**معرفی شاخص جدید شاخص اچ بی[[1]](#footnote-1) (hb)**

**اولن آر بروون[[2]](#footnote-2)**

**ترجمه بهار داوودآبادی، دانشجوی کارشناسی ارشد علم اطلاعات و دانش شناسی دانشگاه شهید بهشتی**

**چکیده**

شاخص h دستاورد/تاثیر انتشار نویسندگان را مشخص می‌کند و توسط جورج هیرش مبتکر اینگونه تعریف می‌شود: «تعداد مقالات با تعداد استنادات ≥h». شاخص h این ویژگی ذاتی را دارد که نویسندگانی که مجموع استنادات آن‌ها متفاوت است شاخص h یکسانی دارند. در واقع، مقالاتی که کمتر از h به آنها ارجاع داده‌اند یا استنادهای یک مقاله که بالاتر از h است، هیچ تاثیری در شاخص h نویسنده ندارد.

چنین استنادی، استنادهای «مازاد» هستند که توسط شاخص h اعتباری ندارند. برای رفع این کمبودها، یک اصلاح ساده و سرراست به نام شاخص اچ بی را پیشنهاد می کنیم:$h\_{b}=h+\sqrt{e}h\_{b}=h+ \sqrt{e} $

که در آن h شاخص h هیرش و e مجموع همه استنادها منهای h2 است. بنابراین، e استنادهای مازاد است که توسط شاخص h اعتبار داده نشده است.

واژه‌های کلیدی: تأثیر نشریه، تأثیر نویسنده، شاخص هیرش، شاخص اصلاح‌شده هیرش، تأثیر کتابشناختی

**مقدمه**

شاخص h اولین بار توسط هیرش در مقاله ای با عنوان: «شاخصی برای تعیین کمیت خروجی تحقیقات علمی یک فرد» توصیف شد. شاخص h در ابتدا برای اندازه گیری موفقیت یک فرد در نظر گرفته‌شده‌بود، اگرچه امروزه معمولاً به عنوان معیار تأثیر کتابشناختی استفاده می‌شود. هیرش شاخص h را به عنوان «تعداد مقالات با تعداد استناد ≥h» تعریف کرد. شاخصh مورد پذیرش و حمایت، انتقاد، اصلاحات، تکمیل قرارگرفته‌است و یک جایگزین تک عددی ابداع شده‌است . ارزیابی گسترده‌ای از شاخص h و انواع آن منتشر شده‌است. اخیراً، شاخص h و حوزه مرتبط با تأثیر مجله، موضوع مقالات منتشر شده در مجلات آموزش عالی بوده‌است و همچنین این شاخص برای ارزیابی تأثیر بیماری و برای زمینه‌های علمی مختلف مقیاس‌بندی شده‌است که نشان داده شده است که هر مقاله میانگین استنادهای متفاوت قابل توجهی دارند. شاخص h را می‌توان با استفاده از پایگاه‌های داده‌ای مثل تامپسون ساینتیفیک وب آو ساینس[[3]](#footnote-3) یا گوگل اسکالر سیتیشن[[4]](#footnote-4) محاسبه کرد. یک کتاب و نرم‌افزار مرتبط عموماً برای موضوع تأثیر انتشار، به‌خصوص برای تجزیه و تحلیل استنادات دانشگاهی و برای محاسبه شاخص h مفید است.

ذاتاً، شاخص h ویژگی نامطلوبی دارد که افراد با تعداد استنادات بسیار متفاوت شاخص h یکسانی دارند. هدف ما در این گزارش اصلاح شاخص h برای بیان تفاوت‌ نتایج ناشی از افزایش کل استنادها و حفظ سایر ویژگی‌های شاخص است که در ارزیابی تأثیر انتشار یک فرد از جمله برای ارتقاء، تصدی، تخصیص دارایی و اهداف اداری مرتبط، مفید بوده‌است.

**نتایج**

 زمانی که دقیقاً h مقاله، h استناد دریافت می‌کند، هر شاخص h می‌تواند با حداقل استنادات h ایجاد شود. اثر استنادهای h، عبارت$\sqrt{e}\sqrt{e}$ در شاخصhb را توجیه می‌کند. همچنین، از لحاظ نظری هر شاخص h می‌تواند با هر تعداد کل استناد ≥h2 ایجاد شود. استنادهای بزرگتر از h2، استنادهای «مازاد» هستند و به نظر منطقی و مطلوب به نظر می‌رسد که شاخص h برای اعتبار دادن به این استنادها اصلاح شود. این امر به ویژه برای دانشمندان جوانی که ممکن است تنها چند مقاله و شاخص h نسبتاً پایین داشته باشند، مهم است. بر اساس مقایسه، مطلوب است که یک شاخص تک عددی "میوه در حال رسیدن" را نشان دهد، در حالی که شاخص h فقط میوه "کاملا رسیده" را در یک رویداد نشان می‌دهد، زیرا در موقعیت خاصی که استنادی به یک مقاله "تقریبا رسیده"می‌شود یک افزایش واحد در h ایجاد می‌کند.

شکل.1. شاخص h را توضیح می‌دهد و محدودیت آن را در عدم محاسبه استنادهای «مازاد» نشان می‌دهد. یک رکورد انتشار خاص برای یک نویسنده فرضی که 90 مقاله را منتشر کرده و 820 بار استناد شده‌است، نشان می‌دهد. استنادها به شرح ذیل توزیع می‌شوند: 1 مقاله با 100; 9 مقاله هر کدام با 30; و گروه‌های 10 مقاله هر کدام با 20، 10، 5، 4، 3، 2، 1 و 0 استناد. با بازبینی، شاخص h، 20 است (تعداد مقالات با ≥ h استناد). با این حال، تعداد مقالات و توزیع‌های استنادی آن‌ها تنها یکی از تعداد زیادی با شاخص h، 20 است. در واقع، هر شاخص h را می‌توان از توزیع‌های متعدد ایجاد کرد (به جز مورد بی‌اهمیت با یک مقاله). شکل 1 همچنین نشان می‌دهد که مقاله‌ای که هیچ استنادی ندارد و 20 استناد مازاد دارد هیچ تاثیری روی شاخص h، 20 نمی‌گذارد. به همین ترتیب، هیچ استنادی از مقالات دارای رتبه پایین (آنهایی که در این مورد زیر مقاله بیستم هستند) هیچ تاثیری روی شاخص h نمی‌گذارند (طبق تعریف، دقیقاً 20 مقاله که 20 بار یا بیشتر استناد شده‌اند، شاخص h را ایجاد می‌کنند). استنادهای دیگر "مازاد" هستند زیرا هیچ تاثیری روی شاخص h نمی‌گذارند.

تعداد استنادات

تعداد مقاله

شکل1.

تصویر گرافیکی برای یک نویسنده فرضی با شاخص h 20 (برای جزئیات بیشتر به متن توجه کنید).

شاخص h نتایج نامطلوب دیگری هم دارد. به عنوان مثال، اولین استناد از هر مقاله منتشر شده یک فرد، شاخص h، 1 را ایجاد می‌کند. با این حال، فرد هرچقدر تعداد مقاله بیشتری منتشر کند (هرچقدر هم تعداد آنها زیاد باشد) تا زمانی که دو مقاله دو بار مورد استناد قرار نگیرند، شاخص h تغییری نمی‌کند و (حداقل) n مقاله باید n بار استناد شوند تا شاخص h به n تبدیل شود. بنابراین، استنادهای اضافی ≤h برای هر تعداد مقاله‌ای که h بار استناد شده‌اند، شاخص h را افزایش نمی‌دهند. همچنین، یک مقاله، صرف نظر از اینکه چقدر استناد زیادی داشته باشد، نمی تواند شاخص h بزرگتر از 1 داشته‌باشد، و مجموعه‌ای از مقالات، هر چقدر هم که دارای استناد زیاد باشند، نمی تواند شاخص h بزرگتر از تعداد مقالات موجود در آن مجموعه داشته‌باشند. همچنین، اگر دانشمندی فقط یک مقاله منتشر کند و آن مقاله چندین بار مورد استناد قرار گیرد، شاخص h همان 1 باقی می‌ماند. نامطلوب است که یک شاخص به این شیوه رفتار کند.

جدول 1 شاخص h و شاخص hb را مقایسه می‌کند. سوابق انتشار با توزیع‌های مختلف استناد ایجاد شد تا مفهوم «استنادهای مازاد» را روشن کند و ریاضی را به اندازه کافی ساده کند که بتوان با بازبینی ارزیابی کرد. با این حال، سایر توزیع‌های استنادی حاوی همین نقص‌ها هستند. جدول 1 توزیع های A-F نویسندگان مختلف را با کل استنادات بسیار متفاوت اما تعداد مقالات یکسان برای جداسازی اثر تعداد استنادها مقایسه می‌کند. برای شفاف‌سازی، استنادهای مازاد برای توزیع های A-H در مقالاتی با استنادهای بیشتر از h رخ می‌دهد. بدیهی است که انتشارات با استناد کمتر از h هیچ تاثیری بر h ندارند.

جدول 1.

اثرات اصلاحی شاخص hb

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| توزیع‌های استنادی\* | کل استنادها | استنادهای مازاد (e)✝ | انتشارات | شاخص h | شاخص hb |
| A | 100 | 0 | 10 | 10 | 10 |
| B | 190 | 90 | 10 | 10 | 19.48 |
| C | 200 | 100 | 10 | 10 | 20.00 |
| D | 230 | 130 | 10 | 10 | 21.40 |
| E | 370 | 270 | 10 | 10 | 26.43 |
| F | 550 | 450 | 10 | 10 | 31.21 |
| G | 1050 | 650 | 20 | 20 | 45.50 |
| H± | 7150 | 5550 | 260 | 40 | 114.50 |

\*توزیع A دارای حداقل تعداد انتشارات و استناد است که شاخص h، 10 را ایجاد می‌کند. سایر توزیع‌های استنادی برای شفاف‌سازی مشخص شده و برای ارائه طیف وسیعی از کل استنادها طراحی شده‌اند، اما منحصر به استدلال ما نیستند. توزیع B دارای 190 استناد با 1 مقاله 100 بار و 9 مقاله هر کدام 10 بار استناد شده‌اند، است. توزیع C، 1 مقاله که 50 بار استناد شده، 2 مقاله هر کدام 40 بار استناد شده و 7 مقاله که هرکدام 10 بار استناد شده است، شامل می‌شود. توزیع D، 1 مقاله 50 بار استناد شده، 3 مقاله که هر کدام 30 بار استناد شده، 3 مقاله که هر مقاله 20 بار استناد شده و 3 مقاله هر کدام 10 بار استناد شده‌است، دارد. توزیع E دارای 1 مقاله که100 بار استناد شده، 3 مقاله که هرکدام 50 بار استناد شده، 3 مقاله که هر مقاله 30 بار استناد شده و 3 مقاله که هرکدام 10 بار استناد شده‌اند، است. توزیع F دارای 1 مقاله با هر یک از تعداد استنادهای 100، 90، 80، 70، 60، 50، 40، 30، 20، و 10 است. توزیع G، 1 مقاله 200 بار استناد شده، 4 مقاله هر یک 100 بار استناد شده، 5 مقاله که هر مقاله 50 بار استناد شده و 10 مقاله که هر کدام20 بار استناد شده است، دارد. توزیع H دارای 1 مقاله 500 بار استناد شده، 1 مقاله 300 بار استناد شده، 3 مقاله که هر کدام 150 بار استناد شده، 10 مقاله که هر مقاله 100 بار استناد شده، 25 مقاله که هرکدام 40 بار استناد شده، 60 مقاله که هر مقاله 30 بار استناد شده، 70 مقاله که هر کدام 20 بار استناد شده، 50 مقاله که هر کدام 10 بار و 40 مقاله که هر کدام 5 بار استناد شده اند، است.

✝مجموع آن استنادها به هر مقاله که بزرگتر یا کمتر از مقدار h است. استنادهای مازاد دارای مقدار e هستند که بدین صورت تعریف می‌شود:

$$h\_{b}=h+\sqrt{e}h\_{b}=h+ \sqrt{e}$$

بنابراین، e استنادهای مازاد است که توسط شاخص h اعتبار داده‌نشده‌است.

±استناد‌ها برای نشان دادن تأثیرات بر نویسنده‌ای با تعداد انتشارات بالا و استناد‌هایی که شامل برخی از انتشارات با استنادهای کمتر از h است، توزیع می‌شوند.

**بحث**

مشخص است (جدول 1) که بسیاری از توزیع‌های استنادی می‌توانند شاخص h و تعداد انتشارات یکسانی داشته باشند اما دارای شاخص‌های hb منحصربه‌فرد باشند. در جدول 1 ردیف اول دارای کمترین استناد (h) برای شاخص h، 10 است و استنادهای «مازاد» (e) آن، صفر است. همه استنادها تاثیر می‌گذارند و شاخص h و شاخص‌ hb یکسان هستند. ( شاخص‌ hb برابر با $10+ \sqrt{0}10+ \sqrt{0}$ است.)

در جدول 1 ردیف دوم شاخص h آن نیز برابر با 10 است ولی شاخص‌ hb آن برابر با $10+ \sqrt{190-100}10+ \sqrt{190-100}$ است.

بنابراین، شاخص hb، نسبت به h، تحت تاثیر 90 استناد مازاد که بر روی شاخص h تاثیری نداشت، افزایش می‌یابد. جدول 1 همچنین نشان می‌دهد که شش نویسنده، هر کدام با شاخص‌های h یکسان 10 و دقیقاً 10 مقاله، اما توزیع‌های منحصر به فرد استنادها (A-F)، دارای شاخص‌های hb هستند که از 10 تا 31.21 متغیر است. اساساً، شاخص h هرگز نمی‌تواند بزرگتر از تعداد کل انتشارات باشد، اما می تواند به سمت پایین تا صفر باشد. شاخص h دقیقاً برابر است با تعداد کل مقالاتی که ≥h بار استناد شده‌اند. به طور خاص، همانطور که توسط توزیع‌های استنادی A-G (جدول 1) نشان داده‌شده‌است، هنگامی که هر مقاله حداقل 10 بار استناد شده‌است، شاخص h آن 10 می شود. صرف نظر از استنادهای مازاد، شاخص h تا زمانی که مقاله دیگری منتشر شود روی 10 باقی می‌ماند و هر یک از یازده مقاله کنونی حداقل 11 بار استناد می‌شود. این نقل‌قول‌های «مازاد» اضافی توسط شاخص hb به عنوان تابع $\sqrt{e}\sqrt{e}$ معلوم و اعتبار داده می‌شوند جایی که e مقدار استنادهای «مازاد» است.

قرار دادن $ \sqrt{e}\sqrt{e}$باعث افزایش ملایمی در شاخص hb می‌شود.

برای مقایسه بیشتر شاخص h با شاخص hb، یک مورد اضافی را در نظر بگیرید (نه در جدول 1) با استنادهایی که به طور یکسان در توزیع B (جدول 1) توزیع شده اند با این تفاوت که همه 90 استناد مازاد بین مقالاتی که کمتر از شاخص h استناد شده‌اند، توزیع شده‌اند (در این مورد 10). این لزوماً تعداد کل مقالات را به بیش از 10 افزایش می‌دهد، اما این استنادها هیچ کمکی به شاخص h نمی‌کنند. برای ارزیابی تأثیرات روی شاخص hb برای این مورد، استناد‌ها باید به طور مشخص اختصاص داده‌شوند. فرض کنید 10 مقاله هر کدام با 10 استناد (برای ایجاد h = 10) و هر مقاله با تعداد استناد ذیل: 9، 8، 7، 6، 5، 4، 3، 2، و 1 است. اکنون 19 مقاله با تعداد کل استناد 145 است. شاخص h، 10 باقی می‌ماند اما شاخص hb اکنون $10+ \sqrt{45}10+ \sqrt{45}$ است که برابر با 16.71 است و به 9 مقاله ای که در مجموع 45 استناد دریافت کرده‌اند اعتبار می‌دهد. به عنوان مثال، اگر 10 مقاله وجود داشت که هر کدام 10 بار استناد شده‌بود و یک مقاله 9 بار "مازاد" استناد شده‌بود، شاخص h البته 10خواهدبود، اما شاخص hb $10+ \sqrt{9}=1310+ \sqrt{9 }=13$ است. به همین ترتیب، برای 1 استناد «مازاد»، شاخص hb برابر با $10+ \sqrt{1}=1110+\sqrt{1}=11$ خواهدبود. نتایج موارد دیگر به راحتی قابل محاسبه است.

می توان استدلال کرد که عدم موفقیت شاخص h در اعتبار مقاله‌ای با استنادهای زیر h توجه کمی دارد. با این حال، محققان در اوایل کار خود به احتمال زیاد چنین استنادهای مازاد را از مقالاتی که هنوز زیاد استناد نشده‌اند، داشته‌باشند. در واقع، شاخص hb این موضوع را تشخیص می‌دهد و بر اساس سهم نسبتاً کم $\sqrt{e}\sqrt{e}$ (اما برای تبعیض بین نویسندگان) توسط مقالات کم استناد، اعتبار مناسب را می‌دهد. با این حال، وقتی مقالاتی وجود دارد که استنادهای زیادی بیش از h دریافت می‌کنند (جدول 1 توزیع‌های B–H؛ برای جزئیات بیشتر به پاورقی مراجعه کنید)، این مقالات به‌طور مناسبی توسط شاخص hb که به مقادیر نسبتاً بزرگی افزایش می‌یابد، اعتبار می‌یابند، زیرا مقدار $\sqrt{e}\sqrt{e}$ نسبتاً بزرگ است.

**نتیجه‌گیری**

شاخص h با محاسبه شاخص hb که شاخص h به علاوه مقدار $(\sqrt{e})(\sqrt{e})$است با e تعریف‌شده به عنوان استنادهای مازاد (استنادهای بیش از حداقل تعداد کل (h) که می‌تواند شاخص h را تعریف کند) اصلاح می‌یابد. شاخص hb به استنادهایی اعتبار می‌دهد که در غیر این صورت وزنی ندارند، بین نویسندگانی که دارای شاخص h یکسانی هستند، فرق می‌گذارند و به ویژه برای تمایز بین سوابق انتشارات محققان جوان با شاخص‌های h متوسط اما مقالات قابل توجهی که در حال رسیدن هستند اما هنوز بسیار مورد استناد قرار نگرفته‌اند، مفید است.

1. . hb-index [↑](#footnote-ref-1)
2. . [Olen R. Brown](https://pubmed.ncbi.nlm.nih.gov/?term=Brown%20OR%5BAuthor%5D) [↑](#footnote-ref-2)
3. . Thompson Scientific Web of Science [↑](#footnote-ref-3)
4. . Google Scholar Citations [↑](#footnote-ref-4)