풀이

STEP II 1989, 문제 15

확률과 통계 • 기하학적 확률

> **STEP II 1989, 15번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/89-s2-q15

핵심 아이디어

"둘레(perimeter) 위에서 임의로" 고른 점은 길이에 대하여 균등(uniform)하게 분포하므로, 모든 계산은 전체 길이 $\pi + 2$에서 출발한다. 삼각형은 언제나 중심 O를 한 꼭짓점으로 가지므로, 그 넓이는 $\frac{1}{2}ab\sin C$이다. 여기서 a와 b는 O로부터의 거리이고 C는 O에서의 각이다. 그러면 각 조건부 경우는 간단한 기댓값(expectation) 계산으로 귀결되고, 전체 기댓값의 법칙(law of total expectation)이 이들을 하나로 이어 준다.

풀이

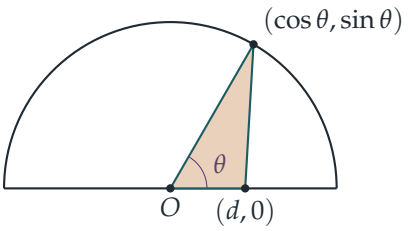
O를 원점에 두고, 지름(diameter)은 x축을 따라 $(-1,0)$에서 $(1,0)$까지 놓으며, 호(arc)는 $y \geq 0$ 영역에 놓는다. 둘레의 길이는 $2+\pi$이다. 지름이 2, 호가 π를 차지한다. 각 점은 둘레를 따라 균등하게 분포하므로

$$\mathrm{P}(\text{점이 지름 위에 있다}) = \frac{2}{\pi+2}, \qquad \mathrm{P}(\text{점이 호 위에 있다}) = \frac{\pi}{\pi+2}.$$

점이 호 위에 있다는 조건 아래에서 그 점은 $(\cos\theta, \sin\theta)$이고, θ는 $[0,\pi]$에서 균등하게 분포한다. 점이 지름 위에 있다는 조건 아래에서 그 점은 $(d,0)$이고, d는 $[-1,1]$에서 균등하게 분포한다.

두 점은 서로 독립(independent)이며, "정확히 한 점이 지름 위에 있다"는 사건은 첫 번째 점이 지름 위에 있는 경우와 두 번째 점이 지름 위에 있는 경우 모두로 일어날 수 있으므로

$$\mathrm{P}(\text{정확히 한 점이 지름 위에 있다}) = 2 \cdot \frac{2}{\pi+2} \cdot \frac{\pi}{\pi+2} = \frac{4\pi}{(\pi+2)^2}.$$



소문항 (i)

한 점을 지름 위의 $(d,0)$, 다른 점을 호 위의 $(\cos\theta,\sin\theta)$라 하자. O에서 나가는 두 변의 길이는 $|d|$와 1이고, 그 사이의 각은 θ이다($d<0$이면 그 각은 $\pi-\theta$인데, 사인값은 같다). 따라서

$$A=\tfrac{1}{2}\,|d|\sin\theta.$$

이 경우라는 조건 아래에서 d와 θ는 독립이므로, 곱의 기댓값은 기댓값의 곱이다. d가 $[-1,1]$에서 균등하게 분포하므로 $|d|$는 $[0,1]$에서 균등하게 분포하고 $\mathrm{E}|d|=\frac{1}{2}$이다. 또한

$$\mathrm{E}(\sin\theta)=\frac{1}{\pi}\int_0^{\pi}\sin\theta\,\mathrm{d}\theta=\frac{2}{\pi}.$$

그러므로

$$\mathrm{E}(A\mid \text{정확히 한 점이 지름 위에 있다})=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{2}{\pi}=\frac{1}{2\pi}$$

이고, 이것이 구하려던 결과이다.

소문항 (ii)

이제 두 점이 모두 호 위에 있고, 각각의 각은 θ_1과 θ_2로서 서로 독립이며 $[0,\pi]$에서 균등하게 분포한다. 두 점 모두 O로부터 거리가 1이고, O에서 두 점 사이의 각은 $X=|\theta_1-\theta_2|$이며 이는 $[0,\pi]$에 속한다. 따라서

$$A=\tfrac{1}{2}\sin X.$$

? **생각해 보기** 왜 X에 대하여 주어진 확률밀도함수(density)를 쓸 수 있는가?

각 θ는 단위 반원을 따라 잰 호의 길이이므로, 호 위에서 두 점을 균등하게 고르는 것은 길이가 π인 선분 위에서 두 점을 균등하게 고르는 것과 같다. 각 X는 바로 그 두 점 사이의 거리이므로, X는 $[0,\pi]$에서 확률밀도함수 $2(\pi-x)/\pi^2$를 가진다. 따라서

$$\mathrm{E}(\sin X)=\frac{2}{\pi^2}\int_0^{\pi}(\pi-x)\sin x\,\mathrm{d}x.$$

적분을 나누면 $\int_0^{\pi}\pi\sin x\,\mathrm{d}x=2\pi$이고, $u=x$로 두고 부분적분(integration by parts)하면

$$\int_0^{\pi}x\sin x\,\mathrm{d}x=\Big[-x\cos x\Big]_0^{\pi}+\int_0^{\pi}\cos x\,\mathrm{d}x=\pi+0=\pi.$$

그러므로 $\int_0^\pi (\pi - x)\sin x\,\mathrm{d}x = 2\pi - \pi = \pi$이고, $\mathrm{E}(\sin X) = \dfrac{2}{\pi}$이며

$$\mathrm{E}(A \mid \text{어느 점도 지름 위에 있지 않다}) = \frac{1}{2}\cdot\frac{2}{\pi} = \frac{1}{\pi}$$

이고, 이것이 구하려던 결과이다.

A의 기댓값

세 번째 경우가 있다. 두 점이 모두 지름 위에 있으면 "삼각형"은 퇴화(degenerate)한다(세 점이 모두 x축 위에 있다). 따라서 $A = 0$이다. 어느 점도 지름 위에 있지 않을 확률은 $(\frac{\pi}{\pi+2})^2$이다. 전체 기댓값의 법칙에 의하여

$$\begin{aligned}\mathrm{E}(A) &= \frac{4\pi}{(\pi+2)^2}\cdot\frac{1}{2\pi} + \frac{\pi^2}{(\pi+2)^2}\cdot\frac{1}{\pi} + \frac{4}{(\pi+2)^2}\cdot 0\\ &= \frac{2+\pi}{(\pi+2)^2}\\ &= \frac{1}{2+\pi}\end{aligned}$$

이고, 이것이 우리가 원하던 결과이다.

검산

- 세 경우의 확률 $\frac{4}{(\pi+2)^2}$, $\frac{4\pi}{(\pi+2)^2}$, $\frac{\pi^2}{(\pi+2)^2}$를 더하면 $\frac{(2+\pi)^2}{(\pi+2)^2} = 1$이다.
- 당연히 그래야 하듯이 $\mathrm{E}(A) \le \frac{1}{2}$이다. 이 값은 가능한 가장 큰 넓이(두 반지름이 직각을 이룰 때)이다. 실제로 $\frac{1}{2+\pi} \approx 0.1945$이다.
- 점의 쌍 2×10^6개로 몬테카를로 시뮬레이션(Monte Carlo simulation)을 한 결과 P(정확히 한 점) ≈ 0.4757(정확한 값 0.4754), 조건부 평균 0.1591과 0.3184(정확한 값 $\frac{1}{2\pi} \approx 0.1592$와 $\frac{1}{\pi} \approx 0.3183$), 그리고 $\mathrm{E}(A) \approx 0.1944$를 얻었다.

흔한 실수

- P(정확히 한 점)에서 인수 2를 빠뜨리는 것. 두 점 중 어느 것이든 지름 위에 있는 점이 될 수 있다.
- "지름 위에 있을" 확률을 $\frac{1}{2}$로 보는 것. 점은 경계의 어느 조각인지가 아니라 길이에 따라 균등하게 고른다.
- 소문항 (ii)에서 반지름 **사이의** 각의 사인 대신 $\sin\theta_1 \sin\theta_2$ 같은 것을 쓰는 것.

STEP 너머

기댓값을 경우들의 분할(partition)에 따라 나누어 $\mathrm{E}(A) = \sum_i \mathrm{E}(A \mid B_i)\,\mathrm{P}(B_i)$로 쓰는 것이 전체 기댓값의 법칙이며, 대학 1학년 확률론의 핵심 도구 가운데 하나이다. 그 일반형 $\mathrm{E}(A) = \mathrm{E}\big(\mathrm{E}(A \mid Y)\big)$는 분기 과정(branching process)에서 마르코프 연쇄(Markov chain)에 이르기까지 어디에서나 등장한다. 소문항 (ii)에서 인용한 X의 확률밀도함수는 그 자체로 좋은 연습 문제이다. 이는 정사각형 $[0, \pi]^2$ 안의 띠 모양 영역의 넓이에서 나온다. 이와 같은 문제는 기하 확률(geometric probability)에 속한다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 조건부 기댓값(conditional expectation)과 결합 확률밀도함수(joint density)를 알맞은 수준에서 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.