



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**APLICACIÓN DE LOS MODELOS LINEALES
GENERALIZADOS PARA EL CÁLCULO DE PRIMAS
DE TARIFA DE LA COBERTURA DAÑOS
MATERIALES DEL RAMO DE AUTOMÓVILES**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

P R E S E N T A :

HEBER TRUJILLO URIOSTE



**DIRECTOR DE TESIS:
ACT. GABRIEL VARGAS VILCHIS
2017**

CIUDAD UNIVERSITARIA, CD.MX.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

1. Datos del alumno
Trujillo
Urioste
Heber
55 85 46 21 83
Universisad Nacional Autónoma de
México
Facultad de Ciencias
Actuaría
308114290
2. Datos del tutor
Actuario
Gabirl
Vargas Vilchis
3. Datos del Sinodal 1
Dra.
Ruth Selene
Fuentes
García
4. Datos del Sinodal 2
Dr.
Ricardo
Ramírez
Aldana
5. Datos del Sinodal 3
Actuario
José Fernando
Soriano
Flores
6. Datos del Sinodal 4
Actuario
Abraham
Chávez
Sánchez
7. Datos del Trabajo Escrito
Aplicación de los modelos lineales
generalizados para el cálculo de
primas de tarida de la cobertura

daños materiales del remo de automóviles

154 p

2017

ÍNDICE

INTRODUCCIÓN	1
CAPÍTULO 1	3
1 Fundamentos de la Teoría de los Modelos Lineales Generalizados	3
1.1 Supuestos de los Modelos Lineales Generalizados	3
1.2 Medidas de Dispersión	4
1.2.1 Severidad: Media y Varianza	5
1.2.2 Frecuencia: Media y Varianza	5
1.3 Modelo Multiplicativo	5
1.4 Modelos de Dispersión Exponencial	6
1.4.1 Función Generadora y Función Cumulante	7
1.4.2 Modelo Poisson para la Estimación de la Frecuencia de los Siniestros	9
1.4.3 Modelo Gamma para la Estimación de la Severidad de los Siniestros	11
1.5 Función Liga	12
1.5.1 Función Liga Logarítmica	13
1.5.2 Función Liga Logit	14
1.6 Estimación de Parámetros	14
1.6.1 Resultado General para el Método de Máximo Log-Verosimilitud	14
1.6.2 Modelo Multiplicativo Poisson para la frecuencia	16
1.6.3 Modelo Multiplicativo Gamma para la Severidad	16
1.7 Análisis de la Devianza del Modelo	17
1.8 Estimación del Parámetro de Dispersión	18
1.8.1 Estimación del Parámetro de Dispersión por el Método de Pearson	18
1.8.2 Estimación del Parámetro de Dispersión por el Método de la Devianza	19
1.9 Prueba de Hipótesis para Modelos Jerárquicos	19
1.10 Intervalos de Confianza de Fisher	20
1.10.1 Matriz de Información de Fisher	20
1.10.2 Intervalos de Confianza contruidos con la información de Fisher	21
1.11 Análisis de Residuos	22
1.11.1 Residuos de Pearson	22
1.11.2 Residuos de Anscombe	23
1.11.3 Residuos de Devianza	23
1.12 Sobre-dispersión	23
1.13 Consideraciones Especiales al Ajustar un Modelo	25
1.13.1 Interacción Entre Variables	25

1.13.2 Offset	27
CAPÍTULO 2	28
2.1 Características de la Cobertura	28
2.1.1 Riesgos Amparados	28
2.1.2 Deducibles	28
2.2 Delimitación de la Información	29
2.3 Ajuste Preliminar de la Estructura de Error y Tratamiento de Datos Atípicos	29
2.3.1 Montos Negativos	30
2.3.2 Ajuste por Pérdidas con Monto Pequeño	31
2.3.2.1 Ajuste de Pérdidas Parciales	32
2.3.2.2 Ajuste de Pérdidas Totales	34
2.3.2.3 Ajuste de Rotura de Cristales	37
2.3.3 Ajuste por Pérdidas con Monto Excesivo	39
2.3.3.1 Ajuste de Pérdidas Parciales	40
2.3.3.2 Ajuste de Pérdidas Totales	42
2.3.3.3 Ajuste Rotura de Cristales	43
2.3.4 Ajuste por Monto de Alta Frecuencia	44
2.4 Unión del Archivo plano de Siniestros con el Archivo Plano de Emisión	46
2.5 Análisis Exploratorio de la Información Estadística del Sector Asegurador Mexicano	47
2.5.1 Variables dependientes y de Ponderación	48
2.5.1.1 Variables de Respuesta	48
2.5.1.1.1 Frecuencia: Número de Siniestros	48
2.5.1.1.2 Severidad: Monto Neto de Siniestros	48
2.5.1.2 Variables Ponderadoras	49
2.5.1.2.1 Frecuencia: Exposición	49
2.5.1.2.2 Severidad: Número de Siniestros	50
2.5.2 Variables Independientes	50
2.5.2.1 Variables Geográficas	50
2.5.2.1.1 Entidad Federativa	51
2.5.2.1.1.1 Pérdidas Parciales	51
2.5.2.1.1.2 Pérdidas Totales	54
2.5.2.1.1.3 Rotura de Cristales	56
2.5.2.2 Variables del Vehículo	58
2.5.2.2.1 Marca, Tipo (Sub-Marca), Segmento y Tipo de Vehículo	59
2.5.2.2.1.1 Pérdidas Parciales	60
2.5.2.2.1.2 Pérdidas Totales	62
2.5.2.2.1.2 Rotura de Cristales	63
2.5.2.2.2 Modelo	65
2.5.2.2.2.1 Pérdidas Parciales	66
2.5.2.2.2.2 Pérdidas Totales	69
2.5.2.2.2.3 Rotura de Cristales	71
2.5.2.2.3 Uso del Vehículo	73
2.5.2.2.3.1 Pérdidas Parciales	73
2.5.2.2.3.2 Pérdidas Totales	75
2.5.2.2.3.3 Rotura de Cristales	77
2.5.2.3 Características de Contratación de la Póliza	78
2.5.2.3.1 Deducible	78
2.5.2.3.1.1 Pérdidas Parciales	79
2.5.2.3.1.2 Pérdidas Totales	80

CAPÍTULO 3	82
3.1 Ajuste de los Modelos	82
3.1.1 Definición de los Niveles Base	82
3.1.2 Información General del Modelo y Estimación de los Parámetros β 's	82
3.1.2.1 Frecuencia de Pérdidas Parciales	82
3.1.2.2 Severidad de Pérdidas Parciales	84
3.1.2.3 Frecuencia de Pérdidas Totales	86
3.1.2.4 Severidad de Pérdidas Totales	87
3.1.2.5 Frecuencia Rotura de Cristales	89
3.1.2.6 Severidad Rotura de Cristales	90
3.2 Pruebas de Bondad de Ajuste	92
3.2.1 Intervalos de Confianza	92
3.2.1.1 Frecuencia Pérdidas Parciales	92
3.2.1.2 Severidad Pérdidas Parciales	95
3.2.1.3 Frecuencia Pérdidas Totales	99
3.2.1.4 Severidad Pérdidas Totales	102
3.2.1.5 Frecuencia Rotura de Cristales	105
3.2.1.6 Severidad Rotura de Cristales	108
3.2.2 Test de Consistencia aleatoria	110
3.2.2.1 Frecuencia Pérdidas Parciales	110
3.2.2.2 Severidad Pérdidas Parciales	114
3.2.2.3 Frecuencia Pérdidas Totales	117
3.2.2.4 Severidad Pérdidas Totales	120
3.2.2.5 Frecuencia Rotura de Cristales	123
3.2.2.6 Severidad Rotura de Cristales	125
3.2.3 Residuos	128
3.2.3.1 Frecuencia Pérdidas Parciales	128
3.2.3.2 Severidad Pérdidas Parciales	129
3.2.3.3 Frecuencia Pérdidas Totales	130
3.2.3.4 Severidad Pérdidas Totales	131
3.2.3.5 Frecuencia Rotura de Cristales	132
3.2.3.6 Severidad Rotura de Cristales	133
3.2.4 Curva de Lift	134
3.2.4.1 Frecuencia Pérdidas Parciales	135
3.2.4.2 Severidad Pérdidas Parciales	136
3.2.4.3 Frecuencia Pérdidas Totales	137
3.2.4.4 Severidad Pérdidas Totales	138
3.2.4.5 Frecuencia Rotura de Cristales	139
3.2.4.6 Severidad Rotura de Cristales	140
3.2.5 Conclusiones	141
CAPÍTULO 4	143
4.1 Modelo de Daños Materiales	143
4.2 Mutualización de los Datos Excluidos	144
4.3 Inclusión de las Hipótesis Financieras	145
4.4 Resumen Numérico	146
CONCLUSIONES Y COMENTARIOS	150

BIBLIOGRAFÍA	152
GLOSARIO DE TÉRMINOS	153

INTRODUCCIÓN

Desde la década de los ochenta los modelos lineales generalizados fueron aceptados como una alternativa para el cálculo de primas de riesgo. Hoy día en la práctica actuarial de los mercados más sofisticados es común el uso de estos modelos; sin embargo, en el Sector Asegurador Mexicano, la metodología y aplicación de la misma están comenzado a explorarse.

Este trabajo escrito tiene como objetivo exponer la manera en que los modelos lineales generalizados pueden ser empleados para la estimación de las primas de riesgo¹. Es conveniente especificar que la aplicación de este trabajo tendrá efecto sobre la cobertura de Daños Materiales del Ramo de Automóviles de pólizas individuales, con vigencia anual.

Como primer paso se expondrá de manera breve el marco teórico de los modelos lineales generalizados, abordándolos desde una perspectiva apropiada para el cálculo de la prima de riesgo de la cobertura de Daños Materiales para el Ramo de Automóviles de pólizas individuales. Luego y con la ayuda de una herramienta especializada para la manipulación y análisis de grandes volúmenes de información, se prepararán los datos, cuya importancia en el proceso es fundamental para garantizar la correcta estimación de cualquier modelo. Cabe señalar que la formación de todo actuario debe envolver el uso de herramientas de programación, pues facilitan el entendimiento y administración de los riesgos del Sector Asegurador Mexicano.

Una vez que los datos se encuentren de manera óptima para ser procesados se efectuará el ajuste de los modelos correspondientes bajo los estándares estadísticos que garanticen un correcto ajuste y consistencia en el poder predictivo de los modelos. Habiendo confirmado el mejor ajuste posible restará recargar todos los costos comunes de una Compañía Aseguradora ajenos al pago de siniestros, como por ejemplo: las comisiones por ventas, los gastos de operación, margen de utilidad, etc. Este último impacto en el modelo de riesgo, debe de efectuarse de manera correcta, ya que terminará de generar la certidumbre necesaria en torno a la naturaleza del negocio de una Compañía Aseguradora para gozar de un ambiente que permita cumplir de rentabilidad.

Finalmente es importante mencionar que la alternativa para el cálculo de la prima de tarifa propuesta en esta tesis cumple con el Estándar de Práctica Actuarial No. 1 de primas, que propone²:

“Principio 1.- La prima de tarifa es la cantidad necesaria para cubrir, al menos, el valor esperado de los costos futuros.”

“Principio 2.- La prima de tarifa debe garantizar suficiencia y solidez.”

¹ La prima de riesgo es equivalente al monto que debe aportar cada elemento asegurado para indemnizar los daños cubiertos en que cada uno puede incurrir

² Los principios se tomaron del Comité de Estándares de Práctica Actuarial de la Asociación Mexicana de Actuarios, A.C. 2003. Estándares de Práctica Actuarial No. 01.

“Principio 3.- La prima de tarifa debe reconocer las características individuales o particulares de la unidad expuesta al riesgo.”

“Principio 4.- La determinación de la prima de tarifa debe sustentarse sobre bases actuariales.”

CAPÍTULO 1

1 Fundamentos de la Teoría de los Modelos Lineales Generalizados³

Los modelos de regresión lineal múltiple y simple son usados en una amplia gama del conocimiento con la finalidad de predecir los posibles resultados de una variable aleatoria, bajo las siguientes restricciones⁴:

1. Se asume una distribución normal para los errores.
2. La media es una función lineal de las variables exógenas.

En la práctica actuarial el uso de los modelos de regresión lineal pudiera, en la mayoría de los casos, no ser una opción para el cálculo de primas de riesgo de una compañía de seguros dados los dos anteriores supuestos; 1: La distribución de los montos de siniestros debería de ser positiva y la distribución normal toma valores en los reales negativos; 2: Es más adecuada una función no lineal cuando se estima la media de la prima de riesgo de la cobertura de Daños Materiales del Ramo de Autos, tal y como lo veremos al comparar el modelo multiplicativo con el aditivo.

Dado lo anterior, en este primer capítulo, se comprobará si los supuestos de los modelos lineales generalizados se ajustan a las características de la cobertura de Daños Materiales para el Ramo de Autos. Como segundo paso se hablará tanto de la media y varianza, como del modelo multiplicativo para finalmente abordar los fundamentos de los Modelos Lineales Generalizados.

1.1 Supuestos de los Modelos Lineales Generalizados

Antes de comenzar a enlistar los supuestos de los Modelos Lineales Generalizados es necesario explicar un poco acerca de la estructura matemática de la frecuencia y severidad.

Ambos componentes, frecuencia y severidad, son el resultado numérico de dividir una variable aleatoria y una ponderadora. La variable ponderadora, desde ahora denotada como ω , derivará de una variable respuesta, también desde ahora denotada por X . Ahora bien el cálculo del valor esperado y la varianza se ejecutarán sobre el siguiente cociente $Y = X/\omega$, en lugar de hacerlo sobre la variable respuesta en sí. Tal como podríamos llegar a sospechar la relación entre variable ponderadora y la volatilidad del modelo es inversamente proporcional. La siguiente tabla resume el caso de la frecuencia y severidad.

Tabla 1.1

Ponderador ω	Variable respuesta X	Cociente $Y = X/\omega$
Exposición	Número de siniestros	Frecuencia de siniestros
Número de siniestros	Costo de siniestros	Severidad de siniestros

³ El primer capítulo incluye traducciones y referencias de los libros [2], [5] y [6] de la bibliografía

⁴ Estas restricciones aplican de un modo más flexible en el caso de los modelos Lineales Generalizados

Ahora bien, los supuestos para el ajuste de un Modelo Lineal Generalizado aplicados al cálculo de primas son:

Supuesto 1:

Independencia entre pólizas: Consideremos n diferentes pólizas individuales. Para cualquier variable respuesta, sea X_i la variable respuesta para la póliza $i = 1, \dots, n$. Entonces X_1, X_2, \dots, X_n son independientes.

El anterior supuesto es equivalente a decir que el número o monto de siniestros es independiente entre las distintas pólizas que posee una misma cartera de estudio. En la realidad puede suceder que dadas dos pólizas pertenecientes a una misma compañía ambas estén involucradas en el mismo siniestro, no obstante el registro de ambos siniestros será en coberturas distintas, los daños propios serán cargados a la cobertura de Daños Materiales, mientras que los daños a terceros pertenecen a la cobertura de Responsabilidad Civil. De tal modo que la operación del seguro favorece a este supuesto.

Supuesto 2:

Independencia con respecto al tiempo: Consideremos n intervalos de tiempo disjuntos. Para cualquier variable respuesta, sea X_i la variable respuesta del intervalo de tiempo i . Entonces X_1, X_2, \dots, X_n son independientes.

Este supuesto dicta que si X_1 es la variable aleatoria que representa el número de siniestros del mes uno, enero, del año en curso y X_2 es la variable aleatoria que representa el número de siniestros del mes dos, febrero, del mismo año, entonces X_1 será independiente de X_2 .

Si sumamos el efecto de los dos anteriores supuestos podríamos concluir que el costo de los siniestros individuales es independiente.

Supuesto 3:

Homogeneidad: Consideremos cualesquiera dos pólizas en una misma celda de tarifa ⁵que posean la misma exposición. Para cualquier variable respuesta, sea X_i la variable respuesta para la póliza i . Entonces X_1 y X_2 tienen la misma distribución de probabilidad.

Tal y como podríamos llegar a pensar, el supuesto de homogeneidad no se cumple cabalmente, sin embargo para hacer frente a las diferencias dentro de una misma celda de tarifa existe el sistema *bonus/malus*, un sistema de recargos y descuentos.

1.2 Medidas de Dispersión

Tal y como lo menciona el principio número 1 del Estándar de Práctica Actuarial No. 1, el análisis se efectuará sobre el valor esperado, pero para ser congruente con el principio número 2

⁵ Es un subconjunto de una cartera de pólizas, cuyos elementos poseen características de riesgo similares entre sí.

es necesario también entender la variación de los datos con respecto al valor esperado, i.e., analizar la varianza.

1.2.1 Severidad: Media y Varianza

Sea ω el número de siniestros y X la suma de las respuestas individuales de los ω siniestros Z_1, \dots, Z_ω , entonces $X = \sum_{k=1}^{\omega} Z_k$ será el costo agregado de todos los siniestros Z_k 's $k = 1, \dots, \omega$. Los supuestos 1 y 2 implican que las Z_k 's son independientes ya que los siniestros provienen de diferentes pólizas o diferentes puntos en el tiempo; el supuesto 3 implica una idéntica distribución, así que podemos escribir a $E[Z_k] = \mu$ y $Var[Z_k] = \sigma^2$, ahora bien, por las reglas básicas de la esperanza y varianza podemos decir que $E[X] = \sum_{k=1}^{\omega} E[Z_k] = \omega\mu$ y $Var[X] = \omega\sigma^2$. Esta sería la media para $Y = \frac{X}{\omega}$ igual a la severidad de siniestros, de tal forma que $E[Y] = E\left[\frac{X}{\omega}\right] = \mu$ y $Var[Y] = Var\left[\frac{X}{\omega}\right] = \sigma^2/\omega$.

1.2.2 Frecuencia: Media y Varianza

Sea ω la exposición, que de manera general y en la práctica, siempre podrá ser expresado de la forma $\omega = \frac{m}{n}$ de tal modo que m es la cantidad de días que estuvo expuesto un grupo de pólizas y n será el número de días que posee un año, puesto que tradicionalmente la prima se calcula de manera anual. Sea Z_i la variable respuesta que representa el número de siniestros que ocurrieron en la fracción $\frac{1}{n}$, donde cada siniestro proviene de intervalos de tiempo distintos (días) o bien de un mismo intervalo de tiempo pero de pólizas distintas, por los tres supuestos anteriores las Z_i 's serán independientes e idénticamente distribuidas, de tal modo que si Z es la suma de n de estas variables, Z poseerá exposición $\omega = 1$, $E[Z] = \mu$ y $Var[Z] = \sigma^2$ además, por las reglas básicas de la esperanza y varianza, sucede que la $E[Z_i] = \mu/n$ y $Var[Z_i] = \sigma^2/n$. Si X es la suma de m intervalos de tiempo entonces $E[X] = \sum_{k=1}^m E[Z_k] = \mu \frac{m}{n} = \mu\omega$ y $Var[X] = \sum_{k=1}^m Var[Z_k] = \sigma^2 \frac{m}{n} = \sigma^2\omega$, simultáneamente sucede que $E[Y] = E\left[\frac{X}{\omega}\right] = \mu$ y $Var[Y] = \sigma^2/\omega$.

1.3 Modelo Multiplicativo

La idea principal de los Modelos Lineales Generalizados, aplicados al cálculo de primas de riesgo, es asignar de acuerdo a las características de riesgo una prima justa. Lo anterior puede lograrse si afectamos el costo por factores que, en el sector asegurador, son conocidos como factores de riesgo. En esta sección describiremos de manera breve el modelo multiplicativo, mismo que nos permite ajustar el costo de cada póliza a las características que ésta posea.

Supongamos, por simplicidad que nuestra tarifa por ahora sólo hace diferencia entre el género del conductor: masculino y femenino y Tipo de vehículo: pequeño y grande. De este modo tendríamos algo similar a la tabla siguiente.

Tabla 1.2

<i>Género/ Tipo de Vehículo</i>	<i>Pequeño</i>	<i>Grande</i>
<i>Hombre</i>	$\mu_{1,1}$	$\mu_{1,2}$
<i>Mujer</i>	$\mu_{2,1}$	$\mu_{2,2}$

Donde $\mu_{1,2}$, por ejemplo, es la media de la frecuencia o severidad para la celda de tarifa Género: Hombre y Tipo de Vehículo: Grande.

La pregunta ahora sería ¿Cómo calcular la media de cada celda? En la mayoría de los casos no contamos con suficiente información como para poder estimar de manera confiable la media de cada celda de la tabla anterior. Es por eso que de manera general definimos el modelo multiplicativo como sigue:

El modelo multiplicativo, aplicable al ejemplo anterior sería:

$$\mu_{i,j} = \gamma_0 \gamma_{1,i} \gamma_{2,j}.$$

Donde $\gamma_{1,i}$ es el i -ésimo nivel del primer factor: Género. La interpretación de γ_0 es debido a la necesidad de factores únicos pues de caso contrario el modelo estará sobre-parametrizado. Para evitar lo anterior fijamos a γ_0 como la celda de tarifa base del modelo multiplicativo y re-escalamos a los demás niveles de riesgo en función de ésta, típicamente se selecciona la celda de tarifa con mayor exposición. Por lo anterior, los $\gamma_{1,i}$ y $\gamma_{2,j}$ que conformen al nivel base serán iguales a 1.

De manera general, el modelo multiplicativo luce de la siguiente forma:

$$\mu_{i_1, i_2, \dots, i_M} = \gamma_0 \gamma_{1, i_1} \gamma_{2, i_2} \cdots \gamma_{M, i_M}, \quad (1.1)$$

Donde γ_{j, i_j} corresponde al efecto del nivel i_j de la j -ésima variable, con $j = 1, \dots, M$.

1.4 Modelos de Dispersión Exponencial

Como mencionamos al principio de este capítulo, los Modelos de Regresión Lineal simple o múltiple no son una vía adecuada para la estimación de la prima de riesgo porque asumen una distribución Normal para los errores. Los Modelos Lineales Generalizados dan la oportunidad de asignar cualquier distribución de la familia exponencial a los errores, i.e., habilitan la posibilidad de usar Modelos de Dispersión Exponencial. En este sentido los Modelos de Regresión Lineal simple o múltiple son un caso particular de los Modelos Lineales Generalizados.

Sea $y^T = (y_1, y_2, \dots, y_n)$ el vector con los resultados de las n distintas y mutuamente excluyentes celdas de tarifa, es decir, y_i será el número o monto de siniestros para cada celda de tarifa. Además cada registro i del vector contiene el peso de exposición o número de siniestros ω_i , así como el detalle de los niveles de los factores de riesgo que componen a cada celda de tarifa. Siguiendo el ejemplo anterior, nuestro vector luciría de la siguiente forma:

Tabla 1.3

i	Género	Tipo de Vehículo	Observación	Peso
1	Hombre	Pequeño	y_1	ω_1
2	Hombre	Grande	y_2	ω_2
3	Mujer	Pequeño	y_3	ω_3
4	Mujer	Grande	y_4	ω_4

Si Y_i $i = 1, \dots, 4$ es la variable aleatoria para la celda de tarifa i entonces sucederán dos cosas:

1. Por los supuestos de la sección 1.1 las Y_i 's serán independientes e idénticamente distribuidas.
2. La distribución de las Y_i 's tendrá la siguiente forma:

$$f_{Y_i}(y_i; \theta_i, \varphi) = \exp \left\{ \frac{y_i \theta_i - b(\theta_i)}{\varphi / \omega_i} + c(y_i, \varphi, \omega_i) \right\}, \quad (1.2)$$

Donde:

θ_i : Es un parámetro que depende de la celda de tarifa, generalmente asociado con la media que está en una vecindad abierta alrededor del 0.

φ : Es el parámetro de dispersión, constante para toda las i 's y es mayor a cero.

$\omega_i > 0$ y $b(\theta_i)$ Es la función cumulante, misma de la que hablaremos más adelante. Es importante mencionar que la relación entre la función cumulante y la distribución de probabilidad de cualquier familia, dados los parámetros θ_i y φ , es única, es decir, dada la elección de $b(\theta_i)$ la distribución quedará completamente especificada y sus dos primeros momentos definidos, por tanto la función $c(y_i, \varphi, \omega_i)$ no es de interés.

1.4.1 Función Generadora y Función Cumulante

Una buena parte del análisis de la tarifa es el estudio del ajuste de la estructura de error. Podemos abordar este tema desde la perspectiva de los momentos de la estructura de error.

Debido al supuesto de que la Y_i 's serán independientes e idénticamente distribuidas, en lo sucesivo y, para simplificar la notación, nos ahorraremos el uso de índices.

Definición 1.1 Para cada entero n , se dice que el n -ésimo momento de Y es $\mu^n = E[Y^n]$ y el n -ésimo momento central de Y es $\mu^n = E[(Y - \mu)^n]$ con $\mu = E[Y]$.

Es importante advertir al lector que el análisis de la tarifa, tema de este trabajo, hace sólo énfasis en el primero momento y el segundo momento central de Y .

Ahora se hablará brevemente de la función generadora de momentos que, como su nombre lo indica, puede ser usada para calcular los distintos momentos de una variable aleatoria; no obstante, en la práctica es más común el uso de ésta para la caracterización de la distribución de una variable aleatoria.

“Definición 1.2 Sea Y una variable aleatoria. La función generadora de momentos de Y denotada como $M_Y(t)$ será igual a la esperanza de e^{tY} , siempre que la esperanza exista para t en una vecindad alrededor de 0. Es decir, que si $h > 0$ para toda t tal que $-h < t < h$, entonces $M_Y(t) = E[e^{tY}]$ ”.

“Teorema 1.1 Sea $M_Y(t)$ la función generadora de momentos de Y , entonces $E[Y^n] = M_Y^{(n)}(0)$ donde $M_Y^{(n)}(0) = \frac{d^n}{dt^n} M_X(t) \Big|_{t=0}$ ”.

Para los modelos de dispersión exponencial continuos la función generadora de momentos queda dada por:

$$\begin{aligned}
 E[e^{tY}] &= \int e^{ty} f_Y(y; \theta, \varphi) dy = \int \exp \left\{ \frac{y\theta - b(\theta)}{\varphi/\omega} + ty + c(y, \varphi, \omega) \right\} dy \\
 &= \int \exp \left\{ \frac{y(\theta + t\varphi/\omega) - b(\theta)}{\varphi/\omega} + c(y, \varphi, \omega) \right\} dy \\
 &= \int \exp \left\{ \frac{y(\theta + t\varphi/\omega) + b(\theta + t\varphi/\omega) - b(\theta + t\varphi/\omega) - b(\theta)}{\varphi/\omega} + c(y, \varphi, \omega) \right\} dy \\
 &= \exp \left\{ \frac{b(\theta + t\varphi/\omega) - b(\theta)}{\varphi/\omega} \right\} \int \exp \left\{ \frac{y(\theta + t\varphi/\omega) - b(\theta + t\varphi/\omega)}{\varphi/\omega} \right. \\
 &\quad \left. + c(y, \varphi, \omega) \right\} dy, \quad (1.3)
 \end{aligned}$$

Dado el hecho de que t está en un vecindario abierto alrededor del cero y ahí $E[e^{tY}]$ converge, entonces $\theta + t\varphi/\omega$ está en el espacio de los parámetros y la última integral de la expresión 1.3 es igual a uno, ya que se integra sobre todo su dominio. Por lo tanto la función generadora de momentos existe para cualquier modelo de dispersión exponencial y está dada por:

$$E[e^{tY}] = \exp \left\{ \frac{b(\theta + t\varphi/\omega) - b(\theta)}{\varphi/\omega} \right\}.$$

Decimos que la función cumulante es el logaritmo de la función generadora de momentos y está denotada por:

$$\Psi(t) = \frac{b\left(\theta + \frac{t\varphi}{\omega}\right) - b(\theta)}{\varphi/\omega}, \quad (1.4)$$

La función cumulante puede usarse, al igual que la función generadora de momentos, para calcular los momentos, tal y como se refiere en el teorema 1.1.

$$\begin{aligned} \Psi'(t) &= \frac{b'\left(\theta + \frac{t\varphi}{\omega}\right)}{\varphi/\omega} \varphi/\omega = b'\left(\theta + \frac{t\varphi}{\omega}\right), \\ E[Y] &= \Psi'(0) = b'(\theta) = \mu, \\ \Psi''(t) &= b''\left(\theta + \frac{t\varphi}{\omega}\right) \varphi/\omega, \\ \text{Var}[Y] &= \Psi''(0) = b''(\theta) \varphi/\omega. \end{aligned}$$

Cuando hablamos del segundo momento central de una función, es decir la varianza, es más conveniente abordarla como una función de la media μ . En párrafos anteriores se muestra que $\mu = E[Y] = b'(\theta)$ y como al inicio se asume que la función b es además invertible podemos establecer la siguiente relación $\theta = b'^{-1}(\mu)$, además podemos sustituir en $b''(\theta)$ la anterior relación y así obtener la función de varianza $v(\mu) = b''(b'^{-1}(\mu))$. Dicho de otro modo podemos expresar a la varianza de la variable aleatoria Y como función de μ , siempre y cuando se asuma $\varphi/\omega = 1$.

¿Cuál es la importancia de hacer lo anterior? Bien, una familia de distribuciones de probabilidad está caracterizada de una manera única por su función de varianza. El siguiente cuadro muestra la función de varianza para algunos de los principales modelos de dispersión exponencial.

Tabla 1.4

Distribución	Normal	Poisson	Gamma	Binomial
$v(\mu)$	1	μ	μ^2	$\mu(1 - \mu)$

1.4.2 Modelo Poisson para la Estimación de la Frecuencia de los Siniestros

Típicamente, cuando se habla de la estructura de error para el recuento de siniestros se sugiere emplear una distribución Poisson o Binomial Negativa, en caso de existir sobredispersión. La elección de una distribución Poisson está plenamente justificada⁶ si denotamos a $K(t)$ como el

⁶ La Completa justificación de la elección de la distribución Poisson se encuentra en el anexo del libro [4] de la bibliografía

número de siniestros en un intervalo semi-cerrado $(0, t]$ con $t > 0$, $K(0) = 0$ y añadimos los siguientes tres supuestos:

- A. Las variables $K(t_2) - K(t_1)$ y $K(t_4) - K(t_3)$ son independientes.
- B. Las variables $K(s + t) - K(s)$ y $K(t)$ tienen la misma distribución.
- C. La probabilidad de que más de un siniestro ocurra al mismo tiempo o que un número infinito de siniestros ocurran en un intervalo finito de tiempo son, ambas, cero.

Con $s > 0$ y $0 \leq t_1 < t_2 \leq t_3 < t_4$.

El supuesto A es una consecuencia del supuesto dos de la sección 1.1 dado que (t_1, t_2) y (t_3, t_4) son intervalos de tiempo disjuntos. El supuesto B es, tanto como el supuesto de homogeneidad de la sección 1.1. Al supuesto C lo podemos dividir en dos; la probabilidad de que en un intervalo de tiempo finito ocurran una cantidad infinita de siniestros para la cobertura de Daños Materiales del Sector Mexicano es igual a cero; mientras que la probabilidad de que dos o más siniestros ocurran al mismo tiempo es tanto como decir que no hay independencia entre pólizas o una casualidad infinitesimal. Por tanto los supuestos A, B y C encajan bastante bien con nuestros supuestos anteriores.

Ahora bien, sea X la variable ⁷ respuesta, en otras palabras, el número de siniestros. Por lo anterior y, dado que $E[X] = \omega\mu$, sabemos que:

$$f_X(x; \mu) = e^{-\omega\mu} \frac{(\omega\mu)^x}{x!}, \quad x = 0, 1, 2, \dots$$

Pero la variable que tenemos como objeto de estudio es la frecuencia, misma que tiene como expresión algebraica $Y = X/\omega$, de tal modo que:

$$f_Y(y; \mu) = P[Y = y] = P\left[\frac{X}{\omega} = y\right] = P[X = \omega y] = e^{-\omega\mu} \frac{(\omega\mu)^{\omega y}}{(\omega y)!}, \quad (1.5)$$

Si $c(y, \varphi, \omega) = \omega y \log(\omega) - \log(\omega y!)$, $\varphi = 1$ y re-parametrizamos a $\theta = \log(\mu)$, que tiene como consecuencia $b(\theta) = e^\theta$, sucede:

$$\begin{aligned} f_Y(y; \theta, \varphi) &= \exp\left\{\frac{y\theta - b(\theta)}{\varphi/\omega} + c(y, \varphi, \omega)\right\} \\ &= \exp\left\{\frac{y \log(\mu) - e^{\log(\mu)}}{\frac{1}{\omega}} + \omega y \log(\omega) - \log(\omega y!)\right\} \\ &= \exp\{y \log(\mu) \omega\} * \exp\{-e^{\log(\mu)} \omega\} * \exp\{\omega y \log(\omega)\} * \exp\{-\log(\omega y!)\} \\ &= (\mu)^{y\omega} * e^{-\omega\mu} * (\omega)^{y\omega} \frac{1}{(\omega y)!} = e^{-\omega\mu} \frac{(\omega\mu)^{\omega y}}{(\omega y)!} = f_Y(y; \mu). \end{aligned}$$

⁷ Continuaremos con el ahorro de los subíndices

Por tanto la distribución Poisson es un Modelo de Dispersión Exponencial y en consecuencia puede usarse para modelar la frecuencia.

1.4.3 Modelo Gamma para la Estimación de la Severidad de los Siniestros

Refiriéndonos a la severidad de siniestros la variable respuesta es ahora el monto neto de siniestros y como variable ponderadora usaremos el número de siniestros, cabe señalar que el número de siniestros ha dejado de ser una variable aleatoria pues la severidad es el monto promedio de siniestros dado un número fijo de éstos.

Pedimos únicamente que la distribución de la estructura de error sea positiva y que sea sesgada a la derecha, de tal modo que la selección de la distribución no es del todo obvia. Asumiendo⁸ que para la distribución del monto de siniestros individuales –usaremos datos del Ramo de Automóviles de pólizas individuales- la varianza de los residuos incrementa de manera proporcional a la media, es decir, un coeficiente constante de variación. El anterior supuesto, sumado a las únicas dos restricciones que pedimos postula a la distribución Gamma como una buena candidata.

La distribución de una variable aleatoria Gamma es:

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}; \quad x > 0, \alpha > 0 \text{ y } \beta > 0.$$

Se resume a continuación algunos de los resultados más relevantes de esta distribución que se usarán en adelante. $E[X] = \alpha/\beta$, $Var[X] = \alpha/\beta^2$ y la suma de variables independientes Gamma con el mismo parámetro β (parámetro de escala) es también una Gamma de parámetro β y el segundo parámetro (parámetro de forma) será la suma de las distintas α 's.

De tal modo que si $Z \sim \text{Gamma}(\alpha, \beta)$ es la distribución individual del monto neto de siniestros, y definimos a $X = \sum_{k=1}^{\omega} Z_k$ como la suma de ω distintos siniestros independientes, pero pertenecientes a una misma celda de tarifa (idénticamente distribuidos) se puede decir que $X \sim \text{Gamma}(\omega\alpha, \beta)$, pues el monto neto esperado de la suma de los siniestros es la suma de los montos netos esperados de los siniestros. Si $Y = X/\omega$ (severidad) entonces:

$$f_Y(y) = \omega f_X(\omega y) = \omega \frac{(\beta)^{\omega\alpha}}{\Gamma(\omega\alpha)} (\omega y)^{\omega\alpha-1} e^{-\beta\omega y} = \frac{(\omega\beta)^{\omega\alpha}}{\Gamma(\omega\alpha)} y^{\omega\alpha-1} e^{-\omega\beta y}, \quad y > 0.$$

Por tanto $Y \sim \text{Gamma}(\omega\alpha, \omega\beta)$ y $E[Y] = \alpha/\beta$.

Ahora bien, sea $\mu = \alpha/\beta$ y $\varphi = 1/\alpha$ luego entonces $\mu > 0$ y $\varphi > 0$ y,

⁸ Baxter, Coutts & Ross en 1980 justificaron este hecho, sin embargo, al momento de ajustar el modelo es importante corroborar este supuesto.

$$f_Y(y) = \frac{1}{\Gamma\left(\frac{\omega}{\frac{1}{\alpha}}\right)} \left(\frac{\omega}{\frac{\alpha}{\beta} * \frac{1}{\alpha}}\right)^{\frac{\omega}{1/\alpha}} y^{\frac{\omega}{1/\alpha}-1} e^{-y\left(\frac{\omega}{\frac{\alpha}{\beta} * \frac{1}{\alpha}}\right)} = \frac{1}{\Gamma\left(\frac{\omega}{\varphi}\right)} \left(\frac{\omega}{\mu\varphi}\right)^{\frac{\omega}{\varphi}} y^{\frac{\omega}{\varphi}-1} e^{-y\left(\frac{\omega}{\mu\varphi}\right)}.$$

Si $c(y, \varphi, \omega) = \log(\omega y / \varphi) \omega / \varphi - \log(y) - \log \Gamma(\omega / \varphi)$ y re-parametrizamos a $\theta = -1/\mu$, que tiene como consecuencia $b(\theta) = -\log(-\theta)$, sucede:

$$\begin{aligned} f_Y(y; \theta, \varphi) &= \exp\left\{\frac{y\theta - b(\theta)}{\varphi/\omega} + c(y, \varphi, \omega)\right\} \\ &= \exp\left\{\frac{-\frac{y}{\mu} + \log\left(\frac{1}{\mu}\right)}{\varphi/\omega} + \log(\omega y / \varphi) \omega / \varphi - \log(y) - \log \Gamma(\omega / \varphi)\right\} \\ &= \exp\left\{-\frac{y\omega}{\mu\varphi}\right\} * \left(\frac{1}{\mu}\right)^{\frac{\omega}{\varphi}} * \left(\frac{\omega y}{\varphi}\right)^{\frac{\omega}{\varphi}} * \frac{1}{y} * \frac{1}{\Gamma(\omega / \varphi)} = \frac{1}{\Gamma\left(\frac{\omega}{\varphi}\right)} \left(\frac{\omega}{\mu\varphi}\right)^{\frac{\omega}{\varphi}} y^{\frac{\omega}{\varphi}-1} e^{-y\left(\frac{\omega}{\mu\varphi}\right)} \\ &= f_Y(y). \end{aligned}$$

Por tanto la distribución Gamma es un Modelo de Dispersión Exponencial y en consecuencia puede usarse para modelar la Severidad.

1.5 Función Liga

Hasta ahora sólo se ha visto la manera de generalizar los resultados de una regresión lineal con respecto a la estructura de error, en otras palabras, se extendió el catálogo de estructuras de error. Aún queda tocar el modo en que los Modelos Lineales Generalizados evitan que la media sea una función lineal de las variables exógenas.

La mejor manera de aterrizar un tema es mediante un ejemplo, por tanto se retomará el ejemplo de la tabla 1.2 y se expresará a la media como función aditiva de los factores de riesgo.

$$\mu_{i,j} = \gamma_0 + \gamma_{1,i} + \gamma_{2,j}.$$

Al igual que en la sección 1.3 se requiere de una restricción para evitar un modelo sobreparametrizado. En este caso forzaremos a que los parámetros del nivel base sean iguales a cero. Digamos pues que la celda (1,1) es la celda base (Género: Hombre y Tipo de Vehículo: Pequeño) por tanto, $\gamma_{1,1} = \gamma_{2,1} = 0$ y $\mu_{1,1} = \gamma_0$. Sea $x_{i,j}$ una variable indicadora tal que:

$$x_{i,j} = \begin{cases} 1, & \text{si } \beta_j \text{ pertenece a } \mu_{i,j} \\ 0, & \text{cualquier otro caso.} \end{cases}$$

$$\text{Y } \beta_1 = \gamma_0, \beta_2 = \gamma_{1,2} \text{ y } \beta_3 = \gamma_{2,2}.$$

Con estas variables se puede re-exresar a los distintos valores esperados de las celdas de tarifa como:

$$\mu_i = \sum_{j=1}^3 \beta_j x_{i,j}, \quad \text{con } i = 1, \dots, 4. \quad (1.6)$$

O bien, de un modo matricial $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$.

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_4 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,3} \\ \vdots & \ddots & \vdots \\ x_{4,1} & \cdots & x_{4,3} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_3 \end{pmatrix}.$$

Donde la primera columna de \mathbf{X} se reserva para la celda base, y las dos restantes serán una para cada variable binaria, es decir, las $x_{i,j}$'s son variables dummies.

1.5.1 Función Liga Logarítmica

Para ligar ambos modelos basta aplicar logaritmo de ambos lados del modelo multiplicativo, como sigue:

$$\log(\mu_{i,j}) = \log(\gamma_0 \gamma_{1,i} \gamma_{2,j}) = \log(\gamma_0) + \log(\gamma_{1,i}) + \log(\gamma_{2,j}).$$

Nuevamente se debe seleccionar una celda base, de igual manera que antes digamos que la celda (1,1) lo es, así, $\gamma_{1,1} = \gamma_{2,1} = 1$ además, al añadir variables indicadoras tal como lo hicimos previamente e igualando como sigue $\beta_1 = \log(\gamma_0)$, $\beta_2 = \log(\gamma_{1,2})$ y $\beta_3 = \log(\gamma_{2,2})$ obtenemos:

$$\log(\mu_i) = \sum_{j=1}^3 \beta_j x_{i,j}, \quad i = 1, \dots, 4.$$

Que es la misma estructura lineal que en la ecuación 1.6.

Los Modelos Lineales Generalizados permiten que del lado izquierdo de la ecuación se escriba cualquier función monótona y diferenciable $g(\mu)$. El resultado general queda de la siguiente forma:

$$g(\mu_i) = \eta_i = \sum_{j=1}^r \beta_j x_{i,j}, \quad i = 1, \dots, n.$$

Por tanto la función liga logarítmica, por ser además monótona y diferenciable, es adecuada cuando se usa un modelo multiplicativo.

1.5.2 Función Liga Logit

Otra función liga de interés es la conocida bajo el nombre de Logit, definida como:

$$\eta_i = g(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right).$$

Esta función liga garantiza que μ_i estará entre cero y uno, tal y como se requiere en un modelo donde μ_i es una proporción. Dichos modelos son llamados de regresión logística.

1.6 Estimación de Parámetros

La palabra estimación es inherente a los modelos estadísticos, pues en todos los casos se cuenta con un conjunto de datos que nos permitirán acotar los posibles resultados de un evento aleatorio, tal y como es el caso de este trabajo.

1.6.1 Resultado General para el Método de Máximo Log-Verosimilitud

Regresando al ejemplo de la tabla 1.2 se podría estimar la frecuencia o severidad simplemente dividiendo el recuento o monto agregado de siniestros de la celda de tarifa de la variable respuesta entre el peso de dicha celda de tarifa, según corresponda. Sin embargo en la práctica no siempre es posible dar demasiada credibilidad a este método, pues la información en algunas celdas de tarifa puede llegar a ser tan escasa que estimar la frecuencia, severidad o prima de riesgo traería consigo errores. Es por tanto que la metodología para la estimación de los modelos lineales generalizados requiere de un método un tanto más exacto, por así llamarlo. El método a emplear es ampliamente conocido en el argot estadístico y recibe como nombre “Máxima Log-Verosimilitud”, que se describirá en breve.

Supongamos que contamos con n observaciones y cada una de éstas posee una distribución perteneciente a los modelos de dispersión exponencial, sumado a esto el supuesto de independencia entre pólizas podemos derivar la log-verosimilitud como función de θ , tal y como sigue:

$$\begin{aligned}
\ell(\theta; \varphi, y) &= \log \left[\prod_{i=1}^n f_{Y_i}(y_i; \theta_i, \varphi) \right] = \log \left[\prod_{i=1}^n \exp \left\{ \frac{y_i \theta_i - b(\theta_i)}{\varphi / \omega_i} + c(y_i, \varphi, \omega_i) \right\} \right] \\
&= \log \left[\exp \left\{ \sum_{i=1}^n \frac{y_i \theta_i - b(\theta_i)}{\varphi / \omega_i} + \sum_{i=1}^n c(y_i, \varphi, \omega_i) \right\} \right] \\
&= \sum_{i=1}^n \frac{y_i \theta_i - b(\theta_i)}{\varphi / \omega_i} + \sum_{i=1}^n c(y_i, \varphi, \omega_i) = \frac{1}{\varphi} \sum_{i=1}^n (y_i \theta_i - b(\theta_i)) \omega_i + \sum_{i=1}^n c(y_i, \varphi, \omega_i).
\end{aligned}$$

Como se está maximizando θ no afectará el remover el parámetro de dispersión φ . Es preferible que la función de log-verosimilitud dependa de las β en lugar de θ y esto se logra mediante dos hechos $\mu_i = b'(\theta_i)$ y $g(\mu_i) = \eta_i = \sum_{j=1}^r \beta_j x_{i,j} = g(b'(\theta_i))$, así que, según la regla de la cadena:

$$\frac{\delta \ell}{\delta \beta_j} = \sum_{i=1}^n \frac{\delta \ell}{\delta \theta_i} \frac{\delta \theta_i}{\delta \beta_j} = \frac{1}{\varphi} \sum_{i=1}^n (\omega_i y_i - \omega_i b'(\theta_i)) \frac{\delta \theta_i}{\delta \beta_j}, \quad j = 1, \dots, r.$$

La segunda parte de la ecuación parece no ser trivial $\theta_i = b'^{-1} \left(g^{-1}(\sum_{j=1}^r \beta_j x_{i,j}) \right)$, nuevamente, por la regla de la cadena:

$$\frac{\delta \theta_i}{\delta \beta_j} = \frac{\delta \theta_i}{\delta \mu_i} \frac{\delta \mu_i}{\delta \eta_i} \frac{\delta \eta_i}{\delta \beta_j}.$$

Donde:

$$\begin{aligned}
\frac{\delta \theta_i}{\delta \mu_i} &= \left(\frac{\delta \mu_i}{\delta \theta_i} \right)^{-1} = \frac{1}{b''(\theta_i)} = \frac{1}{v(\mu_i)'}, \\
\frac{\delta \mu_i}{\delta \eta_i} &= \left(\frac{\delta \eta_i}{\delta \mu_i} \right)^{-1} = \frac{1}{g'(\mu_i)'}, \\
\frac{\delta \eta_i}{\delta \beta_j} &= \frac{\sum_{j=1}^r \beta_j x_{i,j}}{\delta \beta_j} = x_{i,j};
\end{aligned}$$

De tal modo que:

$$\frac{\delta \ell}{\delta \beta_j} = \frac{1}{\varphi} \sum_{i=1}^n (\omega_i y_i - \omega_i b'(\theta_i)) \frac{1}{v(\mu_i)} \frac{x_{i,j}}{g'(\mu_i)} = \frac{1}{\varphi} \sum_{i=1}^n \omega_i \frac{y_i - \mu_i}{v(\mu_i) g'(\mu_i)} x_{i,j}, \quad j = 1, \dots, r.$$

Igualando a cero para las r derivadas parciales, correspondientes a cada uno de los r parámetros del modelo se obtiene:

$$\sum_{i=1}^n \omega_i \frac{y_i - \mu_i}{v(\mu_i) g'(\mu_i)} x_{i,j} = 0, \quad j = 1, \dots, r.$$

Podría parecer a simple vista, como se debe resolver para las distintas μ_i 's, que la solución trivial es $\mu_i = y_i$, sin embargo μ_i es función de las β_j 's y tiene que satisfacer todavía una ecuación más:

$$\mu_i = g^{-1} \left(\sum_{j=1}^r \beta_j x_{i,j} \right).$$

La solución trivial sólo podría tener sentido si y sólo si el número de parámetros a estimar es igual al número de observaciones, en donde claramente $\mu_i = y_i$, modelo saturado.

1.6.2 Modelo Multiplicativo Poisson para la frecuencia

A partir de los resultados anteriores se puede construir de manera sencilla la expresión para el caso particular de una estructura de error Poisson y una función Liga Logarítmica.

Dado que $v(\mu_i) = \mu_i$, $g'(\mu_i) = \frac{1}{\mu_i}$ al tratarse de un modelo logarítmico en combinación con una distribución Poisson.

$$\begin{aligned} \sum_{i=1}^n \omega_i \frac{y_i - \mu_i}{v(\mu_i)g'(\mu_i)} x_{i,j} &= \sum_{i=1}^n \omega_i \frac{y_i - \mu_i}{\mu_i \frac{1}{\mu_i}} x_{i,j} = \sum_{i=1}^n x_{i,j} \omega_i (y_i - \mu_i) = 0 \\ \Leftrightarrow \sum_{i=1}^n x_{i,j} \omega_i y_i &= \sum_{i=1}^n x_{i,j} \omega_i \mu_i, \quad j = 1, \dots, r. \end{aligned}$$

Además:

$$\mu_i = \exp \left(\sum_{j=1}^r \beta_j x_{i,j} \right).$$

1.6.3 Modelo Multiplicativo Gamma para la Severidad

Finalmente queda expresar la ecuación de máximo log-verosimilitud de la función Gamma combinado con la función liga logarítmica.

Dado que $v(\mu_i) = \mu_i^2$, $g'(\mu_i) = \frac{1}{\mu_i}$ al tratarse de una Gamma.

$$\begin{aligned} \sum_{i=1}^n \omega_i \frac{y_i - \mu_i}{v(\mu_i)g'(\mu_i)} x_{i,j} &= \sum_{i=1}^n \omega_i \frac{y_i - \mu_i}{\mu_i^2 \frac{1}{\mu_i}} x_{i,j} = \sum_{i=1}^n \frac{x_{i,j} \omega_i (y_i - \mu_i)}{\mu_i} = 0 \Leftrightarrow \sum_{i=1}^n \frac{x_{i,j} \omega_i y_i}{\mu_i} \\ &= \sum_{i=1}^n x_{i,j} \omega_i, \quad j = 1, \dots, r. \end{aligned}$$

Y nuevamente:

$$\mu_i = \exp\left(\sum_{j=1}^r \beta_j x_{i,j}\right).$$

1.7 Análisis de la Devianza del Modelo

La bondad de ajuste es un punto bastante importante tras cualquier estimación. Un modo de evaluar el ajuste de un modelo es comparándolo con el mejor modelo posible. Éste se obtiene cuando el número de observaciones y de parámetros es el mismo: el cual es conocido como Modelo Saturado y cumple con la condición $y_i = \hat{\mu}_i$. Ahora bien, se define a $\ell(\hat{\boldsymbol{\mu}})$ como la log-verosimilitud en función del vector de medias ajustadas $\hat{\boldsymbol{\mu}}$ a un conjunto de datos.

Luego entonces la devianza escalada D^* quedará definida como:

$$D^* = D^*(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2[\ell(\mathbf{y}) - \ell(\hat{\boldsymbol{\mu}})].$$

Donde $\ell(\mathbf{y})$ es la log-verosimilitud en función del vector de medias ajustadas al modelo saturado.

Algunos puntos importantes a mencionar son:

- Cuando el modelo posee un buen ajuste se espera que $\ell(\hat{\boldsymbol{\mu}})$ sea cercana pero nunca más grande que $\ell(\mathbf{y})$.
- La distribución de D^* será aproximadamente una χ_{n-r}^2 donde n es el número de observaciones y r el número de parámetros independientes.
- El valor esperado de D^* será $n - r$ y comúnmente un indicador a examinar es $\frac{D^*}{n-r}$, un valor mucho mayor a uno indicará un mal ajuste, pues el cociente es una medida de sobredispersión.

Siendo así es bastante aceptable que tanto $\ell(\mathbf{y})$ como $\ell(\hat{\boldsymbol{\mu}})$ tengan el mismo parámetro de dispersión φ . Si se denota como h a la inversa de b' entonces sucederá que $\theta_i = h(\mu_i)$ y:

$$\begin{aligned} D^* &= 2 \left[\frac{1}{\varphi} \sum_{i=1}^n (y_i h(y_i) - b(h(y_i))) \omega_i + \sum_{i=1}^n c(y_i, \varphi, \omega_i) \right. \\ &\quad \left. - \frac{1}{\varphi} \sum_{i=1}^n (y_i h(\hat{\mu}_i) - b(h(\hat{\mu}_i))) \omega_i - \sum_{i=1}^n c(y_i, \varphi, \omega_i) \right] \\ &= \frac{2}{\varphi} \sum_{i=1}^n \omega_i (y_i h(y_i) - b(h(y_i)) - y_i h(\hat{\mu}_i) + b(h(\hat{\mu}_i))). \end{aligned}$$

Multiplicado a D^* por el parámetro de escala φ se obtiene la devianza, también conocida como la devianza des-escalada.

$$D = 2 \sum_i \omega_i (y_i h(y_i) - b(h(y_i)) - y_i h(\hat{\mu}_i) + b(h(\hat{\mu}_i))).$$

El siguiente cuadro resume las distintas expresiones de devianza para algunos de los principales modelos de dispersión exponencial.

Tabla 1.5

Distribución	Devianza
Normal	$\sum_{i=1}^n \omega_i (y_i - \hat{\mu}_i)^2$
Poisson	$2 \sum_{i=1}^n \omega_i \left(y_i \log \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right)$
Gamma	$2 \sum_{i=1}^n \omega_i \left(-\log \left(\frac{y_i}{\hat{\mu}_i} \right) + \frac{(y_i - \hat{\mu}_i)}{\hat{\mu}_i} \right)$
Binomial	$2 \sum_{i=1}^n \omega_i n_i \left(y_i \log \left(\frac{y_i}{\hat{\mu}_i} \right) + (n_i - y_i) \log \left(\frac{n_i - y_i}{1 - \hat{\mu}_i} \right) \right)$

1.8 Estimación del Parámetro de Dispersión

La importancia del parámetro de dispersión viene dada como sigue: si Y_i es la frecuencia o severidad, luego entonces $Var[Y_i] = \varphi v(\mu_i)/\omega_i$. Lo anterior es tanto como decir que la varianza de Y_i es proporcional al valor de φ , también conocido como parámetro de dispersión.

Además, dado que las β 's poseen aproximadamente una distribución normal tal que la precisión de las $\hat{\beta}$'s está afectada por φ , será importante estimar el valor correcto del parámetro de dispersión para obtener buenos intervalos de confianza para las $\hat{\beta}$'s.

1.8.1 Estimación del Parámetro de Dispersión por el Método de Pearson

Comencemos por definir a X^2 como:

$$X^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{Var[Y_i]} = \frac{1}{\varphi} \sum_{i=1}^n \omega_i \frac{(y_i - \hat{\mu}_i)^2}{v(\hat{\mu}_i)}.$$

X^2 Conocida como la Ji-cuadrada de Pearson posee aproximadamente una distribución χ_{n-r}^2 donde r es el número de parámetros independiente y n es el número total de observaciones, por tanto la $E[X^2] = n - r$. Si definimos a $\hat{\varphi}_X$ como:

$$\hat{\varphi}_X = \frac{\varphi X^2}{n - r} = \frac{1}{n - r} \sum_{i=1}^n \omega_i \frac{(y_i - \hat{\mu}_i)^2}{v(\hat{\mu}_i)};$$

Por otro lado:

$$E[\hat{\varphi}_X] = E\left[\frac{\varphi X^2}{n-r}\right] = \frac{\varphi}{n-r} E[X^2] = \varphi.$$

Por tanto $\hat{\varphi}_X$ es un estimador insesgado de φ .

1.8.2 Estimación del Parámetro de Dispersión por el Método de la Devianza

Otra manera de estimar el parámetro de dispersión es mediante la devianza escalada $D^* = D/\varphi$ que igual posee aproximadamente una distribución χ_{n-r}^2 . Si definimos a $\hat{\varphi}_D$ como:

$$\hat{\varphi}_D = \frac{2 \sum_i \omega_i (y_i h(y_i) - b(h(y_i)) - y_i h(\hat{\mu}_i) + b(h(\hat{\mu}_i)))}{n-r};$$

Por otro lado:

$$E[\hat{\varphi}_D] = E\left[\frac{D}{n-r}\right] = \frac{\varphi}{n-r} E[D^*] = \varphi.$$

Por tanto $\hat{\varphi}_D$ también es un estimador insesgado de φ .

1.9 Prueba de Hipótesis para Modelos Jerárquicos

Partiendo del hecho de querer estimar a μ en función de las x_j 's, es conveniente preguntarse si un grupo de x_j 's son relevantes para el modelo o no. Este cuestionamiento es equivalente a preguntarse si las β_j 's correspondientes son o no iguales a cero. Para probar esta clase de hipótesis utilizaremos la prueba de la razón de verosimilitud (LRT por sus siglas en inglés) que se define como sigue:

“Consideremos dos modelos H_r y H_s , tal que $H_s \subset H_r$. Sea $\hat{\mu}^r$ bajo MLE's de H_r y $\hat{\mu}^s$ para H_s del mismo modo. Entonces el estadístico LRT para probar H_s contra H_r es $D^(\mathbf{y}, \hat{\mu}^s) - D^*(\mathbf{y}, \hat{\mu}^r)$ ”.*

De este modo la hipótesis nula $H_0: D^*(\mathbf{y}, \hat{\mu}^s) - D^*(\mathbf{y}, \hat{\mu}^r) = \mathbf{0}$ indicará que el modelo H_r pese a poseer una mayor cantidad de parámetros independientes no es estadísticamente hablando un mejor modelo, mientras que la hipótesis alterna $H_1: D^*(\mathbf{y}, \hat{\mu}^s) - D^*(\mathbf{y}, \hat{\mu}^r) > \mathbf{0}$, señalará lo contrario.

Ahora bien, habiendo ubicado las dos hipótesis complementarias podemos pasar a la cuestión de cómo se distribuye el estadístico LRT, éste se distribuye aproximadamente como una χ^2_{r-s} , donde $r - s$ es el número de parámetros independientes que se añaden al pasar del modelo H_s al modelo H_r , siempre y cuando suceda que ambos modelos posean la misma estructura de error y el mismo parámetro de dispersión.

1.10 Intervalos de Confianza de Fisher

Dada la naturaleza aleatoria del problema que estamos abordando será necesario contar con una herramienta que nos permita decidir si dos o más niveles (relatividades) dentro de una misma variable categórica deberían de ser agrupados, saber cuáles son las fluctuaciones aleatorias que podríamos llegar a considerar como normales antes de implementar algún ajuste de tarifa. Dicha herramienta son los intervalos de confianza.

1.10.1 Matriz de Información de Fisher

La matriz de información de Fisher es una medida de la cantidad de información que proveen las observaciones de una variable aleatoria para la estimación de los parámetros. Se define como la varianza de la primera derivada de la función de log-verosimilitud, así, valores grandes indicarán que la información provista por las observaciones no es la mejor para la estimación de los parámetros. Sin embargo, para simplificar su cálculo usaremos una expresión equivalente definida como el negativo de la esperanza de la segunda derivada de la función de log-verosimilitud. Comencemos por encontrar la expresión para la segunda derivada de la función de log-verosimilitud:

$$\begin{aligned} \frac{\delta^2 \ell}{\delta \beta_j \delta \beta_k} &= \sum_{i=1}^n \frac{\omega_i}{\varphi} \frac{\delta}{\delta \mu_i} \left[\frac{y_i - \mu_i}{v(\mu_i) g'(\mu_i)} \right] x_{i,j} \frac{\delta \mu_i}{\delta \eta_i} \frac{\delta \eta_i}{\delta \beta_k} = \sum_{i=1}^n \frac{\omega_i}{\varphi} \frac{\delta}{\delta \mu_i} \left[\frac{y_i - \mu_i}{v(\mu_i) g'(\mu_i)} \right] x_{i,j} \frac{1}{g'(\mu_i)} x_{i,k} \\ &= \sum_{i=1}^n x_{i,j} \frac{\omega_i}{\varphi g'(\mu_i)} \left[\frac{-1}{v(\mu_i) g'(\mu_i)} \right. \\ &\quad \left. - \frac{(y_i - \mu_i) v'(\mu_i) g'(\mu_i) + (y_i - \mu_i) v(\mu_i) g''(\mu_i)}{v(\mu_i)^2 g'(\mu_i)^2} \right] x_{i,k} \\ &= \sum_{i=1}^n x_{i,j} \frac{-\omega_i}{\varphi v(\mu_i) g'(\mu_i)^2} \left(1 + (y_i - \mu_i) \frac{[v(\mu_i) g''(\mu_i) + v'(\mu_i) g'(\mu_i)]}{v(\mu_i) g'(\mu_i)} \right) x_{i,k}; \end{aligned}$$

Sí

$$a_i = \frac{\omega_i}{\varphi v(\mu_i) g'(\mu_i)^2} \left(1 + (y_i - \mu_i) \frac{[v(\mu_i) g''(\mu_i) + v'(\mu_i) g'(\mu_i)]}{v(\mu_i) g'(\mu_i)} \right);$$

Entonces:

$$\frac{\delta^2 \ell}{\delta \beta_j \delta \beta_k} = - \sum_{i=1}^n x_{i,j} a_i x_{i,k}.$$

Ahora bien, si denotamos como \mathbf{H} a la matriz cuadrada que tiene en la entrada (i, k) a $\frac{\delta^2 \ell}{\delta \beta_j \delta \beta_k}$ podríamos descomponerla como $\mathbf{H} = \mathbf{X}' \mathbf{A} \mathbf{X}$. Donde \mathbf{X} es la matriz de diseño, \mathbf{A} es una matriz con a_i en la diagonal y cero en cualquier otra entrada y \mathbf{X}' es la matriz transpuesta de la matriz diseño.

Un ejemplo ayudará al lector a visualizar este hecho de mejor manera, usemos por simplicidad un a matriz de 2x2:

$$\begin{aligned} \mathbf{H} = \mathbf{X}' \mathbf{A} \mathbf{X} &= \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \end{pmatrix} \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} a_1 & x_{21} a_2 \\ x_{12} a_1 & x_{22} a_2 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \\ &= \begin{pmatrix} x_{11} a_1 x_{11} + x_{21} a_2 x_{21} & x_{11} a_1 x_{12} + x_{21} a_2 x_{22} \\ x_{12} a_1 x_{11} + x_{22} a_2 x_{21} & x_{12} a_1 x_{12} + x_{22} a_2 x_{22} \end{pmatrix}. \end{aligned}$$

No obstante, como mencionamos al inicio, la matriz de información de Fisher es menos la esperanza de la segunda derivada de la función de log-verosimilitud. Cabe señalar que Tanto $E[\mathbf{X}'] = \mathbf{X}'$ como $E[\mathbf{X}] = \mathbf{X}$ y dado que $E[Y_i] = \mu_i$ sucede que $E[a_i] = \frac{\omega_i}{\varphi v(\mu_i) g'(\mu_i)^2} = d_i$. Al igual que definimos a \mathbf{A} definiremos a \mathbf{D} cuyas entradas en la diagonal serán las d_i . Por tanto:

$$\mathbf{I} = -E[\mathbf{H}] = \mathbf{X}' E[\mathbf{A}] \mathbf{X} = \mathbf{X}' \mathbf{D} \mathbf{X}.$$

1.10.2 Intervalos de Confianza construidos con la información de Fisher

Bajo condiciones generales los estimadores de máximo verosimilitud se distribuyen asintótica e insesgadamente como una Normal.

$$\hat{\boldsymbol{\beta}} \approx^d N(\boldsymbol{\beta}, \mathbf{I}^{-1});$$

Donde \approx^d sirve para denotar que con una n lo suficientemente grande se distribuirá normalmente. La demostración formal está fuera de los alcances de este trabajo y es un resultado conocido. Cabe mencionar \mathbf{I}^{-1} que es la matriz de covarianzas de la distribución de $\hat{\boldsymbol{\beta}}$ que continúa siendo proporcional a φ .

Para derivar los intervalos de confianza bastará definir a $c_{j,k}$ como el elemento en el renglón j y la columna k de la matriz $\mathbf{C} = \mathbf{I}^{-1}$, de este modo el intervalo de confianza para β_j , bajo el nivel de confianza α quedará definido como:

$$(\hat{\beta}_j - z\sqrt{c_{j,j}}, \hat{\beta}_j + z\sqrt{c_{j,j}}) = (l, u);$$

Donde z es el percentil $1 - \alpha/2$ asociado a una $N(0,1)$.

Más aún, en el primer párrafo de la sección 1.10 se hizo énfasis en las relatividades que, como el lector recordará, tienen la forma $\hat{\gamma}_j = \exp(\hat{\beta}_j)$, por tanto el intervalo de confianza para las relatividades será $(\exp(l), \exp(u))$.

Finalmente queda el intervalo de confianza para $\hat{\mu}_i$ que dado el valor de un set de variables explicativas tiene la forma $g(\hat{\mu}) = \mathbf{X}\hat{\beta}$. Para generar el intervalo de confianza será necesario antes conocer la expresión para la varianza de $\mathbf{X}\hat{\beta}$ misma que se escribirá a continuación.

$$\text{Var}[\mathbf{X}\hat{\beta}] = \mathbf{X}\mathbf{I}^{-1}\mathbf{X}'.$$

Así el intervalo de confianza para $g(\hat{\mu})$ será:

$$g(\hat{\mu}) \pm z\sqrt{\mathbf{X}\mathbf{I}^{-1}\mathbf{X}'}$$

Donde z nuevamente es el percentil asociado a una $N(0,1)$. El intervalo de confianza es exacto cuando usamos a la función liga identidad y a la distribución Normal como estructura de error. En cualquier otro caso estaremos hablando de una aproximación que mejora conforme incrementa el número de pólizas incluidas en el set de datos. También cabe señalar que en el caso de una estructura de error Gamma, donde el parámetro de dispersión φ es estimado, existe un mayor error en la aproximación.

1.11 Análisis de Residuos

Tal y como en el modelo de regresión lineal simple, se puede expresar a la variable dependiente como:

$$\mathbf{y} = \hat{\mu} + (\mathbf{y} - \hat{\mu});$$

Donde $\hat{\mu}$ será el valor esperado y $(\mathbf{y} - \hat{\mu})$ es conocido como el residuo. Éstos pueden ser usados para medir que tan adecuada fue la elección de la función liga y de la estructura de error. También pueden llegar a indicar la presencia de valores atípicos. Sin embargo, el resultado de arriba sólo hace sentido en el caso de una estructura de error Normal con una función liga identidad, el resultado general puede tener tres distintas formas que expondremos a continuación.

1.11.1 Residuos de Pearson

Definimos al residuo de Pearson como:

$$r_{Pi} = \frac{y_i - \hat{\mu}_i}{\sqrt{v(\hat{\mu}_i)/\omega_i}}$$

Cabe señalar al lector el hecho de que $\sum_i r_{Pi}^2 = \varphi\mathbf{X}^2$.

1.11.2 Residuos de Anscombe

Una de las desventajas de los residuos de Pearson es que para el resto de las distribuciones que no son Normales la distribución de r_{pi} no será exacta y por tanto no tendrá propiedades similares a los residuales de una Normal. La idea es definir a los residuos usando una función $A(y_i)$ en lugar de y_i donde A se escoge para lograr que la distribución de $A(Y_i)$ sea lo más parecida a una Normal. La función A se define como:

$$A(\cdot) = \int \frac{d\mu}{v(\mu)^{\frac{1}{3}}}.$$

De tal modo que los residuos ahora tendrán la forma $A(y_i) - A(\hat{\mu}_i)$ sin embargo deben ser normalizados, es decir, debemos dividirlos entre la raíz cuadrada de la varianza de $A(Y_i)$, quedando como ecuación final:

$$r_{Ai} = \frac{A(y_i) - A(\hat{\mu}_i)}{A'(\hat{\mu}_i)\sqrt{v(\hat{\mu}_i)/\omega_i}}$$

1.11.3 Residuos de Devianza

Existe una última alternativa, que basa su cálculo en la devianza, tal y como la definimos en la sección 1.7. Dicha expresión tiene la forma:

$$r_{Di} = \sqrt{\omega_i 2(y_i h(y_i) - b(h(y_i)) - y_i h(\hat{\mu}_i) + b(h(\hat{\mu}_i)))} * \text{sign}(y_i - \hat{\mu}_i);$$

Donde como recordará el lector h a la inversa de b' y la función $\text{sign}(\cdot)$ regresa el signo, positivo o negativo de su argumento.

Tanto para la expresión de los residuos de la devianza como para los residuos de Pearson es necesario hacer una corrección dentro de la ecuación para estandarizar de manera correcta el resultado. Esta corrección se logra si se divide la expresión del residuo entre $\sqrt{\varphi(1 - h_i)}$ donde h_i es elemento de:

$$D^{\frac{1}{2}} X (X' D X)^{-1} X' D^{\frac{1}{2}}.$$

1.12 Sobre-dispersión

La sobredispersión es un tema que se relaciona con el modelo para el recuento de siniestros. En la sección 1.4.2 sugerimos el uso de una distribución Poisson para la modelación del número de siniestros, sin embargo a menudo sucede que la varianza de los datos en una misma celda de tarifa es mayor que la varianza de la distribución Poisson, a este fenómeno se le conoce con el nombre de sobre-dispersión.

Una manera de modelar la sobre-dispersión es usando un proceso Poisson compuesto. Con este modelo el número de siniestros X_i se distribuye Poisson con media Λ_i , pero Λ_i es una variable aleatoria que existe en el intervalo $(0, \infty)$. De este modo si $\Lambda_i = \lambda_i$ luego entonces X_i se distribuirá como una Poisson con media λ_i . Por tanto la distribución de X_i queda como:

$$f_{X_i}(x_i) = \int_0^{\infty} e^{-\lambda_i} \frac{(\lambda_i)^{x_i}}{x_i!} g(\lambda_i) d\lambda_i.$$

Donde $e^{-\lambda_i} \frac{(\lambda_i)^{x_i}}{x_i!}$ Es la densidad de $X_i | \Lambda_i = \lambda_i$ y $g(\lambda_i)$ es la densidad de Λ_i .

Este resultado empata con el resultado de la sección 1.4.2 pues $\lambda_i = \omega_i \mu_i$, ω_i es siempre conocida y la propiedad aleatoria quedará asociada intrínsecamente a μ_i .

Por las propiedades de la Poisson sucede que:

$$\begin{aligned} E[X_i | \Lambda_i] &= \Lambda_i, \\ Var[X_i | \Lambda_i] &= \Lambda_i, \\ E[X_i] &= E[E[X_i | \Lambda_i]] = E[\Lambda_i], \\ Var[X_i] &= E[Var[X_i | \Lambda_i]] + Var[E[X_i | \Lambda_i]] = E[\Lambda_i] + Var[\Lambda_i]. \end{aligned}$$

Ahora bien, una conveniente elección para Λ_i es una función de probabilidad Gamma, i.e., $\Lambda_i \sim \text{Gamma}(\mu_i, v)$ de tal modo que:

$$Var[\Lambda_i] = vE[\Lambda_i].$$

De tal modo que si $\lambda_i = \omega_i \mu_i$ ocurre entonces:

$$Var[X_i] = (1 + v)\omega_i \mu_i.$$

Sin embargo el análisis de la tarifa es llevado a cabo sobre la frecuencia $Y_i = X_i / \omega_i$, y $Var[Y_i] = Var[X_i] / \omega_i^2$ por tanto:

$$Var[Y_i] = \frac{(1 + v)\mu_i}{\omega_i};$$

Escribiendo a $\varphi = 1 + v > 1$ sucede:

$$Var[Y_i] = \frac{\varphi \mu_i}{\omega_i}.$$

Finalmente queda responder la pregunta. ¿Cuál sería entonces la distribución de X_i si $\Lambda_i \sim \text{Gamma}(\mu_i, v)$? La respuesta es:

$$P[X_i = x_i] = \frac{\Gamma\left(\frac{\omega_i \mu_i}{v} + x_i\right)}{\Gamma\left(\frac{\omega_i \mu_i}{v}\right) x_i!} \left(\frac{1}{1+v}\right)^{\frac{\omega_i \mu_i}{v}} \left(\frac{v}{1+v}\right)^{x_i};$$

Donde $X_i \sim BN\left(\frac{\omega_i \mu_i}{v}, \frac{v}{1+v}\right)$.

Por tanto, si el supuesto de homogeneidad no se cumple cabalmente se puede suponer una distribución Binomial Negativa. En otras palabras, si tenemos grupos de riesgo, caracterizado por una Poisson con una media altamente heterogénea entre los automóviles de pólizas individuales que constituyen a cada grupo, luego entonces, una estructura de error Binomial Negativa ayudará a disminuir la variación del modelo.

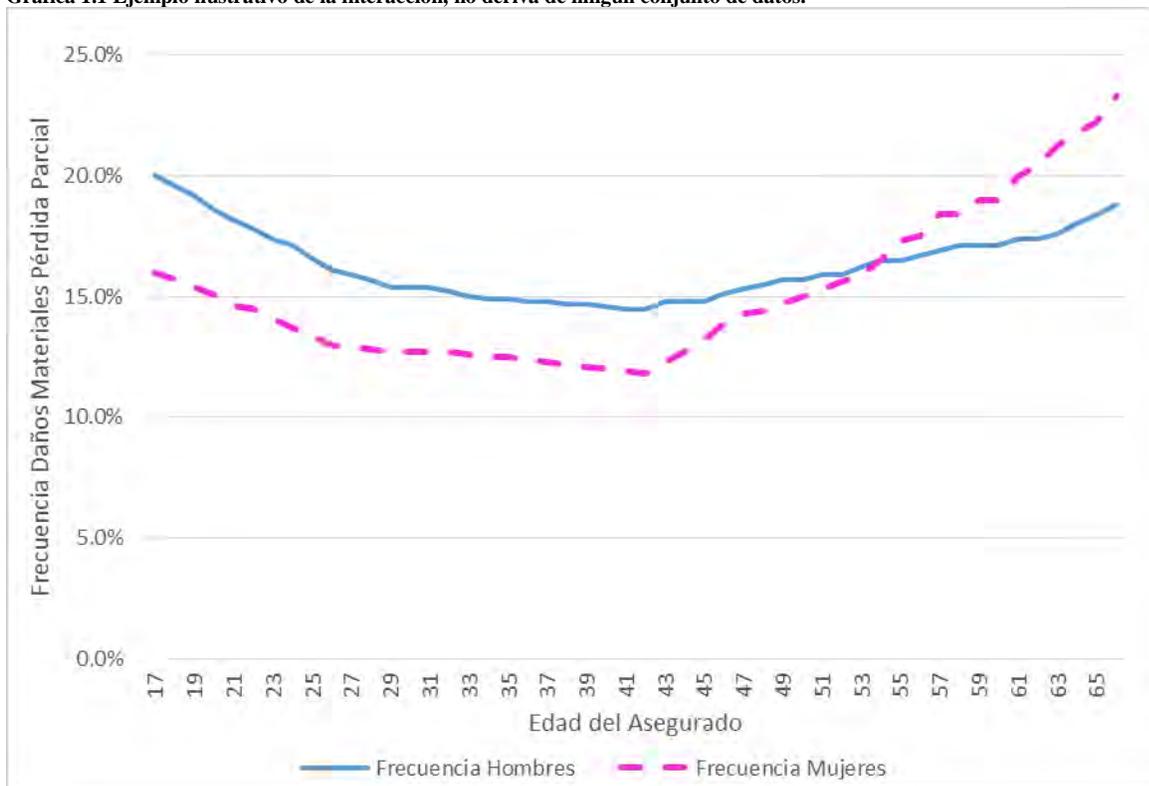
1.13 Consideraciones Especiales al Ajustar un Modelo

Existen en la práctica muchos problemas de los cuales no hemos hablado, hasta ahora. Esta sección pretende ilustrar casos comunes al ajuste de un Modelo Lineal Generalizado a los datos del Ramo de Automóviles del Sector Asegurador Mexicano.

1.13.1 Interacción Entre Variables

En el análisis de tarifa, en un comienzo, se supone que cada factor de riesgo incluido en el modelo posee un impacto no importando el valor de cualquier otro factor de riesgo. Sin embargo, variables como género y edad, en la práctica actuarial del Ramo de Automóviles de Póliza Individual del Sector Asegurador Mexicano, han demostrado que un modelo multiplicativo no es adecuado del todo. Sucede que los hombres jóvenes se ven involucrados con mayor frecuencia que las mujeres jóvenes en siniestros de Daños Materiales de Pérdidas Parciales, mientras que a edades alrededor de los 40 años, el comportamiento entre hombres y mujeres es muy similar, finalmente en las últimas edades la relación entre hombres y mujeres se invierte. Gráficamente ocurre lo siguiente:

Gráfica 1.1 Ejemplo ilustrativo de la interacción, no deriva de ningún conjunto de datos.



Para parametrizar el ejemplo y por simplicidad supongamos un modelo con dos variables: Género y edad. Además supongamos una función liga identidad. Siendo así, el modelo debería lucir de la siguiente forma:

$$y \approx \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2;$$

Donde bien x_1 puede representar la edad del conductor y x_2 , a manera de variable indicadora, representaría el género del conductor. Reacomodando la anterior expresión tendríamos algo de la forma:

$$y \approx \begin{cases} \beta_0 + \beta_1 x_1, & \text{si } x_2 = 0 \\ \beta_0 + \beta_2 + x_1(\beta_1 + \beta_3), & \text{si } x_2 = 1 \end{cases}$$

Esta clase de fenómenos deben de estudiarse y comprobarse estadísticamente. Sin embargo, implementar interacciones en un sistema de cálculo de primas de riesgo puede llegar a complicar la operación de una aseguradora además de ocasionar volatilidad en la estimación de los parámetros al incrementar el número de éstos, por tanto sólo deberemos incluirlas cuando sea estrictamente necesario.

1.13.2 Offset

Algunas veces las áreas técnicas o de suscripción fijan el valor de algún factor de riesgo. Por ejemplo: recargar o descontar en función de la forma de pago, es decir, pago anual o fraccionado; o una estrategia de venta para promover el uso de algún canal comercial. Como sea, los Modelos Lineales Generalizados pueden lidiar con esta clase de situaciones. Si se usa un modelo multiplicativo y las relatividades determinísticas están dadas por $z_i = \log(u_i)$, donde u_i es el valor fijo para algún factor de riesgo. Siendo así el modelo luciría de la siguiente forma:

$$\eta_i = z_i + \sum_{j=1}^r x_{i,j} \beta_j, \quad i = 1, 2, \dots, n.$$

CAPÍTULO 2

2.1 Características de la Cobertura

Esta tesis tiene como objetivo el cálculo de la prima de riesgo de la cobertura de Daños Materiales. Dicha cobertura pertenece al Ramo de Automóviles del Sector Asegurador Mexicano y tiene como objetivo el pago de la indemnización correspondiente a los daños del automóvil, pudiendo ser éstos pérdidas parciales o totales. Cabe señalar que La Ley de Instituciones de Seguros y Fianzas en su Artículo 27 Párrafo X menciona que para el ramo de automóviles, el pago de la indemnización que corresponda a los daños o pérdida del automóvil quedará comprendido dentro de las operaciones y ramos de su séptimo artículo.

2.1.1 Riesgos Amparados

Las pérdidas o daños materiales amparados generalmente incluyen los siguientes riesgos:

- Colisiones y Volcaduras.
- Rotura de cristales. Parabrisas, laterales, aletas, medallón y quemacocos adaptados originalmente por el fabricante.
- Incendio rayo y explosión.
- Ciclón, huracán, granizo, terremoto, erupción volcánica, alud, derrumbe de tierra o piedras, caída o derrumbe de construcciones, edificaciones, estructuras u otros objetos, caída de árboles o sus ramas, hundimiento de tierra, vendavales, maremoto e inundación a causa de fenómenos naturales.
- Daños en y durante su transportación: Ampara los riesgos de varadura, hundimiento, incendio, explosión, Colisión o Volcadura, descarrilamiento o caída del medio de transporte en que el Vehículo sea transportado; caída del Vehículo durante las maniobras de carga, trasbordo o descarga, así como la contribución por Avería gruesa o por cargos de Salvamento.
- El desbielamiento del Vehículo originado por la penetración de agua al interior del motor, distinta a la necesaria para su operación y funcionamiento y por causas ajenas a la voluntad del Asegurado o Conductor.

2.1.2 Deducibles

Se entenderá por deducible al monto, ya sea como porcentaje de la suma asegurada o cantidad fija, con el cuál invariablemente el asegurado participará en cada siniestro.

Por ejemplo, supongamos que se fija un deducible de \$3,000.00 MXN.

Caso 1:

El asegurado incurre en un siniestro cuyo monto asciende a \$1,500.00 MXN, sucederá que la aseguradora no compartirá los daños con el asegurado pues el monto del siniestro no supera el deducible.

Caso 2:

El asegurado incurre en un siniestro cuyo monto asciende a \$7,200.00 MXN, sucederá que la aseguradora compartirá los daños con el asegurado por un monto de \$4,200.00 MXN, esta cantidad es el resultado de restar al monto del siniestro el monto que se ha pactado como deducible.

2.2 Delimitación de la Información

Para la elaboración de esta tesis se tomaron datos reales del Sistema Estadístico del Sector Asegurador (SESA) de Daños de Automóviles de Póliza Individual del periodo 2014, según lo dispuesto en el Título 26 de la Circular Única de Seguros y Fianzas. Considerando el hecho de que la información proviene de todo el sector asegurador, bastará tomar un año. Además, las tendencias de años anteriores pudieran jugar en prejuicio del tercer supuesto, homogeneidad.

Estos datos inicialmente se dividen en tres archivos planos⁹: “Datos Generales”, “Emisión” y “Siniestros”.

El ajuste del modelo se llevará a cabo sobre los tipos de vehículos señalados como: “Automóviles” y “Comerciales”¹⁰, ambos representan un 95.4% de la exposición¹¹ total de la información disponible. Lo anterior es en virtud de lograr un mejor ajuste en los datos, pues tipos de vehículo demasiado distintos como Camiones, Semirremolques, Tracto-camiones, motocicletas o Autobuses podrían distorsionar el resultado final.

2.3 Ajuste Preliminar de la Estructura de Error y Tratamiento de Datos Atípicos

El ajuste preliminar de la estructura de error se lleva a cabo comúnmente sobre el Monto Neto de Siniestros¹², es decir, sobre el archivo plano “Siniestros”. Esto favorecerá un mejor ajuste acotando el monto de siniestros para que su comportamiento sea lo más parecido al de la estructura de error que proponemos para ajustar nuestro modelo.

De las secciones 2.3.1-2.3.4 se usará el hecho de que si $X \sim \text{LogNormal}$ y $Y = \log(X)$, entonces $Y \sim \text{Normal}$. Esta transformación ayudará a estudiar de manera preliminar el ajuste de los datos dada la similitud de la distribución *Gamma* con la distribución *LogNormal*, el lector podría llegar a preguntarse ¿Por qué no ajustar una distribución *LogNormal* en lugar de una

⁹ Las características enteras de estos tres archivos se detallan en el manual del SESA del Ramo de Automóviles de Póliza Individual (Versión 01) 03-dic-2015.

¹⁰ En la sección 2.2.2.2.2 Se especificará a detalle por qué no excluimos este tipo de vehículo.

¹¹ En la sección 2.2.1.2.1 se especifica de manera amplia este concepto.

¹² En la sección 2.4.1.1.2 se señala de manera puntual la forma aritmética y conceptual del monto neto de siniestros.

Gamma? Resulta que la distribución *LogNormal* no es un modelo de dispersión exponencial y por consiguiente no puede ser empleada como distribución en un modelo lineal generalizado. Ambas poseen parámetro de escala y forma, además un buen ajuste de la distribución *LogNormal* generalmente garantiza un buen ajuste de la distribución *Gamma* con los parámetros de forma y escala adecuados. Finalmente, el sentido de esta sección no es ajustar un modelo, sino detectar datos atípicos o tendencias que pudieran degenerar el ajuste del modelo.

No está de más mencionar que en la práctica global la información de los siniestros se prepara en un archivo plano que posee un registro por cada uno de éstos. Sin embargo, la información con la que contamos se encuentra agregada en función de las variables que analizaremos en la sección 2.4.2 pero esto no representa un problema dado que los modelos de dispersión exponencial son cerrados bajo operadores tales como la suma o promedio. Además debemos de partir del supuesto de que dentro de una misma celda de tarifa debe existir homogeneidad. La suma de las dos anteriores aseveraciones demuestra que la información que poseemos sigue siendo adecuada para el ajuste de un modelo lineal generalizado.

Dado que el archivo plano de siniestros con el que contamos separa por tipo de pérdida los siniestros y en virtud del tercer supuesto, analizaremos de manera marginal dicha variable. Los resultados del Ramo de automóviles de póliza individual se despliegan en la siguiente tabla.

Tabla 2.1 Monto promedio de cada tipo de pérdida.

Tipo de Pérdida	Número de Siniestros	Monto Neto de Siniestros	Severidad
<i>Pérdida Parcial</i>	650,185	7,008,954,809	10,780
<i>Pérdida Total</i>	43,073	3,209,053,810	74,503
<i>Rotura de cristales</i>	197,145	512,693,343	2,601

Tal y como esperábamos, el promedio del monto neto de siniestros deja ver que los tipos de pérdida poseen medias completamente distintas. Por tanto deberemos buscar el ajuste de un modelo de frecuencia y severidad para cada tipo de pérdida.

El manual del SESA del Ramo de Automóviles de Póliza Individual dictan que en el archivo plano “Siniestros” deberán reportarse también aquellos siniestros que aunque ocurrieron fuera del periodo de análisis tuvieron un movimiento dentro de él. Por tanto eliminaremos aquellos registros que posean monto distinto de cero y que tengan cero en el campo de número de siniestros.

2.3.1 Montos Negativos

Suponemos una función de densidad de probabilidad *Gamma* para la distribución de los montos de siniestros individuales y como dicha función sólo toma valores en el intervalo $[0, \infty)$, se deberán excluir del análisis todos aquellos registros que posean un monto neto de siniestros negativo, en otras palabras, aquellos registros en donde el monto de recuperaciones y salvamentos es mayor que el monto ocurrido de siniestros más los gastos de ajuste. Cabe señalar que cuando removemos el monto de un registro, dado que éste es negativo, es también necesario remover el número de siniestros correspondientes a este registro. Lo anterior es en virtud de mantener coherencia entre frecuencia y severidad.

Finalmente, el lector no debe preocuparse por una sobreestimación de la prima de riesgo puesto que cuando el modelo esté ajustado se mutualizarán todos los datos excluidos durante la fase del ajuste preliminar, es decir, repartiremos de manera equitativa aquellos siniestros que se han dejado fuera.

El siguiente cuadro muestra el resumen de la información por tipo de pérdida en función de este último cambio.

Tabla 2.2 Monto promedio de cada tipo de pérdida sólo para montos netos de siniestros positivos que ocurrieron durante el periodo de análisis.

<i>Tipo de Pérdida</i>	<i>Número de Siniestros</i>	<i>Monto Neto de Siniestros</i>	<i>Severidad</i>
<i>Pérdida Parcial</i>	544,166	7,769,761,105	14,278
<i>Pérdida Total</i>	42,391	3,193,937,299	75,345
<i>Rotura de cristales</i>	196,642	510,951,024	2,598

2.3.2 Ajuste por Pérdidas con Monto Pequeño

En esta sección se removerán los registros cargados al lado izquierdo del histograma, i.e., los montos muy pequeños. Se muestra a continuación un resumen gráfico y numérico de los datos antes y después de realizar cualquier recorte de información.

Finalmente, de ahora en adelante y hasta el término de la sección 2.3.4 se usará al monto promedio de siniestros en cada celda de tarifa como la variable sobre la que realizaremos los análisis. Lo anterior queda justificado por el párrafo tercero de la sección 2.3.

Cabe mencionar que tanto el resumen numérico de la media y varianza como el resumen gráfico serán desplegados en función de la transformación logarítmica de los datos, pues la idea se sostiene en comprobar si los datos transformados se ajustan a una distribución *Normal* o no. Sin embargo también se mostrará una tabla de cuantiles con la transformación inversa de los datos previamente transformados que permiten explorar las diferencias entre la distribución estimada a partir de los datos y los datos reales. Así el lector tendrá una visión más clara.

2.3.2.1 Ajuste de Pérdidas Parciales

Gráfica 2.1 Ajuste inicial de la transformación logarítmica del monto neto de siniestros para tipo de pérdida total.

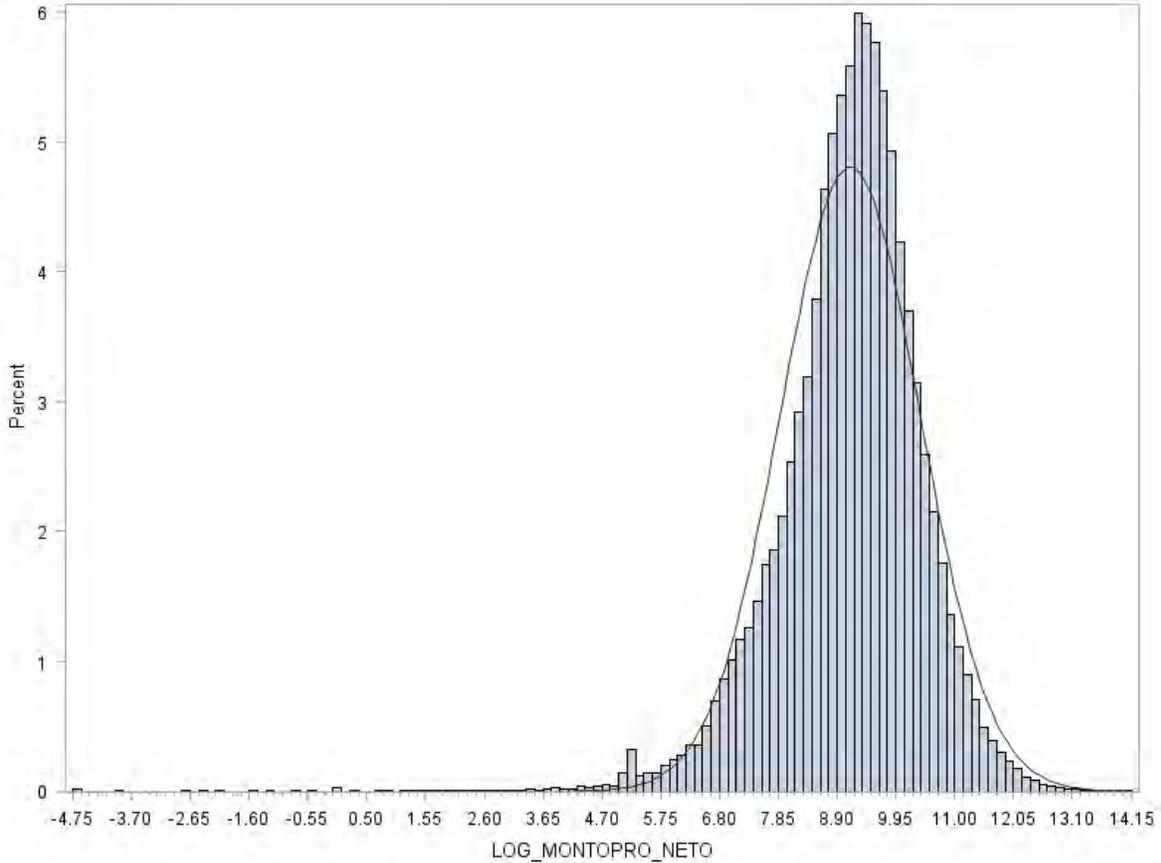


Tabla 2.3 Resumen del logaritmo del parámetro de dispersión inicial para tipo de Pérdida Parcial.

<i>Parámetro de Dispersión Estimado</i>	
<i>Media</i>	9.133
<i>Desviación Estándar</i>	1.246

Tabla 2.4 Resumen de los cuantiles observados contra los Normales estimados para tipo de pérdidas parciales antes de remover datos con montos pequeños.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	231.19	510.69
5%	1,080.10	1,193.43
10%	1,970.00	1,876.36
25%	4,871.56	3,996.60
50%	10,260.63	9,258.52
75%	20,056.13	21,448.07
90%	37,999.95	45,684.30
95%	57,618.56	71,826.59
99%	129,888.28	167,850.67

La diferencia máxima que se observa entre cuantiles es de 120.9% y el promedio de las diferencias es de 27.2%. En adición, el Test Kolmogorov-Smirnov da como resultado un estadístico $D= 0.057923$, tratándose de una muestra con 106,114 registros agregados y $\alpha = 5\%$ el p valor=0.0053.

A continuación se muestra el mismo resumen eliminando todas aquellas observaciones con un monto neto de siniestro menor a 500.

Gráfica 2.2 Ajuste final de la transformación logarítmica del monto neto de siniestros para tipos de pérdida parcial.

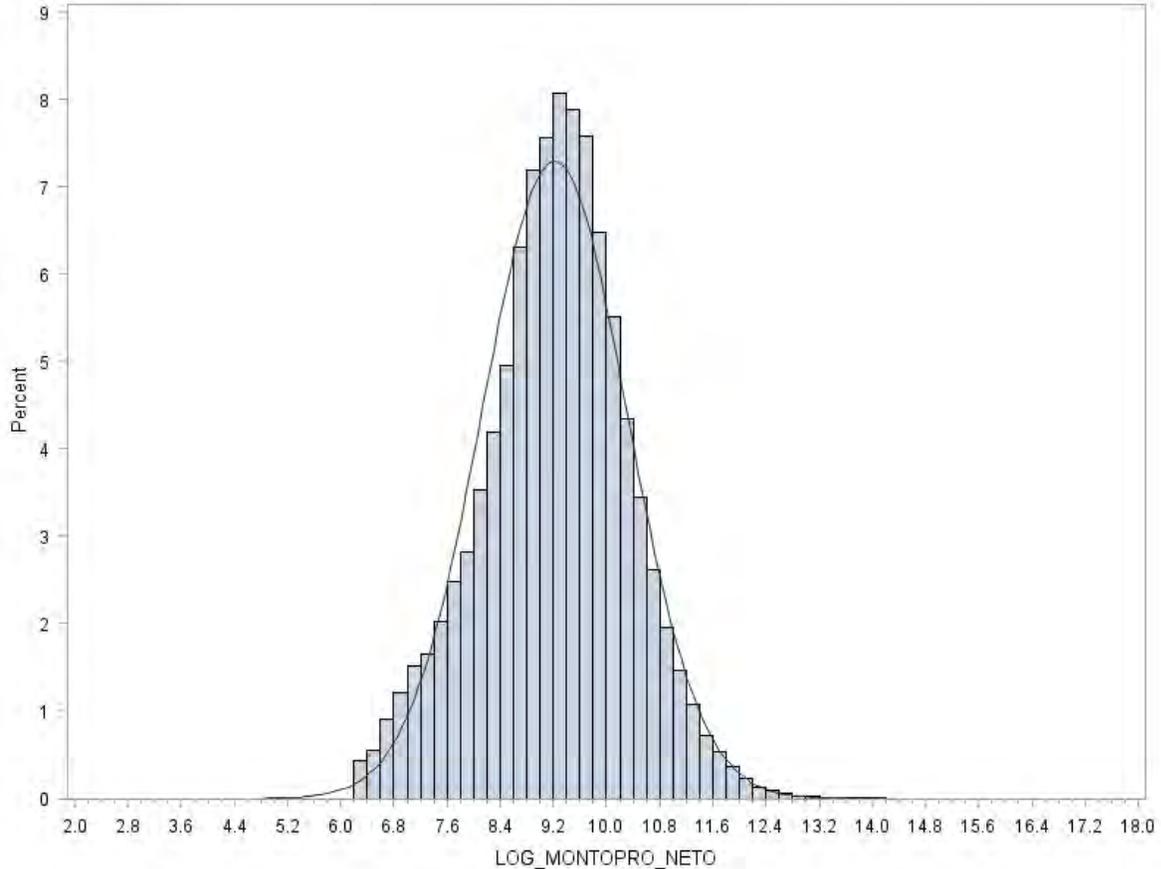


Tabla 2.5 Resumen del logaritmo del parámetro de dispersión final para tipo de Pérdida Parcial.

<i>Parámetro de Dispersión Estimado</i>	
<i>Media</i>	9.217
<i>Desviación Estándar</i>	1.094

Tabla 2.6 Resumen de los cuantiles observados contra los Normales estimados para tipo de pérdida parcial después de remover datos con montos pequeños.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	739.00	789.40
5%	1,409.01	1,664.00
10%	2,289.80	2,476.27
25%	5,173.48	4,811.53
50%	10,533.02	10,064.91
75%	20,367.79	21,054.09

90%	38,526.82	40,909.19
95%	58,300.74	60,878.78
99%	131,134.66	128,328.67

La diferencia máxima que se observa entre cuantiles es de 18.1% y el promedio de las diferencias es de 6.7%. En adición, el Test Kolmogorov-Smirnov da como resultado un estadístico $D= 0.02624$. En general es una mejora bastante considerable y se acerca más al p valor=0.0053. Es importante en este momento mencionar que el ejercicio busca encontrar un valor apropiado x tal que todas las observaciones menores a x queden fuera de los datos a emplear en el modelo, pues mantener estas observaciones actuará en perjuicio del modelo final. Por tanto la prueba sirve como referencia para comparar las distintas propuestas de x y no para demostrar que los datos corresponden a una distribución Normal, esta afirmación es válida en lo sucesivo y hasta antes de la sección 2.4.

2.3.2.2 Ajuste de Pérdidas Totales

Gráfica 2.3 Ajuste inicial de la transformación logarítmica del monto neto de siniestros para tipo de pérdida total.

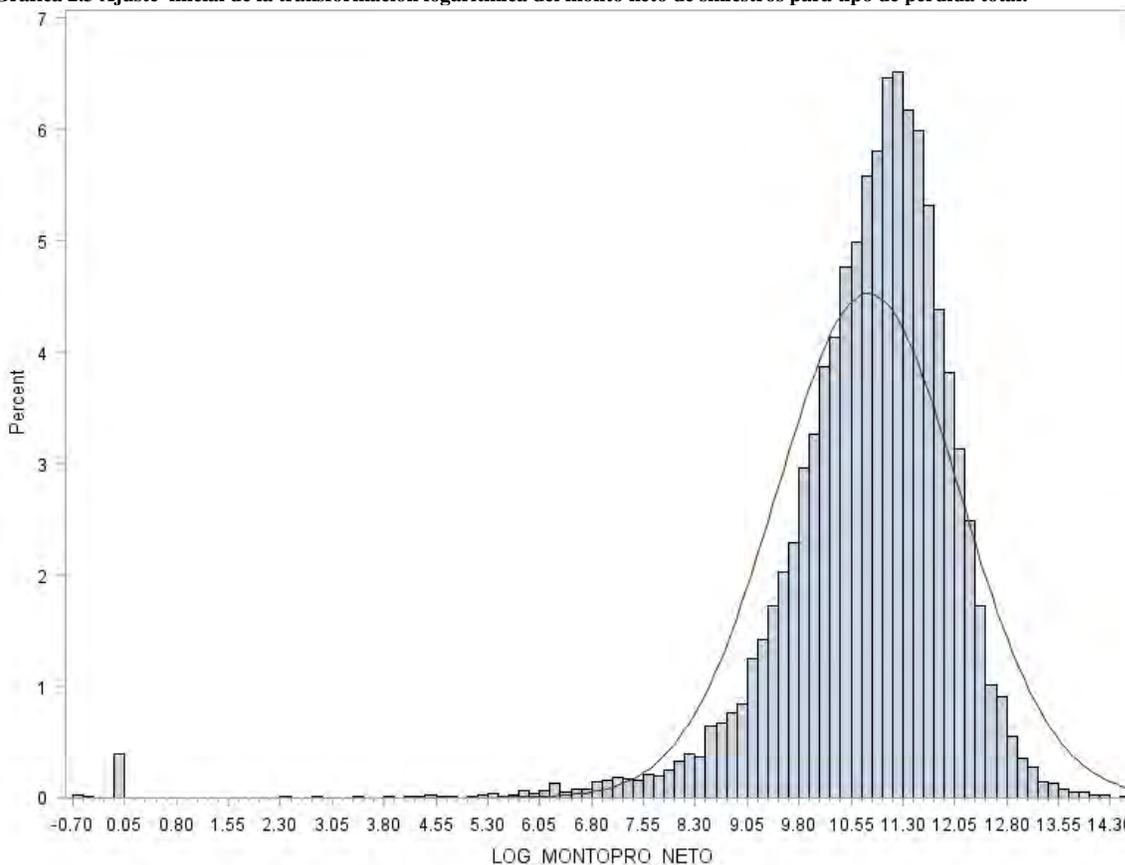


Tabla 2.7 Resumen del logaritmo del parámetro de dispersión inicial para tipo de Pérdida Total.

<i>Parámetro de Dispersión</i>	<i>Estimado</i>
<i>Media</i>	10.801
<i>Desviación Estándar</i>	1.323

Tabla 2.8 Resumen de los cuantiles observados contra los estimados Normales para tipo de pérdida total antes de remover datos con monto pequeño.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	149.96	798.05
5%	1,299.97	2,663.00
10%	6,415.74	5,062.47
25%	26,740.38	14,809.73
50%	67,924.36	48,809.98
75%	142,407.79	160,868.09
90%	266,574.76	470,602.94
95%	395,742.89	894,634.02
99%	827,777.77	2,985,297.60

La diferencia máxima que se observa entre cuantiles es de 432.2% y el promedio de las diferencias es de 123.0%. En adición, el Test Kolmogorov-Smirnov da como resultado un estadístico $D= 0.088762$, tratándose de una muestra con 24,561 registros agregados y $\alpha = 5\%$ el p valor=0.0113.

A continuación se muestra el mismo resumen pero eliminando todas aquellas observaciones con un monto neto de siniestro menor a 10000.

Gráfica 2.4 Ajuste final de la transformación logarítmica del monto neto de siniestros para tipo de pérdida total.

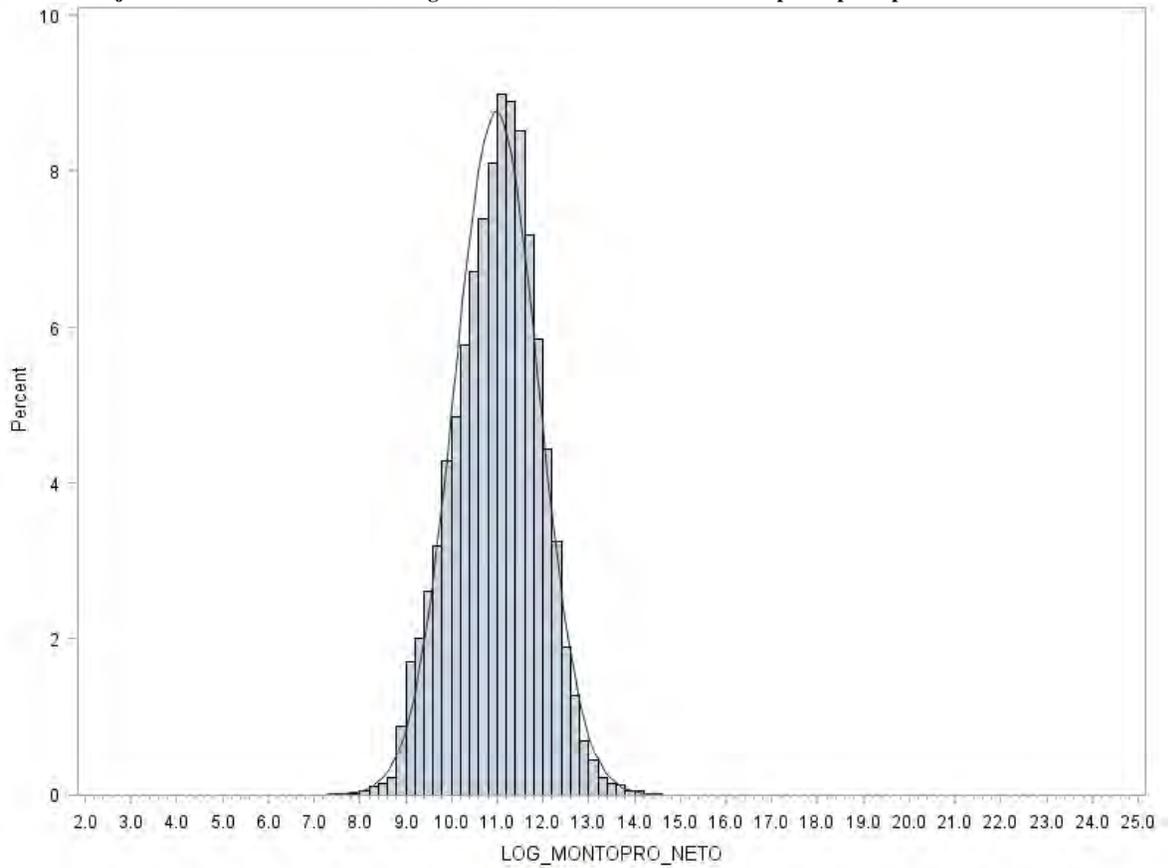


Tabla 2.9 Resumen del logaritmo del parámetro de dispersión final para tipo de Pérdida Total.

<i>Parámetro de Dispersión Estimado</i>	
<i>Media</i>	11.026
<i>Desviación Estándar</i>	0.873

Tabla 2.10 Resumen de los cuantiles observados contra los estimados para tipo de pérdida total después de remover datos con montos pequeños.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	9,333.08	8,068.32
5%	13,931.81	14,626.50
10%	18,844.95	20,084.83
25%	33,618.78	34,120.03
50%	64,173.78	61,476.48
75%	112,199.07	110,766.54
90%	183,029.02	188,167.91
95%	242,784.62	258,388.62
99%	447,230.77	490,985.56

La diferencia máxima que se observa entre cuantiles es de 13.6% y el promedio de las diferencias es de 5.0%. En adición, el Test Kolmogorov-Smirnov da como resultado un estadístico $D=0.0205127$. En general es una mejora bastante considerable y se acerca al p valor=0.0113.

2.3.2.3 Ajuste de Rotura de Cristales

Gráfica 2.5 Ajuste inicial de la transformación logarítmica del monto neto de siniestros para rotura de cristales.

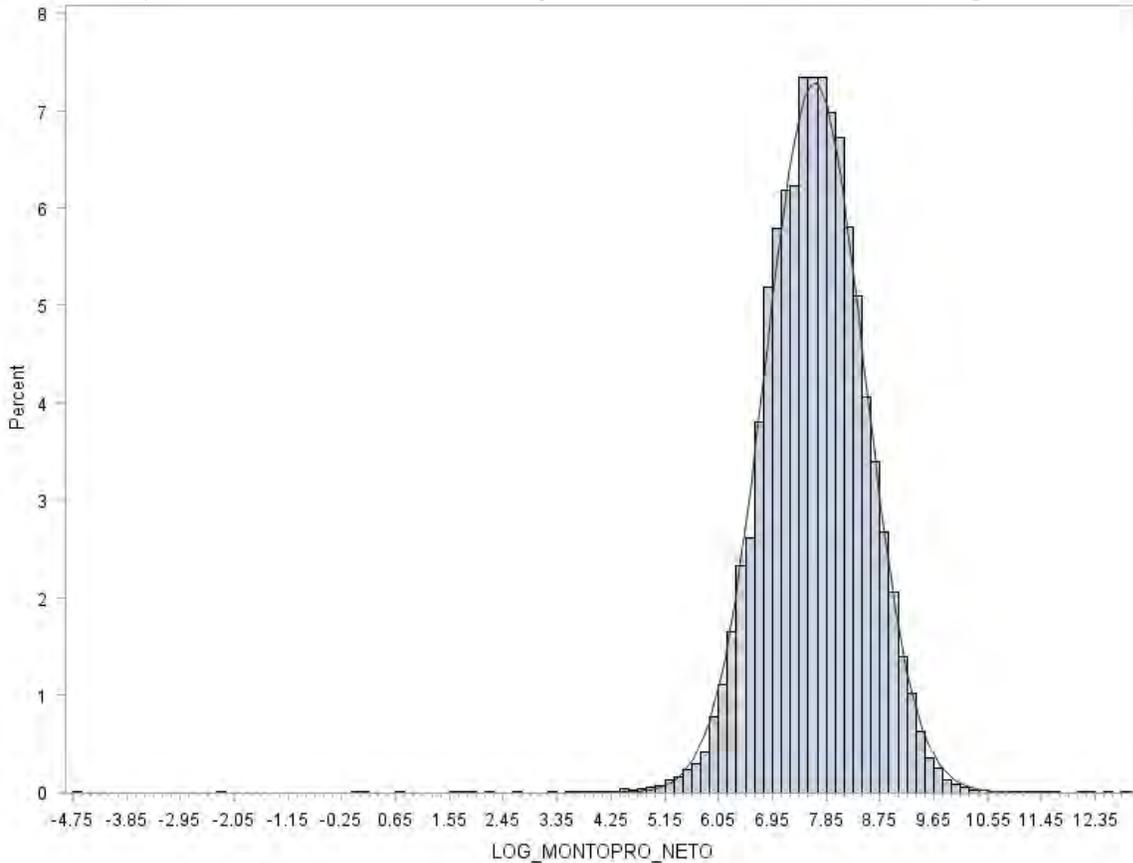


Tabla 2.11 Resumen del logaritmo del parámetro de dispersión inicial para rotura de cristales.

<i>Parámetro de Dispersión</i>		<i>Estimado</i>
<i>Media</i>		7.657
<i>Desviación estándar</i>		0.823

Tabla 2.12 Resumen de los cuantiles observados contra los estimados Normales para rotura de cristales antes de remover los datos con montos pequeños.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	298.92	312.02
5%	572.34	546.63
10%	774.10	737.06
25%	1,220.16	1,214.54
50%	2,126.15	2,115.51

75%	3,653.90	3,684.79
90%	5,954.78	6,071.86
95%	7,911.00	8,187.22
99%	13,258.84	14,343.04

La diferencia máxima que se observa entre cuantiles es de 8.2% y el promedio de las diferencias es de 3.2%. En adición, el Test Kolmogorov-Smirnov da como resultado un estadístico $D=0.0121484$, tratándose de una muestra con 63,764 registros agregados y $\alpha = 5\%$ el p valor=0.0069.

A continuación se muestra el mismo resumen pero eliminando todas aquellas observaciones con un monto neto de siniestro menor a 140.

Gráfica 2.6 Ajuste final de la transformación logarítmica del monto neto de siniestros para rotura de cristales.

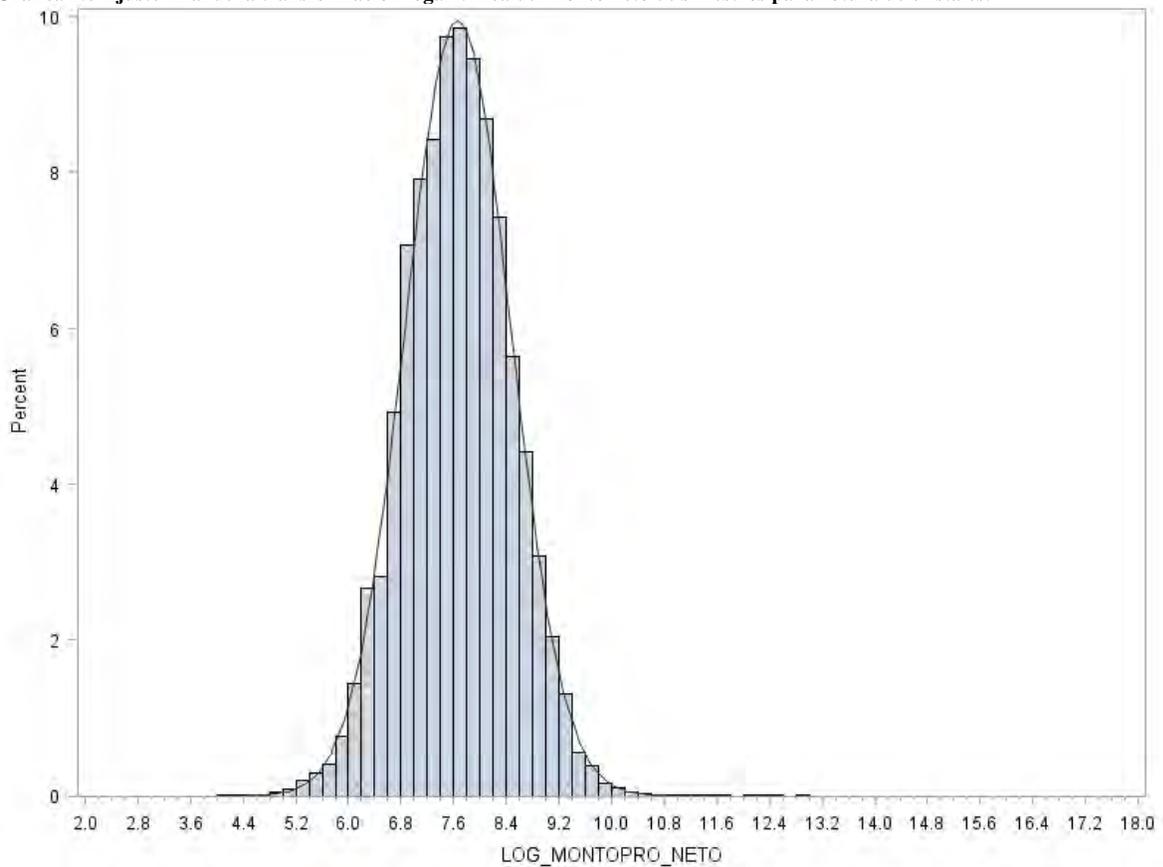


Tabla 2.13 Resumen del logaritmo del parámetro de dispersión final para rotura de cristales.

Parámetro de Dispersión Estimado

<i>Media</i>	7.663
<i>Desviación estándar</i>	0.802

Tabla 2.14 Resumen de los cuantiles observados contra los estimados Normales para rotura de cristales después de remover datos con monto pequeño.

Porcentaje	Cuantil Observado	Cuantil Estimado
1%	321.07	329.58
5%	578.43	569.26
10%	779.70	761.81
25%	1,222.40	1,239.60
50%	2,129.69	2,129.09
75%	3,657.96	3,656.86
90%	5,960.03	5,950.32
95%	7,913.69	7,962.99
99%	13,260.96	13,754.07

La diferencia máxima que se observa entre cuantiles es de 3.7% y el promedio de las diferencias es de 1.4%. En adición, el Test Kolmogorov-Smirnov da como resultado un estadístico $D=0.00637669$, ya menor que el p valor=0.0069.

2.3.3 Ajuste por Pérdidas con Monto Excesivo

Resta analizar los datos atípicos cargados hacia la derecha, es decir, las pérdidas grandes; los montos de dichas pérdidas podrían llegar a desviar el ajuste del modelo, por tanto el tratamiento sugerido es el siguiente:

Sea U la frontera que divide los montos a incluir en el ajuste del modelo de los montos que quedarán fuera por ser muy elevados, entonces:

$$X' = \begin{cases} X & \text{si } X < U \\ U & \text{si } X \geq U \end{cases}$$

Esta nueva variable X' será sobre la que ajustaremos el modelo, el exceso $X - U$ será añadido a la prima base del modelo tal y como se planteó previamente en la sección 2.3.1 para los montos negativos.

Un buen punto de partida para iniciar con este análisis es ver la distancia que hay entre los cuantiles 100%, 99% y 95%. Para aterrizar la idea sugerida supongamos que el cuantil observado al 99% es de 100,000 mientras que el máximo valor observado es de 1,000,000, sería lógico pensar que todo valor por encima de 100,000 es un dato atípico. Bien, de este modo es como procederemos a encontrar candidatos para el valor de U para cada tipo de pérdida.

2.3.3.1 Ajuste de Pérdidas Parciales

El siguiente cuadro muestra una gran distancia entre el máximo y el cuantil del 99%. Sin embargo la gráfica 2.2 sugiere que $\exp(13.2) \approx 540,000$ es un buen candidato para acotar los montos netos grandes de siniestros.

Tabla 2.15 Resumen de los cuantiles para pérdidas parciales antes de remover las pérdidas grandes.

<i>Porcentaje</i>	<i>Cuantil Observado</i>
100%	1,279,936.0
99%	132,868.1
95%	59,075.2
90%	39,149.3
75%	20,755.5
50%	10,873.2
25%	5,543.5
10%	2,722.4
5%	1,830.0
1%	1,155.0
0%	1,000.0

Gráfica 2.7 Ajuste después de remover montos mayores a 540,000 de la transformación logarítmica del monto neto de siniestros para pérdidas parciales.

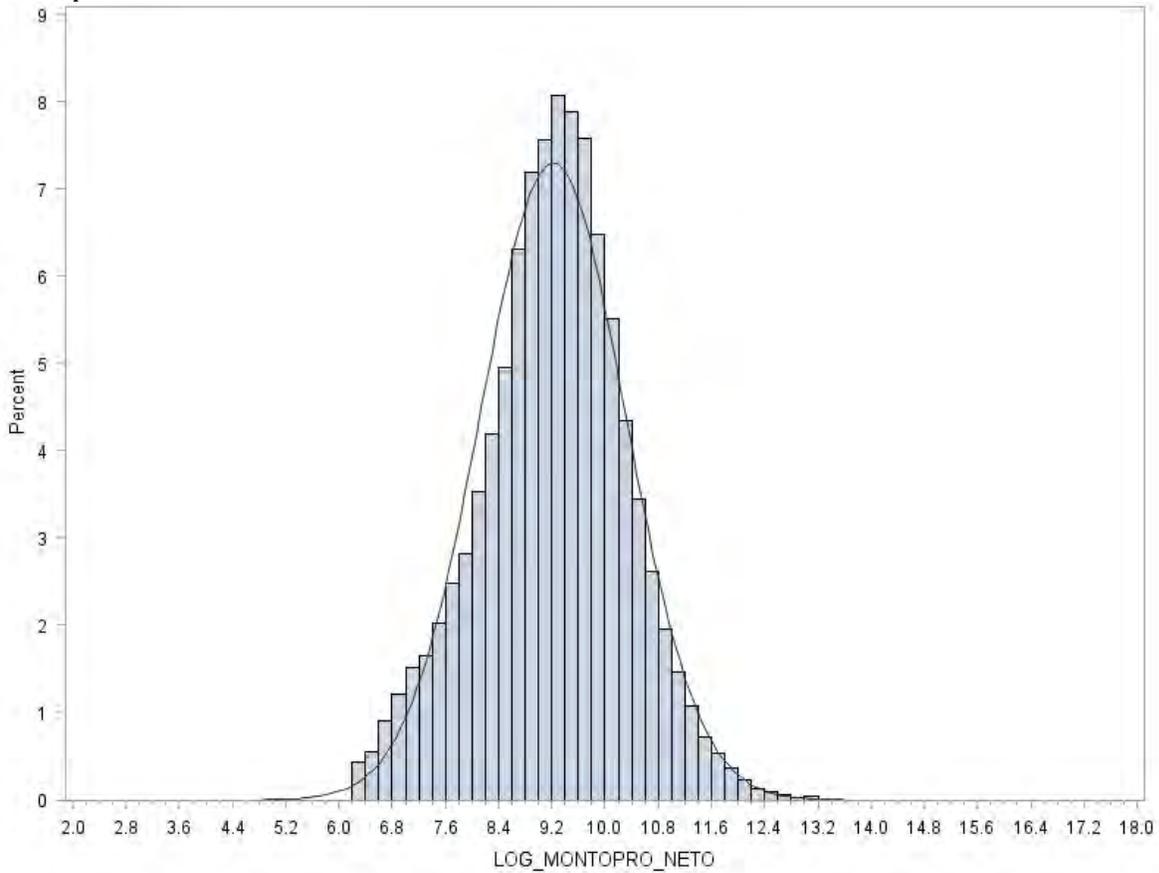


Tabla 2.16 Resumen de los cuantiles observados contra los estimados Normales para pérdidas parciales después de remover las pérdidas grandes.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	739.00	789.88
5%	1,409.01	1,664.67
10%	2,289.80	2,477.02
25%	5,173.48	4,812.11
50%	10,533.02	10,064.10
75%	20,367.79	21,048.40
90%	38,526.82	40,890.78
95%	58,300.74	60,844.70
99%	131,134.66	128,231.18

2.3.3.2 Ajuste de Pérdidas Totales

El siguiente cuadro muestra una distancia amplia entre el máximo y el cuantil del 99%. Sin embargo la gráfica 2.4 sugiere que $\exp(13.5) \approx 730,000$ es un buen candidato para acotar los montos netos grandes de siniestros.

Tabla 2.17 Resumen de los cuantiles para totales antes de remover las pérdidas grandes.

Porcentaje	Cuantil Observado
100%	2,008,905.7
99%	441,093.8
95%	235,299.1
90%	177,067.5
75%	107,462.3
50%	59,152.7
25%	28,093.7
10%	12,433.0
5%	6,567.2
1%	808.3
0%	0.5

Gráfica 2.8 Ajuste después de remover montos mayores a 730,000 de la transformación logarítmica del monto neto de siniestros para pérdidas totales.

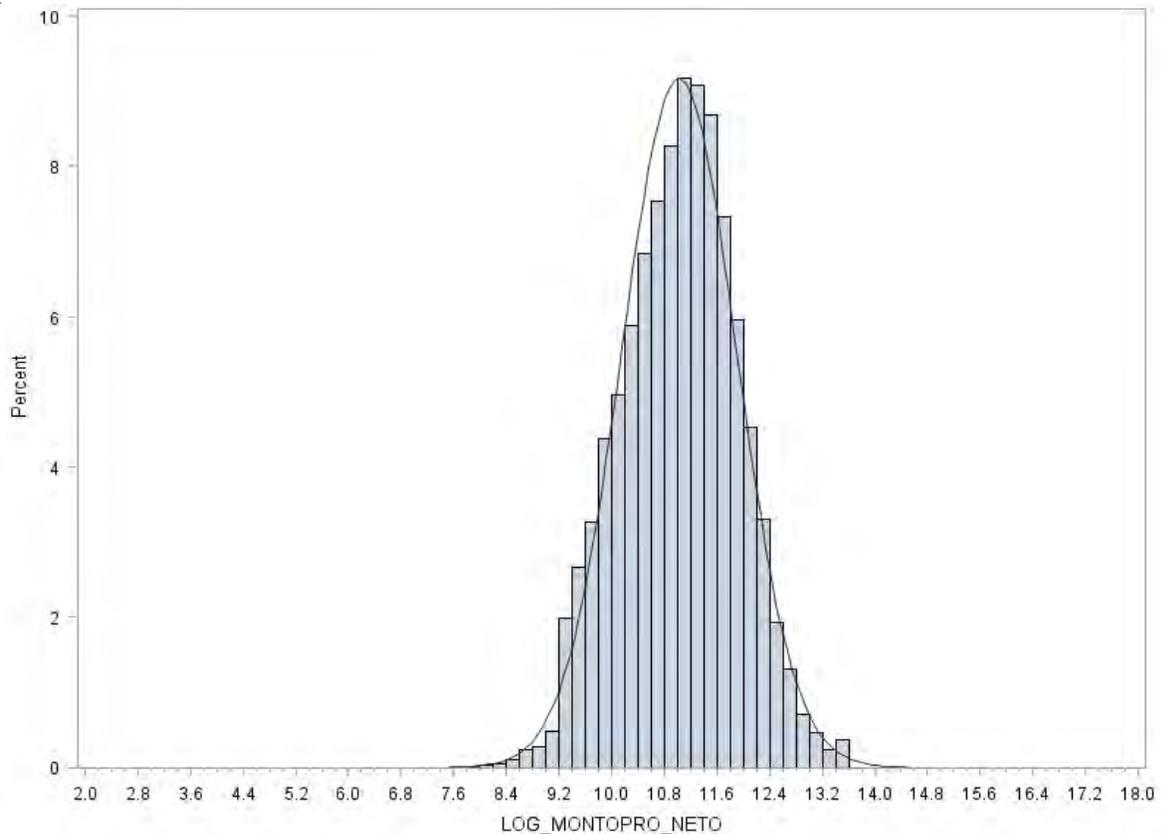


Tabla 2.18 Resumen de los cuantiles observados contra los estimados Normales para pérdidas totales después de remover las pérdidas grandes.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	9,333.08	8,110.62
5%	13,931.81	14,677.05
10%	18,844.95	20,135.30
25%	33,618.78	34,151.44
50%	64,173.78	61,424.87
75%	112,199.07	110,480.02
90%	183,029.02	187,384.89
95%	242,784.62	257,069.05
99%	450,466.95	465,198.74

2.3.3.3 Ajuste Rotura de Cristales

El siguiente cuadro muestra una distancia amplia entre el máximo y el cuantil del 99% Sin embargo la gráfica 2.6 sugiere que $\exp(10.46) \approx 35,000$ es un buen candidato para acotar los montos netos grandes de siniestros.

Tabla 2.19 Resumen de los cuantiles para rotura de cristales antes de remover las pérdidas grandes.

<i>Porcentaje</i>	<i>Cuantil Observado</i>
100%	402,352.7
99%	13,258.8
95%	7,911.0
90%	5,954.8
75%	3,653.9
50%	2,126.1
25%	1,220.2
10%	774.1
5%	572.3
1%	298.9
0%	0.0

Gráfica 2.9 Ajuste después de remover montos mayores a 53,000 de la transformación logarítmica del monto neto de siniestros para rotura de cristales.

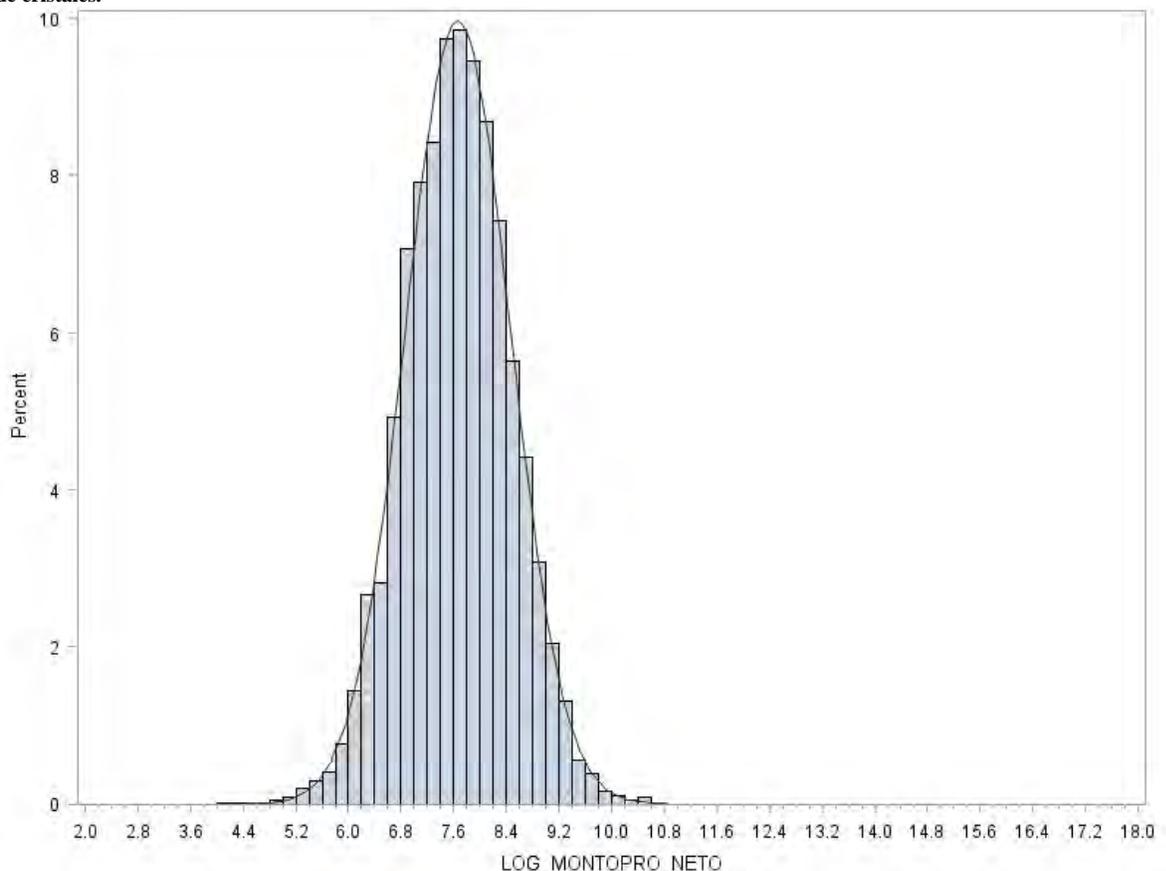


Tabla 2.20 Resumen de los cuantiles observados contra los estimados Normales para pérdidas totales después de remover las pérdidas grandes.

<i>Porcentaje</i>	<i>Cuantil Observado</i>	<i>Cuantil Estimado</i>
1%	321.07	330.85
5%	578.43	570.74
10%	779.70	763.28
25%	1,222.40	1,240.59
50%	2,129.69	2,128.15
75%	3,657.96	3,650.72
90%	5,960.03	5,933.68
95%	7,913.69	7,935.41
99%	13,260.96	13,689.17

Estos ajustes sumarán una mejora considerable para el ajuste del modelo a los datos.

2.3.4 Ajuste por Monto de Alta Frecuencia

Una práctica recomendable dentro del ajuste preliminar de la estructura de error es buscar montos de siniestros con una frecuencia de repetición muy alta. La repetición excesiva de un monto neto de siniestros dentro del conjunto de datos puede indicar, en el caso del sector asegurador, la

existencia de convenios entre las distintas entidades que intervienen en el proceso de pago de la indemnización, e.g., Algunas aseguradoras logran convenios con talleres mecánicos, proveedores o entre ellas (RC Sistema de Pago entre Compañías, SIPAC por sus siglas) con la finalidad de agilizar el proceso de indemnización. En la estimación de la severidad esta clase de situaciones actúan en perjuicio del modelo por ser de carácter determinístico, pues si tratamos de estimar el monto promedio de un siniestro y dicho monto es fijo, luego entonces no es necesario un modelo de severidad.

A continuación se muestran tres tablas, una por cada tipo de pérdida, cada una de ellas despliega los distintos montos de siniestros ordenados de manera descendiente según el número de veces que se han observado a lo largo del archivo plano de siniestros.

Tabla 2.21 Contiene el recuento de los montos netos promedio de siniestros según su aparición en el archivo plano de siniestros para tipo de pérdida parcial.

<i>Promedio Monto Neto Pérdidas Parciales</i>	<i>Recuento</i>	<i>Porcentaje</i>
5,800	342	0.33%
6,700	221	0.21%
1,320	139	0.13%
10,000	138	0.13%
15,000	121	0.12%
5,000	102	0.10%
20,000	91	0.09%
3,000	88	0.08%
7,000	87	0.08%
8,000	86	0.08%

Tabla 2.22 Contiene el recuento de los montos netos promedio de siniestros según su aparición en el archivo plano de siniestros para tipo de pérdida total.

<i>Promedio Monto Neto Pérdidas Totales</i>	<i>Recuento</i>	<i>Porcentaje</i>
730,000	66	0.29%
40,000	8	0.03%
30,000	6	0.03%
15,000	5	0.02%
15,049	5	0.02%
50,000	5	0.02%
11,556	4	0.02%
25,000	4	0.02%
20,000	4	0.02%
35,000	4	0.02%

Tabla 2.23 Contiene el recuento de los montos netos promedio de siniestros según su aparición en el archivo plano de siniestros para rotura de cristales.

<i>Promedio Monto Neto Rotura de Cristales</i>	<i>Recuento</i>	<i>Porcentaje</i>
600	301	0.47%
1,860	102	0.16%
1,200	79	0.12%
760	60	0.09%
2,000	57	0.09%
1,900	56	0.09%
4,000	53	0.08%
960	52	0.08%
1,767	50	0.08%
2,200	48	0.08%

Las tres tablas anteriores indican que no es necesario eliminar ningún registro del archivo plano de siniestro, pues si bien existen montos netos de siniestro cuya frecuencia absoluta de repetición es mayor a uno, la frecuencia relativa de repetición es a lo más 0.47%.

Por último se muestra una tabla igual a la tabla 2.1 pero con las cifras finales antes de amalgamar el archivo plano de siniestros con el de emisión/datos generales y una más con el incremento o decremento para cada campo, derivado de remover registros.

Tabla 2.24 Datos después del ajuste preliminar de siniestros para cada tipo de pérdida.

<i>Tipo de Pérdida</i>	<i>Número de Siniestros</i>	<i>Monto Neto de Siniestros</i>	<i>Severidad</i>
<i>Pérdida Parcial</i>	540,161	7,768,773,745	14,382
<i>Pérdida Total</i>	40,618	3,186,255,236	78,444
<i>Rotura de cristales</i>	196,534	510,942,679	2,600

Tabla 2.25 Tasa de cambio entre el archivo plano de siniestros inicial y final para cada tipo de pérdida.

<i>Tipo de Pérdida</i>	<i>Número de Siniestros</i>	<i>Monto Neto de Siniestros</i>	<i>Severidad</i>
<i>Pérdida Parcial</i>	-0.7%	-1.1%	-0.4%
<i>Pérdida Total</i>	-4.2%	-4.1%	0.1%
<i>Rotura de cristales</i>	-0.1%	-1.0%	-0.9%

2.4 Unión del Archivo plano de Siniestros con el Archivo Plano de Emisión

En esta sección se parte del supuesto de calidad de la información debido a las validaciones del SESA, por lo que para efectos de esta tesis y dado que nuestros archivos planos ya están agrupados por las variables independientes que se describen en la sección 2.5, hemos definido una llave única en función de éstas.

Tabla 2.26 Número de siniestros, monto de siniestros y severidad por tipo de pérdida después de unir el archivo de siniestros con el de emisión.

<i>Tipo de Pérdida</i>	<i>Número de Siniestros</i>	<i>Monto Neto de Siniestros</i>	<i>Severidad</i>
<i>Pérdida Parcial</i>	536,764	7,678,242,654	14,305
<i>Pérdida Total</i>	40,238	3,136,169,513	77,940
<i>Rotura de cristales</i>	194,915	504,644,834	2,589

Tabla 2.27 Tasa de cambio en el número de siniestros, monto de siniestros y severidad por tipo de pérdida después de amalgamar el archivo de siniestros con el de emisión.

<i>Tipo de Pérdida</i>	<i>Número de Siniestros</i>	<i>Monto Neto de Siniestros</i>	<i>Severidad</i>
<i>Pérdida Parcial</i>	-0.6%	-1.2%	-0.5%
<i>Pérdida Total</i>	-0.9%	-1.6%	-0.6%
<i>Rotura de cristales</i>	-0.8%	-1.2%	-0.4%

Ahora se procede a quitar de la tabla de exposición final aquellas celdas de tarifa con exposición igual a cero, puesto que si el conjunto de pólizas que se agrupan en esa celda de tarifa tiene exposición cero no se deberán considerar en el ajuste. Los datos finales para el ajuste del modelo son:

Tabla 2.28 Número de siniestros, monto de siniestros y severidad finales por tipo de pérdida a utilizar en el modelo.

<i>Tipo de Pérdida</i>	<i>Número de Siniestros</i>	<i>Monto Neto de Siniestros</i>	<i>Severidad</i>
<i>Pérdida Parcial</i>	536,543	7,674,868,375	14,304
<i>Pérdida Total</i>	40,218	3,134,535,716	77,939
<i>Rotura de cristales</i>	194,865	504,490,321	2,589

Tabla 2.29 Tasa de cambio final en el número de siniestros, monto de siniestros y severidad por tipo de pérdida.

<i>Tipo de Pérdida</i>	<i>Número de Siniestros</i>	<i>Monto Neto de Siniestros</i>	<i>Severidad</i>
<i>Pérdida Parcial</i>	-0.7%	-1.2%	-0.5%
<i>Pérdida Total</i>	-1.0%	-1.6%	-0.6%
<i>Rotura de cristales</i>	-0.8%	-1.3%	-0.4%

Nuevamente, todas esta clase de pérdidas de información entre un archivo y otro se mutualizarán al final del ajuste.

2.5 Análisis Exploratorio de la Información Estadística del Sector Asegurador Mexicano

Esta sección tiene como objetivo describir el comportamiento marginal de las variables dependientes -frecuencia y severidad-, en función de las variables independientes. La idea es derivar de dicho análisis las variables y sus agrupaciones que compondrán el modelo de frecuencia y severidad. Antes de comenzar a explorar a detalle las cifras es necesario declarar clara y puntualmente el significado de cada variable, tanto dependiente como independiente.

2.5.1 Variables dependientes y de Ponderación

En el capítulo uno definimos tanto a la frecuencia como la severidad. Ambas poseen una expresión algebraica de la forma $Y = X/\omega$ en donde Y también es conocido como indicador clave, X como variable respuesta y ω como variable de ponderación. En esta sección se expone conceptualmente el cálculo y procedencia de cada una de las anteriores variables.

2.5.1.1 Variables de Respuesta

Las variables de respuesta son aquellas cuyo comportamiento se desea estimar con base en los datos reales que se poseen, dicho de otro modo, su comportamiento es aleatorio y deseamos acotar los posibles escenarios con los valores observados de dicha variable.

2.5.1.1.1 Frecuencia: Número de Siniestros

Es el recuento de las eventualidades previstas en el Contrato de Seguro por causa fortuita. Para que dicha eventualidad sea considerada en el recuento es preciso que los daños estén cubiertos en la Póliza, así la Compañía Aseguradora responderá hasta el límite de responsabilidad contratado.

2.5.1.1.2 Severidad: Monto Neto de Siniestros

El monto neto de siniestros tiene la siguiente expresión:

$$\begin{aligned} \text{Monto Neto del Siniestro} \\ = \text{Monto del Siniestro ocurrido} + \text{Gastos de Ajuste} - \text{Salvamentos} \\ - \text{Recuperaciones}; \end{aligned}$$

Donde¹³:

Monto del Siniestro ocurrido: Es el monto del siniestro neto de deducible de los movimientos registrados durante el periodo de análisis. Éste considera los importes de las reservas estimadas más/menos los ajustes a las reservas. Para efectos de los SESAs de Daños de Automóviles de Póliza Individual se consideran todos los montos registrados durante el periodo de análisis, independientemente de la fecha de ocurrencia del siniestro.

Gastos de Ajuste: Monto que considera los gastos directos de ajuste del siniestro generados por la atención del mismo, inclusive para los siniestros improcedentes. Por ejemplo, los pagos de honorarios derivados del ajuste del siniestro.

¹³ Las siguientes definiciones fueron tomadas del Manual del Sistema Estadístico del Ramo de Automóviles de Póliza Individual (Versión 01) 03-dic-2015 con la intención de alinear los resultados del ajuste del modelo a los criterios de extracción de la información.

Salvamentos: Son los restos del Vehículo después de ocurrido el Siniestro de Pérdida Total por daños materiales cuyos derechos han sido cedidos a la Compañía Aseguradora según lo dispuesto por el artículo 116 de la Ley Sobre el Contrato de Seguro.

Recuperaciones: Es el monto que terceros responsables pagan, a consecuencia de un siniestro, como indemnización a la Compañía Aseguradora.

2.5.1.2 Variables Ponderadoras

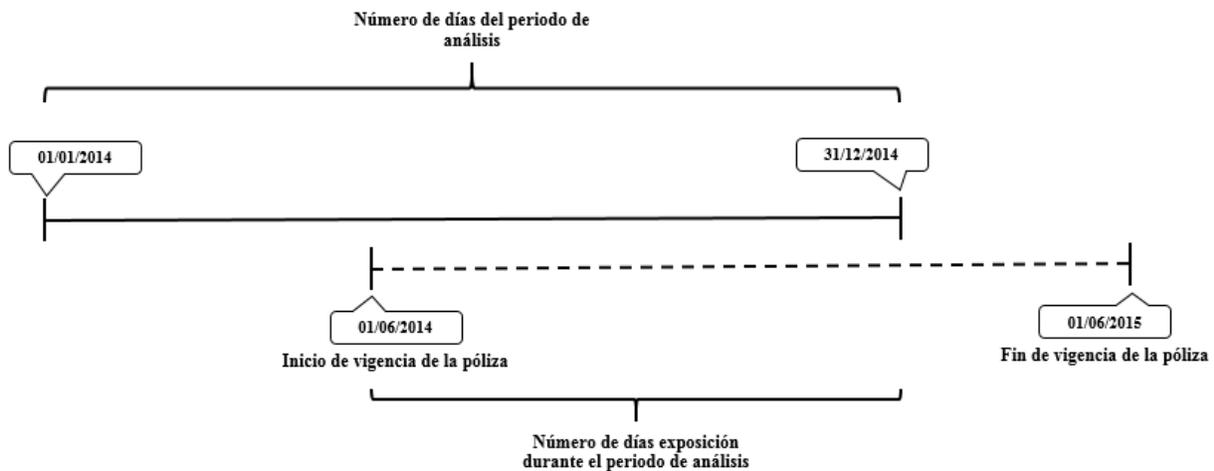
Supongamos que la cartera de una aseguradora está segmentada en tres territorios: Norte, Centro y Sur. Supongamos además que dicha segmentación es consecuencia de un comportamiento distinto entre cada uno de los territorios. Si quisiéramos calcular la frecuencia promedio de toda la cartera deberíamos considerar a cada territorio de manera jerárquica de acuerdo a la participación que cada uno posee en la cartera total, es decir, si el territorio Centro posee el 85% de la participación total luego entonces el promedio de la cartera debería estar sesgado hacia el comportamiento de éste. En virtud de lo anterior los Modelos Lineales Generalizados asignan el peso que cada celda de tarifa aporta a la frecuencia o severidad como se muestra en la siguiente sección.

2.5.1.2.1 Frecuencia: Exposición

La exposición o unidades expuestas representan la proporción de tiempo que la Compañía Aseguradora fue responsable de los riesgos cubiertos en la póliza durante el periodo de análisis. Su forma de cálculo es:

$$\text{Unidades Expuestas} = \frac{\text{Número de días de exposición durante el periodo de análisis}}{\text{Número de días del periodo de análisis}}$$

La siguiente imagen ilustra la el cálculo de las unidades expuestas:



2.5.1.2.2 Severidad: Número de Siniestros

El número de siniestros pasa de ser una variable aleatoria, tal y como lo fue en el estudio de la frecuencia, a ser una variable ponderadora en la severidad: pues la severidad es el costo promedio de los siniestros dado un número fijo de siniestros.

2.5.2 Variables Independientes

Por variables independientes debe entenderse a aquellas variables mediante las cuales intentaremos explicar el fenómeno de estudio, en el caso de esta tesis intentaremos explicar el comportamiento de la frecuencia y luego el de la severidad mediante un conjunto de variables independientes. En esta sección se detalla la forma en que éstas intervienen en el ajuste del modelo.

2.5.2.1 Variables Geográficas

Las variables geográficas son aquellas que describen el comportamiento del riesgo en función de su ubicación, la granularidad de este tipo de variable se recomienda en función de la cantidad de información que se posea, i.e., mientras contemos con más información será preferible una división más exacta de zonas geográficas, pues de caso contrario se pierde la credibilidad en los estimadores para zonas en donde el nivel de información es bajo. El archivo de datos generales contiene como única variable geográfica la entidad federativa, misma que analizaremos a continuación.

2.5.2.1.1 Entidad Federativa

La Entidad federativa corresponde al lugar en donde ocurrió el siniestro. La entidad federativa señalada como “Extranjero” corresponde a todos los siniestros de pólizas contratadas en México pero ocurridos fuera de la República Mexicana.

Recordemos que todo modelo debe cumplir el principio de parsimonia. Éste indica que el modelo debe ser optimizado mediante la minimización del número de parámetros (Número de variables a incluir) y minimización de las diferencias entre los valores reales y los estimados. Por tanto se proponen agrupaciones de los distintos niveles de cada variable para minimizar el número de parámetros finales del modelo. Las agrupaciones buscan reducir el número de parámetros sin comprometer el ajuste de los modelos.

Por ejemplo, buscaremos reducir las distintas entidades federativas de la variable estado en un menor número total de distintas zonas geográficas relevantes para el modelo, buscando siempre dejar en un mismo grupo aquellos estados con un comportamiento similar, i.e., evitando que la dispersión de la frecuencia o severidad observada entre los distintos estados que componen a un mismo grupo no sea mayor a un equis porcentaje que se determinará por tipo de variable respuesta (frecuencia o severidad) y tipo de pérdida (Parcial, total o rotura de cristales). Dicho algoritmo se replicará en todas las variable independientes a incluir en el modelo.

2.5.2.1.1.1 Pérdidas Parciales

Tabla 2.30 Frecuencia de cada entidad Federativa para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

<i>Entidad Federativa</i>	<i>% Exposición</i>	<i>Frecuencia Pérdidas Parciales</i>	<i>Zona</i>
<i>Nuevo León</i>	10.093%	13.25%	1
<i>Sinaloa</i>	2.022%	12.73%	1
<i>Jalisco</i>	9.065%	12.00%	1
<i>Baja California Sur</i>	0.661%	11.63%	1
<i>Coahuila</i>	2.716%	11.42%	2
<i>Tabasco</i>	1.480%	11.24%	2
<i>Tlaxcala</i>	0.371%	10.51%	2
<i>Campeche</i>	0.687%	10.35%	2
<i>Estado de México</i>	9.371%	10.35%	2
<i>Quintana Roo</i>	1.447%	10.34%	2
<i>San Luis Potosí</i>	1.903%	10.26%	2
<i>Yucatán</i>	2.276%	10.25%	2
<i>Veracruz</i>	4.270%	9.98%	2
<i>Guanajuato</i>	3.545%	9.96%	2
<i>Durango</i>	0.957%	8.44%	3
<i>Aguascalientes</i>	1.293%	8.25%	3

<i>Colima</i>	0.719%	8.15%	3
<i>Nayarit</i>	0.639%	7.80%	3
<i>Michoacán</i>	2.184%	7.80%	3
<i>Baja California</i>	2.303%	7.44%	3
<i>Zacatecas</i>	0.641%	7.09%	4
<i>Extranjero</i>	0.346%	0.68%	5
<i>Chihuahua</i>	3.798%	9.74%	5
<i>Tamaulipas</i>	2.654%	9.62%	5
<i>Querétaro</i>	2.723%	9.57%	5
<i>Guerrero</i>	0.804%	9.45%	5
<i>Puebla</i>	3.190%	9.43%	5
<i>Hidalgo</i>	1.152%	9.36%	5
<i>Oaxaca</i>	0.916%	9.33%	5
<i>Sonora</i>	2.306%	9.22%	5
<i>Distrito Federal</i>	20.749%	8.77%	5
<i>Morelos</i>	1.464%	8.65%	5
<i>Chiapas</i>	1.254%	8.60%	5

Nuevo León no sólo posee un porcentaje importante de unidades expuestas sino también el nivel más alto de frecuencia de pérdidas parciales, mientras que el Distrito Federal pese a contener el máximo valor de unidades expuestas tiene uno de los tres valores más pequeños de frecuencia.

Tabla 2.31 Exposición y frecuencia de la agrupación de entidades federativas según los niveles de frecuencia para pérdidas parciales.

<i>Grupos de Zona para P.P.F</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia marginal estimada Zona</i>
1	21.84%	1,162,350	12.63%
2	23.80%	1,266,418	10.45%
3	12.36%	658,026	8.60%
4	0.64%	34,087	7.09%
5	41.36%	2,200,937	9.01%

Tabla 2.32 Severidad de cada entidad Federativa para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

<i>Entidad Federativa</i>	<i>Número Sin Parciales %</i>	<i>Severidad Pérdidas Parciales</i>	<i>Zona</i>
<i>Guerrero</i>	0.6%	24,460	1
<i>Tlaxcala</i>	0.4%	22,294	1
<i>Oaxaca</i>	0.9%	21,351	2
<i>Chiapas</i>	1.1%	21,299	2
<i>Tabasco</i>	1.7%	20,683	2
<i>Zacatecas</i>	0.5%	20,247	2
<i>Hidalgo</i>	1.1%	20,056	2
<i>Quintana Roo</i>	1.5%	19,160	3
<i>Baja California Sur</i>	0.8%	18,735	3

<i>Campeche</i>	0.7%	17,971	3
<i>Michoacán</i>	1.7%	17,595	3
<i>Morelos</i>	1.3%	17,525	3
<i>Puebla</i>	2.9%	16,749	4
<i>Nayarit</i>	0.5%	16,693	4
<i>Veracruz</i>	4.2%	16,620	4
<i>Colima</i>	0.6%	16,561	4
<i>Durango</i>	0.8%	16,385	4
<i>Yucatán</i>	2.3%	16,261	4
<i>Tamaulipas</i>	2.5%	16,176	4
<i>Sinaloa</i>	2.6%	16,164	4
<i>Coahuila</i>	3.1%	15,172	5
<i>Querétaro</i>	2.6%	15,095	5
<i>San Luis Potosí</i>	1.9%	14,786	5
<i>Estado de México</i>	9.6%	14,676	5
<i>Guanajuato</i>	3.5%	14,523	5
<i>Sonora</i>	2.1%	14,264	5
<i>Extranjero</i>	0.1%	43,292	6
<i>Baja California</i>	1.7%	12,868	6
<i>Chihuahua</i>	3.7%	12,804	6
<i>Distrito Federal</i>	18.1%	12,548	6
<i>Aguascalientes</i>	1.1%	12,171	6
<i>Jalisco</i>	10.8%	11,657	6
<i>Nuevo León</i>	13.3%	11,161	6

Si bien Nuevo León fue el estado con la mayor frecuencia de pérdidas parciales es también el estado con el menor monto promedio por siniestro de pérdidas parciales, mientras que Guerrero (sin contar Extranjero cuya participación es casi nula) tiene el mayor monto promedio por cada evento de pérdida parcial, más del doble que el valor mínimo observado.

Tabla 2.33 Número de siniestros en porcentaje y severidad de la agrupación de entidades federativas según los niveles de severidad para pérdidas parciales.

<i>Grupos de Zona para P.P.S.</i>	<i>Número de siniestros %</i>	<i>Número de siniestros</i>	<i>Severidad marginal estimada Zona</i>
1	1.1%	6,121	23,726
2	5.1%	27,303	20,753
3	5.9%	31,644	18,166
4	16.5%	88,467	16,443
5	22.8%	122,473	14,738
6	48.6%	260,535	12,008

2.5.2.1.1.2 Pérdidas Totales

Tabla 2.34 Frecuencia de cada entidad Federativa para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

<i>Entidad Federativa</i>	<i>% Exposición</i>	<i>Frecuencia</i>	<i>Zona</i>
<i>Tabasco</i>	1.5%	1.47%	1
<i>Sinaloa</i>	2.0%	1.42%	1
<i>Guerrero</i>	0.8%	1.33%	1
<i>Baja California Sur</i>	0.7%	1.08%	2
<i>Campeche</i>	0.7%	1.02%	2
<i>Coahuila</i>	2.7%	0.99%	2
<i>Tlaxcala</i>	0.4%	0.97%	2
<i>Chiapas</i>	1.3%	0.90%	3
<i>Yucatán</i>	2.3%	0.89%	3
<i>Hidalgo</i>	1.2%	0.88%	3
<i>Tamaulipas</i>	2.7%	0.84%	3
<i>Oaxaca</i>	0.9%	0.83%	3
<i>Chihuahua</i>	3.8%	0.83%	3
<i>Quintana Roo</i>	1.4%	0.83%	3
<i>Morelos</i>	1.5%	0.83%	3
<i>Veracruz</i>	4.3%	0.83%	3
<i>Nayarit</i>	0.6%	0.81%	3
<i>Nuevo León</i>	10.1%	0.81%	3
<i>Querétaro</i>	2.7%	0.76%	4
<i>Estado de México</i>	9.4%	0.75%	4
<i>Durango</i>	1.0%	0.73%	4
<i>Jalisco</i>	9.1%	0.73%	4
<i>Michoacán</i>	2.2%	0.72%	4
<i>Sonora</i>	2.3%	0.69%	4
<i>Zacatecas</i>	0.6%	0.68%	4
<i>San Luis Potosí</i>	1.9%	0.54%	5
<i>Aguascalientes</i>	1.3%	0.48%	5
<i>Puebla</i>	3.2%	0.66%	6
<i>Colima</i>	0.7%	0.66%	6
<i>Guanajuato</i>	3.5%	0.60%	6
<i>Distrito Federal</i>	20.7%	0.59%	6
<i>Baja California</i>	2.3%	0.58%	6
<i>Extranjero</i>	0.3%	0.01%	6

Llama la atención que los 5 estados con mayor frecuencia de pérdidas parciales sean estados con costa. El Distrito Federal se sigue manteniendo en los niveles más bajos de frecuencia tanto para pérdidas parciales como para pérdidas totales.

Tabla 2.35 Exposición y frecuencia de la agrupación de entidades federativas según los niveles de frecuencia para pérdidas totales.

<i>Grupos de Zona para P.T.F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia marginal estimada Zona</i>
1	4.31%	229,178	1.42%
2	4.44%	236,071	1.00%
3	29.96%	1,594,551	0.83%
4	27.25%	1,450,000	0.74%
5	3.20%	170,085	0.52%
6	0.35%	18,430	0.01%
7	30.51%	1,623,503	0.60%

Tabla 2.36 Severidad de cada entidad Federativa para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

<i>Entidad Federativa</i>	<i>Número Sin Totales %</i>	<i>Severidad Pérdidas Totales</i>	<i>Zona</i>
<i>Extranjero</i>	0.0%	129,392	1
<i>Sinaloa</i>	3.8%	123,969	1
<i>Baja California Sur</i>	0.9%	108,297	1
<i>Chiapas</i>	1.5%	106,820	1
<i>Oaxaca</i>	1.0%	101,552	1
<i>Campeche</i>	0.9%	101,418	1
<i>Tabasco</i>	2.9%	100,215	1
<i>Guerrero</i>	1.4%	97,100	2
<i>San Luis Potosí</i>	1.4%	93,773	2
<i>Michoacán</i>	2.1%	92,097	2
<i>Puebla</i>	2.8%	90,621	2
<i>Nayarit</i>	0.7%	90,516	2
<i>Tlaxcala</i>	0.5%	90,327	2
<i>Colima</i>	0.6%	90,118	2
<i>Veracruz</i>	4.7%	89,702	2
<i>Quintana Roo</i>	1.6%	88,471	2
<i>Sonora</i>	2.1%	86,259	2
<i>Zacatecas</i>	0.6%	85,263	2
<i>Morelos</i>	1.6%	84,838	2
<i>Guanajuato</i>	2.8%	84,676	3
<i>Hidalgo</i>	1.3%	84,566	3
<i>Durango</i>	0.9%	83,659	3
<i>Tamaulipas</i>	2.9%	83,643	3
<i>Estado de México</i>	9.3%	78,664	3
<i>Yucatán</i>	2.7%	78,333	3
<i>Coahuila</i>	3.5%	77,819	3
<i>Querétaro</i>	2.7%	77,375	3
<i>Aguascalientes</i>	0.8%	69,687	4
<i>Baja California</i>	1.8%	68,804	4
<i>Jalisco</i>	8.8%	67,386	4
<i>Chihuahua</i>	4.2%	64,735	4

<i>Nuevo León</i>	10.8%	61,770	5
<i>Distrito Federal</i>	16.3%	59,058	5

Al Distrito Federal le corresponde el menor monto promedio por pérdidas parciales dado que la cantidad de vehículos asegurados de gama baja o antiguos es mayor. Mientras que es más probable que sólo los vehículos más nuevos estén asegurados en otros estados del país.

Tabla 2.37 Número de siniestros y severidad de la agrupación de entidades federativas según los niveles de severidad para pérdidas totales.

<i>Grupos de Zona para P.T.S.</i>	<i>Número de siniestros %</i>	<i>Número de siniestros</i>	<i>Severidad estimada Zona</i>
1	11.1%	4,449	110,206
2	20.0%	8,040	89,957
3	26.3%	10,585	80,061
4	15.5%	6,237	66,955
5	27.1%	10,907	60,139

2.5.2.1.1.3 Rotura de Cristales

Tabla 2.38 Frecuencia de cada entidad Federativa para rotura de cristales, además de una primera reducción categórica de ésta.

<i>Entidad Federativa</i>	<i>% Exposición</i>	<i>Frecuencia</i>	<i>Zona</i>
<i>Sinaloa</i>	2.0%	7.70%	1
<i>Tabasco</i>	1.5%	7.55%	1
<i>Sonora</i>	2.3%	7.16%	1
<i>Guanajuato</i>	3.5%	5.73%	2
<i>Guerrero</i>	0.8%	5.71%	2
<i>Oaxaca</i>	0.9%	5.67%	2
<i>Chiapas</i>	1.3%	5.65%	2
<i>Michoacán</i>	2.2%	5.44%	2
<i>Aguascalientes</i>	1.3%	5.37%	2
<i>Baja California Sur</i>	0.7%	5.29%	2
<i>Colima</i>	0.7%	4.97%	3
<i>Veracruz</i>	4.3%	4.90%	3
<i>Tlaxcala</i>	0.4%	4.63%	3
<i>Hidalgo</i>	1.2%	4.42%	3
<i>Jalisco</i>	9.1%	4.32%	3
<i>Tamaulipas</i>	2.7%	4.31%	4
<i>Coahuila</i>	2.7%	4.26%	4
<i>Nayarit</i>	0.6%	4.02%	4
<i>Querétaro</i>	2.7%	3.93%	4
<i>Baja California</i>	2.3%	3.79%	4
<i>Morelos</i>	1.5%	3.60%	5
<i>Durango</i>	1.0%	3.59%	5

<i>Chihuahua</i>	3.8%	3.43%	5
<i>San Luis Potosí</i>	1.9%	3.16%	5
<i>Zacatecas</i>	0.6%	3.06%	6
<i>Campeche</i>	0.7%	2.86%	6
<i>Estado de México</i>	9.4%	2.82%	6
<i>Puebla</i>	3.2%	2.78%	6
<i>Quintana Roo</i>	1.4%	2.75%	6
<i>Nuevo León</i>	10.1%	2.64%	7
<i>Distrito Federal</i>	20.7%	2.21%	8
<i>Yucatán</i>	2.3%	1.93%	8
<i>Extranjero</i>	0.3%	0.01%	8

El Distrito Federal se mantiene con uno de los niveles más bajos de frecuencia para cualquier tipo de pérdida (Pérdida parcial, total o rotura de cristales), por otro lado Sinaloa es uno de los estados que permanece en el top de frecuencias para cualquier tipo de pérdida.

Tabla 2.39 Exposición y frecuencia de la agrupación de entidades federativas según los niveles de frecuencia para rotura de cristales.

<i>Grupos de Zona para R. Cri. F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia marginal estimada Zona</i>
1	5.81%	309,105	7.45%
2	10.66%	567,202	5.58%
3	15.58%	828,956	4.53%
4	11.04%	587,284	4.08%
5	8.12%	432,237	3.42%
6	15.34%	816,146	2.82%
7	10.09%	537,111	2.64%
8	23.37%	1,243,778	2.15%

Tabla 2.40 Severidad de cada entidad Federativa para rotura de cristales, además de una primera reducción categórica de ésta.

<i>Entidad Federativa</i>	<i>Número de siniestros %</i>	<i>Severidad</i>	<i>Zona</i>
<i>Extranjero</i>	0.0%	3,655	1
<i>Baja California Sur</i>	1.0%	3,322	1
<i>Colima</i>	1.0%	3,135	1
<i>Chiapas</i>	1.9%	2,599	2
<i>Campeche</i>	0.5%	2,548	2
<i>Sonora</i>	4.5%	2,534	2
<i>Aguascalientes</i>	1.9%	2,531	2
<i>Coahuila</i>	3.2%	2,528	2
<i>Tlaxcala</i>	0.5%	2,517	2
<i>Chihuahua</i>	3.6%	2,502	2
<i>Durango</i>	0.9%	2,478	2
<i>Estado de México</i>	7.2%	2,467	2

<i>Veracruz</i>	5.7%	2,448	2
<i>Tamaulipas</i>	3.1%	2,437	2
<i>Morelos</i>	1.4%	2,372	2
<i>Jalisco</i>	10.7%	2,358	3
<i>Nuevo León</i>	7.3%	2,271	3
<i>Yucatán</i>	1.2%	2,203	3
<i>Tabasco</i>	3.1%	2,196	3
<i>Querétaro</i>	2.9%	2,900	4
<i>Puebla</i>	2.4%	2,888	4
<i>Nayarit</i>	0.7%	2,859	4
<i>Distrito Federal</i>	12.5%	2,852	4
<i>Hidalgo</i>	1.4%	2,841	4
<i>Michoacán</i>	3.2%	2,826	4
<i>Guerrero</i>	1.3%	2,797	4
<i>Sinaloa</i>	4.3%	2,793	4
<i>Baja California</i>	2.4%	2,751	4
<i>Zacatecas</i>	0.5%	2,727	4
<i>Oaxaca</i>	1.4%	2,718	4
<i>San Luis Potosí</i>	1.6%	2,702	4
<i>Quintana Roo</i>	1.1%	2,675	4
<i>Guanajuato</i>	5.5%	2,658	4

La variabilidad del costo promedio por rotura de cristales es muy pequeña entre los estados de la República Mexicana, sin embargo si existen diferencias importantes como para incluir la zona geográfica como variable dentro del modelo.

Tabla 2.41 Número de siniestros y la severidad de la agrupación de entidades federativas según los niveles de costo medio para rotura de cristales.

<i>Grupos de Zona para R. Cri. S.</i>	<i>Número de siniestros %</i>	<i>Número de siniestros</i>	<i>Severidad estimada Zona</i>
1	1.9%	3,764	3,228
2	34.5%	67,210	2,488
3	22.2%	43,339	2,299
4	41.3%	80,552	2,799

2.5.2.2 Variables del Vehículo

Las variables del vehículo resultan de vital importancia pues permiten al actuario estudiar el riesgo en función del objeto asegurado, en muchos casos las características de cada vehículo arrojan pistas de cuan grave se debe esperar que sea el costo de un siniestro. En los archivos planos resulta de interés la relación que guardan Tipo, Marca, Segmento y Tipo de vehículo¹⁴. A continuación se hablará a detalle de cada una de las variables disponibles en los SESAS.

¹⁴ En el glosario se incluye la descripción de los distintos niveles que componen a estas variables.

2.5.2.2.1 Marca, Tipo (Sub-Marca), Segmento y Tipo de Vehículo

Dentro del archivo plano de Datos Generales es reportado un campo denominado como Marca Tipo, dicho campo contiene una clave alfanumérica de 4 caracteres de longitud, a cada combinación de Marca y Tipo (Sub-marca) le corresponde una clave única. La relación Marca Tipo contra clave se encuentra en el Catálogo 150 de Autos. La razón por la que he decidido analizar estos cuatro campos de manera conjunta se deriva de los siguientes hallazgos:

- Cada Sub-marca pertenece a una sola marca, salvo tipos genéricos como: “Pick Up”, “Automóviles”, “Comerciales”, “Caja refrigerada” y “Caja seca”- Cuya exposición no supera el 4.3%. Por tanto es recomendable no incluir simultáneamente ambos factores de riesgo al modelo debido a la colinealidad.
 - Con la finalidad de que cada Sub-marca posea una única marca concatenaremos antes de la sub-marca la marca para lograr que cada sub-marca para aquellos casos en donde el nombre de una Sub-marca esté ligada a dos marcas. Por ejemplo, tratándose de la marca Mazda y de la sub-marca CX3 el resultado sería “Mazda-CX3”.
- Cada Sub-marca pertenece solamente a un tipo de Vehículo, es decir, no existe una sub-marca que simultáneamente sea Automóvil y comercial, a excepción de la etiqueta “Otros”.
 - El tratamiento será el mismo que en el punto anterior, i.e., tendremos “Otros-Autos” y “Otros-Com”.
- Cada Sub-marca pertenece sólo a un segmento menos “CAJA SECA” y “RAM PICK UP”.
 - El tratamiento será el mismo que en los dos puntos anteriores.

Con este nuevo campo “Sub-Marca Mejorado”, podemos garantizar el efecto de estas cuatro variables vehiculares en “Sub-Marca Mejorado”. Por tanto y en virtud de un mejor resultado analizaremos a “Sub-Marca Mejorado” como principal variable de vehículo, añadiendo a esta el Modelo y el Uso del Vehículo.

Como el campo “Sub-Marca Mejorado” posee 697 distintas etiquetas será necesario reducir el número de éstas puesto que el número de unidades expuestas en algunas etiquetas es demasiado pequeño y estimar la frecuencia o severidad para cada uno conducirá a errores. Para efectos de este trabajo se han agrupado las etiquetas guardando congruencia con respecto a la Marca y Segmento de cada vehículo. El total de nuevas etiquetas es 218, siendo este campo “Sub-Marca final” sobre

el que se realizará la fase final de agrupación con respecto al valor esperado de la frecuencia o severidad, según aplique para cada modelo. A continuación se muestran los resultados obtenidos para la frecuencia y severidad de cada tipo de pérdida.

2.5.2.2.1.1 Pérdidas Parciales

Tabla 2.42 Extremos de la distribución de la frecuencia de pérdidas parciales y unidades expuestas agrupando por sub-marca final.

<i>Sub-marca final</i>	<i>Unidades Expuestas</i>	<i>Frecuencia Pérdidas Parciales</i>
VENTO	11,036	16.88%
SERIE 1	6,866	16.78%
SWIFT	23,526	15.69%
SAVEIRO	7,928	15.64%
MAZDA 2	7,244	15.01%
ESTACAS	48,111	4.75%
F-350,F450,F550, ESTACAS,CHASIS CAB	29,451	4.68%
C-30,C35,C36,C3500, CHASIS ESTACAS	9,626	4.60%
OTROS-N/D	288	4.16%
AUTOMOVILES TURISTAS	18,891	2.95%

La frecuencia de pérdidas parciales promedio de la información a ocupar en los modelos es de 10.08%, sin embargo al abrir la vista por las distintas sub-marcas se observa una dispersión considerable. Dentro de las sub-marcas con mayor índices de frecuencia de pérdidas parciales predominan los autos, mientras que en las de menor índice de frecuencia predominan las pick ups (vehículos destinados para volúmenes de carga medianos).

Tabla 2.43 Porcentaje de unidades expuestas y la frecuencia de la agrupación para el campo sub-marca final según los niveles de frecuencia para pérdidas parciales.

<i>Grupo sub-marca propuesto P.P. F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	1.1%	56,600	16.0%
2	11.2%	598,555	13.4%
3	27.0%	1,438,735	11.7%
4	12.3%	656,979	8.7%
5	10.7%	570,577	7.4%
6	5.3%	282,796	6.7%
7	3.0%	157,184	5.7%
8	1.9%	102,610	4.7%
9	27.4%	1,457,783	10.1%

Tabla 2.44 Extremos de la distribución de la severidad de pérdidas parciales y número de siniestros agrupando por sub-marca final.

<i>Sub-marca final</i>	<i>Número de Siniestros</i>	<i>Severidad Pérdidas Parciales</i>
ESTACAS	2,285	65,109
AUDI-LUJO	538	45,400
LAND ROVER-MULTIUSOS	426	42,363

<i>BMW-LUJO</i>	545	38,504
<i>MERCEDES BENZ-MULTIUSOS</i>	1,075	35,569
<i>NEON</i>	2,522	6,942
<i>SEDAN</i>	1,442	6,782
<i>WINDSTAR</i>	1,667	6,689
<i>CHEVY</i>	9,948	5,897
<i>ESCORT</i>	535	5,121

La severidad de pérdidas parciales promedio de la información a ocupar en los modelos es de 14,304. Un aspecto que vale la pena señalar es que dentro de los montos más elevados se encuentran en su mayoría vehículos de lujo a excepción de una pick up de la marca Nissan, valorada como un riesgo no preferente en la mayoría de competidores del Sector Asegurador Mexicano. Por otro lado los vehículos con menor monto promedio tienen como común denominador ser de gama baja, i.e., de bajo valor comercial.

Tabla 2.45 Número de siniestros y la severidad de la agrupación para el campo sub-marca final según los niveles de severidad observados para pérdidas parciales.

<i>Grupo sub-marca propuesto P.P.S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1	0.4%	2,285	65,109
2	0.2%	964	44,058
3	0.3%	1,620	36,556
4	2.4%	12,617	30,785
5	1.3%	6,914	25,817
6	3.0%	16,147	22,922
7	4.3%	22,996	20,718
8	2.8%	14,774	19,003
9	7.1%	37,868	17,193
10	11.8%	63,500	15,581
11	18.4%	98,787	14,105
12	13.6%	73,059	11,286
13	11.1%	59,573	8,120
14	1.1%	6,110	6,836
15	2.0%	10,483.0	5,858
16	20.3%	108,846.0	11,804

2.5.2.2.1.2 Pérdidas Totales

Tabla 2.46 Extremos de la distribución de la frecuencia de pérdidas totales y unidades expuestas agrupando por sub-marca final.

<i>Sub-marca final</i>	<i>Unidades Expuestas</i>	<i>Frecuencia Pérdidas Totales</i>
MEGANE	8,464	2.07%
206	11,904	1.99%
VOLVO-LUJO	10,519	1.62%
MONDEO	6,896	1.48%
LEON	8,184	1.42%
JEEP PATRIOT	30,746	0.29%
HIGHLANDER	8,692	0.28%
TRAVERSE	6,518	0.26%
JOURNEY	38,609	0.23%
TUR-AUTOMOVILES	18,891	0.21%

La frecuencia de pérdidas totales promedio de la información a ocupar en los modelos es de 0.76%. De lo anterior podemos observar que los tres valores máximos supera al doble del promedio, el mínimo “TUR-AUTOMOVILES” hace referencia a todos aquellos vehículos que son conducidos en territorio mexicano por turistas, una definición bastante alineada al resultado de la frecuencia de pérdidas totales pues el número de días expuestos de una póliza con dichas características será menor al del resto.

Tabla 2.47 Porcentaje de unidades expuestas y la frecuencia de la agrupación para el campo sub-marca final según los niveles de frecuencia para pérdidas totales.

<i>Grupo sub-marca propuesto P.T. F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	0.4%	20,367	2.0%
2	0.5%	25,599	1.5%
3	4.1%	217,684	1.3%
4	10.5%	557,782	1.1%
5	10.4%	552,650	0.9%
6	20.8%	1,105,259	0.8%
7	10.7%	568,461	0.6%
8	8.0%	424,977	0.5%
9	7.3%	386,242	0.5%
10	5.8%	310,518	0.4%
11	2.7%	142,150	0.3%
12	19.0%	1,010,129	0.7%

Tabla 2.48 Extremos de la distribución de la severidad de pérdidas totales y número de siniestros agrupando por sub-marca final.

<i>Sub-marca final</i>	<i>Número de Siniestros</i>	<i>Severidad Pérdidas Totales</i>
MERCEDES BENZ-MULTIUSOS	186	346,834
AUDI-LUJO	33	327,480
AUDI-MULTIUSOS	183	290,528

<i>ESTACAS</i>	124	234,352
<i>LAND ROVER-MULTIUSOS</i>	92	232,716
<i>POINTER</i>	664	22,004
<i>CHEVY</i>	820	19,941
<i>NEON</i>	281	19,019
<i>KA</i>	108	16,616
<i>ESCORT</i>	47	13,450

La severidad de pérdidas totales de la información a ocupar en los modelos es de 77,939. Nuevamente los líderes en el top de montos promedio mayores son vehículos de lujo y la “ESTACAS” pick up de Nissan, mientras que los montos más bajos los ocupan vehículos de gama baja.

Tabla 2.49 Número de siniestros y la severidad de la agrupación para el campo sub-marca final según los niveles de severidad observados para pérdidas totales.

<i>Grupo sub-marca propuesto P.T.S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1	2.5%	1,013	232,309
2	3.3%	1,313	170,533
3	2.9%	1,176	142,858
4	5.1%	2,045	123,509
5	9.2%	3,719	110,485
6	14.4%	5,811	86,952
7	8.5%	3,437	57,293
8	10.7%	4,292	37,208
9	6.3%	2,525	35,551
10	4.1%	1,663	23,140
11	3.7%	1,484	20,864
12	29.2%	11,740	70,078

2.5.2.2.1.2 Rotura de Cristales

Tabla 2.50 Extremos de la distribución de la frecuencia de rotura de cristales y unidades expuestas agrupando por sub-marca final.

<i>Sub-marca final</i>	<i>Unidades Expuestas</i>	<i>Frecuencia Rotura de Cristales</i>
<i>HIACE</i>	7,988	10.20%
<i>COURIER PICK UP</i>	7,332	8.62%
<i>URVAN PASAJEROS</i>	11,476	7.68%
<i>LEG-COMERCIALES</i>	20,559	7.65%
<i>EXPRESS VAN PASAJEROS</i>	6,275	7.43%
<i>CAVALIER MODELO 2004 AL 1995</i>	10,452	1.05%
<i>CARIBE, BRASILIA, SAFARI</i>	7,765	0.35%

<i>TUR-AUTOMOVILES</i>	18,891	0.34%
<i>AZTEK</i>	9,276	0.20%
<i>CORSAR, VARIANT</i>	6,917	0.01%

La frecuencia de rotura de cristales de la información a ocupar en los modelos es de 3.66%. HIACE, vagoneta de la marca Toyota, se mantiene muy lejos del promedio de frecuencia para rotura de cristales, las características de este vehículo son ideales para el transporte de pasajeros al igual que la “URVAN PASAJEROS”. Los niveles bajos no son una tendencia clara, sin embargo la aparición de “TUR-AUTOMOVILES” no es una sorpresa.

Tabla 2.51 Porcentaje de unidades expuestas y la frecuencia de la agrupación para el campo sub-marca final según los niveles de frecuencia para rotura de cristales.

<i>Grupo sub-marca propuesto R. Cri. F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	0.2%	7,988	10.2%
2	3.5%	187,952	6.8%
3	16.2%	862,808	3.9%
4	21.7%	1,155,333	3.3%
5	13.2%	704,107	2.9%
6	5.9%	312,599	2.5%
7	6.4%	341,708	2.0%
8	2.2%	119,188	1.5%
9	3.4%	180,824	1.2%
10	0.7%	35,932	0.3%
11	26.6%	1,413,379	4.9%

Tabla 2.52 Extremos de la distribución de la severidad de rotura de cristales y número de siniestros agrupando por sub-marca final.

<i>Sub-marca final</i>	<i>Número de Siniestros</i>	<i>Severidad Rotura de Cristales</i>
<i>LAND ROVER-MULTIUSOS</i>	156	13,459
<i>AZTEK</i>	19	8,669
<i>BMW-LUJO</i>	172	8,610
<i>MERCEDES BENZ-LUJO</i>	375	8,548
<i>AUDI-LUJO</i>	131	8,000
<i>CARIBE, BRASILIA, SAFARI</i>	27	853
<i>CHEVY C2</i>	2,716	810
<i>CHEVY</i>	2,702	775
<i>SEDAN</i>	609	717
<i>TSURU</i>	7,245	604

La severidad de rotura de cristales de la información a ocupar en los modelos es de 2,588. Por tercera ocasión los montos más elevados son ocupados por vehículos de lujo, mientras que los más bajos corresponden a vehículos de gama baja. La aparición del vehículo de la marca Pontiac “AZTEK” llama la atención, sin embargo su estructura trasera deja espacio para montos altos.

Tabla 2.53 Número de siniestros y la severidad de la agrupación para el campo sub-marca final según los niveles de severidad observados para Rotura de cristales.

<i>Grupo sub-marca propuesto R. Cri. S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1	0.8%	1,615	8,606
2	1.0%	1,930	6,687
3	2.1%	4,022	5,730
4	4.4%	8,568	4,919
5	6.6%	12,852	4,166
6	9.2%	17,954	3,675
7	11.1%	21,642	3,169
8	8.3%	16,176	2,776
9	13.1%	25,467	2,377
10	5.4%	10,470	2,005
11	9.5%	18,462	1,788
12	9.3%	18,219	1,508
13	6.8%	13,273	686
14	12.4%	24,215	1,170

2.5.2.2.2 Modelo

El modelo corresponde al año de fabricación del vehículo. En la práctica actuarial del Ramo de Automóviles de Póliza Individual es muy común incluir esta variable, pues en muchos casos sucede que vehículos de modelos más recientes poseen una menor frecuencia en comparación con los modelos más antiguos. De cualquier modo, dicha hipótesis será comprobada más adelante. A continuación se estudia el comportamiento de las variables respuesta y de ponderación a lo largo del modelo del vehículo.

Pensando en cortes longitudinales de información es recomendable hacer una transformación lineal a la variable modelo tal y como sigue:

$$Edad_Veh = \begin{cases} -1, & \text{si } A - \text{modelo} < 0 \\ A - \text{modelo}, & \text{si } A - \text{modelo} \geq 0 \end{cases}$$

Donde:

A: es el año del en que comienza la vigencia de la póliza.

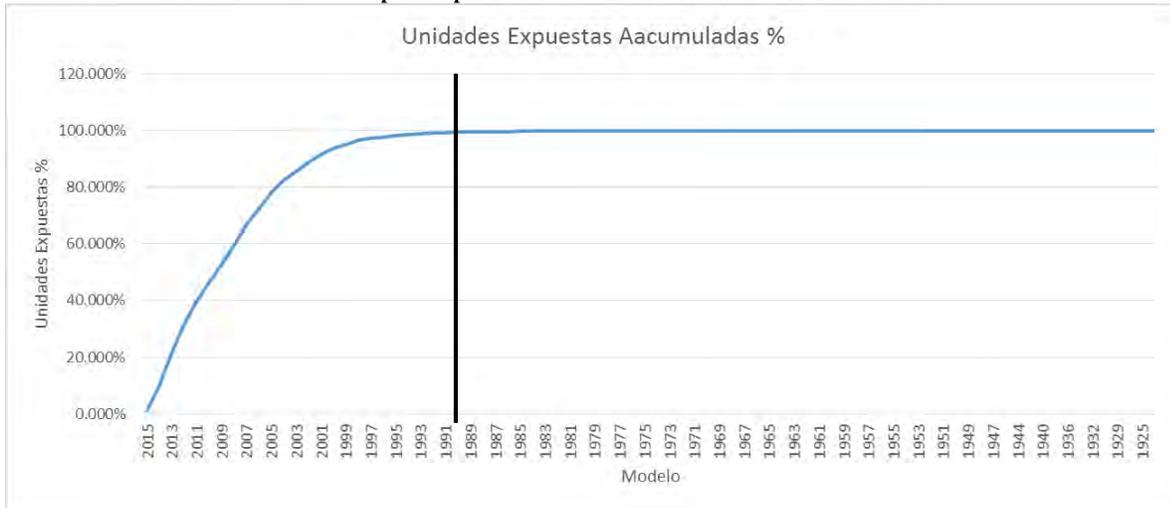
Por ejemplo, tratándose de una póliza con inicio de vigencia el primero de febrero de 2014 que asegura a un vehículo modelo 2008, la edad del vehículo será: $Edad_Veh = 2014 - 2008 = 6$.

Por otro lado para una póliza con mismo inicio de vigencia pero que asegura a un vehículo modelo 2015, la edad del vehículo será: $Edad_{veh} = 2014 - 2015 = -1$.

Una primera versión de esta variable tendría alrededor de 82 distintos niveles, sin embargo si vemos la gráfica de unidades expuestas (gráfica 2.10) acumuladas en porcentaje, tal y como se muestra abajo, podemos decir que después de 1990 (línea negra) el crecimiento es marginal, dicho de otro modo, hay pocas unidades expuestas en modelos anteriores a 1990.

Cabe mencionar que el año modelo promedio es 2008, mientras que el año con mayor cantidad de unidades expuestas es 2013, con alrededor del 11.8% del total de unidades expuestas.

Gráfica 2.10 Acumulación de unidades expuestas por modelo de vehículo ordenado de manera decreciente.



Por tanto, modelos mayores a 1990 los agruparemos en un solo nivel. A Continuación se exponen las distintas agrupaciones a usar en el ajuste de los modelos.

2.5.2.2.1 Pérdidas Parciales

Tabla 2.54 Frecuencia de cada edad de vehículo para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

<i>Edad Veh</i>	<i>% Exposición</i>	<i>Frecuencia Pérdidas Parciales</i>	<i>Grupo</i>
4	6.7%	9.76%	1
7	7.1%	9.74%	1
6	7.2%	9.72%	1
9	5.8%	9.57%	1
13	2.8%	8.56%	1
5	6.2%	9.67%	2
8	5.7%	9.50%	2
10	4.1%	9.41%	2
12	3.2%	9.08%	2
11	3.4%	8.96%	2
14	2.2%	8.02%	3

15	1.4%	7.75%	3
16	1.3%	7.11%	3
17	0.8%	6.68%	4
18	0.5%	6.14%	4
20	0.4%	5.61%	5
19	0.4%	5.59%	5
21	0.3%	4.54%	6
23	0.2%	4.50%	6
24	0.1%	4.40%	6
22	0.3%	4.37%	6
Más de 24	0.5%	2.61%	7
3	7.8%	10.20%	8
-1	1.0%	14.95%	8
0	9.1%	12.65%	8
1	11.8%	11.83%	8
2	10.1%	10.92%	8

Del cuadro anterior podemos concluir que la mayor frecuencia de pérdidas parciales está asociada a los vehículos más recientes, mientras que la menor corresponde a vehículos con los modelos más antiguos.

Tabla 2.55 Exposición y frecuencia de la agrupación de edades de vehículo según los niveles de frecuencia para pérdidas parciales.

<i>Grupo Edad Veh propuesto P.P.F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	1.0%	1,568,631	9.6%
2	20.8%	1,200,831	9.4%
3	17.9%	260,399	7.7%
4	49.3%	66,062	6.5%
5	6.4%	43,624	5.6%
6	2.5%	46,035	4.5%
7	0.8%	24,467	2.6%
8	0.9%	2,111,769	11.5%

Tabla 2.56 Severidad de cada edad de vehículo para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

<i>Edad Veh</i>	<i>Número de siniestros %</i>	<i>Severidad</i>	<i>Grupo</i>
2	10.9%	17,075	1
3	7.9%	15,970	1
4	6.5%	15,322	2
5	5.9%	14,454	2
6	6.9%	13,526	3
7	6.8%	12,606	3

8	5.4%	11,514	4
9	5.5%	10,384	5
10	3.9%	9,453	5
11	3.0%	8,933	6
12	2.8%	8,343	6
13	2.4%	7,634	7
14	1.8%	7,516	7
15	1.1%	6,861	8
17	0.5%	6,683	8
Más de 24	0.1%	6,676	8
16	0.9%	6,615	8
18	0.3%	6,281	8
19	0.2%	5,711	9
20	0.2%	5,533	9
21	0.1%	5,259	9
23	0.1%	4,974	10
24	0.1%	4,921	10
22	0.1%	4,644	10
0	11.4%	19,048	11
1	13.8%	18,097	11
-1	1.5%	17,927	11

Se concluye que la mayor severidad de pérdidas parciales está asociada a los vehículos más recientes, es lógico suponer lo anterior puesto que el costo de los vehículos más recientes es mayor que el de los más antiguos. En el extremo contrario se mantienen los vehículos con los modelos más antiguos.

Tabla 2.57 Número de siniestros y la severidad de la agrupación de edades de vehículo según los niveles de severidad para pérdidas parciales.

<i>Grupo Edad Veh propuesto</i>	<i>Número de Siniestros</i>	<i>Número de</i>	<i>Severidad Grupo</i>
<i>P.P.S.</i>	<i>%</i>	<i>Siniestros</i>	
1	19%	100,852	16,612
2	12%	66,503	14,908
3	14%	73,798	13,070
4	5%	29,044	11,514
5	9%	50,005	9,998
6	6%	31,288	8,646
7	4%	22,130	7,584
8	3%	15,499	6,684
9	1%	3,177	5,539
10	0%	1,319	4,811
11	27%	142,928	18,494

2.5.2.2.2 Pérdidas Totales

Tabla 2.58 Frecuencia de cada edad de vehículo para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

<i>Edad Veh</i>	<i>% Exposición</i>	<i>Frecuencia Pérdidas Totales</i>	<i>Grupo</i>
10	4.1%	0.96%	1
9	5.8%	0.94%	1
8	5.7%	0.93%	1
11	3.4%	0.92%	1
12	3.2%	0.86%	1
7	7.1%	0.85%	1
6	7.2%	0.84%	1
13	2.8%	0.82%	2
14	2.2%	0.76%	2
5	6.2%	0.75%	2
15	1.4%	0.71%	2
18	0.5%	0.59%	3
0	9.1%	0.57%	3
19	0.4%	0.50%	3
20	0.4%	0.43%	4
21	0.3%	0.43%	4
23	0.2%	0.39%	4
-1	1.0%	0.38%	4
22	0.3%	0.24%	5
24	0.1%	0.24%	5
Más de 24	0.5%	0.13%	6
17	0.8%	0.71%	7
4	6.7%	0.71%	7
2	10.1%	0.70%	7
1	11.8%	0.70%	7
3	7.8%	0.68%	7
16	1.3%	0.68%	7

En términos de la frecuencia de pérdidas totales la tendencia no es clara puesto que las edades medias resultan tener un peor resultado, mientras que las edades más altas son aquellas con la frecuencia más baja y finalmente las edades más recientes se sitúan cercanas al promedio de frecuencia para este tipo de pérdida.

Tabla 2.59 Exposición y frecuencia de la agrupación de edades de vehículo según los niveles de frecuencia para pérdidas totales.

<i>Grupo</i>	<i>Edad Veh</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
	<i>propuesto P.T.F.</i>			
1		36.4%	1,937,015	0.9%
2		12.5%	666,868	0.8%
3		10.0%	531,264	0.6%

4	1.9%	99,568	0.4%
5	0.4%	20,506	0.2%
6	0.5%	24,467	0.1%
7	38.4%	2,042,130	0.7%

Tabla 2.60 Severidad de cada edad de vehículo para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

<i>Edad Veh</i>	<i>Número de siniestros %</i>	<i>Severidad</i>	<i>Grupo</i>
-1	0.5%	180,929	1
0	6.9%	171,695	1
3	7.0%	94,962	2
4	6.3%	87,090	3
5	6.1%	78,148	3
6	8.0%	68,250	4
7	7.9%	59,032	5
8	7.1%	50,895	6
9	7.2%	40,065	7
Más de 24	0.1%	39,744	7
10	5.2%	34,410	8
11	4.1%	31,157	8
22	0.1%	25,918	9
12	3.6%	25,814	9
13	3.0%	24,550	9
15	1.3%	22,282	10
14	2.2%	21,781	10
16	1.2%	20,126	10
17	0.7%	18,147	11
18	0.4%	17,016	11
19	0.3%	16,057	11
20	0.2%	15,474	11
21	0.2%	15,297	11
23	0.1%	13,412	11
24	0.0%	10,826	11
1	11.0%	139,942	12
2	9.4%	109,122	12

La severidad por otro lado refleja la tendencia decreciente de los valores de vehículo, dicha tendencia responde a la depreciación que sufren los vehículos conforme el tiempo avanza. Sin embargo aquellos bajo la etiqueta “Más de 24” tienen un comportamiento cercano a edades intermedias, esto es debido a que los vehículos clásicos (vehículos con edad mayor a 30 años) poseen un valor de mercado alto.

Tabla 2.61 Número de siniestros y la severidad de la agrupación de edades de vehículo según los niveles de severidad para pérdidas totales.

<i>Grupo Edad Veh propuesto P.T.S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1	7.4%	2,967	172,311
2	7.0%	2,835	94,962
3	12.4%	4,982	82,668
4	8.0%	3,200	68,250
5	7.9%	3,194	59,032
6	7.1%	2,840	50,895
7	7.2%	2,908	40,061
8	9.4%	3,767	32,980
9	6.7%	2,686	25,245
10	4.7%	1,893	21,507
11	1.9%	762	16,657
12	20.3%	8,184	125,707

2.5.2.2.2.3 Rotura de Cristales

Tabla 2.62 Frecuencia de cada edad de vehículo para rotura de cristales, además de una primera reducción categórica de ésta.

<i>Edad Veh</i>	<i>Exposición %</i>	<i>Frecuencia Pérdidas Totales</i>	<i>Grupo</i>
6	7.2%	4.2%	1
8	5.7%	4.2%	1
7	7.1%	4.1%	1
5	6.2%	4.1%	1
4	6.7%	4.0%	1
9	5.8%	4.0%	1
17	0.8%	3.0%	2
16	1.3%	3.0%	2
0	9.1%	3.0%	2
18	0.5%	2.9%	2
19	0.4%	2.5%	3
-1	1.0%	2.5%	3
20	0.4%	2.2%	3
21	0.3%	2.0%	3
23	0.2%	1.7%	4
24	0.1%	1.4%	4
22	0.3%	1.4%	4
Más de 24	0.5%	0.9%	5
10	4.1%	4.0%	6
3	7.8%	3.9%	6
11	3.4%	3.8%	6
12	3.2%	3.7%	6

2	10.1%	3.6%	6
13	2.8%	3.5%	6
1	11.8%	3.3%	6
14	2.2%	3.2%	6
15	1.4%	3.2%	6

La frecuencia para edades entre 4 y 8 años es la mayor, seguida por una mezcla de vehículo de edades recientes (1 y 2 años) con edades más antiguas (10 a 15 años), para ser las edades mayores a 23 años aquellas cuya frecuencia oscila entre las menores.

Tabla 2. 63 Exposición y frecuencia de la agrupación de edades de vehículo según los niveles de frecuencia para rotura de cristales.

<i>Grupo Edad Veh propuesto R. Cri. F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	38.6%	2,054,392	4.1%
2	11.6%	618,486	3.0%
3	2.1%	112,366	2.4%
4	0.6%	29,853	1.5%
5	0.5%	24,467	0.9%
6	46.6%	2,482,254	3.6%

Tabla 2.64 Severidad de cada edad de vehículo para rotura de cristales, además de una primera reducción categórica de ésta.

<i>Edad Veh</i>	<i>Número de siniestros %</i>	<i>Severidad Rotura de Cristales</i>	<i>Grupo</i>
3	8.2%	2,796	1
4	7.3%	2,726	1
5	6.9%	2,640	2
6	8.3%	2,466	2
7	7.9%	2,333	2
Más de 24	0.1%	2,170	3
8	6.6%	2,166	3
9	6.3%	1,981	3
10	4.5%	1,895	3
11	3.5%	1,825	4
12	3.2%	1,700	4
13	2.6%	1,591	4
14	2.0%	1,501	5
17	0.6%	1,495	5
18	0.4%	1,480	5
15	1.2%	1,476	5
19	0.3%	1,454	5
16	1.1%	1,423	5
21	0.2%	1,182	6
24	0.0%	1,160	6

20	0.2%	1,159	6
23	0.1%	1,145	6
22	0.1%	1,117	6
-1	0.7%	4,031	7
0	7.4%	3,767	7
1	10.6%	3,540	7
2	9.8%	3,120	7

La severidad consecuencia de la rotura de cristales posee una tendencia marcadamente decreciente salvo en el caso de los vehículos potencialmente clásicos, al igual que sucedió con la severidad de pérdidas totales.

Tabla 2.65 Número de siniestros y la severidad de la agrupación de edades de vehículo según los niveles de severidad para rotura de cristales.

Grupo Edad Veh propuesto R. Cri. S	Número de Siniestros %	Número de Siniestros	Severidad Grupo
1	15.5%	30,296	2,763
2	23.1%	45,033	2,472
3	17.4%	33,995	2,030
4	9.3%	18,128	1,716
5	5.5%	10,765	1,476
6	0.6%	1,246	1,157
7	28.4%	55,402	3,466

2.5.2.2.3 Uso del Vehículo

La variable “uso del vehículo” es de interés puesto que el uso Taxi Ruletero se encuentra más expuesto al riesgo de sufrir un siniestro de daños materiales que un vehículo de uso particular, en otras palabras el uso de éste último no es tan constante como el del primero. Dado lo anterior es usual que las aseguradoras ajusten el precio de la cobertura en función del uso para el que el asegurado designe a su vehículo. Dentro de los usos más comunes vemos: Privado, Taxi ruletero, Público federal de carga, Público de carga, Transporte de empleados y escolar. No obstante existen otros menos comunes dentro de la nomenclatura de los SESAs de automóviles y en caso de que las compañías aseguradoras suscriban riesgos fuera de los distintos usos de este catálogo se ha reservado el uso “Otros” para identificarlos.

2.5.2.2.3.1 Pérdidas Parciales

Tabla 2.66 Frecuencia de cada uso del vehículo para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

Uso	% Exposición	Frecuencia Pérdidas Parciales	Grupo
Taxi: Ruletero	1.6%	17.5%	1
Autoescuela	0.0%	14.6%	1
Seguridad Privada	0.0%	14.5%	1

<i>Renta diaria</i>	0.1%	8.5%	2
<i>Seguridad Pública</i>	0.0%	8.0%	2
<i>Público federal de carga</i>	0.6%	7.7%	2
<i>Reparto-Mensajería</i>	0.0%	7.6%	2
<i>Público de carga</i>	2.0%	7.5%	2
<i>Servicios de Emergencia</i>	0.0%	7.3%	2
<i>Transporte público pasajeros: Urbanos</i>	0.0%	4.7%	3
<i>Turistas</i>	0.3%	4.0%	3
<i>Transporte de empleados y escolar</i>	1.3%	3.2%	3
<i>Taxi: Transporte ejecutivo/turismo</i>	0.0%	22.6%	4
<i>Grúas y arrastre</i>	0.0%	12.4%	4
<i>Transporte de turistas (Privado)</i>	0.0%	11.7%	4
<i>Particular</i>	94.0%	10.1%	4
<i>Otros</i>	0.0%	0.3%	4

La frecuencia de pérdidas parciales es considerablemente mayor para aquellos automóviles bajo la etiqueta de “Taxi: Ruletero” que bajo uso “Particular”, por otro lado aquellos automóviles con uso “Público de carga” están asociados a una frecuencia de pérdidas parciales de menor magnitud que el uso “Particular”.

Tabla 2.67 Exposición y frecuencia de la agrupación del uso de vehículo según los niveles de frecuencia para pérdidas parciales.

<i>Grupo Uso Veh Propuesto P.P. F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	1.6%	85,262	17.5%
2	2.7%	145,473	7.6%
3	1.6%	84,166	3.3%
4	94.1%	5,006,917	10.1%

Tabla 2.68 Severidad de cada uso de vehículo para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

<i>Uso</i>	<i>Número de siniestros %</i>	<i>Severidad</i>	<i>Grupo</i>
<i>Transporte de turistas (Privado)</i>	0.0%	35,725	1
<i>Turistas</i>	0.1%	31,673	1
<i>Seguridad Pública</i>	0.0%	30,580	1
<i>Transporte público pasajeros: Urbanos</i>	0.0%	28,788	1
<i>Transporte de empleados y escolar</i>	0.4%	26,621	1
<i>Servicios de Emergencia</i>	0.0%	25,333	1
<i>Público de carga</i>	1.5%	23,157	1
<i>Renta diaria</i>	0.0%	20,754	2
<i>Reparto-Mensajería</i>	0.0%	20,096	2
<i>Seguridad Privada</i>	0.0%	18,408	2
<i>Público federal de carga</i>	0.5%	17,469	2
<i>Autoescuela</i>	0.0%	10,961	3
<i>Taxi: Ruletero</i>	2.8%	10,645	3

Otros	0.0%	168,219	4
Taxi: Transporte ejecutivo/turismo	0.1%	15,401	4
Grúas y arrastre	0.0%	14,745	4
Particular	94.5%	14,172	4

En la severidad de pérdidas parciales para uso “Público de carga” es una de las mayores, un resultado contrario al escenario de la frecuencia de pérdidas parciales, mientras que el uso “Taxi: Ruletero” posee uno de las severidades menores. El uso “Particular” sigue ocupando un sitio cercano a la media.

Tabla 2.69 Número de siniestros y la severidad de la agrupación del uso de vehículo según los niveles de severidad para pérdidas parciales.

Grupo Uso Veh Propuesto P.P.S	Número de Siniestros %	Número de Siniestros	Severidad Grupo
1	2.0%	10,986	24,366
2	0.5%	2,946	17,777
3	2.8%	14,849	10,646
4	94.6%	507,762	14,173

2.5.2.2.3.2 Pérdidas Totales

Tabla 2.70 Frecuencia de cada uso del vehículo para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

Uso	% Exposición	Frecuencia Pérdidas Totales	Grupo
Reparto-Mensajería	0.01%	1.6%	1
Seguridad Pública	0.01%	1.3%	1
Transporte público pasajeros: Urbanos	0.01%	1.2%	1
Seguridad Privada	0.01%	1.2%	1
Grúas y arrastre	0.03%	0.9%	1
Renta diaria	0.05%	0.9%	1
Público de carga	2.01%	0.5%	2
Taxi: Transporte ejecutivo/turismo	0.03%	0.5%	2
Turistas	0.27%	0.3%	3
Transporte de empleados y escolar	1.31%	0.4%	3
Público federal de carga	0.64%	0.8%	4
Servicios de Emergencia	0.02%	0.8%	4
Particular	94.01%	0.8%	4
Taxi: Ruletero	1.59%	0.8%	4
Autoescuela	0.00%	0.0%	4
Otros	0.01%	0.0%	4
Transporte de turistas (Privado)	0.00%	0.0%	4

La frecuencia de pérdidas totales es significativamente menor para el uso “Público de carga” al igual que para el uso “Transporte de empleados y escolar”, mientras que sólo el 0.1% de la exposición total tienen una frecuencia de pérdidas totales por encima del 1%.

Tabla 2.71 Exposición y frecuencia de la agrupación del uso de vehículo según los niveles de frecuencia para pérdidas totales.

<i>Grupo</i>	<i>Uso Veh Propuesto P.T.F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1		0.1%	5,859	1.04%
2		2.0%	108,391	0.54%
3		1.6%	83,827	0.39%
4		96.3%	5,123,741	0.77%

Tabla 2.72 Severidad de cada uso de vehículo para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

<i>Uso</i>	<i>Número de siniestros %</i>	<i>Severidad</i>	<i>Grupo</i>
<i>Transporte público pasajeros: Urbanos</i>	0.0%	176,370	1
<i>Transporte de empleados y escolar</i>	0.7%	148,761	1
<i>Seguridad Pública</i>	0.0%	129,900	1
<i>Turistas</i>	0.1%	121,696	1
<i>Taxi: Transporte ejecutivo/turismo</i>	0.0%	117,560	1
<i>Seguridad Privada</i>	0.0%	116,261	1
<i>Renta diaria</i>	0.1%	106,529	1
<i>Público de carga</i>	1.4%	103,133	1
<i>Servicios de Emergencia</i>	0.0%	95,103	1
<i>Grúas y arrastre</i>	0.0%	83,236	1
<i>Particular</i>	95.3%	77,192	2
<i>Público federal de carga</i>	0.7%	71,438	2
<i>Taxi: Ruletero</i>	1.6%	64,435	2
<i>Reparto-Mensajería</i>	0.0%	38,977	2
<i>Autoescuela</i>	0.0%	-	2
<i>Otros</i>	0.0%	-	2
<i>Transporte de turistas (Privado)</i>	0.0%	-	2

La severidad de pérdidas totales ubica al uso “Taxi: Ruletero” nuevamente como uno de los usos de menor severidad, el uso particular cuya exposición supera en gran medida al resto de los usos también posee una severidad de entre las menores. Existen usos cuya severidad es cero, sin embargo eso no los exenta de sufrir una pérdida de esta naturaleza, por tanto serán agrupados con el nivel de mayor exposición.

Tabla 2.73 Número de siniestros y la severidad de la agrupación del uso de vehículo según los niveles de severidad para pérdidas totales.

<i>Grupo</i>	<i>Uso Veh Propuesto P.T.S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1		2.4%	979	117,976
2		97.6%	39,239	76,940

2.5.2.2.3.3 Rotura de Cristales

Tabla 2.74 Frecuencia de cada uso del vehículo para rotura de cristales, además de una primera reducción categórica de ésta.

Uso	% Exposición	Frecuencia Rotura de Cristales	Grupo
Reparto-Mensajería	0.0%	8.1%	1
Taxi: Ruletero	1.6%	7.9%	1
Seguridad Privada	0.0%	7.6%	1
Taxi: Transporte ejecutivo/turismo	0.0%	7.2%	1
Renta diaria	0.1%	6.9%	1
Transporte de turistas (Privado)	0.0%	6.8%	1
Público federal de carga	0.6%	6.5%	2
Público de carga	2.0%	5.5%	2
Transporte público pasajeros: Urbanos	0.0%	3.8%	3
Particular	94.0%	3.5%	3
Servicios de Emergencia	0.0%	3.2%	3
Seguridad Pública	0.0%	3.2%	3
Transporte de empleados y escolar	1.3%	3.1%	3
Grúas y arrastre	0.0%	2.3%	3
Autoescuela	0.0%	1.3%	3
Turistas	0.3%	0.5%	3
Otros	0.0%	0.3%	3

Tratándose de la frecuencia de rotura de cristales el uso particular se sitúa en las posiciones más bajas, mientras que “Taxi: Ruletero” y “Reparto-Mensajería” son los usos con mayor frecuencia de rotura de cristales.

Tabla 2.75 exposición y frecuencia de la agrupación del uso de vehículo según los niveles de frecuencia para rotura de cristales.

Grupo	Uso Veh Propuesto R. Cri. F.	Exposición %	Exposición	Frecuencia Grupo
1		1.7%	89,803	7.9%
2		2.6%	140,944	5.8%
3		95.7%	5,091,071	3.5%

Tabla 2.76 severidad de cada uso de vehículo para rotura de cristales, además de una primera reducción categórica de ésta.

Uso	Número de siniestros %	Severidad	Grupo
Reparto-Mensajería	0.0%	1,894	1
Seguridad Privada	0.0%	1,832	1
Taxi: Ruletero	3.5%	1,785	1
Otros	0.0%	1,724	1
Transporte público pasajeros: Urbanos	0.0%	1,356	1

<i>Público de carga</i>	3.0%	2,265	2
<i>Público federal de carga</i>	1.1%	2,165	2
<i>Transporte de turistas (Privado)</i>	0.0%	7,312	2
<i>Turistas</i>	0.0%	4,901	2
<i>Taxi: Transporte ejecutivo/turismo</i>	0.1%	4,008	2
<i>Seguridad Pública</i>	0.0%	3,007	2
<i>Autoescuela</i>	0.0%	2,711	2
<i>Transporte de empleados y escolar</i>	1.1%	2,669	2
<i>Particular</i>	91.0%	2,633	2
<i>Servicios de Emergencia</i>	0.0%	2,568	2
<i>Grúas y arrastre</i>	0.0%	2,437	2
<i>Renta diaria</i>	0.1%	2,311	2

El caso de la severidad de rotura de cristales deja al uso “Transporte de Turistas” como el nivel con una mayor severidad asociada, esto puede estar directamente correlacionado con el tipo de sub-marcas que son utilizadas para este fin. La mayoría de los usos poseen una severidad similar al uso “Particular”, mismo que posee un 91% de todos los siniestros, quedando por abajo “Taxi: Ruletero” y otros usos con una menor proporción de siniestros.

Tabla 2.77 Número de siniestros y la severidad de la agrupación del uso de vehículo según los niveles de severidad para rotura de cristales.

<i>Grupo</i>	<i>Uso Veh</i>	<i>Propuesto R. Cri. S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1			3.5%	6,810	1,785
2			96.5%	188,055	2,618

2.5.2.3 Características de Contratación de la Póliza

Las características de contratación de la póliza encierran una gran parte de los acuerdos entre la aseguradora y el asegurado. Dentro del contrato de seguro se estipula el cómo y cuándo se indemnizará al asegurado, es decir, el deducible, el monto al que se pagará la unidad en caso de pérdida total, etc. Tomar en cuenta todas estas características para el entendimiento del riesgo mejorará la aproximación del precio ofertado al cliente en función del cambio en condiciones de contratación.

De hecho, desde el punto de vista regulatorio la Documentación Contractual contiene tanto a las Condiciones Generales como a la carátula de la póliza.

2.5.2.3.1 Deducible

Estudiar el efecto que el deducible tiene sobre la frecuencia y severidad, como factor de riesgo, nos ayudará a comprobar en qué medida es válida la hipótesis: A menor deducible corresponde una mayor frecuencia. Más aún, nos dirá la proporción con la que dichas variables, independiente

y dependiente, se relacionan. Esta variable sólo se considerará para los modelos de frecuencia y severidad de pérdidas totales y parciales. Para la cobertura de cristales es común en la práctica del Sector Asegurador Mexicano la aplicación de un deducible fijo del 20%.

2.5.2.3.1.1 Pérdidas Parciales

Tabla 2.78 Frecuencia de cada deducible para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

<i>Deducible</i>	<i>% Exposición</i>	<i>Frecuencia Pérdidas Parciales</i>	<i>Grupo</i>
3	6.2%	12.7%	1
Resto	4.5%	11.7%	1
10	5.7%	8.1%	2
0	1.7%	4.9%	2
5	81.9%	10.0%	3

Hablando de la frecuencia de pérdidas parciales no se observa que las pólizas con deducible cero sean aquellas con un mayor índice de frecuencia, es más sucede que son aquellas pólizas con el menor índice de frecuencia. El deducible del 3 por ciento es el que posee la mayor frecuencia para este tipo de pérdidas. Un hecho interesante es que el deducible del diez por ciento no posee una frecuencia mucho menor que un deducible del cinco por ciento pese a que el segundo representa la mitad del primero.

Tabla 2.79 Exposición y frecuencia de la agrupación de deducibles según los niveles de frecuencia para pérdidas parciales.

<i>Grupo Deducible Propuesto P.P.F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	10.7%	571,979	12.3%
2	5.7%	295,993	8.1%
3	1.7%	82,405	4.9%
4	81.9%	3,768,347	10.0%

Tabla 2.80 Severidad de cada deducible para tipo de pérdidas parciales, además de una primera reducción categórica de ésta.

<i>Deducible</i>	<i>Número de siniestros %</i>	<i>Severidad</i>	<i>Grupo</i>
10	7.8%	15,542	1
Resto	81.6%	14,438	1
5	4.5%	12,618	2
0	5.2%	12,395	2
3	0.8%	10,643	2

La severidad de pérdidas parciales sitúa al deducible del diez por ciento como aquel con la mayor severidad, contrario a lo que podrías imaginar en una primera instancia (pues si la participación del asegurado es mayor en el siniestro entonces el monto a pagar por la aseguradora

podría ser menor), el 3% de deducible es ahora el de menor nivel de severidad contrario a lo que sucedió en la frecuencia para pérdidas de la misma índole.

Tabla 2.81 Número de siniestros y la severidad de la agrupación de deducibles según los niveles de severidad para pérdidas parciales.

<i>Grupo Deducible Propuesto P.P.S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1	89.4%	479,719	14,535
2	10.6%	56,824	12,356

2.5.2.3.1.2 Pérdidas Totales

Tabla 2.82 Frecuencia de cada deducible para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

<i>Deducible</i>	<i>% Exposición</i>	<i>Frecuencia Pérdidas Totales</i>	<i>Grupo</i>
10	5.7%	0.40%	1
0	1.7%	0.12%	2
3	6.2%	0.82%	3
5	81.9%	0.79%	3
Resto	4.5%	0.77%	3

En el caso de la frecuencia pérdidas totales el deducible cero es también el nivel de menor frecuencia, la conclusión para el deducible del tres por ciento continúa siendo la misma para esta clase de pérdidas, es la de mayor nivel de frecuencia. Lo anterior es prueba de que aquellos conductores que transfieren el total del riesgo cubierto por esta cobertura (deducibles cero) a la aseguradora, son aquellos con mejores hábitos de conducción. Otra conclusión interesante es que aquellas pólizas contratadas con deducible del 5% poseen una frecuencia muy cercana a la frecuencia de pólizas contratadas con cualquier otro porcentaje de deducible distinto a las etiquetas cero, diez y tres por ciento.

Tabla 2.83 Exposición y frecuencia de la agrupación de deducibles según los niveles de frecuencia para pérdidas totales.

<i>Grupo Deducible Propuesto P.T.F.</i>	<i>Exposición %</i>	<i>Exposición</i>	<i>Frecuencia Grupo</i>
1	5.7%	301,023	0.4%
2	1.7%	88,921	0.1%
3	92.7%	4,931,874	0.8%

Tabla 2.84 Severidad de cada deducible para tipo de pérdidas totales, además de una primera reducción categórica de ésta.

<i>Deducible</i>	<i>Número de siniestros %</i>	<i>Severidad</i>	<i>Grupo</i>
Resto	4.6%	96,125	1
0	0.3%	91,856	1
3	6.8%	88,105	2
5	85.3%	76,205	2
10	3.0%	75,184	2

Finalmente la severidad de pérdidas totales indica que en promedio los montos pagados por la aseguradora para deducibles del cinco y diez por ciento son bastante similares, por tanto no existe una razón de peso para asignarlos en grupos de tarifa distintos pese a ser numéricamente muy distintos. En el caso contrario están los deducibles cero, éstos poseen una severidad significativamente mayor pero menor a todos aquellos siniestros que provienen de deducibles distintos a cero, tres, cinco y diez por ciento.

Tabla 2.85 Número de siniestros y la severidad de la agrupación de deducibles según los niveles de severidad para pérdidas totales.

<i>Grupo Deducible Propuesto P.P.S.</i>	<i>Número de Siniestros %</i>	<i>Número de Siniestros</i>	<i>Severidad Grupo</i>
1	4.8%	1,950	95,899
2	95.2%	38,268	77,023

CAPÍTULO 3

3.1 Ajuste de los Modelos

En esta sección se expondrán los resultados del ajuste de los modelos bajo las consideraciones previas, es decir, con el ajuste preliminar del monto de siniestro y con las agrupaciones de los distintos niveles de las variables que compondrán al modelo.

3.1.1 Definición de los Niveles Base

Tal y como se menciona en la sección 1.3 es necesario definir un nivel base para cada una de las variables independientes que acabamos de mencionar. La siguiente tabla resume los resultados del análisis, en donde se asigna a cada variable el nivel base de acuerdo a exposición/número de siniestros.

Tabla 3.1 Resume los niveles base de la cobertura.

<i>Variable</i>	<i>Nivel Base</i>
<i>Entidad</i>	Distrito
<i>Federativa</i>	Federal
<i>Sub-marca</i>	JETTA/CLASICO
<i>Edad Vehículo</i>	1
<i>Uso</i>	Particular
<i>Deducible</i>	5

Los anteriores niveles fueron seleccionados de tal suerte que su exposición y número de siniestros fueran los mayores con respecto al resto de niveles.

3.1.2 Información General del Modelo y Estimación de los Parámetros β 's

A continuación se da un resumen breve de la estructura de error, la función liga, el número de observaciones y la variable dependiente empleados en cada modelo, así como el detalle de la estimación de las β 's, dicha estimación tuvo lugar a través del algoritmo numérico de Newton-Raphson, la representatividad de cada variable dependiente y la vista general del ajuste de bondad para los seis modelos.

3.1.2.1 Frecuencia de Pérdidas Parciales

Tabla 3.2 Información general del modelo de frecuencia de pérdidas parciales.

<i>Información del Modelo</i>	
<i>Estructura de Error</i>	Poisson
<i>Función Liga</i>	Log

Variable Dependiente	Número Siniestros de pérdida parcial
Variable Offset	Logaritmo de las unidades expuestas
Número de Observaciones	408,686

Tabla 3.3 Tabla de representatividad de las variables incluidas en el modelo de frecuencia de pérdidas parciales.

Variable	Grados de Libertad	Chi Cuadrada	P-Value
Entidad Federativa	6	17,814	<.0001
Sub-marca	11	20,269	<.0001
Edad Vehículo	8	6,320	<.0001
Uso Vehículo	6	9,462	<.0001
Deducible	2	9,029	<.0001

Tabla 3.4 Tabla de bondad de ajuste para la frecuencia de pérdidas parciales.

Criterio	Grados de Libertad	Valor	Valor/Grados de Libertad
Devianza	410,000	330,612	0.81
Devianza Escalada	410,000	330,612	0.81
Chi- cuadrada Pearson	410,000	765,582	1.87

Tabla 3.5 Tabla de estimaciones de ajuste para la frecuencia de pérdidas parciales.

Parámetro	Nivel	Estimación $\beta's$	$Exp(\beta's) = \gamma's$
Intercepto	-	2.29	10.13%
Entidad Federativa	1	0.38	1.463
Entidad Federativa	2	0.12	1.128
Entidad Federativa	3	- 0.09	0.917
Entidad Federativa	4	- 0.23	0.791
Entidad Federativa	5	-	1.000
Sub-marca	1	0.37	1.455
Sub-marca	2	0.22	1.250
Sub-marca	3	0.11	1.121
Sub-marca	4	- 0.17	0.845
Sub-marca	5	- 0.25	0.776
Sub-marca	6	- 0.35	0.703
Sub-marca	7	- 0.46	0.634
Sub-marca	8	- 0.69	0.502
Sub-marca	9	-	1.000
Edad Vehículo	1	- 0.14	0.869
Edad Vehículo	2	- 0.15	0.857
Edad Vehículo	3	- 0.27	0.762
Edad Vehículo	4	- 0.40	0.668
Edad Vehículo	5	- 0.53	0.590
Edad Vehículo	6	- 0.70	0.494
Edad Vehículo	7	- 1.16	0.315

Edad Vehículo	8	-	1.000
Uso Vehículo	1	0.84	2.327
Uso Vehículo	2	- 0.06	0.938
Uso Vehículo	3	- 0.95	0.386
Uso Vehículo	4	-	1.000
Deducible	1	0.15	1.161
Deducible	2	- 0.55	0.579
Deducible	3	-	1.000

3.1.2.2 Severidad de Pérdidas Parciales

Tabla 3.6 Información general del modelo de severidad de pérdidas parciales.

Información del Modelo

Estructura de Error	Gamma
Función Liga	Log
Variable Dependiente	Monto de Pérdidas Parciales
Variable Ponderadora	Número Siniestros Parciales
Número de observaciones	95,267

Tabla 3.7 Tabla de representatividad de las variables incluidas en el modelo de severidad de pérdidas parciales.

<i>Variable</i>	<i>Grados de Libertad</i>	<i>Chi Cuadrada</i>	<i>P-Value</i>
Entidad Federativa	5	5,039	<.0001
Sub-marca	15	13,109	<.0001
Edad Vehículo	10	7,924	<.0001
Uso Vehículo	3	543	<.0001
Deducible	1	184	<.0001

Tabla 3.8 Tabla de bondad de ajuste para la severidad de pérdidas parciales.

<i>Criterio</i>	<i>Grados de Libertad</i>	<i>Valor</i>	<i>Valor/Grados de Libertad</i>
Devianza	95,000	20	0.0002
Devianza escalada	95,000	95,267	1
Chi- cuadrada Pearson	95,000	18	0.0002

Tabla 3.9 Tabla de estimaciones de ajuste para la severidad de pérdidas parciales.

<i>Parámetro</i>	<i>Nivel</i>	<i>Estimación $\beta's$</i>	<i>Exp($\beta's$)= $\gamma's$</i>
Intercepto		9.46	12,789.76
Entidad Federativa	1	0.63	1.881
Entidad Federativa	2	0.49	1.626
Entidad Federativa	3	0.39	1.476

Entidad Federativa	4	0.29	1.335
Entidad Federativa	5	0.17	1.183
Entidad Federativa	6	-	1.000
Sub-marca	1	1.53	4.609
Sub-marca	2	1.23	3.411
Sub-marca	3	1.09	2.966
Sub-marca	4	0.90	2.454
Sub-marca	5	0.65	1.913
Sub-marca	6	0.55	1.727
Sub-marca	7	0.44	1.556
Sub-marca	8	0.34	1.412
Sub-marca	9	0.25	1.287
Sub-marca	10	0.18	1.198
Sub-marca	11	0.12	1.123
Sub-marca	12	- 0.09	0.917
Sub-marca	13	- 0.26	0.771
Sub-marca	14	- 0.18	0.835
Sub-marca	15	- 0.42	0.659
Sub-marca	16	-	1.000
Edad Vehículo	1	- 0.09	0.919
Edad Vehículo	2	- 0.17	0.842
Edad Vehículo	3	- 0.26	0.774
Edad Vehículo	4	- 0.33	0.722
Edad Vehículo	5	- 0.40	0.673
Edad Vehículo	6	- 0.52	0.592
Edad Vehículo	7	- 0.63	0.534
Edad Vehículo	8	- 0.76	0.470
Edad Vehículo	9	- 0.95	0.385
Edad Vehículo	10	- 1.05	0.350
Edad Vehículo	11	-	1.000
Uso Vehículo	1	0.26	1.297
Uso Vehículo	2	0.07	1.077
Uso Vehículo	3	- 0.25	0.778
Uso Vehículo	4	-	1.000
Deducible	1	- 0.11	0.896
Deducible	2	-	1.000

3.1.2.3 Frecuencia de Pérdidas Totales

Tabla 3.10 Información general del modelo de frecuencia de pérdidas totales.

<i>Información del Modelo</i>	
Estructura de Error	Poisson
Función Liga	Log
Variable Dependiente	Número Siniestros pérdida Total
Variable Offset	Logaritmo de Unidades Expuestas
Número de Observaciones	408,686

Tabla 3.11 Tabla de representatividad de las variables incluidas en el modelo de frecuencia de pérdidas totales.

<i>Variable</i>	<i>Grados de Libertad</i>	<i>Chi Cuadrada</i>	<i>P-Value</i>
Entidad Federativa	5	2,308	<.0001
Sub-marca	11	4,296	<.0001
Edad Vehículo	6	564	<.0001
Uso Vehículo	3	75	<.0001
Deducible	2	1,308	<.0001

Tabla 3.12 Tabla de bondad de ajuste para la frecuencia de pérdidas totales.

<i>Criterio</i>	<i>Grados de Libertad</i>	<i>Valor</i>	<i>Valor/Grados de Libertad</i>
Devianza	410,000	88,219	0.216
Devianza escalada	410,000	88,219	0.216
Chi- cuadrada Pearson	410,000	714,788	1.749

Tabla 3.13 Tabla de estimaciones de ajuste para la frecuencia de pérdidas totales.

<i>Parámetro</i>	<i>Nivel</i>	<i>Estimación $\beta's$</i>	<i>Exp($\beta's$)= $\gamma's$</i>
Intercepto	-	5.21	0.55%
Entidad Federativa	1	0.91	2.490
Entidad Federativa	2	0.55	1.737
Entidad Federativa	3	0.34	1.399
Entidad Federativa	4	0.20	1.227
Entidad Federativa	5	-	0.18
Entidad Federativa	6	-	1.000
Sub-marca	1	1.00	2.730
Sub-marca	2	0.77	2.154
Sub-marca	3	0.60	1.815
Sub-marca	4	0.45	1.572
Sub-marca	5	0.28	1.327
Sub-marca	6	0.21	1.235
Sub-marca	7	-	0.13
Sub-marca	8	-	0.28

Sub-marca	9	-	0.43	0.651
Sub-marca	10	-	0.54	0.585
Sub-marca	11	-	0.99	0.373
Sub-marca	12		-	1.000
Edad Vehículo	1		0.15	1.164
Edad Vehículo	2		0.06	1.061
Edad Vehículo	3	-	0.15	0.857
Edad Vehículo	4	-	0.43	0.651
Edad Vehículo	5	-	0.72	0.489
Edad Vehículo	6	-	1.31	0.271
Edad Vehículo	7		-	1.000
Uso Vehículo	1		0.23	1.262
Uso Vehículo	2	-	0.11	0.899
Uso Vehículo	3	-	0.44	0.645
Uso Vehículo	4		-	1.000
Deducible	1	-	0.63	0.533
Deducible	2	-	1.92	0.147
Deducible	3		-	1.000

3.1.2.4 Severidad de Pérdidas Totales

Tabla 3.14 Información general del modelo de severidad de pérdidas totales.

Información del Modelo

Estructura de Error	Gamma
Función Liga	Log
Variable Dependiente	Monto de Pérdidas Totales
Variable Ponderadora	Número Siniestros Totales
Número de Observaciones	21,853

Tabla 3.15 Tabla de representatividad de las variables incluidas en el modelo de severidad de pérdidas totales.

<i>Variable</i>	<i>Grados de Libertad</i>	<i>Chi Cuadrada</i>	<i>P-Value</i>
Entidad Federativa	4	1,433	<.0001
Sub-marca	11	9,595	<.0001
Edad Vehículo	11	13,297	<.0001
Uso Vehículo	1	17	<.0001
Deducible	1	22	<.0001

Tabla 3.16 Tabla de bondad de ajuste para la severidad de pérdidas totales.

<i>Criterio</i>	<i>Grados de Libertad</i>	<i>Valor</i>	<i>Valor/Grados de Libertad</i>
<i>Devianza</i>	22,000	9,506	0.436
<i>Devianza escalada</i>	22,000	23,037	1.056
<i>Chi- cuadrada Pearson</i>	22,000	11,446	0.525

Tabla 3.17 Tabla de estimaciones de ajuste para la severidad de pérdidas totales.

<i>Parámetro</i>	<i>Nivel</i>	<i>Estimación $\beta's$</i>	<i>Exp($\beta's$)= $\gamma's$</i>
<i>Intercepto</i>		11.28	79,522.88
<i>Entidad Federativa</i>	1	0.37	1.444
<i>Entidad Federativa</i>	2	0.30	1.344
<i>Entidad Federativa</i>	3	0.21	1.237
<i>Entidad Federativa</i>	4	0.16	1.177
<i>Entidad Federativa</i>	5	-	1.000
<i>Sub-marca</i>	1	1.13	3.095
<i>Sub-marca</i>	2	0.74	2.096
<i>Sub-marca</i>	3	0.55	1.739
<i>Sub-marca</i>	4	0.49	1.624
<i>Sub-marca</i>	5	0.40	1.488
<i>Sub-marca</i>	6	0.25	1.278
<i>Sub-marca</i>	7	- 0.06	0.940
<i>Sub-marca</i>	8	- 0.18	0.833
<i>Sub-marca</i>	9	- 0.38	0.683
<i>Sub-marca</i>	10	- 0.48	0.619
<i>Sub-marca</i>	11	- 0.67	0.511
<i>Sub-marca</i>	12	-	1.000
<i>Edad Vehículo</i>	1	0.24	1.268
<i>Edad Vehículo</i>	2	- 0.22	0.800
<i>Edad Vehículo</i>	3	- 0.34	0.713
<i>Edad Vehículo</i>	4	- 0.52	0.597
<i>Edad Vehículo</i>	5	- 0.60	0.551
<i>Edad Vehículo</i>	6	- 0.71	0.490
<i>Edad Vehículo</i>	7	- 0.86	0.425
<i>Edad Vehículo</i>	8	- 1.02	0.359
<i>Edad Vehículo</i>	9	- 1.28	0.277
<i>Edad Vehículo</i>	10	- 1.40	0.247
<i>Edad Vehículo</i>	11	- 1.68	0.186
<i>Edad Vehículo</i>	12	-	1.000
<i>Uso Vehículo</i>	1	0.09	1.090

Uso Vehículo	2	-	1.000
Deducible	1	- 0.07	0.930
Deducible	2	-	1.000

3.1.2.5 Frecuencia Rotura de Cristales

Tabla 3.18 Información general del modelo de frecuencia rotura de cristales.

<i>Información del Modelo</i>	
Estructura de Error	Poisson
Función Liga	Log
Variable Dependiente	Número Siniestros de rotura de Cristales
Variable Offset	Logaritmo de Unidades Expuestas)
Número de Observaciones	408,686

Tabla 3.19 Tabla de representatividad de las variables incluidas en el modelo de frecuencia rotura de cristales.

<i>Variable</i>	<i>Grados de Libertad</i>	<i>Chi Cuadrada</i>	<i>P-Value</i>
Entidad Federativa	7	24,179	<.0001
Sub-marca	10	16,806	<.0001
Edad Vehículo	5	2,910	<.0001
Uso Vehículo	2	2,555	<.0001

Tabla 3.20 Tabla de bondad de ajuste para la frecuencia rotura de cristales.

<i>Criterio</i>	<i>Grados de Libertad</i>	<i>Valor</i>	<i>Valor/Grados de Libertad</i>
Devianza	410,000	215,778	0.528
Devianza escalada	410,000	215,778	0.528
Chi- cuadrada Pearson	410,000	595,705	1.458

Tabla 3.21 Tabla de estimaciones de ajuste para la frecuencia rotura de cristales.

<i>Parámetro</i>	<i>Nivel</i>	<i>Estimación $\beta's$</i>	<i>Exp($\beta's$)= $\gamma's$</i>
Intercepto	-	3.56	2.85%
Entidad Federativa	1	1.15	3.156
Entidad Federativa	2	0.87	2.376
Entidad Federativa	3	0.67	1.948
Entidad Federativa	4	0.56	1.759
Entidad Federativa	5	0.40	1.492
Entidad Federativa	6	0.22	1.247
Entidad Federativa	7	0.17	1.189
Entidad Federativa	8	-	1.000
Sub-marca	1	0.43	1.535
Sub-marca	2	0.25	1.284
Sub-marca	3	- 0.17	0.843

Sub-marca	4	- 0.33	0.718
Sub-marca	5	- 0.46	0.634
Sub-marca	6	- 0.53	0.588
Sub-marca	7	- 0.74	0.479
Sub-marca	8	- 1.08	0.341
Sub-marca	9	- 1.32	0.267
Sub-marca	10	- 2.49	0.083
Sub-marca	11	-	1.000
Edad Vehículo	1	0.09	1.090
Edad Vehículo	2	- 0.19	0.824
Edad Vehículo	3	- 0.46	0.631
Edad Vehículo	4	- 0.76	0.467
Edad Vehículo	5	- 1.35	0.260
Edad Vehículo	6	-	1.000
Uso Vehículo	1	0.63	1.882
Uso Vehículo	2	0.27	1.304
Uso Vehículo	3	-	1.000

3.1.2.6 Severidad Rotura de Cristales

Tabla 3.22 Información general del modelo de severidad rotura de cristales.

Información del Modelo

Estructura de Error	Gamma
Función Liga	Log
Variable Dependiente	Monto de Rotura de Cristales
Variable Ponderadora	Número Siniestros Cristales
Número de Observaciones	62,288

Tabla 3.23 Tabla de representatividad de las variables incluidas en el modelo de severidad rotura de cristales.

<i>Variable</i>	<i>Grados de Libertad</i>	<i>Chi Cuadrada</i>	<i>P-Value</i>
Entidad Federativa	3	1,123	<.0001
Sub-marca	13	52,681	<.0001
Edad Vehículo	6	7,353	<.0001
Uso Vehículo	1	342	<.0001

Tabla 3.24 Tabla de bondad de ajuste para la severidad rotura de cristales.

<i>Criterio</i>	<i>Grados de Libertad</i>	<i>Valor</i>	<i>Valor/Grados de Libertad</i>
Devianza	62,000	39,052	0.627
Devianza escalada	62,000	66,387	1.066
Chi- cuadrada Pearson	62,000	77,134	1.239

Tabla 3.25 Tabla de estimaciones de ajuste para la severidad rotura de cristales.

<i>Parámetro</i>	<i>Nivel</i>	<i>Estimación $\beta's$</i>	<i>Exp($\beta's$)= $\gamma's$</i>
Intercepto		7.35	1,548.9
Entidad Federativa	1	0.14	1.153
Entidad Federativa	2 -	0.06	0.938
Entidad Federativa	3 -	0.14	0.874
Entidad Federativa	4	-	1.000
Sub-marca	1	1.85	6.386
Sub-marca	2	1.64	5.147
Sub-marca	3	1.48	4.379
Sub-marca	4	1.36	3.906
Sub-marca	5	1.21	3.354
Sub-marca	6	1.05	2.859
Sub-marca	7	0.96	2.612
Sub-marca	8	0.83	2.289
Sub-marca	9	0.73	2.078
Sub-marca	10	0.56	1.747
Sub-marca	11	0.34	1.406
Sub-marca	12	0.27	1.315
Sub-marca	13 -	0.45	0.640
Sub-marca	14	-	1.000
Edad Vehículo	1 -	0.17	0.844
Edad Vehículo	2 -	0.25	0.776
Edad Vehículo	3 -	0.34	0.711
Edad Vehículo	4 -	0.41	0.663
Edad Vehículo	5 -	0.52	0.596
Edad Vehículo	6 -	0.61	0.545
Edad Vehículo	7	-	1.000
Uso Vehículo	1 -	0.19	0.826
Uso Vehículo	2	-	1.000

Es interesante notar que los modelos de severidad son aquellos con presencia casi nula de sobredispersión, mientras que los modelos de frecuencia son aquellos en donde la sobredispersión o subdispersión es un factor latente. El modelo con un mayor número de parámetros no redundantes es el de severidad de pérdidas parciales, con alrededor de 34, mientras que el modelo de severidad de rotura de cristales es el que posee un menor número de parámetros no redundantes, 23 para ser exacto. Por otro lado se confirman todas las conclusiones obtenidas durante el análisis de la sección 2.5.

3.2 Pruebas de Bondad de Ajuste

Una vez ajustados los modelos es necesario hacer pruebas más exhaustivas para tener claro el nivel de ajuste. Esta sección se enfocará en los cuatro principales métodos que brindan al actuario los elementos necesarios para concluir el grado de confianza que deberá depositar en las predicciones.

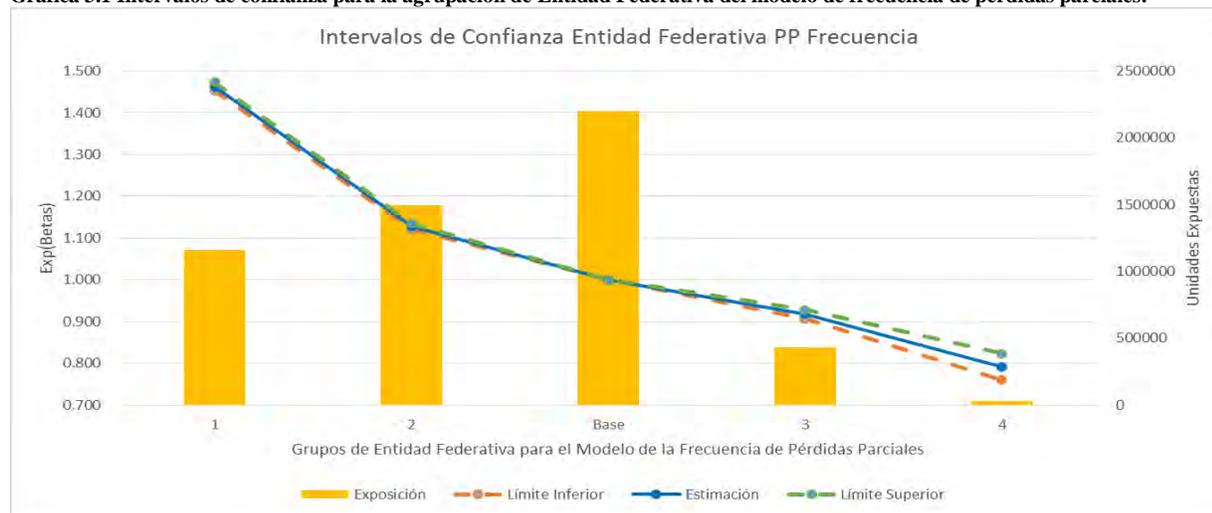
3.2.1 Intervalos de Confianza

3.2.1.1 Frecuencia Pérdidas Parciales

Tabla 3.26 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de frecuencia de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación)= $\gamma's$	Límite Inferior	Exposición
1	1.473	1.463	1.453	1,162,350
2	1.135	1.128	1.120	1,493,639
Base	1.000	1.000	1.000	2,200,937
3	0.928	0.917	0.907	430,805
4	0.823	0.791	0.759	34,087

Gráfica 3.1 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de frecuencia de pérdidas parciales.



El eje “x” de la gráfica 3.1 corresponde a las distintas zonas geográficas de la variable entidad federativa del modelo de frecuencia de pérdidas parciales, las barras amarillas representan la cantidad de unidades expuestas que posee cada zona geográfica y está medida por el eje derecho “y”. El eje izquierdo “y” de la gráfica mide la proporción en que aumenta o descende la frecuencia de cada zona geográfica en comparación con el nivel base, que en la mayoría de los casos es el nivel con mayor cantidad de unidades expuestas (no es una regla). Las líneas punteadas verde y naranja son los intervalos de confianza superior e inferior, la línea continua azul representa a las $\gamma's$. Este diseño de gráfica aplica para toda la sección 3.2.1, en donde los resultados obtenidos al

realizar este test fueron satisfactorios, con intervalos de confianza más amplios en los casos de niveles con poco peso (unidades expuestas o número de siniestros, según corresponda).

Tabla 3.27 Intervalos de confianza para la agrupación de Sub-marca del modelo de frecuencia de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
1	1.486	1.455	1.424	56,600
2	1.261	1.250	1.240	598,555
3	1.129	1.121	1.113	1,187,668
Base	1.000	1.000	1.000	1,708,850
4	0.853	0.845	0.837	656,979
5	0.785	0.776	0.768	570,577
6	0.714	0.703	0.692	282,796
7	0.648	0.634	0.621	157,184
8	0.516	0.502	0.487	102,610

Gráfica 3.2 Intervalos de confianza para la agrupación de Sub-marca del modelo de frecuencia de pérdidas parciales.

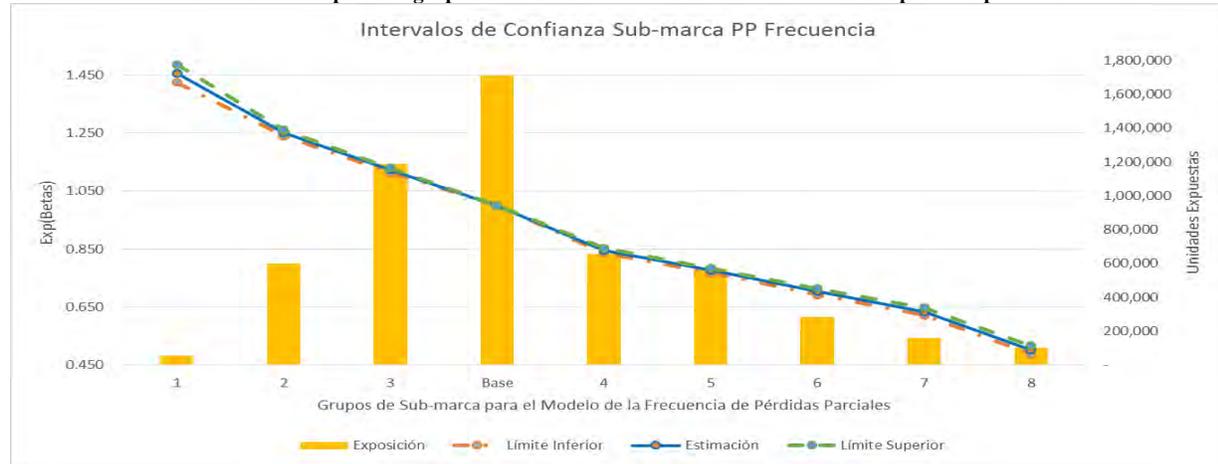


Tabla 3.28 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de frecuencia de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
Base	1.000	1.000	1.000	2,111,769
1	0.875	0.869	0.863	1,568,631
2	0.864	0.857	0.851	1,200,831
3	0.773	0.762	0.751	260,399
4	0.689	0.668	0.648	66,062
5	0.614	0.590	0.567	43,624
6	0.516	0.494	0.473	46,035
7	0.340	0.315	0.291	24,467

Gráfica 3.3 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de frecuencia de pérdidas parciales.

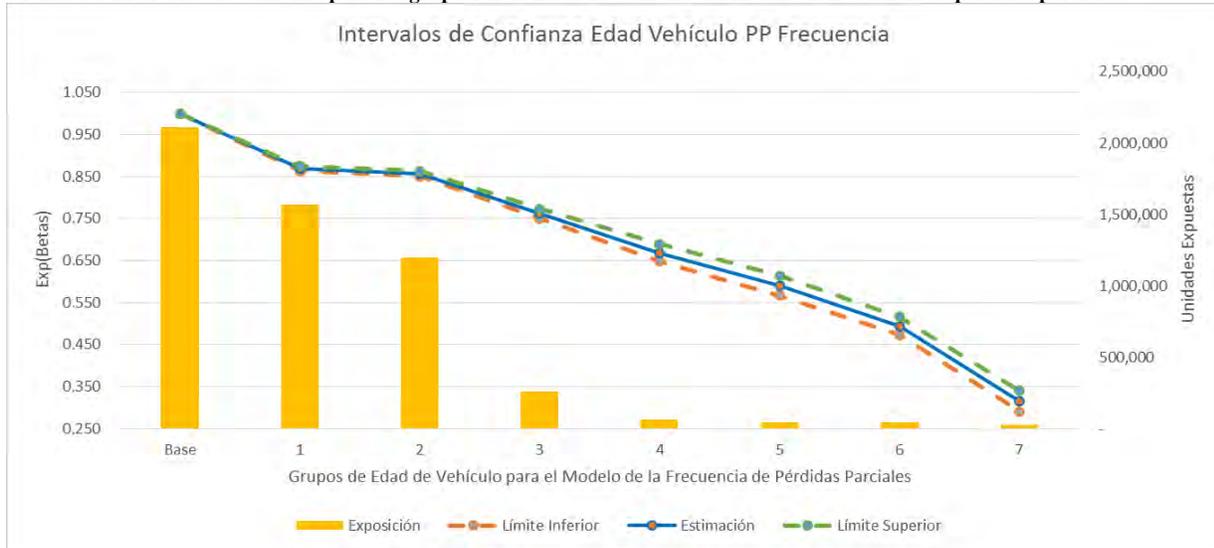


Tabla 3.29 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de frecuencia de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
1	2.372	2.327	2.284	85,262
Base	1.000	1.000	1.000	5,006,917
2	0.957	0.938	0.920	145,473
3	0.401	0.386	0.372	84,166

Gráfica 3.4 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de frecuencia de pérdidas parciales.

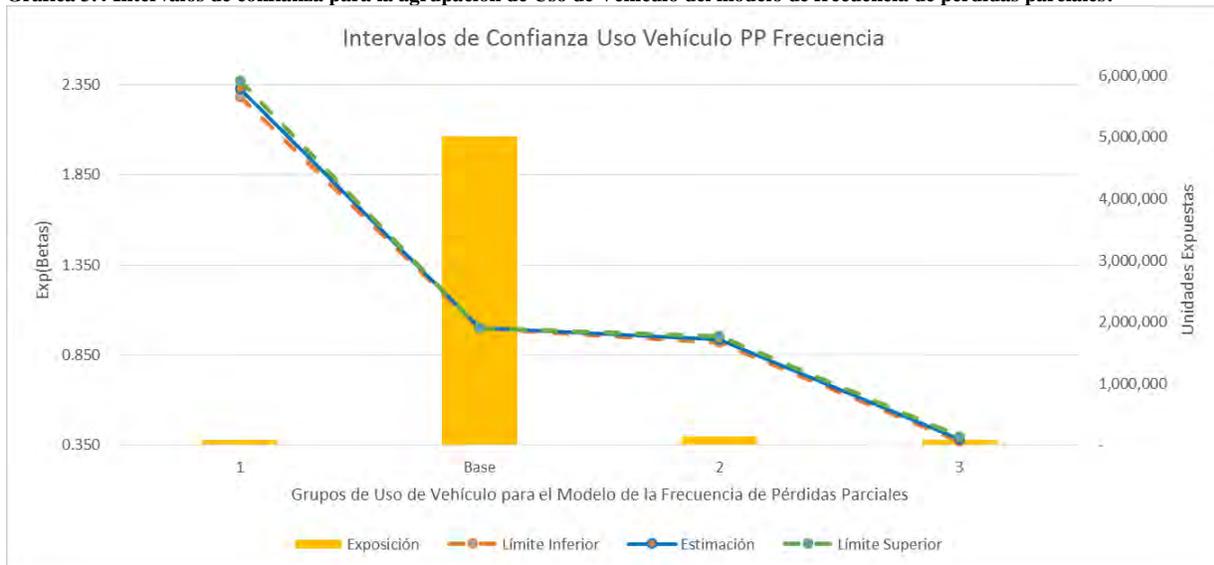
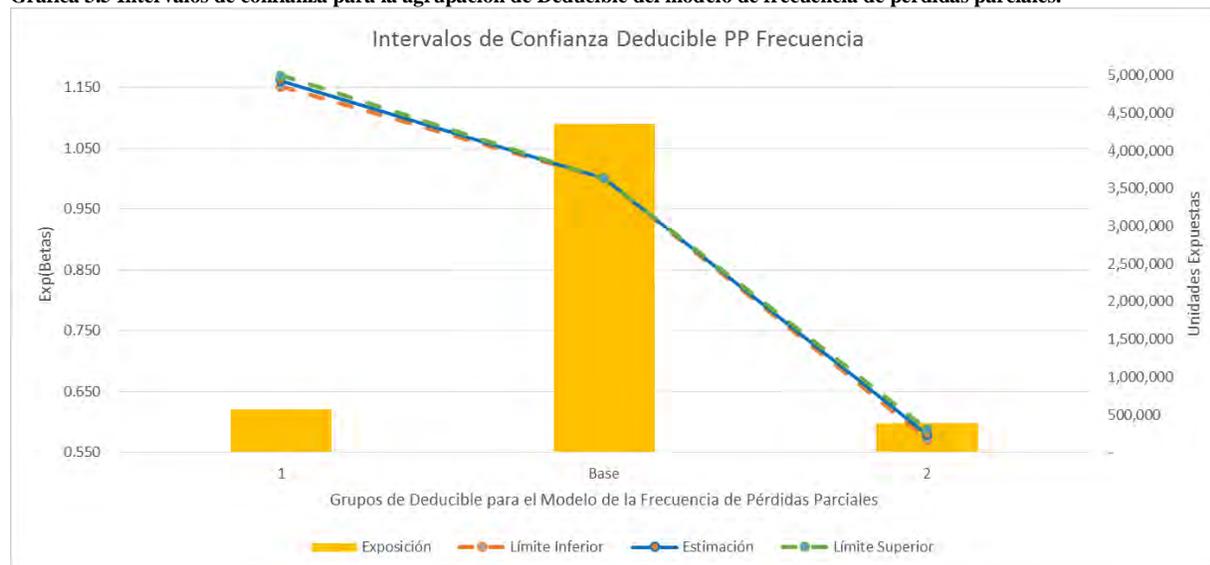


Tabla 3.30 Intervalos de confianza para la agrupación de Deducible del modelo de frecuencia de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
1	1.170	1.161	1.151	571,979
Base	1.000	1.000	1.000	4,359,896
2	0.587	0.579	0.571	389,944

Gráfica 3.5 Intervalos de confianza para la agrupación de Deducible del modelo de frecuencia de pérdidas parciales.



3.2.1.2 Severidad Pérdidas Parciales

Tabla 3.31 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de severidad de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	1.986	1.881	1.786	6,121
2	1.666	1.626	1.588	27,303
3	1.507	1.476	1.445	31,644
4	1.353	1.335	1.319	88,467
5	1.196	1.183	1.170	122,473
Base	1.000	1.000	1.000	260,535

Gráfica 3.6 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de severidad de pérdidas parciales.

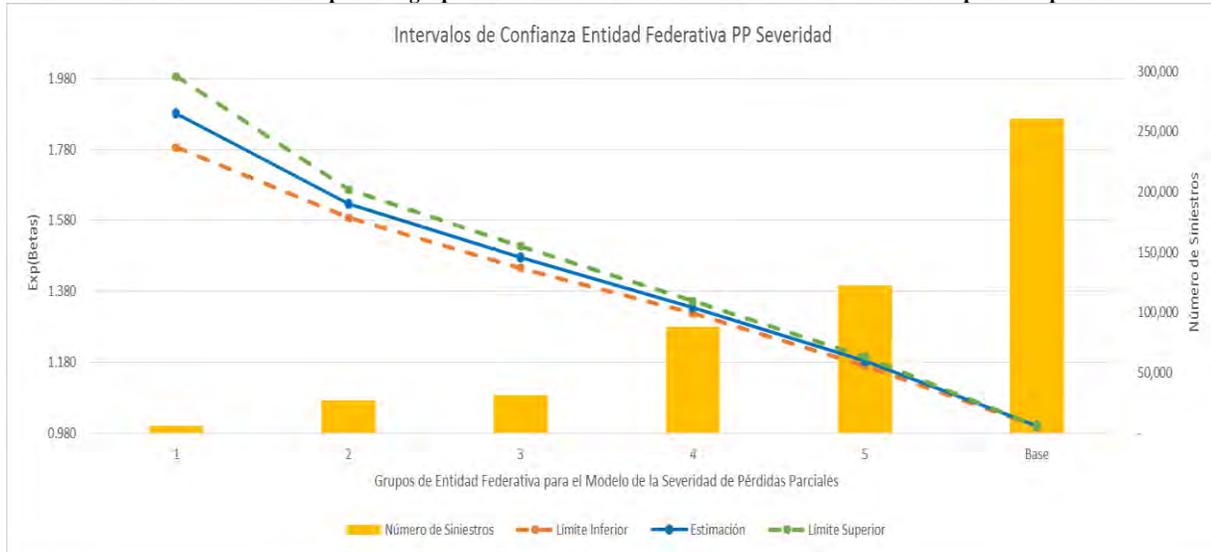


Tabla 3.32 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	5.395	4.609	4.023	2,285
2	4.187	3.411	2.878	964
3	3.415	2.966	2.621	1,620
4	2.567	2.454	2.349	12,617
5	2.022	1.913	1.815	6,914
6	1.788	1.727	1.669	16,147
7	1.600	1.556	1.514	22,996
8	1.459	1.412	1.367	14,774
9	1.314	1.287	1.261	37,868
10	1.218	1.198	1.178	63,500
11	1.139	1.123	1.108	98,787
Base	1.000	1.000	1.000	108,846
12	0.931	0.917	0.904	73,059
13	0.782	0.771	0.760	59,573
14	0.862	0.835	0.810	6,110
15	0.675	0.659	0.644	10,483

Gráfica 3.7 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de pérdidas parciales.

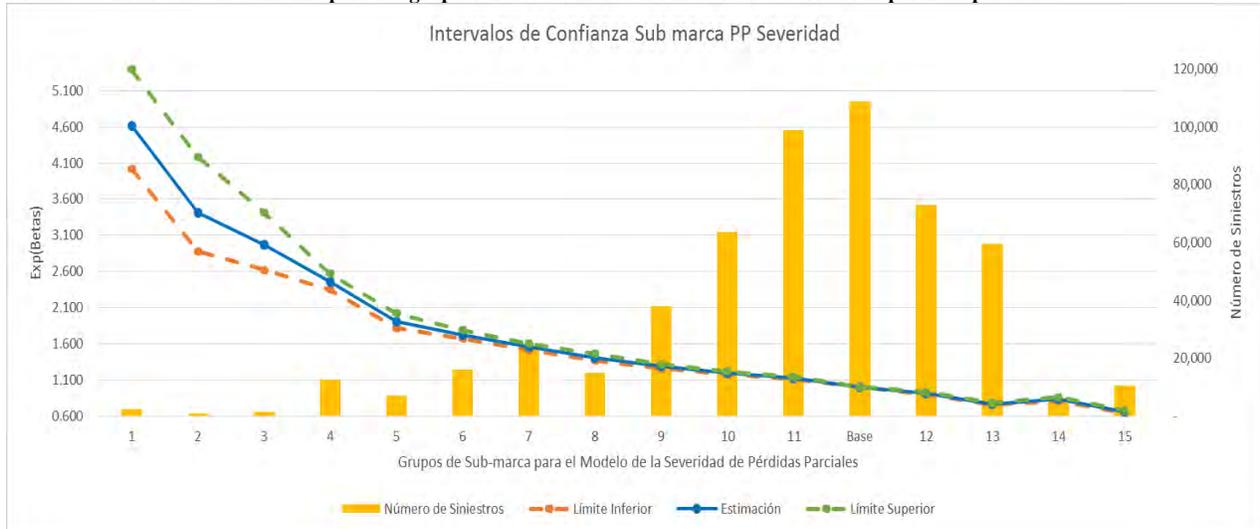


Tabla 3.33 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de severidad de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
Base	1.000	1.000	1.000	142,928
1	0.932	0.919	0.905	100,852
2	0.856	0.842	0.829	66,503
3	0.786	0.774	0.763	73,798
4	0.737	0.722	0.708	29,044
5	0.684	0.673	0.662	50,005
6	0.603	0.592	0.582	31,288
7	0.545	0.534	0.524	22,130
8	0.480	0.470	0.460	15,499
9	0.400	0.385	0.371	3,177
10	0.370	0.350	0.331	1,319

Gráfica 3.8 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de severidad de pérdidas parciales.

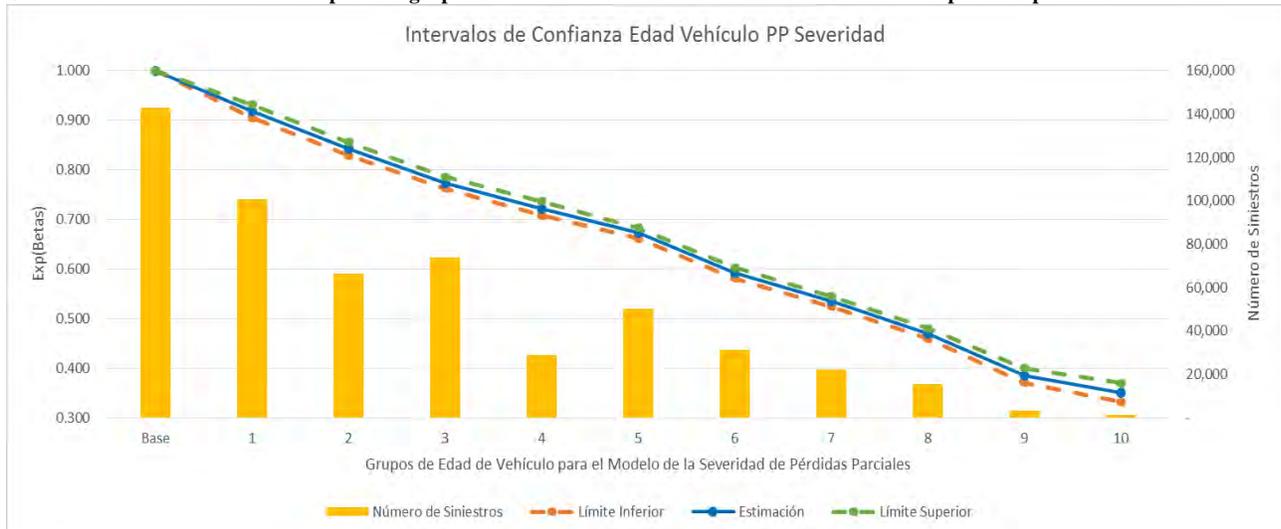


Tabla 3.34 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de severidad de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	1.351	1.297	1.247	10,986
2	1.157	1.077	1.008	2,946
Base	1.000	1.000	1.000	507,762
3	0.798	0.778	0.758	14,849

Gráfica 3.9 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de severidad de pérdidas parciales.

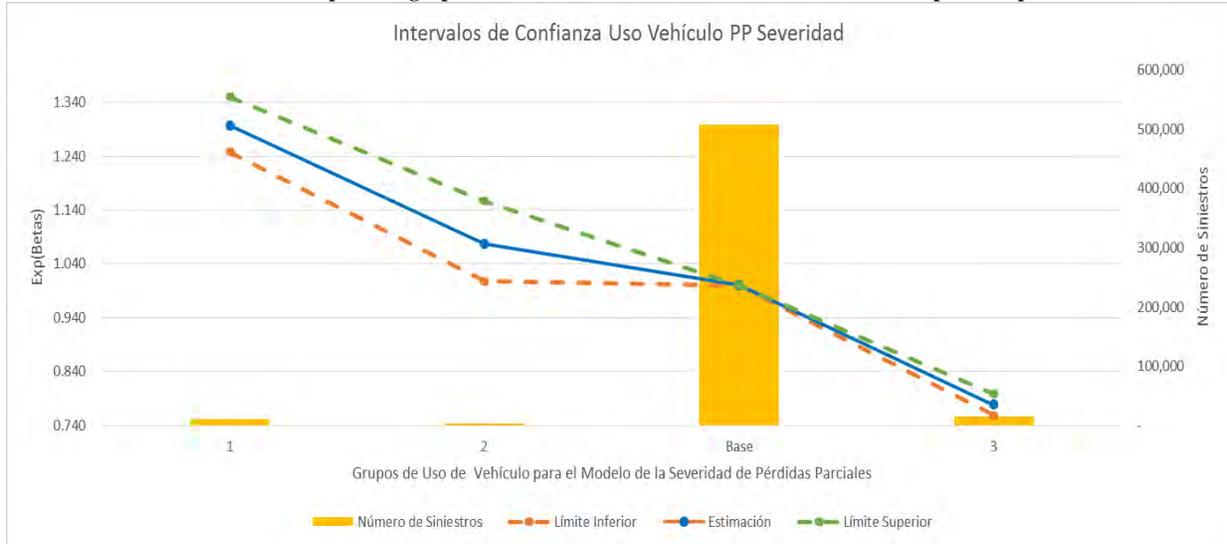
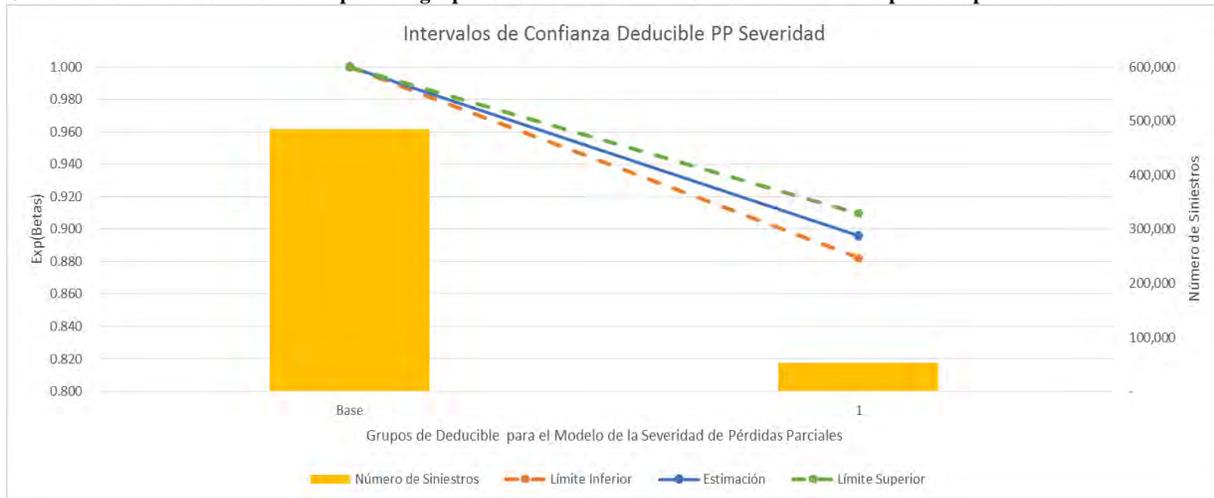


Tabla 3.35 Intervalos de confianza para la agrupación de Deducible del modelo de severidad de pérdidas parciales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
Base	1.000	1.000	1.000	484,103
1	0.910	0.896	0.882	52,440

Gráfica 3.10 Intervalos de confianza para la agrupación de Deducible del modelo de severidad de pérdidas parciales.



3.2.1.3 Frecuencia Pérdidas Totales

Tabla 3.36 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de frecuencia de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
1	2.591	2.490	2.392	229,178
2	1.816	1.737	1.660	236,071
3	1.436	1.399	1.363	1,594,551
4	1.261	1.227	1.194	1,450,000
Base	1.000	1.000	1.000	1,641,933
5	0.893	0.833	0.778	170,085

Gráfica 3.11 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de frecuencia de pérdidas totales.

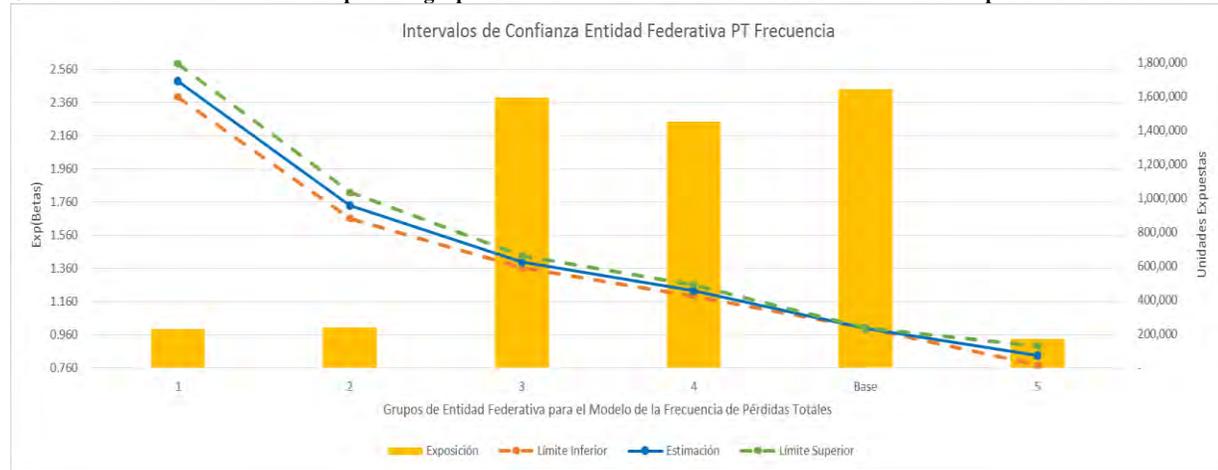


Tabla 3.37 Intervalos de confianza para la agrupación de Sub-marca del modelo de frecuencia de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
1	3.017	2.730	2.470	20,367
2	2.385	2.154	1.944	25,599
3	1.897	1.815	1.736	217,684
4	1.627	1.572	1.519	557,782
5	1.375	1.327	1.280	552,650
6	1.274	1.235	1.197	1,105,259
Base	1.000	1.000	1.000	1,010,129
7	0.911	0.875	0.840	142,150
8	0.794	0.757	0.722	568,461
9	0.686	0.651	0.618	424,977
10	0.622	0.585	0.550	386,242
11	0.414	0.373	0.336	310,518

Gráfica 3.12 Intervalos de confianza para la agrupación de Sub-marca del modelo de frecuencia de pérdidas totales.

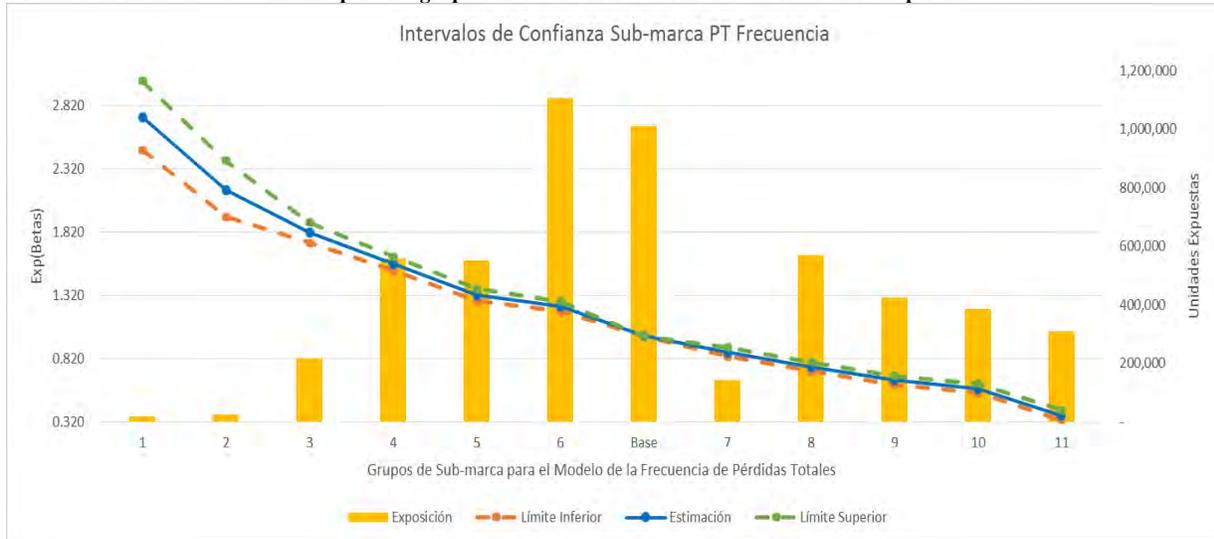


Tabla 3.38 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de frecuencia de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
1	1.190	1.164	1.138	1,937,015
2	1.096	1.061	1.028	666,868
Base	1.000	1.000	1.000	2,042,130
3	0.891	0.857	0.824	531,264
4	0.719	0.651	0.589	99,568
5	0.646	0.489	0.371	20,506
6	0.385	0.271	0.190	24,467

Gráfica 3.13 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de frecuencia de pérdidas totales.

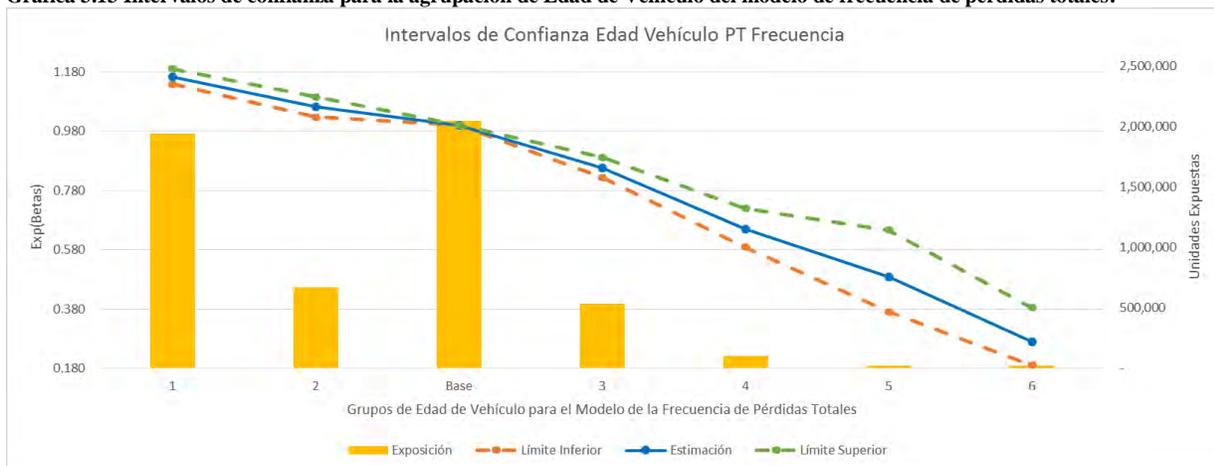


Tabla 3.39 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de frecuencia de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
1	1.544	1.262	1.032	20,202
Base	1.000	1.000	1.000	5,123,741
2	0.976	0.899	0.827	108,391
3	0.724	0.645	0.575	69,485

Gráfica 3.14 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de frecuencia de pérdidas totales.

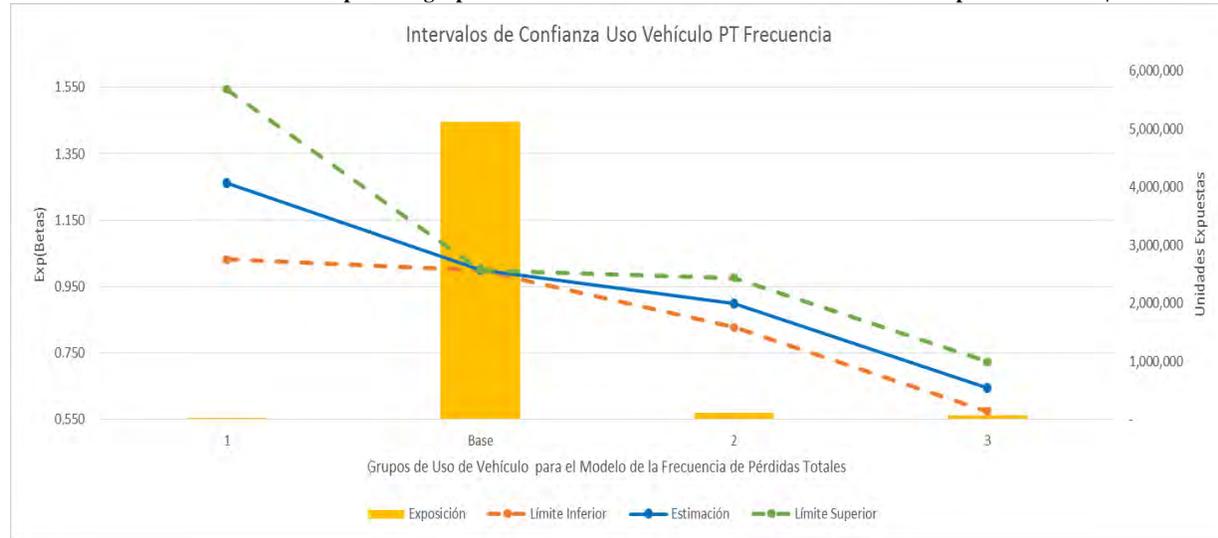
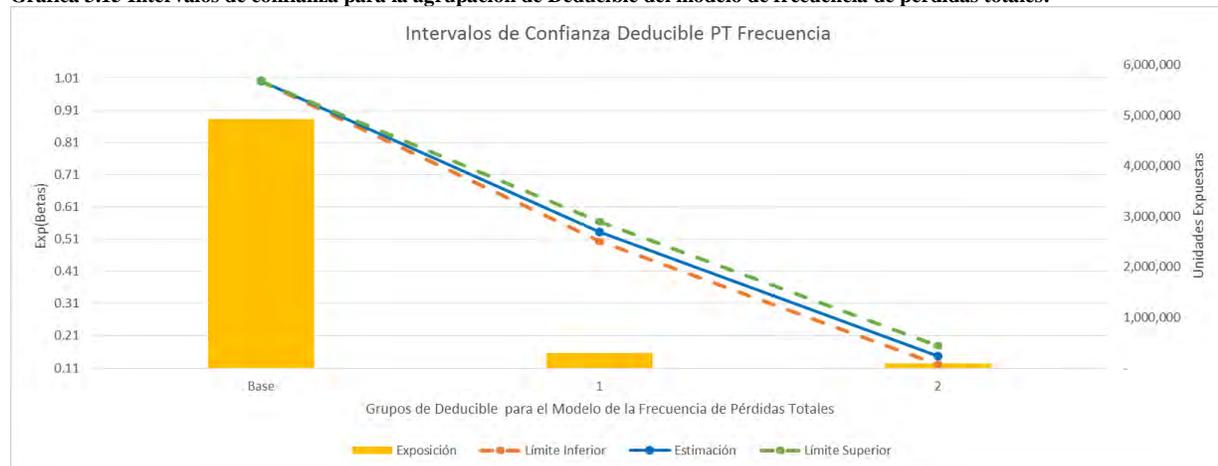


Tabla 3.40 Intervalos de confianza para la agrupación de Deducible del modelo de frecuencia de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Exposición
Base	1.00	1.00	1.00	4,931,874
1	0.564	0.533	0.503	301,023
2	0.179	0.147	0.121	88,921

Gráfica 3.15 Intervalos de confianza para la agrupación de Deducible del modelo de frecuencia de pérdidas totales.



3.2.1.4 Severidad Pérdidas Totales

Tabla 3.41 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de severidad de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	1.477	1.444	1.412	4,449
2	1.369	1.344	1.319	8,040
3	1.258	1.237	1.216	10,585
4	1.201	1.177	1.154	6,237
Base	1.000	1.000	1.000	10,907

Gráfica 3.16 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de severidad de pérdidas totales.

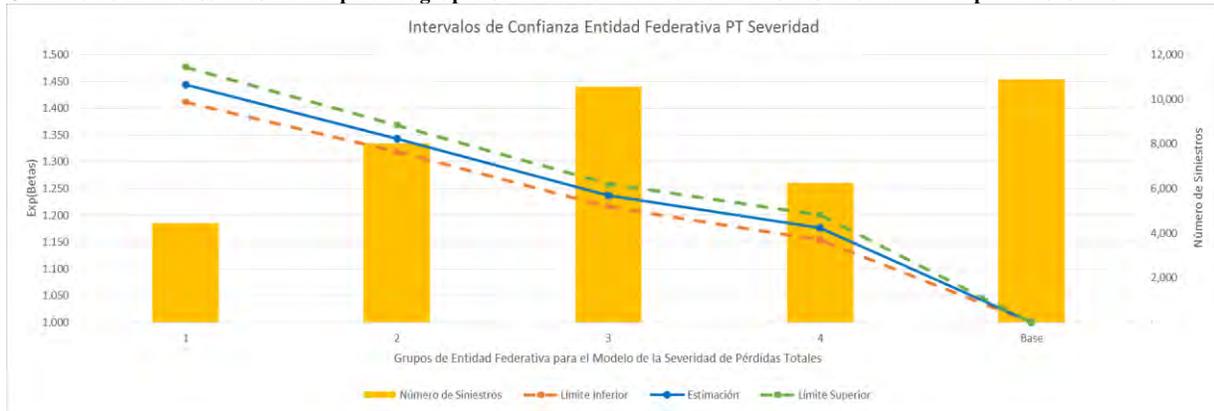


Tabla 3.42 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	3.225	3.095	2.970	1,013
2	2.175	2.096	2.021	1,313
3	1.808	1.739	1.673	1,176
4	1.674	1.624	1.576	2,045
5	1.524	1.488	1.453	3,719
6	1.305	1.278	1.253	5,811
Base	1.000	1.000	1.000	11,740
7	0.963	0.940	0.917	3,437
8	0.852	0.833	0.813	4,292
9	0.702	0.683	0.664	2,525
10	0.640	0.619	0.598	1,663
11	0.530	0.511	0.493	1,484

Gráfica 3.17 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de pérdidas totales.

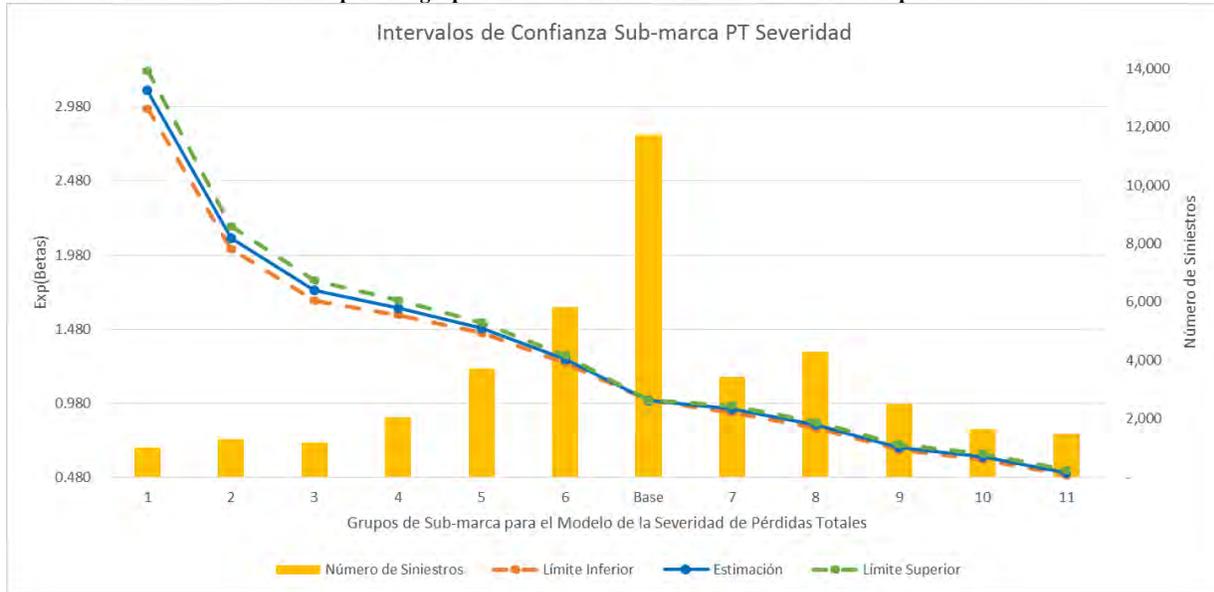


Tabla 3.43 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de severidad de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	1.303	1.268	1.235	2,967
Base	1.000	1.000	1.000	8,184
2	0.823	0.800	0.778	2,835
3	0.729	0.713	0.697	4,982
4	0.613	0.597	0.581	3,200
5	0.566	0.551	0.536	3,194
6	0.503	0.490	0.476	2,840
7	0.437	0.425	0.413	2,908
8	0.368	0.359	0.350	3,767
9	0.285	0.277	0.269	2,686
10	0.255	0.247	0.239	1,893
11	0.196	0.186	0.178	762

Gráfica 3.18 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de severidad de pérdidas totales.

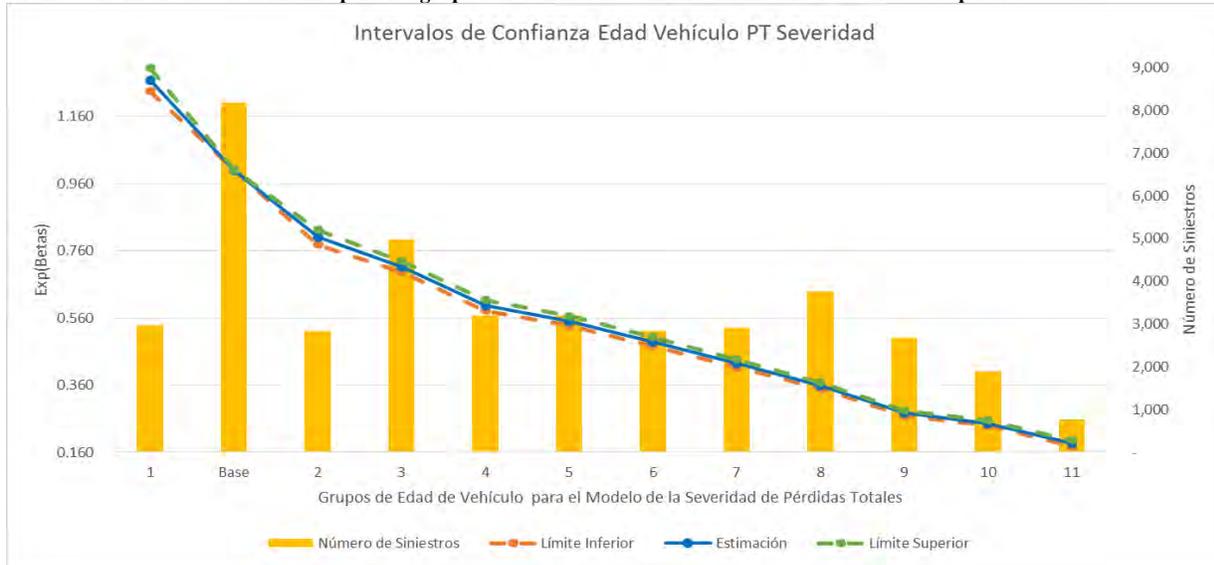


Tabla 3.44 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de severidad de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	1.136	1.090	1.046	979
Base	1.000	1.000	1.000	39,239

Gráfica 3.19 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de severidad de pérdidas totales.

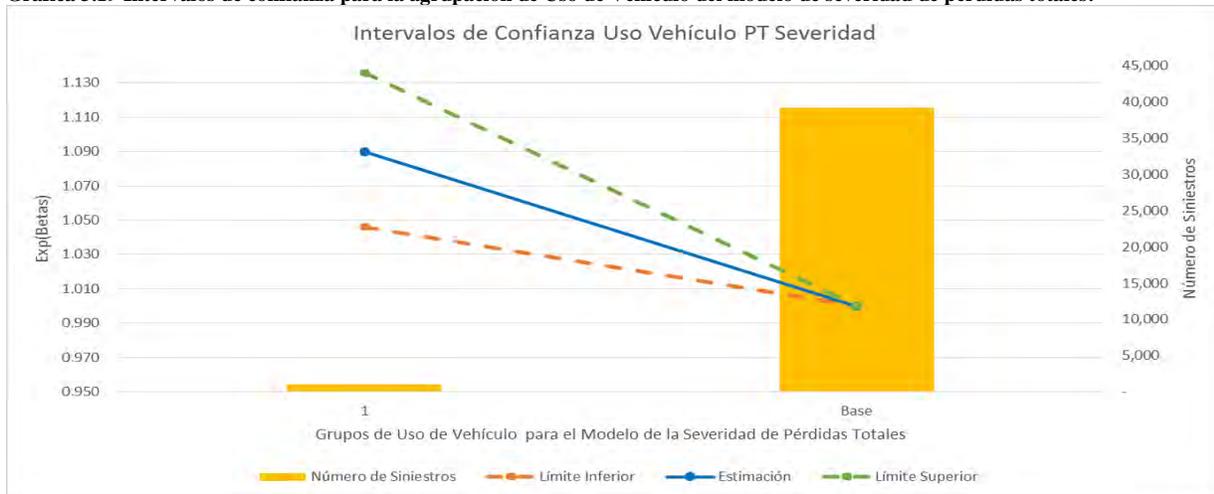
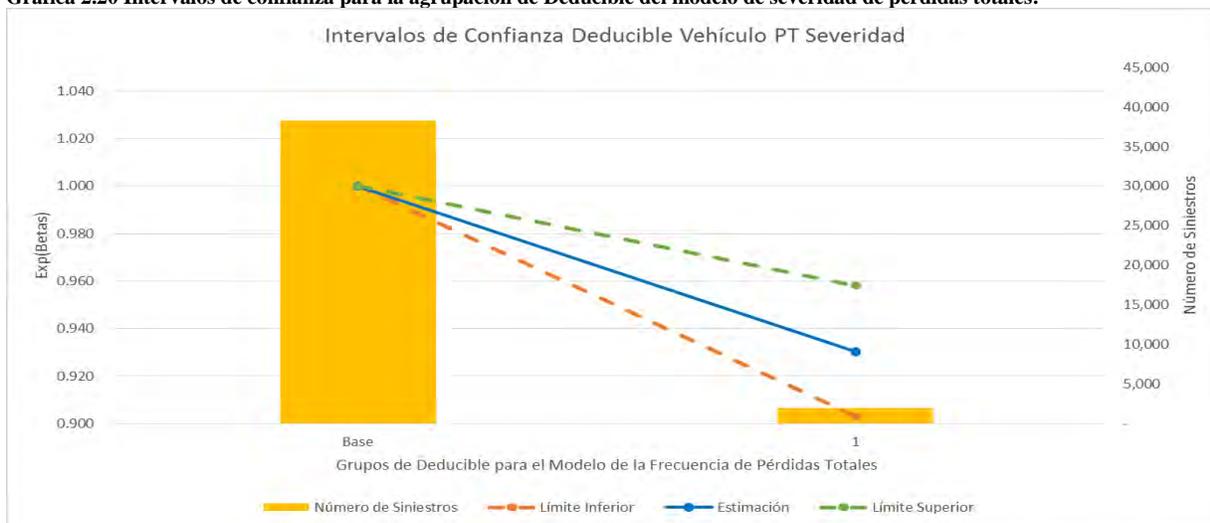


Tabla 3.45 Intervalos de confianza para la agrupación de Deducible del modelo de severidad de pérdidas totales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
Base	1.000	1.000	1.000	38,268
1	0.958	0.930	0.903	1,950

Gráfica 2.20 Intervalos de confianza para la agrupación de Deducible del modelo de severidad de pérdidas totales.



3.2.1.5 Frecuencia Rotura de Cristales

Tabla 3.46 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de frecuencia de rotura de cristales.

Nivel	Límite Superior	Exp(Estimación) γ 's	Límite Inferior	Exposición
1	3.212	3.156	3.101	309,105
2	2.416	2.376	2.338	567,202
3	1.979	1.948	1.918	828,956
4	1.790	1.759	1.728	587,284
5	1.523	1.492	1.463	432,237
6	1.269	1.247	1.225	816,146
7	1.214	1.189	1.165	537,111
Base	1.000	1.000	1.000	1,243,778

Gráfica 3.21 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de frecuencia de rotura de cristales.

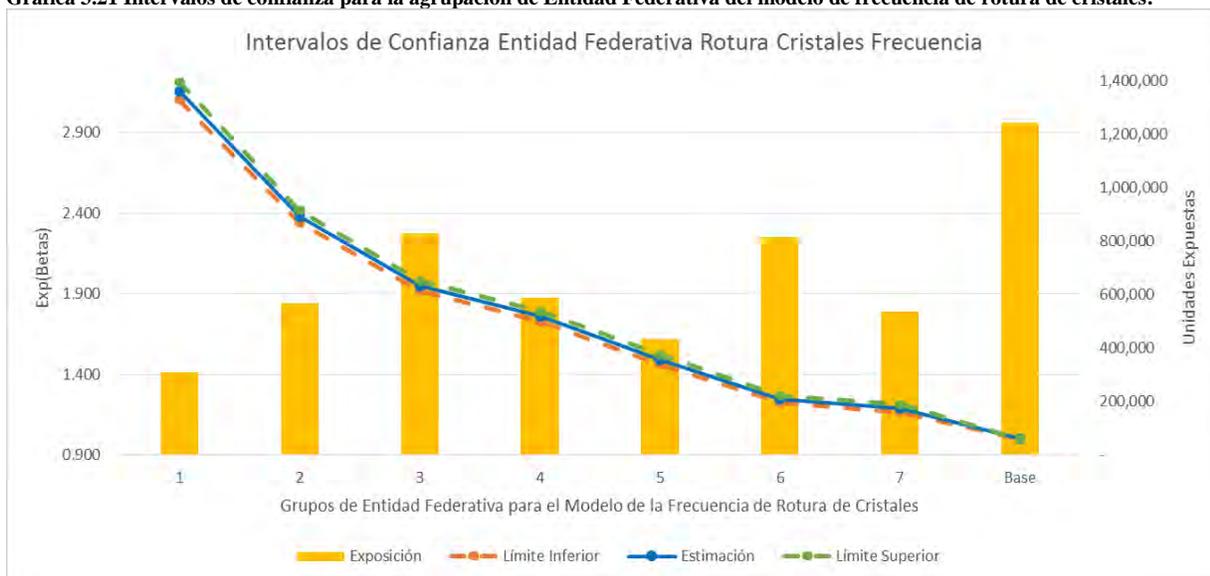


Tabla 3.47 Intervalos de confianza para la agrupación de Sub-marca del modelo de frecuencia de rotura de cristales.

Nivel	Límite Superior	Exp(Estimación) γ 's	Límite Inferior	Exposición
1	1.647	1.535	1.431	7,988
2	1.309	1.284	1.260	187,952
Base	1.000	1.000	1.000	1,413,379
3	0.854	0.843	0.832	862,808
4	0.728	0.718	0.709	1,155,333
5	0.644	0.634	0.624	704,107
6	0.602	0.588	0.575	312,599
7	0.491	0.479	0.467	341,708
8	0.357	0.341	0.325	119,188
9	0.279	0.267	0.255	180,824
10	0.100	0.083	0.069	35,932

Gráfica 3.22 Intervalos de confianza para la agrupación de Sub-marca del modelo de frecuencia de rotura de cristales.

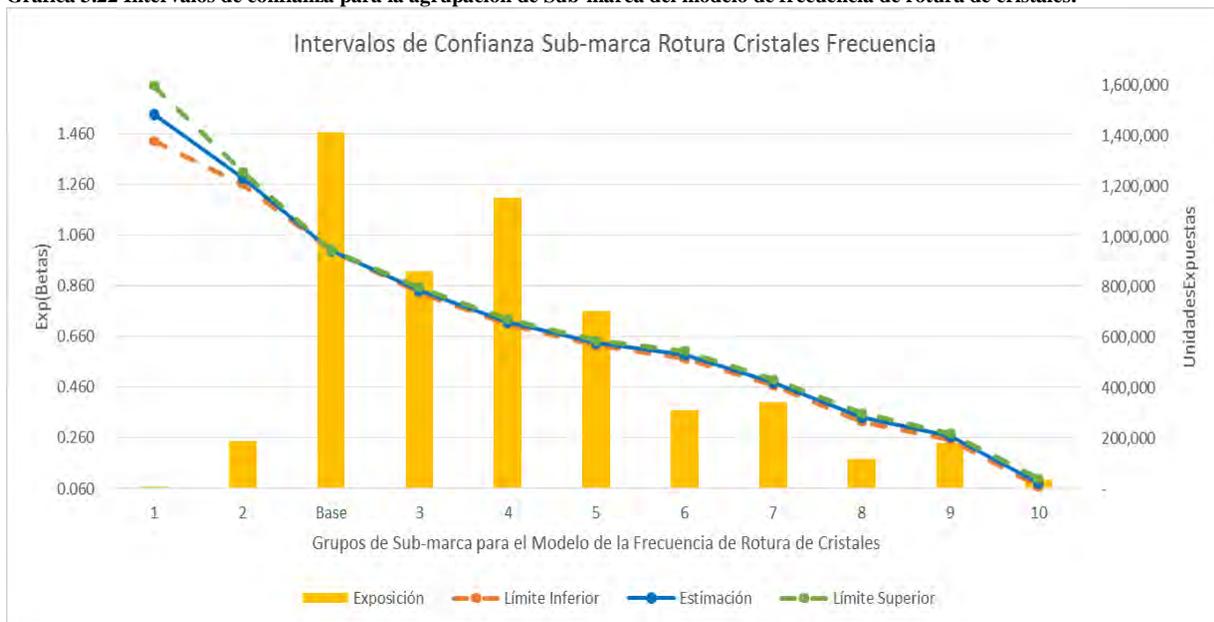


Tabla 3.48 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de frecuencia de rotura de cristales.

Nivel	Límite Superior	Exp(Estimación) γ 's	Límite Inferior	Exposición
1	1.101	1.090	1.080	2,054,392
Base	1.000	1.000	1.000	2,482,254
2	0.837	0.824	0.811	618,486
3	0.656	0.631	0.607	112,366
4	0.513	0.467	0.426	29,853
5	0.297	0.260	0.227	24,467

Gráfica 3.23 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de frecuencia de rotura de cristales.

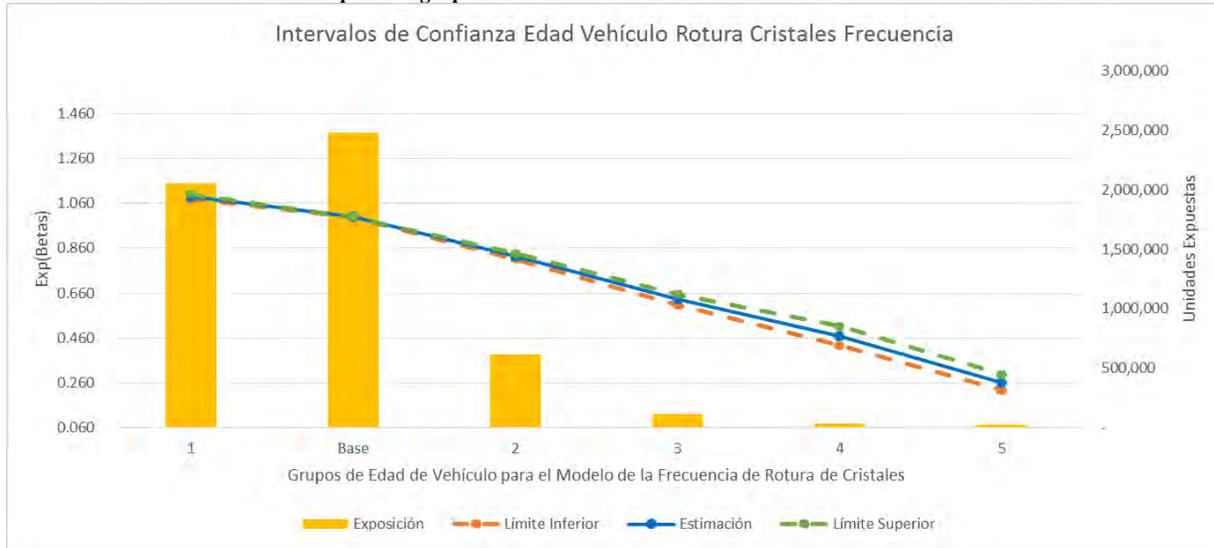
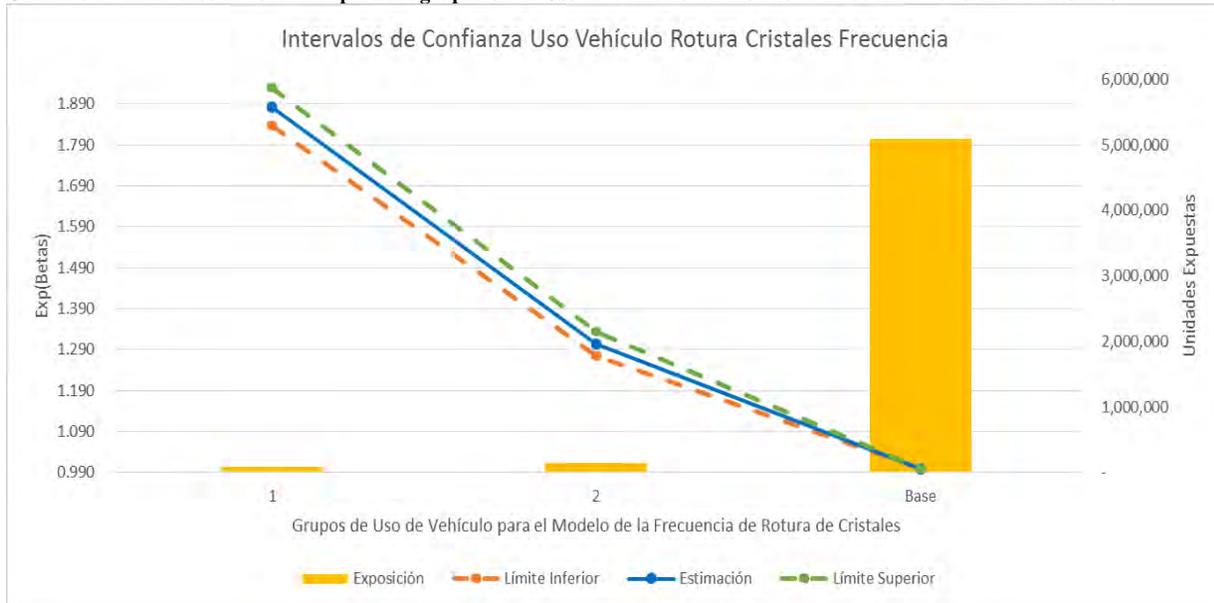


Tabla 3.49 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de frecuencia de rotura de cristales.

Nivel	Límite Superior	Exp(Estimación) γ 's	Límite Inferior	Exposición
1	1.929	1.882	1.837	89,803
2	1.334	1.304	1.275	140,944
Base	1.000	1.000	1.000	5,091,071

Gráfica 3.24 Intervalos de confianza para la agrupación de Uso de Vehículo del modelo de frecuencia de rotura de cristales



3.2.1.6 Severidad Rotura de Cristales

Tabla 3.50 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de severidad de rotura de cristales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	1.182	1.153	1.124	3,764
Base	1.000	1.000	1.000	80,552
2	0.946	0.938	0.931	67,210
3	0.881	0.874	0.866	43,339

Gráfica 3.25 Intervalos de confianza para la agrupación de Entidad Federativa del modelo de severidad de rotura de cristales.

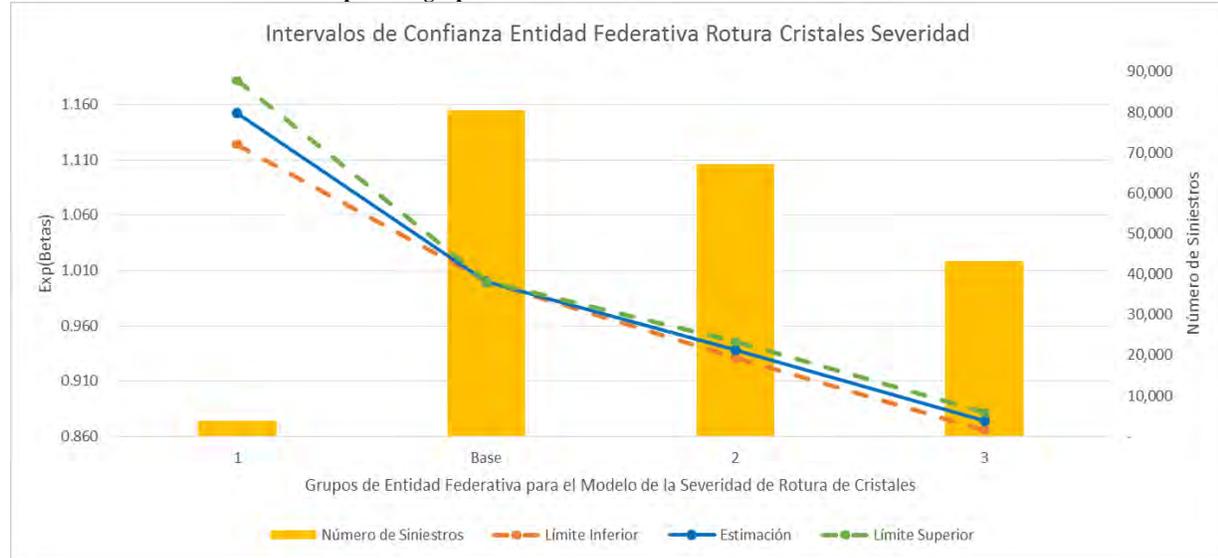


Tabla 3.51 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de rotura de cristales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
1	6.639	6.386	6.143	1,615
2	5.335	5.147	4.967	1,930
3	4.494	4.379	4.268	4,022
4	3.980	3.906	3.832	8,568
5	3.410	3.354	3.299	12,852
6	2.902	2.859	2.817	17,954
7	2.650	2.612	2.576	21,642
8	2.324	2.289	2.254	16,176
9	2.107	2.078	2.051	25,467
10	1.779	1.747	1.717	10,470
11	1.427	1.406	1.386	18,462
12	1.335	1.315	1.296	18,219
Base	1.000	1.000	1.000	24,215
13	0.651	0.640	0.629	13,273

Gráfica 3.26 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de rotura de cristales.

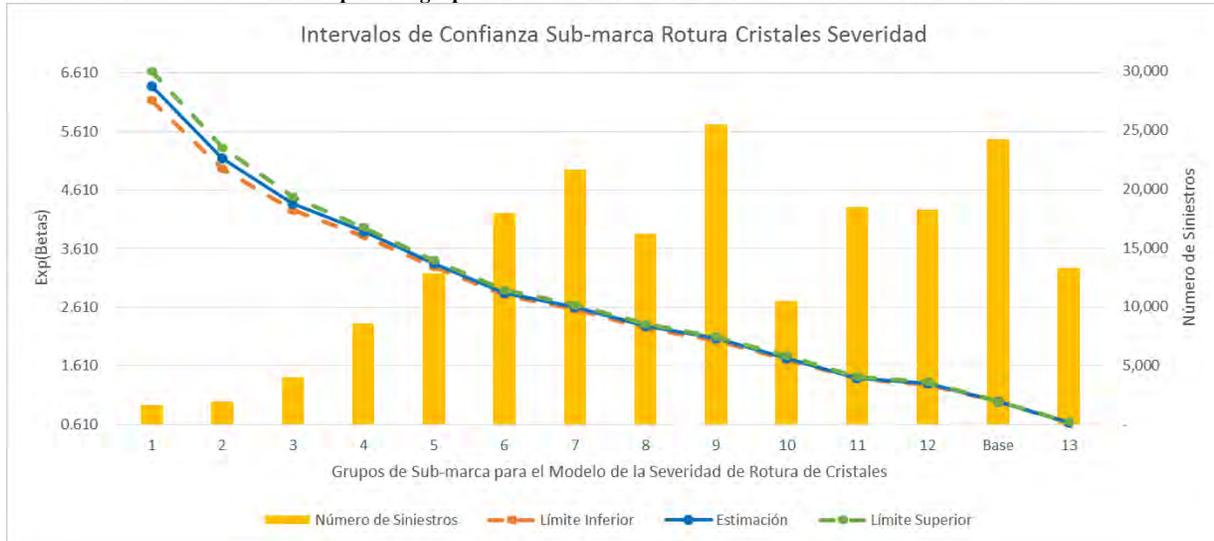


Tabla 3.52 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de severidad de rotura de cristales.

Nivel	Límite Superior	$Exp(\text{Estimación}) \gamma's$	Límite Inferior	Número de Siniestros
Base	1.000	1.000	1.000	55,402
1	0.853	0.844	0.835	30,296
2	0.783	0.776	0.769	45,033
3	0.719	0.711	0.703	33,995
4	0.672	0.663	0.655	18,128
5	0.606	0.596	0.586	10,765
6	0.569	0.545	0.521	1,246

Gráfica 3.27 Intervalos de confianza para la agrupación de Edad de Vehículo del modelo de severidad de rotura de cristales.

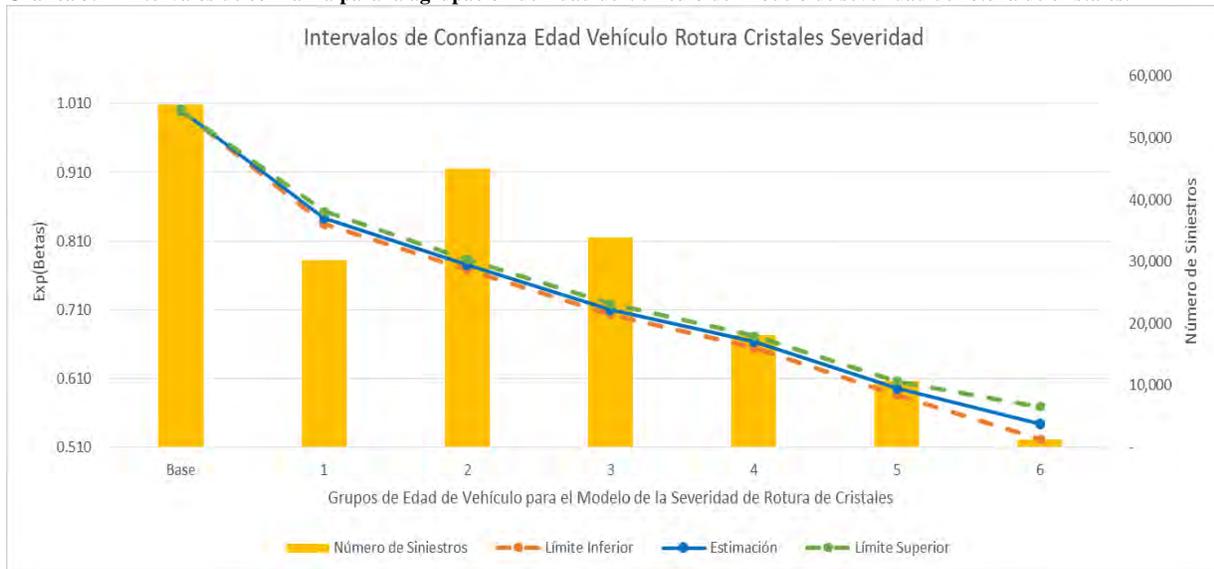
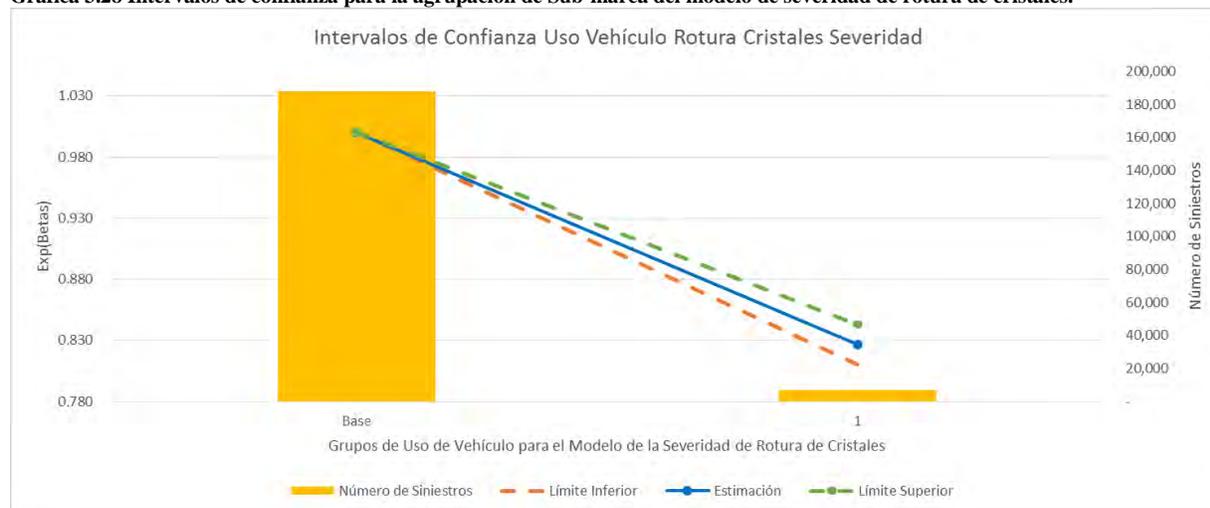


Tabla 3.53 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de rotura de cristales.

Nivel	Límite Superior	Exp(Estimación) $\gamma's$	Límite Inferior	Número de Siniestros
Base	1.000	1.000	1.000	188,055
1	0.843	0.826	0.810	6,810

Gráfica 3.28 Intervalos de confianza para la agrupación de Sub-marca del modelo de severidad de rotura de cristales.



3.2.2 Test de Consistencia aleatoria

Este Test se sostiene sobre la siguiente idea: Supongamos que dividimos a nuestra información de manera aleatoria en n partes de igual número de registros, esta división quedará en función de una nueva variable llamada “Rand”. Ajustamos el modelo con las mismas variables y la misma agrupación de éstas, salvo para la variable a someter a esta prueba en cuyo caso particular haremos una interacción del tipo anidada con la variable “Rand”. Si las n particiones de cada nivel de la variable sometida a este test tuvieran $\beta's$ muy cercanas diremos que es consistente, es decir, el valor de las $\beta's$ se mantiene aún sin considerar toda la información o particiones de ésta y en consecuencia la interacción no será significativa.

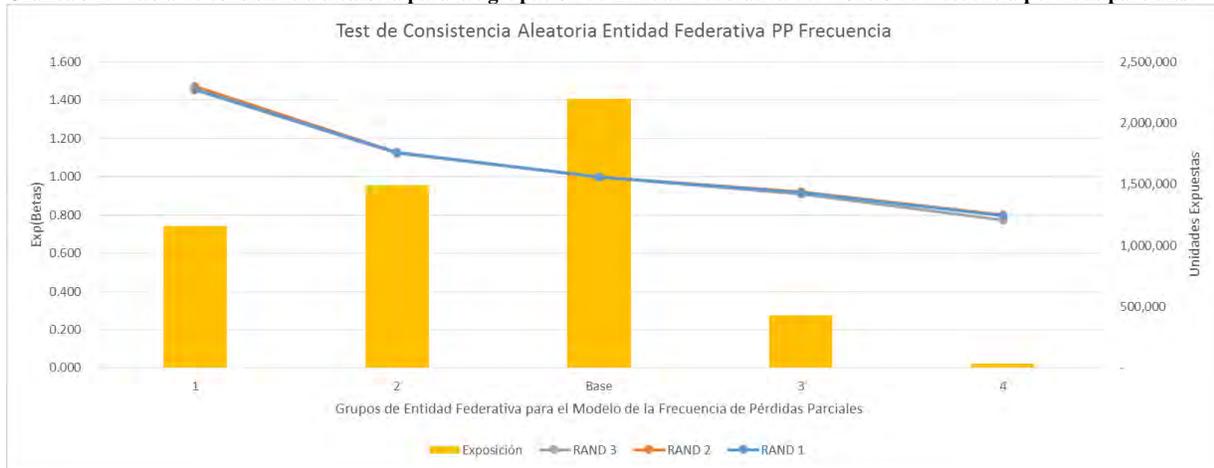
Este test es socorrido en el momento en que se desea saber si un nivel en particular de una agrupación es realmente distinto al resto de niveles.

3.2.2.1 Frecuencia Pérdidas Parciales

Tabla 3.54 Consistencia aleatoria para la agrupación de Entidad Federativa del modelo de frecuencia pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.458	1.472	1.453	1,162,350
2	1.129	1.126	1.125	1,493,639
Base	1.000	1.000	1.000	2,200,937
3	0.919	0.922	0.908	430,805
4	0.798	0.799	0.774	34,087

Gráfica 3.29 Tabla de consistencia aleatoria para la agrupación de Entidad Federativa del modelo de frecuencia pérdidas parciales.



El eje “x” de la gráfica 3.29 corresponde a las distintas zonas geográficas de la variable entidad federativa del modelo de frecuencia de pérdidas parciales, las barras amarillas representan la cantidad de unidades expuestas que posee cada zona geográfica y está medida por el eje derecho “y”. El eje izquierdo “y” de la gráfica mide la proporción en que aumenta o descende la frecuencia de cada zona geográfica en comparación con el nivel base, que en la mayoría de los casos es el nivel con mayor cantidad de unidades expuestas (no es una regla). Las líneas gris, naranja y azul representan cada una al comportamiento de cada partición aleatoria y son medidas por el eje izquierdo “y”. Este diseño de gráfica aplica para toda la sección 3.2.2, en donde los resultados obtenidos al realizar este test fueron satisfactorios, con diferencias notables sólo en los casos de niveles con poco peso (unidades expuestas o número de siniestros, según corresponda).

Tabla 3.55 Consistencia aleatoria para la agrupación de Sub-marca del modelo de frecuencia pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.437	1.462	1.458	56,600
2	1.282	1.243	1.223	598,555
3	1.108	1.125	1.128	1,187,668
Base	0.982	1.015	1.000	1,708,850
4	0.852	0.837	0.842	656,979
5	0.776	0.781	0.769	570,577
6	0.706	0.701	0.699	282,796
7	0.639	0.636	0.626	157,184
8	0.505	0.496	0.502	102,610

Gráfica 3.30 Tabla de consistencia aleatoria para la agrupación de Sub-marca del modelo de frecuencia pérdidas parciales.

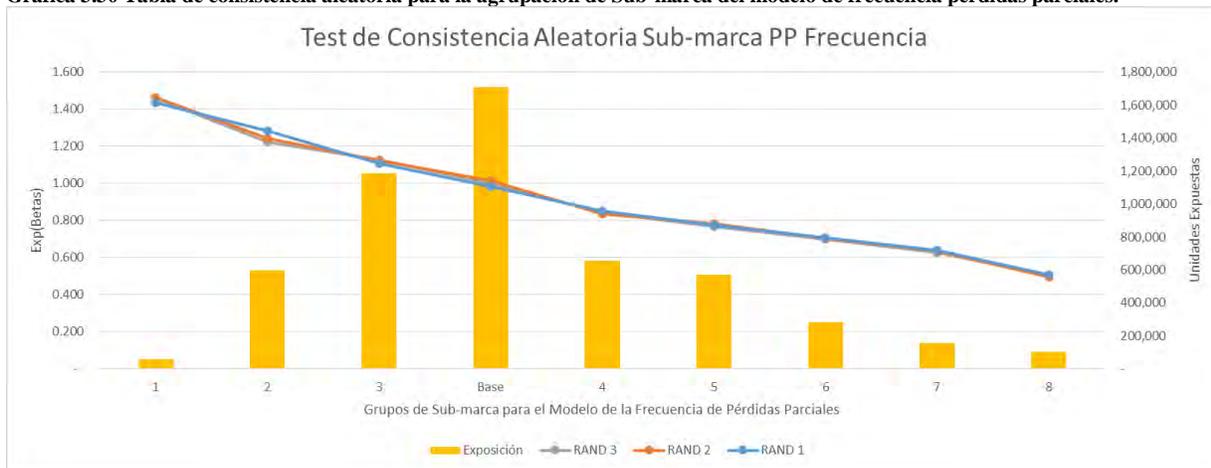


Tabla 3.56 Consistencia aleatoria para la agrupación de Edad de vehículo del modelo de frecuencia pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
Base	1	1	1	2,111,769
1	0.874	0.882	0.875	1,568,631
2	0.872	0.854	0.870	1,200,831
3	0.775	0.746	0.788	260,399
4	0.667	0.668	0.691	66,062
5	0.574	0.599	0.616	43,624
6	0.477	0.506	0.514	46,035
7	0.337	0.269	0.363	24,467

Gráfica 3.31 Tabla de consistencia aleatoria para la agrupación de Edad de vehículo del modelo de frecuencia pérdidas parciales.

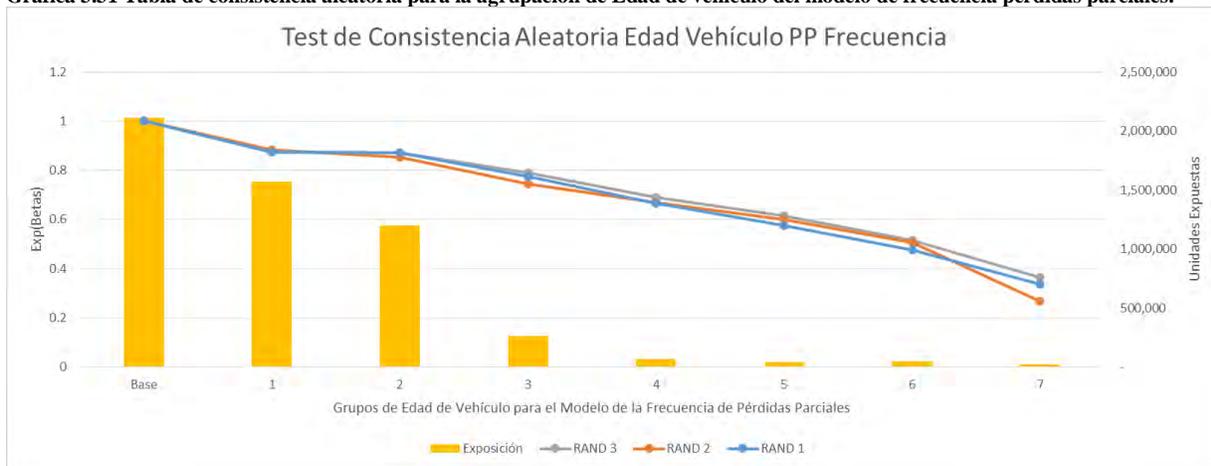


Tabla 3.57 Consistencia aleatoria para la agrupación de Uso de vehículo del modelo de frecuencia pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	2.212	2.425	2.398	85,262
Base	1.000	1.000	1.000	5,006,917
2	0.961	0.934	0.929	145,473
3	0.380	0.386	0.397	84,166

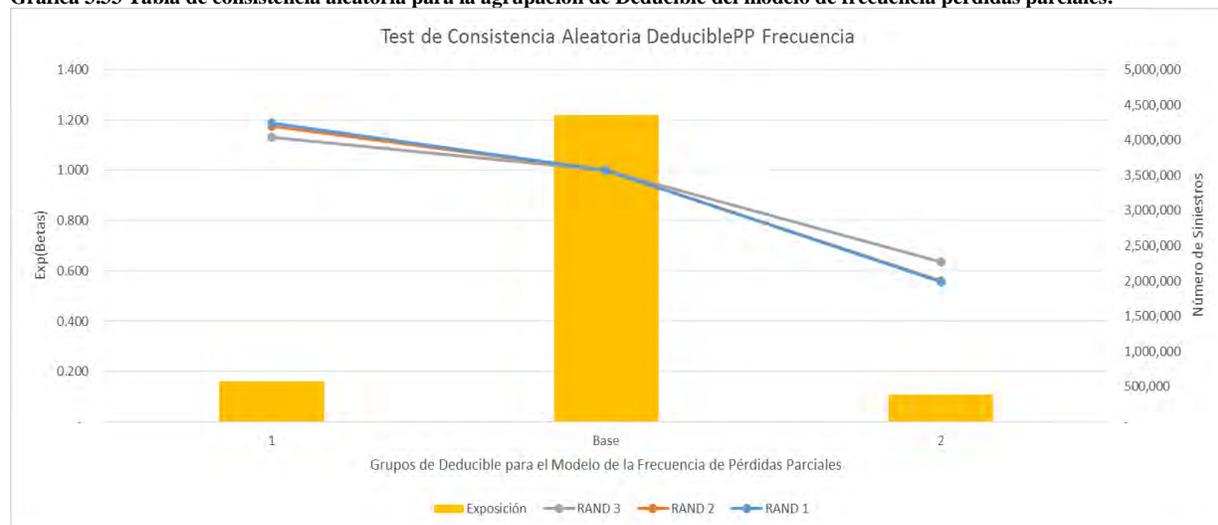
Gráfica 3.32 Tabla de consistencia aleatoria para la agrupación de Uso de vehículo del modelo de frecuencia pérdidas parciales.



Tabla 3.58 Consistencia aleatoria para la agrupación de Deducible del modelo de frecuencia pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.188	1.175	1.132	571,979
Base	1.000	1.000	1.000	4,359,896
2	0.557	0.559	0.636	389,944

Gráfica 3.33 Tabla de consistencia aleatoria para la agrupación de Deducible del modelo de frecuencia pérdidas parciales.



3.2.2.2 Severidad Pérdidas Parciales

Tabla 3.59 Consistencia aleatoria para la agrupación de Entidad Federativa del modelo de severidad pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	1.888	1.922	1.885	6121
2	1.648	1.622	1.655	27303
3	1.480	1.511	1.477	31644
4	1.353	1.342	1.348	88467
5	1.193	1.204	1.185	122473
Base	1.000	1.000	1.000	260535

Gráfica3.34 Tabla de consistencia aleatoria para la agrupación de Entidad Federativa del modelo de severidad pérdidas parciales.

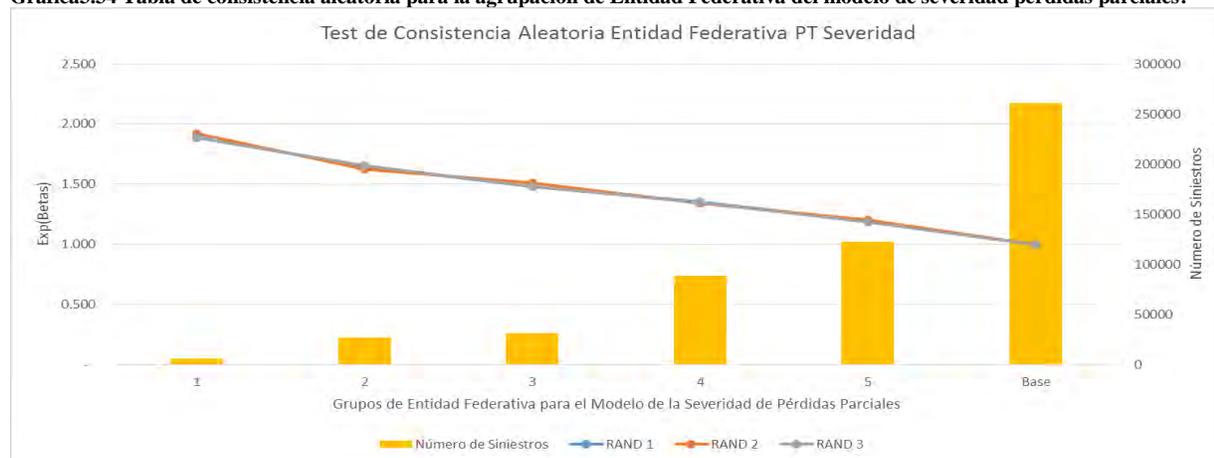


Tabla 3.60 Consistencia aleatoria para la agrupación de Sub-marca del modelo de severidad pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	4.618	4.755	4.611	2,285
2	3.245	3.608	3.525	964
3	3.259	2.983	2.749	1,620
4	2.425	2.561	2.462	12,617
5	1.891	1.941	1.987	6,914
6	1.719	1.748	1.772	16,147
7	1.595	1.559	1.569	22,996
8	1.426	1.382	1.484	14,774
9	1.304	1.291	1.310	37,868
10	1.208	1.212	1.217	63,500
11	1.126	1.136	1.148	98,787
Base	1.000	1.000	1.000	108,846

12	0.934	0.940	0.911	73,059
13	0.781	0.790	0.767	59,573
14	0.801	0.882	0.850	6,110
15	0.680	0.692	0.636	10,483

Gráfica 3.35 Tabla de consistencia aleatoria para la agrupación de Sub-marca del modelo de severidad pérdidas parciales.

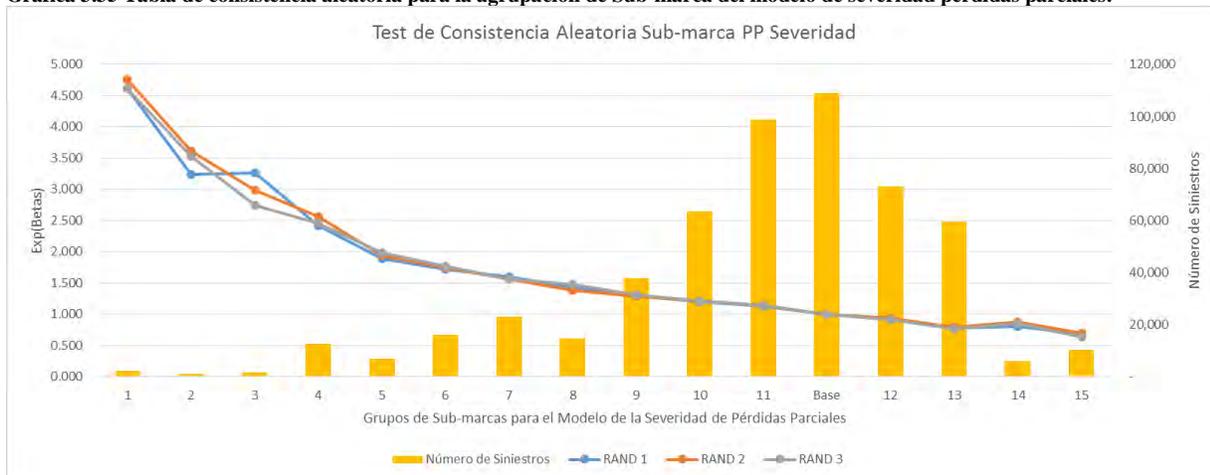


Tabla 3.61 Consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de severidad pérdidas parciales.

<i>Nivel</i>	<i>RAND 1</i>	<i>RAND 2</i>	<i>RAND 3</i>	<i>Número de Siniestros</i>
Base	1.000	1.000	1.000	142,928
1	0.932	0.926	0.911	100,852
2	0.841	0.863	0.834	66,503
3	0.782	0.783	0.767	73,798
4	0.711	0.735	0.732	29,044
5	0.679	0.680	0.669	50,005
6	0.589	0.603	0.593	31,288
7	0.533	0.536	0.541	22,130
8	0.482	0.468	0.467	15,499
9	0.373	0.395	0.393	3,177
10	0.361	0.349	0.347	1,319

Gráfica 3.36 Tabla de consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de severidad pérdidas parciales.

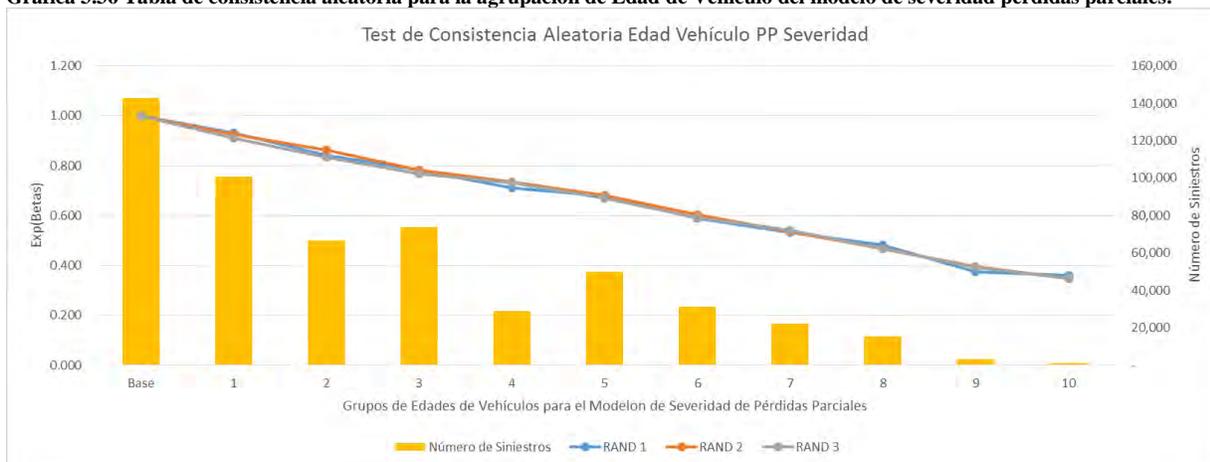


Tabla 3.62 Consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de severidad pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	1.292	1.348	1.269	10,986
2	1.044	1.047	1.155	2,946
Base	1.000	1.000	1.000	507,762
3	0.802	0.814	0.726	14,849

Gráfica 3.37 Tabla de consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de severidad pérdidas parciales.

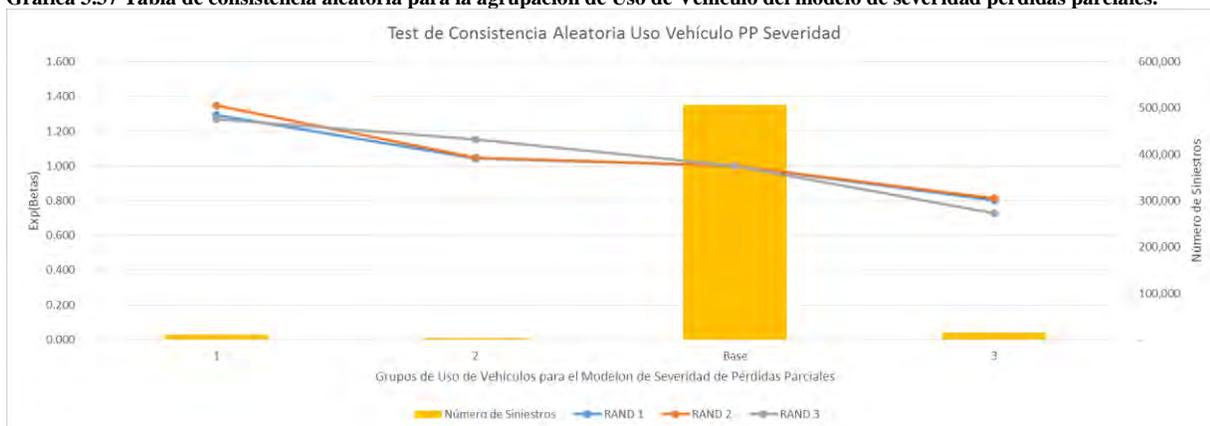
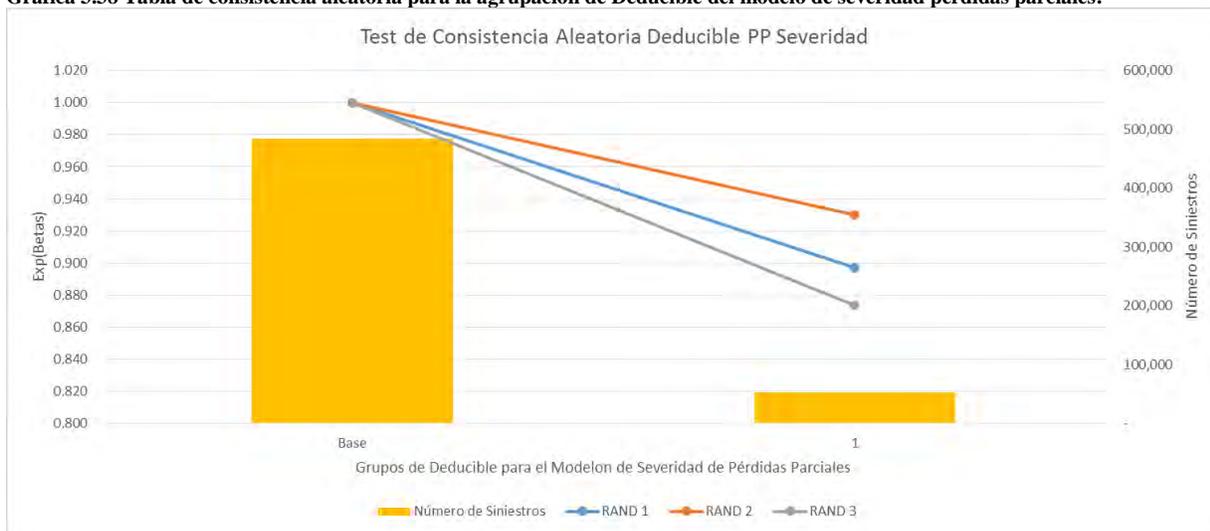


Tabla 3.63 Consistencia aleatoria para la agrupación de Deducible del modelo de severidad pérdidas parciales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
Base	1.000	1.000	1.000	484,103
1	0.897	0.930	0.873	52,440

Gráfica 3.38 Tabla de consistencia aleatoria para la agrupación de Deducible del modelo de severidad pérdidas parciales.



3.2.2.3 Frecuencia Pérdidas Totales

Tabla 3.64 Consistencia aleatoria para la agrupación de Entidad Federativa del modelo de frecuencia pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.888	1.922	1.885	6121
2	1.648	1.622	1.655	27303
3	1.480	1.511	1.477	31644
4	1.353	1.342	1.348	88467
5	1.193	1.204	1.185	122473
Base	1.000	1.000	1.000	260535

Gráfica 3.39 Tabla de consistencia aleatoria para la agrupación de Entidad Federativa de frecuencia pérdidas totales

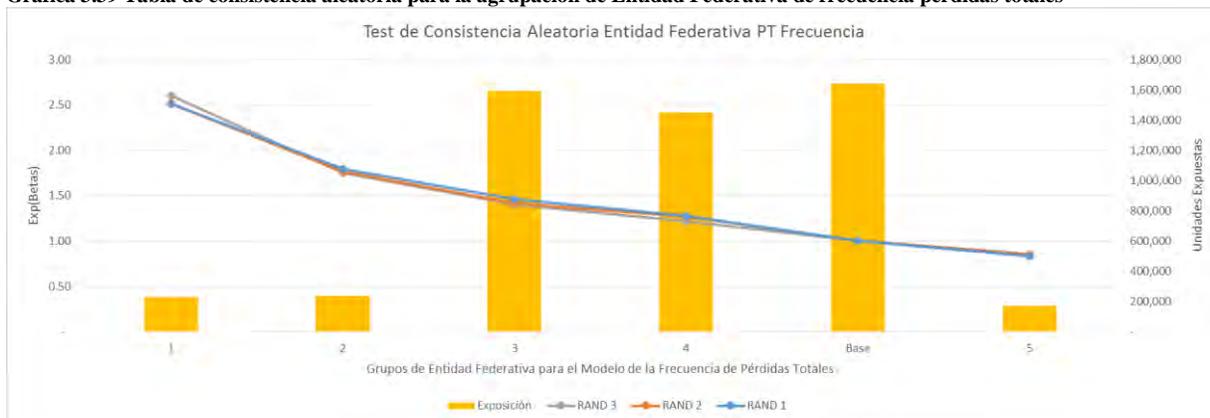


Tabla 3.65 Consistencia aleatoria para la agrupación de Sub-marca del modelo de frecuencia pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	3.15	2.66	2.47	20,367
2	2.39	2.16	1.87	25,599
3	1.83	1.75	1.84	217,684
4	1.57	1.58	1.55	557,782
5	1.35	1.36	1.26	552,650
6	1.23	1.26	1.21	1,105,259
Base	1.00	1.00	1.00	1,010,129
7	0.90	0.88	0.83	142,150
8	0.73	0.75	0.79	568,461
9	0.67	0.63	0.65	424,977
10	0.63	0.60	0.52	386,242
11	0.41	0.34	0.37	310,518

Gráfica 3.40 Tabla de consistencia aleatoria para la agrupación de Sub-marca de frecuencia pérdidas totales.

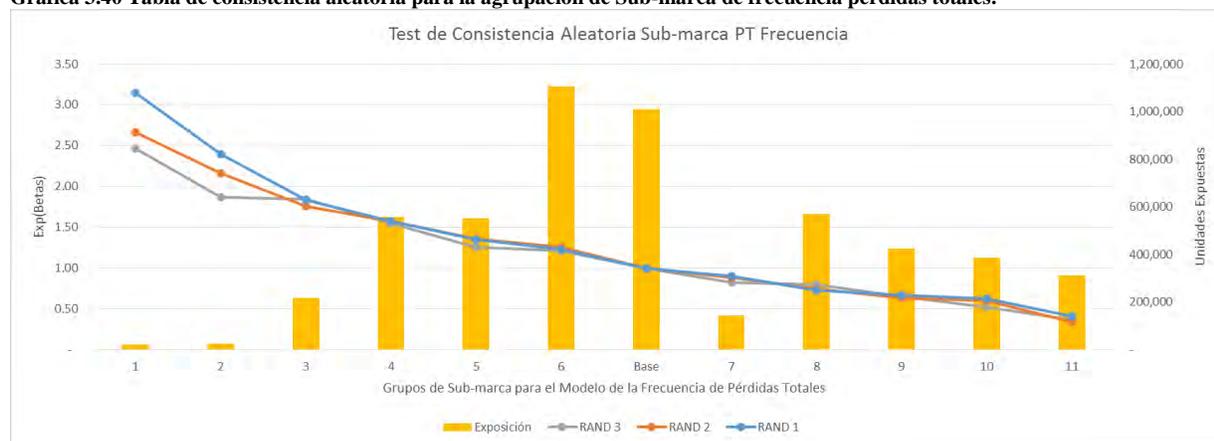


Tabla 3.66 Consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de frecuencia pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.21	1.16	1.15	1,937,015
2	1.09	1.08	1.04	666,868
Base	1.00	1.00	1.00	2,042,130
3	0.84	0.89	0.86	531,264
4	0.72	0.66	0.58	99,568
5	0.46	0.54	0.47	20,506
6	0.24	0.31	0.28	24,467

Gráfica 3.41 Tabla de consistencia aleatoria para la agrupación de Edad de Vehículo de frecuencia pérdidas totales.

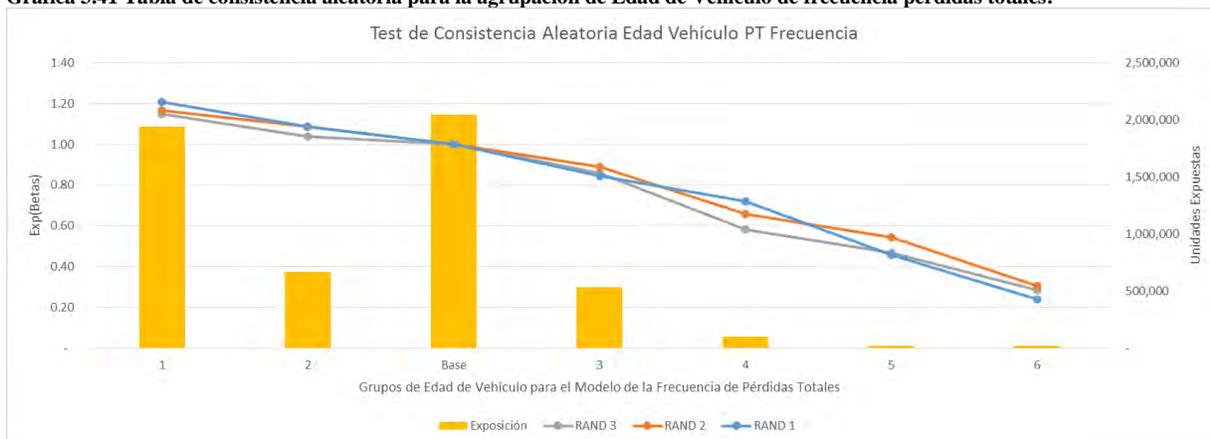


Tabla 3.67 Consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de frecuencia pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.08	1.27	1.53	20,202
Base	1.00	1.00	1.00	5,123,741
2	0.93	0.93	0.88	108,391
3	0.68	0.69	0.61	69,485

Gráfica 3.42 Tabla de consistencia aleatoria para la agrupación de Uso de Vehículo de frecuencia pérdidas totales.

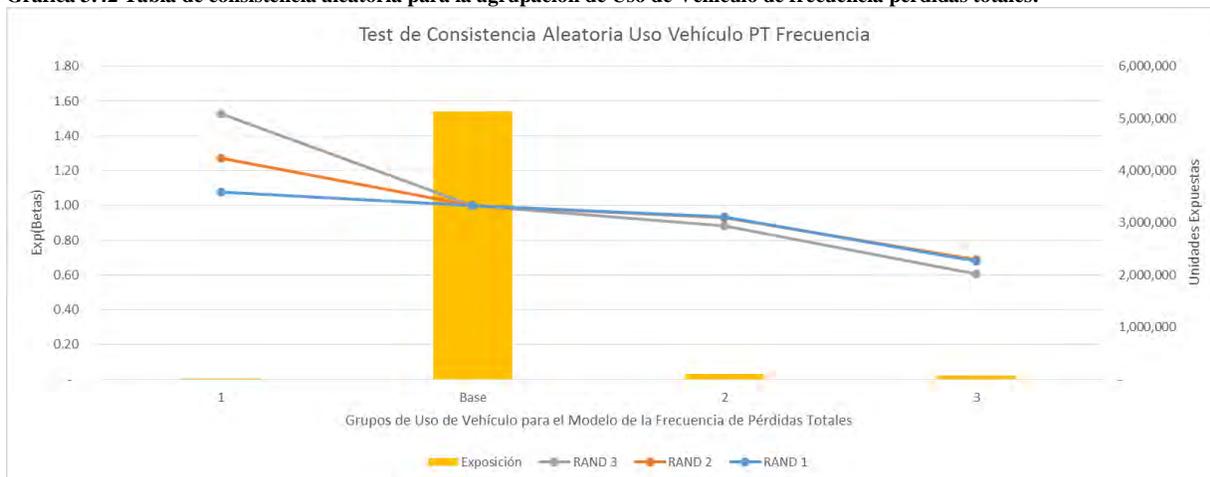
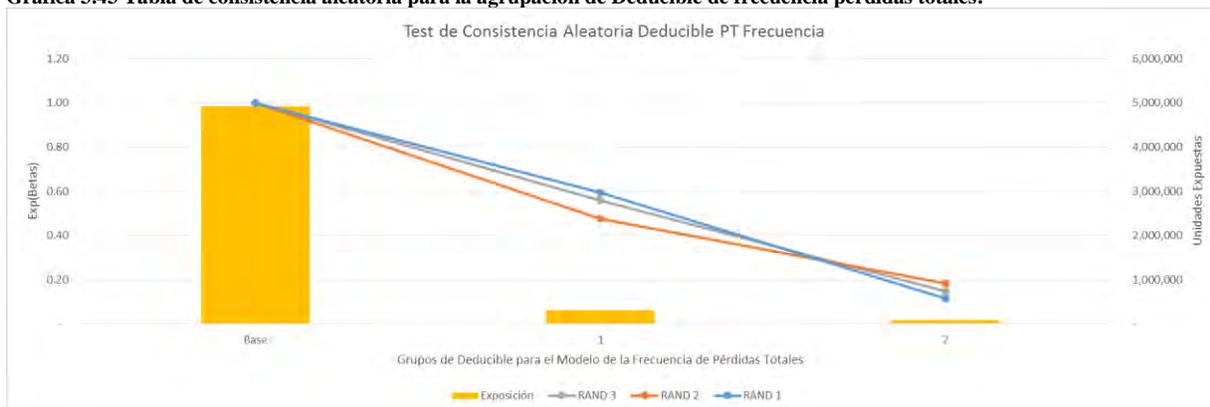


Tabla 3.68 Consistencia aleatoria para la agrupación de Deducible del modelo de frecuencia pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
Base	1.00	1.00	1.00	4,931,874
1	0.59	0.48	0.56	301,023
2	0.11	0.18	0.15	88,921

Gráfica 3.43 Tabla de consistencia aleatoria para la agrupación de Deducible de frecuencia pérdidas totales.



3.2.2.4 Severidad Pérdidas Totales

Tabla 3.69 Consistencia aleatoria para la agrupación de Entidad Federativa del modelo de severidad pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	1.48	1.45	1.50	4,449
2	1.39	1.36	1.37	8,040
3	1.29	1.26	1.25	10,585
4	1.22	1.17	1.21	6,237
Base	1.00	1.00	1.00	10,907

Gráfica 3.44 Tabla de consistencia aleatoria para la agrupación de Entidad Federativa del modelo de severidad pérdidas totales.

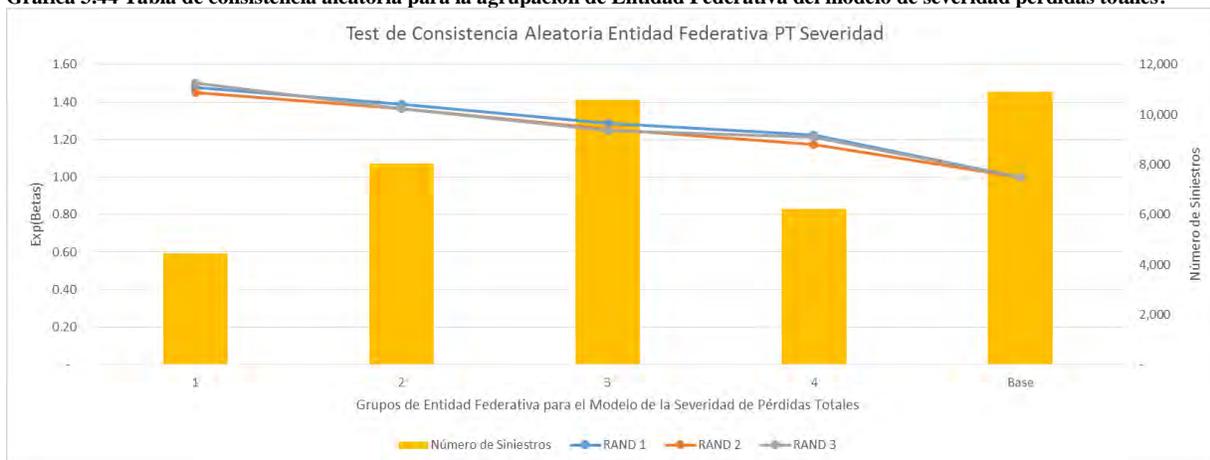


Tabla 3.70 Consistencia aleatoria para la agrupación de Sub-marca del modelo de severidad pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	3.123	3.309	2.905	1,013
2	2.202	2.021	2.101	1,313
3	1.759	1.790	1.707	1,176
4	1.640	1.632	1.634	2,045
5	1.499	1.521	1.475	3,719

6	1.285	1.271	1.306	5,811
Base	1.000	1.000	1.000	11,740
7	0.971	0.929	0.939	3,437
8	0.870	0.821	0.827	4,292
9	0.670	0.706	0.689	2,525
10	0.627	0.620	0.623	1,663
11	0.532	0.516	0.498	1,484

Gráfica 3.45 Tabla de consistencia aleatoria para la agrupación de Sub-marca del modelo de severidad pérdidas totales.

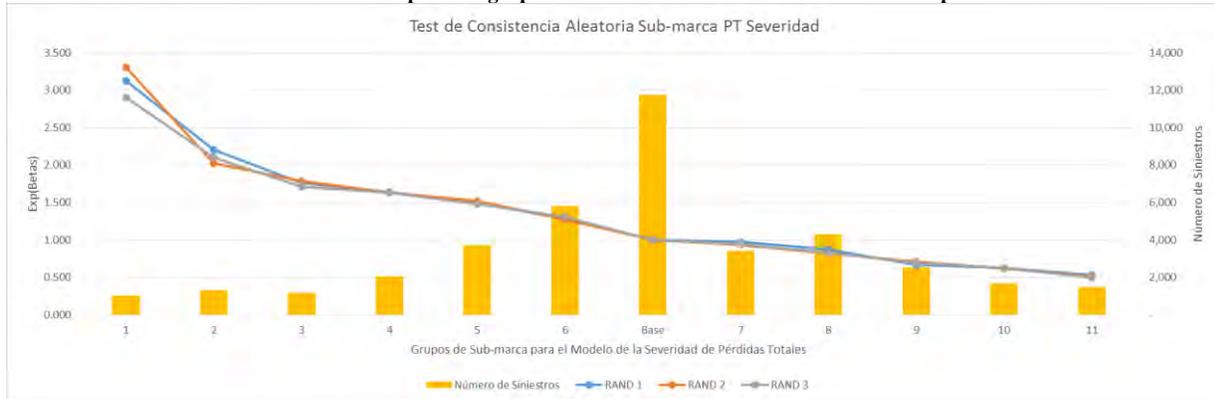


Tabla 3.71 Consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de severidad pérdidas totales.

<i>Nivel</i>	<i>RAND 1</i>	<i>RAND 2</i>	<i>RAND 3</i>	<i>Número de Siniestros</i>
1	1.321	1.285	1.252	2,967
Base	1.000	1.000	1.000	8,184
2	0.814	0.832	0.790	2,835
3	0.721	0.718	0.731	4,982
4	0.614	0.592	0.611	3,200
5	0.575	0.548	0.554	3,194
6	0.495	0.497	0.498	2,840
7	0.427	0.425	0.441	2,908
8	0.355	0.375	0.362	3,767
9	0.277	0.288	0.278	2,686
10	0.257	0.248	0.245	1,893
11	0.204	0.189	0.176	762

Gráfica 3.46 Tabla de consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de severidad pérdidas totales.

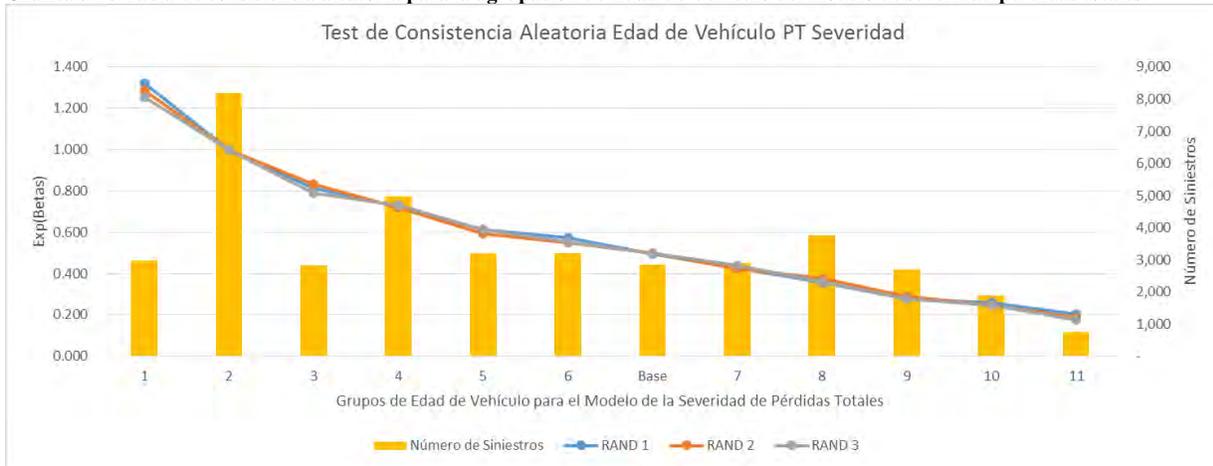


Tabla 3.72 Consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de severidad pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	1.066	1.116	1.113	979
Base	1.000	1.000	1.000	39,239

Gráfica 3.47 Tabla de consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de severidad pérdidas totales.

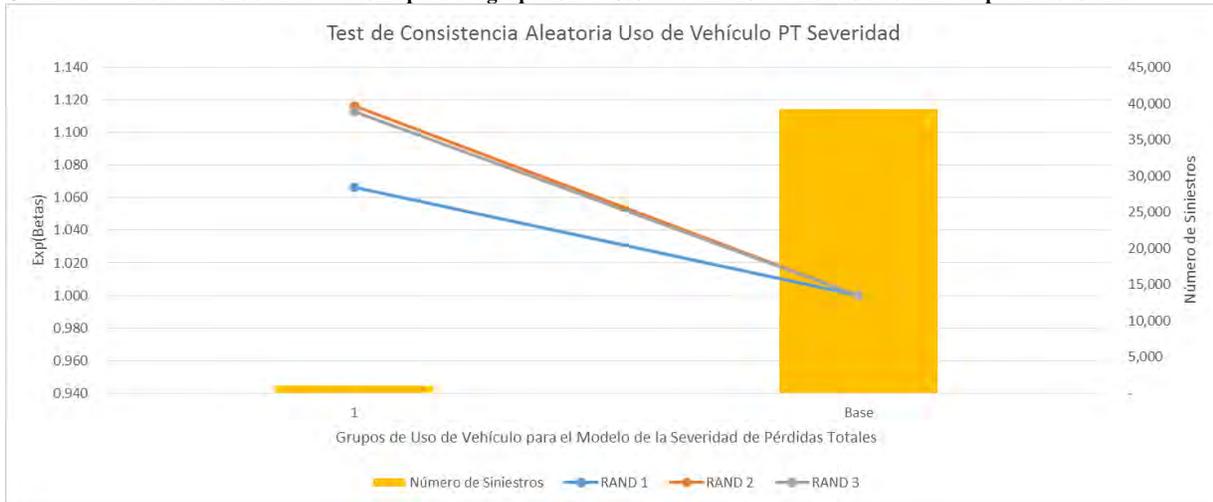
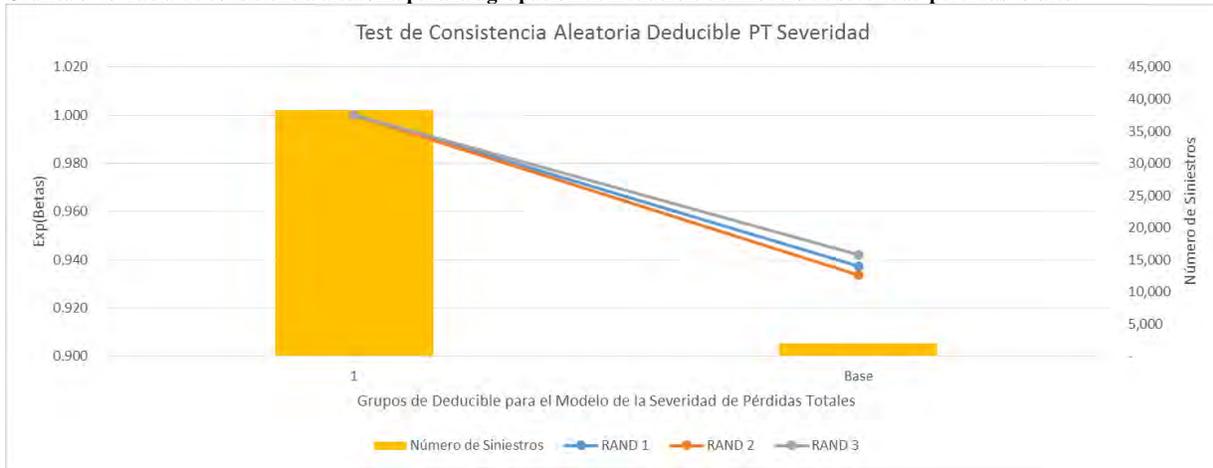


Tabla 3.73 Consistencia aleatoria para la agrupación de Deducible del modelo de severidad pérdidas totales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
Base	1.000	1.000	1.000	38,268
1	0.937	0.934	0.942	1,950

Gráfica 3.48 Tabla de consistencia aleatoria para la agrupación de Deducible del modelo de severidad pérdidas totales.



3.2.2.5 Frecuencia Rotura de Cristales

Tabla 3.74 Consistencia aleatoria para la agrupación de Entidad Federativa del modelo de frecuencia rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	3.158	3.101	3.045	309,105
2	2.308	2.354	2.348	567,202
3	1.900	1.922	1.926	828,956
4	1.733	1.753	1.700	587,284
5	1.447	1.468	1.486	432,237
6	1.238	1.199	1.241	816,146
7	1.154	1.142	1.211	537,111
Base	1.000	1.000	1.000	1,243,778

Gráfica 3.49 Tabla de consistencia aleatoria para la agrupación de Entidad Federativa del modelo de frecuencia rotura de cristales.

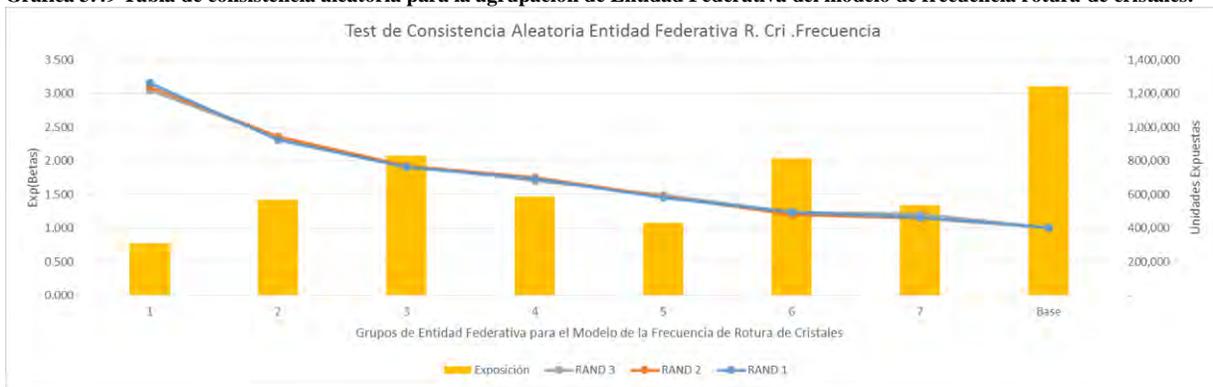


Tabla 3.75 Consistencia aleatoria para la agrupación de Sub-marca del modelo de frecuencia rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.602	1.367	1.637	7,988
2	1.294	1.275	1.304	187,952
Base	1.000	1.000	1.000	1,413,379

3	0.852	0.843	0.849	862,808
4	0.712	0.725	0.731	1,155,333
5	0.634	0.633	0.645	704,107
6	0.571	0.597	0.605	312,599
7	0.474	0.474	0.498	341,708
8	0.331	0.359	0.336	119,188
9	0.236	0.281	0.282	180,824
10	0.088	0.074	0.089	35,932

Gráfica 3.50 Tabla de consistencia aleatoria para la agrupación de Sub-marca del modelo de frecuencia rotura de cristales.

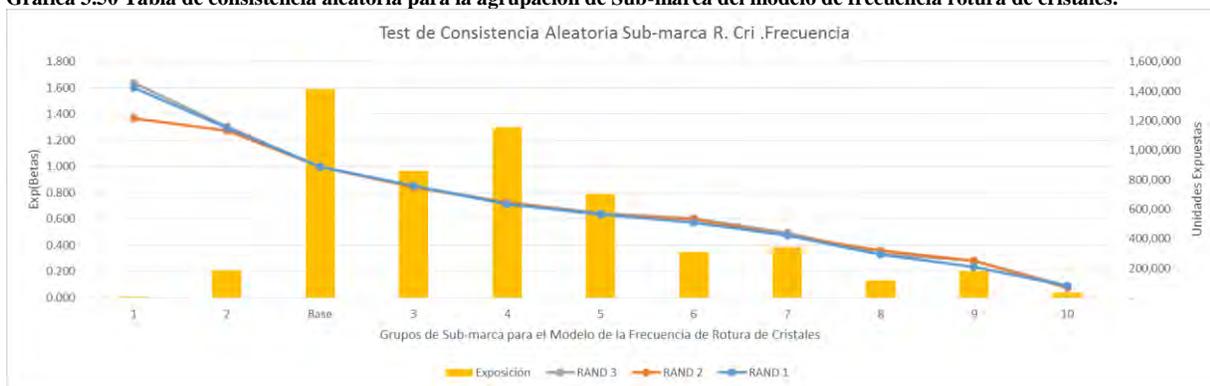


Tabla 3.76 Consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de frecuencia rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	1.084	1.063	1.073	2,054,392
Base	1.000	1.000	1.000	2,482,254
2	0.803	0.837	0.794	618,486
3	0.623	0.604	0.639	112,366
4	0.507	0.443	0.434	29,853
5	0.251	0.264	0.251	24,467

Gráfica 3.51 Tabla de consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de frecuencia rotura de cristales.

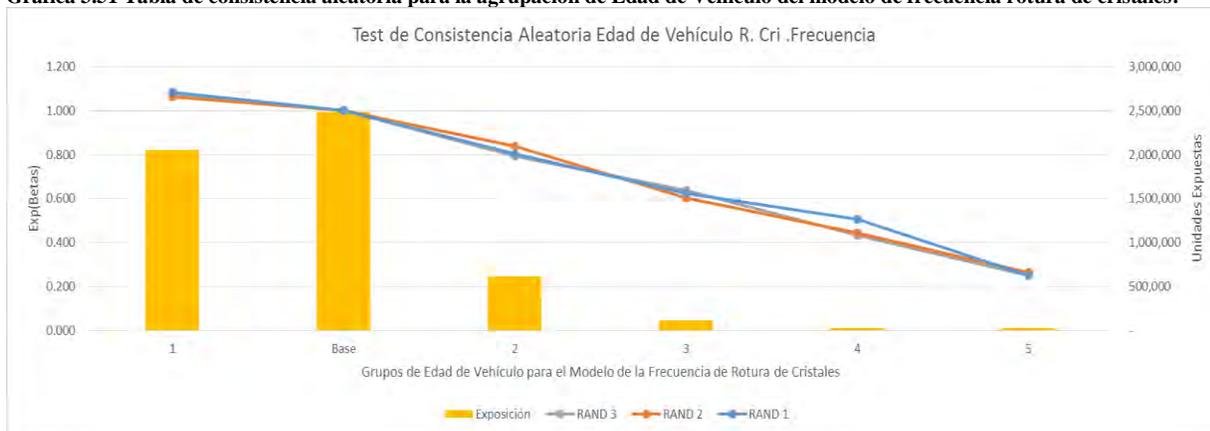
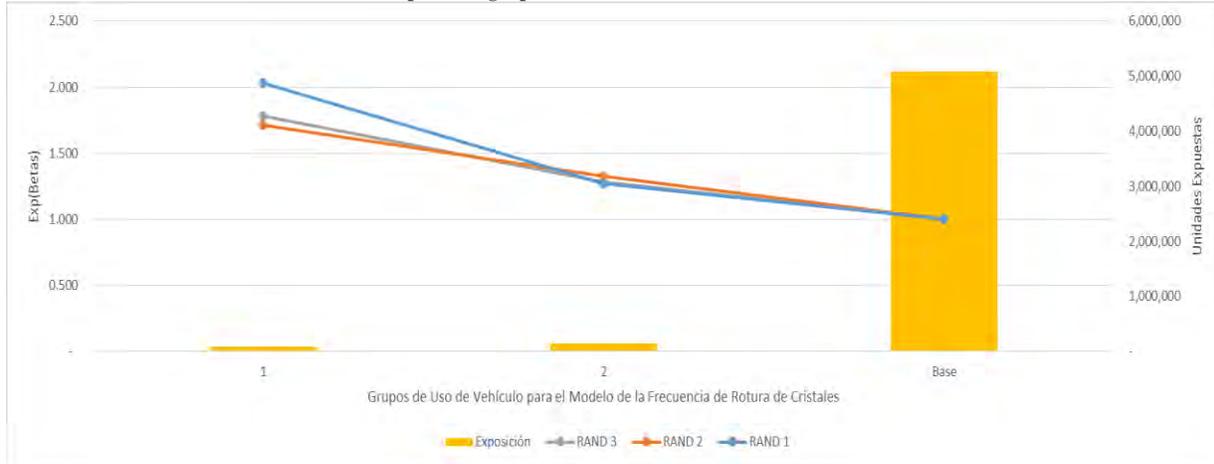


Tabla 3.77 Consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de frecuencia rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Exposición
1	2.030	1.714	1.781	89,803
2	1.269	1.329	1.285	140,944
Base	1.000	1.000	1.000	5,091,071

Gráfica 3.52 Tabla de consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de frecuencia rotura de cristales.



3.2.2.6 Severidad Rotura de Cristales

Tabla 3.78 Consistencia aleatoria para la agrupación de Entidad Federativa del modelo de severidad rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	1.148	1.179	1.107	3,764
Base	1.000	1.000	1.000	80,552
2	0.928	0.946	0.923	67,210
3	0.866	0.868	0.870	43,339

Gráfica 3.53 Tabla de consistencia aleatoria para la agrupación de Entidad Federativa del modelo de severidad rotura de cristales.

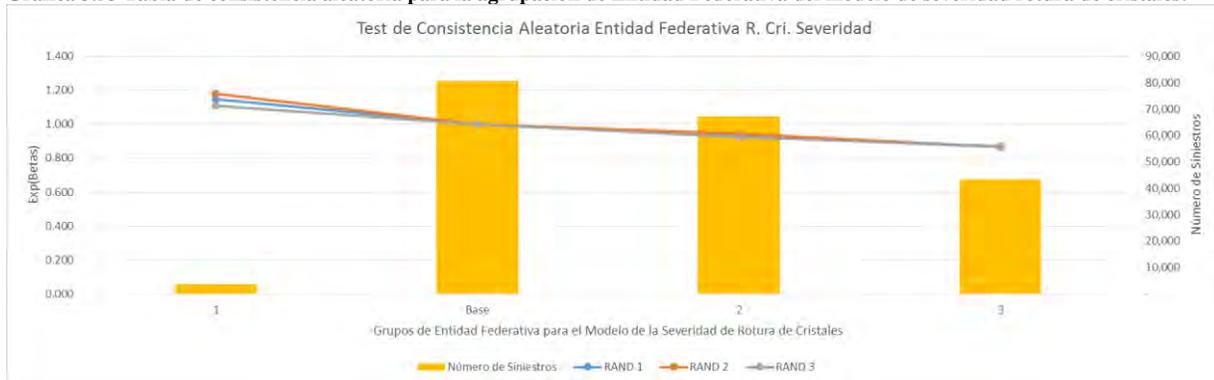


Tabla 3.79 Consistencia aleatoria para la agrupación de Sub-marca del modelo de severidad rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
1	6.365	6.378	6.477	1,615
2	5.327	5.183	5.021	1,930
3	4.305	4.418	4.458	4,022
4	3.901	3.955	3.905	8,568
5	3.281	3.447	3.379	12,852
6	2.903	2.852	2.855	17,954
7	2.575	2.652	2.635	21,642
8	2.288	2.288	2.317	16,176
9	2.118	2.068	2.076	25,467
10	1.798	1.744	1.722	10,470
11	1.378	1.413	1.443	18,462
12	1.326	1.333	1.301	18,219
Base	1.000	1.000	1.000	24,215
13	0.626	0.670	0.637	13,273

Gráfica 3.54 Tabla de consistencia aleatoria para la agrupación de Sub-marca del modelo de severidad rotura de cristales.

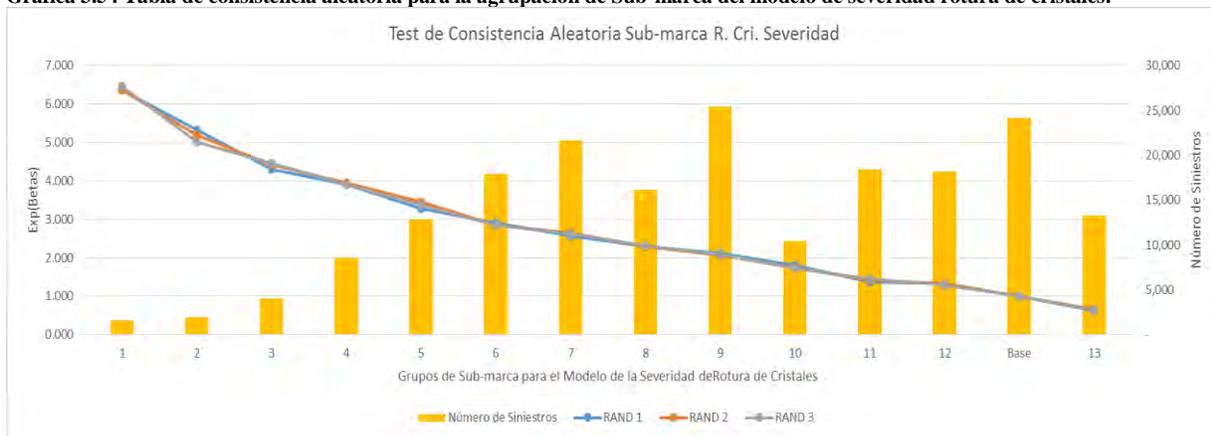


Tabla 3.80 Consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de severidad rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
Base	1.000	1.000	1.000	55,402
1	0.837	0.853	0.850	30,296
2	0.785	0.769	0.780	45,033
3	0.710	0.725	0.704	33,995
4	0.655	0.664	0.676	18,128
5	0.595	0.607	0.592	10,765
6	0.564	0.527	0.546	1,246

Gráfica 3.55 Tabla de consistencia aleatoria para la agrupación de Edad de Vehículo del modelo de severidad rotura de cristales.

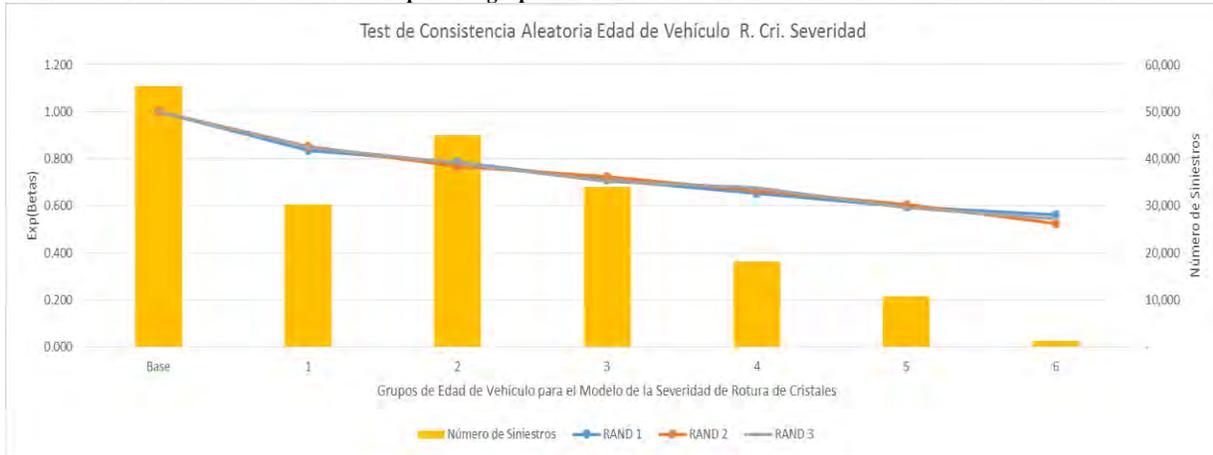
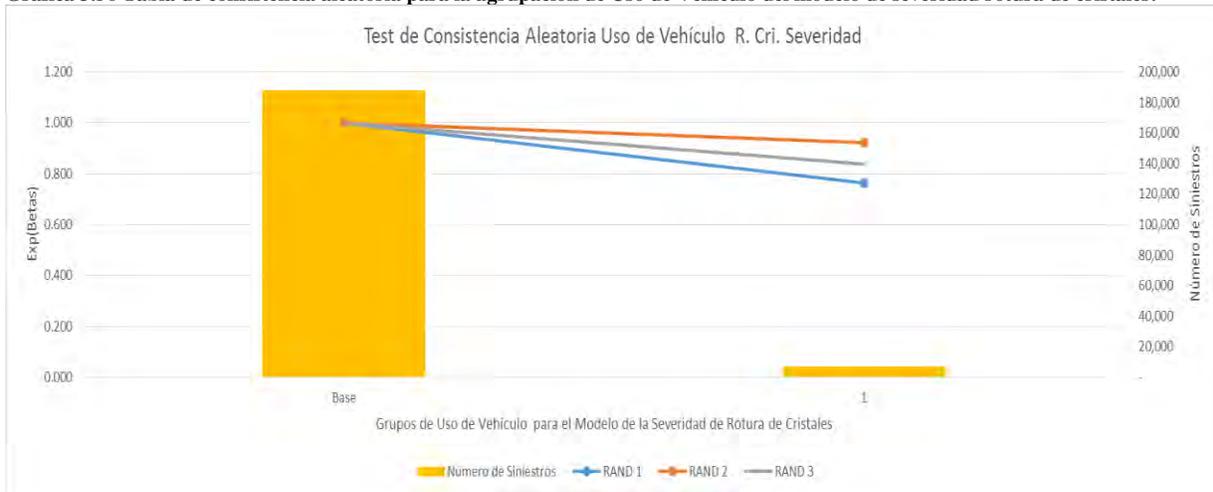


Tabla 3.81 Consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de severidad rotura de cristales.

Nivel	RAND 1	RAND 2	RAND 3	Número de Siniestros
Base	1.000	1.000	1.000	188,055
1	0.764	0.921	0.837	6,810

Gráfica 3.56 Tabla de consistencia aleatoria para la agrupación de Uso de Vehículo del modelo de severidad rotura de cristales.



3.2.3 Residuos

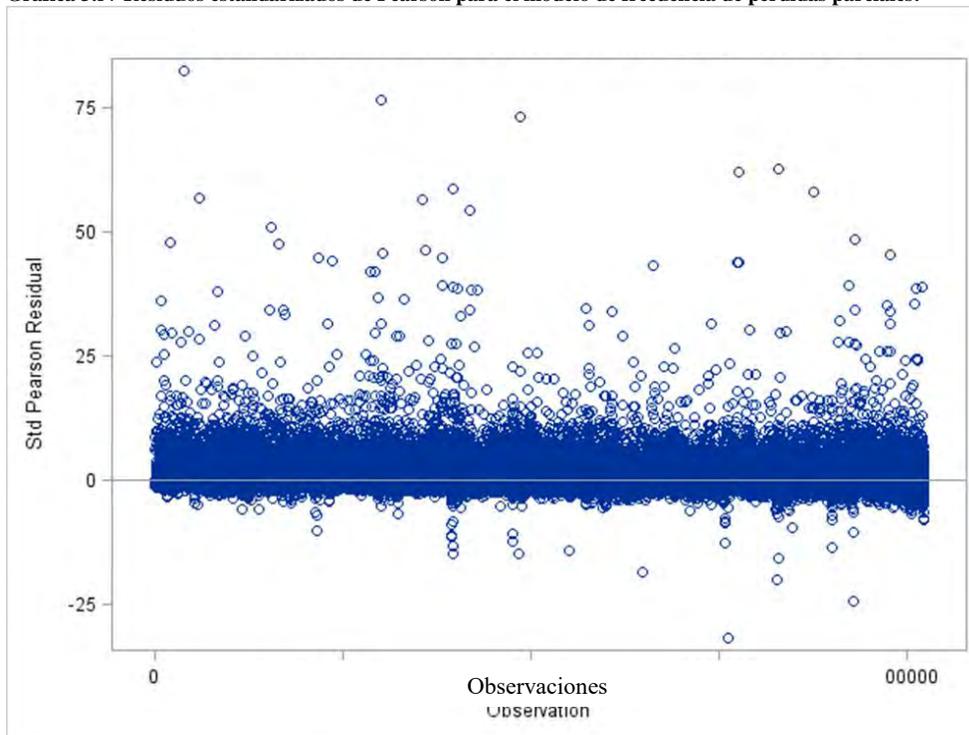
En la sección 1.11 hablamos de los residuos, éstos son una medida de bondad de ajuste pues miden la diferencia que existe entre la estimación y el valor real, de tal modo que entre menor sea la distancia existente entre el valor real y el estimado, mejor será el ajuste del modelo. Como parte del análisis de la bondad de ajuste se desplegarán al lector dos gráficas:

- Gráfica de residuos estandarizados de Pearson.
- Gráfica de residuos estandarizados de Devianza.

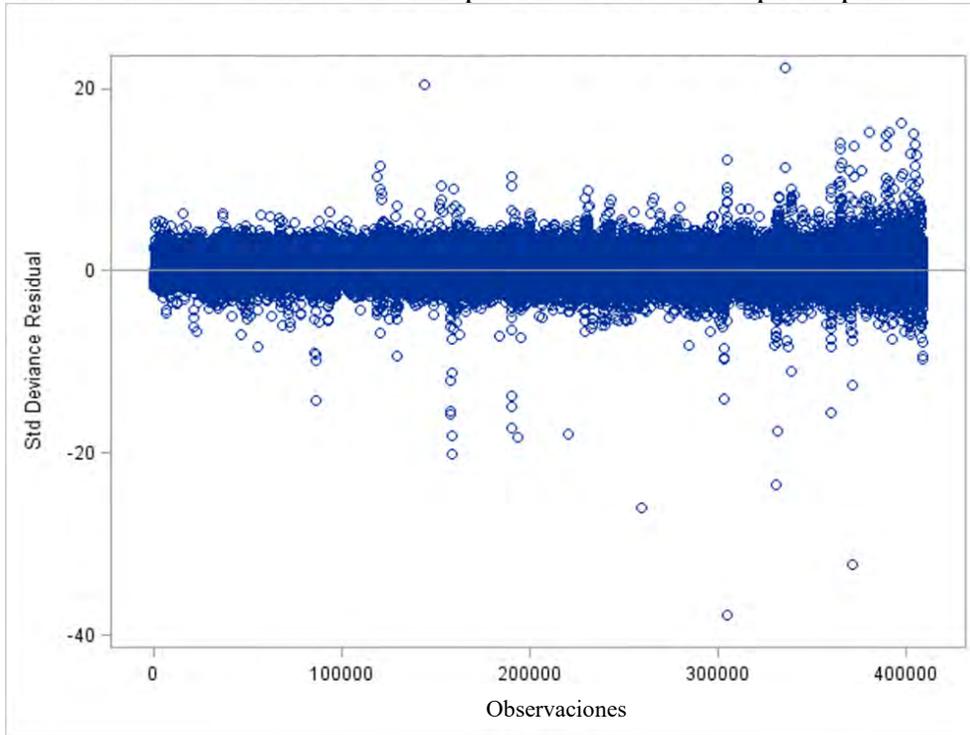
Como se verá a lo largo de esta sección ambas gráficas refleja un buen grado de ajuste en el modelo, pues en su mayoría las diferencias son muy cercanas a cero, tal y como podremos comprobar en las siguientes seis subsecciones.

2.3.3.1 Frecuencia Pérdidas Parciales

Gráfica 3.57 Residuos estandarizados de Pearson para el modelo de frecuencia de pérdidas parciales.

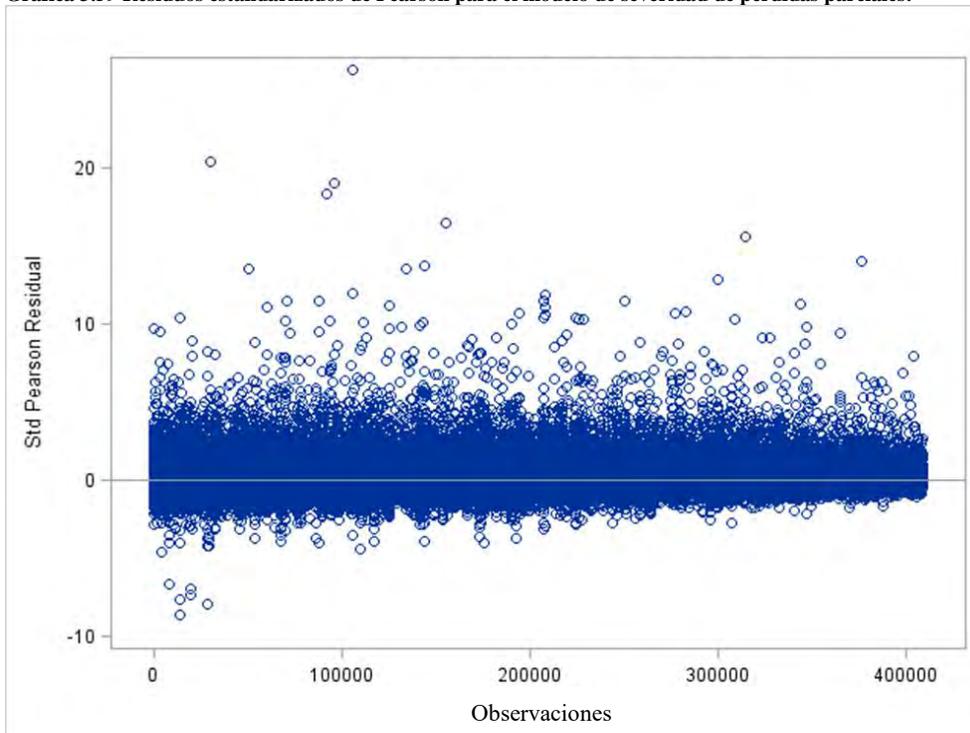


Gráfica 3.58 Residuos estandarizados de Devianza para el modelo de frecuencia de pérdidas parciales.

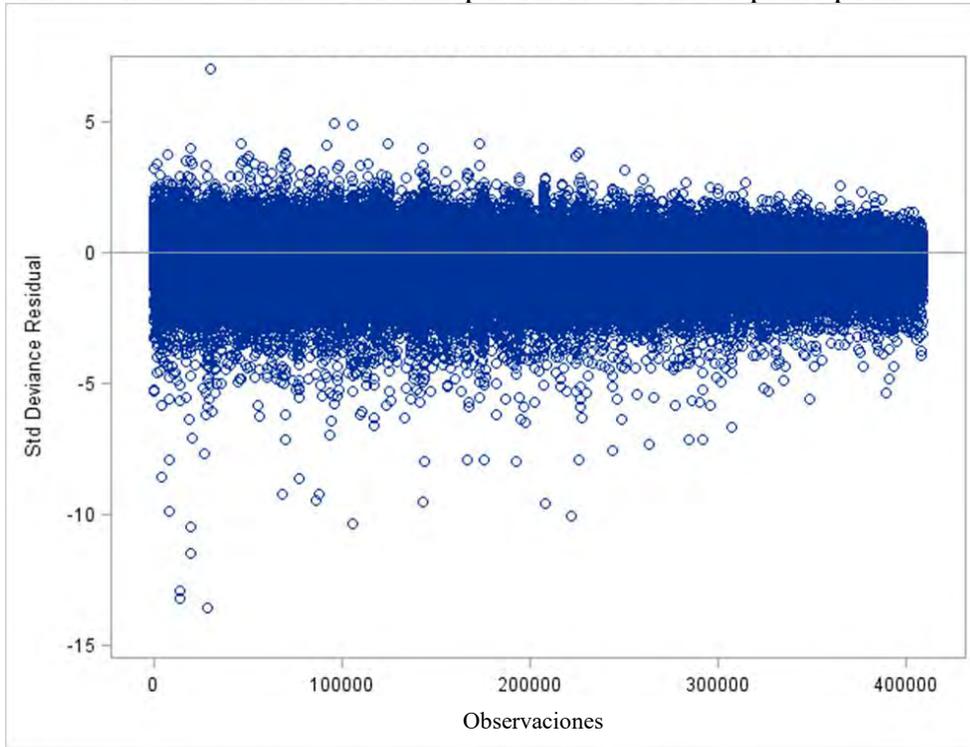


3.2.3.2 Severidad Pérdidas Parciales

Gráfica 3.59 Residuos estandarizados de Pearson para el modelo de severidad de pérdidas parciales.

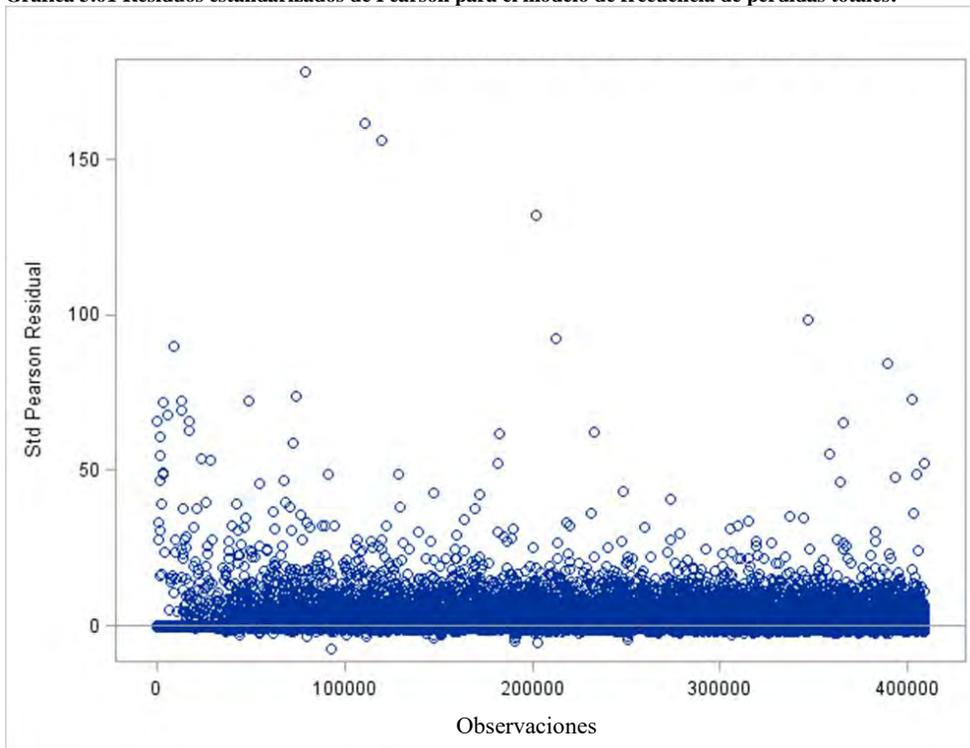


Gráfica 3.60 Residuos estandarizados de Devianza para el modelo de severidad de pérdidas parciales.

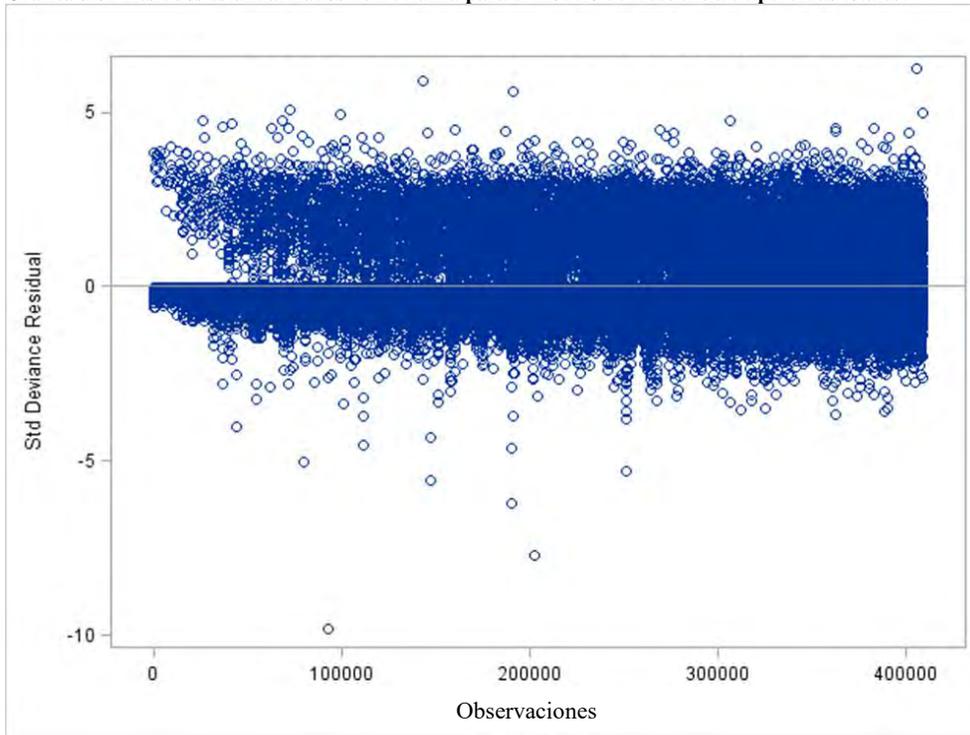


3.2.3.3 Frecuencia Pérdidas Totales

Gráfica 3.61 Residuos estandarizados de Pearson para el modelo de frecuencia de pérdidas totales.

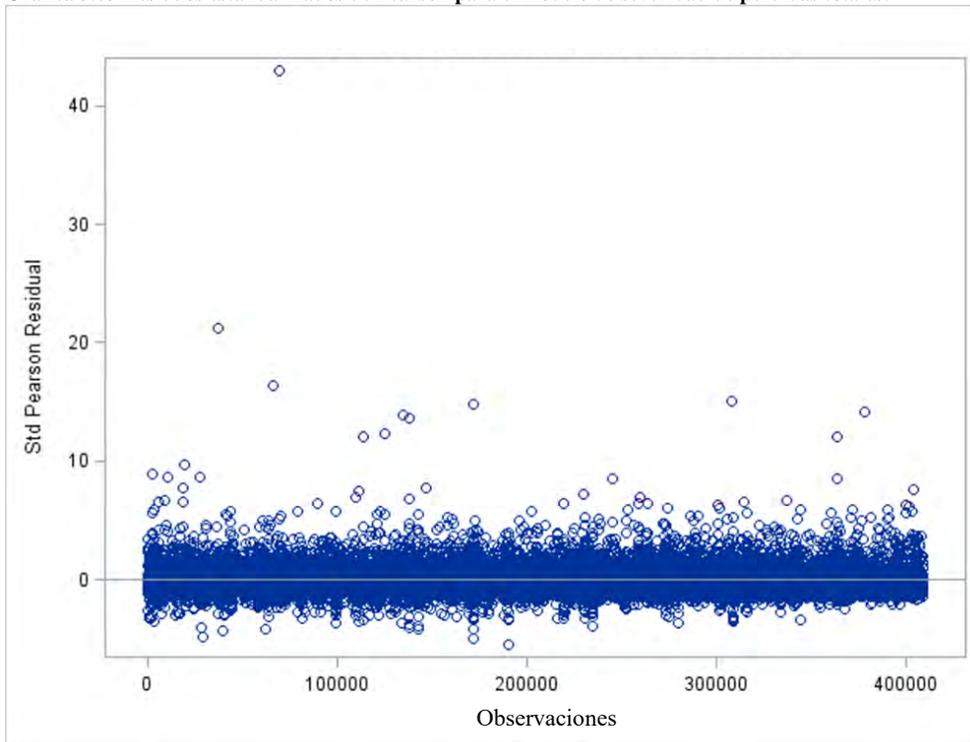


Gráfica 3.62 Residuos estandarizados de Devianza para el modelo de frecuencia de pérdidas totales.

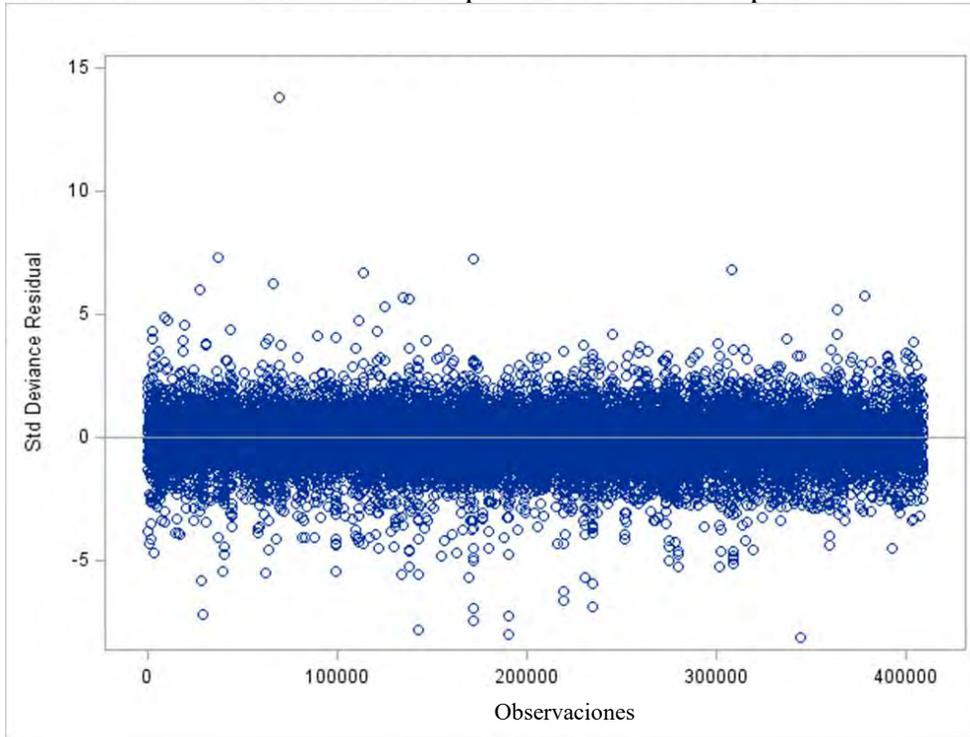


3.2.3.4 Severidad Pérdidas Totales

Gráfica 3.63 Residuos estandarizados de Pearson para el modelo de severidad de pérdidas totales.

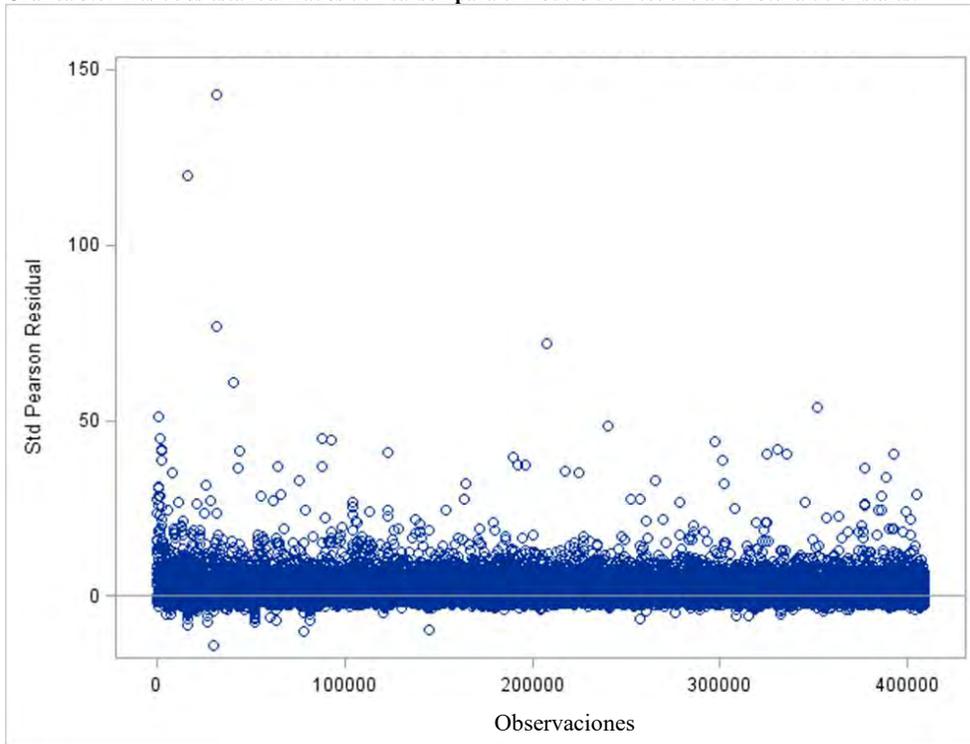


Gráfica 3.64 Residuos estandarizados de Devianza para el modelo de severidad de pérdidas totales.

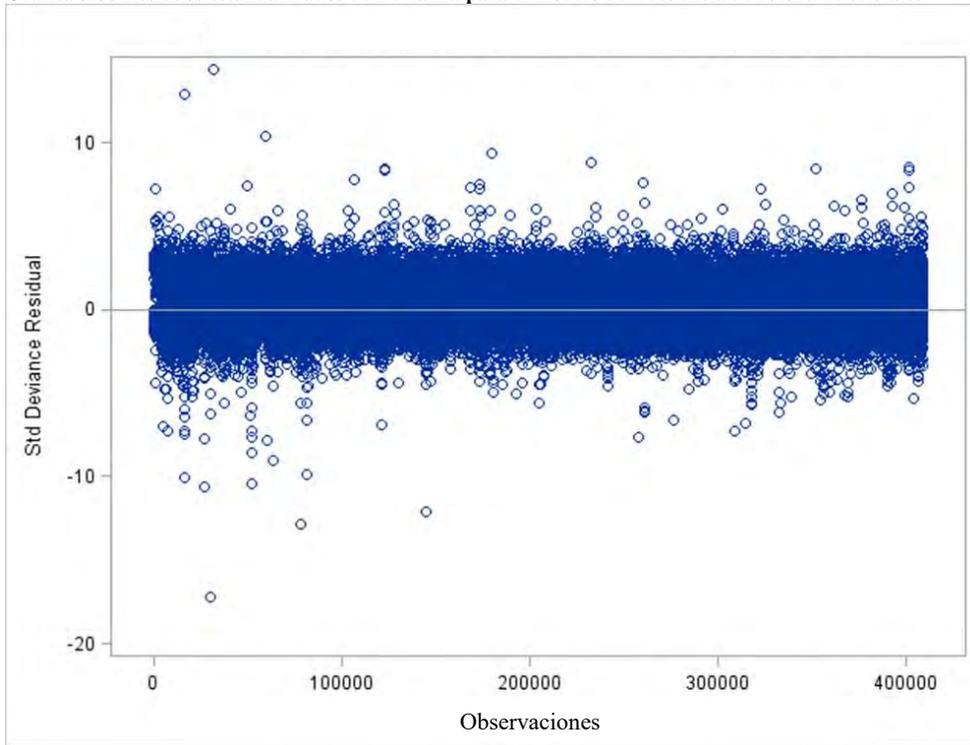


3.2.3.5 Frecuencia Rotura de Cristales

Gráfica 3.65 Residuos estandarizados de Pearson para el modelo de frecuencia de rotura de cristales.

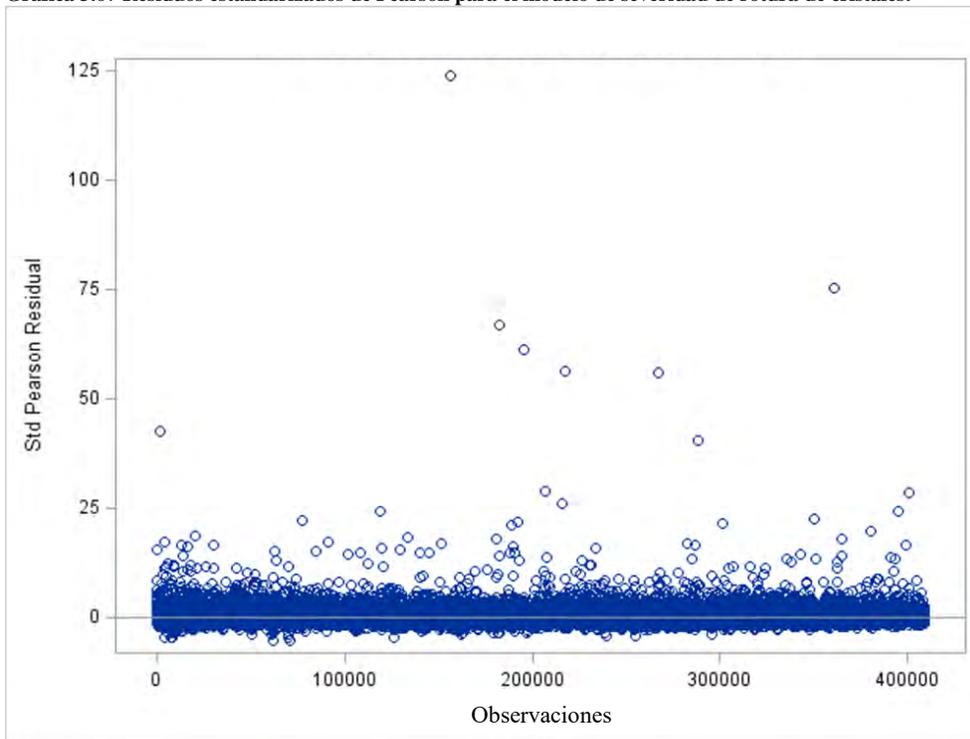


Gráfica 3.66 Residuos estandarizados de Devianza para el modelo de frecuencia de rotura de cristales.

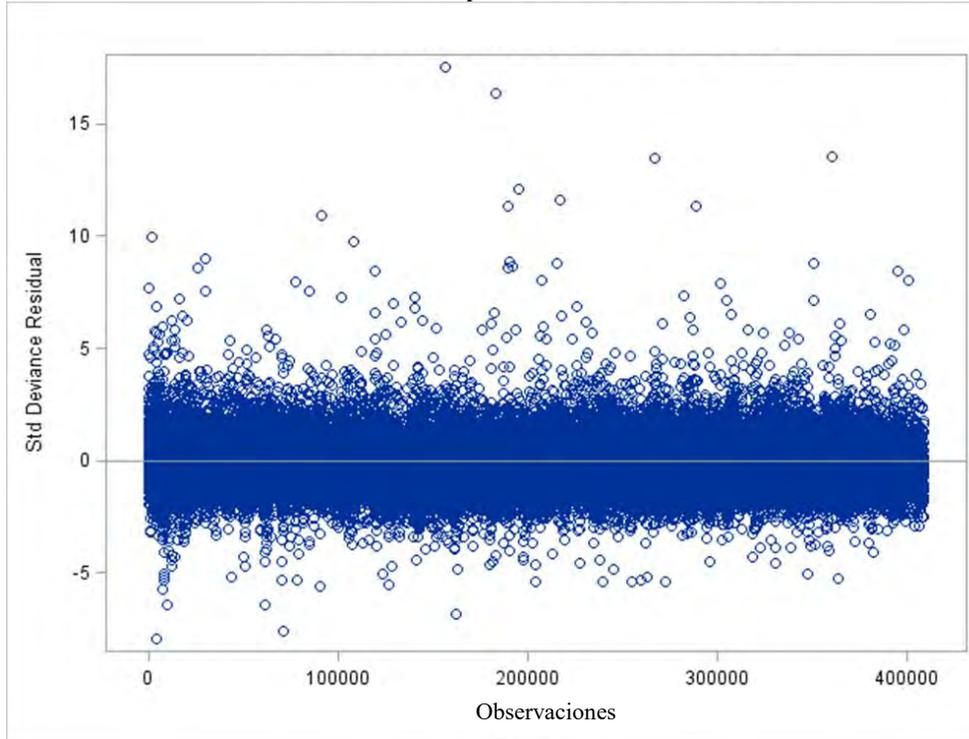


3.2.3.6 Severidad Rotura de Cristales

Gráfica 3.67 Residuos estandarizados de Pearson para el modelo de severidad de rotura de cristales.



Gráfica 3.68 Residuos estandarizados de Devianza para el modelo de severidad de rotura de cristales.



3.2.4 Curva de Lift

Este método compara la curva de predicciones del modelo contra la curva del promedio ponderado (por unidades expuestas o número de siniestros, para frecuencia o severidad, según corresponda) observado de la variable dependiente. Podemos decir que el resultado de este análisis es satisfactorio si los registros con menor frecuencia o severidad de la curva del modelo son también aquellos con menor promedio ponderado observado, es decir, ordenando de manera creciente los datos con respecto a la curva de las predicciones esperaríamos que la curva del promedio ponderado observado fuese creciente en la mayoría de los casos.

Cabe señalar que la curva de Lift no expone de manera puntual alguna desviación, sin embargo pone en evidencia de manera general un buen o mal ajuste.

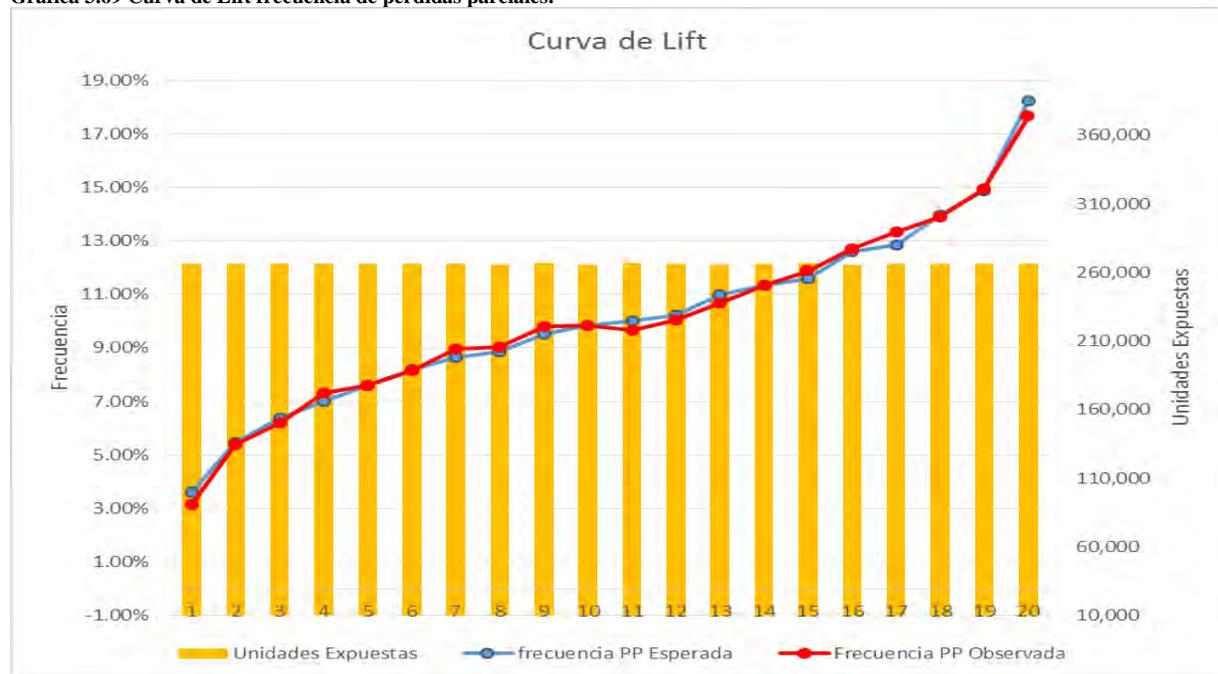
Los resultados obtenidos al realizar este test en los modelos fueron satisfactorios tal y como el lector podrá comprobar a continuación.

3.2.4.1 Frecuencia Pérdidas Parciales

Tabla 3.82 Intervalos de la curva de Lift para la frecuencia de pérdidas parciales.

Grupo	Unidades Expuestas	Número de Sinistros PP	Frecuencia PP Observada	Frecuencia PP Esperada
1	266,086	8,320	3.16%	3.62%
2	266,096	14,452	5.40%	5.45%
3	266,079	16,427	6.19%	6.37%
4	266,021	19,445	7.32%	6.99%
5	266,171	20,386	7.60%	7.59%
6	266,087	21,315	8.17%	8.18%
7	266,076	24,314	8.98%	8.65%
8	265,608	23,504	9.04%	8.85%
9	266,580	25,154	9.81%	9.52%
10	265,274	26,408	9.85%	9.84%
11	266,781	26,954	9.66%	10.00%
12	266,232	27,283	10.06%	10.23%
13	265,675	28,845	10.66%	10.98%
14	266,146	29,111	11.34%	11.33%
15	266,438	30,679	11.87%	11.57%
16	265,632	33,680	12.69%	12.62%
17	266,329	36,046	13.34%	12.83%
18	266,266	36,658	13.91%	13.97%
19	266,138	40,239	14.93%	14.86%
20	266,103	47,323	17.65%	18.25%

Gráfica 3.69 Curva de Lift frecuencia de pérdidas parciales.



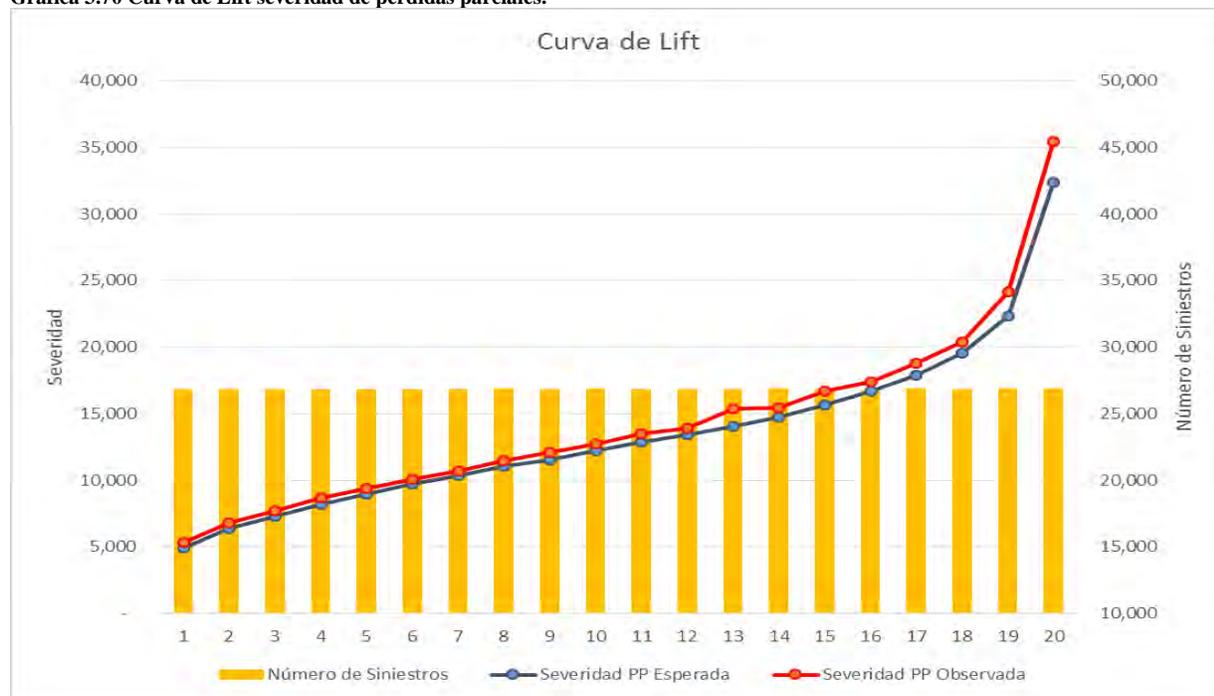
A lo largo del eje “x” de la gráfica 3.69 se levantan los veinte distintos grupos para modelo de frecuencia de pérdidas parciales, las barras amarillas representan la cantidad de unidades expuestas que posee cada grupo, en este caso se procura que todos los grupos posean el mismo número de unidades expuestas, y están medidos por el eje derecho “y”. La línea roja (frecuencia observada del conjunto de datos) y la línea azul (frecuencia derivada de ajustar el modelo) se miden con el eje izquierdo “y” de la gráfica. Claramente y en todos los casos se observa una tendencia creciente en ambas líneas. El número de grupos, en este caso veinte, estará en función de la cantidad de unidades expuestas (frecuencia) o número de siniestros (severidad); sin embargo, una media estándar en la práctica actuarial coincide con el número seleccionado para este trabajo escrito.

3.2.4.2 Severidad Pérdidas Parciales

Tabla 3.83 Intervalos de la curva de Lift para la severidad de pérdidas parciales.

Grupo	Número de Siniestros PP	Monto Siniestros de Pérdida Parciales	Severidad PP Observada	Severidad PP Esperada
1	26,824	142,584,953	5,316	4,946
2	26,829	181,619,985	6,770	6,403
3	26,828	206,663,260	7,703	7,316
4	26,818	232,985,057	8,688	8,197
5	26,830	251,171,958	9,362	8,967
6	26,832	271,075,162	10,103	9,715
7	26,800	287,239,582	10,718	10,355
8	26,855	308,488,968	11,487	11,032
9	26,775	324,317,029	12,113	11,548
10	26,873	341,784,042	12,718	12,221
11	26,821	362,120,138	13,501	12,859
12	26,827	373,251,702	13,913	13,397
13	26,809	411,639,660	15,355	14,050
14	26,849	414,737,015	15,447	14,753
15	26,832	447,990,500	16,696	15,622
16	26,755	465,671,464	17,405	16,677
17	26,872	504,986,786	18,792	17,867
18	26,823	546,860,031	20,388	19,557
19	26,843	647,381,938	24,117	22,351
20	26,848	952,299,145	35,470	32,355

Gráfica 3.70 Curva de Lift severidad de pérdidas parciales.



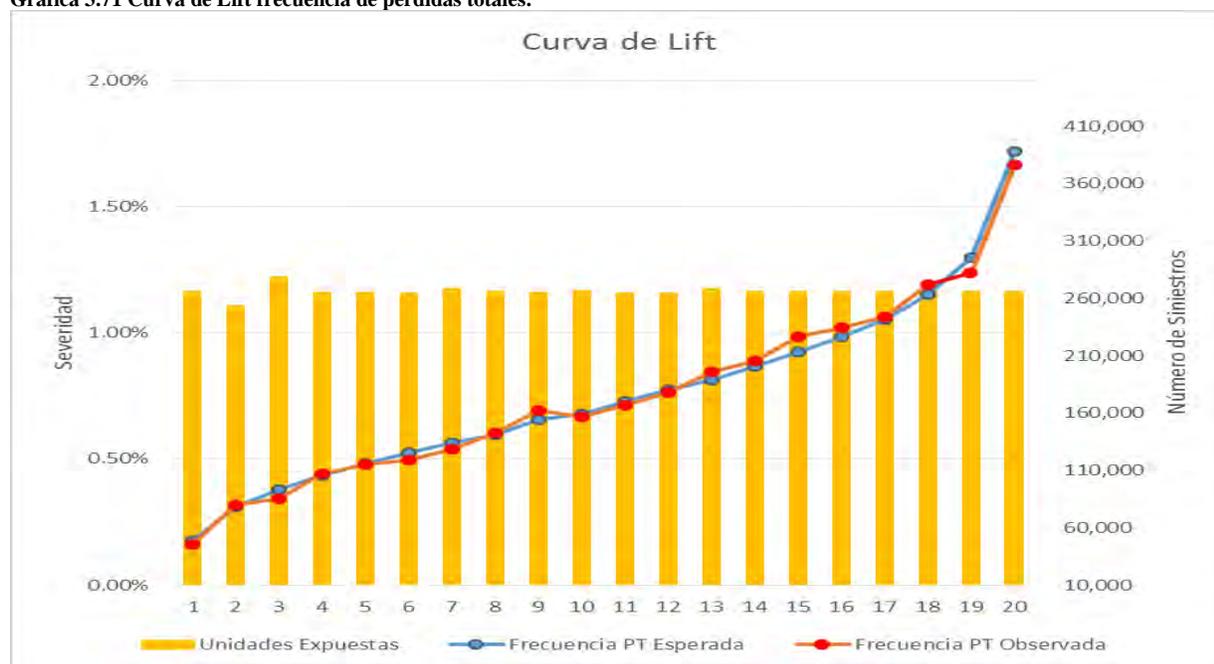
3.2.4.3 Frecuencia Pérdidas Totales

Tabla 3.84 Intervalos de la curva de Lift para la frecuencia de pérdidas totales.

Grupo	Unidades Expuestas	Número de Siniestros PT	Frecuencia PT Observada	Frecuencia PT Esperada
1	266,061	433	0.16%	0.18%
2	253,678	804	0.32%	0.31%
3	278,529	957	0.34%	0.38%
4	265,413	1,176	0.44%	0.43%
5	265,754	1,273	0.48%	0.48%
6	264,502	1,314	0.50%	0.52%
7	268,666	1,445	0.54%	0.56%
8	266,045	1,605	0.60%	0.60%
9	265,492	1,842	0.69%	0.66%
10	266,757	1,781	0.67%	0.68%
11	264,766	1,894	0.72%	0.73%
12	264,779	2,019	0.76%	0.78%
13	268,653	2,267	0.84%	0.81%
14	266,170	2,364	0.89%	0.87%
15	266,060	2,615	0.98%	0.92%
16	266,121	2,710	1.02%	0.99%
17	266,022	2,830	1.06%	1.05%
18	266,109	3,166	1.19%	1.15%

19	266,136	3,293	1.24%	1.30%
20	266,103	4,430	1.66%	1.72%

Gráfica 3.71 Curva de Lift frecuencia de pérdidas totales.



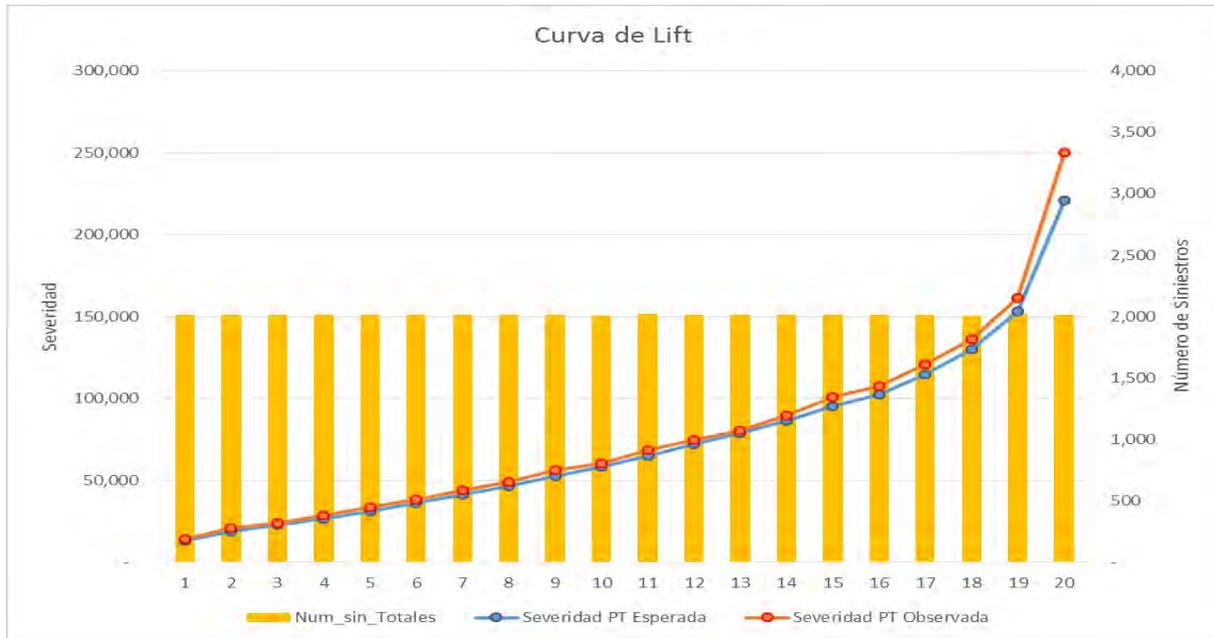
3.2.4.4 Severidad Pérdidas Totales

Tabla 3.85 Intervalos de la curva de Lift para la severidad de pérdidas totales.

Grupo	Número de Siniestros PT	Monto de Siniestros de Pérdida Total	Severidad PT Observada	Severidad PT Esperada
1	2,010	28,739,543	14,298	13,236
2	2,011	41,954,507	20,863	18,762
3	2,011	48,513,143	24,124	22,795
4	2,009	57,943,446	28,842	26,787
5	2,013	67,503,168	33,534	31,359
6	2,011	76,585,048	38,083	36,105
7	2,011	88,145,737	43,832	41,367
8	2,010	98,812,770	49,161	46,752
9	2,009	112,848,933	56,172	52,782
10	2,003	120,951,378	60,385	58,092
11	2,021	138,203,310	68,384	65,208
12	2,011	150,002,375	74,591	72,030
13	2,011	161,283,223	80,201	78,695
14	2,011	180,427,329	89,720	86,369
15	2,010	202,231,929	100,613	95,091
16	2,010	216,273,552	107,599	102,593
17	2,013	242,881,086	120,656	114,805

18	2,005	273,402,525	136,360	129,866
19	2,017	325,236,468	161,248	153,158
20	2,011	502,596,246	249,924	220,566

Gráfica 3.72 Curva de Lift severidad de pérdidas totales.



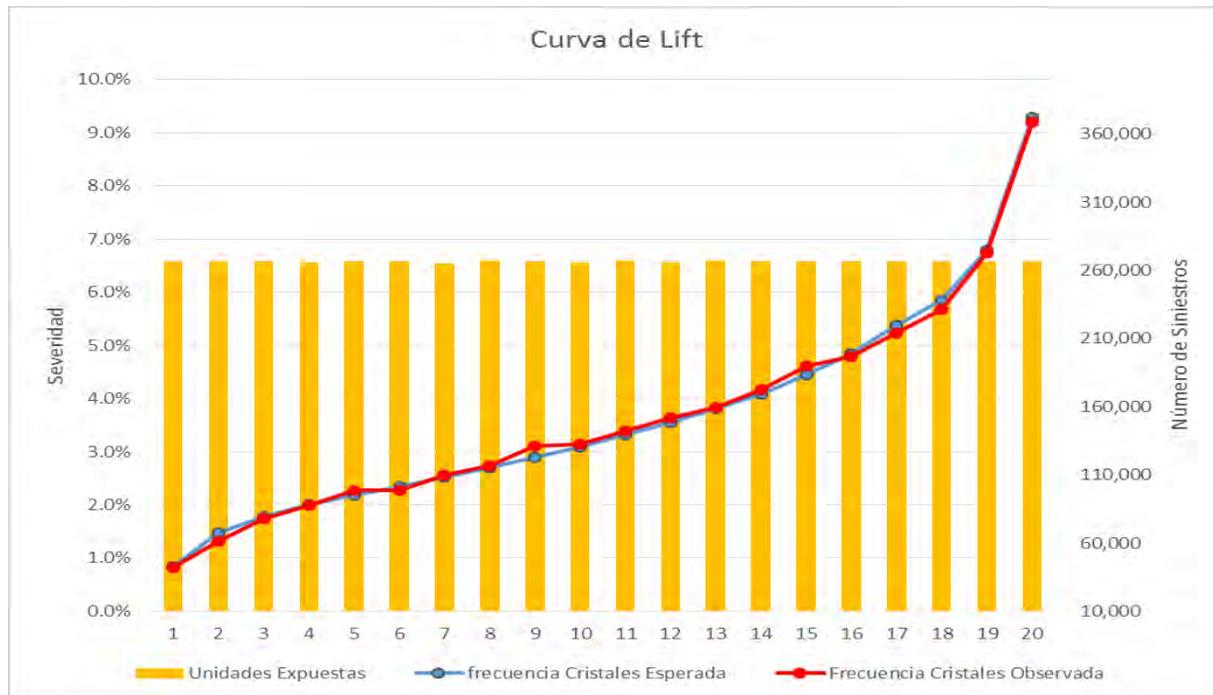
3.2.4.5 Frecuencia Rotura de Cristales

Tabla 3.86 Intervalos de la curva de Lift para la frecuencia de rotura de cristales.

Grupo	Unidades Expuestas	Número de Siniestros de Rotura de Cris.	Frecuencia Rotura de Cris. Observada	Frecuencia Rotura de Cris. Esperada
1	266,068	2,200	0.8%	0.8%
2	266,113	3,522	1.3%	1.5%
3	265,912	4,612	1.7%	1.8%
4	265,838	5,302	2.0%	2.0%
5	266,505	6,074	2.3%	2.2%
6	266,101	6,043	2.3%	2.3%
7	265,142	6,763	2.6%	2.5%
8	267,021	7,279	2.7%	2.7%
9	265,980	8,251	3.1%	2.9%
10	265,560	8,352	3.1%	3.1%
11	266,705	9,042	3.4%	3.3%
12	265,606	9,681	3.6%	3.6%
13	266,628	10,205	3.8%	3.8%
14	266,089	11,092	4.2%	4.1%
15	265,991	12,246	4.6%	4.5%

16	266,081	12,745	4.8%	4.8%
17	266,141	13,937	5.2%	5.4%
18	266,155	15,096	5.7%	5.9%
19	266,006	17,941	6.7%	6.8%
20	266,177	24,482	9.2%	9.3%

Gráfica 3.73 Curva de Lift frecuencia de rotura de cristales.



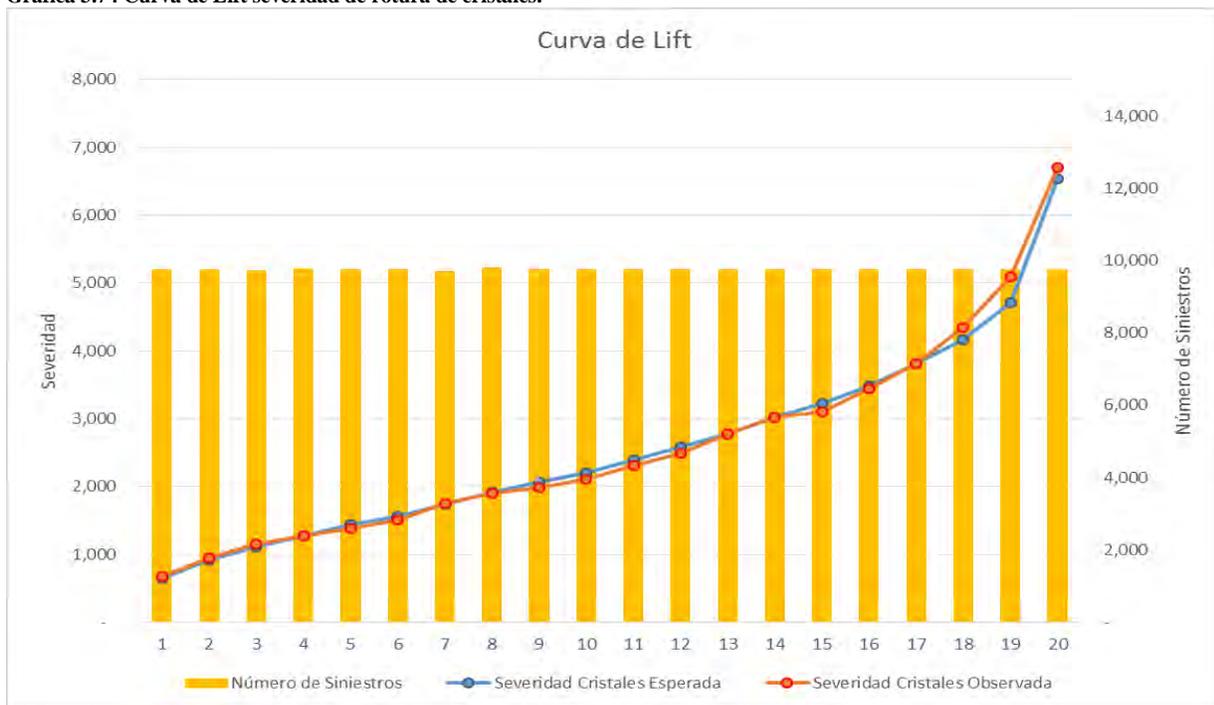
3.2.4.6 Severidad Rotura de Cristales

Tabla 3.87 Intervalos de la curva de Lift para la severidad de rotura de cristales.

Grupo	Número de Siniestros de Rotura de Cris	Monto Siniestros de Rotura de Cris.	Severidad Rotura de Cris. Observada	Severidad Rotura de Cris. Esperada
1	9,739	6,521,129	670	643
2	9,747	9,240,494	948	922
3	9,728	11,146,714	1,146	1,112
4	9,758	12,473,133	1,278	1,280
5	9,741	13,515,650	1,388	1,431
6	9,745	14,649,302	1,503	1,562
7	9,686	16,915,694	1,746	1,736
8	9,797	18,661,794	1,905	1,920
9	9,746	19,301,535	1,980	2,068
10	9,736	20,560,196	2,112	2,198
11	9,746	22,537,923	2,313	2,389
12	9,747	24,203,825	2,483	2,579

13	9,743	27,041,021	2,775	2,777
14	9,743	29,480,354	3,026	3,015
15	9,739	30,257,026	3,107	3,218
16	9,746	33,580,523	3,446	3,480
17	9,744	37,126,625	3,810	3,816
18	9,743	42,281,813	4,340	4,160
19	9,743	49,630,368	5,094	4,715
20	9,748	65,365,201	6,705	6,533

Gráfica 3.74 Curva de Lift severidad de rotura de cristales.



3.2.5 Conclusiones

Las cuatro pruebas analizadas a lo largo de este capítulo fueron satisfactorias:

- Los intervalos de confianza demostraron ser en todos los casos lo suficientemente angostos como para evitar pensar que los parámetros estimados tuviesen un comportamiento impredecible, confirmándose el hecho de que a mayor cantidad de unidades expuestas-número de siniestros, menor será la volatilidad.
- El test de consistencia aleatoria demostró que las variaciones entre la estimación de los parámetros para las tres particiones de los datos corresponden a fluctuaciones del orden aleatorio y que las fluctuaciones aleatorias crecen con forme el peso (unidades expuestas

o número de siniestros) de cada nivel desciende, dicho de otro modo se concluye que cada una de las agrupaciones es consistente con el comportamiento del riesgo.

- Se analizaron dos tipos de residuos, de Pearson y de Devianza en ambos casos los resultados demostraron ser satisfactorios, siendo los modelos de pérdidas totales aquellos con un mayor número de observaciones alejadas del cero. Sin embargo esto no desacredita el poder predictivo de los modelos pues al tratarse de un fenómeno con poco número de observaciones (en comparación con los siniestros de pérdidas parciales) es natural encontrarse con esta clase de comportamiento, mismo que es aceptado en la práctica actuarial.
- Finalmente se exploraron las distintas formas de la Curva de Lift en donde se confirmó la tendencia creciente para los veinte grupos de cada modelo. Por tanto ajustan con un grado de confianza aceptable.

CAPÍTULO 4

En el capítulo anterior se calculó la frecuencia y severidad para las tres particiones de la cobertura de Daños Materiales, finalmente se amalgamarán los seis modelos para el cálculo de la prima de riesgo y eventual derivación de la prima de tarifa.

4.1 Modelo de Daños Materiales

La prima de riesgo para la cobertura de Daños Materiales será la suma de la prima de riesgo de las particiones por tipo de pérdida (pérdidas parciales, pérdidas totales y rotura de cristales) realizadas. Para cada partición la prima de riesgo es el producto del modelo de frecuencia por el modelo de severidad. La fórmula que refleja el cálculo de la prima de riesgo tal y como se acaba de mencionar está dada por:

$$PR_{DM} = (freq_{PP} * sev_{PP}) + (freq_{PT} * sev_{PT}) + (freq_{RCri} * sev_{RCri}), \quad (4.1)$$

Donde:

PR_{DM} : Prima de Riesgo base de la Cobertura Daños Materiales.

$freq_{PP}$: Frecuencia base del modelo de Pérdidas Parciales.

sev_{PP} : Severidad base del modelo de Pérdidas Parciales.

$freq_{PT}$: Frecuencia base del modelo de Pérdidas Totales.

sev_{PT} : Severidad base del modelo de Pérdidas Totales.

$freq_{RCri}$: Frecuencia base del modelo de Rotura de Cristales.

sev_{RCri} : Severidad base del modelo de Rotura de Cristales.

Tabla 4.1 Resumen numérico de la prima base de los modelos.

Indicador	P.P.	P.T.	R. Cri.
Frecuencia	10.13%	0.55%	2.85%
Severidad	12,789.8	79,522.9	1,548.9
Prima de Riesgo=Frecuencia*Severidad	1,295.4	434.9	44.1

Antes de aplicar por completo (4.1) para el cálculo de la prima de riesgo de la cobertura de Daños Materiales es necesario mutualizar los datos que excluimos durante la preparación de los datos.

4.2 Mutualización de los Datos Excluidos

Durante la fase de preparación de los datos se dejaron fuera siniestros con el fin de mejorar el ajuste de los modelos, sin embargo no podemos omitir del cálculo de la prima esa información. Por tanto, esta sección abordará la metodología para la mutualización de estos datos.

Como primer paso debemos ser claros en dos puntos:

- Cuál fue la información inicial previamente definida en el alcance de los modelos.
- Cuál fue la información empleada en cada modelo.

Tabla 4.2 Resumen de información inicial y final para pérdidas parciales.

<i>Pérdidas Parciales</i>	<i>Unidades Expuestas</i>	<i>Número de Siniestros</i>	<i>Monto de Siniestros</i>	<i>Prima de Riesgo</i>
<i>Información Inicial</i>	5,321,818	650,185	7,008,954,809	1,317.0
<i>Información usada en el modelo</i>	5,321,818	536,543	7,674,868,375	1,442.2

En este caso podemos observar que la información excluida por concepto de montos negativos representa 8.68% de los montos incluidos en el modelo. El factor de ajuste a la prima base de pérdidas parciales es de $\frac{1,317.0}{1,442.2} = 0.91323$. Un factor de descuento no es una sorpresa puesto que los casos en donde el monto de recuperación supera al monto ocurrido han aumentado a consecuencia de la entrada en vigor del acuerdo SIPAC (Sistema de Pago entre Compañías).

Tabla 4.3 Resumen de información inicial y final para pérdidas totales.

<i>Pérdida Total</i>	<i>Unidades Expuestas</i>	<i>Número de Siniestros</i>	<i>Monto de Siniestros</i>	<i>Prima de Riesgo</i>
<i>Información Inicial</i>	5,321,818	43,073	3,209,053,810	603.0
<i>Información usada en el modelo</i>	5,321,818	40,218	3,134,535,716	589.0

El factor de ajuste a la prima base de pérdidas totales es de $\frac{603}{589} = 1.02377$.

Tabla 4.4 Resumen de información inicial y final para rotura de cristales.

<i>Rotura de Cristales</i>	<i>Unidades Expuestas</i>	<i>Número de Siniestros</i>	<i>Monto de Siniestros</i>	<i>Prima de Riesgo</i>
<i>Información Inicial</i>	5,321,818	197,145	512,693,343	96.3
<i>Información usada en el modelo</i>	5,321,818	194,865	504,490,321	94.8

El factor de ajuste a la prima base de rotura de cristales es de $\frac{96.3}{94.8} = 1.01626$.

Después de calcular los factores debemos incluirlos a la prima base de cada partición por tipo de pérdida, multiplicando la prima de riesgo resultante de los modelos por el factor de ajuste, según corresponda.

Tabla 4.5 Resumen de la mutualización por tipo de pérdida.

	<i>Prima Base</i>	<i>Pérdida Parcial</i>	<i>Pérdida Total</i>	<i>Rotura de Cristales</i>
<i>Antes de la mutualización</i>	1,295.43	434.89	44.13	
<i>Después de la mutualización</i>	1,183.03	445.23	44.85	

El siguiente paso será incluir la estructura de costos.

4.3 Inclusión de las Hipótesis Financieras

La prima de tarifa se define¹⁵ como el monto unitario necesario para cubrir un riesgo, comprendiendo los costos esperados de siniestralidad y otras obligaciones contractuales, de adquisición y de administración, así como el margen de utilidad previsto.

En virtud de la anterior definición se deben incluir tres conceptos al cálculo previo, mismos que el Estándar de Práctica Actuarial No. 1 define como:

- **Gastos de Administración:** Son los relativos a la suscripción, emisión, cobranza, administración, control y cualquier otra función necesaria para el manejo operativo de una cartera de seguros de corto plazo. Dentro del Sector Asegurador Mexicano éstos oscilan entre 8% y 10%.
- **Costos de adquisición:** Son los relacionados con la promoción y venta de los seguros, que incluyen comisiones a intermediarios, bonos y otros gastos comprendidos dentro de este rubro. Dentro del Sector Asegurador Mexicano éstos oscilan entre 17% y 20%.
- **Margen de utilidad:** Es la contribución marginal a la utilidad bruta general, que se haya definido para el ramo y tipo de seguro en cuestión, de conformidad con las políticas establecidas por la empresa que asume el riesgo. Comúnmente se añade 5%, sin embargo puede variar en función de las políticas de cada compañía aseguradora.

Éstos son calculados como una parte proporcional de la prima de riesgo y la forma de recargarlos está dada por:

$$PT = \frac{PR}{(1 - GA - CA - MU)}, \quad (4.2)$$

Donde:

PT: Es la Prima de Tarifa.

PR: Es la Prima de Riesgo.

GA: Son los Gastos de Administración.

CA: Son los Costos de Adquisición.

MU: Es el Margen de Utilidad.

¹⁵ Definición de Estándares de Práctica Actuarial de primas de la Asociación Mexicana de Actuarios [1]

Aplicando el punto medio de los porcentajes para cada uno de los tres conceptos ($GA = 9\%$, $CA = 18.5\%$ y $MU = 5\%$) a la prima de riesgo se obtiene el siguiente cuadro:

Tabla 4.6 Resumen de la prima de tarifa por tipo de pérdida.

	<i>Prima Base</i>	<i>Pérdida Parcial</i>	<i>Pérdida Total</i>	<i>Rotura de Cristales</i>
<i>Prima de Riesgo Ajustada</i>	1,183.03	445.23	44.85	
<i>Prima de Tarifa</i>	1,752.64	659.60	66.44	

4.4 Resumen Numérico

Al finalizar el cálculo de la prima de tarifa se recomienda elaborar un resumen numérico que despliegue las primas de tarifa promedio por entidad federativa, sub-marca x modelo, etc. Todos estos resultados serán fundamentales en el momento en que el actuario pretenda dar a conocer los niveles de precios al departamento comercial y a los agentes de una compañía aseguradora.

En este punto es recomendable prestar mayor atención a las entidades o sub-marcas con un mayor peso en la cartera.

Tabla 4.7 Resumen de la prima de tarifa de la cobertura Daños Materiales por Entidad Federativa.

<i>Entidad Federativa</i>	<i>Unidades Expuestas</i>	<i>Prima de Tarifa Promedio DM</i>
<i>Distrito Federal</i>	1,104,243	1,772
<i>Nuevo León</i>	537,111	2,452
<i>Estado de México</i>	498,708	2,405
<i>Jalisco</i>	482,410	2,595
<i>Veracruz</i>	227,221	2,988
<i>Chihuahua</i>	202,099	1,904
<i>Guanajuato</i>	188,681	2,600
<i>Puebla</i>	169,779	2,544
<i>Querétaro</i>	144,906	2,538
<i>Coahuila</i>	144,549	2,759
<i>Tamaulipas</i>	141,248	2,467
<i>Sonora</i>	122,710	2,392
<i>Baja California</i>	122,562	1,840
<i>Yucatán</i>	121,105	2,857
<i>Michoacán</i>	116,245	2,774
<i>Sinaloa</i>	107,628	4,508
<i>San Luis Potosí</i>	101,278	2,396
<i>Tabasco</i>	78,767	4,275
<i>Morelos</i>	77,927	2,943
<i>Quintana Roo</i>	77,009	3,519
<i>Aguascalientes</i>	68,808	1,912
<i>Chiapas</i>	66,711	3,299

<i>Hidalgo</i>	61,329	3,108
<i>Durango</i>	50,933	2,323
<i>Oaxaca</i>	48,772	3,480
<i>Guerrero</i>	42,783	4,092
<i>Colima</i>	38,239	2,548
<i>Campeche</i>	36,563	3,538
<i>Baja California Sur</i>	35,202	4,407
<i>Zacatecas</i>	34,087	2,511
<i>Nayarit</i>	34,018	2,551
<i>Tlaxcala</i>	19,757	3,828

Podemos observar que Sinaloa posee el mayor monto de prima promedio, mientras que el Distrito Federal se ubica en el nivel más bajo de prima.

Tabla 4.8 Resumen de la prima de tarifa para las principales marcas.

<i>Marca</i>	<i>Unidades Expuestas</i>	<i>Prima de Tarifa Promedio DM</i>
<i>GENERAL MOTORS</i>	1,014,451	2,001
<i>NISSAN</i>	904,107	2,180
<i>VOLKSWAGEN</i>	641,606	2,348
<i>FORD</i>	573,741	1,963
<i>CHRYSLER</i>	483,384	1,876
<i>TOYOTA</i>	308,485	3,571
<i>HONDA</i>	262,002	2,684
<i>MAZDA</i>	125,554	4,952
<i>RENAULT</i>	95,072	3,365
<i>SEAT</i>	93,097	3,909
<i>BMW</i>	62,282	6,907

El nivel más alto de prima lo ocupa BMW debido a lo costoso que resulta la reparación de estos vehículos, mientras que el segundo lugar lo ocupa MAZDA, marca relativamente nueva en el mercado mexicano. Los niveles más bajos son para las marcas más comerciales cuyo parque vehicular también incluye vehículos de más antigüedad.

Tabla 4.9 Resumen de la prima de tarifa para las principales Sub-marcas.

<i>Sub-marca</i>	<i>Unidades Expuestas</i>	<i>Prima de Tarifa Promedio DM</i>
<i>JETTA/CLASICO</i>	236,827	1,874
<i>AVEO</i>	170,372	2,551
<i>SENTRA</i>	153,006	1,940
<i>TSURU</i>	140,165	1,651
<i>TIIDA</i>	107,725	3,209
<i>CHEVY</i>	93,372	1,039
<i>CHEVY C2</i>	90,272	1,863
<i>FIESTA</i>	85,740	2,288
<i>CRV</i>	83,546	2,607
<i>VERSA</i>	80,695	3,416

SPARK	66,538	3,132
ESTACAS	66,204	1,900
IBIZA	66,174	3,711
ESCAPE	61,158	1,932
FOCUS	60,872	2,817
CIVIC	59,312	2,148
VOYAGER	58,277	1,243
COROLLA	58,264	3,293

El Ibiza es un vehículo que debido a su perfil objetivo, personas jóvenes, posee un particular mal resultado para la cobertura de Daños Materiales, suerte que comparte junto con el Spark. Tanto el Sentra como el Tsuru poseen un nivel bajo de prima en la cobertura de Daños Materiales, muy contrario a lo que se esperaría de la cobertura de Robo Total.

Tabla 4.10 Resumen de la prima de tarifa por modelo.

<i>Modelo</i>	<i>Unidades Expuestas</i>	<i>Prima de Tarifa Promedio DM</i>
2015	52,561	4,554
2014	482,978	4,783
2013	625,536	4,649
2012	536,589	4,321
2011	414,105	3,821
2010	356,533	3,154
2009	327,981	3,085
2008	382,701	2,813
2007	375,454	2,675
2006	305,714	2,422
2005	306,010	2,147
2004	220,242	1,984
2003	179,091	1,775
2002	167,802	1,640
2001	147,933	1,473
2000	118,040	1,249

Se ve una clara tendencia decreciente en los niveles de prima promedio por modelo. Los modelos más recientes parecerían no cumplir con esta regla debido a que el costo de reparación de los modelo más recientes, 2015-2013, sea exactamente el mismo puesto que éstos comparten refacciones, no obstante sucede que los dueños de vehículos último modelo cuidan más a su unidad, mientras que los dueños con vehículos con uno o dos años de antigüedad son más descuidados con su unidad.

Tabla 4.11 Resumen de la prima de tarifa por tipo de vehículo.

<i>Tipo de Vehículo</i>	<i>Unidades Expuestas</i>	<i>Prima de Tarifa Promedio DM</i>
<i>AUTOMOVILES</i>	4,634,270.5	2,916.3
<i>COMERCIALES</i>	687,547.4	2,252.5

El costo promedio de los automóviles es más alto debido a que resultan más afectados al momento de un accidente que los comerciales (Pick Ups).

CONCLUSIONES Y COMENTARIOS

Los Modelos Lineales Generalizados han aportado una alternativa más sofisticada para el cálculo de primas del Sector Asegurador Mexicano por su capacidad de asignar distintos niveles de riesgo a partir de las características de cada póliza suscrita, de este modo el cálculo del nivel de prima para cada póliza será más justo y así se evadirá la anti-selección, situación en donde la Compañía Aseguradora no es capaz de distinguir la calidad del riesgo suscrito. La verosimilitud de lo anterior está siempre en función de la cantidad de observaciones homogéneas empleadas para la creación del modelo y de la inclusión de variables representativas y no correlacionadas al modelo.

Si bien la sofisticación de esta metodología provee una de las mejores vistas posibles para el entendimiento de costo esperado de la cobertura de Daños Materiales para cada póliza, no se está exento de una revisión con el departamento comercial, el equipo suscriptor y la fuerza de ventas de la Compañía Aseguradora para incluir todas aquellas restricciones de negocio y de mercado que garanticen la competitividad de precios antes de materializar los resultados a manos de los potenciales clientes.

El párrafo anterior marca el preámbulo de la siguiente fase del proceso de fijación de precios: la implementación en sistemas. Esta fase es el parteaguas entre el mundo abstracto y sin restricciones de los modelos matemáticos y el criterio comercial y de mercado del negocio necesarios para concretar una venta. Se deberá comenzar respondiendo la pregunta ¿El sistema de cálculo y emisión de primas de la Compañía Aseguradora es apto para implementar un modelo multiplicativo? Si la respuesta es no, estaremos ante una situación en donde la culminación de este trabajo escrito no rebasará el espacio académico, de caso contrario habría que analizar cuáles de las variables independientes (Entidad Federativa, Sub-marca, etc.) pueden ser reconocidas por el sistema como variables de tarificación, en caso de no contar en el sistema con todas aquellas variables incluidas en el cálculo de la prima, se deberán replantear un modelo cuya composición en términos de variables independientes sea completamente compatible con el sistema y reforzar la acción anterior con el monitoreo técnico para cubrir los sesgos relacionados con dicha acción.

Una vez garantizada la compatibilidad del modelo técnico con el sistema de la Compañía Aseguradora, la implementación del modelo derivado del ajuste no sólo consiste en adaptar las salidas que se han obtenido, sino también consiste en minimizar el impacto comercial y técnico por el cambio de tarifa, suavizando los cambios en precios en el mayor número de casos posibles y teniendo a mano una bien elaborada justificación cuando el cambio demande ser radical. Durante la implementación del modelo existirán principalmente tres Riesgos:

- Riesgos operacionales: El sistema necesitará un conjunto de archivos que le dicten cómo y en qué medida deberá recargar o descontar a la prima base en función de las variables de tarifa, un descuido en la carga del conjunto de archivos podría ocasionar un cálculo erróneo al momento de suscribir un negocio. Para mitigar esta posibilidad se recomienda una sesión de pruebas y validaciones que garanticen el correcto funcionamiento tanto de los archivos de tarifa como del algoritmo de cálculo.

- **Riesgos de mercado y comerciales:** Recordemos que el Sector Asegurador Mexicano posee una cantidad amplia de competidores que buscan aumentar su participación en el mercado, por tal razón no suena inverosímil que si los cambios en precios son drásticos la venta de pólizas y la eventual composición del portafolio se modifiquen dando espacio a desviaciones en el resultado. La recomendación es preparar un análisis de impacto por cambio de tarifa que permita identificar los sectores vulnerables de la cartera del Ramo de Automóviles de la Compañía Aseguradora ante el cambio de tarifa.
- **Riesgos Regulatorios:** La Comisión Nacional de Seguros y Fianzas deberá ser notificada ante un cambio en la metodología del cálculo de primas mediante la actualización o registro del producto correspondiente en el que la descripción de esta metodología quede claramente desarrollada y soportada en la Nota Técnica.

No obstante y contrario a lo que el lector pudiera adivinar, el trabajo no termina aquí con la creación del modelo que a sazón bautizaremos como modelo implementable, pues al finalizar todos los pasos previos habrá que tomar en cuenta el hecho de que los datos empleados son siempre un corte transversal de información, de donde surge la necesidad de actualizar los datos y revisar al menos una vez cada año la validez del ajuste de los modelos involucrados en el precio final de la cobertura de Daños Materiales.

La actividad de la que acabamos de hacer mención es conocida en el argot actuarial como monitoreo técnico y su intención es prevenir desviaciones que pudieran afectar la rentabilidad del negocio.

Finalmente se debe remarcar el hecho de que esta es una de las actividades laborales en donde se ponen en práctica y a prueba todas las virtudes aprendidas por el actuario durante su trayectoria académica, que no sólo hace falta limitarse a las lecciones y modelos aprendidos dentro de las aulas, pues el avance y sofisticación del gremio a nivel global nos sitúan en una constante curva de aprendizaje, y de que no existe un único modelo correcto, pues cada modelo dependerá del juicio del actuario y éste deberá hacer uso tanto de su experiencia de mercado como de su conocimiento matemático para garantizar la rentabilidad en la fijación del precio en cada póliza suscrita.

BIBLIOGRAFÍA

- Comité de Estándares de Práctica Actuarial de la Asociación Mexicana de Actuarios, A.C. 2003. Estándares de Práctica Actuarial No. 01.[1].
- Esbjörn Ohlsson and Björn Johansson. 2010. Non-Life Insurance Pricing with Generalized Linear Models. Springer [2].
- George Casella and Roger L Berger. Statistical Inference. Second Edition. DUXBURY pag. 85 [3].
- R.E. Beard. Risk Theory The Stochastic Basis of Insurance. Third Edition Chapman and Hall pag. 349.[4].
- Piet de Jong and Gillian Z. Heller 2007. Generalized Linear Models For Insurance Data [5].
- P. McCullagh and J.A. Nelder FRS. Generalized Linear Models Second Edition. Chapman and Hall [6].
- Noriszura Ismail and Abdul Aziz Jemain. Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models [7].
- Ley de Instituciones de Seguros y Fianzas Art 27. [8].
- Círculo Único de Seguros y Fianzas Título 26 [9].
- Manual del Sistema Estadístico del Ramo de Automóviles de Póliza Individual (Versión 01) 03-dic-2015 [10].
- Ley Sobre el Contrato de Seguro Art. 116[11].

GLOSARIO DE TÉRMINOS

Automóviles: Vehículo ligero destinado al transporte de personas, los cuales se segmentan en subcompacto, compacto, de lujo, deportivos y multiusos (minivans).

Camiones ligeros (CL1): Camiones ligeros (grupo uno) cuyo peso bruto vehicular es de hasta 2,722 Kg. y con peso de prueba (PP) de hasta 1,701 kg.

Camiones ligeros (CL2): Camiones ligeros (grupo dos) cuyo peso bruto vehicular es de hasta 2,722 Kg. y con peso de prueba (PP) mayor de 1,701 y hasta 2,608 Kg.

Camiones ligeros (CL3): Camiones ligeros (grupo tres) cuyo peso bruto vehicular es mayor de 2,722 y hasta 3,856 kg y con peso de prueba (PP1) de hasta 2,608 kg.

Celda de tarifa: Es un subconjunto de una cartera de pólizas, cuyos elementos poseen características de riesgo similares entre sí.

Comerciales: Camiones ligeros (Clase 1 a 3) con capacidad de carga hasta 4.5 toneladas, dentro de esta clasificación se encuentran las pick ups.

Compacto: Unidades con una distancia entre ejes de 2,476 hasta 2,700 mm; con motores de 4 ó 6 cilindros de 2500 hasta cm³ de desplazamiento.

Deportivo: Unidades con una distancia entre ejes de 2,601 hasta 3,000 mm; con motores de 4, 6 u 8 cilindros de 2000 a 4600 cm³ de desplazamiento.

Estimador insesgado: Estimador cuya esperanza es igual al parámetro que se desea estimar.

Exposición de la póliza: Es la proporción de tiempo que una póliza estuvo vigente dentro del periodo de análisis. Para efectos de este trabajo el periodo de análisis es de un año.

Frecuencia de siniestros: Es el número de siniestros dividido entre la exposición, para algún grupo de pólizas vigentes durante un periodo específico de tiempo, i.e., el número promedio de siniestros en un periodo de tiempo (usualmente un año).

Interacción: Sea Y la variable dependiente en un modelo lineal generalizado, X y Z variables independientes, se dice que entre X y Z existe una interacción si el efecto de X sobre Y no es el mismo al margen de los valores de Z , sino que tal efecto depende del valor o nivel de Z .

Lujo: Unidades con una distancia entre ejes de 2,601 hasta 3,000 mm; con motores de 4, 6, 8 ó 12 cilindros de 3200 a 6000 cm³ de desplazamiento.

Modelo saturado: Modelo cuyo número de parámetros es igual al número de observaciones.

Monitoreo técnico: Proceso mediante el cual se verifica el ajuste de los modelos para una muestra de información distinta a la empleada para el ajuste original de los modelos.

Multiusos: Vehículo de uso múltiple o utilitario (SUV o VU): Vehículo automotor diseñado para el transporte de personas y/o productos, con o sin chasis o con equipo especial para operar ocasionalmente fuera del camino.

Mutualizar: Repartir las posibles pérdidas derivadas de un riesgo determinado entre una cantidad finita de participantes expuestas a dicho riesgo.

Póliza: Contrato con vigencia específica que contiene los acuerdos entre asegurado y aseguradora.

Prima de riesgo: Es el monto total de siniestros dividido entre la exposición, i.e., el costo promedio por póliza anual (o cualquier otro periodo de tiempo). Por tanto, la prima de riesgo es el producto de la frecuencia por la severidad de los siniestros.

| *Severidad de siniestro:* Es el monto total de siniestros dividido entre el número de siniestros cuando éstos ocurran, i.e., el costo promedio de los siniestros. Frecuentemente se omiten las reclamaciones nulas de la frecuencia y severidad de los siniestros. Es importante ser consistente en esta parte ya sea al omitir o incluir las reclamaciones nulas en ambos casos.

Siniestro: Es un evento reportado por el asegurado, por el cuál, éste exige una compensación económica que resarza el daño de acuerdo con las Condiciones Generales de la póliza.

Subcompacto: Unidades con una distancia entre ejes hasta 2,475 mm; con un motor de 4 cilindros de hasta 1800 cm³ de desplazamiento.