I am strongly interested in applying for the Multi-disciplinary Doctoral Training Program at the University of Toronto. My research interests include Bayesian Statistics and the interdisciplinary research being conducted between Bayes and other areas. I believe Bayesian Statistics should be further studied given its ability to explore limited data as well as its usefulness in data-intensive areas such as Education, Natural Science and Finance, as it allowed scholars to incorporate prior domain knowledge into statistical models. My ultimate goal is to raise the awareness of Applied Bayes in interdisciplinary research and become a leading researcher in Bayesian Statistics.

My first venture into examining the implications of mathematical assumptions of astrophysics theories affirmed my passion in academic research. Specifically, I investigated and compared two contemporary Supernova theories and their influence on present research directions in the form of literature review, which earned me 100% for the course assignment and an opportunity to publish the article in Arbor Journal, an undergraduate academic journal at the University of Toronto. It has given me the confidence to compare and evaluate mathematical models with available data. My current research explores the use of Bayesian model in Education, largely because of my passion in this area. By replicating a research project that was initially conducted under Frequentist’s model, I argue that Bayesian model may provide unique insights that are equally informative for decision-making in Education. In brief, my research concerns the real-life consequences of mathematical models. These research experiences have given me considerable practical experience in programming and a broad knowledge in interdisciplinary research across many areas, so that I have become better acquainted with the essentials of academic research on a basic level.

To further develop my research skills, I have taken and succeeded a variety of Mathematics/Statistics-related courses, from multivariate calculus to theoretical statistics and applied Bayesian statistics. To learn more about Real Analysis, I self-studied the book Understanding Analysis, solved problems and shared my solutions on my personal website built with R. As a self-motivated learner, I also continue to enhance my skills in R, Python and SQL by leading data science projects ranging from gender equality in the workplace with hierarchy model to modeling infections with differential equations and Bayesian models. In my recent project about the livability of Toronto, I analyzed the rental prices in Toronto/Vancouver using 3 different statistical approaches, and applied Bayesian regression in accordance with a research article about the rental prices in Tokyo. The project was awarded 2nd place for Best Insight in a case competition.

Bayesian Statistics has much potential to be applied in many fields, and this is why I have chosen it as my primary research interest. Based on my undergraduate experience, we emphasized Frequentist's statistics much. In contrast, Bayesian Statistics seemed to only serve as a “side dish”. However, those sources I encountered during my research, which ranged from creative Bayesian textbook Statistical Rethinking by Richard McElreath (2016), to the influential essays on Bayesian vs Frequentist by Bradley Efron (1986), challenged my basic understanding of Statistics in ways not anticipated, especially the ways Bayesian models combine domain knowledge and statistical modeling. Furthermore, as a teaching assistant for an introductory statistics course for undergraduates at UofT, I was very surprised by the application of statistics in many academic articles in scientific fields during my preparation for the tutorials. Specifically, the methodology used in these articles demonstrated that the dominant one-for-all application of Frequentist's theories was not so objective as many scientists expected(Efron, 1985). For instance, many scholars unconsciously utilized significance testing without sufficient justifications and interpreted the method incorrectly, which is known as P-hacking (Head et al., 2015). I believe this implies that beliefs in the research outcomes and knowledge in the data hermeneutics will inevitably inhere in analysis but may not be encompassed by Frequentist’s models. Fortunately, Bayesian statistics provides useful tools including priors and Bayesian update to merge disciplinary knowledge with models formally. Thus, I have strong interest to improve the application of statistical methods in interdisciplinary research and develop context-specific methods from a Bayesian perspective. In particular, I expect that the course of my graduate school will address the issues of applying Bayesian statistics in academic research. With complicated parallels between Frequentist and Bayes in practice, I also wish to offer as a statistician a holistic analysis of the interplay of statistical approaches in decision making and long-term results in different areas.

I am also very excited at the prospect of Bayesian statistics in machine learning. My internship revealed another possible application of Bayesian inference as I worked on modeling the Education data for secondary school enrollment. By employing machine learning models with emphasis on predictive accuracy, I could extrapolate the data effectively and make accurate predictions. However, this type of model does not imply a good casual model in real world, which is also hard to interpret. In Statistics Rethinking (McElreath, 2016), the author shows that Bayesian Statistics is aligned with common logic that allows us to provide sensible interpretation. Hence I believe it is worth investigating the integration of Bayesian statistics and machine learning models, such as Bayesian network in education assessment, which can be powerful in both prediction and reasoning.

Having spent last four years at the University of Toronto (UofT) as an undergraduate student, I am really grateful for the great opportunities provided by the university and would like to contribute to the pivotal research in Bayesian statistics with machine learningAt the same time, I am particularly indebted to Professor Rohan Alexander for supervising my reading course this year and guiding me to critically evaluate Bayesian statistics. I really appreciate the Toronto Data Workshop hosted by Professor Alexander for statistical communication as well. Moreover, I also want to thank Professor Keith Knight for inspiring my research interests in his course and kindly unfolded the mathematical details of Bayesian models for me. Hence I anticipate the greatest pleasure to conduct research as a PhD candidate at UofT.

In addition, UofT has a strong department of statistical science and provides an ideal climate for me to develop my cross-disciplinary interests. In particular, UofT provides the Multi-disciplinary Doctoral Training Program, a unique option that distinguishes UofT’s Statistics program from that of other institutions. The cross-disciplinary nature of UofT’s graduate program would foster, I hope, fruitful discussions with other departments, notably the department of computer science and OISE, thus broadening and enriching my research as well as my general understanding in Bayesian inference and natural science/human science. Together, I am looking forward to a challenging and enjoyable experience at UofT.

Reference

1. Efron, B. (1986). Why Isn't Everyone a Bayesian? The American Statistician, 40(1), 1-5. doi:10.2307/2683105
2. Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. PLoS biology, 13(3), e1002106. <https://doi.org/10.1371/journal.pbio.1002106>
3. McElreath, R. (2016). Statistical Rethinking: A Bayesian Course with Examples in R and Stan (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315372495