详细解答

STEP I 1990，第 16 题

概率与统计 · 三角分布与假设检验

STEP I 1990，第 16 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/90-s1-q16

关键思路

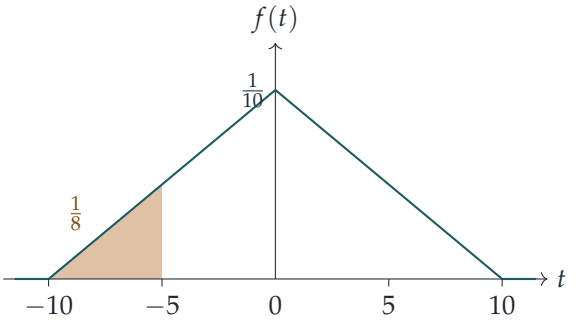
概率密度函数（probability density function）的图像是一个三角形，所以我们需要的每一个概率都是某个小三角形的面积，而累积分布函数（cumulative distribution function）会替我们完成所有工作。第二点要留意的是，只有 8 时的那班巴士才有可能在上午 8 时前后到站：7 时的那班巴士在 7.10 之前已经离开，而 9 时的那班巴士不可能在 8.50 之前到达。所以早上“赶上巴士”的真正意思是“比 8 时的那班巴士更早到站”。最后一部分于是变成一个只有一个参数的二项模型（binomial model）和一个显著性检定（significance test）。

解答

T 的密度函数

密度由 $t=-10$ 处的 0 线性上升至 $t=0$ 处的峰值，然后以相同的速率下降至 $t=10$ 处的 0，所以它在 $[-10,10]$ 上的形式是 $f(t)=k(10-|t|)$，在其他地方则为零。总面积等于一个底为 20、高为 $10k$ 的三角形的面积，所以 $\frac{1}{2}\cdot 20\cdot 10k=1$，由此得 $k=\frac{1}{100}$：

$$f(t)=\begin{cases}\dfrac{10+t}{100}, & -10\le t\le 0,\\ \dfrac{10-t}{100}, & 0\le t\le 10,\\ 0, & \text{其他情况。}\end{cases}$$



为方便稍后使用，先求左半部分的累积分布函数：它是一个底为 $a+10$、高为 $f(a)$ 的三角形的面积，

$$\mathrm{P}(T<a)=\tfrac{1}{2}(a+10)\cdot\frac{a+10}{100}=\frac{(a+10)^2}{200},\qquad -10\le a\le 0,$$

而根据关于 $t=0$ 的对称性，当 $0\le a\le 10$ 时，$\mathrm{P}(T>a)=\dfrac{(10-a)^2}{200}$。图中阴影部分的面积是 $\mathrm{P}(T<-5)=\frac{25}{200}=\frac{1}{8}$。

在上午 8.10 或之前赶上巴士

我们假设我会在车站一直等到有巴士来为止。预定 8 时到站的巴士在上午 8 时之后 T 分钟到达，其中 $-10\le T\le 10$；上一班巴士在 7.10 之前已经离开，而下一班不可能在 8.50 之前到达。所以，我能在 8.10 或之前赶上巴士，当且仅当我比 8 时的那班巴士更早到达车站，而这时 8.10 的期限自然会得到满足。

如果我在上午 8 时之后 a 分钟到达车站，那么当 $T>a$ 时我便能赶上这班车（恰好出现 $T=a$ 的机会为零）。三种情况分别给出

$$\begin{aligned}
7{\cdot}55,\ a=-5:&\quad \mathrm{P}(T>-5)=1-\tfrac{25}{200}=\tfrac{7}{8},\\
7{\cdot}58,\ a=-2:&\quad \mathrm{P}(T>-2)=1-\tfrac{64}{200}=\tfrac{17}{25},\\
8{\cdot}02,\ a=2:&\quad \mathrm{P}(T>2)=\tfrac{64}{200}=\tfrac{8}{25}.
\end{aligned}$$

我在 7·55 到达车站的概率是 0.5，在 7.58 到达的概率是 $0.5\times 0.8=0.4$，在 8.02 到达的概率是 $0.5\times 0.2=0.1$。我睡过头与否和巴士毫无关系，所以根据全概率公式（law of total probability），

$$\begin{aligned}
\mathrm{P}(\text{在 8.10 或之前赶上巴士})&=0.5\cdot\tfrac{7}{8}+0.4\cdot\tfrac{17}{25}+0.1\cdot\tfrac{8}{25}\\
&=0.4375+0.272+0.032=\frac{1483}{2000}=0.7415.
\end{aligned}$$

新时钟是否不准?

首先假设时钟是准确的，那么我每天早上都在 7.55 到达车站。设 T_8 和 T_9 分别为 8 时和 9 时那两班巴士的到站时间，以整点之后的分钟数计算；它们互相独立（independent），各自的密度都是 f。我在上午 9 时之后仍在等车，当且仅当我错过了 8 时的巴士（$T_8 < -5$）而且 9 时的巴士迟到（$T_9 > 0$）。所以在任何一个早上，

$$p = \mathrm{P}(T_8 < -5)\,\mathrm{P}(T_9 > 0) = \tfrac{1}{8} \cdot \tfrac{1}{2} = \tfrac{1}{16}.$$

? **想一想** 检定应该朝哪个方向进行：5 次是“太多”还是“太少”?

时钟调校错误，可以令 p 向任何一个方向偏移。假设时钟快了 d 分钟。那么我到达车站的真实时间是 $7.55 - d$，而时钟显示 9.00 时的真实时间是 $9.00 - d$，所以等到“上午 9 时”之后的概率变成

$$g(d) = \mathrm{P}(T_8 < -5 - d)\,\mathrm{P}(T_9 > -d).$$

时钟快了（$d > 0$）会令我更早到站，所以我很少会错过 8 时的巴士，于是 $g(d) < \frac{1}{16}$；时钟稍为慢了则会产生相反的效果（例如 $g(-2.5) \approx 0.079$）。所以，合理而不偏颇的做法是采用双尾检定（two-sided test）。

设 X 为这 100 个早上之中，按新时钟计算我等到上午 9 时之后的早上数目。不同巴士的到站时间互相独立，而我每天都在同一时间到站，所以各个早上是独立试验，并且

$$H_0：\text{时钟准确，}\quad X \sim \mathrm{B}(100, \tfrac{1}{16});\qquad H_1：\text{时钟调校错误，因此}p \neq \tfrac{1}{16}.$$

在 H_0 之下，我们预期有 $\mathrm{E}(X) = 6.25$ 次，而 $\mathrm{Var}(X) = 100 \cdot \frac{1}{16} \cdot \frac{15}{16} \approx 5.86$，所以标准差（standard deviation）约为 2.42。观察所得的 5 只是略低于平均值。准确地计算，

$$\mathrm{P}(X \leq 5) = \sum_{k=0}^{5} \binom{100}{k} p^k (1-p)^{100-k} \approx 0.40 \qquad \text{其中} p = \tfrac{1}{16},$$

而采用连续性校正（continuity correction）的正态近似（normal approximation）则给出 $\Phi((5.5 - 6.25)/2.42) = \Phi(-0.31) \approx 0.38$。这与 0.025（5% 双尾检定的下尾）相去甚远，事实上与任何常用的显著性水平都相去甚远。所以我们不拒绝 H_0：**没有证据显示新时钟调校错误**。

这并不等于有证据显示时钟是准确的。事实上 $g(d) = g(-5 - d)$，因为密度是对称的（把两班巴士的角色对调即可），所以一个慢了 5 分钟的时钟所给出的 $p = \frac{1}{16}$，与准确的时钟完全相同。这类数据无论收集多少个早上，都无法把这两种情况区分开来。

验算

▶ 密度的面积为 1，而三角形公式与对 f 作数值积分的结果吻合，例如 $\mathrm{P}(T < -2) = 0.32$。

- 答案 0.7415 介乎最好的情况 $\frac{7}{8}$（每次都在 7.55 到达）与最坏的情况 0.32（每次都在 8.02 到达）之间，这是理所当然的。模拟 10^6 个早上，所得结果为 0.741。
- 模拟 10^6 对 (T_8, T_9)，所得结果为 0.0622，对应 $\frac{1}{16} = 0.0625$；而准确的二项尾部概率 $\mathrm{P}(X \leq 5)$ 的数值为 0.4002。

常见错误

- 忘记了要等到 9 时之后，必须先错过 8 时的巴士：答案 $p = \frac{1}{2}$（只考虑“9 时的巴士迟到”）忽略了第一个因子。
- 以为“在 8.10 或之前赶上巴士”需要另外检查 $T \leq 10$：8 时的巴士必定在 8.10 或之前到达，所以只有 $T > a$ 才是关键。
- 数据显示次数略为偏少，却针对“太多”进行单尾检定（one-tailed test）；时钟不准可以令 p 向任何一个方向移动。

STEP 之外

三角形密度正是两个独立均匀随机变量（uniform random variables）之和的分布，这是卷积（convolution）的第一个例子，而大学一年级的概率课程正是用卷积来求和的密度。最后一部分让我们浅尝统计学上的可识别性（identifiability）：对称性 $g(d) = g(-5 - d)$ 意味着两个不同的参数值会产生完全相同的数据，所以没有任何检定能够把它们区分开来。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门）对卷积和随机变量之和有清晰的讲解。

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。