第一章　独立随机变量

随机过程是研究一列随机变量 $X_1, X_2, \ldots$，或者一族随机变量 $\{X(t)\}_{t\geq 0}$ 的学问，其重点会放在随机变量之间的关系上．最基本的情况便是这些随机变量都是互相独立的．在概率论课上，我们学习过经典的大数定律和中心极限定理，这便是独立随机变量序列的经典结果．我们在这一节，介绍一些独立随机变量序列的非渐近结果，它们不是我们这本书的重点，但我们会通过这些讨论引入几个重要的概率模型和数学工具，是未来研究更复杂的随机过程的基础．

概率论课上遇到的大数定律和中心极限定理都是讨论一系列随机变量 $X_1, X_2, \ldots, X_n, \ldots$ 当 $n \to \infty$ 时的渐近性质．我们这里说的**非渐近**性质指的是对于固定（但足够大）的 n，序列 $X_1, \ldots, X_n$ 的性质．

1.1　投球入箱问题（balls-into-bins）

投球入箱是一个简单的随机过程，或者说是随机试验：将 m 个球均匀随机地投入 n 个箱子中．我们用 $[m]$ 来编号所有的球，用 $[n]$ 来编号所有的箱子，并用 X_i 来表示第 i 个球落入的箱子的编号． 那么，$X_1, X_2, \ldots, X_m$ 就是取值为 $[n]$ 的独立随机变量．假设每一个 X_i 都是均匀的随机变量是最简单的情况．对于这个过程，我们可以提出许多有趣的问题．我们这儿介绍俩经典的．

对于自然数 $n \in \mathbb{N}$，我们用 $[n]$ 表示集合 $\{1, 2, \ldots, n\}$．

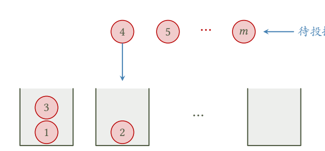

图 1.1: 投球入箱

生日悖论（birthday paradox）

生日悖论指的是在一个有 30 个同学左右的班级中某些同学很可能共享相同生日这一反直觉现象．将箱子视为日期、球视为学生，两个学生生日相同的事件可以建模为某个箱子包含多于一个球的情况．

注意到每个球的投掷是独立的．在已投掷 $k-1$ 个球且无碰撞的条件下，投掷第 k 个球后仍无碰撞的概率为 $\frac{n-k+1}{n}$．因此，

$$\mathbb{P}\left[\text{无相同生日}\right] = \prod_{k=1}^{m} \frac{n-k+1}{n} = \prod_{k=1}^{m-1} \left(1 - \frac{k}{n}\right).$$

由 $1 + x \leq e^x$，上面的概率有上界

$$\exp\left(-\frac{\sum_{k=1}^{m-1} k}{n}\right) = \exp\left(-\frac{m(m-1)}{2n}\right).$$

于是，可以看到当 m 是 $\sqrt{n}$ 的一个足够大的常数倍时，该概率可以任意接近 0．

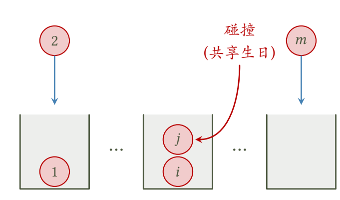

图 1.2: 生日悖论

$1 + x \leq e^x$ 是一个常用的不等式，它来自于对于 e^x 的线性近似，并且对于任何 $x \in \mathbb{R}$ 均成立．我们有的时候也想用另一方向的不等式，一个实用的并且方便记忆的选择是：

$$\forall k > 0,\ 1 + 1/k \geq e^{1/(k+1)};$$
$$\forall k > 1,\ 1 - 1/k \geq e^{-1/(k-1)}.$$

奖券收集（coupon collector）问题

奖券收集问题是如下问题：如果某手机游戏开卡包机制是每一次均匀随机的开出 n 种不同的卡片之一，总共需要开多少包才能收集全部 n 种卡片呢？用投球入箱的语言表述，我们用箱子来表示每一种卡片，用每一个球来表示开启的每一包卡牌，那么即需要投入多少球才能没有空箱子.

需要购买的卡包数是一个随机变量，我们用 Y 来表示．我们首先利用期望的线性性可计算 Y 的期望．设 Y_i 表示在已持有 $i-1$ 种卡片时，收集第 i 种卡片所需的开包次数．那么显然有 $Y=\sum_{i=1}^n Y_i$．由期望线性性：

$$\mathbf{E}[Y]=\sum_{i=1}^{n}\mathbf{E}[Y_i].$$

显然 $Y_i\sim\mathrm{Geom}\left(\frac{n-i+1}{n}\right)$，因此 $\mathbf{E}[Y_i]=\frac{n}{n-i+1}$．从而

$$\mathbf{E}[Y]=\sum_{i=1}^{n}\frac{n}{n-i+1}=n\cdot H_n.$$

对于几何分布 $X\sim\mathrm{Geom}(p)$，有 $\mathbf{E}[X]=p^{-1}$.

上述使用期望的线性性进行计算的技巧是非常常用的．此外，奖券收集问题的结论也值得记住，在我们后续课程中会多次的遇到它.

这儿 $H_n=\sum_{i=1}^n\frac{1}{i}$ 为调和级数的前 n 项，满足 $H_n\sim\log n+\gamma$，其中常数 $\gamma=0.577\ldots$ 被称为**欧拉常数**.

1.2　集中不等式（concentration inequalities）

在奖券收集问题里，从实际的角度来说，我们知道平均 $n\cdot H_n$ 包可以集齐全卡是不够的，因为实际开包的过程可能和所谓的“期望值”相去甚远．我们需要知道关于随机变量 Y 更进一步的信息，比如，如果我想保证以至少 99% 的概率收集全所有卡片，我需要开多少包？这就涉及到随机变量的集中不等式.

马尔可夫不等式（Markov's inequality）

马尔可夫不等式可能是最简单集中不等式，它对于任意非负的随机变量均成立.

定理 1.1　马尔可夫不等式.

对任意非负随机变量 $X\geq 0$ 及 $a>0$，

$$\mathbb{P}[X\geq a]\leq\frac{\mathbf{E}[X]}{a}.$$

证明. 由于 X 非负，由全期望公式，有 $\mathbf{E}[X]\geq a\cdot\mathbb{P}[X\geq a]+0\cdot\mathbb{P}[X<a]$．这便等价于我们想要证明的. □

全期望公式指的是对于样本空间的一个可数划分 $\{A_i\}$，其中每一个 A_i 是一个事件，有 $\mathbf{E}[X]=\sum_i\mathbf{E}[X\mid A_i]\cdot\mathbb{P}[A_i]$.

示例 1.2 奖券收集的集中性.

回忆 Y 为所需球数．应用马尔可夫不等式，对 $c > 0$ 有

$$\mathbb{P}[Y \geq c] \leq \frac{\mathbf{E}[Y]}{c} = \frac{n \cdot H_n}{c}.$$

因此，需要开超过 $100 \cdot n \cdot H_n$ 包的概率小于 0.01.

取 $n = 100$：$H_{100} \approx 5.187$，$\mathbf{E}[Y] \approx 519$．马尔可夫不等式要求 $c \geq 100 \times 519 \approx 51900$ 包才能保证 99% 的概率集齐.

切比雪夫不等式（Chebyshev's inequality）

马尔可夫不等式的成立仅仅要求随机变量 X 是非负的，但实际上，如果我们考虑 X 和它的期望的偏差程度，即 $\mathbb{P}[|X - \mathbf{E}[X]| \geq a]$，那么马尔可夫不等式总能够给出一个上界，即 $\mathbb{P}[|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathbf{E}[|X - \mathbf{E}[X]|]}{a}$．因此，马尔可夫不等式的成立可以看成是无条件的，但这也说明了它在很多随机变量上并不紧，因为它没有利用足够多的随机变量的信息.

我们用一个简单的技巧即可改进它．对于一个（在我们关心的范围内的）递增的函数 $f : \mathbb{R} \to \mathbb{R}$，我们有 $\mathbb{P}[X \geq a] = \mathbb{P}[f(X) \geq f(a)]$．我们便可以对后者使用马尔可夫不等式.

定理 1.3 切比雪夫不等式.

对任意具有有限期望 $\mathbf{E}[X]$ 的随机变量及 $a > 0$，有

$$\mathbb{P}[|X - \mathbf{E}[X]| \geq a] \leq \frac{\mathbf{Var}[X]}{a^2}.$$

证明. 令 $Y = |X - \mathbf{E}[X]|$，显然 $Y \geq 0$．因此

$$\mathbb{P}[|X - \mathbf{E}[X]| \geq a] = \mathbb{P}[Y \geq a] = \mathbb{P}\left[Y^2 \geq a^2\right].$$

利用马尔可夫不等式，我们立即得到

$$\mathbb{P}\left[Y^2 \geq a^2\right] \leq \frac{\mathbf{E}\left[Y^2\right]}{a^2} = \frac{\mathbf{Var}[X]}{a^2}.$$

□

示例 1.4 再探奖券收集.

将切比雪夫不等式应用于奖券收集问题．沿用之前记号，有

$$\mathbb{P}[Y \geq nH_n + t] \leq \mathbb{P}[|Y - \mathbf{E}[Y]| \geq t] \leq \frac{\mathbf{Var}[Y]}{t^2}.$$

我们之前使用 Y_i 表示已有 $i-1$ 种卡时获得一种新卡所需的开包次数．对不同的 i, j，Y_i 与 Y_j 独立．因此

$$\mathbf{Var}[Y] = \mathbf{Var}\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} \mathbf{Var}[Y_i].$$

对于**两两独立**的随机变量，方差和求和可以交换.

对 $i\in[n]$，$Y_i\sim\mathrm{Geom}\left(\frac{n-i+1}{n}\right)$，故

$$\mathbf{Var}\left[Y_i\right]=\frac{1-\frac{n-i+1}{n}}{\left(\frac{n-i+1}{n}\right)^2}=\frac{(i-1)\cdot n}{(n-i+1)^2}\le\frac{n^2}{(n-i+1)^2}.$$

因此，我们只需要控制住 $\sum_{i=1}^n\frac{1}{(n-i+1)^2}=\sum_{i=1}^n\frac{1}{i^2}$．注意到

$$\sum_{i=1}^n\frac{1}{i^2}\le 1+\int_1^\infty\frac{\mathrm{d}x}{x^2}=2$$

成立．因此 $\mathbf{Var}\left[X\right]\le 2n^2$ 且 $\mathbb{P}\left[X\ge nH_n+t\right]\le\frac{2n^2}{t^2}$．切比雪夫不等式告诉我们，需要抽取超过 $\sqrt{200}n+nH_n$ 包的概率小于 0.01．相比马尔可夫不等式（示例 1.2），这距离事实要接近很多．

在这个问题上，通过切比雪夫不等式得到的界比马尔可夫不等式紧得多——为达到相同置信度，马尔可夫不等式说明需选择 $t=\Theta(n\log n)$．仍取 $n=100$：$\sqrt{200}\times 100+519\approx 1933$ 包即可保证 99% 集齐——对比示例 1.2中马尔可夫不等式给出的约 51900 包，改善了约 27 倍．

我们发现，在切比雪夫不等式的证明中，我们通过控制 $\mathbf{E}\left[f(|X-\mathbf{E}\left[X\right]|)\right]$，其中 $f(x)=x^2$，引入了方差．一个很自然的问题是，如果我们选择别的 $f(x)$，能否得到关于“尾分布”$\mathbb{P}\left[|X-\mathbf{E}\left[X\right]|\ge t\right]$ 的更好的不等式？切比雪夫不等式仅仅用到了随机变量二阶矩的信息．

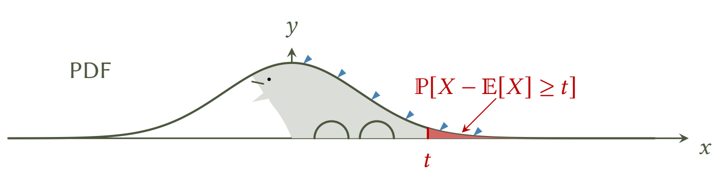

图 1.3: “尾分布”

我们首先要来看一下，在什么情况下，切比雪夫不等式对于尾分布得到的界并不紧．我们考察熟悉的高斯分布 $X\sim\mathcal{N}(0,1)$．使用切比雪夫不等式，我们得到

$$\mathbb{P}\left[X\ge t\right]\le\frac{\mathbf{Var}\left[X\right]}{t^2}=\frac{1}{t^2}.$$

而事实上，我们可以使用高斯分布的概率密度函数直接计算出这个尾分布概率：

$$\mathbb{P}\left[X\ge t\right]=\int_t^\infty\phi(x)\,\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-\frac{x^2}{2}}\,\mathrm{d}x.$$

标准高斯分布的概率密度函数 $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$．

注意到

$$\int_t^\infty e^{-\frac{x^2}{2}}\,\mathrm{d}x\le\int_t^\infty\frac{x}{t}\cdot e^{-\frac{x^2}{2}}\,\mathrm{d}x=t^{-1}e^{-t^2/2},$$

我们有 $\mathbb{P}\left[X\ge t\right]\le\frac{1}{\sqrt{2\pi}t}e^{-\frac{t^2}{2}}$．这说明，实际上，高斯的尾分布是关于 t^2 指数下降的，对于高斯分布来说，切比雪夫不等式得到的界离真相还差距很远．

如果一个分布它长得有点接近高斯，那我们期待能够得到比切比雪夫不等式更加紧的集中不等式．什么样的分布会接近高斯呢？在概率论课中，我们学习

过中心极限定理，我们知道，在一定条件下，很多独立的随机变量之和的分布会收敛到高斯分布．我们接下来两节就讨论这样的随机变量的集中不等式．

切尔诺夫界（Chernoff bound）

我们接着来考虑一系列独立随机变量的和．中心极限定理说，独立随机变量的和的分布（在一定条件下）会收敛到高斯分布，因此，我们接下来的一系列结论可以看成是非渐近版本的中心极限定理，也就是当随机变量的个数 n 是固定的，而不一定趋向于无穷时候的中心极限定理． 最简单的，就是当每个随机变量是伯努利变量的时候，也对应了棣莫弗-拉普拉斯（De Moivre–Laplace）中心极限定理．我们把这个结论称作切尔诺夫界．

棣莫弗-拉普拉斯中心极限定理说的是对于一列 $X_1, X_2, \dots$，如果每一个 $X_i \sim \mathrm{Ber}(p)$，并且是相互独立的，那么

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{D} Y \sim \mathcal{N}(0,1).$$

当然，我们的切尔诺夫界允许每一个 X_i 具有不同的均值 p_i．

定理 1.5 切尔诺夫界.

设 $X_1, \dots, X_n$ 为独立随机变量，其中每个 $X_i \sim \mathrm{Ber}(p_i)$ $(i = 1, 2, \dots, n)$．令 $X = \sum_{i=1}^n X_i$，并记 $\mu := \mathbf{E}[X] = \sum_{i=1}^n p_i$，则有

- 对于任意 $\varepsilon > 0$，$\mathbb{P}[X \geq (1+\varepsilon)\mu] \leq \left(\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right)^{\mu}$.
- 对于 $0 < \varepsilon < 1$，$\mathbb{P}[X \leq (1-\varepsilon)\mu] \leq \left(\frac{e^{-\varepsilon}}{(1-\varepsilon)^{1-\varepsilon}}\right)^{\mu}$.

证明. 我们这儿仅证明第一条（即“上尾界（upper tail bound）”），对于下尾界的证明类似．对任意 $\beta > 0$，有

$$\mathbb{P}[X \geq (1+\varepsilon)\mu] = \mathbb{P}\left[e^{\beta X} \geq e^{\beta(1+\varepsilon)\mu}\right] \leq \frac{\mathbf{E}\left[e^{\beta X}\right]}{e^{\beta(1+\varepsilon)\mu}}.$$

因此需要估计矩生成函数 $\mathbf{E}\left[e^{\beta X}\right]$．由于 $X = \sum_{i=1}^n X_i$ 是独立伯努利变量之和，可得

$$\mathbf{E}\left[e^{\beta X}\right] = \mathbf{E}\left[e^{\beta \sum_{i=1}^n X_i}\right] = \mathbf{E}\left[\prod_{i=1}^n e^{\beta X_i}\right] = \prod_{i=1}^n \mathbf{E}\left[e^{\beta X_i}\right].$$

如果 X 和 Y 独立，那么 $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$．

对 $X_i \sim \mathrm{Ber}(p_i)$，直接计算得：

$$\mathbf{E}\left[e^{\beta X_i}\right] = p_i e^{\beta} + (1-p_i) = 1 + (e^{\beta}-1)p_i \leq \exp\left((e^{\beta}-1)p_i\right)$$

这里又用了 $1 + x \leq e^x$．

因此，

$$\mathbf{E}\left[e^{\beta X}\right] \leq \prod_{i=1}^n \exp\left((e^{\beta}-1)p_i\right) = \exp\left((e^{\beta}-1)\mu\right).$$

从而

$$\mathbb{P}[X \geq (1+\varepsilon)\mu] \leq \left(\frac{\exp(e^{\beta}-1)}{\exp(\beta(1+\varepsilon))}\right)^{\mu}.$$

此式对任意 $\beta > 0$ 成立．选择 β 使分数最小化，令导数为零：

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\left(\frac{\exp(e^{\beta}-1)}{\exp(\beta(1+\varepsilon))}\right) = \exp(e^{\beta}-1-\beta-\beta\varepsilon)\cdot(e^{\beta}-1-\varepsilon) = 0.$$

解得 $\beta = \log(1+\varepsilon)$，代入得：

在本讲义中我们用 log 表示自然对数．

$$\mathbb{P}[X \geq (1+\varepsilon)\mu] \leq \left(\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right)^{\mu}.$$

□

注意到在上述证明中,我们依旧类似切比雪夫不等式的证明,先使用 $\mathbb{P}[X \geq a] = \mathbb{P}[f(X) \geq f(a)]$，然后使用马尔可夫不等式．我们这儿选取的 $f(x) = e^{\beta \cdot x}$ 是指数函数，因此 $\mathbf{E}[f(X)] = \mathbf{E}\left[e^{\beta X}\right]$ 是随机变量 X 的矩生成函数．在一定条件下，矩生成函数唯一确定了随机变量．它可以看成是 X 的所有的 k 阶矩 $\mathbf{E}\left[X^k\right]$ 的加权和，而权重由参数 β 决定．事实上，我们可以选择一个合适的 k，取 $f(x) = x^k$，然后得到类似的结果[1]．但是从证明中可以看到，选用矩生成函数在计算上十分方便．

[1] Terence Tao. **Topics in Random Matrix Theory**. AMS, 2012, §2.1.

事实上，定理 1.5里面给出来的界形式上用起来并不是很方便，也不是很好看出为什么它具有类似于高斯尾分布的下降速度．我们经常使用下面这个推论里给出的界．

推论 1.6.

对任意 $0 < \varepsilon < 1$，有更易用（但稍弱）的形式：

- $\mathbb{P}[X \geq (1+\varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2}{3}\mu\right)$.
- $\mathbb{P}[X \leq (1-\varepsilon)\mu] \leq \exp\left(-\frac{\varepsilon^2}{2}\mu\right)$.

证明. 仅证上尾．需验证当 $0 < \varepsilon < 1$ 时：

$$\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}} \leq \exp\left(-\frac{\varepsilon^2}{3}\right).$$

取对数等价于：

$$\varepsilon - (1+\varepsilon)\log(1+\varepsilon) \leq -\frac{\varepsilon^2}{3}.$$

令 $f(\varepsilon) = \varepsilon - (1+\varepsilon)\log(1+\varepsilon) + \frac{\varepsilon^2}{3}$，其导数满足：

$$f'(\varepsilon) = -\log(1+\varepsilon) + \frac{2}{3}\varepsilon, \quad f''(\varepsilon) = -\frac{1}{1+\varepsilon} + \frac{2}{3}.$$

分析凹凸性可知 $f(\varepsilon) \leq 0$，故原式成立． □

示例 1.7 公平硬币检验.

现有一枚硬币，已知其要么为公平硬币（正面概率 0.5），要么为有偏硬币（正面概率 $0.5+\varepsilon$）．通过抛掷 T 次后计算正面比例 $\hat{p}$ 进行判断是哪种：若 $\hat{p} \leq 0.5 + \varepsilon/2$ 则判为公平，否则判为有偏．现在请问需要投掷多少次硬币，能够保证以至少 $1-\delta$ 的概率正确判断.

我们不妨假设硬币本身是公平的，即 $X_i \sim \mathrm{Ber}(1/2)$．硬币本身是有偏的场

合可以类似说明．那么，总正面数 $X = \sum_{i=1}^{T} X_i$，则误判概率为：

$$\mathbb{P}\left[\hat{p} > 0.5 + \varepsilon/2\right] = \mathbb{P}\left[X > (1+\varepsilon)\mathbf{E}\left[X\right]\right]$$

由切尔诺夫界，当

$$\exp\left(-\frac{\varepsilon^2}{3} \cdot \frac{T}{2}\right) \le \delta$$

时，解得 $T \ge \frac{6}{\varepsilon^2} \log \frac{1}{\delta}$．事实上，在忽略常数倍的意义下，这个界是最优的．

我们将在作业里证明 $\Omega\left(\frac{1}{\varepsilon^2} \log \frac{1}{\delta}\right)$ 次是必要的．

霍夫丁不等式（Hoeffding's inequality）

切尔诺夫界的一个恼人限制是要求每个 X_i 必须是伯努利随机变量．实际上，正如同中心极限定理一样，我们可以大大放松这个要求．正确刻画这类具有类高斯尾分布的概念叫做次高斯分布（*sub-Gaussian distribution*）．我们在本章习题中会讨论这类分布，现在我们先介绍一个特殊情况，也是在本门课中最常用的情况，即有界独立变量的霍夫丁（Hoeffding）不等式．

定理 1.8　霍夫丁不等式.

设 $X_1, \ldots, X_n$ 为独立随机变量，其中每个 X_i 以概率 1 落在区间 $[a_i, b_i]$ 内（$a_i \le b_i$）．令 $X = \sum_{i=1}^{n} X_i$ 则对所有 $t \ge 0$ 有

$$\mathbb{P}\left[|X - \mathbf{E}\left[X\right]| \ge t\right] \le 2\exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

成立．

类似切尔诺夫界的证明，建立此类集中不等式的关键在于对矩生成函数进行有效上界估计．因此，下面的霍夫丁引理将成为证明的主要技术工具．

引理 1.9　霍夫丁引理.

设 X 为满足 $\mathbf{E}\left[X\right] = 0$ 且 $X \in [a, b]$ 的随机变量，则对任意 $\alpha \in \mathbb{R}$ 有

$$\mathbf{E}\left[e^{\alpha X}\right] \le \exp\left(\frac{\alpha^2 (b-a)^2}{8}\right)$$

霍夫丁引理的证明.

我们首先寻找一个线性函数作为 $e^{\alpha x}$ 的上界，以便利用期望的线性性来估计 $\mathbf{E}\left[e^{\alpha X}\right]$．根据指数函数的凸性，对于任何 $x \in [a, b]$，有

$$e^{\alpha x} \le \frac{e^{\alpha b} - e^{\alpha a}}{b - a}(x - a) + e^{\alpha a}.$$

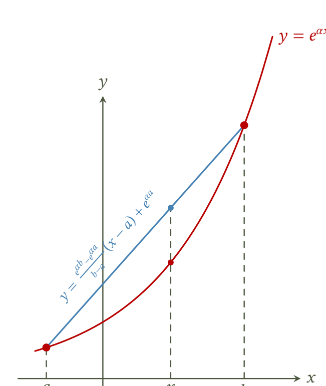

图 1.4: 凸函数 $e^{\alpha x}$ 的线性上界

因此，根据期望的线性性，

$$
\begin{aligned}
\mathbf{E}\left[e^{\alpha X}\right] &\le \frac{e^{\alpha b}-e^{\alpha a}}{b-a}(\mathbf{E}[X]-a)+e^{\alpha a} \\
\triangleright\ \mathbf{E}[X]=0 \qquad &= \frac{-a}{b-a}e^{\alpha b}+\frac{b}{b-a}e^{\alpha a} \\
&= e^{\alpha a}\left(\frac{b}{b-a}-\frac{a}{b-a}e^{\alpha(b-a)}\right) \\
\triangleright\ \theta=-\tfrac{a}{b-a}, t=\alpha(b-a) \qquad &= e^{-\theta t}(1-\theta+\theta e^t) \\
&=:\ e^{g(t)}.
\end{aligned}
$$

其中 $g(t)=-\theta t+\log(1-\theta+\theta e^t)$．根据泰勒定理，对任意实数 t 存在 $\delta\in(0,t)$ 使得

$$
g(t)=g(0)+tg'(0)+\frac{1}{2}g''(\delta)t^2.
$$

计算各阶导数：

$$
\begin{aligned}
&g(0)=0; \\
&g'(0)=-\theta+\left.\frac{\theta e^t}{1-\theta+\theta e^t}\right|_{t=0}=0; \\
&g''(t)=\frac{(1-\theta)\theta e^t}{(1-\theta+\theta e^t)^2}\le\frac{1}{4}.
\end{aligned}
$$

因此 $g(t)\le\frac{1}{8}t^2=\frac{1}{8}\alpha^2(b-a)^2$．即得 $\mathbf{E}\left[e^{\alpha X}\right]\le\exp\left(\frac{\alpha^2(b-a)^2}{8}\right)$． □

借助霍夫丁引理，霍夫丁不等式便很容易证明了．

霍夫丁不等式的证明．不妨设 $\mathbf{E}[X_i]=0$（否则用 $X_i-\mathbf{E}[X_i]$ 代替）．由对称性只需证明 $\mathbb{P}[X\ge t]\le\exp\left(-\frac{2t^2}{\sum(b_i-a_i)^2}\right)$．考虑

$$
\mathbb{P}[X\ge t]\le\frac{\mathbf{E}\left[e^{\alpha X}\right]}{e^{\alpha t}}=\prod_{i=1}^{n}\frac{\mathbf{E}\left[e^{\alpha X_i}\right]}{e^{\alpha t/n}}.
$$

应用霍夫丁引理并取 $\alpha=\frac{4t}{\sum(b_i-a_i)^2}$ 即得结论． □

1.3　独立模型与集中不等式的一些应用

最大负载（max load）问题

投球入箱问题可以被用来建模哈希表中的负载分布．我们考虑有 m 个数据，它们被随机地哈希到 n 个桶中．我们关心的是最大负载，即每个桶中数据的数量的最大值．为了方便讨论，我们接下来假设 $m=n$．

用投球入箱的语言来表述这个问题，我们有 n 个球被随机地投到 n 个箱子中．对于每个 $i\in[n]$，我们定义随机变量 X_i 表示第 i 个桶中的球的数量．我们要计算的是 $\max_{i\in[n]}X_i$ 的值．

在这之前，我们首先计算一下 $\mathbf{E}[X_i]$，即每一个箱子中平均的球的个数．由于每一个箱子都是对称的，我们只需要计算 $\mathbf{E}[X_1]$．我们使用期望的线性性，有

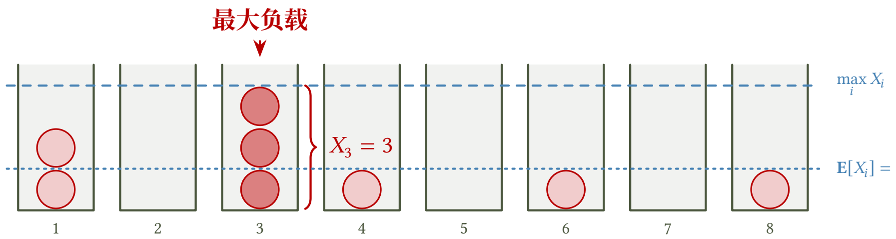

图 1.5: $n = 8$ 时的最大负载问题示意图

$\mathbf{E}[X_1] = \sum_{j=1}^{n} \mathbf{E}[I_j]$,其中 I_j 是一个指示变量,表示第 j 个球是否被投到第 1 个箱子中. 由于每个球被随机地投到 n 个箱子中,因此 $\mathbb{P}[I_j = 1] = \frac{1}{n}$,从而 $\mathbf{E}[I_j] = \frac{1}{n}$. 因此 $\mathbf{E}[X_1] = n \cdot \frac{1}{n} = 1$. n 个球投入 n 个箱子中, 每个箱子平均有 1 个球, 这很合理. 但是, 我们马上可以看到, 最大负载的典型取值就不再是一个常数了, 而是 $\Theta\left(\frac{\log n}{\log\log n}\right)$.

$\Theta\left(\frac{\log n}{\log\log n}\right)$ 是一个看起来有点奇怪的界, 但重要的是, 它是一个当 n 趋向于无穷的时候趋向无穷的量.

我们现在先证明最大负载以高概率为 $O\left(\frac{\log n}{\log\log n}\right)$. 我们在后面的章节会证明其为 $\Omega\left(\frac{\log n}{\log\log n}\right)$. 设 $X = \max_{i\in[n]} X_i$, 我们需要说明存在常数 $c > 0$, 使得 $\mathbb{P}\left[X \geq \frac{c\log n}{\log\log n}\right] = o(1)$. 令 $k = \frac{c\log n}{\log\log n}$. 使用联合界 (union bound), 我们有:

联合界 (union bound) 又称布尔不等式 (Boole's inequality), 指的是对于任意事件 $A_1, A_2, \dots, A_n$, 有 $\mathbb{P}\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \mathbb{P}[A_i]$.

$$\mathbb{P}[X \geq k] = \mathbb{P}[\exists i \in [n], X_i \geq k] \leq \sum_{i=1}^{n} \mathbb{P}[X_i \geq k] = n \cdot \mathbb{P}[X_1 \geq k].$$

注意到 $\{X_1 \geq k\}$ 这个事件意味着在 n 次投球中, 至少有 k 次投入了第 1 个箱子中. 对于 $S \subseteq [n]$, 我们用 A_S 表示事件 $\{$ 对于每一个 $j \in S$, 第 j 个球投到了第 1 个箱子中 $\}$, 因此我们再次使用联合界以及独立性得到:

$$\mathbb{P}[X_1 \geq k] = \mathbb{P}\left[\exists S \subseteq \binom{[n]}{k}, A_S\right] \leq \binom{n}{k} \cdot n^{-k} \leq \left(\frac{e}{k}\right)^k.$$

注意到

$\binom{[n]}{k} := \{S \subseteq [n] : |S| = k\}$.

$$k\log k = \frac{c\log n}{\log\log n}(\log\log n - \log\log\log n + \log c) \sim c\log n.$$

取 $c = 6$, 我们对于充分大的 n, 有

$$\log n + k - k\log k < -\log n.$$

于是, $\mathbb{P}[X \geq k] \leq n \cdot \left(\frac{e}{k}\right)^k < \frac{1}{n} = o(1)$.

流模型计数

假设我们有一台内存极其有限的路由器 (它背后是一个大型网站), 它可以监控访问网站的所有设备 ID. 面对海量的访问数据, 它需要完成一些计算任务, 比如

- 总共有多少个设备访问了?

- 有多少个不同的设备访问了?
- 哪个设备出现的频率最高?

如果我们能存下访问数据本身，那么解决这些计算问题的复杂度并不高．但是由于数据量巨大，我们无法将所有访问记录都存储下来进行离线处理，只能使用有限的内存来处理这些数据流．这便催生了流模型（streaming model）这个计算模型．在流模型中，输入是一个序列 $\sigma = \langle a_1, a_2, \ldots, a_m \rangle$，其中每个 $a_i \in [n]$．顾名思义，"流" 意味着数据是逐个到达的.

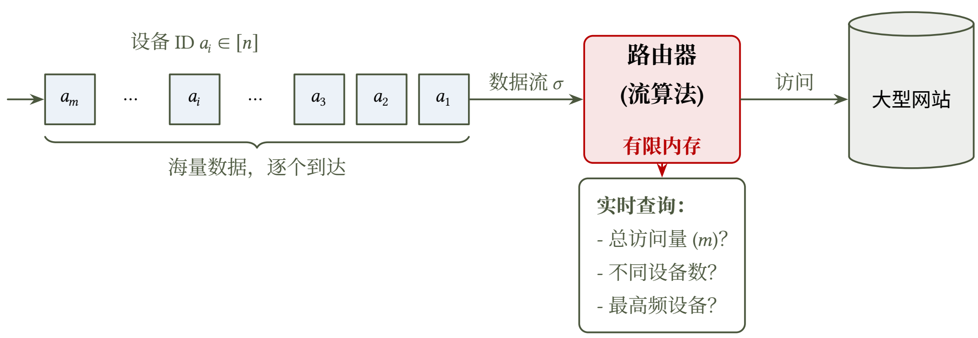

图 1.6: 流模型示意图

我们现在聚焦于最基本的问题：数据流的长度是多少（即 m 的值)? 最简单的想法是维护一个计数器 k．每当一个元素到达时，执行 $k \leftarrow k+1$．不难看出，这需要 $\lceil \log_2 m \rceil$ 比特的内存.

我们能否设计一个更聪明的算法, 仅使用 $o(\log m)$ 的内存? 实际上 $\lceil \log_2 m \rceil - 1$ 比特的内存也是不够的．我们可以使用鸽巢原理来证明这件事情:

> 任何状态数少于 m 的确定性算法，必定在处理某两个不同长度 $i \neq j$ 的前缀时陷入相同的内部状态，从而无法区分出它们.

我们将在练习中请读者严格的证明这个结论.

这个事实说明，如果希望精确的统计 m，那么上面简单维护计数器的方法就是最优的了．但是，如果我们允许近似，情况就很不一样．严格的说，对于给定的误差参数 $\varepsilon > 0$，我们希望算法计算出一个估计值 $\widehat{m}$，使得以高概率有

$$1 - \varepsilon \leq \frac{\widehat{m}}{m} \leq 1 + \varepsilon$$

成立．如果把算法的要求放松成这样，我们能否得到更省内存的算法呢?

Morris 算法（Morris' algorithm）　Robert Morris 在 1978 年提出了一个简单但是巧妙的随机算法.

算法：莫里斯计数算法

输入：数据流序列 $\sigma = \langle a_1, a_2, \ldots, a_m \rangle$，其中 $a_i \in [n]$．

输出：序列长度 m 的估计值．

1. 初始化计数器 $X \leftarrow 0$；
2. 当每个元素到达时：以 2^{-X} 的概率执行 $X \leftarrow X + 1$；
3. 返回估计值 $\widehat{m} = 2^X - 1$．

一个简单的观察是，X 大概是 $\widehat{m}$ 的对数级别大小．因此，如果 $\widehat{m} \approx m$，那么 X 就是一个 $\log m$ 级别大小的数，从而总体占用 $O(\log\log m)$ 比特的内存就可以实现它．这相当于我们之前的精确算法是指数级的提升．我们接下来就是要逐渐在概率的意义下证明 $\widehat{m} \approx m$ 这件事，并且精确的给出实现算法所需要占用的内存．

命题 1.10　莫里斯算法的无偏性.

莫里斯算法的输出 $\widehat{m}$ 是 m 的无偏估计，即 $\mathbf{E}\left[\widehat{m}\right] = m$．

证明. 我们对 m 进行归纳．当 $m = 1$ 时，经过首次处理必定有 $X = 1$，故 $\mathbf{E}\left[\widehat{m}\right] = 2^1 - 1 = 1$，基础情况成立．假设该结论对于 $k < m$ 均成立，这等价于对于 $k < m$，$\mathbf{E}\left[2^{X_k}\right] = k + 1$，其中 X_k 表示处理完第 k 个输入后计数器 X 的值．利用 LOTUS 与归纳假设，有：

LOTUS（Law of the unconscious statistician）指的是对于任意随机变量 X 和（可测）函数 f，有 $\mathbf{E}\left[f(X)\right] = \sum_x f(x) \cdot \mathbb{P}\left[X = x\right]$.

$$
\begin{aligned}
\mathbf{E}\left[\widehat{m}\right] &= \mathbf{E}\left[2^{X_m}\right] - 1 \\
\triangleright \text{ LOTUS} \quad &= \sum_{i=0}^{m} \mathbb{P}\left[X_m = i\right] \cdot 2^i - 1 \\
\triangleright \text{ 算法定义} \quad &= \sum_{i=0}^{m} \left(\mathbb{P}\left[X_{m-1} = i\right] \cdot (1 - 2^{-i}) + \mathbb{P}\left[X_{m-1} = i - 1\right] \cdot 2^{-(i-1)}\right) \cdot 2^i - 1 \\
&= \sum_{i=0}^{m-1} \mathbb{P}\left[X_{m-1} = i\right] \cdot (2^i + 1) - 1 \\
&= \mathbf{E}\left[2^{X_{m-1}}\right].
\end{aligned}
$$

使用归纳假设，我们有 $\mathbf{E}\left[2^{X_{m-1}}\right] = m - 1 + 1 = m$. □

至此，我们证明了其无偏性，也就是说，算法输出的结果，在平均的意义下和正确的结果是一致的．然而，一个实用的随机算法必须额外满足*集中性*，即输出以高概率集中在期望附近，即对于小量 $\varepsilon, \delta > 0$，我们希望满足

$$\mathbb{P}\left[\left|\widehat{m} - m\right| \geq \varepsilon m\right] \leq \delta.$$

这是我们之前学习过的集中不等式的形态，我们可以使用，比如切比雪夫不等式来分析莫里斯算法的集中性．为了应用切比雪夫不等式，我们需要计算 $\widehat{m}$ 的方

差，这需要先计算二阶矩.

引理 1.11.

对于莫里斯算法，有 $\mathbf{E}\left[(2^{X_m})^2\right] = \frac{3}{2}m^2 + \frac{3}{2}m + 1$.

证明. 我们同样对 m 使用归纳法. 当 $m = 1$ 时，$\mathbf{E}\left[(2^{X_1})^2\right] = 4$ ，符合公式. 假设该结论对于所有 $k < m$ 成立：

$$
\begin{aligned}
\mathbf{E}\left[(2^{X_m})^2\right] &= \sum_{i=0}^{m} \mathbb{P}\left[X_m = i\right] \cdot 2^{2i} \\
\triangleright \text{算法定义} \quad &= \sum_{i=0}^{m} \left(\mathbb{P}\left[X_{m-1} = i\right] \cdot (1 - 2^{-i}) + \mathbb{P}\left[X_{m-1} = i - 1\right] \cdot 2^{-(i-1)}\right) \cdot 2^{2i} \\
&= \sum_{i=0}^{m-1} \mathbb{P}\left[X_{m-1} = i\right] \cdot \left(2^{2i} + 3 \cdot 2^{i}\right) \\
&= \mathbf{E}\left[(2^{X_{m-1}})^2\right] + 3\mathbf{E}\left[2^{X_{m-1}}\right].
\end{aligned}
$$

使用归纳假设以及之前证明的无偏性，我们有：

$$
\mathbf{E}\left[(2^{X_m})^2\right] = \left(\frac{3}{2}(m-1)^2 + \frac{3}{2}(m-1) + 1\right) + 3m = \frac{3}{2}m^2 + \frac{3}{2}m + 1.
$$

□

借助这个引理，我们可以直接计算方差：

$$
\mathbf{Var}\left[\widehat{m}\right] = \mathbf{E}\left[\widehat{m}^2\right] - (\mathbf{E}\left[\widehat{m}\right])^2 = \mathbf{E}\left[(2^{X_m} - 1)^2\right] - m^2 \le \frac{m^2}{2}.
$$

应用切比雪夫不等式（定理 1.3），对于任意 $\varepsilon > 0$ ，有：

$$
\mathbb{P}\left[|\widehat{m} - m| \ge \varepsilon m\right] \le \frac{\mathbf{Var}\left[\widehat{m}\right]}{\varepsilon^2 m^2} \le \frac{1}{2\varepsilon^2}. \tag{1.1}
$$

仔细观察式 (1.1), 可以看到这个集中不等式的失败概率上限是 $\frac{1}{2\varepsilon^2}$, 这对于 $\varepsilon < \frac{1}{\sqrt{2}}$ 就已经大于 1 了，从而变得毫无意义. 因此，这是一个很差的界. 究其原因，在于 $\widehat{m}$ 事实上就是一个方差很大的随机变量，我们直接使用它作为算法的输出本身就是很不准确. 为了让算法真正有效，我们必须引入系统性的方法来放大算法的集中度，或者说减小算法输出的方差.

均值技巧（averaging trick） 一个非常直观的让算法变得更稳定的方式是运行多个独立的副本，然后取它们的平均值作为最终输出. 这件事情本质上是减小了输出的方差，我们称之为*均值技巧*. 现在，我们来严格的分析它.

我们设计一个新算法：*独立且并行*地运行莫里斯算法 t 次，记输出分别为 $\widehat{m}_1, \ldots, \widehat{m}_t$. 最终输出取均值：

$$
\widehat{m}^* := \frac{1}{t} \sum_{i=1}^{t} \widehat{m}_i.
$$

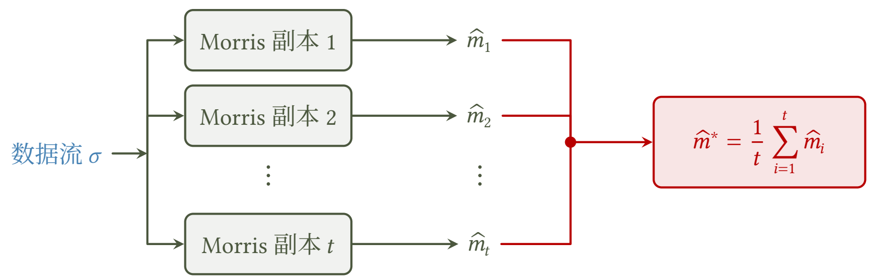

图 1.7: 均值技巧：独立并行运行 t 次，取均值

由独立性，$\mathbf{Var}\left[\widehat{m}^*\right] = \frac{1}{t}\mathbf{Var}\left[\widehat{m}_1\right]$．再次应用切比雪夫不等式可得：

$$\mathbb{P}\left[|\widehat{m}^* - m| \geq \varepsilon m\right] \leq \frac{1}{t \cdot 2\varepsilon^2}. \tag{1.2}$$

在这里的讨论中，我们说算法“成功”是指输出满足 $|\widehat{m}^* - m| < \varepsilon m$ 的事件，否则叫做算法“失败”.

为了将失败概率控制在 δ 以下,我们只需取 $t \geq \frac{1}{2\varepsilon^2\delta}$．这个新算法消耗了 $O\left(\frac{\log\log m}{\varepsilon^2\delta}\right)$ 比特的内存．可以看到，算法的性能（内存消耗）体现了对于失败率 δ 和精度 ε 的权衡.

式 (1.2)的界是很不错的，其导出来的算法对于失败率 δ 的依赖是关于 $1/\delta$ 的线性函数．回忆我们在示例 1.7中，对于公平硬币检验问题，我们可以使用切尔诺夫界得到一个投掷次数关于错误率 δ 的 $\log(1/\delta)$ 级别的依赖．虽然硬币检测问题和我们这里的问题表面上看起来很不一样，但实际上它们具有类似的结构．我们有一个通用的技巧把对于失败率的依赖改进到 $\log(1/\delta)$ 级别，使用的工具依然是切尔诺夫界.

严格来说，每个副本的计数器 X_i 占用的比特数是随机的：存储 X_i 需要 $\lceil\log_2(X_i+1)\rceil$ 比特，只有当 $X_i = O(\log m)$ 时才是 $O(\log\log m)$ 比特．但 X_i 的尾概率随 X_i 指数衰减，对 t 个副本取联合界后，总内存以高概率不超过所述的量．这一额外的失败概率可以并入整体的 δ 中．我们这里不深究这个细节.

中位数技巧（median-of-means trick）　我们在均值技巧中固定 $t = \frac{3}{2\varepsilon^2}$．由前文分析，此时每次运行的失败概率 $\delta \leq 1/3$．我们将这个均值算法作为一个整体，独立并行地运行 s 次，得到估计集合 $\{\widehat{m}_1^*, \ldots, \widehat{m}_s^*\}$．对于任意 $i \in [s]$，均满足 $\mathbb{P}\left[|\widehat{m}_i^* - m| \geq \varepsilon m\right] \leq 1/3$．最终，我们输出这 s 个结果的中位数 $\widehat{M}$.

对于每个 $i \in [s]$，定义指示变量 $Y_i = \mathbb{1}\left[|\widehat{m}_i^* - m| < \varepsilon m\right]$．令 $Y = \sum_{i=1}^{s} Y_i$，显然 $\mathbf{E}[Y] \geq \frac{2}{3}s$．

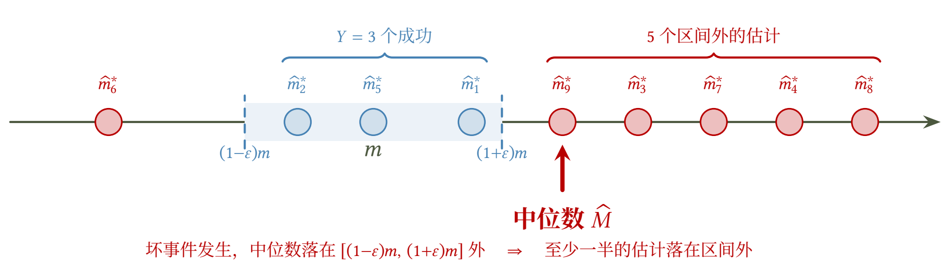

图 1.8: 中位数技巧的 “坏事件”：过半数副本失败时，中位数落在误差区间外

如果最终的中位数 $\widehat{M}$ 是一个 “坏” 估计，必然意味着有一半以上的子算法给出了坏估计．换言之，成功的次数 $Y \leq s/2$．利用推论 1.6 中的下尾切尔诺夫界

（取 $\epsilon' = 1/4$）：

$$\mathbb{P}[Y \le s/2] \le \mathbb{P}[Y \le (1-1/4)\mathbf{E}[Y]] \le \exp\left(-\frac{(1/4)^2}{2}\cdot\frac{2}{3}s\right) = \exp(-s/48).$$

为了让整体失败概率不超过 δ，我们只需令 $\exp(-s/48) \le \delta$．取 $s = 48\log\frac{1}{\delta}$ 即可．

总结起来，结合均值与中位数技巧的最终算法，对于任何 $\varepsilon, \delta \in (0,1)$，以至少 $1-\delta$ 的概率输出一个满足 $|\widehat{M} - m| < \varepsilon m$ 的估计值，并且使用的总内存为 $O\left(\frac{1}{\varepsilon^2}\log\frac{1}{\delta}\log\log m\right)$ 比特.

比较重要的是，我们使用的均值技巧和中位数技巧，本身与莫里斯算法没有什么关系．它们是一种通用的提高随机算法性能的方式.

同前文的讨论，这里的内存上界也是以高概率成立的：需要对所有 $s\cdot t$ 个 Morris 副本的计数器取联合界.

多臂老虎机问题（multi-armed bandit）

在本节中，我们讨论在线优化中的一个经典模型：多臂老虎机问题（multi-armed bandit problem）．我们通过这个问题展示一下集中不等式在算法设计和分析中的作用．假设有一个 n 臂老虎机,每一个臂拉一下会给出一个位于 $[0,1]$ 范围内的随机奖励值．为了方便，我们假设第 i 个臂返回的奖励服从分布 $\mathrm{Ber}(\mu_i)$．我们进一步假设 $\mu_1 > \mu_2 \ge \cdots \ge \mu_n$ 成立．我们现在的可以拉动老虎机 T 轮（每一轮选一个臂），目标是使奖励期望最大化．如果我们已知 $\mu_1, \ldots, \mu_k$，最优策略是始终选择臂 1，此时期望奖励为 $T\mu_1$．然而，当我们不知道每个臂的奖励分布及顺序时，就需要设计一种策略来首先探索老虎机.

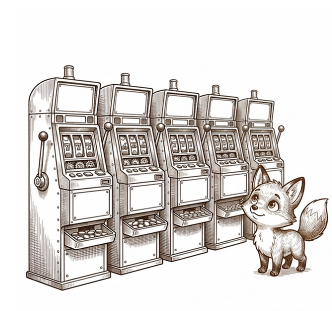

图 1.9: 多臂老虎机问题

记 a_t 为第 t 轮拉动的臂，因此第 t 轮的奖励为 $X_t \sim \mathrm{Ber}(\mu_{a_t})$．一种衡量策略的好坏的方式是使用懊悔（*regret*）．懊悔定义为总轮次 T 中始终选择最佳臂 1 所能获得的期望奖励与策略实际期望奖励之间的差值，即未总是选择最佳臂所带来的代价：

$$R(T) := T\mu_1 - \mathbf{E}\left[\sum_{t=1}^{T} X_t\right] \ge 0.$$

在上述表达式中，期望 $\mathbf{E}[\cdot]$ 的随机性通常来源于两个方面：奖励分布 $\mathrm{Ber}(\mu_i)$ 的随机性以及策略本身的随机性．对于每个 $i \in [n]$，我们定义 $\Delta_i := \mu_1 - \mu_i \ge 0$，表示第 i 臂的奖励均值与最优臂的奖励均值之间的差距．如果我们的策略是平均的拉每一个臂，那么懊悔为 $R(T) = \frac{\sum_{i=1}^{n}\Delta_i}{n}\cdot T$．我们可以选择 Δ_i 使得这是一个关于 T 线性增长的，因此，不是一个好的策略．接下来，我们介绍一种名为“探索然后确定”（explore-then-commit, ETC）的算法，该算法可以实现次线性懊悔.

探索然后确定（explore-then-commit, ETC）算法　为了使懊悔最小化，策略应尽快识别出最佳臂．最直接的方法是给每个臂尝试一定次数，然后选择奖励的**经验均值**（empirical mean）最大的臂．ETC 算法便是实现了这一思想：每个臂 i 被拉动 L 次（总共进行 nL 次探索），并计算 $\hat{\mu}_i$（在 L 次尝试中获得的平均奖励）

. 之后，始终选择 $\hat{\mu}_i$ 最大的臂．其懊悔可以写为：

$$\begin{aligned} R(T) &= L\sum_{i=1}^{n}\Delta_i + \sum_{i=2}^{n}\Delta_i \cdot \sum_{t=nL+1}^{T}\mathbb{P}\left[\hat{\mu}_i > \max_{j\neq i}\hat{\mu}_j\right] \\ &= L\sum_{i=1}^{n}\Delta_i + \sum_{i=2}^{n}\Delta_i \cdot (T-nL)\mathbb{P}\left[\hat{\mu}_i > \max_{j\neq i}\hat{\mu}_j\right]. \end{aligned}$$

当 $i \neq 1$ 时：

$$\mathbb{P}\left[\hat{\mu}_i > \max_{j\neq i}\hat{\mu}_j\right] \leq \mathbb{P}\left[\hat{\mu}_i > \hat{\mu}_1\right].$$

我们用集中不等式来界定上述概率．为此，令 X_j 为探索阶段中第 j 次拉动臂 i 时的奖励, Y_j 为探索阶段中第 j 次拉动臂 1 时的奖励．令 $Z_j = X_j - Y_j \in [-1, 1]$，则 $\mathbf{E}\left[Z_j\right] = -\Delta_i \leq 0$．令 $Z = \sum_{j=1}^{L} Z_j$，则 $\mathbf{E}[Z] = -L\Delta_i$．

根据霍夫丁不等式：

$$\mathbb{P}\left[\hat{\mu}_i > \hat{\mu}_1\right] = \mathbb{P}[Z > 0] = \mathbb{P}\left[Z - \mathbf{E}[Z] > L\Delta_i\right] \leq \exp\left(-\frac{2(L\Delta_i)^2}{\sum_{j=1}^{L}2^2}\right) = \exp\left(-\frac{L\Delta_i^2}{2}\right).$$

因此有：

$$\begin{aligned} R(T) &\leq L\sum_{i=1}^{n}\Delta_i + (T-nL)\sum_{i=2}^{n}\Delta_i\exp\left(-\frac{L\Delta_i^2}{2}\right) \\ \triangleright\ \Delta_i \leq 1 \qquad &\leq \sum_{i=1}^{n}\left(L\Delta_i + T\Delta_i\exp\left(-\frac{L\Delta_i^2}{2}\right)\right). \end{aligned}$$

接下来定义：

$$g(L, \Delta_i) := L + T\Delta_i\exp\left(-\frac{L\Delta_i^2}{2}\right).$$

我们希望找到 L 使得 $R(T)$ 的上界最小化，即 $\min_L \max_{\Delta_i} R(T)$．首先计算 $\max_{\Delta_i} g(L, \Delta_i)$：

$$\frac{\partial}{\partial\Delta_i}g(L, \Delta_i) = T(1 - L\Delta_i^2)\exp\left(-\frac{L\Delta_i^2}{2}\right).$$

当 $0 \leq \Delta_i < \frac{1}{\sqrt{L}}$ 时，$\frac{\partial}{\partial\Delta_i}g(L, \Delta_i) > 0$；当 $1 \geq \Delta_i > \frac{1}{\sqrt{L}}$ 时，$\frac{\partial}{\partial\Delta_i}g(L, \Delta_i) < 0$．

因此，对于所有 $L > 1$：

$$g(L, \Delta_i) \leq g\left(L, \frac{1}{\sqrt{L}}\right) = L + \frac{Te^{-1/2}}{\sqrt{L}}. \tag{1.3}$$

通过令 $L = \Theta(T^{\frac{2}{3}})$，最终可以得到

$$R(T) \leq \sum_{i=1}^{n}\left(L + \frac{Te^{-1/2}}{\sqrt{L}}\right) = \Theta(nT^{\frac{2}{3}}).$$

更好的懊悔　ETC 算法实现了 $\Theta(T^{\frac{2}{3}})$ 的次线性懊悔，这是不错的结果，但仍然不是最优的．其主要缺点在于，在探索阶段，该算法对所有臂一视同仁，无论已经获得的奖励如何，每个臂都固定被尝试 L 次．我们可以更加精妙的应用集中不等式，构造一个称之为上置信界（*upper confidence bound, UCB*）的算法．它可以达到 $\widetilde{O}(T^{\frac{1}{2}})$[1] 的懊悔，我们将在本章习题中讨论这个算法.

[1] $\widetilde{O}(\cdot)$ 忽略了关于 T 的 log 项