詳細解答

STEP II 2005，第 13 題

概率與統計 • 泊松分佈與二項分佈

STEP II 2005，第 13 題

繼續閱讀之前，請先看題目（原題為英文）。題目載於 STEP Support Programme 的試題資料庫：

step.maths.org/questions/05-s2-q13

關鍵思路

每一頁抽樣頁要麼錯誤少於兩個，要麼不是，而且各頁互相獨立，所以 Y 數算的是 n 次獨立試驗（independent trials）中成功的次數：它服從二項分佈（binomial distribution）。其後的一切都在探討 λ 很小時會發生甚麼事，而所需的工具就是 $\mathrm{e}^{-\lambda}$（或 $\ln(1+\lambda)$）的麥克勞林級數（Maclaurin series），只保留第一個沒有被抵消的項。

解答

Y 的分佈

設 X 為某一頁上的錯誤數目，則 $X \sim \mathrm{Po}(\lambda)$。若 $X \leq 1$，便稱該頁為良好頁。於是

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

樣本由 n 頁不同的頁組成。不同頁上的錯誤數目互相獨立，而且全都服從同一個 $\mathrm{Po}(\lambda)$ 分佈，所以無論選中哪 n 頁，「第 j 頁是良好頁」這些事件都互相獨立，各自的概率均為 p。（條件 $n \ll N$ 只是讓我們放心，有足夠多的頁可供抽樣；獨立性並不需要這個條件，它來自模型本身。）

所以 Y 是 n 次獨立試驗中成功的次數，每次成功的概率為 p。某一組指定的 k 頁良好頁和 $n-k$ 頁不良頁出現的概率是 $p^k q^{n-k}$，而選出哪 k 頁為良好頁共有 $\binom{n}{k}$ 種方法，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 $q = 1-p$，正如題目所求。換言之，$Y \sim \mathrm{B}(n, p)$。

第 (i) 部分

展開 $\mathrm{e}^{-\lambda} = 1 - \lambda + \frac{1}{2}\lambda^2 - \frac{1}{6}\lambda^3 + \cdots$，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 的一次項互相抵消，所以當 λ 很小時，

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

因為被略去的各項比 $\frac{1}{2}\lambda^2$ 還要小，相差一個 λ 階的因子。

> **想一想** 為甚麼不是 $q \approx \lambda$，像 $\mathrm{e}^{-\lambda}$ 的首項所暗示的那樣？

因為來自 $\mathrm{e}^{-\lambda}$ 的 $-\lambda$ 剛好被乘以 $1+\lambda$ 所產生的 $+\lambda$ 完全抵消。一頁只有在至少有兩個錯誤時才是不良頁，而 $\mathrm{P}(X \geq 2) \approx \mathrm{P}(X = 2) = \frac{1}{2}\lambda^2\mathrm{e}^{-\lambda} \approx \frac{1}{2}\lambda^2$。這正好從另一個角度獨立地印證了 (i)。

第 (ii) 部分

$\mathrm{P}(Y = n) = p^n$，而我們要找滿足 $p^n \geq 1 - \lambda$ 的最大 n。兩邊都是正數(對於 $0 < \lambda < 1$)，而 ln 是遞增函數，所以這等價於

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

現在 $\ln p = \ln(1+\lambda) - \lambda = (\lambda - \frac{1}{2}\lambda^2 + \cdots) - \lambda \approx -\frac{1}{2}\lambda^2$，而 $\ln(1-\lambda) \approx -\lambda$。由於 $\ln p < 0$，除以它時不等式（inequality）的方向會倒轉：

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

所以滿足條件的最大 n 約為 $2/\lambda$。

等價地，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，當且僅當 $n \leq 2/\lambda$。$(1-q)^n \approx 1 - nq$ 這一步需要 nq 很小；這裏 $nq \approx \frac{2}{\lambda} \cdot \frac{1}{2}\lambda^2 = \lambda$，確實很小，所以這個近似是自洽的。

第 (iii) 部分

由於 $Y > 1$ 蘊含 $Y > 0$，

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

（這裏 $n \geq 2$ 是有作用的：它令 $Y > 1$ 成為可能發生的事件，並令冪 q^{n-1} 的次數至少與 q 相同。）現在固定 n，並設 λ 很小。那麼 $q \approx \frac{1}{2}\lambda^2$ 很小，所以 $p \approx 1$ 及 $1 - q^n \approx 1$，由此得

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

最後一步用了 (i)。這就是所求的結果。

驗算

- 以 $\lambda = 0.8$ 模擬 200,000 個樣本，每個樣本有 $n = 5$ 頁泊松（Poisson）頁，所得 $Y = 0, \ldots, 5$ 的相對頻數與 $\mathrm{B}(5, p)$ 相符，誤差在 0.002 以內。
- 當 $\lambda = 0.01$ 時，$q = 4.967 \times 10^{-5}$，而 $\frac{1}{2}\lambda^2 = 5 \times 10^{-5}$。
- 當 $\lambda = 0.01$ 時，滿足 $p^n \geq 0.99$ 的最大 n 的精確值是 202，而 $2/\lambda = 200$。（在每個級數中多保留一項，會得到 $2/\lambda + \frac{7}{3}$，所以當 $2/\lambda$ 增大時，誤差仍然有界。）
- 當 $\lambda = 0.01$，$n = 3$ 時：精確值為 $1 - \mathrm{P}(Y > 1 \mid Y > 0) = 7.40 \times 10^{-9}$，而 $3(\lambda^2/2)^2 = 7.5 \times 10^{-9}$。

常見錯誤

- 把「少於兩個錯誤」理解為只有 $X = 0$，或理解為 $X \leq 2$。
- 把 $(1 + \lambda)\mathrm{e}^{-\lambda}$ 的展開式停在一次項，然後斷定 $q \approx 0$：第一個沒有被抵消的項其實是二次項。
- 在 (ii) 中除以 $\ln p < 0$ 時，忘記把不等式的方向倒轉。

STEP 之外

第 (ii) 部分是漸近分析（asymptotic analysis）的一個例子：條件由乘積 $n\lambda^2$ 所控制，所以 n 的自然尺度是 $1/\lambda$，而像 p^n 這樣的量，其表現就如 $\mathrm{e}^{-n\lambda^2/2}$。二項分佈的泊松近似，以及概率論中許多「稀有事件」的極限，背後都是同一套推理。Grimmett 與 Welsh 所著的 *Probability: An Introduction*（概率論入門教材）對這些極限定理有嚴謹細緻的闡述。

Beyond STEP 詳細解答：https://beyondstep.keithmcnulty.org/zh-hant/。解答文字採用 CC0 1.0；題目本身版權屬 Cambridge University Press & Assessment 所有（©），本文件不予轉載。由機器從英文翻譯而成；尚未經審閱。