12.19 평면 위의 무작위 점

학습지

문제

STEP II 2012, 13번 문제

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

step.maths.org/questions/12-s2-q13

힌트

먼저 혼자 힘으로 문제를 풀어 보자. 막히면 힌트를 **하나씩** 읽고, 힌트를 하나 읽을 때마다 다시 문제로 돌아가자. **잠깐, 직접 해 보기**라고 표시된 힌트가 핵심 단계이다. 각 힌트 뒤에는 풀이를 적을 공간이 있다.

생각해 보기 평균이 μ인 푸아송 확률변수(Poisson random variable)가 값 0을 가질 확률은 얼마인가?

1단계 가까운 곳에 슈퍼마켓이 없을 확률

다루는 내용: 원 안에 슈퍼마켓이 없을 확률

생각해 보기 반지름이 y인 원 안에 있는 슈퍼마켓의 개수는 푸아송 분포(Poisson distribution)를 따른다. 그 평균(mean)은 얼마인가? 그 개수가 0일 확률(probability)은 얼마인가?

2단계 가장 가까운 거리의 분포

다루는 내용: Y의 분포

잠깐, 직접 해 보기 $Y \geq y$는 거리 y 이내에 슈퍼마켓이 하나도 없다는 뜻이다. 그렇다면 $\mathrm{P}(Y < y)$는 무엇인가? 이를 미분(differentiate)하여 확률밀도함수(probability density function)를 구한다.

3단계 평균

다루는 내용: 평균

생각해 보기 $\mathrm{E}(Y)$를 적분(integral)으로 나타낸다. 지수(exponent)가 $-\frac{1}{2}x^2$이 되도록 치환(substitution)하고, 문제에 주어진 적분을 이용한다.

4단계 분산

다루는 내용: 분산

? **생각해 보기** $\mathrm{E}(Y^2)$의 경우 피적분함수(integrand)에서 y의 지수가 홀수이므로 $u = \pi k y^2$으로 치환한다. 그런 다음 $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \big(\mathrm{E}(Y)\big)^2$을 이용하여 분산(variance)을 구한다.

숨은 수학

평면 위의 푸아송 과정

강도(intensity)가 k인 공간 푸아송 과정이란, 서로소인(disjoint) 영역들에 들어 있는 점의 개수가 서로 독립(independent)이고, 넓이가 A인 영역에 들어 있는 점의 개수가 평균 kA인 푸아송 분포를 따르는 무작위 점 배치이다. 이것은 매우 많은 점을 매우 넓은 영역에 독립적으로, 균일하게, 알맞은 평균 밀도로 떨어뜨렸을 때 얻어지는 것으로, 교란순열(derangement)(Section 9.8)을 설명하는 것과 같은 "희귀 사건의 법칙(law of rare events)"이다. 이는 완전한 공간적 무작위성을 나타내는 자연스러운 모형이며, 실제 패턴은 이것을 기준으로 평가된다. 두 생태학자가 1954년에 고안한 *Clark–Evans* **검정**(Clark–Evans test)은 관측된 평균 최근접 이웃 거리를 $\frac{1}{2\sqrt{k}}$와 비교한다. 이보다 훨씬 작으면 군집(씨앗이 어미 나무 근처에 떨어지는 경우)을, 훨씬 크면 규칙적인 간격(텃세를 부리는 새, 또는 식물들이 물을 두고 벌이는 경쟁)을 시사한다.

3차원에서

같은 계산을 3차원에서 하면, 가장 가까운 별, 분자, 은하까지의 거리에 대해 $\mathrm{P}(Y > y) = e^{-\frac{4}{3}\pi k y^3}$를 얻는다. 찬드라세카르(Subrahmanyan Chandrasekhar)는 1943년에 이를 이용하여 무작위로 배치된 이웃 별들이 한 별에 미치는 중력을 연구했고, 물리학자들은 이를 기체 분자의 평균 자유 행로(mean free path)에 사용한다.

레일리

밀도 $\frac{y}{\sigma^2}e^{-y^2/(2\sigma^2)}$는 레일리 경(Lord Rayleigh)의 이름을 딴 것이다. 그는 1880년에 소리의 이론에서 이 분포를 만났는데, 위상이 무작위인 많은 파동을 합한 것의 진폭을 나타낸다. 이 분포는 또한 두 좌표가 서로 독립인 정규확률변수(normal random variable)인 점과 원점 사이의 거리를 나타내기도 하는데, 이것이 이 분포가 여기에 나타나는 이유를 보는 또 다른 방법이다(연습문제 3). 오늘날에는 무작위 위상을 가진 여러 경로를 따라 수신기에 도달하는 휴대전화 신호의 페이딩(fading)을 모형화하는 데 쓰인다.

더 도전하기

1 **중앙값.** 가장 가까운 슈퍼마켓까지 거리의 중앙값(median)을 구하고, 평균과 비교하여라.

2 **두 번째로 가까운 것.** 두 번째로 가까운 슈퍼마켓까지 거리의 분포를 구하여라. (힌트: $\mathrm{P}(Y_2 > y)$는 y 이내에 슈퍼마켓이 많아야 하나 있을 확률이다.)

3 **정규분포를 따르는 좌표.** U와 V가 서로 독립이고 $\mathrm{N}(0,\sigma^2)$를 따르면, $\sqrt{U^2+V^2}$의 밀도가 $\frac{r}{\sigma^2}e^{-r^2/(2\sigma^2)}$임을 보여라. σ가 어떤 값일 때 슈퍼마켓 거리의 분포가 되는가?

4 **Clark–Evans.** 나무 400그루가 있는 1 km^{2}의 숲을 조사했더니, 나무에서 가장 가까운 이웃 나무까지의 평균 거리가 35 m였다. 이 결과는 완전한 공간적 무작위성과 부합하는가, 아니면 군집이나 규칙적인 간격을 시사하는가?

더 읽을거리

- **Peter Diggle, *Statistical Analysis of Spatial Point Patterns*** (Arnold). 공간 점 패턴의 통계적 분석에 관한 표준 참고서로, Clark–Evans 검정과 그 후속 방법들을 다룬다.
- **John Kingman, *Poisson Processes*** (Oxford University Press). 임의의 차원에서의 푸아송 과정을 다룬 아름다운 소책자이다.
- **Geoffrey Grimmett & David Stirzaker, *Probability and Random Processes*** (Oxford University Press). 푸아송 과정에 관한 장을 보라.

안내에 따른 전체 풀이는 제 12.19절에 있다:

https://beyondstep.keithmcnulty.org/ko/ch12/spatial-poisson.html

매주 STEP 문제 하나를 끝까지 풀어 본다. The STEP Up은 이메일로 무료로 받아 볼 수 있다: keithmcnulty.substack.com