Advice For Graduate Students in Statistics

This page summarizes a discussion at the **Graduate Student Lunch Time Seminar** (1/29/2007). My intention was to put on the table the best practical advice I could muster, spiced up with some general musings about mentoring, writing, and publishing.

First Some Perspective

Every professor has a view about the best way to make use of the years one invests in earning a Ph.D. These views will vary over time, and they will vary from professor to professor. These views can be sources of inspiration and reassurance.

Nevertheless, students listening to such advice do well to keep a firm hold on their own values.

It's common sense that people who offer advice almost inevitably sell the lessons from their own story, and this may or may not be a story that works for you.

Self-Knowledge

Among scientists and other analytical folks it is uncommon to talk about self-knowledge. Still, what MBA candidates typically possess and what Ph.D. candidates often lack is a clear vision of where they expect to be in five years, ten years, or more. This may not be self-knowledge in the refined sense of philosophers or even psychologists, but it's still powerful knowledge.

If a student wants to obtain the most useful advice from a professor, it is an enormous advantage if the student can share with the professor his or her true objective. This takes courage, but it will greatly improve the focus of the advice.

Process versus Product

One spends four, five, or sometimes (sadly) six years in a Ph.D. program. This is a substantial fraction of an adult life, and one surely hopes that the time can be genuinely enjoyable. For many people it is one of the happiest, freest, most creative times of their life.

Specifically, there is no conflict at all between having a lot of fun and obtaining a great Ph.D. In fact, if the process is not fun, then one may want to consider Plan B.

The life of a professor is a lot like the life of a person who does twenty or thirty or more Ph.Ds --- while teaching full time. If the first Ph.D.is not fun, it is unlikely that the "subsequent Ph.Ds" will be fun. It can happen, but it is a long-shot.

Jumping Through Hoops

Any graduate program has a few requirements which may seem stupid, or which may indeed be stupid.

Thus, there are a few "hoops" that one must jump through in order to join the guild. This hoop jumping is not much fun, except perhaps in retrospect. Still, for people who have the basic talent and who can put in regular hours, these obstacles are never really as big as they sometimes seem before the rite of passage is complete.

In fact, the barriers have been coming down for years. At Stanford in the early '70s the quals were closed-book, a three-hour morning session and a three-hour afternoon session for three days --- one day for each core subject. Half of the students did not pass, and it was up or out --- with one re-try. I began my preparation in April and worked seven days a week until I took the quals in early September. I never worked as hard, before or since. We also had exams in French and German, though these were taken less seriously in the '70s than they had been taken even just a few years earlier.

In retrospect, it was intense but it was also a lot of fun.

Still, that is ancient history. For many years universities have been looking hard for ways to make the intermediate tasks of a Ph.D. program more meaningful. At Wharton, I am particularly impressed by the value students get from the "first year paper" and "second year paper" that are done in Finance (and some other departments).

When Professors Discuss "Advising Graduate Students"

When professors discuss among themselves how to best advise students, their thoughts jump almost immediately (and almost exclusively) to coaching about the thesis. Opinions rage --- even the ones that are not said out loud.

The **conventional assumption** is that the student wants to become a professor, and, even without checking with the student, the advice is almost always tempered toward that objective.

This is natural. Historically, most Ph.Ds have pursued careers in university teaching.

It is also psychological. People who have had successful academic careers typically have given very little thought to careers in government, industry, or finance.

Finally, it is at least a little self-serving. Placing a student in a nice academic position has great value to a professor --- it's like an annuity that may provide academic royalty payments for years and years.

Thesis Writing as Preparation for a Career in Academia

There is a consensus that has evolved over the last ten or twenty years about the most efficient way to write a Ph.D. thesis in statistics. The view is that **one should NOT write a thesis**, or at least not write a thesis in the form that was expected a few years back --- or the form that is expected today in less technical fields.

The strategy is instead to focus on writing publishable papers, and these should be submitted as soon as possible.

The hope is that during the third and fourth years of graduate school the student can create several works that can be published. The final thesis then just requires "stapling" these pieces of work together, with or without a small amount of connecting material.

The job candidates that we interview for our open positions have almost all written theses with this design. It is almost a necessity. Assistant professors now face so much pressure to publish that they need to show up on the first day with at least a couple of papers that are well along the path to publication.

One might hope that there might be greater room for choosing one's own way, but in my view there isn't. Still, this shift is really one of style. You can choose another style, but, before you do, please look at the web pages of assistant professors at good places. You will not find many exceptions to the "new" rule.

OK, How To Start ...

Here is the good news. You have already started. You are here. There is a whole faculty of people to (1) help you take the first steps toward identifying a problem, (2) coach you through the creation of a "first result", (3) suggest how the "first result" can be expanded and built upon, (4) suggest when there is enough to "start writing a paper" (5) work closely with you to put the paper into a form that meets publication standards, and (6) vet the new paper through the publication process. For most of their lives, many professors go through this process two or three times a year. By definition this is "educational" and for many people it is also a lot of fun.

What you DON'T need to worry about. You don't need to worry about "finding a problem." That sort of thing went out with professors wearing ties. Right now, January 29, 2007, **it is the responsibility of the professor to identify a suitable problem area** and to work with you on the concrete development of that area. Naturally, if you really have some problem that you are hungry to work on, you can find a professor to work with you on it. Still, this is a somewhat riskier strategy than hopping on a moving boat.

What you DO need to bring to the party. You need to put in an honest days work, say three genuine research hours (pencil in hand or fingers on the keys) and two genuine reading hours. I have never in thirty years known or heard of a student who did this five days a week for two years who did not get a thesis. Working much harder is fine if that is what feels right to you, but working less is risky. Perhaps you can scrape by with less, but the idea does not exalt the sprit.

How I Look for Problems

There is no problem selection strategy that is guaranteed to work for everybody. My own view is tempered by my exposure to Frank Spitzer, a professor I knew at Cornell when I was an undergraduate. Spitzer had a saying that guided him: **"It's not the theorem that is important, but the phenomenon."**

In practical terms we end up with a plan that goes like this:

- **Take almost any problem of current interest.** There is more in this step than meets the eye: the point is that if one works on stuff that is too out of fashion there is the risk of isolating oneself too from the action. Naturally if you can anticipate what WILL become of interest --- well, so much the better.
- **Become familiar with the key results**. If there are proofs, learn these. If there are computations, try them out on examples, perhaps replicating the examples of a paper. Replication is not as easy as it sounds, and it often turns up interesting issues. Incidentally, Gary King has written beautifully and extensively about the value of replication.
- **Strip away as much mumbo jumbo as you can**. Find and fix faulty definitions --- these are often a problem in highly active fields and in applied work. You might also take the popular "story" and ask if it really holds water. Surprisingly often it does not, or, at a minimum some part does not.
- Now, with a little work under your belt (perhaps just a week!), start trying to **identify the core phenomena of the area.** These will probably overlap with themes expressed by others, but put these phenomena in your own words. You will surely **start to draw distinctions that were not drawn before.**
- Remember: focus on the phenomena and let that focus pop you out of any ruts that have already been worn in the path.
- **Next, pose *simple questions* that articulate concrete features of the phenomena of interest.**
- Given each newly articulated question, put a little work into it. Questions almost always have bugs at the beginning. You need to do a little computing and a little concept refining before the real issues emerge.
- Keep good notes of your work. If you are writing code fragments, look for ways to keep these organized. I am personally terrible at this, but I keep trying.
- **If you have a concrete theoretical or empirical result, even a modest one, Latex it out.** Many many times I have started to latex something that I thought was almost trivial, and --- just in the course of being careful ---I would discover that I had to "overcome some objection." This is a very good thing. It deepens the work as one goes along. This phenomenon works very reliably for me . Now I count on it happening.
- Relax a little. You don't have to worry about how little pieces fit into a bigger picture. Eventually they will fit.
- Augment your daily work on your problems by **talking to other people about what you are doing.** We learn a lot when we try to explain our work. New issues, related results, expository "glue" --- these turn up very easily in conversation.
- With three hours concrete work and two hours of more generalized reading (and snooping), you will be astounded by how much progress you can make. It's not instant, and sometimes there are reverses. All of this is part of the process, and none of the effort is ever wasted.
- **That elusive first paper will take longer than the second, but you may be surprised how quickly you can do something that is honest and concrete and that you can be proud to have done.**

Finally, a Few Tricks of the Trade

I have gone over what I think are the core steps. You should be honest with yourself and your professors about where you hope to be in five years. You should jump through the hoops as expeditiously as you can. You should meet with a professor (any professor, anytime) and talk about possible research problems. When you find a topic that interests you, you should start logging your hours. You can count on making progress. It will happen.

So, now you are successfully moving along a path where (in steady state) you can reliably publish two or three nice papers per year. What else should you do to make sure that you are properly acknowledged?

- While keeping your eye on your thesis (i.e. paper writing) goals, you still need to **be the best citizen you can be**. It is essential to do an honorable job on your teaching responsibilities. Outstanding jobs build up favors in the favor bank. Yes, the favor bank is real.
- Get coaching from your friends to help you **build a useful web page**. You can put some personality into it, but don't go overboard. I probably go a bit too far, but I am at a different stage in life.
- The main purpose of an academic web page is to advertise ones academic accomplishments. Many academics are hesitant to indulge in self-promotion. This is a good thing. It increases the market value of the people who tastefully advertise what they have done.

- Specifically, **post your papers on the web** as soon as they are ready to be submitted to a journal. Also, post your papers to the appropriate archive, or archives. Post in PDF and have hyper-reference hot-link to your web page (and mine, thanks!).
- **Win Prizes**. Everyone should do their best to get one of our departmental prizes: Murray and Bursk. These look mighty good on the old C. V. If you have a shot at another prize (Best Bayesian Thesis, etc.) talk to your professors about this. Almost always they will want to help you win. It's great for you and for the department.
- **Also, don't feel that you must win**. Prize committees are always made up of "big cheeses" and every nominated person becomes known to these decision makers. It is victory enough just to be considered.
- **Apply for Grants**. Yes, I am not kidding. As students you can apply for **ASA and IMS travel grants** as well as lots of other stuff. These look good on your C. V. and they are always coupled with a good experience, such as giving a poster, meeting young researchers, learning about how meetings work, etc.
- **Give Great Talks.** You can never practice enough. ***The strange nature of scientific publication is that for most people and most papers --- more people will learn of the work through your talking than through their reading.***
  - Your talk should have a plan. You should show your plan on the first display of your presentation.
  - Your talk should have a beginning, a middle and an end.
  - Toward the end, identify the key "take away's" for the audience. This takes guts, but it's very powerful. If you want people to remember something, tell them to remember it. They will.
  - If you have not got enough "goods" for a full talk, then include (with proper attribution) an exposition of another persons "goods" --- if the exposition is clear and informative, it will be greatly appreciated.
  - You should end on time. The penalty for running over is HUGE (cf John Baez's essay Advice for a Young Scientist).
  - And if you genuinely find yourself ending early, have a reserve module that you can whip out to **exactly** fill the gap.
  - Thank the Audience at the end (otherwise they won't know when to clap, and everyone will feel awkward.)

  Publish What You've Got (Almost) as Soon as You Get It. We all pine for results that are stronger and more interesting than the ones we have in hand, and in times past it may have been reasonable to take things at a slow and steady pace. My own belief is that now the **optimum strategy for the young professional** is simply to publish as quickly and as voluminously as possible. Certainly there is room for judgment here, but, if you write carefully and clearly, and if you have a result that has an honest punch line, I would encourage you to "ship it." This strategy is at least epsilon-optimal, and people who avoid it should ask themselves why. Is it really to squeeze out the last epsilon, or is it something else?

Steele's Four Eccentric Principles of Editor Relationship Management

1. "If it is worth publishing once, it is worth publishing twice." (Now **don't be silly** here. This principle requires a mature interpretation, not a xerox machine. Find yourself some *good examples.* There are many from all the "top people." Rota observed that many of even the greatest mathematicians can be viewed as "one trick ponies.") [Note Added (11/17/2011) in Rota's essay, "Ten Things I Wish I Had Been Taught" he has a section that he unabashedly labels "Publish the Same Result Several Times". Rota's leading example is F. Riesz.]
2. "A paper needs both story and technical content. Too much of either is a bad thing." There is a French quote that goes something like:: "L'art d'ennui consiste en dire tout" --- if you want to bore 'em, tell 'em everything you know.
3. "If an editor tells you to change something, the optimum strategy is simply to ***change it as he suggests*** and send it back to him ASAP --- say within a week." Specifically, to quibble with an editor is pure silliness. The only real consequence is delay, even --- or perhaps especially --- if you are "right", which I am perfectly willing to assume.
4. "If you write a paper that is rejected, don't worry about it. Life is random. Take the referee reports, read them thoughtfully, make any changes that make sense to you, and send the paper to ***another journal.***" Again, you should do this damn fast. I'd hope to close the loop within two weeks, even if co-authors are involved.

> **Incidentally, the last two principles are monumentally important. When I hear anyone suggest *anything* else, it's like hearing the squeal of locked wheels on wet pavement. Always scary, often expensive, sometimes tragic.**
>
> **Final Note: These "principles" are bone-and-gristle pragmatic**. If incorporated into your career plan, they will save you time, save you worry, increase your productivity, and --- I would argue --- benefit society. I could put a rosier spin on these principles, but what would be the fun of that?

Other Resources

The advice of Lasse Pedersen parallels the point of view taken here, but in some areas he goes into more detail. If what I have said makes sense to you, reading Pedersen's essay can put some more meat into the pot.

Harvard Economist Greg Mankiw has a blog page on Advice to Graduate Students with links to a slew of other similarly titled essays. I learned a lot from reading these.

If you read Mankiw's essay and follow the links, you might want to keep in mind that economics seems to be a "meaner" subject than statistics, as least in the sense that many career changing events are powerfully influenced by matters of "taste." Also, economics journals have high rejection rates and a slow turn-around rates.

In mathematics, computer science and statistics the turn-around rates for journal submissions are long compared to biology and medicine, but short compared to economics, other social sciences, or (gasp!) the humanities.

The Academic Archipelago of Statistics

It is interesting to think about academic statistics as if one were an anthropologist contemplating the island cultures of some remote archipelago. There would be the larger "Ph.D. Granting" islands and also a large collection of other islands that are almost too remote and heterogeneous to contemplate.

We can spin this theme out in real time before I dare put a version on the web.

One take-away is that we live in a small community where relationships last for decades.

Our community is far less dispersed than mathematics or computer science. Even ten or even twenty years from now, many of you will regularly see people whom you now see every day.

Navigation: Steele's Home Page or The Always Amusing Rant Page or perhaps Surprise Me!

Appendix:

Varian Suggests: Slash Your Introduction!

"There are three parts to a seminar: the introduction, the content, and the conclusion. My advice about introductions is simple: don't have one. I have seen many seminars ruined by long, pretentious, contentless introductions. My advice: say a few sentences about the big picture and then get down to business: show them what you've got and why it's important." --- Hal Varian from his essay, "How to Build an Economic Model in Your Spare Time".

We Are So Lucky to Be Science Authors

"Thank you for sending me a copy of your book - I'll waste no time reading it."
- Moses Hadas (1900-1966)

This famous publishers' quote hangs like a life sentence over the heads of aspiring authors of fiction --- or biography, or history.

The people who write for the scientific and technical trades are **massively luckier**. If your technical book is up-to-date and professionally executed, then, unless you get stubborn about something silly, your book will find a publisher.

Writers of fiction have no such guarantee. They are perhaps well compensated by getting to live in a world of their own imagining, but golly it takes guts to write 500 or so pages that may never be read by anyone.

View From Olympus

Terence Tao is on almost everyone's list of the greatest mathematicians in the world. Only the thinnest sliver of his work is in probability or statistics, but many statisticians and probabilist would rank him as a leader in either field. The real miracle that Tao is also a prolific (and delightful!) expositor of deep mathematical ideas.

Everyone should review his excellent page of advice for graduate students (and others).

Back to Steele's Homepage.