

**COMPARACIÓN DE METODOLOGÍAS ESTADÍSTICAS PARA EL PRONÓSTICO
DE DEMANDA DE PRODUCTOS DE DIFÍCIL ESTIMACIÓN**

SEBASTIÁN RAMÍREZ AGUDELO

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA DE INGENIERÍA

FACULTAD DE INGENIERÍA INDUSTRIAL

MEDELLÍN

2013

COMPARACIÓN DE METODOLOGÍAS ESTADÍSTICAS PARA EL PRONÓSTICO
DE DEMANDA DE PRODUCTOS DE DIFÍCIL ESTIMACIÓN

SEBASTIÁN RAMÍREZ AGUDELO

Trabajo de grado para optar al título de Ingeniero Industrial

Directora

MARISOL VALENCIA CÁRDENAS

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA DE INGENIERÍA

FACULTAD DE INGENIERÍA INDUSTRIAL

MEDELLÍN

2013

Nota de Aceptación

Firma
Nombre
Presidente del jurado

Firma
Nombre:
Jurado

Firma
Nombre
Jurado

Medellín, Agosto ** de 2013

"... Mal se paga al maestro si se permanece siempre discípulo.

¿Por qué no vais a deshojar vosotros mi corona?

Vosotros me veneráis, mas,

¿qué ocurrirá si vuestra veneración se derrumba?

¡Cuidad de que no os aplaste mi estatua!..."

(Friedrich Nietzsche - Also sprach Zarathustra. Ein Buch für Alle und Keinen)

AGRADECIMIENTOS

Agradezco a la directora de este trabajo de grado, Marisol Valencia Cárdenas por su colaboración y compromiso conmigo y el resto de estudiantes. La misión de docente es una tarea difícil que influye enormemente en las personas y que repercute por ende en la sociedad; dejar huella en los estudiantes y animarlos a que continúen adelante prestando siempre un apoyo incondicional sin esperar nada más a cambio que la satisfacción del otro y la propia, es lo que lleva la labor docente a sus pilares originales y a su razón de existencia. Muchas gracias

También agradezco al resto de docentes que me ayudaron a lo largo de este camino que por cierto, apenas empieza, por haberme mostrado distintas vías del conocimiento, vías que cada vez no hacen más que dejarme percibir la infinitud de éste y el poco tiempo que se tiene para tomar algunas pequeñas partes de él.

No podría faltar agradecer el apoyo prestado por parte de la familia, sin el cual nada de esto sería posible, pues el círculo familiar es incondicional e indiscriminado y no busca más que oportunidades para guiar sus miembros por el camino que se interpreta como el mejor para continuar.

A los amigos muchas gracias por darme un pedazo de su amistad, permitirme hacer parte de un momento de sus vidas y haber permitido que yo entrara en las suyas.

Contenido

RESUMEN	11
INTRODUCCIÓN	12
OBJETIVOS	16
ALCANCE	16
1. ANTECEDENTES.....	17
1.1. Estado del arte	17
1.2. Pronóstico de demanda	19
2. MARCO TEÓRICO	22
2.1. Variable aleatoria	22
2.2. Software generadores de variables aleatorias	24
2.3. Distribución Poisson.....	24
2.4. Distribución Gamma.....	26
2.5. Prueba de bondad de ajuste ji-cuadrado.....	27
2.6. Prueba de Box Pierce	28
2.7. Estadística bayesiana	29
2.8. Monte Carlo por Cadenas de Markov (MCMC)	32
2.9. Regresión gaussiana bayesiana por Cadenas de Markov.	34
2.10. Regresión bayesiana Poisson por cadenas de Markov	35
2.11. Método de optimización Tabú	36
2.12. Error porcentual medio absoluto (MAPE).....	36
3. METODOLOGIA	38
4.1. Prueba de Bondad de ajuste y variable aleatoria.....	39
4.2. Regresión Gaussiana y Poisson bayesiana con MCMC.	39
4.3. Prima Bayesiana	41
4.4. Método de optimización Tabú	45
4. RESULTADOS	47
5. SINTESIS DE RESULTADOS	59
6. PROCEDIMIENTOS PARA REALIZAR PRONÓSTICOS UTILIZANDO LOS MÉTODOS PROPUESTOS	61
CONCLUSIONES.....	64
RECOMENDACIONES	65
REFERENCIAS.....	66

LISTA DE FIGURAS

Gráfico 1. <i>Distribución Poisson para valores de $\lambda=1-9$</i>	25
Gráfico 2. <i>Distribución Gamma para algunos valores α y β</i>	27
Gráfico 3. <i>Tiempo vs ventas 2010 para producto A</i>	47
Gráfico 4. <i>Función de autocorrelación muestral para 2010, producto A</i>	48

LISTA DE TABLAS

Tabla 1. <i>Diferencias entre teoría clásica y bayesiana</i>	18
Tabla 2. <i>Algunas distribuciones conjugadas</i>	32
Tabla 3. <i>Prueba de bondad de ajuste ji-cuadrado para producto A</i>	48
Tabla 4. <i>Prueba de bondad de ajuste ji-cuadrado para todos los productos</i>	48
Tabla 5. <i>Pronóstico mensual de 2010 y APEt para una iteración</i>	49
Tabla 6. <i>MAPE para todos los productor con variable aleatoria simulada</i>	50
Tabla 7. <i>Media de los parámetros y desviación estándar para modelo de regresión Poisson con MCMC</i>	50
Tabla 8. <i>Modelos de regresión bayesiana Poisson para todos los productos y MAPE</i>	51
Tabla 9. <i>Media de los parámetros y desviación estándar para modelo de regresión bayesiana Gaussiana con MCMC</i>	52
Tabla 10. <i>Pronóstico, APEt y MAPE mensual para el 2010 estimado con regresión bayesiana Gaussiana con MCMC</i>	52
Tabla 11. <i>Modelos de regresión bayesiana Gaussiana para todos los productos y MAPE</i>	53
Tabla 12. <i>Valores para variables del modelo de prima bayesiana, APEt y MAPE mensual para 2010</i>	54
Tabla 13. <i>MAPE 2011 para todos los productos obtenidos con prima bayesiana</i>	55
Tabla 14. <i>Valores α y β optimizados con Tabú</i>	55
Tabla 15. <i>Modelos de prima bayesiana con parámetros α y β optimizados para producto A en 2010</i>	56
Tabla 16. <i>Valores para variables del modelo de prima bayesiana con parámetros α y β optimizados, APEt y MAPE mensual para producto A en 2010</i>	57

Tabla 17. <i>MAPE 2011 para todos los productos obtenidos con prima bayesiana con optimización Tabú</i>	58
Tabla 18. <i>Tabla resumen para todos los modelos y productos</i>	59
Tabla 19. <i>Tabla comparativa entre modelos explorados</i>	60

GLOSARIO

Estacionariedad: Proceso cuya media y varianza es constante en el tiempo y cuya covarianza solo depende de k (González & Franco, 2013).

Monte Carlo por Cadenas de Markov: Metodología para generar un muestreo a partir de distribuciones de probabilidad basadas en la construcción de cadenas de Markov donde cada valor simulado, tiene dependencia con el dato anterior, llegando a una convergencia a la distribución deseada (Gill, 2002).

Optimización Tabú: Método que optimiza una función objetivo con ciertas variables de decisión haciendo una caminata alrededor de un punto de forma que encuentre el mínimo valor de la función objetivo y así saltar a otro punto para mejorar cada vez los valores encontrados para la función objetivo (Glover, 1990).

Tendencia: Disposición de una serie de datos en el tiempo que define un camino predecible de manera general. Como el de una ecuación lineal que define una tendencia lineal.

Distribución a priori: Representa la información sobre un parámetro incierto que al combinarse con la función de verosimilitud se halla la distribución a posteriori. Se obtiene por expertos con análisis subjetivo a partir de teoría estadística de probabilidad (Gill, 2002).

Distribución a posteriori: Distribución condicional de la probabilidad, después de la observación de datos, la cual resume el estado de conocimiento sobre cantidades inciertas en el análisis bayesiano. La distribución a posteriori es proporcional al producto de la distribución a priori y la función de verosimilitud (Gill, 2002).

Función de verosimilitud: Distribución conjunta de la muestra aleatoria de los datos (Montgomery, D. C., & Runger, G. C., 2003).

MAPE: *Mean Absolute Percentage Error*. Relación matemática que permite medir la eficiencia de pronósticos realizados a partir del porcentaje del valor absoluto de la diferencia entre el dato real y el pronosticado, todo dividido sobre el dato real.

RESUMEN

Una de las dificultades a las que se enfrentan organizaciones es la existencia de pocos datos históricos de venta o incluso, la falta de ellos para la planeación de sus inventarios y por tanto, su logística. En este estudio se comparan cuatro métodos estadísticos para el pronóstico de ventas futuras a partir de pocos datos, con una aplicación a un caso real. El primero de ellos es una generación de variable aleatoria después de realizar la prueba de bondad de ajuste Ji-cuadrado y con la ayuda de una simulación iterativa, el segundo, la aplicación de la metodología de regresión bayesiana específicamente al caso Normal conjugado es decir, el modelo de regresión gaussiana y Poisson bayesiana con las funciones del paquete *MCMCPack* del software estadístico R, el tercero es la metodología de Prima bayesiana, específicamente aplicado al caso que se obtiene con las distribuciones Gamma-Poisson, y finalmente, se optimiza el pronóstico realizado con el modelo encontrado de prima bayesiano usando el método de optimización Tabú, el cual toma como función objetivo el MAPE, indicador usado como criterio final de comparación de la eficiencia.

INTRODUCCIÓN

Este trabajo propone una investigación sobre métodos estadísticos para el pronóstico de ventas futuras en situaciones de difícil estimación, aplicable a todo tipo de organización aunque enfocado principalmente a productos y empresas con poco tiempo en el mercado, estableciendo una comparación entre sus ventajas, desventajas y criterios de error encontrados en el caso estudiado. En este tipo de productos y empresas los modelos estadísticos clásicos, normalmente los más populares, no se pueden aplicar de una manera efectiva puesto que no se cuenta con una amplia cantidad de datos históricos, haciéndose necesaria la creación y aplicación de modelos para pronosticar con pocos datos.

El aporte de este trabajo, será precisamente para esas empresas nuevas en el mercado que están interesadas en reducir la incertidumbre de su futuro económico y comercial, al tener disponible diferentes métodos de pronóstico que aborden este problema.

En la primera sección de Antecedentes, se hace una exploración bibliográfica sobre algunos tipos de modelos utilizados para pronosticar, teniendo muy en cuenta los más difundidos; y se habla sobre la importancia del pronóstico de la demanda: cuáles son sus ventajas y por qué es necesaria.

La segunda sección explorará el marco teórico que servirá de base para el desarrollo de este estudio, revisándose conceptos como variable aleatoria, partiendo desde sus métodos para generar números aleatorios; también se revisan algunos software generadores de estas variables aleatorias desde ciertos principios de programación computacional. Se repasarán distribuciones de probabilidad como la *Poisson* y la *Gamma*, así como las pruebas de bondad de ajuste Ji-cuadrado (χ^2), necesaria para conocer el tipo de distribución a la que se ajustan los datos utilizados; también se repasa la prueba de incorrelación residual *Box-Pierce* para probar suficiencia de algunos de los modelos analizados como los de regresión gaussiana bayesiana y regresión Poisson bayesiana, estudiados posteriormente.

Se continuará el marco teórico con la definición del teorema de Bayes, desde su aplicación más simple con el uso de variables discretas, hasta su generalización con variables continuas que implicarían el uso de distribuciones de probabilidad y análisis matemáticos más engorrosos con integrales. En esta parte también se

muestra la importancia de los conceptos como la distribución a posteriori, que se define, como proporcional a la función de verosimilitud multiplicada por la distribución a priori, la cual, a su vez, se selecciona subjetivamente a partir de expertos. También se dan razones por las cuales este tipo de estadística no tuvo tanta popularidad sino solo hasta hace pocos años, con el auge de potentes sistemas de cómputo y de software estadístico especializado, que permitieron resolver muchos problemas analíticos que sin esta ayuda resultaban casi imposibles; teniendo también como obstáculo su marca histórica de subjetividad, contaría a la pretensión de un tipo de objetividad reinante en las ciencias, la cual implicó el olvido por años de esta importante teoría estadística-matemática que ha logrado alcanzar protagonismo por sus muy buenos resultados en los últimos tiempos.

Se sigue con la revisión de Monte Carlo con Cadenas de Markov, metodología que permitió resolver muchos de los problemas analíticos que presentaba la estadística bayesiana y la clásica, lo que ha permitido ajustar prácticamente cualquier modelo de probabilidad, así como optimizar modelos para pronóstico. Se abordan luego los modelos ya mencionados de regresión gaussiana bayesiana y de regresión bayesiana Poisson por MCMC. Estos modelos se basan en distribuciones asignadas a priori para los parámetros, así como distribuciones para la verosimilitud de los datos con el fin de generar distribuciones a posteriori condicionales, de las cuales se hace muestreo con la cadena MCMC para encontrar muestras para los coeficientes del respectivo modelo, cuyos parámetros pueden ir acompañados de covariables de tiempo, trigonométricas y ficticias como el mes para ayudar a captar mejor la variabilidad de la serie. Luego, se estudia el método de optimización Tabú, el cual optimiza una función objetivo con ciertas variables de decisión haciendo una caminata alrededor de un punto, con el fin de encontrar el mínimo valor de esta función y así saltar a otro punto y mejorar los valores encontrados. Después se termina esta sección definiendo el error porcentual medio absoluto (MAPE, por sus siglas en inglés), indicador con el cual se evaluará la eficiencia de cada uno de los métodos estudiados, comparando datos pronosticados y los reales.

Más adelante en la tercera sección, se presenta la metodología con la cual se obtendrán los resultados de este estudio, empezando con la Prueba de Bondad de ajuste en relación a la distribución Poisson, por intuirse que ésta es la distribución a la que se justarán los datos, debido a su naturaleza discreta. La aplicación de los modelos de regresión Gaussiana y Poisson bayesiana con MCMC se realiza a partir de la incorporación de covariables de tiempo y trigonométricas. Se continúa con el despliegue analítico de la prima bayesiana a partir de una función de verosimilitud Poisson y distribución a priori Gamma, llegando a un modelo que

tiene en cuenta cambios dinámicos, al incorporar términos autorregresivos, esto con el fin de intentar captar fluctuaciones en las ventas futuras teniendo en cuenta el pasado. En el método Tabú se explica el algoritmo teniendo en cuenta la función objetivo, (que en este caso es el MAPE), sus respectivas restricciones y las variables incorporadas que fueron previamente definidas.

La cuarta sección desplegará los resultados a partir de la metodología propuesta, con datos de una empresa textil que prefirió mantener su confidencialidad. Todos los métodos estudiados se estimarán con 12 datos mensuales del año 2010, con el fin de pronosticar el 2011 y proceder con el cálculo del MAPE para seis tipos de productos. El primero genera pronósticos a partir de una variable aleatoria Poisson, después de encontrar, por medio de la prueba de ji-cuadrado, que a esta distribución se ajustan los datos; luego se realizará el pronóstico con la ayuda de una simulación de 1000 iteraciones.

El modelo de regresión bayesiana gaussiana se estima con la ayuda del paquete MCMC del software estadístico R concretamente con la función MCMCRegress llevando a cabo una simulación de 100.000 iteraciones y un quemado de 20.000.

El modelo de regresión Poisson se estima con el mismo paquete MCMC y con la función MCMC Poisson con la misma cantidad de iteraciones y quemado que el modelo de regresión gaussiana.

El modelo bayesiano se estima con la prima bayesiana que se halla a partir del valor esperado y la distribución a posteriori, que a su vez tiene como función de verosimilitud la distribución Poisson y como a priori la distribución Gamma ya que, en la distribución Poisson normalmente se asume λ como teniendo una distribución gamma (Neelamegham & Chintagunta, 1999).

Por último, en la presentación de resultados, se aplica la optimización Tabú con la cual se intenta optimizar aún más la función MAPE a partir de los parámetros de la distribución Gamma.

En la quinta sección se presentan una síntesis de los resultados.

La sexta sección presenta la manera de utilizar algunos de los métodos estudiados para predecir ventas futuras, con el fin de que el lector interesado pueda aplicarlos, sea en su empresa o como objetivo de investigación.

Por último, se presentan las conclusiones a las que se llegan después de alcanzar los resultados.

Queda decir entonces, que esta exploración de métodos estadísticas para productos de difícil estimación o de pocos datos, no termina aquí y se anima a

otros estudiantes e investigadores a extender este estudio, como lo es explorando otras combinaciones de distribuciones a posteriori partiendo de otras funciones de verosimilitud y distribuciones a posteriori, entre otro tipo de variación y profundización de este tipo de estudio, realizado como labor investigativa y como requisito para obtener el título de grado de ingeniero industrial.

OBJETIVOS

Objetivo General

Comparar metodologías estadísticas para el pronóstico de demanda de productos de difícil estimación.

Objetivos específicos

1. Explorar diferentes técnicas estadísticas para el pronóstico de demandas.
2. Aplicar dichas metodologías en situaciones donde se presenten datos históricos.
3. Comparar los métodos de pronóstico de demanda por medio de su eficiencia.
4. Elegir la metodología que muestre una mejor optimización describiendo el procedimiento para aplicarla en la estimación de la demanda de productos.

ALCANCE

Se proponen alternativas que son poco usuales para pronosticar ventas, dada una situación frecuente en muchas empresas: la existencia de pocos datos. Dos de las técnicas mostradas proponen alternativas bayesianas con los que puede hacerse una planeación de ventas y por tanto, de inventario de manera anticipada con el fin de programar producción con suficiente tiempo de antelación.

Se obtiene al final de la exploración de cuatro técnicas una propuesta estadística eficiente para la estimación de pronóstico de la demanda de producto terminado con pocos datos históricos, aplicado a un caso empresarial, explicando en detalle los pasos a seguir.

Este trabajo constituye entonces un insumo en la planeación y optimización de la demanda para mejorar los inventarios, o en general quienes estén interesados en estimar la venta futura de uno o varios productos.

1. ANTECEDENTES

1.1. Estado del arte

Algunos de los modelos de pronóstico clásicos basan sus estimaciones en distribuciones probabilísticas, y apoyados en pruebas de significancia, intentan concluir sobre el comportamiento de las variables involucradas y pronosticar respuestas. Este tipo de modelos han rechazado fuertemente cualquier uso subjetivo de la información, aunque Ayçaguer & Benavides (2003) apuntan que éstos incluyen de una forma u otra información subjetiva como lo son las escalas de medición y niveles de significancia. En cuanto a los métodos subjetivos, estos basan su inferencia en juicios intuitivos y opiniones de distintas personas; entre estos se encuentra el método Delphi, que consiste en recopilar información a partir de expertos, lo que permitiría pronosticar valores futuros; Otras técnicas, no muy aplicadas hasta ahora para pronósticos que faciliten la optimización de inventario son las asociadas con la Estadística Bayesiana, basada en una combinación de probabilidades objetivas con subjetivas; su desarrollo “se puede comparar al de la estadística clásica, no así su aplicación ni el conocimiento generalizado que se tiene de ella” (Serrano, 2003).

En algunos modelos estadísticos sobre pronósticos de demanda, tradicionalmente utilizados en la optimización de inventarios, se ha evidenciado en su aplicación una continua violación en la validación de supuestos en los errores como se observa en Hillier y Lieberman (2002), donde se hace un énfasis, principalmente en el proceso matemático y estadístico para realizar los pronósticos y no en los supuestos que deben cumplir los errores como lo es que posean normalidad.

Por ejemplo en los modelos de regresión lineal simple y múltiple los cuales caracterizan tendencias lineales para pronosticar usando una o varias variables regresoras o explicativas, Winston (2004) comprueba dichos supuestos de manera gráfica y no haciendo uso de las pruebas de significancia existentes para cada supuesto, cuya comprobación de este modo es mucho más exacta.

En algunos de los modelos clásicos más populares que se utilizan y enseñan a nivel de pregrado como son métodos de suavizamiento exponencial, regresión lineal y promedios móviles simples y ponderados, no capturan adecuadamente toda la influencia que pueden tener los datos históricos, como podrían hacerlo los ARIMA. Sin embargo, los modelos de regresión en series temporales requieren la comprobación de la idoneidad del modelo, que en muchos casos no se cumple, así como puede ocurrir en los modelos ARIMA (Cohen et al, Nov-Dic 1986).

Autores como Synder (2002) proponen correcciones a sistemas por computador para la gestión de inventarios de lenta y rápida rotación donde presenta adaptaciones y modificaciones del método de suavización exponencial para mejorar su aplicación donde generalmente se han hallado dificultades.

Los métodos de series de tiempo realizan estimaciones que tienen la capacidad de capturar la aleatoriedad de la demanda a partir de su comportamiento histórico ya que en estos “se identifican los patrones fundamentales de la demanda que se combinan para producir el patrón histórico observado en la variable dependiente, después de lo cual se elabora un modelo capaz de producir dicho patrón” (Krajewski, Ritzman, & Malhotra, 2008).

Pasando al campo de la Estadística Bayesiana, éste es uno de los métodos subjetivos más usados a nivel general y el cual ha venido tomando fuerza en los últimos años (Correa, 2005), sus inferencias son más fáciles de entender que la tradicional basada en muestreo repetitivo. (Albert, 1997) “Conocer el papel de la subjetividad en la interpretación de información, podría abrir un camino para razonamientos estadísticos flexibles más precisos.” (Berry, 1988)

Correa (2005) hace una diferenciación entre la estadística clásica y la bayesiana, la cual se muestra en la siguiente tabla:

Tabla 1. Diferencias entre teoría clásica y bayesiana

Diferencias entre la teoría clásica y la teoría bayesiana		
<i>Características</i>	<i>Teoría clásica</i>	<i>Teoría bayesiana</i>
Parámetros de interés	Constantes desconocida	Variables aleatorias
Distribución a priori	No existe	Existe y es explícita
Modelo muestral	Se asume	Se asume
Distribución posteriori	No existe	Existe y se deriva
Razonamiento	Inductivo	Deductivo

(Correa, 2005)

Según la tabla 1, es claro que los parámetros en la teoría bayesiana son variables aleatorias, mientras en la teoría clásica son constantes, lo que le permite a éstos adaptarse más a una realidad que puede ser muy variable.

1.2. Pronóstico de demanda

La aleatoriedad de la demanda es frecuente debido a los cambios que se dan en los mercados por cuestiones como crisis económicas, apertura de relaciones comerciales entre otros, es por esto que es importante tener en cuenta modelos estadísticos que logren caracterizarla para hacer un buen pronóstico cada vez más acercado a la realidad futura. En la industria nacional son muy poco frecuentes los métodos para el tratamiento de aleatoriedad de la demanda en los sistemas de diseño y control de inventarios, por su desconocimiento y falta de herramientas propias, necesarias en la optimización de inventarios, mejora de servicio, satisfacción de clientes, aspectos fundamentales en la dinámica productiva del país (Vidal, marzo 2008).

Es por esto que desarrollar prácticas óptimas de distribución en la cadena de abastecimiento en la industria es de significativo valor, bien sea como parte de un proceso productivo o como la entrega del producto final. Un aporte al mejoramiento de la cadena logística consiste en encontrar modelos eficientes para el pronóstico de la demanda de los productos y que a su vez, permitan minimizar los costos del manejo de los inventarios. Para ello es importante reconocer la aleatoriedad de la fluctuante demanda en algunas empresas, tema de numerosos estudios y propuestas metodológicas.

Dependiendo de cuánto se conozca el comportamiento de la demanda, la planeación eficiente va a ser exitosa pues afectará las utilidades de las empresas y por medio de ésta la rentabilidad final, la cual puede ser mejorada a partir del pronóstico mismo, el cual debe contar con un punto de partida que tenga en cuenta la disponibilidad de información lo más cercano posible a la realidad estudiada, desde datos históricos hasta la misma experiencia teórica y práctica, todo esto en miras a predecir lo más cerca posible el futuro (Geweke & Whiteman, 2004).

Los datos históricos de ventas son muy apropiados para realizar pronósticos. Sin embargo puede que no exista esta información para los casos en que hay nuevos inventos o productos como películas, música o libros; ni que tampoco haya una certeza de cuándo habrá información histórica disponible, lo que se necesita en esas situaciones es un patrón que se repita en varias situaciones.

Para productos nuevos es difícil estimar valores de ventas a largo plazo con apenas unas semanas iniciales al lanzamiento del producto, pues se presentan algunos problemas como la intensa promoción publicitaria que hace variar abruptamente los niveles de ventas, haciendo también difícil de encontrar patrones repetitivos de ventas (Fader, Hardie, & Huang, 2004).

Investigaciones en metodologías que den luz a distintas opciones donde no sea necesaria esa amplia cantidad de datos históricos para predecir el futuro en la industria manufacturera, es algo recurrente y necesario para encontrar dicho patrón repetitivo (Valencia, Correa, Díaz, & Ramírez, 2013).

La estadística bayesiana, toma en cuenta información a priori y posee un proceso de actualización de información a medida que se va contando con ésta, es un análisis indicado para este tipo de pronósticos que no cuenta con mayor información histórica pasada.

Esta posibilidad para generar buenos pronósticos sin información de ventas, dependerá de las características de fondo del marco de trabajo que se use para interpretar información histórica pertinente (Lee, Boatwright, & Kamakura, 2003).

En trabajos como el de Lee, Boatwright & Kamakura (2003) se plantea un modelo bayesiano para el pronóstico de ventas semanalmente, previas al lanzamiento de álbumes musicales, los cuales se actualizan a medida que se obtienen los datos reales de ventas en el tiempo.

Algunos investigadores están de acuerdo en que el juicio subjetivo, el cual es propio del análisis bayesiano, juega un papel importante pues permite incluir análisis que no son tan posibles estadísticamente como por ejemplo proyecciones de campañas publicitarias o conflictos internacionales (Fildes & Goodwin, 2007) .

Neelamegham & Chintagunta (1999) han usado una estructura jerárquica bayesiana para pronosticar, llevando a cabo un análisis que predice ventas durante la primera semana de nuevos lanzamientos de películas estadounidenses en mercados internacionales. La metodología permite que los pronósticos se puedan realizar con información proveniente de bases de datos de previas producciones. Entre más información se hace disponible, la metodología utilizada permite combinar la información histórica de mercados internacionales con el natal. Se muestra en este trabajo el mejor desempeño del modelo predictivo bayesiano para este tipo de situaciones.

Urban, Hauser & Roberts (1990) utilizan un método discreto de tiempo análogo de un proceso de Markov de tiempo continuo llamado *macro-flow (discrete time analog of a continuous time Markov process)*. Se describe un modelo para uso de la industria automovilística. Este modelo pronostica el ciclo de vida de un nuevo modelo de carro antes de su introducción.

En trabajos como el de Flora Lu (2005) donde se adoptan técnicas bayesianas para mejorar la calidad estimativa y predictiva del modelo Ohlson, comparándolo con los modelos clásicos, se muestra que los resultados obtenidos con modelos

clásicos contienen errores de predicción en los precios de inventarios y tienen menos poder explicativo que los modelos bayesianos.

2. MARCO TEÓRICO

2.1. Variable aleatoria

Una variable aleatoria es una función que asocia números reales con elementos de un espacio muestral dado.

Las variables aleatorias pueden ser discretas o continuas. Las discretas son en las cuales se pueden contar sus posibles valores, mientras que las continuas pueden tomar valores en una escala de este tipo, es decir, entre dos valores cualesquiera, existen infinitos números (Walpole, 1999).

Una distribución de probabilidad se refiere al conjunto de datos de la variable aleatoria y la manera como en estos están distribuidas las probabilidades.

Una variable aleatoria genera números aleatorios, los cuales pueden ser generados simplemente mediante el retiro de cartas de una baraja, el lanzamiento de un dado o en general por medio de cualquier actividad que implique aleatoriedad. Esta manera de generar números aleatorios es útil cuando se quiera generar pocos números aleatorios y es una buena aproximación a este concepto.

Para generar números aleatorios en el intervalo (0, 1), estos tienen que cumplir algunas propiedades estadísticas como distribuirse uniformemente en este intervalo; no deben estar correlacionados entre sí y debe permitir que se genere un número amplio de ellos.

Lo anterior se hace mediante una relación matemática la cual permite generarlos de manera sucesiva.

Los generadores congruenciales lineales utilizan relaciones multiplicativas y aditivas entre constantes del generador.

Relación aditiva:

$$X_{n+1} = (aX_n + b) \text{ mod } C \quad (1)$$

$$X_{n+1} = (aX_n + C) \text{ mod } m \quad (2)$$

Relación multiplicativa:

$$X_{n+1} = aX_n \text{ mod } m \quad (3)$$

Donde a , b , c y n son números enteros positivos y mod es el operador módulo, el cual da como resultado el residuo de una división.

Existen distintos criterios de selección de los cuales depende la cantidad de números aleatorios que se generan, cuya cantidad se refleja en el periodo del generador, el cual estará dado por la distancia entre la posición de los dos primeros números en repetirse. Es de aclarar que un generador es de periodo completo cuando éste es igual al módulo (Coss Bú, 1992).

Para simular este tipo de variables es necesario contar con un generador de números uniformes o pseudoaleatorios, y una función que, a través de un método específico, transforme estos números en valores de la distribución de probabilidad deseada, sea discreta o continua (López, 2012).

Para generar variables aleatorias existen distintos métodos entre los cuales se encuentran el método de la transformada inversa, el método del rechazo, el método de composición y los procedimientos especiales.

El método de la transformada inversa se basa en la distribución acumulada $F(x)$ de la distribución de la que se desean generar números, se da la siguiente relación

$$F(x) = R \quad (4)$$

La que al obtener la inversa, se tiene:

$$X = F^{-1}(R) \quad (5)$$

La eficacia del proceso dependerá de la posibilidad matemática de transformar la ecuación (Coss Bú, 1992).

Las variables aleatorias son una herramienta de la cual se puede echar mano a la hora de realizar pronósticos cuando se tienen pocos datos o no hay existencia de ellos. La simulación de variable aleatoria ha sido usada en algunos modelos probabilísticos de inventarios propuestos en Taha (2004), donde se asume la demanda como dicha variable aleatoria bajo diferentes tipos de distribución de probabilidad.

Estos modelos pueden ser muy útiles a la hora de pronosticar ventas y optimizar inventarios pero están basados en muchas suposiciones que podrían hacer que no se logre un pronóstico aceptable.

2.2. Software generadores de variables aleatorias

Existe software que facilita la tarea de generar variables aleatorias a través de los métodos expuestos anteriormente. El programa R de: “*The R foundation for statistical computing*”, es un software libre que puede verse como una implementación del lenguaje de programación S. Dicho software simula entre un rango de alrededor 19 distribuciones de probabilidad haciendo uso de funciones programadas, agregando los argumentos respectivos (Venables, Smith, the R Core Team, 2012).

Statgraphics Centurion es otro software estadístico que contiene alrededor de 170 procedimientos estadísticos y funciona bajo licencia pagada. Este software permite generar variables aleatorias de ciertas distribuciones de probabilidad (StatPoint Technologies, Inc., 2009).

Microsoft Excel también es una herramienta con la que se pueden generar números aleatorios a través del uso de sus funciones. Crystal Ball y @ Risk son paquetes complementarios de Excel que sirven para simular variables aleatorias (Eppen, 2000).

También se ha desarrollado software educativo para simular variables aleatorias (López, 2005), el cual cuenta con un generador de números uniformes pseudoaleatorios y una respectiva función que transforma dichos valores en distribuciones de probabilidad. En este software se hace uso del Método de los cuadrados Centrales de Von Neumann, Método Fibonacci, Método aditivo de congruencias, Método multiplicativo de congruencias y Método mixto de congruencias.

2.3. Distribución Poisson

Esta es una función de distribución para variables aleatorias discretas, la cual parte de una frecuencia de ocurrencia media y se dedica al estudio de números de eventos que se logran durante cierto periodo de tiempo y que ocurren de manera independiente con una velocidad constante en tiempo o espacio.

La distribución Poisson puede caracterizar procesos como llamadas a un *call center* por minuto, número de autos que pasan por una vía en un tiempo determinado, el número de bacterias en un cultivo, el número de defectos similares para un material entre otros procesos aleatorios.

Definición:

Sea X una variable aleatoria que representa el número de eventos aleatorios independientes que ocurren a una rapidez constante sobre el tiempo o el espacio. Se dice que X posee una distribución Poisson con función de probabilidad:

$$p(x, \lambda) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} & x = 0, 1, 2, \dots; \lambda > 0 \\ 0 & \text{Para cualquier otro valor} \end{cases} \quad (6)$$

El parámetro de esta función es λ el número promedio de ocurrencias del evento aleatorio por unidad de tiempo (Canavos, 1988).

El gráfico 1 muestra el comportamiento de la distribución para distintos valores de λ .

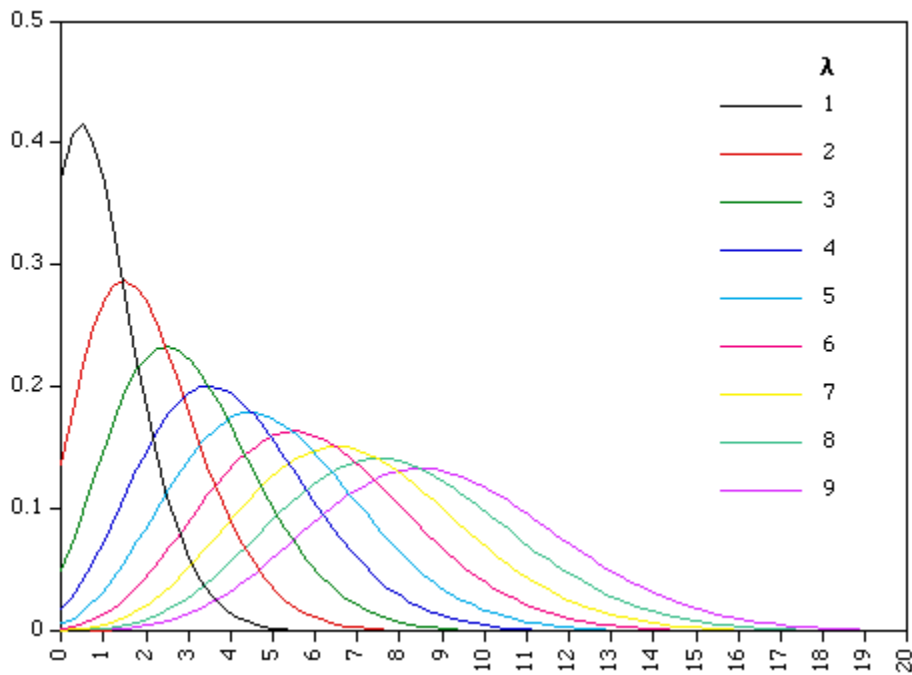


Gráfico 1. Distribución Poisson para valores de $\lambda=1-9$

(Bourke, 2001)

2.4. Distribución Gamma

Distribución de probabilidad continua con parámetro de forma α y de escala β . Esta distribución permite modelar una gran variedad de fenómenos de magnitud positiva. Piénsese un fenómeno en el que se aplica cierta magnitud de fuerza por ciclos que ocurren de manera independiente y a frecuencia promedio, el tiempo que ocurre antes de romperse el material se modela como la distribución Gamma (Canavos, 1988).

Definición:

Si la variable aleatoria X tiene una distribución gamma, tienen una función de densidad de probabilidad dada por:

$$f(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} ; x > 0, \quad \alpha, \beta > 0 \quad (7)$$

En donde $\Gamma(\alpha)$ es la función gamma definida como:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx ; \alpha > 0 \quad (8)$$

El gráfico 2 muestra el comportamiento para esta distribución para un valor constante para β y variación en el parámetro α .

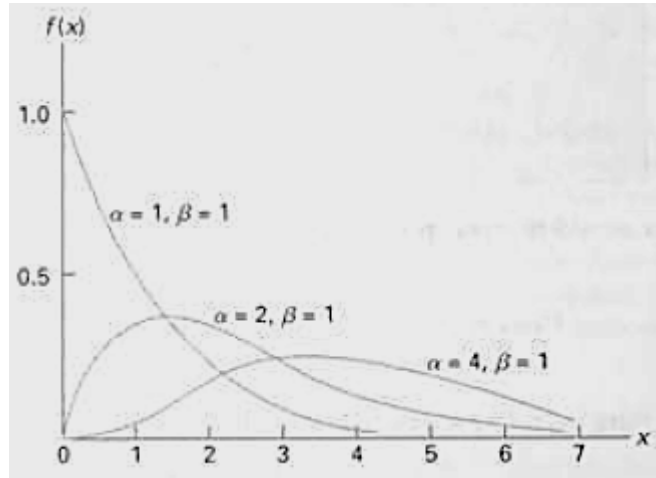


Gráfico 2. Distribución Gamma para algunos valores α y β

(Walpole, 1999)

2.5. Prueba de bondad de ajuste ji-cuadrado

Las pruebas de bondad de ajuste tienen como objetivo determinar si en una muestra dada se puede aceptar la hipótesis de que proviene de una población con una distribución de probabilidad especificada en una hipótesis nula.

H_0 : Los datos vienen de una distribución F_0

Vs

H_1 : Los datos NO vienen de una distribución F_0

Donde F_0 es la distribución de probabilidad hipotética para los datos que se quiere probar. La hipótesis nula se acepta si el valor p es mayor que la significancia asignada α , o si el estadístico de prueba no cae en una región de rechazo.

En el caso de la prueba ji-cuadrado, el estadístico de prueba es el siguiente:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (9)$$

Donde,

O_i es la frecuencia observada i

E_i es la frecuencia esperada (Alea, 2001).

Si dicho estadístico es menor que el valor crítico $\chi_{\alpha, k-1}^2$ se acepta H_0 , de lo contrario se rechaza.

2.6. Prueba de Box Pierce

Prueba de incorrelación de una serie de grado mayor a uno. Este estadístico es muy aplicado en las pruebas de autocorrelación de las series o incluso, residuales en estimaciones de modelos como Box-Jenkins, regresión en series temporales, o simulaciones, el cual determina si las primeras K autocorrelaciones de la serie son cero.

El contraste de hipótesis es:

$H_0: \rho(t, t+k)=0$: hay incorrelación

Vs

$H_1: \rho(t, t+k) \neq 0$: no hay incorrelación

Tiene como estadístico de prueba:

$$Qk = n' \sum_{l=1}^K r_l^2(\hat{a}) \quad (10)$$

Donde $n' = n - d$ siendo n el número de observaciones en la serie y d , el grado de diferenciación no estacional usado para transformar los valores originales de la serie temporal en valores estacionarios. $r_j(\hat{a})$ es la autocorrelación muestral de los residuos separada por un desfaseamiento de k unidades de tiempo (González & Lobato, 2007).

2.7. Estadística bayesiana

La Teoría bayesiana aparece cuando se postula el Teorema de Bayes, el cual toma como base el cálculo de probabilidad de un evento basado en un conocimiento a priori. (Bayes, 1763).

Este acercamiento es a la inversa del acercamiento frecuentista que examina las situaciones sin tener en cuenta los conocimientos previos. En general, la aplicación del teorema permite establecer que la probabilidad de un evento es el resultado condicional de una probabilidad conocida (Marien, 2009). La fórmula del teorema, si $\{A_1, A_2, \dots, A_n\}$ es un sistema completo de sucesos, entonces para cualquier suceso B se tiene lo siguiente:

Se tiene que la probabilidad condicionada es

$$P(A_k/B) = \frac{P(A_k \cap B)}{P(B)} \quad (11)$$

Y que la probabilidad total es

$$P(B) = \sum P(B/A_i) P(A_i) \quad (12)$$

Remplazando (12) en (11) se obtiene:

$$P(A_k/B) = \frac{P(A_k) P(B/A_k)}{\sum_{i=1}^n P(B/A_i) P(A_i)} \quad (13)$$

Donde a la suma de $P(B/A_k)$ se le llama función de verosimilitud (Serrano, 2003)

Para trabajar con variables aleatorias el teorema de Bayes se aplica de la misma manera, pero ya no se utilizan simplemente probabilidades a priori y a posteriori sino distribuciones, lo que hace que se enfrente a cálculos de funciones como productos, sumas o integrales.

Se tiene que una variable aleatoria es una función del espacio muestral de un experimento aleatorio al conjunto de los números reales, de tal modo que a cada resultado del espacio muestral le asigna un número real. La podemos representar como sigue:

$$X : E \rightarrow \text{Reales}$$

$$A_i \rightarrow x_i$$

Donde los x_i son los datos de la variable aleatoria X

Teorema de Bayes para variables aleatorias:

Teorema de la probabilidad total si $X = x_1, x_2, \dots, x_n$ y θ son variables aleatorias donde respectivamente, X representa los datos como escalar, vector o matriz, y θ es un parámetro desconocido de la misma forma.

$$f_x(x) = f(\text{datos}) = \int_{-\infty}^{\infty} f(\text{datos}/\theta)f(\theta) d\theta \quad (14)$$

El teorema de Bayes sería entonces:

$$f(\theta/\text{datos}) = \frac{f(\text{datos}/\theta)f(\theta)}{f_x(x)} \quad (15)$$

Reemplazando la ecuación (14) en (15) se tiene (Serrano, 2003):

$$f(\theta/\text{datos}) = \frac{f(\text{datos}/\theta)f(\theta)}{\int_{-\infty}^{\infty} f(\text{datos}/\theta)f(\theta)d\theta} \quad (16)$$

El denominador será siempre un número constante, que no dependerá, al final, del parámetro θ , por tanto para simplificar se puede utilizar una función proporcional y finalmente el teorema queda como sigue:

$$f(\theta/\text{datos}) \propto f(\text{datos}/\theta)f(\theta) \quad (17)$$

La distribución a priori puede ser subjetivamente determinada por el usuario cuando no se tiene información previa a la “experiencia”, o puede darse más

objetivamente, cuando se tiene información, analizándose estadísticamente qué tipo de distribución posee (Raftery, 1999).

El uso de información conocida a priori es lo que se piensa es la parte destacada de la inferencia bayesiana (Geweke & Whiteman, 2004)

En grandes muestras la distribución a priori tiene muy poca influencia, teniendo una contribución a la media y varianza posteriores del orden $(1/n)$ -th del total, donde n es el tamaño de la muestra. (Raftery, 1999)

Si $f(datos/\theta)$ en (17) se deja fijo X que en este caso representa los datos y se evalúa para cada valor de la variable θ , se tendrá una función en θ , la cual tendrá el nombre de función de verosimilitud de θ . Se denota $l(\theta/datos)$.

Así finalmente se tiene

$$f(\theta/datos) \propto l(\theta/datos)f(\theta) \quad (18)$$

Esto es:

$$\textit{Distribución a posteriori} \propto \textit{función de verosimilitud} \times \textit{distribución a priori}$$

La proporcionalidad simplemente dice que si el lado izquierdo es proporcional al lado derecho, ambas difieren, solamente, de una constante multiplicativa. (Lynch, 2006)

El procedimiento para obtener la distribución a posteriori será entonces el producto de la función de verosimilitud calculada para los datos que se tengan y la distribución a priori de la variable que se determinó; combinándose así la información previa a la “experiencia” con la información objetiva de los datos.

A medida que se obtenga más información, la distribución a posteriori se puede ir actualizando, usándola como distribución a priori, así obteniéndose una nueva distribución a posteriori más actualizada. Esta es una gran ventaja que tiene el análisis bayesiano, ya que el análisis clásico frecuentista implica que a medida que se conozcan más datos, hay que reanalizar todo de nuevo (Bolstad, 2007).

La mezcla del sistema de actualizar y del pronóstico de probabilidad en tiempo real fue llamado precuencial, (combinación entre pronóstico de probabilidad y predicción secuencial) por Dawid (1984). Se distinguió esta forma de pronosticar de otros sistemas, los cuales no actualizan por completo.

También en el análisis bayesiano la información de expertos puede ser incorporada para mejorar los pronósticos (Geweke & Whiteman, 2004).

En el análisis bayesiano los parámetros son considerados cantidades aleatorias, mientras que los datos, habiendo sido ya observados, son considerados como cantidades fijas (Lynch, 2006). De esta manera se pueden dar muchos casos en el que el parámetro de cierta distribución se distribuye como otra o como la misma distribución.

Se dice que cuando la distribución a posteriori es de la misma familia que la distribución a priori, es una familia de distribuciones conjugadas.

A continuación se muestra una tabla con distintas combinaciones de distribuciones que dan lugar a cierta distribución a posteriori.

Tabla 2. Algunas distribuciones conjugadas

Priori	Distribución de datos	Posteriori
Gamma	Poisson	Gamma
Normal	Normal	Normal
beta	Bernoulli	Beta
Gamma	Inversa exponencial	Gamma

(Flores, 2012)

La inferencia bayesiana hasta hace poco tiempo tenía obstáculos técnicos que de alguna manera retardaron su desarrollo, pues generalmente este análisis conlleva a problemas de integración y complejidad en la formulación del problema; sin embargo amplios avances computacionales han permitido su más práctica solución, ampliando así el alcance de este tipo de métodos (Geweke & Whiteman, 2004).

2.8. Monte Carlo por Cadenas de Markov (MCMC)

La metodología de simulación Montecarlo parte de la elección aleatoria de muestras para aproximar un problema combinatorio complejo a uno más simple de resolver (Andrieu, De Freitas, Doucet, & Jordan, 2003). Dicha simulación genera valores independientes tomando como base una distribución de probabilidad deseada. Por otro lado las cadenas de Markov, postulan que los estados anteriores son irrelevantes para la predicción de los estados siguientes, desde que

el estado actual sea conocido, basándose en una estructura de dependencia entre los valores simulados consecutivamente.

Al unir ambas técnicas, se crea la metodología de simulación *Monte Carlo por Cadenas de Markov* que consiste en generar un muestreo a partir de distribuciones de probabilidad basadas en la construcción de cadenas de Markov donde cada valor simulado, tiene dependencia con el dato anterior, llegando a una convergencia a la distribución deseada (Gill, 2002). Después de una larga cantidad de corridas, estos resultados se utilizan como una muestra estacionaria de la distribución deseada. Entre más corridas, la calidad de la muestra es mejor.

Los métodos MCMC son una estrategia para generar valores de la variable aleatoria buscada $x(i)$ mientras explora el espacio X usando cadenas de Markov. Este mecanismo se construye para que la cadena pase más tiempo en las regiones más importantes (Andrieu, De Freitas, Doucet, & Jordan, 2003).

En algunas ocasiones, la generación de variables aleatorias bajo una distribución de probabilidad compleja, se hace muy difícil, para ello se recurre a algoritmos de apoyo basados en MCMC, como el de *Gibbs* o el de *Metropolis Hastings*.

Algunos de estos algoritmos se basan sólo en una caminata aleatoria en la distribución, dándole más importancia a los valores con más probabilidad. Para hacerlo se toma un punto de partida y luego se toma aleatoriamente un punto cercano. Otros como el de *Hastings*, eligen el mejor punto al comparar dos criterios (usualmente el punto de más probabilidad).

Estos algoritmos han jugado un papel importante en estadística, econometría, física y ciencia computacional. Tan importante es, que existen diferentes problemas de altas dimensiones en los que simulación por MCMC es el único método conocido que acerca a una solución en un tiempo razonable (Andrieu, De Freitas, Doucet, & Jordan, 2003).

MCMC no se usa solamente para inferencia bayesiana, sino también en casos en los que, por ejemplo, la verosimilitud no puede ser calculada explícitamente debido a información faltante o dependencia compleja (Geyer, 2011).

Estos métodos MCMC son considerados como el desarrollo más importante en estadística computarizada. Estos métodos han permitido ajustar prácticamente cualquier modelo de probabilidad que antes no era posible por las limitaciones que existían en software para resolver problemas de alto contenido teórico (Martin, Quinn, & Park, 2011)

Muestreador de Gibbs

El muestreador Gibbs es un algoritmo iterativo Monte Carlo diseñado para extraer la distribución a posteriori para la inferencia bayesiana, a partir de la distribución condicional encontrada para el parámetro (Flora Lu, 2005) actualizando el valor en cada iteración i .

Este algoritmo es flexible y confiable para generar cadenas de valores. Este es un *kernel* de transición (Gill, 2002), creado por una serie de distribuciones condicionales en un esquema Markoviano. Si la distribución límite de interés es $f(\beta)$, donde β es un vector k dimensional de coeficientes a estimar, entonces el objetivo es producir una cadena de Markov que genere ciclos alrededor de estas condicionales y que converja a dicha distribución.

2.9. Regresión gaussiana bayesiana por Cadenas de Markov.

En general, las metodologías de estimación de regresión bayesiana se basan en distribuciones asignadas a priori para los parámetros así como distribuciones para la verosimilitud de los datos, con el fin de obtener distribuciones a posteriori condicionales, de las cuales se hace el muestreo con la cadena MCMC, para al final encontrar muestras estacionarias de dichos coeficientes.

Para realizar el modelo específico para este trabajo, el programa R cuenta con un paquete: *MCMCPack*, que con la función *MCMCregress* estima los parámetros, asumiendo el siguiente proceso:

- Aplica el muestreador de Gibbs, utilizando la distribución a priori Normal multivariada para los betas (coeficientes del modelo de regresión).
- Distribución a priori Gamma inversa para la varianza del error condicional.

El modelo de regresión gaussiana bayesiana toma la siguiente forma:

$$y_i = x_i' \beta + \varepsilon_i \quad (19)$$

Donde el error se asume Normal:

$$\varepsilon_i \sim \mathcal{N}(0, \sigma^2) \quad (20)$$

Se asumen a priori semi-conjugada estándar:

$$\beta \sim \mathcal{N}(b_0, B_0^{-1}) \quad (21)$$

$$\sigma^{-2} \sim \mathcal{Gamma}(c_0/2, d_0/2) \quad (22)$$

Donde β y σ^{-2} se asumen como a priori independiente.

2.10. Regresión bayesiana Poisson por cadenas de Markov

El modelo de regresión Poisson bayesiano en este caso se estima con la función *MCMCpoisson*, que genera una muestra de los coeficientes del modelo usando la distribución a priori normal multivariada para los parámetros, distribución Poisson para los datos, y usa una caminata aleatoria con los criterios del algoritmo Metropolis.

El modelo toma la siguiente forma:

$$y_i \sim \mathcal{Poisson}(\mu_i) \quad (23)$$

Donde,

$$\mu_i = \exp(x_i' \beta) \quad (24)$$

Se asume una a priori Normal multivariada en β (Martin, Quinn, & Park, 2013):

$$\beta \sim \mathcal{N}(b_0, B_0^{-1}) \quad (25)$$

En este tipo de modelos de regresión bayesiana, distintos tipos de variables pueden ser incorporadas según la necesidad de ajuste a la demanda histórica. Variables de tiempo lineal y tipo cuadrático son una opción para mejorar el modelo y capturar tendencia. Las variables trigonométricas se utilizan cuando existe variación en el tipo de estacionalidad, sea constante o creciente para así, captar su variabilidad. De esta forma se tiene:

$$\text{Sen}\left(\frac{c\pi t}{L}\right) \quad (26)$$

y

$$\cos\left(\frac{c\pi t}{L}\right) \quad (27)$$

Donde, c es una constante par, t el periodo a pronosticar y L es el número entero positivo que representa el patrón de estacionalidad.

2.11. Método de optimización Tabú

La búsqueda metaheurística Tabú fue propuesta por Glover (1990) como una técnica para sobrellevar el atrapamiento en una solución óptima local en algoritmos heurísticos de búsqueda de base local.

Este algoritmo busca una optimización del valor de una función objetivo f_x con k variables de decisión, de manera que se genere una lista con los valores de los puntos en el espacio k+1 dimensional (k variables y el valor de la función objetivo f_x) llamada lista Tabú. Dicha lista representa la historia del recorrido de la búsqueda, para evitar repetir los puntos ya evaluados, hasta encontrar el mejor valor entre los puntos explorados. Chelouah y Siarry (2005) proponen un algoritmo de exploración-explotación, donde hay una explotación de puntos dentro de un vecindario y una exploración para buscar la mejor solución posible en un conjunto de soluciones, encontrando aproximaciones al óptimo global (Valencia et al, 2013).

2.12. Error porcentual medio absoluto (MAPE)

En aplicaciones prácticas es una forma muy utilizada y común para medir la precisión de un pronóstico a pesar de su relativa simplicidad.

$$\text{Error porcentual} = \frac{A_t - F_t}{A_t} \times 100 \quad (28)$$

$$(APE)_t = \left| \frac{A_t - F_t}{A_t} \right| \times 100 \quad (29)$$

$$MAPE = \sum_{t=1}^N \frac{APE_t}{N} \quad (30)$$

Donde, A_t = Es el valor real en el tiempo t

F_t = Pronóstico en el tiempo t

N = Número de Periodos

$(APE)_t$ = Error porcentual absoluto

$MAPE$ = Error porcentual absoluto medio (McKenzie, 2006)

3. METODOLOGIA

La metodología que se presenta en esta sección busca describir el procedimiento primero de forma general todos los pasos seguidos en el trabajo, luego, en detalle el proceso de cada método estadístico, explicando las variables, distribuciones y premisas que deben ser tenidas en cuenta para llegar a las estimaciones de pronósticos finales respectivas. Finalmente, el desempeño de estos métodos se compara usando el criterio de pronóstico (MAPE), aspecto que se realiza en la sección de resultados, con su respectivo análisis y conclusiones, posterior a la estimación de los pronósticos.

Todos los métodos usados fueron aplicados a un conjunto de 24 datos, que corresponden a las ventas de seis productos de una empresa de confección de 2010 a 2011 (González & Franco, 2013).

Proceso general

Parte I.

Con los primeros 12 datos (2010), se procede así:

1. Hacer un análisis de los datos por medio de la gráfica tiempo vs ventas 2010 y la gráfica de la función de autocorrelación muestral (SAC) (Bowerman, O'Connell, & Koehle, 2007)
2. Aplicar una prueba de bondad de ajuste con el fin de generar la variable aleatoria utilizando una distribución de probabilidad para los datos (explicada en subsección 3.1.).
3. Estimar un modelo de regresión bayesiana Poisson y Gaussiana por Cadenas de Markov (explicados en subsección 3.2.)
4. Con los métodos bayesiana descrita en la subsección 3.3., se encuentra la ecuación de la prima bayesiana genérica para iniciar los pronósticos. El primer paso es estimar las ventas de enero de 2011.

Parte II.

1. Simular una variable aleatoria Poisson, con el parámetro encontrado ($\hat{\lambda} = \bar{x}$), usando un algoritmo para estimar la media de 12 valores de ventas mensuales del año 2011, haciendo 1000 simulaciones y calculando el error relativo por mes y el MAPE anual. Al final se obtienen 12 medias mensuales que representan el pronóstico de cada mes de 2011 y un MAPE final.

2. Partir de la ecuación obtenida para el valor esperado bayesiano, y de los datos actualizados en cada mes de 2011, se emplea la técnica Tabú con el fin de encontrar los parámetros respectivos que minimizan el MAPE.
3. Generar pronósticos con las ecuaciones de regresión bayesiana obtenidos a partir de las simulaciones que se realizan con el paquete de MCMC.
4. Para cada uno de los valores pronosticados mensuales por método estadístico, se estima el APE_t y el MAPE para del año 2011, como criterio final de comparación.

Todos los procesos de estimación para los seis productos siguen los mismos pasos (de las partes I y II). Las técnicas estadísticas serán mostradas en detalle a continuación y al final se muestra el resumen de modelos e indicadores encontrados para todos los productos.

3.1. Prueba de Bondad de ajuste y variable aleatoria

Se realiza la prueba de bondad de ajuste Ji-cuadrado con el fin de determinar a cuál distribución de probabilidad se ajustan los datos. Debido a que los datos son discretos, se le hace la prueba a estos bajo las distribuciones teóricas Binomial, Geométrica, Uniforme y Poisson

El contraste de pruebas de hipótesis, en relación a la distribución Poisson, es:

H₀: Los datos se ajustan a una distribución Poisson con $\hat{\lambda} = \bar{X}$

Vs

H₁: Los datos NO se ajustan a una distribución Poisson con $\hat{\lambda} = \bar{X}$

Para las demás distribuciones, se modifica F₀ y el respectivo parámetro.

Acorde con la distribución de ajuste encontrada, se genera la variable aleatoria con el programa estadístico R y se utiliza el algoritmo mencionado en la segunda parte de la introducción de la metodología.

3.2. Regresión Gaussiana y Poisson bayesiana con MCMC.

Para llevar a cabo este análisis bayesiano, se parte de la propuesta de modelación basada en una exploración de diferentes covariables en cada modelo de los seis productos a analizar, acorde con el tipo de tendencia o estacionalidad que podría presentar el conjunto de datos. Las covariables exploradas son:

- Tiempo lineal : t
 - Tiempo cuadrático : t^2
 - Tiempo cuadrático : t^3
 - Variables trigonométricas: $Sen\left(\frac{c\pi t}{L}\right), Cos\left(\frac{c\pi t}{L}\right)$
 - Ficticias o de nivel, como el mes.
-
- **Modelo de regresión bayesiana Poisson con MCMC.**

Los modelos de regresión bayesiana Poisson se utilizan dada la naturaleza de los datos, y su respectiva prueba de bondad de ajuste, asumiendo los supuestos explicados acerca de normalidad a priori para los parámetros y distribución Poisson para los datos.

La ecuación del modelo de regresión Poisson para el producto A es:

$$\ln(\hat{Y}) = \beta_0 + \beta_1 t^2 + \beta_2 Sen\left(\frac{2\pi t}{12}\right) + \varepsilon_t \quad (31)$$

Donde t es el número que le corresponde al mes, partiendo desde los datos que se usaron para estimar el modelo, es decir, enero del 2010 será $t=1$, por lo que enero del 2011 será $t=13$ y así sucesivamente. ε_t es el error

Sin embargo, la distribución normal de los datos no es inapropiada, ya que se da esta aproximación para la Poisson cuando la media es alta (Montgomery & Runger, 2003), en este caso, las medias de cada producto están alrededor de 100. Por ello se puede explorar el modelo de regresión bayesiana gaussiana con MCMC mostrado a continuación.

- **Modelo de regresión bayesiana gaussiana con MCMC.**

La ecuación empleada es similar a la de un modelo de regresión común, por ejemplo, para el producto A la ecuación general del modelo de regresión gaussiano es:

$$\hat{Y} = \beta_0 + \beta_1 t^2 + \varepsilon_t \quad (32)$$

Los coeficientes finales, para ambos modelos de regresión, se pueden tomar de las estadísticas de tendencia central de los parámetros simulados. Para este trabajo, la estadística que servirá de coeficiente de cada modelo será la media de las simulaciones dicho parámetro.

Estos modelos se estiman utilizando la metodología Monte Carlo por Cadenas de Markov con las funciones MCMCRegress y MCMCPoisson dentro del paquete MCMCpack del software estadístico R.

3.3. Prima Bayesiana

El proceso analítico para generar la ecuación del valor esperado bayesiano (Prima bayesiana), se describe en esta sección, como insumo para presentar la aplicación a los datos que se mostrará en el capítulo de resultados. Aspecto que constituye el aporte principal del trabajo.

Los datos tienen una naturaleza discreta, con distribución probabilística ajustada según la Poisson (resultados presentados en el siguiente capítulo), aspecto que es el punto de partida para emplearla como la verosimilitud y la distribución gamma como la a priori para el parámetro λ , con parámetros α y β que se asumen con distribuciones a priori no informativas uniformes.

El parámetro λ , que representa en este caso una tasa de ventas mensuales, posee una naturaleza continua, y posee una variabilidad alta, por lo cual, la distribución Gamma puede representar adecuadamente dicho comportamiento.

Resumen del método usado para encontrar la Prima Bayesiana de acuerdo a la teoría respectiva.

1. Definir distribuciones de los datos (verosimilitud) y a priori del parámetro
2. Multiplicar la distribución a priori por la verosimilitud
3. Encontrar la forma de la distribución a posteriori
4. Estimar el valor esperado, usando la integral que integra el producto del valor esperado de la distribución de los datos por la distribución a posteriori hallada para el parámetro.

Dicha ecuación es el insumo para efectuar los pronósticos. A continuación se detallará esta metodología, que es aplicable, de manera general a distintos casos de datos con naturaleza Poisson.

- **Proceso de estimación**

Las premisas parten del conocimiento de las distribuciones de los datos y la del parámetro. La función de verosimilitud de los datos es la Poisson (λ).

$$L(\lambda|x_1, x_2, \dots, x_n) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \quad (33)$$

Para λ se tiene la distribución Gamma (α, β) como la distribución a priori, se recuerda a continuación:

$$f(\lambda) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} \quad \lambda > 0, \alpha, \beta > 0 \quad (34)$$

Se calcula la a distribución a posteriori usando la ecuación (15), reemplazando (33) y (34)

$$f(\theta = \lambda|Datos) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \Bigg/ f_x(x) \quad (35)$$

Agrupando y organizando la ecuación anterior se obtiene:

$$f(\theta = \lambda|Datos) = \frac{1}{\prod_{i=1}^n x_i! \Gamma(\alpha)\beta^\alpha} \lambda^{\alpha+\sum_{i=1}^n x_i-1} e^{-\lambda/\beta-n\lambda} \Bigg/ f_x(x) \quad (36)$$

Se reemplaza la ecuación (33) y (34) en (14) teniendo en cuenta que los límites de la integral son enteros positivos por la restricción en el parámetro λ y se agrupa y organiza, lo que al final le da la misma forma que la ecuación (36).

$$f_x(x) = \int_0^\infty \frac{1}{\prod_{i=1}^n x_i! \Gamma(\alpha)\beta^\alpha} \lambda^{\alpha+\sum_{i=1}^n x_i-1} e^{-\lambda\left(\frac{1}{\beta}+n\right)} d\lambda \quad (37)$$

Se asumen constantes los valores que no tienen que ver con el diferencial de la integral.

$$f_x(x) = \frac{1}{\prod_{i=1}^n x_i! \Gamma(\alpha)\beta^\alpha} \int_0^\infty \lambda^{\alpha+\sum_{i=1}^n x_i-1} e^{-\lambda\left(\frac{1+\beta n}{\beta}\right)} d\lambda \quad (38)$$

La integral anterior corresponde a la forma de una gamma, con la derivada del exponencial elevada al valor de alfa:

Se tiene la función de densidad de probabilidad (7) donde su integral es igual a 1, luego, al despejar la integral se obtiene:

$$f_x(x) = \frac{1}{\prod_{i=1}^n x_i! \Gamma(\alpha)\beta^\alpha} \Gamma\left(\alpha + \sum_{i=1}^n x_i\right) \left(\frac{\beta}{1 + \beta n}\right)^{\alpha+\sum_{i=1}^n x_i} \quad (39)$$

Reemplazando el valor anterior en (36) se tiene

$$f(\lambda|Datos) = \frac{\frac{1}{\prod_{i=1}^n x_i! \Gamma(\alpha)\beta^\alpha} \lambda^{\alpha+\sum_{i=1}^n x_i-1} e^{-\lambda/\beta-n\lambda}}{\frac{1}{\prod_{i=1}^n x_i! \Gamma(\alpha)\beta^\alpha} \Gamma(\alpha + \sum_{i=1}^n x_i) \left(\frac{\beta}{1 + \beta n}\right)^{\alpha+\sum_{i=1}^n x_i}} \quad (40)$$

Cancelando y agrupando términos se obtiene finalmente la distribución a priori.

$$f(\lambda|Datos) = \frac{\lambda^{\alpha+\sum_{i=1}^n x_i-1} e^{-\lambda\left(\frac{1+\beta n}{\beta}\right)}}{\Gamma(\alpha + \sum_{i=1}^n x_i) \left(\frac{\beta}{1 + \beta n}\right)^{\alpha+\sum_{i=1}^n x_i}} \quad (41)$$

Son muchas las técnicas bayesianas para pronosticar. Comprenden la inferencia predictiva, el valor esperado de la función predictiva bayesiana (Flórez, 2012), los modelos de regresión vía Monte Carlo por Cadenas de Markov (MCMC) así como los modelos dinámicos bayesianos.

El valor esperado asociado a la predicción, también llamada Prima Bayesiana por Flórez (2012) se utiliza en este trabajo para obtener los pronósticos.

Prima bayesiana:

$$E(X_{t+1}|x) = \int_0^\infty E(x_{t+1}|\lambda) f(\lambda|Datos) d\lambda \quad (42)$$

Donde $E(x_{t+1}|\lambda) = \lambda$ es el valor esperado

Continuando con el análisis y reemplazando el valor esperado y (41) en (42) se tiene después de organizar:

$$E(X_{t+1}|x) = \frac{1}{\Gamma(\alpha + \sum_{i=1}^n x_i) \left(\frac{\beta}{1 + \beta n}\right)^{\alpha + \sum_{i=1}^n x_i}} \int_0^{\infty} \lambda^{1 + \alpha + \sum_{i=1}^n x_i - 1} e^{-\lambda \left(\frac{1 + \beta n}{\beta}\right)} d\lambda \quad (43)$$

Luego de reemplazar términos dentro y fuera de la integral, esta última, nuevamente tiene la forma de la función de densidad Gamma, que vale 1, quedando la siguiente ecuación.

$$E(X_{t+1}|x) = \hat{Y}_t = \left(\alpha + \sum_{i=1}^n x_i\right) \left(\frac{\beta}{1 + \beta n}\right) \quad (44)$$

Valor que constituye la prima bayesiana; que sólo depende de los n datos tomados y de los parámetros α y β .

Un modelo que incorpora un cambio dinámico a la ecuación, incorporando términos autorregresivos a los parámetros, genera el siguiente resultado final, que será utilizado para este trabajo.

$$\hat{Y}_t = \left(\alpha_t + 0.1 \alpha_{t-1} + \sum_{i=1}^n x_i\right) \left(\frac{\beta_t + 0.1\beta_{t-1}}{1 + (\beta_t + 0.1\beta_{t-1})n_{t-1}}\right) \quad (45)$$

Donde α_t y β_t son los parámetros de la distribución Gamma para el tiempo t

α_{t-1} y β_{t-1} son los parámetros para el tiempo t-1

n_t es el número de mes correspondiente.

El valor constante 0.1 corresponde a un valor de correlación muy pequeño que se puede dar entre el valor t versus t-1.

De esta manera se intenta captar fluctuaciones en las ventas futuras, sin embargo los valores de los parámetros α y β , a utilizar en este modelo no serán los óptimos, viéndose al final esto reflejado en el resultado propuesto. En la sección de optimización tabú se encontrarán los valores óptimos para estos parámetros para intentar mejorar mucho más el pronóstico final.

3.4. Método de optimización Tabú

La heurística es diseñada para este trabajo, se exploran los valores de los parámetros de la ecuación (45), de tal manera que los 12 parámetros (uno por mes), sean los que encuentren el menor error medio (MAPE) a partir de los pronósticos generados con este valor esperado.

Función objetivo:
$$MAPE = \sum_{t=1}^N \frac{|A_t - F_t|}{A_t} \cdot \frac{1}{N}$$

Donde, A_t es el valor real y F_t el pronóstico, hallado con la prima bayesiana y la función objetivo es MAPE

Modelo

Minimizar $z=MAPE$

s.a. $\alpha, \beta > 0$

Notaciones:

X_{t+1} : valor a pronosticar en el tiempo $t+1$.

n : Cantidad de datos disponibles para pronosticar.

α : Parámetro de forma

β : Parámetro de escala

α y β son los parámetros de la distribución a priori Gamma del parámetro λ y constituyen las variables de decisión.

Al encontrar los mejores valores de dichos parámetros, deben ser utilizados para la ecuación de la prima bayesiana de forma que pueda efectuarse el pronóstico de manera eficiente. Esto permitirá el uso de la ecuación para planear inventarios, reemplazando los nuevos valores de las variables en cada tiempo t .

- **Algoritmo**

En este análisis se propone una variación en la exploración de puntos nuevos usando una combinación de una distribución uniforme y una normal, de manera que se encuentren parámetros de α y β que minimicen la media de error absoluto

de pronóstico de un producto terminado MAPE. La lista tabú se actualiza cada que se encuentra un punto que genera un valor de la función objetivo menor al mínimo de todos los demás que están en la lista Tabú.

A continuación se explica paso a paso el algoritmo utilizado para optimizar:

1. Crear una función de búsqueda de 12 valores de α y de β alrededor de un punto, de forma que encuentre el mínimo valor de la función objetivo MAPE, estocásticamente, con la función coseno y la distribución normal. Dicho MAPE se calcula con el valor pronosticado con la prima bayesiana actualizada a tiempo t , vs el valor real de la venta.
2. Crear una lista tabú con los mínimos valores de la función objetivo que explotó alrededor de un conjunto de puntos, posteriormente, genera nuevos puntos con la función coseno y la distribución normal, comparándolos con los existentes en la lista, si son iguales a algún conjunto de parámetros, no se incorpora el nuevo punto, de lo contrario, se incorpora si es menor del mínimo que hay en la lista.
3. Al final, el menor valor de la función objetivo de todos será el último de la lista, con los respectivos valores de alfas y betas (12) que lo generaron.

4. RESULTADOS

4.1. Exploración de los datos

Para el primer producto se mostrará el proceso exploratorio.

La siguiente gráfica muestra las ventas del año 2010 del producto A desde el mes de enero hasta diciembre de ese mismo año.

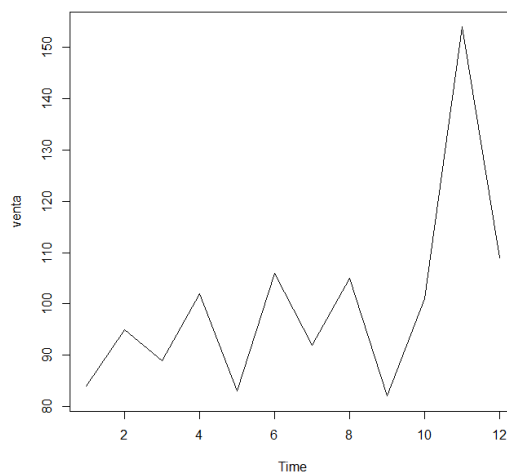


Gráfico 3. Tiempo vs ventas 2010 para producto A

No es clara una tendencia pues al final hay un cambio en la variabilidad muy fuerte que, al menos a simple vista no permitiría hacer una predicción de su futuro comportamiento.

La siguiente gráfica muestra la SAC para la misma serie de tiempo anterior

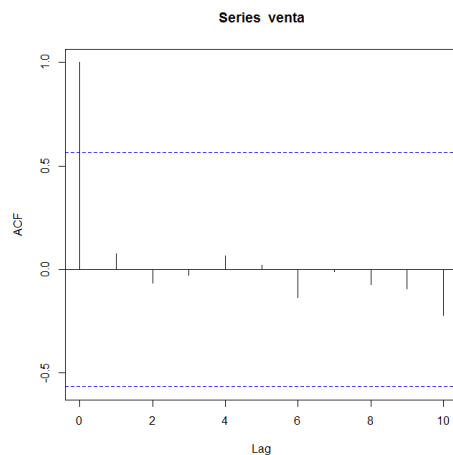


Gráfico 4. Función de autocorrelación muestral para 2010, producto A

No se aprecia alguna dependencia con el pasado de la serie pues todos los valores de autocorrelación están dentro de las bandas; sin embargo concluir con tan pocos datos no es recomendable.

Por lo anterior no existen argumentos teóricos para casarse con un modelo en especial.

4.2. Prueba Bondad de ajuste ji-cuadrado.

Al realizar la prueba sobre los datos del producto A, se obtienen los siguientes valores p para cada una de las distribuciones seleccionadas con el software StatGraphics. En la primera fila se observa el valor del estadístico de prueba, en la segunda los grados de libertad respectivos y en la última el valor P.

Tabla 3. Prueba de bondad de ajuste ji-cuadrado para producto A

	Binomial	Geometric	Poisson	Uniform
Chi-Square	7,31032	49,3153	3,54182	13,2282
D.f.	3	1	3	2
P-Value	0,0626361	2,17959E-12	0,315374	0,00134132

Al comparar los valores P se encuentra que en la prueba de bondad de ajuste para la distribución Poisson se cumple la hipótesis nula, con lo que se concluye que ésta es la distribución a la cual mejor se ajustan los datos del producto A.

A continuación se muestra la tabla con los resultados del resto de productos a los cuales se les aplica el mismo proceso que al producto A. La distribución Poisson siempre fue a la que se ajustó el conjunto de datos de cada producto.

Tabla 4. Prueba de bondad de ajuste ji-cuadrado para todos los productos

	Prueba de bondad de ajuste					
Producto	A	B	C	D	E	F
Distribución	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Valor p	31,54%	13,72%	20,05%	0,04%	16,83%	13,48%

En el producto D, no se cumplió la hipótesis nula al 5%, para ajustar a la distribución Poisson, sin embargo, es el Valor p más alto de varias consideradas, por ello, se selecciona la distribución Poisson por la naturaleza discreta de los datos y por variar en el tiempo.

4.3. Pronósticos con Variable aleatoria

En esta sección se realiza un pronóstico con variable aleatoria Poisson, siendo consecuentes con los resultados obtenidos en la prueba de bondad de ajuste.

Para el producto A se tiene una media, equivalente al valor del parámetro λ de 100.667. Con ayuda de un algoritmo programado en R se lleva a cabo una simulación con 1000 iteraciones, la cual calculará el MAPE para cada una de ellas, para proceder y calcular las estadísticas básicas como mínimo, máximo, mediana, entre otras. De estas estadísticas se toma la media, siendo ésta el MAPE final.

En la tabla 5 se muestran 12 de los valores pronosticados para 2011 y su APe_t correspondiente para una sola iteración.

Tabla 5. Pronóstico mensual de 2010 y APe_t para una iteración

Mes	Pronóstico	APe_t
Ene	123	18,27%
feb	114	14,29%
Mar	110	15,79%
Abr	105	40,34%
May	111	15,27%
Jun	96	25,00%
Jul	94	23,58%
Ago	100	14,53%
Sep	94	6,00%
Oct	75	46,81%
Nov	96	40,00%
Dic	101	41,95%

Nótese la variación del valor del APe_t , moviéndose entre 6% y 47% para este caso

A continuación se muestra la tabla resumen con el MAPE respectivo para cada producto, cuyos valores se obtienen utilizando el mismo proceso que en el producto anterior, luego de las 1000 iteraciones.

Tabla 6. MAPE para todos los productor con variable aleatoria simulada

Producto	Variable aleatoria					
	A	B	C	D	E	F
Distribución	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
MAPE	23,32%	28,86%	25,74%	20,81%	32,86%	24,66%

Se obtiene valores MAPE muy cercanos los unos a los otros, resaltándose el del producto D, que aparece como el mejor.

4.4. Modelo de Regresión bayesiana Gaussiana y Poisson con MCMC

Para el producto A se exploraron dos tipos de modelos, de los cuales el mejor resultado se obtuvo con la regresión bayesiana gaussiana, luego de 100.000 simulaciones, con un quemado de 20.000, usando la función MCMCregress de R.

- **Modelo de regresión bayesiana Poisson**

En la siguiente tabla se presentan los valores de la media de cada uno de los coeficientes y la respectiva desviación estándar para el modelo (31) del producto A, que se retoma a continuación:

$$\ln(\hat{Y}) = \beta_0 + \beta_1 t^2 + \beta_2 \text{Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t$$

Tabla 7. Media de los parámetros y desviación estándar para modelo de regresión Poisson con MCMC

Modelo de regresión bayesiana Poisson con MCMC		
Coeff	Mean	SD
(Intercept)	4.445600	0.0540955
t ²	0.002851	0.0007913
sint	0.059988	0.0547391

Los valores de la media para cada una de las covariables de la columna uno corresponden al valor de los parámetros β_i que completarán el modelo. De esta manera al reemplazarlos en (31) la ecuación sería:

$$\hat{y}_t = 4.445 + 0.00285t^2 + 0.05998 \operatorname{Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t \quad (46)$$

En la siguiente tabla se muestran los modelos respectivos de cada producto usando su modelo respectivo tomando los valores de los parámetros de la misma manera que para el producto A. Ninguno de los parámetros de estos modelos cumplió la hipótesis nula sobre la incorrelación de las simulaciones respectivas, usando la prueba de Box-Pierce.

Tabla 8. Modelos de regresión bayesiana Poisson para todos los productos y MAPE

Modelo de regresión bayesiana Poisson con MCMC		
Producto	Modelo	MAPE
A	$\hat{y}_t = 4.44 + 0.0028t^2 + 0.05998 \operatorname{Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t$	95,73%
B	$\hat{y}_t = 3.8933 + 0.00461 t^2 + 0.3304 \operatorname{Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t$	93,69%
C	$\hat{y}_t = 3.83743 + 0.00445t^2 + \varepsilon_t$	92,73%
D	$\hat{y}_t = 3.6505 + 0.0067t^2 + 0.1728 \operatorname{Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t$	91,09%
E	$\hat{y}_t = 3.6395 + 0.0024 t^2 + 0.1521 \operatorname{Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t$	93,13%
F	$\hat{y}_t = 3.331 + 0.0074 t^2 + 0.167 \operatorname{Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t$	89,62%

Dados los resultados tan poco satisfactorios y la no confiabilidad de los modelos por no cumplir la prueba de incorrelación, se explora el modelo gaussiano, ya que la media de los valores a pronosticar es alta y por ello, la normalidad puede tener un buen desempeño. A continuación se muestran los resultados.

- **Modelo de regresión bayesiana gaussiana**

En la tabla 9 se aprecian los resultados de las estadísticas de la media y desviación estándar para los coeficientes del mejor modelo.

Tabla 9. Media de los parámetros y desviación estándar para modelo de regresión bayesiana Gaussiana con MCMC

Modelo de Regresión bayesiana Gaussiana		
	Mean	SD
(Intercept)	87.4827	8.2195
t²	0.2344	0.1154
Sigma2	342.1060	197.6951

Al igual que en el proceso llevado a cabo anteriormente, de la tabla anterior se toma la media de los valores simulados, para escribir el modelo y reemplazando en (32) queda la siguiente ecuación de pronóstico para el producto A:

$$\hat{y}_t = 87.4827 + 0.2344 t^2 + \varepsilon_t \quad (47)$$

A partir de esta se obtienen los valores finales pronosticados, mostrados en la tabla 10 con su respectivo APE_t y finalmente el valor MAPE de estos 12 valores.

Tabla 10. Pronóstico, APE_t y MAPE mensual para el 2010 estimado con regresión bayesiana Gaussiana con MCMC

Mes	Pronóstico	APE_t	MAPE
Ene	127,10	22,21%	35,06%
Feb	133,43	0,32%	
Mar	140,22	47,60%	
Abr	147,49	16,20%	
May	155,22	18,49%	
Jun	163,43	27,68%	
Jul	172,10	39,92%	
Ago	181,24	54,91%	
Sep	190,85	90,85%	
Oct	200,93	42,51%	
Nov	211,48	32,18%	
Dic	222,50	27,87%	

El porcentaje medio final no es el más adecuado para efectuar pronósticos, para ello se propone otro método bayesiano.

A cada parámetro simulado se realiza la prueba de autocorrelación de orden 2, de Box-Pierce, encontrando para el intercepto, el siguiente resultado:

$$\chi^2 = 7.6849, \text{ Grados de libertad} = 10, \text{ Valor P} = 0.6596$$

Dado el $V_p > 0.05$, se acepta la hipótesis nula (presentada en el ítem 2.5) mostrando que el coeficiente del intercepto no tiene dependencia, siendo de esta forma estacionario. Lo mismo ocurrió para los demás parámetros de este modelo y para todos los parámetros de los modelos del resto de productos. Usando un procedimiento muy similar, se hacen las demás estimaciones y los resultados se muestran en la siguiente tabla.

Tabla 11. Modelos de regresión bayesiana Gaussiana para todos los productos y MAPE

Modelo de regresión bayesiana gaussiana con MCMC					
Producto	Modelo	MAPE	β_0	β_1	β_2
A	$\hat{y}_t = 87.4827 + 0.2344 t^2 + \varepsilon_t$	35,06%	0.65	0.60	-
B	$\hat{y}_t = 58.99955 + 0.09569 t^2 + \varepsilon_t$	18,30%	0.65	0.60	-
C	$\hat{y}_t = 45.621 + 0.273 t^2 - 0.92 \text{ Sen}\left(\frac{2\pi t}{12}\right) + \varepsilon_t$	92,98%	0.95	0.85	0.52
D	$\hat{y}_t = 40.034 + 0.318 t^2 + \varepsilon_t$	127,59%	0.65	0.60	-
E	$\hat{y}_t = 41.51209 + 0.04143 t^2 + \varepsilon_t$	16,27%	0.65	0.60	-
F	$\hat{y}_t = 28.4401 + 0.2813 t^2 + \varepsilon_t$	121,96%	0.65	0.60	-

No se muestra el valor P de la prueba Box-Pierce para ε_t , pues el valor de este error es tan imperceptible, que resulta insignificante en el modelo si no cumpliera la prueba.

Para todos los productos, excepto para el C, de los modelos explorados con las distintas covariables mostradas en la respectiva sección del marco teórico, los modelos con mejor MAPE llevaron siempre el tiempo al cuadrado y no se le incorporó variable trigonométrica. Por ejemplo, al explorar un modelo sin

covariable trigonométrica para el producto C, se obtuvo un MAPE del 96.80% por lo que mejor se eligió el modelo mostrado en la tabla anterior.

4.5. Pronóstico con prima bayesiana

Para llevar a cabo el pronóstico, se utiliza las ecuaciones (44) y (45), la primera, que no incorpora términos autorregresivos puesto que es el primer pronóstico del año, se utiliza para pronosticar el mes de enero de 2011 y la segunda para el resto de meses.

A continuación se muestra el modelo con los valores tomados de la tabla 12.

$$\hat{Y}_{ene=12} = \left(200 + \sum_{i=1}^{12} x_i \right) \left(\frac{200}{1 + 200(12)} \right) \quad (48)$$

Donde $\sum_{i=1}^{12} x_i$ es la suma de las ventas del año 2010.

$$\hat{Y}_{jul=18} = \left(400 + 0.1(200) + \sum_{i=1}^{18} x_i \right) \left(\frac{400 + 0.1(200)}{1 + (400 + 0.1(200))18} \right) \quad (49)$$

Donde $\sum_{i=1}^{18} x_i$ es la suma de las ventas del 2010 y el acumulado de ventas hasta junio del 2011

Los valores de α y β se seleccionan arbitrariamente como semilla de un proceso de optimización posterior con el algoritmo Tabú.

En la siguiente tabla se muestran los α y β arbitrarios, la n_{t-1} correspondiente según el mes, la sumatoria de los x_i , el pronóstico hallado con la ecuación (44) para enero y (45) para el resto de meses y los porcentajes de APE_t y MAPE.

Tabla 12. Valores para variables del modelo de prima bayesiana, APE_t y MAPE mensual para 2010

Producto A							
Mes	α	β	n_{t-1}	$\sum_{i=1}^n x_i$	Pronóstico	APE_t	MAPE
Ene	200	200	12	1202	116,78	12,29%	16,56%

feb	200	200	13	1306	117,34	11,77%	
Mar	200	200	14	1439	118,46	24,70%	
Abr	200	200	15	1534	116,90	33,58%	
May	200	200	16	1710	120,59	7,95%	
Jun	200	200	17	1841	121,20	5,31%	
Jul	400	400	18	1969	132,70	7,89%	
Ago	400	400	19	2092	133,25	13,89%	
Sep	400	400	20	2209	132,43	32,43%	
Oct	400	400	21	2309	130,89	7,17%	
Nov	400	400	22	2450	131,35	17,91%	
Dic	400	400	23	2610	132,60	23,80%	

En la tabla anterior se aprecian cómo los valores de n y la sumatoria se van actualizando a cada tiempo con los nuevos valores del mes anterior, partiendo desde los 12 datos del año anterior.

Se aplica el mismo proceso al resto de productos y se muestra el MAPE obtenido en la próxima tabla.

Tabla 13. MAPE 2011 para todos los productos obtenidos con prima bayesiana

	Prima bayesiana					
MAPE (%)	A	B	C	D	E	F
2011	17,06%	19,59%	20,50%	21,99%	12,94%	16,21%

Se logra mejorar el MAPE de cada producto sin embargo algunos valores aún tienen un porcentaje alto, debido quizás a los α y β no-óptimos que se utilizaron en el modelo.

El siguiente método incorporará esos valores optimizados y se lograrán nuevos resultados.

4.6. Método de optimización Tabú

A continuación se muestra la tabla con los valores óptimos α y β obtenidos con el algoritmo de optimización después de optimizar la función objetivo.

Tabla 14. Valores α y β optimizados con Tabú

Mes	A	B	MAPE
------------	----------	----------	-------------

Ene	45,7	399,8	11,67%
Feb	175,7	398,1	
Mar	(126,4)	404,6	
Abr	161,9	398,9	
May	202,7	389,7	
Jun	193,5	404,2	
Jul	227,2	406,2	
Ago	108,6	392,7	
Sep	79,5	411,3	
Oct	276,3	411,1	
Nov	305,2	426,5	
Dic	249,6	403,6	

Obsérvese cómo varían los valores a lo largo del tiempo muy pocas veces pareciéndose a los valores arbitrarios escogidos en el modelo anterior.

Al utilizar estos parámetros α y β que encuentran el mejor valor del MAPE, hallados con el algoritmo Tabú, se procede a reemplazar estos valores en el modelo bayesiano con ecuación (44) para enero y (45) para los demás meses y se obtienen las ecuaciones que se muestran en la tabla 15, diferenciada por mes. El respectivo error relativo se muestra en la última columna (APE_t).

Tabla 15. Modelos de prima bayesiana con parámetros α y β optimizados para producto A en 2010

Producto A		
Mes	Modelo	APE_t
Ene	$\hat{Y}_{12} = \left(45,7 + \sum_{i=1}^{12} x_i \right) \left(\frac{399,8}{1 + 399,8(12)} \right)$	0,04%
Feb	$\hat{Y}_{13} = \left(175,7 + 0.1 (45,7) + \sum_{i=1}^{13} x_i \right) \left(\frac{398,1 + 0.1(399,8)}{1 + (398,1 + 0.1(399,8))13} \right)$	14,06%
Mar	$\hat{Y}_{14} = \left(-126,4 + 0.1 (175,7) + \sum_{i=1}^{14} x_i \right) \left(\frac{404,6 + 0.1(398,1)}{1 + (404,6 + 0.1(398,1))14} \right)$	0,01%
Abr	$\hat{Y}_{15} = \left(161,9 + 0.1 (-126,4) + \sum_{i=1}^{15} x_i \right) \left(\frac{398,9 + 0.1(404,6)}{1 + (398,9 + 0.1(404,6))15} \right)$	36,25%

Producto A		
Mes	Modelo	APE _t
May	$\hat{Y}_{16} = \left(202,7 + 0.1 (161,9) + \sum_{i=1}^{16} x_i \right) \left(\frac{389,7 + 0.1(398,9)}{1 + (389,7 + 0.1(398,9))16} \right)$	7,98%
Jun	$\hat{Y}_{17} = \left(193,5 + 0.1 (202,7) + \sum_{i=1}^{17} x_i \right) \left(\frac{404,2 + 0.1(389,7)}{1 + (404,2 + 0.1(389,7))17} \right)$	5,58%
Jul	$\hat{Y}_{18} = \left(227,2 + 0.1 (193,5) + \sum_{i=1}^{18} x_i \right) \left(\frac{406,2 + 0.1(404,2)}{1 + (406,2 + 0.1(404,2))18} \right)$	0,06%
Ago	$\hat{Y}_{19} = \left(108,6 + 0.1 (227,2) + \sum_{i=1}^{19} x_i \right) \left(\frac{392,7 + 0.1(406,2)}{1 + (392,7 + 0.1(406,2))19} \right)$	0,00%
Sep	$\hat{Y}_{20} = \left(79,5 + 0.1 (108,6) + \sum_{i=1}^{20} x_i \right) \left(\frac{411,3 + 0.1(392,7)}{1 + (411,3 + 0.1(392,7))20} \right)$	14,95%
Oct	$\hat{Y}_{21} = \left(276,3 + 0.1 (79,5) + \sum_{i=1}^{21} x_i \right) \left(\frac{411,1 + 0.1(411,3)}{1 + (411,1 + 0.1(411,3))21} \right)$	12,43%
Nov	$\hat{Y}_{22} = \left(305,2 + 0.1 (276,3) + \sum_{i=1}^{22} x_i \right) \left(\frac{426,5 + 0.1(411,1)}{1 + (426,5 + 0.1(411,1))22} \right)$	20,95%
Dic	$\hat{Y}_{23} = \left(249,6 + 0.1 (305,2) + \sum_{i=1}^{23} x_i \right) \left(\frac{403,6 + 0.1(426,5)}{1 + (403,6 + 0.1(426,5))23} \right)$	23,80%

Aquí el APE_t se logra llevar a niveles del 0% en algunos meses lo que repercutirá de manera positiva en el MAPE final, en contraste los valores altos como el de abril lo castigarán.

En la tabla 16 se resumen los valores α y β finales, el n_{t-1} respectivo por mes, la sumatoria de los datos hasta el mes anterior, el pronóstico hallado, el error absoluto porcentual y el MAPE, para el producto A.

Tabla 16. Valores para variables del modelo de prima bayesiana con parámetros α y β optimizados, APE_t y MAPE mensual para producto A en 2010

Producto A							
Mes	α	β	n	$\sum_{i=1}^n x_i$	Pronóstico	APE _t	MAPE
Ene	45,7	399,8	12	1202	103,96	0,04%	11,67%
feb	175,7	398,1	13	1306	114,31	14,06%	
Mar	-126,4	404,6	14	1439	94,99	0,01%	
Abr	161,9	398,9	15	1534	112,20	36,25%	
May	202,7	389,7	16	1710	120,54	7,98%	
Jun	193,5	404,2	17	1841	120,85	5,58%	
Jul	227,2	406,2	18	1969	123,07	0,06%	
Ago	108,6	392,7	19	2092	117,00	0,00%	
Sep	79,5	411,3	20	2209	114,95	14,95%	
Oct	276,3	411,1	21	2309	123,48	12,43%	
Nov	305,2	426,5	22	2450	126,48	20,95%	
Dic	249,6	403,6	23	2610	132,60	23,80%	

La tabla anterior cumple el mismo análisis que la número 12 pero con variación de los resultados debido a la incorporación de los α y β optimizados.

Finalmente se muestra la tabla resumen con el MAPE de cada uno del resto de los productos para el año 2011.

Tabla 17. MAPE 2011 para todos los productos obtenidos con prima bayesiana con optimización Tabú

Producto	Prima bayesiana con optimización tabú					
	A	B	C	D	E	F
MAPE (%)	11,67%	12,88%	10,37%	9,78%	7,74%	3,64%

Efectivamente, el MAPE se logró mejorar inmensamente con este método como se pensó en algún momento. Nótese que los valores de MAPE oscilan más o menos entre 3 y 10%, lo que representa un muy buen aporte para los que se interesen en esta técnica.

5. SINTESIS DE RESULTADOS

Se presenta la tabla 18 en la cual se muestran el MAPE respectivo por modelo explorado y por producto para el año 2011.

Tabla 18. Tabla resumen para todos los modelos y productos

Modelo	MAPE 2011 por producto					
	A	B	C	D	E	F
Variable aleatoria	23,32%	28,86%	25,74%	20,81%	32,86%	24,66%
R. Poisson	95,73%	93,69%	92,73%	91,09%	93,13%	89,62%
R. Gaussiana	35,06%	18,30%	92,98%	127,59%	16,27%	121,96%
Prima bayesiana	17,06%	19,59%	20,50%	21,99%	12,94%	16,21%
Prima. Bay. con Opt Tabú	11,67%	12,88%	10,37%	9,78%	7,74%	3,64%

Nótese la inmensa mejora que aporta el método de optimización Tabú en el MAPE de cada producto, especialmente en el F.

- **Comparación entre modelos**

En la tabla 19 se muestra una comparación de ventajas y desventajas de cada modelo utilizado en este estudio, a partir de la percepción de la investigación que implicó la realización de cada modelo.

Tabla 19. Tabla comparativa entre modelos explorados

Modelo	Ventaja	Desventaja
Variable aleatoria	<ul style="list-style-type: none"> • No requiere grandes cantidades de datos. • No se basa en supuestos de residuales. 	<ul style="list-style-type: none"> • No capta tendencia. • Poco exacto • Riesgo de inflación del inventario
Regresión Poisson Bayesiana	<ul style="list-style-type: none"> • Uso de pocos datos • Simulado con paquete R 	<ul style="list-style-type: none"> • Aplicable sólo para datos que se ajusten a una distribución Poisson. • Para este trabajo no tuvo mucha exactitud.
Regresión Gaussiana Bayesiana	<ul style="list-style-type: none"> • Aplicable para datos que se ajusten a una distribución Normal o que se puedan aproximar a ésta. 	<ul style="list-style-type: none"> • La programación para modificar las distribuciones del paquete en R, puede ser compleja.
Prima bayesiana	<ul style="list-style-type: none"> • No requiere pruebas de validación. • Tiene en cuenta los conocimientos previos (a priori). • Fácil actualización de datos a medida que se presenten. • Incorpora información de expertos • Es versátil en el manejo de las distribuciones a priori y de los datos. 	<ul style="list-style-type: none"> • Cálculos analíticos pesados. • Incorpora información subjetiva donde su base de aplicación se puede sustentar en interpretaciones erróneas de teorías estadísticas-matemáticas
Prima con Optimización Tabú	<ul style="list-style-type: none"> • Optimiza sin repetir puntos ya evaluados y actualiza el dato óptimo cada vez que encuentra uno mejor, desechando el anterior. 	

La comparación de los modelos requiere de un análisis más profundo, siendo este el objeto principal del estudio.

6. PROCEDIMIENTOS PARA REALIZAR PRONÓSTICOS UTILIZANDO LOS MÉTODOS PROPUESTOS

Los modelos estudiados pueden ser utilizados para cualquier tipo de datos, pero se hizo énfasis en el pronóstico a partir de pocos datos. El interesado puede tomar alguno de los modelos y estimarlos para evaluar sus resultados, eso sí, partiendo de un análisis previo de sus variables y de la fluctuación de sus ventas a lo largo del tiempo para tomar la decisión más idónea.

Lo más importante es que el o los modelos escogidos, en realidad le generen valor a la empresa que los utilice, siendo siempre dinámicos probando éste y aquél con el fin de obtener el mejor resultado. El valor agregado de este estudio es el pronóstico de difícil estimación a partir de pocos datos, caso que se da en empresas nuevas o que desean lanzar un nuevo producto, sin embargo, como ya se aclaró, puede ser útil para otro tipo de casos.

Como ya se sabe, para el caso de este estudio se utilizaron los datos mensuales del año 2010 para realizar modelos estadísticos y pronosticar el 2011 para luego proceder y hallar eficiencia con el MAPE.

En esta sección se muestra cómo utilizar los diferentes modelos hallados pronosticando algunos meses del 2012, que para este caso, sería el pronóstico futuro pues no se cuenta con los datos de ese año.

El caso del pronóstico utilizando variable aleatoria se omite, pues lo único que se requiere hacer es generar los pronósticos de los 12 meses a partir de una simulación de una variable aleatoria Poisson de parámetro λ de al menos 1.000 iteraciones, lo cual puede lograrse con un algoritmo de programación sencillo que lleve a cabo ciclos repetitivos. Para este caso se utilizó el software R como se comentó en la metodología.

Es muy importante realizar la prueba de bondad de ajuste para evaluar la distribución a la cual los datos mejor se ajustan, así se debe generar la respectiva variable aleatoria con sus parámetros correspondientes.

6.1. Regresión Gaussiana y Poisson bayesiana con MCMC.

Para este trabajo, el modelo Poisson no resultó ser útil pues no se cumplió la prueba de incorrelación Box-Pierce y el MAPE obtenido al final fue demasiado alto, estando alrededor del 90% para todos los productos

El modelo Gaussiano tampoco resultó ser muy eficiente debido a que se obtuvo MAPE muy ineficiente para algunos productos a pesar que otros resultaron mucho mejor que el Poisson.

De todas maneras otro tipo de datos pueden dar lugar a resultados mejores que interesen el uso de este tipo de modelos, por ello se procede a explicar su aplicación.

Primero los datos deben estar ajustados a una distribución Poisson y se debe asumir normalidad a priori para los parámetros, para el caso del modelo de regresión Poisson bayesiana. Para el gaussiano los datos deben estar ajustados a una distribución normal y se asume a priori semi-conjugada estándar para los parámetros, ambos requisitos se encuentran ya programados en el paquete MCMCPack del software R, concretamente en las funciones MCMCRegress para el gaussiano y MCMCPoisson para el Poisson.

Se prosigue estimando un modelo utilizando en lo posible, covariables de tiempo, trigonométricas, de nivel y/o ficticias, mostradas en la metodología.

A continuación se muestra un ejemplo con el modelo hallado para el producto A de la regresión bayesiana Poisson y se omite uno para el gaussiano, ya que después de que el modelo sea estimado, el proceso es idéntico para pronosticar.

Se tendría pues, un modelo como el de la ecuación (51) donde sus parámetros se toman de la tabla donde se muestra la media de cada uno de los valores, como se mostró en la tabla 7.

$$\hat{y}_t = 4.44 + 0.0028t^2 + 0.05998 \operatorname{Sen}\left(\frac{2\pi t}{12}\right) \quad (51)$$

Para pronosticar, por ejemplo, enero del 2012 se reemplaza $t=25$ y se obtiene el pronóstico correspondiente y así sucesivamente.

Se recomendaría actualizar el modelo al ir obteniendo los datos reales, el problema, a diferencia del pronóstico con la prima bayesiana es que se debe modelar todo de nuevo con los nuevos datos y un nuevo modelo con otras covariables podría darse.

6.2. Prima bayesiana con optimización Tabú

Para pronosticar 2012 se recurre a la ecuación (45) y se reemplazan los valores respectivos según el mes a pronosticar. Para el caso de enero, teniendo en cuenta que el mes anterior es diciembre de 2011 se tendría:

$$\hat{Y}_{ene\ 2012} = \left(45,7 + 0.1 (249,6) + \sum_{i=1}^{24} x_i \right) \left(\frac{399,8 + 0.1(403,6)}{1 + (399,8 + 0.1(403,6))24} \right) \quad (52)$$

Donde se utiliza α_t correspondiente al valor optimizado de enero, α_{t-1} sería el valor optimizado para diciembre. Lo mismo se aplicaría para β_t y β_{t-1} .

Se debe tener en cuenta que la sumatoria de los x_i incluiría los valores de 2010 y los ya actualizados para el 2011; en caso de aún no contar con los datos reales se pueden utilizar los pronosticados, pues de alguna forma se está confiando en ellos y por ende se pueden utilizar para pronosticar.

CONCLUSIONES

1. La prima bayesiana usando las distribuciones a priori Gamma y de los datos Poisson, combinado con el método de optimización Tabú, resultó ser el mejor de todos los explorados en este estudio para pronosticar, estando el MAPE entre 3 y 13%, lo que significa una amplia mejora ya que la prima bayesiana por sí sola alcanzó pronósticos alrededor del 20%. Resulta este método como el mejor resultado, al ser aplicado a los datos presentados.
2. En contraste, el modelo de regresión bayesiano Poisson es el que da el MAPE más deficiente, oscilando entre el 90 y 100%. Sin embargo esto no quiere decir que estos modelos se deban desechar, pues también depende mucho del tipo de datos que se tengan y para otro caso podrían ser de mucha utilidad. Se anima entonces al interesado a explorar este modelo cuando tenga datos que se ajusten a una distribución Poisson.
3. El modelo de regresión bayesiano gaussiana, presentó mejores pronósticos que su homólogo de regresión Poisson. De todas formas, al igual que cualquier otro modelo, puede ser útil en otras situaciones particulares.
4. El método de generación de variable aleatoria puede ser una buena opción cuando existan pocos datos, debido a su sencillez, sin embargo puede no ser muy confiable por su carencia teórica para predecir tendencia y fluctuación futura.
5. Las fluctuaciones agresivas que presentan los datos es una de las causas por la que es difícil realizar pronósticos con una técnica estadística, así como la existencia de pocos datos históricos, por ello, el contenido de este estudio aporta alternativas para realizar pronósticos de demanda para productos de difícil estimación, encaminadas a superar este tipo de dificultades.
6. La metodología para estimar la prima bayesiana con la cual se realizan pronósticos, en general, constituye una herramienta confiable estadísticamente para la planeación logística en el lanzamiento de nuevos productos sin recurrir sólo al pensamiento subjetivo. Específicamente, el caso aplicado de la prima bayesiana con las distribuciones Gamma-Poisson y el algoritmo Tabú, mostró su eficiencia para hacer predicción y facilitar la planeación de inventarios de esta empresa.

RECOMENDACIONES

En toda serie de tiempo es importante hacer un análisis exploratorio de la correlación de los datos, así como de covariables que pueden ser determinantes para mejorar toda explicación asociada a su variabilidad y dependencia.

La presencia de pocos datos puede no ser un problema con este tipo de técnicas bayesianas, para lo cual es importante también conocer la opinión de expertos acerca del comportamiento de las ventas.

Las propuestas mostradas son aplicables a numerosos casos en que se requiera pronosticar con datos que tengan una naturaleza discreta o incluso, Poisson, sin embargo, una escala de grandes valores en la variable aleatoria permite también usar las aproximaciones con la distribución normal, lo cual generaría otro tipo de investigaciones donde se combinen diferentes distribuciones a las usadas en este trabajo.

Las metodologías bayesianas presentadas en el trabajo que explican los procesos generales que se pueden llevar a cabo con otro tipo de combinaciones distribucionales, permite que este trabajo pueda aportar procedimientos genéricos para ser replicados en otras situaciones que abarquen un par de distribuciones de probabilidad, a priori y para los datos o investigaciones donde se requiera efectuar pronósticos.

Se logra mostrar en el trabajo que el software R tiene muchas ventajas con respecto a su libre acceso y su versatilidad en cuanto a la programación estadística, lo cual sería una ventaja competitiva para muchas empresas que no posean recursos suficientes para comprar una licencia de otro paquete y requieran establecer programaciones de su producción y ventas a corto y largo plazo.

REFERENCIAS

- Albert, J. (1997). *MATLAB as an environment for Bayesian Computation*. OH: Bowling Green State University.
- Alea, M. V. (2001). *Estadística con SPSS v.10.0*. Barcelona: Edicions Universitat.
- Andrieu, C., De Freitas, N., Doucet, A., & Jordan, M. (2003). An introduction to MCMC for Machine Learning. *Machine Learning*, 50, 5-43.
- Bayes, T. (1763). Essay Towards Solving a Problem in the Doctrine of Chance. *Philosophical Transactions of the Royal Society of London* 53, 370-418.
- Berry, J. O. (1988). Statistical Analysis and the illusion of objectivity. *American Scientist* vol 76, N° 2, 159-165.
- Bolstad, W. M. (2007). *Introduction to Bayesian Statistics* (2da ed.). Hamilton, NZ: Wiley.
- Bourke, P. (2001). *Miscellaneous functions*. Recuperado el 28 de 06 de 2013, de <http://paulbourke.net/miscellaneous/functions/>
- Bowerman, B. L., O'Connell, R. T., & Koehle, A. B. (2007). *Pronóstico, series de tiempo y regresión* (4ta ed.). Monterrey: Cengage Learning.
- Canavos, G. C. (1988). *Probabilidad estadística. Aplicaciones y métodos*. México D.F: McGraw Hill.
- Chelouah, R., & Siarry, P. (2005). A hybrid method combining continuous tabu search and Nelder–Mead simplex algorithms for the global optimization of multimínima functions. *European Journal of Operational Research*, 161(3), 636–654.
- Cohen et al. (Nov-Dic 1986). Forecasting for Inventory Control: An Example of When "Simple" Means "Better". *Interfaces*, 16, 95-99.
- Correa, J. C. (2005). *Estadística Bayesiana*. Medellín: Universidad Nacional - Sede Medellín.
- Coss Bú, R. (1992). *Simulación. Un enfoque práctico*. México D.F: Limusa. Grupo Noriega Editores.
- Dawid, A. (1984). Statistical theory: The prequential approach . *Journal of the Royal Statistical Society Series*, 278-292.

- Eppen, G. D. (2000). *Investigación de operaciones en la ciencia administrativa*. México: Prentice-Hall.
- Fader, P. S., Hardie, B. G., & Huang, C.-Y. (2004). Dynamic Changepoint Model for New Product Sales Forecasting. *Marketing science* Vol. 23, N°1, 50-65.
- Fildes, R., & Goodwin, P. (2007). Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting. *Interfaces*, Vol. 37, No. 6, 570-576.
- Flora Lu, Q. (2005). *Bayesian Forecasting of Stock Prices*.
- Flores, E. (2012). Teoría de Credibilidad en el Cálculo de Reservas. *IX Coloquio Internacional de Estadística*, (pág. 15). Medellín.
- Geweke, J., & Whiteman, C. (2004). Bayesian forecasting. *The Handbook of Economic Forecasting*, 1-78.
- Geyer, C. (2011). Introduction to Markov Chain Monte Carlo. En S. Brooks, A. Gelman, G. L. Jones, & X.-L. Meng, *Handbook of Modern Statistical Methods: Handbook of Markov Chain Monte Carlo* (págs. 3-47). Chapman and Hall/CRC.
- Gill, J. (2002). Bayesian Methods-A social and Behavioral Sciences Approach (p. 459). 459.
- Glover, F. (1990). Tabu search part II. *ORSA J. , . ORSA J. Computing* 2, 4-32.
- González, J. C., & Franco, J. R. (2013). *Propuesta metodológica para la optimización de inventarios utilizando modelos de series temporales*. Medellín.
- González, M., & Lobato, I. (2007). Contrastes de autocorrelación. *Gaceta de Economía Año 7, Número 14*, 41-57.
- Hillier, F. S., & Lieberman, G. J. (2002). *Investigación de operaciones*. México: Mc Graw-Hill.
- Krajewski, L., Ritzman, L., & Malhotra, M. (2008). *Administración de operaciones* (8va ed.). México: Pearson educación.
- Lee, J., Boatwright, P., & Kamakura, W. A. (2003). A Bayesian Model for Prelaunch Sales Forecasting of Recorded Music. *Management science*. Vol. 49, N°2, 179-196.

- López, M. V. (2012). Software para la generación de variables aleatorias empleadas en simulación. *VII Workshop de Investigadores en Ciencias de la Computación*. Corrientes: Universidad Nacional del Nordeste. 9 de Julio nº 1449. CP: 3400.
- Lynch, S. M. (2006). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. Springer.
- Marien, B. (2009). *Faculté de Science Politique et de droit. Département de Science Politique*. Recuperado el 20 de Febrero de 2013, de <http://www.politique.uqam.ca/component/content/article/23-archives-horaires-1er-cycle/10-pol-1800-20-40-50-bruno-marien.html>
- Martin, A. D., Quinn, K. M., & Park, J. H. (2011). MCMCpack: Markov chain Monte Carlo in R. *Journal of Statistical Software Vol. 42 Issue. 9*, 1-21.
- Martin, A. D., Quinn, K. M., & Park, J. H. (2013). *Package MCMCpack version 1.2 - 4.1*.
- McKenzie, J. (2006). Mean absolute percentage error and bias in economic forecasting. *Economics Letters*, 259-262.
- Montgomery, D. C., & Runger, G. C. (2003). *Applied statistics and probability for engineers*. New York: John Wiley & Sons, Inc.
- Neelamegham, R., & Chintagunta, P. (1999). A Bayesian Model to Forecast New Product Performance in Domestic and International. *Marketing science, Vol. 18, N°2*, 115-136.
- Raftery, A. E. (1999). Bayesian Model Selection in Social Research. *Sociological Methodology*, 111-163.
- Serrano, J. (2003). *Iniciación a la Estadística Bayesiana*. Madrid: LA MURALLA, S.A.
- Silva Aycaguer, L. C., & Benavides, R. A. (2003). Apuntes sobre subjetividad y estadística en la investigación en salud. *Revista cubana Salud Pública*, 170-173.
- StatPoint Technologies, Inc. (2009). *STATGRAPHICS® CENTURION XVI. USER MANUAL*.
- Synder, R. (2002). Forecasting sales of slow and fast moving inventories. *European Journal of operational Research*, 648-699.

- Taha, H. A. (2004). *Investigación de operaciones*. México: Pearson education.
- Urban, G. L., Hauser, J. R., & Roberts, J. H. (1990). Prelaunch Forecasting of New Automobiles. *Management Science*, Vol. 36, No. 4, 401-421.
- Valencia, M., Correa, J. C., Díaz, J. F., & Ramírez, S. (2013). Optimización aplicada a predicción con estadística bayesiana. *VI Simposio Internacional de Ingeniería Industrial: Actualida y Nuevas Tendencias 2013. Red Internacional de Investigadores en Ingeniería Industrial*. ISSN: 1856-8351.
- Venables, Smith, the R Core Team. (2012). *An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics*.
- Vidal, V. G. (marzo 2008). Modelos de gestión de inventarios en cadenas de abastecimiento: Revisión literaria. *Facultad de Ingeniería Universidad de Antioquia*(043), 137.
- Walpole, R. E. (1999). *Probabilidad y estadística para ingenieros*. México: Prentice-Hall Hispanoamericana S.A.
- Winston, W. L. (2004). *Ivestigación de operaciones: aplicaciones y algoritmos* (4ta ed.). México: Thomson.