풀이

STEP II 2009, 문제 13

확률과 통계 • 이항분포와 기대 비용

STEP II 2009, 13번 문제

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

step.maths.org/questions/09-s2-q13

핵심 아이디어

고장 나는 엔진의 개수는 이항분포(binomial distribution)를 따르므로, 수리 비용은 이항 확률변수의 간단한 함수이고 그 기댓값(expectation)은 짧은 합으로 구할 수 있다. 깔끔한 형태 $(*)$에 이르는 요령은 "모두 교체하는" 경우들을 하나씩 더하지 않는 것이다. 여사건(complement)을 이용한 다음, $1-q^n = p(1+q+\cdots+q^{n-1})$으로 인수분해하여 p를 인수로 끌어낸다.

풀이

문제의 의도대로 엔진들은 서로 독립적으로(independently) 각각 확률 p로 고장 난다고 가정하고, $q = 1-p$로 쓴다. 마지막 단계 전까지 모든 비용은 K 단위로 잰다.

Andover호

X를 고장 나는 Andover호 엔진의 개수라 하면 $X \sim \mathrm{B}(4, p)$이고, C_A를 수리 비용이라 하자. 그러면

$$C_A = \begin{cases} 0 & X = 0, \\ K & X = 1, \\ 4K & X \geq 2. \end{cases}$$

이다. 이제 $\mathrm{P}(X=0)=q^4$, $\mathrm{P}(X=1)=4pq^3$이므로 $\mathrm{P}(X\geq 2)=1-q^4-4pq^3$이다. 따라서

$$\begin{aligned}\mathrm{E}(C_A) &= K\cdot 4pq^3 + 4K(1-q^4-4pq^3)\\ &= 4K(1-q^4)-12Kpq^3.\end{aligned}$$

이다. $1-q^4=(1-q)(1+q+q^2+q^3)=p(1+q+q^2+q^3)$이므로

$$\begin{aligned}\mathrm{E}(C_A) &= 4Kp(1+q+q^2+q^3)-4Kp\cdot 3q^3\\ &= 4Kp(1+q+q^2-2q^3),\end{aligned}$$

이고, 이것이 바로 구하려던 $(*)$이다.

Basingstoke호

$Y\sim \mathrm{B}(6,p)$를 고장 나는 Basingstoke호 엔진의 개수, C_B를 수리 비용이라 하자. 그러면 $Y\leq 2$일 때 $C_B=YK$이고, $Y\geq 3$일 때 $C_B=6K$이다.

$$\mathrm{P}(Y=0)=q^6,\qquad \mathrm{P}(Y=1)=6pq^5,\qquad \mathrm{P}(Y=2)=15p^2q^4,$$

이므로

$$\begin{aligned}\mathrm{E}(C_B) &= K\cdot 6pq^5 + 2K\cdot 15p^2q^4 + 6K(1-q^6-6pq^5-15p^2q^4)\\ &= 6K(1-q^6)-30Kpq^5-60Kp^2q^4.\end{aligned}$$

을 얻는다. $1-q^6=p(1+q+q^2+q^3+q^4+q^5)$를 이용하여 $6Kp$로 묶으면

$$\mathrm{E}(C_B)=6Kp(1+q+q^2+q^3+q^4+q^5-5q^5-10pq^4).$$

이다.

> **생각해 보기** 괄호 안에 아직 p가 남아 있다. 이것을 어떻게 "$(*)$와 비슷한 형태"로 만들 수 있을까?

형태 $(*)$는 인수 p에 q만의 다항식을 곱한 꼴이므로, 남은 p를 $1-q$로 바꾼다. 즉 $10pq^4=10q^4-10q^5$이다. 항을 정리하면

$$\boxed{\mathrm{E}(C_B)=6Kp(1+q+q^2+q^3-9q^4+6q^5).}$$

이다.

Andover호의 비용이 Basingstoke호 비용의 $\frac{2}{3}$이 되는 때는 언제인가?

$\mathrm{E}(C_A)=\frac{2}{3}\mathrm{E}(C_B)$, 즉

$$4Kp(1+q+q^2-2q^3)=4Kp(1+q+q^2+q^3-9q^4+6q^5).$$

이어야 한다. 양변에 인수 $4Kp$가 있고 $K > 0$이므로, 이 식은 $p = 0$이거나

$$-2q^3 = q^3 - 9q^4 + 6q^5 \iff 3q^3 - 9q^4 + 6q^5 = 0 \iff 3q^3(1-q)(1-2q) = 0.$$

일 때, 그리고 그때에만 성립한다. 따라서 $q = 0$, $q = 1$ 또는 $q = \frac{1}{2}$이고, 이는 각각 $p = 1$, $p = 0$, $p = \frac{1}{2}$에 해당한다. 세 값 모두 $[0,1]$에 속하므로 구하는 값은

$$p = 0, \qquad p = \tfrac{1}{2}, \qquad p = 1.$$

이다. $p = 0$은 시시한 경우이다. 어떤 엔진도 고장 나지 않으므로 두 기대 비용이 모두 0이다. $p = 1$이면 모든 엔진이 고장 나서 비용은 각각 $4K$와 $6K$이다. 흥미로운 값은 $p = \frac{1}{2}$이다.

검산

$p = \frac{1}{2}$일 때 $(*)$에서 $\mathrm{E}(C_A) = 2K(1 + \frac{1}{2} + \frac{1}{4} - \frac{1}{4}) = 3K$이고, Basingstoke호의 공식에서 $\mathrm{E}(C_B) = 3K(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} - \frac{9}{16} + \frac{6}{32}) = \frac{9}{2}K$이며, 실제로 $3 = \frac{2}{3} \cdot \frac{9}{2}$이다. 또한 두 공식을 p의 유리수 값 200개에 대하여 이항분포 위에서 직접 모든 경우를 나열한 결과와 정확히 비교하여 확인했고, 시뮬레이션(simulation)으로도 확인했다. $p = 0.3$에서 2×10^5번의 모의 발사로 얻은 평균 비용은 $1.811K$와 $2.483K$였고, 공식의 값은 $1.805K$와 $2.485K$이다.

흔한 실수

- Andover호 엔진이 두 개 이상 고장 났을 때 비용을 $4K$가 아니라 XK로 계산하는 것. 기준을 넘으면 멀쩡한 엔진까지 포함하여 **모든** 엔진을 교체한다.
- Basingstoke호의 괄호 안에 p를 남겨 두어, 답이 요구된 형태가 되지 않는 것.
- 마지막 방정식에서 p로 나누어 $p = 0$을 잃어버리거나, q^3으로 나누어 $p = 1$을 잃어버리는 것.

STEP 너머

이것은 임계값 정책(threshold policy)에서의 기대 비용을 보여 주는 작은 예로, 신뢰성 이론(reliability theory)과 운용과학(operational research)의 기본 중의 기본이다. 관리자는 멀쩡한 부품을 교체하는 비용과 더 잦은 수리 비용 사이의 균형을 맞추도록 임계값을 정한다. 여기서 사용한 꼬리 합(tail sum) 조작, 즉 $\mathrm{P}(X \geq k)$를 1에서 처음 몇 항을 뺀 것으로 쓰는 방법은, 음이 아닌 정수 값을 갖는 확률변수에 대하여 $\mathrm{E}(X) = \sum_{k \geq 1} \mathrm{P}(X \geq k)$를 주는 것과 같은 아이디어이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 이러한 아이디어를 꼼꼼하게 전개한다.

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.