



UNIVERSIDAD NACIONAL
AUTONOMA DE MÉXICO



FACULTAD DE ESTUDIOS SUPERIORES
ACATLAN



"AJUSTE DE MODELOS DE DISTRIBUCIÓN DE
PROBABILIDAD: SINIESTROS OCURRIDOS EN
HABITACION PARA LA COBERTURA DE INCENDIO"

T E S I N A

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A

JOSE SERGIO MARTINEZ SANTANA

ASESOR: MTRO. IVAN MEJIA GUEVARA

ABRIL 2005

m. 342585



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

A MI HIJO

ARIEL MAXIMILIANO MARTÍNEZ VARGAS

GRACIAS POR LLEGAR A MI VIDA Y
REGALARME TU TERNURA, AMOR Y
CARIÑO Y ALEGRAR CADA MOMENTO
CON TU LINDA SONRISA, TE AMO.

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

NOMBRE: José Sergio

Martínez Santana

FECHA: 5 Abril 2005

FIRMA: [Firma manuscrita]

A DIOS

GRACIAS POR DARMEL DON DE LA VIDA
Y SENTIRTE SIEMPRE JUNTO A MÍ.

A MIS PADRES:

JOSÉ ÁNGEL MARTÍNEZ DÍAZ
MA. GUADALUPE SANTANA VALDESPINO

GRACIAS POR TODO SU AMOR, CARIÑO Y
COMPENSIÓN, SU ESFUERZO NO FUE EN
VANO, ESTE LOGRO TAMBIÉN ES SUYO, LOS
AMO Y RESPETO PROFUNDAMENTE.

A MI HERMANA

MA. DE LOS ÁNGELES MARTÍNEZ SANTANA

ERES LA MEJOR HERMANA DEL MUNDO,
GRACIAS POR CUIDARME Y ESTAR
SIEMPRE A MI LADO, TE ADMIRO Y TE
QUIERO MUCHO.

A MI ESPOSA

ARGELIA VARGAS RANGEL

TE AMO PROFUNDAMENTE, GRACIAS POR
COMPARTIR TU VIDA JUNTO A MÍ Y SOBRE TODO
GRACIAS POR EL REGALO MÁS HERMOSO QUE
ME HAS DADO, EL SER PADRE DE UN NIÑO
MARAVILLOSO.

A MIS SOBRINOS

FÁTIMA CARMÍN ESPINO MARTÍNEZ
CESAR ISAAC ESPINO MARTÍNEZ
VÍCTOR EMANUEL ESPINO MARTÍNEZ

SIEMPRE ESTARÉ A SU LADO, LOS AMO.

A TODOS MIS FAMILIARES Y AMIGOS, Y EN ESPECIAL
A LOS QUE YA NO ESTAN CON NOSOTROS EN ESTA
VIDA.

A MI ASESOR DE TESIS

ACT. IVÁN MEJÍA GUEVARA

GRACIA POR TU AYUDA DESINTERESADA,
PERO SOBRE TODO GRACIAS POR TU
AMISTADA.

A MI QUERIDA UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO Y A TODOS LOS
PROFESORES Y AMIGOS DE LA F. E. S. ACATLÁN Y C. C. H. NAUCALPAN.

CAPÍTULO 1

Modelos Matemáticos

"Was beweisbar ist, soll in der Wissenschaft nicht ohne Beweis geglaubt werden."

"En Ciencia, lo que se puede probar no debe ser creído sin demostración."

R. Dedekind (1887)

La representación de la realidad a través de un modelo que la refleje y explique tiene antecedentes en todas las disciplinas del conocimiento. El uso de modelos con fundamento matemático ha permitido el desarrollo de teorías que describen fenómenos tanto naturales como sociales y permiten su comprensión y regulación, así como la obtención de mejores pronósticos y predicciones sobre sus resultados.

“Un modelo es un esquema teórico, de un sistema o de una realidad compleja, que se elabora para facilitar su comprensión y el estudio de su comportamiento.”¹

Dentro de los modelos nos encontramos con un tipo que ha tomado un gran auge en los últimos tiempos, y en particular, con la aparición de las computadoras, estos son los modelos matemáticos:

“Un modelo matemático es cualquier sistema completo y compatible de ecuaciones (estructuras) matemáticas, diseñadas para que se correspondan con alguna otra entidad: su prototipo. Tal prototipo puede ser una entidad física, biológica, psicológica o conceptual, tal vez, incluso otro modelo matemático.”²

Iniciaremos por revisar el desarrollo que los modelos han tenido a lo largo de la historia de la humanidad.

1.1 Historia de los Modelos Matemáticos

Gran parte del progreso de la humanidad se puede atribuir a la aplicación del método científico³ a los problemas donde la costumbre, inercia y la tradición habían imperado. Ejemplos individuales en los que parece que se usó el método científico para resolver problemas se han encontrado en escritos de miles de años de antigüedad. Podemos enumerar los siguientes acontecimientos más significativos relacionados con la evolución y el uso de modelos a lo largo de la historia de la humanidad:

- A. C.
 - Al suegro de Moisés, Jethro, se le acredita con un tratado sobre organización en el Cáp. 18 del Libro del Éxodo.

¹ Ver <<http://www.rae.es/>>

² Ver <<http://personal.telefonica.terra.es/web/jmora7/Archiv/96zoomc.PDF>>

³ “Estudio sistemático, controlado, empírico y crítico de proposiciones hipotéticas acerca de presuntas relaciones entre varios fenómenos”, ver <http://es.wikipedia.org/wiki/M%C3%A9todo_cient%C3%Adfico>

- Antiguos barcos en la ciudad Italiana de Venecia fueron reacondicionados y reparados en una línea de ensamblaje extremadamente ingeniosa; cada barco era movido a lo largo de esta línea en la que un grupo de expertos trabajadores efectuaban ciertas operaciones específicas en cada estación de la línea.
- 212 A. C.
 - La ciudad de Siracusa empleó a Arquímedes para idear métodos para romper el sitio naval a la ciudad, que estaba bajo el ataque de los romanos.
- 1798.
 - Eli Whitney usó el método científico para analizar la industria del algodón y desarrollar el concepto de partes intercambiables, que permitieron el cambio de la producción ajustada individual a la elaboración de grandes volúmenes de partes estandarizadas.
- 1832.
 - Charles Babbage, hace énfasis en el método científico, división del trabajo, estudio de tiempos y movimientos, especialización y contabilidad de costos.
- Finales de 1800.
 - Frederick Taylor aplicó el método científico al análisis de la productividad en las personas, cavadoras de minas y otros materiales. El estudio del paleo de Taylor es un ejemplo excelente de la aplicación del método científico.
- 1911.
 - Frederick Taylor publicó los *Principios de la Ciencia de la Administración*, los cuales formalizan los tiempos de estudio y los métodos de trabajo-estudio que aún hoy se usan.
 - Frank y Lillian Gilbreth desarrollaron los conceptos básicos de la psicología industrial, ilustrando su investigación sobre el estudio de los movimientos elementales de una operación manual en la industria.
- 1914.
 - F. W. Lanchester intentó tratar cuantitativamente las operaciones militares. Obtuvo ecuaciones que relacionaban el resultado de una batalla tanto a la fuerza relativa de los combatientes como a su capacidad relativa de fuego. Las ecuaciones de Lanchester sugerían que el poder total de las fuerzas de pelea era proporcional al cuadrado de su fuerza numérica. Lanchester modeló una situación que involucraba opciones estratégicas. Probó ese modelo contra una situación conocida en el mundo real.
 - Tomas Alba Edison estudiaba el proceso de la guerra antisubmarina. Reunió estadísticas a ser usadas en el análisis de maniobras mediante las cuales los barcos en la superficie pudieran evadir y destruir submarinos. Ideó un juego de

guerra a ser usado en la simulación de maniobras navales en vez de arriesgar los barcos en condiciones bélicas reales.

- 1915.
 - En el área de control de inventarios, los modelos del tamaño del lote económico tienen una larga genealogía, aunque se ha reportado que G. D. Babcock desarrolló un modelo enunciado en la forma de una ecuación cúbica, su técnica nunca fue publicada. El primer modelo publicado sobre el tamaño del lote económico de inventarios es generalmente atribuido a Ford W. Harris, quien describió y publicó su modelo en este año.
- 1917.
 - A. K. Erlang, un matemático danés que trabajaba con la compañía telefónica de Copenhague, publicó su trabajo más importante, "Solución a Algunos Problemas en la Teoría de Probabilidades Importantes en las Centrales Telefónicas Automáticas". Contenía sus fórmulas de tiempos de espera que había desarrollado sobre la base de principios estadísticos. Estas fórmulas ahora bien conocidas, son de importancia fundamental a la teoría del tráfico telefónico.
- 1924.
 - W. Shewhart realizó las primeras aplicaciones registradas de la deducción estadística, cuando introdujo el concepto de gráficas de control de calidad.
 - La utilización de la deducción estadística y la teoría de probabilidades, fue ayudada por el trabajo de H. F. Dodge y H. G. Romig, compañeros de trabajo con Shewhart en los laboratorios de la Bell Telephone. Desarrollaron la técnica de inspección por muestreo en conexión con el control de calidad y publicaron tablas de muestreo estadístico que son ampliamente usadas hoy en día.
- 1928.
 - Otro ingeniero en los Bell Laboratories, T. C. Fry, hizo importantes contribuciones adicionales a las bases estadísticas de la teoría de colas. Una serie de conferencias dictadas por Fry en este año con relación a las aplicaciones a ingeniería de la teoría de probabilidades, se convirtió en la base de su importante libro sobre la materia.
 - Sir Roland Fisher trabajó sobre diversos métodos estadísticos modernos. Cuando estos fueron publicados tuvieron muy poco efecto; pero ahora, son la base para la mayoría de la teoría estadística aplicada.
- 1930's.
 - John Von Neumann desarrolló un modelo para una economía en expansión. También desarrolló las bases teóricas de la *Teoría de Juegos*.

□ Segunda Guerra Mundial.

- Nacimiento de la *Investigación de Operaciones*⁴. Desde principios de 1937 se pidió a los científicos ingleses cada vez con más frecuencia, que ayudaran a los militares a descubrir la mejor manera de utilizar el radar para localizar aviones enemigos. En septiembre de 1939 los científicos que trabajaban en diferentes aspectos del problema se reunieron en el Cuartel General del Mando de Aviones de Combate. Ese grupo, es considerado como el núcleo del primer grupo de Investigación de Operaciones. Este grupo ampliaba continuamente su área de actividades hasta abarcar más allá del problema original del radar y de su integración con los observadores de tierra. Grupos similares de Investigación de Operaciones se formaron en los Estados Unidos. Estos grupos sentaron las bases para el planteamiento y el desarrollo de soluciones a problemas de programación lineal, transportación, asignación de recursos, etc.

Concluida la Segunda Guerra Mundial los diversos grupos de Investigación de Operaciones que funcionaron para aspectos bélicos, pasaron a servir ahora al terreno industrial. Un poco más tarde, en la década de los cincuenta, la computadora incrementó el cambio, posibilitando el uso de los modelos matemáticos no solo para resolver problemas ocasionales, sino también, para resolver problemas operacionales comunes, rutinarios y aun cotidianos. Es más, con la computadora, es posible el uso de modelos matemáticos para predecir e imaginar problemas, para medir repercusiones si es que se toman ciertas medidas o caminos de acción, para requerir y generar información necesaria para la toma de decisiones.

1.2 Filosofía del Proceso de Modelar

Una de las primeras características que debe determinarse acerca de la situación en estudio es si es más apropiado modelar en términos deterministas o estocásticos. Un modelo es determinista si las predicciones y pronósticos realizados con suficiente información en un instante o etapa en el tiempo pueden determinar exactamente el comportamiento futuro del sistema en estudio. Por otro lado, un modelo es estocástico si este incorpora un comportamiento probabilístico. Para estos modelos no importa que tanto se conozca del sistema en un instante determinado en el tiempo, ya que, es imposible determinar con absoluta certeza la naturaleza del sistema para tiempos futuros. Por lo general, muchos de los modelos usados son de este tipo, es decir, modelos cuya descripción matemática envuelven oportunidad e incertidumbre, esto es totalmente esperado ya

⁴ “Aplicación , por grupos interdisciplinarios, del método científico a problemas relacionados con el control de las organizaciones o sistemas , a fin de que produzcan soluciones que mejor sirvan a los objetivos de la organización”, ver <<http://www.investigacion-operaciones.com/Historia.htm>>

que el mundo real provee una fuerte evidencia de ser un sistema estocástico. La decisión de que tipo de modelo debe ser construido depende de muchos factores y en última instancia simplemente es una elección del investigador. Usualmente un modelo determinista es tomado como una primera aproximación en una situación cuando un modelo estocástico parece ser más apropiado, sin embargo, no se debe de asumir que las predicciones basadas en un tipo de modelo son necesariamente mejores (o peores) que aquellas basadas en otro tipo. Los méritos relativos de los dos tipos de modelos varían de una situación a otra.

Sí un modelo es de uso práctico, se debe de tener un medio para obtener resultados que puedan ser probados o comparados con el mundo real. Muchos de los modelos guían a relaciones matemáticas que involucran uno o varios parámetros. Rastreado estos parámetros de vuelta al mundo real, se puede encontrar que estos están relacionados con una tasa de aprendizaje, una probabilidad de muerte, una razón de activos, etc. Por lo tanto, al comparar el estudio matemático con la realidad, podría ser necesario dar valores numéricos a estos parámetros y para muchos modelos esta es la tarea principal.

El uso de técnicas estadísticas juega un importante rol en varios aspectos de la construcción de modelos. Un cuestionamiento estadístico podría ser, ¿Cuál es la relación entre el estadístico \bar{x} para la muestra y su correspondiente valor medio \bar{X} para la población total? Problemas de este tipo pueden ser resueltos por medio de las técnicas estadísticas de estimación de parámetros. Otra importancia de los métodos estadísticos en la modelación matemática es la de probar el modelo por correspondencia y exactitud: ¿El modelo realmente hace lo que se pretendía que hiciera? Un acercamiento para responder tal pregunta usualmente es por medio de la acumulación de datos, tanto de experimentos controlados o a través de observaciones del mundo real y entonces efectuar la evaluación comparando los resultados predichos por el modelo con aquellos resultados del análisis estadístico de los datos. Existen varias pruebas estadísticas estándar con muchas variantes, las cuales son útiles en esta conexión.

Los dos problemas arriba mencionados no son independientes. Se podría esperar que un modelo prevea resultados confiables si los parámetros han sido cuidadosamente seleccionados. Así, cuando se evalúen las predicciones basadas en un modelo en particular, es esencial tener en cuenta los procesos usados para determinar los parámetros y la exactitud que podría esperarse. También es común que un conjunto de datos se use para determinar los parámetros y para determinar las pruebas del modelo. Variaciones importantes en los datos podrían ser reflejadas en la estimación de los parámetros y en las pruebas, y aunque el modelo pudiera predecir resultados consistentes con los datos, los datos defectuosos podrían oscurecer la eficiencia del modelo.

Una evaluación estadística de la exactitud de un modelo usualmente se lleva a cabo junto con el uso de ciertas medidas estándar de la discrepancia entre lo predicho y los datos observados. De esta forma, se obtienen medidas numéricas de la bondad del ajuste para cada modelo. Naturalmente, si un modelo provee un ajuste más consistente que cualquier otro modelo, entonces este modelo será aceptado y los otros rechazados, sin embargo, frecuentemente sucede que cierto modelo será el que mejor explique y se ajuste a ciertos conjuntos de datos, mientras otro modelo será el que mejor explique y prediga otros conjuntos de datos. Ningún modelo puede ser rechazado, ya que bajo ciertas circunstancias cada uno es el mejor. Asimismo, ningún modelo debe ser completamente aceptado ya que en ciertos casos cada modelo no es el mejor disponible. Estos pueden ser condicionalmente aceptados, estudiados, y usados en aquellas circunstancias donde resulten ser la elección más apropiada.

1.3 Tipos de Modelos

Las diferentes clasificaciones de los modelos nos proporcionan una idea adicional de sus características esenciales, ya que, pueden describirse de muchos modos. Los modelos pueden clasificarse por sus dimensiones, funciones, propósitos o grado de abstracción. La base más común es la de tipos de modelos, que incluyen los siguientes tipos básicos:

- i. Icónicos
- ii. Análogos
- iii. Simbólicos

1.3.1 Modelos Icónicos

Un modelo icónico es una representación física de algunos objetos, ya sea en forma idealizada o en escala distinta. Para expresarlos de otro modo, una representación es un modelo icónico hasta el grado en que sus propiedades sean las mismas que tiene aquello que representa. Los modelos icónicos son muy adecuados para la descripción de acontecimientos en un momento específico del tiempo. Otra característica de un modelo icónico la constituyen sus dimensiones, dos dimensiones (fotografía, plano y mapa), o tres dimensiones (globo, automóvil y avión). Cuando un modelo sobrepasa la tercera dimensión, es imposible construirlo físicamente, y entonces pertenece a otro tipo de categoría llamados simbólicos o matemáticos.

1.3.2 Modelos Análogos

Los modelos análogos pueden representar situaciones dinámicas y se usan más que los icónicos, porque pueden mostrar las características del acontecimiento que se estudia. Las curvas de demanda, las curvas de distribución de frecuencias en las estadísticas y los diagramas de flujo, son ejemplos de modelos análogos. A menudo un modelo análogo es muy adecuado para representar relaciones cuantitativas entre las propiedades de los objetos de varias clases. Al transformar las propiedades en propiedades análogas, con frecuencia podemos incrementar nuestra capacidad de hacer cambios. Una de las ventajas de los modelos análogos sobre los icónicos es que ordinariamente puede hacerse que los primeros representen muchos procesos distintos del mismo tipo, lo que se hace evidente en el flujo de trabajos en proceso y de productos terminados de una fábrica. No podría usarse eficazmente un modelo icónico para estudiar los efectos de ciertos cambios en el control de calidad. Un diagrama de flujo es un modelo análogo muy sencillo y eficaz en esas situaciones.

1.3.3 Modelos Simbólicos

Los modelos simbólicos son representaciones de la realidad y toman la forma de cifras, símbolos y ecuaciones matemáticas. Comienzan como modelos abstractos que formamos en nuestra mente y que luego se registran como modelos simbólicos. Un tipo de modelo simbólico o matemático comúnmente usado es una ecuación. Una ecuación es concisa, precisa y fácil de comprender. Sus símbolos no solo son muchos más fáciles de manipular que las palabras, sino que se escriben más rápidamente. Además de estos atributos, los modelos simbólicos se prestan a las manipulaciones y uso en las computadoras.

Otros tipos incluyen modelos gráficos y pictóricos. Hay que tener en cuenta que pueden presentarse problemas para los que las analogías son más eficientes que los modelos simbólicos. Por ejemplo, un sistema puede ser tan complicado que la cantidad de trabajo requerida para construir un modelo simbólico sea demasiado costoso si se relaciona con las ganancias posibles. A menudo es difícil asignar tan solo un modelo a una clase, y esto es especialmente cierto con respecto a los modelos de simulación, que son modelos análogos y que se describen con símbolos matemáticos.

Los modelos matemáticos usualmente se clasifican en:

- i. Cuantitativos y Cualitativos.
- ii. Estándar y Hechos a la medida.
- iii. Probabilísticos y Determinísticos.
- iv. Descriptivos y de Optimización.
- v. Estáticos y Dinámicos.
- vi. Simulación y No Simulación.

1.4 Proceso para la Construcción de Modelos Matemáticos

El proceso de modelar se ilustra en la figura 1 siguiente,

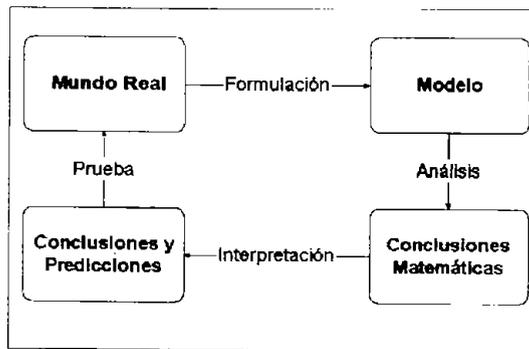


Figura 1. Proceso de Modelar

Algunos de los puntos más importantes de este proceso son:

□ Etapa 1

- **Analizar el problema.** En esta etapa se debe estudiar la situación lo suficiente para identificar el problema en una forma precisa y comprender claramente sus cuestionamientos fundamentales. Se deben determinan los objetivos del problema y por lo general se decide si el problema es del tipo determinista o del tipo estocástico. Solo con una clara y precisa identificación del problema se puede trasladar el problema en conceptos y símbolos matemáticos y de esta forma desarrollar y resolver el modelo.

□ **Etapa 2.**

- **Formular un modelo.** En esta etapa, una vez que el problema ha sido identificado de manera clara y precisa se diseña un modelo formando una abstracción del sistema o fenómeno que se este modelando. Algunas de las tareas principales en esta etapa son:
 - *Obtener datos.* Se deben coleccionar datos relevantes con información benéfica acerca del comportamiento del sistema.
 - *Realizar simplificaciones.* La formulación de un modelo, se debe intentar ser tan simple como razonablemente sea posible. Así, con frecuencia se decide simplificar algunos de los factores que no parecen ser importantes. Muchos problemas son demasiado complejos para considerar cada detalle, y si fueran considerados solo podría hacer del modelo imposible de resolver o consumir un monto importante de tiempo el resolvertlo. Por lo tanto, frecuentemente existen factores que apreciablemente no afectan los resultados. Además simplificando factores, se podría decidir regresar al paso previo para restringir el problema en estudio.
 - *Determinar variables.* Se deben determinar y nombrar las variables. Una variable independiente es la variable de la cual otras dependen. El modelo tratará de explicar las variables dependientes. Para simplificar el modelo, se podrían eliminar algunas variables, tratar ciertas variables como constantes, o agregar varias variables dentro de una. Mientras se decide sobre las variables, se deben establecer sus unidades, tales como: hora, día, semana o año para una unidad de tiempo.
 - *Establecer relaciones entre las variables y submodelos.* Hay que dibujar un diagrama del modelo descomponiéndolo en submodelos e indicar la relación entre las variables. Para simplificar el modelo, se podrían asumir relaciones más simples de lo que realmente son, por ejemplo, se podría asumir que dos variables están relacionadas de una manera lineal en lugar de una relación más compleja.
 - *Determinar ecuaciones.* Mientras se establecen las relaciones entre las variables, se podrían determinar ecuaciones con estas variables. Muchos modelos en las ciencias sociales y de la vida envuelven ecuaciones diferenciales, o las ecuaciones involucran una derivada.

□ Etapa 3

- **Resolver el modelo.** En esta etapa se implementa el modelo. Es importante no brincar a este paso antes de comprender el problema y diseñar el modelo. En otro caso, mucho tiempo puede ser desperdiciado y experimentar grandes frustraciones. Algunas de las técnicas que se pueden emplear en la solución son: álgebra, cálculo, gráficos, programación, paquetes computacionales como MINITAB, STATGRAPHICS, SAS, etc. La solución podría producir una respuesta exacta o podría simular la situación. Si el modelo es muy complicado de resolver, se debe volver a la etapa 2 y realizar simplificaciones adicionales o a la etapa inicial y reformular nuevamente el problema.
- **Verificación e interpretación de la solución del modelo.** Una vez que se tiene la solución, se deben examinar cuidadosamente los resultados y estar seguros que estos tienen sentido y que la solución resuelve el problema original y es útil. Probar la solución y observar si las predicciones concuerdan con los datos reales es sumamente importante. Se debe analizar la solución del modelo para determinar sus implicaciones. Si la solución del modelo muestra debilidad, se deberá volver a la etapa 1 o 2 para determinar si es factible refinar el modelo, si es así, se debe de realizar todo el proceso nuevamente. Así, el ciclo del proceso de modelar es una compensación entre simplificación y refinamiento. Por refinamiento, se podría necesitar extender el alcance del problema en la etapa 1. En la etapa 2, mientras el modelo es refinado, frecuentemente se necesitan reconsiderar las simplificaciones, incluir más variables y submodelos e incluir técnicas más sofisticadas.

□ Etapa 4

- **Publicar los resultados.** Divulgar sobre el modelo es importante por su utilidad. Un reporte contiene los siguientes componentes, los cuales van de la mano con el proceso de modelar:
 - Análisis del Problema
 - Diseño del modelo
 - Solución del modelo
 - Resultados y conclusiones
- **Mantener el modelo.** Como la solución del modelo es usada, podría ser necesario o deseable hacer ciertas correcciones, mejoras o realces. En este caso se debe realizar el proceso nuevamente para revisar la solución desarrollada.

El proceso de modelar es un esfuerzo científico creativo. Como tal, el problema que se pretende modelar usualmente no tiene una respuesta correcta. Los problemas son complejos y muchos modelos proporcionan ser buenos aunque con diferentes soluciones.

Seguiremos el proceso de modelar antes expuesto y realizaremos algunas modificaciones a este, y lo ajustaremos para que nos sirva de base para ajustar modelos matemáticos para fenómenos aleatorios (funciones de distribución). La teoría de probabilidades nos proporciona una gran variedad de modelos que bien podrían ajustarse a nuestros datos, ¿Cómo hacer esta elección? La respuesta a esta pregunta nos la brindara los métodos y técnicas estadísticas los cuales formaran parte en el proceso para ajustar modelos de fenómenos aleatorios a datos generados por el monto de siniestros ocurridos para la cobertura de incendio edificio para casa habitación.

CAPÍTULO 2

Probabilidad y Modelos para Variables Aleatorias

"Solo mediante la aplicación de las leyes del azar a las observaciones pueden determinarse, en un caso concreto, si existe una relación de causa-efecto."

Max Born (1882-1970)

Una gran variedad de experimentos y fenómenos naturales, biológicos, sociales o combinaciones de estos, tienen la característica de generar resultados u observaciones que no son susceptibles de predecirse con certeza, esto es, que aún cuando se estudien repetidamente bajo un mismo conjunto de circunstancias no siempre se observa el mismo resultado. Entre ellos, podemos citar los juegos de azar: lanzamiento de dados o monedas, juegos con barajas, ruletas, loterías, etc. Otra familia de ejemplos también muy conocida, la proporcionan los fenómenos meteorológicos: magnitud, intensidad y extensión de las lluvias; humedad, dirección y velocidad de los vientos; temperaturas máximas y mínimas; y todas sus consecuencias como volúmenes de agua en las presas, magnitud de los daños provocados por inundaciones, sequías, incendios, heladas, etc.

2.1 Fenómenos Aleatorios

Una de las características más notables de nuestra época consiste en el empleo, cada vez mayor, de las ideas de la teoría de probabilidades en una amplia variedad de campos científicos, que abarca temas tan ajenos y diferentes entre sí. ¿Qué es entonces lo que estudia la teoría de probabilidades?, ¿Qué da lugar a esta diversidad de aplicaciones? Para contestar estas preguntas comenzaremos por definir una propiedad en común que tienen ciertos fenómenos.

“Un fenómeno aleatorio (fortuito o al azar) es un fenómeno empírico que se caracteriza por la propiedad de que, al observarlo bajo un determinado conjunto de condiciones, no siempre se obtiene el mismo resultado (no existe regularidad determinista) sino que los diferentes resultados ocurren con regularidad estadística. Esto quiere decir que existen números entre 0 y 1 que representan la frecuencia relativa con la que se observan los diferentes resultados en una serie de repeticiones independientes del fenómeno.”⁵

Dos conceptos estrechamente ligados al de fenómeno aleatorio son los de *evento aleatorio* y *probabilidad de un evento aleatorio*. Un evento aleatorio tiene la propiedad de que la frecuencia relativa con la que aparece en una sucesión muy larga de observaciones realizadas al azar, se acercan a un valor límite estable a medida que el número de observaciones tiende a infinito; el valor límite de la frecuencia relativa se llama probabilidad del evento aleatorio.

Para arrojar más luz sobre el significado de un evento aleatorio, consideremos uno que es característico; por ejemplo, un accidente de automóvil. Es imposible predecir que un automovilista en particular que inicia un viaje por carretera se vea o no envuelto en algún accidente, sin

⁵ Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 18.

embrago, si observamos a todos los automovilistas (o a un gran número de ellos) que viajan por la misma carretera en un mismo día, es posible determinar la proporción de los que sufrirán accidentes, si esta proporción permanece constante de un día a otro, podemos considerar que lo que acontece a un automovilista en particular en esa carretera es un fenómeno aleatorio y que el accidente es un evento aleatorio.

2.2 Modelos Matemáticos de Fenómenos Aleatorios

La teoría de probabilidades estudia los métodos de análisis que son comunes en el tratamiento de fenómenos aleatorios, cualquiera que sea el área en que éstos se presenten. La probabilidad es, pues, la ciencia de los fenómenos aleatorios, en el sentido que estudia las propiedades de estos fenómenos que dependen esencialmente del concepto de aleatoriedad y no de otros aspectos particulares.

Consideramos la teoría de probabilidades como una disciplina matemática y, como tal, se construye mediante el método axiomático. De aquí, mediante la deducción lógica y sin recurrir a la experimentación se obtienen diversas proposiciones llamadas teoremas. Aun cuando éstos no se refieran al mundo real, sino que son consecuencia lógica de los axiomas, representan de hecho, conclusiones sobre fenómenos reales en la medida en la que acepte que dichos fenómenos tienen las propiedades postuladas en los axiomas.

Llegamos así al concepto de lo que es un modelo matemático de un fenómeno real. Si tenemos una regla para convertir proposiciones de una teoría matemática en proposiciones relativas a un fenómeno real, entonces la teoría matemática, que se ha construido axiomáticamente, es un modelo del fenómeno. En general, solo se necesita identificar, los objetos abstractos a que se refieren los axiomas de la teoría con los aspectos del fenómeno real, para estar en condiciones de utilizar esta teoría como su modelo. Una vez que se ha planteado el modelo matemático, uno espera que los teoremas deducidos describan el fenómeno con la misma fidelidad que lo hacen los axiomas, puesto que aquellos no son más que consecuencias lógicas de estos últimos.

2.2.1 Espacio de Descripciones Muestrales

Hemos afirmado que la teoría de probabilidades es el estudio de modelos matemáticos de los fenómenos aleatorios, que trata de las proposiciones que pueden hacerse acerca de un fenómeno aleatorio del que se hacen ciertos postulados, pero ¿Cómo se formulan postulados acerca de un

fenómeno aleatorio?. Para contestar esta pregunta nos valdremos del concepto de espacio de descripciones muestrales.

“Un espacio de descripciones muestrales de un fenómeno aleatorio que denotaremos con la letra Ω , es el espacio de las descripciones (o nombres) de todos los resultados posibles de un experimento.”⁶

Un espacio de descripciones muestrales de un fenómeno aleatorio puede ser definido de más de una manera. Si quienes observan el fenómeno tienen concepciones diferentes de los posibles resultados de la observación, seguramente definirán diferentes espacios de descripciones muestrales.

2.2.2 Eventos

La importancia del concepto de espacio muestral que corresponden a un fenómeno aleatorio se debe a que constituyen un medio para definir el concepto de evento.

“Un evento es un subconjunto del espacio muestral. La clase de todos los eventos asociados con un experimento se define como el espacio de eventos, denotado por Ψ .”⁷

Por lo tanto, un evento lo definimos como un conjunto de descripciones. Al decir que cierto evento A ha ocurrido, significa que el resultado de la situación aleatoria considerada, tiene por descripción un elemento del conjunto A . Obsérvese que se están definiendo dos conceptos: el de *evento* y el de *ocurrencia y realización de un evento*. El primero es un instrumento básico para construir modelos matemáticos de fenómenos aleatorios, mientras que el segundo constituye una base para traducir las proposiciones relativas al modelo matemático a proposiciones sobre el fenómeno real.

Al estudiar un fenómeno aleatorio nos interesan los eventos que puedan ocurrir (o, más precisamente, las probabilidades de que ocurran); por lo tanto, el interés del espacio de descripciones muestrales no radica en sus elementos; que son las descripciones muestrales, sino en sus subconjuntos, ¡que son los eventos!

⁶ Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 24

⁷ Ver Mood, “Introduction to the Theory of Statistics”, p. 15

La definición de espacio muestral es precisa y satisfactoria, mientras que, la definición de evento y espacio de eventos no es totalmente satisfactoria, ya que esta, no dice exactamente que subconjuntos podrían ser eventos y cuales no. Más que desarrollar las matemáticas necesarias para definir precisamente que subconjunto de Ω constituyen el espacio de eventos Ψ , estableceremos algunas propiedades para Ψ que parecerían ser razonablemente requeridas:

- o $\Omega \in \Psi$.
- o Si $A \in \Psi$, entonces $A^c \in \Psi$.
- o Si A_1 y $A_2 \in \Psi$, entonces $A_1 \cup A_2 \in \Psi$

Cualquier colección de eventos con las propiedades arriba mencionadas es llamada un álgebra Booleana, o simplemente un álgebra.

2.3 Definición de Probabilidad

Una de las formas para definir la probabilidad es por medio del concepto *clásico*, que se aplica cuando todos los resultados posibles son igualmente probables.

"Si un experimento aleatorio puede resultar en n resultados igualmente probables y mutuamente excluyentes⁸ y si n_A de estos resultados tienen un atributo A , entonces la probabilidad es la fracción n_A/n ."⁹

Esta definición también se conoce como la probabilidad *a priori*. Una deficiencia importante del concepto clásico de probabilidad es su aplicación limitada, pues hay muchas situaciones en que las posibilidades que se presentan no pueden considerarse igualmente probables. Otro concepto de probabilidad, ampliamente usado es el de *frecuencia relativa* o también conocida como probabilidad *a posteriori*.

⁸ Los subconjuntos A y B se definen ser mutuamente excluyentes o disjuntos si $A \cap B = 0$. Los subconjuntos A_1, A_2, \dots se definen ser mutuamente excluyentes si $A_i A_j = 0$ para cada $i \neq j$, Ver Mood "Introduction to the Theory of Statistics", p. 14.

⁹ Ver Mood, "Introduction to the Theory of Statistics", p. 3.

Si en un experimento aleatorio, se observa que cuando el número n de observaciones aumenta, las frecuencias relativas con las que el número de cierto suceso con el atributo A , la fracción n_A/n , converge hacia cierta cantidad denominada la probabilidad de A , es decir, $P[A] = \lim_{n \rightarrow \infty} (n_A/n)$.

Con esta definición no se puede garantizar lo que ocurrirá en una ocasión en particular, pero si se llevaran registros de un gran número de observaciones durante un largo periodo, se debe de encontrar que la frecuencia observada del suceso con el atributo bajo estudio es bastante confiable.

La *definición axiomática* de la probabilidad se debe al Físico-Matemático ruso Andrei Nikolaevich Kolmogorov.

“Dada una situación aleatoria, descrita por un espacio de descripciones muestrales Ω , la probabilidad es una función¹⁰ que asigna a cada evento un número real no negativo, denotado por $P[A]$, que se llama probabilidad del evento A . Dicha función de probabilidades debe satisfacer los siguientes axiomas:

- i. $P[A] \geq 0$ para todo evento A .
- ii. $P[\Omega] = 1$.
- iii. $P[A_1 \cup A_2] = P[A_1] + P[A_2]$, donde $A_1 \cap A_2 = \emptyset$.¹¹

Un concepto ligado al de función de probabilidad es el de espacio de probabilidades.

“Un espacio de probabilidad es la tripleta $(\Omega, \Psi, P[\cdot])$, donde Ω es un espacio muestral, Ψ es la colección de eventos, cada uno un subconjunto de Ω , y $P[\cdot]$ es una función de probabilidad con dominio Ψ .”¹²

¹⁰ “Una función es una regla que asigna un número real a cada elemento de un conjunto de objetos, llamado dominio de la función. Aquí el dominio de la función de probabilidad $P[\cdot]$ es el conjunto de todos los eventos de Ω ”, Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 34

¹¹ Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 34

¹² Ver Mood, “Introduction to the Theory of Statistics”, p. 25

El espacio de probabilidad es un término único, que permite una forma conveniente de suponer la existencia de los tres componentes de su notación. Los tres componentes están relacionados; Ψ es la colección de subconjuntos de Ω , y $P[\cdot]$ es una función que tiene Ψ como dominio.

2.4 Modelos para Variables Aleatorias

Hemos visto que la probabilidad de un evento aleatorio se puede estudiar únicamente con referencia a un espacio de descripciones muestrales sobre el cual se haya definido una función de probabilidad. Sin embargo, en muchas aplicaciones de la teoría de probabilidad la terminología de los espacios de descripciones muestrales no interviene explícitamente (aunque, el concepto siempre interviene implícitamente). En cambio, muchas aplicaciones de la teoría de probabilidades se basan en el concepto de *variable aleatoria*.

2.4.1 Concepto

Hemos definido un evento como un conjunto de descripciones muestrales; por lo tanto, los eventos de fenómenos aleatorios con resultados numéricos son conjuntos de números reales, sin embargo, no todo conjunto de números reales puede considerarse como evento. Por lo tanto, por la palabra evento entenderemos no solo un conjunto de números reales, sino un conjunto probabilizable de números reales. No nos involucraremos para caracterizar los conjuntos de números reales que son probabilizables, únicamente podemos señalar que la familia \mathfrak{S} de conjuntos probabilizables siempre tiene las propiedades siguientes:

- a) Cualquier intervalo pertenece a \mathfrak{S} (un intervalo es un conjunto de números reales de la forma $\{x: a < x < b\}$, $\{x: a \leq x < b\}$, $\{x: a < x \leq b\}$, $\{x: a \leq x \leq b\}$, en donde a y b pueden ser números finitos o infinitos.
- b) A \mathfrak{S} pertenece el complemento A^c de cualquier conjunto A que pertenezca a \mathfrak{S} .
- c) A \mathfrak{S} pertenece la unión $\bigcup_{n=1}^{\infty} A_n$ de cualquier sucesión de conjuntos $A_1, A_2, \dots, A_n, \dots$ que pertenezcan a \mathfrak{S} .

Una definición más precisa del concepto de evento, se puede hacer de la siguiente manera. Existe en la recta real una familia más pequeña de conjuntos con las propiedades a), b) y c), esta familia se denota por \mathfrak{B} , y cualquier miembro de \mathfrak{B} , se llama conjunto boreliano. Puesto que \mathfrak{B} es la familia más pequeña que posee las propiedades a), b) y c), se deduce que \mathfrak{B} está contenida en

\mathfrak{S} , la familia de conjuntos probabilizables. Entonces, todo conjunto boreliano es probabilizable. Por lo tanto entenderemos por un evento relativo a un fenómeno aleatorio con resultados numéricos, un conjunto boreliano de números reales.

El concepto de variable aleatoria está íntimamente relacionado con el concepto de función.

“Decimos que un objeto X o $X(\cdot)$ es una función definida en un espacio Ω , si para todo elemento ω de Ω hay un número real, denotado por $X(\omega)$, al cual llamamos el valor de la función X en ω .”¹³

Ahora veamos la definición de variable aleatoria:

“Decimos que un objeto X es una variable aleatoria si:

- i. Es una función de valores reales definida en un espacio de descripciones muestrales, sobre una familia de cuyos subconjuntos hayamos definido una función de probabilidad $P[\cdot]$.*
- ii. Para todo conjunto boreliano \wp de números reales, el conjunto $\{\omega : X(\omega) \in \wp\}$ pertenece al dominio de $P[\cdot]$.”¹⁴*

La variable X no recibe el calificativo de aleatoria por el hecho que atribuya de modo imprevisible un valor cualquiera a un elemento ω de Ω , ya que este valor está definido de forma precisa (determinista). Lo que es aleatorio en realidad, es que al hacer el experimento, no sabemos que elemento de Ω puede ocurrir.

2.4.2 Descripción de una Variable Aleatoria

Por definición una variable aleatoria X es una función sobre un espacio de probabilidades, la forma funcional de X raramente nos preocupa, puesto que no nos interesa calcular el valor de $X(\omega)$ que la función X asume en un elemento individual ω del espacio de descripciones muestrales Ω , sobre el cual se ha definido X , más bien, nos interesa la probabilidad de que un valor observado de la variable aleatoria X pertenezca a un conjunto dado \wp .

¹³ Ver, Parzen “Teoría Moderna de Probabilidades y sus Aplicaciones”, p.300

¹⁴ Ver, Parzen “Teoría Moderna de Probabilidades y sus Aplicaciones”, p.300

2.4.2.1 Función de Probabilidad de Una Variable Aleatoria

La función de probabilidad de una variable aleatoria X , comúnmente denotada como $P_X[\cdot]$, es una función de conjunto¹⁵ definida para todo conjunto boreliano \wp de números reales cuyo valor $P_X[\wp]$ es la probabilidad de que X este en \wp . Las siguientes expresiones se usan de manera similar que $P_X[\wp]$. La función de probabilidad $P_X[\wp]$ de la variable aleatoria X se obtiene de la función de probabilidades $P[\cdot]$, que existe en el espacio de descripciones muestrales de Ω sobre el cual se ha definido X como una función, por medio de la siguiente fórmula básica:

$$P_X[\wp] = P[\{\omega : X(\omega) \in \wp\}]$$

La anterior ecuación representa la definición de $P_X[\wp]$. El significado intuitivo de $P_X[\wp]$ anteriormente expresado está implícito en la ecuación, puesto que la función X tendrá un valor observado perteneciente al conjunto \wp si, y sólo si, el valor observado de ω del fenómeno aleatorio es tal que $X(\omega)$ esta en \wp .

2.4.2.2 Ley de Probabilidad de una Variable Aleatoria

La ley de probabilidades de una variable aleatoria X se define como una función de probabilidades $P_X[\cdot]$ en la recta real, que coincida con la función de probabilidades $P_X[\wp]$ de la variable aleatoria X . La teoría de probabilidades estudia las aseveraciones que se pueden hacer acerca de una variable aleatoria, cuando se conoce únicamente su ley de probabilidades. En consecuencia, si enunciamos una proposición acerca de una función de probabilidades $P[\cdot]$, entonces, desde el punto de vista de la teoría de probabilidades, enunciaremos una proposición acerca de todas las variables aleatorias X, Y, \dots , cuyas funciones de probabilidad $P_X[\cdot], P_Y[\cdot], \dots$ coincidan con $P[\cdot]$. Dos variables aleatorias X y Y están idénticamente distribuidas si sus funciones de probabilidad son iguales; es decir, si $P_X[\wp] = P_Y[\wp]$ para todos los conjuntos borelianos \wp .

¹⁵ “Una función de conjunto es una función real cuyo dominio es una colección de conjuntos.” Ver <<http://www.mfc.uclv.edu.cu/>>

2.4.2.3 Función de Distribución¹⁶

"La función de distribución de una variable aleatoria X , denotada por $F_X(\bullet)$, para todo número real x , se define por,

$$F_X(x) = P[X \leq x].^{17}$$

Las propiedades de una función de distribución $F_X(\bullet)$ son:

- i. $F_X(\bullet)$ debe ser no decreciente, en el sentido de que cualesquiera números reales a y b se verifica que,

$$F_X(a) \leq F_X(b) \text{ si } a < b;$$

- ii. Los límites de $F_X(x)$, cuando x tiende a ya sea a más o menos infinito, debe existir y ser

$$\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow +\infty} F_X(x) = 1;$$

- iii. En cualquier punto x ,

$$\lim_{b \rightarrow x^+} F_X(b) = F_X(x)$$

- iv. En cualquier punto x ,

$$\lim_{a \rightarrow x^-} F_X(a) = F_X(x) - p(x)$$

donde $p(x)$ es la probabilidad de que el valor observado del fenómeno aleatorio sea igual a x .

La función de distribución de X determina de manera única la función de probabilidades de X .

La función de masa de probabilidades de una variable aleatoria X , denotada por $p_X(\bullet)$, es una función cuyo valor $p_X(\bullet)$, en cualquier número real x , representa la probabilidad de que el valor observado de la variable aleatoria X sea igual a x , es decir,

$$p_X(x) = P[X = x] = P_X[\{x' : x' = x\}]$$

¹⁶ Frecuentemente también llamada *FUNCIÓN DE DISTRIBUCIÓN ACUMULADA*.

¹⁷ Ver Parzen, "Teoría Moderna de Probabilidades y sus Aplicaciones", p.303

Un punto de masa de probabilidades de la variable aleatoria X es un número real x para el cual $p_X(\bullet)$ es positiva. A partir de la función de distribución $F_X(\bullet)$ obtenemos la función de masa de probabilidades $p_X(\bullet)$ mediante

$$p_X(x) = F_X(x) - \lim_{a \rightarrow x^-} F_X(a)$$

Una variable aleatoria X es discreta si la suma de la función de masa de probabilidades sobre los puntos en los cuales es positiva es igual a 1, es decir,

$$\sum_{x, p_X(x) > 0} p_X(x) = 1$$

En otras palabras, una variable aleatoria X será discreta cuando distribuyamos una unidad de masa en la recta infinita de acuerdo con la función de probabilidades $p_X[\bullet]$, y si al hacer esto asignamos una masa positiva $p_X(x)$ a cada número finito o infinito numerable de puntos.

Si una variable aleatoria X es discreta, basta conocer su función de masa de probabilidades $p_X(\bullet)$ para conocer su función de probabilidades $P_X[\bullet]$, ya que $P_X[\bullet]$ puede expresarse en términos de $p_X(\bullet)$.

"Si X es discreta, entonces, para cualquier conjunto boreliano \wp de números reales, tenemos que:

$$P_X[\wp] = P[X \text{ esté en } \wp] = \sum_{x \in \wp, p_X(x) > 0} p_X(x). \text{ }^{18}$$

Así pues, para evaluar la probabilidad $P_X[\wp]$, con respecto a una variable aleatoria X discreta, de que la variable aleatoria X , tenga un valor observado de esté en \wp , no tenemos sino que enumerar los puntos de masa de probabilidad de X que estén en \wp . Sumamos entonces las masas de probabilidad adjudicadas a estos puntos de masa de probabilidad para obtener $P_X[\wp]$.

La función de distribución de una variable aleatoria discreta X se expresa en términos de su función de masa de probabilidades como:

¹⁸ Ver Parzen, "Teoría Moderna de probabilidades y sus Aplicaciones", p. 304

$$F_X(x) = \sum_{x' \leq x, p_X(x') > 0} p_X(x').$$

La función de distribución $F_X(\cdot)$ de una variable aleatoria X discreta es una función constante por intervalos o función *escalonada*. Consiste en una serie de rectas horizontales en los intervalos comprendidos entre los puntos de masa de probabilidades; en un punto de masa de probabilidades x , la gráfica de $F_X(\cdot)$ da un salto hacia arriba igual a $p_X(x)$.

Se dice que una variable aleatoria X es continua cuando toma valores en cualquier punto de un intervalo (a, b) de la recta real. Si una variable aleatoria X es continua, existe una función no negativa $f_X(\cdot)$, llamada función de densidad de probabilidades de la variable aleatoria X , que tiene la propiedad siguiente, para cualquier conjunto boreliano \wp de números reales.

$$P_X[\wp] = P[X \in \wp] = \int_{\wp} f_X(x) dx$$

Es decir, una vez que se conoce la función de densidad de probabilidades $f_X(\cdot)$, para una variable aleatoria continua X , obtenemos el valor $P_X[\wp]$ de la función de probabilidades en cualquier conjunto boreliano \wp al integrar la función de densidad de probabilidades $f_X(\cdot)$ sobre el conjunto \wp .

La función de distribución $F_X(\cdot)$ de una variable aleatoria continua se obtiene, en términos de su función de densidad de probabilidades, mediante

$$F_X(x) = \int_{-\infty}^x f_X(x') dx'$$

A su vez, la función de densidad de probabilidades de una variable continua se obtiene mediante la derivación de su función de distribución,

$$f_X(x) = \frac{d}{dx} F_X(x)$$

en todos los puntos x en que exista la derivada del miembro derecho en la anterior fórmula.

También la función de distribución sirve para clasificar las variables aleatorias según sus tipos. Diremos que una variable aleatoria X es discreta o continua si su función de distribución $F_X(\cdot)$ es discreta o continua, respectivamente. Se le llama discreta a una ley de probabilidades si corresponde a una función de distribución discreta, y continua si corresponde a una función de distribución continua.

Las leyes de probabilidades pueden clasificarse en familias tomando formas funcionales similares. Por ejemplo, considere la función $b(\cdot; n, p)$ definida para cualquier $n=1, 2, \dots$, y $0 \leq p \leq 1$ por

$$b(x, n, p) = \binom{n}{x} p^x q^{n-x} \text{ para } x=0, 1, \dots, n$$

$$= 0 \text{ en otro caso.}$$

Para valores fijos de n y p la función $b(\cdot; n, p)$ es una función de masa de probabilidades, y por ello define una ley de probabilidades. Las leyes de probabilidades determinadas por $b(\cdot; n_1, p_1)$ y $b(\cdot; n_2, p_2)$ para dos conjuntos diferentes de valores n_1, p_1 y n_2, p_2 son diferentes. No obstante, la forma funcional común de las dos funciones $b(\cdot; n_1, p_1)$ y $b(\cdot; n_2, p_2)$ nos permite tratar simultáneamente las dos leyes de probabilidades que determinan. Llamamos a n y p parámetros, y a $b(\cdot; n, p)$ la función de masa de la ley de probabilidades binomial con parámetros n y p . En los anexos 1 y 2 presentamos un inventario de las leyes de probabilidades discretas y continuas que son más usuales.

2.5 Esperanza de una Variable Aleatoria

Al estudiar las variables aleatorias, es tan importante conocer sus medias y variancias como lo es conocer sus leyes de probabilidades. Definiremos el concepto de esperanza de una variable aleatoria.

2.5.1 Esperanza, Media y Varianza de una Variable Aleatoria

Dada la variable aleatoria X , definimos la esperanza de una variable aleatoria, denotada por $E[X]$, como la media de la ley de probabilidades de X ; es decir,

$$E[X] = \begin{cases} \int_{-\infty}^{\infty} x dF_X(x) \\ \int_{-\infty}^{\infty} x f_X(x) dx \\ \sum_x x p_X(x) \end{cases}$$

Según se especifique X mediante su función de distribución $F_X(\cdot)$, su función de densidad de probabilidad $f(\cdot)$, o su función de masa de probabilidades $p_X(\cdot)$.

Dada una variable aleatoria Y , que surge como una función boreliana¹⁹ de una variable aleatoria X , de manera que,

$$Y = g(X).$$

Para una función boreliana $g(\cdot)$, la esperanza $E[g(X)]$, de acuerdo a la definición de esperanza, está dada por,

$$E[g(X)] = \int_{-\infty}^{\infty} y dF_{g(x)}(y)$$

Por otra parte dada la función de Borel $g(\cdot)$ y la variable aleatoria X , podemos definir la esperanza de $g(x)$ con respecto a la ley de probabilidades de X , denotada por $E_X[g(x)]$, como

$$E_X[g(x)] = \begin{cases} \int_{-\infty}^{\infty} g(x) dF_X(x) \\ \int_{-\infty}^{\infty} g(x) f_X(x) dx \\ \sum_x g(x) p_X(x) \end{cases}$$

Según se defina X mediante su función de distribución $F_X(\cdot)$, su función de densidad de probabilidad $f_X(\cdot)$, o su función de masa de probabilidades $p_X(\cdot)$. Un hecho de gran importancia, que para cualquier variable aleatoria X y función boreliana $g(\cdot)$, se tiene, $E[g(X)] = E_X[g(x)]$.

¹⁹ Una función $f(\cdot)$ es una función boreliana si para cualquier número real c , el conjunto $\{x: f(x) < c\}$ es un conjunto boreliano, ver Parzen "Teoría Moderna de Probabilidades y sus Aplicaciones", p. 175

Si existe alguna de estas esperanzas. La igualdad anterior dice que la esperanza de la variable aleatoria $g(X)$ es igual a la esperanza de la función $g(\cdot)$ con respecto a la variable aleatoria X . Así pues dadas una variable aleatoria X y una función $g(\cdot)$, encontramos dos conceptos distintos representados por $E[g(x)]$ y $E_X[g(x)]$ que, sin embargo, son siempre numéricamente iguales. Se acostumbra usar la notación $E[g(x)]$, por ser más conveniente para las manipulaciones técnicas. No hay que olvidar que aunque se escriba $E[g(x)]$, el concepto que nos interesa suele ser $E_X[g(x)]$, la esperanza de la función $g(x)$ con respecto a la variable aleatoria X .

Dada una variable aleatoria X , denotamos su media por $E[X]$, su media cuadrática por $E[X^2]$, el cuadrado de su media por $E^2[X]$, y su n -ésimo momento con respecto al punto c por $E[(X-c)^n]$, y su n -ésimo momento central por $E[(X-E[X])^n]$. En particular, la varianza de una variable aleatoria, denotada por $Var[X]$, se define como su segundo momento central, de manera que se verifica,

$$Var[X] = E[(X - E[X])^2] = E[X^2] - E^2[X].$$

La desviación estándar de una variable aleatoria, denotada por $\sigma[X]$, se define como la raíz cuadrada de positiva de su varianza, de manera que,

$$\sigma[X] = \sqrt{Var[X]}, \quad \sigma^2[X] = Var[X].$$

Las propiedades de la media y la varianza son, dadas X y Y variables aleatorias y $a, b \in \mathbb{R}$, se cumple para la media:

1. $E[aX + b] = aE[X] + b$;
2. $E[X + Y] = E[X] + E[Y]$.

Y para la varianza:

1. $Var[X] = E[X^2] - (E[X])^2$;
2. $Var[b] = 0$;

3. $Var[aX] = a^2 Var[X]$;
4. $Var[aX + b] = a^2 Var[X]$.

2.5.2 Función Generadora de Momentos

La función generadora de momentos de una variable aleatoria, denotada por $M_X(\bullet)$, se define para cualquier número real t mediante,

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_{x \in \mathbb{R}_x} e^{tx} p_X(x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx & \text{si } X \text{ es continua} \end{cases}$$

Siempre que el valor esperado exista para todo $t \in (-h, h)$. Esta última es una condición técnica necesaria para que $M_X(t)$ sea diferenciable en 0. Se denomina función generadora de momentos porque los momentos ($E(X^n)$) de X , pueden ser obtenidos derivando esta función y evaluando la derivada en $t = 0$, es decir,

$$E(X^n) = \left. \frac{\partial^n}{\partial t^n} M_X(t) \right|_{t=0}.$$

La función generadora de momentos además de permitir calcular los momentos de una variable aleatoria, también permite identificar la función de densidad o de probabilidad de una variable aleatoria debido a la propiedad de unicidad, la cual establece que hay una correspondencia uno a uno entre funciones de densidad o de probabilidad y funciones generadoras de momentos.

2.5.3 Función Característica

La función característica de una variable aleatoria X , denotada por $\phi_X(\bullet)$, se define para cualquier número real t mediante,

$$\phi_X(t) = E(e^{itX}) = E \cos(tX) + i \sin(tX).^{20}$$

Es decir, $\phi_X(t)$ es la esperanza de la variable aleatoria e^{itX} . Algunas de las propiedades más relevantes de esta función son:

1. La función característica está acotada por 1, es decir $|\phi_X(t)| \leq 1$, para todo t ;
2. $\phi_X(0) = 1$;
3. $\overline{\phi_X(t)} = \phi_X(-t)$, donde \bar{z} denota el conjugado complejo de z ;
4. $\phi_X(t)$ es uniformemente continua en \mathbb{R} ;
5. Si X y Y son variables aleatorias independientes, entonces $\phi_{X+Y} = \phi_X \phi_Y$;
6. La función característica determina la función de distribución; por lo tanto, $\phi_X(t) = \phi_Y(t)$ si y solo si $F_X(x) = F_Y(y)$. Esta es una consecuencia de la Fórmula de Inversión²¹.
7. Una variable aleatoria X posee una distribución simétrica si, y solo si, $\phi_X(t) \in \mathbb{R}$, para todo $t \in \mathbb{R}$.
8. Par los números reales a y b , $\phi_{aX+b}(t) = e^{itb} \phi_X(at)$;
9. Si $E|X|^n < \infty$, entonces $\phi_X(t)$ tiene n -derivadas continuas y

$$\frac{d^k \phi_X}{dt^k}(t) = \phi_X^{(k)}(t) = \int_{\mathbb{R}} (it)^k e^{itx} f_X(x) dx, 1 \leq k \leq n.$$

Particularmente, $\phi_X^{(k)}(0) = i^k EX^k$, las funciones características son similares a las funciones generadoras de momentos, en este sentido.

²⁰ Si F_X es la función de distribución asociada a X , por las propiedades de la esperanza tenemos, $\phi_X(t) = \int_{\mathbb{R}} e^{itx} dF_X(x)$, la cual se conoce como la transformada de Fourier-Stieltjes y provee una definición alternativa de la función característica.

²¹ Dada una variable aleatoria X con función característica ϕ y función de distribución F , si x y y son puntos continuos de F , tales que $x < y$, entonces

$$F(x) - F(y) = \frac{1}{2\pi} \lim_{\delta \rightarrow \infty} \int_{-\delta}^{\delta} \frac{e^{-itx} - e^{-ity}}{it} \phi(t) dt.$$

2.5.4 La Ley de los Grandes Números y el Teorema del Límite Central

En las aplicaciones de la teoría de probabilidades a fenómenos reales, hay dos resultados que son muy destacados. Conocemos estos resultados como la *Ley de los Grandes Números* y el *Teorema del Límite Central*.

Decimos que un conjunto de n observaciones X_1, X_2, \dots, X_n constituye una muestra aleatoria de una variable aleatoria X , si X_1, X_2, \dots, X_n son variables aleatoria independientes y están idénticamente distribuidas como X . Sea $S_n = X_1 + X_2 + \dots + X_n$ la suma de las observaciones, llamamos a su media aritmética $\frac{1}{n}S_n$, media muestral. Por medio de las definiciones antes mencionadas se tiene:

$$E[S_n] = nE[X], \quad \text{Var}[S_n] = n\text{Var}[X], \quad \psi_{S_n}(t) = [\psi_X(t)]^n$$

$$E[M_n] = E[X], \quad \text{Var}[M_n] = \frac{1}{n}\text{Var}[X], \quad \psi_{M_n}(t) = \left[\psi_X\left(\frac{t}{n}\right)\right]^n$$

Tenemos entonces que la media muestral $\frac{1}{n}S_n$ tiende a cero cuando el tamaño de la muestra n tiende a infinito. Ahora bien, por la desigualdad de Chebyshev²², sabemos que si una variable aleatoria tiene una varianza pequeña, entonces es aproximadamente igual a su media, en el sentido de que con una probabilidad cercana a 1, una observación de la variable aleatoria dará un valor observado aproximadamente igual a la media; en particular, la probabilidad de que un valor observado de la variable aleatoria esté dentro de tres desviaciones estándar de la media es de 0.997. Así pues, hemos establecido que la media muestral $\frac{1}{n}S_n$ de una muestra aleatoria X_1, X_2, \dots, X_n de una variable aleatoria, es aproximadamente igual a la media, con una probabilidad de que se acercara a 1 tanto como queramos si simplemente tomamos una muestra lo suficientemente grande. Este hecho conocido como la *Ley de los Grandes Números*, lo descubrió por vez primera Jakob Bernoulli a finales del siglo XVII. Por otra parte, además de que la media muestral $\frac{1}{n}S_n$ tiende a ser igual a la media $E[X]$, podemos demostrar otras propiedades de la

²² “Si una variable aleatoria X tiene una media finita μ y una varianza finita σ^2 , entonces para toda $k > 0$, $P(|x - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$ ”, ver <<http://mathworld.wolfram.com/ChebyshevInequality.html>>”

media muestral $\frac{1}{n}S_n$. Podemos evaluar aproximadamente, para cualquier intervalo alrededor de la media $E[X]$, la probabilidad de que la media muestral $\frac{1}{n}S_n$ tenga un valor observado en ese intervalo, puesto que la media muestral está distribuida aproximadamente como una distribución normal cuando el tamaño de la muestra es grande. De manera más general se puede decir que si S_n es la suma de las variables aleatorias independientes e idénticamente distribuidas, X_1, X_2, \dots, X_n , con medias y varianzas finitas, entonces para cualesquiera números reales,

$$P[a \leq S_n \leq b] = P\left[\frac{a - E[S_n]}{\sigma[S_n]} \leq \frac{S_n - E[S_n]}{\sigma[S_n]} \leq \frac{b - E[S_n]}{\sigma[S_n]}\right]$$

$$\doteq \Phi\left(\frac{b - E[S_n]}{\sigma[S_n]}\right) - \Phi\left(\frac{a - E[S_n]}{\sigma[S_n]}\right)$$

La anterior ecuación es una representación de uno de los teoremas más importantes de la teoría de probabilidades, en palabras este dice que si tenemos n variables aleatorias y todas ellas siguen el mismo modelo de distribución (cualquiera que éste sea), la suma S_n se distribuye según una distribución normal estándar, cuando n tienda al infinito. En 1920, George Polya dio a este teorema el nombre de *Teorema Central del Límite de la Teoría de Probabilidades*.

Hemos mencionado que la teoría de la probabilidad se encarga del estudio de los fenómenos aleatorios y propone modelos matemáticos para estos fenómenos, también asigna, por medio de una construcción axiomática, las probabilidades de ocurrencia de los eventos que genera el espacio de descripciones muestrales de estos fenómenos. Por lo tanto la probabilidad nos proporciona los modelos matemáticos los cuales emplearemos para modelar el monto de siniestros ocurridos.

CAPÍTULO 3

Métodos y Técnicas Estadísticas

"Llegaré el día en el que la Estadística será una condición tan necesaria para la convivencia como la capacidad de leer y escribir."

Anónimo

En el Año de 1760, el profesor de la Universidad de Göttingen, Gottfried Achenwall acuñó la palabra estadística, que extrajo del término italiano statista (estadista). La raíz remota de la palabra se halla, por otra parte, en el término latino status, que significa estado o situación; esta etimología aumenta el valor intrínseco de la palabra, por cuanto la estadística revela el sentido cuantitativo de las más variadas situaciones. Tomaremos la siguiente definición de estadística.

*“Ciencia que se ocupa del estudio de fenómenos de tipo genérico, normalmente complejos y enmarcados en un universo variable, mediante el empleo de modelos de reducción de la información y de análisis de validación de los resultados en términos de representatividad”*²³

La estadística para su mejor comprensión y estudio se ha dividido en varias ramas, entre las más destacadas tenemos: Estadística Clásica, Estadística Descriptiva, Estadística Inferencial y la Estadística Bayesiana. En el presente trabajo ahondaremos en la Estadística Descriptiva y la Estadística Inferencial.

*“La Estadística Descriptiva, describe, analiza y representa un grupo de datos utilizando métodos numéricos y gráficos que resumen y presentan la información contenida en ellos”*²⁴

*“La Estadística inferencial apoyándose en el cálculo de probabilidades y a partir de datos muestrales, efectúa estimaciones, decisiones, predicciones u otras generalizaciones sobre un conjunto mayor de datos”*²⁵

3.1 Métodos Descriptivos

La estadística descriptiva es la encargada de resumir, clasificar y ordenar los datos, con el fin de tener una visión más precisa y conjunta de las observaciones, intentando descubrir de esta manera posibles relaciones entre los datos, observando cuales toman valores parecidos, cuales difieren grandemente del resto, destacando hechos de posible interés

²³ Ver <<http://www.geocities.com/ResearchTriangle/Facility/1075/Estad.htm>>

²⁴ Ver <<http://ftp.medprev.uma.es/libro/node3.htm>>

²⁵ Ver <<http://ftp.medprev.uma.es/libro/node3.htm>>

3.1.1 Distribución de Frecuencias

Un arreglo de grandes conjuntos de datos nos presenta una imagen valiosa. Así, se podrían ordenar los datos dentro de clases o categorías. A esta forma de agrupar los datos en forma tabular se le llama una distribución de frecuencia.

*"Una distribución de frecuencia es una tabla resumen en la cual los datos son agrupados dentro de clases o categorías ordenadas numéricamente y convenientemente establecidas"*²⁶

Cuando los datos son ordenados y agrupados dentro de tablas de distribución de frecuencia, el proceso de análisis e interpretación de la información se vuelve mucho más fácil y entendible. Una desventaja de las tablas resumen es que no se puede saber como se distribuyen los datos individuales en un intervalo de clase en particular sin acceder a los datos originales. El número de observaciones en cualquier clase, es la frecuencia para esa clase. Las frecuencias de clase relativas se encuentran dividiendo las frecuencias de clase entre el número total de observaciones. Una distribución porcentaje se obtiene multiplicando las frecuencias de clase relativas por 100. Una distribución de frecuencia acumulada representa en cada intervalo de clase la suma de su predecesor hasta sumar el 100% de las observaciones.

3.1.2 Presentación Gráfica

La presentación de la información por medio de una distribución de frecuencias resulta efectiva para mostrar las características más sobresalientes de un grupo de datos; la presentación gráfica de la misma información, hace a las características más importantes inmediatamente aparentes. Entre las formas gráficas más importantes tenemos:

HISTOGRAMAS. Un histograma es una representación gráfica de una distribución de frecuencia, y se forma construyendo barras sobre los intervalos de clase de una distribución de frecuencias. Cuando se dibuja un histograma, los valores de la variable o fenómeno de interés se registran a lo largo de la escala horizontal, marcando las fronteras de clase. A través de la escala vertical se marca el número, proporción, frecuencia o porcentaje de las observaciones por cada intervalo de clase.

²⁶ Ver Berenson, "Applied Statistics: A First Course", p. 79

POLÍGONOS DE FRECUENCIA. Otra de las formas gráficas de representar una variable aleatoria continua es por medio de un polígono de frecuencias.

*"Un polígono de frecuencias es una representación de la frecuencia o frecuencia relativa asociada con cada intervalo de una variable continua por medio de la unión de los puntos de las alturas en una gráfica de líneas"*²⁷

OJIVAS. Así como una distribución de frecuencias puede representarse gráficamente por un histograma, una distribución de frecuencias acumulada puede representarse gráficamente por medio de una "ojiva". Para construir una ojiva, primero señalamos las fronteras de clase en la escala horizontal, tal como se hace con el histograma. Sobre cada frontera de clase dibujamos un punto a una distancia vertical proporcional a la frecuencia acumulativa. Estos puntos se conectan entonces por líneas rectas. A la ojiva obtenida frecuentemente también se llama "polígono de frecuencias acumuladas".

3.1.3 Medidas de Tendencia Central

Las medidas de tendencia central se refieren a los estadísticos descriptivos, los cuales indican el valor más representativo en una distribución de frecuencias.

LA MODA. Es el valor o intervalo en una distribución de frecuencias con la más alta frecuencia. En una distribución de frecuencia para datos sin agrupar, la moda es el valor que ocurre más comúnmente, el valor asociado con la frecuencia más alta. Para una distribución de frecuencia para datos agrupados la moda es el punto medio del intervalo con la más alta frecuencia.

LA MEDIA. Para las observaciones x_1, x_2, \dots, x_n es el promedio aritmético de éstas y se denota por

$$\bar{x} = \sum_{i=1}^n x_i / n$$

La Media es una medida apropiada de tendencia central para muchos conjuntos de datos, sin embargo, dado que cualquier observación en el conjunto se emplea para su cálculo, el valor de la media puede afectarse de manera desproporcionada por la existencia de algunos valores extremos. Para calcular la media con base en los datos agrupados, sea k el número de clases y x_i el punto medio de la i -ésima clase. Entonces el valor aproximado de la media es,

²⁷ Ver McGrath, "Understanding Statistics: A Research Perspective", p. 52

$$\bar{x} = \sum_{i=1}^k f_i x_i / n$$

en donde f_i es la frecuencia de la i -ésima clase y $n = \sum_{i=1}^k f_i$

LA MEDIANA. Para un conjunto de observaciones es el valor para el cual, cuando todas la observaciones se ordenan de manera creciente, la mitad de éstas es menor que este valor y la otra mitad mayor. Para datos agrupados, la mediana es aquel valor que divide en dos partes iguales la distribución de frecuencia relativa.

$$\text{Mediana} = L + c(j / f_m)$$

en donde L es el límite inferior de la clase donde se encuentra la mediana, f_m es la frecuencia de esa clase, c es la longitud de la clase y j es el número de observaciones en esta clase, necesarias para completar un total de $n/2$.

3.1.4 Medidas de Variabilidad

Las medidas de variabilidad describen el monto de variabilidad en una distribución de frecuencias. Entre las más importantes tenemos:

VARIANZA. La varianza puede ser definida como la diferencia del cuadrado de los promedios y su media. La versión comúnmente más usada para el cálculo de la varianza es con la fórmula,

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

La varianza es una medida razonablemente buena de la variabilidad debido a que si muchas de las diferencias son grandes (o pequeñas) entonces el valor de la varianza será grande (o pequeño). El valor de la varianza puede sufrir un cambio muy desproporcionado, aún más que la media, por la existencia de algunos valores extremos en el conjunto.

DESVIACIÓN ESTÁNDAR. La raíz cuadrada positiva de la varianza recibe el nombre de desviación estándar y se denota por

$$s_x = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}$$

La varianza y la desviación estándar no son medidas de variabilidad distintas, debido a que la última no puede determinarse a menos que se conozca la primera. A menudo se prefiere la desviación estándar en relación con la varianza, por que se expresa en las mismas unidades físicas de las observaciones.

3.2 Inferencia Estadística

La Inferencia Estadística es la parte de la estadística matemática que se encarga del estudio de los métodos para la obtención del modelo de probabilidad (forma funcional y parámetros que determinan la función de distribución) que sigue una variable aleatoria de una determinada población, a través de una muestra (parte de la población) obtenida de la misma. Los dos problemas fundamentales que estudia la inferencia estadística son: el problema de la *estimación* y el problema del *contraste de hipótesis*.

En lo que sigue nos vamos a limitar a problemas de inferencia estadística paramétrica, donde la variable aleatoria objeto de estudio sigue una distribución conocida, y sólo tendremos que tratar de estimar los parámetros que la determinan. Esta situación se presenta con frecuencia debido a que es posible a menudo conocer la forma funcional de la distribución de probabilidad, por consideraciones teóricas, quedando únicamente indeterminados los parámetros que determinan la función de distribución.

Como las poblaciones en las que se pretende estudiar una determinada variable aleatoria son grandes, es muy caro o imposible, estudiar a todos sus individuos; lo que se hace, es estudiar una muestra (una parte) de la población. En todos estos problemas que estudia la inferencia estadística juega un papel fundamental la *Teoría de la Probabilidad* (distintas formas funcionales de las distribuciones de probabilidad) y la *Teoría de Muestras* (procedimientos para tomar muestras de manera apropiada).

3.2.1 Estimación

Existen dos tipos de estimaciones, la estimación puntual y por intervalos. El procedimiento puntual utiliza la información de la muestra para obtener un solo número o punto que estima el parámetro. El procedimiento de estimación por intervalos hace uso de la información de la muestra para obtener un intervalo que se supone va a incluir el parámetro en estudio. En cada caso la estimación real se hace mediante un estimador, que es una regla que establece como utilizar los datos de la muestra para determinar el valor (o valores) que utilizamos como la estimación puntual (o por intervalos).

3.2.1.1 Estimadores y Estimaciones

Llamamos problema de ajuste al problema que consiste en determinar, en una familia de leyes de probabilidad dada, aquella que coincide mejor con la muestra observada. En los casos más usuales, la familia depende de uno o dos parámetros reales desconocidos. Por tanto el problema consiste en determinar el valor del parámetro que se adapta mejor a los datos. Hablamos entonces de estimación paramétrica.

Suponemos que una familia de leyes, que depende de un parámetro desconocido θ , ha sido seleccionada. Ahora es de la muestra, y sólo de ella, que se puede extraer la información. Se llama estimador del parámetro θ a toda función de la muestra, que toma valores en el conjunto de los valores posibles de θ . Es importante distinguir entre los valores aleatorios, asociados a la modelación, y sus realizaciones, identificadas a los datos: Una muestra (teórica) es una n -tupla de variables aleatorias independientes (X_1, \dots, X_n) , que siguen una misma ley P_θ . Para estimar, se propone un estimador en función de la muestra:

$$T = \tau(X_1, \dots, X_n).$$

T es también una variable aleatoria. La selección del modelo y del estimador T , está desconectada de la recolección de los datos. Es, en cierta forma, una planificación que se hace antes de realizar las observaciones y que podrá servir a varias muestras que se recojan del mismo fenómeno.

Una vez que se selecciona el modelo, se considerará a una n -tupla de datos (x_1, \dots, x_n) como una realización de las variables aleatorias (X_1, \dots, X_n) . El valor (real) que toma T :

$$\hat{\theta} = \tau(x_1, \dots, x_n),$$

es la estimación del parámetro (a partir de la muestra observada).

3.2.1.2 Cualidades de un Estimador

Intuitivamente, las características que serían deseables para esta nueva variable aleatoria (que usaremos para estimar el parámetro desconocido) deben ser:

- **Consistencia:** Cuando el tamaño de la muestra crece arbitrariamente, el valor estimado se aproxima al parámetro desconocido.
- **Carencia de sesgo:** También llamado Insesgamiento. El valor medio que se obtiene de la estimación para diferentes muestras debe ser el valor del parámetro.
- **Eficiencia:** Al estimador, al ser v.a., no puede exigírsele que para una muestra cualquiera se obtenga como estimación el valor exacto del parámetro. Sin embargo, podemos pedirle que su dispersión con respecto al valor central (varianza) sea tan pequeña como sea posible.
- **Suficiencia:** El estimador debería aprovechar toda la información existente en la muestra.

"Decimos que el estimador (T_n) es consistente (o convergente), si para todo $\varepsilon > 0$:

$$\lim_{n \rightarrow \infty} P[|T_n - \theta| > \varepsilon] = 0 \text{ }^{28}$$

En consecuencia un estimador consistente se aleja del parámetro con una probabilidad débil, si el tamaño de la muestra es lo suficientemente grande.

La noción de consistencia no da ninguna seguridad práctica de que los valores que toma un estimador estarán efectivamente en un radio fijo alrededor del verdadero valor del parámetro, para un tamaño de muestra dado. La calidad de los estimadores se cuantifica con la noción de error cuadrático.

²⁸ <http://www.math-info.univ-paris5.fr/~ycart/emel/cours/ep/node4.html>

“Sea T cualquier estimador de un parámetro desconocido θ , se define al error cuadrático medio de T como el valor esperado del cuadrado de la diferencia entre T y θ , es decir:

$$ECM(T_n, \theta) = E[(T_n - \theta)^2].”^{29}$$

La razón por la cual el ECM es una cantidad importante para enjuiciar a los posibles estimadores de θ se debe a la siguiente argumento, desarrollando el miembro derecho de la anterior relación, obtenemos:

$$ECM(T_n, \theta) = Var(T_n) + [\theta - E(T_n)]^2$$

El error cuadrático medio de cualquier estimador es la suma de dos cantidades no negativas: una es la varianza del estimador y la otra es el cuadrado del sesgo ³⁰ del estimador. Estas dos cantidades se encuentran relacionadas en forma directa con las propiedades deseables de un estimador. De manera específica, la varianza de un estimador debe ser lo más pequeña posible mientras que la distribución de muestreo debe concentrarse alrededor del valor del parámetro. El error cuadrático está ligado a la consistencia por la siguiente proposición.

“Si el error cuadrático de T_n con respecto a θ tiende a 0 cuando n tiende a infinito, entonces T_n es un estimador consistente de θ .”³¹

Si se dispone de dos estimadores para el mismo parámetro θ , diremos que uno es mejor que el otro si su error cuadrático con respecto a θ es menor. Aún para un estimador consistente, puede suceder que los valores que toma estén desplazados, en promedio, con respecto al verdadero valor del parámetro. Decimos entonces que el estimador es sesgado. El sesgo de θ puede ser positivo, negativo, o cero. Ya que el cuadrado del sesgo es un componente del error cuadrático medio, se desearía, que el valor absoluto de este sea lo más pequeño posible, es decir, es deseable que un estimador tenga un medida igual a al del parámetro que se esta estimando.

“Se dice que una estadística $T = u(X_1, X_2, \dots, X_n)$ es un estimador insesgado del parámetro θ , si la $E(\theta) = \theta$ para todos los posibles valores de θ . De esta forma, para

²⁹ Ver Canavos, “Probabilidad y Estadística. Aplicaciones y Métodos”, p

³⁰ En $ECM(T_n, \theta)$, el término $[\theta - E(T_n)]$ recibe el nombre de sesgo del estimador, denotado como $b(\theta)$

³¹ Ver <<http://www.math-info.univ-paris5.fr/~ycart/emel/cours/ep/node4.html>>

cualquier estimador insesgado de θ , la distribución de muestreo se encuentra centrada alrededor de θ y $ECM(T_n, \theta) = Var(T_n)$."³²

Decimos que un estimador T_n es asintóticamente insesgado, si $b(\theta)$ tiende a 0 cuando n tiende a infinito.

Se dice que un estimador $\hat{\theta}$ es suficiente si utiliza toda la información en una muestra relevante para la estimación de θ , es decir, si todo el conocimiento acerca de θ que se pueda ganar de los valores individuales de la muestra y de su orden se pueden ganar igualmente del valor de $\hat{\theta}$ solo. Podemos describir esta propiedad al referirnos a la distribución o densidad de probabilidad condicional de los valores muestrales dado $\hat{\theta}$, que esta dada por:

$$f(x_1, x_2, \dots, x_n | \hat{\theta}) = \frac{f(x_1, x_2, \dots, x_n, \hat{\theta})}{g(\hat{\theta})} = \frac{f(x_1, x_2, \dots, x_n)}{g(\hat{\theta})}$$

Si depende de θ , entonces los valores particulares de X_1, X_2, \dots, X_n que producen $\hat{\theta}$ serán más probables para algunos valores de θ que para otros, y el conocimiento de estos valores muestrales ayudará en la estimación de θ . Por otra parte, si no depende de θ , entonces los valores particulares de X_1, X_2, \dots, X_n que producen $\hat{\theta}$ serán igualmente probables para cualquier valor de θ , y el conocimiento de estos valores muestrales no ayudarán en la estimación de θ .

*"La estadística $\hat{\theta}$ es un estimador suficiente del parámetro θ si y sólo si para cada valor de $\hat{\theta}$ la distribución o densidad de probabilidad condicional de la muestra aleatoria X_1, X_2, \dots, X_n , dado $\hat{\theta}$, es independiente de θ ".*³³

Un criterio para determinar un estadístico suficiente se conoce como el teorema de factorización de Neyman.

³² Ver Canavos, "Probabilidad y Estadística. Aplicaciones y Métodos", p. 255

³³ Ver Freud, "Estadística Matemática con Aplicaciones", p. 337

“Sea T un estadístico basado en una muestra aleatoria X_1, X_2, \dots, X_n . Entonces T es un estadístico suficiente para la estimación de un parámetro θ si y sólo si la función de verosimilitud se puede factorizar en dos funciones no negativas de la siguiente forma:

$$L(x_1, x_2, \dots, x_n) = g(t, \theta)h(x_1, x_2, \dots, x_n)$$

en donde $g(t, \theta)$ es una función solamente de t y θ , $h(x_1, x_2, \dots, x_n)$ no es una función de θ .”³⁴

La utilidad de la estadística suficiente recae en el hecho de que si un estimador insesgado de un parámetro θ es una función de una estadística suficiente, entonces tendrá la varianza más pequeña de entre todos los estimadores insesgados de θ que no se encuentren basados en una estadística suficiente.

3.2.1.3 Intervalos de confianza

Un estimador por intervalo es una regla que especifica el método que utilizan las mediciones de la muestra para calcular dos números que forman los extremos del intervalo. Sería convenientemente ideal que el intervalo tuviera dos propiedades. Primero, que el intervalo contenga al parámetro θ a estimar. Segundo, que el intervalo sea relativamente estrecho. Nótese que uno o ambos extremos del intervalo variarán de manera aleatoria de una muestra a otra, porque son funciones de las mediciones de la muestra. Así la longitud y la localización del intervalo son cantidades aleatorias, y no podemos estar seguros que el parámetro θ a estimar (fijo) se localice realmente entre los extremos de cualquier intervalo calculado a partir de una sola muestra. Dada esta situación, el objetivo es encontrar un estimador por intervalos que genere intervalos angostos que contengan a θ con una alta probabilidad.

Los estimadores por intervalos se denominan comúnmente intervalos de confianza. Los extremos superior e inferior de un intervalo de confianza se llaman límites de confianza superior e inferior, respectivamente. La probabilidad de que un intervalo de confianza contenga a θ se conoce como coeficiente de confianza. Desde un punto de vista práctico, el coeficiente de confianza indica la fracción de veces, en un muestreo repetitivo, que los intervalos construidos contendrán el parámetro θ . Si se sabe que el coeficiente de confianza asociado a un estimador es alto, estaremos altamente confiados que un intervalo de confianza particular, construido a partir de una sola muestra, contendrá θ .

³⁴ Ver Freud, “Estadística Matemática con Aplicaciones”, p. 340

Suponiendo que $\hat{\theta}_L$ y $\hat{\theta}_U$ son los límites de confianza inferior y superior, respectivamente, para un parámetro θ . Entonces, si

$$P(\hat{\theta}_L < \theta < \hat{\theta}_U) = 1 - \alpha$$

la probabilidad $1 - \alpha$ es el coeficiente de confianza. El intervalo aleatorio resultante, definido por $\hat{\theta}_L$ hasta $\hat{\theta}_U$, se denomina intervalo de confianza bilateral. También es posible construir un intervalo de confianza unilateral tal que $P(\hat{\theta}_L < \theta) = 1 - \alpha$, aunque solo un punto es aleatorio en este caso, el intervalo de confianza es $(\hat{\theta}_L, \infty)$. De manera similar, podríamos tener un intervalo de confianza unilateral superior tal que, $P(\theta < \hat{\theta}_U) = 1 - \alpha$, el intervalo de confianza correspondiente es $(-\infty, \hat{\theta}_U)$.

Un método útil para obtener los intervalos de confianza se denomina método del pivote. Este método depende de la determinación de una expresión pivote que posee dos características:

- Es una función de las mediciones de la muestra y el parámetro desconocido θ , en donde θ es la única cantidad desconocida
- Tiene una distribución de probabilidad que no depende del parámetro θ .

Si se conoce la distribución de probabilidad de la cantidad pivote, entonces se puede utilizar la lógica siguiente para obtener el intervalo deseado de estimación. Si Y es una variable aleatoria, c una constante ($c > 0$) y $P(a \leq Y \leq b) = 0.7$, entonces $P(ca \leq cY \leq cb) = 0.7$. De manera similar, para cualquier constante d , $P(a + d \leq Y + d \leq b + d) = 0.7$, es decir, la probabilidad del evento $P(a < Y < b)$ no se altera por un cambio de escala o por una traslación de Y . Por lo tanto, si conocemos la distribución de probabilidad de una cantidad pivote, podemos aplicar a las operaciones descritas anteriormente para obtener el estimador por intervalos deseado.

3.2.2 Métodos de Estimación

En la sección anterior señalábamos algunas de las propiedades deseables de un estimador, ahora presentaremos dos de los métodos más usados para obtener estos estimadores. Estos métodos son: El método de los momentos y el de máxima verosimilitud.

3.2.2.1 Método de los Momentos

El método de los momentos es un procedimiento muy sencillo y quizá el más antiguo para encontrar un estimador para uno o más parámetros poblacionales. Recuérdese que el k -ésimo momento de una variable aleatoria, tomado con respecto al origen, es $\mu'_k = E(X^k)$. El correspondiente k -ésimo momento de la muestra es el promedio $m'_k = \frac{1}{n} \sum_{i=1}^n X_i^k$. El método de los momentos se basa en el supuesto de que los momentos de la muestra deben proporcionar estimaciones apropiadas para los momentos correspondientes de la población.

*"Eljase como estimaciones aquellos valores de los parámetros que son soluciones de las ecuaciones $\mu'_k = m'_k$, $k=1,2,\dots,t$, en donde t es igual al número de parámetros."*³⁵

Como los estimadores obtenidos por el método de momentos son, obviamente, funciones de los momentos muestrales, los estimadores son generalmente consistentes.

3.2.2.2 El Método de Máxima Verosimilitud

En dos artículos publicados a principios de este siglo, el prominente estadístico R. A. Fisher, propuso un método general de estimación llamado el método de máxima verosimilitud. También mostró las ventajas de este método al demostrar que producía estimadores suficientes siempre que estos existieran y que los estimadores de máxima verosimilitud son estimadores asintóticamente insesgado de varianza mínima.

"Si x_1, x_2, \dots, x_n son los valores de una muestra aleatoria de una población con el parámetro θ , la función de verosimilitud de una muestra está dada por:

$$L(\theta) = f(x_1, x_2, \dots, x_n; \theta)$$

*para los valores de θ dentro de un dominio dado. En este caso $f(x_1, x_2, \dots, x_n; \theta)$, es el valor de la distribución de probabilidad conjunta o de la densidad de probabilidad conjunta de las variables aleatorias X_1, X_2, \dots, X_n en $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$."*³⁶

³⁵ Ver Mendenhall, "Estadística Matemática con Aplicaciones", p. 386

³⁶ Ver Freund, "Estadística Matemática con Aplicaciones", p. 346

En esencia, el método de estimación por máxima verosimilitud selecciona como estimador a aquel valor del parámetro que tiene la propiedad de maximizar el valor de la probabilidad de la muestra aleatoria observada.

“Sea $L(\theta) = (\theta, x_1, \dots, x_n)$ la función de verosimilitud de las variables aleatorias X_1, \dots, X_n de una función (densidad) de probabilidad $f(x, \theta)$. Si $t = u(x_1, \dots, x_n)$ es el valor de θ , el cual maximiza $L(\theta)$, entonces $T = u(X_1, \dots, X_n)$ es el estimador de máxima verosimilitud de θ y t es el estimador correspondiente de máxima verosimilitud”

Las principales propiedades de los estimadores obtenidos por el método antes descrito son:

- Son consistentes
- Son invariantes frente a las transformaciones biunívocas, es decir, si $\hat{\theta}_{MV}$ es el estimador de máxima verosimilitud de θ y $g(\hat{\theta})$ es una función biunívoca de $\hat{\theta}$, entonces $g(\hat{\theta}_{MV})$ es el estimador de máxima verosimilitud de $g(\theta)$.
- Si $\hat{\theta}$ es un estimador suficiente de θ , su estimador de máxima verosimilitud, $\hat{\theta}_{MV}$ es función de la muestra a través de $\hat{\theta}$
- Son asintóticamente normales;
- Son asintóticamente eficientes, es decir, entre todos los estimadores consistentes de un parámetro θ , los de máxima verosimilitud son los de varianza mínima.
- No siempre son insesgados.

3.2.3 Pruebas de Hipótesis

Las pruebas estadísticas de hipótesis se realizan en todos los ámbitos en los cuales puede constatarse la teoría frente a la observación. El problema relacionado con las pruebas de hipótesis estadísticas tiene una fuerte relación con el concepto de estimación. La mayoría de las pruebas estadísticas de hipótesis tienen que ver con los parámetros de las distribuciones, pero algunas veces también tienen que ver con el tipo, o naturaleza, de las distribuciones mismas.

“Una hipótesis estadística es una afirmación o conjetura acerca de la distribución de una o más variables aleatorias. Si una hipótesis estadística especifica completamente

*la distribución, se conoce como hipótesis simple; si no, se conoce como hipótesis compuesta.*³⁷

La esencia de probar una hipótesis estadística es decidir si la información se encuentra apoyada por la evidencia experimental que se obtiene a través de una muestra aleatoria. En forma general, la información involucra ya sea algún parámetro o a alguna forma funcional no conocida de la distribución de interés a partir de la cual se obtiene una muestra aleatoria. La decisión acerca de si los datos muestrales apoyan estadísticamente la afirmación se toma con base en la probabilidad, y, si esta es mínima, entonces será rechazada.

3.2.3.1 Elementos de una Prueba Estadística

Cualquier prueba estadística de hipótesis funciona exactamente de la misma manera y se compone de los mismos elementos esenciales. Los elementos de una prueba estadística de hipótesis son:

- Hipótesis nula, H_0
- Hipótesis alternativa H_1
- Estadístico de la prueba
- Región de rechazo

La hipótesis que se desea probar, es la hipótesis nula denotada como H_0 . La hipótesis alternativa (o de estudio), denotada por H_1 es la hipótesis que debe aceptarse en caso de rechazar H_0 . Las partes funcionales de una prueba estadística son el estadístico de la prueba y la región de rechazo. El estadístico de prueba (como un estimador) es una función de las mediciones muestrales en la cual se fundamenta la decisión estadística. La región de rechazo, denotada como RR , especifica los valores del estadístico de la prueba para los cuales se rechaza la hipótesis nula.

3.2.3.2 Prueba de una Hipótesis Estadística

La prueba de una hipótesis estadística es la aplicación de un conjunto explícito de reglas para decidir si aceptamos la hipótesis nula o la rechazamos a favor de la hipótesis alternativa. El procedimiento de prueba, divide los valores posibles de la estadística de prueba en dos subconjuntos: una región de aceptación para H_0 y una región de rechazo para H_0 . Este procedimiento puede conducir a dos clases de errores.

³⁷ Ver Freund, "Estadística Matemática con Aplicaciones", p. 384

“El rechazar la hipótesis nula H_0 , cuando ésta es verdadera, se llama Error tipo I, y el aceptar la hipótesis nula H_0 cuando ésta es falsa es llamado Error tipo II. El tamaño de un Error tipo I se define como la probabilidad de cometer un Error tipo I (denotada por α), de manera similar el tamaño de un Error tipo II es la probabilidad de se cometa un Error tipo II (denotada como β).”³⁸

Es costumbre referirse a la región de rechazo para H_0 como la región crítica de la prueba y la probabilidad de obtener un valor de la estadística de prueba dentro de la región crítica cuando H_0 es verdadera como el tamaño de la región crítica. Así, el tamaño de una región crítica es justamente la probabilidad α de cometer un Error tipo I. Esta probabilidad también se llama el nivel de significancia de la prueba.

Hemos expuesto el proceso de modelar y como los modelos nos facilitan el estudio y comprensión de los fenómenos reales y en particular de los fenómenos aleatorios con la ayuda de modelos matemáticos, también revisamos como la teoría de probabilidad estudia la propiedades de estos fenómenos aleatorios y proporciona modelos matemáticos para describirlos. En este capítulo revisamos los principales métodos y técnicas estadísticas que nos auxiliarán en nuestro proceso en la elección del modelo o los modelos que podrían ajustarse a los datos, calibrar los parámetros desconocidos y seleccionar el que mejor se ajuste a los datos.

³⁸ Ver Mood, “Introduction to the Theory of Statistics” p. 405

CAPÍTULO 4

Modelo para Siniestros Ocurridos: Incendio en Casa Habitación

*"Comprender las cosas que nos rodean es la mejor
preparación para comprender las cosas que hay mas allá."*

Hipatia (375-415)

Hemos revisados los principios que fundamentan la elaboración de un proceso de modelación y, en específico, el proceso para modelar situaciones bajo incertidumbre, así como algunos de los principios de la teoría de la probabilidad y de la estadística matemática. En este capítulo utilizaremos la teoría expuesta para modelar el monto de los siniestros ocurridos en casa habitación para la cobertura de incendio, cabe señalar que estamos intentando modelar el monto de un solo pago para estos siniestros ocurridos, no estamos intentando modelar los montos agregados. La información analizada se tomó de pólizas de hogar para los siniestros ocurridos en 2003 y únicamente aquellos que afectaron la cobertura de Incendio Edificio. Estos datos provienen de una de las aseguradoras más grandes del país, y líder en el ramo de daños que por políticas de confidencialidad de la información no podemos mencionar su nombre.

4.1 Antecedentes

En un sentido muy general, toda la ciencia actuarial se refiere a distribuciones de pérdidas, ya que esto es precisamente lo que un contrato de seguro cubre. La aseguradora pagará un monto aleatorio en un momento futuro también aleatorio. El cálculo del monto depende de los términos del contrato y de la naturaleza del evento cuando éste sucede. Sin embargo, a través del tiempo, la definición de pérdida se ha restringido a un solo pago, para proveer una distinción importante, haremos las siguientes definiciones.

"Un evento de pérdida o siniestro es un incidente en el cual un asegurado o grupo de asegurados sufren algún daño en sus bienes o en sus personas, el cual es potencialmente cubierto por sus contratos de seguros. La pérdida es el monto medido en términos monetarios del daño sufrido en sus bienes o en sus personas por un asegurado o grupo de asegurados como resultado de un evento de pérdida o siniestro. La pérdida también puede ser de cero. Un evento pagable es un incidente en el cual un asegurado o grupo de asegurados reciben un pago como resultado de un evento de pérdida cubierto por sus contratos de seguros. El monto pagado es el monto actual pagado en términos monetarios a los asegurados como resultado de un evento de pérdida o un 'evento pagable'. Si este es el resultado de un evento pagable, el monto pagado puede ser cero."³⁹

No todos los eventos de pérdida generan un pago y algunos eventos de pérdida generan una pérdida. Cuando un evento de pérdida se presenta, probablemente se incurra en gastos como resultado de atender un potencial siniestro, esto nos lleva a la siguiente definición.

³⁹ Ver Klugman, "Loss Models From Data to Decisions", p. 5

“Los Gastos de Ajuste Asignados al Siniestro (ALAE, Allocated Loss Adjustment Expense) es el monto de los gastos incurridos directamente como resultado de un evento de pérdida”⁴⁰

La diferencia entre la pérdida y el monto pagado casi siempre es atribuible a provisiones específicas en la póliza que reducen las obligaciones de la compañía aseguradora. Algunas de las más comunes modificaciones a las coberturas son:

“Un deducible ordinario es una provisión, la cual establece que cuando la pérdida sea menor que o igual al deducible, no habrá pago alguno y cuando la pérdida exceda al deducible, el monto pagado es la pérdida menos el deducible. El límite de póliza o simplemente el límite es una provisión la cual establece que cuando la pérdida exceda este límite, ningún beneficio adicional será pagado. El factor de coaseguro es la proporción de cualquier pérdida que es pagada por el asegurado después de que cualquier otra modificación (como deducibles o límites) ha sido aplicada. Un coaseguro es una provisión la cual establece que un factor de coaseguro se está aplicando.”⁴¹

4.1.1 Distribución de Perdidas

Ahora definiremos lo que es una distribución de pérdidas.

“Una distribución de perdidas es la distribución de probabilidad ya sea de la pérdida o del monto pagado por un evento de pérdida o del monto pagado de un evento pagable. La distribución puede o no excluir pagos de cero o puede o no incluir los ALAE”⁴²

Ya que la definición es ambigua en un número de puntos, el contexto específico debe ser hecho claro cuando una distribución de perdidas esta siendo determinada o evaluada. Para nuestro caso, la distribución de pérdidas que ajustaremos será del monto de siniestros ocurridos, los cuales son la suma de la reserva inicial, mas los ajustes, mas los gastos de ajuste, menos los salvamentos.

Mientras sea posible para una distribución de pérdidas ser basada en los datos empíricos, distribuciones de perdidas basadas en modelos paramétricos suavizados tienen muchas ventajas. El proceso de modelar supone una distribución paramétrica como la fuente generadora de las

⁴⁰ Ver Klugman, “Loss Models From Data to Decisions”, p. 5

⁴¹ Ver Klugman, “Loss Models From Data to Decisions”, p. 6

⁴² Ver Klugman, “Loss Models From Data to Decisions”, p. 6

pérdidas. Otra ventaja de un modelo paramétrico es la simplicidad. Con una palabra (Lognormal, por ejemplo) y dos números (media y varianza) se ha descrito completamente a la población.

4.2 Elección del Modelo

El primer paso en la determinación de nuestro modelo consiste en seleccionar las distribuciones que podría modelar los datos observados.

4.2.1 Métodos Descriptivos

Los datos a ser ajustados son presentados en la siguiente tabla de frecuencias. En el Anexo 3 presentamos las observaciones muestrales.

Clases	Límite Inferior	Límite Superior	Punto Medio	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Rel. Acum.
<i>en o de bajo</i>		-5,000		0	0.000000	0	0.000000
1	-5000	4,700	-150	30	0.2804000	30	0.2804000
2	4700	14,400	9,550	29	0.2710000	59	0.5514000
3	14400	24,100	19,250	16	0.1485000	75	0.7009000
4	24100	33,800	28,950	8	0.0748000	83	0.7757000
5	33800	43,500	38,650	3	0.0280000	86	0.8037000
6	43500	53,200	48,350	11	0.1028000	97	0.9065000
7	53200	62,900	58,050	5	0.0467000	102	0.9533000
8	62900	72,600	67,750	1	0.0093000	103	0.9626000
9	72600	82,300	77,450	1	0.0093000	104	0.9720000
10	82300	92,000	87,150	3	0.0280000	107	1.0000000
<i>arriba</i>	92000			0	0.0000000		

Tabla 1 Tabla de Frecuencias Siniestros Ocurredos

El primer paso en la determinación de un modelo que se ajuste al monto de los siniestros ocurridos, es producir algunas descripciones de las observaciones muestrales. Empezaremos por las principales medidas de tendencia central, de variabilidad y de forma presentadas en la tabla 2, los valores se encuentran en pesos mexicanos.

<i>Estadístico</i>	<i>Valor</i>
Media	20734.86822
Error típico	2103.07763
Mediana	12757.5
Moda	500
Desviación estándar	21754.40416
Varianza de la muestra	473254100.4
Curtosis	1.096217584
Coefficiente de asimetría	1.330605316
Rango	90410.93
Mínimo	100
Máximo	90510.93
Suma	2218630.9
Cuenta	107

Tabla 2. Estadísticas Descriptivas.

De gran interés son los estadísticos de forma para la asimetría estándar con un valor de 5.61908 y la curtosis estándar con un valor de 2.31464, los cuales pueden ser usados para determinar si las observaciones muestrales son generadas por una distribución normal. Valores para estos estadísticos fuera del rango de -2 a $+2$ indican una desviación importante de normalidad. Los valores para estos estadísticos se encuentran fuera del rango esperado para los datos de una distribución normal. Por lo tanto, como primer punto, podemos aseverar que los datos no los genera una distribución normal.

4.2.2 Métodos Gráficos

Los métodos gráficos son de gran ayuda ya que nos permiten visualizar una posible relación con alguna distribución conocida. La figura 2 presentan el histograma y la ojiva para los datos observados como se había mencionado estas proveen alguna indicación del tipo de distribución que podrían modelar los datos, pero aún así no es evidente. Una mejor vista podría ser obtenida al tomar los logaritmos naturales de los valores observados, en el anexo 3 se presentan los datos puntuales convertidos.

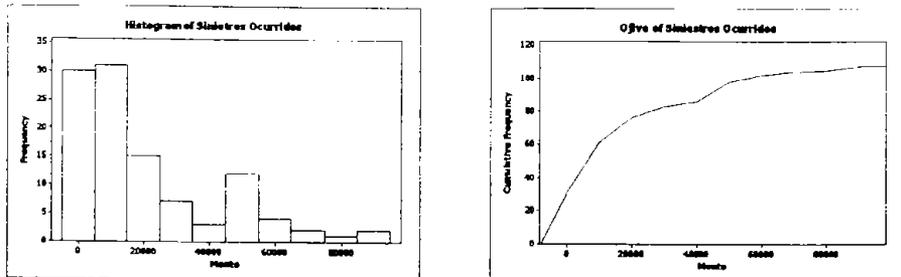


Figura 2. Histograma y ojiva siniestros ocurridos

El histograma y ojiva de los datos transformados se presentan en la figura 3. El histograma revela que la distribución transformada es un tanto simétrica, sugiriendo que una distribución Lognormal podría ser apropiada.

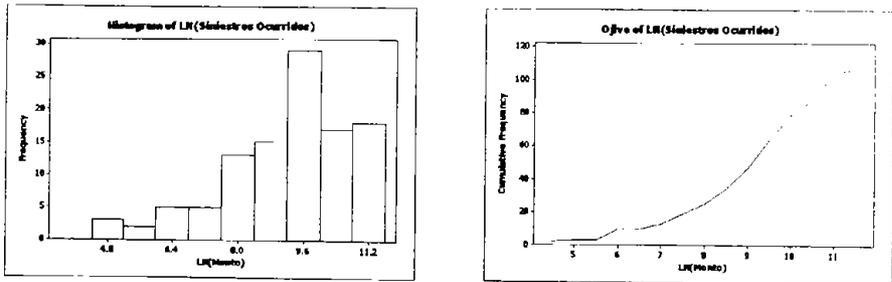


Figura 3. Histograma y Ojiva Ln siniestros ocurridos

Como último paso en la selección del modelo calcularemos los valores de la MRL⁴³. Algún conocimiento de estas funciones puede ser extremadamente valioso para el propósito de modelar.

La MRL se define como:

$$e_x(x) = e(x) = E(X - x | X > x) = \frac{\int_x^{\infty} (t - x) f(t) dt}{S(x)}$$

Intuitivamente, si $e(x)$ toma valores grandes, para valores grandes de x , entonces la distribución tiene una cola pesada ya que la el exceso de la pérdida esperada $X - x$ es grande. Recíprocamente, si $e(x)$ toma valores pequeños para valores grandes de x , tiene una cola ligera. El cálculo de la MRL empírica $e_n(x)$ es relativamente fácil, ya que ésta es simplemente la diferencia del promedio de todas las observaciones mayores o iguales que x . Si ésta pareciera ser una función lineal con pendiente positiva, entonces una distribución Pareto podría ser un modelo apropiado. Si ésta es más parecido a una constante para los valores más grandes de x , entonces se podría tratar el ajuste con una distribución Gamma. Algo entre estas dos, podría sugerir una distribución Lognormal o Weibull con $0 < \tau < 1$. Por otra parte, una función decreciente como $c_1 x^{-\tau+1}$, $\tau > 1$ indica que un modelo Weibull debe ser tomado en cuenta. Si los valores de $e_n(x)$ también pudieran aparentar cambios con varios valores de x , esto podría sugerir que diferentes

⁴³ Mean Residual Life (vida residual promedio)

⁴⁴ Ver Hogg, "Loss Distributions", p. 89

densidades sean usadas en diferentes partes del soporte. En el anexo 3 se presentan los datos puntuales para la MRL.

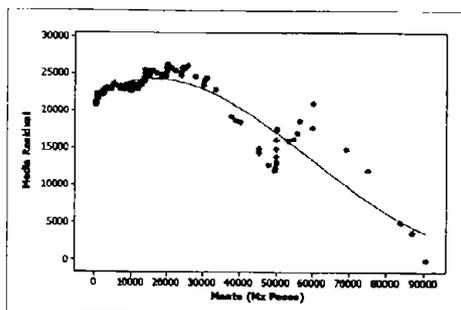


Figura 4. MLR siniestros ocurridos.

La figura 4 muestra los valores para $e_{107}(x)$, la forma de la curva ajustada sugiere que las distribuciones Weibull y Lognormal podrían proporcionar buenos modelos para ajustarse a los datos observados. Por lo tanto tomaremos estas dos distribuciones como posibles generadoras de los valores para los siniestros ocurridos.

4.3 Calibración del Modelo

Con ayuda de la evidencia descriptiva y grafica, hemos concluido que las dos posibles distribuciones generadoras de la información para los datos observados en los montos de siniestros ocurridos las podrían generar las distribuciones Lognormal y Weibull, ahora procederemos a la estimación de los parámetros.

4.3.1 Estimación de Parámetros

Como siguiente paso en nuestro proceso de ajuste es la estimación de los parámetros de las distribuciones que estamos proponiendo como candidatas a modelar los datos empíricos. El método que usaremos para este fin es el bien conocido método de máxima verosimilitud, aunque existen otros métodos de estimación, se eligió este ya que los estimadores encontrados por este método cuentan con muchas características deseables en un estimador y es común encontrar este método de estimación en los paquetes estadísticos. Los resultados de esta estimación son:

Distribución Lognormal				
Parámetro	Estimación	Error Estándar	IC Normal 95%	
			Inferior	Superior
Forma	9.15507	0.149949	8.86118	9.44897
Escala	1.55108	0.10603	1.35659	1.77346

Distribución Weibull				
Parámetro	Estimación	Error Estándar	IC Normal 95%	
			Inferior	Superior
Forma	0.847739	0.0659707	0.727816	0.98742
Escala	19126.5	2293.25	15120.9	24193.3

La figura 5 muestra las probabilidades puntuales ajustadas, al parecer la distribución Weibull se ajusta mejor a los datos que como lo hace la distribución Lognormal, pero aún así esta no es razón suficiente para aceptar una y rechazar la otra, para realizar la selección realizaremos las pruebas de hipótesis necesarias.

4.4 Selección y Validación del Modelo

El siguiente paso en el proceso de ajuste es la selección y validación del modelo lo realizaremos por medio de las pruebas de bondad de ajuste.

4.4.1 Pruebas de Bondad de Ajuste

Ahora probaremos cual de las dos distribuciones se ajusta mejor a los datos por medio de la prueba de Anderson-Darling, esta prueba es usada para probar si una muestra de datos proviene de una población con una distribución específica (al igual que los hacen todas las pruebas de bondad de ajuste), esta es una modificación de la prueba Kolmogorov-Smirnov y da más peso a las colas que la prueba Kolmogorov-Smirnov. La prueba Kolmogorov-Smirnov es de *libre distribución* en el sentido que los valores críticos no dependen de la distribución específica que se esta probando. La prueba de Anderson-Darling hace uso de la distribución específica para calcular los valores críticos. Esto tiene la ventaja de permitir una prueba más sensible y la desventaja de que los valores críticos deben ser calculados para cada distribución. La prueba de Anderson-Darling es una las pruebas de bondad de ajuste alterna a las pruebas Chi-cuadrada y Kolmogorov-Smirnov.

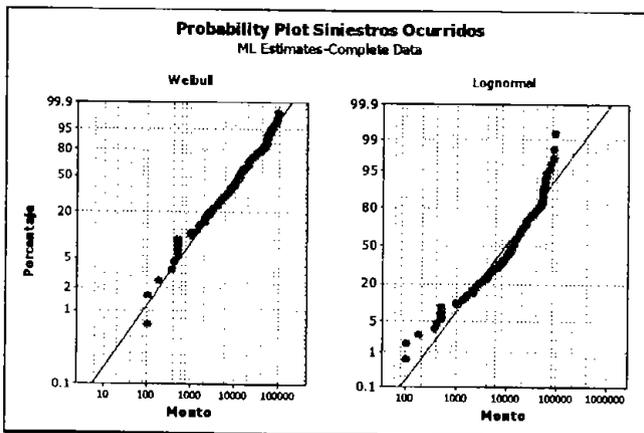


Figura 5. Probabilidades puntuales: siniestros ocurredos

La prueba de Anderson-Darling se define como:

H_0 : Los datos siguen una distribución específica

H_1 : Los datos no siguen una distribución específica

El estadístico de prueba Anderson-Darling es:

$$A^2 = -N - S$$

Donde:

$$S = \sum_{i=1}^N \frac{(2i-1)}{N} \left[\ln F(Y_i) + \ln(1 - F(Y_{N+1-i})) \right]$$

F es la función de distribución acumulativa de la distribución especificada, Y_i son los datos ordenados.

Nivel de significancia. α

Región crítica.

Los valores para la prueba Anderson-Darling dependen de la distribución específica a la que se aplica, estos valores son presentados en la tabla 3 para las distribuciones Lognormal y Weibull. Esta prueba es de una cola y la hipótesis que la distribución es de una forma específica es rechazada si el estadístico de prueba, A , es más grande que el valor crítico.

α	10.00%	5.00%	2.50%	1.00%
Lognormal	0.63100	0.75200	0.87300	1.03500
Weibull	0.63700	0.75700	0.87700	1.03800

Tabla 3. Valores críticos Anderson-Darling

Alternativamente, existen técnicas muy sólidas y pruebas no paramétricas que no conciben fuertes suposiciones sobre la distribución subyacente. Sin embargo, técnicas basadas en suposiciones sobre distribuciones específicas son en general más potentes que las no paramétricas. Por lo tanto, si las suposiciones respecto a la distribución pueden ser validados, entonces estas son generalmente preferidos.

Las hipótesis que tenemos que probar son:

$$H_0 : F(x) = F_n(x)$$

$$H_1 : F(x) \neq F_n(x)$$

donde $F_n(x)$ se distribuye como una Lognormal

y

$$H_0 : F(x) = F_n(x)$$

$$H_1 : F(x) \neq F_n(x)$$

donde $F_n(x)$ se distribuye como una Weibull

Los valores obtenidos para el estadístico de prueba Anderson-Darling son:

Prueba Anderson-Darling	
Lognormal	2.055000
Weibull	0.582000

Ya que el valor del estadístico de prueba Anderson-Darling para la distribución Lognormal es mayor a 1.03500, podemos rechazar la hipótesis nula que los datos observados para el monto de los siniestros ocurridos los genera una distribución Lognormal, con un nivel de significancia del 99%.

De forma similar, ya que el valor del estadístico de prueba Anderson-Darling para la distribución Weibull es menor que el 0.6310, no podemos rechazar la hipótesis nula que los datos observados para el monto de los siniestros ocurridos los genera una distribución Weibull con un nivel de significancia del 90%.

Las graficas presentadas, en la figura 6 y 7 muestran la función de distribución acumulativa empírica junto con función de distribución ajustada, para la distribución Weibull y la distribución Lognormal. Nuevamente, podemos observar que la distribución Weibull es la que mejor se ajusta, pero ahora este hecho también lo respalda la prueba de bondad de ajuste Anderson-Darling realizada a ambas distribuciones.

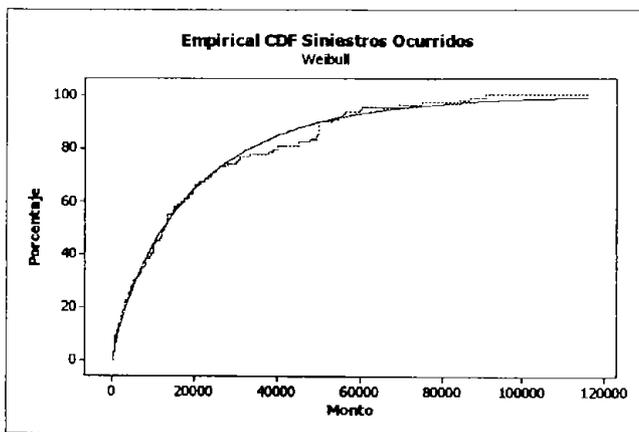


Figura 6. CDF Weibull

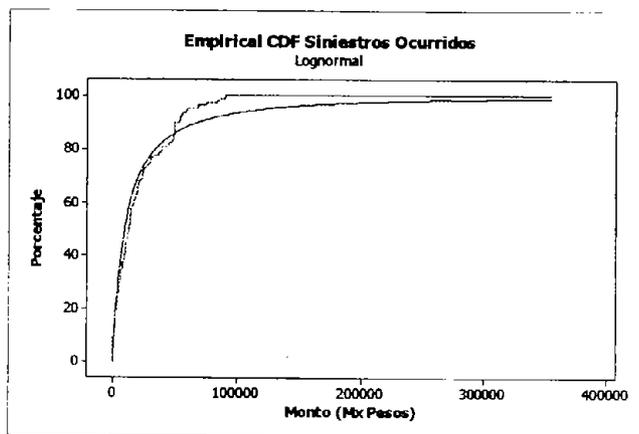


Figura 7. CDF Lognormal

Podemos concluir aseverando que la distribución generadora de la información para el monto de los siniestros ocurredos bien podría ser modelada con una distribución Weibull con un nivel de confianza del 90% y rechazar la distribución Lognormal por medio de la evidencia estadística antes expuesta.

4.5 Aplicación

Hemos presentado el proceso para ajustar distribuciones de probabilidad a los datos de siniestros ocurridos, el cual nos llevó a la conclusión de que la distribución que mejor ajusta el monto de un siniestro ocurrido esta definida por una Weibull. Aplicaremos el modelo para hacer un estimación del monto esperado de un siniestro ocurrido en 2004 así como el calculo de la distribución probabilidades de ocurrencia para valores mayores a un valor específico del monto del siniestro.

4.5.1 El Efecto de la Inflación

Si X es la variable aleatoria que modela los siniestros ocurridos, necesitamos encontrar la distribución de $Z = (1+r)X$, donde r es la inflación en el periodo de estudio. La expresión que relaciona la función de distribución de Z con la de X es:

$$\begin{aligned} F_z(z) &= \Pr[Z \leq z] \\ &= \Pr\left[\frac{Z}{1+r} \leq \frac{z}{1+r}\right] \\ &= F_x\left(\frac{z}{1+r}\right) \end{aligned}$$

Además, si X es una variable aleatoria continua, tenemos la siguiente relación entre sus funciones de densidad de probabilidad:

$$\begin{aligned} f_z(z) &= \frac{d}{dz} F_z(z) \\ &= \frac{f_x\left(\frac{z}{1+r}\right)}{1+r} \end{aligned}$$

Por lo tanto al multiplicar una familia de distribuciones por una constante, se obtiene otro miembro de esa familia. Aplicado a nuestro modelo Weibull tenemos:

Distribución	Parámetros de X	Parámetros de $Z = (1+r)X$
Weibull	η, β	$\eta(1+r), \beta$

La inflación anualizada a Agosto de 2004 es de 4.82 %⁴⁵, por lo tanto, el valor para los nuevos parámetros de nuestro modelo son:

$$\eta = (1 + 0.0482) * 19126.5 = 20048.4$$
$$\beta = 0.847739$$

Recordando las principales características de la distribución Weibull, tenemos:

$$f_x(x) = \frac{\beta}{\eta} \left(\frac{x}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\eta}\right)^\beta\right]$$

Donde:

$$\eta > 0, \beta > 0$$

y,

η es el parámetro de escala

β es el parámetro de forma

$$E[X] = \eta * \Gamma\left[1 + \frac{1}{\beta}\right]$$

Por lo tanto, nuestra estimación para el monto medio de un siniestro ocurrido en 2004 considerando el efecto de la inflación es:

$$E[X] = 20048.4 * 1.08978$$
$$= 21848.17$$

El monto promedio de un siniestro ocurrido en 2003 es de 20,734.87 pesos mientras que para 2004 se espera sea de 21,848.17 pesos, esto representa un incremento del 5.36% con relación a 2003.

4.5.2 Probabilidades de Ocurrencia

Con un modelo estimado pueden calcularse las probabilidades de ocurrencia de un evento, en este caso que produzca perdidas por un monto específico en pesos. Por ejemplo, desearíamos conocer

⁴⁵ Fuente, Banco de México.

cual es la probabilidad de que un evento sea mayor a 10,000 pesos. Recordando que la función de distribución para la Weibull es:

$$F_x(x) = 1 - \exp\left[-\left(\frac{x}{\eta}\right)^{\rho}\right]$$

La probabilidad de tener un evento que supere los 10,000.00 pesos, utilizando nuestro modelo y considerando el efecto de la inflación es de:

$$\begin{aligned} \Pr[X > 10,000] &= 1 - F_x(10,000) \\ &= 1 - \left\{ 1 - \exp\left[-\left(\frac{10,000}{20048.4}\right)^{0.847739}\right] \right\} = \exp\left[-\left(\frac{10,000}{20048.4}\right)^{0.847739}\right] \\ &= 0.57435 \end{aligned}$$

Siguiendo el mismo procedimiento, la tabla 4, presenta la probabilidad de que el monto esperado en 2004 para el monto de un siniestro ocurrido sea mayor al valor de x .

x	$\Pr[X > x]$
100	0.9888824
200	0.9800809
300	0.9720253
400	0.9644379
500	0.9571929
1,000	0.9242834
3,000	0.8188738
5,000	0.7348269
10,000	0.5743493
20,000	0.3686332
30,000	0.2448011
50,000	0.1141778
100,000	0.0201349
500,000	0.0000002

Tabla 4. Probabilidades esperadas en 2004.

Podemos observar que la probabilidad de que el monto de un siniestro sea mayor a un monto pequeños es alta y mientras mayor sea el monto esta decrece. Con una tabla de este tipo también sería posible estimara la siniestralidad agregada esperada para algún periodo de interés y de esta forma estimar el nivel de primas necesario para hacer frente a las obligaciones futuras. En general el tener un modelo a la mano al cual se pueda manipular de una manera relativamente fácil y actualizarlo con información nueva nos brinda una amplia gama de posibilidades para su aplicación.

CONCLUSIONES

El uso de modelos tiene una antiquísima historia ya que estos han ayudado a una mejor comprensión, estudio y control de fenómenos reales. La construcción de modelos se puede decir que es un arte ya que la elaboración de estos se entrelazan muchas variables algunas intrínsecas al problemas a resolver y otras relacionadas a las emociones y sensibilidades del investigador que lleva a cabo el proceso de modelación. El proceso de modelar nos permite encontrar una potencial solución a un problema real, una vez encontrada esta solución se puede tener un mejor control sobre el problema ya que se pueden volver a analizar cada paso del proceso y realizar ajustes o mejoras si fuera necesario.

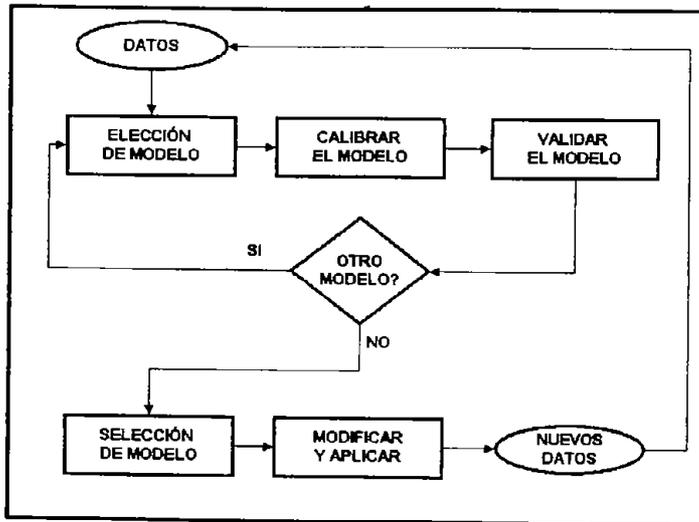
Muchos fenómenos del mundo real tienen un comportamiento aleatorio, la teoría de la probabilidad ha desarrollado varios modelos matemáticos (funciones de distribución) que podrían describir estos fenómenos. La estadística matemática es útil en el proceso de modelar ya que esta provee de métodos y técnicas para poder llevar a cabo, la elección de las función de distribución que podrían ajustarse a los datos de manera adecuada, calibrar los parámetros desconocidos para estas distribuciones, validar y seleccionar el modelo que mejor se ajuste a los datos.

El proceso para ajustar distribuciones de probabilidad a datos se puede resumir en los siguientes pasos:

- 1) Con los datos disponibles, agrupar la información, realizar tablas de frecuencias, calcular medidas numéricas descriptivas.
- 2) Seleccionar los modelos que podrían ser buenos candidatos para modelar la situación en estudio. Se usan histogramas y $e_n(x)$ como guías.
- 3) Calibrar los parámetros desconocidos para los modelos seleccionados.
- 4) Validar los modelos seleccionados para determinar si este se ajusta de forma adecuada a la información empírica. Varias pruebas de diagnostico pueden ser usadas. Algunas de estas pruebas estadísticas son las bien conocidas pruebas de bondad de ajuste como la de Chi-cuadrada, Kolmogorov-Smirnov o la prueba Anderson-Darling. La elección de la prueba esta relacionada directamente con el propósito final del ejercicio de modelar.
- 5) Se pueden considerar otros modelos, este paso es importante en el caso en el que el modelo seleccionado no se ajuste de forma adecuada a la información empírica.
- 6) Seleccionar el modelo que sea mas adecuado.
- 7) El modelo seleccionado es adaptado para aplicaciones futuras. En este trabajo lo usamos para estimar el efecto de la inflación y el calculo de las probabilidades de ocurrencia.

- 8) Con nueva información todos los incisos anteriores deben ser repetidos para mejorar el modelo.

La siguiente figura ilustra este proceso.



Proceso para ajustar distribuciones de probabilidad a datos

Finalmente podemos concluir que la distribución Weibull se ajusta a los datos del monto de un siniestro ocurrido para la cobertura de incendio edificio para pólizas de casa habitación con un nivel de confianza del 90%, llegando a esta conclusión aplicando el proceso antes resumido e ilustrado.

ANEXOS

Anexo 1. Inventario distribuciones discretas más comunes.

DISTRIBUCIONES DISCRETAS				
Distribución	Función de densidad discreta	Media	Varianza	Función generadora de momentos
Uniforme	$f(x) = \frac{1}{N}$ $x = 1, 2, \dots, N$	$\frac{(N+1)}{2}$	$\frac{(N^2 - 1)}{12}$	$\frac{e^t(e^N - 1)}{n(e^t - 1)}$
Bernoulli	$f(x) = p^x(1-p)^{1-x}$ $0 < p < 1$	p	$p(1-p)$	$(1-p) + pe^t$
Binomial	$f(x) = \binom{n}{x} p^x(1-p)^{n-x}$ $x = 0, 1, \dots, n$	np	$np(1-p)$	$[pe^t + (1-p)]^n$
Geométrica	$f(x) = p(1-p)^{x-1}$ $x = 1, 2, \dots$	$\frac{1}{p}$	$\frac{(1-p)}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Hipergeométrica	$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$ $x = 0, 1, \dots, n$ si $n \leq r$, $x = 0, 1, \dots, r$ si $n > r$	$\frac{nr}{N}$	$n \left(\frac{r}{N} \right) \left(\frac{N-r}{N} \right) \left(\frac{N-n}{N-1} \right)$	No útil
De Poisson	$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ $x = 0, 1, 2, \dots$ $\lambda > 0$	λ	λ	$e^{\lambda(e^t - 1)}$
Binomial Negativa	$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$; $x = r, r+1, \dots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left[\frac{pe^t}{1-(1-p)e^t} \right]^r$

Anexo 2. Inventario distribuciones continuas más comunes.

DISTRIBUCIONES CONTINUAS				
Distribución	Función de distribución acumulativa o función de densidad de probabilidad	Media	Varianza	Función generadora de momentos
Uniforme	$f(x) = \frac{1}{b-a}$ $-\infty < a < b < \infty$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt} - e^{at}}{(b-a)t}$
Normal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-(x-\mu)^2/2\sigma^2\right]$ $-\infty < \mu < \infty, \sigma > 0$	μ	σ^2	$\exp\left[\mu t + \frac{1}{2}\sigma^2 t^2\right]$
Exponencial	$f(x) = \lambda e^{-\lambda x}$ $\lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}$ $t < \lambda$
Gamma	$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}$ $\lambda > 0, r > 0$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - t}\right)^r$ $t < \lambda$
Beta	$f(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}$ $a > 0, b > 0$	$\frac{a}{a+b}$	$\frac{ab}{(a+b+1)(a+b)^2}$	No útil
Cauchy	$f(x) = \frac{1}{\pi\beta\left\{1 + \left[\frac{(x-\alpha)}{\beta}\right]^2\right\}}$ $-\infty < \alpha < \infty, \beta > 0$	No existe	No existe	Función característica $e^{-\alpha t - \beta t }$
Lognormal	$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left[-(\log_e x - \mu)^2/2\sigma^2\right]$ $-\infty < \mu < \infty, \sigma > 0$	$\exp\left[\mu + \frac{1}{2}\sigma^2\right]$	$\exp[2\mu + 2\sigma^2]$ $-\exp[2\mu + \sigma^2]$	No útil
Doble exponencial	$f(x) = \frac{1}{2\beta} \exp\left(-\frac{ x-\alpha }{\beta}\right)$ $-\infty < \alpha < \infty, \beta > 0$	α	$2\beta^2$	$\frac{e^{\alpha t}}{1 - (\beta t)^2}$
Weibull	$f(x) = abx^{b-1} \exp[-ax^b]$ $a > 0, b > 0$	$a^{-1/b} \Gamma(1+b^{-1})$	$a^{-2/b} [\Gamma(1+2b^{-1}) - \Gamma^2(1+b^{-1})]$	$E[X^t] = a^{-t/b} \Gamma\left(1 + \frac{t}{b}\right)$
Logistic	$F(x) = \left[1 + e^{-(x-a)/\beta}\right]^{-1}$ $-\infty < \alpha < \infty, \beta > 0$	α	$\frac{\beta^2 \pi^2}{3}$	$e^{\alpha t} \pi \beta t \csc(\pi \beta t)$

Pareto	$f(x) = \frac{\theta x_0^\theta}{x^{\theta+1}}$ $x_0 > 0, \theta > 0$	$\frac{\theta x_0}{\theta - 1}$ $\theta > 1$	$\frac{\theta x_0^2}{(\theta - 1)^2 (\theta - 2)}$ $\theta > 2$	No existe
Gumbler	$F(x) = \exp(-e^{-(x-\alpha)/\beta})$ $-\infty < \alpha < \infty, \beta > 0$	$\alpha + \beta\gamma,$ $\gamma \approx 577216$	$\frac{\pi^2 \beta^2}{6}$	$e^{\alpha t} (1 - \beta t)$ $\text{para } t < 1/\beta$
Distribución <i>t</i>	$f(x) = \frac{\Gamma[(k+1)/2]}{\Gamma(k/2)} \frac{1}{\sqrt{k\pi}} \frac{1}{(1+x^2/k)^{(k+1)/2}}$ $k > 0$	$\mu = 0$ $\text{para } k > 1$	$\frac{k}{k-2}$ $\text{para } k > 2$	No existe
Distribución <i>F</i>	$f(x) = \frac{\Gamma[(m+n)/2]}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{2}\right)^{m/2}$ $\times \frac{x^{(m-2)/2}}{[1+(m/n)x]^{(m+n)/2}}$ $m, n = 1, 2, \dots$	$\frac{n}{n-2}$ $\text{para } n > 2$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $\text{para } n > 4$	No existe
Chi- cuadrada	$f(x) = \frac{1}{\Gamma(k/2)} \left(\frac{1}{2}\right)^{k/2} x^{k/2-1} e^{-(1/2)x}$ $k = 1, 2, \dots$	k	$2k$	$\left(\frac{1}{1-2t}\right)^{k/2}$ $\text{para } t < 1/2$

Anexo 3. Datos muestrales y Media Residual Life.

Ítem	\$ Siniestro Ocurrido	LN Siniestro Ocurrido	M.R.L.	Ítem	\$ Siniestro Ocurrido	LN Siniestro Ocurrido	M.R.L.	Ítem	\$ Siniestro Ocurrido	LN Siniestro Ocurrido	M.R.L.
1	100	4.6052	20,830	37	6,848	8.8317	23,422	73	22,262	10.0106	25,466
2	100	4.6052	21,028	38	7,738	8.9539	22,858	74	23,692	10.0729	24,763
3	180	5.1930	21,149	39	7,984	8.9851	22,945	75	23,955	10.0839	25,266
4	363	5.8955	21,169	40	8,000	8.9872	23,271	76	24,184	10.0934	25,845
5	403	5.9977	21,338	41	8,115	9.0015	23,507	77	25,000	10.1266	25,863
6	500	6.2146	21,450	42	9,158	9.1224	22,809	78	25,600	10.1504	26,134
7	500	6.2146	21,665	43	9,425	9.1511	22,895	79	27,926	10.2373	24,659
8	500	6.2146	21,884	44	10,000	9.2103	22,674	80	29,955	10.3075	23,468
9	500	6.2146	22,107	45	10,000	9.2103	23,040	81	30,160	10.3143	24,158
10	500	6.2146	22,335	46	10,000	9.2103	23,417	82	30,823	10.3360	24,434
11	983	6.8906	22,080	47	10,224	9.2325	23,580	83	33,252	10.4119	22,923
12	1,000	6.9078	22,295	48	10,718	9.2797	23,477	84	37,701	10.5374	19,277
13	1,318	7.1835	22,211	49	11,640	9.3622	22,944	85	39,032	10.5721	18,762
14	1,500	7.3132	22,265	50	11,767	9.3730	23,217	86	40,107	10.5993	18,528
15	1,611	7.3848	22,395	51	11,959	9.3892	23,436	87	45,000	10.7144	14,317
16	2,000	7.6009	22,248	52	12,000	9.3927	23,821	88	45,104	10.7167	14,961
17	2,000	7.6009	22,495	53	12,686	9.4482	23,564	89	47,917	10.7772	12,823
18	2,006	7.6039	22,742	54	12,758	9.4539	23,935	90	49,423	10.8082	11,983
19	2,309	7.7446	22,694	55	13,191	9.4873	23,954	91	49,740	10.8146	12,396
20	2,466	7.8101	22,796	56	13,518	9.5118	24,090	92	49,898	10.8177	13,053
21	2,500	7.8240	23,027	57	13,544	9.5137	24,545	93	50,000	10.8198	13,876
22	2,798	7.9367	22,996	58	13,563	9.5151	25,026	94	50,000	10.8198	14,943
23	3,000	8.0064	23,065	59	13,574	9.5159	25,537	95	50,000	10.8198	16,189
24	3,284	8.0968	23,056	60	14,841	9.6051	24,786	96	50,097	10.8217	17,555
25	3,793	8.2409	22,822	61	14,975	9.6141	25,188	97	53,122	10.8803	15,983
26	3,981	8.2893	22,913	62	15,200	9.6290	25,518	98	54,494	10.9058	16,234
27	4,065	8.3101	23,115	63	15,875	9.6725	25,407	99	55,591	10.9258	17,030
28	4,239	8.3520	23,232	64	16,629	9.7189	25,227	100	56,317	10.9388	18,632
29	4,500	8.4118	23,265	65	17,392	9.7638	25,046	101	59,713	10.9973	17,776
30	4,577	8.4288	23,489	66	18,571	9.8294	24,449	102	59,905	11.0005	21,101
31	5,000	8.5172	23,369	67	18,619	9.8319	25,011	103	69,111	11.1435	14,868
32	5,000	8.5172	23,681	68	19,554	9.8810	24,693	104	74,958	11.2247	12,028
33	5,574	8.6259	23,419	69	19,734	9.8901	25,159	105	83,673	11.3347	4,970
34	6,238	8.7384	23,067	70	19,793	9.8931	25,778	106	86,774	11.3711	3,737
35	6,683	8.8073	22,936	71	20,000	9.9035	26,281	107	90,511	11.4132	0
36	6,842	8.8309	23,098	72	21,133	9.9586	25,867				

BIBLIOGRAFÍA

1. Berenson, M., Levine, D., and Rindsopf, D., (1988), "Applied Statistics: A First Course", United States of America: Prentice Hall.
2. Bowers, N., Gerber, H., Hickman, J., Jones, D., and Nesbitt, C. (1986), "Actuarial Mathematics", Schaumburg, IL: Society of Actuaries.
3. Canavos, C. (1994), "Probabilidad y Estadística. Aplicaciones y Métodos". México: McGraw-Hill.
4. Casualty Actuarial Society (2001), "Foundations of Casualty Actuarial Science", Arlington: Casualty Actuarial Society.
5. Chao, L. (1985), "Introducción a la Estadística", México: Compañía Editorial Continental.
6. Corina Ortega, G. (1974), "Los Modelos Matemáticos como Auxiliares en la Toma de Decisiones", México: Tesis de Licenciatura, Universidad Iberoamericana.
7. Freund, J., Miller, I., y Miller, M. (2000), "Estadística Matemática con Aplicaciones", 6ª ed., México: Pearson Educación
8. Giordano, F. (1997) "A First Course in Mathematical Modeling", Pacific Grove, California: Books/Cole Publishing Company.
9. Hogg R. and Klugman, S (1984), "Loss Distributions", United States of America: John Wiley & Sons, Inc.
10. Huntsberger, D. y Billingsley, P. (1983), "Elementos de Estadística Inferencial", México: Compañía Editorial Continental.
11. Jonson, R. y Kuby, P. (1999), "Estadística Elemental. Lo Esencial", 2da ed., México: International Thomson Editores .
12. Klugman, A., Panjer, H., and Willmont, G. (1998), "Loss Models: From Data to Decisions", United States of America: John Wiley & Sons, Inc.
13. Maki, D. (1983), "Mathematical Models and Applications. With Emphasis on the Social, Life and Management Sciences", United States of America: Prentice-Hall.
14. McGrath, R., (1997), "Understanding Statistics: A Reserch Perspective", Longman: Addison Wesley Publishers. Inc.
15. Mendenhall, W, Wackerly, D., y Scheaffer, R. (1994) "Estadística Matemática con Aplicaciones", 2a ed., México: Grupo Editorial Iberoamericano.
16. Mood, A., Graybill F., and Boes, D. (1974), "Introduction to the Theory of Statistics", 3er ed., United States of America: McGraw-Hills, Inc.
17. Parzen, E. (1987), "Teoría Moderna de Probabilidades y sus Aplicaciones", México: Editorial Limusa.
18. Saaty, T. (1981), "Thinking whit Models: Mathematical Models in the Physical, Biological and Social Sciences", United States of America: Pergamon Press.
19. Thierauf, R. (1972), "Toma de Decisiones por Medio de Investigación de Operaciones", México: Limusa-Wiley.

INTRODUCCIÓN	1
CAPÍTULO 1 MODELOS MATEMÁTICOS	3
1.1 Historia de los Modelos Matemáticos	4
1.2 Filosofía del Proceso de Modelar	7
1.3 Tipos de Modelos	9
1.3.1 Modelos Icónicos	9
1.3.2 Modelos Análogos	10
1.3.3 Modelos Simbólicos	10
1.4 Proceso para la Construcción de Modelos Matemáticos	11
CAPÍTULO 2 PROBABILIDAD Y MODELOS PARA VARIABLES ALEATORIAS	15
2.1 Fenómenos Aleatorios	16
2.2 Modelos Matemáticos de Fenómenos Aleatorios	17
2.2.1 Espacio de Descripciones Muestrales	17
2.2.2 Eventos	18
2.3 Definición de Probabilidad	19
2.4 Modelos para Variables Aleatorias	21
2.4.1 Concepto	21
2.4.2 Descripción de una Variable Aleatoria	22
2.4.2.1 Función de Probabilidad de una Variable Aleatoria	23
2.4.2.2 Ley de Probabilidad de una Variable Aleatoria	23
2.4.2.3 Función de Distribución	24
2.5 Esperanza de una Variable Aleatoria	27
2.5.1 Esperanza, Media y Varianza de una Variable Aleatoria	27
2.5.2 Función Generadora de Momentos	30
2.5.3 Función Característica	30
2.5.4 La Ley de los Grandes Números y el Teorema del Limite Central	32
CAPÍTULO 3 MÉTODOS Y TÉCNICAS ESTADÍSTICAS	34
3.1 Métodos Descriptivos	35
3.1.1 Distribución de Frecuencias	36
3.1.2 Presentación Gráfica	36
3.1.3 Medidas de Tendencia Central	37
3.1.4 Medidas de Variabilidad	38

3.2	Inferencia Estadística	39
3.2.1	Estimación	40
3.2.1.1	Estimadores y Estimaciones	40
3.2.1.2	Cualidades de un Estimador	41
3.2.1.3	Intervalos de confianza	44
3.2.2	Métodos de Estimación	45
3.2.2.1	Método de los Momentos	46
3.2.2.2	El Método de Máxima Verosimilitud	46
3.2.3	Pruebas de Hipótesis	47
3.2.3.1	Elementos de una Prueba Estadística	48
3.2.3.2	Prueba de una Hipótesis Estadística	48
 CAPÍTULO 4 MODELO PARA SINIESTROS OCURRIDOS: INCENDIO EN CASA HABITACIÓN		50
4.1	Antecedentes	51
4.1.1	Distribución de Perdidas	52
4.2	Elección del Modelo	53
4.2.1	Métodos Descriptivos	53
4.2.2	Métodos Gráficos	54
4.3	Calibración del Modelo	56
4.3.1	Estimación de Parámetros	56
4.4	Selección y Validación del Modelo	57
4.4.1	Pruebas de Bondad de Ajuste	57
4.5	Aplicación	61
4.5.1	El Efecto de la Inflación	61
4.5.2	Probabilidades de Ocurrencia	62
 CONCLUSIONES		65
 ANEXOS		67
 BIBLIOGRAFÍA		71

INTRODUCCIÓN

El objetivo principal de este trabajo es exponer el proceso para ajustar modelos de distribución de probabilidad a datos generados por el flujo de efectivo en un sistema asegurador, y en específico, por el flujo hacia fuera del sistema a causa de siniestros ocurridos para la cobertura de incendio en casa habitación; haciendo uso de métodos, teorías y técnicas estadísticas y de probabilidad para llevar acabo el proceso de ajuste.

Iniciaremos el primer capítulo con un recorrido en el que se presenta un esbozo histórico de algunos de los principales acontecimientos que dieron origen a la elaboración y construcción de modelos, y en particular, de los modelos matemáticos. Además se presenta la filosofía sobre la cual éstos se conciben, llegando así a la presentación de algunas de las clasificaciones generalmente aceptadas para los diferentes tipos de modelos. También se expondrán los pasos por los cuales se conduce el proceso de modelar, subrayando la importancia que este tiene como herramienta para predecir, pronosticar y tomar decisiones.

En el segundo capítulo se presentan algunos de los principales conceptos de la teoría de la probabilidad como preámbulo para abordar un concepto muy importante, el de *variable aleatoria*. De la misma forma que se modelan situaciones reales de las más variadas índoles, también podemos modelar situaciones que podemos catalogar como *probabilísticas*, las cuales describen eventos en el mundo real. La noción de variable aleatoria es usada para describir eventos y las funciones de distribución son usadas para dar las probabilidades de ciertos eventos en términos de variables aleatorias, de ahí la importancia de estas. Asimismo se presentarán algunas de las características y propiedades más sobresalientes de los modelos para variables aleatorias

En el tercer capítulo se presentan las principales técnicas y métodos para la recolección y presentación de datos, los cuales servirán como punto de partida para seleccionar los modelos para variables aleatorias, que junto con la evidencia empírica podrían ser buenos candidatos para modelar la situación en estudio. También abordaremos los métodos para la estimación de parámetros y pruebas de hipótesis que sirven para aceptar o rechazar un modelo, y también para verificar la calidad del ajuste

Finalmente, en el cuarto y último capítulo se expone el proceso para ajustar modelos de distribución de probabilidad (distribuciones de pérdidas) a datos generados por el monto de los siniestros ocurridos para la cobertura de incendio en casa habitación, para ello se emplearán las teorías y técnicas desarrolladas en los capítulos anteriores, junto con el uso de sistemas estadísticos como MINITAB, STATGRAPHICS y SAS como herramientas para facilitar el trabajo

proba-estadístico en la presentación gráfica de los datos, estimación de parámetros, selección y validación del modelo.

CAPÍTULO 1

Modelos Matemáticos

"Was beweisbar ist, soll in der Wissenschaft nicht ohne Beweis geglaubt werden."

"En Ciencia, lo que se puede probar no debe ser creído sin demostración."

R. Dedekind (1887)

La representación de la realidad a través de un modelo que la refleje y explique tiene antecedentes en todas las disciplinas del conocimiento. El uso de modelos con fundamento matemático ha permitido el desarrollo de teorías que describen fenómenos tanto naturales como sociales y permiten su comprensión y regulación, así como la obtención de mejores pronósticos y predicciones sobre sus resultados.

“Un modelo es un esquema teórico, de un sistema o de una realidad compleja, que se elabora para facilitar su comprensión y el estudio de su comportamiento.”¹

Dentro de los modelos nos encontramos con un tipo que ha tomado un gran auge en los últimos tiempos, y en particular, con la aparición de las computadoras, estos son los modelos matemáticos:

“Un modelo matemático es cualquier sistema completo y compatible de ecuaciones (estructuras) matemáticas, diseñadas para que se correspondan con alguna otra entidad: su prototipo. Tal prototipo puede ser una entidad física, biológica, psicológica o conceptual, tal vez, incluso otro modelo matemático.”²

Iniciaremos por revisar el desarrollo que los modelos han tenido a lo largo de la historia de la humanidad.

1.1 Historia de los Modelos Matemáticos

Gran parte del progreso de la humanidad se puede atribuir a la aplicación del método científico³ a los problemas donde la costumbre, inercia y la tradición habían imperado. Ejemplos individuales en los que parece que se usó el método científico para resolver problemas se han encontrado en escritos de miles de años de antigüedad. Podemos enumerar los siguientes acontecimientos más significativos relacionados con la evolución y el uso de modelos a lo largo de la historia de la humanidad:

- A. C.
 - Al suegro de Moisés, Jethro, se le acredita con un tratado sobre organización en el Cáp. 18 del Libro del Éxodo.

¹ Ver <<http://www.rae.es/>>

² Ver <<http://personal.telefonica.terra.es/web/jmora7/Archiv/96zoomc.PDF>>

³ “Estudio sistemático, controlado, empírico y crítico de proposiciones hipotéticas acerca de presuntas relaciones entre varios fenómenos”, ver <http://es.wikipedia.org/wiki/M%C3%A9todo_cient%C3%Adfico>

- Antiguos barcos en la ciudad Italiana de Venecia fueron reacondicionados y reparados en una línea de ensamblaje extremadamente ingeniosa; cada barco era movido a lo largo de esta línea en la que un grupo de expertos trabajadores efectuaban ciertas operaciones específicas en cada estación de la línea.
- 212 A. C.
 - La ciudad de Siracusa empleó a Arquímedes para idear métodos para romper el sitio naval a la ciudad, que estaba bajo el ataque de los romanos.
- 1798.
 - Eli Whitney usó el método científico para analizar la industria del algodón y desarrollar el concepto de partes intercambiables, que permitieron el cambio de la producción ajustada individual a la elaboración de grandes volúmenes de partes estandarizadas.
- 1832.
 - Charles Babbage, hace énfasis en el método científico, división del trabajo, estudio de tiempos y movimientos, especialización y contabilidad de costos.
- Finales de 1800.
 - Frederick Taylor aplicó el método científico al análisis de la productividad en las personas, cavadoras de minas y otros materiales. El estudio del paleo de Taylor es un ejemplo excelente de la aplicación del método científico.
- 1911.
 - Frederick Taylor publicó los *Principios de la Ciencia de la Administración*, los cuales formalizan los tiempos de estudio y los métodos de trabajo-estudio que aún hoy se usan.
 - Frank y Lillian Gilbreth desarrollaron los conceptos básicos de la psicología industrial, ilustrando su investigación sobre el estudio de los movimientos elementales de una operación manual en la industria.
- 1914.
 - F. W. Lanchester intentó tratar cuantitativamente las operaciones militares. Obtuvo ecuaciones que relacionaban el resultado de una batalla tanto a la fuerza relativa de los combatientes como a su capacidad relativa de fuego. Las ecuaciones de Lanchester sugerían que el poder total de las fuerzas de pelea era proporcional al cuadrado de su fuerza numérica. Lanchester modeló una situación que involucraba opciones estratégicas. Probó ese modelo contra una situación conocida en el mundo real.
 - Tomas Alba Edison estudiaba el proceso de la guerra antisubmarina. Reunió estadísticas a ser usadas en el análisis de maniobras mediante las cuales los barcos en la superficie pudieran evadir y destruir submarinos. Ideó un juego de

guerra a ser usado en la simulación de maniobras navales en vez de arriesgar los barcos en condiciones bélicas reales.

- 1915.
 - En el área de control de inventarios, los modelos del tamaño del lote económico tienen una larga genealogía, aunque se ha reportado que G. D. Babcock desarrolló un modelo enunciado en la forma de una ecuación cúbica, su técnica nunca fue publicada. El primer modelo publicado sobre el tamaño del lote económico de inventarios es generalmente atribuido a Ford W. Harris, quien describió y publicó su modelo en este año.
- 1917.
 - A. K. Erlang, un matemático danés que trabajaba con la compañía telefónica de Copenhague, publicó su trabajo más importante, "Solución a Algunos Problemas en la Teoría de Probabilidades Importantes en las Centrales Telefónicas Automáticas". Conteníó sus fórmulas de tiempos de espera que había desarrollado sobre la base de principios estadísticos. Estas fórmulas ahora bien conocidas, son de importancia fundamental a la teoría del tráfico telefónico.
- 1924.
 - W. Shewhart realizó las primeras aplicaciones registradas de la deducción estadística, cuando introdujo el concepto de gráficas de control de calidad.
 - La utilización de la deducción estadística y la teoría de probabilidades, fue ayudada por el trabajo de H. F. Dodge y H. G. Romig, compañeros de trabajo con Shewhart en los laboratorios de la Bell Telephone. Desarrollaron la técnica de inspección por muestreo en conexión con el control de calidad y publicaron tablas de muestreo estadístico que son ampliamente usadas hoy en día.
- 1928.
 - Otro ingeniero en los Bell Laboratories, T. C. Fry, hizo importantes contribuciones adicionales a las bases estadísticas de la teoría de colas. Una serie de conferencias dictadas por Fry en este año con relación a las aplicaciones a ingeniería de la teoría de probabilidades, se convirtió en la base de su importante libro sobre la materia.
 - Sir Roland Fisher trabajó sobre diversos métodos estadísticos modernos. Cuando estos fueron publicados tuvieron muy poco efecto; pero ahora, son la base para la mayoría de la teoría estadística aplicada.
- 1930's.
 - John Von Neumann desarrolló un modelo para una economía en expansión. También desarrolló las bases teóricas de la *Teoría de Juegos*.

□ Segunda Guerra Mundial.

- Nacimiento de la *Investigación de Operaciones*⁴. Desde principios de 1937 se pidió a los científicos ingleses cada vez con más frecuencia, que ayudaran a los militares a descubrir la mejor manera de utilizar el radar para localizar aviones enemigos. En septiembre de 1939 los científicos que trabajaban en diferentes aspectos del problema se reunieron en el Cuartel General del Mando de Aviones de Combate. Ese grupo, es considerado como el núcleo del primer grupo de Investigación de Operaciones. Este grupo ampliaba continuamente su área de actividades hasta abarcar más allá del problema original del radar y de su integración con los observadores de tierra. Grupos similares de Investigación de Operaciones se formaron en los Estados Unidos. Estos grupos sentaron las bases para el planteamiento y el desarrollo de soluciones a problemas de programación lineal, transportación, asignación de recursos, etc.

Concluida la Segunda Guerra Mundial los diversos grupos de Investigación de Operaciones que funcionaron para aspectos bélicos, pasaron a servir ahora al terreno industrial. Un poco más tarde, en la década de los cincuenta, la computadora incrementó el cambio, posibilitando el uso de los modelos matemáticos no solo para resolver problemas ocasionales, sino también, para resolver problemas operacionales comunes, rutinarios y aun cotidianos. Es más, con la computadora, es posible el uso de modelos matemáticos para predecir e imaginar problemas, para medir repercusiones si es que se toman ciertas medidas o caminos de acción, para requerir y generar información necesaria para la toma de decisiones.

1.2 Filosofía del Proceso de Modelar

Una de las primeras características que debe determinarse acerca de la situación en estudio es si es más apropiado modelar en términos deterministas o estocásticos. Un modelo es determinista si las predicciones y pronósticos realizados con suficiente información en un instante o etapa en el tiempo pueden determinar exactamente el comportamiento futuro del sistema en estudio. Por otro lado, un modelo es estocástico si este incorpora un comportamiento probabilístico. Para estos modelos no importa que tanto se conozca del sistema en un instante determinado en el tiempo, ya que, es imposible determinar con absoluta certeza la naturaleza del sistema para tiempos futuros. Por lo general, muchos de los modelos usados son de este tipo, es decir, modelos cuya descripción matemática envuelven oportunidad e incertidumbre, esto es totalmente esperado ya

⁴ “Aplicación, por grupos interdisciplinarios, del método científico a problemas relacionados con el control de las organizaciones o sistemas, a fin de que produzcan soluciones que mejor sirvan a los objetivos de la organización”, ver <<http://www.investigacion-operaciones.com/Historia.htm>>

que el mundo real provee una fuerte evidencia de ser un sistema estocástico. La decisión de que tipo de modelo debe ser construido depende de muchos factores y en última instancia simplemente es una elección del investigador. Usualmente un modelo determinista es tomado como una primera aproximación en una situación cuando un modelo estocástico parece ser más apropiado, sin embargo, no se debe de asumir que las predicciones basadas en un tipo de modelo son necesariamente mejores (o peores) que aquellas basadas en otro tipo. Los méritos relativos de los dos tipos de modelos varían de una situación a otra.

Sí un modelo es de uso práctico, se debe de tener un medio para obtener resultados que puedan ser probados o comparados con el mundo real. Muchos de los modelos guían a relaciones matemáticas que involucran uno o varios parámetros. Rastreado de estos parámetros de vuelta al mundo real, se puede encontrar que estos están relacionados con una tasa de aprendizaje, una probabilidad de muerte, una razón de activos, etc. Por lo tanto, al comparar el estudio matemático con la realidad, podría ser necesario dar valores numéricos a estos parámetros y para muchos modelos esta es la tarea principal.

El uso de técnicas estadísticas juega un importante rol en varios aspectos de la construcción de modelos. Un cuestionamiento estadístico podría ser, ¿Cuál es la relación entre el estadístico \bar{x} para la muestra y su correspondiente valor medio \bar{X} para la población total? Problemas de este tipo pueden ser resueltos por medio de las técnicas estadísticas de estimación de parámetros. Otra importancia de los métodos estadísticos en la modelación matemática es la de probar el modelo por correspondencia y exactitud: ¿El modelo realmente hace lo que se pretendía que hiciera? Un acercamiento para responder tal pregunta usualmente es por medio de la acumulación de datos, tanto de experimentos controlados o a través de observaciones del mundo real y entonces efectuar la evaluación comparando los resultados predichos por el modelo con aquellos resultados del análisis estadístico de los datos. Existen varias pruebas estadísticas estándar con muchas variantes, las cuales son útiles en esta conexión.

Los dos problemas arriba mencionados no son independientes. Se podría esperar que un modelo prevea resultados confiables si los parámetros han sido cuidadosamente seleccionados. Así, cuando se evalúen las predicciones basadas en un modelo en particular, es esencial tener en cuenta los procesos usados para determinar los parámetros y la exactitud que podría esperarse. También es común que un conjunto de datos se use para determinar los parámetros y para determinar las pruebas del modelo. Variaciones importantes en los datos podrían ser reflejadas en la estimación de los parámetros y en las pruebas, y aunque el modelo pudiera predecir resultados consistentes con los datos, los datos defectuosos podrían oscurecer la eficiencia del modelo.

Una evaluación estadística de la exactitud de un modelo usualmente se lleva a cabo junto con el uso de ciertas medidas estándar de la discrepancia entre lo predicho y los datos observados. De esta forma, se obtienen medidas numéricas de la bondad del ajuste para cada modelo. Naturalmente, si un modelo provee un ajuste más consistente que cualquier otro modelo, entonces este modelo será aceptado y los otros rechazados, sin embargo, frecuentemente sucede que cierto modelo será el que mejor explique y se ajuste a ciertos conjuntos de datos, mientras otro modelo será el que mejor explique y prediga otros conjuntos de datos. Ningún modelo puede ser rechazado, ya que bajo ciertas circunstancias cada uno es el mejor. Asimismo, ningún modelo debe ser completamente aceptado ya que en ciertos casos cada modelo no es el mejor disponible. Estos pueden ser condicionalmente aceptados, estudiados, y usados en aquellas circunstancias donde resulten ser la elección más apropiada.

1.3 Tipos de Modelos

Las diferentes clasificaciones de los modelos nos proporcionan una idea adicional de sus características esenciales, ya que, pueden describirse de muchos modos. Los modelos pueden clasificarse por sus dimensiones, funciones, propósitos o grado de abstracción. La base más común es la de tipos de modelos, que incluyen los siguientes tipos básicos:

- i. Icónicos
- ii. Análogos
- iii. Simbólicos

1.3.1 Modelos Icónicos

Un modelo icónico es una representación física de algunos objetos, ya sea en forma idealizada o en escala distinta. Para expresarlos de otro modo, una representación es un modelo icónico hasta el grado en que sus propiedades sean las mismas que tiene aquello que representa. Los modelos icónicos son muy adecuados para la descripción de acontecimientos en un momento específico del tiempo. Otra característica de un modelo icónico la constituyen sus dimensiones, dos dimensiones (fotografía, plano y mapa), o tres dimensiones (globo, automóvil y avión). Cuando un modelo sobrepasa la tercera dimensión, es imposible construirlo físicamente, y entonces pertenece a otro tipo de categoría llamados simbólicos o matemáticos.

1.3.2 Modelos Análogos

Los modelos análogos pueden representar situaciones dinámicas y se usan más que los icónicos, porque pueden mostrar las características del acontecimiento que se estudia. Las curvas de demanda, las curvas de distribución de frecuencias en las estadísticas y los diagramas de flujo, son ejemplos de modelos análogos. A menudo un modelo análogo es muy adecuado para representar relaciones cuantitativas entre las propiedades de los objetos de varias clases. Al transformar las propiedades en propiedades análogas, con frecuencia podemos incrementar nuestra capacidad de hacer cambios. Una de las ventajas de los modelos análogos sobre los icónicos es que ordinariamente puede hacerse que los primeros representen muchos procesos distintos del mismo tipo, lo que se hace evidente en el flujo de trabajos en proceso y de productos terminados de una fábrica. No podría usarse eficazmente un modelo icónico para estudiar los efectos de ciertos cambios en el control de calidad. Un diagrama de flujo es un modelo análogo muy sencillo y eficaz en esas situaciones.

1.3.3 Modelos Simbólicos

Los modelos simbólicos son representaciones de la realidad y toman la forma de cifras, símbolos y ecuaciones matemáticas. Comienzan como modelos abstractos que formamos en nuestra mente y que luego se registran como modelos simbólicos. Un tipo de modelo simbólico o matemático comúnmente usado es una ecuación. Una ecuación es concisa, precisa y fácil de comprender. Sus símbolos no solo son muchos más fáciles de manipular que las palabras, sino que se escriben más rápidamente. Además de estos atributos, los modelos simbólicos se prestan a las manipulaciones y uso en las computadoras.

Otros tipos incluyen modelos gráficos y pictóricos. Hay que tener en cuenta que pueden presentarse problemas para los que las analogías son más eficientes que los modelos simbólicos. Por ejemplo, un sistema puede ser tan complicado que la cantidad de trabajo requerida para construir un modelo simbólico sea demasiado costoso si se relaciona con las ganancias posibles. A menudo es difícil asignar tan solo un modelo a una clase, y esto es especialmente cierto con respecto a los modelos de simulación, que son modelos análogos y que se describen con símbolos matemáticos.

Los modelos matemáticos usualmente se clasifican en:

- i. Cuantitativos y Cualitativos.
- ii. Estándar y Hechos a la medida.
- iii. Probabilísticos y Determinísticos.
- iv. Descriptivos y de Optimización.
- v. Estáticos y Dinámicos.
- vi. Simulación y No Simulación.

1.4 Proceso para la Construcción de Modelos Matemáticos

El proceso de modelar se ilustra en la figura 1 siguiente,

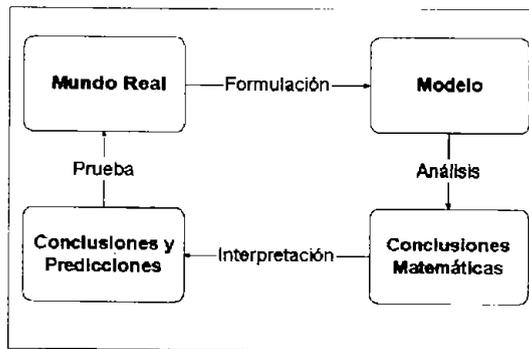


Figura 1. Proceso de Modelar

Algunos de los puntos más importantes de este proceso son:

□ Etapa 1

- **Analizar el problema.** En esta etapa se debe estudiar la situación lo suficiente para identificar el problema en una forma precisa y comprender claramente sus cuestionamientos fundamentales. Se deben determinan los objetivos del problema y por lo general se decide si el problema es del tipo determinista o del tipo estocástico. Solo con una clara y precisa identificación del problema se puede trasladar el problema en conceptos y símbolos matemáticos y de esta forma desarrollar y resolver el modelo.

□ **Etapa 2.**

- **Formular un modelo.** En esta etapa, una vez que el problema ha sido identificado de manera clara y precisa se diseña un modelo formando una abstracción del sistema o fenómeno que se este modelando. Algunas de las tareas principales en esta etapa son:
 - *Obtener datos.* Se deben coleccionar datos relevantes con información benéfica acerca del comportamiento del sistema.
 - *Realizar simplificaciones.* La formulación de un modelo, se debe intentar ser tan simple como razonablemente sea posible. Así, con frecuencia se decide simplificar algunos de los factores que no parecen ser importantes. Muchos problemas son demasiado complejos para considerar cada detalle, y si fueran considerados solo podría hacer del modelo imposible de resolver o consumir un monto importante de tiempo el resolvertlo. Por lo tanto, frecuentemente existen factores que apreciablemente no afectan los resultados. Además simplificando factores, se podría decidir regresar al paso previo para restringir el problema en estudio.
 - *Determinar variables.* Se deben determinar y nombrar las variables. Una variable independiente es la variable de la cual otras dependen. El modelo tratará de explicar las variables dependientes. Para simplificar el modelo, se podrían eliminar algunas variables, tratar ciertas variables como constantes, o agregar varias variables dentro de una. Mientras se decide sobre las variables, se deben establecer sus unidades, tales como: hora, día, semana o año para una unidad de tiempo.
 - *Establecer relaciones entre las variables y submodelos.* Hay que dibujar un diagrama del modelo descomponiéndolo en submodelos e indicar la relación entre las variables. Para simplificar el modelo, se podrían asumir relaciones más simples de lo que realmente son, por ejemplo, se podría asumir que dos variables están relacionadas de una manera lineal en lugar de una relación más compleja.
 - *Determinar ecuaciones.* Mientras se establecen las relaciones entre las variables, se podrían determinar ecuaciones con estas variables. Muchos modelos en las ciencias sociales y de la vida envuelven ecuaciones diferenciales, o las ecuaciones involucran una derivada.

□ Etapa 3

- **Resolver el modelo.** En esta etapa se implementa el modelo. Es importante no brincar a este paso antes de comprender el problema y diseñar el modelo. En otro caso, mucho tiempo puede ser desperdiciado y experimentar grandes frustraciones. Algunas de las técnicas que se pueden emplear en la solución son: álgebra, cálculo, gráficos, programación, paquetes computacionales como MINITAB, STATGRAPHICS, SAS, etc. La solución podría producir una respuesta exacta o podría simular la situación. Si el modelo es muy complicado de resolver, se debe volver a la etapa 2 y realizar simplificaciones adicionales o a la etapa inicial y reformular nuevamente el problema.
- **Verificación e interpretación de la solución del modelo.** Una vez que se tiene la solución, se deben examinar cuidadosamente los resultados y estar seguros que estos tienen sentido y que la solución resuelve el problema original y es útil. Probar la solución y observar si las predicciones concuerdan con los datos reales es sumamente importante. Se debe analizar la solución del modelo para determinar sus implicaciones. Si la solución del modelo muestra debilidad, se deberá volver a la etapa 1 o 2 para determinar si es factible refinar el modelo, si es así, se debe de realizar todo el proceso nuevamente. Así, el ciclo del proceso de modelar es una compensación entre simplificación y refinamiento. Por refinamiento, se podría necesitar extender el alcance del problema en la etapa 1. En la etapa 2, mientras el modelo es refinado, frecuentemente se necesitan reconsiderar las simplificaciones, incluir más variables y submodelos e incluir técnicas más sofisticadas.

□ Etapa 4

- **Publicar los resultados.** Divulgar sobre el modelo es importante por su utilidad. Un reporte contiene los siguientes componentes, los cuales van de la mano con el proceso de modelar:
 - Análisis del Problema
 - Diseño del modelo
 - Solución del modelo
 - Resultados y conclusiones
- **Mantener el modelo.** Como la solución del modelo es usada, podría ser necesario o deseable hacer ciertas correcciones, mejoras o realces. En este caso se debe realizar el proceso nuevamente para revisar la solución desarrollada.

El proceso de modelar es un esfuerzo científico creativo. Como tal, el problema que se pretende modelar usualmente no tiene una respuesta correcta. Los problemas son complejos y muchos modelos proporcionan ser buenos aunque con diferentes soluciones.

Seguiremos el proceso de modelar antes expuesto y realizaremos algunas modificaciones a este, y lo ajustaremos para que nos sirva de base para ajustar modelos matemáticos para fenómenos aleatorios (funciones de distribución). La teoría de probabilidades nos proporciona una gran variedad de modelos que bien podrían ajustarse a nuestros datos, ¿Cómo hacer esta elección? La respuesta a esta pregunta nos la brindara los métodos y técnicas estadísticas los cuales formaran parte en el proceso para ajustar modelos de fenómenos aleatorios a datos generados por el monto de siniestros ocurridos para la cobertura de incendio edificio para casa habitación.

CAPÍTULO 2

Probabilidad y Modelos para Variables Aleatorias

"Solo mediante la aplicación de las leyes del azar a las observaciones pueden determinarse, en un caso concreto, si existe una relación de causa-efecto."

Max Born (1882-1970)

Una gran variedad de experimentos y fenómenos naturales, biológicos, sociales o combinaciones de estos, tienen la característica de generar resultados u observaciones que no son susceptibles de predecirse con certeza, esto es, que aún cuando se estudien repetidamente bajo un mismo conjunto de circunstancias no siempre se observa el mismo resultado. Entre ellos, podemos citar los juegos de azar: lanzamiento de dados o monedas, juegos con barajas, ruletas, loterías, etc. Otra familia de ejemplos también muy conocida, la proporcionan los fenómenos meteorológicos: magnitud, intensidad y extensión de las lluvias; humedad, dirección y velocidad de los vientos; temperaturas máximas y mínimas; y todas sus consecuencias como volúmenes de agua en las presas, magnitud de los daños provocados por inundaciones, sequías, incendios, heladas, etc.

2.1 Fenómenos Aleatorios

Una de las características más notables de nuestra época consiste en el empleo, cada vez mayor, de las ideas de la teoría de probabilidades en una amplia variedad de campos científicos, que abarca temas tan ajenos y diferentes entre sí. ¿Qué es entonces lo que estudia la teoría de probabilidades?, ¿Qué da lugar a esta diversidad de aplicaciones? Para contestar estas preguntas comenzaremos por definir una propiedad en común que tienen ciertos fenómenos.

“Un fenómeno aleatorio (fortuito o al azar) es un fenómeno empírico que se caracteriza por la propiedad de que, al observarlo bajo un determinado conjunto de condiciones, no siempre se obtiene el mismo resultado (no existe regularidad determinista) sino que los diferentes resultados ocurren con regularidad estadística. Esto quiere decir que existen números entre 0 y 1 que representan la frecuencia relativa con la que se observan los diferentes resultados en una serie de repeticiones independientes del fenómeno.”⁵

Dos conceptos estrechamente ligados al de fenómeno aleatorio son los de *evento aleatorio* y *probabilidad de un evento aleatorio*. Un evento aleatorio tiene la propiedad de que la frecuencia relativa con la que aparece en una sucesión muy larga de observaciones realizadas al azar, se acercan a un valor límite estable a medida que el número de observaciones tiende a infinito; el valor límite de la frecuencia relativa se llama probabilidad del evento aleatorio.

Para arrojar más luz sobre el significado de un evento aleatorio, consideremos uno que es característico; por ejemplo, un accidente de automóvil. Es imposible predecir que un automovilista en particular que inicia un viaje por carretera se vea o no envuelto en algún accidente, sin

⁵ Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 18.

embrago, si observamos a todos los automovilistas (o a un gran número de ellos) que viajan por la misma carretera en un mismo día, es posible determinar la proporción de los que sufrirán accidentes, si esta proporción permanece constante de un día a otro, podemos considerar que lo que acontece a un automovilista en particular en esa carretera es un fenómeno aleatorio y que el accidente es un evento aleatorio.

2.2 Modelos Matemáticos de Fenómenos Aleatorios

La teoría de probabilidades estudia los métodos de análisis que son comunes en el tratamiento de fenómenos aleatorios, cualquiera que sea el área en que éstos se presenten. La probabilidad es, pues, la ciencia de los fenómenos aleatorios, en el sentido que estudia las propiedades de estos fenómenos que dependen esencialmente del concepto de aleatoriedad y no de otros aspectos particulares.

Consideramos la teoría de probabilidades como una disciplina matemática y, como tal, se construye mediante el método axiomático. De aquí, mediante la deducción lógica y sin recurrir a la experimentación se obtienen diversas proposiciones llamadas teoremas. Aun cuando éstos no se refieran al mundo real, sino que son consecuencia lógica de los axiomas, representan de hecho, conclusiones sobre fenómenos reales en la medida en la que acepte que dichos fenómenos tienen las propiedades postuladas en los axiomas.

Llegamos así al concepto de lo que es un modelo matemático de un fenómeno real. Si tenemos una regla para convertir proposiciones de una teoría matemática en proposiciones relativas a un fenómeno real, entonces la teoría matemática, que se ha construido axiomáticamente, es un modelo del fenómeno. En general, solo se necesita identificar, los objetos abstractos a que se refieren los axiomas de la teoría con los aspectos del fenómeno real, para estar en condiciones de utilizar esta teoría como su modelo. Una vez que se ha planteado el modelo matemático, uno espera que los teoremas deducidos describan el fenómeno con la misma fidelidad que lo hacen los axiomas, puesto que aquellos no son más que consecuencias lógicas de estos últimos.

2.2.1 Espacio de Descripciones Muestrales

Hemos afirmado que la teoría de probabilidades es el estudio de modelos matemáticos de los fenómenos aleatorios, que trata de las proposiciones que pueden hacerse acerca de un fenómeno aleatorio del que se hacen ciertos postulados, pero ¿Cómo se formulan postulados acerca de un

fenómeno aleatorio?. Para contestar esta pregunta nos valdremos del concepto de espacio de descripciones muestrales.

“Un espacio de descripciones muestrales de un fenómeno aleatorio que denotaremos con la letra Ω , es el espacio de las descripciones (o nombres) de todos los resultados posibles de un experimento.”⁶

Un espacio de descripciones muestrales de un fenómeno aleatorio puede ser definido de más de una manera. Si quienes observan el fenómeno tienen concepciones diferentes de los posibles resultados de la observación, seguramente definirán diferentes espacios de descripciones muestrales.

2.2.2 Eventos

La importancia del concepto de espacio muestral que corresponden a un fenómeno aleatorio se debe a que constituyen un medio para definir el concepto de evento.

“Un evento es un subconjunto del espacio muestral. La clase de todos los eventos asociados con un experimento se define como el espacio de eventos, denotado por Ψ .”⁷

Por lo tanto, un evento lo definimos como un conjunto de descripciones. Al decir que cierto evento A ha ocurrido, significa que el resultado de la situación aleatoria considerada, tiene por descripción un elemento del conjunto A . Obsérvese que se están definiendo dos conceptos: el de *evento* y el de *ocurrencia y realización de un evento*. El primero es un instrumento básico para construir modelos matemáticos de fenómenos aleatorios, mientras que el segundo constituye una base para traducir las proposiciones relativas al modelo matemático a proposiciones sobre el fenómeno real.

Al estudiar un fenómeno aleatorio nos interesan los eventos que puedan ocurrir (o, más precisamente, las probabilidades de que ocurran); por lo tanto, el interés del espacio de descripciones muestrales no radica en sus elementos; que son las descripciones muestrales, sino en sus subconjuntos, ¡que son los eventos!

⁶ Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 24

⁷ Ver Mood, “Introduction to the Theory of Statistics”, p. 15

La definición de espacio muestral es precisa y satisfactoria, mientras que, la definición de evento y espacio de eventos no es totalmente satisfactoria, ya que esta, no dice exactamente que subconjuntos podrían ser eventos y cuales no. Más que desarrollar las matemáticas necesarias para definir precisamente que subconjunto de Ω constituyen el espacio de eventos Ψ , estableceremos algunas propiedades para Ψ que parecerían ser razonablemente requeridas:

- $\Omega \in \Psi$.
- Si $A \in \Psi$, entonces $A^c \in \Psi$.
- Si A_1 y $A_2 \in \Psi$, entonces $A_1 \cup A_2 \in \Psi$

Cualquier colección de eventos con las propiedades arriba mencionadas es llamada un álgebra Booleana, o simplemente un álgebra.

2.3 Definición de Probabilidad

Una de las formas para definir la probabilidad es por medio del concepto *clásico*, que se aplica cuando todos los resultados posibles son igualmente probables.

"Si un experimento aleatorio puede resultar en n resultados igualmente probables y mutuamente excluyentes⁸ y si n_A de estos resultados tienen un atributo A , entonces la probabilidad es la fracción n_A/n ."⁹

Esta definición también se conoce como la probabilidad *a priori*. Una deficiencia importante del concepto clásico de probabilidad es su aplicación limitada, pues hay muchas situaciones en que las posibilidades que se presentan no pueden considerarse igualmente probables. Otro concepto de probabilidad, ampliamente usado es el de *frecuencia relativa* o también conocida como probabilidad *a posteriori*.

⁸ Los subconjuntos A y B se definen ser mutuamente excluyentes o disjuntos si $A \cap B = 0$. Los subconjuntos A_1, A_2, \dots se definen ser mutuamente excluyentes si $A_i A_j = 0$ para cada $i \neq j$, Ver Mood "Introduction to the Theory of Statistics", p. 14.

⁹ Ver Mood, "Introduction to the Theory of Statistics", p. 3.

Si en un experimento aleatorio, se observa que cuando el número n de observaciones aumenta, las frecuencias relativas con las que el número de cierto suceso con el atributo A , la fracción n_A/n , converge hacia cierta cantidad denominada la probabilidad de A , es decir, $P[A] = \lim_{n \rightarrow \infty} (n_A/n)$.

Con esta definición no se puede garantizar lo que ocurrirá en una ocasión en particular, pero si se llevaran registros de un gran número de observaciones durante un largo periodo, se debe de encontrar que la frecuencia observada del suceso con el atributo bajo estudio es bastante confiable.

La *definición axiomática* de la probabilidad se debe al Físico-Matemático ruso Andrei Nikolaevich Kolmogorov.

“Dada una situación aleatoria, descrita por un espacio de descripciones muestrales Ω , la probabilidad es una función¹⁰ que asigna a cada evento un número real no negativo, denotado por $P[A]$, que se llama probabilidad del evento A . Dicha función de probabilidades debe satisfacer los siguientes axiomas:

- i. $P[A] \geq 0$ para todo evento A .
- ii. $P[\Omega] = 1$.
- iii. $P[A_1 \cup A_2] = P[A_1] + P[A_2]$, donde $A_1 \cap A_2 = \emptyset$.¹¹

Un concepto ligado al de función de probabilidad es el de espacio de probabilidades.

“Un espacio de probabilidad es la tripleta $(\Omega, \Psi, P[\cdot])$, donde Ω es un espacio muestral, Ψ es la colección de eventos, cada uno un subconjunto de Ω , y $P[\cdot]$ es una función de probabilidad con dominio Ψ .”¹²

¹⁰ “Una función es una regla que asigna un número real a cada elemento de un conjunto de objetos, llamado dominio de la función. Aquí el dominio de la función de probabilidad $P[\cdot]$ es el conjunto de todos los eventos de Ω ”, Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 34

¹¹ Ver Parzen, “Teoría Moderna de Probabilidades y sus Aplicaciones”, p. 34

¹² Ver Mood, “Introduction to the Theory of Statistics”, p. 25

El espacio de probabilidad es un término único, que permite una forma conveniente de suponer la existencia de los tres componentes de su notación. Los tres componentes están relacionados; Ψ es la colección de subconjuntos de Ω , y $P[\cdot]$ es una función que tiene Ψ como dominio.

2.4 Modelos para Variables Aleatorias

Hemos visto que la probabilidad de un evento aleatorio se puede estudiar únicamente con referencia a un espacio de descripciones muestrales sobre el cual se haya definido una función de probabilidad. Sin embargo, en muchas aplicaciones de la teoría de probabilidad la terminología de los espacios de descripciones muestrales no interviene explícitamente (aunque, el concepto siempre interviene implícitamente). En cambio, muchas aplicaciones de la teoría de probabilidades se basan en el concepto de *variable aleatoria*.

2.4.1 Concepto

Hemos definido un evento como un conjunto de descripciones muestrales; por lo tanto, los eventos de fenómenos aleatorios con resultados numéricos son conjuntos de números reales, sin embargo, no todo conjunto de números reales puede considerarse como evento. Por lo tanto, por la palabra evento entenderemos no solo un conjunto de números reales, sino un conjunto probabilizable de números reales. No nos involucraremos para caracterizar los conjuntos de números reales que son probabilizables, únicamente podemos señalar que la familia \mathfrak{S} de conjuntos probabilizables siempre tiene las propiedades siguientes:

- a) Cualquier intervalo pertenece a \mathfrak{S} (un intervalo es un conjunto de números reales de la forma $\{x: a < x < b\}$, $\{x: a \leq x < b\}$, $\{x: a < x \leq b\}$, $\{x: a \leq x \leq b\}$, en donde a y b pueden ser números finitos o infinitos.
- b) A \mathfrak{S} pertenece el complemento A^c de cualquier conjunto A que pertenezca a \mathfrak{S} .
- c) A \mathfrak{S} pertenece la unión $\bigcup_{n=1}^{\infty} A_n$ de cualquier sucesión de conjuntos $A_1, A_2, \dots, A_n, \dots$ que pertenezcan a \mathfrak{S} .

Una definición más precisa del concepto de evento, se puede hacer de la siguiente manera. Existe en la recta real una familia más pequeña de conjuntos con las propiedades a), b) y c), esta familia se denota por \mathfrak{B} , y cualquier miembro de \mathfrak{B} , se llama conjunto boreliano. Puesto que \mathfrak{B} es la familia más pequeña que posee las propiedades a), b) y c), se deduce que \mathfrak{B} está contenida en

\mathfrak{S} , la familia de conjuntos probabilizables. Entonces, todo conjunto boreliano es probabilizable. Por lo tanto entenderemos por un evento relativo a un fenómeno aleatorio con resultados numéricos, un conjunto boreliano de números reales.

El concepto de variable aleatoria está íntimamente relacionado con el concepto de función.

“Decimos que un objeto X o $X(\cdot)$ es una función definida en un espacio Ω , si para todo elemento ω de Ω hay un número real, denotado por $X(\omega)$, al cual llamamos el valor de la función X en ω .”¹³

Ahora veamos la definición de variable aleatoria:

“Decimos que un objeto X es una variable aleatoria si:

- i. Es una función de valores reales definida en un espacio de descripciones muestrales, sobre una familia de cuyos subconjuntos hayamos definido una función de probabilidad $P[\cdot]$.*
- ii. Para todo conjunto boreliano \wp de números reales, el conjunto $\{\omega : X(\omega) \in \wp\}$ pertenece al dominio de $P[\cdot]$.”¹⁴*

La variable X no recibe el calificativo de aleatoria por el hecho que atribuya de modo imprevisible un valor cualquiera a un elemento ω de Ω , ya que este valor está definido de forma precisa (determinista). Lo que es aleatorio en realidad, es que al hacer el experimento, no sabemos que elemento de Ω puede ocurrir.

2.4.2 Descripción de una Variable Aleatoria

Por definición una variable aleatoria X es una función sobre un espacio de probabilidades, la forma funcional de X raramente nos preocupa, puesto que no nos interesa calcular el valor de $X(\omega)$ que la función X asume en un elemento individual ω del espacio de descripciones muestrales Ω , sobre el cual se ha definido X , más bien, nos interesa la probabilidad de que un valor observado de la variable aleatoria X pertenezca a un conjunto dado \wp .

¹³ Ver, Parzen “Teoría Moderna de Probabilidades y sus Aplicaciones”, p.300

¹⁴ Ver, Parzen “Teoría Moderna de Probabilidades y sus Aplicaciones”, p.300

2.4.2.1 Función de Probabilidad de Una Variable Aleatoria

La función de probabilidad de una variable aleatoria X , comúnmente denotada como $P_X[\cdot]$, es una función de conjunto¹⁵ definida para todo conjunto boreliano \wp de números reales cuyo valor $P_X[\wp]$ es la probabilidad de que X este en \wp . Las siguientes expresiones se usan de manera similar que $P_X[\wp]$. La función de probabilidad $P_X[\wp]$ de la variable aleatoria X se obtiene de la función de probabilidades $P[\cdot]$, que existe en el espacio de descripciones muestrales de Ω sobre el cual se ha definido X como una función, por medio de la siguiente fórmula básica:

$$P_X[\wp] = P[\{\omega : X(\omega) \in \wp\}]$$

La anterior ecuación representa la definición de $P_X[\wp]$. El significado intuitivo de $P_X[\wp]$ anteriormente expresado está implícito en la ecuación, puesto que la función X tendrá un valor observado perteneciente al conjunto \wp si, y sólo si, el valor observado de ω del fenómeno aleatorio es tal que $X(\omega)$ esta en \wp .

2.4.2.2 Ley de Probabilidad de una Variable Aleatoria

La ley de probabilidades de una variable aleatoria X se define como una función de probabilidades $P_X[\cdot]$ en la recta real, que coincida con la función de probabilidades $P_X[\wp]$ de la variable aleatoria X . La teoría de probabilidades estudia las aseveraciones que se pueden hacer acerca de una variable aleatoria, cuando se conoce únicamente su ley de probabilidades. En consecuencia, si enunciamos una proposición acerca de una función de probabilidades $P[\cdot]$, entonces, desde el punto de vista de la teoría de probabilidades, enunciaremos una proposición acerca de todas las variables aleatorias X, Y, \dots , cuyas funciones de probabilidad $P_X[\cdot], P_Y[\cdot], \dots$ coincidan con $P[\cdot]$. Dos variables aleatorias X y Y están idénticamente distribuidas si sus funciones de probabilidad son iguales; es decir, si $P_X[\wp] = P_Y[\wp]$ para todos los conjuntos borelianos \wp .

¹⁵ “Una función de conjunto es una función real cuyo dominio es una colección de conjuntos.” Ver <<http://www.mfc.uclv.edu.cu/>>

2.4.2.3 Función de Distribución ¹⁶

"La función de distribución de una variable aleatoria X , denotada por $F_X(\bullet)$, para todo número real x , se define por,

$$F_X(x) = P[X \leq x].^{17}$$

Las propiedades de una función de distribución $F_X(\bullet)$ son:

- i. $F_X(\bullet)$ debe ser no decreciente, en el sentido de que cualesquiera números reales a y b se verifica que,

$$F_X(a) \leq F_X(b) \text{ si } a < b;$$

- ii. Los límites de $F_X(x)$, cuando x tiende a ya sea a más o menos infinito, debe existir y ser

$$\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow +\infty} F_X(x) = 1;$$

- iii. En cualquier punto x ,

$$\lim_{b \rightarrow x^+} F_X(b) = F_X(x)$$

- iv. En cualquier punto x ,

$$\lim_{a \rightarrow x^-} F_X(a) = F_X(x) - p(x)$$

donde $p(x)$ es la probabilidad de que el valor observado del fenómeno aleatorio sea igual a x .

La función de distribución de X determina de manera única la función de probabilidades de X .

La función de masa de probabilidades de una variable aleatoria X , denotada por $p_X(\bullet)$, es una función cuyo valor $p_X(\bullet)$, en cualquier número real x , representa la probabilidad de que el valor observado de la variable aleatoria X sea igual a x , es decir,

$$p_X(x) = P[X = x] = P_X[\{x' : x' = x\}]$$

¹⁶ Frecuentemente también llamada *FUNCIÓN DE DISTRIBUCIÓN ACUMULADA*.

¹⁷ Ver Parzen, "Teoría Moderna de Probabilidades y sus Aplicaciones", p.303

Un punto de masa de probabilidades de la variable aleatoria X es un número real x para el cual $p_X(\bullet)$ es positiva. A partir de la función de distribución $F_X(\bullet)$ obtenemos la función de masa de probabilidades $p_X(\bullet)$ mediante

$$p_X(x) = F_X(x) - \lim_{a \rightarrow x^-} F_X(a)$$

Una variable aleatoria X es discreta si la suma de la función de masa de probabilidades sobre los puntos en los cuales es positiva es igual a 1, es decir,

$$\sum_{x, p_X(x) > 0} p_X(x) = 1$$

En otras palabras, una variable aleatoria X será discreta cuando distribuyamos una unidad de masa en la recta infinita de acuerdo con la función de probabilidades $P_X[\bullet]$, y si al hacer esto asignamos una masa positiva $p_X(x)$ a cada número finito o infinito numerable de puntos.

Si una variable aleatoria X es discreta, basta conocer su función de masa de probabilidades $p_X(\bullet)$ para conocer su función de probabilidades $P_X[\bullet]$, ya que $P_X[\bullet]$ puede expresarse en términos de $p_X(\bullet)$.

"Si X es discreta, entonces, para cualquier conjunto boreliano \wp de números reales, tenemos que:

$$P_X[\wp] = P[X \text{ esté en } \wp] = \sum_{x \in \wp, p_X(x) > 0} p_X(x). \text{ }^{18}$$

Así pues, para evaluar la probabilidad $P_X[\wp]$, con respecto a una variable aleatoria X discreta, de que la variable aleatoria X , tenga un valor observado de esté en \wp , no tenemos sino que enumerar los puntos de masa de probabilidad de X que estén en \wp . Sumamos entonces las masas de probabilidad adjudicadas a estos puntos de masa de probabilidad para obtener $P_X[\wp]$.

La función de distribución de una variable aleatoria discreta X se expresa en términos de su función de masa de probabilidades como:

¹⁸ Ver Parzen, "Teoría Moderna de probabilidades y sus Aplicaciones", p. 304

$$F_X(x) = \sum_{x' \leq x, p_X(x') > 0} p_X(x').$$

La función de distribución $F_X(\cdot)$ de una variable aleatoria X discreta es una función constante por intervalos o función *escalonada*. Consiste en una serie de rectas horizontales en los intervalos comprendidos entre los puntos de masa de probabilidades; en un punto de masa de probabilidades x , la gráfica de $F_X(\cdot)$ da un salto hacia arriba igual a $p_X(x)$.

Se dice que una variable aleatoria X es continua cuando toma valores en cualquier punto de un intervalo (a, b) de la recta real. Si una variable aleatoria X es continua, existe una función no negativa $f_X(\cdot)$, llamada función de densidad de probabilidades de la variable aleatoria X , que tiene la propiedad siguiente, para cualquier conjunto boreliano \wp de números reales.

$$P_X[\wp] = P[X \in \wp] = \int_{\wp} f_X(x) dx$$

Es decir, una vez que se conoce la función de densidad de probabilidades $f_X(\cdot)$, para una variable aleatoria continua X , obtenemos el valor $P_X[\wp]$ de la función de probabilidades en cualquier conjunto boreliano \wp al integrar la función de densidad de probabilidades $f_X(\cdot)$ sobre el conjunto \wp .

La función de distribución $F_X(\cdot)$ de una variable aleatoria continua se obtiene, en términos de su función de densidad de probabilidades, mediante

$$F_X(x) = \int_{-\infty}^x f_X(x') dx'$$

A su vez, la función de densidad de probabilidades de una variable continua se obtiene mediante la derivación de su función de distribución,

$$f_X(x) = \frac{d}{dx} F_X(x)$$

en todos los puntos x en que exista la derivada del miembro derecho en la anterior fórmula.

También la función de distribución sirve para clasificar las variables aleatorias según sus tipos. Diremos que una variable aleatoria X es discreta o continua si su función de distribución $F_X(\cdot)$ es discreta o continua, respectivamente. Se le llama discreta a una ley de probabilidades si corresponde a una función de distribución discreta, y continua si corresponde a una función de distribución continua.

Las leyes de probabilidades pueden clasificarse en familias tomando formas funcionales similares. Por ejemplo, considere la función $b(\cdot; n, p)$ definida para cualquier $n=1, 2, \dots$, y $0 \leq p \leq 1$ por

$$b(x, n, p) = \binom{n}{x} p^x q^{n-x} \text{ para } x=0, 1, \dots, n$$

$$= 0 \text{ en otro caso.}$$

Para valores fijos de n y p la función $b(\cdot; n, p)$ es una función de masa de probabilidades, y por ello define una ley de probabilidades. Las leyes de probabilidades determinadas por $b(\cdot; n_1, p_1)$ y $b(\cdot; n_2, p_2)$ para dos conjuntos diferentes de valores n_1, p_1 y n_2, p_2 son diferentes. No obstante, la forma funcional común de las dos funciones $b(\cdot; n_1, p_1)$ y $b(\cdot; n_2, p_2)$ nos permite tratar simultáneamente las dos leyes de probabilidades que determinan. Llamamos a n y p parámetros, y a $b(\cdot; n, p)$ la función de masa de la ley de probabilidades binomial con parámetros n y p . En los anexos 1 y 2 presentamos un inventario de las leyes de probabilidades discretas y continuas que son más usuales.

2.5 Esperanza de una Variable Aleatoria

Al estudiar las variables aleatorias, es tan importante conocer sus medias y variancias como lo es conocer sus leyes de probabilidades. Definiremos el concepto de esperanza de una variable aleatoria.

2.5.1 Esperanza, Media y Varianza de una Variable Aleatoria

Dada la variable aleatoria X , definimos la esperanza de una variable aleatoria, denotada por $E[X]$, como la media de la ley de probabilidades de X ; es decir,

$$E[X] = \begin{cases} \int_{-\infty}^{\infty} x dF_X(x) \\ \int_{-\infty}^{\infty} x f_X(x) dx \\ \sum_x x p_X(x) \end{cases}$$

Según se especifique X mediante su función de distribución $F_X(\cdot)$, su función de densidad de probabilidad $f(\cdot)$, o su función de masa de probabilidades $p_X(\cdot)$.

Dada una variable aleatoria Y , que surge como una función boreliana¹⁹ de una variable aleatoria X , de manera que,

$$Y = g(X).$$

Para una función boreliana $g(\cdot)$, la esperanza $E[g(X)]$, de acuerdo a la definición de esperanza, está dada por,

$$E[g(X)] = \int_{-\infty}^{\infty} y dF_{g(x)}(y)$$

Por otra parte dada la función de Borel $g(\cdot)$ y la variable aleatoria X , podemos definir la esperanza de $g(x)$ con respecto a la ley de probabilidades de X , denotada por $E_X[g(x)]$, como

$$E_X[g(x)] = \begin{cases} \int_{-\infty}^{\infty} g(x) dF_X(x) \\ \int_{-\infty}^{\infty} g(x) f_X(x) dx \\ \sum_x g(x) p_X(x) \end{cases}$$

Según se defina X mediante su función de distribución $F_X(\cdot)$, su función de densidad de probabilidad $f_X(\cdot)$, o su función de masa de probabilidades $p_X(\cdot)$. Un hecho de gran importancia, que para cualquier variable aleatoria X y función boreliana $g(\cdot)$, se tiene, $E[g(X)] = E_X[g(x)]$.

¹⁹ Una función $f(\cdot)$ es una función boreliana si para cualquier número real c , el conjunto $\{x: f(x) < c\}$ es un conjunto boreliano, ver Parzen "Teoría Moderna de Probabilidades y sus Aplicaciones", p. 175

Si existe alguna de estas esperanzas. La igualdad anterior dice que la esperanza de la variable aleatoria $g(X)$ es igual a la esperanza de la función $g(\cdot)$ con respecto a la variable aleatoria X . Así pues dadas una variable aleatoria X y una función $g(\cdot)$, encontramos dos conceptos distintos representados por $E[g(x)]$ y $E_X[g(x)]$ que, sin embargo, son siempre numéricamente iguales. Se acostumbra usar la notación $E[g(x)]$, por ser más conveniente para las manipulaciones técnicas. No hay que olvidar que aunque se escriba $E[g(x)]$, el concepto que nos interesa suele ser $E_X[g(x)]$, la esperanza de la función $g(x)$ con respecto a la variable aleatoria X .

Dada una variable aleatoria X , denotamos su media por $E[X]$, su media cuadrática por $E[X^2]$, el cuadrado de su media por $E^2[X]$, y su n -ésimo momento con respecto al punto c por $E[(X-c)^n]$, y su n -ésimo momento central por $E[(X-E[X])^n]$. En particular, la varianza de una variable aleatoria, denotada por $Var[X]$, se define como su segundo momento central, de manera que se verifica,

$$Var[X] = E[(X - E[X])^2] = E[X^2] - E^2[X].$$

La desviación estándar de una variable aleatoria, denotada por $\sigma[X]$, se define como la raíz cuadrada de positiva de su varianza, de manera que,

$$\sigma[X] = \sqrt{Var[X]}, \quad \sigma^2[X] = Var[X].$$

Las propiedades de la media y la varianza son, dadas X y Y variables aleatorias y $a, b \in \mathbb{R}$, se cumple para la media:

1. $E[aX + b] = aE[X] + b$;
2. $E[X + Y] = E[X] + E[Y]$.

Y para la varianza:

1. $Var[X] = E[X^2] - (E[X])^2$;
2. $Var[b] = 0$;

3. $Var[aX] = a^2 Var[X]$;
4. $Var[aX + b] = a^2 Var[X]$.

2.5.2 Función Generadora de Momentos

La función generadora de momentos de una variable aleatoria, denotada por $M_X(\bullet)$, se define para cualquier número real t mediante,

$$M_X(t) = E(e^{tx}) = \begin{cases} \sum_{x \in \mathbb{R}_x} e^{tx} p_X(x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx & \text{si } X \text{ es continua} \end{cases}$$

Siempre que el valor esperado exista para todo $t \in (-h, h)$. Esta última es una condición técnica necesaria para que $M_X(t)$ sea diferenciable en 0. Se denomina función generadora de momentos porque los momentos ($E(X^n)$) de X , pueden ser obtenidos derivando esta función y evaluando la derivada en $t = 0$, es decir,

$$E(X^n) = \left. \frac{\partial^n}{\partial t^n} M_X(t) \right|_{t=0}.$$

La función generadora de momentos además de permitir calcular los momentos de una variable aleatoria, también permite identificar la función de densidad o de probabilidad de una variable aleatoria debido a la propiedad de unicidad, la cual establece que hay una correspondencia uno a uno entre funciones de densidad o de probabilidad y funciones generadoras de momentos.

2.5.3 Función Característica

La función característica de una variable aleatoria X , denotada por $\phi_X(\bullet)$, se define para cualquier número real t mediante,

$$\phi_X(t) = E(e^{itX}) = E \cos(tX) + i \sin(tX).^{20}$$

Es decir, $\phi_X(t)$ es la esperanza de la variable aleatoria e^{itX} . Algunas de las propiedades más relevantes de esta función son:

1. La función característica está acotada por 1, es decir $|\phi_X(t)| \leq 1$, para todo t ;
2. $\phi_X(0) = 1$;
3. $\overline{\phi_X(t)} = \phi_X(-t)$, donde \bar{z} denota el conjugado complejo de z ;
4. $\phi_X(t)$ es uniformemente continua en \mathbb{R} ;
5. Si X y Y son variables aleatorias independientes, entonces $\phi_{X+Y} = \phi_X \phi_Y$;
6. La función característica determina la función de distribución; por lo tanto, $\phi_X(t) = \phi_Y(t)$ si y solo si $F_X(x) = F_Y(y)$. Esta es una consecuencia de la Fórmula de Inversión²¹.
7. Una variable aleatoria X posee una distribución simétrica si, y solo si, $\phi_X(t) \in \mathbb{R}$, para todo $t \in \mathbb{R}$.
8. Par los números reales a y b , $\phi_{aX+b}(t) = e^{itb} \phi_X(at)$;
9. Si $E|X|^n < \infty$, entonces $\phi_X(t)$ tiene n -derivadas continuas y

$$\frac{d^k \phi_X}{dt^k}(t) = \phi_X^{(k)}(t) = \int_{\mathbb{R}} (it)^k e^{itx} f_X(x) dx, 1 \leq k \leq n.$$

Particularmente, $\phi_X^{(k)}(0) = i^k EX^k$, las funciones características son similares a las funciones generadoras de momentos, en este sentido.

²⁰ Si F_X es la función de distribución asociada a X , por las propiedades de la esperanza tenemos, $\phi_X(t) = \int_{\mathbb{R}} e^{itx} dF_X(x)$, la cual se conoce como la transformada de Fourier-Stieltjes y provee una definición alternativa de la función característica.

²¹ Dada una variable aleatoria X con función característica ϕ y función de distribución F , si x y y son puntos continuos de F , tales que $x < y$, entonces

$$F(x) - F(y) = \frac{1}{2\pi} \lim_{\delta \rightarrow \infty} \int_{-\delta}^{\delta} \frac{e^{-itx} - e^{-ity}}{it} \phi(t) dt.$$

2.5.4 La Ley de los Grandes Números y el Teorema del Límite Central

En las aplicaciones de la teoría de probabilidades a fenómenos reales, hay dos resultados que son muy destacados. Conocemos estos resultados como la *Ley de los Grandes Números* y el *Teorema del Límite Central*.

Decimos que un conjunto de n observaciones X_1, X_2, \dots, X_n constituye una muestra aleatoria de una variable aleatoria X , si X_1, X_2, \dots, X_n son variables aleatoria independientes y están idénticamente distribuidas como X . Sea $S_n = X_1 + X_2 + \dots + X_n$ la suma de las observaciones, llamamos a su media aritmética $\frac{1}{n}S_n$, media muestral. Por medio de las definiciones antes mencionadas se tiene:

$$E[S_n] = nE[X], \quad \text{Var}[S_n] = n\text{Var}[X], \quad \psi_{S_n}(t) = [\psi_X(t)]^n$$

$$E[M_n] = E[X], \quad \text{Var}[M_n] = \frac{1}{n}\text{Var}[X], \quad \psi_{M_n}(t) = \left[\psi_X\left(\frac{t}{n}\right)\right]^n$$

Tenemos entonces que la media muestral $\frac{1}{n}S_n$ tiende a cero cuando el tamaño de la muestra n tiende a infinito. Ahora bien, por la desigualdad de Chebyshev²², sabemos que si una variable aleatoria tiene una varianza pequeña, entonces es aproximadamente igual a su media, en el sentido de que con una probabilidad cercana a 1, una observación de la variable aleatoria dará un valor observado aproximadamente igual a la media; en particular, la probabilidad de que un valor observado de la variable aleatoria esté dentro de tres desviaciones estándar de la media es de 0.997. Así pues, hemos establecido que la media muestral $\frac{1}{n}S_n$ de una muestra aleatoria X_1, X_2, \dots, X_n de una variable aleatoria, es aproximadamente igual a la media, con una probabilidad de que se acercara a 1 tanto como queramos si simplemente tomamos una muestra lo suficientemente grande. Este hecho conocido como la *Ley de los Grandes Números*, lo descubrió por vez primera Jakob Bernoulli a finales del siglo XVII. Por otra parte, además de que la media muestral $\frac{1}{n}S_n$ tiende a ser igual a la media $E[X]$, podemos demostrar otras propiedades de la

²² “Si una variable aleatoria X tiene una media finita μ y una varianza finita σ^2 , entonces para toda $k > 0$, $P(|x - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$ ”, ver <<http://mathworld.wolfram.com/ChebyshevInequality.html>>”

media muestral $\frac{1}{n}S_n$. Podemos evaluar aproximadamente, para cualquier intervalo alrededor de la media $E[X]$, la probabilidad de que la media muestral $\frac{1}{n}S_n$ tenga un valor observado en ese intervalo, puesto que la media muestral está distribuida aproximadamente como una distribución normal cuando el tamaño de la muestra es grande. De manera más general se puede decir que si S_n es la suma de las variables aleatorias independientes e idénticamente distribuidas, X_1, X_2, \dots, X_n , con medias y varianzas finitas, entonces para cualesquiera números reales,

$$P[a \leq S_n \leq b] = P\left[\frac{a - E[S_n]}{\sigma[S_n]} \leq \frac{S_n - E[S_n]}{\sigma[S_n]} \leq \frac{b - E[S_n]}{\sigma[S_n]}\right]$$

$$\doteq \Phi\left(\frac{b - E[S_n]}{\sigma[S_n]}\right) - \Phi\left(\frac{a - E[S_n]}{\sigma[S_n]}\right)$$

La anterior ecuación es una representación de uno de los teoremas más importantes de la teoría de probabilidades, en palabras este dice que si tenemos n variables aleatorias y todas ellas siguen el mismo modelo de distribución (cualquiera que éste sea), la suma S_n se distribuye según una distribución normal estándar, cuando n tienda al infinito. En 1920, George Polya dio a este teorema el nombre de *Teorema Central del Límite de la Teoría de Probabilidades*.

Hemos mencionado que la teoría de la probabilidad se encarga del estudio de los fenómenos aleatorios y propone modelos matemáticos para estos fenómenos, también asigna, por medio de una construcción axiomática, las probabilidades de ocurrencia de los eventos que genera el espacio de descripciones muestrales de estos fenómenos. Por lo tanto la probabilidad nos proporciona los modelos matemáticos los cuales emplearemos para modelar el monto de siniestros ocurridos.

CAPÍTULO 3

Métodos y Técnicas Estadísticas

"Llegaré el día en el que la Estadística será una condición tan necesaria para la convivencia como la capacidad de leer y escribir."

Anónimo

En el Año de 1760, el profesor de la Universidad de Göttingen, Gottfried Achenwall acuñó la palabra estadística, que extrajo del término italiano statista (estadista). La raíz remota de la palabra se halla, por otra parte, en el término latino status, que significa estado o situación; esta etimología aumenta el valor intrínseco de la palabra, por cuanto la estadística revela el sentido cuantitativo de las más variadas situaciones. Tomaremos la siguiente definición de estadística.

*"Ciencia que se ocupa del estudio de fenómenos de tipo genérico, normalmente complejos y enmarcados en un universo variable, mediante el empleo de modelos de reducción de la información y de análisis de validación de los resultados en términos de representatividad"*²³

La estadística para su mejor comprensión y estudio se ha dividido en varias ramas, entre las más destacadas tenemos: Estadística Clásica, Estadística Descriptiva, Estadística Inferencial y la Estadística Bayesiana. En el presente trabajo ahondaremos en la Estadística Descriptiva y la Estadística Inferencial.

*"La Estadística Descriptiva, describe, analiza y representa un grupo de datos utilizando métodos numéricos y gráficos que resumen y presentan la información contenida en ellos"*²⁴

*"La Estadística inferencial apoyándose en el cálculo de probabilidades y a partir de datos muestrales, efectúa estimaciones, decisiones, predicciones u otras generalizaciones sobre un conjunto mayor de datos"*²⁵

3.1 Métodos Descriptivos

La estadística descriptiva es la encargada de resumir, clasificar y ordenar los datos, con el fin de tener una visión más precisa y conjunta de las observaciones, intentando descubrir de esta manera posibles relaciones entre los datos, observando cuales toman valores parecidos, cuales difieren grandemente del resto, destacando hechos de posible interés

²³ Ver <<http://www.geocities.com/ResearchTriangle/Facility/1075/Estad.htm>>

²⁴ Ver <<http://ftp.medprev.uma.es/libro/node3.htm>>

²⁵ Ver <<http://ftp.medprev.uma.es/libro/node3.htm>>

3.1.1 Distribución de Frecuencias

Un arreglo de grandes conjuntos de datos nos presenta una imagen valiosa. Así, se podrían ordenar los datos dentro de clases o categorías. A esta forma de agrupar los datos en forma tabular se le llama una distribución de frecuencia.

*"Una distribución de frecuencia es una tabla resumen en la cual los datos son agrupados dentro de clases o categorías ordenadas numéricamente y convenientemente establecidas"*²⁶

Cuando los datos son ordenados y agrupados dentro de tablas de distribución de frecuencia, el proceso de análisis e interpretación de la información se vuelve mucho más fácil y entendible. Una desventaja de las tablas resumen es que no se puede saber como se distribuyen los datos individuales en un intervalo de clase en particular sin acceder a los datos originales. El número de observaciones en cualquier clase, es la frecuencia para esa clase. Las frecuencias de clase relativas se encuentran dividiendo las frecuencias de clase entre el número total de observaciones. Una distribución porcentaje se obtiene multiplicando las frecuencias de clase relativas por 100. Una distribución de frecuencia acumulada representa en cada intervalo de clase la suma de su predecesor hasta sumar el 100% de las observaciones.

3.1.2 Presentación Gráfica

La presentación de la información por medio de una distribución de frecuencias resulta efectiva para mostrar las características más sobresalientes de un grupo de datos; la presentación gráfica de la misma información, hace a las características más importantes inmediatamente aparentes. Entre las formas gráficas más importantes tenemos:

HISTOGRAMAS. Un histograma es una representación gráfica de una distribución de frecuencia, y se forma construyendo barras sobre los intervalos de clase de una distribución de frecuencias. Cuando se dibuja un histograma, los valores de la variable o fenómeno de interés se registran a lo largo de la escala horizontal, marcando las fronteras de clase. A través de la escala vertical se marca el número, proporción, frecuencia o porcentaje de las observaciones por cada intervalo de clase.

²⁶ Ver Berenson, "Applied Statistics: A First Course", p. 79

POLÍGONOS DE FRECUENCIA. Otra de las formas gráficas de representar una variable aleatoria continua es por medio de un polígono de frecuencias.

*"Un polígono de frecuencias es una representación de la frecuencia o frecuencia relativa asociada con cada intervalo de una variable continua por medio de la unión de los puntos de las alturas en una gráfica de líneas"*²⁷

OJIVAS. Así como una distribución de frecuencias puede representarse gráficamente por un histograma, una distribución de frecuencias acumulada puede representarse gráficamente por medio de una "ojiva". Para construir una ojiva, primero señalamos las fronteras de clase en la escala horizontal, tal como se hace con el histograma. Sobre cada frontera de clase dibujamos un punto a una distancia vertical proporcional a la frecuencia acumulativa. Estos puntos se conectan entonces por líneas rectas. A la ojiva obtenida frecuentemente también se llama "polígono de frecuencias acumuladas".

3.1.3 Medidas de Tendencia Central

Las medidas de tendencia central se refieren a los estadísticos descriptivos, los cuales indican el valor más representativo en una distribución de frecuencias.

LA MODA. Es el valor o intervalo en una distribución de frecuencias con la más alta frecuencia. En una distribución de frecuencia para datos sin agrupar, la moda es el valor que ocurre más comúnmente, el valor asociado con la frecuencia más alta. Para una distribución de frecuencia para datos agrupados la moda es el punto medio del intervalo con la más alta frecuencia.

LA MEDIA. Para las observaciones x_1, x_2, \dots, x_n es el promedio aritmético de éstas y se denota por

$$\bar{x} = \sum_{i=1}^n x_i / n$$

La Media es una medida apropiada de tendencia central para muchos conjuntos de datos, sin embargo, dado que cualquier observación en el conjunto se emplea para su cálculo, el valor de la media puede afectarse de manera desproporcionada por la existencia de algunos valores extremos. Para calcular la media con base en los datos agrupados, sea k el número de clases y x_i el punto medio de la i -ésima clase. Entonces el valor aproximado de la media es,

²⁷ Ver McGrath, "Understanding Statistics: A Research Perspective", p. 52

$$\bar{x} = \sum_{i=1}^k f_i x_i / n$$

en donde f_i es la frecuencia de la i -ésima clase y $n = \sum_{i=1}^k f_i$

LA MEDIANA. Para un conjunto de observaciones es el valor para el cual, cuando todas la observaciones se ordenan de manera creciente, la mitad de éstas es menor que este valor y la otra mitad mayor. Para datos agrupados, la mediana es aquel valor que divide en dos partes iguales la distribución de frecuencia relativa.

$$\text{Mediana} = L + c(j / f_m)$$

en donde L es el límite inferior de la clase donde se encuentra la mediana, f_m es la frecuencia de esa clase, c es la longitud de la clase y j es el número de observaciones en esta clase, necesarias para completar un total de $n/2$.

3.1.4 Medidas de Variabilidad

Las medidas de variabilidad describen el monto de variabilidad en una distribución de frecuencias. Entre las más importantes tenemos:

VARIANZA. La varianza puede ser definida como la diferencia del cuadrado de los promedios y su media. La versión comúnmente más usada para el cálculo de la varianza es con la fórmula,

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

La varianza es una medida razonablemente buena de la variabilidad debido a que si muchas de las diferencias son grandes (o pequeñas) entonces el valor de la varianza será grande (o pequeño). El valor de la varianza puede sufrir un cambio muy desproporcionado, aún más que la media, por la existencia de algunos valores extremos en el conjunto.

DESVIACIÓN ESTÁNDAR. La raíz cuadrada positiva de la varianza recibe el nombre de desviación estándar y se denota por

$$s_x = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1)}$$

La varianza y la desviación estándar no son medidas de variabilidad distintas, debido a que la última no puede determinarse a menos que se conozca la primera. A menudo se prefiere la desviación estándar en relación con la varianza, por que se expresa en las mismas unidades físicas de las observaciones.

3.2 Inferencia Estadística

La Inferencia Estadística es la parte de la estadística matemática que se encarga del estudio de los métodos para la obtención del modelo de probabilidad (forma funcional y parámetros que determinan la función de distribución) que sigue una variable aleatoria de una determinada población, a través de una muestra (parte de la población) obtenida de la misma. Los dos problemas fundamentales que estudia la inferencia estadística son: el problema de la *estimación* y el problema del *contraste de hipótesis*.

En lo que sigue nos vamos a limitar a problemas de inferencia estadística paramétrica, donde la variable aleatoria objeto de estudio sigue una distribución conocida, y sólo tendremos que tratar de estimar los parámetros que la determinan. Esta situación se presenta con frecuencia debido a que es posible a menudo conocer la forma funcional de la distribución de probabilidad, por consideraciones teóricas, quedando únicamente indeterminados los parámetros que determinan la función de distribución.

Como las poblaciones en las que se pretende estudiar una determinada variable aleatoria son grandes, es muy caro o imposible, estudiar a todos sus individuos; lo que se hace, es estudiar una muestra (una parte) de la población. En todos estos problemas que estudia la inferencia estadística juega un papel fundamental la *Teoría de la Probabilidad* (distintas formas funcionales de las distribuciones de probabilidad) y la *Teoría de Muestras* (procedimientos para tomar muestras de manera apropiada).

3.2.1 Estimación

Existen dos tipos de estimaciones, la estimación puntual y por intervalos. El procedimiento puntual utiliza la información de la muestra para obtener un solo número o punto que estima el parámetro. El procedimiento de estimación por intervalos hace uso de la información de la muestra para obtener un intervalo que se supone va a incluir el parámetro en estudio. En cada caso la estimación real se hace mediante un estimador, que es una regla que establece como utilizar los datos de la muestra para determinar el valor (o valores) que utilizamos como la estimación puntual (o por intervalos).

3.2.1.1 Estimadores y Estimaciones

Llamamos problema de ajuste al problema que consiste en determinar, en una familia de leyes de probabilidad dada, aquella que coincide mejor con la muestra observada. En los casos más usuales, la familia depende de uno o dos parámetros reales desconocidos. Por tanto el problema consiste en determinar el valor del parámetro que se adapta mejor a los datos. Hablamos entonces de estimación paramétrica.

Suponemos que una familia de leyes, que depende de un parámetro desconocido θ , ha sido seleccionada. Ahora es de la muestra, y sólo de ella, que se puede extraer la información. Se llama estimador del parámetro θ a toda función de la muestra, que toma valores en el conjunto de los valores posibles de θ . Es importante distinguir entre los valores aleatorios, asociados a la modelación, y sus realizaciones, identificadas a los datos: Una muestra (teórica) es una n -tupla de variables aleatorias independientes (X_1, \dots, X_n) , que siguen una misma ley P_θ . Para estimar, se propone un estimador en función de la muestra:

$$T = \tau(X_1, \dots, X_n).$$

T es también una variable aleatoria. La selección del modelo y del estimador T , está desconectada de la recolección de los datos. Es, en cierta forma, una planificación que se hace antes de realizar las observaciones y que podrá servir a varias muestras que se recojan del mismo fenómeno.

Una vez que se selecciona el modelo, se considerará a una n -tupla de datos (x_1, \dots, x_n) como una realización de las variables aleatorias (X_1, \dots, X_n) . El valor (real) que toma T :

$$\hat{\theta} = \tau(x_1, \dots, x_n),$$

es la estimación del parámetro (a partir de la muestra observada).

3.2.1.2 Cualidades de un Estimador

Intuitivamente, las características que serían deseables para esta nueva variable aleatoria (que usaremos para estimar el parámetro desconocido) deben ser:

- **Consistencia:** Cuando el tamaño de la muestra crece arbitrariamente, el valor estimado se aproxima al parámetro desconocido.
- **Carencia de sesgo:** También llamado Inesgamiento. El valor medio que se obtiene de la estimación para diferentes muestras debe ser el valor del parámetro.
- **Eficiencia:** Al estimador, al ser v.a., no puede exigírsele que para una muestra cualquiera se obtenga como estimación el valor exacto del parámetro. Sin embargo, podemos pedirle que su dispersión con respecto al valor central (varianza) sea tan pequeña como sea posible.
- **Suficiencia:** El estimador debería aprovechar toda la información existente en la muestra.

"Decimos que el estimador (T_n) es consistente (o convergente), si para todo $\varepsilon > 0$:

$$\lim_{n \rightarrow \infty} P[|T_n - \theta| > \varepsilon] = 0 \text{ }^{28}$$

En consecuencia un estimador consistente se aleja del parámetro con una probabilidad débil, si el tamaño de la muestra es lo suficientemente grande.

La noción de consistencia no da ninguna seguridad práctica de que los valores que toma un estimador estarán efectivamente en un radio fijo alrededor del verdadero valor del parámetro, para un tamaño de muestra dado. La calidad de los estimadores se cuantifica con la noción de error cuadrático.

²⁸ <http://www.math-info.univ-paris5.fr/~ycart/emel/cours/ep/node4.html>

“Sea T cualquier estimador de un parámetro desconocido θ , se define al error cuadrático medio de T como el valor esperado del cuadrado de la diferencia entre T y θ , es decir:

$$ECM(T_n, \theta) = E[(T_n - \theta)^2].”^{29}$$

La razón por la cual el ECM es una cantidad importante para enjuiciar a los posibles estimadores de θ se debe a la siguiente argumento, desarrollando el miembro derecho de la anterior relación, obtenemos:

$$ECM(T_n, \theta) = Var(T_n) + [\theta - E(T_n)]^2$$

El error cuadrático medio de cualquier estimador es la suma de dos cantidades no negativas: una es la varianza del estimador y la otra es el cuadrado del sesgo ³⁰ del estimador. Estas dos cantidades se encuentran relacionadas en forma directa con las propiedades deseables de un estimador. De manera específica, la varianza de un estimador debe ser lo más pequeña posible mientras que la distribución de muestreo debe concentrarse alrededor del valor del parámetro. El error cuadrático está ligado a la consistencia por la siguiente proposición.

“Si el error cuadrático de T_n con respecto a θ tiende a 0 cuando n tiende a infinito, entonces T_n es un estimador consistente de θ .”³¹

Si se dispone de dos estimadores para el mismo parámetro θ , diremos que uno es mejor que el otro si su error cuadrático con respecto a θ es menor. Aún para un estimador consistente, puede suceder que los valores que toma estén desplazados, en promedio, con respecto al verdadero valor del parámetro. Decimos entonces que el estimador es sesgado. El sesgo de θ puede ser positivo, negativo, o cero. Ya que el cuadrado del sesgo es un componente del error cuadrático medio, se desearía, que el valor absoluto de este sea lo más pequeño posible, es decir, es deseable que un estimador tenga un medida igual a al del parámetro que se esta estimando.

“Se dice que una estadística $T = u(X_1, X_2, \dots, X_n)$ es un estimador insesgado del parámetro θ , si la $E(\theta) = \theta$ para todos los posibles valores de θ . De esta forma, para

²⁹ Ver Canavos, “Probabilidad y Estadística. Aplicaciones y Métodos”, p

³⁰ En $ECM(T_n, \theta)$, el término $[\theta - E(T_n)]$ recibe el nombre de sesgo del estimador, denotado como $b(\theta)$

³¹ Ver <<http://www.math-info.univ-paris5.fr/~ycart/emel/cours/ep/node4.html>>

cualquier estimador insesgado de θ , la distribución de muestreo se encuentra centrada alrededor de θ y $ECM(T_n, \theta) = Var(T_n)$."³²

Decimos que un estimador T_n es asintóticamente insesgado, si $b(\theta)$ tiende a 0 cuando n tiende a infinito.

Se dice que un estimador $\hat{\theta}$ es suficiente si utiliza toda la información en una muestra relevante para la estimación de θ , es decir, si todo el conocimiento acerca de θ que se pueda ganar de los valores individuales de la muestra y de su orden se pueden ganar igualmente del valor de $\hat{\theta}$ solo. Podemos describir esta propiedad al referirnos a la distribución o densidad de probabilidad condicional de los valores muestrales dado $\hat{\theta}$, que esta dada por:

$$f(x_1, x_2, \dots, x_n | \hat{\theta}) = \frac{f(x_1, x_2, \dots, x_n, \hat{\theta})}{g(\hat{\theta})} = \frac{f(x_1, x_2, \dots, x_n)}{g(\hat{\theta})}$$

Si depende de θ , entonces los valores particulares de X_1, X_2, \dots, X_n que producen $\hat{\theta}$ serán más probables para algunos valores de θ que para otros, y el conocimiento de estos valores muestrales ayudará en la estimación de θ . Por otra parte, si no depende de θ , entonces los valores particulares de X_1, X_2, \dots, X_n que producen $\hat{\theta}$ serán igualmente probables para cualquier valor de θ , y el conocimiento de estos valores muestrales no ayudarán en la estimación de θ .

*"La estadística $\hat{\theta}$ es un estimador suficiente del parámetro θ si y sólo si para cada valor de $\hat{\theta}$ la distribución o densidad de probabilidad condicional de la muestra aleatoria X_1, X_2, \dots, X_n , dado $\hat{\theta}$, es independiente de θ ".*³³

Un criterio para determinar un estadístico suficiente se conoce como el teorema de factorización de Neyman.

³² Ver Canavos, "Probabilidad y Estadística. Aplicaciones y Métodos", p. 255

³³ Ver Freud, "Estadística Matemática con Aplicaciones", p. 337

“Sea T un estadístico basado en una muestra aleatoria X_1, X_2, \dots, X_n . Entonces T es un estadístico suficiente para la estimación de un parámetro θ si y sólo si la función de verosimilitud se puede factorizar en dos funciones no negativas de la siguiente forma:

$$L(x_1, x_2, \dots, x_n) = g(t, \theta)h(x_1, x_2, \dots, x_n)$$

en donde $g(t, \theta)$ es una función solamente de t y θ , $h(x_1, x_2, \dots, x_n)$ no es una función de θ .”³⁴

La utilidad de la estadística suficiente recae en el hecho de que si un estimador insesgado de un parámetro θ es una función de una estadística suficiente, entonces tendrá la varianza más pequeña de entre todos los estimadores insesgados de θ que no se encuentren basados en una estadística suficiente.

3.2.1.3 Intervalos de confianza

Un estimador por intervalo es una regla que especifica el método que utilizan las mediciones de la muestra para calcular dos números que forman los extremos del intervalo. Sería convenientemente ideal que el intervalo tuviera dos propiedades. Primero, que el intervalo contenga al parámetro θ a estimar. Segundo, que el intervalo sea relativamente estrecho. Nótese que uno o ambos extremos del intervalo variarán de manera aleatoria de una muestra a otra, porque son funciones de las mediciones de la muestra. Así la longitud y la localización del intervalo son cantidades aleatorias, y no podemos estar seguros que el parámetro θ a estimar (fijo) se localice realmente entre los extremos de cualquier intervalo calculado a partir de una sola muestra. Dada esta situación, el objetivo es encontrar un estimador por intervalos que genere intervalos angostos que contengan a θ con una alta probabilidad.

Los estimadores por intervalos se denominan comúnmente intervalos de confianza. Los extremos superior e inferior de un intervalo de confianza se llaman límites de confianza superior e inferior, respectivamente. La probabilidad de que un intervalo de confianza contenga a θ se conoce como coeficiente de confianza. Desde un punto de vista práctico, el coeficiente de confianza indica la fracción de veces, en un muestreo repetitivo, que los intervalos construidos contendrán el parámetro θ . Si se sabe que el coeficiente de confianza asociado a un estimador es alto, estaremos altamente confiados que un intervalo de confianza particular, construido a partir de una sola muestra, contendrá θ .

³⁴ Ver Freud, “Estadística Matemática con Aplicaciones”, p. 340

Suponiendo que $\hat{\theta}_L$ y $\hat{\theta}_V$ son los límites de confianza inferior y superior, respectivamente, para un parámetro θ . Entonces, si

$$P(\hat{\theta}_L < \theta < \hat{\theta}_V) = 1 - \alpha$$

la probabilidad $1 - \alpha$ es el coeficiente de confianza. El intervalo aleatorio resultante, definido por $\hat{\theta}_L$ hasta $\hat{\theta}_V$, se denomina intervalo de confianza bilateral. También es posible construir un intervalo de confianza unilateral tal que $P(\hat{\theta}_L < \theta) = 1 - \alpha$, aunque solo un punto es aleatorio en este caso, el intervalo de confianza es $(\hat{\theta}_L, \infty)$. De manera similar, podríamos tener un intervalo de confianza unilateral superior tal que, $P(\theta < \hat{\theta}_V) = 1 - \alpha$, el intervalo de confianza correspondiente es $(-\infty, \hat{\theta}_V)$.

Un método útil para obtener los intervalos de confianza se denomina método del pivote. Este método depende de la determinación de una expresión pivote que posee dos características:

- Es una función de las mediciones de la muestra y el parámetro desconocido θ , en donde θ es la única cantidad desconocida
- Tiene una distribución de probabilidad que no depende del parámetro θ .

Si se conoce la distribución de probabilidad de la cantidad pivote, entonces se puede utilizar la lógica siguiente para obtener el intervalo deseado de estimación. Si Y es una variable aleatoria, c una constante ($c > 0$) y $P(a \leq Y \leq b) = 0.7$, entonces $P(ca \leq cY \leq cb) = 0.7$. De manera similar, para cualquier constante d , $P(a + d \leq Y + d \leq b + d) = 0.7$, es decir, la probabilidad del evento $P(a < Y < b)$ no se altera por un cambio de escala o por una traslación de Y . Por lo tanto, si conocemos la distribución de probabilidad de una cantidad pivote, podemos aplicar a las operaciones descritas anteriormente para obtener el estimador por intervalos deseado.

3.2.2 Métodos de Estimación

En la sección anterior señalábamos algunas de las propiedades deseables de un estimador, ahora presentaremos dos de los métodos más usados para obtener estos estimadores. Estos métodos son: El método de los momentos y el de máxima verosimilitud.

3.2.2.1 Método de los Momentos

El método de los momentos es un procedimiento muy sencillo y quizá el más antiguo para encontrar un estimador para uno o más parámetros poblacionales. Recuérdese que el k -ésimo momento de una variable aleatoria, tomado con respecto al origen, es $\mu'_k = E(X^k)$. El correspondiente k -ésimo momento de la muestra es el promedio $m'_k = \frac{1}{n} \sum_{i=1}^n X_i^k$. El método de los momentos se basa en el supuesto de que los momentos de la muestra deben proporcionar estimaciones apropiadas para los momentos correspondientes de la población.

*"Eljase como estimaciones aquellos valores de los parámetros que son soluciones de las ecuaciones $\mu'_k = m'_k$, $k=1,2,\dots,t$, en donde t es igual al número de parámetros."*³⁵

Como los estimadores obtenidos por el método de momentos son, obviamente, funciones de los momentos muestrales, los estimadores son generalmente consistentes.

3.2.2.2 El Método de Máxima Verosimilitud

En dos artículos publicados a principios de este siglo, el prominente estadístico R. A. Fisher, propuso un método general de estimación llamado el método de máxima verosimilitud. También mostró las ventajas de este método al demostrar que producía estimadores suficientes siempre que estos existieran y que los estimadores de máxima verosimilitud son estimadores asintóticamente insesgado de varianza mínima.

"Si x_1, x_2, \dots, x_n son los valores de una muestra aleatoria de una población con el parámetro θ , la función de verosimilitud de una muestra está dada por:

$$L(\theta) = f(x_1, x_2, \dots, x_n; \theta)$$

*para los valores de θ dentro de un dominio dado. En este caso $f(x_1, x_2, \dots, x_n; \theta)$, es el valor de la distribución de probabilidad conjunta o de la densidad de probabilidad conjunta de las variables aleatorias X_1, X_2, \dots, X_n en $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$."*³⁶

³⁵ Ver Mendenhall, "Estadística Matemática con Aplicaciones", p. 386

³⁶ Ver Freund, "Estadística Matemática con Aplicaciones", p. 346

En esencia, el método de estimación por máxima verosimilitud selecciona como estimador a aquel valor del parámetro que tiene la propiedad de maximizar el valor de la probabilidad de la muestra aleatoria observada.

“Sea $L(\theta) = (\theta, x_1, \dots, x_n)$ la función de verosimilitud de las variables aleatorias X_1, \dots, X_n de una función (densidad) de probabilidad $f(x, \theta)$. Si $t = u(x_1, \dots, x_n)$ es el valor de θ , el cual maximiza $L(\theta)$, entonces $T = u(X_1, \dots, X_n)$ es el estimador de máxima verosimilitud de θ y t es el estimador correspondiente de máxima verosimilitud”

Las principales propiedades de los estimadores obtenidos por el método antes descrito son:

- Son consistentes
- Son invariantes frente a las transformaciones biunívocas, es decir, si $\hat{\theta}_{MV}$ es el estimador de máxima verosimilitud de θ y $g(\hat{\theta})$ es una función biunívoca de $\hat{\theta}$, entonces $g(\hat{\theta}_{MV})$ es el estimador de máxima verosimilitud de $g(\theta)$.
- Si $\hat{\theta}$ es un estimador suficiente de θ , su estimador de máxima verosimilitud, $\hat{\theta}_{MV}$ es función de la muestra a través de $\hat{\theta}$
- Son asintóticamente normales;
- Son asintóticamente eficientes, es decir, entre todos los estimadores consistentes de un parámetro θ , los de máxima verosimilitud son los de varianza mínima.
- No siempre son insesgados.

3.2.3 Pruebas de Hipótesis

Las pruebas estadísticas de hipótesis se realizan en todos los ámbitos en los cuales puede constatarse la teoría frente a la observación. El problema relacionado con las pruebas de hipótesis estadísticas tiene una fuerte relación con el concepto de estimación. La mayoría de las pruebas estadísticas de hipótesis tienen que ver con los parámetros de las distribuciones, pero algunas veces también tienen que ver con el tipo, o naturaleza, de las distribuciones mismas.

“Una hipótesis estadística es una afirmación o conjetura acerca de la distribución de una o más variables aleatorias. Si una hipótesis estadística especifica completamente

*la distribución, se conoce como hipótesis simple; si no, se conoce como hipótesis compuesta.*³⁷

La esencia de probar una hipótesis estadística es decidir si la información se encuentra apoyada por la evidencia experimental que se obtiene a través de una muestra aleatoria. En forma general, la información involucra ya sea algún parámetro o a alguna forma funcional no conocida de la distribución de interés a partir de la cual se obtiene una muestra aleatoria. La decisión acerca de si los datos muestrales apoyan estadísticamente la afirmación se toma con base en la probabilidad, y, si esta es mínima, entonces será rechazada.

3.2.3.1 Elementos de una Prueba Estadística

Cualquier prueba estadística de hipótesis funciona exactamente de la misma manera y se compone de los mismos elementos esenciales. Los elementos de una prueba estadística de hipótesis son:

- Hipótesis nula, H_0
- Hipótesis alternativa H_1
- Estadístico de la prueba
- Región de rechazo

La hipótesis que se desea probar, es la hipótesis nula denotada como H_0 . La hipótesis alternativa (o de estudio), denotada por H_1 es la hipótesis que debe aceptarse en caso de rechazar H_0 . Las partes funcionales de una prueba estadística son el estadístico de la prueba y la región de rechazo. El estadístico de prueba (como un estimador) es una función de las mediciones muestrales en la cual se fundamenta la decisión estadística. La región de rechazo, denotada como RR , especifica los valores del estadístico de la prueba para los cuales se rechaza la hipótesis nula.

3.2.3.2 Prueba de una Hipótesis Estadística

La prueba de una hipótesis estadística es la aplicación de un conjunto explícito de reglas para decidir si aceptamos la hipótesis nula o la rechazamos a favor de la hipótesis alternativa. El procedimiento de prueba, divide los valores posibles de la estadística de prueba en dos subconjuntos: una región de aceptación para H_0 y una región de rechazo para H_0 . Este procedimiento puede conducir a dos clases de errores.

³⁷ Ver Freund, "Estadística Matemática con Aplicaciones", p. 384

“El rechazar la hipótesis nula H_0 , cuando ésta es verdadera, se llama Error tipo I, y el aceptar la hipótesis nula H_0 cuando ésta es falsa es llamado Error tipo II. El tamaño de un Error tipo I se define como la probabilidad de cometer un Error tipo I (denotada por α), de manera similar el tamaño de un Error tipo II es la probabilidad de se cometa un Error tipo II (denotada como β).”³⁸

Es costumbre referirse a la región de rechazo para H_0 como la región crítica de la prueba y la probabilidad de obtener un valor de la estadística de prueba dentro de la región crítica cuando H_0 es verdadera como el tamaño de la región crítica. Así, el tamaño de una región crítica es justamente la probabilidad α de cometer un Error tipo I. Esta probabilidad también se llama el nivel de significancia de la prueba.

Hemos expuesto el proceso de modelar y como los modelos nos facilitan el estudio y comprensión de los fenómenos reales y en particular de los fenómenos aleatorios con la ayuda de modelos matemáticos, también revisamos como la teoría de probabilidad estudia la propiedades de estos fenómenos aleatorios y proporciona modelos matemáticos para describirlos. En este capítulo revisamos los principales métodos y técnicas estadísticas que nos auxiliarán en nuestro proceso en la elección del modelo o los modelos que podrían ajustarse a los datos, calibrar los parámetros desconocidos y seleccionar el que mejor se ajuste a los datos.

³⁸ Ver Mood, “Introduction to the Theory of Statistics” p. 405

CAPÍTULO 4

Modelo para Siniestros Ocurridos: Incendio en Casa Habitación

*"Comprender las cosas que nos rodean es la mejor
preparación para comprender las cosas que hay mas allá."*

Hipatia (375-415)

Hemos revisados los principios que fundamentan la elaboración de un proceso de modelación y, en específico, el proceso para modelar situaciones bajo incertidumbre, así como algunos de los principios de la teoría de la probabilidad y de la estadística matemática. En este capítulo utilizaremos la teoría expuesta para modelar el monto de los siniestros ocurridos en casa habitación para la cobertura de incendio, cabe señalar que estamos intentando modelar el monto de un solo pago para estos siniestros ocurridos, no estamos intentando modelar los montos agregados. La información analizada se tomó de pólizas de hogar para los siniestros ocurridos en 2003 y únicamente aquellos que afectaron la cobertura de Incendio Edificio. Estos datos provienen de una de las aseguradoras más grandes del país, y líder en el ramo de daños que por políticas de confidencialidad de la información no podemos mencionar su nombre.

4.1 Antecedentes

En un sentido muy general, toda la ciencia actuarial se refiere a distribuciones de pérdidas, ya que esto es precisamente lo que un contrato de seguro cubre. La aseguradora pagará un monto aleatorio en un momento futuro también aleatorio. El cálculo del monto depende de los términos del contrato y de la naturaleza del evento cuando éste sucede. Sin embargo, a través del tiempo, la definición de pérdida se ha restringido a un solo pago, para proveer una distinción importante, haremos las siguientes definiciones.

"Un evento de pérdida o siniestro es un incidente en el cual un asegurado o grupo de asegurados sufren algún daño en sus bienes o en sus personas, el cual es potencialmente cubierto por sus contratos de seguros. La pérdida es el monto medido en términos monetarios del daño sufrido en sus bienes o en sus personas por un asegurado o grupo de asegurados como resultado de un evento de pérdida o siniestro. La pérdida también puede ser de cero. Un evento pagable es un incidente en el cual un asegurado o grupo de asegurados reciben un pago como resultado de un evento de pérdida cubierto por sus contratos de seguros. El monto pagado es el monto actual pagado en términos monetarios a los asegurados como resultado de un evento de pérdida o un 'evento pagable'. Si este es el resultado de un evento pagable, el monto pagado puede ser cero."³⁹

No todos los eventos de pérdida generan un pago y algunos eventos de pérdida generan una pérdida. Cuando un evento de pérdida se presenta, probablemente se incurra en gastos como resultado de atender un potencial siniestro, esto nos lleva a la siguiente definición.

³⁹ Ver Klugman, "Loss Models From Data to Decisions", p. 5

“Los Gastos de Ajuste Asignados al Siniestro (ALAE, Allocated Loss Adjustment Expense) es el monto de los gastos incurridos directamente como resultado de un evento de pérdida”⁴⁰

La diferencia entre la pérdida y el monto pagado casi siempre es atribuible a provisiones específicas en la póliza que reducen las obligaciones de la compañía aseguradora. Algunas de las más comunes modificaciones a las coberturas son:

“Un deducible ordinario es una provisión, la cual establece que cuando la pérdida sea menor que o igual al deducible, no habrá pago alguno y cuando la pérdida exceda al deducible, el monto pagado es la pérdida menos el deducible. El límite de póliza o simplemente el límite es una provisión la cual establece que cuando la pérdida exceda este límite, ningún beneficio adicional será pagado. El factor de coaseguro es la proporción de cualquier pérdida que es pagada por el asegurado después de que cualquier otra modificación (como deducibles o límites) ha sido aplicada. Un coaseguro es una provisión la cual establece que un factor de coaseguro se está aplicando.”⁴¹

4.1.1 Distribución de Perdidas

Ahora definiremos lo que es una distribución de pérdidas.

“Una distribución de perdidas es la distribución de probabilidad ya sea de la pérdida o del monto pagado por un evento de pérdida o del monto pagado de un evento pagable. La distribución puede o no excluir pagos de cero o puede o no incluir los ALAE”⁴²

Ya que la definición es ambigua en un número de puntos, el contexto específico debe ser hecho claro cuando una distribución de perdidas esta siendo determinada o evaluada. Para nuestro caso, la distribución de pérdidas que ajustaremos será del monto de siniestros ocurridos, los cuales son la suma de la reserva inicial, mas los ajustes, mas los gastos de ajuste, menos los salvamentos.

Mientras sea posible para una distribución de pérdidas ser basada en los datos empíricos, distribuciones de perdidas basadas en modelos paramétricos suavizados tienen muchas ventajas. El proceso de modelar supone una distribución paramétrica como la fuente generadora de las

⁴⁰ Ver Klugman, “Loss Models From Data to Decisions”, p. 5

⁴¹ Ver Klugman, “Loss Models From Data to Decisions”, p. 6

⁴² Ver Klugman, “Loss Models From Data to Decisions”, p. 6

pérdidas. Otra ventaja de un modelo paramétrico es la simplicidad. Con una palabra (Lognormal, por ejemplo) y dos números (media y varianza) se ha descrito completamente a la población.

4.2 Elección del Modelo

El primer paso en la determinación de nuestro modelo consiste en seleccionar las distribuciones que podría modelar los datos observados.

4.2.1 Métodos Descriptivos

Los datos a ser ajustados son presentados en la siguiente tabla de frecuencias. En el Anexo 3 presentamos las observaciones muestrales.

Clases	Límite Inferior	Límite Superior	Punto Medio	Frecuencia	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Rel. Acum.
<i>en o de bajo</i>		-5,000		0	0.000000	0	0.000000
1	-5000	4,700	-150	30	0.2804000	30	0.2804000
2	4700	14,400	9,550	29	0.2710000	59	0.5514000
3	14400	24,100	19,250	16	0.1485000	75	0.7009000
4	24100	33,800	28,950	8	0.0748000	83	0.7757000
5	33800	43,500	38,650	3	0.0280000	86	0.8037000
6	43500	53,200	48,350	11	0.1028000	97	0.9065000
7	53200	62,900	58,050	5	0.0467000	102	0.9533000
8	62900	72,600	67,750	1	0.0093000	103	0.9626000
9	72600	82,300	77,450	1	0.0093000	104	0.9720000
10	82300	92,000	87,150	3	0.0280000	107	1.0000000
<i>arriba</i>	92000			0	0.0000000		

Tabla 1 Tabla de Frecuencias Siniestros Ocurredos

El primer paso en la determinación de un modelo que se ajuste al monto de los siniestros ocurridos, es producir algunas descripciones de las observaciones muestrales. Empezaremos por las principales medidas de tendencia central, de variabilidad y de forma presentadas en la tabla 2, los valores se encuentran en pesos mexicanos.

<i>Estadístico</i>	<i>Valor</i>
Media	20734.86822
Error típico	2103.07763
Mediana	12757.5
Moda	500
Desviación estándar	21754.40416
Varianza de la muestra	473254100.4
Curtosis	1.096217584
Coefficiente de asimetría	1.330605316
Rango	90410.93
Mínimo	100
Máximo	90510.93
Suma	2218630.9
Cuenta	107

Tabla 2. Estadísticas Descriptivas.

De gran interés son los estadísticos de forma para la asimetría estándar con un valor de 5.61908 y la curtosis estándar con un valor de 2.31464, los cuales pueden ser usados para determinar si las observaciones muestrales son generadas por una distribución normal. Valores para estos estadísticos fuera del rango de -2 a $+2$ indican una desviación importante de normalidad. Los valores para estos estadísticos se encuentran fuera del rango esperado para los datos de una distribución normal. Por lo tanto, como primer punto, podemos aseverar que los datos no los genera una distribución normal.

4.2.2 Métodos Gráficos

Los métodos gráficos son de gran ayuda ya que nos permiten visualizar una posible relación con alguna distribución conocida. La figura 2 presentan el histograma y la ojiva para los datos observados como se había mencionado estas proveen alguna indicación del tipo de distribución que podrían modelar los datos, pero aún así no es evidente. Una mejor vista podría ser obtenida al tomar los logaritmos naturales de los valores observados, en el anexo 3 se presentan los datos puntuales convertidos.

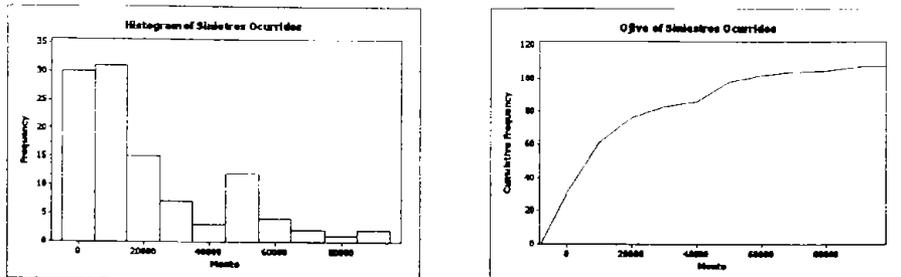


Figura 2. Histograma y ojiva siniestros ocurridos

El histograma y ojiva de los datos transformados se presentan en la figura 3. El histograma revela que la distribución transformada es un tanto simétrica, sugiriendo que una distribución Lognormal podría ser apropiada.

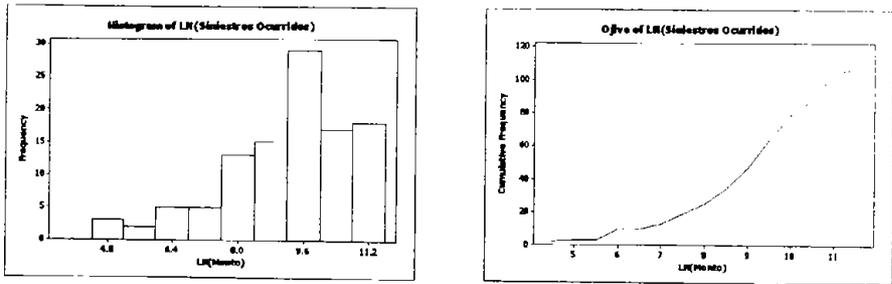


Figura 3. Histograma y Ojiva Ln siniestros ocurridos

Como último paso en la selección del modelo calcularemos los valores de la MRL⁴³. Algún conocimiento de estas funciones puede ser extremadamente valioso para el propósito de modelar.

La MRL se define como:

$$e_x(x) = e(x) = E(X - x | X > x) = \frac{\int_x^{\infty} (t - x) f(t) dt}{S(x)}$$

Intuitivamente, si $e(x)$ toma valores grandes, para valores grandes de x , entonces la distribución tiene una cola pesada ya que la el exceso de la pérdida esperada $X - x$ es grande. Recíprocamente, si $e(x)$ toma valores pequeños para valores grandes de x , tiene una cola ligera. El cálculo de la MRL empírica $e_n(x)$ es relativamente fácil, ya que ésta es simplemente la diferencia del promedio de todas las observaciones mayores o iguales que x . Si ésta pareciera ser una función lineal con pendiente positiva, entonces una distribución Pareto podría ser un modelo apropiado. Si ésta es más parecido a una constante para los valores más grandes de x , entonces se podría tratar el ajuste con una distribución Gamma. Algo entre estas dos, podría sugerir una distribución Lognormal o Weibull con $0 < \tau < 1$. Por otra parte, una función decreciente como $c_1 x^{-\tau+1}$, $\tau > 1$ indica que un modelo Weibull debe ser tomado en cuenta. Si los valores de $e_n(x)$ también pudieran aparentar cambios con varios valores de x , esto podría sugerir que diferentes

⁴³ Mean Residual Life (vida residual promedio)

⁴⁴ Ver Hogg, "Loss Distributions", p. 89

densidades sean usadas en diferentes partes del soporte. En el anexo 3 se presentan los datos puntuales para la MRL.

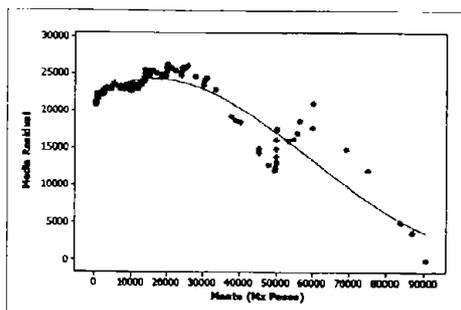


Figura 4. MLR siniestros ocurridos.

La figura 4 muestra los valores para $e_{107}(x)$, la forma de la curva ajustada sugiere que las distribuciones Weibull y Lognormal podrían proporcionar buenos modelos para ajustarse a los datos observados. Por lo tanto tomaremos estas dos distribuciones como posibles generadoras de los valores para los siniestros ocurridos.

4.3 Calibración del Modelo

Con ayuda de la evidencia descriptiva y grafica, hemos concluido que las dos posibles distribuciones generadoras de la información para los datos observados en los montos de siniestros ocurridos las podrían generar las distribuciones Lognormal y Weibull, ahora procederemos a la estimación de los parámetros.

4.3.1 Estimación de Parámetros

Como siguiente paso en nuestro proceso de ajuste es la estimación de los parámetros de las distribuciones que estamos proponiendo como candidatas a modelar los datos empíricos. El método que usaremos para este fin es el bien conocido método de máxima verosimilitud, aunque existen otros métodos de estimación, se eligió este ya que los estimadores encontrados por este método cuentan con muchas características deseables en un estimador y es común encontrar este método de estimación en los paquetes estadísticos. Los resultados de esta estimación son:

Distribución Lognormal				
Parámetro	Estimación	Error Estándar	IC Normal 95%	
			Inferior	Superior
Forma	9.15507	0.149949	8.86118	9.44897
Escala	1.55108	0.10603	1.35659	1.77346

Distribución Weibull				
Parámetro	Estimación	Error Estándar	IC Normal 95%	
			Inferior	Superior
Forma	0.847739	0.0659707	0.727816	0.98742
Escala	19126.5	2293.25	15120.9	24193.3

La figura 5 muestra las probabilidades puntuales ajustadas, al parecer la distribución Weibull se ajusta mejor a los datos que como lo hace la distribución Lognormal, pero aún así esta no es razón suficiente para aceptar una y rechazar la otra, para realizar la selección realizaremos las pruebas de hipótesis necesarias.

4.4 Selección y Validación del Modelo

El siguiente paso en el proceso de ajuste es la selección y validación del modelo lo realizaremos por medio de las pruebas de bondad de ajuste.

4.4.1 Pruebas de Bondad de Ajuste

Ahora probaremos cual de las dos distribuciones se ajusta mejor a los datos por medio de la prueba de Anderson-Darling, esta prueba es usada para probar si una muestra de datos proviene de una población con una distribución específica (al igual que los hacen todas las pruebas de bondad de ajuste), esta es una modificación de la prueba Kolmogorov-Smirnov y da más peso a las colas que la prueba Kolmogorov-Smirnov. La prueba Kolmogorov-Smirnov es de *libre distribución* en el sentido que los valores críticos no dependen de la distribución específica que se esta probando. La prueba de Anderson-Darling hace uso de la distribución específica para calcular los valores críticos. Esto tiene la ventaja de permitir una prueba más sensible y la desventaja de que los valores críticos deben ser calculados para cada distribución. La prueba de Anderson-Darling es una las pruebas de bondad de ajuste alterna a las pruebas Chi-cuadrada y Kolmogorov-Smirnov.

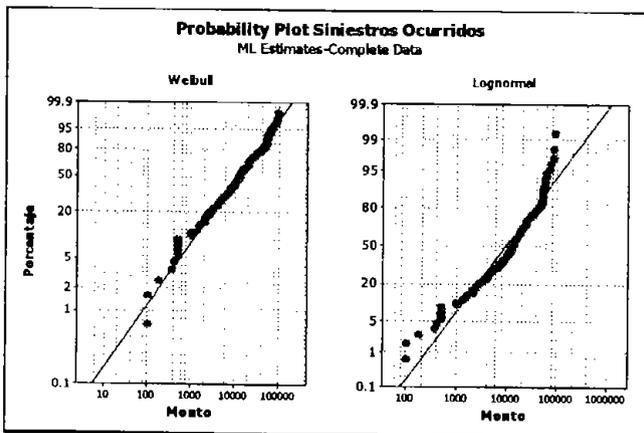


Figura 5. Probabilidades puntuales: siniestros ocurredos

La prueba de Anderson-Darling se define como:

H_0 : Los datos siguen una distribución específica

H_1 : Los datos no siguen una distribución específica

El estadístico de prueba Anderson-Darling es:

$$A^2 = -N - S$$

Donde:

$$S = \sum_{i=1}^N \frac{(2i-1)}{N} [\ln F(Y_i) + \ln(1 - F(Y_{N+1-i}))]$$

F es la función de distribución acumulativa de la distribución especificada, Y_i son los datos ordenados.

Nivel de significancia. α

Región crítica.

Los valores para la prueba Anderson-Darling dependen de la distribución específica a la que se aplica, estos valores son presentados en la tabla 3 para las distribuciones Lognormal y Weibull. Esta prueba es de una cola y la hipótesis que la distribución es de una forma específica es rechazada si el estadístico de prueba, A , es más grande que el valor crítico.

α	10.00%	5.00%	2.50%	1.00%
Lognormal	0.63100	0.75200	0.87300	1.03500
Weibull	0.63700	0.75700	0.87700	1.03800

Tabla 3. Valores críticos Anderson-Darling

Alternativamente, existen técnicas muy sólidas y pruebas no paramétricas que no conciben fuertes suposiciones sobre la distribución subyacente. Sin embargo, técnicas basadas en suposiciones sobre distribuciones específicas son en general más potentes que las no paramétricas. Por lo tanto, si las suposiciones respecto a la distribución pueden ser validados, entonces estas son generalmente preferidos.

Las hipótesis que tenemos que probar son:

$$H_0 : F(x) = F_n(x)$$

$$H_1 : F(x) \neq F_n(x)$$

donde $F_n(x)$ se distribuye como una Lognormal

y

$$H_0 : F(x) = F_n(x)$$

$$H_1 : F(x) \neq F_n(x)$$

donde $F_n(x)$ se distribuye como una Weibull

Los valores obtenidos para el estadístico de prueba Anderson-Darling son:

Prueba Anderson-Darling	
Lognormal	2.055000
Weibull	0.582000

Ya que el valor del estadístico de prueba Anderson-Darling para la distribución Lognormal es mayor a 1.03500, podemos rechazar la hipótesis nula que los datos observados para el monto de los siniestros ocurridos los genera una distribución Lognormal, con un nivel de significancia del 99%.

De forma similar, ya que el valor del estadístico de prueba Anderson-Darling para la distribución Weibull es menor que el 0.6310, no podemos rechazar la hipótesis nula que los datos observados para el monto de los siniestros ocurridos los genera una distribución Weibull con un nivel de significancia del 90%.

Las graficas presentadas, en la figura 6 y 7 muestran la función de distribución acumulativa empírica junto con función de distribución ajustada, para la distribución Weibull y la distribución Lognormal. Nuevamente, podemos observar que la distribución Weibull es la que mejor se ajusta, pero ahora este hecho también lo respalda la prueba de bondad de ajuste Anderson-Darling realizada a ambas distribuciones.

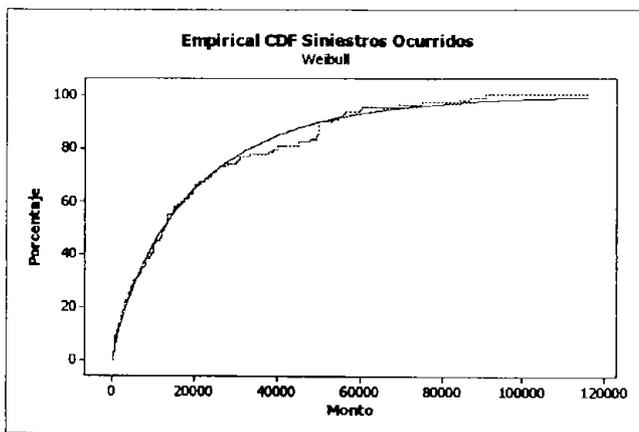


Figura 6. CDF Weibull

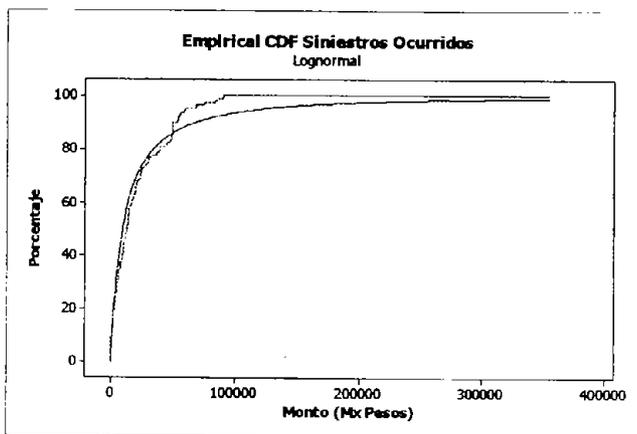


Figura 7. CDF Lognormal

Podemos concluir aseverando que la distribución generadora de la información para el monto de los siniestros ocurredos bien podría ser modelada con una distribución Weibull con un nivel de confianza del 90% y rechazar la distribución Lognormal por medio de la evidencia estadística antes expuesta.

4.5 Aplicación

Hemos presentado el proceso para ajustar distribuciones de probabilidad a los datos de siniestros ocurridos, el cual nos llevó a la conclusión de que la distribución que mejor ajusta el monto de un siniestro ocurrido esta definida por una Weibull. Aplicaremos el modelo para hacer un estimación del moto esperado de un siniestro ocurrido en 2004 así como el calculo de la distribución probabilidades de ocurrencia para valores mayores a un valor específico del monto del siniestro.

4.5.1 El Efecto de la Inflación

Si X es la variable aleatoria que modela los siniestros ocurridos, necesitamos encontrar la distribución de $Z = (1+r)X$, donde r es la inflación en el periodo de estudio. La expresión que relaciona la función de distribución de Z con la de X es:

$$\begin{aligned} F_z(z) &= \Pr\{Z \leq z\} \\ &= \Pr\left[\frac{Z}{1+r} \leq \frac{z}{1+r}\right] \\ &= F_x\left(\frac{z}{1+r}\right) \end{aligned}$$

Además, si X es una variable aleatoria continua, tenemos la siguiente relación entre sus funciones de densidad de probabilidad:

$$\begin{aligned} f_z(z) &= \frac{d}{dz} F_z(z) \\ &= \frac{f_x\left(\frac{z}{1+r}\right)}{1+r} \end{aligned}$$

Por lo tanto al multiplicar una familia de distribuciones por una constante, se obtiene otro miembro de esa familia. Aplicado a nuestro modelo Weibull tenemos:

Distribución	Parámetros de X	Parámetros de $Z = (1+r)X$
Weibull	η, β	$\eta(1+r), \beta$

La inflación anualizada a Agosto de 2004 es de 4.82 %⁴⁵, por lo tanto, el valor para los nuevos parámetros de nuestro modelo son:

$$\eta = (1 + 0.0482) * 19126.5 = 20048.4$$
$$\beta = 0.847739$$

Recordando las principales características de la distribución Weibull, tenemos:

$$f_x(x) = \frac{\beta}{\eta} \left(\frac{x}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\eta}\right)^\beta\right]$$

Donde:

$$\eta > 0, \beta > 0$$

y,

η es el parámetro de escala

β es el parámetro de forma

$$E[X] = \eta * \Gamma\left[1 + \frac{1}{\beta}\right]$$

Por lo tanto, nuestra estimación para el monto medio de un siniestro ocurrido en 2004 considerando el efecto de la inflación es:

$$E[X] = 20048.4 * 1.08978$$
$$= 21848.17$$

El monto promedio de un siniestro ocurrido en 2003 es de 20,734.87 pesos mientras que para 2004 se espera sea de 21,848.17 pesos, esto representa un incremento del 5.36% con relación a 2003.

4.5.2 Probabilidades de Ocurrencia

Con un modelo estimado pueden calcularse las probabilidades de ocurrencia de un evento, en este caso que produzca perdidas por un monto específico en pesos. Por ejemplo, desearíamos conocer

⁴⁵ Fuente, Banco de México.

cual es la probabilidad de que un evento sea mayor a 10,000 pesos. Recordando que la función de distribución para la Weibull es:

$$F_x(x) = 1 - \exp\left[-\left(\frac{x}{\eta}\right)^{\rho}\right]$$

La probabilidad de tener un evento que supere los 10,000.00 pesos, utilizando nuestro modelo y considerando el efecto de la inflación es de:

$$\begin{aligned} \Pr[X > 10,000] &= 1 - F_x(10,000) \\ &= 1 - \left\{ 1 - \exp\left[-\left(\frac{10,000}{20048.4}\right)^{0.847739}\right] \right\} = \exp\left[-\left(\frac{10,000}{20048.4}\right)^{0.847739}\right] \\ &= 0.57435 \end{aligned}$$

Siguiendo el mismo procedimiento, la tabla 4, presenta la probabilidad de que el monto esperado en 2004 para el monto de un siniestro ocurrido sea mayor al valor de x .

x	$\Pr[X > x]$
100	0.9888824
200	0.9800809
300	0.9720253
400	0.9644379
500	0.9571929
1,000	0.9242834
3,000	0.8188738
5,000	0.7348269
10,000	0.5743493
20,000	0.3686332
30,000	0.2448011
50,000	0.1141778
100,000	0.0201349
500,000	0.0000002

Tabla 4. Probabilidades esperadas en 2004.

Podemos observar que la probabilidad de que el monto de un siniestro sea mayor a un monto pequeños es alta y mientras mayor sea el monto esta decrece. Con una tabla de este tipo también sería posible estimara la siniestralidad agregada esperada para algún periodo de interés y de esta forma estimar el nivel de primas necesario para hacer frente a las obligaciones futuras. En general el tener un modelo a la mano al cual se pueda manipular de una manera relativamente fácil y actualizarlo con información nueva nos brinda una amplia gama de posibilidades para su aplicación.

CONCLUSIONES

El uso de modelos tiene una antiquísima historia ya que estos han ayudado a una mejor comprensión, estudio y control de fenómenos reales. La construcción de modelos se puede decir que es un arte ya que la elaboración de estos se entrelazan muchas variables algunas intrínsecas al problemas a resolver y otras relacionadas a las emociones y sensibilidades del investigador que lleva a cabo el proceso de modelación. El proceso de modelar nos permite encontrar una potencial solución a un problema real, una vez encontrada esta solución se puede tener un mejor control sobre el problema ya que se pueden volver a analizar cada paso del proceso y realizar ajustes o mejoras si fuera necesario.

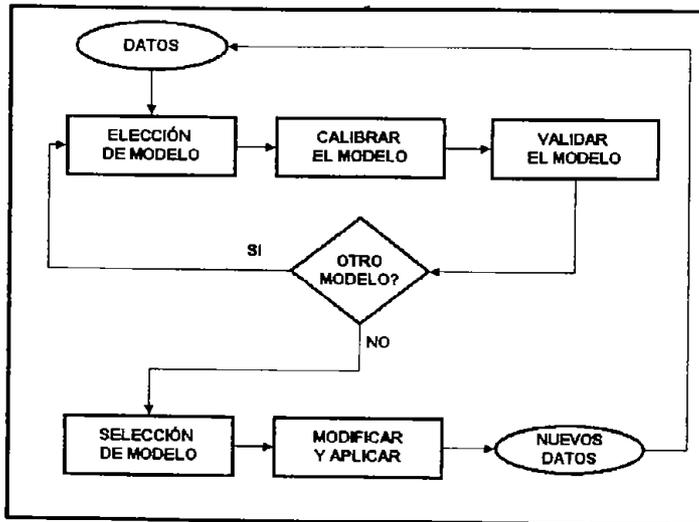
Muchos fenómenos del mundo real tienen un comportamiento aleatorio, la teoría de la probabilidad ha desarrollado varios modelos matemáticos (funciones de distribución) que podrían describir estos fenómenos. La estadística matemática es útil en el proceso de modelar ya que esta provee de métodos y técnicas para poder llevar a cabo, la elección de las función de distribución que podrían ajustarse a los datos de manera adecuada, calibrar los parámetros desconocidos para estas distribuciones, validar y seleccionar el modelo que mejor se ajuste a los datos.

El proceso para ajustar distribuciones de probabilidad a datos se puede resumir en los siguientes pasos:

- 1) Con los datos disponibles, agrupar la información, realizar tablas de frecuencias, calcular medidas numéricas descriptivas.
- 2) Seleccionar los modelos que podrían ser buenos candidatos para modelar la situación en estudio. Se usan histogramas y $e_n(x)$ como guías.
- 3) Calibrar los parámetros desconocidos para los modelos seleccionados.
- 4) Validar los modelos seleccionados para determinar si este se ajusta de forma adecuada a la información empírica. Varias pruebas de diagnostico pueden ser usadas. Algunas de estas pruebas estadísticas son las bien conocidas pruebas de bondad de ajuste como la de Chi-cuadrada, Kolmogorov-Smirnov o la prueba Anderson-Darling. La elección de la prueba esta relacionada directamente con el propósito final del ejercicio de modelar.
- 5) Se pueden considerar otros modelos, este paso es importante en el caso en el que el modelo seleccionado no se ajuste de forma adecuada a la información empírica.
- 6) Seleccionar el modelo que sea mas adecuado.
- 7) El modelo seleccionado es adaptado para aplicaciones futuras. En este trabajo lo usamos para estimar el efecto de la inflación y el calculo de las probabilidades de ocurrencia.

- 8) Con nueva información todos los incisos anteriores deben ser repetidos para mejorar el modelo.

La siguiente figura ilustra este proceso.



Proceso para ajustar distribuciones de probabilidad a datos

Finalmente podemos concluir que la distribución Weibull se ajusta a los datos del monto de un siniestro ocurrido para la cobertura de incendio edificio para pólizas de casa habitación con un nivel de confianza del 90%, llegando a esta conclusión aplicando el proceso antes resumido e ilustrado.

ANEXOS

Anexo 1. Inventario distribuciones discretas más comunes.

DISTRIBUCIONES DISCRETAS				
Distribución	Función de densidad discreta	Media	Varianza	Función generadora de momentos
Uniforme	$f(x) = \frac{1}{N}$ $x = 1, 2, \dots, N$	$\frac{(N+1)}{2}$	$\frac{(N^2 - 1)}{12}$	$\frac{e'(e^n - 1)}{n(e' - 1)}$
Bernoulli	$f(x) = p^x (1-p)^{1-x}$ $0 < p < 1$	p	$p(1-p)$	$(1-p) + pe'$
Binomial	$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ $x = 0, 1, \dots, n$	np	$np(1-p)$	$[pe' + (1-p)]^n$
Geométrica	$f(x) = p(1-p)^{x-1}$ $x = 1, 2, \dots$	$\frac{1}{p}$	$\frac{(1-p)}{p^2}$	$\frac{pe'}{1-(1-p)e'}$
Hipergeométrica	$f(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}$ $x = 0, 1, \dots, n$ si $n \leq r$, $x = 0, 1, \dots, r$ si $n > r$	$\frac{nr}{N}$	$n \left(\frac{r}{N} \right) \left(\frac{N-r}{N} \right) \left(\frac{N-n}{N-1} \right)$	No útil
De Poisson	$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ $x = 0, 1, 2, \dots$ $\lambda > 0$	λ	λ	$e^{\lambda(e'-1)}$
Binomial Negativa	$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$; $x = r, r+1, \dots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left[\frac{pe'}{1-(1-p)e'} \right]^r$

Anexo 2. Inventario distribuciones continuas más comunes.

DISTRIBUCIONES CONTINUAS				
Distribución	Función de distribución acumulativa o función de densidad de probabilidad	Media	Varianza	Función generadora de momentos
Uniforme	$f(x) = \frac{1}{b-a}$ $-\infty < a < b < \infty$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt} - e^{at}}{(b-a)t}$
Normal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-(x-\mu)^2/2\sigma^2\right]$ $-\infty < \mu < \infty, \sigma > 0$	μ	σ^2	$\exp\left[\mu t + \frac{1}{2}\sigma^2 t^2\right]$
Exponencial	$f(x) = \lambda e^{-\lambda x}$ $\lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}$ $t < \lambda$
Gamma	$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}$ $\lambda > 0, r > 0$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - t}\right)^r$ $t < \lambda$
Beta	$f(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}$ $a > 0, b > 0$	$\frac{a}{a+b}$	$\frac{ab}{(a+b+1)(a+b)^2}$	No útil
Cauchy	$f(x) = \frac{1}{\pi\beta\left\{1 + \left[\frac{(x-\alpha)}{\beta}\right]^2\right\}}$ $-\infty < \alpha < \infty, \beta > 0$	No existe	No existe	Función característica $e^{i\alpha t - \beta t }$
Lognormal	$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left[-(\log_e x - \mu)^2/2\sigma^2\right]$ $-\infty < \mu < \infty, \sigma > 0$	$\exp\left[\mu + \frac{1}{2}\sigma^2\right]$	$\exp[2\mu + 2\sigma^2]$ $-\exp[2\mu + \sigma^2]$	No útil
Doble exponencial	$f(x) = \frac{1}{2\beta} \exp\left(-\frac{ x-\alpha }{\beta}\right)$ $-\infty < x < \infty, \beta > 0$	α	$2\beta^2$	$\frac{e^{i\alpha t}}{1 - (\beta t)^2}$
Weibull	$f(x) = abx^{b-1} \exp[-ax^b]$ $a > 0, b > 0$	$a^{-1/b} \Gamma(1+b^{-1})$	$a^{-2/b} \left[\Gamma(1+2b^{-1}) - \Gamma^2(1+b^{-1}) \right]$	$E[X^t] = a^{-t/b} \Gamma\left(1 + \frac{t}{b}\right)$
Logistic	$F(x) = \left[1 + e^{-(x-a)/\beta}\right]^{-1}$ $-\infty < \alpha < \infty, \beta > 0$	α	$\frac{\beta^2 \pi^2}{3}$	$e^{i\alpha t} \pi \beta t \csc(\pi \beta t)$

Pareto	$f(x) = \frac{\theta x_0^\theta}{x^{\theta+1}}$ $x_0 > 0, \theta > 0$	$\frac{\theta x_0}{\theta - 1}$ $\theta > 1$	$\frac{\theta x_0^2}{(\theta - 1)^2 (\theta - 2)}$ $\theta > 2$	No existe
Gumbler	$F(x) = \exp(-e^{-(x-\alpha)/\beta})$ $-\infty < \alpha < \infty, \beta > 0$	$\alpha + \beta\gamma,$ $\gamma \approx 577216$	$\frac{\pi^2 \beta^2}{6}$	$e^{\alpha t} (1 - \beta t)$ $\text{para } t < 1/\beta$
Distribución <i>t</i>	$f(x) = \frac{\Gamma[(k+1)/2]}{\Gamma(k/2)} \frac{1}{\sqrt{k\pi}} \frac{1}{(1+x^2/k)^{(k+1)/2}}$ $k > 0$	$\mu = 0$ $\text{para } k > 1$	$\frac{k}{k-2}$ $\text{para } k > 2$	No existe
Distribución <i>F</i>	$f(x) = \frac{\Gamma[(m+n)/2]}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{2}\right)^{m/2}$ $\times \frac{x^{(m-2)/2}}{[1+(m/n)x]^{\frac{(m+n)/2}}}$ $m, n = 1, 2, \dots$	$\frac{n}{n-2}$ $\text{para } n > 2$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $\text{para } n > 4$	No existe
Chi- cuadrada	$f(x) = \frac{1}{\Gamma(k/2)} \left(\frac{1}{2}\right)^{k/2} x^{k/2-1} e^{-(1/2)x}$ $k = 1, 2, \dots$	k	$2k$	$\left(\frac{1}{1-2t}\right)^{k/2}$ $\text{para } t < 1/2$

Anexo 3. Datos muestrales y Media Residual Life.

Ítem	\$ Siniestro Ocurrido	LN Siniestro Ocurrido	M.R.L.	Ítem	\$ Siniestro Ocurrido	LN Siniestro Ocurrido	M.R.L.	Ítem	\$ Siniestro Ocurrido	LN Siniestro Ocurrido	M.R.L.
1	100	4.6052	20,830	37	6,848	8.8317	23,422	73	22,262	10.0106	25,466
2	100	4.6052	21,028	38	7,738	8.9539	22,858	74	23,692	10.0729	24,763
3	180	5.1930	21,149	39	7,984	8.9851	22,945	75	23,955	10.0839	25,266
4	363	5.8955	21,169	40	8,000	8.9872	23,271	76	24,184	10.0934	25,845
5	403	5.9977	21,338	41	8,115	9.0015	23,507	77	25,000	10.1266	25,863
6	500	6.2146	21,450	42	9,158	9.1224	22,809	78	25,600	10.1504	26,134
7	500	6.2146	21,665	43	9,425	9.1511	22,895	79	27,926	10.2373	24,659
8	500	6.2146	21,884	44	10,000	9.2103	22,674	80	29,955	10.3075	23,468
9	500	6.2146	22,107	45	10,000	9.2103	23,040	81	30,160	10.3143	24,158
10	500	6.2146	22,335	46	10,000	9.2103	23,417	82	30,823	10.3360	24,434
11	983	6.8906	22,080	47	10,224	9.2325	23,580	83	33,252	10.4119	22,923
12	1,000	6.9078	22,295	48	10,718	9.2797	23,477	84	37,701	10.5374	19,277
13	1,318	7.1835	22,211	49	11,640	9.3622	22,944	85	39,032	10.5721	18,762
14	1,500	7.3132	22,265	50	11,767	9.3730	23,217	86	40,107	10.5993	18,528
15	1,611	7.3848	22,395	51	11,959	9.3892	23,436	87	45,000	10.7144	14,317
16	2,000	7.6009	22,248	52	12,000	9.3927	23,821	88	45,104	10.7167	14,961
17	2,000	7.6009	22,495	53	12,686	9.4482	23,564	89	47,917	10.7772	12,823
18	2,006	7.6039	22,742	54	12,758	9.4539	23,935	90	49,423	10.8082	11,983
19	2,309	7.7446	22,694	55	13,191	9.4873	23,954	91	49,740	10.8146	12,396
20	2,466	7.8101	22,796	56	13,518	9.5118	24,090	92	49,898	10.8177	13,053
21	2,500	7.8240	23,027	57	13,544	9.5137	24,545	93	50,000	10.8198	13,876
22	2,798	7.9367	22,996	58	13,563	9.5151	25,026	94	50,000	10.8198	14,943
23	3,000	8.0064	23,065	59	13,574	9.5159	25,537	95	50,000	10.8198	16,189
24	3,284	8.0968	23,056	60	14,841	9.6051	24,786	96	50,097	10.8217	17,555
25	3,793	8.2409	22,822	61	14,975	9.6141	25,188	97	53,122	10.8803	15,983
26	3,981	8.2893	22,913	62	15,200	9.6290	25,518	98	54,494	10.9058	16,234
27	4,065	8.3101	23,115	63	15,875	9.6725	25,407	99	55,591	10.9258	17,030
28	4,239	8.3520	23,232	64	16,629	9.7189	25,227	100	56,317	10.9388	18,632
29	4,500	8.4118	23,265	65	17,392	9.7638	25,046	101	59,713	10.9973	17,776
30	4,577	8.4288	23,489	66	18,571	9.8294	24,449	102	59,905	11.0005	21,101
31	5,000	8.5172	23,369	67	18,619	9.8319	25,011	103	69,111	11.1435	14,868
32	5,000	8.5172	23,681	68	19,554	9.8810	24,693	104	74,958	11.2247	12,028
33	5,574	8.6259	23,419	69	19,734	9.8901	25,159	105	83,673	11.3347	4,970
34	6,238	8.7384	23,067	70	19,793	9.8931	25,778	106	86,774	11.3711	3,737
35	6,683	8.8073	22,936	71	20,000	9.9035	26,281	107	90,511	11.4132	0
36	6,842	8.8309	23,098	72	21,133	9.9586	25,867				

BIBLIOGRAFÍA

1. Berenson, M., Levine, D., and Rindsopf, D., (1988), "Applied Statistics: A First Course", United States of America: Prentice Hall.
2. Bowers, N., Gerber, H., Hickman, J., Jones, D., and Nesbitt, C. (1986), "Actuarial Mathematics", Schaumburg, IL: Society of Actuaries.
3. Canavos, C. (1994), "Probabilidad y Estadística. Aplicaciones y Métodos". México: McGraw-Hill.
4. Casualty Actuarial Society (2001), "Foundations of Casualty Actuarial Science", Arlington: Casualty Actuarial Society.
5. Chao, L. (1985), "Introducción a la Estadística", México: Compañía Editorial Continental.
6. Corina Ortega, G. (1974), "Los Modelos Matemáticos como Auxiliares en la Toma de Decisiones", México: Tesis de Licenciatura, Universidad Iberoamericana.
7. Freund, J., Miller, I., y Miller, M. (2000), "Estadística Matemática con Aplicaciones", 6ª ed., México: Pearson Educación
8. Giordano, F. (1997) "A First Course in Mathematical Modeling", Pacific Grove, California: Books/Cole Publishing Company.
9. Hogg R. and Klugman, S (1984), "Loss Distributions", United States of America: John Wiley & Sons, Inc.
10. Huntsberger, D. y Billingsley, P. (1983), "Elementos de Estadística Inferencial", México: Compañía Editorial Continental.
11. Jonson, R. y Kuby, P. (1999), "Estadística Elemental. Lo Esencial", 2da ed., México: International Thomson Editores .
12. Klugman, A., Panjer, H., and Willmont, G. (1998), "Loss Models: From Data to Decisions", United States of America: John Wiley & Sons, Inc.
13. Maki, D. (1983), "Mathematical Models and Applications. With Emphasis on the Social, Life and Management Sciences", United States of America: Prentice-Hall.
14. McGrath, R., (1997), "Understanding Statistics: A Reserch Perspective", Longman: Addison Wesley Publishers. Inc.
15. Mendenhall, W, Wackerly, D., y Scheaffer, R. (1994) "Estadística Matemática con Aplicaciones", 2a ed., México: Grupo Editorial Iberoamericano.
16. Mood, A., Graybill F., and Boes, D. (1974), "Introduction to the Theory of Statistics", 3er ed., United States of America: McGraw-Hills, Inc.
17. Parzen, E. (1987), "Teoría Moderna de Probabilidades y sus Aplicaciones", México: Editorial Limusa.
18. Saaty, T. (1981), "Thinking whit Models: Mathematical Models in the Physical, Biological and Social Sciences", United States of America: Pergamon Press.
19. Thierauf, R. (1972), "Toma de Decisiones por Medio de Investigación de Operaciones", México: Limusa-Wiley.