## Why a graduate program?

1. I started working in the industry, and continued for some years

* I faced some challenges in the industry about methods
  + Non-linear models, DOE, Bayesian inference, machine learning techniques, etc.

I realized that I could find the answers to those questions by joining a graduate program in statistics.

* Because industry and academia are parallel and incorporated in a way that the answers of industry questions are in the academia

2. The ability to do research independently requires more in-depth knowledge in statistics

* I needed to learn more theory to be able to digest what’s going on behind the methods I was using so that I could apply the methods more appropriately.

Research:

I am interested in biostatistics and genetics problems and learning methods for integrated analysis of -omics data in general like genomics, proteomics, transcriptomics, phenomics, microbiome data.

I have been recently working on a model for identifying gene expressed data using multiple testing correction. Our proposed model is more powerful than the common t-test statistics. It is a good competitive for other available methods for signal recovery.

It works the best when the signal is sparse under dependent data.

The context of my research is genetics data, but the method of my research is all encompassing and can be applied to any type of data.

## Journey plan:

* I am currently a 4th year PhD student
* Plan to extend the method of my research in other areas
* Use network and Bayesian approaches to apply the idea to high dimensional data

That’s why I am a right fit to this position

Courses:

Theory courses made the backbone of statistical inference

Decided to take theory and methods together to understand the theories behind the methods. I can then modify the methods and apply them on my own research.

* Probability theories and advanced probability theories
* Bayesian theory
  + Gibbs sampling
  + Metropolis-Hasting sampling
  + MCMC
* Linear and non-linear models
  + Model assessment
  + Model selection
  + Model diagnostics
* Reliability: important for industry and specifically survival analysis and clinical trial designs
* DOE: elements of DOE, design experiment
* Data visualization courses: develop packages

Summer internship:

An intern at Merck, in the methodology design division

Major task: package development and assessment using simulation

* Simulation of time-to-event endpoints for proportional and non-proportional hazard assumptions. Which is a hot topic these days
* Seen NPH in oncology studies frequently due to delayed effects, crossing survivals, subgroup effects (log-rank test; wlr test;)
* Creating a package which can deal with NPH using normal approximation results.
* Creating a package to deal with NPH and group sequential designs for clinical trials
* Validating packages using unit tests

Other tasks:

* Working on system development life cycle for SAS and R with package development teams
* Assessing R package vs. SAS macro testing strategy

ClinChoice:

* Developing unit test for packages related to group sequential designs
* Simulation studies for oncology clinical trials

Experience:

* Enhanced my communication skills: other interns, other people in different divisions, joining socializing programs held by the company to get to know other interns and colleagues
* Working with a professional statistical team gave me the chance to enrich my teamwork skills. How I can leave my impact on the project when working in a team and how I can help other teammates and get help from them when facing problems.
* I learned if I am spending more than two hours on a piece of code, and I have not made any progress, I should definitely talk to my manager to keep my work productive and don’t waste time and energy.
* Got to work with some real-world problems and using my knowledge in statistics to plan for addressing the problems from a different perspective. What we get in reality is much messier than what we learn from books
* Make a great connection between what I had learned at school and what I can do when designing a clinical trial
* I cherish the experience I gained by doing an internship. I have tried to implement them into my life, my research and my work.

## End questions:

1. do you track your interns? Do you hire them?
2. What are potential challenges of an intern in this position? (want to plan for that)
3. Courses that are more useful for this type of position?
4. Is it more like working on some methodology and theoretic problems or more dealing with data? (is this research oriented?)
5. Timeline for the next steps?
6. What are the typical projects in the seed regulatory statistics team?