详细解答

# STEP II 1998，第 14 题

概率与统计 · 二项分布与泊松分布的期望值

**STEP II 1998，第 14 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/98-s2-q14

## 关键思路

虽然资料库把这题另作归类，但它其实是一道概率（probability）题：只要 Doom 博士在新方案下的期望薪金高于固定薪金 $A$，她就应该接受新方案。不合格人数服从二项分布（binomial distribution），所以我们需要求 $\mathrm{E}[1/(1+X)]$。棘手的因子 $\frac{1}{k+1}$ 可以用恒等式 $\frac{1}{k+1}\binom{N}{k} = \frac{1}{N+1}\binom{N+1}{k+1}$ 吸收掉，于是该和式就变成 $(p+q)^{N+1}$ 几乎完整的二项展开式。同样的技巧配合 $\frac{\lambda^k}{(k+1)!}$，便可处理泊松（Poisson）模型。

## 解答

### 何时接受新方案

设 $X$ 为 Doom 博士的学生中不合格的人数，并记 $q = 1-p$。各学生是否不合格互相独立（independent），每人不合格的概率均为 $p$，所以 $X \sim \mathrm{B}(N,p)$：

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

我们假设 Doom 博士愿意以期望薪金来衡量两个方案。在新方案下，她的薪金是随机变量（random variable）$S = B/(1+X)$，所以

$$\mathrm{E}(S) = B\sum_{k=0}^{N} \frac{1}{k+1}\binom{N}{k} p^k q^{N-k}.$$

? **想一想** 怎样消去和式中的 $\frac{1}{k+1}$？

利用阶乘（factorial）：

$$\frac{1}{k+1}\binom{N}{k} = \frac{N!}{(k+1)!\,(N-k)!} = \frac{1}{N+1}\cdot\frac{(N+1)!}{(k+1)!\,((N+1)-(k+1))!}$$
$$= \frac{1}{N+1}\binom{N+1}{k+1}.$$

把这个结果代入，并同时乘以和除以 $p$，使 $p$ 的幂次与新的二项式系数（binomial coefficient）相配，

$$\begin{aligned}\mathrm{E}(S) &= \frac{B}{(N+1)p}\sum_{k=0}^{N}\binom{N+1}{k+1}p^{k+1}q^{N-k}\\ &= \frac{B}{(N+1)p}\sum_{j=1}^{N+1}\binom{N+1}{j}p^{j}q^{N+1-j}\\ &= \frac{B}{(N+1)p}\left[(p+q)^{N+1}-q^{N+1}\right]\end{aligned}$$

其中先设 $j = k+1$, 再运用二项式定理（binomial theorem）, 因为该和式正是 $(p+q)^{N+1}$ 的完整展开式，只欠 $j=0$ 的一项 $q^{N+1}$。由于 $p+q=1$，

$$\mathrm{E}(S) = \frac{B(1-(1-p)^{N+1})}{(N+1)p}.$$

若 $\mathrm{E}(S) > A$，她就应该接受。两边同乘 $(N+1)p$，而此数为正（公式要有意义，须有 $p>0$；若 $p=0$，她只需直接比较 $B$ 与 $A$），便得

$$A(N+1)p < B(1-(1-p)^{N+1}),$$

这就是所求的条件。

### 泊松模型

当 $N$ 很大而 $p$ 很小，且 $Np$ 大小适中时，二项分布 $\mathrm{B}(N,p)$ 可以用平均值为 $\lambda = Np$ 的泊松分布（Poisson distribution）作良好的近似。因此，如果 Doom 博士任教的班人数众多，每名学生不合格的机会都很低，而且各人是否不合格仍然互相独立，她的学生的不合格人数就可以用泊松分布作模型。（更一般地说，若不合格的情况逐一出现、互相独立，且平均发生率恒定，便适合采用泊松模型。）

设 $X \sim \mathrm{Po}(\lambda)$，$\lambda = Np$，同样的想法依然可行，这次用的是 $\frac{1}{k+1}\cdot\frac{1}{k!} = \frac{1}{(k+1)!}$：

$$\begin{aligned}\mathrm{E}\Big(\frac{B}{1+X}\Big) &= B\sum_{k=0}^{\infty}\frac{1}{k+1}\cdot\frac{\mathrm{e}^{-\lambda}\lambda^{k}}{k!} = \frac{B\mathrm{e}^{-\lambda}}{\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k+1}}{(k+1)!}\\ &= \frac{B\mathrm{e}^{-\lambda}}{\lambda}(\mathrm{e}^{\lambda}-1) = \frac{B(1-\mathrm{e}^{-\lambda})}{\lambda},\end{aligned}$$

因为最后的和式正是指数级数（exponential series），只欠首项 1。所以在这个模型下，Doom 博士的期望薪金为

$$\frac{B(1-\mathrm{e}^{-Np})}{Np} \text{ 英镑.}$$

## 验算

- $N=1$：直接计算得 $\mathrm{E}(S)=B\left(q+\frac{1}{2}p\right)=B\left(1-\frac{1}{2}p\right)$，而公式给出 $B\,\frac{1-q^2}{2p}=B\,\frac{1+q}{2}=B\left(1-\frac{1}{2}p\right)$。
- $p\to 0$：二项分布的表达式趋向 $B$（没有人不合格），而当 $\lambda\to 0$ 时，泊松分布的表达式亦然。
- 在近似理应准确的情况下，两个答案相当吻合：取 $N=1000$，$p=0.003$，由二项分布公式得 $\mathrm{E}(S)/B$ 为 0.31655，由泊松公式则得 0.31674。取 $N=20$，$p=0.3$ 时，逐项精确计算与模拟的结果均为 $\mathrm{E}(S)/B\approx 0.1586$。

## 常见错误

- 写成 $\mathrm{E}\left(\frac{1}{1+X}\right)=\frac{1}{1+\mathrm{E}(X)}=\frac{1}{1+Np}$。期望值（expectation）不能直接穿过非线性函数；事实上 $\frac{1}{1+x}$ 是凸函数（convex function），所以根据延森不等式（Jensen's inequality），真正的值比这个大。
- 忘记了重新编排指标后的和式欠了 $j=0$ 的一项，结果得出 $\mathrm{E}(S)=B/((N+1)p)$。

## STEP 之外

这里的技巧其实是一个特例：计算 $\mathrm{E}\big(1/(1+X)\big)$ 时可以利用 $\frac{1}{1+x}=\int_0^1 t^x\,\mathrm{d}t$，从而把它与概率生成函数（probability generating function）$G(t)=\mathrm{E}(t^X)$ 联系起来：答案就是 $\int_0^1 G(t)\,\mathrm{d}t$。试以 $G(t)=(q+pt)^N$ 和 $G(t)=\mathrm{e}^{\lambda(t-1)}$ 验证一下。关于生成函数，以及二项分布收敛（convergence）至泊松分布，Grimmett 与 Welsh 所著的 *Probability: An Introduction*（概率论入门教材）有很好的讲解。

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。