详细解答

STEP I 1992，第 15 题

概率与统计 • 期望等候时间

STEP I 1992，第 15 题

继续阅读之前，请先看题目（原题为英文）。题目载于 STEP Support Programme 的试题资料库：

step.maths.org/questions/92-s1-q15

关键思路

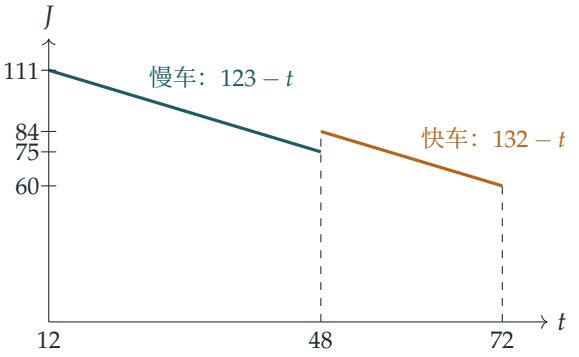
一切每小时重复一次，所以只需考虑一名乘客，其到达时间在一小时内均匀分布（uniformly distributed）。这一小时分成两段间隔：在某班列车之前那段间隔内到达的乘客，全部都会乘搭该班列车，而每人的行程时间就是“等候时间加行车时间”。在同一段间隔内，等候时间是均匀分布的，所以其平均值是间隔长度的一半。这样，整条题目就化为一个以其中一段间隔的长度为变量的二次式（quadratic）。

解答

现行时间表

以正午之后的分钟数量度时间 t（任何一个小时都可以），并设乘客的到达时间 T 在长一小时的区间 $(12, 72]$ 上均匀分布。这段时间内的列车有：在 48 开出的慢车，于 $48+75=123$ 抵达伦敦；以及在 72（即下一个小时的 12 分）开出的快车，于 $72+60=132$ 抵达。

- 若 $12 < T \leq 48$，最先可乘搭的是慢车，所以行程时间为 $J = 123 - T$。
- 若 $48 < T \leq 72$，最先可乘搭的是快车，所以 $J = 132 - T$。



由于 T 在 $(12,72]$ 上的概率密度（density）为 $\frac{1}{60}$，

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

两个被积函数（integrand）都是线性的，所以每个积分都等于区间长度乘以函数在区间中点的值：

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

因此

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

正如所求。

重新编排快车的开出时间

现在设快车在每小时的 x 分开出，慢车则仍在 48 分开出。设 d 为由一班快车开出至下一班慢车开出所相隔的分钟数，所以 $0\le d<60$ 且 $d\equiv 48-x\pmod{60}$。（在原来的时间表中，$d=36$。）于是这一小时分成

- 每班慢车之前一段长度为 d 的间隔：这些乘客的等候时间 W 在 $[0,d)$ 上均匀分布，然后行车 75 分钟；
- 每班快车之前一段长度为 $60-d$ 的间隔：这些乘客的等候时间在 $[0,60-d)$ 上均匀分布，然后行车 60 分钟。

乘客落在第一段间隔的概率是 $\frac{d}{60}$，落在第二段的概率是 $\frac{60-d}{60}$，而在长度为 g 的间隔上均匀分布的等候时间，其平均值为 $\frac{g}{2}$。所以，按乘客在哪一段间隔到达来取条件（conditioning），

可得

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

作为验算，$d=36$ 给出 $1296-1620+5400=5076$，与第一部分吻合。

想一想 为甚么最佳的间隔不是 $d=30$，即把一小时平均分成两半？

在慢车那段间隔内到达的每名乘客，都要因为行车较慢而多付 15 分钟的代价。所以我们希望那段间隔内的乘客少一些，也就是要把它缩短至少于半小时，但又不能短得令快车那段间隔（以及那些乘客的等候时间）变得太长。这个二次式正好平衡了这两种效应。

配方（completing the square）后，

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

它在 $d=22.5$ 时取得最小值，而这个值在 $[0,60)$ 之内。此时 $60\,\mathrm{E}(J)=\frac{19575}{4}$，所以平均行程时间的最小值为

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

快车必须比慢车早 22.5 分钟开出，所以

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

即快车应在每小时的 $25\frac{1}{2}$ 分开出。若开出时间必须是整数分钟，则 $x=25$ 与 $x=26$（$d=23$ 与 $d=22$）不分高下，两者都给出 $\frac{2447}{30}\approx 81.57$ 分钟。

验算

- 当 $d=0$ 时，两班列车同时开出；人人都乘搭快车，平均等候 30 分钟，所以 $\mathrm{E}(J)=90$。公式给出 $\frac{5400}{60}=90$。
- 当 $d\to 60$ 时，慢车刚好在快车之前开出，几乎接载了所有乘客：平均等候 30 分钟，行车 75 分钟，合共 105 分钟。公式给出 $\frac{3600-2700+5400}{60}=105$。
- 以步长 0.1 的网格取遍每个 x，对一小时内的行程时间作数值积分（numerical integration），所得结果是：$x=12$ 时为 84.6，而最小值 81.5625 出现在 $x=25.5$。

常见错误

- 把在每小时 48 分至 60 分之间到达的乘客算作乘搭“12 分”开出的快车，却没有加上一小时：他们的列车是在下一个小时的 12 分开出的。
- 以列车的班数而非间隔的长度作为权重，去平均两个行车时间；决定一班列车接载多少乘客的，是该班列车之前那段间隔的长度。

STEP 之外

这里计算的量，是周期性时间表下的平均等候时间；而“较长的间隔会接载较多乘客”这个观察，正是检查悖论（inspection paradox）的一种形式：随机到达的乘客，落在长间隔的机会比落在短间隔的大。这是更新理论（renewal theory）和排队论（queueing theory）的核心概念，这两门理论研究的是随机的到达与服务。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门）是认识取条件与期望值（expectation）这些基础工具的好起点。

Beyond STEP 详细解答：https://beyondstep.keithmcnulty.org/zh-hans/。解答文字采用 CC0 1.0；题目本身版权属 Cambridge University Press & Assessment 所有（©），本文件不予转载。由机器从英文翻译而成；尚未经审阅。