

船井財団報告書

東京大学計数工学科数理情報コースを卒業して今年から Harvard 大学の統計学部の博士過程に進学する上原雅俊と申します。Harvard 大学では今の所、因果推論の研究を行う予定です。日本はアメリカにある Department of Statistics と同じ規模の統計学部を持つ大学がほぼないので、“統計学部”という言葉に馴染みがない方も多いかもしれません。統計学部は名のとおり、統計を研究する学部ですが情報学部、医学部、生物学部など多様な学部と兼任している先生が多く、学生のバックグラウンドも多様なことが特徴です。また最近だと時代の潮流を考慮しているのか、修士用のプログラムとして Data Science や Machine Learning という名前を冠したものを作っている所も多いです。今回の報告書では何故、海外の大学院への進学を考えたかを簡単に説明しその後、出願準備で思うところについて書こうと思います。

海外の大学院進学を考えるまで

私は大学で統計、機械学習と言われるものを研究・勉強してきました。といっても、大学1年生の頃はそれらの分野に一切、興味はなく数学のゼミを友達と一緒にして過ごしていました。そのような中で最初は数学をかなり使いつつも現実との接点が見えやすい確率論が面白いと思い、数学科に進学しようかなと思っていました。(東大では2年生の後期に進学先を決めます)しかし確率論でメジャーな応用先といえばファイナンスですが、あまり興味が持てないことに気づいて計数工学科という何でもできる所に進学することにしました。学部時代は何をやろうか非常に試行錯誤しましたが(今もしていますが)、統計、機械学習は数学・情報を使えて面白いし、様々な分野と繋がっているのが段々そちらの方へと移行していきました。

統計や機械学習は日本よりアメリカの方が、層がかなり厚いというのは否定しがたい所だと思います。ということで海外の大学院に進学することにあまり躊躇いはありませんでした。実際にアメリカには沢山 Department of Statistics がありますが、日本に統計学部として Department を持っている大学は滋賀大以外ないです。アメリカの統計学部には一応、体系的に教育する環境があり、博士課程に進むぐらいやる気のある学生が多くいて、研究者の層も厚いということが魅力的だと思います。(加えて奨学金も充実している、国際性も獲得できるというメリットも) 勿論、最近、理研で人工知能センターが出来ているなど(研究生として学生も受けれている)、日本でも魅力的な環境は増えてきているのだと思います。

出願準備

先ほど書いた通り、私は Harvard の Department of Statistics に進学します。ということなるべく一般的に書きつつも Department of Statistics の博士課程を受けることを念頭におきつつ出願準備で思ったところを記そうと思います。ちなみに機械学習を大学院でやりたい人は Department of Statistics か Department of Computer Science を受けることになると思いますが、だいぶ選考基準は違うと思います。大まかですが Statistics の場合は数学のバックグラウンド、CS の場合は CS のバックグラウンドが問われるのだと思います。またアメリカには生物統計学部すなわち Department of Biostatistics がありますが、そこもだいぶ違う選考基準を持っていると聞きます。Statistics の場合、そもそも面接はないしあっても形骸的なもので、コミュニケーション

ジョン能力はほぼ見られていない気がしますが、Biostatistics の場合、面接は必須でそれなりに重要視されるとよく言われています。

研究・推薦状

最も重要で大変なところだと思います。基本的には一緒にしたプロジェクト・研究経験を書いてもらうのが推薦状です。ただ統計学部のアプリケーションの場合、数学科出身でしっかり現代数学を学んできたという事が売りであれば研究経験ではなくゼミのパフォーマンスなどが問われるかもしれません。(今度から Berkeley に行く数学科の友達はそのような様子でした。)一方でそれ以外のバックグラウンドの場合、研究経験は必須だと思います。

私の場合、3年生の終わりまでに TOEFL や GRE の subject を受けていたのですが、逆に研究経験が一切ありませんでした。所属していた計数工学科では卒論が4年生の後期からと遅くてこれでは中身のあることは書いて貰えないと思ったので、他の学科の研究室で X 線画像からニューラルネット を使って画像診断の補助を行うという研究に参加していました。その中で、GANs という深層生成モデルの拡張を思いついたので、それで理学部情報学科の先生と論文を書き、機械学習系の国際学会で発表することができました。その学会で何人かの先生と会うことができ、発表を聞いてもらったりしました。(ただその先生がいた大学院は落ちてしまったのですが...) また似たネタで統計的な研究(規格化されていないモデルの漸近分散の改善の研究)にも繋がり面接でも話すことができました。ともかくこれらの研究経験がなければ船井財団の奨学金も大学院も受からなかった気がします。

しっかり研究を頑張っていたら推薦状はついてくると言われますが、学部生の場合、研究を始める時期が遅く指定されている、実は学科内にやりたいことをやっている先生がいなかったなど、敷かれたレールに沿って頑張ってもどうしても推薦状の内容が薄くなってしまう場合もあるでしょう。そのような場合は自分でレールを敷く(例えば早く始めさせてもらう、学部・大学問わず自分で先生を探して研究をする)程度のアクティブさが良い大学院に受かるためには必要なかもしれません。

TOEFL と GRE

TOEFL は 101 点でした。大学3年生の時に受けていたので、本来もっと上げる余地はあったと思われませんが、船井財団の先輩が TOEFL で少々の点を上げるのであれば良い推薦状をもらうために研究をしっかりやったほうが良いと言っていたのでそれに従いました。多分、(受かってから困ることをおいとけば)受かるのには 100 点あれば十分ではないでしょうか。同様に GRE の General はそれほど重要ではないと思います。実際、私は Verbal 147 Math 169 Writing 3.5 という最低限の結果でしたが、一応、博士課程に受かっています。ただ統計学科だと GRE の Subject の Math はそこそこ良い点数が取れているか見られると思います。アジアからの留学生は数学が出来ることが期待されて採用されるので、最低でも 90 パーセンタイルに入るべきでしょう。私の場合、3年生の9月に GRE の subject を受けました。(GRE の subject は年に3回しか受けられない)この策は結構良かったと思っています、おかげで4年生の時、他のことに時間を回す余裕ができました。

成績

統計学は理論的な側面が他の学問と比べて強いので、相当な実績がない限り良い成績を取っておくべきだと思います。例えば Duke 大学(統計学科の ranking だと 10 番以内に入る)では合格者の平均 GPA は 3.8~3.9 と書かれています。私もほぼ大学の授業を聞いた記憶はないですが、上のようなことから成績は A(4)を取っておくようになるべくしていました。勿論、良い研究実績やそれに基づく推薦状があれば、成績が多少悪くても十分にカバーできると一般的に言われています。

奨学金

船井財団の奨学金ともう一つの奨学金を受けました。結局、受かったのは船井財団の奨学金だけでした。いくら実力があっても、条件の良い奨学金は競争率が高く大学院の選考同様、実力に加えてどうしても運的な要素が絡んでくるので、多く受けるに越したことがないのでしょうか。

CV と Statement of Purpose

CV は嘘をつかない範囲で大きめに、かつ簡潔にアピールすることが重要だと思います。大学院の CV でもし見られる所と言え、研究経験、インターン経験、Github のアカウント、web page くらいだと思ったので、私はそこに焦点を当てて簡潔に書きました。また Statement of purpose はかなり誇張して書けるので、あまり合否に関係ないという意見もありますが、それはある程度書けていることが前提だと思われます。例えば A さんのところで B をやりたいと言っているのに、A さんは B を今は実はやっていないということがあったら “こいついい加減なこと言っているな” とおそらく思われるので、受ける大学や志望する先生のことはいっしょに調べてそれに合わせてかくべきなのでしょう。

結果

博士課程で受かったのは Duke University と Harvard University でした。あと補欠や修士で Columbia University, University of Washington, University of British Columbia に受けました。(他には CMU, Stanford University, University of Texas at Austin etc を受けて落ちました(泣)) 今更ながらもっと受けておけば良かったと思いました。例えば University of Pennsylvania や UC Berkeley の統計学科には良い先生が沢山いるので受けとくべきだったかなと後から感じました。先生方には面倒をかけることにはなるとは思いますが、受かってからキャンパスビジットなどで分かることも多いし、受かる前の “どこ行きたいか” より受かってからの “どこ行くべきか” に悩めるよう多く受けとけくべきであるのだと思います。また現実的な話、統計や機械学習の博士課程に受かる人は結構、修士課程や企業を経由してから来ている人が多いので、学部生でいきなり良い所の博士課程に受かることは難しいのかもしれませんが。ただ学部生に期待されているハードルは修士課程の人よりおそらく低いので、そこを逆にチャンスと考えることもできると思います。

最後に

出願準備はかなり大変でしたがこれからのことを考えるとやって良かったと思います。特にアメリカの統計学部に進学する人が日本からほぼいないので、情報集め

にかなり苦労しました。ただ結局の所、推薦状が一番見られてそれ以外のもの(成績, TOEFL, GRE, Statement of Purpose)はそこそこのクオリティが求められるということが重要なことなのかなと思います。最後になりますが、奨学金を含め様々なご支援をして頂いている船井財団の関係者の方には大変感謝しております。お陰様で当面は金銭的な余裕ができるので、その分しっかり研究や勉強に励んでいこうと思います。

おまけ : Harvard の統計学部 の博士課程について

せっかくなので進学する Harvard の統計学部の特徴について書いてみようと思います。一つ目は教員が少ないということです。なのでやりたい分野の教員がいないことは容易にありえると思います。ただ個々のファкультイは最もレベルの高い大学の一つのはずです。またこれからどんどん教員を増やしていく予定と Dean の人が言っていました。二つ目は学生が少ないことです。普通、統計学部の PhD course の学生は一学年 10 人以上いるのが一般的ですが、同期は私を含めて 6 人しかいません。三つ目は機械学習 only の人(統計の Journal に出さず国際学会で発表ばかりしている研究者)、数学マッチョ系の教員(ノンパラの理論的な話、経験過程とかやっている研究者)はいないです。実際にある様々な問題を解くために、統計手法を作るといった意識が重要視されている気がします。また所謂、機械学習 only の人がいないのは Harvard の CS がそもそもあまり強くないということがあると思います。ただ Harvard の CS 内で機械学習に弱いことはかなり意識されているらしくて、Microsoft から Harvard に来た有名な研究者が、これからもっと機械学習の研究者を採用していきたいと言っていました。四つ目は統計学の中の因果推論という分野について世界で最も秀でたグループが集まっていることです。このことを書いてると長くなりそうなので、いつかまた書けたらと思います。五つ目は統計学部以外に統計の強い研究者達が Harvard 内にかなりいることです。例えば公衆衛生学部疫学科の Causal inference の研究グループ、公衆衛生学部生物統計科のグループ、政治学部の計量社会学のグループはどれも世界最高峰のレベルなはずです。(私も今、報告書を書きながら公衆衛生学部の Causal inference の講義を受けにボストンに来ています)