

**Modelos de generación de viajes motivo
Estudio: caso universidad de la costa**

AUTORES:

**Aixa Liliana Gómez Gómez
Jorge Alberto Medina Romero**

**Trabajo de grado para optar título de
INGENIERIA CIVIL**

**Universidad de la costa C.U.C
Departamento de Civil y Ambiental
Ingeniería Civil
Barranquilla
2018**

**Modelos de generación de viajes motivo
Estudio: caso universidad de la costa**

AUTORES:

**Aixa Liliana Gómez Gómez
Jorge Alberto Medina Romero**

Trabajo de grado presentado al tutor

Iván Darío Serrano Arrieta

**UNIVERSIDAD DE LA COSTA
DEPARTAMENTO DE CIVIL Y AMBIENTAL
INGENIERIA CIVIL
BARRANQUILLA
2018**

Nota de aceptación

Firma del Decano

Firma Líder de Grupo de Investigación

Firma de Tutor

Firma de Cotutor

Firma de Juez

Firma de Juez

Dedicatoria

Este proyecto de grado es dedicado a Dios y a mi familia, especialmente a mi adorada madre Ana Oliva y a mi padre Franklin, como también a Jorge por ser parte de mi proyecto de vida.

Aixa Gómez.

Este proyecto de grado es dedicado a Dios, especialmente a mi madre Carmen y mi padre Idelfonso, mis hermanos y demás familiares.

Jorge Medina.

Agradecimientos

Agradecemos principalmente a DIOS por darnos la oportunidad, la capacidad y la disciplina a lo largo de esta carrera, también a Nuestro tutor el ingeniero Iván Darío Serrano Arrieta por ser nuestra guía a lo largo de este proyecto, a nuestros padres por brindarnos lo más valioso que se le puede dejar a un hijo, la educación.

Autores

Resumen

En el presente trabajo se estimaron modelos de generación para viajes motivo estudio de la Universidad de la costa. Los datos utilizados se obtuvieron a partir de una encuesta origen-destino realizada a estudiantes de la universidad. Se estimaron modelos de generación usando al menos tres métodos: Regresión Lineal Múltiple (RLM), análisis por categoría (AC) y Logit Ordinal (LO). De acuerdo a los resultados los modelos (LO) mostraron mayor consistencia econométrica y mejores indicadores de bondad de ajuste. Los modelos utilizados son clave para la planeación estratégica en materia de movilidad para la universidad de la costa. La generación de viajes es un proceso que permite relacionar las actividades de la población con los viajes que se realizan, estos últimos están íntimamente ligados a las características socioeconómicas de la población, esta relación se puede estimar a través de modelos de generación empleados en el presente proyecto. Se analizaron las tasas medias de viajes obtenidas con los modelos de viajes por categorías y Logit Ordinal, así mismo se generaron modelos por el método de Regresión Lineal Múltiple y se evaluaron los modelos obtenidos a través de los diferentes test estadísticos para seleccionar los más confiables.

***Palabras Clave:** Generación de viajes, Análisis por Categorías, Análisis de Clasificación Lineal múltiple, Regresión lineal múltiple, Logit Ordinal*

Abstract

In the present work, generation models were estimated for travel study reasons of the University of the coast. The data used was obtained from a source-destination survey carried out for university students. Generation models were estimated using at least three methods: Multiple Linear Regression (RLM), category analysis (AC) and Ordinal Logit (LO). According to the results, the models (LO) showed greater econometric consistency and better indicators of goodness of fit. The models used are key for strategic planning in terms of mobility for the university of the coast. The generation of trips is a process that allows to relate the activities of the population with the trips that are made, the latter are intimately linked to the socioeconomic characteristics of the population, this relationship can be estimated through generation models used in the present draft. The average travel rates obtained with the travel models by categories and Ordinal Logit were analyzed, as well as models were generated by the Multiple Linear Regression method and the models obtained through the different statistical tests were evaluated to select the most reliable.

Keywords: *Trip generation, Analysis by Categories, Multiple Linear Classification Analysis, Multiple Linear Regression, Ordinal Logit*

Contenido

1	Introducción.....	11
2	Objetivo General.....	13
3	Objetivos Específicos	13
4	Marco Teórico	14
4.1	Modelo de Regresión lineal múltiple (RLM)	14
4.2	Análisis por categorías (AC).....	15
4.3	Modelo Logit ordinal (LO)	17
5	Metodología.....	23
6	Análisis de datos.....	24
6.1	Descripción de las variables dependientes.....	25
6.2	Descripción de variables discretas.....	28
7	CALIBRACIÓN Y SELECCIÓN DE MODELOS	30
7.1	Regresión Múltiple	32
7.2	Análisis por categorías (AC).....	37
7.3	Modelo Logit Ordinal (LO)	41
8	Evaluación de los modelos con fines de predicción.....	43
9	Conclusiones.....	45
10	Anexos: Resultados modelos logit ordinal	47
11	Referencias	51

Lista de tablas y figuras

Tablas.

Tabla 1. Descripción de variables de estudio. 25

Tabla 2. Estadísticos de tendencia central y de dispersión para las variables dependientes 26

Tabla 3. Estadísticos estratos en la muestra 28

Tabla 4. Estadísticos Tasa Motorización 28

Tabla 5. Estadísticos Ocupación Hogar..... 30

Tabla 6. Matriz de Correlación de variables cuantitativas para el modelo RLM con
propósito Estudio..... 32

Tabla 7. Modelo 1 RLM. 34

Tabla 8. Modelo 2 RLM. 35

Tabla 9. Modelo 2 RLM. 36

Tabla 10. No. de Hogares por categoría 39

Tabla 11. Número de Viajes por categoría con motivo Estudio..... 40

Tabla 12. Tasas de generación por categorías para viajes con motivo Estudio..... 40

Tabla 13. Errores de los modelos de RLM calibrados para cada propósito de viaje..... 43

Tabla 14. Cálculo de los NRMSE para cada relación de categorías de variables y propósito
de viaje..... 44

Tabla 15. Errores obtenidos de la validación de los modelos LO generados para cada
propósito de viaje. 44

Figuras.

Figura 1. Interrelaciones básicas. 11

Figura 2. Gráfico de la curva de la función logit..... 19

Figura 3. Diagrama de caja para las distribuciones de viaje. 27

Figura 4. Gráfico de Normalidad en los residuos del modelo RLM para la generación de
 Figura viajes con propósito Estudio 36

Figura 5. Histograma para el Nivel de Ingreso..... 38

Figura 6. Histograma de frecuencias de viajes en el día con propósito Estudio 41

1 Introducción

Los modelos de generación desagregados tienen como unidad de trabajo el hogar e intentan relacionar los viajes generados en los domicilios con características o atributos tales como cantidad de personas que viven en el hogar, tasa de motorización y nivel de ingresos. Para planificar los sistemas de transporte urbano, se requiere el desarrollo de modelos con el fin de predecir las demandas en virtud de los viajes urbanos en un periodo de tiempo determinado. La manera tradicional de abordar el problema de transporte consiste en estudiar la existencia de tres macro-variables: sistema de actividades, sistema de transporte y patrón de flujos, que entre sí exhiben interrelaciones de tres tipos, como se puede ver esquemáticamente en la ilustración 1.

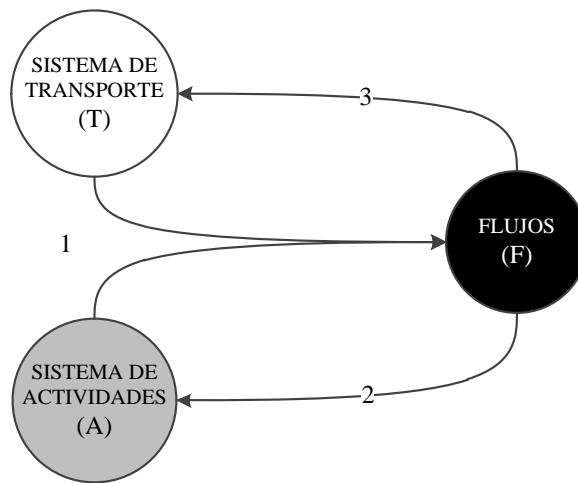


Figura 1. **Interrelaciones básicas.**

Fuente: M. L. Manheim / Fundamentals of Transportation Systems Analysis, Volume 1: Basic Concepts (1979) 13

La relación de tipo 1 indica que el patrón de flujos depende tanto del sistema de transporte como del sistema de actividades, la relación tipo 2 muestra que en el largo plazo los flujos cambian el sistema de actividades a través del nivel de servicio y el consumo de recursos, y la relación tipo 3 revela que a largo y mediano plazo, dados los flujos y sus

niveles de servicio la sociedad reacciona (Manheim, 1979). Entonces, cualquier cambio que afecte el sistema de transporte o el sistema de actividades, o los dos en forma simultánea, producirá una variación en el patrón de flujos.

No obstante, años de experimentación y desarrollo han dado como resultado la definición de una estructura general de modelación denominada modelo clásico de transporte, cuya primera fase, la generación de viajes, tiene por objetivo predecir el número total de viajes producidos y atraídos por cada zona del área de estudio.

En esta línea, el objetivo del presente trabajo no es otro que calibrar, comparar y elegir unos modelos de generación de viajes, con base en una información de encuestas origen-destino aplicadas en la Universidad de la Costa.

Se presentarán calibrados modelos de distintas clases, entre ellos: modelos de regresión lineal, modelos de análisis por categorías y modelos de tipo Logit ordinal. En todo caso se intenta ser rigurosos en los análisis estadísticos y en las comparaciones realizadas entre los distintos modelos obtenidos con el fin de escoger el mejor de ellos y evaluarlo con fines predictivos.

2. Objetivos

2 Objetivo General

- Proponer y estimar modelos de generación de viajes motivo estudio para los estudiantes de la Universidad de la Costa.

3 Objetivos Específicos

- Realizar una revisión sistemática acerca de modelos de generación de viajes para el modelo clásico del transporte.
- Realizar encuestas origen-destino sobre la comunidad de estudiantes de la CUC con el fin de obtener una base de datos del comportamiento de la demanda de viajes motivo estudio.
- Analizar estadísticamente los datos para obtener una medida descriptiva del comportamiento de la demanda.
- Estimar, comparar y validar modelos de regresión línea múltiple, análisis por categorías y el logit ordinal para explicar la generación de viajes motivo estudio de los estudiantes de la Universidad de la Costa.

4 Marco Teórico

El modelo clásico del transporte se ha mantenido desde finales de los años 60 aparentemente inalterado. Este enfoque de modelación parte por considerar una red multimodal de transporte, una zonificación apropiada del área de estudio, y la recolección y codificación de datos tanto para la calibración y validación de los modelos como para su uso en modalidad predictiva. Estos datos incluyen información socioeconómica acerca de la población en cada zona del área de estudio.

Con estos datos obtenidos típicamente de una encuesta origen- destino, se pueden estimar modelos que permitan predecir el total de viajes generados y atraídos por cada zona (generación de viajes) (Ortúzar, 2000).

La literatura exhibe variadas técnicas para estimar los modelos de generación, en los que se procura estimar es el número de viajes producidos sobre la base de relaciones generalmente lineales. A continuación, se presentan las más actuales metodologías.

4.1 Modelo de Regresión lineal múltiple (RLM)

Cuando se desea relacionar un conjunto de datos observados acerca del resultado de un fenómeno o experimento (variable independiente Y), con la cantidad que se agregue de ciertos atributos o características (X), se utilizan técnicas de ajuste (como mínimos cuadrados) que permita calcular una función que explique los valores observados al menor error posible. El modelo de regresión lineal es una herramienta con mucha popularidad en ciencias exactas y naturales (Wonnacott, 1979).

El modelo de regresión lineal en general se expresa de la siguiente forma:

$$Y_i = \alpha + \beta X_i + e_i$$

El modelo presenta los siguientes supuestos a tener en cuenta en la calibración:

- Linealidad.
- Independencia: De la variable aleatoria “residuos”
- Homocedasticidad o igualdad de varianzas de los residuos y los pronósticos.
- Normalidad de los residuos tipificados.
- No colinealidad.

Para modelos de regresión lineal con el objetivo de estimar los viajes originados en una zona, se consideró que la unidad de análisis más adecuada era el hogar, ya que preexisten vínculos interpersonales importantes, al interior del hogar, que un modelo a nivel individual no tendría en cuenta (Ortúzar, 2000). Por lo tanto, se considera al hogar como parte de un vector de datos de entrada, donde cada una de sus características más relevantes y su comportamiento, explicarían la tasa de viajes de la unidad de análisis. Un modelo de regresión a nivel de hogar basado en tasas tendría la siguiente forma:

$$y_i = \theta_0 + \theta_1 x_{1i} + \theta_2 x_{2i} + \dots + \theta_k x_{ki} + e_i$$

En que $y_i = Y_i/H_i$; $x_{ki} = X_{ki}/H_i$; $e_i = E_i/H_i$ y H_i denota el número de hogares en la zona i . En la experiencia investigativa se ha encontrado que los multiplicadores $(1/H_i)$ permite reducir la heterocedasticidad ya que se logra independizar al modelo del tamaño zonal.

4.2 Análisis por categorías (AC)

Este es un procedimiento que se basa en determinar las producciones de viaje en función de las características del hogar. Supone que las tasas de generación (viajes por día)

son constantes en el tiempo para las determinadas categorías de hogares. El método estima las tasas de generación de viajes de manera empírica, el arte del método consiste en elegir las categorías de modo que las desviaciones estándar de las distribuciones de frecuencia sean mínimas. Para construir el modelo se requiere del siguiente modelo:

$$O_i^n = \sum_{h \in H(n)} a_i(h) * T(h)$$

Donde n es el tipo de persona y h el tipo de hogar, $T(h)$ el promedio de viajes con el propósito dado, el cual se establece dividiendo el número de viajes por el número de personas en cada grupo. $a_i(h)$ es el número de hogares tipo h en la zona i . (Ortúzar, 2000).

Este enfoque de modelación requiere conocer el número de hogares por categoría en cada zona, lo cual debe ser determinado o estimado a partir de información socioeconómica independiente, normalmente del Censo Poblacional o de otros catastros urbanos. Es necesario además conocer la distribución de hogares por categoría no sólo en el año base de análisis; también se requiere la distribución futura de los hogares para cada uno de los cortes temporales. Estas proyecciones son parte del ámbito de especialización de otras disciplinas. (FDC 2010)

Finalmente, puede resultar que al estimar las tasas se obtengan resultados contraintuitivos, es decir, tasas que no sean, por ejemplo, crecientes según nivel de ingreso y motorización (variables relevantes de esta metodología y que como se ha indicado, corresponden a una decisión de compromiso). Una forma que se ha demostrado útil en la práctica para evitar este problema es el de agregación de categorías, para efectos del cálculo (FDC Consultores, 2010).

Entre las principales ventajas del uso de este método se destacan lo básico, sencillo e intuitivo que es, y el hecho de que no se establecen relaciones matemáticas entre la tasa de generación de viajes y las variables independientes. Entre las desventajas se encuentra que las variables son escogidas antes de empezar el modelo y no se puede verificar la significancia de estas; además, se requiere de una muestra muy grande; y si se quiere aumentar el número de variables, se debe aumentar mucho la muestra, por lo que resulta difícil predecir el número de hogares en cada categoría a futuro. (Ortúzar 2000).

4.3 Modelo Logit ordinal (LO)

Un modelo de elección discreta describe las elecciones entre diferentes alternativas de los tomadores de decisiones, y se derivan bajo el supuesto de un comportamiento maximizador de la utilidad del tomador de decisiones. El conjunto de alternativas, llamado también conjunto de elección, debe presentar las siguientes características:

- Las alternativas deben ser mutuamente excluyentes desde la perspectiva del tomador de decisiones. Es decir, escoger una alternativa implica necesariamente que no se escoja ninguna de las otras restantes.
- El conjunto de elección debe ser exhaustivo, en el que se incluyan todas las alternativas posibles.
- El número de alternativas debe ser finito.

En esencia, la *elección* puede ser vista como el resultado de un proceso secuencial de toma de decisiones, que consta de cinco pasos (Ben-Akiva & Lerman, 1985):

- Definición del problema de elección
- Generación de alternativas
- Evaluación de atributos de la alternativa
- Elección
- Implementación

Uno de los modelos más ampliamente utilizados en la formulación de modelos de elección discreta es el Logit. Su popularidad se debe a que la fórmula para las probabilidades de elección es fácilmente interpretable. Sea n el individuo (tomador de decisiones), y J el conjunto de alternativas a las que se enfrenta, la utilidad que el mismo obtiene al elegir la alternativa j se descompone en dos partes: una, que es la estimada por el investigador designada por V_{nj} ; y la otra, desconocida, que se denomina error aleatorio ε_{nj} , por tanto $U_{nj} = V_{nj} + \varepsilon_{nj}$.

El modelo logit se obtiene asumiendo que cada error es independiente e idénticamente distribuido en sus valores extremos. (Train, 2009). La probabilidad de que un tomador de decisiones n seleccione la alternativa i está dada por:

$$P_{ni} = Prob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}, \forall j \neq i)$$

O bien,

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} = \frac{e^{\beta'x_{ni}}}{\sum_j e^{\beta'x_{nj}}}$$

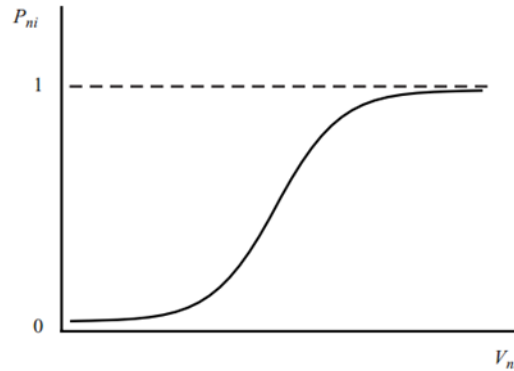
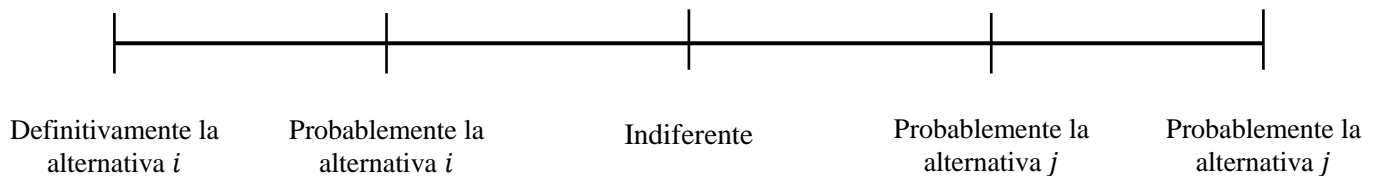


Figura 2. Gráfico de la curva de la función logit.

Fuente: (Kennet E. Train, 2009, Capítulo 3, página 38)

4.3.1 Logit Ordinal

Cuando la variable dependiente tiene más de dos categorías y los valores de cada una tienen un orden secuencial significativo donde un valor es sin duda mayor que el anterior, entonces se tiene un logit ordinal (Torres-Reyna). Un modelo de elección binaria ordinal se deriva cuando hay disponibles respuestas ordenadas, donde el individuo no sólo reporta una preferencia, sino también sobre la fuerza de dicha preferencia. Por ejemplo, si las alternativas i y j están disponibles, el individuo puede reportar (Michel Bierlaire, 2009):



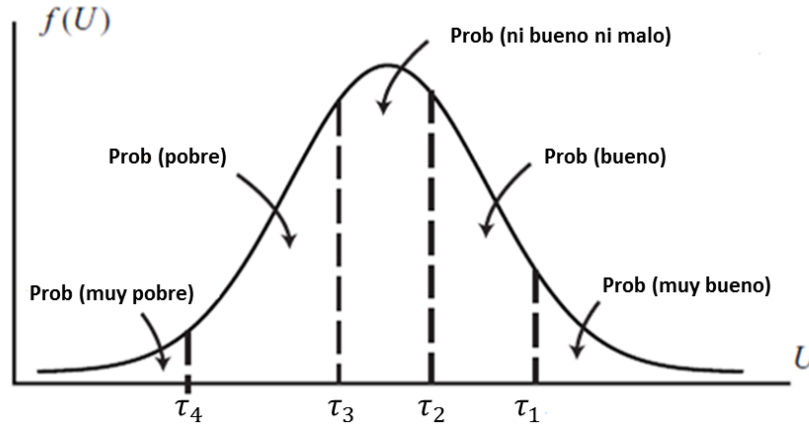
Como es de notar, estas respuestas se encuentran ordenadas, lo cual es una característica principal de este tipo de modelos. Un modelo logit estándar podría

especificarse considerando cada potencial respuesta como una alternativa. Sin embargo, se violaría el supuesto de los modelos logit sobre la independencia de los errores al tener alternativas ordenadas, pues una alternativa es muy similar a las “cercanas” a ella, pero lo es menos con las más “lejanas”.

En el caso de la pregunta sobre la labor que ha desempeñado el Ministerio de Transporte, asumamos que la persona que responde tiene una opinión de qué tan bueno ha sido el trabajo del mismo. La opinión se representa en una variable no observable que se denota con U , donde altos niveles de U significan que la persona considera que el Ministerio está desempeñando una buena labor, y bajos niveles indican una valoración de la labor más pobre.

Aun cuando la opinión de la persona, U , puede tomar muchos niveles diferentes que representen el agrado o desagrado sobre la labor del Ministerio, la pregunta sólo permite 5 posibles respuestas, de manera que la persona elige la respuesta sobre la base del nivel de su U . Es decir, si U está por encima de cierto límite que denominamos τ_1 , el individuo puede escoger la respuesta “Muy buena labor”. Si U está por debajo de τ_1 pero por encima de otro límite τ_2 , entonces contesta “Buena labor”, y así sucesivamente:

RESPUESTA	CONDICIÓN
Muy buena labor	$U > \tau_1$
Buena labor	$\tau_1 > U > \tau_2$
Ni buena ni mala	$\tau_2 > U > \tau_3$
Pobre	$\tau_3 > U > \tau_4$
Muy Pobre	$\tau_4 > U$



Distribución de opinión sobre la labor del Ministerio de Transporte.

Fuente: (Kennet E. Train, 2009)

Ahora bien, la probabilidad de que un individuo escoja “Muy pobre labor” es la probabilidad de que $\tau_4 > U$. Así, la probabilidad de que el individuo seleccione “Pobre labor” es la probabilidad de que U esté por encima de τ_4 (indicando que la persona no cree que la labor es muy pobre), pero por debajo de τ_3 ; esto es, la diferencia entre las funciones de utilidad de dos alternativas $U_{in} - U_{jn}$, dado que se trata de dos alternativas (esto es, si se está o no de acuerdo con el enunciado de la pregunta).

Las probabilidades en el modelo logit ordenado incorporan la fórmula del logit binario. Esta similitud con el logit ordinal es solamente incidental, pues el modelo logit binario tradicional especifica dos alternativas cada una con su función de utilidad, mientras que el modelo logit ordenado tiene una utilidad con múltiples alternativas para representar el nivel de la misma.

Esta semejanza en fórmula, se deriva del hecho de que si dos variables aleatorias son idénticamente distribuidas en valor extremo, luego sus diferencias siguen una distribución logística. Por lo tanto, asumiendo que ambas utilidades en un logit binario son

IID en valor extremo, esto es equivalente a asumir que la diferencia en las utilidades se distribuye logísticamente, al igual que la utilidad en el modelo logit ordinal.

Así, una categoría q está asociada a un intervalo $[\tau_{q-1}, \tau_q]$, la probabilidad de que esa categoría q sea elegida por el encuestado es (Michel Bierlaire, 2009):

$$\begin{aligned}
 P_n(q) &= \Pr(\tau_{q-1} \leq U_{in} - U_{jn} \leq \tau_q) \\
 &= \Pr(\tau_{q-1} \leq (V_{in} - V_{jn}) - (\varepsilon_{jn} - \varepsilon_{in}) \leq \tau_q) \\
 &= \Pr(V_{in} - V_{jn} - \tau_q \leq \varepsilon_n \leq V_{in} - V_{jn} - \tau_{q-1}) \\
 &= F_{\varepsilon_n}(V_{in} - V_{jn} - \tau_{q-1}) - F_{\varepsilon_n}(V_{in} - V_{jn} - \tau_q)
 \end{aligned}$$

5 Metodología

Con el objetivo de realizar las encuestas origen destino a los estudiantes de la Universidad de la Costa con el fin de poder estimar la demanda de viajes desde y hacia la Universidad en los distintos modos de transporte se siguió la siguiente metodología.

Para la realización de la presente encuesta se tuvieron en cuenta los siguientes aspectos: Antes de iniciar la encuesta se registró el nombre del encuestador en el espacio donde dice: *Nombre del Encuestador*. Luego se llegaba donde el usuario y se le presentó formalmente la justificación del estudio definiendo el objetivo general de la encuesta.

El objetivo General de la Encuesta es el siguiente: *Determinar la demanda de viajes desde y hacia la Universidad de la Costa en los distintos modos de Transporte Urbano para hacer un diagnóstico de la movilidad y proponer alternativas de mejoramiento*. Luego se dió inicio a la encuesta comenzando con la Parte 1, que corresponde a la Información General del encuestado así:

5.1 Parte 1: información general del encuestado

Estas preguntas fueron formuladas de manera verbal anotando las respuestas en el formulario. Se obtuvieron las siguientes variables socioeconómicas: género, tasa de motorización del hogar (auto, moto, bicicleta), tamaño del hogar, número de estudiantes en el hogar, número de trabajadores en el hogar y estrato.

Parte 2: información sobre viajes

Una vez conocida la información socioeconómica del entrevistado, se procedió con preguntas relativas a los viajes que el individuo realizó hacia y desde la universidad durante el día inmediatamente anterior. Se procedió a preguntar por el primer viaje hacia la universidad que se reseñó como viaje número 1, luego por el segundo viaje hacia la universidad que se reseñó como viaje número 2 y así consecutivamente.

De cada viaje se preguntó:

- Hora de inicio (desde el lugar de origen).
- Lugar de inicio del viaje.
- Hora de llegada (a la Universidad)
- Modo empleado.
- Ruta: Si el usuario se vino en bus, indicar la ruta que tomó.

Posteriormente se preguntó sobre los viajes realizados desde la Universidad. La información se registró de manera similar a como se hizo anteriormente.

6 Análisis de datos

Las variables del estudio fueron definidas como se describe en la siguiente tabla:

Tabla 1.

Descripción de variables de estudio

Variable	Definición	Tipo de variable	Observaciones
NVEHI	Número de vehículos en el hogar	Cuantitativa - Continua	Autos y Motos
NPERS	Número de personas en el hogar	Cuantitativa - Continua	Tamaño familiar
TRABAJAN	Número de personas que trabajan en el hogar	Cuantitativa - Continua	
ESTUDIA	Número de personas que estudian en el hogar	Cuantitativa - Continua	
NO_TRABAJA	Número de personas cesantes en el hogar	Cuantitativa - Continua	No trabajan ni estudian
IBAJO	Hogares pertenecientes a estratos 1,2	Muda - Discreta	1: si pertenecen; 0: no pertenecen
IMEDIO	Hogares pertenecientes a estratos 3,4	Muda - Discreta	1: si pertenecen; 0: no pertenecen
IALTO	Hogares pertenecientes a estratos 5,6	Muda - Discreta	1: si pertenecen; 0: no pertenecen
SIN_AUTO	Hogares no poseen vehículo/moto	Muda - Discreta	1: si poseen; 0: no poseen
UN_AUTO	Hogares poseen 1 vehículo/moto	Muda - Discreta	1: si poseen; 0: no poseen
DOS_O_MAS_AUTOS	Hogares poseen 2 o más vehículo/moto	Muda - Discreta	1: si poseen; 0: no poseen
Viajes_Estudio_dia	Número de viajes motivo estudio generados hogar/día	Cuantitativa - Continua	No trabajan ni estudian
N válido (según lista)	267		

Fuente: Elaboración propia

6.1 Descripción de las variables dependientes

A continuación, se presenta el análisis estadístico de cada una de las variables dependientes en el estudio, en este caso de los viajes totales en el día para propósitos de estudio.

Tabla 2.

Estadísticos de tendencia central y de dispersión para las variables dependientes

	Mínimo	Máximo	Media	Desv. típ.	Varianza	Asimetría	Curtosis
	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
NVEHI	0	5	0,3933	0,68745	0,473	2,454	9,171
NPERS	1	10	3,9738	1,67109	2,793	0,626	0,53
TRABAJAN	0	5	1,5094	0,96699	0,935	0,891	1,221
ESTUDIA	0	6	1,0787	1,13281	1,283	1,064	1,528
NO_TRABAJA	0	6	1,3858	1,15568	1,336	1,061	1,787
Viajes_Estudio_dia	0	3	1,1386	1,06176	1,127	0,157	-1,464

Fuente: Elaboración de los autores

La primera conclusión que se puede obtener de los datos es que su distribución parece alejarse de una distribución normal ya que tanto su coeficiente de asimetría como su curtosis se alejan considerablemente de 0. Esta situación dificulta la comparación estadística entre medias y desviaciones ya que se podrían estar violando los supuestos requeridos para las pruebas de hipótesis en el marco de la estadística paramétrica.

A continuación, se muestran los diagramas de cajas para los tipos de viajes estudio durante el día, con el propósito de tener una idea acerca de la dispersión de los datos, e identificar posibles datos atípicos.

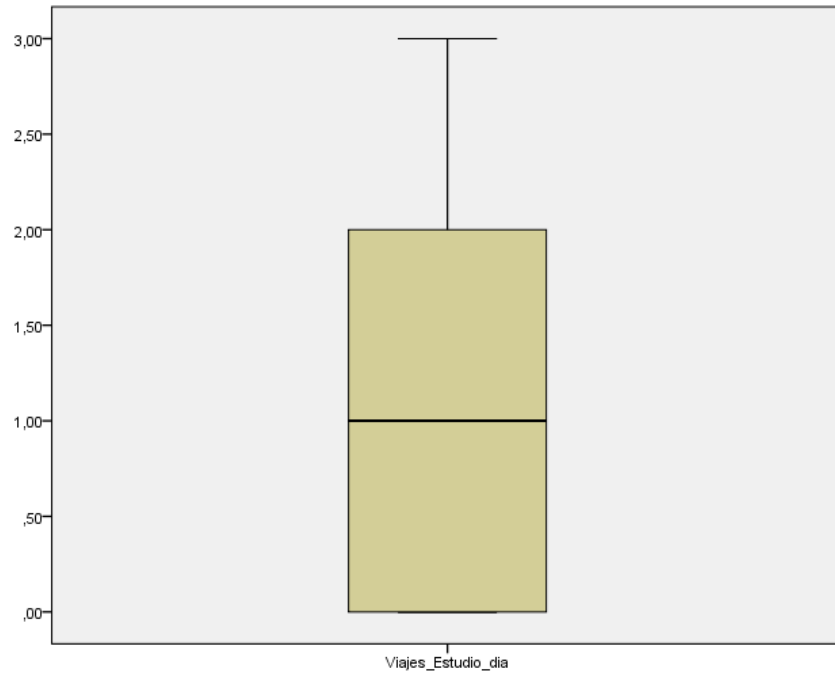


Figura 3. Diagrama de caja para las distribuciones de viaje. Fuente: Elaboración propia

Como se logra apreciar, los datos correspondientes a viajes presentan un sesgo hacia la derecha, indicando que la mayoría de los viajes que se realizan en el hogar oscilan alrededor de valores pequeños, y en algunas pocas ocasiones, se generan hasta 3 viajes por hogar al día con motivo estudio. En general se puede ver que los datos no tienen un comportamiento normal, situación que no permite por ahora la construcción de intervalos de confianza, ni la aplicación de pruebas de hipótesis para evaluar la diferencia o no de los valores medios de las variables. Tampoco es posible hacer comparaciones de varianzas para la validación de supuestos. El problema de normalidad en las observaciones se hace evidente debido a la importante presencia de ceros en la base de datos y al poco número de observaciones que se tiene para los hogares que registran número impar de viajes.

6.2 Descripción de variables discretas.

A partir de la encuesta origen destino se obtuvieron las siguientes variables discretas que describen los viajes motivo estudios realizados durante un día: nivel de ingreso, tasa de motorización, número de personas por hogar y número de personas que estudian o trabajan por hogar.

Nivel de ingreso

En base a lo observado histograma de datos del nivel de ingreso, se las siguientes categorías.

Tabla 3.

Estadísticos estratos en la muestra

Estrato Socioeconómico	Frecuencia	Porcentaje
<i>Ingreso Bajo</i>	190	71.2%
<i>Ingreso Medio</i>	66	24.7%
<i>Ingreso Alto</i>	11	4.1%
Total	267	100.0%

Fuente Tasa de motorización. Elaboracion propia del autor

Esta variable se agrupo de acuerdo al número de hogares pertenecientes a cada una de las categorías que se describen a continuacion:

Tabla 4.

Estadísticos Tasa Motorización

Tasa de Motorización	Frecuencia	Porcentaje
<i>Sin Auto/Motocicleta</i>	183	69%
<i>Un (1) Auto/Motocicleta</i>	70	26%
<i>Dos (2) o más Auto/Motocicleta</i>	14	5%
Total	267	100%

Fuente: Elaboración manual del autor

Como se observa la mayoría de hogares no poseen vehículo, seguido de los que poseen un vehículo. De resto son muy pocos los que poseen dos o más vehículos por lo que se agruparon estos últimos en una categoría:

6.2.1 Ocupación de las personas en el hogar

Dado que respecto a la ocupación de personas se tenían alrededor de 12 variables, se decidió agruparlas en distintas categorías de manera que grupos de ellas tuvieran características comunes a fin de simplificar el análisis y hacerlo más práctico.

En consecuencia, se conformaron las siguientes categorías de variables:

- 1. Personas que trabajan:** este grupo hace referencia a las personas en el hogar cuyos viajes son en su mayoría con propósito trabajo, y a él pertenecen empresarios, trabajadores independientes, empleado, obreros, servicio doméstico, trabajo inestable fuerzas armadas.
- 2. Estudiantes:** este grupo hace referencia a las personas en el hogar cuyos viajes son en su mayoría con propósito estudio. A él pertenecen solamente los que se encuentran tabulados como estudiantes.
- 3. No trabajan:** en esta categoría se agruparon los cesantes, jubilados, no trabaja y dueñas de casa.

Tabla 5.

Estadísticos Ocupación Hogar

Número de personas en el Hogar	Promedio	Porcentaje
<i>Trabajan</i>	1.51	38%
<i>No Trabajan</i>	1.39	35%
<i>Estudian</i>	1.08	27%
Total	3.97	73%

Fuente: Elaboración propia

Los promedios obtenidos indican que el hogar típico corresponde a uno que tiene entre 3 y 4 integrantes, entre 0 y 1 autos, y un ingreso que se ubica en la categoría 1, es decir estratos socioeconómicos 1 y 2.

7 Calibración y selección de modelos

A continuación, se describen los procesos y resultados de la estimación de tasas por los métodos de Regresión Lineal Múltiple (RLM), Análisis por Categorías (AC), Análisis de Clasificación Múltiple (ACM) y Logit Ordinal (LO) asociadas a modelos de generación de viajes motivo estudio.

La metodología usada para calibrar dichos modelos en el caso de RLM fue realizar procedimientos para el análisis de las correlaciones entre las variables, análisis del valor del intercepto y los signos de las variables. Para el caso de los métodos AC y ACM se realizó la categorización de las variables (Ingreso, Tamaño del Hogar y Tasa de Motorización) y

estratificación de hogares por dos variables para cada propósito ya que el comportamiento de los usuarios puede ser notablemente distinta para cada motivo de viaje. Para el método Logit, se evaluó el valor de log-verosimilitud, el R^2 ajustado y la consistencia de signos de los parámetros estimados.

Una vez determinadas las tasas, se procedió a evaluar cada modelo con el fin de determinar cuál modelo presenta mejor ajuste y medida de error. Para dicho fin evaluamos los modelos con la siguiente medida de desempeño:

$$RMSR = \sqrt{\frac{1}{n} \sum (V_{obs} - V_{mod})^2}$$

Donde n es el número de datos, V_{obs} son los viajes observados y V_{mod} son los viajes obtenidos del modelo.

El RMSR llamado error cuadrático medio, indica cuánto es el error en promedio que se está cometiendo el modelo en la estimación de los viajes respecto a los observados en la muestra de validación. En el caso de los modelos por Análisis de categorías y Clasificación múltiple se utilizó el mismo tipo de error para una muestra:

$$RMSE = \sqrt{\frac{\sum (V_{mod} - V_{obs})^2}{No\ Cat - 1}}$$

Con el objetivo de comparar los modelos estimamos el error estandarizado definido así:

$$NRMSE = \frac{RMSE}{\frac{\sum V_{obs}}{No\ Cat}}$$

Donde No Cat es el número de categorías definidas en los métodos por categorías o el número de observaciones en el caso de Regresión Lineal y Logit ordinal.

7.1 Regresión Múltiple

Para explicar los modelos con viajes motivo estudio en primer lugar se realizó un análisis de correlación con el objetivo de observar cuales variables explicaban de mejor forma este tipo de viajes, como se observa en la imagen:

Tabla 6.

Matriz de Correlación de variables cuantitativas para el modelo RLM con propósito Estudio

Correlaciones												
	NVEHI	NPERS	TRABAJAN	ESTUDIA	NO_TRABAJA	IBAJO	IMEDIO	IALTO	SIN_AUTO	UN_AUTO	DOS_O_MAS_AUTOS	Viajes_Estudio_dia
NVEHI	1											
NPERS	,133'	1										
TRABAJAN	,218**	,515**	1									
ESTUDIA	,100	,601**	,121'	1								
NO_TRABAJA	-.088	,426**	-.210**	-.213**	1							
IBAJO	-.503**	-.010	-.144*	-.080	,184**	1						
IMEDIO	,304**	-.001	,093	,075	-.154**	-.900**	1					
IALTO	,485**	,026	,125'	,019	-.086	-.326**	-.119	1				
SIN_AUTO	-.846**	-.030	-.144*	-.017	,094	,548**	-.453**	-.265**	1			
UN_AUTO	,527**	-.057	,038	-.041	-.074	-.429**	,468**	-.038	-.880**	1		
DOS_O_MAS_AUTOS	,722**	,175**	,224**	,117	-.050	-.295**	,021	,628**	-.347**	-.140'	1	
Viajes_Estudio_dia	,147'	,445**	,125'	,635**	-.084	-.104	,097	,026	-.109	,067	,096	1

*. La correlación es significativa al nivel 0,05 (bilateral).
 **. La correlación es significativa al nivel 0,01 (bilateral).

La matriz ilustra que las variables más correlacionadas con los viajes motivo estudio en el día son el número de estudiantes (con un coeficiente de correlación del 63.5%) y el número de personas (con un coeficiente de correlación del 44.5%) a nivel de hogar. Las variables anteriormente citadas están correlacionadas entre si 60.1%, por lo tanto, no deberían estar juntas en el modelo como variables explicativas.

Con respecto a las variables mudas el nivel ingresos (ingreso alto, medio y bajo) posee una correlación positiva, a excepción de los estratos nivel bajo. La tasa de motorización del hogar (sin auto) tiene una correlación negativa, las categorías restantes (un auto y dos o más autos) e positiva, lo cual tiene sentido a nivel de interpretación econométrica. La

significancia en el modelo como variables explicativas se evaluará a partir del *p-value* (*t estadístico*) calibrando el modelo paso a paso:

- Modelo 1: El modelo presenta un R^2 de 0.416. Las variables correlacionadas son el número de vehículos, el número de personas, número de estudiantes en el hogar, y las variables relacionadas al ingreso bajo y medio. Como puede observarse las variables asociadas el estrato o nivel de ingresos tiene un valor t que no es significativo (<1.96) por lo cual deben ser extraídas del modelo de regresión. Además, la alta correlación mostrada entre las variables asociadas al número de personas y estudiantes podrían generar ruido y efectos sobre los coeficientes.

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,645 ^a	,416	,405	,81880

Tabla 7.

Modelo 1 RLM.

Coefficientes^a

Modelo 1	Coefficients no estandarizados		Coefficients tipificados	t	Sig.
	B	Error típ.	Beta		
(Constante)	,150	,324		,461	,645
NVEHI	,123	,093	,080	1,322	,187
NPERS	,059	,038	,093	1,552	,122
ESTUDIA	,534	,056	,569	9,581	,000
IBAJO	,119	,305	,051	,390	,697
IMEDIO	,186	,290	,076	,642	,521

- Modelo 2: El modelo presenta un R^2 de 0.415. Las variables correlacionadas son el número de vehículos, el número de personas, número de estudiantes en el hogar. Como puede observarse las variables tiene un valor t significativo por lo cual extraer las variables anteriormente citada mejoró la capacidad predictiva del modelo. Sin embargo es conveniente estudiar la correlación entre ambas variables y estudiar la variable tasa de motorización especificadas como mudas y capturar efectos de no linealidad.

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
2	,644	,415	,408	,81665

Tabla 8.

Modelo 2 RLM.

Coefficientes^a

Modelo 2	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
(Constante)	,284	,133		2,133	,034
NVEHI	,119	,074	,077	1,620	,106
NPERS	,058	,038	,091	1,528	,128
ESTUDIA	,537	,055	,573	9,704	,000

- Modelo 3: El modelo presenta un R² de 0.41. Las variables correlacionadas son el número de estudiantes en el hogar, y las variables relacionadas a la tasa de motorización especificada como variables mudas para capturar no linealidad. Como puede observarse las variables tienen un valor t significativo. Este modelo se selecciona como el mejor para modelar la generación de viajes con motivo estudio.

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
3	,647 ^a	,419	,410	,81538

Tabla 9.

Modelo 2 RLM.

Coefficientes^a

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
(Constante)	,234	,137		1,705	,089
3 ESTUDIA	,540	,055	,576	9,773	,000
UN_AUTO	,240	,115	,100	2,097	,037
DOS_O_MAS_AU TOS	,119	,229	,025	,517	,606

Para el análisis de los residuos para ver si cumplen homocedasticidad y normalidad se analizan los gráficos de normalidad de tipo Q-Q (cuantiles):

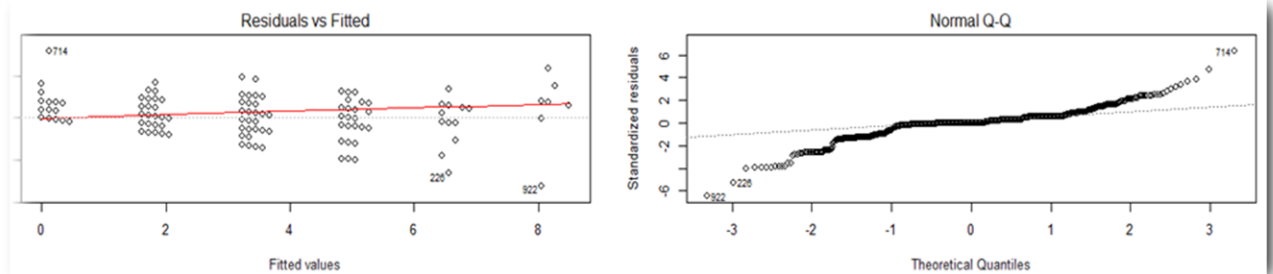


Figura 4. Gráfico de Normalidad en los residuos del modelo RLM para la generación de viajes con propósito Estudio. Fuente: Elaboración propia

Como se observa los residuos si tienen media 0 sin embargo la gráfica normal Q-Q no evidencia normalidad en los mismos en un amplio rango de cuantiles. Se puede decir que los modelos de regresión no se ajustan bien para esta tipología de viajes dado que las variables en las encuestas no logran una buena correlación. Este modelo además de

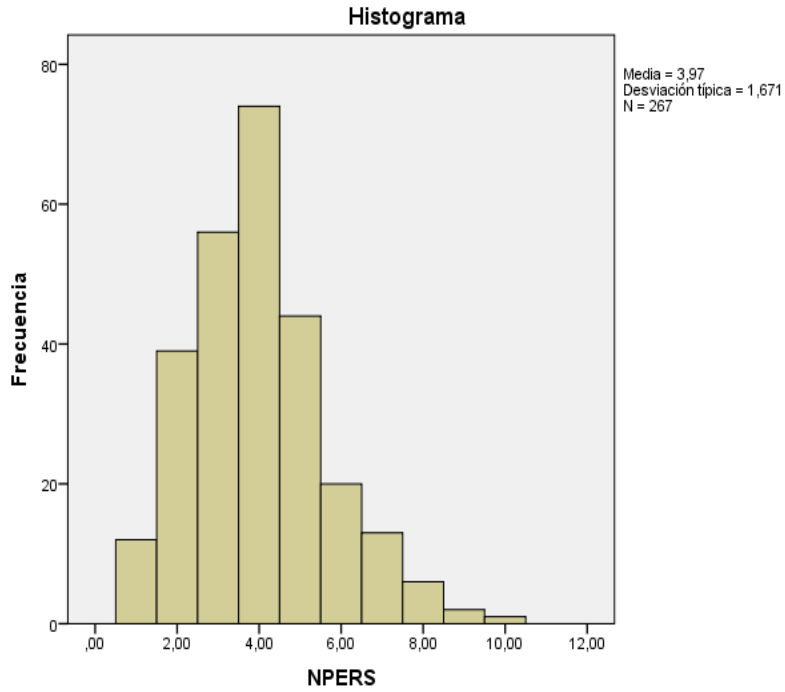
presentar un R^2 pobre, el intercepto posee un coeficiente muy grande respecto al de las otras variables, el cual debería ser 0.

7.2 Análisis por categorías (AC)

El número de categorías de demanda será determinante para definir las dimensiones del modelo, sus requerimientos de información y calibración y finalmente, la precisión de sus resultados. Según la literatura, las variables ingreso, posesión de automóvil y tamaño del hogar son relevantes para efectos de la generación de viajes en los modelos por categorías.

Para cada uno de estos factores mencionados se analizó el histograma de frecuencias en aras de poder agrupar las diferentes categorías en niveles que permitieran disminuir el número total de categorías. Para ello fue necesario analizar el comportamiento de los datos iniciales mediante histogramas de frecuencia con el fin de agrupar datos de cada una de las variables y, así, conformar grupos homogéneos evitando categorías con pocas observaciones.

Así por ejemplo, para el primer factor “*Tamaño del Hogar*”, se muestra el histograma de frecuencias y los niveles por categorías establecidos:



Para el factor “Nivel de Ingreso” se establecieron 3 niveles, a saber: ingreso bajo, medio y alto, agrupando las siguientes categorías iniciales:

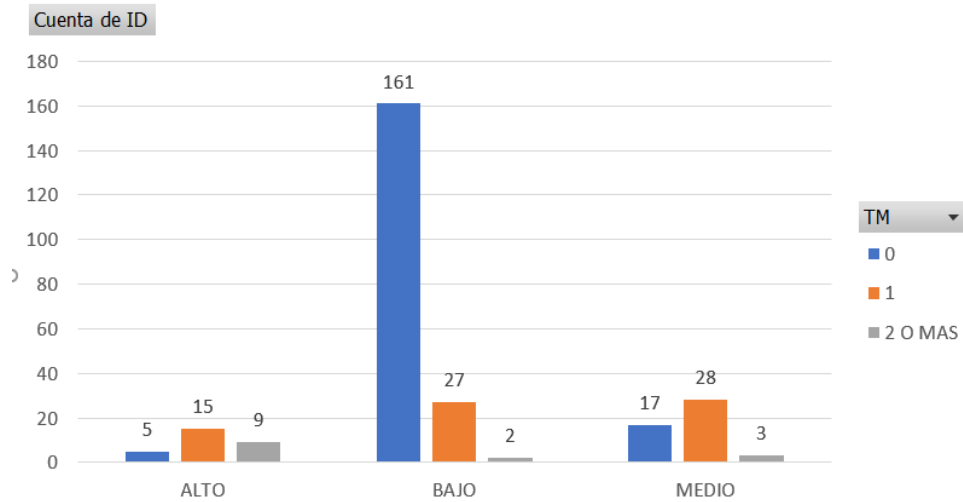


Figura 5. Histograma para el Nivel de Ingreso.

Fuente: Elaboración propia.

Para el factor “Tasa de Motorización” los niveles se establecieron de la siguiente manera: hogares sin posesión de vehículos, hogares con posesión de un vehículo y hogares con posesión de dos o más vehículos.

Teniendo en cuenta los niveles establecidos para cada factor (3 para cada uno), se realizaron modelos para diferentes pares de factores, es decir, cruzando por ejemplo Nivel del Hogar – Posesión de Vehículo; Nivel de Hogar – Nivel de Ingreso; Nivel de Ingreso – Posesión de Vehículos; y bien, considerando los tres factores a la vez. Como resultado, el mejor modelo calibrado fue el obtenido del cruce de los factores Nivel de Ingreso – Posesión de Vehículos, cuya metodología se mostrará a continuación. Para el caso de los otros modelos desarrollados, la metodología es análoga.

Así, teniendo en cuenta los niveles establecidos para cada factor, se tiene un total 9 categorías, para cada una de las cuales debe estimarse las tasas de generación según el tipo de viaje. Para ello se determina primero el número de hogares que se tienen en la muestra para cada categoría, y posteriormente el número de viajes en cada una.

Tabla 10
No. de Hogares por categoría

Número de Hogares				
Nivel De Ingresos	Tasa de Motorización			Total general
	0	1	2 O MAS	
1-BAJO	161	27	2	190
2-MEDIO	17	28	3	48
3-ALTO	5	15	9	29
Total general	183	70	14	267

Viajes con motivo Estudio

Tabla 11.

Número de Viajes por categoría con motivo Estudio

Viajes observados Estudio			
Nivel de Ingresos	Tasa de Motorización		
	0	1	2 O MAS
1-BAJO	282	44	8
2-MEDIO	46	68	10
3-ALTO	8	25	30

Fuente: Elaboración manual del autor

Con esta información es posible calcular las tasas de generación por cada categoría, para cada motivo de viaje.

Tasas de generación para viajes con motivo Estudio

Tabla 12.

Tasas de generación por categorías para viajes con motivo Estudio

Tasas Estudio			
Nivel Ingreso/TM	0	1	2 O MAS
1-BAJO	1,69	2,26	3,00
2-MEDIO	1,84	1,88	1,73
3-ALTO	1,45	1,92	2,32

Fuente: Elaboración manual del autor

Las tasas de generación de viajes estimadas anteriormente fueron empleadas para validar este método empleando una base de datos con datos diferentes, de manera que de la misma se extrajo información sobre el total de hogares por categorías. Así, al estimar los viajes generados por categoría para cada motivo de viaje, se compararon con los viajes reales realizados.

7.3 Modelo Logit Ordinal (LO)

Dado que se desea modelar el total de viajes realizados en el día por cada motivo de viaje, el primer paso a realizar es analizar la distribución de viajes del conjunto de datos para cada propósito.

Como se calibrarán modelos tipo Logit Ordinal, es de saber que se deben establecer categorías que permitan dar respuesta a la pregunta: ¿Cuántos viajes por hogar se realizan en un día con determinado propósito? En consecuencia, es preciso realizar el análisis anteriormente descrito del histograma de frecuencias de viajes, con el objeto de establecer estas categorías.

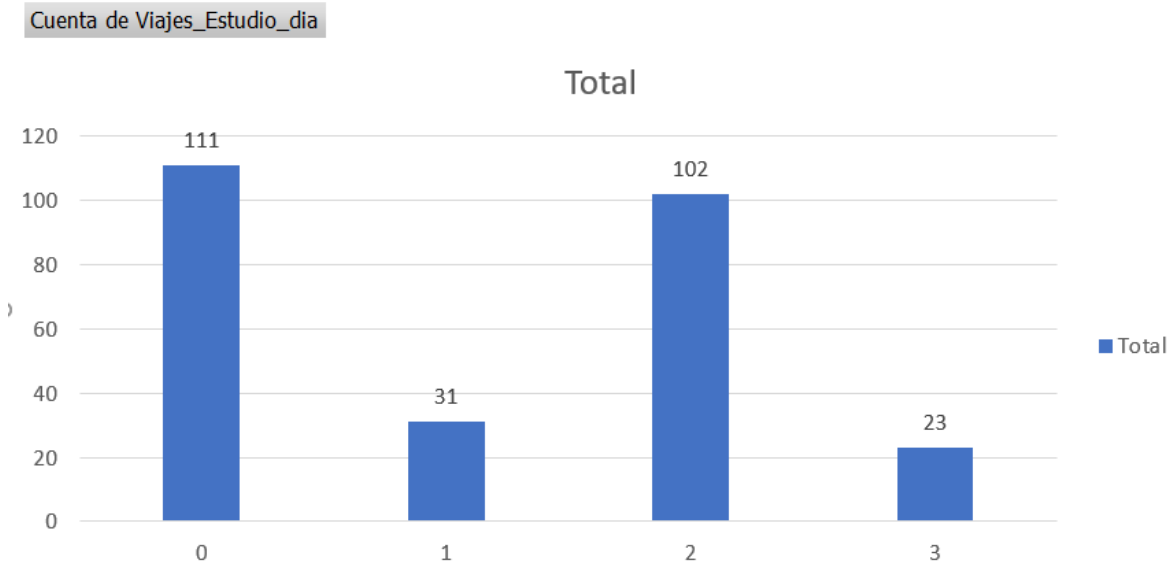


Figura 6. Histograma de frecuencias de viajes en el día con propósito Estudio

Fuente: Elaboración manual del autor

Las categorías propuestas son las siguientes:

Categoría de
Viajes con
propósito Estudio

0

1

2

3

Los mejores modelos encontrados, después de estimar varios de estos analizando diferentes combinaciones de variables explicativas tienen la estructura que se presenta a continuación, de manera conjunta con un resumen estadístico:

Modelos estimados que presentaron consistencia en los signos: (Los resultados se presentan en los anexos):

- Modelo 1

El modelo 1 relaciona las variables asociadas a los estratos, las cuales no resultan ser significativas. El número de personas en el hogar resulta ser altamente significativa.

- Modelo 2

El modelo 2 relaciona las variables asociadas a la tasa de motorización y el número de personas en el hogar, resultan ser altamente significativas. Sin embargo, presenta un log verosimilitud menor al modelo interior.

- Modelo 3

El modelo 3 relaciona las variables asociadas a la tasa de motorización como variables mudas para capturar no linealidad. Los hogares sin auto poseen signo negativo, lo cual tiene consistencia microeconómica. Los hogares sin auto Y el número de personas en el hogar,

resultan ser altamente significativas. Sin embargo, presenta un log verosimilitud menor al modelo interior.

- **Modelo 4**

El modelo relaciona las variables asociadas al número de estudiantes en el hogar y el número de vehículos como variables continuas. El modelo presente significancia estadística en todos los parámetros y un log verosimilitud más positiva que los modelos anteriores.

8 Evaluación de los modelos con fines de predicción

- Modelos de regresión lineal

Para la evaluación de este modelo, se partió de los datos de validación y con estos se obtuvieron las siguientes medidas de desempeño entre el ajuste de los viajes modelados y observados:

Tabla 13.

Errores de los modelos de RLM calibrados para cada propósito de viaje

Viajes	Estudio
RMSR	1,07
NRMSE	0,64

Fuente: Elaboración manual del autor

Se puede decir que los modelos de regresión no se ajustan bien para esta tipología de viajes dado que las variables en las encuestas no logran una buena correlación, además los viajes de este tipo son muy poco en la muestra tomada. Este modelo además de presentar un R^2 pobre, el intercepto posee un coeficiente muy grande respecto al de las otras variables, el cual debería ser 0.

- Modelos por análisis de categorías

Cada modelo fue evaluado, de acuerdo al NRSME, con los datos de la base de validación entregada y los resultados obtenidos fueron los siguientes:

Tabla 14.

Cálculo de los NRMSE para cada relación de categorías de variables y propósito de viaje

Categorías	Estudio
No. Personas – No. Vehículos	0.20462309
Ingreso – No. Personas	0.23792496

De acuerdo a lo anterior, se concluye que el mejor modelo se relaciona con la estratificación No. Personas – No. de Vehículos.

- Modelo Logit Ordinal.

Una vez calibrados los modelos de generación de viajes para cada propósito empleando la modelación logit ordinal anteriormente descrita, se procede a validar con una base de datos ya depurada. Así, se muestran los errores arrojados por cada tipo de modelo calibrado.

Tabla 15.

Errores obtenidos de la validación de los modelos LO generados para cada propósito de viaje.

Medida	Estudio
RMSE	1,71382263
NRMSE	0,87829298

Fuente: Elaboración Propia del autor

9 Conclusiones

Es necesario señalar que los modelos presentados fueron aquellos que exhibieron los mejores indicadores estadísticos de ajuste y además, presentaron valores y tendencias acorde a la percepción de los analistas, por ejemplo, coeficientes de correlación múltiple ajustado adecuados, valores de log-verosimilitud mayores, significancia de variables, menor error cuadrático medio, entre otros. Con el modelo ACM se obtuvieron resultados muy buenos para los casos de propósitos estudio y otros los modelos obtenidos presentan mayor índice de error, pero se explica principalmente porque el número de viajes para dichos propósitos son pocos. Con relación a los valores de las tasas se puede decir que estos cumplen con lo esperado ya que aumentan a medida que se presentan mayor número de personas en el hogar, mayor tasa de motorización e ingresos más altos. Para el caso de los modelos Logit Ordinal, si bien el modelo de generación de viajes arroja mejores resultados considerando la variable “Numero de personas en el hogar que estudian”, cabe mencionar que los estudiantes también pueden realizar viajes con este propósito, sólo que en su mayoría los viajes que realizan tienen como propósito estudio. Esto excluye a las personas que trabajan, quienes también pueden realizar viajes con este propósito, por ejemplo, si realizan una especialización, maestría u otro curso. Estas situaciones pueden influir en los errores del modelo cuando se emplee para fines predictivos. Además, los modelos estimados en el presente trabajo se basaron en los viajes de todo el día, y dado que el comportamiento a lo largo del mismo es diferente para horas punta y las horas fuera de punta, se recomienda mejor realizar calibraciones de modelos para periodos independientes, es decir, periodos punta y fuera de punta para cada uno de los propósitos. De esta manera,

pueden obtenerse mejores ajustes de modelos que a futuro puedan predecir mejor la generación de viajes en las zonas de estudio.

10 Anexos: Resultados modelos logit ordinal

- Modelo 1

```

Ordinal version of the binary logit 00bl.mod
Compare the signs of the coefficients

                                Model: Ordinal Logit
Number of estimated parameters: 6
Number of observations: 267
Number of individuals: 267
Null log-likelihood: -370.141
Cte log-likelihood: -318.720
Init log-likelihood: -364.960
Final log-likelihood: -288.613
Likelihood ratio test: 163.056
Rho-square: 0.220
Adjusted rho-square: 0.204
Final gradient norm: +1.224e-003
Diagnostic: Convergence reached...
Iterations: 14
Run time: 00:00
Variance-covariance: from finite difference hessian
Sample file: BD.txt

Utility parameters
*****
Name      Value  Std err  t-test p-val  Rob. std err  Rob. t-test  Rob. p-val
-----
ASC1      0.00  --fixed--
ASC2      0.00  --fixed--
BETA_IBAJO -0.321 0.560   -0.57 0.57  * 0.480   -0.67   0.50   *
BETA_IMEDIO 0.191 0.589   0.32 0.75  * 0.513   0.37   0.71   *
BETA_NPERS 0.561 0.0800  7.01 0.00  0.0826  6.79   0.00
tau1      1.67  0.630   2.65 0.01  0.570   2.94   0.00
tau2      2.23  0.636   3.51 0.00  0.574   3.89   0.00
tau3      4.75  0.695   6.83 0.00  0.619   7.67   0.00

Utility functions
*****
1      Alt1   one   ASC1 * one + BETA_NPERS * NPERS + BETA_IBAJO * IBAJO + BETA_IMEDIO * IMEDIO
2      Alt2   one   ASC2 * one
    
```

- Modelo 2

```

Number of estimated parameters: 5
Number of observations: 267
Number of individuals: 267
Null log-likelihood: -370.141
Cte log-likelihood: -318.720
Init log-likelihood: -364.960
Final log-likelihood: -288.896
Likelihood ratio test: 162.489
Rho-square: 0.219
Adjusted rho-square: 0.206
Final gradient norm: +2.911e-004
Diagnostic: Convergence reached...
Iterations: 11
Run time: 00:00
Variance-covariance: from finite difference hessian
Sample file: BD.txt

Utility parameters
*****
Name      Value Std err   t-test p-val   Rob. std err Rob. t-test Rob. p-val
-----
ASC1      0.00  --fixed--
ASC2      0.00  --fixed--
BETA_NPERS 0.543 0.0798    6.80  0.00    0.0821    6.62    0.00
BETA_NVEHI 0.297 0.173    1.72  0.09 * 0.164    1.82    0.07 *
tau1      1.89 0.339    5.59  0.00    0.342    5.54    0.00
tau2      2.45 0.351    6.98  0.00    0.347    7.06    0.00
tau3      4.97 0.449   11.07  0.00    0.422   11.76    0.00

Utility functions
*****
1      Alt1      one      ASC1 * one + BETA_NPERS * NPERS + BETA_NVEHI * NVEHI
2      Alt2      one      ASC2 * one
    
```


- Modelo 3

```

Model: Ordinal Logit
Number of estimated parameters: 6
Number of observations: 267
Number of individuals: 267
Null log-likelihood: -370.141
Cte log-likelihood: -318.720
Init log-likelihood: -364.960
Final log-likelihood: -288.640
Likelihood ratio test: 163.001
Rho-square: 0.220
Adjusted rho-square: 0.204
Final gradient norm: +9.039e-004
Diagnostic: Convergence reached...
Iterations: 18
Run time: 00:00
Variance-covariance: from finite difference hessian
Sample file: BD.txt

Utility parameters
*****
Name      Value  Std err  t-test  p-val  Rob. std err  Rob. t-test  Rob. p-val
-----
ASC1      0.00  --fixed--
ASC2      0.00  --fixed--
BETA_NPERS 0.564  0.0817   6.90   0.00   0.0869       6.49       0.00
BETA_SIN_AUTO -0.289 0.506   -0.57  0.57   * 0.434       -0.67       0.50   *
BETA_UN_AUTO 0.223 0.543   0.41  0.68   * 0.486       0.46       0.65   *
tau1      1.72  0.624   2.76  0.01   0.575       2.99       0.00
tau2      2.28  0.630   3.61  0.00   0.584       3.90       0.00
tau3      4.80  0.690   6.95  0.00   0.617       7.78       0.00

Utility functions
*****
1      Alt1   one   ASC1 * one + BETA_NPERS * NPERS + BETA_SIN_AUTO * SIN_AUTO + BETA_UN_AUTO * UN_AUTO
2      Alt2   one   ASC2 * one
    
```

- Modelo 4

```

Model: Ordinal Logit
Number of estimated parameters: 5
Number of observations: 267
Number of individuals: 267
Null log-likelihood: -370.141
Cte log-likelihood: -318.720
Init log-likelihood: -364.960
Final log-likelihood: -255.515
Likelihood ratio test: 229.252
Rho-square: 0.310
Adjusted rho-square: 0.296
Final gradient norm: +1.761e-003
Diagnostic: Convergence reached...
Iterations: 89
Run time: 00:00
Variance-covariance: from finite difference hessian
Sample file: BD.txt

Utility parameters
*****
Name      Value Std err   t-test p-val   Rob. std err Rob. t-test Rob. p-val
-----
ASC1      0.00  --fixed--
ASC2      0.00  --fixed--
BETA_ESTUDIA 1.33  0.142   9.39  0.00   0.189     7.07     0.00
BETA_NVEHI 0.328 0.188   1.75  0.08 * 0.179     1.84     0.07 *
tau1      1.03  0.200   5.16  0.00   0.188     5.50     0.00
tau2      1.74  0.218   8.00  0.00   0.198     8.81     0.00
tau3      4.67  0.366  12.75  0.00   0.341    13.69     0.00

Utility functions
*****
1      Alt1   one   ASC1 * one + BETA_ESTUDIA * ESTUDIA + BETA_NVEHI * NVEHI
2      Alt2   one   ASC2 * one
    
```

11 Referencias

- Alpkokin, P. Historical and critical review of spatial and transport planning in the Netherlands. *Land Use Policy*, Volume 29, Issue 3, July 2012, Pages 536-547
- Chen, B.; Xie, Y.; Tong, W.; Dong, C.; Shi, D.; Wang, B. A comprehensive study of advanced information feedbacks in real-time intelligent traffic systems. *Physica A: Statistical Mechanics and its Applications*, Volume 391, Issue 8, 15 April 2012, Pages 2730-2739, ISSN 0378-4371
- Escobar, D. A. *Instrumentos y metodología de planes de movilidad y transporte en las ciudades medias colombianas*. Tesis doctoral, Universidad Politécnica de Cataluña, Departamento de Infraestructuras del Transporte y del Territorio, Programa de doctorado, “Gestión del territorio e infraestructuras del transporte”. Barcelona: 2008
- Guevara, C. A.; A. Thomas. Multiple Classification Analysis in Trip Production Models. *Transport Policy* 14, (2007): 514-522.
- Hatzopoulou, M.; Miller, E.J. Transport policy evaluation in metropolitan areas: The role of modelling in decision-making. *Transportation Research Part A: Policy and Practice*, Volume 43, Issue 4, May 2009, Pages 323-338.
- Horton N.J.; K.P. Kleinman. Much Ado About Nothing: A Comparison of Missing Data Methods and Software to Fit Incomplete Data Regression Models. *American Statistical Association* 61(1), (2007), 79-90.

Kikuchi, S.; J. Rhee. Adjusting Trip Rate in the Cross-Classification Table by Using the Fuzzy Optimization Method. *Journal of the Transportation Research Board* 1836, (2003): 76–82.

Manheim, M. L. *Fundamentals of Transportation Systems Analysis, Volume 1: Basic Concepts*. MIT Press series in transportation studies, July 1979, Pages 10-57

Montgomery, D. C. *Diseño y análisis de experimentos*. Limusa Wiley, 2a. edición, México, 2013, 686 p.

Mwakalonge, J. L.; Badoe, D. A. Comparison of Alternative Methods for Estimating Household Trip Rates of Cross-Classification Cells. *Journal of the Transportation Research Forum*, Vol. 51, No. 2 (Summer 2012), pp. 5-24.

Nutt, P. Ch. Some guides for the selection of a decision-making strategy. *Technological Forecasting and Social Change*, Volume 19, Issue 2, March 1981, Pages 133-145, ISSN 0040-1625

Ortúzar, J. D. *Modelos de demanda de transporte*, 2a edición, Alfaomega, México (2011).

Ortúzar, J. D.; Willumsen, L. G. *Modelling Transport*, 4th edition, Jhon Wiley & Sons Ltd. Chichester, United Kingdom (2011)

Sheffi, Y. *Urban Transportation Networks: Equilibrium analysis with mathematical programming methods*, Prentice-Hall Inc. New Yersey (1985) p. 2

Stopher, P.; McDonald, KG. Trip generation by cross-classification: an alternative methodology. *Transportation Research Record* 891, 10-17 (1983)

Rengaraju, V.; M. Satyakumar. "Structuring Category Analysis Using Statistical Technique."

Journal of Transportation Engineering 20 (6), (1994): 931-939.

Retherford, R. D.; M.K. Choe. *Statistical Model for Causal Analysis*. John Wiley & Sons,

New York, Inc., NY, 1993.

Wheeler, S.M., Beatley, T. *The Sustainable Urban Development Reader*, Routledge, New

York (2004)