PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA DE TRANSPORTE Y LOGÍSTICA
ICT 2213 – Modelos de Demanda de Transporte
Profesor: Sebastián Raveau

Guía Capítulo 2

1- En el contexto del modelo clásico de transporte de cuatro etapas, explique por qué se dice que las etapas de generación y atracción de viajes suelen ser insensibles a los niveles de servicio en la red. Explique también cómo un modelo de demanda basado en actividades podría corregir dicha limitación.

R: Una respuesta correcta debe referirse a las variables utilizadas en los modelos tradicionales de generación y atracción de viajes y, por lo tanto, de dónde surge la falta de sensibilidad a niveles de servicio. Se deben contrastar estas características con las de un modelo basado en actividades.

2- ¿Por qué un viaje del Trabajo al Supermercado es basado en el Trabajo, pero un viaje del Trabajo al Hogar es basado en el Hogar, si ambos viajes comienzan en el Trabajo? ¿Qué implicancia tiene esta definición en la manera en que se modelan ambos viajes?

R: Una respuesta adecuada debe explicar la categorización de viajes discutida en clases, con foco en el propósito y el origen/destino de los viajes. Adicionalmente se debe explicar la modelación de viajes de ida y vuelta de un tour, según corresponda.

3- Suponga que quiere modelar la generación de viajes con motivos poco frecuentes a nivel zonal, para lo cual recurre a un método de Factor de Crecimiento. Explique claramente cómo obtendría el factor de crecimiento de cada zona, mencionando los pasos a seguir y la información requerida.

R: Se debe explicar el proceso para obtener la expresión analítica de los factores, los datos utilizados, el proceso de ajuste matemático (e.g., obtener los parámetros involucrados) y su uso para predecir a futuro.

4- Explique qué entiende por Zona Cero. Indique cómo identificar una Zona Cero al momento de modelar: i) Generación de viajes a nivel zonal y ii) Atracción de viajes a nivel zonal.

R: La respuesta debe definir claramente en qué consiste una zona cero y cómo pueden identificarse en ambos contextos de modelación (con foco en las características y variables relevantes para generación y para atracción de viajes).

5- En base a todo lo que ha aprendido hasta el momento, explique por qué al momento de predecir y evaluar proyectos de transporte se suele confiar más en los modelos de generación de viajes que en los modelos de atracción de viajes, aun cuando los modelos de atracción de viajes suelen tener mejor ajuste estadístico (e.g. mayor R^2).

R: La respuesta debe hacer referencia al nivel de agregación de ambos modelos y al ajuste resultante en cada caso, enfatizando claramente cómo dichos niveles de agregación implican mayor o menor confianza en los resultados (dada la profundidad con que se modela cada decisión de movilidad).

6- Suponga que para la Tarea 2 ha estimado varios modelos alternativos de generación de viajes. Describa claramente la secuencia de pasos que seguiría para elegir el mejor modelo.

R: Una respuesta correcta debe referirse a los diferentes enfoques de validación, análisis y comparación de modelos, desde perspectivas teóricas y estadísticas, identificando claramente el orden (o prioridad) que se le da a los diferentes enfoques.

7- Indique dos criterios que debiese cumplir una correcta categorización de viajes mediante Árboles de Regresión, para que los promedios obtenidos sean confiables.

R: La respuesta debe referirse a una cantidad mínima de observaciones en cada nodo hoja y a la similitud (e.g., baja varianza) de las observaciones en cada nodo hoja.

8- Usted desea estudiar la atracción de viajes por motivo de salud en la punta mañana en Santiago. Para esto cuenta con datos provenientes de una Encuesta Origen-Destino.

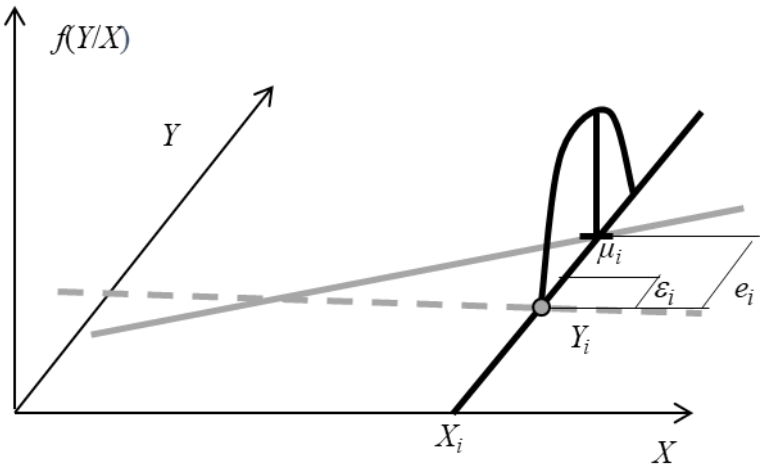
a) ¿En qué contexto(s) preferiría utilizar un enfoque de factores de crecimiento en vez de Regresión Lineal Múltiple para modelar estos viajes?

b) ¿Qué nivel de agregación y variables usaría en un modelo de Regresión Lineal Múltiple para modelar estos viajes?

R: a) Se debe indicar claramente cuándo (i.e., para qué tipos de viaje) el enfoque de factores de crecimiento entregaría mejores resultados que regresión lineal múltiple.

b) Se debe indicar claramente el nivel de agregación tradicional y las variables usualmente utilizadas para modelar atracción de viajes.

9- Considere la representación gráfica de un modelo de Regresión Lineal, vista en clases.



a) Explique la diferencia entre los errores e_i y los errores ε_i.

b) ¿Qué supuestos sobre la distribución de dichos errores es necesario realizar para estimar el modelo de Regresión Lineal mediante Mínimos Cuadrados Ordinarios?

R: a) Es necesario definir claramente en qué consiste y de dónde surge cada tipo de error, explicitando la diferencia entre ellos.

b) Es necesario indicar los supuestos sobre la distribución para la estimación del modelo, asociados a los supuestos II y III de MCO (notar que no se solicitan supuestos para la inferencia estadística de los resultados, asociados al supuesto IV).

10- Suponga que estima un modelo de atracción de viajes mediante Regresión Lineal por Mínimos Cuadrados Ordinarios. ¿Qué haría si el intercepto del modelo es negativo y estadísticamente significativo? ¿Qué haría si el intercepto del modelo es positivo pero no estadísticamente significativo?

R: La respuesta debe caracterizar los potenciales problemas asociados a cada situación, y en caso de requerir una corrección, indicar cómo se llevaría a cabo dicha corrección.

11- Usted desea estimar un modelo de generación de viajes a nivel zonal mediante Mínimos Cuadrados Ordinarios. La especificación del modelo es de la forma:

$$Y_i = \beta_0 + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \varepsilon_i$$

Explique el efecto que tendría en la estimación del modelo cada uno de los siguientes escenarios:

a) Colinealidad perfecta entre X_{1i} y X_{2i}

b) Correlación fuerte entre X_{1i} y X_{2i}

c) Correlación fuerte entre X_{1i} y ε_i

R: a) Responder en referencia al supuesto I de MCO.

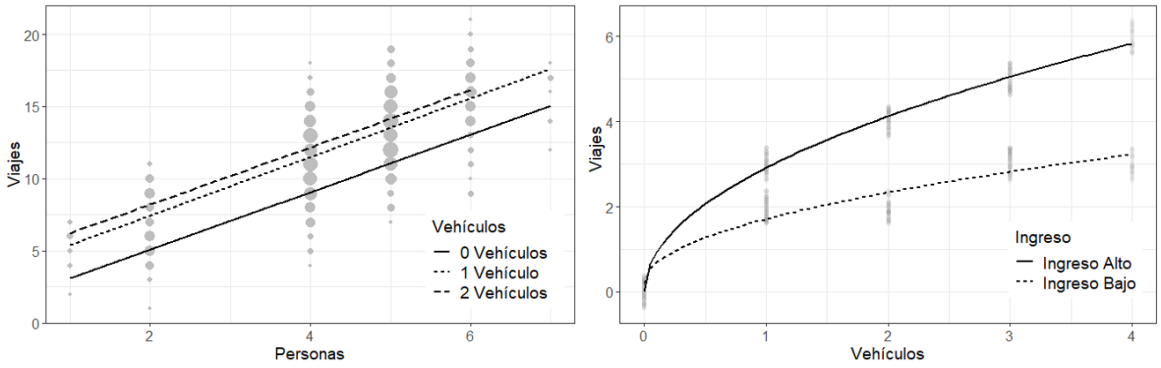
b) Responder en referencia al supuesto I de MCO.

c) Responder en referencia al supuesto II de MCO.

12- Suponga que estima un modelo de generación de viajes a nivel de hogares mediante Regresión Lineal por Mínimos Cuadrados Ordinarios, obteniendo resultados estadísticos satisfactorios (en términos de signos, significancia estadística y ajuste). ¿Cómo comprobaría usted que el modelo no posee problemas de endogeneidad?

R: Una respuesta correcta debe demostrar que se entiende en qué consiste la endogeneidad e indicar métodos visuales y/o estadísticos para analizar la potencial existencia de endogeneidad.

13- Considere los siguientes gráficos de generación de viajes a nivel de hogar para un día laboral. ¿Qué puede decir del impacto que tiene la cantidad de personas y la cantidad de vehículos en la cantidad de viajes generados? ¿Qué puede decir del impacto que tiene el ingreso en el impacto que tienen los vehículos en la cantidad de viajes generados? En base a su análisis, proponga una especificación para el modelo de generación de viajes.



R: En primer lugar, la respuesta debe referirse a los impactos lineales o no lineales, según corresponda, de las variables de personas y vehículos. En segundo lugar, la respuesta debe referirse al potencial efecto conjunto (i.e., efecto doble producto de una interacción) de ingreso y vehículos. Una potencial especificación puede ser:

$$Y_i = \beta_0 + \beta_1 \cdot Personas_i + (\beta_2 + \beta_3 \cdot IngresoAlto) \cdot \sqrt{Vehículos}$$

14- Usted desea aplicar un modelo de Análisis de Clasificación Múltiple para predecir la generación de viajes de una ciudad, con una zonificación dada, a nivel de hogares. El modelo considera cuatro variables: ingreso (3 niveles), tasa de motorización (3 niveles), cantidad de trabajadores (2 niveles) y cantidad de estudiantes (2 niveles).

a) Cómo verificaría que la categorización considerada es adecuada?

b) ¿Cuántos efectos principales, dobles, triples y cuádruples puede tener como máximo el modelo?

R: a) Referirse a la cantidad de observaciones por categoría.

b) 6 principales, 13 dobles, 12 triples y 4 cuádruples.

15- Usted cuenta con los siguientes modelos de generación de viajes a nivel de hogares. Bajo cada parámetro se indica su test-*t*.

Modelo 1 $Y = 1,7 + 0,85 \cdot X_4 + 0,7 \cdot X_3$ $R^2 = 0,395$

$(5,6)$ $(6,6)$ $(2,0)$

Modelo 2 $Y = -0,12 + 0,45 \cdot X_1 + 0,07 \cdot X_2 + 0,6 \cdot X_3$ $R^2 = 0,450$

$(2,1)$ $(1,9)$ $(2,5)$ $(1,9)$

Modelo 3 $Y = 1,3 + 1,75 \cdot X_1 + 0,36 \cdot X_3 - 1,92 \cdot X_4$ $R^2 = 0,495$

$(-2,8)$ $(8,5)$ $(6,3)$ $(-3,9)$

Las variables son: Y son los viajes por hogar en hora punta, X_1 es el número de empleados en el hogar, X_2 es el número de estudiantes en el hogar, X_3 es el número de autos en el hogar, X_4 es el número de residentes en el hogar. Señale qué modelo le parece más adecuado y por qué.

R: El modelo 3 se descarta por signos. Los modelos 1 y 2 se diferencian en el impacto de empleados y estudiantes en la cantidad de viajes. Los parámetros del modelo 2 parecen validar que empleados y estudiantes contribuyen de manera distinta al total de viajes, los parámetros son estadísticamente distintos entre sí, por lo que se prefiere el modelo 2.

16- Usted tiene los siguientes dos modelos de generación de viajes a nivel de hogar, estimados utilizando la misma base de datos:

Modelo 1 $$Y_i = \beta_0 + \beta_1 \cdot P_i + \beta_2 \cdot TM_i + \beta_3 \cdot LA_i$$

Modelo 2 $$Y_i = \theta_0 + (\theta_1 + \delta_1 \cdot LA_i) \cdot P_i + \theta_2 \cdot \sqrt{TM_i}$$

Donde Y_i son viajes por hogar *i*, P_i es el número de personas en el hogar *i*, LA_i toma el valor de 1 si el hogar *i* pertenece a la categoría de ingreso alto, y TM_i es la cantidad de vehículos en el hogar *i*. Los resultados de la estimación de ambos modelos son:

Modelo 1		
$R^2 = 0{,}72$		
$\hat{\beta}_0 = 0{,}53$	$\sigma^2_{\hat{\beta}_0} = 0{,}08$	$t_{\hat{\beta}_0} = 6{,}52$
$\hat{\beta}_1 = 2{,}88$	$\sigma^2_{\hat{\beta}_1} = 0{,}62$	$t_{\hat{\beta}_1} = 4{,}64$
$\hat{\beta}_2 = 0{,}56$	$\sigma^2_{\hat{\beta}_2} = 0{,}29$	$t_{\hat{\beta}_2} = 1{,}90$
$\hat{\beta}_3 = 0{,}33$	$\sigma^2_{\hat{\beta}_3} = 0{,}17$	$t_{\hat{\beta}_3} = 1{,}98$

Modelo 2		
$R^2 = 0{,}76$		
$\hat{\theta}_0 = -0{,}21$	$\sigma^2_{\hat{\theta}_0} = 0{,}23$	$t_{\hat{\theta}_0} = -0{,}91$
$\hat{\theta}_1 = 2{,}54$	$\sigma^2_{\hat{\theta}_1} = 0{,}29$	$t_{\hat{\theta}_1} = 8{,}79$
$\hat{\theta}_2 = 0{,}74$	$\sigma^2_{\hat{\theta}_2} = 0{,}34$	$t_{\hat{\theta}_2} = 2{,}21$
$\delta_1 = 0{,}86$	$\sigma^2_{\delta_1} = 0{,}40$	$t_{\delta_1} = 2{,}17$

a) ¿Qué hipótesis podrían fundamentar una especificación del modelo 2?

b) ¿Cuáles son los signos esperados para los parámetros de los modelos?

c) ¿Cuál de los dos modelos elegiría usted? Justifique claramente su elección. Haga los supuestos que estime convenientes.

d) Un colega le recomienda especificar la cantidad de vehículos utilizando ln(*TM*). ¿Le parece una buena idea? Justifique.

R: a) Referirse a los impactos lineales o no lineales de las variables explicativas.

b) Referirse a los signos que debiesen tener los diferentes parámetros, incluyendo el intercepto.

c) Tras analizar signos, significancias, potenciales problemas de signo al predecir, nivel de ajuste y validez teórica, se prefiere el modelo 2.

d) Referirse a por qué es, numéricamente, una muy mala idea usar una transformación logarítmica para una variable que puede valer 0.