

# ANÁLISIS DE VARIABLES ASOCIADAS EN SINIESTROS VIALES DENTRO DE UN ÁREA GEORREFERENCIADA PARA ESTABLECER PROBABILIDADES DE HOSPITALIZACIÓN. CASO DE APLICACION

Fernanda Martínez Micakoski <sup>1</sup>, José Julián Rivera <sup>2</sup>

<sup>1</sup>Investigadora UTN Trenque Lauquen  
Fac. Reg. Trenque Lauquen, Universidad Tecnológica Nacional  
Racedo 298, (6400) Trenque Lauquen, Bs. As., Argentina  
e-mail: [fer\\_micakoski@yahoo.com](mailto:fer_micakoski@yahoo.com), web: [www.frtl.utn.edu.ar](http://www.frtl.utn.edu.ar)

<sup>2</sup>Subdirector LEMaC  
LEMaC Centro de Investigaciones Viales, Fac. Reg. La Plata, Universidad Tecnológica Nacional  
Calle 60 y 124, (1900) La Plata, Bs. As., Argentina  
e-mail: [lemac@frlp.utn.edu.ar](mailto:lemac@frlp.utn.edu.ar), web: [www.utn.edu.ar/lemac](http://www.utn.edu.ar/lemac)

## RESUMEN

El análisis de asociación permite identificar la existencia de una relación entre dos o más variables, lo cual utilizando un estadístico de prueba posibilita determinar la influencia entre un atributo medido en un evento de tránsito y la hospitalización del participante, posibilitándose la obtención de patrones relacionados a la gravedad del siniestro en vinculación a su necesidad de atención sanitaria.

Dado que las causas y consecuencias de los siniestros difieren de acuerdo al sector que se estudie, en este caso la ciudad de Trenque Lauquen de aproximadamente 40.000 habitantes, la metodología desarrollada contempla la agrupación por cuadrículas y la agrupación de las mediciones de cada incidente atendiendo a las variables de los factores Humano, Vehicular y del Ambiente.

La aplicación del análisis estadístico de referencia se valida en base a la aplicación del modelo sobre datos no utilizados en el análisis. Los resultados en todos los casos permiten concluir que el modelo brinda información precisa para elaborar un diagnóstico o evaluar una intervención.

El presente estudio, que surge del desarrollo de una Tesis para optar al título de Magister en Administración de Negocios, establece la caracterización de cada zona en función de su resultado más crítico, que es el impacto sanitario en la persona involucrada. En base a ésta descripción es posible adecuar estrategias preventivas en forma eficiente y ligar el gasto presupuestario a programas de simple evaluación.

**PALABRAS CLAVES:** seguridad vial, análisis de asociación de variables, estrategias sanitarias preventivas.

## VARIABLES ASSOCIATION ANALYSIS IN ROAD INCIDENTS WITHIN AN GEOREFERENCED AREA TO ESTABLISH ODDS OF HOSPITALIZATION. AN APPLICATION

### ABSTRACT

The association analysis identifies the existence of a relationship between two or more variables, this using a statistical test allows to determine the influence between a attribute measured in an traffic event and the hospitalization of the participant, allowing obtaining patterns related the seriousness of the incident in connection to their need for health care.

Since the causes and consequences of the incident differ according to the sector being studied, in this case Trenque Lauquen city of about 40,000 habitants, the methodology developed use a grouping in zones and a grouping of measurements of each incident in response to the changes of Humans, Vehicles and Environments factors.

The application of statistical analysis is validated based on the application of the model with data not used in the analysis. The results in all cases allows concluded that the model provides accurate information for make a diagnosis or assess an intervention.

The present study, which arises from the development of a thesis of a Master in Business Administration, provides the characterization of each zone according to their most critical result, which is the health impact on the person involved. Based on this description it is possible to tailor preventive strategies efficiently and link the spending with simple evaluation programs.

**KEY WORDS:** road safety, variables association analysis, preventive health strategies.

## 1- INTRODUCCIÓN

Las características viales de una ciudad, en este caso Trenque Lauquen (al oeste de la Provincia de Buenos Aires, Argentina), pueden nuclearse en zonas de acuerdo a diversos factores concurrentes, por ejemplo: vías que poseen semáforos o reductores de velocidad, situadas en el microcentro o en zonas aledañas, con prevalencia de autos, motos o bicicletas, etc.

Contando con una base de datos de los siniestros viales detectados en cada una de estas zonas, se puede analizar la asociación entre las variables medidas y la hospitalización del siniestrado. De esta manera se pueden caracterizar los siniestros viales de cada una de las zonas que conforman la ciudad, tomando como eje la necesidad de hospitalización del siniestrado, brindándose un importante aporte a la construcción e implementación de políticas preventivas eficientes.

Un estudio de este tipo puede conducirse mediante un análisis de asociación de variables, por ejemplo Chi Cuadrado ( $\chi^2$ ), como fuera empleado en el análisis de morbilidad de Hidalgo-Solórzano et al [1].

Al encontrarse la asociación con la hospitalización del siniestrado es posible construir estrategias focalizadas en aspectos críticos que hacen a la colisión, como pueden ser la edad, el modo de desplazamiento, el sexo, el tipo de clima, entre otros. La utilización de éste tipo de herramientas estadísticas a nivel sanitario es ampliamente difundido, como afirma Taucher [2] cuando busca dependencias entre patologías y sus consecuencias.

Estudios internacionales destacan la necesidad de comprender la incidencia de los diversos factores, que influyen al momento de ocurrir un siniestro [3,4,5]. El presente estudio aborda este análisis, arribando a conclusiones que describen los parámetros estratégicos para centrar los esfuerzos de prevención y control en las zonas que estadísticamente pueden ser caracterizadas.

En base a lo expuesto, se pretende establecer una herramienta que permita identificar las variables que se asocian a la hospitalización del siniestrado dentro de un área determinada. Alcanzar éste conocimiento facilita la construcción de estrategias por parte de autoridades municipales, focalizándose en aspectos sanitarios críticos que hacen a la colisión.

## 2- MATERIALES Y MÉTODOS

Previo a proyectar el análisis estadístico se definen las formas que tomarán los datos para lograr acceder a las relaciones buscadas. El análisis estadístico elegido de  $\chi^2$  tiene ciertas especificaciones y requisitos con reglas pre-establecidas. Estadísticamente el enfoque busca encontrar aquellas variables que se asocian a la dependiente [6].

Las variables que conforman los datos son dables vincular según el factor con el que están relacionadas [7], esto le otorga a los responsables del diseño de intervenciones (Gobierno Municipal de Trenque Lauquen, Concejo Deliberante de Trenque Lauquen, etc.) la posibilidad de proyectar actividades preventivas orientadas en el “vector” interviniente.

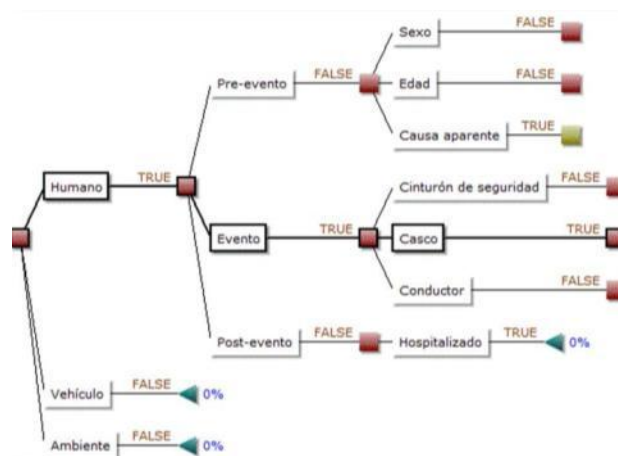


Figura 1. Árbol de decisión: Variables vinculadas al Factor Humano

Para posibilitar el desarrollo del estudio se procede a la construcción de una base de datos que contiene diversas Tablas de interés. Esto conlleva al empleo de un esquema de árbol para el armado de estas Tablas. A manera de ejemplo en la Figura 1 se observa el árbol básico desarrollado para la construcción



cinturon	Cinturón	casco	Casco	conductor	Conductor	hospitalizad	Hospitalizad	AO
No		0 Si	1	1 Si		1 Si		1
No		0 No		0 Si		1 Si		1
No		0 Si		1 Si		1 Si		1
No		0 Si		1 No		1 No		0

Figura 3. Vista parcial de la transformación de variables

Cada variable se transforma en un valor binario agregando las columnas correspondientes, esto resulta muy simple utilizando el condicional de =SI("VALOR COLUMNA X"="SI";1;0).

En el caso de la variable "edad", se crea una columna para cada grupo etario según la clasificación internacional utilizada por la Organización Mundial de la Salud [8]. Esto permite comparar los indicadores con los resultados nacionales e internacionales. La fórmula utilizada para el grupo etario de ejemplo, 15 a 24 años, es =SI(Y("VALOR COLUMNA EDAD">14;"VALOR COLUMNA EDAD"<25);1;0).

Las variables que identifican si el siniestro se produce al intentar embestir un peatón, un animal u otro vehículo, son agrupadas como "evitar embestir" debido a que se intenta analizar la maniobra independientemente del objeto con el que colisiona. La fórmula en éste caso es =SI(O("EMBESTIR PEATÓN"="Si";"EMBESTIR VEHÍCULO"="Si";"EMBESTIR ANIMAL"="Si");1;0).

Estos son ejemplos de los tratamientos encarados en cada caso, siendo dicha metodología de trabajo extrapolable a la totalidad de los aspectos considerados.

### Ordenamiento de tablas en función de $\chi^2$

Una vez divididas las cuadrículas y agrupadas las variables en los factores intervinientes, se genera una estructura de datos que simplifique el cálculo de variables observadas y esperadas en función de  $\chi^2$ . Para ello cada variable recibe un idéntico tratamiento. Por ejemplo para la variable "Conductor", en primer lugar se identifica si la variable analizada se ubica dentro de la cuadrícula de estudio y luego se filtra sobre el total de Hospitalizados aquellos que requirieron atención sanitaria y sufrieron el siniestro en la cuadrícula de estudio.

Agrupada la variable positiva + hospitalizados + cuadrícula de análisis, se incorpora una nueva columna, incluyendo también la no ocurrencia del punto anterior, variable negativa + no hospitalizado + cuadrícula de análisis.

Ésta construcción permite completar la matriz de Frecuencias observadas y esperadas, las cuales constituyen la base del cálculo de  $\chi^2$ .

Como el caso se encuadra como una matriz de dos por dos, si el valor obtenido es mayor a 3,84 (valor crítico para un grado de libertad), entonces se puede afirmar que existe asociación entre la variable independiente y "hospitalizado" (variable dependiente).

### 3- RESULTADOS

Se realiza el registro con una cantidad total de siniestrados en el período determinado de 1.613, de los cuales la cantidad total de siniestrados en condiciones de modelizar es de 1.244 y la cantidad total de siniestrados hospitalizados es de 827.

Se desagregan los participantes de acuerdo a la cuadrícula donde se produjo el evento, Tabla 1, dando lugar a 30 áreas de análisis.

Al ordenar las variables dentro de cada cuadrícula según el factor interviniente (Humano, Vehículo, Ambiente), es posible analizar cada atributo junto a la dependiente en función del modelo de  $\chi^2$ , según de observa en la Tabla 2.

Tabla 1. Cantidad de registros por cuadrícula

Cuadrícula	A0	A1	A2	A3	A4	A5
Registros	17	19	69	44	19	8
Cuadrícula	B0	B1	B2	B3	B4	B5
Registros	25	24	144	93	49	25
Cuadrícula	C0	C1	C2	C3	C4	C5

Registros	25	34	166	163	64	2
Cuadrícula	D0	D1	D2	D3	D4	D5
Registros	3	29	64	69	18	10
Cuadrícula	E0	E1	E2	E3	E4	E5
Registros	2	7	14	32	4	2

Tabla 2. Asociaciones encontradas por cuadrícula

Cuadrícula	Existe asociación con la hospitalización del siniestrado		
	Factor Humano	Factor Ambiente	Factor Vehículo
A2	No	Noche - Día	Moto - Auto
A3	No	No	Auto - Moto
B0	No	No	Auto - Moto
B2	No	Tipo de vía Rotonda	Moto - Auto
B3	No	Tiempo bueno	Auto - Moto
B4	No	No	Moto - Auto
C1	No	Lugar Intersección	Moto - Auto
C2	No	No	Auto - Moto
C3	No	No	Auto – Moto – Colis. lateral
C4	No	Prioridad de paso peatonal	Auto - Moto
D1	No	No	Moto
D2	No	No	Auto - Moto
D3	No	No	Moto - Auto
E3	No	No	Moto

Sobre el plano se puede visualizar claramente el comportamiento de las cuadrículas en función de la asociación con la hospitalización de los participantes del evento. En la Figura 4 se puede observar sobre cuáles cuadrículas impactan las variables del Factor Ambiental asociadas al requerimiento de asistencia sanitaria.



Figura 4 Impacto del Factor Ambiental en la hospitalización del siniestrado

En el caso del Factor Vehicular, se detecta asociación en las cuadrículas que se observan en la Figura 5.





Figura 5 Impacto del Factor Vehicular en la hospitalización del siniestrado

Ninguna de las variables analizadas del Factor Humano se asocia a la hospitalización del participante del evento vial.

#### 4- DISCUSIÓN

##### Validación del análisis

Con el fin de comprobar que los modelos desarrollados se ajustan correctamente a los resultados esperados, la afirmación o rechazo de las Hipótesis, dentro de cada cuadrícula se contrasta cada variable independiente a la hospitalización del siniestrado.

Dicho tratamiento, agrupado según el factor interviniente, permite identificar asociaciones estadísticas que sirvan como base para caracterizar las zonas de la ciudad de Trenque Lauquen.

Bajo el método de  $\chi^2$  se logra generar una Tabla que detalla las variables cuya presencia se vincule a la necesidad de recibir atención sanitaria, parámetro que se acepta como válido para definir consecuencias de mayor gravedad en el resultado de un siniestro vial.

Para validar el análisis se toman datos adicionales no considerados, correspondientes a Noviembre de 2013. Sobre esos registros se descartan aquellos cuyas ubicaciones no se contemplaban dentro de las cuadrículas en las que se afirmaron asociaciones. Se realiza una validación por columna georreferenciada, detallando en una Tabla la presencia o ausencia de cada variable y su consecuencia sanitaria, donde:

- Presencia de variable dependiente<sub>n</sub> = 1
- Ausencia de variable dependiente<sub>n</sub> = 0
- Consecuencia de variable dependiente<sub>n</sub> "hospitalizado" = 1
- Consecuencia de variable dependiente<sub>n</sub> "no hospitalizado" = 0

Para la elección de la fila se opta por aquellas en las que se ha reconocido alguna variable asociada, siendo ésta condición indispensable para ser incluida entre las posibles a validar.

A modo de ejemplo, puede observarse los resultados obtenidos en el Primer Caso. La cuadrícula seleccionada es la A2, la cual se encuentra en el sector centro oeste de la ciudad.

En éste sector se analiza en el registro de siniestrados si se contaba con las características que aportan las variables estadísticamente asociadas: Día – Noche - Moto –Auto, estableciéndose la descripción que se representa en la Tabla 3. Como se puede observar el resultado obtenido es adecuado, ya que cada variable asociada a hospitalización mostró como consecuencia la necesidad de hospitalización del participante.

La validación resulta con similares resultado en el resto del área en estudio, razón por la cual se pueden considerar validadas las conclusiones arribadas.

Tabla 3. Cantidad de variables asociadas contabilizadas/hospitalización - Cuadrícula A2 - Nov 2013

Variable	Cantidad	Hospitalizado
Día	1	1

Noche	0	0
Moto	1	1
Auto	1	1

### Otras formas de análisis aplicables

El análisis de relación entre variables puede lograrse a través de diferentes técnicas, entre las cuales se seleccionó la de  $\chi^2$  por las razones que ya se han descrito, cuestión que no invalida que otra línea de estudio adopte herramientas diferentes.

La técnica de Data Mining (DM) es ampliamente utilizada en numerosos estudios de tráfico a nivel Internacional, contribuyendo al asesoramiento en cuestiones específicas de seguridad vial [9]. Sin embargo, se plantea una dificultad insalvable cuando se pretende avanzar en el conocimiento sobre la siniestralidad vial en ámbitos Municipales: la falta de datos. Las experiencias locales en pequeñas regiones de la Argentina, son incipientes en la creación de bases de datos que constituyan el soporte de modelos predictivos que combinen la trilogía vial (Humano/Vehículo/Ambiente).

Si bien el estudio se enmarca en un proyecto que implementa una plataforma que permite vincular las fuentes de información, se encuentra una limitación en la robustez de la base de datos.

Ésta restricción se observa claramente en la cantidad de siniestros por cuadrícula, por ello se refuerza el enfoque en  $\chi^2$ . Seguramente a medida que se logre salvar ésta barrera se conformen otros estudios que avancen en modelos cuyos rendimientos mejoren los obtenidos en el análisis.

### Otros aspectos de discusión

En general, los modos de desplazamiento “Auto” y “Moto” son las variables que presentan asociación con la hospitalización del siniestrado. Cabe recordar que éstas variables pertenecen al factor Vehículo y si bien se ha trabajado con matrices de dos por dos con una variable independiente y otra dependiente, se debe plantear la posibilidad que nuevos estudios incluyan más de un atributo independiente.

El análisis muestra que la mayoría de las asociaciones ocurren para ambos factores en forma simultánea, es decir, en la cuadrículas en que se encuentra relación entre Moto y Hospitalizado generalmente también se presenta entre Auto/Moto. Por ello se observa que es posible ajustar estrategias de prevención en función del uso más eficiente de los recursos.

Adicionalmente en este sentido, cabe aclarar que la georreferencia de los siniestros muestra que cada zona tiene particularidades que difieren del comportamiento que se observa a nivel general dentro del territorio Municipal, aspecto que debe ser considerado.

## 5- CONCLUSIONES Y RECOMENDACIONES

### Conclusiones

- El análisis de la siniestralidad vial sobre el eje de los factores que intervienen, es un concepto que integra perfectamente las causales posibles de un evento. A nivel local y regional se deben implementar, y en aquellos lugares que existen mejorar, los procesos de centralización de datos comúnmente nucleados bajo Observatorios.
- La georreferenciación de los datos permite estudiar las características de los siniestros cuyos resultados provocan la atención sanitaria del involucrado, dentro de un área de estudio delimitada. Éste método, brinda el conocimiento básico para establecer políticas de prevención más eficientes. El dato sanitario georreferenciado no siempre está disponible en los registros, aspecto que perjudica la conformación de tablas de análisis robustas.
- La metodología utilizada permite estudiar las asociaciones de factores con la hospitalización del siniestrado. Para obtener modelos que aporten mayor información existen otras herramientas basadas en Data Mining. Para avanzar en modelos que exploten la minería de datos es imprescindible mejorar la calidad de los registros y garantizar su continuidad.
- El empleo del árbol de decisión permite verificar que el estudio dispone de variables que midan todos los factores intervinientes, esto posibilita el desarrollo de modelos que se ajustan a las necesidades de información en aplicaciones de éstas características.
- La medición de las variables del siniestro en sí mismo son adecuadas, sin embargo para obtener una comprensión del entorno que contiene cada cuadrícula se requiere de mediciones sociales y de flujo vehicular que la base de información no contiene.
- El adecuado tratamiento a la preparación de las variables, le brinda al modelo la confiabilidad necesaria para el tratamiento estadístico elegido. Así mismo, sienta las bases para futuras aplicaciones de análisis que profundicen el estudio.

- A través de los casos analizados se corrobora que las aplicaciones desarrolladas son confiables, siendo sus empleos sencillos y precisos. Estas dos cualidades son las recomendadas para el desarrollo de modelos estadísticos.
- La búsqueda de asociación, entre variables que describen un siniestro vial y la posible asociación con la hospitalización del participante a través del método de  $\chi^2$  es una buena opción, principalmente en bases de datos cuya robustez no sea la apropiada para técnicas de mayor profundidad de análisis.
- El método de división de cuadrículas sobre un plano, en una base de datos georreferenciada permite caracterizar la siniestralidad vial de diferentes zonas de una misma ciudad, donde las relaciones entre variables permiten acceder a información representativa de un espacio reducido.
- En base a lo expuesto se llega a la conclusión final que si el método se utiliza sobre bases de datos adecuadas, en áreas delimitadas y observando las condiciones de restricción propias del estadístico  $\chi^2$ , los resultados son confiables.

## Recomendaciones

- Se recomienda profundizar el sistema de registro sanitario, para lograr determinar los diferentes niveles de gravedad que son consecuencia de las lesiones causadas por el tránsito, diferenciando las lesiones graves de aquellas leves y aplicando el modelo planteado, mediante lo cual los resultados se podrán establecer sobre el eje de gravedad que se desee estimar.
- Replicar la metodología aquí expuesta en otras ciudades de similares características para obtener puntos de comparación, con los cuales el traslado de experiencias exitosas bajo homogéneas circunstancias anticipen los resultados.
- Incorporar datos socio-económicos a las cuadrículas de análisis, con el fin de medir el impacto que éstos tienen sobre los factores de la trilogía vial permitiendo lograr una caracterización más profunda de las áreas de estudio.
- Experimentar buenas prácticas sobre las variables asociadas a la hospitalización, definidas por cuadrículas, y medir sus variaciones, posibilitándose la obtención de métodos de validación más ajustados a la resolución de la problemática.

## 6- REFERENCIAS BIBLIOGRÁFICAS

1. HIDALGO-SOLÓRZANO EC, HÍJAR M, BLANCO-MUÑOZ J, KAGEYAMA-ESCOBAR ML. "Factores asociados con la gravedad de lesiones ocurridas en la vía pública en Cuernavaca", Salud Pública 47:30-38, Morelos, México. 2005.
2. TAUCHER E. "Bioestadística", Editorial Universitaria, Chile, 1999.
3. MOHAN D, TIWARI G, MELECKIDZEDECK KHAAYESI F, NAFUNKHO M. "Road traffic injury prevention: training manual", Organización Mundial de la Salud, EEUU, 2008.
4. FUENTES FC, HERNANDEZ HV. "La estructura espacial urbana y la incidencia de accidentes de tránsito en Tijuana", ISSN 0187-7372, Frontera Norte V21, Baja California, México, 2009.
5. GELDSTEIN R, BERTONCELLO R. "Aspectos demográficos y sociales de los accidentes de tránsito en áreas seleccionadas de la Argentina diagnóstico y aportes para el diseño de políticas y programas de prevención", CONAPRIS, Ministerio de Salud y Ambiente de la Nación, Argentina, 2006.
6. DOORMAN FJ, MIRANDA F, DE NIE C, OOIJENS J, OVARES L, RAMIREZ C, SAENZ C, SANCHO E. "La metodología del diagnóstico en el enfoque investigación adaptativa", ISBN 90 5187 068X, Costa Rica, 1991.
7. AEC. "Libro Verde de la Seguridad Vial", ISBN: 84-89875-65-0, Asociación Española de la Carretera, España, 2006.



8. OMS. "Informe sobre la situación mundial de la seguridad vial, es hora de pasar a la acción", ISBN 978 92 4 356384 8, Organización mundial de la Salud, Suiza, 2009.
9. PERINETTI A, UBEDA C, UNGARO J. "Lesiones por tránsito en el partido bonaerense de General Pueyrredón: epidemiología, factores de riesgo y letalidad", Revista Argentina de Salud Pública – saludinvestiga.org.ar/rasp/articulos/volumen16/24-30.pdf, Argentina, 2013.

## **SOBRE LOS AUTORES**

**Martínez Micakoski, Fernanda:** Profesora Adjunta Interina de Proyecto Final en la carrera de Ingeniería Industrial UTN Trenque Lauquen. Investigador Categoría V en Programa de Incentivos de la Nación y Categoría D de la UTN. Especialista en Ingeniería Gerencial y Licenciada en Organización Industrial por la UTN. Consejera Docente del Departamento de Ingeniería Industrial de la UTN Trenque Lauquen. Directora de Consultores del Proyecto Observatorios Municipales de Seguridad Vial de la Agencia Nacional de Seguridad Vial Argentina – Banco Mundial.

**Rivera, José Julián:** Subdirector del LEMaC – UTN La Plata. Profesor Adjunto Ordinario de Vías III en la UTN La Plata. Investigador Categoría II en Programa de Incentivos de la Nación y Categoría B de la UTN. Magister en Transporte y Logística e Ingeniero Civil por la UTN. Consejero Docente del Departamento de Ingeniería Civil de la UTN La Plata. Becario de la Universidad de Zaragoza, España.