

# Discrete Choice and Count Models

Antonio Jurlina

4/19/2021

**Problem 1.** Recall that the choice probability under the nested logit model takes the form,

$$P_{ij} = P_{iB_k} \times P_{ij|B_k}$$

where  $P_{iB_k}$  denotes the probability of person  $i$  choosing nest  $k$

$$P_{iB_k} = \frac{e^{W_{ik} + \lambda_k \Phi_{ik}}}{\sum_{l=1}^K e^{W_{il} + \lambda_l \Phi_{il}}}$$

where the *inclusive value* takes the form

$$\Phi_{ik} = \ln \sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}}$$

and  $P_{ij|B_k}$  denotes the probability of person  $i$  choosing alternative  $j$  conditioned on choosing nest  $k$

$$P_{ij|B_k} = \frac{e^{\frac{Y_{ij}}{\lambda_k}}}{\sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}}}$$

a) Here, I use algebra to demonstrate that the nested logit model reduces to the multinomial logit model if  $\lambda_k = 1$  for all  $k$  - the alternatives within every nest are independent of each other. We start with the nested logit choice probability:

$$\begin{aligned} P_{ij} &= P_{iB_k} \times P_{ij|B_k} \\ P_{ij} &= \frac{e^{W_{ik} + \lambda_k \Phi_{ik}}}{\sum_{l=1}^K e^{W_{il} + \lambda_l \Phi_{il}}} \times \frac{e^{\frac{Y_{ij}}{\lambda_k}}}{\sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}}} \\ P_{ij} &= \frac{e^{W_{ik} + \lambda_k \ln \sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}}}}{\sum_{l=1}^K e^{W_{il} + \lambda_l \ln \sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}}}} \times \frac{e^{\frac{Y_{ij}}{\lambda_k}}}{\sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}}} \end{aligned}$$

*Note:*  $e^{x + \ln(b)} = e^x e^{\ln(b)} = e^x e^{\ln b^c} = e^x b^c$

$$P_{ij} = \frac{e^{W_{ik}} \left[ \sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}} \right]^{\lambda_k}}{\sum_{l=1}^K e^{W_{il}} \left[ \sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_l}} \right]^{\lambda_l}} \times \frac{e^{\frac{Y_{ij}}{\lambda_k}}}{\sum_{m \in B_k} e^{\frac{Y_{im}}{\lambda_k}}} \times \frac{e^{\frac{W_{ik}}{\lambda_k}}}{e^{\frac{W_{ik}}{\lambda_k}}}$$

$$P_{ij} = \frac{\left[ \sum_{m \in B_k} e^{\frac{W_{ik} + Y_{im}}{\lambda_k}} \right]^{\lambda_k}}{\sum_{l=1}^k \left[ \sum_{m \in B_k} e^{\frac{W_{il} + Y_{im}}{\lambda_l}} \right]^{\lambda_l}} \times \frac{e^{\frac{W_{ik} + Y_{ij}}{\lambda_k}}}{\sum_{m \in B_k} e^{\frac{W_{ik} + Y_{im}}{\lambda_k}}}$$

Note:  $W_{ik} + Y_{ij} = V_{ik}$

$$P_{ij} = \frac{\left[ \sum_{m \in B_k} e^{\frac{V_{im}}{\lambda_k}} \right]^{\lambda_k}}{\sum_{l=1}^k \left[ \sum_{m \in B_k} e^{\frac{V_{im}}{\lambda_l}} \right]^{\lambda_l}} \times \frac{e^{\frac{V_{ik}}{\lambda_k}}}{\sum_{m \in B_k} e^{\frac{V_{im}}{\lambda_k}}}$$

$$P_{ij} = \frac{e^{\frac{V_{ik}}{\lambda_k}} \left[ \sum_{m \in B_k} e^{\frac{V_{im}}{\lambda_k}} \right]^{\lambda_k - 1}}{\sum_{l=1}^k \left[ \sum_{m \in B_k} e^{\frac{V_{im}}{\lambda_l}} \right]^{\lambda_l}}$$

If we set  $\lambda_k = 1$ , we get

$$P_{im} = \frac{e^{V_{ik}}}{\sum_m e^{V_{im}}}$$

which is the choice probability for the multinomial logit.

- b) Then, I show that the nested logit model also reduces to the multinomial logit model if all the nests  $B_k$  ( $\forall k$ ) are singletons, i.e., each choice alternative is contained in its own nest.

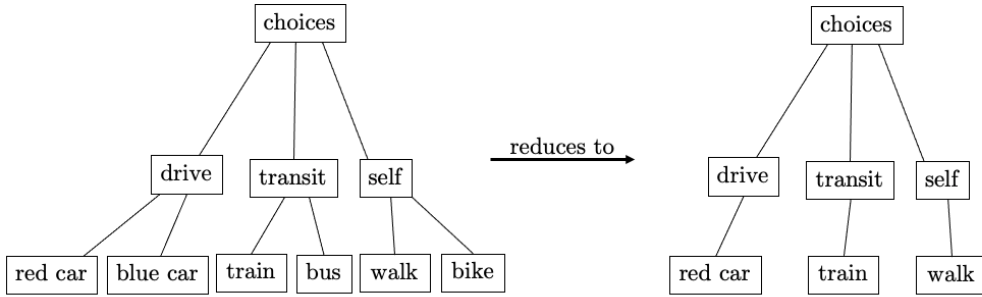


Figure 1: Reduction to singleton nests

Consider a discrete choice model in which each nest contains one choice (Figure 1). This means that when looking at  $P_{ij}$  as derived in part a), we can see that the within-nest summation,  $\sum_{m \in B_k}$ , is actually only summing over one variable, rendering the summation symbol unnecessary. That is,

$$P_{ij} = \frac{e^{\frac{V_{ik}}{\lambda_k}} \left[ \sum_{m \in B_k} e^{\frac{V_{im}}{\lambda_k}} \right]^{\lambda_k - 1}}{\sum_{l=1}^k \left[ \sum_{m \in B_k} e^{\frac{V_{im}}{\lambda_l}} \right]^{\lambda_l}}$$

reduces to

$$\begin{aligned}
P_{ij} &= \frac{e^{\frac{V_{ik}}{\lambda_k}} \left[ e^{\frac{V_{im}}{\lambda_k}} \right]^{\lambda_k - 1}}{\sum_{l=1}^k \left[ e^{\frac{V_{il}}{\lambda_l}} \right]^{\lambda_l}} \\
P_{ij} &= \frac{e^{\frac{V_{ik}}{\lambda_k}} \left[ e^{\frac{V_{ik}\lambda_k - V_{ik}}{\lambda_k}} \right]}{\sum_{l=1}^k e^{V_{il}}} \\
P_{ij} &= \frac{e^{\frac{V_{ik}}{\lambda_k}} e^{V_{ik}} e^{-\frac{V_{ik}}{\lambda_k}}}{\sum_{l=1}^k e^{V_{il}}} \\
P_{ik} &= \frac{e^{V_{ik}}}{\sum_{l=1}^k e^{V_{il}}}
\end{aligned}$$

which is the choice probability for the multinomial logit.

**Problem 2.** The file **count.dta**, which is taken from Gurmu (1997), contains data for 485 household heads who may or may not have visited a doctor during a certain period of time. The variables in the model are:

- $y_i$  - Number of doctor visits
- $C_i$  - Number of children in the household
- $A_i$  - A measure of access to healthcare
- $H_i$  - A measure of health status

### Marginal Effects

First, I estimate a log-linear Poisson regression model to explain the variable  $y_i$  where

$$\lambda_i = e^{\beta_0 + \beta_1 C_i + \beta_2 A_i + \beta_3 H_i}$$

and the results are presented in Tables 1 and 2.

Table 1: Estimates

	Poisson	NB1	NB2
nkids	-0.176*** (0.032)	-0.108** (0.050)	-0.171*** (0.058)
access	0.937*** (0.193)	0.537 (0.332)	0.420 (0.373)
status	0.290*** (0.018)	0.264*** (0.033)	0.315*** (0.052)
cons	0.375*** (0.110)	0.417** (0.186)	0.561*** (0.212)

According to the Poisson model (Table 2), for every additional child, the number of doctor visits decreases by 28%, on average. For every unit increase in the access to health care index, the number of doctor visits increases by 150%, on average. For every unit improvement in health status, the number of doctor visits increases by 47%, on average. According to the Negative Binomial Type 1 model (Table 2), for every additional child, the number of doctor visits decreases by 17%, on average. For every unit increase in the access to health care index, the number of doctor visits increases by 87%, on average. For every unit improvement in health status, the number of doctor visits increases by 43%, on average. According to the Negative Binomial Type 2 model (Table 2), for every additional child, the number of doctor visits decreases by 28%, on average. For every unit increase in the access to health care index, the number of doctor visits increases by 69%, on average. For every unit improvement in health status, the number of doctor visits increases by 52%, on average.

Table 2: Marginal Effects

	Poisson	NB1	NB2
nkids	-0.283*** (0.052)	-0.174** (0.082)	-0.278*** (0.100)
access	1.509*** (0.315)	0.865 (0.538)	0.685 (0.608)
status	0.467*** (0.034)	0.426*** (0.061)	0.515*** (0.112)

## Overdispersion

In the presence of overdispersion, Poisson model estimates from maximum likelihood estimation will still be consistent. However, the standard errors are going to be incorrect, rendering any subsequent hypothesis testing (or p-value interpretation) difficult. This issue stems from the fact that in a Poisson model, the mean and the variance cannot be estimated separately from one another, and if the data set does not match this property, over (or under) dispersion is present.

We start by checking for any presence of overdispersion in the Poisson model with a simple rule of thumb test. If the ratio of variance to the mean (in the unconditional data) is larger than 2, it is likely that overdispersion is present. Indeed, the value of 6.7 suggests that there is likely overdispersion (Table 3). Furthermore, I run the Auxiliary Regression Test (proposed by Cameron and Trivedi in 1996). According to the results, we can reject the null that there is no overdispersion present (Table 3). Finally, we will check for the presence of overdispersion in both Negative Binomial model results. Again, according to the results in Table 3, we can reject the null hypothesis that there is no overdispersion in the simple Poisson specification, with high confidence. This indicates that the estimates from the Negative Binomial models are likely to be more efficient. This set of results consistently confirms that there is likely overdispersion present in the Poisson model which assumes mean and variance are equal and provides no correction for the standard errors.

Table 3: Overdispersion Test Results

test	value	std. error	statistic	distribution	p. value
Rule of Thumb	6.956	-	-	-	-
Auxiliary Regression	2.175	0.597	3.64	Student's t	0.000
NB1 Delta	3.088	0.404	601.24	Chi-square	0.000
NB2 Alpha	1.810	0.201	599.61	Chi-square	0.000

## Conclusion

The estimates from the two Negative Binomial models are more likely to be efficient than the ones from the Poisson. While all three are expected to be consistent, we can notice that statistical significance on the marginal effects across models does not hold for the *access* variable when the standard errors are corrected. For the variables *status* and *nkids*, statistical significance remains, and their values are mostly similar to the Poisson (thus supporting our consistency expectation). We can notice that even while the marginal effects remained mostly equal, standard errors increased slightly with the Negative Binomial Type 1 model. They increased even more with the Negative Binomial Type 2 model, which allows for even higher flexibility than the Type 1, relative to the Poisson model.

## Stata code

```

/******
> @ Authors: Antonio J @
> @ Date: 4/19/2021 @
> @ Filename: problem_set_04.do @
> @*****

/*-----
@ Loading data and initializing log file @
-----*/

// Clear previously stored data and set global options
cls
clear all
cd "antoniojurlina/Projects/advanced_econometrics/data"

use "count.dta" // load data

// Store results in a log file (diary)
cd "antoniojurlina/Projects/advanced_econometrics/output"
log using "project_04_log.txt", replace text

/*-----
@ Models @
-----*/

// Poisson
qui poisson visits nkids access status
estimates store poiss
predict visits_hat
qui margins, dydx(*) post
estimates store me_poiss

// Negative Binomial Type 1 model
qui nbreg visits nkids access status, dispersion(constant)
estimates store nb1
qui margins, dydx(*) post
estimates store me_nb1

// Negative Binomial Type 2 model
qui nbreg visits nkids access status, dispersion(mean)
estimates store nb2
qui margins, dydx(*) post
estimates store me_nb2

/*-----
@ Tests @
-----*/

// Rule of thumb
qui summarize visits
display "Rule of Thumb Test: " round(r(sd)^2/r(mean), 0.001)

// Auxiliary regression
generate aux = ((visits-visits_hat)^2-visits)/visits_hat

```

```

qui regress aux visits_hat, noconstant
display "Auxiliary Regression Test: " _b[visits_hat]

/*-----
@ Results                                     @
-----*/
// Generate a table of estimates
esttab poiss nb1 nb2, ///
    mtitles("Poisson" "NB1" "NB2") b(3) se(3) ///
    star(* 0.1 ** 0.05 *** 0.01) nogaps

// Generate a table of margins
esttab me_pois me_nb1 me_nb2, ///
    mtitles("Poisson" "NB1" "NB2") b(3) se(3) ///
    star(* 0.1 ** 0.05 *** 0.01) nogaps

// Close log file
log close

```

```

## R version 4.0.3 (2020-10-10)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Big Sur 10.16
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
## [1] knitr_1.31      forcats_0.5.0  stringr_1.4.0  dplyr_1.0.3
## [5] purrr_0.3.4     readr_1.4.0    tidyr_1.1.2    tibble_3.0.5
## [9] ggplot2_3.3.3   tidyverse_1.3.0
##
## loaded via a namespace (and not attached):
## [1] Rcpp_1.0.6      highr_0.8       cellranger_1.1.0 pillar_1.4.7
## [5] compiler_4.0.3  dbplyr_2.0.0    tools_4.0.3     digest_0.6.27
## [9] lubridate_1.7.9.2 jsonlite_1.7.2  evaluate_0.14    lifecycle_1.0.0
## [13] gtable_0.3.0    pkgconfig_2.0.3 rlang_0.4.10     reprex_0.3.0
## [17] cli_2.2.0       rstudioapi_0.13 DBI_1.1.1        yaml_2.2.1
## [21] haven_2.3.1     xfun_0.20       withr_2.4.1      xml2_1.3.2
## [25] httr_1.4.2      fs_1.5.0        hms_1.0.0        generics_0.1.0
## [29] vctrs_0.3.6     grid_4.0.3      tidyselect_1.1.0 glue_1.4.2
## [33] R6_2.5.0        fansi_0.4.2     readxl_1.3.1     rmarkdown_2.6
## [37] modelr_0.1.8    magrittr_2.0.1  backports_1.2.1  scales_1.1.1
## [41] ellipsis_0.3.1  htmltools_0.5.1.1 rvest_0.3.6      assertthat_0.2.1
## [45] colorspace_2.0-0 stringi_1.5.3    munsell_0.5.0    broom_0.7.5
## [49] crayon_1.4.1

```