풀이

STEP II 2013, 문제 12

확률과 통계 • 푸아송 분포의 적률

> **STEP II 2013, 12번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/13-s2-q12

핵심 아이디어

X는 U의 홀수 값만, Y는 짝수 값만 취하므로, 모든 기댓값은 홀수 항만 또는 짝수 항만 더한 푸아송(Poisson) 합이 된다. $\mathrm{E}(U)$와 $\mathrm{E}(U(U-1))$을 구할 때 쓰는 요령은 여기서도 그대로 통한다. 즉 k나 $k(k-1)$을 $k!$과 약분하고 λ나 λ^2을 밖으로 꺼낸다. k가 하나 줄어들 때마다 홀짝이 바뀌며, 이 때문에 α와 β가 서로 자리를 바꾼다. 또한 $\alpha+\beta=\mathrm{e}^{\lambda}$이므로 인수 $\mathrm{e}^{-\lambda}$는 $1/(\alpha+\beta)$임에 주목하자.

풀이

이하에서 $k=0,1,2,\ldots$에 대하여 $\mathrm{P}(U=k)=\mathrm{e}^{-\lambda}\lambda^k/k!$이고, $\lambda>0$이다. 두 급수(series)를 항별로 더하면 다음을 얻는다.

$$\alpha+\beta=\sum_{k\geq 0}\frac{\lambda^k}{k!}=\mathrm{e}^{\lambda},\qquad \text{따라서}\qquad \mathrm{e}^{-\lambda}=\frac{1}{\alpha+\beta}.$$

(사실 $\alpha=\cosh\lambda$이고 $\beta=\sinh\lambda$이다.)

소문항 (i)

U가 홀수가 아니면 X는 0이므로, 홀수 k만 기여한다.

$$\mathrm{E}(X)=\sum_{k\ \text{홀수}}k\,\frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}=\mathrm{e}^{-\lambda}\lambda\sum_{k\ \text{홀수}}\frac{\lambda^{k-1}}{(k-1)!}=\mathrm{e}^{-\lambda}\lambda\left(1+\frac{\lambda^2}{2!}+\frac{\lambda^4}{4!}+\cdots\right),$$

여기서 $k-1$은 $0,2,4,\ldots$를 차례로 지나기 때문이다. 따라서

$$\mathrm{E}(X)=\frac{\lambda\alpha}{\alpha+\beta}.$$

마찬가지로 Y는 $k \geq 2$인 짝수 k를 취하고($k = 0$인 항은 어차피 0이다), 이때 $k-1$은 $1, 3, 5, \ldots$를 지난다.

$$\mathrm{E}(Y) = \mathrm{e}^{-\lambda}\lambda \sum_{k \text{ 짝수}, k \geq 2} \frac{\lambda^{k-1}}{(k-1)!} = \frac{\lambda\beta}{\alpha+\beta}.$$

검산해 보면 $\mathrm{E}(X) + \mathrm{E}(Y) = \lambda = \mathrm{E}(U)$이다. $X + Y = U$이므로 당연히 그래야 한다.

소문항 (ii)

? **생각해 보기** k^2이 $k!$과 깔끔하게 약분되지 않을 때 $k^2\lambda^k/k!$은 어떻게 더할까?

$k^2 = k(k-1) + k$로 쓴다. $k(k-1)$ 부분은 $k!$의 인수 두 개와 약분되고, k 부분은 인수 하나와 약분된다. 홀수 k에 대하여 $k(k-1)$의 합은 $k = 3$부터 시작하고($k = 1$에서는 0이 된다), $k-2$는 $1, 3, 5, \ldots$를 지난다.

$$\sum_{k \text{ 홀수}} k(k-1)\frac{\lambda^k}{k!} = \lambda^2 \sum_{k \text{ 홀수}, k \geq 3} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2\beta, \qquad \sum_{k \text{ 홀수}} k\,\frac{\lambda^k}{k!} = \lambda\alpha$$

(두 번째 합은 소문항 (i)에서 구한 것이다.) 그러므로

$$\mathrm{E}(X^2) = \mathrm{e}^{-\lambda}\left(\lambda\alpha + \lambda^2\beta\right) = \frac{\lambda\alpha + \lambda^2\beta}{\alpha+\beta},$$

이고, 여기서 $\mathrm{E}(X)^2$을 빼면

$$\mathrm{Var}(X) = \frac{\lambda\alpha + \lambda^2\beta}{\alpha+\beta} - \frac{\lambda^2\alpha^2}{(\alpha+\beta)^2},$$

가 되어 요구한 결과를 얻는다.

Y의 경우에는 홀짝이 뒤바뀐다. 짝수 k에 대하여 $k(k-1)$의 합은 $k = 2$부터 시작하고 $k-2$는 $0, 2, 4, \ldots$를 지나므로 $\lambda^2\alpha$가 된다. k의 합은 소문항 (i)에서처럼 $\lambda\beta$가 된다. 따라서

$$\mathrm{Var}(Y) = \frac{\lambda\beta + \lambda^2\alpha}{\alpha+\beta} - \frac{\lambda^2\beta^2}{(\alpha+\beta)^2}.$$

마지막 질문. $X + Y = U$이므로 $\mathrm{Var}(X+Y) = \mathrm{Var}(U) = \lambda$이다. 두 분산(variance)을 더하면

$$\begin{aligned}\mathrm{Var}(X) + \mathrm{Var}(Y) &= \frac{\lambda(\alpha+\beta) + \lambda^2(\alpha+\beta)}{\alpha+\beta} - \frac{\lambda^2(\alpha^2+\beta^2)}{(\alpha+\beta)^2} \\ &= \lambda + \lambda^2\,\frac{(\alpha+\beta)^2 - \alpha^2 - \beta^2}{(\alpha+\beta)^2} = \lambda + \frac{2\lambda^2\alpha\beta}{(\alpha+\beta)^2}.\end{aligned}$$

? **생각해 보기** 두 값이 다를 것이라고 예상해야 하는 이유는 무엇일까?

일반적으로 $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$이다. 여기서는 X, Y 중 많아야 하나만 0이 아니므로 항상 $XY = 0$이다. 따라서 $\mathrm{Cov}(X,Y) = -\mathrm{E}(X)\mathrm{E}(Y) = -\lambda^2\alpha\beta/(\alpha+\beta)^2$이고, 이는 위의 계산과 일치한다. 등호가 성립하려면 $\lambda^2\alpha\beta = 0$이어야 한다. 그런데 $\lambda > 0$이면 $\alpha \geq 1$이고 $\beta \geq \lambda > 0$이므로 $\lambda^2\alpha\beta > 0$이다. 그러므로 $\mathrm{Var}(X) + \mathrm{Var}(Y) = \mathrm{Var}(X+Y)$를 만족하는 **0이 아닌 λ의 값은 존재하지 않는다**. 좌변이 항상 엄격히 더 크다. 실제로 $\alpha = \cosh\lambda$, $\beta = \sinh\lambda$이므로 초과분은 $\lambda^2 \sinh 2\lambda / \mathrm{e}^{2\lambda} = \frac{1}{2}\lambda^2\left(1 - \mathrm{e}^{-4\lambda}\right)$이다.

검산

- $\lambda \to 0$이면 U는 거의 항상 0 또는 1이므로 Y는 거의 항상 0이다. 실제로 $\mathrm{E}(Y) \approx \lambda^2 \to 0$이고 $\mathrm{E}(X) \approx \lambda$이다.
- λ가 클 때는 $\alpha \approx \beta \approx \frac{1}{2}\mathrm{e}^{\lambda}$이므로 $\mathrm{E}(X) \approx \mathrm{E}(Y) \approx \frac{1}{2}\lambda$이다. 즉 홀수 값과 짝수 값이 나올 가능성이 같아진다.
- $\lambda = 1$일 때 푸아송 확률을 직접 더하면 $\mathrm{Var}(X) \approx 0.6778$, $\mathrm{Var}(Y) \approx 0.8131$이므로, 그 합은 $\mathrm{Var}(U) = 1$보다 약 $0.4908 = \frac{1}{2}(1 - \mathrm{e}^{-4})$만큼 크다. 두 공식 모두 들어맞는다.

흔한 실수

- 약분한 뒤 홀짝을 잘못 판단하는 것: k가 홀수이면 $k-1$은 짝수(α가 나옴)이고 $k-2$는 홀수(β가 나옴)이다.
- $\mathrm{e}^{-\lambda} = 1/(\alpha+\beta)$임을 잊어서 답을 요구한 꼴로 나타내지 못하는 것.
- 분산이 그냥 더해진다고 가정하는 것. X와 Y는 독립(independent)과는 거리가 멀다. $X \neq 0$임을 알면 반드시 $Y = 0$이다.

STEP 너머

급수를 짝수 부분과 홀수 부분으로 나누는 것은 $\mathrm{e}^{\lambda} = \cosh\lambda + \sinh\lambda$라는 분해에 해당한다. 확률생성함수(probability generating function)를 이용하면 이는 일상적인 계산이 된다. $G(s)$의 짝수 부분은 $\frac{1}{2}\big(G(s) + G(-s)\big)$이다. 마지막에 사용한 공분산(covariance) 항등식은 대학 확률론 강의에서 가장 먼저 배우는 도구 중 하나이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)에서 생성함수와 공분산을 모두 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.