풀이

STEP II 1994, 문제 14

확률과 통계 • 기하학적 확률

> **STEP II 1994, 14번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/94-s2-q14

핵심 아이디어

"원판 위에서 균등(uniform)하다"는 것은 확률이 넓이에 비례한다는 뜻이다. 따라서 중심에서 거리 r 이내에 떨어질 확률은 작은 원판의 넓이를 과녁의 넓이로 나눈 값이고, 이로부터 곧바로 R의 분포함수(distribution function)를 얻는다. 그러면 (ii)는 한 번 던질 때의 확률 p 하나, 기하분포 유형의 "적어도 한 번 성공" 계산, 그리고 정수에서 정의된 함수의 최소화로 귀결되며, 이 최소화는 미적분 대신 이웃한 값의 차를 살펴서 해결한다.

풀이

소문항 (i)

각 다트를 반지름 a인 원판 위에 균등하게 분포하는 점으로 모형화한다. 그러면 다트가 과녁의 어떤 영역에 떨어질 확률은 그 영역의 넓이를 πa^2으로 나눈 값이다.

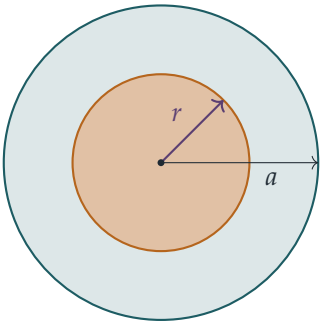
$0 \le r \le a$일 때, 사건 $R \le r$는 다트가 반지름 r인 동심원판 안에 떨어진다는 뜻이므로

$$\mathrm{P}(R \le r) = \frac{\pi r^2}{\pi a^2} = \frac{r^2}{a^2}.$$

이를 미분하면 R의 확률밀도함수(probability density function)는

$$f(r) = \frac{2r}{a^2} \qquad (0 \le r \le a)$$

이고, 그 밖의 범위에서는 $f(r) = 0$이다.



R는 $[0, a]$에서 균등분포를 따르지 **않는다**는 점에 주의한다. 반지름이 클수록 넓이가 더 많기(가는 고리가 더 길기) 때문이며, 그래서 밀도가 r에 비례하여 커진다.

이제

$$\mathrm{E}(R) = \int_0^a r \cdot \frac{2r}{a^2}\,\mathrm{d}r = \frac{2}{a^2}\cdot\frac{a^3}{3} = \frac{2a}{3},$$
$$\mathrm{E}(R^2) = \int_0^a r^2 \cdot \frac{2r}{a^2}\,\mathrm{d}r = \frac{2}{a^2}\cdot\frac{a^4}{4} = \frac{a^2}{2},$$

이므로

$$\mathrm{Var}(R) = \mathrm{E}(R^2) - (\mathrm{E}(R))^2 = \frac{a^2}{2} - \frac{4a^2}{9} = \frac{a^2}{18}$$

이 되어 요구된 결과를 얻는다.

소문항 (ii)

(문제의 "with $a/\sqrt{10}$"은 "within $a/\sqrt{10}$"으로 읽는다.) (i)에 의해, 한 번 던진 다트가 중심에서 $a/\sqrt{10}$ 이내에 떨어질 확률은

$$p = \frac{(a/\sqrt{10})^2}{a^2} = \frac{1}{10}$$

이다. 각 던지기는 독립(independent)이라고 가정한다(매번 과녁 위의 새로운 균등 무작위 점이다). 그러면 m번 던져서 한 번도 $a/\sqrt{10}$ 이내에 들지 못할 확률은 $(0.9)^m$이므로

$$\mathrm{P}(\text{참가자가 돌려받는 상금} £12) = 1 - (0.9)^m.$$

손실은 지불한 $£(12+m)$에서 돌려받는 금액을 뺀 것이므로, 파운드 단위의 기대 손실은

$$L(m) = 12 + m - 12(1 - (0.9)^m) = m + 12(0.9)^m$$

이다. 한 번 더 던질 때마다 확실히 £1이 들지만 £12를 받을 가능성은 커진다. 이 둘이 균형을 이루는 지점을 찾아야 한다.

? 생각해 보기 자연수에서만 정의된 함수는 어떻게 최소화할까?

이웃한 값을 비교한다. $m \geq 1$일 때

$$L(m+1) - L(m) = 1 + 12(0.9)^m(0.9-1) = 1 - 1.2\,(0.9)^m.$$

m이 커질수록 $(0.9)^m$은 감소하므로 이 차는 m에 대하여 증가한다. 따라서 L은 차가 음수인 동안 감소하고, 차가 양수가 되면 증가한다. 차가 음수인 것은 정확히 $(0.9)^m > \frac{1}{1.2} = \frac{5}{6}$일 때이다. 그런데

$$(0.9)^1 = 0.9 > \tfrac{5}{6}, \qquad (0.9)^2 = 0.81 < \tfrac{5}{6}$$

이므로 $L(2) - L(1) < 0$이고, 모든 $m \geq 2$에 대하여 $L(m+1) - L(m) > 0$이다. 따라서

$$L(1) > L(2) < L(3) < L(4) < \cdots$$

이고, 실제로 $L(1) = 11.80$, $L(2) = 11.72$, $L(3) = 11.748$이다. (참가자는 반드시 게임을 하기로 했으므로 $m = 0$은 허용되지 않는다. 어차피 $L(0) = 12$가 더 크다.)

참가자는 $m = 2$번 던지기를 선택해야 하며, 이때 기대 손실은 £11.72이다. 최솟값이 엄격하므로 이러한 값은 $m = 2$뿐이다.

검산

- 단위 원판 위의 균등한 점 4×10^5개로 모의실험을 하면 평균 0.667, 제곱의 평균 0.500, 분산(variance) $0.0556 \approx \frac{1}{18}$을 얻어 $\frac{2}{3}$, $\frac{1}{2}$, $\frac{1}{18}$과 일치한다. 또한 점의 약 10%가 중심에서 $1/\sqrt{10}$ 이내에 떨어졌다.
- 정확한 값은 $L(1) = \frac{59}{5}$, $L(2) = \frac{293}{25}$, $L(3) = \frac{2937}{250}$이며, $1 \leq m \leq 99$ 범위를 모두 조사하면 $m = 2$가 최솟값임이 확인된다.
- 타당성 점검: 분산 $a^2/18$은 $[0, a]$에서의 균등분포 확률변수의 분산 $a^2/12$보다 훨씬 작다. R가 가장자리 쪽에 몰려 있기 때문이다.

흔한 실수

- R가 $[0, a]$에서 균등분포를 따른다고 생각하여 분산을 $a^2/12$로 얻는 것.
- 고정된 £12가 비용의 일부일 뿐이라는 점을 잊는 것. 손실은 단순히 m이 아니라 $12 + m$에서 기대 상금을 뺀 것이다.
- $m + 12(0.9)^m$의 도함수를 0으로 놓으면 $m \approx 2.3$을 얻는데, 이는 답이 아니다. 그 양쪽의 정숫값을 비교해야 한다.

STEP 너머

평면 위의 균등한 점으로부터 R의 분포를 구하는 것은 결합밀도(joint density)를 극좌표(polar coordinates)로 변환하는 가장 간단한 예이다. 여기서 $\mathrm{d}A = r\,\mathrm{d}r\,\mathrm{d}\theta$에 나타나는 야코비안(Jacobian) 인자 r가 바로 R의 밀도를 r에 비례하게 만드는 원인이다. 결합밀도의 변수변환(change of variables)과 이산적인 선택에 대한 최적화는 대학 1학년 확률론의 표준 주제이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)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.