



Universidad de Valladolid



**ESCUELA DE INGENIERÍAS
INDUSTRIALES**

UNIVERSIDAD DE VALLADOLID

ESCUELA DE INGENIERIAS INDUSTRIALES

GRADO EN INGENIERÍA EN ORGANIZACIÓN INDUSTRIAL

**CARACTERIZACIÓN DE COLISIONES FRONTALES EN
CARRETERAS INTERURBANAS DE UNA CALZADA**

Autor:

De Cáceres Riol, Laura

Tutores:

Gordaliza Ramos, Alfonso

Fernández Temprano, Miguel Alejandro
Estadística e Investigación Operativa

Valladolid, Julio de 2017

*A todas las personas que me han acompañado durante estos cuatro años.
A mis tutores, por la ayuda y los conocimientos aportados durante esta etapa.*

RESUMEN

Las colisiones frontales y frontolaterales en carreteras interurbanas de una calzada suponen hasta el 12% del total de fallecidos en este tipo de vías en España, de acuerdo con el último informe de siniestralidad proporcionado por la Dirección General de Tráfico^[1]. Dado que estas carreteras constituyen el 90% de la red de carreteras de nuestro país, es prioritario buscar medidas para prevenir este tipo de accidentes.

La segregación de los carriles de carreteras interurbanas de una calzada (bien manteniendo 1+1 carriles o convirtiéndolas a las denominadas “carreteras 2+1”) supondría el evitar muchos de los fallecimientos ocurridos en este tipo de vías. De acuerdo con datos de la DGT, las llamadas “carreteras 2+1” podrían evitar hasta 338 muertes en España.

El presente Trabajo de Fin de Grado tiene como finalidad caracterizar estadísticamente los escenarios donde se producen las colisiones frontales en carreteras interurbanas de una calzada, así como aquellos tramos susceptibles de convertirse en carreteras 2+1.

De esta manera, se pretende ayudar a las administraciones a las que pertenecen las vías afectadas a conocer dónde actuar para evitar las colisiones frontales.

PALABRAS CLAVE

Colisión frontal, Accidentes de tráfico interurbanos, Estudios Caso-Control, Odds-Ratio, Curva ROC

ABSTRACT

The sum of both head on collisions and side collisions occurred in roads which have a single carriage in each direction amount to 12% of mortal victims in this kind of roads in Spain, according to the accident rates reported by the DGT.

Providing that this kind of roads account for over 90% of overall road network kilometres of our country, it is very important to take actions to prevent the head-on-collisions.

The lane segregation of one-carriage interurban roads, implementing a median (either keeping one lane, or converting them in “2+1” roads) could avoid up to 338 deaths in Spain.

The goal of present study is the identification of the factors associated with higher likelihood of head-on collisions, and to identify which roads could be converted in “2+1 roads”. With this characterization we want to help the administrations that own the roads to know where they should act to avoid this kind of accidents

KEY WORDS

Head-on-collisions, interurban traffic accidents, case-control studies, Odds-ratio, ROC curves.

ÍNDICE

INTRODUCCIÓN	1
ANTECEDENTES	1
JUSTIFICACIÓN	3
OBJETIVOS	4
CONSIDERACIONES ÉTICAS	5
CAPÍTULO 1 . MÉTODOS ESTADÍSTICOS UTILIZADOS	7
1.1. ESTUDIOS CASO-CONTROL	9
1.1.1. DIFERENCIA ENTRE ESTUDIO CASO-CONTROL Y ESTUDIO COHORTE	10
1.1.2. PROCEDIMIENTO DE APLICACIÓN DE UN ESTUDIO CASO-CONTROL RETROSPECTIVO	12
1.1.3. APLICACIÓN DE LA METODOLOGÍA CASO-CONTROL AL PRESENTE ESTUDIO	13
1.2. ODDS-RATIO: MEDIDA DE ASOCIACIÓN	14
1.3. REGRESIÓN LOGÍSTICA BINARIA	17
1.3.1. DEFINICIÓN Y UTILIDAD.....	17
1.3.2. OBTENCIÓN DE LA FUNCIÓN LOGÍSTICA Y ODDS-RATIO	18
1.3.3. UTILIDADES DE LA REGRESIÓN LOGÍSTICA	21
1.3.4. PROCEDIMIENTOS DE SELECCIÓN DE VARIABLES	21
1.4. CURVAS ROC (RECEIVER OPERATING CHARACTERISTIC)	22
1.4.1. SENSIBILIDAD Y ESPECIFICIDAD	22
1.4.2. REPRESENTACIÓN DE LAS CURVAS ROC.....	23
1.4.3. INTERPERTACIÓN DE LAS CURVAS ROC.....	24
1.5. VALIDACIÓN CRUZADA: LEAVE-ONE-OUT (LOOCV)	25
CAPÍTULO 2. BASE DE DATOS DE PARTIDA	27
DESCRIPCIÓN BASE DE DATOS DE PARTIDA	29
CAPÍTULO 3 . OBTENCIÓN Y TOMA DE DATOS DE LOS CASOS Y LOS CONTROLES 37	
3.1. CONSIDERACIONES PREVIAS A LA OBTENCIÓN DE LOS CASOS	39
3.1.1. PERÍODO TEMPORAL.....	39
3.1.2. UBICACIÓN GEOGRÁFICA DE LOS INDIVIDUOS INCLUIDOS EN EL ESTUDIO	39
3.1.3. RED DE CARRETERAS DE ESPAÑA	39
3.2. SELECCIÓN DE LOS CASOS	41
3.3. CONSIDERACIONES POSTERIORES A LA OBTENCIÓN DE LOS CASOS	44
3.4. OBTENCIÓN DE LOS CONTROLES	46
3.4.1. CREACIÓN DEL LISTADO DE CARRETERAS	46
3.4.2. CREACIÓN DE LA DISTRIBUCIÓN UNIFORME.....	47
3.4.3. OBTENCIÓN DE LAS UBICACIONES DE LOS CONTROLES.....	49
3.5. TOMA DE DATOS PARA LOS CASOS Y LOS CONTROLES	52
CAPÍTULO 4. VARIABLES DEL MODELO	55
4.1. VARIABLES INCLUIDAS EN LA BASE DA DATOS ARENA	57
ANCHURA DE LA CALZADA	57

ANCHURA DE CARRIL.....	57
ANCHURA DE ARCÉN	58
BARRERA DE SEGURIDAD	58
HITOS DE ARISTA	58
ACERA	59
NÚMERO DE CARRILES	59
4.2. VARIABLES ADICIONALES	59
RADIO DE CURVATURA.....	59
INTENSIDAD MEDIA DIARIA (IMD).....	61
CAPÍTULO 5. REGRESIÓN LOGÍSTICA SIMPLE	67
5.1. ANCHO DE CALZADA	70
5.1.1. TABLA CRUZADA ANCHO DE CALZADA*CASOCONTROL	70
5.1.2. ODDS-RATIO PARA LA VARIABLE “ANCHO DE CALZADA”	71
5.2. ANCHO DE CARRIL.....	74
5.2.1. TABLA CRUZADA ANCHURA DE CARRIL*CASOCONTROL	74
5.2.2. ODDS-RATIO PARA LA VARIABLE “ANCHO DE CARRIL”	76
5.3. RADIO DE CURVATURA	78
5.3.1. TABLA CRUZADA RADIO*CASOCONTROL.....	78
5.3.2. ODDS-RATIO PARA LA VARIABLE “RADIO DE CURVATURA”	80
5.4. ARCÉN.....	81
5.4.1. TABLA FRECUENCIA ARCÉN*CASOCONTROL	81
5.4.2. ODDS-RATIO PARA LA VARIABLE “ARCÉN”	82
5.5. BARRERA DE SEGURIDAD	84
5.5.1. TABLA CRUZADA BARRERA DE SEGURIDAD*CASOCONTROL	84
5.5.2. ODDS-RATIO PARA LA VARIABLE “BARRERA DE SEGURIDAD”	85
5.6. HITOS DE ARISTA.....	87
5.6.1. TABLA CRUZADA HITOS DE ARISTA*CASOCONTROL.....	87
5.6.2. ODDS-RATIO PARA LA VARIABLE “HITOS DE ARISTA”	88
5.7. ACERA.....	90
5.7.1. TABLA CRUZADA ACERA*CASOCONTROL.....	90
5.7.2. ODDS-RATIO PARA LA VARIABLE “ACERA”	91
5.8. INTENSIDAD MEDIA DIARIA CATEGORIZADA (IMD)	93
5.8.1. TABLA CRUZADA INTENSIDAD MEDIA DIARIA*CASO-CONTROL	93
5.8.2. ODDS-RATIO PARA LA VARIABLE “INTENSIDAD MEDIA DIARIA CATEGORIZADA” (IMD).....	94
5.9. NÚMERO DE CARRILES	97
5.9.1. TABLA CRUZADA NÚMERO DE CARRILES*CASO-CONTROL.....	97
5.9.2. ODDS-RATIO PARA LA VARIABLE “NÚMERO DE CARRILES”	98
5.10. RESUMEN DE INFLUENCIA DE LAS VARIABLES INDEPENDIENTES.....	100
CAPÍTULO 6. MODELOS MULTIVARIANTES: RESULTADOS Y VALIDACIÓN	103
6.1. RESULTADOS MULTIVARIANTES: ODDS-RATIO Y VALIDACIÓN DE LOS MODELOS.....	106
6.1.1. MODELO CURVAS Y RECTAS.....	106
6.1.2. MODELO RECTAS.....	113
6.1.3. MODELO CURVAS.....	119

6.2. MODELOS FINALES SELECCIONADOS	124
6.2.1. MODELO FINAL “COMPLETO” PARA CURVAS Y RECTAS: MODELO FORWARD/BACKWARD	124
6.2.2. MODELO FINAL PARA LAS RECTAS, FORWARD/BACKWARD.....	128
6.2.3. MODELO FINAL PARA LAS CURVAS, FORWARD/BACKWARD.....	130
CONCLUSIONES Y LÍNEAS FUTURAS DEL TRABAJO.....	135
CONCLUSIONES DEL ESTUDIO	137
RECOMENDACIONES.....	139
FUTURAS LÍNEAS DEL TRABAJO.....	142
BIBLIOGRAFÍA	145
ANEXOS	153
ANEXO 1- CUESTIONARIO ESTADÍSTICO DE ACCIDENTES DE CIRCULACIÓN CON VÍCTIMAS	155
ANEXO 2- CÓDIGOS DE PROGRAMACIÓN EN SAS.....	157
MODELO COMPLETO: CURVAS Y RECTAS.....	157
MODELO “SOLO RECTAS”	157
MODELO “SOLO CURVAS”	158
ANEXO 3- TABLAS	159
ANEXO 4- ILUSTRACIONES	161
ANEXO 5. ECUACIONES	163

INTRODUCCIÓN

ANTECEDENTES

Las carreteras interurbanas de una calzada suponen el 90% de la longitud total de carreteras de nuestro país. En ocasiones se presta poca atención al cuidado y diseño de estas carreteras, a pesar de que acaparan un 74% del total de fallecidos en accidentes de tráfico en España. [2]

Una de las medidas cuya implantación está recibiendo más atención recientemente para reducir la accidentalidad en estas carreteras es la segregación de carriles en carreteras de una calzada. Esta alternativa está mostrando excelentes resultados en cuanto a la reducción de cifras de accidentalidad en países como Irlanda, Suecia, Nueva Zelanda, Alemania, o en España, especialmente cuando se realiza la conversión de carreteras “1+1” (un carril por sentido) a “carreteras 2+1”.

El diseño de la vía 2+1 consiste en una sección transversal de tres carriles. Los dos carriles exteriores son carriles generales de tránsito para cada uno de los sentidos de circulación. El carril central se va turnando en un sentido u otro mediante el uso de una mediana, de forma que los conductores que circulan por ambos sentidos puedan adelantar por igual. Generalmente, el carril central se alterna cada 1,2-5 kilómetros, con una zona de transición de hasta 300 metros de longitud.[3][4]

Las colisiones frontales consisten en el choque frontal de dos vehículos que circulan en sentidos opuestos. Sin embargo, en el presente estudio también se considerarán las colisiones frontolaterales que ocurren en carreteras interurbanas fuera de intersección, dado que con motivo de un acto reflejo de uno de los conductores, una colisión frontal podrá ser registrada como colisión frontolateral en la base de datos ARENA (base de datos de la Dirección General de Tráfico utilizada para la gestión de la información sobre accidentes). Explicado esto, emplearemos el término “colisión frontal” para hacer referencia a los dos tipos colisiones que han sido explicados.

A pesar de que en España las colisiones frontales representan tan sólo un 3% del total de accidentes, los fallecidos en colisiones frontales y frontolaterales representan más del 12 % del total de fallecidos en accidentes de tráfico. [1]

De acuerdo con el siguiente gráfico, las colisiones frontales se sitúan como la tercera causa de fallecimiento en las carreteras españolas. Por tanto, cabe destacar la severidad de este tipo de colisiones.

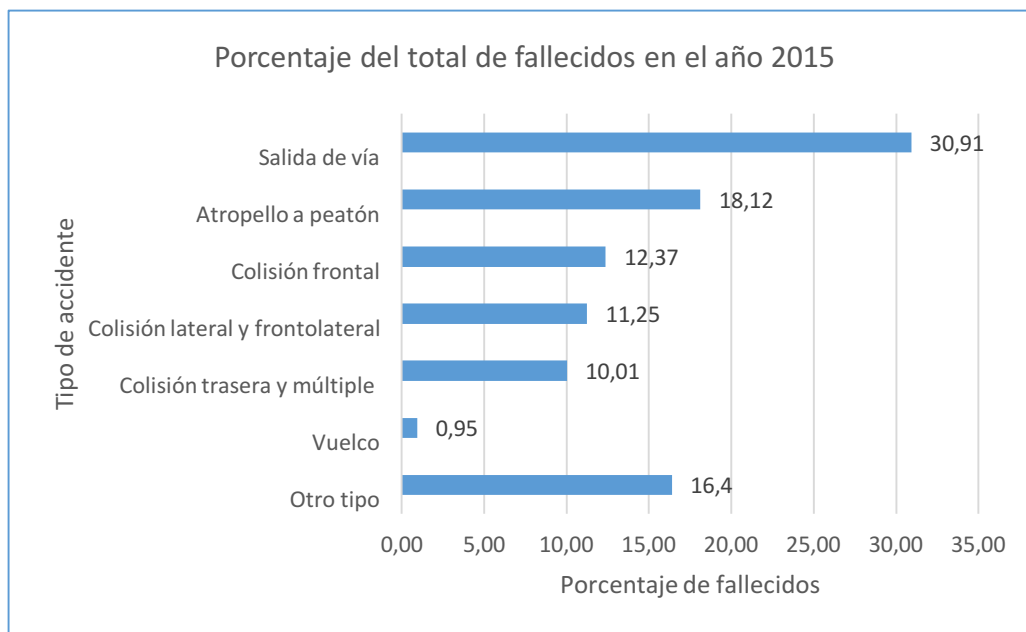


Ilustración 1: Fallecidos en accidentes de tráfico durante el año 2015 [1]

Estos datos han conducido a la realización de múltiples estudios, tanto a nivel nacional como internacional.

A nivel nacional, la Generalitat de Cataluña es pionera en el estudio y la prevención de este tipo de accidentes en España, así como en la instalación de Sistemas de Contención de Vehículos (SCV), y en la segregación de carriles. Recientemente, llevaron a cabo el estudio “Colisiones frontales en vías interurbanas de doble sentido: un problema de salud pública en seguridad vía” [2], en el que se estudiaban los factores que influían en este tipo de colisiones comparándolos con otro tipo de accidentes, empleando como metodología la Regresión de Poisson.

Este estudio realizado por la Generalitat de Cataluña no sólo incluía las características geométricas de diseño de la vía, si no que también estudiaban la influencia de las personas y los vehículos implicados en las colisiones. Además, tan solo se incluían las variables recogidas en la base de datos ARENA. La metodología utilizada permitía conocer las variables que tenían influencia en las colisiones frontales con respecto a otro tipo de colisiones.

Por el contrario, el presente Trabajo Fin de Grado va a emplear la metodología “caso-control” para estudiar la influencia que ejercen sobre la ocurrencia de este tipo de colisiones las variables geométricas de diseño de la vía, así como las variables de tráfico de la misma. Además, no se limitará al uso de las variables recogidas en la Base de Datos ARENA, sino que se añadirán variables complementarias, que puedan contribuir a aumentar la capacidad explicativa del modelo.

A nivel internacional, un reciente estudio desarrollado en Australia: “Verification of Austroads Road Design Criteria Based on Objective Safety Evidence”^[5], analizó la relación entre determinados factores de diseño de las carreteras y la seguridad de las mismas. Sin embargo, este estudio no se centraba en ningún tipo de accidente en particular.

La actualidad del tema no sólo queda demostrada por los diferentes estudios realizados recientemente y por la literatura a la que han dado lugar, sino también por la constatación de la organización en los últimos años de numerosos eventos científico-técnicos dedicados a este asunto, en particular la jornada técnica organizada en Cataluña en el mes de junio de 2017^[6], para dar a conocer este tipo de solución o diferentes ponencias como el Congreso Nacional de Seguridad Vial de Valencia en Julio 2017.

De acuerdo con la información expuesta anteriormente, se hace necesario realizar un estudio estadístico, con el fin de obtener las localizaciones más adecuadas para la instalación de SCV (Sistemas de Contención de Vehículos) que permitan evitar las colisiones frontales.

En el presente Trabajo Fin de Grado se plantea un estudio estadístico inferencial, con el fin de caracterizar los factores geométricos de diseño de la vía que influyen en las colisiones frontales.

JUSTIFICACIÓN

La justificación de la realización del presente estudio radica en que se espera que contribuya a la reducción del número de fallecidos en carreteras interurbanas como consecuencia de colisiones frontales. Esto conllevaría una mejora de la seguridad vial en las carreteras españolas.

La aplicabilidad de las conclusiones obtenidas en el campo de la seguridad vial sería inmediata, dado que determinadas administraciones ya han mostrado interés en esta propuesta, con el objetivo de evitar las colisiones frontales en carreteras de una calzada, así como mejorar la movilidad en las mismas durante determinadas “horas punta”, habilitando carriles reversibles correctamente protegidos .

Estas administraciones son titulares de numerosas carreteras de una calzada, susceptibles a convertirse en carreteras 2+1. Por ello, son coautoras de documentos normativos o recomendaciones como “Criterios de aplicación de SCV” (Ministerio de Fomento)^[7] y “Recomendaciones específicas sobre SCV en Carreteras Locales” (Comunidades Autónomas y Diputaciones Forales) en los que se define dónde proteger y qué tipo de SCV utilizar.^[8]

El impacto socioeconómico de este proyecto está relacionado principalmente con la realización de recomendaciones a las adiministraciones titulares de las carreteras (principalmente administraciones regionales o nacionales), aplicando las conclusiones obtenidas. De esta manera, se informará a estas administraciones de las características geométricas de diseño de la vía de las ubicaciones dónde se producen las colisiones frontales, para que puedan plantear en sus planes directores de seguridad vial en qué sitios concretos se puede actuar de manera eficaz para reducir las colisiones frontales, y con ello, disminuir la cifra de fallecidos en vías interurbanas.

Como ya se ha sido mencionado, el carácter innovador de este estudio con respecto a los realizados anteriormente es que se determinarán los factores de riesgo de ocurrir colisiones frontales frente a no haber ocurrido, empleando técnicas caso-control y analizando un conjunto de variables diferente. Por tanto, se obtendrán resultados diferentes a los obtenidos en estudios similares, en los que se analizaron factores de riesgo de ocurrir colisiones frontales frente al resto de tipos de accidente.

OBJETIVOS

El objetivo del presente Trabajo Fin de Grado es caracterizar estadísticamente las localizaciones donde se producen las colisiones frontales en carreteras interurbanas de una calzada. De este modo, se obtendrán las características de la vía que se comportan como factores de riesgo y de protección, permitiendo así tomar decisiones fundamentadas estadísticamente acerca de en qué ubicaciones es prioritario actuar para evitar las colisiones frontales.

Una de las alternativas para prevenir este tipo de accidentes es la instalación de barreras de segregación seguras, es decir, barreras cuyo comportamiento ante el impacto de un vehículo sea correcto en función de unas normativas de ensayo. Con ello, se contribuirá a la reducción del número de fallecidos en carreteras interurbanas como consecuencia de colisiones frontales.

Como posibles variables de influencia, se incluirán adicionalmente en el estudio características geométricas de diseño de la vía, así como la Intensidad Media Diaria de tráfico, que se obtendrá de las administraciones titulares de carreteras que han mostrado su interés en este proyecto.

El objetivo del presente TFG no es solo la caracterización de las carreteras en las que ocurren accidentes de tráfico, si no que también se pretende que los resultados sean útiles a las administraciones titulares de las vías para que puedan actuar de forma eficaz en la prevención de este tipo de accidentes. Por ello, una vez caracterizados los escenarios en los que se producen colisiones frontales, se aplicará un segundo criterio ("Anchura de la calzada mayor de 7

metros”), detectando aquellos tramos de carretera que, a priori, podrían ser susceptibles de “convertirse en tramos 2+1”. De esta manera, se estaría avalando con información objetiva el hecho de que las administraciones apostasen por una actuación de coste medio pero de alta prioridad. Todo ello conllevaría la mejora de las cifras de accidentalidad y un mejor ratio coste-beneficio.

La Epidemiología ha desarrollado las metodologías necesarias para llevar a cabo estudios y análisis estadísticos sobre una determinada “población”. Se realizará un estudio “caso-control”, en el que los casos representan las ubicaciones afectadas por el evento “colisión frontal” en vías interurbanas de una calzada, y los controles aquellas ubicaciones seleccionadas al azar sobre las mismas vías donde han ocurrido los casos, aunque evitando lógicamente la coincidencia de controles en las mismas ubicaciones que los casos.

Los análisis estadísticos se aplicarán haciendo uso de la Base de Datos ARENA. Concretamente, se utilizarán los datos de colisiones frontales que han ocurrido en el período temporal 2009-2013, en Castilla y León, en vías de titularidad autonómica y estatal, por los motivos que se expondrán posteriormente.

En resumen, el presente estudio nos permitirá determinar qué factores geométricos de diseño y de tráfico de la vía se comportan como factores de riesgo o de protección, afectando al riesgo de que en una determinada ubicación se produzca el desenlace estudiado “colisión frontal”. Además, permitirá determinar en qué ubicaciones es recomendable intervenir, con el objetivo de obtener un ratio coste-beneficio adecuado, pero llevando a cabo actuaciones de alta prioridad.

CONSIDERACIONES ÉTICAS

Para la realización del análisis estadístico, se recurrirá a la Base de Datos Arena de la Dirección General de Tráfico.

En ninguna de estas consultas se manejarán datos que no estén completamente anonimizados (cumpliendo con Ley Orgánica 15/1999 de Protección Datos), de manera que este proyecto no supondrá ningún problema ético.

CAPÍTULO 1 . MÉTODOS ESTADÍSTICOS UTILIZADOS

1.1. ESTUDIOS CASO-CONTROL

La metodología “caso-control” es un potente método que permite la investigación de factores que pueden explicar la ocurrencia de un tipo de evento en particular. Es ampliamente utilizada en estudios epidemiológicos. Sin embargo, su utilización se está extendiendo a muchos otros campos. [9]

Generalmente, se estudia la influencia de uno o varios factores sobre una variable respuesta dicotómica, que toma los siguientes valores:

Valor de la variable respuesta	Caso/control	Estado del individuo
Y=1	Casos	Ha experimentado el desenlace estudiado
Y=0	Controles	No ha experimentado el desenlace estudiado.

Tabla 1. Posibles valores de la variable respuesta dicotómica

Particularizando la tabla anterior al presente estudio:

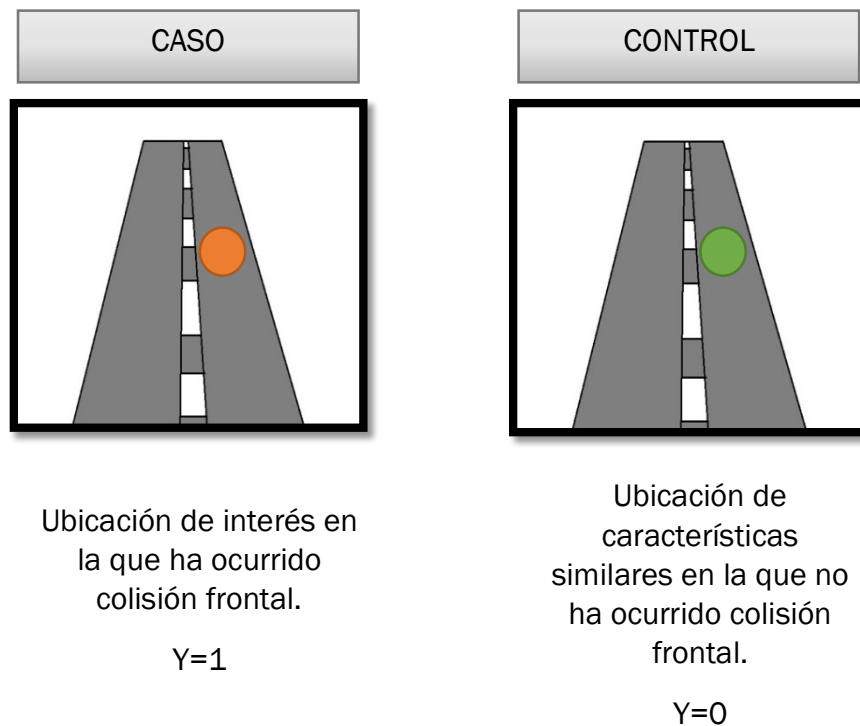


Ilustración 2: Casos y controles particularizados al presente estudio

1.1.1. DIFERENCIA ENTRE ESTUDIO CASO-CONTROL Y ESTUDIO COHORTE

En el presente apartado, se pretende distinguir los estudios caso-control, metodología empleada en el presente Trabajo Fin de Grado, de los estudios cohorte. Además, se clasifican los diferentes tipos de estudios caso-control y estudios cohorte que existen.

ESTUDIOS COHORTE

Los estudios cohorte se caracterizan porque en primer lugar se identifica a la población incluida en el estudio, y posteriormente se desarrolla un resultado de interés en esta población, tras ser expuestos o no a los factores de riesgo.

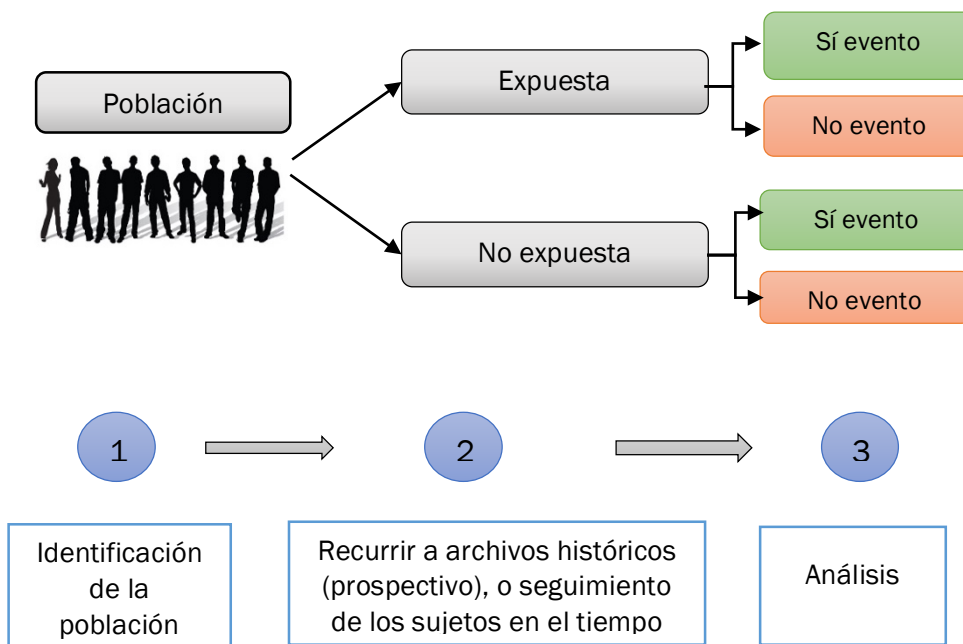


Ilustración 3 Esquema de un estudio cohorte

Una ventaja de este tipo de estudios es que permiten observar simultáneamente varias variables respuesta. Además, son útiles para investigar variables de exposición inusuales. [10]

Sin embargo, entre los inconvenientes se encuentra el hecho de que, cuando la probabilidad de ocurrencia del evento es baja, requieren una muestra de gran tamaño, y que es necesario seguir el transcurso del estudio durante un largo periodo de tiempo. [10]

Dentro de los estudios cohorte, se distinguen dos tipos:

- **Estudios prospectivos.** Son estudios iniciados en el presente, pero en los que es necesario esperar para observar la aparición del resultado estudiado en el futuro. Es decir, se recopilan las variables de exposición, y posteriormente se observa qué resultado aparece. Existen métodos de recolección de datos particulares de este tipo de estudios. Su principal inconveniente es que se alargan en el tiempo hasta que se desarrolle el evento de interés.
- **Estudios retrospectivos.** Son más típicos de los estudios caso-control, pero también existen estudios cohorte retrospectivos. Su principal característica es que se realizan en el presente, pero reconstruyen datos históricos del pasado.

ESTUDIOS CASO-CONTROL

La metodología caso-control se caracteriza porque los sujetos pertenecientes al subgrupo de los “casos” ya han experimentado el resultado de interés en el pasado. Una vez identificados los “casos”, se obtienen los “controles”. Finalmente, se obtiene información de las variables de exposición.

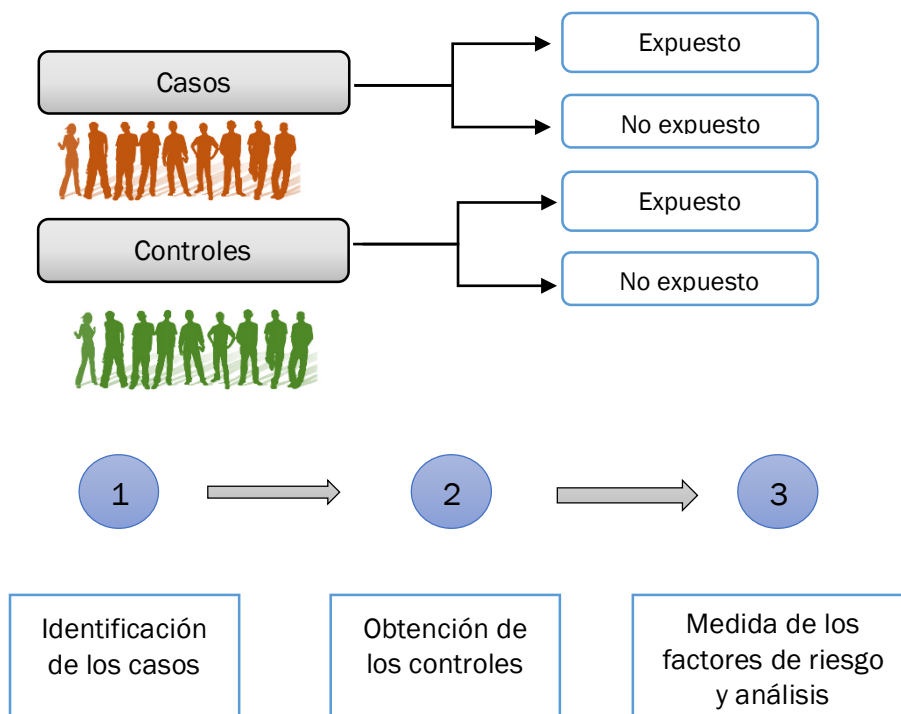


Ilustración 4. Esquema de un estudio caso-control

Se pueden distinguir dos tipos de estudios caso-control:

- **Estudio caso-control prospectivo:** los casos se identifican posteriormente al inicio del estudio, de modo que existe la posibilidad de la inclusión de nuevos casos durante el transcurso del mismo.
- **Estudio caso control retrospectivo:** los casos son identificados previamente al inicio del estudio.

Las ventajas de un estudio caso control es que permiten estudiar los factores de riesgo asociados a eventos con escasa probabilidad de ocurrencia, o con un largo período de latencia, de manera rápida. Además, son relativamente baratos y rápidos, y requieren menos tamaño de muestra que los estudios cohorte, anteriormente explicados. Se pueden estudiar múltiples variables de exposición o factores de riesgo. [10][11]

Sin embargo, este tipo de estudios no permite calcular el riesgo absoluto, por lo que se verá la necesidad de utilizar otras medidas de asociación. Además, hay que tener especial cuidado con los sesgos que pueden surgir al seleccionar los casos y los controles. Por otro parte, son ineficientes en cuanto el estudio de factores de exposición inusuales. [10][11]

1.1.2. PROCEDIMIENTO DE APLICACIÓN DE UN ESTUDIO CASO-CONTROL RETROSPECTIVO

SELECCIÓN DE LOS CASOS

En primer lugar, se identifican los casos que serán incluidos en el estudio. Para ello, es necesario tener presente las siguientes consideraciones: [12]

- Establecer de manera precisa cuál es el desenlace de interés.
- Determinar cuáles son los criterios de inclusión que han de cumplir los individuos del estudio.

SELECCIÓN DE LOS CONTROLES

Una vez obtenidos los casos, se seleccionarán los controles, o individuos de características similares a los casos, en los que no ha ocurrido el desenlace de interés.

La selección de los controles es una parte fundamental del estudio, pues tienen como utilidad estimar la proporción de exposición esperada de un grupo que no ha desarrollado el evento.

Por ello, hay que tener presentes las siguientes consideraciones: ^{[12][13]}

- Generalmente, los controles se corresponden con un grupo de la subpoblación de individuos que está en riesgo de desarrollar el evento estudiado, y de los cuales se puede asegurar que en el caso de que lo desarrollen, quedarán incluidos dentro del grupo de los casos. Es decir, los casos y los controles deben entenderse como dos subgrupos que proceden de una misma población.
- Cuando los casos se obtienen de una población definida en tiempo, espacio y lugar, y estos constituyen una muestra representativa del evento de estudio, la elección de los controles se puede realizar mediante un muestro aleatorio simple de la población base.
- Deben presentar la misma probabilidad de exposición que los casos.

RECOPIACIÓN DE INFORMACIÓN

De los casos y los controles, se recopilará la misma información. De este modo, se podrá estudiar la relación entre los factores de riesgo y el desenlace estudiado, dado que se podrá comparar la frecuencia de exposición a todos los factores entre los casos y los controles.

1.1.3. APLICACIÓN DE LA METODOLOGÍA CASO-CONTROL AL PRESENTE ESTUDIO

En el estudio a realizar en el presente Trabajo Fin de Grado, se estudiarán los factores de riesgo que influyen en el resultado “Colisión Frontal”.

Los casos se tomarán de la Base de Datos ARENA. Por tanto, se tomarán ubicaciones en la que ya han ocurrido colisiones frontales en el pasado, y que se adecuen a las condiciones impuestas en el estudio. Para la selección de los casos que son de interés, se impondrán una serie de condiciones a través de las variables recogidas en la Base de Datos de la DGT.

Por otra parte, mediante un muestreo aleatorio de todas las carreteras de la zona geográfica incluida en el estudio, se obtendrán los controles. Se trata de

ubicaciones en las que no hayan ocurrido colisiones frontales de acuerdo con la Base de Datos, pero que también cumplan las condiciones impuestas en el estudio, ya que los casos y controles han de cumplir características similares, de modo que, si en una determinada ubicación incluida en grupo de los controles ocurre una colisión frontal, estaría incluida en el grupo de los casos.

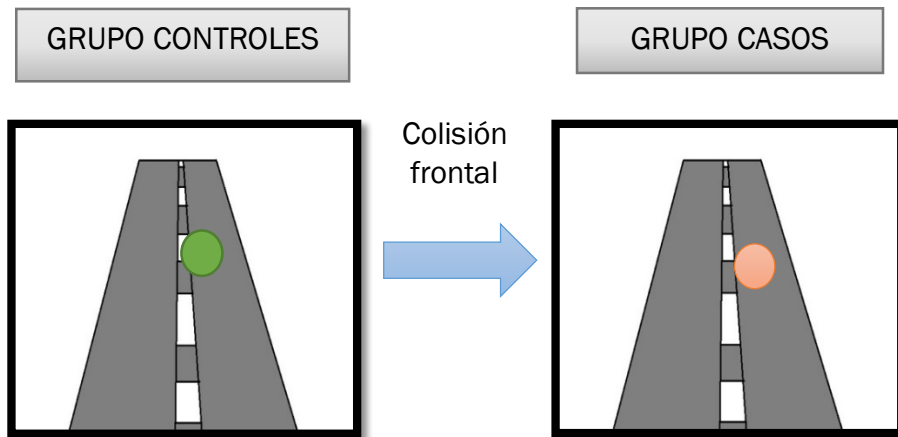


Ilustración 5. Relación entre casos y controles en el presente estudio

Posteriormente, se tomarán los datos de los factores de riesgo, o se utilizan los ya disponibles en la Base de Datos, en función de si se encuentran disponibles o no.

Por tanto, el presente trabajo Fin de Grado se sirve de la metodología “caso-control” para analizar los factores de riesgo que influyen en las colisiones frontales.

Finalmente, se crea la variable caso-control, que tomara el valor de “1” en los casos, y “0” en los controles, y actúa como variable respuesta.

1.2. ODDS-RATIO: MEDIDA DE ASOCIACIÓN

En la metodología caso-control, la proporción de casos en riesgo de toda la población es desconocida, por lo que no se puede medir la incidencia en términos absolutos. Es decir, los controles son una muestra representativa de la población, por lo que no se puede obtener el total de población en riesgo. [14] Por tanto, en este tipo de estudio se recurre a otros tipos de medidas de asociación.

El Odds-ratio es la medida de asociación más comúnmente utilizada en este tipo de estudios. Muestra la asociación entre una variable de exposición y el

desenlace. Por tanto, estima la probabilidad de que un evento se desarrolle dados unos determinados factores de riesgo. [15]

Uno de los motivos por lo que los Odds-ratio son tan usados es debido a que su cálculo no conlleva ninguna suposición adicional, si no que se pueden obtener directamente de la regresión logística. Sirven para expresar la probabilidad de que un evento ocurra respecto a la probabilidad de que no ocurra. Considerando el caso más simple, en el que la variable de exposición toma dos posibles valores, y la variable respuesta es dicotómica, el Odds-Ratio puede expresarse del siguiente modo:

$$\text{Odds} - \text{Ratio} = \frac{\text{Odds de exposición en los casos}}{\text{Odds de exposición en los controles}}$$

Los individuos de un determinado estudio se pueden clasificar en función de la variable de exposición y del resultado, en los siguientes grupos: [15]

		Resultado	
		+ (Casos)	- (controles)
Variable de exposición	+ (expuesto)	A	B
	- (no expuesto)	C	D

Tabla 2. Relación entre la variable de exposición y el resultado para el cálculo del Odds-ratio

Por tanto, se pueden distinguir las siguientes variantes del resultado en función si se encuentra expuesto o no a los factores de riesgo:

- A- Número de **casos expuestos**
- B- Número de **controles expuestos**
- C- Número de **casos no expuestos.**
- D- Número de **controles no expuestos**

Conociendo que el Odds-ratio es la Probabilidad de que el evento se desarrolle, respecto a que no lo haga, se puede calcular del siguiente modo:

$$Odds - Ratio(OR) = \frac{\frac{\text{Casos expuestos}}{\text{Casos no expuestos}}}{\frac{\text{Controles expuestos}}{\text{Controles no expuestos}}} = \frac{A/C}{B/D} = \frac{AD}{BC}$$

Ecuación 1. Cálculo del Odds-ratio

La interpretación de los valores del Odds-ratio es la siguiente:

OR	SIGNIFICADO	TIPO DE ASOCIACIÓN	Tipo de factor
=1	El factor no influye en el transcurso del evento estudiado	No existen asociación entre la variable respuesta y la variable de exposición.	No existe asociación
>1	El factor aumenta el riesgo probabilidad de que se produzca el evento estudiado	Existe asociación positiva entre la variable respuesta y la variable de exposición.	Factor de riesgo
<1	El factor disminuye el riesgo de que se produzca el evento estudiado	Existe asociación negativa entre el factor de exposición y la variable respuesta.	Factor de protección

Tabla 3. Interpretación del Odds-ratio

El Intervalo de Confianza (IC) es utilizado como un indicador de la significación estadística del Odds-ratio. Cuando el IC no contiene al valor 1 se puede afirmar que es estadísticamente significativo. Sin embargo, sería incorrecto interpretar un OR con un $IC_{95\%}$ asociado que contiene al 1 como un indicador de la evidencia de la falta de asociación entre la variable y el resultado. [15]

Para el cálculo de los Intervalos de Confianza se escoge el nivel de significación que se pretende obtener (α), y se obtiene el valor de Z, el cual depende del nivel de significación elegido. [16]

$$IC_{1-\alpha} \text{ Ln OR} = \text{Ln OR} \pm Z \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

$$IC_{1-\alpha}OR = e^{\text{Ln } OR \pm Z \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}} = OR * e^{\pm Z \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}}$$

Ecuación 2 Intervalo de Confianza para el Odds-ratio

Como resumen de la interpretación de los Intervalos de Confianza, se ha elaborado la siguiente tabla:

IC	Tipo de factor
Ambos extremos son mayores que 1	Factor de riesgo
Contiene al 1	No existe evidencia de la asociación
Ambos extremos son menores que 1	Factor de protección

Tabla 4. Interpretación de los Intervalos de Confianza para el Odds-ratio

1.3. REGRESIÓN LOGÍSTICA BINARIA

1.3.1. DEFINICIÓN Y UTILIDAD

La regresión logística es un modelo matemático comúnmente utilizado para determinar la influencia una o varias variables de exposición sobre una variable respuesta dicotómica. En el modelo se distingue:

- **Variable dependiente:** se trata de una variable dicotómica. Los sujetos que experimentan el desenlace estudiado se clasifican como 1, y los que no lo experimentan se clasifican como 0.
- **Variables independientes:** son las variables que podrían tener influencia en el modelo. También se denominan factores de riesgo.

Un aspecto clave de este tipo de estudio es determinar correctamente las variables de exposición. Se han de tener en cuenta todas las variables independientes:

$$(X_1, X_2, X_3 \dots X_k), k \text{ variables independientes.}$$

Si el número de variables es mayor que 1 ($K > 1$), se trata de un problema multivariante. Se pretende encontrar la influencia que las k variables exposición ejercen sobre la variable dependiente (D):



Ilustración 6. Objetivo de la función logística

El interés de los modelos multivariantes reside en su capacidad para evaluar no solo la contribución conjunta de todas las variables independientes para explicar la respuesta, sino también para evaluar la contribución particular de cada una de las variables en presencia de las restantes. [17]

1.3.2. OBTENCIÓN DE LA FUNCIÓN LOGÍSTICA Y ODDS-RATIO

Si se pretende ajustar un determinado modelo multivariante de manera lineal, se puede hacer uso de la siguiente expresión.

$$z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

{

$X_1, X_2 \dots X_k$: variables de exposición

$\beta_1, \beta_2, \dots, \beta_k$: coeficientes de regresión

Ecuación 3. Expresión lineal modelo logístico

Si calculamos el valor esperado de Y dados los valores de X :

$$E(Y|X) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

De acuerdo con la expresión anterior, la variable respuesta podría tomar cualquier valor entre $-\infty$ y $+\infty$. Sin embargo, tal y como ha sido explicado anteriormente, la variable respuesta será de tipo dicotómica, cuyos posibles valores son "0" y "1". Por tanto, el rango de la variable respuesta se encontrará entre dichos valores.

$$0 \leq E(Y|X) \leq 1$$

Los problemas que surgen al ajustar linealmente el modelo son los siguientes:

1. Los errores no cumplirían la hipótesis de seguir una distribución normal.
2. No se cumple la homocedasteicidad de la varianza, dado que no es constante. En la distribución binomial la varianza depende de la media con lo que no puede haber varianza constante si la media varía.
3. El resultado no se expresa en términos de probabilidad, dado que puede tomar valores negativos, o ser mayor que 0.

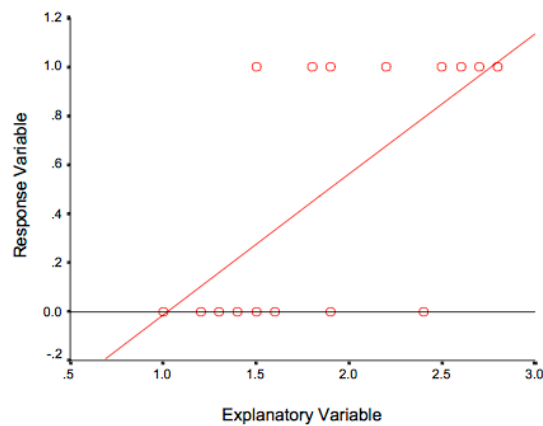


Ilustración 7. Ajuste lineal

Para solucionar este problema se suele recurrir a la función logística $\Phi(z)$, la cual describe la forma matemática del modelo logístico.

$$\phi(z) = \frac{1}{1 + e^{-z}}$$

Ecuación 4. Forma matemática del modelo logístico [20]

En el siguiente gráfico, se puede observar gráficamente la función del modelo logístico.

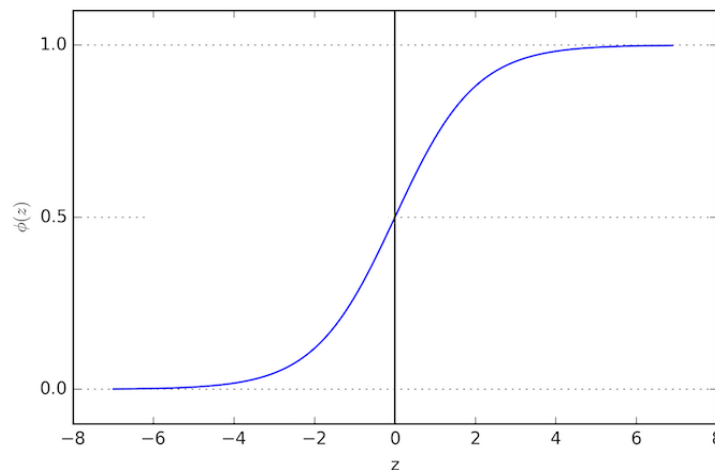


Ilustración 8 Representación Gráfica de la Función Logística

De acuerdo con el gráfico anterior, independientemente del valor que tome z , la función tomará un valor entre 0 y 1. Por tanto, el modelo es ideal para asociarlo con probabilidades.

Sustituyendo la expresión lineal de la ecuación 3 en la ecuación 2, se obtiene:

$$E(Y|X) = P = \frac{1}{1 + e^{-(\alpha + \sum \beta_i X_i)}} = \frac{e^{\alpha + \sum \beta_i X_i}}{1 + e^{\alpha + \sum \beta_i X_i}}$$

La expresión anterior se puede escribir en términos de probabilidad del siguiente modo:

$$\frac{P}{1 - P} = e^{\alpha + \sum \beta_i X_i}$$

Una transformación muy utilizada consiste en tomar logaritmos en ambas partes de la ecuación.

$$g(x) = \text{logit}(P) = \ln\left(\frac{P}{1 - P}\right) = \alpha + \sum \beta_i X_i$$

Ecuación 5. Transformación logarítmica

Este último modelo permite que $E(Y|X)$ se encuentre entre 0 y 1. Además, permite aplicar los principios de la regresión lineal, y los errores se describen mediante una variable binomial.

En la regresión logística, cada uno de los coeficientes de regresión (β_k) se corresponden con el cambio en el logaritmo de la variable respuesta al aumentar una unidad la variable independiente X_k . [15]

Retomando la ecuación 4, y haciendo uso de la ecuación 1, se obtiene:

$$OR = \frac{\frac{P(1)}{1 - P(1)}}{\frac{P(0)}{1 - P(0)}} = e^{\beta_i}$$

Ecuación 6. Odds-ratio

Por tanto, la función exponencial de un coeficiente de regresión (e^{β_k}) es el odds-ratio asociado a una unidad de incremento de la variable de exposición.

[15]

1.3.3. UTILIDADES DE LA REGRESIÓN LOGÍSTICA

La Regresión logística tiene tres utilidades fundamentales [18]:

1. Permite estimar la asociación estadística que existe entre dos o más variables independientes, y una variable dependiente dicotómica (Y).
2. Permite medir el signo de esa relación en caso de que exista.
3. Estima la probabilidad de que se produzca el suceso estudiado en función de los valores que adopten las variables independientes.

1.3.4. PROCEDIMIENTOS DE SELECCIÓN DE VARIABLES

Existen tres modos diferentes de seleccionar aquellas variables que son “influyentes” en el modelo.

- **Método Forward**, o método de selección hacia adelante, consiste en comenzar incluyendo en el modelo una única variable, e ir añadiendo variables en función de si mejoran la capacidad explicativa del modelo o no. El proceso finaliza cuando las variables excluidas del modelo no

aportan explicaciones adicionales, o cuando ya han sido añadidas todas las variables.

- **Método Backward**, o método de selección hacia atrás. Comienza incluyendo todas las variables en el modelo (k variables), y elimina paso a paso variables del modelo cuya influencia sobre la capacidad explicativa es menor. Se eliminan variables paso a paso, hasta que la eliminación de una variable supone una disminución acusada de la capacidad explicativa del modelo.
- **Método Stepwise**, es una combinación de los dos métodos anteriores. Es decir, en cada paso permite eliminar o añadir variables al modelo.

1.4. CURVAS ROC (RECEIVER OPERATING CHARACTERISTIC)

El significado de las curvas ROC está asociado a la capacidad del modelo planteado para discriminar el porcentaje de sujetos que experimentan el desenlace estudiado, respecto a aquellos que no lo hacen. Es decir, muestran la habilidad que posee un determinado modelo para clasificar o predecir una determinada condición. Además, este método proporciona el punto de corte óptimo. [19][20]

1.4.1. SENSIBILIDAD Y ESPECIFICIDAD

Generalmente, la exactitud de las predicciones se expresa a través de la sensibilidad y la especificidad. Estos parámetros se definen del siguiente modo en aquellos casos en los que la variable respuesta es dicotómica: [21]

- **La sensibilidad** es la probabilidad de clasificar correctamente a un sujeto como positivo frente al desenlace estudiado. Se denomina Fracción de verdaderos positivos (FVP)
En nuestro caso, la sensibilidad es la probabilidad de que el modelo clasifique a un determinado individuo como “colisión frontal” cuando realmente lo es.
- **La especificidad**, por el contrario, es la probabilidad de clasificar correctamente a un sujeto como negativo (FVN).

En nuestro estudio, consiste en clasificar a un determinado individuo como “no colisión frontal”, cuando realmente no se ha producido dicho accidente. Es equivalente a restar uno a la Fracción de Falsos Positivos (FFP)

Para cada sujeto a clasificar, se distinguen los siguientes posibles resultados: [20]

	Colisión frontal	No colisión Frontal
Prueba positiva	Verdadero positivo (A)	Falso positivo (C)
Prueba negativa	Falso negativo (B)	Verdadero Negativo (D)

Tabla 5. Valores de los casos en función de si se han clasificado correctamente o no

Por tanto, a partir de la clasificación de los individuos del estudio, se pueden calcular la sensibilidad y la especificidad como:

$$\text{Sensibilidad} = \frac{A}{A+B} \rightarrow \text{Fracción de verdaderos positivos (FVP)}$$

$$\text{Especificidad} = \frac{D}{C+D} \rightarrow 1 - \text{Fracción de falsos positivos (FFP)} = \text{Fracción de verdaderos negativos (FVN)}$$

Ecuación 7. Cálculo de la sensibilidad y especificidad

La magnitud contraria a la especificidad es la inespecificidad, que equivale a la fracción de falsos positivos.

1.4.2. REPRESENTACIÓN DE LAS CURVAS ROC

Como ha sido mencionado anteriormente, las curvas ROC permiten obtener un indicador de la exactitud global del modelo.

Mediante la representación de las dos magnitudes explicadas en el apartado anterior (sensibilidad en el eje “Y” y 1-especificidad en el eje “X”), obtenidas al considerar todos los posibles valores de corte de la prueba, se obtiene la curva ROC. Esta curva proporciona una visión global de la capacidad de clasificación de un determinado modelo.

El punto de corte hace referencia a la probabilidad a partir por la que un determinado caso es considerado como positivo. Es decir, si el punto de corte es 0,75, todos aquellos individuos cuya probabilidad de experimentar el resultado estudiado sea mayor de 0,75 son incluidos en el grupo de los positivos, y aquellos cuya probabilidad sea menor que 0,75 son considerados

como negativos. Por tanto, cuanto mayor sea el punto de corte, más aumenta la probabilidad de falsos negativos, pero menos la de falsos positivos.

1.4.3. INTERPERTACIÓN DE LAS CURVAS ROC

La capacidad de clasificación de la Curva ROC aumenta a medida que la curva se desplaza de la diagonal hacia el vértice. Esto significaría que no existe solapamiento entre los individuos que han experimentado el desenlace estudiado, y aquellos que no lo han hecho. [22] Por ello, cuando la sensibilidad es 1, la inespecificidad es 0, y a viceversa. Es decir, cuando la fracción de verdaderos positivos (FVP) vale 1, la de falsos positivos (FFP) vale 0, y, al contrario.

En la curva sin capacidad de predicción, la especificidad y la sensibilidad son iguales para todos los puntos, por lo que el solapamiento es total.

En función de los diferentes comportamientos “tipo” de la curva ROC, se pueden observar los siguientes casos:

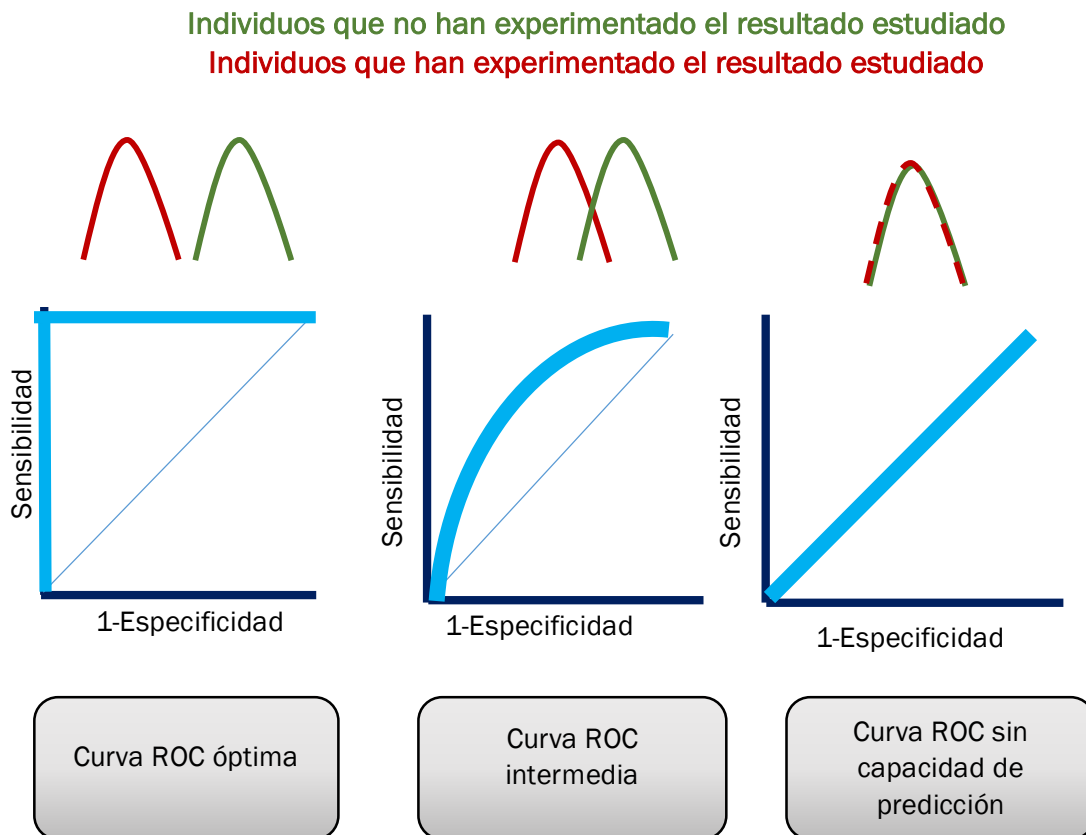


Ilustración 9. Comparación de diferentes curvas ROC

De modo que en la situación de la izquierda de la ilustración “Tipos Curvas ROC”, la sensibilidad y la especificidad serían perfectos, por lo que se trata de la curva óptima. En la práctica, las curvas serán similares a la “curva ROC intermedia” mostrada en Ilustración 9. Comparación de diferentes curvas ROC

El área bajo la curva ROC (AUC) también puede ser empleado como indicador de la exactitud de la prueba.

- La exactitud máxima se corresponde con un área igual a la unidad. Este caso se corresponde con aquel en que el 100% de los individuos, hayan experimentado o no el resultado a estudiar, se clasifican correctamente.
- Exactitud mínima se corresponde con un área igual a 0,5. Se corresponde con aquella situación en que los individuos presentan la misma probabilidad de clasificarse en un grupo que en otro. Por ello, el modelo se considera inútil, pues no se elimina la incertidumbre inicial.

Por tanto, cuanto más se aproxime el área a un valor igual a la unidad, mayor capacidad de discriminación presentará el modelo.

1.5. VALIDACIÓN CRUZADA: LEAVE-ONE-OUT (LOOCV)

El procedimiento Leave-One-Out es un método de validación de utilidad si se busca estimar el error con precisión. Este procedimiento se enmarca dentro de las metodologías de validación cruzada.

Para la validación del modelo se crean dos grupos, que presentan las siguientes características:

- Grupo del “dato excluido”: es aquel con el que se realiza el test. Incluye solo al caso que ha sido excluido en una determina iteración.
- Grupo del “resto de datos”: se incluyen el resto de datos del estudio. Se denominan datos de entrenamiento.

El procedimiento aplicado es bastante aproximado, pero presenta gran variabilidad, pues solo se realiza con un dato. Por eso, el proceso se repite un número de veces igual al número de datos disponibles. De este modo, todos los datos serán usados alguna vez como “dato excluido”.

ITERACIÓN	DATO EXCLUIDO	DATO EXCLUIDO	DATO EXCLUIDO	DATO EXCLUIDO
1º	1	1	1	1
2º	2	2	2	2
3º	3	3	3	3
...
Nº	n	n	n	n

Tabla 6. Caso excluido en cada interacción en el procedimiento Leave-One-Out.

En cada una de las iteraciones se realiza el test con el dato excluido, y se entrena con el resto de ellos, de modo que se obtiene si ha sido correctamente si ha sido clasificado correctamente. El error se calcula como la media aritmética de los errores de cada iteración para obtener un único resultado. [23]

Como ventaja, se encuentra el hecho de que permite obtener una estimación del error muy precisa, pues el error no depende de los individuos que se incluyan en cada grupo, y se repite para cada uno de los datos. Sin embargo, conlleva una gran carga computacional.

Dada la gran precisión de este método de validación (LOOCV), lo utilizaremos para la validación de los modelos considerados en el estudio. Inicialmente, calcularemos el porcentaje de clasificación aparente mediante SPSS, de manera que nos permita prever la capacidad de discriminación de los modelos. Posteriormente, comprobaremos que es correcto empleando la presente metodología, de modo que se obtenga con mayor exactitud el valor del error de todos los modelos considerados

CAPÍTULO 2. BASE DE DATOS DE PARTIDA

DESCRIPCIÓN BASE DE DATOS DE PARTIDA

ARENA es un sistema de captura, almacenamiento y gestión de la información sobre accidentes para la D.G.T, cuya principal finalidad es proporcionar un interfaz único de inserción de datos de accidentes de tráfico por parte de los usuarios autorizados. Asimismo, el sistema permite el tratamiento de la información recogida para la elaboración de estadísticas e informes sobre accidentalidad. [24]

Por tanto, esta base de datos permite acceder a la información acerca de todos los accidentes de tráfico ocurridos en España.

El formulario que completan las autoridades competentes para registrar los datos de los accidentes de tráfico se encuentra incluido en el anexo 1.

A partir de la información recogida a través de la aplicación ARENA, se genera el fichero de accidentes de tráfico. Este fichero contiene:

- La **información general** del accidente, lo que incluye datos de su localización, fecha de ocurrencia, circunstancias, características geométricas, etc.
- La **información particular** de los vehículos involucrados, incluyendo matrícula, características técnicas, condiciones, mercancías peligrosas, etc.
- La **información de las personas** que se ven afectadas por el accidente, distinguiendo entre conductores, pasajeros y peatones.

En el presente estudio se hará uso exclusivamente de la información general de accidentes, pues se considerarán variables relacionadas con las características geométricas de la ubicación en la que han tenido lugar los accidentes, dado que el interés del estudio es encontrar aquellas localizaciones cuyas características geométricas influyan en que exista mayor riesgo de colisiones frontales.

Las variables incluidas en el fichero general, así como las categorías de cada una de ellas, son las siguientes:

CODIFICACIÓN	ETIQUETA	VALORES
<i>provin</i>	Provincia	Adquiere un valor de dos dígitos en función de la provincia.
<i>isla</i>	Isla	Tendrá valor "0" debido a que únicamente toma valores positivos para las distintas islas de España.
<i>año</i>	Año	Es el año en que ha ocurrido el accidente.
<i>mes</i>	Mes	Toma valores de 1 a 12, en función del mes en que ha ocurrido el accidente.
<i>diasem</i>	Día de la semana	Toma valores de 1 a 7, en función del día de la semana en que ha ocurrido el accidente, empezando por "Lunes".
<i>dia</i>	Día del mes	Es el día del mes en que ha ocurrido el accidente.
<i>tipodia</i>	Tipo de día	"L" "Laborable" "F" "Festivo" "P" "Posterior a festivo" "A" "Anterior a festivo"
<i>zona</i>	Zona	1 "Carretera" 2 "Zona Urbana" 3 "Travesía" 4 "Variante"
<i>red</i>	Titularidad de la vía	1 "Estatál" 2 "Autonómica" 3 "Provincial" 4 "Municipal" 5 "Otras"
<i>sent</i>	Sentido	"A" "Ascendente" "D" "Descendente" "N" "No especificado"
<i>codmun</i>	Código municipio de	
<i>habit</i>	Habitantes	1 "Mas de cien mil" 2 "De 50 a 100 mil" 3 "De 5 a 50 mil" 4 "De mil a cinco mil" 5 "Menos de mil"

CODIFICACIÓN	ETIQUETA	VALORES
<i>tipovia</i>	Tipo de vía	1 "Autopista" 2 "Autovía" 3 "Vía rápida" 4 "Vía convencional con carril lento" 5 "Vía convencional" 6 "Camino vecinal" 7 "Vía de Servicio" 8 "Ramal de enlace" 9 "Otro tipo"
<i>numcarr</i>	Número de carriles	
<i>anchcal</i>	Anchura de calzada	1 "Menos de 5,99 metros" 2 "Entre 6 y 6,99 metros" 3 "De 7 metros o más"
<i>anchcarr</i>	Anchura de carril	1 "De más de 3,75 metros" 2 "De 3,25 a 3,75 metros" 3 "Menos de 3,25 metros"
<i>marcvial</i>	Marcas viales	1 "Inexistentes o borradas" 2 "Solo separación de carriles" 3 "Separación carriles y bordes" 4 "Solo separación de bordes"
<i>arcen</i>	Arcén	1 "Inexistente o impracticable" 2 "Menor de 1,5 metros" 3 "De 1,5 a 2,49 metros" 4 "De 2,5 metros en adelante"
<i>arcenpav</i>	Arcén pavimentado	"S" "Si" "N" "No" " "
<i>mediana</i>	Mediana entre calzadas	"S" "Si" "N" "No"
<i>barrera</i>	Barrera de seguridad	"S" "Si" "N" "No"
<i>panel</i>	Paneles direccionales	"S" "Si" "N" "No"
<i>hitos</i>	Hitos de arista	"S" "Si" "N" "No"
<i>captafar</i>	Captafaros	"S" "Si"

CODIFICACIÓN	ETIQUETA	VALORES
		"N" "No"
<i>fueraint</i>	Fuera intersección de	1 "Recta" 2 "Curva suave" 3 "Curva fuerte sin señalizar" 4 "Curva fuerte con señal y sin velocidad señalizada" 5 "Curva fuerte con señal y velocidad señalizada"
<i>tipoint</i>	Tipo de intersección	1 "En T o Y" 2 "En X o +" 3 "Enlace de entrada" 4 "Enlace de salida" 5 "Giratoria" 6 "Otros"
<i>aconint</i>	Acondicionamiento de intersección	1 "Nada especial" 2 "Solo isletas o paso para peatones en vía secundaria" 3 "Paso para peatones o isleta en centro de la vía principal" 4 "Carril central de espera" 5 "Raqueta de giro a izquierda" 6 "Otro tipo"
<i>priori</i>	Prioridad regulada por	1 "Agente" 2 "Semáforo" 3 "Señal de stop" 4 "Señal de ceda el paso" 5 "Solo marcas viales" 6 "Paso para peatones" 7 "Otra señal" 8 "Ninguna (solo norma)"
<i>superf</i>	Superficie	1 "Seca y limpia" 2 "Umbría" 3 "Mojada" 4 "Helada" 5 "Nevada" 6 "Barrillo" 7 "Gravilla suelta" 8 "Aceite" 9 "Otra"

CODIFICACIÓN	ETIQUETA	VALORES
<i>lumin</i>	Luminosidad	1 "Pleno día" 2 "Crepúsculo" 3 "Iluminación suficiente (noche)" 4 "Iluminación insuficiente (noche)" 5 "Sin iluminación (noche)"
<i>facatm</i>	Factores atmosféricos	1 "Buen tiempo" 2 "Niebla intensa" 3 "Niebla ligera" 4 "Lloviznando" 5 "Lluvia fuerte" 6 "Granizando" 7 "Nevando" 8 "Viento fuerte" 9 "Otro"
<i>visibil</i>	Visibilidad	1 "Edificios" 2 "Configuración del terreno" 3 "Vegetación" 4 "Factores atmosféricos" 5 "Deslumbramiento" 6 "Polvo o humo" 7 "Otra causa" 8 "Sin restricción"
<i>se_apel</i>	Señalización de peligro	1 "Señalización existente" 2 "Señalización inexistente" 3 "Señalización innecesaria (no hay peligro)"
<i>acera</i>	Acera	"S" "Si" "N" "No"
<i>se_alver</i>	Visibilidad de la señalización vertical	1 "Buena" 2 "Deficiente" 3 "Nula"
<i>tipoacc</i>	Tipo de accidente	11 "Colisión frontal" 12 "Colisión fronto-lateral" 13 "Colisión lateral" 14 "Colisión por alcance" 15 "Colisión múltiple o en caravana" 21 "Colisión con vehículo estacionado o averiado"

CODIFICACIÓN	ETIQUETA	VALORES
		22 "Colisión con valla de defensa" 23 "Colisión con barrera de paso a nivel" 24 "Colisión con otro objeto o material" 31 "Atropello a peatón sosteniendo bicicleta" 32 "Atropello a peatón reparando el vehículo" 33 "Atropello a peatón aislado o en grupo" 34 "Atropello a conductor de animales" 35 "Colisión con animal conducido o rebano" 36 "Colisión con animales sueltos" 41 "Vuelco en la calzada" 51 "Salida de la vía por la izda. y choque con árbol o poste" 52 "Salida de la vía por la izda. y choque con muro o edificio" 53 "Salida de la vía por la izda. y choque con cuneta o bordillo" 54 "Salida de la vía por la izda. y otro choque" 55 "Salida de la vía por la izda. con despeñamiento" 56 "Salida de la vía por la izda. con vuelco" 57 "Salida de la vía por la izda. en llano"

CODIFICACIÓN	ETIQUETA	VALORES
		58 "Salida de la vía por la izda. otra" 61 "Salida de la vía por la dcha. y choque con árbol o poste" 62 "Salida de la vía por la dcha. y choque con muro o edificio" 63 "Salida de la vía por la dcha. y choque con cuneta o bordillo" 64 "Salida de la vía por la dcha. y otro choque" 65 "Salida de la vía por la dcha. con despeñamiento" 66 "Salida de la vía por la dcha. con vuelco" 67 "Salida de la vía por la dcha. en llano" 68 "Salida de la vía por la dcha. otra" 71 "Otro"
<i>Denscir</i>	Densidad de la circulación	1 "Fluida" 2 "Densa" 3 "Congestionada"

Tabla 7: Variables incluidas en el fichero general de la base datos ARENA

Una vez conocidas todas las variables de la Base de Datos Arena, se seleccionarán aquellas que puedan ser de interés para el presente estudio. Además, se estudiará la posibilidad de incluir variables adicionales que, de estar recogidas en los ficheros de la DGT, puedan influir en las colisiones frontales.

CAPÍTULO 3 . OBTENCIÓN Y TOMA DE DATOS DE LOS CASOS Y LOS CONTROLES

3.1. CONSIDERACIONES PREVIAS A LA OBTENCIÓN DE LOS CASOS

Previamente a la obtención de los casos, se han de tener en cuenta dos consideraciones, que conciernen al período temporal considerado, y a la ubicación geográfica de los individuos incluidos en el estudio. Además, se describirá la Red de Carreteras de España, con objeto de conocer el tipo de vías que transcurren por la geografía española en función de su titularidad.

3.1.1. PERÍODO TEMPORAL

Como ha sido mencionado anteriormente, el presente Trabajo Fin de Grado se restringe al período temporal 2009-2013.

La utilización de datos correspondientes a estos años, en lugar de otros más recientes, se debe a la dificultad que se presentó a la hora de obtener datos de los años posteriores, a pesar de los esfuerzos de contactar con las administraciones pertinentes.

3.1.2. UBICACIÓN GEOGRÁFICA DE LOS INDIVIDUOS INCLUIDOS EN EL ESTUDIO

Dado que el número de colisiones frontales que se han producido en el período temporal 2009-2013 en España es demasiado grande como la realización de un estudio empleando la metodología “caso-control”, en el que se requiere tomar datos adicionales para los casos, y tomar todos los datos para los controles, se decidió establecer como banco de pruebas la Comunidad Autónoma de Castilla y León.

Por tanto, el total de casos se redujo de más de 11000 casos en toda España a 697 en Castilla y León, los cuales representan una muestra del total de individuos afectados lo suficientemente grande como para obtener resultados de interés.

3.1.3. RED DE CARRETERAS DE ESPAÑA

Las carreteras interurbanas que discurren por Castilla y León pueden pertenecer a tres administraciones oficiales. En función de ello, se distinguen carreteras de titularidad estatal, autonómica y provincial.

Titularidad	Red de Carreteras del estado	Red de Carreteras de la Junta	Red de Carreteras de las diputaciones estatales	Total
Longitud de la Red (Km)	5233	11287	15929	32449

Tabla 8 Longitud de la Red de Carreteras de Castilla Y León [25]

RED DE CARRETERAS DEL ESTADO

Son todas aquellas carreteras pertenecientes al Ministerio de Fomento (red de carreteras de titularidad estatal). En Castilla y transcurren 5.233 Kilómetros de este tipo de vía, que se reparten del siguiente modo [25] :

Carretera convencional		3614 Km
Carretera de gran capacidad	Autopistas de peaje	282 Km
	Autovías y autopistas libre	1246 Km
	Carretera desdoblada	91 Km

Tabla 9. Distribución de la RCE en Castilla y León.

RED DE CARRETERAS DE LA JUNTA [25]

La red de titularidad autonómica presenta dos subcategorías:

- **Red básica**, la cual proporciona soporte para el tráfico de largo recorrido, así como la comunicación con los principales municipios de la Comunidad Autónoma.
- **Red Complementaria**, constituida por las carreteras de la Red Regional no catalogadas como básicas. Por ello, es una red de gran longitud y con tráfico muy desiguales. A su vez, la red complementaria se divide en la red complementaria preferente, que comunica los núcleos comarcales; y la red complementaria local, para el tráfico de corto recorrido.

RED DE CARRETERAS DE LAS DIPUTACIONES PROVINCIALES

Las carreteras gestionadas por las Diputaciones Provinciales, se caracterizan por su clara función de accesibilidad a los municipios de cada provincia, comunicando los diferentes municipios con las redes de movilidad de la región.

La longitud de red competencia de las diputaciones alcanza los 15.929 km. En el caso particular de Castilla y León, el conjunto de esta red tiene una Intensidad Media Diaria (IMD) inferior a 180 vehículos/día. [25]

3.2. SELECCIÓN DE LOS CASOS

A partir de los datos recogidos para cada uno de los accidentes registrados en la Base de Datos Arena, se realiza una selección para obtener los casos que se incluirán en el estudio. Utilizando el software SPSS, se realiza un filtrado de casos, imponiendo las siguientes condiciones:

ZONA

Se seleccionarán aquellos accidentes ubicados en carreteras interurbanas. Por tanto, se impondrá la condición de que la variable “zona” se corresponda con una carretera, que son vías interurbanas fuera de población; o con una variante, las cuáles se corresponden con carreteras de circunvalación, ronda y vías similares exteriores a una población.

- Zona =1 (“carretera”)
- Zona=4 (“variante”)

MEDIANA

Las colisiones frontales ocurren en ubicaciones en las que no existe mediana de separación entre los dos carriles, por lo que seleccionamos aquellos casos en lo que no se encuentren presente medianas de separación de carriles.

- Mediana=“NO”

RED

Se estudiarán aquellos casos que ocurren en carreteras interurbanas de una calzada con un carril para cada sentido. Por tanto, se incluyen las carreteras

del ministerio de fomento, de la Junta de Castilla y León y de las Diputaciones Provincial.

- Red=1 (“estatal”)
- Red=2 (“autonómica”)
- Red =3 (“provinciales”)

PROVINCIA

El estudio se limita a la comunidad Autónoma de Castilla y León, por lo que seleccionamos todos los casos que hayan ocurrida en alguna de las provincias pertenecientes a dicha comunidad:

- Provincia=5 (“Ávila”)
- Provincia=9 (“Burgos”)
- Provincia=24 (“León”)
- Provincia=34 (“Palencia”)
- Provincia=37 (“Salamanca”)
- provincia=40 (“Segovia”)
- provincia=42 (“Soria”)
- Provincia=47 (“Valladolid”)
- Provincia=49 (“Zamora”)

TIPO DE ACCIDENTE

Se pretende estudiar todos los casos de accidente frontales. Además, muchos accidentes frontales, son registrados como accidentes frontolaterales en la base de datos, como consecuencia del giro brusco de uno de los vehículos implicados, como acto reflejo ante la inminente colisión frontal. Por tanto:

- Tipo de accidente=”11” (frontal)
- Tipo de accidente=”12”(frontolateral)

FUERA INTERSECCIÓN

Dado que no solo hemos seleccionado accidentes frontales, si no que también los frontolaterales, se requiere excluir todos aquellos accidentes frontolaterales que ocurren como consecuencia de la salida de un vehículo de una intersección, y no se deben a un acto reflejo para evitar una colisión frontal. Por ello, la variable “fuera de intersección” debe tomar cualquier valor distinto de 0.

- Fuera de intersección=1 (rectas)
- Fuera de intersección=2 (curva suave)
- Fuera de intersección=3 (curva fuerte sin señalizar)
- Fuera de intersección=4 (curva fuerte con señal y sin velocidad señalizada)
- Fuera de intersección=5 (curva fuerte con señal y velocidad señalizada)

TIPO DE VÍA

Se seleccionan exclusivamente las vías convencionales y las vías convencionales con carril lento, pues nos interesan aquellos accidentes que ocurren en carreteras convencionales de doble sentido, o de doble sentido con carril lento. De este modo se descartan otro tipo de vías, como autovías o autopistas, irrelevantes para el presente estudio.

- Tipo de vía=4 (“vías convencionales”)
- Tipo de vía=5 (“vías convencionales con carril lento”)

Una vez establecidas todas las condiciones que han de cumplir los casos, escribimos la sentencia lógica, y la implementaremos a través del seleccionador de casos del SPSS.

```
(red=1 | red=2 | red=3) & (tipovia=4 | tipovia=5) &
(tipoacc=11 | tipoacc=12) & (mediana="N") & (zona=1 |
zona=4) & (fueraint=1 | fueraint=2 | fueraint=3 | fueraint=4 |
fueraint=5) & (provin=24 | provin=34 | provin=9 | provin=49 |
provin=47 | provin=37 | provin=5 | provin=40 | provin=42) &
(titulvia=1 | titulvia=2)
```

De este modo se obtiene un fichero, en el que se incluyen 697 casos que cumplen las características impuestas entre los años 2009 y 2013.

3.3. CONSIDERACIONES POSTERIORES A LA OBTENCIÓN DE LOS CASOS

CONSIDERACIONES DE APLICACIÓN DEL ESTUDIO EN FUNCIÓN DEL TIPO DE VÍA

Como ha sido explicado anteriormente, la finalidad del presente estudio es realizar recomendaciones sobre las ubicaciones más adecuadas para que en ellas se realice la conversión de carreteras convencionales con un carril para cada sentido, o carreteras “1+1”, en carreteras “2+1”. Por lo tanto, es importante considerar la anchura de la calzada de los tramos de carretera en los que se va a implantar esta medida.

Como se ha indicado en la tabla 7: Variables incluidas en el fichero general de la base de datos ARENA, la variable “anchura de calzada” presenta las siguientes categorías:

1. Ancho de calzada inferior a 5,99 metros
2. Ancho de calzada mayor que 5,99 metros, y menor que 6,99 metros
3. Ancho de calzada superior a 6,99 metros

Es evidente que anchos de calzada incluidos en la tercera categoría facilitan la segregación de carriles, pues se disminuyen las modificaciones y el coste de las obras a realizar en la vía. Por el contrario, anchos de calzada incluidos en la primera categoría dificultarían considerablemente el proceso.

Por tanto, sería conveniente elegir localizaciones con un ancho superior a 6’99 metros. De este modo, se consigue realizar una actuación de coste medio, pero de alta prioridad, tal y como ha sido explicado en los objetivos del Trabajo Fin de Grado.

Con objetivo de estudiar la distribución de la variable ancho de calzada en los casos obtenidos previamente, se realiza una tabla cruzada con las variables “Titularidad de la Vía” y “Anchura de calzada”. Se obtienen los siguientes resultados:

		Anchura de calzada			Total
		Menos de 5,99 metros	Entre 6 y 6,99 metros	De 7 metros o más	
Titularidad de la vía	Estatal	3	35	254	292
	Autonómica	36	138	122	296
	Provincial	39	57	13	109
Total		78	230	389	697

Tabla 10: Tabla cruzada Titularidad de la Vías * Anchura de la calzada

A priori, se puede observar que el número de casos ocurridos en carreteras provinciales que presentan una anchura mayor que 7 metros son tan solo 13. En el siguiente gráfico se observa la distribución de la variable “Anchura de carretera” para los casos ocurridos en carreteras provinciales.

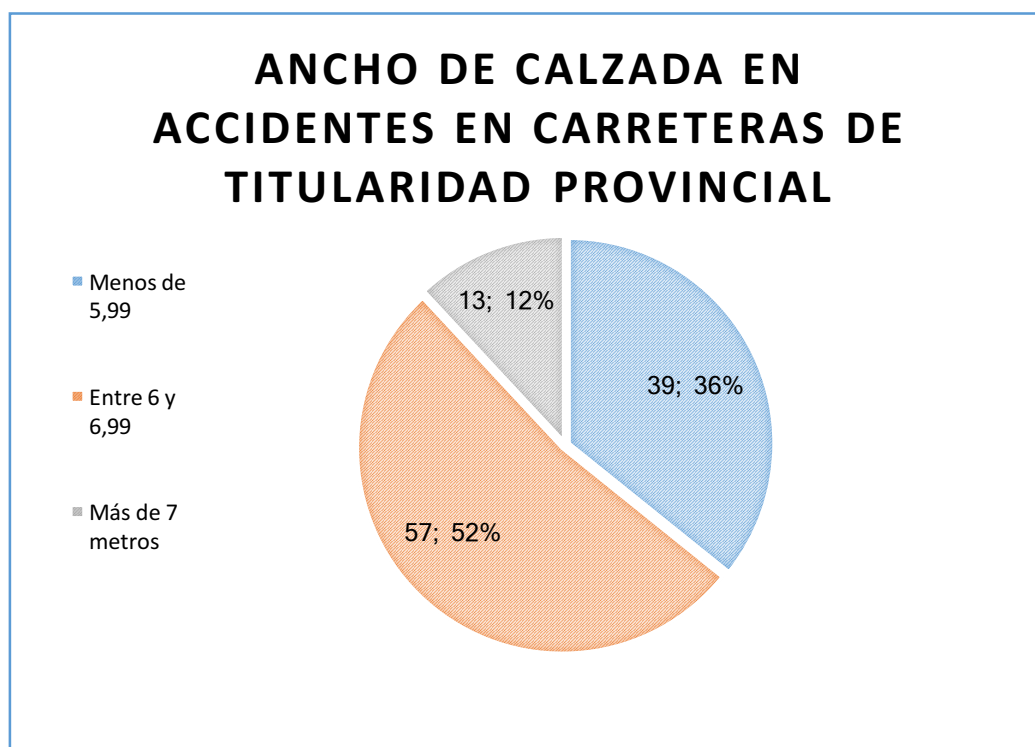


Ilustración 10: Porcentaje de accidentes ocurridos en carreteras de diferente anchura.

Realizando un análisis de los datos, se puede observar en la ilustración anterior que casi el 90% de los casos que ocurren en las carreteras de titularidad de vía provincial, presentan un valor de la variable “anchura de calzada” menor o igual que 6,99.

Dado que, como ha sido justificado, las carreteras más susceptibles en convertirse en carreteras “2+1” son aquellas con una anchura superior a 6,99, se descartarán los 109 casos ocurridos en las carreras de titularidad Provincial.

Además, las carreteras Provinciales, pertenecientes a las diputaciones, son poco transitadas, y resulta casi imposible tomar datos geométricos de algunas de ellas.

Por tanto, finalmente se obtienen 588 casos totales en Castilla y León que cumplen las condiciones impuestas.

3.4. OBTENCIÓN DE LOS CONTROLES

La obtención de los controles se realizó mediante un muestreo aleatorio. Como herramienta de apoyo para su realización, y para identificar a qué carretera y punto kilométrico pertenecían los datos aleatorios obtenidos, fue creado un archivo Excel, cuyo funcionamiento y utilidad serán explicados en los siguientes apartados.

3.4.1. CREACIÓN DEL LISTADO DE CARRETERAS

En primer lugar, se recopiló la información requerida sobre todas aquellas carreteras convencionales de titularidad autonómica y estatal que transcurren por la Comunidad Autónoma de Castilla y León. Para ello se utilizaron los catálogos proporcionados por los organismos titulares de las carreteras (Junta de Castilla y León y Ministerio de Fomento). [25][26]

Finalmente, se elaboró un listado, en el que se incluyeron todas las carreteras convencionales interurbanas de titularidad autonómica y estatal de la provincia de Castilla y León, así como la longitud de cada una de ellas.

Se realizó una tabla, con los mismos campos que se incluyen en Tabla 11: modelo excel para la obtención de los controles.

Nombre carretera	$PK_{inicial}$	PK_{final}	Acumulado
AV-104	0	9,697	9,697
...
N-631	87,11	425,08	16459,533

Tabla 11: modelo excel para la obtención de los controles

El significado de cada uno de estos campos es el siguiente:

- **Nombre de carretera:** se anota el nombre de la carretera (*ejemplo: N-630, CL-650*)
- **$PK_{inicial}$:** punto kilométrico de inicio de la carretera en la Comunidad Autónoma de Castilla y León. Podrá ser distinto de 0 en las carreteras del Ministerio de Fomento (nacionales). Se expresa en kilómetros con una posición decimal, dado que en los datos de los catálogos del Ministerio de Fomento y de la Junta de Castilla y León vienen expresados

de este modo, y en la base de datos ARENA también se recoge el punto kilométrico con una única posición decimal.

- **PK_{final}** : punto kilométrico final de la carretera en la Comunidad Autónoma (podrá no coincidir con la longitud total de la carretera si la misma transcurre por distintas Comunidades Autónomas). Por las mismas razones que han sido expuestas en la explicación del campo “Punto Kilométrico Inicial”, se expresa con una posición decimal.
- **Acumulado**: indica la longitud acumulada de todas las carreteras anteriores incluidas en el listado, incluyendo la longitud de la carretera de la propia fila, medida en kilómetros. Por tanto, se calcula mediante la suma de la longitud de carretera de una determinada fila, y el campo “acumulado” de la fila anterior. La celda “acumulado” de la última fila del listado representa la longitud total de las carreteras incluidas en el estudio.

3.4.2. CREACIÓN DE LA DISTRIBUCIÓN UNIFORME

Una vez se ha obtenido el listado de carreteras de la provincia de Castilla y León, con sus respectivas longitudes, se procede a la realización del muestro aleatorio, para la obtención de los controles.

Para ello, se crea una distribución Uniforme, $U(a,b)$. Dado que los casos se encontraban localizados mediante puntos kilométricos con una posición decimal, en la variable uniforme creada se utilizaron como unidades los hectómetros. De este modo, al realizar la conversión de hectómetros a kilómetros, los controles también presentarán una posición decimal.

Los extremos “a” y “b” de la Distribución Uniforme creada son:

- “a”: el extremo inferior de la distribución uniforme será el hectómetro de inicio del listado de carreteras.

$$a=0 \text{ Hm.}$$

- “b”: el extremo superior de la uniforme será el hectómetro final de la longitud total de la carretera. Se obtiene a través de la siguiente expresión:

$$b= (\text{acumulado de la última fila} *10) \text{ Hm}$$

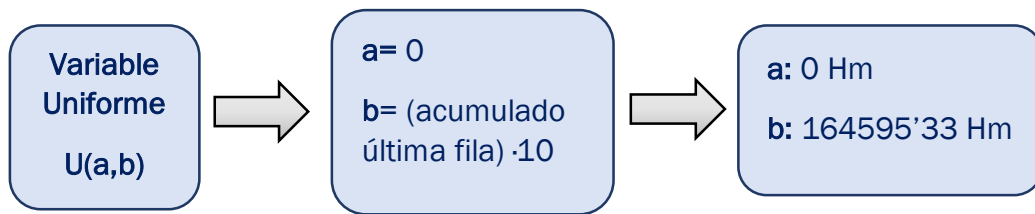


Ilustración 11. Creación de la distribución Uniforme

Para la creación de la uniforme, en Excel se ha utilizado la función ALEATORIO. La representación de la Uniforme creada para obtener los controles se muestra a continuación.

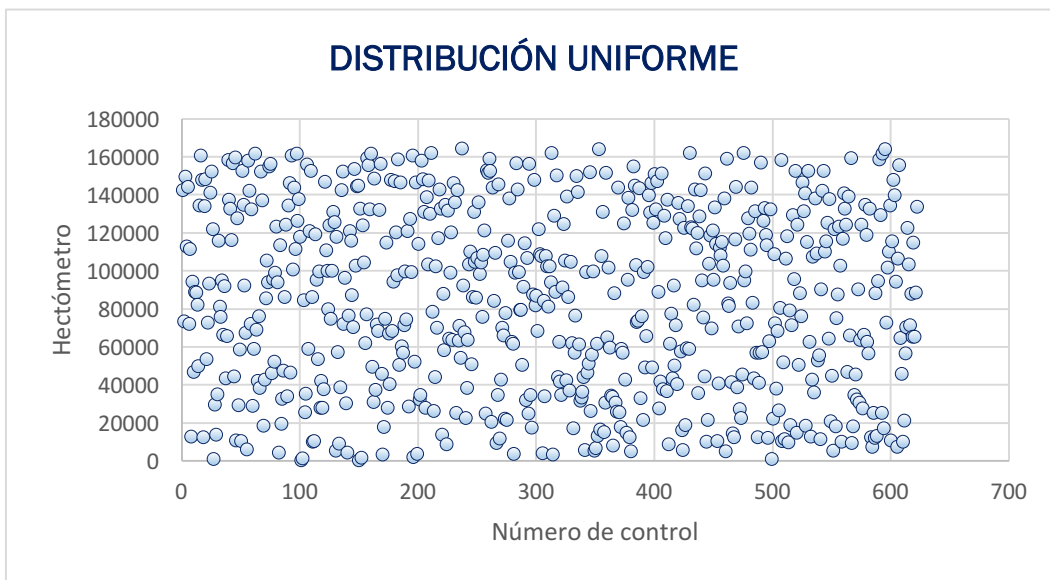


Ilustración 12. Representación de la Distribución Uniforme Creada

La distribución uniforme creada se caracteriza por que cada uno de los n valores de su rango tiene la misma probabilidad [27]

$$f(x_i) = \frac{1}{n}$$

De este modo, la probabilidad de que una determinada ubicación incluida en el listado de carreteras se obtenga como control depende únicamente de la amplitud del intervalo, y no de su posición. Por tanto, la probabilidad se reparte uniformemente.

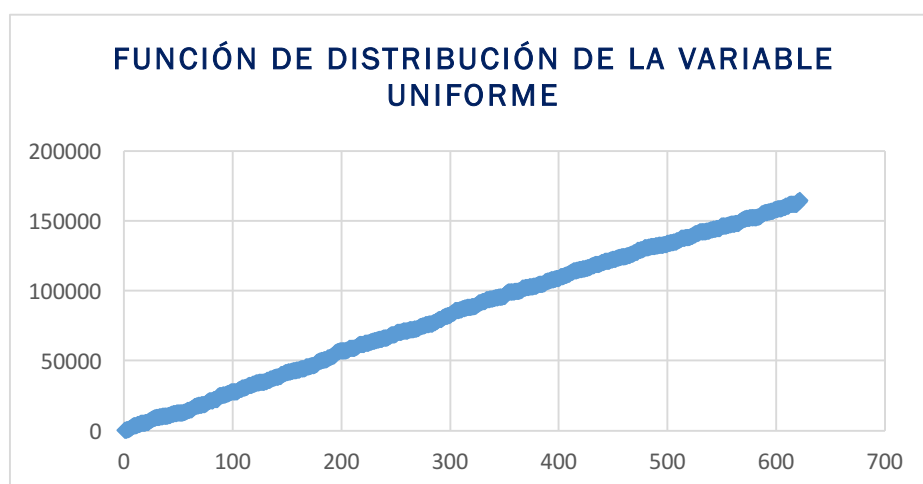


Ilustración 13 Función de distribución de la variable uniforme creada

Como medida de seguridad, debido a que no todos los controles obtenidos tienen por qué cumplir los criterios necesarios para ser incluidos en el estudio, se decidió crear inicialmente más de 588 controles.

De este modo, cuando un control se corresponda con una ubicación que no cumpla los criterios necesarios, podrá ser descartada, sin necesidad de modificar los controles obtenidos anteriormente.

Dado que el hecho de encontrar un control que no se adecúe al estudio es aleatorio, este último paso no introduce ningún sesgo.

3.4.3. OBTENCIÓN DE LAS UBICACIONES DE LOS CONTROLES

Una vez creada la distribución uniforme, se dispone de un listado con todos los controles. Se crea una tabla, en la que se incluye el número de control, y el hectómetro en el que se encuentra, sobre la longitud total de las carreteras que transcurren por la Comunidad Autónoma de Castilla y León, en hectómetros.

Para agilizar su posterior localización, los datos se filtran en Excel, obteniendo como resultado de este proceso los datos ordenados de forma ascendente. Este paso se realiza con el número inicial de controles, y en caso de descartar alguno de ellos, se coge el siguiente control que quedó fuera de los 588 controles incluidos en el filtrado inicial.

Dado que inicialmente la longitud de las carreteras se encontraba en kilómetros, se obtiene el kilómetro de las ubicaciones correspondientes a los controles.

Control	Hectómetro
1	371
2	444
3	892
4	1325
...	...
588	

$\frac{Hm}{10} \rightarrow km$

Control	Kilómetro
1	37,1
2	44,4
3	89,2
4	132,5
...	...
588	

Ilustración 14. Obtención del Kilómetro asociado a cada control, sobre la longitud total de Carreteras de Castilla y León.

A continuación, se obtiene el nombre de la carretera y el punto kilométrico de la misma, para cada uno de los controles. Se realizan los siguientes pasos:

1. Encontrar el nombre de la carretera a la que pertenece el control.

Dado el kilómetro correspondiente a un determinado control sobre la longitud total de las carreteras que transcurren por la provincia de Castilla y León, se acude al listado de carreteras. (Tabla 11: modelo excel para la obtención de los controles)

En la columna “longitud acumulada en kilómetros”, se buscan dos celdas adyacentes, tales que el kilómetro del control sea mayor que la longitud acumulada de la primera celda, pero menor que la longitud acumulada de la segunda.

Provincia	Nombre carretera	Longitud en Km	Longitud acumulada en Km
Ávila	AV-104	9,697	9,697
	AV-114	17,328	27,025
	AV-120	19,657	46,682
	AV-322	2,5	49,182

Control	Kilómetro
1	37,1

37,1€ [27.025,46.682)

Ilustración 15. Localización de la carretera y el punto kilométrico (1)

2. Una vez realizado el paso anterior, se conoce que la carretera en la que se ubica ese control es la carretera con la que se corresponde la celda inferior.

Provincia	Nombre carretera	Longitud en Km	Longitud acumulada en Km
Ávila	AV-104	9,697	9,697
	AV-114	17,328	27,025
	AV-120	19,657	46,682
	AV-322	2,5	49,182

Nombre de la carretera:
Av-120

Ilustración 16. Localización de la carretera y el punto kilométrico (2)

3. Una vez conocida la carretera, se obtiene el punto kilométrico de la misma en la que se encuentra ubicado el control. Para ello, se calcula la diferencia entre el kilómetro del control sobre la longitud total en kilómetros y la longitud acumulada en km de la carretera anterior a la de la fila a la que pertenece.

Provincia	Nombre carretera	Longitud en Km	Longitud acumulada en Km
Ávila	104	9,697	9,697
	114	17,328	27,025
	120	19,657	46,682
	322	2,5	49,182

$$P_{km} = 37.1 - 27.025 = 10.075$$

AV-120, Pkm= 10.1

Ilustración 17. Localización de la carretera y punto kilométrico (3)

Se repite el mismo procedimiento para cada uno de los controles, descartando aquellos que no cumplan las mismas condiciones que fueron impuestas para los casos, pues tal y como se mencionó en la explicación de la metodología “caso-control” (capítulo 1), casos y controles se pueden entender como subgrupos pertenecientes a la misma población.

Evidentemente, en la metodología caso-control, si una determinada ubicación es un caso, no podrá ser utilizada como control. Por tanto, es necesario revisar todos los controles obtenidos, y asegurarse de que ninguna de las localizaciones obtenidas se corresponde con un caso, no solo entendiendo iguales aquellas ubicaciones de la misma carretera con el mismo punto

kilométrico. Los individuos del estudio son los tramos (curvas o rectas), por lo que un mismo individuo (es decir una misma curva o recta con ubicación muy cercana) no podrán considerarse controles si en dicha ubicación ha ocurrido algún caso.

Del mismo modo, es importante comprobar que ninguno de los controles se repite, dado que cada control ha de ser único.

Finalmente, se crea la variable dicotómica “caso-control”, la cual actuará como variable respuesta.

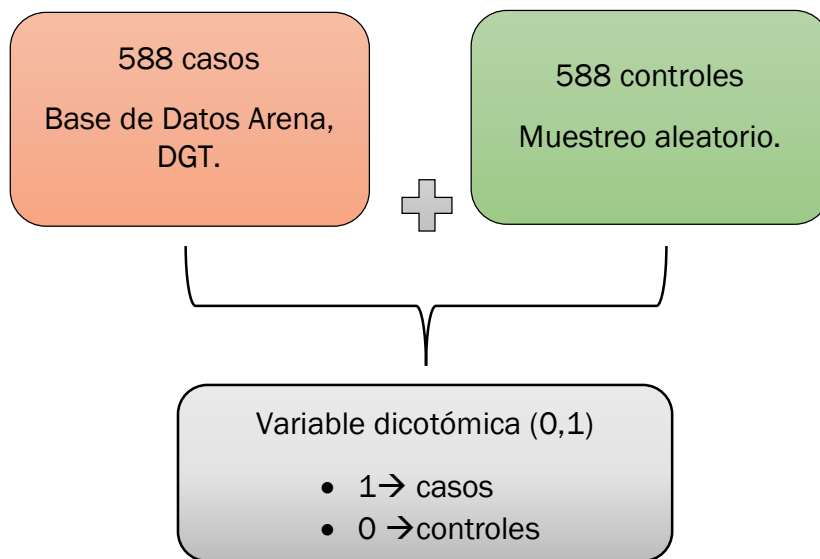


Ilustración 18. Esquema definitivo de casos y controles

3.5. TOMA DE DATOS PARA LOS CASOS Y LOS CONTROLES

Una vez se ha obtenido la carretera y el punto kilométrico de todos los controles, se dispone de un fichero, en el que aparecen los casos y los controles, la carretera en la que se encuentra localizados, y el punto kilométrico exacto de la misma en la que se ubican.

Mediante las herramientas software Iberpix 4 [28] y Google Maps, se procede a la localización de los controles, con la finalidad de obtener las coordenadas geográficas de los mismos, y agilizar el proceso de recogida de datos una vez se hayan decidido las variables a incluir en el estudio (capítulo 4).

De los casos se han de obtener los datos de las variables adicionales incluidas en el modelo. Sin embargo, de los controles no se dispone información, por lo

que se recogerán datos de todas las variables de la base de datos ARENA que se incluyan en el estudio, así como los datos de las variables adicionales.

Generalmente, el proceso de localización consta de dos pasos:

1. Mediante el software “Iberpix 4” [28], creado por el Ministerio de Fomento, se obtiene la localización aproximada del punto kilométrico.

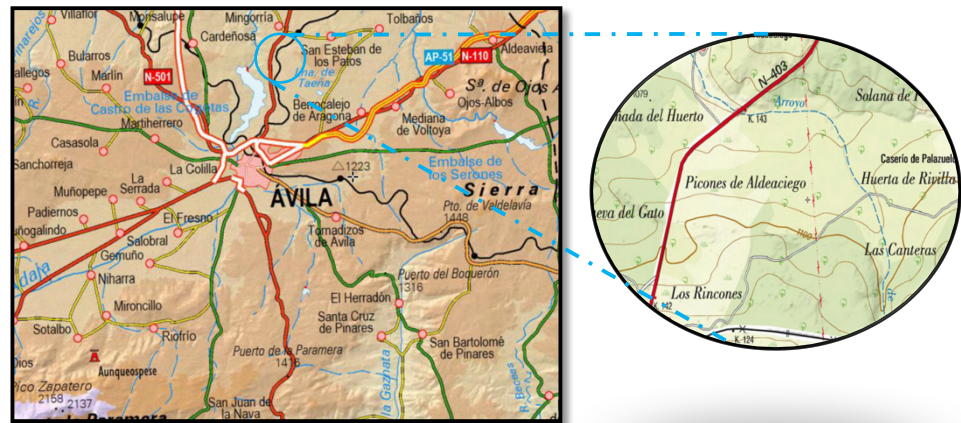


Ilustración 19. Localización Punto Kilométrico mediante Iberpix 4

2. Utilizando el modo “Street view” de la aplicación online Google Maps, se localiza la ubicación exacta del punto kilométrico. En primer lugar, se localiza la señal del punto kilométrico buscado. A continuación, utilizando la herramienta “regla”, se miden los metros a los que se encuentra un caso o un control del punto kilométrico.



Ilustración 20. Localización del punto kilométrico exacto, hito kilométrico+ herramienta regla

CAPÍTULO 4. VARIABLES DEL MODELO

En el presente capítulo, se pretende describir detalladamente las variables que se incluirán posteriormente en el modelo, con el objetivo de definir las características geométricas de diseño de la vía, o de tráfico de la vía a la que se refiere cada una de ellas, y sus categorías.

La Base de Datos Arena recoge muchas de las características geométricas de las ubicaciones en las que ocurren los accidentes. Sin embargo, por decisión previa a la realización del estudio, se recogerán algunas variables a mayores de las incluidas inicialmente en la Base de Datos, que podrían ser factores de riesgo que afectan a la producción de colisiones frontales.

Por tanto, las variables del estudio se pueden clasificar en variables incluidas en la base de datos ARENA y variables adicionales.

- Variables incluidas en la Base de Datos Arena: son variables que se encuentra recogidas en la Base de Datos de la DGT, y que se considera que podrían afectar a las colisiones frontales.
- Variables Adicionales: son variables geométricas o de tráfico que no se encuentran recogidas en la base de datos de la DGT, pero que se cree que podrían afectar a la ocurrencia de las colisiones frontales.

4.1. VARIABLES INCLUIDAS EN LA BASE DA DATOS ARENA

ANCHURA DE LA CALZADA

La calzada es la parte de una carretera dispuesta para la circulación de vehículos. Se trata de una variable categórica, cuyas categorías son, de acuerdo a la Tabla 7: Variables incluidas en el fichero general de la base datos ARENA, las siguientes:

- Carreteras de Menos de 5,99 metros
- Carreteras de 5,99 metros a 6,99 metros.
- Carreteras de más de 7 metros.

ANCHURA DE CARRIL

Los carriles son cada una de las divisiones de una calzada por las que circulan los vehículos. A pesar de que inicialmente se encontraba categorizada tal y como se muestra en la tabla 7: Variables incluidas en el fichero general de la

base datos ARENA, se decidió cambiar el orden de las categorías para facilitar la interpretación de los resultados.

1. Ancho de carril menor 3,25 metros.
2. Ancho de carril entre 3,25 metros y 3,75 metros.
3. Ancho de carril mayor que 3,75 metros.

ANCHURA DE ARCÉN

El arcén es el margen lateral de una carretera. Se considera la medida del arcén del lado de la calzada en el que se ha producido el accidente. Como se ha mencionado anteriormente, se encuentra categorizado del siguiente modo:

1. Inexistente o impracticable.
2. Menor de 1.5 metros.
3. Entre 1,5 y 2,5.
4. Mayor de 2,5 metros

BARRERA DE SEGURIDAD

La barrera de seguridad se trata de un sistema de contención de vehículos situado en unos o los dos márgenes de la carretera. Se trata de una variable que distingue la presencia o ausencia de este elemento, por lo que toma los valores:

1. Sí
2. No

HITOS DE ARISTA

Los hitos de arista son elementos reflectantes situados en los laterales de la plataforma de la carretera, cuyo objetivo es balizar los márgenes de la plataforma de la carretera, así como marcar el hectómetro de la misma. De nuevo, se trata de una variable que indica la presencia o ausencia de estos elementos. Por tanto, presenta las siguientes categorías:

1. Sí
2. No

ACERA

La acera es la parte lateral de una carretera dispuesta para la circulación de peatones. De nuevo, es una variable categórica cuyo propósito es indicar la presencia o ausencia de acera.

1. Sí
2. No

NÚMERO DE CARRILES

El número de carriles hace referencia al número de carriles destinados al tránsito de vehículos, en cada uno de los sentidos de circulación. Las categorías que presenta inicialmente en las carreteras incluidas en el estudio son las siguientes:

- 11: un carril en cada sentido
- 12: un carril en el sentido ascendente de la numeración kilométrica, y dos en el contrario
- 21: dos carriles en el sentido ascendente de la numeración kilométrica, y uno en el contrario
- 22: dos carriles en cada sentido de la circulación.

Se consideró reordenar las categorías, con el objetivo de distinguir solamente si existía alguna diferencia entre que hubiese un solo carril para cada sentido de la circulación, o que algún sentido contase con más de un carril. Por tanto, las categorías finales con las que se trabajó fueron las siguientes:

- “Un carril en cada sentido de la circulación”
- “Más de un carril en al menos un sentido de la circulación”

4.2. VARIABLES ADICIONALES

RADIO DE CURVATURA

Para realizar el estudio, se consideró que la variable geométrica “radio de curvatura” podría tener influencia en las colisiones frontales.

Para la medida del radio se ha utilizado el siguiente método [29]

1. Medir en línea recta la cuerda “L”, desde el punto PC al punto PT.
2. Hallar el punto medio de la cuerda, “c”.

3. Trazar una perpendicular desde “c” a la huella, y obtener la longitud “M” de este segmento.

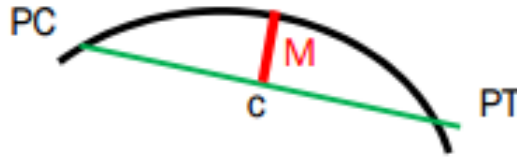


Ilustración 21. Esquema para la obtención del radio de curvatura

Por último, se obtendrá el radio de curvatura aplicando la siguiente fórmula.

$$R = \frac{L^2}{8 * M} + \frac{M}{2}$$

Ecuación 8: Obtención del radio de curvatura

Como ha sido mencionado en el capítulo anterior, se utilizará Google Maps para la medición de las variables. En el caso particular del radio de curvatura, se utiliza la herramienta “regla” de google para obtener las distancias necesarias (L y M).

CATEGORIZACIÓN DE LA VARIABLE “RADIO DE CURVATURA”

Inicialmente, el radio fue medido de manera continua, pero posteriormente se decidió categorizar la variable.

La elección de las categorías de la variable “Radio de Curvatura” se fundamentó en la Instrucción de Carreteras Norma 3.1-IC. [30]

De acuerdo a la norma anterior, las carreteras convencionales se nombran con una C (Carretera convencional), seguido de un número que indica la Velocidad de Proyecto, expresado en Kilómetros por hora.

Conociendo que las carreteras incluidos en el estudio serán del tipo C-100, C-90, C-80 o C-70, se puede intuir el orden de los radios que manejaremos en el estudio.

Carreteras Multicarril C-100 y Carreteras Convencionales C-100	$250 \leq R < 700$
	$700 \leq R < 5000$
	$5000 \leq R < 7500$
	$R \geq 7500$
Carreteras Multicarril C-90, C-80, C-70, C-50, C-40 y carreteras Convencionales C-90, C-80, C-70, C-60, C-50, C-40	$50 \leq R < 350$
	$350 \leq R < 2500$
	$2500 \leq R < 2500$
	$R \geq 2500$

Tabla 12 Radio Mínimos y máximos en función del tipo de carretera [27]

Por otra parte, la Orden Circular sobre sistemas de Contención de Vehículos [7] predispone los siguientes radios límite para la instalación de SCV en carretera:

		Radio (m)
Carretera con calzadas separadas		750
Carretera de calzada única	Arcén $\geq 1,5$	250
	Arcén $< 1,5$	200

Tabla 13. Sistemas de contención de vehículos para motocicletas [7]

Tomando en consideración toda la información mencionada anteriormente, se crearon las siguientes categorías:

1. Recta
2. Curva suave: radio mayor que 750 metros.
3. Curva intermedia: radio entre 750 metros y 250 metros.
4. Curva cerrada: radio menor que 250 metros.

INTENSIDAD MEDIA DIARIA (IMD)

La Intensidad Media Diaria muestra la fluencia del tráfico en los diferentes tramos de carreteras, utilizando como unidad de medida vehículos/día.

A pesar de que la IMD no se trata de una variable geométrica de diseño de la carretera, y no es posible actuar sobre ella, se decidió incluirla en los modelos

multivariante, dado que de esta manera se conseguiría valorar el efecto del resto de variables eliminando su efecto.

Además, el diseño de la geometría de las carreteras podrá estar realizado en función de la IMD, por lo que podría ser determinante en los resultados obtenidos.

Dado que esta variable no se encontraba incluida en la base de datos Arena, la IMD de los diferentes tramos de carretera se obtuvo utilizando los Mapas de Tráfico proporcionados por la Junta de Castilla y León o por el Ministerio de Fomento, en función de la titularidad de la carretera. [31][32]

Para recopilar estos datos, los organismos titulares han establecido estaciones en puntos estratégicos de los diferentes tramos de carretera. La función de estas estaciones de aforos es el conteo de los vehículos que circulan por el punto de la carretera en el que se encuentran situadas.

Para calcular el flujo de vehículos diario en un determinado tramo de carretera, dividen el flujo total anual, entre los 365 días que dispone un año. [33]

Tal y como se muestra en la figura, los distintos tramos de carretera se encuentran etiquetados mediante uno o varios cuadros, en los que aparece la IMD del tramo de carrera.

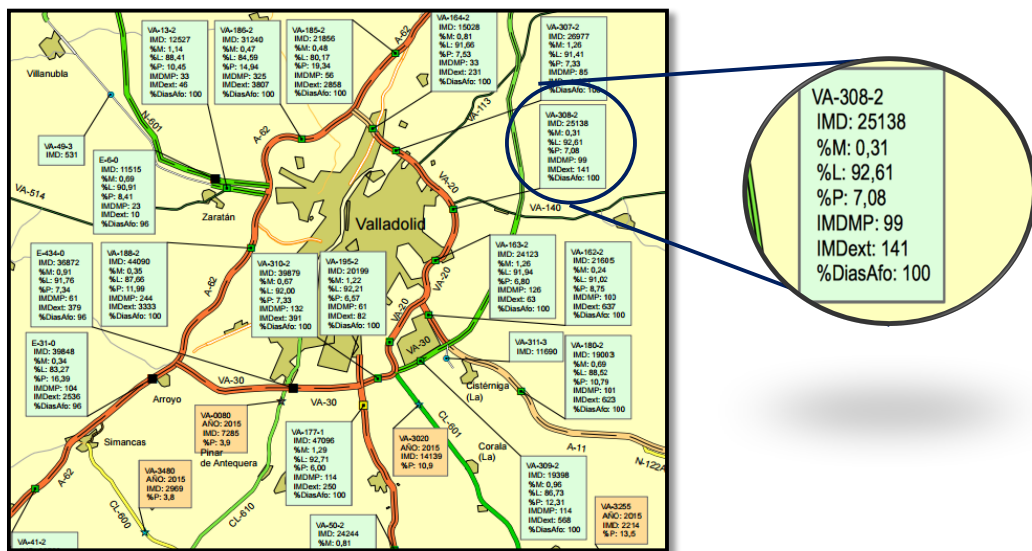


Ilustración 22. Mapa de Tráfico de la Red Carreteras del Estado

Los diferentes colores de los tramos de carretera representan la categoría en la que se encuentra incluida la Intensidad Media Diaria de dicho tramo de acuerdo a los criterios utilizados por el organismo titular de la vía. Por último, el tipo de icono de cada estación de aforamiento se refiere al tipo de estación de la que se trata.



Ilustración 23 leyenda mapa IMD del Ministerio de Fomento (provincia Valladolid)

CATEGORIZACIÓN DE LA INTENSIDAD MEDIA DIARIA

Para la codificación de la variable IMD, se pretendió crear categorías teniendo en cuenta los criterios de categorización utilizados por las administraciones titulares, de manera que se facilitase la aplicabilidad de los resultados posteriores.

El informe RACE sobre el nivel de riesgo de las carreteras del Ministerio de Fomento en función del IMD [34], utiliza las siguientes categorías para clasificar el riesgo de las carreteras en función de la IMD:

IMD
IMD <2000
20000<IMD<10000
10000<IMD<20000
20000<IMD<60000
60000<IMD<80000
IMD>80000

Tabla 14. Categorización de la IMD según el informe Race

Por otra parte, el plan de Carreteras de la Junta de Castilla y León [25], emplea una codificación diferente, dado que la Intensidad Media Diaria de estas carreteras es generalmente menor que la de las carreteras Nacionales, pertenecientes al Ministerio de Fomento.

Tabla 15 IMD de carreteras de titularidad Autonómica, aparece un resumen de la IMD de las diferentes clases de carreteras de la Junta de Castilla y León. Se puede apreciar que la IMD de las carreteras autonómicas se encuentra en el rango 500-3000 vehículos por día.

Tipo de Red	2001 Veh./día	2002 Veh./día	2003 Veh./día	2004 Veh./día	2005 Veh./día	2006 Veh./día
R. Básica	2169	2319	2433	2659	2664	2764
R.Complementaria Preferente	1006	1005	878	921	927	961
R. Complementaria Local	492	521	451	471	453	428
Total	1044	1086	1087	1141	1137	1164

Tabla 15 IMD de carreteras de titularidad Autonómica expresada en vehículos por día [25]

Además, en el mismo informe, aparece una clasificación de la Tabla 16 Clasificación IMD según la Junta de Castilla y León, en la que aparece la velocidad de proyecto en función de los IMD.

TIPO DE RED	IMD (vehículos/día)
Básica y Complementaria Preferente	<2000
	>2000
Complementaria Local	< 500
	De 500 a 2000
	>2000

Tabla 16 Clasificación IMD según la Junta de Castilla y León [25]

En la tabla anterior se puede apreciar que realizan una codificación utilizando como puntos de corte 500 vehículos al día y 2000 vehículos al día.

Para facilitar la posterior aplicabilidad de los resultados obtenidos en el posterior estudio, la variable IMD fue codificada de modo que las categorías resultantes fueran una simplificación de las categorías utilizadas por las Instituciones titulares de las carreteras.

Por tanto, las categorías elegidas fueron las siguientes:

Categoría	IMD (Vehículos por día)
1	$0 \leq IMD < 500$
2	$500 \leq IMD < 2000$
3	$2000 \leq IMD < 10000$
4	$IMD \geq 10000$

Tabla 17 Categorización de la variable IMD

Tal y como se puede apreciar, las categorías podrían ser englobadas dentro de las categorías utilizadas por la Junta de Castilla y León, y por el Ministerio de Fomento.

CAPÍTULO 5. REGRESIÓN LOGÍSTICA SIMPLE

En el presente capítulo, se pretende realizar un primer análisis de carácter univariante, con el objetivo de estudiar la influencia de las variables individuales en el evento “colisión frontal”, para así poder intuir como se comportarán en los modelos multivariantes, en presencia del resto de variables.

Para cada una de las variables incluidas en el estudio, se obtendrá la tabla cruzada entre la misma y la variable respuesta “caso-control”, lo cual nos permitirá comparar la diferencia entre los porcentajes de los casos y los controles para las diferentes categorías de una variable.

Este análisis tendrá una doble utilidad, ya que por una parte nos permitirá intuir qué condiciones geométricas de diseño de la vía, así como de tráfico de la misma, se comportan como factores de riesgo cuando actúan de manera independiente, y nos permitirá determinar las categorías de referencia que se seleccionarán en las regresiones logísticas simples y multivariantes posteriores.

Por último, se incluye en este capítulo la regresión logística simple para cada una de las variables que se incluirán en los modelos finales. Esto permitirá determinar el tipo de asociación estadística entre cada variable individual y las colisiones frontales, en caso de que exista. Adicionalmente, se obtendrán los Odds-ratio y los Intervalos de Confianza, que nos permitirán determinar de modo más preciso las asociaciones existentes, y nos indicarán si dicha asociación es significativa.

Tanto en la regresión logística simple, como en la regresión logística multivariante, se diferenciarán, siempre que la variable lo permita, tres modelos distintos:

- **Modelo completo**, incluyendo todas las ubicaciones, curvas y rectas.
- **Modelo “solo curvas”** seleccionando las ubicaciones que se corresponden con curvas.
- **Modelo “solo rectas”** seleccionando las ubicaciones que se corresponden con rectas.

La distinción de estos tres modelos permitirá obtener una comprensión más profunda de la influencia de las variables, pues en el análisis del modelo completo se obtendrán conclusiones aplicables a todos los casos, mientras que los modelos “solo curvas” y modelo “solo rectas” nos permitirán determinar si existen diferencias entre ambos tipos de ubicaciones.

5.1. ANCHO DE CALZADA

5.1.1. TABLA CRUZADA ANCHO DE CALZADA*CASOCONTROL

	Control	Caso	Total
Menos de 5,99 metros	199 33,8 %	39 6,7%	238
Entre 6 y 6,99 metros	225 38,3 %	173 29,4%	398
De 7 metros o más	164 27,9%	376 63,9 %	540
Total	588	588	1176

Tabla 18 Cruce de las variables "Ancho de Calzada" y "Caso-control"

De acuerdo con la Tabla 18 Cruce de las variables "Ancho de Calzada" y "Caso-control", se intuye que las calzadas cuya anchura es inferior a 5,99 metros son las que menos influyen en las colisiones frontales, pues el número de controles es considerablemente superior al de casos. Sin embargo, cuando el ancho de calzada es "7 metros o más", el número de casos supera al de controles (164 frente a 376), por lo que este nivel parece ser el que más influye en las colisiones frontales de acuerdo al estudio descriptivo.

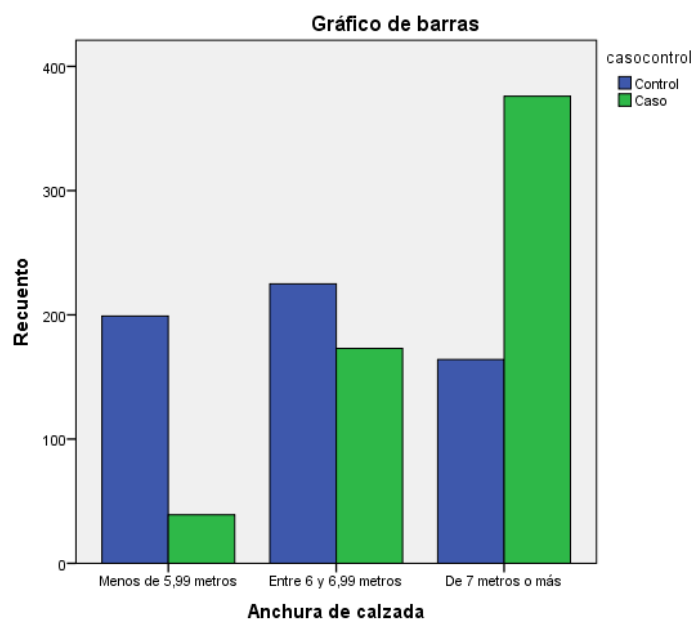


Ilustración 24: Gráfico de barras de la distribución de la variable "Anchura de Calzada" para los casos y los controles

En el gráfico de barras anterior se puede apreciar claramente la diferencia en la distribución de la variable “anchura de calzada” en los casos y en los controles. Visualmente, se puede observar una clara tendencia de aumento de las colisiones frontales (casos), a medida que aumenta la anchura de la calzada.

Dado que para realizar las regresiones logísticas posteriores será necesario establecer como categoría de referencia aquella que se considere de menor riesgo, la referencia se situará en el primer nivel: “*ancho de calzada menor de 5,99 m*”.

5.1.2. ODDS-RATIO PARA LA VARIABLE “ANCHO DE CALZADA”

Se calculará el Odds-ratio y los Intervalos de Confianza al 95% para los tres modelos que se explicaron al inicio del capítulo. De este modo, se podrá apreciar la fuerza de la asociación estadística entre la variable “Ancho de calzada” y la variable “Caso-Control”. Para ello, se realizará una regresión logística binaria, en la que la variable dependiente es “caso-control”, mientras que la variable independiente es “ancho de calzada”.

Tal y como se indicó en el apartado anterior, se situará como categoría de referencia para los tres modelos la categoría “ancho de calzada menor de 5,99 metros”

RECTAS Y CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
$\leq 5,99$ m			170,948	2	,000			
5,99 m -6,99 m	1,367	,202	45,695	1	,000	3,923	2,640	5,831
$\geq 6,99$	2,459	,199	153,437	1	,000	11,699	7,927	17,264
Constante	-1,630	,175	86,612	1	,000	,196		

Tabla 19: Regresión logística para la Variable "Ancho de calzada", modelo completo

Dado que se trata de una variable de tipo categórico, se han creado tantas variables dummies como categorías posee la variable original menos una. Estas nuevas variables son dicotómicas, y poseen el siguiente significado:

- La primera variable creada “5,99 m- 6,99 m” es un variable dicotómica, cuyos posibles valores son:
 - “1” cuando el ancho de calzada mide entre 5,99 m – 6,99 m.
 - “0” en caso contrario.

Por tanto, el coeficiente β correspondiente a esta variable hace referencia a lo que aumenta el logaritmo de la variable respuesta cuando la variable “5,99 m – 6,99 m” tome el valor de 1, con respecto a cuando tome el valor de 0.

- La variable “ $\geq 6,99$ ” también es una variable dicotómica, que puede tomar los siguientes valores:
 - “1” cuando el ancho de calzada es mayor que 6,99 metros
 - “0” en caso contrario.

Por tanto, el término $\beta_i X_i$ correspondiente a la variable “ $\geq 6,99$ ” será distinto de 0 cuando el ancho de calzada sea mayor o igual que 6,99, y el coeficiente del término independiente (β) representará lo que aumenta el logaritmo de la variable respuesta cuando la variable “ancho de calzada” es mayor que 6,99 metros, con respecto a cuándo es menor que 5,99 metros.

De la Tabla 19: Regresión logística para la Variable "Ancho de calzada", modelo completo, se pueden extraer las siguientes conclusiones:

- La presencia de calzadas con anchura entre 5,99 metros y 6,99 metros influye en la ocurrencia del evento “colisión frontal” cuando actúa de manera independiente, comportándose como un factor de riesgo, dado que el Intervalo de Confianza calculado al 95% no incluye al 1. El OR toma el valor de 3,923, lo cual se traduce en un aumento del riesgo de ocurrencia de las colisiones frontales de casi 4 veces cuando la anchura de calzada se encuentra comprendida entre 5,99 y 6,99, en lugar de ser inferior a 5,99.
- La presencia de calzadas con anchura mayor que 6,99 metros influye en la ocurrencia del evento “colisión frontal” cuando actúa de manera independiente, siendo un factor de riesgo que aumenta el riesgo de ocurrencia del evento estudiado, dado que el Intervalo de Confianza calculado al 95% no incluye al 1. El OR toma el valor de 11,699, lo cual se traduce en un aumento del riesgo de ocurrencia de las colisiones frontales de casi 12 veces cuando la anchura de calzada es mayor de 6,99 en lugar de ser menor que 5,99.

El análisis anterior nos permite concluir que, a mayor ancho de calzada, mayor riesgo de ocurrencia de ocurrencia de colisiones frontales.

RECTAS								
	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
$\leq 5,99$ m			106,752	2	,000			
5,99 m -6,99 m	1,576	,305	26,733	1	,000	4,836	2,661	8,788
$\geq 6,99$	2,773	,297	87,187	1	,000	16,000	8,941	28,633
Constante	-2,144	,273	61,717	1	,000	,117		

Tabla 20: Regresión Logística para la variable "Ancho de calzada", solo rectas

Al aplicar la regresión logística a las ubicaciones rectas, obtenemos un resultado muy similar al obtenido al elegir el conjunto completo de datos:

- En las ubicaciones rectas, anchos de calzada entre 5,99 m y 6,99 m, aumentan el riesgo de que se produzcan colisiones frontales casi 5 veces (OR=4,836), respecto a calzadas con una anchura menor que 5,99, pues el p-valor toma el valor de 0, y el intervalo de confianza calculado no contiene al 1.
- Anchos de calzada superiores a 6,99 m aumentan 16 el riesgo de que se produzcan colisiones frontales respecto a calzadas estrechas en ubicaciones rectas, dado que el intervalo de confianza calculado no contiene al 1 y el p-valor toma el valor de 0.

Por tanto, en el modelo "solo rectas" también se puede afirmar que el ancho de calzada actúa como factor de riesgo, aumentando el riesgo de que se produzcan colisiones frontales cuanto mayor es su valor.

CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
$\leq 5,99$ m			64,390	2	,000			
5,99 m -6,99 m	1,244	,284	19,251	1	,000	3,471	1,991	6,052
$\geq 6,99$	2,245	,285	62,238	1	,000	9,444	5,406	16,497
Constante	-1,085	,236	21,101	1	,000	,338		

Tabla 21: Regresión logística para la variable Ancho de Calzada, solo curvas

En el caso de las curvas, podemos extraer conclusiones muy similares a las obtenidas en el “modelo completo” y “modelo solo rectas” para esta misma variable.

- En las ubicaciones curvas, anchos de calzada entre 5,99 y 6,99 metros aumentan el riesgo de que se produzcan colisiones frontales 3,5 veces (OR=3,471) respecto a anchos de calzada menores de 5,99, pues el intervalo de confianza calculado no contiene el 1, y el p-valor es 0.
- Anchos de calzada mayores que 6,99 aumentan aún más el riesgo de que se produzca el desenlace estudiado en ubicaciones curvas, alcanzando el OR el valor de 9,444. De nuevo, el resultado es significativo, dado que el p-valor es 0, y el intervalo de confianza no contiene el valor 1.

Se puede apreciar que la influencia de la variable independiente “ancho de calzada” es mayor en rectas que en curvas cuando actúa de manera independiente, pues los Odds-ratio son superiores, lo que indica un mayor aumento del riesgo para las ubicaciones rectas.

5.2. ANCHO DE CARRIL

5.2.1. TABLA CRUZADA ANCHURA DE CARRIL*CASOCONTROL

A continuación, se muestra la tabla cruzada “ancho de carril” y “Caso-control”.

	Caso-control		Total
	Control	Caso	
Menos de 3,25 metros	363 61,73%	145 24,66%	508 43,20%
De 3,25 a 3,75 metros	211 35,88%	427 72,62%	638 54,25%
Más de 3,75 metros	14 2,38%	16 2,72%	30 2,55%
Total	588	588	1176

Tabla 22: Tabla cruzada "ancho de carril" y "caso-control"

De acuerdo con los resultados de la Tabla 22: Tabla cruzada "ancho de carril" y "caso-control", y con el gráfico de barras que se muestra a continuación, se puede intuir que anchuras de carril intermedias (de 3,25 a 3,75 metros) podrían ser un factor de riesgo, pues el número de casos es considerablemente superior al de controles en esta categoría. Sin embargo, en anchuras de carril menores que 3,25 metros, ocurre lo contrario. Por ello, situaremos como referencia para la realización de los posteriores estudios la categoría "Menor o igual que 3,25 metros".

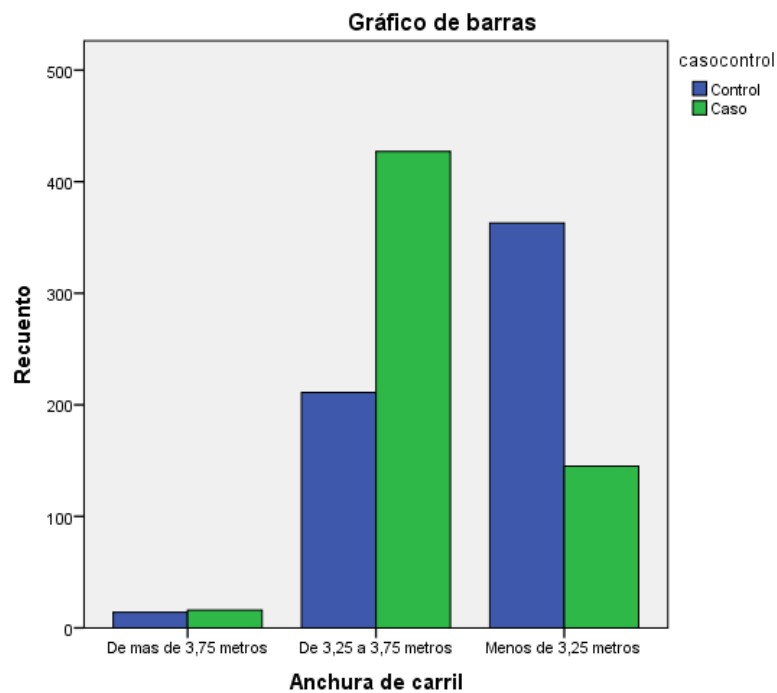


Ilustración 25 Gráfico de barras de la distribución de la variable anchura de carril en los casos y en los controles.

5.2.2. ODDS-RATIO PARA LA VARIABLE “ANCHO DE CARRIL”

RECTAS Y CURVAS

De nuevo, para conocer la fuerza de la asociación estadística, calcularemos los Odds-Ratio, realizando una regresión logística binaria.

En este caso, la variable situada como referencia ha sido “anchura de carril menor o igual que 3,25 metros”.

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
$\leq 3,25$ m			157,442	2	,000			
3,25 m-3,75 m	1,623	,129	157,346	1	,000	5,066	3,932	6,528
$\geq 3,75$ m	1,051	,379	7,696	1	,006	2,861	1,361	6,013
Constante	-,918	,098	87,254	1	,000	,399		

Tabla 23. Odds-ratio para la variable "Ancho de carril", rectas y curvas

De nuevo, se crean tantas variables dicotómicas como categorías tiene la variable original menos una, en las que el valor “0” asociado a estas nuevas variables es la variable establecida como referencia.

De la tabla anterior, se puede deducir:

- La presencia de anchuras de carril entre 3,25 y 3,75 es un factor de riesgo, pues el Intervalo de Confianza creado no incluye el valor “1”. Además, el OR es 5,066, lo que indica que anchos de carril comprendidos entre 3,25 metros y 3,75 metros, se asocian a un aumento de más de cinco veces del riesgo de ocurrencia del evento “colisión frontal”, respecto ubicaciones con ancho de carril inferior a 3,25 metros.
- La presencia de anchuras de carril mayores de 3,75 actúa como un factor de riesgo, pues de nuevo el Intervalo de Confianza creado para el Odds-ratio (OR) no incluye el valor “1”. Además, el OR es 2,861, lo que indica que la presencia del factor “ancho de carril mayor que 3,75 metros” se asocia con un aumento del riesgo de ocurrencia del desenlace estudiado de casi tres veces con respecto a lugares con anchos de carril inferiores a 3,25 metros.

El análisis anterior nos permite concluir que el ancho de carril actúa como un factor de riesgo, aumentando el riesgo de que se produzcan colisiones frontales. Además, los anchos de carril mayores de 3,25 metros, pero menores que 3,75 metros, son los que presentan una mayor asociación con el desenlace estudiado.

RECTAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
$\leq 3,25$ m			106,774	2	,000			
3,25 m-3,75 m	1,822	,177	105,972	1	,000	6,184	4,371	8,749
$\geq 3,75$ m	1,860	,744	6,242	1	,012	6,421	1,493	27,614
Constante	-1,349	,144	88,094	1	,000	,260		

Tabla 24. Odds-ratio para la variable "Ancho de carril, solo rectas"

Los resultados obtenidos tras realizar la regresión logística binaria incluyendo tan solo las ubicaciones rectas son muy similares a los obtenidos incluyendo todos los casos.

- Anchuras de carril intermedias (entre 3,25 metros y 3,75 metros) aumentan el riesgo de que se produzcan colisiones frontales respecto a anchuras de carril inferiores a 3,25 metros (OR=6,184). El intervalo de confianza no contiene al 1, y el p-valor asociado es 0.
- Anchos de calzada mayores que 3,75 metros actúan como factor de riesgo, aumentando más de 6 veces el riesgo de que se produzcan colisiones frontales. El p-valor toma el valor de 0,012, por lo que no sería significativo a nivel $\alpha=0,01$, pero sí a nivel $\alpha=0,025$.

CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
$\leq 3,25$ m			57,743	2	,000			
3,25 m-3,75 m	1,553	,205	57,258	1	,000	4,726	3,161	7,067
$\geq 3,75$ m	,421	,449	,880	1	,348	1,524	,632	3,673
Constante	-,421	,140	8,998	1	,003	,656		

Tabla 25. Odds-ratio para la variable "Ancho de carril", solo curvas"

En el caso de realizar la regresión logística incluyendo tan solo las curvas, el resultado obtenido es diferente.

- Un ancho de carril incluido en el rango 3,25 metros- 3,75 metros, es un factor de riesgo frente a un ancho de carril menor o igual que 3,25 metros. El odds-ratio es igual a 4,726, por lo que se puede afirmar que en ubicaciones en las que el ancho de calzada se encuentre entre 3,25 metros y 3,75 metros, se aumenta casi 5 veces el riesgo de colisiones frontales respecto a aquellas ubicaciones en las que el ancho de carril es menor o igual que 3,25 metros.
- Anchos de carril mayores o iguales que 3,75 metros, no actúan como factor de riesgo, dado que el p-valor toma el valor de 0,348, por lo que no es significativo a los niveles de significación habituales. Además, el Intervalo de Confianza calculado para el Odds-ratio no contiene al 1.

5.3. RADIO DE CURVATURA

5.3.1. TABLA CRUZADA RADIO*CASOCONTROL

	Caso-control		Total
	Control	Caso	
Rectas	390 66,33%	310 52,72%	700 59,52%
Curva suave (≥ 750)	17 2,89%	21 3,57%	38 3,23%
Curvas intermedias (750 metros -250 metros)	72 12,24%	95 16,16%	167 14,20%
Curvas cerradas (≤ 250)	109 18,54%	162 27,55%	271 23,04%
Total	588	588	1176

Tabla 26: Tabla cruzada "Radio de curvatura" y "Caso-control"

De acuerdo con los datos de la tabla cruzada anterior, se puede apreciar que en las rectas los controles son más abundantes que los casos, mientras que,

en todos los tipos de curvas, los casos con más abundantes que los controles. La afirmación anterior se traduce en que las colisiones frontales podrían presentar mayor riesgo de ocurrencia en las curvas.

Además, se puede apreciar que, en las curvas, el porcentaje de casos crece con respecto al de controles a medida que disminuye el radio de curvatura.

En el siguiente gráfico de barras se puede apreciar claramente como aumenta el número de casos frente a los controles a medida que disminuye el radio de curvatura, mientras que en las rectas el número de controles supera al número de casos.

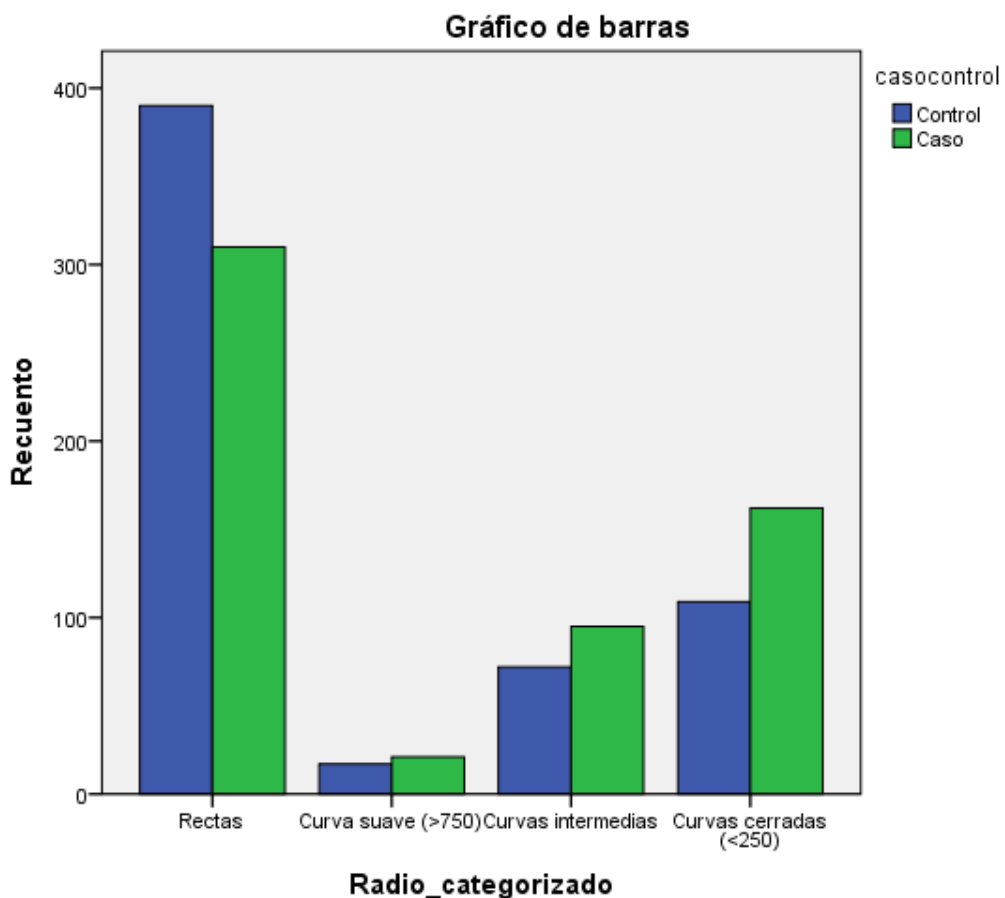


Ilustración 26: Gráfico de barras de la distribución de la variable "Radio de curvatura" en los casos y en los controles.

Dado que en las rectas el número de controles es considerablemente superior al de casos, y en las curvas predominan los casos, se puede decir que a priori las colisiones frontales presentan mayor riesgo de ocurrencia en las curvas, por lo que la categoría situada de referencia para los análisis posteriores serán las rectas, pues es la categoría de menor riesgo.

5.3.2. ODDS-RATIO PARA LA VARIABLE “RADIO DE CURVATURA”

De nuevo, se realizará una regresión logística binaria, de modo que se calcularán los Odds-ratio, y consecuentemente, se conocerá la fuerza de la asociación estadística entre la variable “Radio de curvatura” y la variable “Caso-control”.

Tal y como se indicó en el apartado anterior, se situará como categoría de referencia para los tres modelos la categoría “rectas”.

RECTAS Y CURVAS								
	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Rectas			22,900	3	,000			
≥ 750	,441	,335	1,732	1	,188	1,554	,806	2,997
750 m -250 m	,507	,174	8,503	1	,004	1,660	1,181	2,334
≤ 250	,626	,145	18,529	1	,000	1,870	1,406	2,486
Constante	-,230	,076	9,103	1	,003	,795		

Tabla 27 Regresión Logística para la variable "Radio de Curvatura Categorizado", rectas y curvas

Conocido el significado de las variables Dummy creadas, y sabiendo que la categoría de referencia son las rectas, podemos obtener las siguientes conclusiones:

- Las curvas suaves (radio > 750 metros) no aumentan el riesgo de que se produzcan colisiones frontales respecto a las rectas, dado que el p-valor no es significativo a los niveles habituales, y el Intervalo de confianza creado no contiene al 1.
- Las curvas intermedias, cuyos radios pertenecen al rango 250 metros - 750 metros, actúan como factor de riesgo, pues aumentan 1,660 veces el riesgo de que se produzcan colisiones frontales con respecto a las rectas (OR=1,66). Sin embargo, se puede observar que el p-valor toma el valor de 0,04, y el intervalo no contiene al 1, por lo que la asociación es significativa a nivel $\alpha=0,05$, pero no lo es a nivel $\alpha=0,01$.

- Las curvas pronunciadas, cuyos radios son menores que 250 metros, aumentan el riesgo de que se produzcan colisiones frontales respecto a las rectas, pues el p-valor es 0, y el Intervalo de Confianza calculado al 95% no contiene al 1. Por tanto, las curvas pronunciadas aumentan el riesgo de que se produzcan colisiones frontales casi al doble (OR=1,87) respecto a las rectas.

Por tanto, se puede observar que a las curvas intermedias y pronunciadas aumentan el riesgo de que se produzcan colisiones frontales respecto a las rectas, y que a medida que la curva se hace más pronunciada, es decir, a medida que disminuye el radio de curvatura, esta influencia se hace más significativa.

5.4. ARCÉN

5.4.1. TABLA FRECUENCIA ARCÉN*CASOCONTROL

	Caso-control		Total
	Control	Caso	
Inexistente o impracticable	152 25,85%	101 17,18%	253 21,51%
Menor de 1,5 metros	320 54,42%	351 59,69%	671 57,05%
De 1,5 a 2,49 metros	102 17,34%	127 21,60%	229 19,47%
De 2,5 metros en adelante	14 2,38%	9 1,53%	23 1,96%
Total	588	588	1176

Tabla 28: Tabla cruzada "Arcén" y "Caso-control"

La tabla cruzada anterior muestra que el número de controles es mayor que el número de casos cuando no existe arcén. Sin embargo, el número de casos supera al de controles cuando el arcén presenta una anchura entre 1,5 metros y 2,40 metros. Se puede apreciar que el número de controles supera de nuevo a los casos cuando el arcén es superior a 2,5 metros.

La categoría arcén inexistente o impracticable será elegida como categoría de referencia, pues el porcentaje de controles es considerablemente superior al de casos para esta categoría, por lo que se considera la categoría de menor riesgo.

5.4.2. ODDS-RATIO PARA LA VARIABLE “ARCÉN”

De nuevo, para conocer la fuerza de la asociación estadística, calcularemos los Odds-Ratio, realizando una regresión logística binaria. En este caso, la variable situada como referencia ha sido “arcenes impracticables o insuficientes”.

RECTAS Y CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Inexistente o impracticable			15,356	3	,002			
$\leq 1,5$ m	,501	,150	11,189	1	,001	1,651	1,231	2,214
1,5 m-2,49 m	,628	,185	11,545	1	,001	1,874	1,304	2,692
$\geq 2,49$ m	-,033	,446	,005	1	,941	,967	,404	2,319
Constante	-,409	,128	10,139	1	,001	,664		

Tabla 29 Regresión logística variable "arcén", rectas y curvas

Se pueden extraer las siguientes conclusiones de la tabla anterior:

- La presencia de arcenes cuya anchura es inferior a 1,5 metros actúa como factor de riesgo, dado que el p-valor =0,001, y el Intervalo de confianza creado no contiene al 1. El Odds-ratio toma el valor de 1,651, lo que implica que la presencia de arcenes menores de 1,5 metros aumenta el riesgo de ocurrencia de colisiones frontales 1,651 veces con respecto a la ausencia de arcén.
- Del mismo modo, la presencia de arcenes de anchura entre 1,5 metros y 2,5 metros, actúa como factor de riesgo aumentando el riesgo de que se produzcan las colisiones frontales casi al doble (Odds-ratio=1,874), dado que el p-valor=0,001, y el Intervalo de Confianza no contiene al 1.
- La presencia de arcenes muy anchos ($\geq 2,49$ metros), por el contrario, no afecta al evento “colisiones frontales”, pues el Intervalo de Confianza contiene el valor 1, y el p-valor toma el valor de 0,941. Por tanto, arcenes cuya anchura es mayor o igual que 2,5 metros no aumentan el riesgo de colisiones frontales respecto a arcenes impracticables o insuficientes.

Por tanto, se puede deducir que la presencia de arcén actúa como factor de riesgo aumentando el riesgo de ocurrencia de colisiones frontales respecto aquellas ubicaciones que carecen de arcén, excepto en el caso de arceles muy anchos (>2,5 metros), los cuales son escasos en carreteras convencionales.

RECTAS

	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Inferior	Superior
Inexistente o impracticable			19,289	3	,000			
≤ 1,5 m	,750	,209	12,935	1	,000	2,118	1,407	3,188
1,5 m-2,49 m	1,007	,240	17,581	1	,000	2,737	1,709	4,382
≥ 2,49 m	,295	,587	,254	1	,614	1,344	,426	4,242
Constante	-,883	,181	23,730	1	,000	,413		

Tabla 30 Regresión logística para la variable "Arcén", solo rectas

De nuevo, los resultados obtenidos en el modelo logístico que solo incluye las ubicaciones que se corresponden con rectas son muy similares a los obtenidos en el caso de curvas y rectas:

- Arcenes menores o iguales que 1,5 metros aumentan el riesgo de colisiones frontales más del doble (OR=2,118) respecto a carreteras con arceles inexistentes o impracticables, dado que el p-valor es 0, y el Intervalo de Confianza no contiene al 1.
- La presencia de un arcén cuya anchura es superior a 1,5 metros, pero inferior a 2,5 metros, aumenta el riesgo de colisiones frontales casi el triple (OR=2,737) respecto a aquellas ubicaciones en las que el arcén es impracticable o insuficiente. De nuevo se puede observar que esta asociación es estadísticamente significativa a través del p-valor y del Intervalo de confianza calculado, el cual no contiene al 1.
- Arcenes de más de 2,5 metros de anchura no aumentan el riesgo de colisiones frontales respecto a arceles impracticables o insuficientes, pues el p-valor es mayor que 0,1, por lo que no es significativo a los niveles habituales.

CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Inexistente o impracticable			1,661	3	,646			
$\leq 1,5$ m	,196	,228	,739	1	,390	1,217	,778	1,904
1,5 m-2,49 m	,266	,318	,702	1	,402	1,305	,700	2,432
$\geq 2,49$ m	-,412	,699	,348	1	,555	,662	,168	2,604
Constante	,189	,195	,941	1	,332	1,208		

Tabla 31: Regresión logística para la variable "Arcén", curvas

En el caso de la regresión logística binaria realizada incluyendo solo las curvas, los resultados son bastante diferentes, ya que la presencia de arcén no constituye un factor de riesgo respecto a la ausencia del mismo en las ubicaciones curvas, independientemente de la anchura del mismo. Se puede apreciar que los p-valores no son significativos a los niveles habituales, y el Intervalo de confianza calculado contiene al 1 en todos los casos.

5.5. BARRERA DE SEGURIDAD

5.5.1. TABLA CRUZADA BARRERA DE SEGURIDAD*CASOCONTROL

	Caso-control		Total
	Control	Caso	
No	437 74,32%	342 58,16%	779 66,24%
Si	151 25,68%	246 41,84%	397 33,76%
Total	588	588	1176

Tabla 32 Tabla cruzada "Barrera de seguridad" y "Caso-Control"

De acuerdo con la tabla cruzada anterior, se puede apreciar que la distribución de la variable "Barrera de Seguridad" es diferente en los casos y en los controles. Se aprecia claramente que cuando no existe este sistema de contención, el número de controles es considerablemente superior al de casos

(437 frente a 342). Sin embargo, cuando existe barrera de seguridad, el número de casos supera considerablemente al de controles (246 frente a 151). Dados los datos anteriores, en los posteriores análisis situaremos la categoría “No” de la variable “Barrera de seguridad” como categoría de referencia.

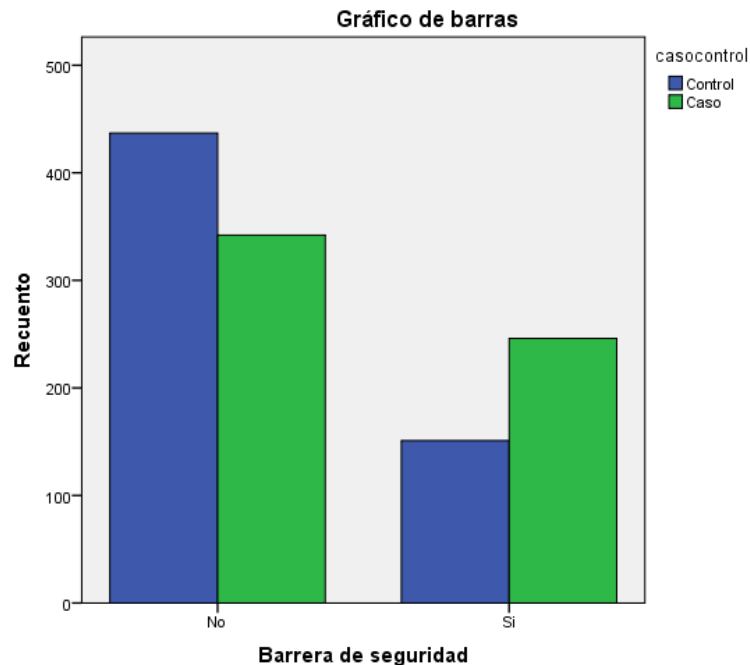


Ilustración 27: Gráfico de Barras, distribución de la variable “Barrera de Seguridad” para los casos y los controles.

Esta información se resume en el gráfico de barras anterior. Visualmente, se aprecia cómo se invierte la predominancia de los casos y los controles para las dos categorías de la variable “Barrera de Seguridad”.

5.5.2. ODDS-RATIO PARA LA VARIABLE “BARRERA DE SEGURIDAD”

De nuevo, para conocer la fuerza de la asociación estadística entre la variable “Barrera de Seguridad” y el desenlace “Colisión Frontal”, calcularemos el Odds-Ratio, realizando una regresión logística binaria.

En este caso, la variable situada como referencia ha sido la ausencia de barreras de seguridad.

RECTAS Y CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Barrera de seguridad	,733	,126	33,808	1	,000	2,082	1,626	2,665
Constante	-,245	,072	11,528	1	,001	,783		

Tabla 33 Regresión Logística para la variable "Barrera de Seguridad", Curvas y Rectas

Dado que el Intervalo de confianza calculado no contiene al 1, y el p-valor toma el valor de 0, se puede afirmar que la presencia de barreras de seguridad en los márgenes de la carretera actúa como factor de riesgo, pues aumenta el riesgo de que se produzca el desenlace estudiado en una cantidad igual al valor que toma el Odds-ratio (OR=2,082) con respecto a aquellas ubicaciones que no están provistas con esta medida de contención.

RECTAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Barrera de seguridad	,728	,188	14,914	1	,000	2,071	1,431	2,996
Constante	-,384	,087	19,693	1	,000	,681		

Tabla 34: Regresión Logística para la variable "Barrera de Seguridad", solo rectas

Los resultados obtenidos en el modelo que incluye solo las ubicaciones que se corresponden con rectas son similares a los obtenidos con el modelo en el que se incluyen todos los casos.

En las ubicaciones rectas la barrera de seguridad actúa como factor de riesgo, aumentando el riesgo de que se produzcan colisiones frontales a más del doble (Odds-ratio=2,071) con respecto a las ubicaciones rectas que carecen de este tipo de elementos de seguridad, dado que el p-valor=0, y el intervalo de confianza al 95% no contiene al 1.

CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Barrera de seguridad	,487	,187	6,753	1	,009	1,627	1,127	2,349
Constante	,089	,133	,442	1	,506	1,093		

Tabla 35. Regresión Logística para la variable "Barrera de Seguridad", Solo curvas

Los resultados obtenidos incluyendo tan solo las ubicaciones curvas son de nuevo similares a los dos modelos anteriores, dado que la barrera de seguridad actúa como factor de riesgo, aumentando el riesgo de que se produzcan colisiones frontales 1,6 veces más con respecto a las ubicaciones curvas que no están provistas de este tipo de sistemas de contención de vehículos. La asociación es significativa, dado que el p-valor es 0, y el intervalo de confianza no contiene al 1.

5.6. HITOS DE ARISTA

5.6.1. TABLA CRUZADA HITOS DE ARISTA*CASOCONTROL

	Caso-control		Total
	Control	Caso	
No	404 68,71%	213 36,22%	617 52,46%
Si	184 31,29%	375 63,78%	559 47,54%
Total	588	588	1176

Tabla 36: Tabla cruzada "Hitos de Arista" y "Caso-Control"

De acuerdo con la tabla anterior, los casos se encuentran mayoritariamente situados en ubicaciones con presencia de hitos de arista. Por el contrario, en los lugares que carecen de hitos de arista, encontramos que el número de controles (ubicaciones en las que no han ocurrido colisiones frontales) es mayor que el número de casos.

El siguiente gráfico de barras resume la información anterior. Se puede observar cómo se invierte la distribución de los casos y los controles en las dos categorías de la variable. Mientras que para la categoría "No" la barra

correspondiente a los controles supera considerablemente a los casos, en la categoría “Sí” ocurre lo contrario.

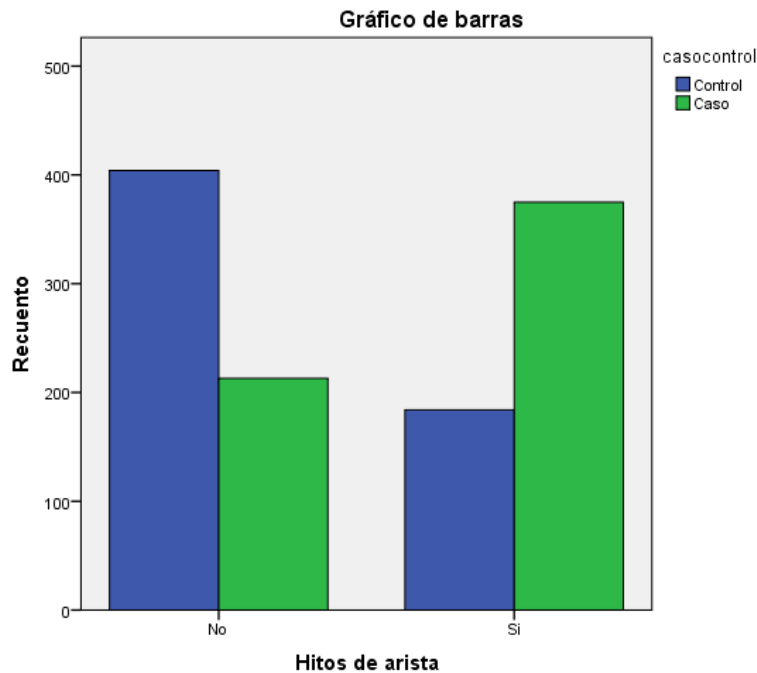


Ilustración 28. Gráfico de barras distribución de los casos y los controles para la variable "Hitos de Arista"

Por tanto, la categoría “No” de la variable “Hitos de arista” será elegida como referencia para los análisis posteriores, pues se trata de la categoría de menor riesgo, de acuerdo con los resultados revelados por la tabla anterior.

5.6.2. ODDS-RATIO PARA LA VARIABLE “HITOS DE ARISTA”

Siguiendo el mismo procedimiento que con el resto de variables, calcularemos los Odds-ratio, realizando una regresión logística binaria.

Tal y como se ha justificado en el apartado anterior, la categoría situada como referencia será la ausencia de hitos de arista (Hitos de arista= “No”).

RECTAS Y CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Hitos de arista	1,352	,124	119,714	1	,000	3,866	3,034	4,925
Constante	-,640	,085	57,148	1	,000	,527		

Tabla 37: Regresión Logística para la variable "Hitos de arista", Rectas y curvas

De acuerdo a la tabla anterior, se comprueba que las ubicaciones provistas con hitos de arista en los laterales de la carretera presentan un mayor riesgo de que se produzcan colisiones frontales respecto a aquellas ubicaciones que carecen de estos elementos.

El intervalo de confianza no contiene al 1, y el p-valor es 0. El Odds-ratio toma el valor de 3,866, lo que se traduce en un aumento del riesgo de colisiones frontales de casi 4 veces en las ubicaciones provistas de hitos de arista, respecto a aquellas que no cuentan con este tipo de elementos.

RECTAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Hitos de arista	1,196	,159	56,596	1	,000	3,307	2,421	4,515
Constante	-,803	,112	51,691	1	,000	,448		

Tabla 38 Regresión Logística para la variable "Hitos de arista", rectas

En el caso del modelo que incluye solo las ubicaciones rectas, los resultados obtenidos fueron muy similares a los obtenidos en el modelo que incluía todos los casos.

De acuerdo con la tabla anterior, la presencia de hitos de arista aumenta más del triple (OR=3,307) el riesgo de que se produzcan colisiones frontales en las rectas, respecto a las rectas que no están provistas de estos elementos. El resultado es estadísticamente significativo, dado que el Intervalo de Confianza creado no contienen al 1, y el p-valor toma el valor 0.

CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Hitos de arista	1,630	,204	63,888	1	,000	5,105	3,423	7,614
Constante	-,402	,131	9,393	1	,002	,669		

Tabla 39 Regresión logística para la variable "Hitos de arista", curvas

En la regresión logística realizada para el modelo que incluye solo las curvas, el resultado muestra, de nuevo, un aumento del riesgo de colisiones frontales en las ubicaciones provistas de hitos de arista, respecto a aquellas ubicaciones que carecen de estos elementos.

En este caso, el Odds-ratio toma el valor de 5,105, lo que indica que la asociación estadística entre los hitos de arista y las colisiones frontales es mayor en ubicaciones curvas que en las rectas. La asociación es significativa, pues el p-valor es 0, y el Intervalo de confianza calculado al 95% no contiene el valor 1.

5.7. ACERA

5.7.1. TABLA CRUZADA ACERA*CASOCONTROL

	Caso-control		Total
	Control	Caso	
No	575 97,79%	566 96,26%	1141 97,02%
Si	13 2,21%	22 3,74%	35 2,98%
Total	588	588	1176

Tabla 40: Tabla cruzada "Acera" * "Caso-control"

La tabla anterior muestra una escasa diferencia entre el número de casos y controles en ambas categorías de la variable "Acera". Por ello, se puede intuir que la asociación entre esta variable y el desenlace estudiado no va a ser significativa.

Sin embargo, con objetivo de obtener evidencias estadísticas de la falta de asociación, se realizará la regresión logística binaria para esta variable.

La categoría "No" será utilizada como referencia, pues para esta categoría el número de controles, o ubicaciones en las que se ha producido colisión frontal, es ligeramente superior.

En el siguiente gráfico de barras, correspondiente a la variable "acera", se puede apreciar la escasa diferencia entre casos y controles en las categorías de la variable.

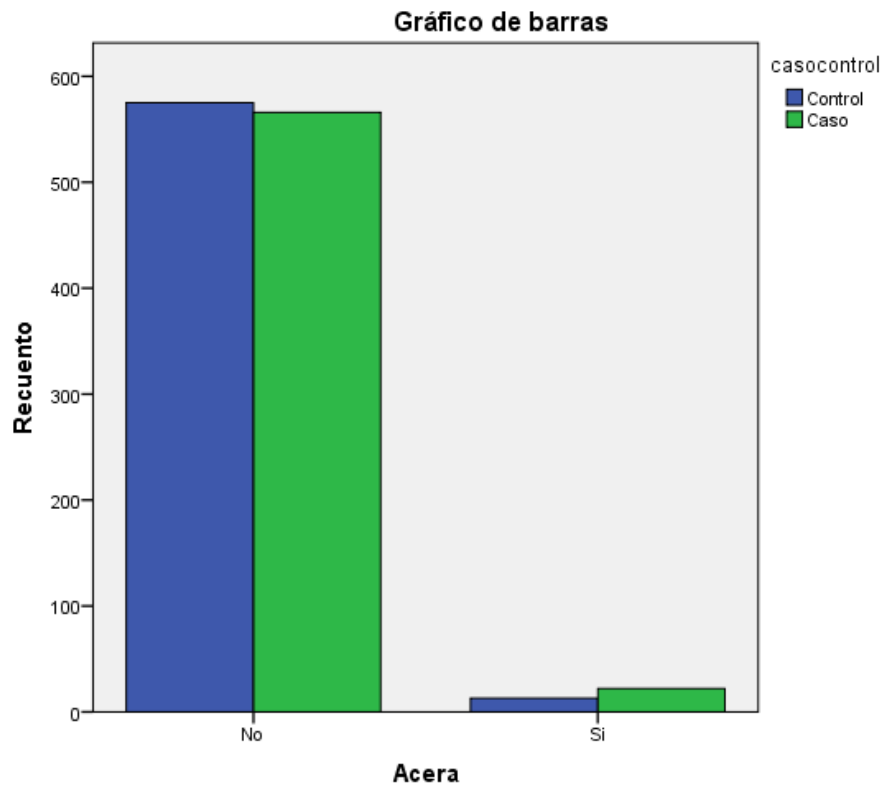


Ilustración 29: Gráfico de barra de la distribución de los casos y los controles para las categorías de la variable "Acera"

5.7.2. ODDS-RATIO PARA LA VARIABLE "ACERA"

De nuevo, se realiza una regresión logística univariante para calcular el Odds-Ratio de la variable "Acera"

En este caso, la categoría elegida como referencia ha sido la ausencia de acera, por los motivos anteriormente explicados.

RECTAS Y CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Acera	,542	,355	2,332	1	,127	1,719	,858	3,446
Constante	-,016	,059	,071	1	,790	,984		

Tabla 41 Regresión Logística para la variable "acera", rectas y curvas

De acuerdo con los resultados obtenidos al realizar la regresión logística binaria incluyendo todos los casos, la presencia de acera no actúa como factor de riesgo, tal y como se intuía en el apartado anterior, dado que el Intervalo de Confianza contiene al 1, y el p-valor es mayor que 0,1, por lo que no es significativo a los niveles habituales.

RECTAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Acera	,835	,424	3,877	1	,049	2,304	1,004	5,287
Constante	-,259	,078	11,151	1	,001	,772		

Tabla 42: Regresión Logística para la variable "Acera", rectas

Los resultados obtenidos para el modelo que incluye solo las rectas muestran que, para dichas ubicaciones, la presencia de acera podría ejercer cierta influencia, aumentando el riesgo de colisiones frontales respecto a las ubicaciones rectas desprovistas de acera.

Sin embargo, podemos ver que esta influencia se encuentra al límite de la significación estadística, pues el p-valor=0,049, por lo que no es significativo a nivel $\alpha=0,01$, pero sí a nivel $\alpha=0,05$. Esto último se puede apreciar también en el Intervalo de Confianza, pues no contiene el valor 1, pero se encuentra justo en el límite de contenerlo.

CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Acera	,068	,652	,011	1	,918	1,070	,298	3,842
Constante	,338	,094	12,932	1	,000	1,402		

Tabla 43. Regresión Logística para la variable "Acera", curvas

Analizando los resultados obtenidos para el modelo "solo curvas", se puede apreciar que no existe evidencia de asociación estadística entre la variable "acera" y el desenlace estudiado "colisión frontal". El Intervalo de Confianza

contiene al 1, y el p-valor no es significativo a los niveles habituales (p -valor=0,918).

5.8. INTENSIDAD MEDIA DIARIA CATEGORIZADA (IMD)

5.8.1. TABLA CRUZADA INTENSIDAD MEDIA DIARIA*CASO-CONTROL

	Caso-control		Total
	Control	Caso	
IMD <500	211 35,88%	56 9,52%	3 0,26%
$500 \leq \text{IMD} < 2000$	262 44,56%	195 33,16%	101 8,59%
$2000 \leq \text{IMD} < 10000$	111 18,88%	303 51,53%	1026 87,24%
$\text{IMD} \geq 10000$	4 0,68%	31 5,27%	46 3,91%
Total	588	588	1176

Tabla 44: Tabla cruzada "Intensidad Media Diaria"*"Caso-control"

La tabla cruzada muestra los siguientes resultados:

- Cuando la intensidad media diaria es menor que 2000 vehículos al día, el número de controles es considerablemente superior al de los casos.
- Cuando la Intensidad media diaria es elevada ($\text{IMD} > 2000$), ocurre el fenómeno contrario. El número de casos supera considerablemente a los controles.

Dado que la categoría $\text{IMD} < 500$ presenta una mayor diferencia entre el número de casos y el de controles, se considera la categoría de menor riesgo, por lo que será establecida como categoría de referencia.

En el siguiente gráfico de barras se aprecia visualmente la inversión en las proporciones del número de casos y controles para categorías altas y bajas del IMD. Es decir para las dos primeras categorías, las cuales se corresponden con Intensidades Medias bajas, el número de controles (barra azul) supera a los

casos, especialmente en la primera categoría. Por el contrario, en las dos últimas categorías, que se corresponden con IMD altas, el número de casos (barra verde) supera considerablemente a los controles.

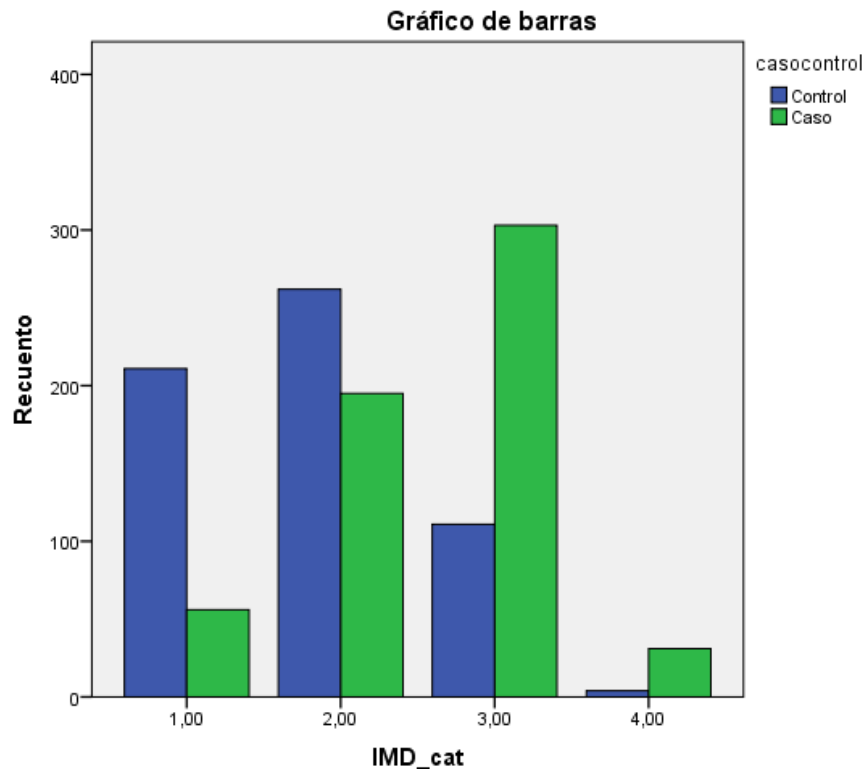


Ilustración 30: Gráfico de Barras para la distribución de la variable "IMD categorizado" en los casos y en los controles

5.8.2. ODDS-RATIO PARA LA VARIABLE "INTENSIDAD MEDIA DIARIA CATEGORIZADA" (IMD)

De nuevo se realiza la regresión logística binaria, situando la variable "IMD Categorizado" como variable independiente, y la variable "Caso-control" como variable respuesta.

Tal y como se justificó en el apartado anterior, la categoría $IMD < 500$ será utilizada como categoría de referencia.

RECTAS Y CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
< 500			184,367	3	,000			
$500 \leq \text{IMD} < 2000$	1,031	,178	33,711	1	,000	2,804	1,980	3,972
$2000 \leq \text{IMD} < 10000$	2,331	,187	155,625	1	,000	10,285	7,132	14,834
$\text{IMD} \geq 10000$	3,374	,552	37,346	1	,000	29,201	9,895	86,174
Constante	-1,327	,150	77,871	1	,000	,265		

Tabla 45: Regresión Logística para la variable "IMD Categorizada", curvas y rectas

De acuerdo con los resultados obtenidos tras la implementación del modelo en SPSS, se pueden realizar las siguientes afirmaciones:

- Una IMD comprendida en el rango de 500 a 2000 vehículos por día aumentan el riesgo de que se produzcan colisiones frontales casi al triple (Odds-ratio=2,804) respecto a aquellas localizaciones en las que la IMD es inferior a 500 vehículos por día. La asociación es estadísticamente significativa, dado que p-valor es 0, y el Intervalo de Confianza creado no contiene al 1.
- Una IMD superior a 2000 vehículos al día, e inferior a 10000 vehículos al día, actúa como factor de riesgo aumentando el riesgo de que se produzcan colisiones frontales más de 10 veces (Odds-ratio=10,285) con respecto a una IMD menor de 500. De nuevo, se puede justificar que la asociación es estadísticamente significativa, pues el p-valor es 0, y el Intervalo de Confianza no contiene al 1.
- Por último, flujos de tráfico muy elevados (IMD superior a 10000 vehículos por día), actúan como factor de riesgo aumentando el riesgo de que se produzca el evento estudio casi 30 veces (Odds-ratio=29,201). De nuevo, se justifica que la asociación es significativa.

En definitiva, se podría afirmar que a medida que aumenta la IMD, también lo hace el riesgo de que se produzcan colisiones frontales. Por tanto, esta variable actúa como factor de riesgo, que aumenta el riesgo de que se produzca el desenlace estudiado, "colisión frontal".

RECTAS

	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Inferior	Superior
< 500			121,370	3	,000			
500 ≤ IMD < 2000	1,230	,262	22,081	1	,000	3,422	2,048	5,716
2000 ≤ IMD < 10000	2,643	,268	96,949	1	,000	14,050	8,303	23,775
IMD ≥ 10000	3,347	,677	24,480	1	,000	28,424	7,548	107,039
Constante	-1,807	,230	61,691	1	,000	,164		

Tabla 46: Regresión Logística para la variable "IMD Categorizada", rectas.

Si se implementa el modelo que incluye tan solo las localizaciones que se corresponden con rectas, se podría afirmar de nuevo que un aumento de la IMD aumenta el riesgo de que se produzcan colisiones frontales en rectas.

- IMD situado en el rango de 500 a 2000 vehículos por día actúa como factor de riesgo, aumentando hasta más del triple (OR=3,422) el riesgo de que se produzcan colisiones frontales.
- Localizaciones con Intensidades Medias Diarias situadas en el rango desde 2000 a 10000 vehículos por día aumentan el riesgo de que se produzcan colisiones frontales catorce veces (OR=14,05) respecto a aquellas localizaciones con IMD menor a 500 vehículos por día.
- Por último, IMD superiores a 10000 vehículos por día actúan como factor de riesgo, pues aumenta el riesgo de que se produzca el evento estudio casi 29 veces (OR=28,424) con respecto a ubicaciones con una IMD menor de 500 vehículos al día.

Todas las asociaciones anteriores son significativas, dado que el p-valor en todos los casos es 0, y el Intervalo de Confianza no contiene al 1.

CURVAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
< 500			64,032	3	,000			
500 ≤ IMD < 2000	,941	,255	13,644	1	,000	2,561	1,555	4,219
2000 ≤ IMD < 10000	2,101	,279	56,712	1	,000	8,178	4,733	14,131
IMD ≥ 10000	3,651	1,049	12,102	1	,001	38,500	4,923	301,095
Constante	-,817	,206	15,760	1	,000	,442		

Tabla 47. Odds-ratio para la variable "IMD categorizada", curvas

De nuevo, se pueden extraer las mismas conclusiones en el caso de las curvas. Es decir, se puede afirmar que una IMD superior a 500 vehículos al día es un factor de riesgo, pues aumenta el riesgo de que se produzcan colisiones frontales. Además, se puede apreciar que cuanto mayor es el IMD, mayor es el OR, por lo que más aumentará dicho riesgo.

En el caso de las curvas, la influencia de una IMD muy elevada (IMD > 10000) es aún mayor que en los dos modelos anteriores. Se podría afirmar que una IMD muy elevada (superior a 10000 vehículos al día), tiene un gran impacto en las colisiones frontales en las ubicaciones curvas, aumentando el riesgo de que se produzcan casi 40 veces respecto a las localizaciones curvas con una IMD menor de 500 vehículos al día.

5.9. NÚMERO DE CARRILES

5.9.1. TABLA CRUZADA NÚMERO DE CARRILES * CASO-CONTROL

	Caso-control		Total
	Control	Caso	
Un carril en cada sentido	577 98,12%	555 94,38%	1132 96,26%
Al menos un sentido con varios carriles	11 1,88%	33 5,61%	44 3,74%
Total	588	588	1176

Tabla 48. Tabla cruzada "Número de carriles" y "Caso-control"

De acuerdo con la tabla anterior, los controles en las ubicaciones que cuentan con un carril para cada sentido superan a los casos. Por el contrario, en la categoría que incluye las ubicaciones con al menos un sentido con más de un carril, el número de casos es superior al de controles.

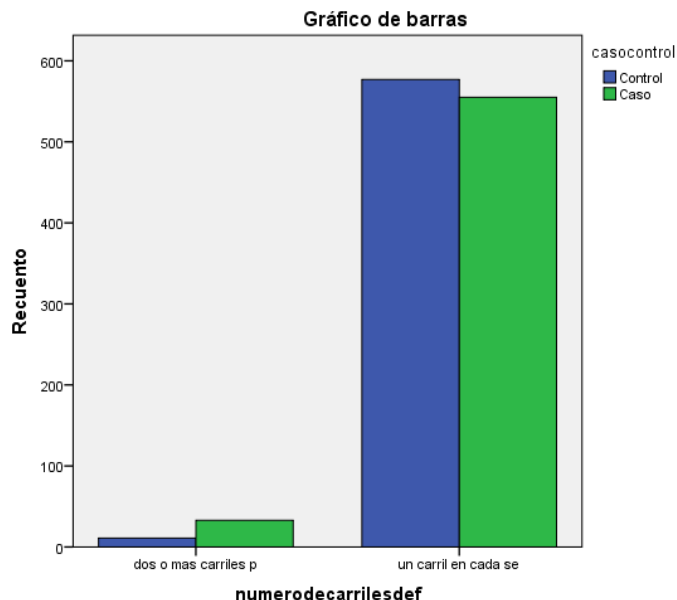


Ilustración 31. Gráfico de barras para la variable "Número de carriles"

En el gráfico de barras anterior se puede observar cómo cambia la distribución de los casos y los controles en las dos categorías de la variable.

En definitiva, se puede intuir que el hecho de que una ubicación cuente con varios carriles para cada sentido puede constituir un factor de riesgo. Acorde a esto, la categoría que será establecida como referencia sería "un carril para cada sentido".

5.9.2. ODDS-RATIO PARA LA VARIABLE "NÚMERO DE CARRILES"

CURVAS Y RECTAS

	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Número de carriles	1,137	,353	10,372	1	,001	3,119	1,561	6,232
Constante	-,039	,059	,428	1	,513	,962		

Tabla 49. Odds-ratio para la variable "Número de carriles"

De acuerdo con la tabla anterior, se puede afirmar que la presencia de un sentido con más de un carril actúa como factor de riesgo, dado que el Odds-ratio toma el valor de 3,119, lo que se traduce en que el riesgo de que se produzca el resultado estudiado se aumenta tres veces cuando hay más de un carril en al menos uno de los sentidos, respecto a cuando solo hay un carril en cada sentido. La asociación es significativa, dado que el p-valor=0,001, y el Intervalo de Confianza para el OR no contiene el valor 1.

RECTAS								
	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Número de carriles	,935	,711	1,726	1	,189	2,546	0,632	10,263
Constante	-,241	,077	9,921	1	,002	,786		

Tabla 50. Odds-ratio e IC para la variable "Número de carriles", rectas

Las conclusiones obtenidas en el modelo "solo rectas" difieren de las obtenidas en el modelo anterior, dado que, en este caso, la presencia de más de un carril en al menos uno de los sentidos, no supone un riesgo adicional respecto al hecho de que solo haya un carril en cada sentido. El-p-valor es >0,1, por lo que la asociación no es significativa.

CURVAS								
	β	Error estándar	Wald	gl	Sig.	Exp(β)	95% C.I. para EXP(β)	
							Inferior	Superior
Número de carriles	,938	,414	5,136	1	,023	2,555	1,135	5,750
Constante	,278	,096	8,383	1	,004	1,321		

Tabla 51. Odds-ratio e IC para la variable "Número de carriles", Curvas

De acuerdo con esta última tabla, se pueden realizar las mismas afirmaciones en el modelo "solo curvas" que en el "modelo completo". La presencia de un sentido con más de un carril multiplica 2,555 veces el riesgo de colisiones frontales. El intervalo de confianza calculado al 95% no contiene al 1, y el p-valor toma el valor de 0,023, por lo que el resultado es significativo para $\alpha=0,025$, pero no a nivel $\alpha=0,01$.

5.10. RESUMEN DE INFLUENCIA DE LAS VARIABLES INDEPENDIENTES

Como resumen del capítulo, se muestra una tabla en el que se resume la influencia de las variables cuando actúan de manera independiente.

CURVAS Y RECTAS

Variable independiente	Categorías	Riesgo de colisión frontal	Tipo de factor
Ancho de calzada	Ancho de calzada >5,99 m	Aumenta	Factor de riesgo
Anchura de carril	Ancho de carril >3,25 m	Aumenta	Factor de riesgo
Radio de curvatura	Radio < 750 m	Aumenta	Factor de riesgo
Arcén	Presencia arcén < 2,49 m	Aumenta	Factor de riesgo
Barrera de seguridad	Presencia de Barrera	Aumenta	Factor de riesgo
Hitos de arista	Presencia de Hitos de Arista	Aumenta	Factor de riesgo
Acera	No existe asociación		
IMD	IMD superior a 500 vehículos al día	Aumenta	Factor de riesgo
Número de carriles	Presencia de más de un carril en cada sentido	Aumenta	Factor de riesgo

Tabla 52. Resumen de la influencia de las variables (regresión logística simple, curvas y rectas)

RECTAS

Variable independiente	Categorías	Riesgo de colisión frontal	Tipo de factor
Ancho de calzada	Ancho de calzada >5,99 m	Aumenta	Factor de riesgo
Anchura de carril	Ancho de carril >3,25	Aumenta	Factor de riesgo
Arcén	Presencia de arcén	Aumenta	Factor de riesgo
Barrera de seguridad	Presencia de Barrera	Aumenta	Factor de riesgo

Variable independiente	Categorías	Riesgo de colisión frontal	Tipo de factor
Hitos de arista	Presencia de Hitos de Arista	Aumenta	Factor de riesgo
Acera	Presencia de acera	Aumenta	Factor de riesgo
IMD	IMD superior a 500 vehículos al día	Aumenta	Factor de riesgo
Número de carriles	No existe asociación		

Tabla 53. Resumen de la influencia de las variables (regresión logística simple, rectas)

CURVAS

Variable independiente	Categorías	Riesgo de colisión frontal	Tipo de factor
Ancho de calzada	ancho de calzada >5,99 m	Aumenta	Factor de riesgo
Anchura de carril	Intermedia (3,25 m - 3,75m)	Aumenta	Factor de riesgo
Radio de curvatura	Radio < 750 m	Aumenta	Factor de riesgo
Arcén	No asociación		
Barrera de seguridad	Presencia de Barrera	Aumenta	Factor de riesgo
Hitos de arista	Presencia de Hitos de Arista	Aumenta	Factor de riesgo
Acera	No existe asociación		
IMD	IMD superior a 500 vehículos al día	Aumenta	Factor de riesgo
Número de carriles	Presencia de más de un carril en cada sentido	Aumenta	Factor de riesgo

Tabla 54. Resumen de la influencia de las variables (regresión logística simple, curvas)

CAPÍTULO 6.
MODELOS
MULTIVARIANTES:
RESULTADOS Y
VALIDACIÓN

A continuación, se plantean los modelos multivariantes, con el objetivo de estudiar la influencia conjunta de todas las variables incluidas en el estudio.

En el capítulo anterior se plantearon los modelos simples, en los que se estudiaba la fuerza de la asociación estadística entre una determinada variable independiente y el desenlace estudiado.

Una vez conocida la influencia de cada una de las variables, y habiendo determinado cuáles actúan como factores de riesgo cuando se consideran de manera independiente, es necesario plantear los modelos multivariantes, dado que es la única manera de analizar con mayor profundidad el evento estudiado, y de considerar la interrelación entre las diferentes variables incluidas en el estudio.

Por ello, se plantea una regresión logística multivariante, incluyendo todas las variables de interés, para los tres modelos que han sido distinguidos en el capítulo anterior:

- Modelo multivariante **curvas y rectas**
- Modelo multivariante **curvas**
- Modelo multivariante **rectas**

Estos modelos nos permitirán estudiar el comportamiento de cada variable cuando el resto se encuentran incluidas en el modelo. Para ello, se realizará una regresión logística binaria, en la que la variable “caso-control” actúa como variable respuesta, y el resto de variables explicativas como variables independientes.

Una vez planteados los modelos, se calcularán los Odds-ratio, de modo que se puedan detectar las asociaciones estadísticas entre los factores estudiados y el desenlace “colisión frontal”, y mediante el p-valor, se determinará si dicha asociación es estadísticamente significativa o no lo es.

Para facilitar los análisis posteriores, se incluye una tabla resumen con todas las variables explicativas del modelo, y la categoría de referencia que ha sido establecida al plantear los modelos, elegida de acuerdo a los análisis realizados en el capítulo previo.

VARIABLE	REFERENCIA
Ancho de calzada	<5,99 metros
Ancho de carril	<3,25 metros
Radio de curvatura	Recta
Arcén	Impracticable o insuficiente
Barrera	No
Hitos de arista	No

VARIABLE	REFERENCIA
Acera	No
IMD	<500
Número de carriles	Un solo carril en cada sentido

Tabla 55. Categorías de referencia para todas las variables del modelo

Finalmente, se procede a realizar la validación de los modelos obtenidos, usando para ello la validación cruzada “Leave-One-Out” y las curvas ROC. Para ello, se utilizó el software SAS. Los códigos de programación empleados para obtener los resultados se encuentran en el anexo 2.

Como ya fue explicado en el primer capítulo, la validación cruzada “Leave-One-Out” se utilizará para obtener el porcentaje correcto de clasificación con mayor precisión. Para ello, se realizarán tantas iteraciones como datos se disponga, realizando el test en cada iteración con un dato diferente. Si el test realizado es un acierto, se suma uno al número total de aciertos. Finalmente, se divide el número de aciertos entre el número de datos, obteniendo así el porcentaje de clasificación correcto.

Además, el Área Bajo la Curva ROC proporcionará una medida de la capacidad global de clasificación del modelo, y permitirán realizar una comparación de todos los modelos obtenidos, de modo que se podrá comprobar estadísticamente si existe diferencia o no entre ellos, fundamentando así la elección de los modelos definitivos.

Para obtener las Curvas ROC, así como la tabla de clasificación según el método LOOCV, se procederá a implementar los modelos multivariantes en SAS.

Una vez obtenidos los modelos finales, se podrán seleccionar aquellos modelos más adecuados en cuanto a la capacidad explicativa de los mismos, y al número de variables que incluyen.

6.1. RESULTADOS MULTIVARIANTES: ODDS-RATIO Y VALIDACIÓN DE LOS MODELOS

6.1.1. MODELO CURVAS Y RECTAS

REGRESIÓN LOGÍSTICA MULTIVARIANTE: ODDS RATIO

A continuación, se presenta la tabla obtenida en el modelo introducción:

	B	Error estándar	Wald	g	Sig.	Exp(B)	Ecuación
5,99m<Ancho de calzada<6,99m	1,880	,305	37,939	1	,000	6,556	X ₁
Ancho de calzada >6,99 m	2,597	,372	48,787	1	,000	13,428	X ₂
Arcén<1,5	-1,597	,271	34,792	1	,000	,203	X ₃
1,5 < Arcén < 2,5m	-2,690	,333	65,270	1	,000	,068	X ₄
Arcén > 2,5 m	-3,412	,565	36,438	1	,000	,033	X ₅
Barrera	-,051	,167	,091	1	,762	,951	X ₆
Hitos kilométricos	,849	,181	21,993	1	,000	2,338	X ₆
Acera	,695	,443	2,465	1	,116	2,004	X ₇
500<IMD < 2000	,628	,213	8,659	1	,003	,628	X ₈
2000<IMD<10000	1,766	,245	52,184	1	,000	1,766	X ₉
IMD>10000	2,944	,606	23,598	1	,000	2,944	X ₁₀
Número de Carriles	-,031	,587	,003	1	,958	,969	X ₁₁
3,25 m <Ancho de carril > 3,75m	,479	,227	4,444	1	,035	,479	X ₁₂
Ancho de carril > 3,75m	-1,078	,634	2,892	1	,089	-1,078	X ₁₃
Radio >750m	,102	,411	,062	1	,804	,102	X ₁₄
250m<Radio>750m	,186	,210	,783	1	,376	,186	X ₁₅
Radio<250	1,276	,204	39,154	1	,000	1,276	X ₁₆
Constante	-2,226	,236	89,329	1	,000	,108	

Tabla 56. Odds-ratio modelo introducción, curvas y rectas.

Tal y como se puede ver en la tabla anterior, en el modelo introducción se mantienen todas las variables introducidas en la ecuación, y se obtienen aquellas que son significativas mediante el p-valor.

En la siguiente tabla se resumen los factores que son significativos, así como el signo de la asociación estadística entre los mismos y el desenlace estudiado. También se incluye el porcentaje de clasificación según SPSS, o porcentaje aparente de clasificación.

Método De Selección De Variables	Porcentaje Correcto De Clasificación	Variables en la ecuación	Variables Significativas	Influencia
Introducción	75,4%	Todas las variables del modelo	Ancho de calzada	Factor de riesgo
			Arcén	Factor de protección
			Hitos	Factor de riesgo
			Ancho de carril intermedio (entre 3,25 m y 3,75 m)	Factor de riesgo
			Radio < 250 m.	Factor de riesgo
			IMD	Factor de riesgo

Tabla 57. Resultados modelo introducción, SPSS

VALIDACIÓN DEL MODELO: CURVA ROC Y LOOCV

Para validar el modelo anterior, se obtiene la curva ROC, y la tabla de clasificación según el método Leave-One Out.

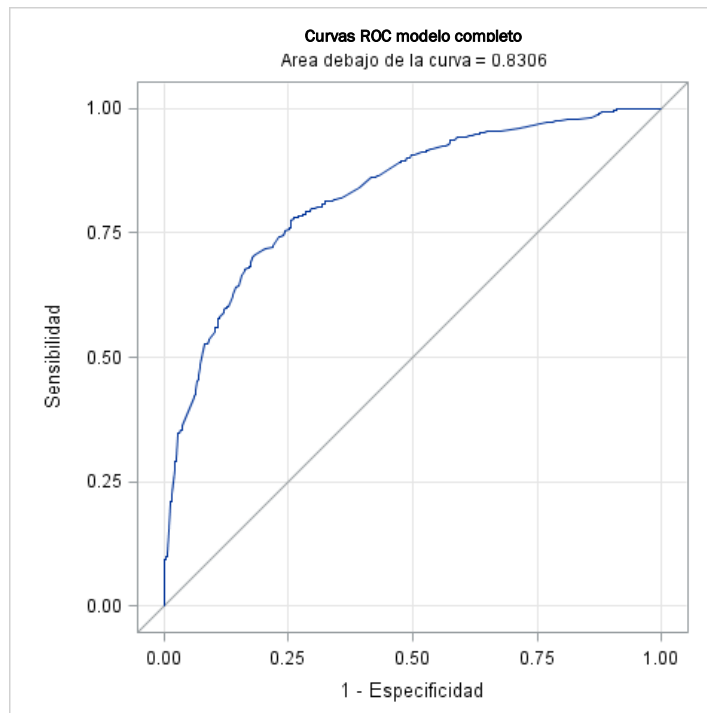


Ilustración 32. Curva ROC, introducción, curvas y rectas

La curva ROC proporciona una medida global de la precisión diagnóstica. Un área bajo la curva de 1 permite clasificar correctamente al 100% de los individuos, y se corresponde con una curva en la que no existe solapamiento entre las colisiones frontales, y la ausencia de colisiones frontales. En este caso, el área bajo la curva es 0,8306. Por tanto, aunque existe cierto solapamiento entre aquellos individuos que han experimentado el desenlace estudiado, y aquellos que no, es un área bajo la curva aceptable.

Dado que SPSS usa 0,5 como punto de corte por defecto, se fija el mismo punto de corte en SAS. A continuación, se muestra la comparativa del porcentaje correcto de clasificación según Leave-One-Out y el porcentaje de clasificación aparente según SPSS.

Punto de corte = 0'5	
SPSS (% de clasificación aparente)	LEAVE-ONE-OUT (% correcto de clasificación)
75'4%	75%

Como se ha explicado, el valor del porcentaje correcto de clasificación es más exacto si se calcula mediante la LOOCV. Por tanto, consideraremos que el porcentaje correcto de clasificación del modelo es un 75%. Sin embargo, se puede apreciar que no hay apenas diferencia entre ambos resultados.

La tabla de clasificación obtenida mostraba errores menores para otros puntos de corte. El valor óptimo obtenido para este modelo es 75,3%. Sin embargo, esto se conseguiría aumentando el punto de corte a 0,580, lo que a su vez conllevaría un aumento de los falsos negativos, y una disminución de los falsos positivos

MODELO BACKWARD/FORWARD

Tras implementar la regresión logística multivariante mediante los procedimientos de selección Backward y Forward, se comprobó que los resultados obtenidos eran idénticos. Por ello, nos referiremos a ellos como “modelo Backward/Forward”.

Consecuentemente, las Curvas ROC obtenidas, así como la tabla de clasificación, serán idénticas.

REGRESIÓN LOGÍSTICA MULTIVARIANTE: ODDS-RATIO

A continuación, se muestra la tabla correspondiente a la regresión logística multivariante que resulta al implementar el modelo Backward/Forward en el modelo completo.

	B	Error estándar	Wald	gl	Sig.	Exp(B)	Ecuación
5,99m<Ancho de calzada<6,99m	1,891	,305	38,383	1	,000	6,626	X ₁
Ancho de calzada >6,99 m	2,616	,372	49,535	1	,000	13,684	X ₂
Arcén<1,5	-1,628	,269	36,563	1	,000	,196	X ₃
1,5 < Arcén < 2,5m	-2,737	,331	68,485	1	,000	,065	X ₄
Arcén > 2,5 m	-3,455	,562	37,757	1	,000	,032	X ₅
Hitos kilométricos	,810	,176	21,121	1	,000	2,247	X ₆
500<IMD < 2000	,646	,213	9,215	1	,002	1,907	X ₇
2000<IMD<10000	1,789	,244	53,848	1	,000	5,983	X ₈
IMD>10000	2,947	,601	24,056	1	,000	19,053	X ₉
3,25 m <Ancho de carril > 3,75m	,473	,226	4,363	1	,037	1,605	X ₁₀
Ancho de carril > 3,75m	-1,082	,514	4,442	1	,035	,339	X ₁₁
Radio >750m	,077	,411	,035	1	,852	1,080	X ₁₂
250m<Radio>750m	,161	,207	,605	1	,437	1,174	X ₁₃
Radio<250	1,238	,193	41,015	1	,000	3,449	X ₁₄
Constante	-2,226	,236	89,329	1	,000	,108	

Tabla 58. Odds-Ratio modelo backward/forward, curvas y rectas

En la siguiente tabla, se resumen las variables significativas del modelo, así como el signo de la asociación estadística entre las mismas y el desenlace estudiado. Además, se presenta el porcentaje de clasificación aparente del modelo.

Método De Selección De Variables	Porcentaje Correcto De Clasificación	Variables en la ecuación	Variables Significativas	Influencia
Forward/ Backward Condicional, LR,Wald	75%	Ancho de calzada	Ancho de calzada	Factor de riesgo
		Arcén	Arcén	Factor de protección
		Hitos	Hitos	Factor de riesgo
		Ancho de carril	Ancho de carril intermedio (entre 3,25 m y 3,75 m)	Factor de riesgo
			Ancho de carril mayor que 3,75 metros	Factor de protección

Método De Selección De Variables	Porcentaje Correcto De Clasificación	Variables en la ecuación	Variables Significativas	Influencia
		Radio de curvatura	Ra dios <250 m	Factor de riesgo
		IMD	IMD	Factor de riesgo

Tabla 59. Resumen modelo backward/forward, curvas y rectas

VALIDACIÓN DEL MODELO: CURVA ROC Y LOOCV

Implementando el modelo anterior en SAS, se obtienen los resultados que se muestran a continuación.

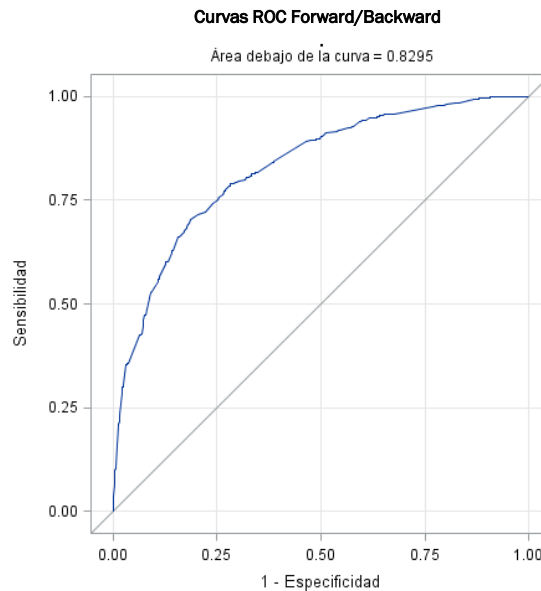


Ilustración 33. Curva ROC, Forward y Backward (Curvas y rectas)

El área bajo la curva toma el valor de 0,8295. Es muy similar a los resultados obtenidos para el modelo introducción, por lo que se puede intuir que las diferencias entre ambos modelos no serán significativas. De nuevo, existe cierto solapamiento entre los individuos que han experimentado el resultado estudiado y aquellos que no lo han hecho, sin embargo, es una cifra correcta.

Se procede a comparar el resultado obtenido con SPSS y el del método Leave-One-Out. Para ello, se obtiene el porcentaje de clasificación de acuerdo a ambos métodos.

Punto de Corte=0'5	
SPSS	LEAVE-ONE-OUT
% de clasificación aparente	% correcto de clasificación
75%	73'9%

De acuerdo con la tabla de comparación anterior, el porcentaje de clasificación correcto es 73,9%. De nuevo, no se aprecia gran diferencia entre ambos errores para un punto de corte de 0,5. Sin embargo, fijaremos el error real del modelo como 73,9%

La tabla de clasificación del método LOOCV permitía obtener para un punto de corte de 0,54, un porcentaje de clasificación de 75,2%. De nuevo, este resultado conllevaría aumentar los falsos negativos, y a disminuir los verdaderos positivos.

COMPARACIÓN DE MODELOS Y SELECCIÓN

Realizando la comparación los dos modelos considerados para las curvas y las rectas, se puede apreciar que apenas existe diferencia entre los modelos Backward y Forward y el modelo Introducción.

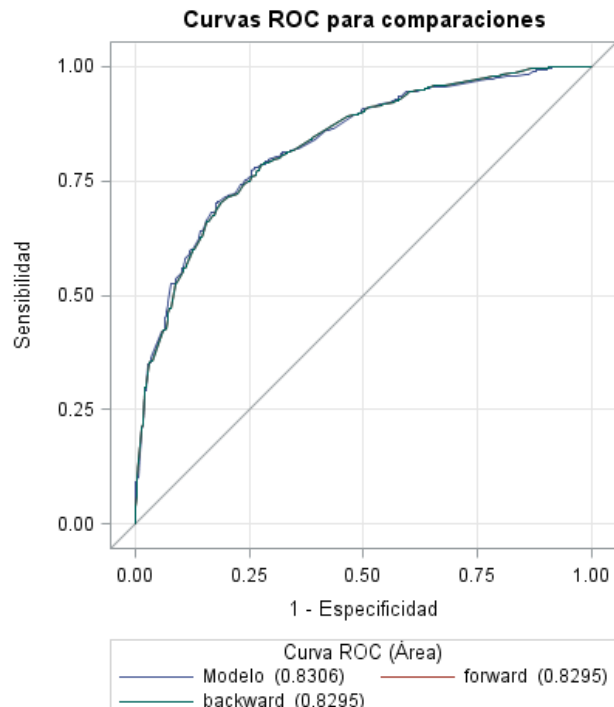


Ilustración 34. Comparativa de modelo, curvas y rectas

Si se resume la información obtenida en SAS y SPSS, se puede observar que tanto el área bajo la curva, como la forma de la curva son muy similares para ambos modelos. Lo mismo ocurre con los porcentajes correctos de clasificación.

	% de clasificación aparente	% Correcto de clasificación LOOCV	Área Bajo la curva
Introducción	75,4%	75%	0,8306
Backward Forward	75%	73,9%	0,8295

Tabla 60. Comparación de resultados, SAS y SPSS

En el código de SAS incluimos una línea que realice la comparativa de modelos, para así poder obtener evidencias estadísticas de la diferencia entre ellos.

La tabla comparativa mostrada compara los modelos Forward/Backward y el modelo Introducción. El p-valor es 0,4631, por lo que se puede afirmar que no existe diferencia entre los modelos Introducción y Backward/Forward. Por tanto, sería recomendable elegir el modelo Forward/Backward, pues incluye menos variables en el modelo, pero posee una capacidad explicativa muy similar.

	Estimador	Error estándar	Límites al 95% de Wald		Chi-cuadrado	Pr>Chicua
Modelo Introducción-Backward/Forward	0,00104	0,00142	-0,00174	0,00382	0,5384	0,4631

Tabla 61. Comparación estadística de modelos según SAS

6.1.2. MODELO RECTAS

A continuación, se presentan los resultados obtenidos para las ubicaciones que se corresponden con rectas.

MODELO INTRODUCCIÓN

REGRESIÓN LOGÍSTICA MULTIVARIANTE: ODDS-RATIO

	B	Error estándar	Wald	gl	Sig.	Exp(B)	Ecuación
5,99m<Ancho de calzada<6,99m	1,492	,401	13,857	1	,000	4,446	X ₁
Ancho de calzada >6,99 m	2,287	,477	23,004	1	,000	9,841	X ₂
Arcén<1,5	-1,112	,347	10,260	1	,001	,329	X ₃
1,5 < Arcén < 2,5m	-2,243	,417	28,968	1	,000	,106	X ₄
Arcén > 2,5 m	-2,967	,715	17,207	1	,000	,051	X ₅
Hitos kilométricos	,392	,235	2,768	1	,096	1,479	X ₆
500<IMD < 2000	,749	,296	6,407	1	,011	2,116	X ₇
2000<IMD<10000	1,832	,332	30,523	1	,000	6,248	X ₈
IMD>10000	3,028	,751	16,261	1	,000	20,648	X ₉
3,25 m <Ancho de carril > 3,75m	,612	,290	4,445	1	,035	1,845	X ₁₀
Ancho de carril > 3,75m	,284	1,029	,076	1	,783	1,328	X ₁₁
Número de carriles	-,947	,962	,969	1	,325	,388	X ₁₂
Barrera	,071	,229	,095	1	,758	1,073	X ₁₃
Acera	,776	,507	2,350	1	,125	2,174	X ₁₄
Constante	-2,318	,320	52,483	1	,000	,098	

Tabla 62. Odds-Ratio modelo introducción, solo rectas

En la siguiente tabla se muestran las variables incluidas en el estudio que han resultado significativas de acuerdo a la tabla anterior, y se muestra el porcentaje aparente de clasificación obtenido según SPSS.

Método De Selección De Variables	Porcentaje Correcto De Clasificación	Variables en la ecuación	Variables Significativas	Influencia
Introducción	75%	Todas las variables	Ancho de calzada	Factor de riesgo
			Arcén	Factor de protección
			Ancho de carril intermedio (entre 3,25 m y 3,75 m)	Factor de riesgo
			IMD	Factor de riesgo

Tabla 63. Resumen modelo Introducción, rectas

VALIDACIÓN DEL MODELO: CURVA ROC Y LOOCV

A continuación, se muestra la curva ROC, en la cual se representan los pares sensibilidad-(1-especificidad). El área bajo la curva es 0,8167. Por tanto, la capacidad de clasificación es buena.

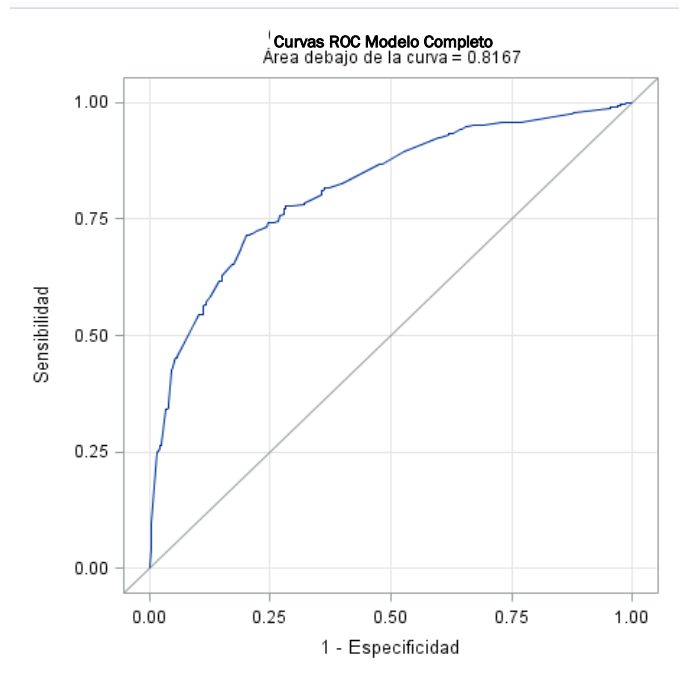


Ilustración 35. Curva ROC para modelo Introducción, rectas

En primer lugar, se compara el porcentaje aparente de clasificación según SPSS y el porcentaje correcto de clasificación de acuerdo al método LOOCV.

Punto de corte=0'5	
SPSS	LEAVE-ONE-OUT
% de clasificación aparente 75%	% correcto de clasificación 73'8%

De nuevo, el valor real es ligeramente inferior al valor calculado por el SPSS. En este caso, se llega a obtener un porcentaje correcto de clasificación de 74,4%, para un punto de corte de 0,48. Por tanto, se aumentarían los falsos positivos, y se disminuirían los falsos negativos.

MODELO BACKWARD, FORWARD

De nuevo, los modelos resultantes al implementar la regresión logística binaria en las ubicaciones rectas fueron idénticos para los procedimientos de selección Backward y Forward. A continuación, se muestran los Odds-Ratio obtenidos para este modelo.

	B	Error estándar	Wald	gl	Sig.	Exp(B)	Ecuación
5,99m<Ancho de calzada<6,99m	1,537	,399	14,811	1	,000	4,651	X ₁
Ancho de calzada >6,99 m	2,451	,469	27,289	1	,000	11,603	X ₂
Arcén<1,5	-1,052	,340	9,579	1	,002	,349	X ₃
1,5 < Arcén < 2,5m	-2,130	,403	27,973	1	,000	,119	X ₄
Arcén > 2,5 m	-2,750	,695	15,672	1	,000	,064	X ₅
500<IMD < 2000	,789	,294	7,216	1	,007	2,200	X ₆
2000<IMD<10000	1,896	,329	33,151	1	,000	6,660	X ₇
IMD>10000	2,860	,726	15,525	1	,000	17,467	X ₈
3,25 m <Ancho de carril > 3,75m	,643	,286	5,033	1	,025	1,902	X ₉
Ancho de carril > 3,75m	-,209	,852	,060	1	,806	,811	X ₁₀
Constante	-2,303	,319	52,194	1	,000	,100	

Tabla 64. Odds-Ratio modelo backward/forward, solo rectas

En la siguiente tabla se resume la influencia de las variables cuya presencia modifica el riesgo de colisión frontal, y se recoge el porcentaje aparente de clasificación.

Método De Selección De Variables	Porcentaje Correcto De Clasificación	Variables en la ecuación	Variables Significativas	Influencia
Forward	74,2%	Ancho de calzada	Ancho de calzada	Factor de riesgo
		Arcén	Arcén	Factor de protección
		IMD	IMD	Factor de riesgo
		Ancho de carril	Ancho de carril intermedio (entre 3,25 y 3,75)	Factor de riesgo

Tabla 65. Resumen modelo Backward/Forward, solo rectas

VALIDACIÓN DEL MODELO: CURVA ROC Y LOOCV

La curva ROC que resulta al representar los pares sensibilidad y (1-especificidad) para el modelo anterior es la siguiente:

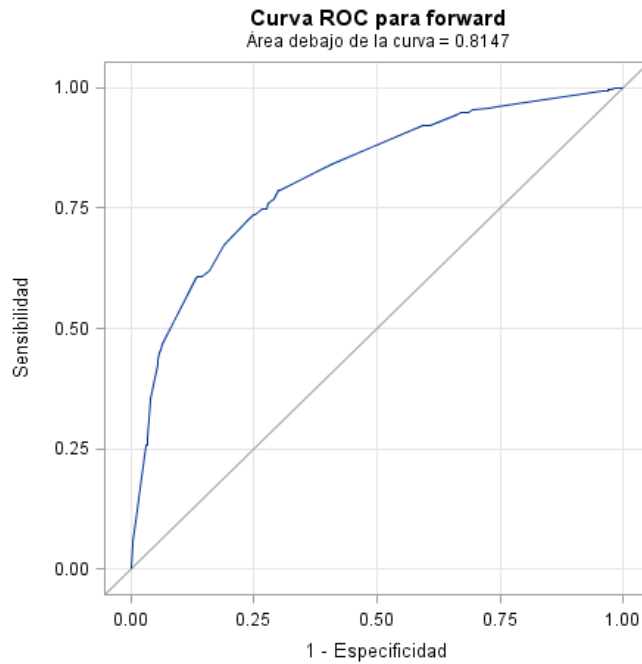


Ilustración 36. Curva ROC, Backward/Forward, Rectas

Los porcentajes de clasificación obtenidos mediante SPSS y LOOCV para un punto de corte de 0,5 son muy similares. El valor real, aportado por el método LOOCV, implica que el 73,1% de los casos se clasifican correctamente.

% Correcto de clasificación	
SPSS	LEAVE-ONE-OUT
% clasificación aparente	% correcto de clasificación
74,2%	73,1%

En la tabla de clasificación del método LOOCV, se pudo observar que el máximo porcentaje correcto de clasificación es 74,7%, y se corresponde con un punto de corte de 0,460.

COMPARACIÓN ENTRE MODELOS Y SELECCIÓN

El gráfico de superposición de las curvas muestra que los modelos son muy similares en todos los casos. Por tanto, se puede intuir que no habrá diferencias entre ellos.

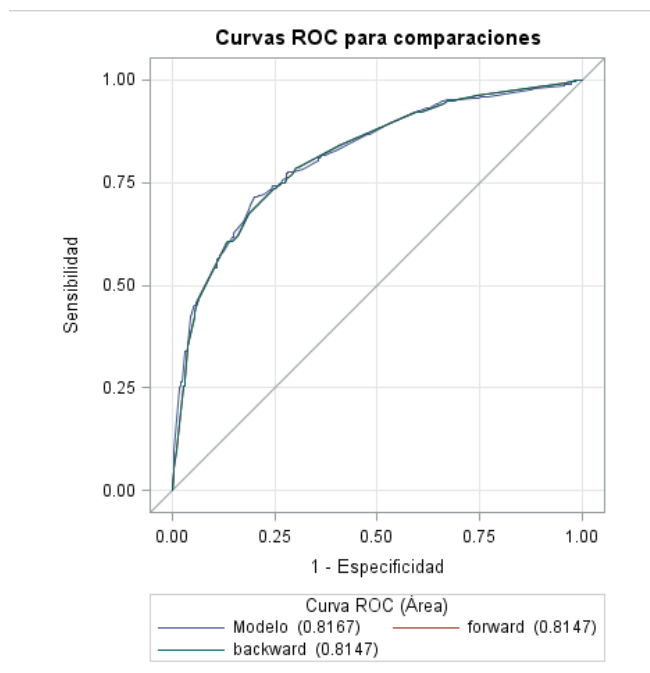


Ilustración 37. Comparación curvas ROC, rectas.

De nuevo, la tabla de comparación de modelos proporcionada por SAS muestra que no existe diferencia entre los modelos, dado que el p-valor toma el valor de 0,6572.

	Estimador	Error estándar	Límites al 95% de Wald		Chi-cuadrado	Pr>Chicua
Modelo Intermedio-Backward	0,00198	0,00446	-0,00676	0,0107	0,1969	0,6572

Tabla 66. Comparativa de modelos según SAS, rectas

Comparando los resultados obtenidos para los diferentes modelos, se decide que se seleccionará el modelo Backward/Forward en el caso de las rectas, dado que el área bajo la curva es casi idéntica en ambos modelos, y los porcentajes de clasificación, también son similares. Sin embargo, el modelo Backward/Forward incluye un menor número de variables.

	% Correcto de Clasificación aparente SPSS	% Correcto de clasificación LOOCV	Área Bajo la curva
Introducción	75%	73,8%	0,8167
Backward Forward	74,2%	73,1%	0,8147

Tabla 67. Comparación de modelos SPSS Y SAS, curvas

6.1.3. MODELO CURVAS

MODELO INTRODUCCIÓN

REGRESIÓN LOGÍSTICA MULTIVARIANTE: ODDS-RATIO

A continuación, se obtiene la tabla de la regresión logística multivariante del modelo “introducción” en las ubicaciones curvas.

	B	Error estándar	Wald	gl	Sig.	Exp(B)	Ecuación
5,99m<Ancho de calzada<6,99m	2,681	,556	23,290	1	,000	14,601	X_1
Ancho de calzada >6,99 m	3,462	,675	26,333	1	,000	31,881	X_2
Arcén<1,5	-2,518	,507	24,660	1	,000	,081	X_3
1,5 < Arcén < 2,5m	-3,441	,620	30,799	1	,000	,032	X_4
Arcén > 2,5 m	-4,366	,991	19,411	1	,000	,013	X_5
Hitos kilométricos	1,622	,306	28,134	1	,000	5,065	X_6
500<IMD < 2000	,502	,326	2,366	1	,124	1,652	X_7
2000<IMD<10000	1,865	,388	23,118	1	,000	6,459	X_8
IMD>10000	3,161	1,147	7,602	1	,006	23,599	X_9
3,25 m <Ancho de carril > 3,75m	,058	,383	,023	1	,880	1,059	X_{10}
Ancho de carril > 3,75m	-2,191	,878	6,225	1	,013	,112	X_{11}
250m<Radio>750m	,173	,466	,138	1	,710	1,189	X_{12}
Radio <250m	1,472	,498	8,739	1	,003	4,358	X_{13}
Número de carriles	,472	,839	,316	1	,574	1,603	X_{14}
Barrera	-,269	,256	1,109	1	,292	,764	X_{15}
Acera	-,030	,978	,001	1	,975	,970	X_{16}
Constante	-2,326	,573	16,458	1	,000	,098	

Tabla 68. Odds-Ratio modelo introducción, curvas

Al igual que en los modelos anteriores, resumimos en una tabla el porcentaje de clasificación aparente, así como las variables significativas, y el tipo de asociación entre las mismas y la variable respuesta.

Método De Selección De Variables	Porcentaje correcto de clasificación	Variables en la ecuación	Variables Significativas	Influencia
Introducción	78,8%	Variables en la ecuación	Ancho de calzada	Factor de riesgo
			Arcén	Factor de protección
			Hitos	Factor de riesgo
			Ancho de carril > 3,75 metros	Factor de protección
			Radio < 250 m.	Factor de riesgo
			IMD>2000	Factor de riesgo

Tabla 69. Resumen modelo introducción, curvas

VALIDACIÓN DEL MODELO: CURVA ROC Y LOOCV

La curva ROC obtenida al implementar el modelo anterior en SAS se muestra en la siguiente ilustración.

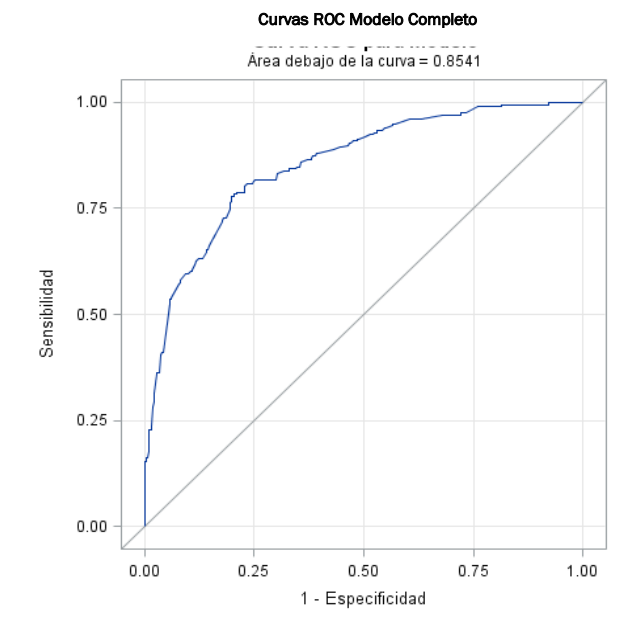


Ilustración 38. Curva ROC, modelo introducción, curvas

La curva ROC obtenida muestra que el área bajo la curva es 0,8541, lo cual indica una exactitud global del modelo adecuada, a pesar de que existe cierto solapamiento entre las ubicaciones en las que tienen lugar colisiones frontales y las que no.

Si se comparan el porcentaje aparente de clasificación y el porcentaje correcto de clasificación según el método LOOCV, se puede observar que no existe apenas diferencia.

Punto de Corte=0'5	
SPSS	LEAVE-ONE-OUT
% de clasificación aparente	% corecto de clasificación
78,8%	76,5%

Esto indica que el porcentaje correcto de clasificación es 76,5%, para el punto de corte 0,5. De nuevo la tabla de clasificación del método LOOCV mostraba % de clasificación correctos más elevados para otros puntos de corte.

MODELO BACKWARD/FORWARD

Al igual que en el caso del modelo completo, y de modelo “solo rectas”, los modelos Backward/Forward son idénticos.

REGRESIÓN LOGÍSTICA MULTIVARIANTE: ODDS-RATIO

	B	Error estándar	Wald	gl	Sig.	Exp(B)	Ecuación
5,99m<Ancho de calzada<6,99m	2,667	0,549	23,608	1	0	14,398	X ₁
Ancho de calzada >6,99 m	3,438	0,669	26,369	1	0	31,118	X ₂
Arcén<1,5	-1,628	0,269	36,563	1	0	0,196	X ₃
1,5 < Arcén < 2,5m	-2,5	0,498	25,202	1	0	0,082	X ₄
Arcén > 2,5 m	-3,451	0,61	31,977	1	0	0,032	X ₅
Hitos kilométricos	1,564	0,298	27,607	1	0	4,776	X ₆
500<IMD < 2000	0,497	0,324	2,342	1	0,126	1,643	X ₇
2000<IMD<10000	1,861	0,384	23,44	1	0	6,429	X ₈
IMD>10000	3,251	1,143	8,092	1	0,004	25,818	X ₉

	B	Error estándar	Wald	gl	Sig.	Exp(B)	Ecuación
3,25 m <Ancho de carril > 3,75m	0,034	0,383	0,008	1	0,93	1,034	$X_{1,0}$
Ancho de carril > 3,75m	-1,974	0,731	7,303	1	0,007	0,139	$X_{1,1}$
250m<Radio>750m	0,157	0,464	0,115	1	0,735	1,17	$X_{1,2}$
Radio <250m	1,421	0,493	8,307	1	0,004	4,14	$X_{1,3}$
Constante	-2,367	0,57	17,222	1	0	0,094	

Tabla 70. Regresión logística multivariante, Backward/Forward, curvas

Se obtiene la siguiente curva ROC:

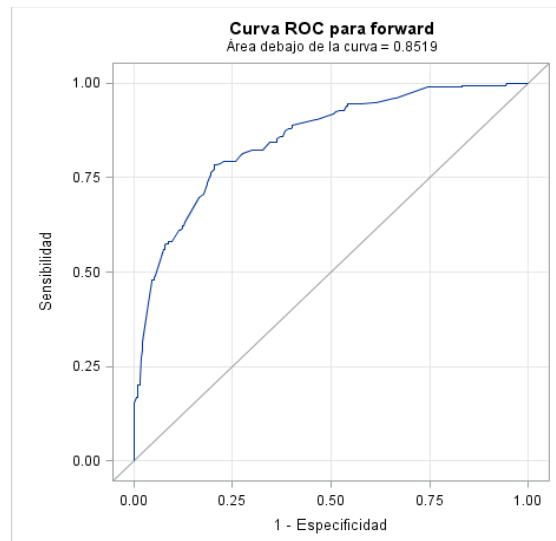


Ilustración 39. Modelo Backward/Forward, curvas

El área bajo la curva se comporta como un indicador de la fiabilidad global del modelo. Dado que el área bajo la curva es 0,8519, presenta fiabilidad global correcta.

Mediante la tabla de clasificación, se obtiene un valor del porcentaje correcto de clasificación de 77,1% para un punto de corte de 0,5. Comparando los resultados obtenidos mediante LOOCV y SPSS, la diferencia es muy pequeña, sin embargo, dada la fiabilidad del método LOOCV, consideraremos el error estimado por este método como error como real.

Punto de corte=0'5	
SPSS	LEAVE-ONE-OUT
% de clasificación aparente 78,8%	% correcto de clasificación 77,1%

Modificando el punto de corte se podría obtener una precisión diagnóstica de 78,2%.

COMPARATIVA DE MODELOS CURVAS

La siguiente gráfica superpone las curvas obtenidas para los distintos modelos.

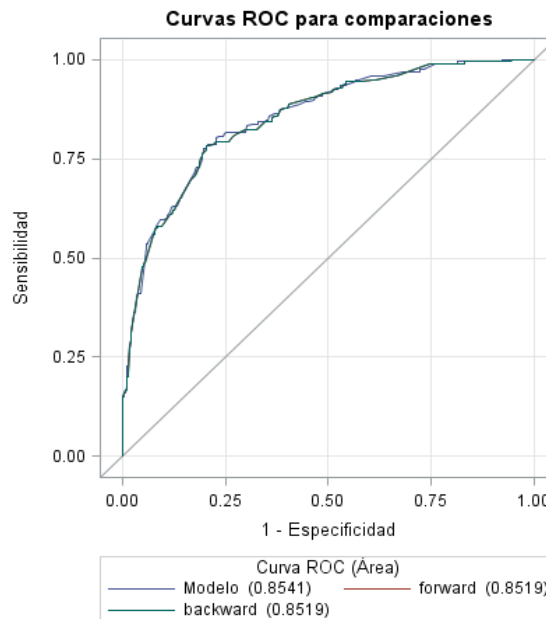


Ilustración 40: Comparativa de modelos, solo curvas

Comparando el modelo Backward/Forward con el modelo introducción, se puede apreciar que apenas existe diferencia entre el área bajo la curva de los diferentes modelos, y que ambas son muy similares en los porcentajes correctos de clasificación.

	% Correcto de Clasificación SPSS	% Correcto de clasificación LOOCV	Área Bajo la curva
Introducción	78,8%	76,5%	0,8541
Backward Forward	78,8%	77,1%	0,8519

Tabla 71 Comparación de modelos SPSS Y SAS, curvas

De nuevo, la tabla de comparación de modelos proporcionada por SAS muestra que no existe diferencia entre los modelos, dado que el p-valor toma el valor de 0,3354

	Estimador	Error estándar	Límites al 95% de Wald		Chi-cuadrado	Pr>Chicua
Modelo Intermedio-Backward	0,00217	0,00225	-0,00225	0,00659	0,9279	0.3354

Tabla 72. Comparación estadística de modelos según SAS, curvas

Por tanto, la decisión final sería optar por el modelo Backward/Forward, ya que el porcentaje correcto de clasificación de acuerdo al método LOOCV es superior al del modelo introducción (77,1% frente a 76,5%). Además, incluye menos variables en el modelo.

6.2. MODELOS FINALES SELECCIONADOS

Recopilando todas las decisiones tomadas en los apartados anteriores, los modelos finales elegidos para la definición del evento estudiado serían:

- Modelo “completo”, curvas y rectas: Backward/forward
- Modelo “solo rectas”: Backward/Forward
- Modelo “solo curvas”: Backward/Forward

6.2.1. MODELO FINAL “COMPLETO” PARA CURVAS Y RECTAS: MODELO FORWARD/BACKWARD

Modelo	% Correcto de clasificación (LOOCV)	Variables en la ecuación	Variables significativas
Modelo Forward/Backward	73,9%	Ancho de carril	Ancho de carril
		Ancho de calzada	Ancho de calzada
		Radio de curvatura	Radio de curvatura
		IMD	IMD
		Arcén	Arcén
		Hitos kilométricos	Hitos kilométricos

Tabla 73. Características modelo curvas y rectas

Para obtener las ecuaciones del modelo logístico, usamos la tabla de la regresión logística binaria obtenida para el modelo Forward/Backward: Tabla 58. Odds-Ratio modelo backward/forward, curvas y rectas.

Por tanto, la ecuación del modelo de regresión obtenido sería:

$$\begin{aligned} \text{Logit}(P) &= \ln\left(\frac{P}{1-P}\right) = \\ &= -2,226 + 1,891X_1 + 2,616X_2 - 1,628X_3 - 2,737X_4 \\ &\quad - 3,445X_5 + 0,81X_6 + 0,646X_7 + 1,789X_8 + 2,947X_9 \\ &\quad + 0,473X_{10} - 1,082X_{11} + 0,077X_{12} + 0,161X_{13} + 1,238X_{14} \end{aligned}$$

Ecuación 9. Ecuación final del modelo logístico, curvas y rectas

Dado que todas las variables incluidas en el estudio son de tipo categórico, las variables independientes ($X_1 \dots X_{14}$) son variables dicotómicas, cuyos valores para un determinado individuo pueden ser:

- “0” cuando el individuo que se pretende evaluar no posee la característica de la categoría a la que hace referencia la variable.
- “1” cuando el individuo evaluado presenta la característica de la variable a la que hacen referencia.

Por tanto, la probabilidad de que en una determinada ubicación ocurra una colisión frontal, en función de las variables estudiadas, es la siguiente:

$$P = \frac{1}{1 + e^{-(-2,226 + 1,891X_1 + 2,616X_2 - 1,628X_3 + \dots + 0,161X_{13} + 1,238X_{14})}}$$

Analizando la influencia que ejercen las variables del modelo, se pueden extraer las siguientes conclusiones:

- Los factores de riesgo con mayor influencia sobre el desenlace estudiado son la **IMD** superior a 10000 vehículos por día, y **calzada cuya anchura** es superior a 6,99 metros. La influencia de ambos factores supone un aumento del riesgo de colisión frontal de 260 veces.

$$OR = e^{\beta_2} * e^{\beta_9} = 13,684 * 19,053 = 260,72$$

Se puede apreciar un gran aumento del riesgo de colisión frontal cuando se junta la influencia de ambos factores, manteniendo el resto

constantes. La explicación podría radicar en que una IMD tan elevada conlleva un gran número de adelantamientos, los cuales podrían estar directamente relacionados con las colisiones frontales. Además, el ancho de carril podría influir en la velocidad de los vehículos, pues los conductores disminuirán la precaución.

- El **arcén** se comporta como factor de protección, a pesar de que, en el estudio de las variables individuales, el arcén era un factor cuya presencia aumentaba el riesgo de colisión frontal. Por tanto, la presencia del resto de variables ha modificado la influencia de la variable arcén. Cuanto mayor es el arcén, más se reduce el riesgo de ocurrencia de las colisiones frontales. De hecho, la presencia de un arcén mayor que 2,5 metros reduce casi 32 veces el riesgo de colisiones frontales.

$$e^{\beta_5 X_5} = e^{-3,455 X_5} = 0,032 = \frac{1}{31,25}$$

- El **ancho de carril** se comporta como factor de riesgo cuando mide entre 3,25 y 3,75 metros, pues aumenta 1,6 veces el riesgo de colisión frontal (OR=1,605). Sin embargo, calzadas mayores de 3,75 metros actúan como factor de protección, reduciendo casi la tercera parte el riesgo de que se produzcan colisiones frontales (OR=0,339)
- El **radio de curvatura** solo influye en las colisiones frontales en el caso de que se trate de curvas muy cerradas, pues en el resto de casos el p-valor es mayor que 0,1. Si el radio es inferior a 250 metros, el riesgo que se produzcan colisiones frontales aumentaría 3,5 veces.

$$OR = e^{\beta_{13} X_{13}} = e^{1,238 X_{13}} = 3,449$$

- Los **hitos kilométricos** aumentan el riesgo de colisiones frontales 2,247 veces. Esto podría deberse a la percepción que los conductores tienen de la carretera, pero habría que estudiar en profundidad la influencia que esta variable ejerce sobre las colisiones frontales.

A continuación, se presentan algunos casos de ubicaciones particulares:

- En primer lugar, consideraremos el peor caso posible, en el que se dan una serie de condiciones que no aparecerán en la práctica. En una ubicación con calzada mayor que 6,99, arcén inexistente, con presencia

de hitos kilométricos, un IMD superior a 10000 vehículos al día, ancho de carril entre 3,25 y 3,75, y radio menor que 250 metros es la siguiente:

$$P = \frac{1}{1 + e^{-(-2,226 + 2,616 + 0,81 + 2,947 + 0,473 + 1,238)}} = 0,9971512$$

Por tanto, si en una determinada ubicación se juntase el efecto de todas las variables mencionadas, el riesgo de experimentar el resultado “colisión frontal” sería muy elevada.

Mediante el Odds-ratio, se puede cuantificar el aumento del riesgo para las condiciones expuestas anteriormente.

$$OR = e^{2,616} * e^{0,81} * e^{2,947} * e^{0,473} * e^{1,238} = 3242,18$$

Por tanto, para las condiciones anteriores, el riesgo de experimentar el desenlace estudiado incrementa más de 3000 veces.

- Sin embargo, es muy extraño que se den todas estas condiciones simultáneamente. Por ejemplo, este tipo de carreteras de gran anchura, y en las que el flujo del tráfico es elevado, suelen presentar un arcén ancho. Como ha sido explicado anteriormente, el arcén constituye un factor de protección. Sería más real, por ejemplo, que se diesen las siguientes condiciones: curva de radio mayor que 750 metros, con ancho de carril entre 3,25 y 3,75, ancho de calzada mayor que 6,99 metros, arcén entre 1,5 y 2,5 metros, presencia de hitos y un IMD entre 2000 y 10000 vehículos/día.

$$P = \frac{1}{1 + e^{-(-2,226 + 2,616 - 2,737 + 0,81 + 1,789 + 0,473 + 0,077)}} = 0,69$$

$$OR = e^{2,616} * e^{2,737} * e^{0,81} * e^{1,789} * e^{(0,473)} * e^{(0,077)} = 20,65$$

Por tanto, la presencia de todos los factores mencionados en este punto supondría un aumento de 20 veces del riesgo de que se produzcan colisiones frontales (OR=20,65)

- Considerando una ubicación sin hitos kilométricos, una anchura de calzada mayor de 6,99, un arcén de entre 1,5 y 2,5 metros, un radio de curvatura superior a 250 metros, un ancho de carril de entre 3,25 y 3,75

metros, y una IMD entre 2000 y 10000 vehículos por día, se obtendría el siguiente riesgo:

$$e^{\beta_2 X_2} e^{\beta_4 X_4} e^{\beta_8 X_8} e^{\beta_{10} X_{10}} e^{\beta_{13} X_{13}} = e^{2,616} e^{-2,737} e^{1,789} e^{0,473} e^{0,161} = 9,99$$

Si esa misma ubicación presentase un ancho de calzada inferior a 5,99, el aumento de riesgo obtenido si se mantienen el resto de variables constantes es el siguiente:

$$e^{\beta_2 X_2} e^{\beta_4 X_4} e^{\beta_8 X_8} e^{\beta_{10} X_{10}} e^{\beta_{13} X_{13}} = e^{-2,737} e^{1,789} e^{0,473} e^{0,161} = 0,73$$

En definitiva, se puede apreciar que el criterio de priorización de carreteras cuya anchura supera los 6,99 metros no solo es adecuado desde el punto de vista económico y de facilidad de modificación de las carreteras, sino que, además, las carreteras cuya anchura supera los 6,99 se consideran de alto peligro comparándolas con las carreteras estrechas.

6.2.2. MODELO FINAL PARA LAS RECTAS, FORWARD/BACKWARD

A continuación, se muestra una tabla resumen del modelo:

Modelo	% Correcto de clasificación (LOOCV)	Variables en la ecuación	Variables significativas
Modelo Forward/Backward	73,1%	Ancho de carril	Ancho de carril
		Ancho de calzada	Ancho de calzada
		IMD	IMD
		Arcén	Arcén

Tabla 74. Características del modelo final para las rectas

Utilizaremos los resultados mostrados en la Tabla 64. Odds-Ratio modelo backward/forward, solo rectas para obtener la ecuación del modelo. Sustituyendo los coeficientes de los términos independientes, se obtiene la ecuación del modelo logístico.

$$\begin{aligned} \text{Logit}(P) &= \ln\left(\frac{P}{1-P}\right) = \\ &= -2,303 + 1,537X_1 + 2,451X_2 - 1,052X_3 - 2,130X_4 - 2,75X_5 \\ &\quad + 0,789X_6 + 1,896X_7 + 2,86X_8 + 0,643X_9 - 0,209X_{10} \end{aligned}$$

Ecuación 10. Ecuación final del modelo logístico, rectas

Operando con la expresión anterior, se obtiene la siguiente expresión:

$$\frac{P}{1-P} = e^{-2,303+1,537X_1+2,451X_2-1,052X_4+\dots+0,643X_9-0,209X_{10}} = e^{-2,303} e^{1,537X_1} \dots e^{0,209X_{10}}$$

De acuerdo a la tabla del modelo, y a las ecuaciones, se puede afirmar que:

- El **ancho de calzada** se mantiene como un factor de riesgo importante. Cuanto mayor es el ancho de calzada, más aumenta el riesgo de que tenga lugar una colisión frontal. En el caso de que el ancho de calzada sea mayor que 6,99, se aumenta el riesgo de colisión frontal 11 veces con respecto a anchuras de calzada menores de 5,99 metros, si el resto de variables permanecen constantes (OR=11,603).
- Los **hitos de arista** dejan de comportarse como un factor de riesgo para las ubicaciones rectas.
- **La IMD** es el factor que más riesgo ejerce sobre las ubicaciones rectas. Además, cuanto mayor es el flujo de vehículos, más aumenta el riesgo. Una intensidad media diaria superior a 10000 vehículos al día supone aumentar el riesgo de colisión frontal más de 17 veces con respecto a un a IMD menor que 500 vehículos al día.
- Solo existe asociación estadística entre el **ancho de carril** y el desenlace estudiado cuando este mide entre 3,25 y 3,75 metros. En este caso, el ancho de carril constituye un factor de riesgo que aumenta casi dos veces el riesgo de colisión frontal.
- Al igual que en los otros dos modelos, el arcén se comporta como un factor de protección. El riesgo de colisión frontal se reduce más cuanto mayor es la anchura del arcén. En el caso de que la anchura del arcén sea mayor que 2,5 metros, el riesgo de que ocurra una colisión frontal disminuye 15,64 veces con respecto a la ausencia de arcén, siempre y cuando el resto de variables de la ecuación permanezcan constantes.

Sobre el modelo, se pueden obtener las siguientes conclusiones:

- Considerando una recta, con una anchura de calzada mayor de 6,99, un arcén de entre 1,5 y 2,5 metros, un ancho de carril de entre 3,25 y

3,75 metros y una IMD entre 2000 y 10000 vehículos por día, se obtendría el siguiente aumento del riesgo:

$$e^{\beta_2 X_2} e^{\beta_4 X_4} e^{\beta_8 X_8} e^{\beta_9 X_9} = e^{2,451} e^{-2,75} e^{1,896} e^{0,643} = 9,39$$

Si esa misma ubicación presentase un ancho de calzada inferior a 5,99, el aumento del riesgo obtenido si se mantienen el resto de variables constantes es el siguiente:

$$e^{\beta_2 X_2} e^{\beta_4 X_4} e^{\beta_7 X_7} = e^{-2,75} e^{1,896} e^{0,643} = 0,81$$

6.2.3. MODELO FINAL PARA LAS CURVAS, FORWARD/BACKWARD

Como ha sido justificado anteriormente, el modelo idóneo para explicar las colisiones frontales en las ubicaciones curvas es el modelo Forward/Backward (ambos eran iguales). En la siguiente tabla se resumen las características básicas de modelo.

Modelo	% Correcto de clasificación (LOOCV)	VARIABLES en la ecuación	VARIABLES significativas
Modelo Forward/Backward	77,1%	Ancho de carril	Ancho de carril
		Ancho de calzada	Ancho de calzada
		Radio de curvatura	Radio de curvatura
		IMD	IMD
		Arcén	Arcén
		Hitos kilométricos	Hitos kilométricos

Tabla 75. Características del modelo final para las curvas

El modelo final para las curvas contiene las mismas variables en la ecuación que el modelo completo (curvas y rectas). Sin embargo, se puede apreciar que en este caso la capacidad correcta de clasificación es superior (aumenta de 73,9% a 77,1%). Por tanto, se podría decir que la capacidad de clasificación global del modelo es superior para las curvas que para las rectas.

A partir de la Tabla 70. Regresión logística multivariante, Backward/Forward, curvas, y conociendo la forma de la ecuación del modelo logístico, se puede obtener la ecuación logística para las ubicaciones curvas.

$$\begin{aligned} \text{Logit}(P) &= \ln\left(\frac{P}{1-P}\right) = \\ &= -2,367 + 3,438X_1 - 1,628X_2 - 2,5X_4 - 3,451X_5 + 1,564X_6 \\ &+ 0,497X_7 + 1,861X_8 + 3,251X_9 + 0,034X_{10} - 1,974X_{11} \\ &+ 0,157X_{12} + 2,421X_{13} \end{aligned}$$

Ecuación 11. Ecuación final del modelo logístico, curvas

Operando con la expresión anterior, se obtiene la siguiente expresión:

$$\begin{aligned} \frac{P}{1-P} &= e^{(-2,367+3,438X_1-1,628X_2--2,5X_4+\dots+3,251X_9+0,157X_{12}+2,421X_{13})} = \\ &= e^{-2,367} e^{3,438X_1} \dots e^{2,421X_{13}} \end{aligned}$$

Las conclusiones obtenidas considerando el modelo “solo curvas” son muy similares a las del modelo completo.

Para las variables individuales, se puede concluir:

- A diferencia del modelo “curvas y rectas”, **la IMD** solo es un factor de riesgo en las ubicaciones curvas cuando el flujo de vehículos es superior a 2000 vehículos al día. Cuanto mayor es el IMD, mayor es el riesgo de que se produzcan colisiones frontales. Considerando el valor extremo de 10000 vehículos al día, se obtiene un aumento del riesgo de casi 26 veces (OR=25,818). De nuevo, el motivo podría deberse a la gran cantidad de adelantamientos.
- El **ancho de calzada** superior a 6,99 metros es el factor de riesgo que ejerce mayor influencia sobre el evento estudiado, pues aumenta más de 31 veces el riesgo de que se produzcan colisiones frontales,

$$e^{\beta_2 X_2} = e^{3,438 X_2} = 31,118$$

La explicación de este brusco aumento del riesgo podría deberse a que en carreteras más anchas la precaución de los conductores es menor, lo que se traduce en mayor velocidad de

circulación. Además, el tráfico es mayor, y se producen mayor número de adelantamientos.

- En este caso, el hecho de un **carril** posea una anchura de entre 3,25 metros y 3,75 metros no constituye un factor de riesgo. Sin embargo, los carriles superiores a 3,75 metros se mantienen como factor de protección. El OR asociado es de 0,139, por lo que la presencia de carril de más de 3,75 metros de ancho reduce el riesgo de colisiones frontales una séptima parte.
- El **arcén** continúa comportándose como un factor de protección en presencia del resto de variables. Se repite el hecho de que, a mayor anchura del arcén, mayor es la protección que ejerce este elemento sobre el desarrollo del evento estudiado. La presencia de arcones de más de 2,5 metros disminuye 12,2 veces el riesgo de colisiones frontales (OR=0,082).
- De nuevo, **el radio de curvatura** constituye un factor de riesgo solo cuando las curvas son muy cerradas (radio<250). Este tipo de curvas aumentan más de 4 veces el riesgo de colisiones frontales si el resto de variables permanecen constantes.
- La presencia **de hitos kilométricos** aumenta el riesgo de colisiones frontales casi 5 veces con respecto a la ausencia de ellos.

Para el modelo curvas se obtuvieron las siguientes conclusiones:

- Considerando una ubicación sin hitos kilométricos, una anchura de calzada mayor de 6,99, un arcén de entre 1,5 y 2,5 metros, un radio de curvatura superior a 250 metros, y un ancho de carril de entre 3,25 y 3,75 metros y una IMD entre 2000 y 10000 vehículos por día, se obtendría el siguiente aumento del riesgo:

$$e^{\beta_2 X_2} e^{\beta_4 X_4} e^{\beta_8 X_8} e^{\beta_{10} X_{10}} e^{\beta_{13} X_{13}} = e^{2,667} e^{-2,5} e^{1,861} e^{0,034} e^{0,157} = 9,2$$

Si esa misma ubicación presentase un ancho de calzada inferior a 5,99, el aumento de riesgo obtenido si se mantienen el resto de variables constantes es el siguiente:

$$e^{\beta_2 X_2} e^{\beta_4 X_4} e^{\beta_8 X_8} e^{\beta_{10} X_{10}} e^{\beta_{13} X_{13}} = e^{-2,5} e^{1,861} e^{0,034} e^{0,157} = 0,64$$

De nuevo, se puede apreciar la gran influencia que ejerce el ancho de carril mayor que 6,99, por lo que de nuevo se cumple que, además de ser utilizado como criterio de priorización, ejerce una gran influencia sobre las colisiones frontales con respecto a anchuras de carril inferiores a 5,99 metros.

CONCLUSIONES Y LÍNEAS FUTURAS DEL TRABAJO

Como ha sido mencionado en el apartado “objetivos del proyecto”, la principal finalidad del presente Trabajo Fin de Grado consistía en caracterizar las localizaciones en las que ocurren colisiones frontales. Para ello, se ha utilizado la metodología “caso-control”, y se han llevado a cabo tres estudios que difieren en las localizaciones o escenarios que se han incluido en los mismos:

1. Modelo completo, en el que se incluyen curvas y rectas
2. Modelos rectas, en el que se incluyen las ubicaciones que se corresponden con rectas
3. Modelos curvas, en el que se incluyen las ubicaciones que se corresponden con curvas.

Tras la utilización de la metodología “caso-control” sobre los tres modelos definidos, se obtuvieron una serie de factores de riesgo y factores de protección en cada uno de los escenarios. Dada la evidente homogeneidad de los resultados en cuanto a variables significativas, variables en la ecuación, y porcentajes correctos de clasificación en los tres escenarios considerados, se considerará el modelo completo obtenido mediante el procedimiento de selección de variables “Backward/Forward” para obtener las conclusiones del estudio y para realizar las recomendaciones pertinentes. Los resultados se utilizarán para realizar recomendaciones a las administraciones titulares de las vías implicadas.

A continuación, se resumen las conclusiones del proyecto, y se plantean las recomendaciones asociadas a una simulación con dos ubicaciones.

CONCLUSIONES DEL ESTUDIO

En los siguientes puntos, se resumen los hallazgos obtenidos en cuanto a aquellas características de diseño y de tráfico de la vía que influyen en el riesgo de que se produzca una colisión frontal:

1. La **anchura de calzada y la Intensidad Media Diaria** son los factores de riesgo que causan mayor influencia sobre la producción de colisiones frontales. Sin embargo, se ha de considerar la IMD como un factor sobre el que no es posible actuar, pues no se trata de una característica geométrica de diseño de la vía. Este último factor fue incluido en el estudio para poder analizar la influencia del resto de variables eliminando el

efecto que podrían ocasionar los diferentes flujos de tráfico que circulan por las carreteras.

Una explicación lógica para este resultado es que un mayor IMD implica un mayor flujo de automóviles en ambos sentidos de la circulación y, por consiguiente, un mayor número de adelantamientos. Además, mayores anchos de calzada podrían reducir la precaución de los conductores, lo que se traduciría en mayor velocidad.

2. El **arcén** es el factor que mayor protección ejerce en este tipo de accidentes. Cuanto mayor es su anchura, más se reduce el riesgo de colisiones frontales. Por ejemplo, un arcén cuya anchura supera los dos metros y medio reduce casi 32 veces el riesgo de colisión frontal, con respecto a una ubicación de las mismas características sin arcén.
3. El **radio de curvatura** ejerce influencia sobre las colisiones frontales cuando la curva presenta un radio menor que 250 metros. En este caso, el riesgo de que en una determinada ubicación de estas características (curva con radio menor de 250 m) ocurra una colisión frontal se multiplica por 4 respecto de una recta que comparta las mismas características en cuanto al resto de variables. Esto podría deberse a que en este tipo de curvas la visibilidad puede llegar a ser muy escasa, considerando las condiciones del terreno o la vegetación.
4. Los **hitos kilométricos** constituyen un factor de riesgo, pues su presencia casi quintuplica el riesgo de colisión frontal, si el resto de factores se mantienen constantes. Se requiere un estudio en profundidad para analizar los motivos reales de la relación entre las colisiones frontales y los hitos kilométricos.
5. El **ancho de carril** constituye un factor de riesgo cuando miden entre 3,25 y 3,75 metros de ancho. Sin embargo, si su anchura es mayor que 3,75 metros, se comporta como un factor de protección. Es importante tener en cuenta que los carriles cuya anchura supera los 3,75 metros son escasos, y suelen presentarse en vías de dos carriles para un sentido.

Considerando el conjunto de todas las conclusiones mencionadas, se podría proceder a realizar recomendaciones.

RECOMENDACIONES

Como ha sido explicado, este proyecto permitirá obtener recomendaciones fundamentadas estadísticamente desde un enfoque inferencial. Por tanto, una de las posibles líneas futuras del presente Trabajo Fin de Grado sería considerar todas las posibles ubicaciones de las carreteras de la Junta de Castilla y León y del Ministerio de Fomento, y elegir aquellas ubicaciones más adecuadas para realizar la segregación de carriles.

Como demostración práctica de las conclusiones obtenidas, se realizará una simulación, en la que se considerarán dos ubicaciones situadas en carreteras interurbanas de una calzada. Para cada una de ellas se considerará la probabilidad de que en una determinada ubicación ocurra una colisión frontal, así como el Odds-Ratio.

Para estas recomendaciones, se tomará en consideración el criterio de priorización de ancho de calzada mayor que 6,99 metros, ya que se ha justificado que las localizaciones que cumplen este criterio son prioritarias, pues en ellas aumenta el riesgo de colisión frontal, y además presentarían un menor coste de implantación de las soluciones consideradas.

A continuación, se presenta una simulación de recomendación entre dos diferentes alternativas de escenarios en los que actuar para evitar las colisiones frontales.

UBICACIÓN NÚMERO 1

A continuación, se presenta otra ubicación que se encuentra en una vía del ministerio de fomento. De nuevo, la titularidad de la misma es estatal, sin embargo, presenta características diferentes a la ubicación anterior:

- Radio de curvatura > 750 metros
- Hitos de arista=sí
- Arcén entre 1,5 y 2,5
- Ancho de calzada mayor que 6,99 metros
- Ancho de carril entre 3,25 y 3,75
- IMD entre 2000 y 10000 vehículos al día.



Ilustración 41. Ubicación 2, simulación. Coordenadas: 41.092529, -5.699564

Sustituyendo en la Ecuación 9. Ecuación final del modelo logístico, curvas y rectas, se obtiene que en la ubicación anterior el riesgo de experimentar una colisión frontal se multiplica casi 20 veces en relación con una ubicación recta que no cuente con hitos de arista, sin arcén, con una anchura de calzada menor 5,99 y una Intensidad Media Diaria menor de 500 vehículos al día.

$$\begin{aligned} \text{Logit}(P) &= \ln\left(\frac{P}{1-P}\right) = \\ &= -2,226 + 2,616 * 1 - 2,737 * 1 + 0,81 * 1 + 1,789 * 1 \\ &\quad + 0,473 * 1 + 0,077 * 1 = 0,802 \end{aligned}$$

$$OR = e^{2,616} * e^{-1,268} * \dots * e^{0,077} = 19,12$$

Por tanto, la probabilidad de que en este escenario ocurra una colisión frontal es la siguiente:

$$P = \frac{2,23}{3,23} = 0,673$$

UBICACIÓN NÚMERO 2

Trataremos la siguiente ubicación de la carretera VA-514. El organismo titular de la mismas es la junta de Castilla y León.



Ilustración 42. Ubicación 4, simulación 41,652452, -4,827669

A continuación, se muestran las características de la última simulación incluida en este apartado:

- Recta
- Anchura de calzada menor que 5,99
- Anchura de calzada menor que 3,25
- Arcén menor que 1,5 metros
- IMD entre 500 y 2000 vehículos al día
- Hitos de arista=no

Ecuación 9. Ecuación final del modelo logístico, curvas y rectas, se obtiene que en la ubicación anterior el riesgo de experimentar una colisión frontal se reduce 2,67 veces en relación con una ubicación recta que no cuente con hitos de arista, sin arcén, con una anchura de calzada menor 5,99 y una Intensidad Media Diaria menor de 500 vehículos al día.

$$\text{Logit}(P) = \ln\left(\frac{P}{1-P}\right) = -2,226 - 1,628 * 1 + 0,646 * 1 = -3,208$$

$$OR = 0,374$$

Por ello, en este caso no sería necesario tomar medidas para reducir el número de colisiones frontales. Esta afirmación se refuerza tras el cálculo de la probabilidad, pues en esta ubicación la probabilidad de que se produzca una colisión frontal toma el valor de 0,039.

$$P = 0,039$$

CONCLUSIONES SIMULACIÓN

Tras considerar cuatro ubicaciones distintas, ocurridas en carreteras de titularidad autonómica y estatal, se podrían obtener las siguientes conclusiones:

- La ecuación del modelo logístico nos permite obtener la probabilidad de que en una determinada ubicación ocurra una colisión frontal. En los dos casos anteriores se observa que la ubicación con mayor probabilidad de experimentar el resultado estudiado es la primera ($P=0,673$). Además, esta primera ubicación cumple con el criterio de priorización (ancho de calzada $> 6,99$ metros), mientras que la segunda ubicación no lo cumple.
- El hecho de conocer la probabilidad y el odds-ratio nos permitirá determinar aquellas ubicaciones en las que la instalación de sistemas de contención de vehículos es prioritaria. En este caso, la primera ubicación presenta un riesgo considerablemente superior a la ubicación 1.
- Del mismo modo, la probabilidad y el odds-ratio asociado a las diferentes ubicaciones nos permitirán determinar en qué ubicaciones no es necesario intervenir debido a que el riesgo de que ocurra una colisión frontal en las mismas no se incrementa de manera sustancial por las condiciones de tráfico y de diseño de la vía.

FUTURAS LÍNEAS DEL TRABAJO

Uno de las finalidades del presente Trabajo Fin de Grado es contribuir a la reducción del número de fallecidos. Para ello, se propone la implantación de Sistemas de Contención de Vehículos seguros y eficientes, que permitan evitar las colisiones frontales, y que a su vez no impliquen costes de reparación elevados en caso de colisión.

Por ello, previamente a la implantación de estos sistemas, se requería caracterizar las localizaciones en las que ocurren este tipo de accidentes. Tras el estudio “caso-control” realizado, se han obtenido las características geométricas de diseño de la vía y de tráfico de la misma,

cuya presencia constituye un factor de riesgo o de protección en el desarrollo del desenlace estudiado.

A pesar de que el estudio ha sido llevado a cabo en vías que transcurrían por el territorio de Castilla y León, el banco de pruebas elegido podría ser representativo del total de carreteras de la geografía española, ya que no parece razonable que el diseño de las vías pueda obedecer a cuestiones locales y, además, Castilla y León presenta un muestrario de escenarios topográficamente variado. Por ello, los resultados podrán ser aplicados a todas las vías de titularidad autonómica y estatal a nivel nacional, si bien sería interesante la realización de un estudio nacional para constatar la reproducibilidad de los resultados.

Estos resultados obtenidos, podrán ser transformados en recomendaciones fundamentadas estadísticamente. De este modo, se podrá informar a las administraciones titulares de las vías implicadas en el estudio de las ubicaciones que presentan riesgo potencial de que se produzcan colisiones frontales. Además, si se aplica el criterio de priorización de anchura superior a 6,99 metros, no solo se conseguirá reducir el número de colisiones, si no que se conseguirá actuar de forma eficiente, consiguiendo un adecuado ratio coste-beneficio.

Por otra parte, una vez caracterizadas las localizaciones, se debería determinar cuál es la forma de segregación de carriles más adecuada en cada caso, de entre las múltiples existentes (marcas viales, medianas, etc.).

Por último, se está trabajando para conseguir datos más recientes, con objetivo de elaborar un artículo científico con datos actualizados para someterlo a publicación en una revista indexada, como “Injury Prevention”, o “Accident and Analysis Prevention”. De este modo, los resultados obtenidos en el estudio serán divulgados a la comunidad científica, de modo que se permita hacer uso de ellos para reducir el número de accidentes, lo cual se considera la finalidad última del presente Trabajo Fin de Grado. Además, estos nuevos datos podrán ser empleados para validar los modelos obtenidos.

BIBLIOGRAFÍA

- [1] Dirección General de Tráfico, "Las principales cifras de la siniestralidad vial, España 2015", 2015.

<http://www.dgt.es/Galerias/seguridad-vial/estadisticas-e-indicadores/publicaciones/principales-cifras-siniestralidad/Las-principales-cifras-2015.pdf>

- [2] M. Olabarria, E. Santamariña-Rubio, M. Marí-Dell’Olmo, M. Gotsens, A. Novoa, C. Borrell and K. Pérez, "Head-on crashes on two-way interurban roads: a public health concern in road safety", *Gaceta Sanitaria*, vol. 29, pp. 16-23, 2015.

- [3] "Carreteras 2+1 para reducir la mortalidad.", *Revista de Tráfico y Seguridad Vial.*, 2015.

<http://revista.dgt.es/es/noticias/nacional/2015/09SEPTIEMBRE/0910Carretera-dos-mas-uno.shtml#.WVdgVcYrygR>

- [4] AUSTROADS, "Guías de Mediana, Tratamiento de Línea Central para Reducir los Choques Frontales", 2016.

- [5] AUSTROADS AP-T320-17, "Verification of Austroads Road Design Criteria Based on Objective Safety Evidence", 2017.

- [6] Jornada técnica: "CARRETERAS 2+1 DEBATE DE UNA SOLUCIÓN CON FUTURO", Barcelona, 2017.

http://www.atc-piarc.com/media/jornadas/233_1.pdf

- [7] *Orden Circular 35/2014 Sobre Criterios De Aplicación De Sistemas De Contención De Vehículos*. Gobierno de España.

<https://www.fomento.gob.es/NR/rdonlyres/11AFD57B-6604-4E4D-8F1E-D4535CB18249/124764/OC352014.pdf>

- [8] "Recomendaciones específicas sobre SCV en Carreteras Locales" (Comunidades Autónomas y Diputaciones Forales)

- [9] D. Kleinbaum, *Logistic regression*. New York [u.a.]: Springer, 1994.
- [10] J. Song and K. Chung, "Observational Studies: Cohort and Case-Control Studies", *Plastic and Reconstructive Surgery*, vol. 126, no. 6, pp. 2234-2242, 2010.
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2998589/pdf/nihms-237355.pdf>
- [11] Susan Lewallen and Paul Courtright "Epidemiology in Practice: Case-Control Studies"
- [12] Pita Fernández S. "Tipos de estudios clínico epidemiológicos"
- http://www.fisterra.com/mbe/investiga/6tipos_estudios/6tipos_estudios.pdf
- [13] E. Lazcano-Ponce, E. Salazar-Martínez and M. Hernández-Avila, "Estudios epidemiológicos de casos y controles. Fundamento teórico, variantes y aplicaciones", *Salud Pública de México*, vol. 43, no. 2, pp. 135-150, 2001.
- http://www.scielosp.org/scielo.php?script=sci_arttext&pid=S0036-36342001000200009
- [14] Lorraine K.Alexander, Brettania Lopesm, Kristen Ricchetti-Masterson, Karin B. Yeatts "Case-Control Studies, Eric Notebook" UNC CH Department of Epidemiology
- https://sph.unc.edu/files/2015/07/nciph_ERIC5.pdf
- [15] Magdalena Szumillas, "Explaining Odds-ratio" *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, bol.19, no 3, pp.227-229, 2010
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2938757/>
- [16] Erik Cobo, Belchin Kostov; Jordi Cortés; José Antonio González; Pilar Muñoz "Intervalos de Confianza", 2014.
- http://ocw.upc.edu/sites/all/modules/ocw/estadistiques/download.php?file=715001/2013/1/54990/08_intervalos_de_confianza-5331.pdf

- [17] "Multivariable Analysis: A Practical Guide for Clinicians, 2nd edition edited by Katz, M. H.", Biometrics, vol. 63, no. 1, pp. 306-306, 2007.
- [18] Manual Salas Veslaco "La regresión logística. Una aplicación a la demanda de estudios universitarios"
- [19] S. Lemeshow and D. Hosmer, Applied Logistic Regression, 2nd ed. Hoboken: John Wiley & Sons, 2017.
http://resource.heartonline.cn/20150528/1_3kOQSTg.pdf
- [20] ROC Curves | NCSS Statistical Software | NCSS.com", Ncss.com, 2017. [Online]. Available:
<https://www.ncss.com/software/ncss/roc-curves-ncss/>.
- [21] López de Ullibarri Galparsoro I, Píta Fernández, S. "Curvas ROC"
- [22] José Blas Navarro, Josep M. Doménech, Núria de la Osa, Lourdes Ezpeleta "El análisis de curvas ROC en estudios epidemiológicos de psicopatología infantil: aplicación al cuestionario CBCL" , Anuario de Psicología, Vol. 29, nº1, pp. 3-15
<http://www.raco.cat/index.php/anuariopsicologia/article/viewFile/61374/88312> super bien explicado
- [23] Oldemar Rodríguez Rojas, "Validación cruzada (Cross-Validation) y remuestreo (Bootstrapping)".
http://www.oldemarrodriguez.com/yahoo_site_admin/assets/docs/Presentación_-_CV.293124233.pdf
- [24] Manual de usuario Dirección General de Tráfico.
<http://www.dgt.es/Galerias/seguridad-vial/estadisticas-e-indicadores/publicaciones/Manual-de-usuario.pdf>
- [25] Junta de Castilla y León, "Plan regional sectorial de Carreteras, 2008-2020", 2008.
http://www.jcyl.es/junta/cf/Plan_carreteras_2008_2020.pdf

- [26] Ministerio de Fomento, "Inventario de la Red de Carreteras del Estado", 2008.

<https://www.fomento.gob.es/NR/rdonlyres/F46372F5-C05B-409C-9BC3-53525ACE341D/137470/Informe3.pdf>

- [27] D. Montgomery and G. Runger, Probabilidad y estadística aplicadas a la ingeniería, 2nd ed. México: Limusa Wiley, 2012.

- [28] Herramienta Software Iberpix 4, Ministerio de Fomento

<http://contenido.ign.es/iberpix2/visor/>

- [29] Manual de reconstrucción de accidentes de tráfico. CESVIMAP.

- [30] Serie Normativas, Instrucciones de Cosntrucción- Instrucción de Construcción Norma 3.1- IC

https://www.fomento.gob.es/NR/rdonlyres/7CDCD3E7-850A-4A9C-813D-B87FAEDE1A7A/136019/Norma_31IC_Trazado_ORDEN_FOM_273_2016.pdf

- [31] Junta de Castilla y León, "Mapas de Tráfico y Velocidades", 2015.

http://www.carreterasymtransportes.jcyl.es/web/jcyl/CarreterasTransportes/es/Plantilla100Detalle/1284349710614/_/1284570582854/Redaccion

- [32] Ministerio de Fomento, "Mapas de Tráfico y Velocidades", 2015.

http://www.fomento.gob.es/NR/rdonlyres/E0E64DD8-8311-439D-AEBB-67F32AAE11B1/139841/MapProvin_2016.pdf

- [33] Luis Bañón Blazquez, *Manual de Carreteras*

https://sirio.ua.es/proyectos/manual_%20carreteras/01020106.pdf

- [34] Informe Race, "Evaluación de la Red de Carreteras del estado", 2016.


http://www.race.es/race.es/wp-content/uploads/2016/12/Informe_RACE-EuroRAP-2016.pdf

- [35] R. Anderson, R. Jin and G. Grunkemeier, "Understanding logistic regression analysis in clinical reports: an introduction", *The Annals of Thoracic Surgery*, vol. 75, no. 3, pp. 753-757, 2003.
- http://www.vhpharmsci.com/decisionmaking/Therapeutic_Decision_Making/Intermediate_files/Understanding%20Logistic%20Regression%20Analysis%20in%20Clinical%20Reports-%20An%20Introduction-JTS.pdf
- [36] Application of European 2+1- Number 275 Roadway Designs (April 2003)
- http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rrd_275.pdf
- [37] <https://www.uoc.edu/in3/emath/docs/RegLogistica.pdf>
- [38] SutterJ.M.; KalivasJ.H., "Comparison of Forward Selection, Backward limination, and Generalized Simulated Annealing for Variable Selection"
- [39] Karl L. Wuensch, "Logistic Regression with SAS", 2015
- <http://core.ecu.edu/psyc/wuenschk/MV/multReg/Logistic-SAS.pdf>
- [40] Karl L. Wuensch, "Logistic Regression with SPSS", 2014
- <http://core.ecu.edu/psyc/wuenschk/MV/multReg/Logistic-SPSS.pdf>
- [41] Identificación de factores de riesgo
- <https://www.bioestadistica.uma.es/baron/apuntes/ficheros/cap08.pdf>

ANEXOS

ANEXO 1- CUESTIONARIO ESTADÍSTICO DE ACCIDENTES DE CIRCULACIÓN CON VÍCTIMAS

CUESTIONARIO ESTADÍSTICO DE ACCIDENTES DE CIRCULACIÓN CON VÍCTIMAS																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
45	1. AÑO					2. NUMERO DE ACCIDENTE <small>(A rellenar por la Jefatura Provincial de Tráfico)</small>					HOJA		 DIRECCIÓN GENERAL DE TRÁFICO		DE CONFORMIDAD CON LO DISPUESTO EN LOS ARTÍCULOS 10, 11 Y 13 DE LA LEY DE LA FUNCIÓN ESTADÍSTICA PÚBLICA, LAS PERSONAS QUE HAYAN INTERVENIDO EN UN ACCIDENTE ESTÁN OBLIGADAS A COLABORAR EN LA CUMPLIMENTACIÓN DE ESTE CUESTIONARIO, CUYOS DATOS PERSONALES SE HALLAN AMPARADOS POR EL SECRETO ESTADÍSTICO.		59																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
46	3. PROVINCIA					ISLA					 DIRECCIÓN GENERAL DE TRÁFICO		DE CONFORMIDAD CON LO DISPUESTO EN LOS ARTÍCULOS 10, 11 Y 13 DE LA LEY DE LA FUNCIÓN ESTADÍSTICA PÚBLICA, LAS PERSONAS QUE HAYAN INTERVENIDO EN UN ACCIDENTE ESTÁN OBLIGADAS A COLABORAR EN LA CUMPLIMENTACIÓN DE ESTE CUESTIONARIO, CUYOS DATOS PERSONALES SE HALLAN AMPARADOS POR EL SECRETO ESTADÍSTICO.		60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
47	4. MES		5. DÍA		6. HORA		7. DÍA SEMANA		8. TIPO DÍA L A F P		9. TOTAL VÍCTIMAS		10. MUERTOS		11. HERIDOS GRAVES		12. HERIDOS LEVES		13. VEHÍCULOS IMPLICADOS		18. HABITANTES (en miles)		19. ZONA		61																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
														Nombre de la calle y número				Entre calle y calle																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
48	20. TIPO DE VÍA					21. ANCHURA DE LA CALZADA					23. MARCAS VIALES					25. ARCÉN PAVIMENTADO					26. ELEMENTOS DE SEGURIDAD DE LA VÍA					62																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
														Número de carriles calzada																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
49	27. FUERA DE INTERSECCIÓN					INTERSECCIÓN					29. Acondicionamiento					30. PRIORIDAD REGULADA POR:					31. SUPERFICIE					63																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
														Con carretera Con calle.					28. Tipo					32. LUMINOSIDAD					33. FACTORES ATMOSFÉRICOS					34. VISIBILIDAD RESTRINGIDA POR:					35. OTRA CIRCUNSTANCIA					36. SEÑALIZACIÓN DE PELIGRO					37. ACERAS					38. ÁRBOLES					39. VISIBILIDAD DE LA SEÑALIZACIÓN VERTICAL (si existe)					40. TIPO DE ACCIDENTE					41. CIRCULACIÓN					42. CIRCULACIÓN BAJO MEDIDAS ESPECIALES					43. MATRICULA Y AÑO DE MATRICULACION					44. MARCA Y MODELO					45. POSIBLES FACTORES CONCURRENTES (Opinión del agente)					46. CROQUIS					47. SÍMBOLOS A UTILIZAR					48. DESCRIPCIONES					49. CROQUIS					50. SÍMBOLOS A UTILIZAR					51. DESCRIPCIONES					52. SÍMBOLOS A UTILIZAR					53. DESCRIPCIONES					54. SÍMBOLOS A UTILIZAR					55. DESCRIPCIONES					56. SÍMBOLOS A UTILIZAR					57. DESCRIPCIONES					58. SÍMBOLOS A UTILIZAR					59. DESCRIPCIONES					60. SÍMBOLOS A UTILIZAR					61. DESCRIPCIONES					62. SÍMBOLOS A UTILIZAR					63. DESCRIPCIONES					64. SÍMBOLOS A UTILIZAR					65. DESCRIPCIONES					66. SÍMBOLOS A UTILIZAR					67. DESCRIPCIONES					68. SÍMBOLOS A UTILIZAR					69. DESCRIPCIONES					70. SÍMBOLOS A UTILIZAR					71. DESCRIPCIONES					72. SÍMBOLOS A UTILIZAR					73. DESCRIPCIONES					74. SÍMBOLOS A UTILIZAR					75. DESCRIPCIONES					76. SÍMBOLOS A UTILIZAR					77. DESCRIPCIONES					78. SÍMBOLOS A UTILIZAR					79. DESCRIPCIONES					80. SÍMBOLOS A UTILIZAR					81. DESCRIPCIONES					82. SÍMBOLOS A UTILIZAR					83. DESCRIPCIONES					84. SÍMBOLOS A UTILIZAR					85. DESCRIPCIONES					86. SÍMBOLOS A UTILIZAR					87. DESCRIPCIONES					88. SÍMBOLOS A UTILIZAR					89. DESCRIPCIONES					90. SÍMBOLOS A UTILIZAR					91. DESCRIPCIONES					92. SÍMBOLOS A UTILIZAR					93. DESCRIPCIONES					94. SÍMBOLOS A UTILIZAR					95. DESCRIPCIONES					96. SÍMBOLOS A UTILIZAR					97. DESCRIPCIONES					98. SÍMBOLOS A UTILIZAR					99. DESCRIPCIONES					100. SÍMBOLOS A UTILIZAR					101. DESCRIPCIONES					102. SÍMBOLOS A UTILIZAR					103. DESCRIPCIONES					104. SÍMBOLOS A UTILIZAR					105. DESCRIPCIONES					106. SÍMBOLOS A UTILIZAR					107. DESCRIPCIONES					108. SÍMBOLOS A UTILIZAR					109. DESCRIPCIONES					110. SÍMBOLOS A UTILIZAR					111. DESCRIPCIONES					112. SÍMBOLOS A UTILIZAR					113. DESCRIPCIONES					114. SÍMBOLOS A UTILIZAR					115. DESCRIPCIONES					116. SÍMBOLOS A UTILIZAR					117. DESCRIPCIONES					118. SÍMBOLOS A UTILIZAR					119. DESCRIPCIONES					120. SÍMBOLOS A UTILIZAR					121. DESCRIPCIONES					122. SÍMBOLOS A UTILIZAR					123. DESCRIPCIONES					124. SÍMBOLOS A UTILIZAR					125. DESCRIPCIONES					126. SÍMBOLOS A UTILIZAR					127. DESCRIPCIONES					128. SÍMBOLOS A UTILIZAR					129. DESCRIPCIONES					130. SÍMBOLOS A UTILIZAR					131. DESCRIPCIONES					132. SÍMBOLOS A UTILIZAR					133. DESCRIPCIONES					134. SÍMBOLOS A UTILIZAR					135. DESCRIPCIONES					136. SÍMBOLOS A UTILIZAR					137. DESCRIPCIONES					138. SÍMBOLOS A UTILIZAR					139. DESCRIPCIONES					140. SÍMBOLOS A UTILIZAR					141. DESCRIPCIONES					142. SÍMBOLOS A UTILIZAR					143. DESCRIPCIONES					144. SÍMBOLOS A UTILIZAR					145. DESCRIPCIONES					146. SÍMBOLOS A UTILIZAR					147. DESCRIPCIONES					148. SÍMBOLOS A UTILIZAR					149. DESCRIPCIONES					150. SÍMBOLOS A UTILIZAR					151. DESCRIPCIONES					152. SÍMBOLOS A UTILIZAR					153. DESCRIPCIONES					154. SÍMBOLOS A UTILIZAR					155. DESCRIPCIONES					156. SÍMBOLOS A UTILIZAR					157. DESCRIPCIONES					158. SÍMBOLOS A UTILIZAR					159. DESCRIPCIONES					160. SÍMBOLOS A UTILIZAR					161. DESCRIPCIONES					162. SÍMBOLOS A UTILIZAR					163. DESCRIPCIONES					164. SÍMBOLOS A UTILIZAR					165. DESCRIPCIONES					166. SÍMBOLOS A UTILIZAR					167. DESCRIPCIONES					168. SÍMBOLOS A UTILIZAR					169. DESCRIPCIONES					170. SÍMBOLOS A UTILIZAR					171. DESCRIPCIONES					172. SÍMBOLOS A UTILIZAR					173. DESCRIPCIONES					174. SÍMBOLOS A UTILIZAR					175. DESCRIPCIONES					176. SÍMBOLOS A UTILIZAR					177. DESCRIPCIONES					178. SÍMBOLOS A UTILIZAR					179. DESCRIPCIONES					180. SÍMBOLOS A UTILIZAR					181. DESCRIPCIONES					182. SÍMBOLOS A UTILIZAR					183. DESCRIPCIONES					184. SÍMBOLOS A UTILIZAR					185. DESCRIPCIONES					186. SÍMBOLOS A UTILIZAR					187. DESCRIPCIONES					188. SÍMBOLOS A UTILIZAR					189. DESCRIPCIONES					190. SÍMBOLOS A UTILIZAR					191. DESCRIPCIONES					192. SÍMBOLOS A UTILIZAR					193. DESCRIPCIONES					194. SÍMBOLOS A UTILIZAR					195. DESCRIPCIONES					196. SÍMBOLOS A UTILIZAR					197. DESCRIPCIONES					198. SÍMBOLOS A UTILIZAR					199. DESCRIPCIONES					200. SÍMBOLOS A UTILIZAR					201. DESCRIPCIONES					202. SÍMBOLOS A UTILIZAR					203. DESCRIPCIONES					204. SÍMBOLOS A UTILIZAR					205. DESCRIPCIONES					206. SÍMBOLOS A UTILIZAR					207. DESCRIPCIONES					208. SÍMBOLOS A UTILIZAR					209. DESCRIPCIONES					210. SÍMBOLOS A UTILIZAR					211. DESCRIPCIONES					212. SÍMBOLOS A UTILIZAR					213. DESCRIPCIONES					214. SÍMBOLOS A UTILIZAR					215. DESCRIPCIONES					216. SÍMBOLOS A UTILIZAR					217. DESCRIPCIONES					218. SÍMBOLOS A UTILIZAR					219. DESCRIPCIONES					220. SÍMBOLOS A UTILIZAR					221. DESCRIPCIONES					222. SÍMBOLOS A UTILIZAR					223. DESCRIPCIONES					224. SÍMBOLOS A UTILIZAR					225. DESCRIPCIONES					226. SÍMBOLOS A UTILIZAR					227. DESCRIPCIONES					228. SÍMBOLOS A UTILIZAR					229. DESCRIPCIONES					230. SÍMBOLOS A UTILIZAR					231. DESCRIPCIONES					232. SÍMBOLOS A UTILIZAR					233. DESCRIPCIONES					234. SÍMBOLOS A UTILIZAR					235. DESCRIPCIONES					236. SÍMBOLOS A UTILIZAR					237. DESCRIPCIONES					238. SÍMBOLOS A UTILIZAR					239. DESCRIPCIONES					240. SÍMBOLOS A UTILIZAR					241. DESCRIPCIONES					242. SÍMBOLOS A UTILIZAR					243. DESCRIPCIONES					244. SÍMBOLOS A UTILIZAR					245. DESCRIPCIONES					246. SÍMBOLOS A UTILIZAR					247. DESCRIPCIONES					248. SÍMBOLOS A UTILIZAR					249. DESCRIPCIONES					250. SÍMBOLOS A UTILIZAR					251. DESCRIPCIONES					252. SÍMBOLOS A UTILIZAR					253. DESCRIPCIONES					254. SÍMBOLOS A UTILIZAR					255. DESCRIPCIONES					256. SÍMBOLOS A UTILIZAR					257. DESCRIPCIONES					258. SÍMBOLOS A UTILIZAR					259. DESCRIPCIONES					260. SÍMBOLOS A UTILIZAR					261. DESCRIPCIONES					262. SÍMBOLOS A UTILIZAR					263. DESCRIPCIONES					264. SÍMBOLOS A UTILIZAR					265. DESCRIPCIONES					266. SÍMBOLOS A UTILIZAR					267. DESCRIPCIONES					268. SÍMBOLOS A UTILIZAR					269. DESCRIPCIONES					270. SÍMBOLOS A UTILIZAR					271. DESCRIPCIONES					272. SÍMBOLOS A UTILIZAR					273. DESCRIPCIONES					274. SÍMBOLOS A UTILIZAR					275. DESCRIPCIONES					276. SÍMBOLOS A UTILIZAR					277. DESCRIPCIONES					278. SÍMBOLOS A UTILIZAR					279. DESCRIPCIONES					280. SÍMBOLOS A UTILIZAR					281. DESCRIPCIONES					282. SÍMBOLOS A UTILIZAR					283. DESCRIPCIONES					284. SÍMBOLOS A UTILIZAR					285. DESCRIPCIONES					286. SÍMBOLOS A UTILIZAR					287. DESCRIPCIONES					288. SÍMBOLOS A UTILIZAR					289. DESCRIPCIONES					290. SÍMBOLOS A UTILIZAR					291. DESCRIPCIONES					292. SÍMBOLOS A UTILIZAR					293. DESCRIPCIONES					294. SÍMBOLOS A UTILIZAR					295. DESCRIPCIONES					296. SÍMBOLOS A UTILIZAR					297. DESCRIPCIONES					298. SÍMBOLOS A UTILIZAR					299. DESCRIPCIONES					300. SÍMBOLOS A UTILIZAR					301. DESCRIPCIONES					302. SÍMBOLOS A UTILIZAR					303. DESCRIPCIONES					304. SÍMBOLOS A UTILIZAR					305. DESCRIPCIONES					306. SÍMBOLOS A UTILIZAR					307. DESCRIPCIONES					308. SÍMBOLOS A UTILIZAR					309. DESCRIPCIONES					310. SÍMBOLOS A UTILIZAR					311. DESCRIPCIONES					312. SÍMBOLOS A UTILIZAR					313. DESCRIPCIONES					314. SÍMBOLOS A UTILIZAR					315. DESCRIPCIONES					316. SÍMBOLOS A UTILIZAR					317. DESCRIPCIONES					318. SÍMBOLOS A UTILIZAR					319. DESCRIPCIONES					320. SÍMBOLOS A UTILIZAR					321. DESCRIPCIONES					322. SÍMBOLOS A UTILIZAR					323. DESCRIPCIONES					324. SÍMBOLOS A UTILIZAR					325. DESCRIPCIONES					326. SÍMBOLOS A UTILIZAR					327. DESCRIPCIONES					328. SÍMBOLOS A UTILIZAR					329. DESCRIPCIONES					330. SÍMBOLOS A UTILIZAR					331. DESCRIPCIONES					332. SÍMBOLOS A UTILIZAR					333. DESCRIPCIONES					334. SÍMBOLOS A UTILIZAR					335. DESCRIPCIONES					336. SÍMBOLOS A UTILIZAR					337. DESCRIPCIONES					338. SÍMBOLOS A UTILIZAR					339. DESCRIPCIONES					340. SÍMBOLOS A UTILIZAR					341. DESCRIPCIONES					342. SÍMBOLOS A UTILIZAR					343. DESCRIPCIONES					344. SÍMBOLOS A UTILIZAR					345. DESCRIPCIONES					346. SÍMBOLOS A UTILIZAR					347. DESCRIPCIONES					348. SÍMBOLOS A UTILIZAR					349. DESCRIPCIONES					350. SÍMBOLOS A UTILIZAR					351. DESCRIPCIONES					352. SÍMBOLOS A UTILIZAR					353. DESCRIPCIONES					354. SÍMBOLOS A UTILIZAR					355. DESCRIPCIONES					356. SÍMBOLOS A UTILIZAR					357. DESCRIPCIONES					358. SÍMBOLOS A UTILIZAR					359. DESCRIPCIONES					360. SÍMBOLOS A UTILIZAR					361. DESCRIPCIONES					362. SÍMBOLOS A UTILIZAR					363. DESCRIPCIONES					364. SÍMBOLOS A UTILIZAR					365. DESCRIPCIONES					366. SÍMBOLOS A UTILIZAR					367. DESCRIPCIONES					368. SÍMBOLOS A UTILIZAR					369. DESCRIPCIONES					370. SÍMBOLOS A UTILIZAR					371. DESCRIPCIONES					372. SÍMBOLOS A UTILIZAR					373. DESCRIPCIONES					374. SÍMBOLOS A UTILIZAR					375. DESCRIPCIONES					376. SÍMBOLOS A UTILIZAR					377. DESCRIPCIONES					378. SÍMBOLOS A UTILIZAR					379. DESCRIPCIONES					380. SÍMBOLOS A UTILIZAR					381. DESCRIPCIONES					382. SÍMBOLOS A UTILIZAR					383. DESCRIPCIONES					384. SÍMBOLOS A UTILIZAR					385. DESCRIPCIONES					386. SÍMBOLOS A UTILIZAR					387. DESCRIPCIONES					388. SÍMBOLOS A UTILIZAR					389. DESCRIPCIONES					390. SÍMBOLOS A UTILIZAR					391. DESCRIPCIONES					392. SÍMBOLOS A UTILIZAR					393. DESCRIPCIONES					394. SÍMBOLOS A UTILIZAR					395. DESCRIPCIONES					396. SÍMBOLOS A UTILIZAR					397. DESCRIPCIONES					398. SÍMBOLOS A UTILIZAR					399. DESCRIPCIONES					400. SÍMBOLOS A UTILIZAR					401. DESCRIPCIONES					402. SÍMBOLOS A UTILIZAR					403. DESCRIPCIONES					404. SÍMBOLOS A UTILIZAR					405. DESCRIPCIONES					406. SÍMBOLOS A UTILIZAR					407. DESCRIPCIONES					408. SÍMBOLOS A UTILIZAR					409. DESCRIPCIONES					410. SÍMBOLOS A UTILIZAR					411. DESCRIPCIONES					412. SÍMBOLOS A UTILIZAR					413. DESCRIPCIONES					414. SÍMBOLOS A UTILIZAR					415. DESCRIPCIONES					416. SÍMBOLOS A UTILIZAR					417. DESCRIPCIONES					418. SÍMBOLOS A UTILIZAR					419. DESCRIPCIONES					420. SÍMBOLOS A UTILIZAR					421. DESCRIPCIONES					422. SÍMBOLOS A UTILIZAR					423. DESCRIPCIONES					424. SÍMBOLOS A UTILIZAR					425. DESCRIPCIONES					426. SÍMBOLOS A UTILIZAR					427. DESCRIPCIONES					428. SÍMBOLOS A UTILIZAR					429. DESCRIPCIONES					430. SÍMBOLOS A UTILIZAR</				

45 VEH. A	TIPO DE VEHÍCULOS IMPLICADOS 0.1 BICICLETA O TRICICLO SIN MOTOR 0.2 CICLOMOTOR 1.0 COCHE DE MINUSVÁLIDO 1.1 MOTOCICLETA 2.1 TURISMO DE SP HASTA 9 PLAZAS 2.2 TURISMO SIN REMOLQUE 2.3 TURISMO CON REMOLQUE 2.4 AMBULANCIA 3.0 MAQUINARIA DE OBRAS Y AGRÍCOLA 3.1 TRACTOR AGRÍCOLA SIN REMOLQUE 3.2 TRACTOR AGRÍCOLA CON REMOLQUE 4.1 CAMIÓN (PM ≤ 3.500 K) SIN REMOLQUE 4.2 CAMIÓN (PM ≤ 3.500 K) CON REMOLQUE 4.3 FURGONETA 5.1 CAMIÓN (PM ≥ 3.500 K) SIN REMOLQUE 5.2 CAMIÓN (PM ≥ 3.500 K) CON REMOLQUE 5.3 CAMIÓN CISTERNA SIN REMOLQUE 5.4 CAMIÓN CISTERNA CON REMOLQUE 5.5 VEHÍCULO ARTICULADO 6.1 AUTOBÚS DE LÍNEA REGULAR 6.2 AUTOBÚS ESCOLAR 6.3 OTRO AUTOBÚS 7.0 TREN 8.0 CARRO 8.1 OTROS VEHÍCULOS (ampliar observaciones) 9.0 DESCONOCIDO	NACIONALIDAD DEL CONDUCTOR 0. ESPAÑA 1. PORTUGAL 2. FRANCIA 3. MARRUECOS 4. ALEMANIA 5. GRAN BRETAÑA 6. ITALIA 7. SUIZA 8. BÉLGICA 9. HOLANDA 10. ESTADOS UNIDOS 11. OTROS PAÍSES DEL MAGREB 12. OTROS PAÍSES 13. DESCONOCIDA	
46 VEH. B		VEH. A 59 VEH. B 60	
47 VEH. A	ESTADO DEL VEHÍCULO 0. APARENTEMENTE NINGÚN DEFECTO 1. NEUMÁTICOS MUY DESGASTADOS 2. PINCHAZO O REVENTÓN 3. PÉRDIDA DE RUEDA 4. LUCES DELANTERAS DEFICIENTES 5. LUCES TRASERAS DEFICIENTES 6. FRENOS DEFICIENTES 7. DIRECCIÓN ROTA O DEFECTUOSA 8. SOBRECARGADO 9. CARGA MAL ACONDICIONADA 10. OTROS DEFECTOS 11. SE IGNORA	ACCIÓN DEL CONDUCTOR 0.1 SIGUIENDO LA RUTA 0.2 ADELANTANDO POR LA DERECHA 0.3 ADELANTANDO POR LA IZQUIERDA 1.1 GIRANDO O SALIENDO HACIA OTRA VÍA O ACCESO POR LA DERECHA 1.2 GIRANDO O SALIENDO HACIA OTRA VÍA O ACCESO POR LA IZQUIERDA 1.3 GIRANDO EN "U" 2.1 INCORPORÁNDOSE DESDE OTRA VÍA O ACCESO 2.2 CRUZANDO INTERSECCIÓN 2.3 ESTACIONANDO O SALIENDO DEL ESTACIONAMIENTO 3.1 CIRCULANDO HACIA ATRÁS 4.1 MANIOBRA SÚBITA PARA SALVAR OBSTÁCULO O VEHÍCULO 4.2 MANIOBRA SÚBITA PARA SALVAR PEATÓN AISLADO O EN GRUPO	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
48 VEH. B		VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64	
49 VEH. A	NÚMERO DE OCUPANTES POR VEHÍCULO 1. PROFESIONAL POR CUENTA PROPIA 2. PROFESIONAL POR CUENTA AJENA 3. DE VEHÍCULO MILITAR 4. DE VEHÍCULO ALQUILADO SIN CONDUCTOR 5. PARTICULAR 6. SE DESCONOCE	PRESUNTAS INFRACCIONES DEL CONDUCTOR 0.1 CONDUCCIÓN DISTRAÍDA O DESATENTA 1.1 INCORRECTA UTILIZACIÓN DEL ALUMBRADO 2.1 CIRCULAR POR MANO CONTRARIA O SENTIDO PROHIBIDO 2.2 INVADIR PARCIALMENTE EL SENTIDO CONTRARIO 2.3 GIRAR INCORRECTAMENTE 2.4 ADELANTAR ANTIREFLEGRAMENTARIAMENTE 2.5 CIRCULAR EN ZIG-ZAG 3.1 NO MANTENER INTERVALO DE SEGURIDAD 3.2 FRENAR SIN CAUSA JUSTIFICADA 4.1 NO RESPETAR LA NORMA GENÉRICA DE PRIORIDAD 4.2 NO CUMPLIR LAS INDICACIONES DE SEMÁFORO 4.3 NO CUMPLIR LA SEÑAL DE "STOP" 4.4 NO CUMPLIR LA SEÑAL DE "CEDA EL PASO" 4.5 NO RESPETAR EL PASO PARA PEATONES 4.6 NO CUMPLIR OTRA SEÑAL DE TRÁFICO O POLICIA 5.1 NO INDICAR O INDICAR MAL UNA MANIOBRA 5.2 ENTRAR SIN PRECAUCIÓN EN LA CIRCULACIÓN 5.3 PARADO O ESTACIONAMIENTO PROHIBIDO O PELIGROSO 6.1. CICLISTAS O CICLOMOTORISTAS EN POSICIÓN PARALELA 6.2 CICLISTA O CICLOMOTORISTA CIRCULANDO FUERA DE PISTA O ARCÉN 7.1 APERTURA DE PUERTAS SIN PRECAUCIÓN 8.1 OTRA INFRACCIÓN 9.1 NINGUNA INFRACCIÓN	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
50 VEH. B		VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64	
51 VEH. A	TIPO DE CONDUCTOR 1. PROFESIONAL POR CUENTA PROPIA 2. PROFESIONAL POR CUENTA AJENA 3. DE VEHÍCULO MILITAR 4. DE VEHÍCULO ALQUILADO SIN CONDUCTOR 5. PARTICULAR 6. SE DESCONOCE		VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
52 VEH. B			VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
53 VEH. A	MOTIVO DEL DESPLAZAMIENTO 1. DURANTE SU JORNADA DE TRABAJO 2. DIRIGIRSE O REGRESAR DEL LUGAR DEL TRABAJO 3. SALIDA O REGRESO DE VACACIONES 4. SALIDA O REGRESO DE PUENTES Y FESTIVOS 5. URGENCIAS 6. OCIO 7. OTRO	 DIRECCIÓN GENERAL DE TRÁFICO	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
54 VEH. B			VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
55 VEH. A	DESPLAZAMIENTO PREVISTO 1. LOCAL (MENOS DE 50 km.) 2. MEDIO (DE 50 A 200 km.) 3. LARGO (MÁS DE 200 km.) 4. SE DESCONOCE	ACCION DEL PEATÓN 1. ATRAVESANDO INTERSECCIÓN 2. CRUZANDO CALZADA FUERA DE INTERSECCIÓN 3. EN ARCÉN POR SU DERECHA 4. EN ARCÉN POR SU IZQUIERDA 5. EN CALZADA POR SU DERECHA 6. EN CALZADA POR SU IZQUIERDA 7. TRABAJAR EN LA CALZADA 8. REPARANDO VEHÍCULO 9. SUBIR O DESCENDER DE UN VEHÍCULO 10. SOBRE ACERA O REFUGIO 11. OTRA	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
56 VEH. B			VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
57 VEH. A	TRANSPORTE DE MERCANCIAS PELIGROSAS 1. EXPLOSIVOS 2. RADIATIVOS 3. INFLAMABLES 4. OTRAS MP	PRESUNTAS INFRACCIONES ADMINISTRATIVAS 1. CARECER DEL PERMISO DE CONDUCCIÓN ADECUADO 2. PERMISO DE CONDUCCIÓN CADUCADO 3. EXCESO DE VIAJEROS O CARGA 4. NO TENER EFECTUADA LA INSPECCIÓN TÉCNICA REGLAMENTARIA DEL VEHÍCULO 5. TACÓGRAFO NO REVISADO 6. NINGUNA DE LAS RELACIONADAS 7. SE IGNORA	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
58 VEH. B			VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
59 VEH. A	LOCALIZACIÓN DE LAS LESIONES MÁS GRAVES 1. CABEZA 2. CARA 3. CUELLO 4. PECHO 5. ESPALDA 6. ABDOMEN 7. EXTREMIDADES SUPERIORES 8. EXTREMIDADES INFERIORES 9. TODO EL CUERPO 10. SE IGNORA	PRESUNTAS INFRACCIONES SOBRE VELOCIDAD 1. VELOCIDAD INADECUADA PARA LAS CONDICIONES EXISTENTES 2. SOBREPASAR LA VELOCIDAD ESTABLECIDA 3. MARCHA LENTA ENTORPECIENDO LA CIRCULACIÓN 4. NINGUNA 5. SE IGNORA	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
60 VEH. B			VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
61 VEH. A	VEHÍCULO EN QUE VIAJABA A B P-PEATÓN	CONDICIONES PSICOFÍSICAS 0. APARENTEMENTE NORMAL 1. ALCOHOL SIN PRUEBA DE ALCOHOLEMIA 2. ALCOHOL CON PRUEBA POSITIVA DE ALCOHOLEMIA 3. DROGAS 4. ENFERMEDAD SÚBITA 5. SUEÑO O SOPOR 6. CANSANCIO 7. PREOCUPACIÓN 8. SE IGNORA	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
62 VEH. B			VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
63 VEH. A	USO DE ACCESORIOS DE SEGURIDAD 1. UTILIZANDO CINTURÓN 2. SISTEMA DE RETENCIÓN INFANTIL 3. UTILIZANDO CASCO 4. CON REFLECTANTES (peatón) 5. NINGUNO 6. SE DESCONOCE	DEFECTO FÍSICO PREVIO 0. SIN DEFECTO CONOCIDO 1. DE VISIÓN 2. DE AUDICIÓN 3. DE EXTREMIDADES SUPERIORES 4. DE EXTREMIDADES INFERIORES 5. OTRO 6. SE IGNORA	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
64 VEH. B			VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
65 VEH. A		LESIVIDAD 1. MUERTO 2. HERIDO GRAVE 3. HERIDO LEVE 4. LISO 5. SE IGNORA	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
66 VEH. B		PERMISOS DE CONDUCCIÓN CLASE AÑO DE EXPE- DICIÓN	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
67 VEH. A		HORAS DE CONDUCCIÓN CONTINUADA 1. MENOS DE 1 2. DE 1 A 3 3. DE 3 A 5 4. MÁS DE 5 5. SE IGNORA	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64
68 VEH. B		EDAD SEXO	VEH. A 61 VEH. B 62 VEH. A 63 VEH. B 64

ANEXO 2- CÓDIGOS DE PROGRAMACIÓN EN SAS

En el siguiente anexo se muestran los códigos implementados en SAS para obtener las curvas ROC, así como las tablas de clasificación “Leave-One-Out”, y la comparación entre los modelos considerados.

En cada uno de ellos, se distinguen el modelo completo, el modelo forward, y el modelo Backward.

MODELO COMPLETO: CURVAS Y RECTAS

```
libname acc 'C:\Users\Desktop\LAURA';
proc logistic data=acc.modelocurvasyrectas_cyl;
class barrera IMD_cat anchcarr_nueva radio_categorizado acera
hitos arcen anchcalz numcarr;
model casocontrol=barrera IMD_cat anchcarr_nueva
radio_categorizado acera hitos arcen anchcalz numcarr/ctable;
*model casocontrol=IMD_cat anchcarr_nueva radio_categorizado
hitos arcen anchcalz/ctable;
*model casocontrol=IMD_cat anchcarr_nueva radio_categorizado
hitos arcen anchcalz/ctable;
roc 'forward' IMD_cat anchcarr_nueva radio_categorizado hitos
arcen anchcalz;
roc 'backward' IMD_cat anchcarr_nueva radio_categorizado
hitos arcen anchcalz;
roccompare reference('backward') / estimate e;
run;
```

MODELO “SOLO RECTAS”

```
libname acc 'C:\Users\Desktop\LAURA';

proc logistic data= acc.modelorectas_cyl;
class barrera IMD_cat anchcarr_nueva acera hitos arcen
anchcalz numcarrfef;
model casocontrol=barrera IMD_cat anchcarr_nueva acera hitos
arcen anchcalz numcarrfef/ctable;
*model casocontrol=IMD_cat anchcarr_nueva arcen
anchcalz/ctable;
*model casocontrol=IMD_cat anchcarr_nueva arcen
anchcalz/ctable;
roc 'forward' IMD_cat anchcarr_nueva arcen anchcalz;
roc 'backward' IMD_cat anchcarr_nueva arcen anchcalz;
roccompare reference('backward') / estimate e;
run;
```

MODELO “SOLO CURVAS”

```
libname acc 'C:\Users\Desktop\LAURA';
proc logistic data= acc.modelocurvas_cyl;
class barrera IMD_cat anchcarr_nueva radio_categorizado acera
hitos arcen anchcalz numcarrfef;
*model casocontrol=barrera IMD_cat anchcarr_nueva
radio_categorizado acera hitos arcen anchcalz
numcarrfef/ctable;
*model casocontrol=IMD_cat anchcarr_nueva radio_categorizado
hitos arcen anchcalz/ctable;
model casocontrol=IMD_cat anchcarr_nueva radio_categorizado
hitos arcen anchcalz/ctable;
roc 'forward' IMD_cat anchcarr_nueva radio_categorizado hitos
arcen anchcalz;
roc 'backward' IMD_cat anchcarr_nueva radio_categorizado hitos
arcen anchcalz;
rocontrast reference('backward') / estimate e;
run;
```

ANEXO 3- TABLAS

TABLA 1. POSIBLES VALORES DE LA VARIABLE RESPUESTA DICOTÓMICA.....	9
TABLA 2. RELACIÓN ENTRE LA VARIABLE DE EXPOSICIÓN Y EL RESULTADO PARA EL CÁLCULO DEL ODDS-RATIO	15
TABLA 3. INTERPRETACIÓN DEL ODDS-RATIO	16
TABLA 4. INTERPRETACIÓN DE LOS INTERVALOS DE CONFIANZA PARA EL ODDS-RATIO	17
TABLA 5. VALORES DE LOS CASOS EN FUNCIÓN DE SI SE HAN CLASIFICADO CORRECTAMENTE O NO	23
TABLA 6. CASO EXCLUIDO EN CADA INTERACCIÓN EN EL PROCEDIMIENTO LEAVE-ONE-OUT.	26
TABLA 7: VARIABLES INCLUIDAS EN EL FICHERO GENERAL DE LA BASE DATOS ARENA.....	35
TABLA 8 LONGITUD DE LA RED DE CARRETERAS DE CASTILLA Y LEÓN ^[25]	40
TABLA 9. DISTRIBUCIÓN DE LA RCE EN CASTILLA Y LEÓN.....	40
TABLA 10: TABLA CRUZADA TITULARIDAD DE LA VÍAS * ANCHURA DE LA CALZADA	44
TABLA 11: MODELO EXCEL PARA LA OBTENCIÓN DE LOS CONTROLES	46
TABLA 12 RADIO MÍNIMOS Y MÁXIMOS EN FUNCIÓN DEL TIPO DE CARRETERA ^[27]	61
TABLA 13. SISTEMAS DE CONTENCIÓN DE VEHÍCULOS PARA MOTOCICLETAS [7]	61
TABLA 14. CATEGORIZACIÓN DE LA IMD SEGÚN EL INFORME RACE.....	64
TABLA 15 IMD DE CARRETERAS DE TITULARIDAD AUTONÓMICA EXPRESADA EN VEHÍCULOS POR DÍA ^[25]	64
TABLA 16 CLASIFICACIÓN IMD SEGÚN LA JUNTA DE CASTILLA Y LEÓN ^[25]	65
TABLA 17 CATEGORIZACIÓN DE LA VARIABLE IMD	65
TABLA 18 CRUCE DE LAS VARIABLES "ANCHO DE CALZADA" Y "CASO-CONTROL"	70
TABLA 19: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "ANCHO DE CALZADA", MODELO COMPLETO.....	71
TABLA 20: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "ANCHO DE CALZADA", SOLO RECTAS	73
TABLA 21: REGRESIÓN LOGÍSTICA PARA LA VARIABLE ANCHO DE CALZADA, SOLO CURVAS.....	74
TABLA 22: TABLA CRUZADA "ANCHO DE CARRIL" Y "CASO-CONTROL"	75
TABLA 23. ODDS-RATIO PARA LA VARIABLE "ANCHO DE CARRIL", RECTAS Y CURVAS	76
TABLA 24. ODDS-RATIO PARA LA VARIABLE "ANCHO DE CARRIL, SOLO RECTAS.....	77
TABLA 25. ODDS-RATIO PARA LA VARIABLE "ANCHO DE CARRIL", SOLO CURVAS	77
TABLA 26: TABLA CRUZADA "RADIO DE CURVATURA" Y "CASO-CONTROL"	78
TABLA 27 REGRESIÓN LOGÍSTICA PARA LA VARIABLE "RADIO DE CURVATURA CATEGORIZADO", RECTAS Y CURVAS	80
TABLA 28: TABLA CRUZADA "ARCÉN" Y "CASO-CONTROL"	81
TABLA 29 REGRESIÓN LOGÍSTICA VARIABLE "ARCÉN", RECTAS Y CURVAS.....	82
TABLA 30 REGRESIÓN LOGÍSTICA PARA LA VARIABLE "ARCÉN", SOLO RECTAS	83
TABLA 31: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "ARCÉN", CURVAS	84
TABLA 32 TABLA CRUZADA "BARRERA DE SEGURIDAD" Y "CASO-CONTROL"	84
TABLA 33 REGRESIÓN LOGÍSTICA PARA LA VARIABLE "BARRERA DE SEGURIDAD", CURVAS Y RECTAS.....	86
TABLA 34: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "BARRERA DE SEGURIDAD", SOLO RECTAS	86
TABLA 35. REGRESIÓN LOGÍSTICA PARA LA VARIABLE "BARRERA DE SEGURIDAD", SOLO CURVAS	87
TABLA 36: TABLA CRUZADA "HITOS DE ARISTA" Y "CASO-CONTROL"	87
TABLA 37: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "HITOS DE ARISTA", RECTAS Y CURVAS.....	88
TABLA 38 REGRESIÓN LOGÍSTICA PARA LA VARIABLE "HITOS DE ARISTA", RECTAS.....	89
TABLA 39 REGRESIÓN LOGÍSTICA PARA LA VARIABLE "HITOS DE ARISTA", CURVAS	89
TABLA 40: TABLA CRUZADA "ACERA" * "CASO-CONTROL"	90
TABLA 41 REGRESIÓN LOGÍSTICA PARA LA VARIABLE "ACERA", RECTAS Y CURVAS	91
TABLA 42: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "ACERA", RECTAS	92
TABLA 43. REGRESIÓN LOGÍSTICA PARA LA VARIABLE "ACERA", CURVAS	92
TABLA 44: TABLA CRUZADA "INTENSIDAD MEDIA DIARIA"*"CASO-CONTROL"	93
TABLA 45: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "IMD CATEGORIZADA", CURVAS Y RECTAS	95

TABLA 46: REGRESIÓN LOGÍSTICA PARA LA VARIABLE "IMD CATEGORIZADA", RECTAS.....	96
TABLA 47. ODDS-RATIO PARA LA VARIABLE "IMD CATEGORIZADA", CURVAS	97
TABLA 48. TABLA CRUZADA "NÚMERO DE CARRILES" Y "CASO-CONTROL".....	97
TABLA 49. ODDS-RATIO PARA LA VARIABLE "NÚMERO DE CARRILES"	98
TABLA 50. ODDS-RATIO E IC PARA LA VARIABLE "NÚMERO DE CARRILES", RECTAS	99
TABLA 51. ODDS-RATIO E IC PARA LA VARIABLE "NÚMERO DE CARRILES", CURVAS	99
TABLA 52. RESUMEN DE LA INFLUENCIA DE LAS VARIABLES (REGRESIÓN LOGÍSTICA SIMPLE, CURVAS Y RECTAS)	100
TABLA 53. RESUMEN DE LA INFLUENCIA DE LAS VARIABLES (REGRESIÓN LOGÍSTICA SIMPLE, RECTAS)	101
TABLA 54. RESUMEN DE LA INFLUENCIA DE LAS VARIABLES (REGRESIÓN LOGÍSTICA SIMPLE, CURVAS)	101
TABLA 55. CATEGORÍAS DE REFERENCIA PARA TODAS LAS VARIABLES DEL MODELO	106
TABLA 56. ODDS-RATIO MODELO INTRODUCCIÓN, CURVAS Y RECTAS.....	107
TABLA 57. RESULTADOS MODELO INTRODUCCIÓN, SPSS	108
TABLA 58. ODDS-RATIO MODELO BACKWARD/FORWARD, CURVAS Y RECTAS	110
TABLA 59. RESUMEN MODELO BACKWARD/FORWARD, CURVAS Y RECTAS	111
TABLA 60. COMPARACIÓN DE RESULTADOS, SAS Y SPSS	113
TABLA 61. COMPARACIÓN ESTADÍSTICA DE MODELOS SEGÚN SAS	113
TABLA 62. ODDS-RATIO MODELO INTRODUCCIÓN, SOLO RECTAS.....	114
TABLA 63. RESUMEN MODELO INTRODUCCIÓN, RECTAS.....	114
TABLA 64. ODDS-RATIO MODELO BACKWARD/FORWARD, SOLO RECTAS	116
TABLA 65. RESUMEN MODELO BACKWARD/FORWARD, SOLO RECTAS	116
TABLA 66. COMPARATIVA DE MODELOS SEGÚN SAS, RECTAS	118
<i>TABLA 67. COMPARACIÓN DE MODELOS SPSS Y SAS, CURVAS</i>	<i>119</i>
TABLA 68. ODDS-RATIO MODELO INTRODUCCIÓN, CURVAS.....	119
TABLA 69. RESUMEN MODELO INTRODUCCIÓN, CURVAS	120
TABLA 70. REGRESIÓN LOGÍSTICA MULTIVARIANTE, BACKWARD/FORWARD, CURVAS	122
TABLA 71 COMPARACIÓN DE MODELOS SPSS Y SAS, CURVAS	123
TABLA 72. COMPARACIÓN ESTADÍSTICA DE MODELOS SEGÚN SAS, CURVAS.....	124
TABLA 73. CARACTERÍSTICAS MODELO CURVAS Y RECTAS	124
TABLA 74. CARACTERÍSTICAS DEL MODELO FINAL PARA LAS RECTAS	128
TABLA 75. CARACTERÍSTICAS DEL MODELO FINAL PARA LAS CURVAS	130

ANEXO 4- ILUSTRACIONES

ILUSTRACIÓN 1: FALLECIDOS EN ACCIDENTES DE TRÁFICO DURANTE EL AÑO 2015 ^[1]	2
ILUSTRACIÓN 2: CASOS Y CONTROLES PARTICULARIZADOS AL PRESENTE ESTUDIO	9
ILUSTRACIÓN 3 ESQUEMA DE UN ESTUDIO COHORTE	10
ILUSTRACIÓN 4. ESQUEMA DE UN ESTUDIO CASO-CONTROL	11
ILUSTRACIÓN 5. RELACIÓN ENTRE CASOS Y CONTROLES EN EL PRESENTE ESTUDIO	14
ILUSTRACIÓN 6. OBJETIVO DE LA FUNCIÓN LOGÍSTICA	18
ILUSTRACIÓN 7. AJUSTE LINEAL	19
ILUSTRACIÓN 8 REPRESENTACIÓN GRÁFICA DE LA FUNCIÓN LOGÍSTICA	20
ILUSTRACIÓN 9. COMPARACIÓN DE DIFERENTES CURVAS ROC.....	24
ILUSTRACIÓN 10: PORCENTAJE DE ACCIDENTES OCURRIDOS EN CARRETERAS DE DIFERENTE ANCHURA.	45
ILUSTRACIÓN 11. CREACIÓN DE LA DISTRIBUCIÓN UNIFORME.....	48
ILUSTRACIÓN 12. REPRESENTACIÓN DE LA DISTRIBUCIÓN UNIFORME CREADA.....	48
ILUSTRACIÓN 13 FUNCIÓN DE DISTRIBUCIÓN DE LA VARIABLE UNIFORME CREADA.....	49
ILUSTRACIÓN 14. OBTENCIÓN DEL KILÓMETRO ASOCIADO A CADA CONTROL, SOBRE LA LONGITUD TOTAL DE CARRETERAS DE CASTILLA Y LEÓN.	50
ILUSTRACIÓN 15. LOCALIZACIÓN DE LA CARRETERA Y EL PUNTO KILOMÉTRICO (1)	50
ILUSTRACIÓN 16. LOCALIZACIÓN DE LA CARRETERA Y EL PUNTO KILOMÉTRICO (2)	51
ILUSTRACIÓN 17. LOCALIZACIÓN DE LA CARRETERA Y PUNTO KILOMÉTRICO (3)	51
ILUSTRACIÓN 18. ESQUEMA DEFINITIVO DE CASOS Y CONTROLES	52
ILUSTRACIÓN 19. LOCALIZACIÓN PUNTO KILOMÉTRICO MEDIANTE IBERPIX 4	53
ILUSTRACIÓN 20. LOCALIZACIÓN DEL PUNTO KILOMÉTRICO EXACTO, HITO KILOMÉTRICO+ HERRAMIENTA REGLA.....	53
ILUSTRACIÓN 21. ESQUEMA PARA LA OBTENCIÓN DEL RADIO DE CURVATURA	60
ILUSTRACIÓN 22. MAPA DE TRÁFICO DE LA RED DE CARRETERAS DEL ESTADO	62
ILUSTRACIÓN 23 LEYENDA MAPA IMD DEL MINISTERIO DE FOMENTO (PROVINCIA VALLADOLID)	63
ILUSTRACIÓN 24: GRÁFICO DE BARRAS DE LA DISTRIBUCIÓN DE LA VARIABLE "ANCHURA DE CALZADA" PARA LOS CASOS Y LOS CONTROLES.....	70
ILUSTRACIÓN 25 GRÁFICO DE BARRAS DE LA DISTRIBUCIÓN DE LA VARIABLE ANCHURA DE CARRIL EN LOS CASOS Y EN LOS CONTROLES.	75
ILUSTRACIÓN 26: GRÁFICO DE BARRAS DE LA DISTRIBUCIÓN DE LA VARIABLE "RADIO DE CURVATURA" EN LOS CASOS Y EN LOS CONTROLES.	79
ILUSTRACIÓN 27: GRÁFICO DE BARRAS, DISTRIBUCIÓN DE LA VARIABLE "BARRERA DE SEGURIDAD" PARA LOS CASOS Y LOS CONTROLES.	85
ILUSTRACIÓN 28. GRÁFICO DE BARRAS DISTRIBUCIÓN DE LOS CASOS Y LOS CONTROLES PARA LA VARIABLE "HITOS DE ARISTA"	88
ILUSTRACIÓN 29: GRÁFICO DE BARRA DE LA DISTRIBUCIÓN DE LOS CASOS Y LOS CONTROLES PARA LAS CATEGORÍAS DE LA VARIABLE "ACERA"	91
ILUSTRACIÓN 30: GRÁFICO DE BARRAS PARA LA DISTRIBUCIÓN DE LA VARIABLE "IMD CATEGORIZADO" EN LOS CASOS Y EN LOS CONTROLES	94
ILUSTRACIÓN 31. GRÁFICO DE BARRAS PARA LA VARIABLE "NÚMERO DE CARRILES"	98
ILUSTRACIÓN 32. CURVA ROC, INTRODUCCIÓN, CURVAS Y RECTAS	108
ILUSTRACIÓN 33. CURVA ROC, FORWARD Y BACKWARD (CURVAS Y RECTAS)	111
ILUSTRACIÓN 34. COMPARATIVA DE MODELO, CURVAS Y RECTAS	112
ILUSTRACIÓN 35. CURVA ROC PARA MODELO INTRODUCCIÓN, RECTAS	115
ILUSTRACIÓN 36. CURVA ROC, BACKWARD/FORWARD, RECTAS.....	117
ILUSTRACIÓN 37. COMPARACIÓN CURVAS ROC, RECTAS.....	118
ILUSTRACIÓN 38. CURVA ROC, MODELO INTRODUCCIÓN, CURVAS	120
ILUSTRACIÓN 39. MODELO BACKWARD/FORWARD, CURVAS	122

ILUSTRACIÓN 40: COMPARATIVA DE MODELOS, SOLO CURVAS	123
ILUSTRACIÓN 41. UBICACIÓN 2, SIMULACIÓN. COORDENADAS: 41.092529, -5.699564	140
ILUSTRACIÓN 42. UBICACIÓN 4, SIMULACIÓN 41,652452, -4,827669	141

ANEXO 5. ECUACIONES

ECUACIÓN 1. CÁLCULO DEL ODDS-RATIO	16
ECUACIÓN 2 INTERVALO DE CONFIANZA PARA EL ODDS-RATIO	17
ECUACIÓN 3. EXPRESIÓN LINEAL MODELO LOGÍSTICO	18
ECUACIÓN 4. FORMA MATEMÁTICA DEL MODELO LOGÍSTICO [20].....	19
ECUACIÓN 5. TRANSFORMACIÓN LOGARÍTMICA	20
ECUACIÓN 6. ODDS-RATIO	21
ECUACIÓN 7. CÁLCULO DE LA SENSIBILIDAD Y ESPECIFICIDAD	23
ECUACIÓN 8: OBTENCIÓN DEL RADIO DE CURVATURA	60
ECUACIÓN 9. ECUACIÓN FINAL DEL MODELO LOGÍSTICO, CURVAS Y RECTAS	125
ECUACIÓN 10. ECUACIÓN FINAL DEL MODELO LOGÍSTICO, RECTAS.....	128
ECUACIÓN 11. ECUACIÓN FINAL DEL MODELO LOGÍSTICO, CURVAS.....	131