

# 피어슨 상관분석을 이용한 계량서지 지수 비교

## A comparison of bibliometric indicators using Pearson's correlation

손경미, 계명대학교 문헌정보학과, real@kmu.ac.kr

이용구, 계명대학교 문헌정보학과, yonggulee@kmu.ac.kr

Gyeong-Mi Son, Dept. of LIS, Keimyung University

Yong-Gu Lee, Dept. of LIS, Keimyung University

연구자의 연구 성과와 저널의 영향력을 평가 할 수 있는 계량서지 지수는 매우 중요하다. 하지만 이들은 특징이 모두 다르기 때문에 저널을 선택할 때 여러 가지 지수를 함께 고려할 필요가 있다. 이를 위해 본 연구에서는 피어슨 상관분석을 통해 각 지수들을 비교 분석하여 이들 간의 관계를 보여주고자 한다. 이를 위해 상관 분석을 주제와 Article Influence Score(AIS) 순위에 따라 수행하였다. Total Cites와 Eigenfactor Score, Article influence Score와 5-Year Impact Factor의 상관 분석 결과를 주제에 따른 평균값으로 살펴보았을 때, 이들은 매우 높은 상관관계가 있었다. 또한 Article influence Score와 5-Year Impact Factor의 경우, 모든 주제에서 AIS 순위에 따른 상위 25% 구간에서 두 지수 간에 상관관계가 매우 높은 것으로 나타났다.

### 1. 서론

계량서지 지수는 연구자의 연구 성과와 저널의 영향력을 평가할 수 있는 척도로 활용되기 때문에 중요하다. 저널 영향력지수(Journal Impact Factor), 총 인용 횟수(Total Cites), 아이겐 팩터 점수(Eigenfactor Score), 논문 영향도 지수(Article Influence Score) 이외에도 많은 계량서지 지수가 저널의 평가척도로 사용되고 있다. 이들은 특징이 모두 다르기 때문에 저널을 선택할 때 단순히 한 가지 지수를 바탕으로 선택하는 것보다 여러 가지 지수를 함께 고려하는 복합적인 평가가 필요하다.

Impact Factor는 저널의 영향력을 평가할

수 있지만 해당 저널에 게재된 각각의 논문의 영향력이나 연구자의 영향력을 평가하는 것은 어려움이 있다. Eigenfactor는 인용패턴에 따라 가중치를 부여하고 자기 인용을 제외하여 저널의 영향력을 평가하기 때문에 Impact Factor보다 보완된 지표이다. 이와 함께 살펴볼 수 있는 지수로는 Eigenfactor Score를 논문 수로 정규화한 Article Influence Score가 있다.

이처럼 각각의 계량서지 지수가 가지고 있는 한계점을 지수 비교를 통해 보완하여 세밀한 영향력 평가가 이루어지도록 해야 한다. 이에 본 연구에서는 피어슨 상관분석을 통해 저널들의 각 지수들을 비교하고 분석하여 이들 간의 관계를 보여주고자 한다.

## 2. 선행연구

계량서지 지수 연구는 인용 패턴을 중심으로 하는 연구와 기존의 계량서지 지수를 보완하는 새로운 지수를 제안하는 연구가 수행되어 졌다. 인용 패턴에 근거한 지수 연구는 이재윤(2011)이 있다. Eigenfactor 지수를 제안한 West(2010) 연구가 있다.

이 연구에서는 기존의 저널 영향력 평가 지수 Impact Factor와 Total Cites, 상대적으로 최근에 제시된 Eigenfactor Score와 Article Influence Score를 분석하여 지수들의 관계와 특징을 이해하고자 한다.

West, Bergstrom, Bergstrom(2010)은 Impact Factor와 Eigenfactor의 차이점이 총 인용 횟수를 계산하기 위한 기간이 2년과 5년으로 다르다고 하였으며 Impact Factor가 총 인용 횟수를 해당 저널의 총 논문 수로 나누었기 때문에 Eigenfactor Score를 해당 저널의 총 논문 수로 나눈 Article Influence Score와

비교할 수 있다고 하였다.

## 3. 연구방법

### 3.1 데이터 수집 및 분석방법

데이터 수집대상은 Web of Science 데이터 베이스의 Science와 Social Science 분야 학술지 인용 정보를 바탕으로 하는 Thomson Reuters의 Journal Citation Reports(JCR)에서 제공하는 232개의 주제 분야 중 저널의 수가 가장 많은 상위 4가지(Biochemistry, Mathematics, Economics, Pharmacology) 주제 분야를 대상으로 하였다.

데이터는 선정된 4가지 주제 분야에서 모든 저널의 Total Cites, Impact Factor, 5-Year Impact Factor, Eigenfactor Score, Article Influence Score와 같은 2013년 기준의 계량서지 지수 값을 추출하였다. 이들 중 한 가지 이상의 지수에서 결측값이 존재하는 저널은

<표 1> 주제별 전체 계량서지 지수 값의 피어슨 상관분석 결과

\* () : Impact Factor와 각 지수의 상관계수

상관계수	Mathmatics	Biochemistry	Economics	Pharmacology	평균
5-Year Impact Factor & Eigenfactor® Score	0.4526 (0.4745)	0.5644 (0.5673)	0.6691 (0.6452)	0.5264 (0.5361)	0.5531 (0.5557)
Total Cites & 5-Year Impact Factor	0.4295 (0.4528)	0.3631 (0.3663)	0.6487 (0.6354)	0.4137 (0.4105)	0.4637 (0.4662)
Total Cites & Eigenfactor® Score	<b>0.9187</b>	0.9268	<b>0.9273</b>	0.9372	0.9275 (0.9275)
Article influence® Score & 5-Year Impact Factor	0.8919 (0.8175)	<b>0.9725</b> (0.9479)	0.8816 (0.8145)	<b>0.9879</b> (0.9817)	0.9335 (0.8904)
Total Cites & Article influence® Score	0.3551	0.3772	0.6769	0.3593	0.4421 (0.4421)
Article influence® Score & Eigenfactor® Score	0.4108	0.6010	0.7113	0.4833	0.5516 (0.5516)

제외 하였다.

데이터 분석 방법에서 각 지수 간의 비교를 위해 상관분석이 필요하다. 이러한 지수 비교를 위해 추출된 전체 데이터를 이용하여 피어슨 상관분석을 수행하였고, 결과는 <표 1>과 같다. 또한 연구 대상인 4가지 분야 전체 저널의 지수별 상관계수를 <표 2>로 작성하였다. 이 결과를 바탕으로 각 지수 간의 관계를 구체적으로 분석하기 위해 주제별로 Article Influence 순위를 기준으로 하여 상위 25%, 25-50%, 50-75%, 75-100% 4구간으로 나누어 상관계수 분석을 수행하였다.

### 3.2 분석 결과

전체 계량서지 지수 값의 피어슨 상관분석 결과를 평균적으로 살펴보면, Article influence Score와 5-Year Impact Factor는 0.9335, Total Cites와 Eigenfactor Score는 0.9275로 상관관계가 상대적으로 매우 높은 것으로 나타났다. 하지만 나머지 지수들 간의 상관관계는 상대적으로 낮은 것으로 나타났다.

주제별로 살펴보면, Total Cites와 Eigenfactor Score는 대체로 상관관계가 매우 높게 일관되게 나왔지만, Article influence Score와 5-Year Impact Factor는 주제에 따라 상관관계가 차이가 있는 것으로 나타났다. Biochemistry와 Pharmacology 두 분야가 매우 높은 상관관계를 보였다.

주제와 무관하게 수집된 전체 저널의 지수별 상관분석(<표 2>)의 상관계수 값에서도 Total Cites와 Eigenfactor Score는 0.9269, Article influence Score와 5-Year Impact Factor는 0.7872로 나타나 이들 지수의 쌍 간에 높은 상관관계가 있음을 알 수 있다.

<표 2> 전체 피어슨 상관계수

전체 피어슨 상관계수	
5-Year Impact Factor & Eigenfactor® Score	0.5368
Total Cites & 5-Year Impact Factor	0.4004
Total Cites & Eigenfactor® Score	<u>0.9269</u>
Article influence® Score & 5-Year Impact Factor	<u>0.7872</u>
Total Cites & Article influence® Score	0.3222
Article influence® Score & Eigenfactor® Score	0.4979

유사한 두 지수인 Article influence Score와 5-Year Impact Factor를 상세히 비교하기 위하여, Article influence Score를 4가지 구간으로 나누어 해당 그룹별로 피어슨 상관계수를 구하였다. 그 결과는 <표 3>과 같다.

Article influence Score와 5-Year Impact Factor는 Article influence Score 순위와 주제에 따라 상이한 결과를 나타내었다. Mathematics와 Economics의 경우, 상위 25% 구간에서는 매우 높은 상관관계를 나타내지만 나머지 구간에서는 상관관계가 낮다. Biochemistry와 Pharmacology의 경우, 0-25% 구간과 75-100% 구간에서는 매우 높은 상관관계를 나타내지만 25-50% 구간과 50-75% 구간에서는 상대적으로 낮은 상관관계를 나타내었다. 모든 주제에서 상위 25% 구간은 두 지수 간에 상관관계가 매우 높은 것으로 나타났다. 이는 Article influence Score가 높은 저널이 동시에 5-Year Impact Factor 높다는 것을 의미한다. 즉, Article influence Score가 높은 저널들은 역시 5-Year Impact Factor도 높다.

〈표 3〉 Article Influence Score 순위에 따른  
Article Influence Score & 5-Year Impact Factor의 피어슨 상관분석 결과

\* ( ) : Impact Factor와 Article Influence Score의 상관계수

Article influence® Score & 5-Year Impact Factor	Mathmatics	Biochemistry	Economics	Pharmacology
상위 0-25%	0.9181 (0.8732)	0.9696 (0.9303)	0.8523 (0.7929)	0.9906 (0.9842)
상위 25-50%	0.2893 (0.2134)	0.6082 (0.5682)	0.4219 (0.3468)	0.6264 (0.4055)
상위 50-75%	0.1345 (0.0661)	0.3649 (0.2015)	0.5109 (0.3218)	0.6807 (0.6303)
상위 75-100%	0.3438 (0.3244)	0.9504 (0.8685)	0.5709 (0.3694)	0.8556 (0.7953)

#### 4. 결론 및 제언

이 연구에서는 계량서지 지수의 상관분석을 통해 각 지수들의 관계와 특징을 알아보고자 하였다. 그 결과는 다음과 같다.

첫째, Total Cites와 Eigenfactor Score, Article Influence Score와 5-Year Impact Factor의 상관 분석 결과를 주제에 따른 평균 값으로 살펴보았을 때, 이들은 매우 높은 상관관계가 있음을 알 수 있다. 또한 5-Year Impact Factor를 Impact Factor로 대입하여 상관분석을 하였을 때, 결과에서는 큰 차이가 없음을 알 수 있다.

둘째, Article Influence Score와 5-Year Impact Factor의 경우, 모든 주제에서 AIS 순위에 따른 상위 25% 구간에서 두 지수 간에 상관관계가 매우 높은 것으로 나타났다. 그러므로 이는 두 지수가 서로 대체될 수 있음을 의미한다.

후속 연구에서는 주제 분야를 확대했을 경우에도 동일한 상관관계를 나타내는지 확인할 필요가 있다. 아울러 주제별 특징을 파악하는

미시적인 연구가 필요하다. 이를 통해 Article Influence Score와 5-Year Impact Factor의 상관 분석을 면밀히 수행하여 두 지수가 서로 대체 될 수 있는지 확인 할 수 있을 것으로 생각된다.

#### 참고문헌

이재윤 (2011). 인용 네트워크 분석에 근거한 문헌 인용 지수 연구. 한국문헌정보학회지, 45(2), 119-143.

Jevin D. West (2010). Eigenfactor: ranking and mapping scientic knowledge. Unpublished doctoral dissertation, University of Washington. Seattle, WA, USA.

Jevin D. West, Theodore C. Bergstrom, and Carl T. Bergstrom (2010). The Eigenfactor Metrics™: A Network Approach to Assessing Scholarly Journals. College & Research Libraries, 71(3), 236-244.