



نقدی بر شاخص ارجاع هِرش: یک مسئله فرمی ترکیبیاتی

برگردان: شیرین حجازیان*

شدن تفاوت h ها و شماره ارجاعاتشان، در بین اشخاص مختلف باشند. در عوض، شخص خواهان یک خط تراز کمی است تا بداند در عمل، معنای «مقایسه‌پذیر»، «بسیار متفاوت» و «مشابه» چیست. اکنون، در حالی که به نظر می‌رسد که این موضوع به آماردان‌ها مربوط است، ولی در اینجا نشان می‌دهیم که چگونه یک کتاب درسی در ترکیبیات می‌تواند رابطه بین h -ایندکس و $(N_{\text{citations}})$ را روشن کند. در اینجا مدل ساده‌ای ارائه می‌دهیم که نگرانی‌هایی را در مورد استفاده نادرست از h -ایندکس برمی‌انگیزد. برای شروع، فهرست تعداد ارجاعات یک پژوهشگر به ازای هر مقاله را به ترتیب نزولی به صورت $(\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N_{\text{papers}}})$ را به عنوان افزایش از $N_{\text{citations}}$ در نظر بگیرید که λ با نمودار یانگ خود متناظر می‌شود. به عنوان مثال

$$\lambda = (5, 3, 1, 0) \leftrightarrow \begin{array}{|c|c|c|c|c|} \hline \bullet & \bullet & \bullet & \bullet & \bullet \\ \hline \bullet & \bullet & \bullet & & \\ \hline \bullet & & & & \\ \hline \end{array}$$

یک متخصص ترکیبیات تأیید می‌کند که h -ایندکس λ ، طول ضلع مربع دارفی است (در شکل با علامت • مشخص شده): این بزرگترین مربع $h \times h$ است که در λ جای گرفته است. این مشاهده ساده تازگی ندارد، و در شیوه‌های محاسبه ارجاعات و همچنین نوشتگان مربوط به ترکیبیات ظاهر می‌شود، برای مثال [۱] را ببینید. از آنجا که نمودار یانگ با اندازه $N_{\text{citations}}$ و با h -ایندکس ماکسیمم تقریباً یک مربع است، از روی شکل ملاحظه می‌کنیم که

$$0 \leq h \leq \lfloor \sqrt{N_{\text{citations}}} \rfloor$$

حال سؤال زیر را در نظر بگیرید:

برای $N_{\text{citations}}$ مفروض بازه برآورد شده برای h چیست؟ در نظر گرفتن فقط $N_{\text{citations}}$ به عنوان تنها داده ورودی برای یافتن یک پاسخ مناسب، به سختی کافی به نظر می‌رسد. دقیقاً به همین دلیل است که ما این سؤال را، در تشابه با مسائل معمول فرمی، یک

این متن ترجمه‌ای از مقاله زیر است:

Alexander Yong, Critique of Hirsch's Citation Index: A Combinatorial Fermi Problem, Notices of the AMS, Volume 61, Number 9, 1040-1050.

۱ پیشگفتار

معیار h -ایندکس توسط فیزیكدانی به نام جی. ای. هرش، در سال ۲۰۰۵ برای اندازه‌گیری کیفی برونداد یک پژوهشگر پیشنهاد گردید. این عدد بزرگترین عدد صحیح n است که فرد دارای n مقاله است که هر کدامشان حداقل دارای n ارجاع است و تعداد ارجاعات به سایر مقالات وی به طورضعیفی کمتر از n است. اگر چه تمرکز اولیه معیار هرش روی فیزیكدان‌ها بود ولی استفاده از این شاخص در حال حاضر مقبولیت عام یافته است. به عنوان مثال پایگاه‌هایی مانند

Google Scholar و Web of Science

در میان دیگر معیارهای مربوط به ارجاعات، مانند تعداد کل ارجاعات، h -ایندکس را نیز در نمایه فرد مشخص می‌کنند.

جذابیت آنچه توسط هرش پدید آمد این است که h -ایندکس وسیله کمکی ساده و مفیدی در کنار شمارش ارجاعات $(N_{\text{citations}})$ است، به دلیل آن که کمیت اخیر می‌تواند به وسیله تعداد کمی مقاله یا کتاب با ارجاعات زیاد، دچار تغییر ناگهانی گردد. به گفته خودش: «بحث من بر سر این است که دو فرد مختلف با h مشابه بر اساس تأثیر علمی کلی‌شان قابل مقیاسه‌اند، حتی اگر تعداد کل مقالات یا تعداد کل ارجاعات بسیار متفاوتی داشته باشند. بالعکس، بین دو نفر (با سن علمی برابر) با تعداد مشابه مقالات و یا تعداد مشابه ارجاعات ولی با مقدار بسیار متفاوت h ، آن که h بیشتری دارد احتمالاً دانشمند مؤثرتری است».

به نظر می‌رسد که کاربران در زمان ارزیابی‌شان، مایل به برجسته

$N_{\text{citations}}$	۵۰	۱۰۰	۲۰۰	۳۰۰	۴۰۰	۵۰۰	۷۵۰	۱۰۰۰	۱۲۵۰
بازه‌های h	[۲, ۵]	[۳, ۷]	[۵, ۹]	[۷, ۱۱]	[۸, ۱۳]	[۹, ۱۴]	[۱۱, ۱۷]	[۱۳, ۲۰]	[۱۵, ۲۲]
$N_{\text{citations}}$	۱۵۰۰	۱۷۵۰	۲۰۰۰	۲۵۰۰	۳۰۰۰	۳۵۰۰	۴۰۰۰	۴۵۰۰	۵۰۰۰
بازه‌های h	[۱۷, ۲۴]	[۱۸, ۲۶]	[۲۰, ۲۸]	[۲۲, ۳۱]	[۲۵, ۳۴]	[۲۷, ۳۶]	[۲۹, ۳۹]	[۳۱, ۴۱]	[۳۴, ۴۳]
$N_{\text{citations}}$	۵۵۰۰	۶۰۰۰	۶۵۰۰	۷۰۰۰	۷۵۰۰	۸۰۰۰	۹۰۰۰	۱۰۰۰۰	
بازه‌های h	[۳۵, ۴۵]	[۳۶, ۴۷]	[۳۷, ۴۹]	[۳۹, ۵۱]	[۴۰, ۵۲]	[۴۲, ۵۴]	[۴۴, ۵۷]	[۴۷, ۶۰]	

جدول ۱.۳: بازه‌های اطمینان برای h -ایندکس

جدول ۱.۳ را با استفاده از اتحاد اوایلر-گاوس برای افزایش:

مسئله فرمی ترکیباتی می‌نامیم، بخش «مسائل فرمی ترکیباتی» را ملاحظه نمایید.

$$\prod_{i=1}^n \frac{1}{1-x^i} = 1 + \sum_{k \geq 1} \frac{x^{k^2}}{\prod_{j=1}^k (1-x^j)^2} \quad (1.3)$$

نوشته‌ایم. اثبات (۱.۳) از طریق مربع‌های دارفی معمولاً به دانشجویان کارشناسی درس ترکیبیات آموزش داده می‌شود. این مطلب در بخش «اتحاد اوایلر-گاوس و کاربرد آن در h -ایندکس» به اختصار توضیح داده شده است. اهداف آموزشی این نوشتار در بخش‌های: «مسائل فرمی ترکیباتی» و «اتحاد اوایلر-گاوس و کاربرد آن در h -ایندکس» گنجانیده شده‌اند.

نتیجه‌مجانبی که مورد استفاده قرار می‌دهیم و منسوب است به C.D. Savage، S. Corteel [۳]، اندازه متوسط مربع دارفی را، وقتی $N_{\text{citations}} \rightarrow \infty$ به دست می‌دهد. از آنجا که فرمول آن‌ها، حتی برای مقادیر $N_{\text{citations}}$ کم تعداد، هم راستا با محاسبات ما است نتیجه آن‌ها را به عنوان قاعده تقریب h -ایندکس بیان می‌کنیم:

$$h = \frac{\sqrt{6} \log 2}{\pi} \sqrt{N_{\text{citations}}} \approx 0.54 \sqrt{N_{\text{citations}}}$$

تمرکز این نوشتار بر ریاضیدانان است. برای اکثریت بزرگی از کسانی که مورد بررسی قرار گرفتند، h -ایندکس واقعی به دست آمده با استفاده از MathScinet و Google Scholar، در بازه‌های اطمینان قانع کننده‌ای قرار می‌گیرند. به علاوه ما دریافتیم که قاعده تقریب به اندازه قابل قبولی برای ریاضیات محض دقیق است. به عنوان مثال، جدول ۲.۳ این مطلب را برای برندگان مدال فیلدز بعد از سال ۱۹۹۸ نشان می‌دهد^۱

در [۶] نشان داده شده است که h -ایندکس برای برنده جایزه نوبل مقداری قابل پیش‌بینی دارد ولی از نظر ما رابطه h -ایندکس با برندگان مدال فیلدز نامشخص است. چرا که h -ایندکس تعدادی MathSciNet از ۲۰۰۰ کامل نیستند. ارجاعات Google Scholar و Thompson and Reuter's Web of Science نیز خطا دارند. ما برآن شدیم که MathSciNet کامل‌ترین گزینه برای تحلیل ارجاعات ریاضی‌دانان می‌باشد. برای برندگان اخیر مدال فیلدز، تأثیر داده‌های گم شده کاهش داده شده است.

Critique of Hirsch's Citation Index: A Combinatorial Fermi Problem

Alexander Yong

In 2005, physicist J. E. Hirsch [H05] proposed the h -index to measure the quality of a researcher's output. This metric is the largest integer n such that the person has n papers with at least n citations each, and all other papers have weakly less than n citations. Although the original focus of Hirsch's index was on physics, the h -index is now widely popular. For example, Google Scholar and the Web of Science highlight the h -index, among other metrics such as total citation count, in their profile summaries. An enticing point made by Hirsch is that the h -index is an easy and useful supplement to a citation count ($N_{\text{citations}}$), since the latter metric may be skewed by a small number of highly cited papers or textbooks. In his words,

"I argue that two individuals with similar h 's are comparable in terms of their overall scientific impact, even if their total number of papers or their total number of citations is very different. Conversely, comparing two individuals (of the same scientific age) with a similar number of total papers or of total citation count and very different h values, the one with the higher h is likely to be the more accomplished scientist."

It seems to us that users might tend to eyeball differences of h 's and citation counts among individuals during their assessments. Instead, one Alexander Yong is associate professor of mathematics at the University of Illinois at Urbana-Champaign. His email address is ayong@uiuc.edu.
DOI: <http://dx.doi.org/10.1090/noti1164>

desires a quantitative baseline for what "comparable," "very different," and "similar" actually mean. Now, while this would appear to be a matter for statisticians, we show how textbook combinatorics sheds some light on the relationship between the h -index and $N_{\text{citations}}$. We present a simple model that raises specific concerns about potential misuses of the h -index.

To begin, think of the list of a researcher's citations per paper in decreasing order $\lambda = (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{N_{\text{citations}}})$ as a partition of size $N_{\text{citations}}$. Graphically, λ is identified with its Young diagram.

For example, $\lambda = (5, 3, 1, 0) \rightarrow$

A combinatorialist will recognize that the h -index of λ is the side-length of the Durfee square (marked with * above): this is the largest $h \times h$ square that fits in λ . This simple observation is nothing new, and appears in both the biometric and combinatorial literature; see, e.g., [Anilakis99, FlS09]. In particular, since the Young diagram of size $N_{\text{citations}}$ with maximum h -index is roughly a square, we see graphically that $0 \leq h \leq \sqrt{N_{\text{citations}}}$.

Next, consider the following question: Given $N_{\text{citations}}$, what is the estimated range of h ? Taking only $N_{\text{citations}}$ as input hardly seems like sufficient information to obtain a meaningful answer. It is exactly for this reason that we call the question a combinatorial Fermi problem, by analogy with usual Fermi problems; see the section "Combinatorial Fermi Problems."

1040

NOTICES OF THE AMS

VOLUME 61, NUMBER 9

از آنجا که هیچ اطلاعات پیشینی‌ای نداریم، نمایه ارجاعات هر فرد را به حالت غیر اریب در نظر می‌گیریم. یعنی هر افزایش $N_{\text{citations}}$ با احتمال برابر در نظر گرفته می‌شود. در واقع، نظریه زیبایی در مورد رفتارهای مجانبی این افزایش‌های تصادفی یکنواخت وجود دارد که توسط A. Vershik و همکارانش بسط و توسعه یافته است، به عنوان مثال [۸] را ملاحظه کنید. ولی علاقه ما، عملاً، به مقادیر پایین (به دردتخور) از $N_{\text{citations}}$ است، جایی که تمامی نتایج مجانبی به‌واقع چندان مرتبط نیستند. در عوض، از سری‌های مولد و رایانه‌های رومیزی مدرن برای محاسبه احتمال آن که یک λ تصادفی، مربع دارفی از اندازه h داشته باشد، استفاده می‌کنیم. به طور مشخص‌تر،^۱ ارجاعات قبل از ۲۰۰۰ MathSciNet کامل نیستند. ارجاعات Google Scholar و Thompson and Reuter's Web of Science نیز خطا دارند. ما برآن شدیم که MathSciNet کامل‌ترین گزینه برای تحلیل ارجاعات ریاضی‌دانان می‌باشد. برای برندگان اخیر مدال فیلدز، تأثیر داده‌های گم شده کاهش داده شده است.

بازه اطمینان	قاعده تقریب	h	$N_{\text{citations}}$	سال دریافت جایزه	برنده مدال
[۱۳, ۲۰]	۱۷/۲	۱۵	۱۰۱۲	۱۹۹۸	T.Gowers
[۱۴, ۲۱]	۱۷/۶	۱۴	۱۰۶۲	۱۹۹۸	R.Borcherds
[۱۸, ۲۶]	۲۲/۵	۲۵	۱۷۳۸	۱۹۹۸	C. McMullen
[۲۲, ۳۲]	۲۷/۶	۲۳	۲۶۰۹	۱۹۹۸	M.Kontsevich
[۴, ۸]	۶/۲	۵	۱۳۳	۲۰۰۲	L.Lafforgue
[۱۶, ۲۳]	۲۰/۰	۲۰	۱۳۸۲	۲۰۰۲	V. Voevodsky
[۷, ۱۲]	۱۰/۰	۸	۳۶۲	۲۰۰۶	G. Perelman
[۱۴, ۲۱]	۱۸/۲	۱۹	۱۱۳۰	۲۰۰۶	W. Werner
[۱۸, ۲۵]	۲۲/۱	۲۴	۱۶۷۷	۲۰۰۶	A. Okounkov
[۳۸, ۵۱]	۴۴/۳	۴۰	۶۷۳۰	۲۰۰۶	T. Tao
[۵, ۱۰]	۸/۲	۹	۲۲۸	۲۰۱۰	C. Ngô
[۹, ۱۴]	۱۲/۰	۱۲	۴۹۰	۲۰۱۰	E. Lindenstrauss
[۹, ۱۵]	۱۲/۳	۱۲	۵۲۱	۲۰۱۰	S. Smirnov
[۲۴, ۳۳]	۲۹/۲	۲۵	۲۹۳۱	۲۰۱۰	C. Villani

جدول ۲.۳: برندگان مدال فیلدز ۱۹۹۸-۲۰۱۰

از دارندگان مدال فیلدز مندرج در جدول ۲.۳ برابر یا حتی مقداری کمتر از رقبایشان در سن آکادمیک مشابه، یا آنهایی که تعداد ارجاعات مشابه دارند، می باشد. شاید این منعکس کننده تفاوت فرهنگی میان جامعه ریاضیات با دیگر جوامع علمی باشد.

در بخش «مقایسه های بیشتر بر اساس داده های تجربی» ریاضیدانان آکادمی ملی علوم (آمریکا) را مورد تجزیه و تحلیل قرار می دهیم و نشان می دهیم که همبستگی بین h -ایندکس واقعی و به دست آمده از روش تقریبی، $R = 0.94$ است که پس از حذف ارجاعات کتابها $R = 0.95$ خواهد بود. همچنین برندگان جایزه آبل و دانشیاران سه مؤسسه پژوهشی را مورد بحث قرار خواهیم داد. سرانجام از خوانندگان می خواهیم که خودشان این تقریبها را مجدداً بررسی نمایند.

در اینجا سه پیامد یا کاربرد احتمالی تحلیل های خود را به بحث می گذاریم.

۲ قاعده تقریب و ارجاعات پر تعداد

روش ما از لحاظ نظری رفتار h -ایندکس را برای پژوهشگران با تعداد پرشمار ارجاعات پیش بینی می کند. گستره ای که این پیش بینی ها درست اند، اطلاعات مفیدی به دست می دهد. این امر نه تنها برای اشخاص مستقل، بلکه برای کل یک حوزه کاری درست است.

هرش، در واقع، یک ضریب تناسب a را به صورت $N_{\text{citations}} = ah^2$ تعریف کرد و متذکر شد «من به طور تجربی دریافته ام که a بین ۳ و ۵ تغییر می کند.» این بدان معناست که h بین $0.45 \approx \sqrt{1/5}$ و $0.58 \approx \sqrt{1/3}$ برابر $\sqrt{N_{\text{citations}}}$ قرار می گیرد.

مقایسه h ها وقتی $N_{\text{citations}}$ ها بسیار متفاوتند.

می دانیم معمولاً h -ایندکس، با افزایش $N_{\text{citations}}$ رشد می کند اما چه وقت تفاوت زیاد تعداد ارجاعات، مقایسه h ها را فاقد اطلاعات مفید می سازد؟

برای مثال، h -ایندکس تائو و آکونکوف به ترتیب برابر

$$h_{\text{Okounkov}} = 24(1677 \text{ citations})$$

می‌کنیم. اما $h_{Stanley}=35$ ، یعنی ۲۰ درصد خطا. حال، ۳۳۶۲ ارجاع به او مربوط به کتاب‌ها است. با کم کردن این تعداد h -ایندکس او ۳۰/۳ برآورد می‌شود، حال آن که h -ایندکس تصحیح شده وی برابر ۳۲ است، فقط ۵ درصد خطا. این پدیده نامعمول نیست؛ جدول ۳.۳ را ملاحظه کنید.

به عنوان مثالی دیگر نمایه T.Tao را در Google Scholar در نظر بگیرید. از آنجا که او ۳۰۵۳ ارجاع دارد، قاعده تقریب h -ایندکس او را ۹۳/۶ پیش‌بینی می‌کند که بسیار دور از h -ایندکس واقعی او یعنی ۶۵ می‌باشد. حال، پنج مورد با بیشترین ارجاعات او (مقاله مشترک با E. Candes روی سنجش تحت فشار) را به کار می‌گیریم. با حذف مقالات در این زمینه، تعداد ۱۳۹۴۲ ارجاع باقی می‌ماند. تقریب جدید برای او برابر ۶۳/۷ است و h -ایندکس تصحیح شده وی ۶۱ می‌باشد. در بسیاری از مواردی که مشاهده کرده‌ایم، با کنار گذاشتن مواردی از نمایه پژوهشگر که موجب انحراف می‌شود، باقیمانده نمایه با قاعده تقریب همخوانی دارد.

۴ نتیجه‌گیری و خلاصه

روشن است که کیفیت یک پژوهشگر را، خواه برندگان مدال فیلدز، دریافت کنندگان جایزه آبل، کاندیدای ارتقاء شغلی یا نامزدهای دریافت پژوهانه باشند، نمی‌توان فقط با اعداد اندازه گرفت. اما، در واقعیت، h -ایندکس به طور رسمی یا غیر رسمی برای مقایسه به کار می‌رود. این نوشتار در نظر دارد تا چارچوبی نظری و قابل آزمودن برای به‌طور کمی فهمیدن محدودیت‌های چنین ارزیابی‌هایی فراهم آورد. برای ریاضیدان‌ها، دقت قاعده تقریب حکایت از این دارد که تفاوت h -ایندکس بین دو ریاضیدان قویاً تحت تأثیر شماره ارجاعات آنهاست.

در حالی که بحث در مورد ریاضیدان‌های معروف و آمارهای مربوط به آن‌ها موضوع گفتگوهای تفریحی در کافی شاپ‌هاست، یک مورد جدی که h -ایندکس در جلسات دانشگاهی مطرح می‌شود مربوط به ریاضیدانان نسبتاً جوان است. پژوهشگر A را با ۱۰۰ ارجاع و h -ایندکس ۶ و پژوهشگر B را با ۵۰ ارجاع و h -ایندکس ۴ در نظر بگیرید. چنین اعدادی برای استادیاران ریاضی که در حال گرفتن شغل دائم هستند غیرطبیعی نیست. مدل ما پیش‌بینی می‌کند که h_A اندکی بیش از h_B است. ولی آیا واقعاً می‌توان گفت که چه کسری از $h_A - h_B$ نشانه‌ای از کیفیت است؟

وقتی A و B زمینه کاری متفاوتی دارند مسئله فراگیرتر می‌شود. انتظار می‌رود که ارجاعات به کارهای اساسی در ریاضیات کاربردی

می‌توان با تلاش برای درک تشابه بین کران بالای تجربی هرش و قاعده تقریب آغاز کرد. حدسی از کن فیلد E.R. Canfield (مکاتبه خصوصی، بخش «اتحاد اوپلر-گاوس...» را ملاحظه نمایید) بیان می‌کند که تمرکز در اطراف مربع دارفی متوسط است. بنابراین، از لحاظ نظری، انتظار می‌رود که قاعده تقریب برای مقادیر بزرگ $N_{citations}$ تقریباً درست باشد.

دریغا که از نظر تجربی، این مطلب حتی برای ریاضیدانان محض، درست نیست. اما نکته مرتبطی مشاهده می‌کنیم: مقدار $\sqrt{N_{citations}}/54$ برای تقریباً همه ریاضیدان‌ها، فیزیکدان‌ها، آماردان‌ها و متخصصین علوم کامپیوتر با ارجاعات پر تعداد ($N_{citations} > 10000$)، که در نظر گرفته‌ایم، بیش از مقدار واقعی h است. در موارد نادری این مطلب روی نمی‌دهد و درصد اندکی (کمتر از ۵٪) از این تقریب درست نیست. از طرف دیگر، تغییر در جهت دیگر اغلب بزرگ است (۵۰٪ یا بیشتر، در حوزه‌هایی از مهندسی یا زیست‌شناسی چندان غیرمعمول نیست).

در بین برندگان جایزه آبل تقریباً برابری اتفاق می‌افتد. همچنین تمام فیزیکدان‌های برجسته که در [۶] ذکر شده‌اند را در نظر گرفته‌ایم (به جز Cohen و Anderson، به دلیل شباهت اسمی در Web of Science). حدس همواره یک کران بالاست (به طور متوسط ۱۴-۲۰ درصد، خیلی بالا است). تساوی توسط D.J. Scalapino (۲۵۸۸۱) ارجاع؛ (۱) C.Vafa (۲۲۹۰۲) ارجاع؛ (۰/۹۹) و J.N.Bachall (۲۷۶۳۵) ارجاع؛ (۰/۹۸) کسب شده است. در اینجا از نسبت $\frac{true\ h}{estimated\ h}$ استفاده کرده‌ایم.

یک دلیل این که افراد با تعداد ارجاع بالا، h -ایندکس کمتری از h -ایندکس مورد انتظار دارند این است که این اشخاص دارای کتاب‌های درسی با تعداد ارجاع بالا هستند. همچنین، دانشمندان معروف اغلب به سوی «اثر متیو» کشیده می‌شوند (مثلاً، ارجاعات بدون دلیل به مقالات یا کتاب‌های بسیار معروف آن‌ها).

۳ h -ایندکس‌های نامتعارف

به طور کلی برآوردهای ما، روشی برای مشخص کردن یک h -ایندکس نامتعارف برای پژوهشگران فعال، یعنی آنهایی که خارج بازه اطمینان، یا آنهایی که قاعده تقریب برای آنها مخصوصاً غیردقیق است، به دست می‌دهد.

برای این که ببینیم ارجاع به کتاب‌ها چه تأثیری روی تقریب‌های ما می‌گذارد، متخصص ترکیبیات R.P. Stanley را در نظر می‌گیریم. از آنجا که استنلی دارای ۶۵۱۰ ارجاع است، ایندکس او را ۴۳/۶ برآورد

نام دریافت کننده	سال دریافت جایزه	$N_{\text{citations}}$	h	قاعده تقریب	بازه تقریب
J.P. Serre	۲۰۰۳	۱۰۱۱۹	۵۳	۵۴/۳	[۴۷, ۶۰]
I. Singer	۲۰۰۴	۲۹۸۲	۲۸	۲۹/۵	[۲۴, ۳۴]
M. Atiyah	۲۰۰۴	۶۵۶۴	۴۰	۴۳/۷	[۳۷, ۴۹]
P. Lax	۲۰۰۵	۴۶۰۱	۳۰	۳۶/۳	[۳۱, ۴۲]
L. Carleson	۲۰۰۶	۱۹۸۰	۱۸	۲۴/۰	[۱۹, ۲۸]
S.R.S. Varadham	۲۰۰۷	۲۸۹۴	۲۸	۲۹/۰	[۲۴, ۳۳]
J. Thompson	۲۰۰۸	۷۸۹	۱۴	۱۵/۲	[۱۱, ۱۸]
J. Tits	۲۰۰۸	۳۴۶۳	۲۸	۳۱/۸	[۲۷, ۳۶]
M. Gromov	۲۰۰۹	۷۶۷۱	۴۱	۴۷/۳	[۴۰, ۵۴]
J. Tate	۲۰۱۰	۲۹۷۹	۳۰	۲۹/۵	[۲۴, ۳۴]
J. Milnor	۲۰۱۱	۷۸۵۶	۴۸	۴۷/۹	[۴۱, ۵۴]
E. Szemerédi	۲۰۱۲	۲۵۳۶	۲۶	۲۷/۲	[۲۲, ۳۱]
P. Deligne	۲۰۱۳	۶۵۷۶	۳۶	۴۳/۸	[۳۷, ۵۰]

جدول ۳.۳: دریافت‌کنندگان جایزه آبل

در شامپین کانتی که حدود ۲۰۰۰۰۰ نفر جمعیت دارد ۱۰ رستوران مک دونالد وجود دارد. فرض کنید تعداد این رستوران‌ها متناسب با جمعیت باشد. چون جمعیت ایالات متحده ۳۰۰ میلیون نفر است، یک محاسبه سر انگشتی تعداد رستوران‌های مک دونالد را ۱۵۰۰۰ برآورد می‌کند. جواب واقعی در سال ۲۰۱۲ برابر ۱۴۱۵۷ است.

استفاده از یک فرض ساده شده، مانند آنچه در بالا فرض کردیم، یک ویژگی یک مسئله فرمی است. واضح است که فرض یکنواخت بودن در بالا واقعاً درست نیست. اما، زمانی که یافتن پاسخ‌های دقیق‌تر یا خیلی زمان‌بر هستند و یا غیر قابل دسترس هستند، تمرکز کار بر تقریب‌های خوب و سریع است. این تقریب در ادامه می‌تواند به عنوان راهنمایی برای کار روی یافتن جواب‌های بهتر و دقیق‌تر مورد بهره‌برداری قرار گیرد.

اگر چه تقریب در اینجا به مقدار واقعی نزدیک است، ولی وقتی تقریب خوب نباشد حتی نتیجه جالب‌تر خواهد شد، زیرا به تشخیص یک فرض نادرست کمک می‌کند. برای مثال، تحلیل مشابهی تعداد فروشگاه‌های هُوَل فود در ایالات متحده را برابر ۰ برآورد می‌کند و این روشن می‌کند که حضور این شرکت متناسب با جمعیت نمی‌باشد.

مسائل فرمی/محاسبات استاندارد سر انگشتی، قسمتی از آموزش‌های فیزیک یا مهندسی است. این مسائل در ساختن مدل‌های ریاضی از ارزش نظری برخوردارند و در حوزه‌هایی مانند مشاوره مدیریت ارزشی «واقع‌گرایانه» دارند. اما شاید به دلیل آن که از ذات نادقیقی برخوردارند، به طور مشخص، قسمتی از ریاضیات محض محسوب نمی‌شوند. این امر به ویژه در مورد ترکیبیات شمارشی درست است، هر چند که هدف این حوزه شمارش

بسیار بیش از ریاضیات محض باشد. در زمینه‌های تجربی، مقالات، نویسندگان متعددی دارند. از آنجایی که h -ایندکس ترتیب نویسندگان را در نظر نمی‌گیرد، به نظر می‌رسد که این نکته در چنین حوزه‌هایی تقریب ما را تحت تأثیر قرار دهد.

ریاضیدانان محض، به طور قابل ملاحظه‌ای تعداد کمتری نویسنده همکار، مقاله و ارجاع دارند. این امری غیرمعمول نیست، برای مثال در اغلب موارد تعداد ارجاعات به حل مسئله‌ای باز که مدت زیادی حل نشده مانده بود، نسبتاً کم است. بنابراین یک مدل توصیفی برای ریاضیدانان محض، به دلایل بنیادین برای دیگر حوزه‌ها ناکارآمد خواهد بود.

اگر این نکته درست باشد، آیا هنوز هم می‌توان h -ایندکس را به طور فراگیر مورد استفاده قرار داد؟ این مسئله از لحاظ نظری پرسیدن این سؤال را توجیه می‌کند: آیا امکان دارد، همانگونه که در [۷] پیشنهاد شده است، حتی پس از بازنگری در معیار، مقایسه ساده‌ای بین حوزه‌های علمی مختلف انجام داد؟

۵ مسائل فرمی ترکیبیاتی

۱-۵ مسائل فرمی معمولی

مسائل فرمی به افتخار انریکو فرمی، که توانایی او در به دست آوردن جواب‌های کمی تقریبی خوب با استفاده از داده‌های اندک در دسترس، زیانزد است، نام‌گذاری شده‌اند. برای روشن شدن موضوع از مثال زیر استفاده می‌کنیم [۴]:

چه تعداد رستوران مک دونالد در ایالات متحده فعالند؟

می‌بریم. اگر Par مجموعه همه افزاها باشد و $\sigma: \text{Par} \rightarrow \mathbb{Z}_{\geq 0}$ تابعی باشد که اندازه افزا را به دست می‌دهد، آنگاه سری مولد برای (Par, σ) عبارت است از

$$P(x) = \prod_{i=1}^{\infty} \frac{1}{1-x^i}.$$

یعنی، $\# \text{Par}(n) = [x^n]P(x)$. کتاب درسی [۲] را به عنوان مرجعی برای این کار، ملاحظه کنید.

اتحاد اوایلر-گوس (۱.۳) را از مقدمه به یاد آورید. برهان ترکیبیاتی شناخته شده آن به این ترتیب است که هر نمودار یانگ λ به طور دوسویی به یک سه تایی $(D_\lambda, R_\lambda, B_\lambda)$ تجزیه می‌شود که در آن D_λ یک مربع $k \times k$ ، R_λ یک نمودار یانگ با حداکثر k سطر و B_λ یک افزا با حداکثر k ستون است. یعنی، D_λ مربع دارفی، R_λ شکل سمت راست مربع و B_λ شکل زیر آن است. برای مثال:

$$\lambda = \begin{array}{|c|c|c|c|} \hline \cdot & \cdot & & \\ \hline \cdot & & & \\ \hline & & & \\ \hline \end{array} \mapsto \left(\begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array}, \begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array}, \begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} \right) = (D_\lambda, R_\lambda, B_\lambda).$$

سری مولد برای افزای با حداکثر k ستون برابر است با $\prod_{j=1}^k \frac{1}{1-x^j}$ ، چون مزدوج («ترانهاده») یک شکل با حداکثر k سطر، شکلی با حداکثر k ستون است، نتیجه می‌گیریم که سری مولد شکلی با حداکثر k سطر نیز به صورت $\prod_{j=1}^k \frac{1}{1-x^j}$ می‌باشد.

از این بحث نتیجه می‌گیریم که سری مولد یک نمودار یانگ با مربع دارفی از اندازه k عبارت است از

$$x^{k^2} \prod_{j=1}^k (1-x^j)^{-2}.$$

با محاسبه برای h و $N_{\text{citations}}$ ثابت داریم:

$$\begin{aligned} \text{Prob}(\lambda: |\lambda| = N_{\text{citations}}, \text{Durfee square of size } k) \\ = \frac{[x^{N_{\text{citations}}}] x^{k^2} \prod_{j=1}^k (1-x^j)^{-2}}{\# \text{Par}(N_{\text{citations}})} \end{aligned}$$

کتاب‌های درسی معمولاً بحث را با به دست آوردن (۱.۳) به پایان می‌رسانند. در کلاس درس، استفاده از کامپیوتر برای بسط تیلور $\sum_{k=a}^b x^{k^2} \prod_{j=1}^k (1-x^j)^{-2}$ و مقایسه ضرائب با اعداد افزا از قبل معلوم، به مدرس این امکان را می‌دهد که اتحاد را به طرزی ملموس برای دانشجویان تشریح کند. تغییرات a و b حوزه تغییرات اندازه

اشیائی است که، براساس تجربه نویسنده این نوشتار، دانشجویان آرزو می‌کنند دارای کاربردی غیرنظری باشند.

۲-۵ یک مشابه ترکیبیاتی

به قیاس، یک مسئله فرمی ترکیبیاتی را تعریف می‌کنیم: یک $\varepsilon > 0$ ثابت را در نظر بگیرید. فرض کنید S یک مجموعه متناهی از اشیای ترکیبیاتی و $\omega: S \rightarrow \mathbb{Z}_{\geq 0}$ یک آماره روی S باشد. در این صورت، مقدار ω را روی هر عضو بازه اطمینان $[a, b]$ ، که احتمال یکنواخت انتخاب عضوی از S خارج از این بازه، کمتر از ε باشد، برآورد کنید.

بنا به تعریف، سری مولد برای مسئله ترکیبیاتی (S, ω) به وسیله $G_{(s, \omega)}(x) = \sum_{s \in S} x^{\omega(s)}$ داده می‌شود. به ازای هر k ، $G_{(s, \omega)}(x) = [x^k] G_{(s, \omega)}(x)$ ؛ یعنی ضریب x^k در $G_{(s, \omega)}(x)$. در کتاب‌های درسی معمولاً کار با استخراج ضریب از طریق به کارگیری فرمول‌های معتبر در حلقه سری‌های توانی صورتی انجام می‌پذیرد. ولی آنچه اغلب در کلاس مورد تأکید قرار نمی‌گیرد این است که این ضریب و همچنین $\#S$ می‌تواند توسط یک سیستم جبر کامپیوتری که برای تعیین سریع بازه $[a, b]$ طراحی گردیده، به سرعت استخراج شود. چون کامپیوتر کار را انجام می‌دهد، این کار را مشابه محاسبه سرانگشتی قلمداد می‌کنیم.

برای مقادیر «معقول» ε (مانند $\varepsilon = 1/2$)، اغلب بازه $[a, b]$ صعب‌الحصول است. در چنین حالت‌هایی، ممکن است قضیه‌ای در مورد تمرکز مجانبی در نزدیکی یک شیء «شاخص» وجود داشته باشد. اما، حتی اگر چنین قضایی وجود داشته باشند، مسئله حالت متناهی را حل نمی‌کنند.

استفاده از توزیع یکنواخت، روشی سریع برای به دست آوردن تقریب‌هایی است که قابل مقایسه با داده‌های تجربی هستند. این کار در نهایت، کاربر را به در نظر گرفتن سایر توزیع‌های احتمالی و تحلیل‌های آماری قانع کننده‌تر دعوت می‌کند (دقیقاً مانند مثال مک دونالد)، مثلاً استفاده از روش‌های زنجیره مارکوف و مونت کارلو.

در اینجا مسئله فرمی ترکیبیاتی دیگری را ذکر می‌کنیم:

تعداد خانواده‌های زبان‌های بومی در قاره آمریکا [۹].

یعنی برای این مسئله، اساساً راهی برای یافتن پاسخ درست، با درجه اطمینان بالا وجود ندارد.

۶ اتحاد اوایلر-گوس و کاربرد آن در h -ایندکس

در اینجا چشم‌انداز بخش ۲ را برای مسئله h -ایندکس، که در آن $S = \text{Par}(n)$ و $\omega: S \rightarrow \mathbb{Z}_{\geq 0}$ اندازه ضلع مربع دارفی افزا است، به کار

بازه تقریب	قاعده تقریب	h	$N_{\text{citations}}$
[۱, ۴]	۳/۰	۳	۳۰
[۲, ۴]	۳/۲	۴	۳۵
[۲, ۵]	۴/۰	۴	۵۶
[۲, ۵]	۴/۰	۵	۵۶
[۳, ۵]	۴/۳	۵	۶۳
[۳, ۵]	۴/۳	۶	۶۳
[۳, ۶]	۴/۸	۳	۷۸
[۳, ۶]	۴/۹	۵	۸۴
[۳, ۶]	۵/۱	۷	۸۸
[۴, ۷]	۶/۰	۸	۱۲۲
[۴, ۷]	۶/۱	۷	۱۲۶
[۴, ۸]	۶/۲	۶	۱۳۳
[۴, ۸]	۶/۲	۷	۱۳۳
[۴, ۸]	۶/۶	۸	۱۵۰
[۵, ۸]	۶/۹	۷	۱۶۳
[۵, ۱۰]	۸/۱	۱۰	۲۲۸

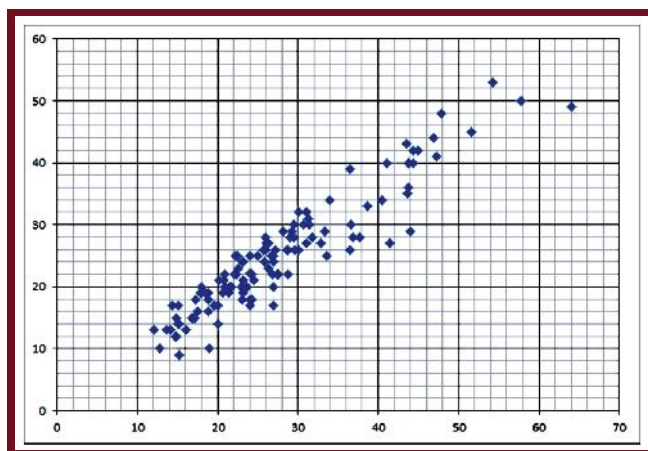
بازه تقریب	قاعده تقریب	h	$N_{\text{citations}}$
[۱, ۳]	۲/۴	۳	۱۹
[۳, ۶]	۴/۸	۶	۸۰
[۴, ۷]	۵/۷	۶	۱۱۳
[۴, ۸]	۶/۱	۴	۱۳۰
[۵, ۱۰]	۷/۷	۶	۲۰۲
[۹, ۱۵]	۱۲/۲	۱۱	۵۱۱

بازه تقریب	قاعده تقریب	h	$N_{\text{citations}}$
[۱, ۲]	۱/۷	۲	۱۰
[۱, ۲]	۱/۸	۲	۱۱
[۱, ۳]	۲/۷	۳	۲۵
[۲, ۵]	۴/۰	۴	۵۴
[۳, ۵]	۴/۳	۵	۶۴
[۳, ۵]	۴/۳	۵	۶۴
[۳, ۵]	۴/۴	۶	۶۷
[۴, ۷]	۵/۵	۶	۱۰۴
[۴, ۸]	۶/۵	۸	۱۴۴
[۶, ۱۱]	۸/۹	۵	۲۶۹

جدول ۴.۳: دانشیاران سه دانشگاه تحقیقاتی

۱-۷ آکادمی ملی علوم (آمریکا)

ما قاعده تقریب خود را برای تمام ۱۱۹ ریاضیدان عضو آکادمی ملی علوم آزمودیم (شکل‌های (۱.۳) و (۲.۳) و ضمیمه را ببینید). ضریب همبستگی $R = ۰/۹۳$ به دست آمد. پس از حذف کتاب‌ها (بر اساس فهرست MathSciNet)، $R = ۰/۹۵$ حاصل شد. یک نگرانی عمده آن است که ارجاعات قبل از سال ۲۰۰۰ در MathSciNet فهرست نشده است. به هر حال، از دیدگاه ما نتایج هنوز قابل اعتمادند. توضیحات بخش «مطالعه بیشتر» را ملاحظه کنید.



مربع دارفی را نشان می‌دهد. مثلاً ظاهر شدن افزایشی از اندازه ۹۸٪ محتمل است. با تعبیر این‌ها برحسب مسئله h -ایندکس، این داده‌ها همان محاسباتی هستند که جدول ۴.۳ را به ما داده‌اند^۲ نشان می‌دهد که نتیجه حاصل از مدل اندازه مربع دارفی تقریباً برابر است با $۰/۵۴\sqrt{N_{\text{citations}}}$. حدس تمرکز کن‌فیلد بیان می‌کند که برای هر $\epsilon > ۰$ ، وقتی $N_{\text{citations}} \rightarrow \infty$ داریم:

$$\frac{\# \text{ partitions with } (\lambda - \epsilon)\mu < h < (\lambda + \epsilon)\mu}{\# \text{Par}(N_{\text{citations}})} \rightarrow 1 \quad (۲.۳)$$

که در اینجا $\mu = \frac{\sqrt{6} \log 2}{\pi} \sqrt{N_{\text{citations}}}$. این نتیجه با جدول ۱ سازگار است. بحث بیشتر در این زمینه در مقاله‌ای دیگر خواهد آمد. همچنین، می‌توان توزیع‌های احتمالی دیگری مانند اندازه پلانچرل را روی نمودار یانگ آزمایش نمود که به هر شکل λ احتمال $(f^\lambda)/|\lambda|!$ را نسبت می‌دهد. در اینجا، f^λ تعداد جدول‌های یانگ استاندارد در شکل λ است.

۷ مقایسه‌های بیشتر با داده‌های تجربی

^۲درواقع، محاسبه $\text{Par}(N_{\text{citations}})$ با استفاده از سری مولد، وقتی $N_{\text{citations}}$ چندین هزار است، روی یک کامپیوتر رومیزی چندان آسان نخواهد بود. در عوض می‌توانیم از تقریب هاردی-راموناگان $\text{Par}(N_{\text{citations}}) \sim \frac{1}{4N_{\text{citations}}} e^{\pi \sqrt{\frac{2N_{\text{citations}}}{3}}}$ استفاده کنیم. حتی به‌طور دقیق‌تر می‌توانیم از Wolfram Alpha، که تعداد افزاها تا یک میلیون را به دست می‌دهد، استفاده کرد، که این چیزی بیش از نیاز ما است.

تقریب فقط می‌تواند یک «حالت تعادل» که سال‌ها پس از انتشار یک نتیجه مهم اتفاق می‌افتد، را منعکس نماید. این نگرانی در مورد برندگان اخیر مدال فیلدز تقریباً کمتر وجود دارد (جدول ۱). اما، هم زمان با آن که MathSciNet ارجاعات قبلی بیشتری را وارد فهرست می‌کند، می‌توان سعی بر کمی کردن این تأثیرات کرد. در حال حاضر استفاده از MathSciNet دارای توجیه عملی است، زیرا در ارتقاء شغلی یا تصمیمات مربوط به پژوهانه، آثار علمی تازه اهمیت دارند. پس برای این موارد، داده‌های پس از سال ۲۰۰۰ کافی هستند.

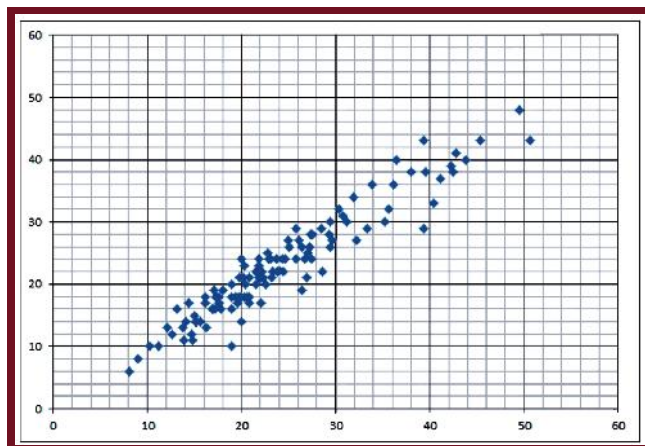
به عنوان یک آزمون دیگر، قاعده تقریب را برای حوزه وسیعی از علوم و با استفاده از Google Scholar به کار بردیم. برای پژوهشگرانی با تعداد متوسطی از ارجاعات، اتکا به این روش اغلب به طور مشابه خوب است. همچنین، این قاعده برای اکثریت وسیعی از پژوهشگران با ارجاعات زیاد، یک کران بالا به دست می‌دهد (ولی، همان‌طور که قبلاً متذکر شدیم، در برخی حوزه‌ها از دقت کمتری برخوردار است). ولی این آزمون‌ها یک اشکال عمده دارند، چرا که فقط افراد دارای نمایه را در نظر می‌گیرند. در نتیجه، از ارائه این نتایج در اینجا اجتناب می‌ورزیم.

ما قاعده تقریب و بازه‌های اطمینان را به عنوان یک روش سیستماتیک مطالعه پیشنهاد کرده و ادعا می‌کنیم که قاعده تقریب یک «پژوهشگر ایدآل» را مشخص می‌کند (این اصطلاح به عاریت گرفته شده از اصطلاح «گاز ایدآل» در مکانیک آماری است. در حقیقت، کاربرد بسیار مفیدی از افزایش‌های تصادفی در مطالعه آماره بولتزمن روی یک شبکه یک بعدی از گاز فرمیون وجود دارد). انحراف از این ایدآل، نتیجه‌ای از ناهنجاری‌هاست. برای یک حوزه کاری آیا از نظر آماری می‌توان پژوهشگرانی را که نزدیک به قاعده تقریب هستند از آنهایی که دور از آن هستند تمیز داد؟

قدردانی

نویسنده از جرج اندروز، راد کن فیلد، جان دو آنجلو، فیلیپ دی فرانچسکو، سرگئی فومین، ایلیا کاپوویچ، الکساندر کاستوچکا، گابریل لاناوه، تام نوینس، الیور پچنیک، جیم پراپ، بروس رزنیک، هال شنک، اندرو سیلز، آرمین استراب، هیو توماس، الکساندر وو، آن یانگ و دیوید یانگ برای توصیه‌های مفیدشان تشکر می‌نماید. به ویژه از جان دو آنجلو و ایلیا کاپوویچ به دلیل تشویق‌شان سپاسگزاریم. این متن در طول یک کلاس ترکیبیات در دانشگاه UIUC تهیه شد؛ نویسنده از اظهارنظرهای دانشجویان کمال تشکر را دارد. این کار توسط پژوهانه NSF مورد حمایت قرار گرفته است.

شکل ۱.۳: قاعده تقریب (محور افقی) در مقابل h واقعی (محور عمودی) برای ریاضیدانان عضو آکادمی ملی علوم



شکل ۲.۳: قاعده تقریب (محور افقی) در مقابل h واقعی (محور عمودی) برای ریاضیدانان عضو آکادمی ملی علوم (بدون کتاب‌ها)

۸ برندگان جایزه آبل

شاید جایزه آبل بیش از مدال فیلدز به جایزه نوبل شباهت داشته باشد، زیرا فاقد محدودیت سنی است. در اینجا نیز سازگاری با بازه‌های تقریب معقول است؛ ولی نگرانی در مورد ارجاعات قبل از سال ۲۰۰۰ به قوت خود باقی است. جدول ۳.۳ را ببینید.

۹ دانشیاران

سر انجام، در جدول ۴.۳ تمام دانشیاران ریاضی سه مؤسسه پژوهشی را مورد بررسی قرار داده‌ایم. در بین سی و دو نفر، همه به جز پنج نفر h -ایندکسی در بازه برآورد شده دارند، و همه آن‌های دیگر حداکثر یک واحد خارج از این حوزه هستند.

۱۰ مطالعه بیشتر

به نظر می‌رسد که مدل ساده ارائه شده، نیروی هدایتگر h -ایندکس را توضیح می‌دهد. ولی نیروها یا منابع اختلال دیگری، بسته به حوزه کاری یا حتی شهرت پژوهشگر، در کار هستند. به علاوه فهم بهتر این کمی سازی نیاز به کار بیشتری دارد تا مدل را دقیق تر سازد.

عدم وجود ارجاعات قبل از سال ۲۰۰۰ در MathSciNet بر چگونگی تعبیر این نتایج برای اعضای آکادمی ملی علوم و برندگان جایزه آبل تأثیرگذار است. به عنوان مثال، توافق تقریبی با قاعده

اعضاء	$N_{\text{citations}}$	قاعده تقریب	h	بدون هیچ کتابی	تخمین اصلاح شده	h اصلاح شده
G. Andrews	۴۸۶۶	۳۷/۷	۲۸	۲۵۷۹	۲۷/۴	۲۴
M. Artin	۲۳۲۶	۲۶	۲۶	۲۰۹۷	۲۴/۷	۲۴
M. Aschbacher	۱۳۸۶	۲۰	۱۷	۹۱۱	۱۶/۳	۱۳
R. Askey	۲۴۸۰	۲۶/۹	۱۷	۱۲۳۵	۱۹/۰	۱۶
M. Atiyah	۶۵۶۴	۴۳/۷	۴۰	۵۳۹۰	۳۹/۶	۳۸
H. Bass	۲۴۷۲	۲۶/۸	۲۲	۱۸۶۹	۲۳/۳	۲۲
E. Berlekamp	۷۶۴	۱۴/۹	۱۲	۳۶۳	۱۰/۳	۱۰
J. Bernstein	۲۵۹۷	۲۷/۵	۲۲	۲۴۸۴	۲۶/۹	۲۱
S. Bloch	۱۴۹۷	۲۰/۹	۲۰	۱۳۶۳	۱۹/۹	۱۸
E. Bombieri	۱۷۴۶	۲۲/۶	۲۳	۱۶۰۸	۲۱/۷	۲۲
J. Bourgain	۶۹۱۹	۴۴/۹	۴۲	۶۵۹۰	۴۳/۸	۴۰
H. Brezis	۱۱۴۶۸	۵۷/۸	۵۰	۸۳۸۶	۴۹/۵	۴۸
F. Browder	۲۸۱۵	۲۸/۷	۲۲	۲۸۰۷	۲۸/۶	۲۲
W. Browder	۶۴۶	۱۳/۷	۱۳	۵۴۷	۱۲/۶	۱۲
R. Bryant	۱۴۸۹	۲۰/۸	۲۱	۱۲۲۸	۱۸/۹	۲۰
L. Cañarelli	۶۷۴۵	۴۴/۳	۴۲	۶۲۸۰	۴۲/۸	۴۱
E. Calabi	۱۲۲۴	۱۸/۹	۱۸	۱۲۲۴	۱۸/۹	۱۸
L. Carleson	۱۹۸۰	۲۴	۱۸	۱۴۸۴	۲۰/۸	۱۷
S.-Y. Alice Chang	۱۸۲۸	۲۳/۱	۲۴	۱۸۰۶	۲۲/۹	۲۴
J. Cheeger	۳۳۸۷	۳۱/۴	۳۰	۳۳۴۸	۳۱/۲	۳۰
D. Christodoulou	۷۸۳	۱۵/۱	۱۷	۵۹۴	۱۳/۲	۱۶
A. Connes	۶۴۷۵	۴۳/۵	۴۳	۵۳۱۸	۳۹/۴	۴۳
I. Daubechies	۴۶۷۴	۳۶/۹	۲۸	۳۰۰۲	۲۹/۶	۲۷
P. Deift	۳۰۰۴	۲۹/۶	۲۶	۲۵۴۵	۲۷/۲	۲۶
P. Deligne	۶۵۶۷	۴۳/۸	۳۶	۵۵۹۲	۴۰/۴	۳۳
P. Diaconis	۳۲۳۳	۳۰/۷	۳۰	۲۹۷۰	۲۹/۴	۳۰
S. Donaldson	۲۷۱۲	۲۸/۱	۲۹	۲۲۷۷	۲۵/۸	۲۹
E. Dynkin	۱۵۸۳	۲۱/۵	۲۰	۱۰۹۰	۱۷/۸	۱۶
Y. Eliashberg	۱۶۲۸	۲۱/۸	۲۰	۱۴۶۰	۲۰/۶	۱۸
L. Faddeev	۱۸۲۰	۲۳	۲۰	۱۲۸۵	۱۹/۴	۱۸
C. Feferman	۳۸۲۸	۳۳/۴	۲۹	۳۸۱۵	۳۳/۴	۲۹
M. Freedman	۱۲۰۷	۱۸/۸	۱۶	۹۹۰	۱۷	۱۶
W. Fulton	۵۸۹۰	۴۱/۴	۲۷	۱۴۲۴	۲۰/۴	۲۰
H. Furstenberg	۲۰۶۴	۲۴/۵	۲۱	۱۶۵۰	۲۱/۹	۲۱
D. Gabai	۱۳۱۴	۱۹/۶	۱۷	۱۳۱۴	۱۹/۶	۱۷
J. Glimm	۱۸۲۶	۲۳/۱	۱۸	۱۴۱۹	۲۰/۳	۱۸
R. Graham	۳۸۸۱	۳۳/۶	۲۵	۲۲۸۰	۲۵/۸	۲۴
U. Grenander	۸۹۵	۱۶/۱	۱۳	۲۲۷	۸/۱	۶
P. Griffiths	۴۵۸۱	۳۶/۵	۲۶	۱۶۹۲	۲۲/۲	۲۲
M. Gromov	۷۶۷۱	۴۷/۳	۴۱	۶۲۰۰	۴۲/۵	۳۸
B. Gross	۱۶۹۲	۲۲/۲	۲۵	۱۶۳۵	۲۱/۸	۲۴

اعضاء	$N_{\text{citations}}$	تقریب قاعده	h	بدون هیچ کتابی	تخمین اصلاح شده	h اصلاح شده
V. Guillemin	۳۷۱۰	۳۲/۹	۲۷	۲۰۳۵	۲۴/۴	۲۲
R. Hamilton	۲۴۹۰	۲۶/۹	۲۰	۲۳۹۲	۲۶/۴	۱۹
M. Hochster	۱۷۲۷	۲۲/۴	۲۲	۱۶۵۷	۲۲	۲۱
H. Hofer	۲۱۴۰	۲۵	۲۵	۱۹۲۸	۲۳/۷	۲۴
MJ. Hopkins	۷۱۴	۱۴/۴	۱۷	۷۱۴	۱۴/۴	۱۷
R. Howe	۱۶۸۰	۲۲/۱	۲۲	۱۵۷۹	۲۱/۵	۲۲
H. Iwaniec	۲۸۲۲	۲۸/۷	۲۶	۱۸۲۵	۲۳/۱	۲۴
A. Jaffe	۷۹۴	۱۵/۲	۹	۲۷۷	۹	۸
P. Jones	۱۱۱۲	۱۸	۱۹	۱۱۱۲	۱۸	۱۹
V. Jones	۲۰۲۵	۲۴/۳	۱۸	۱۶۶۹	۲۲/۱	۱۷
R. Kadison	۱۹۲۲	۲۳/۷	۲۰	۱۰۴۲	۱۷/۴	۱۸
R. Kalman	۵۵۸	۱۲/۸	۱۰	۴۲۸	۱۱/۲	۱۰
N. Katz	۲۳۷۰	۲۶/۳	۲۳	۱۵۸۲	۲۱/۵	۲۰
D. Kazhdan	۲۳۳۲	۲۶/۱	۲۷	۲۳۳۲	۲۶/۱	۲۷
R. Kirby	۹۶۳	۱۶/۸	۱۵	۶۷۸	۱۴/۱	۱۴
S. Klainerman	۲۳۲۴	۲۶	۲۸	۲۱۴۴	۲۵/۰	۲۷
J. Kohn	۱۲۳۱	۱۸/۹	۱۹	۱۰۶۸	۱۷/۶	۱۸
J. Kollár	۳۱۰۰	۳۰/۱	۲۶	۱۹۴۷	۲۳/۸	۲۲
B. Kostant	۲۵۰۹	۲۷	۲۵	۲۵۰۹	۲۷	۲۵
R. Langlands	۱۴۶۶	۲۰/۶	۱۹	۷۷۳	۱۵/۰	۱۵
H.B. Lawson	۲۵۷۶	۲۷/۴	۲۲	۱۸۴۶	۲۳/۲	۲۱
P. Lax	۴۶۰۱	۳۶/۶	۳۰	۲۵۶۰	۳۲/۲	۲۷
E. Lieb	۵۱۴۷	۳۸/۷	۳۳	۴۳۴۹	۳۵/۶	۳۲
T. Liggett	۱۹۷۵	۲۴	۱۷	۹۸۴	۱۶/۹	۱۶
L. Lovasz	۵۶۳۸	۴۰/۵	۳۴	۴۲۵۹	۳۵/۲	۳۰
G. Lusztig	۵۷۸۶	۴۱/۱	۴۰	۴۹۴۵	۳۸/۰	۳۸
R. MacPherson	۲۰۳۱	۲۴/۳	۲۲	۱۶۷۶	۲۲/۱	۲۱
G. Margulis	۲۲۶۷	۲۵/۷	۲۶	۱۷۸۸	۲۲/۸	۲۵
J. Mather	۱۳۹۹	۲۰/۲	۲۱	۱۳۹۹	۲۰/۲	۲۱
B. Mazur	۲۸۴۲	۲۸/۸	۲۶	۲۴۴۰	۲۶/۷	۲۴
D. McDuff	۲۲۸۹	۲۵/۸	۲۴	۱۴۱۷	۲۰/۳	۲۳
H. McKean	۲۴۸۰	۲۶/۹	۲۴	۱۷۰۱	۲۲/۳	۲۱
C. McMullen	۱۷۳۸	۲۲/۵	۲۵	۱۳۶۸	۲۰/۰	۲۴
J. Milnor	۷۸۵۶	۴۷/۹	۴۸	۴۵۵۹	۳۶/۵	۴۰
J. Morgan	۱۹۸۵	۲۴/۱	۲۵	۱۴۸۴	۲۰/۸	۲۱
G. Mostow	۱۱۸۰	۱۸/۵	۱۹	۸۹۶	۱۶/۲	۱۷
J. Nash	۱۳۳۷	۱۹	۱۰	۱۳۳۷	۱۹	۱۰
E. Nelson	۱۰۱۰	۱۷/۲	۱۵	۷۵۳	۱۴/۸	۱۱
L. Nirenberg	۹۱۴۵	۵۱/۶	۴۵	۸۷۸۱	۵۰/۶	۴۳

اعضاء	$N_{\text{citations}}$	تقریب قاعده	h	بدون هیچ کتابی	تخمین اصلاح شده	h اصلاح شده
S. Novikov	۲۳۶۸	۲۶/۳	۲۷	۱۶۷۷	۲۲/۱	۲۱
A. Okounkov	۱۶۷۷	۲۲/۱	۲۴	۱۶۷۷	۲۲/۱	۲۴
D. Ornstein	۱۱۰۰	۱۷/۹	۱۹	۱۰۲۲	۱۷/۳	۱۸
J. Palis	۱۵۷۰	۲۱/۴	۱۹	۸۹۵	۱۶/۲	۱۸
P. Rabinowitz	۶۶۳۳	۴۴	۲۹	۵۳۱۶	۳۹/۴	۲۹
M. Ratner	۵۰۶	۱۲/۱	۱۳	۵۰۶	۱۲/۱	۱۳
K. Ribet	۱۰۲۲	۱۷/۳	۱۸	۱۰۲۱	۱۷/۳	۱۸
P. Sarnak	۳۱۱۴	۳۰/۱	۳۲	۲۷۸۰	۲۸/۵	۲۹
M. Sato	۷۳۸	۱۴/۷	۱۲	۷۳۸	۱۴/۷	۱۲
R. Schoen	۳۹۴۵	۳۳/۹	۳۴	۳۴۹۳	۳۱/۹	۳۴
J. Serre	۱۰۱۱۹	۵۴/۳	۵۳	۴۴۸۱	۳۶/۱	۳۶
C. Seshadri	۹۸۴	۱۶/۹	۱۵	۸۳۱	۱۵/۶	۱۴
Y. Sinai	۳۳۵۷	۳۱/۳	۳۱	۲۵۴۷	۲۷/۳	۲۸
I. Singer	۲۹۸۲	۲۹/۵	۲۸	۲۹۵۱	۲۹/۳	۲۸
Y. Siu	۱۴۹۴	۲۰/۹	۲۲	۱۳۵۰	۱۹/۸	۲۱
S. Smale	۴۵۸۱	۳۶/۵	۳۹	۳۹۴۲	۳۳/۹	۳۶
R. Solovay	۷۸۱	۱۵/۱	۱۴	۷۸۱	۱۵/۱	۱۴
J. Spencer	۷۵۸	۱۴/۹	۱۵	۱۳۳۴	۱۹/۷	۱۸
R. Stanley	۶۵۱۰	۴۳/۶	۲۵	۳۱۴۸	۳۰/۳	۳۲
H. Stark	۶۷۸	۱۴/۱	۱۳	۶۵۳	۱۳/۸	۱۳
C. Stein	۷۶۳	۱۴/۹	۱۲	۶۵۸	۱۳/۹	۱۱
E. Stein	۱۴۰۴۹	۶۴	۴۹	۵۷۸۸	۴۱/۱	۳۷
R. Steinberg	۱۸۵۰	۲۳/۲	۱۹	۱۰۶۸	۱۷/۶	۱۷
S. Sternberg	۲۴۳۸	۲۶/۷	۲۵	۱۴۷۶	۲۰/۸	۱۸
D. Stroock	۳۲۹۹	۳۱/۰	۲۷	۲۰۲۸	۲۴/۳	۲۴
D. Sullivan	۳۳۰۷	۳۱/۱	۳۲	۳۲۴۸	۳۰/۸	۳۱
R. Swan	۱۱۰۹	۱۸	۲۰	۹۹۸	۱۷/۱	۱۹
E. Szemerédi	۲۵۳۶	۲۷/۲	۲۶	۲۵۳۶	۲۷/۲	۲۶
T. Tao	۶۷۳۰	۴۴/۳	۴۰	۶۲۱۴	۴۲/۳	۳۹
J. Tate	۲۹۷۹	۲۹/۵	۳۰	۲۶۱۲	۲۷/۶	۲۸
C. Taubes	۱۸۶۶	۲۳/۲	۲۴	۱۶۲۶	۲۱/۸	۲۳
J. Thompson	۷۸۹	۱۵/۲	۱۴	۷۸۹	۱۵/۲	۱۴
J. Tits	۳۴۶۳	۳۱/۸	۲۸	۲۹۵۸	۲۹/۴	۲۶
K. Uhlenbeck	۱۸۵۲	۲۳/۲	۲۱	۱۷۵۶	۲۲/۶	۲۰
S. Varadhan	۲۸۹۴	۲۹	۲۸	۲۱۵۳	۲۵/۱	۲۶
D. Voiculescu	۲۹۵۲	۲۹/۳	۲۹	۲۳۸۷	۲۶/۴	۲۶
A. Wiles	۱۳۸۷	۲۰	۱۴	۱۳۸۷	۲۰	۱۴
S-T. Yau	۷۵۳۶	۴۶/۹	۴۴	۷۰۶۶	۴۵/۴	۴۳
E. Zelmanov	۱۰۵۵	۱۷/۵	۱۶	۱۰۲۰	۱۷/۲	۱۶

مراجع

- [6] J. E. Hirsch, An index to quantify an individual's scientific research output, *Proc. Natl. Acad. Sci.* 102(46) (2005), November 15, pp. 16569–16572.
- [7] J. E. Iglesias and C. Pecharromán, Scaling the h-index for different scientific ISI fields, *Scientometrics* 73, Issue 3, (2007), December, pp. 303–320.
- [8] Z. Su, Asymptotic analysis of random partitions, *Asymptotic theory in probability and statistics with applications*, Adv. Lect. Math. (ALM) 2, Int. Press, Somerville, MA, 2008, pp. 44–79.
- [9] A. Yong, The Joseph Greenberg problem: Combinatorics and comparative linguistics, preprint 2013. *arxiv:1309.5883*.
- [1] T. Anderson, R. Hankin and P. Killworth, Beyond the Durfee square: Enhancing the h-index to score total publication output, *Scientometrics* 76, no. 3 (2008), 577–588.
- [2] R. A. Brualdi, *Introductory Combinatorics*, fifth edition, Pearson Prentice Hall, Upper Saddle River, New Jersey, 2010.
- [3] E. R. Canfield, S. Corteel and C. D. Savage, Durfee polynomials, *Elect. J. Comb.* 5 (1998), R32.
- [4] L. Cooper, Making Estimates in Research and Elsewhere, <http://physics.illinois.edu/undergrad/SeniorThesis/EstimatesResearch.pdf>.
- [5] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, 2009.

* دانشگاه فردوسی مشهد