



Predicción de Siniestralidad en los Seguros de Autos

Seminario de Investigación

Morales Ramírez Ángel Francisco

**Professor
García Cid Kevin Eric**

Actuaria



Índice general

1. Introducción	3
1.1. Objetivo del Proyecto	3
1.2. Planteamiento del Problema	3
1.3. Técnicas Actuariales	4
1.4. Modelo de Regresión Logístico	4
1.4.1. Introducción	4
1.4.2. Aplicaciones	4
1.4.3. Tipos de Regresión Logística	4
1.4.4. Ventajas y Desventajas	4
1.4.5. Importación de Python	4
1.4.6. Matriz de Confusión	4
1.4.7. Curva ROC	4
1.5. Elección de los Datos	4
1.5.1. Descripción de la Variables	5
2. Justificación	7
3. Metodología	8
4. Marco Teórico	9
5. Desarrollo	10
6. Análisis de Resultados	11
7. Recomendaciones	12
8. Conclusiones	13
9. Anexos	14
10. Bibliografía	15



CAPÍTULO 1

Introducción

1.1 OBJETIVO DEL PROYECTO

El objetivo del presente proyecto es desarrollar un modelo predictivo que estime la probabilidad de siniestros en seguros de automóviles, utilizando técnicas de regresión logística.

1.2 PLANTEAMIENTO DEL PROBLEMA

En las empresas de seguros: el seguro de automóviles es el segundo con más demanda, después de los seguros de vida, por eso resulta de vital importancia identificar con antelación los posibles siniestros para evitar pérdidas. Para hacer esto, podemos implementar nuevos modelos con regresión logística identificando patrones y tendencias en la siniestralidad.

Un modelo que permita estimar la probabilidad de siniestros en función de variables clave, como la edad del conductor, el tipo de vehículo, el historial de siniestros y la zona geográfica, no solo ayudaría a las aseguradoras a fijar primas más justas, sino también a diseñar programas de prevención y reducir costos asociados a los siniestros.

Este proyecto permitirá a las aseguradoras mejorar sus modelos de tarificación, mejorando su rentabilidad y reduciendo sus pérdidas. Desde el punto de vista académico, este trabajo aporta una aplicación práctica de técnicas estadísticas a un problema real

Los problemas específicos a tratar de solucionar en este proyecto son:

- ¿Se pueden identificar patrones en la siniestralidad de autos usando modelos de regresión logística?
- ¿Cuáles son las variables determinantes en la probabilidad de siniestros en seguros de automóviles?
- ¿Se pueden usar modelos de regresión logística para el calculo de primas en seguros de autos?

1.3 TÉCNICAS ACTUARIALES

Las técnicas actuariales adecuadas para analizar estos datos, es la **regresión logística** ya que buscamos predecir la probabilidad de un evento (Que un auto sufra un siniestro).

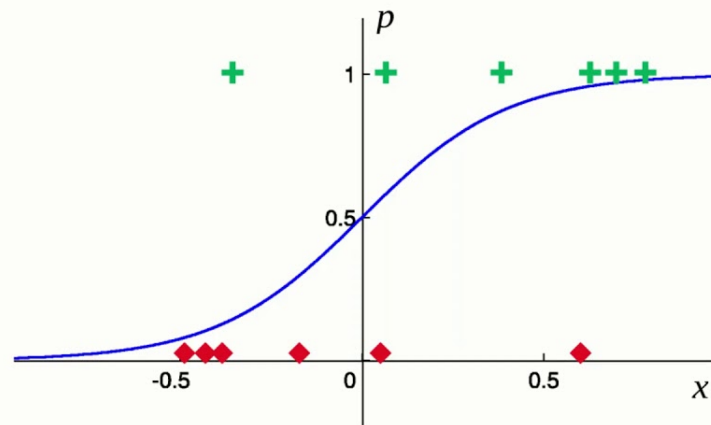
La función logística predice la ocurrencia de un evento entre 0 y 1 donde π es nuestra probabilidad de éxito, en este caso que el auto sufra un siniestro.

REGRESIÓN LOGÍSTICA

INTERPRETACIÓN

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 x$$

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$



1.4 MODELO DE REGRESIÓN LOGÍSTICO

1.4.1 INTRODUCCIÓN

1.4.2 APLICACIONES

1.4.3 TIPOS DE REGRESIÓN LOGÍSTICA

1.4.4 VENTAJAS Y DESVENTAJAS

1.4.5 IMPORTACIÓN DE PYTHON

1.4.6 MATRIZ DE CONFUSIÓN

1.4.7 CURVA ROC

1.5 ELECCIÓN DE LOS DATOS

Se tomará una base de datos de una aseguradora con la siguiente información de sus clientes: duración de la póliza, antigüedad del automóvil, marca y modelo del automóvil, etc para predecir si el



asegurado tendrá una futura reclamación. La base de datos se puede consultar en el siguiente link:

Dataset: Reclamaciones de seguros de automóviles

1.5.1 DESCRIPCIÓN DE LA VARIABLES

1. **policy_id**: Identificador Único del contrato póliza
2. **policy_ternure**: Plazo de la poliza
3. **age_of_car**: Edad del auto en años
4. **age_of_policyholder**: Edad del asegurado en años
5. **area_cluster**:
6. **population density**: Densidad de población de la ciudad del asegurado
7. **make**:
8. **segment**:
9. **model**: Nombre del modelo del auto
10. **fuel.type**: Tipo de combustible que utiliza el auto
11. **max_torque**: Torque máximo generado por el auto (Nm@rpm)
12. **max_power**: Potencia máxima generado por el auto (bhp@rpm)
13. **engine.type**: Tipo de motor utilizado por el auto
14. **airbags**: Número de bolsas de aire en el auto
15. **is_esc**: Valor booleano que indica
16. **is_adjustable_steering**: Valor booleano que indica
17. **is_tpms**: Valor booleano que indica
18. **is_parking_sensors**: Valor booleano que indica
19. **is_parking_camera**: Valor booleano que indica
20. **rear_brakes.type**: Tipo de frenos en la parte trasera del auto
21. **displacement**: Cilindrada del motor del auto
22. **cylinder**: Número de cilindros presentes en el motor del auto
23. **transmission.type**: Tipo de transmisión del auto
24. **gear_box**: Número de marchas / velocidades en el auto
25. **steering.type**:



26. **turning_radius:** El espacio que necesita el auto para realizar un giro (metros)
27. **length:** Longitud del auto (milímetros)
28. **width:** Ancho del auto (milímetros)
29. **height:** Altura del auto (milímetros)
30. **gross_weight:** Peso máximo permitido en el auto completamente cargado, es decir, incluido pasajeros, carga y equipamiento (Kg)
31. **is_front_fog_lights:** Valor booleano que indica si el auto tiene luces antiniebla delanteras
32. **is_rear_window_wiper:** Valor booleano que indica si el auto tiene limpiaparabrisas trasero
33. **is_rear_window_washer:** Valor booleano que indica si el auto tiene lava parabrisas trasero
34. **is_rear_window_defogger:** Valor booleano que indica si el auto tiene desempañador de la ventana trasera
35. **is_bake_assist:** Valor booleano que indica si el auto tiene la función de asistencia de frenado
36. **is_power_door_lock:** Valor booleano que indica si el auto tiene cerradura eléctrica
37. **is_central_locking:** Valor booleano que indica si el auto tiene función de cierre centralizado
38. **is_power_steering:** Valor booleano que indica si el auto tiene dirección asistida
39. **is_driver_seat_height_adjustable:** Valor booleano que indica si el auto tiene altura de asiento ajustable
40. **is_day_night_rear_view_mirror:** Valor booleano que indica si el auto tiene el espejo retrovisor diurno y nocturno
41. **is_ecw:** Valor booleano que indica si el auto tiene advertencia de verificación del motor (Engine Check Warning)
42. **is_speed_alert:** Valor booleano que indica si el auto tiene sistema de alerta de velocidad
43. **ncap_rating:** Calificación de seguridad otorgada por NCAP
44. **is_claim:** Valor booleano que indica si el asegurado sufrió un siniestro



CAPÍTULO 2

Justificación



CAPÍTULO 3

Metodología



CAPÍTULO 4

Marco Teórico



CAPÍTULO 5

Desarrollo



CAPÍTULO 6

Análisis de Resultados



CAPÍTULO 7

Recomendaciones



CAPÍTULO 8

Conclusiones



CAPÍTULO 9

Anexos



CAPÍTULO 10

Bibliografía