

統計学を学んで  
奥井佑 (九州大学)

この度は、計量生物学会の奨励賞に選出いただき誠にありがとうございます。本来計量生物学への思いなどを記す場ということですが、まとまった考えを持ち合わせていないため、この場では、現在までの研究活動を中心に記します。

私は東京大学医学部健康総合科学科、いわゆる“保健学科”の出身であり、学部時代には統計学には全く関心を持っておりませんでした。統計学に関心を持つ前は法律などの社会科学の勉強をなぜか熱心にしておりました。学部3年次に、統計学を受験科目として東大の経済学研究科が受験できることを知ったのが統計学を勉強し始めたきっかけです。その後、学部4年次より疫学・生物統計学研究室に所属し、博士課程まで計6年間同研究室のお世話になりました。当初は、疫学や臨床試験に関連する統計学の方法論を中心に学び、生存時間解析や因果推論について卒業論文や修士論文では取り組みました。博士課程に進学したのち、弘前大学の中路重之先生が主導する研究プロジェクトに参加する機会を得たことにより、自身の方向性が大きく変化しました。

弘前大学医学部では中路重之先生を中心として地域住民を対象とした疫学研究を行っており、住民の方から収集した多角的なデータをどう分析するかということがデータ解析のテーマとしてありました。多変数のデータを統合的に解析できる可能性がある手法として、機械学習が使えるのではないかと当時は漠然と考え、機械学習に関心を持つようになりました。生物統計学の方法論の研究において、複雑な状況下において治療（曝露）の効果をいかに推定・検定するかということやそのための研究デザインが話題の中心になることが多いという印象があります。そのため、変数が多いデータに対して機械学習を利用して探索的なデータ分析を行うというデータマイニングの発想はとても新鮮なものに感じられました。同時期に、ビッグデータ解析や機械学習という語がバズワードとなり空前のデータサイエンスブームが起きており、世の中に機械学習関連の書籍があふれるようになっていたことも機械学習に関心を持つきっかけになったと思います。それら書籍にふれる中で、岩田具治先生が書かれたトピックモデルに関する書籍を通じて初めてトピックモデルの存在を知りました。本の中では今まで見たことがない数式が羅列されており、興味深いモデルだと感じて今後の研究対象に決めたのを覚えています。また、弘前大学における疫学研究では腸内細菌のデータを収集していることから、当初トピックモデルを適用する対象として腸内細菌データに関心を持つようになりました。弘前大学の研究には奥野恭史先生や井元清哉先生といった一流のデータサイエンティスト・バイオインフォマティシャンも参加されており、そういった先生方の存在を知ったことも自身の方向性に大きく影響を与えました。バイオインフォマティクスや機械学習は他分野のものだと当初は思っていたこともあり、情報系の学生が学部で習うような基礎的な内容を博士課程にて1から勉強する羽目になりましたが、そのことが現在のキャリアにも運良くつながっていると考えています。

奨励賞をいただいた腸内細菌に関する論文に関してですが、機械学習手法を細菌データにどう応用するかについて検討した論文になっています。細菌データ解析の方法論については近年、海外の生物統計学の一流雑誌においても掲載事例が増えてきていますが、国内においては疫学研究などで細菌データが収集されることはまだ一般的とはいえ、解析の方法論の研究者も数少ないのが現状です。研究テーマが決まらない学生の方などは今後研究してみてもよいかもしれません。

現在、私は九州大学病院の医療情報部に所属しており、院内の情報関連の業務や観察研究の支援、統計の方法論の研究などを行っています。研究面では、企業や他の研究者との共同研究を通して、院内の電子カルテやレセプトデータを研究活用する方法を検討しています。データを抽出するための環境を整備する、あるいは非構造化データを整形するうえで情報学的知見も必要です。また、実際のデータ解析の上では、多変数を探索的に扱う機械学習と同時に、臨床家にとっては馴染みが深い医学統計も必要だと感じています。

最後に今後の抱負として、元指導教員である大橋靖雄先生が佐久間昭先生の言葉を引用して、統計家は研究の墓堀人ではなく産婆であれということをよくおっしゃっていました。医学分野において、現実の医学的問題に取り組む都合上、問題設定や研究の立案は基本的に医療関係者が行う必要があります。統計家は補助的な存在であることが多くあると思います。ただ、データサイエンスの必要性が叫ばれデータ駆動型の研究が増加する昨今、既存のデータから価値を生み出すうえで、統計の専門家がより主体的に研究に携わるチャンスも増えているのかもしれませんが、私自身は、統計家というよりも主体的に研究を生み出す統計学者を目指して、今後励んでいきたいと考えています。