



**UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MÉXICO**



FACULTAD DE ECONOMÍA

**“SIMULACIÓN Y ANÁLISIS DE COLA MÁXIMA POSITIVA
PARA LA SINIESTRALIDAD DE UN SEGURO DE GASTOS
MÉDICOS PERSONALES EN MÉXICO, 2009 – 2020”.**

TESIS

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN ACTUARÍA

PRESENTA:

JOSÉ GARCÍA FERREGRINO

ASESOR:

DRA. EN C. ANNEL HURTADO JARAMILLO

REVISORES:

DR. EN E. MIGUEL ÁNGEL DÍAZ CARREÑO

L. EN A.F. JAIME ALBERTO RANGEL BERNAL

TOLUCA, ESTADO DE MÉXICO

NOVIEMBRE 2023

ÍNDICE

ÍNDICE DE DIAGRAMAS.....	iv
ÍNDICE DE GRÁFICAS.....	iv
ÍNDICE DE TABLAS.....	vi
INTRODUCCIÓN	7
CAPÍTULO I. El seguro y el riesgo	10
1.1 Conocimientos generales.....	10
1.2 Evolución del seguro en el mundo	11
1.2.1 Edad Antigua	12
1.2.2 Edad media.....	12
1.2.3 Época moderna	13
1.3 Evolución del seguro en México y su presencia en la actualidad.....	14
1.4 El riesgo	17
1.4.1 Características de los riesgos asegurables.	19
1.5 Tipos de seguros.....	21
1.6 Antecedentes del seguro de gastos médicos.....	23
1.7 Características del seguro de gastos médicos.....	24
1.8 Teoría clásica del riesgo	28
1.8.1 Teoría del riesgo individual.....	29
1.8.2 Teoría del riesgo colectivo.....	30
CAPÍTULO II: Teoría de valores extremos	32
2.1 Distribución Generalizada de Valores Extremos	34
2.2 Distribución de Pareto Generalizada.....	37
2.2.1 Distribuciones estables.....	41

CAPÍTULO III: Cadenas de Markov y los métodos Monte Carlo vía Cadenas de Markov.	46
.....	46
3.1 Cadenas de Markov	46
3.2 Los métodos Monte Carlo vía Cadenas de Markov.....	54
3.2.1 Metropolis Hastings	57
3.2.2 Baseline Metropolis Hastings.....	61
3.2.3 Informative Priors Baseline Metropolis Hastings.....	63
CAPÍTULO IV: Aplicación metodológica.....	65
4.1 Descripción de la base de datos	65
4.2 Métodos para definir umbrales	70
4.3 Estudio de simulación	75
4.4 Aplicación para la siniestralidad extrema de un seguro de Gastos Médicos Personales en México, 2009 – 2020.....	81
CONCLUSIONES GENERALES	88
Referencias	90
ANEXOS	97
Anexo 1.....	98
Anexo 2.....	100

ÍNDICE DE DIAGRAMAS

Diagrama 1. Tipos de seguros.....	22
Diagrama 2. Transición entre estados en una Cadena de Markov.....	47

ÍNDICE DE GRÁFICAS

Gráfica 1. Composición de cartera de seguros y fianzas, 2020.....	16
Gráfica 2. Índice CR-5 del seguro de Accidentes y Enfermedades, 2020.	25
Gráfica 3. Índice de siniestralidad del seguro de gastos médicos personales en México, 2009 - 2020.	27
Gráfica 4. Densidad de la distribución Fréchet.	35
Gráfica 5. Densidad de la distribución Gumbel.....	36
Gráfica 6. Densidad de la distribución Weibull.	37
Gráfica 7. Densidad de la GPD.	40
Gráfica 8. Monto de siniestros del seguro de gastos médicos personales en México, 2009 – 2020.	66
Gráfica 9. Cuantiles de la distribución normal para los siniestros de un seguro de gastos médicos personales en México, 2009-2020.	67
Gráfica 10. Excesos Medios para la siniestralidad del seguro de Gastos Médicos Personales en México, 2009 – 2020.	73
Gráfica 11. Excesos medios para los últimos 1000 siniestros registrados en el seguro de gastos médicos personales en México, 2009 – 2020.....	74
Gráfica 12. Distribución a posteriori de γz para diferentes valores de n , cuando la distribución base es $L(0,1)$ izquierda, $C(0,1)$ en medio y $Normal(0,1)$ derecha, por MH y BMH.	76

Gráfica 13. Distribución a posteriori de σz para diferentes valores de n , cuando la distribución base es L (0,1) izquierda, C (0,1) en medio y Normal (0,1) derecha, por MH (gráficas superiores) y BMH (gráficas inferiores). 77

Gráfica 14. Distribución a posteriori de σz para diferentes valores de n , cuando la distribución base es L (0,1) izquierda, C (0,1) en medio y Normal (0,1) derecha, por MH (gráficas superiores) e IPBMH (gráficas inferiores). 78

Gráfica 15. Autocorrelación de la cadena de Markov, cuando la distribución base es C (0,1), MH. 79

Gráfica 16. Distribución a posteriori de la GPD para diferentes valores de n , para el seguro de gastos médicos personales en México, 2009 - 2020, MH. 82

ÍNDICE DE TABLAS

Tabla 1. Porcentaje de asegurados en gastos médicos con respecto a la población total (2012 - 2020).....	26
Tabla 2: Valores para c_1 , c_2 , c_3 para las tres distribuciones subyacentes.....	64
Tabla 3. Estadística Descriptiva de los datos.....	66
Tabla 4. Número de observaciones que superan el umbral.....	74
Tabla 5. MAE y RMSE de MH, BMH e IPBMH con $a = 1$, cuando la distribución base es L (0,1).....	80
Tabla 6. MAE y RMSE de MH, BMH e IPBMH con $a = 1$, cuando la distribución base es C (0,1).....	80
Tabla 7. MAE y RMSE para MH, BMH e IPBMH con $a = 1$, cuando la distribución base es N (0,1).	81
Tabla 8. MAE y RMSE para la GPD mediante ML y MH.....	83
Tabla 9. Intervalo de credibilidad para el parámetro γ , de la GPD cuando se tiene diferente valor de n	85
Tabla 10. Intervalo de credibilidad para el parámetro σ , de la GPD cuando se tiene diferente valor de n	85

INTRODUCCIÓN

La crisis financiera mexicana conocida como “el efecto tequila” o “el error de diciembre” (1994 – 1995) fue un punto de inflexión en la economía mexicana en donde el sistema financiero se vio en la necesidad de iniciar un proceso de revisión y modernización para regular las instituciones financieras, esto con el fin de dar estabilidad y transparencia al mercado. Las empresas aseguradoras no han sido ajenas al impacto económico de esta crisis, el cual ha sido determinado principalmente por medio de reducción de emisión de pólizas, impactos en la operación e infraestructura, impacto en la siniestralidad (aumento o disminución dependiendo el ramo), diferimiento en pago de primas, entre otros, por lo que, el sector asegurador se ve directamente afectado en caso de crisis financieras y económicas.

En este sentido, la Comisión Nacional de Seguros y Fianzas (CNSF) (órgano desconcentrado de la Secretaría de Hacienda y Crédito Público) es la entidad encargada de supervisar que la operación del sector asegurador y afianzador y, demás personas supervisadas, se apeguen al marco normativo en México, preservando la solvencia y estabilidad financiera de las entidades aseguradoras para garantizar los intereses del público usuario, así como promover el sano desarrollo en el sector para extender la cobertura a la mayor parte de la población.

La clasificación general de los tipos de seguros en México, de acuerdo con la Ley de Instituciones de Seguros y Fianzas, son: seguros de vida, seguros de accidentes y enfermedades y seguros de daños. En específico, el seguro de accidentes y enfermedades se encuentra integrado por tres ramos: accidentes personales, gastos médicos y salud. En este proyecto, nos enfocamos en el seguro de gastos médicos personales, el cual tiene como objetivo hacer frente a la posibilidad de quebranto económico al atender problemas de salud, principalmente aquellas que requieran atención médica – quirúrgica en un hospital derivados de un accidente o una enfermedad. Asimismo, es importante mencionar que de acuerdo con la CNSF (2020), el seguro de accidentes y enfermedades representó el seguro con mayor crecimiento real anual con un 7.4%, impulsado principalmente por el ramo de gastos médicos personales.

Por consiguiente, uno de los aspectos más controvertidos en instituciones aseguradoras es definir el costo de las primas acorde a los riesgos asumidos en una cartera de asegurados, así como establecer metodologías eficientes en la constitución de reservas técnicas para hacer frente a sus obligaciones futuras. Esto es posible siempre y cuando se tenga una base sólida de lo que se espera pagar por los siniestros ocurridos en un periodo de tiempo. En este contexto, la hipótesis principal de la investigación es que la simulación Monte Carlo vía Cadenas de Markov (MCMC) mejora la estimación de parámetros en una distribución de valores extremos comparado con los métodos clásicos probabilísticos. Asimismo, el interés del presente proyecto radica en definir la mejor metodología para el análisis de siniestralidad de un seguro de gastos médicos personales en México para el periodo 2009 – 2020, que permita simular la distribución de los montos máximos y poder inferir sobre los resultados, debido a que la realización de un siniestro de magnitud catastrófica puede poner en riesgo la solvencia financiera de una institución aseguradora si no se toman las políticas adecuadas de cesión y retención.

Cabe mencionar que, en la actualidad no se encuentran antecedentes acerca de “metodologías de simulación para los siniestros máximos de un seguro de gastos médicos personales”, y los estudios realizados en el extranjero son limitados enfocándose principalmente en temas medioambientales (ver, por ejemplo, Martín et al., 2022). Por esta situación, es conveniente afianzar un mayor conocimiento en la representación del área de colas pesadas, a través de metodologías MCMC. Asimismo, se hace mención que el presente trabajo es extrapolable a cualquier cartera de asegurados que cuenten con características similares.

La importancia de medir la siniestralidad de eventos extremos, mediante simulación MCMC, se basa principalmente en obtener un estudio que ayude al investigador a describir el comportamiento de cola máxima, a través de distribuciones *a priori* altamente informativas, mediante la relación de parámetros de la distribución de valores máximos y las distribuciones base como son: Lévy, Cauchy y Normal. También se menciona que no existen otras distribuciones que sean consideradas altamente informativas, por lo que sería interesante, la realización de estos estudios en un futuro.

Dicho lo anterior, la estructura del trabajo se conforma de cuatro capítulos: en el primero de ellos, se presentan los conceptos básicos del seguro, los antecedentes de la función aseguradora en México y en el mundo, así como la introducción a la teoría clásica del riesgo.

En el capítulo 2, se aborda la teoría de valores extremos en sus dos vertientes, picos sobre un umbral y el método de bloques; además, se hace referencia a algunos proyectos realizados en materia de riesgos con esta teoría.

En el capítulo 3, encontramos las definiciones de Cadena de Markov, transición entre estados en uno y n pasos, distribución en equilibrio, introducción a la teoría bayesiana y el planteamiento de las metodologías MCMC.

Posteriormente, en el capítulo 4, se desarrollan las metodologías propuestas en el capítulo 3, partiendo de un análisis *a priori* de los datos; asimismo, se presentan las metodologías estadísticas para definir umbrales y un análisis de errores de las metodologías propuestas.

Finalmente se dan las conclusiones obtenidas en este trabajo.

CAPÍTULO I. El seguro y el riesgo

1.1 Conocimientos generales

El campo de los seguros tiene su propia nomenclatura y forma de expresarse; a través de los años diferentes ciencias se han sumado al entendimiento de esta rama, aportando el conocimiento para la toma de decisiones en entidades aseguradoras, instituciones supervisoras y el sector en general. Aunado a esto, para llevar a cabo el presente proyecto, es necesario abordar la teoría y los conceptos básicos que sustentan los procedimientos técnicos que se realizan en la presente investigación.

De acuerdo con las definiciones que se le han dado al contrato de seguro, lo definimos como el instrumento financiero que permite hacer frente a las posibilidades de quebranto económico, ya sea a nivel individual o de negocios; mediante el pago de una prima, una entidad aseguradora se compromete a indemnizar el daño producido en caso de que el objeto o persona asegurable experimente un siniestro.

Otra definición es la de riesgo, la cual se define como la ocurrencia al azar de un acontecimiento capaz de producir daños de tipo material o personal. Algunos de estos son: el riesgo por desastres naturales, el cuál es la pérdida económica por los daños producidos y por las muertes ocasionadas; el riesgo por el fallecimiento de una persona, el cual es que sus descendientes y seres queridos, aún no cuenten con independencia económica; el riesgo de un accidente o una enfermedad, el cuál es la cantidad monetaria que se ha de pagar en la recuperación de la salud, entre otros. Es importante mencionar que en los próximos párrafos se mencionan todos los riesgos a los que se encuentran expuestas las personas y los bienes materiales.

En el mismo sentido, definimos el termino prima como la cantidad monetaria que ha de satisfacer el asegurado a la institución de seguros por la cobertura de los riesgos pactados en el contrato de seguros, siendo estos la cobertura básica y la cobertura adicional. La cobertura básica se encuentra presente en todos los seguros, ya que representa la razón de ser del contrato, y las coberturas adicionales, las cuales son complementarias a la cobertura básica y, brindan mayor protección al asegurado.

Continuando con el análisis, también se menciona el concepto de siniestro, el cuál es la realización o materialización del riesgo amparado en el contrato de seguros, motivando la aparición del principio indemnizatorio, obligando a la entidad aseguradora a satisfacer total o parcialmente al asegurado o a los beneficiarios, el capital garantizado dentro de los límites y exclusiones pactadas.

Finalmente, definimos el termino simulación como el proceso de imitar un sistema real o una situación mediante largos procesos computacionales y numéricos con la finalidad de comprender el sistema o evaluar nuevas estrategias en el aprendizaje automático. Por consiguiente, se presenta la historia del seguro en el mundo y en México, que es de ayuda para entender como fungían las mutualidades de riesgo y los préstamos en la antigüedad, para pasar a seguros con primas como los conocemos hoy en día.

1.2 Evolución del seguro en el mundo

Las primeras formas de seguros surgieron principalmente en los países de India, Egipto, Grecia y Roma a mediados de los 3200 años a.C. y, aunque el seguro no fuese conocido como una institución definida con caracteres propios, se llevaba a cabo estrategias de prevención de riesgos con la finalidad de favorecer el comercio marítimo entre estos países (Consorti, 2005).

La técnica estándar consistía en que los propietarios de barcos, frecuentemente tomarían prestados fondos para adquirir mercancías y financiar sus viajes, estipulando que, al término del viaje, restituirían el total más los intereses, mientras que, si ocurría algún imprevisto, como el robo de mercancías o la pérdida del barco, no habría obligación de devolver la cantidad recibida, denominándose a esta práctica, “Préstamos a la Gruesa” o “Nauticum Foenus”.

Loredana (2005), por su parte, menciona que la evolución del seguro en el mundo está marcada por tres grandes etapas: La Edad Antigua, la Edad Media y la Época Moderna, por lo que a continuación se mencionan los acontecimientos más importantes de cada pauta histórica.

1.2.1 Edad Antigua

La literatura muestra antecedentes de ciertas formas de seguros en civilizaciones tan antiguas como la fenicia; civilización que con el desarrollo que obtuvo en materia naval, se convirtió en uno de los comercios más importantes del Mediterráneo. De esta manera, los “Prestamos a la Gruesa” o “Nauticum Foenus”, mencionados anteriormente, se llevaban a cabo por los mercaderes babilonios, entre los 4000 y 3000 años a.C., en donde asumían el riesgo por pérdidas de caravanas que cruzaban la antigua Babilonia (Loredana, 2005).

Con el paso de tiempo, esta práctica se formalizó a través de los primeros textos legales históricos creados; es el caso del “Código Hammurabi” 1728 a. C., nacido en la antigua Mesopotamia por el rey de Babilonia Hammurabi, por medio del cual, si los navieros perdían sus mercancías durante el trayecto, el gobierno local estaba obligado a devolverles lo que hubieran perdido, similar a los préstamos a la gruesa, pero en este caso eran reconocidos legalmente y con intervención del gobierno. También, dentro de este código, se encuentran antecedentes de ciertas formas de seguros sobre el pago de deudas agrícolas y ganaderas, en donde, si un agricultor o un ganadero tenía alguna deuda relacionada a su cosecha o a sus animales y, si estas eran afectadas por algún desastre natural, se le perdonaba el pago a esa temporada (González, 2003).

1.2.2 Edad media

Partiendo de la decadencia y caída del Imperio romano, hasta el siglo XVII, el seguro con dificultad creció (Gran Enciclopedia del Mundo, 1978). Dicho con otras palabras, a partir de la crisis del siglo III, que es conocida también como “anarquía militar”, y hace referencia a un periodo histórico del Imperio romano con duración de aproximadamente 50 años; el progreso de la actividad aseguradora se vio limitada por la paralización del comercio, el crecimiento de la inflación, el declive de los ingresos económicos, la pérdida de cultivos y la superpoblación; sin embargo, al entrar a esta época nos encontramos con el nacimiento de sociedades mutuales o también las denominadas “Guildas”¹.

¹ Aquellas instituciones formadas por miembros de la misma profesión con el fin de brindarse asistencia mutua, en donde, con el pago de una cuota, los miembros de la asociación estaban protegidos por las

Posteriormente a mediados de la época media y la época moderna, el préstamo a la gruesa evoluciona con el decreto del papa Gregorio IX en el año 1230, quien prohibió los intereses del seguro marítimo y estableció el concepto de prima, de tal forma, una persona asumía el riesgo a cambio del pago incondicionado del siniestro (Varela, 1997).

En este sentido, durante la primera mitad del siglo XIV, los contratos se llevaban a cabo de forma verbal, luego pasan a ser de forma escrita con intervención de notarios y corredores, por lo que el primer contrato de seguro parece surgir en el año 1347 en Génova Italia, haciendo referencia a un seguro marítimo, donde se aseguró un buque llamado “Santa Clara”, cuya ruta cubrió los puertos de Génova en Italia y de Mallorca en España y, de esta manera dio origen la póliza de seguros, la cual más tarde se crearon las norma legislativas.

1.2.3 Época moderna

La explosión del comercio en el Renacimiento y el Mercantilismo en esta época, impulsó la creación de numerosas leyes y tratados en la protección de las propiedades y la vida humana. “En el año 1549, Carlos V dicta la primera ley que regula con carácter obligatorio el contrato de Seguro Marítimo y, de ser normalmente representada por un individuo, pasa a tener forma de Entidades Pluripersonales y Sociedades Anónimas” (Loredana, 2005, p. 13).

Dicho con otras palabras, a finales del siglo XVII el contrato de seguros se profesionalizó en material legal hasta el punto de la inauguración de entidades aseguradoras, tal es el caso de “Fire Office”, la cual se constituyó por el médico Nicholas Barbon, quien fue encomendado en la reconstrucción de las viviendas después del gran incendio de Londres en el año 1666², y de esta manera, se creó el primer seguro contra incendios; primero destinado a cubrir riesgos terrestres.

pérdidas en caso de muerte, enfermedad, accidentes, inundaciones, entre otros riesgos a los que estaban expuestos.

² Existe diferentes teorías de lo sucedido: John Farynor, un panadero de alta reputación del Rey II, el cual no apagó bien los hornos del pan, hasta que el incendio comenzó a extenderse sin límites y arrasó con más de la mitad de la ciudad (Castillo, 2014). Por otro lado, Johnson (2020) menciona que después de la

Posterior al “Gran Incendio”, un acontecimiento que marcó un antes y un después en la historia del seguro, fue el auge que existió en las “cafeterías”, puntos de reunión entre empresarios y políticos, principalmente con alguna relación al riesgo de mar, por lo cual, Edward Lloyds, propietario de una cafetería llamada Lloyd’s, tras encabezar varias reuniones, para el año 1698 publicó la hoja titulada “Lloyd’s News”, que contenía información relevante sobre las embarcaciones, los viajes recientes y los mercados más grandes del mundo. De este modo, la pequeña cafetería Lloyd’s se convirtió en la bolsa de seguros más grande del mundo.

De acuerdo a lo anterior mencionado, a partir del siglo XVIII (año 1701), ya podemos referirnos a seguros de la misma forma en la que hoy los conocemos. Blanco & Carvajal (2001), mencionan que ha sido necesario una evolución de más de seis siglos en el sector, permitiendo pasar de formas financieramente próximas a la amortización de préstamos, para llegar a seguros con primas, en donde el cobro no se realiza como devolución a un préstamo, sino como el pago a un riesgo asumido.

1.3 Evolución del seguro en México y su presencia en la actualidad

La actividad aseguradora en México comenzó de la misma forma que en los países europeos, derivado la actividad mercantil marítima, la cual surgió en los puertos de Veracruz y Acapulco, donde los españoles mantenían comunicación con Europa durante la época de la Colonización, a mediados del siglo XVIII (Castillo, 2014).

En este sentido, durante la etapa del dominio español, México adoptó un proceso de costumbres políticas y sociales que sustituían la forma tradicional de convivencia y, al realizar México su independencia en 1821, se conservó las legislaciones que había en el Virreinato, de modo que cualquier objeto de materia mercantil estaba regulado por el Código de las Ordenanzas de Sevilla, incluyendo la actividad aseguradora, hasta el año 1854. Posterior a esto, el pueblo mexicano se ocupó de reconstruir un nuevo sistema

catástrofe, un relojero francés llamado Lucky Hubert confesó haber iniciado el fuego deliberadamente, razón por la cual, fue condenado enseguida.

político, dando origen al Código de Comercio y, a partir del siglo XIX el seguro fue legislado por las normas: 1854, 1884 y 1889 y por el Código Civil: 1870 y 1884 (Varela, 1997).

Las implementaciones anteriores permitieron la creación de la primera compañía de seguros, situada en el puerto de Veracruz, la cual se denominó “Compañía de seguros marítimos de Nueva España”, y tenía como propósito principal cubrir los riesgos de lo que se denominaba “La Carrera de Las Indias”. Posteriormente, se promulgaron las primeras leyes del seguro mexicano, durante la gobernanza del General Porfirio Díaz, en el siglo XIX y XX; reglamentación por medio de la cual se empieza a dar lineamientos legales a las instituciones de seguros locales y extranjeras que operaban en el país³.

Es a partir de lo anterior, que la evolución del seguro en el mundo y en México, trasciende a través de los años, principalmente con el desarrollo del comercio marítimo, donde la necesidad de las personas de buscar reparo ante los riesgos que las puedan afectar, fue lo que generó el mecanismo de compensación, denominado seguro.

En la actualidad, la incertidumbre se ha convertido en un elemento de la vida cotidiana, donde las entidades aseguradoras constantemente diseñan, modifican y facilitan sus productos, para que las personas o empresas interesadas en adquirir un seguro, tengan la posibilidad de escoger el que mejor se ajuste a sus necesidades. De acuerdo con Focarelli (2017), Chamberlain, Coetze & Camargo (2017), citados en Sabater (2017), un mercado asegurador desarrollado contribuye significativamente al crecimiento económico de cada nación, de tal manera, se puede afirmar que el sector asegurador: 1) Reduce la incertidumbre y volatilidad, 2) Promueve estabilidad financiera a través de inversiones a largo plazo y, 3) Fomenta la buena práctica del ahorro.

En este sentido, la Comisión Nacional de Seguros y Fianzas (CNSF) (órgano desconcentrado de la Secretaría de Hacienda y Crédito Público) es la entidad encargada de supervisar que la operación del sector asegurador y afianzador y, demás personas supervisadas, se apeguen al marco normativo en México, preservando la solvencia y estabilidad financiera de las entidades aseguradoras para garantizar los intereses del

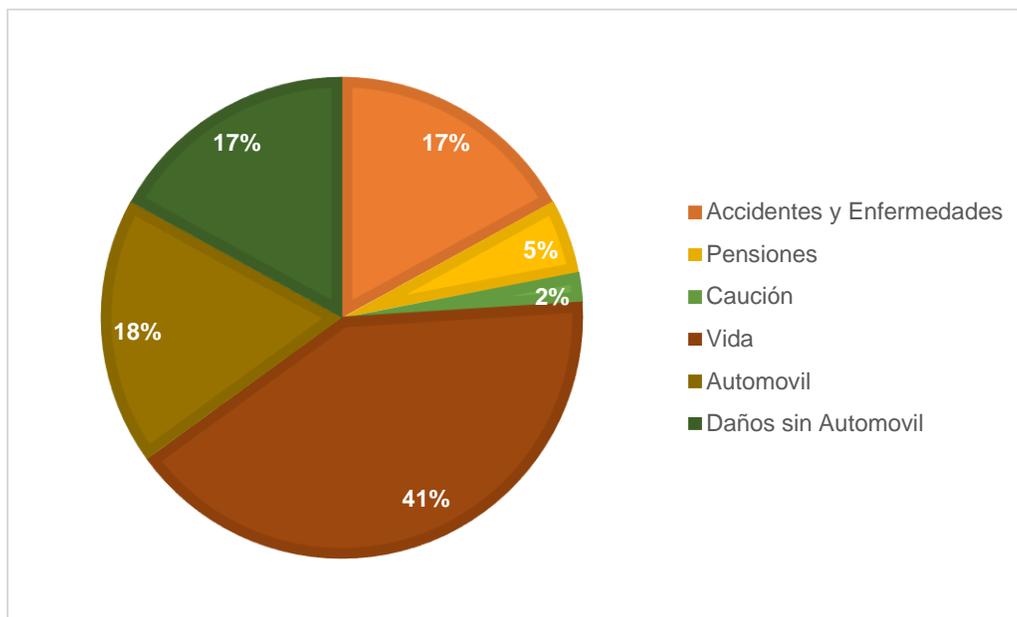
³ Esta Ley fue conocida como Ley del Timbre y tenía la facultad de generar un pequeño impuesto por el uso de papel certificado en cualquier acto comercial o documental.

publico usuario y así, promover el sano desarrollo en el sector para extender la cobertura a la mayor parte de la población.

De acuerdo al Boletín de Análisis Sectorial CNSF (2020), al 31 de diciembre de 2020 el mercado de seguros y fianzas estaba conformado por 112 instituciones, siendo 102 instituciones de seguros y 10 instituciones de fianzas. Del total de instituciones, 56 presentaron capital mayoritariamente extranjero con autorización para operar como filiales de instituciones financieras del exterior, entre ellas, 7 estaban incorporadas a algún grupo financiero. Asimismo, las 56 instituciones restantes, presentaron capital mayoritariamente nacional, de las cuales, 10 pertenecían a algún grupo financiero. El mismo autor menciona que 13 compañías tenían autorización para operar de forma exclusiva el seguro de vida; 63 operaban de forma exclusiva el seguro de no vida, y 36 podían operar de forma combinada seguros de vida y no vida.

Por lo anterior, la composición de cartera al cierre del año 2020, se encuentra conformada por: Vida, 41.3%; Automóviles, 17.8%; Daños sin Automóvil, 17%; Accidentes y Enfermedades, 17.2%, y Pensiones, 4.9% (ver gráfica 1).

Gráfica 1. Composición de cartera de seguros y fianzas, 2020.



Fuente: Elaboración propia con datos de la CNSF (2020).

De acuerdo con la gráfica 1, el mercado de seguros y fianzas es relativamente competitivo y, el seguro de Accidentes y Enfermedades es uno de los seguros con mayor representatividad del sector, ubicándose en la tercera posición; además, como se mencionó anteriormente, este seguro presentó el mayor crecimiento real anual para el sector con un 7.4%, impulsado principalmente por el ramo de gastos médicos. Antes de examinar las metodologías a utilizar en el presente proyecto, es necesario hacer mención a las características principales del riesgo, los diferentes tipos de seguros que se ofrecen en México y el seguro de gastos médicos personales, como se muestra a continuación.

1.4 El riesgo

Se considera parte importante en la investigación analizar las características principales del riesgo, ya que esto ayuda a determinar, a las instituciones involucradas y al sector asegurador en general, si un tipo de riesgo es asegurable o no. De acuerdo a la definición de riesgo, planteada anteriormente, Jones (1998) lo clasifica en dos tipos: el riesgo especulativo y el riesgo puro. El primero, se relaciona con tres tipos de resultados: pérdida, ganancia o ningún cambio. Como ejemplo de esto, tenemos el caso de un inversionista que adquiere acciones en la bolsa de valores, el cual está especulando que el valor de la acción aumentará y obtendrá ganancias, mientras que, también se sabe que el precio de la acción puede caer lo cual representaría pérdidas de forma parcial o en su totalidad. En el mismo sentido, se sabe que el valor de la acción puede mantener el mismo precio; por lo que, en esta situación, no perdería capital ni tampoco obtendría ganancias. En consecuencia, un riesgo especulativo es aquel en donde una inversión puede generar una ganancia o una pérdida.

Por otra parte, el riesgo puro no tiene que ver con la posibilidad de ganancia, sino que se refiere a los acontecimientos que están garantizados para producir algún tipo de pérdida. Como ejemplo, tenemos el caso de una persona que se incapacita después de un accidente, pues debido a este acontecimiento, no podrá seguir generando ingresos económicos ni cubrir los gastos de medicamentos. Esta posibilidad de pérdida financiera sin la opción de ganancia (riesgo puro), es el único tipo de riesgo que se puede asegurar.

En este contexto, existe la administración del riesgo por parte de las instituciones de seguros, que consiste en una serie de metodologías o pasos para identificar y evaluar los riesgos a los que se enfrentan las personas y empresas, las cuales se mencionan a continuación.

- **Evitar el riesgo**

El primer método es evidente en la administración de riesgos, el cual es sencillamente evitarlo por completo. Como ejemplo, se puede evitar la pérdida financiera en el mercado de valores si un inversionista decide no invertir en él. Sin embargo, algunas veces, evitar el riesgo no es eficaz ni práctico.

- **Controlar el riesgo**

Se puede controlar el riesgo llevando a cabo algunos procedimientos para prevenir pérdidas económicas producidas por algún incidente o la ocurrencia de algún evento contingente. En este caso podemos mencionar, por ejemplo, para evitar un incendio en una tienda, se puede prohibir el fumar dentro de las instalaciones o procurar no guardar papeles y cajas innecesarias. Adicionalmente, se pueden instalar sistemas de detectores de humo, disminuyendo así el posible daño ocasionado. Con estas medidas se podría controlar el riesgo, minimizando la posibilidad de pérdida y disminuyendo su severidad.

- **Aceptar el riesgo**

Un tercer método es aceptar o retener el riesgo, lo cual implica asumir toda la responsabilidad financiera del mismo. Algunas personas o negocios, deciden aceptar la responsabilidad total por un riesgo financiero determinado, en lugar de adquirir un seguro para cubrirlo. En esta situación se dice que la persona o el negocio se autoasegura contra el evento contingente. Un autoseguro es, por lo tanto, una técnica de administración de riesgos por medio de la cual una persona o negocio acepta la responsabilidad financiera por las pérdidas asociadas con riesgos específicos.

- **Transferir el riesgo**

Transferir el riesgo es el cuarto método de administración de riesgos y consiste en trasladar la responsabilidad financiera de ese riesgo a la otra parte, generalmente como un intercambio por un pago. La forma habitual de transferencia de riesgo, es a través del contrato de seguro. Cuando una compañía aseguradora acepta otorgar una cobertura, la institución emite una póliza de seguro, el cual es un documento que contiene los términos del acuerdo entre la compañía de seguros y el propietario de la póliza. En el mismo sentido, la aseguradora puede hacer uso de la transferencia de riesgo mediante contratos de reaseguro, cuando las pérdidas superan un límite establecido por la institución de seguros. De acuerdo con Delgado (2018), el reaseguro es el contrato donde el asegurador se protege en caso de que no cuente con los recursos suficientes para pagar a sus clientes ante desequilibrios de capital o pérdida de reservas acumuladas.

En este contexto, se mencionan a continuación las características que se toman en cuenta en la decisión de si un riesgo es asegurable o no.

1.4.1 Características de los riesgos asegurables.

Jones (1998), menciona que los productos de seguros se diseñan con base a los principios básicos que definen, que riesgo es asegurable y cuál no. Para que un riesgo o pérdida probable sea considerada asegurable, debe cubrir las siguientes características:

- **La pérdida debe ser debida al azar.**

Para que una pérdida potencial sea asegurable, debe contener un elemento de casualidad o azar. La pérdida debe ser provocada por un suceso inesperado o por un evento que no sea originado intencionalmente por la persona cubierta en el contrato de seguro.

- **La pérdida debe ser definitiva**

Este principio se refiere a que una pérdida asegurable debe ser definida en términos del tiempo y cantidad; es decir, la institución aseguradora tiene que definir cuando va a pagar el beneficio al asegurado y cuál será la cantidad a pagar.

De acuerdo con la forma en que una póliza establece el monto del beneficio, los seguros se pueden clasificar como un contrato de indemnización o un valor contractual. En un “*contrato de indemnización*”, el monto de beneficio o suma asegurada a pagar, se basa en el monto de la pérdida financiera resultante en el momento de ocurrir el siniestro. En este caso, la póliza declara que el monto del beneficio es igual al monto de la pérdida financiera cubierta o al monto máximo estipulado en el contrato, “lo que resulte menor”. Por otra parte, un “*valor contractual*” especifica el monto del beneficio que se pagará en caso de que acontezca una pérdida cubierta en el contrato de seguro, sin importar el importe actual de la pérdida ocurrida.

- **La pérdida debe ser significativa**

Este principio se relaciona al interés que hay entre el contratante del seguro y el objeto o persona asegurable. El costo administrativo que resultaría mantener el pago de beneficios por pérdidas no significativas, provocaría que el costo de mantener esa protección, sería demasiado elevado con relación al monto de la pérdida potencial y, por consiguiente, la mayoría de las personas no lo podrían cubrir.

- **La tasa de la pérdida debe ser predecible**

Para ofrecer un tipo de cobertura de seguros, una institución aseguradora puede calcular la tasa probable de pérdida que los asegurados sufrirán bajo esta cobertura. En concreto, la compañía pronosticará el número y el momento en que ocurrirá el riesgo asegurado cubierto, en un grupo de asegurados, determinando las primas adecuadas a cobrar en cada caso. En este sentido, en ningún caso se puede predecir la pérdida que una persona en particular experimentará en su vida. Se desconoce cuando una persona puede fallecer, se incapacite o necesite asistencia médica; sin embargo, se puede predecir con un alto grado de precisión, el número de personas que, dentro de un grupo específico, se lleve a cabo dichos eventos.

- **La pérdida no debe de ser catastrófica para la aseguradora**

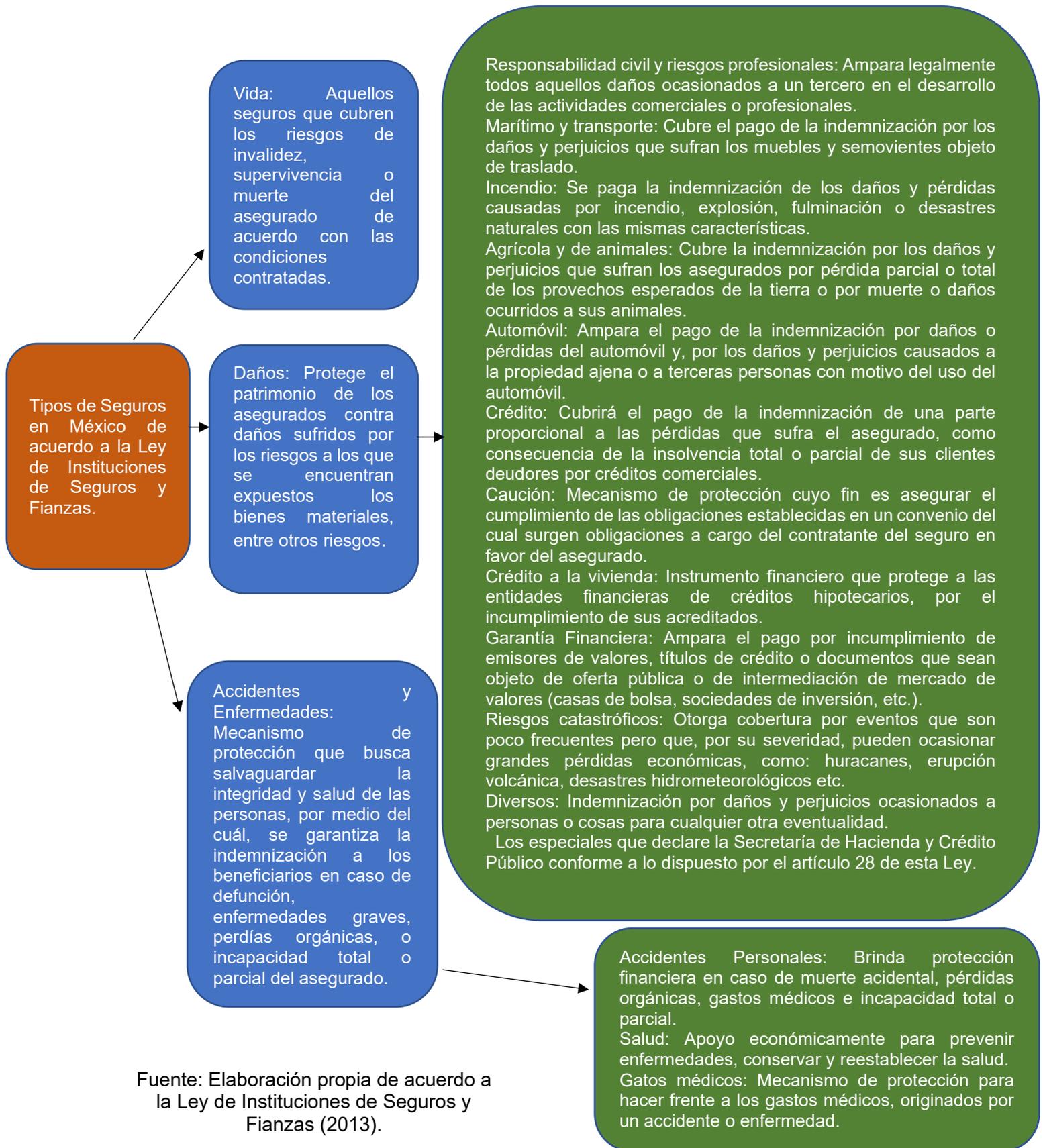
Una pérdida potencial no se considera como asegurable si un solo acontecimiento puede causar un daño financiero catastrófico a la institución de seguros. Alternativamente la institución de seguros puede reducir la posibilidad de sufrir pérdidas catastróficas por medio de la transferencia de riesgos a otra compañía de seguros, como se mencionó anteriormente.

Los principios básicos del riesgo, son aplicables de acuerdo a los riesgos asumidos por cada institución de seguros, pues éstas operan en diferentes ramos comerciales, como se describe en la siguiente sección.

1.5 Tipos de seguros

De acuerdo al artículo 25 de la Ley de Instituciones de Seguros y Fianzas, las operaciones y ramos de seguros que se ofrecen en México, bajo cualquier modalidad (Sociedades Anónimas o Sociedades Mutualistas) son (ver diagrama 1):

Diagrama 1. Tipos de seguros.



Fuente: Elaboración propia de acuerdo a la Ley de Instituciones de Seguros y Fianzas (2013).

De acuerdo con el diagrama 1, los diferentes productos de seguros que se ofrecen en México, permiten abarcar la mayoría de riesgos a los que se encuentran expuestas las personas y las empresas y, hoy en día el adquirir un seguro, es más posible para la población. De esta manera, en la presente investigación, se tiene interés en conocer el comportamiento de siniestralidad para un seguro de gastos médicos personales, por lo que, a continuación, se aborda las características principales y los antecedentes históricos más representativos.

1.6 Antecedentes del seguro de gastos médicos

La historia del seguro de gastos médicos se remonta en la antigüedad en Egipto, a través de acuerdos de defensa y asistencia mutua, por medio del cual, con la aportación de una cuota, quedaban protegidos los miembros en caso de enfermedad. También en Grecia, a los ciudadanos que sufrían un accidente y estuvieran limitados en sus actividades físicas a consecuencia de la defensa al Estado, recibían auxilio y educación para sus descendientes (Lozano & Morgado, 1997).

Posteriormente en la Edad Media, existieron formas de seguros de asistencia médica, para los marinos, en donde se les otorgaba protección contra algunas enfermedades y un seguro de pérdida de miembros para soldados. De acuerdo con Straulino (2014), las primeras formas de seguros de gastos médicos en la época moderna, surge en el año 1840, para los pasajeros que viajaban por tren, donde se les otorgaba una indemnización en caso de fallecimiento o heridas graves, a consecuencia de un accidente ocurrido durante el trayecto.

En México, los antecedentes del seguro de gastos médicos se encuentran en la Ley de Accidentes de Trabajo del Estado de México expedida en 1904 y la Ley de Accidentes de Trabajo del Estado de Nuevo León en 1906, donde se reconocía por primera vez en el país la obligación para los empresarios de atender a sus empleados en caso de enfermedad, accidente o muerte, derivados del cumplimiento de sus labores. Sin embargo, fue hasta la Constitución de 1917 con las sucesivas Leyes Federales del trabajo, que la

protección en materia de riesgo encontró base sólida para su funcionamiento (ISSEMyM, 2021).

En este sentido, los servicios de salud en México, son proporcionados a través de instituciones públicas, como instituciones de seguridad social compuestas por el Instituto Mexicano del Seguro Social (IMSS), el Instituto de Seguridad Social al Servicio de los Trabajadores del Estado (ISSSTE), los servicios médicos de Petróleos Mexicanos (PEMEX), la Secretaría de la Defensa Nacional (SEDENA) y la Secretaría de Marina, e instituciones aseguradoras privadas. En consecuencia, se mencionan las instituciones más representativas en protección de los trabajadores para atender enfermedades y algunos accidentes del sector público.

En 1943, se fundó el Instituto Mexicano del Seguro Social (IMSS), que es una de las instituciones con mayor presencia en el sector de salud y protección social, brindando atención médica para los trabajadores y sus familiares (IMSS, 1943). Posteriormente en el año 1959, surgió el Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), como un órgano público descentralizado con personalidad jurídica y patrimonios propios para administrar gran parte de los seguros, prestaciones y servicios, como la jubilación y las pensiones de los trabajadores del gobierno federal (ISSSTE, 2020), otorgando servicios de salud como atención médica preventiva, atención médica curativa y de maternidad, riesgos de trabajo, invalidez, entre otros.

En este sentido, dentro del sector privado, una de las problemáticas a las que se enfrentaban las instituciones de seguros era en la competitividad del precio de la prima, debido a que este seguro se encuentra sujeto al costo de tratamientos y generalmente, sus costos son elevados. Es así, a finales del siglo XX, el seguro de gastos médicos comienza su expansión y perfeccionamiento junto con los avances médicos y tecnológicos.

1.7 Características del seguro de gastos médicos

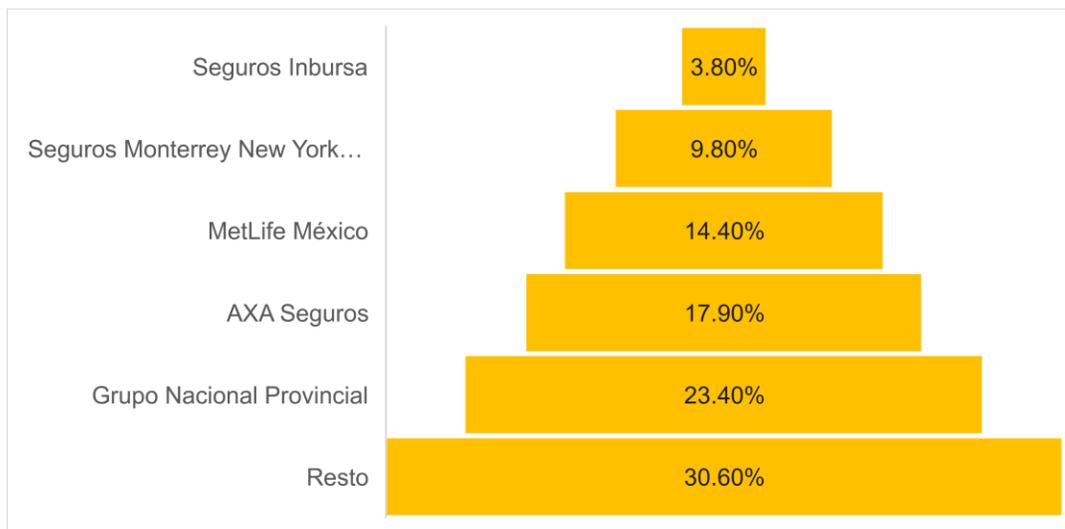
En definitiva, el contar con un seguro de gastos médicos personales en México es de suma importancia, debido a que una enfermedad pone en riesgo la salud de las personas y

frecuentemente el patrimonio familiar; además, el goce de una buena salud es un elemento central para la acumulación de capital humano, es decir, se puede determinar el tiempo de productividad laboral que tiene una persona a lo largo de su vida.

En este sentido, el seguro de gastos médicos tiene su fundamento en el Artículo 27 de la Ley de Instituciones de Seguros y Fianzas, sección IV (publicada el 4 de abril de 2013 en el Diario Oficial de la Federación): “Para el ramo de gastos médicos, estos seguros se toman como los contratos de seguro que tengan por objetivo cubrir los gastos médicos, hospitalarios y demás que sean necesarios para la recuperación de la salud o vigor vital del asegurado, cuando se haya afectado por causa de un accidente o enfermedad”.

Dicho de otra manera, el seguro de gastos médicos es el instrumento financiero que permite hacer frente a la posibilidad de quebranto económico al atender problemas de salud, principalmente aquellos que requieran atención médica – quirúrgica en un hospital derivados de un accidente o una enfermedad. Un aspecto significativo para este seguro, es que su índice CR-5 (*Concentration Ratio*)⁴, se ubicó en 69.4% para el año 2020 en México, representado de la siguiente manera (ver gráfica 2):

Gráfica 2. Índice CR-5 del seguro de Accidentes y Enfermedades, 2020.



Fuente: Elaboración propia con datos de la CNSF (2020).

⁴ Representa la suma de las cinco participaciones más grandes en términos de prima directa.

De acuerdo con la gráfica 2, las empresas con mayor captación de primas para el ramo de gastos médicos personales son: Grupo Nacional Provincial y AXA Seguros, debido a que, como se mencionó en párrafos anteriores, el seguro de accidentes y enfermedades, representó el mayor crecimiento real anual con un 7.4%, impulsado principalmente por el ramo de gastos médicos. En el mismo contexto, el porcentaje de asegurados en gastos médicos con respecto a la población total, de acuerdo con datos de la Asociación Mexicana de Instituciones de Seguros (AMIS) (2020), se muestra en la tabla 1.

Tabla 1. Porcentaje de asegurados en gastos médicos con respecto a la población total (2012 - 2020).

Año	2012	2013	2014	2015	2016	2017	2018	2019	2020
Porcentaje	7.2	7.5	7.7	7.5	8.1	8.6	8.5	9.0	9.5

Fuente: Elaboración propia con datos de la AMIS (2020).

De acuerdo con los datos proporcionados en la tabla 1, en promedio, existe un aumento anual de 0.29 puntos porcentuales de asegurados en gastos médicos personales. Por lo que ofertar este tipo de seguro, conduce a las compañías aseguradoras a enfrentar una competencia cada vez más grande, situación que impulsa a la actualización y modificación de los procesos tradicionales de administración y tarificación, en ocasiones sacrificando sus utilidades y descuidando los principales aspectos técnicos inherentes a este tipo de seguro (Saavedra, 2006).

Continuando con nuestro análisis, es importante mencionar que existen diferentes tipos de seguros de gastos médicos en México, los cuales dependen de las necesidades que se deseen cubrir y del número de personas que se desee asegurar. En cuanto al número de asegurados, se tiene la siguiente clasificación:

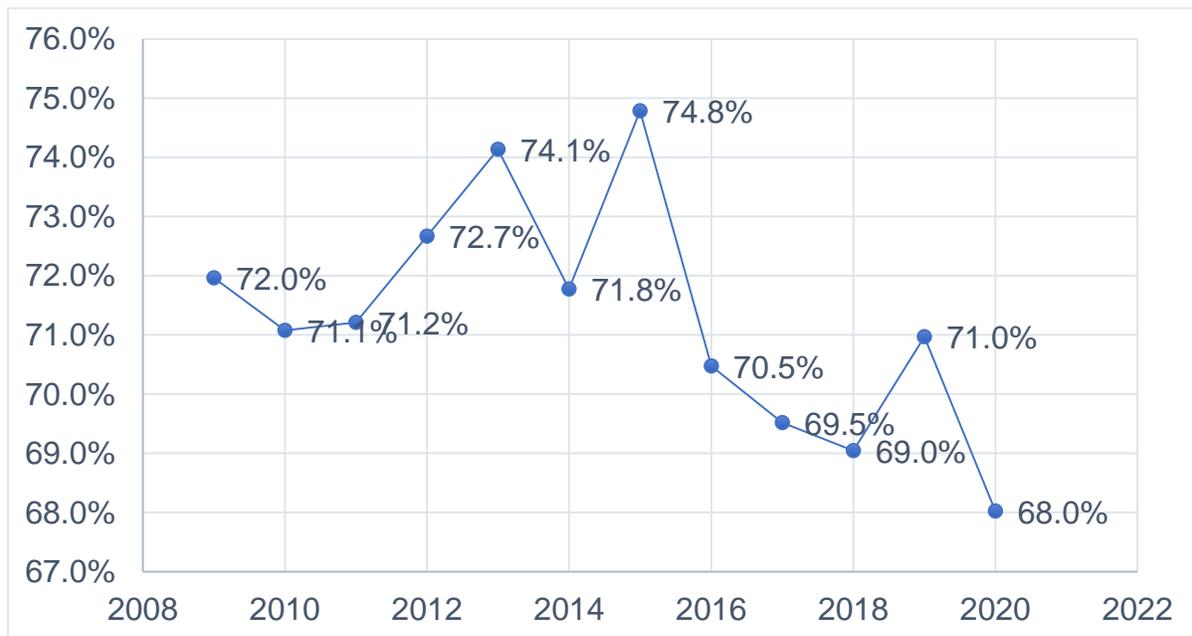
- 1) Individual: Este tipo de seguro, como su propio nombre lo dice, es de carácter personal, el cálculo de la prima se realiza con base a la siniestralidad de cada compañía de seguros y se emite una póliza individual.
- 2) Colectivo: Es un seguro dirigido a personas que desean asegurar a dos o más individuos con alguna relación en común, por ejemplo, cuando se desea

asegurar a una familia o bien para el caso de empresas u organizaciones que enfrentan el riesgo de pérdidas financieras a consecuencia de una enfermedad de sus empleados. Mediante este tipo de seguro quedan protegidos varios asegurados en una misma póliza.

Por otra parte, las coberturas que se ofrecen para este ramo son: cobertura financiera por honorarios médicos, hospitalización, medicamentos y exámenes médicos, que corresponden a la cobertura básica del seguro. En cuanto a las coberturas adicionales, estas pueden variar dependiendo de cada institución; sin embargo, algunas de las más representativas son: cuidados y atención dental especializada, emergencias en el extranjero, gastos funerarios, cobertura básica para cónyuges e hijos.

Otra característica importante del seguro de gastos médicos personales es el índice de siniestralidad (*Loss Ratio*), este valor representa la proporcionalidad que existe entre siniestros pagados y primas emitidas en un periodo de tiempo, por lo que, es de utilidad al evaluar el rendimiento anual de las compañías aseguradoras o bien para predecir el costo del seguro en los próximos años (ver gráfica 3).

Gráfica 3. Índice de siniestralidad del seguro de gastos médicos personales en México, 2009 - 2020.



Fuente: Elaboración propia con datos de la FES, CNSF, 2009 – 2020.

De manera general, en la gráfica 3 se puede observar que la siniestralidad de un seguro de gastos médicos personales en México, para el periodo 2016 - 2018, presenta una tendencia decreciente desde 2015, año en el que alcanza su nivel máximo (74.85), para llegar a un mínimo de 68.0% en 2020, lo cual trajo como resultado que las instituciones aseguradoras tuvieran buen rendimiento durante ese periodo, mientras que ante la presencia de un aumento del *loss ratio*, como en el caso del año 2015, las aseguradoras tienden a incrementar el precio del producto, como mecanismo de protección ante eventos desfavorables como la insolvencia.

Por otra parte, las modalidades de contratación para este tipo de seguro, regularmente es temporal a un año, ya que los factores de riesgo a los que se encuentran expuestas las instituciones de seguros, dificulta el cálculo de primas para mayores temporalidades. Haciendo contraste con un seguro de vida vitalicio, el factor de riesgo es el descenso de la persona y, se sabe que en algún momento determinado, la persona va a fallecer, por lo que el riesgo se puede cuantificar a través de las tablas de mortalidad; no obstante, para un seguro de gastos médicos, no existe estadística que pueda hacer en su conjunto todos los factores de riesgo que conlleva; no se sabe cuándo una persona pueda tener una enfermedad o un accidente y requiera atención médica, por lo que esta idea se puede analizar a mayor detalle en la siguiente sección del presente capítulo.

1.8 Teoría clásica del riesgo

De acuerdo con Markham (1962), las empresas de seguros están expuestas a dos tipos de riesgo: el riesgo comercial y el riesgo de seguros. El riesgo comercial se encuentra presente en todas las compañías que enfrenan riesgos comerciales, operativos, de inversión, reputacional, etc. Asimismo, el riesgo de seguros está relacionado a las fluctuaciones de las medidas de riesgo por la diferencia entre los montos de reclamos y los montos esperados de reclamos. Cramér (1930), de acuerdo con Markham, clasifica a estos riesgos en dos tipos: riesgos externos derivados de un exceso de mortalidad por guerras y epidemias y, en segundo lugar, el riesgo de variación aleatoria, el cual no es atribuible a alguna causa definida.

En este contexto, la teoría del riesgo es una herramienta para analizar este tipo de variaciones, y se puede considerar como una rama de las matemáticas actuariales que tiene como objetivo proporcionar modelos matemáticos para analizar las fluctuaciones de las variables de interés y proteger a las instituciones de seguros de eventos desfavorables que, por su severidad, pueden ocasionar la ruina. Gerber (1979), por su parte, define a la teoría del riesgo como el conjunto de ideas para diseñar, dirigir y regular una empresa de riesgos. Por esta situación, la teoría del riesgo se puede abordar desde dos puntos de vista diferentes: teoría de riesgo individual y teoría del riesgo colectivo.

1.8.1 Teoría del riesgo individual

Esta teoría busca medir el riesgo total de una cartera de asegurados como resultado de lo que pueda suceder a cada póliza admitida por ella, es decir, cada póliza con una cierta probabilidad de ocurrencia de siniestro, en donde la pérdida total será la suma de las contribuciones de las pólizas individuales. Esto es,

$$S = S_1 + S_2 + \dots + S_n = \sum_{i=1}^n S_i.$$

Siendo S la siniestralidad total de la cartera y S_i la siniestralidad de la póliza i -ésima, con $i = 1, 2, \dots, n$. Si se supone que los sumandos son variables aleatorias independientes, con la misma función de distribución $F(x)$, la variable aleatoria suma, S , tendrá por función de distribución:

$$F(x) = F_1(x) * F_2(x) * \dots * F_n(x)$$

Es decir, la n -ésima convolución de $F(x)$. De acuerdo al Teorema del Límite Central, esta suma es aproximadamente normal si el número de pólizas es lo suficientemente grande. No obstante, esta modalidad presenta algunos inconvenientes, como los menciona Cid (2020):

- La hipótesis de independencia de los sumandos que componen la siniestralidad total, no siempre se rechaza⁵.
- La distribución total no siempre se adecua al Teorema del Límite Central, si los grupos homogéneos tienen un número pequeño de pólizas.
- En algunas ocasiones, tiene una mayor complejidad matemática que la teoría del riesgo colectivo, dificultando su aplicación.
- No responde a preguntas como, ¿cuál es la probabilidad de que una institución aseguradora se arruine en un futuro?

1.8.2 Teoría del riesgo colectivo

En contraprestación surge la teoría del riesgo colectivo, la cual examina el comportamiento estadístico de las dos variables aleatorias que intervienen en la determinación de la siniestralidad: el número de siniestros y la cuantía o costo de cada uno de ellos, es decir, las dos variables que intervienen en el proceso de riesgo son: $u(t) \geq 0$ de tipo discreto, que representa el número de siniestros en un periodo de observación t y, $S_i(t) \in \mathbb{R}$, de carácter continuo, que se refiere a la cuantía correspondiente al i -ésimo siniestro, por lo que el total de pérdidas agregadas se obtiene a través de:

$$S = S_1 + S_2 + \dots + S_u = \sum_{i=1}^u S_i.$$

Donde el proceso será $S = 0$, si $u(t) = 0$.

Cid (2020), menciona que relacionado a las dos variables aleatorias del proceso del riesgo colectivo, las funciones que determinan las distribuciones básicas de probabilidad, serán la función de cuantía del número de siniestros $P_n(t)$, que representa la probabilidad de que en el periodo t , se produzcan n siniestros, siendo esto, $P(u(t) = n)$ y, la función de distribución de probabilidad de la cuantía del siniestro S_i , $V_i(x, t)$, que representa la probabilidad de que en el periodo t , la cuantía del i -ésimo siniestro no exceda de x

⁵ Puede suceder por pólizas duplicadas en la misma cartera (varias pólizas para la misma vida), planes de pensiones para trabajadores que laboran en la misma empresa, entre otros.

unidades monetarias, siendo así, $P(S_i \leq x)$, por lo que, la función de distribución del proceso total de riesgo de S será:

$$\begin{aligned}
 F_S(x) &= P(S \leq x) \\
 &= \sum_{n=0}^{\infty} P(u(t) = n) V_i(x/n) \\
 &= \sum_{n=0}^{\infty} P_n(t) P(S \leq x | u(t) = n) \\
 &= \sum_{n=0}^{\infty} P_n(t) P(X_1 + X_2 + \dots + X_n \leq x | u(t) = n) \\
 &= \sum_{n=0}^{\infty} P_n(t) F^{*n}(x)
 \end{aligned}$$

Donde $V_i(x/n) = P(S \leq x | u(t) = n)$, función de distribución de la cuantía del daño total si se han producido n siniestros y $F^{*n}(x) = P(X_1 + X_2 + \dots + X_n \leq x | u(t) = n)$, la n -ésima convolución de F . En la práctica, regularmente la variable aleatoria $u(t)$ sigue una distribución Poisson, por lo que se dice que S tiene una distribución Poisson compuesta.

De manera adicional, los procesos de simulación han tomado importante relevancia en el cálculo del monto de reclamos agregados, por lo que, en este proyecto se busca la aproximación a la cuantía de siniestralidad mediante modelos de simulación Monte Carlo vía Cadenas de Markov para los siniestros extremos, que se lleva a cabo con mayor detalle en los próximos capítulos.

CAPÍTULO II: Teoría de valores extremos

La teoría del riesgo vista en el anterior capítulo es utilizada comúnmente para expresar la siniestralidad esperada de una institución aseguradora a través de la ley de los grandes números, pues este teorema supone que, si se lleva a cabo una prueba repetidas veces, por ejemplo, lanzar una moneda, jugar a la ruleta rusa, etc., la frecuencia con la que se repetirá un determinado suceso, tenderá a una constante. Lo anterior es posible para representar valores cercanos a la media de la distribución; sin embargo, asumir que una distribución es capaz de expresar la siniestralidad total, incluyendo las colas, es erróneo. Por esta situación, surge la necesidad de recurrir a la teoría de valores extremos (EVT), con el fin de cuantificar los casos raros, los cuales no pueden ser medidos a través de la ley de los grandes números.

De acuerdo con Rydman (2018), la EVT es el conjunto de herramientas estadísticas que tienen el objetivo de predecir la ocurrencia de eventos que se desvían extremadamente de la media muestral. Podemos encontrar literatura de su amplia aplicación, por ejemplo, en disciplinas de climatología: eventos máximos de temperaturas Nogaj, et al. (2006), Coelho, et al. (2008), Cardarelo & Luraghi (2019); precipitaciones, Acero, García, & Gallego (2011), García, et al. (2007); e ingeniería para el desarrollo de construcciones modernas en áreas marítimas, Castillo, et al. (2004). Asimismo, es importante mencionar las características principales de los trabajos realizados en la medición de valores extremos en el sector financiero, destacando a Pérez & Tarango (2019), quienes realizaron una estimación de parámetros para predecir la ocurrencia de siniestralidad máxima de una cartera de asegurados de automóvil en un periodo de 2008 – 2017, a través del método de máxima verosimilitud y el método de momentos, y así, determinar el capital necesario para cubrir tales eventos.

Otra aportación es la de Brazauskas & Kleefeld (2009), quienes realizaron un análisis a los seguros daneses por pérdidas de incendios durante un horizonte de tiempo de 10 años, con el fin de mostrar la importancia en la elección de métodos de estimación de parámetros en el ajuste de la distribución de valores extremos. Los autores concluyen que los métodos clásicos probabilísticos ajustan bien a su muestra de datos, pues estas

metodologías se diseñaron para abarcar gran parte de la región paramétrica; sin embargo, al aplicar análisis de sensibilidad bajo escenarios de perturbación de datos, el método robusto, mostro un mejor comportamiento⁶.

También destacamos a Singh, Allen & Powell (2011), quienes realizaron un pronóstico de VaR dinámico a corto plazo, de las series financieras del índice ASX All Ordinaries (Australia) y S&P 500 (E.U.), de manera que, ajustan los residuos de un modelo GARCH (Modelo Autorregresivo Generalizado Condicional Heterocedástico) a una distribución de valores máximos y lo comparan con un modelo GARCH, en donde se supone que los residuos siguen una distribución normal, por lo que la metodología dinámica, mostró mejores resultados para las fluctuaciones de los índices.

Por otra parte, Murillo (2009) realizó una aplicación de teoría de valores extremos al riesgo operacional de una institución financiera por las pérdidas económicas en la línea de banca minorista, usando una metodología de Hill para la estimación de parámetros. Murillo menciona la importancia en la elección de un umbral, debido a que umbrales bajos suponen procesos con mayor número de observaciones, lo que disminuye la varianza.

Por último, se enfatiza a Sámano (2020), quien realizó un análisis a la saturación de vuelos del aeropuerto internacional de la Ciudad de México, bajo la teoría de valores extremos y un pronóstico de árboles de decisión para proponer alternativas en la saturación de llegadas y salidas de pasajeros.

En este contexto, la teoría de valores extremos es de importancia al cuantificar eventos máximos para evitar consecuencias graves en el futuro, por ejemplo, situaciones económicas, medioambientales etc., por lo que puede ser analizada desde dos perspectivas diferentes: la distribución Generalizada de Valores Extremos y, la distribución Generalizada de Pareto, las cuales se detallan a continuación.

⁶ Los métodos utilizados por Brazauskas & Kleefeld (2009) fueron: Máxima Verosimilitud (EMV), método de cuantiles o matching percentil, simulación Monte Carlo, tomando como valores iniciales la EMV y el método de momentos recortados.

2.1 Distribución Generalizada de Valores Extremos

Dada una sucesión de variables aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.), X_1, X_2, \dots, X_m , con función de distribución $F(x)$, y donde M_n representa el máximo de todas ellas, se pretende conocer la distribución de máximos:

$$F_{M_n}(x) = P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = [F(x)]^n.$$

De acuerdo al primer teorema fundamental (Fisher, Tippett, Gnedenko (1928)), la distribución asintótica de valores extremos de v.a.i.i.d., se puede aproximar mediante una distribución generalizada de valores extremos (GEVD), con parámetro de forma, $\varepsilon \in \mathbb{R}$; parámetro de localización, $\mu \in \mathbb{R}$ y, parámetro de escala, $\sigma > 0$. Así, la función de distribución es la siguiente:

$$GEVD(x; \varepsilon, \sigma, \mu) = \exp \left\{ - \left(1 + \varepsilon \left(\frac{x - \mu}{\sigma} \right) \right)^{-\frac{1}{\varepsilon}} \right\}$$

donde $1 + \varepsilon \left(\frac{x - \mu}{\sigma} \right) > 0$.

Como resultado, son tres distribuciones que son consideradas asintóticas a la GEVD, las cuales son: distribución Fréchet, distribución Gumbel y distribución Weibull.

- Cuando el parámetro de forma $\varepsilon > 0$, la función se puede interpretar como la distribución de Fréchet, con parámetro $\varepsilon > 0$, de forma; $\sigma > 0$, de escala y, $\mu \in \mathbb{R}$, de localización. En este caso la función de distribución está dada por:

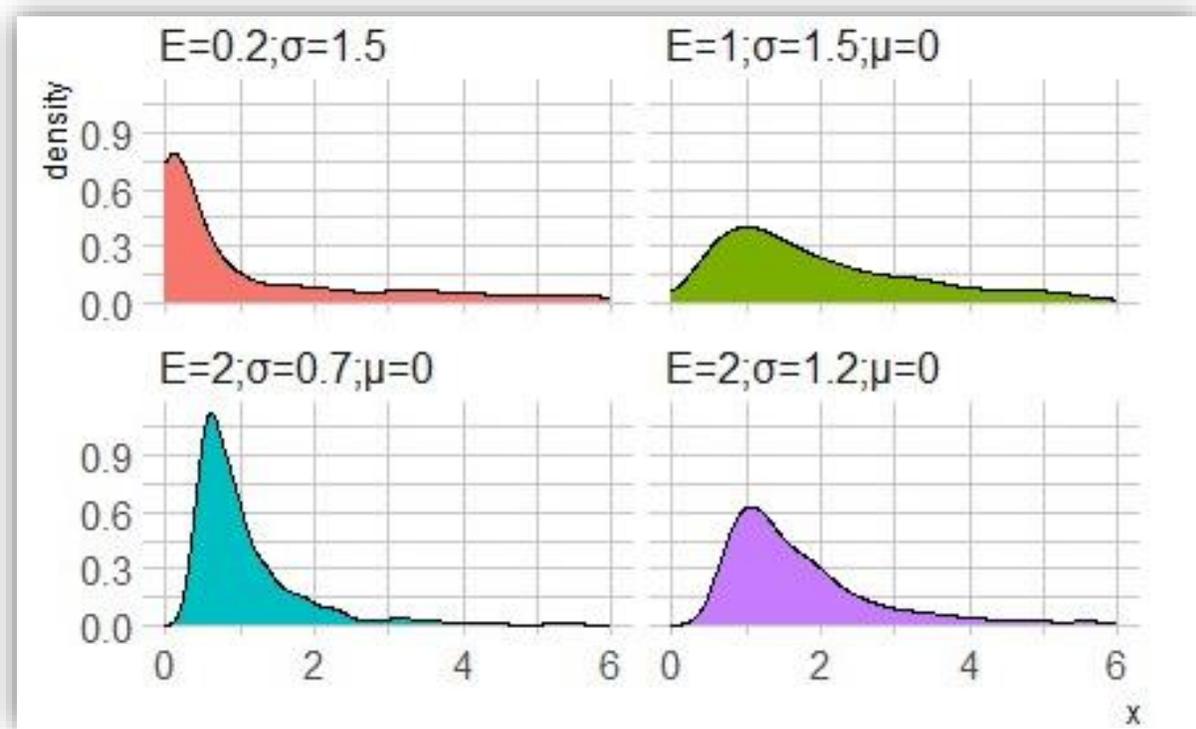
$$G_1(x; \varepsilon, \sigma, \mu) = \exp \left\{ - \left(\frac{x - \mu}{\sigma} \right)^{-\varepsilon} \right\}, \quad \text{para } x > \mu.$$

y, la función de densidad:

$$g_1(x; \varepsilon, \sigma, \mu) = \frac{\varepsilon \exp \{ -(x - \mu)/\sigma \}^{-\varepsilon} \left(\frac{x - \mu}{\sigma} \right)^{-\varepsilon - 1}}{\sigma}, \quad \text{para } x > \mu.$$

Tomando en cuenta diferentes valores para los parámetros de la distribución Fréchet, en la gráfica 4 se muestra el comportamiento de su densidad.

Gráfica 4. Densidad de la distribución Fréchet.



Fuente: Elaboración propia.

- Cuando $\mathcal{E} = 0$, la función que resurge es la distribución Gumbel con parámetro $\mu \in \mathbb{R}$, de localización; $\sigma > 0$, parámetro de escala y, función de distribución dada por:

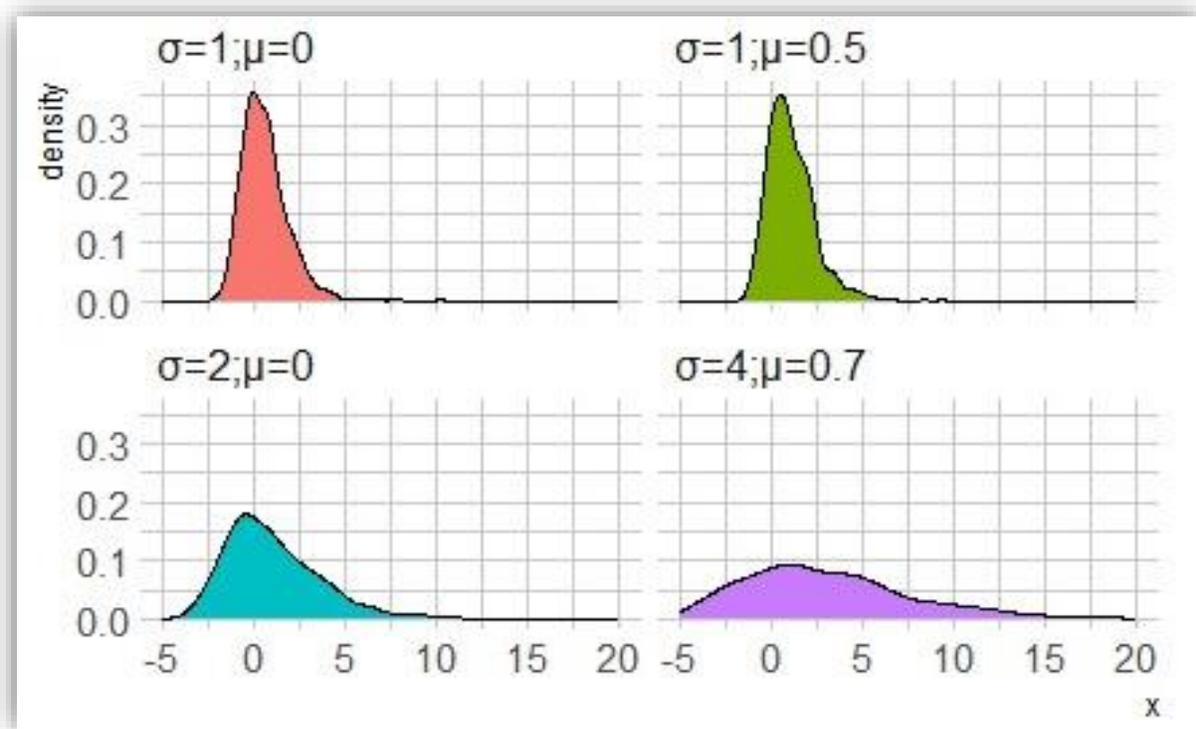
$$G_2(x; \sigma, \mu) = \exp \left\{ -\exp \left\{ \left(\frac{\mu - x}{\sigma} \right) \right\} \right\}$$

En cuanto a su función de densidad, es la siguiente:

$$g_2(x; \sigma, \mu) = \frac{\exp \left\{ \frac{(\mu - x)}{\sigma} - \exp \left\{ \frac{(\mu - x)}{\sigma} \right\} \right\}}{\sigma}, \quad x \in \mathbb{R}$$

De manera similar, teniendo en cuenta diferentes valores para los parámetros de la distribución Gumbel, la gráfica 5 representa el comportamiento de su densidad.

Gráfica 5. Densidad de la distribución Gumbel.



Fuente: Elaboración propia.

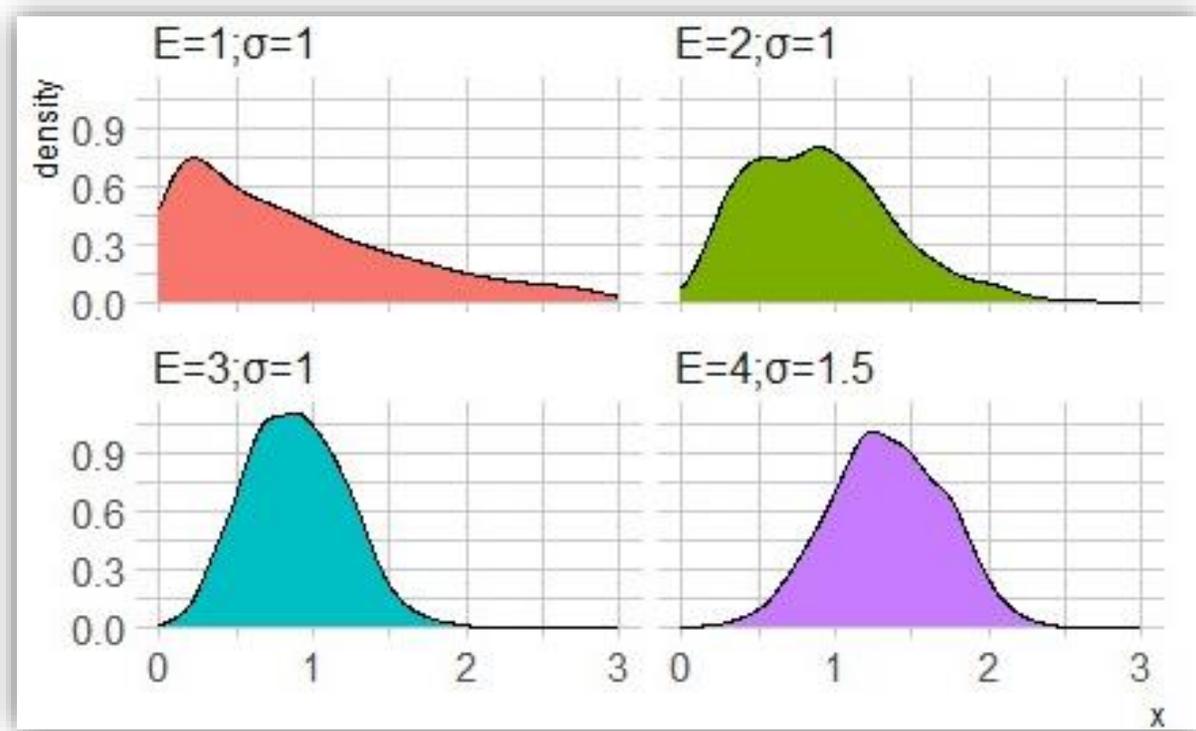
- Por último, cuando el parámetro de forma toma valores negativos ($\varepsilon < 0$), la GEVD, se interpreta como la distribución Weibull con $\varepsilon > 0$, parámetro de forma; $\sigma > 0$, parámetro de escala y, función de distribución:

$$G_3(x; \varepsilon, \sigma) = 1 - \exp\left\{-\left(\frac{x}{\sigma}\right)^\varepsilon\right\}$$

La función de densidad es la siguiente (ver gráfica 6):

$$g_3(x; \varepsilon, \sigma) = \frac{\varepsilon \exp\left\{-\left(\frac{x}{\sigma}\right)^\varepsilon\right\} \left(\frac{x}{\sigma}\right)^{\varepsilon-1}}{\sigma}, \quad x > 0$$

Gráfica 6. Densidad de la distribución Weibull.



Fuente: Elaboración propia.

De acuerdo con las gráficas presentadas, se puede observar que la distribución Fréchet presenta colas aún más pesadas que la distribución Weibull y la distribución Gumbel; sin embargo, en la práctica, la distribución Gumbel es de las más utilizadas, debido a sus propiedades analíticas. También, es importante mencionar que la elección de la distribución en el modelo de valores extremos, depende de las características y naturaleza de los datos que se estén analizando.

2.2 Distribución de Pareto Generalizada

La metodología de excesos sobre un umbral dentro de la teoría de valores extremos (Distribución Generalizada de Pareto) tiene sus orígenes con Pickands (1975), quien establece que los excesos $(X - u)$ sobre un umbral u , que sea lo suficientemente alto, dado $X \geq u$, tienen una distribución de Pareto Generalizada.

De acuerdo con Martín et al. (2022), sea X una variable aleatoria (v.a.) con función de distribución $F(x)$, si x_F es el extremo derecho finito o infinito de $F(x)$, esto es:

$$x_F = \sup \{x : F(x) < 1\}$$

En este contexto, la función de distribución para los excesos por encima de un umbral u , es:

$$F_{x_u}(x) = P[X_u \leq x] = P[X \leq x + u | X > u]$$

y, por definición de probabilidad condicionada, se obtiene:

$$\begin{aligned} &= \frac{P[u < X \leq x + u]}{P[X > u]} \\ &= \frac{F(x + u) - F(u)}{1 - F(u)} \end{aligned}$$

para $0 \leq x \leq x_{F-u}$. Es importante mencionar que X_u es la v.a. que obtenemos cuando consideramos la distribución por encima de un umbral que regularmente es llamada distribución de cola. En este sentido, dado un parámetro de forma $\gamma \in \mathbb{R}$, y un parámetro de escala $\sigma > 0$, la v.a. X sigue una Distribución Generalizada de Pareto (GPD), con función de distribución:

$$G(x|\gamma, \sigma) = \begin{cases} 1 - \left(1 + \frac{\gamma x}{\sigma}\right)^{-\frac{1}{\gamma}} & \gamma \neq 0 \\ 1 - \exp\left(-\frac{x}{\sigma}\right) & \gamma = 0 \end{cases} \quad (2.1)$$

La ecuación (2.1) es válida cuando $x \geq 0$, para $\gamma \geq 0$, y $0 \leq x \leq -\frac{\sigma}{\gamma}$ para $\gamma < 0$. En el caso en donde $\gamma = 0$, corresponde a una distribución exponencial con parámetro $1/\sigma$; sin embargo, cabe mencionar que este caso no presenta una importancia relevante en el presente proyecto, ya que esta distribución es considerada de cola ligera y no respondería al objetivo de encontrar una metodología que sea capaz de explicar los siniestros graves en el sector asegurador.

En el caso en donde $\gamma < 0$, se dice que la distribución corresponde a una cola acotada superiormente y de acuerdo con Zavershynskyy (2016), esta situación aparece

cuando la distribución subyacente también lo está, como es el caso de la distribución beta o la distribución normal. Sin embargo, estas distribuciones tampoco parecen ser útiles para explicar los siniestros máximos de un seguro de gastos médicos personales en México, por lo que, en la presente investigación se asume que la GPD tiene parámetro de forma, γ , estrictamente positivo.

De la función de probabilidad de la GPD podemos obtener su función de densidad como se muestra a continuación:

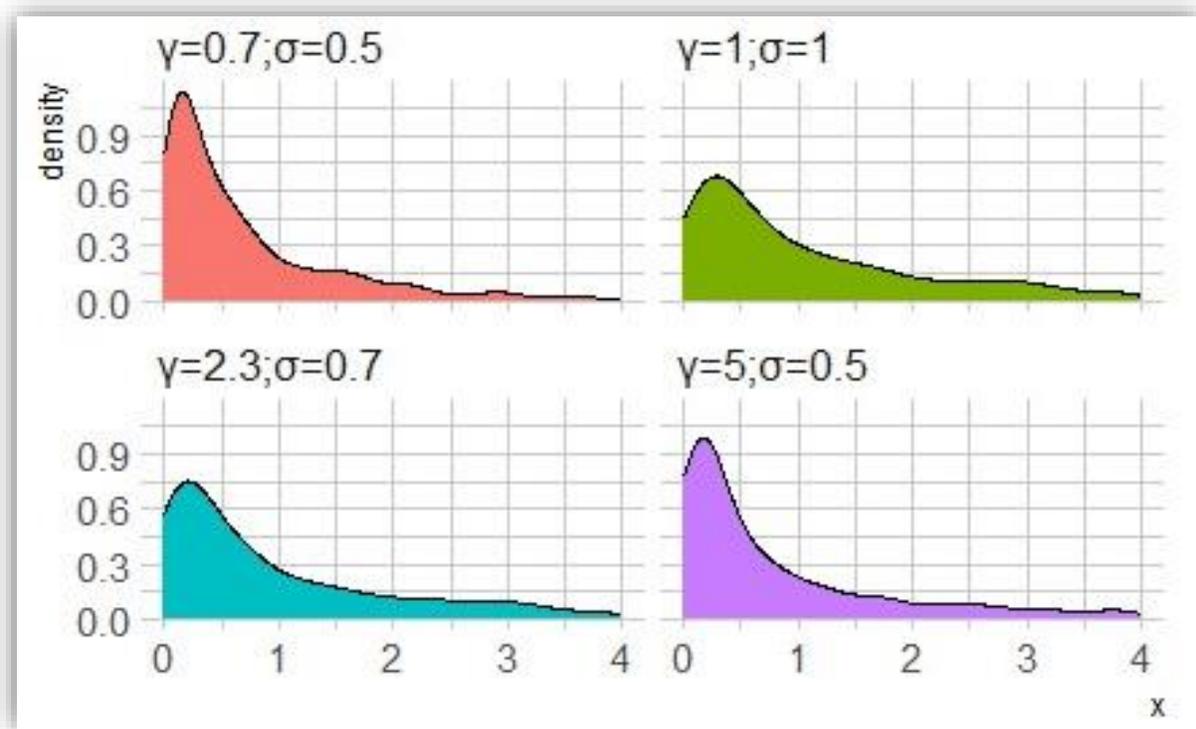
$$\begin{aligned} & \frac{\partial}{\partial x} \left(1 - \left(1 + \frac{\gamma x}{\sigma} \right)^{-\frac{1}{\gamma}} \right) \\ &= \frac{\partial}{\partial x} (1) - \frac{\partial}{\partial x} \left(\left(1 + \frac{\gamma x}{\sigma} \right)^{-\frac{1}{\gamma}} \right) \\ &= - \left(\frac{\partial}{\partial x} \left(\left(1 + \frac{\gamma x}{\sigma} \right)^{-\frac{1}{\gamma}} \right) \right) \\ &= - \left(-\frac{1}{\gamma} \left(1 + \frac{\gamma x}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}} \frac{\gamma}{\sigma} \right) \end{aligned}$$

Entonces, la densidad de la cola es:

$$g(x|\gamma, \sigma) = \sigma^{-1} \left(1 + \frac{\gamma x}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}}$$

A partir de lo anterior, se muestra una representación del comportamiento de la curva de densidad de la GPD, cuando consideramos diferentes valores para sus parámetros (ver gráfica 7).

Gráfica 7. Densidad de la GPD.



Fuente: Elaboración propia.

De acuerdo con la gráfica 7, podemos observar que el parámetro de escala determina la concentración de la distribución, ya que, al aumentar este valor, se presentan densidades más expandidas horizontalmente. En cuanto al parámetro de forma, sin importar el valor que esta toma, la distribución tiene presencia de valores atípicos.

Media y Varianza de la Distribución de Pareto Generalizada

La media y varianza de la GPD, vienen dadas por:

Media

$$E(x) = \frac{\sigma}{1-\gamma}, \quad \Leftrightarrow \gamma < 1.$$

Varianza

$$\text{var}(x) = \frac{\sigma^2}{(1-2\gamma)(1-\gamma)^2}, \quad \Leftrightarrow \gamma < \frac{1}{2}.$$

2.2.1 Distribuciones estables

Las distribuciones estables son una colección de distribuciones de probabilidad cuyas propiedades analíticas han contribuido a la elaboración de metodologías de sistemas físicos, hidrológicos, biólogos, económicos, entre otros, de manera que su uso responde a alternativas eficientes para analizar fenómenos de alta volatilidad; además, cumple con el teorema del límite central, es decir, contiene a la distribución normal como un caso particular de ésta. De acuerdo con Nolan (2013), la teoría de distribuciones estables fue desarrollada por el pionero Paul Lévy en el año 1930, quien examinó los límites que existe cuando se normaliza la suma de v.a.i.i.d.. Siendo más precisos, tenemos lo siguiente:

Una v.a. X es estable si para dos copias independientes de X , por ejemplo, X_1 y X_2 y, cualesquiera constantes a y b , se cumple:

$$aX_1 + bX_2 \xrightarrow{d} cX + d$$

para alguna contante $c > 0$ y algún $d \in \mathbb{R}$ (" \xrightarrow{d} " denota convergencia en distribución o mejor conocida como convergencia débil de v.a.).

Otra definición es la siguiente: se dice que una v.a. X sigue una distribución estable, con parámetros $a > 0$ y $b \in \mathbb{R}$, si se cumple,

$$X \xrightarrow{d} aZ + b \tag{2.2}$$

Donde Z es una v.a. con parámetros definidos por su función característica:

$$\begin{aligned} \varphi(u) &= E(\exp\{iuZ\}) \\ \varphi(u) &= \begin{cases} \exp\left\{-|u|^\alpha \left[1 - i\beta \tan\left(\frac{\pi\alpha}{2}\right) \text{sign}(u)\right]\right\}, & \alpha \neq 1 \\ \exp\left\{-|u| \left[1 + i\beta \frac{2}{\pi} (\text{sig } u) \ln|u|\right]\right\}, & \sigma = 1 \end{cases} \end{aligned} \tag{2.3}$$

$\text{Sign}(u)$ es la función signo de u que toma valores: 1 si $u > 0$; -1 si $u < 0$, y 0 si $u = 0$. En este sentido, una distribución estable depende de cuatro parámetros: el índice de estabilidad, $\alpha \in (0,2]$; el parámetro de asimetría, $\beta \in [-1,1]$; el parámetro de escala, $\gamma > 0$, y $\delta \in \mathbb{R}$, parámetro de localización (González, 2007).

Es importante mencionar que de acuerdo con Nolan (2005), la complejidad matemática de las distribuciones estables, son solo tres casos especiales que tienen función de densidad de probabilidad, mejor dicho, la integral respecto de (u) de la función característica (2.3), solo tiene solución analítica para los casos que se describen a continuación, α -estable $f_{\alpha,\beta}(\cdot|\gamma, \delta)$.

- Cuando $\alpha = \frac{1}{2}$, y $\beta = 1$ obtenemos la distribución Lévy con función de densidad:

$$f_{\frac{1}{2},1}(x|\gamma, \delta) = \left(\frac{\gamma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x - \delta)^{\frac{3}{2}}} \exp\left\{-\frac{\gamma}{2(x - \delta)}\right\}$$

- Cuando $\alpha = 1$, y $\beta = 0$ es una distribución Cauchy con función de densidad:

$$f_{1,0}(x|\gamma, \delta) = \frac{\gamma}{\pi((x - \delta)^2 + \gamma^2)}$$

- Cuando $\alpha = 2$, y $\beta = 0$ se hace presente la distribución Gaussiana con media δ y desviación estándar γ .

$$f_{2,0}(x|\gamma, \delta) = \frac{\exp\{-(x - \delta)^2/(2\gamma^2)\}}{\sqrt{2\pi}\gamma}$$

Continuando con nuestro razonamiento, se puede estandarizar la familia de distribuciones estables, los p -cuantiles q_p de una distribución estable X , pueden expresarse en términos de los p -cuantiles z_p , de la distribución Z como:

$$q_p = az_p + b \tag{2.4}$$

Suponiendo que X sigue una distribución estable con parámetros a y b para la ecuación (2.2), y fijamos el umbral $u = q_p$ para el problema de valores extremos. Entonces, para u suficientemente grande, z_p de (2.4) también es grande, y, consecuentemente se denota $u_z = z_p$. Es importante mencionar que a y b son llamadas constantes de normalización y corresponden a los parámetros de forma y escala de la GPD, por lo que, podemos aprovechar esta relación para realizar estimaciones de la forma $Z_{u_z} \sim GPD(\gamma_z, \sigma_z)$.

En tal sentido, tenemos lo siguiente:

$$\frac{X_u}{a} = Z_{uZ}$$

Por definición de probabilidad acumulativa:

$$\begin{aligned} F_{X_u}(x) &= P(X_u \leq x) = P\left(\frac{X_u}{a} \leq \frac{x}{a}\right) \\ &\approx G\left(\frac{x}{a} \mid \gamma Z, \sigma Z\right) = 1 - \left(1 + \frac{\gamma Z x}{\sigma Z a}\right)^{-\frac{1}{\gamma Z}} \\ &\Rightarrow X_u \sim GPD(\gamma Z, a\sigma Z) \end{aligned}$$

De esto se desprende que, el parámetro γZ permanecerá constante para todas las variables aleatorias del conjunto de distribuciones estables, sin importar los parámetros de la distribución subyacente, mientras que el parámetro de escala es obtenido a través del producto del parámetro de normalización a y, el parámetro de escala σZ para la distribución límite GPD de Z_{uZ} .

$$\gamma = \gamma Z; \sigma = a\sigma Z \quad (2.5)$$

En efecto, en el caso en donde la distribución subyacente sea Cauchy o Lévy, o cualquier distribución estable X con índice de estabilidad $0 < \sigma < 2$, se considera que la cola de la distribución es pesada y conduce a una GPD con $\gamma > 0$. Además, de acuerdo con Lévy (1925) citado en Martín et al. (2022), las distribuciones estables con índice de estabilidad $\alpha \neq 2$, presentan una propiedad interesante: Dada la distribución Z , su función de supervivencia $\bar{F}(x)$, se puede aproximar mediante:

$$\bar{F}(x) \sim (1 + \beta) C_\alpha x^{-\alpha}, \quad x \rightarrow \infty \quad (2.6)$$

Donde $C_\alpha = \frac{1}{\pi} \Gamma(\alpha) \sin\left(\frac{\alpha\pi}{2}\right)$. Consecuentemente se presenta la siguiente proposición.

Proposición 1: Cuando la distribución subyacente es una distribución estable estándar Z con $\alpha < 2$, la relación entre los parámetros de Z y los parámetros de la GPD que estima la distribución por encima del p -cuantil de Z , u_Z , es (Martín et al., 2022):

$$\hat{\gamma}_Z = \frac{1}{\alpha}; \hat{\sigma}_Z = \frac{1}{\alpha} \left(\frac{C_\alpha(1+\beta)}{1-p} \right)^{\frac{1}{\alpha}} \quad (2.7)$$

Dem. Dada la función de supervivencia de la distribución estándar Z , \bar{F} , y la función de supervivencia de la distribución límite GPD, \bar{G} de Z_{u_Z} donde u_Z es lo suficiente grande, por la proposición 1 y (2.6) que también es lo suficientemente grande, tenemos:

$$\bar{F}_{u_Z}(x) \sim \frac{(1+\beta)C_\alpha x^{-\alpha}}{\bar{F}_Z(u_Z)} = \frac{(1+\beta)C_\alpha x^{-\alpha}}{1-p}$$

Igualando ambas ecuaciones:

$$\left(1 + \frac{\gamma_Z}{\sigma_Z} x\right)^{-\frac{1}{\gamma_Z}} = \frac{(1+\beta)C_\alpha x^{-\alpha}}{1-p}$$

La parte derecha de la expresión, puede ser vista como:

$$\left(1 + \frac{\gamma_Z}{\sigma_Z} x\right)^{-\frac{1}{\gamma_Z}} = \left(\left[\frac{(1+\beta)C_\alpha}{1-p} \right]^{\frac{1}{\alpha}} x \right)^{-\alpha}$$

En este sentido, podemos tomar $\hat{\gamma}_Z = \frac{1}{\alpha}$ como estimador de γ_Z . Sustituyendo $\frac{1}{\alpha}$ por γ_Z , obtenemos:

$$\Rightarrow \left(1 + \frac{1}{\alpha \sigma_Z} x\right)^{-\alpha} = \left(\left[\frac{(1+\beta)C_\alpha}{1-p} \right]^{\frac{1}{\alpha}} x \right)^{-\alpha}$$

$$\Rightarrow 1 + \frac{1}{\alpha \sigma_Z} x = \left[\frac{(1+\beta)C_\alpha}{1-p} \right]^{\frac{1}{\alpha}} x$$

$$\Rightarrow \frac{1}{x} \left[1 + \frac{1}{\alpha \sigma_Z} x \right] = \left[\frac{(1+\beta)C_\alpha}{1-p} \right]^{\frac{1}{\alpha}}$$

$$\Rightarrow \frac{1}{x} + \frac{1}{\alpha \sigma_Z} = \left[\frac{(1+\beta)C_\alpha}{1-p} \right]^{\frac{1}{\alpha}}$$

$$\Rightarrow \frac{1}{\alpha \sigma z} = \left[\frac{(1 + \beta) C_\alpha}{1 - p} \right]^{-\frac{1}{\alpha}} - \frac{1}{x}$$

$$\Rightarrow 1 = \left(\left[\frac{(1 + \beta) C_\alpha}{1 - p} \right]^{-\frac{1}{\alpha}} - \frac{1}{x} \right) \alpha \sigma z$$

Ya que $\frac{1}{x}$ puede ser estadísticamente insesgado, entonces:

$$1 \sim \left[\frac{(1 + \beta) C_\alpha}{1 - p} \right]^{-\frac{1}{\alpha}} \alpha \sigma z$$

Por lo que, el estimador de σz es:

$$\hat{\sigma}_z = \frac{1}{\alpha} \left(\frac{C_\alpha (1 + \beta)}{1 - p} \right)^{\frac{1}{\alpha}} \quad \blacksquare$$

Por ende, en la aplicación metodológica nos aseguraremos de la precisión de estos estimadores mediante un amplio estudio de simulación.

CAPÍTULO III: Cadenas de Markov y los métodos Monte Carlo vía Cadenas de Markov.

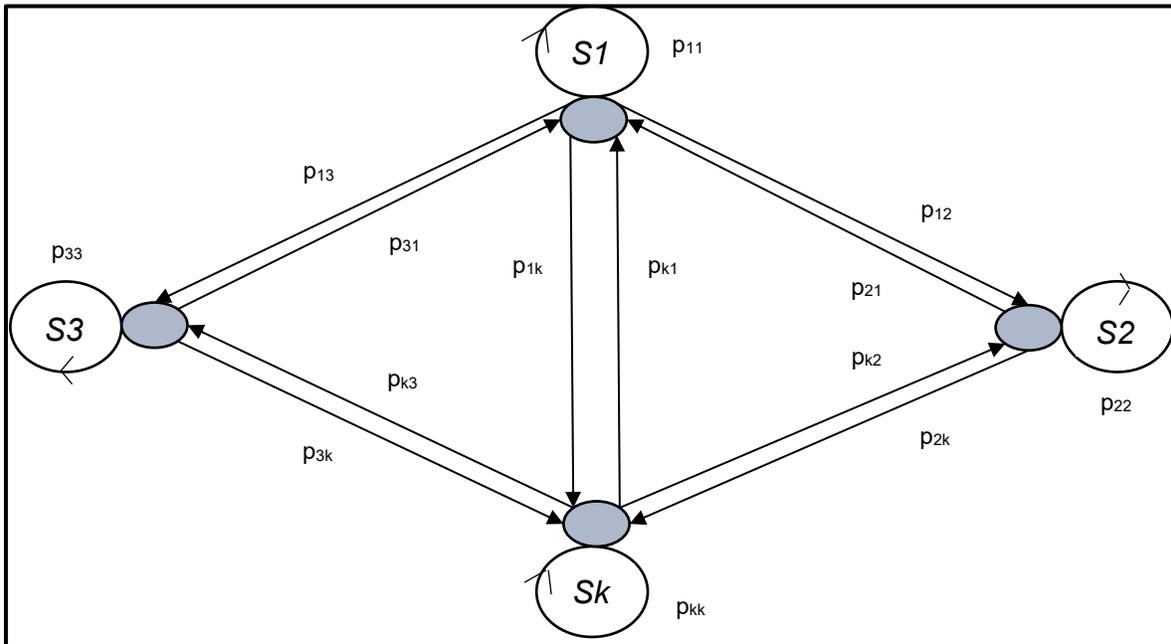
El presente capítulo tiene como objetivo mostrar las definiciones y las principales características referentes a las cadenas de Markov, las cuales son un caso particular de los procesos estocásticos; además, se aborda la teoría bayesiana esencial para el desarrollo de los métodos Monte Carlo vía Cadenas de Markov (MCMC) y así, proceder al estudio de simulación para los siniestros máximos de un seguro de gastos médicos personales en México en el periodo 2009 - 2020.

3.1 Cadenas de Markov

La teoría de los procesos estocásticos ha sido motivada ante la necesidad de modelar el comportamiento de sistemas que evolucionan a lo largo del tiempo o del espacio cuya dinámica se rige por el azar, como, el número de personas que esperan en el supermercado para pagar en un instante t de tiempo o el precio de las acciones de una empresa a lo largo de un año. De acuerdo con Barbosa & Llinás (2016), un proceso estocástico (X_t) con $t \in T$, es el conjunto de variables aleatorias (v.a.) X_t sobre un espacio de probabilidad común (Ω, δ, P) , donde dichas variables toman valores en un espacio medible (S, θ) , llamado espacio de estados. Los elementos de S se denominan estados. El conjunto $T \neq \phi$ se llama espacio de parámetros del proceso, en donde $T \subseteq \mathbb{R}$ o $T \subseteq \mathbb{N}$ y cada $t \in T$ representa el tiempo.

En este sentido, se definen a las cadenas de Markov como un tipo de proceso estocástico a tiempo discreto, donde $S \subseteq \{0,1, \dots\}$, finito o infinito y numerable, en cuyo caso se denomina a la cadena finita o infinita numerable. Entonces, dado el proceso estocástico $\{X_n\}$, la variable aleatoria X_n representa la posición de la cadena al n -ésimo paso o etapa y, X_0 a la posición inicial del proceso (Rincón, 2012) (ver diagrama 2).

Diagrama 2. Transición entre estados en una Cadena de Markov.



Fuente: Elaboración propia.

En el diagrama 2 se observa una representación de un proceso estocástico donde las variables de estado han sido discretizadas en k posibles situaciones (*estado 1, ..., estado k*), y el tiempo ha sido discretizado en etapas (horas, días, meses, etc.). En las cadenas de Markov se dice que la transición de un estado a otro, solo depende de ambos estados, es decir, se puede asociar una probabilidad p_{ij} a la transición del estado E_i en la fecha n , hacia el estado E_j en la fecha $n + 1$. Si $p_j(n)$, la probabilidad de que el proceso se encuentre en el estado E_j en una fecha n , y si las probabilidades de transición p_{ij} se mantienen constantes, se obtiene una Cadena de Markov (Bedoya & Barrera, 2006).

Una definición formal es la siguiente: Una cadena de Markov es un proceso estocástico a tiempo discreto $\{X_n: n = 0, 1, \dots\}$, con espacio de estados $S \subseteq \{0, 1\}$ y, satisface la propiedad de Markov, esto es, que para cualquier entero $n > 0$ y cualesquiera estados X_0, \dots, X_{n+1} , se cumple:

$$p(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = p(X_{n+1} = x_{n+1} | X_n = x_n)$$

En otras palabras, la probabilidad del evento futuro ($X_{n+1} = x_{n+1}$), solamente dependerá del evento presente ($X_n = x_n$), mientras que la información correspondiente al evento pasado ($X_0 = x_0, \dots, X_{n-1} = x_{n-1}$), es irrelevante.

Continuando con el análisis, sea una cadena de Markov $\{x_n\}$, $n \geq 0$, con espacio de estados finito S_k , donde k es el número de elementos y, variando los índices $i, j \in S_k$, se obtiene la matriz de probabilidades de transición en un paso como se puede apreciar a continuación; además, si se cumplen las propiedades de que la probabilidad de pasar del estado i a j es mayor a 0 y, si la suma de probabilidades de transición es 1, se dice que tenemos una **matriz completamente estocástica**.

$$\begin{pmatrix} p_{00} & p_{01} & p_{02} & \dots \\ p_{11} & p_{12} & p_{13} & \dots \\ p_{21} & p_{22} & p_{23} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Consecuentemente a lo anterior, la distribución de probabilidad conjunta de las variables aleatorias X_0, X_1, \dots, X_n , se puede obtener directamente de la propiedad de Markov, la cual es la siguiente:

$$p(x_0, x_1, x_2, \dots, x_n) = p(x_0) p(x_1|x_0) p(x_2|x_1) \dots p(x_n|x_{n-1}).$$

Esta ecuación establece que la distribución conjunta de v.a. $p(x_0, x_1, x_2, \dots, x_n)$, se encuentra determinada por la matriz de probabilidades de transición y por una distribución inicial, la cual representa la proporcionalidad de cada estado k sobre el espacio S_k , y que el sumando debe de ser igual a 1.

Probabilidad de transición en n pasos

La probabilidad de transición en n pasos de una cadena de Markov $p_{ij}^{(n)}$ o $p_{ij}(n)$, corresponde a la probabilidad de pasar del estado i en el tiempo m , al estado j en el tiempo $m + n$, y se representa como $p(X_{n+m} = j | X_m = i)$. En el mismo contexto, haciendo variar los índices $i, j \in S_k$, se obtiene la matriz de probabilidades de transición en n pasos como se muestra a continuación.

$$\begin{Bmatrix} p_{00}(n) & p_{01}(n) & \dots \\ p_{10}(n) & p_{11}(n) & \dots \\ \vdots & \vdots & \ddots \end{Bmatrix}$$

Es necesario mencionar que las cadenas de Markov se encuentran regidas por la ecuación de Chapman Kolmogorov, la cual permite calcular la probabilidad de trayectoria de un estado i hacia el estado j , y que atraviesa por un estado q cualquiera en un tiempo intermedio α . De acuerdo con Medel (2021), la definición es la siguiente: Sea $\{X_n\}, n \geq 0$, una cadena de Markov para cualquier par de tiempos α y n , tales que $0 \leq \alpha < n$, y para cualesquiera estados $i, j \in S_k$, se cumple:

$$p_{ij}(n) = \sum_q p_{iq}(\alpha) p_{qj}(n - \alpha).$$

Dem.

$$\begin{aligned} p_{ij}(n) &= p(X_n = j \mid X_0 = i) \\ &= p(X_n = j, X_0 = i) / p(X_0 = i) \end{aligned}$$

Por el teorema de probabilidad total obtenemos:

$$\begin{aligned} &= \sum_q p(X_n = j, X_\alpha = q, X_0 = i) / p(X_0 = i) \\ &= \sum_q p(X_n = j \mid X_\alpha = q, X_0 = i) p(X_\alpha = q, X_0 = i) / p(X_0 = i) \end{aligned}$$

Y por la propiedad de Markov:

$$\begin{aligned} &= \sum_q p(X_n = j \mid X_\alpha = q) p(X_\alpha = q \mid X_0 = i) \\ &= \sum_q p(X_{n-\alpha} = j \mid X_0 = q) p(X_\alpha = q \mid X_0 = i) \\ &= \sum_q p_{iq}(\alpha) p_{qj}(n - \alpha). \quad \blacksquare \end{aligned}$$

De esto se desprende que, si P es la matriz de probabilidades de transición en un paso, la probabilidad de trayectoria del estado i a j en n pasos, $p_{ij}(n)$, está determinada por la entrada (i, j) , de la matriz de probabilidades de transición P , elevada a la n -ésima potencia.

$$p_{ij}(n) = (P^n)_{ij}.$$

Por lo que, calcular las probabilidades de transición en n pasos, se transforma en obtener la n -ésima potencia de la matriz de probabilidades de transición en un paso:

$$p_{ij}(n) = \begin{pmatrix} p_{00}(n) & p_{01}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}^n = P_{ij}^n.$$

De este modo, si se conoce esta matriz y, si $p_i = p(X_0 = i)$ es una distribución inicial, la distribución de la variable X_n , será:

$$p(X_n = j) = \sum_i p_i p_{ij}(n).$$

Clasificación de estados en una cadena de Markov

Los estados de una cadena de Markov se pueden clasificar en dos tipos, de lo que depende, si una cadena de Markov es capaz de regresar al estado de partida, ya sea en uno o n pasos. Sea $\{X_n\}, n \geq 0$, una cadena de Markov, e i cualquier estado perteneciente a S_k , se pueden clasificar en recurrente o transitorio:

- El estado i es recurrente si $p(X_n = i \text{ para algún } n \geq 1 \mid X_0 = i) = 1$.
- El estado i es transitorio si $p(X_n = i \text{ para algún } n \geq 1 \mid X_0 = i) < 1$.

Adicionalmente, se dice que una cadena de Markov es **irreducible**, o tiene la propiedad de **irreductibilidad**, si todos sus estados se comunican entre sí. Es importante mencionar, que esta comunicación no se refiere a que todos los estados del espacio S_k sean accesibles a los mismos en un solo paso o transición, si no que, existe una probabilidad positiva de pasar del estado i, j en $n \geq 1$ pasos para toda la cadena.

Distribución estacionaria de una cadena de Markov

Sea $\{X_n: n = 0, 1, \dots\}$, una cadena de Markov con espacio de estado S_k , donde $k = 1, 2, \dots$; matriz de probabilidades de transición $P = p_{ij}$, y distribución inicial, $\pi^0 = \pi_0^0, \pi_1^0, \pi_2^0, \dots$, donde el superíndice indica tiempo igual a cero, es decir, la distribución de probabilidad al tiempo cero de la variable aleatoria X_0 . Nos interesa conocer la distribución de la variable aleatoria X_1 , entonces:

$$\pi_j^1 = p(X_1 = j)$$

Por el teorema de probabilidad total, lo anterior lo expresamos como:

$$\begin{aligned} &= \sum_i p(X_1 = j | X_0 = i) p(X_0 = i) \\ &= \sum_i \pi_i^0 p_{ij} \end{aligned}$$

En notación matricial, estas identidades las podemos ver como:

$$(\pi_0^1, \pi_1^1, \dots) = (\pi_0^0, \pi_1^0, \dots) \begin{pmatrix} p_{00} & p_{01} & \dots \\ p_{10} & p_{11} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \Rightarrow \pi^1 = \pi^0 P$$

De esta manera, la distribución al tiempo 1 está determinada por el producto de la distribución al tiempo 0 y la matriz de transiciones P . De manera general, la distribución al tiempo $n + 1$, es igual al producto de la distribución al tiempo n por la matriz P .

$$\pi^{n+1} = \pi^n P$$

Si se expresa la distribución al tiempo n , π^n , en términos de probabilidades anteriores, la ecuación converge al producto de la distribución inicial por la matriz P , en el tiempo $n + 1$:

$$\pi^{n+1} = \pi^0 P^{n+1}$$

De esta manera, la distribución inicial y las potencias de la matriz de probabilidades de transición, determinan la distribución de las variables aleatorias, X_n , de una cadena de Markov, lo que resulta en una sucesión de distribuciones de probabilidad $\pi^0, \pi^1, \pi^2, \dots$,

sobre el espacio de estados S_k . Consecuentemente, se dice que la distribución de probabilidad $\pi = \pi_0, \pi_1, \pi_2, \dots$, es estacionaria o invariante en el tiempo para una cadena de Markov con matriz de probabilidades de transición $P = p_{ij}$, si satisface:

$$\pi = \pi P$$

En otras palabras, si π es la distribución inicial de la Cadena de Markov, y si esta distribución es estacionaria, entonces π es la distribución de la variable aleatoria X_n , para cualquier valor de n .

Es importante mencionar que las distribuciones estacionarias juegan un papel importante en el comportamiento límite de una cadena de Markov, ya que ayudan a determinar en qué momento se ha alcanzado la distribución objetivo, es decir, después de varias transiciones las probabilidades de entrar a ciertos estados, convergen a valores particulares.

Cadenas de Markov reversibles en el tiempo

Sea una cadena de Markov $\{X_n : n \geq 0\}$ irreducible, con probabilidad de transición p_{ij} y, con distribución estacionaria π . También suponga un tiempo $m \geq 1$ fijo y definamos un nuevo proceso, $Y_n = X_{m-n}$, para $n = 0, 1, 2, \dots, m$, es decir, $\{Y_n : n = 0, 1, \dots, m\}$ es la cadena original en sentido inverso en el tiempo. Una cadena de Markov es reversible en el tiempo si para cualesquiera estados i, j se cumple:

$$\pi_i p_{ij} = \pi_j p_{ji}.$$

Avanzando en nuestro razonamiento, se determina las probabilidades del nuevo proceso de la siguiente manera:

$$\begin{aligned} Y_{ij} &= p\{X_m = j | X_{m+1} = i\} \\ &= \frac{p\{X_m = j, X_{m+1} = i\}}{p\{X_{m+1} = i\}} \\ &= p\{X_{m+1} = i | X_m = j\} \frac{p\{X_m = j\}}{p\{X_{m+1} = i\}} \end{aligned}$$

Además, las probabilidades de transición del estado j hacia el estado i dependen del cociente $p\{X_m = j\}/p\{X_{m+1} = i\}$; sin embargo, esta dependencia desaparece cuando se toma la existencia de una distribución estacionaria π , para $\{X_n: n \geq 0\}$, pues en este caso, la ecuación anterior se reduce a:

$$Y_{ij} = p_{ji} \frac{\pi_j}{\pi_i}$$

Por lo que, las probabilidades de transición de la nueva cadena ahora son también estacionarias. Si estas probabilidades son las mismas que la cadena original $\{X_n: n \geq 0\}$, satisface que $\pi_i p_{ij} = \pi_j p_{ji}$. De esta manera, el nuevo proceso resulta también ser una cadena de Markov, pues cumple el criterio de independencia entre pasado y futuro, cuando se conoce el presente (Calva, 2016).

Distribución límite

Todo lo mencionado hasta aquí, nos lleva a la siguiente definición: Toda matriz de probabilidades de transición $P = p_{ij}$ determina una sucesión de distribuciones de probabilidad, $\pi^0, \pi^1, \dots, \pi^n$ sobre el espacio de estados S_k . Entonces de acuerdo a la definición de estacionariedad tenemos:

$$\pi^n = \pi^0 P^n, \quad n \geq 1 \quad (3.1)$$

Es interesante conocer, que bajo ciertas condiciones la sucesión de variables aleatorias es convergente a una distribución de probabilidad límite π . Si se considera esta hipótesis como verdadera, tendremos lo siguiente:

$$\pi = \lim_{n \rightarrow \infty} \pi^n \quad (3.2)$$

De la expresión (3.2), se desprenden algunas incógnitas que analizamos a continuación. Cuando el límite $n \rightarrow \infty$ en la ecuación (3.1) obtenemos:

$$\pi = \pi P \quad (3.3)$$

$$\pi = \pi^0 \left(\lim_{n \rightarrow \infty} P^n \right) \quad (3.4)$$

De acuerdo a la ecuación (3.3) se puede observar que la distribución límite es una distribución estacionaria y asimismo no depende de la distribución inicial; además, la ecuación (3.4), indica que la distribución límite, está determinada por la n -ésima potencia de la matriz de probabilidades de transición P , y finalmente el límite de las potencias de P , es una matriz con todos sus renglones idénticos, siendo este renglón la **distribución límite**.

Tomando en cuenta una cadena de Markov con matriz de probabilidades $P = (p_{ij})$ y distribución inicial π^0 , se denomina distribución límite de la cadena de Markov al vector (Salas, 2013):

$$\pi = \lim_{n \rightarrow \infty} \pi^0 P^n = \lim_{n \rightarrow \infty} \sum_i \pi_i^0 p_{ij}^{(n)}.$$

3.2 Los métodos Monte Carlo vía Cadenas de Markov.

Comúnmente, en la estimación de parámetros de la distribución de valores máximos en cualesquiera de sus dos vertientes (picos sobre el umbral y el método de bloques), suponen desperdiciar gran cantidad de información, y esto podría influir en la precisión de las estimaciones. Por esta situación la teoría bayesiana resulta conveniente para abordar estos casos, y así, aprovechar la información disponible de los datos.

En un estudio clásico de carácter general, la técnica estándar de estimación de densidades, consiste en ajustar el conjunto de datos disponibles a diferentes distribuciones teóricas y elegir la mejor de ellas, según un criterio determinado, siendo esto ideal para representar el área central de las curvas, pero no el área de las colas.

Debido a la importancia de la teoría de valores extremos, se han propuesto varios métodos para estimar los parámetros de la distribución de máximos; sin embargo, no existen muchas aplicaciones mediante técnicas bayesianas. Podemos mencionar a Mora & Perret (2007), quienes aplicaron una técnica de Gibbs y un procedimiento BLUP, que es una variante de los métodos MCMC, en el análisis genético de árboles forestales. Martínez, Alvear, & Tovar (2020), por su parte, estimaron los parámetros de localización y

de escala de la distribución Lévy, mediante el algoritmo Metropolis – Hastings, tomando como distribuciones *a priori* la distribución uniforme y la distribución gamma y, Zhang & Stephens (2009), emplearon la propia distribución de valores máximos como densidad *a priori*.

También se hace mención a Martín et al. (2020), quienes realizaron un estudio de simulación para encontrar la relación de parámetros entre la distribución Gumbel de valores extremos y la distribución subyacente: Exponencial, Normal y Gumbel y, de esta manera, incorporar nuevas metodologías en el análisis bayesiano. De forma simultánea, Martín et al. (2022), realizaron un estudio de simulación para estimar los parámetros de la distribución de Pareto Generalizada sobre valores máximos de contaminación atmosférica en un periodo de 2011 - 2020.

Por lo anterior el uso de la estimación bayesiana a menudo se asocia con dificultades para elegir una distribución *previa*. Para esta situación se han propuesto varias alternativas, por ejemplo, centrarse específicamente en la elección de la distribución *a priori*, o bien mejorando la eficiencia del algoritmo. En este sentido, los métodos MCMC, se basan en la construcción de una Cadena de Markov que converge a una densidad objetivo y equilibrada; además, de acuerdo con Bautista (2007), para poder realizar una correcta estimación de parámetros bajo la teoría bayesiana, la cadena de Markov debe de cumplir las siguientes características:

- Debe de ser aperiódica $d(i) = 1$, es decir, todos sus estados deben de estar comunicados consigo mismos, $p_{ii}(n) > 0$.
- Tiene que ser irreducible (contener una única clase de estados).

Por consiguiente, se propone un estudio de simulación mediante las metodologías MCMC, específicamente Metropolis Hastings (MH), Baseline Metropolis Hastings (BMH), e Informative Priors Baseline Metropolis Hastings (IPBMH), que, asegura una mejora en la precisión de las estimaciones, debido a sus propiedades analíticas por contener distribuciones *a priori* altamente informativas; además, se incorporan tres casos especiales de distribución subyacente: Lévy, Cauchy y Normal.

Conceptos básicos de la estimación bayesiana

En la estimación bayesiana el pilar fundamental es el teorema de Bayes, el cuál surge en la teoría de la probabilidad por el matemático inglés Thomas Bayes (1702 – 1761), quien determinó la utilidad del cálculo de probabilidades condicionales de la forma $P(A_j | B)$ cuando se conocen $P(B | A_j)$ y $P(A_j)$. Sea A_1, A_2, \dots una partición de Ω tal que $P(A_i) \neq 0$, para $i = 1, 2, \dots$ y, sea B un evento cualquiera tal que $P(B) \neq 0$. Entonces para cada $j = 1, 2, \dots$, la propiedad de Bayes, es la siguiente:

$$p(A_j|B) = p(B|A_j) p(A_j) / \sum_{i=1}^{\infty} p(B|A_i) p(A_i).$$

Dem.

Partiendo de la definición de probabilidad condicional:

$$p(A_j|B) = \frac{p(A_j \cap B)}{p(B)}$$

el numerador se puede expresar como:

$$= p(B|A_j) p(A_j)$$

y, el denominador por el teorema de probabilidad total:

$$= \sum_{i=1}^{\infty} p(B|A_i) p(A_i)$$

Entonces, resulta el Teorema de Bayes:

$$p(A_j|B) = p(B|A_j) p(A_j) / \sum_{i=1}^{\infty} p(B|A_i) p(A_i). \quad \blacksquare$$

El teorema de Bayes parte del principio que la información *previa* de la simulación estocástica, se puede cuantificar mediante una distribución de probabilidad que es llamada distribución *a priori*; además, de acuerdo con Zacarías (2006), la información de los datos se cuantifica a través de la función de verosimilitud, la cual, si las variables aleatorias son

independientes, se considera como función de verosimilitud a la densidad conjunta de v.a., que a su vez será el producto de las densidades marginales:

$$L(\theta) = L(x|\theta) = f(x_1, \dots, x_n | \theta)$$

$$\Rightarrow f(x|\theta) = \prod_{i=1}^n f(x_i|\theta)$$

3.2.1 Metropolis Hastings

En la estadística clásica se conoce como muestreo a la técnica de selección de un subconjunto de observaciones pertenecientes a una población; no obstante, en muchas ocasiones el muestreo directo no es simple. El algoritmo MH, es un método MCMC, que permite generar una secuencia de muestras aleatorias (Cadena de Markov), a partir de una distribución de probabilidad donde cada iteración del proceso, solamente depende de dos pasos: proponer un nuevo estado del sistema y, aceptarlo o rechazarlo dado un criterio determinado.

En este sentido, para realizar inferencia estadística en el enfoque bayesiano, tras asumir una densidad *a priori* para los parámetros $\pi(\theta)$ y, combinando esta distribución con la información aportada por los datos que se cuantifican mediante la función de verosimilitud $L(\theta|x)$, se puede determinar la función de densidad *a posteriori* de los parámetros de la forma $\pi(\theta|x) \propto L(\theta|x)\pi(\theta)$, donde \propto significa “es proporcional”.

Dada la definición anterior, el algoritmo MH sigue los pasos siguientes:

Paso 1.- Escoger un valor inicial del parámetro θ^i .

Paso 2.- Generar un parámetro candidato θ^r , a través de una distribución de prueba que determinamos como $q(\theta^r|\theta^{i-1})$.

Paso 3.- Evaluar la probabilidad de aceptación a través de:

$$\alpha = p(\theta^r, \theta^{i-1}) = \min \left\{ \frac{\pi(\theta^r|\theta^{i-1}) q(\theta^r, \theta^{(i-1)})}{\pi(\theta^{(i-1)}|\theta^r) q(\theta^{(i-1)}, \theta^r)}, 1 \right\}$$

Paso 4.- General un valor $u \sim U(0,1)$.

Paso 5.- Evaluar si $u < \alpha$, se fija $\theta^i = \theta^r$. En caso contrario $\theta^i = \theta^{(i-1)}$.

Paso 6.- $i = i + 1$ y regresar al paso 2.

Este proceso representa el núcleo de transición de la metodología MH y, después de varias transiciones, la cadena de Markov converge a una distribución estacionaria $\pi(x)$. Nuestro proceso aleatorio se mueve de θ^r, θ^{i-1} , más a menudo que de θ^{i-1}, θ^r , por lo que, el algoritmo asegura el equilibrio a través de la propiedad de reversibilidad de cadenas de Markov, y así, restringiendo muestras que se encuentran fuera del espacio paramétrico.

Consideraciones adicionales

Avanzando en nuestro razonamiento, el algoritmo MH presenta algunas características importantes en la elección de la distribución de prueba $q(\theta^r, \theta^{(i-1)})$ como son:

- Suele escogerse una distribución simétrica, como la distribución normal o la distribución uniforme y de esta manera simplificar la probabilidad de aceptación a:

$$p(\theta^r, \theta^{i-1}) = \min \left\{ \frac{\pi(\theta^r | \theta^{i-1})}{\pi(\theta^{(i-1)} | \theta^r)}, 1 \right\}$$

- Además, si la distribución *a priori* es altamente informativa, la probabilidad de aceptación se reduce a:

$$p(\theta^r, \theta^{i-1}) = \min \left\{ \frac{L(\theta^r | x)}{L(\theta^{i-1} | x)}, 1 \right\}$$

En este panorama, es importante mencionar que, para hacer inferencia bayesiana con MH para el caso de las distribuciones estables, tenemos que distinguir las situaciones en donde nos encontramos con colas ligeras ($\gamma < 0$), y colas pesadas ($\gamma > 0$). Supongamos que $X_u \sim GPD(\gamma, \sigma)$, y se dispone de m valores. Sea $x = x^1, x^2, x^3, \dots, x^n$, una muestra de n valores de x y, $x_u = x_u^1, x_u^2, \dots, x_u^m$, una muestra de m valores de x_u .

Colas ligeras ($\gamma < 0$)

Considerando $\tau = -\gamma$ y $\delta = -\frac{\sigma}{\gamma}$, entonces $X_u \sim GPD(-\tau, \tau\delta)$, por lo que, la función de verosimilitud, se cuantifica mediante:

$$L(\tau, \delta | x_u^i) = \prod_{i=1}^m f(x_u^i) = \prod_{i=1}^m \left[\tau^{-m} \delta^{-m} \left(1 - \frac{x_u^i}{\delta}\right)^{\frac{-\tau+1}{\tau}} \right]$$

$$L(\tau, \delta | x_u^i) = \tau^{-m} \delta^{-m} \prod_{i=1}^m \left(1 - \frac{x_u^i}{\delta}\right)^{\frac{-\tau+1}{\tau}}$$

Al respecto, para las distribuciones *a priori*, consideramos una distribución $\Gamma(a_0, b_0)$ y $\Gamma(a_1, b_1)$, para ambos parámetros τ y δ de la GPD. Entonces, las distribuciones posteriores condicionales son:

$$\pi(\tau | \delta, x_u) \propto \tau^{a_0-1-m} \exp\{-b_0\tau\} \prod_{i=1}^m \left(1 - \frac{x_u^i}{\delta}\right)^{\frac{-\tau+1}{\tau}}$$

$$\pi(\delta | \tau, x_u) \propto \delta^{a_1-1-m} \exp\{-b_1\delta\} \prod_{i=1}^m \left(1 - \frac{x_u^i}{\delta}\right)^{\frac{-\tau+1}{\tau}}$$

Notar que la probabilidad de aceptación está definida por el mínimo entre uno y alfa, que se calcula mediante:

$$\alpha_\tau = \left(\frac{\tau^*}{\tau^{(j)}}\right)^{a_0-m-1} \exp\left\{b_0(\tau^* - \tau^{(j)}) + \left(\frac{1}{\tau^*} - \frac{1}{\tau^{(j)}}\right) \sum_{i=1}^m \ln\left(1 - \frac{x_u^i}{\delta^{(j)}}\right)\right\}$$

$$\alpha_\delta = \left(\frac{\delta^*}{\delta^{(j)}}\right)^{a_1-m-1} \exp\left\{b_1(\tau^* - \tau^{(j)}) + \left(\frac{-\tau^{(j)} + 1}{\tau^{(j)}}\right) \sum_{i=1}^m \ln\left(1 - \frac{x_u^i}{\delta^*}\right) - \left(\frac{-\tau^{(j)} + 1}{\tau^{(j)}}\right) \sum_{i=1}^m \ln\left(1 - \frac{x_u^i}{\delta^{(j)}}\right)\right\}$$

De esta manera, se obtiene las estimaciones de γ y σ de la GPD con $\gamma < 0$, mediante $\gamma = -\tau$ y $\sigma = \tau\delta$.

Colas pesadas ($\gamma > 0$)

La función de verosimilitud está definida como:

$$L(\gamma, \sigma | x_u^i) = \prod_{i=1}^m f(x_u^i) = \prod_{i=1}^m \left[\sigma^{-1} \left(1 + \frac{\gamma x_u^i}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}} \right]$$

$$L(\gamma, \sigma | x_u^i) = \sigma^{-m} \prod_{i=1}^m \left(1 + \frac{\gamma x_u^i}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}}$$

En cuanto a las distribuciones *a priori*, se considera una distribución Pareto tipo 1(a_0, b_0), para el parámetro γ y, una distribución $inv\Gamma(a_1, b_1)$, para el parámetro de escala σ , debido a que estas distribuciones son comunes en la práctica.

$$\pi(\gamma) \propto \gamma^{-(a_0+1)}, \text{ con } \gamma > b_0$$

$$\pi(\sigma) \propto \exp\left\{-\frac{b_1}{\sigma}\right\} \sigma^{-(a_1+1)}$$

Las distribuciones *posteriores* condicionales son las siguientes:

$$\pi(\gamma | \sigma, x_u) \propto \gamma^{-(a_0+1)} \prod_{i=1}^m \left(1 + \gamma \frac{x_u^i}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}}$$

$$\pi(\sigma | \gamma, x_u) \propto \sigma^{-(m+a_1+1)} \exp\left\{-\frac{b_1}{\sigma}\right\} \prod_{i=1}^m \left(1 + \gamma \frac{x_u^i}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}}$$

En el mismo sentido, aplicando MH, notar que los ratios de aceptación están definidos por el mínimo entre 1 y alfa, que se calculan mediante:

$$\alpha_\gamma = \left(\frac{\gamma^{(j)}}{\gamma^*} \right)^{(a_0+1)} \exp \left\{ \left(\frac{\gamma^{(j)} + 1}{\gamma^{(j)}} \right) \sum_{i=1}^m \ln \left(1 + \gamma^{(j)} \frac{x_u^i}{\sigma^{(j)}} \right) - \left(\frac{\gamma^* + 1}{\gamma^*} \right) \sum_{i=1}^m \ln \left(1 + \gamma^* \frac{x_u^i}{\sigma^{(j)}} \right) \right\}$$

$$\alpha_\sigma = \left(\frac{\sigma^{(j)}}{\sigma^*}\right)^{m+a_1+1} \exp \left\{ b_1 \left(\frac{1}{\sigma^{(j)}} - \frac{1}{\sigma^*} \right) + \left(\frac{\gamma^{(j)} + 1}{\gamma^{(j)}} \right) \sum_{i=1}^m \ln \left(1 + \gamma^{(j)} \frac{x_u^i}{\sigma^{(j)}} \right) - \left(\frac{\gamma^{(j)} + 1}{\gamma^{(j)}} \right) \sum_{i=1}^m \ln \left(1 + \gamma^{(j)} \frac{x_u^i}{\sigma^*} \right) \right\}$$

Es así, se reiteran los pasos anteriores y de esta manera se construye la Cadena de Markov, que después de varias iteraciones, converge a una distribución objetivo de los parámetros γ y σ , con GPD $\gamma > 0$.

3.2.2 Baseline Metropolis Hastings

La siguiente propuesta es el método Baseline Metropolis Hastings, el cual fue diseñado con el objetivo de aprovechar toda la información disponible de los datos y, de este modo hacer uso de la relación de parámetros de la distribución base y la distribución de valores extremos. De manera general, el proceso es el siguiente:

Paso 1.- Aplicar Metropolis Hastings para estimar los parámetros de la distribución subyacente θ .

Paso 2.- Utilizar la relación de parámetros de la distribución subyacente θ , con los parámetros de la distribución $GPD(\gamma, \sigma)$ y, calcular las estimaciones de γ y σ .

En el caso de las distribuciones estables, dicha relación se ha explicado con anterioridad y viene dada por la ecuación (2.5). Entonces, para las tres distribuciones seleccionadas, el proceso detallado es el siguiente:

1.- Aplicar el algoritmo MH para estimar el parámetro de escala a partir de la distribución subyacente estable.

2.- Hacer uso de la relación (2.5) para calcular las estimaciones de γ y σ , utilizando estimadores de $\widehat{\gamma z}$ y $\widehat{\sigma z}$, que detallamos a continuación.

Por la proposición 1 y (2.7) tenemos las siguientes estimaciones para $\widehat{\gamma z}$ y $\widehat{\sigma z}$:

- Para la distribución Lévy:

$$\widehat{\gamma_Z} = 2; \widehat{\sigma_Z} = \frac{4}{\pi(1-p)^2} \quad (3.1)$$

- Para la distribución Cauchy.

$$\widehat{\gamma_Z} = 1; \widehat{\sigma_Z} = \frac{1}{\pi(1-p)} \quad (3.2)$$

- Para la distribución Normal, la propiedad (2.7) no se encuentra verificada; sin embargo, se aproxima mediante el siguiente estimador, de acuerdo con Martín et al., 2022:

$$\widehat{\gamma_Z} = -0.7 + 0.61p; \widehat{\sigma_Z} = 0.34 + 3.18(1-p) - 12.4(1-p)^2 \quad (3.3)$$

Consecuentemente, el siguiente paso es estimar el parámetro de normalización a de la distribución estable X , por el método de distribución subyacente. Asimismo, como se mencionó con anterioridad, el parámetro b en la ecuación (2.2), de la distribución estable X , no tiene ninguna influencia en la estimación de los parámetros de la GPD, por lo que se considera $b = 0$.

- Distribución Lévy

La función de máxima verosimilitud queda determinada de la siguiente manera:

$$L(a | x) = \left(\frac{a}{2\pi}\right)^{\frac{n}{2}} \left[\prod_{i=1}^n x_i^{-\frac{3}{2}} \right] \exp \left\{ -\frac{a}{2} \sum_{i=1}^n \frac{1}{x^i} \right\}$$

Tomando como distribución *a priori* una distribución $\Gamma(a_0, b_0)$, y haciendo uso de (2.5) y (3.1), obtenemos las estimaciones de γ y σ .

- Distribución Cauchy

La función de máxima verosimilitud es:

$$L(a|x) = \pi^{-n} a^{-n} \prod_{i=1}^n \left[\left(\frac{x_i}{a} \right)^2 + 1 \right]^{-1}$$

En el mismo sentido, tomando como distribución *a priori* una distribución $\Gamma(a_0, b_0)$, y haciendo uso de (2.5) y (3.2), obtenemos las estimaciones de γ y σ .

- Distribución normal

En este caso la distribución *a priori* considerada, es una $inv\Gamma(a_0, b_0)$ para a^2 , y haciendo uso de (2.5) y (3.3), se obtienen las estimaciones de γ y σ .

Es necesario mencionar que, para ampliar el catálogo de distribuciones de referencia, la estrategia es encontrar dicha relación entre los parámetros de la distribución subyacente y los parámetros de la GPD de cola, a través de estudios de simulación. Por el momento, no existen estudios en la bibliografía sobre este tema, siendo interesante realizarlos en un futuro.

3.2.3 Informative Priors Baseline Metropolis Hastings

Finalmente, el ultimo método propuesto es Informative Priors Baseline Metropolis Hastings (IPBMH) que se caracteriza por proporcionar estimaciones más precisas que los métodos anteriores debido a que sus propiedades analíticas permiten aportar mayor peso a la cola de la GPD. En este sentido, se dice que la distribución *a priori* es altamente informativa por emplearse las estimaciones obtenidas en BMH. Simultáneamente, este método puede ser empleado para otras distribuciones subyacentes una vez obtenida la relación de parámetros con la GPD. Para las distribuciones estables seleccionadas, las distribuciones de prueba son:

$$\gamma \sim N(\gamma z, b_1); \sigma \sim N(\alpha \gamma z, b_2).$$

Estas estimaciones son las vistas en la sección anterior; a es estimado por BMH y, γz , σz se obtienen a través de 3.1 – 3.3. Además, los valores b_1 y b_2 quedan determinados de la siguiente manera (Martín et al., 2022):

- b_1 es constante con 0.03, 0.065 y 0.1 para la distribución subyacente Normal, Cauchy y Lévy respectivamente.
- Considerando los valores de la tabla 2 (ver tabla 2), el estimador de b_2 es: $b_2 = \exp\{c_1 p^2 + c_2 p + c_3\}$.

Tabla 2: Valores para c_1 , c_2 , c_3 para las tres distribuciones subyacentes.

Distribución	c_1	c_2	c_3
Lévy	500.2	-900.9	408.2
Cauchy	323.57	-588.51	266.13
Normal	-46.24	83.55	-41.58

Fuente: Martín *et al.*, (2022).

En este panorama, aplicamos Metropolis Hastings con las distribuciones *a priori* altamente informativas y de esta manera, se asegura una mejora en la precisión de las estimaciones, por lo que, la distribución *a posteriori* condicionada queda determinada por:

$$\pi(\gamma, \sigma | x_u) \propto \sigma^{-m} \exp \left\{ -\frac{1}{2b_1^2} (\gamma - \gamma z)^2 - \frac{1}{2b_2^2} (\sigma - a\sigma z)^2 \right\} \prod_{i=1}^m \left(1 + \gamma \frac{x_u^i}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}}$$

Y las distribuciones marginales:

$$\pi(\gamma | \sigma, x_u) \propto \exp \left\{ -\frac{1}{2b_1^2} (\gamma - \gamma z)^2 \right\} \prod_{i=1}^m \left(1 + \gamma \frac{x_u^i}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}}$$

$$\pi(\sigma | \gamma, x_u) \propto \sigma^{-m} \exp \left\{ -\frac{1}{2b_2^2} (\sigma - a\sigma z)^2 \right\} \prod_{i=1}^m \left(1 + \gamma \frac{x_u^i}{\sigma} \right)^{-\frac{(\gamma+1)}{\gamma}}$$

En este caso, el ratio de aceptación es determinado por:

$$\alpha_\gamma = \exp \left\{ \frac{1}{2b_1^2} \left((\gamma^{(j)} - \gamma z)^2 - (\sigma^* - \sigma z)^2 \right) - \frac{(\gamma+1)}{\gamma} \sum_{i=1}^m \ln \left(1 + \gamma^* \frac{x_u^i}{\sigma^{(j)}} \right) + \frac{(\gamma+1)}{\gamma} \sum_{i=1}^m \ln \left(1 + \gamma^{(j)} \frac{x_u^{i*}}{\sigma^{(j)}} \right) \right\}$$

$$\alpha_\sigma = \left(\frac{\sigma^{(j)}}{\sigma^*} \right)^m \exp \left\{ \frac{1}{2b_2^2} \left((\sigma^{(j)} - a\sigma z)^2 - (\sigma^* - a\sigma z)^2 \right) - \frac{(\gamma^{(j)}+1)}{\gamma^{(j)}} \sum_{i=1}^m \ln \left(1 + \gamma^{(j)} \frac{x_u^i}{\sigma^*} \right) + \frac{(\gamma^{(j)}+1)}{\gamma^{(j)}} \sum_{i=1}^m \ln \left(1 + \gamma^{(j)} \frac{x_u^i}{\sigma^{(j)}} \right) \right\}$$

CAPÍTULO IV: Aplicación metodológica

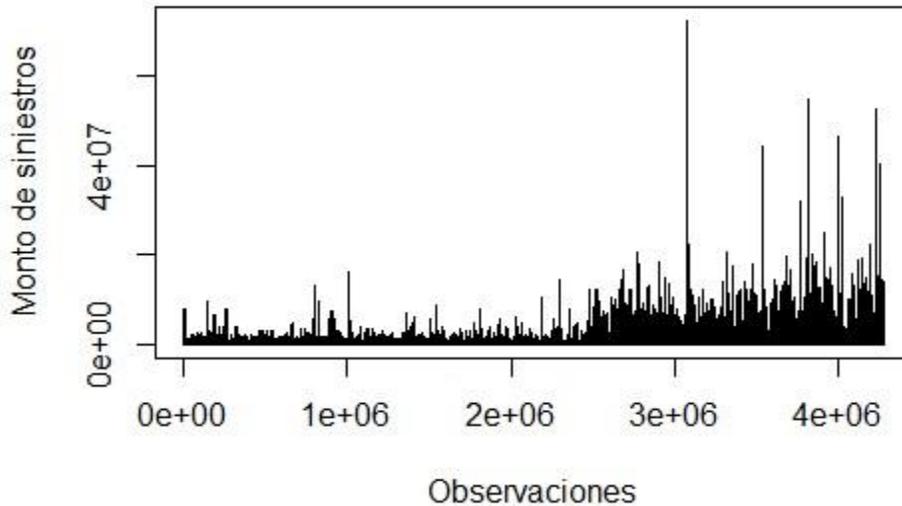
En la presente sección se realiza la aplicación a las metodologías planteadas en anteriores capítulos sobre los siniestros de un seguro de gastos médicos personales en México en el periodo 2009 – 2020 y, un estudio de simulación para comprobar la precisión de las estimaciones. De manera puntual comenzamos con un análisis estadístico para conocer el comportamiento natural de los datos.

4.1 Descripción de la base de datos

Los datos utilizados para la elaboración de los métodos Monte Carlo vía Cadenas de Markov (MCMC), corresponden a los siniestros ocurridos de un seguro de gastos médicos personales en México en el periodo 2009 – 2020, obtenidos de la Comisión Nacional de Seguros y Fianzas⁷, donde el costo por siniestro pagado por las instituciones aseguradoras corresponde a la suma de los montos generados por honorarios médicos, gastos hospitalarios, gastos por medicamentos, estudios auxiliares y otros gastos, menos el deducible y el coaseguro. Las cifras se manejan en moneda nacional y se obtiene un total de 4,283,724 observaciones. En cuanto al software utilizado es RStudio en su versión 4.2.2 (ver gráfica 8) (ver tabla 3):

⁷cnsf.gob.mx/EntidadesSupervisadas/InstitucionesSociedadesMutualistas/Paginas/AccidentesEnfermedades.aspx

Gráfica 8. Monto de siniestros del seguro de gastos médicos personales en México, 2009 – 2020.



Fuente: Elaboración propia.

Tabla 3. Estadística Descriptiva de los datos.

Observaciones	4,283,724	Media	49736.9
Mínimo	1	Desviación estándar	242435.9
Máximo	72562802	Coefficiente de asimetría	38.94293
Mediana	6751.59	Curtosis	5002.025

Fuente: Elaboración propia.

De acuerdo con la gráfica 8 y a la tabla 3, la media del conjunto de datos se ubica en \$49,736.9. La desviación estándar es de \$242,435.9, lo que indica, que hay gran dispersión de datos con respecto a la media muestral. El coeficiente de asimetría, tiene un valor positivo, por lo que, la distribución de los datos se encuentra sesgada a la derecha, es decir, hay más valores a la derecha de la media que a la izquierda y, en cuanto, al valor de curtosis, es de 5,002.025, lo que se interpreta como el hecho de que la distribución tiene más valores en las colas en comparación con una distribución normal.

Continuando con el análisis, las primeras pruebas a realizar son conocidas como “pruebas de normalidad”, las cuales son necesarias para saber si los datos estudiados siguen una distribución Gaussiana.

Pruebas de normalidad para el conjunto de datos

▪ Método Gráfico (gráfico de cuantiles teóricos Q-Q)

El Método gráfico Q-Q o “método de cuantiles”, es de utilidad para determinar si un conjunto de observaciones proviene de una variable con un modelo de probabilidad teórico específico. De acuerdo con Walpole et al., (2012) un cuantil de una muestra $q(f)$, es un valor para el que una fracción específica f de los valores de los datos es menor o igual a $q(f)$.

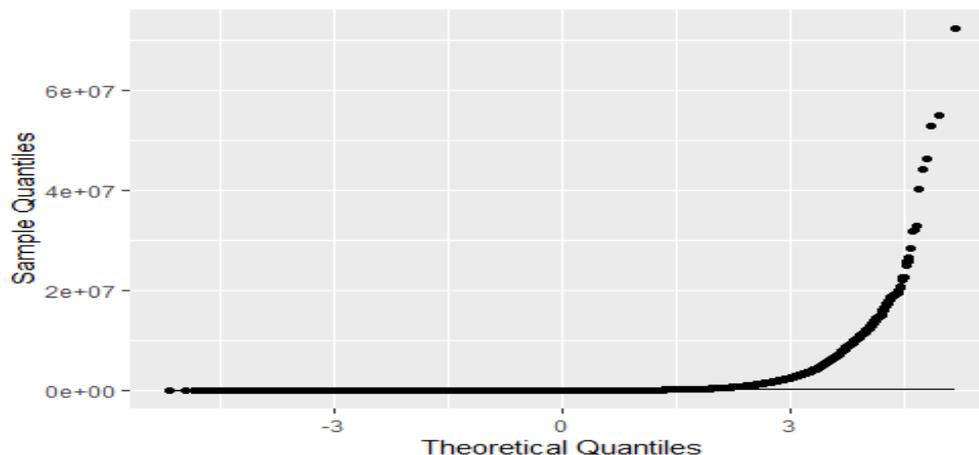
En tal sentido, las gráficas Q-Q nos permiten identificar similitudes visuales entre los valores de una variable aleatoria numérica y una evaluación empírica de la fracción de observaciones excedidas por los valores de los datos, o entre dos variables numéricas.

Para efectos prácticos del presente proyecto, esta fracción se calcula como:

$$f_i = \frac{i - \frac{3}{8}}{n + \frac{1}{4}}$$

donde i es el orden de las observaciones dadas de menor a mayor y, n el tamaño de la muestra. En la gráfica 8 se realiza la aplicación de cuantiles para determinar si nuestros datos siguen una distribución normal (ver gráfica 9).

Gráfica 9. *Cuantiles de la distribución normal para los siniestros de un seguro de gastos médicos personales en México, 2009-2020.*



Fuente: Elaboración propia.

De acuerdo con la gráfica 9, la muestra de los datos no se ajusta a una distribución normal, ya que, los siniestros extremos no siguen la recta, y, por lo tanto, denota otra distribución de probabilidad. Además, es necesario sustentar el argumento con las siguientes pruebas estadísticas.

- **Test de Kolmogorov–Smirnov (K-S) con corrección de Lilliefors**

Conocida como prueba *K-S*, es una prueba de significancia estadística no paramétrica que verifica si los datos de la muestra siguen un modelo con una distribución normal; además, es utilizada en la práctica para determinar si dos muestras de datos son iguales en distribución. En específico, contrasta la distribución de los datos con la distribución teórica y el resultado que obtenemos es el estadístico de *K-S* que se calcula con la siguiente fórmula:

$$T = \sup_x |F^*(x) - F_n(x)|$$

Donde $F^*(x)$ es la función de distribución teórica, mientras que $F_n(x)$ es la función de distribución empírica estimada a partir de las muestras aleatorias. En este contexto, se plantean las siguientes hipótesis:

$H_0: F(x) = F^*(x)$ para $-\infty \leq x \leq \infty$ (Los datos siguen una distribución normal).

$H_A: F(x) \neq F^*(x)$ para al menos un valor de x (Los datos no siguen una distribución normal).

Si T supera el cuantil $1 - \alpha$, como se indica en la tabla de cuantiles del estadístico *K-S* (véase anexo 1), rechazamos la H_0 al nivel de significancia α ; no obstante, de acuerdo con Steinskog, Tjøstheim, & Kvamastø (2007), esta prueba presenta algunas inconsistencias como la sensibilidad al tamaño de la muestra, por lo que, Lilliefors (1967), propuso la “corrección de Lillefors”, la cual utiliza la misma estadística de *K-S*, pero con valores críticos ajustados (véase anexo 2), lo cual lleva a conclusiones diferentes sobre la normalidad de los datos. Dada una muestra de observaciones, el estadístico Lilliefors se define como:

$$D = \max_x |F^*(X) - S_n(X)|$$

Donde $S_n(X)$ es la función de distribución de la muestra y $F^*(X)$ es la función de distribución acumulada normal con media muestral μ y desviación estándar σ . Realizando la prueba para el conjunto de datos, obtenemos los siguientes resultados:

$$D= 0.417$$

$$P\text{-value: } < 0.000000000000000022$$

Se debe agregar que el planteamiento de las hipótesis es el mencionado con anterioridad y, en general, en las pruebas realizadas. En este sentido, rechazamos la hipótesis nula e inferimos que nuestros datos no proceden de una distribución normal. A su vez, existen situaciones en donde las pruebas estadísticas pueden ser contradictorias, es decir, se puede rechazar la hipótesis nula en el test *K-S*, y, en *Lilliefors* no o, simplemente se desea verificar que las pruebas son adecuadas para el conjunto de datos obtenidos.

- **Prueba Jarque-Bera (J-B)**

Otra prueba estadística es Jarque-Bera (*J-B*) la cual se basa en las medidas clásicas de asimetría y curtosis de una distribución, que son el tercer y cuarto momento respectivamente y, de acuerdo con Carmona & Carrión (2015), su principal característica es que se utilizada para grandes muestras a partir de 2000. Sea cualquier distribución F con momentos centrales finitos, μ_k , donde $k \leq 3$, la asimetría se mide como:

$$\text{asimetría} = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = \frac{E(x - \bar{x})^3}{[E(x - \bar{x})^2]^{\frac{3}{2}}}$$

En cuanto a la curtosis, se define como: sea cualquier distribución F con momentos centrales finitos, μ_k , donde $k \leq 4$:

$$\text{curtosis} = \frac{\mu_4}{\mu_2^2} = \frac{E(x - \bar{x})^4}{[E(x - \bar{x})^2]^2}$$

En principio, la prueba de normalidad *J-B* utiliza los coeficientes de asimetría igual a 0, es decir se asume una distribución simétrica, y, curtosis igual a 3. Entonces, el estadístico se obtiene a través de:

$$JB = n \left[\frac{S^2}{6} + \frac{(K - 3)^2}{24} \right]$$

Donde n representa el tamaño de la muestra; S es el coeficiente de asimetría, y, K es el coeficiente de curtosis. Realizando la prueba $J-B$ para el conjunto de datos obtenemos los siguientes resultados:

JB= 4461600000000

P-value: < 0.000000000000000022

Lo que indica que evidencia a favor de la hipótesis alternativa, es decir, rechazamos la hipótesis nula e inferimos que nuestros datos no proceden de una distribución normal.

4.2 Métodos para definir umbrales

En el estudio de valores extremos, en cualesquiera de sus dos vertientes (método de bloques y picos sobre un umbral), la dificultad radica en fijar el número de valores máximos para cada modelo, ya que si este número es elevado se estarían analizando valores que no corresponden a la cola de la distribución. En caso de que el número sea pequeño se necesitaría una metodología capaz de modelar los datos de manera eficiente. En este sentido, existen diferentes alternativas en la elección del umbral, por ejemplo, suele escogerse un estadístico de orden o un cuantil determinado. Asimismo, en el ámbito de siniestralidad de seguros, es usual ver el método de “excesos medios” o “vida media residual”, debido a que su utilidad tiene relación directa con reaseguro, en donde se fija una prima para los siniestros que excedan dichas prioridades, es decir, el reaseguro cubre las cantidades $X - u$ condicionado a que $X > u$. En tal sentido, la definición de vida media residual es la siguiente.

Vida Media Residual

La metodología de vida media residual o de excesos medios (EM) es comúnmente utilizada para revelar el grosor de la cola de la distribución de probabilidad. La cuantía $e_{(u)}$

es el denominado valor medio en exceso del umbral u ; además, dejando el umbral como variable independiente y el valor de exceso medio como variable dependiente se obtiene, $e_F(u)$, la función media de excesos.

De acuerdo con Pérez & Tarango (2019), si X es una variable aleatoria continua no negativa, con media finita y con función de distribución $F(x)$, la función de excesos medios está dada por:

$$e_F(u) = E(X - u | X > u) = \frac{1}{1 - F(u)} \int_u^{\infty} (1 - F(y)) dy, \quad \text{con } u \in [0, u_k)$$

Donde $u_k = \sup \{u: F(u) < 1\}$. Además, una característica interesante que presenta la distribución de Pareto generalizada (GPD), es que su función de excesos medios está determinada por:

$$e_{(u)} = \frac{\sigma + \gamma u}{1 - \gamma}, \quad \text{donde } \gamma < 1,$$

Es decir, una función lineal de u con el gradiente $\frac{\gamma}{1-\gamma}$ y, el intercepto $\frac{\sigma}{1-\gamma}$. En este contexto, la función de excesos medios es el primer acercamiento a la delimitación de un umbral en el análisis de valores extremos mediante la GPD, por lo que, es necesario su representación, a partir de la muestra disponible y, que permitirá conocer la distribución de los datos.

Siguiendo a Embrechts, Klüppelberg, & Mikosch (1997), la función de excesos medios es estimada mediante:

$$\hat{e}_n(u) = \frac{1}{\text{card}\Delta_n(u)} \sum_{i \in \Delta_n(u)} (x_{(i)} - u)$$

Donde $x_{(i)}$ es la i -ésima estadística de orden de una muestra aleatoria de tamaño n ; $\Delta_n(u) = \{i = 1, \dots, n; x_i > u\}$ y $\text{card}\Delta_n(u)$ es la cardinalidad del conjunto $\Delta_n(u)$. De esto se desprende, que la función de excesos medios será obtenida a través de las coordenadas $(u, \hat{e}_n(u))$.

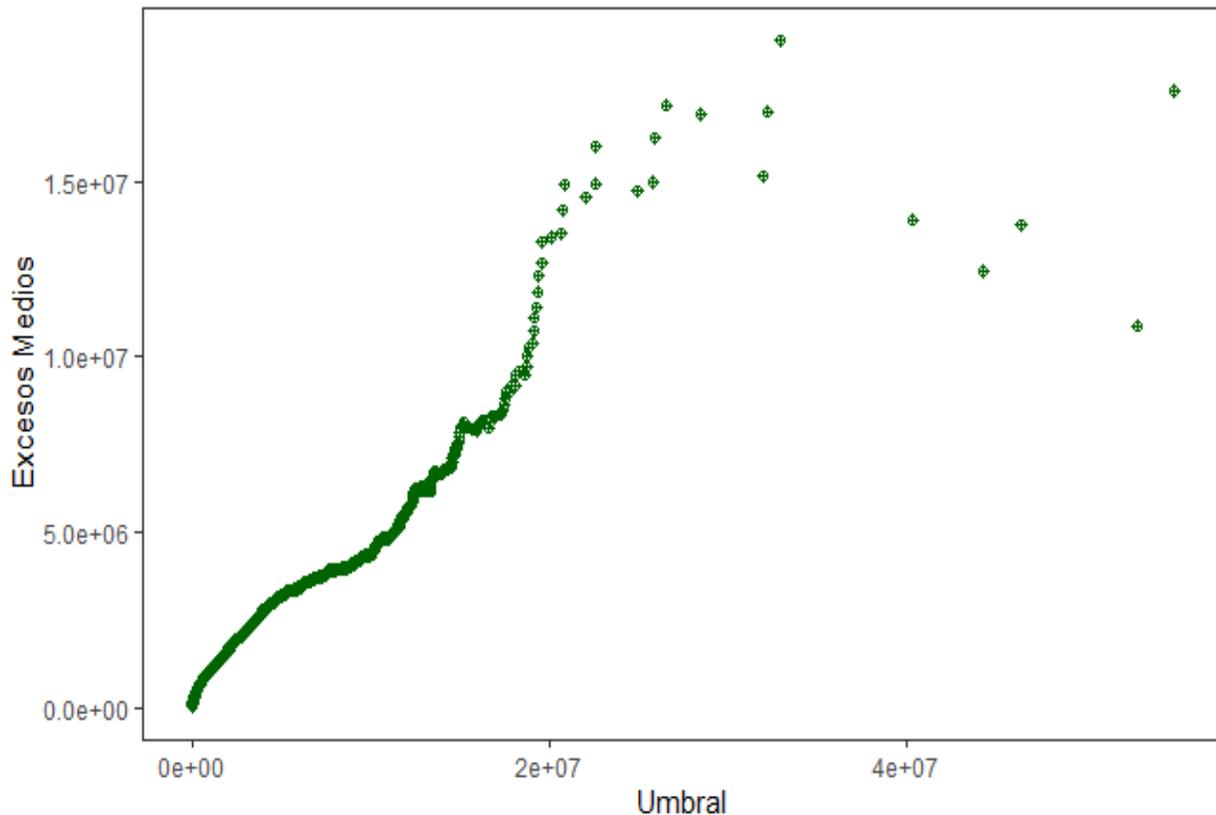
Formas de la función de excesos medios.

Por lo anteriormente mencionado, la gráfica de excesos medios presenta las siguientes características dependiendo la naturaleza de los datos:

- En el caso en donde la muestra de observaciones provenga de una distribución exponencial, que es considerada una distribución de colas ligeras, el gráfico de excesos medios mostrará una recta horizontal, debido a que la función de excesos es constante e independiente de u , siendo esta un caso especial de (EM) al contener esta característica.
- En el caso en donde se presenten distribuciones aún más ligeras que la distribución exponencial, como la distribución Weibull, la gráfica de EM mostrará una tendencia descendiente, lo que indica que, en algún momento determinado se presentará un punto final por la derecha y, por lo tanto, los valores extremos juegan un papel ínfimo en el análisis de valores extremos.
- Finalmente, cuando se presentan distribuciones de cola pesada, por ejemplo, la GPD o la distribución Lévy, la gráfica de EM será creciente conforme aumenta la prioridad (tendencia ascendente).

En este contexto, en la gráfica 10 se muestra la función de excesos medios para los siniestros de un seguro de gastos médicos personales en México en el periodo 2009 – 2020 (ver gráfica 10). Es importante mencionar que, si se considera un percentil al 95% como prioridad, para estos datos, se corresponde a un umbral de \$191,445.6, dejando por encima 214,187 observaciones, lo que indica que la distribución de probabilidad contiene valores extremos demasiado alejados de la media muestral y, se estaría considerando valores que no representan la cola de la distribución.

Gráfica 10. Excesos Medios para la siniestralidad del seguro de Gastos Médicos Personales en México, 2009 – 2020.



Fuente: Elaboración propia.

De acuerdo con la gráfica 10, la función de EM indica tendencia ascendente a medida que aumenta la prioridad o el valor del umbral, lo que quiere decir, distribución de cola pesada. Asimismo, se observa que a partir del umbral \$10,940,266.00, la gráfica muestra una tendencia lineal y pendiente positiva mucho más pronunciada, por lo que sería ideal considerar esta cantidad como prioridad. Asimismo, para efectos prácticos de la investigación, se selecciona \$14,569,008.71 como valor de umbral. En este sentido, el número de observaciones que cumplen ser mayores a las prioridades se muestran en la Tabla 4 (ver tabla 4).

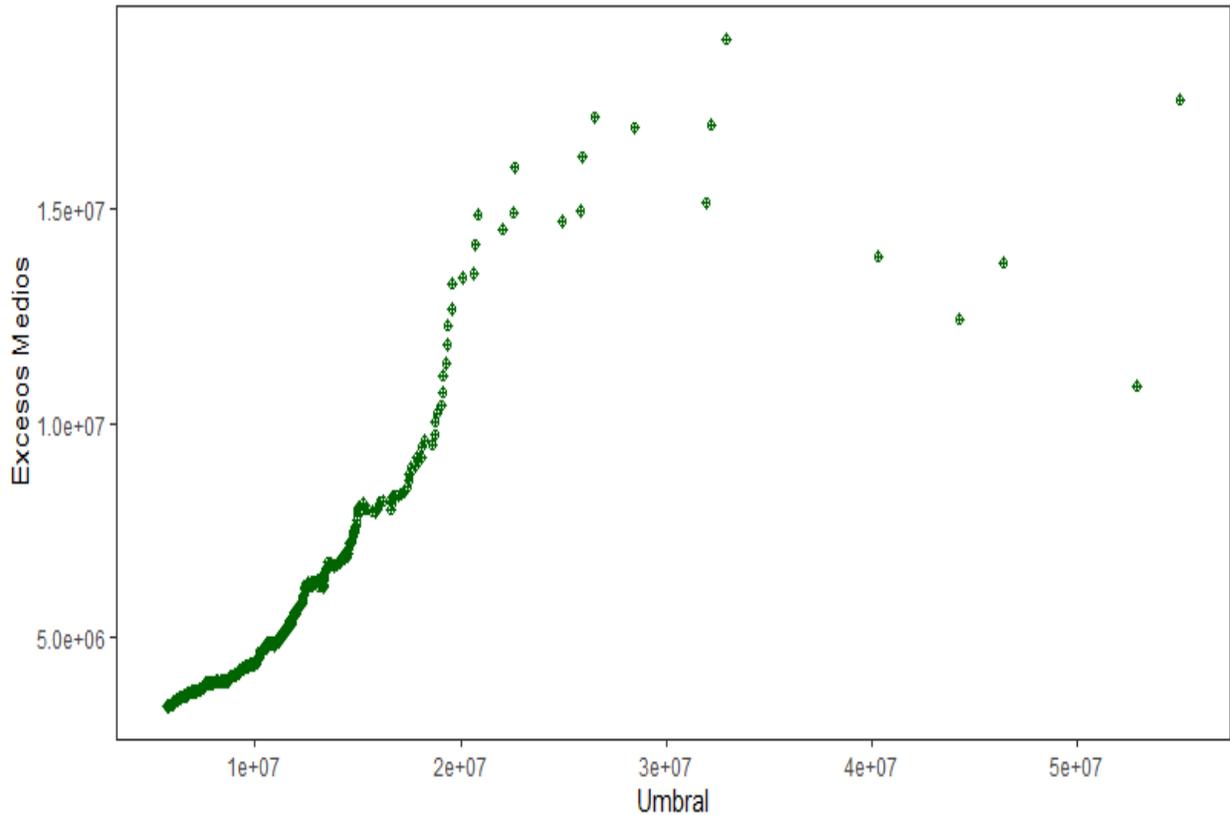
Tabla 4. Número de observaciones que superan el umbral.

Umbral	Observaciones
10,940,266	192
14,569,008.71	71

Nota: Elaboración propia.

Lo anterior mencionado, se puede observar a mayor detalle en la gráfica 11, en donde se realiza la gráfica de excesos medios para los siniestros de un seguro de gastos médicos personales en México, 2009 – 2020, considerando las últimas 1000 observaciones.

Gráfica 11. Excesos medios para los últimos 1000 siniestros registrados en el seguro de gastos médicos personales en México, 2009 – 2020.



Fuente: Elaboración propia.

4.3 Estudio de simulación

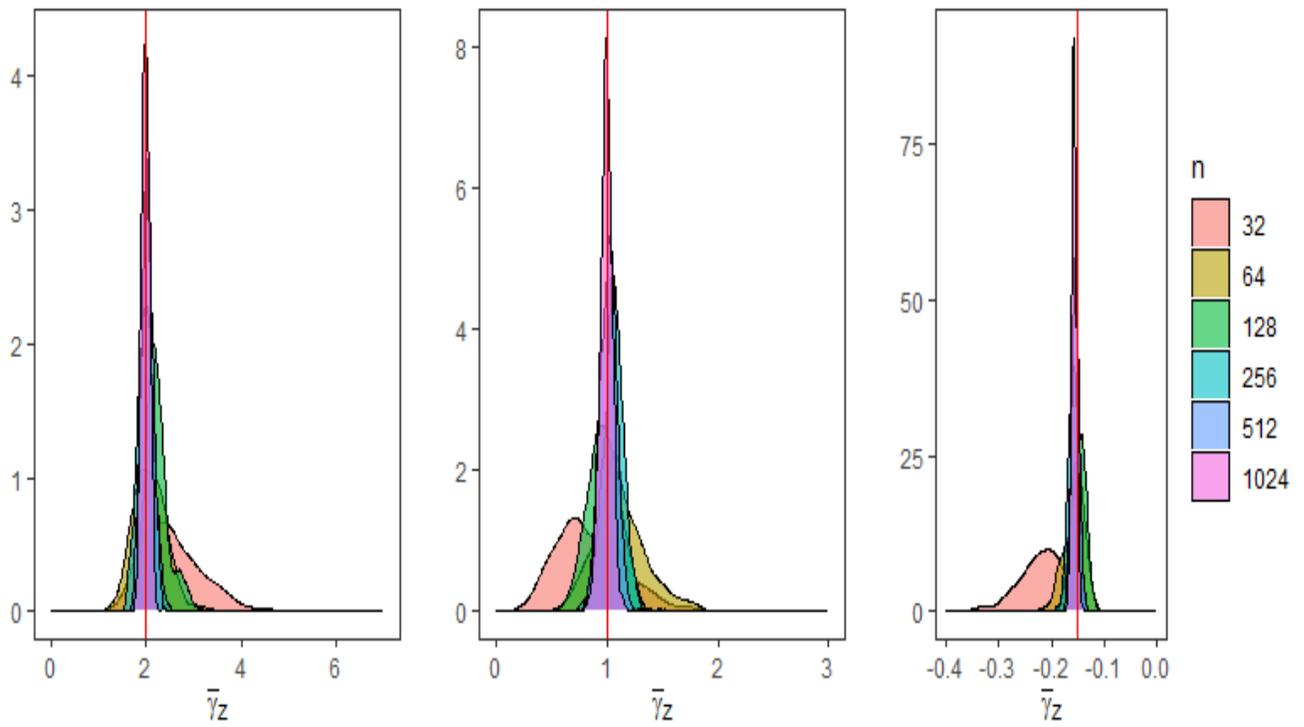
El presente apartado tiene como objetivo evaluar la precisión de las estimaciones por las metodologías propuestas en anteriores capítulos mediante un amplio estudio de simulación. En específico, las metodologías son: Metropolis Hastings (MH), Baseline Metropolis Hastings (BMH) e Informative Priors Baseline Metropolis Hastings (IPBMH). Es importante mencionar que existen otras metodologías MCMC, pero el objetivo también se centra en comparar IPBMH con el método clásico MH que son las dos metodologías que se aproximan a la Distribución Generalizada de Pareto (GPD).

Para garantizar la convergencia del algoritmo MCMC, debemos estar seguros de que se ha alcanzado la distribución *a posteriori* de los parámetros. Fijamos $p = 0.9$, y, el umbral $u = q_p$ y consideramos $b = 0$. El tamaño de la muestra es $n = 2^i$ con $i = 5, \dots, 10$ para las tres distribuciones seleccionadas, con escala $a = 1$. Además, realizamos los siguientes procedimientos para cada iteración:

- Periodo de rodaje (Burn In): Eliminar los primeros valores generados.
- Inicializar el algoritmo con diferentes valores iniciales y seleccionarlos para cada muestra.
- Adelgazamiento (thinning): Seleccionar una submuestra cada periodo para asegurar la ausencia de autocorrelación.

Estos procedimientos se realizan a través de la paquetería CODA de *R – Studio*; en específico: 1000 valores para el periodo Burn In, 25 valores para el thinning, y, se obtiene una Cadena de Markov de longitud 10,000 para cada parámetro. La secuencia se realiza 100 veces, y, se considera la media como estimador (ver gráfica 12).

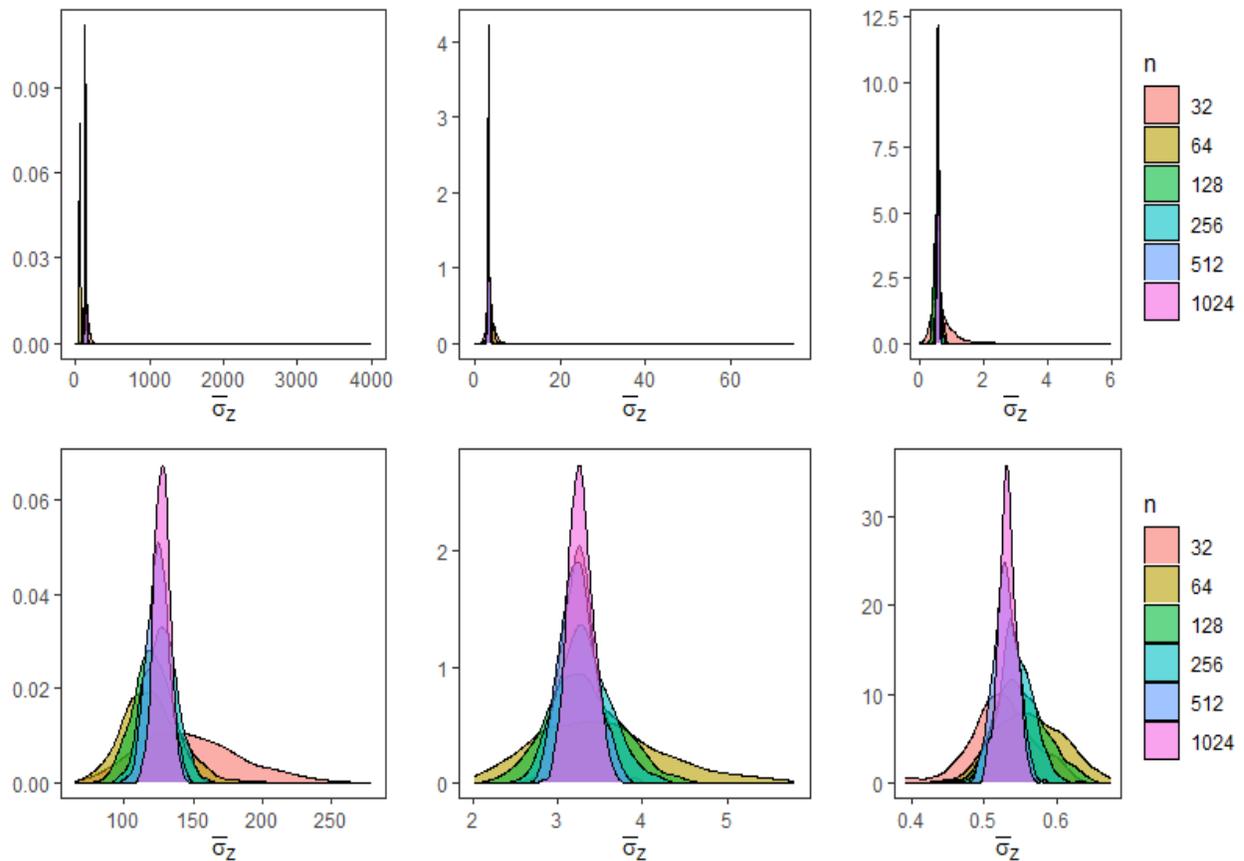
Gráfica 12. Distribución a posteriori de γz para diferentes valores de n , cuando la distribución base es $L(0,1)$ izquierda, $C(0,1)$ en medio y $Normal(0,1)$ derecha, por MH y BMH.



Fuente: Elaboración propia.

De acuerdo con la gráfica 12 (la estimación por BMH es la línea roja vertical), la primera distribución se encuentra sesgada a la derecha, y, a medida que n aumenta, la asimetría se reduce. En el caso de BMH ofrece una estimación puntual (ver gráfica 13).

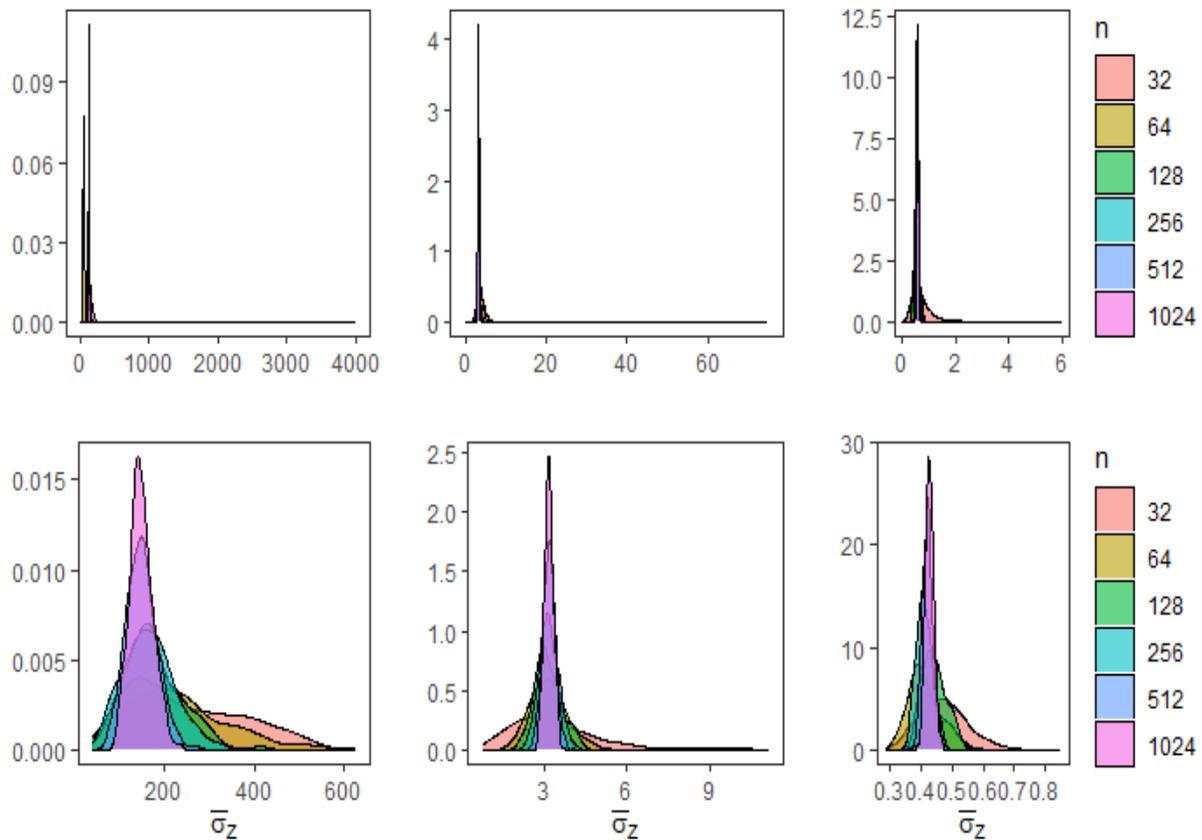
Gráfica 13. Distribución a posteriori de σ_z para diferentes valores de n , cuando la distribución base es $L(0,1)$ izquierda, $C(0,1)$ en medio y Normal $(0,1)$ derecha, por MH (gráficas superiores) y BMH (gráficas inferiores).



Fuente: Elaboración propia.

En la gráfica 13 se muestran las estimaciones por ambos métodos, MH y BMH. Se puede observar que BMH otorga distribuciones *a posteriori* menos sesgadas cuando n es pequeño a diferencia de MH, que se dificulta llegar a las estimaciones deseadas (ver gráfica 14).

Gráfica 14. Distribución a posteriori de σ_z para diferentes valores de n , cuando la distribución base es $L(0,1)$ izquierda, $C(0,1)$ en medio y Normal $(0,1)$ derecha, por MH (gráficas superiores) e IPBMH (gráficas inferiores).



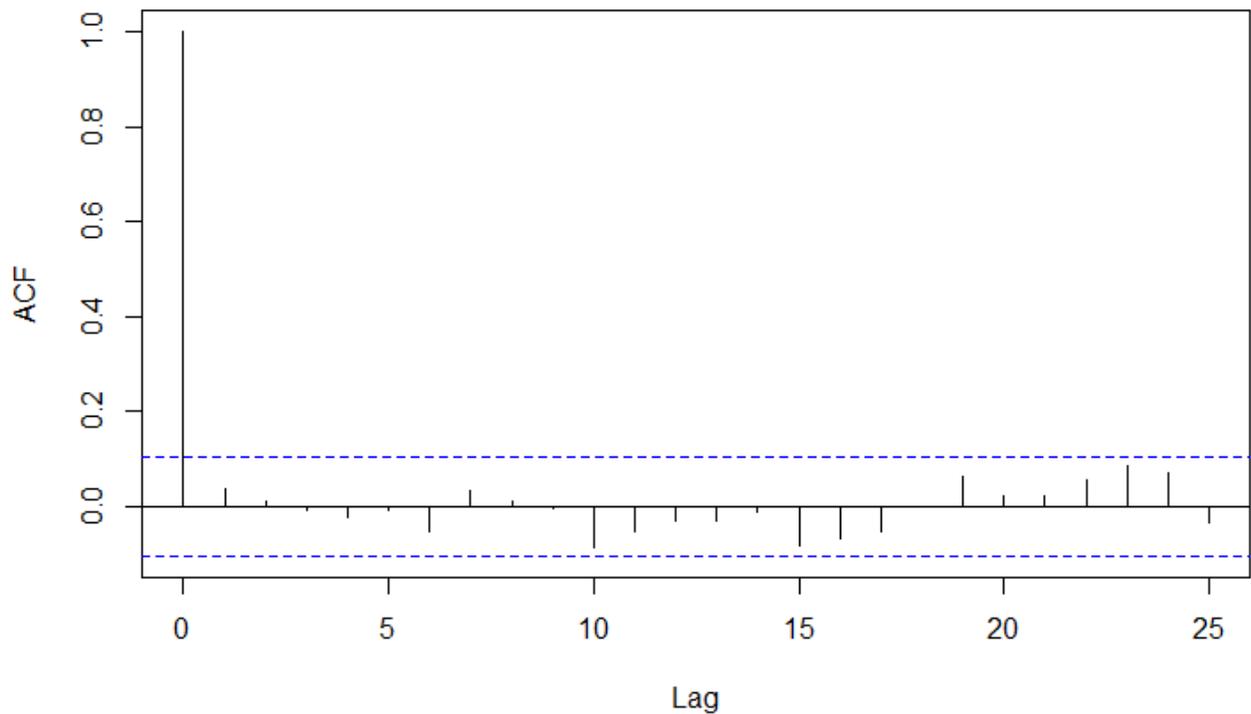
Fuente: Elaboración propia.

Finalmente, en la gráfica 14 se presentan las distribuciones *a posteriori* para el parámetro $\bar{\sigma}_z$ por los dos métodos que se aproximan a la distribución Generalizada de Pareto, MHM e IPBMH. Se puede observar que IPBMH otorga distribuciones *a posteriori* menos sesgadas cuando el número de observaciones es pequeño y, cuando el número de observaciones son elevadas, ambos métodos tienen buen desempeño, como cuando la distribución base es Cauchy. En este sentido, se considera la media como estimador, ya que las distribuciones *a posteriori* son aproximadamente normales.

Adicionalmente, independientemente de la distribución seleccionada y el método empleado, el adelgazamiento de la cadena de Markov permite la ausencia de autocorrelación, es decir, se considera que la cadena contiene un conjunto de muestras

independientes e idénticamente distribuidas de la distribución *a posteriori*. En este sentido, una vez que se aplicó el adelgazamiento y el periodo de rodaje, se realiza la instrucción ACF en R Studio y, se obtiene la siguiente gráfica (ver gráfica 15).

Gráfica 15. Autocorrelación de la cadena de Markov, cuando la distribución base es $C(0,1)$, MH.



Fuente: Elaboración propia.

De acuerdo con la gráfica 15, se obtiene una convergencia favorable en la cadena de Markov de los procesos realizados (El rezago 0, su valor siempre es 1). Continuando con nuestro razonamiento, con el fin de evaluar la precisión de las estimaciones, realizamos la medida del “error absoluto medio” (MAE) y el “error cuadrático medio” (RMSE), donde $\hat{\theta}_i$ es la estimación obtenida a partir de la muestra, y θ es el valor real del parámetro estimado (ver tabla 5 - 7).

$$MAE = \frac{1}{M} \sum_{i=1}^M |\hat{\theta}_i - \theta|$$

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{\theta}_i - \theta)^2}$$

Donde M es el número de muestras.

Tabla 5. MAE y RMSE de MH, BMH e IPBMH con $a = 1$, cuando la distribución base es $L(0, 1)$.

	n	MHM	BMH	IPBMH
MAE	32	0.0023	0.0021	0.0022
RMSE		0.003	0.0029	0.0029
MAE	128	0.0023	0.0024	0.0021
RMSE		0.0031	0.0032	0.0028
MAE	512	0.0025	0.0024	0.0022
RMSE		0.0032	0.0032	0.003

Fuente: Elaboración propia.

Tabla 6. MAE y RMSE de MH, BMH e IPBMH con $a = 1$, cuando la distribución base es $C(0, 1)$.

	n	MHM	BMH	IPBMH
MAE	32	0.1	0.1046	0.099
RMSE		0.1254	0.1326	0.1267
MAE	128	0.1065	0.1037	0.1041
RMSE		0.134	0.131	0.132
MAE	512	0.1052	0.1024	0.1064
RMSE		0.1329	0.1301	0.1335

Fuente: Elaboración propia.

Tabla 7. MAE y RMSE para MH, BMH e IPBMH con $\alpha = 1$, cuando la distribución base es $N(0,1)$.

	n	MHM	BMH	IPBMH
MAE	32	0.6661	0.6183	0.7195
RMSE		0.7988	0.7578	0.8801
MAE	128	0.6139	0.6011	0.6955
RMSE		0.7523	0.7373	0.8488
MAE	512	0.586	0.6006	0.6571
RMSE		0.7185	0.7377	0.802

Funete: Elaboración propia.

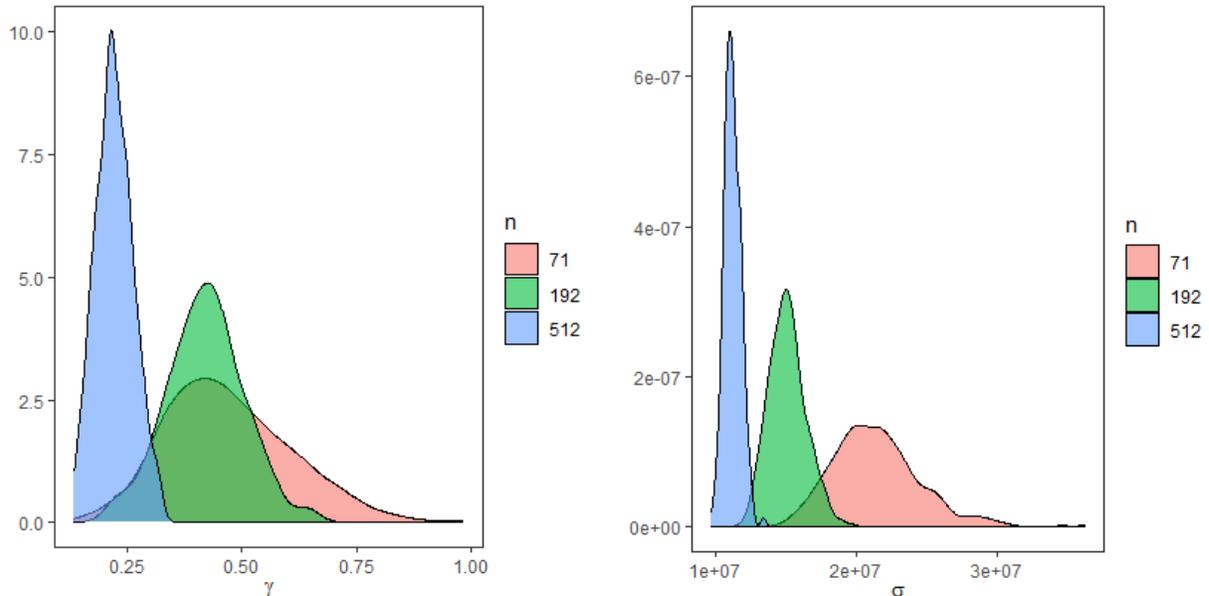
De acuerdo con las tablas 5 – 7, podemos inferir que la metodología IPBMH proporciona estimaciones entre MHM y BMHM, cuando el índice de estabilidad de la distribución estable es menor a 2 sin descartar que BMHM, lo que estima, son los parámetros de la distribución base, a diferencia de MHM e IPBMH que aproximan a la GPD. De acuerdo Martín et al. (2022), en la práctica los datos podrían ajustarse mejor a una mezcla de distribuciones mediante el uso de las distribuciones *a priori* altamente informativas construidas por la información del conjunto total de los datos, por lo que, IPBMH proporciona mejores aproximaciones a la GPD cuando los datos difieren de las distribuciones simples.

4.4 Aplicación para la siniestralidad extrema de un seguro de Gastos Médicos Personales en México, 2009 – 2020.

Para el caso práctico se considera los umbrales establecidos por excesos medios; además, un estadístico de orden superior de 512 valores que corresponde a \$7,637,124.00 como umbral, con el fin de evaluar diferentes escenarios en la estimación de parámetros de la GPD. En este sentido, es importante mencionar que nuestros datos no se ajustan a una distribución simple, por lo que, para el presente proyecto se considera la metodología MH para el comportamiento futuro de los siniestros extremos de un seguro de gastos médicos personales en México.

De igual modo, como lo hicimos en la sección anterior, obtenemos las siguientes distribuciones *a posteriori* para los parámetros de la GPD como se muestra en la gráfica 16 (ver gráfica 16).

Gráfica 16. Distribución *a posteriori* de la GPD para diferentes valores de n , para el seguro de gastos médicos personales en México, 2009 - 2020, MH.



Fuente: Elaboración propia.

Como se puede observar en la gráfica 16, para el umbral \$7,637,124.00 obtenemos una distribución poco variable, ya que, como lo vimos en excesos medios, se están considerando valores menos atípicos de los reales; sin embargo, para el umbral \$10,940,266.00 el comportamiento es adecuado ya que no existe mucha variabilidad y se representan todos los valores que corresponden a la cola de la distribución de probabilidad. En el caso del umbral \$14,569,008.71, la distribución *a posteriori* muestra demasiada variabilidad y sesgo como se mostró en el estudio de simulación y, aunque la estimación del parámetro de forma es similar a la estimación cuando se consideran 512 observaciones, la incertidumbre es mayor (distribución expandida horizontalmente). Continuando con el análisis, no se pudo realizar el MAE y el RMSE con el valor real de la distribución, por lo que se realiza una réplica Monte Carlo de longitud 10,000 utilizando las estimaciones por Máxima Verosimilitud para aseverar la precisión de las estimaciones (ver tabla 8).

Tabla 8. MAE y RMSE para la GPD mediante ML y MH.

GPD		
	n	MH
MAE	512	0.00000012
RMSE		0.00000015
MAE	192	0.00000012
RMSE		0.00000015
MAE	71	0.00000008
RMSE		0.00000001

Fuente: Elaboración propia.

De acuerdo con la tabla 8, las estimaciones mediante MH no difieren de las estimaciones por ML casi en su totalidad para los tres umbrales elegidos. Es importante mencionar que sería interesante realizar las estimaciones de la GPD para los siniestros de un seguro de gastos médicos personales en México mediante mixtura de distribuciones y el uso de las distribuciones *a priori* altamente informativas en futuras investigaciones.

Asimismo, con base al estudio realizado, el estimador bajo la teoría bayesiana, no muestra un mejor comportamiento a diferencia de la estimación clásica de valores extremos; sin embargo, mediante la convergencia de cadenas de Markov, obtenemos la distribución *a posteriori* del parámetro, y, no solo una estimación, por lo que, en muchas situaciones es mejor contar con información adicional; además, salva el problema de las derivadas parciales de Máxima Verosimilitud que en ocasiones puede presentar problemas de convergencia.

Uno de los principales problemas de la teoría de valores extremos es que se dispone de poca información en cualesquiera de sus dos vertientes (excesos de un umbral y máximo de bloques), por lo que, una alternativa para suavizar el problema es extraer información del conjunto total de los datos. En este sentido, en el presente proyecto se muestra la relación de parámetros de las distribuciones estables Normal, Cauchy y Lévy con la distribución límite GPD.

Asimismo, independientemente de los métodos seleccionados, las distribuciones *a posteriori* muestran menor sesgo y variabilidad a medida que el número de observaciones aumenta. En el caso cuando disponemos de poca información, BMH e IPBMH proporcionan distribuciones *a posteriori* menos sesgadas y más concentradas a diferencia de MH. Cuando el número de observaciones es grande, la cadena de Markov converge de manera inmediata y la probabilidad máxima se vuelve más prominente.

Continuando con nuestro razonamiento, considerando diferentes valores iniciales para los métodos propuestos y variando el número de umbrales, el porcentaje de aceptación de la cadena de Markov es variable, por ejemplo, mediante el uso de las distribuciones *a priori* poco informativas, se obtiene un porcentaje de aceptación entre 27% - 55%, adecuado para la convergencia de la cadena de Markov; sin embargo, con IPBMH, el porcentaje de aceptación aumenta, debido a que la mayor parte del tiempo se encuentra en el espacio paramétrico de la distribución de probabilidad.

Además, la propuesta de excesos medios al delimitar un umbral muestra un comportamiento relativamente bueno, ya que, en el caso de 192 observaciones, las distribuciones *a posteriori* fueron aproximadamente normales, para ambos parámetros sin encontrar demasiada variabilidad y sesgo como en el caso de 71 observaciones o de 512 donde se consideran valores más cercanos a la media de la distribución.

Es importante mencionar que en la práctica el uso de las distribuciones *a priori*, en la estimación bayesiana, representa un reto, ya que el mal uso de estas llevaría al análisis de un espacio paramétrico no perteneciente a la distribución. En el presente proyecto para el caso de MH, se utilizan distribuciones *a priori* poco informativas y mediante la relación de parámetros de las distribuciones estables y la distribución límite de la GPD se construyen las distribuciones *a priori* altamente informativas.

Inferencia y predicción

Continuando con nuestro análisis, las distribuciones *a posteriori* obtenidas para la siniestralidad extrema de un seguro de gastos médicos personales, mediante procesos de simulación, son aproximadamente normales, por lo que se considera la media como un buen estimador. Una ventaja de contener la distribución *a posteriori* de los parámetros

estudiados, es la posibilidad de hacer inferencia bayesiana mediante un intervalo de credibilidad, que se considera ideal cuando existe holgura en la toma de decisión. En este sentido, en las tablas 9 y 10 se muestra una credibilidad entre 0.025 y 0.975 para diferentes valores de umbral (ver tablas 9 y 10).

Tabla 9. *Intervalo de credibilidad para el parámetro γ , de la GPD cuando se tiene diferente valor de n .*

n	2.50%	97.50%
512	0.150746	0.308242
192	0.243467	0.58588
71	0.23845	0.766486

Fuente: Elaboración propia.

Tabla 10. *Intervalo de credibilidad para el parámetro σ , de la GPD cuando se tiene diferente valor de n .*

n	2.50%	97.50%
512	10162521	12401118
192	12887731	17803256
71	16393389	28603405

Fuente: Elaboración propia.

De acuerdo con las tablas 9 y 10, cualquier valor que se encuentre entre los intervalos de credibilidad establecidos se considera una correcta estimación para el modelo de valores extremos; sin embargo, en esta investigación es de interés evaluar la estimación puntual para el caso en donde se tienen 192 observaciones (\$10,940,266.00 de umbral) en un seguro de gastos médicos personales en México, por lo que, tenemos los siguientes valores de parámetros: $\gamma = 0.42138690$ y $\sigma = 15105594$. Asimismo, de acuerdo al análisis de los datos realizado, se observa un comportamiento anormal a partir de un valor de 16,000,000. En este sentido, evaluamos en la función de distribución de la GPD.

$$G(x | 0.4213869, 15,105,594) = 1 - \left(1 + \frac{0.4213869 * x}{15105594}\right)^{-\frac{1}{0.4213869}}$$

$$\Rightarrow G(16000000 | 0.4213869, 15,105,594) = 0.5834553$$

Dicho de otra manera, el resultado se interpreta como que, la probabilidad de experimentar un siniestro menor o igual a \$16,000,000.00 dado que se ha superado una cantidad de \$10,940,266.00, es de 58%. Lo anterior, es lo mismo que evaluar su complemento, es decir, la probabilidad de que un evento de magnitud catastrófica exceda \$16,000,000.00, dado que ha superado \$10,940,266.00, es de 42% o en términos actuariales, la probabilidad de ruina de una institución aseguradora dado que se ha materializado un siniestro mayor a \$10,940,266.00 es de 42%, por lo que, se sugiere hacer uso de contratos de reaseguro.

Otra aplicación es la respuesta a los siguientes planteamientos.

¿Cuál es la probabilidad de experimentar un evento de magnitud \$14,569,008.71?

$$g(14,569,008.71 | 0.4213869, 15,105,594)$$

$$= 0.4213869^{-1} \left(1 + \frac{14596008.71 * 0.4213869}{15105594}\right)^{-\frac{(0.4213869+1)}{0.4213869}}$$

$$\Rightarrow g(14,569,008.71 | 0.4213869, 15,105,594) = 0.0000000210$$

¿Cuál es la probabilidad de experimentar un siniestro de magnitud entre \$11,000,000.00 y \$17,000,000.00?

$$P(11,000,000 \leq x \leq 17,000,000)$$

$$= G(17,000,000 | 0.4213869, 15,105,594)$$

$$- G(11,000,000 | 0.4213869, 15,105,594)$$

$$\Rightarrow P(11,000,000 \leq x \leq 17,000,000) = 0.1317976$$

Estas y muchas aplicaciones se pueden llevar a cabo una vez obtenidos los parámetros de la GPD. Por lo que, se espera que el presente proyecto sea de utilidad para futuras investigaciones dentro del gremio actuarial y del sector en general, debido a que

es un área muy dinámica que goza diversas actualizaciones en temas investigación de metodologías.

En este sentido, los siniestros del seguro de gastos médicos personales en México para el periodo 2009 – 2020, pueden ser modelados mediante la metodología empleada, en donde el proceso de rechazar o no una propuesta, solo depende del valor inmediato anterior, debido a que, al llevar a cabo un determinado número de iteraciones, la cadena de Markov converge a una distribución *a posteriori*.

CONCLUSIONES GENERALES

El presente proyecto fue redactado ante la necesidad de cuantificar el riesgo que se genera al asumir elevadas cuantías en instituciones aseguradoras para un seguro de gastos médicos personales en México en el periodo 2009 - 2020, mediante procesos de simulación.

En este sentido, después de conocer el comportamiento natural de los datos, y al aplicar las metodologías planteadas, se rechaza la hipótesis nula, es decir, se infiere que la simulación Monte Carlo vía Cadenas de Markov (MCMC), no mejora la estimación de parámetros en una distribución de valores extremos comparado con los métodos clásicos probabilísticos, por ejemplo, el método de máxima verosimilitud; sin embargo, las metodologías propuestas, presentan la ventaja de proporcionar información adicional a través de la distribución *a posteriori* de los parámetros estudiados y, no solo una estimación; además, si el número de observaciones es pequeño, el uso de distribuciones *a priori* altamente informativas otorga mejores aproximaciones.

Asimismo, es importante mencionar que los objetivos de la presente investigación se cumplieron satisfactoriamente, ya que, las metodologías propuestas permitieron simular la siniestralidad máxima de un seguro de gastos médicos personales en México en el periodo 2009 – 2020; en específico, a través de Metropolis Hastings por lo que sería interesante realizar en futuras investigaciones la mixtura de distribuciones de probabilidad del conjunto total de datos, haciendo uso de las distribuciones *a priori* altamente informativas.

En el mismo sentido, la propuesta de excesos medios al delimitar un umbral para la Distribución de Pareto Generalizada (GPD), muestra un comportamiento relativamente bueno, ya que, permite identificar en qué momento, la distribución de probabilidad hace presencia de valores atípicos y, de esta manera tomar una decisión adecuada en la delimitación de un umbral. En este sentido, si el número de observaciones son escasos, las distribuciones *a posteriori* muestran demasiada variabilidad y sesgo; en el caso en donde se analizan datos que no corresponden a la cola de la distribución, se pierde el enfoque de analizar los valores graves para el sector.

Es importante mencionar que en la práctica el uso de las distribuciones *a priori*, en la estimación bayesiana, representa un reto, ya que el mal uso de estas llevaría al análisis de un espacio paramétrico no perteneciente a la distribución de probabilidad. En el presente proyecto para el caso de MH, se utilizan distribuciones *a priori* poco informativas, y, mediante la relación de parámetros de las distribuciones estables y la GPD, se construyen las distribuciones *a priori* altamente informativas.

Finalmente, considerando diferentes valores iniciales por los métodos propuestos y, variando el número de observaciones, el porcentaje de aceptación de la cadena de Markov fue variable por cada método, es decir, mediante el uso de las distribuciones *a priori* poco informativas, se obtuvo un porcentaje de aceptación entre 27% - 55%, adecuado para la convergencia de la cadena de Markov; sin embargo, con el uso de distribuciones *a priori* altamente informativas el porcentaje de aceptación aumenta, debido a que la cadena de Markov se encuentra la mayor parte del tiempo en el espacio paramétrico estudiado, lo que ayuda a explorar áreas difíciles de alcanzar, construyendo así la distribución *a posteriori* ajustada.

REFERENCIAS

- Acero, F., García, J., & Gallego, M. (2011). Peaks-over-Threshold Study of Trends in Extreme Rainfall over the Iberian Peninsula. *Journal of Climate*, 24(4), 1089-1105. <https://journals.ametsoc.org/view/journals/clim/24/4/2010jcli3627.1.xml>
- AMIS. (2020). *Centro estadístico del Sector Asegurador*. AMIS: <https://centroestadisticoamis.mx/>
- Barbosa, R., & Llinás, H. (2016). *Procesos estocásticos con aplicaciones*. Universidad del Norte.
- Bautista, P. (2007). *Experimentación con un algoritmo de MCMC multiescala y autoajustable [Tesis de Licenciatura, CIMAT]*. Repositorio Institucional. <https://ciamat.repositorioinstitucional.mx/jspui/bitstream/1008/81/2/TE%20238.pdf>
- Bedoya, J., & Barrera, M. (2006). Convergencia de las cadenas de markov. *Scientia et Technica*, 12(32). <https://www.redalyc.org/articulo.oa?id=84911652014>
- Blanco, S., & Carvajal, P. (01 de 02 de 2001). *Antecedentes del seguro a prima*. Researchgate: https://www.researchgate.net/publication/24140184_Antecedentes_del_seguro_a_prima
- Brazauskas, V., & Kleefeld, A. (2009). Robust and Efficient Fitting of the Generalized Pareto Distribution with Actuarial Applications in View. *Insurance: Mathematics and Economics*, 45 (3), 424-435.
- Calva, G. (2016). *Cadenas de Markov reversibles: Algunas aplicaciones [Tesis de Licenciatura, UCE]*. Repositorio Institucional. <http://www.dspace.uce.edu.ec/handle/25000/8124>
- Cámara de diputados del H. Congreso de la Unión. (2013, 04 de abril). *Ley de Instituciones de Seguros y Fianzas*. Diario Oficial de la Federación. https://www.senado.gob.mx/comisiones/finanzas_publicas/docs/LISF.pdf

- Cardarello, M., & Luraghi, L. (2019). *Análisis de valores extremos: Una aplicación a temperaturas mínimas en Uruguay [Informe de pasantía, Udelar]*. Repositorio Institucional. <https://hdl.handle.net/20.500.12008/31396>
- Carmona, M., & Carrión, H. (2015). *Potencia de prueba estadística de normalidad Jarque-Bera frente a las pruebas Anderson Darling, Jarque Bera robusta, chi-cuadrada, chen-shapiro y shapiro-wilk [Tesis de Licenciatura, UAEMex]*. Repositorio Institucional. <https://ri.uaemex.mx/handle/20.500.11799/94337>
- Castillo, A. L. (2014). *Modelo de Regresión y Estadística, para la siniestralidad de autos de una empresa, dentro de la vigencia 2012 - 2013 [Tesis de Licenciatura, UAEMex]*. Repositorio Institucional. <https://ri.uaemex.mx/handle/20.500.11799/68039>
- Castillo, E., Hadi, A., Balakrishnan, N., & Sarabia, J. (2004). *Extreme Value and Related Models with Applications in Engineering and Science*. Hoboken: WILEY.
- Cid, A. (2000). *Siniestralidad en Seguros no vida: Provisión para prestaciones. Un nuevo método [Tesis doctoral, Universidad Complutense de Madrid]*. Repositorio institucional. <https://eprints.ucm.es/id/eprint/3642/1/T24021.pdf>
- CNSF. (31 de 12 de 2020). *Boletín de Análisis Sectorial*. Gobierno de México: <https://www.gob.mx/cnsf/documentos/boletin-de-analisis-sectorial-23209>
- CNSF. (2020). *Información Consolidada del Sector Asegurador*. Gobierno de México: <https://www.cnsf.gob.mx/EntidadesSupervisadas/InstitucionesSociedadesMutualistas/Paginas/InformacionConsolidada.aspx>
- Coelho, C., Ferro, C., Stephenson, D., & Steinskog, D. (2008). Methods for Exploring Spatial and Temporal Variability of Extreme Events in Climate Data. *Journal of Climate*, 21(10), 2072-2092. https://journals.ametsoc.org/view/journals/clim/21/10/2007jcli1781.1.xml?tab_body=pdf
- Consorti, A. M. (08 de 2005). *Crónicas de dos siglos del Seguro en México*. Gobierno de México.

https://www.gob.mx/cms/uploads/attachment/file/74253/Cronica_de_dos_siglos_d_el_seguro_en_Mexico.pdf

Delgado, M. (2018). *¿Cómo funciona el Reaseguro?* Rastreador.mx: <https://www.rastreador.mx/articulos-destacados/caracteristicas-del-reaseguro>

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer Science & Business Media.

Fisher, R., & Tippett, L. (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proceedings Cambridge Philosophical Society*, 24, 180-190.

García, J., Gallejo, M., Serrano, A., & Vaquero, J. (2007). Trends in Block-Seasonal Extreme Rainfall over the Iberian Peninsula in the Second Half of the Twentieth Century. *Journal of Climate*, 20(1), 113-130. <https://journals.ametsoc.org/view/journals/clim/20/1/jcli3995.1.xml>

Gerber, H. U. (1979). *An Introduction to Mathematical Risk Theory*. S.S. Huebner Foundation, University of Pennsylvania, Philadelphia.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44, 423-453.

González, E. (2007). *Pruebas de Bondad de Ajuste para distribuciones estables [Tesis de Doctorado, CP]*. Repositorio Institucional. <http://colposdigital.colpos.mx:8080/xmlui/handle/10521/1536>

González, G. (2003). *Manual Básico del Seguro*. Asunción: Intercontinental. Calameo.

Gran Enciclopedia del Mundo. (1978). *Antecedentes históricos del seguro*. Durvan S.A. Bilbao.

IMSS. (1943). *Conoce al IMSS*. Gobierno de México: <http://www.imss.gob.mx/>

ISSEMyM. (2021). *Antecedentes de la Seguridad Social*. Gobierno del Estado de México. <https://www.issemym.gob.mx/node/19>

- ISSSTE. (2020). *60 años de actividad ininterrumpida*. Gobierno de México. <https://www.gob.mx/issste/es/articulos/60-anos-de-actividad-ininterrumpida?idiom=es#:~:text=El%2030%20de%20diciembre%20de,se%20encargaba%20de%20las%20prestaciones>.
- Johnson, B. (01 de 07 de 2020). *Great Fire of London 1666*. Historic UK. <https://www.historic-uk.com/HistoryUK/HistoryofEngland/The-Great-Fire-of-London/>
- Jones, H. E. (1998). *Principios del Seguro de Vida, Salud y Anualidades*. Atlanta: LOMA.
- Loredana, A. (2005). *Análisis Jurídico de la Intermediación del Contrato de Seguro [Tesis de Licenciatura, Universidad de las Américas Puebla]*. Repositorio Institucional. http://catarina.udlap.mx/u_dl_a/tales/documentos/ledf/specia_j_al/
- Lozano, N., & Morgado, E. (1997). *Instituciones de Derecho del trabajo y de la seguridad social*. México: AIADTSS.
- Markham, P. (1962). An introduction to collective risk theory and its application to stop-loss reinsurance. *Transactions of Society of Actuaries*, 14(40). <https://www.soa.org/globalassets/assets/library/research/transactions-of-society-of-actuaries/1962/january/tsa62v14pt1n4022.pdf>
- Martín, J., Parra, M., Pizarro, M., & Sanjuán, E. (2022). Baseline Methods for the Parameter Estimation of the Generalized Pareto Distribution. *Entropy*, 24(2), 1-16. <https://www.mdpi.com/1099-4300/24/2/178>
- Martín, J., Parra, M., Pizarro, M., & Sanjuán, E. (2020). Baseline Methods for Bayesian Inference in Gumbel Distribution. *Entropy*, 22(11), 1-18. <https://www.mdpi.com/1099-4300/22/11/1267>
- Martínez, J., Alvear, C., & Tovar, J. (2020). Uso de la distribución Lévy para ajustar datos con marcada asimetría y valores extremos. *Revista Investigación Operacional*, 41(1), 93-107. <https://rev-inv-ope.pantheonsorbonne.fr/sites/default/files/inline-files/41120-07.pdf>

- Medel, M. C. (2021). *Caminata Aleatoria: Ruina del Jugador* [Tesis de Licenciatura, BUAP]. Repositorio Institucional. <https://repositorioinstitucional.buap.mx/handle/20.500.12371/15355>
- Mora, F., & Perret, S. (2007). Aplicación de técnicas bayesianas en el análisis genético de árboles forestales. *Bosque (Valdivia)*, 28(3), 198-206. https://scielo.cl/scielo.php?script=sci_arttext&pid=S0717-92002007000300003
- Murillo, G. (2009). La teoría de valor extremo y el riesgo operacional: Una aplicación en una entidad financiera. *Ingenierías Universidad de Medellín*, 8(15), 59-70. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1692-33242009000300007
- Nogaj, M., Yiou, P., Parey, S., Malek, F., & Naveau, P. (2006). Amplitude and frequency of temperature extremes over the North Atlantic region. *Geophysical Research Letters*, 33.
- Nolan, J. (2013). Financial modeling with heavy-tailed stable distributions. *WIREs*, 6(1), 1-11. <https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.1286>
- Nolan, J. (2005). *Modeling Financial Data with stable distributions*. Department of Mathematics and Statistics, American University.
- Pérez, K., & Tarango, M. (2019). *Modelación de los montos máximos de siniestros en el seguro de automoviles en México (2008 - 2017)* [Tesis de licenciatura, UAEMex]. Repositorio Institucional. <http://hdl.handle.net/20.500.11799/105232>
- Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3, 119-131.
- Rincón, L. (2012). *Introducción a los procesos estocásticos*. Departamento de ciencias UNAM. Repositorio Institucional.
- Rydman, M. (27 de 06 de 2018). *Application of the Peaks-Over-Threshold Method on Insurance Data*. Uppsala University: <http://uu.diva-portal.org/smash/record.jsf?pid=diva2%3A1231783&dswid=-7010>

- Saavedra, I. d. (2006). *Fundamentos y Tarificación del Seguro de Vida Grupo*. CNSF: http://www.cnsf.gob.mx/Eventos/Premios_2014/1er%20Lugar%20Seguros%202006.pdf
- Sabater, A. (07 de 2017). El desarrollo del sector asegurador y el crecimiento económico: Un análisis empírico para 47 países. *Economía y Sociedad*, 22(52). https://www.scielo.sa.cr/scielo.php?script=sci_arttext&pid=S2215-34032017000200088
- Salas, J. (2013). *Cadenas de Markov desde un punto de vista de aplicaciones [Tesis de Licenciatura, UV]*. Repositorio Institucional.
- Sámano, F. (2020). *Modelación estadística de las salidas y llegadas del aeropuerto internacional de la ciudad de México mediante la distribución de pareto generalizada [Tesis de Licenciatura, UAEMex]*. Repositorio Institucional. <http://ri.uaemex.mx/handle/20.500.11799/109461>
- Singh, A., Allen, D., & Powell, R. (2011). Value at Risk Estimation Using Extreme Value Theory. In *19th International Congress on Modelling and Simulation*. Perth, Australia.
- Steinskog, D., Tjøstheim, D., & Kvamastø, N. (2007). A Cautionary Note on the Use of the Kolmogorov - Smirnov Test for Normality. *Monthly Weather Review*, 135(3), 1151-1157. <https://journals.ametsoc.org/view/journals/mwre/135/3/mwr3326.1.xml>
- Straulino, M. (2014). *Modelación de pérdida agregada aplicada a siniestros de cáncer de mama: caso de una empresa aseguradora [Tesina de Licenciatura, UAEMex]*. Repositorio Institucional. <http://hdl.handle.net/20.500.11799/68072>
- Varela, L. (01 de 02 de 1997). *El contrato de Seguro y su Perfeccionamiento*. Comisión Nacional de Seguros y Fianzas: <https://www.cnsf.gob.mx/Difusion/OtrasPublicaciones/DOCUMENTOS%20DE%20TRABAJO%20DESCRIPTIVOS/DdT67conimag%20BV%20ok.pdf>
- Walpole, R., Myers, R., Myers, S., & Ye, K. (2012). *Probabilidad y estadística para ingeniería y ciencias*. México: PEARSON.

Zacarías, A. (2006). *Métodos de integración y simulación Monte Carlo en la teoría bayesiana [Tesis de Licenciatura, UTM]*. Repositorio Institucional. http://jupiter.utm.mx/~tesis_dig/9944.pdf

Zavershynskyy, O. (2016). *Análisis comparativo de tres modelos de densidad compuesta para el estudio de siniestros de cola pesada con R [Tesis de Maestría, Universidad Complutense]*. Repositorio Institucional. <https://eprints.ucm.es/id/eprint/42142/>

Zhang, J., & Stephens, M. A. (2009). A New and Efficient Estimation Method for the Generalized Pareto Distribution. *Technometrics*, 51, 316-325. <https://www.tandfonline.com/doi/abs/10.1198/tech.2009.08017>

ANEXOS

Anexo 1

Tabla de cuantiles del estadístico Kolmogorov – Smirnov.

n	Nivel de significación α							
	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001
1	0.90000	0.95000	0.97500	0.99000	0.99500	0.99750	0.99900	0.99950
2	0.68337	0.77639	0.84189	0.90000	0.92929	0.95000	0.96838	0.97764
3	0.56481	0.63604	0.70760	0.78456	0.82900	0.86428	0.90000	0.92065
4	0.49265	0.56522	0.62394	0.68887	0.73424	0.77639	0.82217	0.85047
5	0.44698	0.50945	0.56328	0.62718	0.66853	0.70543	0.75000	0.78137
6	0.41037	0.46799	0.51926	0.57741	0.61661	0.65287	0.69571	0.72479
7	0.38148	0.43607	0.48342	0.53844	0.57581	0.60975	0.65071	0.67930
8	0.35831	0.40962	0.45427	0.50654	0.54179	0.57429	0.61368	0.64098
9	0.33910	0.38746	0.43001	0.47960	0.51332	0.54443	0.58210	0.60846
10	0.32260	0.36866	0.40925	0.45562	0.48893	0.51872	0.55500	0.58042
11	0.30829	0.35242	0.39122	0.43670	0.46770	0.49539	0.53135	0.55588
12	0.29577	0.33815	0.37543	0.41918	0.44905	0.47672	0.51047	0.53422
13	0.28470	0.32549	0.36143	0.40362	0.43247	0.45921	0.49189	0.51490
14	0.27481	0.31417	0.34890	0.38970	0.41762	0.44352	0.47520	0.49753
15	0.26589	0.30397	0.33750	0.37713	0.40420	0.42934	0.45611	0.48182
16	0.25778	0.29472	0.32733	0.36571	0.39201	0.41644	0.44637	0.46750
17	0.25039	0.28627	0.31796	0.35528	0.38086	0.40464	0.43380	0.45540
18	0.24360	0.27851	0.30936	0.34569	0.37062	0.39380	0.42224	0.44234
19	0.23735	0.27136	0.30143	0.33685	0.36117	0.38379	0.41156	0.43119
20	0.23156	0.26473	0.29408	0.32866	0.35241	0.37451	0.40165	0.42085
21	0.22517	0.25858	0.28724	0.32104	0.34426	0.36588	0.39243	0.41122
22	0.22115	0.25283	0.28087	0.31394	0.33666	0.35782	0.38382	0.40223
23	0.21646	0.24746	0.27490	0.30728	0.32954	0.35027	0.37575	0.39380
24	0.21205	0.24242	0.26931	0.30104	0.32286	0.34318	0.36787	0.38588

25	0.20790	0.23768	0.26404	0.29518	0.31657	0.33651	0.36104	0.37743
26	0.20399	0.23320	0.25908	0.28962	0.30963	0.33022	0.35431	0.37139
27	0.20030	0.22898	0.25438	0.28438	0.30502	0.32425	0.34794	0.36473
28	0.19680	0.22497	0.24993	0.27942	0.29971	0.31862	0.34190	0.35842
29	0.19348	0.22117	0.24571	0.27471	0.29466	0.31327	0.33617	0.35242
30	0.19032	0.21756	0.24170	0.27023	0.28986	0.30818	0.33072	0.34672
31	0.18732	0.21412	0.23788	0.26596	0.28529	0.30333	0.32553	0.34129
32	0.18445	0.21085	0.23424	0.26189	0.28094	0.29870	0.32058	0.33611
33	0.18171	0.20771	0.23076	0.25801	0.27577	0.29428	0.31584	0.33115
34	0.17909	0.21472	0.22743	0.25429	0.27271	0.29005	0.31131	0.32641
35	0.17659	0.20185	0.22425	0.25073	0.26897	0.28600	0.30597	0.32187
36	0.17418	0.19910	0.22119	0.24732	0.26532	0.28211	0.30281	0.31751
37	0.17188	0.19646	0.21826	0.24404	0.26180	0.27838	0.29882	0.31333
38	0.16966	0.19392	0.21544	0.24089	0.25843	0.27483	0.29498	0.30931
39	0.16753	0.19148	0.21273	0.23785	0.25518	0.27135	0.29125	0.30544
40	0.16547	0.18913	0.21012	0.23494	0.25205	0.26803	0.28772	0.30171
41	0.16349	0.18687	0.20760	0.23213	0.24904	0.26482	0.28429	0.29811
42	0.16158	0.18468	0.20517	0.22941	0.24613	0.26173	0.28097	0.29465
43	0.15974	0.18257	0.20283	0.22679	0.24332	0.25875	0.27778	0.29130
44	0.15795	0.18051	0.20056	0.22426	0.24060	0.25587	0.27468	0.28806
45	0.15623	0.17856	0.19837	0.22181	0.23798	0.25308	0.27169	0.28493
46	0.15457	0.17665	0.19625	0.21944	0.23544	0.25038	0.26880	0.28190
47	0.15295	0.17481	0.19420	0.21715	0.23298	0.24776	0.26600	0.27896
48	0.15139	0.17301	0.19221	0.21493	0.23059	0.24523	0.26328	0.27611
49	0.14987	0.17128	0.19028	0.21281	0.22832	0.24281	0.26069	0.27339
50	0.14840	0.16959	0.18841	0.21068	0.22604	0.24039	0.25809	0.27067

n > 50	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$	$\frac{1.73}{\sqrt{n}}$	$\frac{1.85}{\sqrt{n}}$	$\frac{1.95}{\sqrt{n}}$
--------	-------------------------	-------------------------	-------------------------	-------------------------	-------------------------	-------------------------	-------------------------	-------------------------

Anexo 2

Tabla de corrección de Lilliefors

n	Nivel de significación α				
	0.2	0.15	0.1	0.05	0.01
4	0.300	0.319	0.352	0.381	0.417
5	0.285	0.299	0.315	0.337	0.405
6	0.265	0.277	0.294	0.319	0.364
7	0.247	0.258	0.276	0.300	0.348
8	0.233	0.244	0.261	0.285	0.331
9	0.223	0.233	0.249	0.271	0.311
10	0.215	0.224	0.239	0.258	0.294
11	0.206	0.217	0.230	0.249	0.284
12	0.206	0.217	0.230	0.249	0.284
13	0.190	0.202	0.214	0.234	0.268
14	0.183	0.194	0.207	0.227	0.261
15	0.177	0.187	0.201	0.220	0.257
16	0.173	0.182	0.195	0.213	0.250
17	0.169	0.177	0.189	0.206	0.245
18	0.166	0.173	0.184	0.200	0.239
19	0.163	0.169	0.179	0.195	0.235
20	0.160	0.166	0.174	0.190	0.231
21	0.157	0.162	0.170	0.186	0.226
2	0.154	0.158	0.167	0.182	0.220
23	0.150	0.154	0.164	0.179	0.213
24	0.146	0.151	0.161	0.176	0.207
25	0.142	0.147	0.158	0.173	0.200
26	0.139	0.145	0.155	0.171	0.197
27	0.137	0.143	0.153	0.168	0.194

28	0.135	0.140	0.150	0.166	0.192
29	0.133	0.138	0.147	0.163	0.190
30	0.131	0.136	0.144	0.161	0.187
n > 30	$\frac{0.736}{\sqrt{n}}$	$\frac{0.768}{\sqrt{n}}$	$\frac{0.805}{\sqrt{n}}$	$\frac{0.886}{\sqrt{n}}$	$\frac{1.031}{\sqrt{n}}$
