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

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

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

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

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:

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

$$P_{ij} = \frac{e^{W_{ik}+\lambda_k \Phi_{ik}}}{\sum_{l=1}^{K} e^{W_{il}+\lambda_l \Phi_{il}}} \times \frac{e^{\frac{Y_{ij}}{\lambda_k}}}{\sum_{m \epsilon B_k} e^{\frac{Y_{im}}{\lambda_k}}}$$

$$P_{ij} = \frac{e^{W_{ik}+\lambda_k ln \sum_{m \epsilon B_k} e^{\frac{Y_{im}}{\lambda_k}}}}{\sum_{l=1}^{K} e^{W_{il}+\lambda_l ln \sum_{m \epsilon B_k} e^{\frac{Y_{im}}{\lambda_k}}}} \times \frac{e^{\frac{Y_{ij}}{\lambda_k}}}{\sum_{m \epsilon B_k} e^{\frac{Y_{im}}{\lambda_k}}}$$

Note: $e^{x+cln(b)} = e^x e^{cln(b)} = e^x e^{lnb^c} = e^x b^c$

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

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

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

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

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

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.

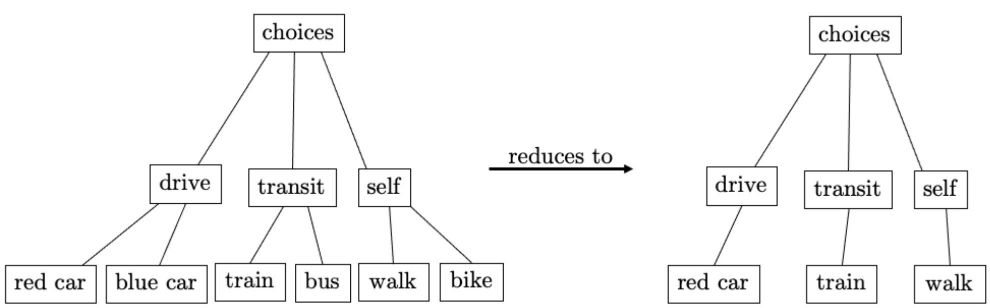

Figure 1: Reduction to singleton nests

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, $\sum_{m \epsilon B_k}$, is actually only summing over one variable, rendering the summation symbol unnecessary. That is,

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

reduces to

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

$$P_{ij} = \frac{e^{\frac{V_{ik}}{\lambda_k}} \left[e^{\frac{V_{ik}\lambda_k - V_{ik}}{\lambda_k}} \right]}{\sum_{l=1}^{k} e^{V_{il}}}$$

$$P_{ij} = \frac{e^{\frac{V_{ik}}{\lambda_k}} e^{V_{ik}} e^{-\frac{V_{ik}}{\lambda_k}}}{\sum_{l=1}^{k} e^{V_{il}}}$$

$$P_{ik} = \frac{e^{V_{ik}}}{\sum_{l=1}^{k} e^{V_{il}}}$$

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:

```
## 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        cellranger_1.1.0  pillar_1.4.7      compiler_4.0.3
##  [5] dbplyr_2.0.0      tools_4.0.3       digest_0.6.27     lubridate_1.7.9.2
##  [9] jsonlite_1.7.2    evaluate_0.14     lifecycle_1.0.0   gtable_0.3.0
## [13] pkgconfig_2.0.3   rlang_0.4.10      reprex_0.3.0      cli_2.2.0
## [17] rstudioapi_0.13   DBI_1.1.1         yaml_2.2.1        haven_2.3.1
## [21] xfun_0.20         withr_2.4.1       xml2_1.3.2        httr_1.4.2
## [25] fs_1.5.0          hms_1.0.0         generics_0.1.0    vctrs_0.3.6
## [29] grid_4.0.3        tidyselect_1.1.0  glue_1.4.2        R6_2.5.0
## [33] fansi_0.4.2       readxl_1.3.1      rmarkdown_2.6     modelr_0.1.8
## [37] magrittr_2.0.1    backports_1.2.1   scales_1.1.1      ellipsis_0.3.1
## [41] htmltools_0.5.1.1 rvest_0.3.6       assertthat_0.2.1  colorspace_2.0-0
## [45] stringi_1.5.3     munsell_0.5.0     broom_0.7.5       crayon_1.4.1
```