详细解答

STEP II 2009，第 13 题

概率与统计 · 二项分布的期望成本

STEP II 2009，第 13 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/09-s2-q13

关键思路

发生故障的引擎数目服从二项分布 (binomial distribution)，所以维修费用是一个二项随机变量(random variable)的简单函数，而其期望值(expectation)只是一个简短的和。要得出简洁的形式 $(*)$，窍门是不要把“全部更换”的情况逐一相加：改用补集 (complement)，然后利用因式分解 $1-q^n = p(1+q+\cdots+q^{n-1})$ 抽出因子 p。

解答

我们按照题目的原意，假设各引擎是否发生故障互相独立（independent），每个引擎发生故障的概率均为 p，并记 $q = 1-p$。所有费用一律以 K 为单位计算，直到最后为止。

Andover 型

设 X 为 Andover 型引擎中发生故障的数目，则 $X \sim \mathrm{B}(4, p)$；又设 C_A 为维修费用。于是

$$C_A = \begin{cases} 0 & X = 0, \\ K & X = 1, \\ 4K & X \geq 2. \end{cases}$$

现在 $\mathrm{P}(X=0) = q^4$ 及 $\mathrm{P}(X=1) = 4pq^3$，所以 $\mathrm{P}(X \geq 2) = 1 - q^4 - 4pq^3$。因此

$$\begin{aligned} \mathrm{E}(C_A) &= K \cdot 4pq^3 + 4K(1 - q^4 - 4pq^3) \\ &= 4K(1-q^4) - 12Kpq^3. \end{aligned}$$

由于 $1-q^4=(1-q)(1+q+q^2+q^3)=p(1+q+q^2+q^3)$，

$$\begin{aligned}\mathrm{E}(C_A)&=4Kp(1+q+q^2+q^3)-4Kp\cdot 3q^3\\&=4Kp(1+q+q^2-2q^3),\end{aligned}$$

这正是所要求的 $(*)$。

Basingstoke 型

设 $Y\sim \mathrm{B}(6,p)$ 为 Basingstoke 型引擎中发生故障的数目，C_B 为维修费用。则当 $Y\le 2$ 时 $C_B=YK$，而当 $Y\ge 3$ 时 $C_B=6K$。由

$$\mathrm{P}(Y=0)=q^6,\qquad \mathrm{P}(Y=1)=6pq^5,\qquad \mathrm{P}(Y=2)=15p^2q^4,$$

可得

$$\begin{aligned}\mathrm{E}(C_B)&=K\cdot 6pq^5+2K\cdot 15p^2q^4+6K(1-q^6-6pq^5-15p^2q^4)\\&=6K(1-q^6)-30Kpq^5-60Kp^2q^4.\end{aligned}$$

利用 $1-q^6=p(1+q+q^2+q^3+q^4+q^5)$ 并抽出 $6Kp$，

$$\mathrm{E}(C_B)=6Kp(1+q+q^2+q^3+q^4+q^5-5q^5-10pq^4).$$

> ? **想一想** 括号内仍然有一个 p。怎样才能把它化成“与 $(*)$ 相似的形式”?

形式 $(*)$ 是因子 p 乘以一个只含 q 的多项式（polynomial），所以我们把余下的 p 换成 $1-q$：$10pq^4=10q^4-10q^5$。合并各项后，

$$\boxed{\mathrm{E}(C_B)=6Kp(1+q+q^2+q^3-9q^4+6q^5).}$$

Andover 型的费用何时是 Basingstoke 型费用的 $\frac{2}{3}$?

我们需要 $\mathrm{E}(C_A)=\frac{2}{3}\mathrm{E}(C_B)$，即

$$4Kp(1+q+q^2-2q^3)=4Kp(1+q+q^2+q^3-9q^4+6q^5).$$

两边都有因子 $4Kp$，而 $K>0$，所以此式成立当且仅当 $p=0$ 或

$$-2q^3=q^3-9q^4+6q^5\iff 3q^3-9q^4+6q^5=0\iff 3q^3(1-q)(1-2q)=0.$$

所以 $q=0$、$q=1$ 或 $q=\frac{1}{2}$，即 $p=1$、$p=0$ 或 $p=\frac{1}{2}$。三个值都在 $[0,1]$ 之内，故所求的值为

$$p=0,\qquad p=\tfrac{1}{2},\qquad p=1.$$

$p=0$ 这个值最乏味：引擎永不发生故障，两个期望费用都是零。当 $p=1$ 时，每个引擎都发生故障，费用分别为 $4K$ 和 $6K$。真正有趣的值是 $p=\frac{1}{2}$。

验算

当 $p = \frac{1}{2}$ 时，$(*)$ 给出 $\mathrm{E}(C_A) = 2K(1 + \frac{1}{2} + \frac{1}{4} - \frac{1}{4}) = 3K$，而 Basingstoke 型的公式给出 $\mathrm{E}(C_B) = 3K(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} - \frac{9}{16} + \frac{6}{32}) = \frac{9}{2}K$，果然 $3 = \frac{2}{3} \cdot \frac{9}{2}$。我们亦取了 200 个有理数 p 值，把两条公式与直接按二项分布逐一列举所得的结果作精确比对，并以模拟方式验证：当 $p = 0.3$ 时，2×10^5 次模拟发射所得的平均费用为 $1.811K$ 和 $2.483K$，而公式所得的则是 $1.805K$ 和 $2.485K$。

常见错误

- 当有两个或以上 Andover 型引擎发生故障时，把费用算作 XK 而不是 $4K$：一旦超过门槛，*所有*引擎都要更换，包括完好的那些。
- 在 Basingstoke 型的括号内留下一个 p，以致答案不符合题目要求的形式。
- 在最后的方程中除以 p 而漏掉 $p = 0$，或除以 q^3 而漏掉 $p = 1$。

STEP 之外

这是在门槛策略（threshold policy）下计算期望费用的一个小例子，而这类计算正是可靠性理论（reliability theory）和运筹学（operational research）的家常便饭：管理者要选定门槛，在更换完好零件的费用与更频密维修的费用之间取得平衡。我们所用的尾和（tail sum）处理手法，即把 $\mathrm{P}(X \geq k)$ 写成一减去首几项，与得出非负整数值变量的公式 $\mathrm{E}(X) = \sum_{k \geq 1} \mathrm{P}(X \geq k)$ 背后的想法如出一辙。Grimmett 与 Welsh 合著的 *Probability: An Introduction*（概率论入门教材）对这些想法有严谨细致的阐述。

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。