

Computer-Intensive Statistics

Assessment Exercises

Students should talk to their supervisors to find out whether or not their department requires some of this work as part of any formal accreditation process.

Ex 1 There is another type of bootstrap-based confidence interval not discussed in the lectures, called *ABC*. Use the references to find out about it, and write up a one-page description of the method and a comparison of its strengths and weaknesses (compared to the methods included in the lecture notes).

Ex 2 Thursday's practical analysed only part of the Australian AIDS survival data. Extend the analysis to the complete dataset: you will need to take 'state' into account (for a classical analysis see Venables & Ripley, 2002, 379ff), and should expect to use many minutes of computation.

You should look at several (at least two) runs to convince yourself that the chains have converged. Do so in parallel using all the cores on your computer. (You might like to look at package `parallel`, which is part of R and has a PDF manual.)

Write up the analysis you did (and for those of you who have not written up a Bayesian analysis before, you need to document the choice of priors).

Ex 3 Consider a probit model for LD50 for the following data

dose	animals	deaths
50	5	1
100	10	4
200	5	2
300	10	6
2000	5	4

with a suitable vague prior. Find a suitable HPD region for LD50. How does it compare to a logit model and to a classical GLM analysis?

Ex 4 (For those with computer-intensive projects.)

Think about how you could organize the algorithms that are used in your field to make use of multiple CPUs, either a shared-memory computer with 24 conventional CPU cores or a GPU-based 'personal supercomputer'. Prepare a page or two of write-up for a non-specialist audience on which approaches are most promising in such an environment (and why).