

한국프로야구에서의 투수평가지표

이장택¹

¹단국대학교 응용통계학과

접수 2014년 1월 22일, 수정 2014년 2월 12일, 게재확정 2014년 3월 12일

요약

투수를 평가할 때 중요한 요소는 일반적으로 다승과 방어율을 사용하지만 이 지표들은 팀의 도움 또는 운과 같은 요소의 영향을 받는다. 그래서 야구통계학자들은 투수 개인의 능력만을 측정하는 많은 지표들을 제안하였는데 이와 같은 평가지표들은 가짓수가 너무 많고 복잡하기 때문에 팬들을 때때로 당황하게 만든다. 본 연구에서는 대표적인 투수평가지표들을 이용하여 지표들의 특성을 반영하는 주성분을 찾아보고 한국프로야구에 적합한 투수들의 능력을 객관적으로 평가할 수 있는 투수지표를 제안하였다.

주요용어: 세이버메트릭션, 주성분분석, 케이-평균군집, 투수평가지표.

1. 머리말

야구에서 투수의 경기결과와 상관성이 있는 대표적인 통계량은 승 (W), 패 (L), 홀드 (H), 세이브 (SV), 던진 이닝 (IP), 피홈런 (HR), 피안타 (H), 피1루타 (1B), 피2루타 (2B), 피3루타 (3B), 허용 볼넷 (BB), 탈삼진 (K), 허용희생타 (SF), 타수 (AB), 승률 (W%) 등이 있다. 그러나 이런 단편적인 양적 수치는 투수의 능력평가에 도움은 되지만 객관적인 평가를 하는데 사용하는 것은 무리가 따르는데 아무리 뛰어난 투수도 투수에 관한 모든 부분에서 뛰어나기가 매우 힘들기 때문이다.

일반적으로 어떤 통계량들이 투수를 평가하는데 가장 큰 항목으로 적용되는 지를 살펴보면 매스컴에서는 올해도 투수를 평가할 때, 다승, 방어율, 승률 등을 말한다. 하지만 이와 같은 고전적인 투수통계량은 일단 다승은 타력과 볼펜이 강한 팀의 선발이면 매우 유리하며, 방어율은 팀의 수비력이 상당한 영향력을 미치며, 승률은 운의 영향을 많이 받을 수 있다. 이런 이유로 다년간 쌓인 통계데이터를 분석하는 세이버메트릭션 (sabermetrician)들은 투수가 제어할 수 있는 것은 공을 던지는 것뿐이므로 팀의 도움과 무관한 투수 자신이 책임을 지는 통계량들을 연구하기 시작했으며 이와 같은 노력은 투수들의 평가에 중요한 잣대가 되기 시작하고 있다.

세이버메트릭션들이 고안한 중요한 투수평가지표들은 WHIP (walks plus hits divided by innings pitched), BABIP (batting average on balls in play), DICE (defense independent component era) 및 FIP (fielding independent pitching) 등이 있는데, 특정 지표를 사용하는 것에 따라 투수들의 평가가 서로 상이하게 될 수 있기 때문에 지표들에 대한 종합적인 판단을 할 필요가 있다. 하지만 야구팬들에게 이와 같은 세이버메트릭션 지표들이 너무 어렵게 여겨질 수도 있으며 또 인터넷 검색을 통한 각 지표가 매년 야구데이터의 업데이트 여부로 조금씩 차이가 있음을 알 수 있다. 따라서 여러 가지 투수지표들을 통합하는 지표를 연구할 필요성을 느끼는데, 그 중 한 가지 방법은 제안된 여러 가지 투수지표를 골고루

¹ (448-701) 경기도 용인시 죽전동 126번지, 단국대학교 응용통계학과, 교수. E-mail: jtlee@dankook.ac.kr

잘 반영하는 주성분을 만드는 것이다. 따라서 본 연구에서는 투수지표를 1개의 주성분으로 요약하고 만들어진 지표를 이용하여 우리나라 대표투수들의 능력을 비교하고, 또한 몇 개의 군집으로 나누어 선수들의 수준을 평가하여 본다.

한편 한국프로야구에 대한 최근 연구들을 살펴보면 시계열모형을 이용하여 관중 수 예측을 다룬 Lee와 Bang (2010), 인공지능망을 이용하여 포스트시즌 진출 예측을 살펴본 Chea 등 (2010), 한국프로야구 타자들에 대한 세이버메트릭스 지수 값을 이용하여 선수들의 경기력과 연봉간의 패턴을 분석한 Seung과 Kang (2012), 출루율과 장타율이 득점에 미치는 연구를 한 Kim (2012)이 있다. 본 논문은 다음과 같이 구성되어 있다. 2절에서는 데이터의 구성 및 여러 가지 투수평가지표에 대한 설명과 기술 통계 값을 살펴보았으며, 3절에서는 제안된 투수지표와 K-평균 군집분석의 결과를 설명하였다. 끝으로 4절에서는 본 연구의 결론에 대해 언급하였다.

2. 연구방법

2.1. 데이터의 구성

본 연구에 사용된 데이터는 한국야구위원회 (KBO)에 기록되어 있는 2000년부터 2013년 사이에 있었던 경기 중 규정투수이닝 수를 채운 270명의 경기결과를 분석대상으로 하였다. 선수들의 경기 자료는 2000년부터 2006년 KBO에서 발행한 한국프로야구연감 및 홈페이지 (<http://www.koreabaseball.com/record>)에서 구하였으며, 통계패키지는 SPSS 18K를 이용하였다. 조사된 선수들의 수는 팀명이 우리, 히어로즈, 현대를 넥센에, 해태를 KIA에 포함시키면 KIA 32명, LG 26명, NC 3명, SK 33명, 넥센 36명, 두산 31명, 롯데 35명, 삼성 38명, 그리고 한화 36명으로 신생팀 NC를 제외하면 삼성이 가장 많았으며, LG가 가장 적었다.

2.2. 투수지표들의 정의

투수들의 평가를 위한 고전적인 지표와 세이버메트릭션들이 고안해 놓은 지표들은 상당수가 있다. 하지만 그중에서도 보편화되어있고 많이 사용하는 평가지표를 본 연구에서 사용하였으며 각 지표들에 대한 자세한 계산 방법은 Table 2.1에 정리되어 있다. 계수가 포함된 DICE나 FIP 지표는 모두 과거의 기록에 기초한 것이므로 매년 업데이트되어 계수가 변할 수 있기 때문에 인터넷에서 볼 수 있는 공식들은 조금씩 차이가 있다. 또한 이 수치들은 메이저리그 야구를 기반으로 하기 때문에 한국프로야구를 이용하여 재추정하면 다르게 나타날 수도 있다.

Table 2.1 Formulae for pitching statistics

Pitching Statistics	Formulae
ERA	$ERA = 9(ER)/IP$
K/9	$K/9 = 9(K)/IP$
BB/9	$BB/9 = 9(BB)/IP$
WHIP	$WHIP = (BB + H)/IP$
BABIP	$BABIP = (H - HR)/(AB - K - HR + SF)$
DICE	$DICE = 3.00 + \frac{13(HR) + 3(BB + HBP) - 2(K)}{IP}$
FIP	$FIP = \frac{13(HR) + 3(BB) - 2(K)}{IP} + 3.20$

1. ERA (earned run average)

투수의 잘못으로 내준 점수를 자책점 (earned run; ER)이라고 하고, 방어율 (ERA)은 9이닝 당 ER의 비율을 의미한다. ERA는 가장 보편화된 투수평가의 기준이지만 팀 전체의 문제인 실점을 투수만의 통계량으로 취급하는 모순이 있다.

2. K/9

9이닝 동안의 탈삼진 개수를 뜻한다. 실질적인 투수의 구위를 평가하는 기준으로 9로 표시하는 이유는 9이닝을 기준으로 하기 위해서다.

3. BB/9

9이닝 동안의 허용한 볼넷의 개수를 뜻한다. 투수의 제구력을 평가하는 기준이 되며 낮을수록 좋다.

4. WHIP (walks plus hits divided by innings pitched)

이닝 당 볼넷과 안타를 투수가 내주는 것을 의미하는 통계량으로 가장 보편화된 세이버메트릭션 투수지표이다. WHIP의 값이 적을수록 좋은 투수를 의미한다.

5. BABIP (batting average on balls in play)

인플레이로 이어진 타구에 대한 타율을 계산하는 용어인데, 홈런을 제외하는 건 홈런은 그 자체로 플레이가 종료된 상황이라고 간주할 수 있기 때문이다. 기본적으로 투수의 성적에 운이 얼마나 개입되는지에 대한 척도로 사용된다.

6. DIPS (defense independent pitching stats)

투수가 통제할 수 있는 영역인 삼진, 볼넷, 홈런, 데드볼을 갖고 평균자책점의 형태로 나타낸 것으로, 계수가 너무 복잡한 공식들이 많아서 최근에는 근사계산하는 간편화된 공식을 많이 사용하는데, 그 중 하나가 DICE (defense independent component era)이다.

7. FIP (fielding independent pitching)

DIPS의 개량된 형태로 투수의 미래 성적을 예측하는데 매우 유용하며 평균자책점과 같은 형태로 산출되기 때문에 비교해서 살펴보기가 편하다.

2.3. 데이터 분석

Table 2.2는 7가지 투수평가지표에 대해 수집된 데이터를 이용하여 계산한 기술통계 값을 보여준다. 가장 변화가 큰 측도는 K/9이며 평균값이 비슷한 ERA, DICE, FIP 중에서는 ERA가 팀의 수비력 등의 영향을 많이 받기 때문에 변동 폭이 DICE와 FIP보다 크며, 평균, 최대 및 최소를 각각 비교하면 FIP의 경우가 ERA와 DICE의 경우보다 약간 크다.

Table 2.2 Descriptive statistics for pitching statistics

	Mean	Variance	Maximum	Minimum
ERA	3.745	.567	5.890	1.820
K/9	6.277	1.920	10.347	2.556
BB/9	3.240	.837	6.248	1.408
WHIP	1.328	.019	1.706	1.000
BABIP	0.248	.000	.321	.196
DICE	3.965	.386	6.014	2.392
FIP	4.011	.375	6.065	2.469

Table 2.3은 7개의 투수지표들에 대한 상관행렬을 보여준다. K/9와 BABIP의 관계를 제외한 모든 지표들의 상관관계가 유의수준 1%에서 유의함을 알 수 있다.

Table 2.3 Correlations for pitching statistics

	ERA	K_9	BB_9	WHIP	BABIP	DICE	FIP
ERA	1**	-.200**	.351**	.753**	.314**	.663**	.661**
K/9		1**	.181**	-.203**	-.004**	-.484**	-.475**
BB/9			1**	.685**	-.291**	.398**	.403**
WHIP				1**	.346**	.611**	.617**
BABIP					1**	-.190**	-.176**
DICE						1**	.991**
FIP							1**

** : Correlation is significant at the 0.01 level (2-tailed).

3. 데이터분석

3.1. 주성분분석

프로야구 전구단의 2000년부터 2013년도 투수들의 기록을 이용하여 7개의 투수지표를 계산하고 상관관계가 큰 근거를 이용하여 변수 개수의 축약을 위해 요인분석 중주성분분석을 선택하였다. 표본적합도 KMO 값은 0.43으로 권장치인 0.5에 근접하고, Bartlett의 구형성 검정은 유의확률이 0.000으로 나타나 요인분석이 가능하다고 판단되었다. 그리고 7개의 투수평가지수 (ERA, K/9, BB/9, WHIP, BABIP, DICE, FIP)에 대해 상관행렬을 이용하여 주성분분석을 실시하였고 그 결과 설명된 총분산은 Table 3.1과 같다.

Table 3.1 Total variance explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	3.624	51.774	51.774
2	1.403	20.037	71.811
3	1.317	18.813	90.625
4	.456	6.515	97.140
5	.186	2.661	99.801
6	.009	.131	99.932
7	.005	.068	100.000

Table 3.1의 고유값 중에서 3개가 1보다 크며, 이에 대응하는 주성분이 전체 변동의 90.63%를 설명함으로써 7개의 변수에 대한 정보는 3개의 주성분으로 축약될 수 있으며 이는 스크리 플롯 (scree plot)을 이용하여도 마찬가지이다.

Table 3.2 Component matrix

Component	ERA	K/9	BB/9	WHIP	BABIP	DICE	FIP
1	.823	-.425	.583	.855	.019	.920	.921
2	.372	.119	-.212	.358	.976	-.253	-.243
3	.031	.789	.729	.269	-.142	-.189	-.182

Table 3.2의 성분행렬을 보면 주성분1에 대하여 7개 변수의 공통성은 DICE와 FIP이 크며 BABIP이 작게 나타나는데, 수치가 해당 변수와 주성분 사이의 상관계수를 의미하므로 주성분1은 투수의 능력을 나타내는 성분으로 이해할 수 있으며 작은 값을 갖는 것이 바람직함을 알 수 있다. 주성분2는 총분산의 20% 정도를 설명하는데, 공통성이 DICE와 FIP이 작고 BABIP이 월등히 큰 것으로 보아서 투수의 운

을 측정하는 요인으로 설명된다. 주성분3은 공통성이 K/9와 BB/9이 매우 큰 것으로 보아서 삼진과 볼넷에 연관되는 성분이다. 하지만 실제로 주성분3의 요인점수는 선수들의 성적과 연관성이 뚜렷하지 않기 때문에 이상을 정리하면 주성분1이 운을 배제한 투수의 능력에 대한 종합적 지표로 가장 타당하다고 간주된다.

3.2. 투수등급지표 (pitching grade index)

주성분분석의 결과를 이용하기 위하여 표본이 갖는 각 요인에 대한 측정치인 요인점수를 산출할 필요가 있는데, Z_1 부터 Z_7 을 ERA, K/9, BB/9, WHIP, BABIP, DICE, FIP의 표준화된 변수라고 할 때 Table 3.3은 각 표본의 변수값들을 구해진 요인들의 값으로 바꾸어 준 요인점수의 계수로서 주성분1에 대한 어떤 투수의 요인 점수 S 는 다음과 같이 계산된다.

$$S = 0.227Z_1 - 0.117Z_2 + 0.161Z_3 + 0.236Z_4 + 0.005Z_5 + 0.254Z_6 + 0.254Z_7$$

Table 3.3 Component score coefficient matrix

Component	ERA	K/9	BB/9	WHIP	BABIP	DICE	FIP
1	.227	-.117	.161	.236	.005	.254	.254
2	.266	.085	-.151	.255	.696	-.181	-.173
3	.024	.599	.554	.204	-.108	-.144	-.138

그리고 Figure 3.1은 계산된 투수들의 요인점수를 오름차순으로 정렬하여 그림으로 나타낸 결과이다.

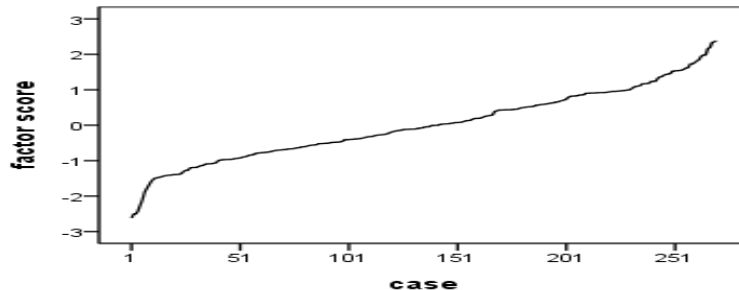


Figure 3.1 Factor score diagram

Figure 3.1을 이용하면 규정투구이닝 수를 채운 한국프로야구 투수들의 경기결과는 대략 급격한 요인점수의 기울기 변화가 있는 요인점수 -1 이하를 1개의 집단, 요인점수 2 이상을 1개의 집단, 완만한 요인점수의 기울기 변화가 있는 요인점수 -1부터 2까지 집단을 주관적으로 1개에서 3개 정도의 집단으로 나눌 수 있을 것 같기 때문에 전체 집단을 3개에서 5개의 그룹으로 나눌 수 있겠다. 아울러 S 는 대략 -3부터 3사이의 값을 가지며, 투수능력을 우수한 순서로 그룹번호를 G_1, G_2, G_3, \dots 로 나타낼 때, 다음 Table 3.4는 각 그룹에 속한 선수들의 빈도수를 나타낸다. 그룹의 분할은 K -평균 군집분석을 이용하였으며, 이 경우 K 는 군집의 개수를 나타낸다. 만일 군집의 개수를 3개로 하면 군집의 특색이 별로 없고 5개로 하면 최우수 군집에 속하는 선수들의 수가 너무 적다. 따라서 군집의 개수는 4개가 현실적으로 타당하다고 간주되며 Table 3.5는 군집수의 개수를 4개로 한 경우의 주요 투수 통계량의 평균값을 군집 별로 나타낸 결과이다. 운을 측정하는 BABIP을 제외하고 모든 통계량들이 상식적인 순위로 정렬되어 있다.

Table 3.4 Cluster distribution table

K	G_1	G_2	G_3	G_4	G_5
3	78	108	84		
4	41	106	84	39	
5	9	28	75	73	85

그런데 우수한 투수는 $K/9$ 은 크며, ERA, BB/9, WHIP, DICE, FIP은 작다. 따라서 뛰어난 투수일 수록 요인점수 S 는 작은 값을 가지게 되는데, 사용의 편리함을 위하여 투수등급지표 (pitching grade index; PGI)를 $PGI = -S$ 로 정의하기로 한다. 따라서 PGI 값이 클수록 우수한 투수이며, Table 3.6은 PGI를 이용하여 구한 상위 10위에 속하는 선수들이다. 그리고 Table 3.7은 PGI 기준에 따른 연도별 최우수 투수이다.

Table 3.5 Mean of pitching statistics by groups

Group	W	L	ERA	$K/9$	BB/9	WHIP	BABIP	DICE	FIP
G_1	13.32	6.95	2.78	7.50	2.67	1.14	0.24	3.12	3.19
G_2	11.05	7.98	3.46	6.33	2.86	1.27	0.25	3.70	3.75
G_3	10.60	8.31	4.11	5.95	3.55	1.40	0.25	4.28	4.32
G_4	8.56	9.54	4.73	5.51	4.16	1.51	0.24	4.86	4.90

Table 3.6 A player listing of the top 10 PGI scores

Rank	Name	Year	Team	PGI
01	Ryu hyeonjin	2010	Hanhwa	2.620
02	Ryu hyeonjin	2012	Hanhwa	2.510
03	Ryu hyeonjin	2006	Hanhwa	2.501
04	Yun seokmin	2011	KIA	2.419
05	Park myeonghwan	2004	Doosan	2.258
06	Yun seokmin	2012	KIA	2.108
07	Greisinger	2006	KIA	1.897
08	Rios	2007	Doosan	1.791
09	Yun seokmin	2008	KIA	1.694
10	Bae yeongsu	2005	Samsung	1.595

Table 3.7 A best player of the year by PGI scores

Year	Name	Team	PGI
2000	Harriger	LG	1.448
2001	Park sukjin	Lotte	0.971
2002	Song jinwoo	Hanhwa	1.271
2003	Kim jinwoo	KIA	0.891
2004	Park myeonghwan	Doosan	2.258
2005	Bae yeongsu	Samsung	1.595
2006	Ryu hyeonjin	Hanhwa	2.501
2007	Rios	Doosan	1.791
2008	Yun seokmin	KIA	1.694
2009	Lopez	KIA	1.411
2010	Ryu hyeonjin	Hanhwa	2.620
2011	Yun seokmin	KIA	2.419
2012	Ryu hyeonjin	Hanhwa	2.510
2013	Lee jaehak	NC	1.068

PGI는 거의 투수 자신만의 실력에 관한 지표라고 할 수 있는데, Table 3.6을 참고하면 2000년대 한국 프로야구는 투수부분에서 류현진과 윤석민, 2명으로 대변된다고 할 수 있다. 그 중 류현진은 소속팀의 전력이 약체임에도 불구하고 발군의 기량을 갖춘 투수로 평가된다. PGI의 상위 3개 기록이 류현진이며 가장 높은 기록도 류현진 선수가 세운 2.62로 나타났다. PGI 2위인 2012년 류현진 기록은 고전적인 승패기준으로는 겨우 9승 9패였으며 평범한 기록이었다. 한편 PGI 값이 2를 초과할 확률은 규정투구이닝 수를 채운 투수 중에서 2.2% 정도로 나타났다.

4. 결론

투수가 공을 던짐으로써 야구는 시작되는 스포츠다. 그만큼 투수의 비중이 크고 중요하다. 따라서 프로야구 각 구단의 투수들도 제대로 평가가 이루어져야 하며, 특히 큰 활약을 한 선수나 저평가된 선수들을 배려해야 할 것이다. 하지만 투수를 평가하는 지표들은 특성과 중요성이 각각 여러 부분에서 나타나고 중요하지 않은 지표가 거의 없기 때문에 종합적으로 판단하는 것은 매우 주관적인 문제라고 간주된다.

본 연구에서는 지금까지 사용된 핵심적인 7개의 투수평가지표들을 이용하여 통합투수평가지표인 PGI를 제안했다. 팀의 수비능력이나 운과 같은 요인들을 최대한 배제한 순수한 투수의 능력을 측정하는 PGI는 주성분분석을 이용하여 작성되었으며, K -평균 군집분석을 활용하여 한국프로야구 투수력을 4개의 군집으로 나누었다. 본 연구의 PGI는 보다 객관적이고 타당한 투수들의 연봉 책정 도구가 되리라 여겨지며, 현대야구의 특성 중 하나가 투수의 분업화인데 분업화된 투수에 대해 통합된 투수지표를 만드는 것도 향후 연구과제로 의미가 있으리라 간주된다.

References

- Chea, J. S., Cho, E. H. and Eom, H. J. (2010). Comparisons of the outcomes of statistical models applied to the prediction of post-season entry in Korean professional baseball. *The Korean Journal of Measurement and Evaluation in Physical Education and Sport Science*, **12**, 33-48.
- Kim, H. J. (2012). Effects of on-base and slugging ability on run productivity in Korean professional baseball. *Journal of the Korean Data & Information Science Society*, **23**, 1065-1074.
- Korea Baseball Organization (2001-2006). *2000-2005 official baseball guide*, Korea Baseball Organization, Seoul.
- Korea Baseball Organization (2013). <http://www.koreabaseball.com/Record>.
- Lee, J. T. and Bang, S. Y. (2010). Forecasting attendance in the Korean professional baseball league using GARCH models. *Journal of the Korean Data & Information Science Society*, **21**, 1041-1049.
- Seung, H. B. and Kang, K. H. (2012). A study on relationship between the performance of professional baseball players and annual salary. *Journal of the Korean Data & Information Science Society*, **23**, 285-298.

Pitching grade index in Korean pro-baseball

Jang Taek Lee¹

¹Department of Applied Statistics, Dankook University

Received 22 January 2014, revised 12 February 2014, accepted 12 March 2014

Abstract

In baseball, the traditional measure of pitchers are wins and ERA. But these statistics are influenced by luck or team power. So sabermetrician proposes a number of indicators that predict future performance. We determine a new measure, which we call pitching grade index (PGI) that efficiently summarizes a pitcher's performance on a numerical scale using principal components analysis. The PGI statistic can often be useful to assessing a pitcher's individual contribution. Also K-means clustering algorithm are used for segmentation of players into groups.

Keywords: K-means clustering, pitching grade index, principal components analysis, sabermetrician.

¹ Professor, Department of Applied Statistics, Dankook University, Yongin 448-701, Korea.
E-mail: jtlee@dankook.ac.kr