Introduction

Fragility describes how a system is likely to respond to small perturbations to its parameters. I investigated fragility in biased random walks. I considered stochastic models such as birth and death chains that have the potential to describe a variety of scenarios. I made heavy use of computer simulation to guide my intuitions, and applied analytical methods. In this writeup, I reflect on my findings and the process of math research.

Project Phases

Mathematical Context: At first, my conception of mathematical fragility was vague. This was due to my inexperience with probability theory. I needed a better understanding of the context, especially stochastic models. I worked through exercises and referred to a number of textbooks, introducing myself to the key concepts of recurrence and hitting-time. This took longer than originally expected, yet proved critically important as I would have been left without appropriate language or definitions necessary to consider birth and death chains.

Initial Simulations: I next turned my attention towards some specific probability distributions. In particular, I needed to develop my intuitions about birth and death chains, and the property of recurrence—which describes the chain's potential to return to a certain point. I used Python with the `NumPy` and `Matplotlib` libraries to simulate birth and death chains defined by various probability distributions, and analyze and plot return times to 0. I had previous experience with similar forms of data analysis, and found that this came naturally to me. However, I found it challenging to set appropriate parameters for each simulation—too many data points caused simulations to run unnecessarily long, but too few (or poor sampling technique) produced inconclusive results. I learned techniques that can be used to investigate hypotheses about return times.

In particular, my first simulations led me to define the *power-distributed chain with parameter* β, whose behavior is governed by the probability functions $p(n)$ and $q(n)$. When walking on the natural numbers, $p(n)$ represents the probability of the chain stepping in the positive direction, and $q(n)$ the negative. These functions are defined by

$$p(n) = \begin{cases} 1 & \text{if } n = 0 \\ \frac{1}{2+n^{-\beta}} & \text{if } n > 0 \end{cases}, \qquad q(n) = 1 - p(n).$$

I hypothesized that the power-distributed chain is *positive recurrent* (i.e. returns to 0 infinitely often and in finite expected time) if $\beta < 1$, and is *null recurrent* (i.e. returns to 0 infinitely often but with infinite expected return time) if $\beta \geq 1$. Simulations allowed me to compare return times values of β, and the results supported my hypothesis.

Analytic Proof: In mathematics, however, numerical results from simulation do not constitute sufficient evidence. As such, I turned to analytic methods to prove my conjecture about the power-distributed chain. While I drew upon established results to prove that some infinite sums converge, I needed to be creative with how I applied approximations. I often became frustrated feeling like I was working in circles. Completing this phase required my persistence as a mathematician. Eventually, I was able to make some key observations that allowed me to succeed in proving my conjecture. Overall, the time spent here felt too long for the results established, but it challenged both my mathematical ability and my perserverance.

The Right Question: Through my analytic investigation of the power-distributed birth and death chain, I had to consider if recurrence properties offered the best information, or if other properties about birth and death chains would be more enlightening. This led me to consider the more general *mean hitting time*—the expected number of steps for a chain in a particular state to reach a different state. Mean hitting times are more general than recurrence times and better represent a system's "efficiency" in reaching a particular state, while recurrence times are a more direct representation of whether a system frequently "gets lost." I am still unsure that mean hitting times should be the object of study.
Exploration of Hitting Times: Let $h_s(n)$ be the expected hitting time to 0 from position n with step size s. I was advised that calculating hitting times was difficult, and so investigated using alternative methods. I returned to simulations, updating the necessary parameters to instead vary the step-size of the power-distrbuted chain. This helped enlighten quantities of particular interest, such as $\partial h/\partial s$, the change in hitting time with respect to step-size. I believe that these quantities hold important information about the fragility of a system.

Further Research

The work completed thus far has led me to investigate hitting times. I think that hitting times are still valuable in uncovering more about mathematical fragility, especially as it relates to step-size. A number of quantities—such as specific limits or partial derivatives—may illuminate facts about how a birth and death chain responds to perturbations in its parameters. Further work to determine the specific shapes and classes of hitting time functions $h_s(n)$ for various probability distributions will help understanding the restrictions of these hitting time functions. This could, eventually, allow for complicated analytical methods to deduce properties of these functions.

However, it could be that a different quantity offers more suitable information. For example, it might be more enlightening to consider the probability that a birth and death chain reaches a particular point. This may allow a broader class of probability distributions to be considered and as such could direct further research. I expect direct computation of these probabilities to be difficult, and so further simulation may be the best way to uncover preliminary information about these quantities.

Final Thoughts

Above all, this project helped me to better understand my strengths and limits as a mathematician. It gave me the invaluable opportunity to experience research early in my career. I found it difficult to predict how long answering a mathematical research question would take, and I had to persist even when I felt behind schedule or lacking in ideas. I found research very different from my academic work, as I had to creatively look for new directions. While the answer may not be known when working on problem sets, I usually find I know the angle to best tackle the problem. In contrast, this research was challenging in that I both needed to ask the right questions to best guide my ideas, as well as deduce the best techniques and approaches to use when answering those questions. This is not easy, and gaining this experience so early in my career will surely aid me in the future. I would like to thank Northwestern and the Office of Undergraduate Research for giving me this opportunity, as well as Dr. Jason Siefken for his incredible mentorship and support throughout this project, even when separated by great distances.