



**UNIVERSIDAD ESAN**

**FACULTAD DE CIENCIAS ECONÓMICAS Y ADMINISTRATIVAS**

**Determinantes de la siniestralidad de autos en una empresa aseguradora peruana: Un enfoque bayesiano**

**Tesis para optar el Título de Licenciada en Economía y Negocios Internacionales que presenta:**

**AUTOR: Liesel Nathaly Condori Ancasi**

**ASESOR: Jorge Cortez Cumpa**

**LIMA – PERÚ**

**2020**



**DETERMINANTES DE LA SINIESTRALIDAD DE AUTOS EN UNA EMPRESA  
ASEGURADORA PERUANA: UN ENFOQUE BAYESIANO**

## **DEDICATORIA**

Gracias a Dios por permitirme recorrer con perseverancia el camino del aprendizaje. A mis padres Ángel y Maruja por confiar y creer en mí. A mis hermanos por sus deseos de hacer realidad mis sueños. A mis abuelos por su cariño, confianza y ayuda. A mis familiares por su apoyo moral y consejos.

Mi agradecimiento y gratitud a mi asesor Jorge Cortez Cumpa, por la paciencia, constancia y asertividad para apoyarme en este proceso sobre todo cuando las cosas se ponían difíciles. Asimismo, mi reconocimiento a los docentes de la Universidad Esan, por compartir sus conocimientos durante mi etapa de estudiante, cada uno de sus aportes fue importante para realizar la tesis.

A mis amigos por siempre darme ánimos para continuar y estar felices por cada logro que tengo. Finalmente, a mi porque a pesar que el camino fue largo me enseñó que cualquier meta es alcanzable si realmente te esfuerzas por lograrla. Todo es cuestión de voluntad.

## INDICE

INDICE.....	5
ÌNDICE DE GRÀFICOS.....	8
ÍNDICE DE TABLAS.....	8
RESUMEN.....	9
ABSTRACT.....	10
INTRODUCCION.....	11
CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA.....	12
1.1. DESCRIPCIÓN DE LA SITUACIÓN PROBLEMÁTICA.....	12
1.2. FORMULACIÓN DEL PROBLEMA.....	13
1.2.1. Problema General.....	13
1.2.2. Problemas Específicos.....	13
1.3. DETERMINACION DE OBJETIVOS.....	13
1.3.1. Objetivo General:.....	13
1.3.2. Objetivos Específicos.....	14
1.4. JUSTIFICACIÓN DE LA INVESTIGACIÓN.....	14
1.5. DELIMITACIÓN DE LA INVESTIGACIÓN.....	15
CAPÍTULO II: MARCO TEÓRICO.....	15
2.1. BASES TEÓRICAS.....	15
2.1.1. La Economía de la Información.....	15
2.1.2. La Asimetría de la Información:.....	17
2.1.3. La información Asimétrica en el mercado de seguros.....	19
a) Los Problemas de Riesgo Moral en el Mercado de Seguros.....	19
b) Los Problemas de Selección Adversa en el Mercado de Seguros.....	20
- Los Problemas de Signaling.....	22
- Los Problemas de Screening.....	22
2.1.4. La teoría de la Decisión.....	23
a) Decisiones Bajo Certidumbre.....	23
b) Decisiones Bajo Riesgo.....	23
c) Decisiones Bajo Incertidumbre.....	24

2.1.5.	Probabilidad de siniestro .....	24
2.1.6.	Factores relativos a la conducta aleatoria del individuo.....	24
2.1.7.	Factores relativos al vehículo asegurado.....	25
2.1.8.	Factores relativos a la póliza .....	26
2.1.9.	Factores relativos al conductor.....	26
2.1.10.	Términos relacionados a seguros .....	27
a)	Contrato de Seguro.....	27
b)	Riesgo.....	27
c)	Siniestro.....	28
d)	Prima de seguros .....	28
2.2.	ANTECEDENTES DE LA INVESTIGACIÓN.....	29
2.2.1.	Antecedentes empíricos: .....	29
2.2.2.	Otros antecedentes de la Investigación .....	31
a)	Historia del Seguro.....	31
b)	Historia del Seguro en el Perú.....	32
c)	Historia del Seguro Vehicular .....	33
d)	Historia del Seguro Vehicular en el Perú.....	33
e)	El seguro obligatorio de accidentes de tránsito (SOAT).....	34
2.3.	CONTEXTO DE LA INVESTIGACIÓN .....	35
2.3.1.	El Mercado Asegurador Peruano .....	35
2.3.2.	El mercado del Seguro Vehicular Peruano.....	36
2.4.	HIPÓTESIS.....	37
2.4.1.	Argumentación.....	37
2.4.2.	Enunciado de la Hipótesis .....	38
a)	Hipótesis General .....	38
b)	Hipótesis Especificas .....	38
CAPÍTULO III: METODOLOGIA .....		39
3.1.	DISEÑO DE LA INVESTIGACIÓN .....	39
3.2.	POBLACIÓN Y MUESTRA.....	39
3.3.	MÉTODO DE RECOLECCIÓN DE DATOS.....	39
3.4.	MÉTODO DE ANÁLISIS DE DATOS .....	40

3.4.1.	El Modelo Logit .....	40
a)	Modelo de Regresión Logística Binaria.....	41
-	Bondad de Ajuste del Modelo.....	42
-	Contraste de bondad de ajuste global de Hosmer-Lemeshow .....	42
3.5.	TÉCNICAS DE RECOLECCIÓN DE DATOS: .....	43
3.6.	TÉCNICAS PARA EL PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN .....	43
3.6.1.	La probabilidad Condicional.....	43
3.6.2.	El teorema de Bayes.....	44
3.6.3.	Inferencia Bayesiana .....	45
a)	Caso Inferencia Bayesiana con a priori Informativa .....	45
b)	Caso Inferencia Bayesiana a priori no Informativa.....	46
3.6.4.	El Algoritmo MCMC: Muestro de Gibbs y Metropolis-Hanstings.....	47
a)	El Algoritmo de Muestreo de Gibbs .....	47
b)	El Algoritmo de Metropolis-Hastings.....	48
3.6.5.	Paquete Estadístico de Análisis Econométrico STATA:.....	49
CAPITULO IV: ANALISIS E INTERPRETACION DE LAS PRUEBAS ESTADÍSTICAS, FÓRMULAS Y RESULTADOS .....		50
4.1.	ANÁLISIS DE LA RELACIÓN ENTRE LAS VARIABLES DE ESTUDIO.....	50
4.2.	ANÁLISIS DE DATOS CON INFERENCIA BAYESIANA.....	58
4.2.1.	Distribución de Probabilidad A priori.....	58
4.2.2.	Función de Máximo Verosimilitud .....	59
4.2.3.	Distribución de Probabilidad Posterior .....	63
CAPÍTULO V: DISCUSIÓN, CONCLUSIONES Y RECOMENDACIONES.....		66
5.1.	DISCUSIÓN DE RESULTADOS Y CONCLUSIONES .....	66
5.2.	LIMITACIONES Y RECOMENDACIONES.....	69
ANEXOS .....		77
Anexo 1: Matriz de Consistencia.....		77
Anexo 2: Empresas de Seguros Autorizadas al 31 de diciembre del 2019.....		79
Anexo 3: Estructura de las Primas de Seguros Netas por Ramos .....		79
Anexo 4: Evolución de las Primas de Seguros Netas por Ramo.....		80
Anexo 5: Ranking del Total de Primas de Seguros.....		80

Anexo 6: Ranking del Total de Primas de Seguros para los ramos generales, accidentes y enfermedades .....	81
Anexo 7: Archivo do en software econométrico STATA.....	82
Anexo 8: Diagrama del Repositorio de Autos .....	85

## ÍNDICE DE GRÁFICOS

Gráfico 1: Composición del Mercado según Primas de Seguros Netas de Principales Riesgos – Año 2018 .....	36
Gráfico 2: Composición del Mercado según Primas de Seguros Netas de Principales Riesgos – Año 2019 .....	37
Gráfico 3: Periodo de fabricación de los vehículos asegurados.....	52
Gráfico 4: Marca de los vehículos asegurados.....	53
Gráfico 5: Obtención de GPS de los vehículos asegurados .....	54
Gráfico 6: Sexo del propietario del vehículo asegurado .....	55
Gráfico 7: Tipo de vehículo asegurado .....	56
Gráfico 8: Estado civil del propietario del vehículo asegurado .....	57
Gráfico 9: Histograma de los parámetros del modelo LOGIT Bayesiano .....	65

## ÍNDICE DE TABLAS

Tabla 1: Descripción de la ocurrencia del siniestro vehicular .....	50
Tabla 2: Descripción de las variables explicativas según la ocurrencia de siniestro vehicular .....	50
Tabla 3: Periodo de fabricación de los vehículos asegurados.....	51
Tabla 4: Marca de los vehículos asegurados.....	52
Tabla 5: Obtención de GPS de los vehículos asegurados .....	53
Tabla 6: Sexo del propietario del vehículo .....	54
Tabla 7: Tipo de vehículo de la empresa aseguradora .....	55
Tabla 8: Estado civil del propietario del vehículo asegurado .....	56
Tabla 9: Estadísticos de la edad del propietario del vehículo asegurado .....	57
Tabla 10: Estadísticos de la prima pagada del seguro del vehículo .....	58
Tabla 11: Estadísticos para comparar modelos PROBIT y LOGIT.....	59
Tabla 12: Estimación del modelo LOGIT.....	61
Tabla 13: Estimación del modelo LOGIT Bayesiano .....	63



## RESUMEN

La presente investigación tiene como objetivo encontrar la relación que existe entre los determinantes no observables, asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana. Para lograr el objetivo planteado se hizo una revisión de la teoría relacionada a la economía de la información y se analizaron los problemas de asimetría de información en el mercado asegurador. Asimismo, se demuestra que el problema existente de selección adversa en el mercado de seguros de autos, evidenciado por la diferencia de información que poseen el principal y el agente respecto al riesgo de este último, impacta en la probabilidad de siniestro. En tal sentido, la existencia de factores de riesgo no observables modifica la probabilidad de siniestro de un auto asegurado. El contexto anteriormente mencionado hizo propicia la aplicación del instrumento bayesiano como medio de análisis en términos de multiplicación de probabilidades permitiendo tomar en cuenta la información que no puede conocer la aseguradora. Luego de realizar el planteamiento del modelo logit bayesiano se pudo identificar que la explicación del modelo para encontrar los determinantes de la probabilidad de siniestro mejoraba al incluir los factores de riesgo no observables mediante una distribución a priori. Adicionalmente, se pudo identificar que los determinantes de la probabilidad de siniestro relevantes son: la marca del vehículo, la obtención del GPS, el tipo de vehículo, el estado civil del propietario del vehículo, la prima pagada por el seguro del vehículo.

**ABSTRACT**

The objective of this research is to find the relationship that exists between the unobservable determinants associated with the random behavior of the individual and the occurrence of vehicle accidents in a Peruvian insurance company. To achieve the stated objective, a review of the theory related to the information economy was made and the problems of information asymmetry in the insurance market were analyzed. Likewise, it is shown that the existing problem of adverse selection in the auto insurance market, evidenced by the difference in information held by the principal and the agent regarding the latter's risk, impacts the probability of a claim. In this sense, the existence of unobservable risk factors modifies the probability of a claim of an insured car. The aforementioned context favored the application of the Bayesian instrument as a means of analysis in terms of multiplication of probabilities, allowing to take into account the information that the insurer cannot know. After carrying out the approach of the Bayesian logit model, it was possible to identify that the explanation of the model to find the determinants of the probability of a claim improved by including the unobservable risk factors through an a priori distribution. Additionally, it was possible to identify that the relevant determinants of the probability of a claim are: the vehicle brand, obtaining the GPS, the type of vehicle, the marital status of the vehicle owner, the premium paid for the vehicle insurance.

## INTRODUCCION

La presente investigación busca encontrar la relación que existe entre los determinantes no observables, asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana. El objetivo planteado se realizó mediante una revisión teórica de la economía de la información haciendo énfasis en el análisis de los problemas de asimetría de información en el mercado asegurador.

La investigación se enfoca en analizar el problema de selección adversa que ocurre entre el principal y el agente por la dificultad de conocer toda la información relacionada al riesgo que tiene este último, cuando las primas no reflejan la heterogeneidad de los riesgos del asegurado. Por lo tanto, al momento de estudiar los determinantes de la probabilidad de siniestro en el mercado asegurador de autos, existe factores no observables; es decir, que la aseguradora no puede conocer.

En consecuencia, la existencia de factores de riesgo no observables podría modificar la probabilidad de ocurrencia de siniestro de un auto asegurado; motivo por el cual, de todos los determinantes no observables que pueden existir se hizo especial énfasis en los que están relacionados a la conducta aleatoria del individuo. La teoría de decisión refuerza este planteamiento ya que demuestra que los individuos no siempre actúan racionalmente.

El contexto antes mencionado hace propicia la aplicación del método bayesiano como medio de análisis en términos de multiplicación de probabilidades permitiendo tomar en cuenta la información que no puede conocer la aseguradora. Luego de realizar el planteamiento del modelo logit bayesiano, el ajuste del modelo para encontrar los determinantes de la probabilidad de siniestro mejora luego de incluir los determinantes de riesgo no observables mediante una probabilidad.

Como parte de respuesta a los objetivos específicos se identifica que los determinantes de la probabilidad de siniestro relevantes son: la marca del vehículo, la obtención del GPS, el tipo de vehículo, el estado civil del propietario del vehículo, la prima pagada por el seguro del vehículo. Estos resultados guardan relación con las investigaciones revisadas como parte del marco teórico en la sección de los antecedentes de la investigación.

Finamente, el tema de investigación analizado es importante debido a que la selección adversa produce ineficiencias en los mercados, dando lugar a que las transacciones no se realicen de forma beneficiosa, obteniéndose mercados con niveles de actividad inferiores o la inexistencia de los mismos.

## **CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA**

### **1.1. DESCRIPCIÓN DE LA SITUACIÓN PROBLEMÁTICA**

En el Perú, la penetración del seguro de autos es baja, sólo el 23% de vehículos están asegurados, es decir que de las 2.7 millones de unidades que circulan en el país, 621 mil tienen un seguro y poco más de 2 millones no tienen protección. (Bedoya, 2019) Por otro lado, sólo el 40% de vehículos cuenta con el Seguro Obligatorio de Accidentes de tránsito, SOAT exponiendo a sus pasajeros y a peatones al desamparo en caso de un accidente de tránsito. (APESEG, 2019)

Analizando los siniestros vehiculares se tiene que el de mayor recurrencia es el choque en circulación, que representa el 55% del total de accidentes, seguido de la rotura de lunas (13%), robo parcial (7%) y choque estacionado (7%) entre otros. (Puelles, 2019) Para el año 2018 se reportaron 3,244 accidentes de tránsito fatales y 87,480 no fatales. Asimismo, hubo 13,787 denuncias por robo de vehículos. (INEI, 2018)

Las aseguradoras para calcular la prima del seguro de autos toman en cuenta dos valores: la cantidad de accidentes que ocurren con respecto a cada 100 y el costo de los accidentes en el lugar donde se contrata el seguro. (Díaz, 2018) En tal sentido, las aseguradoras intentan clasificar sus pólizas en clases tarifarias homogéneas, es decir, asignando a las mismas todas aquellas pólizas que pertenezcan a la misma clase a fin de conseguir cobrar una prima adecuada. (Segovia Vargas, Camacho Miñano, Pascual Ezama y Rodríguez Piñero, 2015)

Sin embargo, determinados aspectos de la toma de decisiones del asegurador, como las variables que deben utilizar y la manipulación del valor de las mismas para llegar a una decisión final están delimitados o definidos de forma imprecisa. (De Andrés Sánchez y Terceño Gomez, 2002) Cuando las primas no reflejan la heterogeneidad de los riesgos de los asegurados, se genera un fenómeno de selección adversa. Para los riesgos bajos, la prima que refleja el riesgo promedio parece alta, mientras que, para los altos riesgos, la cobertura parece baja generando que los riesgos altos se aseguran en mayor proporción que los riesgos bajos. (Bardey, 2008)

En tal sentido, al analizar los determinantes de riesgo de los seguros de autos, algunos no pueden ser identificados, como, por ejemplo: el comportamiento agresivo al volante u otras variables como la agresividad, la personalidad, el estrés o la propensión al riesgo que deberían tenerse en cuenta y, sin embargo, al menos de forma explícita, no parece que las compañías les den mucha importancia. (Segovia Vargas, Camacho Miñano, Pascual Ezama y Rodríguez Piñero, 2015)

En consecuencia, analizar la relación que existe entre la ocurrencia de siniestros y los determinantes no observables asociados a la conducta aleatoria del individuo cobra relevancia en el

estudio de la información asimétrica en el mercado de seguros de autos y hace posible contrastar la presunción que estos factores no observables si están correlacionados con la probabilidad de ocurrencia de siniestro.

## **1.2. FORMULACIÓN DEL PROBLEMA**

### **1.2.1. Problema General**

¿Cuál es la relación entre los determinantes no observables, asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana?

### **1.2.2. Problemas Específicos**

- ¿Cuál es la relación de los determinantes observables, asociados al conductor con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018?
- ¿Cuál es la relación de los determinantes observables, asociados al vehículo con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018?
- ¿Cuál es la relación de los determinantes observables, asociados a la póliza con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018?

## **1.3. DETERMINACION DE OBJETIVOS**

### **1.3.1. Objetivo General:**

Encontrar la relación que existe entre los determinantes no observables, asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana.

### **1.3.2. Objetivos Específicos**

- Demostrar que los determinantes observables, asociados al conductor tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018
- Demostrar que los determinantes observables, asociados al vehículo tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018
- Demostrar que los determinantes observables, asociados a la póliza tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018

### **1.4. JUSTIFICACIÓN DE LA INVESTIGACIÓN**

La selección adversa produce ineficiencias en los mercados, ya que da lugar a que no se realicen transacciones que serían beneficiosas, obteniéndose mercados con niveles de actividad inferiores, donde sólo se acaban vendiendo productos de baja calidad, o incluso la desaparición o inexistencia de los mercados. En tal sentido, en el mercado de seguros de autos, se busca minimizar el problema de ventaja informacional de los asegurados sobre los aseguradores.

La evidencia empírica demuestra que la ocurrencia del siniestro puede ser explicada por factores de riesgo que no se pueden determinar a priori, es decir factores escondidos que la aseguradora no puede conocer. Del conjunto infinito de posibles factores que no pueden ser conocidos por el problema de ventaja informacional de los asegurados frente a la aseguradora, se analizará a los determinantes no observables asociados a la conducta aleatoria del individuo.

La aleatoriedad se asocia a todo proceso cuyo resultado no es previsible más que en razón de la intervención del azar; esto sucede con las acciones buenas o malas que puede tomar el individuo al conducir un vehículo. La teoría de la decisión, refuerza este argumento porque demuestra que los individuos no siempre actúan racionalmente. Estos hechos son importantes pero las aseguradoras no los toman en cuenta por la imposibilidad de conocer a ciencia cierta cuál sería su resultado.

Con la presente investigación se impulsa el estudio de la economía de la información mediante el análisis de la asimetría de la información en el mercado de seguros y el problema de selección adversa que se da entre los participantes del contrato de seguro: el asegurado y el asegurador. Asimismo, la

metodología aplicada en la investigación contribuye con el empleo del método bayesiano y la estadística actuarial para el análisis de los factores de riesgo que no pueden ser determinados a priori que se presume están correlacionados con la ocurrencia de siniestro.

Con todas estas premisas, el objetivo de este trabajo es comprobar la capacidad de explicación de los determinantes no observables asociados a la conducta aleatoria del individuo sobre la variable “ocurrencia de siniestro”, luego con los resultados obtenidos y la metodología aplicada se aporta información relevante para otras investigaciones relacionadas así como para los entes reguladores de la actividad de seguros en Lima y el Perú que busquen la eficiencia y aumento de la penetración de los seguros de autos.

## **1.5. DELIMITACIÓN DE LA INVESTIGACIÓN**

El problema planteado en la presente investigación se analiza tomando los siguientes detalles de tiempo y espacio:

- Años analizados 2016, 2017 y 2018
- Datos proporcionados por una empresa aseguradora peruana con amplia participación en el mercado asegurador.

## **CAPÍTULO II: MARCO TEÓRICO**

### **2.1. BASES TEÓRICAS**

#### **2.1.1. La Economía de la Información**

Durante más de 100 años, la modelización de la ciencia económica estuvo centrada en modelos de información perfecta, aunque se reconocía que en realidad la información era imperfecta. De acuerdo a Marshall “Natura non facit saltum”, se confiaba en que las economías con información no tan imperfecta se asemejaban a las economías con información perfecta. (Stiglitz, 2002)

En el siglo XX, los creadores del modelo neoclásico, olvidaron de cómo podrían alterar sus análisis las cuestiones relacionadas con la información tal vez porque no pudieron ver la forma de incorporarlas en sus modelos aparentemente precisos, quizás porque hacerlo los hubiera llevado

a unas conclusiones incómodas acerca de la eficiencia de los mercados. (Stiglitz, 2002)

En 1961, algunos economistas tolemaicos reconocieron a la información como un recurso valioso ignorado por la economía. Asimismo, se dio razón que para muchos problemas en la economía este descuido de la ignorancia era permisible o incluso deseable, pero para otros no y que algunos aspectos importantes de la organización económica adquirirían un nuevo significado cuando eran considerados desde el punto de vista de la búsqueda de información. (Stigler, 1961)

Los problemas de asignación, redistribución y crecimiento de los recursos con la globalización proporcionaron un marco donde no sólo se cuestionan teóricamente los modelos de interpretación de la realidad, sino también los espacios de la realidad susceptibles de análisis. Los economistas pasaron de observar y analizar la producción a interesarse por la naturaleza de los servicios, la calidad o las capacidades y habilidades de negociación de los agentes económicos. (Rausell Köster y Carrasco Arroyo, 2003)

A partir de 1990, la sociedad de la información surge como un enfoque caracterizado por un cambio de paradigma en las estructuras industriales y en las relaciones sociales. La expresión “sociedad de la información” designa una forma nueva de organización de la economía y la sociedad. Los esfuerzos por convertir la información en conocimiento es una característica que la identifica. Cuanto mayor es la cantidad de información generada por una sociedad, mayor es la necesidad. (Castells, 2000)

Luego en el año 2000, se analizó el término de sociedad del conocimiento referido a la transformación social, cultural, económica, política e institucional, y una perspectiva más pluralista y desarrolladora. Asimismo, se señala que es preferible al de la “sociedad de la información” ya que expresa mejor la complejidad y el dinamismo de los cambios que se están efectuando sobre el conocimiento y su importancia para empoderar y desarrollar todos los sectores de la sociedad y el crecimiento económico. (Khan, 2003)

Por lo antes expuesto, uno de los nuevos temas importantes en el funcionamiento de la economía es el papel que juega la información. Así la aproximación de las ciencias sociales coincide en considerar que los procesos de adquisición y difusión de la información son factores directamente explicativos del desarrollo social y del crecimiento económico y de la configuración de las relaciones entre individuos, organizaciones e instituciones. (Rausell Köster y Carrasco Arroyo, 2003)

Es así que la economía de la información también conocida como teoría de contratos estudia las consecuencias de la existencia de asimetría de información entre diversos agentes económicos, principal-agente, sobre la eficiencia de la relación que establecen. En estas situaciones los agentes intentan sobrellevar la ignorancia de cierta información relevante, tomando decisiones



diseñadas para adquirir nueva información o evadir los costos de su ignorancia. (Loteró, 2005)

De acuerdo al sistema de clasificación de la literatura académica en el campo de la economía adoptado por la Revista de Literatura Económica (JLE) por sus siglas en inglés, la economía de la información se encuentra dentro de la categoría D – Microeconomía específicamente en la categoría D8 - Información, conocimiento e incertidumbre. (American Economic Association, 2019)

### **2.1.2. La Asimetría de la Información:**

En economía, la información es determinante en la toma de decisiones. A veces, consumidores y productores no tienen la misma información sobre las variables involucradas en su elección. Así, se define la información asimétrica como “la situación en la que el comprador y el vendedor tienen información diferente sobre una transacción”. (Villaseca, Torrent y Lladós, 2001)

Akerlof (1970) fue el primero en analizar la información asimétrica mediante el análisis del mercado de automóviles de segunda mano, en el cual, el vendedor posee mejor información que el comprador sobre el estado del coche. El comprador debe suponer la calidad del producto y deducir que hay coches de buena y mala calidad; en tal sentido, orientar sus ofertas a un precio de calidad intermedia provocando que la proporción de coches de baja calidad vendidos sea mayor a los de alta calidad y con un mayor precio. Mientras haya más coches de baja calidad vendidos, los compradores potenciales aumentarán su certeza que todos los coches disponibles son de baja calidad y no estarán dispuestos a pagar un precio elevado por ellos. Este proceso continuaría hasta que los coches defectuosos expulsan del mercado de segunda mano a los de mejor calidad, que sólo se venderían a un menor precio. En consecuencia, la información asimétrica sobre la calidad de los coches ocasiona la venta de menos productos de calidad. En cambio, si los consumidores podrían identificar la calidad del producto antes de la compra, los productos de calidad desplazarían a los de baja calidad.

Arrow (1973) identificó otro efecto de la información asimétrica sobre la evolución futura de los precios. Se analizó que el agente económico tiene una probabilidad de distribución sobre posibles valores de las variables interesantes para él, como los precios actuales y futuros o las cualidades de los bienes. Estas variables económicas se observan sobre alguna otra variable llamada señal. En tal sentido, las decisiones tomadas dependen, de la distribución de variables económicas, pero si esta distribución es a su vez modificada por señales, entonces el comportamiento económico depende no solo de las variables que se consideran relevantes, sino

también de las señales que pueden tener poca importancia económica pero que pueden reducir la incertidumbre en la predicción de variables no observadas.

Spence (1973) analizó como en algunos mercados evitaban los problemas de información asimétrica mediante señales sobre la calidad de los productos o servicios que querían intercambiar. Por ejemplo, en el mercado de bienes de consumo duraderos, las empresas que fabrican un producto de mejor calidad ofrecen una garantía generosa con el objetivo de indicar a los consumidores que se trata de un buen producto a fin de que estos puedan pagar un precio más elevado. Para que funcione una señal de mercado, su provisión tiene que ser mucho más costosa para los productores de baja calidad. Además, las señales que transmiten información acerca de la calidad de un bien o de un servicio obligan a todos los competidores a revelar información, aunque ésta les sea poco favorable.

Asimismo, Spence (1973) analizó la información asimétrica en el mercado de trabajo. Por ejemplo, un postulante sabe mejor su capacidad profesional que la empresa que considera su contratación; para estos casos, la señalización es muy frecuente. En este escenario, la política salarial es importante debido a que las empresas no conocen la productividad del trabajador, pero asumen que a mayor salario la productividad aumenta. La asimetría de información, puede ocasionar que los trabajadores persigan sus objetivos propios en lugar de los objetivos empresariales, este es un caso denominado relaciones de agencia. En este tipo de relación, el bienestar de un agente económico depende de las acciones de otro. La diferencia de objetivos conlleva a la aparición de incentivos que posibiliten la ejecución de los objetivos marcados por el principal.

Stiglitz (1975), extendió el estudio de la asimetría de la información al mercado del crédito. En el mercado de los créditos la persona que se endeuda sabe mucho mejor que la entidad financiera si está dispuesto y en condiciones de repagar el dinero prestado. En tal sentido, reconoce que hay importantes asimetrías en la información y su tamaño puede ser afectado por las acciones de las empresas y de los individuos. Asimismo, manifiesta que los modelos económicos basados en la idea de mercados perfectos están alejados de la realidad y que, si bien pueden ser aceptados para países de Norteamérica o de Europa occidental, son totalmente inadecuados para países en desarrollo ya que la realidad de estos países está caracterizada por la imperfección de la información, la ausencia de mercados y la frecuencia y persistencia de instituciones disfuncionales.

Rothschild y Stiglitz (1976) hacen una aproximación de la información asimétrica al mercado asegurador mostrando que las ofertas de ventas que sobreviven al proceso competitivo, no especifican un precio al que los clientes puedan comprar todo el seguro que desean, sino más bien un precio y una cantidad (una cantidad particular de seguro que el individuo puede comprar a ese precio). Además, si los individuos estuvieran dispuestos o pudieran revelar su información,

todos podrían estar mejor. Tal es el caso de los individuos de alto riesgo que causan una externalidad: los individuos de bajo riesgo se encuentran en una peor situación de la que estarían en ausencia de los individuos de alto riesgo. Sin embargo, los individuos de alto riesgo tampoco están mejor de lo que estarían en ausencia de las personas de bajo riesgo.

El que la información es imperfecta era, por supuesto, un hecho muy conocido por los economistas. La razón por la que los modelos con información imperfecta no se desarrollaron antes fue que no estaba claro cómo hacerlo: mientras que hay una única forma en la que la información es perfecta, hay un número infinito de maneras en las que la información puede ser imperfecta. Una de las claves del éxito fue formular unos modelos sencillos en los que el conjunto de información relevante podía especificarse plenamente y en los que también se podía especificar totalmente las formas concretas en las que la información era imperfecta. (Stiglitz, 2002)

### **2.1.3. La información Asimétrica en el mercado de seguros**

La información asimétrica y los problemas de “azar moral” y “selección adversa” que la misma plantea se han erigido en uno de los pilares fundamentales de la teoría financiera en su vertiente contractual. La aplicación de dicho enfoque de la teoría financiera a los contratos de seguros tiene una especial relevancia, debido a que la información asimétrica que generalmente tienen asegurador y asegurado, respecto a la siniestralidad particular de este último, que, sin duda difiere de la siniestralidad media de la población asegurada, que forma la base de cálculo de las primas por las compañías aseguradoras. (Lobera, 1994)

#### **a) Los Problemas de Riesgo Moral en el Mercado de Seguros**

El riesgo moral es el comportamiento del agente que no es observable para el principal, o que, aun siendo observable, no es verificable. “La no verificabilidad del esfuerzo tiene como consecuencia inmediata que éste no puede ser incluido entre los términos del contrato ya que, en caso de conflicto, ningún tribunal de justicia podría pronunciarse sobre si el contrato fue satisfecho o no” (Macho y Pérez, 1994)

Estos comportamientos no provienen de informaciones privadas sino de la incapacidad de las compañías de observar algunas acciones de sus asegurados que influyen en su riesgo. De manera general, el riesgo moral en el sector asegurador se define como el aumento de los gastos reembolsados generados por el hecho que los individuos tienen un seguro. Es importante distinguir lo que releva el cambio de riesgo antes que este ocurra, por el aumento de la frecuencia de los siniestros (i.e. el riesgo moral ex ante) del cambio que ocurre por comportamientos después

de que el riesgo ocurre (riesgo moral ex post). (Bardey, 2008)

La problemática del riesgo moral ex ante se da cuando el seguro desmotiva las acciones de prevención de los asegurados. Por ejemplo, un seguro completo, donde no se perciben costos o beneficios de acciones de prevención porque se obtiene el mismo nivel de riqueza, sin importar lo que suceda. Además, como los comportamientos preventivos no pueden ser observados, los aseguradores no pueden fijar las primas en función de estos, lo que hace que los asegurados no se beneficien de una reducción de sus primas por ‘comportamientos positivos’. Al final, la ineficiencia del riesgo moral ex ante se traduce en riesgos más frecuentes y en primas mayores. (Bardey, 2008)

Para contrarrestar estos efectos, se puede incluir un deducible en el contrato para reducir las ineficiencias del riesgo moral antes de su firma. El deducible permite que los asegurados pueden percibir beneficios en realizar acciones preventivas. (Shavell, 1979) En contraparte, para algunos tipos de seguros, incluir un deducible en el contrato de seguro puede no ser eficiente en especial al existir “incentivos naturales”, por ejemplo, el seguro de salud y donde el asegurado intenta siempre mantener mejor salud. (Bardey y Lesur, 2005)

El riesgo moral ex post plantea que el reembolso pueda aumentar una vez que ocurrió el daño cuando los asegurados se benefician de una cobertura generosa. Esta forma de riesgo moral se observa cuando existe acumulación de pequeños accidentes que son reportados en un solo evento o mediante la colusión con el taller mecánico. En el caso del seguro de salud, tener una cobertura generosa disminuye el precio de los cuidados de salud ocasionando aumento de los gastos por parte de los asegurados. (Bardey, 2008)

## **b) Los Problemas de Selección Adversa en el Mercado de Seguros**

La selección adversa se manifiesta previa a la firma de un contrato y se da ante la desigualdad de información entre los entes involucrados sobre la calidad del bien u otros aspectos. Donde el principal solo puede conocer las características exógenas para determinar qué tipo es el agente; sin embargo, el agente realmente conoce de sí mismo; por ello, con qué tipo el principal está tratando. Este tipo de problema de asimetría de información se le denomina “tipo oculto” (Dammert Lira, Molinelli Aristondo y Carbajal Navarro, 2013)

Este fenómeno se origina en una ventaja informacional de los asegurados sobre los aseguradores. Los asegurados disponen de información relativa a su riesgo que el asegurador no puede observar, como su patrimonio genético, su estado de salud, sus habilidades para manejar, entre otros. La selección adversa surge entonces de la incapacidad de los aseguradores para identificar

perfectamente la clase de riesgos al cual pertenecen sus asegurados. Esta incapacidad obliga a fijar primas más o menos en función del riesgo promedio calculado a nivel de toda la población o al interior de categorías de riesgo fácilmente observables. (Bardey, 2008)

Rothschild y Stiglitz (1976) plantearon un modelo básico para analizar la asimetría de información en el mercado asegurador. Este modelo considera que un individuo posee un ingreso de tamaño  $W$  si tiene la suerte de evitar un accidente. En caso de que se produzca un accidente, sus ingresos serán sólo  $W - d$ . El individuo puede asegurarse contra este accidente pagando a una compañía de seguros una prima  $\alpha_1$ , a cambio de la cual se le pagará  $\alpha_2^*$  si se produce un accidente. Sin seguro sus ingresos en los dos estados, "accidente", "sin accidente", eran  $(W, W - d)$ ; con el seguro es ahora  $(W - \alpha_1, W - d + \alpha_2)$ , donde  $\alpha_2 = \alpha_2^* - \alpha_1$ .

Para describir cómo funciona el mercado, es necesario describir las funciones de oferta y demanda de los participantes en el mercado. Sólo hay dos tipos de participantes, individuos que compran seguros y compañías que lo venden. Determinar la demanda individual de contratos de seguro es sencillo. Un individuo compra un contrato de seguro para alterar su patrón de ingresos en todos los estados de la naturaleza. Sea  $W_1$  el ingreso sin accidente y  $W_2$  el ingreso si ocurre un accidente; el teorema de utilidad esperado establece que bajo supuestos relativamente leves sus preferencias de ingresos en estos dos estados de la naturaleza se describen por una función de la forma,

$$(1) \quad \widehat{V}(p, W_1, W_2) = (1 - p)U(W_1) + pU(W_2),$$

Donde  $U(\cdot)$  representa la utilidad de los ingresos monetarios y  $p$  la probabilidad de un accidente. Las demandas individuales pueden derivarse de (1). Un contrato  $\alpha$  vale  $= \widehat{V}(p, W - \alpha_1, W - d + \alpha_2)$ . De todos los contratos ofrecidos al individuo, este elige el que maximiza  $V(p, \alpha)$ . Puesto que siempre tiene la opción de no comprar el seguro, un individuo comprará un contrato  $\alpha$  sólo si  $V(p, \alpha) \geq V(p, 0) = \widehat{V}(p, W, W - d)$ . Se asume que las personas son idénticas en todos los aspectos salvo su probabilidad de tener un accidente y que son reacias al riesgo ( $U'' < 0$ ); por lo tanto  $V(p, \alpha)$  es cuasi-cóncavo.

Las compañías de seguros deciden qué contratos deben ofrecer para la venta y a qué personas. El retorno de un contrato de seguro es una variable aleatoria. Se asume que las empresas son neutrales al riesgo ya que sólo se preocupan por los beneficios esperados, de modo que cuando el contrato  $\alpha$  se vende a un individuo que tiene una probabilidad de incurrir en un accidente de  $p$ , es beneficioso.

$$(2) \quad \pi(p, \alpha) = (1 - p)\alpha_1 - p\alpha_2 = \alpha_1 - p(\alpha_1 + \alpha_2)$$

Incluso si las empresas no maximizan sus beneficios esperados, en un mercado competitivo bien organizado es probable que se comporten como si se maximizara (2). Las compañías de seguros tienen recursos financieros de tal manera que están dispuestas y capaces de vender cualquier número de contratos que crean que obtendrán un beneficio esperado. El mercado es competitivo en el sentido de que hay entrada libre. En conjunto, estos supuestos garantizan que cualquier contrato que se exija y que se espera que sea rentable sea sustituido.

Asimismo, para conocer cómo los clientes y las empresas llegan a conocer o estimar el parámetro  $p$ , que desempeña un papel tan crucial en las fórmulas de valoración (1) y (2). Se realiza la suposición de que los individuos conocen sus probabilidades de accidente, mientras que las empresas no. Dado que los compradores de seguros son idénticos en todos los aspectos, salvo su propensión a tener accidentes, la fuerza de esta suposición es que las empresas no pueden discriminar entre sus clientes potenciales en función de sus características.

Una empresa puede utilizar el comportamiento de mercado de sus clientes para inferir sobre sus probabilidades de accidente. Por consiguiente, aquellos clientes con altas probabilidades de accidentes exigirán más seguro que aquellos que son menos propensos a accidentes; aunque posiblemente sea preciso, esta no es una manera rentable de averiguar acerca de las características del cliente. Las empresas de seguros quieren conocer las características de sus clientes para decidir qué términos deben ofrecer para permitirles comprar un seguro. A veces es posible obligar a los clientes a tomar decisiones de mercado de tal manera que ambos revelen sus características y tomen las decisiones que la empresa hubiera querido que tomaran, este dispositivo de mercado es un mecanismo de autoselección.

#### - **Los Problemas de Signaling**

Los problemas de señales se presentan cuando el agente, previa negociación del contrato, identifica su tipo y por lo tanto toma ciertas decisiones que le reflejan al principal dichas características. (Loteró, 2005)

Los mercados con información imperfecta difieren de los de información completa en que, con información imperfecta, las acciones o elecciones realizadas en el mercado pueden transmitir u ocultar información. Por ejemplo, una persona suscribe una póliza de seguros para trasmitirle a la compañía de seguros su creencia de que la probabilidad de accidentarse es baja. (Stiglitz, 2002)

#### - **Los Problemas de Screening**

Los problemas de screening surgen cuando la parte no informada debe diseñar un conjunto de alternativas entre las que permite elegir a la parte informada. En este caso el diseño de las alternativas debe ser tal que los distintos tipos de agentes existentes en la parte informada se auto clasifiquen o se separen según sus elecciones entre las distintas alternativas. (Usategui, 1999)

#### **2.1.4. La teoría de la Decisión**

La teoría de la decisión se ocupa de analizar cómo elige una persona aquella acción que, de entre un conjunto de acciones posibles, le conduce al mejor resultado dadas sus preferencias. Por otro lado, en los últimos años ha sido tan grande su influencia en disciplinas como la psicología y la economía, que han contribuido con la matemática aplicada, a su desarrollo. (Aguiar, 2004)

En tal sentido, esta teoría busca explicar la razón del comportamiento de un individuo y si fue el resultado de un proceso razonado o si se trató de una reacción emocional. Para diferenciar frente a las causas del comportamiento, muchos autores usan el término “elección individual” que sugiere algún nivel de razonamiento. No obstante, importa que el comportamiento sea estable ya que sino la información pasada no resulta útil para predecir el efecto de las nuevas circunstancias sobre el comportamiento. (Vial y Zurita, 2011)

Becker (1977) señala que las familias y empresas no maximizan, por lo menos no consistentemente, ni coherentemente y que las preferencias no son bien ordenadas, además que la teoría no es útil en un comportamiento explicativo. Asimismo, teoremas económicos han mostrado también ser consistentes con una amplia clase de conducta irracional. Es decir que la teoría económica en muchos casos es más compatible con la conducta irracional de lo que se había sospechado previamente.

##### **a) Decisiones Bajo Certidumbre**

Es aquella en la que los tomadores de decisión tienen certeza del estado de la naturaleza que va a ocurrir y las consecuencias de cada una de las alternativas que implica la selección de la decisión. En tal sentido, escogen la alternativa que maximice su bienestar o minimice su pérdida. Se conoce con exactitud el resultado para cada decisión, por lo tanto, se pueden enumerar todos los rendimientos en una sola columna y considerarla como la representación de un estado de la naturaleza del cual estamos seguros de que ocurrirá. (Eppen, Gould, Schmidt, Moore y Weatherferd, 2000)

##### **b) Decisiones Bajo Riesgo**

Este modelo supone que hay muchos estados en la naturaleza y que quien toma la decisión puede obtener una estimación de probabilidades de la ocurrencia de cada uno de estos estados. Asimismo, no se enfoca sólo en los resultados, sino también en la cantidad de riesgo de cada una de las decisiones. Las acciones se basan en resultados esperados y el tomador de decisiones tiene una cierta información sobre la posibilidad de los estados de la naturaleza y traduce esto en una distribución de probabilidad, siendo el estado de la naturaleza una variable aleatoria. (Eppen, Gould, Schmidt, Moore y Weatherferd, 2000)

### **c) Decisiones Bajo Incertidumbre**

Este modelo también supone que hay muchos estados posibles en la naturaleza, pero en este caso el tomador de decisiones no puede especificar las probabilidades de ocurrencia de esos estados. Hay información deficiente para tomar decisiones y no se tiene ningún control sobre la situación ni se conoce la interacción de las variables. El decisor conoce cuáles son los posibles estados de la naturaleza, pero no dispone de información alguna sobre qué posibilidades de ocurrencia y además no puede cuantificar de ninguna manera esta incertidumbre. (Eppen, Gould, Schmidt, Moore y Weatherford, 2000)

#### **2.1.5. Probabilidad de siniestro**

El siniestro cuyo acaecimiento se protege con la póliza debe «poder suceder». Tal posibilidad o probabilidad tiene dos limitaciones extremas: de un lado, la frecuencia; de otro, la imposibilidad. La excesiva reiteración del riesgo y su materialización en siniestros atenta contra el principio básico del seguro que es la aleatoriedad referido la incertidumbre relativa respecto a su existencia. Una gran frecuencia de siniestros en el seguro de automóviles convertiría a una empresa aseguradora en una de servicios de conservación o reparación de vehículos que, lógicamente, podría ser prestado, pero en tal caso su precio no sólo sería más elevado, sino que tendría una naturaleza completamente distinta. Del mismo modo, la absoluta imposibilidad de que el riesgo se manifieste en siniestro situaría a las entidades aseguradoras en una posición privilegiada, al percibir unos ingresos no sujetos a contraprestación, lo cual resultaría tan absurdo como la reiteración continua de siniestros. (Fundación MAPFRE, 2020)

#### **2.1.6. Factores relativos a la conducta aleatoria del individuo**

La teoría económica tradicional supone que las personas tienen un buen acceso a la información y pueden procesarla perfectamente. También, se asume que los agentes económicos siempre ejecutan sus planes y que actúan a beneficio personal. Este modelo simplificado de comportamiento humano ha ayudado a los economistas a formular teorías que han proporcionado soluciones a problemas económicos importantes y complejos. La evidente diferencia entre la teoría y la realidad hizo posible la introducción de factores distintos a los tradicionales tales como racionalidad limitada, percepciones sobre la equidad y falta de autocontrol. (Castaño, 2018)

Según la economía del comportamiento, el modelo de comportamiento humano basado en la premisa de que las personas optimizan es y siempre ha sido inverosímil. (Thaler, 2018) Gran parte de la economía del comportamiento o economía de la conducta surge del hecho de que las personas tienen



un lado emocional e irracional que los modelos económicos tradicionales no logran incorporar. En tal sentido, en las ciencias modernas el hombre moderno constituye un tomador de decisiones asediado por la incertidumbre en todos los frentes (Levine, 2010)

Sea la aleatoriedad definida como la característica fundamental de fenómenos imprevisibles y la variable aleatoria como la que puede asumir distintos valores según el azar estadístico. (Diaz, 1991) De lo antes mencionado, ante un escenario de falta de información, no se puede tener la certeza que las acciones de un individuo serán necesariamente estables es decir puede existir una aleatoriedad intrínseca en el comportamiento.

### 2.1.7. Factores relativos al vehículo asegurado

Se define al carro como el vehículo o armazón con ruedas que se emplea para transportar objetos diversos como el cesto de la compra, libros, comida y equipaje. (Real Academia Española, 2020)

Todos los vehículos deben de contar con un seguro así ordena la ley (SOAT). Sin embargo, existen otro tipo de seguros particulares de autos cuyo precio depende de una serie de factores relacionados con el conductor, el vehículo y la cobertura demandada, entre otros. (APESEG, 2019)

- **Año del vehículo:** los vehículos recién matriculados pagan menores primas que los que están matriculados hace años. En otras palabras, el precio de seguro de un vehículo será mayor en el mercado si se compra un vehículo de segunda mano.
- **Monto asegurado:** como resulta lógico cuando más haya costado el vehículo y las coberturas incluidas en la póliza de seguro abarquen más, mayor será el precio del seguro del vehículo, es necesario conocer el precio del vehículo o el monto asegurado ya que la aseguradora corre mayor riesgo cubriendo este vehículo. Ya que tiene que ver el valor, los costes de reparación, la potencia máxima o si se considera un robo deseable.
- **Tipo de vehículo:** hace referencia a la capacidad del vehículo y uso que le dará el asegurado, de acuerdo a sus necesidades, se afrontará una prima superior según se elija un vehículo con más caballos fuerza, ya que se considera que con la potencia del vehículo se incrementan las posibilidades de sufrir un accidente, por lo que se eleva el precio del seguro del vehículo en esos casos.

- **Marca de vehículo:** la marca, modelo, número de asientos, puertas, plazas, tipo de vehículo. Según sea sus características, las aseguradoras prestan atención a esos detalles a la hora de calcular el precio de seguro de vehículo, cuando el vehículo sea de más cilindros mayor será la prima de seguro.

### 2.1.8. Factores relativos a la póliza

Se define a la póliza de seguro como el conjunto de documentos en los que se describen las condiciones del contrato de seguro. Está compuesto por:

- Condiciones generales: Cláusulas establecidas por la aseguradora para cada tipo de seguro que son iguales para todos los contratantes.
  - Condiciones particulares: Contienen información específica sobre el contratante y el riesgo asegurado, como: identificación de las partes, designación del asegurado y el beneficiario, descripción de la materia asegurada, la suma asegurada o el alcance de la cobertura y monto.
  - Condiciones especiales: Disposiciones adicionales que amplían, reducen, aclaran y en general, modifican el contenido o efectos de las condiciones generales o particulares.
  - Anexos de la Póliza: Por ejemplo, declaraciones de salud, entre otras.
  - Endosos de la póliza: Modificaciones o nuevas declaraciones del contratante efectuadas luego de la emisión de la póliza. (Superintendencia de Banca y Seguros, 2020)
- **Antigüedad de la póliza:** Las compañías aseguradoras ofrecen bonificaciones en el precio de seguro del vehículo a los clientes antiguos que siempre renuevan y cuentan con el seguro de esta compañía y son buenos conductores.
  - **Vigencia de la póliza:** Las pólizas de seguro por lo general mantienen una vigencia anual si la duración periódica del tiempo previsto en la póliza de seguro y sus efectos de cobertura se extiende en su periodo el coste será mayor.

### 2.1.9. Factores relativos al conductor

Se define al conductor como la persona que con mayor frecuencia utiliza el vehículo asegurado su experiencia siniestral es un elemento para determinar la prima del seguro. (Pacífico Seguros, 2020)

- **Sexo del asegurado:** Este factor también es importante para las compañías porque, estadísticamente, el sexo masculino tiene más accidentes o mayor frecuencia de conducir en

estado étlico, bien debido a que requieren mayor número de desplazamientos porque está asociado a nivel de frecuencia que utilizan el vehículo.

- **Edad del asegurado:** Por lo general, los conductores maduros tienen menos accidentes que los conductores menos experimentados, si la edad del conductor es relativamente menor a 25 años, el precio del seguro del vehículo que se debe abonar será más elevado, ya que las estadísticas revelan que los jóvenes entre 18 y 25 años frecuentan tener más accidentes.

Los seguros vehiculares se han convertido en lo indispensable antes de adquirir un vehículo, los conductores experimentados antes de entrar al volante deben considerar que la máquina puesta en marcha cuenta con una póliza de seguro, por lo general los propietarios de estos vehículos buscan ahorrar en el seguro buscando compañías que le ofrezcan primas reducidas. Pero las compañías aseguradoras antes de ofrecer su servicio buscan conocer una serie de factores relacionados con el vehículo asegurado, el conductor y la póliza de seguro. (Kelisto Seguros, 2019)

#### **2.1.10. Términos relacionados a seguros**

##### **a) Contrato de Seguro**

Contratar un seguro tiene por finalidad proteger tu patrimonio ante el riesgo que algún evento afecte tus bienes, salud o vida, de modo que sea una empresa aseguradora quien se encargue de cubrir los gastos generados o pagar una indemnización, según lo acordado en el contrato y lo dispuesto por la Ley del Contrato de Seguros, Ley N° 29946. (Superintendencia de Banca y Seguros, 2020)

##### **b) Riesgo**

El riesgo es la posibilidad de que ocurra una situación imprevista que pueda ocasionar daños a la persona o afectar sus bienes. La posibilidad de sufrir una pérdida o daño, puede o no ocurrir, pero si se diera, esta ocasionaría gastos para los que no se estaría preparado por consiguiente se vería afectada su economía. (Superintendencia de Banca y Seguros, 2020)

Picos y Valero (2007) señalan que un riesgo asegurable debe cumplir ciertas condiciones:

- Posible e incierto y aleatorio o que depende del azar.
- Su realización debe producir una necesidad patrimonial.
- Debe afectar por igual a todos los elementos del colectivo asegurable.

- Deber ser lícito y no producir lucro al asegurado.
- Susceptible al tratamiento estadístico.
- El riesgo que se cubre debe estar perfectamente determinado.
- Los riesgos a cubrir deben presentar una regularidad en su comportamiento y frecuencia, como en cuanto a la intensidad del daño económico que causan en su acaecimiento. Tal comportamiento ha de adaptarse a una determinada "Ley de regularidad estadística", que permita su tratamiento actuarial. Los riesgos "esporádicos" y los "catastróficos o extraordinarios" requieren un tratamiento muy especial.
- Agrupación de riesgos dada la infinidad a los que se puede estar expuesto. La delimitación hace que se elimine la acumulación de estos ante un siniestro ocurrido que haga que varios puedan estar afectos a pérdidas y haga que se originen pérdidas no previstas.
- Debe haber homogeneidad en la cartera en cuanto suma asegurada de manera que se pueda poner límites máximos de cobertura o bien cediendo en reaseguro los excesos de su capacidad.

Fortiana, Claramount y Boj del Val (1999) definen al riesgo como como el acontecimiento incierto, independiente de la voluntad exclusiva de las partes y cuya realización implica, normalmente, consecuencias desfavorables para el asegurado.

### **c) Siniestro**

El siniestro es la manifestación concreta del riesgo asegurado, a consecuencia de la cual, se producen daños cuya indemnización está garantizada por el contrato de seguro, motivando que la aseguradora satisfaga total o parcialmente al asegurado o a sus beneficiarios, el monto del capital asegurado. (Superintendencia de Banca y Seguros, 2020)

### **d) Prima de seguros**

La prima de seguros es el precio del seguro que paga el asegurado en el momento de la emisión de la póliza. La prima es por lo general para una vigencia anual del seguro, aunque puede excepcionalmente pagarse la prima por una sola vez para una cobertura de varios años (prima única en seguros de vida) y también por una vigencia menor de un año (prima a corto plazo, como para un viaje, seguro de transportes de mercancías, entre otros). (APESEG, 2019)

## 2.2. ANTECEDENTES DE LA INVESTIGACIÓN

La literatura sobre los aspectos teóricos de la asimetría de información fue muy amplia desde los años setenta. Sin embargo, fue necesario esperar el final de los años ochenta y el principio de los noventa para que los economistas se dedicaran a la medición de estos fenómenos. (Bardey D. , 2008)

### 2.2.1. Antecedentes empíricos:

En el mercado asegurador, el problema principal - agente al momento de decidir otorgar o no un seguro implica tomar en cuenta factores exógenos que no podrán ser controlados por el agente y que no forman parte del conocimiento del principal (Vásquez, 2015) .

A inicios del siglo XXI, en el auge de los mercados de automóviles en los países europeos, los estudios empíricos sobre la asimetría de la información se centraron en este mercado. Melgar y Guerrero (2005), en “Los siniestros en el seguro del automóvil: un análisis econométrico aplicado”, tuvieron como objetivo determinar los principales factores de la cantidad de siniestros del seguro automovilístico, así verificar el cumplimiento de la teoría de la relación directa entre altas coberturas y mayor cantidad de ocurrencia de siniestros. Se analizaron datos de una aseguradora española mediante un modelo binomial negativo inflado de ceros donde se encontró que los factores más relevantes para explicar la siniestralidad vehicular son: algunas categorías de vehículos (camiones, remolques, ciclomotores y motos); algunos usos (agrícola propio, industrial y el transporte escolar); la edad; la experiencia en la conducción; la prima pagada y el grado de cobertura. Finalmente, se demostró la previsión del mercado sobre la correlación positiva entre la siniestralidad de los asegurados y la cobertura que eligen.

Mateus (2008) en su investigación “Análisis de los seguros de automóviles en Colombia entre 1997 y 2005 de la empresa Liberty Seguros S.A. basado en el nivel de riesgo y tasa de siniestralidad” busca mantener los márgenes de utilidad y atención controlados a través de la medición del riesgo y la siniestralidad. Entre los resultados encontrados, se puede verificar estadísticamente mediante un modelo de tipo *count data* que los factores que influyen más en la culpabilidad de los asegurados en los siniestros son: el sexo, la zona de circulación, el tramo de la edad, la antigüedad del vehículo y la cobertura de la póliza; de donde se puede derivar la existencia de un componente de riesgo moral y/o selección adversa. Este último explica que cuando el asegurado ya obtiene las garantías de protección del auto muestra una conducta de descuido, siempre y cuando la cobertura elegida por la aseguradora cubra el riesgo de siniestro.

Melgar y Ordaz (2013) en su trabajo de investigación titulado “Aplicación del análisis discriminante al estudio de la siniestralidad en el ramo del seguro de automóviles” explican que indagar en las características de los asegurados que estén relacionadas con la incidencia de siniestros en el sector de seguros de autos pueden ser muy beneficioso para sí mismo. Este estudio fue realizado mediante los datos registrados de una aseguradora multinacional que opera en el mercado español. La metodología empleada para procesar dichos datos fue a través del análisis discriminante, entre las variables utilizadas en las estimaciones del modelo se encuentra: la edad, el sexo, el tipo y el uso del auto, la antigüedad de la licencia de conducir y la póliza, la prima y el grado de cobertura. Finalmente, entre las principales conclusiones se muestra que el grado de cobertura, los tipos de vehículos y la prima pagada guardan mayor relación con la función discriminante, esto indica que estas variables explican significativamente la adecuada calificación del seguro. Asimismo, los asegurados que tengan edad por encima del promedio que conduzcan camión, moto, ciclomotor o vehículos especiales son más propensos a no incurrir en un siniestro.

García (2014) en su estudio titulado “Modelo de regresión y estadística, para la siniestralidad de autos de una empresa, dentro de la vigencia 2012-2013” pudo identificar mediante los resultados de un modelo de regresión con variables dicotómicas que las variables que explican de manera significativa las causas más frecuentes de siniestralidad tales como colusiones y roturas de cristales, de un cliente de la cartera de Lockton México, fueron: la edad, el tipo de pérdida (si la pérdida del vehículo es total el costo del siniestro es mayor) y el modelo del vehículo, para dichas variables las conclusiones pudieron evidenciar estadísticamente que los parámetros eran distintos a cero a un nivel de significación de 5%. Por otro lado, las variables explicativas como sexo, tipo de vehículo y algunos modelos no cumplieron la significación individual en el modelo estimado; por ello se puede inferir que dichas variables no explican la incidencia de siniestro en el sector de autos.

Quishpe (2015) en su trabajo de tesis “Factores de riesgo de siniestralidad y cálculo de primas de los vehículos asegurados en el Ecuador mediante modelos lineales generalizados”, en base a un análisis multivariante de regresión logística que asume como “individuo Tipo” o Categoría Base a “una mujer entre 18 y 25 años de edad, que posee un automóvil marca Chevrolet, con años de fabricación entre el 2012 al 2014, con un valor asegurado de USD 10.000 a USD 15.000, que circula en la ciudad de Quito, con una póliza renovada, suscrita durante un año”; ha encontrado que los factores que aumentan la probabilidad de ocurrencia de siniestros son, en orden de influencia: tener vehículos con años de fabricación menores al 2012, de marcas diferentes a Chevrolet y contratar seguros por montos entre USD 15.001 y USD 50.000. Mientras que los factores que disminuyen dicha probabilidad son: tener vehículos diferentes a los automóviles, la edad del conductor del vehículo es superior a los 25 años, y que las zonas de circulación sean Cuenca, Santo Domingo, Riobamba, Loja, Ambato, Coca, Ibarra, Manta, Portoviejo o Machala. Cabe resaltar que la variable “Sexo del asegurado” en los resultados econométricos no tiene significancia estadística.

Segovia Vargas, Camacho Miñano, Pascual Ezama y Rodríguez Piñero (2015) en su investigación: “Análisis de los factores de riesgo en el seguro de automóvil mediante ecuaciones estructurales” estudiaron la capacidad de explicación de los factores de riesgo cualitativos y cuantitativos tales como: género, región, combustible, asientos, edad del vehículo, edad del asegurado, bonus malus (BM) entre otros sobre la variable accidente. El estudio planteó que los factores escondidos son revelados, en parte, por el número de siniestros acreditados por los asegurados y se reflejan en la variable BM. Por lo tanto, se plantearon dos modelos de ecuaciones estructurales sin considerar la variable BM y posteriormente consideraron la variable BM. Los resultados de la investigación fueron que el primer modelo, en el que no se incluye la variable BM, tiene una capacidad predictiva sobre los siniestros del 84%. Para el segundo modelo, en el que se incluye la variable BM, aumenta su capacidad explicativa hasta el 86%. Por lo tanto, la inclusión de la variable BM incrementa el poder explicativo del modelo en un 2%. Finalmente se concluye que existen muchas variables importantes que podrían incrementar la probabilidad de predecir accidentes y que no se tienen en cuenta de forma directa.

Vargas (2016) en su estudio: “La siniestralidad del seguro vehicular y las reservas técnicas de Sweden Seguros sucursal Ambato en el año 2014”, estableció en qué medida la siniestralidad impacta en los beneficios de las empresas aseguradoras, al obligar algunas veces a sustraer fondos de sus reservas técnicas y cómo la ocurrencia de siniestros provoca consecuencias negativas en los balances financieros de las empresas aseguradora cuando el riesgo no ha sido estimado adecuadamente en los recursos y finanzas. El modelo de análisis fue el de frecuencia y severidad que da como resultados las tasas más acertadas para el otorgamiento de las pólizas, tomando en cuenta las variables: condiciones de la póliza, deducibles, uso, edad y género. Las conclusiones obtenidas mostraron que la edad y el sexo del asegurado tienen mayor impacto en la siniestralidad ante el uso de 2 métodos, también se pudo evidenciar cómo el acertado cálculo de la del riesgo contribuye a gestionar mejor las utilidades de la empresa.

## **2.2.2. Otros antecedentes de la Investigación**

### **a) Historia del Seguro**

La historia del seguro se remonta a las antiguas civilizaciones griegas, romanas, babilónicas e hindúes quienes realizaban contratos a la gruesa financiando pérdidas. Años después, en la edad media, las asociaciones religiosas colectaban y distribuían fondos entre sus miembros en caso de muerte de alguno de ellos, (Specia Jimenez, 2005)

La presencia de esquemas para la protección de la vida del hombre se advierte desde el Imperio Babilónico y su Rey Hammurabi (1810 – 1750 a.c.) mediante el Código de Hammurabi que

indemnizaba a las esposas y descendientes, en caso de la muerte del cónyuge. (Junguito, s.f)

Durante la Edad Media los piratas vagaban por los mares capturando a las tripulaciones de los barcos para cobrar rescate a fin de salvaguardar sus vidas; de lo contrario, estos eran obligados a arrojarse al mar. (Gran Enciclopedia del Mundo, 1978). En 1347, aparece el primer contrato de seguro marítimo en Italia y cubría posibles accidentes, naufragios o acciones de piratas. En 1385 aparece la primera póliza, escrita en italiano. (APESEG, 2019)

La etapa del seguro moderno, aparece en Hamburgo en el año 1500 con la existencia de unas cajas especiales de propietarios cuyo fin era agruparse para socorrerse en caso de incendio. (Enciclopedia ENCARTA, 2004). En 1710, tras el gran incendio en Londres, se fundó Fire Office. En este momento surgió la idea de compensar las pérdidas financieras en escenarios como estos. Se considera a Fire Office la primera compañía aseguradora contra incendios. (APESEG, 2019)

Para mediados del siglo XIX, el concepto del rol de los seguros y del contrato del seguro estaba muy avanzado, se destaca: el carácter condicional y aleatorio del seguro, la necesidad de establecer el valor de la cosa asegurada y la prima prometida a los aseguradores, se señala que todos los bienes y cosas con probabilidad de dañarse pueden asegurarse. Finalmente, el contrato de seguros establece las obligaciones para la aseguradora y el asegurado advirtiendo problemas de fraude y sus implicancias. (Junguito, s.f)

En tal sentido, el hombre siempre se ha visto amenazado en distintos escenarios de su vida, ya sea en su integridad, salud, dinero o propiedades. Por eso, ha buscado soluciones que puedan ayudarlo a apaciguar las consecuencias de siniestros e imprevistos. (APESEG, 2019)

## **b) Historia del Seguro en el Perú**

En el Perú, una de las primeras Cías. que ofreció seguros de vida fue "La Providencia", en 1863. Años más tarde, el gobierno de Nicolás de Piérola (1895-1899), proporcionó un marco económico para las Cías. a fin de nacionalizar los seguros de incendio y marítimo. En 1937, se delega la fiscalización de las aseguradoras a la Superintendencia de Bancos y en el año 1971, se designa al Banco de la Nación la realización de operaciones de reaseguro. En el año 1987, el Senado frena la estatización del sistema financiero decisión del presidente Alan García, entre ellos 17 Cías. de Seguros. (López Martínez, 2004)

Luego de la liberalización el sector asegurador se benefició de externalidades positivas como una inflación controlada, el crecimiento económico del país, la exoneración del Impuesto General a las Ventas (IGV) a los seguros de vida individual y la creación del Sistema Privado de Pensiones (SPP), las normas dadas por la Superintendencia de Banca y Seguros (SBS), la creación del Seguro



Obligatorio contra Accidentes de Transite (SOAT) el 2002 y la aparición de nuevas empresas que modernizaron los canales de distribución que permitieron a las compañías orientarse hacia la oferta de estas pólizas y aumentar la penetración del sector de seguros. (Andrea, Cerruti, y Tapia, 2007)

### **c) Historia del Seguro Vehicular**

En 1821, se tiene el primer antecedente relacionado a una ordenanza dictada en París donde cada cochero pagaba 20 céntimos diarios de su salario para un fondo de cobertura de multas y reparaciones de daños a terceros. En 1825, se creó La Automedon que era una sociedad para cubrir la responsabilidad civil de los cocheros franceses. Sin embargo, en 1844, el Tribunal de Comercio del Sena declaró nulos estos seguros argumentando que el hecho de estar asegurados provocaría falta de atención en los conductores. (Rodríguez Basurto, 2007)

Los seguros reaparecieron y se expandieron con fuerza hasta finales del siglo XIX, momento en el que otro descubrimiento revolucionó el mundo. En 1886, Karl Benz registró el automóvil, en la Oficina de Patentes de Berlín. Poco a poco, las pólizas se fueron adaptando al nuevo sistema de transporte. (CMS Seguros, 2013)

El primer seguro de coches registrado tuvo lugar en Estados Unidos el 1 de febrero de 1898, pese a que en Reino Unido ya tenía una póliza exclusiva para autos. Sin embargo, los motorizados eran un lujo raro, tanto que todavía se experimentaba con motores a vapor. La compañía a cargo de la venta habría sido Travellers Insurance. (Seguros TV, 2016)

El primer comprador de este seguro habría sido Truman J. Martin, un doctor de Nueva York, pagando 12,25 dólares por una cobertura de hasta 5.000 dólares. El coche estaba asegurado, frente a colisiones con los 18 millones de carruajes y jinetes existentes en Estados Unidos y no frente a otros coches ya que en ese momento tan sólo había 4.000 en esa región. (St. Onge, 2008)

En España, por ejemplo, hubo que esperar hasta 1912 para poder ver la primera normativa de seguros privados y no fue hasta la década de 1930 cuando en Reino Unido se introdujo un primer experimento de seguro de automóvil obligatorio, extendiéndose a partir de entonces por todo el mundo. (Revista de Historia , 2019)

### **d) Historia del Seguro Vehicular en el Perú**

En el Perú, en el año 1900 llega el primer auto a vapor. Años más tarde en 1915, se emite la primera póliza de un seguro vehicular por la compañía de seguros Internacional, (hoy Rimac). En el

2002 se creó el SOAT con un fin netamente social que asegura la atención, de manera inmediata e incondicional a las víctimas de accidentes de tránsito en todo el territorio peruano. (Agois, 2016)

El ramo de seguros vehiculares peruano tiene por objeto la prestación de indemnización derivadas de accidentes producidos a consecuencia de la circulación de vehículos. La legislación de algunos países diferencia el Seguro Obligatorio, destinado a la cobertura, dentro de los límites legalmente establecidos, de los daños personales o materiales causados a terceras personas, y el Seguro Voluntario, que cubre el exceso de los límites del Seguro Obligatorio, así como otras garantías. Se trata realmente de una cobertura combinada, en la que se incluyen riesgos de daños, Responsabilidad Civil y Defensa Jurídica. (Guardiola Lozano, 1990)

#### **e) El seguro obligatorio de accidentes de tránsito (SOAT)**

El Seguro Obligatorio de Accidentes de Tránsito, vigente desde el 2002, es el seguro que garantiza de manera incondicional el pago de gastos médicos e indemnizaciones a causa de un accidente de tránsito según lo establece la Ley General de Transporte y Tránsito Terrestre. Asimismo, señala que no sólo cubre al asegurado sino también a los ocupantes del vehículo y no se requiere pronunciamiento de ninguna autoridad para que los hospitales y clínicas atiendan a los heridos; es decir, el SOAT se activa sin investigación de por medio y al margen de las causas de un accidente. (Ministerio de Transportes y Comunicaciones, 2019)

El SOAT es un seguro de responsabilidad hacia terceros que cubre al vehículo registrado en la póliza. Pero no cubre daños materiales –por ejemplo, un auto impactado- solo daños a las personas. El SOAT tiene una cobertura límite de 5 UIT para gastos médicos, 4 UIT como indemnización por invalidez permanente o fallecimiento y de 1 UIT si la invalidez es temporal o para gastos de sepelio. (APESEG, 2019)

Las indemnizaciones del SOAT pueden ser cobradas hasta dos años después del accidente. Vencido ese plazo, las indemnizaciones no cobradas son transferidas al Fondo de Compensación del SOAT bajo dominio del Ministerio de Transportes y Comunicaciones. Este fondo se usa para cubrir gastos médicos y de sepelio cuando el causante de un accidente huye del lugar. En el 2015, este fondo cubrió 324 siniestros por casi un millón y medio de soles. (APESEG, 2019)

Las aseguradoras que ofrecen el SOAT son siete: La Positiva, Rímac, Pacífico, Interseguro, Mapfre, Protecta, Cardif.

## 2.3. CONTEXTO DE LA INVESTIGACIÓN

El mercado de seguros en los países latinoamericanos está directamente relacionado con el crecimiento del PBI, es decir, el comportamiento positivo de la economía genera que los agentes realicen más actividades empresariales; las mismas que deberán ser cubiertas ante un posible riesgo, mediante la adquisición de un seguro. Esto sumado al aumento del poder adquisitivo explica el crecimiento del mercado de seguros. No obstante, el crecimiento de primas implica también crecimiento de la siniestralidad. Asimismo, una desventaja de las economías latinoamericanas es la informalidad, las mismas que dificultan la penetración del seguro. (Equilibrium Clasificadora de Riesgo, 2018)

### 2.3.1. El Mercado Asegurador Peruano

Para el año 2018, el mercado asegurador peruano se recuperó y las primas de seguros crecieron 13.6% mostrando un panorama alentador para el sector como resultado del impulso de la economía por parte del gobierno mediante la aceleración de la reconstrucción de los daños causados por el fenómeno del Niño. Asimismo, la siniestralidad se redujo en 15.4% para el año 2018 con 6,879 millones por debajo del pico histórico registrado el año anterior debido a las indemnizaciones pagadas por el fenómeno del Niño Costero que afectó Lima y las regiones de la Costa Norte. (Rindebro, 2018)

En el año 2019, “La desaceleración económica, así como la incertidumbre política, hacen que se reduzcan las compras de servicios, productos o la necesidad de generar un ahorro a largo plazo; así como también hace que se frenen las inversiones públicas y privadas. Este contexto no ha sido ajeno a la industria aseguradora que, si bien ha mostrado dinamismo, el crecimiento no alcanzó al del año anterior”, señaló Eduardo Morón, presidente de la Asociación Peruana de Empresas de Seguros, APESEG.

Para cierre del 2019, el sistema de seguros peruano estaba compuesto por 20 empresas<sup>1</sup> de las cuales 7 se encuentran dedicadas a ramos generales y de vida, 8 exclusivamente a ramos generales y 5 exclusivamente a ramos de vida. Asimismo, la composición de la producción estuvo liderada por los Ramos Generales con una participación de 36,8%. A su vez, los Seguros de Vida registraron una participación de 34,1%.<sup>2</sup> (Superintendencia de Banca y Seguros, 2019)

---

<sup>1</sup> Anexo 2: Empresas de Seguros Autorizadas al 31 de diciembre del 2019. Tomado de Evolución del Sistema de Seguros al IV trimestre de 2019 por Superintendencia de Banca y Seguros (SBS)

<sup>2</sup> Anexo 3: Estructura de las Primas de Seguros Netas por Ramos. Tomado de Evolución del Sistema de Seguros al IV trimestre de 2019 por Superintendencia de Banca y Seguros (SBS)

Las Primas de Seguros Netas (PSN) del sistema asegurador ascendieron a S/ 14 113,5 millones a setiembre de 2019, lo que representa un aumento de 9,7% respecto de lo registrado en el mismo mes de 2018, cuando las PSN ascendieron a S/ 12 868,7 millones<sup>3</sup>. (Superintendencia de Banca y Seguros, 2019)

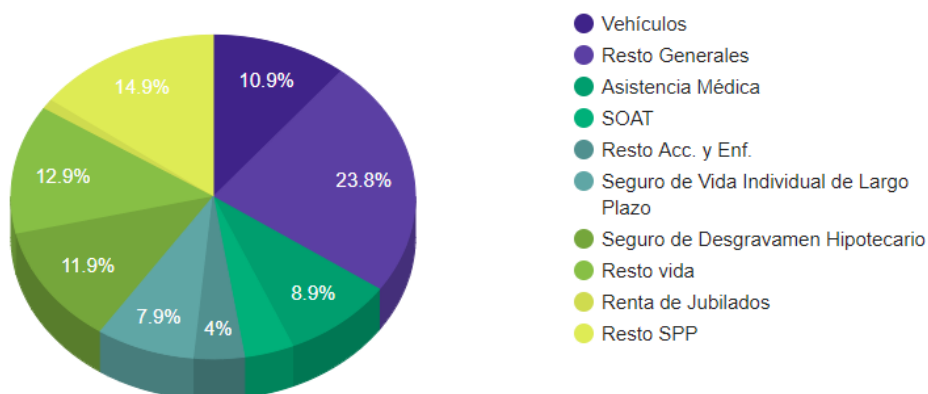
El ranking de primas de seguros netas muestra que el porcentaje de participación por empresa se encuentra concentrado entre las empresas Rimac, Pacífico Seguros y Mapfre Perú con 30.69%, 25.88% y 7.54% respectivamente<sup>4</sup>. De la misma forma sucede para los ramos generales, accidentes y enfermedades con 34.2%, 24.80% y 15.4% respectivamente<sup>5</sup>. Finalmente, para el ramo de vida la participación se concentra entre las empresas Rimac, Pacifico Seguro e Interseguros con 27.3%, 26.9% y 10.95% respectivamente. (Superintendencia de Banca y Seguros, 2020)

### 2.3.2. El mercado del Seguro Vehicular Peruano

El Mercado peruano de seguros vehiculares está conformado por 5 empresas principales Rímac, Pacifico Seguros, Mapfre, La Positiva y HDI Seguros con una pequeña participación. La participación de mercado de 4 empresas principales se ha mantenido casi constante en los últimos años, liderando el mercado Rímac, con 38% de participación con 521 millones de soles en primas emitidas.

#### Gráfico 1:

*Composición del Mercado según Primas de Seguros Netas de Principales Riesgos – Año 2018*

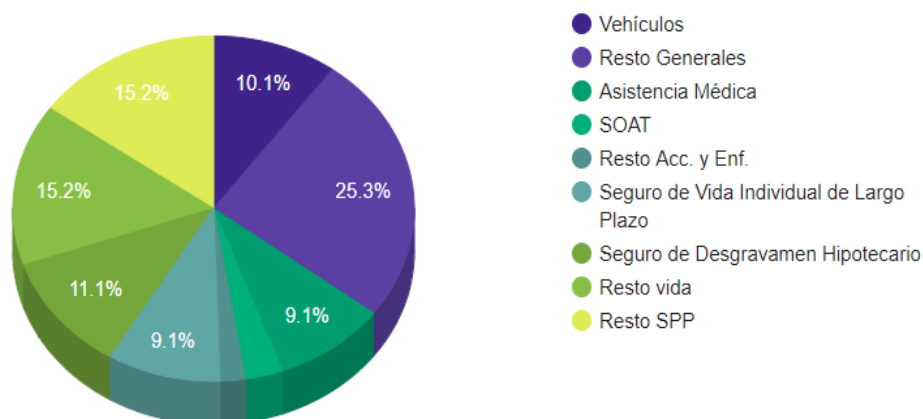


Fuente: Asociación Peruana de Empresas de Seguros - APESEG

<sup>3</sup> Anexo 4: Evolución de las Primas de Seguros Netas por Ramo. Tomado de Evolución del Sistema de Seguros al IV trimestre de 2019 por Superintendencia de Banca y Seguros (SBS)

<sup>4</sup> Anexo 5: Ranking del Total de Primas de Seguros al 31 de enero del 2020. Participación en el Mercado enero 2020 por Superintendencia de Banca y Seguros (SBS)

<sup>5</sup> Anexo 6: Ranking del Total de Primas de Seguros ramos generales, accidentes y enfermedades al 31 de enero del 2020. Participación en el Mercado enero 2020 por Superintendencia de Banca y Seguros (SBS)

**Gráfico 2:****Composición del Mercado según Primas de Seguros Netas de Principales Riesgos – Año 2019**

Fuente: Asociación Peruana de Empresas de Seguros – APESEG

Tal y como se observa en los gráficos Número 1 y Número 2, los seguros vehiculares conforman gran parte del mercado de primas con 10.9% y 10.1% en los años 2018 y 2019 respectivamente.

## 2.4. HIPÓTESIS

### 2.4.1. Argumentación

Los siniestros vehiculares constituyen una problemática en el Perú, de acuerdo a las estadísticas del Instituto Nacional de Estadística e Informática INEI, 9 de cada 100 mil habitantes, mueren en accidentes de tránsito. A pesar que las proyecciones de siniestros vehiculares sean altas, sólo el 25% del total del parque automotor sin contar motos se encuentra asegurado; en cifras representa sólo 640000 vehículos del total de 2600000. (Iparraguirre, 2018)

De acuerdo a Guillermo Tham, Gerente Técnico de Qualitas Perú Compañía de Seguros, una sociedad poco bancarizada, con poca cultura de la prevención, con inclinación a la informalidad y fraude ocasiona que las primas no resulten económicamente atractivas para los potenciales asegurados. En consecuencia, existe poca penetración de seguros vehiculares y gran cantidad de siniestros no cubiertos.

En tal sentido, para una aseguradora la decisión de otorgar o no un seguro a un individuo, así como la decisión de adquirir o no un seguro por parte del cliente, responde a un problema de diferencia de intereses entre el principal y el agente como consecuencia de la asimetría de la información. (Vásquez, 2015)

Es preciso señalar que, según la economía del comportamiento, el modelo de comportamiento humano basado en la premisa de que las personas optimizan es y siempre ha sido inverosímil. (Thaler, 2018) Además, gran parte de la economía del comportamiento o economía de la conducta surge del hecho de que las personas tienen un lado emocional e irracional que los modelos económicos tradicionales no logran incorporar. En tal sentido, en las ciencias modernas el hombre moderno constituye un tomador de decisiones asediado por la incertidumbre en todos los frentes (Levine, 2010)

En consecuencia, parte de la literatura revisada concluye que existen componentes de selección adversa y riesgo moral evidenciados en un cambio de conducta que se da cuando el asegurado ya obtiene las garantías de protección del auto y muestra descuido, siempre y cuando la cobertura elegida por la aseguradora cubra el riesgo de siniestro. (Mateus, 2008)

Asimismo, en la literatura revisada respecto a los determinantes de la probabilidad de siniestro en los seguros de autos, las investigaciones concluyen que los factores asociados al conductor, al vehículo, a la póliza tienen una correlación positiva con el aumento de la probabilidad de siniestro de un vehículo. (Quishpe, 2015)

Finalmente, con la metodología aplicada, se mostrará una de las alternativas econométricas - estadísticas que puede seguirse para evaluar las relaciones entre la probabilidad de siniestro con diversos factores como son los determinantes asociados a la conducta aleatoria del individuo, los asociados al conductor, los asociados al vehículo y la póliza. Finalmente, el emplear datos de una aseguradora de Lima puede brindar información relevante para otros investigadores.

#### **2.4.2. Enunciado de la Hipótesis**

##### **a) Hipótesis General**

Existe una relación positiva entre los determinantes no observables asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana.

##### **b) Hipótesis Específicas**

- Existe una relación entre los determinantes observables asociados al conductor con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018
- Existe una relación entre los determinantes observables asociados al vehículo con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018
- Existe una relación entre los determinantes observables asociados a la póliza con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018

## **CAPÍTULO III: METODOLOGIA**

### **3.1. DISEÑO DE LA INVESTIGACIÓN**

La investigación es no experimental de tipo correlacional. No experimental debido a que las variables independientes ya acontecieron y no han sido modificadas premeditadamente. Por último, es de tipo correlacional ya que describe la relación entre los determinantes y la probabilidad de ocurrencia de siniestro.

### **3.2. POBLACIÓN Y MUESTRA**

La población de la presente investigación está conformada por el mercado asegurador de autos y los clientes potenciales a asegurarse.

La muestra es el total de pólizas, de una compañía aseguradora de lima, que se encuentran vigentes para los años 2016, 2017 y 2018. La cantidad de datos evaluados en la presente investigación es de 315,194.

### **3.3. MÉTODO DE RECOLECCIÓN DE DATOS**

Los datos correspondientes a las pólizas del seguro de autos en la empresa aseguradora elegida se obtienen mediante de dos formas:

- Datos proporcionados por el cliente al momento de la suscripción del seguro.
- Módulo de Siniestros del Sistema de seguros donde se registra y actualiza los datos relacionados a los siniestros ocurridos a los clientes como fecha de ocurrencia, causa, tipo de siniestros, cobertura, monto estimado de atención.

Los datos son consolidados en hojas Excel o en un archivo de texto mes a mes y luego se cargan a una base de datos en un repositorio de tablas llamado Datamart de Autos mediante la herramienta Internet Information Services (IIS). Este repositorio es una base de datos Oracle diseñado bajo el modelo estrella (Fact Table) que permite relacionar los atributos más importantes de la póliza y el siniestro con tablas asociadas denominadas dimensiones que aportan mayor información para el

análisis, algunas dimensiones son: tiempo, marca, modelo de auto, ubigeo, categorías de clientes, brokers entre otros. (Ver imagen adjunta)<sup>6</sup>

Para el caso específico del Datamart de Autos, esta herramienta permite incluir información actualizada y asociarla según la póliza contratada. Asimismo, ayuda en el manejo de información dado que se puede llevar un mejor control.

Para la presente investigación la compañía aseguradora compartió datos sobre las pólizas mas no datos que puedan vulnerar la privacidad del cliente, es decir, los datos proporcionados por la empresa aseguradora no permiten identificar al cliente por lo tanto no pone en riesgo la vulneración de las cláusulas de confidencial que manejan y que están estipuladas por la SBS y por la “Ley de Protección de Datos Personales” reglamentados desde el año 2011.

### **3.4. MÉTODO DE ANÁLISIS DE DATOS**

#### **3.4.1. El Modelo Logit**

El modelo econométrico elegido es el Logit el cual también ha sido usado en diversos estudios previos concernientes a los factores de la probabilidad de siniestro. La regresión logística forma parte del conjunto de métodos estadísticos que caen bajo tal denominación y corresponde al caso en que se valora la contribución de diferentes factores en la ocurrencia de un evento simple.

En general, la regresión logística es adecuada cuando la variable de respuesta Y es politómica (admite varias categorías de respuesta, tales como: mejora, mejora mucho, empeora, se mantiene), pero es especialmente útil en particular cuando solo hay dos posibles respuestas (cuando la variable de respuesta es dicotómica), que es el caso más común. (Hosmer & Lemeshow, 2013)

El modelo de regresión logística tiene principalmente dos objetivos:

- 1) Investigar cómo influye en la probabilidad de ocurrencia de un suceso, la presencia o no de diversos factores y el valor o nivel de los mismos.
- 2) Determinar el modelo más parsimonioso y mejor ajustado que describa el comportamiento entre la variable respuesta y las variables regresoras (Salcedo, 2002)

Sea Y una variable dependiente binaria (con dos posibles valores: 0 y 1). Sean un conjunto de k variables independientes,  $(x_1, x_2, \dots, x_k)$ , observadas con el fin de predecir/explicar el valor de Y.

---

<sup>6</sup> Anexo 6: Diagrama de Repositorio de Autos



El objetivo consiste en determinar:

$$P[Y = 1/x_1, x_2, \dots, x_k] \rightarrow P[Y = 0/x_1, x_2, \dots, x_k] = 1 - P[Y = 1/x_1, x_2, \dots, x_k]$$

Para ello, se construye el modelo  $P[Y = 1/x_1, x_2, \dots, x_k] = p(x_1, x_2, \dots, x_k; \beta)$  donde:  $p(x_1, x_2, \dots, x_k; \beta): R^k$  función de enlace  $\rightarrow [0,1]$  que depende de un vector de parámetros  $\beta = (\beta_1, \beta_2, \dots, \beta_k)$

#### - Función de Verosimilitud

Con el fin de estimar  $\beta = (\beta_1, \beta_2, \dots, \beta_k)$  y analizar el comportamiento del modelo estimado se toma una muestra aleatoria de tamaño  $n$  dada por  $(x_i, y_i)_{i=1,2,\dots,n}$  donde el valor de las variables independientes es  $x_i = (x_{i1}, x_{i2}, \dots, x_{ik})$  e  $y_i \in [0,1]$  es el valor observado de  $Y$  en el  $i$ -ésimo elemento de la muestra.

Como  $(Y/X_1, X_2, \dots, X_k) \in B[1, p(X_1, X_2, \dots, X_k; \beta)]$ , la función de verosimilitud viene dada por:

$$L[\beta/(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)] = \prod_{i=1}^n p_i^{y_i} (1-p_i)^{1-y_i} \text{ donde } p_i = p(x_i; \beta) = p[(x_{i1}, x_{i2}, \dots, x_{ik}); \beta]_{i=1,2,\dots,n}$$

#### - Modelo Lineal

$$\begin{cases} 0 & \text{si } \beta_1 X_1 + \dots + \beta_k X_k < c_0 \\ \beta_1 X_1 + \dots + \beta_k X_k & \text{si } c_0 < \beta_1 X_1 + \dots + \beta_k X_k \leq c_1 \\ 1 & \text{si } \beta_1 X_1 + \dots + \beta_k X_k > c_1 \end{cases} \quad c_0, c_1 \text{ son constantes}$$

#### - Modelo Logit (modelo de regresión logística binaria)

$p(x_1, x_2, \dots, x_k; \beta) = G[\beta_1 x_1 + \dots + \beta_k x_k]$  donde  $G(x) = \frac{e^x}{1+e^x}$  es la función de distribución logística.

#### a) Modelo de Regresión Logística Binaria

El modelo logístico establece la siguiente relación entre la probabilidad de que ocurra el suceso, dado que el individuo presenta los valores  $(X = x_1, X = x_2, \dots, X = x_k)$

$$P[Y = 1/x_1, x_2, \dots, x_k] = \frac{1}{1 + e^{(-\beta_0 - \beta_1 x_1 - \beta_2 x_2 - \dots - \beta_k x_k)}}$$

El objetivo es hallar los coeficientes  $(\beta_0, \beta_1, \dots, \beta_k)$  que mejor se ajusten a la expresión funcional.

Se conoce como odds (ratio del riesgo) al cociente de probabilidades:

$$\text{Odds (ratio de riesgo)} = \frac{P[Y=1/X_1, X_2, \dots, X_k]}{1-P[Y=1/X_1, X_2, \dots, X_k]} = \frac{p(X_1, X_2, \dots, X_k; \beta)}{1-p(X_1, X_2, \dots, X_k; \beta)} = e^{\beta_1 + \beta_2 X_2 + \dots + \beta_k X_k}$$

se toma como primera variable explicativa a la variable constante que vale 1.

Tomando logaritmos neperianos en la expresión anterior, se obtiene una expresión lineal para el modelo:

$$\text{Logit}[P(Y=1)] = \text{Ln} \left[ \frac{P[Y=1/X_1, X_2, \dots, X_k]}{1-P[Y=1/X_1, X_2, \dots, X_k]} \right] = \beta_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Aquí se aprecia que el estimador del parámetro  $\beta_2$  se podrá interpretar como la variación en el término Logit (logaritmo neperiano del cociente de probabilidades) originada por una variación unitaria en la variable  $x_2$  (suponiendo 2 constantes el resto de variables explicativas).

Cuando se hace referencia al incremento unitario en una de las variables explicativas del modelo, aparece el concepto de odds-ratio como el cociente entre los dos odds asociados (el obtenido al realizar el incremento y el anterior al mismo). Suponiendo que ha habido un incremento unitario en la variable X

$$\text{Odds ratio} = \frac{\text{Odds2}}{\text{Odds1}} = e^{\beta_i} \quad \text{OR} = e^{\beta_i}$$

De donde se desprende que, un coeficiente  $\beta_i$  cercano a cero, es decir, un odds-ratio próximo a 1, indicará que cambios en la variable explicativa  $x_i$  asociada no tendrán efecto alguno sobre la variable dependiente Y.

#### - **Bondad de Ajuste del Modelo**

Se utilizan dos tipos de contrastes: (a) Contrastes que analizan la bondad de ajuste desde un punto de vista global. (b) Contrastes que analizan la bondad de ajuste paso a paso.

#### - **Contraste de bondad de ajuste global de Hosmer-Lemeshow**

- **El índice de bondad de ajuste**

$z^2 = \sum_{i=1}^n \frac{(y_i - \hat{p}_i)^2}{\hat{p}_i(1-\hat{p}_i)}$  donde  $\hat{p}_i = p(x_{i1}, x_{i2}, \dots, x_{ik}; \hat{\beta})_{i=1,2,\dots,n}$ ,  $z^2 \approx \chi_{n-k}^2$  si el modelo ajustado es cierto

- **El estadístico desviación viene dado por la expresión**

$$D = 2 \sum_{i=1}^n y_i \ln \left[ \frac{y_i}{\hat{p}_i} \right] + 2 \sum_{i=1}^{n-m} (1 - y_i) \ln \left[ \frac{(1-y_i)}{1-\hat{p}_i} \right]$$
 Donde  $m$ =número de observaciones con  $y_i = 1$  y  $D \approx \chi_{n-k}^2$  si el modelo ajustado es cierto.

### 3.5. TÉCNICAS DE RECOLECCIÓN DE DATOS:

Para esta investigación se ha tomado como fuente los datos de una compañía aseguradora peruana. La recolección de los datos se realiza considerando los siguientes criterios:

- Se consideran las pólizas vigentes en los años 2016, 2017 y 2018.
- Se excluyen las pólizas de personas jurídicas es decir sólo se consideran las pólizas otorgadas a personas naturales contratadas a título personal.
- Dado que población objetivo del presente estudio se centra en Lima Metropolitana, sólo se toma en cuenta la muestra de dicho lugar.

### 3.6. TÉCNICAS PARA EL PROCESAMIENTO Y ANÁLISIS DE LA INFORMACIÓN

De acuerdo a (Maddala, 1985) la interpretación de la probabilidad como una frecuencia relativa de largo alcance es una de las posibles interpretaciones. Existe también la probabilidad en grados de creencia razonable. Es así que existen eventos en la realidad que serán deducidos a partir de experiencias anteriores o de situaciones similares. En tal sentido, también existirá una frecuencia en la que se basen estas probabilidades, pero todo dependerá de la situación que se evalúe y como se considere la misma (si como un caso especial o dentro de la misma clase que los eventos anteriores). Sea el análisis de la probabilidad que un individuo tenga un accidente de auto durante el próximo año y se sabe la frecuencia de accidentes en personas dentro de un rango de edad es de  $\frac{1}{10}$ . Entonces si no se tiene mayor información del individuo, la probabilidad subjetiva será  $\frac{1}{10}$ . Pero si se sabe algo más del individuo como conductor como por ejemplo la cantidad de accidentes que ha tenido en el último año o si se sabe el modelo y color de auto que maneja y se tiene información que las personas que tienen las mismas características del individuo tienen mayor probabilidad de accidentarse entonces la probabilidad de accidentarse subjetiva del individuo aumentará sustancialmente. La cuestión reside en asignar ponderaciones a todos estos factores, lo cual implica una nueva elección subjetiva.

#### 3.6.1. La probabilidad Condicional

La probabilidad condicional describe la ocurrencia simultánea de dos sucesos  $E_1$  y  $E_2$ . Sea  $P(E_1 \cap E_2)$  o bien  $P(E_1 E_2)$  a la probabilidad que ocurran conjuntamente los dos sucesos  $E_1$  y  $E_2$  del espacio muestral. Entonces la probabilidad condicionada de  $E_1$  dado  $E_2$  es  $P(E_1|E_2)$ , y es igual a,

$$\frac{P(E_1 \cap E_2)}{P(E_2)} \quad \text{o} \quad \frac{P(E_1 E_2)}{P(E_2)}$$

siempre que  $P(E_2)$  sea distinta de cero. (Maddala, 1985)

Dos sucesos  $E_1$  y  $E_2$ , se dice que son independientes si las probabilidades condicionada y no condicionada son iguales, es decir,

$$P(E_1|E_2) = P(E_1) \quad \text{o bien} \quad P(E_2|E_1) = P(E_2)$$

Puesto que

$$P(E_1|E_2) = \frac{P(E_1 E_2)}{P(E_2)} = P(E_1)$$

Se deduce  $P(E_1 E_2) = P(E_1) \cdot P(E_2)$

### 3.6.2. El teorema de Bayes

El teorema de Bayes está basado en la definición de probabilidad condicionada dada anteriormente. Sean A y B dos sucesos. Entonces por la definición de probabilidad condicionada, tenemos.

$$P(A|B) = \frac{P(AB)}{P(B)} \quad \text{y} \quad P(B|A) = \frac{P(AB)}{P(A)}$$

Así,

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Si se sustituye H (hipótesis en torno al modelo generado de los datos) por A y D (datos observados) por B, entonces se tiene:

$$P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)}$$

Aquí  $P(D|H)$  es la probabilidad de observar los datos. Supuesto que H sea cierto. Esto es habitualmente conocido con el nombre de verosimilitud (likelihood).  $P(H)$  es la probabilidad de que

ocurra H antes de la observación de los datos (usualmente denominada probabilidad a priori).  $P(H|D)$  es la probabilidad de que ocurra H después de observar los datos (normalmente llamada probabilidad a posteriori).  $P(D)$  es la probabilidad no condicionada de observación de los datos (habiendo o no ocurrido H). A menudo  $P(D)$  es difícil de calcular. De ahí que se escribe la relación anterior como sigue:

$$P(H|D) \propto P(D|H) \cdot P(H)$$

Esto es la probabilidad a posteriori es igual a la probabilidad a priori multiplicada por la verosimilitud.

A través de la combinación de ambas, el analista conforma un juicio de probabilidad que sintetiza su nuevo grado de convicción al respecto. Esta probabilidad a priori, una vez incorporada la evidencia que aportan los datos, se transforma así en una probabilidad a posteriori. (Maddala, 1985)

### 3.6.3. Inferencia Bayesiana

La inferencia bayesiana tiene como referencia el Teorema de Bayes, de acuerdo a este método se afirma que la probabilidad a posteriori varía por el producto de la probabilidad a priori y la función de verosimilitud. En tal sentido, se puede o no tener nociones a priori sobre el parámetro  $\theta$ .

#### a) Caso Inferencia Bayesiana con a priori Informativa

Sea el primer caso, en el que se tiene alguna noción sobre el parámetro entonces se utiliza una información adicional conocida como la distribución a priori que incluyen algunas nociones previas sobre el parámetro  $\theta$  que se combina con la función de verosimilitud que representa la información muestral para obtener la distribución a posteriori.

$$\text{Probabilidad a posteriori} = \frac{\text{f. de veros.} \times \text{probabil. a priori}}{\sum(\text{f. de veros.} \times \text{probabil. a priori})}$$

Esta división se llama normalización y para distribuciones continuas se denota:

$$f(\theta|y) = \frac{L(y|\theta) \cdot h(\theta)}{\int L(y|\theta) \cdot h(\theta) d\theta}$$

$f(\theta|y)$  es la distribución a posteriori de  $\theta$  dadas unas observaciones muestrales  $y$ ,  $L(y|\theta)$  es la función de verosimilitud,  $h(\theta)$  es la función de densidad de probabilidad a priori y  $\int L(y|\theta) \cdot h(\theta) d\theta$  es el factor de normalización. La forma  $L(y|\theta)$  dependerá de la función de

probabilidad asumida. A menudo  $h(\theta)$  se elige de tal forma que la integral  $\int L(y|\theta) \cdot h(\theta) d\theta$  pueda ser fácilmente representada y la forma funcional de la distribución a posteriori  $f(\theta|y)$  es la misma que la de  $h(\theta)$ . Esta distribución  $h(\theta)$  se llama conjugada a priori. La ventaja de la conjugada a priori es que siempre que se observe una nueva muestra, la revisión de  $\theta$  puede llevarse a cabo mediante el mismo procedimiento analítico. Esto es,  $f(\theta|y)$  siendo la misma forma que  $h(\theta)$ , se puede usar como una nueva función a priori para analizar la nueva información muestral, y las fórmulas obtenidas anteriormente pueden ser usadas para obtener la nueva distribución a posteriori.

### b) Caso Inferencia Bayesiana a priori no Informativa

Sea el segundo caso en el que se carece de una noción a priori sobre  $\theta$ , se sugiere el uso de una difusa o no informativa a priori, cabe resaltar que estas funciones son no informativas en sentido relativo. La función a priori asumida es uniforme sobre un rango no especificado. Estas a priori también se denominan impropias porque las probabilidades no suman uno. Sea una distribución normal con media  $\mu$  y varianza  $\sigma^2$  las funciones a priori son:

$$h(\mu) \propto \text{constante} \quad -\infty < \mu < \infty \text{ (valores estables de los estadísticos)}$$

$$h(\sigma) \propto \frac{1}{\sigma} \quad 0 < \sigma < \infty$$

Si la muestra es muy grande, o si las distribuciones a priori son difusas, la distribución a posteriori será proporcional a la función de máximo verosimilitud.

En la inferencia bayesiana, así como en la inferencia clásica se considera la estimación puntual, la estimación por intervalos y los tests de significatividad. Todos ellos basados en la distribución a posteriori. Asimismo, en la mayoría de aplicaciones económicas se da el caso que hay mucha información subjetiva a priori en los problemas de decisión de empresas individuales, para estas situaciones el método bayesiano es el apropiado para el análisis. (Maddala, 1985).

Por lo general, al contar con una distribución marginal posterior no resulta analíticamente tratable o presenta un difícil problema de solución analítica que involucra tratar integrales de elevada dimensión (Ntzoufras, 2011). Por eso es necesario implementar el método de Monte Carlo vía Cadenas de Markov (MCMC) que fueron enfoques basados en muestreo para calcular densidades marginales por Gelfand y Smith (1990), para una simulación de distribución posterior de los parámetros. El trabajo de

inferencia se da condicionando los datos de observación y cuantificando la incertidumbre existente con probabilidades.

#### 3.6.4. El Algoritmo MCMC: Muestro de Gibbs y Metropolis-Hanstings

En la metodología del algoritmo MCMC fue llevada a cabo en sus inicios por el físico Nicholas Metropolis (Metropolis, 1953) en una versión simplificada que después fue estandarizada por Hastings (1970) y por último es utilizado por investigadores bayesianos Gelfand y Smith (1990). La metodología del MCMC se popularizo con el muestreo de Gibbs implementado por ser un caso especial de algoritmo Metropolis-Hanstings. Los algoritmos son una clase de procedimientos y se basan en la integración de Monte Carlo, empleando cadenas de Markov para crear muestras de distribución que es necesario para el interés de conocer la función de densidad posterior exacta. Las cadenas de Markov deben de converger a una función de distribución posterior estacionaria o equilibrada, por medio de método iterativo que en cada paso nos origina un valor condicionado al valor de un paso antes. Este procedimiento de implementación de algoritmo gestionada para la generación de muestras y estimar parámetros  $\theta$  que se da a partir de un valor inicial  $\theta^{(0)}$  (Gelman y Rubin, 1992) para el caso de generar diferentes muestras (cadenas) operando con diferentes valores iniciales y originar un gran número de iteraciones. Los resultados obtenidos en las primeras iteraciones se recomiendan que se quiten para no caer en influencia de los valores iniciales fijados, reconocido como fase inicial de calentamiento (burning phase). Ya que los valores de la muestra final generados no son independientes, a la vez es necesario examinar la autocorrelación que se presentan entre estos valores. Entre las formas de examinar la autocorrelación es elegir valores a cada k iteraciones eliminando las otras.

##### a) El Algoritmo de Muestreo de Gibbs

La técnica de muestreo de Gibbs para conseguir y generar cadenas de Markov ha sido largamente utilizada en procedimientos de inferencia sobre los parámetros en modelos bayesianos. El procedimiento del método se da por el uso de distribuciones condicionales completas de los parámetros como la distribución de transición en cadena de Markov. Este muestreo de Gibbs es más usado cuando una distribución condicional posterior tiene la forma de una densidad de probabilidad con un patrón ya conocido.

Se plantea que se tiene  $\mathcal{Y}$  datos observados con un vector de tres parámetros ajenos a estimar  $[\theta_1, \theta_2, \theta_3]$  y distribuciones condicionadas

$$f(\theta_1 | y, \theta_2, \theta_3), f(\theta_2 | y, \theta_1, \theta_3), f(\theta_3 | y, \theta_1, \theta_2),$$

Entre los valores rescatados de forma arbitraria para obtener cada uno de nuestros parámetros  $[\theta_1^{(0)}, \theta_2^{(0)}, \theta_3^{(0)}]$ . Uno de los primeros pasos del muestreo de Gibbs es generar una muestra nueva  $[\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}]$  que parte de las distribuciones condicionadas producto de los datos observados con sus valores iniciales. Este procedimiento busca la generación de observaciones aleatorias  $\theta_1^{(1)}$  partiendo de una distribución condicionada  $\theta_1$  y fijando todos los otros parámetros en valor inicial  $\theta_1^{(1)} \sim f(\theta_1 | y, \theta_2^{(0)}, \theta_3^{(0)})$ , para luego generar la observación aleatoria de  $\theta_2^{(1)}$  a partir de  $f(\theta_2 | y, \theta_1^{(1)}, \theta_3^{(0)})$  y para el otro parámetro  $\theta_3^{(1)}$  a partir de  $f(\theta_3 | y, \theta_1^{(1)}, \theta_2^{(1)})$ . Al obtener los parámetros ya actualizados se elabora una secuencia de Gibbs  $[\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}]$  con la finalidad de volver empezar con el procedimiento, cosa que en lugar de condicionar los parámetros a los valores iniciales, se condicionan los valores obtenidos en el último procedimiento de iteración, este procedimiento es sucesivo hasta obtener  $k$  número de muestras de los parámetros no conocidos, este procedimiento genera la cadena de Gibbs en la forma siguiente:

$$[\theta_1^{(0)}, \theta_2^{(0)}, \theta_3^{(0)}], [\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}], \dots, [\theta_1^{(k)}, \theta_2^{(k)}, \theta_3^{(k)}]$$

### b) El Algoritmo de Metropolis-Hastings

Este algoritmo de Metropolis-Hastings es una técnica de simulación estocástica que faculta al muestreo de cualquier función de densidad  $f(\theta)$ , sin la necesidad de saber las distribuciones condicionales (Chib y Greenberg, 1995). Por ejemplo, si una distribución condicional completa de análisis no presenta la forma de densidad, no se puede generar directamente los valores en esta distribución, lo que permite que el algoritmo Metropolis-Hastings se convierta en una opción, este método está basado en la generación de valores próximos de una densidad auxiliar que se presta para el muestreo, la densidad auxiliar debe presentar dos condiciones: la facilidad de generación de muestras aleatorias con versatilidad y rapidez; esta densidad auxiliar deberá encontrarse en el mismo dominio que la densidad objetivo, la utilización de valores candidatos que fueron simulados de la densidad auxiliar se aceptan o se rechazan, dependiendo de cierto grado de probabilidad. Chib y Greenberg (1995), encuentran varias familias de densidades auxiliares que se pueden utilizar para obtener valores candidatos.

Se supone que dada una distribución de probabilidad a posteriori  $f(\theta | y)$  para poder generar una muestra de tamaño  $T$  con el parámetro de interés desconocido  $\theta$ , en estos casos el algoritmo de Metropolis-Hastings implementado en base a la perspectiva bayesiana se podría resumir en los siguientes pasos:



- Se establece el valor inicial de  $\theta^{(0)}$ .
- Para  $t = 1, \dots, T$  se itera los siguientes pasos:
  - a. Se establece  $\theta = \theta^{(t-1)}$
  - b. Se genera un valor nuevo del parámetro candidato  $\theta'$  de una propuesta de distribución condicional  $q(\theta'|\theta)$ .
  - c. Se calcula

$$\alpha = \min \left( 1, \frac{f(\theta'|y)q(\theta|\theta')}{f(\theta|y)q(\theta'|\theta)} \right)$$

- d. Se Actualiza  $\theta^{(t)} = \theta'$  con probabilidad  $\alpha$ ; en caso contrario se establece  $\theta^{(t)} = \theta$

En el algoritmo se busca la convergencia a su distribución de equilibrio independientemente cual sea la distribución  $q(\cdot)$ . Una de las características de importancia del algoritmo es que no hace falta evaluar la normalización constante  $f(y)$  implicada en  $f(\theta|y)$  ya que se omite en  $\alpha$ . Donde la probabilidad de aceptación es simplificada en

$$\alpha = \min \left( 1, \frac{f(y|\theta')f(\theta')q(\theta|\theta')}{f(y|\theta)f(\theta)q(\theta'|\theta)} \right)$$

Este método está implementado en diferentes softwares, una vez terminado todos los pasos, se debe corroborar de que el procedimiento se ha implementado correctamente. Para cada una de las cadenas de MCMC se debe corroborar convergencia y si el número de valores obtenidos es suficiente para culminar de manera segura del procedimiento. Es decir, que las muestras obtenidas deriven de la verdadera distribución estacionaria.

### 3.6.5. Paquete Estadístico de Análisis Econométrico STATA:

Los datos recopilados se procesaron en el software estadístico STATA versión 15 debido a que este permite realizar diversos tipos de análisis estadísticos utilizando el enfoque bayesiano. En tal sentido, el paquete logit-bayesiano hizo posible el estudio de la relación entre las variables proporcionadas en la base de datos de la compañía aseguradora peruana. Asimismo, este programa se eligió debido a las siguientes ventajas:

- Buen ajuste: Existen 45 comandos aplicables a modelos de regresión bayesianos.
- Completo: Debido a la gama de herramientas de gestión de datos que posee.
- Diverso: Por la cantidad de técnicas estadísticas tradicionales y actuales que posee.
- Sencillo: El lenguaje de programación es amigable por la sintaxis intuitiva que tienen los comandos.

## CAPITULO IV: ANALISIS E INTERPRETACION DE LAS PRUEBAS ESTADÍSTICAS, FÓRMULAS Y RESULTADOS

### 4.1. ANÁLISIS DE LA RELACIÓN ENTRE LAS VARIABLES DE ESTUDIO

En este acápite del trabajo se realiza una descripción de la muestra que servirá para construir la función de máximo verosimilitud. A continuación, se describe la ocurrencia del siniestro vehicular durante el periodo 2016-2018 según los datos obtenidos por la empresa aseguradora peruana. En la tabla 1, se puede observar que existe registrado una cantidad de 266,181 clientes asegurados que no han sufrido ningún accidente vehicular, porcentualmente esto representa un 84.45% del total. Asimismo, la cantidad de clientes asegurados que han sufrido algún accidente vehicular registró 49,013, dicho número representa un 15.55% del total.

**Tabla 1:**

#### *Descripción de la ocurrencia del siniestro vehicular*

Siniestro del vehículo	Frecuencia	Porcentaje	Acumulado
No	266,181	84.45%	84.45%
Si	49,013	15.55%	100%
<b>Total</b>	315,194	100%	

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

En la tabla 2 se puede conocer la cantidad de siniestrados y no siniestrados según las categorías de las variables explicativas como: periodo de fabricación, marca del vehículo, obtención de GPS vehicular, tipo de vehículo, sexo y estado civil del cliente.

**Tabla 2:**

#### *Descripción de las variables explicativas según la ocurrencia de siniestro vehicular*

Variable	Categoría	Siniestro del vehículo		Total
		No	Si	
Periodo de fabricación	Otro	136,520	13,641	150,161
	2013-2018	129,652	35,381	165,033
	Otro	136,964	24,492	161,456

<b>Marca del vehículo</b>	NISSAN, KIA, HYUNDAY o TOYOTA	129,208	24,530	153,738
<b>GPS del vehículo</b>	No	239,998	37,426	277,424
	Si	26,174	11,596	37,770
<b>Sexo</b>	Femenino	56,786	17,122	73,908
	Masculino	209,386	31,900	241,286
<b>Tipo de vehículo</b>	Otro	138,333	28,489	166,822
	SEDAN, PICK UP o HATCHBACK	127,839	20,533	148,372
<b>Estado civil</b>	Otro	189,647	31,996	221,643
	Soltero	76,525	17,026	93,551

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

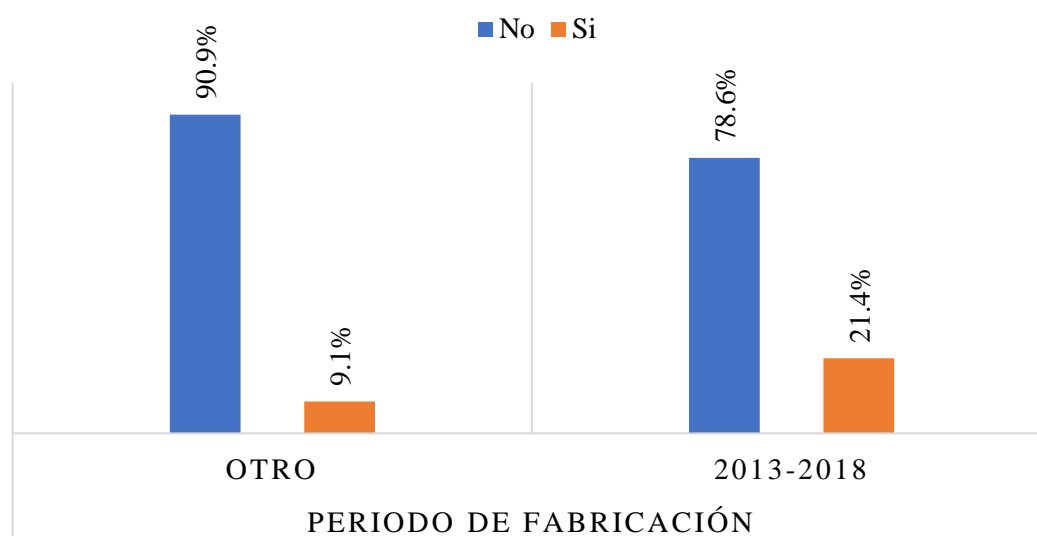
En tabla 3, se describe la temporalidad de fabricación de los vehículos asegurados con la ocurrencia de siniestro para el periodo 2013-2018. Se observa que para los periodos comparados el número de los vehículos que tuvieron siniestros se incrementó de 9.1% a un 21.4%. Por otro lado, para los que no presentaron siniestro, se observa que la proporción fue disminuyendo para el periodo 2013-2018 reflejando una cifra de 78.6% con respecto a otros periodos en el que fue 90.9% (Ver gráfico 3).

**Tabla 3:**

**Periodo de fabricación de los vehículos asegurados**

<b>Siniestro del vehículo</b>	<b>Periodo de fabricación</b>	
	<b>Otro</b>	<b>2013-2018</b>
<b>No</b>	90.9%	78.6%
<b>Si</b>	9.1%	21.4%
<b>Total</b>	100%	100%

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

**Gráfico 3:****Periodo de fabricación de los vehículos asegurados**

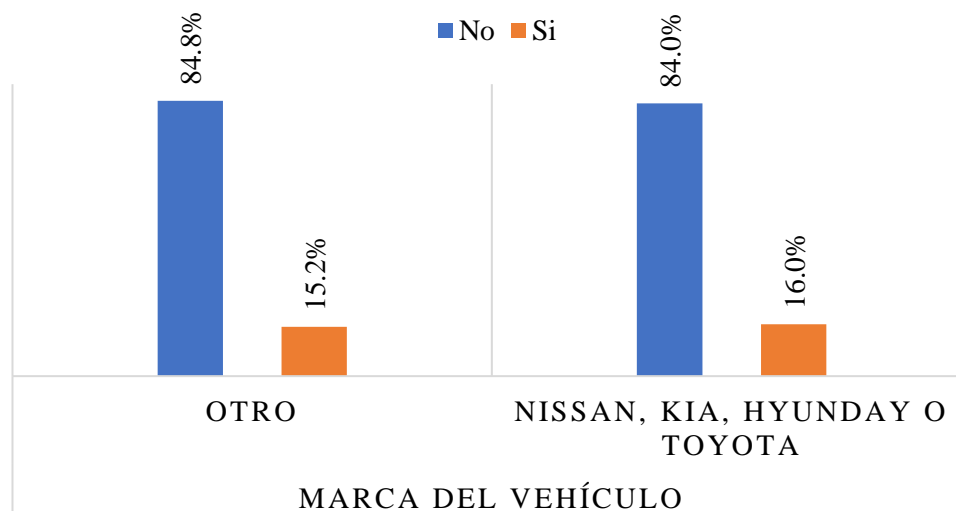
Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

A continuación, la Tabla 4 muestra la marca de los vehículos asegurados y la incidencia de siniestralidad en comparación con otras marcas. Para el periodo analizado las marcas Nissan, Kia, Hyundai o Toyota presenta una mayor incidencia de siniestralidad con respecto a otras marcas. Asimismo, el índice de siniestro de las marcas de análisis representó el 16.0% del total de vehículos asegurados y para los otros tipos de marcas de vehículos la siniestralidad resultó ser 15.2% del total de vehículos asegurados. (Ver gráfico 4).

**Tabla 4:****Marca de los vehículos asegurados**

Siniestro del vehículo	Marca del vehículo	
	Otro	NISSAN, KIA, HYUNDAY o TOYOTA
No	84.8%	84.0%
Si	15.2%	16.0%
Total	100%	100%

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

**Gráfico 4:****Marca de los vehículos asegurados**

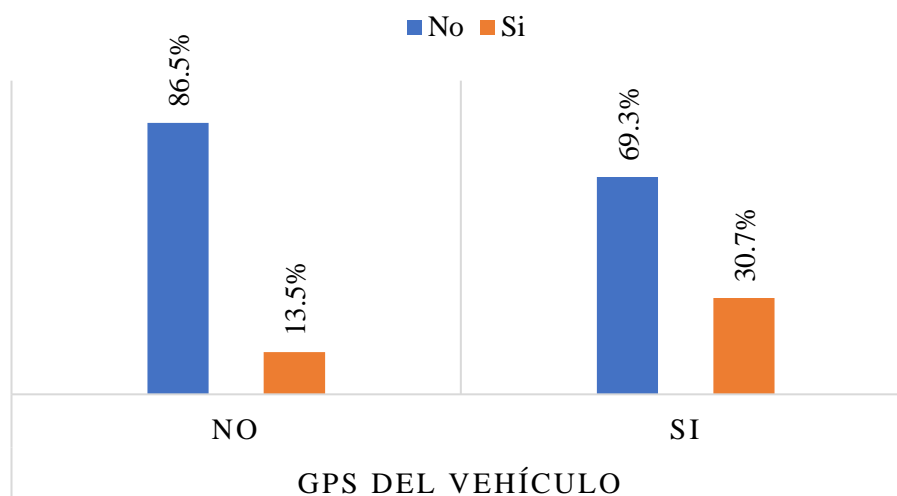
Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

La tabla 5, analiza la obtención del GPS de vehículos asegurados, el 30.7% del total de vehículos que cuentan con GPS registra siniestro vehicular y el 69.3% restante no presenta siniestro vehicular. Asimismo, entre los vehículos que no cuentan con el GPS, el 13.5% si presentan siniestro vehicular y el 86.5% no fueron afectados por algún siniestro vehicular. (Ver gráfico 5).

**Tabla 5:****Obtención de GPS de los vehículos asegurados**

Siniestro del vehículo	GPS del vehículo	
	No	Si
No	86.5%	69.3%
Si	13.5%	30.7%
<b>Total</b>	100%	100%

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

**Gráfico 5:****Obtención de GPS de los vehículos asegurados**

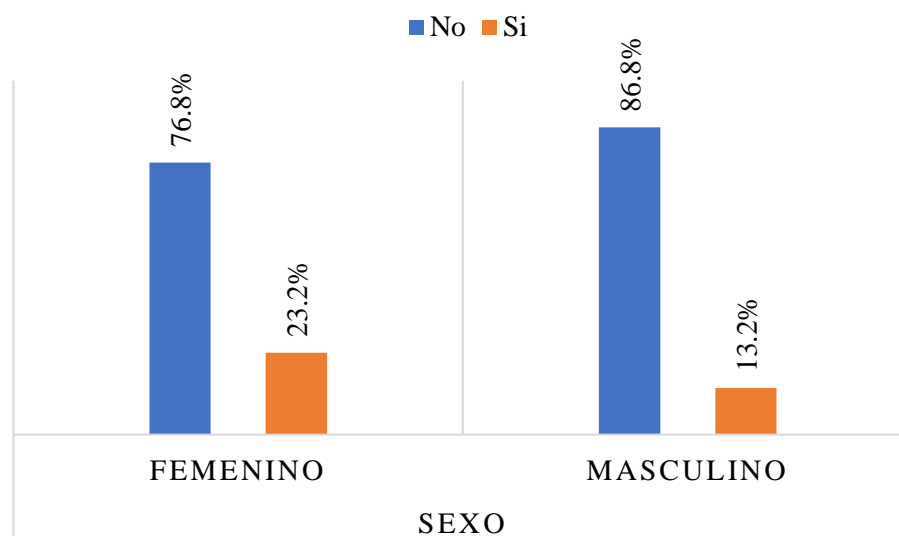
Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

En la Tabla 6 se observan los datos del sexo del propietario del vehículo y la ocurrencia de siniestro. Se tiene que, el 13.2% de propietarios de sexo masculino presentan siniestralidad y el 86.8 % de propietarios masculinos no sufrieron un siniestro vehicular. En el mismo sentido, el 23.2% de propietarias mujeres presentan siniestro vehicular por encima del sexo opuesto, y el 76.8% restante no presenta un siniestro vehicular. (Ver gráfico 6).

**Tabla 6:****Sexo del propietario del vehículo**

Siniestro del vehículo	Sexo	
	Femenino	Masculino
No	76.8%	86.8%
Si	23.2%	13.2%
<b>Total</b>	100%	100%

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

**Gráfico 6:****Sexo del propietario del vehículo asegurado**

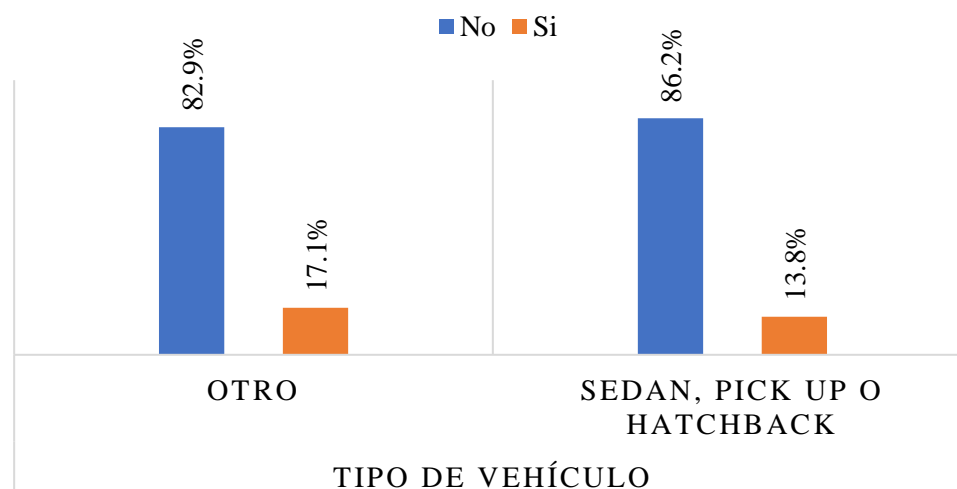
Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

Por otro lado, en la Tabla 7 se analiza el tipo de vehículo asegurado y su incidencia en la siniestralidad. Los datos muestran que los vehículos de tipo sedán, pick up o hatchback tuvieron menor cantidad de siniestros que otros tipos de vehículos del parque automotor, sólo el 13,8% de estos tuvieron siniestros y el 86,2% restante no presenta algún fatal suceso. Los otros tipos de vehículos que existen en el parque automotor tuvieron mayor incidencia en la siniestralidad, el 17,1% de estos tuvo un siniestro y el 82,9% no tuvo alguno. (Ver gráfico 7).

**Tabla 7:****Tipo de vehículo de la empresa aseguradora**

Siniestro del vehículo	Tipo de vehículo	
	Otro	SEDAN, PICK UP O HATCHBACK
No	82.9%	86.2%
Si	17.1%	13.8%
<b>Total</b>	100%	100%

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

**Gráfico 7:****Tipo de vehículo asegurado**

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

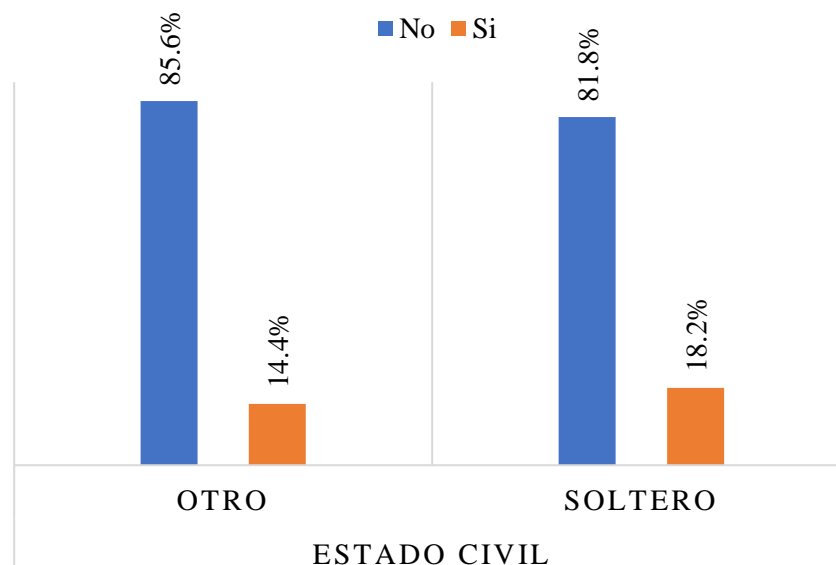
En la Tabla 8 se muestran los datos correspondientes al estado civil del propietario del vehículo asegurado, el 18.2% de propietarios en estado civil soltero registraron accidente vehicular y el 81.8% restante no presentaron algún accidente vehicular. En contraste, el 14.4% de propietarios con un estado civil diferente al soltero si presenta siniestro vehicular superior en 3 puntos porcentuales a los solteros y el 85.6% de propietarios con estado civil otros no tuvieron siniestro vehicular superior en 3 puntos porcentuales a los solteros tal como se aprecia (Ver gráfico 8).

**Tabla 8:****Estado civil del propietario del vehículo asegurado**

Siniestro del vehículo	Estado civil	
	Otro	Soltero
No	85.6%	81.8%
Si	14.4%	18.2%
Total	100%	100%

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana



**Gráfico 8: Estado civil del propietario del vehículo asegurado**

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

En la Tabla 9 se muestra una perspectiva de la edad. La edad promedio del propietario del vehículo que más incurrió en siniestro vehicular es de 43 años, a comparación de los que no presentaron siniestro, que tienen una edad promedio de 45 tal como se observa a continuación:

**Tabla 9:****Estadísticos de la edad del propietario del vehículo asegurado**

Siniestro del vehículo	Estadísticos de la edad del propietario del vehículo			
	Mediana	Media	Máximo	Desviación estándar
No	44	45.38	118	14.42
Si	42	43.89	118	13.70

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

Finalmente, en la Tabla 10 se muestra el análisis de la prima pagada por el seguro del vehículo se obtuvo, la prima promedio de 3822.915 pagada por el seguro del vehículo, si tiene siniestro vehicular y la menor prima promedio 471.8029 pagada por el seguro del vehículo, no tiene siniestro tal como se detalla a continuación:

**Tabla 10:****Estadísticos de la prima pagada del seguro del vehículo**

Siniestro del vehículo	Estadísticos de la prima pagada del seguro			
	Mediana	Media	Máximo	Desviación estándar
No	69.26	471.8029	64871.01	2167.03
Si	724.64	3822.915	74580	6131.304

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

## 4.2. ANÁLISIS DE DATOS CON INFERENCIA BAYESIANA

### 4.2.1. Distribución de Probabilidad A priori

La selección de la distribución a priori toma en consideración la presunción inicial del investigador sobre el fenómeno; para este estudio, la presunción se encuentra planteada en las hipótesis. Es importante identificar el tipo de distribución a priori que corresponde a la presunción del investigador para determinar si se trata de una distribución a priori informativa o no informativa dependiendo de la información disponible. Para esta investigación, se descarta el uso de una distribución a priori informativa dado que no se tiene datos históricos o fuentes de investigaciones pasadas del fenómeno analizado en la aseguradora estudiada. De acuerdo al marco teórico revisado se tiene una presunción o idea vaga respecto al fenómeno; motivo por el cual, se hizo la selección de una distribución a priori no informativa; cabe resaltar que se considera una a priori no informativa relativa por lo que se le asigna el término de vaga o difusa.

La distribución a priori vaga o difusa se basa en el razonamiento de Laplace para la densidad a priori uniforme en el que indica que si no se conoce nada sobre el parámetro entonces no hay razón para asignar probabilidades diferentes. En tal sentido, en equivalencia se considera que todos los valores posibles de los parámetros son igualmente probables. (E. Train, 2009) La teoría presume que una variable tiene una distribución difusa cuando es difícil de asignar a los eventos una probabilidad diferenciada y es la poca disponibilidad de información que se tiene en la presente investigación que hace que no se puede asignar una distribución a priori teórica.

Al realizar el estudio del marco teórico y analizar los factores determinantes de la probabilidad de siniestro, se discutió en algún momento sobre los determinantes no observables relacionados a la conducta aleatoria del individuo; a partir de estos se podía construir la distribución de probabilidad a

priori, pero lo cierto es que antes de este lapso analizado y en general en cualquier situación conocer este tipo de información en específico resulta difícil y hasta imposible por todas las razones detalladas en el marco teórico como puede ser: el problema de selección adversa que se da entre el principal y el agente en el mercado asegurador, la aleatoriedad en el comportamiento del individuo y hasta en su irracionalidad. En tal sentido, se opta por determinar esta información de forma un tanto subjetiva.

#### 4.2.2. Función de Máximo Verosimilitud

La muestra sirve de base para la construcción de la función de máximo verosimilitud porque corresponde a lo ocurrido en la realidad en un periodo donde los determinantes no observables asociados a la conducta aleatoria del individuo han sido los factores principales de cambio en la asignación de probabilidades. En tal sentido, se optó por la regresión logística (modelo logit) para el análisis de los datos debido al tipo de variables dicotómicas que se dispone. (Yupanqui, 2005)

La importancia de los modelos de elección discreta radica en que permiten la modelización de variables cualitativas, característica que exige la codificación de la variable como paso previo a la modelización. En este proceso, los distintos estados de la variable se transforman en códigos o valores susceptibles de ser tratados utilizando técnicas de regresión. (Martinez, 2007)

Estos modelos son planteados en situaciones donde los individuos se enfrentan a procesos de decisión dicotómicos, es decir, en los que únicamente hay dos posibles alternativas que representa de la forma  $y = 1$  o  $y = 0$ . (Mc Fadden, 1974)

Dado que los modelos probit también permiten la modelización de variables cualitativas se utiliza el estadístico de contraste de Mc Fadden. Según se detalla en la Tabla 11, se puede concluir que las variables utilizadas en el modelo econométrico se ajustan mejor en un modelo LOGIT, donde la función densidad sigue una distribución logística. Esto se concluye debido a los valores de los criterios de información de ambos modelos, donde el modelo LOGIT presenta menores valores de AIC, AIC\*n, BIC y BIC' respecto al modelo PROBIT; esta conclusión se refuerza con los coeficientes de determinación de cada modelo, donde el modelo LOGIT presenta mayores valores respecto al modelo PROBIT, esto nos indica que las variables predictoras explican en mayor porcentaje el comportamiento de la incidencia de los siniestros vehiculares

**Tabla 11:**

#### **Estadísticos para comparar modelos PROBIT y LOGIT**

	Current	Saved	Difference
Model:	probit	logit	
McFadden's R2:	0.133	0.134	-0.001

McFadden's Adj R2:	0.133	0.134	-0.001
Count R2:	0.860	0.861	-0.001
Adj Count R2:	0.100	0.105	-0.005
AIC:	0.749	0.749	0.001
AIC*n:	236205.152	235974.873	230.278
BIC:	-3.754e	-3.755e+06	230.278
BIC':	-36145.136	-36375.414	230.278

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

A continuación, se presenta el modelo econométrico en forma de ecuación

$$y_i = \beta X_i + u_i$$

Con  $E(u_i) = 0$ . La expectativa condicional está dada por:  $E(y_i|X_i) = \beta X_i$ . Asumiendo que  $E(U_i) = 0$  la probabilidad del término de error está dada por  $(1 - \beta X_i)$  y  $\beta X_i$ .

Para la presente investigación el modelo tiene las siguientes variables:

$$\sin_i = \beta_0 + \beta_1 fab_i + \beta_2 mak_i + \beta_3 gps_i + \beta_4 edc_i + \beta_5 sex_i + \beta_6 tveh_i + \beta_7 esc_i + \beta_8 pra_i + \varepsilon_i$$

Donde:

$\sin_i$ : Ocurrencia de siniestro del vehículo, toma el valor 1 si ocurre y 0, no ocurre.

$fab_i$ : Periodo de fabricación del vehículo, toma el valor de 1 dentro del periodo 2013-2018 y 0, en otros años.

$mak_i$ : Marca de los vehículos, toma el valor de 1 si el vehículo es NISSAN, KIA, HYUNDAI o TOYOTA y 0, si es otra marca.

$gps_i$ : Obtención de GPS del vehículo, toma el valor de 1 si cuenta y 0, si no cuenta.

$edc_i$ : Edad del propietario del vehículo.

$sex_i$ : Sexo del propietario del vehículo, toma valor de 1 si es masculino y 0, si es femenino.

$tveh$ : Tipo de vehículo, toma 1 si es SEDAN, PICK UP o HATCHBACK y 0, si es de otro tipo.

$esc_i$ : Estado civil del propietario del vehículo, toma valor de 1 si es soltero y 0, si es otro.

$pra_i$ : Prima pagada por el seguro del vehículo.

$\varepsilon_i$ : Error del modelo.

Cabe resaltar que para la presente investigación no es conveniente el uso de MCO (Mínimos cuadrados ordinarios) debido a que la variable ocurrencia de siniestro del vehículo, toma el valor 1 si ocurre y 0, si no. Pero en MCO podría tomar valores negativos o por encima de 1. Asimismo, debido a que la variable Y es binaria, el término del error en el modelo también es binario en tal sentido sigue una distribución binomial y no normal. Finalmente, en MCO se asume que la probabilidad de Y se incrementa con las variables explicativas en el modelo de regresión logística se asumen efectos marginales de las variables explicativas.

**Tabla 12:**

*Estimación del modelo LOGIT*

Logistic regression				Number of obs	=	315,194
				LR chi2(8)	=	36476.70
				Prob > chi2	=	0.0000
Log likelihood = -117978.44				Pseudo R2	=	0.1339
Sin	Coef.	Std. Err.	Z	P>z	[95% Conf. Interval]	
Fab	0.47517	0.01255	37.87	0.0000	0.45058	0.49976
Mak	0.03145	0.01155	2.72	0.0060	0.00882	0.05408
Gps	0.33428	0.01639	20.4	0.0000	0.30217	0.36640
Edc	0.00160	0.00039	4.16	0.0000	0.00085	0.00236
Sex	-0.65761	0.01147	-57.35	0.0000	-0.68008	-0.63513
Tveh	-0.14812	0.01141	-12.98	0.0000	-0.17049	-0.12574
Esc	0.14847	0.01145	12.97	0.0000	0.12603	0.17091
Pra	0.00020	0.00000	119.85	0.0000	0.00020	0.00020
_cons	-1.90363	0.02380	-80	0.0000	-1.95026	-1.85699

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

Del reporte obtenido detallado en la Tabla 12 se muestra que se analizaron 315,194 datos. Asimismo, es preciso señalar que la finalidad de esta regresión es encontrar la relación inicial de las variables de la muestra con el fenómeno de la probabilidad de ocurrencia de siniestro, sin tomar en cuenta la presunción inicial sobre los determinantes no observables asociados a la conducta aleatoria del individuo.

En tal sentido, según el modelo logit estimado se tiene que:

- La fabricación de los vehículos en el periodo 2013-2018 tiene un impacto diferencial positivo en la ocurrencia de siniestros del vehículo.
- La marca de los vehículos como NISSAN, KIA, HYUNDAI o TOYOTA tiene un impacto diferencial positivo en la ocurrencia de siniestros del vehículo.
- La obtención de GPS del vehículo tiene un impacto diferencial positivo en la ocurrencia de siniestros del vehículo.
- Los aumentos unitarios en la edad del propietario del vehículo tienen un impacto positivo en la ocurrencia de siniestros del vehículo.
- Si el cliente de la aseguradora es de sexo masculino, esto tendría un impacto diferencial negativo en la ocurrencia de siniestros del vehículo.
- Si el tipo de vehículo es SEDAN, PICK UP o HATCHBACK, esto tendría un impacto diferencial negativo en la ocurrencia de siniestros del vehículo.
- Si el cliente de la aseguradora es de estado civil soltero, esto tendría un impacto diferencial positivo en la ocurrencia de siniestros del vehículo.
- Los aumentos unitarios en la prima pagada del vehículo tienen un impacto diferencial positivo en la ocurrencia de siniestros del vehículo.

El análisis de la significancia individual de cada variable explicativa, se hizo mediante la observación de la probabilidad Z en el reporte, debido a que esta es menor al nivel de significación de 5% permite rechazar la hipótesis nula, donde el coeficiente de análisis es igual a cero. Por el lado de la significancia global, se cumple dicha significancia debido al estadístico de contraste Chi<sup>2</sup> que es menor al nivel de significación de 5%, por lo que se rechaza la hipótesis nula, donde el vector de parámetros es igual al vector de ceros.

El pseudo R cuadrado, que es el coeficiente de determinación del modelo logit que refleja la bondad de ajuste del modelo planteado, indica que las variables independientes explican en 13.39% el comportamiento de la incidencia de los siniestros vehiculares.

### 4.2.3. Distribución de Probabilidad Posterior

En la siguiente tabla se presenta la estimación del modelo LOGIT bajo el enfoque bayesiano, donde la distribución a priori de los parámetros es difusa o vaga debido a la ausencia de información de dicha distribución respecto a sus parámetros. Mediante la iteración del modelo de máxima verosimilitud, modelo LOGIT estimado en la tabla 12, y la densidad de los parámetros a priori se utiliza la técnica de simulación estocástica del algoritmo de Metropolis-Hasting que incorpora los métodos Monte Carlo vía Cadenas de Markov para muestrear la distribución posterior de los parámetros, esto se da mediante la generación de valores próximos de una densidad auxiliar que permite realizar este muestreo. En tal sentido el algoritmo Metropolis Hastings lo que hace es generar una secuencia de muestras, analizar su distribución y si esta concuerda con la distribución deseada.

La aplicación de este algoritmo genera la reducción de la cantidad de observaciones respecto a la muestra inicial debido a:

- La etapa quemada (Burn in) reduce la muestra debido a que existe la posibilidad que las primeras muestras tomadas aleatoriamente correspondan a áreas de baja densidad por lo que el algoritmo las elimina. Según el reporte se eliminaron 5000 muestras.
- Las iteraciones realizadas también reducen la muestra debido a que primero se toma un punto aleatorio para ser la primera muestra y escoge una función de densidad de probabilidad que sugiere un siguiente valor de la muestra que normalmente se encuentra cercano al punto inicial. Al repetir este proceso sucesivamente se seleccionan nuevas muestras. La admisión de la nueva muestra, depende del ratio de aceptación que contrasta su división con la muestra anterior y compara este resultado con el valor que se tiene antes de realizar la operación; por lo tanto, es posible que muchos valores se rechacen.
- El algoritmo aplicado elimina muestras que se encuentren correlacionadas a fin de obtener un conjunto de muestras independientes.

**Tabla 13:**

#### *Estimación del modelo LOGIT- Bayesiano*

Bayesian logistic regression	MCMC iterations	=	25,000
Random-walk Metropolis-Hastings sampling	Burn-in	=	5,000
	MCMC sample size	=	20,000
	Number of obs	=	10,661
	Acceptance rate	=	0. 2398

Efficiency: min = 0.001131  
 avg = 0.02064  
 max = 0.08469

Log marginal likelihood = -4432.4085

Sin	Mean	Std. Dev.	MCSE	Median	Equal-tailed	
					[95% Cred. Interval]	
Fab	-0.51309	0.05493	0.01155	-0.50854	-0.60176	-0.41698
Mak	0.06313	0.02922	0.00107	0.06332	0.00495	0.11932
Gps	0.00599	0.05767	0.00887	0.00867	-0.10377	0.11256
Edc	-0.01546	0.00152	0.00006	-0.01542	-0.01840	-0.01253
Sex	-0.31344	0.04714	0.00218	-0.31414	-0.40501	-0.22213
Tveh	0.26434	0.02494	0.00402	0.26565	0.21570	0.31060
Esc	0.10615	0.05452	0.00654	0.10703	0.00109	0.20994
Pra	0.00001	0.00000	0.00000	0.00001	0.00000	0.00002
_cons	2.78025	0.12837	0.02051	2.78185	2.54133	3.02708

Fuente: Elaboración propia en base a los datos de la Empresa Aseguradora Peruana

La distribución posterior es el último resultado respecto al análisis bayesiano, y mediante estos resultados se puede obtener la estimación de los parámetros del modelo logit bajo el enfoque bayesiano. En tal sentido, de la tabla anterior se puede concluir que la tasa de aceptación del modelo estimado es de 23.98%, lo cual significa que el algoritmo admite este porcentaje de la cantidad de parámetros propuestos, dicho valor es aceptable para tomar en cuenta que no hay problemas de convergencia, dado que es mayor a 10%, se puede concluir que la eficiencia de este modelo es buena. Asimismo, el paquete de análisis bayesiano del modelo logit del software estadístico STATA valida las variables utilizadas dado que el reporte obtenido no contiene alguna alerta de omisión de variables que indique que se deba eliminar del modelo. Por otro lado, se validan los resultados obtenidos mediante la observación de los intervalos de confianza para cada una de las variables, debido a que estos no pasan por el número 1 (uno) entonces, se concluye que se obtuvieron resultados significativos para el modelo planteado. Con lo antes mencionado, se admite la validez del modelo estimado.



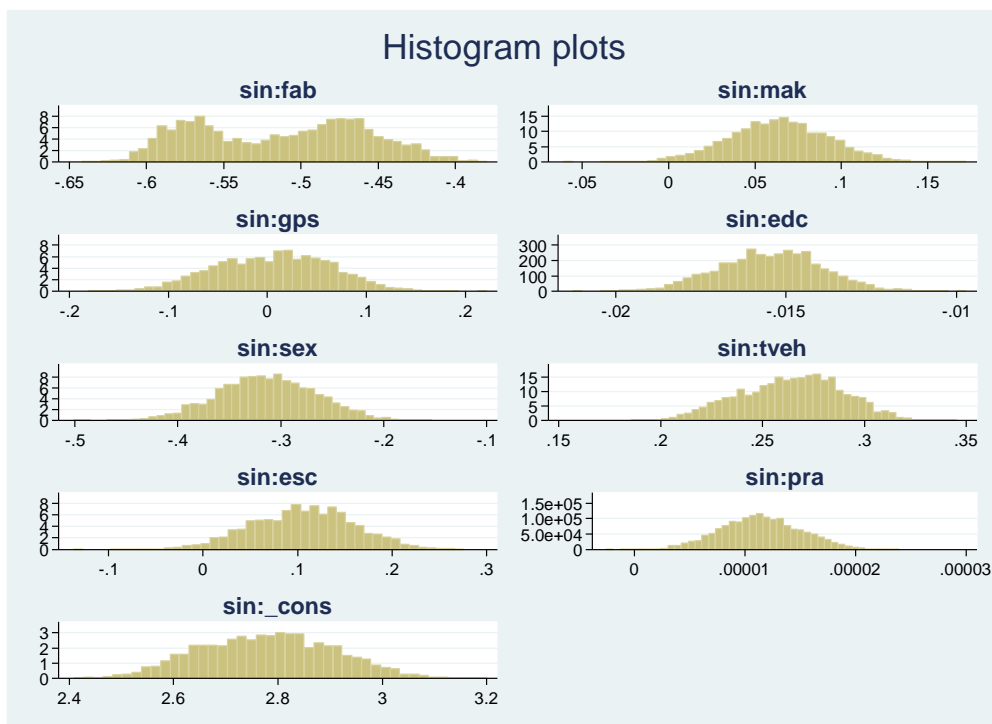
Por otro lado, el modelo logit estimado con el método bayesiano determina estimadores menores en valor absoluto que las estimaciones realizadas respecto al modelo logit, a excepción de los predictores fab, mak, edc, tveh y la constante cuyos valores son mayores al estimado por el método bayesiano. Es decir, el modelo incluyendo la distribución a priori vaga o difusa referida a los determinantes no observables de la conducta aleatoria del individuo obtiene un mejor ajuste.

Respecto a la diferencia de los estimadores de las explicativas del modelo de máxima verosimilitud respecto al modelo bayesiano, podemos mencionar lo siguiente:

- La fabricación de los vehículos en el periodo 2013-2018 tiene un impacto diferencial aproximado de 0.98825 en la ocurrencia de siniestros del vehículo.
- La marca de los vehículos como NISSAN, KIA, HYUNDAI o TOYOTA tiene un impacto diferencial aproximado de -0.03168 en la ocurrencia de siniestros del vehículo.
- La obtención de GPS del vehículo tiene un impacto diferencial aproximado de 0.32829 en la ocurrencia de siniestros del vehículo.
- Aumentos unitarios en la edad del propietario del vehículo tiene un impacto aproximado de 0.01706 en la ocurrencia de siniestros del vehículo.
- Si el cliente de la aseguradora es de sexo masculino, esto tendría un impacto diferencial aproximado de -0.34417 en la ocurrencia de siniestros del vehículo.
- Si el tipo de vehículo es SEDAN, PICK UP o HATCHBACK, esto tendría un impacto diferencial aproximado de -0.41245 en la ocurrencia de siniestros del vehículo.
- Si el cliente de la aseguradora es de estado civil soltero, esto tendría un impacto diferencial aproximado de 0.04232 en la ocurrencia de siniestros del vehículo.
- Aumentos unitarios en la prima pagada del vehículo tiene un impacto aproximado de 0.00019 en la ocurrencia de siniestros del vehículo.

En la siguiente figura se presenta la distribución posterior mediante histogramas según la muestra para los coeficientes del modelo estimado.

***Gráfico 9: Histograma de los parámetros del modelo LOGIT Bayesiano***



Según la figura anterior se puede inferir que los coeficientes estimados del modelo logit bayesiano siguen una distribución normal. Además, se puede concluir que para todas las medias muestrales posteriores y modas muestrales posteriores no hay mucha diferencia respecto al modelo logit en términos absolutos.

## **CAPÍTULO V: DISCUSIÓN, CONCLUSIONES Y RECOMENDACIONES**

### **5.1. DISCUSIÓN DE RESULTADOS Y CONCLUSIONES**

El objetivo fundamental de la tesis es evaluar los determinantes de la probabilidad de siniestro no observables asociados a la conducta del individuo del seguro de autos en una empresa aseguradora peruana bajo el enfoque bayesiano, esto con la finalidad de analizar los problemas de asimetría de información en el mercado asegurador, enfatizando en el estudio del problema de selección adversa en el mercado de seguro de autos. También, se busca impulsar el empleo del método bayesiano y el empleo de la estadística para la asignación de probabilidades a variables desconocidas.

Asimismo, los resultados hallados guardan relación con algunos de los antecedentes empíricos estudiados en la presente investigación tal es el caso de Melgar y Guerrero (2005), que señalan que algunos determinantes tales como la edad, la prima pagada y el tipo de vehículo están relacionados con la ocurrencia de siniestros. Mateus (2008), en su investigación considera que el sexo y la edad están relacionados con la ocurrencia de siniestro. García (2014), indica que las variables como la edad y el

modelo del vehículo están relacionadas con la ocurrencia de siniestro. Quishpe (2015), considera que el género, la edad, la marca son variables que definen al individuo tipo que tiene mayor probabilidad de ocurrencia de siniestro. Segovia Vargas, Camacho Miñano, Pascual Ezama y Rodríguez Piñero (2015), señalan que existen factores ocultos, normalmente no cuantificados relacionados con la ocurrencia de siniestro.

Por otro lado, se puede observar que en algunos casos señalan lo opuesto, por ejemplo: Melgar y Guerrero (2005), señalan que el tipo de vehículo y la prima pagada tienen un impacto positivo en la ocurrencia de siniestro. Mateus (2008), señala que el sexo, la antigüedad del vehículo no tienen un impacto positivo en la ocurrencia de siniestro. García (2014), indica que el impacto de la variable sexo y de la variable tipo de vehículo es negativo en la ocurrencia de siniestro. Quishpe (2015), señala que los factores que determinan la ocurrencia de siniestros son, en orden de influencia: tener vehículos con años de fabricación menores al 2012, de marcas diferentes a Chevrolet y contratar seguros por montos entre USD 15.001 y USD 50.000. Mientras que los factores que disminuyen dicha probabilidad son: tener vehículos diferentes a los automóviles, la edad del conductor del vehículo es superior a los 25 años, y que las zonas de circulación. Asimismo, la variable “Sexo del asegurado” en los resultados econométricos no tiene significancia estadística. Vargas (2016), señala que la edad y el sexo del asegurado tienen mayor influencia en la ocurrencia de siniestro. Por último, Melgar y Ordaz (2013) señalan que el grado de cobertura, los tipos de vehículos y la prima pagada, la edad de los asegurados que conduzcan camión, moto, ciclomotor o vehículos especiales tienen mayor influencia en la ocurrencia de siniestro.

Luego de plantear el modelo logit bajo el enfoque bayesiano, de la iteración de la distribución a priori y la función de máximo verosimilitud se obtienen las siguientes conclusiones:

1. Se encuentra que existe relación entre los determinantes no observables asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana. Sustentado en la estimación del modelo logit bayesiano cuyo valor de la tasa de aceptación de 23.98% es mayor al 10%. Ver tabla N° 13.
2. Se demuestra que los determinantes observables asociados al conductor tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018. Sustentado en la estimación del modelo logit bayesiano dado que los intervalos de confianza para las variables edad, sexo y estado civil del propietario; explican que los resultados son estadísticamente significativos. Ver tabla N° 13.

- La edad del propietario de vehículo tiene un impacto diferencial positivo en la ocurrencia de siniestro, a medida que la edad del propietario aumente mayor será la probabilidad de ocurrencia de siniestro.
  - El sexo del cliente tiene un impacto diferencial negativo en la ocurrencia de siniestro, los clientes de género masculino no tienen mayor probabilidad de ocurrencia de siniestro que las de género femenino.
  - El estado civil del propietario del vehículo tiene un impacto diferencial positivo en la ocurrencia de siniestro, los clientes solteros tienen mayor probabilidad de ocurrencia de siniestro que los casados.
3. Se demuestra que los determinantes observables asociados al vehículo tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018. Sustentado en la estimación del modelo logit bayesiano dado que los intervalos de confianza para las variables periodo de fabricación, marca y obtención del GPS del vehículo; explican que los resultados son estadísticamente significativos. Ver tabla N° 13.
- El año de fabricación del vehículo tiene un impacto diferencial positivo en la ocurrencia de siniestro, es decir cuando la variable años de fabricación del vehículo pertenece al rango de años 2013 al 2018 tiene mayor probabilidad de ocurrencia de siniestro que si perteneciera a otro rango de años.
  - La marca del vehículo tiene un impacto diferencial positivo en la ocurrencia de siniestro, es decir cuando la variable modelo de vehículo es Nissan, Kia, Hyundai o Toyota entonces tiene mayor probabilidad de ocurrencia de siniestro que si fuera de otra marca.
  - La obtención del GPS tiene un impacto diferencial positivo en la ocurrencia de siniestro, los clientes que obtengan el GPS tienen mayor probabilidad de ocurrencia de siniestro que si no lo obtuvieran.
  - El tipo de vehículo tiene un impacto diferencial negativo en la ocurrencia de siniestro, los vehículos de tipo sedán, pick up o hatchback no tienen mayor probabilidad de ocurrencia de siniestro que los vehículos de otros tipos.
4. Se demuestra que los determinantes observables asociados a la póliza tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018. Sustentado en la estimación del modelo logit bayesiano dado que el intervalo de confianza para la variable prima pagada, explica que los resultados son estadísticamente significativos. Ver tabla N° 13.
- La prima pagada por el seguro del vehículo tiene un impacto diferencial positivo en la ocurrencia de siniestro, a medida que la prima pagada aumente entonces la probabilidad de ocurrencia de siniestro será mayor.

## 5.2. LIMITACIONES Y RECOMENDACIONES

1. Se recomienda que las empresas aseguradoras realicen un estudio a fin de asignar un porcentaje, para los determinantes no observables asociados a la conducta aleatoria del individuo, dentro del precio del seguro, debido a que estos factores tienen relación con la ocurrencia de siniestros.
2. Se recomienda a las aseguradoras conocer las características de los conductores que tengan mayor probabilidad de ocurrencia de siniestro dado que esto permitirá elaborar distintos esquemas eficientes y efectivos de administración, gestión y control de los riesgos a los que se encuentren expuestos en el desarrollo de su negocio. Esto es fundamental debido a que forma parte del proceso de toma de decisiones de una aseguradora en especial en casos de análisis de nuevos tipos de seguros o nuevas coberturas. De manera que, las aseguradoras reducen la incertidumbre respecto al riesgo asociado a los conductores de los vehículos que aseguran.
3. Se recomienda a las aseguradoras conocer las características de los vehículos más propensos a tener un siniestro ya que permite tener una mejor administración del riesgo y cuantificar la infraestructura adecuada al momento de hacer frente a un siniestro en función al tipo de vehículo, marca de vehículo, periodo de fabricación y tecnología del GPS. Asimismo, es beneficioso para los asegurados conocer esta información dado que es un factor importante al momento de escoger la cobertura que se comprará para un tipo de vehículo en específico.
4. Se recomienda a la aseguradora analizar y evaluar las primas debido a que estas pueden influenciar en el comportamiento de los conductores. Es decir, cuando hay primas de cobertura total, los conductores son propensos a tener comportamientos indiferentes respecto al cuidado del vehículo. Asimismo, es importante señalar que esta relación se cumple previa a la ocurrencia del siniestro. Si es que ya ocurrió el siniestro, este determina el nuevo valor de la prima.
5. Existen algunas críticas y limitaciones del modelo bayesiano, por ejemplo, la que está referida al supuesto de la distribución a priori y es debido a que el supuesto de un investigador puede diferir completamente del supuesto de otro, por lo tanto, las conclusiones que ambos tengan respecto a un fenómeno son diferentes.

## BIBLIOGRAFÍA

- Agois, E. (14 de diciembre de 2016). *Asociación Automotriz del Perú*. Obtenido de <https://aap.org.pe/descarga/ea2016/RIMAC-Enrique-Agois.pdf>
- Aguiar, F. (2004). Teoría de la decisión e incertidumbre: modelos normativos y descriptivos. *EMPIRIA: Revista de metodología de Ciencias Sociales*, 139-160. Obtenido de <http://espacio.uned.es/fez/eserv/bibliuned:Empiria-2004-BA384079-87D3-BD53-AED2-BCE23A58A02C/Documento.pdf>
- Akerlof, G. (Agosto de 1970). *Università degli studi di Bergamo*. Obtenido de <http://www.data.unibg.it/dati/corsi/8906/37702-Akerlof%20-%20Market%20for%20lemons.pdf>
- American Economic Association. (01 de Junio de 2019). *American Economic Association*. Obtenido de <https://www.aeaweb.org/econlit/jelCodes.php?view=jel#D>
- Andrea, R., Cerruti, O., & Tapia, A. (2007). *Pontificia Universidad Católica del Perú*. Obtenido de <http://revistas.pucp.edu.pe/index.php/strategia/article/view/17524/17800>
- APESEG. (17 de Enero de 2019). *Asociación Peruana de Empresas de Seguros*. Obtenido de Asociación Peruana de Empresas de Seguros: <https://www.apeseg.org.pe/2019/01/una-breve-historia-de-los-seguros/>
- APESEG. (15 de Abril de 2019). *Asociación Peruana de Empresas de Seguros*. Obtenido de <https://www.apeseg.org.pe/2019/04/el-40-de-vehiculos-en-el-peru-aun-no-cuenta-con-soat/>
- Arrow, K. (1973). *Information and Economic Behavior*. Cambridge: Harvard University.
- Bardey, D. (2008). Asimetrías de Información en los Mercados de Seguros: Teoría y Evidencia. *Revista FASECOLDA*, 14-18.
- Bardey, D., & Lesur, R. (2005). Optimal Health Insurance Contract: Is a deductible useful? *Economics Letters*, 313-317.
- Becker, G. (1997). Conducta Irracional y Teoría Económica. *Derecho y Sociedad*, 76-85. Obtenido de <http://revistas.pucp.edu.pe/index.php/derechosociedad/article/view/16667/17001>
- Bedoya, L. (24 de Octubre de 2019). Solo el 23% de vehículos se encuentran asegurados en el Perú. (M. Ruiz, Entrevistador)
- Castaño, M. (2018). La economía del comportamiento: La Racionalidad limitada de los agentes económicos. *Información Comercial Española*, 31- 42.
- Castells, M. (2000). *La Era de la Información*. Madrid: Alianza.

- Chahuara, P. (Julio de 2017). *Repositorio Tesis PUCP*. Obtenido de [http://tesis.pucp.edu.pe/repositorio/bitstream/handle/20.500.12404/9449/CHAHUARA\\_PAULO\\_ROBERTO\\_ESTIMACION\\_BAYESIANA.pdf?sequence=1&isAllowed=y](http://tesis.pucp.edu.pe/repositorio/bitstream/handle/20.500.12404/9449/CHAHUARA_PAULO_ROBERTO_ESTIMACION_BAYESIANA.pdf?sequence=1&isAllowed=y)
- Chib & Greenberg, S. (1995). *Understanding the Metropolis-Hastings Algorithm*. Obtenido de <http://web1.sph.emory.edu/users/hwu30/teaching/statcomp/papers/chibGreenbergMH.pdf>
- CMS Seguros. (8 de Noviembre de 2013). *CMS Seguros*. Obtenido de <https://www.cmsseguros.es/antecedentes-e-historia-del-seguro-de-vehiculos/>
- Dammert Lira, A., Molinelli Aristondo, F., & Carbajal Navarro, M. (2013). *Teoría de la Regulación económica*. Lima: USMP.
- Dávila, N., García, M., Gomez, E., & Perez, J. (2012). *Dialnet*. Obtenido de <https://dialnet.unirioja.es/descarga/articulo/6014739.pdf>
- De Andrés Sánchez, J., & Terceño Gomez, A. (2002). *Dialnet*. Obtenido de <https://dialnet.unirioja.es/descarga/articulo/2469965.pdf>
- Diaz, F. (1991). *Probabilidad y variable aleatoria, nociones y aplicaciones*. Medellín: Universidad Nacional de Colombia.
- Díaz, F. (12 de diciembre de 2018). Mapfre: La demanda de seguros vehiculares solo llega al 23% en Perú, ¿por qué? (F. Gil Mena, Entrevistador)
- E. Train, K. (2009). *Métodos de Elección Discreta con Simulación*.
- Enciclopedia ENCARTA. (2004). *Historia de Londres*.
- Eppen, G., Gould, F., Schmidt, C., Moore, J., & Weatherferd, L. (2000). *Investigación de operaciones en la ciencia administrativa*. Mexico: Prentice Hall.
- Fortiana, J., Claramount, M., & Boj del Val, E. (1999). *Una Alternativa en la Selección de los Factores de Riesgo a Utilizar en el Cálculo de Primas*. Barcelona: Universidad de Barcelona.
- Fundación MAPFRE. (2020). *Fundación MAPFRE*. Obtenido de [https://www.fundacionmapfre.org/fundacion/es\\_es/publicaciones/diccionario-mapfre-seguros/r/riesgo.jsp](https://www.fundacionmapfre.org/fundacion/es_es/publicaciones/diccionario-mapfre-seguros/r/riesgo.jsp)
- García, A. (2014). *Modelo de regresión y estadística, para la siniestralidad de autos de una empresa, dentro de la vigencia 2012-2013*. Obtenido de <https://core.ac.uk/download/pdf/154795747.pdf>
- Gelfand & Smith, A. (1990). *Enfoques basados en muestreo para calcular densidades marginales*. Obtenido de <https://www.tandfonline.com/doi/abs/10.1080/01621459.1990.10476213>

- Gelman & Rubin, A. (1992). *Inference From Iterative Simulation Using Multiple Sequences*.  
Obtenido de [https://projecteuclid.org/download/pdf\\_1/euclid.ss/1177011136](https://projecteuclid.org/download/pdf_1/euclid.ss/1177011136)
- Gran Enciclopedia del Mundo. (1978). *Antecedentes Históricos del Seguro*. Ediciones Bilbao.
- Hachuel, L., Boggio, G., Harvey, G., & Marfetán, D. (Noviembre de 2013). *Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística*. Obtenido de [https://www.fcecon.unr.edu.ar/web-nueva/sites/default/files/u16/Decimocuarta/hachuel\\_y\\_otros\\_estimacion\\_de\\_un\\_modelo\\_logit.pdf](https://www.fcecon.unr.edu.ar/web-nueva/sites/default/files/u16/Decimocuarta/hachuel_y_otros_estimacion_de_un_modelo_logit.pdf)
- Hosmer, D., & Lemeshow, S. (2013). *Applied Logistic Regression*. New York: John Wiley & Sons.
- INEI. (diciembre de 2018). *INEI*. Obtenido de <http://m.inei.gob.pe/estadisticas/indice-tematico/seguridad-ciudadana/>
- Iparraguirre, L. (04 de Noviembre de 2018). *Sólo el 25% de los vehículos en el país cuentan con un seguro vehicular*. Obtenido de <https://andina.pe/agencia/noticia-solo-25-de-vehiculos-el-pais-cuentan-un-seguro-vehicular-731665.aspx>
- Junguito, R. (s.f). *Fundación MAPFRE*. Obtenido de [https://www.fundacionmapfre.org/documentacion/publico/i18n/catalogo\\_imagenes/imagen\\_id.cmd?idImagen=1065519](https://www.fundacionmapfre.org/documentacion/publico/i18n/catalogo_imagenes/imagen_id.cmd?idImagen=1065519)
- Kelisto Seguros. (23 de Septiembre de 2019). *Kelisto*. Obtenido de <https://www.kelisto.es/seguros-coche/consejos-y-analisis/que-factores-afectan-al-precio-del-seguro-de-coche-6107>
- Khan, A. (2003). Abdul Waheed Khan Towards knowledge societies. *A World of Science*, 8 - 9.  
Obtenido de [http://portal.unesco.org/ci/ev.php?URL\\_ID=11958&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201](http://portal.unesco.org/ci/ev.php?URL_ID=11958&URL_DO=DO_TOPIC&URL_SECTION=201)
- Levine, D. (2010). ¿Está la Economía del Comportamiento Condenada a Desaparecer? Lo ordinario frente a lo extraordinario. *El Trimestre Económico*, 509-531.
- Lobera, A. (1994). El seguro en la moderna teoría financiera. *Revista Vasca de Economía*, 94-109.
- López Martínez, H. (2004). *APESEG. 100 años en la historia del seguro*. Lima: Asociación Peruana de Empresas de Seguros.
- Lotero, R. (2005). Teoría de Contratos: Un enfoque económico. *Cuadernos Latinoamericanos de administración*, 11-24. Obtenido de <https://www.redalyc.org/pdf/4096/409634371004.pdf>
- Macho, I., & Pérez, D. (1994). *Introducción a la Economía de la Información*. España: Ariel Economía.



- Maddala, G. (1985). *Econometría*. México: McGraw-Hill .
- Martinez, E. (2007). *Depósito de Investigación de la Universidad de Sevilla*. Obtenido de [https://idus.us.es/bitstream/handle/11441/84128/Logic\\_model.pdf?sequence=1&isAllowed=y#:~:text=Este%20economista%20fue%20pionero%20en,las%20preferencias%20de%20los%20individuos.&text=McFadden%20plante%C3%B3%20inicialmente%20el%20caso,y%20%3D%201%20o%20y%2](https://idus.us.es/bitstream/handle/11441/84128/Logic_model.pdf?sequence=1&isAllowed=y#:~:text=Este%20economista%20fue%20pionero%20en,las%20preferencias%20de%20los%20individuos.&text=McFadden%20plante%C3%B3%20inicialmente%20el%20caso,y%20%3D%201%20o%20y%2)
- Mateus, J. (2008). *Análisis de los seguros de automóviles en Colombia entre 1997 y 2005 de la empresa Liberty Seguros S.A. basado en el nivel de riesgo y tasa de siniestralidad*. Obtenido de <http://repository.lasalle.edu.co/bitstream/handle/10185/12131/T10.08%20M419a.pdf;jsessionid=C209345418E87EC1EE2E30B2AF97F403?sequence=1>
- Mc Fadden, D. (1974). *Universidad de California*. Obtenido de <https://eml.berkeley.edu/reprints/mcfadden/zarembka.pdf>
- Melgar, M., & Guerrero, M. (2005). *Los siniestros en el seguro del automóvil: un análisis econométrico aplicado*. Obtenido de <http://www.redalyc.org/pdf/301/30123117.pdf>
- Melgar, M., & Ordaz, J. (2013). *Aplicación del análisis discriminante al estudio de la siniestralidad en el ramo del seguro de automóviles*. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=4749631>
- Metropolis, N. (1953). *Equation of State Calculations by Fast Computing Machines*. Obtenido de <https://bayes.wustl.edu/Manual/EquationOfState.pdf>
- Ministerio de Transportes y Comunicaciones. (07 de Enero de 2019). *Ministerio de Transportes y Comunicaciones*. Obtenido de <https://www.gob.pe/202-ministerio-de-transportes-y-comunicaciones-seguro-obligatorio-de-accidentes-de-transito-soat>
- Murga, V. (30 de Junio de 2019). *El mercado de seguros en Perú*. Obtenido de <https://culturadeseguros.pe/el-mercado-de-seguros-en-peru/>
- Ntzoufras, I. (2011). *BAYESIAN MODELING USING WINBUGS*. Obtenido de <http://civil.colorado.edu/~balajir/CVEN6833/bayes-resources/chapter05.pdf>
- Pacífico Seguros. (2020). *Manual del Seguro de Autos*. Lima: Pacífico Seguros.
- Picos, J., & Valero, E. (2007). *Frecuencia de la Siniestralidad y Coste Medio del Siniestro como base del Cálculo de la Prima Pura en Seguros Contra Incendios Forestales en Propietarios Forestales en Galicia*. Sevilla.

- Puelles, L. (2 de Enero de 2019). Seguro vehicular: ¿Qué día y en qué horarios se registran mayor cantidad de siniestros? (M. Ríos, Entrevistador)
- Quishpe, I. (2015). *Factores de Riesgo de Siniestralidad y el Cálculo de Primas de los Vehículos asegurados en el Ecuador mediante Modelos Lineales Generalizados*. Quito.
- Rausell Köster, P., & Carrasco Arroyo, S. (2003). *Algunos apuntes sobre la economía de la información, la comunicación y la cultura*. Tirant Lo Blanch.
- Real Academia Española. (2020). *Diccionario de la Real Academia Española*. Obtenido de <https://dle.rae.es>
- Revista de Historia . (21 de Mayo de 2019). *Revista de Historia*. Obtenido de <https://revistadehistoria.es/la-historia-del-primer-seguro-de-coche/>
- Rodriguez Basurto, I. (22 de Mayo de 2007). *Convencion Nacional de Aseguradores*. Obtenido de <https://docplayer.es/596449-El-seguro-obligatorio-de-vehiculos-automotores-en-el-mundo-act-ignacio-rodriguez-basurto.html>
- Rothschild, M., & Stiglitz , J. (Noviembre de 1976). *The Quarterly Journal of Economy*. Obtenido de <https://uh.edu/~bsorensen/Rothschild%26Stiglitz.pdf>
- Salcedo, C. (2002). *Universidad Nacional Mayor de San Marcos*. Obtenido de [https://sisbib.unmsm.edu.pe/bibvirtualdata/Tesis/Basic/Salcedo\\_pc/enPDF/Cap2.PDF](https://sisbib.unmsm.edu.pe/bibvirtualdata/Tesis/Basic/Salcedo_pc/enPDF/Cap2.PDF)
- Sanchez Torres, J., Gonzáles Zavala, M., & Sánchez Muñoz, M. (2012 de Junio de 2012). La Sociedad de la Información: génesis, iniciativas, concepto y su relación con las TIC. *Revista UIS Ingenierías*, 113-128. Obtenido de <https://www.redalyc.org/pdf/5537/553756873001.pdf>
- Segovia Vargas, M., Camacho Miñano, M., Pascual Ezama, D., & Rodriguez Piñero, P. (Diciembre de 2015). *Innovar: Revista de Ciencias Administrativas y Sociales* . Obtenido de [https://www.researchgate.net/publication/286440894\\_Analisis\\_de\\_los\\_factores\\_de\\_riesgo\\_en\\_el\\_seguro\\_de\\_automovil\\_mediante\\_ecuaciones\\_estructurales](https://www.researchgate.net/publication/286440894_Analisis_de_los_factores_de_riesgo_en_el_seguro_de_automovil_mediante_ecuaciones_estructurales)
- Seguros TV. (2 de Agosto de 2016). *Seguros TV*. Obtenido de <https://blog.segurostv.es/conoces-la-historia-del-primer-seguro-de-coche/>
- Shavell, S. (1979). On Moral Hazard and Insurance. *The Quarterly Journal of Economics*, 541-562.
- Specia Jimenez, A. L. (28 de Noviembre de 2005). *Universidad de las Américas Puebla*. Obtenido de [http://catarina.udlap.mx/u\\_dl\\_a/tales/documentos/ledf/specia\\_j\\_al/](http://catarina.udlap.mx/u_dl_a/tales/documentos/ledf/specia_j_al/)

- Spence, M. (Agosto de 1973). *The Quarterly Journal of Economics*. Obtenido de <http://links.jstor.org/sici?sici=0033-5533%28197308%2987%3A3%3C355%3AJMS%3E2.0.CO%3B2-3>
- St. Onge, K. (27 de Febrero de 2008). *Insurance Journal*. Obtenido de <https://revistadehistoria.es/lahistoria-del-primer-seguro-de-coche/>
- Stigler, G. (1961). *Universidad de Chicago*. Obtenido de <https://home.uchicago.edu/~vlima/courses/econ200/spring01/stigler.pdf>
- Stiglitz, J. (Junio de 1975). *Columbia University Libraries*. Obtenido de <https://academiccommons.columbia.edu/doi/10.7916/D89W0RG5/download>
- Stiglitz, J. (2002). La información y el cambio en el paradigma de la ciencia económica. *Revista Asturiana de Economía*, 95-164.
- Superintendencia de Banca y Seguros. (2019). *Evolución del Sistema de Seguros al IV Trimestre del 2019*. Lima: Superintendencia de Banca y Seguros.
- Superintendencia de Banca y Seguros. (2020). *Boletín Estadístico de Seguros Participación en el Mercado*. Lima: Superintendencia de Banca y Seguros.
- Superintendencia de Banca y Seguros. (Junio de 2020). *Superintendencia de Banca y Seguros*. Obtenido de <https://www.sbs.gob.pe/usuarios/informacion-de-seguros/contratacion-de-seguros/contrato-de-seguro>
- Thaler, R. (2018). Economía del Comportamiento: Pasado, presente y futuro. *Revista de Economía Institucional*, 9-43.
- Usategui, J. (1999). Información Asimétrica y Mecanismos de Mercado. *Ekonomiaz: Revista Vasca de Economía*, 116-141.
- Vargas, T. (2016). *La Siniestralidad del Seguro Vehicular y las Reservas Técnicas de Sweaden Seguros Sucursal Ambato en el año 2014*. Obtenido de <http://repo.uta.edu.ec/bitstream/123456789/22948/1/T3628e.pdf>
- Vásquez, R. (2015). *La teoría de agencia en el ámbito de los seguros de salud*. Obtenido de [https://ruc.udc.es/dspace/bitstream/handle/2183/14632/VazquezNunez\\_Raquel\\_TFG\\_2015.pdf?sequence=2&isAllowed=y](https://ruc.udc.es/dspace/bitstream/handle/2183/14632/VazquezNunez_Raquel_TFG_2015.pdf?sequence=2&isAllowed=y)
- Vial, B., & Zurita, F. (2011). *Microeconomía*. Santiago del Chile: Pontificia Universidad Católica de Chile.

Villaseca, J., Torrent, J., & Lladós, J. (2001). De la economía de la información a la economía del conocimiento: Algunas consideraciones conceptuales y distintivas . *Tendencias*.

Yupanqui, R. (2005). *Tesis Digitales UNMSM*. Obtenido de [http://sisbib.unmsm.edu.pe/bibvirtualdata/Tesis/Basic/yupanqui\\_pr/yupanqui\\_pr.pdf](http://sisbib.unmsm.edu.pe/bibvirtualdata/Tesis/Basic/yupanqui_pr/yupanqui_pr.pdf)

## ANEXOS

## Anexo 1: Matriz de Consistencia

<b>TÍTULO</b>				
<b>DETERMINANTES DE LA SINIESTRALIDAD DE AUTOS EN UNA EMPRESA ASEGURADORA PERUANA: UN ENFOQUE BAYESIANO</b>				
<b>PROBLEMA</b>	<b>OBJETIVO</b>	<b>HIPÓTESIS</b>	<b>VARIABLES</b>	<b>MARCO METODOLÓGICO</b>
<b>PROBLEMA GENERAL.</b>	<b>OBJETIVO GENERAL.</b>	<b>HIPÓTESIS GENERAL.</b>	<b>VARIABLE DEPENDIENTE</b>	<b>ENFOQUE DE LA INVESTIGACIÓN</b>
¿Cuál es la relación entre los determinantes no observables asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana?	Encontrar la relación que existe entre los determinantes no observables asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana.	Existe una relación positiva entre los determinantes no observables asociados a la conducta aleatoria del individuo y la ocurrencia de siniestros vehiculares en una compañía aseguradora peruana.	Ocurrencia del siniestro del vehículo	Cuantitativo
<b>PROBLEMAS ESPECÍFICOS.</b>	<b>OBJETIVOS ESPECÍFICOS.</b>	<b>HIPÓTESIS ESPECÍFICAS.</b>	<b>VARIABLES INDEPENDIENTES</b>	<b>TIPO DE INVESTIGACIÓN</b>
1. ¿Cuál es la relación de los determinantes observables asociados al conductor con la ocurrencia de siniestros vehiculares durante los	1. Demostrar que los determinantes observables asociados al conductor tienen una relación con la ocurrencia	1. Existe una relación entre los determinantes observables asociados al conductor con la ocurrencia de siniestros	Periodo de Fabricación del Vehículo Marca del Vehículo Obtención del GPS	No experimental

<p>años 2016, 2017 y 2018?</p> <p>1. ¿Cuál es la relación de los determinantes observables asociados al vehículo con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018?</p> <p>3. ¿Cuál es la relación de los determinantes observables asociados a la póliza con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018?</p>	<p>de siniestros vehiculares durante los años 2016, 2017 y 2018</p> <p>2. Demostrar que los determinantes observables asociados al vehículo tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018</p> <p>3. Demostrar que los determinantes observables asociados a la póliza tienen una relación con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018</p>	<p>vehiculares durante los años 2016, 2017 y 2018</p> <p>2. Existe una relación entre los determinantes observables asociados al vehículo con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018</p> <p>3. Existe una relación entre los determinantes observables asociados a la póliza con la ocurrencia de siniestros vehiculares durante los años 2016, 2017 y 2018</p>	<p>Edad del Propietario Sexo del Propietario Tipo de vehículo de la aseguradora Estado civil del propietario del vehículo Prima Pagada No observables relacionados a la conducta del individuo</p>	
--	---	---	--	--

## Anexo 2: Empresas de Seguros Autorizadas al 31 de diciembre del 2019

### Empresas de Seguros Autorizadas

(Al 31 de diciembre del 2019)

Ramos Generales y de Vida	Ramos Generales	Ramos de Vida
<b>Chubb Perú</b>	La Positiva	La Positiva Vida
<b>Cardif</b>	Mapfre Perú	Mapfre Perú Vida
<b>Interseguro</b>	Secrex	Rigel
<b>Rímac</b>	Insur	Vida Cámara
<b>Pacífico Seguros</b>	Qualitas	Ohio National Vida
<b>Protecta</b>	Avla Perú	
<b>Crecer Seguros</b>	Coface	
	Liberty	

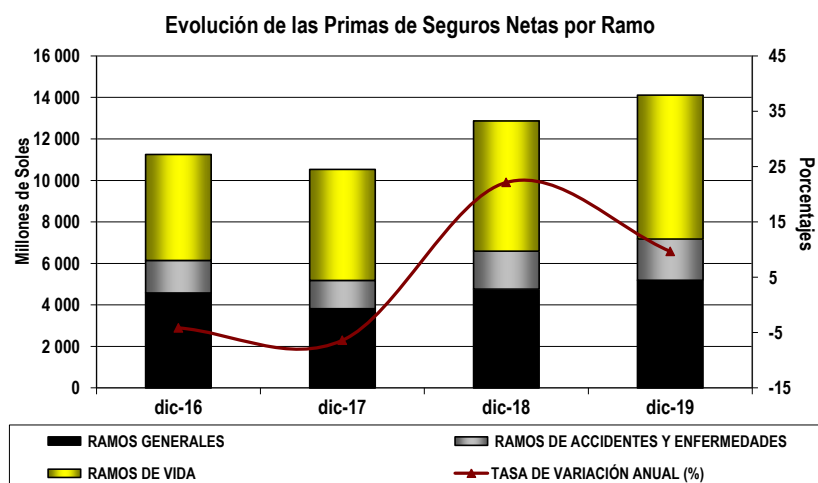
## Anexo 3: Estructura de las Primas de Seguros Netas por Ramos

### Estructura de las Primas de Seguros Netas, por Ramos

(En porcentajes)

	dic-16	dic-17	dic-18	dic-19
Ramos Generales	40,7	36,3	37,0	36,8
Accidentes y Enfermedades	13,9	12,9	14,2	14,0
Seguros de Vida	24,5	32,3	33,5	34,1
Seguros del SPP	21,0	18,5	15,3	15,0
<b>TOTAL</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>

#### Anexo 4: Evolución de las Primas de Seguros Netas por Ramo



#### Anexo 5: Ranking del Total de Primas de Seguros

<b>Total Primas de Seguros</b>				
(al 31 de enero del 2020)				
	<b>Empresas</b>	<b>Monto</b>	<b>Participación (%)</b>	<b>Porcentaje Acumulado</b>
1	Rimac	384 978	30.69	30.69
2	Pacifico Seguros	324 656	25.88	56.57
3	Mapfre Perú	94 536	7.54	64.10
4	La Positiva	84 033	6.70	70.80
5	Interseguro	80 214	6.39	77.20
6	La Positiva Vida	66 477	5.30	82.49
7	Mapfre Perú Vida	55 936	4.46	86.95
8	Protecta	40 516	3.23	90.18
9	Cardif	32 258	2.57	92.75
10	Chubb Seguros	17 885	1.43	94.18
11	Ohio National Vida	17 737	1.41	95.59
12	AVLA Perú	14 172	1.13	96.72
13	Secrex	10 109	0.81	97.53
14	Crece Seguros	9 408	0.75	98.28
15	Vida Cámara	8 867	0.71	98.99
16	Insur	8 645	0.69	99.68
17	Qualitas	2 296	0.18	99.86
18	Liberty Seguros	1 785	0.14	100.00
19	Rigel	6	0.00	100.00
20	Coface Seguros	( 16)	(0.00 )	100.00
	<b>Total</b>	<b>1 254 498</b>	<b>100.00</b>	



**Anexo 6: Ranking del Total de Primas de Seguros para los ramos generales, accidentes y enfermedades**

<b>Ramos Generales y Accidentes y Enfermedades</b>				
<b>(al 31 de enero del 2020)</b>				
<b>Empresas</b>		<b>Monto</b>	<b>Participación (%)</b>	<b>Porcentaje Acumulado</b>
1	Rímac	208 898	34.23	34.23
2	Pacífico Seguros	151 327	24.80	59.03
3	Mapfre Perú	94 536	15.49	74.52
4	La Positiva	84 033	13.77	88.29
5	Chubb Seguros	16 419	2.69	90.98
6	AVLA Perú	14 172	2.32	93.30
7	Secrex	10 109	1.66	94.96
8	Interseguro	9 657	1.58	96.54
9	Insur	8 645	1.42	97.96
10	Cardif	5 567	0.91	98.87
11	Qualitas	2 296	0.38	99.24
12	Liberty Seguros	1 785	0.29	99.54
13	Creceer Seguros	1 507	0.25	99.78
14	Protecta	677	0.11	99.89
15	Mapfre Perú Vida	647	0.11	100.00
16	La Positiva Vida	17	0.00	100.00
17	Rigel	-	-	100.00
18	Vida Cámara	-	-	100.00
19	Ohio National Vida	-	-	100.00
20	Coface Seguros	( 16)	(0.00 )	100.00
<b>Total</b>		<b>610 277</b>	<b>100.00</b>	

**Anexo: Ranking del Total de Primas de Seguros para el ramo vida**

<b>Ramos de Vida</b>				
<b>(al 31 de enero del 2020)</b>				
<b>Empresas</b>		<b>Monto</b>	<b>Participación (%)</b>	<b>Porcentaje Acumulado</b>
1	Rímac	176 080	27.33	27.33
2	Pacífico Seguros	173 328	26.91	54.24
3	Interseguro	70 557	10.95	65.19
4	La Positiva Vida	66 459	10.32	75.51
5	Mapfre Perú Vida	55 289	8.58	84.09
6	Protecta	39 840	6.18	90.27
7	Cardif	26 691	4.14	94.42

8	Ohio National Vida	17 737	2.75	97.17
9	Vida Cámara	8 867	1.38	98.55
10	Crecer Seguros	7 901	1.23	99.77
11	Chubb Seguros	1 465	0.23	100.00
12	Rigel	6	0.00	100.00
13	Insur	-	-	100.00
14	La Positiva	-	-	100.00
15	Mapfre Perú	-	-	100.00
16	Secrex	-	-	100.00
17	Qualitas	-	-	100.00
18	AVLA Perú	-	-	100.00
19	Coface Seguros	-	-	100.00
20	Liberty Seguros	-	-	100.00
<b>TOTAL</b>		<b>644 221</b>	<b>100.00</b>	

### Anexo 7: Archivo do en software econométrico STATA

\*Modelo de Siniestros -Seguros

clear all

cd "G:\ARCHIVOS\ARCHIVOS LIESEL\resultados"

use DATA\_SINIESTROS27.dta, clear

\*Variable dependiente

gen sin=.

replace sin=1 if num\_sin!=.

replace sin=0 if num\_sin==.

\*Variables independientes

gen inc\_gps=.

replace inc\_gps=1 if ind\_gps\_slip=="SI"

replace inc\_gps=0 if ind\_gps\_slip=="NO"

codebook inc\_gps

gen fab\_veh=0

replace fab\_veh=1 if ano\_fabricacion\_veh>=2013 & ano\_fabricacion\_veh<=2018 // 2 veces

replace fab\_veh=0 if ano\_fabricacion\_veh!=2013 | ano\_fabricacion\_veh!=2014 | ///

ano\_fabricacion\_veh!=2015 | ano\_fabricacion\_veh!=2016 | ano\_fabricacion\_veh!=2017 |

ano\_fabricacion\_veh!=2017

```

gen marca_hknt=.

replace marca_hknt=1 if marca_veh=="NISSAN" | marca_veh=="KIA" | marca_veh=="HYUNDAI" |
marca_veh=="TOYOTA" // 2 veces

replace marca_hknt=0 if marca_veh!="NISSAN" | marca_veh!="KIA" | marca_veh!="HYUNDAI" |
marca_veh!="TOYOTA"

codebook marca_hknt

gen tipo_vehiculo=.

replace tipo_vehiculo=1 if tipo_veh=="HATCHBACK " | tipo_veh=="SEDAN" | tipo_veh=="PICK UP"
// 2 veces

replace tipo_vehiculo=0 if tipo_veh!="HATCHBACK " | tipo_veh!="SEDAN" | tipo_veh!="PICK UP"

codebook tipo_vehiculo

gen edad_cliente=edad_cli_vig

gen gen_masculino=.

replace gen_masculino=1 if genero_cliente_cont=="MASCULINO"
replace gen_masculino=0 if genero_cliente_cont!="MASCULINO"

codebook gen_masculino

gen ec_soltero=.

replace ec_soltero=1 if estado_civil_cliente_cont=="SOLTERO"
replace ec_soltero=0 if estado_civil_cliente_cont!="SOLTERO"

codebook ec_soltero

gen prima_neta=prima_neta_cc_cert

*****

keep sin fab_veh marca_hknt tipo_vehiculo inc_gps edad_cliente gen_masculino ///
ec_soltero prima_neta tipo_cliente_cont cliente_region

*Renombrar variables independientes

gen fab=fab_veh

gen mak=marca_hknt

gen gps=inc_gps

gen edc=edad_cliente

gen sex=gen_masculino

gen esc=ec_soltero

gen pra=prima_neta

gen tveh=tipo_vehiculo

```

\*Tablas descriptivas

```

tabulate sin if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA"
table sin fab if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", row sc
table sin mak if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", row sc
table sin gps if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", row sc
table sin sex if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", row sc
table sin tveh if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", row sc
table sin esc if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", row sc
table sin if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", c(edc med edc mean edc
max edc sd edc)
table sin if cliente_region=="REGION LIMA" & tipo_cliente_cont=="PERSONA", c(pra med pra mean pra
max pra sd pra)

```

\*Modelo estimado

```

logit sin fab mak gps edc sex tveh esc pra if cliente_region=="REGION LIMA" &
tipo_cliente_cont=="PERSONA"

quietly fitstat, saving(Logi)

probit sin fab mak gps edc sex tveh esc pra if cliente_region=="REGION LIMA" &
tipo_cliente_cont=="PERSONA", iter(10)

fitstat, using(Logi) force

```

\*Modelo Bayesiano

```

set seed 123

bayesmh sin fab mak gps edc sex tveh esc pra if cliente_region=="REGION LIMA" ///
& tipo_cliente_cont=="PERSONA", ///
likelihood(logit) prior({sin:}, normal(0,1)) ///
mcmcsize(20000) burnin(5000)

bayesgraph histogram _all, combine(rows(5))

```

