풀이

STEP I 2019, 문제 11

확률과 통계 • 기하분포

STEP I 2019, 11번 문제

더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 OCR에서 내려받을 수 있는 2019년 STEP 시험지 묶음(zip 파일) 가운데 STEP I 시험지에 있다:

www.ocr.org.uk/Images/696590-2019-past-papers.zip

이 문제는 *Beyond STEP*의 12.9 절 **치우친 동전으로 공정하게 정하기**에서 단계별 안내와 함께 자세히 다룬다: https://beyondstep.keithmcnulty.org/ko/ch12/von-neumann.html

핵심 아이디어

두 사람 절차의 각 라운드는 결정이 나거나 나지 않는 독립(independent) 시행이므로, 결정이 나는 라운드는 기하분포(geometric distribution)를 따른다. 그러면 "p의 값에 관계없이" 성립하는 부등식은 $p+q=1$일 때 pq가 얼마나 커질 수 있는가 하는 문제로 귀결되고, 이는 산술-기하 평균 부등식(AM–GM inequality)으로 해결된다. 소문항 (ii)에서는 둘째 라운드가 더 이상 처음부터 다시 시작하는 것이 아니므로, 첫째 라운드에서 일어난 일에 따라 경우를 나누고, 같은 pq의 상한으로 마무리한다.

풀이

소문항 (i)

어느 한 라운드에서 두 동전은 독립이다. 둘 다 앞면이 나오거나(확률 p^2) 둘 다 뒷면이 나오면(확률 q^2) 결정이 이루어지므로

$$\mathrm{P}(\text{주어진 한 라운드에서 결정됨}) = p^2+q^2, \qquad \mathrm{P}(\text{결정 안 됨}) = pq+qp = 2pq.$$

검산하면 $p^2+q^2+2pq=(p+q)^2=1$이다.

매번 새로 동전을 던지므로 각 라운드는 독립이다. n번째 라운드에서 결정이 난다는 것은 결정이 나지 않는 라운드가 $n-1$번 이어진 뒤 결정이 나는 라운드가 한 번 온다는 뜻이므로

$$\mathrm{P}(\text{결정 라운드가} n) = (2qp)^{n-1}(q^2+p^2),$$

이고, 이것이 구하는 결과이다.

둘째 결과는 여사건(complement)을 이용하는 것이 가장 쉽다. n번째 라운드까지 결정이 나지 않는다는 것은 결정이 나지 않는 라운드가 연속으로 n번 나온다는 뜻이고, 그 확률은 $(2pq)^n$이다. 따라서

$$\mathrm{P}(\text{결정이 늦어도 라운드}n) = 1 - (2pq)^n.$$

($p^2 + q^2 = 1 - 2pq$를 이용하여 등비급수 $\sum_{r=1}^{n}(p^2+q^2)(2pq)^{r-1}$의 합을 구해도 같은 결과를 얻는다.)

이제 모든 p에 대하여 성립하는 pq의 상한이 필요하다. $p, q \geq 0$이므로 산술-기하 평균 부등식에 의해

$$\sqrt{pq} \leq \frac{p+q}{2} = \frac{1}{2}, \qquad \text{따라서} \qquad pq \leq \frac{1}{4},$$

이고, 등호는 정확히 $p = q = \frac{1}{2}$일 때 성립한다. (같은 말로, $pq = p(1-p) = \frac{1}{4} - (p - \frac{1}{2})^2$이다.) 그러므로 $0 \leq 2pq \leq \frac{1}{2}$이고, 따라서 $(2pq)^n \leq \frac{1}{2^n}$이며

$$\mathrm{P}(\text{결정이 늦어도 라운드}n) = 1 - (2pq)^n \geq 1 - \frac{1}{2^n}$$

이 모든 p의 값에 대하여 성립하여, 증명이 끝난다. 공정한 동전이 최악의 경우이다.

소문항 (ii)

D_1을 라운드 1에서 결정이 나는 사건, D_2를 라운드 2에서 결정이 나는 사건이라 하자. 이 두 사건은 서로 배반(mutually exclusive)이므로 $\mathrm{P}(D_1) + \mathrm{P}(D_2)$를 구하면 된다.

라운드 1. 셋 모두 앞면이거나 셋 모두 뒷면인 경우이므로 $\mathrm{P}(D_1) = p^3 + q^3$이다.

라운드 2. 라운드 1에서 결정이 나지 않았다면, 동전은 앞면 둘과 뒷면 하나이거나 뒷면 둘과 앞면 하나였다.

- **앞면 둘, 뒷면 하나** (세 사람 중 누구든 뒷면이 나온 사람이 될 수 있으므로 확률은 $3p^2q$). 앞면이 나온 두 사람은 다시 던지고, 셋째 사람은 뒷면을 뒤집어 앞면이 보이게 한다. 이제 모두 뒷면인 경우는 불가능하고, 모두 앞면인 경우는 새로 던진 두 동전이 모두 앞면일 때, 즉 확률 p^2로 일어난다. 이 경로의 기여는 $3p^2q \cdot p^2 = 3p^4q$이다.
- **뒷면 둘, 앞면 하나** (확률 $3pq^2$). 같은 논리로 셋째 동전은 이제 뒷면을 보이고, 새로 던진 두 동전이 모두 뒷면이어야 하며 그 확률은 q^2이다. 이 경로의 기여는 $3pq^2 \cdot q^2 = 3pq^4$이다.

라운드 2에서 던지는 것은 라운드 1과 독립이므로 곱할 수 있다. 따라서

$$\mathrm{P}(D_2) = 3p^4q + 3pq^4 = 3pq\,(p^3 + q^3).$$

? 생각해 보기 $p^3+q^3+3pq(p^3+q^3)$를 깔끔하게 정리할 방법이 있을까?

p^3+q^3으로 묶어 낸 다음, $p+q=1$과 세제곱의 합(sum of cubes) 공식을 이용한다.

$$p^3+q^3=(p+q)(p^2-pq+q^2)=(p+q)^2-3pq=1-3pq.$$

따라서

$$\begin{aligned}\mathrm{P}(\text{결정이 늦어도 라운드 2까지}) &= (p^3+q^3)(1+3pq)\\ &= (1-3pq)(1+3pq)\\ &= 1-9p^2q^2.\end{aligned}$$

소문항 (i)에서 $0\le pq\le\frac{1}{4}$이므로 $p^2q^2\le\frac{1}{16}$이고

$$\mathrm{P}(\text{결정이 늦어도 라운드 2까지})\ge 1-\frac{9}{16}=\frac{7}{16}$$

이 p의 값에 관계없이 성립하며, 등호는 $p=\frac{1}{2}$일 때 성립한다. 이로써 증명이 끝난다.

검산

- $p=1$(양면이 모두 앞면인 동전)이면 즉시 결정이 난다. 소문항 (i)에서 $1-(2pq)^n=1$, 소문항 (ii)에서 $1-9p^2q^2=1$이 되어 마땅히 그래야 하는 결과와 일치한다.
- 소문항 (ii)에서 $p=\frac{1}{2}$이면, 라운드 1에서 확률 $\frac{1}{4}$로 결정이 난다. 그렇지 않은 경우($\frac{3}{4}$)에는 라운드 2에서 확률 $\frac{1}{4}$로 결정이 나므로 $\frac{1}{4}+\frac{3}{16}=\frac{7}{16}$이 되어, 하한이 실제로 달성된다.
- $p=0.3$이면 공식은 $1-9(0.21)^2=0.6031$을 준다. 세 사람 절차를 200 000번 컴퓨터로 시뮬레이션한 결과는 0.602였다.

흔한 실수

- 소문항 (ii)의 라운드 1에서 누가 홀로 다른 면이 나온 사람인지에 대한 인수 3을 빠뜨린다.
- 소문항 (ii)의 라운드 2를 동전 세 개를 새로 던지는 라운드처럼 취급한다. 뒤집어 놓은 동전은 고정되어 있으므로 "모두 앞면"과 "모두 뒷면" 중 하나만 가능하다.
- 부등식을 $p=\frac{1}{2}$에서만 증명한다. **모든** p에 대하여 $pq\le\frac{1}{4}$가 필요하며, 산술-기하 평균 부등식 같은 부등식이 이를 보장해 준다.

STEP 너머

소문항 (i)은 흔히 폰 노이만(von Neumann)의 것으로 알려진, 편향된 동전으로 공정한 결정을 얻는 고전적인 기법과 관련이 있다. 두 사람 절차에서 "호텔"과 "찻집"은 $p = \frac{1}{2}$일 때에만 같은 확률로 나오지만, 한 사람이 두 번 던져 HT와 TH를 결과로 사용하면 모든 p에 대하여 공정한 결과를 얻는다. 결정이 날 때까지의 라운드 수는 기하 확률변수(geometric random variable)이며, 이런 절차가 얼마나 빨리 멈추는가에 대한 질문은 정지 시간(stopping time)과 마르코프 연쇄(Markov chain)의 연구로 이어진다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문서)이 두 주제를 모두 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.