阅读材料：耦合（Coupling）

宗语轩

2024 年 2 月 11 日

我们先看下面一个例子:

例 0.1. 有两个玩家 A, B 投掷两枚两面硬币, 玩家 A 投掷到正面的概率是 0.75, 玩家 B 投掷到正面的概率是 0.5. 现两个玩家独立投掷硬币各 100 枚, 记 N_A, N_B 分别表示玩家 A, B 投掷到正面的个数, 证明: 对任意固定的 $k(0<k<100)$, 均有

$$\mathbb{P}(N_A \geq k) \geq \mathbb{P}(N_B \geq k)$$

问题与思考: 我们当然可以显示地写出两者的概率表达式:

$$\mathbb{P}(N_A \geq k) = \sum_{n=k}^{100} \binom{100}{n} 0.75^n 0.25^{100-n},$$
$$\mathbb{P}(N_B \geq k) = \sum_{n=k}^{100} \binom{100}{n} 0.5^n 0.5^{100-n} = \sum_{n=k}^{100} \binom{100}{n} 0.5^{100}.$$

但是这样的话我们不能比较直观地得到待证命题 (读者可以进行尝试), 而命题直观上看又是对的, 那么我们需要怎么解决呢？这里我们要引入一个耦合 (Coupling) 的技巧:

注意到随机变量的定义 $X : \Omega \to \mathbb{R}$ 是 Borel 实值可测函数, 两个随机变量的分布相同并不代表随机变量相同, 也即不代表他们的样本空间相同 (反例请读者自行思考). 换句话说, 同一个的" 输出端" 对应的" 输入端" 不一定唯一. 借助这个想法, 我们通过匹配合适的样本空间来达到上述目的:

我们进行如下机制: 考察两个玩家分别掷一次硬币的情形. 先让玩家 B 投掷, 若结果为正面, 则令玩家 A 结果也为正面; 若结果为反面, 则令玩家 A 结果为正面的概率为 0.5.

按照上述机制, 我们可以计算玩家 A 投掷到正面的概率亦为 0.75. 我们记 $X_{A,i}$ 表示玩家 A 在第 i 次投掷硬币正面的个数 (结果只可能为 1 或 0), $X_{B,i}$ 表示玩家 B 在第 i 次投掷硬币正面的个数. 由上述机制可知: 对任意 $1 \leq i \leq 100$, 均有 $X_{A,i} \geq X_{B,i}$, 因此

$$\sum_{i=1}^{100} X_{A,i} \geq \sum_{i=1}^{100} X_{B,i} \Rightarrow \mathbb{P}\left(\sum_{i=1}^{100} X_{A,i} \geq k\right) \geq \mathbb{P}\left(\sum_{i=1}^{100} X_{B,i} \geq k\right).$$

而 N_A 和 $\sum_{i=1}^{100} X_{A,i}$ 两者分布相同, N_B 和 $\sum_{i=1}^{100} X_{B,i}$ 两者分布相同. 因此

$$\mathbb{P}(N_A \geq k) \geq \mathbb{P}(N_B \geq k).$$

更一般的证明见下:

[0]个人主页: http://home.ustc.edu.cn/~zyx240014/.
发现错误欢迎联系: zyx240014@mail.ustc.edu.cn.

证明. 设随机变量 $Z \sim U[0,1]$, 即 $[0,1]$ 上的均匀分布, 对任意 $1 \le i \le 100$ 引入随机变量 $Y_{A,i}, Y_{B,i}$, 使其相互独立且满足

$$Y_{A,i} = \begin{cases} 1, & 0 \le Z \le 0.75, \\ 0, & 0.75 < Z \le 1. \end{cases}, \quad Y_{B,i} = \begin{cases} 1, & 0 \le Z \le 0.5, \\ 0, & 0.5 < Z \le 1. \end{cases}$$

因此对任意 $1 \le i \le 100$, 均有 $Y_{A,i} \ge Y_{B,i}$, 故

$$\sum_{i=1}^{100} Y_{A,i} \ge \sum_{i=1}^{100} Y_{B,i} \Rightarrow \mathbb{P}\left(\sum_{i=1}^{100} Y_{A,i} \ge k\right) \ge \mathbb{P}\left(\sum_{i=1}^{100} Y_{B,i} \ge k\right).$$

而 N_A 和 $\sum\limits_{i=1}^{100} Y_{A,i}$ 两者分布相同, N_B 和 $\sum\limits_{i=1}^{100} Y_{B,i}$ 两者分布相同. 因此

$$\mathbb{P}(N_A \ge k) \ge \mathbb{P}(N_B \ge k).$$

$\square$

上述例子帮助我们对耦合这个概念有了一个直观地了解, 现在我们给出它的严格定义:

定义 0.1. 概率空间 $\mathcal{X}$ 上两个概率测度 μ, v 的**耦合 (coupling)**, 是指 $\mathcal{X} \times \mathcal{X}$ 上的联合分布 π. 离散性下若随机向量 $(X,Y) \sim \pi$, 则对任意 x, y, 满足

$$\mathbb{P}(X = x) = \mu(x), \qquad \mathbb{P}(Y = y) = \nu(x)$$

注. 联合分布可推出边缘分布, 反之不成立.

考虑有限离散型的情形, μ, ν 的耦合可被矩阵 $q \in \mathbb{R}^{|\mu| \times |\nu|}$ 刻画, 其中 $q(x,y) = \mathbb{P}(X = x, Y = y)$ 满足

$$\sum_y q(x,y) = \mu(x), \qquad \sum_x q(x,y) = \nu(y).$$

注. 耦合的存在性: 对任意 μ, ν, 均存在耦合 $q(x,y) = \mu(x)\nu(y)$, 此时 X 和 Y 独立.

例 0.2. 耦合伯努利分布 $B(0.75)$ 和 $B(0.5)$.

解. 记耦合的矩阵

$$q = \begin{pmatrix} q(0,0) & q(0,1) \\ q(1,0) & q(1,1) \end{pmatrix} := \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

由定义知,

$$\begin{cases} a + b = 0.25, \\ c + d = 0.75, \\ a + c = 0.5, \\ b + d = 0.5, \end{cases} \Rightarrow q = \begin{pmatrix} a & 0.25 - a \\ 0.5 - a & 0.25 + a \end{pmatrix}, 0 \le a \le 0.25.$$

$\square$

注. 例 0.1 中用到的耦合方法即为上述取 $a = 0.25$ 的情形.

例 0.3. 耦合伯努利分布 $B(0.5)$ 和 $B(0.5)$.

解. 由定义知,

$$q = \begin{pmatrix} a & 0.5-a \\ 0.5-a & a \end{pmatrix}, 0 \leq a \leq 0.5.$$

- **Case 1:** $a = 0.5$, $q = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$, identical coupling.
- **Case 2:** $a = 0$, $q = \begin{pmatrix} 0 & 0.5 \\ 0.5 & 0 \end{pmatrix}$, negative coupling.
- **Case 3:** $a = 0.25$, $q = \begin{pmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{pmatrix}$, independent coupling.

□

习题. 对 $0 < p \leq q < 1$, 耦合伯努利分布 $B(p)$ 和 $B(q)$, 并给出当矩阵 q 存在元素 0 时的耦合方法.

提示. 参考**例 0.1** 中的证明方法.

参考资料

[1] 王冠扬, 马尔科夫链里的耦合方法, 中科大研究生创新计划高水平和华罗庚科技英才班前沿短课程.
https://wvpn.ustc.edu.cn/http/77726476706e69737468656265737421e7fb4a8869257b447d468ca88d1
video/detail_5722_30386.htm

[2] Markov chains and mixing times, David A. Levin and Yuval Peres, American Mathematical Society, 2017.