de los de las estimaciones para el modelo ajustado hasta diciembre del 2005, lo que indica que en lo meses del 2006 se mantiene la estructura de los meses precedentes. El siguiente cuadro muestra un comparativo de los valores reales vs los valores estimados con el modelo 8.

Año 2006 MES	**ene**	**Feb**	**mar**	**abr**	**May**	**jun**	**jul**	**ago**	**sep**	**oct**	**nov**	**dic**	**tot**
Fact Real	15,229	52,614	98,340	62,572	102,712	129,098	108,692	243,591	220,942	477,811	460,630	41,492	2,013,723
Estimado ARIMA	13,052	60,573	101,968	86,737	118,370	135,743	130,456	181,123	188,435	476,645	422,616	45,703	1,961,422
Diferencia	2,177	- 7,959	- 3,628	- 24,165	- 15,658	- 6,645	- 21,764	62,468	32,507	1,166	38,014	- 4,211	52,301
% Variación	**16.68%**	**-13.14%**	**-3.56%**	**-27.86%**	**-13.23%**	**-4.90%**	**-16.68%**	**34.49%**	**17.25%**	**0.24%**	**8.99%**	**-9.21%**	**2.67%**

A continuación mostramos los valores obtenidos y su gráfica de pronóstico con su posible variación máxima y mínima, estos valores acumulados serán comparados contra las proyecciones de la compañía

FORECAST 2007

Modelo		Ene 2007	Feb 2007	Mar 2007	Abr 2007	May 2007	Jun 2007	Jul 2007	Ago 2007	Sep 2007	Oct 2007	Nov 2007	Dic 2007
Fac.-Modelo_1	Previsión	21307.63	65355.54	112892.8	79643.63	120738.4	146194.2	128548.1	247241.2	229895.7	507140.9	479637.1	51384.79
	LCS	50845.66	114499.8	176347.1	133876.4	186717.3	218598.6	197002.9	340621.3	320460.2	639597.3	609069.1	97189.85
	LCI	2821.24	27345.18	60654.94	36709.89	66141.10	85254.05	71640.05	165490.4	151043.2	386479.1	362082.2	17539.43

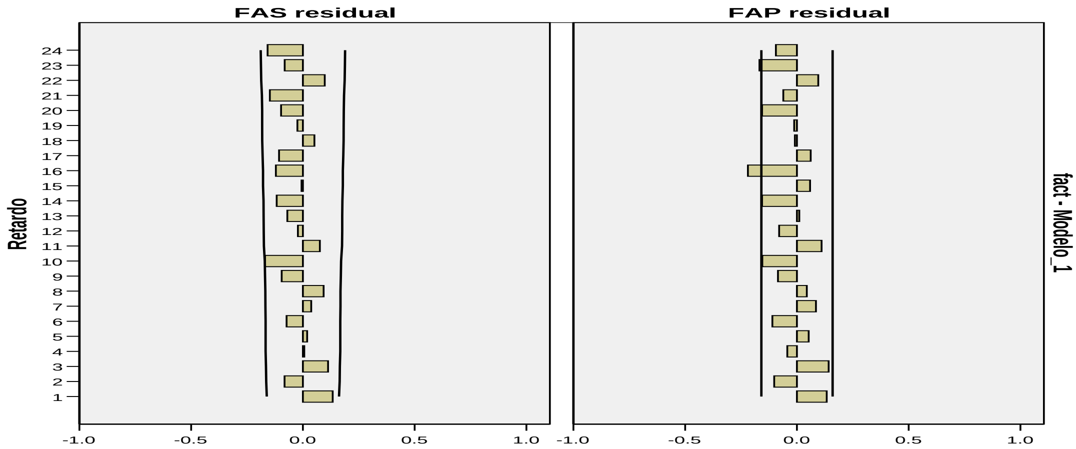
Para cada modelo, las predicciones comienzan después del último valor no perdido del rango del período de estimación solicitado y finalizan en el último período para el que hay disponibles valores no perdidos de todos los predictores o en la fecha de finalización del período de predicción solicitado, lo que ocurra antes.

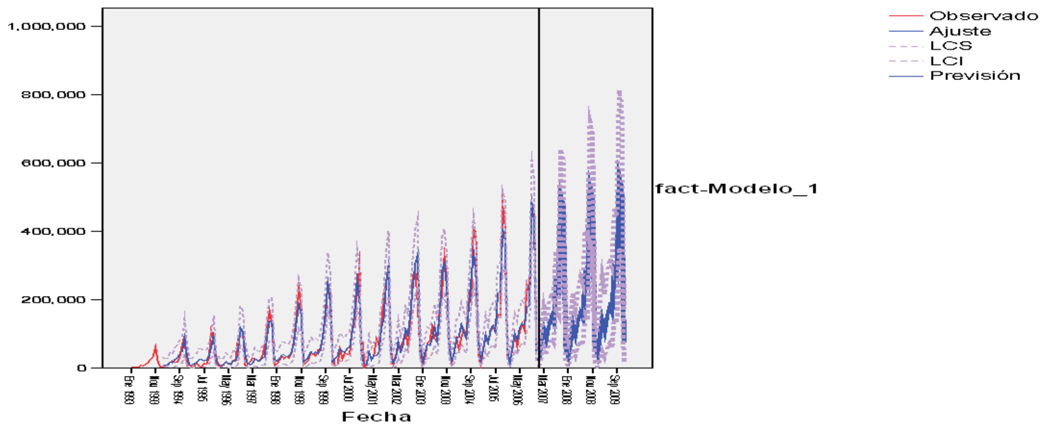
FORECAST 2008

Modelo		Ene 2008	Feb 2008	Mar 2008	Abr 2008	May 2008	Jun 2008	Jul 2008	Ago 2008	Sep 2008	Oct 2008	Nov 2008	Dic 2008
Fac.-Modelo_1	Previsión	28135.33	76305.43	126825.5	91627.70	135147.4	161908.9	143418.0	267211.3	249245.1	535004.6	506815.3	61536.22
	LCS	73753.13	148125.9	217933.0	170723.1	230207.5	265874.1	242383.5	400042.4	378639.9	721436.9	689619.0	131139.8
	LCI	1883.22	24098.48	55579.26	32641.32	60444.07	78548.23	65304.93	155480.4	141198.4	370168.2	345855.3	14024.07

FORECAST 2009

Modelo		Ene 2009	Feb 2009	Mar 2009	Abr 2009	May 2009	Jun 2009	Jul 2009	Ago 2009	Sep 2009	Oct 2009	Nov 2009	Dic 2009
Fac.-Modelo_1	Previsión	36141.94	88449.41	141967.5	104836.3	150796.2	178878.5	159558.2	288466.9	269895.3	564184.3	535324.8	73034.16
	LCS	101943.5	187509.1	265659.2	213635.2	280102.5	319790.8	294429.9	466752.5	444157.7	811612.4	778553.2	171772.9
	LCI	1027.31	20578.22	49965.79	28229.03	54182.94	71160.97	58382.71	144379.0	130332.2	351957.2	327798.9	10499.46





La Compañía planea crecer en facturación un 11% en 2007, un 9% en 2008 y un 9% en 2009, por lo que de acuerdo con esta expectativa de crecimiento y los datos obtenidos por el modelo de ARIMA, podemos encontrar el siguiente cuadro comparativo:

Año	Pronóstico Compañía	Pronóstico ARIMA	% Var
2007	$ 2,235,233	$ 2,189,980	2.07%
2008	$ 2,436,403	$ 2,383,181	2.23%
2009	$ 2,655,680	$ 2,591,534	2.48%

Es decir, nuestro pronóstico esta ligeramente por abajo del esperado por la Compañía, y se ve que entre más a futuro, va creciendo el diferencial, algo lógico por la naturaleza del fenómeno y la limitación que explicamos que tiene este tipo de metodología que es efectiva a corto plazo. Además los Ejecutivos de la Compañía tienen Licencias, Contratos, Inversiones, que nosotros aún no sabemos, y que no consideramos en el cálculo, lo que si queda claro es que no esta fácil para esta compañía crecer ya que la competencia esta cada vez, más cerrada y se deben enfocar a construir nuevos segmentos de negocio.

5.4 Recomendaciones al área comercial sobre que lineamientos se deben tomar en cuenta para obtener los resultados proyectados.

Las áreas que componen el área comercial en cualquier empresa deben de realizar pronósticos adecuados a las necesidades de mercado.

El personal de Ventas debe conocer además de la información general de sus Clientes, aspectos financieros, económicos y culturales, para poder hacer un mejor plan de negocio, que les permitan realizar mejores propuestas ya que las grandes empresas es lo que están desarrollando.

El área de Mercadotecnia, deberá recurrir cada vez más a las estadística para la elaboración de estudios de mercado, el consumidor esta más sensible al precio, y las marcas por si solas, ya no se imponen, por que cada vez el precio se vuelve un factor decisivo en un consumidor con más opciones de compra, lo que le permite sustituir fácilmente cualquier bien o servicio. Quizá la recomendación de un servidor sea crear conceptos propios, y no sólo adaptar los que vienen de E.U. y Europa. Un ejemplo de esto es por ejemplo la Serie del Chavo del 8, que no ha sido a mi perspectiva bien explotada, y que sería lago de lo que se hacia antes (desarrollos y producción nacionales)

Estás dos áreas deberán de concienciar a las áreas de Operaciones y Finanzas, a la primera que el producto a tiempo ya que la serie o película tienen cada día menos espacio en la mente del consumidor, y esto es vital en este negocio (como lo hemos mencionado todo el juguete es importado). Al área de Finanzas, deberá de buscar permitirle al área comercial extender los plazos de pago y adelantar los plazos de temporada, para tener antes el producto en el punto de venta, lo que le dará mayor posibilidad de venta, y mejor lugar en el anaquel.
Creo que una adecuada planeación viene de un pronóstico de ventas adecuado, lo que se debe de implementar y trabajar en comité por estas cuatro áreas de la empresa.

5.5 Elementos de planeación futura que se deben compartir a otras áreas de la empresa, basados en un pronóstico de ventas confiable

Para poder hacer un adecuado pronóstico de ventas las compañías deberán de desarrollar una metodología donde cierta información deberá de ser compartida, los reportes principales de cada área deberán contener información veraz y fácil de entender por todos los involucrados, por lo que podemos resumir en los siguientes formatos por área:

Ventas

- Histórico de ventas por cliente-mes
- Pronóstico de ventas por cliente y por mes
- Asignación por producto-Cliente
- Cuota/Forecast por mes
- Reporte de desplazamientos por producto-cliente
- Asignación de producto por canal de distribución.
- Barómetro de los precios de venta de los principales clientes
- Prioridades de Surtido por Cliente

Mercadotecnia

- Pauta publicitaria por línea-artículo
- Histórico de ventas en piezas de los principales productos continuos
- Estrategia de salida por canal de distribución
- Calendario de lanzamientos y principales objetivos por mes por lanzamiento
- Plan de salida de Close Outs (juguetes obsoletos)
- Expectativa de participación de mercado por marca
- Plan de productos exclusivos por Cliente-Canal de Distribución

Operaciones

- Sourcing por mes artículo (Reporte de llegadas de producto)
- Disponible actual y de obsoletos
- Tiempos de acondicionamiento, disponible, etiquetado y de envió a CEDIS
- Calendarización de cierres de venta, inventarios y

Finanzas

- Reportes Macroeconómicos Inflación, PIB, TIR
- Incremento promedio de lista de precios
- Cierre de la cobranza, reporte que permita saber si es posible incrementar algún Cliente, basándonos en su rentabilidad como tal
- Reporte del Costeo del LAB

Conclusiones

Conclusiones

Los pronósticos de ventas son cada vez más usados por todas la compañías, ya que facilitan la planeación de los egresos y estiman los posibles ingresos, y un mercado como el juguetero eso vale mucho, ya que como todo en la compañía cuyos datos utilizamos, requieren saber cuanto vendrán por producto, por mes, por Cliente para pedir con antelación a Oriente (6 meses o antes), las cantidades requeridas por la filial, el mes de llegada, coordinar todas la áreas para realizar los pedidos, los embarques, la publicidad, la promotoría etc.

Los pronósticos de ventas cada vez son más sofisticados y complejos, antes se basaban en el juicio ejecutivo, y nadie ponía en duda su capacidad para estimar, basados en métodos poco confiables. La estadística y lo sistemas cibernéticos han cambiado estas formas y cada vez se usan modelos matemáticos en Finanzas, Economía, Negocios, etc. El camino esta cada vez más avanzado, aunque es importante destacar que pronósticos complejos como la Metodología de Box-Jenkins, y otros métodos complejos, son usados por muy pocas compañías por su costo y tiempo de elaboración, sin embargo, con la economía de mercado, serán vitales para sobrevivir ya que la planeación es lo que hace que las grandes empresas sigan creciendo, y los pronósticos son la base de la planeación de cualquier empresa.

Cada vez se encuentra más similitud en los pronósticos de ventas de las compañías con las variables macroeconómicas. Estas variables como las ventas de juguetes, sufren de grandes variaciones y las amplias transformaciones de las economías del mundo, las hace cada vez más difíciles de estimar. El precio del petróleo, los picos de inflación, las políticas de los gobiernos, etc. Nos obligan a preguntarnos, ¿Deberemos de cambiar los pronósticos generados bajo un modelo basado generalmente en el juicio de quien pronostica? A menudo, el trabajo actual sobre metodología de pronósticos implica esta cuestión. No se puede decir con exactitud cual es el mejor método, sin embargo, dependerá de la naturaleza del bien a pronosticar el nivel de información, el tiempo y presupuesto al proyecto lo que puede definir la metodología para estimar.

El caso de la serie de tiempo del histórico de ventas del 1994 al 2006, tenía los suficientes datos para aplicar la metodología Box-Jenkins, pero por otro lado el no tener los Costos de Publicidad más que anuales, sólo nos permitió hacer la regresión lineal con las ventas anuales, y se demostró la alta correlación entre los Costos de Publicidad y las Ventas, incluso su importancia con las variables macroeconómicas, y al hacer una regresión lineal múltiple se incremento dicha correlación, y pienso que es correcto que estén muy relacionadas ya que el pico en el consumo de juguete es a final de año, y es donde las variables macroeconómicas están más presentes que nunca en la mente de todos los consumidores, y como necesariamente compraran juguetes, y existe una economía estable, una inflación controlada, y un expectativa de crecimiento económico, hay más créditos, menos tasas de interés, y el consumo interno se encuentra más animado a consumir. El histórico de ventas muestra la caída de facturación en los años de cambio de gobierno, por eso desde el último sexenio, se ha mantenido un crecimiento sostenido de la empresa (arriba de la inflación).

La compañía ha tenido un crecimiento en ventas muy controlado, han estado creciendo muy cercano a la inflación, sin embargo, el mercado ha crecido en mayor proporción, por lo que nuevos competidores se han adueñado de ese diferencial, y con marcas que la compañía en estudio ha dejado su representación. Al 31 de diciembre del 2006 estimamos tiene un participación de mercado alrededor del 18%, contra una del 25% que tenían en el 2004.

Otra cuestión importante para ellos ha sido mantener una fuerte presencia con el líder del mercado, Wal Mart, dándole muchas promociones, exclusivas de producto, lanzamientos, etc. sin embargo este cliente ha crecido a su proveedor más importante, el mismo, por lo que le quita a Hasbro la posibilidad de crecer con otros como Comercial Mexicana, Gigante, Soriana, Liverpool, etc.

Mattel que es el líder de esta categoría se maneja muy agresivo y cada vez más se separa de los demás, MGA, Leep Frog y otros crecen en sus categorías y hasta hoy el que pierde participación es la compañía juguetera analizada.

Hasbro deberá de desarrollar un nuevo modelo de negocio, con nuevas líneas de producto, disminuyendo sus márgenes de utilidad y porcentajes de operación, creemos ya lo esta haciendo y un adecuado pronóstico de los productos hot que tenga le salvará conservar el segundo lugar (generalmente de lo que tiene buenas ventas se le agotan y deja mucha hambre en el mercado), lo malo es que requiere de esos one-shot (películas, series de t.v., modas, etc), para mantenerse en el gusto del público, ya que su marcas propias, no le ayudan de la misma forma que a Mattel.

La metodología de Box-Jenkins por su metodología es la más adecuada para pronosticar las ventas mensuales de la venta de juguetes, los factores de estacionalidad, tendencia, ciclaje que se presentan en la venta juguetera son fácilmente eliminados y reconstruidos por este método, sin embargo por su complejidad y costo que representa utilizarlos en toda la planeación comercial, no es posible utilizarlo en todos los casos, y mucho menos por cualquier persona del área, por lo que deberemos en la práctica adaptar otros métodos más simples.

El mercado del juguete, requiere, y siempre requerirá de personas con un agudo juicio ejecutivo, ya que a pesar de lo completo de estos métodos, es necesario seguir con las opiniones de lo expertos, este mercado es tan cambiante que cada año aparecen nuevas variables que explican el comportamiento del mercado, y es imposible poderlas considerar todas en un modelo matemático, pero creo que vale la pena intentarlo.

Anexo 1

Fórmulas clave utilizadas en el presente trabajo:

Media de la muestra:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Desviación estándar de la muestra

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

Varianza de la muestra

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$

Valor Esperado

$$E(X) = \sum_{i=1}^{n} [X_i P(X_i)]$$

Estadístico Normal Z, como número de varianzas

$$Z = \frac{X - \mu}{\sigma}$$

Estadístico de prueba ***t*** (distribución ***t*** de Student)

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

Intervalo de confianza para la media de la población (muestra grande)

$$\bar{X} \pm Z \frac{S}{\sqrt{n}}$$

Coeficiente de correlación de la muestra

$$r = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{\sqrt{n\sum_{i=1}^{n}(X_i)^2 - (\sum_{i=1}^{n} X_i)^2}\sqrt{n\sum_{i=1}^{n}(Y_i)^2 - (\sum_{i=1}^{n} Y_i)^2}}$$

El método de mínimos cuadrados determina β_0 y β_1 que minimizan esta expresión para:

$$SSE = \sum_{i=1}^{n}(Y_i - \hat{Y})^2 = \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2$$

Ecuación de la pendiente de la recta de regresión ajustada:

$$\beta_1 = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^{\,2} - (\sum_{i=1}^{n} X_i)^2}$$

Ecuación de la intersección con el eje Y de la recta de regresión ajustada:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} = \frac{\sum_{i=1}^{n} Y_i}{n} - \frac{\beta_1 \sum_{i=1}^{n} X_i}{n}$$

Coeficiente de autocorrelación del k-ésimo orden:

$$r_k = \frac{\sum_{t=1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$$

Error Estándar del coeficiente de autocorrelación

$$SE(r_k) = \frac{1 + 2\sum_{i=1}^{n} r_i^2}{n}$$

Estadística de Q Ljung-Box

$$Q = n(n+2) = \frac{\sum_{k=1}^{n} r_k^2}{n-k}$$

Modelo Aleatorio

$$Y_t = \hat{Y}_t + e_t$$

Estadística para probar la significancia de la autocorrelación del retraso 1

$$t = \frac{r_1}{SE(r_1)}$$

Error del pronóstico:

$$e_t = Y_t - \hat{Y}_t$$

Desviación media absoluta:

$$MAD = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right|$$

Error cuadrático medio:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2$$

Error porcentual absoluto medio:

$$MAPE = \frac{1}{n}\frac{\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right|}{Y_t}$$

Error porcentual Medio

$$MPE = \frac{1}{n}\frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)}{Y_t}$$

Modelo Informal de Estimación:

$$\hat{Y}_{t+1} = Y_t$$

Modelo informal con tendencia:

$$\hat{Y}_{t+1} = Y_t + (Y_t - Y_{t-1})$$

Modelo informal de tasa de cambio:

$$\hat{Y}_{t+1} = Y_t \frac{Y_t}{Y_{t-1}}$$

Modelo informal estacional para datos trimestrales:

$$\hat{Y}_{t+1} = Y_{t-3}$$

Modelo informal estacional y con tendencia para datos trimestrales:

$$\hat{Y}_{t+1} = Y_{t-3} + \frac{Y_t - Y_{t-4}}{4}$$

Modelo de promedio simple

$$\hat{Y}_t = \frac{1}{t}\sum_{i=1}^{t} Y_i$$

Promedio simple actualizado, nuevo periodo:

$$\hat{Y}_{t+2} = \frac{t\hat{Y}_{t+1} + Y_{t+1}}{t+1}$$

Modelo de promedio móvil para k periodos:

$$\hat{Y}_{t+1} = \frac{1}{k}[Y_t + Y_{t-1} + Y_{t-2} + \ldots + Y_{t-k+1}]$$

Promedio móvil doble:

$$\hat{M}_t = \frac{1}{k}[M_t + M_{t-1} + M_{t-2} + \ldots + M_{t-k+1}]$$

$$a_t = 2M_t - \hat{M}_t$$

$$b_t = \frac{2}{k-1}(M_t - \hat{M}_t) \Rightarrow$$

$$\hat{Y}_{t+p} = a_t + pb_t$$

$$p = periodo$$

Suavizamiento exponencial simple:

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$$

$$\alpha = parámetro\_de\_suavizamiento < 1$$

Expresión alterna equivalente:

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + \ldots.$$

$$\alpha = parámetro\_de\_suavizamiento < 1$$

Suavizamiento lineal de Holt

Estimado del nivel actual de la serie suavizada exponencialmente:

$$L_t = \alpha Y_t + (1-\alpha)(L_t + Y_{t-1})$$

Estimado de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1})$$

$$\beta = Parámetro\_de\_tendencia$$

Pronóstico del periodo ***p*** en el futuro:

$$\hat{Y}_{t+p} = L_t + pT_t$$

$$p = periodo$$

Suavizamiento exponencial de Winters con estacionalidad multiplicativa:

Nivel estimado de serie suavizada exponencialmente:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

Estimado de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

Estimado de estacionalidad:

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma)S_{t-s}$$

Pronóstico para el ***p*** en el futuro:

$$\hat{Y}_{t+p} = (L_t + pT_t)S_{t-s+p}$$

$$p = periodo$$

Descomposición aditiva de las series de tiempo:

$$Y_t = T_t + S_t + I_t \,\_\_\_\, donde:$$

$$T_t : Tendencia$$

$$S_t : Estacional$$

$$I_t : Irregular$$

Descomposición multiplicativa de las series de tiempo:

$$Y_t = (T_t) * (S_t) * (I_t)$$

Tendencia Lineal:

$$\hat{T}_t = b_0 + b_1 t$$

Tendencia Cuadrática:

$$\hat{T}_t = b_0 + b_1 t + b_2 t^2$$

Tendencia Exponencial:

$$\hat{T}_t = b_0 b^t{}_1$$

Datos ajustados estacionalmente (descomposición multiplicativa)

$$\frac{Y_t}{S_t} = T_t * I_t$$

Componente irregular-cíclico (descomposición multiplicativa)

$$C_t * I_t = \frac{Y_t}{T_t * S_t}$$

Componente irregular (descomposición multiplicativa)

$$I_t = \frac{C_t * I_t}{C_t}$$

Poder actual de compra de un peso:

$$\frac{100}{IPC}$$

(IPC=Índice de Precios al Consumidor)

Valor de un Peso deflactado:

Valor del Peso X Poder de compra de un peso.

Regresión Lineal Simple

Método de Mínimos Cuadrados: fórmula de la pendiente:

$$\beta_1 = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^{\,2} - (\sum_{i=1}^{n} X_i)^2}$$

Método de Mínimos Cuadrados: fórmula para la intersección con el eje Y:

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} = \frac{\sum_{i=1}^{n} Y_i}{n} - \frac{\beta_1 \sum_{i=1}^{n} X_i}{n}$$

Relación entre la pendiente de la línea de regresión y el coeficiente de correlación:

$$\beta_1 = \frac{\sqrt{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}\,(r)}{\sqrt{\sum_{t=1}^{n}(X_t - X_t)^2}}$$

Ecuación de Regresión ajustada:

$$\hat{Y}_t = \beta_0 + \beta_1 X$$

Error estándar de la estimación: fórmula que lo define:

$$S_{y.x} = \sqrt{\frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}{n-2}}$$

Error estándar de la estimación: fórmula que lo define:

$$S_{y.x} = \sqrt{\frac{\sum_{t=1}^{n}(Y_t)^2 - \beta_0\sum_{t=1}^{n}Y_t - \beta_1\sum_{t=1}^{n}X_t Y_t}{n-2}}$$

Error estándar del pronóstico:

$$S_f = S_{y.x}\sqrt{1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{\sum_{t=1}^{n}(X - \bar{X})^2}}$$

Intervalo para la predicción:

$$\hat{Y} \pm ts_f$$

Intervalo para l_a predicción del 95% para muestra grande:

$$\hat{Y} \pm 2s_f$$

Descomposición de la suma de cuadrados y grados de libertad:

SST = SSR + SSE
Donde:
SST = Variabilidad total de la variable Y
SSR = Variabilidad explicada por la relación lineal
SSE = Residual o variabilidad no explicada

$$\sum_{t=1}^{n}(Y_t-\bar{Y}_t)^2=\sum_{t=1}^{n}(\hat{Y}_t-\bar{Y})^2+\sum_{t=1}^{n}(Y_t-\hat{Y}_t)^2$$
$$g.l:\_\_n-1=\_\_\_\_\_1\_\_\_\_\_\_\_\_\_n-2$$

Coeficiente de determinación:

$$r^2{}_1=\frac{\sum_{t=1}^{n}(\hat{Y}_t-\bar{Y})^2}{\sum_{t=1}^{n}(Y_t-\bar{Y})^2}=1-\frac{\sum_{t=1}^{n}(Y_t-\hat{Y})^2}{\sum_{t=1}^{n}(Y_t-\bar{Y})^2}$$

Estadística *t* para probar:

$$H_0:\beta_1=0$$
$$t=\frac{\beta_1}{S_{\beta_1}}$$

Error estándar para el coeficiente de regresión:

$$S_{\beta_1}=\frac{S_{y.x}}{\sqrt{\sum_{t=1}^{n}(X-\bar{X})^2}}$$

Estadística F

$$F=\frac{Cuadrado\_promedio\_de\_la\_regresión}{Error\_Cuadrático\_medio}=\frac{MSR}{MSE}$$

Relación de la estadística F con el coeficiente de determinación:

$$F=\frac{r^2(n-2)}{1-r^2}$$

Coeficiente de autocorrelación de los residuales:

$$r_k(e)=\frac{\sum_{t=k+1}^{n}e_te_{t-k}}{\sum_{t=1}^{n}e_t^2}$$

Regresión Lineal Múltiple

$$\hat{Y}_t=\beta_0+\beta_1X_1+\beta_2X_2+\ldots+\beta_kX_k$$

Descomposición de la suma de cuadrados y grados de libertad asociados:

$$\sum_{t=1}^{n}(Y_t - \bar{Y}_t)^2 = \sum_{t=1}^{n}(\hat{Y}_t - \bar{Y})^2 + \sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2$$
$$SST \_\_\_\_\_\_ = \_\_\_ SSR \_\_\_ + \_\_\_ SSE$$
$$g.l: \_\_ n-1 = \_\_\_\_\_ k \_\_\_\_\_\_\_\_\_ n-k-1$$

Error estándar de la estimación:

$$S_{y.x's} = \sqrt{\frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}{n-k-1}} = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{MSE}$$

Estadística *F* para probar la significación de la regresión:

$$F = \frac{MSR}{MSE}$$

Coeficiente de determinación:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{\sum_{t=1}^{n}(Y_t - \hat{Y})^2}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$$

Coeficiente de correlación múltiple:

$$R^m = \sqrt{R^2}$$

Relación de la estadística F con el coeficiente de determinación:

$$F = \frac{R^2}{1-R^2}\left(\frac{n-k-1}{k}\right)$$

Coeficiente ajustado de determinación:

$$\bar{R}^2 = 1 - (1-R^2)\left(\frac{n-1}{n-k-1}\right)$$

Estadística *t* para probar:

$$H_0 : \beta_j = 0$$
$$t = \frac{\beta_j}{S_{\beta_j}} \_\_ j = 1,2,3..,k$$

Pronóstico de un valor futuro:

$$\hat{Y}_t = \beta_0 + \beta_1 X^*{}_1 + \beta_2 X^*{}_2 + \ldots + \beta_k X^*{}_k$$

Intervalo de confianza para la predicción al utilizar una muestra grande para una respuesta futura:

$$(\hat{Y}^* - t_{\frac{\alpha}{2}} S_{y.x's}, \hat{Y}^* + t_{\frac{\alpha}{2}} S_{y.x's})$$

Factor de expansión de la varianza:

$$VIF_j = \frac{1}{1 - R_j^{\,2}} \_\_\_\_ \; j = 1,2,\ldots,k$$

Valores estandarizados de la variable independiente:

$$\bar{X} = \frac{X_{ij} - \bar{X}_j}{\sqrt{\sum_{i=1}^{n} (X_{ij} - \bar{X}_j)^2}} \qquad \text{j=1,2,…,k} \qquad \text{i=1,2,…,n}$$

Residual estandarizado:

$$\frac{e_i}{S_{e_i}} = \frac{e_i}{S_{y.x's}\sqrt{1 - h_{ij}}} \qquad \text{Donde } h_{ij} = \frac{1}{n} + \frac{(X_i - \bar{X})^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

Anexo 2

Anexo 2 Aplicaciones Estimación por diferentes métodos

Ejemplo 1

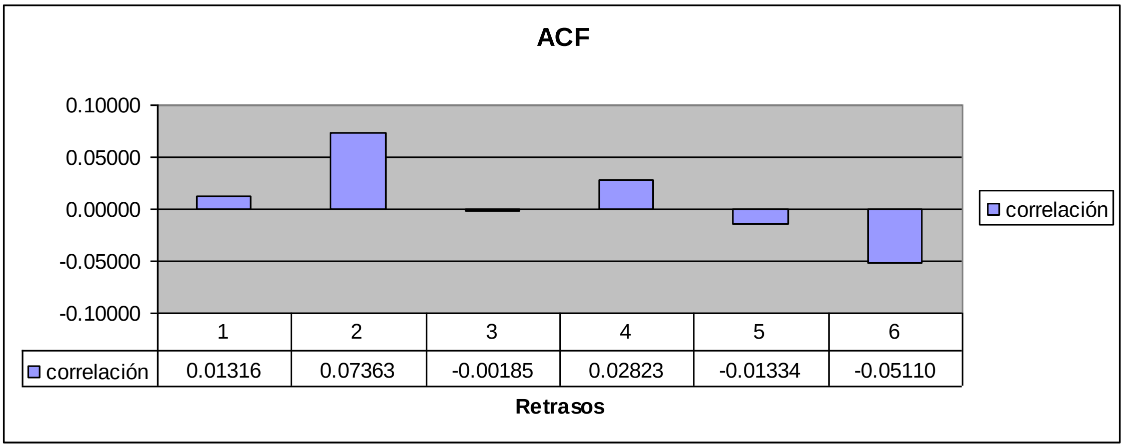
Cálculo de las ACF de un juguete preescolar cuyas ventas por unidad en los meses de ene-dic, son muy constantes, ya que todos los meses en proporción nacen bebés, y tienen un pico en diciembre por ser el mes de más venta de juguetes.

APLICACIÓN DE ACF

Ventas de Juguete Preescolar:
Número de Piezas Vendidas
2005

Mes	Y	Z
Ene	223	0.896
Feb	130	-2.209
Mar	175	-0.707
Abr	188	-0.273
May	195	-0.039
Jun	192	-0.139
Jul	191	-0.173
Ago	196	-0.006
Sep	197	0.028
Oct	207	0.362
Nov	200	0.128
Dic	260	2.131

LAG	ACF
1	0.01316
2	0.07363
3	-0.00185
4	0.02823
5	-0.01334
6	-0.05110

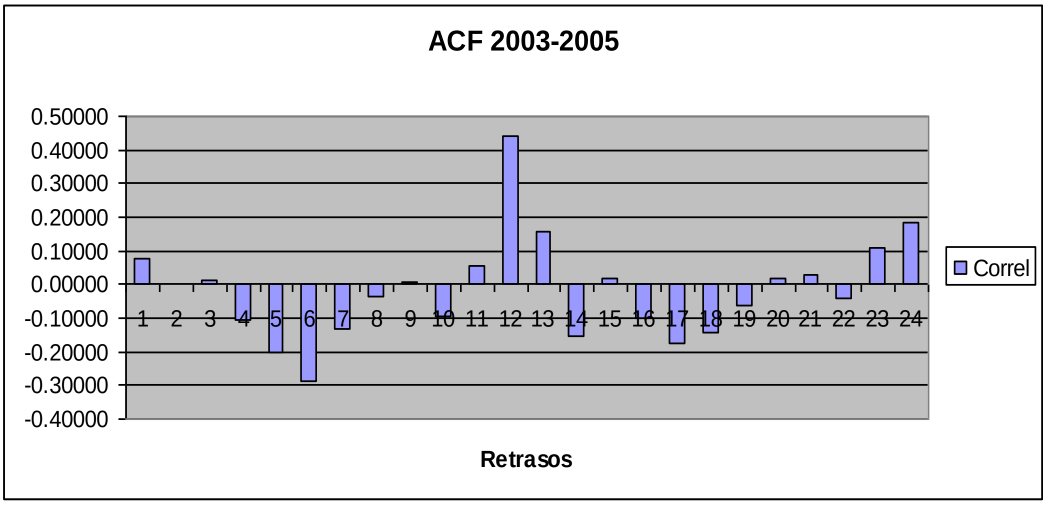


Ejemplo 2

El comportamiento de la ACF de este mismo juguete, pero con más períodos (24) de seguimiento y su gráfica.

Ventas de Juguete Preescolar:
Número de Piezas Vendidas
2003-2005

Mes	Pzas Vends	Dif
Ene	223	0.243
Feb	190	-0.498
Mar	232	0.445
Abr	296	1.883
May	142	-1.576
Jun	251	0.872
Jul	129	-1.868
Ago	179	-0.745
Sep	187	-0.565
Oct	168	-0.992
Nov	218	0.131
Dic	279	1.501
Ene	242	0.670
Feb	159	-1.194
Mar	210	-0.049
Abr	264	1.164
May	181	-0.700
Jun	192	-0.453
Jul	185	-0.610
Ago	177	-0.790
Sep	223	0.243
Oct	201	-0.251
Nov	237	0.558
Dic	297	1.905
Ene	284	1.613
Feb	196	-0.363
Mar	229	0.378
Abr	260	1.074
May	219	0.153
Jun	164	-1.082
Jul	197	-0.341
Ago	176	-0.812
Sep	149	-1.419
Oct	205	-0.161
Nov	207	-0.116
Dic	290	1.748

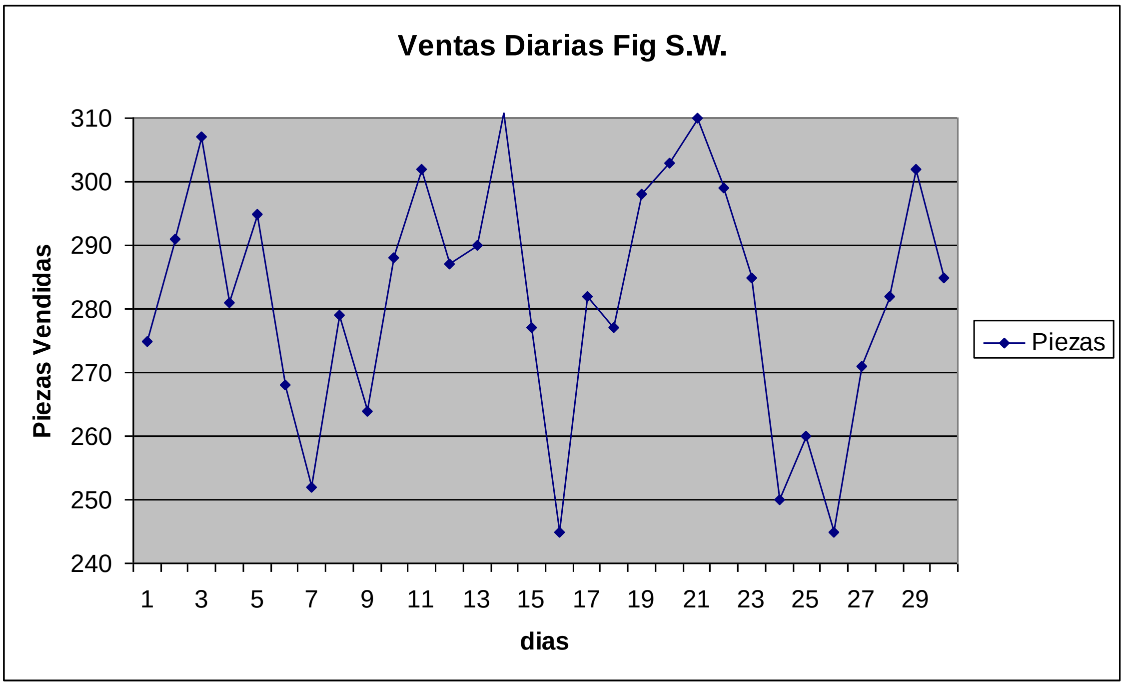


Ejemplo 3

El siguiente ejemplo muestra como algunos sistemas de administración tipo Retail Link, calculan el resurtido automático, en base a las ventas diarias de un producto, que para este ejemplo fueron las figuras de S.W., donde se hizo un seguimiento diario del comportamiento, y se estimo cuanto venderían los siguientes 3 días, para anticiparse a un posible desabasto.

Promedio Simple
Número de Piezas Vendidas diarias de Figuras de Novedad (Película)

T	Yt	Yt acum
1	275	275
2	291	566
3	307	873
4	281	1154
5	295	1449
6	268	1717
7	252	1969
8	279	2248
9	264	2512
10	288	2800
11	302	3102
12	287	3389
13	290	3679
14	311	3990
15	277	4267
16	245	4512
17	282	4794
18	277	5071
19	298	5369
20	303	5672
21	310	5982
22	299	6281
23	285	6566
24	250	6816
25	260	7076
26	245	7321
27	271	7592
28	282	7874
29	302	8176
30	285	8461



Dado que los datos son aparentemente estacionarios, usaremos los valores de los días 1 al 28 para pronosticar durante los días 29 y 30

El pronóstico para 29 es:

Y28+1=	281.21	
El error del pronóstico es:	Y29 -Y28+1=	20.79
El pronóstico para 30 es:		
Y28+2=	28*Y28+1+y29/29 =	281.93
El error del pronóstico es:	Y30 -Y28+2=	3.07
La venta para el día 31 sería:		
Y30+1	=sum(Y1:Y30)/30	282.03
La venta para el día 32, basándonos en 30 sería:		
Y30+2=	30*Y30+1+y30)/30 =	291.53

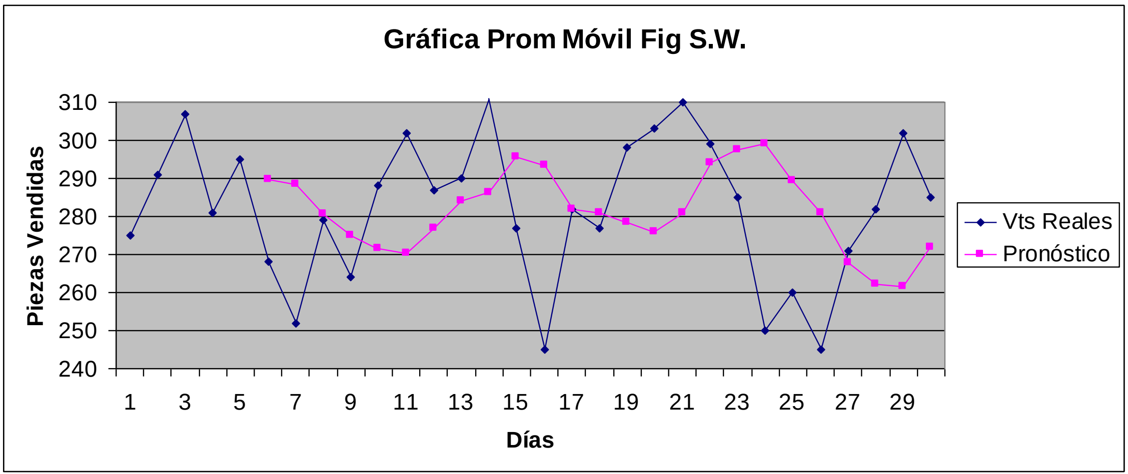
Con estos datos nuestro sistema de resurtido estima las posibles ventas y de continuar con esa tendencia de ventas, será necesario revisar inventarios, para en dado caso solicitar al Proveedor un resurtido, que alcance a cubrir las necesidades de venta de por lo menos dos semanas.

Ejemplo 4

Otra forma de calcular y analizar las ventas de productos es calcular el promedio de figuras vendidas por semana, y con esto calcular las ventas para toda la semana.

Promedios Móviles
Número de Piezas Vendidas diarias de Figuras de Novedad (Película)
Jun-05
Tomemos el promedio movil dela venta de 5 días

T	Yt	Yt (est)	Et
1	275		-
2	291		-
3	307		-
4	281		-
5	295		-
6	268	289.80	- 21.80
7	252	288.40	- 36.40
8	279	280.60	- 1.60
9	264	275.00	- 11.00
10	288	271.60	16.40
11	302	270.20	31.80
12	287	277.00	10.00
13	290	284.00	6.00
14	311	286.20	24.80
15	277	295.60	- 18.60
16	245	293.40	- 48.40
17	282	282.00	-
18	277	281.00	- 4.00
19	298	278.40	19.60
20	303	275.80	27.20
21	310	281.00	29.00
22	299	294.00	5.00
23	285	297.40	- 12.40
24	250	299.00	- 49.00
25	260	289.40	- 29.40
26	245	280.80	- 35.80
27	271	267.80	3.20
28	282	262.20	19.80
29	302	261.60	40.40
30	285	272.00	13.00
31	277.00	277.00	
32	283.40	283.40	
33	285.88	285.88	
34	286.66	286.66	
35	283.59	283.59	



Ejemplo 5
Cálculo de los Promedios Móviles Dobles
Importe de Venta de una cadena Departamental…Liverpool
Promedios Móviles considerando tres años

Día	Yt	Mt=prom M	et	et(2)
1	7		-	-
2	6		-	-
3	11		-	-
4	11	8.00	3.00	9.00
5	28	9.33	18.67	348.44
6	32	16.67	15.33	235.11
7	32	23.67	8.33	69.44
8	41	30.67	10.33	106.78
9	48	35.00	13.00	169.00
10	69	40.33	28.67	821.78
11	102	52.67	49.33	2,433.78
12	103	73.00	30.00	900.00
13	112	91.33	20.67	427.11
14	113	105.67	7.33	53.78
15			-	-
16				
MSE:				**464.52**

Como se observa tienen una tendencia Creciente, el pronóstico esta por debajo de las ventas reales a la cadena

Se procederá calcular los promedios móviles por segunda ocasión, para obtener un pronóstico más acertado.

Promedios Móviles Dobles…Liverpool

Día	Yt	Mt=prom M	Doble Mt=prom M	# Periodos k a	3 b	1 Pronóstico a+b p=1	et	2 Pronóstico a+b p=2	3 Pronóstico a+b p=3	et(2)
1	7						-			-
2	6						-			-
3	11	8.00					-			-
4	11	9.33					-			-
5	28	16.67	11.33	22.00	2.67		-			-
6	32	23.67	16.56	30.78	7.11	24.67	7.33	27.33	30.00	53.78
7	32	30.67	23.67	37.67	7.00	37.89	- 5.89	45.00	52.11	34.68
8	41	35.00	29.78	40.22	5.22	44.67	- 3.67	51.67	58.67	13.44
9	48	40.33	35.33	45.33	5.00	45.44	2.56	50.67	55.89	6.53
10	69	52.67	42.67	62.67	10.00	50.33	18.67	55.33	60.33	348.44
11	102	73.00	55.33	90.67	17.67	72.67	29.33	82.67	92.67	860.44
12	103	91.33	72.33	110.33	19.00	108.33	- 5.33	126.00	143.67	28.44
13	112	105.67	90.00	121.33	15.67	129.33	- 17.33	148.33	167.33	300.44
14	113	109.33	102.11	116.56	7.22	137.00	- 24.00	152.67	168.33	576.00
15						**123.78**		**131.00**	**138.22**	
16										-

MSE: **222.22**

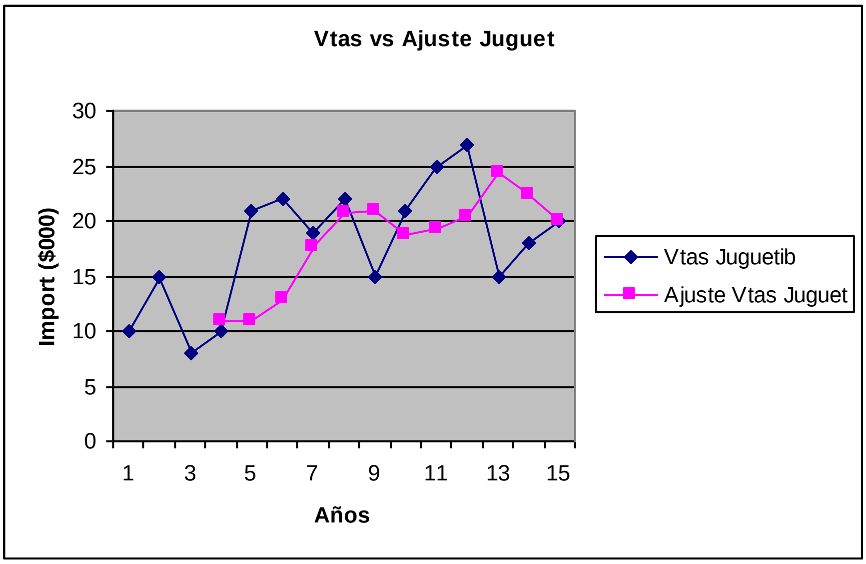
Ejemplo 6, Pronóstico de Ventas
Promedios Móviles Dobles Tres años

Importe de Venta de una cadena Especializada…Juguetibici
Promedios Móviles

Día	Yt	Mt=prom M	et	et(2)
1	10		-	-
2	15		-	-
3	8		-	-
4	10	11.00	- 1.00	1.00
5	21	11.00	10.00	100.00
6	22	13.00	9.00	81.00
7	19	17.67	1.33	1.78
8	22	20.67	1.33	1.78
9	15	21.00	- 6.00	

				36.00
10	21	18.67	2.33	5.44
11	25	19.33	5.67	32.11
12	27	20.33	6.67	44.44
13	15	24.33	- 9.33	87.11
14	18	22.33	- 4.33	18.78
15	20.00	20.00		-
16	17.67	17.67		
17	18.56	18.56		

MSE: **34.12**



Como se observa tienen una tendencia Creciente, el pronóstico esta por debajo de las ventas reales a la cadena

Día	Yt	Mt=prom M	Doble Mt=prom M	# Periodos k a	3 b	1 Pronóstico a+b p=1	et	2 Pronóstico a+b p=2	3 Pronóstico a+b p=3	et(2)
1	10						-			-
2	15						-			-
3	8	11.00					-			-
4	10	11.00					-			-
5	21	13.00	11.67	14.33	0.67		-			-
6	22	17.67	13.89	21.44	3.78	15.00	7.00	15.67	16.33	49.00
7	19	20.67	17.11	24.22	3.56	25.22	- 6.22	29.00	32.78	38.72
8	22	21.00	19.78	22.22	1.22	27.78	- 5.78	31.33	34.89	33.38
9	15	18.67	20.11	17.22	- 1.44	23.44	- 8.44	24.67	25.89	71.31
10	21	19.33	19.67	19.00	- 0.33	15.78	5.22	14.33	12.89	27.27
11	25	20.33	19.44	21.22	0.89	18.67	6.33	18.33	18.00	40.11
12	27	24.33	21.33	27.33	3.00	22.11	4.89	23.00	23.89	23.90
13	15	22.33	22.33	22.33	-	30.33	- 15.33	33.33	36.33	235.11

14	18	20.00	22.22	17.78	- 2.22	22.33	- 4.33	22.33	22.33	18.78
15						15.56		13.33	11.11	
16										-

MSE: 53.76

Esta cadena cada vez se le venderá menos con el paso del tiempo

Tres años

Ejemplo 8

Importe de Venta de una cadena Bod Comercial Mexicana

Promedios Móviles

Año	**Yt**	**Mt=prom M**	**et**	**et(2)**
1993	1		-	-
1994	4		-	-
1995	4		-	-
1996	7	3.00	4.00	16.00
1997	11	5.00	6.00	36.00
1998	10	7.33	2.67	7.11
1999	13	9.33	3.67	13.44
2000	8	11.33	- 3.33	11.11
2001	20	10.33	9.67	93.44
2002	22	13.67	8.33	69.44
2003	35	16.67	18.33	336.11
2004	34	25.67	8.33	69.44
2005	38	30.33	7.67	58.78
2006	41	35.67	5.33	28.44
2007				-
2008				

MSE: 61.61

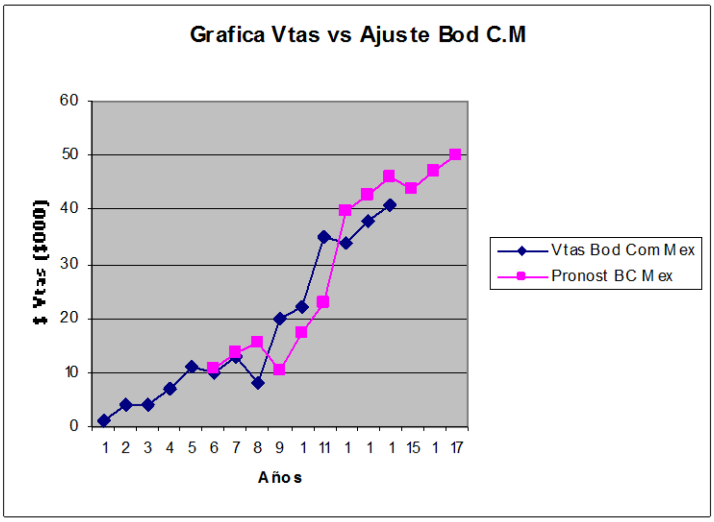
Como se observa tienen una tendencia Creciente, el pronóstico esta por debajo de las ventas reales a la cadena, se hará el cálculo por medias dobles:

Promedios Móviles Dobles

Año	Yt	**Mt=prom M**	**Doble Mt=prom M**	# Periodos k **a**	**3** **b**	**1** **Pronóstico a+b p=1**	**et**	**2** **Pronóstico a+b p=2**	**3** **Pronóstico a+b p=3**	**et(2)**
1	1						-			-
2	4						-			-
3	4	3.00					-			-
4	7	5.00					-			-
5	11	7.33	5.11	9.56	1.11		-			-
6	10	9.33	7.22	11.44	2.11	10.67	- 0.67	11.78	12.89	0.44
7	13	11.33	9.33	13.33	2.00	13.56	- 0.56	15.67	17.78	0.31
8	8	10.33	10.33	10.33	-	15.33	- 7.33	17.33	19.33	53.78
9	20	13.67	11.78	15.56	1.89	10.33	9.67	10.33	10.33	93.44

10	22	16.67	13.56	19.78	3.11	17.44	4.56	19.33	21.22	20.75
11	35	25.67	18.67	32.67	7.00	22.89	12.11	26.00	29.11	146.68
12	34	30.33	24.22	36.44	6.11	39.67	- 5.67	46.67	53.67	32.11
13	38	35.67	30.56	40.78	5.11	42.56	- 4.56	48.67	54.78	20.75
14	41	37.67	34.56	40.78	3.11	45.89	- 4.89	51.00	56.11	23.90
15						**43.89**		**47.00**	**50.11**	
16										-

MSE: **39.22**



Bibliografía

BIBLIOGRAFIA

María del Carmen González Videgaray
Modelos de Decisión con Procesos Estocásticos II
(Metodología de Box-Jenkins)
UNAM, ENEP Acatlán
México, 1990

White, R. Harry
Sales Executives Club of New Cork
Pronóstico de Ventas
Estrategías funcionales que ahorran tiempo y generan utilidades
Compañía Editorial Continental, S.A de C.V.
México 1989

Hurwood D.L, Grossman E.S & Bailey L.
Sales Forecasting
The Conference Board 845 Third ave New York 10022
New York, 1978

SPSS para Windows
Análisis Estadístico
Magdalena Ferrán Aranaz
Mc Graw Hill
México, 2001

Investigación de Mercados
Aplicación de Nuevas Técnicas
Paul E. Green
Ronald E. Frank
Limusa
México, 1985

Introducción a la Probabilidad e Inferencia Estadística
William Mendenhall
Robert J. Beaver
Bárbara M. Beaver
Internacional Thomson Editores, S.A. de C.V.
México, 2004

Estadística
Murray R. Spiegel
Segunda Edición
Mc Graw Hill
México, 2000

Probabilidad e Inferencia Estadística
Larzen, Harold
Limusa
México, 1993

Introducción a la Econometría
D. Gujaratí
México, 1992

Hasbro de México, S. de R.L. de C.V.
Reportes de Ventas y Sourcing de Producto
México, 1993-2006

Pronósticos en los negocios
John E. Hanke
Dean W. Wichern
Pearson Printice Hall
México, 2006

Introducción a la estadística para negocios
Ronald M. Weiers
Thomson Learning
México, 2006

Sistema Financiero de México
Eduardo Villegas Hernández
Rosa María Ortega Ochoa
Mc Graw Hill
México, 2005

Tiendas de Autoservicio
Alfredo D
Información, Análisis y Consultoría
Infobasic, S.A. de C.V.
Mexico D.F., Marzo 2007