



しもかわ あさなお
下川 朝有



東京理科大学理学部第
二部数学科

[研究室]下川研 [分野]応用統計

研究内容：生物統計学

より質の高い医療のために、生物統計学ができること

世の中には知りたいことが沢山あります。身近な例として、貴方は「ある湖の中に何匹の魚がいるか？」に興味を持ったとしましょう。さてこの疑問を解決するには、どのようにアプローチしていけば良いのでしょうか？当然真の値を知りたいければ、その湖の中にいる魚を全て捕まえる（印をつける）しかありません。さてこのような調査を行うことは可能でしょうか？残念ながらこれは現実的には実現不可能です。魚を確実に全て捕まえるためには湖の水を全て抜く方法などが考えられますが、そのためにかかる膨大なコストや時間を考えればそれは明らかです。

そこで思いつく代替法として、「全部を調べるのは無理だけど、特定の調査方法で数匹について調べることで（サンプリングすることで）、全体の匹数がある程度の精度で分かるのではないか？」というものではないでしょうか。そしてその際に、どのようにしてデータを集めるべきか？何匹程度について調べるべきか？どのようにして真の値を推測するべきか？そして誤差はどの程度か？といった疑問が生じます。統計学はこれらの疑問に対して、数学の理論を背景に答えを探していく学問となります。

前置きが長くなってしまいましたが、統計学の一分野であり、医薬や医療の研究に大きく関わっているのが「生物統計学」です。例えば、ある特定の病気に効果があると期待される新薬を開発したときに、その新薬が従来の薬と比べて有効であるかをどのように調べて結論を下すか？、また別な例として、性別や年齢、既往歴といった様々な背景に基づいて、ある患者さんの死亡までの時間は確率的にどのように推移していくのか？といった問題を考えていく学問になります。私はその中でも特に、がんを患った患者が、再発もしくは死亡するまでの時間を確率的に表現するモデルをどのように構築するか、といった研究を行っています。このような研究は実際の患者さんに対して重要な情報を与えてくれるばかりでなく、新たな治療法を確立するために大いに役立つものとなっています。

研究内容：統計的機械学習

膨大なデータから、価値を生み出す

統計学は、実用性の高い学問といわれています。特に近年、多くの研究やビジネスの現場では、非常に膨大で複雑に入り組んだデータ、すなわち「ビッグデータ」をいかに有効に扱うか、またそのようなデータを扱うことができる人材が求められています。皆さんもAIや機械学習といった言葉を様々な場面で目にすることがあるかと思いますが、その理論の多くの部分で、統計学が役に立っています。

私はそのような膨大なデータから、機械学習を用いることで、人間の目では判断できない新たな価値を探し出す研究を行っています。一例として実際にある製薬会社と行う共同研究では、「機械学習を利用した新薬開発の研究」があります。薬は化合物であるため、それを構成する原子と原子間の関係性で表現することができますが、その組み合わせは膨大に存在し、実際に薬として役立つ化合物を探索するには膨大な時間とコストがかかります。そこで過去の膨大な成功と失敗の化合物のデータを用いて機械に学習させることで、薬として役立つ可能性がある化合物を予測する、といった研究を行っています。

最後に、統計学は数学を背景に、世の中の様々な場所で役に立っています。そしてデータがあふれる現代社会で、それはより顕著に表れてきます。統計学に興味がある皆さんは是非、一緒に研究をしていきましょう。