

## گفتگو با بردلی افرون\*

سوزان هلمز، کارل موریس و راب تیبشیرانی

مترجم: کسری علیشاهی

بردلی افرون استاد آمار و آمارزیستی در دانشگاه استنفورد است. کارهای تحقیقاتی او ترکیبی از زمینه‌های نظری و کاربردی را در برمی‌گیرد که شامل رویکرد بی‌زی تجربی، تحلیل داده‌های بقا، خانواده‌های نمایی، روش‌های بوت‌استرپ<sup>۱</sup> و جک‌نایف<sup>۲</sup> و بازه‌های اطمینان می‌شود. بیشتر کارهای عملی او از همکاری با پروژه‌های زیست‌پزشکی در دانشکده پزشکی استنفورد نشأت گرفته است و تعدادی مقاله مرتبط با اخترشناسی و فیزیک نیز نوشته است. مقالات نظری او نیز اغلب با مسائل مشخص و ملموس عملی آغاز می‌شوند. در متن پیش‌رو، هر سه مصاحبه‌کننده از همکاران علمی نزدیک او هستند.

بردلی در می‌سال ۱۹۳۸ در سن پائول در مینسوتا به دنیا آمد. پدر و مادر او، اشرو مایلز افرون، مهاجران یهودی روس‌تبار بودند. برّد با بورس تحصیلی به کلیتیک آمد و در سال ۱۹۶۰ در رشته ریاضی از آنجا فارغ‌التحصیل شد. او از پاییز همان سال دانشجوی استنفورد شد و در نهایت، دکترای خود را زیر نظر روپرت میلر<sup>۳</sup> و هرب سولومون<sup>۴</sup> از دانشکده آمار دریافت کرد. از دیگر اعضای هیات علمی دانشکده آمار در آن زمان می‌توان از چارلز استاین<sup>۵</sup>، هرمان چرنف، مانی پارزن، لینکلن موزس و اینگرام ال‌کین<sup>۶</sup> نام برد. برّد از سال ۱۹۶۰ تاکنون صرف‌نظر از گذراندن دوره‌های فرصت مطالعاتی در هاروارد، امپریال کالج و برکلی، در استنفورد بوده است. او مسئولیت‌های متعددی در دانشگاه داشته است: رئیس دانشکده آمار، معاون گروه علوم، دبیر شورای دانشگاه و رئیس کمیته مشاور دانشگاه و در حال حاضر مسؤول برنامه ریاضیات کاربردی دوره کارشناسی است.

\*) Holmes, S. & Morris, C. & Tibshirani, R., "Bradley Efron: A Conversation with Good Friends", *Statistical Sciences*, 18(2003), 268-281.

1) Bootstrap method      2) Jackknife      3) Rupert Miller      4) Herb Solomon

5) Charles Stein      6) Ingram Olkin

افتخارات علمی افرون شامل دکترای افتخاری از شیکاگو، مادرید و اسلو، جایزه مک آرتور، عضویت در آکادمی ملی علوم و آکادمی علم و هنر آمریکا، عضویت در مؤسسه آمار ریاضی و انجمن آمار آمریکا، مدال ویلکزن، جایزه پارزن، جایزه راثو، و جایزه آماردان برجسته از شاخه شیکاگوی انجمن آمار آمریکا می‌شود. او در سلسله سخنرانی‌های ریتس، والد و فیشر سخنران مدعو بوده است و صاحب کرسی ماکس استاین<sup>۱</sup> به‌عنوان استاد علوم تجربی و انسانی در استنفورد است. افرون ویراستار مجله انجمن آمار آمریکا و رئیس مؤسسه آمار ریاضی بوده است و به‌عنوان رئیس آینده انجمن آمار آمریکا انتخاب شده است که از ۲۰۰۴ رسماً در این سمت قرار خواهد گرفت. بخشی از این مصاحبه برگرفته از گفتگوی ضبط شده در انجمن آمار آمریکا در ۵ نوامبر ۲۰۰۱ با حمایت مالی مرکز تحقیقات فیزر است و بقیه آن در دانشکده آمار دانشگاه استنفورد انجام شده است.

## ۱. سال‌های آغازین

تیبشیرانی: بیا از خیلی قبل شروع کنیم، اولین بار چطور با آمار مواجه شدی؟

افرون: پدر من در سن پائول، مینسوتا، فروشنده و راننده کامیون بود ولی عشق عجیبی به ریاضیات داشت. او به دقت، نتایج مسابقات بیسبال و بولینگ را دنبال می‌کرد و این اثری بیش از آنچه تصور می‌کردم بر من گذاشت. وقتی من به کلتک رفتم، در آنجا با آدم‌های فوق‌العاده باهوشی آشنا شدم ولی تقریباً هیچ خبری از آمار نبود. ما فقط یک درس آمار از کتاب سای درمن<sup>۲</sup> گذرانیدیم. من از یکی از اساتیدم، مورگان وارد<sup>۳</sup>، پرسیدم که آیا چیز بیشتری هست که بتوانم بخوانم و او کتاب کرامر را به من داد که از اول تا آخرش را خواندم. به نظرم کتاب مناسبی بود. کرامر این کتاب را در انزوا و در طول جنگ جهانی دوم نوشته است و من هم کتاب را در انزوا در کلتک خواندم. تصمیم گرفتم رشته آمار را انتخاب کنم چون آینده‌ای به‌عنوان یک ریاضیدان قرن بیستمی برای خودم نمی‌دیدم. من برای قرن نوزدهم مناسب بودم ولی توانایی ذهنی درک تجربیدی که ریاضیات مدرن را در بر گرفته نداشتم، برای همین به سمت آمار رفتم. با برکلی وارد مذاکره شدم و در مصاحبه خیلی خوبی با دو نفر به نام جرزی و اریک هم شرکت کردم که خیلی با من مهربان بودند ولی در نهایت به استنفرد متمایل شدم. وقتی آنجا رفتم، متوجه شدم که در دانشکده ریاضی ثبت نام شده‌ام چون استاد راهنمایم به آن‌ها گفته بود که من احتمالاً علاقه‌مند هستم که ریاضیات را ادامه بدهم. سال اول را در دانشکده ریاضی گذراندم و بعد به آمار تغییر رشته دادم. همان موقع بود که من و کارل، که اولین بار در ۱۹۵۷ به‌عنوان دانشجوی سال اول در کلتک همدیگر را دیده بودیم، دوباره هم‌رشته شدیم. هیچ‌کس در علم، تنها کار نمی‌کند و من بیش از استحقاق خودم از موهبت داشتن همکاران عالی بهره‌مند بوده‌ام. در رشته‌هایی مثل آمار، شما به‌تنهایی به جایی نمی‌رسید. باید آدم‌های هوشمندی دور و برتان باشند که شما را به چالش بکشند. یکی از مزایای بزرگ دانشکده آمار استنفورد این است که پراز آدم‌هایی است که مشتاقند شما را به روشی خوشایند به چالش بکشند.

1) Max H. Stein 2) Cy Derman 3) Morgan Ward

تیبشیرانی: درست است که در دوران دانشجویی، تو را یک بار از دانشگاه اخراج کردند؟  
 افرون: یکی از دلایل آمدن من به استنفورد، مجله طنز دانشگاه بود. من در کلتک یک ستون طنز می‌نوشتم و همیشه دوست داشتم که برای یک مجله طنز کار کنم. استنفورد یک مجله طنز عالی به نام چاپارل<sup>۱</sup> داشت. در ماه‌های اولی که من آنجا بودم، ویراستار مجله دچار مشکل روانی شد و کارش به بیمارستان کشید. من به جای او ویراستار شدم. در یکی از همان شماره‌های مجله، شعری به نام جوان عیاش چاپ کردیم که در آن یک مقدار زیاده‌روی کرده بودیم. من تعلیق شدم و اگر پادرمیانی افرادی مثل آل بوکر<sup>۲</sup>، هلسی رویدن<sup>۳</sup> و هرب سولومون که در دانشگاه مسئولیت بالایی داشتند نبود، نزدیک بود برای همیشه اخراجم کنند. من شش ماه معلق بودم و بعد دوباره به دانشگاه برگشتم و در این دوران مشهورتر از هر زمان دیگری در طول عمرم شده بودم. عکس‌ها هر روز در روزنامه چاپ می‌شد.

موریس: من به یاد دارم که برد به عنوان دانشجوی کارشناسی نویسنده واقعا خوبی بود. چیزی که تو نوشته بودی خیلی معمولی بود. الان مردم هیچ فکر بدی درباره چنین نوشته‌ای نمی‌کنند. تواز قریحه سرشار نویسندگی‌ات استفاده کرده بودی. من فکر می‌کنم همین قریحه یکی از دلایل موفقیت تو در آمار بوده است. توانایی‌ات در این که مطالب را طوری کنار هم قرار دهی که مردم موضوع را بهتر بفهمند.

تیبشیرانی: کمی برایمان از خانواده‌ات بگو. آن طور که من می‌دانم سه تا از برادرهای تو کار دانشگاهی می‌کنند و پسرت هم به زودی وارد کار دانشگاهی خواهد شد.

افرون: پدرم به ما نشان داد که به درد کارهای سنگین نمی‌خوریم. برادر بزرگترم، آرتور، استاد بازنشسته زبان انگلیسی است. یک متخصص برجسته در ادبیات نمایشی. او یک مجله به اسم پانچ<sup>۴</sup> در ادبیات نمایشی منتشر می‌کند که عنوانش از سانچو پانزا گرفته شده است. من پانچ را می‌خوانم. ترکیبی است از نظریات بغرنج انگلیسی و ادبیات سیاه. دو برادر کوچکترم، که دوقلو هم هستند، همیشه خیلی نزدیک بوده‌اند. دان به خدمت سربازی فراخوانده شد و به کانادا رفت و آنجا زندگی خوبی دارد. هر دوی آن‌ها در روانشناسی اجتماعی کار می‌کنند. دان مجله خودش را با موضوع خانواده درمانی اداره می‌کند و رون متخصص در زمینه احساسات ناخوشایند است. او کتابی نوشته با عنوان «همیشه عصبانی<sup>۵</sup>». پسر مایلز از آن مواردی است که برای ما جالب‌اند: یک دانشجوی علوم انسانی که در چند سال اخیر ناگهان به آمار علاقه مند شده است. او الان در چپل هیل در زمینه بازیابی اطلاعات کار می‌کند. گاهی به من تلفن می‌زند و سوالات سختی درباره تجربه مقادیر تکین می‌پرسد.

موریس: در دهه شصت اتفاقات زیادی هم‌زمان در حال وقوع بود. من نمی‌توانستم در مورد انتخاب رشته تصمیم بگیرم و آمار راهی بود که به ریاضیات و تقریباً هر چیز دیگری در کنارش پردازی.

1) Chaparral 2) Al Bowker 3) Halsey Royden 4) Paunch 5) Angry All the Time

خیلی از دانشجویهای ما همین الان هم این را یکی از جذابیت‌های رشته آمار می‌دانند. این نه فقط جذابیت که در واقع یک ضرورت است. در حوالی دهه شصت، بیشتر دانشکده‌ها در حال شکل‌گیری بودند. آمار هم موضوعی بود که ناگهان قد علم کرده بود. آن موقع آمار، حداقل آماری که ما می‌خواندیم، خیلی مجرد بود. بخشی از کار تو آمار زیستی بود ولی درس‌های زیادی هم گذراندی که عمدتاً ریاضیات محض بودند. من تصمیم گرفتم که به RAND بروم. پیش خودم فکر کردم که یک سال آنجا می‌مانم و هر چیزی که در مورد آمار کاربردی لازم است را یاد می‌گیرم. ولی یازده سال آنجا ماندم. کاربردها، ریاضیات و محاسبه، مثلث ایده‌های آمار را می‌سازند. این که این سه تا چطور به هم گره خورده‌اند، محور اصلی رشته ما است. به نظر می‌رسد خیلی مهم است که ما علاقه‌مان را به هر سه جنبه حفظ کنیم.

## ۲. شروع بوت‌استرپ

تیبشیرانی: برد، تصور من این است که کارهای تو روز به روز کاربردی‌تر می‌شوند. به نظر می‌رسد که علاقه‌ات به مسائل واقعی بیشتر شده است. آیا این مشاهده درستی است؟

افرون: می‌شود قضیه را به شکل دیگری هم بیان کرد و آن این است که من دیگر ایده جدیدی ندارم و این اتفاقی است که از تقریباً بیست سال پیش افتاده است. چیزی که من الان دارم، همکارانم و کارهای کاربردی هستند. ما در آمار موقعیتی استثنائی داریم؛ ما آخرین بقایای دانشمندان اصیل هستیم. به این معنی که می‌توانیم به رشته‌های متعددی سرک بکشیم و با آدم‌های برجسته آن رشته‌ها همکاری کنیم. این برای من روش فوق‌العاده‌ای برای ورود به یک موضوع جدید است. بعضی‌ها به کاربرد، صرفاً به خاطر خود کاربرد، علاقه دارند ولی من همیشه به خاطر آمار به کاربرد علاقه‌مند بوده‌ام. مثلاً من و راب با هم روی میکروآرایه‌ها کار می‌کنیم. این موضوع خیلی هیجان‌انگیز است، ولی من به هیچ عنوان علاقه‌ای به زیست‌شناسی میکروآرایه‌ها ندارم. البته به عنوان یک آماتور برایم جالب است ولی دغدغه اصلی‌ام این است که نظریه استنباط چطور در این مورد به کار می‌آید.

تیبشیرانی: یک موضوع عجیب برای من این است که با این که برای ما آمار موضوع ساده‌ای است ولی برای بقیه دانشمندان اصلاً این طور نیست. یکی از همکاران من که خیلی هم دانشمند خوبی است، معتقد است که برای من خیلی سخت‌تر است که به او آمار یاد بدهم تا برای او که به من زیست‌شناسی بیاموزد. زیست‌شناسی برای من خیلی ترسناک به نظر می‌رسد؛ انبوهی از حقایق اسرارآمیز. ولی آمار یک روش فکر کردن است که یاد گرفتنش برای دیگران، اگر از قبل در این رشته آموزش ندیده باشند، بسیار دشوار است. به نظر می‌آید ما از مهارتی یکتا برخورداریم.

موریس: فکر می‌کنی آمار موضوع سختی است یا ما آن را به موضوع سختی تبدیل کرده‌ایم؟  
تیبشیرانی: البته این هم بخشی از ماجراست ولی مفاهیم اساسی به ظاهر ساده آنقدر هم که به نظر می‌رسد ساده نیستند.

موریس: مثلاً ... ؟

تیبشیرانی: مثلاً آزمون جایگشت برای به دست آوردن  $p$  - مقدار از مجموعه‌ای از داده‌ها. موریس: فکر می‌کردم بلافاصله خود  $p$  - مقدار را مثال بزنم. من تصور می‌کنم  $p$  - مقدار مفهوم دشواری است. محاسبه‌اش ساده است ولی گیج‌کننده است به این معنی که مردم معمولاً تعبیر اشتباهی از آن دارند.  $p$  - مقدار قطعاً گزاره‌ای درباره داده‌ها، به فرض درستی فرض صفر، است ولی خیلی‌ها فکر می‌کنند احتمال درستی فرض صفر به شرط داده‌ها است.

تیبشیرانی: مثال دیگر، اختلاط است. تعداد بسیار زیادی از دانشمندان واقعاً خوب آزمایش‌های بدی طراحی می‌کنند که در آن‌ها عوامل مهم با هم مخلوط می‌شوند. اختلاط در رشته‌ها اساسی است. ما این پدیده را می‌فهمیم و می‌دانیم چطور آن را تشخیص دهیم.

افرون: بسیاری از همکاران علمی من در احتمال خیلی خوب هستند؛ مثلاً می‌توانند محاسبات احتمالاتی مربوط به مدل‌های ظریف و پیچیده را انجام دهند، ولی در استدلال معکوس از داده‌ها به مدل احتمالاتی مناسب، کاملاً ضعیف هستند. به یاد دارم سال اولی که وارد آمار شدم، فکر می‌کردم آمار باید برای من ساده باشد چون من از رشته ریاضی می‌آمدم که به مراتب سخت‌تر است. ولی آمار در شروع برای من خیلی سخت‌تر از هر رشته دیگری بود. سال‌ها طول کشید تا واقعاً احساس راحتی کنم. فهمیدن این که مردم چرا در آمار چنین کارهایی می‌کنند ساده نیست؛ چرا از  $p$  - مقدار یا چیزهای دیگر استفاده می‌کنند. شما باید وارد مسائل عملی بشوید و باید حسی نسبت به مسئله پیدا کنید. آمار تنها جایی است که در آن استنباط آماری انجام می‌شود. ما خدمتی واقعی ارائه می‌کنیم که تفکر معکوس است. شما از فکر کردن بر روی یک مسئله مشخص جزئی شروع می‌کنید و بعد به حالت کلی برمی‌گردید. این ممکن است از نظر فلاسفه، نشدنی به نظر آید ولی کاری است که ما هر روز انجام می‌دهیم.

تیبشیرانی: من سال ۱۹۸۱ به استنفورد آمدم. چند سالی بعد از آن که تو بوت استرپ را ابداع کرده بودی. برایمان از تفکرات و وقایعی که منجر به ابداع بوت استرپ شد بگو.

افرون: داستان بوت استرپ شاهی است بر اهمیت داشتن همکاران خوب. روپرت میلر مقاله‌ای با عنوان «یک جکنایف قابل اعتماد» نوشته بود که تلاش خوبی بود برای توجیه نظری جکنایف. آن موقع من و روپرت هر دو در یک سال، سال ۱۹۷۲، برای فرصت مطالعاتی در امپریال کالج پیش دیوید کاکس بودیم. روپرت یک سخنرانی درباره جکنایف انجام داد. بعد از آن دیوید پیش من آمد و پرسید تو فکر می‌کنی مطلب مهمی اینجا وجود داشته باشد؟ سال‌ها بعد فهمیدم که او داشت به من نشانه‌ای قوی برای کار در این زمینه می‌داد. در سال ۱۹۷۷ از من دعوت شد که در قالب سخنرانی‌های ریتس، یک سخنرانی ارائه کنم. من یک خط یادداشت نوشتم: جکنایف، تقریبی از چیست؟ همان موقع که آن خط را می‌نوشتم، در مسیر درست جواب قرار گرفتم. در ابتدا با چیزی شروع کرده بودم که واقعاً پیچیده بود. من به آن توزیع ترکیبی می‌گویم چون

به جای جایگشت‌ها، ترکیب‌ها را در نظر گرفته بودم. بعد متوجه شدم که می‌توانم از شر بخشی از پیچیدگی‌ها بگریزم. بعدتر از شر بخش بیشتری از پیچیدگی‌ها شدم و خیلی زود، دیگر هیچ پیچیدگی‌ای باقی نمانده بود. کمی بی‌معنی به نظر می‌رسید ولی من سخنرانی را ارائه دادم و تقریباً همه خیلی خوششان آمد. از آن به بعد، هیچ وقت به قضاوت خودم درباره موضوعاتی که روی آن‌ها کار می‌کنم اطمینان نمی‌کنم.

تیبشیرانی: مقاله به آنالز فرستاده شد. بازخوردشان چه بود؟

افرون: روپرت میلر آن موقع ویراستار آنالز بود. من متن سخنرانی ریتس را فرستاده بودم و مقاله رد شد. ویراستار همکار، که نامش مشخص نمی‌شود، گفته بود که مقاله هیچ قضیه‌ای ندارد. در نتیجه، من چند قضیه اضافه کردم و فشار زیادی به روپرت آوردم و او هم بالاخره مقاله را چاپ کرد. من قبلاً ویراستار مجله انجمن آمار آمریکا بودم و این ماجرا من را به یاد قانونی انداخت که آن موقع برای خودم داشتم. وقتی مقاله‌ای مردم را عصبانی می‌کند باید آن را با دقت بیشتری بررسی کرد. مقالات اعصاب خرد کن، جامعه‌ای دوقله‌ای تشکیل می‌دهند: بدترین مقالاتی که تا به حال دیده‌اید، که زیرجامعه بزرگی است، و تعداد کمی مقاله واقعاً خوب. از آن موقع تا الان، من تعداد زیادی مقاله نوشته‌ام. اگر فکر می‌کنید هر مقاله‌ای که می‌فرستم فوراً پذیرفته می‌شود، اشتباه می‌کنید. بسیاری از مقالاتم رد شده‌اند. من معمولاً برای بازنویسی مقالات زحمت زیادی می‌کشم. سخت تلاش می‌کنم و نظرات داوران را جدی می‌گیرم، ولی هیچ وقت از این که داور از چیزی خوشش نیاید ناامید نمی‌شوم چون ممکن است دلیلش این باشد که ایده جدیدی مطرح کرده‌ام.

تیبشیرانی: من وقتی بعد از فارغ‌التحصیلی تعدادی سخنرانی درباره بوت‌استرپ انجام دادم متوجه چیز غریبی شدم: صحبت در این مورد برای افراد رشته خودمان خیلی سخت‌تر است. وقتی برای فیزیک‌دان‌ها یا شیمی‌دان‌ها صحبت می‌کردم، می‌گفتند آهان! این همان شبهه‌سازی است. ما این کار را زیاد انجام می‌دهیم. ولی به عنوان ابزاری برای استنباط آماری، پذیرفتنش خیلی سخت‌تر بود چون نیاز به تولید اعداد تصادفی داشت که هنوز خوب جا نیفتاده بود.

موریس: این ماجرا خیلی به موقع اتفاق افتاد. ما ۳۰۰ دلار در ساعت می‌پرداختیم تا از یک کامپیوتر استفاده کنیم و ۳۰۰ دلار آن موقع الان احتمالاً ۱۰۰۰ دلار ارزش داشت. کامپیوترها کند بودند. هر کدام از این محاسبات از چند دقیقه تا یک ساعت طول می‌کشید و این مشکل بزرگی بود. ابداع بوت‌استرپ همزمان با آمدن کامپیوترهای شخصی رخ داد. همه ما می‌دانیم که الان محاسبات، مسأله عمده‌ای حداقل از نظر هزینه نیست، ولی از جنبه مفهومی هنوز مسأله‌ای جدی است. بوت‌استرپ مثالی است از این که تو وقتی تلاش می‌کردی نوشته‌جات مربوط به موضوعی که اطلاع زیادی درباره‌اش نداشتی یعنی جکنایف را مطالعه کنی، نظریه‌اش را به وجود آوردی. این رویکرد، خواندن نتایج دیگران و تلاش برای این که آن‌ها را یک قدم بهبود ببخشی، می‌تواند کاملاً موفقیت‌آمیز باشد. ولی تصور می‌کنم عموماً ثمربخشی آن، کمتر از رویکردی است که گفتی در بیست سال گذشته در

پیش گرفته‌ای: درگیر شدن در یک مسأله واقعی و خیلی زود به چیزهایی برمی‌خوری که قبلاً هیچ موقع متوجه‌اش نشدی.

افرون: من به خودی خود، خواننده خوبی نیستم. ولی وقتی شروع به کار کردن روی موضوعی می‌کنم، مایلیم هر چیزی درباره آن وجود دارد را بخوانم. وقتی جای پای در یک زمینه پیدا کرده باشم و بتوانم بفهمم که مردم چرا فلان کار را انجام می‌دهند، خواندن برایم خیلی ساده‌تر می‌شود. این دقیقاً سخت‌ترین قسمت نوشتن است: این‌که توضیح دهید چرا کاری را انجام می‌دهید نه این‌که چه کاری انجام می‌دهید. وقتی کسی بفهمد که شما چرا کاری را انجام می‌دهید به احتمال زیاد حس خوبی نسبت به آن پیدا خواهد کرد.

موریس: در مورد بوت‌استرپ، چرا تو به جای تلاش برای حل مسأله‌ای مشخص به کمک یک مجموعه داده‌ها، از ایده‌های نظری شروع کردی؟

افرون: واقعیت این بود که یک همکار مسأله جالبی مطرح کرده بود که با شهود من سازگار بود. حجم انبوهی از مقالات بیمایه و کسل‌کننده‌اند، عمدتاً به این دلیل که هیچ چیز شگفت‌انگیزی ندارند. وقتی آن‌ها را مطالعه می‌کنید از همان ابتدا کاملاً می‌دانید که جواب قرار است چه باشد. خیلی به‌ندرت به چیزی برمی‌خورید که برایتان عجیب باشد. یادم می‌آید قضیه بنجامینی - هوشبرگ<sup>۱</sup> درباره نرخ‌های اکتشاف غلط، واقعاً مرا متعجب کرد. من هنگامی که شگفت‌زده می‌شوم، با اشتیاق بسیار بیشتری مطالعه می‌کنم.

تیپشیرانی: چیز دیگری که تو را متمایز می‌کند این است که تو با افراد زیادی مقاله مشترک نوشته‌ای. ما از معدود کسانی هستیم که با تو مقاله مشترک داریم. تو در روش کارت مستقل و فردگرا هستی.

افرون: فردگرایی واژه مودبانه‌ای است. راب می‌توانست بگوید من به‌عنوان یک همکار در نوشتن مقاله مشترک آدم سختی هستم. من در شروع کار هیچ چیز را نمی‌فهمم و بعد وقتی که می‌فهمم اصرار دارم که آن را به روش خودم بیان کنم. توجهم به سرعت از مطلبی به مطلب دیگر منحرف می‌شود. یکی از ویژگی‌های فوق‌العاده آمار این است که می‌توانی به تعداد زیادی از شاخه‌ها سر بکشی. برای کسی با دامنه توجه کوتاه مدت، این رشته آیدآل است چون هر وقت از داده‌های بافت‌برداری خسته شدی می‌توانی سراغ داده‌های اخترشناسی بروی. رشته ما خیلی متراکم نیست. قاعدتاً تصویری از ریاضیات باید از یک لکه مرکزی بسیار چگال و تعدادی لکه کوچکتر که کمی از مرکز خارج شده‌اند، تشکیل شده باشد. ولی تصویری از آمار بسیار پراکنده‌تر خواهد بود، با تعداد زیادی فضای خالی در قسمت‌های کشف نشده.

موریس: کار دیگری که بوت‌استرپ انجام داد این بود که به تعداد زیادی از آماردان‌ها بهانه‌ای داد تا بخواهند که برایشان کامپیوتر خریده شود و این یک تحول بود. بیزگرایان برای مدتی عقب ماندند

1) Benjamini - Hochberg

ولی بعد آن‌ها هم با MCMC وارد صحنه شدند و ناگهان همه در آمار صاحب کامپیوتر شدند.

### ۳. بیزگرایی و بیزگرایی تجربی

تیشیرانی: در سال ۱۹۸۱ تو این سؤال را مطرح کردی که چرا همه بیزگرا نیستند؟ بیست و دو سال بعد از آن زمان، فکر می‌کنی نسبت بیزگراها در میان ما افزایش یافته است؟

افرون: بله، تصور می‌کنم علاقه به آمار بیزی، به‌ویژه در انگلستان بیشتر شده است. انجمن سلطنتی آمار، هر شماره به بیز می‌پردازد. این مسأله بی‌دلیل نیست. یکی از دلایلش این است که آمار بیزی اکنون نسبت به بیست سال پیش تغییر کرده است. واقع‌گراتر شده است و همتش را به جای بحث‌های فلسفی دربارهٔ این‌که چرا فراوانی‌گرایی اشتباه است، صرف حل مسائل واقعی می‌کند. قطعاً یک انقلاب محاسباتی در آمار بیزی رخ داده است. روش نمونه‌گیری گیبس یک مثال واقعاً برجسته است. من معتقدم که بیزگرایی تجربی فصل مشترک طبیعی فراوانی‌گرایی و بیزگرایی است، ولی این نظر آنقدر که من امیدوار بودم مورد توجه قرار نگرفته است. روش‌های بیزی از نوع MCMC یک مشکل دارند، این‌که ناچارند از توزیع‌های پیشین ساده‌ای استفاده کنند که با MCMC هماهنگ باشد. انگار که با همه پیشرفت‌های ریاضی و محاسباتی، مردم همچنان مسیری را انتخاب می‌کنند که کمترین مقاومت را برانگیزد. از یک نظر این کار، مسألهٔ اصلی آمار بیزی را که انتخاب توزیع پیشین است مخفی می‌کند. نکتهٔ جذاب بیزگرایی تجربی، اگر درست استفاده شود، دور زدن مسألهٔ انتخاب توزیع پیشین در فضای با بُعد بالای پارامترها است، که اساس اختلاف میان آمار بیزی و آمار فراوانی‌گرا است.

موریس: از نظر تو بیزگرایی تجربی ناکام بوده است؟

افرون: بیزگرایی تجربی فقط در مقایسه با چیزهایی مانند آزمون ویلکاکسون<sup>۱</sup> که میلیون‌ها بار مورد استفاده قرار می‌گیرند، ناکام بوده است. استفاده از ایده‌های بیزگرایی تجربی یا مدل‌های سلسله مراتبی، عمومیت بیشتری پیدا کرده است، ولی این ایده‌ها واقعاً به جامعه کاربردی نفوذ نکرده‌اند. من به فواید ممکن فکر می‌کنم و فایدهٔ استفاده از ابزارهایی مثل آزمون ویلکاکسون در عمل آنقدرها بیشتر از آزمون  $t$  نیست. ولی فواید عملی بیزگرایی تجربی می‌تواند حتی آماردان‌ها را هم شگفت‌زده کند. شما می‌توانید به راحتی ۵۰٪ تا ۷۵٪ ریسک را از بین ببرید. پس چرا از آن‌ها بیش از این استفاده نمی‌شود؟ دلیلش این است که ما به اندازهٔ کافی از بنیادهای نظری این روش‌ها و این‌که چه مواقعی باید از آن‌ها استفاده کرد مطمئن نیستیم. تحلیل واریانس به‌شکل باورنکردنی پر استفاده و برای موقعیت‌های متعددی مناسب است. یکی از علل این امر این است که فیشربه دانشمندان یاد داد که آماردان‌ها به خوبی از عهدهٔ تحلیل واریانس برمی‌آیند، بنابراین محققان در طراحی آزمایش‌ها ما را در نظر می‌گیرند. فکر می‌کنم اگر ما در تحلیل مسائل بیزگرایی تجربی متبحر شویم و به هر دو

1) Wilcoxon test



جنبه نظری و کاربردی آن اطمینان پیدا کنیم، آن وقت آزمایشگران شروع به طراحی آزمایش‌هایی خواهند کرد که از آن نوع ساختارهای موازی که در بیزگرایی تجربی مورد نیاز است استفاده می‌کند. میکروآرایه‌ها مثال خوبی از ساختار موازی مفید هستند.

هلمز: هنگام استفاده از بیزگرایی تجربی، مشکل سازگاری به وجود نمی‌آید؟ چه توجیهی برای ترکیب کردن پارادایم‌ها وجود دارد، که رویکرد بیزی در پیش بگیریم و بعد با داده‌ها مانند یک فراوانی‌گرا برخورد کنیم؟

افرون: سازگاری پاسخی است از بیزگرها در مقابل بهینگی. فراوانی‌گرها از بهینگی صحبت می‌کنند. بیزگرها برای مقابله، به سازگاری متوسل می‌شوند و می‌گویند که فراوانی‌گرایی ناسازگار است، زیرا اطلاعات به دست آمده از موقعیت‌های مختلف را به نحو منطقی ترکیب نمی‌کند و این انتقاد کاملاً درستی است، به ویژه وقتی مجبوریم اطلاعات را با هم تلفیق کنیم. بیزگرایی جذابیت‌های دیگری هم دارد. مثلاً این که بسیار بیشتر از فراوانی‌گرایی نسبت به مدل‌سازی خوشبین است. فراوانی‌گرایی در این مورد، موضعی تدافعی دارد و تلاش می‌کند ادعایی نکند که احتمال اشتباه بودنش زیاد باشد. چیزهای زیادی در آمار بیزی هست که من می‌پسندم. چیزی که نمی‌پسندم انتخاب فی‌البداهه یک توزیع پیشین و ادعای یافتن جواب است. این خیلی خطرناک است، به خصوص در مورد مسائل در ابعاد بالا. نظریه بیزی بسیار جذاب خواهد شد وقتی ایده خوبی داشته باشید که توزیع پیشین دست‌کم خیلی بد نیست. ممکن است در موقعیت پیچیده‌ای قرار بگیرید که فراوانی‌گرایی در آن گم می‌شود. مثلاً تعداد زیادی مقایسه همزمان؛ و آن وقت رویکرد بیزی حرف‌های جالبی برای گفتن خواهد داشت.

تیبشیرانی: من فکر می‌کنم نکته مهم دیگر این است که مردم وقتی به استفاده از یک ابزار گرایش پیدا می‌کنند که به سؤالاتی جواب دهد که قبلاً جوابی برای آن وجود نداشته است. آمار پایدار در دهه شصت خیلی گسترش پیدا کرده بود، ولی ما الان چقدر از آن استفاده می‌کنیم؟ آمار پایدار در موقعیت‌هایی که ما از قبل نتایجی داشتیم نتیجه‌ای با کیفیت بهتر به دست می‌دهد. بوت‌استرپ به مسائلی جواب می‌دهد که قبلاً جوابی برای آن‌ها وجود نداشت. این از آن نوع ابزارهایی است که مردم از آن استفاده خواهند کرد. تحلیل واریانس مثال خوبی است. ابزاری بنیادی که به سؤالاتی پاسخ می‌دهد که از نظر علمی اهمیت دارند.

افرون: نکته مثبت بوت‌استرپ آن است که انعطاف‌پذیر است و استفاده از آن ساده است و هر چه زمان می‌گذرد ساده‌تر می‌شود. استفاده از نظریه‌هایی مثل مینیمم واریانس یک‌نواخت نااریب که شما ناچارید در هر مورد از نو فکر کنید و کلک تازه‌ای پیدا کنید، ساده نیست. اما موردی مثل تخمین بیشترین درستی‌مایی که یک الگوریتم همه‌جا کار می‌کند، فوق‌العاده موفق است. بنابراین شاید آنچه تلاش داشتیم بگویم این بود که لازم است بیزگرایی تجربی خودکار شود.

تیبشیرانی: فکر می‌کنم نکته‌ای که برد سعی دارد بگوید این است که یک روش برای فراگیر شدن باید نیمه خودکار شده باشد. اما اگر برای به کار بردن آن در هر مورد یک تز دکترای آمار لازم باشد،

در آن صورت تعداد ما کافی نیست تا آن را به ابزاری فراگیر تبدیل کنیم.

موریس: خیلی از نرم‌افزارها در حال حاضر روش‌هایی دارند که به خلاصه شدن مدل‌ها کمک می‌کنند. آیا این به نظر تو کافی نیست؟

افرون: وقتی شما از یک تخمین بیزی یا بیزی تجربی استفاده می‌کنید، برخلاف نظریه کلاسیک، مطمئن نیستید که هر  $\theta$  به‌طور ناریب به‌وسیله  $x$  متناظرش تخمین زده شود. با تخمین بیشترین درست‌نمایی، هر پارامتر به روشی کم و بیش ناریب تخمین زده خواهد شد. اگر از یک تخمین بیزی تجربی استفاده کنید همه چیز به سمت مرکز کشیده می‌شود. شما باید دندان روی جگر بگذارید و باور داشته باشید که گرچه در هر کدام از تخمین‌ها به‌تنهایی ممکن است کمی از هدف دور شوید ولی در مجموع اوضاع به مراتب بهتر است. این چیزی است که ما باید به مردم که شامل خودمان هم می‌شوند، بقبولانیم.

موریس: من معتقدم که عملاً همه برآوردها بهتر می‌شوند. نه به تعبیر فراوانی‌گرایانه، اما به این معنی که بر مبنای اطلاعات موجود، هر کدام از آن‌ها بیشتر احتمال دارد که بهتر شده باشند. البته بعد از این که مقادیر واقعی مشخص شد، متوجه خواهیم شد که بعضی از برآوردها بهتر از بقیه‌اند.

افرون: ولی ممکن است بدانی که برای یک پارامتر غالباً بزرگ، مثلاً ۸۰٪ احتمال دارد که تخمینی کمتر از مقدار واقعی به دست آید.

موریس: من اصلاً این طور فکر نمی‌کنم. اگر تو هر دو مرحله مدل را درست انجام داده باشی، با نزدیک کردن مقادیر تخمین‌ها به یکدیگر، وضع بهتری خواهی داشت. نکته اینجاست که باید در مورد مرحله دوم مدل آگاه باشی، که شامل تبادل‌پذیری پارامترها یا شاید چیزهای پیچیده‌تری بشود.

افرون: بنابراین همان‌طور که ما نوشته‌ایم، باید به ارتباط میان پارامترها اعتقاد داشته باشی. مثلاً اگر روی پنی‌سیلین و آمپی‌سیلین و ده نوع آنتی‌بیوتیک دیگر آزمایش می‌کنی و برای هر کدام تخمینی داری، باید باور داشته باشی که همه داده‌ها، و نه فقط داده‌های پنی‌سیلین، می‌توانند بر تخمین پنی‌سیلین تأثیر بگذارند.

موریس: تو باید در شروع، تصمیم‌گیری که آیا مایل هستی تخمین‌ها را با هم ترکیب کنی یا نه. داده‌ها باید تصمیم بگیرند که یک انقباض بزرگ ممکن است یا نه. مثلاً در بررسی داده‌های بیمارستان‌ها، من تخمین بیمارستان‌های ویرجینیا که مقدار زیادی روند مشابه در آن‌ها انجام می‌شود را منقبض خواهم کرد. تقریباً ۱۶۰ تا از این بیمارستان‌ها وجود دارد. من معتقدم اطلاعات یکی از این بیمارستان‌ها به بقیه آن‌ها هم ربط دارد. نمی‌دانم چقدر، ولی داده‌ها کمک خواهند کرد که تصمیم بگیرم.

افرون: اما اگر تو به بیمارستانی که بهترین امتیاز را کسب کرده بگویی که امتیاز آن‌ها را به پایین منقبض کرده‌ای چون فکر می‌کنی بخشی از آن ناشی از شانس بوده است، آن‌ها با تو موافق نخواهند بود.

موریس: آن‌ها خوششان نخواهد آمد. اما در بیشتر موارد حق با من است. به‌علاوه، یاد دادن این ایده‌ها در اولین درس آمار خیلی مشکل است ولی اکثر کسانی که ما با آن‌ها طرفیم بیش از یک درس نگذرانده‌اند.

افرون: خوب، من فکر می‌کنم نکته‌ی اساسی این است که آمار قرن بیستم به ما یاد داد که هر پارامتری را جداگانه بررسی کنیم، تلاش کنیم آن را بدون اربسی برآورد کنیم، یا با یک تخمین نااریب آزمون کنیم. شاید قرن بیست و یکم باید در مسیر برگشت حرکت کند. ما باید بپذیریم که دیگر این پناهگاه را نخواهیم داشت.

تیشیرانی: تو قبلاً به نکته‌ی دیگری هم اشاره کردی، این که دانشمندان از ما آن سؤالاتی را می‌پرسند که ما به آن‌ها یاد داده‌ایم که پرسند. مثلاً آن‌ها درباره‌ی آزمون  $t$  سؤال می‌کنند. چون این چیزی است که فکر می‌کنند آماردانان قادرند انجام دهند. بعضی وقت‌ها فکر می‌کنم این همه چیز است که آن‌ها فکر می‌کنند ما می‌توانیم انجام دهیم. همین طور که آن‌ها درباره‌ی علم آمار چیزهای بیشتری یاد می‌گیرند، اگر ما بتوانیم مسائل موزی را با روش‌های بی‌زنی تجربی به خوبی حل کنیم، آن وقت سؤال‌های بیشتری از این نوع از ما خواهند پرسید.

موریس: پس شاید ما باید درس‌های مقدماتی آمار را به روش متفاوتی ارائه کنیم. برای مثال اگر به کسانی درس می‌دهیم که قرار نیست آماردان شوند، می‌توانیم به آن‌ها یاد بدهیم که آمار چه قابلیت‌هایی دارد. چه نوع مسائلی ممکن است پیش بیاید و چه موقع باید یک آماردان استخدام کرد. آن‌ها هم ممکن است چنین درسی را بیشتر دوست داشته باشند، چون می‌بینند که به چیزهایی که لازم دارند بدانند نزدیک است، نه این که فقط یک ترم شکنجه شوند.

افرون: در حال حاضر درس دادن ما به شدت تاریخی است. ما با روش‌های نظریه‌ی نرمال از اوایل قرن بیستم شروع می‌کنیم و کم‌کم روش‌های پیچیده‌تر پارامتری را بیان می‌کنیم. بعد شاید در فصل سوم به روش‌های ناپارامتری بپردازیم. اگر رشته‌ی آمار در جهت معکوس توسعه یافته بود، اگر کامپیوترها پیش از ریاضیات در دسترس بودند، ما احتمالاً با روش‌های ناپارامتری شروع می‌کردیم که اساساً ساده‌ترند. بعد در اواخر ترم به مطالب خیلی سخت‌تر مثل آزمون  $t$  و نظریه‌ی نرمال می‌رسیدیم.

موریس: کاری که ما الان می‌کنیم مثل این است که به کسانی که می‌خواهند رانندگی یاد بگیرند درسی بدهیم که با نحوه‌ی عملکرد موتورهای مدل  $T$  شروع می‌شود.

#### ۴. فیشر و دیگر بزرگان

تیشیرانی: تو درباره‌ی فیشر زیاد صحبت کرده‌ای و گفته‌ای که او یکی از قهرمانان فکری تو بوده است. چه کسان دیگری در ۵۰ سال گذشته بر تو تأثیر گذاشته‌اند؟

افرون: فیشر قهرمان همه است. ما فوق‌العاده خوش‌شانس بوده‌ایم که ذهنی تا این حد قوی در رشته‌ی خودمان داشته‌ایم. سخت است کسی را قهرمان بخوانیم ولی افرادی در این رشته هستند که آن‌ها را

هم از نظر قدرت ذهنی و هم به جهت کارهایی که انجام داده‌اند، ستایش می‌کنم: نیمن<sup>۱</sup>، هتلینگ<sup>۲</sup> و روبرت میلر. اگر بخواهم ببینم چه کسی بیشترین تأثیر را بر کارهای من گذاشته: چارلز استاین، که او را هم به شدت ستایش می‌کنم و هرب رابینز.

تیبشیرانی: دیوید کاکس<sup>۳</sup> چطور؟

افرون: خیلی زیاد و بدون هیچ جور ارتباط نزدیک شخصی. کاکس تجسم سنت فیشراس است و من گمان نمی‌کنم برای تفکر روشن در استنباط مثال بهتری وجود داشته باشد. سخت است بدانیم این روزها حال و روز فیشری‌ها چگونه است. در حال حاضر، کاکس دیگری در آن سطح دیده نمی‌شود. امیدوارم که این جریان از بین نرود چون سنتی بسیار عالی است که خیلی خوب خود را با جامعه کسانی که با مسائل عملی آمار استنباطی درگیرند، هماهنگ کرده است.

موریس: جالب است. من قبلاً از تو نشنیده بودم که بگویی ما دیگر کسی مانند دیوید کاکس نخواهیم داشت.

افرون: در این مورد مطمئن نیستم. هنوز آنقدر نگذشته که بتوانم با اطمینان بگویم. میراث فیشری‌ها عمدتاً از انگلستان سرچشمه می‌گیرد. این نگاه واقعاً جالبی به آمار است که در حقیقت نه فراوانی‌گرایانه است و نه بیزگرایانه. نوعی روح انعطاف‌پذیری در آن وجود دارد و نیز مقدار زیادی هوشمندی الگوریتمی و من در حال حاضر چنین خرد عملی‌ای را حتی در خاستگاه این سنت، انگلستان، نمی‌یابم.

موریس: ما همه فیشری‌ها را از انگلیس به آمریکا برده‌ایم.

افرون: آمریکا بستر خیلی مناسبی برای تفکر فیشری نبوده است. اینجا بیش از هر چیز، قلمرو تفکر فراوانی‌گرایانه و چیزی که شاید بتوان آن را پسا فراوانی‌گرایانه نامید، بوده است؛ نوعی توافق فلسفی الحادی که در یادگیری ماشینی نمود می‌یابد. مردم به یاد ندارند که توکی<sup>۴</sup> و موستلر<sup>۵</sup> کتابی نوشتند، همان کتاب سبز معروف که در آن حتی از احتمال هم خبری نیست چه رسد به نظریه استنباط.

موریس: به نظر من می‌آید که فیشر، حتی با همه نبوغی که داشت، اگر به مسائل واقعی نمی‌پرداخت، فیشر نمی‌شد. او متخصص بزرگی در علم ژنتیک بود، اما در کشاورزی هم کار کرد و طراحی آزمایش‌ها را توسعه داد. تصور می‌کنم که اکنون افرادی وجود دارند که این کار را بکنند، اما من نگران حذف شدن جنبه نظری هستم.

تیبشیرانی: من نگران حذف شدن مدل‌ها هستم. کاری که توکی در دهه شصت کرد این بود که گفت دیگر احتیاجی به مدل نیست. من فکر می‌کنم او بیش از حد بر تحلیل اکتشافی داده‌ها تأکید کرد. اتفاق مشابهی در یادگیری ماشینی دارد رخ می‌دهد. مردم طوری عمل می‌کنند که انگار ما فقط به الگوریتم‌های سریع دقیق نیاز داریم و احتیاجی به مدل نیست. به اعتقاد من برای فهمیدن این که یک الگوریتم چطور کار می‌کند لازم است بدانیم چه مدلی مناسب آن است. فکر می‌کنم این

1) Neyman 2) Hotelling 3) David Cox 4) Tukey 5) Mosteller

زمینه‌ای بارور است. معتقدم هسته اصلی رشته ما مدل‌سازی است.

## ۵. مدل و محاسبات

موریس: دانش عمیق‌تر ممکن است در خود مدل باشد. حتی اگر مدل اشتباه باشد، ممکن است سال بعد کسی بتواند از آن استفاده کند و چیز بهتری به دست بیاورد.

افرون: مفهوم مدل به شدت به مفهوم بهینگی گره خورده است. در واقع شما تا وقتی مدل نداشته باشید نمی‌توانید در مورد بهینگی صحبت کنید. با وجود این که در حال حاضر بهینگی برای کسانی که در یادگیری ماشینی کار می‌کنند چندان جالب نیست، در بلند مدت اگر آن را به نظریه برنگردانید علمی به وجود نخواهد آمد. من هم امیدوارم که مدل‌ها به صحنه بازگردند. مدل‌ها را به خوبی می‌توان نقد کرد، زیرا مردم بارها و بارها از آن‌ها استفاده می‌کنند.

موریس: راب، می‌بینم که از مدل‌ها دفاع می‌کنی، تصور می‌کنم تو بیش از من ناپارامتری هستی، ولی با وجود این، از به حاشیه رفتن مدل‌ها متأسفی؟

تیبشیرانی: مدل‌ها می‌توانند هیجان‌انگیزتر از آنچه در ابتدا به نظر می‌رسد باشند. مدل لزومی ندارد که فقط ساده خطی باشد. باید همیشه مدلی در ذهن داشته باشی تا بدانی در یک موقعیت مشخص بهترین کاری که باید بکنی چیست. چون برای دانستن این که چه موقع یک روش موفق نیست، باید بدانی بهترین روش کدام است. در این صورت می‌توانی پیش بروی و بگویی روش، در این مورد، کار نخواهد کرد. من معتقدم برای فهمیدن ویژگی‌های عملیاتی، مدل لازم است.

موریس: من فکر می‌کنم مدل غالباً یک ساده‌سازی بیش از اندازه است.

افرون: اصلاً خود احتمال یک ساده‌سازی عظیم است. ما با انبوهی از اتفاقات غیر قابل توصیف و نویزی مواجه‌ایم. احتمال، مفهوم نویز را بی‌اندازه ساده می‌کند و بعد مدل احتمال همه چیز را باز هم ساده‌تر می‌کند. مغزهای کوچک ما در مواجهه با این جهان پیچیده همین که بتوانند محاسبه و پیش‌بینی کنند، کار بزرگی انجام داده‌اند و ما به هر جور کمکی نیازمندیم. این که اکنون کامپیوترهای بزرگی داریم که می‌توانند هر محاسبه‌ای را که بخواهیم انجام دهند، قطعاً کمک بزرگی است. من یک بار در انجمن ریاضی آمریکا سخنرانی کردم. اولین نکته‌ای که توجهم را جلب کرد این بود که همه به شکل باور نکردنی پیر هستند (این مربوط به زمانی می‌شود که من پیر نبودم). با این سؤال شروع کردم که چه اتفاقی برای ریاضیات می‌افتاد اگر کسی یک کامپیوتر بی‌نهایت سریع اختراع می‌کرد؟ آن وقت شما آن را به خانه می‌آوردید و از جعبه خارج می‌کردید و تا ظهر می‌توانستید تکلیف فرضیه ریمان یا حدس گلدباخ را مشخص کنید و پرسیدم: آیا این پایان ریاضیات خواهد بود؟ و بعد وقتی همه خیلی نگران به نظر می‌رسیدند، من جواب سؤال خودم را دادم: نه، این پایان ریاضیات نخواهد بود، زیرا مردم تازه شروع به استفاده از ماشین‌ها برای جواب دادن به سؤالات سخت‌تر می‌کنند. چیزی شبیه این در آمار اتفاق افتاده است. ما می‌توانیم به سؤالات قدیمی که

خیلی سخت به نظر می آمدند جواب دهیم؛ تقریباً همه سوالات. پس یعنی رشته ما به تاریخ پیوسته است؟ نه، ما تازه شروع به پرسیدن سوالات واقعی تر کرده ایم. در حقیقت الان آماردان های بیشتری وجود دارند و آمار شاخه علمی مهم تری شده است.

تیبشیرانی: صحبت از پیری شد؛ تعداد زیادی از آماردانان وقتی به سن ۶۰ یا ۶۵ سالگی می رسند، به فلسفه رو می آورند. سعی می کنند نظریه بزرگ وحدت همه چیز را کشف کنند، ولی تو این کار را نکردی. به نظر می رسد تو هنوز به مسائل کوچکتر ولی واقعاً کاربردی، علاقه مندی. آیا این یک انتخاب آگاهانه بوده است؟

افرون: من معمولاً درباره چنین چیزهایی فکر نمی کنم، ولی وقتی وارد سن ترسناک ۶۰ سالگی شدم، فکر کردم که تا به حال هیچ برنامه ای نداشته ام و فقط روی هر موضوع لذت بخشی که پیش آمده فکر کرده ام؛ با هر همکاری که خوشم آمده یا روی هر مقاله ای که جالب به نظر رسیده، کار کرده ام. من واقعاً باید تمرکز کنم و تلاش کنم کار بزرگی انجام دهم. ولی بعد که در این باره بیشتر فکر کردم، دیدم اصلاً نمی توانم به این پند عمل کنم. نمی شود سر جای خود نشست و یک کار بزرگ انجام داد، یا دست کم من نمی توانم. بنابراین به سراغ کارهای متعدد کوچک برگشتم به این امید که بعضی از آن ها خوب از آب در بیاید. آمار، رشته فوق العاده دست و دل بازی است، لزومی ندارد باهوش ترین آدم دنیا باشی یا صبح و شب کار کنی؛ همه آن چه باید بکنی این است که به ایده ای برسی و آن را رها نکنی. همان طور که گفتم، بیشتر این رشته به شکل متراکم پر نشده است و بنابراین من به کار کردن روی مسائل کوچک ادامه می دهم.

## ۶. آمار و علم

تیبشیرانی: یکی از چالش هایی که من به آن برخورد کرده ام این است که ما رشته عجیبی داریم از این نظر که خیلی از کسانی که رشته شان آمار نیست کار آماری می کنند. ما کار شیمی یا زیست شناسی نمی کنیم. ما داخل آزمایشگاه نمی رویم و لوله های آزمایش را پر نمی کنیم. ولی آمار چیزی است که برای انجام آن فقط یک کامپیوتر شخصی لازم است. همین باعث می شود افراد زیادی تصور کنند که می توانند آن را به خوبی انجام دهند، در حالی که نمی توانند. ما نه تنها باید به خوبی آمار انجام دهیم، بلکه باید به دانشمندان سایر علوم راه درست انجام امور را بیاموزیم.

موریس: بنابراین آمار اگر بخواهد باقی بماند باید اساساً میان رشته ای باشد. من می خواهم مطمئن شوم که ما زمان کافی برای صحبت درباره یک موضوع دیگر که تواز آن تصویری واقعی داری، خواهیم داشت. من می دانم که تو در استنفورد بعضی مسئولیت های اجرایی داشته ای که باعث شده دید جامع تری از نقش آمار پیدا کنی. استنفورد جای واقعاً جالبی است. شما دانشکده ای خیلی قوی و دانشگاهی خیلی قوی دارید با موقعیت های بین رشته ای. چرا ما باید تلاش کنیم که دانشکده های آمار را حفظ کنیم؟ من مطمئنم که آمار به حیات خود ادامه خواهد داد. ما چه کار می توانیم بکنیم که دانشکده هایی سالم و رشته ای قوی داشته باشیم؟

افرون: من برای مدتی رئیس دانشکده بودم، به قول یکی از همکارانم مثل موشی خانگی که تعلیم می‌بیند تا تبدیل به موش صحرایی شود. من رئیس دانشکده علوم بودم. تجربه واقعاً جالبی بود. انجام این کارها ساده نیست، ولی آماردان‌ها برای این مقام مزیتی واقعی دارند چون ما با رشته‌های زیادی سروکار داریم، در حالی که اکثر دانشگاهیان فقط با رشته خودشان در تماس‌اند. آماردان‌ها در مقایسه چیزها خیلی خوب‌اند و این کاری است که یک رئیس به‌وفور انجام می‌دهد. بعضی وقت‌ها نمی‌شود گفت «الف» یا «ب» خوب‌اند یا نه ولی می‌شود گفت «الف» بهتر است یا «ب».

من زمان زیادی صرف صحبت با دانشمندان دیگر کردم. آن‌ها عالی‌اند ولی من به این نتیجه رسیدم که آمار رشته واقعاً خوشبختی است. اولاً ما گروه کوچکی هستیم و فشار زیادی تحمل نمی‌کنیم. ما برای پول در آوردن تحت فشار وحشتناک نیستیم. شیمی‌دان‌ها و زیست‌شناسان زیر فشار شدیدی هستند تا آزمایشگاه‌های بزرگ راه بیاندازند، چون این تنها راه برای انجام کار علمی برای آنان است. این روزها شما می‌توانید با قیمت خیلی کم یک کامپیوتر بخرید و اگر بخواهید سنتی کار کنید حتی لازم نیست کامپیوتر را هم داشته باشید. آماردان‌ها روابط صمیمی‌ای با هم دارند. بعضی از رشته‌ها به شکل مخوفی رقابتی‌اند. چون در آمار جوایز عظیم یا شهرت وجود ندارد، افراد با هم کاملاً خوب‌اند.

ما با هم زیاد مجادله می‌کنیم، ولی اصولاً آماردان‌های دیگر را دوست داریم و کار یکدیگر را، اگر نه در صفحات مجله‌هایمان دست‌کم در قلبمان، ستایش می‌کنیم. من با خوشحالی به آمار برگشتم، با آرزوی این که دانشکده کوچکمان به خوبی به پیش برود. همان‌طور که قبلاً گفتیم، دانشکده‌های آمار تنها جایی در دنیا هستند که در آن‌ها استنباط به شکل جدی مطالعه می‌شود. اگر دانشکده‌های آمار از بین بروند، مردم همچنان قادر خواهند بود کارهایی که ما کرده‌ایم را انجام دهند، ولی در این صورت تا وقتی فیشر دیگری نیاید هیچ ایده جدیدی در استنباط به وجود نخواهد آمد. فیشر آماردان نبود، ولی تا قبل از فیشر، آماردان بودن تقریباً محال بود.

تیبشیرانی: تو در مورد رشته ما خوشبینی؟

افرون: بله، و من از آن دسته آدم‌های خوشحالی نیستم که نسبت به همه چیز خوشبین‌اند. به نظر من می‌رسد که اگر به آمار در قرن بیستم نگاه کنیم، خمی صعودی با شیب یکنواخت می‌بینیم. ممکن است به سادگی تأثیری که داشته‌ایم را دست‌کم بگیریم، اما هیچ رشته‌ای به عنوان روش اصلی انجام علم، بر این همه رشته دیگر تسلط نداشته است. آمار پدیده‌ای قرن بیستمی است. می‌توان گفت تاریخ جدید آمار دقیقاً از ۱۹۰۱ و با پیرسن<sup>۱</sup> و بیومتريکا<sup>۲</sup> آغاز می‌شود. در شروع، آماردان‌های خیلی کمی وجود داشتند، ولی بعد به تدریج رشته‌های بیشتر و بیشتری شروع به استفاده از آمار به عنوان راهی برای تبادل اطلاعات کردند. همان کاری که پزشکی امروز انجام می‌دهد: آیا آزمایش بالینی انجام شده است؟ آزمایش تصادفی بوده است؟ آزمایش کور بوده است؟

1) Pearson      2) Biometrika

سطح معناداری چه بوده است؟ این، در مقایسه با روش قدیمی بررسی موردی، یک گام بزرگ رو به جلو بوده است! روشی که پزشکان پیش از آن برای آزمایش به کار می‌بردند: من مریضی را معاینه کردم و به او نیتروگلیسیرین دادم و حالش خیلی بهتر شد! رشته‌ها یکی بعد از دیگری به روش‌شناسی آماری اعتماد کردند. البته این به‌طور خاص مناسب رشته‌هایی است که در آن یک قسمت کوچک از داده‌ها به خودی خود حاوی اطلاعات قطعی نیست. اگر شما از یک نفر بپرسید که آیا طرفدار دموکرات‌ها است یا جمهوری‌خواهان، چندان مهم نیست. اما اگر از هزار نفر بپرسید، آن وقت یک نظرسنجی مفید خواهید داشت. علوم دقیق بیشترین مقاومت را در برابر آمار کرده‌اند، چون به آن نیاز نداشته‌اند. اطلاعات آن‌ها سخت به دست می‌آید. اندازه‌گیری می‌کنند و به‌طور حتم نظریهٔ اینشتین، انتقال نور را بهتر از نظریهٔ نیوتن پیش‌بینی می‌کند.

موریس: خوب! بله، ما با کاربردها به شکلی متفاوت از ریاضیات برخورد می‌کنیم. ریاضی‌دان‌ها می‌گویند این به درد رشته‌هایی مثل فیزیک می‌خورد. امیدوارم که ما همچنان به این ارتباط‌های متنوع و به کسانی که می‌توانند چنین ارتباط‌هایی برقرار کنند، ارزش بدهیم.

افرون: من بر مبنای تجربهٔ استنفورد فکر می‌کنم یک چیزی اتفاق بیفتد. ما در کنار دانشکدهٔ آمار، دانشکدهٔ آمار زیستی داریم. ممکن است دانشکدهٔ آمار زمین‌شناسی یا آمار اخترشناسی هم به وجود آید. این رشته‌ها شروع به استفادهٔ بیشتر از آمار کرده‌اند. در فیزیک، در هر آزمایش  $10^{26}$  ذره وجود دارد و نیازی به آمار احساس نمی‌شود. ولی وقتی موقعیت‌هایی با  $10$  یا  $100$  ذره پیش می‌آید، ناگهان کارایی استنباط اهمیت پیدا می‌کند. پاییز سال آینده یک کنفرانس فیزیک و آمار در مرکز شتاب‌دهندهٔ خطی استنفورد برگزار می‌شود.

تیشیرانی: یک مثال دیگر، میکروآرایه‌های DNA است. در بسیاری از حوزه‌های ژنتیک نیازی به آماردان نیست تا بگوید که یک اثر عمده وجود دارد، ولی اگر شما به  $6000$  عامل بالقوه مؤثر نگاه کنید، برای جدا کردن علائم واقعی از نویز، نیاز به کمک آماری خواهید داشت.

موریس: تا اینجا چند نکته را خلاصه کنم: من فکر می‌کنم اگر دانشکده‌های آمار نبودند، دنیا به عقب برمی‌گشت. دانشکدهٔ آمار محل برقراری پیوندهای بین رشته‌ای است. به علاوه جای خوشایندی است. فکر می‌کنم افرادی هستند که عاشق چنین کاری هستند و قرار نیست همه این کار را بکنند. دانشکده‌های آمار می‌توانند روند تزییق روش‌های خوب آماری به سایر رشته‌ها را ساده‌تر کنند. ما باید این را جدی بگیریم. باید هیأت علمی و دانشجویانی را که به این کار علاقه‌مندند، جذب کنیم. اگر منزوی شویم و تلاش نکنیم خودمان را به دیگران معرفی کنیم، ارتباطات میان‌رشته‌ای را از دست خواهیم داد.

## ۷. جهت‌های آینده

افرون: در صنعت و خدمات هم همین‌طور ولی به‌طور خاص در دانشگاه، ایده‌ها حرف اول را می‌زنند. آنچه ما باید واقعاً انجام دهیم این است که به تولید ایده‌های خوب ادامه دهیم. سابقهٔ ما



خیلی خوب بوده است. هر چند سال یک بار، ایده واقعاً مفیدی از آمار برآمده است. اگر ما به این روند ادامه دهیم، نگرانی‌ای بابت آینده دانشکده‌های آمار وجود نخواهد داشت.

هلمز: تو رئیس بعدی انجمن آمار آمریکا خواهی بود. جهت ویژه‌ای وجود دارد که آماردان‌ها علاقه‌مند باشند به طور گروهی، به سمت آن حرکت کنند؟

افرون: در انجمن آمار همین سؤال را از من پرسیدند. این سنت دلپسندی است که رئیس باید برای نشست مشترک آمار، موضوعی انتخاب کند. موضوعی که من، پس از مدتی تعمق، انتخاب کردم آمار به‌عنوان نظامی یکپارچه بود. هیچ‌کس نگران فیزیک به‌عنوان نظامی یکپارچه یا اخترشناسی به‌عنوان نظامی یکپارچه نیست، اما این رشته‌ها از مزیت داشتن سنتی هزار ساله و موضوعی کاملاً مشخص برخوردارند. آمار رشته‌ای با یک یا دو قرن سابقه است با موضوع استنباط که حتی در علوم طبیعی رخ نمی‌دهد. در آمار، نیروی گریز از مرکز وحشتناکی وجود دارد، چون ما در جبهه‌های متعددی کار می‌کنیم و تعدادمان آنقدرها زیاد نیست. به راحتی می‌توان تصور کرد که این رشته به ریاضی‌دانان، آماردان‌های شرکت‌های داروسازی، تحلیل‌گران داده‌های بقاء، محققان نمونه‌گیری و ... تجزیه شود. وقتی از من خواسته شد که کاندیدای ریاست انجمن شوم، خوشحال شدم و احساس افتخار کردم، شاید چون آماردان‌های کرانه غربی نقش بزرگی در انجمن آمار آمریکا نداشته‌اند و بیشتر جذب مؤسسه آمار ریاضی می‌شوند. من خوشحالم که ما بیش از یک تشکیلات آماری داریم، اما انجمن آمار آمریکا چتر ما است. من مایلم افراد در دانشکده‌های برکلی، استنفورد، شیکاگو و سیاتل احساس کنند که در همان رشته‌ای هستند که آماردان‌های مرک<sup>۱</sup>، فایزر<sup>۲</sup>، DOE و پرودنتیال<sup>۳</sup> در آن کار می‌کنند. من خوشحالم که آماردان‌ها در زمینه‌های بسیار متنوعی کار می‌کنند. این برای ما سطح تماس وسیع‌تری با دانشمندان دیگر ایجاد می‌کند. آمار سابقه طولانی در به‌دست آوردن ایده‌ها از افرادی خارج از رشته دارد، شاید حتی فیشر چنین مثالی باشد، ویلکاکسون قطعاً هست. رمز ماجرا در این است که در رشته خودمان یک هسته مرکزی قوی حفظ کنیم ولی در عین حال راه را برای مسائل و ایده‌هایی که از بیرون می‌آیند باز نگاه داریم. از مزایای انجمن آمار، بزرگی و سنت آن است. بر دیوار دفتر انجمن دست‌نوشته‌هایی از سال ۱۸۳۹ وجود دارد (به‌نظر می‌رسد آن‌ها در مورد عضوگیری خیلی نگران بوده‌اند). به‌علاوه، انجمن مجله JASA را دارد، مجله‌ای عالی که طیف بسیار وسیعی از افراد را جذب می‌کند و نشست‌های مشترک آماری که تعداد فوق‌العاده زیادی شرکت‌کننده دارد.

## ۸. نصیحت‌ها و دغدغه‌ها

هلمز: به نظر تو بهترین آموزش ممکن برای دانشجویانی که به رشته آمار می‌آیند چیست؟  
افرون: برای موفقیت در آمار چه چیزهایی لازم است؟ باید مقدار مشخصی ریاضیات بدانی و

1) Merck 2) Pfizer 3) Prudential

واقعاً عاشق اعداد باشی، در غیر این صورت از پس حجم عظیم کارهای عددی که باید انجام داد بر نخواهی آمد. ریاضی‌دان‌های زیادی وجود ندارند که عاشق اعداد باشند. آن‌ها از اعداد فرار می‌کنند. اگر به یک مجله ریاضی نگاه کنی، خیلی کم و به ندرت عدد می‌بینی. مقدار کمی علوم تجربی قطعاً مفید است، چون موضوع رشته ما استنباط علمی است. مردم می‌توانند پیش‌زمینه‌های بسیار متنوعی داشته باشند. زمینه سنتی ریاضی به تنهایی آنقدر هم مطلوب نیست. ما زمان زیادی صرف این می‌کنیم تا به دانشجویانی که بیش از حد زمینه ریاضی دارند از نو بیاموزیم که کمتر اصل موضوعی و دقیق باشند و در عوض مسائل را بیشتر از زاویه‌ای که با روح استنباط آماری سازگار باشد ببینند. باید در مورد هر مسأله‌ای میزان درست دقت را پیدا کرد.

تیبشیرانی: ما امروز به دانشجویانی نیاز داریم که توانایی برنامه‌نویسی بهتری داشته باشند.

افرون: قطعاً مفید است، چون یکی از تجهیزات اصلی ما است. شما دانشجوی زیست‌شناسی‌ای که نداند چطور با پیپت کار کند نمی‌خواهید. من خوشحالم که ما الان دانشجویانی داریم که از فیزیک یا زیست‌شناسی آمده‌اند. آن‌ها ثمرات زیادی برای آمار دارند. ممکن است راه‌های جدیدی برای مواجهه با مسائل داشته باشند. آموزش ریاضی برای اهداف عمومی علم عالی نیست. شما باید حسی نسبت به علوم و عالمان داشته باشید. اخترشناسان ستاره‌ها را دارند، زمین‌شناسان صخره‌ها را دارند، و ما علم را داریم. این ماده خامی است که آماردان‌ها با آن کار می‌کنند. شروع کردن فی‌البداهه از تحلیل داده‌ها کار ساده‌ای نیست. شما با حجم انبوهی داده مواجه‌اید. اول از همه باید چکار کرد؟ بعضی اوقات حتی مقدار کمی آمار هم می‌تواند مفید باشد. من این روزها با آمارزیستی‌ها زیاد کار می‌کنم. می‌دانید که آن‌ها برای حمله به مسائل پیچیده روشی دارند که چندان هم ریاضی محض واقعی نیست. آماردان‌های خوب کمک می‌کنند که آن‌ها به روشی منطقی و روشن فکر کنند.

هلمز: گاهی اوقات دانشمندان، ما را متهم می‌کنند که وقتی دنبال جوابی ساده هستند، سعی می‌کنیم به آن‌ها جواب درست بدهیم. آن‌ها انتظار میزان بیشتری تقریب را دارند.

افرون: چیزی که از نظر ما درست است، از نظر آن‌ها گیج‌کننده است. بعضی وقت‌ها حق دارند و واقعاً گیج‌کننده است. به یاد می‌آورم که در یکی از شکست‌های واقعی که به عنوان مشاور آماری تجربه کردم، خانم دانشمندی نتایج یک آزمایش بزرگ دوجمله‌ای با تعداد زیادی عامل را برایم آورده بود. من به دقت همه چیز را با رگرسیون منطقی توضیح دادم. او به هیچ وجه نمی‌توانست logitها را به عنوان جواب قبول کند و در نهایت، نتایج من را دور انداخت و درصدهای ساده را چاپ کرد که احتمالاً برای مخاطب او درست بود. برای او این کار درست بود. من همیشه متأسفم که چرا تلاش بیشتری برای رسیدن به تفاهم انجام ندادم. به سادگی می‌توانستم نتایجم را به زبان درصدها بیان کنم. فکر می‌کنم همین کافی بود. از آن به بعد، همیشه در نوشتن توضیحات برای مراجعین یا همکارانم دقیق بوده‌ام تا از زبانی قابل قبول یا حداقل به شکل قابل قبولی شبیه به زبانی که آن‌ها عادت دارند با آن فکر کنند، استفاده کنم. همیشه در این کار موفق نبوده‌ام، برای این کار گاهی

دانش علمی بیش از سواد من لازم بوده است. این موضوع، یک سؤال جالب پیش می‌آورد: به‌عنوان آماردان همکار، چه اندازه باید از علوم دیگر مطلب بدانیم؟ جواب‌های متنوعی به این سؤال وجود دارد. این جواب که «هر چقدر بیشتر بدانیم بهتر است»، لزوماً درست نیست. چون اگر خیلی زیاد بدانیم بیش از حد به موضوع نزدیک می‌شویم و ممکن است فکر کنیم که متخصص آن علم هستیم. بعضی آدم‌ها می‌توانند در زمان کوتاه ماهیت اصلی یک موضوع را به‌خوبی یاد بگیرند. ولی اگر قرار باشد برای این که به زیست‌شناسان کمک آماری بدهیم خودمان زیست‌شناس شویم، در این صورت آماردانی وجود نخواهد داشت؛ فقط زیست‌شناسانی وجود دارند که با اعداد، بهتر کار می‌کنند. من قویاً احساس می‌کنم که جوهری از استدلال آماری وجود دارد که از میان رشته‌های متعددی می‌گذرد و این همان چیزی است که ما یاد می‌گیریم. البته ما علاقه‌مندیم افرادی، احیاناً خیلی مطلع، هم باشند که متخصص کمک در رشته‌های خاصی باشند. این الگوی دیگری از همکاری علمی است. هلمز: غالباً اگر زبان دانشمندان سایر علوم را ندانی، نمی‌توانی به سوالات آن‌ها پاسخ دهی.

افرون: زبان قطعاً بسیار مهم است؛ دست‌کم در حد دانستن نام چیزهای مهم. واهه پتروسیان<sup>۱</sup> خیلی خوب ایده‌های پیچیده اخترشناسی مربوط به تبدیلات هوشمندانه نسبیتی را برای من توضیح می‌دهد. خوب، شاید من نتوانم واقعاً فیزیک یاد بگیرم اما حداقل او به من شکل یک تابع را نشان می‌دهد و بعد ما می‌توانیم در این باره با هم صحبت کنیم. این روزها موضوع زیست‌شناسی خیلی پیچیده است، به اندازه یک عمر کار برای انجام وجود دارد ولی ما می‌توانیم به اندازه کافی یاد بگیریم تا بتوانیم مفید باشیم. آیا ما دانشجویانمان را درست تربیت می‌کنیم؟ نمی‌دانم. ما به آن‌ها نمی‌گوییم که بروند و یک سال در یک آزمایشگاه زیست‌شناسی یا چیزی شبیه آن کار کنند. شاید باید این کار را بکنیم، ولی من فکر نمی‌کنم این بهترین کار باشد. راه بهتر این است که اینجا آمار یاد بگیرند ولی دست‌کم آگاه باشند که چطور با دانشمندان سایر رشته‌ها ارتباط برقرار کنند. آن وقت اگر از اینجا بیرون بروند و شغلی بگیرند که قرار باشد برای ۲۰ سال آینده با میکروبیولوژیست‌ها کار کنند، خیلی معقول است که بیشتر از من میکروبیولوژی یاد بگیرند. ولی فکر نمی‌کنم وظیفه ما یاد دادن میکروبیولوژی باشد.

تیبشیرانی: شاید استانداردهای ریاضی ما بیش از حد بالا است و به همین دلیل، تعداد زیادی از دانشجویان را که ممکن است آماردان‌های خوبی بشوند اما آمار ریاضی‌شان خوب نیست، از دست می‌دهیم.

افرون: در استنفورد در این مورد نگرانی وجود دارد ولی شک دارم که دانشگاه‌های زیادی این دغدغه را داشته باشند. حقیقت این است که نظریه، به زبان ریاضیات بیان می‌شود. اخیراً روی مقاله‌ای از هرب رابینسون که در ۱۹۵۶ نوشته شده کار می‌کردم و به یاد آورده بودم که قبلاً عرف چقدر متفاوت بوده است. مقاله با طوفانی از تعاریف سیگما میدان، تابع زیان، تابع خطر،

1) Vahe Petrosian

قانون تصمیم و ... شروع می‌شود. بعد یگراست به سراغ چیزهای جالبی می‌رود که مقاله واقعاً درباره آن نوشته شده است. همه سخنرانی‌ها در سکویا هال همین‌طور شروع می‌شدند. رشته ما الان نسبت به سال ۱۹۶۰ کمتر ریاضی است. ولی این چیز وحشتناکی نیست. در دهه ۶۰ ما همچنان سعی می‌کردیم نتایج بیشتری از نظریه استنباط زیبایی که فیشر، نویمان، والد و غول‌های دیگر پایه‌گذاری کرده بودند، استخراج کنیم. بیشتر مسائل، یک یا دو یا چند پارامتر داشتند. بعد از آن دنبال کردن موضوع دیگر واقعاً برای ذهن انسان سخت می‌شد. این‌طور نیست که ما الان در آمار بهتر شده باشیم و در حقیقت پیشرفت کمی در پایه‌های استنباط آماری رخ داده است. همین یک ساعت پیش داشتم به یک مسأله ژنتیکی با ۴۴۴ اثر اصلی نگاه می‌کردم که با استانداردهای موجود خیلی کوچک است. امید چندانی برای یک راه‌حل در قالب نظریه تصمیم به شیوه اصل موضوعی وجود ندارد. حداقل من چنین راه‌حلی ندارم، اما با یک کامپیوتر خوب و تعدادی ابزار مدرن آماری مثل الگوهای تعمیم یافته خطی، اعتبارسنجی متقابل، بوت استرپ، Splus، نمودارهای ساده، هموارکننده‌ها و چیزهای دیگری از این دست، می‌شود به مسأله حمله کرد و حقیقت این است که این تحلیل من را متوجه نقصی در استنباط در کارهای قبلی کرد که روی میکروآرایه‌ها انجام داده بودم. شاید همه این کارهای روش‌شناسانه‌ای که ما انجام داده‌ایم، حرکتی است به سمت دور جدیدی از پیشرفت در نظریه استنباط. می‌توان به سادگی پذیرفت که ۱۰ میلیون برابر شدن توان محاسباتی که ما شاهد آن بوده‌ایم، وسعت و عمق بیشتری به آمار خواهد داد.

---

مترجم: کسری علیشاهی

دانشکده علوم ریاضی - دانشگاه صنعتی شریف

alishahi@sharif.ir