풀이

STEP I 1993, 문제 14

확률과 통계 • 지수분포를 따르는 대기 시간

> **STEP I 1993, 14번 문제**
>
> 더 읽어 나가기 전에 문제부터 읽어 보자. 이 문제는 STEP Support Programme의 문제 데이터베이스에 있다:
>
> step.maths.org/questions/93-s1-q14

핵심 아이디어

중요한 것은 토드 씨(Mr Toad)의 속력이 아니라 마지막 1킬로미터를 가는 데 걸리는 시간 T이며, $T = 1/V$이다. V의 밀도함수(density)는 다루기 까다로워 보이지만 누적분포함수(cumulative distribution function)는 매우 간단하다. 이를 T에 대한 명제로 바꾸면 T가 모수 α인 지수분포(exponential distribution)를 따름을 알 수 있다. 그다음부터는 각 초록불 구간의 확률이 두 지수함수의 차가 되고, 모든 초록불 구간의 확률을 더하면 등비급수(geometric series)가 된다.

풀이

도착 시간의 분포

V를 분당 킬로미터 단위의 속력이라 하고, $T = 1/V$를 신호등이 초록불로 바뀐 순간부터 재어 그가 신호등에 도착하기까지 걸리는 시간(분)이라 하자.

$$\frac{\mathrm{d}}{\mathrm{d}v}\,\mathrm{e}^{-\alpha/v} = \alpha v^{-2}\mathrm{e}^{-\alpha/v},$$

이고 $v \to 0^+$일 때 $\mathrm{e}^{-\alpha/v} \to 0$임을 알아차리면, V의 분포함수(distribution function)는

$$\mathrm{P}(V \le v) = \int_0^v \alpha u^{-2}\mathrm{e}^{-\alpha/u}\,\mathrm{d}u = \mathrm{e}^{-\alpha/v} \qquad (v > 0).$$

이다. ($v \to \infty$일 때 이 값은 1로 수렴하므로, 이것은 제대로 된 밀도함수이다.) $V > 0$이므로 $t > 0$에 대하여 $T > t \iff 1/V > t \iff V < 1/t$이고, 따라서

$$\mathrm{P}(T > t) = \mathrm{P}\left(V < \tfrac{1}{t}\right) = \mathrm{e}^{-\alpha t}.$$

이다. 그러므로 T는 모수 α인 지수분포를 따른다.

? 생각해 보기 왜 V 대신 T를 다루는가?

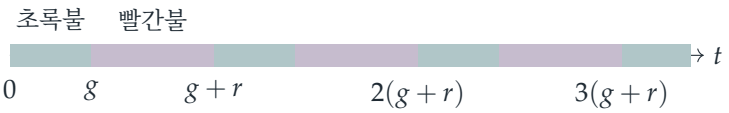
신호등은 시간으로 기술되므로, 우리가 관심 있는 모든 사건은 T의 값의 구간이다. V의 밀도함수를 직접 T의 밀도함수로 옮기려면 변수변환(change of variable)이 필요하지만, $\mathrm{P}(T > t)$를 거치면 그럴 필요 없이 한 줄로 답이 나온다.

한 번의 초록불 구간

시각 0에 신호등이 초록불로 바뀐다. 신호등은 $[0, g)$에서 초록불, $[g, g+r)$에서 빨간불, $[g+r, (g+r)+g)$에서 초록불이며, 이와 같이 계속된다. $(n+1)$번째 초록불 구간은 $n(g+r) \le t < n(g+r)+g$이다. T는 연속확률변수이므로 끝점의 포함 여부는 상관이 없고,

$$\begin{aligned}\mathrm{P}(n(g+r) < T < n(g+r)+g) &= \mathrm{P}(T > n(g+r)) - \mathrm{P}(T > n(g+r)+g) \\ &= \mathrm{e}^{-\alpha n(g+r)} - \mathrm{e}^{-\alpha(n(g+r)+g)},\end{aligned}$$

이 되어 요구된 결과를 얻는다. (인쇄된 시험지에는 두 번째 지수의 닫는 괄호가 잘못된 위치에 있다. 지수는 여기서처럼 $-\alpha(n(g+r)+g)$이어야 한다.) 이 계산은 $n = 0$일 때, 즉 신호등이 한 번도 빨간불로 바뀌기 전인 첫 번째 초록불 구간에서도 똑같이 성립한다.



초록불일 확률

그가 초록불에 통과하는 것은 T가 이 초록불 구간들 중 하나에 속할 때이고, 그때뿐이다. 이 구간들은 서로소이므로 $n = 0$을 포함하여 각 확률을 더하면 된다.

$$\begin{aligned}\mathrm{P}(\alpha) &= \sum_{n=0}^{\infty} \left(\mathrm{e}^{-\alpha n(g+r)} - \mathrm{e}^{-\alpha n(g+r)}\mathrm{e}^{-\alpha g}\right) = (1 - \mathrm{e}^{-\alpha g}) \sum_{n=0}^{\infty} (\mathrm{e}^{-\alpha(g+r)})^n \\ &= \frac{1 - \mathrm{e}^{-\alpha g}}{1 - \mathrm{e}^{-\alpha(g+r)}},\end{aligned}$$

여기서 등비급수의 합을 구했는데, ($\alpha > 0$이고 $g + r > 0$이므로) $0 < \mathrm{e}^{-\alpha(g+r)} < 1$이어서 이 급수는 수렴한다.

극한

$\alpha \to \infty$일 때 ($g > 0, g + r > 0$이므로) $\mathrm{e}^{-\alpha g} \to 0$이고 $\mathrm{e}^{-\alpha(g+r)} \to 0$이다. 따라서

$$\mathrm{P}(\alpha) \to \frac{1-0}{1-0} = 1.$$

이다. 이는 타당하다. α가 크다는 것은 토드 씨가 매우 빠르다는 뜻이고, 그는 거의 항상 첫 번째 초록불 구간이 끝나기 전에 통과한다.

$\alpha \to 0$일 때는 분자와 분모가 모두 0으로 수렴하므로 더 신중해야 한다. 각각에서 $1-\mathrm{e}^{-x} = \mathrm{e}^{-x}(\mathrm{e}^{x}-1)$로 쓰면

$$\mathrm{P}(\alpha) = \frac{\mathrm{e}^{-\alpha g}\left(\mathrm{e}^{\alpha g}-1\right)}{\mathrm{e}^{-\alpha(g+r)}\left(\mathrm{e}^{\alpha(g+r)}-1\right)} = \mathrm{e}^{\alpha r}\cdot\frac{\dfrac{\mathrm{e}^{\alpha g}-1}{\alpha g}}{\dfrac{\mathrm{e}^{\alpha(g+r)}-1}{\alpha(g+r)}}\cdot\frac{g}{g+r},$$

이다. 여기서 분자와 분모를 α로 나누고 인수 g와 $g+r$를 끼워 넣었다. $\alpha \to 0$일 때 $\mathrm{e}^{\alpha r}\to 1$이고, $\alpha g \to 0$이며 $\alpha(g+r)\to 0$이므로, $(\mathrm{e}^{x}-1)/x \to 1$에 의해 두 분수는 각각 1로 수렴한다. 그러므로

$$\mathrm{P}(\alpha) \to \frac{g}{r+g} \qquad \text{단, } \alpha \to 0.$$

이다. 이 경우 토드 씨는 매우 느리고 그의 도착 시간은 아주 많은 주기에 걸쳐 퍼져 있으므로, 초록불일 확률은 단순히 각 주기에서 초록불이 차지하는 비율이 된다.

검산

- $\alpha = 0.7, g = 1, r = 2$일 때 공식은 $\mathrm{P}(\alpha) \approx 0.5737$을 준다. $\mathrm{e}^{-\alpha/v}$의 역함수를 이용해 V를 추출하여 4×10^{5}번의 주행을 모의실험(simulation)한 결과는 0.5744였으며, 개별 초록불 구간 $n = 0, 1, 2$의 확률도 일치했다.
- $g = 1, r = 2$일 때 $\mathrm{P}(10^{-4}) \approx 0.33337$로 $\frac{1}{3}$에 가깝고, $\mathrm{P}(10) \approx 0.99995$이다.
- $r = 0$이면 신호등은 항상 초록불이고, 공식은 마땅히 그래야 하듯 모든 α에 대해 $\mathrm{P}(\alpha) = 1$을 준다.

흔한 실수

- $n = 0$을 빠뜨리는 것. 빨간불이 켜지기 전인 첫 번째 초록불 구간의 확률이 가장 크다.
- 인쇄된 지수 $-\alpha(n(g+r))+g$를 그대로 베끼는 것. 이는 $-\alpha\big(n(g+r)+g\big)$의 오식이다.
- V 자체를 도착 시간으로 다루거나, $\mathrm{E}(T)$를 구하려 하는 것. 사실 여기서 $\mathrm{E}(V)$는 무한대이지만, 문제의 어디에서도 그것이 필요하지 않다.

STEP 너머

분포함수를 통해 V의 분포를 $T = g(V)$의 분포로 바꾸는 것은 확률변수(random variable)를 변환하는 표준적인 방법이다. 여기서 V는 역지수분포(inverse-exponential distribution)를 따른다. 극한 $\alpha \to 0$은 일반적인 아이디어의 작은 예이다. 즉, 많은 주기에 걸쳐 퍼져 있는 무작위 시각은 각 단계에 그 단계의 길이에 비례하는 확률로 떨어지며, 이것이 재생 이론(renewal theory)의 출발점이다. Grimmett과 Welsh의 *Probability: An Introduction*(확률론 입문)은 확률변수의 변환과 지수분포를 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.