풀이

STEP III 2004, 문제 13

확률과 통계 • 기하분포

STEP III 2004, 13번 문제

더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:

step.maths.org/questions/04-s3-q13

핵심 아이디어

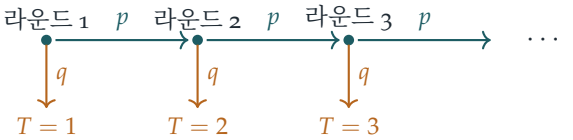
잠시 대회 전체는 잊고 참가자 한 명만 따라가 보자. 그 참가자가 탈락하는 라운드는 기하 확률변수(geometric random variable)이고, 세 참가자의 탈락 라운드는 서로 독립(independent)이다. 그러면 모든 소문항은 독립인 세 기하 확률변수에 관한 문제가 된다. 대상(grand prize)은 **가장 늦은** 탈락 라운드가 겹치는지 아닌지에만 달려 있음이 밝혀지므로, 마지막 소문항은 앞의 두 소문항을 합친 사건의 여사건(complement)이다.

풀이

$q = 1 - p$로 쓰고, $0 < p < 1$이라 가정한다. 참가자 j ($j = 1, 2, 3$)에 대하여 T_j를 그 참가자가 실패하는 라운드라 하자. 참가자가 라운드 k에서 실패하는 것은 라운드 $1, \ldots, k-1$을 통과한 뒤 라운드 k에서 실패하는 경우와 정확히 같으므로

$$\mathrm{P}(T_j = k) = p^{k-1}q, \qquad k = 1, 2, 3, \ldots$$

이다.



각 라운드의 통과 여부는 독립적으로 정해지고 참가자들도 서로 독립적으로 행동하므로, T_1, T_2, T_3는 독립이다. 또한 다음에 주목하자.

$$\mathrm{P}(T_j \leq r-1) = 1 - \mathrm{P}(\text{다음 라운드를 모두 통과: } 1, \ldots, r-1) = 1 - p^{r-1}.$$

소문항 (i)

독립성에 의하여 세 명이 모두 라운드 k에서 탈락할 확률은 $(p^{k-1}q)^3$이다. 이를 k에 대하여 더하면 공비가 $p^3 < 1$인 등비급수(geometric series)가 되므로

$$\mathrm{P}(T_1 = T_2 = T_3) = \sum_{k=1}^{\infty} q^3 p^{3(k-1)} = \frac{q^3}{1-p^3} = \frac{(1-p)^2}{1+p+p^2}.$$

소문항 (ii)

먼저 탈락하는 참가자를 고르는 방법은 3가지이다. 한 가지 선택에 대하여, 나머지 두 명이 모두 라운드 r에서 실패할 확률은 $(p^{r-1}q)^2$이고, 먼저 탈락하는 참가자가 r 이전의 어떤 라운드에서 실패할 확률은 $1-p^{r-1}$이다. 선택이 다르면 이 사건들은 서로 배반이므로

$$\mathrm{P}(\text{두 명은 라운드}r,\ \text{나머지는 더 일찍}) = 3p^{2r-2}q^2\left(1-p^{r-1}\right).$$

소문항 (iii)

? **생각해 보기** T_1, T_2, T_3로 나타낼 때, 대상은 정확히 언제 수여되는가?

T_j 중 최댓값(maximum)을 단 한 명의 참가자만 가진다고 하자. 예를 들어 $T_3 > T_1, T_2$라 하고 $m = \max(T_1, T_2)$라 하자. 라운드 m에서 참가자 3은 아직 남아 있고 그 라운드를 통과한다. 그 라운드에 남아 있는 다른 모든 참가자는 $T_j = m$이므로 그 라운드에서 실패한다. 따라서 참가자 3이 라운드 m에서 상을 받는다.

역으로, 최댓값을 여러 명이 공유한다고 하자. 예를 들어 $T_2 = T_3 = M \geq T_1$이라 하자. 라운드 M 이전에는 참가자 2와 3이 모두 남아 있고 매 라운드를 둘 다 통과하므로, 생존자가 단 한 명인 라운드는 없다. 라운드 M에서는 남은 모든 참가자가 실패하므로 생존자가 아예 없다. 따라서 상은 수여되지 않는다.

그러므로 상이 수여되지 **않는** 것은 가장 늦은 두 탈락 라운드가 같을 때와 정확히 일치한다. 즉 세 값이 모두 같거나(소문항 (i)), 두 값이 어떤 라운드 r에서 같고 세 번째가 그보다 이른 경우이다. 두 번째 경우는 $r \geq 2$이어야 하며, 이는 소문항 (ii)의 사건이다. 소문항 (ii)의 결과를 $r \geq 2$에 대하여 더하고 $s = r-1$로 쓰면

$$\begin{aligned}\sum_{r=2}^{\infty} 3q^2 p^{2r-2}\left(1-p^{r-1}\right) &= 3q^2\sum_{s=1}^{\infty}\left(p^{2s}-p^{3s}\right) = 3q^2\left(\frac{p^2}{1-p^2}-\frac{p^3}{1-p^3}\right)\\ &= 3q^2\,\frac{p^2(1-p^3)-p^3(1-p^2)}{(1-p^2)(1-p^3)} = 3q^2\,\frac{p^2(1-p)}{(1-p^2)(1-p^3)}\\ &= \frac{3p^2q^2}{(1+p)(1-p^3)}.\end{aligned}$$

여기에 소문항 (i)의 결과를 더하고 $1-p^3 = q(1+p+p^2)$를 이용하면

$$\mathrm{P}(\text{상 없음}) = \frac{q^3(1+p)+3p^2q^2}{(1+p)(1-p^3)} = \frac{q^2(1+2p^2)}{(1+p)(1-p^3)} = \frac{(1-p)(1+2p^2)}{(1+p)(1+p+p^2)}.$$

$(1+p)(1+p+p^2) = 1+2p+2p^2+p^3$이고 $(1-p)(1+2p^2) = 1-p+2p^2-2p^3$이므로, 둘의 차는 $3p+3p^3$이다. 따라서

$$\mathrm{P}(\text{대상 수여}) = \frac{3p(1+p^2)}{(1+p)(1+p+p^2)}.$$

검산

- $p \to 0$이면 모두가 라운드 1에서 실패하고, (iii)의 답은 0으로 수렴한다. $p \to 1$이면 동률이 일어나기 어려워지고, 답은 $\frac{3\cdot 2}{2\cdot 3} = 1$로 수렴한다.
- 직접 계산하면, 상은 어떤 T_j 하나가 나머지 둘보다 클 때 수여되므로 그 확률은 $3\sum_k p^{k-1}q\,(1-p^{k-1})^2$이다. $p = \frac{1}{2}$일 때 정확히 합을 구하면 $\frac{5}{7} \approx 0.714$이고, 이는 $\frac{3\cdot\frac{1}{2}\cdot\frac{5}{4}}{\frac{3}{2}\cdot\frac{7}{4}} = \frac{5}{7}$과 일치한다.
- 대회 자체를 시뮬레이션한 결과($p = 0.6$에서 2×10^5회 실행), 전체 실행의 77.9%에서 상이 수여되었으며, 공식의 값은 0.7806이다.

흔한 실수

- 소문항 (ii)에서 어느 참가자가 먼저 탈락하는지를 고르는 인수 3을 빠뜨린다.
- 상이 수여되려면 마지막 두 명이 서로 다른 라운드에서 탈락해야 한다고 생각한다. 두 명이 라운드 m에서 함께 탈락하고 세 번째 참가자가 그 라운드를 통과하면, 그 참가자는 여전히 상을 받는다. 상을 막는 것은 **맨 위**에서의 동률뿐이다.

STEP 너머

이 문제는 독립인 기하 확률변수들의 최댓값과 그 최댓값이 유일한지에 관한 것이다. 참가자가 n명일 때, 맨 위에서 동률이 일어날 확률은 $n \to \infty$일 때 극한값에 수렴하지 않는다. 그 확률은 $\log n$에 따라 아주 미세하게 진동하는데, 이는 리더 선출(leader election) 같은 알고리즘 분석에서 잘 알려진 놀라운 사실이다. 여기서 쓴 도구, 즉 한 변수에 대한 조건부(conditioning) 계산과 등비급수의 합은 Grimmett and Welsh, *Probability: An Introduction*(확률론 입문)에서 자세히 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.