

# The prediction of the Korean Basketball League (KBL) winning percentage: Applying the possession concept

Jiho Ryu<sup>1</sup>, Jin-Wook Han<sup>1\*</sup>, & Minsoo Kim<sup>2</sup>

<sup>1</sup>The Kyung Hee University & <sup>2</sup>Korea Institute of Sport Science

**[Purpose]** The purpose of this study was to predict winning percentage in the Korean Basketball League(KBL) by applying the possession concept. **[Method]** The model to estimate possession was utilized by formula used in NBA(National Basketball Association). Data consist of 20 seasons of the KBL (1997-2017). For data analysis, multiple regression analysis, Fisher's Z transformation and stepwise multiple regression analysis were performed using SPSS 22.0. **[Results]** The result indicated that average team stats per possession had more explanatory power in predicting the KBL teams' winning percentage than average team stats per game. The most important factor for winning in the KBL was defensive rebounds and followed by 2-point field goal percentage, steals, 3-point field goal percentage, free throw attempts, turnovers, offensive rebounds, blocked shots, free throw percentage, and assists. The results of this study provided fundamental information for the data analysis of Korean basketball games. It might be useful for basketball coaches to manage and instruct their teams. **[Conclusions]** Practical implications and future research direction were also suggested.

**Key Words:** Basketball, Possession, Basketball analysis, Prediction of Winning Percentage, KBL

## 서 론

### 1. 연구의 필요성 및 목적

한국보다 스포츠가 먼저 발전한 미국에서는 일찍부터 스포츠 통계에 대한 연구가 활발하게 진행되어 왔다. 특히 야구는 150년이 넘는 역사만큼이나 경기 기록에 대한 분석도 오래되었는데, 누적된 경기 자료를 바탕으로 통계적 관점에서 야구를 분석하는 세이버메트릭스(sabermetrics)는 야구를 바라보는 관점에 있어서 새로운 지평을 연 것으로 평가된다(Lee & Kim, 2016).

세이버메트릭스는 다른 종목에도 영향을 미쳐 미국에서는 농구 경기 데이터를 활용한 양적 분석이 21세기에 들어서 본격적으로 시작되었다(Baumer & Zimbalist, 2014). 농구 데이터 분석은 여러 전문가들을 통해 발전을 거듭해왔는데, 그 과정에 있어서 포제션(possession)이라는 개념은 그 중심에 있다.

농구 경기에서 포제션은 한 팀의 '볼 소유권', '공격권', '공격기회'로 볼 수 있다. 포제션은 한 팀이 볼을 소유했을 때 시작하고, 그 소유를 잃었을 때 종료된다. 포제션의 종료는 (1) 필드골 성공이나 자유투의 성공, (2) 수비 팀의 리바운드, (3) 턴 오버 상황에서 발생한다. 농구 경기는 상대보다 많은 득점을 하는 팀이 승리하고, 득점을 하기 위해서는 공격기회를 얼마나 효율적으로 활용하느냐가 중요하다. 따라서 농구 경기 데이터를 분석함에 있어서 포제션은 매우 핵심적인 역할을 한

논문 투고일 : 2018. 01. 04.

논문 수정일 : 2018. 02. 13.

게재 확정일 : 2018. 02. 26.

\* 교신저자 : 한진욱(hjw5893@khu.ac.kr).

다(Kubatko, Oliver, Pelton, & Rosenbaum, 2007).

일찍이 저명한 농구 코치인 Dean Smith는 경기를 분석함에 있어서 포제션당 득점과 실점으로 경기에 대한 분석을 해야 한다고 주장하였다(Smith, 1981). Smith(1981)는 3점 슈트가 없고 2점 슈트만 있던 시절에 공격에서 완벽한 경기는 한 번의 포제션 당 2점을 올리는 것으로 보았고, 수비 측면에서의 완벽한 경기는 0점을 헌납하는 것으로 보았다. 이러한 그의 주장은 경기 템포에 따라 최종 스코어는 얼마든지 달라질 수 있으므로 최종 스코어로 팀의 공격과 수비의 효율성을 평가하기 어렵다는 판단이었기 때문이다(Smith, 1981).

Smith(1981)의 주장처럼 최근에 농구 데이터 분석에서 포제션의 중요성이 더욱 강조되고 있다. Kubatko et al.(2007)의 연구에서는 포제션은 야구의 아웃카운트와 비슷한 개념이며, 농구 경기 데이터 분석의 기초를 마련한다고 주장했다. 공수교대 방식으로 이루어진 야구에서 아웃카운트를 기준으로 양 팀에 동일한 공격기회를 제공하지만 농구의 경우 시간제한 방식을 택하고 있어서 매 경기 양 팀이 얻는 공격기회의 횟수가 다르다. 그동안 야구에서 데이터 분석이 발달한 이유로 플레이마다 선수들이 독립적이고 수비와 공격이 명확히 구분되는 특징을 지적한다(Baumer & Zimbalist, 2014). 따라서 포제션을 기준으로 하는 농구 경기 데이터 분석은 한 번의 기회에서 창출하는 가치를 탐색하는 것으로서 농구 경기 데이터 분석에 중심 개념이 될 수 있다(Kubatko et al., 2007).

하지만 미국의 농구 경기 데이터 분석 방법의 흐름과 달리 국내의 농구 기록 관련 연구에서는 포제션을 중심으로 하는 연구가 이루어지지 않았다. 그동안 국내의 농구 기록 관련 연구는 선수 개별 능력이나 팀의 경기력 요소에 초점을 맞춰왔다. 개별 선수에 대한 연구에서는 선수의 경기력과 연봉에 대한 연구(Han & Lee, 2003; Kim, 2009; Lee & Choi, 2002)가 주를 이루었고, 팀과 관련한 연구에서는 팀 경기력 요소와 승·패에 관한 연구(Cho & Chae 2016; Jhang, Kwak & Choi, 2015; Kim, 2008, 2010, 2012, 2013, 2015; Lee 2004; Lee & Do, 2005; Park, 2009, 2011)가 주로 수행되어져 왔다.

농구의 팀 경기력 요소와 승리에 대한 국내 연구는 방법론적인 측면에서 크게 두 개의 유형으로 나눌 수 있다. 첫째, 개별 경기 결과에 따라 승리 팀과 패배 팀을 나누고 팀 경기력 요소를 살피는 방법이 있다(Kim 2008, 2012, 2015; Lee & Do, 2005; Park, 2009, 2011). 이 방법은 개별 경기를 분석하기 때문에 승·패를 결정짓는 요소를 파악하기에 용이하다. 둘째, 시즌 기록을 바탕으로 승률에 미치는 영향 요인을 분석하는 방법(Lee, 2004)으로 전체 시즌의 맥락에서 장기적으로 승률에 미치는 영향 요인을 파악하기에 유용하다.

Lee(2004)는 한국남자프로농구의 팀 승률에 미치는 영향 요인에 대한 연구에서 각 팀이 일정 기간 동안 거둔 누적기록의 평균과 승률을 적용하여 팀별 평균기록이 승률을 예측함에 있어 중요한 변수임을 밝혔다. 이러한 연구에서는 팀별 평균기록이 팀 간 동일한 조건에서 얻어진 결과인지를 고려해야 하는데, 기존의 연구에서는 경기수를 그 기준으로 삼았다. 그러나 경기 수가 같은 것만으로는 동일한 조건에서 얻어진 데이터라고 보기에는 한계가 있다. 이러한 한계의 주된 이유는 포제션 개념이 적용되지 않았기 때문이다. 한 팀의 포제션 수치는 그 팀의 공격기회가 얼마만큼 있었는가를 의미한다. 그러므로 평균적으로 포제션 수치가 높은 팀은 그렇지 않은 팀에 비해 공격기회가 많았다는 것을 의미하며, 이는 경기 중에 공수전환이 많이 이루어졌다고 볼 수 있다. 반대로 평균적인 포제션 수치가 낮은 팀은 상대적으로 공격기회가 적었다는 것을 의미하고, 경기의 공수전환이 적었다고 해석할 수 있다. 자연히 포제션 수치가 높은 팀은 그렇지 않은 팀에 비해 경기 기록 지표상으로도 더 큰 수치를 기록하게 된다. 그러나 그 수치는 포제션이 많았기 때문에 더 높게 나타난 것이지 절대적인 능력을 의미하지 않는다. 따라서 포제션의 차이가 경기 중 발생하는 기록에 영향을 미칠 수 있으므로 경기 기록 분석에 있어 팀별 포제션을 고려하는 것이 중요하다.

Smith(1981)와 Kubatko et al.(2007)의 연구에서 포제션의 중요성을 주장한 이유는 포제션을 바탕으로 한 기록이 일반적 상황에서의 누적기록이나 경기당 평균기록에 비해 경기 기록이 의미하는 실질적 가치를 나타내기 때문이다. 그러나 그동안의 국내의 농구 경기 데이터 관련 연구에서는 경기 결과에 따른 누적기록이나 경기당

평균기록을 사용했지만 이러한 기록들은 효율성에 대한 정보를 제공하지는 못한다는 한계가 있다. 그러므로 이처럼 포제션 기준 기록이 실제적 효율성을 내포한다는 점에서 경기당 평균기록과 포제션당 평균기록 중에 어떠한 것이 승률을 설명함에 있어서 적합한지를 살펴볼 필요성이 있다. 포제션당 팀 평균기록은 한 팀의 공격기회당 발생한 기록으로써 공격기회를 얼마나 효율적으로 사용했는지 나타낸다. 따라서 포제션당 팀 평균기록이 경기당 팀 평균기록보다 승률에 대해 높은 설명력을 보인다면, KBL 경기 기록을 분석함에 있어서 포제션당 팀 평균기록의 활용을 우선적으로 고려해봐야 할 것이다.

추가적으로 팀별 포제션당 평균기록이 경기당 평균기록에 비해 승률을 예측함에 있어 상대적으로 높은 설명력을 갖는다는 것이 밝혀진다면, 포제션당 팀 평균기록을 구성하는 요소 중 어떠한 지표가 승률에 미치는 영향이 강한지도 살펴보아야 할 것이다. 따라서 본 연구의 목적은 첫째, 포제션이라는 개념을 적용하여 팀 승률을 예측함에 있어서 경기당 팀 평균기록과 포제션당 팀 평균기록의 상대적인 설명력을 비교하고 둘째, 도출된 팀 승률에 영향을 미치는 요인들을 찾는 데 있다. 연구목적 달성을 위해 다음과 같은 연구문제를 수립하였다.

연구문제 1. 경기당 팀 평균기록과 포제션당 팀 평균기록 중 어떠한 기록이 팀 승률을 예측하기에 더 적합한 기록인가?

연구문제 2. '연구문제1'을 통해 도출된 결과에 따라 경기당 팀 평균기록 혹은 포제션당 팀 평균기록의 어떠한 변수가 팀 승률에 영향을 미치는가?

## 연구방법

### 연구대상

포제션 개념을 적용하여 한국남자프로농구의 승률을 예측하기 위한 목적으로 한국남자프로농구가 시행된 전체 시즌 중에서 10개의 팀이 참가한 1997-1998시즌부터 2016-2017시즌의 자료를 연구대상으로 하였다.

### 연구절차

남자프로농구 KBL 홈페이지에서 10개 팀의 20시즌(1997-2017)의 자료를 수집하였다. 수집한 자료를 토대로 포제션 수치를 추정하는 공식을 통해 시즌별 각 팀의 포제션 수치를 산출하였다. 포제션을 산출하는 방법은 농구관련 데이터를 다루는 전문가들에 따라 차이가 있으나 본 연구에서는 미국프로농구 NBA 홈페이지(www.nba.com)에서 제공하는 포제션 산출 공식을 활용하였다. NBA에서 사용하는 이 공식은 실제 포제션 수치와 유사한 값을 나타내고, 경기 결과를 나타내는 기본적인 지표만을 활용하여 수식에 대입하기 때문에 사용이 용이하다는 장점이 있다(Kubatko et al., 2007). 본 연구에서 사용되는 포제션 산출 공식은 다음과 같다.

$$POSS_t = FGA_t + 0.44 \times FTA_t - OR_t + TO_t$$

$POSS_t$  = 팀 포제션

$FGA_t$  = 팀 필드골 시도

$FTA_t$  = 팀 자유투 시도

$OR_t$  = 팀 공격 리바운드

$TO_t$  = 팀 턴오버

1차 자료 수집을 통해 각 팀의 시즌별 필드골 시도, 자유투 시도, 공격 리바운드, 턴오버 수치를 얻을 수 있었다. 이러한 기록을 위 공식에 적용하여 팀별 포제션 값을 산출하였다. 산출된 포제션 값을 바탕으로 포제션당 팀 평균기록을 구했다.

### 연구변수

본 연구에서 수립된 연구문제를 검증하기 위해 변수로 사용되는 경기기록 지표들은 2점슛 성공률(2P%), 3점슛 성공률(3P%), 자유투 성공률(FT%), 자유투 시도(FTA), 공격 리바운드(OR), 수비 리바운드(DR), 어시스트(AST), 스틸(STL), 블록슛(BLK), 턴 오버(TO) 등이다. 이들 변수는 한국프로농구 승패결정요인에 대한 다년간의 국내 연구에서 공통적으로 사용된 변수들(Kim, 2013)과 승률의 영향 요인을 분석한 선행연

구(Lee, 2004)에서 사용된 변수들을 중심으로 본 연구의 목적에 적합하게 선정되었다. 팀 실책, 수비성공, 팀 속공 등의 변수는 연구대상의 기간에서 공통적으로 집계되지 않았기 때문에 변수에서 제외되었다.

다만, 기존의 관련 연구에서 변수로 고려되지 않았던 자유투 시도는 자유투가 가지는 특성에 근거하여 연구자의 판단에 따라 연구변수로 포함되었다. 자유투를 획득한다는 것은 상대팀에게 반칙을 안김과 동시에 우리 팀에게는 득점할 기회가 생긴다는 이점이 있기 때문에 처해진 상황에 따라 상대적으로 자유투를 유도하는 전략을 사용하는 팀들이 많기 때문이다. 경기당 팀 평균기록과 포제션당 팀 평균기록의 승률에 대한 설명력 비교에서는 슛 성공률이라는 기록은 시도한 슛에 대한 성공한 슛의 비율이므로 포제션당 기록으로 환산하더라도 그 비율의 값은 변하지 않기 때문에 2점슛 성공률, 3점슛 성공률, 자유투 성공률은 변수로 포함되지 않았다.

## 자료 분석

본 연구에서 자료 정리와 포제션 수치 산출 및 포제션당 평균기록 환산은 Excel 2013로 정리했으며, 통계처리에는 SPSS 22.0을 사용하여 분석하였다. 경기당 팀 평균기록과 포제션당 팀 평균기록 중 어떠한 지표가 팀 승률에 대한 설명력이 높은지를 규명하기 위해 다중회귀분석 및 Fisher의 Z 변환을 통한 검증을 실시하였다. 앞선 분석을 통해 도출된 결과에 따라 경기당 팀 평균기록 혹은 포제션당 팀 평균기록의 어떠한 요소가 팀 승률에 영향을 미치는지 알아보기 위하여 단계적 다중회귀분석(stepwise multiple regression)을 실시하였다.

## 연구결과

### 1. 경기당 팀 평균기록과 포제션당 팀 평균기록의 승률에 대한 설명력 비교

KBL의 경기당 팀 평균기록과 포제션당 팀 평균기록의 승률에 대한 설명력을 비교하기 위해 스틸, 리바운

드, 블록슛, 어시스트 등의 경기당 팀 평균기록과 포제션당 팀 평균기록을 각각 독립변수로 투입하고 팀 승률을 종속변수로 설정한 다중회귀분석을 실시하였다. 각 기록에 대한 다중회귀분석을 실행한 다음 각 회귀식에서 산출된 설명력(R)을 Fisher의 Z변환 방식으로 비교하여 어떠한 기록이 승률을 예측함에 있어 설명력이 높은지를 검증하였다.

다중회귀분석을 시행하기 전에 본 연구의 자료가 회귀분석에 적합한 자료인지를 살펴보기 위해 자료의 사례수와 독립변수의 선형성, 다중공선성 위반 여부, 잔차의 독립성, 잔차의 등분산성, 잔차의 정규성의 만족 여부를 확인하였다.

첫째, 자료의 사례 수는 일반적으로 독립변수 개수의 10배 이상을 권장한다(Nam, 2015). 분석에 있어 사용된 경기 지표는 7개이며, 표본의 수는 10개 팀의 20시즌 기록이므로 200개가 되어 사례 수를 충족하였다.

둘째, 독립변수의 선형성과 다중공선성 위반의 여부는 각각 산점도 행렬과 분산팽창요인(VIF)을 통해 파악했다. 산점도 행렬을 분석한 결과 종속변수인 승률에 대해 각 독립변수가 선형성을 띄는 것이 확인되었다. VIF 값은 경기당 팀 평균기록이 1.127부터 2.2024 사이의 값을, 포제션당 팀 평균기록이 1.089부터 1.687 사이의 값을 나타내며 10을 넘기지 않아 다중공선성의 가정을 만족하는 것으로 나타났다.

셋째, 잔차의 독립성, 등분산성, 정규성은 각각 Durbin-Watson 지수, 표준화 등분산 그래프, P-P plot로 가정의 만족 여부를 확인하였다. 잔차의 독립성에 있어서 Durbin-Watson 지수는 경기당 팀 평균기록이 1.857, 포제션당 팀 평균기록이 1.930으로 모두 2에 가까우므로 자기상관이 없이 독립적이다. 표준화 등분산 그래프에서는 표준화 잔차의 값이 -3부터 3 사이에 위치하고 직사각형을 이루고 있어 잔차의 등분산성이 만족되는 것으로 나타났다. 잔차의 정규성은 P-P plot 상에서 산점도가 45도의 직선상에 밀집하게 분포되어 정규성을 만족하는 것으로 확인되었다.

본 연구에서 경기당 팀 평균기록과 팀 승률, 포제션당 팀 평균기록과 승률에 대한 다중회귀분석을 실시한 결과는 다음과 같다.

Table 1. Comparisons of Two Models in terms of the Explanatory Power

	<i>R</i>	<i>R</i> <sup>2</sup>	<i>adj R</i> <sup>2</sup>	Durbin-Watson
Stats per game	.648	.419	.398	1.857
Stats per possession	.761	.580	.564	1.930

경기당 팀 평균기록(경기당 자유투시도, 경기당 공격 리바운드, 경기당 수비 리바운드, 경기당 어시스트, 경기당 스틸, 경기당 블록슛, 경기당 턴 오버)은 승률의 총 분산 중 약 41.9%를 설명한 반면, 포제션당 팀 평균기록(포제션당 자유투시도, 포제션당 공격 리바운드, 포제션당 수비 리바운드, 포제션당 어시스트, 포제션당 스틸, 포제션당 블록슛, 포제션당 턴 오버)은 승률의 총 분산의 약 58%를 설명하는 것으로 나타났다.

경기당 팀 평균기록과 포제션당 팀 평균기록의 회귀식 설명력을 통계적으로 검증하기 위해 Cohen & Cohen(1983)의 제안에 근거하여 Fisher's Z transformation을 활용하여 두 회귀식의 R을 비교한 결과 포제션당 팀 평균기록의 설명력이 통계적으로 더 높은 것으로 나타났다( $Z=2.2504$ ,  $p<.05$ ). 따라서 포제션당 팀 평균기록이 승률을 예측함에 있어서 더 적합하다는 것이 규명되었다.

## 2. 포제션당 팀 평균기록의 요인이 팀 승률에 미치는 영향

경기당 팀 평균기록보다 포제션당 팀 평균기록이 팀 승률을 예측함에 있어서 보다 효과적인 지표로 밝혀졌다. 이러한 결과를 바탕으로 포제션당 팀 평균기록의 승률 영향 요인을 알아보고자 단계적 회귀분석을 실시했으며, 실시에 앞서 회귀분석 기본가정의 만족을 파악했다.

첫째, 표본의 크기에 있어서 표본은 200개로 10개의 독립변수의 10배 이상이므로 표본의 크기는 만족되는 것으로 나타났다. 또한, 독립변수의 선형성 만족 여부는 산점도 행렬로 파악한 결과 만족하는 것으로 나타났으며, 다중공선성은 VIF로 판별한 결과 그 값이 1.107에

서 1.774 사이에서 형성되므로 10미만의 값을 나타내 다중공선성의 가정을 만족했다.

둘째, 잔차의 독립성은 Durbin-Watson 지수로 확인하였으며, 2에 근접한 1.640으로 자기상관이 없는 것으로 나타났다. 한편, 잔차의 등분산성은 표준화 등분산 그래프로 확인했으며 분석결과 3 이상의 극단값이 발견되었고, 극단값이 발생한 문제를 해결하기 위하여 해당 케이스(2009-2010시즌 창원 LG 세이커스)가 입력 오류의 여부인지 점검하고, 입력 오류가 아닌 것으로 나타나 극단값을 제거한 후에 다시 분석한 결과 이후에는 극단값이 발견되지 않았다. 따라서 잔차의 등분산성이 최종적으로 성립되었다. 잔차의 정규성은 P-P plot을 통해 살펴보았으며 산점도가 45°의 직전상에 밀접하게 분포되어 잔차의 정규성이 만족되는 것으로 나타났다.

포제션당 팀 평균기록이 팀 승률에 미치는 영향을 확인하기 위하여 10개의 독립변수와 팀 승률을 종속변수로써 단계적 회귀분석을 실시하였다. <Table 2>와 같이 10단계의 모든 모형에서 유의확률이  $p<.001$ 로 나타나 회귀모형이 유의미한 것으로 확인되었으며, 제외된 독립변수가 없이 모든 변수는 팀 승률에 대해 유의한 영향을 주는 것으로 나타났다.

Table 2. The result of stepwise multiple regression

	STEP	<i>B</i>	S.E.	$\beta$	<i>t</i>	<i>p</i>
		-3.590	.228		-15.732	.000
DR	1	4.112	.350	.551	11.752	.000
2P%	2	1.839	.223	.367	8.237	.000
STL	3	4.474	.480	.356	9.330	.000
3P%	4	1.674	.278	.245	6.019	.000
TO	5	-2.200	.400	-.203	-5.495	.000
FTA	6	1.109	.215	.213	5.149	.000
BLK	7	2.320	.569	.159	4.077	.000
OR	8	1.812	.409	.197	4.432	.000
FT%	9	.650	.164	.152	3.953	.000
AST	10	.590	.242	.106	2.438	.016
$R^2 = .767$ $F = 62.021$						

각 독립변수가 종속변수에 대해 어느 정도 영향을 주는지는 비표준화 계수(B)의 값을 통해 살펴볼 수 있고, 이를 통해 다음과 같이 한국남자프로농구 팀 승률 예측의 회귀방정식을 세울 수 있다.

$$\begin{aligned} \text{팀 승률} = & -3.590 + 4.112 \times DR + 1.839 \times 2P\% \\ & + 4.474 \times STL + 1.674 \times 3P\% - 2.200 \times TO \\ & + 1.109 \times FTA + 2.320 \times BS + 1.812 \times OR \\ & + .650 \times FT\% + .590 \times AST \end{aligned}$$

단계적 다중회귀분석에서 최종 모형인 10단계의 회귀모형에서 설명력은 76.7%를 나타냈고,  $F$  값은 62.021( $p < .001$ )로 제시되어 이 회귀식은 통계적으로 유의한 것으로 나타났다. 회귀모형에서 독립변수의 상대적 중요도는 표준화 계수인 베타( $\beta$ )의 절댓값 크기에 따라 판단한다. 그러므로 KBL 팀 승률을 예측함에 있어서 경기 지표의 상대적 중요도는 수비 리바운드, 2점슛 성공률, 스틸, 3점슛 성공률, 자유투 시도, 턴 오버, 공격 리바운드, 블록슛, 자유투 성공률, 어시스트 순서로 나타났다.

## 논 의

본 연구 결과에 따르면 포제션당 팀 평균기록이 경기당 팀 평균기록보다 승률의 분산을 약 16.1% 더 많이 설명할 수 있으며, 통계적으로도 유의미한 차이가 있는 것으로 나타났다. 이는 한 팀의 경기 기록 지표를 통해 승률을 예측할 경우 경기당 팀 평균기록으로 승률을 예측하는 것보다 포제션당 팀 평균기록으로 예측하는 것이 보다 효과적임을 의미한다. 따라서 '경기 템포에 따라 최종 스코어와 기록은 얼마든지 달라질 수 있으므로 포제션당 득점과 실점으로 경기에 대한 해석을 해야 한다'(Smith, 1981)라는 주장과 같은 맥락으로 볼 수 있다.

경기당 팀 평균기록과 포제션당 팀 평균기록이 차이가 발생하는 이유는 포제션의 개념적 특성에 기인한다. 포제션은 한 번의 공격기회를 의미하므로 포제션당 팀 평균기록은 한 팀이 한 번의 공격기회당 얻어낸 기록이다. 그러므로 경기당 팀 평균기록은 한 팀의 경기력 이

외에도 경기를 진행하는 속도에 따라 공격기회의 양이 얼마든지 달라질 수 있고, 그 차이가 경기당 팀 평균기록에 영향을 미친다. 반면에 포제션당 팀 평균기록은 공격기회를 기준으로 삼기 때문에 경기를 진행하는 속도를 보장하여 한 팀의 경기력을 보다 명확하게 나타낸다. 즉, 포제션당 평균기록은 경기의 템포와 무관한 객관적인 데이터를 얻을 수 있으며(Ibáñez et al., 2008), 포제션당 평균기록은 포제션이라는 개념에서 출발하기 때문에 농구 경기 데이터 분석에서 포제션 개념이 기초가 된다(Kubato et al., 2007). 그러므로 팀의 경기력 요소를 객관적으로 평가할 때, 경기당 평균기록보다는 포제션당 평균기록을 기준으로 삼는다면 현장에서는 더욱 효과적인 경기 분석을 수행할 수 있을 것이다. 게다가 팀 기록은 선수들 개별 기록의 합으로써 선수의 기록 역시 자신이 소속된 팀의 포제션의 횟수에 영향을 받는다. 따라서 선수의 기록 또한 포제션을 기준으로 하는 기록으로 나타내면 선수 기량에 대한 객관적인 정보를 제공할 것이다.

이러한 정보는 특히 프로구단의 선수단 구성에 있어서 합리적인 의사결정을 도울 수 있는 자료로 활용 가치가 높다. 구단에서는 선수의 기록이 선수의 연봉 책정 및 선수 수급에 있어서 큰 요소로 작용하기 때문이다. 따라서 포제션을 적용한 경기 데이터 분석은 팀 전력에 대한 분석과 선수의 가치를 평가함에 있어서 경기당 평균기록보다 객관적인 자료로 활용될 수 있다는 것을 시사한다.

한편, 포제션당 팀 평균기록을 기반으로 승률을 예측한 결과 독립변수의 상대적 중요도는 수비 리바운드, 2점슛 성공률, 스틸, 3점슛 성공률, 자유투 시도, 턴 오버, 공격 리바운드, 블록슛, 자유투 성공률, 어시스트 순서로 나타났다. 이러한 결과는 농구 승률 및 승패에 대한 경기력 요인과 관련한 선행연구들과 유사점과 차이점을 동시에 나타내고 있다.

그 중에서 승률에 가장 큰 영향을 미치는 것으로 나타난 수비 리바운드의 경우, Kim(2013)이 기존의 25편의 농구 승패 결정 요인의 연구들을 메타분석을 활용한 연구와 유사한 결과를 도출하였다. 또한, Lee(2004)의 연구에서 리바운드가 승률을 예측함에 있어서 가장 큰 영향을 미치는 것을 밝혀낸 것과 비슷한 결과를 나타내

었다. 이와 같이 본 연구와 선행연구에서 수비 리바운드가 농구 경기에서 승패를 결정짓는데 중요한 요소로 꼽히는 이유는 수비 리바운드가 상대팀의 공격기회를 차단할 뿐만 아니라 우리 팀의 공격기회를 가져오는 중요한 역할을 하기 때문인 것으로 판단된다.

본 연구에서 슛 성공률에 해당하는 2점슛 성공률, 3점슛 성공률, 자유투 성공률이 승률에 유의한 영향을 미친 결과는 기존 연구(Kim, 2008; Kim, 2013; Lee, 2004)의 결과와 그 맥을 같이 한다. 또한, 득실점의 차이에 영향을 주는 경기력 요소를 살펴본 연구(Jhang et al., 2015)에서도 2점슛 성공률, 3점슛 성공률, 자유투 성공률이 승률에 긍정적인 영향을 미치는 것으로 나타나 본 연구의 결과를 지지해주고 있다. 더불어 스틸, 어시스트, 블록슛, 턴 오버는 승패 요인에 대해 분석한 다른 연구들과 비슷한 결과를 나타냈다. 스틸과 어시스트, 블록슛은 대체로 긍정적인 영향을 미치며 턴 오버의 경우에는 부정적 영향 요인으로 나타났다.

한편, 본 연구에서는 공격 리바운드가 승률에 긍정적인 영향을 미치는 것으로 나타났지만, 선행연구(Kim, 2013)에서는 공격 리바운드가 승리에 긍정적인 요소가 아닌 것으로 나타나 본 연구와 상반된 결과를 제시하였다. 이와 같이 본 연구와 기존 연구 결과에 차이가 발생하는 이유는 다음과 같은 내용으로 설명이 가능하다. 공격 리바운드는 공격 팀이 실패한 슛을 다시 획득하는 것을 의미한다. 따라서 공격 리바운드를 획득할수록 득점을 실패한 횟수가 늘어나기 때문에 패배한 팀이 공격 리바운드의 빈도가 높게 나타날 가능성이 있다고 본다(Park, 2013). 또한 팀의 전략에 따라서는 선수들이 공격 리바운드에 참여하는 대신 수비 전환을 더 우선시하여 상대방의 속공을 사전에 차단하는 방법을 선택하기도 한다(Wiens, Balakrishnan, Brooks, & Guttag, 2013). 따라서 공격 리바운드가 많다는 것은 슛 실패가 많았다는 것을 의미하기도 하고, 전략적으로 공격 리바운드를 시도하지 않는 경향이 있기 때문에 공격 리바운드가 반드시 승리에 긍정적인 요소가 아닐 수 있다. 하지만 공격 리바운드의 획득은 공격기회의 연장이며, 상대팀에게 수비 리바운드를 헌납하여 포제션을 잃는 것보다 더 나은 결과를 기대할 수 있다. 그러므로 본 연구에서 공격 리바운드가 승률에 긍정적인 영향을 끼치는 것

으로 나타난 것은 공격 리바운드의 긍정적인 측면이 더 작용한 결과로 해석할 수 있다.

## 결론 및 제언

본 연구는 기존의 농구 기록 관련 연구에서 다루지 않았던 포제션의 개념을 적용하여 한국남자프로농구의 승률을 예측하는 목적으로 수행되었으며 연구결과는 다음과 같다.

첫째, 포제션당 팀 평균기록은 경기당 팀 평균기록보다 팀 승률에 대한 설명력이 더 큰 것으로 나타났다.

둘째, 포제션당 팀 평균기록으로 팀 승률을 예측했을 때, 수비 리바운드, 2점슛 성공률, 스틸, 3점슛 성공률, 자유투시도, 턴 오버, 공격 리바운드, 블록슛, 자유투 성공률, 어시스트 순으로 영향을 미치는 것으로 나타났다.

연구문제1에서 다루었던 포제션당 팀 평균기록과 경기당 팀 평균기록의 비교에 대한 결과는 포제션을 중심으로 하는 농구 데이터 분석의 중요성을 언급한 선행연구(Kubato et al., 2007; Oliver, 2004; Smith, 1981)의 주장을 뒷받침한다. 따라서 향후 농구 데이터 분석에서 포제션을 기준으로 삼는 방향으로의 변화가 필요하다. 두 번째 연구 문제는 해당 연구와 유사한 형태의 선행연구(Lee, 2004)에서 없었던 포제션 개념을 적용하여 승률을 예측했다. 팀 경기력 요소와 승·패에 관한 연구(Cho & Chae 2016; Jhang, Kwak & Choi, 2015; Kim, 2008, 2012, 2013, 2015; Lee 2004; Lee & Do, 2005; Park, 2009, 2011)에서 나타난 결과와 비교했을 때, 포제션당 팀 평균기록에서는 크게 리바운드와 득점의 성공률의 중요하다고 나타났다.

이와 같은 본 연구의 결과는 해외의 농구 데이터 분석 동향에서 중요하게 여기는 포제션 개념을 국내에 적용했다는 점에서 시의성이 있으며 향후 스포츠 경기의 양적 분석이 중요하게 여겨지고 발전하는 만큼 국내의 농구 데이터 분석에 기초적인 자료로 활용될 수 있을 것이다. 또한, 본 연구가 이론에만 머무는 것이 아니라 남녀프로 농구리그와 구단 등에서 전략·전술을 위한 실무적인 활용이 가능하다는 점에서 활용가치가 있다고 할 수 있다.

본 연구 결과에 따라 후속 연구에서 고려해야 할 문제

를 제언하면 다음과 같다. 첫째로 본 연구에서는 KBL 웹사이트에 제공된 경기 데이터를 바탕으로 승률 영향 요인에 대해 분석하였기에 감독의 전략, 선수들의 심리 상태, 관중의 응원, 경기 일정 등 기록 외적인 요인들은 고려되지 않은 한계가 있다. 따라서 후속 연구에서는 이러한 부분을 고려한 연구가 필요하다. 둘째, 이 연구에서는 KBL의 1997-1998시즌부터 2016-2017시즌까지의 데이터만을 대상으로 하였다. 따라서 KBL의 기록이 더해지는 만큼 앞으로 진행될 연구에서는 본 연구보다 많은 표본을 토대로 추가 연구를 지속할 필요가 있다. 셋째, 본 연구의 KBL 팀별 포제션은 NBA의 포제션을 추정하는 공식을 통해 구한 값이다. 따라서 팀들이 실제로 기록한 포제션 수치와는 다를 한계가 있다. 다만, 아직까지 KBL에서 경기에 발생한 포제션을 공식적으로 집계하고 있지 않은 관계로 향후 보다 면밀한 연구를 위해서는 KBL 특성에 맞는 포제션 추정식을 고안하거나 KBL 모든 경기의 포제션을 집계할 필요가 있다. 마지막으로 국내연구에서 아직 다루지 않은 포제션을 적용한 이 연구로 말미암아 농구 경기의 데이터 분석 연구에서 포제션을 토대로 하는 후속연구가 활발히 이루어질 수 있기를 기대한다.

## 참고문헌

- Baumer, B., & Zimbalist, A. (2014). *The Sabermetric Revolution*. Philadelphia: UPenn.
- Cho, E. H., & Chae, J. S. (2016). Comparison of the outcomes of statistical models applied to the prediction of play-off entry in Korean professional basketball. *The Korean Journal of Physical Education*, 55(1), 733-745.
- Cohen, J. & Cohen, P. (1983). *Applied Multiple regression/correlation analysis for the behavioral sciences* (2nd ed.), LEA: Lawrence Erlbaum Associates, Inc., New Jersey.
- Han, P. S., & Lee, S. I. (2003). Predicting annual salary model and creating variables for professional basketball players using data in 2001-2002 season. *The Korean Journal of Physical Education*, 42(3), 477-486.
- Ibáñez, S. J., Sampaio, J., Feu, S., Lorenzo, A., Gómez, M. A., & Ortega, E. (2008). Basketball game-related statistics that discriminate between teams' season-long success. *European journal of sport science*, 8(6), 369-372.
- Jhang, H. J., Kwak, H., & Choi, S. H. (2015). Analysis of the outcome for the Korean professional basketball games using regression models. *Journal of Korean Institute of Intelligent Systems*, 25(1), 139-142.
- Kim, S. H. (2008). *Developing estimate model of victory and defeat through analyzing the record of the pro-basketball*. Unpublished Master Thesis. Korea National Sport University.
- Kim, S. H. (2012). Comparison of estimating method of victory and defeat through analyzing records of Korean professional basketball. *The Korean Journal of Sports Science*, 21(3), 1347-1360.
- Kim, S. H. (2013). A meta-analysis to estimate of victory and defeat through analyzing records of professional basketball. *The Korean Journal of Measurement and Evaluation in Physical Education and Sport Science*, 15(3), 35-53.
- Kim, S. H. (2015). Analyzing records of Korean professional basketball using general linear model. *Journal of the Korea Data & Information Science Society*, 26(4), 957-970.
- Kim, S. J. (2009). An analysis relation of the salary, state basketball game score Korea performance of the players basketball player. *The Korea Journal of Sports Science*, 18(3), 53-62.
- Kim, S. J. (2010). *Analysis on the main pattern and factor in victory and defeat in Korean Basketball League based on data mining*. Unpublished Doctoral Dissertation, Kook Min University, Seoul.
- Kubatko, J., Oliver, D., Pelton, K., & Rosenbaum, D. T. (2007). A starting point for analyzing basketball statistics. *Journal of Quantitative Analysis in Sports*, 3(3). Article 1.
- Lee, G. B. (2004). The factors of KBL team's playoff pass and winning percent. *Korean Journal of Sports Science*, 15(3), 41-50.
- Lee, J. H., & Do, J. H. (2005) Analysis of victory or defeat depend on cause of offense and defense in basketball game - Focusing on 2004 Basketball Festival. *The Journal of Physical Education*, 33, 71-79.
- Lee, J. T., & Kim, Y. T. (2006). Estimation of winning percentage in Korea pro-sports, *Journal of the Korean*

- Data Analysis Society*, 5, 2105-2116.
- Lee, J. Y., & Choi, B. M. (2002). Annual salary and ability of Korean Professional Basketball players. *Journal of Sport and Leisure Studies*, 17, 879-887.
- Lee, J. Y., & Kim, H. G. (2016). Suggestion of batter ability index in Korea baseball - focusing on the sabermetrics statistics WAR. *The Korean Journal of Applied Statistics*, 29(7), 1271-1281.
- Nam, S. S. (2015). *Statistical analysis of sports medicine and science*, Seoul: Hannarae Publishing Company.
- Oliver, D. (2004). *Basketball on paper: rules and tools for performance analysis*. Washington, D.C. : Brassey's, Inc.
- Park, D. K. (2013). *Estimating determinants of victory or defeat through analyzing contribution ability of professional basketball*. Unpublished Master Thesis, Korea National Sport University.
- Park, J. Y. (2009). The analysis of the factors for winning a game in the 29th Beijing Olympic men's basketball game, *Journal of Sport and Leisure Studies*, 37(2), 1425-1432.
- Park, J. Y. (2011). The analysis of the factors for winning a game in the 16th Asian men's basketball. *Korean Journal of Sports Science*, 20(3), 1731-1741.
- Smith, D. (1981). *Basketball: Multiple offense and defense*, San Francisco, CA: Benjamin Cummings.
- Wiens, J., Balakrishnan, G., Brooks, J., & Guttag, J. (2013, March). *To crash or not to crash: A quantitative look at the relationship between offensive rebounding and transition defense in the NBA*. The Seventh MIT Sloan Sports Analytics Conference, Boston, MA.

## 포제션 개념을 적용한 한국남자프로농구 승률 예측 분석

류지호 · 한진욱(경희대학교), 김민수(한국스포츠개발원)

**[목적]** 본 연구는 기존의 농구 기록 관련 연구에서 다루지지 않았던 포제션의 개념을 적용하여 한국남자프로농구의 승률을 예측하는데 목적이 있다. **[방법]** 한국남자프로농구 1997-1998시즌부터 2016-2017시즌의 자료를 대상으로 선정하여 분석했다. 1차 자료 수집으로 얻은 각 팀의 기록에 NBA에서 사용하는 포제션 산출 공식을 적용하여 팀별 포제션당 팀 평균기록을 산출하였다. 이러한 데이터를 바탕으로 상관분석, 다중회귀분석, Fisher의 Z 변환 및 단계적 다중회귀분석을 실시했으며 다음과 같은 결과를 도출하였다. **[결과]** 첫째, 포제션당 팀 평균기록은 경기당 팀 평균기록보다 팀 승률에 대한 설명력이 더 큰 것으로 나타났다. 둘째, 포제션당 팀 평균기록으로 팀 승률을 예측했을 때, 수비 리바운드, 2점슛 성공률, 스틸, 3점슛 성공률, 자유투 시도, 턴오버, 공격 리바운드, 블록슛, 자유투 성공률, 어시스트 순으로 영향을 미치는 것으로 나타났다. **[결론]** 이러한 연구를 고려해볼 때 농구 경기의 데이터 분석 연구에서 포제션을 토대로 하는 후속연구가 활발히 이루어고, 다양한 관점에서 농구 경기 데이터를 분석하고 해석하는 시도가 진행될 것으로 기대된다.

**주요어:** 포제션, 농구 데이터 분석, 승률 예측, KBL