**Program Applied: Bioinformatics/Biostatistics**

Frankly, when first exposed to biostatistics at seminars (in the second semester of my junior year), I found myself quite ignorant of this academic discipline. To be sure, as a student from the Department of Mathematics, xxxx University, I had sound foundation in statistics; but it required an equally sound knowledge base in biology in order to perform well in a host of seminars on bioinformatics. To develop my interdisciplinary knowledge, I tried desperately to be educated in biology—through self-study, consulting a large amount of relevant technical literature, attending lectures, and discussing with classmates.

Indeed, that was a period of overwhelming fatigue and excitement. Sometimes, going to sleep at 3 o’clock in the morning and getting up 4 hours later was habitual. Nevertheless, the feeling of learning something new and making important progress was itself a source of constant pride. Looking back, this experience has proved decisive in determining my academic interest in biostatistics and in the related bioinformatics. It has further shaped my research interest in the Ph.D. program at your esteemed university.

The emergence and development of an academic interest is always a slow process and my understanding of biostatistics has been deepened through a series of academic activities. I started by reading three important books. The Analysis of Longitudinal Data (Clarendon Press, Oxford), with its refreshing examples of biostatistics, practical probability models and approaches of vertical data processing, gave me basic understanding of the application of statistical methods in biology. My knowledge of bioinformatics derives chiefly from two other books—Biological Sequence Analysis – Probabilistic Models of Proteins and Nucleic Acids (Cambridge University Press) and Bioinformatics – A Practical Guide to the Analysis of Genes and Proteins (Interscience). The input from those books enabled me to propose constructive suggestions in the latter-day seminars and summer workshops, and my ability to do independent study and to learn new knowledge is fully demonstrated.

Prof. xxxx, whom I met early in the junior year, has exerted a defining influence over my academic interest. An expert on probability, he differs from other probability teachers by his emphasis on practical application rather than on pure analysis. Under his instructions, I came to realize the special power of mathematics and statistics in solving practical problems, a power totally beyond other intellectual endeavors. Hence, I have paid special attention to the practical application of statistical theories.

With the two lectures delivered by Prof. xxxx from xxxx University—Semparametric Regression for xxxxxxxxxx, my attention was called to the application of relatively cutting-edge methodologies to problems in biostatistics. In addition, in the second half of my junior year, I attended two semester-long seminars on bioinformatics, one held by the Department of Mathematics on xxxx analysis (with special emphasis on the application of probability models) and one by Prof. Xxxxx  from the Department of Biology on the application of xxxx analysis. Both of the seminars exposed me to the areas of general research and the standard scholarly procedures in bioinformatics.

By far, the greatest impetus that has driven me to bioinformatics is the summer workshop jointly held by our department and the Department of Automation and hosted by Prof. Xxx  from Harvard University. Most of the papers I read featured finding gene modules, DNA sequences, and protein structures. The issues such as how to study protein structures through optimal and statistic methods were also addressed.

Having acquired adequate knowledge in biostatics and bioinformatics, I am on my way to developing my research capacity. In collaboration with Prof. xxxx from the Department of Automation, I am trying to work out methodological predicament faced by Prof. xxxx in her study of the disease-causing genes. In the research she has carried out, conventional empirical algorithms such as perturbation analysis were employed but, for lack of mathematical foundations, optimal solutions have been difficult to find. I have proposed to introduce standard Bayesian theories into the research to reconstruct probability models so that more satisfying statistical results can be achieved.

My research experiences also come from preparing my B.S. thesis on the construction of three dimensional structure of the protein. Conventional research aims at finding proteins with similar structures that can match. While mathematical models take global optimal into full account, they tend to overlook the local optimal. The models might be optimal mathematically but not necessarily so biologically, deviating from the ultimate objective of bioinformatics. My task is to discover proper algorithms that can bridge the variations between the two and to achieve some pioneering results.

Though venturing into bioinformatics and biostatistics only in the middle of my undergraduate program, I have proved myself a fast and smart learner, with tentative research experiences. This potential convinces me that, with all my preparations, I can become a strong candidate for a Ph.D. program in the Department of Statistics, University of xxx. My careful investigation reveals that Prof. xxxx can teach me statistical theory and methods, survival and event history analysis, and mixture models; Prof. xxxx can instruct me on Bayesian inference, missing data, and survey sampling. Prof. xxxx specializes in sequential design & analysis and in probability application. I would like to identify any of them as my potential advisor and work fruitfully under his/her guidance. I am ready to study either biostatistics or bioinformatics, depending on the conveniences of the faculty.

In undertaking your program, my background in mathematics and statistics would be my greatest asset. My academic transcript indicates that I have achieved straight A’s in almost all the mathematics courses and also highest scores of the class in statistics courses. Familiar with classical books such as Probability Theory by Olav Kallenberg and Probability Essentials by J. Jacod and P. Protter, I am also familiar with computer languages such as C and C++ and the application of statistical software.

“Something attempted, something done,” Henry Longfellow asserts. In attempting your program, I have the following study plan. As my knowledge and skills in biostatistics and bioinformatics are basic, I would spend the first two years deepening this knowledge, with special emphasis on learning biology and more advanced statistical methods. Computer applications will also be a major concern. I will participate in research projects whereby to apply theories to practice. Based on intensive coursework and extensive research, I hope to focus on a specific topic, preferably one that falls into my prospective advisor’s research interests to complete my dissertation. I believe that by planning my future study in this way I may achieve expected results through your rewarding program.