详细解答

STEP II 2005，第 13 题

概率与统计 · 泊松分布与二项分布

STEP II 2005，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/05-s2-q13

关键思路

每一页抽样页要么错误少于两个，要么不是，而且各页互相独立，所以 Y 数算的是 n 次独立试验（independent trials）中成功的次数：它服从二项分布（binomial distribution）。其后的一切都在探讨 λ 很小时会发生甚么事，而所需的工具就是 $\mathrm{e}^{-\lambda}$（或 $\ln(1+\lambda)$）的麦克劳林级数（Maclaurin series），只保留第一个没有被抵消的项。

解答

Y 的分布

设 X 为某一页上的错误数目，则 $X \sim \mathrm{Po}(\lambda)$。若 $X \leq 1$，便称该页为良好页。于是

$$\mathrm{P}(\text{良好}) = \mathrm{P}(X=0) + \mathrm{P}(X=1) = \mathrm{e}^{-\lambda} + \lambda\mathrm{e}^{-\lambda} = (1+\lambda)\mathrm{e}^{-\lambda} = p.$$

样本由 n 页不同的页组成。不同页上的错误数目互相独立，而且全都服从同一个 $\mathrm{Po}(\lambda)$ 分布，所以无论选中哪 n 页，“第 j 页是良好页”这些事件都互相独立，各自的概率均为 p。（条件 $n \ll N$ 只是让我们放心，有足够多的页可供抽样；独立性并不需要这个条件，它来自模型本身。）

所以 Y 是 n 次独立试验中成功的次数，每次成功的概率为 p。某一组指定的 k 页良好页和 $n-k$ 页不良页出现的概率是 $p^k q^{n-k}$，而选出哪 k 页为良好页共有 $\binom{n}{k}$ 种方法，所以

$$\mathrm{P}(Y=k) = \binom{n}{k} p^k q^{n-k}, \qquad k = 0, 1, \ldots, n,$$

其中 $p = (1+\lambda)\mathrm{e}^{-\lambda}$ 及 $q = 1-p$，正如题目所求。换言之，$Y \sim \mathrm{B}(n, p)$。

第 (i) 部分

展开 $\mathrm{e}^{-\lambda} = 1 - \lambda + \frac{1}{2}\lambda^2 - \frac{1}{6}\lambda^3 + \cdots$，再乘以 $1+\lambda$：

$$\begin{aligned} p = (1+\lambda)\mathrm{e}^{-\lambda} &= (1 - \lambda + \tfrac{1}{2}\lambda^2 - \tfrac{1}{6}\lambda^3 + \cdots) + (\lambda - \lambda^2 + \tfrac{1}{2}\lambda^3 - \cdots) \\ &= 1 - \tfrac{1}{2}\lambda^2 + \tfrac{1}{3}\lambda^3 - \cdots . \end{aligned}$$

λ 的一次项互相抵消，所以当 λ 很小时，

$$q = 1 - p = \tfrac{1}{2}\lambda^2 - \tfrac{1}{3}\lambda^3 + \cdots \approx \tfrac{1}{2}\lambda^2$$

因为被略去的各项比 $\frac{1}{2}\lambda^2$ 还要小，相差一个 λ 阶的因子。

> **想一想** 为甚么不是 $q \approx \lambda$，像 $\mathrm{e}^{-\lambda}$ 的首项所暗示的那样?

因为来自 $\mathrm{e}^{-\lambda}$ 的 $-\lambda$ 刚好被乘以 $1+\lambda$ 所产生的 $+\lambda$ 完全抵消。一页只有在至少有两个错误时才是不良页，而 $\mathrm{P}(X \geq 2) \approx \mathrm{P}(X = 2) = \frac{1}{2}\lambda^2\mathrm{e}^{-\lambda} \approx \frac{1}{2}\lambda^2$。这正好从另一个角度独立地印证了 (i)。

第 (ii) 部分

$\mathrm{P}(Y = n) = p^n$，而我们要找满足 $p^n \geq 1 - \lambda$ 的最大 n。两边都是正数（对于 $0 < \lambda < 1$），而 ln 是递增函数，所以这等价于

$$n \ln p \geq \ln(1 - \lambda).$$

现在 $\ln p = \ln(1+\lambda) - \lambda = (\lambda - \frac{1}{2}\lambda^2 + \cdots) - \lambda \approx -\frac{1}{2}\lambda^2$，而 $\ln(1-\lambda) \approx -\lambda$。由于 $\ln p < 0$，除以它时不等式（inequality）的方向会倒转：

$$n \leq \frac{\ln(1-\lambda)}{\ln p} \approx \frac{-\lambda}{-\frac{1}{2}\lambda^2} = \frac{2}{\lambda}.$$

所以满足条件的最大 n 约为 $2/\lambda$。

等价地，利用 (i)：$p^n = (1-q)^n \approx 1 - nq \approx 1 - \frac{1}{2}n\lambda^2$，而 $1 - \frac{1}{2}n\lambda^2 \geq 1 - \lambda$ 成立，当且仅当 $n \leq 2/\lambda$。$(1-q)^n \approx 1 - nq$ 这一步需要 nq 很小；这里 $nq \approx \frac{2}{\lambda} \cdot \frac{1}{2}\lambda^2 = \lambda$，确实很小，所以这个近似是自洽的。

第 (iii) 部分

由于 $Y > 1$ 蕴含 $Y > 0$，

$$\mathrm{P}(Y > 1 \mid Y > 0) = \frac{\mathrm{P}(Y > 1)}{\mathrm{P}(Y > 0)} = \frac{\mathrm{P}(Y > 0) - \mathrm{P}(Y = 1)}{\mathrm{P}(Y > 0)} = 1 - \frac{npq^{n-1}}{1 - q^n}.$$

（这里 $n \geq 2$ 是有作用的：它令 $Y > 1$ 成为可能发生的事件，并令幂 q^{n-1} 的次数至少与 q 相同。）现在固定 n，并设 λ 很小。那么 $q \approx \frac{1}{2}\lambda^2$ 很小，所以 $p \approx 1$ 及 $1 - q^n \approx 1$，由此得

$$\mathrm{P}(Y > 1 \mid Y > 0) \approx 1 - nq^{n-1} \approx 1 - n\left(\frac{\lambda^2}{2}\right)^{n-1},$$

最后一步用了 (i)。这就是所求的结果。

验算

- 以 $\lambda = 0.8$ 模拟 200,000 个样本，每个样本有 $n = 5$ 页泊松（Poisson）页，所得 $Y = 0, \ldots, 5$ 的相对频数与 $\mathrm{B}(5, p)$ 相符，误差在 0.002 以内。
- 当 $\lambda = 0.01$ 时，$q = 4.967 \times 10^{-5}$，而 $\frac{1}{2}\lambda^2 = 5 \times 10^{-5}$。
- 当 $\lambda = 0.01$ 时，满足 $p^n \geq 0.99$ 的最大 n 的精确值是 202，而 $2/\lambda = 200$。（在每个级数中多保留一项，会得到 $2/\lambda + \frac{7}{3}$，所以当 $2/\lambda$ 增大时，误差仍然有界。）
- 当 $\lambda = 0.01$, $n = 3$ 时：精确值为 $1 - \mathrm{P}(Y > 1 \mid Y > 0) = 7.40 \times 10^{-9}$，而 $3(\lambda^2/2)^2 = 7.5 \times 10^{-9}$。

常见错误

- 把“少于两个错误”理解为只有 $X = 0$，或理解为 $X \leq 2$。
- 把 $(1 + \lambda)\mathrm{e}^{-\lambda}$ 的展开式停在一次项，然后断定 $q \approx 0$：第一个没有被抵消的项其实是二次项。
- 在 (ii) 中除以 $\ln p < 0$ 时，忘记把不等式的方向倒转。

STEP 之外

第 (ii) 部分是渐近分析（asymptotic analysis）的一个例子：条件由乘积 $n\lambda^2$ 所控制，所以 n 的自然尺度是 $1/\lambda$，而像 p^n 这样的量，其表现就如 $\mathrm{e}^{-n\lambda^2/2}$。二项分布的泊松近似，以及概率论中许多“稀有事件”的极限，背后都是同一套推理。Grimmett 与 Welsh 所著的 *Probability: An Introduction*（概率论入门教材）对这些极限定理有严谨细致的阐述。

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。