풀이

STEP I 1992, 문제 15

확률과 통계 • 기대 대기 시간

> **STEP I 1992, 15번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/92-s1-q15

핵심 아이디어

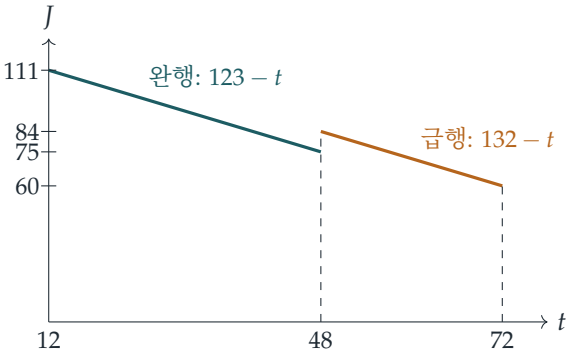
모든 것이 한 시간마다 반복되므로, 한 시간 동안 균등분포(uniform distribution)를 따르는 시각에 도착하는 승객 한 명만 살펴보면 충분하다. 그 한 시간은 두 개의 간격으로 나뉜다. 어떤 열차 직전의 간격에 도착한 승객은 모두 그 열차를 타고, 각 승객의 이동 시간은 "대기 시간 더하기 운행 시간"이다. 한 간격 안에서 대기 시간은 균등분포를 따르므로 그 평균은 간격의 절반이다. 이렇게 하면 문제 전체가 한 간격의 길이에 대한 이차식 문제로 바뀐다.

풀이

현재 시간표

시각 t를 정오로부터 지난 분 단위로 재고(어느 시각이든 상관없다), 승객의 도착 시각 T가 길이 한 시간인 구간 $(12, 72]$에서 균등분포를 따른다고 하자. 이 구간에 있는 열차는 48에 출발하여 런던에 $48 + 75 = 123$에 도착하는 완행열차와, 72(즉 다음 시의 12분)에 출발하여 $72 + 60 = 132$에 도착하는 급행열차이다.

- $12 < T \leq 48$이면 처음 탈 수 있는 열차는 완행이므로 이동 시간은 $J = 123 - T$이다.
- $48 < T \leq 72$이면 처음 탈 수 있는 열차는 급행이므로 $J = 132 - T$이다.



T는 $(12, 72]$에서 확률밀도(density) $\frac{1}{60}$을 가지므로

$$\mathrm{E}(J) = \frac{1}{60}\left[\int_{12}^{48}(123-t)\,\mathrm{d}t + \int_{48}^{72}(132-t)\,\mathrm{d}t\right].$$

각 피적분함수는 일차식이므로, 각 적분은 구간의 길이에 구간 중점에서의 함숫값을 곱한 것과 같다.

$$\int_{12}^{48}(123-t)\,\mathrm{d}t = 36\times(123-30) = 3348, \qquad \int_{48}^{72}(132-t)\,\mathrm{d}t = 24\times(132-60) = 1728.$$

따라서

$$\mathrm{E}(J) = \frac{3348+1728}{60} = \frac{5076}{60} = 84.6\text{분},$$

이고, 이것이 구하려던 결과이다.

급행열차 시각 조정하기

이제 급행열차가 매시 x분에 출발하고, 완행열차는 여전히 48분에 출발한다고 하자. 급행열차 출발부터 다음 완행열차 출발까지의 시간을 d분이라 하면 $0 \le d < 60$이고 $d \equiv 48 - x \pmod{60}$이다. (원래 시간표에서는 $d = 36$이다.) 그러면 한 시간은 다음과 같이 나뉜다.

- 각 완행열차 직전의 길이 d인 간격: 이 승객들은 $[0, d)$에서 균등분포를 따르는 시간 W만큼 기다린 뒤 75분 동안 이동한다.
- 각 급행열차 직전의 길이 $60-d$인 간격: 이 승객들은 $[0, 60-d)$에서 균등하게 기다린 뒤 60분 동안 이동한다.

승객이 첫 번째 간격에 도착할 확률은 $\frac{d}{60}$, 두 번째 간격에 도착할 확률은 $\frac{60-d}{60}$이고, 길이 g인 간격에서 균등분포를 따르는 대기 시간의 평균은 $\frac{g}{2}$이다. 따라서 승객이 어느 간격에 도착

하는지에 대해 조건부로 나누어(conditioning) 계산하면

$$\begin{aligned}60\,\mathrm{E}(J) &= d\left(\frac{d}{2}+75\right)+(60-d)\left(\frac{60-d}{2}+60\right)\\ &= \frac{d^2}{2}+75d+\frac{3600-120d+d^2}{2}+3600-60d\\ &= d^2-45d+5400.\end{aligned}$$

검산해 보면, $d=36$일 때 $1296-1620+5400=5076$이 되어 앞부분의 결과와 일치한다.

생각해 보기 왜 한 시간을 똑같이 나누는 $d=30$이 가장 좋은 간격이 아닐까?

완행열차 간격에 있는 승객은 모두 더 느린 운행 때문에 15분의 손해를 본다. 그러므로 그 간격에 있는 승객 수를 줄이고 싶고, 이는 그 간격을 30분보다 짧게 만드는 것을 뜻한다. 그러나 너무 짧게 하면 급행열차의 간격이(따라서 그 승객들의 대기 시간이) 길어진다. 이차식이 이 두 효과 사이의 균형을 잡아 준다.

완전제곱식(completing the square)으로 고치면

$$d^2-45d+5400=\left(d-\tfrac{45}{2}\right)^2+5400-\tfrac{2025}{4},$$

이고, 이는 $d=22.5$일 때 최소이며 이 값은 $[0,60)$ 안에 있다. 이때 $60\,\mathrm{E}(J)=\frac{19575}{4}$이므로 평균 이동 시간의 최솟값은

$$\mathrm{E}(J)=\frac{19575}{240}=\frac{1305}{16}=81\tfrac{9}{16}\approx 81.56\text{분}.$$

급행열차는 완행열차보다 22.5분 먼저 출발해야 하므로

$$x=48-22.5=25.5,$$

즉 급행열차는 매시 $25\frac{1}{2}$분에 출발해야 한다. 출발 시각이 정수 분이어야 한다면 $x=25$와 $x=26$($d=23$과 $d=22$)이 같은 값을 주며, 둘 다 $\frac{2447}{30}\approx 81.57$분이다.

검산

- $d=0$이면 두 열차가 함께 출발한다. 모두가 급행을 타고 평균 30분을 기다리므로 $\mathrm{E}(J)=90$이다. 공식도 $\frac{5400}{60}=90$을 준다.
- $d\to 60$이면 완행열차가 급행열차 직전에 출발하여 거의 모든 승객을 태운다. 평균 대기 시간 30분, 이동 75분, 합계 105분이다. 공식도 $\frac{3600-2700+5400}{60}=105$를 준다.
- 간격 0.1의 격자 위의 모든 x에 대해 한 시간 동안의 이동 시간을 수치 적분하면 $x=12$에서 84.6, $x=25.5$에서 최솟값 81.5625를 얻는다.

흔한 실수

- 매시 48분과 60분 사이에 도착한 승객을 한 시간을 더하지 않고 "12분" 급행열차에 태우는 실수. 이들이 타는 열차는 **다음** 시의 12분에 출발한다.
- 두 이동 시간을 간격의 길이가 아니라 열차의 수로 가중 평균하는 실수. 열차가 태우는 승객 수를 결정하는 것은 그 열차 직전 간격의 길이이다.

STEP 너머

여기서 계산한 양은 주기적인 시간표에서의 평균 대기 시간이며, 긴 간격이 더 많은 승객을 잡는다는 관찰은 **검사 역설**(inspection paradox)의 한 형태이다. 무작위로 도착하는 승객은 짧은 간격보다 긴 간격에 떨어질 가능성이 더 크다. 이는 무작위 도착과 서비스를 연구하는 재생 이론(renewal theory)과 대기행렬 이론(queueing theory)의 핵심 아이디어이다. Grimmett와 Welsh의 *Probability: An Introduction*(확률론 입문)은 그 바탕이 되는 조건부 확률과 기댓값(expectation)의 도구를 접하기에 좋은 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.