

Computational Statistics

PROJECT 2

2023-2024

A researcher wants to work with a distribution with density function given by

$$f(x; \theta) = \frac{\theta^3}{\theta + 1} x \left(1 + \frac{x}{2}\right) \exp(-\theta x), \quad x, \theta > 0$$

You can find in file project2.txt 100 observations that he believes that come from this distribution

1. Estimate parameter θ using Maximum Likelihood method. You shall describe all steps in full detail and write your own code for this.
2. Estimate the standard error using parametric bootstrap. This implies that you need to find a way to simulate from this distribution.
3. Run a Monte Carlo test to see whether the data come from this distribution.
4. An alternative model for the data could be one with pdf

$$f(x) = \pi f(x; \theta_1) + (1 - \pi) f(x; \theta_2), \quad x, \lambda, \theta_1, \theta_2 > 0$$

Estimate parameters $\pi, \theta_{1,2}$ using ML method and in particular an EM algorithm. Give all the steps of the algorithm in full detail. Then implement this algorithm in R. Derive standard errors for the estimates.

5. Replace your E-step with a simulated one to derive an MCEM algorithm. Implement this and compare your results with that of the previous one.
6. Finally, which one of the two distributions you find as more appropriate for this data set? Explain.

Notes: I expect to see the full derivations for the algorithms together with R code to implement them. I also expect to see the numerical results.

You will upload 2 files

- a pdf file with all your derivations. to save time you can scan hand written manuscripts and create a pdf with them
- a txt file with the R code