



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA
DIVISIÓN DE ESTUDIOS DE POSGRADO

T E S I S

**MODELOS DE SELECCIÓN DISCRETA
PARA DETERMINAR LA DEMANDA
DE VIAJES URBANOS**

QUE PARA OBTENER EL GRADO DE
MAESTRO EN INGENIERÍA
(TRANSPORTE)

P R E S E N T A:
ING. JOSÉ MARÍA SÁNCHEZ GUZMÁN



DIRECTOR DE TESIS:
DR. RICARDO ACEVES GARCÍA

CD. UNIVERSITARIA

ENERO DE 2006



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas

Tesis Digitales

Restricciones de uso

DERECHOS RESERVADOS ©

PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



AGRADECIMIENTOS.

*MI MÁS SINCERO AGRADECIMIENTO A MI **FAMILIA Y AMIGOS** POR EL APOYO Y MOTIVACIÓN PARA SEGUIR ADELANTE Y CULMINAR ESTE TRABAJO.*

***A MI DIRECTOR DE TESIS: DR. RICARDO ACEVES GARCÍA.**
POR SU DEDICACIÓN, PACIENCIA Y APORTACIÓN DE
CONOCIMIENTOS PARA LA REALIZACIÓN DE ESTE
TRABAJO.*

***A LA UNAM**
POR PERMITIRME SER PARTE DE TAN GRANDE
INSTITUCIÓN, Y EN ESPECIAL AL DEPARTAMENTO DE
ESTUDIOS DE POSGRADO DE LA FACULTAD DE
INGENIERÍA, QUE FUE LA PARTE IMPORTANTE EN MI
FORMACIÓN PARA OBTENER MAYOR CONOCIMIENTO Y
PREPARACIÓN COMO PROFESIONISTA.*

*Y UN ESPECIAL AGRADECIMIENTO A **DIOS**, POR DARME LA
OPORTUNIDAD DE VIVIR ESTA VIDA RODEADO DE
PERSONAS QUE ME HACEN CRECER Y SER MEJOR CADA
DÍA.*



RESUMEN.....	2
INTRODUCCIÓN.....	3
Objetivo.....	5
Capítulo 1.....	6
Planeación del transporte.....	6
1.1 Concepto de transporte.....	6
1.2 Características de los problemas de transporte.....	6
1.3 Modelos matemáticos y su función.....	9
1.4 Modelos convencionales de planeación de transporte urbano.....	10
1.5 Deficiencias o limitaciones de los modelos convencionales de planeación de transporte urbano.....	13
Capítulo 2.....	14
Modelos de selección discreta.....	14
2.1 Consideraciones generales.....	14
2.2 Proceso de decisión de un viaje.....	17
2.3 Concepto de utilidad.....	22
2.4 Análisis de oferta.....	25
2.5 Análisis de demanda.....	28
2.6 Teoría económica y conducta del consumidor.....	33
2.7 Base teórica para generar modelos de selección discreta.....	37
2.8 Características generales del modelo Probit y del modelo Logit.....	40
Capítulo 3.....	42
El Modelo Logit.....	42
3.1 Modelo Logit Binario.....	42
3.2 Modelo Logit Multinomial.....	44
3.3 Modelo Logit Jerárquico.....	46
3.4 Modelo Logit Mixto.....	50
3.5 Especificaciones requeridas por el modelo logit.....	54
3.6 Estimación estadística de los parámetros del modelo logit.....	54
Capítulo 4.....	61
Software.....	61
4.1 Software para resolver modelos de planeación del transporte.....	61
4.2 Introducción al sistema e instalación.....	61
4.3 Modelo logit desagregado.....	63
4.4 Modelo logit agregado.....	65
4.4.1 Obtención de parámetros de logit agregado.....	65
4.4.2 Aplicación de parámetros del modelo logit agregado.....	66
4.5 Ejemplo de aplicación del software.....	69
SUMARIO.....	74
COCLUSIONES.....	76
BIBLIOGRAFÍA.....	78



RESUMEN

En el pasado, el concepto de demanda de viajes se enfocaba únicamente a proporcionar un sistema de transporte, su ejecución no tomaba en cuenta las preferencias o la conducta de las personas que usaban algún modo de transporte y se utilizaban para estos fines los modelos llamados de primera generación. Como respuesta a la actual variedad de modos de transporte, después de 1981, surge una nueva forma de determinar la demanda de viajes, donde se comienza a emplear los modelos de segunda generación, utilizados por primera vez por Williams¹, quien consideraba la posibilidad de una selección de viaje como una función de sus características socioeconómicas y la relativa atracción de la opción de transporte, permitiendo así considerar, en la determinación de la demanda, las decisiones individuales desde el punto de vista de “preferencias del viajero”. Esto da lugar a una interesante herramienta para orientar el rumbo de la plantación de transporte, definiendo los objetivos a seguir de manera concreta.

Recientemente, se han realizado esfuerzos para determinar la demanda de viajes en las zonas urbanas y así planear adecuadamente los modos de transporte que representan una mejor opción, para esto se requiere trabajar con modelos matemáticos que se apeguen a la realidad y que consideren el mayor número de variables que tienen influencia al determinar el modo de transporte a utilizar y además apoyarse en herramienta sofisticadas, que permitan ahorrar tiempo en la resolución de dichos modelos matemáticos, los cuales son presentados en este trabajo con la finalidad de proporcionar un enfoque distinto, de los aspectos que llevan a determinar la demanda de viajes urbanos y de esta manera realizar una correcta planeación del transporte urbano.

¹ Williams, H.C.W.L., *On the formation of travel demand models and economic evaluation measures of user benefit*. Environment and Planning 9A(3), 285-344, 1977.



INTRODUCCIÓN

Determinar la demanda de viajes en una zona urbana constituye un esfuerzo altamente prioritario en el gobierno o empresas particulares (en algunos casos), dadas las condiciones que imperan en el sistema de actividad urbana y donde el enfoque debe ser dirigido a tener un sistema equilibrado. Hasta hace poco el concepto de decisiones de viaje del consumidor no se consideraba aplicable en la predicción de demanda ni en la investigación para implementar nuevos modos de transporte. Hoy en día se está comenzando a tomar en cuenta las necesidades del viajero como son: lugar de residencia, lugar de trabajo, lugar donde estudian, estatus familiar y tamaño de la familia.

El desarrollo de una teoría de demanda para una población de consumidores, la cual es una lógica y natural generalización de la teoría tradicional, envuelve la selección entre alternativas discretas, el resultado de todo esto es una teoría consistente de selección del consumidor entre alternativas discretas, las cuales proporcionan una herramienta para el análisis de la conducta del viajero, tratando de proporcionar una base para el estudio de la amplia variedad de decisiones del consumidor que son un factor importante en la determinación de la demanda observada. Dos aspectos adicionales de las decisiones de viaje del consumidor requieren una reexaminación del modelo del consumidor tradicional; 1) el primero es que el viaje representa un principio y un fin para el consumidor con el fin de cubrir actividades básicas. Esto requiere de un análisis de selección de viaje con un modelo de “producción de casa” o “actividad de consumo” de las decisiones del consumidor. 2) el segundo aspecto es la extraordinaria complejidad de las decisiones de viaje: modo, frecuencia, destino, tiempo de viaje, finalidad y propósitos del viaje.

Por tal motivo, en este trabajo se describe una visión pragmática de la determinación de demanda de viajes en las zonas urbanas, en el cual se analizan la forma tradicional y la forma actual de ver esta demanda, y en donde se utilizan los modelos de segunda generación, esto sufre adaptaciones importantes, adicionalmente el valor implícito que guarda el trabajo para la determinación de demanda de viajes y a su vez para la planeación del transporte urbano.

Este trabajo consta de cuatro capítulos, en el capítulo 1 se presenta el concepto de plantación de transporte, hablando de las características mas comunes que se presentan en los problemas de transporte y la ventaja que se tiene al apoyarse en los modelos matemáticos además de presentar los modelos matemáticos convencionales utilizados para determinar la demanda de viajes. En el capítulo 2 se explican los aspectos generales de los modelos matemáticos de segunda generación (Logit y Probit) para determinar la demanda de viajes y se determinara el modelo mas conveniente para trabajar. El capítulo 3 muestra las variaciones que tiene el modelo Logit, la aplicación de este modelo y las consideraciones para determinar la demanda de transporte. Mientras que en el capítulo 4 se presenta una herramienta muy útil para trabajar con los modelos



Logit, hablamos de un software que facilita la resolución de este tipo de problemas. Por ultimo se dará paso a las conclusiones finales y a un sumario.



Objetivo

El objetivo de este trabajo es analizar y desarrollar los modelos que se utilizan para determinar la demanda de transporte urbano, que consideran los conceptos actuales de preferencias o conducta del viajero, denominados de selección discreta, así como aportar un software para facilitar la solución de este tipo de problemas.



Capítulo 1

Planeación del transporte

1.1 Concepto de transporte

De acuerdo con William W¹. “El transporte es el movimiento de personas y mercancías, por los medios que se utilizan para ese fin”. Plantea que algunos autores consideran el transporte de pasajeros como el de mayor importancia, especialmente en áreas urbanas; pero agrega que el transporte de mercancías es quizá el de mayor importancia, para el funcionamiento adecuado y económico de la sociedad, por lo que concluye que ambos deben considerarse esenciales.

El transporte es útil en dos aspectos: utilidad de lugar y utilidad de tiempo, esto en términos económicos significa que se debe contar con las mercancías en el lugar y momento que se necesitan, lo mismo se aplica para el transporte de pasajeros.

En general, el transporte es una actividad que tiene fuerte influencia en las condiciones sociales, económicas, políticas, militares, etc., de un país, por lo que el concepto puede variar de acuerdo al punto de vista que se trate, es decir;

- a) Mecánico: el transporte significa desplazar personas o bienes de un punto del espacio a otro, a través de una trayectoria determinada.
- b) Económico: el transporte significa el desplazamiento de los bienes y servicios indispensables en las actividades de los diferentes sectores económicos de un país, y por lo tanto sirve para medir el grado de desarrollo económico del mismo.
- c) Comercial: el transporte es el medio para obtener mercancías de las fuentes u orígenes de producción, así como el medio para distribuir las a sus compradores.

1.2 Características de los problemas de transporte.

El transporte de pasajeros ha ido evolucionando a través del tiempo, esto debido a las necesidades y demandas de la población, a pesar de la evolución y a nuevas formas de transportación masiva, se siguen teniendo los mismos problemas del pasado (tráfico, contaminación, accidentes, etc.) y siguen incrementándose, ya que muchas de las actividades a realizar (trabajo, compras, recreación, etc.) se encuentran centralizadas, este tipo de problemas se observa de igual forma en países desarrollados, lo cual según M. L. Manheim², los problemas de transporte deben apoyarse en:

¹ W. Hay William, “*Ingeniería de Transporte*”, edit. LIMUSA, México, 2000.

² Manheim L. Marvin, “*Fundamentals of Transportation Systems Analysis*, Volume 1: Basic Concepts.

- El sistema global de transporte de una región debe ser visto como un sistema multimodal simple.
- El estudio del sistema de transporte no puede separarse del análisis del sistema social, económico y político de la región.

Por lo tanto en el estudio del transporte deben considerarse cuatro aspectos, los cuales se muestran en la figura 1:

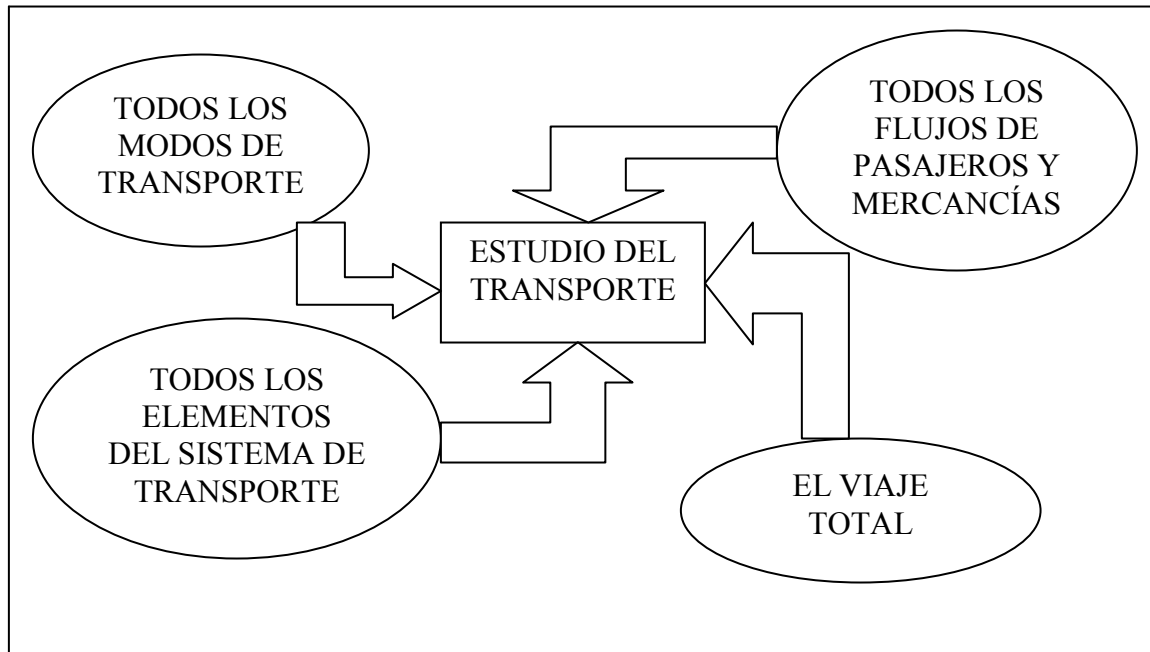


Figura 1³. Elementos a considerarse en el estudio del transporte.

Para complementar la figura anterior se presenta un diagrama que consta de tres variables básicas:

- El sistema de transporte, denotado como **T**.
- El sistema de actividades, señalado como **A** y que representa las actividades sociales y económicas que se realizan en la región.
- La estructura de flujos, indicado como **F** y que son los orígenes, destinos, rutas y volúmenes de personas y cargas que se mueven a través del sistema.

³ Fuente: Manheim, *Fundamental Prospectus of System of Demand Models*, MIT, 1970.

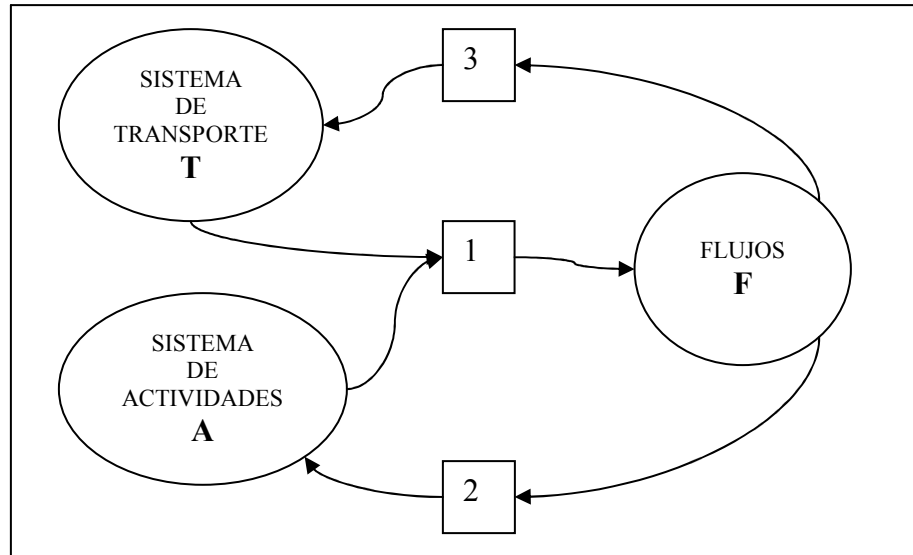


Figura 2⁴: Interrelación entre el transporte, las actividades y los flujos.

En el diagrama se identifican tres clases de relaciones entre las variables:

- La relación 1; indica que los flujos **F** que se presentan en el sistema son el producto de las interrelaciones entre el sistema de transporte **T** y el sistema de actividades **A**.
- La relación 2; señala que los flujos **F** causan cambios en el sistema de actividades **A** en el largo plazo, y por ultimo;
- La relación 3 advierte que los flujos **F** observados en el tiempo generan cambios en el sistema de transporte **T**, obligando a desarrollar nuevos servicios de transporte.

Siguiendo la clasificación de los servicios de transporte de pasajeros y carga se señala que este se lleva a cabo en cuatro grandes sistemas⁵: carretero, ferroviario, aéreo y acuático, donde cada uno de ellos se divide en dos o más modos específicos a saber:

⁴ Fuente: Manheim, *Fundamental Prospectus of System of Demand Models*, MIT, 1970.

⁵ Homburger W. S., Keil J. H., *Fundamentals of Traffic Engineering 20th edition*. University of California, Berkeley, 2000.

MEDIO DE TRANSPORTE	SISTEMA	MODOS DE TRANSPORTE
Por tierra	Carretero	Camión, autobús, automóvil, motocicleta.
	Ferrovioario	Ferrocarril, metro, tren suburbano.
Por aire	Aéreo	Aviación en general.
Por agua	Acuático	Barcos, lanchas, bacazas y canoas

Tabla 1⁶. Clasificación del transporte.

Para la realización del transporte como tal estos grandes sistemas deben contar con una red compuesta de una estructura (terminales, estaciones de servicio, etc.) que le permita llevar a cabo el fin encomendado y para desarrollar dicha estructura donde se puedan soportar todos estos servicios de transporte, es necesario apoyarse en modelos matemáticos para la planeación de transporte y así buscar dar un equilibrio entre la oferta y demanda de transporte.

1.3 Modelos matemáticos y su función.

Un modelo puede ser definido, según Juan de Dios Ortúzar⁷, como: “una representación simplificada del mundo real del sistema de interés, el cual contiene ciertos elementos considerados importantes para el análisis desde un particular punto de vista”. Así de este modo se puede decir que, la función de un modelo matemático es representar un problema dado, con una perspectiva o punto de vista particular y brindar una o más soluciones, siendo útil para hacer previsiones condicionales sobre lo que puede ocurrir con el sistema, si ciertos cambios son introducidos para dicho sistema, aquí hay tres suposiciones básicas en los modelos del sistema de transporte, que tienen un impacto significativo sobre la validación de un modelo.

1. las características clave del sistema pueden ser especificadas o descritas en términos de juegos de observaciones variables.
2. los modelos solo explican un sentido correlativo estadístico.
3. asumir que la estructura funcional o naturaleza del modelo es la misma para todos los individuos y es constante sobre el tiempo.

Dadas esas suposiciones, esto es necesario para describir algunas de las teorías o propósitos conductuales, que algunos modelos de demanda de transporte incorporan y que son encontrados en la práctica hoy en día.

En la actualidad, el número de personas que tienen la necesidad de transportarse a las áreas céntricas de la ciudad, ya sea por trabajo, compras o recreación, es muy alto y esto ha generado que la oferta de transporte se presente de forma desregulada y en consecuencia se de una saturación en ciertos puntos

⁶ Fuente: elaboración propia a través de bibliografía consultada.

⁷ Juan de Dios Ortúzar, Luis G. Willumsen, “*Modelling Transport third edition*”, edit WILEY, 2001.

de la ciudad, designados para terminales de transporte público, por tal motivo es necesario buscar una herramienta útil para que los planeadores de transporte urbano puedan apoyarse, y así reordenar el transporte urbano creando nuevas alternativas y mejorando la ya existentes.

1.4 Modelos convencionales de planeación de transporte urbano.

Los modelos convencionales de demanda de viajes urbanos separan una función que es la de demanda dentro de la generación y atracción de viajes, distribución de viajes, selección modal (selección del modo de transporte) y asignación de rutas, considerando que las observaciones están realizadas en zonas agregadas.

Los modelos convencionales son conocidos como modelos de primera generación o modelos de las cuatro etapas, dichos modelos son clasificados como:

1. Generación y atracción de viajes
2. Distribución de viajes
3. Selección modal y;
4. Asignación de rutas

La estructura básica de estos modelos parte de una forma parecida a la utilizada por Manheim⁸, que en general se evoca a trabajar en una sola parte del sistema dado, conservando las propiedades fundamentales, como se describe a continuación.

Generación de viajes: $N_i = f_1(SE_i)$

El número de viajes partiendo de una zona i es modelado como una función de características socioeconómicas de la zona i . Las variables de transporte no son incorporadas dentro de este modelo, pero por suposición el modelo afirma que la frecuencia de viajes, es totalmente independiente de cambios en el sistema de transporte. El modelo es no-conductual y no-casual, puesto que este no representa las decisiones enfrentadas por personas, tratando de decidir como viajar frecuentemente, como es presentado en la ecuación de la generación de viajes, las variables políticas no son incluidas dentro de los modelos. Esto implica que no cambia para el sistema de transporte las políticas oficiales. En un esfuerzo para rectificar esta deficiencia, algunos estudios tienen incluidos una variable de accesibilidad general⁹ en la ecuación de generación de viajes, porque esta variable mide la accesibilidad general de la zona dada, para alguna otra zona en el área, en lugar de medir el acceso para los destinos que son relevantes para la zona en particular.

⁸ Manheim L. Marvin, *Fundamental prospectus of system of demand models*, MIT, 1970.

⁹ Williams, H.C.W.L., *On the formation of travel demand models and economic evaluation measures of user benefit*. Environment and Planning 9A(3), 1977.

Atracción de viajes: $N_j = f_2(E_j, LU_j)$

En este modelo, el número de viajes terminados en cualquier zona j , esta hecho en función de las características de la zona, tales como empleo o descripción del uso de suelo en dicha zona. Para dejar las variables de transporte fuera de la ecuación, este modelo afirma que la accesibilidad de la zona no tiene efecto sobre el número de viajes. Este modelo, como el primero, no tiene una orientación política, es mecánico en lugar de conductual. Esto proporciona los medios para medir los efectos de las cargas en los estacionamientos, en los centros de la ciudad o un imprevisto en el servicio de tránsito, en el número total de viajes de compras al centro de la ciudad. Esto afirma por suposición, que tales variables no tienen efecto en el número total de viajes de compras al centro de la ciudad.

Distribución: $N_{ij} = f_3(SE_i, E_j, T_{ij})$

Este modelo, también llamado *modelo gravitacional*, distribuye los números preestablecidos de viajes generados por cada zona de origen i a cada zona de destino j . La distribución esta basada en la atracción relativa de la zona j como medida por la ecuación de atracción de viajes y la impedancia de viaje T_{ij} entre i y j . En principio este modelo permite emparejar el origen y el destino para ser sensible al tiempo de viaje a través de la variable de impedancia. Porque de la estructura de los dos modelos anteriores, el proceso de distribución no permite cambiar el número total de viajes partiendo una zona o dirigiéndose a una zona como un resultado de cambios en la ejecución del sistema de transporte, pero solo es posible alterar la localización entre zonas.

En la practica, los factores de impedancia de viaje se basan solo en viajes de auto, entonces estos cambios en el servicio de transporte publico (TP), frecuentemente no tendrán influencia en distribución de viajes predeterminados, el rango completo de tiempo de viaje y factores de costo para todos los modos relevantes de transporte son raramente incluidos en el factor de impedancia, por consiguiente, en la practica, la distribución de viajes es básicamente un procedimiento mecánico de colocación, en el cual la instalación compleja de variables políticas (costos de los diferentes modos de transporte y variables de servicio) son representadas por una medida simplificada de impedancia de auto. De hecho, una consideración más que descontrola la ecuación es que el mecanismo de la impedancia de viaje se basa en una distribución arreglada de longitudes de viaje¹⁰ (en minutos).

El mecanismo de distribución es descriptivo en lugar de causal, y de acuerdo con esto, no predeterminara con precisión el resultado de distribución de viajes, desde las condiciones nuevas de equilibrio que prevalecerían, si el sistema de transporte fuera cambiado.

¹⁰ Como Brand (1972): "la distribución de viajes es modelada como una función de una distribución simple de longitud de viajes el cual prevalecerá en el equilibrio entre oferta y demanda en el archivo de base de datos".

Selección modal:

La función de selección modal probablemente varía más ampliamente en la práctica, que los elementos previos del paquete de plantación de transporte urbano y de las descripciones de los acercamientos convencionales para el modelado de la selección modal urbana¹¹. En los estudios más sofisticados de la plantación de transporte urbano, la selección modal es representada después la distribución de viajes con un modelo de la forma siguiente:

$$\frac{N_{ijauto}}{N_{ijTP}} = f_4(T_{ijauto} T_{ijTP} SE_i, LU_j).$$

Esto es, el número de viajes distribuidos entre pares de zonas, son colocados entre auto y TP, sobre la base de tiempos relativos de viaje y costos entre modos, y en algunos casos, sobre la base de las características socioeconómicas de la zona de origen y las características de uso de suelo de la zona de destino, la diversidad de modos es usualmente ignorado o manejado para combinar todos los modos dentro de los dos modos dichos, TP y auto. Este es el único modelo el cual es fundamentalmente conductual y orientado políticamente, por esta razón esté es adecuado sobre el final del gran proceso mecánico, las variables de política son solo capaces de cambiar la selección entre modos, el número total de viajes y de distribución entre zonas.

Otras limitaciones para estos modelos, los cuales limitan prácticamente la utilidad, es que las técnicas de estimación son muy primitivas, y muy poco pensadas, además son dadas para la forma funcional de el modelo, los modelos frecuentemente incluyen como variables explicatorias, información descriptiva acerca de zonas (tal como densidad residencial), que solo es indirectamente relacionada para los tiempos relativos de viaje y costos, y frecuentemente falta incluir los componentes puerta a puerta del tiempo de viaje y costos.

Asignación de rutas: $N_{ija} = f_5$ (camino de tiempo mínimo, capacidad)

El modelo de asignación de viajes de auto a rutas de vialidades, considera el tiempo mínimo de camino y restricciones de capacidad. Las asignaciones de TP no son modeladas en pequeñas áreas urbanas, este paso es el más cerrado de estos modelos, pero para un proceso de equilibrio no es muy cerrado. El modelo de asignación de rutas, en efecto, genera una medida de impedancia de auto T_{ija} el cual podría ser retroalimentado, dentro de los modelos de distribución de viajes y/o selección modal, para proporcionar algunos aspectos de un proceso de equilibrio. En la práctica no se realiza el efecto de retroalimentación.

Estos modelos asignan viajes de auto para rutas, sobre la suposición de un camino de tiempo mínimo. El proceso de asignación de rutas es un gran procedimiento mecánico, con pequeñas provisiones para la interacción entre la selección de rutas y las otras decisiones de demanda.

¹¹ Son dados por : Fertalet y Weiner, *Modal Split U.S. department of commerce*, Washington, 1966.



1.5 Deficiencias o limitaciones de los modelos convencionales de planeación de transporte urbano.

La descripción hecha, anteriormente, de los modelos convencionales de plantación de transporte de demanda de viajes urbanos puede tener numerosas fallas y limitaciones, en los acercamientos convencionales para el modelado de demanda de viajes urbanos:

- Los modelos son básicamente no-conductuales, reconocen los resultados de las condiciones existentes al tiempo de estudio y guían para conocer los efectos sobre decisiones de viaje, de los cambios en las circunstancias del viajero o en los términos dentro de los cuales ellos ofrecen alternativas en el medio ambiente del transporte.
- Los modelos presentados no son políticamente orientados, excepto la función de selección modal. Los efectos de las variables hechas-políticamente son capaces de controlar, se excluyen de la función de generación y atracción de viajes y son aplicadas muy mecánicamente en la función de distribución de viajes, esto es, esencialmente no interactúan entre ejecución del sistema y las selecciones de frecuencia de viajes o destino de viajes.
- La decisión de viajar de día dependiendo del tiempo, es raramente modelada.
- El equilibrio es esencialmente ignorado, excepto para la extensión limitada que los modelos de asignación de rutas de autos, que toman en cuenta las restricciones de capacidad en asignación de rutas.
- Los modelos son establecidos sobre datos de representación zonal de viajes y condiciones socioeconómicas. Esto oculta mucha de la información en los datos, y juntos con la falta de una estructura conductual, hace los modelos muy difíciles para generalizar de ciudad a ciudad.

La revisión a los puntos anteriores nos indica que, para determinar la demanda de viajes se necesitan modelos más completos, donde se tome en cuenta que el viajero toma una decisión de viaje de una serie finita de opciones, los modelos que nos permiten trabajar con estas opciones son los modelos de selección discreta o modelos de segunda generación, ya que toman en cuenta que el viajero toma la opción que le deja mayor rango de utilidad, tales modelos son:

- Modelo logit y;
- Modelo probit

La explicación de cada uno de estos modelos se tratara en el capítulo siguiente.

Capítulo 2

Modelos de selección discreta

2.1 Consideraciones generales.

Los modelos de demanda agregada de transporte (primera generación), se basan en observaciones para un grupo de viajeros o sobre relaciones promedio del nivel de la zona, por otra parte los modelos de demanda desagregada (segunda generación) son establecidos a partir de observaciones hechas por viajeros individuales y se espera que el uso de éstos, posibilite el desarrollo de modelos más realistas.

Los modelos de primera generación fueron continuamente usados al menos hasta principios de 1980, de hecho fue entonces cuando los modelos de segunda generación empezaron a ser considerados como una seria opción de modelación¹. Los modelos de selección discreta postulan que: *la posibilidad de selecciones individuales a una opción dada, es una función de sus características socioeconómicas y la relativa atracción de la opción.*

Para representar la atraktividad de las alternativas el concepto de utilidad² es utilizado, las alternativas por si mismas, no producen utilidad, esta es derivada de sus características individuales³, por ejemplo, la utilidad observable es definida como una combinación lineal de variables, como se puede ver en la siguiente ecuación:

$$V_{car} = 0.25 - 1.21VT - 2.5ACC - 0.3C/I + 1.1NCAR \quad (1)$$

Donde cada variable representa un atributo de la opción o del viajero, la influencia relativa de cada atributo, en términos de contribución para la completa satisfacción producida por la alternativa, y esta es dada por sus coeficientes, por ejemplo, una unidad de cambio sobre *el tiempo de acceso* (ACC), en la ecuación tiene aproximadamente dos veces el impacto de la unidad de cambio en *el tiempo de viaje en el vehículo* (IVT) y más que siete veces el impacto de una unidad de cambio *costo/ingreso* (C/I). Las variables también pueden representar las características del individuo; por ejemplo, se esperaría que un individuo de un hogar con un gran *número de carros* (NCAR), sería más probable que seleccione la opción del auto que otro perteneciente a una familia de solo un vehículo. La *constante de alternativa específica* 0.25 en la ecuación es interpretada normalmente como una representación de la influencia de toda la red no observada o no incluida explícitamente, por ejemplo, las características del

¹ Williams, H.C.W.L., *On the formation of travel demand models and economic evaluation measures of user benefit*. Environment and Planning 9A(3), 285-344, 1977.

² Thomas A. Domencich, Daniel McFadden, *URBAN TRAVEL DEMAND A behavioral analysis*, edit. North-Holland, 1975.

³ Ortúzar, J de D, *Modelos de Demanda de Transporte*. Ediciones Universidad Católica, Santiago, 1994.

individuo o la opción dentro de la función de utilidad, que podría incluir elementos tales como: comodidad y conveniencia, la cual no es fácil de medir u observar.

En el orden de predecir si una alternativa será seleccionada, de acuerdo con el modelo, el valor de esta utilidad debe ser contrastada con tales alternativas y transformarla dentro de un valor de probabilidad entre 0 y 1, para esto hay una gran variedad de transformaciones matemáticas tales como:

Logit
$$P_1 = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)}$$

Probit
$$P_1 = \int_{-\alpha}^{\alpha} \int_{-\alpha}^{V_1 - V_2 + x_1} \frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1}{\sigma_1}\right)^2 - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} + \left(\frac{x_2}{\sigma_2}\right)^2\right]\right\}}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} dx_2 dx_1$$

Donde la matriz de covarianza de la distribución normal asociada para el último modelo tiene la siguiente forma:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

Los modelos de selección discreta no pueden ser calibrados en general usando técnicas como mínimos cuadrados, porque su variable dependiente P_i es una probabilidad inobservada (entre 0 y 1) y las observaciones son de selecciones individuales (las cuales son 0 ó 1), la única excepción para estos modelos son los grupos homogéneos de individuos, o cuando la conducta de cada individuo es registrada en varias ocasiones, porque las frecuencias observadas de selección son también variables entre 0 y 1.

Algunas de estas propiedades son⁴:

1. Los modelos de demanda desagregada (DM) están basados en teorías de conductas individuales y no constituyen analogías físicas de ningún tipo, por consiguiente, es un esfuerzo hecho para explicar la conducta individual, una importante ventaja sobre los modelos convencionales es que los modelos DM son estables (o transferibles) en tiempo y espacio.
2. Los modelos DM son estimados usando datos individuales y tienen las implicaciones siguientes:
 - Los modelos DM pueden ser más eficientes que los modelos convencionales en términos de la información usada, pocos puntos de

⁴ Ampt, E. y Bonsall P., *Current issues in travel and transport demand surveys*, 1996.

datos son requeridos puesto que cada selección individual es usada como una observación. En la modelación agregada una observación es el promedio de cientos de observaciones individuales.

- Son usados como datos individuales, toda la variabilidad inherente en la información puede ser utilizada.
- Los modelos DM pueden ser aplicados, en principio, a cualquier nivel de agregación, el proceso de agregación no es trivial.
- Los modelos DM son menos probables para sufrir de prejuicios de correlación entre unidades agregadas. Un serio problema de la información agregada es que la conducta individual, puede ser escondida por características asociadas para las zonas, esto es conocido como *correlación ecológica*, la figura 3 muestra que si un modelo de generación de viajes fue estimado usando datos zonales, se obtendría que el número de viajes decrece con el ingreso, lo opuesto sería mostrado para sujetar si los datos fueron considerados a un nivel de hogares. Este fenómeno, el cual por su puesto exagerado en la figura, puede ocurrir por ejemplo si las características del uso de suelo de la zona B es conducente para más viajes a pie.

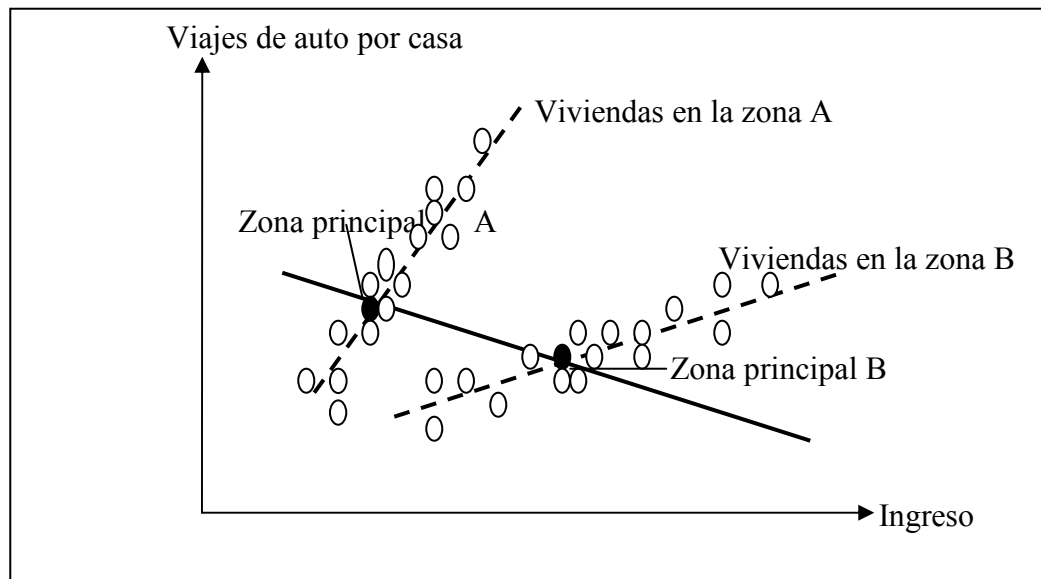


Figura 3⁵.

3. Los modelos desagregados son probabilísticos, además, si ellos aprovechan la probabilidad de cada alternativa y no indica cual se selecciona, el uso debe ser hecho de base en conceptos probabilísticos tales como:

⁵ Fuente: Ortúzar, J de D y Willumsen, *Modelling Transport*, John Wiley & Sons, Chichester, 1990.

- El número esperado de personas que usan una cierta opción de viaje iguales a la suma sobre cada individuo de las probabilidades de seleccionar que alternativa:

$$N_i = \sum_n P_{in}$$

- Un juego independiente pueden ser modelado separadamente considerando cada una como una selección condicional, entonces las probabilidades resultantes pueden ser multiplicadas para el rendimiento de las probabilidades conjuntas para el juego, tales como en:

$$P(f, d, m, r) = P(f)P(d / f)P(m / d, f)P(r / m, d, f)$$

Con f = frecuencia, d = destino, m = modo y r = ruta.

4. Las variables incluidas en el modelo pueden tener coeficientes explícitamente estimados, en principio, la función de utilidad permite cualquier número de especificaciones de variables explicatorias, como opuesto al caso de las funciones de costos generalizados en los modelos convencionales, los cuales son generalmente limitados y tienen varios parámetros arreglados y con las siguientes implicaciones:

- Los modelos DM permiten, para una representación más flexible, las variables políticas consideradas relevantes para el estudio.
- Los coeficientes de las variables explicatorias tienen una interpretación de utilidad marginal directa (reflejan la importancia relativa de cada atributo).

2.2 Proceso de decisión de un viaje

La teoría de la conducta de selección racional, afirma que tomar una decisión puede contar con alternativas posibles en orden de preferencia, y siempre se seleccionará de las alternativas disponibles la cual se considera más deseable, dando sus pruebas y las restricciones relevantes colocadas en la toma de su decisión, tal como su nivel de ingresos o el tiempo disponible. Modificados adecuadamente para tomar en cuenta el fenómeno psicológico de aprendizaje y percepción de errores, esta teoría ha sido usada sucesivamente para analizar y predecir la conducta económica del consumidor en una amplia variedad de aplicaciones, y esto forma la fundación de análisis de economía moderna.

En el modelo de conducta del consumidor que sigue, su teoría es elaborada con el objetivo de la relación entre conducta del consumidor y el transporte, el consumidor se asume que tiene una función de utilidad definida en ambas

concepciones y atributos de transporte. La instalación de alternativas disponibles para el viajero individual, es determinada no solo por las usuales restricciones de presupuestos, sino también por las tecnológicas para llevarlo al trabajo, las actividades de consumo en varias localidades, y los atributos de los modos de transporte para esas localidades. El transporte frecuentemente aparece como una carga arreglada de actividades de consumo e involucra selecciones discretas, la puesta de alternativas disponibles no será una simple instalación de presupuestos del tipo ordinario encontrado en la teoría del consumidor. Consecuentemente, no se obtendrán las implicaciones usuales de la teoría del consumidor sacadas del análisis marginal.

El modelo inicial considerara la selección del consumidor en lo abstracto, sin referencia de especificación para el transporte, el rango que se debe tener en mente son las decisiones hechas por el consumidor en relación con el transporte:

- Localidades de residencia y trabajo
- Disponibilidad de vehículos
- Frecuencia de viajes de trabajo, compras, recreación, y otros viajes
- Destino de viajes
- Hora del día del viaje
- Modo de viaje

Para abarcar esas decisiones, se involucra una selección de corta y larga corrida y la dinámica de consumación de actividades, en general esto es necesario para considerar una completa teoría de la conducta. Para formular una descripción de la economía del consumidor, con el marco conceptual del modelo de Court-Griliches-Becker-Lancaster actividad-consumo-casa-producción. En este modelo se asume que el viajero individual tiene una serie de requerimientos básicos, o conductores, por ejemplo, hambre, sed y descanso.

Sobre su tiempo de vida, el viajero individual tiene disponible un juego **A** de exclusivas alternativas de selección, con cada selección se representa un programa de vida de actividades o actos, cada selección de una actividad de consumo determina los niveles de satisfacción de lo que quiere el viajero individual, cada consumo en la actividad del tiempo de vida, determina un vector de atributos describiendo las compras de las comodidades del mercado, tomando viajes, ejecución del trabajo, etcétera. Las selecciones individuales de una actividad **A** la cual maximiza la utilidad derivada; el vector correspondiente de atributos define su demanda observada. En particular, este vector de atributos especificara la conducta **d** demanda de transporte, a lo largo de las dimensiones mencionadas anteriormente.

Se asume que el viajero individual tiene un tiempo de vida extendida sobre una secuencia finita de cortos periodos, indicando $v = 1, 2, \dots, H$. donde w_v denota un vector finito de niveles de satisfacción de lo requerido en periodo v , $w = (w_1, \dots, w_H)$ denota el vector del tiempo de vida de los niveles de satisfacción requerida.

El proceso de decisión de un viaje tiene que ver con las alternativas disponibles, esto es algo similar a lo que en algunas literaturas se refieren como el “acercamiento de la conducta”, porque en contraste con el acercamiento tradicional, este no esta basado en un modelo descriptivo, pero si en un principio explicito de conducta humana. Específicamente, los viajeros individuales toman en cuenta para realizar un viaje, la alternativa que sea la mejor para ellos.

El proceso de decisión de un viajero individual, junto con varias probabilidades relacionadas, jugaran un rol central en subsecuentes desarrollos y será estructurado de la siguiente manera:

- Dada su localización geográfica i , un tiempo dado en un periodo (hora, día, etcétera), y una actividad (compras, trabajo, recreación, etcétera), una primera decisión individual será si tomar el viaje o no. La probabilidad incondicional de que el viajero realice un viaje en el periodo de tiempo es entonces P_i , por su puesto que el no viaje es $(1 - P_i)$.
- Dada la selección hecha al primer nivel de decisión, y dada su ubicación presente i , el viajero entonces seleccionara una localización j para la conducta de la actividad dada. La probabilidad condicional de esta decisión será $P_{j/i}$.
- Dados los resultados de las primeras dos decisiones, el viajero entonces decidirá cual modo de transporte m utilizar, de las varias alternativas de modos disponibles, entre su localización i y la localización j seleccionada por la conducta de la actividad. Entonces la probabilidad de esta decisión será $P_{m/ij}$.
- Dados los resultados de todas las decisiones anteriores, el viajero finalmente seleccionara una ruta r , entre las disponibles para viajar tan lejos como lo decida. La probabilidad condicional de esta decisión será $P_{r/ijm}$.

El proceso jerárquico puede ser representado por la figura 4:

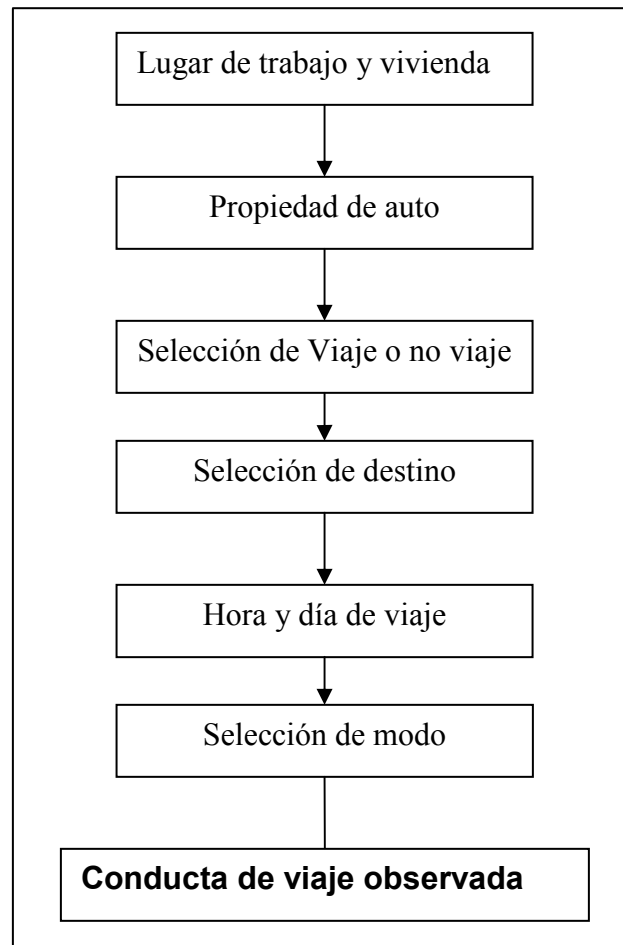


Figura 4⁶.

Muchos puntos deben ser resaltados, primero, el viajero individual en origen de viaje i , no es discernible en termino de probabilidad de su selección de viaje. Esto no implica que todo el viajero individual sea similar, es solo que las probabilidades están sobre el promedio de probabilidades, para viajeros en una zona dada i .

También, para los propósitos de análisis, las selecciones respectivas sobre las que están estructuradas secuencialmente, cada individuo será condicionado sobre esos procedimientos, en realidad lo estarán realizando simultáneamente, esto es, ninguna de esas selecciones es separable de la otra en la mente del viajero, de hecho, el orden de todo esto es arbitrario. Por otra parte, para los residentes que no tienen auto, o no pueden conducir, la selección del modo puede ser superior y podría tomar precedencia sobre las otras selecciones, en tales casos, esto puede ser más lógico para ordenar la selección del modo antes que el destino.

⁶ Fuente: Fleet y Robertson, *Trip generation in the transportation planning process*, 1968.

La relación entre las probabilidades del viajero individual de seleccionar cada uno de los respectivos niveles, y las respectivas demandas de viajes T_i , T_{ij} , T_{ijm} , T_{ijmr} definidos anteriormente, en cada zona de origen i , es un número de viajeros potenciales N_i (por ejemplo los residentes). Como el proceso de predicción de demanda puede iniciar algo, este número es dado virtualmente, si no, tendría que ser estimado por una instancia de los modelos de localización residencial o empleo, para nuestro propósito será externo a los modelos de demanda de viajes⁷.

Desde que se asume que el viajero individual en una zona dada i , es similar en términos de sus probabilidades de selección, la proporción de esos viajeros individuales que decidirán viajar durante el periodo dado de tiempo de análisis, es igual a la probabilidad P_i , así las demanda para el viaje originado en la zona de demanda i , en número de viajes o viajeros,⁸ es igual a:

$$T_i = N_i P_{t/i}; \forall i \quad (2)$$

Dentro de esos viajeros, la proporción de éstos que seleccionará zona/localización j como su destino, es igual a $P_{j/i}$. consecuentemente, la demanda para el viaje del origen i al destino j , otra vez medido en número de viajes o viajeros, es igual a:

$$T_{ij} = T_i P_{j/i} = N_i P_{t/i} P_{j/i}; \forall i, j \quad (3)$$

La proporción de los viajeros T_{ij} que seleccionan un modo m como su modo de transporte es igual a $P_{m/ij}$. Consecuentemente a la demanda para el viaje de la zona i a la zona j en un modo m es:

$$T_{ijm} = T_{ij} P_{m/ij} = N_i P_{t/i} P_{j/i} P_{m/ij}; \forall i, j, m \quad (4)$$

Finalmente la proporción de esos viajeros T_{ijm} que seleccionan una ruta dada r entre i y j en un modo m su itinerario es igual a $P_{r/ijm}$. Consecuentemente, la demanda para el viaje de la zona i a la zona j en un modo m y con ruta r es:

$$T_{ijmr} = T_{ijm} P_{r/ijm} = N_i P_{t/i} P_{m/ij} P_{r/ijm}; \forall i, j, m, r \quad (5)$$

Colocando en un marco general, para el modelado de demanda de viajes, todo lo anterior quiere decir que las respectivas probabilidades $P_{t/i}$, $P_{j/i}$, $P_{m/ij}$, y $P_{r/ijm}$, juegan un rol central. Particularmente, esto es conveniente cuando existe un buen desarrollo de la teoría en microeconomía, a la cual podemos recurrir para evaluar esas probabilidades. Esta teoría ha sido desarrollada para tratar

⁷ Sin embargo, los modelos en los cuales esas variables son determinadas endógenamente, por ejemplo; internamente junto con las variables de viaje, puede ser desarrollado. En tales casos N_i refleja los resultados de una decisión anterior por viajeros individuales para seleccionar i como una localización de origen.

⁸ La decisión involucra un viaje por periodo de tiempo, en particular no se consideran tours.

precisamente las situaciones analíticas como la presentada aquí, en las cuales las selecciones que enfrenta el viajero para cada uno de los cuatro niveles de decisión, como decisión discreta o cualitativa, y puede ser referida con indicadores enteros. Esta teoría es llamada “selección discreta” o teoría de “utilidad al azar”, así estos conceptos representan un prerequisite para desarrollar el modelo de demanda de viajes.⁹

Antes de entrar al concepto de utilidad, se debe resaltar que la definición de las selecciones enfrentadas por un viajero, no puede ser tan obvia, porque los viajeros pueden tener diferentes oportunidades para viajar¹⁰.

2.3 Concepto de utilidad

La utilidad es una oferta dada al viajero individual por una selección de viaje dada, o alternativa (por ejemplo; realizar un viaje, de un destino dado, modo, ruta, o todas sus combinaciones), puede ser tomada de una medida de preferencia de la unión de un viajero hacia una selección en particular, o combinación de selecciones. La utilidad de un modo dado de transporte por un viaje dado puede ser medida por el total de los atributos de los modos de transporte, tales como velocidad, comodidad, seguridad y costo, trasladado dentro de su valor monetario, o el valor para el viajero. La utilidad de un destino de viaje dado puede depender en el número de localizaciones y/o tamaño de estacionamientos, costos promedios de la actividad para ser ejecutados (por ejemplo; precios al menudeo en el caso de viajes de compras, accesibilidad, etcétera).

La forma específica en el cual los diferentes atributos de una alternativa combinada, es dada para definir la utilidad total especificada por la “función de utilidad”, podemos especificar que si la alternativa j (ejemplo; un destino dado) es seleccionada, entonces la utilidad recibida por un viajero individual dado en la zona de demanda i , hacia una selección j , para un viaje, es una función lineal:

$$\tilde{U}_{ij} = b_i - c_{ij} + \sum_k a_j^k x_j^k + \sum_l b_i^l Y_i^l + \sum_m d_{ij}^m Z_{ij}^m; \forall i, j \quad (6)$$

Donde: b_i es el ingreso del viajero, c_{ij} es el costo o precio de la alternativa de i , X_j mide otros atributos, en sus propias unidades (ejemplo; el número de destinos de tiendas), Y_i mide otros atributos del viajero (ejemplo; ingreso) y Z_{ij} caracteriza las diferentes combinaciones viajero/alternativas, incluyendo en particular la distancia o costo de viaje de una zona de origen i a una zona de destino j . La utilidad de una alternativa depende no solo de los atributos de las alternativas, si no también de las características del viajero, porque los dos viajeros puede valorar el mismo juego de atributos de manera diferente. Por otro

⁹ Para un análisis más detallado de los fundamentos teóricos de la teoría de la utilidad del azar y de su papel en microeconomía, ver: Anderson, T.W., *Introduction to Multivariate statical Analysis*, Wiley, New Cork, 1992.

¹⁰ Estos aspectos son desarrollados por: Ben-Akiva, ME y Lerman, *Discrete Choice Analysis*. The MIT Press, Cambridge Mass, 1985.

lado, viajeros con ingresos bajos pueden tomar más en cuenta el costo de viaje, y los viajeros con más altos ingresos podrían considerar más el tiempo de viaje.

La presencia del ingreso del viajero b_i , también como el costo de las alternativas c_{ij} , serán particularmente observadas como cantidades que jugaran importantes roles. La tilde ($\sim$) sobre el símbolo U remarca que la utilidad es una ganancia condicional (o indirecta), así, la utilidad recibida es porque la selección correspondiente ha sido hecha. Esto es, una utilidad “posterior”. También se debe utilizar una utilidad incondicional (directa).

Es importante resaltar que esta especificación de utilidad, la cual es común para todos los viajeros individuales en la zona de origen i , implica que todos los viajeros individuales en una zona dada de demanda i están representados por una “característica” o viajero promedio, esta suposición es muy conveniente cuando se desarrollan modelos de demanda agregada, subsecuentemente esto permite la agregación de viajeros individuales dentro de una zona i geográficamente definida.

Las a_j^k , b_i^l y d_{ij}^m son parámetros cuyos valores pueden ser identificados de observaciones actuales de una muestra de viajeros. La razón para estar interesados en el concepto de utilidad es que es la clave para la predicción de la selección del viajero individual, y consecuentemente, la determinación de demanda de viajes. Específicamente, se asumirá que los viajeros siempre seleccionarán la alternativa de viaje (o combinación de alternativas cuando se realizan muchas selecciones juntas) la cuales ofrecerán la máxima utilidad. Aplicando este principio, tan obvio como pueda parecer, principalmente que las alternativas del viajero las cuales serán seleccionadas por un viajero dado puede ser identificado, si las utilidades condicionales individuales de todas las variadas alternativas puedan ser medidas o por lo menos puedan ser clasificadas, por ejemplo, comparadas numéricamente. Entonces si todas las utilidades tienen diferentes valores numéricos, solamente la alternativa con la más alta utilidad debería ser seleccionada.

Trasladando todo en términos de probabilidad de selección de varias alternativas quiere decir que:

$$\begin{aligned} P_{j/i} &= 1.0; & \text{para } j \text{ tal que } \tilde{U}_{ij} &= \text{Max}_k \{ \tilde{U}_{ij} \} \\ P_{k/i} &= 0; & \text{para } k &\neq j \end{aligned}$$

Donde $P_{j/i}$ es la probabilidad condicional de seleccionar j dado que el viajero esta en una zona i . Si la alternativa m es de más alta utilidad, la probabilidad de selección puede ser especificada como $P_{ij} = 1/m$ para esas alternativas, y 0 para las otras alternativas con menor utilidad. Este resultado quiere decir que si el modelado ha proporcionado la información exacta, en relación con las utilidades del viajero $\tilde{U}_{ij}$, entonces la persona que va a realizar el viaje podrá determinísticamente predecir su decisión.

Esto parece bien hasta que uno se da cuenta de que esto implica, que cada viajero individual en una zona de demanda dada i , sería predecida para hacer la misma selección, esto es claramente irrealista, porque es aparente que cada viajero no esta dentro de una unidad geográfica dada, porque si es pequeña esta unidad utilizará el mismo modo de transporte para ir a trabajar. Desde un punto de vista teórico, el resultado de demanda agregada por las variadas alternativas son discontinuas, esto es, una alternativa dada no atrae a ningún viajero individual hasta que esta ofrece la más alta utilidad, entonces la suposición de que la especificación de utilidad es común para todos los viajeros individuales en el origen de la zona i , es defectuosa y necesita ser revisada.

Existen dos aproximaciones alternativas para esto. La primera, es no agregar a los viajeros dentro de grupos, esto es, realizar la demanda de viajes sobre una base individual, viajero por viajero. Las localizaciones de la demanda i define, en este caso, igualar la ubicación de cada viajero individual (lugar de residencia, dirección, etcétera.), esto es prácticamente infactible en términos prácticos. La otra es, revisar lo relacionado con la formulación de utilidad.

La especificación de la utilidad recibida (condicional) en la fórmula podría ser redefinida y generalizada, por introducir un término aleatorio, el cual representaría la diferencia, entre la especificación de utilidad de un viajero individual y el promedio de los viajeros.¹¹ Especificando la utilidad en la cual el promedio de los viajeros en la zona i , decide cuando seleccionar la alternativa j como:

$$\tilde{V}_{ij} = \bar{U}_{ij} + \xi_{ij}; \forall i, j \quad (7)$$

Donde $\bar{U}_{ij}$ es el promedio de la utilidad de los viajeros, como fue especificado en la formula de utilidad, y el término ξ_{ij} representa la incertidumbre sobre la parte del modelo, en relación con el valor de la función de utilidad para un viajero individual. Este término puede representar lo no observable, o los factores no medibles de utilidad (“fuerza de costumbre”), o errores en las medidas de los factores los cuales hayan sido incluidos.

Si el principal de los términos elegidos ξ_{ij} es el conjunto igual a cero, por conveniencia, el componente de la utilidad $\bar{U}_{ij}$, el cual al mismo tiempo llamado la utilidad “sistemática” o “arreglada”, iguala el valor esperado o principalmente del rango de utilidad $\tilde{V}_{ij}$. Este representa la parte observable de la utilidad individual. El valor del rango de utilidad del término ξ_{ij} puede no ser observado, así entonces el valor de la utilidad total $\tilde{V}_{ij}$ de la alternativa j es desconocida, pero también es elegido, como la suma de una constante $\bar{U}_{ij}$ y un término elegido ξ_{ij} .

¹¹ Este término representa la heterogeneidad en utilidades individuales.

Es claro que el modelo no puede ser más largo en relación con la alternativa que seleccionará el viajero individual, es claro que la predicción de selección por el modelo ahora tomara la forma de las probabilidades para un viajero individual dado que, es observado por el modelador seleccionando una alternativa dada. Esas probabilidades dependerán no solo del valor de $\bar{U}_{ij}$, sino también del rango natural de los términos ξ_{ij} , esto es, las características de la función de distribución de probabilidad conjunta de $\xi_{i1}, \xi_{i2}, \dots, \xi_{ij}$. Finalmente esas probabilidades deben ser consistentes con el mismo principio de maximización de utilidades como fue dicho anteriormente.

2.4 Análisis de oferta

Un sistema de transporte puede ser caracterizado por su ejecución, (tiempos de viaje, frecuencia de servicio, seguridad y rentabilidad) los impactos correspondientes sobre el medio ambiente y los costos utilizados para la construcción, mantenimiento y uso del sistema. Las redes, facilidades y servicios que son parte de este sistema y sus características, son referidas como oferta del transporte.

Desarrollar y manejar la oferta de transporte, es un objetivo principal en un proceso de plantación del transporte y se debe contar con los siguientes procesos:

Análisis de redes a nivel metropolitano para la investigación estratégica:

El nivel de análisis involucra generalmente un examen de redes modales alternativas, hecho típicamente a escala regional o a lo ancho de la zona. Decisiones a este nivel, relacionan cosas como adición de nuevos caminos o facilidades de acceso a la red carretera, adoptando un uso de suelo regional o políticas de impuestos, apuntando a influenciar la conducta del viajero, o añadiendo nuevas tecnologías de sistemas de transporte inteligente, que influirán las selecciones individuales de realizar un viaje. La plantación del periodo de tiempo para esas decisiones, son usualmente medidas anuales. El análisis de oferta en tales casos determina el flujo de equilibrio de los modelos sobre una red modal usando una de las técnicas de asignación. Dado un flujo modal origen-destino, como el determinado por el análisis de demanda y un juego de funciones de ejecución para las uniones comprendidas en la red, los modelos pueden estimar toda la ejecución del sistema de transporte. La complejidad de la forma y tamaño de las redes pueden involucrar miles de nodos y de uniones.

Planeación táctica u operacional.

El análisis de las rutas individuales, uniones o terminales, es un objetivo muy común del análisis de oferta, particularmente por agencias operativas, en esos casos, solo una pequeña "red" es considerada, usualmente consiste de una instalación individual o servicio tal como una ruta de transito, una autopista y sus uniones con el resto del sistema. El periodo de tiempo de planeación es frecuentemente en corto plazo (de uno a tres años). Los tipos de decisiones en

este objetivo del ambiente sobre cambios operacionales, reflejados en las preguntas, ¿Cómo pueden ser mejor acondicionadas las señales de tráfico para reducir los retrasos en la circulación de los vehículos? ó ¿Cómo debe ser modificada una ruta de tránsito para mejorar su ejecución? El rango de la técnica de análisis para la plantación operacional es enfocado para un análisis de una red estratégica e incluye todo desde simples cálculos hasta programas de simulación computacionales. La selección de la técnica para una aplicación dada depende sobre la naturaleza específica de los problemas, los datos disponibles, las capacidades del planeador, el tiempo y presupuesto disponible para el análisis y la exactitud y nivel de detalle de la información requerida para el proceso de decisión de un viaje.

Horarios de servicio de transporte.

Los horarios de los vehículos de transporte público para las rutas y asignación de conductores, son tareas grandes para los operadores de tránsito. El objetivo mayor en ambos casos es minimizar los costos de operación mientras que en una ruta larga es minimizar la flota vehicular, sujeta a las restricciones de los estándares del servicio y a los acuerdos laborales. El periodo de tiempo para dicho proceso de horarios es muy corto, generalmente de 3 a 6 meses. La escala del análisis es la red completa de tránsito o al menos las subredes largas, como pueden ser todas las rutas de servicio proporcionadas por un autobús. La técnica usada para horarios involucra una combinación de procesos manuales y computarizados para encontrar heurísticamente “buenos” horarios¹².

Uno de los componentes más importantes del sistema de transporte para el análisis de la oferta es la infraestructura de transporte. Esta infraestructura incluye las rutas sobre las cuales será su trayecto, las señales que controlan los flujos, terminales y todas las facilidades requeridas para operar y mantener el sistema de transporte, en particular almacenamiento de vehículos, servicio, y facilidades para su mantenimiento. También incluye las rutas y horarios que gobiernan la operación de los servicios y los procedimientos para operar el sistema, el cual incluye todo lo relacionado a regulaciones gubernamentales, como límites de velocidad, requerimientos para licencias, estándares de servicio y las condiciones laborales para el conductor.

En la figura 5 se muestra el rol del análisis de oferta en el proceso de plantación del transporte. El análisis de oferta relaciona cuatro puntos importantes para el proceso.

¹² Para ver más del análisis de horarios de servicio ver Pine et al., 1998

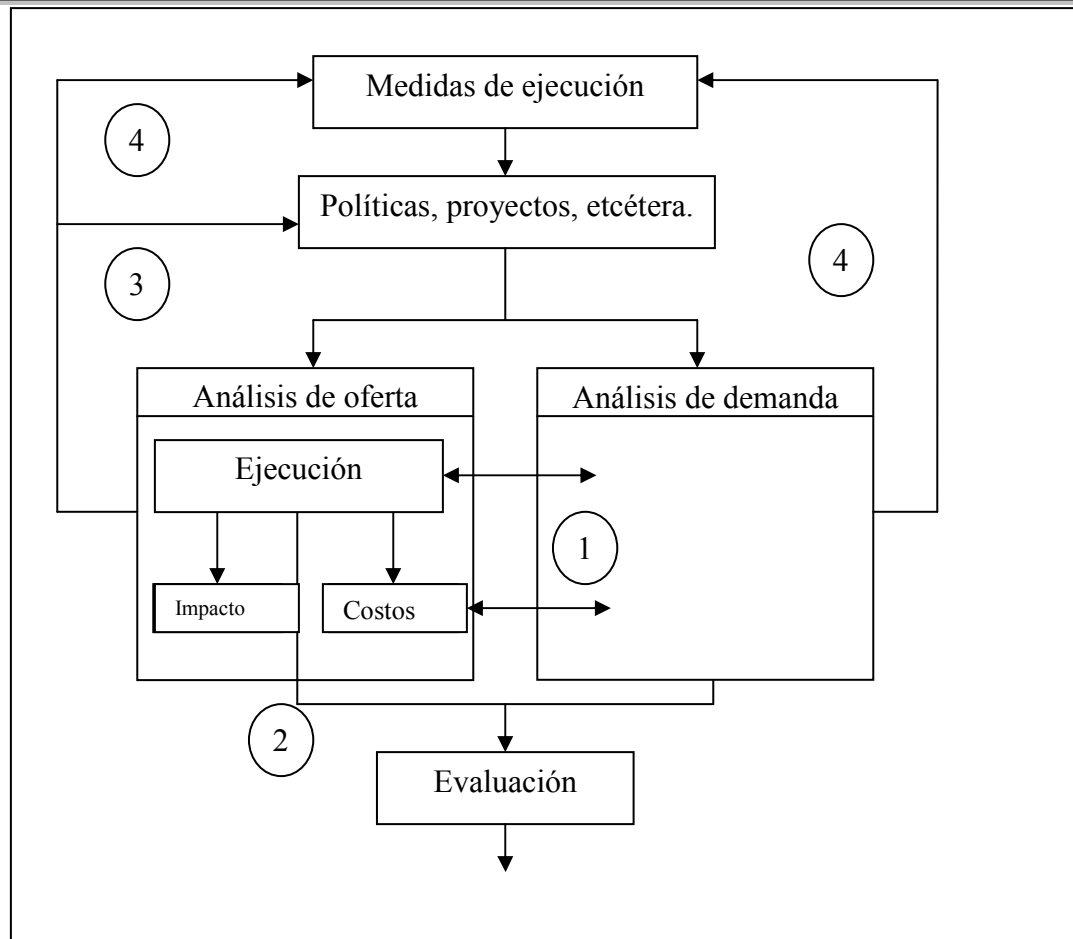


Figura 5¹³.

Unión 1: la ejecución y costos del servicio de transporte son determinaciones importantes de la demanda para esos servicios. La ejecución del sistema de transporte depende sobre el nivel de características de la demanda de viajes. El análisis de demanda y oferta son así inherentemente unidas. En la modelación de la demanda de viajes, esta unión es la más explícita en asignación de viajes, en la cual los flujos de unión (demanda) y tiempos de viaje (ejecución) alcanzan valores de equilibrio después de muchas iteraciones. La ejecución del sistema puede ser incorporada dentro de otros componentes de la demanda, por ejemplo, generación de viajes, distribución de viajes y selección de modo de transporte. En esos casos la ejecución del sistema y costos son utilizados como variables que afectan cuantos viajes ocurrirán y a que lugar están destinados.

Unión 2: sobre la influencia en la demanda para transporte, medidas del sistema de ejecución, impactos y costos constituyen importantes criterios de evaluación en su propia adecuación.

Unión 3: la estimación y previsión del sistema de ejecución, impactos y características de costos conduce a la identificación de proyectos alternativos.

¹³ Fuente: Michael D. Meyer, Eric J. Miller, *Urban Transportation Planning second edition*, Mc Graw Hill, 2001.

Sistemáticamente explorando relación de oferta y medidas de ejecución predichas proveerán a los planeadores un mejor entendimiento de la operación del sistema de transporte y las alternativas posibles que podrían surgir en su operación.

Unión 4: esta unión es relativamente nueva para los planeadores de transporte y refleja la tendencia hacia el monitoreo continuo de la ejecución del sistema de transporte. La retroalimentación del análisis de oferta y demanda es usada para determinar donde se encuentran las diferencias en la red de transporte que conduzcan a la identificación de estrategias para resolver problemas potenciales. Más medidas de ejecución reflejan las características de lado-oferta el sistema de transporte.

La figura anterior también indica el orden para estimar impactos y costos, uno debe primero estimar la ejecución del sistema, por ejemplo, el consumo de energía o niveles de la contaminación del aire en una sección de carretera dependiendo de la velocidad de los vehículos, el número de cambios de velocidad en los ciclos de viaje y el número de vehículos que utilizan la carretera por unidad de tiempo. La operación de los costos que proporciona el servicio de transporte a una frecuencia dada para conocer la demanda cierta dependerá sobre la ruta del tiempo de viaje redondo, y en la vuelta del número de autobuses y conductores necesarios para conocer los horarios.

2.5 Análisis de demanda

La conducta de un estudio de plantación de transporte o examinar los probables impactos de un nuevo sitio en desarrollo, estimar la demanda de viajes esperados para algunos datos futuros, es un punto crítico del departamento encargado de planear el transporte. Investigaciones considerables se han enfocado sobre como describir mejor y predecir la demanda de viajes, en conjunto con la ejecución del sistema, para desarrollar modelos de predicción de demanda, que proporcionen estimaciones reales de viajes futuros sobre redes de transporte. Sin el análisis de demanda, la plantación del transporte no podría ser posible, el desarrollo de predicciones realistas de los viajes futuros es problemático. Puede suceder que la dispersión espacial de viajes de origen, cree significativos retos metodológicos.

El análisis de demanda consiste en muchas tareas que son largamente independientes de las técnicas actuales, usadas para predecir los flujos de viajes. La figura 6 muestra seis tareas básicas que son parte de cualquier análisis de demanda. Las tareas que esto involucra son, de hecho, altamente interrelacionadas. En particular, la definición del problema, selección de la técnica de análisis, colecta de datos y calibración de modelos, son directamente interconectados en la mayoría de los casos. Datos confiables, a menudo determinan la técnica de análisis empleada, de igual forma, el problema del modelo de calibración ayuda a determinar las formas funcionales y especificaciones. De esta forma la figura 6 es así una simplificación de la

naturaleza iterativa de más análisis de demanda e intenta ilustrar algunos de los aspectos claves del proceso de análisis de demanda.

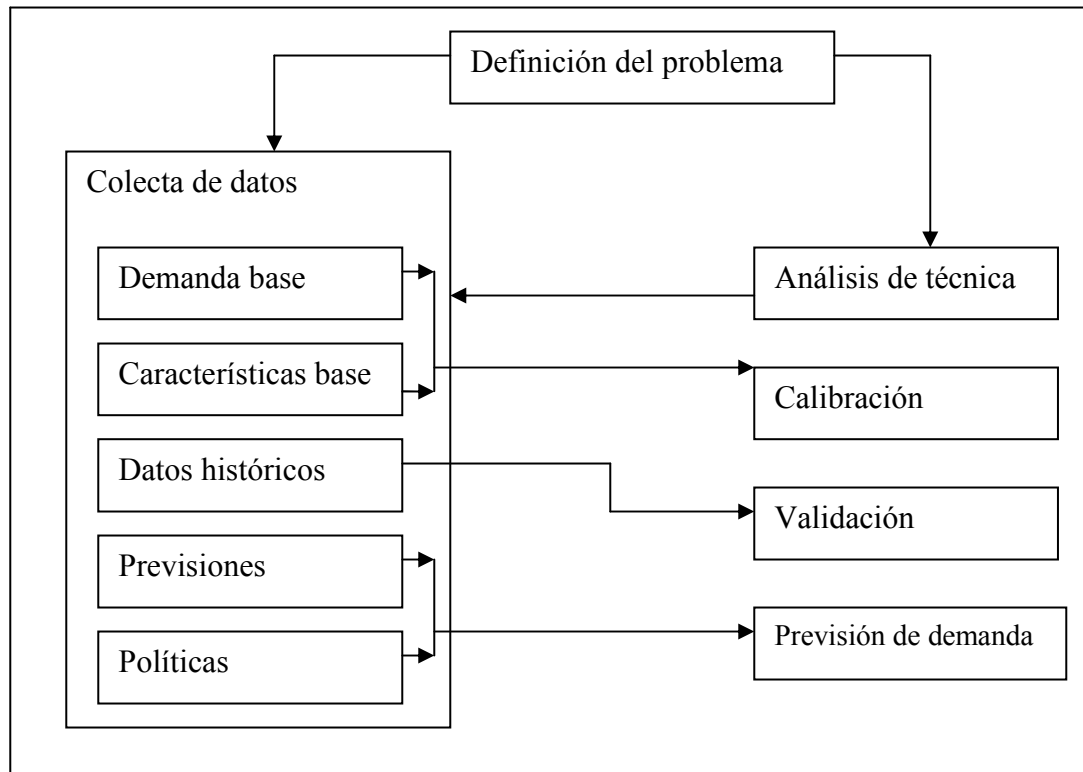


Figura 6¹⁴.

Definición del problema:

La definición del problema es determinado afuera del análisis de demanda, por ser estos problemas usualmente identificados en la fase de diagnostico del proceso de planeación. Tales definiciones de problemas, tienden a ser justamente generales y abstractos en naturaleza, en el orden para ejecutar un análisis de demanda, la definición del problema necesita estar hecho más específico y debe estar definido por lo siguiente:

- *el periodo de planeación específica para el análisis.* El periodo de plantación es típicamente realizado con base anual. Muchos análisis proporcionan una previsión con horizonte a un año. Si anualmente se requieren la previsión para los siguientes 5 años, entonces debe ser hecho explícitamente a la salida, dado que tendría implicaciones para los requerimientos de datos y posiblemente para el análisis de la técnica usada.

¹⁴ Fuente: Michael D. Meyer, Eric J. Miller, *Urban Transportation Planning second edition*, Mc Graw Hill, 2001.

- *el estudio de la región y la estructura de la zona para el análisis.* El límite de la región de estudio, el sistema zonal con la región de estudio, y si los viajes cruzan el límite de estudio para y desde zonas “externas”, son de interés y todo debe ser definido. Esas decisiones dependen obviamente de un contexto de un problema en particular, y típicamente sobre la existencia de datos previamente definidos para la zona (censo de la zona, zonas de tráfico, plantación del distrito, etcétera.)
- *la unidad temporal de demanda.* Son demandas estimadas designadas para el periodo pico, el día típico, un día del fin de semana, y/o del año. La respuesta de esto depende del contexto del problema y disponibilidad de datos.
- *las variables políticas.* Los datos deben estar disponibles para caracterizar las alternativas de interés, en términos de las variables políticas relevantes y la técnica de demanda usada que debe ser capaz de incorporar esas variables. Por ejemplo, un modelo de demanda que no incluya costos de viaje, no será útil para el análisis potencial del impacto de los costos por incremento al precio del combustible.
- *las medidas de efectividad requerida para la evaluación.* Así como las variables políticas o alternativas pueden ser vistas como “las entradas” dentro del análisis de demanda, las salidas es información sobre la demandad de viajes que relacione los impactos de esas alternativas sobre el sistema urbano. Tales medidas están usadas con la fase de evaluación, para asistir la selección entre las alternativas propuestas.

Selección de la técnica de análisis:

La selección de la técnica de análisis será fuertemente influenciada por la definición del problema, el análisis de marco de tiempo, la escala espacial, las variables políticas y las medidas de efectividad. Una topología de “técnicas apropiadas” para varias escalas espaciales y el marco de tiempo, son difícil de construir porque una técnica dada, puede ser típicamente aplicada a varios tipos de escala. Modelos que predicen conductas de viaje individuales, pueden ser empleados virtualmente a cualquier nivel, proporcionado un conveniente procedimiento de agregación. Los modelos agregados, son usados en un amplio rango de actividades de plantación, desde la evaluación de los impactos inmediatos de abrir un nuevo centro de compras, para estimar el flujo de la vialidad en periodo de hora-pico, sobre un horizonte de 20 o 30 años. Dos tendencias parecen existir en este caso, la primera es la escala espacial del incremento del periodo de tiempo y/o análisis, así se tiene la complejidad y la comprensividad de la técnica de análisis requerida, para direccionar adecuadamente el problema de planeación.

La otra tendencia es una escala espacial y el aumento en el periodo de tiempo, los datos, costos y requerimiento de tiempo de análisis, para observar detalladamente las iteraciones, para tener hipotéticamente la relación de conducta y para predecir el futuro como una escala espacial.

Estimación de parámetros:

Virtualmente todos los modelos de transporte consisten de una variable dependiente –la demanda para transporte- expresada como una función de una o más variables independientes o explicatorias. Si D denota la demanda y X denota el vector de variables explicatorias, entonces el modelo de demanda puede ser representado por la siguiente expresión:

$$D = f(X, \theta) \quad (8)$$

Donde θ es un vector de parámetros o coeficientes, que determina la “forma” de la curva de demanda y los “pesos” de los términos de la función de demanda. La elasticidad en un análisis, el exponente de tiempo de viaje en un modelo gravitacional, y los coeficientes de la función de utilidad en un modelo logit, son ejemplos de parámetros en los modelos. Los parámetros de un modelo son supuestos para ser conocidos, en una situación actual, esos parámetros no son conocidos, y deben ser estimados para un problema dado, con aplicación de datos históricos. Este proceso de estimación, conocido como calibración de un modelo, involucra una comparación de niveles de demanda observados con los niveles predeterminados por el modelo, dando valores supuestos a los parámetros. Esos valores de parámetros son entonces ajustados hasta que los niveles de demanda predeterminados, hacen juego con la situación real de la demanda observada, tanto como sea posible.

Los procedimientos de estimación estadística, deben ser usados siempre que sea posible para determinar los valores de los parámetros, debido a la eficiencia y capacidad que proporcionan. Un modelo desarrollado no puede ser estadísticamente estimado, un ejemplo muy común de esto es donde los factores gráficos son usados, en lugar de una función matemática para representar algunas características del sistema, como puede ser la impedancia del viaje.

Validación:

Los parámetros de estimación producen un modelo, que reproduce mejor los datos históricos usados. Antes de usar un modelo de calibración, los analistas deben satisfacer tan bien como sea posible, que el modelo sea capaz de predecir lo mas real posible. Probar las capacidades predicativas del modelo, es conocido como la validación.

Como mínimo, un modelo de demanda debe ser probado en términos de razonabilidad, muchas de estas pruebas cortas son ejecutadas durante la fase de calibración, por ejemplo, tiempo de viaje y costos. El objetivo de dichas pruebas es simplemente para medir que:

1. el modelo no viola las explicaciones teóricas
2. que no exhibe ninguna tendencia patológica
3. que es internamente consistente

Previsiones:

La estación final en el análisis de demanda, es el uso de técnicas de análisis calibradas y estimadas, para generar previsiones de demanda para cada alternativa política. Esto por supuesto es el último propósito del entero proceso de análisis de demanda. El tiempo y esfuerzo dedicado a esta tarea debe ser conmensurado con importancia, para el proceso completo de análisis.

La sensibilidad de la previsión para las suposiciones clave del modelo, debe también ser examinado, esto depende de lo hecho sobre el problema de estudio, el presupuesto y la disponibilidad del personal para ejecutar el trabajo. La necesidad de investigar un amplio rango de alternativas y escenarios supuestos representa una mayor fuerza debajo de los movimientos hacia el desarrollo de las técnicas del esquema de plantación, métodos de respuesta rápidos, que están intentando eliminar o reducir las técnicas de maximización del tiempo y los esfuerzos gastados en investigación de análisis alternativos.

Colecta de datos:

Dos tipos de datos son necesarios para el análisis de demanda:

1. *datos históricos*: sobre la conducta de viaje y la economía asociada, sistema de actividad y variables del sistema de transporte de interés.
2. *previsiones*: o especificaciones de valores futuros de variables socioeconómicas, sistema de actividad y sistema de transporte requerido para predecir la demanda futura para cada alternativa bajo consideración.

Históricamente datos son requeridos para las fases de calibración y modelación de análisis de demanda. De interés inmediato es la generación de características de los datos esperados, para la fase de previsión del análisis. Esos datos pueden ser divididos dentro de dos tipos: características de variables políticas caracterizando las alternativas bajo ciertas consideraciones y políticas no variables.

2.6 Teoría económica y conducta del consumidor.

Algunos de los conceptos teóricos más importantes que forman la base del análisis de la demanda vienen de la teoría económica. Existen cuatro importantes conceptos que son importantes para el análisis de transporte: la teoría de la conducta del consumidor, demanda derivada, la curva de oferta y equilibrio¹⁵.

La conducta del viaje del consumidor: la premisa de la teoría de la conducta del consumidor es que **“un individuo seleccionara un paquete de bienes sobre otros paquetes económicos si este le brinda la mayor utilidad”**, esto es, satisfacción. Las decisiones individuales realizan un proceso consistente de maximizar un concepto de utilidad con función U sujeta a un paquete de restricciones llamadas:

$$\begin{aligned} \max U &= U(X_1, \dots, X_n) \\ \text{sueto a } Y &= P_1 X_1 + \dots + P_n X_n \end{aligned}$$

Donde:

$X_1, X_2, \dots, X_n$ = bienes que esta consumiendo

$P_1, P_2, \dots, P_n$ = precio de los bienes

Y = ingresos

La solución de este problema puede ser ilustrado gráficamente usando una curva de inferencia y una línea de paquetes. Una diferencia en la curva representa varias combinaciones de bienes, en particular que permitirán al individuo mantener un nivel dado de utilidad. Para el caso de dos bienes, las siguientes ecuaciones caracterizan la utilidad del consumidor, maximizando el consumo de X_1 y X_2 , dados para un nivel de ingreso Y . la maximización de los valores de utilidad de X puede ser expresado en una relación funcional como:

$$\begin{aligned} X_1^* &= f(P_1, \dots, P_n, Y) \\ X_2^* &= f(P_1, \dots, P_n, Y) \\ X_3^* &= f(P_1, \dots, P_n, Y) \end{aligned} \tag{9}$$

Estas funciones son conocidas como funciones de demanda, y denotan la cantidad de utilidad maximizada de un bien que un individuo adquiere, dando el precio del bien y el ingreso de consumidor.

El nivel de equilibrio de consumo que fue obtenido por el bien, es denotado por E , tiene una importante interpretación económica. La validación del consumidor esta representada por un intercambio entre las comodidades que van

¹⁵ más de estos conceptos utilizados pueden ser encontrados en: Ortúzar, J de D, *Modelos de Demanda de Transporte*. Ediciones Universidad Católica, Santiago, 1994.

obteniendo para realizar a lo largo de la curva de inferencia¹⁶, mientras que las valuaciones del mercado de las comodidades esta representado por un precio.¹⁷

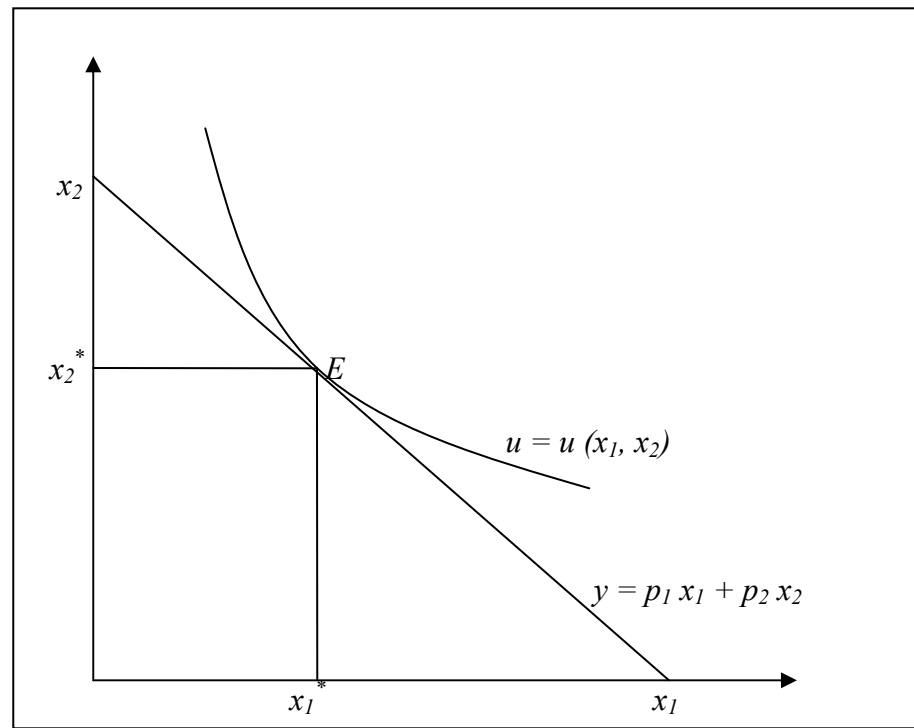


Figura 7¹⁸.

Esto ha sido supuesto que la utilidad es simplemente una función de cantidad de bienes consumidos, sin embargo, algunos, si no es que muchos bienes incluyen los servicios de transporte, no son consumidos por su cantidad, si no por sus atributos¹⁹, esto es, la comida es consumida por sus nutrientes, la ropa por su comodidad, valor y durabilidad, en otras palabras, estos son los atributos por los bienes que generan su utilidad y no la cantidad de los bienes por sí mismos, así que, la demanda de un producto depende sobre el precio, sus características relativas y las características del consumidor que adquiere el bien.

En el caso de transporte, el “bien” demandado es el servicio de transporte por sus varios modos entre puntos y espacio para propósitos particulares. El “precio” de este bien es generalmente tomado para no ser simplemente un costo monetario del viaje, pero otros costos percibidos por el usuario son también el tiempo gastado en la realización del viaje. Estos costos son definidos como costos “percibidos” del viaje en lugar de costo “verdadero”, sobre la suposición que los

¹⁶ Este intercambio es técnicamente referido hacia como la proporción marginal de sustitución.

¹⁷ En términos técnicos, el equilibrio de consumo ocurre en los casos de dos-bienes cuando la proporción marginal sobre los bienes es igual a su precio relativo.

¹⁸ Thomas A. Domencich and Daniel McFadden, Urban Travel Demand A behavioral analysis, North-Holland Publishing Company, Amsterdam, 1975.

¹⁹ Ortúzar, J de D y Willumsen, *Modelling Transport*. John Wiley & Sons, Chichester, 1990.

realizadores del viaje raramente usaran un costo verdadero para evaluar sus opciones de viaje. Si el valor monetario de tiempo es conocido, entonces el tiempo y precio puede ser combinado para el rendimiento, una generalización de costos de viaje, la utilidad de un viaje y la demanda para este, depende sobre un amplio rango de características de modos disponibles, y características de los individuos que realizan estos viajes.

Demanda derivada: el concepto de demanda derivada es la base para el más importante análisis de la demanda de transporte, esto es, el sistema de transporte es el mediador entre las localidades donde se va a realizar una actividad por conexiones físicas y habilita a las personas para moverse de una actividad a otra.

Dos implicaciones mayores pueden surgir de la naturaleza derivada de la demanda de transporte. Primero, el análisis de demanda de transporte no puede ser ejecutado sin consideraciones explícitas del sistema de actividad socioeconómico, que sirve al sistema de transporte y que genera la demanda de viajes. Sobre los más largos periodos de tiempo, la accesibilidad proporcionada por el sistema de transporte puede influenciar donde la gente vive y donde realizan sus actividades económicas, por ejemplo, la previsión de demanda de viajes sobre un tiempo de 20 a 30 años, necesariamente requiere una predicción del uso de suelo en el futuro.

Segundo, porque los viajes ocurren para la causa de “llegar ahí”, esto puede ser caracterizado en términos de tiempo, costo monetario, convivencia, incomodidad, etcétera., asociados con el viaje, estas características representan la desutilidad o generaliza el costo de viaje. Es razonable suponer que las decisiones de viaje son basadas sobre el potencial realizador del viaje de los “pros y contras” o las utilidades netas asociadas con los varios modos de transporte que parten de las opciones de movilidad disponibles. Este concepto de utilidad de viaje, es una herramienta indispensable para iniciar un sin número de técnicas de modelación de demanda de transporte.

La curva de oferta: la curva de oferta expresa la cantidad de un bien dado (Q_s) que se produce como una función del bien, expresado matemáticamente es $Q_s = Q_s(P)$. La función de oferta es frecuentemente hacia arriba (o de modo no decreciente) hacia la derecha

En el análisis de servicios de transporte, la oferta es definida a lo largo de tres dimensiones. El primero de ellos es el concepto de ejecución del sistema, que es, el tiempo de viaje y capacidades que el sistema de transporte proporciona para un capital invertido (representando la infraestructura y vehículos que comprenden el sistema), estrategia operada, y nivel de demanda. Una función de ejecución es una función de oferta inversa de la forma $P_s = P_s(Q)$ don del precio P_s , es interpretado un camino generalizado, el ejemplo clásico de la función de ejecución es el volumen como medida para el tiempo de viaje.

Equilibrio: el punto de intersección entre la curva de demanda y oferta es conocida como el punto de equilibrio. A este punto el precio del bien es tal que $Q_D = Q_S = Q_E$, esto es, la cantidad demanda es igual a la cantidad ofertada. El equilibrio no es preciso cuando el sistema para la demanda o la oferta se mueven a otro precio.

El equilibrio es un concepto fundamental en el análisis del sistema de transporte en que este es que une la oferta y la demanda. El equilibrio del nivel de tráfico, Q_E , debe dar un levantamiento para un nivel de ejecución P_s , sobre el lado de la oferta que en giro da un levantamiento para un nivel de demanda Q_D , sobre el lado demanda, esto es, mientras es frecuentemente conveniente tomar la demanda como arreglada y exógenamente con un análisis de oferta de transporte (y viceversa con el análisis de demanda), finalmente esta suposición es relajada, y un equilibrio es establecido entre demanda y oferta en el orden de lograr una descripción del sistema esperado bajo un juego de condiciones operativas. La suposición de que el mercado de viajes tiende hacia posiciones de equilibrio ha proporcionado un trabajo para ser usado como hipótesis y formas de las bases sobre la cual todas las técnicas de análisis de demanda y oferta son discutidas.

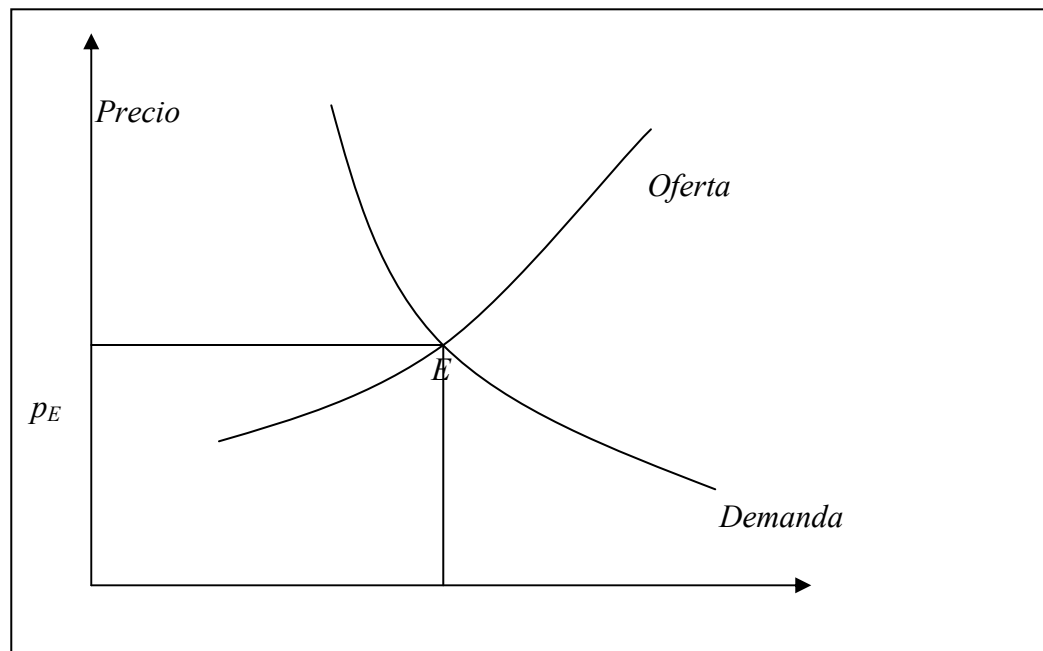


Figura 8²⁰.

²⁰ Thomas A. Domencich and Daniel McFadden, *Urban Travel Demand A behavioral analysis*, North-Holland Publishing Company, Amsterdam, 1975.

2.7 Base teórica para generar modelos de selección discreta.

La base teórica más común para generar modelos de selección discreta es la teoría del rango de utilidad²¹, los cuales postulan que:

1. Los individuos que pertenecen a una población homogénea dada (Q) actúan racionalmente y poseen información perfecta ya que en ellos siempre se selecciona una opción que maximiza su utilidad personal (identificada como “*homo economicus*”) sujeto a las restricciones legales, sociales, físicas y/o presupuestarias.
2. Un juego $A = \{A_1, \dots, A_j, \dots, A_N\}$ de alternativas posibles y un juego de vectores X de atributos medidos de los individuos y sus alternativas. Un individuo dado q es dotado con un juego de atributos $x \in X$ y a su vez mostrara un juego de selecciones $A(q) \in A$.
3. Cada opción $A_j \in A$ tiene asociada una red de utilidad U_{jq} para el individuo q . El modelador que es un observador del sistema no posee la información completa en relación con todos los elementos considerados por el individuo que realiza una selección, además, el modelador asume que U_{jq} puede ser representado por dos componentes:
 - a) Una parte medible, sistemática o representativa V_{jq} la cual es una función de las medidas de los atributos x ; y.
 - b) Una parte del rango ξ_{jq} el cual refleja la idiosincrasia y las pruebas particulares de cada individuo, junto con cualquier medida u observaciones de errores hechos por el modelador, entonces se puede postular que:

$$U_{jq} = V_{jq} + \xi_{jq} \quad (10)$$

Lo cual permite dos aparentes “irracionalidades” para ser explicadas; *i)* dos individuos con los mismos atributos y enfrentando el mismo juego de selecciones, puede tomar una opción diferente, *ii)* algunos individuos no siempre pueden seleccionar la mejor alternativa (desde el punto de vista de atributos considerados por el modelador).

Para la descomposición de la formula anterior y hacerla correcta se necesita una cierta homogeneidad en el estudio de la población. En principio se requiere que *todos los individuos proporcionen el mismo juego de alternativas y enfrente las mismas restricciones*²², y para conocer esto puede necesitarse segmentar el mercado. Si se tiene un termino V representativo esto lleva el subíndice q porque esto es una función de atributos x y puede variar de individuo a individuo, también,

²¹ McFadden y Reid, “*Aggregate travel demand forecasting from disaggregated behavioral models*”, libro de trabajo No 228, 1974.

²² Ver: Ortúzar, J de D y Willumsen, *Modelling Transport. John Wiley & Sons*, Chichester, 1990.

sin perder generalidad esto puede suponer que los residuos ξ son variables de rango, principalmente, 0 y una cierta probabilidad de distribución por ser especificada.

$$V_{jq} = \sum_k \theta_{kj} x_{jkq} \quad (11)$$

Donde los parámetros θ son supuestos para ser constantes para todos los individuos (modelo de coeficientes arreglados), pero pueden cruzar varias alternativas. Esto es importante para enfatizar la existencia de dos puntos de vista en la formulación de lo relacionado con el problema; primero, del individuo se pesan los elementos de interés y se selecciona la opción más conveniente, segundo, de los modeladores, por observación, solo eligen algunos de los elementos necesarios de los elementos ξ , para explicar que por otra parte sería la cantidad de conducta no racional.

4. Los individuos q seleccionan la alternativa de máxima utilidad, que es la selección individual A_j si y solo si:

$$U_{jq} \geq U_{iq}, \forall A_i \in A(q)$$

y es

$$V_{jq} - V_{iq} \geq \xi_{iq} - \xi_{jq}$$

Como el análisis ignora el valor de $(\xi_{iq} - \xi_{jq})$ esto no es posible para determinar con certeza si la ecuación (última) lo tiene. La probabilidad de elegir A_j está dada por:

$$P_{jq} = \Pr \{ \xi_{iq} \leq \xi_{jq} + (V_{jq} - V_{iq}), \forall A_i \in A(q) \} \quad (12)$$

y como la distribución de los residuos ξ no son conocidos, no es posible derivar una expresión analítica para el modelo. Si los residuos son variables de rango con una cierta distribución la cual se puede denotar por $f(\xi) = f(\xi_1, \dots, \xi_N)$. Nos permite saber que la distribución de $U, f(U)$, es la misma pero con diferente significado. Lo que da una ecuación más concisa:

$$P_{jq} = \int_{R_N} f(\xi) d\xi \quad (13)$$

donde

$$R_N = \begin{cases} \xi_{iq} \leq \xi_{jq} + (V_{jq} - V_{iq}), \forall A_i \in A(q) \\ V_{jq} + \xi_{jq} \geq 0 \end{cases}$$

y diferentes formas del modelo pueden ser generadas dependiendo de la distribución de los residuos ξ .

Una clase importante modelos de rango de utilidad son aquellos generados por las funciones de utilidad con residuos independientes y distribuidos idénticamente (IID), en este caso $f(\xi)$ puede ser descompuesto como:

$$f(\xi_1, \dots, \xi_N) = \prod_n g(\xi_n)$$

Donde $g(\xi_n)$ es la función de utilidad distribuida asociada con opción A_n , y la expresión general se reduce a:

$$P_j = \int_{-\alpha}^{\alpha} g(\xi_j) d(\xi_j) \prod_{i \neq j} \int_{-\alpha}^{V_j - V_i + \xi_j} g(\xi_i) d\xi_i \quad (14)$$

Donde se tiene extendido el rango de ambas integrales hasta $-\alpha$ en orden a resolverlo. Una interpretación de dos dimensiones de este modelo, junto con extensión para el caso más general de correlación y varianzas desiguales²³, también la ecuación anterior puede ser expresada como:

$$P_j = \int_{-\alpha}^{\alpha} g(\xi_j) d(\xi_j) \prod_{i \neq j} G(\xi_j + V_j - V_i) \quad (15)$$

con

$$G(x) = \int_{-\alpha}^{\alpha} g(x) dx$$

Y es interesante mencionar que una gran cantidad de esfuerzo ha sido empleado en tratar de encontrar la forma apropiada para g , la cual pertenece a la última ecuación para ser resuelta. Note que los residuos de IID requieren principalmente que las alternativas deban ser independientes. Las opciones de modo mezclado, por ejemplo, combinaciones de tren y auto, usualmente violan estas condiciones.

²³ Son presentadas por: Ortúzar, J de D y Willumsen, *Modelling Transport*. John Wiley & Sons, Chichester, 1990.

2.8 Características generales del modelo Probit y del modelo Logit.

En el modelo Probit²⁴ lo que se hace es trabajar con probabilidades, es decir, la probabilidad de ser un individuo del tipo 1 ó del tipo 0 dadas las características del individuo. Estas probabilidades se hacen depender del siguiente modo de las variables explicativas o características del individuo:

$$P(d_i = 1) = F(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}),$$
$$P(d_i = 0) = 1 - F(\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki})$$

Donde la función matemática F transforma a $\beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}$ en una probabilidad, y por tanto tiene que estar entre 0 y 1, es una función de distribución *normal* de probabilidad, y lo que obtenemos trabajando con este modelo es la probabilidad de que el individuo tome o no tome la opción planteada en el modelo.

El modelo logit²⁵ es usado para modelar una relación entre una variable dependiente Y y una o mas variables independientes X . La variable dependiente Y , es una variable discreta que representa una alternativa o categoría, de un conjunto mutuamente exclusivo de alternativas o categorías.

Las variables independientes son para afectar la alternativa, categoría o selección, y representar una opinión de *prioridad*, a cerca de la causa o elementos asociados importantes, en la selección o en el proceso de clasificación, así de esta forma se tiene que trabajando con el modelo logit se tienen una serie de alternativas

El modelo Logit aparece como una alternativa de modelación razonable en términos prácticos, al permitir su estimación y uso con un computador personal en un tiempo moderado, permitiendo modelar correlación y heteroscedasticidad. Su estructura está determinada por la forma del término que se agrega a la función de utilidad y puede llegar a ser tan general como se desee. En este contexto, puede constituirse en un "competidor" del modelo Probit, que hasta hace poco era considerado como el único o principal camino de flexibilizar la modelación de selecciones discretas considerando correlación y heteroscedasticidad, y también un competidor para el modelo Probit son los modelos Logit Multinomial y Logit Jerárquico, que son los más usados hasta ahora.

La diferencia central entre el Probit y el Logit está en la forma en que en cada uno especifica su estructura. En el caso del Probit, se hace directamente en la matriz de covarianza, y en el caso del Logit se hace directamente en la función de utilidad. En términos de estimación el modelo Logit parece requerir un tiempo de procesamiento un poco menor que el Probit.

²⁴ Ver: Ortúzar, J de D., "Fundamentals Of Discrete Multimodal Choice Modelling", Transport Reviews, 1982.

²⁵ Ver: Ortúzar, J de D., "Fundamentals Of Discrete Multimodal Choice Modelling", Transport Reviews, 1982.



Desde un particular punto de vista, el modelo Logit es una alternativa de modelación interesante y útil más que el modelo Probit, y lo importante sigue siendo tener claras las propiedades del modelo que se está utilizando y justificar adecuadamente, en base a consideraciones teóricas, una cierta estructura de modelo previo a la estimación de los parámetros, por lo cual este trabajo se enfocara a el modelo logit y sus variaciones, como son:

- El modelo logit binario,
- El modelo logit multinomial,
- El modelo logit jerárquico y;
- El modelo logit mixto.

Modelos explicados en el capitulo 3.

Capítulo 3

El Modelo Logit

3.1 Modelo Logit Binario

El modelo logit binario (solo dos alternativas 0 y 1) es utilizado para determinar la demanda de viajes entre dos alternativas y es de costos generalizados de viaje, este modelo esta calibrado solo para encontrar los valores de la mejor opción para un parámetro de dispersión λ y una penalidad δ^1 (supuesta para el segundo modo).

El concepto de máxima entropía² puede ser usado para generar modelos de distribución y selección modal simultáneos. Entonces considerando el problema de máxima entropía en términos de dos modos tenemos que:

$$\text{Max. } \ln W \{T_{ij}^k\} = - \sum_{ijk} (T_{ij}^k \ln T_{ij}^k - T_{ij}^k) \quad (16)$$

$$\text{s.a. } \sum_{jk} T_{ij}^k - O_i = 0 \quad (17)$$

$$\sum_{ik} T_{ij}^k - D_j = 0 \quad (18)$$

$$\sum T_{ij}^k C_{ij}^k - C = 0 \quad (19)$$

Donde: k : es el modo de transporte.

Y la solución al problema es:

$$T_{ij}^k = A_i O_i B_j D_j \exp(-\beta C_{ij}^k) \quad (20)$$

$$P_{ij}^k = \frac{T_{ij}^k}{T_{ij}} = \frac{\exp(-\beta C_{ij}^k)}{\exp(-\beta C_{ij}^1) + \exp(-\beta C_{ij}^2)} \quad (21)$$

donde:

P_{ij}^1 : es la proporción de viajes de i a j por el modo 1.

β : es el parámetro de dispersión a calibrar.

C_{ij}^k : costo, tiempo o distancia de viaje; de i a j , por el modo de transporte k .

¹ Ortúzar, J de D y Willumsen, "Modelling Transport", John Wiley & Sons, Chichester, 1990.

² Ver: Ortúzar j. de D. y Willmsen, "Modelling Transport .third edition", edit: John Wiley & Sons, 2001.

La expresión (21) es conocida como Modelo Logit Binario, la cual tiene las siguientes propiedades:

- Genera una curva de tipo S, similar a algunas curvas empíricas (figura 9).

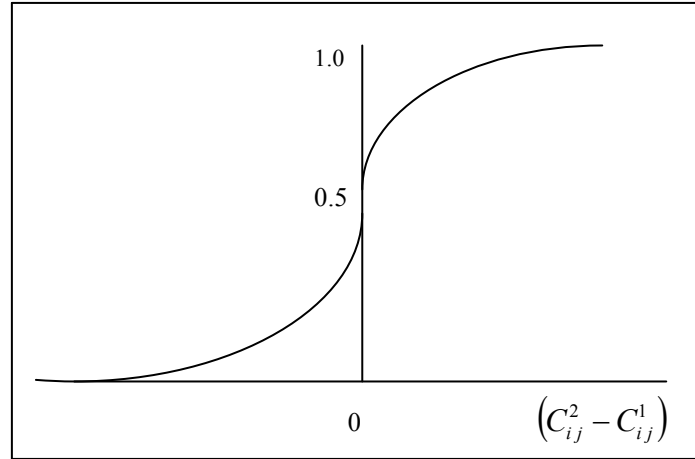


Figura 9³: curva tipo s.

- Si $C_1 = C_2$, entonces $P_1 = P_2 = 0.5$
- Si $C_1 \ll C_2$, entonces P tiende a 1.0
- El modelo se puede generalizar para varios modos.

En esta formulación el parámetro β tiene un doble papel:

1. Actúa como parámetro que controla la dispersión, en la selección de modo.
2. Controla la dispersión entre destinos, para diferentes distancias desde el origen.

Ejemplo del modelo logit binario:

Se tienen datos agregados de los costos entre dos modos de transporte de 5 pares de zonas, se requiere determinar la proporción de viajeros que elegirán cada modo, y el modelo logit binario tendrá los parámetros: $\beta = 0.72, \delta = 0.055$ (la obtención de parámetros se explicara mas adelante).

El planteamiento y solución del problema es:

$$P_{ij}^k = \frac{\exp(-\beta C_{ij}^2)}{\exp(-\beta C_{ij}^2) + \exp(-\beta C_{ij}^1)} =$$

$$P_1 = \frac{1}{1 + \exp\{-\beta(C_2 + \delta - C_1)\}} = \frac{1}{1 + \exp\{0.72(C_2 + 0.055 - C_1)\}}$$

³ Fuente: Ortúzar, J de D y Willumsen, *Modelling Transport third edition*.. John Wiley & Sons, Chichester, 2001.

Sustituyendo los valores de los costos que se tienen por cada zona, nos da como resultado las siguientes probabilidades:

Zona	$P_1(\%)$	$P_2(\%)$	C_1	C_2
1	51	49	21	18
2	57	43	15.8	13.1
3	80	20	15.9	14.7
4	71	29	18.2	16.4
5	63	37	11.0	8.5

3.2 Modelo Logit Multinomial.

El modelo más conocido y utilizado de selección modal es el llamado *logit multinomial*, en este caso, se puede establecer la expresión analítica para determinar las proporciones seleccionadas, definida por:

$$P_{ij}^1 = \frac{\exp(-\beta C_{ij}^1)}{\sum_k \exp(-\beta C_{ij}^k)} \quad (22)$$

Donde:

P_{ij}^1 : es la proporción de viajes de i a j por el modo 1.

β : es un parámetro de dispersión a calibrar.

C_{ij}^k : es costo, tiempo o distancia de viaje; de i a j , por el modo de transporte k .

Los datos analizados por el modelador generalmente no contienen mezcla de los atributos específicos del modo de viaje o alternativa y del viajero, lo cual sería muy costoso, por lo que se hace necesario analizar los dos tipos de datos en forma separada, y considerar aspectos del modelo que son específicos para cada uno de ellos.

De esta forma cuando se tienen datos que consideran las características específicas de los viajeros (características que no son comunes a todos), el modelo apropiado es el logit multinomial, el cual se puede establecer como:

$$P_{ij} = \frac{\exp(-\beta_j X_i)}{1 + \sum_k \exp(-\beta_k X_i)}$$

donde $k = 0, 1, 2, 3, \dots$

X_i : son las características diferentes que tienen los viajeros, por ejemplo; ocupación, nivel de estudios, nivel de ingreso, sexo, edad, etc.

Ahora bien, cuando se cuenta con datos que consideran los atributos específicos del modo de viaje o alternativa de selección, el modelo apropiado es el logit condicional, el cual se puede establecer como:

$$P_{ij}^1 = \frac{\exp(-\beta Z_{ij}^1)}{\sum_j \exp \sum_j (-\beta Z_{ij}^k)}$$

donde $j = 1, 2, 3, \dots$

Z_{ij} : son los atributos del modo o alternativa, por ejemplo: tiempo de viaje, costo del pasaje, tiempo de caminar, costo de estacionamiento, tiempo en el vehículo, tiempo de espera, etc.

Ejemplo del modelo logit multinomial:

Se tiene una muestra de 3 viajeros, cada uno de los cuales, tiene solo 2 alternativas (modos) de viaje, se debe considerar que solo se tiene un atributo para el costo de viaje (C_{ij}), y suponer que el valor del parámetro $\beta = 0.756$.

Las observaciones hechas fueron:

Viajero	Modo elegido	C_{ij}^1	C_{ij}^2
1	1	5	3
2	1	1	2
3	2	3	4

Entonces desarrollando la ecuación del modelo logit multinomial tenemos que:

$$P_{ij}^1 = \frac{\exp(-\beta C_{ij}^1)}{\sum_k \exp(-\beta C_{ij}^k)} = P_{ij}^1 = \frac{\exp(-\beta C_{ij}^1)}{\exp(-\beta C_{ij}^1) + \exp(-\beta C_{ij}^2)}$$

$$P_{ij}^2 = 1 - P_{ij}^1$$

Sustituyendo los valores en la formula, nos da como resultado las siguientes probabilidades:

Viajero	P_{ij}^1	P_{ij}^2
1	0.82	0.18
2	0.32	0.68
3	0.32	0.68

3.3 Modelo Logit Jerárquico

Una de las hipótesis que conduce a la formulación del modelo logit multinomial, es que las variables aleatorias, asociadas a cada alternativa, son independientes entre sí.

Cuando un subconjunto de alternativas presenta mayor similitud (por ejemplo, modos de transporte público frente al automóvil) esta hipótesis pierde su validez. En estos casos algunas de las propiedades del logit multinomial empiezan a considerarse como problemas.

Para manejar situaciones como ésta, hay una formulación más genérica llamada modelo *logit jerárquico* o *anidado*. Su estructura se caracteriza por agrupar los subconjuntos de alternativas correlacionadas (o más semejantes) en un nivel jerárquico o nido. Por ejemplo, modos de transporte público, como el microbús y el metro, pueden constituir un nido.

Cada nido, en su caso, está representado por un "*modo compuesto*" o *súper-modo*, que compite con los demás modos en un mismo nivel jerárquico. Retomando el ejemplo anterior, el *súper-modo* "*transporte público*" (compuesto por ómnibus y metro) compite con el automóvil. El siguiente diagrama presenta la estructura de decisión, semejante a un árbol invertido, adoptada para representar el ejemplo anterior.

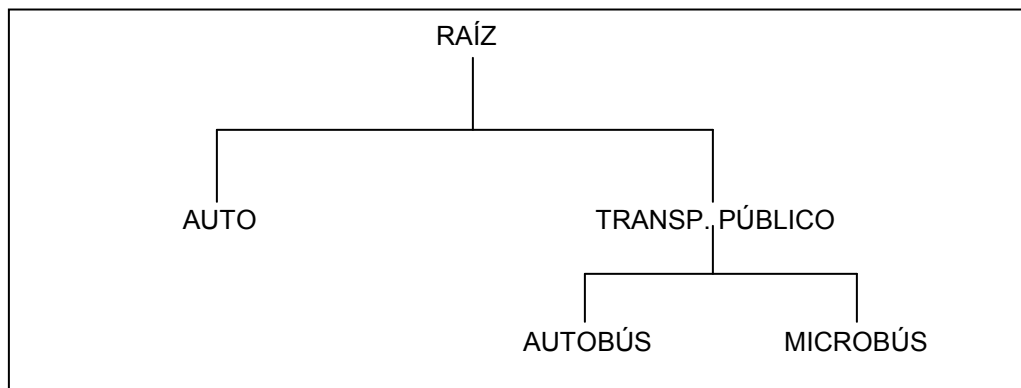


Figura 10⁴: Árbol con dos niveles jerárquicos o nidos

El siguiente diagrama presenta un ejemplo de estructura de decisión con más nidos y más niveles jerárquicos.

⁴ Fuente: Ortúzar, J de D y Willumsen, *Modelling Transport third edition*, John Wiley & Sons, Chichester, 2001.

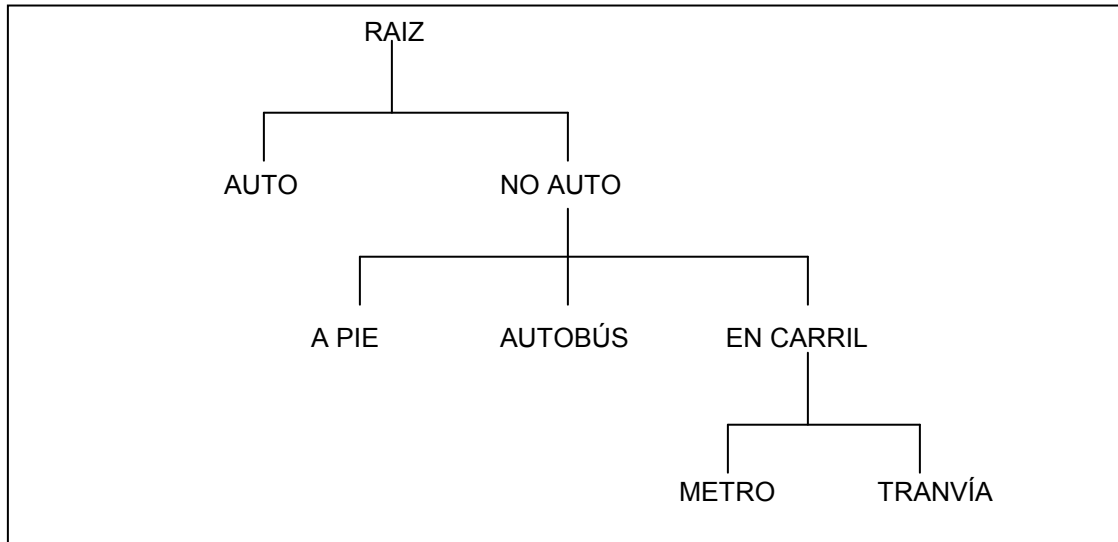


Figura 11⁵: Árbol con tres niveles jerárquicos o nidos

Una ventaja del modelo logit jerárquico, es la posibilidad de calibrarlo usando la forma desarrollada para el modelo logit multinomial.

Para la solución del modelo logit jerárquico se procede de la siguiente forma:

- 1.- Primero, se estima un modelo logit múltiple para las alternativas (modos), por ejemplo metro y tranvía, omitiendo todas las variables (W) que tengan el mismo valor, para este subconjunto de alternativas.

Estas variables se reintegran posteriormente en la jerarquía superior, ya que afectan las elecciones de las alternativas de los nidos o jerarquías superiores.

- 2.- Segundo, la introducción del nido inferior a la jerarquía superior, se hace a través de la alternativa compuesta, a la que se asocia una utilidad representativa de los nidos, la cual tiene dos componentes:

a).- El valor esperado de la utilidad máxima entre alternativas del nido o jerarquía (EMU), considerada como variable.

b).- El vector de atributos comunes (W) a todas las alternativas del nido o jerarquía.

⁵ Ortúzar, J de D y Willumsen, *Modelling Transport third edition*, John Wiley & Sons, Chichester, 2001.

donde (EMU) tiene la siguiente expresión: $EMU = \ln \sum \exp(U_j)$,

con U_j = utilidad de la alternativa A_j del nido, exceptuando a las variables que toman el mismo valor.

Entonces la utilidad compuesta del nido o jerarquía, estará dada como:

$$U_i = \beta (EMU) + \alpha W \quad (23)$$

donde β y α son los parámetros a estimar.

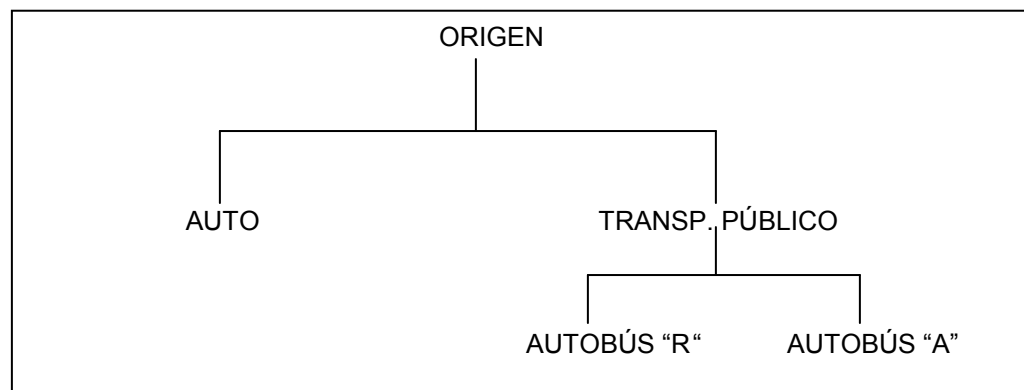
3.- Tercero, hecho lo anterior en el nivel jerárquico superior se debe estimar otro modelo logit múltiple (o binario), que debe incluir la alternativa compuesta y las demás alternativas no contenidas en el nido o jerarquía inferior.

4.- Por ultimo, la probabilidad de que un viajero seleccione la alternativa A_j , estará dada por la probabilidad marginal de elegir la alternativa compuesta (en la jerarquía superior) y la probabilidad condicional de que elija la alternativa A_j , en la jerarquía inferior, dado que se seleccionó la alternativa compuesta.

Ejemplo del modelo logit jerárquico:

Considerar una ciudad donde los viajes se hacen en auto particular (C) y en autobús (B). En términos del modelo, para cualquier alternativa se tiene que el costo de utilizar auto, es igual al costo de utilizar autobús, esto es: $C_c = C_B$.

Ahora se considera que la compañía de autobuses decide pintar la mitad con color rojo y la otra mitad con color azul, pero dando el mismo servicio. Esto significa que el costo de los autobuses rojos será igual al costo de los autobuses azules, esto es: $C_{BR} = C_{BA}$. Además el costo de utilizar el auto particular no se modifica. Entonces la partición modal nos queda de la siguiente forma:



Para este ejemplo se tienen las siguientes expresiones:

$$P_C = \frac{1}{1 + \exp\{-\beta_1(C_B - C_C)\}}; \text{ y } P_B = 1 - P_C$$

$$P_{BR} = \frac{1}{1 + \exp\{-\beta_2(C_{BR} - C_{BA})\}}; \text{ y } P_{BA} = 1 - P_{BR}$$

$$\text{con: } C_B = \frac{-1}{\beta_2} \ln[\exp(-\beta_2 C_{BR}) + \exp(-\beta_2 C_{BA})]$$

Donde:

P_C = Probabilidad de elegir auto

$(1 - P_C)P_{BR}$ = Probabilidad de elegir autobús rojo

$(1 - P_C)P_{BA}$ = Probabilidad de elegir autobús azul

β_i = Parámetro de separación primaria y secundaria, donde $i = 1, 2$

Resolviendo la primera parte y considerando que $C_{BA} = C_{BR}$, se tiene que:

$$P_{BR} = \frac{1}{1 + \exp\{-\beta_2(C_{BR} - C_{BA})\}} = 0.5$$

$$P_{BA} = 1 - P_{BR} = 1 - 0.5 = 0.5$$

Resolviendo la segunda parte y teniendo la misma consideración anterior que: $C_B = C_C$, entonces:

$$P_C = \frac{1}{1 + \exp\{-\beta_1(C_B - C_C)\}} = 0.5$$

$$P_B = 1 - P_C = 1 - 0.5 = 0.5$$

Entonces tendríamos como resultado que: $P_C = 0.5, P_{BR} = 0.25, P_{BA} = 0.25$

Sin embargo si el costo compuesto para autobús C_B , no es igual que para C_{BR} y C_{BA} , lo cual depende de β_2 y para el problema del autobús se tendría la siguiente expresión:

$C_B = C_{BA} - \frac{1}{\beta_2} \ln 2$, entonces el costo compuesto de C_B será mas bajo que para C_{BR} ó C_{BA} .

Si por ejemplo se considera $C_{BR} = C_{BA} = 50$ y $\beta_2 = 0.9$ se tiene que:

$$C_B = C_{BA} - \frac{1}{\beta_2} \ln 2 = 50 - \frac{1}{0.9} \ln 2 = 49.23$$

Y la probabilidad de escoger el auto dependerá del valor de β_1 , entonces podríamos tener:

β_1	P_C
0.001	0.5
0.005	0.499
0.05	0.490
0.1	0.481
...	...
0.9	0.333

3.4 Modelo Logit Mixto

La idea de los modelos logit mixtos como tal, no es una idea nueva, modelos de estas características han sido propuestos con anterioridad⁶, en donde un modelo equivalente a los actuales modelos logit mixtos, es descrito con el nombre de modelo Hedónico. Su reciente reaparición con otro nombre puede deberse a que los avances tecnológicos en computación y métodos numéricos, permiten ahora su estimación en menor tiempo. Recientemente este tipo de modelos ha sido utilizado para modelar diversas situaciones⁷.

En términos más específicos, los modelos logit mixtos nacen de suponer una función de utilidad U_{in} conformada por una componente determinística V_{in} , una componente aleatoria ξ_{in} independiente e idénticamente distribuida, y uno o más términos aleatorios adicionales.

Estos términos de error adicionales pueden ser agrupados en un término aditivo η_{in} , que puede ser función de datos observados de la alternativa, y que permite recoger la presencia de correlación (no impedancia entre valores sucesivos y heteroscedasticidad (varianza constante). Así, la función de utilidad queda definida como:

$$U_{in} = V_{in} + \eta_{in} + \xi_{in}$$

Se puede ver que se trata de una forma distinta de justificar un determinado modelo. La forma usual es hacer supuestos directamente sobre la distribución del

⁶ Por ejemplo se puede citar los trabajos de Cardell y Dunbar (1980) y Boyd y Melman (1980)

⁷ Stopher, P y McDonald, KG, *Trip generation by cross-classification: an alternative methodology*. Transportation Research Record 891, 10-17, (1983).

término de error ξ_{in} , como por ejemplo en el caso del Probit. En cambio, en un modelo Logit Mixto lo que se hace es construir una estructura de error diferente agregando términos que sean fuente de correlación y/o heteroscedasticidad.

Se asume que ξ es IID Gumbel, mientras que η sigue una función de distribución cualquiera definida por una densidad $f(\eta/\theta^*)$, donde θ^* son parámetros fijos que la describen (media y varianza)⁸. Como θ es IID Gumbel, la probabilidad condicional en η que el individuo escoja la alternativa i esto corresponde al modelo Logit Multinomial (o Logit Simple):

$$P_n(i/\eta) = L_{in}(\eta) = \frac{e^{V_{in} + \eta_{in}}}{\sum_j e^{V_{jn} + \eta_{jn}}}$$

Por lo tanto, la probabilidad de elegir la alternativa corresponde a la integral de la probabilidad condicional sobre todos los posibles valores de η , lo que depende de los parámetros de su distribución, esto es:

$$P_{in} = \int L_{in}(\eta) f(\eta/\theta^*) d\eta$$

Como caso particular, puede suponerse una función de utilidad con la siguiente especificación⁹:

$$U_{in} = \underbrace{\beta^t x_{in}}_{V_{in}} + \underbrace{\mu_{in}^t z_{in}}_{\eta_{in}} + \xi_{in} \quad (24)$$

En esta expresión se asume que la componente determinística de la utilidad es lineal en los parámetros β que ponderan los atributos x_{in} . Por otro lado, se asume que η depende de ciertos parámetros (μ_{in}) y datos observados relacionados con la alternativa $i(z_{in})$, relación que también se supone lineal en parámetros. Un supuesto adicional es que el término μ es propio del individuo, sin variar entre alternativas, es decir:

$$\eta_{in} = \mu_n^t z_{in} \quad (25)$$

⁸ En términos prácticos, la distribución de los términos aleatorios generalmente se asume normal, existiendo diversos argumentos detrás de este supuesto. Otra distribución que ha sido utilizada es la log-normal, especialmente en aquellos casos en que se quiere restringir el signo de un determinado parámetro.

⁹ b es un vector de parámetros de dimensión L (hay L variables explicativas en la componente determinística de la función

de utilidad); x es un vector de atributos de dimensión L ; μ_{in} es un vector aleatorio de dimensión K cuyas componentes tienen media cero y con matriz de covarianza Σ ; z_{in} es un vector de atributos asociados con la alternativa i y el individuo n ,

y tiene dimensión K ; finalmente, ξ_{in} es una variable aleatoria que representa el error estocástico.

Esta especificación es la que ha sido utilizada en la mayor parte de los estudios previos¹⁰.

Dada una función de utilidad como en la ecuación (24) y considerando además un supuesto usual (25), sea z_n la matriz de dimensión $K \times J$ que contiene a los vectores z_{in} para cada alternativa perteneciente al conjunto de selección del individuo ($i \in C_n$) y ξ_n un vector aleatorio IID Gumbel con matriz de covarianza $\sum_{\xi}$ y que contiene a los elementos ξ_{in} . Si se asume que cada término de μ_n tiene una función de densidad con media cero y varianza σ_k^2 y que el vector en su conjunto tiene una matriz de covarianza Ω , entonces la matriz de covarianza del modelo ($\sum$), puede escribirse como:

$$\sum = z_n^t \cdot \Omega \cdot z_n + \sum_{\xi} = z_n^t \Omega z_n + \sigma_{\xi}^2 I$$

Es claro que la dimensión de la matriz de covarianza obtenida está bien definida¹¹ y de esta expresión general se concluye que el modelo es capaz de modelar correlación y heteroscedasticidad entre alternativas. En efecto, si obtenemos la covarianza entre dos alternativas se observa que para $i, j \in C_n$, con i o j :

$$\text{cov}(U_{in}, U_{jn}) = \sum_{k=1}^K z_{kin} z_{kjn} \sigma_k^2$$

Que en general será distinto de cero, si es que para al menos algún $k, \sigma_k^2 > 0$ y $z_{kin}, z_{kjn} \neq 0$. En tal caso se asegura la presencia de correlación entre las alternativas i y j .

En cuanto a la varianza;

$$\text{var}(U_{in}) = \sum_{k=1}^K z_{kin}^2 \sigma_k^2 + \frac{\pi}{6\lambda^2}$$

Luego si $\text{var}(U_{in}) \neq \text{var}(U_{jn})$, se asegura la heteroscedasticidad entre dichas alternativas.

¹⁰ Ben-Akiva, ME y Lerman, *Discrete Choice Analysis*. The MIT Press, Cambridge Mass, 1985.

¹¹ La matriz de covarianza es de dimensión $J \times J$. en efecto, como Ω es de dimensión $K \times K$ (con K el número de componentes aleatorias), y z_n tiene dimensión $K \times J$, entonces $z_n^t \cdot \Omega \cdot z_n$ es una matriz de dimensión $J \times J$; luego al sumarse esta ultima matriz $\sum_{\xi}$, que es de dimensión $J \times J$, se obtiene finalmente que $\dim \sum = \dim(z_n^t \cdot \Omega \cdot z_n + \sum_{\xi}) = J \times J$.

Es así como, un modelo logit mixto con parámetros aleatorios con distribución normal, puede aproximar a un modelo Probit.

Además el modelo logit mixto, al permitir modelar la presencia de correlación entre alternativas, es capaz de levantar el supuesto de independencia de alternativas irrelevantes propio del modelo Logit Multinomial. En otras palabras, los patrones de sustitución entre alternativas son flexibles. Dada una probabilidad tipo logit mixto, se puede demostrar que la razón entre probabilidades de dos alternativas, depende de todo el conjunto de alternativas disponibles.

$$P_{in} = \int_{-\alpha}^{\alpha} \dots \int_{-\alpha}^{\alpha} \left\{ \frac{\exp(\beta^t x_{in} + \mu_n^t z_{in})}{\sum_j \exp(\beta^t x_{in} + \mu_n^t z_{in})} \right\} f(\mu_{1n}) \dots f(\mu_{Kn}) d\mu_{1n} \dots d\mu_{Kn}$$

Entonces, si se considera R valores de η obtenidos de una función de densidad $f(\eta / \theta^*)$, para cada una de las repeticiones es posible calcular:

$$P_n(i / \eta^r) = L_{in}(\eta^r) = \frac{e^{V_{in} + \eta_{in}^r}}{\sum_j e^{V_{jn} + \eta_{jn}^r}} \quad (26)$$

con $r = 1, \dots, R$. Luego es posible obtener la probabilidad promedio:

$$\tilde{P}(i) = \frac{1}{R} \sum_{r=1}^R L_{in}(\eta^r) \quad (27)$$

y con ella construir la función de verosimilitud simulada:

$$SL = \sum_q \sum_{i=1}^j y_{qi} \ln \tilde{P}(i) \quad (28)$$

La probabilidad de selección de un modelo logit mixto como la expresada en la ecuación (27), no posee una expresión matemática cerrada a diferencia del Logit Multinomial o del Jerárquico. Es más, como la integral no puede resolverse analíticamente es preciso utilizar simulación. Sin embargo, se puede aprovechar el hecho de que la probabilidad condicional (21) tiene una expresión de forma logit multinomial. Bajo condiciones de regularidad, el estimador resulta asintóticamente equivalente al método de máxima verosimilitud¹².

¹² Ortúzar, J de D y Willumsen, *Modelling Transport*. John Wiley & Sons, Chichester, 1990.

3.5 Especificaciones requeridas por el modelo logit

1.- Identificar el conjunto C de alternativas para su selección. Estas pueden ser diferentes dependiendo de la localización geográfica, población, características socioeconómicas, atributos de las alternativas y factores que influyen el contexto de la selección.

2.- Identificar el subconjunto factible C_n para viajeros de la muestra. Nótese que existen dos conjuntos de selección, uno para la población completa C , y otro para los individuos seleccionados en la muestra. Es importante que en los conjuntos de selección estén representados todos los modos, y no se incluyan modos que no están considerados.

En la práctica puede ser difícil hacer estimación con conjuntos de selección restringidos, sin embargo, el modelo resultante puede ser mejorado si las restricciones del conjunto de solución son conocidas por los viajeros.

3.- Identificar variables y características. El analista debe identificar cuáles variables influyen en el proceso de decisión, qué características de los viajeros son importantes en el proceso de selección y cómo realizar sus mediciones.

4.- Diseñar y administrar una serie de instrumentos (incluyendo dispositivos y esquemas de muestreo), necesarios para recolectar información o para observar y registrar la selección hecha por los viajeros.

5.- Finalmente, los modelos logit múltiple son estimados y redefinidos para seleccionar “lo mejor”, usando todos los datos obtenidos en las etapas previas.

3.6 Estimación estadística de los parámetros del modelo logit

Para estimar los parámetros que aparecen en la función de utilidad del modelo logit, existen varios métodos. Uno de ellos es “el método de regresión lineal” y otro es “el método de máxima verosimilitud”.

El método de regresión lineal:

La técnica de regresión lineal es aplicada cuando se pretende explicar una variable respuesta cuantitativa, en función de una variable explicativa cuantitativa, también llamada variable independiente, variable regresora o variable predictora.

El método de regresión lineal trabaja bajo los siguientes supuestos para los modelos de selección discreta:

- a) La distribución de probabilidades de Y_i para un valor dado X_i , tiene la varianza σ^2 para todos los valores X_i .
- b) Las medias $\mu_i = E(Y_i)$ forman una línea recta $\mu_i = E(Y_i) = a + bX_i$ a y b son los parámetros a estimar.

c) Las variables Y_i son independientes.

La estimación de los parámetros a y b se pueden estimar con las siguientes expresiones:

1. $\hat{a} = \bar{Y} - \hat{b}\bar{X}$; para esta ecuación la media de X y Y es:

$$\bar{Y} = \frac{\sum Y_i}{n}; \quad \bar{X} = \frac{\sum X_i}{n}$$

$$2. \quad \hat{b} = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{n \sum X_i^2 - (\sum X_i)^2}$$

Los estimadores anteriores tienen las siguientes propiedades:

$$a) \quad E(\hat{a}) = a; \quad Var(\hat{a}) = \frac{\sigma^2}{n}$$

$$b) \quad E(\hat{b}) = b; \quad Var(\hat{b}) = \frac{\sigma^2}{\sum X_i^2}$$

Para verificar la relación de los parámetros obtenidos con los datos dados la ecuación del coeficiente de correlación es:

$$r_{x,y} = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{n \sum X_i^2 - (\sum X_i)^2} \sqrt{n \sum Y_i^2 - (\sum Y_i)^2}}$$

El valor del coeficiente de correlación va de $-1 \leq r_{x,y} \leq 1$; y se presentan 3 casos:

- I. cuando $r_{x,y} = -1$ se tiene relación negativa perfecta
- II. cuando $r_{x,y} = 1$ se tiene relación positiva perfecta
- III. cuando $r_{x,y} = 0$ se tiene nula regresión.

Ejemplo:

Calibrar el parámetro β para el modelo logit binario:

$$\text{Ecuación del modelo logit binario: } P_{ij}^k = \frac{T_{ij}^1}{T_{ij}} = \frac{\exp(-\beta C_{ij}^2)}{\exp(-\beta C_{ij}^2) + \exp(-\beta C_{ij}^1)}$$

Donde:

k: es el modo de transporte (1= a modo 1 y 2= a modo 2)

P_{ij}^1 : es la proporción de viajes de i a j por el modo 1.

β : es el parámetro de dispersión a calibrar.

C_{ij}^k : Costo, tiempo o distancia de viaje; de i a j, por el modo de transporte k.

δ : Representa atributos no incluidos como: comodidad, conveniencia, seguridad, etc.

Se supondrá que el costo de transporte es un valor ya conocido, así de esta forma, el modelo solo tendrá dos parámetros a calibrar β y δ .

Para poder trabajar con facilidad se eliminan subíndices y nos queda:

$$P_1 = \frac{1}{1 + \text{Exp}\{-\beta(C_2 + \delta - C_1)\}}; \text{ entonces } P_2 = 1 - P_1, \text{ y sustituyendo se tiene:}$$

$P_2 = 1 - P_1 = \text{Exp}\{-\beta(C_2 + \delta - C_1)\}$; y se considera la razón entre dos modos nos da:

$$\frac{P_1}{1 - P_1} = \frac{1}{\text{Exp}\{-\beta(C_2 + \delta - C_1)\}}; \text{ obteniendo logaritmo de esta función:}$$

$$\ln\left(\frac{P_1}{1 - P_1}\right) = \beta(C_2 - C_1) + \beta\delta; \text{ que es equivalente a: } Y = ax + b$$

Donde:

$$Y = \ln\left(\frac{P_1}{1 - P_1}\right), \quad a = \beta, \quad X = (C_2 - C_1), \quad \text{y } b = \beta\delta.$$

Los datos que se tienen son datos agregados de selección entre 5 pares de zonas:

Zona	$P_1(\%)$	$P_2(\%)$	C_1	C_2	$\frac{Y}{\ln\left(\frac{P_1}{1-P_1}\right)}$	$\frac{X}{(C_2 - C_1)}$	X^2	XY
1	51	49	21	18	0.04	-3	9	-0.12
2	57	43	15.8	13.1	0.29	-2.7	7.29	-0.78
3	80	20	15.9	14.7	1.39	-1.2	1.44	-1.66
4	71	29	18.2	16.4	0.90	-1.8	3.24	-1.62
5	63	37	11.0	8.5	0.53	-2.5	6.25	-1.325
Totales					3.15	-11.2	27.22	-5.513

Sacando los valores de a y b se tiene que:

$$a = \bar{Y} - b\bar{X}, \quad (1)$$

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}, \quad (2)$$

si $n = 5$; entonces la media de X es: $\bar{X} = \frac{-11.2}{5} = -2.24$,

y la media de Y es: $\bar{Y} = \frac{3.15}{5} = 0.63$.

Resolviendo la ecuación (2) y (1) tenemos:

$$b = \frac{(-5.513) - (5)(-2.24)(0.63)}{(27.22) - (5)(-2.24)^2} = \frac{1.53}{38.42} = 0.0401 = \beta\delta$$

$$a = 0.63 - 0.0401(-2.24) = 0.7196;$$

Como $\beta = a = .72$ y $b = \beta\delta$ entonces tenemos:

$$\delta = \frac{b}{\beta} = \frac{0.0401}{0.72} = 0.055; \text{ por lo tanto el modelo logit binario tendrá los}$$

parámetros:

$$\beta = 0.72$$

$$\delta = 0.055$$

En consecuencia el modelo calibrado queda:

$$P_{ij}^k = \frac{\exp(-\beta C_{ij}^k)}{\exp(-\beta C_{ij}^2) + \exp(-\beta C_{ij}^1)} =$$

$$P_1 = \frac{1}{1 + \text{Exp}\{-\beta(C_2 + \delta - C_1)\}} = \frac{1}{1 + \text{Exp}\{0.72(C_2 + 0.055 - C_1)\}}$$

El método de máxima verosimilitud:

De los procedimientos estadísticos probablemente el más versátil, ya que se puede aplicar en gran cantidad de situaciones es el método de “máxima verosimilitud”, la función de verosimilitud esta dada por la multiplicación de probabilidades de seleccionar la opción realmente elegida por el viajero y queda:

$$L(\beta) = P_{11} \cdot P_{32} \cdot P_{13} \cdot P_{24} \dots etc.$$

La forma general de la función de máxima verosimilitud es:

$$L(\beta) = \prod_{i,j} P_{jq}$$

Donde: j = modo elegido
 q = viajero

Y obteniendo el logaritmo natural de la expresión anterior se tiene:

$\ln L(\beta) = l(\beta) = \sum \ln P_{jq}$; que es la función a maximizar para encontrar el parámetro β .

Ejemplo:

Calibrar el modelo logit multinomial para una muestra de 3 viajeros, cada uno de los cuales, tiene solo 2 alternativas (modos) de viaje, trabajar bajo la suposición que solo se tiene un atributo para el costo de viaje (C_{ij}). Además suponer que las observaciones hechas fueron:

Viajero	Modo elegido	C_{ij}^1	C_{ij}^2
1	1	5	3
2	1	1	2
3	2	3	4

$$P_{ij}^1 = \frac{\exp(-\beta C_{ij}^1)}{\sum_k \exp(-\beta C_{ij}^k)} = P_{ij}^1 = \frac{\exp(-\beta C_{ij}^1)}{\exp(-\beta C_{ij}^1) + \exp(-\beta C_{ij}^2)}$$

$$P_{ij}^2 = 1 - P_{ij}^1$$

Entonces la función de máxima verosimilitud queda:

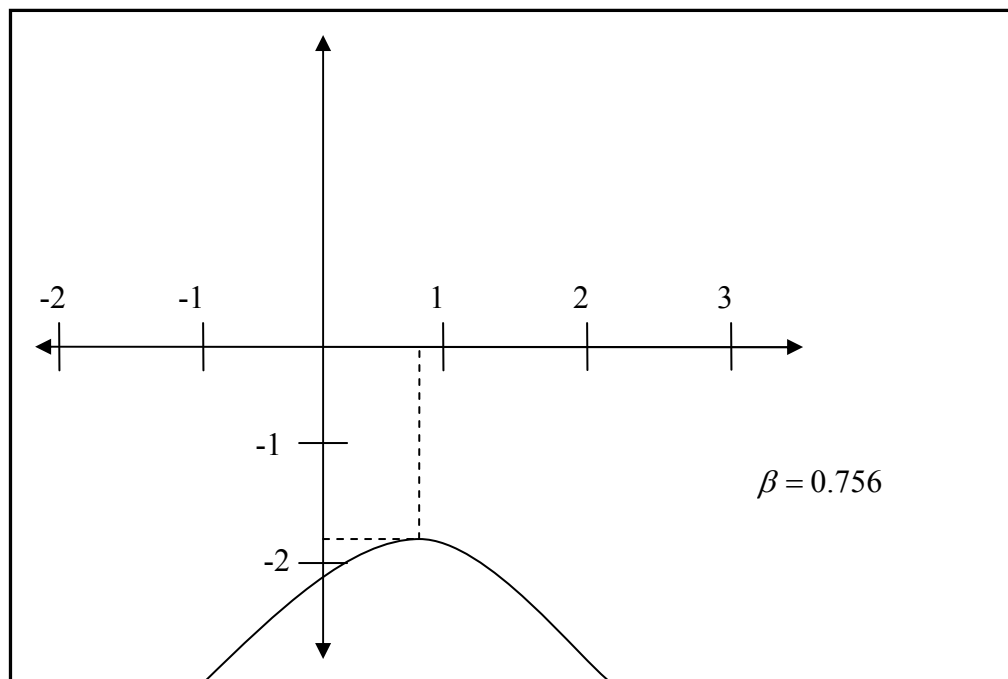
$$L(\beta) = P_{11} \cdot P_{12} \cdot P_{23}$$

$$L(\beta) = \frac{\exp(-\beta 5)}{\exp(-\beta 5) + \exp(-\beta 3)} \cdot \frac{\exp(-\beta)}{\exp(-\beta) + \exp(-\beta 2)} \cdot \frac{\exp(-\beta 3)}{\exp(-\beta 3) + \exp(-\beta 4)};$$

Tomando ln en ambos lados de la ecuación se tiene:

$$\ln(L(\beta)) = l(\beta) = 5\beta - \ln(e^{-5\beta} + e^{-3\beta}) + \beta - \ln(e^{-\beta} + e^{-2\beta}) + 3\beta - \ln(e^{-3\beta} + e^{-4\beta});$$

Se debe graficar $l(\beta)$ para obtener los distintos valores de β , y así se obtiene el máximo de la función:



Con el valor de $\beta = 0.756$ que es el máximo de la función $l(\beta)$, se obtienen las siguientes probabilidades:

Viajero	P_{ij}^1	P_{ij}^2
1	0.82	0.18
2	0.32	0.68
3	0.32	0.68



Capítulo 4

Software

4.1 Software para resolver modelos de planeación del transporte.

Durante los últimos 50 años se ha desarrollado y consolidado una metodología para elaborar modelos para estimar la demanda del transporte, así como de la oferta representada por los sistemas de transporte, para facilitar la resolución de los modelos se ha implementado un sistema, desarrollado en Visual Basic con interfase en Excel, donde se ejecutan los principales modelos logit para la Planeación del Transporte.

El programa se realizó en Visual Basic ya que incluye varias herramientas de documentación, cada una designada para ayudar a aprender y usar un aspecto particular del programa, y al poder alternar su uso con Excel facilita la aplicación y lo hace muy accesible a cualquier usuario, ya que cualquier equipo de cómputo cuenta con esta herramienta.

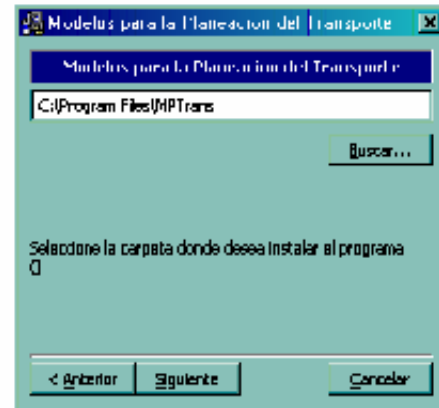
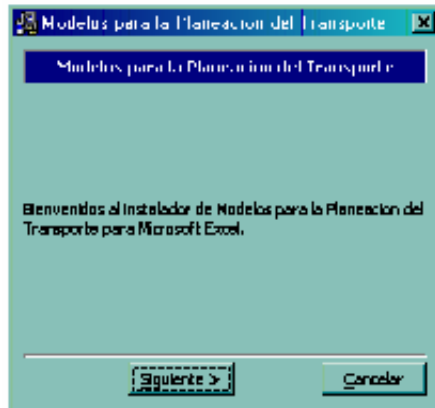
4.2 Introducción al sistema e instalación

El sistema de modelos para la planeación del transporte consta de un complemento para el programa EXCEL, que permite usar los modelos logit para la planeación del transporte a partir de datos almacenados en tablas de Excel.

Antes de instalar el sistema debe asegurarse que el nivel de seguridad de macros en EXCEL sea medio o bajo. En el caso de que el nivel de seguridad sea medio se le preguntará si desea habilitar macros durante el proceso de instalación y al abrir el sistema. Asegúrese de habilitar las macros para que el sistema funcione.

Para instalar el programa se utiliza el programa **INSTALADOR.EXE** que se encarga de copiar todos los archivos requeridos por el sistema al disco duro.

Para abrir el sistema se requiere abrir el archivo **MPtrans.XLA**. Primero se muestra la pantalla de bienvenida, pulse siguiente> para continuar. Se le preguntará la carpeta donde desea que se instale el programa.



Por ultimo seleccione siguiete>. Durante el proceso de instalación se abrirá EXCEL, si tiene el nivel de seguridad en MEDIO asegúrese de habilitar las macros para que se instale el sistema correctamente.

Para activar o desactivar el sistema desde Excel, se utiliza el menú Herramientas —> complementos, donde aparecerá el sistema como una opción dentro del programa.

Al activar el sistema aparecerá dentro de la barra de menús de Excel un nuevo menú llamado **Transporte**, desde el cual pueden utilizarse los diferentes algoritmos para la planeación del transporte. Este menú aparece como se muestra en la figura.



Este menú se encuentra subdividido en 4 categorías que indican cada una de las etapas de la planeación del transporte:

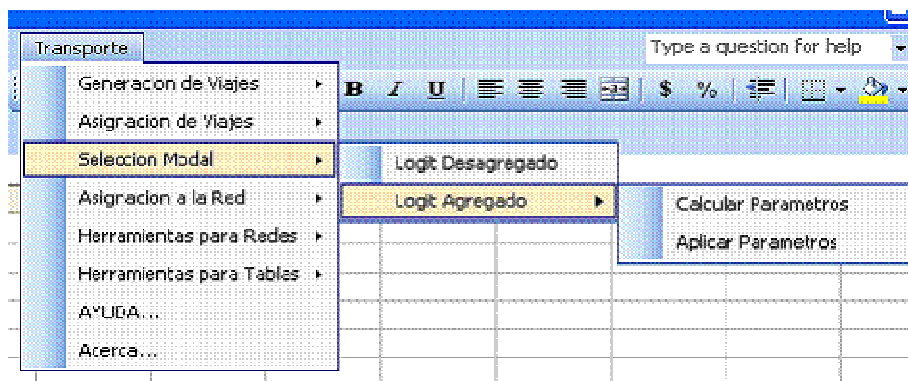
- Generación de Viajes
- Asignación de Viajes
- **Selección Modal**
- Asignación a la Red

Adicionalmente se cuenta con dos categorías de herramientas adicionales que pueden ser de utilidad en el manejo de datos:

- Herramientas para Redes
- Herramientas para Tablas

La parte de interés como ya se había mencionado anteriormente es la de **selección modal**, ya que es la función que nos ayudara a resolver los modelos tipo Logit.

El menú se presenta de la siguiente manera:

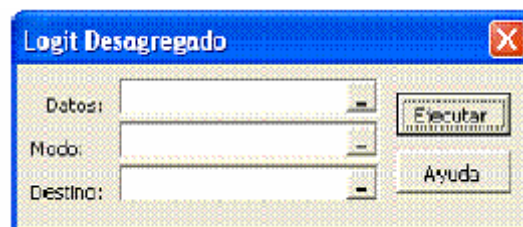


Al elegir la opción de selección modal nos brinda dos opciones para trabajar que son:

- Logit desagregado
- Logit agregado

4.3 Modelo logit desagregado

El modelo Logit desagregado se utiliza cuando se tiene el resultado de encuestas individuales, y se utilizan para obtener las probabilidades de realizar el viaje en cada uno de los modos de transporte, a los que tiene acceso la persona encuestada. Estas probabilidades pueden posteriormente alimentarse al logit agregado, para obtener los factores que indican el peso de cada una de las variables en la selección de modos.



En el campo DATOS se introducen los resultados de cada uno de las encuestas, mientras en el campo MODO se introducen el medio de transporte que utilizo el encuestado, por medio de un numero tomado a partir de cero.

Por ultimo en el campo DESTINO se indica el lugar en que se almacenaran las probabilidades de seleccionar cada uno de los modos de acuerdo a los datos de las encuestas.

Ejemplo

Se tienen 15 encuestas y se desea calcular la probabilidad de que de acuerdo a los resultados de la encuesta un usuario utilizara autobús (0) o automóvil (1) como modo de viaje elegido

Costo Auto	Costo Autobús	Tiempo en Auto	Tiempo en Autobús	Dueño auto	Ingresos	Modo Elegido (*)	Prob. Autobús	Prob. Automóvil
5.1	0.5	7.6	11.4	0	5	0	0.77	0.23
6.5	0.5	9.42	11.4	0	5	0	1.00	0.00
4.5	1	16.57	35.4	0	5	1	0.25	0.75
4.5	0.75	12.56	27.6	1	5	1	0.08	0.92
4.5	0.75	13.16	29.1	1	5	1	0.05	0.95
4.5	1	16.57	35.4	1	15	0	0.08	0.92
4.5	0.75	12.56	27.6	1	15	1	0.10	0.90
4.5	0.75	13.16	29.1	1	15	1	0.07	0.93
4.5	1	16.57	35.4	1	28	1	0.12	0.88
4.5	0.75	12.56	27.6	0	150	0	0.98	0.02
4.7	1	13.98	26.1	1	5	1	0.32	0.68
4.7	0.5	13.94	28.5	1	43	1	0.00	1.00
5.1	0.75	15.86	33.3	1	43	1	0.18	0.82
6.5	1	22.18	53.1	1	5	0	0.99	0.01
6.5	1	22.18	53.1	0	75	0	1.00	0.00

(*) 0- Utilizo Autobús
1- Utilizo Automóvil

Para calcular los viajes producidos se seleccionan los datos de acuerdo a los siguientes colores:

DATOS

MODO

Se selecciona como DESTINO la celda debajo de Prob. Autobuses, por ultimo se selecciona EJECUTAR.

El programa generara las probabilidades de que un usuario con las características de los encuestados escogiera uno u otro medio de transporte.

4.4 Modelo logit agregado

El modelo de logit agregado da como resultado una función que trata de explicar la influencia de diferentes factores en la selección entre diferentes medios de transporte.

Este algoritmo consta de dos partes:

1.- Cálculo de los parámetros de regresión el cual calcula el vector de coeficientes de los parámetros, el cual indica el peso de diferentes factores en la selección de modos. Se requiere un conjunto de parámetros para todos menos el último de los modos existentes.

2.- Aplicación de los parámetros en el cual se aplican los pesos calculados previamente para estimar el nuevo número de viajes de acuerdo a la forma en que se modificaron los factores analizados.

4.4.1 Obtención de parámetros de logit agregado.

Esta ventana se utiliza para obtener los parámetros del modelo Logit agregado. En el campo de DATOS se seleccionan los valores de las variables que se utilizan para estimar las probabilidades de que un usuario utilice diferentes modos de transporte. En el campo MODO se seleccionan las probabilidades de tomar cada uno de los modos de acuerdo a cada uno de los conjuntos de características de usuario. La suma de cada uno de los renglones debe ser igual a 1 para que funcione el algoritmo.



Por último en destino se selecciona la primera celda a partir de la cual se almacenarán los resultados.

Ejemplo.

Se tiene la siguiente tabla de datos, que incluyen el tiempo y el costo de viaje en autobús y automóvil y se desea calcular el peso de cada una de estas variables en la selección de modo (autobús u automóvil).

Tiempo Auto	Costo Auto	Tiempo Autobús	Costo Autobús	Prob. Auto	Prob. Autobús
2.50	1.00	5.00	1.25	0.90	0.10
3.00	1.05	5.00	1.25	0.67	0.33
3.50	1.10	6.00	1.25	0.71	0.29
4.00	1.15	5.50	1.25	0.38	0.63
5.00	1.25	7.50	1.25	0.50	0.50
6.00	0.60	9.00	1.50	0.50	0.50
6.00	0.60	10.00	1.50	0.67	0.33
7.00	0.70	10.00	1.25	0.43	0.57
7.50	1.50	11.00	1.25	0.47	0.53
8.00	0.80	12.50	1.25	0.56	0.44
9.00	1.65	13.50	1.50	0.50	0.50
11.00	1.80	16.00	1.50	0.45	0.55
12.50	1.25	19.00	1.50	0.52	0.48
13.00	1.30	20.00	1.50	0.54	0.46
16.00	1.60	22.50	2.00	0.41	0.59

Para calcular los parámetros se seleccionan los datos de acuerdo a los siguientes colores:

DATOS

MODO

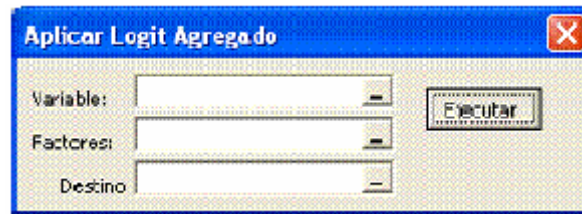
Se selecciona como **DESTINO** la celda donde se deseen almacenar los parámetros y se selecciona EJECUTAR.

El programa generara los parámetros del modelo logit, que posteriormente se pueden utilizar para estimar probabilidades para diferentes usuarios.

	parámetros Auto
CONSTANTE	-1.78760882
Tiempo Auto	-1.3632598
Costo Auto	0.375112027
Tiempo Autobús	0.849788956
Costo Autobús	1.553353692

4.4.2 Aplicación de parámetros del modelo logit agregado

Esta ventana se utiliza para obtener la proporción de viajes por cada modo a partir de los parámetros del modelo Logit agregado. En el campo de VARIABLES se seleccionan los valores de las variables que se utilizan para estimar las probabilidades de que un usuario utilice diferentes modos de transporte. En el campo FACTORES se seleccionan los parámetros previamente calculados del modelo logit.



Por ultimo en destino se selecciona la primera celda a partir de la cual se almacenaran los resultados.

Ejemplo

Se tiene la siguiente tabla de datos, que incluyen el tiempo y el costo de viaje en autobús y automóvil y se han calculado previamente los pesos de cada una de estas variables en la selección de modo (autobús u automóvil). Se desea calcular las probabilidades que se obtienen a través de estos parámetros para compararlas con las probabilidades originales.

Para calcular las probabilidades se seleccionan los datos de acuerdo a los siguientes colores:

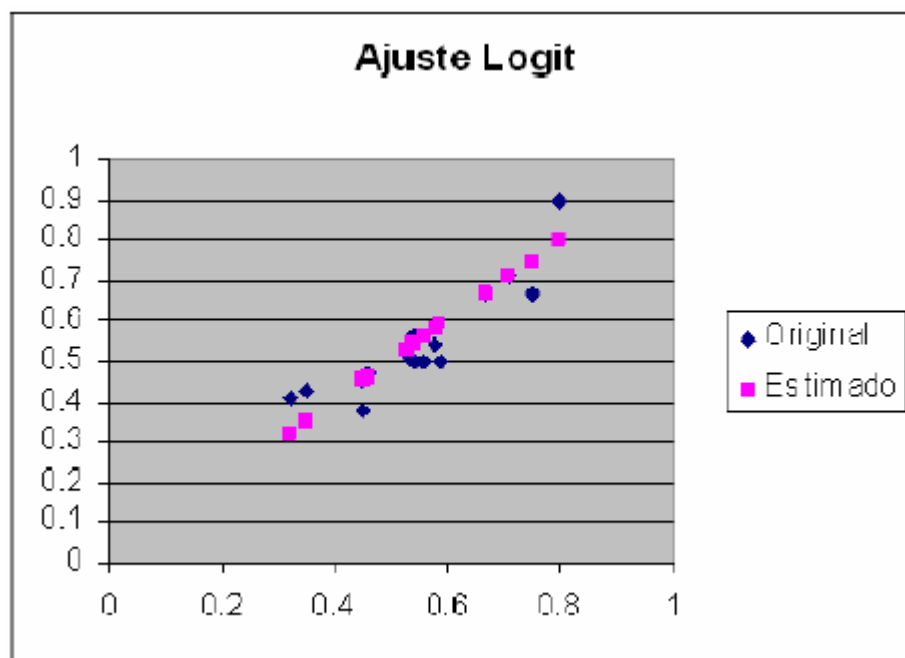
VARIABLE
FACTORES

Se selecciona como DESTINO la celda debajo de Prob. Auto (Estimado), por ultimo se selecciona EJECUTAR.

Tiempo Auto	Costo Auto	Tiempo Autobús	Costo Autobús	Originales		Estimadas	
				Prob. Auto	Prob. Autobús	Prob. Auto	Prob. Autobús
2.50	1.00	5.00	1.25	0.90	0.10	0.80	0.20
3.00	1.05	5.00	1.25	0.67	0.33	0.67	0.33
3.50	1.10	6.00	1.25	0.71	0.29	0.71	0.29
4.00	1.15	5.50	1.25	0.38	0.63	0.45	0.55
5.00	1.25	7.50	1.25	0.50	0.50	0.54	0.46
6.00	0.60	9.00	1.50	0.50	0.50	0.56	0.44
6.00	0.60	10.00	1.50	0.67	0.33	0.75	0.25
7.00	0.70	10.00	1.25	0.43	0.57	0.35	0.65
7.50	1.50	11.00	1.25	0.47	0.53	0.46	0.54
8.00	0.80	12.50	1.25	0.56	0.44	0.54	0.46
9.00	1.65	13.50	1.50	0.50	0.50	0.59	0.41
11.00	1.80	16.00	1.50	0.45	0.55	0.45	0.55
12.50	1.25	19.00	1.50	0.52	0.48	0.53	0.47
13.00	1.30	20.00	1.50	0.54	0.46	0.58	0.42
16.00	1.60	22.50	2.00	0.41	0.59	0.32	0.68

	parámetros Auto
CONSTANTE	-1.78760882
Tiempo Auto	-1.3632598
Costo Auto	0.375112027
Tiempo Autobús	0.849788956
Costo Autobús	1.553353692

El programa generara las probabilidades estimadas a partir del modelo, las cuales se pueden comparar con las probabilidades originales obtenidas de la encuesta. En la grafica se muestra la comparación entre ambas probabilidades.



4.5 Ejemplo de aplicación del software

Consideremos la decisión de un viajero de seleccionar entre los modos: auto particular, metro y autobús; para viajar a su trabajo. La siguiente tabla informa sobre el número de viajes/zona, que se realizan por cada modo.

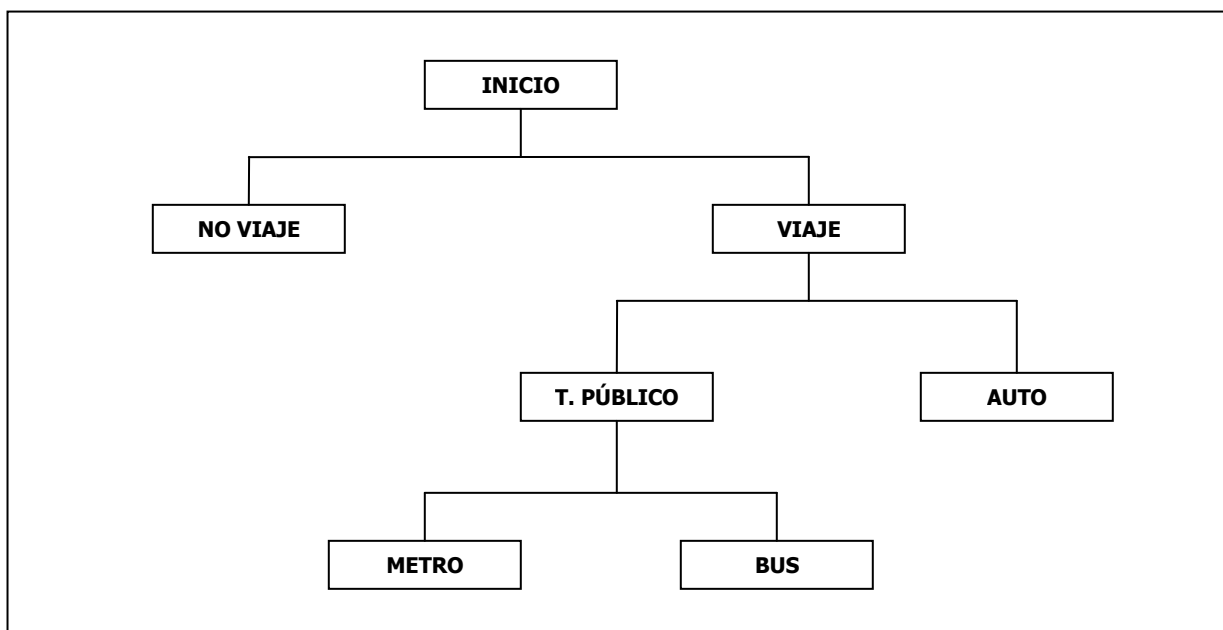
ZONA	BUS	METRO	AUTO	TOTAL
1	3	61	232	296
2	80	137	400	617
3	216	131	301	548
4	268	76	203	547
5	197	50	188	435
6	150	24	164	338
7	91	10	183	284
TOTAL	1005	489	1671	3165

Se desea saber la probabilidad de que los viajeros utilicen los diferentes modos de transporte de su zona.

El proceso del modelo logit es muy atractivo, si se supone que los viajeros toman sus decisiones en forma secuencial.

En términos del ejemplo, el viajero primero decide si realiza o no el viaje, después considera la decisión de usar el modo de transporte público o auto particular, y por último si decide usar el transporte público, debe seleccionar entre viajar en tren o en bus.

De tal forma que el diagrama jerárquico se puede establecer con la siguiente estructura:



El proceso de solución a través de cálculos matemáticos sería de la siguiente forma:

La partición entre las dos últimas opciones, es un modelo logit binario, con probabilidades de selección:

$$P(M/TP) = \frac{\exp(U_M)}{\exp(U_M) + \exp(U_B)}$$

Donde:
$$P(B/TP) = 1 - P(M/TP)$$

En estas utilidades U , solo se deben incorporar elementos que no son comunes con el modo de auto particular.

Para separar entre auto y transporte público, se requiere nuevamente usar el modelo logit binario, con las probabilidades de selección.

$$P(A) = \frac{\exp U_A}{\exp U_A + \exp U_{TP}}$$

Donde la utilidad U_A incorpora todos los atributos del auto particular.

La utilidad del modo transporte público, estará dada por:

$$U_{TP} = \beta(EMU) + \sum_k \alpha_k X_k ;$$

donde, $EMU = \ln(\exp U_M + \exp U_B)$ y la $\sum_k$ considera a todos los elementos comunes dejados fuera en la estimación, del nivel anterior.

Para calcular las probabilidades de cada selección modal, se tendrían que realizar las siguientes operaciones:

$$\begin{aligned} P_A &= P(A) \\ P_M &= P(M/TP) * P(TP) = P(M/TP) * (1 - P_A) \\ P_B &= P(B/TP) * P(TP) = [1 - P(M/TP)] [1 - P_A] \end{aligned}$$

Por último, para asegurar la consistencia interna del modelo, se requeriría que:

$$0 < \beta \leq 1$$

Consideraciones para el valor de β :

Si:

a).- $\beta \leq 0 \Rightarrow$ Aumento de la utilidad de una alternativa (modo), disminuye su probabilidad de selección.

b).- $\beta = 0 \Rightarrow$ Aumento de la utilidad de una alternativa (modo), no altera su probabilidad de selección.

c).- $\beta > 1 \Rightarrow$ Aumento de la utilidad de una alternativa (modo), aumenta su probabilidad de selección.

d).- $\beta = 1 \Rightarrow$ Considera al modelo jerárquico como uno de tipo logit múltiple.

Como el modelo jerárquico puede tener más de dos niveles, para cumplir con el requisito de consistencia interna del modelo, se requiere que:

$$0 < \beta_1 \leq \beta_2 \leq \dots \leq \beta_n \leq 1$$

donde: β_1 = Parámetro del nido o jerarquía más inferior o interna, de cada rama del árbol.

β_n = Parámetro del nido o jerarquía más superior o externa, de cada rama del árbol.

En cambio dándole solución al ejemplo con el software tendríamos:

Los datos que se tienen son por zona, entonces tenemos que aplicar el modelo logit agregado.

ZONA	BUS	METRO	AUTO	TOTAL
1	3	61	232	296
2	80	137	400	617
3	216	131	301	548
4	268	76	203	547
5	197	50	188	435
6	150	24	164	338
7	91	10	183	284
TOTAL	1005	489	1671	3165

1. Trasladar la tabla a Excel para darle solución y calcular los parámetros



Microsoft Excel - Ejemplo logit

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ? Transporte

Escriba una pregunta

Arial 10

C16 -6.15304180124761

ZONA	BUS	METRO	AUTO	TOTAL	PROB.BUS	PROB. METRO	PROB. AUTO	PROB. ESTI MADA BUS	PROB. ESTI MADA METRO	PROB. ESTIMADA AUTO
1	3	61	232	296	0.02	0.2	0.78			
2	80	137	400	617	0.14	0.22	0.64			
3	216	131	301	548	0.39	0.06	0.55			
4	268	76	203	547	0.5	0.13	0.37			
5	197	50	188	435	0.45	0.11	0.44			
6	150	24	164	338	0.44	0.07	0.49			
7	91	10	183	284	0.32	0.03	0.65			
TOTAL	1005	489	1671	3165						

parametros del modelo

-6.1530418	-0.96670198
0.020770053	-0.00205527
-0.04279509	0.019028581
0.022204862	-0.00752295

Parametros Logit Agregado

Datos: Hoja2!\$C\$6:\$E\$12

Modo: Hoja2!\$G\$6:\$I\$12

Destino: Hoja2!\$C\$16

Ejecutar

Ayuda

2. ya que se tiene los parámetros calculados se aplican en el submenú *Aplicar parámetros*.

Microsoft Excel - Ejemplo logit

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ? Transporte

Escriba una pregunta

Arial 10

C16 -6.15304180124761

Logit Desagregado

Logit Agregado

Calcular Parametros

Aplicar Parametros

Generacion de Viajes

Asignacion de Viajes

Seleccion Modal

Asignacion a la Red

Herramientas para Redes

Herramientas para Tablas

AYUDA...

Acerca...

ZONA	BUS	METRO	AUTO	TOTAL	PROB.BUS	PROB. METRO	PROB. AUTO	PROB. ESTI MADA BUS	PROB. ESTI MADA METRO	PROB. ESTIMADA AUTO
1	3	61	232	296	0.02	0.2	0.78			
2	80	137	400	617	0.14	0.22	0.64			
3	216	131	301	548	0.39	0.06	0.55			
4	268	76	203	547	0.5	0.13	0.37			
5	197	50	188	435	0.45	0.11	0.44			
6	150	24	164	338	0.44	0.07	0.49			
7	91	10	183	284	0.32	0.03	0.65			
TOTAL	1005	489	1671	3165						

parametros del modelo

-6.1530418	-0.96670198
0.020770053	-0.00205527
-0.04279509	0.019028581
0.022204862	-0.00752295



Microsoft Excel - Ejemplo logit

Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ? Transporte

Escriba una pregunta

Arial 10

ZONA	BUS	METRO	AUTO	TOTAL	PROB. BUS	PROB. METRO	PROB. AUTO	PROB. ESTI MADA BUS	PROB. ESTI MADA METRO	PROB. ESTI MADA AUTO
1	3	61	232	296	0.02	0.2	0.78	0.02	0.17	0.81
2	80	137	400	617	0.14	0.22	0.64	0.16	0.15	0.69
3	216	131	301	548	0.39	0.06	0.55	0.30	0.16	0.54
4	268	76	203	547	0.5	0.13	0.37	0.62	0.06	0.32
5	197	50	188	435	0.45	0.11	0.44	0.46	0.07	0.47
6	150	24	164	338	0.44	0.07	0.49	0.37	0.07	0.56
7	91	10	183	284	0.32	0.03	0.65	0.33	0.06	0.61
TOTAL	1005	489	1671	3165						

parametros del modelo

-6.1530418	-0.36670198
0.020770053	-0.00205527
-0.04279509	0.019028581
0.022204862	-0.00752295

Aplicar Logit Agregado

Variable: Hoja2!\$C\$6:\$E\$12

Factores: Hoja2!\$C\$16:\$D\$19

Destino: Hoja2!\$J\$6

Ejecutar Ayuda

3. y se obtiene los resultados de las probabilidades de elegir los modos de transporte dados y trasladamos la tabla a Word nuevamente:

ZONA	BUS	METR O	AUT O	TOTA L	PROB. BUS	PROB. METRO	PROB · AUTO	PROB. ESTI MADA BUS	PROB. E STIMAD A METRO	PROB. ESTIM ADA AUTO
1	3	61	232	296	0.02	0.2	0.78	0.02	0.17	0.81
2	80	137	400	617	0.14	0.22	0.64	0.16	0.15	0.69
3	216	131	301	548	0.39	0.06	0.55	0.30	0.16	0.54
4	268	76	203	547	0.5	0.13	0.37	0.62	0.06	0.32
5	197	50	188	435	0.45	0.11	0.44	0.46	0.07	0.47
6	150	24	164	338	0.44	0.07	0.49	0.37	0.07	0.56
7	91	10	183	284	0.32	0.03	0.65	0.33	0.06	0.61
TOTAL	1005	489	1671	3165						

Como se puede observar el tiempo de solución del problema a través del software es muy corto, lo que representa una gran ventaja al tener modelos de este tipo.

CONCLUSIONES

El esfuerzo de predecir la demanda de viajes a través de modelos matemáticos, se ha convertido hoy en día en un proceso totalmente diferente, obedece básicamente a conformar una estrategia para determinar la demanda de viajes a largo plazo y al mismo tiempo se ha sofisticado.

Los modelos de demanda se han aplicado sin considerar que el viajero utiliza los medios de transporte, desde el punto de vista microeconómico, para obtener un bien y satisfacer una necesidad recurrente, omisión grave que ha provocado diferencias significativas en el momento de llevar a la práctica los resultados obtenidos.

Los modelos de demanda basados en la conducta individual, son una herramienta que aparece en fechas recientes, con el propósito de ayudar a implantar las estrategias de predicción de demanda, pero la metodología o procedimientos a seguir resultan un tanto confusos. Es por ello que este trabajo se enfoca a definir los modelos de selección discreta, aplicables a este tipo de problemas ya que estos establecen la probabilidad de que los viajeros elijan una opción dada, considerando sus características socioeconómicas y la atracción relativa de la operación, y lo que determina lo atractivo de la operación es el *concepto de utilidad*, que nos indica que el viajero tomara una opción de acuerdo a la satisfacción o beneficio que le deje dicha selección.

Los modelos de selección discreta de tipo "Logit", se muestran hoy en día como una fuerte opción para ser usados ya que tienen una fuerte ventaja ante los modelos de selección discreta tipo "Probit", esto debido a la bondad que presenta en general el modelo logit para trabajar cuando se tiene una gran cantidad de datos. Además poder contar con una herramienta extra para la resolución de estos al haber implementado un software.

Algo que no está muy claro en la literatura que aborda los modelos de selección discreta tipo logit, es la forma en que estiman los parámetros de calibración, en este trabajo se presentan dos métodos para estimarlos como son: a) el método de regresión lineal y b) el método de máxima verosimilitud.

El método de regresión lineal se aplica cuando se pretende explicar una variable respuesta cuantitativa en función de una variable explicativa cuantitativa, también llamada variable independiente, variable regresora o variable predictora. Por ejemplo, se podría intentar explicar el peso en función de la altura. El modelo intentaría aproximar la variable respuesta mediante una función lineal de la variable explicativa. El método de máxima verosimilitud (en inglés "method of maximum likelihood"). Aunque para aquellos que tienen una formación estadística este método es perfectamente conocido y comprendido, muchos de los usuarios de los programas estadísticos, que están habituados a calcular modelos de



regresión logística, modelos de Poisson, y muchos otros, desconocen cómo se efectúa la estimación de los coeficientes de este modelo.

El software implementado para la solución de los problemas tipo Logit, del cual se habla en este trabajo, abre la posibilidad de contar con una herramienta práctica y de fácil acceso, tanto para el uso didáctico como para el uso comercial.

Finalmente, podemos afirmar que es posible predecir la demanda de viajes en las zonas urbanas, mediante el uso de los modelos de selección discreta en especial el modelo tipo Logit ya que en problemas grandes se presenta como una herramienta importante para los planeadores de transporte.

SUMARIO

Los modelos llamados "Logit Mixtos" han irrumpido con mucha fuerza en el ambiente teórico de la modelación de demanda por transporte en los últimos años. Se trata de una alternativa de modelación que podría situarse entre el modelo Logit y el Probit. Sus promotores claman que tiene la flexibilidad del Probit, manteniendo parte de la simpleza del Logit. En este trabajo se analiza su formulación en detalle, con una óptica de imparcialidad, verificando la consecuencia de sus hipótesis.

En el contexto de la modelación de elecciones discretas en demanda de transporte, el enfoque más utilizado en la actualidad es el basado en la teoría de la utilidad aleatoria. Según esta teoría, cada individuo n tiene una función de utilidad U_{in} asociada a cada una de las alternativas i , escogiendo aquella alternativa que maximiza su utilidad. Esta función, propia del individuo, puede dividirse en una componente sistemática V_{in} , que recoge el efecto de las variables explicativas (atributos medibles u observables por parte del modelador), y una componente aleatoria ξ_{in} que intenta recoger todos aquellos efectos no incluidos en la componente sistemática de la función de utilidad; por ejemplo, la incapacidad del modelador para observar todas las variables que influyen en la decisión, errores de medición, diferencias entre individuos, percepciones incorrectas entre atributos y la aleatoriedad inherente a la naturaleza humana. De los supuestos que se asumen sobre la distribución del error estocástico se derivan los distintos modelos planteados en la literatura.

Los modelos más utilizados en la actualidad son el Logit Multinomial, que se deriva a partir de asumir que los términos de error ξ_{in} e ϵ son IID Gumbel, y el Logit Jerárquico, que se deriva como una extensión del anterior, en que se considera que existe una componente de error adicional y que representa correlación en un grupo de alternativas. En síntesis, se trata de estructuras de covarianza (del término de error) muy simples, lo cual es un supuesto simplificador que no siempre es sostenible, pero que permite tener modelos simples de entender y usar. El modelo Probit, en cambio, se deriva a partir de suponer errores aleatorios con distribución normal multivariada, aceptando en teoría cualquier estructura de error (matriz de covarianza) que los datos permitan estimar, lo cual implica un grado de dificultad de estimación considerable. Este modelo, que aparece como tan deseable desde ese punto de vista, ha sido incorporado tímidamente a la práctica, pese a existir desde hace algún tiempo herramientas poderosas que permiten su estimación mediante simulación.

Es en este contexto que en los últimos años aparecen los modelos Logit mixto (también conocidos como modelos de Error Compuesto o Probit con Kernel Logit), como una alternativa intermedia que se sitúa en algún punto entre el Logit y el Probit. La idea central de este tipo de modelos es considerar más de una componente aleatoria; de esta forma, además de una componente Gumbel IID,



con lo cual el modelo básico es Logit, se agregan otros componentes que permiten modelar correlación y/o heteroscedasticidad. Esto permite ganar generalidad, pero la estimación deja de ser simple como en el caso del modelo Probit, se requiere simulación.

Ya se ha dicho que la distribución del término de error estocástico juega un rol fundamental en los modelos de selección discreta, y que los más utilizados suponen una distribución Gumbel homoscedástica e independiente. Si el punto es incorporar modelos que permitan estructuras de error más generales, es importante analizar qué estructuras sería deseable poder estimar y por qué. Estamos hablando de la posible existencia de correlación y heteroscedasticidad (distinta varianza) en los términos de error. En ambos casos, éstas se pueden dar entre alternativas y entre observaciones. El caso de correlación entre alternativas (presente por ejemplo cuando el usuario percibe algunas alternativas como más similares entre sí que otras) es asimilado bajo ciertas restricciones por el modelo Logit Jerárquico, que permite una matriz de covarianza diagonal por bloques y homoscedástica. Sin embargo, muchos casos de correlación y heteroscedasticidad, fácilmente asociables a situaciones prácticas, no se pueden tratar adecuadamente con los modelos tradicionales. Por ello, resulta interesante encontrar un modelo más general que se adapte a situaciones más sofisticadas.



BIBLIOGRAFÍA

- Ampt, E. y Bonsall P., Current issues in travel and transport demand surveys, 1996.
- Amemiya, T. Advanced Econometrics. Harvard University Press, Cambridge Mass, 1985.
- Anderson, T.W., Introduction to Multivariate Statical Analysis, Wiley, New Cork, 1992.
- Ben-Akiva, ME y Lerman, Discrete Choice Analysis. The MIT Press, Cambridge Mass, 1985.
- Ettema, DF y Timmermans HJP, Activity Based Approaches to Travel Analysis. Elsevier, 1997.
- Fleet y Robertson, Trip generation in the transportation planning process, 1968.
- Kraft, G, Demand for intercity passenger travel in the Washington –Boston corridor project, departamento de comercio U.S., 1969.
- Hartgen y Tanner, Behavioral model of mode choice, reporte preliminar, Albany, N.Y. 1972.
- Jara Díaz, General Micromodel of Users' Behavior: Basic Issues, In Travel Behavior Research: Updating the State of the Play, J.D. Ortúzar, D. Hensher and S. Jara-Díaz (eds), 19 – 33, 1998.
- Jara-Díaz, S y Martínez F. On the specification of Indirect Utility and Willingness to pay for discrete residential location models. Journal of Regional Science 39 (4), 675-688, 1999.
- Garling, T, Laitila, T y Westin, K Theoretical Foundations of Travel Choice Modelling Elsevier, 1998.
- Manheim, Fundamental prospectus of system of demand models , MIT, 1970.
- Michael D. Meyer, Eric J. Miller, Urban Transportation Planning second edition, Mc Graw Hill, 2001.
- MIDEPLAN, Metodología Revisada para Análisis de Transporte en Ciudades de Tamaño Medio. Comisión de Planificación de Inversiones en Infraestructura de Transporte, 1997.



Munizaga, M y Álvarez R, Modelos Mixed Logit: uso y potencialidades. Actas del XI Congreso Chileno de Ingeniería de Tránsito y Transporte, 2000.

Ortúzar, J de D, Modelos de Demanda de Transporte. Ediciones Universidad Católica, Santiago, 1994.

Ortúzar, J de D Modelos Econométricos de Selección Discreta. Ediciones Universidad Católica, Santiago, 2000.

Ortúzar, J de D y Daly, A, Agregación de datos y predicción de demanda. Apuntes de Ingeniería, 39, 51-72, 1990.

Ortúzar, J de D y Willumsen, Modelling Transport, John Wiley & Sons, Chichester, 1990.

Ortúzar, J y otros. “Encuesta Origen Destino de Viajes del Gran Santiago: Diseño y principales resultados”, 1992.

Ortúzar, J de D y Willumsen, Modelling Transport third edition,. John Wiley & Sons, Chichester, 2001.

Pindyck, R. y Rubinfeld, D., Modelos econométricos, Editorial Labor S.A. Barcelona, 1980.

Lisco, The value of commuters travel time: A study in urban transportation, Departamento de Economía, Universidad de Chicago, 1967

Stopher, P y Lee-Goselin, Understanding Travel Behaviour in an Era of Change, 1996.

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

Thomas A. Domencich and Daniel McFadden, Urban Travel Demand A behavioral analysis, North-Holland Publishing Company, Amsterdam, 1975.

Wilson, A, Urban and regional Models in Geography and Planning. John Wiley & Sons, London, 1974.

Wilson, A y Bennet, Mathematical Methods in Human geography and Planning, John Wiley Wiley & Sons, London, 1985.

Williams, H.C.W.L., On the formation of travel demand models and economic evaluation measures of user benefit. Environment and Planning 9A(3), 285-344, 1977.