University of Lisbon ISEG
2025/2026 Paulo M.D.C. Parente

Advanced Topics in Econometrics
(PhD in Mathematics Applied to Economics and Management)
2025/2026
Exercise Sheet 9 - The bootstrap
(version 28/9/2025)

1. The file `vietnam.xlsx` (available in fenix), contains data from the World Bank 1997 Vietnam Living Standards Survey[1].

   (a) The variable `lhhexp1` represents the log of the household total expenditure. Obtain a 95% confidence interval for the mean of `lhhexp1` using the following methods:

      i. Standard asymptotic theory.
      ii. Standard asymptotic theory using the bootstrap mean and the bootstrap standard errors.
      iii. The bootstrap percentile method.
      iv. The bootstrap t-statistics (use both methods taught in the lectures).

   (b) Suppose X_1,...,X_n are i.i.d. continuous random variables from distribution with cumulative distribution function $F_X(.)$ and probability density function $f_X(.)$. Let M_n be the sample median and m_0 be the population median. In this case $\sqrt{n}(M_n - m_0) \xrightarrow{d} N(0, 1/\left(4\{f_X(m_0)\}^2\right))$.

      i. Propose a consistent estimator for the asymptotic variance of $\sqrt{n}(M_n - m_0)$.
      ii. Obtain a 95% confidence interval for the median of `lhhexp1` using the following methods:
         A. Standard asymptotic theory using the consistent estimator of the asymptotic variance of $\sqrt{n}(M_n - m_0)$ that you proposed in question $(b)\, i$.
         B. Standard asymptotic theory using the bootstrap mean and the bootstrap standard errors.
         C. The bootstrap percentile method.

   (c) Consider the following regression model

$$\texttt{lnrlfood}_i = \beta_1 + \beta_2 \texttt{lhhexp1}_i + \beta_3 \texttt{age}_i + \beta_4 \texttt{age}_i^2 + u_i,$$
$$E(u_i | \texttt{lhhexp1}_i, \texttt{age}_i) = 0.$$

   Obtain a 95% confidence interval for β_2 using the following methods:

      i. Standard asymptotic theory (use robust standard errors).
      ii. Standard asymptotic theory using the bootstrap mean and the bootstrap standard errors.

[1] This data set is used by Cameron, A.C. and Trivedi, P.K. (2005). "Microeconometrics: Methods and Applications", Cambridge University Press. The data are available here: http://cameron.econ.ucdavis.edu/mmabook/mmadata.html.

iii. The bootstrap percentile method based on the bootstrap of the pairs.

iv. The bootstrap t-statistics based on the bootstrap of the pairs (use both methods taught in the lectures).

2. The data for this exercise are contained in the Excel file `mrozdat.xlsx`, available for downloading from fenix. The file contains labour supply data for a sample of American married women, as used by Mroz, T. (1987, *Econometrica*, "A The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions," Vol. 55, No. 4, 765-799). See the data description document in fenix for variable definitions.

   (a) Estimate a Probit model for lfp using $prin$, $kl6$, wa, we as explanatory variables (use robust standard errors).

   (b) Obtain a 95% confidence interval for the true coefficient of $kl6$ using the following methods:

      i. Standard asymptotic theory using the bootstrap mean and the bootstrap standard errors.

      ii. The bootstrap percentile method based on the bootstrap of the pairs.

      iii. The bootstrap t-statistics based on the bootstrap of the pairs (use both methods taught in the lectures).