12.23 확률로 얻는 스털링 공식

학습지

문제

STEP II 2016, 13번 문제

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

step.maths.org/questions/16-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 μ이고 분산이 σ^2인 이산 확률변수(discrete random variable) X가 근사적으로 정규분포를 따른다면, $\mathrm{P}(X = \mu)$는 대략 얼마일까? (μ를 중심으로 하는 폭 1인 구간의 확률을 생각해 보자.)

1단계 동전 열여섯 개

다루는 내용: 소문항 (i)

> **생각해 보기** $\mathrm{B}(16, \frac{1}{2})$의 평균(mean)과 표준편차(standard deviation)를 구한다. $\mathrm{P}(X = 8)$을 정규분포(normal distribution) 확률밀도함수의 봉우리 높이로 근사한다.

2단계 중앙 이항계수

다루는 내용: 소문항 (ii)

잠깐, 직접 해 보기 $\mathrm{B}(2n, \frac{1}{2})$에 대하여 $\mathrm{P}(X = n)$을 정확한 값으로 쓰고, 같은 방법으로 근삿값으로도 쓴다. 두 식을 같다고 놓는다.

3단계 스털링 공식

다루는 내용: 소문항 (iii)

생각해 보기 모수가 n인 푸아송 분포(Poisson distribution)에 대해서도 같은 과정을 거친다. 이 분포의 평균과 분산(variance)은 무엇인가? $\mathrm{P}(X = n)$을 정확한 값과 근삿값으로 쓴다.

숨은 수학

근사는 얼마나 정확한가?

이 논증은 휴리스틱(heuristic)이다. 정규근사를 **한 점에서** 사용하는데, 보통의 중심극한정리(central limit theorem)는 이를 정당화해 주지 않는다. 이를 엄밀하게 다룬 **국소 중심극한정리**(*local central limit theorem*)가 바로 그 일을 해 주며, 이로써 이 논증은 진정한 증명이 된다. 사실 스털링 공식은 놀라울 만큼 정확하다. 비 $\frac{n!}{\sqrt{2\pi n}(n/e)^n}$은 대략 $1 + \frac{1}{12n}$이므로, $n = 1$일 때조차 1 대신 0.92를 주고, $n = 10$일 때는 0.8%만큼 작다. 전체 급수 $1 + \frac{1}{12n} + \frac{1}{288n^2} - \cdots$는 **점근급수**(*asymptotic series*)의 한 예이다. n이 클수록 점점 더 좋은 근삿값을 주지만, n을 고정하면 결국 발산(diverge)한다.

계승과 정규분포가 연결되는 이유

스털링 공식의 $\sqrt{2\pi}$와 정규분포의 확률밀도에 나오는 $\sqrt{2\pi}$는 같은 상수이다. $\int_{-\infty}^{\infty} e^{-x^2/2}\,\mathrm{d}x = \sqrt{2\pi}$이기 때문이다. 드무아브르(de Moivre)는 1733년에 바로 이런 방식으로 이항분포의 정규근사를 발견했는데, 그 과정에서 큰 n에 대한 $\binom{n}{k}$의 근삿값이 필요했다. 그는 $n! \approx Cn^{n+1/2}e^{-n}$ 꼴은 찾았지만 상수는 찾지 못했다. 스털링(James Stirling)이 $C = \sqrt{2\pi}$임을 보였다. 정규분포는 계승의 점근적 성질에서 태어난 것이다.

드무아브르와 스털링

드무아브르(Abraham de Moivre)는 런던으로 망명한 프랑스 개신교도로, 개인 교습과 도박꾼들에게 조언을 해 주며 생계를 꾸렸다. 그는 1718년에 *The Doctrine of Chances*(우연의 이론)를 출간했고, 1733년의 소책자에서 이항분포의 정규근사를 발견했는데, 이것이 정규곡선이 처음 등장한 곳이다. 스코틀랜드의 수학자 스털링은 *Methodus Differentialis*(차분법, 1730)에서 이 상수의 값을 발표했고, 드무아브르는 너그럽게 그의 공로를 인정했다. 계승의 근사는 지금도 활발히 연구되는 주제이다. 아주 큰 n에 대해 $n!$을 계산하는 일은 조합론과 암호학에서 필요하다.

더 도전하기

1 **공식 확인하기.** 스털링 공식으로 $10!$과 $20!$을 추정하고, 참값 $3\,628\,800$, $2\,432\,902\,008\,176\,640\,000$과 비교하여라.

2 **(iii)에서 (ii)로.** 소문항 (iii)의 스털링 공식으로부터 소문항 (ii)의 근사가 나옴을 보여라.

3 **$100!$의 자릿수.** $\log_{10} n! \approx n\log_{10}\frac{n}{e} + \frac{1}{2}\log_{10}(2\pi n)$을 이용하여 $100!$이 158자리 수임을 보여라.

4 **동전 던지기에서의 동점.** 소문항 (ii)를 이용하여, 공정한 동전을 1000번 던질 때 앞면이 정확히 500번 나올 확률이 약 0.025임을 보여라.

더 읽을거리

- **William Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1** (Wiley), 제II장. 스털링 공식을 증명한다.
- **Ronald Graham, Donald Knuth & Oren Patashnik, *Concrete Mathematics*** (Addison-Wesley), 제9장. 점근적 방법을 다룬다.
- **Stephen Stigler, *The History of Statistics*** (Harvard University Press). 드무아브르와 정규곡선의 탄생을 다룬다.

안내에 따른 전체 풀이는 제 12.23절에 있다:

https://beyondstep.keithmcnulty.org/ko/ch12/stirling-probability.html

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