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 analytic 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 in part 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 laid the mathematical foundation of my project.

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 the Python programming language with the `NumPy` library to simulate birth and death chains defined by varying transition probabilities. Later, I used the `Matplotlib` library to plot return times to zero. 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 techniques) produced inconclusive results.

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 zero infinitely often and within finite expected return time) if $\beta < 1$, and is *null recurrent* (i.e. returns to 0 infinitely often but with infinite expected return time) if $\beta > 1$. Simulations allowed me to compare return times and values of β. The simulation 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 a particular infinite sum converges, 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 return times and better represent a system's "efficiency" in reaching a particular state, while return times are a more direct representation of whether a system frequently "gets lost." I am still unsure that mean hitting times are the best 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 calculating hitting times was difficult and so used numerical methods. I updated my simulation to instead vary the step size of the power-distrbuted chain. This showed me how quantities of interest, such as $\partial h/\partial s$, vary with respect to step size. I believe that these quantities can be used to classify the fragility of a random walk.

Further Research

The results I have gathered so far show that the hitting time of the power-distributed chain decreases as step size increases. Further, simulations also reveal subsets of step sizes where the mean hitting time instantaneously drops. I suspect that these sequences are closely related to the division algorithm, and an analysis of these sequences may give information about the shape of hitting time functions.

Additionally, many choices of transition probabilities lead to birth and death chains with infinite mean hitting times. Thus, mean hitting times cannot be used to compare these chains (for example, mean hitting times cannot compare the power-distributed chain when $\beta > 1$). Instead, I would like to investigate another quantity: the probability that the chain reaches a particular point. This quantity is always well defined, and would allow comparison of chains with infinite mean hitting times. I would then broaden my study of chains. The power-distributed chain is a single family, and so I wish to continue my investigation by considering chains with other families of transition probabilities.

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 was lacking in ideas. I found research very different from my academic work. While working on problem sets, I typically do not know the answers but usually find I know potential approaches. In contrast, this research was challenging in that I both needed to ask the right questions, as well as find appropriate techniques and methods that answer 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.