

НАУЧНАЯ ЖИЗНЬ

Самарская Лука: проблемы региональной и глобальной экологии.
2017. – Т. 26, № 3. – С. 5-14.

ИНДЕКС ДЛЯ КОЛИЧЕСТВЕННОЙ ОЦЕНКИ ИНДИВИДУАЛЬНОЙ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ ПРОДУКЦИИ

© 2017 Х. Хирш¹

Физический факультет Калифорнийского университета в Сан-Диего,
Ла-Хойя, Калифорния, США
E-mail: jhirsch@ucsd.edu.

Представлена 1 сентября 2005 г. Мануэлем Кардона²,
Институт физики твердого тела им. Макса Планка, Штутгарт (Германия)
(получена после исправлений 15 августа 2005 г.)

Предлагаю индекс h , который определяется как число статей с цитированием $>h$, в качестве полезного показателя для характеристики результатов научной деятельности исследователя.

Ключевые слова. Цитаты, результативность, непредвзятость.

Hirsch J.E. An index to quantify an individual's scientific research output. – I propose the index h , defined as the number of articles with citations $>h$, as a useful index to characterize the scientific output of a researcher.

Key words. Quotes, effectiveness and impartiality.

Для тех немногих ученых, получивших Нобелевскую премию, результативность и актуальность своих исследований несомненна. Среди остальных, – как можно количественно оценить совокупное влияние и актуальность научных исследований конкретного естествоиспытателя? В мире ограниченных ресурсов, такое количественное описание (даже если это потенциально неприятно) часто необходимо для оценки и сравнения (например, при найме преподавателей университета, продвижении по службе, присуждении грантов и т. д.).

Список публикаций исследователя и цитирование его работ явно являются данными, которые содержат полезную информацию. Эта информация включает в себя общее число статей (N_p), опубликованных за n лет, количество цитирований (N_c^j) для каждой статьи (j), журналы, где опубликованы эти статьи, их [журналов] рейтинг и т. д. Этот большой объем информации оценивается по различным критериям разными людьми. Здесь я хотел бы предложить один показатель, « h -индекс»,

¹ **Hirsch J.E.** An index to quantify an individual's scientific research output // PNAS (Proceedings of the National Academy of Sciences of the United States of America). 2005. V. 102, No. 46. P. 16569-16572. Перевод Г.С. Розенберга.

² Кардона Мануэл (Manuel Cardona Castro; 1934-2014) – испанский физик, специализирующийся по физике твердого тела; по данным ISI Citations Web Database (Филадельфия, США) был одним из 8 самых цитируемых физиков с 1970 г.

как очень простой и полезный способ характеристики результатов научной деятельности исследователя.

Ученый имеет индекс h , если h из его N_P статей процитирована не менее h раз, а каждая из оставшихся $(N_P - h)$ процитирована менее h раз ($\leq h$).

Примеры, которые я привожу в этой статье, – это цитируемость работ физиков; однако, я полагаю, что индекс h должен быть полезным и для других научных дисциплин. (В конце статьи я обсужу некоторые особенности использования индекса h в биологических науках). Самый высокий h среди физиков получен для Э. Виттена (Edward Witten; $h = 110$). То есть, Виттен написал 110 статей с не менее, чем 110 цитирований каждая. Это задает нижнюю границу общего числа цитирований статей Виттена $h^2 = 12100$. Конечно, общее число цитирований ($N_{C_{tot}}$), как правило, значительно больше, чем h^2 ; значение h^2 занижает общее число цитирований, т. к. максимально цитируемые работы, часто, цитируются значительно чаще, чем h , а публикации с $< h$ игнорируются. Связь между $N_{C_{tot}}$ и h будет зависеть от формы конкретного распределения (Laherrere, Sornette, 1998), и это позволяет определить коэффициент пропорциональности a из выражения

$$N_{C_{tot}} = a h^2 . \quad [1]$$

Я эмпирически определил, что a колеблется между 3 и 5.

Другие видные физики с высоким h -индексом – это Alan Jay Heeger ($h = 107$), Marvin L. Cohen ($h = 94$), Arthur C. Gossard ($h = 94$), Philip Warren Anderson ($h = 91$), Steven Weinberg ($h = 88$), Michael E. Fisher ($h = 88$), Manuel Cardona ($h = 86$), Pierre-Gilles de Gennes ($h = 79$), John Norris Bahcall ($h = 77$), Zachary Fisk ($h = 75$), Douglas James Scalapino ($h = 75$), Giorgio Parisi ($h = 73$), Steven G. Louie ($h = 70$), Roman Wladimir Jackiw ($h = 69$), Frank Anthony Wilczek ($h = 68$), Cumrun Vafa ($h = 66$), M. Brian Maple ($h = 66$), David Jonathan Gross ($h = 66$), Mildred S. Dresselhaus ($h = 62$) и Stephen William Hawking ($h = 62$). Я утверждаю, что h является более предпочтительным показателем по сравнению с другими «одно числовыми» критериями (single-number criteria), которые обычно используются для оценки результативности научной деятельности исследователя, а именно:

- (i) Общее число статей (N_P). Преимущество: измеряет общую «производительность» ученого. Недостатки: не измеряет важность или влияние статьи [на другие исследования].
- (ii) Общее число цитирований ($N_{C_{tot}}$). Преимущество: измеряет общую результативность и важность статьи. Недостатки: сложность нахождения и возможность завышения за счет малого числа «больших хитов [big hits]» с большим числом соавторов. В таких случаях, в выражении [1] возможно получение нетипичного значения a , большего 5. Еще одним недостатком является то, что $N_{C_{tot}}$ придает чрезмерный вес высокоцитируемым обзорным работам в сравнении с оригинальными исследованиями.
- (iii) Средняя цитируемость одной статьи (т. е. $N_{C_{tot}}$ к N_P). Преимущество: позволяет сравнивать ученых разных возрастов. Недостатки: сравнительно редкий показатель, отдает предпочтение «малопродуктивным» исследователям по сравнению с «высокопродуктивными».
- (iv) Число «важных статей» (определяется как количество статей с цитируемостью, например, $y > 50$). Преимущество: устраняет недостатки критериев *i*, *ii* и *iii*, и дает представление о широте и устойчивости результативности.

Недостаток: y является произвольным и случайно выбранным граничным значением, что позволяет манипулирование в пользу тех или иных физических лиц; кроме того, число y должно быть скорректировано на разных уровнях иерархии.

- (v) Число цитирований каждой из q наиболее цитируемых статей (например, $q = 5$). Преимущество: преодолевает многие из недостатков вышеперечисленных критериев. Недостаток: это не «одно числовой» критерий, что затрудняет его получение и сравнение. Кроме того, q является произвольным и случайным образом выбранной величиной, что также несет элемент субъективности по отношению к конкретным исследователям.

Напротив, предлагаемый h -индекс позволяет измерять широкий спектр параметров результативности исследований и избежать всех недостатков выше перечисленных критериев; кроме того, необходимую информацию легко найти, используя, например, Thomson ISI Web of Science Database (<http://isiknowledge.com>)*, что позволит получить примерную оценку общего количества цитирований (выражение [1]).

Таким образом, я утверждаю, что два исследователя, со сходными значениями h , сопоставимы с точки зрения их воздействия на научное сообщество, даже если их общее число научных публикаций или общее число цитирований очень разнятся. Напротив, сравнивая двух исследователей (одного и того же научного возраста) с одинаковым общим количеством статей или общим числом цитирований, но с разным значением h -индекса, можно делать вывод о том, что тот из них, у кого h больше, вероятно, будет более видным ученым.

Для конкретного человека, можно ожидать, что h -индекс должен примерно линейно возрастать с течением времени. В простейшей модели предполагается, что исследователь публикует p статей в год и что каждая публикация «зарабатывает» c новых цитирований каждый последующий год. Тогда общее число цитирований за $n + 1$ год:

$$N_{c,tot} = \sum_{j=1}^n pcj = \frac{pcn(n+1)}{2} . \quad [2]$$

Если предположить, что все работы y этого года определяют h , то мы имеем

$$(n - y) c = h , \quad [3a]$$

$$py = h . \quad [3b]$$

Левая часть выражения [3a] – это число ссылок на наиболее «свежие» статьи, вносящие вклад в h ; в левой части выражения [3b] – общее количество работ, вносящих вклад в h . Из [3] следует

$$h = \frac{c}{1 + c/p} n . \quad [4]$$

Тогда общее количество цитирований (для не слишком малых n) приблизительно равно

$$N_{c,tot} \approx \frac{(1 + c/p)^2}{2c/p} h^2 . \quad [5]$$

* Естественно, используемая база данных должна быть достаточно полной, чтобы охватить весь период публикаций ученого.

с учетом [1]. Коэффициент a зависит от числа статей и числа ежегодных цитирований в расчете на статью, задаваемых уравнением [5]. Как уже говорилось ранее, мы эмпирически нашли, что $a \sim 3-5$. Линейная зависимость

$$h \sim m n \quad [6]$$

должна восприниматься достаточно обычно для ученых, которые генерируют научные результаты близкого качества на достигнутом в течение карьеры уровне; конечно, m будет достаточно широко варьировать для разных исследователей. В целом, параметр m при оценке взаимосвязи h с n должен стать полезным критерием для сравнения ученых с разным научным долголетием.

Минимальное значение a в модели [1] будет $a \sim 2$ для случая $c = p$, когда статьи с числом цитирований $> h$ и статьи с цитированием $< h$ в равной степени вносят свой вклад в $N_{C_{tot}}$. Значение a будет больше и при $c > p$, и при $c < p$. Для $c > p$, чаще всего, наибольший вклад в общую цитируемость вносят «высокоцитируемые статьи» (h статей с $N_c > h$), тогда как для $c < p$ это «редкоцитируемые статьи» ($N_p - h$ статей с цитированием каждой $< h$), которые вносят наибольший вклад в $N_{C_{tot}}$. Мы видим, что первая ситуация имеет место в подавляющем большинстве, если не во всех случаях. Для линейной модели данного примера $a = 4$ соответствует $c/p = 5,83$ (другое значение, которое задает $a = 4$, $c/p = 0,17$, нереально).

Линейная модель, определенная выше, соответствует распределению

$$N_c(y) = N_0 - \left(\frac{N_0}{h} - 1 \right) y, \quad [7]$$

где $N_c(y)$ – число цитирований статьи y (от наиболее цитируемых, к наименее) и N_0 – число цитирований наиболее часто цитируемой статьи ($N_0 = cn$ в приведенном выше примере). Общее количество статей y_m получаем из соотношения $N_c(y_m) = 0$; следовательно,

$$y_m = \frac{N_0 h}{N_0 - h}. \quad [8]$$

Мы можем записать N_0 и y_m в терминах выражения [1]:

$$N_0 = h \left[a \pm \sqrt{a^2 - 2a} \right] \quad [9a]$$

$$y_m = h \left[a \mp \sqrt{a^2 - 2a} \right]. \quad [9b]$$

Для $a = 2$, $N_0 = y_m = 2h$. Для больших a верхний знак в уравнении [9a] соответствует случаю, когда доминируют высокоцитируемые статьи (наиболее реалистичный случай); нижний знак соответствует случаю, когда в общем числе цитирований доминируют менее часто цитируемые работы.

В более реалистичной модели $N_c(y)$ не является линейной функцией y . Заметим, что $a = 2$ можно смело считать нижней границей в целом, так как меньшее значение a потребует вблизи больших значений y отрицательности второй производной $\partial^2 N_c / \partial y^2$, что также не реально. Общее число цитирований определяется площадью под кривой $N_c(y)$, которая проходит через точку $N_c(h) = h$. В линейной модели нижней границе $a = 2$ соответствует линия с наклоном -1 , как это показано на рис. 1.

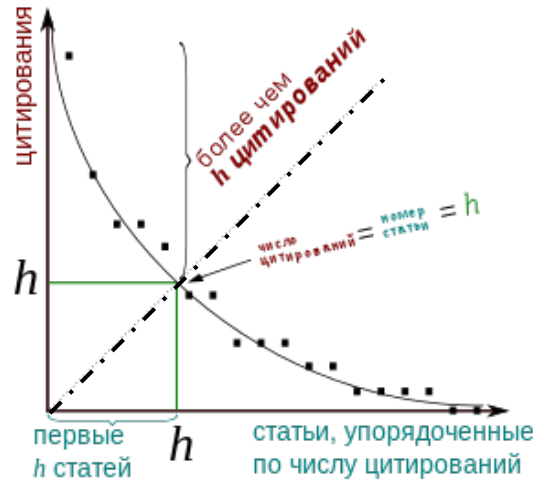


Рис. 1. Схематическая кривая числа цитирований в сравнении с количеством статей, ранжированных в порядке убывания цитируемости.

Пересечение штрихпунктирной линии под углом 45° и кривой дает h . Общее количество цитирований – это площадь под кривой. Предполагая, что вторая производная неотрицательна всюду, минимальная площадь определяется распределением, которое показано пунктирной линией $a = 2$ в выражении [1].

Более реалистическая модель будет иметь экспоненциально растянутый вид:

$$N_c(y) = N_0 \exp\left(-\left\{\frac{y}{y_0}\right\}^\beta\right). \quad [10]$$

Заметим, что для $\beta \leq 1$, $N''_c(y) > 0$ для всех y ; следовательно, $a > 2$. Мы можем записать распределение в терминах h и a в следующем виде:

$$N_c(y) = \frac{a}{\alpha I(\beta)} h \exp\left(-\left\{\frac{y}{h\alpha}\right\}^\beta\right) \quad [11]$$

где

$$I(\beta) = \int_0^\infty dz e^{-z^\beta}, \quad [12]$$

а значение параметра α определяется из выражения:

$$\alpha \exp(\alpha^{-\beta}) = a / I(\beta). \quad [13]$$

Максимально цитируемая работа имеет цитируемость

$$N_0 = \frac{a}{\alpha I(\beta)} h, \quad [14]$$

И общее число статей (хотя бы один раз процитированных) определяется через $N(y_m) = 1$ и имеет вид:

$$y_m = h [1 + \alpha^\beta \ln(h)]^{1/\beta}. \quad [15]$$

Таким образом, распределение для конкретного исследователя может быть смоделировано путем подбора наиболее подходящих β и a в каждом отдельном

случае. Например, для $\beta = 1$, если $a = 3$, $\alpha = 0,661$, $N_0 = 4,54h$ и $y_m = h [1 + 0,66 \ln h]$. Для $a = 4$, $\alpha = 0,4644$, $N_0 = 8,61h$ и $y_m = h [1 + 0,46 \ln h]$. Для $\beta = 0,5$, наименьшее возможное значение $a = 3,70$; в этом случае $N_0 = 7,4h$ и $y_m = h [1 + 0,5 \ln h]^2$. Большие значения a влекут рост N_0 и уменьшение y_m . Для $\beta = 2/3$, минимально возможным будет $a = 3,24$; в этом случае $N_0 = 4,5h$ и $y_m = h [1 + 0,66 \ln h]^{3/2}$.

Линейная зависимость между h и n (выражение [6]) может нарушаться в том случае, когда исследователь «замедляется» в написании статей или вообще перестает их писать. Между этими двумя событиями существует некоторый временной промежуток. В линейной модели предполагается, что исследователь прекращает научную работу через n_{stop} лет; при этом, h продолжает некоторое время расти с той же скоростью:

$$n_{lag} = \frac{h}{c} = \frac{1}{1+c/p} n_{stop} , \quad [16]$$

а потом остается постоянной, потому что все работы уже внесли свой вклад в h . В более реалистичной модели h будет плавно выравниваться с ростом n и не претерпевать дискретные изменения наклона. Однако, в целом, временной лаг будет больше для тех ученых, которые [более равномерно] публикуются на протяжении многих лет (см. выражение [16]).

Кроме того, на самом деле, естественно, не все [научные] работы будут в конечном итоге вносить свой вклад в рост h . Какие-то статьи с низкой цитируемостью никогда не будут способствовать росту h , (особенно, написанные в конце научной карьеры, когда h уже достаточно хорошо сформировался). Как показал С. Реднер (Redner, 1998, 2005), большинство статей «зарабатывают» цитированность за весьма ограниченный период своей популярности и в дальнейшем больше не цитируются. Следовательно, работы, которые способствовали росту h исследователя на ранних стадиях его карьеры, в дальнейшем могут оказаться бесполезными для роста h . Научная работа (или работы), которые в некоторый момент времени имеют ровно h цитирований, рискуют быть исключены (быть исключенными) из h , так как они «вытесняются» другими работами, цитируемость которых растет быстрее. Возможно также, что работы «выпадают (drop out)», но потом могут возвращаться в h и место такой статьи называют «спящей красавицей (sleeping beauties)» (van Raan, 2004).

Для отдельных исследователей, упомянутых выше, я определил n в зависимости от времени, прошедшего с момента их первой опубликованной работы по настоящее время, что позволило определить m в соответствии с уравнением [6]: Witten, $m = 3,89$; Heeger, $m = 2,38$; Cohen, $m = 2,24$; Gossard, $m = 2,09$; Anderson, $m = 1,88$; Weinberg, $m = 1,76$; Fisher, $m = 1,91$; Cardona, $m = 1,87$; de Gennes, $m = 1,75$; Bahcall, $m = 1,75$; Fisk, $m = 2,14$; Scalapino, $m = 1,88$; Parisi, $m = 2,15$; Louie, $m = 2,33$; Jackiw, $m = 1,92$; Wilczek, $m = 2,19$; Vafa, $m = 3,30$; Maple, $m = 1,94$; Gross, $m = 1,69$; Dresselhaus, $m = 1,41$ и Hawking, $m = 1,59$. После изучения цитируемости многих физиков, я пришел к следующим выводам:

- (i) Значение $m \approx 1$ (т. е. h -индекс 20 после 20 лет активной научной деятельности) характеризует успешного ученого.
- (ii) Значение $m \approx 2$ (т. е. h -индекс 40 после 20 лет активной научной деятельности) характеризует выдающегося ученого, которого можно найти только в

лучших университетах мира или крупных научно-исследовательских лабораториях.

(iii) Значение $m \approx 3$ или выше (т. е. h -индекс 60 после 20 лет или 90 после 30 лет активной научной деятельности) характеризует поистине уникальных людей.

Параметр m перестает быть полезным, если ученый не поддерживает свой научный уровень, в то время как параметр h остается значимым в качестве накопительного показателя достижений ученого, который может продолжать расти с течением времени еще долго после того, как ученый прекратил свою публикационную деятельность.

На основе типовых значений h и m , я предлагаю (с большой погрешностью), что для преподавателей в крупных научно-исследовательских университетах, $h \approx 12$ может быть принято за то значение, с которого возможна постановка вопроса о повышении в должности (доцент [associate professor]), а $h \approx 18$ – для полного профессора (full professor). Членство в Американском физическом обществе (American Physical Society) может быть определено индексом $h \approx 15-20$. Членство в Национальной Академии наук США (National Academy of Sciences of the United States of America), как правило, предполагает $h \approx 45$ и выше (кроме исключительных обстоятельств). Замечу, что эти оценки примерно соответствуют стандартному числу лет устойчивой научной деятельности (при $m \approx 1$); временная шкала, конечно, будет короче для ученых с более высоким m . Обращу внимание и на то, что оценки времени берутся от публикации первой научной работы, появление которой происходит за несколько лет до получения, например PhD-степени.

Однако существует ряд особенностей (предостережений, caveats) непосредственного использования h -индекса, которые следует иметь в виду. Очевидно, одно число может дать не более чем грубую оценку многогранного профиля конкретной личности и для ее (личности) полной оценки следует привлекать многие другие факторы. Кроме того, этот факт всегда следует иметь в виду («исключение из правил»), особенно при принятии жизненно важных решений, таких, например, как предоставление или отказ занимать ту или иную должность. Кроме того, наблюдается заметный разброс значений h в различных сферах научной деятельности; в частности, определяют среднее количество ссылок на статьи в данной области знаний, среднее количество научных работ, подготовленных каждым ученым в этой области, и количество ученых, «функционирующих» в этой области (хотя в первом приближении, в широкой области знаний число ученых, работы которых будут много цитироваться, характерное h -значение не обязательно должно быть большим). Ученые, работающие в нетрадиционных (non-mainstream) областях вряд ли смогут достичь столь же высоких значений h , как специалисты высшего звена из тех, кто работает в крайне актуальных направлениях. Хотя я и утверждаю, что большое значение h является надежным показателем высоких достижений, обратное не всегда верно. Существуют значительные различия в асимметрии распределений цитирования даже в пределах данной научной области, и для автора с относительно низким h , который имеет несколько основополагающих работ с чрезвычайно высоким индексом цитирования, индекс h не будет отражать в полной мере его научные достижения. И наоборот, ученый с высоким h , достигнутым, в основном, за счет статей с многочисленными соавторами, был бы «обласкан» (treat overly kindly). Подобласти с неизбежно большим числом совместных исследований в тесном сотрудничестве (например, эксперименты с высокими энергиями), как правило, будут демонстрировать большие h -значения, и я полагаю, что в случаях боль-

шого различия в числе соавторов было бы полезно при сравнении различных лиц нормализовать h на коэффициент, который отражает среднее число соавторов. Для определения научного «возраста» ученого при расчете m , первая статья может иногда не быть отправной точкой, если она представляет собой относительно незначительный вклад задолго до устойчивой производительности исследователя.

Наконец, в какой-то мере при измерении [качества научной деятельности] числом цитат, в идеале хотелось бы исключить самоцитирование. Хотя самоцитирование, очевидно, может увеличить h ученого, его влияние на h намного меньше, чем на общее количество цитирования. Во-первых, все ссылки на статьи с числом цитирования $< h$ не имеют никакого значения, как и самоцитирование для статей с очень высоким h -индексом. Для корректной оценки h с учетом самоцитирования, можно было бы рассмотреть статьи с количеством цитирований просто $> h$ и подсчитать количество самоцитирований в каждой из них. Если статья с $h + n$ цитированиями имеет $> n$ самоцитирований, она будет исключена из расчета h и h будет уменьшено на 1. Как правило, эта процедура требует очень мало каких-либо новых документов. Другая сторона этой медали: ученые заинтересованы в увеличении своего h -индекса за счет самоцитирования и, естественно, цель таких статей просто выйти из зоны $< h$.

В качестве интересного примера я вычислил h и m для физиков, получивших Нобелевскую премию за последние 20 лет (для расчета m , я использовал либо год первой опубликованной работы, либо 1955 год в базе данных ISI Database). Однако мой выбор был ограничен только именами ученых, которые можно было однозначно определить в ISI Database, что сузило наше множество до 76% от общего числа. Это, однако, кроме того несмещенные оценки, потому что общность названия не должна коррелировать с h и m . h -индексы [в нашем случае] изменялись в диапазоне от 22 до 79, а индексы m – в диапазоне от 0,47 до 2,19. Средние значения и стандартные отклонения $\langle h \rangle = 41$, $\sigma_h = 15$ и $\langle m \rangle = 1,14$, $\sigma_m = 0,47$. Распределение h -индексов показано на рис. 2; значение медианы $h_m = 35$, ниже, чем средняя ($\langle h \rangle$) из-за «хвоста» высоких значений h . Интересно, что лауреаты Нобелевской премии имеют существенно более высокие показатели h (для 84% h был не менее 30), что указывает на то, что Нобелевские премии не даются только за «одну удачу», а присуждаются в контексте всей научной работы. Примечательно, что найденные значения m часто не высоки по сравнению с другими успешными учеными (для 49% нашей выборки $m < 1$); это связано, скорее всего с тем, что Нобелевские премии зачастую присуждаются уже после периода максимальной продуктивности исследователей.

Рассмотрим еще один пример. Среди вновь избранных членов Национальной Академии наук США по физике и астрономии в 2005 году, я нахожу $\langle h \rangle = 44$, $\sigma_h = 14$, наивысший $h = 71$, наименьший $h = 20$, медиана $h_m = 46$. Среди всех членов Национальной Академии наук США по физике и астрономии, только для фамилий, которые начинаются с «А» и «В» получаем $\langle h \rangle = 38$, $\sigma_h = 10$ и $h_m = 37$. Эти примеры свидетельствуют о том, что индекс h является стабильным и сопоставимым показателем оценки научных достижений.

Интересная идея – это расширение h -индекса на группу лиц^{**}. База данных литературы по физике высоких энергий (SPIRES database; www.slac.stanford.edu_spIRES_hep) недавно внедрила индекс h в резюме цитирова-

^{**} Впервые было представлено в базе данных SPIRES Database.

ния, а также вычисление h для группы ученых. Общий h -индекс группы, как правило, будет больше, чем у каждого из членов группы, но меньше, чем сумма отдельных показателей h , потому что некоторые из статей, которые способствуют росту h каждого человека, не способствуют росту h для группы. Например, общий h -индекс группы конденсированных сред (condensed matter group) физического факультета Калифорнийского университета в Сан-Диего (University of California at San Diego) $h = 118$, из которых наибольший индивидуальный вклад в 25, наивысший индивидуальный h составляет 66, а сумма всех индивидуальных h больше 300.

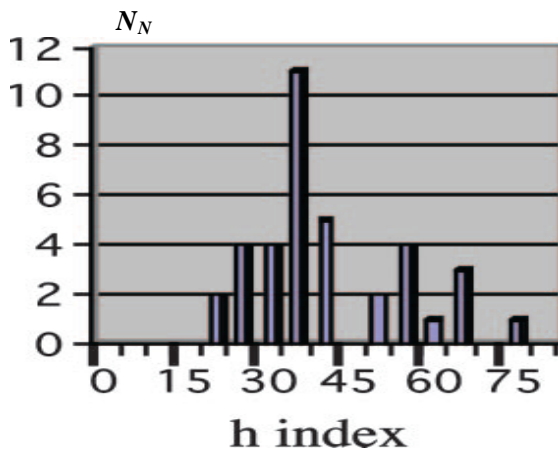


Рис. 2. Гистограмма, демонстрирующая распределение числа лауреатов Нобелевской премии в области физики (N_N) за последние 20 лет в зависимости от их h -индекса. Пик наблюдается для h -индекса между 35 и 39.

Вклад каждого в групповой h не обязательно пропорционален индивидуальному h , и самый высоко цитируемый участник группы h не обязательно специалист с наивысшим значением h . На самом деле, в принципе (хотя редко на практике), специалист с самым низким h -индексом в группе может стать крупнейшим вкладчиком в групповой h . Для потенциального аспиранта с учетом разных программ магистратуры, рейтинг группы в соответствии с их общим h -индексом, вероятно, будет представлять интерес при выборе определенной области знаний; аналогично, и для администраторов, занимающихся этими

вопросами, рейтинги их отделов или целых учреждений в зависимости от общего h также могут быть интересны.

В заключение я хочу обсудить некоторые наблюдения в области биологических и биомедицинских наук. Из списка, составленного Кристофером Кингом (Christopher King) по базе данных Thomson ISI Database, среди самых высоко цитируемых ученых в период 1983-2002 гг. (King, 2003), я определил h -показатели для топ-10 в этом списке (все в науках о жизни), которыми являются, в порядке убывания h : Solomon Halbert Snyder ($h = 191$); David Baltimore ($h = 160$); Robert Charles Gallo ($h = 154$); Pierre Chambon ($h = 153$); Bert Vogelstein ($h = 151$); Salvador Moncada ($h = 143$); Charles A Dinarello ($h = 138$); Tadimitsu Kishimoto ($h = 134$); R. Evans ($h = 127$); Axel Ullrich ($h = 120$). Можно видеть, что не удивительно, что все эти высоко цитируемые исследователи также обладают высокими h -индексами и наивысшие h показатели в науках о жизни значительно выше, чем в физике. Среди 36 новых членов Национальной Академии наук США в области биологических и биомедицинских наук в 2005 году, я определил $\langle h \rangle = 57$, $\sigma_h = 22$, наивысший $h = 135$, самый низкий $h = 18$, а медиана $h_m = 57$. Эти результаты подтверждают тот факт, что h показатели в области биологических наук, как правило, выше, чем в физике; однако, они также показывают, что разница становится значительно выше при максимальных оценках, чем в среднем. Несомненно, дополнительные исследования для понимания сходства и различия h -распределений индекса в различных областях науки были бы интересны.

В заключении еще раз повторю. Я предложил легко вычисляемый индекс h , который дает оценку значимости, важности и широты воздействия совокупного научного вклада ученого. Я полагаю, что этот индекс может служить полезным критерием для непредвзятого сравнения разных исследователей, конкурирующих за один и тот же ресурс, получение которого зависит от важного критерия оценки научных достижений.

Я благодарен многим коллегам из группы конденсированных сред в университете Калифорнии в Сан-Диего и особенно Ивану Шуллер (Ivan Schuller) за стимулирующие дискуссии на эти темы и поощрение опубликовать эти идеи. Я хотел бы также поблагодарить многих читателей, которые писали интересные комментарии, так как эта статья была впервые опубликована в arXiv.org (Hirsch, 2005); рецензентов, которые внесли конструктивные предложения, приведшие к улучшению работы; Трэвиса Брукса (Travis Brooks) и администрацию SPIRES Database за быстрое внедрение h -индекса в базу данных.

СПИСОК ЛИТЕРАТУРЫ

Hirsch, J. E. (2005) arXiv.org E-Print Archive (Aug. 3, 2005). Available at <http://arxiv.org/abs/physics/0508025>.

King, C. (2003) *Sci. Watch* **14**, no. 5, 1.

Laherrere, J. & Sornette, D. (1998) *Eur. Phys. J. E Soft Matter* **B2**, 525–539.

Redner, S. (1998) *Eur. Phys. J. E Soft Matter* **B4**, 131–134. – **Redner, S.** (2005) *Phys. Today* **58**, 49–54.

van Raan, A. F. J. (2004) *Scientometrics* **59**, 467–472.