Readme for reproducing results in "One-inflation and unobserved heterogeneity in population size estimation"

Ryan Godwin

email: ryan.godwin@umanitoba.ca

Session information:

```
R version 3.2.1 (2015-06-18)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 7 x64 (build 7601) Service Pack 1

locale:
[1] LC_COLLATE=English_Canada.1252  LC_CTYPE=English_Canada.1252
[3] LC_MONETARY=English_Canada.1252 LC_NUMERIC=C
[5] LC_TIME=English_Canada.1252

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
[1] maxLik_1.2-4     miscTools_0.6-16

loaded via a namespace (and not attached):
[1] zoo_1.7-12      sandwich_2.3-3  grid_3.2.1      lattice_0.20-31
```

Note: Despite using the `set.seed()` function, large entries in the tables (e.g. > 1×10^5) are not reproducible across all machines. For example, out of three different computers, one produced different values for the large numbers in Tables 1 and 2. I have no explanation for this.

Functions

Many functions required for running various programs are in the file "functions.r", which must be in the working directory.

Reproducing Table 1

The file "table1.r" will reproduce the three top-left entries of Table 1. The other entries can be reproducing by altering the data generating process appropriately.

Reproducing Table 2

The file "table2.r" will reproduce the first row of Table 2. The other rows can be reproduced by altering the sample size "n".

Reproducing Table 3

The file "table3.r" will reproduce the first row of Table 3. The other rows can be reproduced by altering the sample size "n".

Reproducing Figure 1 (right)

The file "tableA8.r" will reproduce the first row of Table A8 in the Web Appendix. Subsequent rows can be reproduced by changing the data generating process, i.e. by assigning the appropriate values to the parameters "n", "a", and "li".

Table A8, reformatted in the file “data\sizeztnb.csv”, is then used to generate Figure 1 (right) in the main paper, by running the code “figure1right.r”. The file “sizeztnb.csv” must be in the “data” folder in the working directory. As a result of running the code, a graphic file “sizeztnb.png” is produced in the working directory.

Reproducing Figure 2 (right)

The file “tableA13.r” will reproduce the top-left entry of each of the three panels in Table A13 in the Web Appendix. The other entries can be reproduced by changing the sample size "n" and the value of "omega". Table A14 can be reproduced by changing the parameter value of “a” from 0.5 to 2 in the file “tableA13.r”, and repeating the above process.

Table A13, reformatted in the file “data\powerztnbl1a5.csv”, and table A14 reformatted in the file “data\powerztnbl1a2.csv”, are then used to generate Figure 2 (right) in the main paper, by running the code “figure2right.r”. Both “.csv” files must be in the “data” folder in the working directory. As a result of running the code, a graphic file “powerztnb.png” is produced in the working directory.

Reproducing Tables 4-8

Table 4 can be reproduced by running the file “table4-8.r”. The data is from Cruyff and van der Heijden (2008), and is already defined in the file “table4-8.r”. Tables 5-8 can be reproduced by altering the “y” variable for the appropriate data. Details, and the data, are in the Web Appendix, but the data for each table is reproduced here for convenience:

Table 5: `y <- c(rep(1,121),rep(2,13),rep(3,5),rep(4,2))`

Table 6:

```
y <- c(rep(1,2176), rep(2,1600), rep(3,1278), rep(4,976), rep(5,748), rep(6,570), rep(7,455),
rep(8,368), rep(9,281), rep(10,254), rep(11,188), rep(12,138), rep(13,99), rep(14,67),
rep(15,44), rep(16,34), rep(17,17), rep(18,3), rep(19,3), rep(20,2), rep(21,1))
```

Table 7: `y <- c(rep(1,516), rep(2,146), rep(3,52), rep(4,24), rep(5,14), rep(6,5))`

Table 8: `y <- c(rep(1,541), rep(2,169), rep(3,95), rep(4,37), rep(5,21), rep(6,23))`

Web Appendix

A text document, “dutch.txt”, is also included in the “data” folder. It is data from van der Heijden et al. (2003), and is necessary in order to run the R code provided in the Web Appendix.