풀이

# STEP III 2012, 문제 12

확률과 통계 • 기하학적 확률

> **STEP III 2012, 12번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/12-s3-q12

## 핵심 아이디어

삼각형을 꼭짓점 $P$를 공유하는 세 개의 작은 삼각형으로 나누면, 세 거리의 합이 항상 높이와 같다는 것을 알 수 있다. 최솟값을 다룰 때는 꼬리 확률(tail probability) $\mathrm{P}(X > x)$를 이용하는 것이 깔끔하다. 모든 거리가 $x$보다 크다는 사건은 원래 삼각형과 닮은(similar) 더 작은 삼각형에 해당하며, 닮은 도형의 넓이는 길이의 비의 제곱에 비례한다. 3차원에서도 같은 아이디어가 부피에 대해 성립하며, 이때 부피는 길이의 비의 세제곱에 비례한다.

## 풀이

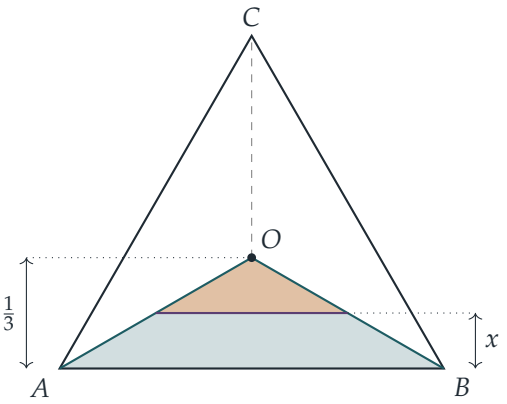
### 소문항 (i)

삼각형의 한 변의 길이를 $s$라 하자. 높이가 1이므로 $ABC$의 넓이는 $\frac{1}{2}s$이다. $P$를 세 꼭짓점과 이으면 $ABC$는 삼각형 $PAB$, $PBC$, $PCA$로 나뉜다. 이 삼각형들의 밑변은 각각 길이가 $s$인 변이고, 높이는 각각 $x_1$, $x_2$, $x_3$이다. 넓이를 더하면

$$\tfrac{1}{2}sx_1 + \tfrac{1}{2}sx_2 + \tfrac{1}{2}sx_3 = \tfrac{1}{2}s, \qquad \text{따라서} \qquad x_1 + x_2 + x_3 = 1,$$

이 되어 원하는 결과를 얻는다.

**$X_1$이 가장 작은 영역.** $AB$와 $CA$에서 같은 거리에 있는 점들은 각 $A$의 이등분선(angle bisector) 위에 있다. 이 이등분선의 $B$ 쪽에서는 $P$가 $CA$보다 $AB$에 더 가깝다. 마찬가지로 $P$가 $BC$보다 $AB$에 더 가까운 것은 정확히 그 점이 각 $B$의 이등분선의 $A$ 쪽에 있을 때이다. 정삼각형에서 각의 이등분선은 중선(median)이기도 하며, 이들은 중심 $O$에서 만난다. 따라서 $X_1 = \min(X_1, X_2, X_3)$인 것은 정확히 $P$가 삼각형 $ABO$(아래 그림에서 색칠한 부분) 안에 있을 때이다. $O$에서 $AB$까지의 거리는 높이의 3분의 1인 $\frac{1}{3}$이다.




**$X$의 확률밀도함수.** 세 삼각형 $ABO$, $BCO$, $CAO$는 합동(congruent)이므로 각각의 확률은 $\frac{1}{3}$이다. (거리가 같아지는 경우는 이등분선 위에서만 일어나는데, 이등분선의 넓이는 0이다.) 대칭성에 의해 $X$는 $P$가 $ABO$ 안에 있다는 조건 아래에서의 $X_1$과 같은 분포(distribution)를 가진다. 이 경우 $X = X_1$은 $P$에서 $AB$까지의 거리이다.

? **생각해 보기** 삼각형 $ABO$ 안에서 $X_1 > x$를 만족하는 점은 어떤 점들인가?

그 점들은 $AB$에 평행하고 $AB$로부터 거리가 $x$인 직선보다 위에 있는 점들로, 꼭짓점이 $O$인 작은 호박색 삼각형을 이룬다. 이 삼각형은 $ABO$와 닮았고, 높이는 $\frac{1}{3}$에 대하여 $\frac{1}{3} - x$이므로, 그 넓이는 $ABO$의 넓이의 $((\frac{1}{3} - x)/\frac{1}{3})^2 = (1-3x)^2$배이다. 따라서 $0 \le x \le \frac{1}{3}$일 때

$$\mathrm{P}(X > x) = (1-3x)^2, \qquad \mathrm{F}(x) = \mathrm{P}(X \le x) = 1 - (1-3x)^2.$$

또한 세 거리의 합이 1이므로 $X$는 $\frac{1}{3}$을 넘을 수 없다. 미분하면

$$\mathrm{f}(x) = \mathrm{F}'(x) = 6(1-3x) \quad (0 \le x \le \tfrac{1}{3}),$$

이고, 그 밖에서는 $\mathrm{f}(x) = 0$이므로 원하는 결과를 얻는다.

**기댓값.**

$$\mathrm{E}(X) = \int_0^{1/3} 6x(1-3x)\,\mathrm{d}x = \left[3x^2 - 6x^3\right]_0^{1/3} = \tfrac{1}{3} - \tfrac{2}{9} = \tfrac{1}{9}.$$

**소문항 (ii)**

점에서 네 면까지의 거리를 $Y_1, \ldots, Y_4$라 하고, $Y = \min(Y_1, \ldots, Y_4)$라 하자. 각 면의 넓이는 모두 $S$로 같고, 부피는 $\frac{1}{3}S \cdot 1$이다. 점을 네 꼭짓점과 이으면 사면체(tetrahedron)는 네 개의 작은 사면체로 나뉜다. 각각은 한 면을 밑면으로, $Y_i$ 중 하나를 높이로 가지므로

$$\tfrac{1}{3}S(Y_1 + Y_2 + Y_3 + Y_4) = \tfrac{1}{3}S, \qquad \text{따라서} \qquad Y_1 + Y_2 + Y_3 + Y_4 = 1.$$

무게중심(centroid)은 모든 면에서 거리가 $\frac{1}{4}$이므로, 내접구(inscribed sphere)의 반지름은 $\frac{1}{4}$이다.

**?** **생각해 보기** 이번에는 집합 $\{Y > y\}$가 어떤 모양인가?

이 집합은 네 면 모두로부터 거리가 $y$보다 큰 점들의 집합이다. 각 면을 자신과 평행하게 안쪽으로 $y$만큼 옮기면, 무게중심이 같은 더 작은 정사면체(regular tetrahedron)가 생긴다. 무게중심에서 각 면까지의 거리가 $\frac{1}{4}$에서 $\frac{1}{4} - y$로 줄어들므로 닮음비(scale factor)는 $(\frac{1}{4} - y)/\frac{1}{4} = 1 - 4y$이다. 부피는 닮음비의 세제곱에 비례하므로, $0 \le y \le \frac{1}{4}$일 때

$$\mathrm{P}(Y > y) = (1 - 4y)^3.$$

(같은 논증을 소문항 (i)의 삼각형에 적용하면 곧바로 $(1 - 3x)^2$를 얻는다. 안쪽 삼각형의 내접원의 반지름(inradius)은 $\frac{1}{3} - x$이기 때문이다.)

$Y \ge 0$이므로 꼬리 확률을 적분하여 $\mathrm{E}(Y)$를 구할 수 있다.

$$\mathrm{E}(Y) = \int_0^{1/4} \mathrm{P}(Y > y)\,\mathrm{d}y = \int_0^{1/4} (1 - 4y)^3\,\mathrm{d}y = \left[-\tfrac{1}{16}(1 - 4y)^4\right]_0^{1/4} = \tfrac{1}{16}.$$

확률밀도함수를 쓰고 싶다면, 확률밀도함수는 $[0, \frac{1}{4}]$에서 $12(1 - 4y)^2$이고, 이 경우에도 $\int_0^{1/4} 12y(1 - 4y)^2\,\mathrm{d}y = \frac{1}{16}$이다.

따라서 가장 가까운 면까지의 거리의 기댓값은 $\frac{1}{16}$이다.

## 검산

- $\int_0^{1/3} 6(1 - 3x)\,\mathrm{d}x = 1$이므로 f는 실제로 확률밀도함수이다.
- 규칙성이 있다. 높이가 1인 삼각형에서는 $\mathrm{E}(X) = \frac{1}{9} = \frac{1}{3^2}$이고, 높이가 1인 사면체에서는 $\mathrm{E}(Y) = \frac{1}{16} = \frac{1}{4^2}$이다. 1차원에서는 $[0, 1]$의 임의의 점에서 더 가까운 끝점까지의 거리의 평균이 $\frac{1}{4} = \frac{1}{2^2}$이므로 규칙에 들어맞는다.
- 삼각형 안의 임의의 점 $4 \times 10^5$개로 몬테카를로 시뮬레이션(Monte Carlo simulation)을 하였더니 평균은 $0.1110$($\frac{1}{9} \approx 0.1111$과 비교)이었고, 히스토그램은 $6(1 - 3x)$와 잘 맞았다. 사면체 안의 점 $3 \times 10^5$개로 하였을 때 평균은 $0.0625$($\frac{1}{16} = 0.0625$와 비교)였다.

## 흔한 실수

- $\mathrm{P}(X > x)$를 미분하면서 음의 부호를 빠뜨리는 것. 확률밀도함수는 $-\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{P}(X > x)$이다.
- 안쪽 삼각형의 높이가 $1 - x$가 아니라 $1 - 3x$라는 것을 잊는 것. 세 변이 각각 $x$만큼 안쪽으로 이동한다.

▶ 소문항 (ii)에서 닮음비의 제곱을 쓰는 것. 부피는 세제곱에 비례한다.

## STEP 너머

이 문제에 나온 두 가지 도구는 대학의 첫 확률론 강의에서 끊임없이 등장한다. 하나는 $\mathrm{P}(\min > x)$를 통해 최솟값의 분포를 구하는 것이고, 다른 하나는 음이 아닌 확률변수(random variable)에 대하여 $\mathrm{E}(Y) = \int_0^\infty \mathrm{P}(Y > y)\,\mathrm{d}y$를 계산하는 것이다. 규칙성은 계속 이어진다. 높이가 1인 $n$차원 단체(simplex)에서 가장 가까운 면까지의 거리의 기댓값은 $1/(n+1)^2$이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 이 두 도구를 꼼꼼하게 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.