

학술지의 피인용횟수 순위를 적용한 tapered h -지수의 변형지표 “ $Kor-h_T$ ”에 관한 연구*

A Study on the “ $Kor-h_T$ ”, a Modified Tapered h -index, by Applying the Ranking According to the Number of Citations of Journals in Evaluating Korean Journals

고영만 (Young Man Ko)**

조수련 (Soo-Ryun Cho)***

박지영 (Ji Young Park)****

초 록

본 연구에서는 ‘학술지의 피인용횟수 순위’를 적용한 tapered h -지수의 변형지표 $Kor-h_T$ 를 고안하여 제안하였다. $Kor-h_T$ 의 의미를 평가하기 위하여 $Kor-h_T$ 지수 값의 중복률 및 지수 값과 평가요소 사이의 연관성 변화를 다른 학술지 평가지표인 h -지수, tapered h -지수 및 IF 와 비교 분석하였다. 지수 값의 중복률 분석은 지표의 변별력을 살펴보기 위한 것이며, 지수 값과 평가요소와의 상관관계 분석은 지표의 평가요소인 피인용횟수 및 논문 수가 지수 값에 각각 어떻게 반영되는지를 알아보기 위한 것이다. 분석을 위해 2008년부터 2010년까지의 한국학술지인용색인(KCI) 3개년 데이터를 사용하였다. 분석 결과 본 연구에서 제안한 $Kor-h_T$ 가 비교 대상 지표에 비해 지수 값의 중복률이 가장 낮아 높은 변별력을 보였으며, 지수 값과 평가요소와의 상관관계에 있어서도 피인용횟수와 논문 수 모두에서 상관관계가 가장 높은 것으로 나타났다.

ABSTRACT

This study describes the meaning of and the formula for $Kor-h_T$, which is a modified index built on the tapered h -index by applying ‘the ranking according to the number of citations of journals’. This study evaluated the de-duplication rate of index values of $Kor-h_T$ and analyzed the change in the correlation between the index values and evaluation elements using the Korea Citation Index data from 2008 to 2010. $Kor-h_T$ is compared with h -index, tapered h -index, and IF . As a result, $Kor-h_T$ appeared to be superior to other indexes on de-duplication rate. It is also shown that there is a very strong positive correlation between the evaluation elements, the number of citations and the number of articles of journals, and the index values of $Kor-h_T$.

키워드: 한국학술지인용색인, tapered h -index, $Kor-h_T$, Kor -Factor, 피인용횟수 순위, 중복제거 KCI, tapered h -index, $Kor-h_T$, Kor -Factor, citation rate ranking, de-duplication

* 이 논문은 한국연구재단의 정책연구용역과제(정책연구-2011-002-학술진흥)의 수행 결과를 토대로 작성되었으며, 한국연구재단의 공식적인 견해와는 다를 수 있습니다.

** 성균관대학교 문과대학 문헌정보학과 교수(ymko@skku.edu) (제1저자)

*** 성균관대학교 부설 정보관리연구소 연구원(chosr@skku.edu) (공동저자)

**** 성균관대학교 부설 정보관리연구소 연구원(bjy1532@skku.edu) (공동저자)

■ 논문접수일자: 2013년 11월 18일 ■ 최초심사일자: 2013년 12월 4일 ■ 게재확정일자: 2013년 12월 9일

■ 정보관리학회지, 30(4), 111-131, 2013. [http://dx.doi.org/10.3743/KOSIM.2013.30.4.111]

1. 서론

학술지를 평가하는 데 있어서 초기에 사용된 대표적인 계량적 지표는 학술지의 총 피인용횟수 또는 총 논문 수였다. 총 논문 수는 학술지의 양적 생산성을, 총 피인용횟수는 학술지의 학술적 영향력을 평가하기 위해 사용되었다. 특히 질 높은 학술지가 인용이 많이 될 것이라는 전제를 바탕으로 총 피인용횟수를 기준으로 학술지의 질을 평가하였다. 그러나 총 논문 수 또는 총 피인용횟수를 기준으로 학술지를 평가할 경우 규모가 큰 학술지에 유리할 수 있다. 이러한 우려에서 규모는 작지만 질적으로 우수한 학술지가 제대로 평가받을 수 있도록 하기 위하여 고안된 것이 Impact Factor(이하 IF)이다(Garfield, 2006).

IF 가 피인용횟수를 기반으로 학술지의 영향력을 평가하는 데 있어서 객관적이고 효과적인 지표 중의 하나임에는 분명하다. 그러나 데이터 베이스의 유형 및 규모의 영향을 많이 받기 때문에 피인용 패턴이나 장서량이 다른 주제 분야를 비교할 때는 IF 를 그대로 적용해서는 안 된다는 것에 주의해야 한다(Ko, Cho, & Park, 2011). 특히 Web of Science(WoS), Scopus 등의 해외 인용색인과 비교하면 소규모이며, 등록 학술지 게재 논문의 2/3 이상이 발행 후 2년 동안 1회 이하로 인용될 정도로 피인용빈도가 작게 나타나고 있는 한국학술지인용색인(Korean Citation Index, 이하 KCI)에 IF 를 적용할 경우 영향력이 실제보다 낮게 나타날 수 있어 KCI에 수록된 학술지를 제대로 평가하기 어려울 수 있다.

이와 관련하여 피인용횟수뿐만 아니라 게재

논문 수를 함께 고려하여 학술지 평가에 적용하는 방안으로 h -지수를 비롯하여 다양한 Hirsch 유형 지수들이 제안되었다. h -지수는 원래 Hirsch(2005)가 개별 연구자의 연구업적 평가를 위해 개발한 지표이며, h -지수를 학술지 평가에 적용하기 위한 방안은 Braun, Glänzel, Schubert(2006)에 의해 제안되었다. h -지수가 제안된 이래 이른바 Hirsch 유형 지수로 불리는 g -지수(Egghe, 2006), A -지수(Jin, 2006), R -지수(Jin, Liang, Rousseau, & Egghe, 2007), AR -지수(Jin, 2007), tapered h -지수(Anderson, Hankin, & Killworth, 2008), f -지수(Tol, 2009), t -지수(Tol, 2009), hg -지수(Alonso, Cabrerizo, Herrera-Viedma, & Herrera, 2010) 등이 발표되었다.

국내에서도 학술지 평가를 위한 다양한 Hirsch 유형 지수들이 제안되었다. 대표적으로 학술지 호 단위로 산출된 h -지수를 합산하여 얻어지는 호 기반 h -지수, 논문별 피인용횟수의 제곱근 값을 활용한 h_s -지수 및 f_s -지수(김관준, 이재운, 2010), 논문 수를 이용하여 h -지수의 표준화를 시도한 $h-j$ 지수(조은성, 송재도, 2011) 등이 있다. 이 지수들은 모두 h -지수에 비해 변별력이 향상된 것으로 나타났으나, 피인용횟수가 h -지수 값 이상인 논문 h 편, 즉 h -core만을 평가의 대상으로 하는 한계를 안고 있다. 특히 h_s -지수와 f_s -지수는 변별력을 향상시키기 위한 목적에서 뚜렷한 인과 관계를 밝히지 않은 채 피인용횟수의 제곱근을 취하고 있다.

KCI와 같이 데이터베이스의 규모가 작고 수록된 학술지의 피인용빈도가 낮은 경우에는 학술지 내 모든 논문의 피인용횟수를 바탕으로 학술지를 평가하는 것이 좋은 방법일 수 있다. 이

점에 있어서 점감(漸減) h -지수(이재윤, 정은경, 유소영, 2012)로 번역되는 tapered h -지수는 h -지수를 비롯한 여타의 Hirsch 유형 지수와 차이를 보인다. tapered h -지수는 h -core를 중심으로 학술지를 평가하는 대신 학술지 내 모든 논문을 대상으로 하여 학술지를 평가하며, 피인용횟수 기준 논문 순위와 해당 논문의 피인용횟수를 바탕으로 새로운 값을 생성하여 지수 값을 산출한다.

따라서 본 연구에서는 피인용횟수가 아주 많거나 적은 논문도 평가요소에 반영될 뿐 아니라 변별력의 문제도 개선할 수 있다는 점에서 tapered h -지수를 면밀히 검토하였다. 그리고 학술지 게재 논문 순위와 피인용횟수를 바탕으로 새로운 값을 생성하여 그 값들을 합산하는 방식으로 학술지를 평가하는 지표로서 tapered h -지수와 유사한 방식의 $Kor-h_T$ 를 고안하였다. 또한 KCI에 $Kor-h_T$ 의 적용가능성을 평가하기 위하여 지수 값의 중복률 및 지수 값과 평가요소인 피인용횟수와 논문 수와의 상관관계를 분석하여 h -지수, tapered h -지수, IF 등과 비교하였다.

본 연구에 사용된 데이터는 2010년 KCI 수록 학술지에 의해 인용된 2008, 2009년 학술지 데이터이다. $Kor-h_T$ 를 개발하고 검증하기 위하여 3년 연속 등록된 학술지 1,861종 중 총 피인용횟수가 0인 학술지 28종, 논문별 피인용횟수 누락 학술지 14종, 피인용횟수 오류 학술지 25종을 제외한 1,794종이 분석되었다.

2. KCI 데이터 분석

2.1 KCI 데이터 일반 현황

2.1.1 2008~2010년 대분류별 KCI 수록

학술지, 논문, 저자, 참고문헌 수 현황
한국연구재단의 KCI 데이터베이스는 인문학, 사회과학, 자연과학, 공학, 의학약학, 농수해양, 예술체육, 복합학의 8개 주제 분야로 나누어져 관리되고 있다. 2008년부터 2010년까지 논문 수, 학술지 수, 참고문헌 수 모두 매년 증가하고 있는 것으로 나타났으며, 특히 학술지 수는 매년 비슷한 양이 증가하고 있고 참고문헌 수는 각 주제 분야에서 고르게 증가하고 있는 것으로 나타났다(〈표 1〉 참조).

2.1.2 KCI 수록 학술지의 발간횟수, 논문 수, 참고문헌 수, 피인용횟수 현황

연도별로 학술지 현황을 살펴보면 평균발간횟수, 평균논문수, 평균참고문헌수는 꾸준한 증가 추세를 보였다. 발간횟수가 가장 높은 분야는 공학이었으며, 평균논문수 역시 공학 분야에서 가장 높게 나타났다. 전체적으로 평균피인용횟수는 2008년 45.10에서 2009년 40.53으로 감소한 것으로 나타났으며, 평균자기인용횟수는 2009년 17.75에서 2010년 15.71로 감소한 것으로 나타났다. 여기서 평균피인용횟수, 평균자기인용횟수는 해당 년도에 이전 2년 논문을 인용한 횟수에 기초한 것이다. 특히 평균피인용횟수는 2008년 농수해양이 76.21, 2009년 자연과학이 69.65, 2010년 농수해양이 67.77로 높게 나타났다. 평균자기인용횟수의 경우, 매년 농수해양, 자연과학, 공학이 높게 나타났다(〈표 2〉 참조).

2.1.3 KCI 수록 학술지의 논문수, 공저자수, 평균피인용횟수, 평균참고문헌수 현황
 논문 수는 사회과학 분야가 2008년 20,934, 2009년 22,812, 2010년 21,436으로 가장 많은 것으로 나타났으며, 평균공저자수는 의약학 분야가 2008년 4.94명, 2009년 4.76명, 2010년 4.68명으로 가장 많은 것으로 나타났다. 또한

평균피인용횟수, 평균참고문헌수는 전반적으로 꾸준한 증가 추세를 보였다. 평균피인용횟수는 농수해양 분야가 2008년 2.45, 2009년 4.63, 2010년 5.77로 가장 크며 평균참고문헌수는 의약학 분야가 2008년 34.61, 2009년 38.22, 2010년 39.00으로 가장 큰 것으로 나타났다(<표 3> 참조).

<표 1> KCI 수록 학술지, 논문, 저자, 참고문헌 수 현황(2008~2010)

연도	대분류	학술지수	논문수	저자수	참고문헌수
2008	인문학	470	14,539	15,569	322,113
	사회과학	605	19,266	31,303	575,140
	자연과학	104	8,723	30,081	169,222
	공학	233	17,691	55,910	247,958
	의약학	206	12,362	59,888	276,409
	농수해양	73	4,562	20,880	103,712
	예술체육	100	4,771	8,404	113,489
	복합학	54	1,831	3,208	42,288
	전체	1,845	83,745	225,243	1,850,331
2009	인문학	503	15,603	16,956	380,171
	사회과학	649	21,726	35,694	722,476
	자연과학	104	8,927	30,664	215,575
	공학	236	19,275	61,738	290,901
	의약학	216	13,007	62,872	305,203
	농수해양	74	4,963	22,938	122,165
	예술체육	104	5,229	9,612	132,544
	복합학	59	2,303	4,242	53,603
	전체	1,945	91,033	244,716	2,222,638
2010	인문학	526	16,591	18,173	426,850
	사회과학	696	23,130	38,614	781,451
	자연과학	109	8,901	30,902	218,845
	공학	242	20,164	63,416	305,824
	의약학	228	13,305	64,915	311,482
	농수해양	73	5,266	24,439	130,263
	예술체육	113	5,534	10,004	141,307
	복합학	67	2,632	5,076	62,562
	전체	2,054	95,523	255,539	2,378,584

〈표 2〉 KCI 수록 학술지의 발간, 논문, 참고문헌, 피인용 현황(2008~2010)

연도	대분류	학술지수	평균 발간횟수	평균 논문수	평균 참고문헌수	평균 피인용횟수	평균 자기인용횟수
2008	인문학	470	2.20	44.70	559.88	20.94	4.19
	사회과학	605	2.65	45.21	832.86	60.90	11.43
	자연과학	104	4.51	124.09	1,238.72	72.05	38.17
	공학	233	5.16	125.02	1,031.92	32.40	20.88
	의약학	206	4.14	114.40	1,235.45	32.08	17.76
	농수해양	73	4.40	117.29	1,419.36	76.21	41.80
	예술체육	100	2.21	59.93	772.69	39.28	11.27
	복합학	54	2.49	39.63	557.08	26.92	8.15
전체	1,845	3.47	83.78	956.00	45.10	19.21	
2009	인문학	503	2.37	51.76	664.05	17.73	4.03
	사회과학	649	2.93	51.47	1,026.11	45.96	9.50
	자연과학	104	4.82	140.89	1,718.37	69.65	43.65
	공학	236	5.44	134.34	1,287.68	49.72	30.72
	의약학	216	4.35	119.30	1,371.08	27.62	14.01
	농수해양	74	4.55	121.73	1,625.41	65.40	27.85
	예술체육	104	2.48	66.40	886.17	26.19	7.23
	복합학	59	2.70	43.75	658.61	21.96	4.99
전체	1,945	3.71	91.20	1,154.68	40.53	17.75	
2010	인문학	526	2.51	54.96	742.58	21.26	4.57
	사회과학	696	3.16	57.79	1,125.65	52.83	9.53
	자연과학	109	4.91	144.76	1,752.51	59.01	29.93
	공학	242	5.60	154.31	1,315.77	49.82	27.04
	의약학	228	4.57	124.91	1,364.17	30.71	14.38
	농수해양	73	4.62	125.86	1,346.10	67.77	25.23
	예술체육	113	2.63	66.78	1,030.93	34.33	8.93
	복합학	67	3.05	49.71	798.86	26.94	6.04
전체	2,054	3.88	97.39	1,184.57	42.83	15.71	

〈표 3〉 KCI 수록 학술지의 논문, 공저자, 평균피인용, 평균참고문헌 현황(2008~2010)

연도	대분류	논문수	평균공저자수	평균피인용횟수	평균참고문헌수
2008	인문학	11,614	1.08	0.78	27.73
	사회과학	20,934	1.63	1.26	27.47
	자연과학	6,985	3.38	1.99	24.23
	공학	10,990	3.22	2.04	22.56
	의약학	7,986	4.94	1.87	34.61
	농수해양	4,244	4.35	2.45	24.44
	예술체육	3,792	1.84	1.32	29.93
	복합학	1,494	1.72	0.95	28.31
전체	68,039	2.77	1.58	27.41	

연도	대분류	논문수	평균공저자수	평균피인용횟수	평균참고문헌수
2009	인문학	13,265	1.08	1.04	32.73
	사회과학	22,812	1.65	2.23	34.51
	자연과학	7,684	3.41	3.74	30.86
	공학	12,508	3.27	2.90	26.47
	의약학	8,384	4.76	3.86	38.22
	농수해양	4,482	4.51	4.63	28.79
	예술체육	4,213	1.86	2.09	34.95
	복합학	1,717	1.77	1.68	35.88
	전체	75,065	2.79	2.77	32.80
2010	인문학	12,724	1.08	1.34	36.75
	사회과학	21,436	1.67	2.77	37.33
	자연과학	6,602	3.33	3.60	31.33
	공학	11,791	3.31	3.27	27.83
	의약학	7,071	4.68	5.07	39.00
	농수해양	3,956	4.66	5.77	30.69
	예술체육	4,048	1.88	3.06	37.26
	복합학	1,729	1.85	2.36	41.88
	전체	69,357	2.81	3.40	35.26

2.2 KCI 수록 학술지의 피인용 특징

2.2.1 연구 분야 대분류별 학술지 수, 논문 수 및 피인용 현황

2008, 2009 2개년 간 발행된 총 170,358편의 논문 중 1회 이상 인용된 논문은 51,804편, 비율로는 약 30.4%인 것으로 나타나 전반적으로 인용률이 매우 낮은 것으로 분석되었다. 사회과학 분야가 43.3%로 가장 높았으며, 의약학 분야가 20.7%로 가장 낮은 비율을 보였다(〈표 4〉 참조).

2.2.2 연구 분야 대분류별 1회 이상 피인용 논문 수 및 비율

〈표 5〉는 1회 이상 인용된 논문 51,804편을 피인용 횟수에 따라 대분류별로 정리한 것으로,

전체 피인용논문수의 61.3%가 1회만 인용된 것으로 나타났다. 특히 1회만 인용된 논문수는 공학 분야가 72.9%로 가장 높았으며, 사회과학 분야가 53.1%로 가장 낮게 나타났다.

2.2.3 연구 분야 대분류별 1회 이상 피인용, 자기피인용 횟수 및 비율

〈표 6〉은 총 피인용횟수 90,249회 중 자기피인용횟수가 24,358회로 약 27.0%에 이르고 있음을 보여준다. 2008, 2009년도 2개년 간 KCI에 등록된 학술지의 논문 중 2010년 동안 한 번이라도 인용된 논문의 비율은 매우 낮은 것으로 나타났다(의약학 20.7%, 사회과학 43.3%, ..., 전체 30.4%). 또한 대부분의 논문(사회과학 53.1%, 공학 72.9%)에서 피인용횟수가 1회에 머물고 있으며, 자기피인용횟수 비율 또한 높아 공학

〈표 4〉 2010 분석 대상 학술지의 대분류별 논문 수

대분류	학술지 수	게재 논문 수			2010년 한 번이라도 인용된 논문 수		
		2008년 발행	2009년 발행	합계	2008년 발행 (비율)	2009년 발행 (비율)	합계 (비율)
인문학	459	14,511	14,978	29,489	4,235 (29.2%)	5,130 (34.3%)	9,365 (31.8%)
사회과학	592	19,047	20,759	39,806	8,506 (44.7%)	8,735 (42.1%)	17,241 (43.3%)
자연과학	99	8,391	8,618	17,009	2,244 (26.7%)	2,246 (26.1%)	4,490 (26.4%)
공 학	227	17,465	18,924	36,389	3,867 (22.1%)	4,334 (22.9%)	8,201 (22.5%)
의약학	201	12,313	12,718	25,031	2,692 (21.9%)	2,492 (19.6%)	5,184 (20.7%)
농수해양	71	4,442	4,824	9,266	1,407 (31.7%)	1,469 (30.5%)	2,876 (31.0%)
예술체육	97	4,591	4,959	9,550	1,551 (33.8%)	1,647 (33.2%)	3,198 (33.5%)
복합학	48	1,763	2,055	3,818	571 (32.4%)	678 (33.0%)	1,249 (32.7%)
전 체	1,794	82,523	87,835	170,358	25,073 (30.4%)	26,731 (30.4%)	51,804 (30.4%)

〈표 5〉 2010 분석 대상 학술지의 한 번이라도 인용된 논문 수 및 비율(대분류별)

대분류	x회 인용된 논문 수(비율 %)									합계
	1회	2회	3회	4회	5회	6~9	10~14	15~19	20~	
인문학 (459종)	5,728 (61.2%)	2,197 (23.5%)	840 (9.0%)	336 (3.6%)	138 (1.5%)	116 (1.2%)	9 (0.1%)	0	1 (0.0%)	9,365
사회과학 (592종)	9,163 (53.1%)	4,087 (23.7%)	1,895 (11.0%)	935 (5.4%)	504 (2.9%)	550 (3.2%)	95 (0.6%)	11 (0.1%)	1 (0.0%)	17,241
자연과학 (99종)	2,833 (63.1%)	919 (20.5%)	399 (8.9%)	174 (3.9%)	71 (1.6%)	72 (1.6%)	20 (0.4%)	2 (0.0%)	0	4,490
공 학 (227종)	5,977 (72.9%)	1,512 (18.4%)	471 (5.7%)	146 (1.8%)	59 (0.7%)	32 (0.4%)	4 (0.0%)	0	0	8,201
의약학 (201종)	3,632 (70.1%)	1,005 (19.4%)	312 (6.0%)	131 (2.5%)	43 (0.8%)	54 (1.0%)	5 (0.1%)	2 (0.0%)	0	5,184
농수해양 (71종)	1,810 (62.9%)	609 (21.2%)	252 (8.8%)	114 (4.0%)	51 (1.8%)	31 (1.1%)	8 (0.3%)	1 (0.0%)	0	2,876
예술체육 (97종)	1,847 (57.8%)	672 (21.0%)	305 (9.5%)	154 (4.8%)	76 (2.4%)	109 (3.4%)	26 (0.8%)	9 (0.3%)	0	3,198
복합학 (48종)	756 (60.5%)	298 (23.9%)	115 (9.2%)	49 (3.9%)	24 (1.9%)	7 (0.6%)	0	0	0	1,249
전 체 (1,794종)	31,746 (61.3%)	11,299 (21.8%)	4,589 (8.9%)	2,039 (3.9%)	966 (1.9%)	971 (1.9%)	167 (0.3%)	25 (0.0%)	2 (0.0%)	51,804

〈표 6〉 2010 분석 대상 학술지의 대분류별 피인용횟수

대분류	학술지 수	2010년 피인용횟수			2010년 자기피인용횟수		
		2008년 발행	2009년 발행	합계	2008년 발행 (비율)	2009년 발행 (비율)	합계 (비율)
인문학	459	6,772	8,835	15,607	1,078 (15.9%)	1,439 (16.3%)	2,517 (16.1%)
사회과학	592	17,768	16,583	34,351	2,793 (15.7%)	3,327 (20.1%)	6,120 (17.8%)
자연과학	99	3,974	3,735	7,709	1,514 (38.1%)	1,649 (44.1%)	3,163 (41.0%)
공 학	227	5,496	6,130	11,626	2,521 (45.9%)	3,285 (53.6%)	5,806 (49.9%)
의약학	201	4,157	3,616	7,773	1,480 (35.6%)	1,448 (40.0%)	2,928 (37.7%)
농수해양	71	2,402	2,427	4,829	831 (34.6%)	1,038 (42.8%)	1,869 (38.7%)
예술체육	97	3,169	3,121	6,290	688 (21.7%)	816 (26.1%)	1,504 (23.9%)
복합학	48	935	1,129	2,064	165 (17.6%)	286 (25.3%)	451 (21.9%)
전 체	1,794	44,673	45,576	90,249	11,070 (24.8%)	13,288 (29.2%)	24,358 (27.0%)

의 경우 49.9%, 자연과학은 41.0%에 이르는 것으로 나타났다. 이는 피인용횟수만을 기반으로 하는 단일지표가 KCI 기반의 학술지 평가지표로 적합하지 않음을 나타내며 자기피인용횟수에 대한 통제의 필요성을 시사하는 것이다.

3. KCI 기반의 “*Kor-h_T*” 개발

3.1 KCI 기반 복합평가지표

Kor-Factor1, *Kor-Factor2*, F_{10}

한국연구재단에서는 해외 기관의 평가지표가 지니는 문제점을 극복하고 국내 연구 환경에 적합한 학술지 평가의 계량적 기준을 마련하기

위하여 2006년부터 재단의 등재(후보)지의 서지 정보와 인용정보를 제공하는 KCI 데이터베이스를 구축하기 시작하였다. 그리고 2007년에는 KCI 데이터베이스를 기반으로 국내 학술지의 질을 계량적으로 평가하기 위한 지표로 *Kor-Factor*의 첫 번째 알고리즘(이하 *Kor-Factor1*)이 홍종선과 조수련(2007)에 의해 개발되었다.

*Kor-Factor1*은 지수 값들이 정규분포를 이루도록 하기 위하여 의도적으로 총 논문 수에 자연로그를 적용한 점, 데이터가 변화하면 지수 값의 분포가 달라져 정규분포를 이루지 않을 수 있다는 점, 측정기간이 길어질수록 총 논문 수를 통제하기 어려워지는 점 등이 문제점으로 지적되었다. 이에 따라 Ko, Cho, Park(2011)은 자연로그를 적용하지 않고 측정기간의 변화에

$$Kor-Factor1 = \frac{(\text{피인용횟수} - \text{자기피인용횟수}/2) \text{의 평균}}{(\text{피인용횟수} - \text{자기피인용횟수}/2) \text{의 표준편차}} + \ln(\text{총 논문수})$$

$$Kor-Factor2 = \frac{(\text{피인용횟수} - \text{자기피인용횟수}/2) \text{의 평균}}{(\text{피인용횟수}) \text{의 표준편차}} \times (\text{연평균 논문수})$$

$$F_{10} = \frac{(\text{피인용횟수} - \text{자기피인용횟수}/2) \text{의 평균}}{(\text{피인용횟수}) \text{의 표준편차}} \times (\text{총 논문수의 십분위수})$$

도 총 논문 수를 통제할 수 있도록 $Kor-Factor1$ 을 개선한 이른바 $Kor-Factor2$ 를 제안하였다.

$Kor-Factor2$ 가 학술지 논문 수를 통제하면서 지수 값과 평가요소 사이의 연관성 문제를 개선하였으나 여전히 지수 값과 질적 평가요소 사이의 연관성에 비해 지수 값과 양적 평가요소 사이의 연관성이 높은 것으로 평가되었다(조은성, 송재도, 2011). 지수 값 전체에서 논문 수가 차지하는 비중도 피인용횟수의 평균이나 표준편차에 비해 훨씬 큰 편으로 나타났다. 이러한 문제점을 개선하기 위하여 고영만과 박지영(2012)은 ‘연평균 논문 수’ 대신 ‘총 논문 수의 십분위수’를 적용한 평가지표 F_{10} 을 제안하였다. 이때 십분위수는 측정기간 동안 각 학술지에 게재된 논문 수를 크기에 따라 10개의 구간으로 나누었을 때 특정 학술지의 논문 수가 속하는 구간의 경계값을 의미한다.

3.2 tapered h -지수의 변형지표 “ $Kor-h_T$ ”

3.2.1 h -지수와 tapered h -지수

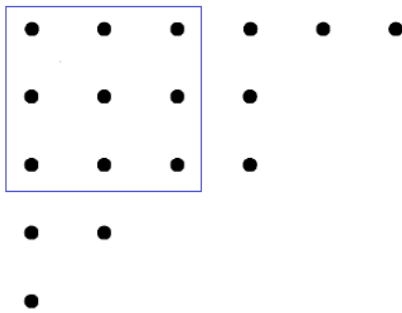
$Kor-Factor$ 가 지니는 또 하나의 문제는 복

합지표로 구성되어 지표의 해석과 지표 값의 산출이 쉽지 않다는 점이다. 이에 비해 h -지수는 산출 방법이 단순하며 지수 값에 대한 직관적인 해석이 가능하다. 예를 들어, 학술지 h -지수 값이 5라는 것은 학술지 게재 논문 중 피인용횟수가 5 이상인 논문이 최소 5편이라는 것을 의미한다.

h -지수의 주요한 특징은 피인용횟수가 매우 크거나 작은 몇몇 논문의 영향을 받지 않는다는 것이다. 이것은 지수의 강건성(robustness)이라는 장점으로 작용하기도 하지만 한편으로는 피인용횟수가 큰 논문들에 대한 추가 인용이 발생해도 지수 값에 반영되지 않는다는 점에서 단점으로 지적되기도 한다(Egghe, 2006). 따라서 동일한 값이 빈번하여 변별력이 낮은 문제점이 있으며(Batista, Campiteli, Konouchi, & Martinez, 2006), 규모가 작은 학술지에 불리하게 작용할 수 있다(Braun, Glänzel, & Schubert, 2006).

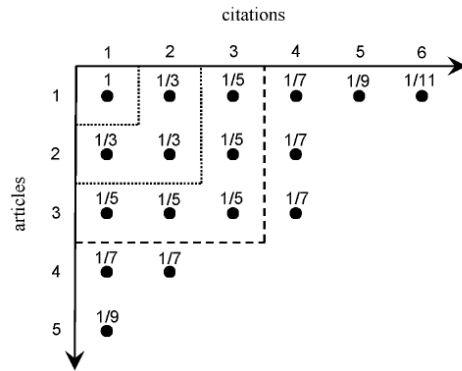
h -core를 벗어난 논문들이 반영되지 않는 h -지수의 단점을 해결하기 위한 지표로 제안된 tapered h -지수는 정수론(Number Theory)의 ‘Durfee square’ 개념을 h -지수에 적용한 것이

다. 'Durfee square'는 어떤 자연수를 그보다 작은 자연수로 분할하였을 때 나타나는 가장 큰 정사각형을 말한다. 예를 들어, $17=6+4+4+2+1$ 로 표현하면 아래 <그림 1>과 같이 나타나는 가장 큰 정사각형은 한 변의 길이가 3인 정사각형이며, 이 정사각형이 'Durfee square'이다. 'Durfee square'의 한 변의 길이 3은 발표한 논문의 피인용횟수가 각각 6, 4, 4, 2, 1인 학술지의 h -지수와 동일하다(h -지수 = 3).



<그림 1> Durfee square 예시

tapered h -지수는 h -지수가 <그림 1>의 정사각형 부분만 반영하는 점, 즉 피인용횟수가 매우 큰 논문이나 매우 작은 논문의 피인용횟수를 배제함으로써 전체 피인용횟수를 제대로 반영하지 못하는 점을 개선하고자 정사각형의 바깥 부분에 위치한 피인용횟수도 지수 값에 반영한다. 이때 논문 순위와 피인용횟수를 이용하여 새로운 값을 적용한다. 어떤 학술지의 발표 논문 5편의 피인용횟수가 각각 6, 4, 4, 2, 1이라고 하면 이 학술지의 총 피인용횟수 17에 대해 <그림 2>와 같이 새로운 값을 부여할 수 있으며 이 값들을 모두 더한 것이 학술지의 tapered h -지수 값이 된다.



<그림 2> tapered h -지수에서 논문별 피인용횟수 적용 방식 (Anderson, Hankin, & Killworth, 2008)

위 예에서 각 논문별 새로운 값을 살펴보면

1위 논문의 값 $h_{T(1)} = 1 + 1/3 + 1/5 + 1/7 + 1/9 + 1/11 \approx 1.87821$

2위 논문의 값 $h_{T(2)} = 1/3 + 1/3 + 1/5 + 1/7 \approx 1.00952$

3위 논문의 값 $h_{T(3)} = 1/5 + 1/5 + 1/5 + 1/7 \approx 0.74286$

4위 논문의 값 $h_{T(4)} = 1/7 + 1/7 \approx 0.28571$

5위 논문의 값 $h_{T(5)} = 1/9 \approx 0.11111$

$\therefore h_T = h_{T(1)} + h_{T(2)} + \dots + h_{T(5)} \approx 4.02742$

(h -지수는 3, g -지수는 4임)

피인용횟수 기준 순위가 r 인 논문의 피인용횟수를 c_r 라고 하면, 학술지에 대한 tapered h -지수 값, 즉 h_T 의 산출식은 다음과 같다.

$h_T = \sum_{r=1}^n h_{T(r)}$ (n : 논문 수)

$c_r \leq r$ 이면, 해당 논문의 값 $h_{T(r)} = \frac{c_r}{2r-1}$

$c_r > r$ 이면, 해당 논문의 값 $h_{T(r)} = \frac{r}{2r-1} + \sum_{i=r+1}^{c_r} \frac{1}{2i-1}$

3.2.2 KCI 기반의 tapered h-지수 변형지표:
“Kor-h_T”

국내 학술지의 질적 수준 평가를 위한 KCI 데이터베이스 기반의 학술지 평가지표는 피인용횟수이라는 단일 요소를 기반으로 하는 해외 기관의 학술지 평가지표가 지니고 있는 불합리성을 극복할 수 있어야 한다. 따라서 본 연구에서는 학술지 평가요소의 하나인 논문 수를 반영하기 위해 먼저 tapered h-지수(이하 h_T 지수)를 변형한 수정h_T를 고안하였다. h_T 지수에 피인용횟수가 없는 논문은 반영되지 않는 것에 비해, 수정h_T에서는 대분류별 최대 피인용논문수를 기준으로 일괄적용값을 부여함으로써 논문 수를 반영하고자 하였다. 여기서 ‘대분류별 최대 피인용논문수’는 8개 분야로 구분된 KCI의 각 대분류 내에서 측정기간 동안 한 번이라도 인용된 학술지의 최대 피인용논문수를 의미한다. 예를 들어, 대분류 A에서 측정기간 동안 한 번이라도 인용된 논문이 가장 많은 학술지가 B이고 그 수(최대 피인용논문수)가 45라면, 대분류 A에 속한 각 학술지에 적용되는 일괄적용값은

$$\frac{1}{2 \times (45 + 1) - 1}$$

이다. 이것은 피인용횟수가 0인 논문들에 동일하게 적용되는 값으로, 대분류 내에서 피인용횟수가 없는 논문들이 일괄적으로 같은 값을 가지면서 한 번이라도 피인용된 논문에 비해서는 낮은 값을 가지도록 하기 위한 것이다.

$$\begin{aligned} \text{수정}h_T &= (\text{측정기간의 } h_T \text{ 값}) + \\ &\quad (\text{'대분류별 최대 피인용논문수+1'의} \\ &\quad \text{tapered식 일괄적용값}) \times \\ &\quad (\text{피인용 0인 논문수}) \end{aligned}$$

그렇지만 인용이 되지 않는 논문이 다수를 차지하는 KCI 데이터베이스에서는 측정기간이 길어질수록 평가 대상 논문이 많아져 수정h_T에 피인용횟수의 반영률이 낮아지고 논문 수의 영향력이 지나치게 커질 수 있다. 이러한 문제점을 극복하기 위하여 본 연구에서는 측정기간을 1년 단위로 하여 원하는 기간만큼 반복 측정 후 각 측정치를 합산하여 평균을 취하는 지수로 h_T평균과 “Kor-h_T”를 고안하였다.

Kor-h_T: 1년 단위 측정 수정h_T 값의 합산 평균
(* 본 연구의 경우 2008년, 2009년 각 지수 값의 합산 평균)

4. “Kor-h_T” 검증

4.1 비교 대상 지표 간 순위상관계수

KCI 데이터를 토대로 h-지수, h_T 지수, 수정 h_T, Kor-h_T를 적용하여 각각의 변별력을 비교하고, 이들 Hirsch 유형 지수의 평가요소인 총 피인용횟수와 총 논문 수와의 상관관계를 비교 분석하여 Kor-h_T의 적합성을 검증하였다. 1년 치 값을 각각 계산한 후 합산된 값의 평균을 구하는 Kor-h_T의 변별력 효과를 보다 상세하게 검증하기 위하여, 1년 단위 측정 h_T 지수 값을 각각 계산한 후 합산 값의 평균을 구하는 h_T 평균을 추가하여 분석하였다. 이때 총 피인용횟수를 평가요소로 하는 Hirsch 유형 지수와는 달리 피인용횟수의 평균값을 구하는 IF와도 비교 분석하였다. 비교 대상 지표 간 스피어만 순위상관관계를 분석한 결과 비교 분석이 가능한 값으로 측정되었다(〈표 7〉 참조).

〈표 7〉 h_T 관련 지표 사이의 스피어만 순위상관계수 - 전체 1,794종

	h -지수	h_T 지수	h_T 평균	수정 h_T	$Kor-h_T$	IF
h -지수	1	.938**	.921**	.930**	.908**	.806**
h_T 지수		1	.995**	.996**	.987**	.821**
h_T 평균			1	.993**	.994**	.800**
수정 h_T				1	.994**	.785**
$Kor-h_T$					1	.756**
IF						1

** p<0.01, * p<0.05

4.2 대분류별 tapered식 일괄적용값

KCI 등록 학술지 게재 논문의 2/3 이상이 발행 후 2년 동안 1회 이하로 인용될 정도로 피인용빈도가 작게 나타나고 있는 현황은 피인용횟수만을 기반으로 하는 단일지표가 KCI 기반의 학술지 평가지표로 적합하지 않음을 시사한다(〈표 4〉, 〈표 5〉 참조). h_T 지수는 피인용횟수가 매우 큰 논문이나 매우 작은 논문의 피인용횟수를 배제함으로써 전체 피인용횟수를 제대로 반영하지 못하는 h -지수의 문제점을 개선하고자 제안된 것이다. h_T 지수와 마찬가지로 $Kor-h_T$

역시 h -지수의 문제점을 개선하고, 더불어 KCI 등록 학술지의 피인용 특징을 반영하기 위하여 대분류별 피인용횟수 0인 논문 1편에 대한 일괄적용값을 적용하여 지수 값을 산출하였으며, 그 값은 〈표 8〉과 같다([부록] 참조).

4.3 기술통계량 및 중복률

h_T 관련 학술지 평가지표의 기술통계량과 변별력 검증을 위해 각 지수 값의 빈도수와 중복률을 측정하였다. 지수 값의 빈도수가 많을수록 중복률은 낮아지며, 지수 값의 중복률이 낮을

〈표 8〉 $Kor-h_T$ 산출 시 대분류별 피인용횟수 0인 논문 1편에 대한 일괄적용값

대분류	2008 발행 논문		2009 발행 논문		2008-09 2년 발행 논문	
	최대 피인용 논문수	일괄적용값	최대 피인용 논문수	일괄적용값	최대 피인용 논문수	일괄적용값
인문학	45	0.01099	46	0.01075	81	0.00613
사회과학	86	0.00578	98	0.00508	182	0.00274
자연과학	177	0.00282	178	0.00280	316	0.00158
공학	190	0.00262	185	0.00270	375	0.00133
의약학	72	0.00690	62	0.00800	134	0.00372
농수해양	147	0.00339	153	0.00326	300	0.00166
예술체육	227	0.00220	234	0.00213	461	0.00108
복합학	128	0.00389	181	0.00275	309	0.00162

• 적용 시에는 반올림 없이 원래의 값을 그대로 적용함

경우 지표의 변별력이 높다고 할 수 있다. h_T 지수의 지수 값은 변별력 측면에서 h-지수에 비해 크게 개선되었으며 피인용횟수가 없는 논문에 일괄적용값을 적용한 Kor-h_T의 변별력은 더욱 개선되어 가장 우수한 것으로 나타났다. h-지수는 중복률이 99.5%에 달해 KCI 데이터 평가에 적합하지 않은 것으로 나타났다(〈표 9〉 참조).

4.4 지표 값과 평가요소 간 순위상관계수

〈표 10〉은 h_T 관련 지표 값들과 학술지의 평가요소인 총 피인용횟수 및 총 논문수 간의 순위상관계수를 산출한 결과이다. 스피어만 순위상관계 분석 결과 Kor-h_T 지표 값과 평가요소인 총 피인용횟수 및 논문 수와의 순위상관

관계가 다른 지표에 비해 모두 높은 것으로 나타났다.

Hirsch 유형 지수는 피인용횟수를 비교적 잘 반영하며 논문 수와의 상관관계는 약하다. 피인용논문이 모두 반영되는 h_T 지수와 수정h_T에서는 h-지수에 비해 논문 수의 반영률이 향상되었다. 인용되지 않은 논문에도 일괄값을 부여하는 수정h_T는 h_T 지수에 비해 논문 수와의 관계가 좀 더 높아진 것으로 나타났다. h_T 지수, h_T 평균, 수정h_T, Kor-h_T 모두 피인용횟수가 논문 수보다 지표 값과의 상관관계가 더 높으면서 두 평가요소 모두와 일정한 상관관계 이상인 것으로 나타나 양자를 잘 반영하고 있는 것으로 평가할 수 있다. 이 중에서도 Kor-h_T가 평가요소인 피인용횟수와 논문 수 측면 모두에서

〈표 9〉 h_T 관련 학술지 평가지표의 기술통계량 및 중복률

	N	최소값	최대값	평균	표준편차	빈도수(중복률)
h-지수	1794	1	9	2.523	1.240	9(99.5%)
h _T 지수	1794	1.00000	19.13554	4.603	2.325	1,237(31.0%)
h _T 평균	1794	0.50000	15.25061	3.494	1.840	1,418(21.0%)
수정h _T	1794	1.01192	19.40206	4.792	2.354	1,774(1.1%)
Kor-h _T	1794	0.51189	15.51684	3.673	1.876	1,792(0.1%)
IF	1794	0.00980	3.21186	0.554	0.457	1,293(27.9%)

〈표 10〉 h_T 관련 지표값과 평가요소와의 스피어만 순위상관계수 - 전체 1,794종

	총 피인용횟수	피인용 평균(IF)	논문수
h-지수	.880**	.806**	.403**
h _T 지수	.978**	.821**	.512**
h _T 평균	.987**	.800**	.548**
수정h _T	.981**	.785**	.556**
Kor-h _T	.989**	.756**	.599**
IF	.741**	1	.001

** p<0.01, * p<0.05

상관관계가 가장 높은 것으로 나타났다. 또한 $Kor-h_T$ 는 논문 수뿐만 아니라 피인용횟수의 평균인 IF 와도 일정 수준 이상의 상관관계가 있는 것을 확인할 수 있다.

5. 결론

연구 업적을 평가하는 목적과 상황이 항상 같을 수는 없다. WoS와 Scopus와 같은 국제적인 학술지인용색인에 비해 KCI는 국내 학술지를 대상으로 하는 소규모의 데이터베이스이고 학술지의 피인용도가 낮기 때문에 피인용횟수만을 기준으로 학술지를 평가하는 것에는 한계가 있을 수 있다. 더 많은 논문을 게재하는 학술지가 학술 커뮤니티에 더 많은 영향을 미칠 수 있다는 측면에서 학술지의 생산성을 반영하는 논문 수도 학술지 평가에 함께 고려해야 할 것이다.

본 연구에서는 2008년에서 2010년까지 3개년의 KCI 데이터를 기반으로 학술지의 피인용횟수와 논문 수를 함께 고려하는 평가지표로서 tapered h -지수(h_T 지수)를 변형한 $Kor-h_T$ 를 고안하였다. Hirsch 유형 지수의 하나인 h_T 지수는 h -core 논문만을 평가요소로 삼는 h -지수와는 달리 학술지에 게재된 피인용논문과 그 피인용횟수를 적용하여 지수 값을 산출한다. $Kor-h_T$ 는 피인용 논문뿐만 아니라 학술지에 게재된 모든 논문을 평가 대상으로 한다.

본 연구에서는 $Kor-h_T$ 의 의미를 알아보기 위해 지수 값의 중복률 및 지수 값과 평가요소 사이의 연관성 변화를 분석하여 h -지수, h_T 지수, 수정 h_T 및 IF 와 비교하였다. 지수 값의 중

복률 분석은 지표의 변별력을 살펴보기 위한 것이며, 지수 값과 평가요소와의 상관관계 분석은 지표의 평가요소인 학술지의 피인용횟수 및 논문 수가 지수 값에 각각 어떻게 반영되는지를 알아보기 위한 것이다.

연구 결과 지표의 변별력 측면에서는 h_T 지수에 비해 수정 h_T 지수 값의 중복률이 월등히 낮았으며, $Kor-h_T$ 의 중복률이 가장 낮아 변별력이 가장 우수한 것으로 나타났다. h_T 지수, h_T 평균, 수정 h_T , $Kor-h_T$ 모두 피인용횟수가 논문 수보다 지표와의 상관관계가 더 높으면서 두 평가요소 모두 일정 수준 이상의 상관관계를 갖고 있는 것으로 나타나 우수한 지표로 평가할 수 있다. 특히 이 중에서도 $Kor-h_T$ 가 피인용횟수와 논문 수 측면 모두에서 상관관계가 가장 높은 것으로 나타났다. 그리고 $Kor-h_T$ 는 피인용횟수의 평균인 IF 와도 일정 수준 이상의 상관관계가 있는 것을 확인할 수 있었다.

Hirsch 유형 지수들은 보통 자기피인용횟수를 100% 반영하여 산출된다. 그러나 KCI 데이터 분석에서 언급한 바와 같이 KCI에 수록된 학술지들의 자기피인용률이 매우 높기 때문에 자기피인용횟수에 대한 통제가 필요하다. 본 연구에서는 Hirsch 유형 지수들에 대해 자기피인용횟수를 일부만 반영하거나 모두 제외하는 것에 대한 검증은 수행하고자 하였으나, 본 연구를 위해 제공된 KCI 데이터에서 논문별 자기피인용횟수를 확보하는 것이 불가능하여 이를 수행하지 못하였다. 이는 본 연구의 한계점으로, 추후 KCI 데이터의 보완과 후속 연구를 통해 자기피인용횟수에 대한 통제 및 자기피인용횟수의 적정 반영 정도를 검증할 필요가 있다.

참 고 문 헌

- 고영만, 박지영 (2012). 한국 학술지 평가를 위한 KCI 기반 복합지표의 지수 값과 질적·양적 평가요소 사이의 연관성 및 학술지 등재 상태 구별 능력에 관한 연구. 한국문헌정보학회지, 46(2), 245-260. <http://dx.doi.org/10.4275/KSLIS.2012.46.2.245>
- 고영만, 조수련 (2012). 국내 연구 환경에 적합한 KCI 기반 학술지 복합인용지표 개발 및 연구업적 평가 적용방안 연구. 대전: 한국연구재단.
- 김관준, 이재운 (2010). 학술지 영향력 측정을 위한 h -지수의 응용에 관한 연구. 정보관리학회지, 27(1), 269-287. <http://dx.doi.org/10.3743/KOSIM.2010.27.1.269>
- 이재운, 정은경, 유소영 (2012). 계량서지적 평가지표(R&D Activities). 대전: 한국연구재단.
- 조은성, 송재도 (2011). 국내외 마케팅 학술지의 영향력: Kor-Factor와 Impact Factor의 문제점을 중심으로. 마케팅관리연구, 16(2), 53-82.
- 홍중선, 조수련 (2007). KCI 기반 Kor-Factor(Korea Factor) 평가지표 개발 및 시범적용. 서울: 한국학술진흥재단.
- Alonso, S., Cabrerizo, F. J., Herrera-Viedma, E., & Herrera, F. (2010). hg-index: A new index to characterize the scientific output of researchers based on the h - and g -indices. *Scientometrics*, 82(2), 391-400.
- Anderson, T. R., Hankin, R. K. S., & Killworth, P. D. (2008). Beyond the Durfee square: Enhancing the h -index to score total publication output. *Scientometrics*, 76(3), 577-588. <http://dx.doi.org/10.1007/s11192-007-2071-2>
- Batista, P. D., Campiteli, M. G., Konouchi, O., & Martinez, A. S. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics*, 68(1), 179-189. <http://dx.doi.org/10.1007/s11192-006-0090-4>
- Braun, T., Glänzel, W., & Schubert, A. (2006). A Hirsch-type index for journals. *Scientometrics*, 69(1), 169-173. <http://dx.doi.org/10.1007/s11192-006-0147-4>
- Egghe, L. (2006). Theory and practise of the g -index. *Scientometrics*, 69(1), 131-152. <http://dx.doi.org/10.1007/s11192-006-0144-7>
- Garfield, E. (2006). The history and meaning of the Journal Impact Factor. *Journal of the American Medical Association*, 295(1), 90-93. <http://dx.doi.org/10.1001/jama.295.1.90>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572. <http://dx.doi.org/10.1073/pnas.0507655102>

- Jin, B. (2006). h-index: An evaluation indicator proposed by scientist. *Science Focus*, 1(1), 8-9.
- Jin, B. (2007). The AR-index: Complementing the h-index. *ISSI Newsletter*, 3(1), 6.
- Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The R- and AR-indices: Complementing the h-index. *Chinese Science Bulletin*, 52(6), 855-863.
- Ko, Y. M., Cho, S.-R., & Park, Y. S. (2011). A study on the optimization of KCI-based index (Kor-Factor) in evaluating Korean journals. *Scientometrics*, 88(1), 61-71.
<http://dx.doi.org/10.1007/s11192-011-0384-z>
- Tol, R. S. J. (2009). The h-index and its alternatives: An application to the 100 most prolific economists. *Scientometrics*, 80(2), 317-324. <http://dx.doi.org/10.1007/s11192-008-2079-7>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Cho, Eun Seong, & Song, Jae Do (2011). The influence of Korean and international marketing journals: Focused on the problems of Kor-Factor and Impact Factor. *Journal of Marketing Management*, 16(2), 53-82.
- Hong, Chong Sun, & Cho, Soo-Ryun (2007). Development and preliminary application of evaluation index, Kor-Factor, based on KCI. Seoul: Korea Research Foundation.
- Kim, Pan-Jun, & Lee, Jae Yun (2010). A study on journal impact measurement with Hirsch-type indices, 27(1), 269-287. <http://dx.doi.org/10.3743/KOSIM.2010.27.1.269>
- Ko, Young Man, & Cho, Soo-Ryun (2012). A study on the optimization of KCI-based index with composite elements and the application to evaluate the Korean domestic academic journals. Daejeon: National Research Foundation of Korea.
- Ko, Young Man, & Park, Ji Young (2012). A study on the correlation between the index values and qualitative · quantitative evaluation elements and the distinction ability of the journals' registration condition of the KCI-based composite index in evaluating Korean journals. *Journal of the Korean society for Library and Information Science*, 46(2), 245-260.
<http://dx.doi.org/10.4275/KSLIS.2012.46.2.245>
- Lee, Jae Yun, Chung, EunKyung, & Yu, So-Young (2012). A bibliometric indicators for research policy (R&D Activities). Daejeon: National Research Foundation of Korea.

[부록] Kor-h_T 지수 값 산출 방식

가. 산출식

Kor-h_T: 2008년, 2009년(학술지 발행연도) 각각의 수정h_T 값을 구한 다음, 두 값의 평균 산출
 수정h_T : (측정기간의 h_T 지수값) + ('대분류별 최대피인용논문수+1'의 tapered식 일괄적용값) x
 (피인용 0인 논문수)

나. 산출 절차

1. 평가기간 동안의 각 학술지의 논문별 피인용횟수 데이터 수집하여 정렬

- 평가기간 동안 한 번이라도 인용된 논문들의 피인용횟수 기준 정렬
- 대부분의 h 계열 지표 산출에서 필요한 과정임
- 논문별 피인용횟수 기준으로 내림차순 정렬하여 각 논문에 순위 부여함, h 계열 지표 산출시 논문 순위는 동일 값에 대해 동일 순위를 부여하지 않고 1, 2, 3, 4, ... 순으로 하나씩 부여함
- 수정h_T 평균을 산출하는 것이므로 2008년과 2009년(발행연도 기준) 데이터를 구분하여 피인용횟수 기준으로 내림차순으로 정렬

* 예: 학술지 '도서문화' 2008년 발행논문의 2010년 피인용횟수 기준 정렬

L	M	N	O	P	Q	R	S	T	U	V	W
학술지명	권	호	발행년	저자구분	저자	저자 소속	논문제목	등재구분	피인용횟수	논문순위	피인용 정렬
『도서문화』		32	2008	주저자	김준	전남발전연	섬관광 실태 등재		3	1	3
『도서문화』		32	2008	주저자	홍선기	목포대학교	島嶼文化研; 등재		2	2	2
『도서문화』		31	2008	주저자	정철환	목포대학교	한반도 남서 등재		2	3	2
『도서문화』		31	2008	주저자	박종오	목포대학교	'삿구덩이'를 등재		1	4	1
『도서문화』		32	2008	주저자	이헌중	목포대학교	서남해안 도 등재		1	5	1
『도서문화』		31	2008	주저자	이헌중	목포대학교	신석기시대 등재		1	6	1
『도서문화』		31	2008	주저자	김준	전남발전연	젓새우잡이; 등재		1	7	1
『도서문화』		31	2008	주저자	나승만	목포대학교	조도 조기잡 등재		1	8	1
『도서문화』		31	2008	주저자	고석규	목포대학교	조선시기 표 등재		1	9	1
『도서문화』		31	2008	주저자	홍선기	목포대학교	지리정보시; 등재		1	10	1
『도서문화』		32	2008	주저자	김명진	경북대학교	太祖王建의 등재		1	11	1
『도서문화』		32	2008	주저자	이경엽	목포대학교	한국 도서; 등재		1	12	1
19세기 영어권 문	12	2	2008	주저자	이선주	이화여자대	디킨즈의 『; 등재		1	1	1
!Canon&Culture(?;2		1	2008	주저자	김정우	충신대학교	클란의 성서 등재후보		1	1	1
!Canon&Culture(?;2		1	2008	주저자	E. Tov	The HebrewThe Contrib	등재후보		1	2	1

* 예: 학술지 '도서문화' 2009년 발행논문의 2010년 피인용횟수 기준 정렬

L	M	N	O	P	Q	R	S	T	U	V	W
학술지명	권	호	발행년	저자구분	저자	저자 소속	논문제목	등재구분	피인용횟수	논문순위	피인용 정렬
『도서문화』		34	2009	주저자	윤영숙	목포대학교	강화도 첫새 등재		2	1	2
『도서문화』		33	2009	주저자	강봉룡	목포대학교	한국 해양사 등재		2	2	2
『도서문화』		34	2009	주저자	홍순일	목포대학교	한반도 도 등재		1	3	1
『도서문화』		33	2009	주저자	홍석준	목포대학교	『도서해양』 등재		1	4	1
『도서문화』		33	2009	주저자	김경옥	목포대학교	16~17세기 등재		1	5	1
『도서문화』		33	2009	주저자	조경만	목포대학교	갯벌과 인간 등재		1	6	1
18세기 영문학	6	2	2009	주저자	최주리	이화여자대	Anxiety, Res 등재		1	1	1
19세기 영어권 문학	13	1	2009	주저자	육은정	국민대학교	Sympathy w 등재		1	1	1
Comparative Korean S	17	1	2009	주저자	박옥수	영남대학교	단편소설 변 등재		3	1	3
Comparative Korean S	17	1	2009	주저자	이강록	우송대학교	한국어 읽기 등재		3	2	3
Comparative Korean S	17	3	2009	주저자	김경훈	연변대학교	디아스포라 등재		2	3	2
Comparative Korean S	17	3	2009	주저자	김종희	경희대학교	재외 하위 C 등재		2	4	2

2. '평가기간 동안의 tapered h-지수 값' 산출

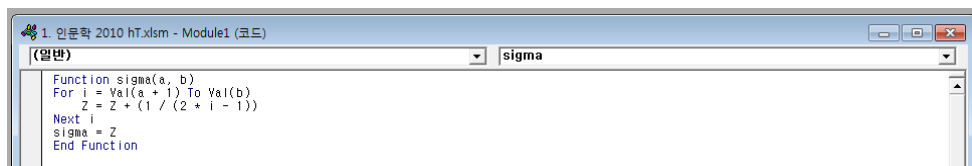
▪ $h_T = \sum_{r=1}^n h_{T(r)}$ (n: 논문 수)

$c_r \leq r$ 이면, 해당 논문의 값 $h_{T(r)} = \frac{c_r}{2r-1}$

$c_r > r$ 이면, 해당 논문의 값 $h_{T(r)} = \frac{r}{2r-1} + \sum_{i=r+1}^{c_r} \frac{1}{2i-1}$

▪ 본 연구 과정에서는 엑셀의 매크로와 함수를 사용하여 위 수식에 따른 학술지 h_T 지수 값을 산출함

① “1”에서 얻은 논문순위와 피인용횟수를 바탕으로 하는 매크로 작성



* sigma(논문순위, 피인용횟수)

* 예: 학술지 ‘도서문화’의 2008년 발행논문에 대한 sigma값 계산
(논문별 h_{T(r)}값 산출을 위한 중간 과정)

f_s =sigma(V2, W2)

L	M	N	O	P	Q	R	S	T	U	V	W	X
학술지명	권	호	발행년	저자구분	저자	저자 소속	논문제목	등재구분	피인용횟수	논문순위	피인용 정렬	sigma
『도서문화』		32	2008	주저자	김준	전남발전연·섬관광 실태 등재			3	1	3	0.5333333333
『도서문화』		32	2008	주저자	홍선기	목포대학교 島嶼文化硏 등재			2	2	2	0
『도서문화』		31	2008	주저자	정철환	목포대학교 한반도 남서 등재			2	3	2	0
『도서문화』		31	2008	주저자	박종오	목포대학교 '삿구덩이'를 등재			1	4	1	0
『도서문화』		32	2008	주저자	이현중	목포대학교 서남해안 도 등재			1	5	1	0
『도서문화』		31	2008	주저자	이현중	목포대학교 신석기시대 등재			1	6	1	0
『도서문화』		31	2008	주저자	김준	전남발전연·첫새우잡이 등재			1	7	1	0
『도서문화』		31	2008	주저자	나승만	목포대학교 조도 조기잡 등재			1	8	1	0
『도서문화』		31	2008	주저자	고석규	목포대학교 조선시기 표 등재			1	9	1	0
『도서문화』		31	2008	주저자	홍선기	목포대학교 지리정보시 등재			1	10	1	0
『도서문화』		32	2008	주저자	김명진	경북대학교 太祖王建의 등재			1	11	1	0
『도서문화』		32	2008	주저자	이경엽	목포대학교 한국 도서허 등재			1	12	1	0

② 논문순위와 피인용횟수, sigma값을 이용하여 각 논문의 h_{T(r)}값 계산: (엑셀 수식 'IF' 이용)

* 예: 학술지 ‘도서문화’의 2008년 발행논문에 대한 h_{T(r)}값 계산

f_s =IF(V2>=W2,W2/(2*V2-1),(V2/(2*V2-1))+X2)

L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
학술지명	권	호	발행년	저자구분	저자	저자 소속	논문제목	등재구분	피인용횟수	논문순위	피인용 정렬	sigma	hT@
『도서문화』		32	2008	주저자	김준	전남발전연·섬관광 실태 등재			3	1	3	0.5333333333	1.5333333333
『도서문화』		32	2008	주저자	홍선기	목포대학교 島嶼文化硏 등재			2	2	2	0	0.6666666667
『도서문화』		31	2008	주저자	정철환	목포대학교 한반도 남서 등재			2	3	2	0	0.4
『도서문화』		31	2008	주저자	박종오	목포대학교 '삿구덩이'를 등재			1	4	1	0	0.142857143
『도서문화』		32	2008	주저자	이현중	목포대학교 서남해안 도 등재			1	5	1	0	0.1111111111
『도서문화』		31	2008	주저자	이현중	목포대학교 신석기시대 등재			1	6	1	0	0.090909091
『도서문화』		31	2008	주저자	김준	전남발전연·첫새우잡이 등재			1	7	1	0	0.076923077
『도서문화』		31	2008	주저자	나승만	목포대학교 조도 조기잡 등재			1	8	1	0	0.066666667
『도서문화』		31	2008	주저자	고석규	목포대학교 조선시기 표 등재			1	9	1	0	0.058823529
『도서문화』		31	2008	주저자	홍선기	목포대학교 지리정보시 등재			1	10	1	0	0.052631579
『도서문화』		32	2008	주저자	김명진	경북대학교 太祖王建의 등재			1	11	1	0	0.047619048
『도서문화』		32	2008	주저자	이경엽	목포대학교 한국 도서허 등재			1	12	1	0	0.043478261

③ 논문별 h_{T(r)}값을 모두 더하여 학술지의 h_T(tapered h-지수)값 계산

→ 각 학술지의 (평가기간 동안의 tapered h-지수 값) 얻음

* 예: 학술지 ‘도서문화’의 2008년(발행연도 기준)분 학술지 h_T값 계산

f_s =SUM(Y2:Y13)

L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
학술지명	권	호	발행년	저자구분	저자	저자 소속	논문제목	등재구분	피인용횟수	논문순위	피인용 정렬	sigma	hT@
『도서문화』		32	2008	주저자	김준	전남발전연·섬관광 실태 등재			3	1	3	0.5333333333	1.5333333333
『도서문화』		32	2008	주저자	홍선기	목포대학교 島嶼文化硏 등재			2	2	2	0	0.6666666667
『도서문화』		31	2008	주저자	정철환	목포대학교 한반도 남서 등재			2	3	2	0	0.4
『도서문화』		31	2008	주저자	박종오	목포대학교 '삿구덩이'를 등재			1	4	1	0	0.142857143
『도서문화』		32	2008	주저자	이현중	목포대학교 서남해안 도 등재			1	5	1	0	0.1111111111
『도서문화』		31	2008	주저자	이현중	목포대학교 신석기시대 등재			1	6	1	0	0.090909091
『도서문화』		31	2008	주저자	김준	전남발전연·첫새우잡이 등재			1	7	1	0	0.076923077
『도서문화』		31	2008	주저자	나승만	목포대학교 조도 조기잡 등재			1	8	1	0	0.066666667
『도서문화』		31	2008	주저자	고석규	목포대학교 조선시기 표 등재			1	9	1	0	0.058823529
『도서문화』		31	2008	주저자	홍선기	목포대학교 지리정보시 등재			1	10	1	0	0.052631579
『도서문화』		32	2008	주저자	김명진	경북대학교 太祖王建의 등재			1	11	1	0	0.047619048
『도서문화』		32	2008	주저자	이경엽	목포대학교 한국 도서허 등재			1	12	1	0	0.043478261

3. {(대분류별 최대 피인용논문수+1의 tapered식 일괄적용값)×(피인용횟수 0인 논문 수)} 산출

- 8개 대분류로 구분하여 인용되지 않은 논문에 대해 대분류 내에서 일괄적으로 같은 값을 부여하는 방식으로, 이 과정에서 대분류별 차이가 일부 고려되면서 원래 h_T 지수에 비해 수정 h_T 에서 대분류별 차이가 다소 상쇄됨
- “2”는 평가기간 동안 한 번 이상 인용된 논문에 대한 값을 구하는 것이고, “3”은 평가기간 동안 한 번도 인용되지 않은 논문에 대한 값을 구하는 것임

① ‘대분류별 최대 피인용논문수+1의 tapered식 일괄적용값’ 산출

- .. 8개 대분류로 구분하여 각 대분류 내에서 학술지의 최대 피인용논문수를 확인할 수 있으며 각 학술지의 논문별 피인용횟수 기준 정렬 데이터 (1)에서 논문순위의 마지막 값을 통해 얻을 수 있음
- .. 앞의 예에서 인문학 분야 학술지 ‘도서문화’의 2008년 발행논문의 논문순위 마지막 값 12는 2008년 발행논문 중 2010년 피인용논문수가 12라는 의미이며, 각 분야 학술지들의 마지막 논문순위를 모아서 가장 큰 값을 찾으면 됨
- .. 예: 인문학 분야의 경우, 2008년 발행 기준으로 학술지 ‘漢文學報’의 피인용논문수가 45로 가장 커서

$$(\text{인문학 최대 피인용논문수}+1 \text{의 tapered식 일괄적용값}) = \frac{1}{2 \times (45+1) - 1}$$

- .. 이때 분야별 ‘최대 피인용논문수’에 1을 더하는 것은 평가기간 동안 한 번도 인용되지 않은 논문이 해당 분야에서 한 번이라도 인용된 논문에 비해 더 적은 값을 갖게 하기 위한 것이며, 이는 한 번이라도 인용된 논문이 전혀 인용되지 않은 논문에 비해 높은 평가를 받도록 하기 위한 조치임

〈수정 h_T 평균 산출 시 대분류별 피인용횟수 0인 논문 1편에 대한 일괄적용값〉

대분류	2008 발행 논문		2009 발행 논문		2008, 2009 2년 발행 논문	
	최대 피인용 논문수	일괄적용값	최대 피인용 논문수	일괄적용값	최대 피인용 논문수	일괄적용값
인문학	45	0.01099	46	0.01075	81	0.00613
사회과학	86	0.00578	98	0.00508	182	0.00274
자연과학	177	0.00282	178	0.00280	316	0.00158
공학	190	0.00262	185	0.00270	375	0.00133
의약학	72	0.00690	62	0.00800	134	0.00372
농수해양	147	0.00339	153	0.00326	300	0.00166
예술체육	227	0.00220	234	0.00213	461	0.00108
복합학	128	0.00389	181	0.00275	309	0.00162

.. 위 표의 일괄적용값은 $\frac{1}{2 \times \{(\text{최대 피인용 논문수}) + 1\} - 1}$ 의 값을 소수 5째 자리까지만 나타낸 것이며, 실제 계산 과정에서는 반올림하지 않고 원값을 사용함

② 각 학술지의 평가기간 동안 한 번도 인용되지 않은 논문 수를 구하여 ①에서 얻은 해당 분야의 일괄적용값과 곱해 줌

→ {(대분류별 최대 피인용논문수+1의 tapered식 일괄적용값) × (피인용횟수 0인 논문 수)} 얻음

.. 예: 학술지 ‘도서문화’는 인문학 분야에 속하므로 2008년 발행 기준의 일괄적용값은

$\frac{1}{2 \times (45 + 1) - 1}$ 이고, 이 학술지의 2008년 발행논문 중 2010년 동안 한 번도 인용되지

않은 논문 수는 8이므로 인용되지 않은 논문에 대한 평가 값은

$$\frac{1}{2 \times (45 + 1) - 1} \times 8 \approx 0.08791$$

4. “2”에서 얻은 값과 “3”에서 얻은 값을 더하여 평가기간의 수정h_T 산출

$$\text{수정}h_T = (\text{측정기간의 } h_T \text{ 지수값}) + (\text{'대분류별 최대 피인용논문수+1'의 tapered식 일괄적용값}) \times (\text{피인용 0인 논문수})$$

- 예: 2008년 발행 기준의 학술지 ‘도서문화’의 수정h_T 값
 = ((2)에서 얻은 h_T값) + (인문학 분야 일괄적용값) × (인용되지 않은 논문 수)
 = 3.29101... + 0.08791...
 ≈ 3.37893

5. 각 학술지의 수정h_T 평균 산출

- 각 학술지의 2008년 발행논문의 2010년도 피인용횟수 데이터를 이용하여 수정h_T값을 산출하고, 2009년 발행논문의 2010년도 피인용횟수 데이터를 이용하여 수정h_T값을 산출한 다음, 두 값의 평균을 계산
 → “Kor-h_T (수정h_T 2년 평균)” 값을 얻음
- 소수 5째 자리로 반올림한 값 사용
- 평가기간이 3년, 4년, 5년으로 늘어날 경우 각 발행연도별로 수정h_T값을 산출한 다음, 그 평균을 계산하면 됨
- 평가기간을 1년으로 하여 즉시성지수와 유사한 개념으로 사용할 수도 있음

