$$
\begin{aligned}
&= \sum_{k=1}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k} + np \\
&= (1-p)^n \sum_{k=2}^{n} k(k-1)\binom{n}{k}\left(\frac{p}{1-p}\right)^k + np
\end{aligned}
$$

对二项展开式 $(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k$ 两边同时求导两次可得

$$
n(n-1)(1+x)^{n-2} = \sum_{k=2}^{n}\binom{n}{k}k(k-1)x^{k-2} \;\Rightarrow\; n(n-1)x^2(1+x)^{n-2} = \sum_{k=2}^{n}\binom{n}{k}k(k-1)x^k,
$$

将 $x = p/(1-p)$ 带入有

$$
E(X^2) = n(n-1)p^2 + np = n^2p^2 + np(1-p),
$$

从而得到 $\mathrm{Var}(X) = E[X^2] - (E[X])^2 = np(1-p)$.

例 3.9 有 5 个选择题, 每个选择题有 4 种答案, 只有一种正确, 求一学生随机猜对 4 个选择题的概率?

解 将每一个选择题看作一次 Bernoulli 试验, 事件 A 表示猜正确, 则有 $P(A) = 1/4$. 整个问题等价于 5 重 Bernoulli 试验, 用 X 表示学生猜对题的个数, 则 $X \sim B(5, 1/4)$, 从而得到

$$
P(X=4) = \binom{5}{4}\cdot\frac{1}{4^4}\cdot\frac{3}{4} = \frac{15}{4^5}.
$$

3.3.4 几何分布

定义 3.7 在多重 Bernoulli 试验, 设事件 A 发生的概率为 p. 用随机变量 X 表示事件 A 首次发生时的试验次数, 则 X 的取值为 $1, 2, \cdots$, 其分布列为

$$
P(X=k) = (1-p)^{k-1}p \qquad (k \geqslant 1)
$$

称 X 服从参数为 p 的几何分布, 记为 $X \sim G(p)$.

首先可知 $P(X=k) = (1-p)^{k-1}p \geqslant 0$ 以及

$$
\sum_{k=1}^{\infty} P(X=k) = p\sum_{k=1}^{\infty}(1-p)^{k-1} = p\frac{1}{1-(1-p)} = 1,
$$

从而验证了几何分布构成一个分布列. 对几何分布, 我们有

性质 3.8 若随机变量 $X \sim G(p)$ $(0<p<1)$, 则有

$$E(X)=\frac{1}{p} \qquad 和 \qquad \mathrm{Var}(X)=\frac{1-p}{p^2}.$$

证明 根据期望的定义有

$$E(X)=\sum_{k=1}^{\infty} kP(X=k)=\sum_{k=1}^{\infty} kp(1-p)^{k-1}=p\sum_{k=1}^{\infty} k(1-p)^{k-1}.$$

对级数展开式 $(1-x)^{-1}=\sum_{k=0}^{\infty} x^k$ 两边求导有

$$\sum_{k=1}^{\infty} kx^{k-1}=\frac{1}{(1-x)^2}.$$

令 $x=1-p$ 可证 $E(X)=1/p$. 对于随机变量 X 的方差, 首先计算

$$E(X^2)=\sum_{k=1}^{\infty} k^2p(1-p)^{k-1}=p\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-1}+1/p.$$

对级数展开式 $(1-x)^{-1}=\sum_{k=0}^{\infty} x^k$ 两边求二阶导有

$$\sum_{k=2}^{\infty} k(k-1)x^{k-2}=\frac{2}{(1-x)^3} \quad\Longrightarrow\quad \sum_{k=2}^{\infty} k(k-1)x^{k-1}=\frac{2x}{(1-x)^3}.$$

令 $x=1-p$ 可得 $E(X^2)=(2-p)/p^2$, 于是有 $\mathrm{Var}(X)=E(X^2)-(EX)^2=(1-p)/p^2$.

下面给出几何分布的一个重要性质: 无记忆性 (memoryless property).

定理 3.4 设随机变量 $X \sim G(p)$, 对任意正整数 m, n, 有

$$P(X>m+n|X>m)=P(X>n).$$

几何分布无记忆性的直观解释: 假设现在已经历 m 次失败, 从当前起至成功的次数与 m 无关. 例如, 一人赌博时前面总输, 觉得下一次应该赢了, 然而无记忆性给出下一次是否赢与前面输了多少次无关.

证明 根据几何分布的定义, 对任何正整数 k 有

$$P(X>k)=\sum_{i=k+1}^{\infty} p(1-p)^{i-1}=p\sum_{i=k+1}^{\infty}(1-p)^{i-1}=p\frac{(1-p)^k}{1-(1-p)}=(1-p)^k.$$

根据条件概率的定义有

$$P(X>m+n|X>m)=\frac{P(X>m+n)}{P(X>m)}=\frac{(1-p)^{m+n}}{(1-p)^m}=(1-p)^n=P(X>n)$$

这里利用事件 $\{X>m+n\}\cap\{X>m\}=\{X>m+n\}$.

例 3.10 古人非常重视生男孩且资源有限, 规定每个家庭可生一个男孩, 如果没男孩则可以继续生育直至有一个男孩; 若已有一个男孩, 则不再生育. 多年后男女比例是否会失衡?

解 对一个家庭而言, 用随机变量 X 表示该家庭的小孩个数, 则 $X=1,2,\cdots$, 以及

$$P(X=k)=p(1-p)^{k-1},$$

这里 $p=1/2$ 表示生男孩的概率. 根据几何分布的期望有一个家庭小孩数的期望为 $E[X]=1/p=2$, 由此可得一个家庭的小孩男女比例 $1:1$.

3.3.5 Pascal/负二项分布

几何分布考虑在多重试验中事件 A 首次发生时所进行的试验次数, 可以这个问题进一步推广到事件 A 第 r 次发生时所进行的试验次数. 具体而言, 在多次 Bernoulli 试验中, 随机事件 A 发生的概率为 $p\in(0,1)$. 用 X 表示事件 A 第 r 次成功时发生的试验次数, 则 X 取值 $r,r+1,r+2,\cdots$, 其分布列为

$$P(X=k)=\binom{k-1}{r-1}p^{r-1}(1-p)^{k-r}\cdot p=\binom{k-1}{r-1}p^r(1-p)^{k-r},\quad k=r,r+1,r+2,\cdots,$$

称随机变量 X 服从参数为 r 和 p 的 **负二项分布**.

易知 $P(X=k)\geqslant 0$, 下面证明

$$\sum_{k=r}^{\infty}P(X=k)=p^r\sum_{k=r}^{\infty}\binom{k-1}{r-1}(1-p)^{k-r}=1.$$

设 $q=1-p$, 根据泰勒展式有

$$p^{-r}=(1-q)^{-r}=\sum_{t=0}^{\infty}\binom{t+r-1}{r-1}q^t=\sum_{k=r}^{\infty}\binom{k-1}{r-1}(1-p)^{k-r}\quad(令\ k=t+r).$$

定理 3.5 设随机变量 X 服从参数为 $p\in(0,1)$ 和 $r>0$ 的负二项分布, 则有

$$E(X)=\frac{r}{p}\quad 和\quad \mathrm{Var}(X)=\frac{r(1-p)}{p^2}.$$

证明 对于期望 $E(X)$ 有

$$\begin{aligned} E(X) &= \sum_{k=r}^{\infty} k \cdot P(X=k) = \sum_{k=r}^{\infty} k \binom{k-1}{r-1} p^r (1-p)^{k-r} \\ &= \frac{r}{p} \sum_{k=r}^{\infty} \binom{k}{r} p^{r+1} (1-p)^{k-r} = \frac{r}{p} \sum_{k=r}^{\infty} \binom{k+1-1}{r+1-1} p^{r+1} (1-p)^{k-r} = \frac{r}{p}, \end{aligned}$$

这里利用

$$\sum_{k=r}^{\infty} \binom{k+1-1}{r+1-1} p^{r+1} (1-p)^{k-r} = 1.$$

对 $E(X^2)$ 的计算, 类似有

$$\begin{aligned} E(X^2) &= \sum_{k=r}^{\infty} k^2 P(X=k) = \sum_{k=r}^{\infty} k^2 \binom{k-1}{r-1} p^r (1-p)^{k-r} \\ &= \frac{r}{p} \sum_{k=r}^{\infty} (k+1-1) \binom{k}{r} p^r (1-p)^{k-r} \\ &= \frac{r(r+1)}{p^2} \sum_{k=r}^{\infty} \binom{k+1}{r+1} p^{r+1} (1-p)^{k-r} - \frac{r}{p} \sum_{k=r}^{\infty} \binom{k}{r} p^r (1-p)^{k-r} \\ &= \frac{r(r+1)}{p^2} - \frac{r}{p}, \end{aligned}$$

由此可得

$$Var(X) = E(X^2) - (E(X))^2 = \frac{r(r+1)}{p^2} - \frac{r}{p} - \frac{r^2}{p^2} = \frac{r(1-p)}{p^2}.$$

3.3.6 泊松分布

定义 3.8 若随机变量 X 的分布列为

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, 2, \cdots)$$

其中 $\lambda > 0$ 是一个给定的常数, 称随机变量 X 服从参数为 λ 的泊松分布, 记为 $X \sim P(\lambda)$.

容易验证 $P(X=k) = \lambda^k e^{-\lambda}/k! \geqslant 0$, 并根据泰勒展式 $e^{\lambda} = \sum_{k=0}^{\infty} \lambda^k/k!$ 有

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1.$$

泊松分布用于描述大量试验中稀有事件出现次数的概率模型, 例如: 在一段时间内电话收到的呼叫次数, 放射物在一段时间内放射的粒子数, 一段时间内通过某路口的出租车数, 一本书中一页出现的语法错误数, 一天内到一所银行办理业务的顾客数等.

性质 3.9 对任意给定的 $\lambda > 0$, 若 $X \sim P(\lambda)$, 则

$$E(X) = \lambda \qquad \text{和} \qquad \mathrm{Var}(X) = \lambda.$$

证明 根据期望的定义和幂级数 $e^\lambda = \sum_{k=0}^\infty \lambda^k/k!$ 有

$$E(X) = \sum_{k=0}^\infty k \cdot P(X = k) = \sum_{k=1}^\infty k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^\infty \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

对于随机变量的方差, 首先计算

$$\begin{aligned} E[X^2] &= \sum_{k=0}^\infty k^2 P(X = k) = \sum_{k=1}^\infty k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^\infty k \cdot \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda^2 e^{-\lambda} \sum_{k=2}^\infty \frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda} \sum_{k=1}^\infty \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda^2 + \lambda. \end{aligned}$$

从而得到 $\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \lambda$.

例 3.11 设随机变量 X 服从参数为 λ 的泊松分布, 且 $P(X = 1) = P(X = 2)$, 求 $P(X \geqslant 4)$.

解 根据泊松分布的定义可知 $P(X = k) = \lambda^k e^{-\lambda}/k!$ 和 $P(X = 1) = P(X = 2)$ 可得

$$\lambda e^{-\lambda} = \lambda^2 e^{-\lambda}/2 \quad \Longrightarrow \quad \lambda = 2,$$

进一步得到

$$P(X \geqslant 4) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) = 1 - 5e^{-2} - 4e^{-2}/3.$$

下面研究二项分布和泊松分布的关系, 即泊松定理:

定理 3.6 对任意给定的常数$\lambda > 0$, n 为任意正整数, 设 $np_n = \lambda$, 则对任意给定的非负整数 k, 有

$$\lim_{n\to\infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

证明 由 $p_n = \lambda/n$, 有

$$\begin{aligned} \binom{n}{k} p_n^k (1 - p_n)^{n-k} &= \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} (1 - \frac{1}{n}) \cdots (1 - \frac{k-1}{n}) \left(1 - \frac{\lambda}{n}\right)^{n-k} \end{aligned}$$

$$= \quad \frac{\lambda^k}{k!}(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})\left(1-\frac{\lambda}{n}\right)^{\frac{n}{\lambda}\frac{n-k}{n}\lambda}$$

当 $n\to\infty$ 时有 $(1-\frac{\lambda}{n})^{\frac{n}{\lambda}}\to e^{-1}$ 以及 $\frac{n-k}{n}\lambda\to\lambda$, 从而完成证明.

泊松分布的应用: 若随机变量 $X\sim B(n,p)$, 当 n 比较大而 p 比较小时, 令 $\lambda=np$, 有

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda}.$$

即利用泊松分布近似计算二项分布.

例 3.12 射击训练每次命中目标的概率为 0.002, 现射击 1000 次, 求命中目标在 10 次与 50 次之间的概率. (用泊松近似计算)

解 将 1000 次射击可看作 1000 重 Bernoulli 试验, 设随机变量 X 表示 1000 射击训练中射中目标的次数, 则 $X\sim B(1000,0.002)$, 利用泊松分布近似, 则可以看作 $X\sim P(2)$, 于是有

$$P(500\leqslant X\leqslant 600)=\sum_{k=10}^{50}\binom{1000}{k}(0.002)^k 0.998^{1000-k}\approx\sum_{k=10}^{50}\frac{2^k}{k!}e^{-2}.$$

例 3.13 有 80 台同类型设备独立工作, 发生故障的概率是 0.01, 一台设备发生故障时只能由一人处理, 考虑方案: I) 由四人维护, 每人单独负责 20 台; II) 由三人共同维护 80 台. 哪种方案更为可取?

解 首先讨论方案 I), 用事件 A_i 表示第 i 人负责的设备发生故障不能及时维修, 用 X_i 为第 i 人负责的 20 台设备同一时刻发生故障的台数, 则有 $X\sim B(20,0.01)$, 根据泊松定理有近似有 $X\sim P(0.2)$, 进一步有

$$P(A_i)=P(X_i\geqslant 2)=1-P(X=0)-P(X=1)\approx 1-\sum_{k=0}^{n}\frac{(0.2)^k}{k!}e^{-0.2}\approx 0.0175.$$

因四人独立维修, 有设备发生故障时而不能及时的概率

$$P(A_1\cup A_2\cup A_3\cup A_4)\geqslant P(A_1)\approx 0.0175.$$

对方案 II): 设随机变量 Y 为 80 台设备中同一时刻发生故障的台数, 则 $Y\sim B(80,0.01)$, 根据泊松定理有近似有 $X\sim P(0.8)$, 则有设备发生故障不能及时维修的概率为

$$P(Y\geqslant 4)=1-\sum_{k=0}^{3}P(Y=k)\approx 1-\sum_{k=0}^{3}\frac{(0.8)^k}{k!}e^{-0.8}\approx 0.0091.$$

由此比较可知方案 II) 更优.

3.4 案例分析: 随机二叉树叶结点的平均高度

初始为一个根结点, 再每一次迭代过程中, 随机选择一个叶子结点, 将该叶子结点分裂为左、右叶子结点, 由此重复进行 k 次, 求此随机树一个叶结点的平均高度.