풀이

STEP I 2000, 문제 14

확률과 통계 • 균등분포의 적률

> **STEP I 2000, 14번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/00-s1-q14

핵심 아이디어

여기서 다루는 모든 것은 균등분포(uniform distribution)의 적률(moment)이며, $[-1,1]$ 위의 X의 적률은 구하기 쉽다. 홀수 차 적률은 대칭성에 의해 0이 되고, 짝수 차 적률은 $1/(k+1)$이다. $Z = Y - X$에 대해서는 Z의 분포를 전혀 구할 필요가 없다. $(Y-X)^2$과 $(Y-X)^4$을 이항정리로 전개하고 독립성(independence)을 이용하면, 모든 기댓값이 이미 알고 있는 적률들의 곱으로 바뀐다.

풀이

X의 적률

X는 $[-1,1]$에서 확률밀도함수(density)가 $\frac{1}{2}$이므로, 음이 아닌 정수 k에 대하여

$$\mathrm{E}(X^k) = \frac{1}{2}\int_{-1}^{1} x^k \,\mathrm{d}x = \begin{cases} 0, & k\text{가 홀수 (피적분함수가 기함수)}, \\ \dfrac{1}{k+1}, & k\text{가 짝수}. \end{cases}$$

특히 $\mathrm{E}(X) = \mathrm{E}(X^3) = 0$, $\mathrm{E}(X^2) = \frac{1}{3}$, $\mathrm{E}(X^4) = \frac{1}{5}$이다. 따라서

$$\mathrm{E}(X^2) = \frac{1}{3}, \qquad \mathrm{Var}(X^2) = \mathrm{E}(X^4) - (\mathrm{E}(X^2))^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.$$

Y도 같은 분포를 따르므로 적률도 같다.

Z^2의 평균

전개하면 $Z^2 = Y^2 - 2XY + X^2$이다. X와 Y는 독립이므로 $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y) = 0$이고, 따라서

$$\mathrm{E}(Z^2) = \frac{1}{3} - 0 + \frac{1}{3} = \frac{2}{3}.$$

Z^2의 분산

$\mathrm{E}(Z^4)$가 필요하다. 이항정리(binomial theorem)에 의해

$$Z^4 = (Y - X)^4 = Y^4 - 4Y^3X + 6Y^2X^2 - 4YX^3 + X^4.$$

? **생각해 보기** Y^3X의 기댓값을 정말 $\mathrm{E}(Y^3)\,\mathrm{E}(X)$로 놓아도 되는가?

그렇다. X와 Y가 독립이므로 Y^3과 X처럼 이들의 함수끼리도 독립이고, 독립인 확률변수들의 곱의 기댓값은 기댓값들의 곱과 같다. 두 교차항은 각각 X 또는 Y의 홀수 거듭제곱을 포함하므로 기댓값이 0이다. 그러므로

$$\begin{aligned}\mathrm{E}(Z^4) &= \mathrm{E}(Y^4) + 6\,\mathrm{E}(Y^2)\,\mathrm{E}(X^2) + \mathrm{E}(X^4)\\ &= \frac{1}{5} + 6\cdot\frac{1}{3}\cdot\frac{1}{3} + \frac{1}{5} = \frac{2}{5} + \frac{2}{3} = \frac{16}{15}.\end{aligned}$$

따라서

$$\mathrm{Var}(Z^2) = \mathrm{E}(Z^4) - (\mathrm{E}(Z^2))^2 = \frac{16}{15} - \frac{4}{9} = \frac{48-20}{45} = \frac{28}{45}.$$

$\frac{28}{45} = 7 \times \frac{4}{45}$이므로 $\mathrm{Var}(Z^2) = 7\,\mathrm{Var}(X^2)$이고, 이로써 증명이 끝난다.

검산

- Z는 $[-2, 2]$에서 삼각형 모양의 확률밀도함수 $f(z) = \frac{1}{4}(2 - |z|)$를 가진다. 직접 계산하면 $\mathrm{E}(Z^4) = 2\int_0^2 \frac{1}{4}z^4(2-z)\,\mathrm{d}z = \frac{1}{2}\left(\frac{64}{5} - \frac{32}{3}\right) = \frac{16}{15}$이고, 전개로 얻은 결과와 일치한다.
- $\mathrm{E}(Z) = 0$이므로 $\mathrm{E}(Z^2) = \mathrm{Var}(Y - X) = \mathrm{Var}(Y) + \mathrm{Var}(X) = \frac{1}{3} + \frac{1}{3}$이다.
- (X, Y) 쌍 백만 개를 시뮬레이션한 결과 $\mathrm{Var}(X^2) \approx 0.0889$, $\mathrm{Var}(Z^2) \approx 0.623$을 얻었으며, 이는 $4/45 \approx 0.0889$, $28/45 \approx 0.622$와 잘 맞는다.

흔한 실수

- $\mathrm{Var}(Z^2) = \mathrm{Var}(Y^2) + \mathrm{Var}(X^2)$로 쓰는 것. 이는 Z^2 안의 $-2XY$ 항을 무시한 것인데, 이 항은 평균이 0이지만 $4\,\mathrm{Var}(XY) = \frac{4}{9} = \frac{20}{45}$만큼 기여한다. 이 실수를 하면 $\frac{28}{45}$ 대신 $\frac{8}{45}$를 얻게 된다.

- Y^2X^2 항의 이항계수 6을 빠뜨리는 것.
- $[-1,1]$에서 확률밀도함수를 $\frac{1}{2}$ 대신 1로 놓는 것. 이렇게 하면 모든 적률이 두 배가 된다.

STEP 너머

$\mathrm{E}(X^k)$를 X의 적률이라 하며, 독립인 확률변수들의 합을 다루는 깔끔한 방법은 적률생성함수(moment generating function) $M_X(t) = \mathrm{E}(\mathrm{e}^{tX})$를 이용하는 것이다. 이때 $M_{Y-X}(t) = M_Y(t)M_X(-t)$가 성립한다. 이 곱을 t에 대한 멱급수로 전개하면 Z의 모든 적률을 한꺼번에 얻는다. 적률생성함수와 특성함수(characteristic function)는 대학 첫 확률론 강의에서 핵심적인 위치를 차지한다. 이들은 중심극한정리(central limit theorem)의 증명에서 주된 도구이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)에서 이를 명료하게 전개한다.

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.