MSDS 7333: Quantifying the World

Case Study 6: Simulation Study of a Branching Process

Instructions

In this chapter, we are studying the behavior of programs/jobs running on a CPU (or multiple CPUs in parallel). We were interested in the distribution of the completion time for the entire system. In particular, we wanted to know which combinations of job creation and job completion rates leads to processes that definitely complete in finite time, and those that might never complete.

For this case study, choose one (and only one) of the following problems from the back of the chapter: 6, 9, 10, 11 (page 307 in NTL text). Write a report on your analysis, including introduction, background, methods, results, and conclusions/discussions. Your write-up need only to focus on answering the question of your choice. In other words, you do not have to include code that is already in the chapter or in other class materials.

1. Choose your group
2. You can use any platform/package you would like
3. You can alternatively come up with a question of interest regarding making a text-based classifier of Wikipedia articles à la the TMNT example. I’ve included the code “caseStudy.r” for getting a document term matrix from a Wikipedia article. f you do this option, make sure you describe the scientific problem you are trying to solve and the processing you do to get a document term matrix, what methods you are using for supervised (e.g. logistic lasso) or unsupervised (e.g. clustering) learning, and the implementation of the algorithms you are using (e.g. glmnet).
4. Write a report on your analysis, including an abstract, introduction/background, methods, results, conclusions/discussions, and references. I’m not planning on including a rubric for this case study due to its unstructured nature.