

Máster Universitario Ciencias Actuariales y Financieras
2017-2018

Trabajo Fin de Máster

**“Modelo actuarial sobre la
percepción de la calidad del servicio
en las prestaciones de hogar e
influencia en la fidelización del
cliente”**

Sabina Martín Martínez

Tutor/es

Dr. D. José Miguel Rodríguez - Pardo del Castillo

Dr. D. Jesús Ramón Simón del Potro

Madrid, julio 2018

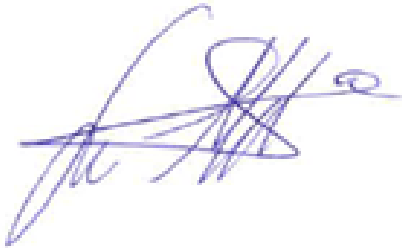
Esta tesis es propiedad del autor. No está permitida la reproducción total o parcial de este documento sin mencionar su fuente. El contenido de este documento es de exclusiva responsabilidad del autor, quien declara que no se ha incurrido en plagio y que la totalidad de referencias a otros autores han sido expresadas en el texto.

En caso de obtener una calificación igual o superior a 8.0 Notable, autorizo la publicación de este trabajo en el centro de Documentación de la Fundación Mapfre.

Sí, autorizo a su publicación.

No, desestimo su publicación.

Firmado:

A handwritten signature in blue ink, consisting of several overlapping loops and a long horizontal stroke extending to the right.

RESUMEN

Desde hace unos años el cliente ha tomado relevancia como parte de la gestión integral de las asociaciones. Esta importancia también se ha visto reflejada en las entidades aseguradoras, donde los clientes son cada vez más exigentes.

El mercado de no vida español es muy competitivo, con un número importante de compañías aseguradoras, muchas de ellas de ámbito internacional. Como consecuencia de ello se están elevando los costes de hacer nuevos clientes y especialmente, de retener a los actuales. Mantener una cartera de clientes rentables y satisfechos es uno de los elementos clave del negocio asegurador.

En esta tesis nos centraremos en analizar la experiencia del cliente en las prestaciones de hogar para ver qué factores son clave a la hora de recomendar y conseguir que el cliente quede satisfecho con el servicio.

Palabras clave: Modelo Lineal General, GLM, seguro de hogar, prestaciones de hogar, recomendación, cliente, experiencia del cliente, satisfacción del cliente.

ÍNDICE

Introducción	1
Capítulo 1. Marco teórico del seguro	3
1.1. Introducción	3
1.2. El seguro Multirriesgo de Hogar	3
Capítulo 2. Enfoque actuarial	7
2.1. Metodología y muestra para el estudio	7
2.2. Análisis descriptivo de la información	10
2.2.1. Análisis univariante	10
2.2.1.A. Variable dependiente	10
2.2.1.B. Variables independientes	12
2.1.2 Análisis bivariante	30
2.1.2. A. Relación variable dependiente vs independientes	30
2.1.2. B. Relación entre variables independientes	36
2.3. Modelo Lineal Generalizado (GLM)	38
2.3.1. Especificación del modelo teórico	38
2.3.2. Selección y evaluación del modelo	43
2.3.3. Estimación de parámetros	45
2.3.4. Análisis de los residuos	46
2.3.5. Validación del modelo	48
2.4. Modelo Lineal Generalizado en función de la tipología del proceso. Reparado vs Indemnizado	50
2.4. A Proceso reparatorio	50
2.4. B Proceso indemnizatorio	53
2.5. Modelo Lineal Generalizado de Alerta Temprana	55
Capítulo 3. Análisis de resultados y conclusiones	57
Capítulo 4. Referencias bibliográficas	59
Capítulo 5. ANEXOS	
Anexo I. Tablas de contingencia	
Anexo II. GLM	
Modelo Global	
Modelo proceso reparatorio	
Modelo proceso indemnizatorio	
Modelo de alerta temprana	
Anexo III. Código SPSS	

INTRODUCCIÓN

En el presente trabajo de fin de máster se va a analizar la percepción por parte del cliente de la calidad del servicio en las prestaciones del seguro de hogar y cómo ésta influye en la fidelización del cliente.

Día a día se manifiesta de manera más clara la importancia del cliente a la hora de potenciar el negocio en un entorno cada vez más cambiante. El sector asegurador forma parte de este ambiente que se caracteriza por una alta competitividad de precios y coberturas, donde es el trato diferencial en el servicio lo impacta de manera decisiva en la relación con el cliente y por consiguiente en la fidelización del mismo. Se analizará, adicionalmente si existe una influencia en esta fidelización de variables sociodemográficas, veremos si el impacto de éstas prevalece sobre la propia actuación de la compañía en la prestación del servicio.

El cliente es la herramienta necesaria para potenciar el negocio, la idea es poder controlar la lealtad del cliente como objetivo para crecer de forma rentable en el tiempo.

Con los resultados de este análisis la compañía puede ver que “teclas” tocar para llegar al cliente y poder satisfacer las necesidades del mismo.

Empresas como Starbucks, Inditex o Apple han tomado como práctica de gestión empresarial las relacionadas con la experiencia del cliente, haciendo hincapié en la importancia de conceptos relacionados con la satisfacción del mismo, la calidad del servicio ofrecido y la atención ofrecida en todo el proceso. Con la crisis mundial iniciada con la caída de Lehman Brothers en 2008 se ponen en duda cada vez más los modelos de negocio destacando la importancia de ofrecer una experiencia diferencial que dé como resultado una alta satisfacción con el servicio y un mayor índice de recomendación de la compañía.

En un entorno cada vez más digitalizado, donde está demostrado que los comentarios, opiniones y valoraciones de otros consumidores son clave a la hora de tomar la decisión de compra, el cliente fidelizado toma un papel cada vez más importante en el desarrollo de la imagen de la compañía, saber gestionar las experiencias vividas por éstos en el proceso de iteración con la compañía debe convertirse en una de las palancas relevante en el modelo de negocio.

El sector asegurador español se enfrenta a un nuevo marco competitivo desencadenado por la era digital, que está transformando el entorno tanto tecnológica como socialmente. El seguro de hogar no es ajeno a esta situación, y se coloca en un entorno en el cual aparecen nuevos jugadores que general y aprovechan nuevos modelos de negocio con ofertas y servicio novedosos. Es fundamental ofrecer al cliente un producto que se adapte a sus necesidades y conseguir, de ese modo, aumentar la fidelización y atraer nuevos clientes.

El trabajo se va a focalizar en el análisis de las prestaciones del seguro de hogar. Para poder anticiparse a las necesidades del cliente, las compañías deben contar con suficiente información y elaborar herramientas de análisis específicas. Con el objetivo de facilitar esta herramienta se va a desarrollar un modelo predictivo que permita analizar las variables que impactan de manera decisiva en la recomendación del cliente.

Se van a seguir los siguientes pasos:

- En primer lugar, se desarrollará a nivel teórico las características principales del seguro Multirriesgo de hogar.
- Desarrollo del enfoque práctico, analizando las variables disponibles para ver el perfil del cliente y las características del producto a analizar, llevando a cabo un análisis univalente, multivalente y finalmente desarrollando el modelo lineal generalizado (GLM). Para llevar a cabo estos análisis emplearemos el software estadístico SPSS.
- La última etapa será el análisis de resultados y conclusiones sobre las implicaciones en la fidelización del cliente.

Al final del documento se indican las referencias bibliográficas siguiendo el estilo UNE-ISO 690:2013 y los anexos con las tablas auxiliares y la programación llevada a cabo para elaborar el presente trabajo.

CAPITULO 1. MARCO TEÓRICO DEL SEGURO

1.1. Introducción

El concepto de seguro está relacionado con un acuerdo por el que se establece el traspaso de las consecuencias económicas desfavorables producidas como consecuencia del riesgo asegurado, a otra persona llamada asegurador, que puede ser distinta de la que puede sufrirlas (asegurado).

En función del tipo de riesgo que se quiere cubrir, podemos clasificar los seguros en:

- Seguros personales. Aseguran o cubren los riesgos que afectan a las personas y que puedan derivarse respecto a la vida o integridad física del asegurado. Seguros de este tipo son el de vida, enfermedad, accidentes, asistencia sanitaria.
- Seguros materiales. Éstos tienen como finalidad cubrir los riesgos que afectan a los animales o a las cosas. Por ejemplo, el incendio, el robo, rotura de cristales, pedrisco (agrícolas).
- Seguros patrimoniales. Este grupo de seguros cubren los riesgos que amenazan al patrimonio de las personas y producen una necesidad dineraria al asegurado. Por ejemplo, el seguro de responsabilidad civil, paralización de trabajos, seguro de crédito.

En el artículo 1 de la ley 50/1980, de 8 de octubre, de Contrato de seguro aparece la siguiente definición: *“El contrato de seguro es aquel por el que el asegurador se obliga, mediante el cobro de una prima y para el caso de que se produzca el evento cuyo riesgo es objeto de cobertura, a indemnizar a otra llamada asegurado o a la persona que este designe, beneficiario, dentro de los límites pactados, un perjuicio o daño que le pueda causar un suceso incierto o a satisfacer un capital, una renta u otras prestaciones convenidas”*.

1.2. El seguro Multirriesgo de Hogar

El seguro de hogar se enmarca dentro de los conocidos como seguros multirriesgo, éstos nacen de la idea de marketing del “cliente integral”, haciendo necesario crear productos con unas coberturas más amplias que las pólizas tradicionales que contenían los riesgos de manera aislada. Se trata de un contrato de seguro en el que se han reunido las garantías y coberturas de distintas pólizas de los ramos tradicionales (incendios, robo, responsabilidad civil, asistencia, etc.) cubriendo diferentes riesgos, mediante un único documento o póliza.

Los seguros multirriesgo son seguros de daños cuya finalidad principal es cubrir las pérdidas económicas que afecten al asegurado en sus bienes, con motivo de un siniestro; este tipo de seguros también pueden cubrir riesgos patrimoniales y personales.

Los seguros multirriesgo se caracterizan por:

- Al cubrir varios riesgos pueden ofrecer una cobertura más amplia.
- Están destinados a un segmento concreto del mercado.
- Tienen una administración simplificada.
- El coste es más reducido que el de otros seguros. La principal razón de que sea más económico que las pólizas de individuales es porque sólo se pagan impuestos una vez. El consorcio se paga en los riesgos individuales por cada póliza, y en los seguros multirriesgo se paga una sola vez.
- Ofrecen garantías especiales que no podrían ser ofrecidas de manera aislada por su elevado coste.

Dentro de los seguros multirriesgo podemos hacer la siguiente clasificación:

- Pólizas sencillas:
 - o Multirriesgo de Hogar
 - o Multirriesgo de Comunidades
 - o Multirriesgo de Comercios
- Pólizas industriales:
 - o Multirriesgo Industrial
 - o Otras pólizas concretas para un sector o actividad específica

Para hacernos una idea de la estructura del ramo de no vida y del peso que ocupa el seguro de multirriesgo dentro del sector, podemos analizar los datos del análisis técnico de los seguros multirriesgo publicado por ICEA de 2017. En este estudio se indica que las primas de multirriesgo sobre el ramo de no vida suponen un 20,2%, concretamente 6.882,56 millones de euros de los 34.002,71 del total (figura 1.1).

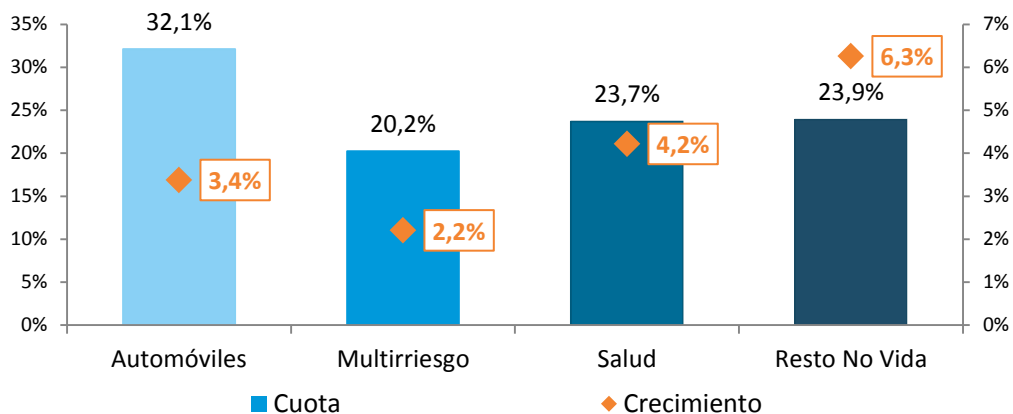


Figura 1.1. Composición del negocio de No Vida. Fuente: Asociación ICEA, análisis técnico de los seguros multirriesgo. Año 2017.

A su vez, en este estudio también se publica la composición dentro del seguro de multirriesgo. Podemos ver que el seguro de hogar ocupa un lugar destacado dentro de dicha composición, con el 60,8% de las primas y un crecimiento de 3% anual, siendo el producto que más crece dentro de los multirriesgo.

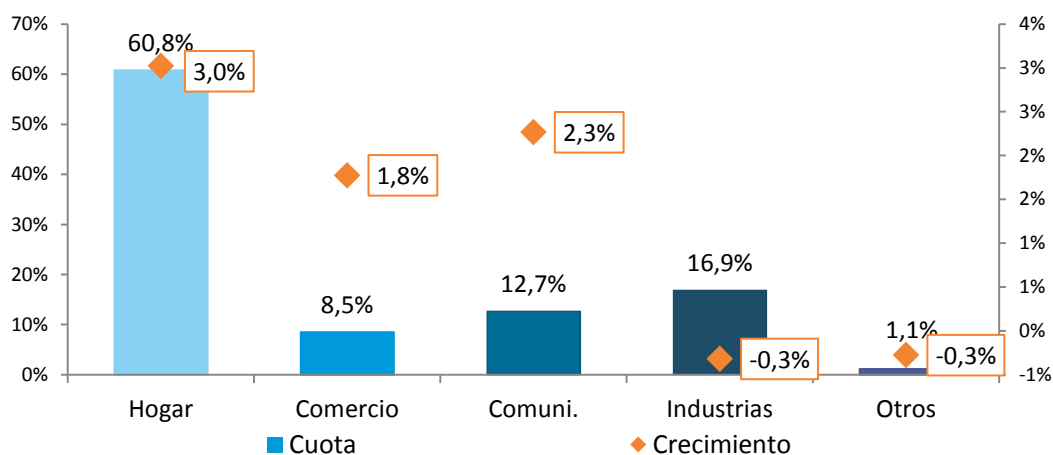


Figura 1.2. Composición del ramo Multirriesgo. Fuente: Asociación ICEA, análisis técnico de los seguros multirriesgo. Año 2017.

Históricamente, se puede destacar que el seguro multirriesgo de hogar es el primero dentro de este ramo que aparece en España, concretamente en los años sesenta. Su antecedente es la póliza de seguro combinado de incendio, explosión, robo y expoliación, exclusivamente para mobiliarios particulares. Normalmente este tipo de seguro parte habitualmente de la cobertura básica de incendio y sus riesgos complementarios, a la que se pueden ir añadiendo otras garantías según las necesidades de cobertura de cada caso concreto y dependiendo de cada entidad aseguradora.

Un primer análisis de las coberturas del seguro de hogar permite distinguir dos conceptos fundamentales:

- **Continente**, edificio o vivienda. Se refiere a aquellos bienes que no puedan separarse del inmueble sin pérdida de valor, garajes y trasteros, así como la parte proporcional que corresponda de las zonas comunes del edificio o la urbanización. La domótica también forma parte del inmueble. El valor que a declarar en la póliza es el de coste de reconstrucción (no se incluye el valor del suelo, ya que no es asegurable) distinto del de compra o venta. Dicho valor se calcula multiplicando los metros cuadrados construidos de la vivienda por el coste medio de reconstrucción de otra vivienda con similares características.
- **Contenido** o mobiliario. Con carácter general se considera contenido: muebles, electrodomésticos, ropa, otros enseres personales, provisiones, víveres y ajuar. Salvo pacto en contrario, las aseguradoras suelen excluir

de la cobertura vehículos, animales definidos como potencialmente peligrosos, plantas, joyas y objetos de valor, pieles, antigüedades.

En cuanto a estos últimos objetos, las condiciones generales de la póliza suelen requerir una valoración expresa e individualizada de cuadros y obras de arte, colecciones filatélicas y numismáticas, objetos de plata y metales preciosos que no sean joyas, alfombras y tapices, y abrigos de piel. La inclusión en el seguro de estos bienes de especial valor podrá dar lugar a un incremento de la prima.

De todos modos, lo común es establecer en la póliza que elementos se consideran continente o contenido a los efectos de la misma. Los límites de las coberturas calculan como porcentajes sobre el valor asegurado o cantidades máximas para cada riesgo asegurable, o ambas a la vez.

A nivel de mercado, según datos de ICEA publicados en el análisis técnico de los seguros multirriesgo, el 89% de las pólizas de negocio incluyen tanto continente como contenido.

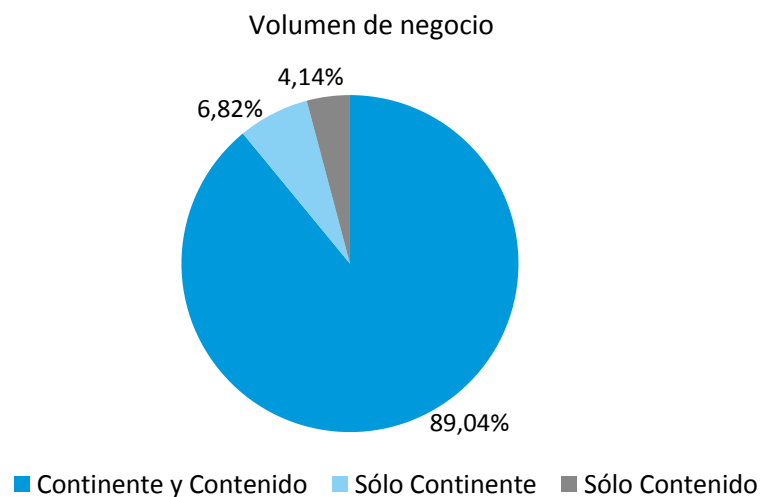


Figura 1.3. Distribución de las Pólizas en Función de la Cobertura. Fuente: Asociación ICEA, análisis técnico de los seguros multirriesgo. Año 2017.

CAPITULO 2. ENFOQUE ACTUARIAL

2.1. Metodología y muestra para el estudio

Para desarrollar el modelo planteado se ha tomado de partida una base de datos real de clientes encuestados sobre la prestación de su seguro de hogar. Se trata de un estudio online realizado a través de un panel a clientes con al menos una póliza en vigor de autos, hogar, salud, decesos o vida (riesgo y ahorro).

Por petición de la entidad propietaria de la base de datos y para conservar la confidencialidad de los mismos se ha realizado un muestreo aleatorio estratificado por tramos de edad, sexo y zona nielsen para conseguir una muestra más reducida pero representativa que refleje las principales características sociodemográficas. Finalmente tomamos de partida una base de datos con 1804 registros (índice de error de 2,6%) y las siguientes variables:

Tabla 2.1. Resumen de variables.

Variable	Tipo Variable	Categorías	
Edad	Cuantitativa continua		
Número de seguros contratados	Cuantitativa discreta		
Antigüedad	Cuantitativa discreta		
Zona Nielsen	Cualitativa nominal	- Norte - Sur - Centro - Galaica	- Levante - Catalana - Canarias
Sexo	Cualitativa binaria	- Hombre	- Mujer
Hábitat	Cualitativa nominal	- Urbano (Más de 50.000 habitantes)	- Rural (Hasta 50.000 habitantes)
Recomendación	Cualitativa ordinal	- Valores 1 - 10	
Intención de continuar	Cualitativa ordinal	- Valores 1 - 10	
Opinión general	Cualitativa ordinal	- Valores 1 - 10	
Vinculado a préstamo hipotecario	Cualitativa binaria	- Sí	- No
Cobertura del siniestro	Cualitativa nominal	- Totalmente - Parcialmente	- No lo ha cubierto / Ha rechazado el siniestro
Estado de siniestro	Cualitativa nominal	- Ya lo ha reparado la compañía - Ya lo he reparado con el dinero de la indemnización	- Estoy pendiente de reparación - Ya me han indemnizado - Pendiente de indemnización

Fuente: Elaboración propia

Tabla 2.2. Resumen de variables.

Variable	Tipo Variable	Categorías
Tipología del siniestro	Cualitativa nominal	<ul style="list-style-type: none"> - Robo - Daños por agua - Cristales - Robo - Subida de tensión / tormenta / daños eléctricos - Rotura saneamientos - Rotura electrodomésticos - Incendio - Viento - Cerrajería - Otros
Rapidez en la comunicación	Cualitativa ordinal. Valores 1 – 10	- Valores 1 - 10
Facilidad de contacto con la entidad	Cualitativa ordinal. Valores 1 - 10	- Valores 1 - 10
Información de los siguientes pasos a seguir en la tramitación	Cualitativa ordinal. Valores 1 - 10	- Valores 1 - 10
Interés por parte de la compañía en solucionar la prestación	Cualitativa ordinal. Valores 1 - 10	- Valores 1 - 10
Sencillez del lenguaje utilizado por el personal de la compañía	Cualitativa ordinal. Valores 1 - 10	- Valores 1 - 10
Canal de comunicación de la prestación	Cualitativa nominal	<ul style="list-style-type: none"> - En la oficina de la compañía - En la sucursal del banco / caja - Agente - Corredor - Teléfono de siniestros de la compañía - Internet - App de la compañía - Otros
Valoración del alcance del siniestro o de los daños producidos	Cualitativa nominal	<ul style="list-style-type: none"> - Sí, envió un perito - Sí, hubo una tele peritación o foto peritación - No hubo peritación - No recuerdo
Rapidez en la peritación	Cualitativa ordinal	- Valores 1 - 10
Rapidez en la reparación	Cualitativa ordinal	- Valores 1 - 10
Rapidez en la tramitación completa	Cualitativa ordinal	- Valores 1 - 10
Calidad de la reparación	Cualitativa ordinal	- Valores 1 - 10
Trato personalizado	Cualitativa ordinal	- Valores 1 - 10
Iniciativa por parte de la compañía	Cualitativa ordinal	- Valores 1 - 10
Adaptación del seguro a tus necesidades	Cualitativa ordinal	- Valores 1 - 10
Adecuada relación entre precio de la póliza y coberturas incluidas	Cualitativa ordinal	- Valores 1 - 10
Esfuerzo realizado en la gestión de la prestación	Cualitativa ordinal	- Valores 1 - 10
Expectativas de servicio	Cualitativa ordinal	<ul style="list-style-type: none"> - Mejor de los esperado - Igual, es lo esperado - Peor de los esperado

Fuente: Elaboración propia

Con esta información, el objetivo es analizar si la recomendación del cliente se ve afectada tanto por las propias características del mismo como por el proceso de prestación del servicio, analizando que variables y factores son claves en la prescripción del cliente y cuales tiene una menor incidencia.

Para ello, primero se va a realizar un análisis descriptivo de la información. Con ello se van a conocer las variables que pueden influir en el servicio y la fidelización del cliente, así como las principales características de las mismas perfilando así, al cliente que ha sufrido una prestación en el seguro de hogar.

Después se analizará la relación de cada variable con la fidelización para analizar las características de los clientes que recomiendan la entidad, para finalmente estudiar de que variables la recomendación es dependiente o independiente.

Todo este análisis previo descriptivo y de asociación de variables, adicionalmente va a servir de apoyo a la hora de emplear la técnica de modelización de Modelos Lineales Generalizados (GLM).

2.2. Análisis descriptivo de la información

En este apartado se desarrollará un trabajo de estadística descriptiva añadiendo las ideas desarrolladas por Jonh W. Tukey en el libro “Exploratory Data Analysis”, se trata de un enfoque de análisis de datos exploratorio que se emplea para analizar la información disponible en el trabajo e identificar la estructura y distribución de los datos.

2.2.1. Análisis univariante

Antes de aplicar cualquier técnica de análisis multivariante hay que llevar a cabo un análisis previo de los datos, el primer paso es analizar las variables que disponemos en la base de datos. Para ellos se va a proceder a realizar un estudio univariante de las variables, tanto independientes como dependientes. Primero analizaremos la variable dependiente para después proceder al análisis de los posibles factores o variables independientes que pueden afectar a la recomendación.

2.2.1.A. Variable dependiente.

¿Recomendarías a familiares y amigos la contratación de un seguro en esta compañía?

El índice de recomendación es la variable cualitativa ordinal que se quiere estimar a través del GLM para identificar que variables inciden en la recomendación o no de la compañía. Para ello se transformará en una variable cualitativa binaria (recomendación y no recomendación). Veamos la frecuencia de la variable original:

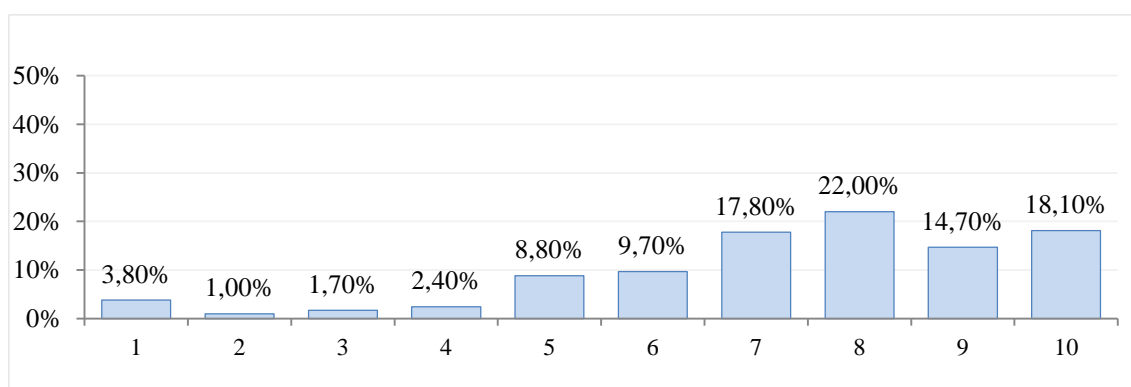


Figura 2.1. Distribución de frecuencias de la variable recomendación. Fuente: Elaboración propia.

Las valoraciones que se muestran en la figura 2.1. van desde 1 (no lo recomendaría en absoluto) hasta 10 (lo recomendaría totalmente).

La distribución de frecuencias y porcentajes de la variable recomendación, aparecen reflejados de forma más detallada en la siguiente tabla.

Tabla 2.3. ¿Recomendarías a familiares y amigos la contratación de un seguro en esta compañía?.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	No lo recomendaría en absoluto 1	66	3,8	3,8	3,8
	2	17	1,0	1,0	4,8
	3	29	1,7	1,7	6,5
	4	42	2,4	2,4	8,9
	5	153	8,8	8,8	17,7
	6	168	9,7	9,7	27,4
	7	308	17,8	17,8	45,2
	8	381	22,0	22,0	67,2
	9	254	14,7	14,7	81,9
	Lo recomendaría totalmente 10	313	18,1	18,1	100,0
Total		1731	100,0	100,0	

Fuente: Elaboración propia.

Se aprecia una mayor concentración en las valoraciones por encima de cinco. Para poder realizar el análisis predictivo se ha reconvertido en una variable binaria:

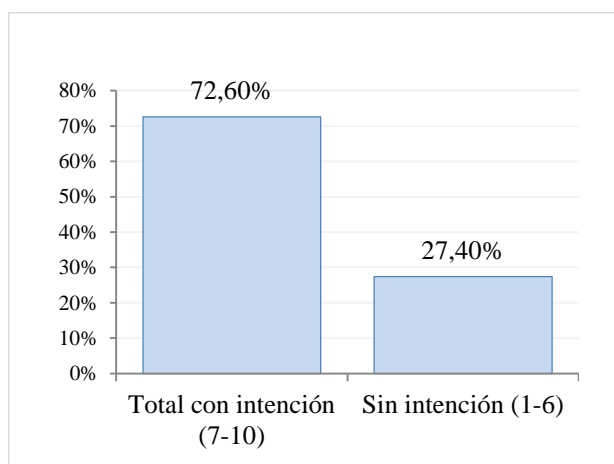


Figura 2.2. Distribución de frecuencias de la variable dicotómica Recomendación.

Fuente: Elaboración propia.

En la figura 2.2. se refleja claramente que hay un peso significativamente mayor de los clientes con intención de recomendar, se va analizar qué factores y en qué medida afectan a este segmento de clientes.

2.2.1.B. Variables independientes

I. Antigüedad

Ésta variable nos indica el número de años que lleva asegurado el cliente con éste seguro de hogar. En la figura 2.3 se aprecia la existencia de datos atípicos (indicados visualmente como “*”) y valores extremos (“o”).

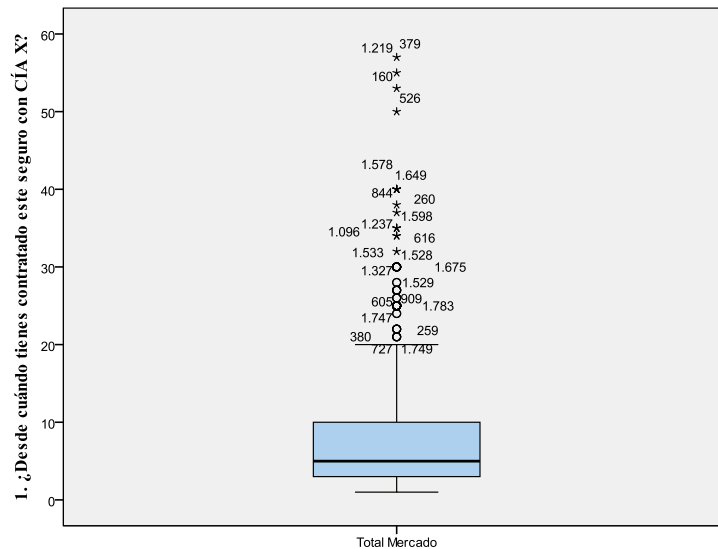


Figura 2.3. Gráfico de cajas con atípicos y extremos, variable antigüedad. Fuente: Elaboración propia.

Estos casos no se han tenido en cuenta a la hora de realizar los análisis posteriores, reduciendo la muestra a 1.731 encuestas.

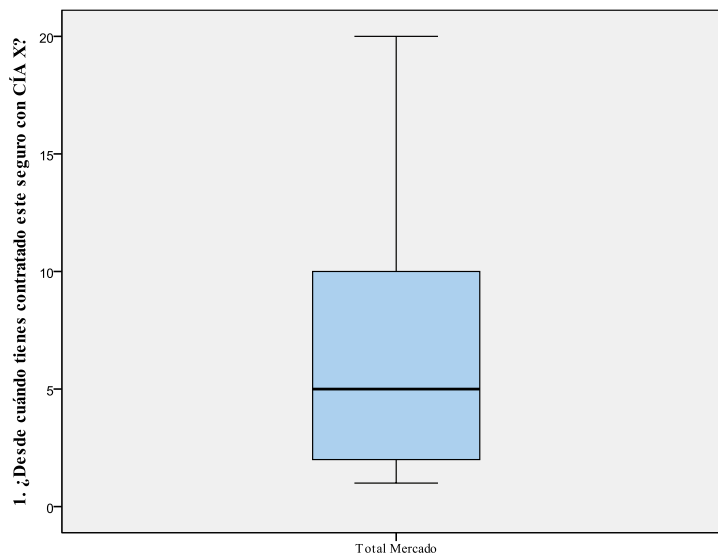


Figura 2.4. Gráfico de cajas, variable antigüedad corregida. Fuente: Elaboración propia.

La distribución de frecuencias resultante quedaría:

Tabla 2.4. ¿Desde cuando tienes contratado este seguro con cia x?.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	1	209	12,1	12,3	12,3
	2	223	12,9	13,1	25,4
	3	202	11,7	11,9	37,3
	4	141	8,1	8,3	45,6
	5	212	12,2	12,5	58,1
	6	90	5,2	5,3	63,4
	7	55	3,2	3,2	66,7
	8	72	4,2	4,2	70,9
	9	30	1,7	1,8	72,7
	10	181	10,5	10,7	83,3
	11	25	1,4	1,5	84,8
	12	46	2,7	2,7	87,5
	13	10	,6	,6	88,1
	14	15	,9	,9	89,0
	15	92	5,3	5,4	94,4
	16	6	,3	,4	94,8
	17	15	,9	,9	95,6
	18	9	,5	,5	96,2
	19	7	,4	,4	96,6
	20	58	3,4	3,4	100,0
	Total	1698	98,1	100,0	
Perdidos	Sistema	33	1,9		
	Total	1731	100,0		

Fuente: Elaboración propia.

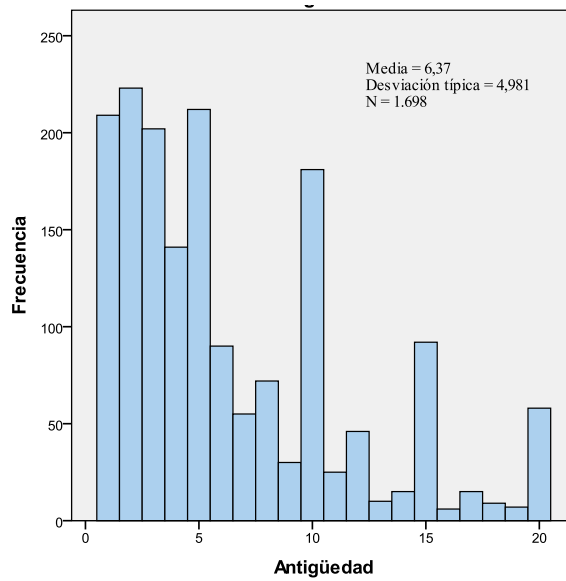


Figura 2.5. Histograma variable antigüedad.

Fuente: Elaboración propia.

La mayoría de los asegurados tienen una antigüedad por debajo de diez años, siendo la antigüedad media de este tipo de producto de 6,37 años. Se aprecia un ligero repunte en pólizas con antigüedades comprendidas entre quince y veinte años.

Visualmente ya se intuye que esta variable no parece presentar normalidad, pero el contraste de normalidad nos lo confirma estadísticamente. Para realizar este contraste nos fijaremos, por ejemplo, en el estadístico de Kolmogorov-Smirnov, se trata de un procedimiento de bondad de ajuste con el que medimos el grado de concordancia existente entre la distribución de la antigüedad y la distribución normal. Su objetivo es señalar si los datos provienen de una población que tiene la distribución teórica especificada, es decir, contrasta si las observaciones podrían razonablemente proceder de la distribución especificada.

Tabla 2.5. Pruebas de normalidad para la variable antigüedad

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Antigüedad	,190	1698	,000	,873	1698	,000

Fuente: Elaboración propia.

Tabla 2.6. Pruebas de Kolmogorov – Smirnov para una muestra, variable antigüedad.

		Antigüedad
N		1698
Parámetros normales ^{a,b}	Media	6,37
	Desviación típica	4,981
Diferencias más extremas	Absoluta	,190
	Positiva	,190
	Negativa	-,140
Z de Kolmogorov-Smirnov		7,819
Sig. asintót. (bilateral)		,000

Fuente: Elaboración propia.

El estadístico de prueba Kolmogorov-Smirnov presenta un nivel de significación igual a cero, rechazamos así la hipótesis nula de normalidad. El siguiente gráfico Q-Q normal ratifica este resultado, ya que los valores observados no se sitúan sobre la recta esperada bajo el supuesto de normalidad.

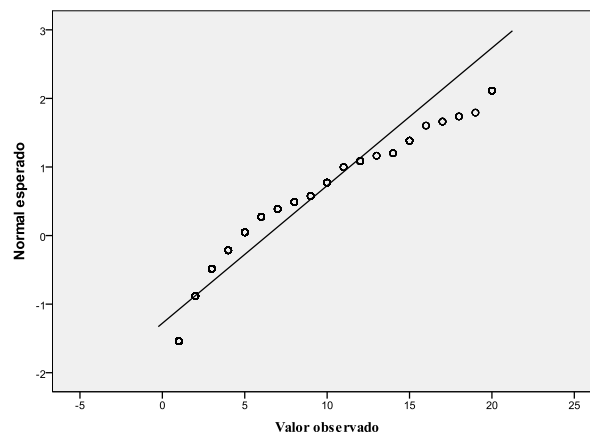


Figura 2.6. Gráfico Q-Q normal variable antigüedad.

Fuente: Elaboración propia.

II. Edad

Esta variable cuantitativa discreta recoge la edad del encuestado, que en este estudio puede ser asegurado o tomador del seguro. Si analizamos el gráfico de cajas no se aprecian datos atípicos o valores extremos.

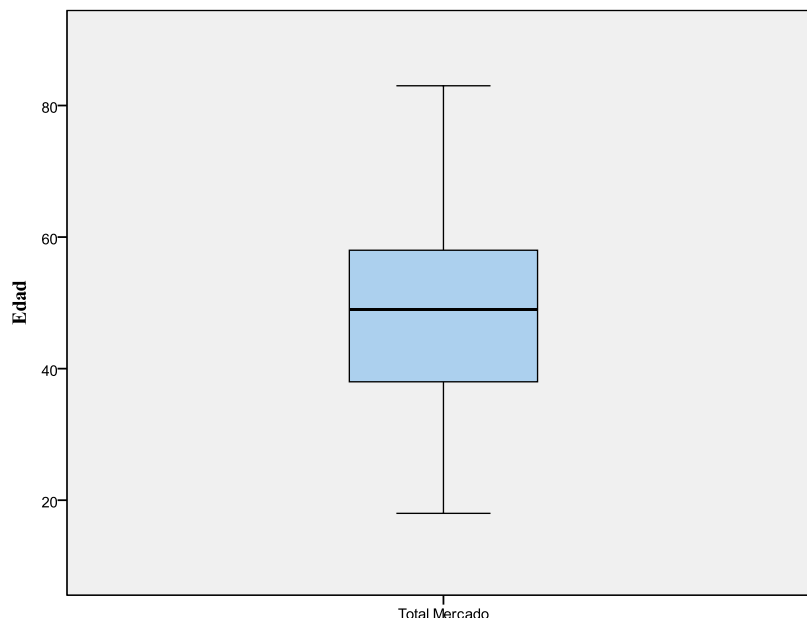


Figura 2.6. Gráfico de cajas de la variable edad.
Fuente: Elaboración propia.

Para analizar la distribución de la edad la he agrupado de la siguiente forma:

Tabla 2.7. Frecuencias y porcentaje de la variable agrupada edad.

	Frecuencia	Porcentaje válido	Porcentaje acumulado
Válidos 25 ó menos	167	9,6	9,6
De 26 a 35 años	209	12,1	21,7
De 36 a 45 años	356	20,6	42,3
De 46 a 55 años	419	24,2	66,5
De 56 a 65 años	408	23,6	90,1
Más de 65 años	172	9,9	100,0
Total	1731	100,0	

Fuente: Elaboración propia.

La mayoría de los clientes que están en posesión de un seguro de hogar y han tenido una prestación tienen edades comprendidas entre 36 y 65 años.

De nuevo, al tratarse de una variable cualitativa estudiaremos las características más relevantes como hicimos con la antigüedad del cliente (apartado I).

En primer lugar, con el análisis visual del histograma vemos que la muestra presenta una edad media de 47,5 años. La prueba de normalidad de Kolmorov-Smirnov nos indica de nuevo que esta variable no sigue una distribución normal.

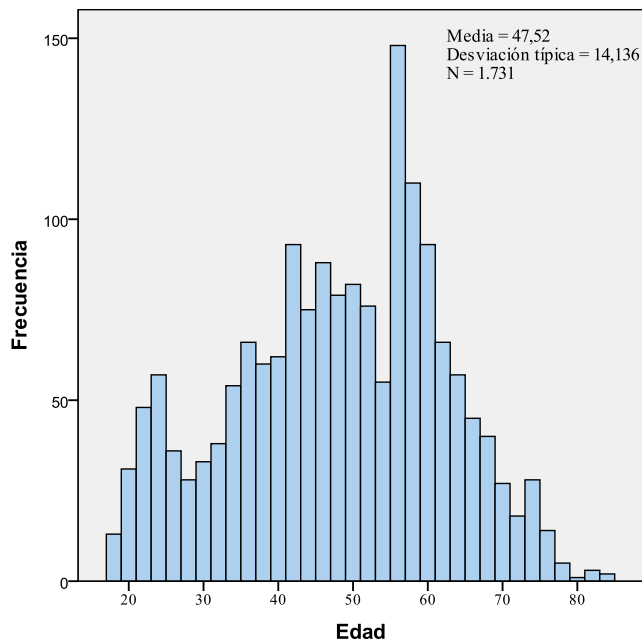


Tabla 2.8. Prueba de Kolmogorov – Smirnov para una muestra, variable edad.

		Edad
N		1731
Parámetros normales,a,b	Media	47,52
	Desviación típica	14,136
Diferencias más extremas	Absoluta	,081
	Positiva	,043
	Negativa	-,081
Z de Kolmogorov-Smirnov		3,382
Sig. asintót. (bilateral)		,000

Fuente: Elaboración propia.

Figura 2.8. Histograma de la variable edad.
Fuente: Elaboración propia.

Tabla 2.9. Pruebas de normalidad para la variable edad.

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Edad	,081	1731	,000	,981	1731	,000

Fuente: Elaboración propia.

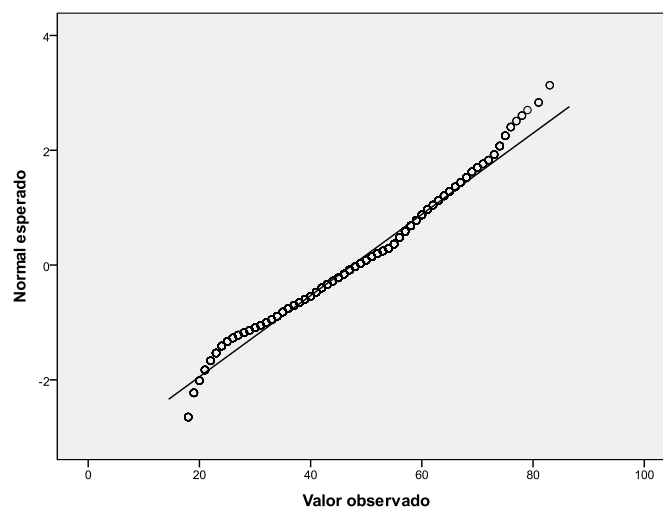


Figura 2.9. Gráfico Q-Q normal variable edad.
Fuente: Elaboración propia.

III. Número de seguros contratados

Esta variable indica el número de seguros que tiene contratado este cliente. La variable está limitada hasta tres seguros de hogar, pudiendo indicar otras tipologías que no son objeto de éste análisis. La distribución de frecuencias es la siguiente:

Tabla 2.10. Frecuencia y porcentaje de la variable número de seguros contratados.

		Frecuencia	Porcentaje válido	Porcentaje acumulado
Válidos	1	1264	73,0	73,0
	2	349	20,2	93,2
	3	118	6,8	100,0
	Total	1731	100,0	

Fuente: Elaboración propia.

Tres de cada cuatro encuestados tienen únicamente un seguro de hogar, un 20% tiene dos y apenas hay clientes que estén en posesión de tres seguros de hogar. El número medio de seguros es 1,3.

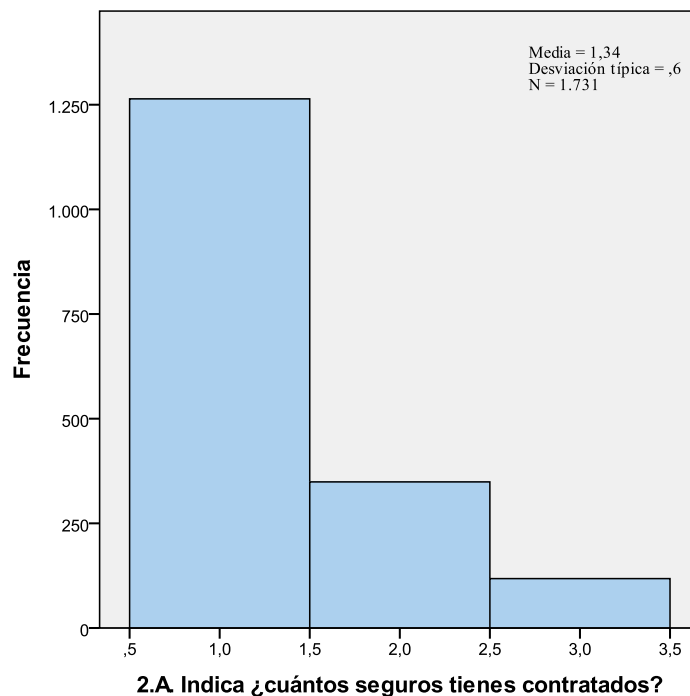


Figura 2.10. Histograma de la variable número de seguros.
Fuente: Elaboración propia.

IV. Sexo

Se trata de una variable dicotómica que nos indica que el 52% de los encuestados son mujeres.

Tabla 2.11. Frecuencia y porcentaje de la variable sexo.

		Frecuencia	Porcentaje válido	Porcentaje acumulado
Válidos	Hombre	833	48,1	48,1
	Mujer	898	51,9	100,0
	Total	1731	100,0	

Fuente: Elaboración propia.

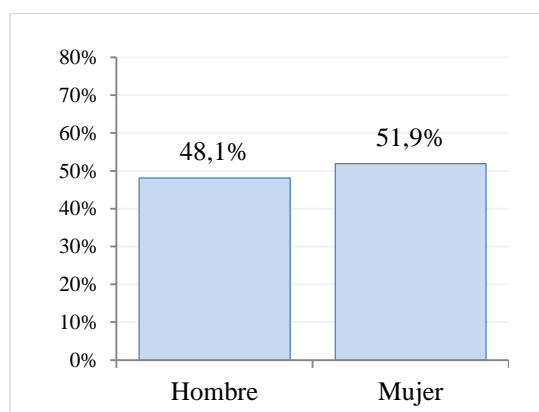


Figura 2.11. Distribución de frecuencias por sexo.
Fuente: Elaboración propia.

V. Zona Nielsen

Para tratar el ámbito geográfico, se ha decidido agrupar los ámbitos territoriales en distintas zonas en función de la geolocalización.

Tabla 2.12. Frecuencia y porcentaje de la variable Zona Nielsen.

		Frecuencia	Porcentaje válido	Porcentaje acumulado
Válidos	Norte	163	9,4	9,4
	Sur	312	18,0	27,4
	Centro	499	28,8	56,3
	Galaica	152	8,8	65,0
	Levante	239	13,8	78,9
	Catalana	323	18,7	97,5
	Canarias	43	2,5	100,0
	Total	1731	100,0	

Fuente: Elaboración propia.

A continuación, se muestra un mapa que indica que provincias corresponden a cada zona:

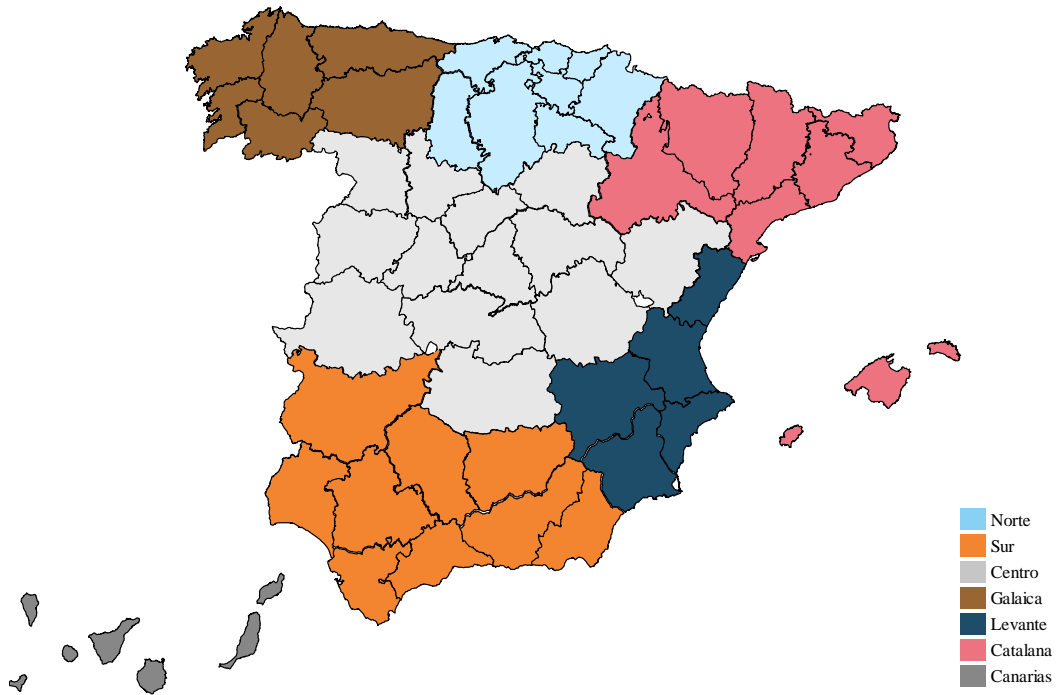


Figura 2.12. Mapa de Zonas nielsen. Fuente: Elaboración propia.

La distribución resultante de esta agrupación geográfica sería:

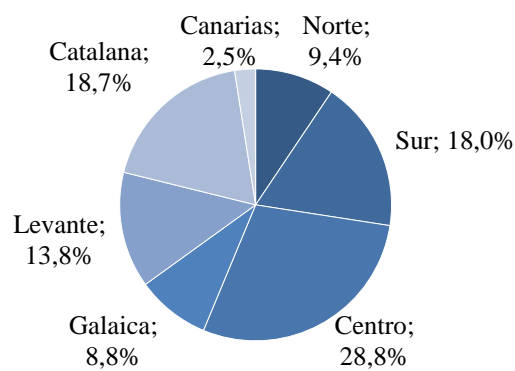


Figura 2.13. Mapa de Zonas nielsen. Fuente: Elaboración propia.

Vemos que las zonas con mayor peso en la muestra son centro, catalana y sur.

VI. Hábitat

Adicionalmente se ha realizado una agregación geográfica distinguiendo entre hábitat rural (menos de 50.000 habitantes) y urbano (más de 50.000 habitantes).

Tabla 2.13. Frecuencia y porcentaje de la variable hábitat.

		Frecuencia	Porcentaje válido	Porcentaje acumulado
Válidos	Urbano (Más de 50.000 habitantes)	1166	67,4	67,4
	Rural (Hasta 50.000 habitantes)	565	32,6	100,0
	Total	1731	100,0	

Fuente: Elaboración propia.

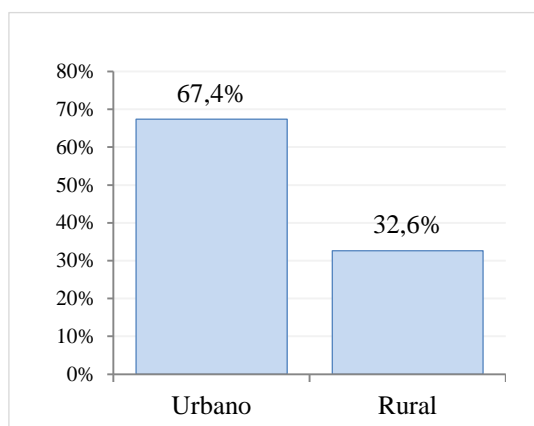


Figura 2.14. Distribución de frecuencias por hábitat.
Fuente: Elaboración propia.

Dos tercios de la muestra proceden de un ámbito urbano, será interesante ver si la recomendación será mayor en poblaciones de más de 50.000 habitantes dado el peso que tiene sobre la muestra.

VII. Intención de continuar

Esta variable recoge la intención de continuar del asegurado en la compañía, es decir, la prescripción del cliente. Se trata de una variable cualitativa ordinal que recoge valoraciones de 1 a 10, siendo 1 ninguna intención de continuar y 10 completa intención de continuar asegurado. La distribución de frecuencias sería la siguiente:

Tabla 2.14. Frecuencia y porcentaje de intención de recomendar.

		Frecuencia	Porcentaje válido	Porcentaje acumulado
Válidos	1	86	5,0	5,0
	2	24	1,4	6,4
	3	41	2,4	8,7
	4	49	2,8	11,6
	5	145	8,4	19,9
	6	144	8,3	28,2
	7	260	15,0	43,3
	8	329	19,0	62,3
	9	272	15,7	78,0
	10	381	22,0	100,0
	Total	1731	100,0	

Fuente: Elaboración propia.

Al igual que ocurría con nuestra variable dependiente (recomendación), a éste tipo de variables vamos a transformarlas en una dicotómica donde el valor 1 corresponda a los clientes satisfechos / con intención (valoraciones de 7 a 10) y el valor 2 a clientes no satisfechos / sin intención (valoraciones de 1-6).

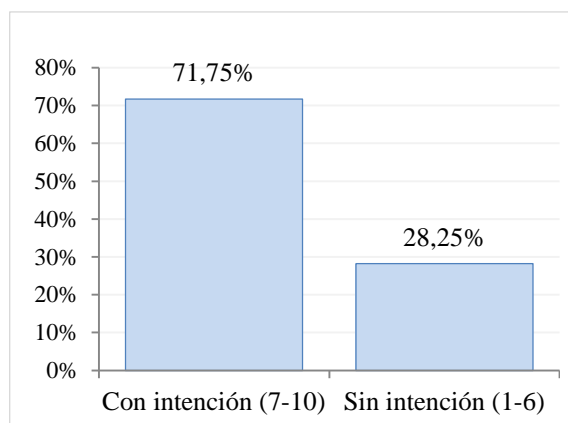


Figura 2.15. Distribución de frecuencias según la intención de continuar en la compañía.

Fuente: Elaboración propia.

En este caso el 72% de los clientes que han hecho uso del seguro de hogar tienen intención de continuar asegurados en la compañía

VIII. Opinión general

Otra variable disponible para el análisis es la opinión general del cliente sobre el servicio recibido. La distribución de frecuencias y el gráfico que nos muestra la transformación en variable dicotómica se muestran a continuación:

Tabla 2.15. Frecuencia y porcentaje de la opinión general sobre la compañía.

		Frecuencia	Porcentaje válido	Porcentaje acumulado
Válidos	1	65	3,8	3,8
	2	26	1,5	5,3
	3	43	2,5	7,7
	4	58	3,4	11,1
	5	132	7,6	18,7
	6	153	8,8	27,6
	7	277	16,0	43,6
	8	366	21,1	64,7
	9	307	17,7	82,4
	10	304	17,6	100,0
Total		1731	100,0	

Fuente: Elaboración propia.

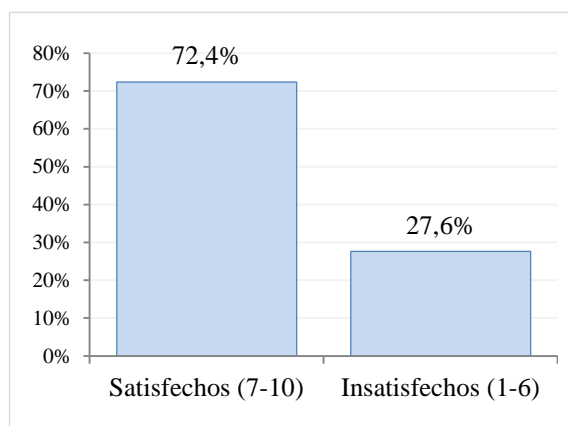


Figura 2.16. Distribución de frecuencias según opinión general sobre la compañía.

Fuente: Elaboración propia.

Vemos que el 72% de los clientes están satisfechos con el servicio recibido, porcentaje más bajo que el de la prescripción (apartado VII).

IX. Vinculación a préstamo hipotecario

Un tema que puede tener un efecto importante en el entorno del seguro de hogar, es si este viene impuesto por un préstamo hipotecario, de ahí que analicemos la relación del seguro con dicho préstamo.

Tabla 2.16. Frecuencia y porcentaje de vinculado a préstamos hipotecario.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Sí	471	27,2	27,6	27,6
	No	1238	71,5	72,4	100,0
	Total	1709	98,7	100,0	
Perdidos	Sistema	22	1,3		
Total		1731	100,0		

Fuente: Elaboración propia.

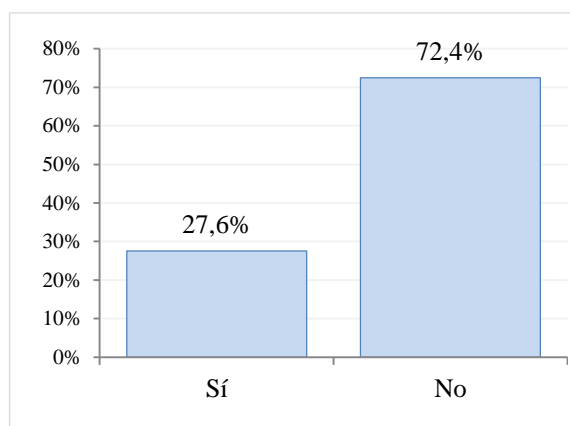


Figura 2.17. Distribución de frecuencias según si el cliente está vinculado a préstamos hipotecarios.

Fuente: Elaboración propia.

En esta muestra se aprecia que la gran mayoría de los clientes no tiene el seguro vinculado a un préstamo hipotecario si bien, hay un 27,6% de la muestra que sí está relacionado con la hipoteca de la vivienda.

X. Cobertura del siniestro

Variable que recoge la información referente a si el siniestro ha sido cubierto totalmente, parcialmente o ha sido rechazado por la compañía. El resultado es el siguiente:

Tabla 2.17. Frecuencia y porcentaje de la variable cobertura del siniestro.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Totalmente	1156	66,8	68,4	68,4
	Parcialmente	445	25,7	26,3	94,7
	No lo ha cubierto / Ha rechazado el siniestro completo	89	5,1	5,3	100,0
	Total	1690	97,6	100,0	
Perdidos	Sistema	41	2,4		
Total		1731	100,0		

Fuente: Elaboración propia.

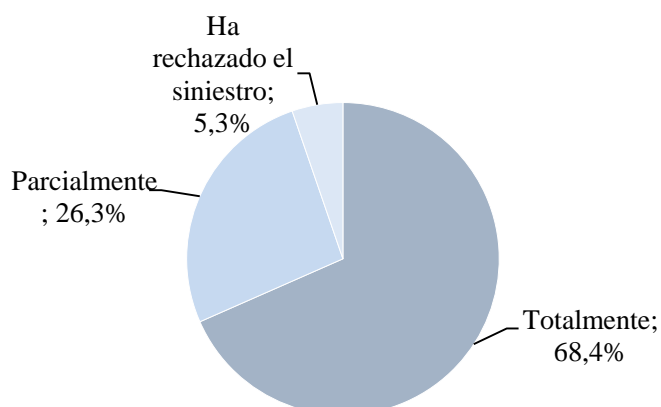


Figura 2.18. Distribución según la cobertura del siniestro.
Fuente: Elaboración propia.

Apenas existen siniestros rechazados por la compañía (5,3%), la mayoría de los siniestros han sido cubiertos, concretamente el 68% han resultado con una cobertura completa.

XI. Estado de siniestro

Las prestaciones pueden seguir una vía reparatoria o indemnizatoria, la tabla de frecuencias muestra las posibles situaciones de la prestación:

Tabla 2.18. Frecuencia y porcentaje de la variable estado del siniestro.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Ya lo ha reparado la compañía de seguros	1041	57,7	62,3	62,3
	Ya lo he reparado con el dinero de la indemnización	226	12,5	13,5	75,9
	Estoy pendiente de reparación	157	8,7	9,4	85,3
	Ya me han indemnizado	182	10,1	10,9	96,2
	Pendiente de indemnización	64	3,5	3,8	100
	Total		1670	92,6	100,0
Perdidos	Sistema	134	7,4		
Total		1804	100,0		

Fuente: Elaboración propia

Para incluir la variable en la regresión que se realizara más adelante, se ha procedido a realizar una agrupación de las prestaciones, de tal forma que el 71,7% de las mismas han seguido un proceso reparatorio.

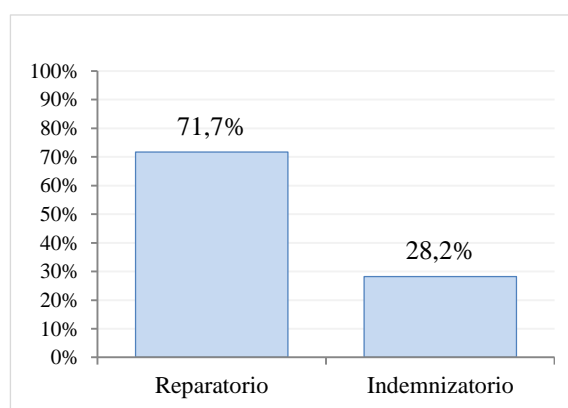


Figura 2.19. Distribución según la cobertura del siniestro.

Fuente: Elaboración propia.

XII. Tipología del siniestro

Esta variable recoge la tipología de la prestación del siniestro de hogar. La tabla de frecuencias de la variable original es la siguiente:

Tabla 2.19. Frecuencia y porcentaje de la variable tipología del siniestro.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Daños por agua	593	34,3	37,0	37,0
	Cristales	199	11,5	12,4	49,5
	Robo	83	4,8	5,2	54,7
	Subida de tensión / tormenta / daños eléctricos / cortes de electricidad	120	6,9	7,5	62,1
	Rotura saneamientos	177	10,2	11,1	73,2
	Rotura electrodomésticos	191	11,0	11,9	85,1
	Incendio	29	1,7	1,8	86,9
	Viento	44	2,5	2,7	89,7
	Cerrajería	88	5,1	5,5	95,2
	Otros	77	4,4	4,8	100,0
	Total	1601	92,5	100,0	
Perdidos	Sistema	130	7,5		
Total		1731	100,0		

Fuente: Elaboración propia

En función de la tipología de la prestación, el proceso de servicio varía notablemente y este hecho puede influir en la precepción del cliente del mismo e incidir en la recomendación de la compañía. Para realizar el análisis se han realizado las siguientes agrupaciones:

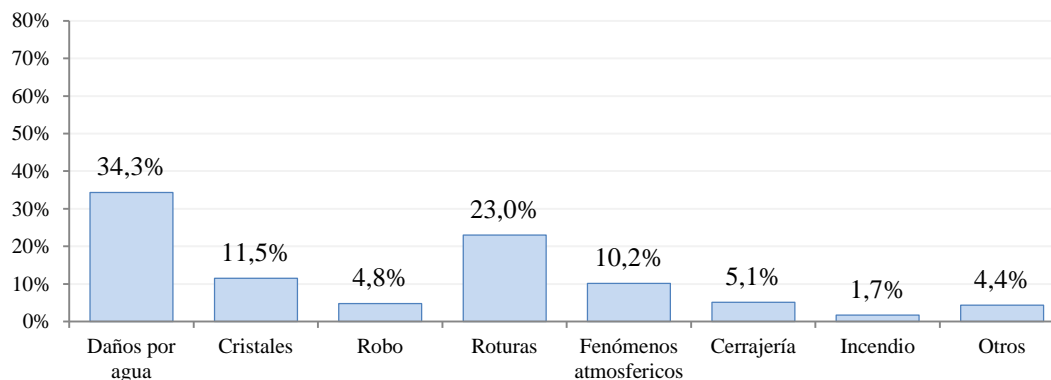


Figura 2.18. Distribución según la tipología del siniestro.

Fuente: Elaboración propia.

La categoría “roturas” incluye tanto roturas de electrodomésticos como de saneamientos. Fenómenos atmosféricos recoge las prestaciones de subida de tensión/ tormenta/ daños eléctricos/ cortes de electricidad y viento. El otros recoge otras tipologías no identificadas o combinación de las anteriores, al ser una variable de respuesta única no se pueden incluir en las categorías correspondientes.

Las prestaciones más frecuentes vienen motivadas por siniestros de daños por agua o roturas.

XIII. Esfuerzo realizado en la gestión de la prestación

Un aspecto que ha tomado relevancia en los últimos tiempos a la hora de analizar la relación del cliente con la compañía es el esfuerzo que tiene que realizar el cliente para gestionar su prestación. En numerosos textos aparece reflejado como *Customer Effort Score (CES)*.

El Customer Effort Score mide como incrementa una experiencia positiva la satisfacción mediante la eliminación de las barreras y facilidad de resolución de incidencias. La puntuación se establece en una escala del 1 al 10, como en el resto de aspectos si bien en este caso 10 representa poco esfuerzo y 1 mucho esfuerzo.

Tabla 2.20. Frecuencia y porcentaje del esfuerzo realizado en la gestión de la prestación.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Mucho esfuerzo 1	75	4,3	4,7	4,7
	2	61	3,5	3,8	8,5
	3	58	3,4	3,6	12,1
	4	77	4,4	4,8	16,9
	5	170	9,8	10,6	27,6
	6	133	7,7	8,3	35,9
	7	277	16,0	17,3	53,2
	8	320	18,5	20,0	73,2
	9	227	13,1	14,2	87,4
	Muy poco esfuerzo 10	202	11,7	12,6	100,0
Total	1600	92,4	100,0		
Perdidos	Sistema	131	7,6		
Total		1731	100,0		

Fuente: Elaboración propia

Vemos que el porcentaje de clientes que han tenido que hacer un alto esfuerzo para gestionar su prestación es el 35,9%. Más adelante veremos cómo influye este aspecto en la recomendación del cliente.

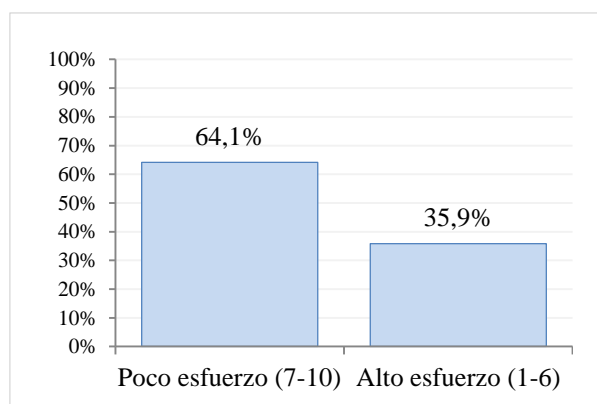


Figura 2.19. Distribución del esfuerzo realizado en la gestión de la prestación.

Fuente: Elaboración propia.

XIV. Expectativas de servicio

Es necesario saber si el servicio ha cubierto las expectativas del cliente y en qué medida lo ha hecho. Este cumplimiento se recoge en esta variable cualitativa, clasificándolas como mejor, igual o peor de lo esperado.

Hay que tener presente que esta variable viene influenciada tanto por la generación de expectativas previa a la realización del servicio (durante la contratación), como las que se generan como consecuencia de la prestación.

Tabla 2.21. Frecuencia y porcentaje del cumplimiento de expectativas.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Mejor de lo que esperaba	493	28,5	30,8	30,8
	Igual (es lo que esperaba)	921	53,2	57,5	88,3
	Peor de lo que esperaba	187	10,8	11,7	100,0
	Total	1601	92,5	100,0	
Perdidos	Sistema	130	7,5		
Total		1731	100,0		

Fuente: Elaboración propia

En la mayoría de los casos no se sorprende al cliente (57,5% igual a lo esperado), un 31% si considera que el proceso ha sido mejor de lo esperado y un 12% no han visto cumplidas las expectativas que tenían sobre el seguro de hogar.

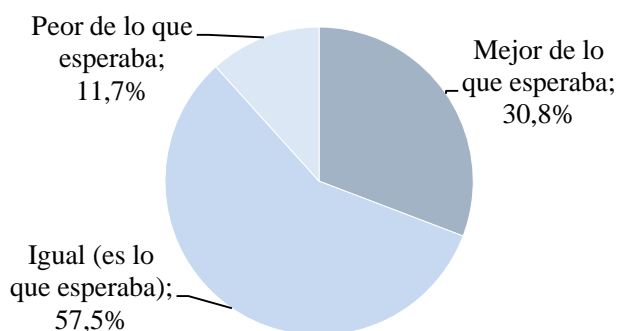


Figura 2.20. Distribución del cumplimiento de expectativas.

Fuente: Elaboración propia.

XV. Canal de comunicación de la prestación

La comunicación se puede realizar a través de los canales recogidos en la tabla 2.35. Lo más frecuente es comunicarle el siniestro a la compañía a través del teléfono.

Tabla 2.22. Frecuencia y porcentaje del canal de comunicación.

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	En la oficina de la compañía	130	7,5	8,1	8,1
	En la sucursal del banco / caja de ahorros	27	1,6	1,7	9,8
	Me puse en contacto con el agente	254	14,7	15,9	25,7
	Me puse en contacto con el corredor	173	10,0	10,8	36,5
	Llamé al teléfono de siniestros de la compañía, del tipo 900 o similar	915	52,9	57,2	93,7
	A través de internet	78	4,5	4,9	98,6
	A través de la App de la compañía	18	1,0	1,1	99,7
	Otros	5	,3	,3	100,0
	Total	1600	92,4	100,0	
Perdidos	Sistema	131	7,6		
Total		1731	100,0		

Fuente: Elaboración propia

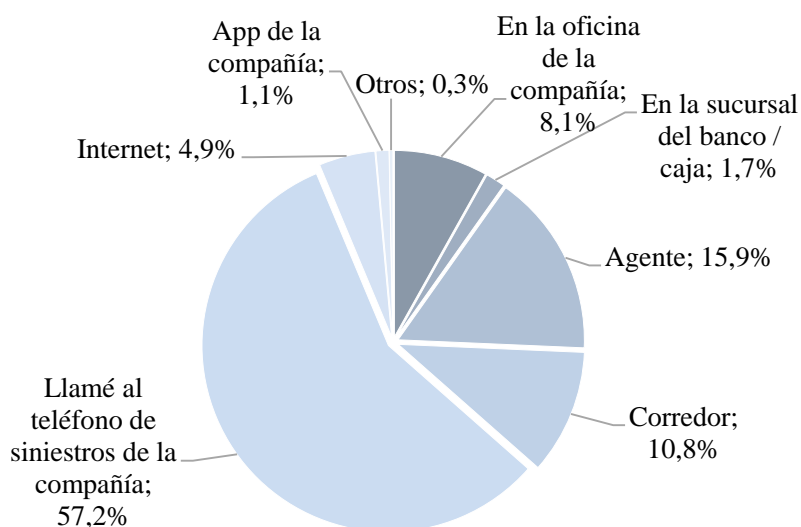


Figura 2.21. Distribución del canal de comunicación.

Fuente: Elaboración propia.

Veremos si el empleo de un canal u otro influye en la recomendación final del cliente.

Nota: para el resto de aspectos no se va a mostrar este análisis debido a la confidencialidad de datos.

2.1.2 Análisis bivalente

Después de ver las características de cada variable de forma individual hay que analizar las relaciones entre las mismas y de cada una con la variable dependiente. De ésta forma se puede identificar que variables tienen una alta relación entre sí para poder detectar posibles problemas de multicolinealidad.

2.1.2. A. Relación variable dependiente vs independientes

Para analizar la relación que puede existir entre nuestra variable dependiente y el resto de información disponible vamos a dividir el trabajo en dos fases:

A.1. Contrastación estadística de la relación de dependencia

Para comprobar si existe dependencia o asociación entre variables podríamos limitarnos a observar la correlación, pero no correlación no implica dependencia por lo que es necesario llevar a cabo una comprobación estadística, en este caso se va a realizar el contraste chi-cuadrado para independencia de variables, si el resultado es próximo a cero se rechaza hipótesis de independencia.

Ya hemos visto que la mayoría de las variables son de tipo cualitativas o categóricas, incluida la propia variable dependiente. Para estudiar la relación entre dos variables categóricas primero debemos preguntarnos si ésta existe (si son dependientes) y en caso de que la respuesta sea afirmativa, el siguiente paso es determinar qué tipo de relación presentan.

Como ya se ha indicado, para estudiar la existencia de relación entre variables vamos a utilizar el contraste de independencia chi-cuadrado. Este contraste es uno de los más utilizados en aplicaciones estadísticas, fue introducido por Karl Pearson en 1900. Con esta estadístico vamos a poder decidir si dos criterios de clasificación sobre el mismo conjunto muestral son independientes o no, es decir, si existe o no relación entre dos variables categóricas que se han medido sobre el mismo conjunto de individuos.

Las hipótesis del contraste serían:

H_0 : las variables son independientes

H_1 : las variables no son independientes (son dependientes)

El estadístico del contraste se obtiene de la siguiente ecuación:

$$\chi^2 \rightarrow Q = \sum_{i=1}^k \sum_{j=1}^r \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

Donde e_{ij} son las frecuencias esperadas si los dos criterios de clasificación fueran independientes. Se trata de una medida de concordancia (distancia) entre las frecuencias observadas (n_{ij}) y las frecuencias esperadas.

El estadístico Q tomará valores mayores cuanto menos se parezcan las frecuencias observadas y esperadas. Lo que se pretende es intentar medir cuánto se parecen las frecuencias observadas (reales) a las esperadas bajo el supuesto de que ambas variables sean independientes. Si toma valores grandes será poco creíble que ambas frecuencias se parecen, pero si el valor del estadístico es pequeño la hipótesis será más creíble. Para entender este contraste hay que establecer que es grande y que es pequeño.

Bajo la hipótesis de independencia el estadístico Q tiene una ley $\chi^2_{(k-1)(r-1)}$. Se utiliza como umbral de decisión el percentil $(1-\alpha)100$ de esta ley de probabilidad, siendo α el nivel de significación.

Una vez que se establece que existe dependencia entre dos variables, para determinar la intensidad de la misma se utilizan las siguientes medidas de asociación:

➤ Variables nominales:

- Basadas en el estadístico chi-cuadrado. Con estas medidas se intenta eliminar la influencia del tamaño muestral, para ello se efectúan correcciones para normalizar y conseguir que tomen un valor entre 0 y 1, éstas medidas son:

- Coeficiente de contingencia de Pearson

$$C = \sqrt{\frac{Q/n}{1 + Q/n}} \rightarrow 0 \leq C \leq 1$$

$C = 0 \rightarrow$ no existe asociación

$C \approx 1 \rightarrow$ asociación completa

- V de Cramer

$$V = \sqrt{\frac{Q/n}{\min\{k-1, r-1\}}} \rightarrow 0 \leq V \leq 1$$

$V = 0 \rightarrow$ no existe asociación

$V \approx 1 \rightarrow$ asociación completa

La V de Cramer se puede expresar en porcentaje de asociación.

Ambas medidas no suelen obtener resultados muy elevadas, aunque según el estadístico de chi-cuadrado indique que si existe relación.

- Basadas en la reducción proporcional del error. Una de las dos variables que aparecen en la tabla de contingencia es función de la otra, miden como de bien se puede predecir el valor de la variable dependiente (recomendación) conocido el valor de la independiente.
 - Coeficiente lambda de Goddman-Kruskal.

$$\lambda = \frac{P(\text{error, sin información de } X) - P(\text{error, con información de } X)}{P(\text{error, sin información de } X)}$$

$$0 \leq \lambda \leq 1$$

Si X e Y son independientes $\lambda = 0$, pero $\lambda = 0$ no implica independencia.

Si $\lambda = 0 \Rightarrow$ la información sobre X no ayuda en la predicción de Y.

Si $\lambda = 1 \Rightarrow$ la información sobre X sí ayuda en la predicción de Y.

➤ Variables ordinales.

Las medidas para analizar la relación de este tipo de variables se basan en la concordancia (C = pares de valores concordantes) o la discordancia (D = valores discordantes) de los pares de valores.

- γ de Goodman-Kruskal

$$\gamma = \frac{C - D}{C + D} \rightarrow -1 \leq \gamma \leq 1$$

Si $\gamma = 0 \Rightarrow$ No asociación.

Si $\gamma = 1 \Rightarrow$ Asociación completa positiva.

Si $\gamma = -1 \Rightarrow$ Asociación completa negativa.

- \mathcal{D} de Sommer

$$\mathcal{D} = \frac{C - D}{\frac{n(n-1)}{2} - T_x} \rightarrow -1 \leq \mathcal{D} \leq 1$$

$$T_x = \sum_{i=1}^k \frac{n_{i\cdot}(n_{i\cdot} - 1)}{2}$$

- T_B y T_C de Kendall

$$T_B = \frac{C - D}{\sqrt{\frac{n(n-1)}{2} - T_x(\frac{n(n-1)}{2} - T_y)}} \rightarrow -1 \leq T_B \leq 1$$

$$T_y = \sum_{j=1}^k \frac{n_{\bullet j}(n_{\bullet j} - 1)}{2}$$

$$T_C = \frac{\min\{k, r\} (C - D)}{\min\{k - 1, r - 1\} n^2} \rightarrow -1 \leq T_C \leq 1$$

Éstas medidas se interpretan de manera análoga la γ de Goodman-Kruskal.

En la siguiente hoja se muestra un resumen de las variables disponibles par el análisis con el resultado del contraste de independencia respecto a la recomendación y el resultado de las medidas de asociación.

Tabla 2.23. Resumen de variables, contraste chi cuadrado y medidas de asociación con la variable dependiente (recomendación).

	Contraste chi cuadrado					Asociación variables nominales	Asociación variables ordinales		
	Q	gl	P-valor	Nivel de significación: 0,05		V cramer	Tb Kendall	Tc de Kendall	Gamma
Edad	13,407	5	0,020	Rechazo Ho	Dependientes	8,8%			
Antigüedad	7,770	3	0,051	No se puede rechazar Ho	Independientes	6,8%			
Número de seguros contratados	3,405	2	0,182	No se puede rechazar Ho	Independientes	4,4%			
Sexo	0,640	1	0,424	No se puede rechazar Ho	Independientes	1,9%			
Zona Nielsen	4,740	8	0,785	No se puede rechazar Ho	Independientes	52,0%			
Hábitat	0,122	1	0,727	No se puede rechazar Ho	Independientes	0,8%			
Intención de continuar	1.042,097	1	0,000	Rechazo Ho	Dependientes		0,776	0,623	0,975
Opinión general	1.068,255	1	0,000	Rechazo Ho	Dependientes		0,786	0,786	0,977
Vinculado a préstamo hipotecario	29,872	1	0,000	Rechazo Ho	Dependientes	13,2%			
Cobertura del siniestro	307,365	2	0,000	Rechazo Ho	Dependientes	42,6%			
Tipología del siniestro	12,182	10	0,273	No se puede rechazar Ho	Independientes	8,8%			
Rapidez en la comunicación	505,675	1	0,000	Rechazo Ho	Dependientes		0,540	0,423	0,868
Facilidad de contacto con la entidad	437,906	1	0,000	Rechazo Ho	Dependientes		0,503	0,405	0,835
Información de los siguientes pasos a seguir en la tramitación	587,183	1	0,000	Rechazo Ho	Dependientes		0,852	0,461	0,894
Interés por parte de la compañía en solucionar la prestación	515,557	1	0,000	Rechazo Ho	Dependientes		0,546	0,437	0,869
Sencillez del lenguaje utilizado por el personal de la compañía	428,525	1	0,000	Rechazo Ho	Dependientes		0,497	0,396	0,832
Valoración del alcance del siniestro o de los daños	11,806	3	0,008	Rechazo Ho	Dependientes	8,6%			
Rapidez en la peritación	141,542	1	0,000	Rechazo Ho	Dependientes		0,286	0,255	0,601
Rapidez en la reparación	325,320	1	0,000	Rechazo Ho	Dependientes		0,434	0,381	0,780
Rapidez en la tramitación completa	516,512	1	0,000	Rechazo Ho	Dependientes		0,546	0,454	0,868
Calidad de la reparación	365,284	1	0,000	Rechazo Ho	Dependientes		0,459	0,399	0,802
Trato personalizado	516,278	1	0,000	Rechazo Ho	Dependientes		0,546	0,442	0,868
Iniciativa por parte de la compañía	508,703	1	0,000	Rechazo Ho	Dependientes		0,542	0,456	0,866
Adaptación del seguro a tus necesidades	633,706	1	0,000	Rechazo Ho	Dependientes		0,605	0,499	0,906
Adecuada relación entre precio de la póliza - coberturas	620,700	1	0,000	Rechazo Ho	Dependientes		0,599	0,504	0,905
Esfuerzo realizado para la gestión de la prestación	192,462	1	0,000	Rechazo Ho	Dependientes		0,333	0,292	0,647
Expectativas de servicio	323,618	2	0,000	Rechazo Ho	Dependientes		0,374	0,336	0,739

Fuente: Elaboración propia

A.2. Análisis de tablas de contingencia entre las variables independientes con la dependiente.

Para poder analizar la relación de dependencia o independencia entre dos variables cualitativas nominales, hay que estudiar su distribución conjunta o tabla de contingencia.

La tabla de contingencia es una tabla de doble entrada, donde en cada espacio figura el número de casos o individuos que poseen un nivel de uno de los factores o características analizadas y otro nivel del otro factor analizado. La tabla de contingencia se define por el número de variables que se analizan conjuntamente y el número de clases o niveles de los mismos. El formato de la misma es el siguiente:

		Recomendación (variable j)		
		Total con intención (7-10)	Total sin intención (1-6)	Marginal
Variable i	Categoría a	n_{11}	n_{12}	$n_{1.}$
	Categoría b	n_{21}	n_{22}	$n_{2.}$
	Marginal	$n_{.1}$	$n_{.2}$	N

n_{ij} = número de observaciones que tienen el atributo i y j

$n_{i.}$ = número de individuos que tienen el atributo i (marginal i)

$n_{.j}$ = número de individuos que tienen el atributo j (marginal j)

Las tablas de contingencia tienen dos objetivos fundamentales:

- Organizar la información contenida en una muestra cuando está referida a dos factores (variables cualitativas).
- Analizar si existe alguna relación de dependencia o independencia entre los niveles de las variables cualitativas objeto de estudio. El hecho de que dos variables sean independiente significa que los valores de una de ellas no están influidos por la modalidad o nivel que adopte la otra.

El listado completo de estas tablas de contingencia está incluido en el anexo, como ejemplo se analiza la siguiente, referente al estado del siniestro:

Tabla 2.24. Tabla de contingencia entre la variable recomendación y estado de siniestro.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
4. ¿En qué estado se encuentra tu siniestro?	Reparados	Recuento	926	223	1149
		% dentro de 4. ¿En qué estado se encuentra tu siniestro?	80,6%	19,4%	100,0%
	Indemnizados	Recuento	297	155	452
		% dentro de 4. ¿En qué estado se encuentra tu siniestro?	65,7%	34,3%	100,0%
Total	Recuento	1223	378	1601	
	% dentro de 4. ¿En qué estado se encuentra tu siniestro?	76,4%	23,6%	100,0%	

Fuente: Elaboración propia

Las prestaciones que han seguido la vía reparatoria (80,6%) dan como resultado un nivel de recomendación mayor que las prestaciones indemnizatorias (65,7%).

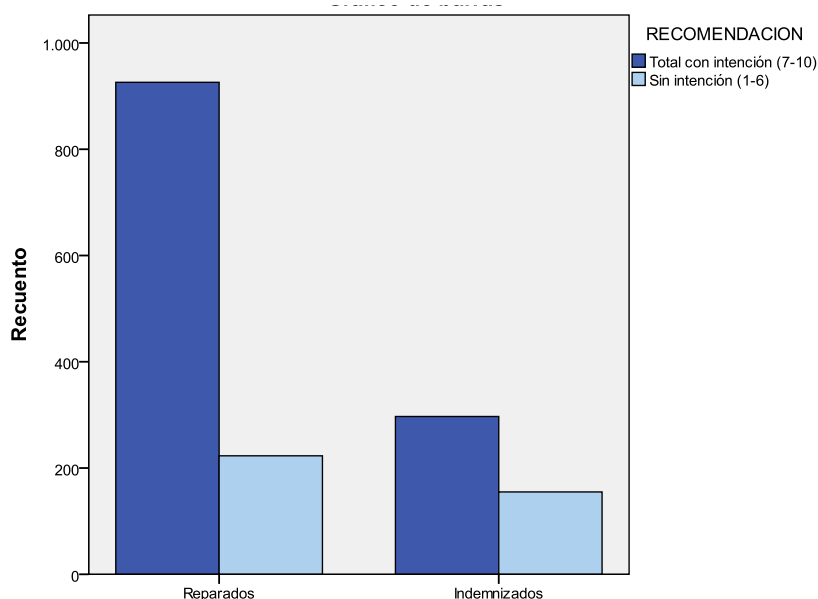


Figura 2.22. Gráfico de barras de la variable recomendación y estado de siniestro.
Fuente: Elaboración propia.

Para ver qué características presentan los clientes que finalmente recomiendan la entidad se hace el mismo ejercicio con el resto de tablas del anexo.

Se ha elegido esta tabla como muestra para introducir la posibilidad de analizar en el apartado 2.3. la posibilidad de establecer modelizaciones diferentes en función de esta tipología.

2.1.2. B. Relación entre variables independientes

Una vez analizado la relación entre la variable dependiente y las variables independientes hay que analizar la relación existente entre los propios factores. Es necesario analizar si hay variables independientes muy relacionadas entre sí ya que incluir variables en el modelo con una alta relación entre si supone poder descartar variables que no son necesarias y que aportan la misma información.

Para analizar esta relación se ha optado por emplear el coeficiente de correlación de spearman, se han resaltado en azul las correlaciones más altas. Para mostrar la información se han seleccionado únicamente las variables que presentaban alguna correlación elevada y aparecen en la figura 2.38.

	RECOMENDACIÓN	Intención de continuar	Opinión general	Rapidez en la comunic.	Facilidad de contacto con la entidad	Información de los siguientes pasos a seguir	Interés por parte de la compañía en solucionar la prestación	Rapidez en la peritación	Valoración del alcance del siniestro	Rapidez en la reparación	Rapidez en la tramitación completa	Calidad de la reparación	Trato personalizado	Iniciativa por parte de la compañía	Adaptación del seguro a tus necesidades	Adecuada relación entre precio de la póliza y coberturas
RECOMENDACIÓN	1,000															
Intención de continuar	,861	1,000														
Opinión general	,893	,909	1,000													
Rapidez en la comunicación	,644	,651	,683	1,000												
Facilidad de contacto con la entidad	,636	,662	,685	,840	1,000											
Información de los siguientes pasos a seguir	,673	,675	,703	,829	,826	1,000										
Interés por parte de la compañía en solucionar la prestación	,671	,680	,706	,838	,832	,851	1,000									
Sencillez del lenguaje utilizado por el personal	,632	,648	,679	,799	,835	,836	,833									
Rapidez en la peritación	,638	,644	,675	,776	,777	,759	,761	1,000	-,011							
Rapidez en la reparación	,599	,622	,666	,733	,703	,730	,740	,780	,124	1,000						
Rapidez en la tramitación completa	,636	,637	,681	,783	,755	,776	,782	,825	,110	,868	1,000					
Calidad de la reparación	,625	,624	,667	,691	,682	,710	,725	,730	,091	,777	,763	1,000				
Trato personalizado	,646	,642	,681	,742	,754	,755	,769	,781	,079	,761	,783	,835	1,000			
Iniciativa por parte de la compañía	,664	,660	,696	,748	,759	,768	,794	,778	,071	,771	,803	,824	,863	1,000		
Adaptación del seguro a tus necesidades	,722	,728	,759	,724	,739	,746	,759	,747	,035	,705	,732	,704	,740	,768	1,000	
Adecuada relación entre precio de la póliza y coberturas	,715	,726	,759	,708	,712	,728	,731	,714	,035	,693	,727	,684	,723	,747	,877	1,000

Figura 2.23. Matriz de correlaciones de variables. Fuente: Elaboración propia.

2.3. Modelo Lineal Generalizado (GLM)

2.3.1. Especificación del modelo teórico

Los apartados anteriores nos han servido tanto para conocer las variables que pueden o no influir en la recomendación así como la relación que presentan entre ellas y con la variable dependiente. El siguiente paso es elaborar un modelo que nos permita saber cuáles de éstas tienen un impacto significativo en la recomendación de la compañía. Para ello se ha optado por elaborar un modelo lineal generalizado (Generalized Linear Model).

Los pioneros de estos modelos fueron Nelder y Wedderburn (1972), y se trata de una extensión de los habituales modelos lineales. Éstos buscan explicar una variable dependiente (y), a través de un conjunto de variables independientes o explicativas (x), con la salvedad de que con los GLM no se necesita que ambas variables tengan entre ellas una relación de carácter lineal. Los GLM son, por tanto, una prolongación de los modelos lineales que permiten utilizar distribuciones no normales de los errores (binomiales, poisson, gamma, etc) y varianzas no constantes.

Esta modelización tiene una serie de características que la hace diferentes al modelo lineal general. Sabemos que los modelos lineales se basan en los siguientes supuestos:

- a) La variable dependiente se relaciona linealmente con las independientes.
- b) Los errores se distribuyen siguiendo la ley normal.
- c) La varianza es constante.

Éstos modelos asumen que la variable respuesta es una variable aleatoria que sigue una distribución normal con media y varianza constante. La mayoría de las veces estos supuestos no se cumplen y se necesita realizar alguna transformación de la variable respuesta. Como alternativa se presentan los modelos lineales generalizados, que presentan una serie de características propias:

- a) La variable dependiente no sigue una distribución normal.
- b) La relación entre la variable dependiente y las independientes no tiene por qué ser lineal.
- c) Los residuos no tienen por qué tener varianza constante (homocedasticidad)
- d) Existe una relación lineal entre las variables explicativas y una transformación de la media de la variable respuesta.

En este caso nos encontramos con una variable respuesta no cuantitativa, sino, como ya hemos visto, se trata de una variable cualitativa binaria (recomendación – no recomendación).

Se trata de elaborar un modelo de recomendación de cliente en el cual la variable respuesta va a ser si el cliente va a recomendar a la entidad o no. En esta situación un modelo de regresión lineal estimado por mínimos cuadrados no garantiza que la

predicción tome valores en el intervalo $[0,1]$. La forma de asegurarnos que la predicción de la variable recomendación tome valores en este intervalo, es transformar la variable, lo que se asimila a establecer una relación no lineal entre la variable respuesta y las explicativas. El modelo se reformularía como:

$$E[y_i / x = x_i] = p_i = F(x_i' \beta)$$

Donde F es cualquier función de distribución, de garantiza que $\hat{p}_i \in [0,1]$ para cualquier valor de $x_i' \hat{\beta}$.

A esta F se la denomina *función vínculo*. La función vínculo se encarga de *linealizar* la relación entre la variable dependiente y las variables independientes mediante la transformación de la variable respuesta.

Debido a que apenas hay diferencias con otros modelos lineales generalizados y que los resultados son más fáciles de interpretar, nos vamos a centrar en el modelo *logit*, que aparece cuando F es la función de distribución logística que hace que la probabilidad de ocurrencia venga definida por la siguiente función:

$$p_i = \frac{1}{1 + e^{-x_i' \beta}} \rightarrow (1 - p_i) = \frac{1}{1 + e^{x_i' \beta}}$$

Y si despejamos $x_i' \beta$, se tiene que

$$x_i' \beta = \log\left(\frac{p_i}{1 - p_i}\right) = \text{Logit}(p_i)$$

Obteniendo de la transformación un modelo lineal que se denomina *logit*, la variable *Logit* (p_i) representa la diferencia entre las probabilidades de pertenecer a ambas clases (recomendación ($y_i = 1$) o no recomendación ($y_i = 0$)).

$$p_i = P(y_i = 1 | x = x_i), \quad (1 - p_i) = P(y_i = 0 | x = x_i)$$

La estimación de estos modelos se realiza por el método de máxima verosimilitud, dado un conjunto de n observaciones independientes, el proceso de estimación se compone den los siguientes pasos:

1. En primer lugar hay que establecer cómo va a ser la función de verosimilitud:

$$L(\beta; X, y) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{(1-y_i)} = \prod_{i=1}^n F(x_i' \beta)^{(y_i)} [1 - F(x_i' \beta)]^{(1-y_i)}$$

Donde F es la función de distribución logística.

2. El siguiente paso es calcular el logaritmo neperiano de $L(\beta; X, y)$, (el punto que maximiza L también minimiza $\log L$)

$$\log L(\beta; X, y) = \sum_{i=1}^n (y_i \log[F(x'_i \beta)] + (1 - y_i) \log[1 - F(x'_i \beta)])$$

3. Posteriormente hay que calcular las derivadas de $\log L(\beta; X, y)$ respecto a los β_j , igualamos a cero estas derivadas y se obtiene un sistema de $k + 1$ ecuaciones (tantas como parámetros a estimar).

$$\frac{d \log L(\beta; X, y)}{d\beta_j} = \sum_{i=1}^n \left(y_i \frac{\frac{dF(x'_i \beta)}{d\beta_j}}{F(x'_i \beta)} + (1 - y_i) \frac{-\frac{dF(x'_i \beta)}{d\beta_j}}{1 - F(x'_i \beta)} \right) x_{ji} = 0$$

4. Ya que el sistema de ecuaciones obtenido, como hemos visto, no es lineal en los parámetros β_j la solución $\hat{\beta}$ se obtendrá a través de algún algoritmo iterativo. En nuestro caso será el programa SPSS el que realizará los cálculos necesarios. A partir de la matriz de segundas derivadas se obtendrá la varianza estimada de $\hat{\beta}$.

Cuando se emplean Modelos Lineales Generalizados, la medida equivalente que se interpreta como la suma residual de cuadrados en los modelos de regresión lineal clásico, toma el nombre de *deviance* (D). Ésta se obtiene a partir de la siguiente ecuación:

$$D = 2 \sum_{i=1}^N \left[y_i \log \left(\frac{y_i}{\hat{y}_i} \right) + (n_i - y_i) \log \left(\frac{n_i - y_i}{n_i - \hat{y}_i} \right) \right]$$

De tal forma que:

$$D = 2 \sum_{i=1}^N [o \log(o/e)]$$

Donde o son los “éxitos” (y_i) y “fallos” ($n_i - y_i$) observados, y los e denotan las correspondientes frecuencias esperadas estimadas o valor ajustado $\hat{y}_i = n_i \hat{\pi}_i$ y $(n_i - y_i) = n_i - n_i \hat{\pi}_i$.

Resumiendo, este tipo de modelos tienen tres componentes básicos:

- Componente aleatoria: se encarga de identificar a la variable respuesta y a su distribución de probabilidad. Este componente es una variable aleatoria Y con n observaciones independientes.

Como ya se ha comentado, en nuestro caso, Y_i indica el número de éxitos de un número fijo de ensayos, y se modeliza como una binomial. En otras situaciones se puede asignar a una distribución de poisson, binomial negativa o incluso a una normal. Todos estos modelos se pueden incluir dentro de la llamada familia exponencial.

Centrándonos en el caso que nos ocupa, el proceso toma su nombre de aquel que lo definió, Bernoulli. Éste proceso consiste en realizar un experimento aleatorio una sola vez y observar si cierto suceso ocurre o no, definiendo a p como la probabilidad de que ocurra ese suceso (éxito) y a $q = 1 - p$, como la probabilidad de que no ocurra (fracaso), por lo que la variable sólo puede tomar dos posibles valores $(p, 1 - p)$

La distribución binomial es una generalización de la distribución de Bernoulli, que tiene lugar cuando en lugar de realizar el experimento aleatorio una sola vez, se realiza n veces, siendo cada prueba independiente de la anterior. Se trata de una distribución de probabilidad discreta que mide el número de éxitos cuando tenemos una variable aleatoria discreta, de tal forma que sólo puede tomar los valores $0, 1, 2, 3, 4, \dots, n$, (suponiendo que se han realizado n pruebas). La distribución binomial fue desarrollada por Jakob Bernoulli (Suiza, 1654-1705) y es la principal distribución de probabilidad discreta para variables que sólo pueden tomar dos posibles resultados (dicotómicas).

Se dice que X sigue una distribución Binomial de parámetros n y p , que se representa con la siguiente notación:

$$X \sim B(n, p)$$

Su función de probabilidad viene definida por:

$$F(X = x) = \binom{n}{x} * p^x * (1 - p)^{n-x}$$

Donde n debe ser un entero positivo y p debe pertenecer al intervalo $[0, 1]$. Su media y su varianza se obtienen y a partir de las siguientes expresiones:

$$E(x) = n * p$$

$$\sigma^2 = n * p * (1 - p)$$

- Componente sistemática: es la parte que nos indica cuales van a ser las variables explicativas (independientes) utilizadas en la función predictora lineal, a esta combinación lineal se denomina *predictor lineal* y toma la siguiente forma:

$$\alpha + \beta_1 x_1 + \dots + \beta_k x_k$$

- Función link: Es una función del valor esperado de Y (recomendación), y se define como una combinación lineal de las variables independientes. Tenemos el valor esperado, $\mu = E(Y)$, la función link especifica una función $g(\cdot)$ que relaciona μ con el predictor lineal de la siguiente forma:

$$g(\mu) = \alpha + \beta_1 x_1 + \dots + \beta_k x_k$$

De este modo, la función link g relaciona las componentes aleatoria y sistemática. La función g más simple es $g(\mu) = \mu$, esto es, la identidad que da lugar al modelo de regresión lineal clásico.

Las funciones de ligadura (link) más empleadas en este tipo de modelización son:

Tabla 2.25. Funciones link

Función vínculo	Fórmula
Identidad	μ
Logarítmica	$\text{Log}(\mu)$
Logit	$\text{Log}\left(\frac{\mu}{n - \mu}\right)$
Recíproca	$1/\mu$
Raíz cuadrada	$\sqrt{\mu}$
Exponencial	μ^n

Fuente: Elaboración propia

Como ya se ha comentado anteriormente, en vista de que es el modelo que mejor se va a ajustar a los datos y puesto que el empleo de otro tipo de modelos más complicados refleja resultados similares, se va a emplear la función logit para una variable respuesta binaria.

2.3.2. Selección y evaluación del modelo

El SPSS permite emplear varios métodos que ayudan a realizar la selección de variables y no tener que realizar manualmente todo el proceso, para ellos se ha escogido el *método por pasos hacia delante a razón de verosimilitud*, se trata de un proceso que contrasta la entrada de factores basándose en la significación del estadístico de puntuación y contrasta la eliminación según la probabilidad del estadístico a razón de verosimilitud, que se sustenta en estimaciones de máxima verosimilitud parcial.

De todas las iteraciones que realiza el programa, bajo el principio de parsimonia (el modelo más reducido que explica la recomendación), se obtiene el modelo final que vamos a analizar. En la siguiente tabla se muestran los pasos del algoritmo iterativo para estimar la verosimilitud del modelo ajustado.

Tabla 2.26. Historial de iteraciones.

Iteración	-2 log de la verosimilitud	Coeficientes								
		Constant	Ligado a hipoteca	Información de los siguientes pasos a seguir	Calidad de la reparación	Relación entre precio -coberturas	Esfuerzo realizado para la gestión	Sí perito	Intención de continuar	Satisfacción general
1	642,484	-4,820	-,171	,369	,203	,216	-,181	-,198	1,405	1,455
2	532,937	-6,420	-,388	,664	,448	,427	-,428	-,495	1,956	1,840
3	513,746	-7,022	-,581	,824	,649	,563	-,661	-,803	2,257	2,058
Paso 8 4	512,441	-7,173	-,655	,867	,720	,602	-,752	-,929	2,362	2,150
5	512,431	-7,185	-,663	,870	,727	,605	-,761	-,942	2,373	2,160
6	512,431	-7,185	-,663	,870	,727	,605	-,762	-,942	2,373	2,160

a. Método: Por pasos hacia adelante (Razón de verosimilitud)

b. En el modelo se incluye una constante.

c. -2 log de la verosimilitud inicial: 1298,867

d. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: Elaboración propia

A partir de aquí puede contrastarse la significación global del modelo mediante el estadístico:

$$L_{exp} = -2 \log L_0 - (-2 \log L) = 1.298,867 - 512,431 = 786.44 \sim \chi_8^2$$

La hipótesis a contrastar es:

$$H_0 : \beta_1 = \dots = \beta_k = 0$$

Que indica que los coeficientes no son significativos conjuntamente. La hipótesis alternativa sería pues:

$$H_0 : \beta_1 \neq \dots \neq \beta_k \neq 0$$

Que supone que los coeficientes si son significativos conjuntamente y por lo cual el modelo pasa a ser significativo a nivel global.

Se rechaza H_0 al 5% de significación ya que $L_{exp} = 786.44 > x_{8, 95\%}^2 = 15,507$, por lo que no se puede aceptar la hipótesis nula de no significación conjunta de los parámetros. Los resultados vienen reflejados en la siguiente tabla:

Tabla 2.27. Pruebas omnibus sobre los coeficientes del modelo.

		Chi cuadrado	gl	Sig.
	Paso	4,054	1	,044
Paso 8	Bloque	786,436	8	,000
	Modelo	786,436	8	,000

Fuente: Elaboración propia.

Ya se ha comprobado que el modelo es significativo globalmente. El R^2 de Cox y Snell vale 48,6% y el de Nagelkerke (versión mejorada R^2 de Cox y Snell) vale 72,9%.

La verosimilitud del modelo (-2log de la verosimilitud, deviance) se calcula mediante el ratio entre la verosimilitud del modelo propuesto sobre un modelo que contiene únicamente la constante (en el que las variables independientes no tienen efecto). Así, la verosimilitud es una medida relativa que nos permite contrastar modelos, de tal forma que el “mejor modelo” será aquel que minimice su valor.

Tabla 2.28. Resumen del modelo.

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
8	512,431 ^a	,486	,729

Fuente: Elaboración propia.

Los indicadores de la tabla 2.28 nos dan una idea de lo que ganamos utilizando el modelo, pero no nos dicen qué capacidad tenemos de predecir los datos. Si analizamos la tabla de clasificación que aparece más abajo (tabla 2.29), el 92,2% de los casos está correctamente clasificado y la contra diagonal está bastante equilibrada (celdas sombreadas en naranja en la tabla), por lo que dejamos el punto de corte en el 0.5.

Tabla 2.29. Tabla de clasificación.

Observado		Pronosticado		
		RECOMENDACION		Porcentaje correcto
		Total con intención (7-10)	Sin intención (1-6)	
Paso 8	Total con intención (7-10)	857	43	95,2
	Sin intención (1-6)	49	233	82,6
Porcentaje global				92,2

El valor de corte es ,500

Fuente: Elaboración propia.

La tabla 2.30 muestra los aspectos que finalmente se han seleccionado ya que son significativos a nivel individual, incluso al 5% de nivel de significación. Identificamos efectos negativos y positivos sobre la recomendación:

- Las variables con efectos negativos ordenadas de más a menos serían:
 - Envío de perito para valorar los daños.
 - Esfuerzo realizado para gestionar la prestación.
 - Seguros ligados a préstamo hipotecario.

- Las variables que afectan de manera positiva son:
 - Intención de continuar
 - Satisfacción general
 - Información de los siguientes pasos a seguir en la tramitación
 - Calidad de la reparación
 - Adecuada relación entre precio de la póliza y coberturas incluidas

2.3.4. Análisis de los residuos

Los residuos son la diferencia entre el valor observado y el pronosticado. Este componente nos informa del grado de exactitud de los pronósticos.

El primer paso es analizar el histograma de los residuos. En la figura 2.25 se aprecia que la parte central acumula más casos que los que existen en una normal.

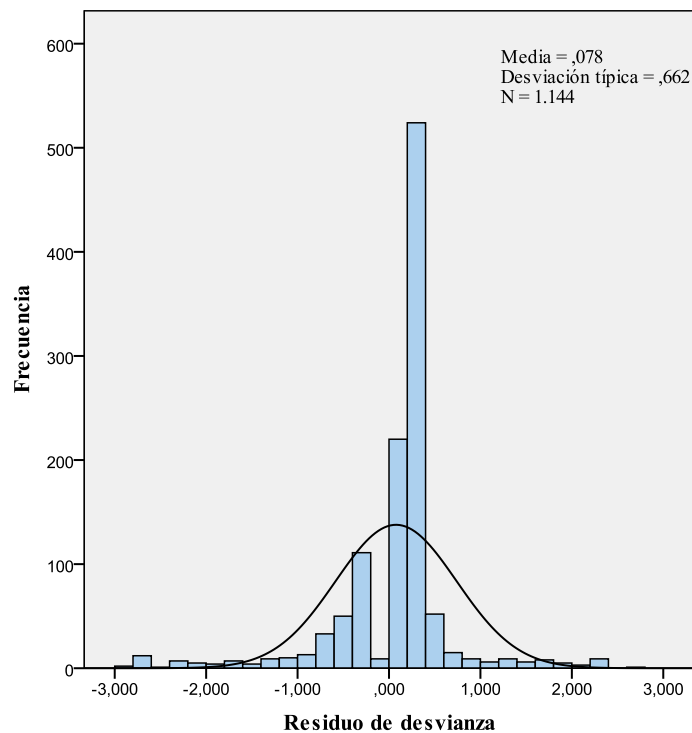


Figura 2.25. Histograma de los residuos de desviación.
Fuente: Elaboración propia.

Para asegurarnos se analizan los contrastes de normalidad de Kolmoronov – Smirnov y de Shapiro Wilk.

Tabla 2.31. Prueba de normalidad Kolmogorov-Smirnov.

	Kolmogorov-Smirnova		
	Estadístico	gl	Sig.
Residuo de desviación	,277	1144	,000

Fuente: Elaboración propia.

Tabla 2.32. Prueba de normalidad Shapiro-Wilk.

	Shapiro-Wilk		
	Estadístico	gl	Sig.
Residuo de desviación	,774	1144	,000

Fuente: Elaboración propia.

El nivel de significación de ambos contrastes indica que el residuo de la desviación del modelo no sigue una distribución normal.

Otra forma de ver la estructura de los residuos de la desviación es con el gráfico Q-Q Normal, cuanto más explique el modelo más cerca estarán los dos extremos y más concentrados los casos en esos extremos. En la figura 2.26 se muestra la relación entre los residuos de la desviación y las predicciones del modelo. De nuevo, cuanto más explique el modelo más cerca estarán los dos extremos.

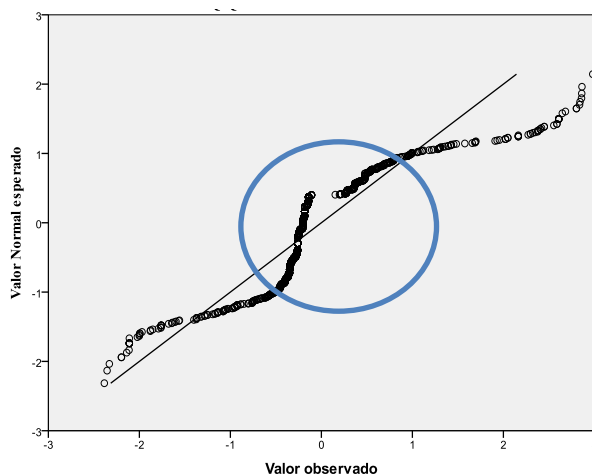


Figura 2.26. Gráfico Q-Q Normal de valor de desviación. Fuente: Elaboración propia.

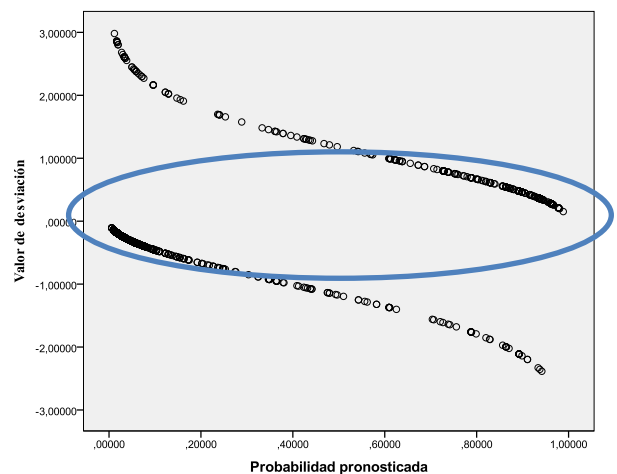


Figura 2.27. Gráfico de dispersión de los residuos de la desviación – predicción del modelo. Fuente: Elaboración propia.

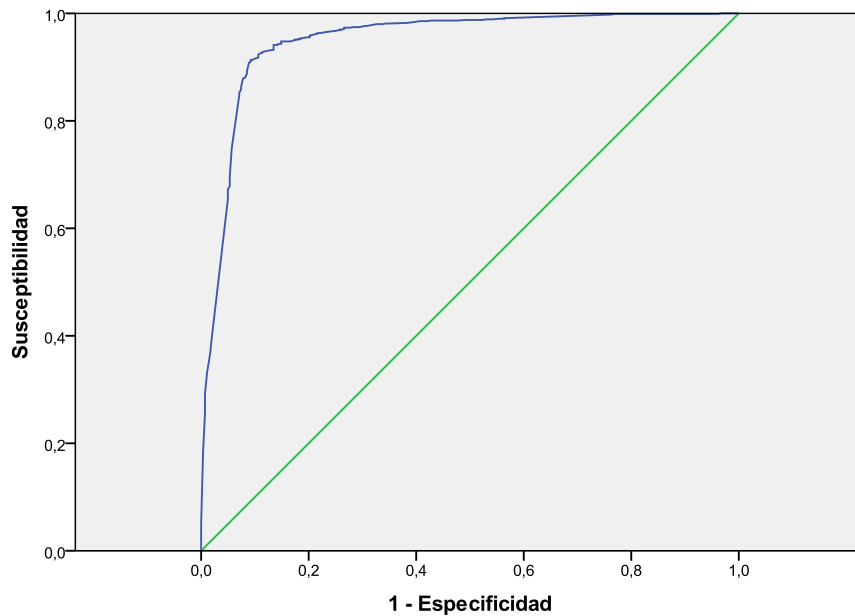
2.3.5. Validación del modelo

Para ver si el modelo escogido es una herramienta válida vamos a fijarnos en varios aspectos.

El primero de ellos es analizar la curva ROC. Se trata de una herramienta estadística utilizada para medir la capacidad discriminante de una prueba basada en una variable de decisión (recomendación o no), la finalidad es clasificar a los individuos de una población en dos grupos: uno que presente un evento de interés y otro que no, en nuestro caso el cliente que va a recomendar y el que no. Esta capacidad de diferenciar sujetos está sujeta al *valor umbral* elegido de entre todos los posibles resultados de la variable de decisión, es decir, la variable por cuyo resultado se clasifica a cada individuo en un grupo u otro. La curva es el gráfico resultante de representar, para cada *valor umbral*, las medidas de sensibilidad y especificidad de la prueba.

En nuestro caso la sensibilidad mide el número de clientes que verdaderamente recomienda, es la relación entre los que de verdad recomiendan y el total clientes. La especificidad mide el grupo de clientes que no recomiendan (verdaderos negativos / total clientes).

A continuación, se muestra el dibujo de la curva ROC ajustada a nuestros datos:



Los segmentos diagonales son producidos por los empates.

Figura 2.28. Curva COR. Fuente: Elaboración propia.

Tabla 2.33. Área bajo la curva COR.

Variables resultado de contraste: Valor pronosticado del predictor lineal				
asintótico al 95%				
Área	Error típ.a	Sig. asintóticab	Límite	
			Límite inferior	superior
,949	,008	,000	,933	,965

Fuente: Elaboración propia.

Sabemos que cuanto más alejada este la curva ROC de la diagonal principal mejor es el método de clasificación, ya que la curva ROC ideal sería la que con una especificidad de 1 tuviera una sensibilidad de 1, y cuanto más cercana esté a dicha diagonal peor será el método de clasificación. La diagonal principal es la que corresponde a la peor clasificación y que tienen un área bajo ella de 0.5. En nuestro caso el método escogido parece ser bastante correcto ya que la curva ROC está muy alejada de la diagonal y el área bajo la curva es muy próximo a 1 (0,949).

El segundo paso realizado para validar fehacientemente si el modelo consigue predecir o no la recomendación del cliente es contrastar este resultado sobre el 20% de la muestra reservada para este fin.

Para ello hemos formulado el modelo predictivo con las variables y coeficientes que nos obteníamos en los apartados anteriores de tal forma que:

$$P = \frac{1}{1 + e^{-(c+x_i' \beta)}}$$

Sabiendo que:

$$c + x_i' \beta = -7,185 + (-0,662 * \text{ligado a hipoteca}) \\ + (0,870 * \text{informacion de los pasos a seguir}) \\ + (0,726 * \text{calidad de la reparación}) + (0,604 * \text{relacion precio coberturas}) \\ + (-0,761 * \text{esfuerzo en la gestión}) + (-0,941 * \text{períto}) \\ + (2,372 * \text{intención de continuar}) + (2,159 * \text{opinión general})$$

El resultado se muestra en la siguiente tabla (2.34), el modelo logra clasificar al 91% de los clientes correctamente.

Tabla 2.34. Tabla de clasificación para validar el modelo.

		RECOMENDACION		
		Total con intención (7-10)		Porcentaje correcto
		Recuento	Recuento	
RECOMENDACION	Total con intención (7-10)	262	9	96,7%
	Sin intención (1-6)	22	52	70,3%
Porcentaje global				91,0%

Fuente: Elaboración propia.

En vista de estas comprobaciones parece confirmarse que el modelo obtenido con el 80% de la muestra sirve para ajustar el resto de los datos con bastante fiabilidad.

2.4. Modelo Lineal Generalizado en función de la tipología del proceso. Reparado vs Indemnizado

Resulta interesante analizar la posibilidad de realizar una modelización distinta en función del proceso seguido por la prestación.

Hemos visto en el apartado 2.1.2. A.2 que existe una diferencia significativa en la recomendación en función del proceso seguido en el servicio. Las prestaciones que asume la compañía y como consecuencia tiene lugar una reparación, por norma general, obtienen unos niveles más altos de recomendación, mientras que las que han sido indemnizadas son más críticas en este sentido. En el siguiente gráfico se aprecia la diferencia en la recomendación que existe en las prestaciones de hogar en función del proceso:

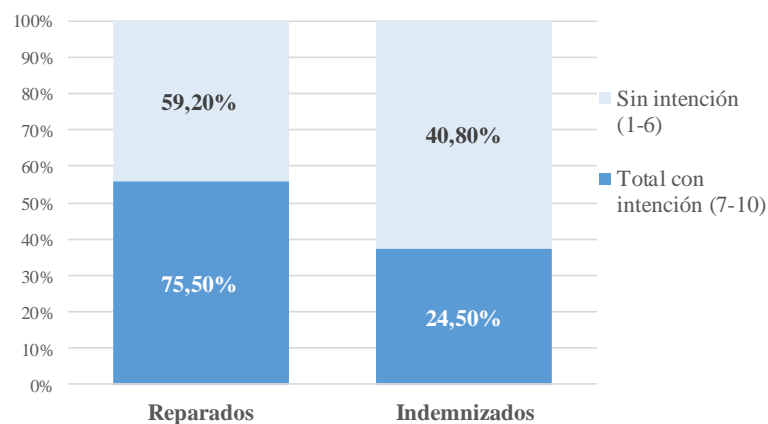


Figura 2.29. Recomendación en función del proceso seguido.
Fuente: Elaboración propia.

2.4. A Proceso reparatorio

En primer lugar, vamos a analizar las prestaciones del seguro de hogar que han tenido como consecuencia la realización de algún tipo de reparación.

Para realizar la modelización he seguido los pasos explicados en el apartado 2.3. de este trabajo, en esta ocasión el modelo con el cual se obtiene la deviance más pequeña con el menor número posible de parámetros significativos es el siguiente:

$$P = \frac{1}{1 + e^{-(c + x_i' \beta)}}$$

Sabiendo que:

$$c + x_i' \beta = -6,241 + (-0,759 * \text{ligado a hipoteca}) + (1,056 * \text{tratao personalizado}) \\ + (-0,868 * \text{esfuerzo en la gestión}) \\ + (2,2364 * \text{intención de continuar}) + (2,588 * \text{opinión general}) \\ + (-0.750 * \text{perito})$$

El modelo es significativo a nivel global ya que obtenemos un estadístico:

$$L_{exp} = 449,446 \sim \chi_6^2$$

La hipótesis a contrastar es:

$$H_0 : \beta_1 = \dots = \beta_k = 0$$

Se rechaza H_0 si $L_{exp} > \chi_{6, 95\%}^2 = 12,592$, no se puede aceptar la hipótesis nula, por lo que todos los parámetros a nivel global seleccionados son significativos.

El R^2 de Cox y Snell vale 42,7% y el de Nagelkerke (versión mejorada R^2 de Cox y Snell) vale 68,1%, más bajos que en el modelo global que hemos tomado como bueno anteriormente. Sin embargo, la tabla de clasificación nos da un resultado bastante acertado de nuevo:

Tabla 2.35. Tabla de clasificación para modelo reparatorio.

		Pronosticado			
		RECOMENDACION		Porcentaje correcto	
Observado		Total con intención (7-10)	Sin intención (1-6)		
Paso 6	RECOMENDACION	Total con intención (7-10)	623	27	95,8
		Sin intención (1-6)	35	122	77,7
Porcentaje global					92,3

a. El valor de corte es ,500

Fuente: Elaboración propia.

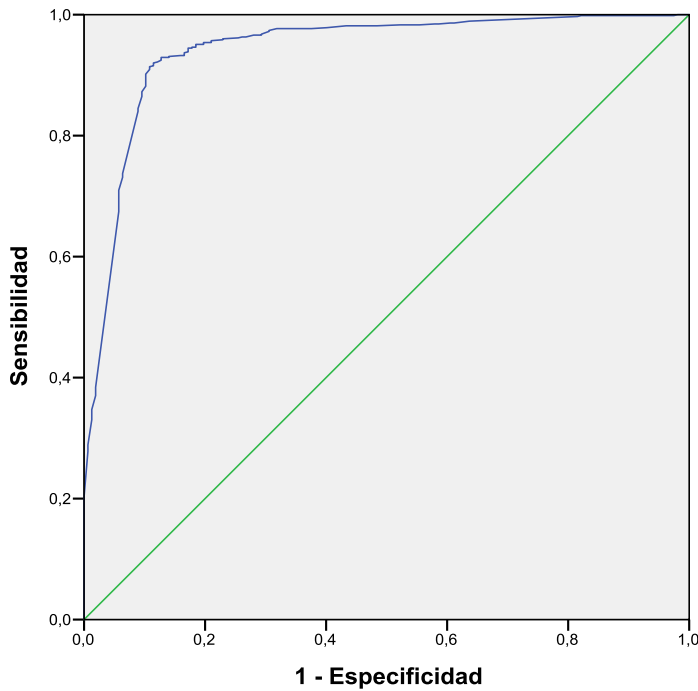
En este modelo clasifica correctamente al 92,3% de los encuestados, manteniendo así un alto porcentaje de aciertos en el nuevo modelo realizado para siniestros reparados. La contra diagonal también se mantiene equilibrada, por lo que parece indicar que el modelo es bastante fiable, para comprobarlo aplicamos los coeficientes al 20% de la muestra reservada:

Tabla 2.36. Tabla de clasificación para validar el modelo reparatorio.

		Pronosticado		
		RECOMENDACION		Porcentaje correcto
Observado		Total con intención (7-10)	Sin intención (1-6)	
RECOMENDACION	Total con intención (7-10)	201	7	96,6%
	Sin intención (1-6)	15	32	68,1%
Porcentaje global				91,4%

Fuente: Elaboración propia.

En el análisis de la curva COR vemos que el método clasifica de manera bastante correcta a los clientes ya que la curva está alejada de la diagonal principal y el área bajo ella es muy próxima a 1. (0.942).



Los segmentos de diagonal se generan mediante empates.

Figura 2.30. Curva COR recomendación. Fuente: Elaboración propia.

Tabla 2.37. Resumen del procesamiento de casos.

RECOMENDACION	N válido (por lista)
Positivo ^a	651
Negativo	157
Perdidos	70

Los valores más grandes de las variables de resultado de prueba indican una prueba mayor para un estado real positivo.

a. El estado real positivo es Total con intención (7-10).

Fuente: Elaboración propia.

Tabla 2.38. Área bajo la curva COR

Variables de resultado de prueba: Valor pronosticado del predictor lineal				
Área	Error estándar ^a	Significación ^b asintótica	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,942	,011	,000	,919	,964

Las variables de resultado de prueba: Valor pronosticado del predictor lineal tienen, como mínimo, un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Las estadísticas podrían estar sesgadas.

a. Bajo el supuesto no paramétrico
b. Hipótesis nula: área verdadera = 0,5

Fuente: Elaboración propia.

2.4. B Proceso indemnizatorio

Ahora se va a estudiar la posibilidad de realizar un modelo que ayude a predecir si el cliente que ha sufrido una prestación de hogar, solucionada vía indemnización, va a recomendar a la compañía o no.

La probabilidad de recomendar o no va a venir determinada por la siguiente función:

$$P = \frac{1}{1 + e^{-(c+x'_i \beta)}}$$

Sabiendo que:

$$c + x'_i \beta = -10,175 + (1,406 * \text{información de los siguientes pasos a seguir}) \\ + (1,989 * \text{relación precio coberuras}) \\ + (3,327 * \text{intención de continuar})$$

Comprobamos que el modelo es significativo a nivel global:

$$L_{exp} = 280,280 \sim \chi^2_3$$

La hipótesis a contrastar es:

$$H_0 : \beta_1 = \dots = \beta_k = 0$$

Se rechaza H_0 si $L_{exp} > \chi^2_{6, 95\%} = 0,352$, no se puede aceptar la hipótesis nula, por lo que todos los parámetros a nivel global seleccionados son significativos.

El R^2 de Cox y Snell vale 56,6% y el de Nagelkerke (versión mejorada R^2 de Cox y Snell) vale 77,7%, más altos que los dos modelos planteados anteriormente.

La tabla de clasificación es este proceso nos da un porcentaje de acierto en 91,1%.

Tabla 2.39. Tabla de clasificación para modelo indemnizatorio.

			Pronosticado		
			RECOMENDACION		Porcentaje correcto
Observado			Total con intención (7-10)	Sin intención (1-6)	
Paso 3	RECOMENDACION	Total con intención (7-10)	209	7	96,8
		Sin intención (1-6)	23	97	80,8
Porcentaje global					91,1

a. El valor de corte es ,500

Fuente: Elaboración propia.

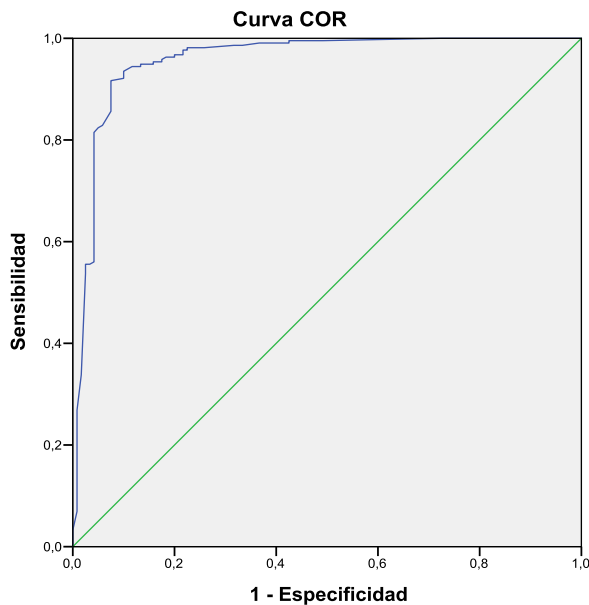
Aplicándolo al 20% reservado para la validación:

Tabla 2.40. Tabla de clasificación para modelo indemnizatorio.

Observado	RECOMENDACION	Pronosticado		Porcentaje correcto
		RECOMENDACION		
		Total con intención (7-10)	Sin intención (1-6)	
Total con intención (7-10)		61	2	96,8%
Sin intención (1-6)		10	17	63,0%
Porcentaje global				86,7%

Fuente: Elaboración propia.

De nuevo tenemos una curva alejada de la diagonal principal por lo que parece que el método escogido si está realizando una buena clasificación.



Los segmentos de diagonal se generan mediante empates.

Tabla 2.41. Resumen de procesamiento de casos.

RECOMENDACION	N válido (por lista)
Positivo ^a	216
Negativo	120
Perdidos	21

Los valores más grandes de las variables de resultado de prueba indican una prueba mayor para un estado real positivo.

a. El estado real positivo es Total con intención (7-10).

Fuente: Elaboración propia.

Figura 2.31. Curva COR recomendación. Fuente: Elaboración propia.

Tabla 2.42. Área bajo la curva COR

Variables de resultado de prueba: Valor pronosticado del predictor lineal

Área	Error estándar ^a	Significación ^b asintótica	95% de intervalo de confianza asintótico	
			Límite inferior	Límite superior
,957	,013	,000	,932	,982

Las variables de resultado de prueba: Valor pronosticado del predictor lineal tienen, como mínimo, un empate entre el grupo de estado real positivo y el grupo de estado real negativo. Las estadísticas podrían estar sesgadas.

a. Bajo el supuesto no paramétrico

b. Hipótesis nula: área verdadera = 0,5

Fuente: Elaboración propia

2.5. Modelo Lineal Generalizado de Alerta Temprana

Saber con antelación a la resolución de la prestación si el cliente es un potencial prescriptor o no puede resultar muy útil para las compañías, identificar a éstos clientes en las fases tempranas del proceso puede permitir cambiar el resultado del mismo.

Para ello se ha abordado la posibilidad de elaborar un modelo de “alerta temprana”. La idea es emplear la misma metodología que se viene desarrollando a lo largo de este trabajo pero en lugar de tomar como covariables independientes las que componen todo el proceso, se va a seleccionar una serie de aspectos previos a la prestación. De esta forma, se quiere conseguir analizar qué aspectos van a impactar más en la recomendación del cliente.

Las variables que se van a incluir como posibles factores causantes de recomendación en las primeras fases son los que aparecen en la siguiente tabla:

Tabla 2.43. Listado de variables seleccionadas para el análisis.

Variables independientes	
Edad	Tipología del siniestro
Número de seguros contratados	Rapidez en la comunicación
Antigüedad	Facilidad de contacto con la entidad
Zona Nielsen	Información de los siguientes pasos a seguir en la tramitación
Sexo	Interés por parte de la compañía en solucionar la prestación
Hábitat	Sencillez del lenguaje utilizado por el personal de la compañía
Vinculado a préstamo hipotecario	Canal de comunicación de la prestación

Fuente: Elaboración propia.

El mejor modelo obtenido con estas variables para predecir la recomendación o no del cliente pasa a ser el siguiente:

$$P = \frac{1}{1 + e^{-(c + x_i' \beta)}}$$

Sabiendo que:

$$\begin{aligned}
 c + x_i' \beta = & -4,655 + (-0,840 * \text{Ligado a hipoteca}) \\
 & + (0,758 * \text{rapidez en la comunicación}) \\
 & + (1,361 * \text{información en los siguientes pasos a seguir}) \\
 & + (0,966 * \text{interés por parte de la comoañía en solucionar el siniestro}) \\
 & + (0,687 * \text{sencillez en el lenguaje utilizado})
 \end{aligned}$$

La tabla de coeficientes sería:

Tabla 2.44. Variables en la ecuación.

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Paso 5a Ligado a hipoteca	-,840	,172	23,762	1	,000	,432	,308	,605
Rapidez en la comunicación	,758	,231	10,795	1	,001	2,134	1,358	3,354
Información de los siguientes pasos a seguir en la tramitación	1,361	,233	34,063	1	,000	3,901	2,470	6,162
Interés por parte de la compañía en solucionar la prestación	,966	,232	17,345	1	,000	2,629	1,668	4,142
Sencillez del lenguaje utilizado por el personal	,687	,210	10,705	1	,001	1,987	1,317	2,998
Constante	-4,655	,373	155,406	1	,000	,010		

a. Variable(s) introducida(s) en el paso 5: Sencillez del lenguaje utilizado por el personal.

Fuente: Elaboración propia.

La tabla de clasificación, con un porcentaje de acierto del 85,1%, sería la siguiente:

Tabla 2.45. Tabla de clasificación.

Observado	RECOMENDACION	Total con intención (7-10)	Pronosticado		
			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Porcentaje correcto
Paso 5			873	71	92,5
		Sin intención (1-6)	124	244	66,3
		Porcentaje global			85,1

Fuente: Elaboración propia.

Si el modelo lo extrapolamos al 20% reservado para comprobar el porcentaje de acierto, obtenemos la tabla 2.46, que se sitúa el porcentaje de acierto en el 84%, muy similar al obtenido con la muestra de referencia.

Tabla 2.46. Tabla de clasificación.

Observado	RECOMENDACION	Total con intención (7-10)	Pronosticado		Porcentaje correcto
			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	
			268	25	91,5%
		Sin intención (1-6)	38	62	62,0%
		Porcentaje global			84,0%

Fuente: Elaboración propia.

CAPITULO 3. ANÁLISIS DE RESULTADOS Y CONCLUSIONES

El objetivo de del análisis desarrollado en el trabajo es analizar los factores que inciden en la recomendación de los clientes que sufren una prestación en su seguro de hogar. Este mismo análisis y reflexión pueden extrapolarse a cualquier seguro / proceso. Las principales conclusiones obtenidas serían:

1. En primer lugar se ha obtenido un modelo global que nos ha dado como resultado un porcentaje elevado de clasificación correcta de clientes. De todas las variables de las que se dispone en el estudio, finalmente se establece que las características sociodemográficas no inciden de manera significativa en la recomendación del cliente, lo que realmente afecta es el propio proceso de prestación.

Mantener al cliente informado y que éste realice el menor esfuerzo en la gestión es tan importante como la propia calidad de la reparación. También veíamos que el envío del perito y el propio hecho de que el seguro de hogar venga ligado con un préstamo hipotecario inciden negativamente en la recomendación.

Por otro lado, las propias características del producto, como la relación precio coberturas también resulta una variable decisiva a la hora de recomendar la entidad, cuanto mejor relación mayor será la probabilidad de que el cliente recomiende el seguro.

Finalmente, cuanto mayor es la satisfacción con el proceso y por tanto la intención de continuar, más probable será que el cliente recomiende el servicio.

2. En el seguro de hogar existen dos vías de resolución, vía reparación o indemnización. Hay numerosos estudios que analizan las ventajas y desventajas de ambos procesos, aquí hemos comprobado que sí existen diferencias en la recomendación en ambos procesos, siendo más usual que el cliente que ha tenido una reparación recomiende más que aquel que ha tenido indemnización.

En las prestaciones que han seguido la vía *reparatoria*, finalmente, la personalización en trato (impacto positivo) como el envío del perito para valorar los daños (impacto negativo), son los factores que inciden de manera más significativa en la recomendación del cliente. Como ocurría en el modelo global, el hecho de que el seguro esté ligado a un préstamo hipotecario (impacto negativo), así como la opinión general del servicio y como resultado la intención de continuar con la entidad (impactos positivos) son los otros aspectos que nos ayudan a predecir si el cliente va a recomendar o no la entidad.

En las prestaciones que han seguido la vía *indemnizatoria* vemos que los factores que van a incidir significativamente en la recomendación son la información de los pasos a seguir en la tramitación, la relación precio coberturas y la propia intención de continuar el cliente en la entidad (todas con impacto positivo).

Como conclusión, en las prestaciones que han seguido la vía reparatoria cada fase del proceso no es tan importante como el resultado global del mismo, que se verá reflejado en la propia opinión general y en la intención de recomendar del cliente. En las prestaciones sujetas a indemnización la compañía debe hacer especial hincapié en transmitir al cliente la información sobre el propio proceso de tramitación.

3. Todos estos análisis se han realizado después de la prestación del servicio, son datos y modelos útiles para poder saber que está impactando en la prescripción de los clientes de la entidad y poder influir en las futuras prestaciones para lograr que se incremente la probabilidad de recomendación.

Pero ¿y si se pudiera saber de antemano a la finalización el proceso, si el cliente es un potencial prescriptor o no de la entidad? Sería interesante establecer unas características de “alerta temprana” que permita a la entidad saber si la probabilidad del cliente de recomendar es baja, y de esta forma incidir en el resto de proceso para poder cambiar el resultado previsto.

Para analizar esta posibilidad se han tomado como covariables las que corresponden a las primeras fases de la prestación y las propias a las características del cliente. En éste entorno se ha llegado a la conclusión de que la rapidez en la comunicación, la información sobre el proceso a seguir, el interés mostrado así como la sencillez del lenguaje son factores clave en las primeras fases del proceso. De nuevo las características inherentes al cliente no son determinantes a la hora de la recomendación. En la siguiente tabla se muestra las variables que se han empleado para modelar este aspecto, marcadas en azul están las que finalmente son significativas para el modelo.

Tabla 2.47. Variables incluidas en el modelo de alerta temprana.

Variables disponibles para modelizar “alerta temprana”	
Edad	Tipología del siniestro
Número de seguros contratados	Rapidez en la comunicación
Antigüedad	Facilidad de contacto con la entidad
Zona Nielsen	Información de los siguientes pasos a seguir en la tramitación
Sexo	Interés por parte de la compañía en solucionar la prestación
Hábitat	Sencillez del lenguaje utilizado por el personal de la compañía
Vinculado a préstamo hipotecario	Canal de comunicación de la prestación

Fuente: Elaboración propia.

CAPITULO 4. REFERENCIAS BIBLIOGRÁFICAS

- ADECOSE, 2018. *Estudio sobre la calidad del servicio en las Compañías Aseguradoras en relación con las corredurías asociadas*. Instituto de Actuarios Españoles
- AGUILERA KEYSER, Marcos, 2014. *Introducción a la Tarificación de la Cartera de Autos con SAS*
- Asociación para el desarrollo de la experiencia de cliente (DEC), 2017. *La Experiencia de Cliente Rentable*.
- BERLANGA SILVENTE, Vanesa y VILÀ BAÑOS, Ruth, 2014. Cómo obtener un Modelo de Regresión Logística Binaria con SPSS. *Revista d'Innovació i Recerca en Educació*, [en línea]. 7 (2), 105-118 [consulta: abril 2018]. Disponible en: <http://www.ub.edu/ice/reire.htm>
- CERDA, Jaime, CIFUENTES, Lorena, 2010. Uso de tests diagnósticos en la práctica clínica (Parte 1). Análisis de las propiedades de un test diagnóstico. *Rev Chil Infectol*.
- CERDA, Jaime, CIFUENTES, Lorena, 2010. Uso de tests diagnósticos en la práctica clínica (Parte 2). Aplicación clínica y utilidad de un test diagnóstico. *Rev Chil Infectol*.
- DOBSON, Annette y BARRET, Adrian, 1945. *An introduction to linear models*. 3rd edition.
- ICEA, 2018. *Análisis técnico de los seguros multirriesgo. Año 2017 (Informe nº1.508)*
- MARTÍN CABELLO, Juan Alfonso, 2015. *Análisis e inclusión de variables exógenas en la tarificación de autos mediante modelización por GLM*. Fundación Mapfre.
- MARTÍNEZ RUIZ-ANDREU, Pilar, 2016. *Gestión integral del multirriesgo de hogar de la cobertura de robo y modelización actuarial aplicando técnica GLM (variables intrínsecas, exógenas y de comportamiento)*. Fundación Mapfre.
- MCCULLAGAH, Peter y NELDER, John Ashworth. 1952. *Generalized Linear Models*. 2nd edition.
- PARDO, Antonio, RUIZ, Miguel Ángel y SAN MARTÍN, Rafael, 2007. *Psicothema*, [en línea]. Vol. 19, nº 2, pp. 308-321 [consulta: abril 2018]. Disponible en: <http://www.psycothema.com/psycothema.asp?id=3365>
- PAVÍA, Jose M, 2016. *Modelos lineales generalizados en economía y ciencias sociales*.

- PÉREZ LÓPEZ, Cesar, 2004. *Técnicas de análisis multivariante de datos*.
- WILDER Tukey, John, 1977. *Exploratory Data Analysis*.

CAPITULO 5. ANEXOS

Anexo I. Tablas de contingencia

Edad

Tabla A.1. Tabla de contingencia edad - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
Edad	25 ó menos	Recuento	108	59	167
		% dentro de Edad	64,7%	35,3%	100,0%
	De 26 a 35 años	Recuento	142	67	209
		% dentro de Edad	67,9%	32,1%	100,0%
	De 36 a 45 años	Recuento	252	104	356
		% dentro de Edad	70,8%	29,2%	100,0%
	De 46 a 55 años	Recuento	309	110	419
		% dentro de Edad	73,7%	26,3%	100,0%
	De 56 a 65 años	Recuento	314	94	408
		% dentro de Edad	77,0%	23,0%	100,0%
	Más de 65 años	Recuento	131	41	172
		% dentro de Edad	76,2%	23,8%	100,0%
Total		Recuento	1256	475	1731
		% dentro de Edad	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Antigüedad

Tabla A.2. Tabla de contingencia antigüedad - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
Antigüedad	2 años	Recuento	295	137	432
		% dentro de Antigüedad	68,3%	31,7%	100,0%
	De 3 a 5 años	Recuento	402	153	555
		% dentro de Antigüedad	72,4%	27,6%	100,0%
	De 6 a 10 años	Recuento	317	111	428
		% dentro de Antigüedad	74,1%	25,9%	100,0%
	Más de 10 años	Recuento	219	64	283
		% dentro de Antigüedad	77,4%	22,6%	100,0%
Total		Recuento	1233	465	1698
		% dentro de Antigüedad	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Número de seguros contratados

Tabla A.3. Tabla de contingencia tipo de seguro - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
Tipo Seguro	Hogar	Recuento	1136	432	1568
	NSeguro 1	% dentro de Tipo Seguro	72,4%	27,6%	100,0%
	Hogar	Recuento	106	33	139
	NSeguro 2	% dentro de Tipo Seguro	76,3%	23,7%	100,0%
	Hogar	Recuento	14	10	24
	NSeguro 3	% dentro de Tipo Seguro	58,3%	41,7%	100,0%
Total		Recuento	1256	475	1731
		% dentro de Tipo Seguro	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Zona Nielsen

Tabla A.4. Tabla de contingencia zona nielsen - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
Zona Nielsen	Noreste / Catalunya y Balears	Recuento	129	39	168
		% dentro de Zona Nielsen	76,8%	23,2%	1,0
	Levante	Recuento	188	66	254
		% dentro de Zona Nielsen	74,0%	26,0%	100,0%
	Sur / Andalucía	Recuento	240	80	320
		% dentro de Zona Nielsen	75,0%	25,0%	100,0%
	Centro	Recuento	125	57	182
		% dentro de Zona Nielsen	68,7%	31,3%	100,0%
	Noroeste	Recuento	115	45	160
		% dentro de Zona Nielsen	71,9%	28,1%	100,0%
	Norte Centro	Recuento	119	50	169
		% dentro de Zona Nielsen	70,4%	29,6%	100,0%
	Canarias	Recuento	33	11	44
		% dentro de Zona Nielsen	75,0%	25,0%	100,0%
	AMB (Área Metropolitana Barcelona)	Recuento	120	48	168
		% dentro de Zona Nielsen	71,4%	28,6%	100,0%
	AMM (Área Metropolitana Madrid)	Recuento	249	90	339
		% dentro de Zona Nielsen	73,5%	26,5%	100,0%
Total		Recuento	1318	486	1804
		% dentro de Zona Nielsen	73,1%	26,9%	100,0%

Fuente: Elaboración propia.

Sexo

Tabla A.5. Tabla de contingencia sexo - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
Sexo	Hombre	Recuento	597	236	833
		% dentro de Sexo	71,7%	28,3%	100,0%
	Mujer	Recuento	659	239	898
		% dentro de Sexo	73,4%	26,6%	100,0%
Total		Recuento	1256	475	1731
		% dentro de Sexo	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Hábitat

Tabla A.6. Tabla de contingencia hábitat - recomendación

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
habitat	Urbano (Más de 50.000 habitantes)	Recuento	843	323	1166
		% dentro de habitat	72,3%	27,7%	100,0%
	Rural (Hasta 50.000 habitantes)	Recuento	413	152	565
		% dentro de habitat	73,1%	26,9%	100,0%
Total		Recuento	1256	475	1731
		% dentro de habitat	72,6%	27,4%	100,0%

. Fuente: Elaboración propia.

Intención de continuar

Tabla A.6. Tabla de contingencia intención de recomendar - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
De continuar el servicio como hasta ahora ¿mantendrías tu póliza con esta compañía?	Total con intención (7-10)	Recuento	1171	71	1242
		% dentro de De continuar el servicio como hasta ahora ¿mantendrías tu póliza con esta compañía?	94,3%	5,7%	100,0%
	Sin intención (1-6)	Recuento	85	404	489
		% dentro de De continuar el servicio como hasta ahora ¿mantendrías tu póliza con esta compañía?	17,4%	82,6%	100,0%
Total		Recuento	1256	475	1731
		% dentro de De continuar el servicio como hasta ahora ¿mantendrías tu póliza con esta compañía?	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Opinión general

Tabla A.7. Tabla de contingencia opinión general - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
¿Cuál es tu opinión general sobre el servicio de tu compañía de seguros?	Total con intención (7-10)	Recuento % dentro de ¿Cuál es tu opinión general sobre el servicio de tu compañía de seguros?	1181 94,2%	73 5,8%	1254 100,0%
	Sin intención (1-6)	Recuento % dentro de ¿Cuál es tu opinión general sobre el servicio de tu compañía de seguros?	75 15,7%	402 84,3%	477 100,0%
Total		Recuento % dentro de ¿Cuál es tu opinión general sobre el servicio de tu compañía de seguros?	1256 72,6%	475 27,4%	1731 100,0%

Fuente: Elaboración propia.

Vinculación a préstamo hipotecario

Tabla A.8. Tabla de contingencia vinculación a préstamo hipotecario - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
2.A. ¿Tienes este seguro ligado a una hipoteca?	Sí	Recuento % dentro de 2.A. ¿Tienes este seguro ligado a una hipoteca?	297 63,1%	174 36,9%	471 100,0%
	No	Recuento % dentro de 2.A. ¿Tienes este seguro ligado a una hipoteca?	944 76,3%	294 23,7%	1238 100,0%
Total		Recuento % dentro de 2.A. ¿Tienes este seguro ligado a una hipoteca?	1241 72,6%	468 27,4%	1709 100,0%

Fuente: Elaboración propia.

Cobertura del siniestro

Tabla A.9. Tabla de contingencia cobertura del siniestro - recomendación

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
3. ¿CÍA X ha cubierto el siniestro...?	Totalmente	Recuento % dentro de 3. ¿CÍA X ha cubierto el siniestro...?	996 86,2%	160 13,8%	1156 100,0%
	Parcialmente	Recuento % dentro de 3. ¿CÍA X ha cubierto el siniestro...?	227 51,0%	218 49,0%	445 100,0%
	No lo ha cubierto / Ha rechazado el siniestro	Recuento % dentro de 3. ¿CÍA X ha cubierto el siniestro...?	25 28,1%	64 71,9%	89 100,0%
Total		Recuento % dentro de 3. ¿CÍA X ha cubierto el siniestro...?	1248 73,8%	442 26,2%	1690 100,0%

. Fuente: Elaboración propia.

Estado del siniestro

Tabla A.9. Tabla de contingencia estado del siniestro - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
4. ¿En qué estado se encuentra tu siniestro?	Reparados	Recuento	926	223	1149
		% dentro de 4. ¿En qué estado se encuentra tu siniestro?	80,6%	19,4%	100,0%
	Indemnizados	Recuento	297	155	452
		% dentro de 4. ¿En qué estado se encuentra tu siniestro?	65,7%	34,3%	100,0%
Total	Recuento		1223	378	1601
	% dentro de 4. ¿En qué estado se encuentra tu siniestro?		76,4%	23,6%	100,0%

Fuente: Elaboración propia.

Tipología de la prestación

Tabla A.10. Tabla de contingencia tipología del siniestro - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
5. ¿Qué tipo de siniestro es el que ha tenido lugar?	Robo	Recuento	4	1	5
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	80,0%	20,0%	100,0%
	Daños por agua	Recuento	473	149	622
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	76,0%	24,0%	100,0%
	Cristales	Recuento	154	52	206
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	74,8%	25,2%	100,0%
	Robo	Recuento	53	26	79
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	67,1%	32,9%	100,0%
	Subida de tensión / tormenta / daños eléctricos / cortes de	Recuento	100	24	124
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	80,6%	19,4%	100,0%
	Rotura saneamientos	Recuento	143	44	187
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	76,5%	23,5%	100,0%
	Rotura electrodomésticos	Recuento	164	33	197
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	83,2%	16,8%	100,0%
	Incendio	Recuento	22	8	30
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	73,3%	26,7%	100,0%
	Viento	Recuento	35	10	45
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	77,8%	22,2%	100,0%
	Cerrajería	Recuento	69	23	92
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	75,0%	25,0%	100,0%
	Otros (Anotar)	Recuento	68	15	83
		% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?	81,9%	18,1%	100,0%
Total	Recuento		1285	385	1670
	% dentro de 5. ¿Qué tipo de siniestro es el que ha tenido lugar?		,8	,2	1,0

Fuente: Elaboración propia.

Rapidez en la comunicación

Tabla A.11. Tabla de contingencia rapidez en la comunicación – recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
7.1.1. Rapidez en la comunicación	Satisfecho	Recuento	1112	168	1280
		% dentro de 7.1.1. Rapidez en la comunicación	86,9%	13,1%	100,0%
	no satisfecho	Recuento	144	307	451
		% dentro de 7.1.1. Rapidez en la comunicación	31,9%	68,1%	100,0%
Total		Recuento	1256	475	1731
		% dentro de 7.1.1. Rapidez en la comunicación	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Facilidad de contacto con la entidad

Tabla A.12. Tabla de contingencia facilidad de contacto con la entidad - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
7.1.2. Facilidad de contacto con la entidad	Satisfecho	Recuento	1073	164	1237
		% dentro de 7.1.2. Facilidad de contacto con la entidad	86,7%	13,3%	100,0%
	no satisfecho	Recuento	183	311	494
		% dentro de 7.1.2. Facilidad de contacto con la entidad	37,0%	63,0%	100,0%
Total		Recuento	1256	475	1731
		% dentro de 7.1.2. Facilidad de contacto con la entidad	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Información de los siguientes pasos a seguir en la tramitación

Tabla A.13. Tabla de contingencia información de los siguientes pasos a seguir en la tramitación - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
7.1.3. Información de los siguientes pasos a seguir en la tramitación	Satisfecho	Recuento	1118	148	1266
		% dentro de 7.1.3. Información de los siguientes pasos a seguir en la tramitación	88,3%	11,7%	100,0%
	no satisfecho	Recuento	138	327	465
		% dentro de 7.1.3. Información de los siguientes pasos a seguir en la tramitación	29,7%	70,3%	100,0%
Total		Recuento	1256	475	1731
		% dentro de 7.1.3. Información de los siguientes pasos a seguir en la tramitación	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Interés por parte de la compañía en solucionar la prestación

Tabla A.14. Tabla de contingencia interés por parte de la compañía en solucionar la prestación - recomendación.

			RECOMENDACION		
			Total con intención (7-10)		Sin intención (1-6)
7.1.4. Interés por parte de la compañía en solucionar la prestación	Satisfecho	Recuento	1094	153	1247
		% dentro de 7.1.4. Interés por parte de la compañía en solucionar la prestación	87,7%	12,3%	100,0%
	no satisfecho	Recuento	162	322	484
		% dentro de 7.1.4. Interés por parte de la compañía en solucionar la prestación	33,5%	66,5%	100,0%
Total	Recuento	1256	475	1731	
	% dentro de 7.1.4. Interés por parte de la compañía en solucionar la prestación	72,6%	27,4%	100,0%	

Fuente: Elaboración propia.

Sencillez del lenguaje utilizado por el personal de la compañía

Tabla A.15. Tabla de contingencia sencillez del lenguaje utilizado por el personal de la compañía- recomendación.

			RECOMENDACION		
			Total con intención (7-10)		Sin intención (1-6)
7.1.5. Sencillez del lenguaje utilizado por el personal de la compañía	Satisfecho	Recuento	1084	174	1258
		% dentro de 7.1.5. Sencillez del lenguaje utilizado por el personal de la compañía	86,2%	13,8%	100,0%
	no satisfecho	Recuento	172	301	473
		% dentro de 7.1.5. Sencillez del lenguaje utilizado por el personal de la compañía	36,4%	63,6%	100,0%
Total	Recuento	1256	475	1731	
	% dentro de 7.1.5. Sencillez del lenguaje utilizado por el personal de la compañía	72,6%	27,4%	100,0%	

Fuente: Elaboración propia.

Valoración del alcance del siniestro o de los daños producidos

Tabla A.16. Tabla de contingencia valoración del alcance del siniestro - recomendación.

			RECOMENDACION		
			Total con intención (7-10)		Sin intención (1-6)
8. ¿La entidad realizó alguna valoración del alcance del siniestro o de los daños producidos?	Sí, envió un perito	Recuento	706	243	949
		% dentro de 8. ¿La entidad realizó alguna valoración del alcance del siniestro o de los daños producidos?	74,4%	25,6%	100,0%
	Sí, hubo una teleperitación o fotoperitación	Recuento	95	37	132
		% dentro de 8. ¿La entidad realizó alguna valoración del alcance del siniestro o de los daños producidos?	72,0%	28,0%	100,0%
	No hubo peritación	Recuento	353	76	429
		% dentro de 8. ¿La entidad realizó alguna valoración del alcance del siniestro o de los daños producidos?	82,3%	17,7%	100,0%
	No recuerdo	Recuento	69	22	91
		% dentro de 8. ¿La entidad realizó alguna valoración del alcance del siniestro o de los daños producidos?	75,8%	24,2%	100,0%
	Total	Recuento	1223	378	1601
		% dentro de 8. ¿La entidad realizó alguna valoración del alcance del siniestro o de los daños producidos?	76,4%	23,6%	100,0%

Fuente: Elaboración propia.

Canal de comunicación de la prestación

Tabla A.17. Tabla de contingencia canal de comunicación - recomendación.

			RECOMENDACION		
			Total con intención		Total
			(7-10)	Sin intención (1-6)	
6. ¿A través de qué canal comunicaste la prestación a CÍA X?	En la oficina de la compañía	Recuento	100	30	130
		% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	76,9%	23,1%	100,0%
	En la sucursal del banco / caja de ahorros	Recuento	18	9	27
		% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	66,7%	33,3%	100,0%
	Me puse en contacto con el agente	Recuento	207	47	254
		% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	81,5%	18,5%	100,0%
	Me puse en contacto con el corredor	Recuento	137	36	173
		% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	79,2%	20,8%	100,0%
	Llamé al teléfono de siniestros de la compañía, del tipo 900 o similar	Recuento	695	220	915
		% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	76,0%	24,0%	100,0%
A través de internet	Recuento	47	31	78	
	% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	60,3%	39,7%	100,0%	
A través de la App de la compañía	Recuento	14	4	18	
	% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	77,8%	22,2%	100,0%	
Otros (Anotar)	Recuento	4	1	5	
	% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	80,0%	20,0%	100,0%	
Total	Recuento	1222	378	1600	
	% dentro de 6. ¿A través de qué canal comunicaste la prestación a CÍA X?	76,4%	23,6%	100,0%	

Fuente: Elaboración propia.

Rapidez en la peritación

Tabla A.18. Tabla de contingencia rapidez en la peritación- recomendación.

			RECOMENDACION		
			Total con intención		Total
			(7-10)	Sin intención (1-6)	
9.1.1. Rapidez en la peritación	Satisfecho	Recuento	706	115	821
		% dentro de 9.1.1. Rapidez en la peritación	86,0%	14,0%	100,0%
	no satisfecho	Recuento	550	360	910
		% dentro de 9.1.1. Rapidez en la peritación	60,4%	39,6%	100,0%
Total	Recuento	1256	475	1731	
	% dentro de 9.1.1. Rapidez en la peritación	72,6%	27,4%	100,0%	

Fuente: Elaboración propia.

Rapidez en la reparación

Tabla A.19. Tabla de contingencia rapidez en la reparación - recomendación

			RECOMENDACION		
			Total con		Total
			intención (7-10)	Sin intención (1-6)	
9.1.2. Rapidez en la reparación	Satisfecho	Recuento	900	113	1013
		% dentro de 9.1.2. Rapidez en la reparación	88,8%	11,2%	100,0%
	no satisfecho	Recuento	356	362	718
		% dentro de 9.1.2. Rapidez en la reparación	49,6%	50,4%	100,0%
Total		Recuento	1256	475	1731
		% dentro de 9.1.2. Rapidez en la reparación	72,6%	27,4%	100,0%

. Fuente: Elaboración propia.

Rapidez en la tramitación completa

Tabla A.20. Tabla de contingencia rapidez en la tramitación completa - recomendación.

			RECOMENDACION		
			Total con		Total
			intención (7-10)	Sin intención (1-6)	
9.1.3. Rapidez en la tramitación completa	Satisfecho	Recuento	1054	128	1182
		% dentro de 9.1.3. Rapidez en la tramitación completa	89,2%	10,8%	100,0%
	no satisfecho	Recuento	202	347	549
		% dentro de 9.1.3. Rapidez en la tramitación completa	36,8%	63,2%	100,0%
Total		Recuento	1256	475	1731
		% dentro de 9.1.3. Rapidez en la tramitación completa	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Calidad de la reparación

Tabla A.21. Tabla de contingencia calidad de la reparación - recomendación.

			RECOMENDACION		
			Total con		Total
			intención (7-10)	Sin intención (1-6)	
10.1.1. Calidad de la reparación	Satisfecho	Recuento	942	118	1060
		% dentro de 10.1.1. Calidad de la reparación	88,9%	11,1%	100,0%
	no satisfecho	Recuento	314	357	671
		% dentro de 10.1.1. Calidad de la reparación	46,8%	53,2%	100,0%
Total		Recuento	1256	475	1731
		% dentro de 10.1.1. Calidad de la reparación	72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Trato personalizado

Tabla A.22. Tabla de contingencia calidad de la reparación - recomendación.

			RECOMENDACION		
			Total con		Total
			intención (7-10)	Sin intención (1-6)	
10.1.2. Trato personalizado	Satisfecho	Recuento	1085	147	1232
		% dentro de 10.1.2. Trato personalizado	88,1%	11,9%	100,0%
	no satisfecho	Recuento	171	328	499
		% dentro de 10.1.2. Trato personalizado	34,3%	65,7%	100,0%
Total	Recuento		1256	475	1731
	% dentro de 10.1.2. Trato personalizado		72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Iniciativa por parte de la compañía

Tabla A.23. Tabla de contingencia iniciativa por parte de la compañía - recomendación.

			RECOMENDACION		
			Total con		Total
			intención (7-10)	Sin intención (1-6)	
10.1.3. Iniciativa por parte de la compañía	Satisfecho	Recuento	1036	120	1156
		% dentro de 10.1.3. Iniciativa por parte de la compañía	89,6%	10,4%	100,0%
	no satisfecho	Recuento	220	355	575
		% dentro de 10.1.3. Iniciativa por parte de la compañía	38,3%	61,7%	100,0%
Total	Recuento		1256	475	1731
	% dentro de 10.1.3. Iniciativa por parte de la compañía		72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Adaptación del seguro a tus necesidades

Tabla A.24. Tabla de contingencia adaptación del seguro a tus necesidades - recomendación.

			RECOMENDACION		
			Total con		Total
			intención (7-10)	Sin intención (1-6)	
11.1.1. Adaptación del seguro a tus necesidades	Satisfecho	Recuento	1085	113	1198
		% dentro de 11.1.1. Adaptación del seguro a tus necesidades	90,6%	9,4%	100,0%
	no satisfecho	Recuento	171	362	533
		% dentro de 11.1.1. Adaptación del seguro a tus necesidades	32,1%	67,9%	100,0%
Total	Recuento		1256	475	1731
	% dentro de 11.1.1. Adaptación del seguro a tus necesidades		72,6%	27,4%	100,0%

Fuente: Elaboración propia.

Adecuada relación entre precio de la póliza y coberturas incluidas

Tabla A.25. Tabla de contingencia adecuada relación entre precio de la póliza y coberturas incluidas - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
11.1.2. Adecuada relación entre precio de la póliza y coberturas incluidas	Satisfecho	Recuento	1056	99	1155
		% dentro de 11.1.2. Adecuada relación entre precio de la póliza y coberturas incluidas	91,4%	8,6%	100,0%
	no satisfecho	Recuento	200	376	576
		% dentro de 11.1.2. Adecuada relación entre precio de la póliza y coberturas incluidas	34,7%	65,3%	100,0%
Total	Recuento	1256	475	1731	
	% dentro de 11.1.2. Adecuada relación entre precio de la póliza y coberturas incluidas	72,6%	27,4%	100,0%	

Fuente: Elaboración propia.

Esfuerzo realizado para la gestión de la prestación del seguro

Tabla A.26. Tabla de contingencia esfuerzo realizado para la gestión de la prestación del seguro - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
12. ¿Cómo valorarías el esfuerzo realizado para la gestión de la prestación del seguro?	Satisfecho	Recuento	871	155	1026
		% dentro de 12. ¿Cómo valorarías el esfuerzo realizado para la gestión de la prestación del	84,9%	15,1%	100,0%
	no satisfecho	Recuento	385	320	705
		% dentro de 12. ¿Cómo valorarías el esfuerzo realizado para la gestión de la prestación del	54,6%	45,4%	100,0%
Total	Recuento	1256	475	1731	
	% dentro de 12. ¿Cómo valorarías el esfuerzo realizado para la gestión de la prestación del	72,6%	27,4%	100,0%	

Fuente: Elaboración propia.

Expectativas de servicio

Tabla A.27. Tabla de contingencia expectativas de servicio - recomendación.

			RECOMENDACION		
			Total con intención (7-10)	Sin intención (1-6)	Total
13. Los servicios ofrecidos hasta ahora por CÍA X han resultado...	Mejor de lo que esperaba	Recuento	457	36	493
		% dentro de 13. Los servicios ofrecidos hasta ahora por CÍA X han resultado...	92,7%	7,3%	100,0%
	Igual (es lo que esperaba)	Recuento	715	206	921
		% dentro de 13. Los servicios ofrecidos hasta ahora por CÍA X han resultado...	77,6%	22,4%	100,0%
	Peor de lo que esperaba	Recuento	51	136	187
		% dentro de 13. Los servicios ofrecidos hasta ahora por CÍA X han resultado...	27,3%	72,7%	100,0%
Total	Recuento	1223	378	1601	
	% dentro de 13. Los servicios ofrecidos hasta ahora por CÍA X han resultado...	76,4%	23,6%	100,0%	

Fuente: Elaboración propia.

Anexo II. GLM

Modelo Global

Tabla A.28. Resumen de los pasos.

Paso	Mejora			Modelo			% de clas. correcta
	Chi cuadrado	gl	Sig.	Chi cuadrado	gl	Sig.	
1	669,856	1	,000	669,856	1	,000	91,8%
2	63,986	1	,000	733,841	2	,000	92,0%
3	17,096	1	,000	750,937	3	,000	91,8%
4	10,608	1	,001	761,545	4	,000	92,3%
5	8,398	1	,004	769,943	5	,000	92,3%
6	6,355	1	,012	776,298	6	,000	92,4%
7	6,083	1	,014	782,382	7	,000	92,1%
8	4,054	1	,044	786,436	8	,000	92,2%

a. No se pueden eliminar ni añadir más variables al modelo actual. b. Bloque final: 1

Fuente: Elaboración propia.

Tabla A.29. Prueba de Hosmer y Lemeshow.

Paso	Chi cuadrado	gl	Sig.
8	1,923	7	,964

Fuente: Elaboración propia.

Tabla A.30. Tabla de contingencia para la prueba de Hosmer y Lemeshow .

	RECOMENDACION = Total con intención (7-10)		RECOMENDACION = Sin intención (1-6)		Total
	Observado	Esperado	Observado	Esperado	
	Paso 8				
1	156	154,936	1	2,064	157
2	114	113,558	2	2,442	116
3	41	41,761	2	1,239	43
4	239	240,972	10	8,028	249
5	126	124,854	6	7,146	132
6	115	114,636	9	9,364	124
7	80	79,330	39	39,670	119
8	20	21,124	94	92,876	114
9	9	8,829	119	119,171	128

Fuente: Elaboración propia.

Tabla A.31. Variables que no están en la ecuación.

		Puntuación	gl	Sig.	
	Número de seguros	,270	1	,603	
	Prestación por Agua	,466	1	,495	
	Prestación por Cristales	,982	1	,322	
	Prestación por Robo	1,196	1	,274	
	Prestación por Subida de tensión / viento	1,210	1	,271	
	Prestación por Rotura de saneamientos / electrodomésticos	,772	1	,380	
	Prestación por Cerrajería	,215	1	,643	
	Expectativas mejor de lo esperado	1,928	1	,165	
	Expectativas peor de lo esperado	,048	1	,827	
	Mujer	,512	1	,474	
	Hombre	,512	1	,474	
	habitat_Urbano	,140	1	,708	
	habitat_Rural	,140	1	,708	
	nielsen_NORTE	,468	1	,494	
	nielsen_SUR	,596	1	,440	
	nielsen_CENTRO	,261	1	,609	
	nielsen_GALAICA	,003	1	,959	
	nielsen_LEVANTE	1,090	1	,296	
Paso 8	Variables	nielsen_CATALANA	,017	1	,895
	Repadros	1,668	1	,197	
	Indemnizados	1,668	1	,197	
	Cobertura total del siniestro	1,158	1	,282	
	Rapidez en la comunicación	2,355	1	,125	
	Facilidad de contacto con la entidad	1,465	1	,226	
	Interés por parte de la compañía en solucionar la prestación	,404	1	,525	
	Sencillez del lenguaje utilizado por el personal de la compañía	,094	1	,759	
	Facilidad de contacto con la entidad / Es de fácil uso, con accesos amigables y cómodos	,361	1	,548	
	Sencillez del lenguaje utilizado por el personal de la compañía / Los formularios son sencillos y claros	,010	1	,920	
	Rapidez en la peritación	,481	1	,488	
	Rapidez en la reparación	,002	1	,967	
	Rapidez en la tramitación completa	1,164	1	,281	
	Trato personalizado	,997	1	,318	
	Iniciativa por parte de la compañía	,125	1	,723	
	Adaptación del seguro a tus necesidades	,794	1	,373	
	Antigüedad	3,297	1	,069	
	Edad	1,039	1	,308	

a. No se calculan los chi-cuadrado residuales a causa de las redundancias.

Fuente: Elaboración propia.

Modelo proceso reparatorio

Tabla A.32. Historia de iteraciones.

Iteración		-2 log de la verosimilitud	Coeficientes						
			Constant	Ligado a hipoteca	Trato personalizado	Esfuerzo realizado para la gestión	Intención de continuar	Satisfacción general	Perito
Paso 6	1	434,109	-4,483	-,199	,353	-,185	1,335	1,667	-,159
	2	359,602	-5,789	-,444	,710	-,451	1,911	2,173	-,390
	3	346,757	-6,176	-,664	,965	-,729	2,235	2,464	-,633
	4	345,860	-6,238	-,750	1,048	-,853	2,352	2,576	-,738
	5	345,853	-6,241	-,759	1,056	-,867	2,364	2,588	-,749
	6	345,853	-6,241	-,759	1,056	-,868	2,364	2,588	-,750

a. Método: Por pasos hacia adelante (Razón de verosimilitud)

b. En el modelo se incluye una constante.

c. -2 log de la verosimilitud inicial: 795,299

d. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: Elaboración propia.

Tabla A.33. Pruebas omnibus sobre los coeficientes del modelo.

		Chi cuadrado	gl	Sig.
Paso 6	Paso	4,854	1	,028
	Bloque	449,446	6	,000
	Modelo	449,446	6	,000

Fuente: Elaboración propia.

Tabla A.34. Resumen del modelo.

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
6	345,853 ^a	,427	,681

a. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: Elaboración propia.

Tabla A.35. Prueba de Hosmer y Lemeshow.

Paso	Chi cuadrado	gl	Sig.
6	3,821	6	,701

Fuente: Elaboración propia.

Tabla A.36. Tabla de contingencia para la prueba de Hosmer y Lemeshow.

		RECOMENDACION = Total con intención (7-10)		RECOMENDACION = Sin intención (1-6)		Total
		Observado	Esperado	Observado	Esperado	
		Paso 6	1	86	85,927	
	2	26	26,314	1	,686	27
	3	324	322,675	8	9,325	332
	4	26	26,029	1	,971	27
	5	109	107,513	5	6,487	114
	6	56	61,664	28	22,336	84
	7	19	15,324	64	67,676	83
	8	4	4,556	49	48,444	53

Fuente: Elaboración propia.

Tabla A.37. Tabla de clasificación

Observado	RECOMENDACION	Total con intención (7-10)	Pronosticado		
			RECOMENDACION		Porcentaje correcto
			Total con intención (7-10)	Sin intención (1-6)	
Paso 6					
		Total con intención (7-10)	623	27	95,8
		Sin intención (1-6)	35	122	77,7
		Porcentaje global			92,3

a. El valor de corte es ,500

Fuente: Elaboración propia.

Tabla A.38. Variables en la ecuación.

	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
							Inferior	Superior
Paso 6 ^a Ligado a hipoteca	-,759	,313	5,881	1	,015	,468	,253	,864
Trato personalizado	1,056	,362	8,511	1	,004	2,874	1,414	5,843
Esfuerzo realizado para la gestión de la prestación	-,868	,387	5,014	1	,025	,420	,197	,897
Intención de continuar	2,364	,442	28,553	1	,000	10,634	4,468	25,308
Opinion general	2,588	,430	36,149	1	,000	13,302	5,722	30,926
Perito	-,750	,348	4,643	1	,031	,473	,239	,934
Constante	-6,241	,830	56,498	1	,000	,002		

a. Variable(s) introducida(s) en el paso 6: Perito.

Fuente: Elaboración propia.

Modelo proceso indemnizatorio

Tabla A.41. Historia de iteraciones.

Iteración		-2 log de la verosimilitud	Coeficientes			
			Constant	P713_PRES_D M	P1112_PRES_D M	p5s_DM
Paso 3	1	190,749	-5,696	,660	1,002	2,124
	2	161,836	-8,357	1,110	1,591	2,839
	3	157,866	-9,785	1,347	1,904	3,220
	4	157,701	-10,154	1,403	1,984	3,321
	5	157,700	-10,175	1,406	1,989	3,327
	6	157,700	-10,175	1,406	1,989	3,327

a. Método: Por pasos hacia adelante (Razón de verosimilitud)

b. En el modelo se incluye una constante.

c. -2 log de la verosimilitud inicial: 437,980

d. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: Elaboración propia.

Tabla A.42. Pruebas omnibus sobre los coeficientes del modelo.

		Chi cuadrado	gl	Sig.
Paso 3	Paso	8,542	1	,003
	Bloque	280,280	3	,000
	Modelo	280,280	3	,000

Fuente: Elaboración propia.

Tabla A.43. Resumen del modelo.

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
3	157,700 ^a	,566	,777

a. La estimación ha finalizado en el número de iteración 6 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: Elaboración propia.

Tabla A.44. Prueba de Hosmer y Lemeshow.

Paso	Chi cuadrado	gl	Sig.
3	,012	2	,994

Fuente: Elaboración propia.

Tabla A.45. Tabla de contingencia para la prueba de Hosmer y Lemeshow.

		RECOMENDACION = Total con intención (7-10)		RECOMENDACION = Sin intención (1-6)		Total
		Observado	Esperado	Observado	Esperado	
		Paso 3				
	1	180	179,859	13	13,141	193
	2	10	10,141	6	5,859	16
	3	17	17,141	20	19,859	37
	4	9	8,859	81	81,141	90

Fuente: Elaboración propia.

Tabla A.46. Tabla de clasificación.

Observado			Pronosticado			
			RECOMENDACION			Porcentaje correcto
			Total con intención (7-10)	Sin intención (1-6)		
Paso 3	RECOMENDACION	Total con intención (7-10)	209	7	96,8	
		Sin intención (1-6)	23	97	80,8	
Porcentaje global					91,1	

a. El valor de corte es ,500

Fuente: Elaboración propia.

Tabla A.47. Variables en la ecuación.

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 3 ^a	Información de los siguientes pasos a seguir en la tramitación	1,406	,477	8,688	1	,003	4,081	1,602	10,395
	Adecuada relación entre precio de la póliza y coberturas incluidas	1,989	,459	18,748	1	,000	7,307	2,970	17,978
	Intención de continuar	3,327	,429	60,243	1	,000	27,854	12,023	64,529
	Constante	-10,175	,979	108,053	1	,000	,000		

a. Variable(s) introducida(s) en el paso 3: Información de los siguientes pasos a seguir.

Fuente: Elaboración propia.

Tabla A.48. Matriz de correlaciones.

		Constant	Información de los siguientes pasos a seguir en la tramitación	Adecuada relación entre precio de la póliza y coberturas incluidas	Intención de continuar
Paso 3	Constant	1,000	-,377	-,451	-,605
	Información de los siguientes pasos a seguir en la tramitación	-,377	1,000	-,372	-,044
	Adecuada relación entre precio de la póliza y coberturas incluidas	-,451	-,372	1,000	-,003
	Intención de continuar	-,605	-,044	-,003	1,000

Fuente: Elaboración propia.

Tabla A.49. Resumen de los pasos.

Paso	Mejora			Modelo			% de clas. correcta
	Chi cuadrado	gl	Sig.	Chi cuadrado	gl	Sig.	
1	233,600	1	,000	233,600	1	,000	90,8%
2	38,138	1	,000	271,738	2	,000	90,8%
3	8,542	1	,003	280,280	3	,000	91,1%

a. No se pueden eliminar ni añadir más variables al modelo actual. b. Bloque final: 1

Fuente: Elaboración propia.

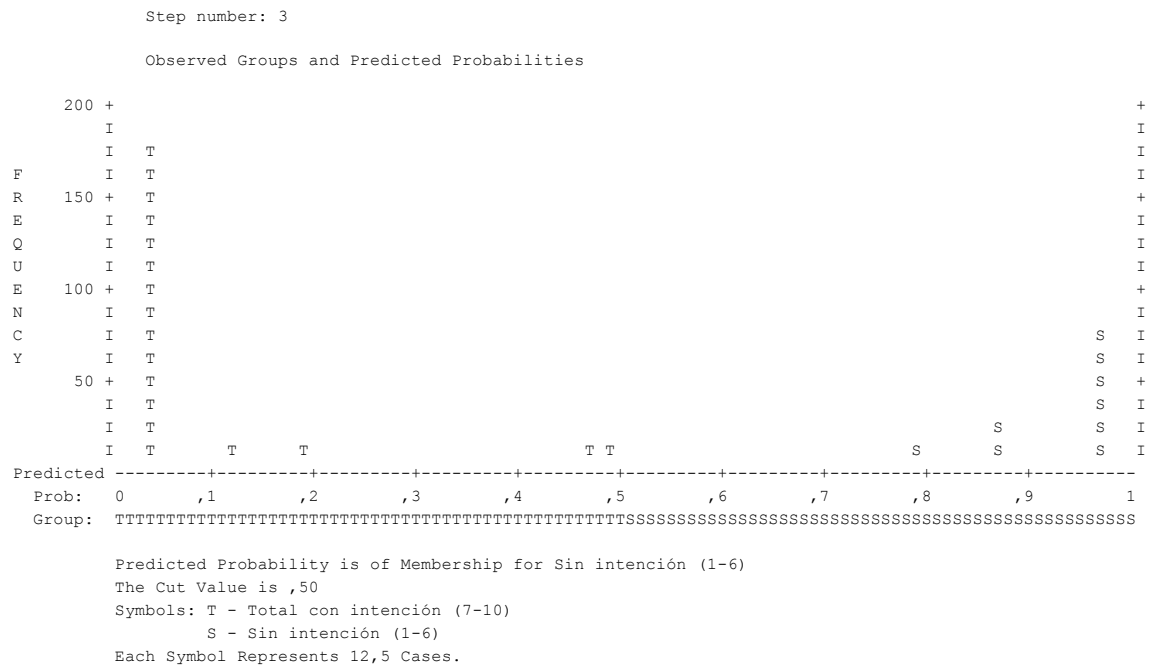


Figura A.2. Gráfico de clasificación. Fuente: Elaboración propia.

Modelo de alerta temprana

Tabla A.50. Historia de iteraciones.

Iteración		Coeficientes						
		-2 log de la verosimilitud	Constant	Ligado a hipoteca	Rapidez en la comunicación	Información de los siguientes pasos a seguir en la tramitación	Interés por parte de la compañía en solucionar la prestación	Sencillez del lenguaje utilizado por el personal
Paso 5	1	1075,834	-3,739	-,426	,575	1,083	,687	,470
	2	1033,279	-4,517	-,730	,727	1,304	,908	,645
	3	1031,530	-4,649	-,833	,756	1,358	,963	,685
	4	1031,524	-4,655	-,840	,758	1,361	,966	,687
	5	1031,524	-4,655	-,840	,758	1,361	,966	,687

a. Método: Por pasos hacia adelante (Razón de verosimilitud)

b. En el modelo se incluye una constante.

c. -2 log de la verosimilitud inicial: 1557,117

d. La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Fuente: Elaboración propia.

Tabla A.51. Pruebas omnibus sobre los coeficientes del modelo.

		Chi cuadrado	gl	Sig.
Paso 5	Paso	10,167	1	,001
	Bloque	525,593	5	,000
	Modelo	525,593	5	,000

Fuente: Elaboración propia.

Tabla A.52. Resumen del modelo.

Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
5	1031,524	,330	,475

Fuente: Elaboración propia.

Tabla A.53. Prueba de Hosmer y Lemeshow.

Paso	Chi cuadrado	gl	Sig.
5	2,451	4	,653

Fuente: Elaboración propia.

Tabla A.54. Matriz de correlaciones.

		Constant	Ligado a hipoteca	Rapidez en la comunicación	Información de los siguientes pasos a seguir en la tramitación	Interés por parte de la compañía en solucionar la prestación	Sencillez del lenguaje utilizado por el personal
Paso 5	Constant	1,000	-,639	-,133	-,098	-,126	-,195
	Ligado a hipoteca	-,639	1,000	-,033	-,055	-,033	-,034
	Rapidez en la comunicación	-,133	-,033	1,000	-,283	-,360	-,159
	Información de los siguientes pasos a seguir en la tramitación	-,098	-,055	-,283	1,000	-,315	-,283
	Interés por parte de la compañía en solucionar la prestación	-,126	-,033	-,360	-,315	1,000	-,195
	Sencillez del lenguaje utilizado por el personal	-,195	-,034	-,159	-,283	-,195	1,000

Fuente: Elaboración propia.

Tabla A.55. Variables que no están en la ecuación.

			Puntuación	gl	Sig.
Paso 5	Variables	Número de seguros	,106	1	,745
		p5_Agua_DM	,140	1	,708
		p5_Cristales_DM	2,615	1	,106
		p5_Robo_DM	2,532	1	,112
		p5_Fenomenos_Atmos_DM	2,536	1	,111
		p5_Rotura_DM	,022	1	,882
		p5_Cerrajería_DM	1,362	1	,243
		Mujeres_DM	,364	1	,546
		Hombres_DM	,364	1	,546
		habitat_Urbano_DM	,062	1	,803
		habitat_Rural_DM	,062	1	,803
		nielsen_NORTE_DM	1,839	1	,175
		nielsen_SUR_DM	,776	1	,378
		nielsen_CENTRO_DM	,003	1	,956
		nielsen_GALAICA_DM	,009	1	,926
		nielsen_LEVANTE_DM	,098	1	,755
		nielsen_CATALANA_DM	,000	1	1,000
		Facilidad de contacto con la entidad	2,633	1	,105
		Antigüedad	1,027	1	,311
		EDAD	1,136	1	,286

Fuente: Elaboración propia.

Tabla A.56. Resumen de los pasos

Paso	Mejora			Modelo			% de clas. correcta	Variable
	Chi cuadrado	gl	Sig.	Chi cuadrado	gl	Sig.		
1	426,625	1	,000	426,625	1	,000	83,3%	IN: Información de los siguientes pasos a seguir
2	50,981	1	,000	477,606	2	,000	83,9%	IN: Interés por parte de la compañí
3	23,087	1	,000	500,693	3	,000	84,1%	IN: Ligado a hipoteca
4	14,732	1	,000	515,425	4	,000	84,8%	IN: Rapidez en la comunicación
5	10,167	1	,001	525,593	5	,000	85,1%	IN: Sencillez del lenguaje utilizado

a. No se pueden eliminar ni añadir más variables al modelo actual. b. Bloque final: 1

. Fuente: Elaboración propia.

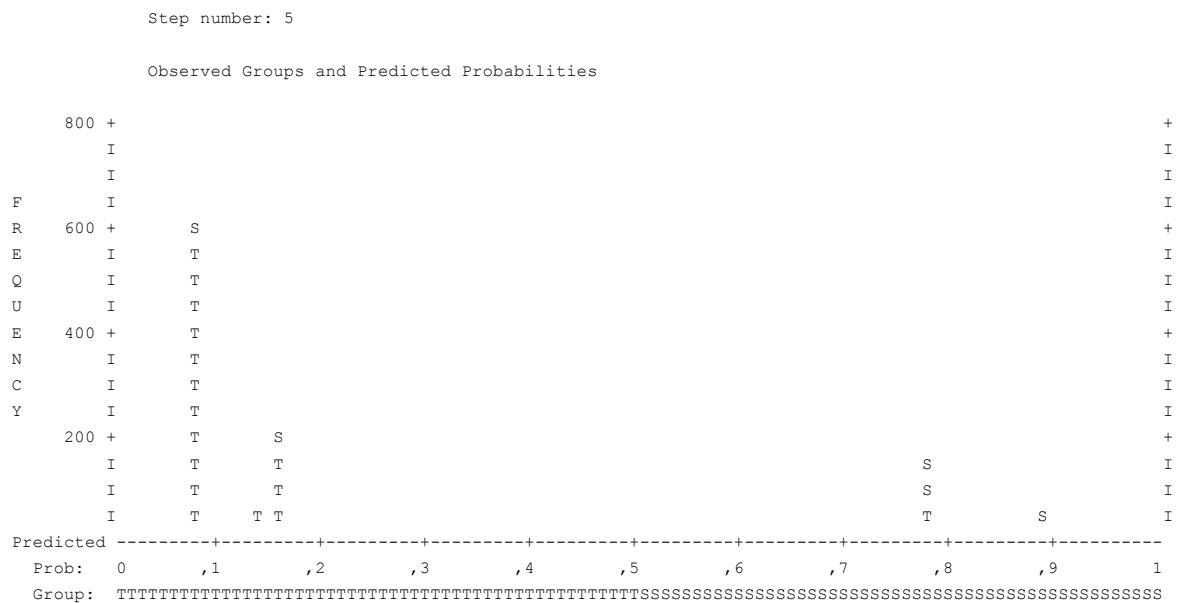


Figura A.3. Gráfico de clasificación. Fuente: Elaboración propia.

Anexo III. Código SPSS

➤ Muestra aleatoria estratificada

```
CSPLAN SAMPLE
/PLAN FILE='D:\master\TFM\plan22.csplan'
/PLANVARS SAMPLEWEIGHT=SampleWeight_Final_
/PRINT PLAN MATRIX
/DESIGN STRATA=SEXO Edad_tramos nielsen
/METHOD TYPE=SIMPLE_WOR ESTIMATION=DEFAULT
/SIZE MATRIX=SEXO Edad_tramos nielsen;1 1.00 1.00 2;
1 1.00 2.00 13;1 1.00 3.00 10;1 1.00
4.00 3;1 1.00 5.00 8;1 1.00 6.00 4;1 1.00 7.00 1;
1 2.00 1.00 7;1 2.00 2.00 17;1 2.00 3.00 20;1 2.00
4.00 7;1 2.00 5.00 7;1 2.00 6.00 9;1 2.00 7.00 3;
1 3.00 1.00 16;1 3.00 2.00 28;1 3.00 3.00 37;1
3.00 4.00 9;1 3.00 5.00 22;1 3.00 6.00 26;
1 3.00 7.00 5;1 4.00 1.00 20;1 4.00 2.00 48;1 4.00 3.00
64;1 4.00 4.00 20;1 4.00 5.00 40;1 4.00 6.00 40;
1 4.00 7.00 4;1 5.00 1.00 21;1 5.00 2.00 41;1 5.00
3.00 76;1 5.00 4.00 23;1 5.00 5.00 36;1 5.00 6.00 47;
1 5.00 7.00 3;1 6.00 1.00 8;1 6.00 2.00 17;1
6.00 3.00 40;1 6.00 4.00 9;1 6.00 5.00 26;1 6.00 6.00 31;
1 6.00 7.00 2;2 1.00 1.00 8;2 1.00 2.00
30;2 1.00 3.00 33;2 1.00 4.00 11;2 1.00 5.00 17;
2 1.00 6.00 25;2 1.00 7.00 5;2 2.00 1.00 16;2 2.00
2.00 28;2 2.00 3.00 43;2 2.00 4.00 11;2 2.00 5.00 18;
2 2.00 6.00 22;2 2.00 7.00 3;2 3.00 1.00 25;2
3.00 2.00 38;2 3.00 3.00 62;2 3.00 4.00 23;2 3.00 5.00 26;
2 3.00 6.00 33;2 3.00 7.00 8;2 4.00 1.00
26;2 4.00 2.00 28;2 4.00 3.00 60;2 4.00 4.00 20;
2 4.00 5.00 20;2 4.00 6.00 42;2 4.00 7.00 5;2 5.00
1.00 16;2 5.00 2.00 20;2 5.00 3.00 60;2 5.00 4.00 18;
2 5.00 5.00 26;2 5.00 6.00 48;2 5.00 7.00 4;2
6.00 1.00 4;2 6.00 2.00 12;2 6.00 3.00 16;2 6.00 4.00 6;
2 6.00 5.00 8;2 6.00 6.00 9;2 6.00 7.00 1
/STAGEVARS INCLPROB(InclusionProbability_1_)
CUMWEIGHT(SampleWeightCumulative_1_).
CSSELECT
/PLAN FILE='D:\master\TFM\plan22.csplan'
/CRITERIA STAGES=1 SEED=RANDOM
/CLASSMISSING EXCLUDE
/PRINT SELECTION.
```

➤ Recodificación variables

- Edad

```
RECODE
Edad
  (Lowest thru 25=1) (26 thru 35=2) (36 thru 45=3) (46
thru 55=4) (56 thru 65=5) (66 thru Highest=6) INTO
Edad_tramos.
EXECUTE .
value labels Edad_tramos
1 '25 o menos'
2 'De 26 a 35 años'
3 'De 36 a 45 años'
4 'De 46 a 55 años'
5 'De 56 a 65 años'
6 ' Más de 65'.
EXECUTE.
```

- Provincia

```
if provincia='A CORU?A' nielsen=4.
if provincia='A CORUÑA' nielsen=4.
if provincia='ÁLAVA' nielsen=1.
if provincia='ALAVA' nielsen=1.
if provincia='ALBACETE' nielsen=5.
if provincia='ALICANTE' nielsen=5.
if provincia='ALMERÍA' nielsen=2.
if provincia='ALMERIA' nielsen=2.
if provincia='ASTURIAS' nielsen=4.
if provincia='ÁVALA' nielsen=3.
if provincia='ÁVILA' nielsen=3.
if provincia='AVILA' nielsen=3.
if provincia='BADAJOZ' nielsen=2.
if provincia='BALEARES' nielsen=6.
if provincia='BALEARs' nielsen=6.
if provincia='ILLES BALEARS' nielsen=6.
if provincia='BARCELONA' nielsen=6.
if provincia='BIZKAIA' nielsen=1.
if provincia='BURGOS' nielsen=1.
if provincia='CÁCERES' nielsen=3.
if provincia='CACERES' nielsen=3.
if provincia='CÁDIZ' nielsen=2.
if provincia='CADIZ' nielsen=2.
if provincia='CANTABRIA' nielsen=1.
if provincia='CASTELLÓN' nielsen=5.
if provincia='CASTELLON' nielsen=5.
if provincia='CASTELLON DE LA PLANA' nielsen=5.
if provincia='CEUTA' nielsen=2.
if provincia='CIUDAD REAL' nielsen=3.
if provincia='CÓRDOBA' nielsen=2.
if provincia='CORDOBA' nielsen=2.
if provincia='CUENCA' nielsen=3.
if provincia='GUIPUZCOA' nielsen=1.
if provincia='GUIPÚZCOA' nielsen=1.
if provincia='GIPUZKOA' nielsen=1.
if provincia='GIPÚZCOA' nielsen=1.
if provincia='GIRONA' nielsen=6.
if provincia='GERONA' nielsen=6.
```



```

if provincia='GRANADA' nielsen=2.
if provincia='GUADALAJARA' nielsen=3.
if provincia='HUELVA' nielsen=2.
if provincia='HUESCA' nielsen=6.
if provincia='JAÉN' nielsen=2.
if provincia='JAEN' nielsen=2.
if provincia='LA CORUÑA' nielsen=4.
if provincia='LA RIOJA' nielsen=1.
if provincia='LAS PALMAS' nielsen=7.
if provincia='LAS PALMAS DE GRAN CANARIA' nielsen=7.
if provincia='LEÓN' nielsen=4.
if provincia='LEON' nielsen=4.
if provincia='LLEIDA' nielsen=6.
if provincia='LUGO' nielsen=4.
if provincia='MADRID' nielsen=3.
if provincia='MALAGA' nielsen=2.
if provincia='MÁLAGA' nielsen=2.
if provincia='MELILLA' nielsen=2.
if provincia='MURCIA' nielsen=5.
if provincia='NAVARRA' nielsen=1.
if provincia='ORENSE' nielsen=4.
if provincia='OURENSE' nielsen=4.
if provincia='PALENCIA' nielsen=1.
if provincia='PONTEVEDRA' nielsen=4.
if provincia='SALAMANCA' nielsen=3.
if provincia='SANTA CRUZ DE TENERIFE' nielsen=7.
if provincia='SEGOVIA' nielsen=3.
if provincia='SEVILLA' nielsen=2.
if provincia='SORIA' nielsen=3.
if provincia='TARRAGONA' nielsen=6.
if provincia='TENERIFE' nielsen=7.
if provincia='TERUEL' nielsen=3.
if provincia='TOLEDO' nielsen=3.
if provincia='VALENCIA' nielsen=5.
if provincia='VALÈNCIA' nielsen=5.
if provincia='VALLADOLID' nielsen=3.
if provincia='VIZCAYA' nielsen=1.
if provincia='ZAMORA' nielsen=3.
if provincia='ZARAGOZA' nielsen=6.
EXECUTE.

```

```

value label nielsen
1 'Norte'
2 'Sur'
3 'Centro'
4 'Galaica'
5 'Levante'
6 'Catalana'
7 'Canarias'.

```

- Recomendación, satisfacción general e intención de continuar

```

Recode
p4
p5
p6
(1 thru 6 = 1)
(7 thru 10 = 2)
(missing=sysmis)

```

```

    (else =copy)
    into
    p4s
    p5s
    p6s.
    exe.
    value label p6s
    1 'No Satisfechos (1-6)'
    2 'Total Satisfechos (7-10)'.
    value label p4s to p5s
    1 'Sin intención (1-6)'
    2 'Total con intención (7-10)'.

```

➤ Creación de variables dummy

- Tipología de la prestación

```

if P5_PRES_AH=6      p5_Agua=1.
VALUE LABELS p5_Agua
1 'Agua'.
EXECUTE.

if P5_PRES_AH=7      p5_Cristales=1.
VALUE LABELS p5_Cristales
1 'Cristales'.
EXECUTE.

if P5_PRES_AH=8      p5_Robo=1.
VALUE LABELS p5_Robo
1 'Robo'.
EXECUTE.

if (P5_PRES_AH=9 | P5_PRES_AH=14)  p5_Fenomenos_Atmos=1.
VALUE LABELS p5_Fenomenos_Atmos
1 'Subida de tensión / viento'.
EXECUTE.

if (P5_PRES_AH=10 | P5_PRES_AH=11)  p5_Rotura=1.
VALUE LABELS p5_Rotura
1 'Rotuta de saneamientos / electrodomésticos'.
EXECUTE.

if P5_PRES_AH=15      p5_Cerrajería=1.
VALUE LABELS p5 Cerrajería
1 'Cerrajería'.
EXECUTE.

```

- Ligado a hipoteca

```

RECODE P2_A_PRES_AH (1=1) (ELSE=0) INTO P2_ligado_DM.
EXECUTE.
VALUE LABELS P2_ligado_DM
1 'Ligado a hipoteca'.
EXECUTE.

```

- Expectativas de servicio

```
RECODE P13_PRES_AH (1=1) (ELSE=0) INTO P13_mejor_DM.  
EXECUTE.
```

```
VALUE LABELS P13_mejor_DM  
1 'Expectativas mejor de lo esperado'.  
EXECUTE.
```

```
RECODE P13_PRES_AH (1=1) (ELSE=0) INTO P13_peor_DM.  
EXECUTE.
```

```
VALUE LABELS P13_peor_DM  
1 'Expectativas peor de lo esperado'.  
EXECUTE.
```

- Sexo

```
RECODE sexo (1=1) (ELSE=0) INTO Hombres_DM.  
EXECUTE.
```

```
VALUE LABELS Hombres_DM  
1 'Hombre'.  
EXECUTE.
```

```
RECODE sexo (1=1) (ELSE=0) INTO Mujeres_DM.  
EXECUTE.
```

```
VALUE LABELS Mujeres_DM  
1 'Mujer'.  
EXECUTE.
```

- Hábitat

```
RECODE habitat_cod (1=1) (ELSE=0) INTO habitat_Urbano_DM.  
EXECUTE.
```

```
VALUE LABELS habitat_Urbano_DM  
1 'Urbano >50.000hab'.  
EXECUTE.
```

```
RECODE habitat_cod (1=1) (ELSE=0) INTO habitat_Rural_DM.  
EXECUTE.
```

```
VALUE LABELS habitat_Rural_DM  
1 'Urbano <50.000hab'.  
EXECUTE.
```

- Zona Nielsen

```
RECODE nielsen (1=1) (ELSE=0) INTO nielsen_NORTE_DM.  
RECODE nielsen (2=1) (ELSE=0) INTO nielsen_SUR_DM.  
RECODE nielsen (3=1) (ELSE=0) INTO nielsen_CENTRO_DM.  
RECODE nielsen (4=1) (ELSE=0) INTO nielsen_GALAICA_DM.  
RECODE nielsen (5=1) (ELSE=0) INTO nielsen_LEVANTE_DM.  
RECODE nielsen (6=1) (ELSE=0) INTO nielsen_CATALANA_DM.  
EXECUTE.
```

```

VALUE LABELS nielsen_NORTE_DM
1 'Norte'.
VALUE LABELS nielsen_SUR_DM
1 'Sur'.
VALUE LABELS nielsen_CENTRO_DM
1 'Centro'.
VALUE LABELS nielsen_GALAICA_DM
1 'Galaica'.
VALUE LABELS nielsen_LEVANTE_DM
1 'Levante'-
VALUE LABELS nielsen_CATALANA_DM
1 'Catalana'.
EXECUTE.

```

- Canal de comunicación

```

if P6_PRES_AH =1 P6_PRES_AH_OFI=1.
if P6_PRES_AH <>1 P6_PRES_AH_OFI=2.
if P6_PRES_AH =2 P6_PRES_AH_BANCO=1.
if P6_PRES_AH <>2 P6_PRES_AH_BANCO=2.
if P6_PRES_AH =3 P6_PRES_AH_AGENTE=1.
if P6_PRES_AH <>3 P6_PRES_AH_AGENTE=2.
if P6_PRES_AH =4 P6_PRES_AH_CORR=1.
if P6_PRES_AH <>4 P6_PRES_AH_CORR=2.
if P6_PRES_AH =5 P6_PRES_AH_TLF=1.
if P6_PRES_AH <>5 P6_PRES_AH_TLF=2.
if P6_PRES_AH =6 P6_PRES_AH_WEB=1.
if P6_PRES_AH <>6 P6_PRES_AH_WEB=2.
if P6_PRES_AH =9 P6_PRES_AH_APP=1.
if P6_PRES_AH <>9 P6_PRES_AH_APP=2.
EXECUTE.

```

Modelos lineales generalizados.

```

GENLIN RECOMENDACION (REFERENCE=FIRST) BY
p5_Agua_DM p5_Cristales_DM p5_Robo_DM p5_Fenomenos_Atmos_DM
p5_Rotura_DM p5_Cerrajeria_DM P6_PRES_AH_OFI P6_PRES_AH_BANCO
P6_PRES_AH_AGENTE P6_PRES_AH_CORR P6_PRES_AH_TLF P6_PRES_AH_WEB
P6_PRES_AH_APP P2_ligado_DM P13_mejor_DM P13_peor_DM Mujeres_DM
Hombres_DM habitat_Urbano_DM habitat_Rural_DM nielsen_NORTE_DM
nielsen_SUR_DM nielsen_CENTRO_DM nielsen_GALAICA_DM
nielsen_LEVANTE_DM nielsen_CATALANA_DM p4_Repadaros
p4_Indemnizados P3_PRES_AH_DM P711_PRES_DM P712_PRES_DM
P713_PRES_DM P714_PRES_DM P715_PRES_DM P716_PRES_DM P717_PRES_DM
P911_PRES_DM P912_PRES_DM P913_PRES_DM P1011_PRES_DM
P1012_PRES_DM P1013_PRES_DM P1111_PRES_DM P1112_PRES_DM
P12_PRES_ADM p5s_DM p6s_DM P8_PRES_AH_DM (ORDER=ASCENDING)
WITH P2A P1_PRES_AH EDAD
/MODEL P2_ligado_DM P713_PRES_DM P1011_PRES_DM P1012_PRES_DM
P12_PRES_ADM p5s_DM p6s_DM P8_PRES_AH_DM P12_PRES_ADM*p5s_DM
P12_PRES_ADM*p6s_DM P12_PRES_ADM*P8_PRES_AH_DM p5s_DM*p6s_DM

```

```
p5s_DM*P8_PRES_AH_DM p6s_DM*P8_PRES_AH_DM
P12_PRES_ADM*p5s_DM*p6s_DM P12_PRES_ADM*p5s_DM*P8_PRES_AH_DM
P12_PRES_ADM*p6s_DM*P8_PRES_AH_DM p5s_DM*p6s_DM*P8_PRES_AH_DM
P12_PRES_ADM*p5s_DM*p6s_DM*P8_PRES_AH_DM INTERCEPT=YES
DISTRIBUTION=BINOMIAL LINK=LOGIT
/CRITERIA METHOD=FISHER(1) SCALE=DEVIANCE COVB=ROBUST
MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006 (ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3 (LR)
CILEVEL=95 CITYPE=WALD LIKELIHOOD=FULL
/MISSING CLASSMISSING=EXCLUDE
/PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION
(EXPONENTIATED) COVB CORB LMATRIX GEF HISTORY(1)
/SAVE MEANPRED CIMEANPREDL CIMEANPREDU PREDVAL XBPRED XBSTDERROR
COOK LEVERAGE RESID PEARSONRESID STDPEARSONRESID DEVIANCERESID
STDDEVIANCERESID LIKELIHOODRESID.
```