Applied Microeconometrics - Assignment 3

Walter Verwer (589962) & Bas Machielsen (590049)

October 8, 2021

1. Estimate a multinomial logit model with a choice probability . . ..

We estimate such that each school has a unique intercept and unique β-coefficient. It seems that individual specific factors have little predictive power, no matter what school. The only exception is school number 7, which is decidedly more attractive for students with a higher test score.

```r
dataset_rdy <- mlogit::mlogit.data(dataset,
                    choice = "choice",
                    shape = "long",
                    alt.var = "school",
                    chid.var = "id"
                    )

privet <- mlogit::mlogit(formula = choice ~  0 | 1 + gender + testscore , data = dataset_rdy)
```

```r
modelsummary(privet,
             stars = c("*" = 0.1, "**" = 0.05, "***" = 0.01),
             gof_omit = c("BIC|rho"),
             title = "Multinomial Logit",
             out = "kableExtra") %>%
  kableExtra::kable_styling(latex_options = c("hold_position"), font_size = 7)
```

Table 1: Multinomial Logit

	Model 1
(Intercept) $\times$ 2	29.043
	(29.687)
(Intercept) $\times$ 7	-63.955**
	(30.896)
(Intercept) $\times$ 8	-121.599
	(137.519)
(Intercept) $\times$ 9	12.135
	(31.489)
(Intercept) $\times$ 14	3.838
	(36.458)
(Intercept) $\times$ 24	41.401
	(93.355)
(Intercept) $\times$ 25	-18.484
	(32.529)
(Intercept) $\times$ 28	33.928
	(42.547)
(Intercept) $\times$ 29	-43.143
	(37.071)
gender $\times$ 2	-0.469
	(0.442)
gender $\times$ 7	-0.419
	(0.436)
gender $\times$ 8	-0.255
	(1.468)
gender $\times$ 9	-0.217
	(0.466)
gender $\times$ 14	-0.013
	(0.538)
gender $\times$ 24	-0.195
	(1.468)
gender $\times$ 25	-0.232
	(0.471)
gender $\times$ 28	0.491
	(0.672)
gender $\times$ 29	-0.359
	(0.519)
testscore $\times$ 2	-0.051
	(0.055)
testscore $\times$ 7	0.120**
	(0.057)
testscore $\times$ 8	0.218
	(0.251)
testscore $\times$ 9	-0.021
	(0.058)
testscore $\times$ 14	-0.007
	(0.067)
testscore $\times$ 24	-0.081
	(0.172)
testscore $\times$ 25	0.035
	(0.060)
testscore $\times$ 28	-0.064
	(0.078)
testscore $\times$ 29	0.080
	(0.068)
Num.Obs.	422
AIC	1672.6
Log.Lik.	-809.325

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

2. Estimate a conditional logit model with the choice probability ....

```
zdravstvuy <- mlogit::mlogit(formula = choice ~ distance + sibling, data = dataset_rdy)
```

```
modelsummary(zdravstvuy,
             stars = c("*" = 0.1, "**" = 0.05, "***" = 0.01),
             gof_omit = c("BIC|rho"),
             title = "Conditional Logit",
             out = "kableExtra") %>%
  kableExtra::kable_styling(latex_options = c("hold_position"), font_size = 6)
```

Table 2: Conditional Logit

	Model 1
(Intercept) × 2	1.427***
	(0.239)
(Intercept) × 7	1.236***
	(0.232)
(Intercept) × 8	−2.143***
	(0.743)
(Intercept) × 9	0.901***
	(0.256)
(Intercept) × 14	−0.152
	(0.278)
(Intercept) × 24	−2.615***
	(0.739)
(Intercept) × 25	0.844***
	(0.257)
(Intercept) × 28	−0.590*
	(0.334)
(Intercept) × 29	0.281
	(0.279)
distance	−0.246***
	(0.030)
sibling	1.689***
	(0.190)
Num.Obs.	422
AIC	1501.8
Log.Lik.	−739.893

* p < 0.1, ** p < 0.05, *** p < 0.01

2.2 How do you interpret the estimated coefficients? Based on your estimates, do you think that students prefer schools that are closer to their home? Do students prefer the high school(s) that their sibling(s) attend?

We can interpret the estimates either by constructing an odds ratio, comparing the choice of an individual of two schools, but we can also calculate the point estimate, and compute the marginal effects, which are dependent on the β-coefficient magnitudes. Hence, the precise magnitude of the derivatives changes according to individual and school, but the sign of the derivative can be easily observed. That leads us to conclude that yes and yes, the closer the school, the higher the probability of an individual choosing that school. The effect's magnitude is, however, much more limited in comparison to the sibling effect, which is very strong, and in the hypothesized direction: if a sibling is in a particular school, the probability increases by a large magnitude.

3. Now combine the previous specification, and estimate a mixed logit model that includes both individual-specific covariates (such as test score and gender) and alternative-varying regressors (such as distance and sibling).

```
hoi <- mlogit::mlogit(formula = choice ~ distance + sibling | gender + testscore, data = dataset_rdy)
```

```
modelsummary(hoi,
             stars = c("*" = 0.1, "**" = 0.05, "***" = 0.01),
             gof_omit = c("BIC|rho"),
             title = "Mixed Logit",
             out = "kableExtra") %>%
  kableExtra::kable_styling(latex_options = c("hold_position"), font_size = 5)
```

Table 3: Mixed Logit

	Model 1
(Intercept) × 2	15.680
	(30.881)
(Intercept) × 7	−80.108**
	(32.740)
(Intercept) × 8	−138.617
	(137.454)
(Intercept) × 9	−13.277
	(34.200)
(Intercept) × 14	−7.090
	(38.171)
(Intercept) × 24	11.942
	(95.417)
(Intercept) × 25	−49.272
	(34.930)
(Intercept) × 28	16.530
	(43.405)
(Intercept) × 29	−75.215*
	(39.585)
distance	−0.261***
	(0.031)
sibling	1.675***
	(0.193)
gender × 2	−0.617
	(0.464)
gender × 7	−0.594
	(0.468)
gender × 8	−0.441
	(1.488)
gender × 9	−0.619
	(0.508)
gender × 14	−0.002
	(0.560)
gender × 24	−0.375
	(1.481)
gender × 25	−0.451
	(0.507)
gender × 28	0.401
	(0.692)
gender × 29	−0.669
	(0.554)
testscore × 2	−0.026
	(0.057)
testscore × 7	0.150**
	(0.060)
testscore × 8	0.250
	(0.251)
testscore × 9	0.027
	(0.063)
testscore × 14	0.013
	(0.070)
testscore × 24	−0.026
	(0.176)
testscore × 25	0.092
	(0.064)
testscore × 28	−0.032
	(0.080)
testscore × 29	0.139*
	(0.073)
Num.Obs.	422
AIC	1502.9
Log.Lik.	−722.444

* p < 0.1, ** p < 0.05, *** p < 0.01

4. Explain the independence of irrelevant alternatives (IIA) assumption and discuss if it is relevant in this context

In any multinomial logit model, if we take the ratio of the probabilities of choosing two alternatives (either for a given individual, or in the case in which individual effects are absent), we find that this ratio is only dependent on the the characteristics of the individuals *and/or* the characteristics of the alternatives, but not on the characteristics of other alternatives. We think it is relevant in this situation: we are dealing with more than two situations, and even while IIA is required by theoretical choice models, behaviorally, it can be violated. For example, empirically, it is plausible that one's (reported) preference might change from alternative schools A to B if you augment the choice set from $\{A, B\}$ to $\{A, B, C\}$. This situation is ruled-out a priori by the logit model error structure of probabilistic choice. Theoretically, the objective is to accurately explain school choice, so if IIA is routinely being violated in empirical choices, our modeling strategy should take that into account.

A concrete example in this situation when IIA could be violated is as follows. Suppose there are three schools: A, B, and C. School A and school B are very close to each other, and school C is further away. If now school A is removed from the set of alternatives, then by IIA the students attending school A should divide themselves proportionally over school B and C. However, it is more likely that a greater fraction of students from school A now choose school B since these schools are closer to eachother and distance is an important characteristic for the choice. What this indicates is that in our setting it might be the case that alternatives (schools) are in relation to eachother, which is a violation of the independent and irrelevant alternatives assumption.