

夢見る頃を過ぎても

統計学と物理学の間で

伊庭幸人

統計数理研究所 予測制御系

スランプ脱出のきっかけは、よき研究仲間との出会いだった。着眼の独自性をいかに発展させるか、いささかの自負を胸にスタートした研究。「夢」はすぐにも実現しそうだった。院生以来10数年の研究生活を振り返りつつ、いま改めて大学院教育に求められているものは何かを考える。

統計学との出会い

最初に面白いと思ったのは、分割表の解析の話です。当時、私は統計数理研究所（統数研）の共同利用研究員で、友人が所員だったので、彼の部屋で遊んでいて、所員の坂元先生の本を見つけたんです。社会調査とかで100項目くらいデータを取ったのを、どう整理して自動的に最適な分割表を作るかという話なんです。が、「最適」をどう定義するかが自明じゃないんですね。

分割表より簡単な例として、ヒストグラムの切り方で説明しましょう。ヒストグラムを粗く切りすぎて豆腐みたいにすると元の情報がなくなってしまうし、かといって際限なく細かくしていくと元のデータに逆戻りしてしまいます。どこかでちょうどいいところがあるような気がしますが、それを決める根拠は何なのか。統数研の赤池先生のまわりに集まったグループは、こうした疑問を予測の最適化という観点から追求していたのです。数理的なこととか、議論の中身以前に、こういうことが学問になるとは物理学科では習わなかったのでびっくりしました。

最適化からサンプリングへ

もともと、私の専門は統計物

理で、当時はセル・オートマトンによる疫学モデルとか、いまでいう「複雑系」みたいなことをやっていたんです。それで、統計学に開眼したといっても、具体的に研究者としてどう貢献するかは、また別です。

そこで、モンテカルロ法の話になります。'80年代、シミュレーション・アニーリング（模擬焼きなまし法）が流行りました。LSIの設計等をコンピュータで解くのに、乱数を入れて徐々に「冷やす」計算をしながら、いい解に辿り着くやり方です。以前から「揺らぎが大事」という「精神論」は聞いていましたが、実際にいろんな問題が解けるっていうのはいいな、と思いました。しかし、統計物理をやる立場からいうと、揺らぎをもっと積極的に使えないか、というのが気になります。もともと確率分布を扱う学問である統計学への応用を考えたら、冷やさない方向、揺らぎのある答えが出てくる方向に可能性があるんじゃないか、と思ったんです。

統計の世界でも、すでに、統数研の尾形先生、種村先生をはじめとして、冷やさない方向、つまりサンプリング手法としてのモンテカルロ法を使って研究されている人たちがいましたが、これはもっと全面的に展開

できるはずだと考えました。

統計的推論のわかりやすい例として、遺伝子の伝播を推定する問題があります。親の持っている遺伝子がこうなら孫がその遺伝子を持つ確率はこれこれと、上から辿る「順問題」に対して、子や孫などを含めてわかっている部分——劣性遺伝子の場合とはとくに部分的な情報しかないわけですが——から出発して、家系図の全部の人たちについて確率を求めるのが「逆問題」です。データから始めるんですね。

こういう推論をモンテカルロ法でやると、答えは順列組み合わせで作出した「ありうる世界」を動き回りながら出てくるんです。ありうる世界をその度合いに応じて、次から次へとスキップしていく。最適解から離れるほうにもある確率で動かしてやる、つまり冷やさないの、ずーっと全体が動いています。イメージ的にいえば、方向も時間の向きもない、ランダムに揺らめく巨大な事象の世界です。

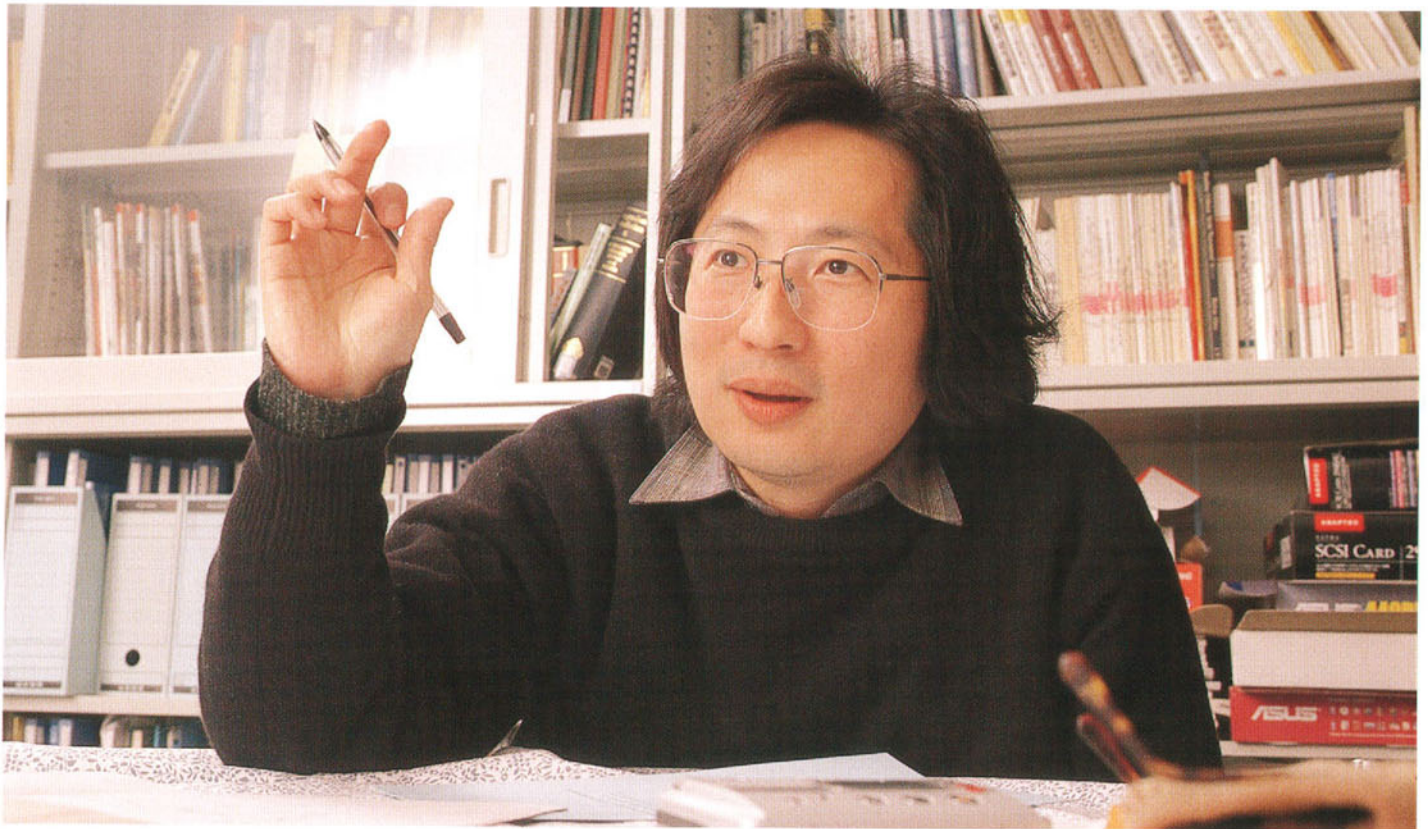
人間、どこかで卒業しないと

まあ、言うのは簡単ですが、なんでもできそうなのに、どうい問題を解けばいいのかわからない。いま家系図の例を出しましたが、これは後になって知った話です。複雑な推論の例と

して、じゃんけんの三すくみに関係した問題を英語の論文に書きましたが、パズルみたいだと言われて落ちてしまったんです。

相手が大きすぎてテーマがしばらく切れない、とモタついているうちに、海外の統計家たちが一齐にこの手法での論文を発表し始めたんですよ。統計学では投稿から雑誌掲載まで時間がかかるということもあるのですが、こちらからみると、ほんの3、4年くらいの中にすっかり普及してしまったという感じでした。その頃には、統数研に就職していたのですが、海外出張から帰ってくる人が、みんな、「Gibbs samplerとかいうのがはやってるけど、あれって何」って。Gibbs samplerっていうのは、さっきから話してるモンテカルロ法の一種です。なんだ、みんなもう知っているのか、と正直がっかりしました。

当時みなに言われたのですが、今思えばそのとおりで、恥をかいてでも学位論文を出してしまえばよかったんです。英語の論文も書かなきゃだめですね。はじめ、太平洋はおれのものと錯覚していたのに、瀬戸内海になったらがっかりですよ。本当は瀬戸内海でも食べきれないくらい魚はいるんだけど。あるところで人間、卒業し



伊庭幸人 (いば・ゆきと) : 中学校の授業で原子論を聞いて以来、「微細なものの集まりとしての全体」のイメージが頭に住みつき、物理から統計物理、統計へと進む原点となっている。東京大学理学系研究科博士課程を経て、'88年、統数研入所。総研大教官として、平成10年度国際シンポジウム「複雑系への戦略—構成と記述」のオーガナイザーを務めた。
伊庭幸人ホームページ= <http://www.ism.ac.jp/~iba>

撮影：由利修一

ないといけなく、そこどころが私の場合うまくいかなかったんですよ。しばらくスランプが続きました。

パートナーシップの成果

まず、仲間を作らないとだめだ、と思って。阪大の物理でゼミをやったのがきっかけで、阪大の人たちといっしょに仕事をすることになって、その頃からですね、やっと再起し始めたんです。

新しい研究方向の1つは、モンテカルロ法の新しい手法の開発です。以前は、物理ですでに用いられている手法を統計に使うということでしたが、今度は方法論自体の研究をしよう。その中でも割合うまくいったのは、格子ポリマー・格子たんぱく模型のためのモンテカルロ法の開発ですね。格子たんぱく模型というのは、たんぱく質の性

質を定性的に調べるために作られたおもちやです。格子の上に横たわった紐みたいなもので、紐の形に応じてエネルギーが決まるのですが、それで重みのついた紐の形の分布から公平なサンプリングをしろというのが問題です。

モンテカルロ法では、紐が自分自身に重なってはいけないという条件が障害になります。私たちが考えたのは、重なってもいいけど、重なったら文句(罰金)が出るようにして、その文句の程度を調節してやるというものでした。文句の度合いをコンピュータが自動的に調節してくれるというのがポイントです。阪大の院生の1人とその助教と組んで、共同研究をすることになったんです。途中、むずかしいところもあったし、近いアイデアの先行研究も見つかったんですが、パートナー

の頑張りにも助けられて、これならという結果が出ました。山中湖か井の頭公園の池かわからなくても、1つ仕事できたわけです。

こういった仕事をいくつか集めて、40歳でようやく理学博士になれました。でも、論文の前半は、10年前に出そうとしたものと内容的にはほとんど同じなんです。その部分を外したほうが、物理学科に出すものとしてはまともまりが良かったのですが、やっぱり意地もあるし、数を頼むというのと、両方ですね(笑)。

形容詞なしの「科学者」を夢見て

学生はいろいろな価値観、研究観、気質の先生や研究室に出会うことが大事ではないでしょうか。手でいじれる問題をやりながらゆっくり考えるタイプの学生が、才気煥発に思想を語る先生しか知らなかったら、自分

は馬鹿だと思い込んでしまうかもしれませんね。反対に、グローバルな状況を見通せる学生が、問題解き専門の先生とずっと付き合うのも苦痛でしょう。

また、より大きなレベルでも、価値観の違いというのは大事です。数学・科学・工学と教条的に分けて済みます人もいますが、統計学みたいにカテゴリー不明のものに親しんでいる立場からいうと、もっと多面的に柔軟に考えていいと思う。そのためにも、若いうちに複数の価値観を体験してほしいですね。

分野防衛的な考えが染み付いている「〇〇学者」ではなく、分野や価値観による違いを体感したうえで、それを主体的に選んだ「科学者」を育てることを今後の総研大には期待したいと思います。

(取材構成 白石厚郎)