Homework 3

Computation Camp

Due August 16, 2023

Question 1

This question will take you through estimating a normal linear model using maximum likelihood. You will also conduct inference using three methods. The lwage.csv dataset (Dropbox link (https://www.dropbox.com/scl/fi/d9s4zyozbcr1kwosw8kr1/lwage.csv?rlkey=11d9jocmqlodcpvlp4efllbf8&dl=0)) contains a large sample of simulated worker's log hourly wage (lwage), college attainment (educ), and years of experience (exp). Suppose the data-generating process is
$\log(wage_i) = \beta_0 + \beta_1 educ_i + \beta_2 exp_i + \beta_3 exp_i^2 + \varepsilon_i$, where ε_i is an error term distributed $N\left(0, \sigma^2\right)$ and is assumed exogenous.

A. Estimate the model using OLS in your preferred statistical software. You could also try out the GLM (https://juliastats.org/GLM.jl/v0.11/) package in Julia. Report standard errors and t-statistics.

B. Derive the log-likelihood of the model for any arbitrary guess of parameters $\left(b_0, b_1, b_2, b_3, s^2\right)$, where s^2 is a guess of the variance of the error term. Hint: recall that the PDF for $N\left(0, \sigma^2\right)$ is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}$$

C. Estimate the model via log likelihood and report your results. I recommend following the example set at https://julianlsolvers.github.io/Optim.jl/stable/\#examples/generated/maxlikenlm/ (https://julianlsolvers.github.io/Optim.jl/stable/\#examples/generated/maxlikenlm/).

D. We will now obtain standard errors using a Bootstrap method. For a single bootstrap run, you will draw a random sample with replacement from the original data set that is equal in size. Then, obtain parameter estimates according to the procedure in part C). Do this 100 times to obtain 100 different estimates of the model parameters before reporting the sample standard errors of the parameter estimates. Each time you do this, use a different random seed so as to induce randomness in which observations are kept and which are left out. Run this program on a single core, and then parallelize it to run on Linstat with 20 cores. Comment on how much time parallelization saves here.

Question 2

Consider the following matching problem. Suppose you have n people, numbered 1 to n along with n slips of paper, also numbered 1 to n. The slips of paper are put in a hat, the hat is passed around, and everyone takes a slip (ignore things like slips being on top of the pile and assume that the drawing here is perfectly random). What is the distribution of the number of people who draw the slip with their number?

A. Conduct this exercise for $n = 10$ via 10,000 Monte-Carlo simulations and plot the resultant distribution.

B. Conduct this exercise for $n = 20$ and plot the resultant distribution. Comment on how it compares to the distribution you found in part A).

Question 3

One rule-of-thumb for retirement savings is that you want to have 10x your earnings in savings at age 67 if you'd like to sustain your current lifestyle in retirement. This problem studies the extent to which uncertainty makes it difficult to achieve this.

Suppose that an agent has earnings $E = 100$ and savings $S = 100$ at age 30. Assume all savings are invested in stocks. Yearly percentage returns on stocks are drawn from a normal distribution with mean 6% and standard deviation 6%. Yearly percentage raises are drawn uniformly from $[0, 6]$. The only decision the agent makes is a single proportion P of their earnings to put in savings every year between the ages 30 and 67.

A. Suppose there is no uncertainty, so that returns on stocks and raises are equal to their distributional mean every year. What level of saving allows the agent to have $10x$ their earnings in savings by age 67 ?

B. Now run 10,000 simulations of the outcomes of an agent following the savings rule found in part A) with uncertainty in stock returns and wages. How often does the agent fall short of the $10x$ goal?

C. What level of savings gives at least a 90% chance of meeting the $10x$ goal with uncertainty?