

## 월드컵 축구 예제를 통한 통계교육

김혁주<sup>1</sup> · 김영일<sup>2</sup>

<sup>1</sup>원광대학교 수학·정보통계학부 및 기초자연과학연구소, <sup>2</sup>중앙대학교 정보시스템학과

(2010년 7월 접수, 2010년 9월 채택)

### 요약

확률과 통계 교육에서 학생들의 관심을 집중시켜 강의 효과를 높일 수 있는 예제들을 발굴하는 데 있어 많은 학자들의 꾸준한 노력이 뒤따라야 한다. 여러 분야의 예제를 제시할 수 있지만, 특히 다양한 예제들을 찾을 수 있는 분야가 스포츠이며, 그 중에서도 월드컵 축구는 여러 가지 흥미로운 문제들의 보고(寶庫)이다. 월드컵 축구대회는 전 국민적 관심사이기 때문에, 이 분야의 예제를 통한 학습 역시 확률과 통계에 대한 학생들의 관심을 진작시킬 수 있는 효과적인 방법이라 본다. 본 논문에서는 월드컵 축구의 조별 리그와 16강 토너먼트 등의 경기 방식과 승점제 및 조편성과 관련된 경우의 수와 확률 문제들을 제시하였으며, 통계모형을 이용한 시뮬레이션을 통해 2010 남아공 월드컵을 앞둔 시점이라 가정하고 본선 진출 각 팀의 16강 진출 확률과 우승 확률을 계산하는 논리적 방법을 제안하였다.

주요용어: 경우의 수, 확률, 승점분포, 조건부기댓값, 통계모형, 모의실험.

### 1. 서론

통계가 타학문 분야는 물론이고 정치, 경제, 산업 등 국가의 주요 분야에 중요하게 이용되고 있다는 사실은 새삼 언급할 필요가 없을 것이다. 최근에는 스포츠에서 담당하는 역할도 커지고 있는데, 이를테면 (1) 경기를 앞두고 최적의 전략을 수립하거나, (2) 선수의 최적 컨디션을 얻는 데 통계가 이용되고 있다. 또한 (3) 경기 또는 대회 결과의 예측하거나, (4) 각종 규정과 기록의 합리성을 추구하는 데에도 통계적 개념과 방법론이 큰 역할을 한다. 그리고 이 통계의 이론적 바탕이 되는 것은 경우의 수와 확률이다.

따라서 확률과 통계 교육에서 학생들의 이해를 돕기 위하여 제시할 수 있는 적절한 예제들이 스포츠 분야에서 많이 나오는 것은 자연스럽다. 특히 월드컵 축구는 대회 시작 전의 조 추첨에서 우승팀 결정에 이르기까지 모든 단계가 경우의 수 및 확률·통계와 깊이 연관되어 있다. 본 연구에서는 확률분야의 경우의 수 모형으로 설명할 수 있는 몇 가지의 스포츠 예제들을 월드컵 축구를 중심으로 소개한다. 또한 통계 모형을 이용한 시뮬레이션을 통해 2010 남아공 월드컵을 앞둔 시점이라 가정하고 본선 진출 팀 각각의 16강 진출 확률과 우승 확률을 계산하는 논리적 방법을 제안하였다.

통계적 기법을 사용하여 월드컵을 예측하는 연구는 국내의 일부 학자들에 의해서도 수행된 바 있다. 강명욱과 김영일 (2002)은 모의실험을 이용하여 2002년의 한·일 월드컵 예측을 처음으로 시도하였으며, 김도현 등 (2007)은 Bradley-Terry 모형을 이용하여 2006 독일 월드컵을 예측하였다. 또한 홍종선 등 (2010)은 Bradley-Terry 모형과 Newton-Raphson 알고리즘을 사용하여 2010 남아공 월드컵을 예측하

이 논문은 2008년도 원광대학교의 교비 지원에 의해서 수행됨.

<sup>2</sup>교신저자: (456-756) 경기도 안성시, 대덕면 내리 산 40-1, 중앙대학교 산업과학대학 정보시스템학과, 교수.

Email: yik01@cau.ac.kr

였다. 축구이외의 예제로 국내에서 다룬 논문은 테니스를 주제로 한 조대현 (2008, 2009)등을 들 수 있으나 외국의 경우와는 달리 발표되는 논문 편수는 매우 미미한 편이다. 일반적인 스포츠모형을 다룬 책으로는 Winston (2009)을 들 수 있다. 이 책 역시 축구예제는 들고 있지 않는데, 세계적으로 인기 있는 스포츠 종목인 축구에 관련된 통계논문이 많지 않다는 점은 특이한 점이다.

본 논문의 2절에서는 경우의 수 및 확률 관련 예제들을, 3절에서는 기초 통계학의 지식만 가지고도 구성할 수 있는 통계모형을 제안하여 보았다. 그리고 4절에서는 향후 스포츠 분야에서의 통계교육의 필요성을 언급하였다.

**2. 경우의 수와 확률 관련 예제들**

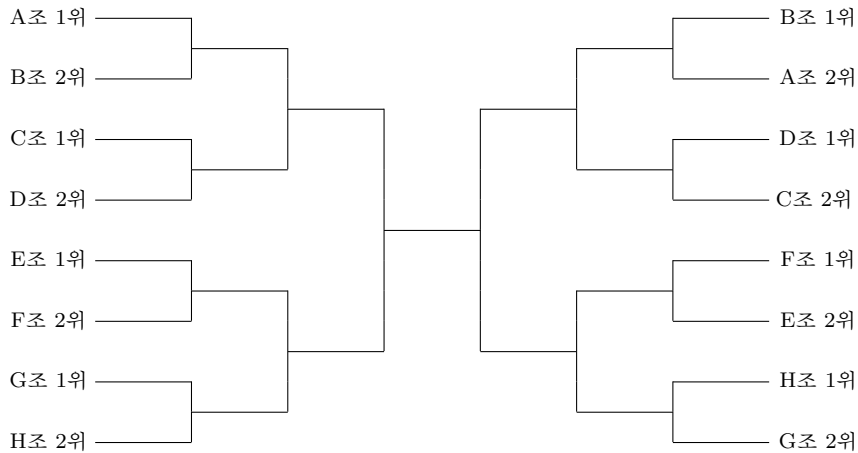
**2-1.** 월드컵 축구대회의 경기 방식은 다음과 같다. 32개 팀이 참가하여 4개 팀씩 8개조로 나뉜다. 각 조에서 4개 팀이 풀리그(다른 팀과 전부 한 번씩 경기)를 벌여 각 조의 1, 2위 팀이 16강에 진출한다. 16강전부터는 패하면 그대로 탈락하는 토너먼트제로 우승팀을 가린다. 단, 4강전(즉 준결승전)에서 패한 두 팀은 3, 4위전을 치른다.

- (1) 월드컵 축구대회의 경기 수는 모두 몇 경기인가?
- (2) 팀별 치르는 경기 수에 대한 팀 수의 분포를 구하라.

경기수	( )	( )	...
팀 수	( )	( )	...

● 이 문제는 조합(combination)의 개념과 토너먼트의 개념(한 경기당 한 팀 탈락)에 바탕을 둔 경기 수 공식을 이용하여 설명하면 학생들을 쉽게 이해시킬 수 있는 문제이다.

**2-2.** (2-1의 월드컵 경기 방식 참조) 월드컵 축구대회에서 아래의 각 내용을 고려할 때, 가능한 경우의 수를 구하라. 단, 16강전 이후의 대진은 다음과 같이 이루어진다.

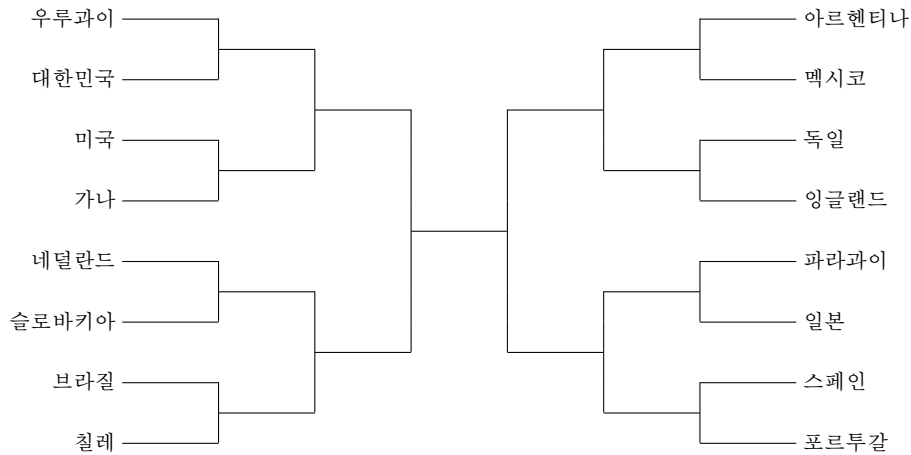


- (1) 16강 진출 팀
- (2) 8강 진출 팀
- (3) 4강 진출 팀

- (4) 결승 진출 팀
- (5) 우승 팀
- (6) 1, 2위 팀(순위를 따져서)
- (7) 1, 2, 3, 4위 팀(순위를 따져서)

● 이 문제는 순열(permutation)과 조합, 그리고 곱의 법칙(law of multiplication)을 이용하고 대진표의 특성을 고려하여 풀 수 있는 문제이다. (7)번 문제에서 학생들이 일반적인 순열의 공식을 사용하는 것으로 생각할 가능성이 있고 실제 2010년 1학기에 수업을 해 보니 그런 학생들이 상당히 나왔는데, 이것은 잘못된 것이고 대진표의 특성을 고려하여 풀어야 한다는 점을 잘 설명할 필요가 있다.

**2-3.** 2010년에 남아프리카공화국에서 열린 월드컵 축구대회 16강 대진표가 다음과 같다.



- (1) 결승전은 어느 나라와 어느 나라의 대결이 될 것인지를 생각할 때 모든 가능한 경우는 몇 가지인가?
- (2) 준결승전에서 패한 두 팀은 3, 4위전을 치르고, 준결승전에서 승리한 두 팀은 결승전을 치른다. 즉 1, 2, 3, 4위를 반드시 가린다. 1, 2, 3, 4위가 어느 팀이 될 것인지를 생각할 때, 가능한 경우는 모두 몇 가지인가?
- (3) 이 16개국 중 우루과이, 브라질, 칠레, 아르헨티나, 파라과이는 남미에 있고, 대한민국과 일본은 아시아에, 미국과 멕시코는 북중미에 있으며, 가나는 아프리카에 있고, 나머지 6개국은 유럽에 있다. 결승전 대진을 나라가 속한 대륙으로만 구별한다면(즉 어느 대륙과 어느 대륙이 맞붙는지) 가능한 경우는 몇 가지인가?
- (4) 어느 대륙의 팀이 1, 2위가 될 것인지를 생각할 때(즉 팀들을 대륙으로만 구별), 가능한 경우는 몇 가지인가?

● 이 문제에서는 곱의 법칙을 이용하되, 중복되어 카운트된 경우는 빼야 한다는 점을 주지시켜 설명할 필요가 있다.

**2-4.** 1986 멕시코, 1990 이탈리아, 1994 미국 월드컵의 경기 방식은 다음과 같다. 24개 팀이 참가하여 4개 팀씩 6개조로 나뉜다. 각 조에서 4개 팀이 풀리그를 벌여 각 조의 1, 2위 팀이 16강에 진출하고, 6개

조의 3위 팀들 중 성적이 좋은 4개 팀이 와일드카드로 16강에 합류한다. 어느 팀들이 16강에 진출할 것인가를 고려할 때, 가능한 경우는 모두 몇 가지인가?

● 현재의 월드컵 경기 방식은 1998년의 프랑스 월드컵부터 사용되고 있는 방식이라는 점에서, 1986년, 1990년, 1994년 대회에 관한 이 예제는 주목할 만하다. 특히 이들 대회에서는 두 팀이 16강에 진출하는 조도 있고 세 팀이 16강에 진출하는 조도 있을 뿐만 아니라 그것이 어느 조가 될지도 모르기 때문에 교역용으로 상당한 가치가 있는 문제이다. 조합과 곱의 법칙이 유용한 도구로 사용된다.

**2-5.** 월드컵 축구에서 같은 조에 편성된 네 팀이 경기를 하여 조 1, 2위 팀이 16강에 진출한다. 순위를 결정하는 최우선 기준은 승점(승 3점, 무 1점, 패 0점)이며, 승점이 같은 경우 골 득실차, 다득점, 승자승, 추첨 순으로 우선순위가 주어져 순위를 결정한다. 즉 어떠한 경우에도 조의 1, 2, 3, 4위 팀이 정해진다.

- (1) 조의 1, 2, 3, 4위 팀의 승점분포로 가능한 것은 모두 몇 가지인가?(예를 들면, 1위 팀 3승, 2위 팀 2승1패, 3위 팀 1승2패, 4위 팀 3패면 승점분포가 9, 6, 3, 0이며, 이것이 한 가지 경우이다.)
- (2) 어느 팀이 몇 위가 되느냐 하는 것까지 고려한다면, 조의 1, 2, 3, 4위 팀의 승점분포로 가능한 것은 모두 몇 가지인가?(예를 들면, 승점분포가 똑같이 9, 6, 3, 0이더라도, 가 팀이 1위이고 나 팀이 2위인 경우와, 나 팀이 1위이고 가 팀이 2위인 경우를 구별한다.)
- (3) (2)에서 승점이 같은 팀들은 구별하지 않는다면, 가능한 경우는 모두 몇 가지인가?(예를 들면, 승점분포가 9, 3, 3, 3인 경우는 4가지로 센다.)

● 이 문제는 학생들뿐 아니라 축구에 관심이 있는 일반인과 축구 전문가들도 관심을 가질 만한, 그러나 너무 복잡해서 답을 구할 엄두를 내지 못하는 문제가 아닐까 생각한다. 우선 (1)번 문제는 특별한 공식이 없고, 적절한 기준(예를 들면, 1위 팀의 승점)에 따라 경우를 나눈 뒤 끈기 있게 생각하여 풀 문제라고 생각한다. (2)번 문제에서는 (1)번의 결과와 순열을 고려해야 할 것이며, (3)번 문제에서는 (1)번의 결과와, 사물을 여러 그룹으로 나누는 경우의 수를 이용하면 될 것이다.

**2-6.** 2010년의 남아공 월드컵 축구대회에 참가할 32개 팀의 조를 편성하려고 한다. 각 조에 4개 팀씩 8개조(A조~H조)로 편성한다. 우선 32개 팀은 4개의 포트에 배치되어 있는데, 각 포트에는 8개씩의 팀이 포함되어 있다. 포트 1은 시드 배정을 받은 8개 팀(남아공, 이탈리아, 독일, 잉글랜드, 스페인, 네덜란드, 브라질, 아르헨티나)이며, 남아공은 개최국이므로 A조에 편성된다. 각 조에는 4개의 포트 각각에서 한 팀씩이 편성된다.

- (1) 조 편성 결과 나올 수 있는 모든 경우의 수를 구하라.
- (2) 서로 다른 포트에 있는 특정한 두 팀(두 팀 모두 시드 비배정팀. 예: 한국과 나이지리아)이 같은 조에 속하는 경우의 수를 구하라.
- (3) 서로 다른 포트에 있는 특정한 두 팀(두 팀 모두 시드 비배정팀)이 같은 조에 속할 확률을 구하라.
- (4) 특정한 두 팀(한 팀은 시드 배정팀, 한 팀은 비배정팀. 예: 브라질과 북한)이 같은 조에 속하는 경우의 수를 구하라.
- (5) 특정한 두 팀(한 팀은 시드 배정팀, 한 팀은 비배정팀)이 같은 조에 속할 확률을 구하라.

● 이것은 시드 배정과 포트 사용으로 인해 제한된 랜덤화가 사용되었다는 것을 설명하게 되는 문제이다. 순열과 곱의 법칙을 이용하여 풀 수 있을 것이다. 그리고 경우의 수를 바탕으로 구한 확률이 직관적 의미의 확률과 일치한다는 것을 학생들이 확인하게 하는 의미도 있는 문제이다.

**2-7.** 2002년에 개최된 한·일 월드컵의 조 편성 문제를 생각해 보자. 문제의 유형을 다양화하기 위하여 4개의 포트를 사용하지 않는다고 가정하자. 단, 서로 다른 조에 속해야 하는 8개의 시드 배정 팀은 정해져 있다. 시드 배정 팀 중 전 대회 우승팀인 프랑스는 A조, 한국은 D조, 일본은 H조에 속해야 하며, 브라질이 A, B, C, D조 중 하나에 속하면 아르헨티나는 E, F, G, H조 중 하나에 속해야 하고, 브라질이 E, F, G, H조 중 하나에 속하면 아르헨티나는 A, B, C, D조 중 하나에 속해야 한다.

- (1) 조 편성 결과 나올 수 있는 모든 경우의 수를 구하라.
- (2) 특정한 두 팀(두 팀 모두 시드 비배정팀. 예: 잉글랜드와 나이지리아)이 같은 조에 속하는 경우의 수를 구하라.
- (3) 특정한 두 팀(두 팀 모두 시드 비배정팀)이 같은 조에 속할 확률을 구하라.
- (4) 특정한 두 팀(한 팀은 시드 배정팀, 한 팀은 비배정팀. 예: 아르헨티나와 잉글랜드)이 같은 조에 속하는 경우의 수를 구하라.
- (5) 특정한 두 팀(한 팀은 시드 배정팀, 한 팀은 비배정팀)이 같은 조에 속할 확률을 구하라.
- (6) 한국, 폴란드, 미국, 포르투갈이 같은 조에 편성되는 경우의 수를 구하라.
- (7) 한국, 폴란드, 미국, 포르투갈이 같은 조에 편성될 확률을 구하라.

● 이 문제는 순열과 곱의 법칙, 그리고 사물을 여러 그룹으로 나누는 경우의 수를 복합적으로 이용하면 될 것이다. 문제 2-6과 2-7에서 나오는 천문학적인 경우의 수를 통하여, 학생들은 월드컵 조 추첨에서 얼마나 많은 경우가 나올 수 있는지 실감하게 될 것이다.

**2-8.** 두 축구팀 A와 B가 축구 경기를 했는데 연장전까지 해도 승부가 가려지지 않아 승부차기를 하고 있다. 승부차기는 양 팀이 페널티킥을 5개씩 차서 골을 많이 넣은 팀이 승리하는 제도이다. 5개씩 차는 때 동점이면 승부가 가려질 때까지 양 팀에서 계속해서 번갈아 1개씩 더 차게 된다. 각 선수가 승부차기에서 골을 넣을 확률은 독립적으로  $p$ 라 하자. A팀은 5개를 차서 4개를 성공시켰고 B팀은 4개를 차서 3개를 성공시킨 상태에서 5번째 키커가 찰 준비를 하고 있다.

- (1) 현 시점에서 A, B 각 팀이 승리할 확률을 구하라.
- (2) 현 시점부터 승부가 가려질 때까지 필요한 키커의 수(양 팀 합하여)를  $X$ 라 할 때  $E(X)$ 를 구하라.

● (1)번 문제는 조건부확률(conditional probability)과 전확률법칙(law of total probability)을 이용해서 풀도록 설명할 수 있으며, (2)번 문제는 승부차기가 조건부확률분포와 조건부기댓값(conditional expectation)의 설명에 사용될 수 있음을 보여 주는 예제이다.

### 3. 통계모형

**3-1.** 국제축구연맹(FIFA)에서 운영하는 인터넷 사이트인 [www.fifa.com](http://www.fifa.com)에서는 2개월 주기로 회원국의 축구실력을 점수화하여 랭킹을 매기고 있다. 이 점수가 각국의 축구실력을 절대적으로 대표한다고 볼 수는 없지만 공신력을 가진 기관에서 나온 점수인 만큼 세계의 많은 미디어들이 인용하는 정보인 것만큼은 틀림없다.

- (1) 축구실력이라는 변수를 만들어 내는 과정을 생각하여 보라. 월드컵에 진출한 각국의 축구실력이 객관적으로 공정하게 평가되는 과정이 축구에 관련한 통계적인 변수들에 의해 결정된다는 사실을 다른 학생들과 함께 토론해 보라.

- (2) 같은 실력을 가진 두 축구팀이 경기를 한다 하더라도 어떤 때는 이기고 어떤 때는 지기도 한다. 이런 의미에서 축구팀의 실력은 확률변수이다. 그럼 이런 확률적인 현상은 어느 확률변수를 선택하는 것이 좋은가?
- (3) A 팀의 축구실력을  $X$ 라는 정규분포를 따르는 확률변수, B 팀의 축구실력을  $Y$ 라는 정규분포를 따르는 확률변수로 나타내자. 그리고 두 팀의 평균 축구실력은 동일하다. 즉, 평균과 표준편차가 같은 정규분포라고 가정한다. 그렇다면 두 팀 간의 경기결과는 새로운 확률변수로 탄생이 된다. 즉,  $X - Y$ 란 새로운 확률변수가 된다. 이 확률변수는 어떤 확률변수인가? 구체적으로 답하라. 두 확률변수는 독립임을 가정하라. 만약 독립이 아님을 가정하게 되는 경우 이는 실제 어떤 의미인가?
- (4) (3)에서 정의한 확률변수의 값이 양의 아주 큰 값이 나오면 A팀이 이기는 결과가 되는 것이고 음의 아주 큰 값이 나오면 B팀이 이기는 결과가 나온다.  $|X - Y|$ 의 값이 그렇게 크지 않다면 두 팀 간의 경기는 무승부로 끝날 것이다. 무승부의 확률은 영국이나 스페인의 일급 리그인 경우 25%, 그리고 한국 K-리그 축구인 경우는 35%이다. 무승부의 확률을 이렇게 설정할 때,  $|X - Y|$ 의 값이 표준편차의 몇 배 이내이면 두 팀의 경기가 무승부로 끝나는지 구하라. 즉  $P(|X - Y| < k\sigma) = 0.25$ 에서  $k$ 값을 구하라.
- (5) 그렇다면 각 팀의  $\sigma$ 를 측정하는 문제가 발생된다. 즉, 평균을 [www.fifa.com](http://www.fifa.com)에서 제공하는 점수로 가정하는 것은 문제없으나 표준편차의 크기를 정하는 문제는 생각하여 볼 필요가 있다. 월드컵의 예선 각 조는 4개 팀이 배정된다. 실력 및 대륙별 안배를 통해 배정된다. 배정된 팀 중 제일 랭킹이 높은 팀이 제일 낮은 팀에 질 가능성은 희박하긴 하지만 항상 존재한다. 이런 경과를 이번이라 칭한다. 배정된 4개 팀 중 최상위 팀이 최하위 팀의 평균 점수까지 떨어질 수 있다고 가정하자. 그렇다면  $\sigma$ 에 대한 추정값을 짐작할 수 있겠는가? 기초통계학에서 배운 정규분포의 성질에 의거한 경험의 규칙(rule of thumb)에 따르면 평균을 중심으로  $\pm 2\sigma$  혹은  $\pm 3\sigma$ 에는 전체 자료의 95% 내지 99%가 포함된다는 사실을 기억하라.
- (6)  $\sigma$ 의 추정이 마무리 되었다면 각 조에 속한 4개 팀이 서로 경기를 했을 경우 상대간의 전적을 계산할 수 있다. 2010년 대회의 B조에는 아르헨티나, 그리스, 나이지리아, 한국 등 4개 팀이 속해 있다. 나이지리아가 한국과 경기를 했을 때 지거나 비길 확률이 얼마나 되는지, 혹은 아르헨티나가 그리스와 경기를 했을 때 이길 확률과 같은 값을 구할 수 있다. 물론 이 경우는 [www.fifa.com](http://www.fifa.com)에서 발표한 점수를 각 팀의 평균 실력으로 가정한다.

**3-2.** 3-1에서 언급한 통계적인 모형과 가정을 이해했다면 이제는 이를 종합적으로 계산하여 월드컵 모형을 만들 필요가 있다. 즉 각국의 16강, 8강, 4강, 결승 및 우승 확률을 월드컵 규칙에 의거하여 모형을 수립하여 계산하는 종합적인 절차이다. 그러나 팀 간 성적에 대한 통계적인 계산이 마무리가 되었다 하더라도 이를 월드컵 규칙에 의해 종합화하는 작업은 매우 복잡한 단계를 거쳐야 하기 때문에 이 시점에서 필요한 지식이 확률변수 값을 구현한 다음 모의실험을 하는 방법이다. 다행히 엑셀과 같은 프로그램에는 난수 생성기를 위시하여 정규분포를 따르는 확률변수 값을 생성하는 루틴이 이미 내장되어 있는 상태이다.

- (1) 엑셀에는 `=rand()`와 같은 난수 생성기와 `=norminv(mu, sigma)`와 같은 정규분포를 따르는 확률변수 값을 생성하는 루틴이 존재한다. 2010년 월드컵 B조에 속한 4개국 중 아르헨티나와 나이지리아가 경기를 했을 때 아르헨티나가 이길 확률, 나이지리아가 이길 확률, 그리고 비길 확률을 각각 모의실험을 통해 구하라.  $p(|X - Y| > k\sigma)$ 가 발생이 되면 승부가 결정되고 그렇지 않으면 무승부가 난다는 가정을 한다. 상용 프로그램인 @RISK나 Crystall Ball과 같은 상용프로그램을 사용해도 좋다. 단 모의실험 횟수는 사용자의 입장에서 정한다.

- (2) 결과적으로 아르헨티나와 우리나라가 2010년 월드컵에서는 16강에 진출하였다. 이러한 결과가 나올 확률이 얼마나 되었는지 역으로 구하라. 마찬가지로 방법으로 스페인과 네덜란드가 결승에서 경기를 할 확률이 얼마나 되었는지 구하라.
- (3) 예선 첫 경기에서 우리나라는 그리스를, 아르헨티나는 나이지리아를 이겼다. 이런 정보가 주어졌을 때 우리나라와 아르헨티나가 동시 16강에 진출할 확률은 얼마나 되었는지 구하라. 조건부 확률이다.
- (4) 개별적인 모의실험에 의해 얻은 결과를 종합할 차례이다. 월드컵은 8개조로 나누어 같은 방식으로 경기를 치른다. 위에서 논한 논리로 모의실험을 시행한 다음 이를 종합하여 각 팀의 16강, 8강, 4강, 그리고 결승, 우승 확률을 계산하라. 그리고 이를 표로 작성하고 도박사의 확률과 비교하여 기말보고서로 작성한다.
- (5) 지난 2002년, 2006년, 그리고 2010년 3개 년도를 비교하여 [www.ffa.com](http://www.ffa.com) 에 의해 배정된 팀들이 정상적으로 16강에 진출하였는지 점그림(dot plot)과 같은 도시적인 방법을 통해 분석하라.

#### 4. 결론

본 연구에서는 학교 수업 시간에 월드컵 축구의 예제를 통하여 확률·통계나 경우의 수를 설명할 수 있는 사례를 살펴보았다. 또한 통계모형을 이용한 시뮬레이션을 통해 2010 남아공 월드컵을 앞둔 시점이거나 가정하고 본선 진출 팀 각각의 16강 진출 확률과 우승 확률을 계산하는 논리적 방법을 제안하였다. 2010년 1학기 교양수업에서 인문사회계열 학생들에게 본 논문에서 소개한 예제들을 사용하여 수업을 하였는데, 복잡한 문제는 이해하는 데 어려움을 느끼지만, 평소에 흥미를 많이 가지고 있던 월드컵 축구에 관한 예제를 다룬다는 점에서 타 분야의 예제들에 비해 관심을 보이는 학생들이 많았고, 수업이 끝날 때에는 아쉬움을 표시하는 학생들도 있었다. 요즘 학생들 중에는 월드컵을 포함하여 스포츠에 관심이 있는 학생들이 많으므로, 스포츠 분야의 예제를 종종 사용하여 설명을 하고 가끔 스포츠에 얽힌 일화를 들려주면 수업에 좀 더 높은 관심과 흥미를 끌 수 있지 않을까 생각한다. 효과적인 교육뿐 아니라 스포츠 분야와 통계의 접목을 통해 스포츠 분야로의 진출도 학생들은 생각하여 볼 수 있을 것이다. 경륜과 경마 등 확률이 직접적으로 구현되는 현장에서의 모형개발, 각 스포츠 프로팀들의 전력분석 및 전력 강화 모형, 개개인의 데이터 분석을 통한 경기전략 수립, 경기협회 차원에서의 새로운 경기 규칙 개발 등 많은 응용분야가 있음을 알리는 것도 하나의 부수적인 효과라 볼 수 있다. 물론 이는 스포츠의 과학화 및 대중화를 책임지고 있는 대한체육회의 총괄적인 업무라고 본다. 본 연구는 통계학을 통한 스포츠의 대중화를 위한 한 걸음이라 생각한다.

#### 참고문헌

- 강명욱, 김영일 (2002). Sports statistical model and simulation, <한국분류연구학회지>, **5**, 11-24.  
 국제축구연맹 공식 홈페이지. [www.ffa.com](http://www.ffa.com).  
 김도현, 이상인, 김응대 (2007). BRADLEY-TERRY 모형을 이용한 2006 독일 월드컵 예측, <응용통계연구>, **20**, 205-218.  
 조대현 (2008). 라운드로빙 방식을 응용한 복식조 편성방법, <응용통계연구>, **21**, 1015-1026.  
 조대현 (2009). 실력이 순서화된 경우에 대한 복식조 편성방법, <응용통계연구>, **22**, 1331-1343.  
 홍종선, 정민섭, 이재형 (2010). 2010 남아공 월드컵 축구 예측모형 연구, <한국통계학회 2010년 춘계 학술논문발표회 프로시딩>, 74.  
 CrystalBall 홈페이지. [www.crystalball.com](http://www.crystalball.com).  
 @RISK 홈페이지. [www.palisade.com](http://www.palisade.com).  
 Winston, W. L. (2009) *Mathletics*, Princeton University Press.

# Teaching Statistics through World Cup Soccer Examples

Hyuk Joo Kim<sup>1</sup> · Youngil Kim<sup>2</sup>

<sup>1</sup>Division of Mathematics & Informational Statistics and Institute of Basic Natural Sciences,  
Wonkwang University

<sup>2</sup>Department of Information System, ChungAng University

(Received July 2010; accepted September 2010)

---

## Abstract

In teaching probability and statistics classes, we should increase efforts to develop examples that enhance teaching methodology in delivering more meaningful knowledge to students. Sports is one field that provides a variety of examples and World Cup Soccer events are a treasure house of many interesting problems. Teaching, using examples from this field, is an effective way to enhance the interest of students in probability and statistics because World Cup Soccer is a matter of national interest. In this paper, we have suggested several examples pertaining to counting the number of cases and computing probabilities. These examples are related to many issues such as possible scenarios in the preliminary round, victory points necessary for each participant to advance to the second round, and the issue of grouping teams. Based on a simulation using a statistical model, we have proposed a logical method for computing the probabilities of proceeding to the second round and winning the championship for each participant in the 2010 South Africa World Cup.

**Keywords:** Number of cases, probability, distribution of victory points, conditional expectation, statistical model, simulation.

---

---

This paper was supported by Wonkwang University in 2008.

<sup>2</sup>Corresponding author: Professor, Department of Information System, School of Industrial Study, ChungAng University, Ansong, Kyunggi 456-756, Korea. E-mail: yik01@cau.ac.kr