풀이

STEP II 2017, 문제 12

확률과 통계 • 푸아송 과정

> **STEP II 2017, 12번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/17-s2-q12

핵심 아이디어

서로 독립(independent)인 두 물고기 포아송 흐름(Poisson stream)은 합쳐져 하나의 포아송 흐름이 되고, 그 비율은 두 비율의 합이다. (i)의 합성곱(convolution)이 보여 주는 것이 바로 이것이다. 시간에 관한 문항에서는 첫 물고기가 **언제** 잡히는지에 관한 명제를 물고기가 **몇 마리** 잡혔는지에 관한 명제로 바꾼다. "아담의 첫 물고기는 시각 t 이후에 잡힌다"는 정확히 "아담은 $[0,t]$에서 물고기를 한 마리도 잡지 못한다"는 뜻이며, 그 확률(probability)은 포아송 공식에서 곧바로 얻어진다.

풀이

이 풀이 전체에서 X와 Y는 길이가 T인 구간에서 아담과 이브가 잡은 물고기의 수이다. 따라서 $X \sim \mathrm{Po}(\lambda T)$이고 $Y \sim \mathrm{Po}(\mu T)$이며, 둘은 서로 독립이다.

소문항 (i)

정수 $r \geq 0$에 대하여, 사건 $X+Y=r$는 $j=0,1,\ldots,r$에 대한 서로소인 사건 $\{X=j,\ Y=r-j\}$들로 나뉜다. 독립성을 이용하면

$$\begin{aligned}\mathrm{P}(X+Y=r) &= \sum_{j=0}^{r} \mathrm{P}(X=j)\,\mathrm{P}(Y=r-j) = \sum_{j=0}^{r} \frac{e^{-\lambda T}(\lambda T)^j}{j!}\cdot\frac{e^{-\mu T}(\mu T)^{r-j}}{(r-j)!} \\ &= \frac{e^{-(\lambda+\mu)T}}{r!}\sum_{j=0}^{r}\binom{r}{j}(\lambda T)^j(\mu T)^{r-j} = \frac{e^{-(\lambda+\mu)T}((\lambda+\mu)T)^r}{r!},\end{aligned}$$

이다. 여기서 $r!$을 곱하고 나눈 다음 이항정리(binomial theorem)를 이용하였다. 이것은 $\mathrm{Po}((\lambda+\mu)T)$의 확률이므로, 시간 T 동안 잡은 물고기의 총수는 모수가 $(\lambda+\mu)T$인 포아송

분포를 따른다.

소문항 (ii)

$0 \le r \le k$에 대하여

$$\mathrm{P}(X = r \mid X + Y = k) = \frac{\mathrm{P}(X = r,\ Y = k - r)}{\mathrm{P}(X + Y = k)} = \frac{\dfrac{e^{-\lambda T}(\lambda T)^r}{r!} \cdot \dfrac{e^{-\mu T}(\mu T)^{k-r}}{(k-r)!}}{\dfrac{e^{-(\lambda+\mu)T}((\lambda+\mu)T)^k}{k!}},$$

이다. 분모에는 소문항 (i)을 이용하였다. 지수함수들이 약분되고, T의 거듭제곱도 약분되어(분자에 $r + (k-r) = k$개, 분모에 k개가 있다)

$$\mathrm{P}(X = r \mid X + Y = k) = \binom{k}{r} \left(\frac{\lambda}{\lambda+\mu}\right)^r \left(\frac{\mu}{\lambda+\mu}\right)^{k-r}.$$

이 남는다. 따라서 총수가 k로 주어졌을 때 아담이 잡은 물고기의 수는 $\mathrm{B}(k,\ \frac{\lambda}{\lambda+\mu})$를 따른다. 사실상 k마리의 물고기 각각이 독립적으로 확률 $\frac{\lambda}{\lambda+\mu}$로 아담의 것이 되는 셈이다. T가 사라졌다는 점에 주목하자.

소문항 (iii)

처음부터 잰 시각으로, 아담과 이브가 첫 물고기를 잡는 시각을 각각 A와 E라 하자. 아담의 첫 물고기가 시각 t 이후에 잡히는 것은 정확히 그가 $[0, t]$에서 아무것도 잡지 못할 때이므로

$$\mathrm{P}(A > t) = \mathrm{P}(\mathrm{Po}(\lambda t) = 0) = e^{-\lambda t}, \qquad \text{이고, 마찬가지로} \qquad \mathrm{P}(E > t) = e^{-\mu t}.$$

이다. 따라서 A의 확률밀도함수(probability density function)는 $t \ge 0$에서 $\lambda e^{-\lambda t}$이고($1 - e^{-\lambda t}$를 미분한다), 두 사람의 어획이 독립이므로 A와 E도 독립이다.

? 생각해 보기 A의 밀도와 E에 관한 확률을 어떻게 결합할까?

A의 값에 대하여 조건을 걸자. 아담의 첫 물고기가 시각 t에 잡힌다면, 그가 먼저 잡는 것은 정확히 $E > t$일 때이고, 독립성에 의해 그 확률은 $e^{-\mu t}$이다. 모든 t에 대하여 합하면

$$\mathrm{P}(A < E) = \int_0^\infty \lambda e^{-\lambda t} e^{-\mu t}\, \mathrm{d}t = \lambda \left[\frac{-e^{-(\lambda+\mu)t}}{\lambda+\mu}\right]_0^\infty = \frac{\lambda}{\lambda+\mu}.$$

이다. (E가 연속분포를 따르므로 동시에 잡을 확률은 0이다.) 따라서 첫 물고기가 아담의 것일 확률은 $\dfrac{\lambda}{\lambda+\mu}$이며, 이는 (ii)의 이항분포 모수와 같다.

소문항 (iv)

구하려는 것은 $\mathrm{E}(M)$이다. 여기서 $M = \max(A, E)$는 두 사람 모두 물고기를 한 마리씩 잡게 되는 시각이다. 둘 중 누구든 첫 물고기를 잡는 시각을 $m = \min(A, E)$라 쓰자. 임의의

두 수에 대하여 max + min은 두 수의 합이므로

$$M = A + E - m \qquad \Longrightarrow \qquad \mathrm{E}(M) = \mathrm{E}(A) + \mathrm{E}(E) - \mathrm{E}(m).$$

이다.

$\mathrm{E}(A)$와 $\mathrm{E}(E)$. 처음부터 본 아담의 어획은 어느 한 마리를 잡은 순간부터 본 것과 같은 종류의 과정이므로, 첫 물고기까지 기다리는 시간의 기댓값(expectation)은 주어진 간격의 기댓값과 같다. 즉 $\mathrm{E}(A) = \lambda^{-1}$이고, 마찬가지로 $\mathrm{E}(E) = \mu^{-1}$이다. (직접 계산하면, 부분적분(integration by parts)으로 $\mathrm{E}(A) = \int_0^\infty t\lambda e^{-\lambda t}\,\mathrm{d}t = \lambda^{-1}$이다.)

? 생각해 보기 "둘 중 누구든 잡는 첫 물고기"는 어떤 종류의 과정일까?

소문항 (i)에 의해, 길이가 t인 임의의 구간에서 두 사람이 합쳐 잡은 물고기의 수는 $\mathrm{Po}((\lambda+\mu)t)$를 따른다. 따라서 합친 어획은 아담의 것과 정확히 같은 종류의 과정이되 비율이 λ 대신 $\lambda+\mu$이며, 첫 물고기까지 기다리는 시간의 기댓값은 $\mathrm{E}(m) = (\lambda+\mu)^{-1}$이다. (직접 계산하면 $\mathrm{P}(m > t) = \mathrm{P}(A > t)\,\mathrm{P}(E > t) = e^{-(\lambda+\mu)t}$이다.) 그러므로

$$\mathrm{E}(M) = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda+\mu} = \frac{\lambda^2+\lambda\mu+\mu^2}{\lambda\mu(\lambda+\mu)}.$$

이다.

검산

- $\lambda = \mu$이면 소문항 (iii)은 대칭성(symmetry)이 요구하는 대로 $\frac{1}{2}$을 주고, 소문항 (iv)는 $\frac{3}{2\lambda}$를 준다. 같은 "새 출발" 논리에 따라 첫 물고기까지 $\frac{1}{2\lambda}$을 기다리고, 다른 사람을 위해 $\frac{1}{\lambda}$을 더 기다리는 것이다.
- $\mu \to \infty$이면(이브가 물고기를 즉시 잡으면) 소문항 (iv)의 값은 아담 혼자 필요한 시간인 $\frac{1}{\lambda}$로 수렴(tend to)한다.
- $\frac{1}{\mu} > \frac{1}{\lambda+\mu}$이므로 (iv)의 답은 당연히 $\max(\frac{1}{\lambda}, \frac{1}{\mu})$ 이상이다.
- $\lambda = 1.3$, $\mu = 0.7$일 때 공식은 0.65와 1.6978을 주며, 2×10^5번의 낚시를 컴퓨터로 시뮬레이션한 결과는 0.651과 1.700이었다.

흔한 실수

- (iv)에서 $\frac{1}{\lambda} + \frac{1}{\mu}$을 더하거나 $\max(\frac{1}{\lambda}, \frac{1}{\mu})$를 취하는 것. 최댓값의 기댓값은 이 둘 중 어느 것도 아니다.
- (ii)에서 $\mathrm{P}(X+Y=k)$로 나누는 것을 잊는 것. 결합확률을 조건부확률(conditional probability)로 바꾸어 주는 것이 바로 이 나눗셈이다.

STEP 너머

여기서 만난 것은 **포아송 과정**(Poisson process)이다. 사건 사이의 대기 시간은 지수분포(exponential distribution)를 따르고 무기억성(memoryless)을 가지며, 독립인 포아송 과정들은 중첩(superpose)되어 또 하나의 포아송 과정이 되고, 각 사건은 그 비율에 비례하는 확률로 독립적으로 표시된다(그 역연산인 "솎아내기(thinning)"도 성립한다). 이 사실들은 대기행렬 이론(queueing theory)과 연속시간 마르코프 연쇄(continuous-time Markov chain)의 토대이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 이 문제에서 자연스럽게 이어지는 방식으로 포아송 과정과 지수분포를 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.