풀이

STEP II 1997, 문제 13

확률과 통계 • 뷔퐁의 바늘

STEP II 1997, 13번 문제

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

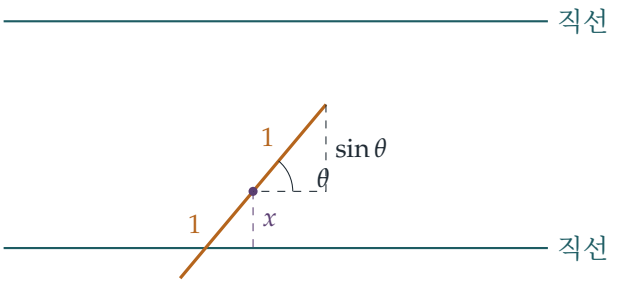
step.maths.org/questions/97-s2-q13

핵심 아이디어

이것은 뷔퐁(Buffon)의 바늘(Buffon's needle) 문제이다. 바늘의 위치를 서로 독립(independent)인 두 확률적인 양으로 나타낸다. 하나는 바늘의 중심에서 가장 가까운 직선까지의 거리이고, 다른 하나는 바늘이 직선들과 이루는 예각이다. 중심을 고정하고 나면, 모든 것은 바늘이 직선에 수직인 방향으로 얼마나 멀리 뻗는지에 달려 있다. 그 길이는 바늘 길이의 절반에 그 각의 사인을 곱한 값이다. 그러면 각 소문항은 균등분포(uniform distribution)를 따르는 각에 관한 문제가 된다.

풀이

"무작위로 떨어뜨린다"를 보통의 방식으로 모형화한다. 중심에서 가장 가까운 직선까지의 거리 X는 $[0,1]$에서 균등분포를 따르고(직선 사이의 간격이 2 cm이므로 가장 가까운 직선은 많아야 1 cm 떨어져 있다), 바늘과 직선이 이루는 예각 θ는 X와 독립으로 $[0,\frac{\pi}{2}]$에서 균등분포를 따른다. (각을 $[0,\pi)$에서 균등하게 잡아도 같은 답이 나온다. 어느 쪽이든 $\sin\theta$의 분포가 같기 때문이다.)



바늘의 각 절반은 길이가 1이므로, 양 끝은 중심의 위와 아래로 수직 거리 $\sin\theta$만큼 떨어져 있다.

소문항 (i)

중심이 가장 가까운 직선에서 x만큼 떨어져 있을 때, 바늘이 그 직선과 만나는 것은 한쪽 끝이 그 직선을 넘어설 때, 곧

$$\sin\theta > x \iff \theta > \arcsin x,$$

일 때와 정확히 같다. $[0, \frac{\pi}{2}]$에서 sin이 증가하기 때문이다. θ가 $[0, \frac{\pi}{2}]$에서 균등분포를 따르므로

$$\mathrm{P}(\text{가장 가까운 직선과 만남} \mid X = x) = \frac{\frac{\pi}{2} - \arcsin x}{\frac{\pi}{2}} = 1 - \frac{2}{\pi}\arcsin x = \frac{2}{\pi}\arccos x.$$

소문항 (ii)

바늘이 가장 가까운 직선과 만난다고 하자. 즉 $\theta > \arcsin x$이다. 바늘을 따라 재면, 교점은 중심에서 $x/\sin\theta$만큼 떨어져 있다(θ의 대변이 x인 직각삼각형의 빗변이다). 따라서 두 조각의 길이는 $1 + x/\sin\theta$와 $1 - x/\sin\theta$이고, 짧은 쪽은

$$S = 1 - \frac{x}{\sin\theta}.$$

이다. θ가 $\arcsin x$에서 $\frac{\pi}{2}$까지 움직이면 S는 0에서 $1 - x$까지 증가하므로, S는 $[0, 1 - x]$의 값을 가진다.

만난다는 조건(conditioning) 아래에서 θ는 길이가 $\frac{\pi}{2} - \arcsin x = \arccos x$인 구간 $(\arcsin x, \frac{\pi}{2}]$에서 균등분포를 따른다. $0 \le s \le 1 - x$에 대하여

$$S \le s \iff \frac{x}{\sin\theta} \ge 1 - s \iff \sin\theta \le \frac{x}{1-s} \iff \theta \le \arcsin\frac{x}{1-s},$$

이다. 여기서 $1 - s > 0$과 $x/(1-s) \le 1$($s \le 1 - x$이므로)을 사용했으므로 아크사인이 정의된다. 따라서 짧은 조각의 누적분포함수(cumulative distribution function)는

$$F(s) = \mathrm{P}(S \le s) = \begin{cases} 0, & s < 0, \\ \dfrac{\arcsin(\frac{x}{1-s}) - \arcsin x}{\arccos x}, & 0 \le s \le 1 - x, \\ 1, & s > 1 - x. \end{cases}$$

이다. 확인해 보면, 가운데 식은 $s = 0$에서 0이고 $s = 1 - x$에서 $(\frac{\pi}{2} - \arcsin x)/\arccos x = 1$이므로 F는 연속이다.

생각해 보기 바늘이 직선과 만난다는 사실을 알고 나서도 θ가 여전히 균등분포를 따르는 이유는 무엇인가?

균등분포를 따르는 확률변수에 부분구간에 놓인다는 조건을 주면, 그 부분구간에서 여전히 균등분포를 따른다. $\arcsin x \le a \le \frac{\pi}{2}$에 대하여 $\mathrm{P}(\theta \le a \mid \theta > \arcsin x) = (a - \arcsin x)/(\frac{\pi}{2} - \arcsin x)$이고, 이는 a에 대한 일차식이다. 이것이 바로 위에서 사용한 사실이다.

소문항 (iii)

가장 가까운 직선은 거리 $X \le 1$에 있고, 다른 모든 직선은 적어도 거리 $2 - X \ge 1$에 있다. 바늘은 중심에서 많아야 $\sin\theta \le 1$만큼 뻗으므로, 가장 가까운 직선이 아닌 다른 직선과 만나려면 정확히 $X = 1$이고 $\theta = \frac{\pi}{2}$이어야 하며, 그 확률은 0이다. 따라서 바늘이 모든 직선과 만나지 않는 것은 가장 가까운 직선과 만나지 않는 것과 정확히 같고, 소문항 (i)에 의하여

$$\mathrm{P}(\text{만나지 않음} \mid X = x) = \frac{2}{\pi}\arcsin x.$$

이다. X는 $[0,1]$에서 균등분포를 따르므로(확률밀도 1), x에 대하여 평균을 낸다.

$$\mathrm{P}(\text{만나지 않음}) = \int_0^1 \frac{2}{\pi}\arcsin x \,\mathrm{d}x.$$

$u = \arcsin x$, $\mathrm{d}v = \mathrm{d}x$로 두고 부분적분(integration by parts)하면

$$\begin{aligned}\int_0^1 \arcsin x \,\mathrm{d}x &= \Big[x\arcsin x\Big]_0^1 - \int_0^1 \frac{x}{\sqrt{1-x^2}}\,\mathrm{d}x \\ &= \frac{\pi}{2} - \Big[-\sqrt{1-x^2}\Big]_0^1 = \frac{\pi}{2} - 1.\end{aligned}$$

이다. 그러므로

$$\mathrm{P}(\text{바늘이 모든 직선과 만나지 않음}) = \frac{2}{\pi}\Big(\frac{\pi}{2} - 1\Big) = 1 - \frac{2}{\pi} \approx 0.363.$$

검산

- 소문항 (i)의 결과는 양 끝에서 타당하게 행동한다. $x = 0$이면 중심이 직선 위에 있고 확률은 1이다. $x \to 1$이면 수직으로 놓인 바늘만 직선에 닿으므로 확률은 0으로 다가간다.
- 만날 확률은 $2/\pi \approx 0.637$이고, 이는 뷔퐁의 고전적인 답 $2\ell/(\pi d)$에 $\ell = d = 2$를 넣은 값이다.
- 무작위 투하 10^6번을 시뮬레이션한 결과 만나지 않는 비율은 0.3626으로, $1 - 2/\pi = 0.3634$와 비교된다. $x = 0.2, 0.5, 0.7$일 때 시뮬레이션으로 얻은 $F(s)$의 값은 공식과 약 0.002 이내로 일치했다.

흔한 실수

- 소문항 (ii)에서 조건을 거는 것을 잊는 실수. $\arccos x$ 대신 $\frac{\pi}{2}$로 나누면 결코 1에 도달하지 않는 "분포함수"가 나온다.
- 짧은 조각을 $1 - x\sin\theta$나 $1 - x/\cos\theta$로 잡는 실수. 삼각형의 어느 변이 x이고 어느 각이 θ인지 확인하라.
- 소문항 (iii)에서 만나는 조건을 조정하지 않은 채 $\frac{1}{2}$을 곱하거나 X가 $[0,2]$에서 균등분포를 따른다고 두는 실수. 가장 가까운 직선까지의 거리는 $[0,1]$에서 균등분포를 따른다.

STEP 너머

뷔퐁의 바늘(1777)은 기하학적 확률(geometric probability)의 최초의 문제 가운데 하나이며, π가 나타난다는 것은 바늘을 떨어뜨려서 π의 (느린) 몬테카를로(Monte Carlo) 추정값을 얻을 수 있다는 뜻이다. 가장 깔끔한 현대적 접근은 기댓값의 선형성(linearity of expectation)을 이용한다. 교점 개수의 기댓값은 바늘의 길이에 비례하며, 이는 구부러진 바늘("뷔퐁의 국수(Buffon's noodle)")에 대해서도 성립한다. 이 사실은 적분기하학(integral geometry)과 크로프톤 공식(Crofton formula)으로 이어진다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 결합분포(joint distribution)와 이런 종류의 기하학적 확률 문제를 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.