UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO

Unidad Académica Profesional Cuautitlán Izcalli

TESIS

“Modelo de cálculo del límite máximo de retención óptimo para las obligaciones que asume una compañía de seguros en los ramos de daños sin autos”

PARA OBTENER EL TÍTULO DE:

LICENCIADO EN ACTUARÍA

PRESENTA:

ERICK GERARDO CAMARENA SOSA

ASESOR:

MTRO. EN FINANZAS ROBERT HERNÁNDEZ MARTÍNEZ

Cuautitlán Izcalli, Edo. de México, junio de 2016.

UAEM | Universidad Autónoma del Estado de México

Cuautitlán Izcalli, Estado de México, a 04 de octubre de 2016
Oficio No. UAPCI/DA/DEP/076/2016

DG CUyUAP
Dirección General de Centros Universitarios y Unidades Académicas Profesionales

DR. ROLANDO HEREDIA DOMINICO
SUBDIRECTOR ACADÉMICO DE LA UAP
CUAUTITLÁN IZCALLI
P R E S E N T E

ASUNTO: VOTO APROBATORIO ASESOR Y REVISORES DE TRABAJO DE INVESTIGACIÓN, Y AUTORIZACIÓN PARA IMPRESIÓN.

Los que suscriben, por este medio manifestamos que el trabajo de investigación en la modalidad de **TESIS** titulada "**MODELO DE CÁLCULO DEL LÍMITE MÁXIMO DE RETENCIÓN ÓPTIMO PARA LAS OBLIGACIONES QUE ASUME UNA COMPAÑÍA DE SEGUROS EN EL RAMO DE "DAÑOS SIN AUTOS"**. Del alumno **ERICK GERARDO CAMARENA SOSA**, con número de cuenta **1028847**, de la Licenciatura en **Actuaría**; cumple con los requisitos y cualidades que corresponden a esta opción de evaluación profesional.

Por lo anterior, **OTORGAMOS** nuestro **VOTO APROBATORIO** en términos del Reglamento de Evaluación Profesional de la Universidad Autónoma del Estado de México; asimismo, manifestamos que estamos de acuerdo en la impresión del mismo.

Sin otro particular, nos reiteramos a sus órdenes.

ATENTAMENTE
PATRIA, CIENCIA Y TRABAJO
"2016, Año del 60 Aniversario de la Universidad Autónoma del Estado de México"
"2016, Año de Leopoldo Flores Valdés"

ASESOR

MTRO. ROBERT HERNÁNDEZ MARTÍNEZ

REVISOR

ACT. ESPARTACO MARTÍNEZ TOLENTINO

REVISOR

ACT. LETICIA JIMENÉZ NORIEGA

ESTADOS UNIDOS MEXICANOS · UNIVERSIDAD AUTÓNOMA · PATRIA · CIENCIA · Y TRABAJO · ESTADO DE MÉXICO
U.A.P. CUAUTITLÁN IZCALLI
SUBDIRECCIÓN ACADÉMICA

UAP CUAUTITLÁN IZCALLI

2013-2017

C.C.P. MTRO. EN FIN. ROBERT HERNÁNDEZ MARTÍNEZ.- JEFE DE LA UNIDAD DE EVALUACIÓN PROFESIONAL
ACT. ESPARTACO MARTÍNEZ TOLENTINO.-PROFESOR DE ASIGNATURA DE ACTUARÍA DE LA UAP C. IZCALLI
ACT. LETICIA JIMENÉZ NORIEGA.- PROFESORA DE ASIGNATURA DE ACTUARÍA DE LA UAP C. IZCALLI
ERICK GERARDO CAMARENA SOSA.- ALUMNO INTERESADO.
ARCHIVO RHD/RHM/mdbr*

www.uaemex.mx
Av. Prolongación Islas s/n, Col. Atlanta 2ª. Sección, Cuautitlán Izcalli, Estado de México. C.P. 54740
Tel. (0155) 11 13 40 60 y 11 13 40 62 Ext. 103 y 104 Conmutador 7211
E-mail: duapci@gmail.com @UAP Cuautitlan Iz. UAP Cuautitlán Izcalli http://www.uaemex.mx/uapci/

AGRADECIMIETOS

A la Universidad Autónoma del Estado de México a través de la Unidad Académica Profesional Cuautitlán Izcalli, por las facilidades brindadas para la realización de esta meta, y por abrirme las puertas de lo que ahora es mi segundo hogar.

A la Coordinación de Actuaría, por enseñarme que el camino de la Actuaría nunca es sencillo, pero con constancia y dedicación todo es posible; y que más que una profesión, es un estilo de vida.

Al Maestro en Finanzas Robert Hernández Martínez, por confiar en mí para la elaboración del presente trabajo de investigación, y por enseñarme que la vida académica conlleva muchas responsabilidades.

Al Doctor Rolando Heredia Dominico, porque gracias a su ejemplo, comentarios y enseñanzas aprendí que la ciencia demanda compromiso, y que con perseverancia y entrega se cumplen las metas que se proponen.

A la Doctora Eva Marta Chaparro por apoyarme de manera incondicional en todo momento; y atender mis problemas personales y académicos para buscar posibles soluciones durante mi estancia en la universidad.

Al Maestro Rodrigo Díaz Infante Pesquera por sus valiosos comentarios al presente trabajo, y por mostrarse siempre accesibles y dispuestos a colaborar.

Al Actuario Espartaco Martínez Tolentino, por aceptar ser revisor de este proyecto de investigación, y tomarse el tiempo necesario para examinar todos los elementos técnicos y formales del mismo.

A mis compañeros y amigos de la Empresa Mapfre seguros por acceder a involucrarme en este proyecto, brindarme su conocimiento obtenido con la experiencia y permitir desarrollarlo con toda libertad.

Agradezco, especialmente, a mis compañeros y amigos que estuvieron presentes a lo largo de mi formación como Actuario; en la Unidad Académica Cuautitlán Izcalli: a Cristian, Edgar, Andrew, Carlos, Denisse, Irving, Marlon, Karen y a todo mi grupo por compartir tantos momentos y experiencias.

A mis profesores de cada asignatura para formar un profesionista integral, de igual manera a todo el personal que labora en la Unidad Académica Profesional Cuautitlán Izcalli por brindarme su apoyo y amistad.

A todos aquellos que en algún momento pusieron un pie en mi camino para que tropezara, porque gracias a ellos aprendí a levantarme y saltar obstáculos cada vez mayores.

MUCHAS GRACIAS

DEDICATORIAS

A mi viejita, María Elena Galindo, por siempre confiar en mí y enseñarme que podría conseguir todo lo que quisiera si me lo proponía y que a pesar de que ahora ya no este conmigo sigue mostrándome el camino del triunfo.

A mis padres, Claudia y Erick, porque su amor incondicional e infinito siempre me han acompañado, aún en los momentos más difíciles.

A mi querida hermana, Yatziri, por ser mi razón para luchar en la vida y formar un modelo a seguir del cual se enorgullezca en unos años.

A mi tía, Minelia, quien se convirtió en mi mejor amiga y que ha estado conmigo en los momentos que más lo he necesitado, apoyándome de manera incondicional.

A mi abuela Amelia, porque a pesar de que estamos peleando todo el tiempo siempre desea lo mejor para mí.

A mis primos con los cuales compartí grandes historias e hicieron de mi infancia una etapa increíble de mi vida.

A mis tíos porque siempre vieron por mi bienestar

A mi tío, Juan Carlos Camarena, por permitirme ser parte de su hermosa familia y que donde está sigue guiando el camino de toda mi familia

Al equipo de Thematika Eventos con los cuales maduré de manera personal y formé grandes amistades, con una mención especial a Claudia y Benjamín que siempre me brindaron su confianza y amistad.

A mis amigos de candelaria que formamos un vínculo más allá de la amistad y que siempre han correspondido al cariño que les tengo.

A mis amigos del CEB que, a pesar de vernos poco por las actividades de cada uno, siempre se hacen presentes en los momentos importantes.

Resumen ejecutivo

Debido a que la normatividad de la actividad aseguradora en México establece la obligación de las compañías de seguros de realizar el cálculo del límite máximo de retención, establecido en la Ley de Instituciones de Seguros y Fianzas y la Circular Única de Seguros y Fianzas; en esta investigación se propone un modelo actuarial para estimar dicho valor límite de retención de las obligaciones contingentes adquiridas por una Compañía de seguros; aplicado a los ramos de Daños sin autos, de tal suerte que garantice la solvencia financiera de la aseguradora.

Abstract

Due to the regulation of insurance business in Mexico establishes the obligation of insurance companies to perform the calculation of the maximum retention limit established in the *Ley de Instituciones de Seguros y Fianzas* and the *Circular Única de Seguros y Fianzas*; in this research an actuarial model is proposed to estimate the limit value retention of contingent liabilities acquired by an insurance company; applied to property lines without cars, in such a way to ensure the financial solvency of the insurer.

Keywords:

Riesgo, solvencia, límite máximo de retención, seguro de daños, simulaciones, modelo, prueba de bondad de ajuste, siniestros, severidad, frecuencia, probabilidad de perdida

Clasificación JEL (Journal Economic Literature):G22

ÍNDICE GENERAL

INTRODUCCIÓN

Debido a que la normatividad establece que es obligación de las compañías de seguros realizar el cálculo del ***límite máximo de retención,*** establecido en el artículo 235 de la Ley de Instituciones de Seguros y Fianzas y en el capítulo 9 de la Circular Única de Seguros y Fianzas. Se pretende demostrar que dicho estudio, realizado comúnmente por las consultorías externas, no se comporta únicamente como un proceso estocástico. Para esto, fundamentándonos en que la regulación vigente en la materia, establece que el límite máximo de retención se deberá calcular bajo las reglas determinadas por la SHCP (Secretaría de Hacienda y Crédito Público); y será regulado por la CNSF (Comisión Nacional de Seguros y Fianzas) para evitar que las obligaciones adquiridas por una compañía aseguradora sean mayores al capital reservado para los posibles siniestros cubiertos en los contratos, sin necesidad de recurrir a los esquemas de reaseguro.

Es necesario tener claros todos los conceptos actuariales necesarios para el entendimiento del estudio. Las compañías aseguradoras operan en varias áreas y a su vez varios ramos, las clasificaciones y sus entidades reguladoras definen varios conceptos formales de cada una de sus áreas de trabajo.

La primera parte del documento se dedica a explicar de manera detallada algunos de los conceptos indispensables del sector asegurador, se comienza con definiciones básicas que se emplean de manera ordinaria en todos los seguros hasta conceptos más complejos que sólo se tratan en sitios exclusivos de las compañías aseguradoras.

Ya dominado el lenguaje formal que se usa en el sector asegurador se deberá emplear un lenguaje técnico actuarial el cual aplica métodos estadísticos y probabilísticos que se desarrollan durante todo el estudio.

Posteriormente, se explica de manera teórica y práctica cómo elaborar un modelo de simulación, cómo ampliar el razonamiento lógico necesario y el modo de clasificar los componentes del mismo.

El diseño y la metodología de modelo de límite máximo de retención se desglosan paso a paso durante el documento de tal manera que se identifique cada uno de los elementos del método a emplear y lograr interpretar de manera precisa, los resultados que se obtengan con el modelo propuesto.

Para terminar, se pretende obtener una cifra concreta, a una fecha determinada; del límite máximo de retención de las obligaciones contingentes adquiridas para una Compañía de seguros en los ramos de Daños sin autos, que represente un valor óptimo para garantizar la solvencia financiera de dicha aseguradora y comprobar la eficiencia de nuestro modelo.

Planteamiento del problema

Existe una ausencia de metodología y de modelos que estimen eficientemente las obligaciones contingentes adquiridas por las compañías de seguros que les permita cumplir a cabalidad la legislación de seguros y la normatividad emitida para la Comisión Nacional de Seguros y Fianzas, en materia de criterios prudenciales para la administración de riesgos de dichas aseguradoras.

Justificación

Que la compañía cuente con una metodología que le garantice asumir de manera óptima su ***límite máximo de retención*** permitiéndole generar un crecimiento financiero sustentable, y así contar con la solidez que la compañía de seguros requiere; garantizando la solvencia y evitando mantener capital ocioso, el cual se podría invertir en instrumentos que generen rendimientos, incrementando los productos que la compañía puede ofrecer y expandir el sector al que se dirige; ganando confianza y reconocimiento en el mercado; por lo tanto, el profesionista en Actuaría es fundamental para el diseño óptimo de dicha metodología y su implementación a través de un modelo actuarial.

Antecedentes

Una de las áreas de oportunidad en las que puede participar el Actuario en una Compañía de Seguros es el área técnica, en la cual se analiza; entre otras cosas, las condiciones financieras en que se encuentra la institución, con el principal objetivo de que la compañía tenga el capital suficiente para hacer frente a sus obligaciones futuras. En este sentido, se determina estratégicamente asumir a retención sólo una parte de los riesgos de su cartera, garantizando así, la solvencia y estabilidad financiera de la misma, previniendo que en el futuro esta condición pueda convertirse en un tema de supervivencia de la compañía en el sector.

Para aportar elementos que permitan garantizar dicha estabilidad financiera, se realiza un análisis denominado ***límite máximo de retención,*** en el cual mediante modelos matemáticos y aplicación de herramientas estadísticas, se analiza el comportamiento de los siniestros. Con el objetivo de controlar que las obligaciones futuras que acepte la compañía no sean mayores al capital asignado al ramo en cuestión.

Hipótesis

La metodología de cálculo de las obligaciones que puede asumir una compañía de seguros puede ser estimada de manera óptima a través de un modelo de simulación que se ajuste a la cartera en vigor y a la normatividad aplicable; obteniendo como resultado el límite máximo de retención en los ramos Daños sin autos de una compañía de seguros.

Objetivos

Aportar una metodología novedosa que cumpla con los requerimientos prudenciales de la Comisión Nacional de Seguros y Fianzas, mediante el diseño de un modelo de cálculo óptimo del límite máximo de retención en el ramo “Daños sin autos” de una Compañía de seguros.

Identificar cómo se distribuye la siniestralidad histórica de la compañía, midiendo la frecuencia y severidad de los siniestros de cada uno de los ramos de “Daños sin autos”.

Diseñar un modelo de cálculo mediante la metodología aplicable, del límite máximo de retención por ramo.

Emitir conclusiones que permitan la toma de decisiones asertivamente, a partir de los resultados derivados del modelo para la estimación del límite máximo de retención por ramo.

Capítulo I. Terminología de seguros

1.1 CONOCIMIENTOS GENERALES

La actividad aseguradora, como cualquier otra que supone una especialidad, tiene su propia forma de expresarse, su lenguaje particular y su fisonomía técnica y doctrinaria peculiar. Por lo que es indispensable para la elaboración de este modelo conocer la terminología que se usa en el campo de los seguros, con el objeto de entender algunos vocablos que se utilizarán en el desarrollo de este documento.

1.1.1 Antecedentes del seguro

El Código de Hammurabi fue el primer registro en la historia que constituyó formas elementales de seguro mutual hace más de tres mil setecientos años, con el riesgos que implicaba del tráfico comercial por tierra o agua; y la necesidad de estimular el comercio y la inversión se comenzaron a tomar las primeras medidas para minimizar dicho riesgo, por ejemplo: si un viajero babilonio era asaltado, el gobierno local estaba obligado a devolverle lo que hubiera perdido (en la actualidad se podría asemejar a un seguro de viaje); si un granjero tenía una deuda, pero un desastre natural le afectaba, se le perdonaba el pago correspondiente a esa temporada (seguro agropecuario sería su similar); y si las mercancías de un comerciante no llegaban sanas y salvas, se les perdonaban las deudas en que hubiere incurrido para financiar su caravana (cumple como un seguro de carga).”Reglamentaba un aspecto de la seguridad social, si un hombre adoptaba a un hijo, éste estaba obligado a mantener al padre en su vejez. Si no lo hacía, era castigado con la muerte” (Código Hammurabi).

“Mil años después, los mercaderes de Rodas, verdadera potencia naval de entonces, que comerciaban en el mediterráneo gracias a su flota marítima, crearon un sistema de protección mutual” (Osorio, 2003). En esa época se pedían préstamos para financiar sus viajes, con la garantía de sus barcos y su cargamento, y si el viaje tenía éxito, se pagaban los prestamos más los intereses fijados de antemano; si se perdía el barco o el cargamento en el mar, se le perdonaba el préstamo y los intereses al perjudicado por el siniestro, lógico que este interés era bastante elevado para compensar el riesgo de pérdida, en esencia, tenían una cobertura primitiva que podría asimilarse a un seguro de casco y mercaderías. Los rodios ricos, actuaban como banqueros y como aseguradores, cuando financiaban estos viajes, en la medida en que distribuían sus riesgos entre gran número de barcos (o cascos), tenían probabilidades de salir adelante, aunque sufrieran pérdidas ocasionales.

“La ley de Rodas establecía el principio de las contribuciones a un fondo común, si para aligerar un navío se lanza por la borda alguna mercancía, lo que uno ha perdido le será devuelto mediante la contribución de todos” (Osorio, 2003), un principio de la avería gruesa.

De igual manera los romanos tenían sociedades funerarias organizadas para solventar los gastos funerarios comportándose como un seguro de vida. “Los soldados de las legiones contribuían a un fondo (collegia militum) que ayudaban a sufragar gastos de mudanzas de los militares trasladados de guarnición (collegia funeratitia) que ayudaba para gastos de sepelio y con una suma a la viuda y huérfanos del fallecido.” (Barrera, 1957). Este fondo era constituido de dos maneras, una mediante contribución fija y la otra se repartía el gasto entre todos sus miembros.

En Roma la transferencia de riesgo era era un acto común, como cláusula accesoria a un contrato, así, por ejemplo, el artífice que engarzaba mal una piedra preciosa era responsable por su pérdida. Sin lugar a dudas, en Roma existían muchas asociaciones de asistencia mutua, pero no conocieron la institución jurídica del seguro. El estado romano supo asumir riesgos marítimos, en principio como cosa excepcional, luego como norma, en resguardo de sus intereses.

Posteriormente, en la edad media existe un cambio en el sistema económico. “Un nuevo sistema implicaba un grado más avanzado de especialización, lo que da origen al establecimiento de relaciones comerciales entre el campo y la ciudad, derivadas de la división del trabajo, que se intensifica en el ámbito internacional.” (Osorio, 2003)

Una institución precursora del seguro lo constituyeron las guildas medievales la que desarrollaron mecanismos de asistencia, en la que las aportaciones (primas) no estaban con relación a las prestaciones (ANTECEDENTES HISTÓRICOS DEL SEGURO).

“El seguro marítimo apareció hasta el siglo XIV con el desarrollo del comercio marítimo, fue la primera vez que abonaba una prima de manera formal, el origen de los primero documentos son italianos” (Sánchez, 2000), el primer contrato de seguro conocido hace referencia a un seguro marítimo establecido en Génova, y parece datar del año 1317. El seguro marítimo es la modalidad más antigua de la actividad aseguradora. Se conocen prácticas del puerto de Cagliari en 1318, un decreto del Dogo de Génova en el año 1336, los libros de comercio de Francesco del Bene y Cía. de 1318 a 1350 en Florencia, si se discuten algunos de estos antecedentes históricos, existe unanimidad en aceptar uno, datado el 23 de octubre de 1347, y se reconoce que en la segunda mitad de ese siglo estaba ampliamente difundido en Pisa, Florencia y Génova, que son las primeras ciudades en darles normas legislativas. Que prohibían la usura, asegurar a extranjeros y asegurar el objeto del seguro por el monto total de su valuación

Por último, la ciudad de Amberes en el siglo XVI toma el liderazgo asegurador que las ciudades españolas habían tenido hasta ese momento y allí se producen varias leyes entre 1563 y 1570, esta última fue publicada por el Duque de Alba y en ella se regula por primera vez la supervisión del estado sobre el negocio asegurador. El clásico Guidon de la Mer, compilación anónima sirvió de base para que Luis XIV promulgara la célebre Orden nance de la Marina en 1681, que abrió rumbos a la legislación marítima moderna. La primera ley inglesa de seguros fue promulgada en el año 1601 “El punto de partida de la moderna industria aseguradora tuvo su inicio el 2 de septiembre de 1666, a consecuencia del gran incendio de Londres, con la creación de la Fire Office apareció el seguro contra incendio o seguro de incendio.” (Barrera, 1957), ya que es el primer seguro destinado a cubrir riesgos terrestres.

1.1.2 Definición del seguro

“El seguro es un sistema de protección del hombre y de su patrimonio frente a diversos hechos que amenazan su integridad, su vida, su interés y su propiedad” (Palacios, 1996) esta definición es para cuando se considera al seguro como institución, Los hechos nocivos que causan pérdidas o daños son inciertos pero previsibles. El seguro garantiza el resarcimiento de un capital para reparar o cubrir la pérdida o daño que aparezca en cualquier momento, recibiendo como contraprestación un precio por adelantado por el servicio de protección que ofrece.

Si se considera al seguro como contrato; el seguro es el convenio entre dos partes, la compañía o entidad aseguradora y el asegurado o contratante, mediante el cual la primera se compromete a cubrir económicamente la pérdida o daño que el segundo puede sufrir durante la vigencia del contrato. La obligación del asegurado o contratante es pagar, al firmar el contrato, el precio del seguro, parcial o totalmente.

1.1.3 Clasificación del seguro

El seguro se puede clasificar desde diversos puntos de vista (Osorio, 2003):

A. Según la finalidad del servicio: Se puede distinguir por un lado el Seguro Social, cuya finalidad es proteger a los llamados sectores populares (principalmente los trabajadores) contra riesgos socialmente extendidos, como son la vejez, la salud, el desempleo, etc.; y el Seguro Privado o Comercial habitualmente explotados por empresas comerciales.

B. Según su origen histórico: Desde este punto de vista se ha dividido el seguro en marítimo, terrestre y aéreo.

C. Según el objeto asegurado: Desde el punto de vista del objeto al que se dirige la protección del seguro, este se clasifica en seguros de daños (que a su vez se divide en seguros de bienes, incluyendo incendio, robo, automotores, etc., y de responsabilidades), y el seguro de personas (incluyendo los de vida, accidentes y salud).

D. Según la Ley de Instituciones de Seguros y Fianzas (LISF) que se aplica en México, clasifica a los seguros en distintos ramos que pertenecen a las siguientes operaciones:

1. Vida: están orientados a proteger la economía familiar y a garantizar su bienestar a lo largo de los años. Los seguros de vida es uno de los tipos del seguro en el que el pago por parte de la compañía de seguros de la suma asegurada del contrato depende del fallecimiento o sobrevivencia del asegurado en un momento determinado. En este tipo de seguro el pago de la indemnización no guarda relación con el valor del daño producido por la concurrencia del siniestro, debido a que la persona no es valuable económicamente. De ahí que este tipo de seguro no constituya un contrato de indemnización propiamente dicho, diferenciándose así, de los seguros de daños.
2. Accidentes y enfermedades: Son los que cubren la lesión o incapacidad que afecte la integridad personal o la salud del Asegurado, ocasionada por un accidente o enfermedad. En esta clase de seguros la Aseguradora, mediante el pago de la prima correspondiente, cubre los gastos hospitalarios, atención médica, intervenciones quirúrgicas, alimentos, medicamentos, análisis clínicos, rayos x, etc. a los asegurados y en su caso a los dependientes económicos cuando así quede convenido en la póliza del seguro.
3. Daños: Bajo esta denominación se recogen todos los seguros cuyo fin principal es reparar la pérdida sufrida, a causa de siniestro, en el patrimonio del tomador del seguro. Los seguros de daños pueden dividirse en dos grandes grupos: los seguros de cosas destinados a resarcir al asegurado de las pérdidas materiales directamente sufridas en un bien integrante de su patrimonio y los seguros de responsabilidad que garantizan al asegurado contra la responsabilidad civil en que pueda

incurrir ante terceros por actos en los que sea responsable.

Esta última clasificación será la adecuada a lo largo del documento, debido que durante el estudio se hará referencia a la LISF.

1.2 EL RIESGO

La incertidumbre se ha considerado como uno de los principales elementos de la vida, ninguna persona o empresa tiene un conocimiento cierto y total de lo que pueda pasar en el futuro y la historia proporciona muchos ejemplos de cómo la humanidad ha tratado de eliminar o, al menos, reducir la incertidumbre. En la misma medida en que se ha producido el desarrollo económico, se ha hecho más complejo también el grado de incertidumbre con que se enfrentan las entidades y ha aumentado en consecuencia de demanda de su cobertura.

1.2.1 Concepto y características del riesgo

"Es la posibilidad de pérdida o daño, es decir, la constante amenaza que pesa sobre el hombre y su patrimonio." (Osorio, 2003) Desde la cuna hasta la tumba El hombre está en constante riesgo por enfermedad, accidente y muerte prematura. Los bienes igualmente pueden sufrir incendios, robo, merma, deterioro, en fin, toda suerte de riesgos o eventos dañinos. El seguro tiene como fin proteger al hombre contra estos riesgos indemnizando su pérdida con una suma equitativa previamente convenida.

En la terminología aseguradora, el concepto de riesgo se emplea para interpretar dos ideas diferentes: en la primera se define al riesgo como objeto asegurado y en la segunda idea se considera al riesgo como posible ocurrencia por azar de un acontecimiento que produce una necesidad económica y cuya aparición real o existencia se previene y garantiza en la póliza y obliga al asegurador a efectuar la prestación o indemnización, que le corresponde. Esta posibilidad de pérdida o daño, que pesa sobre el ser humano y su patrimonio, está denominado riesgo y tiene una extensión para cualquier individuo.

Las características esenciales del riesgo, para ser objeto del seguro, son los siguientes (Palacios, 1996):

a) Incierto y aleatorio: Debe haber una relativa incertidumbre, pues el conocimiento de su existencia real haría desaparecer la aleatoriedad, principio básico del seguro. En algunos casos, se conoce con certeza que ocurrirá, pero se ignora cuándo, como en la cobertura de los seguros de vida. También ocurre que dicha incertidumbre se apoya en el dilema si ha ocurrido o no, como a veces sucede en los seguros de transporte, en que es técnicamente posible la suscripción de una póliza que asegure el riesgo de hundimiento de un buque desaparecido, desconociendo ambas partes contratantes si en el momento de la suscripción de la póliza, el barco ha naufragado o no.
b) Posible: Ha de existir posibilidad de riesgo, tal posibilidad tiene dos limitaciones extremas, de un lado, la frecuencia y del otro la imposibilidad. La excesiva reiteración del riesgo y su materialización en siniestro atenta contra la aleatoriedad del suceso cubierto, así, una gran frecuencia en la sección automóviles, aparte de resultar antieconómico para la empresa, exigirá que las primas de seguros necesariamente tendrán un incremento, del mismo modo la imposibilidad de que el riesgo se manifieste en siniestro, situaría a la empresa aseguradora en una posición de presentar una cobertura absurda que haría inviable la comercialización del producto, como ofrecer una cobertura de daños por nevadas, en un país tropical.
c) Concreto: El riesgo debe ser analizado y valorado en sus dos aspectos cualitativo y cuantitativo, antes de proceder a asumirlo. Solo de esta forma la entidad podrá decidir sobre la factibilidad de la cobertura y luego fijar la prima adecuada para la misma.
d) Lícito: El riesgo que se asegure no debe estar en contra de las reglas morales, del orden público ni en perjuicio de terceros, pues conforme a la legislación, la póliza será nula automáticamente. Esta regla tiene dos excepciones, en los seguros de vida, la cobertura del riesgo de suicidio, pasados los tres años de cobertura; y los seguros de responsabilidad civil, en

donde puede garantizarse los daños causados a terceros cometidos por imprudencia.

e) Fortuito: El riesgo debe provenir de un acto o acontecimiento ajeno a la voluntad humana de producirlo. No obstante, es indemnizable el siniestro producido a consecuencia de actos realizados por un tercero, ajeno al vínculo contractual que une a la aseguradora con el asegurado, aunque en tal caso la entidad aseguradora se reserva el derecho de ejercitar las acciones pertinentes contra el responsable del daño(principio de subrogación), como también es indemnizable el siniestro causado intencionadamente por el asegurado, siempre que los daños se hayan producido con ocasión de fuerza mayor o para evitar otros graves
f) Contenido económico: La realización del riesgo ha de producir una necesidad tasable en valores económicos, que se satisface con la indemnización correspondiente.

De dicha clasificación, existe una división de los riesgos para el manejo de la actividad aseguradora y es de la siguiente manera:

a) Riesgos puros: son aquellos que al materializarse siempre originan pérdidas y con la intervención del seguro se recupera al estado anterior: riesgos de incendio, inundación, accidentes, etc., son típicamente asegurables.
b) Riesgos especulativos: son aquellos que al materializarse pueden originar indistintamente beneficio o pérdida: inversiones en divisas ante una posible devaluación o revaluación, aventura comercial, juego de azar, etc., normalmente no son asegurables.
c) Riesgos dinámicos: son los relacionados con las incertidumbres producidas por una sociedad en cambio permanente: Condiciones ambientales, necesidades del consumidor, nuevas tecnologías, etc.
d) Riesgos estáticos: son los riesgos puros, que no se ven influidos por el factor humano: terremoto, huracanes, rayo, etc.
e) Riesgo objetivo: es un concepto técnico - estadístico que se refiere a la variación probable entre una pérdida real y pérdida probable.
f) Riesgo subjetivo: es una incertidumbre psicológica que proviene de la actitud o estado del individuo. Una persona puede tener un riesgo objetivo

importante, pero no tener riesgo subjetivo porque ignora el peligro o lo subestima.

g) Riesgos asegurables: se clasifican en leves, graves y catastróficos.

1.2.2 Administración de riesgo

La administración de riesgos es la disciplina que combina los recursos financieros, humanos, materiales y técnicos de una empresa, para identificar o evaluar los riesgos potenciales y decidir cómo manejarlos con la combinación optima de costo-efectividad.

La administración de riesgos en un marco amplio implica que las estrategias, procesos, personas, tecnología y conocimiento están alineados para manejar toda la incertidumbre que una organización enfrenta.

Por otro lado, los riegos y oportunidades van siempre a la mano, y la clave es determinar los beneficios potenciales de estas sobre los riesgos.

También, cabe mencionar, que es una función empresarial cuyo objetivo es la conservación de los activos y del poder de generación de beneficios mediante la minimización a largo plazo del efecto financiero de las pérdidas accidentales.

La administración de riesgos busca cumplir con ciertos objetivos que son.

1. Acontecimientos de previsión:
    - Identificación de los recursos materiales, humanos y financieros de las empresas.
    - Identificación de los riesgos que están expuestos los recursos de la empresa.
    - Evaluación del posible impacto financiero de un accidente atreves de su medición adecuada.
    - Jerarquización de los riesgos identificados y evaluados
    - Elaboración de programas de prevención.

En las compañías aseguradoras el riesgo más latente que se pretende evitar es el de las reservas insuficientes.

2. Evaluación de los eventos corrientes:
    - Revisión de contratos, reingeniería; es decir, nueva versión a para la mejora.
    - Elaboración de manuales de seguridad de higiene.
    - Elaboración de programas de capacitación en el manejo de equipo de seguridad.
    - Elaboración de planes de emergencia y evaluación.
    - Realización de simulacros.
3. Procedimiento para la solución de accidentes consecuenciales:
    - Supervisión de empresa
    - Cuidado de la planta productiva
    - Supervivencia de la empresa
    - Financiamiento para la normalización de las operaciones
    - Recuperación de seguros, fianzas y otros contratos.
    - Evaluación de los planes de emergencia para su validación y mejoramiento.
    - Conservación de la planta productiva.

1.2.3 Funcionamientos del seguro como instrumento para mitigar el riesgo

Conviene indicar que lo adquirido por los asegurados es una promesa plasmada en un contrato de adhesión, es decir, un conjunto de condiciones hechas solamente por una de las partes, al que la otra presta su conformidad. En muchos casos la parte que se adhiere (el Asegurado) no conoce muy bien el contrato, sus consecuencias ni la interpretación de sus cláusulas, por lo que podría suscribir, sin saberlo, un compromiso contrario a sus intereses o, simplemente no adecuado a sus necesidades.

Hay, pues, una necesidad real de que el contenido del contrato sea conocido y aprobado por quienes tienen a su cargo el funcionamiento de las instituciones económicas, que de ese modo protegen al público y lo representan ante los profesionales aseguradores.

A los objetivos ya mencionados de proteger los intereses de los asegurados, beneficiarios o terceros, debe unirse el de la regulación y supervisión del seguro desde el punto de vista de los intereses económicos y sociales de carácter general, que puede contemplarse desde varias perspectivas:

1. Coordinación de las inversiones de las empresas de seguros con la política general de la administración del país.

2. Necesidad de impedir la salida de divisas originada por la utilización excesiva de los servicios de entidades extranjeras de seguros y reaseguros.

3. Adopción y ejecución de medidas para establecer y reforzar el mercado nacional de seguros, que es reconocidamente requisito esencial para el logro del desarrollo económico nacional. Dentro de las funciones de la Autoridad de Control son:

- Exigir el cumplimiento de las leyes vigentes en cada momento por parte de quienes intervienen en el mercado asegurador; entidades, agentes y corredores de seguros, reaseguradores, corredores de reaseguros y liquidadores de siniestros; en concordancia a la política social de mercado implementada en el país.
- Conocer constantemente la situación económica-financiera de las entidades que actúan en el mercado.
- Detectar y corregir las situaciones irregulares y las anomalías que comprometan la solvencia de las entidades para satisfacer el pago de siniestros e indemnizaciones al asegurado.
- Procurar que la actuación técnica de las entidades se base en principios actuariales, estadísticos y cálculos correctos.
- Elaborar informes y estadísticas que reflejen la situación real y constituyan una guía para cuantos trabajan en él.
- El papel de la Superintendencia de Seguros es dejar que las empresas de seguros decidan todo lo referente al giro comercial del negocio del seguro,

pero salvaguardando la solvencia económica financiera de la misma, precautelando de esa manera los intereses de los asegurados.

Para cumplir dichos objetivos, sus actividades se orientan en las siguientes líneas:

I. Análisis de las bases técnicas del funcionamiento de los aseguradores: tarifas suficientes, modelos de pólizas, tipo de contratos, etc.
II. Análisis de los balances, cuentas de resultados y coberturas de las reservas de las entidades aseguradoras.
III. Inspecciones, el organismo de control tiene la facultad de realizar inspecciones a los aseguradores cuando lo considere oportuno.
IV. Capacidad coercitiva: Dispone normalmente de posibilidades de actuación respecto a los aseguradores que, en la mayoría de los casos, van desde la simple recomendación sobre algún aspecto concreto que se ha de mejorar o adaptar al contenido de las leyes, hasta la disolución de la entidad aseguradora cuando su situación no le permita cumplir con los fines para los que fue creada.

1.2.4 Mercado de seguros en México

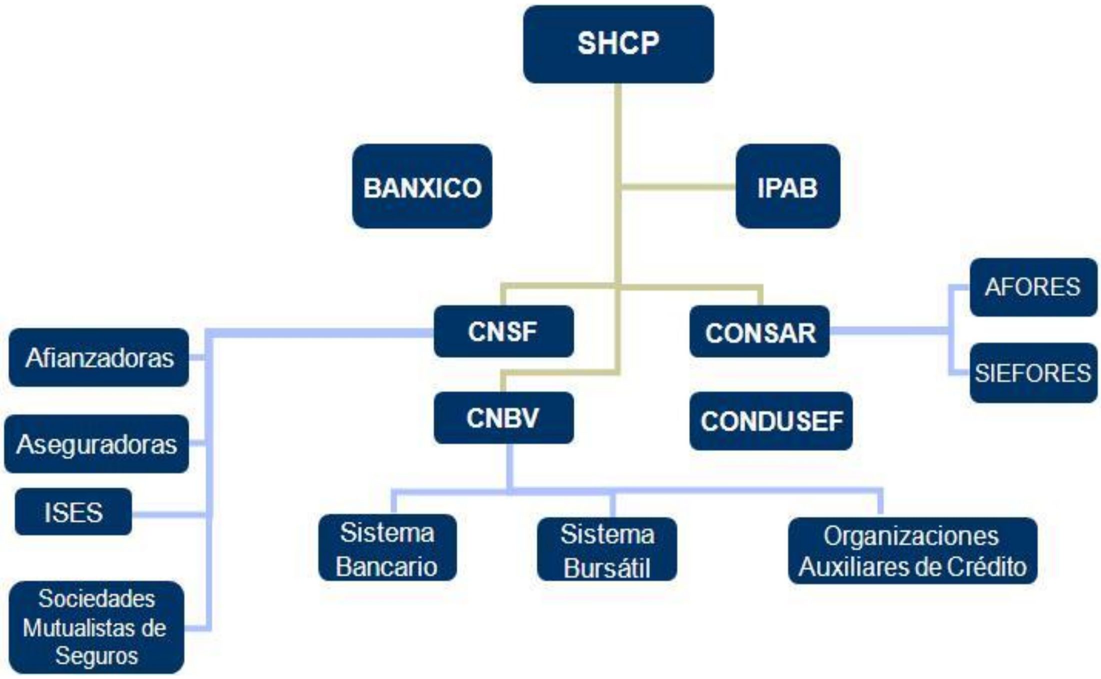
El término “Mercado de Seguros” se usa para denotar los medios disponibles para efectuar seguros y los diversos aseguradores que están dispuestos a aceptar riesgos.

Como en todo mercado, en el de los seguros se presentan también la oferta y la demanda. La oferta está representada por las empresas de seguros existentes en el mercado, así como sus intermediarios y sus auxiliares (agentes, corredores, liquidadores); y por las ramas en las cuales se ofrecen protección. Según diferentes políticas económicas de estados la oferta de seguros puede ser libre o monopólica. En nuestro país la oferta es libre ya que existen varias empresas aseguradoras.

La demanda de seguros está constituida por el conjunto de personas y empresas que requieran los servicios de los aseguradores, salvo cuando la ley exige seguros obligatorios para precautelar a terceros. El Estado juega en todos los países una función reguladora y fiscalizadora del seguro, a través de una legislación específica

1.2.4.1 Organismos reguladores

Sistema financiero mexicano



Fuente: Secretaría de Hacienda y Crédito Publico

1.2.4.2 Composición del mercado

Al finalizar el cuarto trimestre de 2015, el sector asegurador estaba conformado por 98 compañías, de las cuales una era institución nacional de seguros, dos sociedades mutualistas de seguros y el resto instituciones de seguros privadas. De las 98 instituciones, 15 empresas estaban incorporadas a algún grupo financiero, mientras que 58 presentaban capital mayoritariamente extranjero, con autorización para operar como filiales de instituciones financieras del exterior.

Mercado Asegurador Mexicano

Compañía Aseguradora	Accidentes y Enfermedades	Autómoviles	Daños s/ autos	Pensiones	Vida	Operaciones de la compañía
A.N.A. Compañía de Seguros S.A. de C.V.		✓	✓			2
ABA Seguros S.A. de C.V.		✓	✓			2
ACE Seguros S.A.	✓	✓	✓		✓	4
Agroasemex S.A.			✓		✓	2
AIG Seguros México S.A. de C.V.	✓	✓	✓		✓	4
AIG United Guaranty México S.A.			✓			1
Allianz México S.A. Compañía de Seguros	✓	✓	✓		✓	4
Armour Secure Insurance S.A. de C.V.			✓			1
Aseguradora Interacciones S.A. de C.V. Grupo Financiero Interacciones	✓	✓	✓		✓	4
Aseguradora Patrimonial Daños S.A.		✓	✓			2
Aseguradora Patrimonial Vida S.A.					✓	1
Aserta Seguros de Vida S.A. de C.V. Grupo Financiero Aserta	✓				✓	2
Assurant Daños México S.A.			✓			1
Assurant Vida México S.A.	✓				✓	2
Atradius Seguros de Crédito S.A.			✓			1
AXA Seguros S.A. de C.V.	✓	✓	✓		✓	4
BUPA México Compañía de Seguros S.A. de C.V.	✓					1
Cardif México Seguros de Vida S.A. de C.V.	✓				✓	2
Cardif México Seguros Generales S.A. de C.V.		✓	✓			2
CESCE México S.A. de C.V.			✓			1
Chubb de México Compañía de Seguros S.A. de C.V.	✓	✓	✓		✓	4
COFACE Seguro de Crédito México S.A. de C.V.			✓			1
Deco Seguros S.A. de C.V.					✓	1
Der Neue Horizont Re S.A.			✓			1
El Águila Compañía de Seguros S.A. de C.V.		✓	✓			2
Euler Hermes Seguro de Crédito S.A.			✓			1
FM Global de México S.A. de C.V.			✓			1
General de Seguros S.A.B.	✓	✓	✓		✓	4
Genworth Seguros Daños S.A. de C.V.		✓	✓			2
Genworth Seguros de Crédito a la Vivienda S.A. de C.V.			✓			1
Genworth Seguros Vida S.A. de C.V.					✓	1
Grupo Mexicano de Seguros S.A. de C.V.			✓			1
Grupo Nacional Provincial S.A.B.	✓	✓	✓		✓	4
HDI Seguros S.A. de C.V.	✓	✓	✓		✓	4
HDI-Gerling de México Seguros S.A.			✓			1
HIR Compañía de Seguros S.A. de C.V.	✓				✓	2
Horizontes Banorte Generali S.A. de C.V.				✓		1
HSBC Pensiones S.A.				✓		1
HSBC Rentas Vitalicias S.A.				✓		1
HSBC Seguros S.A. de C.V. Grupo Financiero HSBC	✓	✓	✓		✓	4
HSBC Vida S.A. de C.V.					✓	1
Insignia Life S.A. de C.V.					✓	1
Istmo México Compañía de Reaseguros S.A. de C.V.	✓	✓	✓			3
La Latinoamericana Seguros S.A.	✓	✓	✓		✓	4
Landamerica Title Insurance Company Of México S.A.			✓			1
Mapfre Tepeyac S.A.	✓	✓	✓		✓	4
MBIA México S.A. de C.V.			✓			1
MetLife Más S.A. de C.V.	✓				✓	2
Metlife México S.A.	✓				✓	2
Metlife Pensiones México S.A.				✓		1
Metropolitana Compañía de Seguros S.A.	✓	✓	✓		✓	4
Nezter Seguros S.A. de C.V			✓			1
Old Mutual Life S.A. de C.V.					✓	1
Pan-American México Compañía de Seguros S.A. de C.V	✓					1
Patrimonial Inbursa S.A.	✓	✓	✓		✓	4
Pensiones Banamex S.A. de C.V. Grupo Financiero Banamex				✓		1
Pensiones Banorte S.A. de C.V. Grupo Financiero Banorte				✓		1
Pensiones BBVA Bancomer S.A. de C.V. Grupo Financiero BBVA Bancomer				✓		1
Pensiones Inbursa S.A. Grupo Financiero Inbursa				✓		1
Pensiones Sura S.A. de C.V.				✓		1
Prevem Seguros S.A. de C.V.	✓					1
Primero Seguros S.A. de C.V.		✓	✓			2
Principal Pensiones S.A. de C.V. Principal Grupo Financiero				✓		1
Principal Seguros S.A. de C.V. Principal Grupo Financiero	✓				✓	2
Profuturo GNP Pensiones S.A. de C.V.				✓		1
Protección Agropecuaria Compañía de Seguros S.A.			✓			1
Prudential Seguros México S.A.					✓	1
QBE de México Compañía de Seguros S.A. de C.V.	✓	✓	✓		✓	4
Quálitas Compañía de Seguros S.A. de C.V.		✓	✓			2
REASEGURADORA PATRIA S.A.	✓	✓	✓		✓	4
Royal & Sunalliance Seguros (México) S.A. de C.V.	✓	✓	✓	✓	✓	5
Seguros Afirme S.A. de C.V. Afirme Grupo Financiero		✓	✓		✓	3
Seguros Argos S.A. de C.V.					✓	1
Seguros Atlas S.A.	✓	✓	✓		✓	4
Seguros Azteca S.A. de C.V.	✓				✓	2
Seguros Azteca Daños S.A. de C.V.		✓	✓			2
Seguros Banamex S.A. de C.V. Grupo Financiero Banamex	✓	✓	✓		✓	4
Seguros Banorte S.A. de C.V. Grupo Financiero Banorte	✓	✓	✓		✓	4
Seguros BBVA Bancomer S.A. de C.V. Grupo Financiero BBVA Bancomer	✓	✓	✓		✓	4
Seguros de Crédito a la Vivienda SHF S.A. de C.V.			✓			1
Seguros de Crédito Inbursa S.A.			✓			1
Seguros de Vida Sura México S.A. de C.V.	✓				✓	2
Seguros El Potosí S.A.	✓	✓	✓		✓	4
Seguros Inbursa S.A. Grupo Financiero Inbursa	✓	✓	✓		✓	4
Seguros Monterrey New York Life S.A. de C.V.	✓				✓	2
Seguros Multiva S.A. Grupo Financiero Multiva	✓	✓	✓		✓	4
Seguros Priza S.A. de C.V.					✓	1
Solunion México Seguros de Crédito S.A.			✓			1
Sompo Japan Nipponkoa Insurance de México S.A. de C.V.		✓	✓			2
SPT Sociedad Mutualista de Seguros		✓	✓			2
Stewart Title Guaranty de México S.A. de C.V.			✓			1
Thona Seguros S.A. de C.V.	✓				✓	2
Tokio Marine Compañía de Seguros S.A. de C.V.	✓	✓	✓		✓	4
Torreón Sociedad Mutualista de Seguros			✓			1
XL Seguros México S.A. de C.V.			✓			1
Zurich Compañía de Seguros S.A.	✓	✓	✓			3
Zurich Santander Seguros México S.A.	✓	✓	✓		✓	4
Zurich Vida Compañía de Seguros S.A.	✓				✓	2
Total	45	40	63	12	50	

Fuente: Comisión Nacional de Seguros y Fianzas

1.2.5 Principios de dispersión del riesgo

El ser humano puede hacer frente a los riesgos de varias maneras:

A. Transferencia del riesgo: es la forma de trasmitir el riesgo de una persona a otra en virtud del contrato de seguro, el asegurado a cambio del pago correspondiente transfiere sus riesgos al asegurador.

B. Autoseguro: situación en la que una persona, física o jurídica, soporta con su patrimonio las consecuencias económicas derivadas de sus propios riesgos, sin intervención de ninguna entidad aseguradora.

C. Subsidio: es cuando el gobierno o el estado a través de una política social ayuda o auxilia económicamente a las personas como consecuencia de una situación de emergencia para el ciudadano, una catástrofe o presiones económicas.

1.3 PRIMAS, SINIESTROS Y CONTRATOS DE SEGURO

El negocio de las compañías aseguradoras es recibir un pago en una o varias exhibiciones para asumir el riesgo de pérdida de un bien o una persona, resarciendo económicamente a un beneficiario el patrimonio perdido o el ingreso que dicha persona obtenía antes del accidente, al evento trágico se le conoce como siniestro y el convenio entre la aseguradora y asegurado debe estar claramente plasmado en un contrato. De manera más particular se abordará cada uno de estos conceptos

1.3.1 Concepto de primas

La Prima es el precio del seguro que paga el asegurado, contratante o tomador en el momento de la emisión de la póliza. La prima es, por lo general, para una vigencia anual del seguro, aunque puede excepcionalmente pagarse la prima por una sola vez, para una cobertura de varios años (prima única en seguros de vida) y también por una vigencia menor de un año (prima a corto plazo, como para el caso de un viaje, seguro de transportes de mercancías, etc.).

La prima se determina mediante sólidas bases estadísticas referidas a la frecuencia, intensidad y probabilidad de pérdidas o daños frente a un cúmulo de bienes o personas expuestas al riesgo (Osorio, 2003). De esto se ocupa precisamente el actuario de seguros, utilizando los recursos matemáticos que posee.

1.3.2 Tipos de primas

La prima puede ser de varias clases según la cuantía monetaria que tiene que satisfacer el tomador del seguro. Los siguientes conceptos no tratan de ser rígidos ni absolutos, pudiendo tener variantes. Sólo cumplen el propósito de difusión a que aspira este documento:

A. Prima de riesgo, llamada también prima pura, natural, matemática- o estadística, es la cantidad necesaria y suficiente que el asegurador debe percibir para cubrir el riesgo. Nace precisamente de la base estadística antes referida. Por ejemplo, si un grupo de 1,000 propietarios de un barrio sufre 3 incendios de sus viviendas al año como promedio, la prima de riesgo sería 3/1000, es decir, 3 al millar. Si estos siniestros ocasionan daños o pérdidas por 100.000 el total de 300.000 sería cubierto en un sistema de mutualidad a razón de 300 por cada propietario. Ésta sería la prima de riesgo de este caso sencillo y referencias.

B. Prima de tarifa, llamada también prima comercial, es la prima de riesgo más los recargos para la administración o gestión del seguro. Los recargos son los gastos de adquisición, formada básicamente por la comisión de agenciamiento que se paga al corredor, broker o intermediario; los gastos de administración, que vienen a ser los gastos en que incurre el asegurador para el manejo de la cartera de seguros, como son sueldos y gastos generales de gestión; los recargos asignados a la utilidad razonable del asegurador, llamado también el margen de beneficio.

Si se asumen sólo los tres tipos de recargo indicados como una proporción de la prima comercial o prima bruta, para efectos de simplicidad, puede señalarse el siguiente esquema para la obtención de la prima de tarifa:

$$B = \frac{P}{1 - (a + k + b)}$$

$$B \; = \; Prima\ comercial, de\ tarifa\ o\ prima\ bruta.$$

$P \; = \; Prima\ de\ riesgo.$

$a \; = \; Gastos\ de\ administración, un\ \%\ de\ B.$

$k \; = \; Gastos\ de\ adquisición, un\ \%\ de\ B.$

$b \; = \; Margen\ de\ beneficio, un\ \%\ de\ B.$

C. Prima de facturación. Es la prima de tarifa más los recargos de ley, como son los impuestos sobre la prima, los derechos de emisión y otros agregados, ordenados por disposiciones legales, así como los intereses de financiación en el caso de que el asegurador otorgue facilidades de pago fraccionado de la prima anual.

1.3.3 Concepto de siniestros

Es la concreción del riesgo, es decir, su materialización, como el incendio que devora una fábrica, el robo de mercancías, el hundimiento de una nave, la rotura de una maquinaria, el terremoto, la muerte prematura de un padre de familia, etc. Es en este momento en el que el seguro también materializa su acción de protección e indemnización. Hay personas que, al verse frente a un siniestro como el choque de su automóvil o el incendio de su negocio, dicen con razón: si no fuera por el seguro estos bienes que son el producto de mi esfuerzo o de mi sacrificio de muchos años hubieran quedado en la nada; otro habría dicho: más vale tener un seguro y no necesitarlo que necesitar un seguro y no tenerlo.

Los conceptos que se deben medir para mitigar el riesgo son los de frecuencia y severidad; la frecuencia nos indica la periodicidad con la que ocurren los siniestros y la severidad la intensidad de los mismos.

Los riesgos que el asegurador se disponga a cubrir deben presentar una regularidad en su comportamiento, tanto en cuanto a la frecuencia con que se presentan, como en cuanto a la severidad del daño económico que causan en su acaecimiento. Tal comportamiento ha de adaptarse a una determinada de regularidad estadística, que permita su tratamiento actuarial. Los riesgos esporádicos y los catastróficos o extraordinarios también deben ser considerados.

Es precisa la agrupación en una cartera de la más amplia posible masa de riesgos, no solo porque el mayor volumen de negocio permite realizar mejor la compensación, entre toda la masa de expuestos al riesgo, de los siniestros que ocurran, sino además porque a mayor número de riesgos cubiertos, menores serán las diferencias que se produzcan entre las probabilidades teóricas del sufrir pérdidas y el número efectivo de siniestros ocurridos. Además, el asegurador debe cubrir riesgos homogéneos adecuadamente diseminados en el territorio, eliminando en lo posible el peligro de cúmulos, que se presentan cuando un solo suceso puede afectar a varios bienes, aparentemente distintos e independientes. Un siniestro con cúmulos puede producir pérdidas que rebasen las previsiones máximas de los aseguradores

Uno de los principios fundamentales de la técnica aseguradora exige que los riesgos que se cubran sean homogéneos cuantitativa y cualitativamente, con el objeto de permitir una adecuada compensación entre los riesgos que se agrupan. Los riesgos asumidos deben ser de un mismo tipo o clase, concretamente del tipo a que corresponde el modelo estadístico que se pretende aplicar. La compensación está prevista para riesgos de la misma naturaleza ya que el asegurador no puede prever en cuales de los riesgos de su cartera va a producirse el siniestro. Si afecta a un riesgo cuantitativamente elevado, puede producirse un desequilibrio económico, y si afectase a uno pequeño existiría una desviación positiva.

1.3.4 Indemnización

La indemnización es el desembolso monetario que efectúa el asegurador al producirse un siniestro amparado por la póliza; es decir, La indemnización es el pago que realiza el asegurador al asegurado como consecuencia del siniestro. Sin embargo, en el lenguaje contable y estadístico de los seguros se suele utilizar la palabra siniestro como sinónimo de indemnización.

Tratándose de seguros de daños, el propósito del seguro es restituir al asegurado a su situación patrimonial inmediatamente anterior al momento del siniestro. De ahí que el seguro no pueda ser fuente de lucro para el asegurado.

La indemnización asume tres formas básicas:

1) Reparación del bien siniestrado.

2) Reemplazo del bien siniestrado por otro similar.

3) Entrega al asegurado de una suma de dinero equivalente a la pérdida sufrida.

Cuando la suma asegurada fijada en la póliza sea menor del valor real de los bienes afectados en el momento del siniestro, se trata de un típico caso de “infraseguro”. En este caso la indemnización se reducirá en la misma proporción del “infraseguro” mediante la aplicación de la fórmula:

$$I = (valor\ asegurado\ /\ valor\ real) * importe\ de\ la\ pérdida$$

1.3.5 Contrato de seguro y normatividad aplicable

El contrato corriente y característico de seguro privado es aquel en que una parte, el asegurador, contra el pago de una prima, se obliga a indemnizar al asegurado dentro de los límites convenidos, del daño que experimente a consecuencia de un siniestro o pagarle un capital, o una renta, al verificarse un evento atinente a la vida humana.

Sin embargo, el seguro puede originarse también en una norma legal. En este supuesto, la constitución de la relación jurídica asegurativa es obligatoria para el asegurado y las prestaciones de las partes no son interdependientes, sino de cumplimiento automático.

1.3.6 Características y principios del contrato de seguro

Por el contrato de seguro el asegurador se obliga mediante una prima, a indemnizar el daño causado por un acontecimiento incierto, o a suministrar una prestación al producirse un evento relacionado con la vida humana.

Puede tener por objeto toda clase de riesgos si existe interés asegurable, salvo prohibición de la ley.

Se señala que esta definición tiene las ventajas de no comprometer opinión acerca de la naturaleza del contrato y de comprender a todas las especies de seguros.

El contrato de seguro reúne las siguientes características (Palacios, 1996);

a) Sustentabilidad: La obligación del asegurador de soportar las consecuencias económicas del riesgo debe ser consecuencia de un pacto autónomo, es decir, distinto de todo otro negocio jurídico, si se tratase de una obligación legal25 o accesoria a otro contrato26, no habría seguro en el sentido estricto.

b) Formal: El contrato de seguro se consignará por escrito en la póliza, llegar a la conclusión de que la falta de póliza tiene por consecuencia que el contrato no produzca obligación ni acción en el juicio; En los seguros colectivos, la póliza de seguro queda en manos del contratante y se emite los certificados de coberturas que se entrega a cada asegurado, que trasmite la validez del contrato de seguro a los poseedores de dichos certificados.

c) Bilateral: Esto se deduce de la estructura del contrato del seguro; el tomador del seguro se obliga a pagar la prima y el asegurador se obliga a la prestación pecuniaria; si bien esta prestación está subordinada a un evento incierto, cual es la realización del siniestro.

d) Oneroso: A la futura prestación del asegurador se opone la actual prestación del tomador del seguro, bajo la forma del premio. A la obligación de pagar el premio se contrapone, como equivalente la promesa o asunción de la obligación de pagar la indemnización o el capital convenido.

e) Aleatorio: el contrato del seguro debe ser un acontecimiento incierto, en el seguro de daños, por acontecimiento incierto se comprende que en momento de contratar la póliza el asegurado no sabe cuándo ocurrirá el siniestro y el asegurador si tendrá que pagar algo, cuándo y cuánto. Continúa definiendo nuestro Código Civil que, en los seguros de personas, suministrará una prestación al producirse un evento relacionado con la vida humana, que el asegurador no sabe cuándo tendrá que pagar el capital convenido o la renta, en caso de muerte. En el caso de seguros de sobrevivencia, se obliga a pagar una renta mientras viva el asegurado, que no sabe cuándo dejará de pagar la renta.

f) Duración: El seguro es un contrato de crédito, puesto que no hay simultaneidad, en las prestaciones de los contratantes. Frente a la prestación del asegurado, consistente en el pago del premio, se ofrece la contraprestación por la otra parte en el futuro y para el caso de que se produzca un acontecimiento llamado siniestro, que sea incierto en el cuándo, como es la muerte, y en cuanto como en los demás seguros; es decir, si se produce el siniestro, entonces vencerá la obligación del asegurador de realizar su contraprestación. El contrato de seguro es un contrato cuyo contenido no se agota en un momento por el cambio de las prestaciones (como en la compraventa): el seguro es un contrato de ejecución continuada, de tracto sucesivo continuo; así el seguro engendra la relación del seguro (el estar asegurado), como vinculo continuo de las partes por un periodo más o menos largo.

g) Buena fe: El seguro es un contrato celebrado en masa en el que se ofrece la característica propia del contrato de adhesión, esto es la subordinación del contratante a las condiciones contractuales redactadas unilateralmente por la empresa aseguradora.

Esta característica exige por parte de la empresa, una observancia estricta de la máxima de buena fe, que es incompatible con las cláusulas lesivas para el asegurado, o simplemente obscuras. Contra los abusos posibles en esta forma de contratación, en nuestro país el derecho ha reaccionado con medidas preventivas; al exigir la registración previa, de los modelos de pólizas a ser comercializados en los riesgos comunes, en la Superintendencia de Seguros. En resumen, desde el punto de vista de las empresas de seguros, la buena fe consistirá en cerciorarse que el otro contratante conoce y entiende todas las cláusulas del contrato y que ninguna de ellas es peligrosa, lesiva u onerosa, ni están redactadas en términos obscuros.

1.3.7 Elementos formales del contrato de seguro

Se comenzará con las partes indispensables del contrato. Los sujetos del seguro, es decir aquellas personas sobre las cuales recaen los derechos y obligaciones del contrato, son, en principio, el asegurador y el asegurado.

El asegurador: Debe ser una entidad autorizada por la Comisión Nacional de Seguros y Fianzas. El asegurado es el titular del interés asegurado y, por lo tanto, contratante del asegurador y titular de todos los derechos, obligaciones y cargas de la relación asegurativa. En los seguros de daños patrimoniales puede ocurrir que quien contrate con el asegurador no sea el titular del interés asegurado, configurándose así el seguro por cuenta ajena donde, junto a las figuras del asegurador y del asegurado, aparece la del tomador o estipulante, que es quien celebra el contrato de seguro con el asegurador.

Tratándose del asegurador, el mismo debe ser una entidad autorizable para operar en seguros.

El asegurado: En el caso del asegurado, se debe tener en cuenta que la contratación de un seguro es en general un acto de administración, excepto en el caso del seguro de personas cuando el asegurado no es el beneficiario de la indemnización, en cuyo caso se requiere capacidad para disponer. El asegurado es siempre el titular del interés asegurado.

En el seguro de vida el asegurado puede ser el destinatario de la prestación (seguro de supervivencia) o puede ocurrir que contrate el seguro para el caso de su muerte, apareciendo entonces y sin ser parte en el contrato la figura del beneficiario. Asimismo, puede ocurrir que la vida sobre cuya existencia o duración se celebre el contrato no sea la del asegurado ni la del beneficiario, surgiendo la figura del tercero. De esta forma en el seguro de vida pueden aparecer hasta cuatro personas: el asegurador, el asegurado, el tercero cuya muerte o supervivencia determina la prestación y el beneficiario.

La LISF establece que no podrán contratarse seguros para el caso de muerte de los interdictos y de los menores de 14 años, y que los menores de edades mayores de 18 años pueden contratar seguros sobre su propia vida sólo si designan beneficiarios a sus ascendientes, descendientes, cónyuge o hermanos que se hallen a su cargo

El tomador o el contratante: Es la persona que celebre el contrato con la empresa aseguradora y puede hacerlo por cuenta propia (en este caso es asegurado en tanto sea el titular del interés asegurable) o por cuenta ajena (como en el caso de que tome un préstamo financiero y se vea obligado por el contrato de préstamo a tomar un seguro de vida para cancelación de deudas). En ambos casos es el obligado al pago de las primas.

Beneficiario: Es la persona que recibirá la indemnización en caso de siniestro. Generalmente es el mismo asegurado o contratante. Pero no siempre es así, en el caso de seguros de vida, al fallecer el asegurado, el beneficiario puede ser algún miembro de su familia, sus herederos legales o cualquier persona previamente designada en el contrato por el asegurado o contratante, en estos casos el beneficiario no es parte del contrato, ya que no participa en la celebración, pero tiene derechos sobre las prestaciones del asegurador.

El contrato de seguro debe probarse por escrito, mediante la póliza de seguro. Sin embargo, se admiten los demás medios de prueba, siempre que exista principio de prueba por escrito.

Propuesta o solicitud del asegurado: Dado el carácter consensual entre las partes para dar inicio a la vigencia de un contrato, el acuerdo de voluntades se manifiesta en la propuesta, la sola firma de la propuesta por el posible asegurado, no obliga aún a la aseguradora, ya que se está hablando de una oferta de negocio y aceptación por la otra parte. En consecuencia, no constituye un precontrato. La propuesta del contrato, cualquiera sea su forma, no obliga al asegurado ni al asegurador. La propuesta puede subordinarse al conocimiento previo de las condiciones generales.

Si bien, la oferta o la propuesta del seguro no forma parte integrante de la póliza, en la práctica se adjunta como antecedente del contrato y es esencial para la interpretación del contrato, ya que la propuesta debe indicar e identificar al riesgo con los mayores detalles tales como: capital a ser asegurado, su ubicación, identificación y calificación del riesgo, etc.; en los seguros de personas la propuesta es mucha más detallada ya que normalmente se adjunta a la propuesta la declaración de salud, ésta última puede ser bajo la característica de declaración jurada o la declaración de un médico contratado por la compañía que le haya evaluado. De ahí lo mencionado en el párrafo anterior donde se exceptúa en los seguros de personas la propuesta de prorroga sin una verificación previa del asegurado y en todo caso si el asegurado tiene un contrato individual de un seguro de personas, en la misma está contemplado cuánto tiempo más puede prorrogarse dicho seguro.

La propuesta no obliga al asegurador, aunque la propuesta lleve el membrete de la aseguradora.

La póliza: La palabra póliza, deriva del italiano *póliza*, es el instrumento en el que se hace constar la totalidad de las condiciones y formalidades de un contrato. Comúnmente se refiere al contrato de seguro, aunque también puede aplicarse a otros tipos de contratos, como la póliza de fletamento, etc.

Como contrato de seguro, es un documento impreso redactado por el asegurador, su texto debe estar en castellano y ser registrado previamente ante la Superintendencia de Seguros, salvo aquellas con cláusulas de riesgos muy específicos, las que tienen un plazo de 30 días para su registración después de emitirse la póliza.

Normalmente, la póliza comprende dos partes, las Condiciones Generales y las Condiciones Particulares y en algunos casos específicos tienen las denominadas condiciones particulares específicas. Normalmente forman parte del contrato de seguro, los certificados, que se emplean habitualmente en pólizas flotantes para amparar cada una de sus aplicaciones o en las pólizas colectivas, para poner en manos del asegurado individual un elemento probatorio del contrato.

Condiciones Generales: La función importante de las condiciones Generales se deduce del hecho de que todos los contratos de seguros se pactan sobre la base de las condiciones generales, que las compañías aseguradoras proponen para cada sección del ramo en que están autorizados a operar y que forman parte del contrato, De esta forma, se ofrecen como configuración típica del contenido del contrato y tienen gran valor como Derecho del seguro no codificado. En este sentido constituyen verdadero derecho consuetudinario. Por ello han de pasar por la inspección de la Autoridad de Control para ser registrados y verificar que se adecuen a lo dispuesto en las leyes vigentes.

Condiciones Particulares: Recogen aspectos concretamente relativos al riesgo individualizado que se asegura y en particular los siguientes;

- Nombre y domicilio de las partes contratantes, y designación del asegurado y beneficiario, en su caso.
- Concepto en el cual se asegura.
- Naturaleza del riesgo cubierto.
- Designación de los objetos asegurados y de su situación.
- Suma asegurada y alcance de la cobertura.
- Importe de la prima, recargos e impuestos.
- Vigencia del contrato, con expresión de cuando comienzan y terminan.
- Forma de pago, entrega inicial y cantidad de cuotas.

Condiciones Específicas: Cuya misión es matizar o perfilar el contenido de algunas normas recogidas en aquellas, el establecimiento de franquicias a cargo del asegurado, la supresión de algunas exclusiones y la inclusión de otras nuevas.

Exclusiones: Hay riesgos que el asegurador no asume, y que deben figurar en el contrato, ya sea por ser contrarios al espíritu del seguro, ya sea por su carácter catastrófico, ya sea por no corresponder al ramo del seguro que pertenece la póliza. Entre ellas algunas son generales, es decir que se presentan en todos los ramos del seguro, y otras que son particulares, o sea que afectan solo a determinadas secciones del seguro, citando algunas:

- El dolo del asegurado son riesgos que no se pueden asegurar.
- El vicio propio (o sea el germen de destrucción o deterioro que llevan en sí mismas las cosas), como por ejemplo la merma de los cereales no son asegurables por no tener origen fortuito, sin embargo, se las cubre en la parte que supere lo previsto, mediante pacto expreso.

- El desgaste es una pérdida previsible y no fortuita, sin embargo, los siniestros ocurridos como consecuencia de estos, pueden ser cubiertos
- Los daños que tienen su origen en actos de autoridad política, civil o judicial no suelen ser cubiertos.
- Los daños causados por hechos de guerra internacional o civil, revoluciones, etc.; y los provocados por fenómenos de la naturaleza como erupciones volcánicas, terremotos y los riesgos nucleares, salvo pacto en contrario.

Anexos, endosos y cláusulas adicionales: contienen ampliaciones, modificaciones o restricciones a las condiciones principales establecidas en las pólizas.

1.4 RAMOS Y TIPOS DE SEGUROS

Como ya se mencionó anteriormente la Comisión Nacional de Seguros y Fianzas atreves de la ley general de seguros y fianzas clasifica a los seguros por operación y ramo agrupando a los seguros de la siguiente manera:

I. Vida.

II. Accidentes y enfermedades:

- Accidentes personales.
- Gastos médicos.
- Salud.

III. Daños:

- Responsabilidad civil y riesgos profesionales;
- Marítimo y transportes.
- Incendio.
- Agrícola y de animales.
- Automóviles.
- Crédito.
- Caución.

- Crédito a la vivienda.
- Garantía financiera.
- Riesgos catastróficos.
- Diversos.
- Los especiales que declare la Secretaría, conforme a lo dispuesto por LGSF.

1.4.1 Seguro de Incendio

Cubre los daños que ocasione el fuego o llama. Puede cubrir, bajo la misma póliza, una gama de otras coberturas adicionales llamadas líneas aliadas, tales como explosión, terremoto, daños por agua, daños por humo, inundación, hasta huelgas y conmociones civiles, daño malicioso y vandalismo, impacto de vehículos, terrorismo, etc.

1.4.2 Seguro de Automóvil

Cubre los riesgos a los que está sometido un vehículo de motor con motivo de su circulación por vías oficialmente autorizadas.

Los daños cubiertos pueden ser al propio vehículo (choque o vuelco accidental, incendio, rotura de cristales, robo) y daños terceros (personales o materiales).

El capital asegurado por daño propio debe ser hasta por el valor real o comercial del vehículo, mientras que para responsabilidad civil (daños a terceros) puede ser una suma convenida, independiente del valor del vehículo. Hay seguros de responsabilidad civil obligatorios y voluntarios.

1.4.3 Seguro de Riesgos Diversos

Los riesgos diversos se dividen en dos: diversos técnicos y diversos misceláneos.

Los diversos misceláneos o generales cubre los daños por rotura de cristales, cualquier daño a anuncios luminosos, robo con violencia a casa habitación y mercancías, robo de dinero y documento de valor, y robo de objetos personales.

Los diversos técnicos o de ingeniería cubre los daños a calderas y recipientes sujetos a presión, rotura de maquinaria, equipo para contratistas y maquinaria pesada móvil, equipo electrónico, y montaje de maquinaria.

1.4.4 Seguro contra Accidentes Personales

El seguro de accidentes y enfermedades cubre los daños por lesión o incapacidad que afecte la integridad personal, salud o vigor vital del asegurado, como consecuencia de un evento externo, violento, súbito y fortuito.

1.4.5 Seguro de Gastos Médicos

El seguro de gasto medico tiene como objeto cubrir los gastos médicos, hospitalarios y demás que sean necesarios para la recuperación de la salud o vigor vital del asegurado, cuando se hayan afectado por causa de un accidente o enfermedad.

1.4.6 Seguro de Transporte

Llamado también seguro de carga, cubre las mercancías o carga conducida por la nave si el transporte es marítimo. También la carga cubierta por el seguro puede ser aérea, terrestre, fluvial o lacustre. Los riesgos inherentes al transporte abarcan todo daño o pérdida, falta de entrega, merma, robo, etc., desde el almacén de origen hasta el del importador.

1.4.7 Seguro de Responsabilidad Civil

Cubre los daños que una persona o empresa puede causar a terceros, para lo cual aquella fija en la póliza un límite de cobertura hasta el cual puede la entidad aseguradora indemnizar a la víctima por el daño que el asegurado le ha causado. Hay diferentes modalidades de responsabilidad civil: general, de productos, patronal, es decir, de tipo contractual y extracontractual.

1.4.8 Seguro de Aeronavegación

Ampara a las naves, buques, yates y toda embarcación contra los llamados riesgos del mar, como son el naufragio o hundimiento, la varadura, encallamiento o choque con elementos naturales; la colisión, abordaje o choque con otra nave, incendio, etc.

1.4.9 Seguro Agropecuario

También conocido como seguro agrícola y de animales, este seguro pretende indemnizar o resarcir inversiones, por los daños o perjuicios que sufran los asegurados por pérdida parcial o total de las ganancias esperadas de la tierra o por muerte, pérdida o daños ocurridos a sus animales.

1.4.10 Seguro de Vida

Debería llamarse seguro contra el riesgo de muerte (prematura), pero, por razones de orden psicológico y comercial, la palabra muerte se elimina en el nombre del seguro, como en toda la extensión de la póliza.

Este seguro garantiza el pago de un capital a los deudos o beneficiarios del asegurado en el momento en que éste fallece. Hay diferentes formas, modalidades, clases de seguros de vida.

1.5 CONTRATOS DE REASEGURO

Una de las formas de diversificar los riesgos adquiridos en contratos de seguros es la del reaseguro y el coaseguro en el cual compartes y cedes una parte del riesgo con otras instituciones para evitar caer en el caso de insuficiencia de pago a causa de siniestros catastróficos.

1.5.1 Coaseguro

Es la distribución del riesgo entre dos o más aseguradores. El que emite la póliza se llama entidad líder o administradora y las demás son coaseguradoras. Hay tres tipos de coaseguro:

a) Técnicamente el coaseguro es típico o puro cuando el propio asegurado es quien solicita por escrito que el riesgo sea distribuido en coaseguro, señalando las proporciones respectivas, en cuyo caso cobrará de cada coasegurador el importe proporcional del siniestro que ocurra.

b) Coaseguro interno es el que resulta de la iniciativa o convenio de la entidad líder y demás coaseguradoras. El asegurado no conoce este convenio y se entiende únicamente con quien emitió la póliza.

c) Coaseguro pactado es una distribución del riesgo entre el asegurador y el propio asegurado, con el objeto de obtener una rebaja en la prima según normas tarifarías. Puede haber fórmulas tales como 80/20; 70/30, etc., donde el asegurado se convierte en su propio asegurador por una proporción generalmente menor del riesgo.

1.5.2 Concepto de reaseguro

Es la cesión del riesgo que efectúa el asegurador (cedente) a otra entidad llamada reaseguradora, según las prescripciones señaladas en un convenio especial llamado tratado de reaseguro.

Sí el coaseguro es la distribución horizontal del riesgo, el reaseguro es la distribución vertical del riesgo. Una participación del coaseguro puede también reasegurarse. La porción asumida por el asegurador o coasegurador se llama retención. El reasegurador asume el excedente conjuntamente con la prima que corresponde a este exceso. El reaseguro es toda una institución, eminentemente internacional, sujeta a normas técnicas, jurídicas, comerciales y doctrinas propias. Es válido en estos comentarios afirmar categóricamente que el reaseguro es el seguro del seguro y es la prueba más evidente de la solidaridad internacional del seguro.

1.5.3 Clasificación del reaseguro

Dada la heterogeneidad cuantitativa de los riesgos asegurables y la imposibilidad de que la intensidad de los siniestros sea uniforme, una compañía aseguradora debe encontrar algún recurso para lograr eliminar los factores que pueden desequilibrar tanto el capital como la reserva acumulada.

Una alternativa a esta situación es el reaseguro, es un contrato en virtud del cual, el reasegurador toma a su cargo los riesgos de la cedente ,en una proporción de las obligaciones de ésta frente a su cliente, de una manera autónoma e independiente, y por lo cual recibe la parte proporcional de la primas correspondientes a los riesgos asumidos, o bien cubre a la cedente resarciéndole, en su caso, por las desviaciones de siniestralidad esperada, cobrándose una prima convenida a la celebración del contrato.

En sus principios, el reaseguro se practicaba negociando la trasferencia riesgo por riesgo, conocida ésta como reaseguro facultativo. Con el desarrollo industrial y comercial del siglo XIX, el seguro tomó un auge que hizo necesario buscar formas más flexibles de cobertura. Por esta razón, se establecieron contratos automáticos de reaseguro, para cubrir prácticamente todos los negocios aceptados por una compañía en un determinado ramo.

Estos dos grupos de contratos, facultativo y automático, son usados por el asegurador para transferir ya sea el riesgo o el siniestro al reasegurador. Estos contratos se dividen, a su vez, en contratos proporcionales y no proporcionales.

En el reaseguro automático, la compañía que trasfiere el riesgo o el siniestro da un porcentaje determinado de todos los contratos de un tipo y la reaseguradora tiene la obligación de aceptarlo.

En cambio, el reaseguro facultativo es un convenio donde la compañía que transfiere el riesgo o el siniestro no tiene la obligación de dar el contrato, sino conserva la libertad de decidir que negocios y en que amplitud desea reasegurar y, asimismo, el reasegurador está en la libertad de aceptar o no el contrato. Es decir, tanto la cedente como el reasegurador tienen la libertad de proponer, aceptar o rechazar un negocio determinado.

En los contratos proporcionales, el reasegurador acepta una parte fija de la responsabilidad asumida sobre un riesgo suscrito por la cedente, haciéndose cargo tanto de las obligaciones (siniestro), como de los derechos (prima, previa deducción de una comisión de reaseguro destinada a cubrir los gastos de adquisición y administración). En este tipo de contratos, se hace una trasferencia proporcional de riesgos y primas.

En los contratos proporcionales se pueden distinguir los siguientes tipos:

Contrato de cuota parte: En el contrato de cuota parte, el reasegurador acepta una proporción fija de todos los riesgos aceptados por la compañía cedente. De esta forma, participa proporcionalmente en todos los siniestros y recibe a cambio la misma proporción de todas las primas netas. En dichos contratos se estipula que la compañía cedente, cederá automáticamente y el reasegurador aceptará la participación acordada en todos los riesgos suscritos que se ajusten al contrato.

Contrato de excedente. Otro de los contratos proporcionales, es el contrato con base en excedentes de retención, donde la compañía cedente no está obligada a ceder todos los riesgos que acepte de sus asegurados; solo cede aquella parte de los riesgos que superen su propia capacidad. En estos contratos la cedente adopta límites de retención variables, relacionada directamente con los niveles de los distintos riesgos asegurados.

El contrato ordinario es llamado primer excedente, lo que significa que los riesgos que “exceden” el límite de retención, alimentan a este contrato antes que a cualquier otro. Los contratos subsecuentes son convenidos como segundo excedente, tercer excedente, etcétera, recibiendo éstos la parte correspondiente después que el contrato anterior haya recibido el monto completo al cual tiene derecho.

El reasegurador recibe la prima proporcional al riesgo que asume y pagará al riesgo que asume y pagará los eventuales siniestros en la misma proporción. La cobertura siempre se expresa en un múltiplo pleno de retención (límite de retención) y se indica también el monto máximo que puede ser cedido al reasegurador.

Contrato open cover. En este tipo de contrato, el límite de responsabilidad no se establece con base en múltiplos de pleno de retención de la cedente, de tal suerte que, independientemente del importe retenido, la cedente puede llenar el contrato a su máxima capacidad, es decir, no tiene límites precisos (se maneja como un reaseguro facultativo). Este tipo de contratos no se otorgan fácilmente por parte de los reaseguradores, debido a las características tan abiertas del contrato, que pueden provocar grandes desviaciones en siniestralidad.

Al igual que en los contratos proporcionales, existen diversos contratos no proporcionales, entre los que se encuentran:

Working cover: En este contrato la cedente busca incrementar el volumen de primas retenidas, sin exceder de una suma determinada su aportación en cada siniestro por riesgo. Con frecuencia se emplea para esta clase de cobertura el término inglés working excess of loss (WXL). Este tipo de contrato protege contra siniestros que sobrepasen determinada parte del importe que decidió conservar la institución de seguro por cuenta propia de un riesgo dado.

Contrato XL Catastrófico: Otra cobertura de los contratos no proporcionales es la de exceso de pérdida catastrófico, cuya nomenclatura es XL Catastrófico, el cual cubre el riesgo en caso de la acumulación o agregación de pérdidas derivadas de un suceso o acontecimiento de naturaleza catastrófica (tempestades, terremoto, etcétera).

Stop loss: Este contrato protege los resultados anuales de una compañía en un ramo contra una desviación negativa debido a una incidencia de siniestros mayor a la esperada, ya sea por el número o la importancia de estos eventos.

En estos convenios, el reasegurador no es responsable del pago de ningún siniestro hasta que la tasa de siniestralidad exceda el porcentaje convenido de las primas. A partir de este punto, el reasegurador paga todos los siniestros, grandes o pequeños, pero sin rebasar el límite de responsabilidad establecido en el contrato.

El uso de contratos no proporcionales se formaliza fijando montos de prioridad y cobertura, siendo la prioridad el monto máximo de pérdida que corre a cargo de la cedente y el remanente de la pérdida que correrá a cargo del reasegurador del exceso de pérdida hasta por la cantidad fijada como límite de cobertura. En compensación del compromiso de asumir los montos de siniestros que sobre pasan el límite fijado como prioridad a cargo de la cedente, el reasegurador recibe un monto del volumen de primas generadas por el negocio cubierto.

1.5.4 El uso del reaseguro en una cartera de riesgo

Ahora que se ha detallado el funcionamiento de los diferentes tipos de contrato de reaseguro, se puntualizara su funcionalidad, ventajas y desventajas que ofrecen al emplearse en una cartera de riesgo.

Los contratos proporcionales cumplen la función de compartir una parte del riesgo expuesto en la cartera a cambio de la proporción de prima correspondiente con la finalidad de reducir la responsabilidad de la compañía y así tener un requerimiento de capital menor. Este tipo de reaseguro es conveniente emplearse cuando el volumen de la cartera es grande y la suscripción de los riesgos sigue políticas claras y precisas, ya que es lo que permite a la reaseguradora confiar en la transferencia automática del riesgo sin un análisis riesgo por riesgo.

Algunas de las desventajas son el no poder variar la retención del asegurador en riesgos muy expuestos a un siniestro, y el transferir gran parte de los riesgos que pudieran quedar a retención por tener poca siniestralidad, esto refuerza el argumento de que para que el uso de contratos proporcionales resulte benéfico para ambas partes se debe contar con una cartera grande y con un equilibrio en la exposición de los riesgos.

Dentro de los contratos proporcionales, el empleo de contratos cuota parte resulta menos costoso en términos operativos debido a que todos los riesgos se reparten de la misma manera entre la cedente y la reaseguradora, por otro lado, el uso de excedente implica el registro de porcentajes de retención riesgo por riesgo tanto para la cesión de primas como de siniestros.

Por otra parte, los contratos no proporcionales tienen como objeto resarcir las pérdidas de la cedente cuando se excede el monto que desea asumir a retención, ya sea por riesgo, por evento o por siniestralidad total en un ejercicio. El empleo de estos contratos requiere de un adecuado análisis ya que implican un costo cierto y la aplicación de los contratos es incierto.

Se empleará un contrato en exceso de pérdida cuando se tengan riesgos expuestos a una siniestralidad que desviará el resultado de la compañía de forma significativa, encontrar el punto en el que los montos de siniestros son significativos o no dependen tanto de la frecuencia y severidad de los que se puede considerar como siniestros atípicos como del comportamiento del resto de la cartera. Si la mayor parte de la siniestralidad de la cartera proviene de frecuencia y no de severidad, entonces resulta conveniente adoptar esquemas no proporcionales, ya que estos estarán protegiendo a la compañía de una desviación por un siniestro muy grande, no así por incremento en la frecuencia, esta medida debe ser controlada con la política de suscripción de la compañía, ya que dicho incremento implicaría la aceptación de riesgos altamente expuestos.

Por su parte los contratos Stop Loss se emplearán cuando se determine un monto máximo de siniestralidad en el año por parte de la compañía para garantizar el resultado de la misma. Este tipo de contratos resultan complicados de obtener en el mercado reaseguro debido a que no es sencillo para las reaseguradoras cuantificar el impacto que podría tener un contrato de este tipo y por tanto el costo que deberían tener, ya que la siniestralidad en todo el ejercicio obedece tanto a factores de mercado como a las políticas propias de la cedente.

1.6 SOLVENCIA

Solvencia financiera es la capacidad de una empresa para cumplir todas sus obligaciones sin importar su plazo. En ocasiones es referida como liquidez, pero ésta es solo uno de los grados de solvencia. Se dice que una empresa cuenta con solvencia cuando está capacitada para liquidar los pasivos contraídos al vencimiento de los mismos y demuestra que podrá conservar dicha situación en el futuro.

1.6.1 La solvencia del asegurador

Hay dos razones por las que la solvencia es un objetivo de especial significación para el sector asegurador. La primera obedece a la naturaleza de las prestaciones a las que se compromete quien cubre el riesgo; dichas prestaciones vienen habitualmente asociadas a situaciones de necesidad del asegurado, y es natural que quiera reducirse todo lo posible el riesgo de incapacidad del asegurador para responder a su compromiso. Por otra parte la quiebra de la solvencia puede obedecer a la simple fluctuación la solvencia adecuada aleatoria de la siniestralidad sin que necesariamente hayan de concurrir en su aparición factores ajenos al puro azar. Más aún, el riesgo de insolvencia es susceptible de tratamiento estadístico-actuarial y su garantía entre así de lleno en las obligaciones del asegurador, que ha de garantizar su propia solvencia como un aspecto más de su actividad.

Cabe añadir aun otro argumento para justificar la hegemonía que la solvencia y los medios de salvaguardarla presentan en la economía del empresario asegurador, y es el papel fundamental que la confianza juega en esta actividad; un deterioro de dicha confianza puede apartar al público de los mecanismos del seguro, por lo que el empresario es el primer interesado en afrontar la cuestión de la solvencia con criterios técnicos rigurosos.

El concepto de solvencia presenta dos acepciones distintas según se refiera a la solvencia estática o a la solvencia dinámica .la primera es la capacidad del asegurador para hacer frente a las obligaciones derivadas de los compromisos ya adquiridos; se trata, por tanto, de que las provisiones para riesgos en curso, matemáticas y para prestaciones pendientes, estén bien calculadas e invertidas en valores realizables por importe suficiente. La solvencia dinámica es la capacidad del asegurador para cumplir los compromisos que puedan aparecerle como consecuencia de su actividad futura. La solvencia estática contempla la capacidad del asegurador en un momento dado para pagar las indemnizaciones derivadas de las primas contabilizadas. Si dichas primas representan el valor medio de la siniestralidad y esta no se ha apartado de dicho valor medio, el asegurador contara con las disponibilidades necesarias para compensar la siniestralidad si determinó correctamente en beneficio; por lo tanto, el cálculo de las provisiones técnicas de primas y prestaciones sean correctas, e invirtió en aquellas disponibilidades de forma satisfactoria.

Ahora bien, el importe de la siniestralidad puede experimentar fluctuaciones alrededor de su valor medio; que debe coincidir con la prima de riesgo, y de ahí surge el segundo aspecto, dinámico, de la solvencia. Los medios para lograr esta solvencia dinámica consisten en la exigencia de garantías financieras por encima de las provisiones técnicas de prima y de prestaciones y son fundamentalmente el margen de solvencia y las provisiones para desviación de la siniestralidad, es decir, elementos del patrimonio asegurador que no están afectados a los compromisos contraídos en virtud de las primas ya emitidas. El margen de solvencia se proyecta hacia el porvenir y considera la solvencia en relación con la evolución prevista de la siniestralidad, el dinamismo de la empresa y del medio en que opera. A este concepto se denomina como margen mínimo de solvencia dinámica. (Prieto, 2005)

Que el asegurador cuente con un patrimonio libre no comprometido para garantizar su solvencia es necesario para el desempeño de su actividad y para la obtención de los ingresos que la justifican. Resulta lógico por tanto que la legislación y la práctica tributaria reconozcan adecuadamente aquellas necesidades y atribuyan a las dotaciones correspondientes el carácter de partida deducible en la determinación de la renta sujeta a imposición.

De otro punto de vista, la solvencia adecuada del sector asegurador proporciona beneficios para la economía en su conjunto al aumentar la capacidad de retención de riesgos del país considerado. Caso típico que ilustra lo dicho es Finlandia, en donde se introdujo en 1953 una avanzada legislación en materia de solvencia. El consiguiente reforzamiento de la misma ha sido considerado como causa de un notable incremento de la capacidad de retención de riesgos del mercado finlandés: en 1950 se cedía al extranjero casi un tercio de las primas del seguro directo, sin que las aceptaciones del exterior representaran importancia alguna; en 1980 solo se cedió el 12% de las primas fuera del país y las aceptaciones en reaseguro alcanzaron el 30% del negocio.

1.6.2 Los riesgos de desviación de siniestralidad

La suficiencia de las primas para cubrir la siniestralidad es una hipótesis de la teoría del riesgo que siempre se pone de manifiesto con la claridad que sería de desear, se supone que las primas de riesgo representan el valor medio de la siniestralidad y ésta es una provisión que en la práctica debe controlarse, tanto por la empresa misma como por las autoridades correspondientes de supervisión.

Si las primas son insuficientes, las conclusiones derivadas de los cálculos actuariales para la determinación del margen mínimo de solvencia o de la provisión técnica de estabilización, han de ser revisados a la luz de tal realidad.

Ahora bien, aun suponiendo que las primas sean suficientes, la siniestralidad anual es una variable aleatoria y como tal la fluctuará alrededor de su valor medio. Este riesgo de fluctuación justifica la necesidad del margen de solvencia adicional a las provisiones técnicas de primas y siniestros. Pero hay otros riesgos que comprometen la estabilidad del asegurador. Uno de ellos es el de alteraciones en las probabilidades básicas del proceso de riesgo. La probabilidad de que suceda un siniestro puede variar año tras año, así como la distribución de probabilidades de tamaño o cuantía de cada siniestro; la inflación juega aquí un papel importante, al contribuir que la probabilidad de que el valor del siniestro supere una cierta cuantía sea cada vez mayor. Otros riesgos importantes que corre el asegurador son el de sufrir pérdidas en sus inversiones, el de insolvencia de sus reaseguradores, los de tipo de cambio, etc. Estos riesgos son de tal naturaleza que su eliminación resulta prácticamente imposible; en unos casos porque no se puede prever o cuantificar económicamente, y en otro caso, porque aun siendo medibles, su reducción a límites inapreciables elevaría las garantías financieras a tales niveles que serían prácticamente insoportables para la economía del asegurador. Por ello, hay que aceptar la imposibilidad del empeño de lograr la seguridad total, aunque también es cierto que los compromisos razonables entre riesgos y exigencias financieras son perfectamente posibles.

Otras características de los riesgos aludidos es el de su carácter particular o individual para cada empresa: la política de reaseguro, la de selección de riesgos y la de inversiones, y otros tantos campos de la decisión empresarial son factores fundamentales que configuran la posibilidad de solvencia. Por tal razón, el único enfoque que en rigor seria satisfactorio al plantear la solidez del ente asegurador es el que exigiera la evaluación individual de sus riesgos, a la vista de sus particulares circunstancias. Este punto de vista exigirá un estudio estadístico actuarial que, en base a la retención, distribución del tamaño o cuantía del siniestro y probabilidad de ruina, determinará el margen mínimo de solvencia. Según este planteamiento, bastaría que la legislación estableciera las pautas para los pertinentes estudios actuariales y, quizá, la probabilidad máxima de ruina. El enfoque seguido en las prácticas es, sin embargo, el opuesto: las normas generales establecen reglas precisas para determinar el margen de solvencia o las dotaciones mínimas a la provisión de estabilización o súper siniestralidad; estas normas han de ser cumplidas por la generalidad de los aseguradores que operan en el mercado.

El segundo de los planteamientos anteriores tiene la ventaja de su mayor simplicidad y facilidad de control. Un tercer punto de vista, mixto de los dos anteriores, consistiría en que la legislación establecería unas reglas precisas que sirvieran para determinar las exigencias financieras generales; las empresas podrían apartarse de la regla general mediante las justificaciones técnica-actuariales oportunas y a su vez, las autoridades de control, exigirían niveles superiores cuando pudieran probar la insuficiencia de la regla general para el caso concreto. Sin embargo, esta tercera vía presenta, al menos, dos inconvenientes: uno, el de la seguridad jurídica que acarrearía y otro, el de las posibles distorsiones en la concurrencia que podrían derivarse de ella; precisamente uno de los méritos que se acreditan al margen de solvencia del mercado internacional de seguros es el de favorecer la sana competencia y promover la libertad de establecimiento.

1.6.3 La garantía de la solvencia

La solvencia estática descansa sobre el adecuado cálculo e inversión de las provisiones de primas y siniestros. Entre las primeras deben concluirse las que complementan la simple periodificación de éstas, cuando son insuficientes para cubrir el riesgo. Dentro de las provisiones de siniestros quedan comprendidas las exigidas por aquellos que han ocurrido al cierre del ejercicio pero que a tal fecha aún no se habían comunicado al asegurado: es el caso de las llamadas IBNR. (incurred, but not reported).

La solvencia dinámica se fundamenta ante todo en la suficiencia de las primas, y después de la disponibilidad de recursos no afectos a los compromisos derivados de las devengadas. Se trata, obviamente del margen de solvencia o patrimonio propio no comprometido, y de las provisiones técnicas de desviación de siniestralidad, que sirven para compensar excess anuales de siniestralidad sobre las previsiones que fundamentaron el cálculo de la prima. Un factor decisivo en la evaluación de la solvencia dinámica es el reaseguro.

En el cálculo de las exigencias financieras de la solvencia dinámica pueden adaptarse tres enfoques diferentes, que construyen algunos temas de la teoría del riesgo: pueden adoptarse como *horizonte temporal* el año o ejercicio económico, el periodo de varios años o bien el periodo de duración infinita.

Según el primer punto de vista, el margen de solvencia o las provisiones de equilibrio se calculan de forma que garanticen, con una determinada probabilidad, la solvencia de asegurador durante un año. Se supone que ello es satisfactorio, habida cuenta del control que sobre la solvencia ejercen las autoridades administrativas, las cuales, si detectan insuficiencia de las garantías exigidas, prohíben al asegurador que continúe sus operaciones o le obligan a la liquidación de la empresa.

El segundo punto de vista es más ambicioso y trata de garantizar la solvencia durante un mínimo de *n* años. Este enfoque tiene que ver más con la continuidad de la empresa que con la protección de los intereses del asegurado, propia del primer planteamiento.

El tercer ángulo para el planteamiento de la solvencia es, no solo el más exigente, sino también el que mayores dificultades presenta en cuanto a su formulación rigurosa; algunas de las hipótesis sobre las que descansa dicha formulación están, desgraciadamente, demasiado alejadas de la realidad; por ejemplo, la acumulación ilimitada de reservas.

Cabe preguntarse qué ventajas e inconvenientes presentan respectivamente, las dos formas de las *exigencias financieras* anteriormente mencionadas: el margen de solvencia y las provisiones para la desviación de siniestralidad. El margen de solvencia es una magnitud global para el conjunto de la empresa, compuesta por partidas de su pasivo significativas de los recursos propios, a las que se añaden la infravaloración de activos y otros elementos más de parecida significación. Su ventaja es la sencillez, y la homogeneidad con que afecta a los diversos aseguradores. Tiene, no obstante, varios inconvenientes, que no presentan las provisiones para desviación de siniestralidad. El primero de estos inconvenientes consiste en que el margen de solvencia no separa suficientemente las características técnicas de los diversos tipos de riesgos cubiertos, puesto que no considera al negocio como formado por varios ramos, sino en su totalidad. La experiencia demuestra que la variabilidad estadística del seguro de crédito es superior a la del seguro de accidentes, por ejemplo. El segundo inconveniente es de orden fiscal: si el margen está compuesto por el capital y las reservas libres del asegurador, entre otras partidas, no es fácil argumentar antes las autoridades tributarias que las dotaciones anuales necesarias para mantener el margen dentro de los mínimos establecidos es partida deducible en la determinación del beneficio. En cambio, a las provisiones de equilibrio les presenta su naturaleza técnica un carácter más compatible con un adecuado tratamiento fiscal. Un tercer inconveniente del margen de solvencia es el siguiente: no es fácil aceptar que ley regule la inversión de capital y las reservas libre del asegurador, es decir, los activos que materializan elementos del margen; y sin embargo, la seguridad de las inversiones es tan importante como el cálculo del margen mínimo de solvencia. No deja de ser extraña la asimetría que se observa en las reglamentaciones de muchos países. Determinan minuciosamente los elementos constitutivos del margen de solvencia, así como el importe mínimo de este; las autoridades de control adoptan severas medidas en relación con el asegurador que presente un déficit del margen. Y, sin embargo, nada se dice en relación con los activos en que se materializa el margen.

Una pregunta que lícitamente cabe formular es si resulta necesario el control de la administración pública sobre la solvencia, y se concluye que la necesidad de control depende de la propia entidad aseguradora: si ésta se propone objetivos a largo plazo, el control no es necesario, pues sus gestores marcarán unos niveles de solvencia superiores a los que, con el exclusivo fin de proteger a los aseguradores, exijan las normas generales en vigor.

La garantía de la solvencia tiene como finalidad primordial la protección de los asegurados y terceros perjudicados. Quizá esta reflexión justifique el hecho de que en ciertos países no se ejerce control sobre los reaseguradores.

1.7 RESUMEN

El primer capítulo se dedicó a identificar los términos más comunes relacionados con el sector asegurador.

Se comenzó por definir cada uno de los conceptos indispensables de manera jerárquica desde lo básico y general hasta lo más específico y particular; es decir, responder a todas esas incógnitas conceptuales como son ¿Qué es un seguro?, ¿Para qué sirve?, ¿Cómo opera los seguros en México?, ¿Cómo se clasifican?, ¿Qué tipo de seguro existen?, ¿Cómo funciona el mercado asegurador en México?, ¿Cuáles son los órganos reguladores?, ¿Bajo qué leyes se rigen las compañías aseguradoras?, ¿Qué característica tiene un contrato de seguro?, etc.

De tal manera que al obtener una clara noción de los elementos anteriores y ver su relación con la participación actuarial de manera teórica. Se cumple con el objetivo de este capítulo.

Capitulo II. Ajuste de curvas a funciones de distribución

2.1 INTRODUCCION A LOS CONCEPTOS DE PROBABILIDAD Y ESTADÍSTICA

“La palabra probabilidad (*probabilité*) se utiliza por primera vez para denotar algo medible, en la obra de Antonie Arnauld-Pierre Nicole. Este autor usa la palabra al considerar la diferencia que existe entre la posible ocurrencia de eventos.” (Calva, 2005) Esta diferencia la hacía de manera empírica. Al igual que Ar- nauld, se utilizará la palabra probabilidad para atribuir un grado de ocurrencia a determinados eventos.

Para obtener la probabilidad de un evento, lo correcto será expresarla de tal forma que pueda ser aplicable a diversas situaciones. No servirá de mucho decir que es probable ganar un juego, ya que lo que realmente se desea saber es qué tan probable lo es. Para facilitar las comparaciones, se deberá llevar el concepto a términos numéricos y para esto necesitamos alguna fórmula o método que nos arroje un número. Para lo cual, a continuación, se mostrarán 3 fórmulas para el cálculo de probabilidades que se desarrollaron en los inicios de esta teoría.

Galileo, Cardano, Pascal y Fermat, mostraron por qué algunos eventos ocurrirían con mayor frecuencia que otros, para lo cual contaban todos los posibles resultados del juego o experimento, y los comparaban con los que favorecían la ocurrencia del evento de interés, y es precisamente esta idea la que da origen a la primera fórmula que se muestra para el cálculo de probabilidades

$$P(\text{"Ocurrencia de una evento"}) = \frac{Casos\ favorables\ al\ evento}{Casos\ totales}$$

La fórmula anterior fue desarrollada y utilizada de manera empírica más de un siglo, pero fue hasta los principios del siglo XIX que Pierre Simon de Laplace, en su “Ensayo filosófico sobre las probabilidades”, sintetizo todo lo desarrollado en probabilidad y la enuncio. Esta fórmula se conoce como fórmula de probabilidad clásica o formula de Laplace. Un ejemplo clásico de cómo se aplica esta fórmula es en el lanzamiento de una moneda. Se puede intuir que al lanzar una moneda equilibrada la probabilidad de obtener sol es un medio.

$$P(\mathrm{sol}) = \frac{Casos\ favorables\ al\ evento}{Casos\ totales} = \frac{1}{2}$$

Analizando el cociente de casos favorables y totales, se puede deducir que el número de casos totales siempre será mayor o igual que el de casos favorables. Por lo tanto, el valor máximo de la probabilidad de un evento será 1 y el menor 0. Es decir, si A es un evento cualquiera entonces:

$$0 \leq P(A) \leq 1$$

Al utilizar la fórmula de probabilidad clásica se debe ser capaces de identificar y contar los posibles resultados del experimento, verificar que sea un número de posibilidades finito, y, además, suponer igual probabilidad de ocurrencia

Entre los muchos intentos por tener alguna ventaja al jugar con los dados había, o hay, personas que introducían bolitas de plomo en estos para favorecer a alguna de las caras.

Si se lanza un dado cargado, los posibles resultados son 1, 2, 3, 4, 5 y 6, pero ya no se puede asignar igual probabilidad a cada resultado, es decir, si la intención es conocer la probabilidad de obtener un “1”, no bastara con pensar en casos favorables y totales.

En caso de que el dado esté cargado, lo que se puede hacer para asignar probabilidades, por ejemplo, es lanzar el dado cien veces, y observar en cuantas de estas se obtuvo el 1 y asignar la probabilidad con la fórmula

$$P(\text{"}Obtener\ 1\text{"}) = \frac{x}{n}$$

Donde x es el número de veces que se obtuvo el 1, y n el número de veces que se lanzó el dado. Esta forma de asignar probabilidad nos debe resultar natural pues se usan en la vida cotidiana, por ejemplo, cuando se está esperando a una persona y esta suele llegar tarde, se puede suponer que llegará tarde por experiencias pasadas.

La asignación de probabilidades de acuerdo con el comportamiento a largo plazo en los experimentos es conocida como probabilidad frecuentista o teoría frecuentista.

Entre la teoría frecuentista y la clásica existe una relación. Se espera que si el número de ensayos que se realizan para obtener la probabilidad con el método frecuentista crece indefinidamente, la probabilidad que se obtenga para el evento será la misma que si se pudiera calcular mediante la fórmula clásica.

2.1.1 Independencia de dos sucesos

Uno de los principios involucrados en las formulas antes mencionadas es el de independencia de dos sucesos, cada evento de éxito no depende del evento anterior y se explicará a continuación.

"Dos sucesos S_1 y S_2 son independientes si $P\left(\frac{S_1}{S_2}\right) = P(S_1)$." (Hossack, I.B.; Pollard, J.H.; Zehnwirth, B, 2001) Esto es lógico, debido a que la igualdad implica que la probabilidad de que ocurra S_1 no viene afectada por la ocurrencia del suceso S_2 . Entonces se puede observar que, para dos sucesos independientes S_1 y S_2,

$$P(S_1 \cup S_2) = P(S_1) \cdot P(S_2)$$

La presencia de sucesos independientes generalmente simplifica la manera considerable el trabajo estadístico. Es costumbre, por ello, efectuar hipótesis de independencia, si las mismas son asumibles. En ocasiones, no es adecuado efectuar tal tipo de hipótesis. por ejemplo, un incendio desatado en un fabrica puede extenderse a una fábrica vecina eventualmente asegurada por el mismo asegurador.

Por ejemplo, si la probabilidad de que una casa sea robada es de 1 por cada 100, y la probabilidad de que una casa sea alcanzada por un rayo es de 1 por cada 10000, calcular la probabilidad de que una casa sea robada o alcanzada por un rayo o ambos sucesos. Representando el suceso “robo” por A y el suceso “alcanzada por un rayo “por B y suponiendo que ambos sucesos son independiente. Entonces:

$$(A \cap B) = P(A)\ x\ P(B)$$

$$= 0.01\ X\ 0.0001$$

$$= 0.000001$$

La probabilidad de que una casa sea robada o sea alcanzada por un rayo (o ambas circunstancias simultáneamente) se representa de la siguiente manera:

$$P(A\ \cup B) = P(A) + P(B) - P(A\ \cap B\ )$$

= 0.01 + 0.0001 – 0.000001

= 0.010099

El concepto de espacio muestral es indispensable en la teoría de probabilidades y que se mencionará constantemente. “El espacio muestral se define como el conjunto de todos los posibles resultados de un fenómeno o experimento a observar y es denotado con la letra griega Ω $(omega)$.” (Hossack, I.B.; Pollard, J.H.; Zehnwirth, B, 2001) Por ejemplo, al lanzar un dado se tiene:

$$\Omega = \{1,2,3,4,5,6\}$$

2.1.2 Variables aleatorias

Los experimentos se conciben de manera que los resultados del espacio muestral son cualitativos o cuantitativos. Como un ejemplo de resultado cualitativo se tiene el lanzamiento de una moneda es *"cara"* o *"cruz"*. Puede ser útil la cuantificación de los resultados cualitativos de un espacio muestral y, mediante el empleo de medidas numéricas, estudiar su comportamiento aleatorio. El concepto de variable aleatoria proporciona un medio para relacionar cualquier resultado con una medida cuantitativa.

Se comienza por definir: “Sea S un espacio muestral sobre el que se encuentra definida una función de probabilidad. Sea X una función de valor real definida sobre S, de manera que transforme los resultados de S en puntos sobre la recta de los reales. Se dice entonces que X es una variable aleatoria.” (Canavos, 1988)

Para ilustrar la noción de variable aleatoria, considérese el lanzamiento de una moneda. El espacio muestral está constituido por dos posibles resultados, *"cara"* y *"cruz".* Sea $X(cruz) = O$ y $X(cara) = 1$; de esta manera se han transformado los dos posibles resultados del espacio muestral en puntos sobre la recta de los reales.

Por $P(X = O)$ se entenderá la probabilidad de que la variable aleatoria tome el valor cero o, de manera equivalente, la probabilidad de que caiga cruz cuando se lance la moneda.

En el ejemplo anterior, el número de posibles valores de la variable aleatoria es finito. Sin embargo, se pueden definir variables aleatorias cuyos valores sean contables o no. Ya que una variable aleatoria es una caracterización cuantitativa de los resultados de un espacio muestral, ésta posee básicamente la naturaleza discreta o continua de este espacio.

Se determina qué una variable aleatoria X es discreta si el número de valores que puede tomar es contable (ya sea finito o infinito), y si éstos pueden arreglarse en una secuencia que corresponde con los enteros positivos.

Se establece qué una variable aleatoria X es continua si sus valores consisten en uno o más intervalos de la recta de los reales.

2.1.3 Distribuciones discretas y distribuciones continuas

En general, una variable aleatoria discreta X representa los resultados de un espacio muestral en forma tal que por $P(X = x)$ se entenderá la probabilidad de que X tome el valor de x. De esta forma, al considerar los valores de una variable aleatoria es posible desarrollar una función matemática que asigne una probabilidad a cada realización x de la variable aleatoria X. Esta función recibe el nombre de función de probabilidad de la variable aleatoria X. El término más general, distribución de probabilidad. Se refiere a la colección de valores de la variable aleatoria y a la distribución de probabilidades entre éstos. Sin embargo, hacer referencia a la distribución de probabilidad de X no sólo implica la existencia de la función de probabilidad, sino también la existencia de la función de distribución acumulativa de X.

Se define como: “Sea X una variable aleatoria discreta. Se llamará a $p(x) \equiv P(X = x)$ función de probabilidad de la variable aleatoria X, si satisface las siguientes propiedades:” (Canavos, 1988)

1. $p(x) \geq 0$ para todos los valores x de X
2. $\Sigma_x \, p(x) = 1.$

La función de distribución acumulativa de la variable aleatoria X es la probabilidad de que X sea menor o igual a un valor específico de x y está dada por:

$$F(x) \equiv P(X \leq x) = \sum_{x_i \leq x} p(x_i)$$

Por lo tanto, en el caso discreto, una variable aleatoria X está caracterizada por la función de probabilidad puntual $p(x)$, la cual determina la probabilidad puntual de que $X = x$, y por la función de distribución acumulativa $F(x)$, la que representa la suma de las probabilidades puntuales hasta el valor x de X inclusive. Nótese que las definiciones anteriores son consistentes con los axiomas de probabilidad, los cuales indican que esta función no es negativa para cualquier valor de la variable aleatoria y la suma de las probabilidades para todos los valores de X es igual a uno.

Considérese como ejemplo el lanzamiento de dos dados. Si X es la variable aleatoria que representa la suma de las caras, la función de probabilidad de X es

$$p(x) = \begin{cases} \dfrac{6 - |7 - x|}{36} & x = 2,3, \dots ,12, \\ 0 & \text{para cualquier otros caso} \end{cases}$$

Con la función anterior, pueden determinarse las probabilidades para varios valores de X. Además, puede evaluarse la función de distribución acumulativa de X de la siguiente forma:

$$F(1) \equiv P(X \leq 1) = 0$$
$$F(2) \equiv P(X \leq 2) = {}^{1}\!/_{36}$$
$$F(3) \equiv P(X \leq 3) = {}^{3}\!/_{36}$$
$$F(4) \equiv P(X \leq 4) = {}^{6}\!/_{36}$$
$$F(5) \equiv P(X \leq 5) = {}^{10}\!/_{36}$$
$$F(6) \equiv P(X \leq 6) = {}^{15}\!/_{36}$$
$$F(7) \equiv P(X \leq 7) = {}^{21}\!/_{36}$$
$$F(8) \equiv P(X \leq 8) = {}^{26}\!/_{36}$$
$$F(9) \equiv P(X \leq 9) = {}^{30}\!/_{36}$$
$$F(10) \equiv P(X \leq 10) = {}^{33}\!/_{36}$$
$$F(11) \equiv P(X \leq 11) = {}^{35}\!/_{36}$$
$$F(12) \equiv P(X \leq 12) = 1$$

Nótese que

$$P(X > 7) = 1 - P(X \leq 7) = 1 - F(7) = {}^{15}\!/_{36}$$
$$P(X = 7) = P(X \leq 7) - P(X \leq 6) = F(7) - F(6) = {}^{6}\!/_{36}$$

$$P(5 \leq X \leq 9) = P(X \leq 9) - P(X \leq 4) = F(9) - F(4) = {}^{24}\!/_{36}$$

En general, "la función de distribución acumulativa $F(x)$ de una variable aleatoria discreta es una función no decreciente de los valores de X, de tal manera que:" (Canavos, 1988)

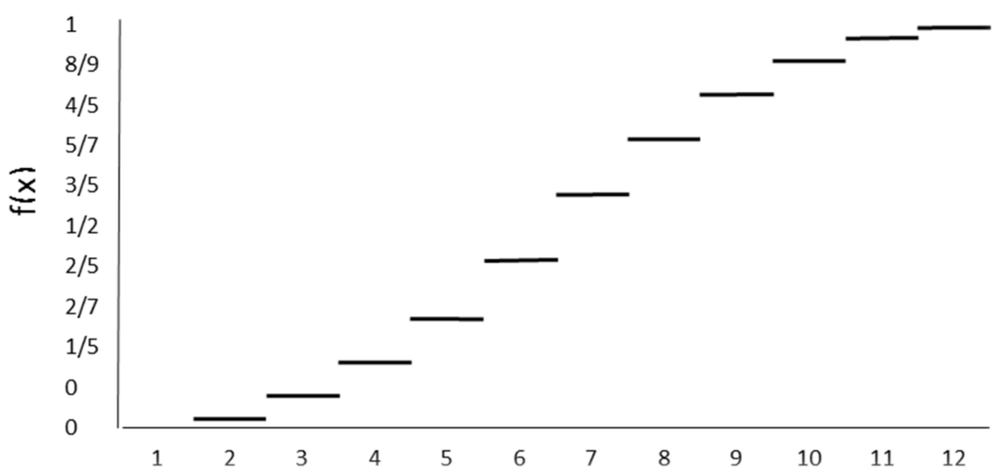
1. $0 \leq F(x) \leq 1$ para cualquier x
2. $F(x_i) \geq f(x_j) \quad si\ x_i \geq x_j$
3. $P(X > x) = 1 - F(x)$

Además, puede establecerse que para variables aleatorias de valor entero se tiene que:

4. $P(X = x) = F(x) - F(x-1)$
5. $P(x_i \leq X \leq x_j) = F(x_j) - F(x_i - 1)$

En la siguiente grafica es evidente que la función de distribución acumulativa de una variable aleatoria discreta, es una función escalón, que toma un valor superior en cada salto.

Función de distribución acumulativa discreta



Fuente: Canavos, G. (1988). Probabilidad y Estadística. Aplicaciones y métodos. Estado de México

En el caso discreto, se asignan probabilidades positivas a todos los valores puntuales de la variable aleatoria, pero la suma de todas ellas es uno aún a pesar de que el conjunto de valores sea infinito contable. Para el caso continuo, lo anterior no es posible.

Por esta razón, la probabilidad de que una variable aleatoria continua X tome un valor específico x es cero.

Se ilustrará el sentido de este resultado mediante el siguiente ejemplo: supóngase que se observa el intervalo entre dos llegadas consecutivas a un servicio. Si el dispositivo de medición puede medir el tiempo basta una décima de segundo, entonces un intervalo de 83.4 segundos puede realmente tomarse como la media y el verdadero valor puede encontrarse entre 83.35 y 83.45 segundos, Por lo tanto, en el caso continuo es más lógico visualizar las probabilidades de intervalos que de puntos en particular.

La distribución de probabilidad de una variable aleatoria continua X está caracterizada por una función $f(x)$ que recibe el nombre defunción de densidad de probabilidad. Esta función $f(x)$ no es la misma función de probabilidad que para el caso discreto. Como existe la probabilidad de que X tome el valor específico x es cero, la función de densidad de probabilidad no representa la probabilidad de que $X = x$.

Más bien, ésta proporciona un medio para determinar la probabilidad de un intervalo $a \leq X \leq b$.

Para ilustrar lo que se entiende como función de densidad de probabilidad, supóngase que se miden los tiempos, entre dos llegadas consecutivas, de 100 clientes a una tienda y se agrupan en diez intervalos de un minuto cada uno, como se muestra en la siguiente. En este punto se grafican las frecuencias relativas para cada intervalo por medio de rectángulos, como se muestra en la gráfica de tipo histograma posteriormente, para indicar que la frecuencia se refiere al intervalo completo más que a un punto en particular del mismo. Nótese que, puesto que la base tiene una longitud igual a uno, el área de cada rectángulo es la frecuencia relativa del correspondiente intervalo y, por lo tanto, la suma de las áreas de todos los rectángulos es igual a uno.

Tabla de frecuencias de distribución continua

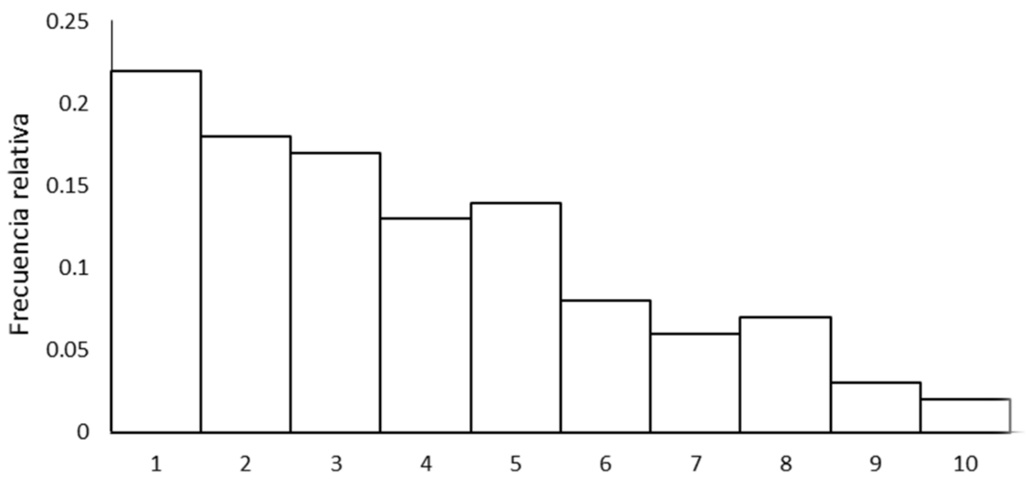
Intervalo	*Número de llegadas*	*Frecuencia relativa*
$0 < x \leq 1$	22	0.22
$0 < x \leq 2$	18	0.18
$0 < x \leq 3$	17	0.17
$0 < x \leq 4$	13	0.13
$0 < x \leq 5$	14	0.14
$0 < x \leq 6$	8	0.08
$0 < x \leq 7$	6	0.06
$0 < x \leq 8$	7	0.07
$0 < x \leq 9$	3	0.03
$0 < x \leq 10$	2	0.02

Fuente: Canavos, G. (1988). Probabilidad y Estadística. Aplicaciones y métodos. Estado de México

Supóngase que en lugar de observar los tiempos entre dos llegadas consecutivas de 100 clientes. Se observan los tiempos para 1 000 clientes y se agrupan en 20 intervalos de medio minuto cada uno; o bien pueden observarse los tiempos para 10 000 clientes agrupándolos en 40 intervalos de 15 segundos cada uno. Cada vez que esto se hace. Se produce un histograma que es cada vez menos irregular. Pero en el que la frecuencia sigue siendo prácticamente la misma. Al continuar' este proceso de aumento del número de observaciones mientras se disminuye la amplitud de los intervalos de clase. Se llegará a una curva límite. Esto es. Cuando el número observado de tiempos. Entre dos llegadas consecutivas. Sea muy grande y la amplitud de los intervalos de clase sea muy pequeña, la frecuencia relativa aparecerá. En esencia, como una curva lisa. Con base a la gráfica de tipo histograma, puede especularse que la curva límite para este ejemplo es la que se muestra en la gráfica exponencial.

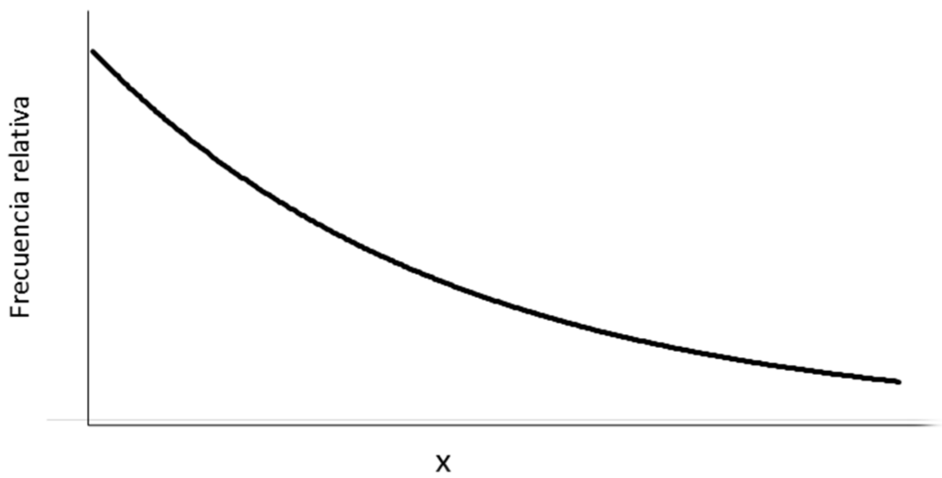
La función $j(x)$, cuya gráfica es la curva límite que se obtiene para un número muy grande de observaciones y para una amplitud de intervalo muy pequeña, es la función de densidad de probabilidad para una variable aleatoria continua X. Ya que la escala vertical se elige de manera que el área total bajo la curva es igual a uno.

Histograma de la función de distribución continua



Fuente: Canavos, G. (1988). Probabilidad y Estadística. Aplicaciones y métodos. Estado de México

Representación gráfica de la tendencia de una distribución continua



Fuente: Canavos, G. (1988). Probabilidad y Estadística. Aplicaciones y métodos. Estado de México

La función de densidad de probabilidad de una variable aleatoria continua X se define formalmente de la siguiente manera:

"Si existe una función $f(x)$ tal que

1. $f(x) \geq 0, -\infty < x < \infty$

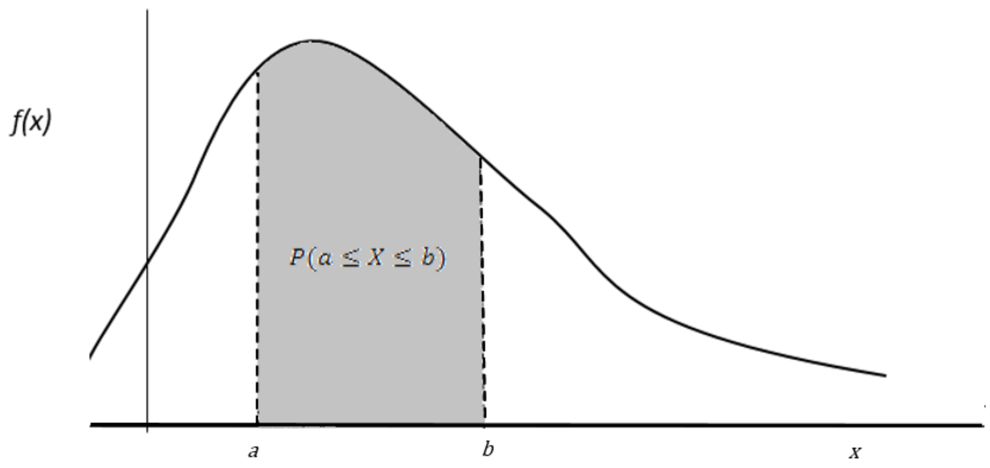
2. $\int_{-\infty}^{\infty} f(x)dx = 1,\ \ y$

3. $P(a \leq x \leq b) = \int_{a}^{b} f(x)dx$

Para cual cualquier *a* y *b*, entonces $f(x)$ es la función de densidad de probabilidad de la variable aleatoria continua X.” (Canavos, 1988)

Puesto que área total bajo f(x) es uno, la probabilidad del intervalo $a \leq X \leq b$ es el área acotada por la función de densidad y las rectas $X = a$ y $X = b$, como se muestra a continuación

Representación gráfica de la función de probabilidad continua



Fuente: Canavos, G. (1988). Probabilidad y Estadística. Aplicaciones y métodos. Estado de México.

Al igual que en el caso de una variable discreta, la función de distribución acumulativa de una variable aleatoria continua X es la probabilidad de que X tome un valor menor o igual a algún x específico .esto es:

$$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)dt$$

En donde t es una variable artificial de integración .por lo tanto, la función de distribución acumulativa $F(x)$ es el área acotada por la función de densidad que se encuentra a la izquierda de la recta $X = x$, como se muestra en la gráfica a continuación.

Dado que para cualquier variable aleatoria continua X.

$$P(X = x) \int_x^x f(t)dt = 0$$

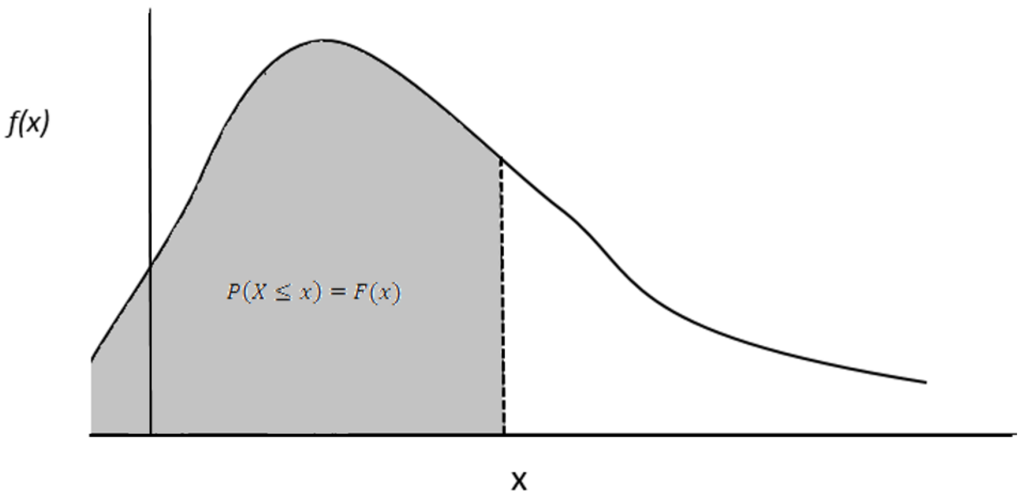
Entonces

$$P(X \leq x) = P(X < x) = F(x)$$

"La distribución acumulativa F(x), es una función lisa no decreciente de los valores de la variable aleatoria con las siguientes propiedades:" (Canavos, 1988)

1. $F(-\infty) = 0$
2. $F(\infty) = 1$
3. $P(a < x < b) = F(b) - f(a)$
4. $dF(x) \,/\, dx = f(x)$

Representación gráfica de la función de distribución acumulativa continua



Fuente: Canavos, G. (1988). Probabilidad y Estadística. Aplicaciones y métodos. Estado de México

La propiedad de que la derivada de la función de distribución acumulativa es la función de densidad de probabilidad, es una consecuencia del teorema fundamental del cálculo integral.

2.1.4 Medidas de tendencia central y dispersión de datos estadísticos.

Al realizar un estudio de investigación, se pretende generalmente inferir o generalizar resultados de una muestra a una población. Se estudia en particular a un reducido número de individuos a los que se tienen acceso con la idea de poder generalizar los hallazgos a la población de la cual esa muestra procede. Este proceso de inferencia se efectúa por medio de métodos estadísticos basados en la probabilidad.

"La población representa el conjunto grande de individuos que se desean estudiar y generalmente suele ser inaccesible." (Belinchón, 2007) Es decir, en definitiva, un colectivo homogéneo que reúne unas características determinadas.

La muestra es el conjunto menor de individuos (subconjunto de la población accesible y limitado sobre el que se realizan las mediciones o el experimento con la idea de obtener conclusiones generalizables a la población). El individuo es cada uno de los componentes de la población y la muestra. La muestra debe ser representativa de la población y con ello se pretende decir que cualquier individuo de la población en estudio debe haber tenido la misma probabilidad de ser elegido.

Las razones para estudiar muestras en lugar de poblaciones son diversas y entre ellas se señalan (Belinchón, 2007):

1. Ahorrar tiempo. Estudiar a menos individuos es evidente que lleva menos tiempo.
2. Como consecuencia del punto anterior ahorraremos costes.
3. Estudiar la totalidad de los pacientes o personas con una característica determinada en muchas ocasiones puede ser una tarea inaccesible o imposible de realizar.
4. Aumentar la calidad del estudio. Al disponer de más tiempo y recursos, las observaciones y mediciones realizadas a un reducido número de individuos pueden ser más exactas y plurales que si se tuvieran que realizar a una población.

5. La selección de muestras específicas nos permitirá reducir la heterogeneidad de una población al indicar los criterios de inclusión y/o exclusión.

El objeto de estudio en cada individuo de la muestra son las variables (edad, sexo, peso, talla, tensión arterial sistólica, etcétera). Los datos son los valores que toma la variable en cada caso. Lo que se va a realizar es una medición, es decir, asignar valores a las variables incluidas en el estudio. Además, se debe determinar la escala de medida que se aplicará a cada variable.

La naturaleza de las observaciones será de gran importancia a la hora de elegir el método estadístico más apropiado para abordar su análisis. Con este fin, se clasifican las variables, a grandes rasgos, en dos tipos variables:

Las Variables cuantitativas son las variables que pueden medirse, cuantificarse o expresarse numéricamente. Las variables cuantitativas pueden ser de dos tipos:

Variables cuantitativas continuas, si admiten tomar cualquier valor dentro de un rango numérico determinado (edad, peso, talla).

Variables cuantitativas discretas, si no admiten todos los valores intermedios en un rango. Suelen tomar solamente valores enteros (número de hijos, número de partos, número de hermanos, etc.).

Variables cualitativas. Este tipo de variables representan una cualidad o atributo que clasifica a cada caso en una de varias categorías. La situación más sencilla es aquella en la que se clasifica cada caso en uno de dos grupos (hombre/mujer, enfermo/sano, fumador/no fumador). Son datos dicotómicos o binarios. Como resulta obvio, en muchas ocasiones este tipo de clasificación no es suficiente y se requiere de un mayor número de categorías (color de los ojos, grupo sanguíneo, profesión, etcétera).

En el proceso de medición de estas variables, se pueden utilizar dos escalas:

Escalas nominales: ésta es una forma de observar o medir en la que los datos se ajustan por categorías que no mantienen una relación de orden entre sí (color de los ojos, sexo, profesión, presencia o ausencia de un factor de riesgo o enfermedad, etcétera).

Escalas ordinales: en las escalas utilizadas, existe un cierto orden o jerarquía entre las categorías (grados de disnea, estadiaje de un tumor, etcétera).

Una vez que se han recogido los valores que toman las variables de nuestro estudio (datos), se procederá al análisis descriptivo de los mismos. Para variables categóricas, como el sexo o la estatura, se quiere conocer el número de casos en cada una de las categorías, reflejando habitualmente el porcentaje que representan del total, y expresándolo en una tabla de frecuencias.

Para variables numéricas, en las que puede haber un gran número de valores observados distintos, se ha de optar por un método de análisis distinto, es oportuno mencionar que dichas variables cuantitativas serán las utilizadas a lo largo de este capítulo.

Para comenzar a determinar la frecuencia en la que los datos observados están, el primer paso a seguir es organizar los datos; la cual consiste en agrupar aquellos datos que se repiten varias veces. La manera adecuada de trabajar los datos es la construcción de tablas de frecuencia, para ello es necesario definir sus elementos.

Frecuencia absoluta (n_i): es el número de veces que se repite un determinado valor (x_i) de la variable. La suma de todas las frecuencias absolutas es igual al tamaño muestral.

Este tipo de frecuencias no son comparables con las obtenidas en otras muestras de distinto tamaño.

Frecuencia relativa (f_i): es igual a la frecuencia absoluta dividida por el número total de datos, es decir por el tamaño muestral

$$f_i = \frac{n_i}{n}.$$

La suma de todas las frecuencias relativas es igual a la unidad.

Frecuencia acumulada (N_i): Nos dice el número de datos que hay igual o inferiores a uno

$$N_i = \sum_{j=1}^{i} N_i + n_i$$

La última frecuencia acumulada absoluta es el tamaño muestral.

Frecuencia relativa acumulada (F_i): Es el resultado de dividir cada frecuencia acumulada por el número total de datos

$$F_i = \frac{N_i}{n} = \sum_{j=1}^{i} f_j$$

La última frecuencia relativa acumulada es la unidad.

Se le nombra tabla de frecuencias a la tabla que contiene al conjunto de diferentes valores que ha tomado una variable (los datos sin repetir), ordenados de menor a mayor con sus correspondientes frecuencias.

Tomando un caso sencillo .se desea saber cuántos carros están asegurados en una colonia, clasificándolos por colores (en las aseguradoras el color del auto es un dato a considerar para estudios estadísticos)

$$x_1 = Azul$$

$$x_2 = Blanco$$

$$x_3 = Negro$$

$$x_4 = Gris$$

$$x_5 = Rojo$$

$$x_6 = Amarillo$$

$$x_0 = otro\ color$$

Con la información anterior se construye la siguiente tabla de frecuencias

Tabla de frecuencias

x_i	n_i	f_i	N_i	F_i
0	2	0.04	2	0.04
1	4	0.08	6	0.12
2	21	0.42	27	0.54
3	15	0.3	42	0.84
4	6	0.12	48	0.96
5	1	0.02	49	0.98
6	1	0.02	50	1

Fuente: Nania, L. (2003). Estadística aplicada a la hidrología. Granada: Universidad de Granada. España.

¿Cuál es el número de autos asegurados con colores deportivos (rojo y amarillo)?

Observando la tabla anterior se resolverá la pregunta de la siguiente manera

$$AUTOD = n_5 + n_6 = 1 + 1 = 2$$

¿Cuál es el número de autos asegurados con colores no deportivos?

$$AUTONOD = \; n_0 + n_1 + n_2 + n_3 + n_4 = 2 + 4 + 21 + 15 + 6 = 48$$

Otra manera de obtener este resultado es usando el complemento

$$AUTONOD \leq n_4 = N_6 - AUTOD$$

$$AUTONOD = 50 - 2 = 48$$

¿Qué porcentaje de autos de color deportivo están asegurados?

$$\%AUTOD = f_5 + f_6 = 0.02 + 0.02 = 0.04$$

De igual manera

$$\%AUTOD = F_6 - (\%AUTONOD \leq f_4)$$

$$\%AUTOD = 1 - (0.04 + 0.08 + 0.42 + 0.3 + 0.12)$$

$$\%AUTOD = 1 - 0.96 = 0.04$$

Existen ocasiones que el rango de los valores distintos que toman la variable es enorme, es decir N es demasiado grande. En el caso en que la variable a estudiar sea continua. La solución es agrupar los diferentes valores de la variable en intervalos de clase. Teniendo en cuenta que lo que se gana en manejabilidad se perderá en información, con lo que los resultados serán aproximados.

Agrupar en intervalos de clase consiste en agrupar los datos en un número relativamente pequeño de intervalos que cumplan:

Prescindir que exista ambigüedad con respecto a la clase a que pertenece una observación particular; es decir determinar bien los intervalos para evita duplicar o discriminar datos

Cubran todo el rango de valores que se tienen en la muestra.

Para crear los intervalos es necesario definir el proceso:

- A las fronteras del intervalo, límites inferior y superior de la clase se denotará como L_{i-1} , L_i.
- Marca de clase (c_i) será el punto medio del intervalo, es decir, al promedio aritmético entre el límite inferior y superior:

$$C_i = \frac{L_i + L_{i-1}}{2}$$

Dicho valor será tomado como representativo.

- Amplitud (a_i) a la diferencia entre el extremo superior e inferior:

$$a_i = L_i - L_{i-1}.$$

- Al número de observaciones de una clase se le llama frecuencia de clase (n_I), si se divide esta frecuencia por el número total de observaciones, se llama frecuencia relativa de clase (f_I), y del mismo modo que lo se hace para datos sin agrupar se definirían N_I, y F_I.

Para construir una distribución de frecuencias agrupada en intervalos se realiza lo siguiente:

1. Se Empezará determinando el recorrido de la variable o rango de valores que se tienen en la muestra. Se define como la diferencia entre el mayor y el menor valor de la variable. $Recorrido = x_k - x_1$

2. Asignar número de clases: depende del tamaño de la muestra. Para muestras de tamaño moderado, n <50, se suele elegir un número de clases igual a$\sqrt{n}$, o bien se usa la fórmula de Sturtges, (se toma el resultado de calcular el logaritmo de n, dividir por el Logaritmo de 2 y sumar 1:

$$m = \frac{\log(n)}{\log(2)} + 1;$$

En general el número de clases no debe sobrepasar de 15 ó 20, en caso de muestras grandes.

3. Se determina la amplitud de lo intervalos. Es más cómodo que la amplitud de todas las clases sea la misma (siempre que sea posible), si es

$$a_i = \frac{Recorrido}{n^a\ de\ intervalos}$$

Se tomará como regla, a no ser que se indique lo contrario, utilizar el intervalo cerrado por la izquierda y abierto por la derecha.

Para ejemplificarse tiene que, el área de estadística de una aseguradora tiene como estudio el número y monto de siniestros del seguro de incendio para futuros análisis y toma de decisiones por lo que le es conveniente elaborar una tabla de frecuencias por intervalos .El monto menor de siniestros pagados del años 2000 es 3.3 unidades monetarias y el mayor 6.1 unidades monetarias, la diferencia es 2.8 unidades monetarias y por tanto $Recorrido\ =\ 2.8$.

La cantidad total de siniestros ocurridos durante el mismo fue de 40; por lo tanto $N = 40$, de tal manera se asignaran $6\,clases$.

Por lo anterior la amplitud la es el cociente de $a = 2.8/6 = 0.467$.

Como la amplitud es un número con muchos decimales, sería complicado la asignación de los datos en cada clase, por lo que se puede usar la siguiente técnica que facilitará dicha asignación: Para que los intervalos resulten con amplitud de 0.5 unidades monetarias se tomará como primer valor 3.25 en vez de 3.3 y como último 6.25 en vez de 6.1 de esta manera: $Recorrido = 6.25 - 3.25 = 3$ y $amplitud =\ 3/6 = 0.5$, respetando el número de clases y eliminar que la amplitud considere muchos decimales se puede ampliar el recorrido de manera que se ajuste al número de clases

Una posible construcción de la tabla de frecuencias por intervalos sería la siguiente:

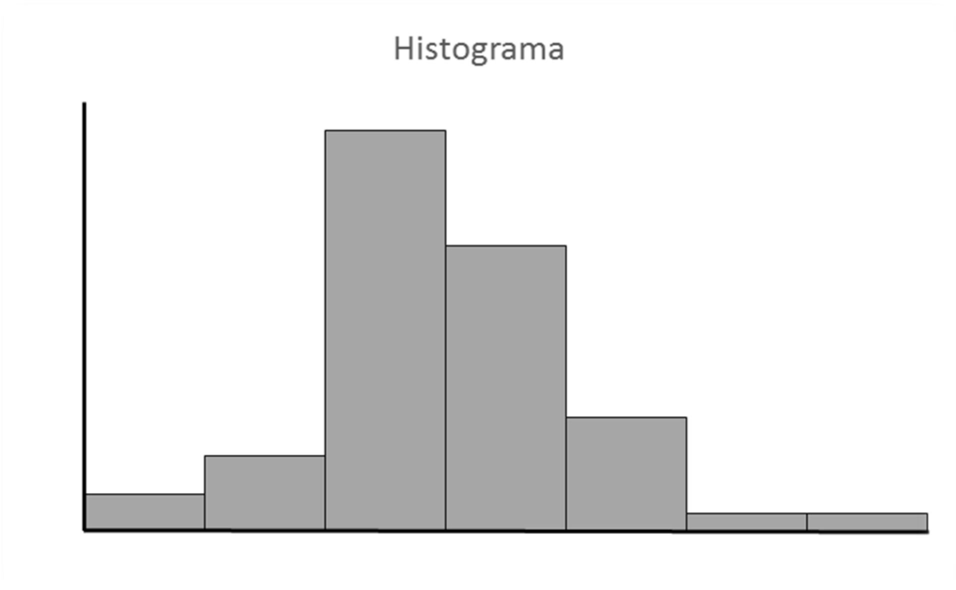
Tabla de frecuencias por intervalos

(L_{i-1}, L_i)	c_i	n_i	f_i	N_i	F_i
(3.25-3.75)	3.5	3	0.075	3	0.075
(3.75-4.25)	4	8	0.2	11	0.275
(4.25-4.75)	4.5	14	0.35	25	0.625
(4.75-5.25)	5	6	0.15	31	0.775
(5.25-5.75)	5.5	4	0.1	35	0.875
(5.75-6.25)	6	5	0.125	40	1

Fuente: Nania, L. (2003). Estadística aplicada a la hidrología. Granada: Universidad de Granada. España.

La forma de la distribución de frecuencias se percibe más rápidamente y quizás se retiene durante más tiempo en la memoria si se representa gráficamente.

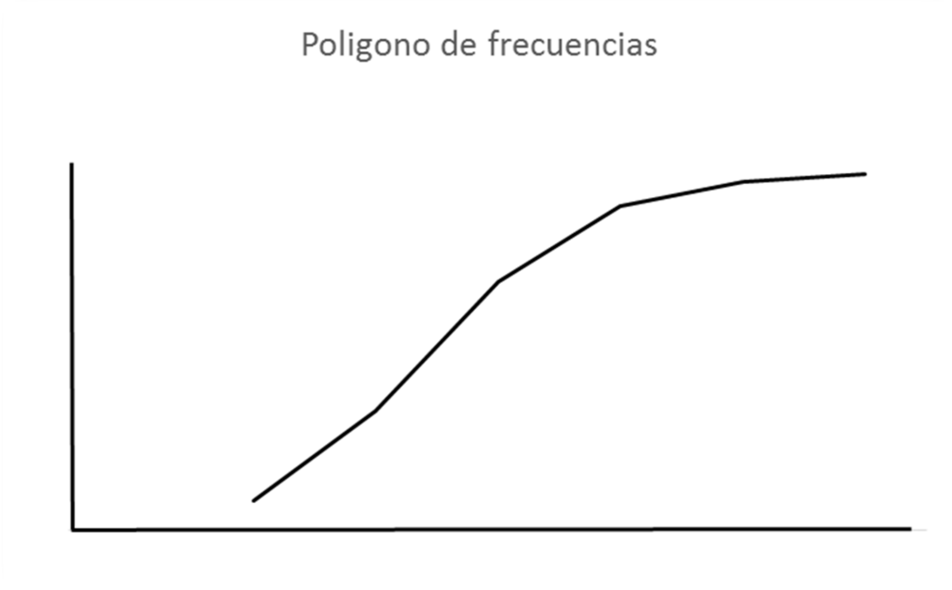
*histograma***:** Es la representación gráfica equivalente al diagrama de barras para datos agrupados, en el eje de ordenadas representarnos las clases y levantarnos sobre cada clase rectángulos unidos entre sí de altura igual a la frecuencia de la clase (absolutas o relativas)



Fuente: Nania, L. (2003). Estadística aplicada a la hidrología. Granada: Universidad de Granada. España.

El histograma o diagrama de barras proporcionan mucha información respecto a la estructura de los datos (y si la muestra es representativa de la población, respecto a la estructura de la población): el valor central de la distribución, su dispersión y la forma de la distribución. Cuando se encuentran distribuciones donde los intervalos no tienen la misma amplitud, las barras del histograma tienen un área proporcional a la frecuencia que se quieran representar.

Polígono de frecuencias: Es la representación habitual para datos cuantitativos agrupados de las frecuencias acumuladas (absolutas o relativas), mediante puntos se representan las frecuencias en el eje de ordenadas y la marca de clase en el de abscisas. Después se unen estos puntos por trozos de rectas.



Fuente: Nania, L. (2003). Estadística aplicada a la hidrología. Granada: Universidad de Granada. España.

Las medidas tendencia central se utilizan para encontrar un valor que represente a todos los datos. La medida más evidente que calcular para describir un conjunto de observaciones numéricas es su valor medio. La *media* no es más que la suma de todos los valores de una variable dividida entre el número total de datos de los que se dispone.

Formalmente, se denota como $(X_1, X_2, \ldots, X_n)$ a los n datos recogidos de la variable en cuestión, el valor medio vendrá dado por:

$$Media(\bar{X}) = \sum_{i=1}^{n} \frac{x_i}{n} = \mu$$

Otra medida de tendencia central que se utiliza habitualmente es la mediana. *La mediana* (Me) es el valor que separa por la mitad las observaciones ordenadas de menor a mayor, de tal forma que el 50% de estas son menores que la mediana y el otro 50% son mayores. Si el número de datos es impar la mediana será el valor central, si es par se tomará como mediana la media aritmética de los dos valores centrales. Es la observación equidistante de los extremos.

Tomando como ejemplo los siguientes números ordenados 15, 21, 32, 59, 60, 60, 61, 64, 71, 80. El número de observaciones es par (10 individuos), los dos valores que se encuentran en el medio son 60 y 60. Si se realiza el cálculo de la media de estos dos valores nos dará a su vez 60, que es el valor de la mediana.

Si la *media* y la *mediana* son iguales, la distribución de la variable es simétrica. La media es muy sensible a la variación de las puntuaciones. Sin embargo, la mediana es menos sensible a dichos cambios.

En el caso de datos agrupados, La mediana se encuentra en el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas, es decir se tienen que buscar el intervalo en el que se encuentre $N/_2$.

$$Me = L_i + \frac{N/_2 - F_{i-1}}{f_i} + a_i$$

L_i es el límite inferior de la clase donde se encuentra la mediana.

$N/_2$ es la semisuma de las frecuencias absolutas.

F_{I-1} es la frecuencia acumulada anterior a la clase mediana.

a_i es la amplitud de la clase.

La *mediana* es independiente de las amplitudes de los intervalos.

Por último, otra medida de tendencia central, no tan usual como las anteriores, es la *moda* (Mo), siendo éste el valor de la variable que presenta una mayor frecuencia.; es decir la moda es el valor que más se repite. Puede suceder que haya más de una moda o ninguna (si todos los valores tienen igual frecuencia).

Utilizando el ejemplo anterior de las 10 observaciones ordenadas el valor que más se repite es 60, que es la moda.

De igual manera que en la mediana, en el caso de datos agrupados, la *moda* se obtiene diferente con la siguiente formula:

$$Mo = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} + a_i$$

L_i es el límite inferior de la clase modal.

$N/_2$ es la semisuma de las frecuencias absolutas.

F_I es la frecuencia absoluta de la clase modal.

F_{I-1} es la frecuencia absoluta inmediatamente inferior a la clase modal.

F_{I+1} es la frecuencia absoluta inmediatamente posterior a la clase modal.

a_i es la amplitud de la clase.

Otro aspecto a tener en cuenta al describir datos continuos es la dispersión de los mismos. Existen distintas formas de cuantificar esa variabilidad. De todas ellas, la *varianza* (σ^2) de los datos es la más utilizada. Es la media de los cuadrados de las diferencias entre cada valor de la variable y la media aritmética de la distribución.

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

Esta varianza muestral se obtiene como la suma de las diferencias de cuadrados y por tanto tiene como unidades de medida el cuadrado de las unidades de medida en que se mide la variable estudiada.

En el mismo ejemplo anterior de los190 datos ordenados la varianza sería:

$$\sigma^2 = \frac{(15-52.3)^2 + (21-52.3)^2 + \cdots (80-52.3)^2}{10} = 427.61$$

Cabe recordar que para este ejemplo la media es 52.3

$$\bar{X} = \frac{15+21+32+\cdots+80}{10} = 52.3$$

La *desviación típica* o *desviación estándar* (σ) es la raíz cuadrada de la varianza. Expresa la dispersión de la distribución y se expresa en las mismas unidades de medida de la variable. La desviación típica es la medida de dispersión más utilizada en estadística.

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

Aunque esta fórmula de la desviación típica muestral es correcta, en la práctica, la estadística nos interesa para realizar inferencias poblacionales, por lo que en el denominador se utiliza, en lugar de n, el valor $n - 1$.

Por tanto, la medida que se utiliza es la cuasidesviación típica, dada por:

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

Aunque en muchos contextos se utiliza el término de desviación típica para referirse a ambas expresiones.

En los cálculos del ejercicio previo, la desviación típica muestral, que tiene como denominador n, el valor sería 20.678.

$$\sigma = \sqrt{\frac{(15-52.3)^2+(21-52.3)^2+\cdots(80-52.3)^2}{10}} = 20.678$$

A efectos de cálculo se hará como $n-1$ y el resultado seria 21,79.

El haber cambiado el denominador de n por $n-1$ está en relación al hecho de que esta segunda fórmula es una estimación más precisa de la desviación estándar verdadera de la población y posee las propiedades que se necesitan para realizar inferencias a la población.

Cuando se quieren señalar valores extremos en una distribución de datos, se suele utilizar la amplitud como medida de dispersión. La amplitud es la diferencia entre el valor mayor y el menor de la distribución.

Por ejemplo, utilizando los datos del ejemplo previo se tiende 80-15 =65.

Como medidas de variabilidad más importantes, conviene destacar algunas características de la varianza y desviación típica:

- Son índices que describen la variabilidad o dispersión y por tanto cuando los datos están muy alejados de la media, el numerador de sus fórmulas será grande y la varianza y la desviación típica lo serán.
- Al aumentar el tamaño de la muestra, disminuye la varianza y la desviación típica. Para reducir a la mitad la desviación típica, la muestra se tiene que multiplicar por 4.
- Cuando todos los datos de la distribución son iguales, la varianza y la desviación típica son iguales a 0.
- Para su cálculo se utilizan todos los datos de la distribución; por tanto, cualquier cambio de valor será detectado.

Otra medida que se suele utilizar es el coeficiente de variación (CV). Es una medida de dispersión relativa de los datos y se calcula dividiendo la desviación típica muestral por la media y multiplicando el cociente por 100. Su utilidad estriba en que nos permite comparar la dispersión o variabilidad de dos o más grupos. Así, por ejemplo, si se asume el reporte de 5 siniestros de del seguro de responsabilidad civil n la zona centro (70, 60, 56, 83 y 79 unidades monetarias) cuya media es de 69.6 unidades monetarias y su desviación típica = 10.44 y el monto del seguro de incendio de los mismos siniestros (150, 170, 135, 180 y 195 unidades monetarias) cuya media es de 166 unidades monetarias y su desviación típica de 21.3. La pregunta sería: ¿qué distribución es más dispersa, el peso o la tensión arterial? Si se comparan las desviaciones típicas se observa que la desviación típica de la tensión arterial es mucho mayor; sin embargo, no se puede comparar dos variables que tienen escalas de medidas diferentes, por lo que se calculan los coeficientes de variación:

Para calcular el coeficiente de variación se utiliza la siguiente formula:

$$CV = \frac{\sigma}{\bar{X}}$$

Regresando al ejemplo anterior:

$$CV\ del\ seguro\ de\ RC = \frac{10.44}{69.6} = 15\%$$

$$CV\ del\ seguro\ de\ incendio = \frac{21.30}{166} = 12.8\%$$

A la vista de los resultados, se observa que los montos de los siniestros del seguro de RC tienen mayor dispersión a comparación que el seguro de incendio.

2.1.5 Cuantiles

Cuando los datos se distribuyen de forma simétrica (esto ocurre cuando los valores de su media y mediana están próximos), se usan para describir esa variable su media y desviación típica. En el caso de distribuciones asimétricas, la mediana y la amplitud son medidas más adecuadas. En este caso, se suelen utilizar además los cuantiles. Los cuantiles no son medidas de tendencia central sino medidas de posición.

Los cuantiles son valores de la distribución que la dividen en partes iguales, es decir, en intervalos, que comprenden el mismo número de valores. Los más usados son los cuartiles, los deciles y los percentiles

El percentil es el valor de la variable que indica el porcentaje de una distribución que es igual o menor a esa cifra.

Los percentiles son 99 valores que dividen en cien partes iguales el conjunto de datos ordenados.

El *percentil* de orden (P_p) es el menor valor superior al $p\%$ de los datos (ordenados de menor a mayor los datos, deja el $p\%$ de datos por delante). La forma más cómoda de calcularlos es a partir de las frecuencias acumuladas:

Para las distribuciones no agrupadas el percentil p es aquel valor cuya frecuencia acumulada más se acerca por arriba al $p\%$ de n, es decir:

$$P_p = X_i \ \ tal\ que \ \ N_{i-1} < \frac{p_n}{100} N_i$$

En las distribuciones agrupadas se parte de la misma idea que cuando se calculó la mediana, se busca en primer lugar el intervalo (L_{i-1}, L_i) cuya frecuencia acumulada sea $N_{i-1} < \frac{P_n}{100} \leq N_i$, a continuación para hallar el percentil se aplica la siguiente fórmula:

$$P_p = L_{i-1} + \frac{\left(\frac{p_n}{100} - N_{i-1}\right) a_i}{n_i}$$

Los cuartiles (C_i): son los tres valores que dividen al conjunto de datos ordenados en cuatro partes iguales, son un caso particular de los percentiles:

$$C_1 = P_{25}$$

$$C_2 = P_{50}$$

$$C_3 = P_{75}$$

Los *deciles* (D_i): Son los nueve valores que dividen al conjunto de datos ordenados en diez partes iguales, son también un caso particular de los percentiles.

$$D_1 = P_{25}$$

$$D_2 = P_{20} \ \ldots$$

$$D_9 = P_{90}$$

La Mediana también es un caso particular de percentil: Me=$P_{50.}$

2.1.6 Esperanza (Valor esperado)

El valor esperado (o esperanza) de una variable aleatoria es un concepto muy importante en el estudio de las distribuciones de probabilidad. La esperanza de una variable aleatoria tiene sus orígenes en los juegos de azar, debido a que los apostadores deseaban saber cuál era su esperanza de ganar repetidamente un juego. En este sentido, el valor esperado representa la cantidad de dinero promedio que el jugador está dispuesto a ganar o perder después de un número muy grande de apuestas. Este significado también es válido para una variable aleatoria. Es decir, el valor promedio de una variable aleatoria después de un número grande de experimentos, es su valor esperado.

Para ilustrar la esencia de la esperanza, se analizará el siguiente juego de azar.

Supóngase que se tiene moneda normal y el jugador tiene tres oportunidades para que al lanzarla aparezca una "cara". El juego termina en el momento en el que cae una "cara” o después de tres intentos, lo que suceda primero. Si en el primero, segundo o tercer lanzamiento aparece "cara” el jugador recibe 2 unidades monetarias, 4 unidades monetarias, y 8 unidades monetarias respectivamente. Si no cae "cara" en ninguno de los tres lanzamientos, pierde 20 unidades monetarias. Para determinar la ganancia o pérdida, promedio después de un número muy grande de juegos, sea X la variable aleatoria que representa la cantidad que se gana o se pierde cada vez que se juega. Los posibles valores de X junto con sus respectivas probabilidades se encuentran en la tabla siguiente.

Tabla del ejemplo de valor esperado

X	$P(X)$
2	$P(X=2)=\frac{1}{2}$
3	$P(X=4)=\frac{1}{4}$
8	$P(X=8)=\frac{1}{8}$
-20	$P(X=-20)=\frac{1}{8}$

Fuentes: Wackerly, D., Mendenhall, W., & Scheaffer, R. (2009). Estadística matemática con aplicaciones.

Después de un número grande de juegos se espera ganar 2 unidades monetarias en cualquiera de los dos lanzamientos, 4 unidades monetarias en cualesquiera de los cuatro lanzamientos, 8 unidades monetarias una vez cada ocho lanzamientos y se espera perder 20 unidades monetarias una vez en cada ocho intentos. El valor esperado, o la cantidad promedio que se ganaría en cada juego después de un número muy grande de éstos, se determina multiplicando cada cantidad que se gana o se pierde por su respectiva probabilidad y sumando los resultados. De acuerdo con la anterior, la esperanza de ganar es:

$$(2)(1/2)\ +\ (4)(1/4)\ +\ (8)(1/8)\ +\ (-20)(1/8) =\ 0.50$$

Por juego. Nótese que el valor esperado de 50 centavos no es ninguno de los posibles valores de la variable aleatoria; de esta forma, es completamente posible que una variable aleatoria nunca tome el valor de su esperanza.

El ejemplo anterior sugiere la siguiente definición de la esperanza matemática de una variable aleatoria:

"El valor esperado de una variable aleatoria X es el promedio o valor medio de X y está dado por:" (Wackerly, Mendenhall, & Scheaffer, 2009)

$$E(X) = \sum_x xp(x) \qquad si\ x\ es\ discreta, o$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad si\ X\ es\ continua.$$

En donde $p(x)$ y $f(x)$ son las funciones de probabilidad y de densidad de probabilidad, respectivamente.

La esperanza de una variable aleatoria X no es una función de X sino un número fijo y una propiedad de la distribución de probabilidad de X. Por otra parte, el valor esperado puede no existir dependiendo de si la correspondiente suma o integral no converge en un valor finito.

Para ejemplificar. Si la variable aleatoria X representa la suma de las caras de dos dados cuando éstos se lanzan, demostrar que el valor esperado de X es siete.

Con la función de probabilidad de X dada en el apartado de distribuciones discretas se tiene:

$$E(X) = \sum_{x=2}^{12} xp(x) = (2)\left(\frac{1}{36}\right) + (3)\left(\frac{2}{36}\right) + \cdots + (12)\left(\frac{1}{36}\right) = 7$$

En el caso de una distribución continua se tiene que; una variable aleatoria que representa la proporción de accidentes automovilísticos fatales en Estados Unidos se comporta con la siguiente función de distribución

$$f(x) = \begin{cases} 42x(1-x)^5, 0 < x \leq 1 \\ 0, en\ otro\ caso \end{cases}$$

Y como se definió anteriormente el valor esperado de la proporción de accidentes es:

$$E(X) = 42\int_0^1 xf(x) = 42\int_0^1 x^2(1-x)^5$$

$$= 42x^3\left(\frac{1}{3} - \frac{5x}{4} + 2x^2 - \frac{5x^3}{3} + \frac{5x^4}{7} - \frac{x^5}{8}\right)\Big|_0^1$$

$$= 0.25$$

2.2 DISTRIBUCIONES DISCRETAS ÚTILES EN EL TRABAJO DE SEGUROS GENERALES

Como se mencionó anteriormente, una variable aleatoria es una función de valor real definida sobre un espacio muestral. Por lo tanto, una variable aleatoria se puede usar para identificar eventos numéricos que son de interés en un experimento, en este caso, el evento de interés es la frecuencia con la que se siniestran las obligaciones adquiridas de una aseguradora en los ramos de daños sin autos. Tomado a X como nuestra variable aleatoria que oscila entre 0 y 1.

Como la frecuencia de siniestralidad tiende a ser un número finito de valores se determina que dichas variables aleatorias son de tipo discreta.

Por lo que es necesario indagar la estructura y los parámetros que son empleados en las distribuciones discretas utilizadas comúnmente en los seguros generales.

2.2.1 Distribución Binomial

Se dice que se encuentra frente a una variable aleatoria con distribución binomial cuando se considera el número de veces que sucede un evento A en n ejecuciones independientes de un experimento aleatorio.

Si se comienza con un único ensayo. En este caso, suponiendo que la probabilidad de que se produzca el suceso en un único ensayo sea denotada por p, se puede expresar la probabilidad de la ocurrencia del suceso A como:

$$P\ (A)\ =\ p$$

Y del mismo modo, se puede decir que la probabilidad de que no ocurra A será $1-p$, es decir:

$$P\ (\text{"}Complementario\ de\ A\text{"}) =\ 1-p\ =\ q$$

Que también se expresa del siguiente modo:

$$P\ (no(A)) = q$$

Considerando la variable aleatoria X el número de veces que se produce el suceso A y bajo el supuesto de que el experimento se realiza 1 vez.

En este caso, el número de valores diferentes que puede tomar la variable X serán dos: el 0 y el 1. Es decir, puede ocurrir que el suceso se produzca, en cuyo caso $X = 1$ o que no se produzca, en cuyo caso será $X = 0$ y no hay más posibilidades. El espacio muestral se reduce a dos elementos: el 0 y el 1.

Las probabilidades correspondientes serán:

$$P(X = 0) = q$$

$$P(X = 1) = p$$

por lo tanto, la función de probabilidad tomará los valores:

$$f\ (0) = q$$

$$f\ (1) = p$$

Las dos expresiones anteriores pueden agruparse del siguiente modo:

$$f(x) = p^x q^{1-x} \quad x = 0{,}1$$

Visto de otra manera:

$$f(0) = p^0 q^{1-0} = q$$

$$f(1) = p^1 q^{1-1} = p$$

Además, si se necesita, se puede suponer que la función toma el valor 0 para cualquier otro valor de x.

Ahora supóngase que el experimento anterior se lleva a cabo n veces. En esta situación la variable aleatoria X es el número de veces que se produce el suceso A puede tomar los valores desde 0 hasta n. Obtengamos cada una de estas probabilidades.

Para la determinación de estas probabilidades consideremos el suceso: "El número de veces que se produce el suceso A es exactamente x al repetir n veces el experimento. Es decir que la variable aleatoria X tome el valor x, es decir que el suceso A se haya producido x veces y el Complementario(A) se haya producido $n - x$ veces.

Si representamos mediante una secuencia de letras que identifiquen el suceso que se ha producido, tendríamos una expresión similar a la siguiente:

$$A, A, A, \ldots, A, C, C, \ldots C$$

En la cual la letra A significa que se ha producido el suceso A y la letra C significa que se ha producido el suceso complementario de A.

Si suponemos que A se ha producido x veces y C se ha producido $n-x$ veces, la probabilidad de que se produzca el suceso anterior será:

$$p^x q^{n-x}$$

Si además consideramos todos los casos similares a los anteriores, es decir que A haya sucedido x veces, pero no todos seguidos sino en los posibles órdenes en que puedan producirse, y los sumamos tendremos el valor de la función de probabilidad para el valor x de la variable aleatoria X.

El número de posibles combinaciones es: $\binom{n}{x}$

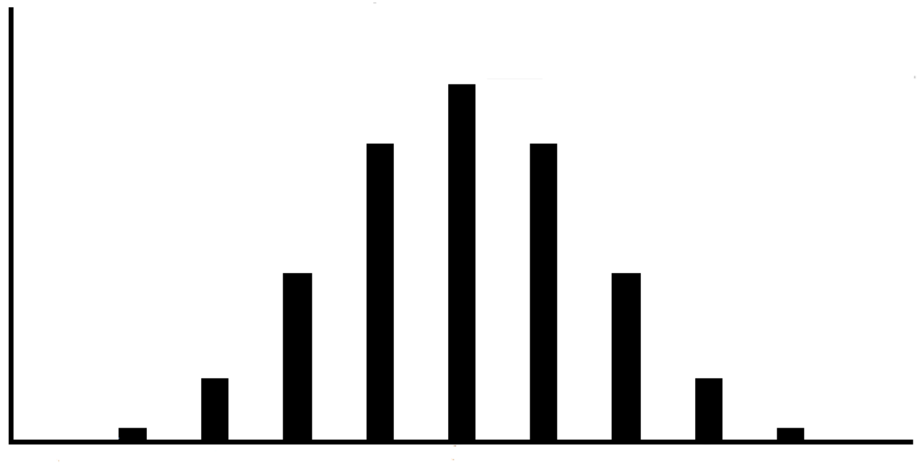
La función de probabilidad es el conjunto probabilidades para los diferentes posibles valores de x; en este caso las probabilidades son las siguientes:

$$f(x) = \binom{n}{x} p^x q^{n-x} \quad x = 0,1,2,\dots,n$$

La función de distribución será por definición:

$$F(x) = \sum_{k \le x} \binom{n}{k} p^k q^{n-k}$$

Representación gráfica de la función de distribución Binomial



Fuente: Spiegel, M. (1991). Probabilidad y Estadística.

Como se muestra, gráficamente así es la función de densidad de probabilidad de la distribución binomial para diferentes valores de n y p.

Hacer que la función dependa de los valores de n y de p, de manera que la representación gráfica sirva para diferentes valores de p.

Para obtener μ y σ se utilizara el valor esperado, donde $\mu = E(X)$ y $\sigma^2 = Var(X)$, como se definió anteriormente

$$E(X) = \sum_X xp(x) = \sum_{x=1}^{n} x \frac{n!}{(n-x)!x!} p^x q^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n!}{(n-x)!(x-1)!} p^x q^{n-x}$$

Dado que el primer término de la suma es 0.

Los sumandos de esta última expresión presentan un sorprendente parecido con las probabilidades binomiales. De hecho, si se factoriza np en cada término de la suma y se hace un cambio de variable $z = x - 1$,

$$E(X) = np \sum_{x=1}^{n} \frac{n-1!}{(n-x)!(x-1)!} p^{x-1} q^{n-x}$$

$$= np \sum_{z=0}^{n-1} \frac{n-1!}{(n-1-z)!z!} p^z q^{n-1-z}$$

$$= np \sum_{z=0}^{n-1} \binom{n-1}{z} p^z q^{n-1-z}$$

Observe que $p(z) = \binom{n-1}{z} p^z q^{n-1-z}$ es la función de probabilidad binomial basada en $(n - 1)$ pruebas, y se deduce que

$$\mu = np$$

Como ya se mencionó $\sigma^2 = Var(X)$ y $Var(X) = E(X^2) - \mu^2$. Entonces, σ^2 se puede calcular si se encuentra $E(X^2)$. Hallar $E(X^2)$ directamente es difícil por lo tanto se utilizará la siguiente expresión.

$$E(X^2) = E[X(X-1)] + E(X) = E[X(X-1)] + \mu$$

En este caso

$$E[X(X-1)] = \sum_{x=0}^{n} x(x-1) \frac{n!}{(n-x)!x!} p^x q^{n-x}$$

Los términos primero y segundo de esta suma son iguales a cero (cuando $x = 0$ y $x = 1$). Entonces

$$E[X(X-1)] = \sum_{x=0}^{n} \frac{n!}{(n-x)!(x-2)!} p^x q^{n-x}$$

De nuevo, los sumandos de la última expresión se asemejan mucho a probabilidades binomiales. Se Factoriza $n(n-1)p^2$ en cada término de la suma y sea $z = x - 2$ para obtener

$$E[X(X-1)] = n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(n-x)!(x-2)!} p^{x-2} q^{n-x}$$

$$= n(n-1)p^2 \sum_{z=2}^{n-2} \frac{(n-2)!}{z!(n-2-z)!} p^z q^{n-2-z}$$

$$= n(n-1)p^2 \sum_{z=2}^{n-2} \binom{n-2}{z} p^z q^{n-2-z}$$

De nuevo se observa que $p(z) = \binom{n-2}{z} p^z q^{n-2-z}$es la función de probabilidad binomial basada en $(n-2)$ pruebas.por lo tanto

$$E[X(X-1)] = n(n-1)p^2$$

y

$$E(X^2) = E[X(X-1)] + \mu = n(n-1)p^2 + np$$

Entonces

$$\sigma = E(X^2) - \mu^2 = n(n-1)p^2 + np - n^2p^2$$

$$= np[(n-1)p + 1 - np] = np(1-p) = npq$$

2.2.2 Distribución Binomial negativa

La función de probabilidad para una distribución binomial negativa está dada por:

$$f(x) = \binom{k+x-1}{x} p^k q^x \quad k > 0, x = 0,1,2 \ldots$$

Siendo k el número total de éxitos en una sucesión de $k + x$ ensayos y siendo x el número de fallos que ocurren antes de la obtención de los $\boldsymbol{k}$ éxitos. El éxito k-ésimo se obtiene en última prueba, es decir en la prueba $k + x$.

De la fórmula anterior se lee que ha habido k éxitos y x fracasos. Además, como se obliga a que el último éxito sea el de la última prueba el número combinatorio se reduce.

Obsérvese que cuando k es igual a 1 se tiene que:

$$f(x) = \binom{x}{x} pq^x \quad x = 0,1,2, \ldots$$

o lo que es lo mismo:

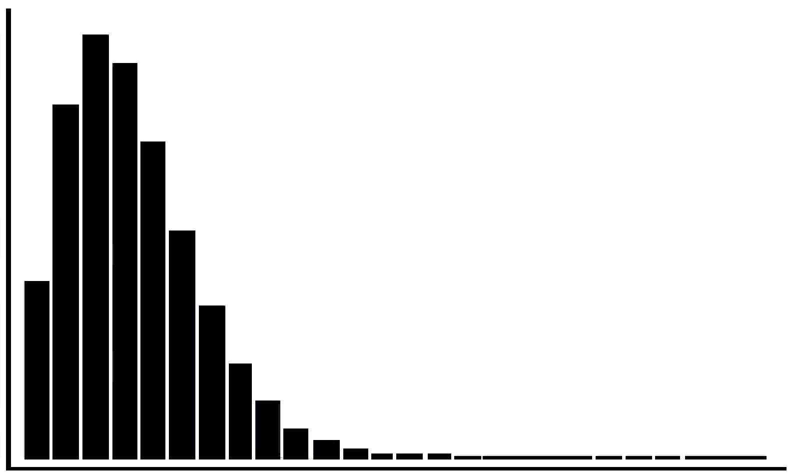
$$f(x) = pq^x \;\; x = 0,1,2, \ldots$$

Si X es una variable aleatoria con una distribución binomial negativa. Calculando el valor esperado para obtener μ y σ^2 como en el tema anterior se tiene que

$$\mu = \frac{r}{p}$$

$$\sigma^2 = \frac{r(1-p)}{p^2}$$

Representación gráfica de la función de distribución Binomial Negativa.



Fuente: Spiegel, M. (1991). Probabilidad y Estadística.

Es decir que cuando el número de éxitos es igual a 1 se trata del caso particular de la distribución geométrica.

2.2.3 Distribución Geométrica

Considerando los posibles valores de una variable aleatoria que se pueden obtener cuando, al ejecutar un experimento aleatorio, se cuenta el número de fallos antes de conseguir el primer acierto. Por ejemplo, el número de veces (menos uno) que se han obtenido al lanzar un dado hasta obtener el número 6. Es decir, la variable aleatoria X representará el número de fallos, intentos o pruebas antes del primer éxito, es decir que X podrá tomar valores en el conjunto: {0, 1, 2, 3,..., x,...}

Suponiendo que p es la probabilidad de éxito y q la probabilidad de fracaso. Se supone que todas las pruebas tienen la misma probabilidad de éxito y que las pruebas son totalmente independientes unas de otras, esto quiere decir que no se están considerando fenómenos en los que el resultado de un experimento influya en el resultado del siguiente, es decir que no se consideran fenómenos de aprendizaje.

Por ejemplo, X podría ser el número de disparos que preceden al primer acierto, en el caso de una prueba olímpica de tiro al arco. Se supone que la probabilidad de acierto de cada uno de los jugadores es constante para cada uno de ellos, ya que cada uno de ellos tiene una habilidad adquirida tras muchas horas de entrenamiento, por ejemplo, p. La probabilidad de fallo será la complementaria, es decir: $1 - p$.

Las probabilidades correspondientes a cada uno de los posibles sucesos serán las siguientes, es decir, la probabilidad de cometer 0 fallos será:

$$P(X = 0) = p$$

La probabilidad de cometer exactamente un fallo y luego tener éxito, será:

$$P(X = 1) = q.p$$

La probabilidad de cometer exactamente dos fallos y luego tener éxito, será:

$$P(X = 2) = q.q.p\ldots$$

La probabilidad de hacer exactamente x fallos y luego tener éxito, será:

$$P(X = x) = q.q \ldots q.p = p.q^{x-1}$$

De esta manera se pudo deducir la función de probabilidad, es decir, la probabilidad de que la variable aleatoria X tome un valor particular.

Se dice que una variable aleatoria X tiene una distribución de probabilidad geométrica si y sólo si

$$P(x) = q^{x-1}p, para\ x = 1,2,3,\ldots\ tal\ que\ 0 \leq p \leq 1$$

En la fórmula anterior puede verse que la distribución geométrica tiene una función de probabilidad que se corresponde con la probabilidad conjunta de obtener un éxito, p, tras la obtención de x fracasos, cuya probabilidad es q^{x-1}

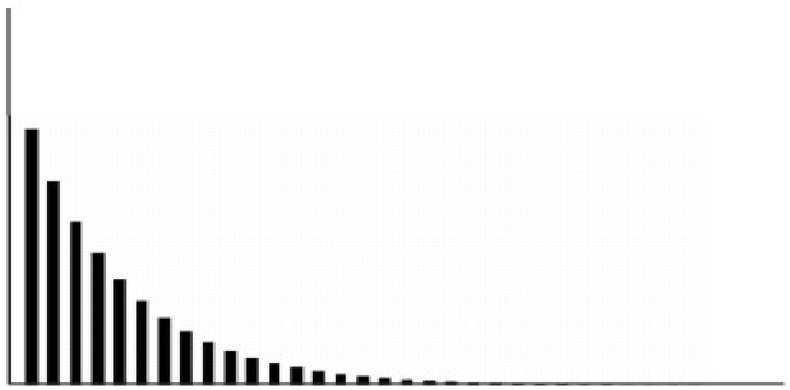
Dado que X es una variable aleatoria con una distribución geométrica y repitiendo el proceso que se hizo con las distribuciones expuestas anteriormente se determina que

$$\mu = \frac{1}{p}$$

$$\sigma^2 = \frac{1-p}{p^2}$$

Una serie de datos que se comportan como una distribución geométrica se pueden ver gráficamente de la siguiente manera.

Representación gráfica de la función de distribución Geométrica



Fuente: Spiegel, M. (1991). Probabilidad y Estadística. Madrid: McGraw-Hill.

2.2.4 Distribución Poisson

Se utiliza para representar diferentes fenómenos aleatorios; por ejemplo: El número de accidentes de tráfico producidos en un determinado punto kilométrico en un determinado intervalo de tiempo; por ejemplo, en un día, en un mes, etc. El número de vehículos que llegan a una estación de servicio en un determinado periodo de tiempo, etc.

Se le suele denominar, distribución de llegadas ya que en muchas ocasiones las llegadas a un sistema siguen aproximadamente esta distribución.

Si se realiza una serie muy grande de ensayos de Bernoulli, en los cuales la probabilidad de que se verifique un determinado evento sea muy pequeña, la probabilidad de que se produzca x veces el evento considerado está dada por la distribución de Poisson.

A la siguiente expresión

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!}; para\ \lambda \neq 0\ y\ x = 0,1,2,\ldots$$

Se le denomina función de probabilidad de Poisson.

Esta distribución es una buena aproximación a la distribución Binomial cuando n es grade y p es pequeño, por ejemplo, $n = 1000$ y $p = 1/100$. En general, cuando el producto $n \cdot p$ es superior a 5 la aproximación es adecuada.

La distribución Binomial se aproxima a la de Poisson cuando el número de observaciones n es grande y la probabilidad p en pequeña.

Demostración:

Para demostrarlo se tiene que a partir de la función de probabilidad de la distribución de Binomial

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \quad x = 0,1,2,\ldots,n$$

Se observará que ocurre con la distribución de probabilidad cuando n tiende a infinito y la probabilidad p tiende a cero, para cada uno de los valores de x.

En algunas situaciones el producto $n \cdot p$ podría estabilizarse; en estos casos se observa que existe el límite.

Téngase en cuenta que se consideran casos en los cuales p es muy pequeño y n es muy grande. Por lo tanto, el cociente m entre n debe ser muy pequeño; lo cual significa que el número de casos λ es muy pequeño en relación al número n. Por ejemplo, si n representa el número de nacimientos en un año, m podría representar el número de nacimientos de niños siameses; es decir, el cociente λ entre n representará la probabilidad de que se produzca un parto de niños siameses (hermanos gemelos unidos por alguna parte de su cuerpo)

Suponiendo que el producto sea un valor determinado λ, es decir que la probabilidad es $p = {}^{\lambda}/_{n}$, entonces el límite puede calcularse del siguiente modo:

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)(n-x)(n-x-1)\ldots}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)}{x!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)}{n^x} \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)}{n \cdot n \cdot \ldots \cdot n} \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

Obsérvese que en numerador del primer cociente consta de x términos, es decir, desde el primer factor, que es n, hasta el factor x, que es el n-(x-1) suman en total x factores. Del mismo modo en el denominador se tiene n elevado a x, o lo que es lo mismo el producto de n tomado x veces. Expresando cada uno de los factores en fracciones independientes y haciendo algunas equivalencias se tiene que:

$$= \frac{n}{n} \frac{n-1}{n} \frac{n-2}{n} \frac{n-3}{n} \cdots \frac{n-x+1}{n} \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n}{n}\frac{n-1}{n}\frac{n-2}{n}\frac{n-3}{n}\cdots\frac{n-x+1}{n}\frac{\lambda^x}{x!}\left(1+\frac{1}{-\frac{n}{\lambda}}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n}{n}\frac{n-1}{n}\frac{n-2}{n}\frac{n-3}{n}\cdots\frac{n-x+1}{n}\frac{\lambda^x}{x!}\left(1+\frac{1}{-\frac{n}{\lambda}}\right)^{-\frac{n}{\lambda}(-\lambda)}\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n}{n}\frac{n-1}{n}\frac{n-2}{n}\frac{n-3}{n}\cdots\frac{n-x+1}{n}\frac{\lambda^x}{x!}\left[\left(1+\frac{1}{-\frac{n}{\lambda}}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}\left(1-\frac{\lambda}{n}\right)^{-x}$$

Haciendo el límite cuando n tiende a infinito, nos queda:

$$P(X=x)=\frac{\lambda^x}{x!}e^{-\lambda}$$

Que es exactamente la probabilidad de la distribución de Poisson para $X=x$.

Obsérvese que los x primeros factores tienen por límite el valor 1 y lo que se encuentra entre corchetes es la definición del número e. Recuérdese que en matemáticas se demuestra que:

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e^{-\lambda}$$

También se puede demostrar que el límite anterior es el número e cuando la x varía en el campo de los números reales negativos; es decir, cuando x tiende a menos infinito.

Haciendo lo mismo para todos los valores de la variable aleatoria X se tiene

$$f(x)=e^{-\lambda}\frac{\lambda^x}{x!}\ \ para\ x=0,1,\ldots$$

que es la función de probabilidad de la distribución de Poisson.

La esperanza de la distribución de Poisson tiene el siguiente valor

$$E(X)=\sum_x xf(x)=\sum_{x=o}^{\infty}x\frac{e^{-\lambda}\lambda^x}{x!}=\lambda\sum_{x=0}^{\infty}\frac{\lambda^{x-1}e^{-\lambda}}{(x-1)!}$$

Ahora bien, sea $z=x-1$ se obtiene que

$$E(X)=\lambda\sum_{z=0}^{\infty}\frac{e^{-\lambda}\lambda^z}{z!}$$

Observando que con $z=x-1$ es semejante a la distribución Poisson por lo tanto

$$\mu = \lambda$$

En el caso de la varianza se tiene que $Var(X) = E(X^2) - [E(X)]^2$, para este caso se usara la expresión $E(X^2) = E[X(X-1)] + \mu$

$$E[X(X-1)] = \sum_x x(x-1)f(x) = \sum_{x=o}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}e^{-\lambda}}{(x-2)!}$$

Ahora bien, sea z=x-2 por lo que se obtiene

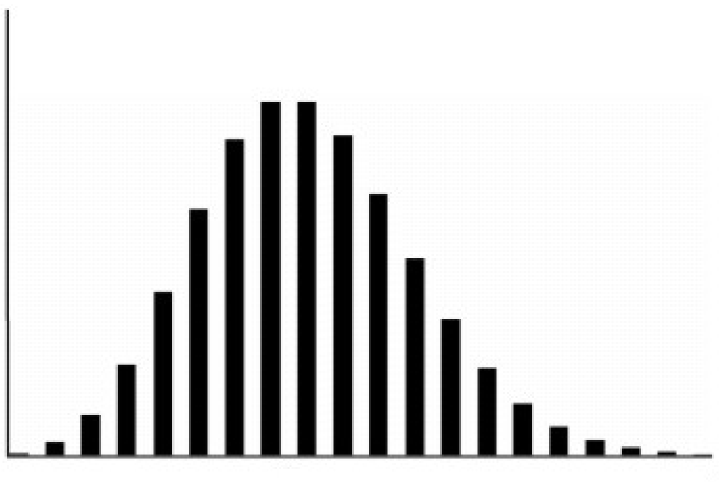
$$E[X(X-1)] = \lambda^2 \sum_{x=0}^{\infty} \frac{\lambda^{x-2}e^{-\lambda}}{(x-2)!} = \lambda^2$$

Así que $E(X^2) = E[X(X-1)] + \mu = \lambda^2 + \lambda$

Por lo tanto

$$Var(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Representación gráfica de la función de distribución Poisson



Fuente: Spiegel, M. (1991). Probabilidad y Estadística. Madrid: McGraw-Hill.

2.3 DISTRIBUCIONES CONTINUAS ÚTILES EN EL TRABAJO DE SEGUROS GENERALES

De la misma manera que en el apartado anterior, se estudiaran las variables aleatorias pero con un interés de experimento diferente, ahora se estudiara la severidad de la siniestralidad, variable que está relacionada con el costo de siniestralidad de las obligaciones adquiridas de una compañía aseguradora, de tal manera que la variable que explica la severidad requerida se comporta como una distribución continua al no tener un valor finito.

Así que a continuación se expondrá el proceso de tipificación de una variable continua y se profundizará en la importancia del uso estas distribuciones en los seguros generales.

2.3.1 Distribución normal

La función Normal es el modelo más utilizado y con mayor importancia en el campo de la estadística. Sin embargo, su uso es muy limitado en los seguros generales, dado que las variables raramente se comportan de esta forma.

La función de distribución de probabilidad normal es:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Donde μ la media de la población y σ Desviación estándar de la población.

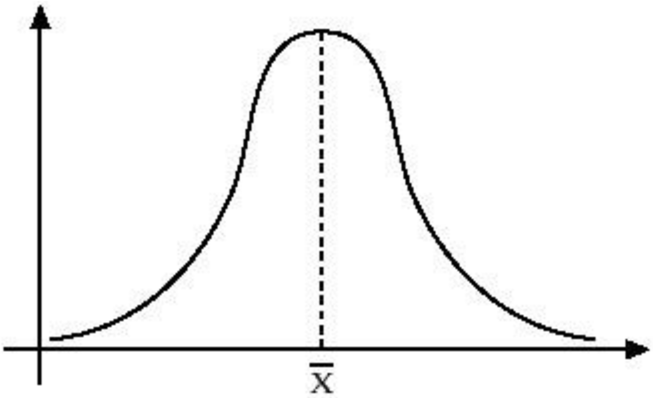
Para resolver esta función se recurren a métodos numéricos para evaluarla, y para hacer esto más sencillo se le ha asignado una variable estandarizada, cuya expresión es la siguiente:

$$z = \frac{x-\mu}{\sigma}$$

La cual está normalmente distribuida con media 0 y desviación estándar unitaria. Así, la función principal queda como:

$$f(x) = f(z) = \int_{-\infty}^{z} \frac{1}{2\pi} e^{\frac{-z^2}{2}} dz$$

Representación gráfica de la función de distribución Normal



Fuente: Armitage, P., & Berry, G. (1997). Estadística Para La Investigación Biomédica.

En la mayoría de los seguros generales la distribución del costo de siniestralidad se comporta de manera asimétrica con una larga cola a la derecha por eso es importante estudiar otras distribuciones con estas características.

.

2.3.2 Distribución Logarítmico-Normal

Esta distribución es frecuentemente utilizada para modelar la distribución del costo de siniestralidad debido a que las variables relacionadas a la siniestralidad son generalmente positivas, por lo cual es usual que presenten distribuciones asimétricas. Así, se ha propuesto aplicar una transformación logarítmica donde $Y = Ln\, X$, está normalmente distribuida; luego X está distribuida en forma Normal, y su función de densidad de probabilidad es

$$f(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{x\beta} e^{-\frac{1}{2}\left(\frac{\ln x - \alpha}{\beta}\right)^2}$$

Donde los parámetros de la función son α y β, que son la media y la desviación estándar de los logaritmos de la variable aleatoria, y están definidos como sigue:

$$\alpha = \sum_{i=1}^{n} \frac{\ln x_i}{n}$$

$$\beta = \left[\sum_{i=1}^{n} \frac{(\ln x_i - \alpha)^2}{n}\right]^{1/2}$$

Luego la función de distribución de probabilidad es:

$$F(x) = \int_0^x \frac{1}{\sqrt{2\pi}} \frac{1}{x\beta} e^{-\frac{1}{2}\left(\frac{\ln x - \alpha}{\beta}\right)^2} dx$$

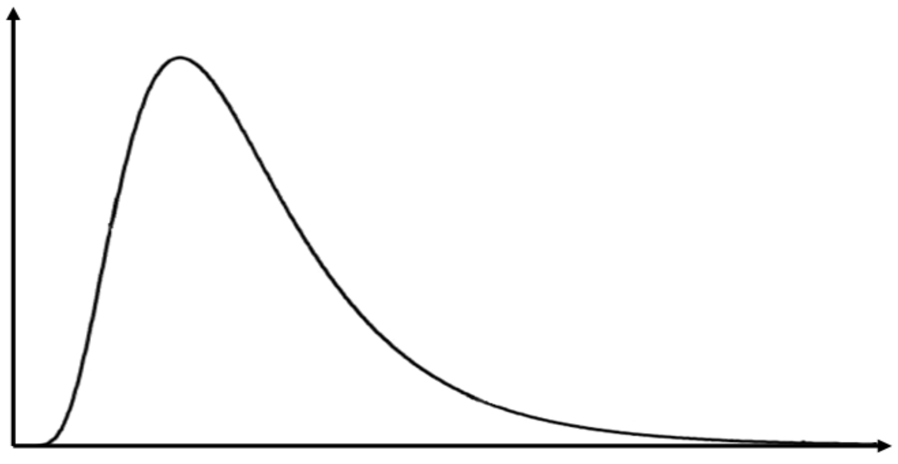
donde e corresponde a la constante de Neper

Al igual que en la distribución normal, se le asigna a z los siguientes valores:

$$z = \frac{Ln\, x - \alpha}{\beta}$$

La función de distribución de probabilidad entonces, sigue la siguiente tendencia:

Representación gráfica de la función de distribución Log-Normal



Fuente: Armitage, P., & Berry, G. (1997). Estadística Para La Investigación Biomédica.

Estudios realizados por Poblete *et al*., (2002), identifican a la función Log-Normal, entre otras funciones, como la que presenta mejor bondad de ajuste a series de caudales anuales, por sobre un 90% para el test de Kolmogorov-Smirnov y ji cuadrado.

2.3.3 Distribución Exponencial

Se suele utilizar para modelizar el tiempo transcurrido entre dos sucesos aleatorios cuando la tasa de ocurrencia λ se supone constante.

La distribución exponencial es el equivalente continuo de la distribución geométrica discreta.

Nos interesa saber el tiempo hasta que ocurre determinado evento, sabiendo que, el tiempo que pueda ocurrir desde cualquier instante dado t, hasta que ello ocurra en un instante t_f , no depende del tiempo transcurrido anteriormente en el que no ha pasado nada.

Concretando, si una variable aleatoria continua X distribuida a lo largo de los reales, es tal que su función de densidad es

$$f(x) = \lambda e^{-\lambda x} \; si \; x > 0$$

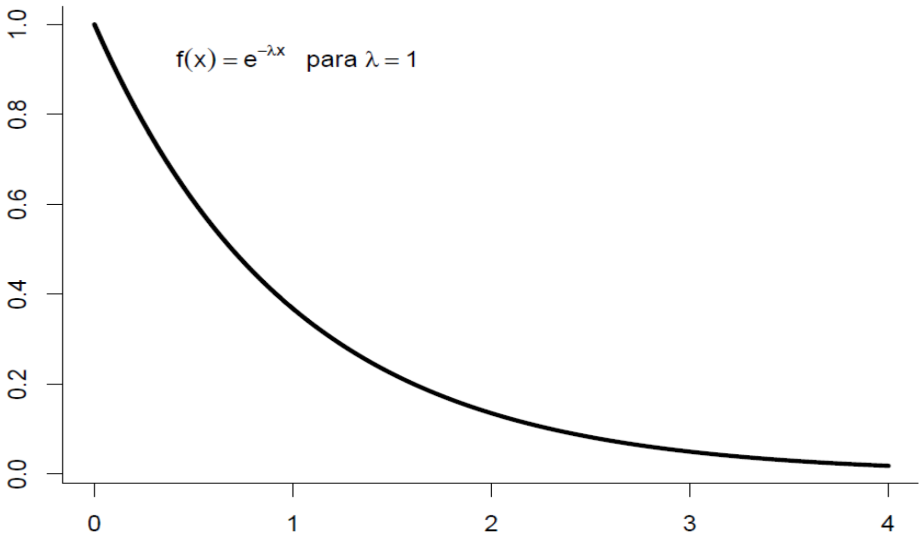
Se dice que sigue una distribución exponencial de parámetro λ.

Un cálculo inmediato nos dice que si $x > 0$

$$\int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x}$$

La tendencia se puede ver en la siguiente gráfica:

Representación gráfica de la función de distribución Exponencial



Fuente: Casella, G., & Berger, R. (1990). Statistical inference.

Por lo que la función de distribución se denota como:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & si\ x > 0 \\ 0 & en\ otro\ caso \end{cases}$$

Los parámetros del valor esperados son:

$$E[X] = \frac{1}{\lambda} \text{ y } V[X] = \frac{1}{\lambda^2}$$

2.3.4 Distribución Gamma

La distribución gamma tiene aplicación en el costo de siniestralidad sobre todo si los riegos de los seguros generales son heterogéneos

Una generalización de la distribución exponencial es la distribución Gamma. Una variable aleatoria de tiempo de vida se dice que se distribuye según una Gamma de parámetros α y λ, $G(\alpha, \lambda)$.

Es una función que extiende el concepto de factorial a los números complejos.

Sea una función $\Gamma\ (0, \infty)$ que pertenece a los reales enteros donde

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

Con el fin de observar algunos resultados o propiedades de esta función, se procede a integrar por partes. Tomando $u = x^{\alpha-1}$ y $dv = e^{-x}dx$,así se obtiene

$$\Gamma(\alpha) = -e^{-x}x^{\alpha-1}|_{0}^{\infty} + \int_0^{\infty} e^{-x}(\alpha-1)x^{\alpha-2}dx = (\alpha-1)\int_0^{\infty} x^{\alpha-2}e^{-x}dx$$

Para $\alpha > 1$, lo cual ocasiona la formula $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

Al aplicar reiteradamente la formula anterior se tiene:

$$\Gamma(\alpha) = (\alpha-1)(\alpha-2)\Gamma(\alpha-2) = (\alpha-1)(\alpha-2)(\alpha-3)\Gamma(\alpha-3)$$

Y así sucesivamente, se evidencia que cuando $\alpha = n$ donde n es un entero positivo,

$$\Gamma(n) = (n-1)(n-2)\ldots\Gamma(1).$$

Sin embargo, por la definición de $\Gamma(\alpha), \Gamma(1) = \int_0^{\infty} e^{-x}dx = 1$ y de aquí $\Gamma(n) = (n-1)!$

Es una distribución de probabilidad continua adecuada para modelizar el comportamiento de variables aleatorias con asimetría positiva y/o los experimentos en donde está involucrado el tiempo.

Una variable aleatoria X tiene una distribución gamma si su función de densidad está dada por:

$$F(x,\alpha,\beta) = f(x) = \begin{cases} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta} & para\; x > 0; \alpha, \beta > 0 \\ 0 & en\; otro\; caso \end{cases}$$

Para demostrar que la función f es una función de densidad se debe probar que $\int_{-\infty}^{\infty} f(x)dx = 1$ .para este caso $\int_{-\infty}^{\infty} f(x,\alpha,\beta)dx = 1.$|

De esta manera $\int_{-\infty}^{\infty} f(x,\alpha,\beta)dx = \int_{-\infty}^{\infty} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta}dx$

$$\int_{-\infty}^{0} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta}dx + \int_0^{\infty} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta}dx$$

Pero según la definición de $f(x,\alpha,\beta), \int_{-\infty}^{0} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta}dx = 0$

Así, $\int_{-\infty}^{\infty} f(x,\alpha,\beta)dx = \int_0^{\infty} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1}e^{-x/\beta}dx = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-x/\beta}dx$

Realizando el siguiente cambio de variable. Sea $u = {}^{x}/_{\beta}$, entonces $x = \beta u$, así $dx = \beta du$. Por lo cual se tendría que:

$$\frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x^{\alpha-1}e^{-x/\beta}dx = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty}(\beta u)^{\alpha-1}e^{-u}\beta du$$

$$\frac{1}{\beta\Gamma(\alpha)}\int_0^{\infty} \beta^{\alpha}u^{\alpha-1}e^{-u}du = \frac{1}{\beta}\int_0^{\infty} u^{\alpha-1}e^{-u}du$$

Ahora bien, recordando que $\Gamma\ (0,\infty)$ pertenece a los reales enteros donde

$$\Gamma(\alpha) = \int_0^{\infty} u^{\alpha-1}e^{-u}du\ , \forall\ \alpha > 0.$$

Así que,

$$\frac{1}{\Gamma(\alpha)}\int_0^{\infty} u^{\alpha-1}e^{-u}du = \frac{1}{\Gamma(\alpha)}\Gamma(\alpha) = 1$$

Por lo tanto es f una función de densidad.

Para obtener la esperanza se utiliza la expresión:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^{\infty} x\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}dx = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x^{\alpha}e^{-x/\beta}dx$$

Ahora bien, sea $u = {}^{x}/_{\beta}$ $: x = u\beta$; con lo que $dx = \beta du$, así

$$E(X) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty}(u\beta)^{\alpha}e^{-u}\beta dx = \frac{\beta}{\Gamma(\alpha)}\int_0^{\infty} u^{\alpha}e^{-u}du = \frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\beta\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha\beta$$

Por lo tanto

$$E(X) = \alpha\beta$$

Para la varianza se usará la siguiente formula $V(X) = E(X^2) - [E(X)]^2$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^{\infty} x^2\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}dx = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty} x^{\alpha+1}e^{-x/\beta}dx$$

Ahora bien, sea $u = {}^{x}/_{\beta}$ $: x = u\beta$; con lo que $dx = \beta du$, así

$$E(X^2) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\int_0^{\infty}(u\beta)^{\alpha+1}e^{-u}\beta dx = \frac{\beta^2}{\Gamma(\alpha)}\int_0^{\infty} u^{\alpha+1}e^{-u}du = \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)}$$

$$= \frac{\beta^2\alpha(\alpha+1)\Gamma(\alpha)}{\Gamma(\alpha)}$$

$$= (\alpha+1)\alpha\beta^2$$

Terminando la fórmula de la varianza se tiene que

$$V(X) = E(X^2) - [E(X)]^2 = (\alpha + 1)\alpha\beta^2 - (\alpha\beta)^2 = \alpha^2\beta^2 + \alpha\beta^2 - \alpha^2\beta^2 = \alpha\beta^2$$

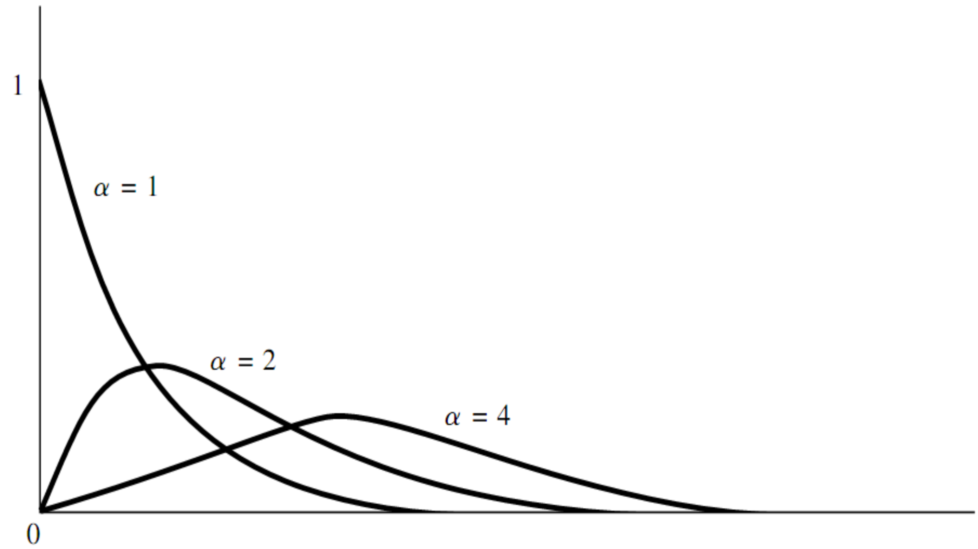
La distribución gamma tiene relación con otras distribuciones. Si se tiene un parámetro α de valores elevados y β pequeña, entonces la función gamma converge con la distribución normal. De media $\mu = \alpha\beta$, y varianza $\sigma^2 = v]$ entonces la variable aleatoria se distribuye como una ji cuadrada con v grados de libertad.

Si $\alpha = 1$, entonces se tiene la distribución exponencial de parámetro $\lambda = {}^{1}/_{\beta}$.

De esta forma, la distribución gamma es una distribución flexible para modelizar las formas de la asimetría positiva, de las más concentradas y puntiagudas, a las más dispersas y achatadas.

Para valorar la evolución de la distribución al variar los parámetros se tienen los siguientes gráficos. Primero se comprueba que para $\alpha = 1$ la distribución tiene similitudes con la exponencial.

Representación gráfica de la función de distribución Gamma



Fuente: Ruiz–Maya, L. (1977). Métodos Estadísticos de investigación

2.3.5 Distribución Pearson Tipo III

Esta distribución posee una gran flexibilidad y diversidad de forma, dependiendo de los valores de sus parámetros, asimilando su utilización para valores extremos de siniestralidad. La función de densidad de probabilidad se define como

$$f(x) = \frac{1}{\alpha\Gamma(\beta)}\left\{\frac{x-\delta}{\alpha}\right\}^{\beta-1} e^{\frac{x-\delta}{\alpha}}$$

Donde α, β, δ son los parámetros de la función y $\Gamma(\beta)$ es la función de Gamma. Los parámetros α, β, δ se evalúan a partir de *n* datos medidos. Asimismo los parámetros de la distribución pueden ser estimados en función del promedio (*x*) y de la desviación estándar (S) de la muestra, por medio de las siguientes expresiones:

$$\alpha = \frac{S}{\sqrt{\beta}}\,;\ \beta = \left(\frac{2}{\gamma}\right)^2\ ;\ \delta = \bar{x} - \alpha\beta$$

Donde γ Coeficiente de sesgo y e Constante de Neper

El coeficiente de sesgo, se define como,

$$\gamma = \sum_{i=1}^{n} \frac{(x_i-\bar{x})^3/n}{S^3}$$

La función de distribución de este modelo es:

$$F(x) = \frac{1}{\alpha\Gamma(\beta)}\int_0^x e^{-\left(\frac{x-\delta}{\delta}\right)}\left(\frac{x-\delta}{\delta}\right) dx$$

Entonces, sustituyendo se alcanza la siguiente expresión:

$$y = \frac{x-\delta_1}{\alpha_1}$$

Finalmente la ecuación queda como:

$$F(y) = \frac{1}{\Gamma(\beta_1)}\int_0^y y^{\beta-1}e^{-y}dy$$

Siendo la anterior una función *ji* cuadrada con $2\beta_1$ grados de libertad y $^{\chi^2}$= 2*y:*

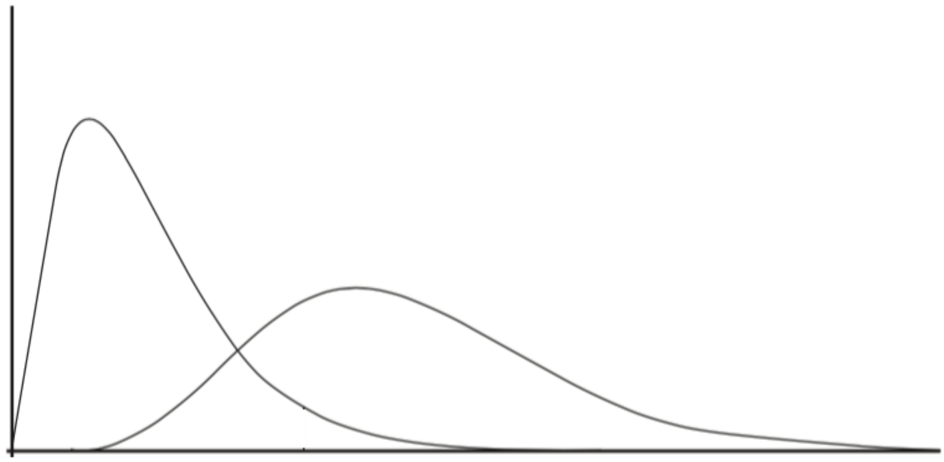
$$F(y) = F(^{\chi 2}|v) = F_{x^2}(2y|2\beta_1)$$

Para ajustar está distribuciones de tres parámetros, se necesita estimar el coeficiente de asimetría de la distribución; para ello es necesario disponer de una serie con longitud de registros larga.

Las distribuciones con dos parámetros, son usualmente preferidas cuando se dispone de pocos datos, ya que reducen la varianza de la muestra.

Lo dicho anterior me su pude visualizar gráficamente en la siguiente imagen mostrando dos ejemplos con parámetros diferentes de la distribución Pearson III

Representación gráfica de la función de distribución Pearson



Fuente: Ruiz–Maya, L. (1977). Métodos Estadísticos de investigación

2.3.6 Distribución Gumbel

Si se tienen N muestras, cada una de las cuales contienen n eventos y si se selecciona el máximo de x de los n eventos de cada muestra, es posible que, a medida que n aumenta, la función de distribución de probabilidad de x tiende a:

$$F(x) = e^{-e^{-d(x-\mu)}} \quad para \; -\infty \leq x \leq \infty$$

Donde e es la Constante de Neper.

Los parámetros de la distribución de una muestra de tamaño infinito, tienden a los siguientes valores, en base a la media aritmética y la desviación estándar de la muestra:

$$d = \frac{1}{0.779696 \cdot S} \quad ; \quad \mu = \bar{x} - 0.450047 \cdot S$$

2.3.7 Distribución Goodrich

La función de Goodrich elimina los valores extremos, en que la probabilidad de ocurrencia es muy pequeña. Por lo mismo, consigue suprimir las que puede provocar un solo valor anómalo. Así, la función de distribución de Goodrich queda definida por:

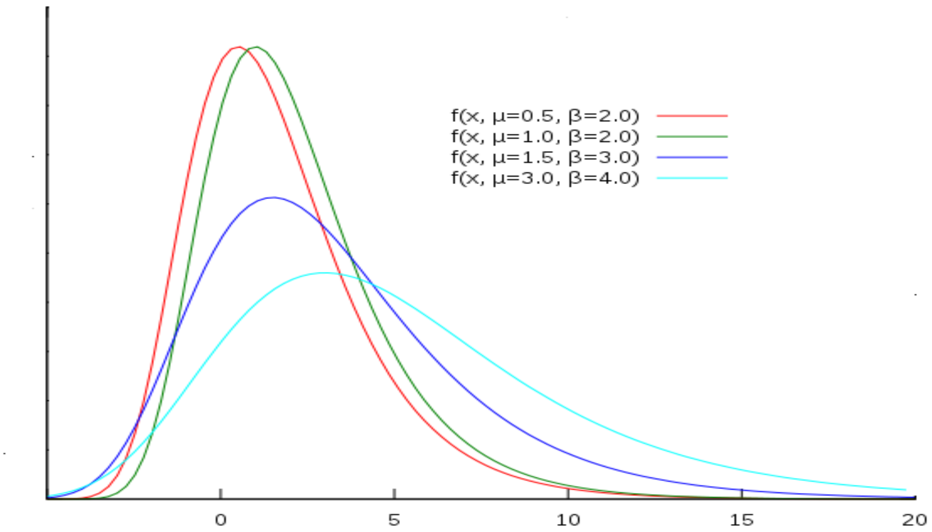
$$F(X) = 1 - e^{-a(x-x_1)^{1/p}} \quad Para\, X_1 < X \leq \infty$$

Donde los parámetros se determinan a partir del siguiente sistema de ecuaciones:

$$\frac{m^3}{S^3} P(p)\ ;\ a^{2p} = \frac{1}{S^2}\left[\Gamma(2p+1) - \Gamma^2(p+1)\right]\ ;\ X_1 = \bar{x} - \frac{\Gamma(p+1)}{a^p}$$

Donde m_3 es el momento central de orden tres, S^3 es la desviación típica al cubo, $P(p)$, la función auxiliar de Goodrich, S^2 es la varianza muestral, Γ es la Función Gamma, $\bar{x}$ es Media muestral y e es la Constante de Neper

Representación gráfica de la función de distribución Goodrich



Fuente: O'Reilly, F., & Rueda, R. (1999). Tests of fit for discrete distributions based on the probability generating function.

2.3.8 Distribución de Erlang

La distribución de Erlang sucede cuando el parámetro α , en la distribución gamma, es un entero positivo. Es decir,

$$f(x,\alpha,\beta) = \begin{cases} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, para\ x > 0, \alpha\ \in enteros\ positivos, \beta > 0 \\ 0\ , en\ otro\ caso \end{cases}$$

Es la función de densidad de probabilidad para la variable aleatoria X que tiene una distribución de Erlang.

Ahora bien, si el número de eventos aleatorios independientes que ocurren en un lapso específico es una variable de Poisson, con una frecuencia constante de ocurrencia igual a $1/\beta$, entonces para un α , el tiempo de espera hasta que ocurre el $\alpha - ésimo$ evento de Poisson tiene una distribución de Erlang. Lo anterior se evidencia al comparar las funciones de distribución acumulativas de las distribuciones Poisson y Erlang. Así, siendo F_P la función de distribución acumulada para la distribución de Poisson, se tiene que:

$$P(X \leq x) = F_p(x, \lambda) = \sum_{n=0}^{x} \frac{e^{-\lambda}\lambda^n}{n!}$$

Con lo cual, la probabilidad de que ocurran a lo más $\alpha - 1$ eventos de Poisson en un tiempo x , a una frecuencia constante $^1/_\beta$, está dada por:

$$P(X \leq \alpha - 1) = F_p\left(\alpha - 1, {}^x/_\beta\right) = \sum_{n=0}^{\alpha-1} \frac{e^{-x/\beta}(^x/_\beta)^n}{n!}$$

Por otro lado, si se supone que el tiempo de espera sigue el modelo de Erlang, la probabilidad de que el tiempo de espera hasta que ocurra el $\alpha - \acute{e}simo$ evento exceda un lapso x específico, está determinado por

$$P(X > x) = 1 - F_E(x, \alpha, \beta)$$

$$= 1 - \left\{1 - \left[1 + \frac{x}{\beta} + \frac{1}{2!}\left(\frac{x}{\beta}\right)^2 + \cdots + \frac{1}{(\alpha-1)!}\left(\frac{x}{\beta}\right)^{\alpha-1}\right] e^{-x/\beta}\right\}$$

$$= \left[1 + \frac{x}{\beta} + \frac{1}{2!}\left(\frac{x}{\beta}\right)^2 + \cdots + \frac{1}{(\alpha-1)!}\left(\frac{x}{\beta}\right)^{\alpha-1}\right] e^{-x/\beta}$$

$$= \sum_{n=0}^{\alpha-1} \frac{e^{-x/\beta}\left(^x/_\beta\right)^n}{n!} = F_P\left(\alpha - 1, {}^x/_\beta\right)$$

En otras palabras, la probabilidad de que el tiempo que transcurre hasta el $\alpha - \acute{e}simo$ evento exceda el valor x es igual a la probabilidad de que el número de eventos de Poisson observados en x no sea mayor que $\alpha - 1$. De esta forma, la distribución de Erlang es el modelo para el tiempo de espera hasta que ocurre el $\alpha - \acute{e}simo$ evento de Poisson, y la distribución de Poisson es el modelo para el número de eventos independientes que ocurren en un tiempo, encontrándose éste distribuido de acuerdo con el modelo de Erlang. En este contexto, $\lambda = {}^1/_\beta$ es la frecuencia constante de ocurrencia y es el tiempo promedio entre dos ocurrencias sucesivas. Análisis basado en Canavos

2.3.9 Distribución Pareto

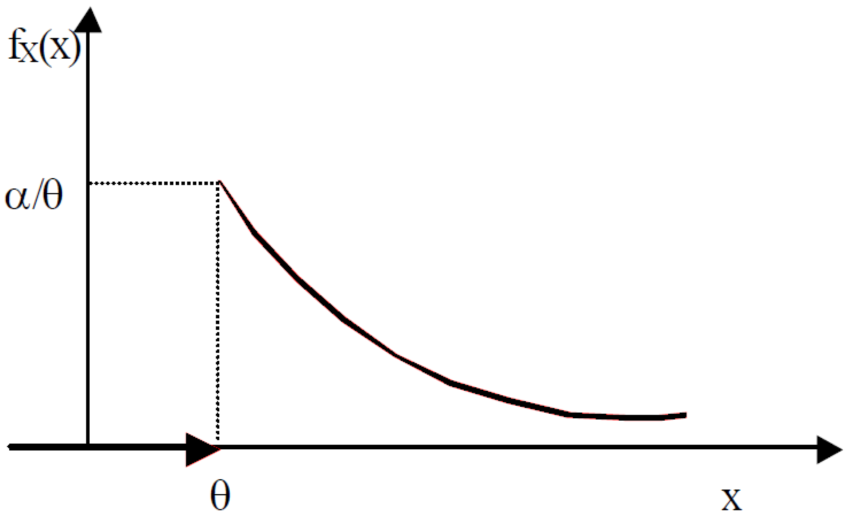
La *densidad de Pareto* se introduce para modelizar la distribución del ingreso cuando ésta es fuertemente inequitativa, y se es común utilizarla cuando el costo de siniestralidad es muy grande y genera una cola muy larga. La forma funcional de la densidad se presenta a continuación:

$$f(x) = f(x) = \begin{cases} \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}} & si\ x \geq \theta \\ 0 & en\ otro\ caso \end{cases}$$

Donde $\alpha > 1$ y $\theta > 0$ (espacio paramétrico)

De la fórmula anterior resulta el siguiente gráfico de la función de densidad:

Representación gráfica de la función de distribución Pareto

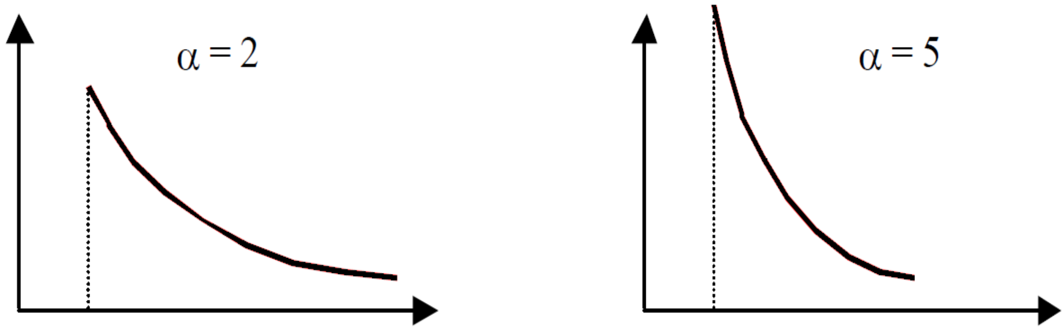


Fuente: O'Reilly, F., & Rueda, R. (1999). Tests of fit for discrete distributions based on the probability generating function.

El parámetro θ puede interpretarse como las primas anuales de un ramo. Se trata de un indicador de posición. Si la población es el conjunto de asegurados en el país, entonces θ es el parámetro de prima anual del mercado de ese ramo.

El parámetro α tiene que ver con la dispersión. A mayores valores de α, se obtienen densidades de Pareto más concentradas en las proximidades del mínimo, es decir, menos dispersas.

Función de distribución Pareto de parámetros diferentes



Fuente: O'Reilly, F., & Rueda, R. (1999). Tests of fit for discrete distributions based on the probability generating function.

La ***función de distribución*** de Pareto puede obtenerse mediante primitivación de la función de densidad.

$$F(x) = f(x) = \begin{cases} 0 & si\ x \leq 0 \\ \int_{\theta}^{x} \frac{\alpha\theta^{\alpha}}{t^{\alpha+1}} dt = -\frac{\theta^{\alpha}}{t^{\alpha}}\Big|_{\theta}^{x} = 1 - \left(\frac{\theta}{x}\right)^{\alpha} & si\ x > 0 \end{cases}$$

Se observa que $1 - F(x) = \left(\theta/x\right)_{\alpha}$ representa la proporción de valores extremos mayores de x.

Calculando las clásicas medidas de posición de la distribución: media, mediana y modo.

$$\mu = E(X) = \int_{\theta}^{\infty} x f(x) dx = \int_{\theta}^{\infty} x \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}} dx = \alpha\theta^{\alpha} \int_{\theta}^{\infty} x^{-\alpha}\, dx = \alpha\theta^{\alpha} \frac{x^{-\alpha+1}}{-\alpha+1}\Big|_{\theta}^{\infty} = \frac{\alpha}{\alpha-1}\theta$$

Para valores grandes de α, la media tiende a aproximarse a θ (por derecha), lo que confirma que la distribución tiende a concentrarse cerca de θ para valores grandes de α. Para que exista la media se requiere que α sea mayor que la unidad.

$$x_{0.5}: F(x_{0.5}) = 0.5 \rightarrow 1 - \left(\frac{\theta}{x_{0.5}}\right)^{\alpha} = 0.5 \rightarrow \frac{\theta}{x_{0.5}} = 0.5^{1/\alpha} \rightarrow x_{0.5} = \theta 2^{-1/\alpha}$$

Para α grande, la mediana tiende a θ.

Como $f(x)$es decreciente a partir de θ, y la densidad es nula hasta θ, entonces el modo de la distribución está en θ.

Para calcular la varianza se aplica la relación entre el segundo momento centrado y los momentos ordinarios

Partiendo de la definición de $V(X) = E(X^2) - E^2(X)$

$$E(X^2) = \int_{\theta}^{\infty} x^2 \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}} dx = \alpha\theta^{\alpha} \int_{\theta}^{\infty} x^{-\alpha+1} dx = \cdots = \frac{\alpha}{\alpha-2}\theta^2 \ \ si\ \alpha > 2$$

$$V(X) = \frac{\alpha}{\alpha-2}\theta^2 - \left(\frac{\alpha}{\alpha-1}\theta\right)^2 = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$$

Entonces:

$$\sigma = \sqrt{\frac{\alpha}{\alpha-2}}\frac{\theta}{\alpha-1}$$

2.3.10 Distribución Weibull

Se trata de una distribución continúa asociado a variables del tipo tiempo de vida, tiempo hasta que un mecanismo falla, etc. La función de densidad de este modelo viene dada por:

$$f(x) = \begin{cases} 0, \ x < 0 \\ \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, \ x \geq 0 \end{cases}$$

Como se observa, depende de dos parámetros: $\alpha > 0$ y $\beta > 0$, donde α es un parámetro de escala y β es un parámetro de forma (lo que proporciona una gran flexibilidad a este modelo).

La función de distribución se obtiene por la integración de la función de densidad y es:

$$F(x) = 1 - e^{-(\frac{x}{\alpha})^{\beta}}$$

2.3.11 Distribución Beta generalizada

La función de densidad beta es una función de densidad de dos parámetros definida sobre el intervalo cerrado $0 \leq y \leq 1$. Frecuentemente se usa como modelo para proporciones.

Se dice que una variable aleatoria Y tiene una distribución de probabilidad beta con parámetros $\alpha > 0$ y $\beta > 0$ si y sólo si la función de densidad de x es

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, 0 \leq x \leq 1 \\ 0\,, \quad en\ otro\ caso \end{cases}$$

Donde

$$B = (\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

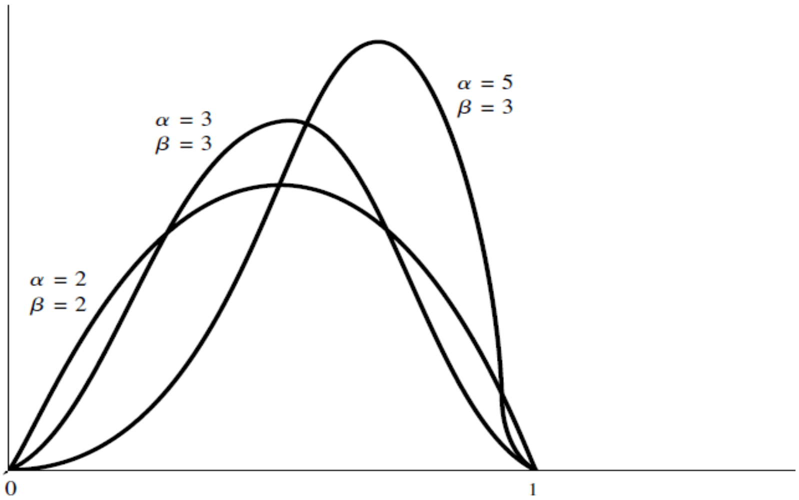
Las gráficas de funciones de densidad beta toman formas muy diferentes para diversos valores de los dos parámetros α y β. Algunos de éstos se muestran a continuación

Si x es una variable aleatoria con distribución beta α > 0 y β > 0, entonces

$$E[X] = \frac{\alpha}{\alpha+\beta}$$

$$V[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Representación gráfica de la función de distribución Beta generalizada



Fuente: Ruiz–Maya, L. (1977). Métodos Estadísticos de investigación

2.3.12 Distribución Uniforme

Se dice que una variable aleatoria X posee una distribución uniforme en el intervalo [a,b],si su función de densidad es la siguiente:

$$f(x) = \frac{1}{b-a}\ si\ a \leq x \leq b$$

La probabilidad de que al hacer un experimento aleatorio, el valor de X esté comprendido en cierto subintervalo de [a, b] depende únicamente de la longitud del mismo, no de su posición.

Cometiendo un pequeño abuso en el lenguaje, se puede decir que en una distribución uniforme la probabilidad de todos los puntos del soporte es la misma.

Por la definición de esperanza se tiene que

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{a}^{b} x\left(\frac{1}{b-a}\right)dx$$

$$= \left(\frac{1}{b-a}\right)\frac{x^2}{2}\Big|_{a}^{b} = \frac{b^2-a^2}{2(b-a)} = \frac{b-a}{2}$$

Así que $E[X] = \frac{b-a}{2}$

Por definición para determinar la varianza se tiene que:

$$V[X] = E(X^2) - [E(X)]^2$$

Por lo que se comienza con desarrollando la esperanza.

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_{a}^{b} x^2 \left(\frac{1}{b-a}\right) dx$$

$$= \left(\frac{1}{b-a}\right) \frac{x^3}{3}\bigg|_{a}^{b} = \frac{b^3-a^3}{3(b-a)} = \frac{(b-a)^2}{3}$$

Regresando al cálculo por definición:

$$V[X] = E(X^2) - [E(X)]^2$$

$$\frac{(b-a)^2}{3} - \left(\frac{b-a}{2}\right)^2 = \frac{4(b-a)^2-3(b-a)^2}{12} = \frac{(b-a)^2}{12}$$

Por lo tanto $V[x] = \frac{(b-a)^2}{12}$

2.4 ESTIMACIÓN

Después de describir las distribuciones estadísticas más importantes desde el punto de vista de los seguros generales, y destacar que cada distribución depende de uno o más parámetros que determinan su tendencia, dispersión e incluso forma; de esta manera, ahora se estudiará la importancia de estimar.

Rara vez son conocidos a priori los parámetros de una distribución, y se necesita estimarlos a partir de los datos disponibles de siniestros producidos antes de poder aplicar la distribución al problema concreto. La función de las observaciones realizadas que se eligen para estimar un parámetro es conocida con el nombre de estimador, y el valor numérico que toma el estimador utilizando una determinada base de datos será una estimación del parámetro.

A menudo se pueden utilizar varias funciones diferentes que darán lugar a los respectivos estimadores, por ejemplo; la media muestra y la mediana muestral, ambas pueden ser utilizadas para estimar la media de la población y se debe decidir cuál utilizar.

2.4.1 Estimadores y sus características

En todo este tema se harán suposiciones que se está estudiando una población cuya distribución es conocida excepto en un parámetro $(\mu, \sigma^2, \lambda, \dots)$ al que se llamará θ. A la distribución de la población se denotará por $f(x)$.

Se dice que se encuentra ante un problema de estimación cuando, dada una población con una distribución $f(x)$ donde θ es un parámetro desconocido, se hará inferencia en base a los datos muestrales $X_1. X_2, \dots, X_n$ el valor de θ. Si al inferir el parámetro se da un único valor se estará ante un problema de estimación puntual. (Rincón, 2007)

Estimador puntual: será una función de la muestra aleatoria (un estadístico) que se utilizará para estimar el valor del parámetro θ.

Estimación: valor obtenido del estimador al sustituir por los valores de una muestra completa.

Un estimador es, por tanto, un estadístico y, por ello, es una variable aleatoria con una determinada distribución de probabilidad llamada distribución muestral.

Dado un parámetro, se podrían utilizar distintos estimadores puntuales para estimarlo. Por ejemplo, para estimar la varianza de la población se puede utilizar la varianza muestral o la cuasi-varianza muestral. ¿Cuál es mejor? Se puede observar a continuación como comprobar si un estadístico es un buen estimador de un parámetro. Para ello se exige el cumplimiento de una serie de propiedades. Como el estadístico es una variable aleatoria, las propiedades se le tienen que exigir a su distribución de probabilidad.

2.4.2 Sesgo y error cuadrático medio

El sesgos el valor medio que puede tomar un estimador para las distintas muestras de igual tamaño que pueden extraerse de una población.

Si se denomina a $\hat{\theta}$ como el estimador del parámetro poblacional θ. Entonces, la media del estimador t se obtendrá como:

$$E(\hat{\theta}) = \frac{\sum \hat{\theta}_i}{C_n^N}$$

Donde $\hat{\theta}_i$ Valor que toma el estadístico en la muestra i, y ${}_N C_n$ (combinatoria de N en n), que representa la cantidad total de muestras de tamaño n que se pueden obtener en una población de tamaño N.

La varianza de un estimador da idea de la variación entre los valores que puede tomar el estimador en las distintas muestras respecto la esperanza o media del estimador.

$$Var(\hat{\theta}) = \frac{\sum[\hat{\theta}_i - E(\hat{\theta})]^2}{C_n^N}$$

La esperanza de la media muestral es el valor medio de los valores que toma el estadístico media muestral para las distintas muestras de n elementos que pueden extraerse de una población de N elementos. Si $\bar{x}$ es el estimador media muestral.

Entonces: $E(\bar{x})$ representa la esperanza de la media muestral

Varianza de la media muestral: Si $\bar{x}$ es el estimador media muestral, entonces $Var(\bar{x})$ representa la varianza de la media muestral

Existen más de una formula ("estimador") para estimar el parámetro θ y entre las diferentes fórmulas (estimadores) se debe elegir aquel que mejores propiedades presente. Las propiedades de los estimadores no requieren el conocimiento de la muestra, sino la fórmula del estimador. (Spiegel, 1991)

El error muestral es la diferencia existente entre el valor del estimador y el valor del parámetro. Para obtener dicho error se utiliza el concepto de Error Cuadrático Medio (MEC) y se define como:

$$ECM = E(\hat{\theta} - \theta)^2$$

2.5 PROPIEDADES DE UN ESTIMADOR PUNTUAL Y MÉTODO DE ESTIMACIÓN

Los estadísticos han ideado diferentes métodos para obtener estimaciones puntuales de los parámetros. Cuando el tamaño de la muestra es grande todos los métodos tienden a dar lugar a las mismas estimaciones.

"El método de máxima verosimilitud normalmente define estimadores que son bastante satisfactorios en relación a una serie de propiedades necesarias que debe cumplir el estimador." (Hossack, I.B.; Pollard, J.H.; Zehnwirth, B, 2001)

2.5.1 Insesgado

Un estimador $\hat{\theta}$ es insesgado sí $E\left[\hat{\theta}\right] = \theta$. Es decir, su distribución está centrada en el parámetro a estimar.

Ejemplos: Sea una muestra aleatoria simple $X_1.X_2, \ldots, X_n$ tal que $E[X_i] = \mu$ y $Var[X_i] = \sigma^2$.

a) Considerando como estimador de la media poblacional a la media muestral. Es decir $\hat{\theta} = \bar{X}_n$. Por el tema anterior, sabemos que $E[\bar{X}_n] = \mu$. Por tanto, la media muestral es estimador insesgado de la media poblacional.

b) Suponiendo como estimador de la varianza poblacional a la varianza muestral $\hat{\sigma}^2 = {S_n}^2$. Del tema anterior, se sabe que $E\left[{S_n}^2\right] = \frac{n-1}{n}\sigma^2$. Por tanto, la varianza muestral no es un estimador insesgado de la varianza poblacional.

c) Considerando ahora como estimador de la varianza poblacional a la cuasi-varianza muestral ,$\hat{\sigma}^2 = {S_n}^2$. Del tema anterior se sabe qué $E\left[{S_n}^2\right] = \sigma^2$. Por tanto, la cuasi-varianza muestral es un estimador insesgado de la varianza poblacional.

2.5.2 Eficiencia relativa

Dados dos estimadores de θ; $\hat{\theta}_1$ y $\hat{\theta}_2$ se dice que $\hat{\theta}_1$es más eficiente que $\widehat{\theta_2}$, si $Var[\hat{\theta}_1] < Var\ [\hat{\theta}_2]$. Nos interesa el que tenga menos dispersión. Para comparar la eficiencia se construye el cociente $\frac{Var[\hat{\theta}_1]}{Var\ [\hat{\theta}_2]}$. Si es mayor que 1, entonces $\hat{\theta}_2$ es más eficiente; si es igual a 1, entonces ambos estimadores son igual de eficientes; si es menor que 1, entonces $\hat{\theta}_1$ es más eficiente.

Por ejemplo Considerando como estimadores de σ^2 a ${S_n}^2$ y ${S_n}^{*2}$. Se Calculas el cociente de las varianzas:

$\frac{Var[{S_n}^2]}{Var[{S_n}^{*2}]} = \frac{\frac{2(n-1)}{n^2}\sigma^4}{\frac{2}{(n-1)}\sigma^4} = \frac{(n-1)^2}{n^2} < 1$ Por tanto, ${S_n}^2$ es más eficiente.

2.5.3 Consistencia

Se dice que $\hat{\theta}$ es un estimador consistente de θ si cumple:

$$\lim_{n\to\infty} E[\hat{\theta}] = \theta \ y \lim_{n\to\infty} Var[\hat{\theta}] = 0 \ \leftrightarrow \ \lim_{n\to\infty} P[\hat{\theta} = \theta] = 1$$

Esto significa que si se toma la mayor muestra posible, el estimador coincidiría con el valor del parámetro.

Por ejemplo obsérvese que los estimadores de los cuales se han hablado en el apartado anterior son consistentes.

a) Si se considera $\hat{\mu} = \bar{X}_n$. Se cumple que $E[\bar{X}_n] = \mu$ y $Var[\bar{X}_n] = \frac{\sigma^2}{n}$. Pero tomando, $\lim_{n\to\infty} \frac{\sigma^2}{n} = 0$. Por tanto, $\bar{X}_n$ es estimador consistente.

b) Si se considera $\hat{\sigma}^2 = {S_n}^2$. Se cumple que:

$$E[{S_n}^2] = \frac{n-1}{n}\sigma^2 \ y \ Var[{S_n}^2] = \frac{2(n-1)}{n^2}\sigma^4$$

Tomando límites,

$$\lim_{n\to\infty}\frac{n-1}{n}\sigma^2 = \sigma^2 \ y \ \lim_{n\to\infty}\frac{2(n-1)}{n^2}\sigma^4 = 0$$

Por tanto, ${S_n}^2$ es un estimador consistente de σ^2.

c) Si se considera $\hat{\sigma}^2 = {S_n}^{*2}$. Se cumple que

$$E\left[{S_n}^{*2}\right] = \sigma^2 \ y \ Var\left[{S_n}^{*2}\right] = \frac{2}{n-1}\sigma^2$$

Tomando límites,

$$\lim_{n\to\infty}\sigma^2 = \sigma^2 \ y \ \lim_{n\to\infty}\frac{2}{n-1}\sigma^4 = 0$$

Por tanto, ${S_n}^{*2}$ es un estimador consistente de la varianza poblacional.

2.5.6 Método de máxima verosimilitud

El método de la máxima verosimilitud permite obtener estimadores que generalmente son satisfactorios. Tienen las deseables propiedades de ser consistentes, asintóticamente normales y asintóticamente eficientes para muestras grandes, en condiciones bastante generales.

Sea X una variable aleatoria con distribución $f(x;\theta)$ donde θ es el parámetro desconocido. Sean $X_1, X_2, \dots, X_n$ de n variables aleatorias independientes con la misma distribución que X; es decir, sea $()(X_1, X_2, \dots, X_n)$ una muestra aleatoria simple bajo estas condiciones la distribución conjunta de las variables aleatorias $X_1, X_2, \dots, X_n$ será igual al producto de las marginales.

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

Si se considera $X_1, X_2, \dots, X_n$ $x_1, x_2, \dots x_n$ fijos y se estudia esta función como función de θ recibe el nombre de función de verosimilitud y se denota por $V(\theta)$.

Sean $\hat{\theta}_1 = u_1(X_1, X_2, \dots, X_n), \hat{\theta}_2 = u_2(X_1, X_2, \dots, X_n), etc.$ diversos estimadores de θ. De todos ellos se pretende elegir el que haga máxima la función de verosimilitud. Es decir, un estimador $\hat{\theta}$ será estimador máximo-verosímil (EMV) de θ si maximiza $V(\theta)$.

Debido a que la función de verosimilitud es no negativa, continua y creciente, alcanzará su máximo en los mismos puntos que su logaritmo y por ello, y por razones de cálculo se suele maximizar $\ln V(\theta)$ cuando esta depende de exponenciales .Así pues, se deberá resolver la siguiente ecuación:

$$\frac{d(\ln V(\theta))}{d\theta} = 0$$

En el caso de dos o más parámetros desconocidos, el procedimiento es el mismo. Por ejemplo, si se tuviera $V(\theta_1, \theta_2, \theta_3)$ $V(\theta_1, \theta_2, \theta_3)$ los tres estimadores máximo verosímiles serán los que maximizan la función $V(\theta_1, \theta_2, \theta_3)$ $V(\theta_1, \theta_2, \theta_3)$ o su logaritmo. Se obtendrían al resolver las ecuaciones siguientes:

$$\frac{d(\ln V(\theta_1,\theta_2,\theta_3))}{d\theta_1}$$

$$\frac{d(\ln V(\theta_1,\theta_2,\theta_3))}{d\theta_2}$$

$$\frac{d(\ln V(\theta_1,\theta_2,\theta_3))}{d\theta_3}$$

Las propiedades de los estimadores son aplicables en el método de máxima verosimilitud, de tal manera que para la obtención de dicho método se deben cumplir con estos aspectos:

a) Son consistentes.

b) Son asintóticamente eficientes. Es decir, tienen la varianza mínima cuando el tamaño muestral tiende a infinito.

c) Si $\hat{\theta}$ es estimador suficiente de θ , el EMV de θ es función de $\hat{\theta}$.

d) Son asintóticamente normales. Es decir, su distribución tiende a la distribución normal cuando tiende a infinito el tamaño de la muestra.

e) Si $\hat{\theta}$ es EMV de θ , entonces $g(\hat{\theta})$ es EMV de $g(\theta)$, siendo g una aplicación biyectiva.

Por ejemplo para obtener el EMV del parámetro de una variable aleatoria X que sigue una distribución de $Bernouilli$, $X \rightarrow Be(p)$. Su función de cuantía es

$$f(x;p) = p^x(1-p)^{1-x}$$

Si se elige una muestra de tamaño n, la función de verosimilitud correspondiente será:

$$V(p) = f(x_1;p)\cdot f(x_2;p)\cdot \ldots \cdot f(x_n;p)$$

$$= p^{x_1}(1-p)^{1-x_1}\cdot p^{x_2}(1-p)^{1-x_2}\cdot \ldots \cdot p^{x_n}(1-p)^{1-x_n}$$

$$= p^{\sum x_i}(1-p)^{n-\sum x_i}$$

Tomando logaritmos se tiene:

$$\ln V(p) = \sum x_i \ln p + \left(n \sum x_i\right)\ln(1-p)$$

Para obtener el EMV de p se debe resolver la ecuación

$$\frac{d(\ln p)}{dp} = 0$$

En este caso

$$\frac{d(\ln V(p))}{dp} = \sum x_i \frac{1}{p} + (n - \sum x_i)\frac{(-1)}{(1-p)} = 0$$

Haciendo operaciones e igualando denominadores se obtiene:

$$(1-p)\sum x_i - p(n - \sum x_i)$$

$$= \sum x_i - p\sum x_i - np + p\sum x_i = \sum x_i - np = 0$$

Despejando el valor de p se obtiene el estimador

$$\hat{\mu} = \frac{\sum x_i}{n}$$

Por tanto $EMV(p) = \frac{\sum x_i}{n}$. Es decir, la proporción de éxitos de la muestra.

2.6 PRUEBA DE HIPÓTESIS Y SIGNIFICANCIA

La estadística inferencial es el proceso de usar la información de una muestra para describir el estado de una población. Sin embargo, es frecuente que se use la información de una muestra para probar un reclamo o conjetura sobre la población. El reclamo o conjetura se refiere a una *hipótesis*. El proceso que corrobora si la información de una muestra sostiene o refuta el reclamo se llama *prueba de hipótesis*.

En inferencia estadística, hay dos problemas centrales, la estimación de parámetros y las pruebas de hipótesis. Una hipótesis estadística es una suposición que se hace acerca de la distribución de una variable aleatoria

Una prueba estadística de una hipótesis es un procedimiento en el cual se usa una muestra con el fin de determinar cuándo se puede rechazar la hipótesis, es decir, actuar como si fuera falsa o cuándo no se debe rechazar esta, es decir, actuar como si fuera verdadera.

Con mucha frecuencia se tiene que tomar decisiones en las que el azar juega un papel muy importante. Si se debe elegir entre dos posibilidades, la decisión entre ambas se debería basarse en el resultado de una prueba estadística.

Se asume que se tiene una muestra aleatoria $X_1,\ldots,X_n$ de alguna función de densidad $f(x;\,\theta)$. Una hipótesis estadística será una hipótesis sobre la distribución de la población. (Rincón, 2007)

Una hipótesis estadística es una aseveración o conjetura acerca de una distribución de una o más variables aleatorias. Si la hipótesis estadística está totalmente especificada se tiene el llamado caso simple, en otro caso es referido como caso compuesto.

Se denota a la hipótesis estadística, por la letra H, seguida de dos puntos que a su vez es seguida por la afirmación que especifica la hipótesis.

Una prueba de una hipótesis estadística H es una regla o procedimiento para decidir o no, si se rechaza H.

En la prueba de hipótesis se pone a prueba un reclamo hecho sobra la naturaleza de una población a base de la información de una muestra. El reclamo se llama hipótesis estadística.

Una *hipótesis estadística* es un reclamo hecho sobre la naturaleza de una población.

Por ejemplo, la premisa formulada por una compañía aseguradora de que su seguro de incendio se siniestra con una frecuencia promedio de 48 horas, es una hipótesis estadística.

Si surgieran reclamos de siniestros de parte de los clientes, entonces se pone a prueba el reclamo del siniestro. La hipótesis estadística sometida a prueba se llama la *hipótesis nula*, y se denota como H_0**.**

2.6.1 Elementos de prueba de hipótesis

La Hipótesis Nula (H_0) es la premisa, reclamo, o conjetura que se pronuncia sobre la naturaleza de una o varias poblaciones.

Por ejemplo, para aprobar o desaprobar la frecuencia de la siniestralidad del seguro de daños se debe probar la hipótesis estadística de que $\mu \geq 48$. Por lo tanto, la hipótesis nula es:

$$H_0 : \mu \geq 48.$$

Luego se procede a tomar una muestra aleatoria de siniestros y medir su frecuencia media. Si la información obtenida de la muestra no apoya el reclamo en la hipótesis nula (H_0), entonces otra cosa es cierta. La premisa alterna a la hipótesis nula se llama *hipótesis alterna* y se representa por H_1.

Hipótesis Alterna es una premisa que es cierta cuando la hipótesis nula es falsa.

Por ejemplo, para la frecuencia de siniestros:

$$\boldsymbol{H_0} \geq 48 \; y \; H_1 < 48$$

Para probar si la hipótesis nula es cierta, se toma una muestra aleatoria y se calcula la información, como el promedio, la proporción, etc. Esta información muestral se llama *estadística de prueba*.

Una estadística de prueba se basa en la información de la muestra como la media o la proporción.

A base de la información de una muestra es común cometer dos tipos de errores en nuestra decisión.

1. Se puede rechazar un $\boldsymbol{H_0}$ que es cierto.
2. Se puede aceptar un $\boldsymbol{H_0}$ que es falso.

El primero se llama error tipo 1, cuando se rechaza una Hipótesis Nula que es cierta se comete error tipo 1.

Y el segundo error se llama error tipo 2, cuando se acepta una Hipótesis Nula que es falsa se comete error tipo 2.

Para ser muy cuidadosos en no cometer el error tipo 1, se debe especificar la probabilidad de rechazar H_0, denotada por α. A ésta se le llama *nivel de significancia.*

El nivel de Significancia es la probabilidad ($\boldsymbol{\alpha}$) más alta de rechazar H_0 cuando H_0 es cierto.

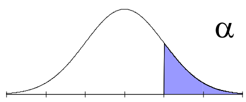
Para mantener la probabilidad de cometer el error tipo 1 baja, se debe escoger un valor pequeño de ($\boldsymbol{\alpha}$).

Usando un valor preasignado de $\boldsymbol{\alpha}$ se construye una región de rechazo o región crítica en la curva normal estándar o en la curva t que indica si se debe rechazar H_0.

Una región crítica o de rechazo es una parte de la curva de z o de la curva t donde se rechaza H_0.

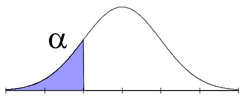
La región puede ser de una cola o de dos dependiendo de la hipótesis alterna.

Para $H_1\colon \mu\ >$ valor aceptado, la región de rechazo está dada por:



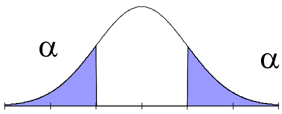
(Cola derecha, z o t)

Para H_1: $\mu <$ valor aceptado, la región de rechazo está dada por:



(Cola izquierda, z o t)

Para H_1: $\mu \neq$ valor aceptado, la región de rechazo es de dos colas y está dada por:



(2-colas, z o t)

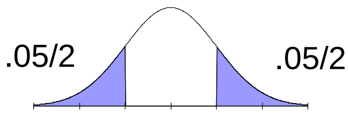
Para otro caso**;** Determine si la región de rechazo es de la cola derecha, de la cola izquierda o de dos colas.

a. $H_0 : \mu = 15, H_1 : \mu \neq 15, \alpha = .05$

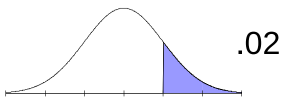
b. $H_0 : p \leq 0.7, H_1 : p > 0.7, \alpha = .02$

Solución: La forma de la región de rechazo está determinada por la hipótesis alterna.

$H_1 : \mu \neq 15$ significa que la región está en ambas colas.



$H_1 : p > 0.7,$ significa que la región está en la cola derecha.

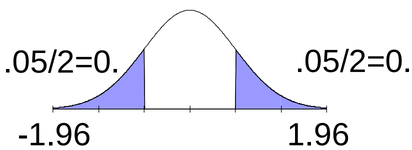


En el Ejemplo 1a, se presume que la región de rechazo es parte de la curva normal estándar. Complete el dibujo de la región crítica para los valores α siguientes:

a. α = .05

La solución es la siguiente:

a. Del ejemplo 1(a), se tiene que:



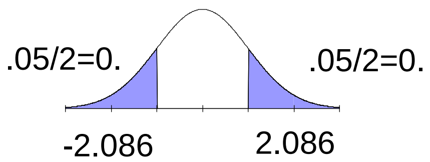
De la tabla de la distribución normal, la $P(Z < z) = 0.025$ corresponde a un valor $Z = -1.96$. Por simetría la $P(Z > z) = .025$ corresponde a $Z = 1.96$.

En el ejemplo 1a, se presume que la región de rechazo es parte de la curva t. Complete el dibujo de la región de rechazo para:

$\alpha = .05\, y\, v = 14$

La solución es la siguiente:

a. Del ejemplo 1(a), $\alpha = .05\, y\, v = 14$, tenemos:



De la tabla de la distribución t, la $P(T<t) = 0.025$ corresponde a un valor $t = -2.086$. Por simetría la $P(T > t) = 0.025$ corresponde a $t = 2.086$.

En el último ejemplo establezca las hipótesis nula y alterna.

a. Las millas por galón promedio de un nuevo modelo de automóvil es 32.

$$H_0: \mu = 32, H_1 : \mu \neq 32$$

b. Más del 65% de los empleados de un colegio aportan a Fondos Unidos.

$$H_0: p \geq .65, H_1: p < .65$$

c. En promedio, los empleados de cierta compañía viven a no más de 15 millas de la misma.

$$H_0: \mu \leq 15, H_1 : \mu > 15$$

d). Al menos un 60% de la población adulta de una comunidad votará en las próximas elecciones Presidenciales.

$$H_0: p \geq .6, H_1: p < .6$$

e). El peso promedio de un pollo para asar es de al menos cuatro libras.

Solución:

$$H_0: \mu \geq 4, H_1 : \mu < 4$$

2.7 PRUEBAS DE BONDAD DE AJUSTE

Como menciono anteriormente En inferencia estadística, con un conjunto de datos muéstrales, se plantea inferir sobre la población. Esas inferencias o decisiones que se pueden tomar son referentes a estimación de parámetros o prueba de hipótesis. Por ejemplo, cuando se hace el supuesto sobre la variable de interés asociada a los elementos de una población, que sigue una distribución particular.

Estos supuestos pueden ser ciertos o no, los cuales se llaman hipótesis estadísticas. Al procedimiento mediante el cual se investiga la verdad o falsedad de una hipótesis se le llama prueba de hipótesis. Las pruebas de bondad de ajuste son un caso particular de pruebas de hipótesis; la particularidad radica que en las pruebas de hipótesis clásicas, se supone que tanto en la hipótesis nula como en la hipótesis alternativa es conocida la forma funcional de la densidad; y en el caso de las pruebas de bondad de ajuste se conoce la forma funcional en la hipótesis nula, pero en la hipótesis alternativa se desconoce la función de densidad, y por ello, el conjunto de posibilidades puede ser muy grande.

En un fenómeno de interés, es importante verificar la suposición de que los datos experimentales se ajustan, o no, a cierto modelo probabilístico. En este caso se está ante un problema de bondad de ajuste. Las pruebas de bondad de ajuste tienen por objetivo determinar si los datos se generaron o se pueden modelar con una determinada función distribución.

Una teoría de bondad de ajuste para el caso continuo muy desarrollada es la que se basa en la función de distribución empírica, y para el caso discreto, existe un símil.

Con las facilidades de computo que se tienen en la actualidad se ha propuesto, además del uso del remuestreo paramétrico (bootstrap), la simulación de valores de la estadística de prueba, pero la simulación no se hace con la función de distribución de la estadística de prueba (que depende de los parámetros) sino que se hace con la distribución condicional dada una estadística suficiente

Ajuste de curvas a funciones de distribución

En la construcción del modelo de simulación es importante decidir si un conjunto de datos se ajusta apropiadamente a una distribución específica de probabilidad. Al probar la bondad del ajuste de un conjunto de datos, se comparan las frecuencias observadas realmente en cada categoría o intervalo de clase con las frecuencias esperadas teóricamente.

2.7.1 Generalidades de las pruebas de bondad de ajuste

Una familia paramétrica de funciones de distribución es el conjunto de funciones que depende de un parámetro θ; el parámetro puede ser un número real o un vector y la familia se denota como (González, 2007):

$$F_0 = \{F(x;\, \theta) : \theta \in \boldsymbol{\Theta}\}.$$

Donde Θ es una familia paramétrica de $\theta_1, \ldots, \theta_n$

Dadas las observaciones $(X_1, \ldots, X_n)$ independientes, con distribución F, se desea probar la hipótesis nula $H_0: F = F_0$. En principio, la hipótesis alternativa será $H: F \neq F_0$, pero es posible que dentro de esta alternativa múltiple haya algunas distribuciones para las que nos interese especialmente que la prueba tenga una buena potencia.

A la hipótesis H_0 se la llama hipótesis de ajuste de la distribución F_0 al modelo del cual proviene la muestra. Las pruebas de H_0 se llaman pruebas de ajuste.

“A lo largo del Siglo XIX, los modelos aleatorios se volvieron cada vez más frecuentes y cada vez más necesarios para describir la naturaleza. Un modelo se consideraba adecuado en tanto no presentara incoherencias evidentes con los resultados de la experiencia." (González, 2007)

Recién en 1999 surgió la primera prueba de ajuste, a partir de la cual los científicos pudieron poner a prueba sus modelos e incluso seleccionar entre varios modelos propuestos para un mismo fenómeno, cuales con adecuados y cuales no lo son.

Para decidir si se rechaza $H_0\colon F = F_0$ a partir de la información dada por la muestra aleatoria simple $X_0,\ldots,X_n$ de F, resulta natural estimar F por medio de la muestra, y comparar la estimación con F_0.

El estimador de máxima verosimilitud de F es la distribución de probabilidades $\widehat{F}$ para la que, si $Y_1,\ldots,Y_n$ es una muestra de F, entonces la probabilidad de que resulte $(Y_1,\ldots,Y_n) = (\mathrm{X}_1,\ldots,\mathrm{X}_n)$ es máxima.

Sea X una variable aleatoria poblacional $f_0(x)$ la distribución (o densidad) de probabilidad especificada o supuesta para X

Se desea probar la hipótesis:

$$H_0\colon f(x) = f_0(x)$$

En contraste con la hipótesis alterna:

$$H_1\colon f(x) \neq f_0(x)(\text{negación de } \mathrm{H}_0)$$

2.7.2 Prueba de ji-cuadrada

La prueba Ji cuadrada hace uso de la distribución del mismo nombre para probar la bondad del ajuste al comparar el estadístico de prueba $\chi_0{}^2$ con el valor en tablas de la mencionada distribución Ji cuadrada con *v* grados de libertad y un nivel de significancia α.

Esta prueba es aplicable para variables aleatorias discretas o continuas. Sea una muestra aleatoria de tamaño n tomada de una población con una distribución especificada $f_0(x)$ que es de interés verificar. (Beltrán, 2012)

Suponer que las observaciones de la muestra están agrupadas en k clases, siendo fo_i la cantidad de observaciones en cada clase $i = 1, 2, \ldots, k$

Con el modelo especificado $f_0(x)$ se puede calcular la probabilidad p_i que un dato cualquiera pertenezca a una clase i.

Con este valor de probabilidad se puede encontrar la frecuencia esperada fe_i para la clase i, es decir, la cantidad de datos que según el modelo especificado deberían estar incluidos en la clase i: $fe_i = p_i\, n\,;\ i = 1, 2, \ldots, k$

Se tiene entonces dos valores de frecuencia para cada clase i

fo_i es la frecuencia observada (corresponde a los datos de la muestra)

fe_i es la frecuencia esperada (corresponde al modelo propuesto)

La teoría estadística demuestra que la siguiente variable es apropiada para realizar una prueba de bondad de ajuste:

Se define al estadístico para la prueba de bondad de ajuste Ji-cuadrado como:

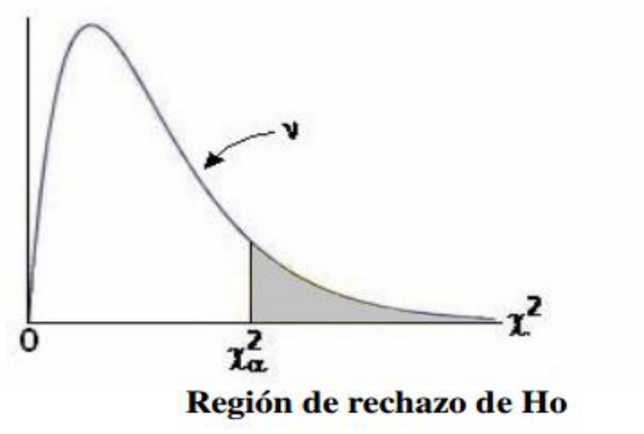
$\chi^2 = \sum_{i=1}^{k} \frac{(fo_i - fe_i)^2}{fe_i}$, distribución Ji-cuadrado con $\nu = k - r - 1$ grados de libertad donde r es la cantidad de parámetros de la distribución que deben estimarse a partir de la muestra

Es una condición necesaria para aplicar esta prueba que $\forall i, fe_i \geq 5$

Dado un nivel de significancia α se define un valor crítico ${\chi_\alpha}^2$ para el rechazo de la hipótesis propuesta H_0: $f(x) = f_0(x)$.

Si las frecuencias observadas no difieren significativamente de las frecuencias esperadas calculadas con el modelo propuesto, entonces el valor de estadístico de prueba χ^2 será cercano a cero, pero si estas diferencias son significativas, entonces el valor del estadístico χ^2 estará en la región de rechazo de H_0

Representación gráfica de la región de rechazo de una prueba



Fuente: González, E. (2007). Pruebas de bondad de ajuste para distribuciones estables.

Para ejemplificarlo se tomará el siguiente caso. Se ha tomado una muestra aleatoria de montos de siniestros y se ha registrado su factor de siniestralidad diaria. Estos resultados se los ha agrupado en 7 clases en el siguiente cuadro

Tabla de frecuencias para prueba de bondad ji-cuadrada

i	$clase(duración)$	$frecuencia\ observada\ fo_i$
1	1.45-1.95	2
2	1.95-2.45	1
3	2.45-2.95	4
4	2.95-3.45	15
5	3.45-3.95	10
6	3.95-4.45	5
7	4.45-4.95	3

Fuente: Villers, S. (2001). Programa para cálculo de p-values de estadísticas EDF para pruebas de bondad de ajuste.

Verificar con 5% de significancia que factor de siniestralidad de un seguro de incendio en una compañía se seguros se distribuye normalmente con media 3.5 y desviación estándar 0.7

Entonces sea X la frecuencia diaria de siniestralidad (variable aleatoria) se deben considerar los siguientes puntos.

1. H_0:X se comporta como una distribución normal con $\mu = 3.5\ y\ \sigma = 0.7$
2. H_1: hipótesis alternativa que rechaza H_0
3. El nivel de significancia es $\alpha = .05$

Utilizando la función estándar de la distribución normal $z = \frac{x-\mu}{\sigma}$ se realizan los cálculos.

$p_1 = P(X \leq 1.95) = P(Z \leq (1.95 - 3.5)/0.7) = 0.0136$

$p_2 = P(1.95 \leq X \leq 2.45) = P((1.95 - 3.5)/0.7 \leq Z \leq (2.45 - 3.5)/0.7) = 0.0532$

$p_3 = P(2.45 \leq X \leq 2.95) = P((2.45 - 3.5)/0.7 \leq Z \leq (2.95 - 3.5)/0.7) = 0.135$

... (etc.)

Por otra parte, se calcula de las frecuencias esperadas

$$e_1 = p_1\, n = 0.0136\,(40) \approx 0.5$$

$$e_2 = p_2\, n = 0.0532\,(40) \approx 2.1$$

$$e_3 = p_3\, n = 0.135\,(40) \approx 5.4\ \ldots\ (etc)$$

Los resultados se resumen en la siguiente tabla:

Tabla de frecuencias para prueba de bondad ji-cuadrada

$montos(diarios)$	$frecuencia\ observada\ fo_i$	$frecuencia\ esperada\ fe_i$
1.45-1.95	2	0.5
1.95-2.45	1	2.1
2.45-2.95	4	5.4
2.95-3.45	15	10.3
3.45-3.95	10	10.7
3.95-4.45	5	7
4.45-4.95	3	3.5

Fuente: Villers, S. (2001). Programa para cálculo de p-values de estadísticas EDF para pruebas de bondad de ajuste.

En el caso de la frecuencia esperada $n = 40$, redondeando los decimales de cada una de las clases esto es correcto.

Es necesario que se cumpla la condición $\forall i, e_i \geq 5$ por lo que se deben agrupar clases adyacentes. Como resultado se tienen cuatro clases $k = 4$

Tabla simplificada de frecuencias para prueba de bondad ji-cuadrada

$montos(diarios)$	$frecuencia\ observada\ fo_i$	$frecuencia\ esperada\ fe_i$
1.45-2.95	7	8.5
2.95-3.45	15	10.3
3.45-3.95	10	10.7
3.95-4.95	8	10.5

Fuente: Villers, S. (2001). Programa para cálculo de p-values de estadísticas EDF para pruebas de bondad de ajuste.

Ahora se puede definir la región de rechazo de H_0

Observando este ejemplo, la media y la desviación estándar de la distribución normal no se estimaron, sino que están propuestas, de donde $r = 0$

$\alpha = 0.05, v = k - 1 = 3 \leftrightarrow {\chi_{0.05}}^2 = 7.815$ (valor se encuentra en la tabla de χ^2)

Recordando la propuesta de la hipótesis, se rechaza H_0 si $\chi^2 > 7.815$

Al realizar el cálculo del estadístico de prueba

$$\chi^2 = \sum_{i=1}^{k} \frac{(fo_i - fe_i)^2}{fe_i} = \left[\frac{(7-8.5)^2}{8.5} + \frac{(15-10.3)^2}{10.3} + \frac{(10-10.7)^2}{10.7} + \frac{(8-10.5)^2}{10.5}\right] = 3.05$$

Como $3.05 < 7.815$, se dice que no hay evidencia suficiente para rechazar el modelo propuesto para la población.

Para ejemplificar una distribución discreta se usará el caso de una Poisson.

La siguiente tabla presenta información de cantidades sobre el número de plantas Larrea divaricata halladas en cada uno de los 48 cuadrantes de nuestro desierto fronterizo, como se publica en el artículo “Some Sampling Characteristics of Plants and Arthropods of the Arizona Desert” (Ecology,1962: 567-571)

Tabla de frecuencias para prueba de bondad ji-cuadrada

i	$numero\ de\ plantas$	$frecuencia\ observada\ fo_i$
1	0	9
2	1	9
3	2	10
4	3	14
5	4	2
6	5	2
7	6	2

Fuente: Villers, S. (2001). Programa para cálculo de p-values de estadísticas EDF para pruebas de bondad de ajuste.

y se desea determinar datos se ajustan a una distribución de Poisson, Utilizando un nivel 0.05 de significancia.

El valor de λ en este caso debe estimarse

$$\lambda = \frac{\sum x_i \cdot fo_i}{n} = \frac{101}{48} = 2.10$$

Se declara la hipótesis

1. H_0:X se comporta como una distribución Poisson con λ= 2.10
2. H_1: hipótesis alternativa que rechaza H_0
3. El nivel de significancia es $\alpha = .05$

Se calculan la probabilidad correspondiente a cada intervalo

$p_1 = P(X = 0) = \frac{e^{-2.1}2.1^0}{0!} = e^{-2.1}$

$p_2 = P(X = 1) = \frac{e^{-2.1}2.1^1}{1!} = 0.25725$

… (etc)

Ahora se Calcula de las frecuencias esperadas

$fe_1 = p_1\, n = e^{-2.1}\,(48) = 5.88$

$fe_2 = p_2\, n = (.25725)(48) = 12.34$ ... (etc.)

Tabla simplificada de frecuencias para prueba de bondad ji-cuadrada

$numero\ de\ plantas$	$frecuencia\ observada\ fo_i$	$frecuencia\ observada\ fe_i$
0	9	5.88
1	9	12.34
2	10	12.96
3	14	9.07
4	2	7.75

Fuente: Villers, S. (2001). Programa para cálculo de p-values de estadísticas EDF para pruebas de bondad de ajuste.

Recordando que la condición $\forall i, e_i \geq 5$, se tiene $k = 5$

Ahora se puede definir la región de rechazo de H_0, se observa que en este ejemplo se estimó el parámetro de la distribución, de donde $r = 1$

$\alpha = 0.05, \nu = 5 - 1 - 1 = 3 \leftrightarrow \chi_{0.05}{}^2 = 7.815$ (valor se encuentra en la tabla de χ^2)

Recordando la propuesta de la hipótesis, se rechaza H_0 si $\chi^2 > 7.815$

Al realizar el cálculo del estadístico de prueba

$$\chi^2 = \sum_{i=1}^{k} \frac{(fo_i - fe_i)^2}{fe_i} = \left[\frac{(9-5.88)^2}{5.88} + \cdots + \frac{(6-7.75)^2}{7.75}\right] = 6.31$$

Como $6.31 < 7.815$, se dice que no hay evidencia suficiente para rechazar el modelo propuesto para la población, de modo que al nivel de 5%, la distribución de Poisson da un ajuste razonable a los datos. (Villers, 2001)

2.7.3-Estadística kolmogorov-Smirnov

De igual manera que en la prueba de ji cuadrada se tiene las hipótesis de $H_0\ y\ H_1$

H_0: Los datos analizados siguen una distribución M.

H_1: Los datos analizados no siguen una distribución M

Donde el estadístico de contrates está dado por:

$$D = sup\left|\hat{F}_n(x_i) - F_0(x_i)\right|$$

x_i es el $i - ésimo$ valor observado en la muestra (cuyos valores se han ordenado previamente de menor a mayor).

$\hat{F}_n(x_i)$ es un estimador de la probabilidad de observar valores menores o iguales que x_i.

$F_0(x)$ es la probabilidad de observar valores menores o iguales que x_i cuando H_0 es cierta.

Así pues, D es la mayor diferencia absoluta observada entre la frecuencia acumulada observada $\hat{F}_n(x_i)$ y la frecuencia acumulada teórica $F_0(x)$, obtenida a partir de la distribución de probabilidad que se especifica como hipótesis nula.

Si los valores observados $\hat{F}_n(x_i)$ son similares a los esperados $F_0(x)$, el valor de D será pequeño. Cuanto mayor sea la discrepancia entre la distribución empírica ˆ $\hat{F}_n(x_i)$ y la distribución teórica, mayor será el valor de D. (González, 2007)

Por tanto, el criterio para la toma de la decisión entre las dos hipótesis será de la forma:

$$Si\ D \leq D_\alpha \ \Rightarrow \ No\ se\ rechaza\ H_0$$

$$Si\ D > D_\alpha \ \Rightarrow \ Rechazar\ H_0$$

Siendo α el nivel de significancia

Para el cálculo práctico del estadístico D deben obtenerse:

$$D^+ = max\left\{\frac{i}{n} - F_0(x_i)\right\}, D^- = max\left\{F_0(x_i) - \frac{i-1}{n}\right\}$$

y a partir de estos valores:

$$D = max\{D^+, D^-\}$$

A su vez, el valor de D_α depende del tipo de distribución a probar y se encuentra tabulado. En general es de la forma:

$$D_\alpha = \frac{c_\alpha}{k(n)}$$

Donde c_α y $k(n)$ se encuentran en las tablas siguientes:

Tabla de nivel de significancia para prueba Kolmogorov-Smirnov

C_α	α		
Modelo	**0.1**	**0.05**	**0.01**
General	1.224	1.358	1.628
Normal	0.819	0.895	1.035
Exponencial	0.99	1.094	1.308
Weibull n=10	0.76	0.819	0.944
Weibull n=20	0.779	0.843	0.973
Weibull n=50	0.79	0.856	0.988
Weibull n= ∞	0.803	0.874	1.007

Fuente: González, E. (2007). Pruebas de bondad de ajuste para distribuciones estables.

Tabla de distribuciones para prueba Kolmogorov-Smirnov

Distribución que se contrasta	$\boldsymbol{k(n)}$
General.*parametros conocidos*	$k(n) = \sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}$
Normal	$k(n) = \sqrt{n} + 0.01 + \frac{0.85}{\sqrt{n}}$
Exponencial	$k(n) = \sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}$
Weibull	$k(n) = \sqrt{n}$

Fuente: González, E. (2007). Pruebas de bondad de ajuste para distribuciones estables.

Se entenderá mejor con el siguiente caso

Determinar si los valores 6, 2.3, 4.8, 5.6, 4.5, 3.4, 3.3, 1.9, 4.8, 4.5 conforman a una distribución normal con $\mu = 4.1\ y\ \sigma^2 = 1.82$

Para determinar el valor D se crea la siguiente tabla, dicha tabla ayudará a que los cálculos se realicen en menor tiempo.

Tabla de frecuencias para prueba Kolmogorov-Smirnov

$\boldsymbol{x}$	$\boldsymbol{x_i}$*(ordenados)*	$\boldsymbol{i}$	$\boldsymbol{z = \frac{x_i - \mu}{\sigma}}$	$\boldsymbol{F_0}$	$\boldsymbol{D^+}$	$\boldsymbol{D^-}$
6	1.9	1	-1.631	0.051	0.049	0.051
2.3	2.3	2	-1.334	0.091	0.109	-0.009
4.8	3.3	3	-0.593	0.276	0.024	0.076
5.6	3.4	4	-0.519	0.302	0.098	0.002
4.5	4.5	5	0.296	0.616	-0.116	0.216
3.4	4.5	6	0.296	0.616	-0.016	0.116
3.3	4.8	7	0.519	0.698	0.002	0.098
1.9	4.8	8	0.519	0.698	0.102	-0.002
4.8	5.6	9	1.112	0.867	0.033	0.067
4.5	6	10	1.408	0.92	0.08	0.02

Fuente: González, E. (2007). Pruebas de bondad de ajuste para distribuciones estables.

El primer paso es ordenar los datos de menor a mayor, después se obtiene el valor de z ya que se conoce μ y σ .

Posteriormente se debe encontrar valor obtenido de z en la tabla de la normal el cual se empleará como F_0

Y finalmente se obtiene $D^+ \ y \ D^-$:

$$D^+ = max\left\{\frac{i}{n} - F_0(x_i)\right\} , D^- = max\left\{F_0(x_i) - \frac{i-1}{n}\right\}$$

Que en el caso anterior $D^+ = 0.109 \ y \ D^- = 0.216$

Obtenidos estos valores se establece D

$$D = max\{D^+, D^-\} = max\{0.109,0.216\} = 0.216$$

Por otra parte se determina D_α, como los datos observados constituyen una distribución normal y el nivel de significancia es $\alpha = 0.05$, entonces $C_{0.05} = 0.895$ y $k(10) = \sqrt{10} + 0.01 + \frac{0.85}{\sqrt{10}}$

Por lo tanto:

$$D_\alpha = \frac{c_\alpha}{k(n)} = \frac{0.895}{k(10)=\sqrt{n}+0.01+\frac{0.85}{\sqrt{10}}} = \frac{0.895}{3.42} = 0.262$$

Recordando el criterio para la tomar la decisión como el valor $D = 0.216 \leq 0.262$, no se rechaza H_0 y se acepta que los datos se distribuyen normalmente

Para otro caso continuo Considérese que las diez observaciones siguientes son una muestra aleatoria de una distribución continua. Probar la hipótesis de que esos datos provienen de una distribución exponencial con promedio 2, en el nivel de significación 0.05.

Los datos son: 0.406, 2.343, 0.538, 5.088, 5.587, 2.563, 0.023, 3.334, 3.491, 1.267

Replicando el proceso anterior, se ordenan las diez observaciones ascendentemente y entonces se calcula, para cada x_i, el valor de $F_0(x_i)$, donde H_0 establece que $F(x)$ es exponencial con $\boldsymbol{\lambda} = 2$. Por lo tanto, $F(x) = 1 - e^{-x_i}/_2$

Tabla de frecuencias ordenadas para prueba Kolmogorov-Smirnov

x	x_i(ordenados)	i	$F_0 = 1 - e^{-x_i/2}$	D^+	D^-
0.406	0.023	1	0.011	0.0886	0.0114
2.343	0.406	2	0.184	0.0163	0.0837
0.538	0.538	3	0.236	0.0641	0.0359
5.088	1.267	4	0.469	-0.0693	0.1693
5.587	2.343	5	0.690	-0.1901	0.2901
2.563	2.563	6	0.722	-0.1224	0.2224
0.023	3.334	7	0.811	-0.1112	0.2112
3.334	3.491	8	0.825	-0.0254	0.1254
3.491	5.088	9	0.921	-0.0214	0.1214
1.267	5.587	10	0.939	0.0612	0.0388

Fuente: González, E. (2007). Pruebas de bondad de ajuste para distribuciones estables.

Obtenidos estos valores se establece D

$$D = max\{D^+, D^-\} = max\{0.0886, 0.2901\} = 0.2901$$

Por otra parte se determina D_α, como los datos observados constituyen una distribución exponencial y el nivel de significancia es $\alpha = 0.05$, entonces $C_{0.05} = 1.094$ y $k(10) = \sqrt{10} + 0.12 + \frac{0.11}{\sqrt{10}}$

Por lo tanto:

$$D_\alpha = \frac{c_\alpha}{k(n)} = \frac{0.895}{k(10) = \sqrt{n} + 0.12 + \frac{0.11}{\sqrt{10}}} = \frac{1.094}{3.317} = 0.329$$

Recordando el criterio para la tomar la decisión como el valor $D = 0.2901 \leq 0.329$, no se rechaza H_0 y se acepta que los datos se distribuyen normalmente

2.8-RESUMEN

En este capítulo se utilizó un lenguaje técnico actuarial.

El propósito en este apartado fue aprender a identificar de manera teórica y práctica, varios elementos probabilísticos que servirán para explicar algunos procesos estadísticos.

En necesario realizar un análisis estadístico para interpretar el comportamiento de los datos que se van a estudiar, por lo que es indispensable tener muy presente los conocimientos relacionados a la materia mencionada.

Partiendo de ahí, se recapitularon algunos temas requeridos de dicha materia, iniciando por las variables aleatorias, continuando con examinar cada una de las distribuciones continuas y discretas; y, por último, señalando qué es un estimador, las propiedades que debe cumplir y las pruebas de bondad a las que se deben ajustar bajo las reglas de la prueba de hipótesis .

Capitulo III. Modelos y simulación

3.1 ¿QUÉ ES UN MODELO?

Un modelo es una representación simplificada de un sistema elaborado para comprender, predecir y controlar el comportamiento de dicho sistema.

Un sistema se define como una sección de la realidad que es foco primario de un estudio y está compuesto por variables que interactúan entre sí para alcanzar el éxito de un fin lógico. Un sistema puede realizar una función que no es realizable por sus componentes individuales. También se puede definir como la porción del Universo que será objeto de la simulación (Banks, Carson, & Nelso, 1996).

La representación de modelos puede adoptar distintas formas:

- Mentales: visión personal de un país o ideología
- Físicas: de una casa, un puente, un ordenador, etc.
- Simbólicas: grafico o esquemas.

Aplicaciones de la simulación

La simulación es conveniente cuando no existe una formulación matemática analíticamente resoluble. Muchos sistemas reales no pueden ser modelados matemáticamente con las herramientas actualmente disponibles, por ejemplo, la conducta de un cliente de un banco (Fishman, 1978).

Tenemos el ejemplo cuando existe una formulación matemática, pero es difícil obtener una solución analítica, los modelos matemáticos utilizados para modelar un reactor nuclear o una planta química son imposibles de resolver en forma analítica sin realizar serias simplificaciones.

De igual manera si no existe el sistema real. Es problema del ingeniero que tiene que diseñar un sistema nuevo. El diseño del sistema mejorará notablemente si se cuenta con un modelo adecuado para realizar experimentos.

Si los experimentos son imposibles debido a impedimentos económicos, de seguridad, de calidad o éticos. En este caso el sistema real está disponible para realizar experimentos, pero la dificultad de los mismos hace que se descarte esta opción. Un ejemplo de esto es la imposibilidad de provocar fallas en un avión real para evaluar la conducta del piloto, tampoco se puede variar el valor de un impuesto para evaluar la reacción del mercado.

El sistema evoluciona muy lentamente o muy rápidamente. Un ejemplo de dinámica lenta es el problema de los científicos que estudian la evolución del clima. Ellos deben predecir la conducta futura del clima dado las condiciones actuales, no pueden esperar a que un tornado arrase una ciudad para luego dar el mensaje de alerta. Por el contrario, existen fenómenos muy rápidos que deben ser simulados para poder observarlos en detalles, por ejemplo, una explosión.

Entre las posibles desventajas de la simulación se pueden citar (Fishman, 1978):

- El desarrollo de un modelo puede ser costoso, laborioso y lento.
- Existe la posibilidad de cometer errores. No se debe olvidar que la experimentación se lleva a cabo con un modelo y no con el sistema real; entonces, si el modelo está mal o se cometen errores en su manejo, los resultados también serán incorrectos.
- No se puede conocer el grado de imprecisión de los resultados. Por lo general el modelo se utiliza para experimentar situaciones nunca planteadas en el sistema real, por lo tanto, no existe información previa para estimar el grado de correspondencia entre la respuesta del modelo y la del sistema real.

Actualmente la simulación presta un invalorable servicio en casi todas las áreas posibles, algunas de ellas son (Shannon, 1988):

a) Procesos de manufacturas: Ayuda a detectar cuellos de botella, a distribuir personal, determinar la política de producción.

b) Plantas industriales: Brinda información para establecer las condiciones óptimas de operación, y para la elaboración de procedimientos de operación y de emergencias.

c) Sistemas públicos: Predice la demanda de energía durante las diferentes épocas del año, anticipa el comportamiento del clima, predice la forma de propagación de enfermedades.

d) Sistemas de transportes: Detecta zonas de posible congestionamiento, zonas con mayor riesgo de accidentes, predice la demanda para cada hora del día.

e) Construcción: Predice el efecto de los vientos y temblores sobre la estabilidad de los edificios, provee información sobre las condiciones de iluminación y condiciones ambientales en el interior de los mismos, detecta las partes de las estructuras que deben ser reforzadas.

e) Diseño: Permite la selección adecuada de materiales y formas. Posibilita estudiar la sensibilidad del diseño con respecto a parámetros no controlables.

f) Educación: Es una excelente herramienta para ayudar a comprender un sistema real debido a que puede expandir, comprimir o detener el tiempo, y además es capaz de brindar información sobre variables que no pueden ser medidas en el sistema real. Capacitación: Dado que el riesgo y los costos son casi nulos, una persona puede utilizar el simulador para aprender por sí misma utilizando el método más natural para aprender: el de prueba y error.

La importancia de la Simulación es evidente al considerar el impacto que tuvieron algunos trabajos, como son (Law & Kelton, 1991):

a) La Perestroyka: Estudios de simulación efectuados en Rusia en las décadas del 70 y 80 convencieron a los dirigentes de la necesidad de plantear un fuerte cambio en la economía de ese país.

b) La caída de la bolsa de New York en 1988: La utilización de programa s de simulación por parte de los corredores de la bolsa causó una falsa inestabilidad que provocó la caída.

c) El regreso del Apolo 13: La simulación jugó un rol fundamental en la determinación del plan de emergencia. La nave retornó con éxito a pesar de las graves averías.

d) Los Voyagers: Gracias a la simulación se pudieron establecer los itinerarios óptimos para estas naves con un mínimo consumo de energía aprovechando la atracción gravitacional de los planetas.

e) Proyecto Monte Carlo: Von Newman y Ulam (1945) emplearon simulación para estudiar reacciones nucleares.

f) Los modelos del planeta: Algunos plantean la posibilidad de un calentamiento global debido al efecto invernadero. Otros plantean la posibilidad de un enfriamiento y predicen una nueva era glaciar.

g) Capacitación de tropas: En el operativo “Tormenta del desierto” llevado a cabo en la guerra contra Irak, las tropas de todas las fuerzas estadounidenses que participaron (fuerza aérea, marina y ejército) fueron entrenadas con simuladores.

h) Capacitación de policías: Se utiliza entornos virtuales para que el policía aprenda a conducirse en situaciones de riesgo.

i) Simuladores de vuelos: Fue una de las primeras aplicaciones de los simuladores. Actualmente se utilizan para entrenar pilotos de aviones comerciales y de combate.

Las variables o parámetros (P) son atributos que se fijaron durante el diseño del sistema ya sea por el diseñador o por la naturaleza, por ejemplo: la cilindrada del motor, la aceleración de la gravedad. Las variables se clasifican a su vez en (Kelton, Sadowski, & Sadowski, 1998):

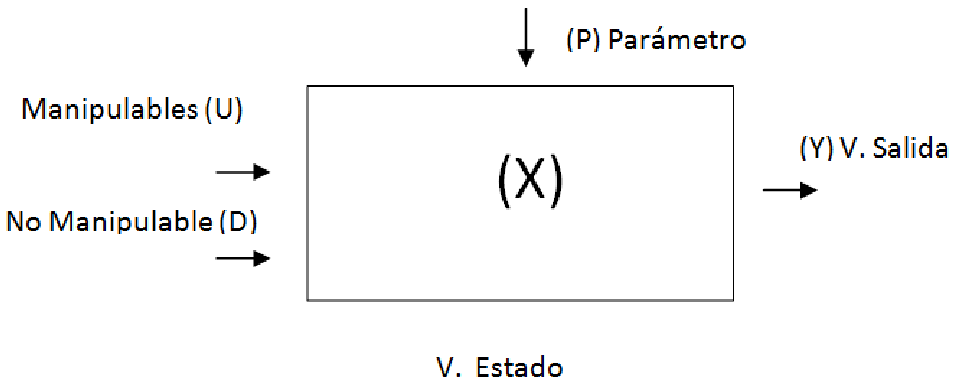
Variables de entrada o exógenas: Son fijadas por el medioambiente del sistema. Pueden ser manipulables (U), se fijan a voluntad, o no (D). Un ejemplo del primer caso es la posición del pedal del acelerador, y del segundo caso es la velocidad del viento. Una variable de entrada no manipulable se denomina perturbación.

Variables de salida (Y): Son las variables de estado, o combinación de ellas, que son medidas o traspasan la frontera del sistema.

Variables internas: Son las variables del sistema que no son ni de entrada, ni de salida, ni parámetros.

Variables de estado (X): Conforman el conjunto mínimo de variables internas del sistema necesarias para describir completamente su estado interno.

Representación gráfica de las variables de un modelo



Fuente: Shannon, R. (1988). Simulación de Sistemas. Diseño, desarrollo e implementación.

3.1.1 Clasificación y Tipos de los modelos

De acuerdo a su naturaleza, un sistema puede ser (Law & Kelton, 1991):

Determinístico: Si el sistema no contiene ningún elemento aleatorio es un sistema determinístico. En este tipo de sistema, las variables de salidas e internas quedan perfectamente determinadas al especificar las variables de entrada, los parámetros y las variables de estado. Es decir, las relaciones funcionales entre las variables del sistema están perfectamente definidas. El calentador eléctrico estudiado es un sistema determinístico.

Estocástico: En este caso algún elemento del sistema tiene una conducta aleatoria. Entonces, para entradas conocidas no es posible asegurar los valores de salida. Un ejemplo de sistema estocástico es una máquina tragamonedas en la cual una misma acción (tirar la palanca) genera un resultado incierto (ganar o perder). Cuando un sistema determinístico es alimentado con entradas estocásticas, la respuesta del sistema es también estocástica. Por ejemplo, la temperatura ambiente es una variable estocástica que afecta la respuesta del calentador eléctrico. En el mundo real, los sistemas siempre tienen elementos estocásticos ya sea por su propia naturaleza o porque son fenómenos no comprendidos actualmente; por ejemplo, a un cavernícola le podía parecer que los eclipses eran fenómenos aleatorios, hoy son predichos. Sin embargo, se puede considerar a un sistema real con un sistema determinístico si su incertidumbre es menor que un valor aceptado.

Continuo: Se tiene un sistema continuo cuando las relaciones funcionales entre las variables del sistema sólo permiten que el estado evolucione en el tiempo en forma continua (basta que una variable evolucione continuamente). Matemáticamente, el estado cambia en infinitos puntos de tiempo. El recipiente del calentador es un subsistema continuo porque tanto M como T evolucionan en forma continua durante la operación del sistema.

Discreto: Se tiene un sistema discreto cuando las relaciones funcionales del sistema sólo permiten que el estado varíe en un conjunto finito (contable) de puntos temporales. Las causas instantáneas de los cambios de estados se denominan eventos.

El interruptor del calentador es un subsistema discreto porque la intensidad I sólo puede variar en los instantes que se abre o se cierra el interruptor. La apertura y el cierre del interruptor son eventos. Un sistema continuo puede comportarse en forma discreta si las entradas son discretas. Los sistemas reales son combinaciones de continuos y discretos. La forma de tratarlos se adopta de acuerdo a la característica dominante.

De acuerdo a la naturaleza del modelo empleado, la simulación puede ser por (Fishman, 1978):

Identidad: Es cuando el modelo es una réplica exacta del sistema en estudio. Es la que utilizan las empresas automotrices cuando realizan ensayos de choques de automóviles utilizando unidades reales.

Cuasi-identidad: Se utiliza una versión ligeramente simplificada del sistema real. Por ejemplo, los entrenamientos militares que incluyen movilización de equipos y tropas, pero no se lleva a cabo una batalla real.

Laboratorio: Se utilizan modelos bajo las condiciones controladas de un laboratorio. Se pueden distinguir dos tipos de simulaciones:

- Juego operacional: Personas compiten entre ellas, ellas forman parte del modelo, la otra parte consiste en computadoras, maquinaria, etc. Es el caso de una simulación de negocios donde las computadoras se limitan a recolectar la información generada por cada participante y a presentarla en forma ordenada a cada uno de ellos.
- Hombre-Máquina: Se estudia la relación entre las personas y la máquina. Las personas también forman parte del modelo. La computadora no se limita a recolectar información, sino que también la genera. Un ejemplo de este tipo de simulación es el simulador de vuelo.

Simulación por computadora: El modelo es completamente simbólico y está implementado en un lenguaje computacional. Las personas quedan excluidas del modelo. Un ejemplo es el simulador de un sistema de redes de comunicación donde la conducta de los usuarios está modelada en forma estadística.

3.1.2 Estructura de un modelo de simulación

En el desarrollo de una simulación se pueden distinguir las siguientes etapas (Banks, Carson, & Nelso, 1996):

1. Formulación del problema: En este paso debe quedar perfectamente establecido el objeto de la simulación. El cliente y el desarrollador deben acordar lo más detalladamente posible los siguientes factores: los resultados que se esperan del simulador, el plan de experimentación, el tiempo disponible, las variables de interés, el tipo de perturbaciones a estudiar, el tratamiento estadístico de los resultados, la complejidad de la interfaz del simulador, etc. Se debe establecer si el simulador será operado por el usuario o si el usuario sólo recibirá los resultados. Finalmente, se debe establecer si el usuario solicita un trabajo de simulación o un trabajo de optimización. Definición del sistema: El sistema a simular debe estar perfectamente definido. El cliente y el desarrollador deben

acordar dónde estará la frontera del sistema a estudiar y las interacciones con el medioambiente que serán consideradas.

2. Formulación del modelo: Esta etapa es un arte y será discutida más adelante. La misma comienza con el desarrollo de un modelo simple que captura los aspectos relevantes del sistema real. Los aspectos relevantes del sistema real dependen de la formulación del problema; para un ingeniero de seguridad los aspectos relevantes de un automóvil son diferentes de los aspectos considerados por un ingeniero mecánico para el mismo sistema. Este modelo simple se irá enriqueciendo como resultado de varias iteraciones.

3. Colección de datos: La naturaleza y cantidad de datos necesarios están determinadas por la formulación del problema y del modelo. Los datos pueden ser provistos por registros históricos, experimentos de laboratorios o mediciones realizadas en el sistema real. Los mismos deberán ser procesados adecuadamente para darles el formato exigido por el modelo.

4. Implementación del modelo en la computadora: El modelo es implementado utilizando algún lenguaje de computación. Existen lenguajes específicos de simulación que facilitan esta tarea; también, existen programas que ya cuentan con modelos implementados para casos especiales.

5. Verificación: En esta etapa se comprueba que no se hayan cometidos errores durante la implementación del modelo. Para ello, se utilizan las herramientas de debugging provistas por el entorno de programación.

6. Validación: En esta etapa se comprueba la exactitud del modelo desarrollado. Esto se lleva a cabo comparando las predicciones del modelo con: mediciones realizadas en el sistema real, datos históricos o datos de sistemas similares. Como resultado de esta etapa puede surgir la necesidad de modificar el modelo o recolectar datos adicionales.

7. Diseño de experimentos: En esta etapa se decide las características de los experimentos a realizar: el tiempo de arranque, el tiempo de simulación y el número de simulaciones. No se debe incluir aquí la elaboración del conjunto de alternativas a probar para seleccionar la mejor, la elaboración de esta lista y su manejo es tarea de la optimización y no de la simulación. Debe quedar claro

cuando se formula el problema si lo que el cliente desea es un estudio de simulación o de optimización.

8. Experimentación: En esta etapa se realizan las simulaciones de acuerdo el diseño previo. Los resultados obtenidos son debidamente recolectados y procesados.

9. Interpretación: Se analiza la sensibilidad del modelo con respecto a los parámetros que tienen asociados la mayor incertidumbre. Si es necesario, se deberán recolectar datos adicionales para refinar la estimación de los parámetros críticos.

Implementación: conviene acompañar al cliente en la etapa de implementación para evitar el mal manejo del simulador o el mal empleo de los resultados del mismo

10. Documentación: incluye la elaboración de la documentación técnica y manuales de uso. La documentación técnica debe contar con una descripción detallada del modelo y de los datos; también, se debe incluir la evolución histórica de las distintas etapas del desarrollo. Esta documentación será de utilidad para el posterior perfeccionamiento del simulador.

De igual manera que en las simulaciones, la construcción de un modelo debe seguir ciertas etapas

1. Análisis de un sistema a través de un modelo, esto implica que la estructura del modelo que representa un sistema sea manipulable.

2. Se determina al modelo conceptual como una representación lógica aproximada del sistema real que como tal constituye una abstracción simplificada del mismo

3. Identificación de las entidades principales del sistema y de sus atributos característicos

4. Identificación en representación de las reglas que gobiernan el sistema que se pretende simular

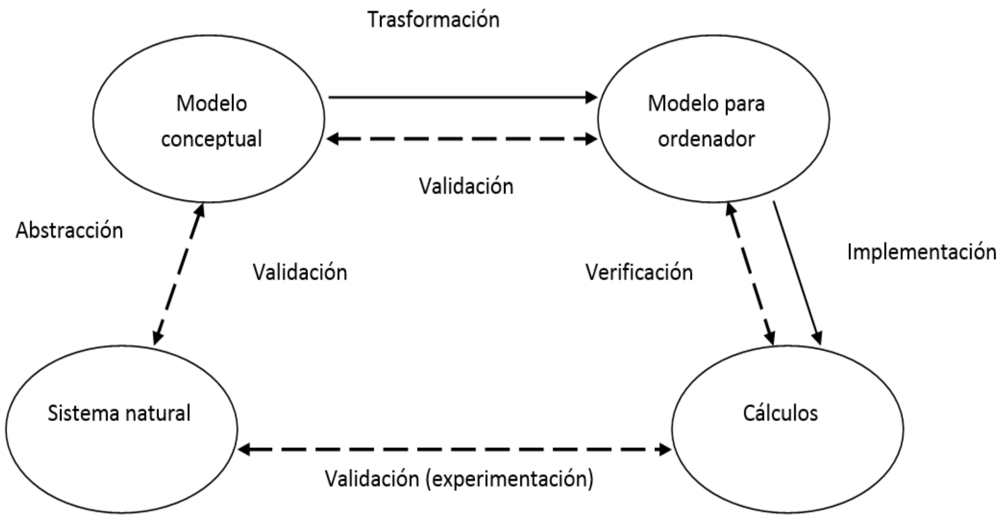
5. Captación de la naturaleza de las interacciones lógicas del sistema que se modeliza.

6. Verificación de las reglas incorporadas al modelo, son una representación valida de las del sistema que se modeliza.

7. Formulación de hipótesis de modelización que traducen adecuadamente la información sobre el sistema y las relaciones entre sus variables y su papel en los cambios de estado del sistema

8. Representación del comportamiento aleatorio, inerte a los comportamientos del sistema que los exhiban.

Representación gráfica de la construcción de un modelo



Fuente: Shannon, R. (1988). Simulación de Sistemas. Diseño, desarrollo e implementación.

La representación matemática de la estructura de un modelo de simulación es (Law & Kelton, 1991):

$$E = f(x_i : y_i)$$

Donde:

E: es el efecto del componente del sistema.

x_i : son las variables y los parámetros que se pueden controlar

y_i : son las variables y los parámetros que no se pueden controlar

F : es la relación entre x_i & y_i , que da origen a E

Cualquier modelo se puede combinar de los siguientes elementos

- Componentes: son partes constituyentes que en conjunto forman un sistema. También llamadas elementos o subsistemas
- Variables: toman valores que la forma de la función permite. Existen de tipos exógenas y endógenas

Exógenas o independientes: surgen por causas externas

Endógenas o dependientes: surgen dentro del problema.

- Parámetros: pueden tomar valores arbitrarios.
- Relaciones o funciones: describen el comportamiento de las variables y los parámetros dentro de un componente o dentro de los componentes de un sistema.

Pueden ser determinísticas (una salida es determinada por una entrada dada) o estocásticas (una es indefinida por una entrada)

- Restricciones: limitaciones impuestas a los valores de las variables o la manera en la cual los recursos pueden asignarse o consumirse.

Pueden ser autoimpuestas por el diseñador o impuestas por el sistema mediante la naturaleza del mismo.

- Funciones de objetivo: Es una definición explicita delos objetivos o metas del sistema y de cómo se evaluarán.

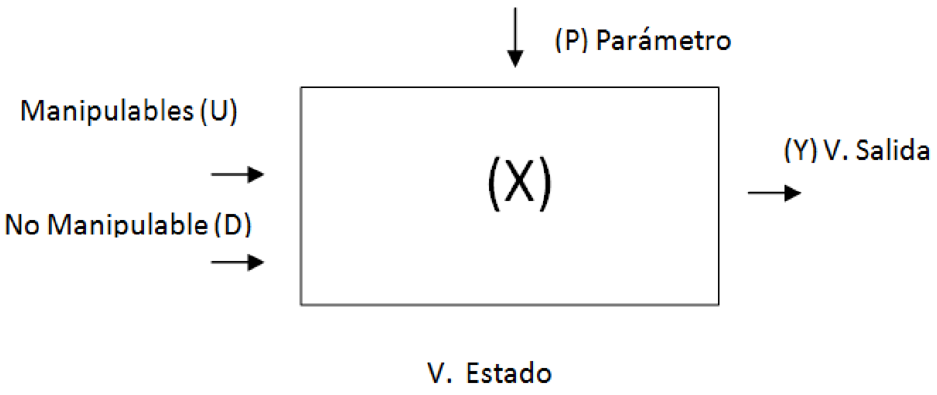
3.1.3 Criterios para realizar un modelo óptimo

Como ya se planteó anteriormente, cuando se simula se experimenta con un modelo para obtener ciertos resultados. Un modelo es también un sistema, y de acuerdo al tipo de variables de salida del modelo, el modo de simulación será (Shannon, 1988):

Análisis: Es el modo más empleado, en él las variables de salida del modelo representan a las variables de salida del sistema real. Este modo se utiliza para estimar la respuesta del sistema real ante entradas especificadas.

Debido a que imita un sistema que realmente funciona, el modelo es matemáticamente más estable y se asegura la existencia de una solución.

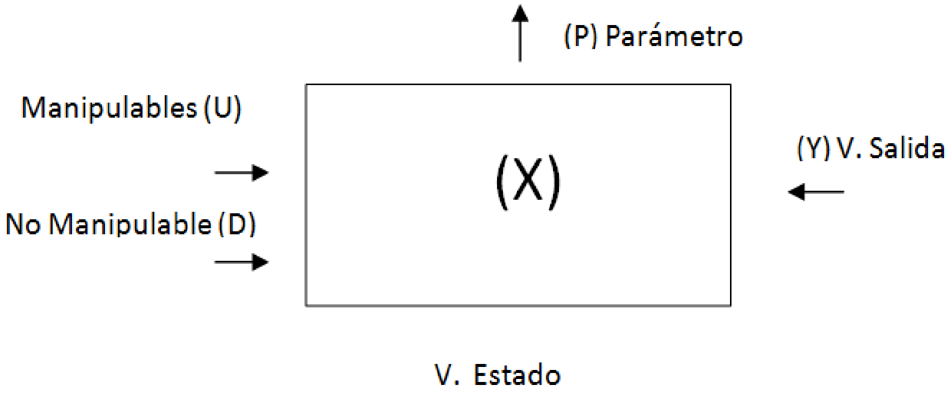
Representación del modelo de análisis



Fuente: Shannon, R. (1988). Simulación de Sistemas. Diseño, desarrollo e implementación.

Diseño: En este modo las salidas del modelo representan a los parámetros del sistema real. Se utiliza en la etapa de diseño de un equipo donde el problema es determinar los parámetros para los cuales el sistema producirá las salidas deseadas para las entradas especificadas.

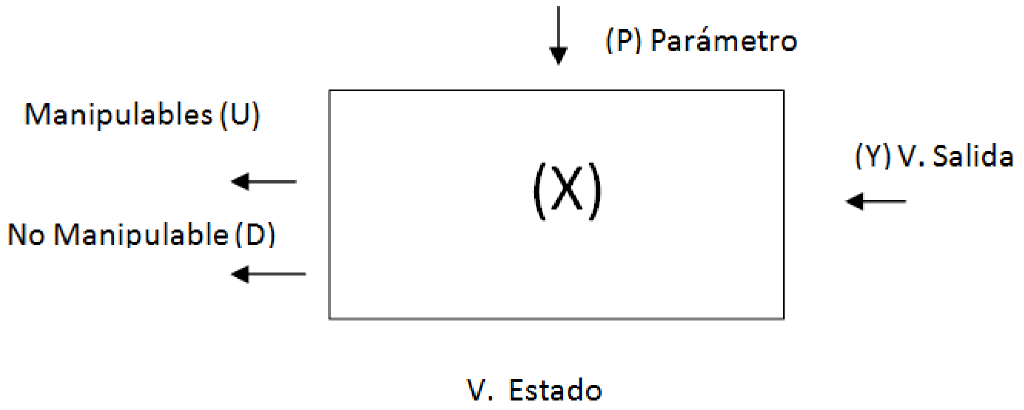
Representación del modelo de diseño



Fuente: Shannon, R. (1988). Simulación de Sistemas. Diseño, desarrollo e implementación.

Control: Las variables de salida del modelo representan a las variables de entrada del sistema real. Este modo sirve para determinar los valores que deberán adoptar las entradas del sistema para producir los resultados deseados. Se utiliza cuando se desea determinar las condiciones de operación de un sistema.

Representación del modelo de control



Fuente: Shannon, R. (1988). Simulación de Sistemas. Diseño, desarrollo e implementación.

Por lo general, los simuladores se operan en modo diseño dejando al usuario la tarea de iterar para obtener los resultados provistos por los otros modos. Por ejemplo, para estimar el voltaje requerido para el calentador (modo diseño), se pueden realizar varias simulaciones en modo análisis para un conjunto de valores de voltaje, y se selecciona el que produce la salida deseada

Modelado es el proceso de construcción de un modelo. Un modelo es una representación de un objeto, sistema, o idea. Usualmente, su propósito es ayudar explicar, entender o mejorar un sistema (Shannon, 1988). Los modelos son útiles para:

El pensamiento: Al construir un modelo necesariamente se debe ordenar y completar el conocimiento que del sistema real se posee.

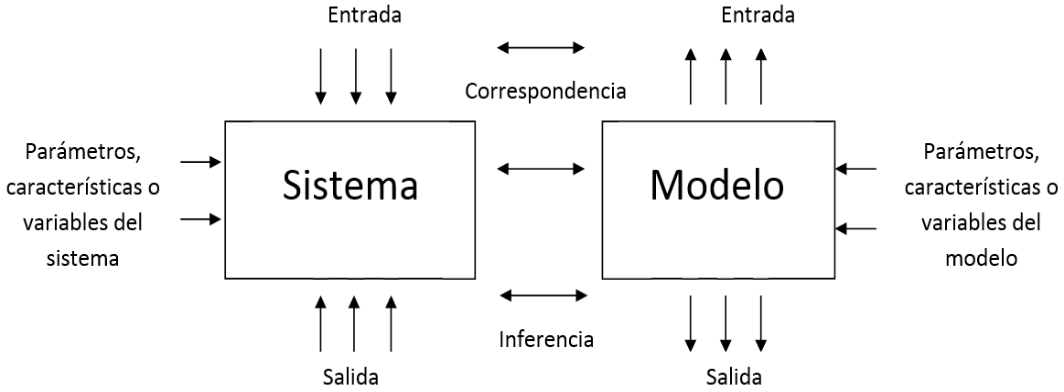
La comunicación: Un modelo elimina la ambigüedad del lenguaje para comunicarse con expertos.

El entrenamiento y la instrucción: Un modelo puede ser utilizado para entrenar con costo y riesgo casi nulos. Por ejemplo, los submarinos a escala utilizados por la marina alemana para entrenar en secreto antes de la segunda guerra mundial; o también, el sistema de barcos a escalas utilizados actualmente en Francia para entrenar a los capitanes de barcos petroleros.

La predicción: Un modelo sirve para predecir la conducta del sistema real. Es el caso de los modelos utilizados para predecir, mediante simulación, la evolución del clima mundial. El modelo de la teoría de la relatividad predice, sin hacer una simulación, que no es posible superar la velocidad de la luz.

La experimentación: La experimentación con un modelo es barata y segura. Se emplea frecuentemente en el diseño de un sistema; por ejemplo, las pruebas que se realizan en un túnel de viento con un modelo a escala de un avión o de un automóvil.

Representación gráfica de la interacción de los componentes de un modelo



Fuentes: Ross, S. M. (2006). Simulation.

El modelado es un arte. Cualquier conjunto de reglas para desarrollar modelos tiene una utilidad limitada y sólo puede servir como una guía sugerida. El arte de modelar consiste en la habilidad para analizar un problema, resumir sus características esenciales, seleccionar y modificar las suposiciones básicas que caracterizan al sistema, y luego enriquecer y elaborar el modelo hasta obtener una aproximación útil. Los pasos sugeridos para este proceso son (Ogunnaike B.A., 1994):

1. Establecer una definición clara de los objetivos.

2. Analizar el sistema real.

3. Dividir el problema del sistema en problemas simples.

4. Buscar analogías.

5. Considerar un ejemplo numérico específico del problema.

6. Determinar las variables de interés.

7. Escribir los datos obvios.

8. Escribir las ecuaciones teóricas o empíricas que describen los fenómenos presentes y relacionan las variables de interés.

9. Si se tiene un modelo manejable, enriquecerlo. De otra manera, simplificarlo.

Generalmente, simplificar un modelo implica: Convertir variables en constantes. Eliminar o combinar variables. Suponer linealidad.

Agregar suposiciones más potentes y restricciones. Restringir los límites del sistema.

Para enriquecerlo se procede de la forma contraria. Durante el proceso de modelado se debe alcanzar un equilibrio entre el grado de detalle y el riesgo de falta de exactitud. El mejor modelo, es el modelo más simple que puede resolver el problema con el grado de exactitud requerido.

Un modelo debe ser:

- Fácil de entender por parte del usuario. Dirigido a metas u objetivos.
- Sensato, no de respuestas absurdas.
- Fácil de manipular y controlar por parte del usuario. Es decir, debe ser sencillo comunicarse con el modelo.
- Completo, en lo referente a asuntos importantes.
- Adaptable, con un sencillo procedimiento para modificar o actualizar el modelo.

- Evolutivo, debe ser sencillo al principio y volverse más complejo en el tiempo.

Se consideran 4 pasos para convertir un sistema en un modelo matemático.

1. especificación del propósito del modelo

2. especificación de los componentes que se incluirán en el modelo.

3. especificación de los parámetros y variables asociado con los componentes

4. especificación de las relaciones funcionales entre los componentes, parámetros y variables.

Alguno de los propósitos para los que se usan los modelos de simulación son:

1. evaluación: determinar la realización del proyecto del sistema, el más correcto y apropiado.

2. comparación: para tomar decisiones entre políticas y procedimientos operativos propuestos.

3. predicción: estimación del rendimiento del sistema

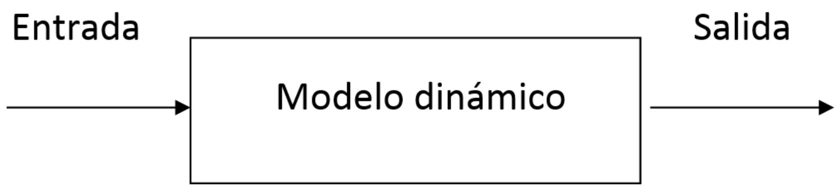
4. Análisis de sensibilidad: determinar los factores que afectan el rendimiento del sistema.

5. optimización: determinar qué factores producirán una mejor respuesta

6. relaciones funcionales: establecer la naturaleza de la relación uno o más factores significativos y la respuesta del sistema.

Los componentes de un modelo estarán determinados por 3 entidades distintas, la entrada, el sistema y la respuesta (Ross, 2006).

Representación gráfica de los componentes de un modelo



Fuentes: Ross, S. M. (2006). Simulation.

Se deben conocer al menos dos de los 3 componentes involucrados y se pueden dar algunos de estos casos.

1. cuando se conocen las ecuaciones que describen el sistema y se necesita conseguir las salidas a partir de entradas dadas (son las más sencillas para modelar).

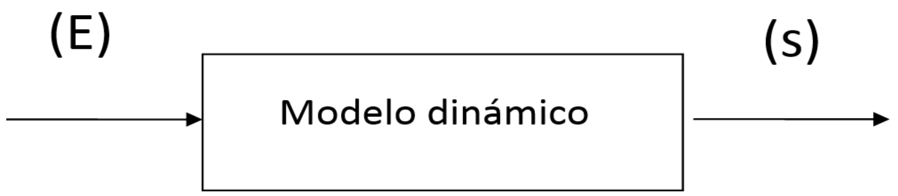
2. cuando se conocen las ecuaciones que describen el sistema y se necesita una salida deseada, el problema es conseguir las entradas necesarias para producir dicha respuesta. Estos casos se clasifican en problemas de control.

3. cuando se conocen las entradas y salidas del sistema y se debe conseguir la descripción matemática del sistema. A este caso se le conoce como problema de identificación de la estructura.

En general los componentes convierten entradas en salidas clasificándose de la siguiente manera (Ross, 2006):

Elementos de transformación: se opera sobre una o más entradas de alguna manera determinada y se transforma en una o más salidas.

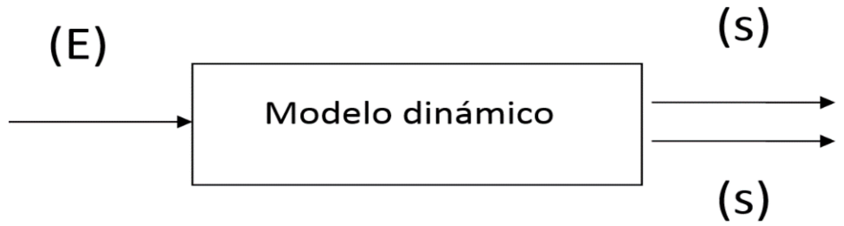
Representación de los elementos de transformación



Fuentes: Ross, S. M. (2006). Simulation.

Elementos de clasificación: se separan o clasifican una o más entradas en dos o más salidas diferentes.

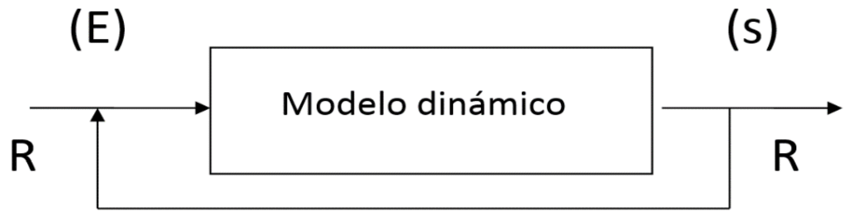
Representación de los elementos de clasificación



Fuentes: Ross, S. M. (2006). Simulation.

Elementos de retroalimentación: se modifica la entrada de alguna manera como una función de salida.

Representación de los elementos de retroalimentación



Fuentes: Ross, S. M. (2006). Simulation.

3.1.4 Riesgo en la elaboración de un modelo

Si el modelo que se construye sólo se orienta a reproducir las salidas del sistema real sin intentar modelar su comportamiento interno; entonces, será un modelo experimental o de caja negra (Banks, Carson, & Nelso, 1996). En cambio, si el modelo también intenta reproducir las relaciones funcionales del sistema será un modelo con base teórica. Un modelo experimental requiere una gran cantidad de datos para poder calibrarlo o ajustarlo correctamente, y su rango de validez está limitado a este conjunto de datos. En contraposición, un modelo teórico requiere una cantidad menor de datos y puede ser utilizado fuera del rango de los mismos ya que el rango de validez del modelo está dado por la teoría utilizada y no por los datos.

Entre las desventajas de las simulaciones se encuentran estos casos específicos:

- No se debe utilizar cuando existan técnicas analíticas que permitan plantear, resolver y Optimizar todo el sistema o alguna parte del mismo. Existe un gran deseo de incluir todo en la simulación donde los resultados son visibles y comprendidos por todas las personas (especialistas o no), las cuales gustan hacer ensayos.
- No es posible asegurar que el modelo sea válido: Se corre el riesgo de tomar medidas erróneas basadas en aplicar conclusiones falsas obtenidas mediante un modelo que no representa la realidad.
- No existe criterio científico de selección de alternativas a simular (Estrategia). Es posible omitir una buena sugerencia de innovación simplemente porque a nadie se le ocurrió ensayarla.
- Existe el riesgo de utilizar un modelo fuera de los límites para el cual fue construido, queriendo realizar ensayos para el cual el modelo no es válido. Es posible elaborar todo un gran andamiaje de pruebas y resultados falsos, basados en un modelo confiable y válido bajo otras condiciones.

3.2 SIMULACIONES MONTE CARLO

El método fue llamado así por el principado de Mónaco por ser “la capital del juego de azar”, al tomar una ruleta como un generador simple de números aleatorios. El nombre y el desarrollo sistemático de los métodos de Monte Carlo datan aproximadamente de 1944 con el desarrollo de la computadora. Sin embargo, hay varias instancias (aisladas y no desarrolladas) en muchas ocasiones anteriores a 1944 (Ybnias, 2009).

El uso real de los métodos de Monte Carlo como una herramienta de investigación, proviene del trabajo de la bomba atómica durante la Segunda Guerra Mundial. Este trabajo involucraba la simulación directa de problemas probabilísticos de hidrodinámica concernientes a la difusión de neutrones aleatorios en material de fusión.

Aún en la primera etapa de estas investigaciones, John von Neumann y Stanislao Ulam refinaron esta curiosa “Ruleta rusa'' y los métodos de división. Sin embargo, el desarrollo sistemático de estas ideas tuvo que esperar el trabajo de Harris y Herman Kahn en 1948. Aproximadamente en el mismo año, Fermi, Metropolos y Ulam obtuvieron estimadores para los valores característicos de la ecuación de Schrödinger para la captura de neutrones a nivel nuclear (Ybnias, 2009).

Alrededor de 1970, los desarrollos teóricos en complejidad computacional comienzan a proveer mayor precisión y relación para el empleo del método Monte Carlo. La teoría identifica una clase de problemas para los cuales el tiempo necesario para evaluar la solución exacta al problema crece con la clase, al menos exponencialmente con M. La cuestión a ser resuelta era si MC pudiese o no estimar la solución al problema de tipo intratable con una adecuación estadística acotada a una complejidad temporal polinomial en M. Karp(1985) muestra esta propiedad para estimar en una red plana multiterminal con arcos fallidos aleatorios. Dyer(1989) utiliza MC para estimar el volumen de un convexbody en el espacio Euclidiano M-dimensional. Broder(1986), Jerrum y Sinclair (1988) establecen la propiedad para estimar la persistencia de una matriz o en forma equivalente, el número de matching perfectos en un grafo bipartito.

Los orígenes de esta técnica están ligados al trabajo desarrollado por Stan Ulam y John Von Neumann a finales de los 40 en el laboratorio de Los Alamos, cuando investigaban el movimiento aleatorio de los neutrones. En años posteriores, la simulación de Monte Carlo se ha venido aplicando a una infinidad de ámbitos como alternativa a los modelos matemáticos exactos o incluso como único medio de estimar soluciones para problemas complejos. Así, en la actualidad es posible encontrar modelos que hacen uso de simulación MC en las áreas informática, empresarial, económica, industrial e incluso social. En otras palabras, la simulación de Monte Carlo está presente en todos aquellos ámbitos en los que el comportamiento aleatorio o probabilístico desempeña un papel fundamental - precisamente, el nombre de Monte Carlo proviene de la famosa ciudad de Mónaco, donde abundan los casinos de juego y donde el azar, la probabilidad y el comportamiento aleatorio conforman todo un estilo de vida.

Son muchos los autores que han apostado por utilizar hojas de cálculo para realizar simulación MC. La potencia de las hojas de cálculo reside en su universalidad, en su facilidad de uso, en su capacidad para recalcular valores y, sobre todo, en las posibilidades que ofrece con respecto al análisis de escenarios (“what-ifanalysis”). Las últimas versiones de Excel incorporan, además, un lenguaje de programación propio, el Visual Basic for Applications, con el cual es posible crear auténticas aplicaciones de simulación destinadas al usuario final. En el mercado existen de hecho varios complementos de Excel (Add-Ins) específicamente diseñados para realizar simulación MC, siendo los más conocidos: @Risk,CrystallBall, Insight.xla, SimTools.xla, etc..

La simulación de Monte Carlo es una técnica cuantitativa que hace uso de la estadística y los ordenadores para imitar, mediante modelos matemáticos, el comportamiento aleatorio de sistemas reales no dinámicos (por lo general, cuando se trata de sistemas cuyo estado va cambiando con el paso del tiempo, se recurre bien a la simulación de eventos discretos o bien a la simulación de sistemas continuos).

La clave de la simulación MC consiste en crear un modelo matemático del sistema, proceso o actividad que se quiere analizar, identificando aquellas variables (inputs del modelo) cuyo comportamiento aleatorio determina el comportamiento global del sistema. Una vez identificados dichos inputs o variables aleatorias, se lleva a cabo un experimento que consistente en: generar, con ayuda del ordenador, muestras aleatorias (valores concretos) para dichos inputs, y analizar el comportamiento del sistema ante los valores generados. Tras repetir n veces este experimento, se dispondrán de n observaciones sobre el comportamiento del sistema, lo cual nos será de utilidad para entender el funcionamiento del mismo, nuestro análisis será tanto más preciso cuanto mayor sea el número n de experimentos que se lleven a cabo.

3.2.1 Método de Monte Carlo

La simulación es el proceso de diseñar y desarrollar un modelo computarizado de un sistema o proceso y conducir experimentos con este modelo con el propósito de entender el comportamiento del sistema o evaluar varias estrategias con las cuales se puede operar el sistema (Shannon, 1988).

- Modelo de simulación: conjunto de hipótesis acerca del funcionamiento del sistema expresado como relaciones matemáticas y/o lógicas entre los elementos del sistema.
- Proceso de simulación: ejecución del modelo a través del tiempo en un ordenador para generar muestras representativas del comportamiento.

Simulación estadística o Monte Carlo está basada en el muestreo sistemático de variables aleatorias. Las aplicaciones de simulaciones se pueden expresar de maneras continuas, discretas y autómatas.

Simulación continua: Los estados del sistema cambian continuamente su valor. Estas simulaciones se modelan generalmente con ecuaciones diferenciales.

Simulación por eventos discretos: Se define el modelo cuyo comportamiento varía en instantes del tiempo dados. Los momentos en los que se producen los cambios son los que se identifican como los eventos del sistema o simulación.

Simulación por autómatas celulares: Se aplica a casos complejos, en los que se divide al comportamiento del sistema en subsistemas más pequeños denominadas células. El resultado de la simulación está dado por la interacción de las diversas células.

3.2.2 Procesos de Monte Carlo

Para el proceso Monte Carlo es necesario definir correctamente la problemática, se debe realizar una descripción cualitativa y cuantitativa del problema; por lo tanto, para comenzar a realizar un modelo Monte Carlo es indispensable hacer un planteamiento general.

Lo primero que se debe realizar es la formulación del modelo, lo cual nos indicara la lógica del plan a emplear, deberá ser detallado y con coherencia.

Teniendo el planteamiento lógico se procederá a la programación. Es necesario declarar cual será el lenguaje necesario de programación a implementar en el modelo, deberá evitarse en un principio tener limitantes con las variables y parámetros.

La Verificación y Validación del modelo es necesaria para tener un nivel de confianza alto.

El Diseño de experimentos y plan de corridas es el lado creativo de la elaboración del modelo, la intención de este paso en el proceso es realizar un modelo de fácil entendimiento, automatizado y óptimo en tiempo de realización y ejecución de las simulaciones. Al aplicar los pasos anteriores los procesos del ordenador no deben interrumpirse y deberán imprimir información veraz.

El Análisis de resultados está relacionado con la toma de decisiones, los modelos computables solo nos muestran información, es necesario hacer un reporte final con los resultados obtenidos para explicar la eficiencia del modelo con información correcta

3.2.3-Software auxiliar

El software que se utilizará para la creación del modelo de simulación es @RISK.

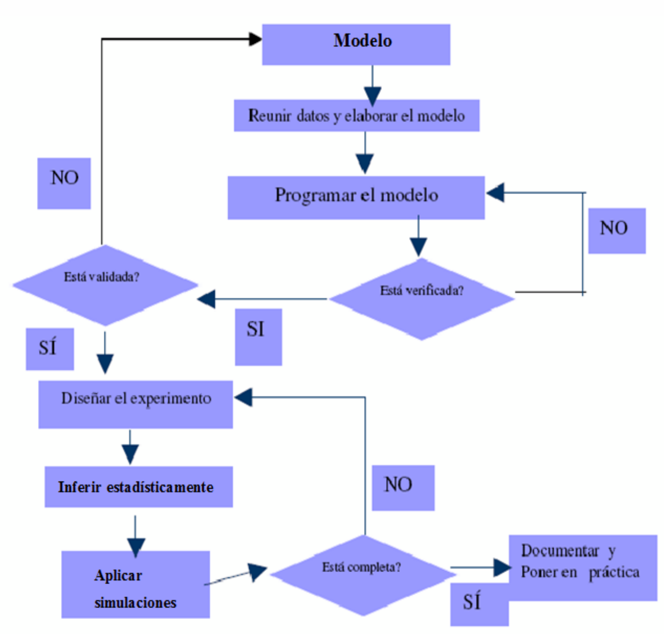
El software @RISK incorpora técnicas avanzadas de modelación y de análisis de riesgo a Microsoft Excel. La fácil interfaz y manejo de los datos para la creación de nuestro modelo a realizar, permitirá que las simulaciones del modelo den un claro panorama del análisis de riesgo.

En un sentido amplio, análisis de riesgo es cualquier método, cualitativo y/o cuantitativo, de estimar el impacto del factor riesgo en situaciones de decisión. Existen miles de métodos que combinan las técnicas cuantitativa y cualitativa en mayor o en menor grado. El objetivo de cualquiera de estos métodos es ayudar a la persona a elegir la acción que se debe tomar, teniendo en cuenta los posibles resultados de cada acción. El análisis de riesgo de @RISK es un método de análisis cuantitativo diseñado para definir los resultados de una decisión en forma de distribución de probabilidad. En general, las técnicas de análisis de riesgo de @RISK comprenden cuatro pasos:

- Desarrollo de un modelo: mediante la definición del problema o situación en el formato de la hoja de cálculo de Excel
- Identificación de la incertidumbre: en las variables de la hoja de cálculo de Excel, especificación de los posibles valores con distribuciones de probabilidad, e identificación de los resultados inciertos que desea analizar
- Análisis del modelo mediante simulación: para determinar el rango y las probabilidades de todas las conclusiones posibles de los resultados de la hoja de trabajo
- Toma de decisión: basada en los resultados obtenidos y en las preferencias personales @RISK ayuda en los tres primeros pasos de este proceso ofreciéndole una eficaz y flexible herramienta que se incorpora a Excel para facilitar la generación de modelos y el análisis de riesgo. Los resultados obtenidos por @RISK se pueden utilizar para orientar la decisión que se va a tomar.

3.2.4 Diagrama

Diagrama del algoritmo de simulación Monte Carlo



Fuente: Ybnias., G. (2009). Métodos cuantitativos para los negocios.

El algoritmo de Simulación Monte Carlo Crudo o Puro está fundamentado en la generación de números aleatorios por el método de Transformación Inversa, el cual se basa en las distribuciones acumuladas de frecuencias:

- Determinar la/s variable aleatoria y sus distribuciones acumuladas (F)
- Iterar tantas veces como muestras se necesiten
  - Generar un número aleatorio
  - Uniforme ℵ (0,1).
  - Determinar el valor de las variables aleatorias para el número aleatorio generado de acuerdo a las clases que se tengan.
- Calcular media, desviación estándar error y realizar el histograma.
- Analizar resultados para distintos tamaños de muestra.

Otra opción para trabajar con Monte Carlo, cuando la variable aleatoria no es directamente el resultado de la simulación o se tiene relación entre variables es la siguiente:

- Diseñar el modelo lógico de decisión
- Especificar distribuciones de probabilidad para las variables aleatorias relevantes.
- Incluir posibles dependencias entre variables.
- Muestrear valores de las variables aleatorias.
- Calcular el resultado del modelo según los valores del muestreo (iteración) y registrar el resultado
- Repetir el proceso hasta tener una muestra estadísticamente representativa
- Obtener la distribución de frecuencias del resultado de las iteraciones
- Calcular media, desvío.
- Analizarlos resultados

Las principales características a tener en cuenta para la implementación o utilización del algoritmo son:

- El sistema debe ser detallado por una o más funciones de distribución de probabilidad (fdp)
- Generador de números aleatorios: como se generan los números aleatorios es importante para evitar que se produzca correlación entre los valores muestrales.
- Establecer límites y reglas de muestreo para las fdp: conocemos que valores pueden adoptar las variables.
- Definir Scoring: Cuando un valor aleatorio tiene o no sentido para el Modelo a simular.
- Estimación Error: Con que error trabajamos, cuanto error podemos aceptar para que una corrida sea válida
- Técnicas de reducción de varianza.
- Paralelización y vectorización: En aplicaciones con muchas variables se estudia trabajar con varios procesadores paralelos para realizar la simulación.

3.2.5 Números aleatorios y simulación

Muchos problemas que surgen en el ámbito del análisis estadístico pueden ser resueltos utilizando números aleatorios. En el proceso de muestreo se utilizan tales números para seleccionar aleatoriamente los elementos de una muestra de entre los que forman parte de una población, a efecto de estimar los parámetros de la misma. Así mismo, los agricultores que experimentan con distintas técnicas, encuentran necesario practicar distintos tratamientos para las distintas parcelas de una manera aleatoria a efectos de evitar comparaciones sesgadas entre los mismos. Tales asignaciones se efectúan a través de números aleatorios.

Los investigadores pueden utilizar los números aleatorios para estudiar la distribución de una variable aleatoria compleja que venga en función de otras variables aleatorias cuyas distribuciones sean conocidas. Las observaciones aleatorias de las variables con distribuciones conocidas son generadas utilizando números aleatorios, y tales observaciones aleatorias son entonces utilizadas para obtener las observaciones aleatorias de la variable aleatoria compleja. A partir de los valores simulados, los analistas pueden efectuar inferencia en relación con las distribuciones conocidas (forma, media, varianza, simetría, etc.).

Para saber que numero de simulaciones deben realizar para la elaboración del modelo se utiliza la siguiente Técnica. Se plantea que el número de simulaciones son lo bastante grandes para generar confianza en los datos que dichas simulaciones proyectan, partiendo de esta primicia la manera de determinar la cantidad de simulaciones se basara en el teorema de límite central.

El teorema del límite central dice que si una muestra es lo bastante grande $(n > 30)$, sea cual sea la distribución de la variable de interés, la distribución de la media muestral será aproximadamente una normal. Además, la media será la misma que la de la variable de interés, y la desviación típica de la media muestral será aproximadamente el error estándar.

Los números aleatorios pueden ser generados ordenadores y calculadoras programables. Estos números son generalmente fracciones distribuidas uniformemente a lo largo del intervalo unitario $(0,1)$. De nuevo, normalmente se considera la distribución como continua, aunque es evidente que solo múltiplos de algunos números pequeños son factibles de ser utilizados en un ordenador digital. Quizás resultara más conveniente referirse a dichos números como seudoaleatorios, toda vez que son generados por una regla completamente especificada. Una vez que ha sido elegido el primero, la totalidad de la secuencia queda determinada. El procedimiento matemático para generar la secuencia es tan adecuado que no existe contraste estadístico razonable que detecte falta de aleatoriedad. La gran ventaja de una regla explicita es que la secuencia puede ser reproducida a efectos de comprobación. El método del centro del cuadrado fue el primer algoritmo ideado para generar números seudoaleatorios: los dígitos centrales del cuadrado de un número aleatorio eran utilizados para definir el siguiente número aleatorio. Este método, de cualquier forma, resulto pronto insatisfactorio. Todos los ordenadores modernos incluyen tablas de números aleatorios, y en la actualidad se utilizan más frecuentemente los generadores lineales congruentes o generadores lineales de congruencias. Estos generadores operan con números enteros y producen observaciones aleatorias en el intervalo unitario mediante división.

Todos los ordenadores modernos y muchas calculadoras programables incorporan tablas de números aleatorios y no es necesario que el usuario conozca cómo han sido realmente generados dichos números (Hossack, I.B.; Pollard, J.H.; Zehnwirth, B, 2001). El método lineal congruente, de cualquier forma, es sencillo y bastante interesante, y se va a describirlo en esta sección. Solo en raras ocasiones necesitara el usuario definir su propio generador.

Cuando se utiliza un generador lineal congruente de primer orden, se procederá a multiplicar el número entero aleatorio W_{n-1} , correspondiente al paso n -1, por una constante entera k_1 , y a continuación el producto se dividirá por otra constante entera p. El resto será un entero W_n , y tal número será tomado como el entero aleatorio correspondiente al paso n. El proceso viene representado por la siguiente ecuación:

$$W_{\boldsymbol{n}} = k_1 W_{n-1} \pmod p$$

El numero entero aleatorio W_n es divido por p a efectos de producir una observación aleatoria dentro del intervalo unitario. El usuario deberá definir una semilla inicial. Las constantes p y k_1 han de ser elegidas cuidadosamente .en notas de J.G.Skellam se recomienda utilizar las constantes p= 999,563 y k= 470,001. Se obtiene un ciclo de longitud total p-1.

Los generadores lineales congruentes de segundo y tercer orden vienen de

$$W_n = K_1 W_{n-1} + K_2 W_{n-2} \ (mod\ p)$$

$$W_n = K_1 W_{n-1} + K_2 W_{n-2} + K_3 W_{n-3} \ (mod\ p)$$

En cada caso, el entero W_n es divido por p a efectos de obtener una observación aleatoria del intervalo unitario. Las constantes p, k_1, k_2 y k_3 de la siguiente tabla son remondadas por Skellam (Hossack, I.B.; Pollard, J.H.; Zehnwirth, B, 2001).

Constantes enteras recomendadas por J.G.Skellam para utilizar en generadores lineales congruentes. En cada caso se realizara el ciclo completo de p^{s-1} (S es el orden del generador)

Tabla de Skellam para generar números aleatorios

Orden del generador *(s)*	**Constantes recomendadas**			
	p	k_1	k_2	k_3
1	999,564	470,001	-	-
2	998,917	366,528	508,531	-
2	999,563	254,754	529,562	-
3	997,783	360,137	519,815	616,087
3	997,783	286,588	434,446	388,251

Fuente: Hossack, I.B.; Pollard, J.H.; Zehnwirth, B. (2001). Introducción a la Estadística con aplicaciones a los seguros generales.

Utilizar el generador lineal congruente de primer orden de Skellam para generar tres números aleatorios del intervalo unitario (es decir, cinco fracciones aleatorias en el intervalo de 0 a 1). Utilizar como semilla W_0 = 671,800.

W_1 es calculado de la siguiente forma:

470,001 x 671,800 = 315, 746, 671,800

315, 746, 671,800 ÷ 999,563 = 315,884 + 713,108/999,563

W_1 =713,108 y la primera fracción aleatoria es

713,108/999,563 = 0.71341976

El proceso se repite:

470,001 x 713,108 =335, 161, 473,108

335, 161, 473,108 ÷ 999,563 = 335,308 + 2,704/999,563

W_2 = 2,704 y la segunda fracción aleatoria es

2,704/999,563 = 0.00270518

470,001 x = 2,704 = 1, 270, 882,704

1, 270, 882,704 ÷ 999,563 = 1,271 + 438,131/999563

W_3 = 438,131 y la tercera fracción aleatoria es

438,131/999563 = 0.043832255

3.2.6 Cuando se debe simular

La técnica de la simulación puede ser una muy útil y poderosa herramienta para el estudio de problemas que surgen en los seguros generales. No debería, sin embargo, ser utilizada de manera indiscriminada. Existen dos principales situaciones en las cuales puede resultar la mejor aproximación al problema.

Cuando el problema no puede ser resuelto matemáticamente de manera exacta.

Cuando es posible obtener una solución matemática exacta pero resulta extremadamente difícil o tediosa, y una adecuada aproximación al resultado puede ser obtenida rápidamente y de manera sencilla mediante simulación.

3.3 RESUMEN

En este capítulo se determinó paso a paso lo que implica la creación de un modelo. Las ventajas y desventajas de los componentes del mismo, que factores van a ser determinantes para establecer que un modelo sea óptimo y los elementos a tomar en cuenta para minimizar el error del modelo.

En la segunda parte del capítulo se explicó ¿qué es un modelo de simulación?; particularmente este trabajo se dedicó a explicar de manera detallada el método Monte Carlo, explicando ¿qué es?, ¿Cómo funciona? y la razón por la que se elige, y por último se expone cual serán las herramientas auxiliares en los que se aplicaron las simulaciones en un ordenador.

Capitulo IV. Metodología para el cálculo del límite máximo de retención

4.1 LÍMITE MÁXIMO DE RETENCIÓN

Cuando en una cartera de pólizas donde se asegura un mismo tipo de riesgo, las sumas aseguradas no son homogéneas, existe la posibilidad de que se produzca una desviación en la siniestralidad esperada, debido ello a la falta de homogeneidad. Imagine el lector, el caso de una cartera de pólizas de seguros de los ramos daños sin autos, donde se tienen 900 pólizas de suma asegurada de 100,000, y se tiene cinco pólizas con suma asegurada de 2 millones. Lo más seguro es que los siniestros ocurran en forma aleatoria sin que la suma asegurada tenga ninguna influencia en ello. Si se siniestra alguna de las pólizas con suma asegurada de 2 millones, se produce un efecto de pérdida para la compañía aseguradora. Ahí la importancia de considerar un límite de retención para evitar problemas de solvencia.

4.1.1 Introducción

La forma de evitar el riesgo de pérdida debido a la desviación que puede producir la falta de homogeneidad en sumas aseguradas, es reteniendo sólo una parte de las sumas aseguradas grandes, a esa cantidad que se retiene se le llama “límite de retención”. Resulta entonces evidente la necesidad de contar con elementos que permitan determinar cuál puede ser el límite de retención en una cartera de pólizas, de manera que evite la pérdida en el supuesto de que los siniestros puedan ocurrir en algún escenario adverso, es decir, que afecten a pólizas con sumas aseguradas grandes.

El tema, sin embargo, es abordado en la literatura actuarial de una manera que incorpora bases teóricas y supuestos de tal naturaleza que son de muy poca utilidad para la práctica. El abismo entre la teoría y la práctica aparece inevitablemente desde el momento en que se presentan metodologías que parten de suponer que se cuenta con la función de probabilidad de esto y de aquello, que si esto se distribuye gama o beta, etc., etc., etc. Siendo consciente de ello, los elementos de análisis que presento, se concentran en atender las limitaciones y supuestos propios de la práctica, buscando las herramientas más apropiadas para tener una solución razonable y coherente.

4.1.2 Antecedentes

El pasado 24 de mayo de 2010 la Secretaria de Hacienda y Crédito Público (SHCP) publicó en el Diario Oficial de la Federación, las nuevas reglas para la determinación del límite máximo de retención, que deberán seguir todas las compañías de seguros. Mismas que implican fundamentalmente los siguientes cambios

La aprobación por parte del Consejo de Administración. De la metodología y los límites máximos de retención propuestos resultado de su aplicación.

La necesidad de contar con un soporte técnico que justifique la determinación de estos límites aun cuando no excedan el 5% de los activos computables para capital.

Anteriormente el Límite Máximo de Retención se calculaba de acuerdo a las reglas establecidas en 1996 y el Acuerdo de 1998 que indicaba: para Accidentes y Enfermedades, y Daños dicho límite se calculaba como el 5% de los Activos Computables al Capital Mínimo de Garantía.

Debido a los cambios que han existido en el sector y aunado a la búsqueda de un valor de retención tal que, en escenarios adversos probables de ocurrencia de siniestro, no se ponga en riesgo la solvencia de las compañías, se ha llegado a la necesidad de buscar mediante métodos más robustos un límite más preciso y confiable.

4.1.3 Normatividad aplicable

En cumplimiento a lo dispuesto en los artículos 258, 259, 260, 261 , 352 Y 487 de la Ley de Instituciones de Seguros y de Fianzas, y la Quinta, Sexta , Séptima, Octava y Novena de las “Reglas para fijar los Límites Máximos de Retención de las Instituciones y Sociedades Mutualistas de Seguros en las Operaciones de Seguro y Reaseguro”, según la reforma publicada en el Diario Oficial de la Federación el 24 de mayo de 2010, que define que los límites máximos serán establecidos como parte de los esquemas de suscripción y administración de riesgos de cada institución o sociedad mutualista de seguros y dichos limites deberán estar sustentados en un método técnico que los calcule.

> *ARTÍCULO 258.- La Comisión, mediante disposiciones de carácter general, con acuerdo de su Junta de Gobierno, establecerá los procedimientos para determinar, en cada operación o ramo, o bien en cada ramo o subramo, según sea el caso, los límites máximos de retención de las Instituciones.*
>
> *ARTÍCULO 259.- La Comisión, en la emisión de las disposiciones de carácter general a que se refiere el artículo 258 de esta Ley, deberá propiciar la consecución de cualquiera de los objetivos siguientes:*
>
> *I. La seguridad de las operaciones de las Instituciones;*
>
> *II. La diversificación técnica de los riesgos y de las responsabilidades que asuman las Instituciones;*
>
> *III. El aprovechamiento de la capacidad de retención de los sistemas asegurador y afianzador;*
>
> *IV. El desarrollo de políticas adecuadas para la cesión y aceptación de reaseguro o reafianzamiento interno y externo, o*
>
> *V. La conveniencia de dispersar los riesgos y las responsabilidades que por su naturaleza puedan provocar una inadecuada acumulación y afectar la estabilidad de los sistemas asegurador y afianzador.*

ARTÍCULO 260.- Las Instituciones fijarán anualmente, con sujeción a las disposiciones de carácter general a que se refiere el artículo 258 de la presente Ley, sus límites máximos de retención, atendiendo a las operaciones, ramos o subramos que tengan autorizados, así como a los riesgos o responsabilidades que asuman. Para ello, tomarán en cuenta, como mínimo, lo siguiente:

I. El volumen de las operaciones de la Institución;

II. El monto de los Fondos Propios Admisibles de la Institución;

III. El monto y características de los riesgos o responsabilidades asumidos por la Institución; IV. La composición de la cartera de riesgos o responsabilidades de la Institución;

V. La experiencia obtenida respecto al comportamiento de la siniestralidad, o bien respecto al incumplimiento de fiados y al pago de reclamaciones;

VI. La suficiencia, calidad y liquidez de las garantías de recuperación recabadas por la Institución;

VII. La capacidad financiera, técnica y operativa de los contratantes de seguros o de los fiados;

VIII. El grado de avance en el cumplimiento de las obligaciones legales o contractuales del contratante del seguro materia del riesgo asegurado, o bien del cumplimiento de las responsabilidades garantizadas;

IX. La acumulación de riesgos por contratante o grupos de contratantes de seguros, o bien de responsabilidades por fiado o grupos de fiados, y

X. Las políticas que aplique la Institución para ceder o aceptar reaseguro o reafianzamiento.

Las Instituciones informarán a la Comisión, en la forma y términos que ésta establezca en las disposiciones de carácter general a que se refiere el artículo 258 de este ordenamiento, los límites máximos de retención que hayan determinado.

ARTÍCULO 261.- Los excedentes que las Instituciones tengan sobre los límites máximos de retención a que se refiere el artículo 260 de esta Ley, deberán distribuirlos, mediante su cesión a través de contratos de reaseguro o reafianzamiento, a otras Instituciones o a Reaseguradoras Extranjeras, o bien mediante contratos de coaseguro o coafianzamiento con otras Instituciones.

Para dar cumplimiento a lo previsto en este artículo, de manera previa a la expedición de una póliza de seguros o de fianzas que exceda los límites máximos de retención de las Instituciones a que se refiere el artículo 260 de este ordenamiento, dichas Instituciones deberán contar con evidencia de la aceptación de las otras Instituciones o Reaseguradoras Extranjeras que participarán, según sea el caso, en el reaseguro, coaseguro, reafianzamiento o coafianzamiento respectivos.

ARTÍCULO 352.- Las Sociedades Mutualistas deberán diversificar y dispersar los riesgos y las responsabilidades que asuman al realizar sus operaciones, a través de la celebración de contratos de reaseguro con otras Instituciones de Seguros o con Reaseguradoras Extranjeras, empleando en su caso los servicios de Intermediarios de Reaseguro.

Las Sociedades Mutualistas deberán practicar las operaciones de reaseguro, en su carácter de cedentes, en términos que les permitan una adecuada diversificación de los riesgos o responsabilidades que asuman. A tal efecto, en la realización de operaciones de cesión de reaseguro, las Sociedades Mutualistas deberán procurar una adecuada dispersión en el uso de reaseguradores.

ARTÍCULO 487.- Las infracciones a esta Ley, así como a las disposiciones de carácter general que de ella emanen o a los reglamentos respectivos, que a continuación se señalan, serán sancionadas conforme a lo siguiente:

La Comisión, con acuerdo de su Junta de Gobierno, mediante disposiciones de carácter general, determinará, en cada operación o ramo, los límites máximos de retención de las Sociedades Mutualistas en un solo riesgo.

Además, será aplicable a las Sociedades Mutualistas, en lo conducente, lo previsto en los artículos 259 a 261 y 264 de esta Ley.

V. Multa por el equivalente del 5% al 15% de los excedentes que tengan las Instituciones y Sociedades Mutualistas sobre sus límites máximos de retención, conforme a lo previsto en los artículos 256 y 258 de la presente Ley;

Considerando los artículos anteriores de la Ley de Instituciones de Seguros y de Fianzas se determinan “Reglas para fijar los Límites Máximos de Retención de las Instituciones y Sociedades Mutualistas de Seguros en las operaciones de Seguro y Reaseguro”

CAPÍTULO PRIMERO

Disposiciones Generales

PRIMERA.- Para efectos de las presentes reglas, se entenderá por:

I. Comisión, a la Comisión Nacional de Seguros y Fianzas.

II. Consejo de Administración, al consejo de administración de las instituciones y sociedades mutualistas de seguros.

III. Institución, en singular o plural, a las instituciones y sociedades mutualistas de seguros.

IV. Ley, a la Ley General de Instituciones y Sociedades Mutualistas de Seguros.

V. Secretaría, a la Secretaría de Hacienda y Crédito Público.

SEGUNDA.- La fijación de los límites máximos de retención que en cada operación o ramo asuman las Instituciones, será responsabilidad del Consejo de Administración y se sujetará a lo previsto en los artículos 29 Bis, 37 y 86 de la Ley, así como a las presentes reglas.

TERCERA.- La Secretaría será el órgano competente para interpretar, aplicar y resolver para efectos administrativos todo lo relacionado con estas reglas.

CUARTA.- La Comisión, en ejercicio de las facultades de inspección y vigilancia que le otorga la Ley y sin perjuicio de lo previsto en las presentes reglas, podrá establecer la forma y términos en que las Instituciones deberán informarle y comprobarle todo lo concerniente a la fijación de sus límites máximos de retención que asuman en las operaciones o ramos de seguro y reaseguro.

Las Instituciones deberán presentar, en la forma y términos que al efecto establezca la Comisión mediante disposiciones de carácter general, el acuerdo del Consejo de Administración en el que se hayan fijado los límites de retención aprobados para cada operación o ramo.

QUINTA.- En la fijación de los límites máximos de retención, las Instituciones deberán procurar el aprovechamiento de su capacidad de retención, el desarrollo de políticas adecuadas para la cesión y aceptación, según corresponda, de reaseguro interno o externo, así como la dispersión de aquellos riesgos que por su naturaleza puedan afectar su solvencia y estabilidad.

El límite máximo de retención será la cantidad máxima que las Instituciones podrán retener en cada uno de los riesgos asegurados en las pólizas en vigor, una vez deducida la parte cedida en los diversos contratos de reaseguro en que participen, considerando como parte de dicho límite: los deducibles, franquicias, corredores o cualquier otro elemento que los contratos de reaseguro establezcan y que puedan resultar en responsabilidad que la Institución que cede el riesgo, deba asumir.

SEXTA.- Los límites máximos de retención que asuman las Instituciones por cada operación o ramo que tengan autorizado practicar, serán los que apruebe su Consejo de Administración, los cuales deberán ser fijados mediante procedimientos técnicos de valoración de riesgos, atendiendo a los principios establecidos en los artículo 37 y 86 de la Ley, así como a lo previsto en las presentes reglas.

Los límites máximos de retención deberán fijarse anualmente, al momento en que se realice el diseño de los planes anuales de reaseguro de las Instituciones. Adicionalmente, las Instituciones podrán realizar ajustes a su límite máximo de retención durante el año, siempre y cuando existan cambios importantes en su cartera de riesgos, los cuales deberán ser autorizados previamente por su Consejo de Administración y hacerlos del conocimiento de la Comisión.

El límite máximo de retención deberá fijarse al menos por cada una de las operaciones que tenga autorizada la Institución a practicar. Sin perjuicio de lo anterior, atendiendo la naturaleza de los riesgos asegurados, las Instituciones podrán establecer límites de retención por cada ramo, subramo o tipo de seguro que operen.

SÉPTIMA.- Los límites máximos de retención por cada operación, ramo, subramo o tipo de seguro que practiquen las Instituciones y que apruebe su Consejo de Administración, deberán fijarse mediante la aplicación de un método técnico que tome en cuenta lo siguiente: el volumen que represente en el ejercicio de su actividad la operación, ramo, subramo o tipo de seguro que corresponda; la calidad y el monto de sus recursos; así como el monto de las sumas en riesgo; las características de los riesgos que asuma; la composición de su cartera; la experiencia obtenida respecto al comportamiento de la siniestralidad, y sus políticas de reaseguro.

El método técnico para fijar los límites máximos de retención, deberá permitir que la Institución conozca con un alto grado de confiabilidad, que el límite de retención adoptado es un valor tal que en escenarios adversos probables de ocurrencia de siniestros, no pone en riesgo su solvencia. Se entenderá como escenarios adversos probables, aquellos en los que se considere la ocurrencia simultánea de siniestros de las pólizas con las mayores sumas aseguradas contenidas en el portafolio de pólizas en vigor de la Institución.

La fijación del límite máximo de retención deberá realizarse con la información de pólizas en vigor de la Institución, pudiendo incorporar al *cálculo carteras hipotéticas de pólizas que correspondan a los planes de negocio del año de que se trate o negocios en donde la Institución conozca su futura realización.*

En el caso de las operaciones y ramos de daños y de accidentes y enfermedades, el límite máximo de retención obtenido conforme a la aplicación del método técnico mencionado en la presente regla, deberá corresponder, por cada riesgo asegurado, a una cantidad que no sea superior al 5% de los activos computables al capital mínimo de garantía de acuerdo a los límites de inversión establecidos en las Reglas para el Capital Mínimo de Garantía de las Instituciones de Seguros (AcCMG). Cuando la Institución presente un margen de solvencia global, podrá considerar además los activos computables al capital mínimo de garantía en exceso a los límites de inversión previstos en las citadas Reglas (AcExc CMG):

*LMR = 5% * (AcCMG + AcExc CMG)*

En el caso de sociedades mutualistas de seguros, el límite máximo de retención obtenido conforme a la aplicación del método técnico mencionado en la presente regla, deberá corresponder, por cada riesgo asegurado a una cantidad que no sea superior al 5% de su patrimonio.

OCTAVA.- El método técnico a que se refiere la Séptima de estas reglas, deberá ser aprobado por el Consejo de Administración de la Institución debiendo contar previamente con la opinión favorable de un actuario certificado para el registro de notas técnicas de la operación de que se trate, quien se pronunciará respecto al impacto que puede tener sobre la solvencia de la Institución, la adopción de los límites máximos de retención que resulten del procedimiento aplicado, y de la observancia que se ha dado a los lineamientos establecidos en las presentes reglas, pudiendo sugerir cualquier cambio sobre el procedimiento técnico, escenario e hipótesis utilizadas en la determinación de dicho límite.

NOVENA.- El método técnico mencionado en la Séptima y Octava de estas reglas, los límites máximos de retención aprobados por el

Consejo de Administración y la opinión del actuario certificado a que se refieren las reglas anteriores, deberán ser resguardados por la Institución y estar disponibles en caso de que sean requeridos por la Comisión, en cualquier tiempo para efectos de inspección y vigilancia.

DÉCIMA.- Cuando la Comisión, en el ejercicio de sus facultades de inspección y vigilancia, determine que una Institución han incumplido con los límites máximos de retención aprobados por su Consejo de Administración o bien, que dichos límites no se apegan a lo establecido en los artículos 37 y 86 de la Ley, así como a lo indicado en las presentes reglas, procederá a requerir a la Institución de que se trate un plan de regularización en términos de lo señalado en el artículo 74 Bis de la Ley.

4.1.4 Información requerida para el modelo

Se utiliza la información histórica de expuestos, primas retenidas, suma asegurada, número de siniestros y monto de siniestro por operación y ramo por un período de al menos un año. La información se agrupa por rangos de suma asegurada para captar de mejor manera las diferencias en severidad que presentan las diferentes categorías de riesgos.

4.1.5 Objetivo y descripción del modelo

El objetivo es obtener la distribución de la siniestralidad anual bruta y retenida considerando diferentes niveles de retención para las diferentes operaciones y ramos operados por la compañía, mediante un proceso de simulación. Para ello se generan un número n de simulaciones y con el resultado de las mismas se obtienen los percentiles correspondientes a diferentes niveles de confianza para determinar la desviación en la siniestralidad a retención correspondiente para cada combinación. Con los resultados anteriores se analiza cual es el límite de retención adecuado considerando el nivel de tolerancia al riesgo establecido por la administración.

El método técnico para fijar los límites máximos de retención, deberá permitir que la Institución conozca con un alto grado de confiabilidad, que el límite de retención adoptado es un valor tal que, en escenarios adversos probables de ocurrencia de siniestros, no pone en riesgo su solvencia. Se entenderá como escenarios adversos probables, aquellos en los que se considere la ocurrencia simultánea de siniestros de las pólizas con las mayores sumas aseguradas contenidas en el portafolio de pólizas en vigor de la Institución.

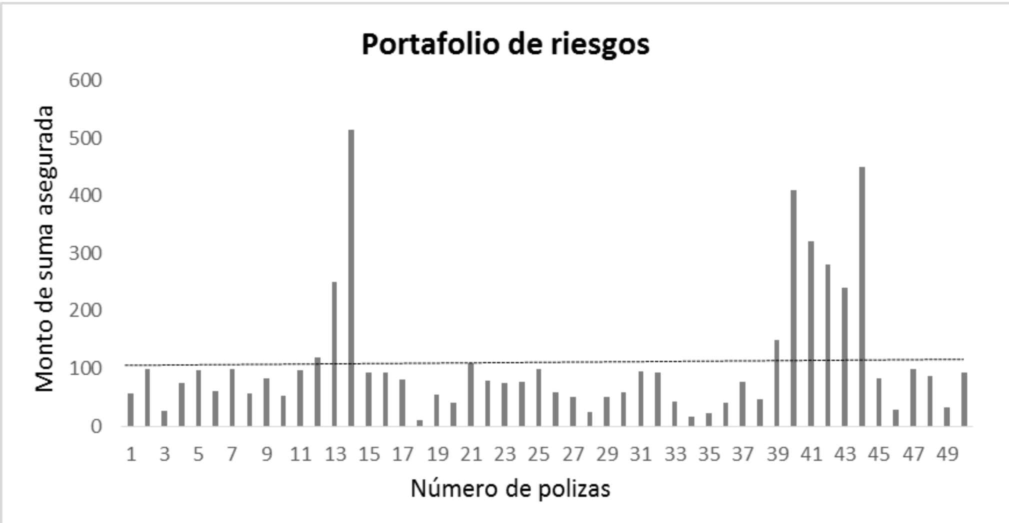
La fijación del límite máximo de retención deberá realizarse con la información de pólizas en vigor de la Institución, pudiendo incorporar al cálculo carteras hipotéticas de pólizas que correspondan a los planes de negocio del año de que se trate o negocios en donde la Institución conozca su futura realización.

4.1.6 Fuentes de incertidumbre y limitaciones de los resultados futuros

Durante el estudio, es necesario tomar como base la información proporcionada por la compañía aseguradora sin haber auditado o realizado validaciones o evaluaciones independientes. La exactitud del análisis a realizar depende directamente de la exactitud e integridad de la información entregada por la misma, por lo tanto cualquier desviación significativa que la compañía pudiera detectar, debe ser reportada para su inmediata corrección y sustitución de los resultados finales del modelo propuesto. Se entiende con lo anterior que toda la información disponible que es relevante para este análisis y que dicha información es completa y exacta.

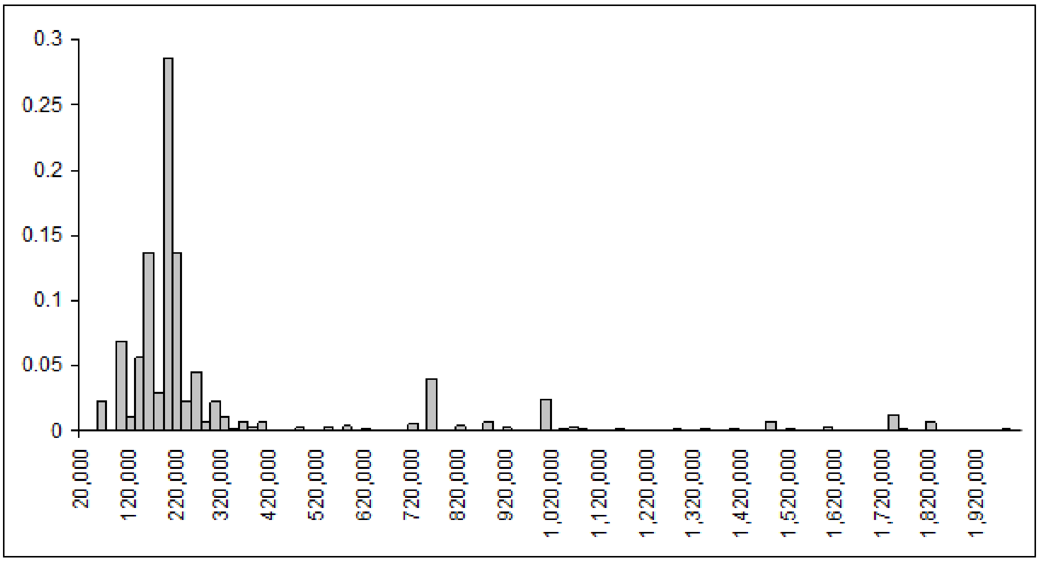
Existen algunas limitaciones respecto a la exactitud de nuestras proyecciones. Esto se debe a que la siniestralidad última está sujeta a los resultados que se presenten en los eventos futuros, como por ejemplo cambios en el perfil de cartera, cambios en la distribución de sumas aseguradas, la ocurrencia de un siniestro, cambios en los patrones de reporte de siniestros, variaciones en los alcances de las coberturas, entre otros. Por lo anterior, debe tenerse en cuenta que los resultados futuros pueden presentar desviaciones significativas respecto a la siniestralidad esperada y los percentiles de desviación de la siniestralidad esperada. En nuestra opinión, sin embargo, se han empleado técnicas y supuestos que se consideran apropiados, por lo que las conclusiones que se presentan se consideran razonables dada la información disponible.

A medida que una cartera de pólizas se va formando, se produce un desequilibrio natural en el grado de homogeneidad de las sumas aseguradas. La homogeneidad entendida como el parecido que existe entre las sumas aseguradas en riesgo de las pólizas que conforman la cartera en vigor.



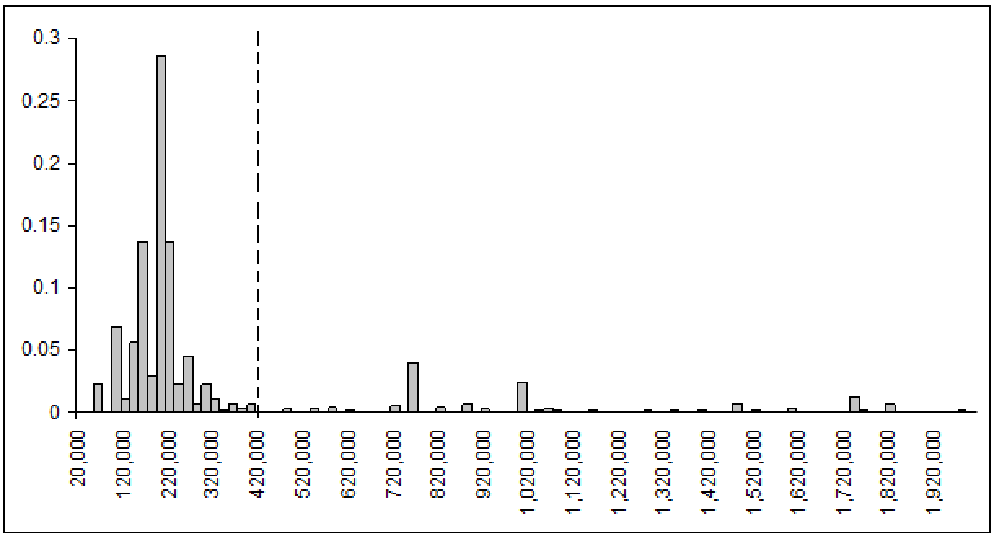
Representación de homogeneidad de riesgos de una cartera de seguros

El grado de homogeneidad se puede estudiar mediante la función de probabilidad asociada a las sumas aseguradas en riesgo. Específicamente en términos de la media y la varianza de dicha función. Se presenta a continuación la función de probabilidad de las sumas aseguradas de una cartera de 877 pólizas.



Representación gráfica de la distribución de pérdidas de una cartera de seguros

La probabilidad de pérdida está en función de la magnitud de la varianza y desviación de dicha función. Empíricamente significa que la probabilidad de que un siniestro ocurra en valores extremos, es significativa. En nuestro ejemplo existen 114 pólizas (clase 2) de un total de 877, que se encuentran alejadas de la media y que pueden generar pérdidas.



Representación gráfica de la distribución de pérdidas máximas de una cartera de seguros

4.2 DESARROLLO TÉCNICO DE LA METODOLOGÍA

El proceso de simulación se realiza considerando la información histórica de la operación y ramo que se esté analizando y ajustando funciones de distribuciones de probabilidad que describan el comportamiento del número de siniestros y severidad de los mismos de una manera adecuada.

4.2.1 Selección de la información

Como se mencionó en los requerimientos para la realización del estudio del límite máximo de retención, se requiere que la compañía aseguradora adapte las bases de datos con la información adecuada para un mejor análisis en el momento de elaborar nuestro modelo. La primera base de datos tendrá la información de siniestralidad y la segunda base la información de las primas.

En el caso de las bases de datos de siniestralidad se visualiza así:

NUMERO_SINIESTRO	FECHA_OCURRENCIA	COSTO_SINIESTRO	SUMA_SEGURADA	RAMO	FACTOR_SINIESTRALIDAD
311341130100464	12/04/2011	500	50,000,000	RESPONSABILIDAD CIVIL	0.001%
311355130100078	13/07/2011	114,028	760,183,817	INCENDIOO	0.015%
314392230000114	20/02/2014	59,765	239,061,069	INCENDIOO	0.025%
312392230000482	17/06/2012	3,005,372	300,537,213	DIVERSOS	1%
312392230000228	02/03/2012	34,946,668	174,733,339	TRANSPORTE	20%

Tabla de campos de información necesaria para el cálculo del límite de retención

Estos son los campos mínimos necesarios que se requieren para nuestro modelo, el campo de numero de siniestro no es significativo pero se utiliza para examinar todos los detalles de dicho siniestro, los siguientes campos si son significativos ya que son los que nos servirán para determinar la frecuencia y la severidad con la que se siniestran las obligaciones adquiridas. Es necesario mencionar que dichos campos poseen la información de los últimos 5 años.

Para el caso de la base de datos de primas se observa de la siguiente manera

NUMERO_POLIZA	FECHA_VIGENCIA	SUMA_SEGURADA	PRIMA	RAMO	PORCENTAJE _RETENCION
AS113411301003	12/04/2015	2,000,000	60,000	DIVERSOS	80%
AS315510100078	13/07/2015	44,598,400	1,337,952	AGROPECUARIO	15%
AS564392230014	20/02/2015	13,390,102	401,703	INCENDIOO	10%
AS812923000482	17/06/2015	4,000,000	120,000	DIVEROS	80%
AS012392200228	02/03/2015	1,000,000	30,000	TRANSPORTE	35%

Tabla de campos de información necesaria para el cálculo del límite de retención

De igual manera que en la base anterior, el primer campo no es significativo; es necesario que la fecha de vigencia este dentro de año inmediato anterior ya que al realizar la proyección de nuestro modelo, se espera que los riesgos del próximo año sean; por lo menos, los mismos que este año, es decir que se pueda renovar el 100% de la cartera actual.

En ambas bases la información requerida puede ser mayor para un mejor análisis a la hora de tomar decisiones de negocio, por lo que serían campos de información cualitativa, en el caso de ser cuantitativa solo será información complementaria como porcentaje de cesión, porcentaje de coaseguro etc.

4.2.2 Determinación de la siniestralidad

Después tener nuestras bases de datos con los campos adecuados, se procederá a determinar nuestra siniestralidad.

Se debe comenzar por filtrar todos los siniestros por ramo, debido que en nuestro modelo cada ramo se comportara de manera diferente. Se debe contemplar para nuestro modelo de simulación el comportamiento de siniestralidad de nuestros datos, es decir cómo se distribuyen nuestros riesgos

El factor de siniestralidad nos representa que porcentaje de nuestra suma asegurada se está siniestrando por lo tanto es el riego que al final se está cubriendo. Por lo que es necesario determinar la función de distribución y los parámetros con los que se están comportando los datos, previamente se realizó un análisis de nuestras sumas aseguradas conforme al factor de siniestralidad, debido a que existen cuentas en la cartera con sumas aseguradas muy grandes pero basándonos en la experiencia registrada, el factor que se siniestra es pequeño para dichas sumas aseguradas. Es necesario aclarar que aunque no se exista registro de que una póliza de suma asegurada muy grande no se siniestre con un factor de siniestralidad alto esté exento de que ocurra, se deberá considera la posibilidad de que dicho evento pueda ocurrir, es poco probable pero posible.

Consecuentemente para tener un mejor ajuste de nuestros datos a una función distribución teórica se clasifican en rangos según la suma asegurada y para un ajuste mayor, dichos rangos pueden clasificarse en distintas funciones de distribución según el porcentaje del factor de siniestralidad.

Para simular el número de siniestros por simulación se utilizará alguna distribución discreta de estas:

- *Poisson*
- *Geométrica*
- *Binomial negativa*

Para establecer dicha distribución se observará cual es el comportamiento de la siniestralidad diaria, lo cual nos determinara la frecuencia con la que se siniestra cada operación.

Para simular la severidad de los siniestros se utiliza la información de estos, clasificada por operación y ramo, y sus correspondientes sumas aseguradas. Con esta información se procede a obtener la severidad por siniestro.

$$Severidad = \frac{Siniestro}{Suma\ Asegurada}$$

Con la información de severidad se procede a encontrar una función de distribución que se ajuste a la severidad para el rango de suma asegurada seleccionado.

Entre las funciones utilizadas para encontrar la que mejor se ajuste a la severidad real de la compañía son las siguientes

- *Pareto*
- *Logística*
- *Log-logistic*
- *Exponencial*
- *Valor extremo generalizada*

- *Normal*
- *Log-normal*
- *Pearson*
- *Gamma*
- *Gumbel*
- *Weibull*
- *Goodrich*
- *Uniforme*
- *Beta*
- *Erlang*

Los parámetros de cada una de las funciones se determinan por el método máxima verosimilitud

Para demonstrar dicho análisis se tomará la información de siniestralidad de una compañía aseguradora de daños.

En la información obtenida se estudia el ramo de responsabilidad civil. Durante los últimos 5 años la compañía registro 6210 siniestros con sumas aseguradas a partir de los 10,000 hasta los 190,000,000 unidades monetarias pero con factores de siniestralidad diferentes así que para determinar los rangos se hacer cuadros comparativos para verificar en que rango son menos variantes los factores de siniestralidad.

Suma asegurada	Máx. factor siniestralidad	Mín. factor siniestralidad	Número de siniestros
0-10,000,000	100%	0%	4680
10,000,000-20,000,000	9%	0%	270
20,000,000-30,000,000	24%	0%	103
30,000,000-40,000,000	2%	0%	219
40,000,000-50,000,000	1%	0%	6
50,000,000-60,000,000	2%	0%	542
60,000,000-70,000,000	1%	0%	262
70,000,000-80,000,000	0%	0%	23
80,000,000-90,000,000	0%	0%	1
100,000,000-110,000,000	4%	0%	11
120,000,000-130,000,000	1%	0%	89
>130,000,000	0%	0%	4
Total general	100%	0%	6210

Tabla de retención de siniestralidad del ramo Responsabilidad Civil

En este caso se observa como en sumas aseguradas menores a 10,000,000 el factor de siniestralidad varía entre el .001% hasta el 100% y que en sumas aseguradas mayores el factor de siniestralidad es bajo.

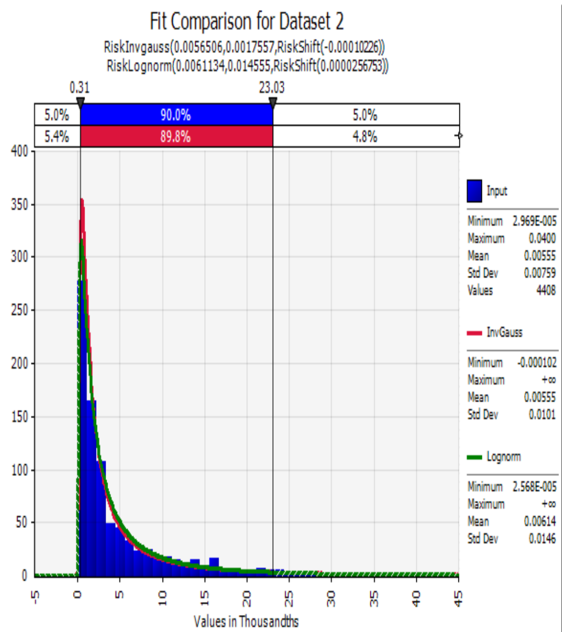
Después de establecer nuestros rangos con el análisis previo se tiene que establecer cuál es la función de distribución de nuestro factor de siniestralidad y sus parámetros. Utilizando el software mencionado en el capítulo anterior se verifica con cuál de las distintas distribuciones continuas teóricas existe un mayor ajuste.

Como se desgloso en el capítulo 2, se utilizara el método de bondad de ajuste, el cual es un caso particular de las prueba de hipótesis, para asignar nuestra distribución a trabajar. La prueba que se tomara como primer referencia es la prueba de bondad de ajuste kolmogorov-Smirnov, para considerar que el factor de siniestralidad se ajusta a dicha función de distribución el nivel de significancia α debe ser mayor al 95%, es decir que para que no se rechace la hipótesis de que los datos se comportan como esa distribución teórica la prueba de bondad de ajuste kolmogorov-Smirnov debe ser $H_0\colon F(x) \leq 1-\alpha$ así que $H_0\colon F(x) \leq 0.05$, por lo tanto las funciones de distribuciones que arrojen un resultando menor al 0.05 podrán ser consideradas para nuestro modelo, a menor valor de la prueba de bondad de ajuste nos indica que el ajuste a dicha función de distribución es mejor, en caso de que varias distribuciones tengan un valor menor al .05 también se puede escoger examinado medidas de tendencia central , medidas de dispersión y el valor del percentil al 90%, 95% y 99%. Las funciones de distribución que tengan un valor mayor al 0.05 serán discriminadas en automático.

Considerando el mismo caso del ramo de responsabilidad civil, el software @Risk realiza el paso anterior con los rangos propuestos.

Fit Ranking

Fit	K-S
InvGauss	0.0479
Lognorm	0.0504
Pearson5	0.0577
Expon	0.1855
ExtValue	0.1907
Normal	0.2352
Logistic	0.2441
LogLogistic	0.2441
Pareto	0.3655
Triang	0.4609
Uniform	0.5846
BetaGeneral	*N/A*
Gamma	*N/A*
Pearson6	*N/A*
Weibull	*N/A*



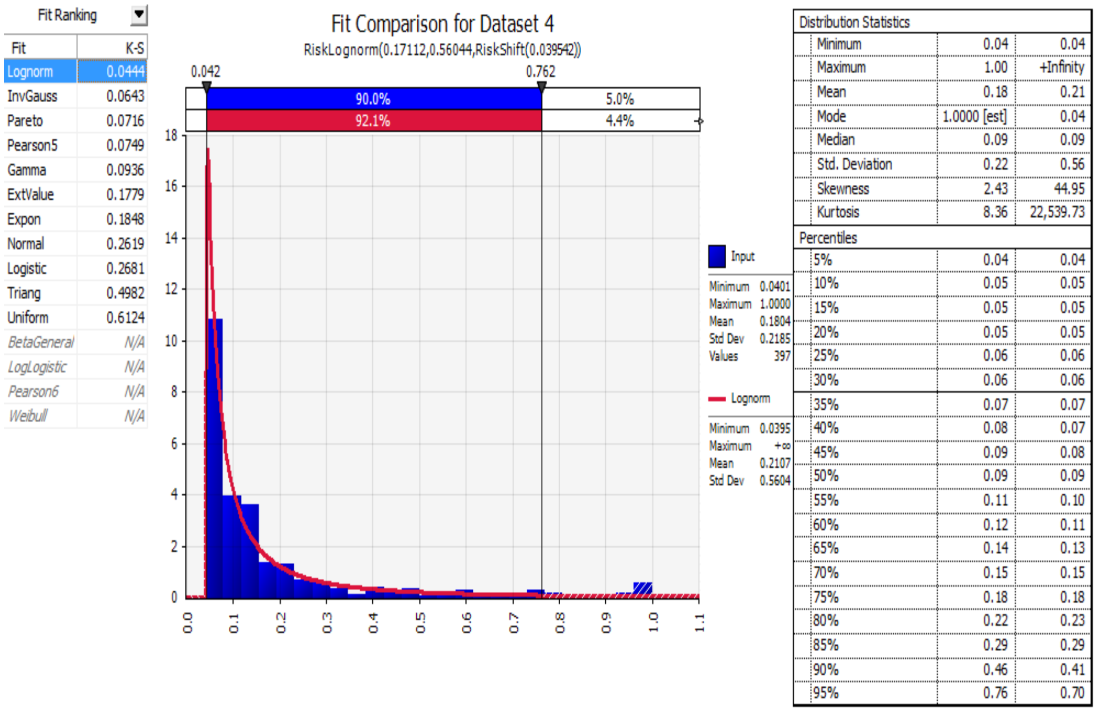
	Input	InvGauss	Lognorm
Distribution Statistics			
Minimum	2.97E-05	- 0.00	2.57E-05
Maximum	0.04	+Infinity	+Infinity
Mean	0.01	0.01	0.01
Mode	0.000315 [est]	0.00	0.00
Median	0.00	0.00	0.00
Std. Deviation	0.01	0.01	0.01
Skewness	2.25	5.38	20.64
Kurtosis	8.08	51.27	2,700.99
Percentiles			
5%	0.0003	0.0003	0.0003
10%	0.0004	0.0004	0.0004
15%	0.0005	0.0006	0.0006
20%	0.0007	0.0007	0.0008
25%	0.0009	0.0009	0.0010
30%	0.0011	0.0011	0.0012
35%	0.0015	0.0013	0.0014
40%	0.0016	0.0015	0.0017
45%	0.0018	0.0018	0.0020
50%	0.0024	0.0022	0.0024
55%	0.0029	0.0026	0.0028
60%	0.0032	0.0031	0.0034
65%	0.0041	0.0037	0.0041
70%	0.0052	0.0045	0.0049
75%	0.0067	0.0056	0.0060
80%	0.0089	0.0072	0.0076
85%	0.0117	0.0096	0.0099
90%	0.0157	0.0136	0.0139
95%	0.0230	0.0223	0.0228

Representación gráfica del ajuste de curvas del ramo de Responsabilidad Civil en @Risk

Como se puede observar la funciones de distribución de Lognormal y la Inverse gaussian son las únicas que pueden ser consideradas por la prueba de kolmogorov-Smirnov y las demás distribuciones que se muestran son descartadas por no cumplir con el nivel de significancia por caso práctico la función de distribución Lognormal se asemeja más al percentil 95% a los datos de entrada por lo que se tomará prioridad a dicha distribución.

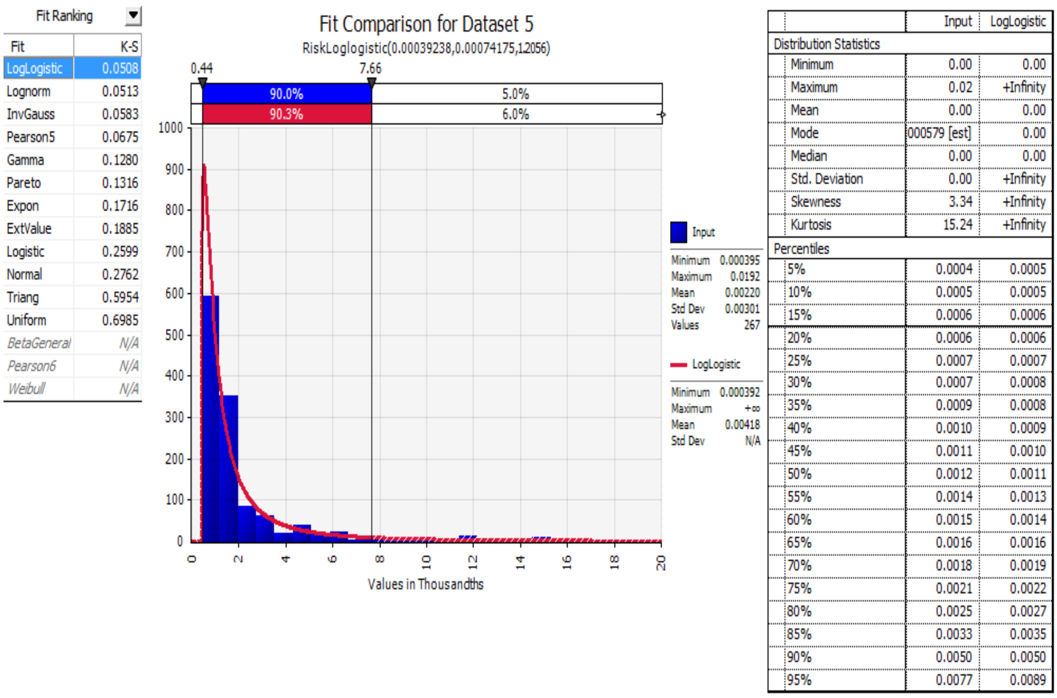
Recordando que en primer instancia del análisis de los datos se clasifico en rangos por suma asegurada, la pantalla anterior corresponde al 40% de los datos de la suma asegurada menor o igual a 10,000,000 unidades monetarias que se comportan como un función de distribución Lognormal con parámetros $\mu = 0.0061134$ y $\sigma = 0.000025675$.

El otro 60% de los datos también se comportaron como un función de distribución Lognormal pero con distintos parámetros $\mu = 0.17112$y $\sigma = 0.039542$.



Representación gráfica del ajuste de curvas del ramo de Responsabilidad Civil en @Risk

Para el rango de la sumas aseguradas mayores a 10,000,000 unidades monetarias son similares a la función de distribución Loglogistic con los parámetros $\gamma = 0.0003923$, $\alpha = 1.205$ y $\beta = 0.00074175$



Representación gráfica del ajuste de curvas del ramo de Responsabilidad Civil en @Risk

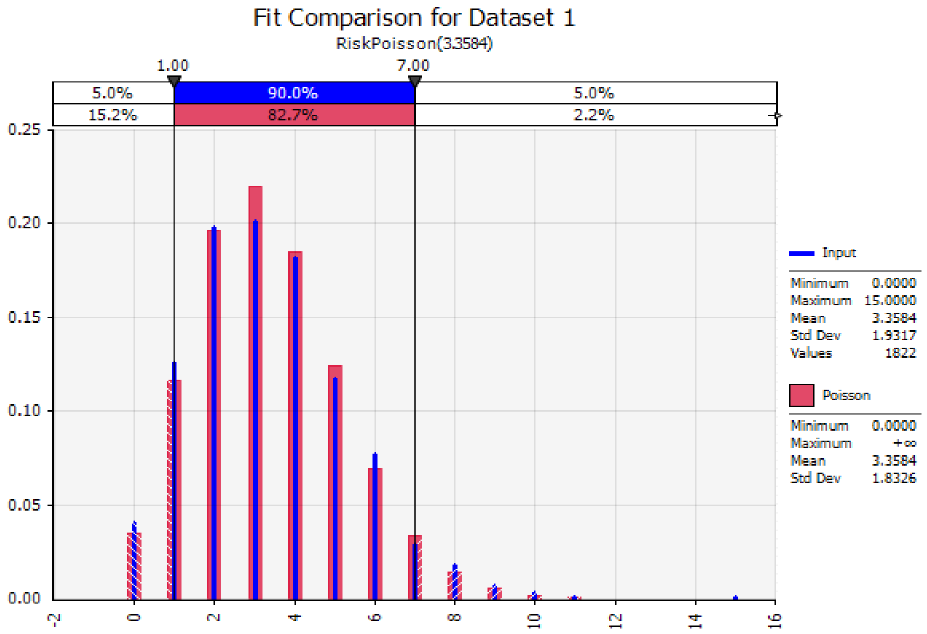
Un pequeño porcentaje de los datos de este rango son atípicos por lo que no se ajustan correctamente a la distribución anterior pero no se deben discriminar del estudiar porque eso no significa que exista un siniestro con un factor de siniestralidad alto en este rango así que se tomará en cuenta el 1% de los datos como una función de distribución uniforme con parámetros $a = .01$ y $b = 1$, entonces nuestro cuadro resumen de las distribuciones a considerar en la severidad de siniestralidad queda de la siguiente manera .

RC	Datos Total	6210
S.A.<= 10,000,000	fact < = .04	
	Datos	4408
	Dist	log norm
	fact > .04	
	Datos	397
	Dist	log norm
S.A.>10,000,000	fact < .01	
	Datos	1380
	Dist	Loglogistic
	fact >= .01	
	Datos	25
	Dist	Uniforme U(0.01.1)

Tabla de distribuciones por rango del ramo de Responsabilidad Civil

Posteriormente para establecer la frecuencia se utilizarán las fechas de ocurrencia del siniestro, como es un numero finito de datos la función de distribución de la frecuencia de siniestralidad debe ser una distribución discreta.

Siguiendo con el caso del ramo anterior la distribución que se empleará solo en este ramo es función distribución Poisson con parámetro $\lambda = 3.3584$.



Para los ramos de Daños sin autos se puede generalizar la función de distribución discreta ,con normalidad las distribuciones utilizadas es la Poisson o la binomial negativa debido a que la diferencia entre los parámetros de cada distribución es pequeña, para efectos de nuestro modelo se decidió emplear la distribución que más se ajuste. El valor del parámetro que se muestra en nuestra distribución es la frecuencia de siniestralidad diaria, en nuestro modelo se debe multiplicar por los 365 días del año para mostrar la frecuencia de siniestralidad anual.

Entonces se obtuvo la frecuencia de siniestralidad la cual nos indica el número de siniestros que ocurren por año ,tomando como referencia los últimos 5 años, para modelar un predicción del número de siniestros para nuestro próximo año; de igual manera se mide la severidad con la que se siniestran nuestras obligaciones adquiridas.

Con la información obtenida por la compañía aseguradora se replicó el análisis estadístico anterior con otros ramos que opera la compañía para seguro de daños sin autos como son: incendio, transporte, diversos y un caso poco común agropecuario. El cuadro resumen de las distribuciones que se obtuvieron por ramo se muestran a continuación.

Transporte	Datos Total	7453
S.A.< 10,000,000	fact < = .5	
	Datos	6816
	Dist	Pearson
	fact > .5	
	Datos	345
	Dist	Beta general
S.A.>= 10,000,000	fact < .01	
	Datos	266
	Dist	Beta general
	fact >= .01	
	Dist	Uniforme
	23	U(.01,0.05)
	3	U(.05,1)

Tabla de distribuciones por rango del ramo de Transporte

Incendio	Datos Total	4086
S.A.< 20,000,000	Datos	3639
	Dist	Lognorm
S.A.>= 20,000,000	fact < .00016	
	Datos	284
	Dist	Beta general
	0.00016 <= fact < = .01	
	Datos	116
	Dist	Beta general
	fact > .01	
	Dist	Uniforme
	16	U(.01,0.02)
	9	U(.02,0.15)
	2	U(.15, 1)

Tabla de distribuciones por rango del ramo de Incendio

Agropecuario	Datos Total	7487
S.A.< 30,000,000	Datos	7322
	Dist	Lognorm
S.A.>= 30,000,000	fact < .08	
	Datos	156
	Dist	Beta general
	fact => .08	
	Dist	Uniforme
	9	U(.01,1)

Tabla de distribuciones por rango del ramo de Agropecuario

Diversos		
	Datos Total	33259
S.A.< =12,000,000	fact < .6	
S.A.< =12,000,000	Datos	26490
S.A.< =12,000,000	Dist	Lognorm
S.A.< =12,000,000	0.6 <= fact < = .9	
S.A.< =12,000,000	Datos	846
S.A.< =12,000,000	Dist	Beta general
S.A.< =12,000,000	fact > .9	
S.A.< =12,000,000	Datos	759
S.A.< =12,000,000	Dist	Weibull
S.A.> 12,000,000	fact < = .01	
S.A.> 12,000,000	Datos	5142
S.A.> 12,000,000	Dist	exponencial
S.A.> 12,000,000	fact > .01	
S.A.> 12,000,000	Dist	Uniforme
S.A.> 12,000,000	19	U(.01,0.05)
S.A.> 12,000,000	3	U(.05,1)

Tabla de distribuciones por rango del ramo de Diversos

Como ya fue señalado, para la elaboración de este modelo se manejaran distribuciones discretas distintas por ramo que son las siguientes.

ramo	Diaria		Distribucion
	# siniestros	maximo	
RC	6210	15	Poisson
Carga	7453	15	Bin.Neg
Incendio	4066	21	Bin.Neg
Agropecuario	7487	175	Geometrica
Diversos	33259	64	Bin.Neg

Tabla resumen de las distribuciones de frecuencia por ramo

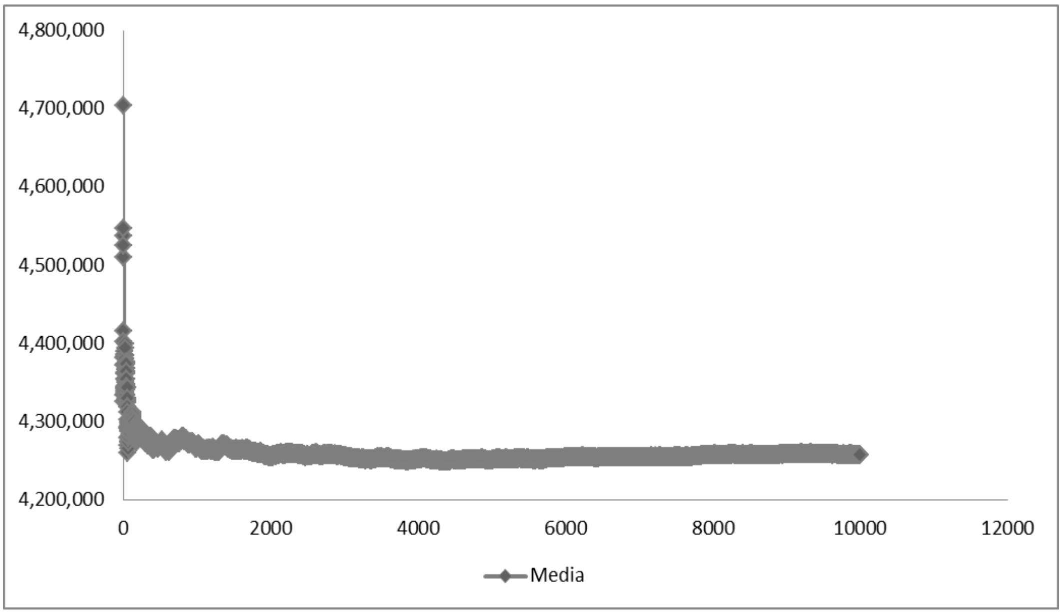
Los parámetros y las gráficas de los datos de entrada de cada una de las distribuciones por ramo se ven reflejadas en las pantallas del anexo 1 al final del documento, con esto ya se puede proceder a la elaboración del modelo de simulación.

4.2.3 Simulaciones de los montos de siniestralidad

Con los parámetros obtenidos con las técnicas estadísticas anteriores, se tiene el conocimiento de la tendencia de los datos de entrada por cada ramo según los rangos definidos. Entonces ya se puede elaborar las simulaciones correspondientes para medir el riesgo de los montos siniestrados a futuro y los elementos principales a considerar para el modelo de simulación son los números de siniestros y montos por siniestro.

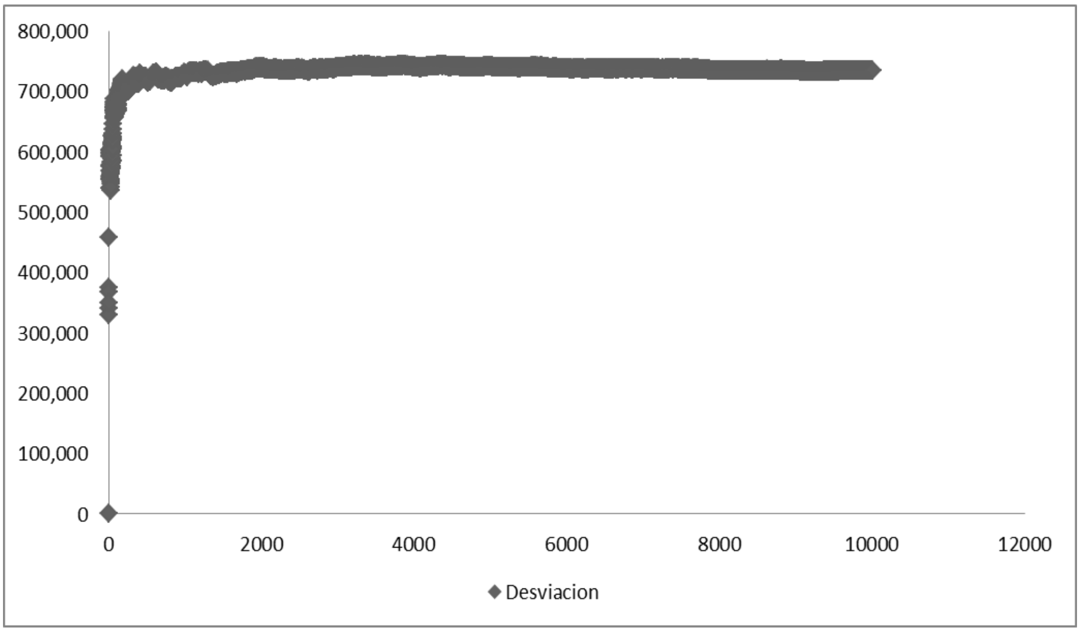
Como se hizo mención en el apartado anterior, el número de siniestros a simular está dado por la función de frecuencia seleccionada con base en la experiencia histórica y los montos de siniestros se simulan con la función de severidad, la cual generará números aleatorios afectando el factor de siniestralidad.

Para saber cuántas iteraciones se necesitan para confiar que el modelo de simulación es veraz, es necesario hacer un ejercicio previo donde se toma una cifra cualquiera, que este caso se utilizará un siniestro bruto al azar y se generarán números aleatorios que se comporten como una función de distribución uniforme, conforme se generen los números aleatorios se calculará la media y la desviación respecto a la media de la función acumulativa de cada conjunto de datos y en el momento donde se observe que converge estas medidas de tendencia se determinará el número de iteraciones que se aplicarán en dicho modelo.



Representación gráfica de la convergencia simulaciones

Como se explicó anteriormente al calcular la media a cada conjunto de datos, se observa como la media converge a un número conforme la cantidad de datos va incrementando. En este caso se nota como la media converge a los 10,000 datos.



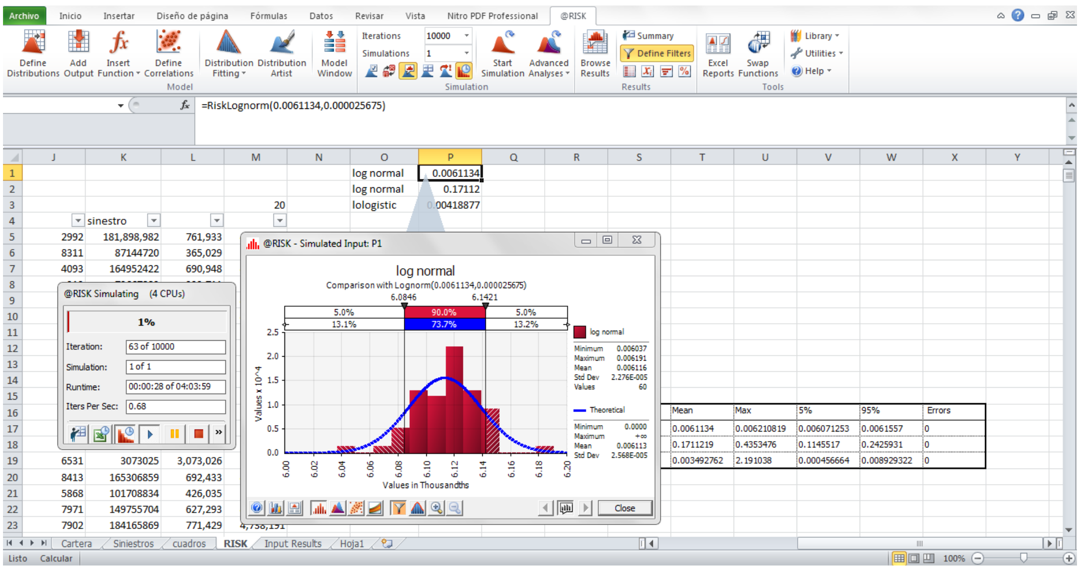
Representación gráfica de la convergencia simulaciones

De igual manera, se examina la tendencia de la desviación respecto a la media y la convergencia se comporta de manera similar, por lo tanto 10,000 serán las iteraciones que se aplicarán en cada ramo del modelo de simulación.

Dándole seguimiento al ramo de responsabilidad civil en la estimación de pérdidas se toma una frecuencia de 1230 siniestros que resulta del cálculo de nuestra frecuencia de siniestralidad diaria de una distribución Poisson por los días del año y se simularan los factores de siniestralidad con las distribuciones y parámetros establecidos previamente que generarán 10,000 iteraciones de dichos factores que se aplicarán a montos brutos de siniestralidad seleccionados al azar de nuestra cartera en vigor.

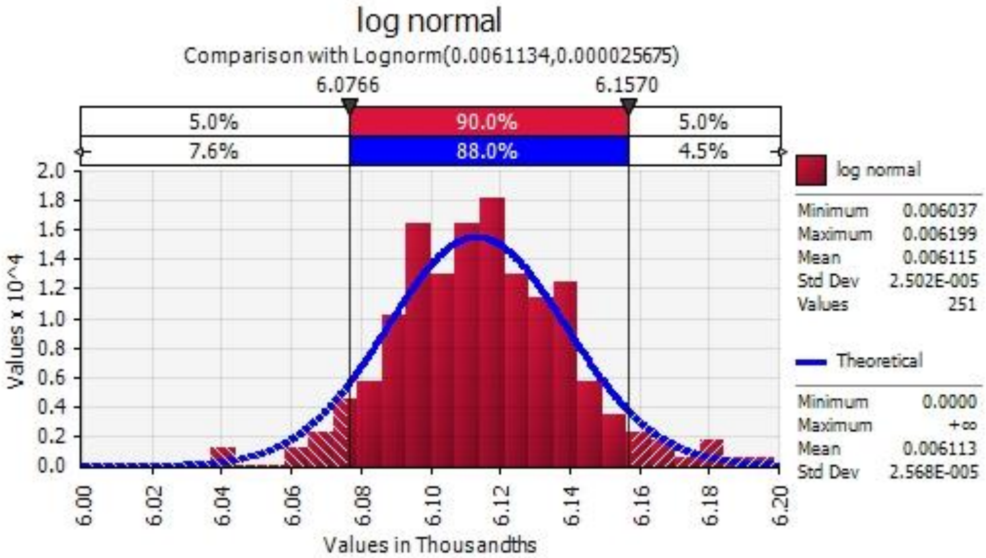
Recordando las funciones de distribución empleadas en este ramo, las simulaciones serán ejecutadas en el software @Risk en el cual solo define la distribución a simular estableciendo los parámetros obtenidos. Cada simulación se empleará según el rango al que pertenece la suma asegurada y se realizará el proceso de las iteraciones en los campos del factor de siniestralidad que nos proporcionarán lo montos de siniestralidad para la estimación de perdidas, el diseño de modelo se emplea en Excel constituyendo un modelo dinámico.

Al ejecutar las simulaciones se debe espera que el proceso complete el 100% de ellas, una pantalla de la ejecución del inicio de las iteraciones en @risk para el ramo de responsabilidad civil es la siguiente:

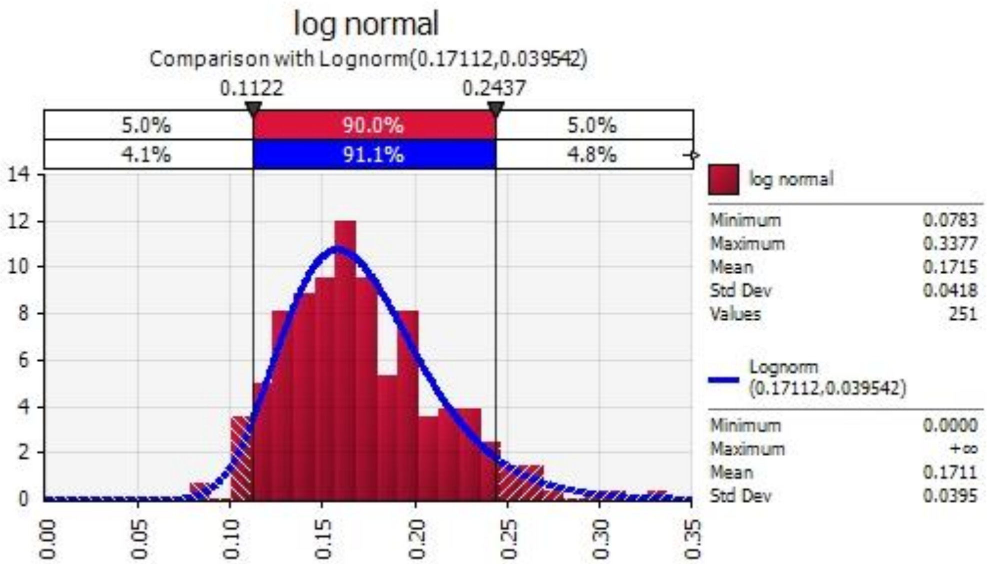


Representación gráfica del proceso de simulación en @Risk

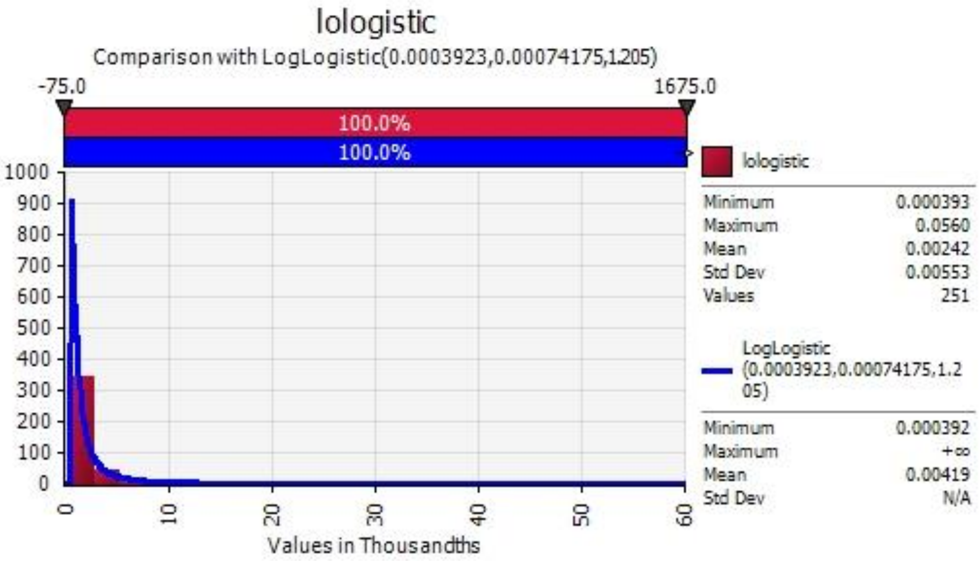
Como se muestra en la pantalla se están generando iteraciones de los factores de siniestralidad de una distribución Lognormal antes definida, el proceso va comenzando y debe cumplir con el 100% , dicho proceso se debe repetir según las distribuciones propuestas por ramo, en este caso en particular son tres distribuciones y los resultados son los siguientes.



Representación gráfica de los resultados de las simulaciones en @Risk en el ramo de Responsabilidad civil



Representación gráfica de los resultados de las simulaciones en @Risk en el ramo de Responsabilidad civil



Representación gráfica de los resultados de las simulaciones en @Risk en el ramo de Responsabilidad civil

Al final de todos los procesos de simulación el software @risk ilustra un cuadro resumen de las funciones de distribución simuladas.

Distribución	Cell	Graph	Min	Mean	Max	5%	95%	Errors
log normal	P1	6.00ths 6.25ths	0.006009581	0.0061134	0.006210819	0.006071253	0.0061557	0
log normal	P2	0.00 0.45	0.06418912	0.1711219	0.4353476	0.1145517	0.2425931	0
lologistic	P3	-0.5 2.5	0.0003924	0.003492762	2.191038	0.000456664	0.008929322	0

Cuadro resumen de las simulación en @Risk para el ramo de Responsabilidad Civil

Obtenida nuestra estimación de pérdidas simuladas se procede al diseño de varios escenarios para obtención de límite máximo de retención óptimo.

4.2.4 Aplicación de los límites de retención en diversos escenarios

A cada monto de siniestro correspondiente a la simulación se le aplica el límite de retención que se esté probando.

Una vez aplicado dicho límite a cada uno de los siniestros de la simulación, se suma el total de siniestros por simulación, obteniendo así el vector de monto de siniestros retenidos totales de la simulación.

Con las simulaciones se obtienen la perdida media y los valores extremos de las pérdidas (cuantiles 95%, 99.5%). Con la base de datos de la cartera en vigor se muestra cuáles son los riegos que estuvieron por lo menos un día en vigor y se tiene el riego de perdida de dicha obligación. Como se definió para estimar lar pérdidas a futuro se espera que por lo menos las obligaciones de la cartera actual de la aseguradora se renueven, de tal manera que bajo este vigor se emplearan las simulaciones de los montos de siniestralidad. El objetivo siguiente es saber que monto del siniestro es pertinente retener según la solvencia de la compañía.

La pérdida media se constituye con el valor esperado, las perdidas mayores al ingreso por prima de la cartera anual en vigor y los valores extremos con el cuantil al 95% y 99.5%. Para determinar los límites de retención es necesario obtener la desviación en la pérdidas esperadas el cual se calculó como $Desviación = PM - VM$, cumpliendo con las reglas que estableció la Comisión Nacional de Seguros y Fianzas la desviación de pérdidas esperadas, debe ser mayor al 95%.

Culminado con el caso del ramo de responsabilidad civil, se establecen 5 escenarios con diferentes límites de retención, en cada escenario se calcula la perdida media y los valores extremos para obtener la desviación, en cada escenario del límite de retención propuesto va incrementado y consecuentemente la retención del siniestro bruto también incrementa. Los resultados de cada escenario son los siguientes

límite de retención propuesto	Desviación 95%	Desviación 99.5%
10,000,000	-1,968,047	-3,033,548
25,000,000	-18,876,579	-21,070,272
50,000,000	-50,504,089	-54,017,945
70,000,000	-73,501,254	-77,449,298
80,000,000	-83,689,097	-88,393,402

Cuadro resumen de la desviación de siniestralidad del ramo de Responsabilidad Civil

En caso de que se cuenten con pólizas partículas cedidas al 100%, estas no estarán incluidas en el modelo.

4.2.5 Determinar el límite máximo de retención de una compañía hipotética

Cumpliendo con las reglas para determinar el límite de retención para cada una de las operaciones de dicho ramo, el límite de retención se fijará de acuerdo a los siguientes criterios.

- Que la desviación de la siniestralidad retenida sea menor que el capital asignado a la operación.
- Que la desviación se encuentre dentro de los rangos aceptables a los diferentes niveles de confianza, alrededor de la media representan la desviación que se puede observar en periodos entre 2 o 10 años y la desviación al percentil 99.5% representa la necesidad de capital para asumir pérdidas y que son comparables con los requerimientos locales de capital.

Cumpliendo con la regla 7 para determina el límite máximo de retención, dice que limite no deben superar al 5% de los activos computables. Para la compañía aseguradora que ese está trabajando los activos computables al capital es de 1,200,000,000 así el límite debe ser menor a 60,000,000 en caso específicos que sea mayor y se quiera establecer dicho límite se necesita la aprobación de un actuario independiente.

De tal manera que el límite debe ser menor al 60,000,000 y la desviación de pérdidas esperadas tampoco debe supera esta requerimiento de solvencia . Utilizando estos escenarios se analizan los porcentajes de la desviación respecto al 5% de los activos computables y después establecer el límite máximo de retención óptimo.

límite de retención propuesto	10,000,000	25,000,000	50,000,000	70,000,000	80,000,000
Desviación 95%	4,058,154	15,951,597	26,292,536	39,503,296	52,844,984
Desviación 99.5%	3,033,548	21,146,443	54,333,385	77,450,455	89,032,026
Desv 95%/AC	7%	27%	44%	66%	88%
Desv 99.5%/AC	5%	35%	91%	129%	148%

Cuadro resumen de la desviación de siniestralidad del ramo de Responsabilidad Civil con diferentes límites de retención propuestos

Para concluir se puede observar que a mayor porcentaje de perdida esperada la desviación es mayor, si se tomara como referencia la desviación al 95% la estimación de pérdidas no supera los 60,000,000 unidades monetarias, en cambio si la referencia es la desviación al 99.5% la estimación de pérdidas se ajusta de mejor manera al límite de retención. Por lo tanto el límite máximo de retención óptimo es de 50,000,000 unidades monetarias para el ramo de responsabilidad civil.

4.3 ANÁLISIS DE RESULTADOS: CONCLUSIONES Y SUGERENCIAS.

A lo largo de esta investigación se derivaron varias conclusiones. En cumplimiento a la normatividad del cálculo del límite máximo de retención de los ramos de Daños sin autos, el diseño del modelo de simulación propuesto cumple con la hipótesis previamente establecida, y se puede comprobar con los resultados finales que se presentan en este apartado de los requerimientos formales del estudio.

El procedimiento de aplicar los límites de retención en varios escenarios para determinar el límite máximo de retención se debe repetir con los otros ramos desarrollados de manera alterna al ramo de responsabilidad civil.

límite de retención propuesto	10,000,000	25,000,000	50,000,000	70,000,000	80,000,000
Responsabilidad Civil	3,033,548	21,146,443	54,333,385	77,450,455	89,032,026
Transporte	6,144,534	17,221,981	51,442,171	81,324,645	90,097,146
Incendio	5,983,023	24,875,690	61,946,785	87,345,672	96,783,456
Agropecuario	3,147,588	18,374,650	48,093,495	66,849,384	76,029,381
Diversos	2,093,847	15,637,409	42,304,950	65,748,320	78,930,347

Cuadro resumen de la desviación de siniestralidad de los ramos de Daños sin autos

Como se puede apreciar, las desviaciones en la siniestralidad de cada ramo se aproximan al límite de retención de 50,000,000 unidades monetarias y cumple perfectamente con ser menor de 5% de los activos computables, aquí no fue necesario solicitar la aprobación de un actuario independiente porque ningún límite propuesto rebasó la regla de los activos computables.

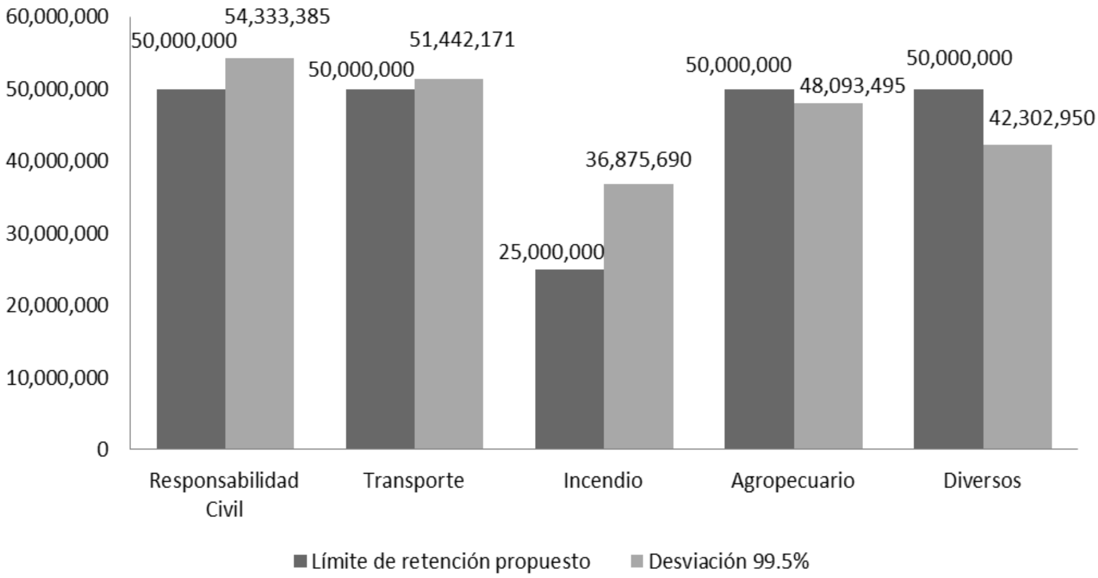
límite de retención propuesto	10,000,000	25,000,000	50,000,000	70,000,000	80,000,000
Responsabilidad Civil	5%	35%	91%	129%	148%
Transporte	10%	29%	86%	136%	150%
Incendio	17%	61%	110%	146%	161%
Agropecuario	5%	31%	80%	111%	127%
Diversos	3%	26%	71%	110%	132%

Cuadro resumen del porcentaje de desviación de siniestralidad respecto a los activos computables de los ramos de Daños sin autos

Tomando una postura conservadora en el caso exclusivo de incendio se dispuso que el límite máximo de retención sería de 25,000,000 unidades monetarias porque la desviación de siniestralidad excedía el 5% de los activos computables, quizá en una postura agresiva cabía la posibilidad de emplear el límite máximo de retención como 50,000,000 unidades monetarias

Ramo	Límite de retención propuesto	Desviación 99.5%
Responsabilidad Civil	50,000,000	54,333,385
Transporte	50,000,000	51,442,171
Incendio	25,000,000	36,875,690
Agropecuario	50,000,000	48,093,495
Diversos	50,000,000	42,304,950

Cuadro resumen de la desviación de siniestralidad de los ramos de Daños sin autos con límite de máximo de retención óptimo sugerido.



Representación gráfica de la desviación de siniestralidad de los ramos de Daños sin autos con límite de máximo de retención óptimo sugerido.

En concreto, se expone en el cuadro anterior el límite máximo de retención óptimo de cada ramo, al mismo tiempo se presenta la desviación de siniestralidad al 95% correspondiente.

Operación	Requerimiento Bruto de Solvencia	Desviación 99.5%
Daños sin autos	305,912,030	233,049,690

Cuadro comparativo entre la desviación de siniestralidad de los ramos de Daños sin autos contra el requerimiento de capital correspondiente

Por último se compacta todo el estudio de la desviación de siniestralidad total por operación, es decir se suman todas las desviaciones de siniestralidad de cada ramo correspondiente a la operación de daños sin autos y debe ser menor al requerimiento bruto de solvencia de la misma, que es el monto de recursos que las Instituciones deben mantener para enfrentar la exposición a desviaciones en la siniestralidad esperada de las distintas operaciones del seguro, la exposición a quebrantos por insolvencia de reaseguradores, y la exposición a las fluctuaciones adversas en el valor de los activos que respaldan a las obligaciones contraídas con los asegurados.

Por lo tanto, se puede decir hipótesis que se planteó al principio de esta investigación, la cual establece que ***la metodología de cálculo de las obligaciones que puede asumir una compañía de seguros puede ser estimada de manera óptima a través de un modelo de simulación que se ajuste a la cartera en vigor, y a la normatividad aplicable; obteniendo como resultado el límite máximo de retención en los ramos de Daños sin autos de una compañía de seguros es cierta,*** y refleja eficientemente el nivel de obligaciones que podrá enfrentar una Compañía aseguradora en su operación anual. No obstante debe señalarse que dicho modelo debe estar en permanente actualización para reflejar fielmente la realidad y la dinámica de la actividad aseguradora.

Como ya se mencionó, una correcta metodología del cálculo óptimo del Límite Máximo de Retención busca preservar la seguridad de las Compañías aseguradoras, ahí radica la importancia de la práctica adecuada de dicho estudio. Los beneficios de la elaboración del modelo propuesto son aprovechar la capacidad de retención de éstas y proteger su solidez financiera, a fin de garantizar el cumplimento de sus obligaciones con los asegurados. Por lo general el estudio del límite máximo de retención lo realizan las consultorías actuariales externas, debido a que las compañías aseguradoras le delegan dicha responsabilidad como un proceso de auditoría para medir su solvencia, se validan los resultados con las áreas técnicas de la aseguradora y se entregan los resultados al consejo de administración de la misma para la toma de decisiones.

Como se mencionó en los antecedentes, el Límite Máximo de Retención se establecía como el 5% de los activos computables, en comparación a la metodología desarrollada en esta investigación, las decisiones para tomar el Límite Máximo de Retención se ajustan a la siniestralidad histórica de la compañía, hecho que genera un pronóstico más exacto.

Dentro de las compañías aseguradoras el modelo óptimo determinado para el Límite Máximo de Retención debe ser de fácil operación para el área de suscripción, el área técnica y el área de administración de riesgos porque los datos requeridos para la metodología propuesta es información que dichas áreas deben manejarla y adaptarla a las necesidades del modelos, ya que el proceso de elaboración de esta metodología se facilitaría si trabajaran dichas áreas de la compañía en conjunto para evitar contratar alguna consultaría actuarial externa a la compañía para que realice el estudio.

Las conclusiones se basan en una serie de supuestos de condiciones y eventos futuros. Dichos supuestos, que están documentados en secciones subsecuentes de la presente investigación, deben ser comprendidos a fin de poder entender nuestras conclusiones en el contexto apropiado.

4.4 RESUMEN

En este capítulo se desarrolló el **modelo óptimo** del *límite máximo de retención* de las obligaciones que asume una Compañía de Seguros.

Una vez establecidos los fundamentos requeridos a lo largo de los capítulos anteriores se procedió a su aplicación práctica en nuestro modelo de optimización, de acuerdo a lo siguiente.

El primer paso fue analizar y estandarizar las bases de datos históricas de los ramos de Daños sin autos para determinar la frecuencia y el factor de severidad, que nos indicará que porcentaje de la suma asegurada se afectó.

Después se identificó qué distribución se ajusta para los casos de severidad y que distribución se ajusta para los casos de la frecuencia, de tal manera que se puedan obtener los parámetros que se utilizarán en las simulaciones.

Posteriormente al ajuste de curvas se ejecutaron las simulaciones al modelo.

Finalmente, se determinaron los límites máximos de retención de las obligaciones que asume una Compañía de Seguros para observar cuál es el óptimo y se presentarán las conclusiones del modelo, aceptándose la hipótesis planteada en esta investigación.

ANEXOS

I. Resultados de factores de siniestralidad en @Risk y definición de distribuciones.

Responsabilidad civil

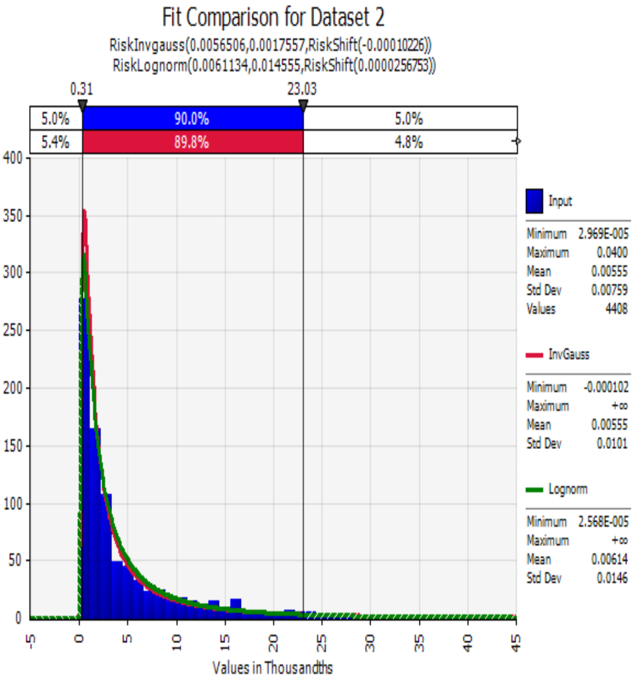
RANGOS		MÁXIMO	MÍNIMO	PROMEDIO	#SINIESTROS
Valores menores a	10,000,000	100%	0%	2%	4680
10,000,000	20,000,000	9%	0%	0%	270
20,000,000	30,000,000	24%	0%	1%	103
30,000,000	40,000,000	2%	0%	0%	219
40,000,000	50,000,000	1%	0%	0%	6
50,000,000	60,000,000	2%	0%	0%	542
60,000,000	70,000,000	1%	0%	0%	262
70,000,000	80,000,000	0%	0%	0%	23
80,000,000	90,000,000	0%	0%	0%	1
100,000,000	110,000,000	4%	0%	1%	11
120,000,000	130,000,000	1%	0%	0%	89
Valores mayores a	130000000	0%	0%	0%	4
Total general					6210

RC	Datos Total	6210
S.A.<= 10,000,000	fact < = .04	
	Datos	4408
	Dist	log norm
	fact > .04	
	Datos	397
	Dist	log norm
S.A.>10,000,000	fact < .01	
	Datos	1380
	Dist	Loglogistic
	fact >= .01	
	Datos	25
	Dist	Uniforme U(0.01.1)

Fact < = .04   70.9%

Fit Ranking

Fit	K-S
InvGauss	0.0479
Lognorm	0.0504
Pearson5	0.0577
Expon	0.1855
ExtValue	0.1907
Normal	0.2352
Logistic	0.2441
LogLogistic	0.2441
Pareto	0.3655
Triang	0.4609
Uniform	0.5846
BetaGeneral	*N/A*
Gamma	*N/A*
Pearson6	*N/A*
Weibull	*N/A*

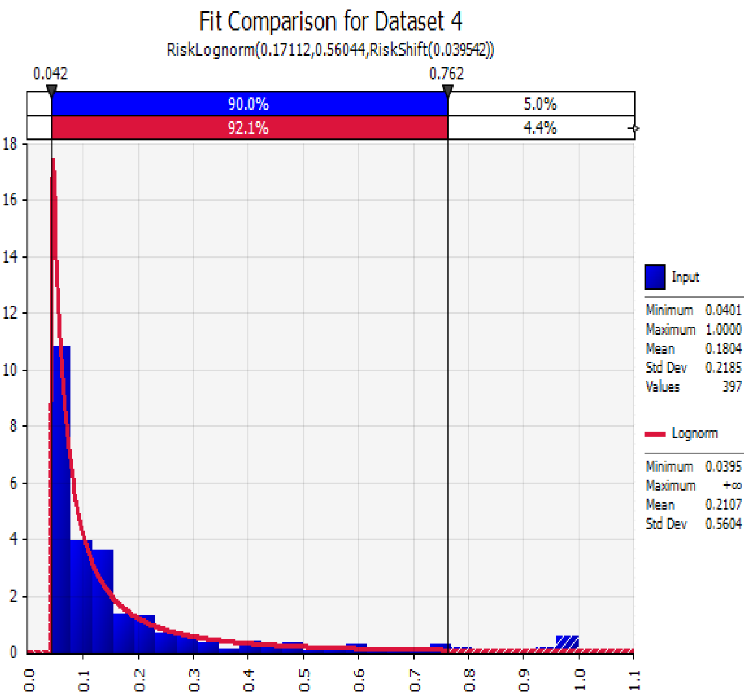


	Input	InvGauss	Lognorm
Distribution Statistics			
Minimum	2.97E-05	-0.00	2.57E-05
Maximum	0.04	+Infinity	+Infinity
Mean	0.01	0.01	0.01
Mode	.000315 [est]	0.00	0.00
Median	0.00	0.00	0.00
Std. Deviation	0.01	0.01	0.01
Skewness	2.25	5.38	20.64
Kurtosis	8.08	51.27	2,700.99
Percentiles			
5%	0.0003	0.0003	0.0003
10%	0.0004	0.0004	0.0004
15%	0.0005	0.0006	0.0006
20%	0.0007	0.0007	0.0008
25%	0.0009	0.0009	0.0010
30%	0.0011	0.0011	0.0012
35%	0.0015	0.0013	0.0014
40%	0.0016	0.0015	0.0017
45%	0.0018	0.0018	0.0020
50%	0.0024	0.0022	0.0024
55%	0.0029	0.0026	0.0028
60%	0.0032	0.0031	0.0034
65%	0.0041	0.0037	0.0041
70%	0.0052	0.0045	0.0049
75%	0.0067	0.0056	0.0060
80%	0.0089	0.0072	0.0076
85%	0.0117	0.0096	0.0099
90%	0.0157	0.0136	0.0139
95%	0.0230	0.0223	0.0228

Fact > .04 6.3% Datos

Fit Ranking

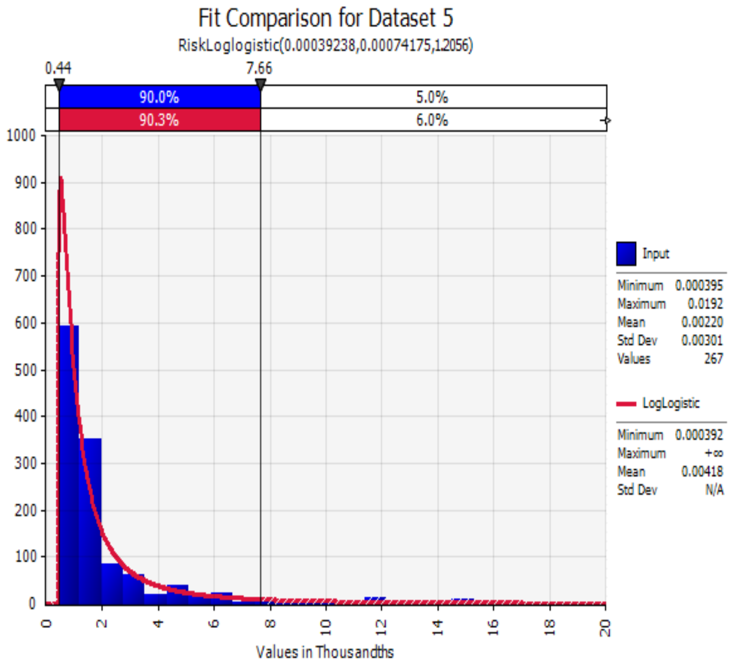
Fit	K-S
Lognorm	0.0444
InvGauss	0.0643
Pareto	0.0716
Pearson5	0.0749
Gamma	0.0936
ExtValue	0.1779
Expon	0.1848
Normal	0.2619
Logistic	0.2681
Triang	0.4982
Uniform	0.6124
BetaGeneral	*N/A*
LogLogistic	*N/A*
Pearson6	*N/A*
Weibull	*N/A*



Distribution Statistics		
Minimum	0.04	0.04
Maximum	1.00	+Infinity
Mean	0.18	0.21
Mode	1.0000 [est]	0.04
Median	0.09	0.09
Std. Deviation	0.22	0.56
Skewness	2.43	44.95
Kurtosis	8.36	22,539.73
Percentiles		
5%	0.04	0.04
10%	0.05	0.05
15%	0.05	0.05
20%	0.05	0.05
25%	0.06	0.06
30%	0.06	0.06
35%	0.07	0.07
40%	0.08	0.07
45%	0.09	0.08
50%	0.09	0.09
55%	0.11	0.10
60%	0.12	0.11
65%	0.14	0.13
70%	0.15	0.15
75%	0.18	0.18
80%	0.22	0.23
85%	0.29	0.29
90%	0.46	0.41
95%	0.76	0.70

Fact < .01 22.2% Datos

Fit	K-S
LogLogistic	0.0508
Lognorm	0.0513
InvGauss	0.0583
Pearson5	0.0675
Gamma	0.1280
Pareto	0.1316
Expon	0.1716
ExtValue	0.1885
Logistic	0.2599
Normal	0.2762
Triang	0.5954
Uniform	0.6985
BetaGeneral	N/A
Pearson6	N/A
Weibull	N/A



	Input	LogLogistic
Distribution Statistics		
Minimum	0.00	0.00
Maximum	0.02	+Infinity
Mean	0.00	0.00
Mode	000579 [est]	0.00
Median	0.00	0.00
Std. Deviation	0.00	+Infinity
Skewness	3.34	+Infinity
Kurtosis	15.24	+Infinity
Percentiles		
5%	0.0004	0.0005
10%	0.0005	0.0005
15%	0.0006	0.0006
20%	0.0006	0.0006
25%	0.0007	0.0007
30%	0.0007	0.0008
35%	0.0009	0.0008
40%	0.0010	0.0009
45%	0.0011	0.0010
50%	0.0012	0.0011
55%	0.0014	0.0013
60%	0.0015	0.0014
65%	0.0016	0.0016
70%	0.0018	0.0019
75%	0.0021	0.0022
80%	0.0025	0.0027
85%	0.0033	0.0035
90%	0.0050	0.0050
95%	0.0077	0.0089

Transporte

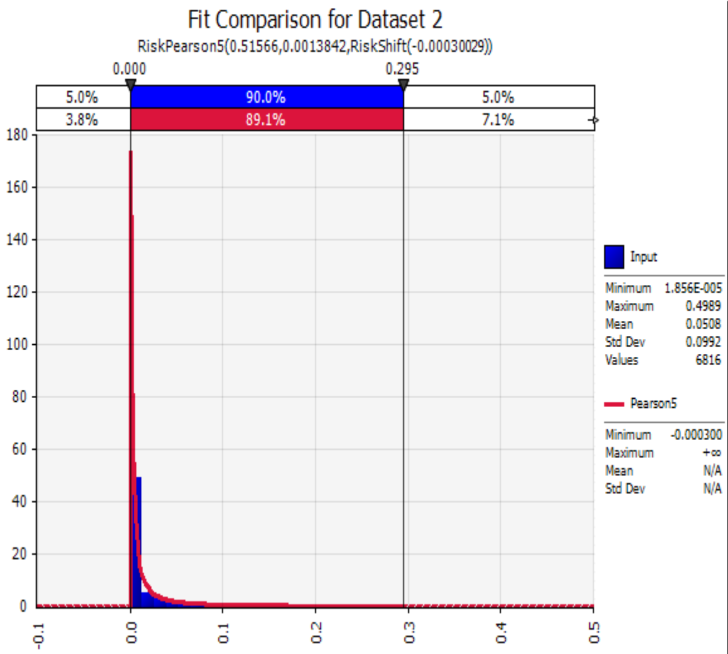
RANGOS		MÁXIMO	MÍNIMO	PROMEDIO	#SINIESTROS
Valores menores a	10,000,000	100%	0%	8%	7161
10,000,000	20,000,000	4%	0%	1%	78
20,000,000	30,000,000	21%	0%	1%	161
30,000,000	40,000,000	2%	2%	2%	1
70,000,000	80,000,000	0%	0%	0%	51
100,000,000	110,000,000	0%	0%	0%	1
Total general					7453

Transporte	Datos Total	7453
S.A.< 10,000,000	fact < = .5	
	Datos	6816
	Dist	Pearson
	fact > .5	
	Datos	345
	Dist	Beta general
S.A.>= 10,000,000	fact < .01	
	Datos	266
	Dist	Beta general
	fact >= .01	
	Dist	Uniforme
	23	U(.01,0.05)
	3	U(.05,1)

Fact < = .5 91.4%Datos

Fit Ranking

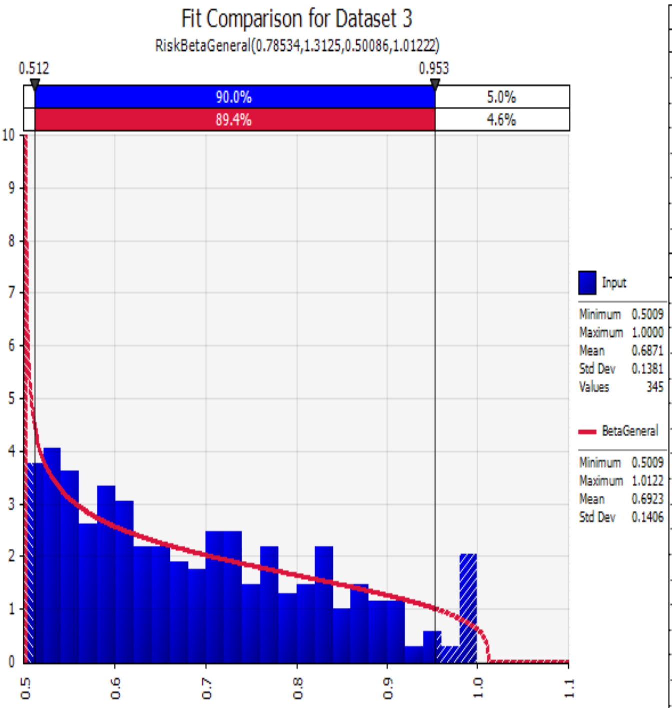
Fit	K-S
Pearson5	0.0546
InvGauss	0.0799
Lognorm	0.1051
ExtValue	0.3097
Normal	0.3134
Logistic	0.3374
LogLogistic	0.3374
Pareto	0.3430
Expon	0.4549
Triang	0.6131
Uniform	0.6775
BetaGeneral	*N/A*
Gamma	*N/A*
Pearson6	*N/A*
Weibull	*N/A*



	Input	Pearson5
Distribution Statistics		
Minimum	1.86E-05	- 0.00
Maximum	0.50	+Infinity
Mean	0.05	+Infinity
Mode	0.00700 [est]	0.00
Median	0.01	0.01
Std. Deviation	0.10	+Infinity
Skewness	2.51	+Infinity
Kurtosis	8.86	+Infinity
Percentiles		
5%	0.0003	0.0004
10%	0.0006	0.0007
15%	0.0009	0.0010
20%	0.0012	0.0013
25%	0.0017	0.0017
30%	0.0023	0.0022
35%	0.0030	0.0027
40%	0.0040	0.0034
45%	0.0053	0.0043
50%	0.0069	0.0054
55%	0.0070	0.0070
60%	0.0085	0.0091
65%	0.0140	0.0121
70%	0.0226	0.0168
75%	0.0382	0.0245
80%	0.0673	0.0384
85%	0.1164	0.0680
90%	0.1941	0.1507
95%	0.2951	0.5814

Fact > .5   4.63% Datos

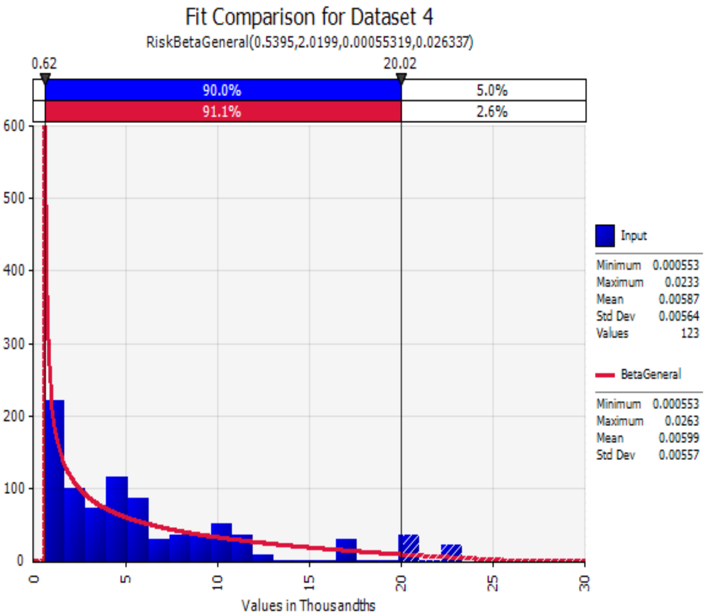
Fit	K-S
BetaGeneral	0.0436
Triang	0.0476
Weibull	0.0687
Pearson5	0.0712
ExtValue	0.0764
LogLogistic	0.0771
Lognorm	0.0781
InvGauss	0.0810
Pareto2	0.0907
Expon	0.0917
Logistic	0.1034
Rayleigh	0.1063
Normal	0.1159
Pareto	0.1206
Uniform	0.1969
Erf	0.7626
ChiSq	N/A
Erlang	N/A
Gamma	N/A
Pearson6	N/A
Student	N/A



	Input	BetaGeneral	Triang
Distribution Statistics			
Minimum	0.50	0.50	0.50
Maximum	1.00	1.01	1.06
Mean	0.69	0.69	0.69
Mode	1.0000 [est]	0.50	0.50
Median	0.66	0.67	0.66
Std. Deviation	0.14	0.14	0.13
Skewness	0.58	0.45	0.57
Kurtosis	2.34	2.06	2.40
Percentiles			
5%	0.51	0.51	0.52
10%	0.53	0.52	0.53
15%	0.54	0.54	0.54
20%	0.56	0.55	0.56
25%	0.56	0.57	0.58
30%	0.59	0.59	0.59
35%	0.60	0.61	0.61
40%	0.62	0.63	0.63
45%	0.64	0.65	0.65
50%	0.66	0.67	0.66
55%	0.69	0.69	0.69
60%	0.71	0.72	0.71
65%	0.73	0.74	0.73
70%	0.76	0.77	0.75
75%	0.78	0.80	0.78
80%	0.82	0.83	0.81
85%	0.85	0.87	0.84
90%	0.88	0.90	0.88
95%	0.95	0.95	0.94

Fact < .01   3.57% Datos

Fit	K-S
BetaGeneral	0.0528
Pearson5	0.1026
ExtValue	0.1053
Expon	0.1142
Pareto2	0.1195
InvGauss	0.1202
Lognorm	0.1310
Lognorm2	0.1310
Rayleigh	0.1650
Normal	0.1732
Logistic	0.1793
Pareto	0.2257
Triang	0.2594
Uniform	0.4362
ChiSq	0.5047
Erf	0.5271
Erlang	N/A
Gamma	N/A
LogLogistic	N/A
Pearson6	N/A
Student	N/A
Weibull	N/A



	Input	BetaGeneral
Distribution Statistics		
Minimum	0.0006	0.0006
Maximum	0.0233	0.0263
Mean	0.0059	0.0060
Mode	000642 [est]	0.0006
Median	0.0045	0.0040
Std. Deviation	0.0056	0.0056
Skewness	1.5016	1.1736
Kurtosis	4.7053	3.6152
Percentiles		
5%	0.0006	0.0006
10%	0.0006	0.0007
15%	0.0009	0.0009
20%	0.0011	0.0011
25%	0.0015	0.0015
30%	0.0022	0.0018
35%	0.0025	0.0023
40%	0.0031	0.0028
45%	0.0039	0.0034
50%	0.0045	0.0040
55%	0.0047	0.0048
60%	0.0053	0.0057
65%	0.0059	0.0066
70%	0.0071	0.0077
75%	0.0081	0.0090
80%	0.0097	0.0105
85%	0.0102	0.0123
90%	0.0130	0.0146
95%	0.0200	0.0178

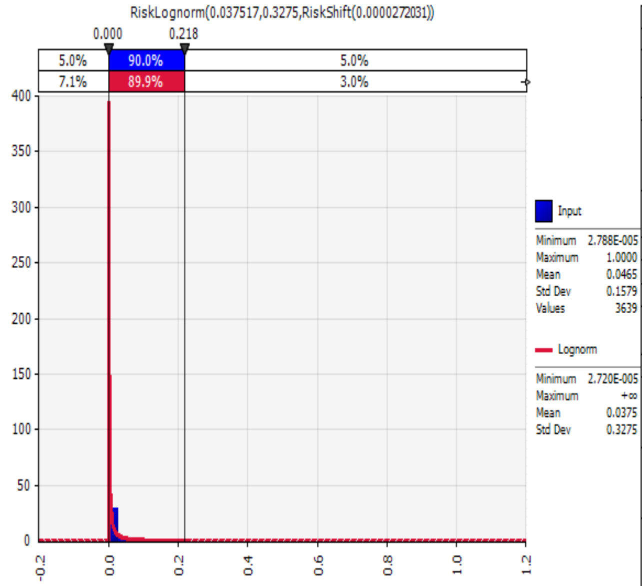
Incendio

RANGOS		MÁXIMO	MÍNIMO	PROMEDIO	#SINIESTROS
Valores menores a	20,000,000	100%	0%	5%	3639
20,000,000	120,000,000	21%	0%	0%	281
120,000,000	220,000,000	96%	0%	1%	73
220,000,000	320,000,000	0%	0%	0%	30
320,000,000	420,000,000	2%	0%	0%	14
420,000,000	520,000,000	8%	8%	8%	1
520,000,000	620,000,000	2%	0%	0%	10
620,000,000	720,000,000	2%	0%	1%	2
720,000,000	820,000,000	1%	0%	0%	2
820,000,000	920,000,000	1%	0%	1%	3
1,020,000,000	1,120,000,000	0%	0%	0%	2
1,120,000,000	1,220,000,000	0%	0%	0%	1
1,920,000,000	2,020,000,000	0%	0%	0%	1
2,120,000,000	2,220,000,000	0%	0%	0%	2
2,320,000,000	2,420,000,000	0%	0%	0%	3
2,520,000,000	2,620,000,000	0%	0%	0%	1
3,420,000,000	3,520,000,000	0%	0%	0%	1
Total general					4066

incendio	Datos Total	4086
S.A.< 20,000,000	Datos	3639
	Dist	Lognorm
S.A.>= 20,000,000	fact < .00016	
	Datos	284
	Dist	Beta general
	0.00016 <= fact < = .01	
	Datos	116
	Dist	Beta general
	fact > .01	
	Dist	Uniforme
	16	U(.01,0.02)
	9	U(.02,0.15)
	2	U(.15, 1)

89.5% Datos

Fit	K-S
Lognorm	0.0407
InvGauss	0.0436
Pearson5	0.0437
Pareto	0.3058
ExtValue	0.3408
Normal	0.3843
Rayleigh	0.4444
Expon	0.4931
Triang	0.7852
Uniform	0.8358
BetaGeneral	*N/A*
Erlang	*N/A*
Gamma	*N/A*
Logistic	*N/A*
LogLogistic	*N/A*
Pearson6	*N/A*
Student	*N/A*
Weibull	*N/A*

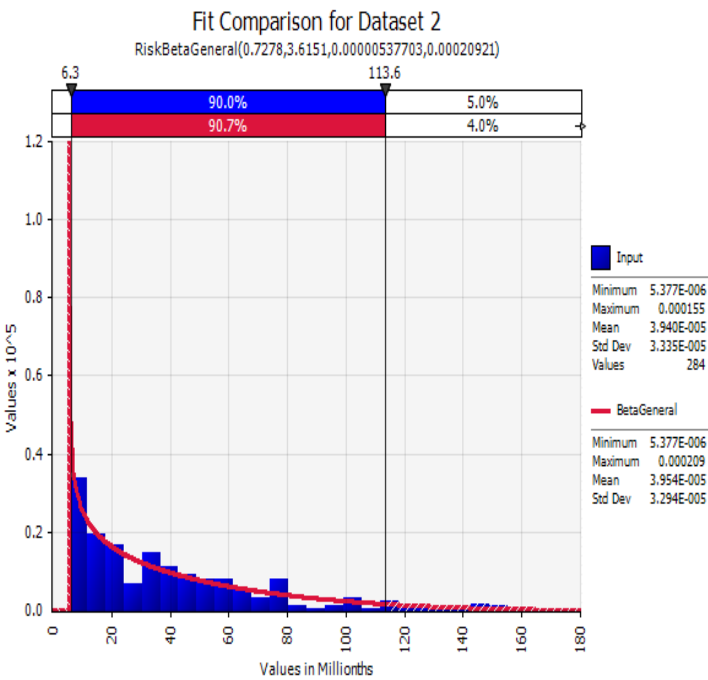


	Input	Lognorm	InvGauss	Pearson5
Distribution Statistics				
Minimum	2.79E-05	2.72E-05	- 0.00	- 0.00
Maximum	1.00	+Infinity	+Infinity	+Infinity
Mean	0.05	0.04	0.05	+Infinity
Mode	1.0000 [est]	8.25E-05	0.00	0.00
Median	0.00	0.00	0.00	0.00
Std. Deviation	0.16	0.33	0.26	+Infinity
Skewness	4.82	691.39	16.42	+Infinity
Kurtosis	26.40	36,462,418	452.19	+Infinity
Percentiles				
5%	0.0002	0.0002	0.0003	0.0003
10%	0.0004	0.0003	0.0004	0.0004
15%	0.0005	0.0005	0.0006	0.0006
20%	0.0007	0.0008	0.0008	0.0008
25%	0.0010	0.0011	0.0010	0.0010
30%	0.0013	0.0015	0.0013	0.0012
35%	0.0016	0.0019	0.0016	0.0016
40%	0.0022	0.0026	0.0019	0.0019
45%	0.0029	0.0033	0.0024	0.0024
50%	0.0037	0.0043	0.0030	0.0030
55%	0.0047	0.0056	0.0039	0.0038
60%	0.0060	0.0073	0.0050	0.0050
65%	0.0080	0.0096	0.0065	0.0066
70%	0.0111	0.0128	0.0089	0.0090
75%	0.0157	0.0174	0.0126	0.0130
80%	0.0231	0.0247	0.0189	0.0201
85%	0.0360	0.0371	0.0312	0.0349
90%	0.0645	0.0618	0.0605	0.0755
95%	0.2181	0.1318	0.1657	0.2785

Fact < .00016 6.98%

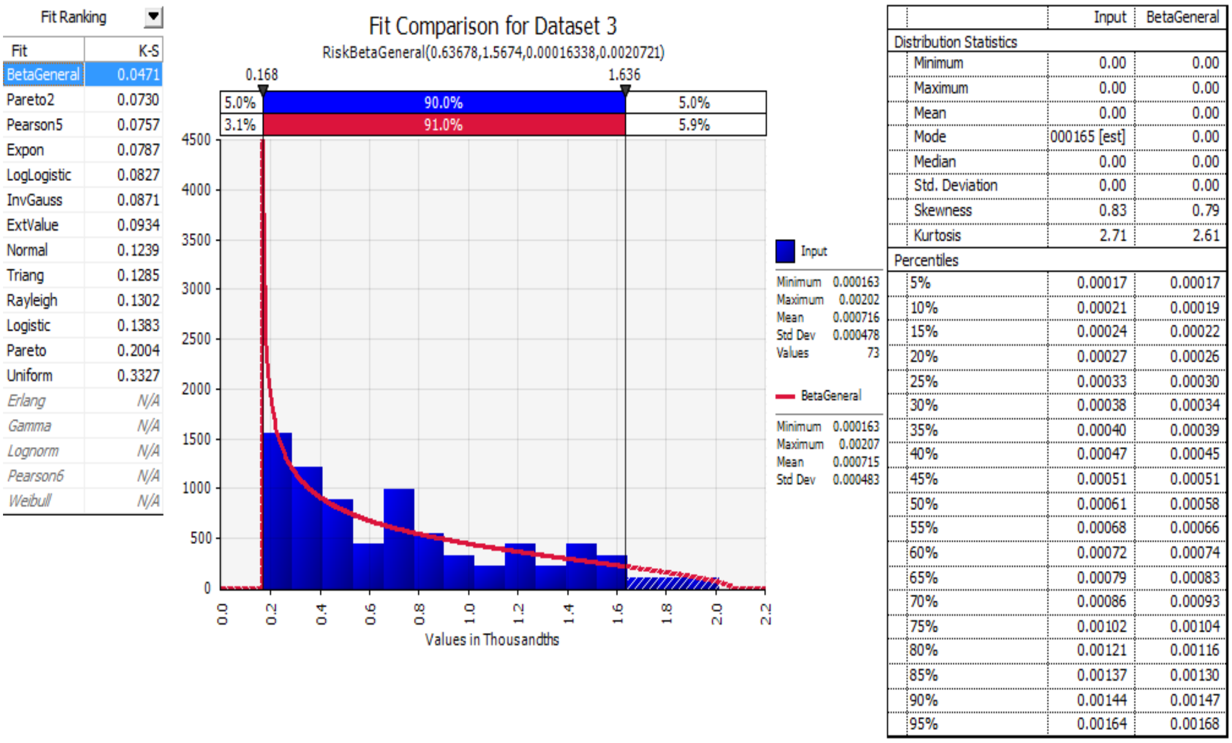
Fit Ranking

Fit	K-S
BetaGeneral	0.0328
Expon	0.0657
LogLogistic	0.0832
Lognorm	0.0953
InvGauss	0.0969
ExtValue	0.0996
Normal	0.1539
Logistic	0.1604
Pareto	0.2058
Triang	0.2222
Uniform	0.4397
Erlang	*N/A*
Gamma	*N/A*
Pearson5	*N/A*
Pearson6	*N/A*
Weibull	*N/A*



	Input	BetaGeneral
Distribution Statistics		
Minimum	5.38E-06	5.38E-06
Maximum	0.00	0.00
Mean	3.94E-05	3.95E-05
Mode	292E-06 [est]	5.38E-06
Median	3.15E-05	2.92E-05
Std. Deviation	3.34E-05	3.29E-05
Skewness	1.39	1.30
Kurtosis	4.68	4.36
Percentiles		
5%	6.29E-06	6.23E-06
10%	7.64E-06	7.60E-06
15%	8.79E-06	9.31E-06
20%	1.05E-05	1.13E-05
25%	1.30E-05	1.36E-05
30%	1.57E-05	1.61E-05
35%	1.83E-05	1.89E-05
40%	2.20E-05	2.20E-05
45%	2.66E-05	2.54E-05
50%	3.15E-05	2.92E-05
55%	3.32E-05	3.34E-05
60%	3.85E-05	3.81E-05
65%	4.37E-05	4.34E-05
70%	4.88E-05	4.93E-05
75%	5.48E-05	5.63E-05
80%	6.15E-05	6.45E-05
85%	7.15E-05	7.46E-05
90%	7.92E-05	8.79E-05
95%	1.14E-04	1.08E-04

0.00016 <= Fact < = .01 2.85% Datos



Agropecuario

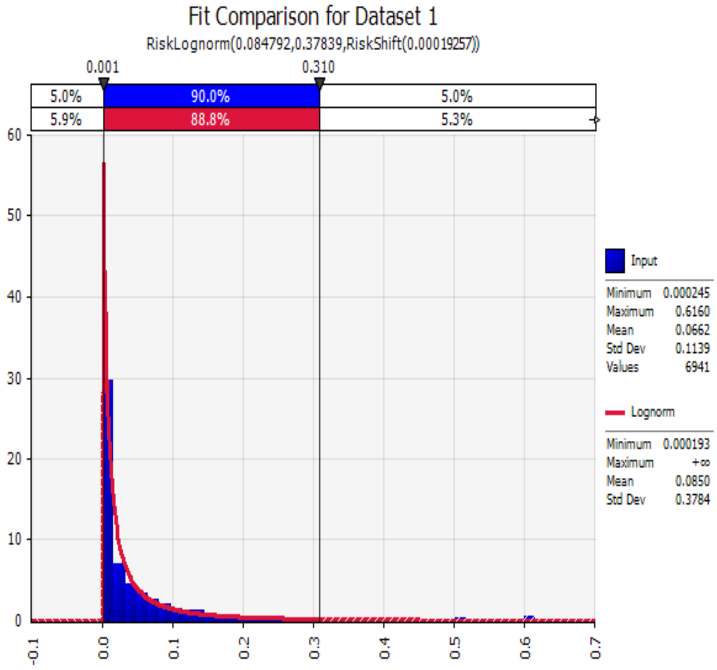
RANGOS		MAXIMO	MINIMO	PROMEDIO	#SINIESTROS
Valores menores a	30,000,000	100%	0%	9%	7322
30,000,000	60,000,000	8%	0%	1%	86
60,000,000	90,000,000	5%	0%	1%	32
90,000,000	120,000,000	33%	1%	12%	6
120,000,000	150,000,000	5%	0%	3%	4
150,000,000	180,000,000	4%	1%	2%	2
180,000,000	210,000,000	43%	1%	16%	7
210,000,000	240,000,000	38%	1%	10%	5
240,000,000	270,000,000	3%	1%	2%	8
330,000,000	360,000,000	7%	0%	2%	4
360,000,000	390,000,000	2%	1%	1%	4
420,000,000	450,000,000	4%	0%	1%	5
450,000,000	480,000,000	12%	1%	6%	2
Total general					7487

Agropecuario	Datos Total	7487
S.A.< 30,000,000	Datos	7322
	Dist	Lognorm
S.A.>= 30,000,000	fact < .08	
	Datos	156
	Dist	Beta general
	fact => .08	
	Dist	Uniforme
	9	U(.01,1)

97.8% Datos

Fit Ranking

Fit	K-S
Lognorm	0.0567
InvGauss	0.0754
Pearson5	0.0770
Gamma	0.1031
ExtValue	0.2052
Expon	0.2688
Normal	0.2813
Logistic	0.2862
LogLogistic	0.2862
Pareto	0.3204
Triang	0.5314
Uniform	0.6478
BetaGeneral	N/A
Pearson6	N/A
Weibull	N/A

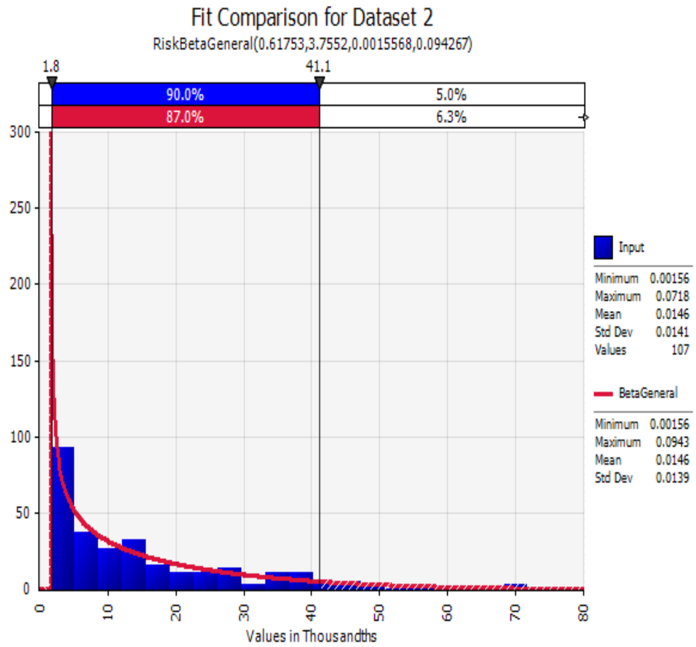


	Input	Lognorm
Distribution Statistics		
Minimum	0.00	0.00
Maximum	0.62	+Infinity
Mean	0.07	0.09
Mode	).00627 [est]	0.00
Median	0.02	0.02
Std. Deviation	0.11	0.38
Skewness	2.90	102.26
Kurtosis	11.91	210,945.30
Percentiles		
5%	0.001	0.001
10%	0.003	0.002
15%	0.003	0.003
20%	0.004	0.004
25%	0.005	0.006
30%	0.006	0.008
35%	0.007	0.010
40%	0.010	0.012
45%	0.014	0.015
50%	0.018	0.019
55%	0.025	0.023
60%	0.033	0.029
65%	0.044	0.037
70%	0.056	0.047
75%	0.073	0.060
80%	0.095	0.081
85%	0.129	0.113
90%	0.184	0.173
95%	0.310	0.327

Fact < .08 2.08% Datos

Fit Ranking

Fit	K-S
BetaGeneral	0.0368
LogLogistic	0.0938
Expon	0.1047
Pearson5	0.1057
Lognorm	0.1112
InvGauss	0.1165
ExtValue	0.1250
Normal	0.1762
Logistic	0.1880
Pareto	0.1902
Triang	0.3298
Uniform	0.4880
Gamma	*N/A*
Pearson6	*N/A*
Weibull	*N/A*



	Input	BetaGeneral
Distribution Statistics		
Minimum	0.00	0.00
Maximum	0.07	0.09
Mean	0.01	0.01
Mode	.00449 [est]	0.00
Median	0.01	0.01
Std. Deviation	0.01	0.01
Skewness	1.48	1.50
Kurtosis	5.15	5.10
Percentiles		
5%	0.00	0.00
10%	0.00	0.00
15%	0.00	0.00
20%	0.00	0.00
25%	0.00	0.00
30%	0.00	0.00
35%	0.01	0.01
40%	0.01	0.01
45%	0.01	0.01
50%	0.01	0.01
55%	0.01	0.01
60%	0.01	0.01
65%	0.01	0.02
70%	0.02	0.02
75%	0.02	0.02
80%	0.03	0.02
85%	0.03	0.03
90%	0.04	0.03
95%	0.04	0.04

Diversos

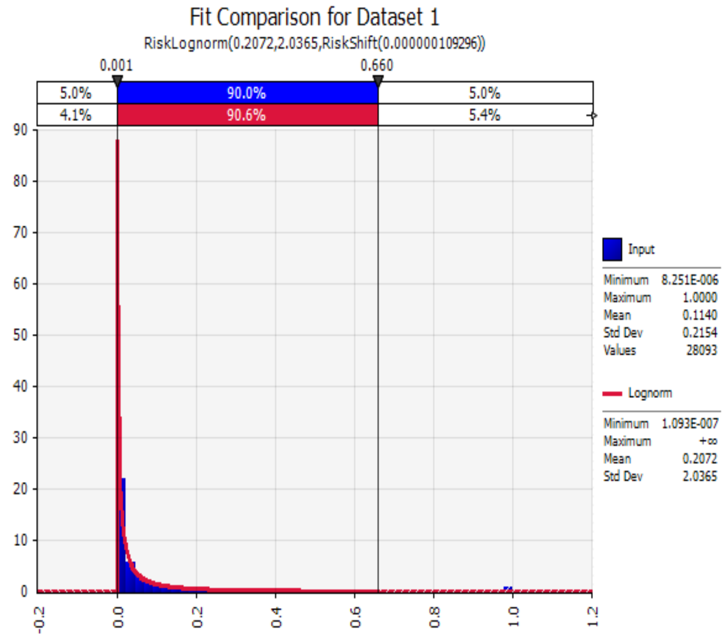
<table>
<tr><td>Diversos</td><td>Datos Total</td><td>33259</td></tr>
<tr><td rowspan="9">S.A.< =12,000,000</td><td colspan="2">fact < .6</td></tr>
<tr><td>Datos</td><td>26490</td></tr>
<tr><td>Dist</td><td>Lognorm</td></tr>
<tr><td colspan="2">0.6 <= fact < = .9</td></tr>
<tr><td>Datos</td><td>846</td></tr>
<tr><td>Dist</td><td>Beta general</td></tr>
<tr><td colspan="2">fact > .9</td></tr>
<tr><td>Datos</td><td>759</td></tr>
<tr><td>Dist</td><td>Weibull</td></tr>
<tr><td rowspan="7">S.A.> 12,000,000</td><td colspan="2">fact < = .01</td></tr>
<tr><td>Datos</td><td>5142</td></tr>
<tr><td>Dist</td><td>exponencial</td></tr>
<tr><td colspan="2">fact > .01</td></tr>
<tr><td>Dist</td><td>Uniforme</td></tr>
<tr><td>19</td><td>U(.01,0.05)</td></tr>
<tr><td>3</td><td>U(.05,1)</td></tr>
</table>

RANGOS		MAXIMO	MINIMO	PROMEDIO	#SINIESTROS
Valores menores a	12,000,001	100%	0%	11%	28095
12,000,001	32,000,001	9%	0%	0%	1206
32,000,001	52,000,001	18%	0%	0%	1168
52,000,001	72,000,001	1%	0%	0%	172
72,000,001	92,000,001	0%	0%	0%	151
92,000,001	112,000,001	1%	0%	0%	169
112,000,001	132,000,001	0%	0%	0%	176
132,000,001	152,000,001	0%	0%	0%	562
152,000,001	172,000,001	1%	0%	0%	48
172,000,001	192,000,001	1%	0%	0%	63
192,000,001	212,000,001	0%	0%	0%	226
212,000,001	232,000,001	0%	0%	0%	56
232,000,001	252,000,001	0%	0%	0%	107
252,000,001	272,000,001	0%	0%	0%	7
272,000,001	292,000,001	0%	0%	0%	77
292,000,001	312,000,001	0%	0%	0%	244
312,000,001	332,000,001	0%	0%	0%	11
332,000,001	352,000,001	0%	0%	0%	228
372,000,001	392,000,001	0%	0%	0%	4
452,000,001	472,000,001	0%	0%	0%	260
472,000,001	492,000,001	0%	0%	0%	1
Valores mayores a	512,000,001	0%	0%	0%	228
Total general					33259

Fact < .6 79.6% Datos

Fit Ranking

Fit	K-S
Lognorm	0.0355
Pearson5	0.0681
InvGauss	0.0748
ExtValue	0.2480
Normal	0.2983
Logistic	0.3183
Expon	0.3210
Pareto	0.3775
Triang	0.5693
Uniform	0.6556
BetaGeneral	*N/A*
Gamma	*N/A*
LogLogistic	*N/A*
Pearson6	*N/A*
Weibull	*N/A*

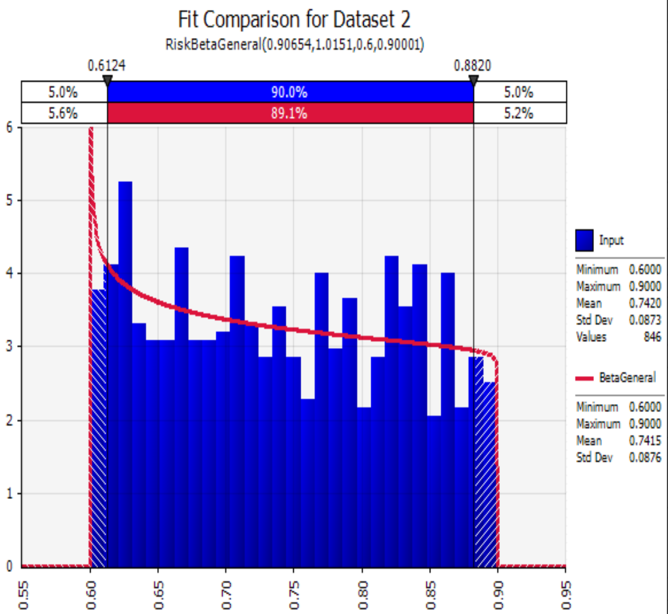


	Input	Lognorm
Distribution Statistics		
Minimum	8.25E-06	1.09E-07
Maximum	1.00	+Infinity
Mean	0.11	0.21
Mode	1.0000 [est]	0.00
Median	0.02	0.02
Std. Deviation	0.22	2.04
Skewness	2.71	978.96
Kurtosis	9.91	#######
Percentiles		
5%	0.00	0.00
10%	0.00	0.00
15%	0.00	0.00
20%	0.00	0.00
25%	0.01	0.00
30%	0.01	0.01
35%	0.01	0.01
40%	0.01	0.01
45%	0.02	0.02
50%	0.02	0.02
55%	0.03	0.03
60%	0.04	0.04
65%	0.05	0.05
70%	0.07	0.06
75%	0.10	0.09
80%	0.15	0.13
85%	0.22	0.19
90%	0.36	0.33
95%	0.66	0.71

0.6 <= fact < = .9 2.5% Datos

Fit Ranking

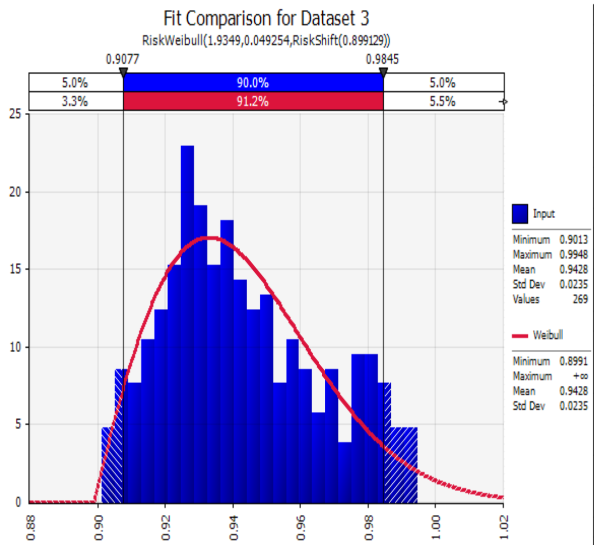
Fit	K-S
BetaGeneral	0.0168
Uniform	0.0461
Normal	0.0720
LogLogistic	0.0720
Logistic	0.0727
Lognorm	0.0748
InvGauss	0.0759
Weibull	0.0763
ExtValue	0.0835
Triang	0.1284
Expon	0.1422
Pareto	0.1634
Gamma	*N/A*
Pearson5	*N/A*
Pearson6	*N/A*



	Input	BetaGeneral	Uniform
Fit			
Minimum	0.60	0.60	0.60
Maximum	0.90	0.90	0.90
Mean	0.74	0.74	0.75
Mode	0.8188 [est]	0.60	0.60
Median	0.74	0.74	0.75
Std. Deviation	0.09	0.09	0.09
Skewness	0.07	0.10	-
Kurtosis	1.77	1.79	1.80
Percentiles			
5%	0.61	0.61	0.61
10%	0.62	0.62	0.63
15%	0.63	0.64	0.64
20%	0.65	0.65	0.66
25%	0.67	0.66	0.67
30%	0.68	0.68	0.69
35%	0.70	0.69	0.70
40%	0.71	0.71	0.72
45%	0.72	0.72	0.74
50%	0.74	0.74	0.75
55%	0.76	0.75	0.77
60%	0.77	0.77	0.78
65%	0.79	0.78	0.80
70%	0.80	0.80	0.81
75%	0.82	0.82	0.83
80%	0.83	0.83	0.84
85%	0.85	0.85	0.86
90%	0.86	0.87	0.87
95%	0.88	0.88	0.89

Fact > .9 2.2% Datos

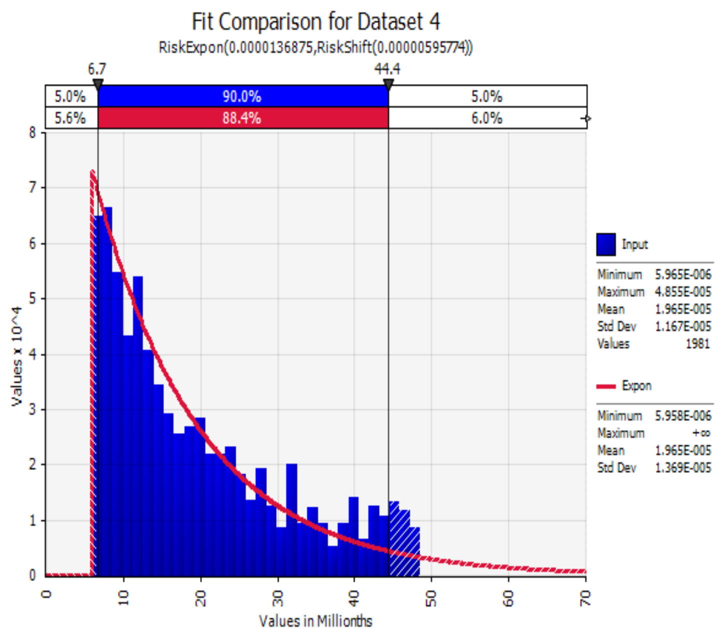
Fit Ranking	
Fit	K-S
Weibull	0.0476
LogLogistic	0.0484
ExtValue	0.0485
Rayleigh	0.0487
Triang	0.0494
InvGauss	0.0514
Lognorm	0.0523
Pearson5	0.0535
BetaGeneral	0.0635
Logistic	0.0685
Normal	0.0786
Uniform	0.1523
Expon	0.2057
Pareto	0.2089
Erlang	*N/A*
Gamma	*N/A*
Pearson6	*N/A*



	Input	Weibull	LogLogistic	ExtValue	Rayleigh
Distribution Statistics					
Minimum	0.90	0.90	0.87	-Infinity	0.90
Maximum	0.99	+Infinity	+Infinity	+Infinity	+Infinity
Mean	0.94	0.94	0.94	0.94	0.94
Mode	0.9208 [est]	0.93	0.93	0.93	0.93
Median	0.94	0.94	0.94	0.94	0.94
Std. Deviation	0.02	0.02	0.03	0.03	0.02
Skewness	0.42	0.68	2.66	1.14	0.63
Kurtosis	2.25	3.33	35.94	5.40	3.25
Percentiles					
5%	0.9077	0.9097	0.9095	0.9100	0.9098
10%	0.9138	0.9145	0.9154	0.9152	0.9147
15%	0.9189	0.9184	0.9196	0.9190	0.9187
20%	0.9214	0.9218	0.9230	0.9223	0.9222
25%	0.9259	0.9250	0.9260	0.9252	0.9254
30%	0.9273	0.9280	0.9287	0.9280	0.9284
35%	0.9297	0.9310	0.9314	0.9307	0.9314
40%	0.9329	0.9339	0.9340	0.9333	0.9343
45%	0.9355	0.9369	0.9366	0.9361	0.9373
50%	0.9387	0.9399	0.9392	0.9388	0.9402
55%	0.9412	0.9430	0.9420	0.9418	0.9433
60%	0.9458	0.9462	0.9449	0.9449	0.9465
65%	0.9495	0.9496	0.9481	0.9482	0.9499
70%	0.9551	0.9533	0.9517	0.9519	0.9535
75%	0.9594	0.9574	0.9558	0.9562	0.9576
80%	0.9657	0.9621	0.9609	0.9612	0.9621
85%	0.9722	0.9677	0.9673	0.9674	0.9676
90%	0.9800	0.9749	0.9767	0.9759	0.9746
95%	0.9845	0.9860	0.9939	0.9901	0.9853

Fact < = .01 15.5% Datos

Fit Ranking	
Fit	K-S
Expon	0.0445
BetaGeneral	0.0551
Lognorm	0.0640
InvGauss	0.0668
LogLogistic	0.0829
ExtValue	0.0928
Triang	0.1155
Normal	0.1263
Logistic	0.1370
Pareto	0.1679
Uniform	0.2812
Gamma	*N/A*
Pearson5	*N/A*
Pearson6	*N/A*
Weibull	*N/A*

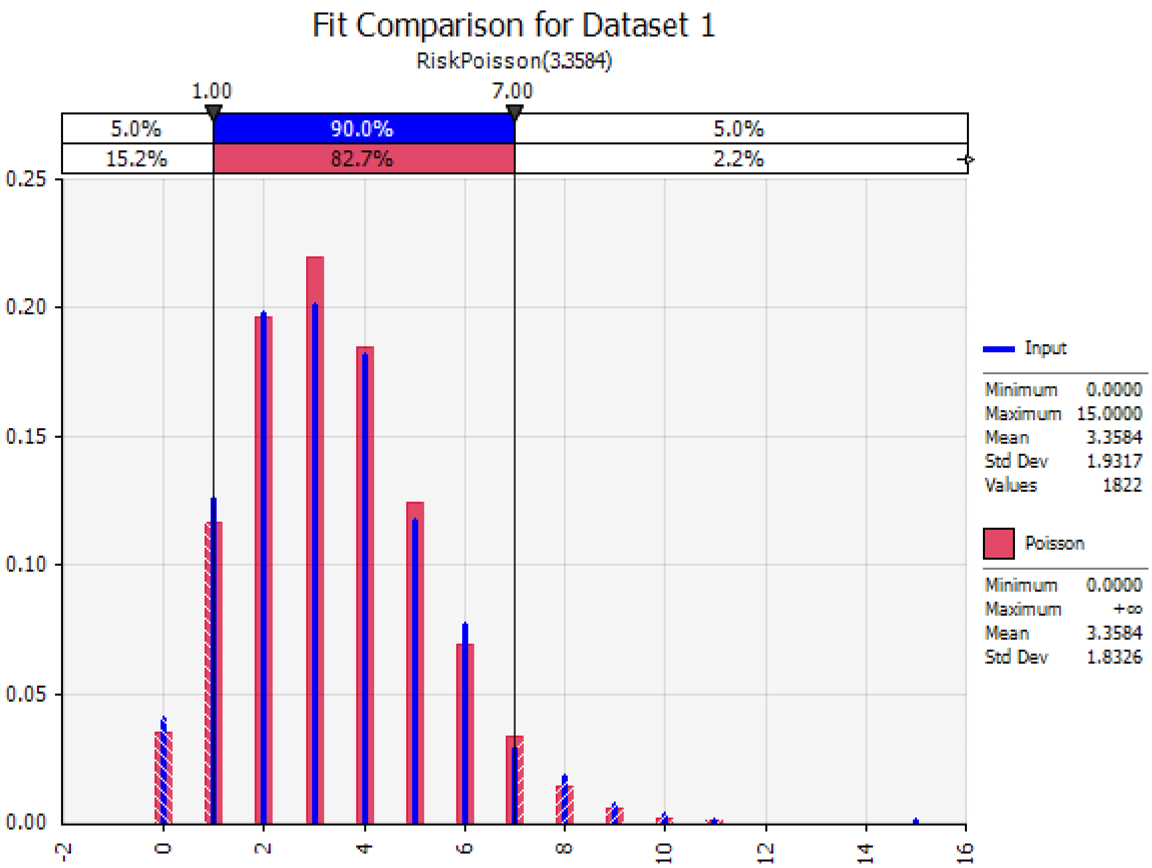


	Input	Expon
Distribution Statistics		
Minimum	5.97E-06	5.96E-06
Maximum	4.86E-05	+Infinity
Mean	1.97E-05	1.97E-05
Mode	210E-05 [est]	5.96E-06
Median	1.61E-05	1.55E-05
Std. Deviation	1.17E-05	1.37E-05
Skewness	0.87	2.00
Kurtosis	2.67	9.00
Percentiles		
5%	6.743E-06	6.660E-06
10%	7.480E-06	7.400E-06
15%	8.128E-06	8.182E-06
20%	9.234E-06	9.012E-06
25%	1.002E-05	9.895E-06
30%	1.115E-05	1.084E-05
35%	1.210E-05	1.185E-05
40%	1.306E-05	1.295E-05
45%	1.441E-05	1.414E-05
50%	1.613E-05	1.545E-05
55%	1.795E-05	1.689E-05
60%	1.987E-05	1.850E-05
65%	2.180E-05	2.033E-05
70%	2.403E-05	2.244E-05
75%	2.689E-05	2.493E-05
80%	3.022E-05	2.799E-05
85%	3.402E-05	3.192E-05
90%	3.933E-05	3.747E-05
95%	4.443E-05	4.696E-05

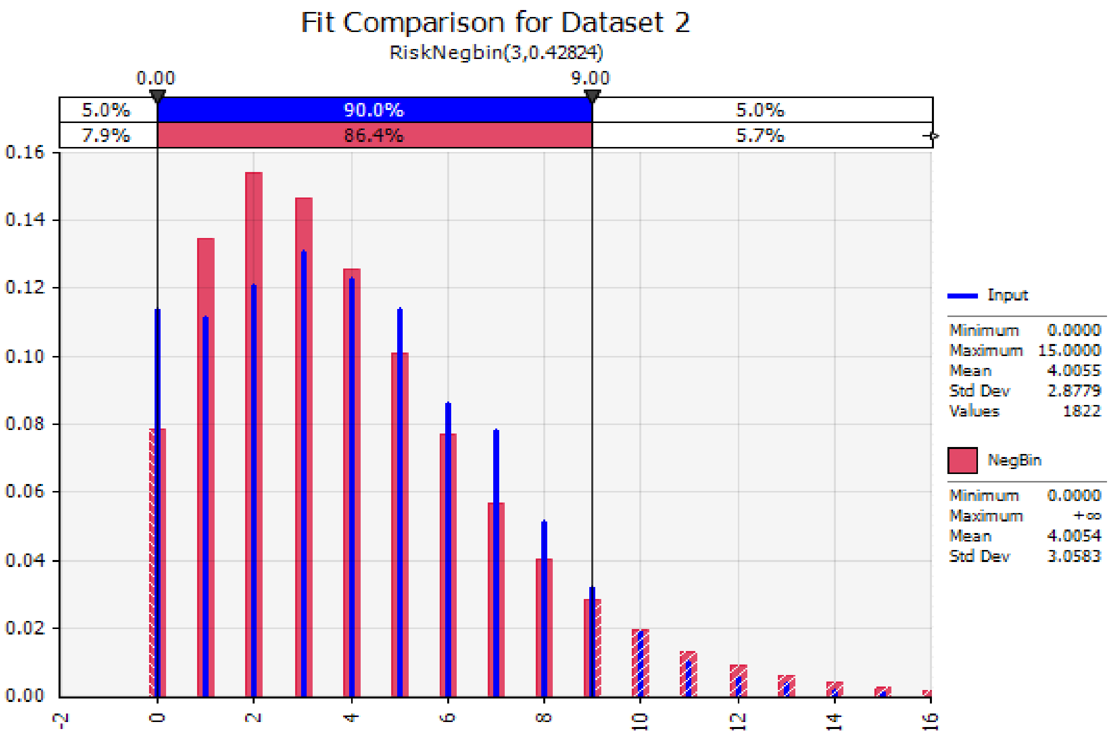
Frecuencia

ramo		Diaria		Distribución
		# siniestros	máximo	
040	RC	6210	15	Poisson
050	Carga	7453	15	BIN.Neg
060	Incendio	4066	21	BIN.Neg
080	Agropecuario	7487	175	Geometrica
110	Diversos	33259	64	BIN.Neg

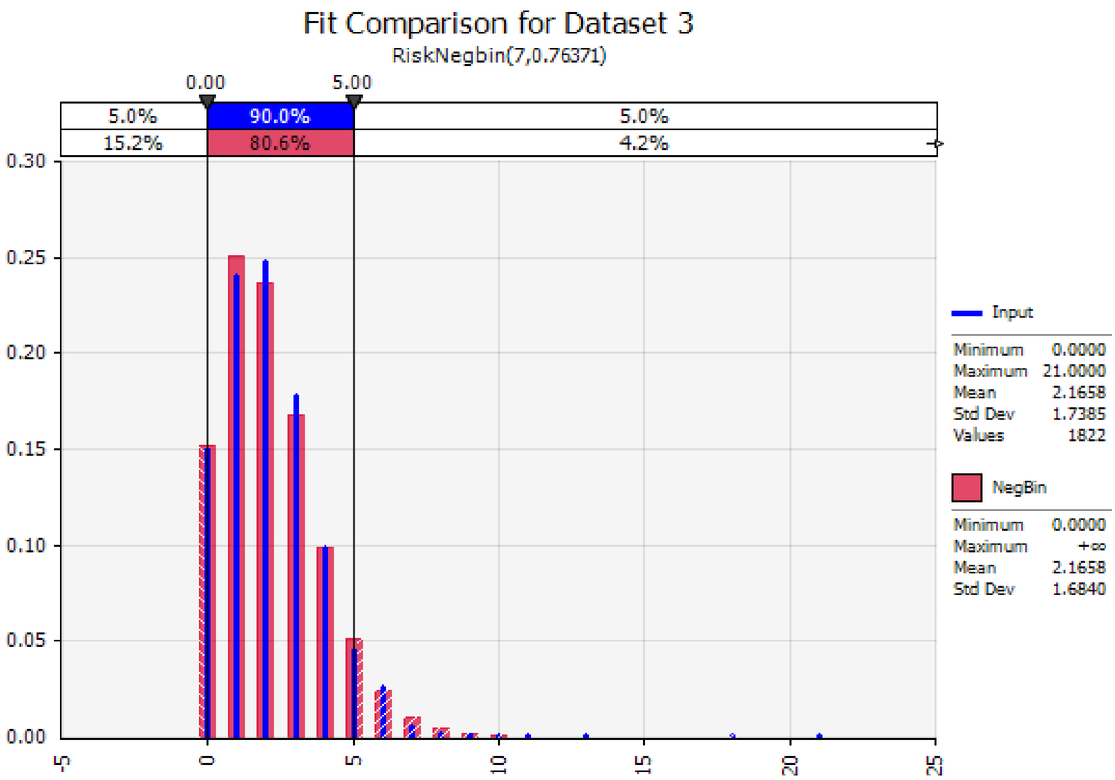
Responsabilidad civil



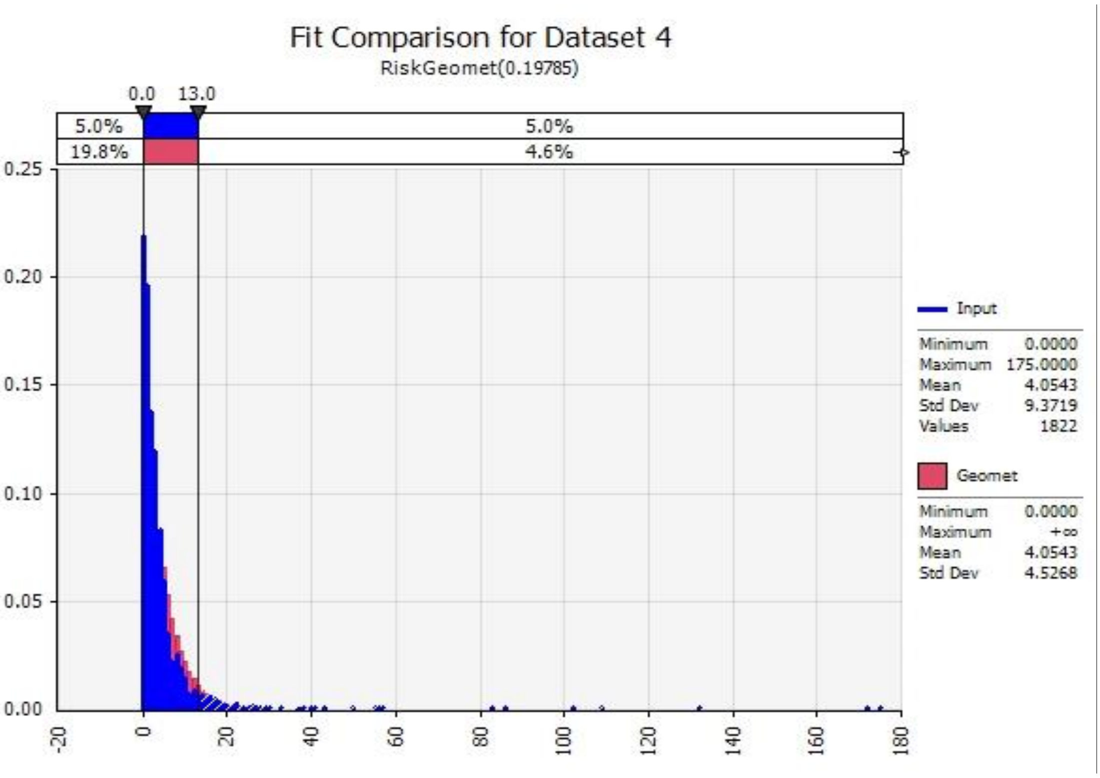
Transporte



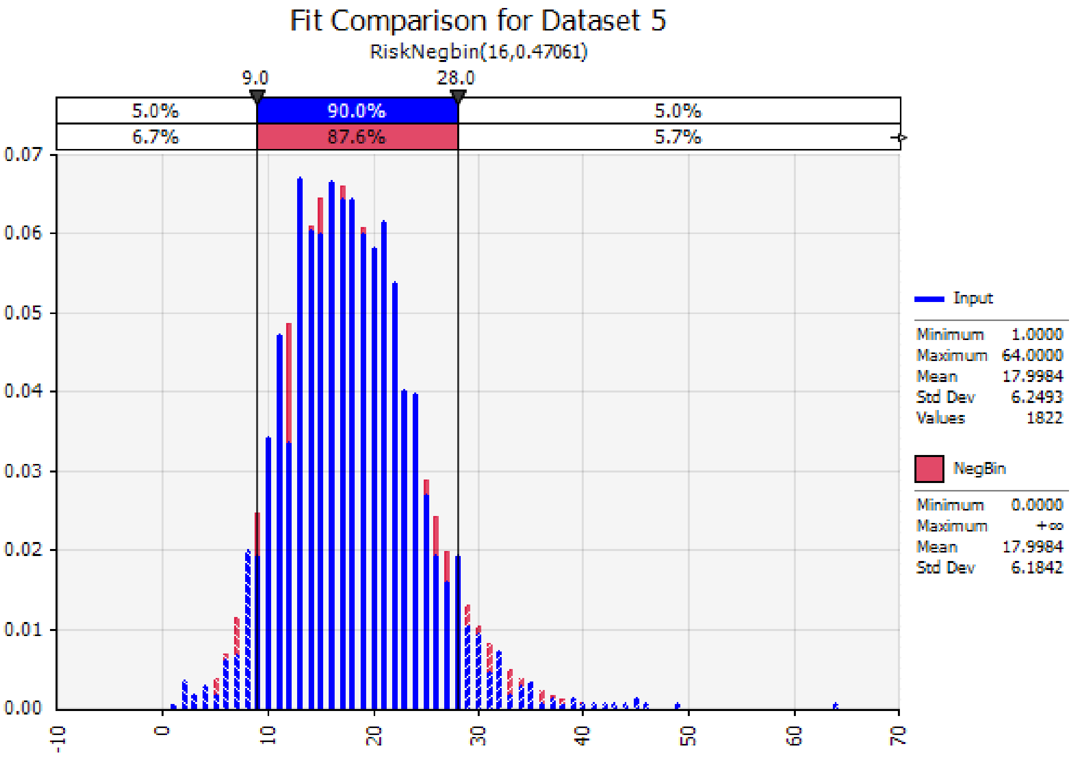
Incendio



Agropecuario

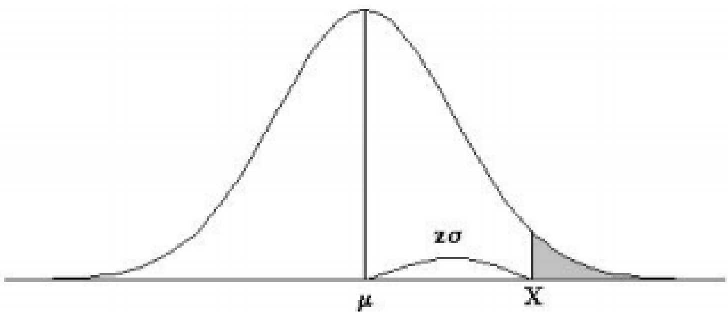


Diversos



II. Tabla de distribución normal estandarizada.

Áreas bajo la curva normal



Ejemplo:

$$Z = \frac{X - \mu}{\sigma}$$

$P\,[Z > 1] = 0.1587$

$P\,[Z > 1.96] = 0.0250$

Desv. normal x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010

REFERENCIAS BIBLIOGRÁFICAS

Aguilar, P. (2011). Límite de Retención. Elementos de análisis técnico. *AMIS*.

ANTECEDENTES HISTÓRICOS DEL SEGURO. (s.f.). *Gran enciclopedia del Mundo.* Ediciones Bilbao.

Armitage, P., & Berry, G. (1997). *Estadística Para La Investigación Biomédica. Edición en español.* Madrid: Harcourt Brace.

Banks, J., Carson, J., & Nelso, B. (1996). *Discrete-Event System Simulation. Second Edition.* New Jersey: Prentice-Hall.

Barrera, J. (1957). *Tratado de Derecho Mercantil vol.I.* México D.F.: Porrúa.

Belinchón, J. A. (2007). *Apuntes de Estadística.* Mexico D.F.: UNAM.

Beltrán, J. I. (2012). *Cálculo del p-valor en pruebas de bondad de ajuste .* México, D.F.: Tesis de Maestría en Ciencias Matemáticas Aplicadas e Industriales, UAM .

Calva, L. E. (2005). *Consideraciones sobre algunos conceptos básicos de probabilidad.* México D.F.

Canavos, G. (1988). *Probabilidad y Estadística. Aplicaciones y métodos.* Estado de México: McGraw-Hill.

Casella, G., & Berger, R. (1990). Statistical inference. *Duxbury Press*.

Chow, V., Maidment, D., & Mays, L. (1994). *Hidrología Aplicada. Colombia.* Bogota: McGraw- Hill Interamericana.

Código Hammurabi. (s.f.). *Gran enciclopedia del Mundo.* Marín S.A. Tomo 10.

Comisíon Nacional de Seguros y Fianzas . (2014). *Circular Única de Seguros y Fianzas.* México D.F.

Comisión Nacional de Seguros y Fianzas. (2014). *Ley de Instituciones de Seguros y Fianzas.* México D.F.

Eco, U. (2000). *Como hacer una tesis.Tecnicas y procediientos de estudio, investigacion y escritura.* Salamanca : Universidad de Salamanca.

Fishman, G. (1978). *Conceptos y Métodos en la simulación digital de eventos discretos.* Mexico D.F.: Limusa.

García, A. (2002). *Teoría del valor extremo: una aplicación al sector asegurador.* Madrid: Fundacion MAPFRE.

González, E. (2007). *Pruebas de bondad de ajuste para distribuciones estables.* Estado de México: Tesis de doctorado. Centro de postgrados.

Hossack, I.B.; Pollard, J.H.; Zehnwirth, B. (2001). *Introducción a la Estadística con aplicaciones a los seguros generales.* Madrid: Fundación Mapfre.

Kelton, W., Sadowski, R., & Sadowski, D. (1998). *Simulation with Arena.* Boston: Mc Graw Hill.

Kroll, C., & Vogel, R. (2002). *Probability Distribution of low Streamflow Series in the United States.* United States: Journal of Hydrology Engineering.

Law, A., & Kelton, W. (1991). *Simulation Modeling & Analysis 2a edition.* New York: McGraw-Hill.

Lillegard, M., & S., E. (1997). Stochastic simulations conditioned on sufficient statistics. *Biometrika*, 235–240.

Linsley, R., Kohler, M., & Paulus, J. (1988). *Hidrología para ingenieros 2 ed.* Mexico D.F.: McGraw-Hill Interamericana S.A.

López-Martínez., L. (2010). *Una Prueba de razón de verosimilitudes para discriminar entre la distribución Poisson, Binomial y Binomial negativa.* Montecillo: Tesis de maestría. Colegio de Postgraduados. Campus Montecillo.

Mapfre, F. (2013). *Introducción al Reaseguro.* Madrid: Fundación Mapfre.

Mood, A., Graybill, F., & Boe, D. (1974). *Introduction to the Theory of Statistics.* Colorado: McGraw-Hill.

Muñoz, M. (2004). *Análisis de algunas variables hidrológicas y su ajuste a funciones de distribución de probabilidad, en tres cuencas de la Región del Maule.* Talca: Tesis de Ing. Forestal. Talca, Chile. Universidad de Talca.

Nania, L. (2003). *Estadística aplicada a la hidrología.* Granada: Universidad de Granada. España.

O'Reilly, F. (1990). Algunas consideraciones sobre la inferencia estadística. *Ciencia*, 29–43.

O'Reilly, F., & Rueda, R. (1999). Tests of fit for discrete distributions based on the proba- bility generating function. *Comm. Statist. Sim. Comp*, 259–274.

O'Reilly, F., & Rueda, R. (1999). Tests of fit for discrete distributions based on the probability generating function. *Comm. Statist. Sim. Comp*, 259–274.

Ogunnaike B.A., H. R. (1994). *Process Dynamics, Modeling and Control.* New York: Oxford.

Osorio, G. (2003). *Manual Básico del Seguro.* Asunción.

Palacios, H. (1996). *Introducción al Cálculo Actuarial.* Madrid: Fudación MAPFRE.

Peña, J. C. (1982). *El Método Estadístico en la Investigacíon Médica.* Madrid: Karpus.

Picos, J. (2003). *Los seguros de Incendios Forestales: Antecedentes y estudio de Viabilidad de su aplicación a Galicia.* vigo: Tesis universidad de vigo.

Picos, J. (2006). *Los seguros contra Incendios Forestales y su aplicación a Galicia.* madrid: fundacion mapfre.

Pizarro, R., & Novoa, P. (1996). *Análisis comparativo de modelos matemáticos precipitación-escorrentía en cuencas de la España peninsular.* Madrid: Tesis Doctoral.Universidad Politécnica de Madrid.

Prieto, E. (2005). *Problemas Estadisticos en Seguros Generales.* Madrid.

Rincón, L. (2007). *Curso Elemental de Probabilidad y Estadística.* Mexico D.F.: UNAM.

Ross, S. M. (2006). *Simulation.* California: University of Southern California.

Ruiz–Maya, L. (1977). *Métodos Estadísticos de investigación (Introduccíon al Análisis de la Varianza).* Madrid: I.N.E. Artes Gráficas.

Sánchez, O. (2000). *La Institucción del Seguro en México.* México D.F.: Porrúa.

Schmelkes, C. (2009). *Manual para la presentación de anteproyectos e informes de investigación 2a edición .* New York: Oxford.

Shannon, R. (1988). *Simulación de Sistemas. Diseño, desarrollo e implementación.* México D.F.: Trillas.

Spiegel, M. (1991). *Probabilidad y Estadística.* Madrid: McGraw-Hill.

Varas, E., & Bois, P. (1998). *Hidrología probabilística.* Santiago: Universidad Católica de Chile.

Villers, S. (2001). *Programa para cáculo de p-values de estadísticas EDF para pruebas de bondad de ajuste.* Mexico D.F: Tesis de maestría.Universidad Autonoma de México.

Wackerly, D., Mendenhall, W., & Scheaffer, R. (2009). *Estadística matemática con aplicaciones 7ª edición.* Cengage Learning.

Ybnias, G. (2009). *Métodos cuantitativos para los negocios.* Huancayo.