풀이

STEP I 2007, 문제 14

확률과 통계 • 푸아송 모수의 유일성

> **STEP I 2007, 14번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/07-s1-q14

핵심 아이디어

각 정보는 $h(\lambda) =$ 상수 꼴의 방정식을 주며, 여기서 h는 평균에 대한 어떤 함수이다. 이 방정식이 λ를 유일하게 결정하는지는 전적으로 $\lambda > 0$에서 h의 **모양**에 달려 있다. 순단조(strictly monotonic)인 h는 각 값을 많아야 한 번 지나지만, 증가하다가 감소하는 h는 대부분의 값을 두 번 지나고 최댓값만 한 번 지난다. 따라서 이 문제는 결국 각 함수가 어디에서 단조이고 어디에서 최대가 되는지를 찾는 문제이다.

풀이

풀이 전체에서 $\lambda > 0$이고(푸아송(Poisson) 분포의 평균은 양수이다), $k = 0, 1, 2, \ldots$에 대하여 $\mathrm{P}(X = k) = \mathrm{e}^{-\lambda}\lambda^k/k!$이다.

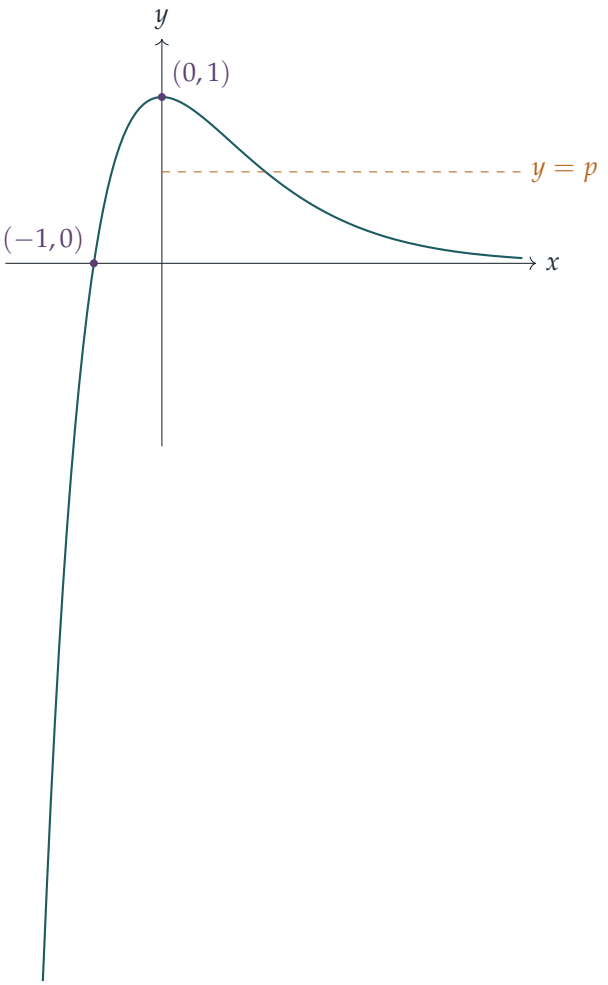
소문항 (i)

$f(x) = (x+1)\mathrm{e}^{-x}$라 하자. 곱의 미분법(product rule)에 의하여

$$f'(x) = \mathrm{e}^{-x} - (x+1)\mathrm{e}^{-x} = -x\,\mathrm{e}^{-x}.$$

$\mathrm{e}^{-x} > 0$이므로 $f'(x)$의 부호는 x의 부호와 반대이다. 즉 f는 $x < 0$에서 증가하고 $x > 0$에서 감소한다. 따라서 유일한 극점(turning point)은 $(0, 1)$에서의 극대이며, 이 점은 곡선이 y축과 만나는 점이기도 하다. 곡선은 $x + 1 = 0$인 곳, 즉 $(-1, 0)$에서 x축과 만난다.

양 끝의 행동을 보자. $x \to \infty$일 때 e^{-x}가 일차 인수를 압도하므로 $f(x) \to 0$이고, 위쪽에서 다가간다(x축이 점근선(asymptote)이다). $x \to -\infty$일 때는 두 인수의 크기가 모두 커지므로 $f(x) \to -\infty$이다.



이제 확률을 생각하자.

$$\mathrm{P}(X \geq 2) = 1 - \mathrm{P}(X = 0) - \mathrm{P}(X = 1) = 1 - \mathrm{e}^{-\lambda} - \lambda \mathrm{e}^{-\lambda} = 1 - f(\lambda),$$

이므로, 주어진 정보는 정확히

$$f(\lambda) = (1 + \lambda)\mathrm{e}^{-\lambda} = p.$$

라는 뜻이다. $\lambda > 0$에서 함수 f는 연속이고 순감소하며($f'(\lambda) = -\lambda \mathrm{e}^{-\lambda} < 0$이기 때문이다), $f(0) = 1$에서 시작하여 $\lambda \to \infty$일 때 0을 향해 감소한다. 따라서 $(0, 1)$의 모든 값을 정확히 한 번씩 취한다. 사잇값 정리(intermediate value theorem)에 의하여 해가 적어도 하나 존재하고, 순단조성 때문에 두 번째 해는 없다. 그림에서 $0 < p < 1$인 수평선 $y = p$는 오른쪽 가지와 정확히 한 번 만난다. 따라서 λ는 유일하게 결정된다.

소문항 (ii)

여기서는 $\mathrm{P}(X=1)=\lambda \mathrm{e}^{-\lambda}=q$이다. $g(\lambda)=\lambda \mathrm{e}^{-\lambda}$라 두면

$$g'(\lambda)=(1-\lambda)\mathrm{e}^{-\lambda},$$

이므로 g는 $(0,1]$에서 순증가하고 $[1,\infty)$에서 순감소하며, 최댓값은 $g(1)=\mathrm{e}^{-1}$이다. 또한 $\lambda \to 0^+$일 때와 $\lambda \to \infty$일 때 모두 $g(\lambda)\to 0$이다. (실제로 $g(\lambda)=\mathrm{e}^{-1}f(\lambda-1)$이므로, g의 그래프는 소문항 (i)의 곡선을 오른쪽으로 한 칸 평행이동하고 세로 방향으로 e^{-1}배 한 것이다.)

> **생각해 보기** 정점까지 올라갔다가 내려오는 곡선과 수평선은 몇 번 만날 수 있는가?

단조인 각 부분과는 많아야 한 번 만나므로 다음이 성립한다.

- $0<q<\mathrm{e}^{-1}$이면 $(0,1)$에 해가 하나, $(1,\infty)$에 또 하나 있으므로 가능한 λ의 값은 두 개이다.
- $q>\mathrm{e}^{-1}$이면 해가 없다.
- $q=\mathrm{e}^{-1}$이면 유일한 해는 정점 자체인 $\lambda=1$이다.

따라서 주어진 정보가 λ를 유일하게 결정하는 q의 값은 정확히 하나뿐이며, 그 값은 다음과 같다.

$$q=\mathrm{e}^{-1}, \qquad \lambda=1.$$

소문항 (iii)

조건부 확률(conditional probability)의 정의에 의하여, 그리고 사건 $\{X=1\}$이 $\{X\le 2\}$에 포함되므로

$$\mathrm{P}(X=1 \mid X\le 2)=\frac{\mathrm{P}(X=1)}{\mathrm{P}(X\le 2)}=\frac{\lambda \mathrm{e}^{-\lambda}}{\mathrm{e}^{-\lambda}\left(1+\lambda+\frac{1}{2}\lambda^2\right)}=\frac{2\lambda}{2+2\lambda+\lambda^2}.$$

이를 $k(\lambda)$라 하자. 분자와 분모를 λ로 나누면($\lambda>0$이므로 가능하다)

$$k(\lambda)=\frac{2}{\lambda+\dfrac{2}{\lambda}+2}.$$

> **생각해 보기** $\lambda>0$일 때 $\lambda+2/\lambda$의 최솟값은 무엇인가?

산술-기하 평균 부등식(AM–GM inequality)에 의하여 $\lambda+\dfrac{2}{\lambda}\ge 2\sqrt{\lambda\cdot\dfrac{2}{\lambda}}=2\sqrt{2}$이고, 등호는 정확히 $\lambda=2/\lambda$일 때, 즉 $\lambda=\sqrt{2}$일 때 성립한다. 더욱이 $\lambda+2/\lambda$는 $(0,\sqrt{2}]$에서 순감

소하고 $[\sqrt{2}, \infty)$에서 순증가하며(도함수 $1 - 2/\lambda^2$의 부호는 $\sqrt{2}$에서만 바뀐다), 양 끝에서 ∞로 발산한다. 따라서 k는 소문항 (ii)의 g와 같은 증가 후 감소 모양을 가지며, 최댓값은 다음과 같다.

$$k(\sqrt{2}) = \frac{2}{2\sqrt{2}+2} = \frac{1}{\sqrt{2}+1} = \sqrt{2} - 1.$$

앞에서와 똑같이, $0 < r < \sqrt{2} - 1$이면 λ의 값이 두 개이고, $r > \sqrt{2} - 1$이면 하나도 없으며, $r = \sqrt{2} - 1$이면 정확히 하나이다. 따라서 λ가 유일하게 결정되는 것은 다음 경우뿐이다.

$$r = \sqrt{2} - 1, \qquad \lambda = \sqrt{2}.$$

다른 방법으로, $k(\lambda) = r$의 분모를 없애면 이차방정식 $r\lambda^2 - 2(1-r)\lambda + 2r = 0$을 얻는다. 두 근의 곱은 $2 > 0$이고 합은 $2(1-r)/r > 0$이므로, 실근이 있다면 둘 다 양수이다. 근이 정확히 하나인 것은 판별식(discriminant) $4\left[(1-r)^2 - 2r^2\right] = 4(1 - 2r - r^2)$이 0이 될 때뿐이며, 이때 $r = \sqrt{2} - 1$(양의 근을 택한다)이고 중근은 $\lambda = (1-r)/r = \sqrt{2}$이다.

검산

- 소문항 (ii): $\lambda = 1$이면 $\mathrm{P}(X = 1) = \mathrm{e}^{-1} \approx 0.3679$이다. $0 < \lambda < 40$ 범위에서 $\lambda \mathrm{e}^{-\lambda}$를 수치적으로 훑어보면 이 값을 넘는 경우가 없고, $q = 0.3$이면 교점이 두 개이다.
- 소문항 (iii): $\lambda = \sqrt{2}$이면 $\dfrac{2\sqrt{2}}{2 + 2\sqrt{2} + 2} = \dfrac{\sqrt{2}}{2 + \sqrt{2}} = \sqrt{2} - 1 \approx 0.4142$이고, 수치 탐색으로 이것이 최댓값임을 확인할 수 있다. $r = 0.35$이면 λ의 값이 두 개이고, $r = 0.45$이면 하나도 없다.
- 타당성 점검: $\lambda \to 0$일 때는 확률 질량이 거의 모두 0에 몰리고, $\lambda \to \infty$일 때는 $X \le 2$라는 조건 아래의 질량이 2에 쌓이므로, 두 극한 모두에서 $\mathrm{P}(X = 1 \mid X \le 2) \to 0$이다. 이는 공식이 말해 주는 바와 같다.

흔한 실수

- $\lambda > 0$임을 잊는 것: 소문항 (i)에서 곡선은 $x < 0$에서도 존재하지만, 확률과 관련 있는 것은 $x > 0$인 가지뿐이다.
- 소문항 (i)에서 이유를 밝히지 않고 유일성을 주장하는 것: 해가 존재한다는 것(연속성, 그리고 값이 1에서 0까지 내려간다는 것)과 두 번째 해가 없다는 것(순단조성)을 모두 보여야 한다.
- 소문항 (iii)에서 $\mathrm{P}(X = 1 \text{ 이고} X \le 2) = \mathrm{P}(X = 1)\,\mathrm{P}(X \le 2)$를 사용하는 것: 두 사건은 독립이 아니며, $\{X = 1\}$이 $\{X \le 2\}$에 포함될 뿐이다.

STEP 너머

이 문제는 **식별가능성**(*identifiability*)을 처음 맛보게 해 준다. 통계학에서는 관측된 양이 모형의 모수를 결정하는지를 묻는데, 그 답은 모수에서 관측량으로 가는 사상이 일대일일 때 정확히 "그렇다"이다. 여기서 $\mathrm{P}(X \geq 2)$는 λ를 식별하지만, $\mathrm{P}(X = 1)$과 $\mathrm{P}(X = 1 \mid X \leq 2)$는 일반적으로 그렇지 않다. 같은 증가 후 감소 모양은 가능도(likelihood)를 최대화할 때 다시 나타난다. $\lambda \mathrm{e}^{-\lambda}$는 관측값 하나 $X = 1$을 본 뒤의 λ의 가능도이며, $\lambda = 1$에서의 정점이 바로 최대가능도 추정값(maximum likelihood estimate)이다. Blitzstein과 Hwang의 *Introduction to Probability*(확률론 입문)는 푸아송 분포와 조건부 확률에 대해 더 읽어 보기에 생동감 넘치는 책이다.

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.