

---

## Microeconometrics - Assignment 02

---

Submission deadline: 31.01.2021, 11:59 p.m.

- Assignments handed in too late will not be corrected.
- This assignment is part of your exam, therefore, cheating (illegal cooperation between groups) will result in a failing final grade.
- Your pdf document cannot contain more than 2000 words. After 2000 words, I will stop reading.

Every assignment consists of:

- (1) A pdf document (named: "MIC\_WS2021\_A02\_lastname1\_lastname2\_lastname3") containing your answers as well as explanations. The plots and tables to which you are referring, must be included in the document. Do not include any parts of your code in the pdf document. Your solution needs to be typed on the computer, not hand-written and scanned. You can use Word, LaTeX or any other text-producing software. You have to apply proper mathematical notation (for example  $x_1$ ,  $\beta^2$ ,  $\sum_{i=1}^n Y_i$ , not x1, B^2, sum(Y\_i)). All solutions need to be fully transparent.
- (2) A well-documented R script (named: "MIC\_WS2021\_A02\_lastname1\_lastname2\_lastname3"), which replicates your results.
- (3) Every assignment needs to include a title page with your full Names, student ID's, your course of study, university, and your email address.

In sum, each group must submit only 2 documents via email to t.hainbach(at)tu-berlin.de (subject: "MIC\_WS2021\_A02\_lastname1\_lastname2\_lastname3"): a pdf containing your answers, plots and explanations and a R document containing your code and comments explaining your code.

Problem 1: Logit Model (Polynomials and Bootstrap)

- a) Download the dataset "titanic\_train\_data.csv", which contains real data of 891 Titanic travelers and load the data into R. Estimate the following logit model:  
$$P(\text{Survived} = 1|\mathbf{x}) = \Lambda(\beta_0 + \beta_1 \text{Age} + \beta_2 \text{Age}^2).$$
- b) Plot the estimated logistic regression function and interpret it. Does the estimated logistic regression function give some information about the rescuing strategy? Explain.
- c) Now, estimate the following logit model:  $P(\text{Survived} = 1|\mathbf{x}) = \Lambda(\beta_0 + \beta_1 \text{Fare})$ . Plot the estimated logistic regression function and interpret it.
- d) Apply the bootstrap idea to create and plot nonparametric pointwise 95% confidence intervals of the prediction for the logistic regression model in c) (use 100 bootstrap replications).
- e) Interpret the nonparametric 95% confidence interval of the prediction (created in d) ) for a fare value of 100 in detail.

Problem 2: Probit Model (Interpretation)

- a) Download the dataset “titanic.dta”, which contains real data of 2201 Titanic travelers and load the data into R. The data includes whether the travelers survived, in which class they travelled, their gender and their age (coded as adult or kid). Estimate a probit model that links the survival probability to these characteristics. Then interpret the effect of the variable sex properly. Does the estimated effect conform to your expectation?
- b) Program the marginal effect of being a man on the survival probability yourself (not using any package). Draw this marginal probability effect of gender as a function of  $x_i'\beta$ .
  - (i) Is the resulting function linear? Explain.
  - (ii) For which  $x_i'\beta$  is the resulting marginal effect on the probability to survive largest? Explain.
  - (iii) Does estimating the marginal probability effect make sense in this particular case? Shortly explain.
- c) Find out what the effect of being in the 1st class is on the survival probability?
- d) Test the hypothesis that being a man had different probability effects throughout the cabin classes. Your solution needs to be fully transparent.