풀이

STEP I 1988, 문제 15

확률과 통계 • 이항 가설검정

> **STEP I 1988, 15번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/88-s1-q15

핵심 아이디어

"점수의 절반 넘게를 파울로 얻는다"는 조건을 파울 수와 골 수 사이의 간단한 부등식(inequality)으로 바꾼다. 그다음 골 수에 대해 조건부(conditioning)로 생각하면, 같은 확률로 일어나는 파울 값의 개수를 세는 문제만 남는다. 둘째 부분에서는 **예전** 분포를 따른다면 300번의 시도에서 파울 아홉 번인 경우가 몇 번쯤 나올 것으로 기대되는지, 그리고 아무것도 바뀌지 않았다면 관찰된 횟수가 얼마나 놀라운지를 묻는다.

풀이

확률

어떤 경기에서 특정한 한 팀에 대하여, 그 팀이 넣은 골 수를 G, 그 팀을 상대로 범한 파울 수를 F라 하자. 그러면 G와 F는 독립(independent)이고,

$$\mathrm{P}(G=0),\ \ldots,\ \mathrm{P}(G=4) \;=\; \tfrac{1}{9},\ \tfrac{2}{9},\ \tfrac{3}{9},\ \tfrac{2}{9},\ \tfrac{1}{9}, \qquad \mathrm{P}(F=f)=\tfrac{1}{10} \quad (f=0,1,\ldots,9).$$

이다. 팀은 골로 $2G$점, 파울로 F점을 얻으므로 총점은 $2G+F$이다. 점수의 절반 넘게를 파울로 얻는 것은 정확히

$$F > \tfrac{1}{2}(2G+F) \iff F > 2G.$$

일 때이다. (팀이 점수를 전혀 얻지 못하면 $F=0=2G$이므로 이 경우는 세지 않는데, 엄격한 부등식이 이를 올바르게 처리한다.)

$G=g$일 때 유리한 F의 값은 $2g+1,\ldots,9$이고, 이는 같은 확률을 갖는 열 개의 값 가운데 $9-2g$개이다. 따라서 F와 G의 독립성을 이용하면

$$\mathrm{P}(F>2G \mid G=g)=\frac{9-2g}{10}, \qquad g=0,1,2,3,4,$$

이고, 그 값은 차례로 $\frac{9}{10}, \frac{7}{10}, \frac{5}{10}, \frac{3}{10}, \frac{1}{10}$이다. 전확률 법칙(law of total probability)에 의하여

$$\mathrm{P}(F > 2G) = \frac{1\cdot 9 + 2\cdot 7 + 3\cdot 5 + 2\cdot 3 + 1\cdot 1}{9\times 10} = \frac{45}{90} = \frac{1}{2}.$$

답이 깔끔하게 나오는 데에는 이유가 있다. 골의 분포는 2에 대하여 대칭이고 $(9-2g)/10$은 g에 대한 일차식이므로, 그 평균은 $g=2$에서의 값, 곧 $\frac{1}{2}$과 같다.

파울 분포가 바뀌었는가?

한 팀이 범하는 파울 수는 그 상대 팀을 상대로 범한 파울 수이므로, 예전 규칙에서는 각 팀이 한 경기에서 파울을 9번 범할 확률이 $\frac{1}{10}$이다. 300개의 팀-경기 가운데 팀이 파울을 9번 범한 경우의 수를 X라 하자. 우리는

$$H_0:\ p = \tfrac{1}{10}\quad (\text{변화 없음}) \qquad \text{대} \qquad H_1:\ p < \tfrac{1}{10},$$

를 검정한다. 여기서 p는 한 팀이 한 경기에서 파울을 9번 범할 확률이다. 규칙 변경은 파울이 많이 나오는 경우를 덜 일어나게 하려는 것이었으므로, 대립가설(alternative hypothesis)은 단측(one-sided)이다.

? 생각해 보기 여기서 이항(binomial) 모형을 쓰는 것이 정당한가?

300개의 팀-경기가 서로 독립이고, 각각 같은 확률 p를 가져야 한다. 문제에서 한 경기 안에서의 독립성은 주어져 있고, 서로 다른 경기들도 독립이라고 가정하는 것이 합리적이다. 따라서 H_0 아래에서 $X \sim \mathrm{B}(300, \frac{1}{10})$이고,

$$\mathrm{E}(X) = 30, \qquad \mathrm{Var}(X) = 300\times\tfrac{1}{10}\times\tfrac{9}{10} = 27, \qquad \text{표준편차} \sqrt{27} \approx 5.20.$$

이다. 관찰값은 $X=12$로, 기댓값 30보다 18만큼 작으며 이는 약 3.5 표준편차(standard deviation)에 해당한다. 정규근사(normal approximation)와 연속성 수정(continuity correction)을 쓰면

$$\mathrm{P}(X \le 12) \approx \Phi\left(\frac{12.5-30}{\sqrt{27}}\right) = \Phi(-3.37) \approx 0.0004.$$

이다. 정확한 이항 확률은 이보다 더 작아서 약 0.0001이다. 어느 쪽이든, 아무것도 바뀌지 않았다면 이만큼 극단적인 결과는 천 시즌에 한 번도 채 일어나지 않는다. 이는 흔히 쓰는 5% 또는 1% 유의수준(significance level)조차 훨씬 넘어서므로, H_0을 기각한다. 이는 파울 수의 분포가 바뀌었다는 좋은 증거가 **맞으며**, 바뀐 방향도 기대했던 대로이다. 파울 아홉 번인 경기가 더 드물어진 것이다.

두 가지 주의할 점을 짚어 둘 만하다. 이 검정은 P(9번 파울)이 줄어들었다는 것만 보여 줄 뿐, 분포의 나머지 부분이 어떻게 바뀌었는지는 보여 주지 않는다(그렇게 보는 것이 자연스러운 설명이기는 하지만). 또한 시즌 사이에 다른 요인들도 바뀌었을 수 있으므로, 통계만으로는 규칙 변경이 감소를 **일으켰다**고 증명할 수 없다.

검산

- 골 확률 $\frac{1}{9}(1+2+3+2+1)$의 합은 당연히 1이다.
- 5×10개의 순서쌍 (g, f)를 모두 정확한 분수로 나열하면 $\mathrm{P}(F > 2G) = \frac{1}{2}$을 얻고, 200,000경기를 시뮬레이션하면 0.4997이 나왔다.
- 이항 확률을 직접 더하면 $X \sim \mathrm{B}(300, 0.1)$에 대하여 $\mathrm{P}(X \leq 12) = 9.2 \times 10^{-5}$이다. 정규 근사($3.8 \times 10^{-4}$)와 평균이 30인 푸아송 근사(Poisson approximation)(1.7×10^{-4}) 모두 이 값이 매우 작다는 데 일치한다.

흔한 실수

- F를 $2G$가 아니라 G와 비교하는 것: 한 골은 2점이다.
- $F > 2G$ 대신 $F \geq 2G$를 세는 것. "절반 넘게"는 엄격한 조건이고, 점수가 없는 팀은 파울로 얻은 점수도 없다.
- 표본 크기로 300개의 팀-경기 대신 150경기를 쓰는 것, 또는 계산하기 전에 귀무가설(null hypothesis)과 이항 모형을 밝히는 것을 잊는 것.

STEP 너머

둘째 부분은 이항 비율에 대한 가설검정(hypothesis test)이고, 우리가 구한 꼬리 확률은 p-값(p-value)이다. 대학에서는 우도비 검정(likelihood ratio test)과 카이제곱 검정(chi-squared test)을 배우게 되는데, 이들은 한 값만이 아니라 관찰된 파울 수의 분포 전체를 예전 분포와 비교한다. Blitzstein and Hwang, *Introduction to Probability*(확률론 입문)는 확률 쪽 내용(조건부 확률, 이항분포와 정규근사)을 매우 명료하게 다룬다.

Beyond STEP 풀이: https://beyondstep.keithmcnulty.org/ko/. 풀이 글은 CC0 1.0이다. 문제 자체의 저작권은 © Cambridge University Press & Assessment에 있으며 여기에 싣지 않는다. 영어 원문을 기계 번역하였으며 아직 검토되지 않았다.