



**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE INGENIERIA**



Elaboración del pronóstico de llamadas atendidas en un Centro de Atención Telefónica,
mediante Simulación Monte Carlo

ACT. ARTURO LOPEZ SANCHEZ



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

A mi Madre,
"Rosita La Bonita" que siempre me ha impulsado, hasta en lo que no le parece

A mi Padre,
"Aguilita" por el gran ejemplo que sigue siendo y nunca lo ha notado

A mi Hermano,
"El Pollo" por que a pesar de todo siempre me ha dado su apoyo

A mis amigos y amigas,
Por el gusto de compartir y tener un pretexto mas para "chelear"

“Un grado de estudios mas elevado que los demás,
no te hace mas inteligente,
pero sí mas intransigente”

INDICE

	Pag
Introducción	I
Capitulo I. Funcionamiento de un Centro de Atención Telefónica	
1.1 Funcionamiento de un Centro de Atención Telefónica (CAT)	1
1.1.1 Llamadas Inbound	
1.1.2 Llamadas Outbound	
1.2 Principales Indicadores de un CAT	7
1.2.1 Medidores de Servicio	
1.2.2 Medidores de Calidad	
1.2.3 Medidores de Eficiencia	
1.3 Administración de un CAT	11
Capitulo II. Análisis de Datos, Modelos Estáticos	
2.1 Exploración de datos	14
2.1.1 Análisis de Frecuencias	
2.1.2 Descripción Gráfica de los Datos	
2.2 Funciones de Distribución	17
2.2.1 Distribución Normal	
2.2.2 Distribución F	
2.2.3 Distribución t	
2.3 Análisis de Regresión	20
2.3.1 Conceptos Básicos	
2.3.2 Validar el modelo	
Capitulo III. Análisis de Datos, Modelos de Series de Tiempo	
3.1 Conceptos preliminares de las Series de Tiempo	27
3.1.1 Procesos Estocásticos	
3.1.2 Pruebas de estacionariedad	
3.2 Modelos Lineales de Series de Tiempo	30
3.2.1 Promedios Móviles	
3.2.2 Suavización Exponencial Simple	
3.3 Pronóstico con Modelos de Series de Tiempo	33
3.3.1 Consideraciones generales	
3.3.2 Técnicas de Pronóstico	

Capitulo IV. Modelación y Simulación, Técnica Monte Carlo	
4.1 Modelos	38
4.1.1 Conceptos Preliminares	
4.1.2 Tipos de Modelos	
4.2 Simulación	42
4.2.1 Conceptos Básicos	
4.2.2 Pasos para realizar un estudio de simulación	
4.2.3 Verificación del modelo y validación	
4.3 Simulación de Monte Carlo	48
4.3.1 Números Aleatorios	
4.3.2 Técnica de simulación Monte Carlo	
Capitulo V. Aplicación práctica de la simulación, caso sistema de TV de paga	
5.1 Análisis de la información	56
5.1.1 Estadísticas de un CAT	
5.1.2 Pronóstico en un CAT	
5.2 Modelación en un CAT	61
5.2.1 Llamadas entrantes	
5.3 Simulación de Monte Carlo en hoja de calculo	63
5.3.1 Simulación operación In Bound	
5.3.2 Validación	
5.3.3 Interpretación de los resultados	
Conclusiones	72
Anexo	
Anexo I. Generador para Variables Uniformes	75
Anexo 2A. Cálculo para la Validación del Modelo de Simulación	76
Anexo 2B. Cálculo para la Validación del Modelo de Simulación II	77
Anexo 2C. Programación Visual Basic para obtener los resultados de validación	78
Anexo 3. Resumen validación del modelo de Simulación	85
Anexo 4A. Fórmulas para derivar los modelos de pronóstico, Promedios Móviles, tercer rezago	86
Anexo 4B. Suavización exponencial Simple	87
Anexo AC. Suavización exponencial Doble	88
Anexo 4D. Mínimos Cuadrados Ordinario	89
Bibliografía	90

INTRODUCCION

La atención personalizada “uno a uno” que solía ofrecerse a los usuarios de los servicios financieros, seguros, de correos, tramites gubernamentales, etc., ha dejado de ser una realidad para convertirse en una pesadilla por las largas horas en una fila esperando a ser atendidos por un ejecutivo de servicio, esto se debe al constante crecimiento de los clientes y usuarios que exigen agilidad por parte del personal que los atiende, rapidez y lo mas importante calidad en el servicio. A su vez la constante expansión de las empresas, la diversificación y especialización en los servicios y la complejidad tecnológica, tienen un efecto negativo en la atención de las demandas de los clientes y usuarios.

En un intento por resolver esta problemática las empresas optaron por un nuevo proceso que pudiera brindar atención a los usuarios en menos tiempo y con la misma calidad, lo que resulto en incorporar un Centro de Atención Telefónica (CAT) que representa un canal eficiente de comunicación entre los usuarios y las Instituciones prestadoras de servicios sin la necesidad de asistencia presencial.

Un CAT es una plataforma telefónica que tiene la función de facilitar la comunicación entre las personas, por medio del teléfono. Permite el contacto directo de una Institución privada o Gubernamental con el cliente o usuario que permite a su vez que sean contactados para solucionar sus problemas e inquietudes. El mayor valor agregado sustancial es el registro del historial de contactos y operaciones realizadas, su análisis proporciona los elementos suficientes para desarrollar estrategias en pos de mejorar la atención a sus clientes.

A través del mismo, se ofrecen servicios de recepción de llamadas (inbound) y llamadas de salida (outbound), estos servicios es susceptible dividirlos en grupos siguientes: Atención a Clientes, Campañas de Venta y Telemarketing,

Gestión de Cobranza, Mesas de Ayuda e Internet Contact Center, por mencionar algunas.

Si bien los proyectos de un CAT se conciben en términos de llamadas telefónicas y computadores, el objetivo final que es la optimización del servicio de atención a clientes como un medio para elevar la rentabilidad Institucional. Para llevar a cabo tal actividad, los administradores realizan medidas para conocer el desempeño de sus sistemas y asegurar el óptimo nivel de servicio a los clientes.

Las mediciones mas comunes incluyen: nivel de servicio, nivel de abandono, tiempo de respuesta a un requerimiento, duración promedio de llamada, tiempo promedio de abandono. El nivel de servicio es afectado por tres variables que son el tiempo promedio de la llamada, el volumen de llamadas recibidas y número de operadores disponibles para atenderlas. Como se podrá advertir, si el volumen de llamadas se incrementa pero el número de agentes y el tiempo promedio de la llamada permanece inalterado, entonces el nivel de servicio se verá afectado.

Aunque se conoce el efecto por una variación en las mediciones por un cambio en cualquiera de estas, en la práctica no hay cabida para experimentar sin afectar el rendimiento de la operación, no obstante el presente estudio muestra que un modelo de simulación simboliza una amplia gama de posibilidades resultantes por practicar del ensayo y error con la bondad de no afectar la operación.

La revisión general del desempeño de los CAT se basa en sus resultados así como la detección de sus oportunidades de mejorar en sus principales aspectos: niveles de servicio, productividad de los operadores, calidad en la atención telefónica, utilización adecuada de recursos, y aseguramiento de calidad de la operación. Aspectos importantes a considerar para mejorar el desempeño son:

Análisis de tráfico: Días y horarios pico

- Revisión estadística del desempeño histórico del volumen de llamadas.
- Identificación de semanas, días y horas pico.
- Distribución de llamadas por día y hora de la semana: pronóstico de llamadas por hora y día de la semana
- Criterios para manejo de días y periodos críticos

Manejar en forma óptima los periodos y días críticos

- Existen fechas específicas que requieren una planeación especial: días festivos, puentes, periodos vacacionales, entrada y regreso a clases, etc. Un manejo adecuado de estos días permite ahorros importantes y manejar productivamente el tiempo del personal.

Si adicionalmente se agregan nuevos servicios o el volumen de llamadas es estacional, las horas picos en los volúmenes de llamadas deberán ser gestionadas con el número de agentes apropiado para cumplir con un nivel de servicio adecuado. Deberá entonces predecirse las necesidades de agentes a base de información histórica de llamadas para no perturbar el nivel de servicio, y una técnica aplicable para predecir con un buen grado de certeza es el análisis de Series de Tiempo.

Un punto importante en el tratamiento de la información histórica del volumen de las llamadas, tiempo promedio de llamada y el número de operadores asignados, es la experiencia que se tiene en la gestión de datos orientado a detectar donde se están creando los problemas y sobre esta base plantear una solución para tomar la mejor decisión.

Además, el objetivo principal de llevar a cabo reseñas históricas, medición de datos, pronóstico del volumen esperado de llamadas, control de los ejecutivos

en horas pico, niveles de servicio u otros, es lograr la prosecución de la calidad del servicio a través de una mejora continua en el seno de la Institución, de los administradores y supervisores de la operación, de sus ejecutivos y del personal en general.

Conjuntamente el acelerado crecimiento de los usuarios de los diversos servicios ha traído consigo un elemento de improvisación en la instrumentación y puesta en marcha de los CAT viéndose obligados a dedicar horas extras de trabajo tratando de investigar y detectar los motivos por los cuales, el servicio no alcanza los niveles y objetivos establecidos. Como los índices de demora en la atención de la llamada, y de abandono, como el nivel de calidad de servicio que superan ampliamente los parámetros establecidos.

Dentro de las tendencias actuales en el manejo de un Centro de Atención Telefónica se destaca la precisión en las métricas y análisis histórico de sucesos con relevada importancia, basado en el volumen de llamadas recibidas por tal evento.

Esto es posible solamente si la estrategia del planeamiento se armoniza con las opciones con respecto tecnología, procesos de organización y de personal. El planeamiento correcto traduce inmediatamente en ahorros de espacio, ventajas operacionales y costes futuros reducidos.

En un orden alternativo respecto a la administración de un CAT se contrasta el modelo de administración actual, es decir, lo antes mencionado con un modelo que involucra mayor complejidad como lo es la simulación Monte Carlo. En este estudio no se analizan todas las particularidades en la administración de un CAT debido a lo extenso que resultaría abarcar cada una de las Áreas, pero si resulta favorable desarrollar la aplicación de esta técnica a todas las áreas con sus peculiaridades.

En lo respectivo a la simplicidad con la que se aborda el análisis de las series históricas, se debe a que uno de los objetivos de este trabajo es mostrar que través de herramientas de exploración sencillas es posible generar planes de estrategias en ascenso de la calidad del CAT. Así mismo se debe vencer en primera instancia la renuencia en la aplicación de modelos de análisis histórico en este segmento empresarial.

En completa disposición de cumplir con el objetivo del presente estudio refiriéndose a la incorporación de modelos cuantitativos y de simulación en la gestión del CAT para la toma de decisiones, no obstante lo elemental de las herramientas utilizadas se alcanzo un buen desempeño de la puesta en marcha de técnicas matemáticas para medir la efectividad del ejercicio diario en el manejo de llamadas, sugiriendo que las modificaciones experimentales sin afectar la operación contrastan positivamente a las modificaciones ejecutadas por puro conocimiento empírico.

Adicionalmente la falta de programas de cómputo especializados en el análisis de series de tiempo, obliga a valerse de las herramientas disponibles como la hoja de cálculo con la finalidad de aplicar los instrumentos de exploración de datos históricos con las limitantes naturales de los elementos que componen una hoja de cálculo. No obstante, por experiencia del analista del presente estudio que ha comprobado la efectividad práctica de tal herramienta con excelentes resultados y muy cercanos a los resultados que presenta un programa especializado en métodos cuantitativos.

Inicialmente se describe de forma breve y en términos generales el funcionamiento de un CAT dentro del Capítulo I, así como los conceptos básicos en la gestión del mismo. Dentro del Capítulo II que se refiere a la descripción en la aplicación de análisis estadístico necesario para efectuar la validación resultante al aplicar los modelos de Mínimos Cuadrados, Suavización Exponencial, Promedios Móviles y Simulación Monte Carlo.

En el Capitulo III se expone el manejo de los modelos de Series de Tiempo y las métricas para medir el error y la exactitud del pronóstico. El Capitulo IV comprende una exposición de la Modelación y la Simulación pretendiendo esclarecer que se trata de técnicas de uso práctico con amplias aplicaciones, sin embargo posee ciertas limitaciones naturales, como el número de variables adaptables de la realidad a la simulación.

El resultado de la aplicación de lo temas teóricos expuestos se concentra en el Capitulo V, presentando los resultados de salida del modelo de simulación y un comparativo de los efectos producidos tras la aplicación de modelos de series de tiempo.

Capítulo I

Funcionamiento de un Centro de Atención Telefónica

La atención personalizada que solía ofrecerse a los usuarios de los servicios financieros, de seguros, de correos, tramites gubernamentales, etc., ha dejado de ser una realidad para convertirse en una pesadilla por el tiempo que se pasa en fila esperando a ser atendido por un ejecutivo de servicio, esto se debe al constante crecimiento del número de clientes y usuarios de los diferentes servicios, que exigen agilidad por parte de la persona que los atiende, rapidez y lo mas importante calidad en el servicio. A su vez, la constante expansión de las empresas, la diversificación y especialización en los servicios y la complejidad tecnológica, tienen un efecto negativo en la atención de las demandas de los clientes y usuarios.

1.1 Funcionamiento de un Centro de Atención Telefónica

La alta competitividad en los mercados ha obligado a las empresas a incorporar nuevos métodos para integrar los avances tecnológicos mas recientes dirigidos hacia la excelencia en la Gestión Empresarial, enfocados en la atención del creciente número de clientes que exigen un servicio de mayor calidad en el menor tiempo que sea posible y por encima de todo que resulte eficiente, respondiendo con exactitud a las demandas de los usuarios para evitar realizar la misma operación más de una vez.

Por tal motivo las empresas han decidido buscar nuevos canales de comunicación para hacer mas eficiente el contacto con sus clientes enfocando sus estrategias para lograra la satisfacción y prolongada permanencia como cliente en la empresa, dando lugar a que la industria del Centro de Atención Telefónica (CAT) surja con un gran auge como herramienta de apoyo para el logro de las estrategias planteadas.

A través de un CAT se puede dar servicio no solo a clientes, sino también atención a colaboradores internos y proveedores dentro de la misma empresa, a quienes se les designara en adelante como clientes internos y externos respectivamente. El propósito principal del CAT

es lograr que la comunicación entre la empresa y sus clientes sea constante para gestionar de forma eficiente el negocio que representa la empresa, maximizando recursos, reduciendo costos e incrementando los beneficios.

Pero la duda surge al cuestionarse ¿cómo es que un CAT puede incrementar las ganancias de una empresa o mejor aun elevar los niveles de productividad en una empresa o negocio por el hecho de realizar o recibir una llamada?. Antes de responder a esta pregunta, lo primero es definir la función principal de un CAT que se resume en: hacer llamadas (outbound) o recibir llamadas (inbound) con el fin de satisfacer determinados aspectos de la relación con los clientes internos y externos, teniendo como objetivo prestar atención a través del aparato telefónico con altos niveles de calidad y eficiencia tanto como si la persona atendida estuviera presente. Las llamadas outbound pueden ser: promociones (telemarketing), cobranza y las llamadas inbound pueden ser: información, consultas, reclamaciones, quejas, entre otras.

1.1.1 Llamadas In Bound

Las llamadas de entrada o in bound se generan cuando los clientes llaman para solicitar asistencia, solicitar una orden o simplemente información de servicios o por respuesta directa de medios publicitarios, promociones, puntos de venta, etc; las llamadas recibidas en el CAT son atendidas por operadores de servicio que responden oportuna¹ y cortésmente haciendo uso de la tecnología disponible, que asegura la satisfacción total de los clientes.

Para ilustrar el proceso de las llamadas de entrada, se explica brevemente en términos técnicos de telefonía los pasos que se siguen para que una llamada de entrada se conecte al CAT. Se considera a los clientes externos que se comunican a la empresa por medio de un número de servicio gratuito 01 800 otorgado a la empresa y asignado por la Red Pública Telefónica Conmutada (PSTN)², que se enlaza frecuentemente al número de Identificación Automática de Números (ANI)³; el número marcado se etiqueta por el número del Servicio de

¹ El término respuesta oportuna se refiere al tiempo que tarda un operador en contestar la llamada

² Public Service Telephone Network

³ Automatic Number Identification

Identificación del Número Marcado (DNIS)⁴. El servicio PSTN utiliza el ANI y DNIS para conectarse con el centro.

El CAT de la empresa posee un Ruteador de Intercambio Privado (PBX)⁵, y el número identificado del cliente por el DNIS enlaza el PBX con el PSTN. Si el Centro tiene mas de una llamada la combinación del ANI y el DNIS se pueden utilizar para enrutar la llamada. El PBX se conecta al PSTN a través de un número de líneas telefónicas, llamadas líneas troncales, si estas troncales se encuentran libres entonces la llamada se conectará con el PBX, de lo contrario se experimentara un tono ocupado.

Las llamadas conectadas al CAT a través del PBX se enlazan con el Centro Interactivo de Voz (IVR)⁶ también conocido como contestador automático, que forma la cola de las llamadas según las necesidades con la opción de completar el servicio por medio del IVR, sin la necesidad de hablar con un agente. En el caso de querer comunicar una necesidad a un Representante de Servicio de Atención al Cliente (CRS)⁷, entonces la llamada se maneja por el Distribuidor Automático de Llamadas (ACD)⁸ que conecta la llamada según el criterio marcado desde el IVR, a su vez el ACD concentra las estadísticas de la operación.

En el caso de que un CRS no se encuentre disponible entonces el ACD puede “mantener” la llamada hasta ser atendida mientras se escucha música por el auricular, anuncios comerciales o alguna otra información, pero para los clientes la espera puede no valer la pena y colgar la llamada antes de ser atendido, lo que se refiere como abandono. Una vez que se conecta con el cliente, los operadores pueden atender la llamada mientras consultan la terminal de computadora con la información de la empresa que contiene los datos del cliente.

Los servicios de ACD e IVR son particularmente importantes en la operación de un CAT debido a que estos sistemas generan gran cantidad de información del flujo de las llamadas, tales como: cuándo se recibió; quién llamo; qué acciones tomo en el IVR y cuánto tiempo le llevo cada acción; si espero en cola y cuánto tiempo; si un CSR en atender la llamada y cuánto

⁴ Dialed Number Identification Service

⁵ Private Branch Exchange

⁶ Interactive Voice Response

⁷ Customer Service representative

⁸ Automatic Call Distributor

tiempo tardó; la clave del CSR que atendió la llamada, etc. Los datos generados por estos sistemas se almacenan para realizar el análisis histórico detallado.

Con el fin de ampliar el concepto del funcionamiento del ACD y del IVR, se introducen definiciones alternas y los beneficios por contar con esta tecnología para mejorar la capacidad de un CAT.

El ACD es un sistema telefónico especializado que puede manejar llamadas entrantes o realizar llamadas salientes, puede reconocer y responder una llamada entrante, buscar en la base de datos instrucciones sobre qué hacer con la llamada, reproducir locuciones, grabar respuestas del usuario y enviar la llamada a un operador, cuando haya uno libre o cuando termine la locución.

Beneficios del ACD:

- Mejora la imagen de la corporación
- Mejora el servicio al cliente
- Mejor manejo de las llamadas
- Mejor control del sistema

El IVR consiste en un conjunto de mensajes de voz y marcación de tonos desde un teléfono, de este modo se obtiene información del usuario llamante que en el destino sirve para la autenticación e identificación del mismo. También permite realizar transacciones totalmente automatizadas.

Servicios proporcionados por IVR:

- Información y transacciones financieras
- Información de eventos deportivos, recreativos y culturales
- Requisitos para trámites oficiales
- Contratación de películas en TV de paga
- Información de horarios y costos, etc.

Ventajas del IVR:

- Servicio las 24 hrs. Los 365 días del año

- Capacidad de realizar transacciones en forma automática
- Alta capacidad de almacenamiento de información
- Reducción del personal del centro Telefónico
- Anuncios y avisos de la compañía
- Acceso desde cualquier teléfono
- Optimizar el número de troncales
- Aumentar el número de llamadas

Desventajas del IVR

- Los clientes no están convencidos de hablar con una máquina
- Si no existe la información requerida no se ofrece un buen servicio
- Atención no personalizada
- La navegación es larga y tediosa
- Todavía hay teléfonos de disco

En lo que concierne al IVR con áreas in bound (recepción de llamadas), es que las llamadas que entran por unidad de tiempo fluctúan durante el día, esto dificulta la asignación de operadores por periodo de tiempo, por ejemplo, cada media hora.

Otra desventaja de las áreas in bound es el tiempo que a un operador le toma atender a un cliente durante la llamada y darle solución, si suponemos que el tiempo que le toma a los operadores es de larga duración se espera que el abandono de llamadas se incremente, por consiguiente baja el nivel de servicio y la cola de espera se satura, por ende se bloquean las llamadas entrantes, es decir, se experimenta un tono ocupado o la llamada se corta.

1.1.2 Llamadas Out Bound

Las llamadas de salida o out bound representan un mayor grado de dificultad que las llamadas in bound, ya que no se puede presuponer la atención del receptor. Un factor clave para tener éxito en una campaña out bound es la calidad de la información contenida en los registros de la base de datos.

Los servicios que se ofrecen en las campañas out bound proporcionan un lazo de contacto permanente entre el consumidor y la empresa, los centros de llamadas de salida pueden ayudar a aumentar lealtad de marca de fábrica y hacer clientes “de por vida” proponiendo varios programas de ventas en práctica que se pueden utilizar para los lanzamientos y las mejoras del nuevo producto, así como ventas de la garantía extendida. Este contacto continuo con los clientes fomenta estrechas relaciones, lealtad de marca de fábrica mejorada.

La constitución de las bases de datos para obtener los resultados esperados en cualquier tipo de campaña de salida juega un papel preponderante para asegurar que cada llamada genere un contacto exitoso con el cliente. Los servicios que se ofrecen en llamadas de salida en CAT son los siguientes:

Cobranza Preventiva: Recordar a su cliente el próximo vencimiento de pago.

Avisos de Cobranza: Una vez vencido el plazo de pago, se recuerda al cliente el saldo que se adeuda.

Telecobranza: Se establecen compromisos de pago con el cliente moroso, especificando día y sucursal de pago, así como, ejecutivo en sucursal a contactar.

Agendamiento de Entrevistas: Consiste en recorrer bases de datos ya sea proporcionadas por nuestro mandante o propias a efecto de concertar entrevistas para vendedores a terreno de nuestro cliente.

Llamadas de Bienvenida : Todos los nuevos clientes de nuestro mandante son llamados para darles una cordial bienvenida, detallarles sus derechos y deberes como clientes. Además se realiza una fiscalización a la gestión de venta de los vendedores.

Llamadas de Cumpleaños : En la fecha mencionada, los clientes son llamados para desearles un feliz aniversario de vida.

1.2 Principales Indicadores de un CAT

Para conocer el desempeño de un CAT se necesita una gran cantidad de información, por ejemplo, se necesitan datos sobre las necesidades y expectativas de los clientes, el tiempo de espera y la tolerancia del cliente, la carga al sistema, las actividades de los agentes y su rendimiento, el patrón de llamadas, los componentes de los gastos, las actividades en otras áreas de la empresa, y las condiciones en el entorno externo. Pero también se debe de ser capaz de ir más allá de los detalles y evaluar el rendimiento en general, sin la necesidad de revisar docenas de informes. Entonces, la pregunta es, ¿qué medidas pueden resumir adecuadamente las numerosas actividades del centro?, pero hay que ser cauteloso ya que cualquier medida puede llevar a una conclusión equivocada.

1.2.1 Medidores de Servicio

La medición del servicio se refiere a que tan bien se esta atendiendo a los clientes, que tan oportuna es la atención de los representantes del servicio a clientes para atender a las demandas de sus propios clientes, entre las medidas o métricas para establecer el patrón de servicio se encuentran las siguientes:

Porcentaje de nivel de servicio

El nivel de servicio se mide por el porcentaje de llamadas que se atendieron con éxito en el día y es una medida de alto nivel de cuán rápido los clientes se ponen en contacto con los agentes. Los mejores centros de llamadas toman esta medida muy en serio, y tratan de apegarse a la misma con la mayor consistencia posible.

Tasa de Abandono

El abandono de llamadas se da por parte de los clientes y representa un problema en los centros. Si los clientes cuelgan el teléfono antes de poder hablar con ellos, perdemos la oportunidad de atenderlos o de resolver su problema con el servicio. Sin embargo, el abandono de llamadas es difícil de predecir y es un indicador poco confiable para estimar el tiempo que tuvieron que esperar los clientes. En última instancia, nosotros podemos controlar cuan accesible somos — cuantos troncales tenemos, y cuantos agentes experimentados están disponibles. Pero resulta inevitable controlar la reacción del cliente, o la infinidad de circunstancias que influyen en su comportamiento.

Tiempo promedio de contestación

El tiempo que tarda un operador en contestar una llamada puede resultar en un cliente satisfecho por haber atendido su llamado oportunamente o en un cliente que se reserva el derecho de replica por el tiempo que tardaron en atender a su llamada sin razón aparente. Esta medida tiene consecuencias negativas tanto para la empresa como para los representantes del servicio al cliente, ya que un cliente al cortar una llamada por considerar que esperar no vale la pena, significa para la empresa una rellamada por parte de un cliente no atendido, por ende mayor carga de trabajo.

Tiempo promedio de llamada

El tiempo promedio de llamada incluye: el tiempo de conversación con el cliente y el trabajo posterior a la llamada. El tiempo de conversación va desde al saludo de entrada hasta la despedida de la llamada. El trabajo posterior es el que debe realizarse inmediatamente después de la llamada. Tiempos promedio de llamada altos indican la necesidad de intervención o retroalimentación por el supervisor, especialmente en la forma de usar el tiempo de trabajo posterior a la llamada. Para representar el verdadero tiempo de trabajo posterior, se sugiere seguir los siguientes criterios:

1. el trabajo está relacionado con una llamada entrante
2. se realiza después de la llamada y,
3. sigue a la llamada inmediatamente o muy poco después, es decir, el trabajo es un componente importante de la carga de llamadas aleatorias que se reciben.

1.2.2 Medidores de Calidad

La calidad es un conjunto de cualidades en una persona que la acercan a la excelencia y crecimiento profesional, significa identificar y satisfacer las expectativas de los clientes. La calidad en un CAT considera diversos factores de índole cualitativa, como la atención del CRS al cliente, eficiencia con la que fue atendido, rapidez en la solución propuesta por el operador, exactitud de la resolución dada, por lo que la medición de la calidad genera una serie de resultados subjetivos y cuestionables en su aplicación práctica.

Actualmente se aplica un esfuerzo adicional para perfeccionar las mediciones de calidad, para asegurar que los representantes estén no solo dedicando el tiempo necesario a la atención de

las llamadas, sino superando las expectativas de los clientes. En un CAT el estándar de calidad se mide por el desempeño del ejecutivo durante la llamada, desde el saludo, identificación del operador y nombre de la empresa a la que se está comunicando hasta la frase de cortesía con el que cierra la llamada, para ejemplificar este proceso se enumeran los aspectos básicos de una llamada de buena calidad:

1. Se logra cumplir y superar las expectativas del cliente
2. Todos los datos se ingresan correctamente en sistema
3. El operador brinda resultados y soluciones correctas
4. El cliente recibe información correcta y completa
5. No se transfiere a un cliente de un teléfono a otro
6. No se apura al cliente
7. No se deja en espera al cliente sin una explicación
8. El operador sondea y obtiene información necesaria
9. El cliente confía que fue una llamada eficaz
10. El operador cumple con las promesas que hizo al cliente
11. Seguimiento del proceso que el cliente solicitó
12. La orden puede interpretarse por todo el personal

Asegurar a los clientes buen trato y solución a sus necesidades, requiere que sus operadores y supervisores, así como áreas de calidad un trabajo conjunto dirigido a mejorar e implantar estrategias enfocadas a una mejora continua de la imagen en la empresa o negocio. Las herramientas para observar el comportamiento de los representantes consisten en el monitoreo en sombra o paralelo, esta actividad consiste en conectar un auricular a una de las entradas del aparato telefónico y escuchar (sin intervención) la llamada, observando el desempeño del operador durante la llamada. Esta misma actividad puede llevarse a cabo vía remoto desde otro aparato telefónico a través de otra línea, con la opción de poder grabar la llamada. En ambos casos la retroalimentación por el desempeño de la llamada, ya sea positiva o negativa, sirve para mejorar o mantener la calidad en el servicio.

En la mayoría de los centros de llamadas, el criterio cualitativo, el cual enfoca conocimiento de los productos y servicios, habilidades de interacción con el público y atención de llamadas, y la política de la empresa, continua su proceso de evolución, haciéndose cada vez más refinado y específico. La mayoría utiliza algún tipo de monitoreo para evaluar el desempeño individual y para identificar necesidades de adiestrar y entrenar.

1.2.3 Medidores de Eficiencia

Existe una estrecha relación entre las medidas para cuantificar la eficiencia y la medición del servicio, ya que en ambos casos el objetivo que se persigue es la completa satisfacción del cliente.

Las medidas que comúnmente se utilizan para medir la eficiencia en un CAT son:

Tiempo Promedio de llamada

A diferencia del tiempo promedio de llamada que considera el tiempo total de la llamada en los indicadores de servicio, se da mayor énfasis al tiempo invertido que el operador consume al incluir cualquier actividad relativa a la misma fuera del teléfono, es decir, consulta con el supervisor, con alguna área en particular, información en archivo o cualquier fuente externa. Este tiempo puede afectarse según el giro de la empresa o negocio o por las políticas impuestas para dar seguimiento a una solicitud e incluso por problemas de desempeño del sistema.

Llamadas por hora

Esta medida es útil para observar el desempeño de las llamadas que están siendo atendidas durante un día, así como observar el comportamiento histórico del flujo de las llamadas contestadas en horas pico en la operación. Existe la posibilidad de obtener esta medida por operador y llevar el record del desempeño de cada uno por separado, con implicaciones positivas...

Porcentaje de trabajo posterior

Esta medida se refiere al tiempo que un operador invierte en actividades administrativas fuera del teléfono, esta medida puede ser directamente afectada por el desempeño de los sistemas de cómputo y servidores o por la falta de conocimiento de las políticas operativas.

Porcentaje de no disponible

Esta medida está dirigida a cuantificar el tiempo que un operador permanece fuera de línea sin autorización del supervisor, es un indicador que se monitorea constantemente durante la operación con la ayuda del sistema adquirido por la empresa que muestra en tiempo real (con 2 segundos de desfase) quien se encuentra fuera de la operación sin justificación, esta puede ser una causa de abandono ya que no hay operadores suficientes para contestar las llamadas.

Las medidas cuantitativas acostumbradas y tomadas como un credo para calcular la productividad de mayor importancia, está desapareciendo. Con mayor frecuencia, está siendo reemplazada por medidas más enfocadas y específicas, como medidas cualitativas y de adherencia. Los agentes se pueden concentrar en estar disponibles, y en manejar cada transacción de acuerdo a las necesidades de cada una. Si son bien implementadas, las medidas cuantitativas y de adherencia pueden cultivar un entorno de trabajo mejor, mejorar la calidad y aumentar la productividad.

Se incluye en el Anexo una síntesis de la manera como se llevan a cabo los calculos para cada uno de los medidores antes mencionados, bajo el nombre de Métricas CAT.

1.3 Administración de un CAT

La creciente complejidad de las llamadas y la tecnología aplicada requiere más supervisión, adiestramiento e intercambio de información, los operadores reciben llamadas mas exigentes y requieren de elementos innovadores en capacitación para adquirir habilidades en el manejo de clientes difíciles. El asunto se complica aún más, cuando las tareas de los supervisores son tan variadas, por ejemplo, el tiempo que dedican al monitoreo y retroalimentación es muy variado. Además, cuando las llamadas en espera se acumulan, algunos supervisores se dedican a atender llamadas, mientras que otros están disponibles para ayudar a los agentes.

El intercambio de información es casi nulo, ya que no hay tiempo disponible para estas actividades. Y los supervisores que tienen grupos demasiado grandes pueden verse inundados con tareas que los alejan cada vez más del trabajo principal, que es el de atender llamadas y a los operadores con llamadas difíciles.

Ya que en los centros de llamada no están únicamente dedicados a resolver solo los problemas relacionados con las llamadas, calidad y servicio, las recomendaciones concernientes a dar solución a los problemas con el personal encargado de la supervisión tienen un rango amplio de acción, que va desde la capacitación en materia de supervisión hasta cursos de liderazgo, pero hay que tener en mente enfoques más estructurados para lograr ocupar el tiempo en labores de supervisión mas que desarrollar habilidades directivas.

En paralelo tenemos la dificultad de atender a los clientes y saturar su tolerancia con excusas que afectan la imagen de la empresa acerca de las deficiencias del servicio y la baja calidad con la que son atendidos, se calcula que existen siete factores que afectan la tolerancia del cliente. Estos factores influyen todo, desde cuánto tiempo el cliente esperará, hasta cuántos abandonarán la llamada, cuántos tratarán de nuevo cuando reciben un tono de ocupado, y como reaccionarán al contestador automático, tal como el mensaje de voz. También afectan la percepción del cliente en cuanto al servicio que el centro les brinda.

Pero no todo está perdido siempre parece que hay más problemas que soluciones, más llamadas que recursos, más quejas que soluciones. La verdad es que la administración en tiempo real necesita ser un complemento de un buen plan. Son las decisiones del momento las que le permiten mantener el nivel de servicio.

Las claves para mantener el nivel de servicio en tiempo real son las siguientes:

- Establecer una buena base de antemano, de manera de no crear Ud. mismo, muchas de las crisis a las que está respondiendo.
- Proveer información en tiempo real a los agentes y supervisores, y entrenarlos para que sepan interpretar esta información.
- Planear su procedimiento de escalación por anticipado, y establecer los límites que definen cuando se utilizan las alternativas.
- Establecer un equipo o una persona para coordinar las tácticas de tiempo real.
- Revisar y refinar su plan de escalación continuamente.
- Mejorar su proceso de planificación constantemente. La administración en tiempo real nunca será un buen sustituto para la planificación exacta de recursos.
- Recuerde, la administración eficaz de tiempo real requiere planificación, coordinación, y práctica.

En lo que respecta al nivel de calidad que debe manejarse dentro de la organización, se debe realizar un análisis para evaluar el perfil de la demanda del mercado, del cliente y sus necesidades. No es un asunto que deba llevarse a cabo esporádicamente, es un esfuerzo cotidiano que permita ir tomando pequeñas decisiones que incidan en la mejora del servicio y

la consecución de los objetivos planificados, en estricta relación con los representantes del servicio a clientes, que mantienen comunicación constante con los clientes y son ellos quienes mejor podrían dar cuenta de las necesidades e inquietudes de los clientes propios y los clientes potenciales, así como las tendencias en el mercado y características de la competencia.

Naturalmente, se presupone una evaluación en un sentido más amplio que terminará por implicar cambios estructurales y estratégicos, en respuesta a los cambios permanentes en los requerimientos del servicio.

Por último, la evaluación cuantitativa del desempeño en el CAT tiene que estar dirigida a la determinación de medidores que resulten de mayor utilidad, en vez de crear un sin fin de reportes que no permiten visualizar el entorno de la operación y que pueden crear falsas expectativas. Para elevar la calidad en la medición se requiere de:

- a. Un elevado grado de objetividad en la medición
- b. Homogeneidad
- c. Fiabilidad de la información resultante
- d. Segmentación adecuada
- e. Garantizar una muestra representativa y fiable

Además de conformar bases de datos confiables y contar con el personal capacitado para llevar a cabo el análisis e interpretación, con habilidades para determinar las tendencias que motivan el cambio por movilidad de los factores internos y ajenos a la operación. Pero hay que ser cuidadoso y evitar mal interpretar los resultados, que si bien están sustentados por modelos matemáticos, esto no quiere decir que se trate de una verdad absoluta

Capítulo II

Análisis de Datos Modelos Estáticos

Los modelos de regresión de una sola variable, simples en forma, pero de amplia aplicación en términos de la gran variedad de prácticas en cuestiones de negocios. En este tipo de modelos, la variable en estudio se expresa en términos de variables explicatorias, que representan el comportamiento de la realidad.

Los modelos de regresión son importantes, no solo por que son útiles para la aplicación de pruebas de hipótesis y pronóstico, sino por que forman parte del campo para el análisis de modelos de ecuaciones simultáneas y de series de tiempo.

2.1 Exploración de Datos

La estadística se ha definido como la ciencia que describe fenómenos numéricos y estudio de fenómenos aleatorios y la mayor contribución es la obtención de conclusiones basadas en la obtención de datos provenientes de la experimentación, Canavos [2]. Para comprender mejor la naturaleza de la estadística es necesario comprender las nociones de población y muestra, definidas de la manera siguiente por Canavos [2]: “La población es la colección de toda la posible información que caracteriza a un fenómeno. Por otro lado, la muestra es un subconjunto representativo seleccionado de la población” (p. 3). La información extraída del conjunto poblacional se analiza con el propósito de realizar deducciones acerca del comportamiento poblacional proyectando de lo específico a lo general. En este procedimiento nunca podrá tenerse una certeza del 100%, siempre existirá la posibilidad de error. Es importante hacer notar que los resultados derivados del análisis de la información, son estrictamente de carácter deductivo y las variaciones con respecto a la realidad pueden llegar a representar diferencias significativas. Sin embargo, también existe una medida de confiabilidad que se mide en términos de probabilidad, que indica que la inducción es significativa.

2.1.1 Análisis de frecuencias

Antes de comenzar a realizar cualquier operación matemática¹ es recomendable hacer una descripción gráfica del conjunto de datos para observar la frecuencia de repetición de las observaciones, así como identificar patrones en el conjunto de observaciones. Las observaciones se agrupan por clases, evitando que se superpongan entre si para evitar ambigüedades con respecto a la clase que pertenece un dato del conjunto. El número de observaciones en una clase recibe el nombre de frecuencia de clase y el cociente de una frecuencia de clase respecto a la combinación de observaciones se el conoce como frecuencia relativa. Las fronteras de clase se denominan límites y el promedio de los límites superior e inferior es el punto medio de la clase. A continuación se proporciona la definición de elementos básicos para la obtención de la distribución de frecuencias².

Límites de Clase. Son los valores mayor y menor que son factibles de encontrarse como datos y pertenecer a una clase. El límite superior difiere del límite inferior de la clase siguiente para evitar ambigüedades.

Fronteras de clase. Son los puntos medios entre límites de clases adyacentes, los límites de una clase estarán siempre contenidos entre las fronteras de la misma clase.

Marca de clase. Es el punto medio del intervalo de clase y su notación mas conocida es x_i , es importante dejar claro que debido a la construcción tanto de límites como fronteras, es indistinto calcular las marcas de clase a partir de los límites de clase o de las fronteras, sin embargo, se deben utilizar unos o los otros, pero nunca un límite o una frontera.

Frecuencia. Es el número de datos de la muestra que corresponde a la clase tratada, su notación es f_i y se determina realizando el conteo del número de observaciones en la muestra, dentro del rango determinado por lo límites o las fronteras.

1 Se hace esta recomendación debido a que en muchas Empresas, se comienza por obtener medidas de dispersión sin conocer con antelación la distribución que siguen los datos, por ende se obtienen resultados poco significativos

2 Obtenidas de: UNAM (Mayo 2001), [Propedéutico de Probabilidad y Estadística]. División de estudio de posgrado. México, D.F., Junio 2001

Frecuencia acumulada. Denotado por F_i , y es el número de datos en la muestra cuyo valor no debe exceder la frontera superior de la clase, es decir, el número total de las frecuencias observadas.

Frecuencia relativa. Se distingue por el símbolo f_i^* y es la proporción de los datos de la muestra que pertenece a una clase definida, se expresa como sigue:

$$f_i^* = f_i / n = f_i / \sum f_i$$

siendo n el número de datos de la muestra e i el número de la clase,

Frecuencia relativa acumulada. Es la proporción de datos en la muestra que no exceden la frontera de la clase, es decir, se asemeja a la función de distribución y la semejanza será aun mayor si la muestra es grande.

$$F_i^* = F_i / n = F_i / \sum f_i$$

Longitud de clase. Es la diferencia entre la frontera superior e inferior de la clase.

Rango de la muestra. Es la diferencia entre los valores mayor y menor de la muestra.

2.1.2 Descripción Gráfica de los Datos

Cuando se representa gráficamente la información el conjunto de datos se elige una forma gráfica, esta representación resulta útil para comprender su comportamiento. Existe una gran variedad de gráficas, sin embargo se utilizara alguna de ellas.

Histograma de frecuencias (Figura 1). Es una gráfica formada por barras rectangulares, su base se encuentra centrada en las marcas de clase de la distribución y sus áreas representan las frecuencias absolutas y relativas correspondientes.

Polígono de frecuencias. Es una gráfica poligonal que representa para cada marca de clase la frecuencia de dicha clase y se construye uniendo mediante líneas rectas los puntos medios superiores de las barras del histograma de frecuencias.

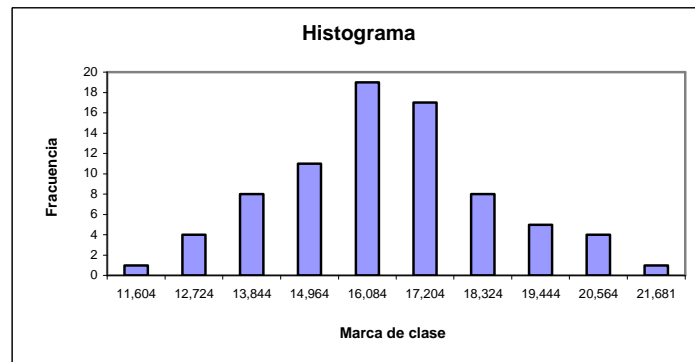


Figura 1. Histograma de frecuencias

Ojiva. Es una gráfica poligonal que representa la frecuencia acumulada de cada frontera de la clase, cuando se representa la frecuencia relativa acumulada se le llama ojiva porcentual. La ojiva porcentual es una aproximación a la gráfica de la función de distribución de la variable aleatoria que representa la población³.

2.2 Funciones de Distribución

2.2.1 Distribuciones Continuas

Distribución Normal

La distribución Normal o Gaussiana es la mas importante y de mayor utilidad de las distribuciones continuas de probabilidad y la que mas se utiliza. La apariencia gráfica de la distribución tiene forma de campana y se extiende sin límite hacia en ambos sentidos; positivo y negativo. Se puede aplicar en el estudio de fenómenos aleatorios de la disciplinas como la ingeniería, negocios y economía.

³ Ibid., p.15

Definición Canavos[2], por DeMoivre (1733) función de densidad de probabilidad de la distribución Normal. Se dice que una variable aleatoria X se distribuye normalmente si su función de densidad de probabilidad esta dada por:

$$f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp - \left[\frac{1}{2} (x - \mu / \sigma)^2 \right]$$

Los parámetros de la distribución normal son μ y σ , estos parámetros son la media y la desviación estándar de X .

Cuando la media de la distribución es 0 y la varianza es 1 se denomina "normal tipificada" o "normal estandarizada", y su ventaja reside en que hay tablas donde se recoge la probabilidad acumulada para cada punto de la curva de esta distribución dado que existe un número infinito de valores para μ y σ , simplificandose mediante el empleo de la siguiente transformación: sea Z una v.a. definida como:

$$Z = (X - \mu) / \sigma$$

Así, para $z = (x - \mu) / \sigma$, $P(X \leq x) = P(Z \leq z)$ y

$$F_x(x; \mu, \sigma) = F_z(z; 0, 1)$$

Donde $F_z(z; 0, 1)$ es la función de distribución acumulativa de la función de probabilidad normal estandarizada.

Se expone un ejemplo de Canavos [2], para ilustrar su funcionalidad. Sea X una v.a. que representa la inteligencia medida por medio de pruebas IQ. Si X es $N(100,10)$, obtener las probabilidades de que X sea mayor que 100, menor que 85, a los mas 112. por lo menos 108, mas grande que 90 y entre 95 y 120, (p. 137).

Solución. Las probabilidades se obtienen de la siguiente forma:

$$P(X < 85) = P(Z < 85 - 100 / 10) = P(Z < -1.5)$$

$$= F_z(-1.5; 0, 1) = 0.0668$$

$$P(X \leq 112) = P(Z \leq 1.2) = F_z(1.2; 0, 1) = 0.8849$$

$$P(X \geq 108) = P(Z \leq 0.8) = 1 - F_z(0.8; 0, 1) = 0.2119$$

$$P(X > 90) = P(Z > -1) = 1 - F_z(-1; 0, 1) = 0.8413$$

$$P(95 \leq X \leq 120) = P(-0.5 \leq Z \leq 2) = F_z(2; 0, 1) - F_z(-0.5; 0, 1) = 0.6687$$

Distribución F

Si Z_1 y Z_2 son dos variables Chi cuadrado⁴ independientemente distribuidas con k_1 y k_2 grados de libertad respectivamente, la variable:

$$F = Z_1/k_1 / Z_2/k_2$$

Sigue la distribución F (de Fisher)⁵ con k_1 y k_2 grados de libertad. Una variable distribuida como F se denota como F_{k_1, k_2} , donde los subíndices indican los grados de libertad asociados con las variables Z.

Las propiedades de la distribución, son:

- Esta distribución esta sesgada hacia la derecha. Pero puede demostrarse que a medida que k_1 y k_2 aumentan, la distribución F se aproxima a la normal.
- El valor medio de una variable distribuida como F es $k_2/(k_2 - 2)$ definido para $k_2 > 4$ y su varianza es:

$$\frac{2k_2^2(k_1 + k_2 - 2)}{k_1(k_2 - 2)^2(k_2 - 4)}$$

que se defina para $k_2 > 4$.

- El cuadrado de una variable aleatoria distribuida como t con k grados de libertad sigue una distribución F con 1 y k grados de libertad; simbólicamente,

$$t_k^2 = F_{1, k}$$

4 Distribución chi cuadrado. Sean las variables estandarizadas normales e independientes (con media cero y varianza unitaria), entonces la cantidad $Z = \sum Z_i$ se dice que tiene una distribución chi cuadrada con k grados de libertad, donde grados de libertad significa el número de cantidades independientes de la suma total.

5 Gujarati [6] 1981, p. 398

Distribución t

Si Z_1 es una variable normal estandarizada y otra variable Z_2 sigue la distribución chi cuadrado con k grados de libertad además esta independientemente distribuida de Z_1 , entonces la variable definida como:

$$T = Z_1 \sqrt{k} / \sqrt{Z_2}$$

Sigue la distribución t de Student⁶ con k grados de libertad. Una variable distribuida como t se suele designar como t_k .

Propiedades de la distribución t

- La forma geométrica de la distribución t es parecida a la forma de la distribución Normal y a medida que aumentan los grados de libertad la distribución t se aproxima a la normal.
- La medida de la distribución t es cero y su varianza es $k/(k-2)$.

2.3 Análisis de Regresión

2.3.1 Conceptos Básicos

Regresión

Al analizar una serie de datos una cuestión importante es estudiar la probable asociación cuantitativa entre el número de variables en estudio, a lo que se le conoce en estadística como análisis de regresión. El término regresión se utilizó por primera vez por Francis Galton, en un artículo publicado en Londres en 1886. Galton observó que a pesar de que existe una tendencia de los padres a tener hijos altos y los padres de baja estatura a tener hijos bajitos (Gujarati [6], 1981, p. 3), la distribución de las estaturas de una población no cambia sustancialmente de generación a generación. Explica que había una tendencia de la estatura

⁶ Gujarati [6] 1981, p. 397

promedio de los niños con padres de determinada estatura a moverse o **regresar** a la estatura promedio de toda la población.

“El análisis de regresión trata de la dependencia de una variable, variable dependiente, en una o mas variables, variables explicatorias, con el objeto de estimar o predecir la media o valor promedio (pobacional) de la primera con base en los valores conocidos o fijados (en muestras repetidas) de las segundas”. (Gujarati [6], 1981, p. 4) Es decir, obtener una ecuación empírica de predicción que proporcione un modelo teórico que no esta disponible, es decir una función matemática sencilla o un polinomio que describa, el comportamiento de la variable dependiente o respuesta, dados los valores de las variables independientes o de predicción. De lo que resulta obvio, la imposibilidad de predecir un valor particular o exacto en el contexto real. Sin embargo es posible predecir un valor promedio para el conjunto de datos en estudio.

Un procedimiento útil cuando se tiene una sola variable de predicción es graficar la variable respuesta contra la variable de predicción. Si este refleja una tendencia lineal⁷ podrá suponerse un modelo lineal general con la siguiente ecuación:

$$Y = a + b X$$

dónde Y es la variable dependiente de la que se pretende predecir los movimientos y X la variable independiente. Para obtener los valores de a y b de la ecuación la técnica mas utilizada para este propósito es el método de mínimos cuadrados (MC), que consiste en encontrar las estimaciones para los parámetros de la ecuación seleccionada mediante la minimización de la suma de los cuadrados de las diferencias entre los valores observados de la variable respuesta y de aquellos proporcionados por la ecuación de regresión⁸.

Método de Mínimos Cuadrados Ordinario

El criterio en la técnica de MC puede explicarse formalmente como:

$$\text{Minimizar } \sum_{i=1}^N (Y_i - \hat{Y}_i)^2$$

⁷ No es condicional en la aplicación del modelo de Mínimos Cuadrados

⁸ Pindyck, Rubinfeld [7], 1981, p. 8

dónde $\hat{Y}_i = \alpha + \beta X_i$ representa la ecuación del modelo lineal estimado con intersección⁹ α y pendiente β . En esta notación Y_i representa el valor de Y para cada observación i y que corresponde al valor de X para esa observación, mientras que N es el número de observaciones. \hat{Y}_i es el valor ajustado o de predicción de Y_i , es el valor de Y sobre la línea recta asociado con la observación X_i . Las desviaciones se calculan restando cada valor ajustado a cada valor de las observaciones y como resultado de la suma de las desviaciones, cero. Los valores de a y b se calculan a partir de las ecuaciones:

$$\beta = \frac{\sum (X_i - \mu_x) (Y_i - \mu_y)}{\sum (X_i - \mu_x)^2}$$

$$\alpha = \mu_y - \beta \mu_x$$

donde μ_x y μ_y son las medias muestrales de Y y X respectivamente.

$$\hat{Y}_i = \alpha + \beta X_i$$

es la ecuación de la línea de regresión de mínimos cuadrados muestral.

Supuestos del Método de Mínimos Cuadrados Ordinario

Supuesto de Normalidad

El modelo clásico de regresión lineal supone que cada desviación e_i está distribuido normalmente $e_i \sim N(0, \sigma^2)$ con¹⁰:

Media: $E(e_i) = 0$

Varianza: $E(e_i^2) = \sigma^2$

Covarianza: $E(e_i e_j) = 0 \quad i \neq j$

⁹ si $\alpha=0$, entonces la intersección se encuentra en el origen

¹⁰ e_i es una variable aleatoria, no observable, que puede tomar valores positivos o negativos, mejor conocida como perturbación estocástica o término de error estadístico

Cabe mencionar que para dos variables normalmente distribuidas, la covarianza igual a cero significa independencia de las variables y no correlación¹¹.

Supuesto 1

El valor medio de las desviaciones o perturbaciones estocásticas (e_i) es igual con cero

$$E(e_i | X_i) = 0$$

Supuesto 2

Se dice que las perturbaciones no están correlacionadas, mejor conocido como supuesto de no autocorrelación, que significa que dada cada observación X_i las desviaciones de 2 cualesquiera valores calculados \hat{Y}_i de su media no presentan patrones.

$$\text{Cov}(e_i, e_j) = E[e_i - E(e_i)] E[e_j - E(e_j)] = E(e_i e_j) = 0$$

Supuesto 3

La varianza de e_i para cada X_i es un número positivo o constante, igual a la σ^2 , se conoce mejor como supuesto de homocedasticidad o igual varianza. En el caso de que

$$\text{Var}(e_i | X_i) = E[e_i - E(e_i)]^2 = E(e_i)^2 = 0$$

la varianza sufriera un cambio conforme aumenta el número de observaciones se le conoce como supuesto de heterocedasticidad o varianza desigual¹².

$$\text{Var}(e_i | X_i) = \sigma_i^2$$

11 Gujarati [6] 1981, p. 65

2.3.2 Validar el modelo

Bondad de Ajuste

Para determinar que tan bien se ajusta la línea de regresión a los datos se utiliza el indicador R^2 o prueba de bondad de ajuste, para calcularlo se requiere de tres conceptos la suma de los cuadrados totales SCT, la suma de los errores al cuadrado SEC y la suma de los cuadrados de la regresión SCR. SCT esta dado por $\sum (Y_i - \mu_y)^2$ que es la variación de las Y_i 's de la μ_y . SEC esta dado por $\sum (Y_i - \hat{Y}_i)^2 = \sum e_t^2$, si la línea de regresión se ajusta a todos los datos entonces $SEC = 0$. $SCR = \sum (\hat{Y}_i - \mu_y)^2$.

Tenemos entonces que

$$SCT = SCR + SCE$$

para un buen ajuste SCE sería pequeño, se define entonces a R^2 como

$$R^2 = SCR / SCT = \text{fracción de variación en Y explicada por X}^{13}$$

Medidas de confiabilidad

En virtud de que los datos pueden cambiar fácilmente de una muestra a otra y que por lo tanto los estimadores experimentaran cambios, se requiere de una medida de confiabilidad¹⁴ para comprobar que los estimadores de la línea de regresión α y β sean los más precisos, en estadística existe una medida de precisión para los estimadores llamada error estándar (es) y se obtiene de la manera siguiente:

$$\begin{aligned} \text{Var}(\beta) &= \sigma^2 / (\sum x_i^2) \\ \text{es}(\beta) &= \sigma / (\sqrt{\sum x_i^2}) \end{aligned}$$

12 Ibid, p. 8

13 Winston, Albright, 1997, p. 758

14 En estadística la confiabilidad de un estimador puntual se mide por su error estándar o por su varianza. En lugar de confiar en la estimación puntual, se verá cual es la probabilidad de que un estimador puntual se encuentre dentro de cierto rango o intervalo alrededor del verdadero parámetro.

$$\text{Var}(\alpha) = (\sum x_i^2 / N \sum x_i^2) (\sigma^2)$$

$$\text{es}(\alpha) = (\sqrt{\sum x_i^2 / N \sum x_i^2}) (\sigma)$$

donde x_i^2 es el cuadrado de las desviaciones.

Pruebas de significancia

Una prueba de significancia es un procedimiento para verificar la veracidad o falsedad de una hipótesis nula H_0 ¹⁵, la idea principal consiste en contrastar un estadístico de prueba versus la distribución muestral del estadístico bajo la hipótesis nula H_0 . Quizás la prueba más común es la prueba de la t para la pendiente, se lleva a cabo la prueba contrastando $H_0: \beta = 0$ versus $H_a: \beta \neq 0$

Utilizar la distribución t para construir un intervalo de confianza con 95% de confianza para los parámetros, primero se estandariza el parámetro de la regresión que sigue la distribución t con $N - 2$ grados de libertad (gl.)

$$t_{N-2} = \frac{\hat{\beta} - \beta}{s_{\hat{\beta}}} = \frac{\hat{\beta} - \beta}{s / \sqrt{\sum x_i^2}}$$

Como el estadístico de prueba sigue la distribución t es posible construir intervalos de confianza¹⁶, seleccionando un valor apropiado, llamado valor crítico denotado por t_c ,

$$\text{Pr}(t_c < t_{N-2} < t_c)$$

15 La hipótesis nula también se le conoce como hipótesis propuesta o hipótesis sometida a análisis y suele probarse contra la hipótesis alterna que se designa como H_a . La teoría de las pruebas de hipótesis se enfoca al diseño de pruebas o procedimientos que permitan decidir cuando aceptar o rechazar la hipótesis nula. En general una hipótesis nula es una declaración o hipótesis que considera ciertos efectos que no están presentes en la muestra.

16 El enfoque del intervalo de confianza en la realización de pruebas de hipótesis consiste, en primer lugar, en encontrar el intervalo de confianza apropiado y luego verificar si el valor de la hipótesis nula cae dentro o fuera de él.

El valor crítico varía según el número de gl. El valor crítico del estadístico t será central en nuestra decisión si rechazar o no la hipótesis nula. El intervalo de confianza para el verdadero valor del coeficiente para la pendiente es:

$$\Pr(\beta - t_c s_\beta < \beta_0 < \beta + t_c s_\beta) = .95$$

Y el valor para el intercepto α es:

$$\Pr(\alpha - t_c s_\alpha < \alpha_0 < \alpha + t_c s_\alpha) = .95$$

Es posible determinar intervalos de confianza para cada nivel de significancia mientras el valor crítico de la distribución t sea elegido correctamente¹⁷.

¹⁷ Pindyck, Rubinfeld [7], 1981, p. 57

Capítulo III

Análisis de Datos

Modelos de Series de Tiempo

Hoy en día la variable tiempo toma un nuevo significado en la vida de muchos profesionistas, quienes en un esfuerzo por predecir el futuro dedican largas jornadas de trabajo en el análisis de series de datos observados en diferentes momentos del tiempo (mejor conocido como series de tiempo), con la finalidad de encontrar un patrón de comportamiento que otorgue un ligero indicio de lo que pasara el día de mañana. Aunque existe cuantiosa información disponible para la investigación de series de tiempo, desafortunadamente la mayor parte de esta información es de naturaleza no experimental. Por consiguiente, existe una posibilidad de incurrir en errores de observación, ya sea por acción u omisión. No obstante, mas que un desaliento representa una oportunidad para mejorar los modelos de análisis que existen actualmente.

3.1 Conceptos preliminares de las Series de Tiempo

3.1.1 Procesos Estocásticos

Un proceso estocástico o aleatorio es una colección de variables aleatorias ordenadas en el tiempo, denotándose para una variable continua $Y(t)$ y para una variable discreta Y_t . De la misma forma que se utilizan datos muestrales para hacer inferencias respecto a la población, la serie de tiempo se emplea para llevar a cabo inferencias respecto al proceso estocástico subyacente. (Gujarati [6a], 2004, p. 771)

Proceso Estocástico Estacionario

Se dice que es un Proceso Estocástico Estacionario si la media y la varianza son constantes en el tiempo y si el valor de la covarianza entre los dos periodos depende solamente de la distancia o rezago entre estos dos periodos de tiempo y no en el tiempo en el cual se ha

calculado la covarianza (Gujarati [6a], 2004, p. 772), para explicar lo antes dicho; sea Y_t una serie de tiempo con la siguiente propiedad:

Media: $E(Y_t) = \mu$

Varianza: $\text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$

Covarianza: $\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$

donde γ_k , la covarianza (o autocovarianza) al rezago k, es la covarianza entre los dos valores Y_t y Y_{t+k} , es decir, separados por k periodos.

Si una serie de tiempo es estacionaria su media, varianza y covarianza permanecen constantes sin importar el momento en el cual se midan, es decir permanecen invariantes respecto al tiempo.

Una serie de tiempo no estacionaria tendrá una media que varía en el tiempo o una varianza que cambia con el tiempo o ambas. Si la serie no es estacionaria su comportamiento solo puede estudiarse bajo el periodo considerado o particular, como consecuencia no podrá generalizarse a otros periodos.

3.1.2 Pruebas de estacionariedad

Prueba gráfica

El método gráfico proporciona una clave inicial respecto a la naturaleza de la serie de tiempo, puede mostrar una tendencia creciente con lo cual se sugiere que esta variando, posiblemente resulte un indicio que la serie es estacionaria.

Función de Autocorrelación (FAC) y Correlograma

La función de autocorrelación FAC al rezago k, denotada por ρ_k , definida como:

$$\rho_k = \gamma_k / \gamma_0$$

es decir, covarianza al rezago k entre la varianza

El valor de ρ_k se encuentra entre -1 y 1 al igual que cualquier coeficiente de correlación, si se gráfica ρ_k frente a k, la gráfica se le conoce como correlograma poblacional.

Dado que en la práctica solo se cuenta con un extracto de la población solo se podrá calcular la función de correlación muestral ρ_k teniendo primero que calcular primero la covarianza muestral al rezago k γ_k y la varianza muestral γ_0 definidas como:

$$\gamma_k = \frac{\sum (Y_t - \mu_y) (Y_{t+k} - \mu_y)}{n}$$

$$\gamma_0 = \frac{\sum (Y_t - \mu_y)^2}{n}$$

donde n es el tamaño de la muestra y μ_y es la media muestral.

Por lo tanto la función de autocorrelación muestral al rezago k es¹:

$$\rho_k = \gamma_k / \gamma_0$$

Consideraciones de los coeficientes de autocorrelación

Su importancia radica en el error estándar (ee), Bartlett demostró que si una serie de tiempo es puramente aleatoria es decir ruido blanco², los coeficientes de autocorrelación muestrales son aproximadamente:

$$\rho_k \sim N(0, 1/n)$$

que quiere decir que en muestras grandes los coeficientes de autocorrelación están normalmente distribuidos con media cero y varianza uno sobre el tamaño de la muestra, entonces siguiendo los pasos de la distribución Normal estándar con un intervalo de confianza del 95% para cualquier ρ_k es:

¹ Gujarati [6a], 2004, p. 782

² Si se tiene un proceso estocástico de media cero, varianza constante, donde los diferentes términos no presentan correlación entre sí, esta serie se le llama ruido blanco. La importancia radica en que el ruido blanco al no depender de su pasado las nuevas observaciones son solo innovaciones debidas puramente al azar

$$\rho_k \pm 1.96 (ee)$$

es decir, $\Pr [\rho_k - 1.96 (ee) < \rho_k < \rho_k + 1.96(ee)]$ en términos de pruebas de hipótesis, tenemos la hipótesis nula $H_0 = 0$ y la hipótesis alternativa $H_a \neq 0$, en otras palabras si el intervalo incluye el valor cero no se rechaza H_0 , o sea que $\rho_k \approx 0$, pero si el intervalo incluye el valor cero entonces rechazamos H_0 es decir $\rho_k \neq 0$.

Para evitar realizar pruebas de significancia de cualquier coeficiente de autocorrelación individualmente, la tarea se facilita mucho si se aplica una prueba de hipótesis conjunta para todos los coeficientes de autocorrelación simultáneamente y probar que son iguales a cero, aplicando la prueba Q desarrollada por Box y Pierce definida como:

$$Q = n \sum \rho_k^2$$

donde n es el tamaño de la muestra y m la longitud del rezago, el estadístico Q se prueba versus la distribución Ji cuadrada, es decir, si la Q calculada excede el valor crítico de la tabla Ji cuadrada el nivel de significancia seleccionado se rechaza $H_0 = 0$ de que todos los ρ_k son iguales con cero, no obstante algunos de ellos serán cero³.

3.2 Modelos Lineales de Series de Tiempo

3.2.1 Promedios Móviles

Sea El modelo Y

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1}$$

donde μ es una constante y u , el término de error estocástico con ruido blanco. En este modelo Y en el periodo t es igual a una constante mas un promedio móvil de los términos de errores presente y pasado. Así, en el tiempo presente se dice que Y sigue un proceso de media móvil de primer orden o MA(1)

³ Gujarati [6a], 2004, p. 787

pero si Y se expresa:

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2}$$

es un proceso MA(2) y en general;

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q}$$

Es un proceso MA(q), un proceso de media móvil es simplemente una combinación lineal de términos de error con ruido blanco "hacia atrás". En este modelo se suponen los errores u_t distribuidos independientemente a lo largo del tiempo y distribuidos normalmente con media 0 y varianza σ^2 y la covarianza igual a cero. En el proceso MA(1) se tiene una covarianza de cero cuando el desplazamiento es mayor a un periodo, es decir, que el proceso tiene memoria de un solo periodo y ningún valor de Y_t es correlacionado con Y_{t-1} y Y_{t+1} , el proceso olvida los eventos del pasado cuando toman lugar mas de un periodo hacia atras, los eventos ocurridos mas de un periodo en el pasado no tienen efecto en los eventos presentes.

Se sugiere que el modelo de promedios móviles provee de información al pronóstico únicamente en cierto número de eventos en el futuro.

La correlación se determina por la función para k periodos

$$\rho = \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}$$

que prueba que el proceso de promedios móviles de orden q tiene memoria de exactamente q periodos⁴

⁴ Pindyck, Rubinfeld, 1981, p. 515

3.2.2 Suavización Exponencial Simple

Al igual que los promedios móviles se utiliza para pronósticos a corto y mediano plazo. Un problema con los promedios móviles es que se vale únicamente de valores pasados y actuales de la serie Y_t para obtener el valor de la nueva serie \tilde{Y}_t lo cual puede remedirse al utilizar un promedio móvil centrado, es decir:

$$\tilde{y}_t = 1/5(y_{t-2} + y_{t-1} + y_t + y_{t+1} + y_{t+2})$$

La suavización exponencial⁵ involucra el uso del promedio exponencial ponderado para suavizar. La suavización exponencial para la serie \tilde{y}_t esta dado por:

$$\tilde{y}_t = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots$$

y la fórmula recursiva para el calculo de la serie \tilde{y}_t :

$$\tilde{y}_t = \alpha y_t + (1 - \alpha)y_{t-1}$$

el valor de α se mantiene entre los valores 0 y 1.

Suavización Exponencial Doble

Si la tendencia y la media varían lentamente en el tiempo, se recomienda utilizar un modelo de suavizamiento exponencial de orden mayor, como el modelo de suavización exponencial doble, también conocido como modelo de suavización exponencial de Brown, que consiste en aplicar por segunda ocasión el alisado a la serie a la que se aplico el modelo simple, es decir:

calculo para el suavizamiento simple,

$$\tilde{y}_t = \alpha y_t + (1 - \alpha)y_{t-1}$$

aplicando nuevamente el alisado, tenemos \tilde{y}'

⁵ Pindyck, Rubinfeld, 1981, p. 484

$$\tilde{y}'_t = \alpha \tilde{y}_t + (1 - \alpha) \tilde{y}_{t-1}$$

la fórmula recursiva para calcular el doble alisado exponencial⁶.

3.3 Pronóstico con Modelos de Series de Tiempo

3.3.1 Consideraciones generales

Pronosticar valores futuros de una serie de tiempo observada es un problema importante en muchas áreas, incluyendo la economía, planeación de la producción, pronóstico de ventas y control de almacenamiento.

Suponga que se tiene una serie de tiempo observada, x_1, \dots, x_n , el problema entonces es estimar x_{n+1} o en general x_{n+q} . La predicción de x_{n+q} realizada en el tiempo n , q periodos adelante denotado por $x(n, q)$. El entero q es llamado periodo o plazo futuro. No hay un proceso universal para llevar a cabo un pronóstico, en su lugar debe buscarse el procedimiento mas apropiado para un conjunto de condiciones específico⁷.

Los procedimientos de pronóstico se pueden clasificar en tres categoría generales:

- Subjetivos. El pronóstico se puede realizar sobre una base utilizando la intuición, conocimiento del mercado y cualquier otra información relevante.
- Univariados. El pronóstico se puede basar por completo en una serie de observaciones históricas, ajustando un modelo a la serie y extrapolando. Estos métodos también se conocen como “naive” o “proyección”.
- Multivariados. El pronóstico se puede realizar tomando en cuenta las observaciones y variables participantes, por ejemplo, las ventas dependen del inventario. Los modelos de regresión entran en esta clasificación, conocidos como modelos econométricos. La

⁶ Ibid, p. 29

⁷ Evans, Olson [3], p. 335 - 337

utilización de indicadores de plazos futuros pertenecen también a esta categoría y son llamados en ocasiones “causales” o “predictivos”⁸.

En la práctica el procedimiento de pronóstico involucra en muchas ocasiones una combinación de las categorías antes mencionadas.

Antes de elegir un procedimiento para pronosticar, es esencial considerar el objetivo del pronóstico, la exactitud requerida, costo por desarrollar un modelo, variables a considerar, disponibilidad de información y periodos a pronosticar. El proceso de pronóstico depende en cierta medida del plazo futuro, ya sea de corto plazo (menos de nueve meses) o largo plazo. Entre mayor sea el plazo futuro, es muy probable que el modelo sufra cambios y el pronóstico no sea representativo o confiable. En muchos casos el proceso de pronóstico reproduce un pronóstico puntual, sin embargo no indica la incertidumbre asociada con la estimación de la demanda futura. Por esta razón es deseable producir un intervalo pronosticado.

3.3.2 Técnicas de Pronóstico

Promedios móviles

Este modelo se ajusta mejor a datos con mucha volatilidad o datos sin tendencia o estacionalidad. El pronóstico para el periodo siguiente se calcula como el promedio de las k observaciones mas recientes:

$$Y_{t+1} = (A_t + A_{t-1} + \dots + A_{t+1-k}) / k$$

El valor de k se puede elegir arbitrariamente, aun su elección afecta la exactitud del pronóstico. Entre mayor sea el valor de k el pronostico depende en mayor medida de la ultima observación; entre menor sea el valor de k , mayor sensibilidad del valor del pronostico por cambios en la serie de tiempo. El mejor valor de k se puede identificar analizando los errores de pronóstico. En otras palabras, el promedio de lo que ha ocurrido en el pasado se empleara par pronosticar el futuro, si Y_t es el promedio móvil calculado en l periodo t , el pronóstico de valor en el periodo $t+1$ se toma como Y_t . Con un promedio móvil, únicamente se utilizan las N observaciones.

⁸ Apuntes de Econometría, 27/06/2002

Cada vez que se dispone de una nueva observación, el promedio se desplaza para incluir la observación mas reciente y dejar rezagada la mas vieja de las observaciones, como consecuencia la perdida de datos mientras mas se amplia el rango de los datos, es decir, el valor de k^9 .

Suavización Exponencial Simple

El modelo de suavización exponencial simple pondera las observaciones pasadas con ponderaciones exponencialmente decrecientes; si el dato es muy reciente, la ponderación será mayor. El modelo básico se expone a continuación:

$$Y_{t+1} = Y_t + \alpha(A_t - Y_t)$$

Donde Y_{t+1} es el pronostico para el periodo $t+1$, Y_t es el pronostico para el periodo t y A_t es el valor observado en el periodo t , α es la constante entre 0 y 1 llamada constante de suavizamiento.

Al igual que en los promedios móviles, el promedio calculado hasta el periodo t se utiliza para pronosticar el periodo $t+1$. Al realizar tal pronóstico el requerimiento primordial es la constante de suavizamiento, la observación más reciente y el pronóstico anterior, sin la necesidad de almacenar las observaciones previas. En este modelo no se desprecia en absoluto ninguna de las observaciones, con regularidad queda incorporado el pronóstico anterior.

En lo que respecta al valor de la constante de suavizamiento se puede ajustar de tal forma para que el pronóstico responda rápida o lentamente a los cambios en el sistema. Por ejemplo, si $\alpha = 0$ simplemente se repetirá el pronostico del periodo anterior; $\alpha = 1$ se pronosticara el ultimo periodo de la demanda actual. Entre mas cerca se encuentre α del valor 1, el modelo responderá mas rápidamente a cambios en la serie de tiempo por que se pone mayor peso en la observación actual que en el pronostico. Y como en el modelo de promedios móviles, esta técnica se ajusta mejor a series con datos volátiles sin tendencia ni estacionalidad

⁹ DUKE UNIVERSITY: "Introduction to forecasting. Forecasting". Decision 411 Course Outline 2005, [en línea]. Recuperado el 15 de noviembre del 2006. <<http://www.duke.edu/~rnau/411home.htm...>>

Suavización Exponencial Doble

La fórmula para el pronóstico se basa en la extrapolación de una línea a través de 2 centros, es decir:

$$Y_{t+1} = a(t) + b(t)$$

donde

$$a(t) = 2\tilde{y}_t - \tilde{y}'_t$$

$$b(t) = (\alpha/(1-\alpha))(\tilde{y}_t - \tilde{y}'_t)$$

cabe aclarar que para la utilización de este modelo de pronóstico, se debe utilizar el mismo parámetro α , para el alisado simple y doble¹⁰.

3.3.3 Métricas de Error y Exactitud en el pronóstico

La calidad del pronóstico depende en la exactitud para predecir valores futuros de una serie de tiempo. El término error de pronóstico se define como la diferencia entre el pronóstico y el valor actual de la serie de tiempo que se conoce. Para el modelo de promedios móviles, para los diferentes valores de k se generaran diferentes pronósticos, como para diferentes valores de α para el modelo de suavización exponencial.

Para analizar la exactitud de estos modelos con mayor precisión, se definen las métricas de error¹¹, para realizar comparaciones cuantitativas con la serie de datos observada. Las tres métricas más utilizadas son la media absoluta de desviación (MAD), raíz del error cuadrado medio (RMSE) y porcentaje de error de la media absoluta (MAPE). La MAD es el promedio de las diferencias entre el valor actual y el pronóstico, ponderado sobre el rango de valores pronosticados:

$$MAD = \sum_{t=1}^n |A_t - F_t| / n$$

¹⁰ Ibid, p. 32

¹¹ Véase, Evans, Olson [3], p. 338–339

donde A_t es el Valor Actual de la serie de tiempo en el tiempo t y F_t es el valor del pronóstico en el tiempo t , n es el número del valor del pronóstico. Esta medida es confiable y mucho mayor efectiva al tratarse de datos distribuidos normalmente.

La medida del RMSE es probablemente la más utilizada, esta métrica penaliza los errores más grandes, ya que elevar al cuadrado cifras grandes tiene mayor impacto que elevar al cuadrado cifras pequeñas.

$$\text{RMSE} = \sqrt{\sum_{t=1}^n (A_t - F_t)^2 / n}$$

Nuevamente, n representa el número de valores pronosticados utilizado para calcular el promedio.

Por último, el MAPE que representa el promedio de los errores absolutos dividido por el número actual de observaciones.

$$\text{MAPE} = \frac{\sum_{t=1}^n |A_t - F_t| / A_t}{n} \times 100$$

Una ventaja del MAPE es que permite comparar la exactitud del pronóstico entre diferentes escalas de series de tiempo.

Aunque las métricas representan una buena herramienta para determinar la exactitud del pronóstico, es necesario ubicar el intervalo factible en el que se moverán los valores futuros determinados por el modelo. Esto se puede llevar a cabo calculando un límite de confianza o intervalo de confianza para el pronóstico, usualmente determinado con el 95% de confianza. Se determina sumando y restando al valor pronosticado de cada valor en el “futuro”, el valor de la desviación estándar de los errores de pronóstico.

El intervalo de confianza se amplía mientras el horizonte de tiempo va incrementando, debido a incremento del error en el proceso, cabe aclarar que un intervalo de confianza con poca variación no es siempre el mejor modelo.

Capítulo IV

Modelación y Simulación Técnica Monte Carlo

Hay quienes se dan el tiempo necesario y cualquier cantidad de precauciones le son insuficientes antes de tomar una decisión y aquellos que a la ligera sin más preámbulo deciden el curso que se debe llevar. Aunque no hay mayor diferencia entre uno u otro, una decisión precipitada sin fundamento puede conducir al inevitable fracaso y tomarse demasiado tiempo para decidir deriva en elevados costos y pérdida de múltiples oportunidades para mejorar la posición actual. Estar preparado para lo que no se espera y evitar cometer los errores del pasado es parte de lo que se debe tener en mente como estrategia al enfrentar situaciones alternantes en estado y forma, es claro que es una tarea difícil pero no imposible. La tendencia actual que permite explorar y experimentar, en completo dominio de la situación y reduciendo el espacio de consecuencias, pudiera parecer un cuento de hadas pero gracias a aquellos que se dieron a la tarea de crear una metodología que hiciera posible manipular el tiempo y espacio se les merece respeto.

4.1 Modelos

4.1.1 Conceptos Preliminares

En el mundo real los problemas tienden a ser muy complejos, pues hay una gran variedad de hechos inherentes que afectan directamente a cualquier situación, resultando una tarea casi imposible tratar de controlar todos, pero si resulta mucho mas sencillo influir sobre los factores mas relevantes de un problema. En referencia a la determinación de llamadas de salida en un CAT la dificultad para determinar cuantas personas llamaran en un evento determinado es un fenómeno incontrolable, no obstante es factible deducir el número de llamadas que se recibirán, deducidos por experiencia empírica.

La practica y experiencia supone mayor conocimiento sobre ciertas situaciones empíricas, un mejor dominio de las causas y efectos. No olvidemos que cuanto mas aprendemos, más comprendemos nuestra ignorancia. Este trabajo pretende precisamente abatir o combatir el

desconocimiento que enfrenta la gestión de un CAT con herramientas ampliamente utilizadas con éxito en otros sectores de la industria, la comunicación y el comercio relacionados con la atención vía telefónica.

Después de haber identificado los factores críticos a partir del conocimiento empírico; y combinados de manera lógica de modo que puedan representar un conjunto generalizado se tendrá entonces un modelo conceptual del problema real que dan cabida a particularizar la complejidad del fenómeno.

Las ventajas de crear un modelo son:

- Ahorro de tiempo y esfuerzo mental
- Permite tomar decisiones con rapidez
- Permite modificaciones rápidas y efectivas

Tras haber construido un modelo, el comportamiento de este se puede deducir por medio del análisis lógico, causa y efecto. Si el análisis de los resultados del modelo se comporta acorde a la realidad, la solución del modelo servirá como solución efectiva del problema empírico, por ejemplo; determinar las llamadas de entrada en un CAT a través de un modelo de series de tiempo.

No se descarta la inclusión en errores al crear el modelo de los errores mas comunes se tienen: la exclusión de variables importantes, definir erróneamente la relación entre las variables y no menos importante el planteamiento del modelo.

Un modelo puede definirse como la simplificación de un problema empírico y el primer paso para construir uno es; elección de los factores o variables más significativos en un proceso (la mejor herramienta para llevar a cabo esta fase es la observación) que pueden clasificarse de la manera siguiente:

Variables de Decisión

Son aquellas que se encuentran bajo el control de quien toma la decisión

- Calidad en el servicio
- Calidad en la llamada

Variables exógenas

O externas, son aquellas que son importantes para el problemas de decisión, condicionadas por factores externos fuera del alcance del decisor

- Demanda futura

Políticas y restricciones

Son factibles de modificarse, son aquellas que regulan el entorno e imponen restricciones y limitaciones de carácter cualitativo

- Incidencias del personal
- Políticas de la Institución

Medidas de desempeño

Son los criterios o medidas de carácter cuantitativo

- Índices de productividad del CAT

Variables Intermedias

Se utilizan para relacionar las variables de decisión y las variables exógenas con las medidas de desempeño¹

- Desempeño por intervalos de tiempo

El modelo construido se alimenta de las variables de decisión, exógenas, políticas y restricciones, así como las intermedias dando como resultado las medidas de desempeño

A lo largo del tiempo los modelos han contribuido al desarrollo de numerosas áreas dentro de las empresas, fábricas, etc. Modelos que se han construido para describir actividades en particular y sacar mayor provecho de esta, a lo que se conoce como análisis de sensibilidad. Este tipo de análisis muestra lo que sucede si se hace determinado cambio en una de las variables exogenas de decisión, cuyo objetivo es ver como se afecta la medida de desempeño. Haciendo uso del análisis de sensibilidad, es posible determinar un aumento o reducción en la plantilla de operadores a cargo de atender las llamadas, considerando el tiempo que les toma dar la atención a un usuario.

¹ Naylor [4], p. 52

La finalidad de crear un modelo es la búsqueda de la aportación a la toma racional de una decisión que permita contar con información objetiva acerca del comportamiento de un caso en estudio, para efectos del planteamiento actual pretendiéndose demostrar la utilidad en la aplicación de modelos de Series de Tiempo para determinar la demanda futura en contraste con los resultados de un modelo de simulación orientado a establecer un valor aproximado al método mencionado, asimismo se ensaya en los resultados de salida, con alcance de aplicación práctica, para ser tomados en cuenta en la determinación de estrategias auxiliares que den paso a la renovación de la gestión y toma de decisiones.

Las faltas u omisiones al momento de crear un modelo se sugiere deben detectarse antes de la puesta en marcha, aunque no siempre es posible debido a varias razones que el analista ira descubriendo con la experiencia y por mencionar uno de esos factores es la variable tiempo, aunado a esto tenemos el grado de complejidad con el que se trate objeto de investigación.

4.1.2 Tipos de Modelos

Modelo simbólico

Es una representación gráfica que se utiliza para representar secuencias de actividades y la relación en los componentes del modelo. Comúnmente se expresan a través de diagramas de flujo y diagramas de distribución en planta. Constituye una forma fácil y sencilla en el desarrollo de un modelo, y son claros para el usuario que los interpreta. Su principal ventaja es que logra sintetizar los procesos, dada la forma global del sistema. Su limitación se tiene en que no provee elementos para el análisis cuantitativo y por su carácter estático no logra capturar la dinámica del sistema.

Modelo analítico

Es una fórmula matemática que da como resultado una solución cuantitativa. Esta técnica logra concentrar operaciones aritméticas hasta algoritmos complejos de programación matemática que proponen soluciones óptimas a los problemas. La ventaja de los modelos analíticos es la de brindar soluciones rápidas sin la necesidad de experimentación. Algunas de sus desventajas son:

- Requieren de supuestos simplificados

- No se logra incorporar el comportamiento aleatorio de muchos problemas en el mundo real.
- A medida que la complejidad del sistema crece, encontrar una solución se vuelve muy complicado

Modelo de simulación

Es una técnica que logra capturar las relaciones causa y efecto de un sistema, con que se espera representar el problema real².

4.2 Simulación

4.2.1 Conceptos Básicos

Para cada persona es común evaluar situaciones sin importar la naturaleza de la misma, tratando siempre de obtener algún beneficio o provecho por mínimo que este sea, ya sea por avanza mas rápido que los demás en las horas de mayor trafico vehicular o adquirir antes que los demás un stock de acciones con los mas altos rendimientos, incluso cambiar a la mas corta para ser el primero en ser atendido, ya sea en un supermercado o en el banco.

Para ciertas situaciones ya se cuenta con modelos matemáticos para resolver problemas complejos de decisión, tales como la estadística, programación lineal, análisis de decisiones, teoría de colas y pronostico, muchas de estas herramientas requiere de los analistas para realizar simplificaciones del modelo.

Dada su utilidad han sido ampliamente utilizados para resolver los problemas a los que se enfrentan los analistas, sin embargo proveen de poca información acerca del comportamiento dinámico del sistema en el corto plazo que se esta modelando, esto es, los cambio a lo largo del tiempo.

² Rafael E. Rivera, Michael Bernald y Rolando Salas: "Antología Simulación Industrial". Manuscrito no publicado. Universidad Fidélitas IIC-2002. Febrero 2002

Es entonces cuando se piensa en la simulación como una representación alternativa de la realidad, para lograr este propósito es necesario hacer uso de una computadora para facilitar los cálculos y en su caso hacer representaciones gráficas de su comportamiento.

En virtud que la simulación esta basada en modelar matemáticas, se hará notar la diferencia entre determinístico y estocástico; estático versus dinámico. La simulación determinista incluye parámetros y variables que están fijos y conocidos con certeza, mientras la simulación estocástica asigna distribuciones de probabilidad a alguna o todas variables o parámetros.

En la práctica ambos tipos de simulación son importantes, aunque muchos analistas se inclinan a reducir la definición de aquellos experimentos basados solo en procesos estocásticos.

Hoy en día la simulación es ampliamente aceptada y utilizada como herramienta para análisis de sistemas y problemas de decisión, a continuación se presentan algunos factores que motivan la simulación:

1. La simulación es una alternativa de solución al problema bajo estudio
2. Los modelos a simular pueden representar el proceso del mundo real en un enfoque mas realista, en virtud de que se pueden incluir restricciones
3. Se pueden implantar cambios en la estructura del modelo para responder a la pregunta ¿que pasara si?
4. La simulación es menos costosa que la experimentación de laboratorio
5. Se puede utilizar con fines pedagógicos para ilustrar un modelo o mejorar la comprensión de un proceso
6. Para procesos dinámicos. La simulación es el único medio para la observación directa dentro de los limites preestablecidos por el analista³

En un CAT existen índices para representar su productividad en términos de:

- Servicio
- Calidad
- Eficiencia
- Rentabilidad

³ lab TA Joy Woller: "The Basics of Monte Carlo Simulations",[en línea]. University of Nebraska-Lincoln Physical Chemistry Lab (Chem 484). Recuperado el 15 de noviembre del 2006. <<http://www.chem.unl.edu/zeng/joy/mclab/mcintro.html> 2005>

No obstante, estas mediciones se emplean para determinar el desempeño por día sin la posibilidad de generar escenarios experimentales, expresando en términos cuantitativos opciones representativas para gestionar la operación del CAT, posibilidad intrínseca que ofrece un modelo de simulación.

4.2.2 Pasos para realizar un estudio de simulación

Actualmente no existe una guía práctica para crear un modelo de simulación ya que incluye mayor experimentación que otros métodos, según algunos autores se identifican dos fases en un estudio de simulación; la primera es la relacionada con la formulación del problema, codificación y validación del sistema real, es decir, que el modelo desarrollado sea una representación del sistema real.

La segunda etapa esta relacionada con el desarrollo del modelo, la experimentación y optimización, durante el proceso se determinan diversas alternativas que se encuentran cerca del resultado óptimo. En caso de existir fallas, la aplicación del modelo de simulación en la práctica revelara errores que llevara a la tarea de la revisión completa en la construcción del modelo y relación entre las variables.

A continuación se enuncian los pasos relevantes en la construcción de un modelo de simulación:

- a) Formulación del problema. La observación es el elemento primario para conceptualizar la realidad y proponer los ajustes necesarios para mejorar el sistema en forma particular. Usualmente las declaraciones de las metas se expresan en forma de pregunta, hipótesis probable y efectos estimados.
 - ¿Cómo determinar el arribo de llamadas en un periodo de tiempo dado?
 - ¿Qué efectos habrán de experimentarse por un cambio en la plantilla del personal?

- b) Construcción del modelo. Se requiere mucha habilidad por parte del analista para traducir el lenguaje abstracto de un modelo en forma más simple y sencilla, se requiere la habilidad para seleccionar y modificar los supuestos básicos que caracterizan el sistema.
 - Determinar las variables del modelo

- c) Colección de datos I. Es posiblemente la actividad que mayor tiempo consume, los datos cuantitativos son necesarios por muchas razones una de ellas es la descripción del sistema que se va a simular, si no se entiende bien el sistema real es muy probable que los resultados carezcan de veracidad.
- Se tomaron bases de datos de 2003 a 2006
- d) Colección de datos II. Los datos deben juntarse para generar las variables estocásticas en el sistema
- La variable mas importante en el modelo es el arribo de llamadas por día de la semana
- e) Análisis de datos. Los datos coleccionados deben organizarse y posteriormente analizarse. Los datos generalmente se utilizan para estimar el valor de los parámetros, para fenómeno aleatorio y construcción de distribuciones acumulativas, a los que se le puede dar tratamiento por las herramientas estadísticas conocidas.
- La información se filtro por día de la semana y se agrupo por intervalos para determinar su frecuencia y generar las variables aleatorias
- f) Diseño de experimentos. Se realiza el análisis de los resultados que se generan a partir del modelo de simulación por cambios o movimientos en las variables.
- Se considera un incremento del 3% en el incremento de las llamadas por cancelación, dato promedio observado
- g) Implementación. Puesta en marcha⁴

4.2.3 Verificación del modelo y validación

El primer paso en la comprobación del modelo es la verificación con los profesionales en la materia, para asegurarse que no existen errores de lógica en la estructura o el diseño del modelo. Las técnicas implementadas para llevar a cabo este proceso son los cálculos manuales y diagramas de flujo del sistema para asegurar que los datos de entrada en el modelo son los correctos. En caso de haber utilizado algún programa de computadora es

⁴ Ibid, p. 39

necesario verificar que el mismo no tiene errores de programación antes de llevar a cabo cualquier tipo de análisis. Una vez que se ha determinado que los parámetros de entrada y la lógica del modelo son representativos, se ha completado la verificación.

El siguiente paso para determinar si los datos de salida o resultados del modelo se ajustan correctamente al sistema real es la validación, a través de la calibración del modelo, un proceso iterativo de comparación entre los datos reales y los resultados que arroja el modelo, este proceso se repite hasta que existe confianza plena en los resultados. Los pasos que se incluyen en la validación son:

Prueba de generación de variables. El proceso de análisis involucra la aplicación de pruebas no paramétricas de bondad de ajuste, así como pruebas de hipótesis correspondientes a la distribución de variables estocásticas. Estas pruebas se deben aplicar a los resultados de varios generadores para asegurar que las variables reales y simuladas se distribuyen de la misma de igual manera.

- Los resultados obtenidos tras aplicar el modelo de simulación, las variables se consideran estadísticamente significativas, una explicación mas detallada se expone en el Capítulo V.

Validación subjetiva. El diseño, así como los resultados del modelo de simulación, se debe sujetar a revisión por personal familiarizado con el sistema real, pero no directamente involucrado con el proceso de estudio de simulación.

- Por experiencia propia, se considera sumamente importante validar y calibrar el modelo con personas de amplia experiencia y conocimiento sobre el tema de estudio, esto es sumamente útil para contrarrestar errores u omisiones en el modelo, tales como suprimir, ingresar o asociar variables, factores internos o externos no considerados, o eventos relevantes susceptibles de influenciar el resultado obtenido.

Validación Histórica. La serie resultante al aplicar el modelo de simulación comparado con los datos del sistema real puede representar diferencias, que al ser mínimas se puede decir que el modelo ha sido validado, sin embargo no es una garantía que el modelo predice con exactitud el comportamiento del sistema real en diferentes condiciones.

- La condición de ajuste con la realidad, depende en gran medida de la variabilidad de los datos reales, sin embargo el resultado de un modelo de simulación pudiera representar las condiciones
- “óptimas” o libres de errores, para el desempeño de una actividad cualquiera que esta sea.

Bondad de Ajuste

Es necesario considerar varios criterios para indicar si la trayectoria de la serie simulada concuerda con la trayectoria de la serie observada, de modo que la concordancia pueda atribuirse a la casualidad. Se debe tener en cuenta medidas y técnicas para comprobar la bondad de ajuste de un modelo de simulación. Aunque existen varias técnicas para llevar a cabo esta tarea, los especialistas en administración y economía se han ocupado la mayor parte de las veces en utilizar la prueba gráfica para validar los modelos.

A continuación se enumeran las técnicas de referencia utilizadas en el presente estudio y aquellas susceptibles de aplicación que profundizan más en la validación de modelos de simulación:

1. *Análisis de varianza*. Es un conjunto de técnicas para analizar los datos que se pueden usar para comprobar que la hipótesis de que la media (o varianza) de una serie generada por un experimento es igual a la media (o varianza) de la serie observada.
2. *Prueba Ji Cuadrada*. Es la prueba estadística clásica que se puede emplear para comprobar la hipótesis de que el conjunto de datos generados por un modelo de simulación posee la misma distribución de frecuencia que el conjunto de datos históricos observados.
3. *Prueba Kolmogorov Smirnov*. Es del tipo de prueba libre de distribución (no paramétrica) relacionada con el grado de concordancia que hay entre la distribución de un conjunto de valores de muestra (serie simulada) y alguna distribución teórica especificada. La prueba implica la especificación de la distribución de frecuencia acumulativa de datos simulados y reales. Maneja las observaciones individuales por separado y, al contrario de la prueba Ji cuadrada, no pierde información debido a la combinación de las categorías.
4. *Análisis de regresión*. Se sugiere la posibilidad de efectuar regresiones de series reales sobre series generadas y comprobar si las ecuaciones de regresión resultante tienen intersecciones que no difieran notablemente de cero y cuenten con pendientes bastante cercanas a la unidad.

5. *Análisis espectral.* Tiene como objetivo principal la cualificación y evaluación de los datos correlacionados una vez que estos se transforman dentro del dominio de la frecuencia.
6. *Coefficiente de desigualdad de Theil.* Denominado U, proporciona un índice que determina el que un modelo de simulación proporcione predicciones retrospectivas P_i de datos históricos observados A_i . U varía entre 0 y 1. Si $U=0$, se obtiene predicciones perfectas; si $U=1$, las predicciones son deficientes.

$$U = \frac{\sqrt{1/n \sum (P_i - A_i)^2}}{\sqrt{1/n \sum P_i^2 + \sum A_i^2}}$$

A pesar de la discusión del éxito o fracaso de un modelo de simulación se debe medir a través de lo bien que el modelo predice el fenómeno del sistema real, no se ha dicho que el cuidado que se pone en la selección de las suposiciones y en la comprobación estadística de estas carezca de propósito o sean actividades inútiles⁵.

La serie resultante después de aplicar la simulación, fue sujeta a las 6 pruebas antes mencionadas para sancionar si la información histórica se ajusta con la información real, concluyendo la significancia estadística de los datos simulados, no se omite mencionar que un grado de exactitud al 100% requiere de un enorme esfuerzo que quizá no valga la pena, es decir, un grado medianamente alto de conformidad con la realidad y una basta experiencia en el tema son las mejores herramientas para analizar una serie de datos y emitir un juicio sumamente aceptable.

4.3 Simulación de Monte Carlo

Como introducción a la Simulación Monte Carlo, se comienza con una breve ejemplificación en la obtención de los números aleatorios, ya que por la facilidad en el acceso a generadores de números aleatorios se desconoce la mecánica para generarlos.

⁵ Naylor [4], p. 168-169

4.3.1 Números Aleatorios

Para evitar el uso de tablas para generar números aleatorios existen diversas técnicas para generarlos, tal como el **método de cuadrado medio** propuesto por J. Von Neuman en 1946. En la práctica los números generados por métodos matemáticos se denominan como números pseudo aleatorios.

Este método opera utilizando como siguiente número aleatorio los dígitos intermedios del cuadrado del número aleatorio precedente, la metodología a seguir se presenta a continuación:

1. Elegir un número aleatorio de n dígitos (a este número se le denomina semilla en términos computacionales)
2. Elevar al cuadrado el número elegido y añadir un cero a la derecha del número para producir un número con 2n dígitos.
3. El nuevo número aleatorio se determina eligiendo los n dígitos de la parte media del número en la segunda etapa.
4. Repetir los pasos 2 y 3 para obtener números aleatorios adicionales.

Para ejemplificar esta metodología, se elige un número aleatoriamente 2356, se eleva al cuadrado dando como resultado 5550736, se le añade un cero al final 55507360 y el nuevo número aleatorio es 5073, repitiendo el proceso tenemos que 5073 al cuadrado resulta 25735329(0) el nuevo número es 7353, si se requiere un número mayor de aleatorios se continua el proceso. Ya que por lo general los números aleatorios se encuentran entre 0 y 1, se antepone un punto decimal a la primera cifra del número aleatorio obtenido o en su defecto se divide entre 10 mil si los aleatorios se crean por medio de una hoja de cálculo.

Es pertinente aclarar que este método no es el más adecuado para generar una gran cantidad de aleatorios antes de que se repitan o se de lugar a incongruencias con el método, así mismo al aplicar pruebas de aleatoriedad cabe la posibilidad que no se pasen con satisfacción.

Otra técnica mas confiable para generar números aleatorios es la de los **métodos congruenciales** en la que se tiene:

$$X_{n+1} = aX_n + c \text{ (modulo } m)$$

en esta se genera el siguiente entero aleatorio X_{n+1} utilizando el entero precedente X_n , las constantes a y c y el modulo m . Después de que se genera el entero $aX_n + c$ se ejecuta el modulo m , para producir el nuevo aleatorio entero X_{n+1} . Para comenzar se necesita el número inicial llamado semilla designado como X_0 que puede obtenerse de una tabla de números aleatorios. Nótese que el resultado siguiente X_{n+1} depende del aleatorio precedente, esta característica de este tipo de generadores minimiza los recursos de almacenamiento, a su vez impone restricciones para el periodo siguiente.

Al calcular X_n se generan los números aleatorios distribuidos en el intervalo $[0,1)$, al dividir por m , es decir, $R_n = X_n / m$. Si se desea que los números R_n se distribuyan en el intervalo $[0,1]$ se divide entre $(m - 1)$, así que $R_n = X_n / (m - 1)$ la recursividad para R_n es un nuevo número aleatorio con un valor medio aproximado de .5. Aunque existen otros métodos por computadora que utilizan el dígito semilla no se hará mención a la manera de calcular números aleatorios, pero efectos en este estudio se utilizara la función de hoja de cálculo ALEATORIO para la generación de los mismos⁶.

Generador de variables aleatorias

Para elaborar un generador de proceso se trabaja con la función de densidad probabilística y la función de densidad acumulada, que se define:

$$P(x) = 1 / (b - a) \quad a \leq x \leq b$$

La representación gráfica todos los valores a y b tienen densidades iguales a $1/(b-a)$. La función de densidad acumulada se calcula como se muestra a continuación:

$$P(x) = \int_a^x p(x) dx$$

sustituyendo $P(x) = 1/(b-a)$, tenemos:

$$P(x) = \int_a^x 1/(b-a) dx$$

⁶ Ibid, p. 39

$$\begin{aligned}
 &= 1/(b-a) \int_a^x dx \\
 &= 1/(b-a) [x]_a^x \\
 &= 1 / (b - a) [x - a] \\
 P(x) &= x - a / b - a^7
 \end{aligned}$$

Para obtener un número aleatorio con la función de distribución conocida se conoce como procedimiento de generación inversa. El procedimiento implica igualar la variable aleatorio uniforme R ($0 \leq R \leq 1$) a $P(x)$ y se despeja x , como se muestra enseguida:

$R = P(x) = x - a / b - a$ despejando x tenemos $R(b - a) = x - a$, $x = a + R(b - a)$, esta ecuación representa el generador de proceso para la distribución uniforme, es posible generar variables con distribución uniforme entre a y b , simplemente identificando los valores de a y b y generando un número aleatorio R y sustituyendo en la ecuación. Por ejemplo una variable entre 10 y 20 la relación es:

$$x = 10 + R(20 - 10)$$

para el número aleatorio .6134 el valor de la variable aleatoria generada es 16.134.

4.3.2 Técnica de simulación Monte Carlo

El método de Monte Carlo (MC) tiene origen en la ciudad de Monaco, el nombre que toma esta técnica le pertenece originalmente a uno de los mas conocidos casinos de esta ciudad, idea que surge gracias al juego de ruleta (un generador simple de números aleatorios). El método MC es considerado una técnica estocástica, ya que se basa en el uso de números aleatorios y distribuciones de probabilidad para investigar problemas.

⁷ Evans, Olson [3], p. 71

Los métodos numéricos conocidos como MC pueden ser descritos como modelos de simulación estadística, donde la simulación estadística es en general cualquier métodos que utiliza una secuencia de números aleatorios para llevar a cabo una simulación, en contraste se encuentran los métodos numéricos discretos convencionales, que se aplican a ecuaciones ordinarias o ecuaciones diferenciales que describen algún sistema físico o matemático.

En muchas aplicaciones de MC en un proceso físico se simula directamente, ya que incluso no hay necesidad de escribir la ecuación diferencial que describe el comportamiento del sistema, el único requisito es que el sistema físico o matemático sea descrito por una función de distribución de probabilidad (fdp), es entonces cuando se procede a simular haciendo uso de los números aleatorios de la fdp. El resultado de la simulación se toma como el promedio del número de observaciones, es decir, de las distintas corridas que se llevaron a cabo, que pueden llegar a ser hasta de millones. En la práctica, se puede predecir el error estadístico en la media de las observaciones y estimar el número de corridas por el método MC con un nivel de error dado.

La simulación se lleva a cabo a través del muestreo de la fdp, que requiere una forma efectiva para generar números aleatorios distribuidos uniformemente en el intervalo (0,1). El resultado de este muestreo o corridas de simulación, se debe acumular o llevar un registro apropiado para producir el resultado deseado, para proponer una solución del problema a la mano. En contraste con una solución obtenida por un método numérico convencional que comienza con un modelo matemático del sistema, que se simplifica y se resuelve el conjunto algebraico de ecuaciones para las incógnitas desconocidas del sistema.

Se debe de tener en mente que la descripción general de la aplicación práctica de la técnica MC puede no contar con elementos estocásticos, como la evaluación de una integral definida o un sistema de ecuaciones lineales. Sin embargo, se puede confiar en la solución en términos de la fdp, a pesar de parecer una solución artificial este paso permite la designación de un proceso estocástico para el propósito de la simulación⁸.

⁸ Op. Cit., p. 39

Componentes de la técnica Monte Carlo

A continuación se describen brevemente los componentes de la técnica de simulación MC que forman parte de muchas de las aplicaciones desde su desarrollo, sin embargo los componentes aquí considerados forman parte de las instrucciones generales para construir un modelo de simulación mediante la técnica MC y no forman parte de la construcción de modelos matemáticos. Los componentes primarios del modelo incluyen:

- a) Funciones de distribución de probabilidad, sistemas físicos o matemáticos se deben describir a través de un conjunto de fdp.
- b) Generador de números aleatorios, una fuente de números aleatorios uniformemente distribuidos debe estar disponible.
- c) Muestreo, descripción de la muestra proveniente de fdp conocidas.
- d) Registro, los resultados generados por el modelo de simulación se deben registrar para obtener el resultado deseado.
- e) Estimación del error, estimar la varianza como función del número de corridas o intentos.
- f) Técnicas de reducción de varianza, métodos para reducir la varianza en la solución estimada.
- g) Paralelización y vectorización, algoritmos que permiten a la técnica MC mayor eficiencia a través del uso de programas de computadora.

Su aplicación práctica se refiere a:

- a) El conjunto de llamadas de entrada en un CAT, que se describe a través de conjuntos de fdp.
- b) Los números aleatorios se generaron a través de la función ALEATORIO de la hoja de cálculo de Excel.
- c) Las muestras se obtuvieron de bases de datos de la operación de un CAT y trabajo de campo (monitoreo de llamadas).
- d) El registro de los resultados se almacena en hoja de cálculo para su análisis.
- e) y f) Se aplicaron las técnicas de análisis expuestas en los capítulos que conforman el presente estudio.
- f) La simulación a través del uso de la hoja de cálculo de Excel se detalla en el anexo al final de este trabajo.
- g) Para el presente caso, se utilizan un sencillo algoritmo de programación diseñado en Visual Basic para efectos de reducir el tiempo en las corridas de simulación.

Los componentes anteriores dan la pauta para examinar y analizar problemas de sistemas complejos con cientos de variables, que resultaría casi imposible resolverlo a través de un modelo matemático. Con la técnica de simulación MC, un sistema grande se puede muestrear a través de configuraciones de números aleatorios, y los datos pueden describir el sistema como un todo, con el que no se obtendrá un resultado específico, si no una serie de posibles resultados derivados de la simulación de un problema⁹.

En conjunto analistas expertos en modelos cuantitativos y personal con amplia experiencia en la Gestión del CAT, se desarrolla un importante equipo de trabajo capaz de proporcionar soluciones con elevado de efectividad para optimizar la operación y uso de los recursos.

⁹ Op. Cit., p. 39

Capítulo V

Aplicación práctica de la simulación

Caso sistema de Televisión de paga

A partir del último bienio de la década de los veinte, y de manera especial a partir de los años cincuenta en que la televisión comienza a funcionar de manera cotidiana y a transformarse en una presencia de singular importancia en la vida cultural, política y económica de México, ocurren una gran cantidad de acontecimientos que van conformando el complejo fenómeno en que la TV se ha convertido en nuestros días. En las siguientes líneas se hace descripción de los acontecimientos más relevantes de la TV de paga en nuestro país.

El 20 de mayo de 1969, la empresa Cablevisión, S.A., filial de Telesistema Mexicano, obtiene la concesión para prestar el servicio de TV por cable en la Ciudad de México, el cual se empieza a proporcionar un año después, en 1970.

El 1 Septiembre de 1989 inicia sus transmisiones en México un nuevo sistema de TV por suscripción. Se trata de MVS Multivisión que emplea una tecnología nunca antes utilizada en nuestro país denominado MMDS, sigla que significa Multichannel Multipoint Distribution System, es decir, Sistema de Distribución Multicanal Multipunto. El sistema consiste en el envío, a través del espacio aéreo, de señales de televisión codificadas hacia los hogares de los suscriptores a quienes se dota previamente del equipo necesario para la recepción.

El 9 de Marzo de 1995 las empresas Hughes Communications, de Estados Unidos, Organización Cisneros, de Venezuela, Televisión Abril, de Brasil, y MVS Multivisión de México, dan a conocer la creación del Consorcio Galaxy Latin América. Se trata de una alianza estratégica para lanzar al mercado DiecTV, un servicio de televisión vía satélite “directo al hogar” (DTH)¹ con cobertura Latinoamericana.

¹ Direct To Home, servicio de televisión vía satélite directa al hogar. Se trata de un nuevo sistema que permite la transmisión de señales de TV a los hogares directamente desde un satélite. La diferencia entre el sistema DTH y otros servicios de TV directa al hogar vía satélite es que permite transmitir audio y video digitales, es decir, el sonido tiene calidad de un compact disc, y la imagen una definición mucho mayor que la normal.

Las empresas que proveen el servicio de TV por suscripción manejan un CAT para proporcionar atención a sus clientes en la comodidad de sus hogares para cualquier asunto relacionado con servicios de activación de eventos especiales, saldo, dudas generales, reparación, en este último caso en ocasiones se requiere de la visita del personal especializado para reparar cualquier falla con el receptor digital, conexión, cableado, etc., misma que se agenda desde la oficina matriz.

La continua expansión de este servicio al interior de los distintos Estados de la República, representa un excesivo gasto de capital para establecer un centro de atención en cada Estado, por tal motivo el servicio a los clientes se hace por comunicación telefónica desde la oficina matriz, atendido por los representantes dispuestos a dar una respuesta eficiente en el menor tiempo posible, garantizando así la satisfacción del cliente.

5.1 Análisis de la información

5.1.1 Estadísticas de un CAT

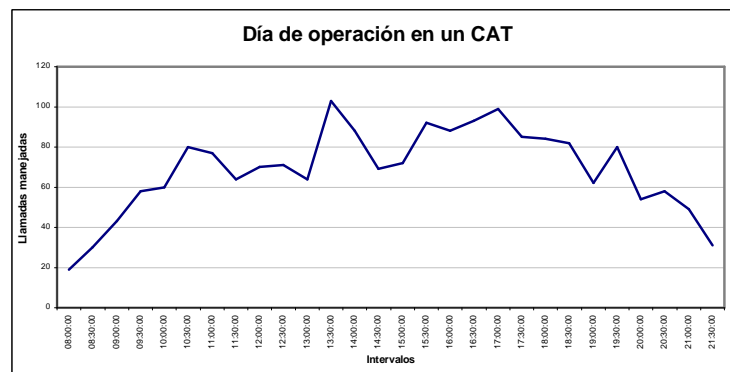
En un CAT se genera una cantidad inimaginable de datos cada segundo y hora al momento de entrar en operación, en muchos casos la operación permanece constante las 24 horas del día, en el caso del presente estudio se observó que la operación comienza a las 6 de la mañana y hasta las 2 de la mañana del otro día. Por su naturaleza particular, se requiere dedicar especial atención a ciertos puntos en la operación para garantizar la satisfacción de los clientes y usuarios, identificando la carga de trabajo por hora y día. La operación a lo largo del día describe rasgos distintivos propios, identificándose distintas magnitudes denominadas picos y valles, como se observa en la Gráfico 1.

La obtención de los datos para realizar el análisis provienen de reportes generados en la operación del CAT de forma diaria comprendiendo el periodo de 2001 a 2005, así como muestras de datos de trabajo de campo al que se le conoce como Monitoreo de Calidad dividido en Monitoreo en Sombra, que consiste en conectar un auricular al aparato telefónico del ejecutivo de servicio observando el desempeño del mismo durante la llamada. El otro tipo de

Monitoreo es el denominado Remoto (no presencial) en el que se da seguimiento al desarrollo de una llamada sin que el ejecutivo tenga conocimiento que esta siendo escuchado.

Los datos que describen la curva corresponden a las llamadas recibidas durante el día, concentrándose en intervalos de media hora; ya que resultaría poco práctico almacenar los registros en lapsos de 1 minuto o menores.

Gráfico 1



Así mismo, otros registros de vital importancia se almacenan en el mismo intervalo de tiempo para su análisis, como lo son: llamadas respondidas (1), abandonadas (2), tiempo promedio de respuesta (3), tiempo promedio en comunicación con el cliente (4), nivel de servicio (5), tiempo en espera de llamada (6), tiempo total del grupo de trabajo (7), a esta concentración de datos se le conoce como Reporte ACD por grupo de trabajo que se actualiza cada media hora en tiempo real y con quince minutos de retraso para consultas históricas.

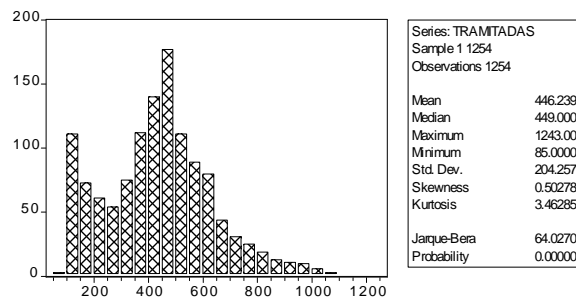
El reporte que se usa actualmente ha influido de manera determinante en el aumento de la calidad al cliente, ofreciendo información uniforme para lograr un nivel de servicio eficaz y eficiente, además de representar un insumo vital para el desarrollo y fortalecimiento empresarial. Enseguida se presenta el resumen de un día en la operación:

INTERVALO	CALLS HNDL	(1) CALLS ANSWD	(2) CALLS ABAND	(3) ANSWD ACW	(4) AGT TALK	% ABAND	AVG AGT	TOTAL CONPERS	(5) SERVICE LEVEL	(6) WAIT TIME	(7) ACD TALK
08:00:00	19	19	0	2.68	328	0.0	13	01:44:07	100	04:37:15	01:43:46
08:30:00	30	30	0	2.43	403	0.0	13	03:21:22	100	03:06:43	03:21:22
09:00:00	43	43	0	2.58	231	0.0	13	02:57:55	100	03:28:53	02:45:38
09:30:00	58	58	0	5.4	310	0.0	13	05:03:58	100	01:22:10	05:00:08
10:00:00	60	51	9	42.9	347	15.0	13	04:59:18	100	01:27:21	04:55:20
10:30:00	80	67	13	57.15	327	16.25	13	06:06:46	17	00:19:29	06:04:37
11:00:00	77	55	22	154.38	401	28.57	12	06:10:42	11	00:00:06	06:07:11
11:30:00	64	52	12	178.52	372	18.75	12	05:49:56	28	00:03:03	05:22:28
12:00:00	70	55	15	186.89	414	21.43	13	06:23:01	0	00:00:42	06:19:34
12:30:00	71	53	18	142.64	364	25.35	11	05:23:13	27	00:04:06	05:21:14
13:00:00	64	36	28	301.36	573	43.75	12	05:52:44	3	00:26:20	05:43:33
13:30:00	103	85	18	80.75	428	17.48	21	10:12:38	12	00:05:26	10:06:42
14:00:00	88	80	8	26.3	464	9.09	22	10:28:24	22	00:13:22	10:19:02
14:30:00	69	68	1	2.53	416	1.45	21	08:11:12	100	02:29:01	07:51:43
15:00:00	72	71	1	2.54	334	1.39	22	06:42:41	100	04:12:53	06:34:57
15:30:00	92	91	1	4.25	340	1.09	22	08:50:21	100	02:08:10	08:35:29
16:00:00	88	76	12	76.82	414	13.64	19	08:48:50	20	00:37:21	08:44:53
16:30:00	93	86	7	50.92	258	7.53	18	06:41:58	79	00:20:56	06:10:19
17:00:00	99	87	12	78.31	395	12.12	22	10:29:29	2	00:50:57	09:32:02
17:30:00	85	85	0	2.58	299	0.0	27	08:00:23	100	04:44:32	07:03:34
18:00:00	84	84	0	2.63	302	0.0	23	07:50:07	100	04:08:10	07:02:55
18:30:00	82	79	3	17.57	417	3.66	21	09:51:40	53	01:10:42	09:08:50
19:00:00	62	61	1	3.61	397	1.61	22	08:03:08	97	03:29:20	06:44:00
19:30:00	80	78	2	2.63	364	2.5	23	08:55:14	100	03:00:24	07:52:50
20:00:00	54	54	0	2.54	424	0.0	22	07:13:04	100	04:09:23	06:22:02
20:30:00	58	58	0	3.12	352	0.0	19	06:41:40	100	03:29:09	05:40:26
21:00:00	49	45	4	21.67	444	8.16	18	06:09:29	67	03:17:43	05:32:40
21:30:00	31	31	0	2.65	414	0.0	17	04:12:41	100	04:54:20	03:33:39
TOTAL	1,925	1,738	187	47.83	373	9.7	17	191:16:01	61.16	58:17:57	179:40:54

La importancia relativa del contenido del reporte es la concentración de información indispensable para identificar la carga de trabajo por horas pico, días y semanas. Sin embargo, aglutinar cúmulos de datos sin tratamiento formal estadístico lo único que se posee es una enorme lista de cifras históricas ocupando memoria en el disco duro del servidor.

En vías hacia la mejora continua se comienza por agrupar la información en intervalos para facilitar su análisis y deducir los indicadores iniciales para comenzar el diagnóstico y exponer el

dictamen. Para este caso, el análisis se aplica a las llamadas tramitadas, del que se cuenta con información diaria y semanal correspondiente de diciembre del 2001 a 2005, a continuación se presenta un informe completo que resume los indicadores estadísticos mas relevantes.



Se detecta dispersión en los datos, que se atribuye a la variabilidad de la carga de trabajo de los días de la semana, así como de eventos fortuitos que elevan enormemente el arribo de llamadas para solicitar algún trámite por el servicio. Se observa además asimetría positiva y nivel elevado en el coeficiente de kurtosis; por ende se asume que los datos no se comportan normalmente, que se confirma con la probabilidad de la prueba Jarque Bera.

5.1.2 Pronóstico en un CAT

Planificar las llamadas entrantes en función del volumen es una tarea diaria que permite asignar el número de estaciones y representantes requerido en los distintos horarios de la operación del CAT, con la finalidad de garantizar el mejor servicio a los clientes de la empresa.

El análisis estadístico de las bases de datos y el pronóstico respectivo se realizó en hoja de cálculo de Excel, esto con la finalidad de mostrar la utilidad y beneficio del análisis histórico de las llamadas entrantes y registro de trámites administrativos generados; para determinar la probable carga de trabajo esperada en el futuro y determinar la asignación óptima del personal en los días y horarios con mayor volumen de llamadas.

Las técnicas de predicción empleadas para determinar el pronóstico de las llamadas atendidas, manejadas y tramitadas, son:

- Promedios Móviles
- Suavización Exponencial
- Mínimos Cuadrados ordinarios

A continuación se presenta una tabla comparativa del valor pronosticado que se obtuvo al aplicar promedios móviles con rezago de 3 MA(3) y 5 periodos MA(5), suavización exponencial simple (SES), doble (SED) y regresión lineal (MCO), así mismo, se exponen las métricas para determinar el error del valor futuro obtenido, la desviación estándar de las perturbaciones estocásticas, bondad de ajuste (R^2) e intervalos de confianza para cada uno de los modelos. Adicionalmente, en el **Anexo 4A, 4B, 4C, 4D** se incluyen las fórmulas ocupadas para efectuar los cálculos correspondientes a cada modelo de pronóstico, así como las fórmulas para deducir las métricas de error.

Pronóstico para llamadas Tramitadas

Modelo	Pronóstico	MAD	RMSE	MAPE	R^2	Desviación Estándar	Intervalo de Confianza	
							Máximo	Mínimo
MA (3)	2639	341	440	13.6%	0.463	441	3521	1758
MA (5)	2773	357	454	14.2%	0.377	455	3682	1863
SES	2792	439	553	17.6%	0.205	554	3901	1684
SED	2826	457	578	18.2%	0.306	579	3984	1668
MCO	2883	458	566	18.6%	0.038	568	4019	1748

Adviértase que el valor pronosticado con menor variación en los términos de error y mejor bondad de ajuste y menor desviación estándar se produce con el modelo MA(3), le sigue el MA (5) con variaciones mínimas en las métricas de error y la R^2 . Nótese que el valor del pronóstico se subestima al aplicar el modelos MCO, debido a la elevada variación que experimentan los datos que se confirma con el pobre ajuste de la línea de tendencia con los datos originales.

Cabe mencionar que la situación que se enfrenta es un fenómeno social con muchos altibajos y alguien poco familiarizado con modelos de pronóstico podría sugerir o en su caso, destacar que el valor mínimo del intervalo de confianza para el modelo MA (3) incluye el valor del pronóstico obtenido por la técnica MCO y los valores obtenidos para el pronóstico en MA(5), SES, SED no

superan el valor máximo del intervalo para MA(3), así mismo, mientras el valor del pronóstico decrece el error porcentual muestra tendencia inversa.

El cálculo del pronóstico de los diferentes métodos utilizados en este estudio se realizaron en hoja de calculo manualmente, así como el calculo de las métricas de error para su validación y verificación, el desarrollo de las fórmulas se puede consultar en el Anexo.

5.2 Modelación en un CAT

5.2.1 Llamadas entrantes

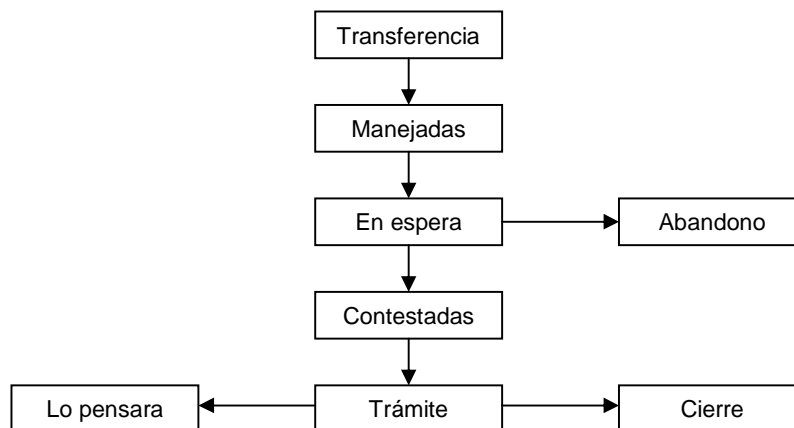
El presente estudio evalúa el desempeño del Área de Anticancelaciones, su principal enfoque es la recepción de llamadas, es un área especializada en asesorar y atender las solicitudes de trámite de cancelación. Se atiende 6 días por semana en horario de 8 am a las 8 pm, con 30 representantes al servicio de los clientes. La llamada recibida por un representante del Área de Anticancelaciones sondea el motivo real por el cual el cliente solicita cancelar el servicio, con la finalidad de convencer al cliente que cuenta con el mejor servicio de TV de paga, la mejor programación, calidad de imagen y audio, posteriormente se ofrece una serie de promociones y descuentos que el cliente puede elegir, válidos solo en caso de continuar con su servicio activo.

En promedio un ejecutivo atiende 38 llamadas en su turno de trabajo, de los cuales evita que cancelen 55%, cancelan 30.5% y le da seguimiento a los tramites que no respondieron inmediatamente 14.5%, estas cifras varían positiva y negativamente acorde a la habilidad en el manejo de llamada de los representantes.

Modelo Anecdótico

El trámite de solicitud de cancelación inicia en el Área de Atención a Clientes, donde se genera un reporte de cancelación y se transfiere la llamada al Área de Anticancelaciones para continuar el seguimiento de la petición del cliente. En caso de tráfico de llamadas el cliente tendrá que esperar su turno para ser atendido, si el cliente considera que el tiempo en línea es

demasiado y que no vale la pena esperar entonces corta la llamada, esta acción se registra como abandono. Si la llamada es contestada por uno de los representantes del Área, se solicita el motivo por el cual requiere cancelar con la finalidad de identificar la razón real por el cual se pide el trámite, e iniciar el proceso de negociación para persuadir al cliente que continúe con el servicio destacando las ventajas y promociones vigentes aplicables a la fecha, por su antigüedad y buen historial de pagos, no obstante la decisión final dependerá del cliente. Cualquiera que sea el caso, cancelación o no cancelación, se continúa trámite administrativo para establecer el estatus final, para asegurarse no dejar ninguno sin resolver, sobre todo si los clientes mencionan comunicarse posteriormente para indicar su decisión, en estos casos el personal administrativo tiene que comunicarse con el cliente y confirmar la continuidad con el servicio. Enseguida se presenta un diagrama que reúne las particularidades del proceso para realizar el trámite de cancelación:



Las variables identificadas de mayor importancia por la volatilidad que experimentan son:

- Llamadas manejadas, transferencias directas del Área de Atención a Clientes
- Llamadas contestadas, atendido por un representante
- Clientes cancelados, se prescinde del servicio definitivamente
- Clientes no cancelados, continúan con el servicio por tiempo indefinido
- Trámite administrativo, cierre del trámite

Modelo de Simulación

El modelo esta orientado a explorar el mantenimiento o disminución de la base de datos por trámite de cancelación del servicio de TV por suscripción. Las llamadas de entrada al Área de Anticancelaciones son transferencias del Área de Atención a Clientes, que generan la carga de trabajo denominada "llamadas manejadas", que se asume siguen una distribución normal. Los clientes que deciden evitar la espera para ser atendidos por un representante y cuelgan, generan las llamadas abandonadas.

Si la comunicación es exitosa con uno de los representantes se genera: un trámite de cancelación, no cancelado o en su defecto sin respuesta inmediata, este último caso constituye en el futuro un exceso en las llamadas que entran por el Area de Atención a Clientes para notificar la decisión de continuar o prescindir del servicio que presta la empresa (e igualmente se asume siguen un comportamiento normal), como consecuencia se espera un aumento de llamadas para el Area de Anticancelaciones y por ende un incremento en el nivel de abandono, lo que se traduce en bajos niveles de servicio.

El siguiente y último paso es el cierre administrativo de los trámites que se generan en el día, la acumulación determina el rezago de trabajo, y se presenta debido a que la capacidad máxima para el cierre es de 3,500 casos semanales. El cierre de trámites cancelados define la reducción real de la base de clientes de la empresa, su crecimiento promedio observado es de .3 casos por semana.

5.3 Simulación Monte Carlo en hoja de calculo

5.3.1 Simulación operación In Bound

El modelo de simulación analítico genera valores aleatorios de los tramites cancelados, no cancelados, llamadas atendidas y manejadas en función de la función de distribución de probabilidad del histórico de los datos observados por día de la semana. Como se conoce el valor máximo y mínimo, la distribución uniforme para generar los números aleatorios asume por

definición que todos los valores intermedios tienen exactamente la misma probabilidad de ocurrencia, aplicando el métodos de la transformada inversa se obtiene que el generador para variables uniformes se determino a partir de:

$$X = a + (b - a) * \text{RND}$$

donde a es el valor mínimo, b el valor máximo y RND el número aleatorio. Este tratamiento se aplico para generar las variables canceladas, no canceladas, atendidas y manejadas, acorde a la frecuencia acumulada relativa que define los límites máximo y mínimo para cada intervalo, consultar Anexo 1 en el que se muestra la mecánica para el calculo de las variables en Excel . En complemento a lo anteriormente dicho cabe mencionar brevemente el procedimiento aplicado para obtención de las variables.

En primera instancia se realizo una separación de los datos históricos en una semana típica de 5 días por semana, es decir, Lunes a Viernes en virtud de que la operación para Sábados y Domingos no es representativa en términos de generación de información. Posterior a la segmentación, se agruparon los datos por intervalos de frecuencia para obtener la fdp por día, aplicando la fórmula anteriormente mencionada para deducir el registro resultante por su frecuencia de aparición, mismo tratamiento que se aplico a los datos relacionados con Cancelados, No Cancelados, Retenidos + Cancelados, Atendidas, Manejadas(de esta se deduce el abandono).

La fórmula para la obtención de los resultados de salida para Cancelados, No Cancelados, Retenidos + Cancelados, es la siguiente:

Acumulado registro por frecuencia = $\sum d$; donde d es el día de la semana

Abandono = Manejadas – Atendidas

No Canceladas = (Recibidas – Manejadas) * % No Cancelados

Cancelados = Recibidas – Manejadas – No Canceladas

De estas ecuaciones se deducen los siguientes indicadores:

Carga de trabajo = Cancelados + No Cancelado

Incremento en la plantilla = Base Total + No Cancelado

Decremento en la plantilla = Base Total – Cancelado

En el seguimiento de la operación del CAT se detectó un patrón de comportamiento el cual era factible dar continuidad a través de ecuaciones simples mostrando visulmente por medio de gráficos, los efectos tanto positivos como negativos, al ejercer una influencia externa controlada (como reducir el tiempo de llamada) y conocer los efectos que pudieran llegar a tenerse instituyendo estrategias operativas para incrementar la productividad o reducir en la medida de lo posible los efectos por una deficiente administración.

En la tabla siguiente se muestra la tabla de frecuencias, detallando los valores máximo y mínimo de la base de datos correspondiente al día lunes de lo trámites cancelados, los intervalos, frecuencia de los datos, frecuencias relativa y acumulada. La variable aleatoria se deduce haciendo uso del generador de números aleatorios en Excel, que designa un aleatorio que se compara con los límites máximo y mínimo de la frecuencia acumulada de cada intervalo para asignar el valor de la variable aleatoria aplicando la transformada inversa para calcularlo.

CANCELADOS					
Max	750	Min	50	Aleatorio	0.3112
Intervalo	Frecuencia	Frecuencia Relativa	Frecuencia Relativa Acumulada	Variable Aleatoria	
50	120	52	0.249	0.249	
121	191	65	0.311	0.560	140
192	262	38	0.182	0.742	
263	333	17	0.081	0.823	
334	404	5	0.024	0.847	
405	475	12	0.057	0.904	
476	546	6	0.029	0.933	
547	617	4	0.019	0.952	
618	688	8	0.038	0.990	
689	755	2	0.010	1.000	
Total		209			140

El abandono se obtiene de restar a las llamadas manejadas; las llamadas atendidas. La carga de trabajo se obtiene de la diferencia del máximo de solicitudes que puede cerrarse administrativamente y los trámites generados en el día, el excedente que llegara a suscitarse, ya sea por exceso de trámites generados o dejar de cerrar casos, cualquiera que fuere la causa será la carga de trabajo acumulada para un nuevo día. Nuevamente se aplica la transformada inversa que genera variables aleatorias de la carga de trabajo utilizando los límites máximo y mínimo conocidos para el cierre administrativo, los valores correspondientes son 3,250 y 2,500, la fórmula se presenta a continuación:

$$X = 2,500 + (3,250 - 2,500) * RND$$

El paso del tiempo en todos los casos se integra en la simulación al contemplar el crecimiento promedio en los tramites por semana: cancelados de .3 y no cancelados de .8, así como en las llamadas atendidas y manejadas de .8 y 11.2 llamadas respectivamente.

5.3.2 Validación

Antes de realizar cualquier explicación cuantitativa, el primer paso para comprobar los resultados del modelo simulado con el sistema real es verificar el comportamiento del sistema simulado con las personas estrechamente relacionadas con el mismo, en virtud de que pueden señalar las inconsistencias producidas por una combinación única de circunstancias que nunca fueron identificadas en el modelo, en este caso se debe reestructurar el modelo propuesto. Recuérdese que el propósito de crear un modelo es la representación con mayor ajuste al sistema real, la clave se encuentra en la total aceptación de que el modelo se comporta como estaba planeado.

La validación consiste en aplicar una serie de pruebas cuantitativas para garantizar que el modelo simulado es estadísticamente significativo, es decir, la comparación detallada de las corridas de simulación con los datos del sistema real para mostrar si existen grandes diferencias entre uno y otro, en caso registrarse grandes diferencias, se recomienda revisar el planteamiento del modelo original.

Las medidas aplicadas en el presente estudio para confirmar la bondad de ajuste del modelo simulado se comentan a continuación, y las que se obtuvieron utilizando la hoja de cálculo de Excel se muestran en el **Anexo 2A, 2B, en el Anexo 2C** se incluye la programación empleada para automatizar la obtención de los resultados :

Análisis de Varianza

El análisis de varianza consiste en comprobar que no existen diferencias significativas de muestras aleatorias provenientes de la misma población, en las cuales se considera simultáneamente todos los datos y se contrasta una hipótesis nula general que consiste en afirmar que no existe diferencia alguna entre las medias de los distintos grupos. En el tipo mas sencillo de análisis de la varianza se consideran dos tipos de variación, la varianza entre grupos que se refiere a la dispersión de las medias de los grupos respecto a la media total, la varianza dentro de los grupos que se refiere a la dispersión de las puntuaciones en cada grupo particular.

Con base en los conceptos expuestos en los Capítulos II y III, el análisis de varianza se deben satisfacer las hipótesis siguientes:

1. Los elementos de los diversos grupos se suponen elegidos por muestreo aleatorio de poblaciones de distribución normal.
2. La varianza de los grupos ha de ser homogénea $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$
3. las muestras que constituyen los subgrupos han de ser independientes. En estas condiciones, las estimaciones son independientes de la varianza y la razón de las varianzas entre y dentro de los grupos tienen una distribución F de Snedecor.

El procedimiento general consiste en realizar una tabla como se muestra a continuación:

Origen de variación	Grados de libertad GI	Suma de cuadrados	Cuadrados medios
<i>Entre grupos</i>	numero de grupos (k) - 1	$\Sigma X^2_e = [\Sigma(\Sigma X)^2/n] - (\Sigma X_T)^2/N$	ΣX^2_e
<i>Dentro de grupos</i>	Suma del numero de casos dentro de cada subgrupo (n) -1 $(n_1-1)+(n_2-1)+\dots+(n_k-1)$	$\Sigma X^2 = \Sigma X^2 - (\Sigma X)^2/n$	$\Sigma X^2/\Sigma(n_k -1)$
<i>Cuadrados Totales</i>	$(k-1) + \Sigma(n_k -1)$	$\Sigma X^2_e + \Sigma X^2$	

El contraste de la diferencia de las dos varianzas se puede realizar mediante el contraste de la F de Snedecor, evaluando de la tabla de análisis de varianza el siguiente contraste de la F:

$$F = \text{cuadrado medio entre grupos} / \text{cuadrado medio dentro de los grupos}$$

Los valores de F se interpretan mediante la tabla de distribución F, si el valor calculado es superior al valor de tablas se rechaza la hipótesis nula, es decir, existe diferencia significativa de las medias al nivel 5 por cien. A veces el valor de la F es menor que 1, en tal caso no tiene objeto continuar puesto que estas razones no son significativas.

Regresión lineal

Se hace referencia al tema en el Capítulo IV, sección 4.2, sub sección 4.2.3

Coefficiente de desigualdad de Theil

Se hace referencia al tema en el Capítulo IV, sección 4.2, sub sección 4.2.3

Prueba Chi Cuadrada

Cuando alguno de los elementos del sistema se comportan estocásticamente, una prueba frecuente se aplica en el estudio de simulación, de compatibilidad de un conjunto de frecuencias observadas con alguna frecuencia teórica. La prueba Chi cuadrada sirve para realizar tal comparación, utilizando el estadístico de prueba:

$$\chi^2 = \sum_{i=1}^n (O_i - E_i)^2 / E_i$$

donde O_i es el valor observado en la clase i -ésima, E_i es el valor que se espera de la distribución de densidad en dicha clase y n es el número de clases (no el tamaño de la muestra denotado por N).

Bondad de Ajuste (R^2)

Para determinar que tanto se ajustan los datos simulados en cada corrida a los datos reales se utiliza el indicador R^2 . El cálculo se realiza obteniendo los conceptos SCT, SEC y suma de los cuadrados del modelo de simulación SCMS, este concepto es análogo a SCR. Esta medida indica el porcentaje de ajuste del modelo simulado respecto a los datos reales.

5.3.3 Interpretación de los resultados

Se realizó un total de 108 corridas, un total aproximado a 1,080,000 registros resultantes tras aplicar el modelo de simulación propuesto, los resultados se exponen en el **Anexo 3**. Se generó un total de 209 resultados de salida, es decir, se agruparon los registros de salida para obtener una base mas compacta y aplicar las técnicas de validación, comparándose con los datos del sistema real que contiene 209 registro históricos y se intenta revelar el grado en que se ajustan haciendo uso de la pruebas anteriormente mencionadas, así como de las métricas de error.

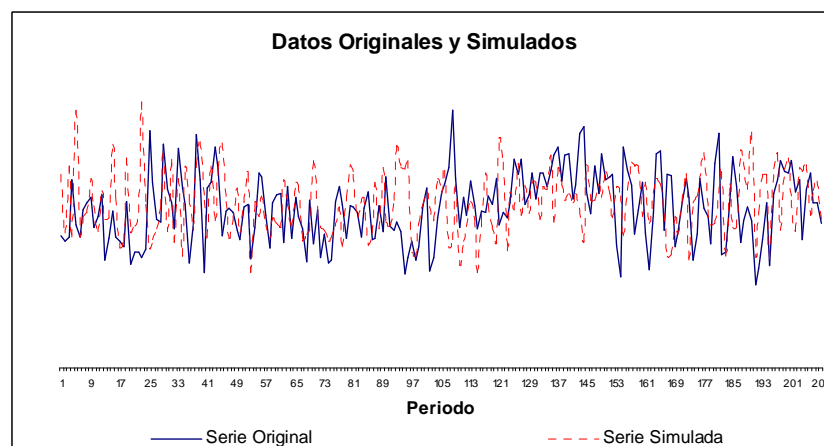
Tabla 1

Métricas de Validación y Bondad de Ajuste

Límites	R ²	X ²	MAD	RMSE	MAPE	MCO	THEIL	ANOVA F _p
Max	0.36	124.63	731.3	878.0	0.55	0.243	0.156	8.56
Min	0.25	9.63	536.7	688.6	0.46	-0.200	0.124	0.03
Aceptar	N A	6	N A	N A	N A	N A	108	63
Rechazar	N A	101	N A	N A	N A	N A	0	45

El grado de ajuste con la prueba R² es pobre y se imputa a la enorme variación de los datos originales

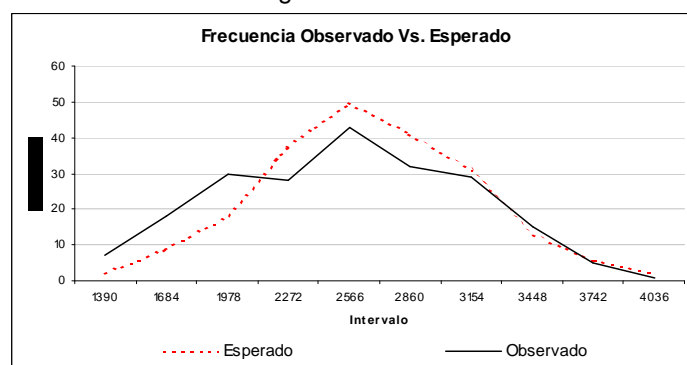
Figura 1.



En lo que respecta a la prueba Ji Cuadrada, de las 108 corridas generadas con el modelo de simulación se acepta únicamente un 6 ocasiones al realizar el contraste de la fdp de los datos generados por el modelo de simulación y la fdp observada, esperando una frecuencia similar para cada intervalo. El valor crítico de tablas aplicable en la prueba es de 15.51, superado en 101 ocasiones al realizar la prueba y se crea error en una ocasión.

La falla al aplicar la prueba se atribuye nuevamente a la variación de la serie original, de hecho se observó que en numerosas ocasiones se subestimó el valor de los dos primeros intervalos al comparar la frecuencia con la que se presentan, causado por la reducida frecuencia en los intervalos de la serie original, por ende al generar las variables aleatorias se subestiman los valores con respecto a los originales, en este caso se recomienda reducir el horizonte de tiempo.

Figura 2.



Al aplicar la prueba MCO, debiendo obtenerse en teoría un valor de la pendiente cercano a uno e intercepto cercano a cero. Con base en lo anterior, se infiere que los datos generados por el modelo no se ajustan a la serie real. No obstante, gráficamente se comprueba que se mueven en la misma dirección y oscilaciones muy semejantes a los datos reales (ver figura 1).

En cuanto al coeficiente de desigualdad Theil se confirma la obtención de buenas predicciones, en virtud de que en las 108 corridas generadas el valor máximo obtenido es 0.156, muy cercano al valor cero; por lo que se concluye que el modelo de simulación genera valores muy cercanos a los valores reales.

Para finalizar con los comentarios en relación a las pruebas de bondad de ajuste, la prueba ANOVA² es significativa en más del 58% de las veces. Al realizar el contraste de la diferencia de medias entre el valor calculado, denotado F_p y el valor de tablas F_T equivalente a 3.86, se espera $F_p < F_T$; en caso contrario, se rechaza la hipótesis nula H_0 de que no existe diferencia significativa de medias.

En términos de métricas de error, se presentan diferencias que podrían inducir a tomar la decisión de rechazar el modelo de simulación, sin embargo debe tomarse en cuenta la variación que experimentan los datos originales y las pruebas de bondad de ajuste que indican un buen grado de “compatibilidad” entre la realidad y los resultados de la simulación. Así mismo, tómesese en cuenta que la simulación es una herramienta de análisis para diseñar, evaluar y predecir el comportamiento de procesos y sistemas con alto grado de complejidad con grado de error, que el decisor habrá de considerar al dar a conocer su determinación de aprobar o rechazar los resultados del modelo, o en su defecto, solicitar una revisión del modelo en su totalidad.

En suma, se considera que el modelo de simulación cumple favorablemente con las pruebas de bondad de ajuste y métricas de error, por consiguiente se considera estadísticamente significativo y se ajusta satisfactoriamente con los datos del sistema real. Se presume que los resultados de salida pueden explicar el comportamiento presente y futuro con buen grado de certeza.

Adicionalmente, la comparación de las diversas medidas de certeza aplicadas en esta investigación para aceptar o rechazar el modelo ya sea de simulación o de pronóstico ejemplifica que cualquiera que se aplique independientemente arrojará un resultado positivo o negativo, pero en conjunto proporcionan amplio espectro de posibilidades para tomar una decisión.

² Analysis of Variance

Conclusiones

Debido a la enorme variación que experimentan los datos en la serie histórica original su manipulación requirió un esfuerzo adicional en la obtención de muestras de varianza mínima para obtener un modelo que se ajustara favorablemente, no obstante se advierte que los resultados obtenidos en el pronóstico se desempeñan mejor en el corto plazo por lo antes mencionado, pues el valor de los intervalos de confianza a periodos de mediano y largo plazo para pronosticar su comportamiento se extienden excesivamente, mas allá de 2 desviaciones estándar.

Los datos de salida fueron obtenidos mediante el uso de la hoja de cálculo Excel y el motivo por el cual se utilizó esta herramienta tuvo 2 atribuciones principales; uno debido a que es un instrumento de fácil acceso, bajo costo y no requiere de equipos de alta tecnología, segundo impulso por la bondad de su manejo y cómoda maniobrabilidad en la ejecución de fórmulas y comandos. Adicionalmente, se pueden aplicar modelos complejos de investigación de operaciones, esto dependiendo de la habilidad, conocimientos y experiencia del analista.

En lo que respecta al resultado obtenido al aplicar el modelo de simulación se considera que la serie simulada se ajusta satisfactoriamente a la serie histórica original, ya que así lo muestran las medidas de validación empleadas. A pesar de los múltiples ajustes al modelo de simulación, las métricas utilizadas para calibrar el modelo resultan estadísticamente significativas cuando se sobrevalúan los datos de la serie histórica original que cuando se obtienen valores inferiores, lo anterior se atribuye a la no estacionalidad de la serie advirtiéndose una ligera tendencia en sentido ascendente, misma que se integro en el modelo de simulación al añadir el porcentaje de crecimiento semanal.

Aunque el modelo de simulación es una representación teórica de la realidad bajo condiciones controladas para acceder a la experimentación, sin embargo la variación de información conduce invariablemente a resultados con un grado representativamente bajo

de certidumbre, lo que conduce a pensar que la aplicación de un modelo de pronóstico en estas condiciones no es lo más recomendable, sino una revisión completa de los procesos y procedimientos aplicando el enfoque sistémico para la gestión actual del CAT.

El primer contacto en el CAT utilizando herramientas estadísticas para determinar la tendencia de la operación, en el presente estudio el caso de las cuentas canceladas, sugiere una opción alternativa en la gestión del CAT, además del control de la operación en el sentido estrictamente empírico basado en la experiencia. La ventaja de iniciarse en un análisis más detallado confiere la posibilidad de integrar un mayor número de variables para generar escenarios factibles, mostrando las trayectorias probables sobre las que pudiera continuar el flujo de la operación en condiciones normales o extraordinarias, según sea el caso.

En contraste con el conocimiento empírico y la experiencia en las que no hay cabida para la experimentación y disposición de acciones dispuestas con antelación para el caso de eventos extraordinarios durante la operación, y la resistencia a modelos estadísticos que consienten una amplia gama de aplicaciones para examinar distintas situaciones, determinan un área de oportunidad para explorar.

El ambiente relacionado con modelos cuantitativos como en este trabajo se hizo presente, requieren además de un amplio conocimiento y experiencia del enfoque sistémico, es decir, obtener un pronóstico para una serie de tiempo puede requerir de unos momentos para analizar los datos con las herramientas con las que cuenta la hoja de cálculo y obtener un resultado el cual se considera satisfactorio.

No obstante, abordar un problema según sean los requerimientos de una Institución o Empresa de invertir mayor o menor cantidad de tiempo para efectuar una evaluación y/o aplicar una reingeniería de procesos a los procedimientos para obtener conclusiones exactas o adecuadas acerca del funcionamiento, todo estriba en la importancia relativa que se imponga por parte de los Funcionarios o Directivos.

Lo referente al uso de herramientas de análisis cuantitativo y su precisión en numerosas ocasiones resulta indiferente o relativo, es decir, hay que valerse de lo que se tiene a

mano ya que no siempre se tendrán las mejores condiciones o las mas optimas para llevar a cabo una evaluación al igual que la colección de datos en simulación, a efecto de reducir tales carencias la experiencia y los conocimientos con los que cuenta un analista se convertirán en las variables mas valiosas para lograr el éxito y satisfacción profesional que un trabajo llegue a su fin, no obstante las insuficiencias en la información e instrumentos de análisis.

Bibliografía

- [1] Downie, N. M. y Heath, R. W. Métodos Estadísticos Aplicados. México, D. F.: Harla, 1973. Ostle, Bernard. Estadística Aplicada. México, D. F.: Limusa, 1986.
- [2] Canavos, George C. Probabilidad y Estadística, Aplicaciones y Métodos. México, D. F.: Mc Graw Hill, 1988
- [3] Evans, James R. & Olson, David L.. Introduction to Simulation and Risk Analysis. New Jersey: Prentice Hall, 2002
- [4] Naylor, Thomas H. Experimentos de Simulación en computadoras con modelos de sistemas económicos. México, D. F.: 1982.
- [5] Senge, Peter. La Quinta Disciplina. México, D. F.: Ediciones Granica, 1998.
- [6] Gujarati, Damodar N. Econometría. México, D. F.: Mc Graw Hill, 2004.
- [7] Pindyck, Robert S. & Rubinfeld, Daniel L. Econometric Models ad Economic Forecast. Singapore: Mc Graw Hill international, 1981.
- [8] Jensen Raul A., Bard F. Jonathan. Operations Research, models and Methods. USA: John Wiley & Sons, Inc., 2003.
- [9] Sasieni M., Yaspan A., Friedman L. Investigación de Operaciones, Métodos y Problemas. México, D. F.: Limusa, 1982.
- [10] Winston Wayne L., Albrtight Christian S. Practical Management Science, Spreadsheet modeling and applications. California: Duxbury Press, 1997.
- [11] Davis, K. Roscoe, McKeown, Patrick G. modelos Cuantitativos para Administración. México, D. F.: Grupo Editorial Iberoamérica, 1986.

[12] Gallagher, Charles A., Watson Hugh J. Métodos Cuantitativos para la Toma de decisiones en Administración. México, D. F.: Mc Graw Hill, 1982.

Consultas en Internet

[13] Arellano, M. (2001): "Introducción al Análisis Clásico de Series de Tiempo", [en línea] 5campus.com, Estadística. Recuperado el 15 de noviembre del 2006. <<http://www.5campus.com/leccion/seriest>>

[14] Biz/ed time web: "Examples of Analysing Time Series Data [TimeWeb]", [en línea] Time Series. Recuperado el 15 de noviembre del 2006. <http://www.bized.ac.uk/timeweb/crunching/crunch_analysis_illus.htm>

[15] Biz/ed time web: "Background to Analysing Time Series [TimeWeb]", [en línea] Time Series. Recuperado el 15 de noviembre del 2006. <http://www.bized.ac.uk/timeweb/crunching/crunch_analysis_expl.htm>

[16] Universidad Autónoma de Madrid: "Curso de predicción económica y empresarial edición 2004", [en línea]. Fondo de Investigación e Innovación Richard Stone del Instituto L.R. Klein-Centro Stone. Recuperado el 15 de Noviembre de 2006. <<http://www.uam.es/docencia/predysim/predysim/descargas>>

[17] lab TA Joy Woller: "The Basics of Monte Carlo Simulations", [en línea]. University of Nebraska-Lincoln Physical Chemistry Lab (Chem 484). Recuperado el 15 de noviembre del 2006. <<http://www.chem.unl.edu/zeng/joy/mclab/mcintro.html> 2005>

[18] Instituto Mexicano de Telemarketing: "Simulación: factor clave en la toma de decisiones de un Call Centres". Centros Telefónicos, [en línea]. Recuperado el 15 de noviembre del 2006. <<http://www.imt.com.mx/recontact/11/simulacion.php>>

[19] DUKE UNIVERSITY: "Introduction to forecasting. Forecasting". Decision 411 Course Outline 2005, [en línea]. Recuperado el 15 de noviembre del 2006. <<http://www.duke.edu/~rnau/411home.htm...>>

[20] EINSTEINNET: "Análisis Clásico de Series Temporales 1". Series de Tiempo, [en línea]. Recuperado el 15 de noviembre del 2006.

<<http://www.einsteinnet.com/econometria/ARIMA/arima2.htm>>

Otras referencias

[21] Rafael E. Rivera, Michael Bernald y Rolando Salas: "Antología Simulación Industrial". Manuscrito no publicado. Universidad Fidélitas IIC-2002. Febrero 2002

[22] Instituto Mexicano de Telemarketing A. C. (13 y 14 de Marzo de 2001), [5º. Congreso Internacional, centros Telefónicos / Centros de Contacto con Clientes]. México, D.F.

[22] UNAM (Mayo 2001), [Propedeutico de Probabilidad y Estadística]. División de estudio de posgrado. México, D.F., Junio 2001

[23] Weiss, Carol H. Investigación Evaluativa, Métodos para determinar la eficiencia de los programas de acción , México D.F., Editorial Trillas 1978.

Anexo 1

Generador para variables uniformes

DIA	Día_Sem					
MAX	1243					
MIN	107					
INTERVALO	=ENTERO((CL4-CL5)/10)					
VALOR	=CQ20					
ALEATORIO					=ALEATORIO()	
	INTERVALO	FRECUENCIA	FRECUENCIA RELATIVA	FRECUENCIA ACUMULADA		
	107	220	31	=CN10/\$CN\$20	=CO10	=SI(Y(CP9<\$CQ\$8,\$CQ\$8<CP10),(CL10+ALEATORIO()*(CM10-CL10)),0)
	221	334	17	=CN11/\$CN\$20	=CO11+CP10	=SI(Y(CP10<\$CQ\$8,\$CQ\$8<CP11),(CL11+ALEATORIO()*(CM11-CL11)),0)
	335	448	50	=CN12/\$CN\$20	=CO12+CP11	=SI(Y(CP11<\$CQ\$8,\$CQ\$8<CP12),(CL12+ALEATORIO()*(CM12-CL12)),0)
	449	562	53	=CN13/\$CN\$20	=CO13+CP12	=SI(Y(CP12<\$CQ\$8,\$CQ\$8<CP13),(CL13+ALEATORIO()*(CM13-CL13)),0)
	563	676	28	=CN14/\$CN\$20	=CO14+CP13	=SI(Y(CP13<\$CQ\$8,\$CQ\$8<CP14),(CL14+ALEATORIO()*(CM14-CL14)),0)
	677	790	9	=CN15/\$CN\$20	=CO15+CP14	=SI(Y(CP14<\$CQ\$8,\$CQ\$8<CP15),(CL15+ALEATORIO()*(CM15-CL15)),0)
	791	904	11	=CN16/\$CN\$20	=CO16+CP15	=SI(Y(CP15<\$CQ\$8,\$CQ\$8<CP16),(CL16+ALEATORIO()*(CM16-CL16)),0)
	905	1018	4	=CN17/\$CN\$20	=CO17+CP16	=SI(Y(CP16<\$CQ\$8,\$CQ\$8<CP17),(CL17+ALEATORIO()*(CM17-CL17)),0)
	1019	1132	3	=CN18/\$CN\$20	=CO18+CP17	=SI(Y(CP17<\$CQ\$8,\$CQ\$8<CP18),(CL18+ALEATORIO()*(CM18-CL18)),0)
	1133	1243	3	=CN19/\$CN\$20	=CO19+CP18	=SI(Y(CP18<\$CQ\$8,\$CQ\$8<CP19),(CL19+ALEATORIO()*(CM19-CL19)),0)
		209				=SUMA(CQ10:CQ19)

Anexo 2A

Cálculo para la Validación del Modelo de Simulación

Cuadrado Serie Simulada	Serie Original	Serie Simulada	Error de Simulación	Error Desfazado	Valor Absoluto
=Y221^2	2206	2988	=X221-Y221		=ABS(Z221)
=Y222^2	2111	2895	=X222-Y222	=Z221	=ABS(Z222)
=Y223^2	2187	1845	=X223-Y223	=Z222	=ABS(Z223)
=Y224^2	3137	2917	=X224-Y224	=Z223	=ABS(Z224)
=Y225^2	2392	3247	=X225-Y225	=Z224	=ABS(Z225)
=Y226^2	2172	2178	=X226-Y226	=Z225	=ABS(Z226)
=Y227^2	2646	1823	=X227-Y227	=Z226	=ABS(Z227)
=SUMA(V221:V429)		=PROMEDIO(Y221:Y429)	=CONTAR.SI(Z221:Z429)		=PROMEDIO(AB221:AB429)
		=DESVEST(Y221:Y429)	=CONTAR.SI(Z221:Z429)		
+		=ESTIMACION.LINEAL(Y221:Y429,W221:W429,1,0)	=ESTIMACION.LINEAL(Y221:Y429,W221:W429,1,0)		
		=DECIMAL(Y432,2,1)	=DECIMAL(Z432,2,1)		
	=MAX(X221:X429)	=MAX(Y221:Y429)			
	=MIN(X221:X429)	=MIN(Y221:Y429)			
	=ENTERO((X436-X437)/10)	=ENTERO((Y436-Y437)/10)			
=X437+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Y437	=Y440+Y438	esp	obs
=U441+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z440+1	=Y441+\$Y\$438	=X440	=W440
=U442+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z441+1	=Y442+\$Y\$438	=X441	=W441
=U443+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z442+1	=Y443+\$Y\$438	=X442	=W442
=U444+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z443+1	=Y444+\$Y\$438	=X443	=W443
=U445+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z444+1	=Y445+\$Y\$438	=X444	=W444
=U446+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z445+1	=Y446+\$Y\$438	=X445	=W445
=U447+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z446+1	=Y447+\$Y\$438	=X446	=W446
=U448+\$X\$438	=FRECUENCIA(Y221:Y429,V440:V449)	=Z447+1	=Y448+\$Y\$438	=X447	=W447
=Y436	=FRECUENCIA(Y221:Y429,V440:V449)	=Z448+1	=Y449+\$Y\$438	=X448	=W448
				=X449	=W449
	=ESTIMACION.LINEAL(X221:X429,Y221:Y429,1,0)	=ESTIMACION.LINEAL(X221:X429,Y221:Y429,1,0)			
	=DECIMAL(X451,3)	=DECIMAL(Y451,1)			
sumas/n^2	1	1			
sumas/n	=(SUMA(Y221:Y429)+SUMA(X221:X429))^2/(W429^2)	=(SUMA(Y221:Y429)+SUMA(X221:X429))^2/(W429^2)			
entre gps	=(SUMA(X221:X429))^2/W429	=(SUMA(X221:X429))^2/W429			
cdrds tts	=(Y455+Z455)-Y454	=(Y455+Z455)-Y454			
	=(U430+V430)-Y454	=(U430+V430)-Y454			
gl ntre gps	1	1			
gl dentro gps	=(W429-1)+(W429-1)	=(W429-1)+(W429-1)			
			=Y456	=Z459/Y459	=AA459/AA460
			=AC456	=Z460/Y460	
			=SUMA(Z459:Z460)		

Anexo 2B

Cálculo para la Validación del Modelo de Simulación II

Cuadrado Error de Simulación

=Z221^2	=AB221/X221	=(Y221-PROMEDIO(\$X\$220:\$X\$429))^2	=Y221^2	=X221^2	=CONTAR.SI(A18:A1216,">0")	=PROMEDIO(A18:A1216)
=Z222^2	=AB222/X222	=(Y222-PROMEDIO(\$X\$220:\$X\$429))^2	=Y222^2	=X222^2		
=Z223^2	=AB223/X223	=(Y223-PROMEDIO(\$X\$220:\$X\$429))^2	=Y223^2	=X223^2		
=Z224^2	=AB224/X224	=(Y224-PROMEDIO(\$X\$220:\$X\$429))^2	=Y224^2	=X224^2		
=Z225^2	=AB225/X225	=(Y225-PROMEDIO(\$X\$220:\$X\$429))^2	=Y225^2	=X225^2		
=Z226^2	=AB226/X226	=(Y226-PROMEDIO(\$X\$220:\$X\$429))^2	=Y226^2	=X226^2		
=Z227^2	=AB227/X227	=(Y227-PROMEDIO(\$X\$220:\$X\$429))^2	=Y227^2	=X227^2		
=RAIZ(PROMEDIO(AC221:AC429))	=RAIZ(PROME	=CONCATENAR(X452,V432,Y452)	=RAIZ((1/2(=RAIZ((1/2	=RAIZ((1/209)*SUMA(AC221:/	
=SUMA(AC221:AC429)		=SUMA(AE221:AE431)		=AH430/(A		
=AE432/(AC432+AE432)						

=SI(AA440<5,0,((AB440-AA440)^2)/AA440)
 =SI(AC440=0,(((AB440+AB441)-(AA440+AA441))^2)/AA440,(((AB440+AB441)-(AA440+AA441))^2)/AA440)
 =((AB442-AA442)^2)/AA442
 =((AB443-AA443)^2)/AA443
 =((AB444-AA444)^2)/AA444
 =((AB445-AA445)^2)/AA445
 =((AB446-AA446)^2)/AA446
 =((AB447-AA447)^2)/AA447
 =(((AB448+AB449)-(AA448+AA449))^2)/(AA448+AA449)
 =SUMA(AC440:AC449)

=Y457-Y456

F tab
3.86

Anexo 2C

Programación Visual Basic para obtener los resultados de validación

```
1 Generador de valores simulados
2 Sub SIMULAR()
3 ' COPIARXDIA Macro
4 ' Macro grabada el 22/12/2005 por RET168
5 Range("A8:I\2000").Select Determinar el rango de entrada de los datos
6 Selection.ClearContents Limpiar el área de trabajo
7 For K = 8 To 216 Rango de la variable
8 Row = K Asignar valor a variable artificial
9 Cells(6, 1) = K - 7 Determinar rango de salida
10 Cells(Row, 1) = K - 7 Determinar rango de salida
11 Range("C6:I6,K6:Q6,S6:Y6,AA6:AG6,AI6,AM6:AS6,AU6,AW6:BB6,BD6:BI6,BQ6:BT6"). _Select Determinar rango de salida

13 Selection.Copy Copiar selección
14 Cells(Row, 3).Select Determinar rango de salida
15 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False Pegar selección

17 Next
18 Range("J8").Select Seleccionar rango de salida
19 Range(Selection, Selection.End(xlDown)).Select Seleccionar hasta el último registro
20 Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha
21 'Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha

22 Range("R8").Select Seleccionar rango de salida
23 Range(Selection, Selection.End(xlDown)).Select Seleccionar hasta el último registro
24 Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha
25 'Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha

26 Range("Z8").Select Seleccionar rango de salida
27 Range(Selection, Selection.End(xlDown)).Select Seleccionar hasta el último registro
28 Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha
29 'Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha

30 Range("AH8").Select Seleccionar rango de salida
31 Range(Selection, Selection.End(xlDown)).Select Seleccionar hasta el último registro
32 Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha
33 'Selection.INSERT Shift:=xlToRight Insertar celdas a la derecha
```

```

34 Range("AJ8:AL8").Select      Seleccionar rango de salida
35 Range(Selection, Selection.End(xlDown)).Select  Seleccionar hasta el último registro
36 Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha
37 'Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha

38 Range("AT8").Select      Seleccionar rango de salida
39 Range(Selection, Selection.End(xlDown)).Select  Seleccionar hasta el último registro
40 Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha
41 'Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha

42 Range("AV8").Select      Seleccionar rango de salida
43 Range(Selection, Selection.End(xlDown)).Select  Seleccionar hasta el último registro
44 Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha
45 'Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha

46 Range("BC8").Select      Seleccionar rango de salida
47 Range(Selection, Selection.End(xlDown)).Select  Seleccionar hasta el último registro
48 Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha
49 'Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha

50 Range("BJ8:BP8").Select  Seleccionar rango de salida
51 Range(Selection, Selection.End(xlDown)).Select  Seleccionar hasta el último registro
52 Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha
53 'Selection.INSERT Shift:=xlToRight  Insertar celdas a la derecha

54 Range("AG1").Select      Seleccionar rango
55 End Sub  Fin de la instrucción

56 Calculo de las métricas de error bondad de ajuste, medidas de dispersión y ANOVA
57 Sub ESTADISTICAS()
58 '
59 ' ESTADISTICAS Macro
60 ' Macro grabada el 14/03/2006 por RET168
61 Sheets("Hoja2").Select  Seleccionar hoja de trabajo
62 Range("U219:AH459").Select  Seleccionar rango
63 Selection.Copy  Acción copiar
64 ActiveSheet.Next.Select  Continuar siguiente selección
65 Selection.End(xlDown).Select  Seleccionar hasta el último registro
66 Range("U220").Select  Rango de salida
67 ActiveSheet.Paste  Acción copiar
68 Range("R216").Select  Rango de salida
69 Selection.End(xlUp).Select  Seleccionar hasta el último registro
70 Range("Y8").Select  Rango de salida
71 Range(Selection, Selection.End(xlDown)).Select  Seleccionar hasta el último registro
72 Application.CutCopyMode = False  Liberar acción copiado
73 Selection.Copy  Acción copiar
74 Selection.End(xlDown).Select  Seleccionar hasta el último registro

```

75 Range("Y220").Select Rango de salida
76 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False Pegar selección
78 Selection.End(xlDown).Select Seleccionar hasta el último registro
79 Range("Y447").Select Rango de salida

82 Range("AI220").Select Seleccionar rango
83 ActiveCell.FormulaR1C1 = "=COUNTIF(R[-212]C:R[-4]C,"">0"")" Insertar fórmula CONTAR.SI, para determinar llamadas atendidas
84 Range("BB220").Select Rango de salida
85 ActiveCell.FormulaR1C1 = "=COUNTIF(R[-212]C:R[-4]C,"">0"")" Insertar fórmula CONTAR.SI, para determinar carga de trabajo
86 Range("BI220").Select Seleccionar rango
87 ActiveCell.FormulaR1C1 = "=COUNTIF(R[-212]C:R[-4]C,"">0"")" Insertar fórmula CONTAR.SI, para determinar carga de trabajo por incremento en la plantilla

88 Range("AV8").Select Seleccionar rango
89 ActiveCell.FormulaR1C1 = "=1-(RC[3]/RC[-1])" Insertar fórmula para determinar porcentaje de abandono por incremento en la plantilla
90 Range("AV8").Select Seleccionar rango
91 Selection.AutoFill Destination:=Range("AV8:AV216") Insertar fórmula para determinar porcentaje de abandono por incremento en la plantilla al rango seleccionado

92 Range("BC8").Select Seleccionar rango
93 ActiveCell.FormulaR1C1 = "=1-(RC[-8]/RC[3])" Insertar fórmula para determinar porcentaje de abandono por decremento en la plantilla
94 Range("BC8").Select Seleccionar rango
95 Selection.AutoFill Destination:=Range("BC8:BC216") Insertar fórmula para determinar porcentaje de abandono por decremento en la plantilla al rango seleccionado

96 Range("BK8").Select Seleccionar rango
97 ActiveCell.FormulaR1C1 = "=1-(RC[-54]/RC[-10])" Fórmula para determinar incremento porcentual en la plantilla por cancelación de cuenta
98 Range("BK8").Select Seleccionar rango de salida
99 Selection.Copy Acción copiar rango
100 Range("BK9:BK216").Select Seleccionar rango de salida
101 ActiveSheet.Paste Acción pegar
102 Application.CutCopyMode = False Liberar acción copiado

104 Range("BL8").Select Seleccionar rango
105 ActiveCell.FormulaR1C1 = "=1-(RC[-47]/RC[-12])" Fórmula para determinar incremento en la plantilla por no cancelar activa
106 Range("BL8").Select Seleccionar rango
107 Selection.AutoFill Destination:=Range("BL8:BL216") Aplicar al rango seleccionado

108 Range("BN8").Select Seleccionar rango
109 ActiveCell.FormulaR1C1 = "=1-(RC[-6]/RC[-57])" Fórmula para determinar decremento porcentual en la plantilla por cancelación de cuenta
110 Range("BN8").Select Seleccionar rango
111 Selection.Copy Acción copiar
112 Range("BN9:BN216").Select Seleccionar rango
113 ActiveSheet.Paste Acción pegar
114 Application.CutCopyMode = False Liberar acción copiado

116	Range("BO8").Select	Seleccionar rango	
117	ActiveCell.FormulaR1C1 = "=1-(RC[-8]/RC[-50])"		Fórmula para determinar decremento porcentual en la plantilla por no cancelar de cuenta
118	Range("BO8").Select	Seleccionar rango	
119	Selection.AutoFill Destination:=Range("BO8:BO216")		Aplicar al rango seleccionado
121	Range("AU220").Select	Seleccionar rango de salida	
122	ActiveCell.FormulaR1C1 = "=SUM(R[-212]C:R[-1]C)"		Fórmula para determinar el total acumulado del abandono
123	Range("AY220").Select	Seleccionar rango de salida	
124	ActiveCell.FormulaR1C1 = "=SUM(R[-212]C:R[-1]C)"		Fórmula para determinar el porcentaje global del abandono
125	Range("BF220").Select	Seleccionar rango de salida	
126	ActiveCell.FormulaR1C1 = "=SUM(R[-212]C:R[-1]C)"		Fórmula para determinar el abandono por decremento en la plantilla
127	Range("AV220").Select	Seleccionar rango de salida	
128	ActiveCell.FormulaR1C1 = "=1-(RC[3]/RC[-1])"		Fórmula para la variación porcentual del abandono por incremento en la plantilla
130	Range("BC220").Select	Seleccionar rango de entrada	
131	ActiveCell.FormulaR1C1 = "=1-(RC[-8]/RC[3])"		Cálculo variación porcentual de la plantilla
133	Range("BK220").Select	Seleccionar rango de entrada	
134	ActiveCell.FormulaR1C1 = "=AVERAGE(R[-212]C:R[-4]C)"		Incremento de la Plantilla (Promedio Canceladas)
135	Range("BL220").Select	Seleccionar rango de entrada	
136	ActiveCell.FormulaR1C1 = "=AVERAGE(R[-212]C:R[-4]C)"		Incremento de la Plantilla (Promedio No Canceladas)
137	Range("BN220").Select	Seleccionar rango de entrada	
138	ActiveCell.FormulaR1C1 = "=AVERAGE(R[-212]C:R[-4]C)"		Decremento de la Plantilla (Promedio Canceladas)
139	Range("BO220").Select	Seleccionar rango de entrada	
140	ActiveCell.FormulaR1C1 = "=AVERAGE(R[-212]C:R[-4]C)"		Decremento de la Plantilla (Promedio No Canceladas)
141	Range("BT220").Select	Seleccionar rango de entrada	
142	ActiveCell.FormulaR1C1 = "=AVERAGE(R[-212]C:R[-4]C)"		Variación porcentual de movimiento en la base de Clientes
144	Range("A18").Select	Seleccionar rango de entrada	
145	Range(Selection, Selection.End(xlDown)).Select	Seleccionar rango de entrada	
146	Selection.Replace What:="0", Replacement:="", LookAt:=xlPart, _	SearchOrder:=xlByRows, MatchCase:=False	Reemplazar fórmula por valores
149	Range("BB8").Select	Seleccionar rango de entrada	
150	Range(Selection, Selection.End(xlDown)).Select	Seleccionar rango de entrada	
151	Selection.Replace What:="0", Replacement:="", LookAt:=xlPart, _	SearchOrder:=xlByRows, MatchCase:=False	Reemplazar fórmula por valores
153	Range("AJ220").Select	Seleccionar rango de entrada	
154	ActiveCell.FormulaR1C1 = "=AVERAGE(R[-212]C[-1]:R[-4]C[-1])"		Promedio de la carga de trabajo
155	Range("BA220").Select	Seleccionar rango de entrada	
156	ActiveCell.FormulaR1C1 = "=AVERAGE(R[-212]C[1]:R[-4]C[1])"		Promedio de Cancelados
157	Range("AG1").Select	Rango final	


```

160 End Sub Fin

161 Reporteador de estadísticas
162 Sub RESSTATS()
163 ' RESSTATS Macro
164 ' Macro grabada el 15/03/2006 por RET168
165 Sheets("VALIDACION").Select Seleccionar Hoja
166 Rows("4:4").Select Selección de fila
167 Selection.Insert Shift:=xlDown Insertar una fila

169 'METRICAS PRONOSTICO Comandos para cálculo de indicadores por bloque
170 Sheets("SIMULACION").Select Seleccionar Hoja
171 Range("AB429:AD429").Select Seleccionar rango
172 Selection.Copy Acción copiar
173 Sheets("VALIDACION").Select Seleccionar Hoja
174 Range("D4").Select Seleccionar rango
175 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False Acción pegar

178 'MCO Comandos para cálculo de indicadores por bloque
179 Sheets("SIMULACION").Select Seleccionar Hoja
180 Range("AE429").Select Seleccionar rango
181 Selection.Copy Acción copiar
182 Sheets("VALIDACION").Select Seleccionar Hoja
183 Range("G4").Select Seleccionar rango
184 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False Acción pegar

187 ' JI CUADRADA Comandos para cálculo de indicadores por bloque
188 Sheets("SIMULACION").Select Seleccionar Hoja
189 Range("AC449").Select Seleccionar rango
190 Selection.Copy Acción copiar
191 Sheets("VALIDACION").Select Seleccionar Hoja
192 Range("C4").Select Seleccionar rango
193 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False Acción pegar

196 'R^2 Comandos para cálculo de indicadores por bloque
197 Sheets("SIMULACION").Select Seleccionar Hoja
198 Range("AC433").Select Seleccionar rango
199 Selection.Copy Acción copiar
200 Sheets("VALIDACION").Select Seleccionar Hoja
201 Range("B4").Select Seleccionar rango
202 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False Acción pegar

204 ' THEILComandos para cálculo de indicadores por bloque
205 Sheets("SIMULACION").Select Seleccionar Hoja

```

206	Range("AG431").Select	Seleccionar rango	
207	Selection.Copy	Acción copiar	
208	Sheets("VALIDACION").Select	Seleccionar Hoja	
209	Range("H4").Select	Seleccionar rango	
210	Selection.PasteSpecial	Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _	False, Transpose:=False Acción pegar
213	' MEDIA	Comandos para cálculo de indicadores por bloque	
214	Sheets("SIMULACION").Select	Seleccionar Hoja	
215	Range("Y429").Select	Seleccionar rango	
216	Selection.Copy	Acción copiar	
217	Sheets("VALIDACION").Select	Seleccionar Hoja	
218	Range("I4").Select	Seleccionar rango	
219	Selection.PasteSpecial	Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _	False, Transpose:=False Acción pegar
222	' DESVEST	Comandos para cálculo de indicadores por bloque	
223	Sheets("SIMULACION").Select	Seleccionar Hoja	
224	Range("Y430").Select	Seleccionar rango	
225	Selection.Copy	Acción copiar	
226	Sheets("VALIDACION").Select	Seleccionar Hoja	
227	Range("J4").Select	Seleccionar rango	
228	Selection.PasteSpecial	Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _	False, Transpose:=False Acción pegar
231	' SUB ESTIMADO	Comandos para cálculo de indicadores por bloque	
232	Sheets("SIMULACION").Select	Seleccionar Hoja	
233	Range("Z429").Select	Seleccionar rango	
234	Selection.Copy	Acción copiar	
235	Sheets("VALIDACION").Select	Seleccionar Hoja	
236	Range("K4").Select	Seleccionar rango	
237	Selection.PasteSpecial	Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _	False, Transpose:=False Acción pegar
240	' SOBRE ESTIMADO	Comandos para cálculo de indicadores por bloque	
241	Sheets("SIMULACION").Select	Seleccionar Hoja	
242	Range("Z430").Select	Seleccionar rango	
243	Selection.Copy	Acción copiar	
244	Sheets("VALIDACION").Select	Seleccionar Hoja	
245	Range("L4").Select	Seleccionar rango	
246	Selection.PasteSpecial	Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _	False, Transpose:=False Acción pegar
249	' F PRUEBA	Comandos para cálculo de indicadores por bloque	
250	Sheets("SIMULACION").Select	Seleccionar Hoja	
251	Range("AB458").Select	Seleccionar rango	
252	Selection.Copy	Acción copiar	
253	Sheets("VALIDACION").Select	Seleccionar Hoja	

```

254 Range("M4").Select      Seleccionar rango
255 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False  Acción pegar

259 ' DEDUCCIONES INCREMENTO Y DECREMENTO PLANTILLA  Comandos para cálculo de indicadores por bloque
260 Sheets("SIMULACION").Select      Seleccionar Hoja
261 Range("AI220,AJ220,AU220:AV220,AY220,BA220:BC220,BF220,BK220:BL220,BN220:BO220,BT220"). _ Select  Seleccionar rango

263 Selection.Copy  Acción copiar
264 Sheets("VALIDACION").Select      Seleccionar Hoja
265 Range("N4").Select      Seleccionar rango
266 Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _ False, Transpose:=False  Acción pegar

269 Sheets("SIMULACION").Select      Seleccionar Hoja
270 Range("AG1").Select      Seleccionar rango

272 End Sub  Fin

```

Anexo 3

Resumen validación del modelo de Simulación

BONDAD DE AJUSTE		METRICAS DE ERROR					MEDIDAS DE DISPERSION		DESVIACIONES		ANOVA
VALOR CRITICO $J_i^2 = 14.07 - 15.51$											F TABLA 3.86
R^2	Ji Cuadrada	MAD	RMSE	MAPE	MCO	THEIL	MEDIA	DESVEST	SUB ESTIMADO	SOBRE ESTIMADO	ANOVA F PRUEBA
0.313573929	24.4315217	634.430622	773.018207	0.50470225	0.016	0.14088471	2697.83732	523.32674	104	105	0.142865406
0.308274777	46.142066	564.555024	706.920825	0.47815494	0.129	0.12810918	2739.45455	468.955282	94	115	1.448534358
0.31404056	38.5031983	593.282297	733.601087	0.49167054	0.086	0.13293092	2735.60766	494.130811	105	104	1.221062312
0.327804093	31.2249194	564.521531	709.952788	0.47922086	0.157	0.1281075	2759.88995	490.048744	104	105	2.470347921
0.308375517	64.8997051	617.669856	760.000085	0.50226242	0.027	0.13710233	2759.2201	502.04839	100	109	2.381091789
0.311247522	34.355809	641.655502	804.58493	0.51268719	-0.03	0.1443829	2783.50239	531.641798	91	118	3.806570685
0.342590396	19.1430272	620.272727	776.369693	0.49506207	0.072	0.14062646	2725.17225	559.755664	101	108	0.734558359
0.335429741	23.2723458	579.813397	722.701964	0.48240094	0.147	0.13012024	2769.42105	506.345575	90	118	2.990051057
0.292709696	53.3876272	600.779904	741.369175	0.49344458	0.025	0.13428218	2741.54545	473.735685	90	119	1.535384807
0.303842632	31.8367725	626.464115	785.114836	0.50657795	-0.02	0.14248763	2722.45455	517.968405	102	106	0.702065214
0.299875582	69.3955524	614.392344	761.123892	0.50149527	0.005	0.1364144	2799.12919	484.190887	88	121	5.434957472
0.302514101	51.7494498	577.4689	717.041213	0.48382905	0.095	0.13066806	2707.80383	472.379212	104	105	0.345308153
0.328310144	13.6784137	567.712919	740.116065	0.47951811	0.1	0.13438091	2720.15311	516.908911	108	101	0.633269888
0.314024107	28.9696429	590.69378	734.607043	0.48583051	0.083	0.13386635	2703.43062	497.540949	107	101	0.242417916
0.288699236	55.785083	601.483254	740.667338	0.49827226	0.018	0.13281291	2800.99522	456.495645	89	120	5.871517176
0.284820127	21.8834331	638.674641	786.407229	0.50799698	-0.07	0.14284665	2721.73684	495.483692	102	107	0.706543692
0.324515899	18.022699	619.937799	774.41417	0.50463287	0.04	0.13876188	2793.13397	525.405266	100	109	4.578142321
0.276004415	68.9845807	637.196172	810.984699	0.51268807	-0.15	0.14582198	2779.57895	491.377112	93	116	3.782400519
0.30123697	13.5713603	655.066986	798.351672	0.50705292	-0.05	0.14555513	2695.44019	525.132078	103	106	0.110920858
0.322565705	44.1995168	615.368421	778.110163	0.49805752	0.029	0.14090864	2731.33971	535.498621	102	106	0.976643813
0.333528753	28.8002579	601.421053	735.889772	0.4884078	0.119	0.13271465	2758.37799	515.48498	93	116	2.279232167
0.295570117	42.7511117	649.736842	812.094259	0.51899413	-0.09	0.14518133	2808.84689	510.583231	94	115	6.058319545
0.336746706	26.1106861	571.38756	700.746492	0.48165322	0.185	0.12741785	2715.33014	499.068345	102	107	0.513809839
0.296289072	46.8107233	630.110048	771.030445	0.507821	-0.02	0.13753838	2826.89474	478.602059	90	119	8.276151425
0.333186301	32.4008507	605.995215	744.429594	0.49271414	0.106	0.13361548	2785.21053	516.298721	97	112	4.035198917
0.289776707	48.6038748	605.483254	750.992759	0.49473789	0	0.13514538	2778.64115	470.0286	90	119	3.850345271
0.272824263	46.0332233	623.665072	755.645521	0.50205128	-0.05	0.13641997	2762.91866	455.979396	91	118	2.812772375

Anexo 4A

Fórmulas para derivar los modelos de pronóstico, Promedios Móviles, tercer rezago

DATOS ORIGINALES			METRICAS DE ERROR			
NO CANCELADOS - CANCELADOS	MA 3	ERROR DE PRONOSTICO	MAD	RMSE	MAPE	SSR
2206						
2111						
2187		=D10-E11	=ABS(F10)	=H10^2	=H10/D10	
3137	=PROMEDIO(D8:D10)	=D11-E12	=ABS(F11)	=H11^2	=H11/D11	=(E11-PROMEDIO(\$C\$9:\$C\$218))^2
2392	=PROMEDIO(D9:D11)	=D12-E13	=ABS(F12)	=H12^2	=H12/D12	=(E12-PROMEDIO(\$C\$9:\$C\$218))^2
2172	=PROMEDIO(D10:D12)	=D13-E14	=ABS(F13)	=H13^2	=H13/D13	=(E13-PROMEDIO(\$C\$9:\$C\$218))^2
2646	=PROMEDIO(D11:D13)	=D14-E15	=ABS(F14)	=H14^2	=H14/D14	=(E14-PROMEDIO(\$C\$9:\$C\$218))^2
2747	=PROMEDIO(D12:D14)	=D15-E16	=ABS(F15)	=H15^2	=H15/D15	=(E15-PROMEDIO(\$C\$9:\$C\$218))^2
2841	=PROMEDIO(D13:D15)	=D16-E17	=ABS(F16)	=H16^2	=H16/D16	=(E16-PROMEDIO(\$C\$9:\$C\$218))^2
2336	=PROMEDIO(D14:D16)	=D17-E18	=ABS(F17)	=H17^2	=H17/D17	=(E17-PROMEDIO(\$C\$9:\$C\$218))^2
2527	=PROMEDIO(D15:D17)	=D18-E19	=ABS(F18)	=H18^2	=H18/D18	=(E18-PROMEDIO(\$C\$9:\$C\$218))^2
2897	=PROMEDIO(D16:D18)	=D19-E20	=ABS(F19)	=H19^2	=H19/D19	=(E19-PROMEDIO(\$C\$9:\$C\$218))^2
1804	=PROMEDIO(D17:D19)	=D20-E21	=ABS(F20)	=H20^2	=H20/D20	=(E20-PROMEDIO(\$C\$9:\$C\$218))^2
2149	=PROMEDIO(D18:D20)	=D21-E22	=ABS(F21)	=H21^2	=H21/D21	=(E21-PROMEDIO(\$C\$9:\$C\$218))^2
2603	=PROMEDIO(D19:D21)	=D22-E23	=ABS(F22)	=H22^2	=H22/D22	=(E22-PROMEDIO(\$C\$9:\$C\$218))^2
2186	=PROMEDIO(D20:D22)	=D23-E24	=ABS(F23)	=H23^2	=H23/D23	=(E23-PROMEDIO(\$C\$9:\$C\$218))^2
2122	=PROMEDIO(D21:D23)	=D24-E25	=ABS(F24)	=H24^2	=H24/D24	=(E24-PROMEDIO(\$C\$9:\$C\$218))^2
2012	=PROMEDIO(D22:D24)	=D25-E26	=ABS(F25)	=H25^2	=H25/D25	=(E25-PROMEDIO(\$C\$9:\$C\$218))^2
2777	=PROMEDIO(D23:D25)	=D26-D27	=ABS(F26)	=H26^2	=H26/D26	=(E26-PROMEDIO(\$C\$9:\$C\$218))^2
<i>INTERVALO DE CONFIANZA INFERIOR</i>	<i>PRONOSTICO MA3</i>	<i>INTERVALO DE CONFIANZA SUPERIOR</i>	<i>CALCULO METRICAS DE ERROR</i>			
=E213-(\$D\$223*2)	=PROMEDIO(C210:C212)	=D213+(\$D\$223*2)	=PROMEDIO(H6:H213)	=RAIZ(PROMEDIO(I6:I213))	=PROMEDIO(J6:J213)	=SUMA(K6:K214)

Anexo 4B

Suavización exponencial Simple

α = 0.132						
DATOS ORIGINALES	SUAVIZACION EXPONENCIAL	METRICAS DE ERROR				
		ERROR DE PRONOSTICO	MAD	RMSE	MAPE	
	=T9					
2206	=SU\$4*T8+(1-SU\$4)*U7 =U7	=T8-V8	=ABS(W8)	=W8^2	=Y8/T8	
2111	=SU\$4*T9+(1-SU\$4)*U8 =U8	=T9-V9	=ABS(W9)	=W9^2	=Y9/T9	
2187	=SU\$4*T10+(1-SU\$4)*U9 =U9	=T10-V10	=ABS(W10)	=W10^2	=Y10/T10	
3137	=SU\$4*T11+(1-SU\$4)*U10 =U10	=T11-V11	=ABS(W11)	=W11^2	=Y11/T11	
2392	=SU\$4*T12+(1-SU\$4)*U11 =U11	=T12-V12	=ABS(W12)	=W12^2	=Y12/T12	
2172	=SU\$4*T13+(1-SU\$4)*U12 =U12	=T13-V13	=ABS(W13)	=W13^2	=Y13/T13	
2646	=SU\$4*T14+(1-SU\$4)*U13 =U13	=T14-V14	=ABS(W14)	=W14^2	=Y14/T14	
2747	=SU\$4*T15+(1-SU\$4)*U14 =U14	=T15-V15	=ABS(W15)	=W15^2	=Y15/T15	
2841	=SU\$4*T16+(1-SU\$4)*U15 =U15	=T16-V16	=ABS(W16)	=W16^2	=Y16/T16	
2336	=SU\$4*T17+(1-SU\$4)*U16 =U16	=T17-V17	=ABS(W17)	=W17^2	=Y17/T17	
2527	=SU\$4*T18+(1-SU\$4)*U17 =U17	=T18-V18	=ABS(W18)	=W18^2	=Y18/T18	
2897	=SU\$4*T19+(1-SU\$4)*U18 =U18	=T19-V19	=ABS(W19)	=W19^2	=Y19/T19	
1804	=SU\$4*T20+(1-SU\$4)*U19 =U19	=T20-V20	=ABS(W20)	=W20^2	=Y20/T20	
2149	=SU\$4*T21+(1-SU\$4)*U20 =U20	=T21-V21	=ABS(W21)	=W21^2	=Y21/T21	
2603	=SU\$4*T22+(1-SU\$4)*U21 =U21	=T22-V22	=ABS(W22)	=W22^2	=Y22/T22	
1981	=SU\$4*T31+(1-SU\$4)*U30 =U30	=T31-V31	=ABS(W31)	=W31^2	=Y31/T31	
3957	=SU\$4*T32+(1-SU\$4)*U31 =U31	=T32-V32	=ABS(W32)	=W32^2	=Y32/T32	
3089	=SU\$4*T33+(1-SU\$4)*U32 =U32	=T33-V33	=ABS(W33)	=W33^2	=Y33/T33	
2469	=SU\$4*T34+(1-SU\$4)*U33 =U33	=T34-V34	=ABS(W34)	=W34^2	=Y34/T34	
2428	=SU\$4*T35+(1-SU\$4)*U34 =U34	=T35-V35	=ABS(W35)	=W35^2	=Y35/T35	
3722	=SU\$4*T36+(1-SU\$4)*U35 =U35	=T36-V36	=ABS(W36)	=W36^2	=Y36/T36	

INTERVALO DE CONFIANZA INFERIOR	PRONOSTICO SES	INTERVALO DE CONFIANZA SUPERIOR	CALCULO METRICAS DE ERROR		
=V214-(SV\$223*2)	=#jREF!*T211+(1-#jREF!)*L=V212+(SV\$221*2)		=PROMEDIO(Y3:Y211)	=RAIZ(PROMEDIO(Z3:Z211))	=PROMEDIO(AA3:AA211)

Anexo 4C

Suavización exponencial Doble

$\alpha = 0.132$

DATOS SUAVIZADOS	SUAVIZACION EXPONENCIAL DOBLE	METRICAS DE ERROR				
		ERROR DE PRONOSTICO	E-1	MAD	RMSE	MAPE
	=AD12					
2111	=\$AE\$6*AD11+(1-\$AE\$6)*AE10	=AE10	=AD11-AF11	=ABS(AG11)	=AG11^2	=AI11/AD11
2123.54	=\$AE\$6*AD12+(1-\$AE\$6)*AE11	=AE11	=AD12-AF12	=AG12	=AG12^2	=AI12/AD12
2121.88472	=\$AE\$6*AD13+(1-\$AE\$6)*AE12	=AE12	=AD13-AF13	=AG13	=AG13^2	=AI13/AD13
2130.47993696	=\$AE\$6*AD14+(1-\$AE\$6)*AE13	=AE13	=AD14-AF14	=AG14	=AG14^2	=AI14/AD14
2263.34058528128	=\$AE\$6*AD15+(1-\$AE\$6)*AE14	=AE14	=AD15-AF15	=AG15	=AG15^2	=AI15/AD15
2280.32362802415	=\$AE\$6*AD16+(1-\$AE\$6)*AE15	=AE15	=AD16-AF16	=AG16	=AG16^2	=AI16/AD16
2266.02490912496	=\$AE\$6*AD17+(1-\$AE\$6)*AE16	=AE16	=AD17-AF17	=AG17	=AG17^2	=AI17/AD17
2316.18162112047	=\$AE\$6*AD18+(1-\$AE\$6)*AE17	=AE17	=AD18-AF18	=AG18	=AG18^2	=AI18/AD18
2373.04964713257	=\$AE\$6*AD19+(1-\$AE\$6)*AE18	=AE18	=AD19-AF19	=AG19	=AG19^2	=AI19/AD19
2434.81909371107	=\$AE\$6*AD20+(1-\$AE\$6)*AE19	=AE19	=AD20-AF20	=AG20	=AG20^2	=AI20/AD20
2421.77497334121	=\$AE\$6*AD21+(1-\$AE\$6)*AE20	=AE20	=AD21-AF21	=AG21	=AG21^2	=AI21/AD21
2435.66467686017	=\$AE\$6*AD22+(1-\$AE\$6)*AE21	=AE21	=AD22-AF22	=AG22	=AG22^2	=AI22/AD22
2496.56093951463	=\$AE\$6*AD23+(1-\$AE\$6)*AE22	=AE22	=AD23-AF23	=AG23	=AG23^2	=AI23/AD23
2405.14289549869	=\$AE\$6*AD24+(1-\$AE\$6)*AE23	=AE23	=AD24-AF24	=AG24	=AG24^2	=AI24/AD24
2798.05040511992	=\$AE\$6*AD45+(1-\$AE\$6)*AE44	=AE44	=AD45-AF45	=AG44	=AG45^2	=AI45/AD45
2790.12375164409	=\$AE\$6*AD46+(1-\$AE\$6)*AE45	=AE45	=AD46-AF46	=AG45	=AG46^2	=AI46/AD46
2651.90341642707	=\$AE\$6*AD47+(1-\$AE\$6)*AE46	=AE46	=AD47-AF47	=AG46	=AG47^2	=AI47/AD47
2620.2361654587	=\$AE\$6*AD48+(1-\$AE\$6)*AE47	=AE47	=AD48-AF48	=AG47	=AG48^2	=AI48/AD48
<i>INTERVALO DE CONFIANZA INFERIOR</i>	<i>PRONOSTICO SED</i>	<i>INTERVALO DE CONFIANZA SUPERIOR</i>	<i>CALCULO METRICAS DE ERROR</i>			
=AF215-(\$AF\$223*2)	=\$AE\$2*AD214+(1-\$AE\$2)*AE213	=AF215+(\$AF\$223*2)	=PROMEDIO(AI5:AI214) =RAIZ(PROMEDIO =PROMEDIO(AK5:AK214)*100			

Anexo 4D

Mínimos Cuadrados Ordinario

PARAMETROS	α	β
PARAMETROS ESTIMADOS	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)
ERROR ESTANDAR	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)
R²	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)
F	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)
SUMA CUADRADOS	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)	=ESTIMACION.LINEAL(C4:C212,A4:A212,,1)