Travel Mode - Multinomial Logit Model

February 5, 2020

For multinomial models that include category–specific as well as global effects the function "mlogit" from the library "mlogit" can be used.

```
> library(mlogit)
```

The "Travel Mode"–data are stored in the "Edcat"–package and can be loaded by the following command.

```
> data(ModeChoice, package="Ecdat")
```

For the use of the function "mlogit" an appropriate data set has to be built. This is done by use of the function "mlogit.data".

```
> travel.long <- mlogit.data(ModeChoice, choice="mode", shape="long", alt.levels=
+ c("air","train","bus","car"))
```

Now the model can be fitted. In the formula first the category–specific effects and then, separated by "|", the global effects are specified.

```
> travel.kat.id <- mlogit(mode ~ invt + gc|hinc, data=travel.long)
> summary(travel.kat.id)

Call:
mlogit(formula = mode ~ invt + gc | hinc, data = travel.long,
    method = "nr")

Frequencies of alternatives:
    air    train      bus      car
0.27619 0.30000 0.14286 0.28095

nr method
4 iterations, 0h:0m:0s
g'(-H)^-1g = 0.00216
successive function values within tolerance limits

Coefficients :
                  Estimate Std. Error z-value  Pr(>|z|)
train:(intercept)  3.5250366  0.6549825  5.3819 7.371e-08 ***
bus:(intercept)    2.2782769  0.7176686  3.1746  0.001501 **
car:(intercept)    1.5334957  0.7065856  2.1703  0.029985 *
invt              -0.0031266  0.0009548 -3.2746  0.001058 **
```

```
gc                -0.0016225  0.0055279 -0.2935  0.769130
train:hinc        -0.0569409  0.0124103 -4.5882 4.471e-06 ***
bus:hinc          -0.0355771  0.0131492 -2.7056  0.006817 **
car:hinc          -0.0023652  0.0104475 -0.2264  0.820898
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -250.17
McFadden R^2:  0.11839
Likelihood ratio test : chisq = 67.186 (p.value = 3.9423e-13)
```

Now the same model is fitted with the package "VGAM".

```
> library(VGAM)
```

At first the data need to be prepared adequately to be ready for use with the function "vglm".

```
> travelmode <- matrix(ModeChoice$mode, byrow = T, ncol = 4)
> colnames(travelmode) <- c("air","train","bus","car")
> travelhinc <- matrix(ModeChoice$hinc, byrow = T, ncol = 4)
> travelhinc <- travelhinc[,1]
> travelinvt <- matrix(ModeChoice$invt, byrow = T, ncol = 4)
> colnames(travelinvt) <- c("invtair","invttrain","invtbus","invtcar")
> travelgc <- matrix(ModeChoice$gc, byrow = T, ncol = 4)
> colnames(travelgc) <- c("gcair","gctrain","gcbus","gccar")
> travelinvt <- sweep(travelinvt[,-1], 1, travelinvt[,1])
> travelgc <- sweep(travelgc[,-1], 1, travelgc[,1])
> Invt <- travelinvt[,1]
> Gc <- travelgc[,1]
> traveldat <- cbind(travelhinc, travelinvt, Invt, travelgc, Gc)
> traveldat <- as.data.frame(traveldat)
```

Now the model can be fitted.

```
> fit <- vglm(travelmode ~ Invt + Gc + travelhinc,
+             multinomial(parallel = FALSE ~ travelhinc, refLevel = 1),
+             xij = list(Invt ~ invttrain + invtbus + invtcar,
+                        Gc ~ gctrain + gcbus + gccar),
+             form2 = ~ Invt + invttrain + invtbus + invtcar +
+                        Gc + gctrain + gcbus + gccar + travelhinc,
+             data = traveldat, trace = TRUE)

VGLM    linear loop  1 :  deviance = 501.46294
VGLM    linear loop  2 :  deviance = 500.33384
VGLM    linear loop  3 :  deviance = 500.33167
VGLM    linear loop  4 :  deviance = 500.33167

> summary(fit)

Call:
vglm(formula = travelmode ~ Invt + Gc + travelhinc, family = multinomial(parallel = FALSE
```

```
       travelhinc, refLevel = 1), data = traveldat, form2 = ~Invt +
       invttrain + invtbus + invtcar + Gc + gctrain + gcbus + gccar +
       travelhinc, xij = list(Invt ~ invttrain + invtbus + invtcar,
       Gc ~ gctrain + gcbus + gccar), trace = TRUE)

Pearson residuals:
                          Min      1Q  Median      3Q   Max
log(mu[,2]/mu[,1]) -2.782 -0.5490 -0.2578  0.7185 3.302
log(mu[,3]/mu[,1]) -2.139 -0.4076 -0.2213 -0.1314 4.308
log(mu[,4]/mu[,1]) -2.005 -0.6589 -0.2444  0.8660 5.128

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept):1  3.5250538  0.6549818    5.382 7.37e-08 ***
(Intercept):2  2.2782935  0.7176672    3.175  0.00150 **
(Intercept):3  1.5334984  0.7065854    2.170  0.02998 *
Invt          -0.0031266  0.0009548   -3.275  0.00106 **
Gc            -0.0016225  0.0055279   -0.294  0.76913
travelhinc:1  -0.0569415  0.0124103   -4.588 4.47e-06 ***
travelhinc:2  -0.0355781  0.0131491   -2.706  0.00682 **
travelhinc:3  -0.0023652  0.0104474   -0.226  0.82090
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

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

Residual deviance: 500.3317 on 622 degrees of freedom

Log-likelihood: -250.1658 on 622 degrees of freedom

Number of Fisher scoring iterations: 4

No Hauck-Donner effect found in any of the estimates


Reference group is level  1  of the response

> summary(travel.kat.id)

Call:
mlogit(formula = mode ~ invt + gc | hinc, data = travel.long,
    method = "nr")

Frequencies of alternatives:
    air   train     bus     car
0.27619 0.30000 0.14286 0.28095

nr method
4 iterations, 0h:0m:0s
```

```
g'(-H)^-1g = 0.00216
successive function values within tolerance limits

Coefficients :
                  Estimate Std. Error z-value  Pr(>|z|)
train:(intercept)  3.5250366  0.6549825   5.3819 7.371e-08 ***
bus:(intercept)    2.2782769  0.7176686   3.1746  0.001501 **
car:(intercept)    1.5334957  0.7065856   2.1703  0.029985 *
invt              -0.0031266  0.0009548  -3.2746  0.001058 **
gc                -0.0016225  0.0055279  -0.2935  0.769130
train:hinc        -0.0569409  0.0124103  -4.5882 4.471e-06 ***
bus:hinc          -0.0355771  0.0131492  -2.7056  0.006817 **
car:hinc          -0.0023652  0.0104475  -0.2264  0.820898
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -250.17
McFadden R^2:  0.11839
Likelihood ratio test : chisq = 67.186 (p.value = 3.9423e-13)
```

At last we compare the coefficients of the two fitted models.

```
> summary(travel.kat.id)$CoefTable

                     Estimate    Std. Error      z-value       Pr(>|z|)
train:(intercept)  3.525036582 0.6549824780   5.3818792 7.371222e-08
bus:(intercept)    2.278276953 0.7176686437   3.1745527 1.500676e-03
car:(intercept)    1.533495711 0.7065856497   2.1702899 2.998489e-02
invt              -0.003126577 0.0009547997  -3.2745899 1.058154e-03
gc                -0.001622510 0.0055279076  -0.2935124 7.691305e-01
train:hinc        -0.056940856 0.0124103406  -4.5881784 4.471305e-06
bus:hinc          -0.035577091 0.0131492259  -2.7056415 6.817260e-03
car:hinc          -0.002365193 0.0104474472  -0.2263895 8.208985e-01

> summary(fit)@coef3

                Estimate    Std. Error   z value     Pr(>|z|)
(Intercept):1  3.525053759 0.6549818022  5.381911 7.369921e-08
(Intercept):2  2.278293544 0.7176672465  3.174582 1.500525e-03
(Intercept):3  1.533498382 0.7065854109  2.170294 2.998455e-02
Invt          -0.003126583 0.0009547997 -3.274596 1.058133e-03
Gc            -0.001622528 0.0055279012 -0.293516 7.691278e-01
travelhinc:1  -0.056941460 0.0124102999 -4.588242 4.469941e-06
travelhinc:2  -0.035578093 0.0131491108 -2.705741 6.815211e-03
travelhinc:3  -0.002365198 0.0104474488 -0.226390 8.208981e-01
```