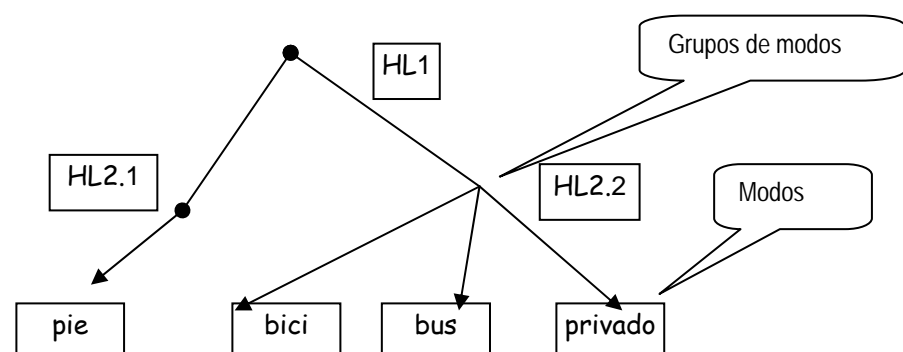


Anexo. Modelos de demanda: Reparto Modal

Introducción

La **prognosis de la demanda** a corto-medio plazo (2008 y 2015) requiere incorporar el reparto modal en el transporte privado-público en escenarios con el tranvía: la estimación de la captación OD del transporte público (tranvía más autobús urbano) a partir de la movilidad total y a su vez determinar la matriz OD en vehículo privado para cada escenario. El reparto de la cuota de mercado entre el autobús y el tranvía en el transporte público se determinará a partir del modelo de asignación transit del entorno TransCAD (ligado automáticamente al modelo de la red del entorno).

Los modelos agregados de reparto modal se van a adaptar para su utilización en los modelos MACRO de transporte privado y público en las redes TransCAD. Se han tanteado distintas alternativas de modelización del reparto modal, aunque se acabe construyendo un modelo *nested logit* o *logit* jerárquico, a veces denominado *logit* anidado, que se estimará con el método VGLM de R dentro del paquete adicional VGAM.



Para facilitar la tarea de la prognosis, se han elaborado indicadores a partir de la grabación de la encuesta de movilidad 2006 de APPEND, los valores agregados de número de viajes por modo principal a nivel de pareja OD, de motivo principal (4) y de segmento de población (4). La estimación de los modelos se ha efectuado con un subconjunto consistente de observaciones recogidas en la grabación de la Encuesta de Movilidad 2006, concretamente se han filtrado los registros según las siguientes restricciones:

- Se consideraran viajes INTERIOR-INTERIOR (entre zonas de transporte de 1 a 38). Sólo se han considerado desplazamientos con etapas en modos: a pie, en bicicleta, en autobús urbano de TUVISA y en coche (conductor y acompañante).
- Se eliminarán todos los viajes con origen y final en la misma zona. Los viajes por modo principal se han de codificar en formato entero y corresponden a la expansión poblacional, no a la muestra.
- Los desplazamientos se han procesado para estudiar las cadenas modales y asignarles un **MODO PRINCIPAL**. La lógica ha sido asignar el modo de mayor jerarquía dentro de la ordenación *a pie < bici < autobús < coche*.

- El **MOTIVO PRINCIPAL** del viaje se ha obtenido del procesado de la grabación. Los casos donde se habían mezclado desplazamientos por varios motivos dentro de un único registro han sido suprimidos, así como los de duración excesiva en muchos casos debidos a la codificación de desplazamientos de ida y vuelta dentro de un mismo registro de viaje.

| MOTIVO APPEND | A | Codificación M OTIVO PRINCIPAL |
|--------------------------|----|--------------------------------|
| 1. Casa | 1 | - |
| 2. Trabajo | 2 | 1 |
| 3. Estudios | 3 | 1 |
| 4. Compras | 4 | 3 |
| 5. Médico/hospital | 5 | 2 |
| 6. Visita amigo/familiar | 6 | 4 |
| 7. acompañar personas | 7 | 2 |
| 8. Gestiones de trabajo | 8 | 1 |
| 9. Gestiones personales | 9 | 2 |
| 10. Ocio, diversión | 10 | 4 |
| 11. Comida/cena | 11 | 4 |
| 12. Sin destino fijo | 12 | 4 |
| 13. 2ª residencia | 13 | 4 |
| 14. Ir a parada | 14 | - |
| 99. No indica | 99 | - |

| MOTIVO PRINCIPAL | CODIFICACIÓN NUMERICA |
|------------------|-----------------------|
| Obligada | 1 |
| Gestiones | 2 |
| Compras | 3 |
| Ocio | 4 |

| Modo APPEND | Codificación MODO PRINCIPAL |
|---------------------------|-----------------------------|
| 1. A pie | 1 |
| 2. Autobús | Eliminar |
| 3. Urbano/TUVISA | 3 |
| 4. Bus interurbano | Eliminar |
| 5. Ferrocarril cercanías. | Eliminar |
| 6. RENFE | Eliminar |
| 7. BUS EMPRESA | Eliminar |
| 8. BUS ESCOLAR | Eliminar |
| 9. AUTOCAR | Eliminar |
| 10. Taxi | Eliminar |
| 11. Coche conductor | 4 |
| 12. Coche acompañante | 4 |
| 13. Moto conductor | 4 |
| 14. Moto acompañante | 4 |
| 15. Bicicleta | 2 |
| 16. Camión/Furgoneta | Eliminar |
| 17. Bus no indicia | Eliminar |
| 18. Otros | Eliminar |

| MODO PRINCIPAL | CODIFICACIÓN NUMERICA |
|----------------|-----------------------|
| Pie | 1 |
| Bici | 2 |
| Bus | 3 |
| Auto | 4 |

| PERFIL DE INDIVIDUO Porcentaje | Codificación SEGMENTO para reparto modal |
|--|--|
| Menor de 16 años 8,6% | 1 |
| Hombre de 16 a 44 años activo con estudios secundarios 11,7% | 2 |
| Mujer de 16 a 44 años activa con estudios secundarios 13,0% | 2 |
| De 16 a 44 años activo/a con estudios primarios 4,2% | 2 |
| De 44 y más años activo/a con estudios secundarios 11,0% | 2 |
| De 44 y más años activo/a con estudios primarios 8,3% | 2 |
| De 16 a 64 años inactivo/a 11,8% | 3 |
| De 16 a 64 años inactivo/a con estudios secundarios 9,6% | 3 |
| Mayor de 64 años 17,3% | 4 |
| Sin clasificar 4,5% | |

Clasificación por segmentos:

| SEGMENTO | CODIFICACIÓN NUMERICA |
|------------------------|--------------------------|
| Menor de 16 años 8,6% | 1 |
| Entre 16-64 activo | 2 |
| Entre 16-64 inactivo | 3 |
| Mayor de 64 años 17,3% | 4 |

Síntesis de columnas:

- *i* : pareja OD de 1 a *I*
- *ori*: zona interior origen de los viajes
- *desti*: zona interior destino de los viajes
- *gi1* hasta *giJ* : J campos con nº de viajes por la pareja *i* en modo *j*
- *X11* hasta *X1J* : J campos con variables de alternativa *X1*
- ...
- *Xk1* hasta *XkJ*: J campos con variables de alternativa *Xk*
- **Motivo**: L variables dummies cada una con 1 o 0 según el registro corresponda al motivo indicado. Tendrán por Número *dmot1*, *dmot2*, ..., *dmotL*. Adicionalmente una columna llamada *motivo* contiene el número correspondiente al motivo.
- **Segmento**: G variables dummies cada una con 1 o 0 según corresponda al segmento.

| | |
|-----------------|--|
| od | Identificador de la pareja OD del desplazamiento o viaje (no etapa) |
| ori | Origen del viaje |
| desti | Destino del viaje |
| g1 | Número viajes modo 1 para el motivo y segmento |
| g2 | Número viajes modo 2 para el motivo y segmento |
| ... | |
| gJ | Número viajes modo J para el motivo y segmento |
| motivo | Motivo del desplazamiento |
| dista | Distancia Manhattan en km entre origen y destino . Formato con 2 decimales |
| segment | Segmento socioeconómico en que se clasifican los individuos vinculados al presente registro |
| dmot1 | 1 si motivo =1 o 0 sinó |
| ... | |
| dmotL | 1 si motivo =L o 0 sinó |
| dseg1 | 1 si segmento =1 o 0 sinó |
| ... | |
| dsegG | 1 si segmento =G o 0 sinó |
| ttinveh1 | Tiempo de viaje OD ponderado para el modo 1 in-vehicle (vem o piu 0) en minutos con dos decimales |
| ... | |
| ttinvehJ | Tiempo de viaje OD ponderado para el modo J in-vehicle ... |
| ttt1 | Tiempo total de viaje OD ponderado para el modo 1 (des d'inici 1a Etapa hasta final darrera Etapa) en minutos con dos decimales |
| ... | |
| tttJ | Tiempo total de viaje OD ponderado para el modo J ... |
| tespera1 | Tiempo de espera OD ponderado para el modo 1 (vem o bus o coche 0) en minutos con dos decimales . En bus està lligat al 50% de la freqüència en minutos (dos decimales) |
| ... | |
| tesperaJ | Tiempo de espera OD ponderado para el modo J ... |
| tacces1 | Tiempo de acceso hasta el modo OD ponderado para el modo 1 (a pie 0) en minutos con dos decimales . En viajes donde hay una etapa de bus esta vinculado al tiempo de las etapas a pie en minutos (dos decimales) . En viajes donde hay una etapa de coche està vinculado al tiempo de acceso (ida o vuelta) al lugar de aparcamiento y búsqueda de aparcamiento. |
| ... | |
| taccesJ | Tiempo de acceso hasta el modo OD ponderado para el modo J ... |
| ppark1 | Porcentaje del aparcamiento de pago para el modo 1 (tiene sentido en auto en resto és 0.00), |
| ... | |
| pparkJ | Porcentaje del aparcamiento de pago para el modo J ... |

Las variables OD por modo principal consideradas en el presente estudio son *a priori*: *ttinveh* (tiempo de viaje en el vehículo), *ttt* (tiempo total de viaje), *tespera* (tiempo de espera en transporte público), *tacces* (tiempo de acceso a una parada de bus o al lugar de aparcamiento), *ppark* (porcentaje de tipo aparcamiento en destino) y *tcoste* (coste total del viaje). Adicionalmente, se dispone de las variables socioeconómicas (comunes por tanto a todas las alternativas de) : segmento y motivo principal del desplazamiento.

Especificación archivo de datos para calibrar el reparto modal con la movilidad actual

Notación anglosajona: decimales con separador punto y sin separador-indicador de miles o millones. Archivo ASCII con primera fila con los nombres de las variables: cortos, sin mayúsculas ni acentos. Campos separados por un blanco (NO tabuladores) o bien formato .csv. Archivo con filas: ordenado por zonas de transporte, donde se consideran únicamente el ámbito interior del estudio.

Codificación indexada de las columnas: índice por modos *j* nº de modo (gran modo) numerados de 1 a *J* (*J*=4) y *I* índice por motivo numerado de 1 a *L* (*L*=4), ordenado el archivo por pareja origen-destino (*i,j*) .

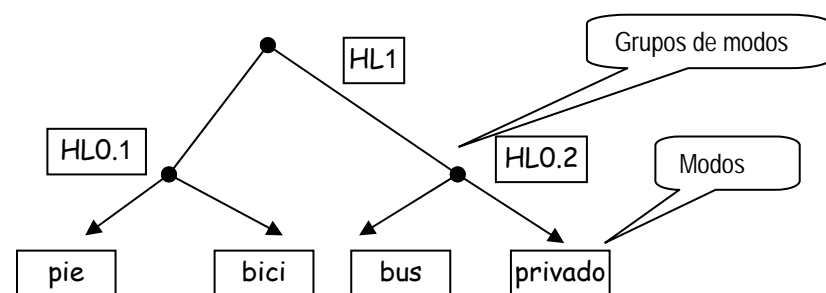
Las variables OD de alternativa modal son: *ttinveh*, *ttt*, *tespera*, *tacces*, *ppark*, *tcoste* (*X1* a *X6*). La presentación de los datos es agregada y por tanto los valores han de corresponder a las medias ponderadas (según coeficientes de expansión de la Encuesta de Movilidad 2006) para cada OD-motivo-segmento.

| | |
|---------------|--|
| tcost1 | Coste del viaje en modo 1 . A pie 0... vem o bus precio 1 viaje en bonus tipo T10 . Coche considera el gasto del combustible más mantenimiento aproximadamente 0,2€/km |
| ... | |
| tcostJ | Coste del viaje en modo J ... |

Las variables explicativas individuales o por alternativa se tendrán que elaborar para el horizonte de estudio 2015 para poder establecer las matrices OD modales.

Modelo de Reparto Modal: Tanteo de Modelos

El modelo final propuesto responde a un modelo *logit* jerárquico cuya estructura se muestra a continuación.



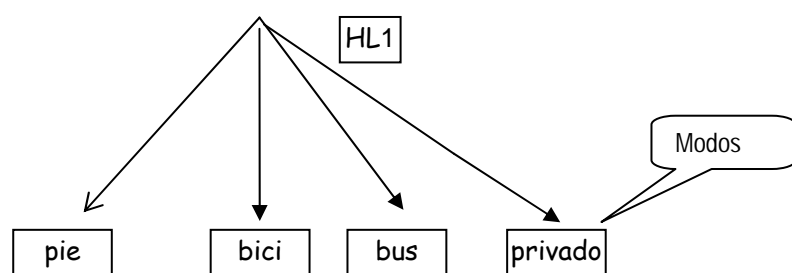
R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Los modelos *logit* condicionales incluyen exclusivamente variables de alternativa (explicativas de características de los modos de transporte en competencia) y se han estimado a partir del paquete VGAM de R:

Thomas W. Yee (2007). VGAM: Vector Generalized Linear and Additive Models. R package version 0.7-3. <http://www.stat.auckland.ac.nz/~yee/VGAM>

Modelo Logit Multinomial Condicional

Situación: hipótesis de independencia entre alternativas irrelevantes que conlleva un Modelo *Logit* Multinomial homocedástico.



El paquete estadístico empleado para el análisis de datos es de libre distribución y gran reputación en el mundo académico.

La estadística descriptiva sintética de las variables caracterizadas a partir de la Encuesta de APPEND para la calibración de un modelo de reparto modal en la situación actual se resumen en:

```

> summary(rmod)
  od          ori          desti          g1
Min.   : 102.0   Min.   : 1.00   Min.   : 1.00   Min.   : 0.00
1st Qu.: 729.8   1st Qu.: 7.00   1st Qu.: 8.00   1st Qu.: 0.00
Median :1634.5   Median :16.00   Median :17.00   Median : 35.88
Mean   :1759.2   Mean   :17.42   Mean   :17.54   Mean   : 51.66
3rd Qu.:2713.0   3rd Qu.:27.00   3rd Qu.:27.00   3rd Qu.: 70.82
Max.   :4338.0   Max.   :43.00   Max.   :43.00   Max.   :889.48

  g2          g3          g4          motiu
Min.   : 0.000   Min.   : 0.000   Min.   : 0.00   Min.   :1.000
1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.: 0.00   1st Qu.:1.000
Median : 0.000   Median : 0.000   Median : 0.00   Median :2.000
Mean   : 3.658   Mean   : 8.946   Mean   : 27.88   Mean   :2.457
3rd Qu.: 0.000   3rd Qu.: 0.000   3rd Qu.: 47.70   3rd Qu.:4.000
Max.   :333.450   Max.   :285.470   Max.   :596.82   Max.   :4.000

  dista          segment          dmot1          dmot2
Min.   :0.340   Min.   :1.000   Min.   :0.0000   Min.   :0.0000
1st Qu.:1.190   1st Qu.:2.000   1st Qu.:0.0000   1st Qu.:0.0000
Median :1.900   Median :2.000   Median :0.0000   Median :0.0000
Mean   :2.194   Mean   :2.549   Mean   :0.2818   Mean   :0.2507
3rd Qu.:2.930   3rd Qu.:3.000   3rd Qu.:1.0000   3rd Qu.:1.0000
Max.   :9.290   Max.   :4.000   Max.   :1.0000   Max.   :1.0000

  dmot3          dmot4          dseg1          dseg2
Min.   :0.0000   Min.   :0.0000   Min.   :0.00000   Min.   :0.0000
1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.0000
Median :0.0000   Median :0.0000   Median :0.00000   Median :0.0000
Mean   :0.1965   Mean   :0.2710   Mean   :0.06005   Mean   :0.4638
3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:1.0000
Max.   :1.0000   Max.   :1.0000   Max.   :1.00000   Max.   :1.0000

  dseg3          dseg4          ttinveh1          ttinveh2          ttinveh3
Min.   :0.0000   Min.   :0.0000   Min.   :0         Min.   :0         Min.   : 2.21
1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0         1st Qu.:0         1st Qu.: 7.75
Median :0.0000   Median :0.0000   Median :0         Median :0         Median :12.51
Mean   :0.3432   Mean   :0.1329   Mean   :0         Mean   :0         Mean   :14.84
3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:0         3rd Qu.:0         3rd Qu.:20.00
Max.   :1.0000   Max.   :1.0000   Max.   :0         Max.   :0         Max.   :75.00

  ttinveh4          ttt1          ttt2          ttt3
Min.   : 1.00   Min.   : 3.00   Min.   : 2.08   Min.   : 5.00
1st Qu.: 6.63   1st Qu.:15.00   1st Qu.: 7.26   1st Qu.:20.78
Median :10.28   Median :25.00   Median :11.60   Median :25.41
Mean   :12.77   Mean   :29.48   Mean   :13.45   Mean   :27.58
3rd Qu.:15.42   3rd Qu.:38.37   3rd Qu.:17.88   3rd Qu.:31.98
Max.   :50.00   Max.   :113.52   Max.   :56.70   Max.   :79.00

  ttt4          tesperal          tespera2          tespera3          tespera4
Min.   : 1.00   Min.   :0         Min.   :0         Min.   : 6.000   Min.   :0
1st Qu.: 9.79   1st Qu.:0         1st Qu.:0         1st Qu.: 9.750   1st Qu.:0
Median :13.23   Median :0         Median :0         Median : 9.750   Median :0
Mean   :15.16   Mean   :0         Mean   :0         Mean   : 9.826   Mean   :0
3rd Qu.:18.32   3rd Qu.:0         3rd Qu.:0         3rd Qu.: 9.750   3rd Qu.:0
Max.   :60.00   Max.   :0         Max.   :0         Max.   :25.000   Max.   :0

  taces1          taces2          taces3          taces4          ppark1
Min.   :0         Min.   :0         Min.   : 3.000   Min.   : 2.000   Min.   :0
1st Qu.:0         1st Qu.:0         1st Qu.: 4.650   1st Qu.: 3.000   1st Qu.:0
Median :0         Median :0         Median : 4.650   Median : 3.270   Median :0
Mean   :0         Mean   :0         Mean   : 4.675   Mean   : 3.271   Mean   :0
3rd Qu.:0         3rd Qu.:0         3rd Qu.: 4.650   3rd Qu.: 3.270   3rd Qu.:0
Max.   :0         Max.   :0         Max.   :75.000   Max.   :45.000   Max.   :0
    
```

```

  ppark2          ppark3          pparktipus1_4          pparktipus2_4          pparktipus3_4
Min.   :0         Min.   :0         Min.   :0.0000   Min.   :0.00000   Min.   :0.0000
1st Qu.:0         1st Qu.:0         1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.0000
Median :0         Median :0         Median :0.0000   Median :0.00000   Median :0.0000
Mean   :0         Mean   :0         Mean   :0.1448   Mean   :0.01535   Mean   :0.2282
3rd Qu.:0         3rd Qu.:0         3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:0.0000
Max.   :0         Max.   :0         Max.   :1.0000   Max.   :1.00000   Max.   :1.0000

  tcost1          tcost2          tcost3          tcost4          f.mot
Min.   :0         Min.   :0         Min.   :0.0700   Min.   :0.060     MOT.obl:1206
1st Qu.:0         1st Qu.:0         1st Qu.:0.4300   1st Qu.:0.210     MOT.ges:1073
Median :0         Median :0         Median :0.4300   Median :0.340     MOT.com: 841
Mean   :0         Mean   :0         Mean   :0.3832   Mean   :0.395     MOT.oci:1160
3rd Qu.:0         3rd Qu.:0         3rd Qu.:0.4300   3rd Qu.:0.530
Max.   :0         Max.   :0         Max.   :0.8600   Max.   :1.670

  f.seg          f.grp          totgrp          velo4.4.1
SEG.<16 : 257   Min.   : 1.000   Min.   : 11.20   Min.   : -20.408
SEG.act :1985   1st Qu.: 3.000   1st Qu.: 46.12   1st Qu.: 1.900
SEG.inact:1469   Median : 8.000   Median : 68.02   Median : 3.693
SEG.+64 : 569   Mean   : 8.376   Mean   : 92.14   Mean   : 4.114
          3rd Qu.:14.000   3rd Qu.:112.88   3rd Qu.: 5.083
          Max.   :16.000   Max.   :889.48   Max.   : 55.810

  velo2.2.1          velo2.3.1          velo.4.1          velo.2.1
Min.   : -20.396   Min.   :0         Min.   : -20.408   Min.   : -20.396
1st Qu.: 4.820   1st Qu.:0         1st Qu.: 1.900   1st Qu.: 4.820
Median : 4.921   Median :0         Median : 3.693   Median : 4.921
Mean   : 4.923   Mean   :0         Mean   : 4.114   Mean   : 4.923
3rd Qu.: 6.066   3rd Qu.:0         3rd Qu.: 5.083   3rd Qu.: 6.066
Max.   : 41.770   Max.   :0         Max.   : 55.810   Max.   : 41.770

  velo.3.1          ttinveh.4.1          ttinveh.2.1          ttinveh.3.1
Min.   : -24.2767   Min.   : 1.00   Min.   :0         Min.   : 2.21
1st Qu.: -1.4816   1st Qu.: 6.63   1st Qu.:0         1st Qu.: 7.75
Median : -0.1756   Median :10.28   Median :0         Median :12.51
Mean   : -0.1995   Mean   :12.77   Mean   :0         Mean   :14.84
3rd Qu.: 0.9315   3rd Qu.:15.42   3rd Qu.:0         3rd Qu.:20.00
Max.   : 59.5901   Max.   :50.00   Max.   :0         Max.   :75.00

  ttt.4.1          ttt.2.1          ttt.3.1          ettt3.4.1          ettt3.2.1
Min.   : -101.32   Min.   : -87.92   Min.   : -80.000   Min.   :0         Min.   :0
1st Qu.: -20.88   1st Qu.: -21.15   1st Qu.: -7.600   1st Qu.:0         1st Qu.:0
Median : -11.03   Median : -13.70   Median : 1.140   Median :0         Median :0
Mean   : -14.32   Mean   : -16.03   Mean   : -1.898   Mean   :0         Mean   :0
3rd Qu.: -4.19   3rd Qu.: -7.59   3rd Qu.: 7.410   3rd Qu.:0         3rd Qu.:0
Max.   : 54.00   Max.   : 21.07   Max.   : 55.000   Max.   :0         Max.   :0

  ettt3.3.1          tespera.4.1          tespera.2.1          tespera.3.1          taces.4.1
Min.   : -22.500   Min.   :0         Min.   :0         Min.   : 6.000   Min.   : 2.000
1st Qu.: 0.000   1st Qu.:0         1st Qu.:0         1st Qu.: 9.750   1st Qu.: 3.000
Median : 0.000   Median :0         Median :0         Median : 9.750   Median : 3.270
Mean   : 1.762   Mean   :0         Mean   :0         Mean   : 9.826   Mean   : 3.271
3rd Qu.: 0.000   3rd Qu.:0         3rd Qu.:0         3rd Qu.: 9.750   3rd Qu.: 3.270
Max.   : 38.750   Max.   :0         Max.   :0         Max.   :25.000   Max.   :45.000

  taces.2.1          taces.3.1          tcost.4.1          tcost.2.1          tcost.3.1
Min.   :0         Min.   : 3.000   Min.   :0.060   Min.   :0         Min.   :0.0700
1st Qu.:0         1st Qu.: 4.650   1st Qu.:0.210   1st Qu.:0         1st Qu.:0.4300
Median :0         Median : 4.650   Median :0.340   Median :0         Median :0.4300
Mean   :0         Mean   : 4.675   Mean   :0.395   Mean   :0         Mean   :0.3832
3rd Qu.:0         3rd Qu.: 4.650   3rd Qu.:0.530   3rd Qu.:0         3rd Qu.:0.4300
Max.   :0         Max.   :75.000   Max.   :1.670   Max.   :0         Max.   :0.8600
    
```

El modo de referencia seleccionada es a pie, por tanto se crean las variables explicativas de alternativa como la diferencia entre el valor de la alternativa y el valor correspondiente al modo a pie.

Se han definido variables de tanteo como la velocidad (en km /h) por modo (respecto el a pie) y el tiempo teórico de viaje total en autobús suponiendo que el tiempo de espera es la mitad de la frecuencia.

```
rmod <-transform(rmod,
  velo4.4.1 = 60*((dista/ttt4)-(dista/ttt1)),velo4.2.1 = zeros,velo4.3.1 =zeros,
  velo3.4.1 = zeros,velo3.2.1 = zeros,velo3.3.1 =60*((dista/ttt3)-(dista/ttt1)),
  velo2.4.1 = zeros,velo2.2.1 = 60*((dista/ttt2)-(dista/ttt1)),velo2.3.1 =zeros,
  velo.4.1 = 60*((dista/ttt4)-(dista/ttt1)),velo.2.1 = 60*((dista/ttt2)-
(dista/ttt1)),velo.3.1 = 60*((dista/ttt3)-(dista/ttt1)),
  ttinveh.4.1=ttinveh4-ttinveh1, ttinveh.2.1=zeros,ttinveh.3.1=ttinveh3-ttinveh1,
  ttt.4.1=ttt4-ttt1, ttt.2.1=ttt2-ttt1,ttt.3.1=ttt3-ttt1,
  ettt3.4.1 = zeros,ettt3.2.1 = zeros,ettt3.3.1 =(tacc3+tespera3+ttinveh3-ttt3),
  tespera.4.1=tespera4-tespera1, tespera.2.1=tespera2-tespera1,tespera.3.1=tespera3-
tespera1,
  tacc3.4.1=tacc34, tacc3.2.1=zeros,tacc3.3.1=tacc33,
  tcost.4.1=tcost4-tcost1, tcost.2.1=tcost2-tcost1,tcost.3.1=tcost3-tcost1
)
```

Un primer análisis de consistencia de los datos muestra diferencias significativas entre las distancias de los desplazamientos por motivo y por segmento:

```
> tapply(dista,f.mot,summary)
$MOT.obl
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.340  1.300   2.120   2.478  3.290   9.290

$MOT.ges
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.340  1.190   1.850   2.063  2.660   8.470

$MOT.com
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.340  1.030   1.720   1.968  2.650   7.210

$MOT.oci
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.340  1.160   1.900   2.184  2.960   9.110

> tapply(dista,f.seg,summary)
$`SEG.<16`
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.340  0.970   1.610   1.939  2.650   6.650

$SEG.act
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.340  1.330   2.070   2.374  3.170   9.290
```

| | | | | | | |
|-------------|-------|---------|--------|-------|---------|-------|
| \$SEG.inact | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| | 0.340 | 1.160 | 1.890 | 2.145 | 2.890 | 8.470 |
| \$`SEG.+64` | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| | 0.340 | 0.940 | 1.510 | 1.807 | 2.400 | 6.660 |

Gráficamente se pueden representar la tendencia central de la longitud de los desplazamientos según el motivo principal y el segmento de población que lo realice y realizar contrastes de hipótesis estadísticas que argumentan diferencias significativas de las longitudes de los desplazamientos por Motivo Principal y según el Segmento de la población que lo realice.

```
> oneway.test(dista~ f.mot, data=rmod)

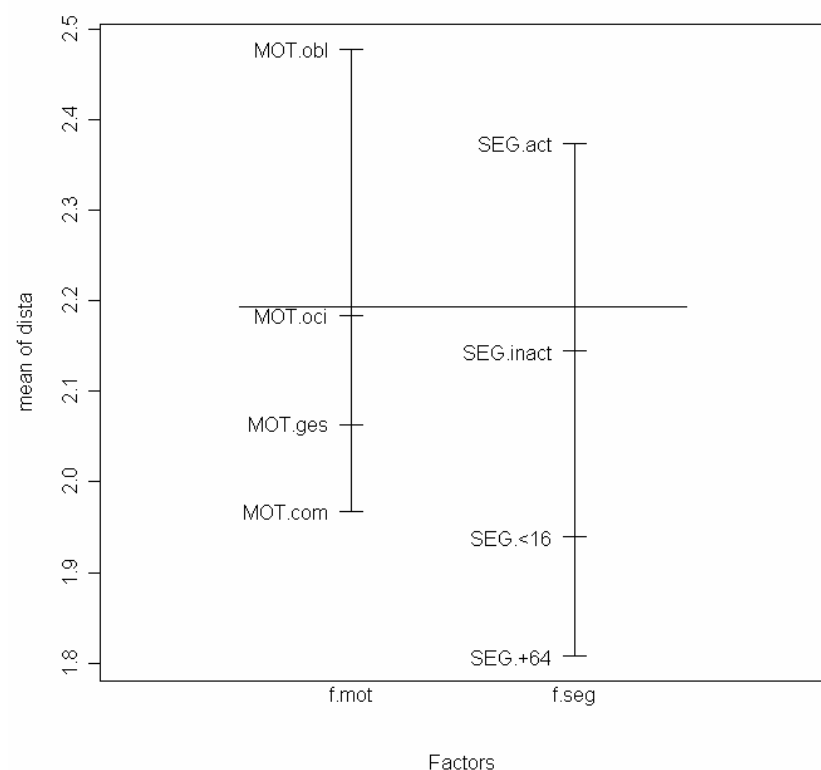
One-way analysis of means (not assuming equal variances)

data: dista and f.mot
F = 26.9983, num df = 3.000, denom df = 2322.474, p-value < 2.2e-16

> oneway.test(dista~ f.seg, data=rmod)

One-way analysis of means (not assuming equal variances)

data: dista and f.seg
F = 35.8674, num df = 3.000, denom df = 986.424, p-value < 2.2e-16
```

La variable característica de los modos mecanizados de tiempo de viaje en vehículo medio modal por registro (OD-segmento-motivo), se ha visto con mucha frecuencia que coincide con el tiempo total de viaje modal por registro, de ahí que no haya sido posible sacarle partido a esta importante variable: debido a las características de la grabación efectuada por APPEND.

El modelo nulo con grupo de referencia a pie y por tanto variables explicativas diferencia respecto al pie facilita unos resultados:

```
> m0 <- vglm( cbind(g2, g3, g4, g1) ~ 1, fam = multinomial, data=rmod)
> summary(m0)
```

```
Call:
vglm(formula = cbind(g2, g3, g4, g1) ~ 1, family = multinomial,
      data = rmod)
```

```
Pearson Residuals:
           Min       1Q   Median       3Q      Max
log(mu[,1]/mu[,4]) -8.8046 -2.4529 -1.6963 -0.54724 89.255
log(mu[,2]/mu[,4]) -12.9995 -3.6216 -2.4835 -0.90146 46.316
log(mu[,3]/mu[,4]) -21.2132 -5.8669 -3.8048  8.14555 32.490
```

```
Coefficients:
           Value Std. Error t value
(Intercept):1 -2.64875  0.0082744 -320.11
(Intercept):2 -1.75382  0.0055366 -316.77
(Intercept):3 -0.61688  0.0035925 -171.71
```

```
Number of linear predictors: 3
```

```
Names of linear predictors:
log(mu[,1]/mu[,4]), log(mu[,2]/mu[,4]), log(mu[,3]/mu[,4])
```

```
Dispersion Parameter for multinomial family: 1
```

```
Residual Deviance: 686554.2 on 12837 degrees of freedom
```

```
Log-likelihood: -410234.1 on 12837 degrees of freedom
```

```
Number of Iterations: 5
```

La tabla de confusión resultante muestra que todos los desplazamientos se clasifican como a pie:

```
> # Tabla de confusión: filas obs x columnas pred
> # Predicción de probabilidades modales
> pm0 <- predict.vglm( m0, type="res" )
> apply(pm0,2,mean)
           g2           g3           g4           g1
0.03966378 0.09706440 0.30256884 0.56070298
> dgm0
$r2n
[1] 0

$r2
[1] 0

$r2e
[1] 0

$aic
[1] 820474.2

$roc
[1] 56.0703

$tc
      obici  obus  ocar  opeu
pbici      0     0     0     0
pbus       0     0     0     0
pauto      0     0     0     0
ppeu 15638 38269 119292 221065
$pbci
obici  obus  ocar  opeu
  0     0     0   100
```

Se tantean diversos modelos y para garantizar la consistencia de los estimadores con la teoría de la utilidad aleatoria se suprimen variables que muestren inconsistencia y se añade la variable explicativa distancia común a todos los modos de transporte en competencia. El primer modelo tanteado contiene todas las variables explicativas de alternativa (exclusivamente): tiempo de viaje en vehículo (*ttinveh*), tiempo total de viaje (*ttt*), tiempo de espera (*tespera*), tiempo de acceso (*taces*) y coste del desplazamiento (*tcost*). Las variables de tiempo se miden en minutos y el coste en euros: son variables explicativas diferencia respecto el modo de referencia que es a pie.

```

> # Modelo elemental: condicional puro sin var motivo-segmento ni var usuario
> m1 <- vglm( cbind(g2, g3, g4, g1) ~
+ ttinveh.2.1 + ttinveh.3.1+ttinveh.4.1+
+ ttt.2.1 + ttt.3.1+ttt.4.1+
+ tespera.2.1 + tespera.3.1+tespera.4.1+
+ taces.2.1 + taces.3.1+taces.4.1+
+ tcost.2.1 + tcost.3.1 + tcost.4.1 ,
+
+      xij = list(
+ ttinveh ~ ttinveh.2.1 + ttinveh.3.1 + ttinveh.4.1,
+ ttt ~ ttt.2.1 + ttt.3.1+ ttt.4.1,
+ tespera ~ tespera.2.1 + tespera.3.1+tespera.4.1,
+ taces ~ taces.2.1 + taces.3.1 + taces.4.1 ,
+ tcost ~ tcost.2.1 + tcost.3.1 + tcost.4.1 ) ,
+
+      fam = multinomial, data=rmod)
>
> coef(m1, matrix=TRUE)
      log(mu[,1]/mu[,4]) log(mu[,2]/mu[,4]) log(mu[,3]/mu[,4])
(Intercept)      -3.36574241      -5.20555381      -4.07875717
ttinveh           0.14123626           0.14123626           0.14123626
ttt              -0.04885012          -0.04885012          -0.04885012
tespera          0.01055856           0.01055856           0.01055856
taces           0.07871864           0.07871864           0.07871864
tcost            1.88061031           1.88061031           1.88061031
> summary(m1)

Call:
vglm(formula = cbind(g2, g3, g4, g1) ~ ttinveh.2.1 + ttinveh.3.1 +
  ttinveh.4.1 + ttt.2.1 + ttt.3.1 + ttt.4.1 + tespera.2.1 +
  tespera.3.1 + tespera.4.1 + taces.2.1 + taces.3.1 + taces.4.1 +
  tcost.2.1 + tcost.3.1 + tcost.4.1, family = multinomial,
  data = rmod, xij = list(ttinveh ~ ttinveh.2.1 + ttinveh.3.1 +
  ttinveh.4.1, ttt ~ ttt.2.1 + ttt.3.1 + ttt.4.1, tespera ~
  tespera.2.1 + tespera.3.1 + tespera.4.1, taces ~ taces.2.1 +
  taces.3.1 + taces.4.1, tcost ~ tcost.2.1 + tcost.3.1 +
  tcost.4.1))

Pearson Residuals:
      log(mu[,1]/mu[,4])  Min      1Q  Median      3Q      Max
log(mu[,2]/mu[,4])  -52.141 -2.5732 -1.5399 -0.74853 104.197
log(mu[,3]/mu[,4]) -109.665 -4.2605 -2.3210  4.54395  49.983

Coefficients:
      Value Std. Error  t value
(Intercept):1 -3.365742 0.00946507 -355.5960
(Intercept):2 -5.205554 0.02992128 -173.9750
(Intercept):3 -4.078757 0.01313456 -310.5362
ttinveh         0.141236 0.00046215  305.6082
ttt            -0.048850 0.00029879 -163.4926
tespera         0.010559 0.00262363   4.0244
taces           0.078719 0.00196134  40.1352
tcost            1.880610 0.01867294 100.7132

Number of linear predictors: 3

Names of linear predictors:
log(mu[,1]/mu[,4]), log(mu[,2]/mu[,4]), log(mu[,3]/mu[,4])

Dispersion Parameter for multinomial family: 1

```

```
Residual Deviance: 502269.4 on 12832 degrees of freedom
```

```
Log-likelihood: -318091.7 on 12832 degrees of freedom
```

```
Number of Iterations: 5
```

```
>
> dgml <- diagnosi( m1, m0, M, n, nobs, Noms )
> dgml
```

```
$r2n
[1] 0.5183557
```

```
$r2
[1] 0.2246093
```

```
$r2e
[1] 0.2681348
```

```
$aic
[1] 636199.4
```

```
$roc
[1] 68.90028
```

```
$tc
      obici  obus  ocar  opeu
pbici      0      0      0    443
pbus      503  6281  2381  1698
pauto    6072 15942 68880 22436
ppeu     9063 16046 48031 196488
```

```
$pbc
      obici  obus  ocar  opeu
0.00000 16.41276 57.74067 88.88246
```

```
$pgof
[1] 0
```

```
$tg
[1] 184284.8
```

```
$pg
[1] 0
```

```
>
```

No se consigue predecir usos modales de bicicleta. El modelo no resulta satisfactorio según la Teoría de la Utilidad Aleatoria ya que existen coeficientes como el coste o el tiempo de espera que deberían intervenir con signo negativo: a mayor tiempo de espera o a mayor coste menor utilidad del modo.

A nivel puramente estadístico el coeficiente de determinación generalizado se sitúa cerca del 27% lo que según los estudios de simulación para modelos de utilidad aleatoria no resulta excesivamente bajo y se debe considerar equivalente a un coeficiente de determinación clásico de más del 50%, concretamente el coeficiente de determinación de Naglekerke indica una explicabilidad del casi el 53% de la elección modal agregada. Globalmente la capacidad predictiva del modelo resultante es de cerca del 70%, con una gran asimetría de calidad modal en la predicción. Concretamente no se consigue predecir ningún desplazamiento en bicicleta, sólo un 16% en autobús y un 58% en auto privado. La mayor parte de los desplazamientos en autobús están mal

clasificados por el modelo como desplazamientos en automóvil y la mala clasificación del automóvil es porque se vincula a una predicción de a pie.

Para limitar el efecto de las características diferenciales entre los modos no mecanizados y los mecanizados se ha incluido una variable individual, la distancia, que es común para todas las alternativas. El modelo de selección discreta de alternativas agregado deja de ser condicional puro y se convierte en mixto. Se procede a introducir la variable individual distancia y a suprimir todos las variables explicativas que en el proceso de estimación no faciliten coeficientes consistentes con la Teoría de la Utilidad.

```
> summary( m3)
Call:
vglm(formula = cbind(g2, g3, g4, g1) ~ dista + ttt.2.1 + ttt.3.1 +
  ttt.4.1 + tespera.2.1 + tespera.3.1 + tespera.4.1 + taces.2.1 +
  taces.3.1 + taces.4.1 + tcost.2.1 + tcost.3.1 + tcost.4.1,
  family = multinomial, data = rmod, xij = list(ttt ~ ttt.2.1 +
  ttt.3.1 + ttt.4.1, tespera ~ tespera.2.1 + tespera.3.1 +
  tespera.4.1, taces ~ taces.2.1 + taces.3.1 + taces.4.1,
  tcost ~ tcost.2.1 + tcost.3.1 + tcost.4.1))

Pearson Residuals:
           Min          1Q      Median          3Q          Max
log(mu[,1]/mu[,4]) -115.35 -2.1214 -1.3296 -0.70634 79.500
log(mu[,2]/mu[,4]) -135.05 -2.8055 -1.7890 -1.00159 61.805
log(mu[,3]/mu[,4]) -139.57 -4.6816 -2.7750  5.27426 49.883

Coefficients:
           Value Std. Error  t value
(Intercept):1 -4.3641743  0.01887773 -231.1811
(Intercept):2 -3.8685890  0.03534773 -109.4438
(Intercept):3 -3.3023946  0.01284057 -257.1844
dista:1        0.9085022  0.00772969  117.5341
dista:2        1.3560284  0.00547045  247.8822
dista:3        1.5479595  0.01019771  151.7948
ttt            -0.0051308  0.00025275  -20.3001
tespera       -0.0110007  0.00260060   -4.2301
taces         -0.0106919  0.00228264   -4.6840
tcost        -1.3965947  0.04956298  -28.1782

Number of linear predictors:  3

Names of linear predictors:
log(mu[,1]/mu[,4]), log(mu[,2]/mu[,4]), log(mu[,3]/mu[,4])

Dispersion Parameter for multinomial family:  1

Residual Deviance: 536384.2 on 12830 degrees of freedom

Log-likelihood: -335149.1 on 12830 degrees of freedom

Number of Iterations: 5
>
> Dm3.m2<-2*( logLik(m3) - logLik(m2) )
> Dm3.m2
[1] 78242.31
> dfm3.m2<-df.residual(m2)-df.residual(m3)
> 1-pchisq(Dm3.m2,dfm3.m2)
> dgm3 <-diagnosi( m3, m0, M, n, nobs, Noms )
```

```
> dgm3
$r2n
[1] 0.426011

$r2
[1] 0.1830296

$r2e
[1] 0.2183037

$aic
[1] 670318.2

$roc
[1] 67.05811

$tc
      obici  obus  ocar  opeu
pbici    0     0     0     0
pbus     0     0     0     0
pauto  6610 23567 67092 23771
ppeu   9028 14702 52200 197294

$pbci
      obici  obus  ocar  opeu
0.00000  0.00000 56.24183 89.24705

$pgof
[1] 0

$tg
[1] 150170

$pg
[1] 0
>
```

La explicabilidad del modelo es bastante pobre y se sitúa alrededor del 43%, según el coeficiente de determinación de Naglekerke (el más semejante en magnitud al equivalente de la regresión clásica). La capacidad predictiva es del 67% con un porcentaje del 100% de mal clasificados en autobús y en bicicleta y un 44% en automóvil. Los ciclistas son predichos principalmente como modo a pie (en casi un 75% de los casos). Los usuarios del autobús son etiquetados (predichos) mayoritariamente como usuarios del coche en más de un 60% de los casos y como peatones en cerca del 40% restante. Globalmente el modelo incluye variables que son estadísticamente significativas, es decir explicativas de la elección modal, reduciendo en más de 150000 unidades la desviación del modelo nulo, lo que resulta estadísticamente significativo con 7 parámetros.

El contraste de la variable individual distancia resulta significativa. Obsérvese que participa en la utilidad de manera diferencial según el modo. Los coeficientes de las variables explicativas incluidas en la utilidad son consistentes y reproducen una característica conocida y aceptada por los planificadores del transporte: el tiempo de acceso a pie y el tiempo de espera tienen del orden del doble de impacto negativo en la utilidad modal que el tiempo de viaje global.

La introducción de variables explicativas individuales (no de alternativa) relativas al motivo del desplazamiento y del segmento de la población refina los estimadores sin añadir una calidad significativamente mayor a la capacidad predictiva del modelo.

```
> idev<-( deviance(m3) - deviance(m6) )
> gll <-df.residual(m3)-df.residual(m6 )
> idev
[1] 27120.78
> gll
[1] 15
> 1-pchisq(idev,gll)
[1] 0
> summary(m6)
Call:
vglm(formula = cbind(g2, g3, g4, g1) ~ dista + dmot2 + dmot3 +
      dmot4 + dseg1 + dseg2 + ttt.2.1 + ttt.3.1 + ttt.4.1 + tespera.2.1 +
      tespera.3.1 + tespera.4.1 + taces.2.1 + taces.3.1 + taces.4.1 +
      tcost.2.1 + tcost.3.1 + tcost.4.1, family = multinomial,
      data = rmod, xij = list(ttt ~ ttt.2.1 + ttt.3.1 + ttt.4.1,
      tespera ~ tespera.2.1 + tespera.3.1 + tespera.4.1, taces ~
      taces.2.1 + taces.3.1 + taces.4.1, tcost ~ tcost.2.1 +
      tcost.3.1 + tcost.4.1))

Pearson Residuals:
           Min          1Q      Median          3Q          Max
log(mu[,1]/mu[,4]) -130.78 -1.9951 -1.1768 -0.61494 94.761
log(mu[,2]/mu[,4]) -149.44 -2.7456 -1.7362 -0.89949 56.456
log(mu[,3]/mu[,4]) -155.22 -4.4316 -2.4144  4.88568 48.593

Coefficients:
           Value Std. Error  t value
(Intercept):1 -3.6907229 0.02457935 -150.15541
(Intercept):2 -4.0470046 0.03855745 -104.96037
(Intercept):3 -3.9866100 0.01695908 -235.07231
dista:1      0.8604134 0.00787341  109.28095
dista:2      1.3749382 0.00561658  244.79977
dista:3      1.4375898 0.01088361  132.08755
dmot2:1     -1.6629099 0.03455345  -48.12572
dmot2:2     -0.0986040 0.01853135   -5.32093
dmot2:3      0.6330811 0.01210616   52.29414
dmot3:1     -1.4435493 0.02869451  -50.30751
dmot3:2     -0.2876663 0.01760662  -16.33853
dmot3:3     -0.1885000 0.01285787  -14.66028
dmot4:1     -0.5226533 0.02147532  -24.33739
dmot4:2     -0.0601575 0.01698109   -3.54262
dmot4:3      0.1111334 0.01246680    8.91435
dseg1:1      0.0981150 0.03247488    3.02126
dseg1:2      0.3524936 0.02502133   14.08772
dseg1:3      0.0560107 0.02270756    2.46661
dseg2:1     -0.0143949 0.01894526   -0.75982
dseg2:2     -0.2137293 0.01421636  -15.03404
dseg2:3      1.0232591 0.00963205  106.23479
ttt         -0.0059421 0.00025608  -23.20454
tespera     -0.0020506 0.00262403   -0.78145
taces       -0.0180051 0.00224883   -8.00643
```

```
tcost      -0.7983270 0.05323345 -14.99672

Number of linear predictors: 3

Names of linear predictors:
log(mu[,1]/mu[,4]), log(mu[,2]/mu[,4]), log(mu[,3]/mu[,4])

Dispersion Parameter for multinomial family: 1

Residual Deviance: 509263.4 on 12815 degrees of freedom

Log-likelihood: -321588.7 on 12815 degrees of freedom

Number of Iterations: 7
> dgm6
$r2n
[1] 0.4993747

$r2
[1] 0.2160849

$r2e
[1] 0.2569593

$aic
[1] 643227.4

$roc
[1] 68.21546

$tc
      obici  obus  ocar  opeu
pbici    0    0    0    0
pbus     0  1422  426  740
pauto  5300 20271 71621 24419
ppeu  10338 16576 47245 195906

$pbci
      obici      obus      ocar      opeu
0.000000  3.715801 60.038393 88.619184

$pgof
[1] 0

$tg
[1] 177290.8

$pg
[1] 0
```

El mejor modelo *logit* multinomial añade 22 parámetros al modelo nulo, incluye las variables de tiempo de viaje total, tiempo de espera y tiempo de acceso, así como el coste del viaje (sin considerar el tema aparcamiento, ya que no se disponía específicamente de este coste a nivel de encuesta). El modelo no satisface los contrastes de bondad, es decir todavía queda por explicar casi el 50% de la sistemática de la elección modal. La calidad predictiva es del 68% de los desplazamientos bien clasificados, aunque su composición sigue siendo muy desigual: el 88% de los viajes a pie como modo principal están bien clasificados (predichos), aunque sólo un 60% de los viajes en coche se predicen clasificados en ese modo, con cerca del 40% restante clasificado como a pie.

La selección modal del autobús resulta predicha de manera precaria (sólo un 4%), siendo mayoritariamente clasificados como usuarios del coche privado (55%) o a pie (45%). Tampoco hay previsiones satisfactorias de elección modal de la bicicleta: masivamente, cerca del 80% de los casos, son clasificados como peatones o sinó como usuarios del vehículo privado.

El impacto modal de los grandes motivos de desplazamiento resulta consistente con el sentido y los coeficientes se interpretan en la siguiente tabla en función de su impacto en la utilidad, en el incremento o decremento porcentual de sus odds (tasa relativa de elección respecto el modo de referencia a pie) y su impacto aproximado en términos porcentuales:

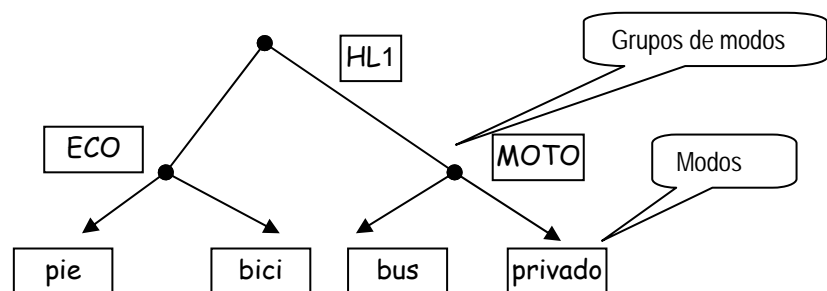
| Impacto de variables | Utilidad | Incremento/Decremento Porcentual en los odds | Incremento/Decremento Porcentual en la probabilidad absoluta de selección modal | Comentarios |
|---|-----------|--|---|--|
| B <i>Media utilidad no observada</i> | - | -98 | | |
| <i>c</i> <i>Motivo gestiones</i> | 3.6907229 | - | -6% | |
| <i>i</i> <i>Motivo compras</i> | 1.6629099 | - | -5% | |
| <i>c</i> <i>Motivo compras</i> | - | -76 | -5% | |
| <i>l</i> <i>Motivo Ocio</i> | 1.4435493 | - | -2% | |
| <i>e</i> <i>Motivo Ocio</i> | - | -41 | -2% | |
| <i>t</i> <i>Segmento población activa</i> | 0.5226533 | - | 0% | |
| <i>a</i> <i>Segmento población activa</i> | 0.0143949 | -1 | 0% | |
| <i>Segmento niños</i> | 0.098115 | 10 | 0% | |
| <i>Distancia (+1km)</i> | 0.8604134 | 136 | 3% | Por cada km de incremento de longitud del viaje el uso de la bici sobre a pie se incrementa en un 136% o en absoluto se incrementa en un 3% la probabilidad de selección modal |
| A <i>Media utilidad no observada</i> | - | -98 | | |
| <i>u</i> <i>observada</i> | 4.0470046 | | | |
| <i>t</i> <i>Motivo compras</i> | - | -25 | -3% | |
| <i>o</i> <i>Segmento población activa</i> | 0.2876663 | - | -2% | |
| <i>b</i> <i>Segmento población activa</i> | - | -19 | -2% | |
| <i>ú</i> <i>población activa</i> | 0.2137293 | | | |
| <i>s</i> <i>Motivo gestiones</i> | -0.098604 | -9 | -1% | |
| <i>Motivo ocio</i> | - | -6 | -1% | |
| <i>Segmento niños</i> | 0.0601575 | | | |
| <i>Segmento niños</i> | 0.3524936 | 42 | 3% | |
| <i>Distancia (+1km)</i> | 1.3749382 | 295 | 12% | Por cada km de incremento de longitud del viaje el uso del autobús sobre el a pie se incrementa en un 295% o en absoluto se incrementa en un 12% la probabilidad de |

| selección modal | | | | |
|-----------------|------------------------------------|-----------|-----|------|
| A | <i>Media utilidad no observada</i> | -3.98661 | -98 | |
| u | <i>Motivo compras</i> | -0.1885 | -17 | -5% |
| t | <i>Segmento niños</i> | 0.0560107 | 6 | 1% |
| o | <i>Motivo Ocio</i> | 0.1111334 | 12 | 3% |
| p | <i>Motivo gestiones</i> | 0.6330811 | 88 | 16% |
| r | <i>Segmento población activa</i> | 1.0232591 | 178 | 26% |
| j | <i>Distancia (+1km)</i> | 1.4375898 | 321 | 30% |
| v | | | | |
| a | | | | |
| d | | | | |
| o | | | | |
| G | <i>Ttt (+5min)</i> | - | | |
| l | | 0.0059421 | -3 | -1% |
| o | <i>Tespera (+5min)</i> | - | | |
| b | | 0.0020506 | -1 | 0% |
| a | <i>Tacces (+5min)</i> | - | | |
| l | | 0.0180051 | -9 | -2% |
| | <i>Tcost (+1Euro)</i> | -0.798327 | -55 | -20% |

Modelo Logit Jerárquico

Reparto modal en la jerarquía de nivel 1 : ECOLOGICA vs MOTORIZADA. Se definen variables explicativas diferencias entre utilidades esperadas en los nidos.

Se pueden estimar los parámetros del reparto modal con el método glm() estándar del paquete R, no es necesario recurrir a métodos de estimación para modelos condicionales y se pueden emplear los diagnósticos estándar para modelos lineales generalizados de respuesta binaria (regresión logística).



Función de Utilidad en el Nido ECO (a pie más bicicleta)

En la jerarquía ECO (de Ecològica, HL0.1) se toma a pie como referencia y uso de bici como la respuesta positiva en el modelo logístico agregado a considerar. Las variables explicativas consideradas reflejan las diferencias entre bici y a pie (es decir, variable en bici menos variable en a pie).

Los tanteos demuestran una baja capacidad predictiva del modelo, sin duda marcada por la gran esparsidad de los datos, hay muy pocas observaciones en los datos recogidos en la Encuesta Domiciliaria 2006 de uso modal de la bicicleta. El mejor modelo estimado consistente con la teoría de la utilidad aleatoria es:

```
> summary(m01.1)
Call:
glm(formula = cbind(round(g2), round(g1)) ~ dista + ttt.2.1 +
    f.mot + dseg2 + dseg3, family = binomial, data = df01)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-16.814  -3.318  -2.187  -1.312   34.258

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.2079196  0.0349562 -120.377 < 2e-16 ***
dista        0.7832765  0.0083053   94.311 < 2e-16 ***
ttt.2.1     -0.0040482  0.0006031  -6.713 1.91e-11 ***
f.motMOT.ges -1.7225259  0.0344486  -50.003 < 2e-16 ***
f.motMOT.com -1.4362739  0.0287076  -50.031 < 2e-16 ***
f.motMOT.oci -0.4860587  0.0215305  -22.575 < 2e-16 ***
```

```
dseg2      0.6928883  0.0313540  22.099 < 2e-16 ***
dseg3      0.9164345  0.0313232  29.257 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 102774 on 2644 degrees of freedom
Residual deviance: 85426 on 2637 degrees of freedom
AIC: 85728

Number of Fisher Scoring iterations: 8
```

El modelo contiene 8 variables explicativas. La variable de alternativa es el tiempo de viaje total en bici (diferencia respecto al pie) y se puede observar que a medida que aumenta el tiempo total en bicicleta disminuye la utilidad de este modo, concretamente cada 1 min de aumento del tiempo total de viaje en bicicleta se reduce en 4 milésimas la utilidad del modo (siendo este valor estimado estadísticamente significativo. Las variables características del desplazamiento (pero no del modo) estadísticamente significativas son la distancia (o longitud del desplazamiento), el indicador del motivo y la pertenencia del desplazamiento al segmento de población adulta, activa o inactiva. Concretamente, se interpretaría:

- Por cada km de incremento de longitud en el desplazamiento la utilidad del modo bicicleta se incrementa en términos relativos al modo a pie en 0.78 unidades. En términos de odds (probabilidad relativa de bicicleta sobre a pie), por cada km de incremento de longitud los odds se multiplican por $\exp(0.78)$, es decir se incrementan en casi un 120%.
- Por cada minuto de incremento del tiempo total del viaje en bicicleta (respecto al pie), los odds (probabilidad relativa de bicicleta sobre a pie) se multiplican por $\exp(-0.0040482)$, es decir se reducen en menos del 0,5%.
- En promedio, la utilidad de la bicicleta se reduce en 1.73 unidades por motivo gestiones respecto al motivo de referencia trabajo, fijadas el resto de las variables. En términos de probabilidad relativa de bicicleta sobre a pie, ésta se reduce en un 82% por motivo gestiones sobre la referencia de motivo trabajo.
- En promedio, la utilidad de la bicicleta se reduce en 1.44 unidades por motivo compras respecto al motivo de referencia trabajo, fijadas el resto de las variables. En términos de probabilidad relativa de bicicleta sobre a pie, ésta se reduce en un 76% por motivo compras sobre la referencia de motivo trabajo.
- En promedio, la utilidad de la bicicleta se reduce en 0.49 unidades por motivo ocio respecto al motivo de referencia trabajo, fijadas el resto de las variables. En términos de probabilidad relativa de bicicleta sobre a pie, ésta se reduce en cerca de un 40% por motivo ocio sobre la referencia de motivo trabajo.
- En lo que respecta a los grupos de población, en promedio la utilidad de la bicicleta para la población adulta (entre 16 y 64 años) activa se incrementa en 0.69 unidades respecto el grupo de referencia que es población infantil o tercera edad. En términos de probabilidad relativa de bicicleta sobre a pie, ésta se incrementa en un 100% por en la población adulta activa sobre la referencia de población infantil o ancianos.
- En lo que respecta a los grupos de población, en promedio la utilidad de la bicicleta para la población adulta (entre 16 y 64 años) inactiva se incrementa en 0.92 unidades respecto el grupo de referencia que es población infantil o tercera edad. En términos de probabilidad relativa de bicicleta sobre a pie, ésta se incrementa en un 150% en la población adulta inactiva sobre la referencia de población infantil o ancianos.

| Variable | | Odds (probabilidad relativa sobre a pie) | Incremento/decreme nto de los odds sobre la referencia (a pie) en % |
|----------|--------------|---|--|
| | Coefficiente | | |

| | | | |
|------------------------------|------------|-----------|--------|
| Ttt.2.1 | -0.0040482 | 0.99596 | 0.40% |
| Motivo | | | |
| Gestiones | -1.7225259 | 0.1786144 | 82.1% |
| Motivo | | | |
| Compras | -1.4362739 | 0.2378122 | 76.2% |
| Motivo Ocio | -0.4860587 | 0.6150457 | 38.5% |
| Adultos y población activa | 0.6928883 | 1.9994823 | 100.0% |
| Adultos y población inactiva | 0.9164345 | 2.5003594 | 150.0% |
| Longitud del viaje | 0.7832765 | 2.1886316 | 118.9% |

| | | | | |
|------------------|--------|------|--------|--------------|
| Bicicleta | 1007 | 141 | 1148 | 0.7% |
| Total | 161216 | 9846 | 171062 | 93.7% |

La movilidad diaria media a pie observada en día laborable supone casi un 95% del total de viajes en modos ecológicos. El porcentaje de selecciones modales acertadas sobre el total en modos ecológicos es de casi el 94%, pero el modelo actúa con una sobrepredicción del modo a pie, lo cual no resulta de extrañar dada la baja incidencia del modo bicicleta: más del 99% de las predicciones se efectúan imputando el modo a pie. Hay que considerar que la capacidad predictiva está limitada al NIDO ECO.

El anterior modelo constituye una caracterización detallada de la función de utilidad de la bicicleta respecto al pie, pero resulta poco práctica para la prognosis ya que en los modelos de demanda propuestos, la distribución se realiza a nivel agregado por motivos y segmentos de población y posteriormente se aplica el reparto modal, por tanto la función de utilidad más apta para prognosis queda recogida en el siguiente modelo:

```
> summary(sm01.2)
Call:
glm(formula = cbind(round(g2), round(g1)) ~ dista, family = binomial,
    data = df01)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-15.721  -3.565  -2.557  -1.902   35.648

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.175571   0.018956  -220.3  <2e-16 ***
dista        0.846025   0.008023   105.4  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 102774 on 2644 degrees of freedom
Residual deviance: 91736 on 2643 degrees of freedom
AIC: 92026
```

La probabilidad de uso de la bicicleta aumenta proporcionalmente a la distancia del desplazamiento a razón de una pérdida de utilidad de 0.84 unidades por km, lo que se traduce en términos de probabilidad relativa de uso de la bicicleta respecto al pie con un incremento de un 133% de los odds de la bicicleta por cada km de aumento de la longitud del viaje.

| Variable | | Odds (probabilidad relativa sobre a pie) | Incremento/decremento de los odds sobre la referencia (a pie) en % |
|----------------|-----------|--|--|
| Constante bici | -4.175571 | | |

El coeficiente de determinación generalizado en tanto por uno toma el valor 0.1687 y reescalado por el número de parámetros (coeficiente de McFadden) se fija a 0.1666. Hay que evitar comparar este coeficiente con el coeficiente de determinación de la regresión clásica, ya que el primero depende en magnitud de la probabilidad de respuesta positiva (uso de bicicleta). El más semejante desarrollado hasta la fecha, en tema de coeficientes de determinación, es el de Naglekerke (1991) que ronda en el modelo descriptivo el 20%. Son valores bajos, pero no están alejados de los coeficientes que se publican en estudios de calibración de funciones de utilidad aleatorias; adicionalmente, en nuestro caso se ha partido de datos agregados, no individualizados precisamente para facilitar la aplicabilidad inmediata a la prognosis de la demanda y evitar dificultades prácticas.

```
> # Coef determinacio de Naglekerke modelo descriptivo:
> R2N<- (1-exp((deviance(m01.1)-deviance(m01.0))/sum(df01$g1+df01$g2))) / (1-exp(-deviance(m01.0)/sum(df01$g1+df01$g2)))
> # Coeficient determinación tipo Mc Fadden
> R2e<- (1-( ( deviance(m01.1)/df.residual(m01.1) ))/( deviance(m01.0)/df.residual(m01.0) ))
> R2<- (1-( ( deviance(m01.1)/1 ))/( deviance(m01.0)/1 ))
> R2;
[1] 0.1687942
> R2e;
[1] 0.1665878
> R2N
[1] 0.2006399
```

La explicabilidad del modelo en el nido ECO se detalla en la tabla de confusiones parcial que aparece a continuación:

| MODELO LOGIT JERÁRQUICO | | NIDO ECOLÓGICO | | |
|-------------------------|---------------------|----------------|--------|-----------|
| Predicciones | Observaciones (EMD) | | | |
| | A pie | Bicicleta | Total | % s Todos |
| A pie | 160209 | 9705 | 169914 | 99.3% |

| | | | |
|----------------|----------|------|--------|
| Longitud viaje | 0.846025 | 2.33 | 133.0% |
|----------------|----------|------|--------|

- La explicabilidad del modelo de reparto modal del nido ECO (bici vs a pie) ofrece en la modalidad descriptiva y en la simplificada unos indicadores de bondad del ajuste bastante deficientes:

```
> # Coef determinacio de Naglekerke :
> R2N<-(1-exp((deviance(sm01.2)-deviance(sm01.0))/sum(df01$g1+df01$g2) )) / (1-exp(-
deviance(sm01.0)/sum(df01$g1+df01$g2)))
> # Coeficient determinación tipo Mc Fadden
> R2e<- (1-( ( deviance(sm01.2)/df.residual(sm01.2) )/( deviance(sm01.0)/df.residual(sm01.0)
)))
> R2<- (1-( ( deviance(sm01.2)/1 )/( deviance(sm01.0)/1 ) ) )
> R2;
[1] 0.1073995
> R2e;
[1] 0.1070618
> R2N
[1] 0.1293578
```

- La hipótesis nula que el modelo se adapta bien a los datos se rechazaría en ambos casos ya que la desviación residual resulta muy elevada (p valores 0).

```
Call:
glm(formula = cbind(round(g2), round(g1)) ~ dista + ttt.2.1 +
     f.mot + dseg2 + dseg3, family = binomial, data = df01)
```

```
Null deviance: 102774 on 2644 degrees of freedom
Residual deviance: 85426 on 2637 degrees of freedom
```

```
Call:
glm(formula = cbind(round(g2), round(g1)) ~ dista, family = binomial,
     data = df01)
```

```
Null deviance: 102774 on 2644 degrees of freedom
Residual deviance: 91736 on 2643 degrees of freedom
AIC: 92026
```

- Las variables socioeconómicas comunes a las alternativas son globalmente explicativas según el test de la desviación aplicable a estos dos modelos encajados.
- La variable tiempo de viaje total (ttt.2.1) significativa en el modelo m01.1 no es estadísticamente significativa en el modelo simplificado sm01.1 y por tanto se suprime en beneficio del modelo descrito m01.2.

```
> summary(sm01.0)
Call:
glm(formula = cbind(round(g2), round(g1)) ~ 1, family = binomial,
     data = df01)
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-11.024  -3.734  -2.888  -2.367   42.539
```

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.648753   0.008275  -320.1  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 102774 on 2644 degrees of freedom
Residual deviance: 102774 on 2644 degrees of freedom
AIC: 103062
```

```
Number of Fisher Scoring iterations: 5
```

```
> summary(sm01.1)
```

```
Call:
glm(formula = cbind(round(g2), round(g1)) ~ dista + ttt.2.1,
     family = binomial, data = df01)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-15.637  -3.561  -2.558  -1.904   35.603
```

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.1879114   0.0205124  -204.165  <2e-16 ***
dista        0.8459300   0.0080208  105.467  <2e-16 ***
```

```
ttt.2.1      -0.0009115   0.0005758   -1.583    0.113
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 102774 on 2644 degrees of freedom
Residual deviance: 91733 on 2642 degrees of freedom
AIC: 92026
```

```
Number of Fisher Scoring iterations: 5
```

```
> summary(sm01.2)
```

```
Call:
glm(formula = cbind(round(g2), round(g1)) ~ dista, family = binomial,
     data = df01)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-15.721  -3.565  -2.557  -1.902   35.648
```

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.175571   0.018956  -220.3  <2e-16 ***
dista        0.846025   0.008023   105.4  <2e-16 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)
```



```

Null deviance: 102774 on 2644 degrees of freedom
Residual deviance: 91736 on 2643 degrees of freedom
AIC: 92026

Number of Fisher Scoring iterations: 5

> # Variable tiempo de viaje total : contraste de modelos
> idev<-( deviance(sm01.2) - deviance(sm01.1) )
> gll <-df.residual(sm01.2)-df.residual(sm01.1)
> idev
[1] 2.490399
> gll
[1] 1
> 1-pchisq(idev,gll)
[1] 0.1145427
> # Contraste de modelo simplificado (sm01.2) y m01.1
> idev<-( deviance(sm01.2) - deviance(m01.1) )
> gll <-df.residual(sm01.2)-df.residualsm01.1)
Error: error de sintaxis en "gll <-df.residual(sm01.2)-df.residualsm01.1)"
> idev
[1] 6309.767
> gll <-df.residual(sm01.2)-df.residual(sm01.1)
> gll
[1] 1
> 1-pchisq(idev,gll)
[1] 0
>

```

En todas las observaciones de la muestra, el modelo simplificado sm01.2 ofrece una utilidad mayor para el modo a pie y por tanto, por consistencia con la teoría de la utilidad : todas las observaciones agregadas que usan modos ECO serán imputadas por la prognósis del modelo a pie. Este aspecto se aliviará empleando el suavizado *logit* que a la práctica tiende a facilitar siempre probabilidades no nulas para los modos enfrentados, por tanto, captación de la bicicleta, aunque mínima, está garantizada. Este aspecto se describe en apartados posteriores bajo la denominación de Predicción Agregada por Utilidad Proporcional.

auto privado, por tanto la función de utilidad estimada refleja la utilidad sistemática del autobús menos la del auto.

| Variable | Coefficiente | Odds (probabilidad relativa sobre a pie) | Incremento/decremento de los odds sobre la referencia (a pie) en % | z valor (estadístico contraste H0 coeficiente 0) |
|--------------|--------------|--|--|--|
| tcost.3.4 | -0.8857913 | 0.41 | 58.8% | -10.149 |
| ttt.3.4 | -0.0470243 | 0.95 | 4.6% | -93.15 |
| dista | 0.0722192 | 1.07 | 7.5% | 4.371 |
| f.motMOT.com | -0.3131843 | 0.73 | 26.9% | -16.681 |
| f.motMOT.ges | -0.9287199 | 0.40 | 60.5% | -47.634 |
| f.motMOT.oci | -0.3186608 | 0.73 | 27.3% | -17.868 |
| dseg2 | -1.7178352 | 0.18 | 82.1% | -65.428 |
| dseg3 | -0.4557372 | 0.63 | 36.6% | -17.077 |

El mejor modelo estimado muestra unos coeficientes altamente significativos y consistentes con la teoría de la utilidad ya que tiempo y coste tienen un impacto negativo sobre la función de utilidad y los motivos compras, gestiones y ocio también tienen un efecto promedio negativo sobre la utilidad del autobús tomando como referencia el motivo de movilidad obligada. En los segmentos de población adulta entre 16 y 64 años, bien sean laboralmente activos o no, el impacto medio en la utilidad del autobús es negativo tomando como referencia los no adultos. La interpretación de los coeficientes es la siguiente:

- **Coste del viaje:** Cada euro de aumento del coste del viaje (exclusivamente tarifa o combustible, no se tiene en cuenta el coste del aparcamiento o su dificultad en el caso del coche privado) comporta un decremento en la utilidad de 0.886 unidades. En la escala de los *odds*: cada euro de incremento implica un decremento del uso del autobús relativo al auto privado de casi el 60%.
- **Tiempo total de viaje:** cada minuto de incremento del tiempo de viaje (tiempo/s de acceso a pie, más tiempo de espera (si autobús) más tiempo de trayecto/s) implica un decremento de 0.047 unidades de la utilidad, lo que se traduce en términos de probabilidad relativa de uso del autobús sobre el auto en una reducción de cerca del 5% por minuto.
- **Distancia:** cada km de incremento de la longitud del viaje se traduce en un incremento de 0.0722 unidades de la utilidad del autobús respecto el auto privado. En términos de *odds* o probabilidades relativas de autobús sobre auto privado, los *odds* se incrementan en un 7.5% por cada km de incremento de la longitud total del viaje.
- **Motivo Compras:** el efecto promedio en la utilidad del modo autobús se decrementa en 0.3132 unidades en viajes por compras con los mismos parámetros que el mismo viaje por movilidad obligada. Los *odds* de selección modal del autobús sobre el auto privado se reducen en casi un 27% al efectuarse un viaje por motivo compras que por trabajo, bajo los mismos valores del resto de variables explicativas.
- **Motivo Gestiones:** el efecto promedio en la utilidad del modo autobús se decrementa en 0.9287 unidades en viajes por gestiones personales o laborales con los mismos parámetros que el mismo viaje por movilidad

Función de Utilidad en el Nido MOTORIZADO (autobús más auto privado)

La calibración de la función de utilidad se ha realizado a partir de la selección de todas las relaciones por motivos y segmentos tales que tuvieran descritos viajes en modo autobús y/o auto privado.

El mejor modelo de utilidad descriptivo se muestra a continuación, en él se detalla el impacto promedio de los distintos motivos (grandes) en la utilidad y grandes segmentos (los cuatro ya descritos).

```
Call:
glm(formula = cbind(g3, g4) ~ dista + ttt.3.4 + tcost.3.4 + f.mot +
    dseg2 + dseg3, family = binomial, data = df02)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-21.712  -5.845  -3.602   5.215  23.621

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.4515749  0.0331320  13.630 < 2e-16 ***
dista        0.0722192  0.0165206   4.371 1.23e-05 ***
ttt.3.4     -0.0470243  0.0005048 -93.150 < 2e-16 ***
tcost.3.4   -0.8857913  0.0872752 -10.149 < 2e-16 ***
f.motMOT.ges -0.9287199  0.0194969 -47.634 < 2e-16 ***
f.motMOT.com -0.3131843  0.0187751 -16.681 < 2e-16 ***
f.motMOT.oci -0.3186608  0.0178341 -17.868 < 2e-16 ***
dseg2       -1.7178352  0.0262554 -65.428 < 2e-16 ***
dseg3       -0.4557372  0.0266871 -17.077 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 149718  on 2118  degrees of freedom
Residual deviance: 127459  on 2110  degrees of freedom
AIC: 128138
```

El modelo emplea 9 parámetros y se definen 4 variables genéricas de las alternativas y el resto son variables socioeconómicas y por tanto comunes para todas las alternativas. La alternativa de referencia considerada es

obligada. Los *odds* de selección modal del autobús sobre el auto privado se reducen en más de un 60% al efectuarse un viaje por motivo gestiones que por trabajo, bajo los mismos valores del resto de variables explicativas.

- **Motivo Ocio:** el efecto promedio en la utilidad del modo autobús se decreta en 0.3186 unidades en viajes por ocio con los mismos parámetros que el mismo viaje por movilidad obligada. Los *odds* de selección modal del autobús sobre el auto privado se reducen en más de un 27% al efectuarse un viaje por motivo ocio que por trabajo, bajo los mismos valores del resto de variables explicativas.
- **Segmento de Población Activa entre 16 y 64 años:** el efecto medio en la utilidad del autobús para este segmento conlleva una reducción de 1.7178 unidades, lo que equivale a una reducción de los *odds* de uso de autobús sobre coche respecto el segmento de niños y ancianos de más del 82% manteniendo el resto de variables constantes.
- **Segmento de Población No Activa entre 16 y 64 años:** el efecto medio en la utilidad del autobús para este segmento conlleva una reducción de 0.4557372 unidades, lo que equivale a una reducción de los *odds* de uso de autobús sobre coche respecto el segmento de niños y ancianos de más del 35% manteniendo el resto de variables constantes.

La desviación residual del modelo de utilidad para el nido MOTORIZADO es muy elevada y por tanto queda todavía por explicar gran parte de las causas de la selección modal.

El coeficiente de determinación del modelo de utilidad agregada para el nido MOTORIZADO es de 0.1480 (en tanto por uno), lo que reescalado por el número de parámetros disminuye hasta 0.1454 por tanto rayando la precariedad, pero ninguna otra variable ha podido ser introducida en el modelo siendo significativamente distinta de cero y consistente su valor con la Teoría de la Utilidad. El coeficiente de determinación de Naglekerke se sitúa en 0.2375, lo que determina un modelo dentro de los umbrales habituales en este tipo de modelización, pero con una gran parte de la selección modal todavía pendiente de sistematización.

La explicabilidad del modelo en el nido MOTORIZADOS se detalla en la tabla de confusiones parcial que aparece a continuación:

| MODELO LOGIT JERÁRQUICO | | NIDO MOTORIZADOS | | |
|----------------------------|---------------------|------------------|-------|-----------|
| Predicciones | Observaciones (EMD) | | | |
| | Autobús | Auto Privado | Total | % s Todos |
| Autobús | 1572 | 5044 | 6616 | 11% |
| Auto Privado | 15 460 | 38177 | 53637 | 89% |
| Total | 17032 | 43221 | 60253 | |
| % Modal observado | 28.3% | 71.7% | | 63% |

La tasa de predicciones correctas por Utilidad Máxima es del 63% globalmente, sin embargo la tasa de predicción correcta entre usuarios reales del autobús está por debajo del 10%, mientras que para los usuarios del vehículo privado está por encima del 88%. El modelo resulta bastante precario. Este aspecto se aliviará empleando el suavizado *logit* que a la práctica tiende a facilitar siempre probabilidades no nulas para los modos enfrentados, por tanto, la captación de la bicicleta, aunque mínima, está garantizada. Este aspecto se describe en apartados posteriores bajo la denominación de Predicción Agregada por Utilidad Proporcional.

Modelo reducido para prognosis: HL02 Bus vs Auto

Este modelo se denomina sm02.1 y contiene como variables explicativas la distancia, el tiempo total de viaje y el coste total del viaje (sólo combustible, excluye aparcamiento).

```
> summary(sm02.1)

Call:
glm(formula = cbind(g3, g4) ~ dista + ttt.3.4 + tcost.3.4, family = binomial,
    data = df02)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-16.942  -5.913  -4.222   5.493  27.628

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.0261219  0.0279216  -0.936   0.350
dista        -0.3524887  0.0117475 -30.005 <2e-16 ***
ttt.3.4      -0.0437620  0.0004822 -90.749 <2e-16 ***
tcost.3.4    -3.1356381  0.0624084 -50.244 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 149718  on 2118  degrees of freedom
Residual deviance: 138567  on 2115  degrees of freedom
AIC: 139237
```

El modelo es significativamente menos explicativo que el modelo descriptivo anteriormente descrito m02.2.

```
> # Contraste modelo m02.2 descriptivo versus simplificado sm02.1
> idev<-( deviance(sm02.1) - deviance(sm02.2) )
> gll <-df.residual(sm02.1)-df.residual(sm02.2)
> idev
[1] 11108.72
> gll
[1] 5
> 1-pchisq(idev,gll)
[1] 0
>
```

El modelo sm02.1 contiene las tres variables más la constante y son estadísticamente distintas de cero y a la usanza del test global de ajuste usado en regresión clásica, conjuntamente las tres variables son estadísticamente explicativas de la selección modal. La cuestión es que no lo son suficiente y la discriminación dentro de los modos MOTORIZADOS del autobús y el coche contiene una gran variabilidad no explicada de manera sistemática en las variables empleadas en el modelo. No ha podido introducirse ninguna variable adicional disponible al encontrarse comportamientos (coeficientes estimados) inconsistentes con la teoría de la utilidad.

```
> summary(sm02.0)

Call:
glm(formula = cbind(g3, g4) ~ 1, family = binomial, data = df02)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.136934   0.005875  -193.5  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 149718  on 2118  degrees of freedom
Residual deviance: 149718  on 2118  degrees of freedom
AIC: 150382

Number of Fisher Scoring iterations: 5

Number of Fisher Scoring iterations: 5

> # Contraste modelo nulo versus simplificado sm02.1
> idev<-( deviance(sm02.0) - deviance(sm02.1) )

> g11 <-df.residual(sm02.0)-df.residual(sm02.1)
> idev
[1] 11150.89
> g11
[1] 3
> 1-pchisq(idev,g11)
[1] 0
>
```

El coeficiente de determinación de Naglekerke y los generalizados de Mc Fadden con y sin reescalado son bajos, pero dentro de la tónica general que suelen presentar estos modelos.

```
> # Coef determinacio de Naglekerke modelo prognosis:
> R2N<-(1-exp((deviance(sm02.1)-deviance(sm02.0))/sum(df01$g3+df01$g4) )) / (1-exp(-
deviance(sm02.0)/sum(df01$g3+df01$g4)))
> # Coeficient determinación tipo Mc Fadden
> R2e<- (1-( ( deviance(sm02.1)/df.residual(sm02.1) )/( deviance(sm02.0)/df.residual(sm02.0)
)))
> R2<- (1-( ( deviance(sm02.1)/1 )/( deviance(sm02.0)/1 ) ) )
> R2;
[1] 0.07447914
> R2e;
[1] 0.07316635
> R2N
[1] 0.2645275
```

Jerarquía MOTORIZADOS vs ECO

Las conclusiones de calidad de los modelos condicionales por nido es bastante desfavorable: hay dos modos dominantes, A pie en modos ECO y Coche en modos MOTORIZADOS. Las variables disponibles de la encuesta de movilidad 2006 no han podido determinar modelos de mejor calidad que los ya expuestos en los apartados anteriores. En este apartado se analizará el reparto entre la jerarquía de primer nivel ECO vs MOTORIZADOS a partir del cálculo de las utilidades máximas esperadas por nido, ECO y MOTORIZADOS. El modelo resultante de reparto entre modos ECO y MOTORIZADOS tiene gran explicabilidad desde el punto de vista estadístico y podrá compensar parcialmente los desatinos de los modelos condicionales por nidos.

La utilidad máxima esperada se puede calcular para cada uno de los NIDOS a partir de los modelos descritos en los apartados anteriores :

EmuEco: logaritmo natural de la suma de las exponenciales de las utilidades de los modos ECO (a pie y bicicleta).

EmuNeco: logaritmo natural de la suma de las exponenciales de las utilidades de los modos MOTORIZADOS (autobús y auto privado).

Se toma de referencia el NIDO ECO y por tanto se calculan las diferencias de utilidades esperadas entre MOTORIZADOS y ECO. Se añaden a la jerarquía las variables específicas de alternativa tiempo de acceso a pie para desplazamientos OD realizados en auto y tiempo de espera en la parada para desplazamientos OD que se quisieran realizar en autobús.

El modelo resultante muestra un coeficiente para la variable que modela las diferencias esperadas en la utilidad en el nido MOTORIZADOS menos ECO inferior a 1 y por tanto consistente con el uso de un modelo logia jerárquico. Los coeficientes estimados se indican a través de la salida de resultados del paquete R a continuación:

```
> summary(m1.0)
Call:
glm(formula = cbind(g3 + g4, g1 + g2) ~ tespera3 + taces4 +
demuneco, family = binomial, data = rmod)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-23.2395  -0.4277  -0.2787   0.3494  24.8079

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.912741   0.040830  22.355  <2e-16 ***
tespera3    -0.043982   0.002972  -14.797  <2e-16 ***
taces4      -0.074198   0.008820   -8.413  <2e-16 ***
demuneco     0.750613   0.005504  136.378  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 434062  on 4279  degrees of freedom
Residual deviance: 18176  on 4276  degrees of freedom
```

AIC: 20719

El modelo globalmente presenta una bondad aceptable, es decir la desviación residual ha disminuido notoriamente aunque todavía se detecte un desajuste que ha sido imposible de corregir con las variables explicativas disponibles:

| Variable | Coefficiente | Odds (probabilidad relativa sobre a pie) | Incremento/decremento de los odds sobre la referencia (a pie) en % | z valor (estadístico contraste H0 coeficiente 0) |
|----------|--------------|--|--|--|
| tespera3 | -0.043982 | 0.96 | 4.3% | 14.797 |
| taces4 | -0.074198 | 0.93 | 7.2% | -8.413 |
| demuneco | 0.750613 | 2.12 | 111.8% | 136.378 |

- Todos los coeficientes estimados en el modelo son significativamente distintos de cero a juzgar por el estadístico z empleado para contrastar las hipótesis nulas individuales de coeficiente igual a cero, cuya distribución asintótica es una normal estándar.
- El tiempo de espera tiene un impacto negativo en la utilidad del nido MOTORIZADOS con una reducción de 0.044 unidades por minuto, que se transcribe en un decremento de los odds, probabilidad relativa de uso de modos MOTORIZADOS sobre uso de modo ECO de más de 4% por minuto de espera.
- El tiempo de acceso tiene un impacto negativo en la utilidad de los modos motorizados con una reducción de 0.074 unidades por minuto adicional de acceso manteniendo el resto de las variables constantes y por tanto un decremento del uso de modos MOTORIZADOS sobre ECO de más del 7% por minuto adicional de caminata.
- La diferencia de las utilidades máximas esperadas indica a mayor utilidad esperada de los modos motorizados respecto los ecológicos mayor utilidad global de los modos motorizados y concretamente por unidad de incremento en la utilidad esperada de los modos motorizados la probabilidad relativa de elegir estos modos frente al pie o la bicicleta (modos ecológicos) se incrementa en más de un 110%. El coeficiente indica una correlación no despreciable entre las alternativas modales intranodales y por tanto la propuesta *logit* jerárquica resulta más adecuada que el modelo *logit* multinomial básico descrito en un apartado anterior. El contraste de hipótesis de coeficiente de la diferencia de utilidades esperadas igual a 1 se rechaza con comodidad (el error estándar del coeficiente es de 0.0055), por tanto, el modelo *Logit* Jerárquico no colapsa a un *Logit* Multinomial.

En las observaciones se observa un 60% de los viajes en modos ecológicos (nido ECO) y el 40% restante en modos motorizados (autobús y auto privado):

```
> apply(rmod[, 4:7], 2, sum) / sum(rmod[, 4:7])
      g1      g2      g3      g4
0.56070298 0.03966378 0.09706440 0.30256884
> apply(rmod[, 70:71], 2, sum) / sum(rmod[, 70:71])
      geco      gneco
0.6003668 0.3996332
```

El modelo ilustrado facilita unas predicciones de reparto modal sobre la muestra completa de trabajo consistentes tal y como se muestra a continuación:

```
> apply(rmod[, 74:75], 2, sum) / sum(rmod[, 74:75])
```

```
pgneco  pgeco
0.3996722 0.6003278
```

Dentro de cada NIDO, las predicciones de probabilidades (condicionales) individuales de cada modo son:

```
> apply(pcm01.1, 2, mean)
[1] 0.93044169 0.06955831
> apply(pcm02.2, 2, mean)
[1] 0.2530469 0.7469531
```

- En el nido ECO, dado que un individuo usa un modo ecológico, la probabilidad condicional de ir a pie es del 93%, y en bici del 7%.
- En el nido MOTORIZADOS, dado que un individuo usa un modo motorizado, la probabilidad condicional de ir en autobús es del 25%, y en auto privado del 75%.
- En la jerarquía superior, la probabilidad absoluta de ir en modo ecológico predicha por el modelo m1.0 es del 60% y de usar un modo motorizado del 40%.

Las anteriores probabilidades son a nivel de individuo, pero no todas las relaciones OD tienen el mismo peso, ya que los datos de partida son agregados y por tanto las anteriores probabilidades únicamente reflejan el reparto de probabilidades medio en función de las utilidades suavizado por la transformación *logit*.

Ahora cabría considerar dos maneras de realizar predicciones con datos agregados que se va a ilustrar sobre la observación 504 de la siguiente tabla:

1. El grupo de observaciones 504 corresponde a 69+48+194= 311 individuos. Las probabilidades condicionales predichas individuales dentro del Nido ECO son 0.87 y 0.13 para a pie y bicicleta respectivamente. Las probabilidades condicionales predichas individuales dentro del Nido MOTORIZADOS son 0.12 y 0.88 para autobús y coche, respectivamente. Las probabilidades predichas individuales en modos ECO es del 60% y por tanto, del 40% para modos MOTORIZADOS. Por tanto, las probabilidades modales predichas individuales pueden calcularse como:
 - a. $P(a\ pie) = P(ECO)P(a\ pie/ECO) = 0.6 \times 0.83 = 0.50$
 - b. $P(bici) = P(ECO)P(bici/ECO) = 0.6 \times 0.17 = 0.10$
 - c. $P(bus) = P(MOTORIZADOS)P(bus/MOTORIZADOS) = 0.4 \times 0.12 = 0.05$
 - d. $P(auto) = P(MOTORIZADOS)P(auto/MOTORIZADOS) = 0.4 \times 0.88 = 0.35$
2. Predicción Agregada de Máxima Utilidad. Ahora bien, si se requiere consistencia con la teoría de la utilidad, las probabilidades reflejan la relación de las utilidades por tanto en el ejemplo sería el modo a pie el que tuviera la mayor utilidad sistemática con el modelo jerárquico elegido y de ahí que los 311 individuos fuesen predichos como individuos que elegirían como modo de transporte en los viajes del grupo A PIE.
3. Predicción Agregada de Utilidad Proporcional. Otra interpretación más próxima a la vertiente estadística de la cuestión sería interpretar que de los 311 individuos, un 50% elegirían a pie, un 10% bici, un 5% autobús y un 35% coche.

```
> round(rmod[ 500:505, c(4:7,67:69,72:73,76:79) ], dig=2) # Per claretat
      g1 g2 g3  g4 emueco emuneco demuneco peco pneo ppeu pcbici pcbus pauto
500  0  0 55 130 -9.31 -0.49  8.81  0.0  1.0  0.00  0.00  0.15  0.85
501  0  0 88  0 -9.31  0.07  9.37  0.0  1.0  0.00  0.00  0.60  0.40
502  0  0  0  52 -9.31 -0.50  8.80  0.0  1.0  0.00  0.00  0.18  0.82
```


| | | | | | | | | | | | | | |
|-----|----|---|----|-----|-------|-------|-------|-----|-----|------|------|------|------|
| 503 | 0 | 0 | 88 | 0 | -9.31 | 0.15 | 9.46 | 0.0 | 1.0 | 0.00 | 0.00 | 0.63 | 0.37 |
| 504 | 69 | 0 | 48 | 194 | 0.07 | -0.87 | -0.94 | 0.6 | 0.4 | 0.83 | 0.17 | 0.12 | 0.88 |
| 505 | 0 | 0 | 81 | 0 | -9.31 | -0.62 | 8.68 | 0.0 | 1.0 | 0.00 | 0.00 | 0.12 | 0.88 |

El coeficiente de determinación generalizado sin corregir vale 0.96 y corregido por los grados de libertad de los modelos toma por valor 0.96 (en tantos por uno). El coeficiente de determinación de Naglekerke toma por valor 0.98 (en tanto por uno). Por tanto el modelo de jerarquía superior que discrimina entre tipo ECO y MOTORIZADOS resulta muy satisfactorio.

```
> R2;
[1] 0.9581261
> R2e;
[1] 0.9580967
> R2N
[1] 0.9764923
```

Las Tablas de Confusión ofrecen una valoración clara de la capacidad predictiva de los modelos y en este caso del modelo *logit* jerárquico propuesto. En filas se ponen las predicciones por modo y en columnas las observaciones por modo. Las diagonales de la matriz indican los individuos observado por modo que han sido predichos como de ese modo y de ahí, que la proporción de diagonales en la matriz respecto al total de individuos indique la capacidad predictiva global del modelo (porcentaje de aciertos) y su complementario el porcentaje de fallos. El análisis detallado de las distintas columnas permite detectar el porcentaje de aciertos por modo. Los porcentajes marginales por fila indican la distribución modal en la muestra y los porcentajes marginales en columnas indican la distribución modal en la predicción.

La Tabla de Confusión para el modelo *logit* jerárquico descriptivo resultante con las predicciones agregadas de Máxima Utilidad es:

| Predicciones | Observaciones (EMD) | | | | | |
|---------------------|---------------------|-----------|---------|--------------|---------------|-----------|
| | A pie | Bicicleta | Autobús | Auto privado | Total | % s Todos |
| A pie | 215556 | 14276 | 6387 | 24049 | 260268 | 66% |
| Bicicleta | 321 | 300 | 96 | 212 | 929 | 0% |
| Autobús | 131 | 0 | 8548 | 3490 | 12169 | 3% |
| Auto privado | 5057 | 1062 | 23238 | 91541 | 120898 | 31% |
| Total | 221065 | 15638 | 38269 | 119292 | 394264 | 100% |
| | 56% | 4% | 10% | 30% | | |
| % Aciertos | 98% | 2% | 22% | 77% | 80% | |

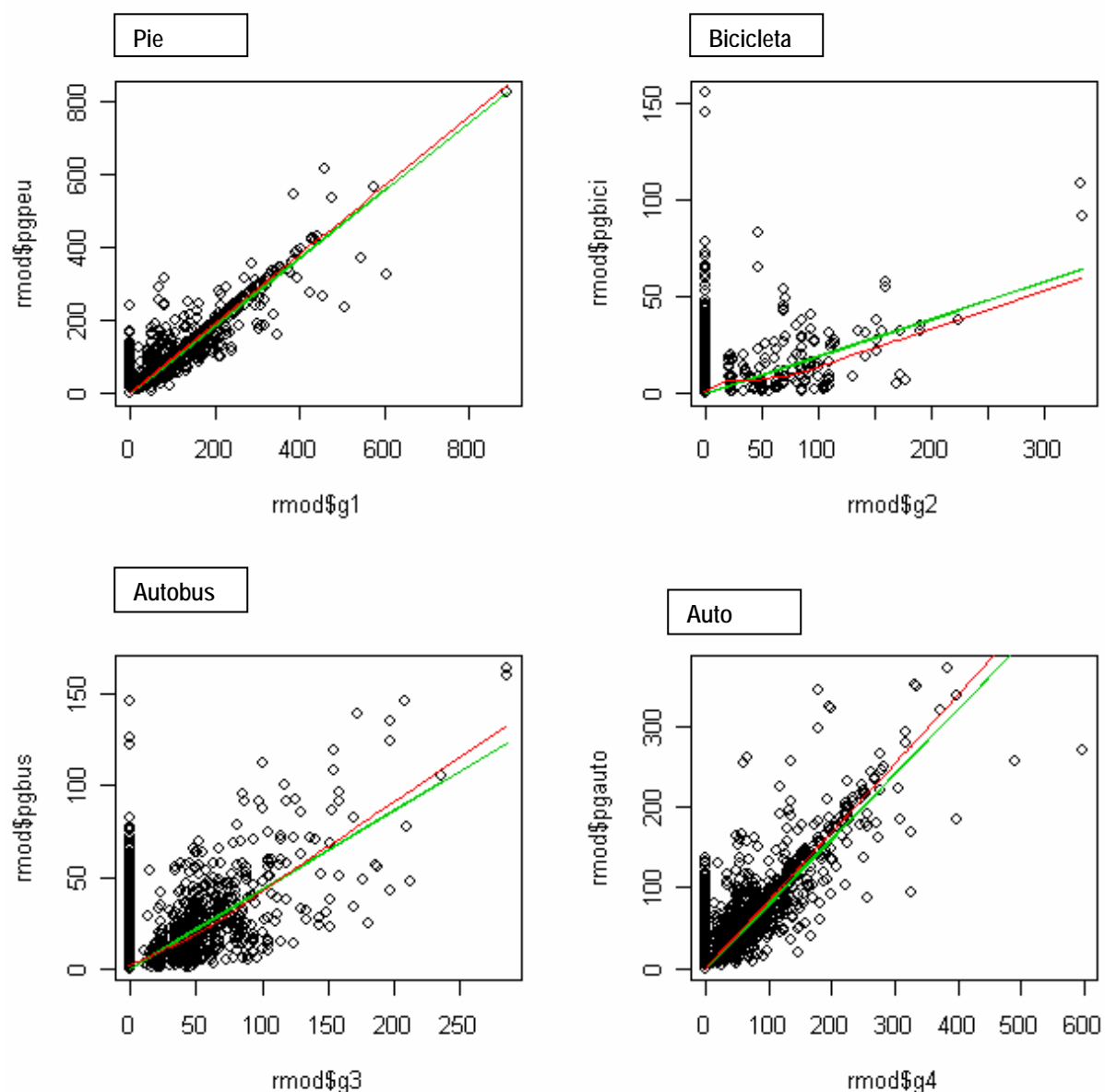
La Tabla de Confusión para el modelo *logit* jerárquico descriptivo resultante con las predicciones agregadas de Utilidad Proporcional es:

| Predicciones | Observaciones (EMD) | | | | | |
|---------------------|---------------------|-----------|---------|--------------|---------------|-----------|
| | A pie | Bicicleta | Autobús | Auto privado | Total | % s Todos |
| A pie | 190803 | 11787 | 3383 | 14354 | 220328 | 56% |
| Bicicleta | 11752 | 2116 | 533 | 1827 | 16227 | 4% |
| Autobús | 4310 | 376 | 12549 | 21237 | 38473 | 10% |
| Auto privado | 13981 | 1345 | 21772 | 81757 | 118856 | 30% |
| Total | 220846 | 15624 | 38237 | 119176 | 393883 | 100% |
| | 56% | 4% | 10% | 30% | 100% | |
| % Aciertos | 86% | 14% | 33% | 69% | 73% | |

Las diferencias son notables y dado que la aproximación es fundamentalmente estadística se empleará como herramienta de predicción agregada la denominada de Utilidad Proporcional, dado que el modelo jerárquico global estimado secuencialmente no resulta plenamente consistente con la teoría de la utilidad aleatoria.

Si nos centramos en el análisis de las observaciones de viajes que seleccionaron el autobús (un 10% del total diario) y en cifras absolutas unos 38250 viajes, solamente unos 12550 son predichos por el modelo correctamente e imputados a una selección modal del autobús, esto da una tasa de aciertos del 33% para este modo. Globalmente y reforzada por la alta incidencia del modo a pie, la tasa global de aciertos se sitúa en el 73%, considerada satisfactoria para este tipo de modelos. Globalmente, la tasa de mercado predicha para el autobús será idéntica a la muestral, es decir del 10% y de algún modo el ruido aleatorio colabora para compensar los valores modales totales agregados: puede observarse que la gran mayoría de predicciones efectuadas de modo autobús proceden de observaciones originariamente pertenecientes al auto privado, por otro lado una confusión muy razonable dadas las características urbanas de Vitoria-Gasteiz y la grabación de partida para este análisis procedente de la Encuesta de APPEND 2006.

Si se enfrentan las observaciones modales observadas y predichas agregadas, los gráficos correspondientes a cada uno de los modos refleja el típico suavizado logístico en las predicciones modales absolutas agrupadas.



En rojo se muestra el ajuste por mínimos cuadrados de la recta $y = bx$, donde x son observaciones e y son predicciones. En verde se muestra el ajuste de observaciones vs predicciones por técnicas de regresión local no paramétricas: se observa la total consistencia entre ambos enfoques a pesar que la aproximación mínimo cuadrática no es ciertamente adecuada para efectuar el presente análisis por temas de heterocedasticidad.

Los indicadores estadísticos para los ajustes mínimo cuadráticos se muestran a continuación en el mismo orden: a pie, bicicleta, autobús y auto privado.

```
> summary(a1)
Call:
lm(formula = rmod$pgpeu ~ -1 + rmod$g1)

Residuals:
    Min       1Q   Median       3Q      Max
-239.224   0.000   0.000   3.041  242.028

Coefficients:
```

```
Estimate Std. Error t value Pr(>|t|)
rmod$g1 0.928028  0.003923  236.5  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22.54 on 4279 degrees of freedom
Multiple R-Squared: 0.929, Adjusted R-squared: 0.9289
F-statistic: 5.595e+04 on 1 and 4279 DF, p-value: < 2.2e-16

> summary(a2)

Call:
lm(formula = rmod$pgbici ~ -1 + rmod$g2)

Residuals:
    Min       1Q   Median       3Q      Max
-27.5535   0.0000   0.7611   3.3825  155.7973

Coefficients:
Estimate Std. Error t value Pr(>|t|)
rmod$g2 0.193526  0.006389  30.29  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.543 on 4279 degrees of freedom
Multiple R-Squared: 0.1766, Adjusted R-squared: 0.1764
F-statistic: 917.5 on 1 and 4279 DF, p-value: < 2.2e-16

> summary(a3)

Call:
lm(formula = rmod$pgbus ~ -1 + rmod$g3)

Residuals:
    Min       1Q   Median       3Q      Max
-53.992   0.000   0.000   7.702  146.268

Coefficients:
Estimate Std. Error t value Pr(>|t|)
rmod$g3 0.433772  0.007606  57.03  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.59 on 4279 degrees of freedom
Multiple R-Squared: 0.4319, Adjusted R-squared: 0.4317
F-statistic: 3253 on 1 and 4279 DF, p-value: < 2.2e-16

> summary(a4)

Call:
lm(formula = rmod$pgauto ~ -1 + rmod$g4)

Residuals:
    Min       1Q   Median       3Q      Max
-209.364   0.000   0.000   4.386  207.326

Coefficients:
Estimate Std. Error t value Pr(>|t|)
rmod$g4 0.803922  0.005894  136.4  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 22.06 on 4279 degrees of freedom
Multiple R-Squared: 0.813, Adjusted R-squared: 0.813
F-statistic: 1.86e+04 on 1 and 4279 DF, p-value: < 2.2e-16
>
```

Predicciones Reparto Modal del Autobús y el Auto Privado

Las predicciones deberán realizarse para el horizonte 2008 y para el horizonte 2015:

1. El horizonte 2008 después de reestructurar la red de autobús de Tuvisa, donde no se esperan cambios remarcables en la generación/distribución de viajes respecto los datos del 2006, pero tiene consecuencia para la movilidad la introducción del tranvía.
2. El horizonte 2015 mantiene el tranvía y la reestructuración de la red de autobús urbano, pero ha de reflejar los cambios en la generación/atracción de viajes por zonas debidos a los cambios urbanísticos. Se aplicarán previamente al reparto modal, los modelos de generación y atracción por motivos de manera que faciliten los totales generados y atraídos por todos los modos y motivos a nivel de zona de transporte.

La aplicación del modelo de reparto modal requiere primeramente de la definición de las variables explicativas en los horizontes 2008 y 2015.

Especificación del archivo de datos para calibrar el reparto modal con la movilidad actual

Notación anglosajona: decimales con separador punto y sin separador-indicador de miles o millones. Archivo ASCII con primera fila con los nombres de las variables: cortos, sin mayúsculas ni acentos. Campos separados por un blanco (NO tabuladores) o bien formato .csv. Archivo con filas: ordenado por zonas de transporte, donde se consideran únicamente el ámbito interior del estudio.

Codificación indexada de las columnas: índice por modos j nº de modo (gran modo) numerados de 1 a J (J=4) ordenado el archivo por pareja origen-destino (i,j) .

Las variables OD de alternativa modal son: *ttt*, *tespera*, *tacces*, *tcoste*, *dista* (longitud del desplazamiento OD, en métrica norma 1 (Manhattan)). La presentación de los datos es agregada y por tanto los valores han de corresponder a las medias ponderadas OD, sin segregación por segmento, ni por motivo (éstos sólo se han empleado en los denominados modelos descriptivos).

Síntesis de columnas de entrada:

i : pareja OD de 1 a I

gi : Viajes totales pareja OD de 1 a I

ori: zona interior origen de los viajes

desti: zona interior destino de los viajes

X11 hasta X1J : J campos con variables de alternativa X1

...

Xk1 hasta XkJ: J campos con variables de alternativa Xk

| | |
|-----------------|--|
| od | Identificador de la pareja OD del desplazamiento o viaje (no etapa) |
| ori | Origen del viaje |
| desti | Destino del viaje |
| god | Viajes totales diarios (modos todos y motivos todos) en escenario por pareja OD |
| dista | Distancia Manhattan en km entre origen y destino . Formato con 2 decimales |
| ttt1 | Tiempo total de viaje OD ponderado para el modo 1 (des d'inici 1a Etapa hasta final darrera Etapa) en minutos con dos decimales |
| ... | |
| tttJ | Tiempo total de viaje OD ponderado para el modo J ... |
| tespera1 | Tiempo de espera OD ponderado para el modo 1 (vem o bus o coche 0) en minutos con dos decimales . En bus està lligat al 50% de la frecuencia en minutos (dos decimales) |
| ... | |
| tesperaJ | Tiempo de espera OD ponderado para el modo J ... |
| tacces1 | Tiempo de acceso hasta el modo OD ponderado para el modo 1 (a pie 0) en minutos con dos decimales . En viajes donde hay una etapa de bus esta vinculado al tiempo de las etapas a pie en minutos (dos decimales) . En viajes donde hay una etapa de coche està vinculado al tiempo de acceso (ida o vuelta) al lugar de aparcamiento y búsqueda de aparcamiento. |
| ... | |
| taccesJ | Tiempo de acceso hasta el modo OD ponderado para el modo J ... |
| tcost1 | Coste del viaje en modo 1 . A pie 0... vem o bus precio 1 viaje en bonus tipo T10 . Coche considera el gasto del combustible más mantenimiento aproximadamente 0,2€/km |
| ... | |
| tcostJ | Coste del viaje en modo J ... |

Variables de salida (un registro por pareja OD):

| | |
|----------------|---|
| od | Identificador de la pareja OD del desplazamiento o viaje (no etapa) |
| ori | Origen del viaje |
| desti | Destino del viaje |
| g1 | Número viajes modo 1 en pareja OD |
| g2 | Número viajes modo 2 en pareja OD |
| ... | |
| gJ | Número viajes modo J en pareja OD |
| emueco | Utilidad máxima esperada en el nido no motorizado ECO en pareja OD |
| emuneco | Utilidad máxima esperada en el nido MOTORIZADOS en pareja OD |
| demueco | Diferencia emuneco-emueco |
| geco | Número viajes modos 1 y 2 (a pie más bici, ECO) en pareja OD |
| gneco | Número viajes modos 3 y 4 (MOTORIZADOS) en pareja OD |
| p1 | Probabilidad marginal del modo 1 en pareja OD |
| p2 | Probabilidad marginal del modo 2 en pareja OD |

| | |
|---------------|---|
| ... | |
| pJ | Probabilidad marginal del modo J = 4 en pareja OD |
| peco | Probabilidad marginal modos 1 y 2 (a pie más bici, ECO) en pareja OD |
| pneco | Probabilidad marginal de modos 3 y 4 (MOTORIZADOS) en pareja OD |
| pcbus | Probabilidad condicional del modo 3 (bus) dado uso modo motorizado en pareja OD |
| pcauto | Probabilidad marginal del modo 4 (auto) dado uso modo motorizado en pareja OD |

En este punto se recuerdan la construcción de variables imprescindibles para obtener el buen funcionamiento del reparto modal:

- Modelo sm01.2: Modelo *logit* binario para el reparto modal en el nido ECO a pie y bus con referencia a pie y por tanto el predictor lineal puede interpretarse como la diferencia de utilidad de la bicicleta con respecto al a pie. Permite construir la utilidad esperada del nido ECO (variable emueco) y las probabilidades OD predichas condicionales de los modos a pie y bici (pcpie y pcbici, no relevantes para la modelística global y por tanto no forman parte de las variables de salida).
- Modelo sm02.1: Modelo *logit* binario para el reparto modal en el nido MOTORIZADOS bus y coche con referencia el auto privado y por tanto el predictor lineal puede interpretarse como la diferencia de utilidad del autobús con respecto al vehículo privado. Permite construir la utilidad esperada del nido MOTORIZADOS (variable emuneco) y las probabilidades OD predichas condicionales de los modos bus y auto (pcbus y pcauto).
- Modelo sm1.0: Modelo logia binario para el reparto modal entre modo motorizados (MOTORIZADOS) y no motorizados (ECO) con referencia los modos no motorizados (ecológicos). Permite obtener las predicciones de probabilidades OD marginal de nido para selección modal de modos ECO y MOTORIZADOS en el escenario. Permite a su vez obtener las probabilidades OD absolutas predichas a partir de las probabilidades condicionales por nido y las marginales de nido.

```
> summary(sm01.2)

Call:
glm(formula = cbind(round(g2), round(g1)) ~ dista, family = binomial,
     data = df01)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-15.721  -3.565  -2.557  -1.902   35.648

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.175571   0.018956  -220.3  <2e-16 ***
dista        0.846025   0.008023   105.4  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 102774 on 2644 degrees of freedom
Residual deviance:  91736 on 2643 degrees of freedom
AIC: 92026

Number of Fisher Scoring iterations: 5

> summary(sm02.1)
```

```
Call:
glm(formula = cbind(g3, g4) ~ dista + ttt.3.4 + tcost.3.4, family = binomial,
     data = df02)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-16.942  -5.913  -4.222   5.493  27.628
```

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.0261219  0.0279216  -0.936   0.350
dista        -0.3524887  0.0117475 -30.005  <2e-16 ***
ttt.3.4      -0.0437620  0.0004822 -90.749  <2e-16 ***
tcost.3.4    -3.1356381  0.0624084 -50.244  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
    Null deviance: 149718 on 2118 degrees of freedom
Residual deviance: 138567 on 2115 degrees of freedom
AIC: 139237
```

Number of Fisher Scoring iterations: 5

```
> summary(sm1.0)
```

```
Call:
glm(formula = cbind(g3 + g4, g1 + g2) ~ tespera3 + taces4 +
     dsemuneco, family = binomial, data = rmod)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-22.2266  -0.3980  -0.2570   0.3190  26.2131
```

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.983728   0.041612  23.641  <2e-16 ***
tespera3     -0.044760   0.002971 -15.065  <2e-16 ***
taces4       -0.077025   0.009187  -8.384  <2e-16 ***
dsemuneco     0.763411   0.005954  128.212  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(Dispersion parameter for binomial family taken to be 1)

```
    Null deviance: 434062 on 4279 degrees of freedom
Residual deviance:  17763 on 4276 degrees of freedom
AIC: 20305
```

Number of Fisher Scoring iterations: 6

La manera más cómoda de aplicar los modelos detallados sería utilizar el paquete estadístico de distribución gratuita R, sin embargo por comodidad se van a indicar los cálculos a efectuar para la obtención de las variables de salida por el orden adecuado y como si se estuviera en una hoja de cálculo tipo excel.

| | |
|----------------|---|
| od | Identificador de la pareja OD del desplazamiento o viaje (no etapa) |
| ori | Origen del viaje |
| desti | Destino del viaje |
| pcbus | Probabilidad condicional del modo 3 (bus) dado uso modo motorizado en pareja OD |
| pcauto | Probabilidad marginal del modo 4 (auto) dado uso modo motorizado en pareja OD |
| emuneco | Utilidad máxima esperada en el nido MOTORIZADOS en pareja OD |
| pcpie | Probabilidad condicional del modo 1 dado uso modo no motorizado en pareja OD |
| pcbici | Probabilidad marginal del modo 2 dado uso modo no motorizado en pareja OD |
| emueco | Utilidad máxima esperada en el nido no motorizado ECO en pareja OD |
| demueco | Diferencia emuneco-emueco |
| peco | Probabilidad marginal modos 1 y 2 (a pie más bici, ECO) en pareja OD |
| pneco | Probabilidad marginal de modos 3 y 4 (MOTORIZADOS) en pareja OD |
| p1 | Probabilidad marginal del modo 1 (a pie) en pareja OD |
| p2 | Probabilidad marginal del modo 2 (bicicleta) en pareja OD |
| p3 | Probabilidad marginal del modo 3 (autobús) en pareja OD |
| p4 | Probabilidad marginal del modo 4 (auto privado) en pareja OD |
| | |
| geco | Número viajes modos 1 y 2 (a pie más bici, ECO) en pareja OD |
| gneco | Número viajes modos 3 y 4 (MOTORIZADOS) en pareja OD |
| g1 | Número viajes modo 1 (a pie) en pareja OD |
| g2 | Número viajes modo 2 (bicicleta) en pareja OD |
| ... | |
| gJ | Número viajes modo J (=4, auto privado) en pareja OD |

Detalle de cálculo:

| | <p>Modelo sm02.1:</p> <pre>glm(formula = cbind(g3, g4) ~ dista + ttt.3.4 + tcost.3.4, family = binomial, data = df02)</pre> <p>Coefficients:</p> <table border="1"> <thead> <tr> <th></th> <th>Estimate</th> <th>Std. Error</th> <th>z value</th> <th>Pr(> z)</th> </tr> </thead> <tbody> <tr> <td>(Intercept)</td> <td>-0.0261219</td> <td>0.0279216</td> <td>-0.936</td> <td>0.350</td> </tr> <tr> <td>dista</td> <td>-0.3524887</td> <td>0.0117475</td> <td>-30.005</td> <td><2e-16 ***</td> </tr> <tr> <td>ttt.3.4</td> <td>-0.0437620</td> <td>0.0004822</td> <td>-90.749</td> <td><2e-16 ***</td> </tr> <tr> <td>tcost.3.4</td> <td>-3.1356381</td> <td>0.0624084</td> <td>-50.244</td> <td><2e-16 ***</td> </tr> </tbody> </table> | | Estimate | Std. Error | z value | Pr(> z) | (Intercept) | -0.0261219 | 0.0279216 | -0.936 | 0.350 | dista | -0.3524887 | 0.0117475 | -30.005 | <2e-16 *** | ttt.3.4 | -0.0437620 | 0.0004822 | -90.749 | <2e-16 *** | tcost.3.4 | -3.1356381 | 0.0624084 | -50.244 | <2e-16 *** |
|-----------------|--|------------|----------|------------|---------|----------|-------------|------------|-----------|--------|-------|-------|------------|-----------|---------|------------|---------|------------|-----------|---------|------------|-----------|------------|-----------|---------|------------|
| | Estimate | Std. Error | z value | Pr(> z) | | | | | | | | | | | | | | | | | | | | | | |
| (Intercept) | -0.0261219 | 0.0279216 | -0.936 | 0.350 | | | | | | | | | | | | | | | | | | | | | | |
| dista | -0.3524887 | 0.0117475 | -30.005 | <2e-16 *** | | | | | | | | | | | | | | | | | | | | | | |
| ttt.3.4 | -0.0437620 | 0.0004822 | -90.749 | <2e-16 *** | | | | | | | | | | | | | | | | | | | | | | |
| tcost.3.4 | -3.1356381 | 0.0624084 | -50.244 | <2e-16 *** | | | | | | | | | | | | | | | | | | | | | | |
| pcbus | <pre>ubus= -0.02612 -0.35249 dista -0.04376 (ttt3) -3.13564 (tcost3) uauto=-0.04376 (ttt4) -3.13564 (tcost4) pcbus = exp(ubus)/(exp(uauto)+ exp(ubus))</pre> | | | | | | | | | | | | | | | | | | | | | | | | | |
| pcauto | 1-pcbus | | | | | | | | | | | | | | | | | | | | | | | | | |
| emuneco | ln(exp(uauto)+ exp(ubus)) | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <pre>> summary(sm01.2) glm(formula = cbind(round(g2), round(g1)) ~ dista, family = binomial, data = df01) Coefficients: Estimate Std. Error z value Pr(> z) (Intercept) -4.175571 0.018956 -220.3 <2e-16 *** dista 0.846025 0.008023 105.4 <2e-16 ***</pre> | | | | | | | | | | | | | | | | | | | | | | | | | |
| pcpie | <pre>ubici = -4.17557 + 0.84603 dista upie = 0 pcpie = exp(upie) / (exp(upie) + exp(ubici))</pre> | | | | | | | | | | | | | | | | | | | | | | | | | |
| pcbici | 1-pcpie | | | | | | | | | | | | | | | | | | | | | | | | | |
| emueco | ln(exp(upie) + exp(ubici)) | | | | | | | | | | | | | | | | | | | | | | | | | |
| demuneco | emuneco - emueco | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <pre>> summary(sm1.0) glm(formula = cbind(g3 + g4, g1 + g2) ~ tespera3 + taces4 + demuneco, family = binomial, data = rmod) Coefficients: Estimate Std. Error z value Pr(> z) (Intercept) 0.983728 0.041612 23.641 <2e-16 *** tespera3 -0.044760 0.002971 -15.065 <2e-16 *** taces4 -0.077025 0.009187 -8.384 <2e-16 *** demuneco 0.763411 0.005954 128.212 <2e-16 ***</pre> | | | | | | | | | | | | | | | | | | | | | | | | | |
| pneco | uneco = 0.98373 -0.04476 tespera3 -0.07703 taces4 + 0.76341 demueco | | | | | | | | | | | | | | | | | | | | | | | | | |

| | |
|--------------|--|
| | ueco = 0 pcneco = $\exp(\text{uneco}) / (\exp(\text{uneco}) + \exp(\text{ueco}))$ |
| peco | 1-pneco |
| p1 | peco * pcpie |
| p2 | peco * pcbici |
| p3 | pneco * pcbus |
| p4 | pneco * pcauto |
| | # god es el total de viajes en la pareja OD (motivos y modos todos) |
| geco | god * peco |
| gneco | god * pneco |
| g1 | god * p1 |
| g2 | god * p2 |
| g3 | god * p3 |
| g4 | god * p4 |

El modelo de cálculo de reparto modal se aplicará a la movilidad total OD procedente de la Encuesta de Movilidad de APPEND 2006 y estimará la matriz OD modal ante la reestructuración de la red autobuses de Tuvisa con el tranvía (horizonte 2008). El reparto modal debe aplicarse en la submatriz interior del ámbito (zonas 1 a 38), en el resto del ámbito el reparto modal se contemplará idéntico al existente en la actualidad.

El modelo de generación/atracción más distribución debe aplicarse previamente al reparto modal en el horizonte 2015.