Lecture Notes for Probability Theory – Class 4

Zhanghao Wu

问题

1 问题 (矩阵乘法) 给定三个矩阵 $A, B, C \in \mathbb{F}_2^{n\times n}$，通过下方算法检验，是否有 $AB = C$ 成立，试求算法犯错的概率 $P(wrong)$，是否有 $P(wrong) \leq \frac{1}{2}$

- 随机选取一个向量 $r \in \mathbb{F}_2^n$
- 检验 ABr 与 Cr 是否相等
- 若不相等，说明结论不成立，如果相等，则可能犯错

评注 该问题的关键在于 $P(wrong)$ 具体是什么意思，下面是两种不同的看法：

1. $P(wrong)$ 是指在某一种情况下，如 AB 与 C 不相等时，给出相反的结果，即 $ABr = Cr$，的概率

   在这种看法下，分为两种可能，

   (a) 当 $AB = C$ 时，$P(wrong) = P(ABr \neq Cr) = 0$；

   (b) 当 $AB \neq C$ 时，$P(wrong) = P(ABr = Cr) = \frac{|ker(AB-C)|}{\mathbb{F}_2^n} \leq \frac{1}{2}$（r 随机选取时，落在 $ker(AB - C)$ 中的概率）

   因此，在这种看法下，犯错的概率是小于 $\frac{1}{2}$ 的，那么在多次验证后，错误的概率会指数级迅速减小

2. $P(wrong)$ 是指在 $ABr = Cr$ 的情况下 $AB \neq C$ 的条件概率，亦即 $P(wrong) = P(AB \neq C|ABr = Cr)$ 那么根据贝叶斯原理可以通过概率传递的方式，计算出第一次错误的后验概率，再将其作为先验概率计算第二次验证的后验概率……

两种看法都有合理性，也说明，模型不同会导致不同的结果

2 问题 (信封与钱) 现有两个信封，一个信封中装有 200 元，另一个信封中只有 100 元，此时拿起一个信封，试问更换信封能否使使拿到 200 元的概率更大？下面的说法是否存在问题？

假设，拿起的信封中钱的数额为 D 元，那么更换信封时，有一半的概率使 D 变为 $2D$,另有一半的概率是 D 变为 $\frac{D}{2}$,因此,更换信封收益的期望 $E=\frac{1}{2}\times D-\frac{1}{2}\times\frac{D}{2}=\frac{D}{4}>0$，更换信封更划算。

评注 ~~（笔者注）这个论断所做的序贯树并不正确，由于所谓两个等概率的情况所对应的D并不相同，也就是说在拿起的信封的中装有 *200* 元和 *100* 元时，更换信封的收益并不是等概率的，序贯树的修改如下图1所示，而右侧计算出的收益期望为$E =$~~ $\frac{1}{2}\times(-100)+\frac{1}{2}\times 100=0$。

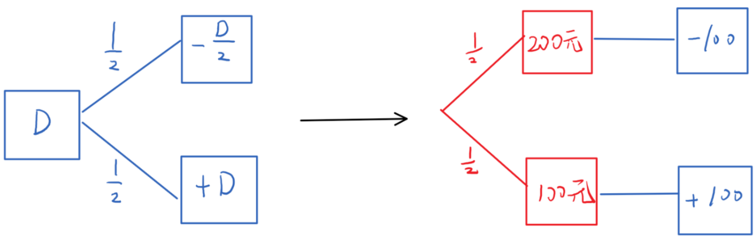

图 1: 收益序贯树

感谢，李沐阳同学对此处的指正，这一段是我的个人观点，我自认为比较直观，但在笔记中出现有些不妥，因此，关于这道问题，详情请参考《概率导论》第 94 页，例 2.18

随机变量与期望

在之前的我们定义了，样本空间上的概率函数（不严格），有 (Ω, P)，下面定义有限或可数无限的样本空间上的随机变量（不严格）

3 定义 (随机变量 (random variable))

定义在样本空间 Ω 上的实值函数 $\mathbb{R}^\Omega$，称为随机变量 (random variable)。

4 例 对于某个事件 $A \subseteq \Omega \to 1_A$ 或 $X_A = \begin{cases} 0, & x \notin A \\ 1, & x \in A \end{cases} \in \mathbb{R}^\Omega$，此处的 1_A 或 X_A 即是一种特殊的随机变量（伯努利随机变量 笔者注）。这个随机变量还有一个很好的性质 $P(A) = E(1_A)$

5 定义 (期望 (不严格))

$E[X] = \sum_{x_i} P(X = x_i)$（1657 年，对无穷的情况可能存在不收敛的情况）

6 推论 (期望的性质)

期望满足如下一些性质：

- $E[f(X)] = \sum_{y_i} P(f(x) = y_i) = \sum_{x_i} f(x_i)P(X = x_i)$
  (由于 $P(f(x) = y_i) = \sum_{x \in f^{-1}(y_i)} P(X = x)$)
- $E[X + X'] = E[X] + E[X']$
- $E[aX + bX'] = aE[X] + bE[X']$
- 全概率公式：$P(A) = \sum P(B_i)P(A|B_i) \to$
  全期望公式 $E[X] = E(E[X|Y])$
  (i.e$\sum_y p_Y(y)E[X|Y = y]$，随机变量的函数仍是随机变量 笔者注)

7 问题 (朋友更受欢迎)

表述："朋友的朋友总比你的朋友多"

更严格的表述:任意一张图 $G = (V, E_d)$, 对于 $\mathbf{x}_0 \in V$,及与 x_0 x_1, 必有 $E[deg(x_0)] \leq E[deg(x_1)]$

证明 x_1 度数的期望为：

$$\begin{aligned} E[deg(x_1)] &= \frac{1}{|V|} \sum_{x_0} \frac{1}{deg(x_0)} \sum_{x_0 x_1 \in E} deg(x_1) \\ &= \frac{1}{|V|} \sum_{x_0 x_1 \in E} (\frac{deg(x_0)}{deg(x_1)} + \frac{deg(x_1)}{deg(x_0)}) \\ &\geq \frac{2|E|}{|V|} = E[deg(x_0)] \end{aligned}$$

只有在所有点的度数相同时，才能取到等号

评注 此题为 x_0 相邻的半径为 1 的球，如果球的半径增大又会有怎样的结论呢？（也就是朋友的朋友的朋友……会如何）

8 定义 (方差与标准差 (variance, standard variance))

- 方差：$var(X) = E[(X - E[X])^2]$
- 标准差：$\sigma_X = \sqrt{var(X)}$

9 推论

- $var(X) = E[(X - E[X])^2] = E[X^2] + (E[X])^2 - 2E[X]E[X] = E[X^2] - E[X]^2$
- $var(aX + b) = a^2 var(X)$

10 例 (期望不等式)

1. 马氏不等式 (Markov Inequality)

   $P(|X - E[X]| \geq a) \leq \frac{E[E[X]]}{a}, \forall a > 0, X$ 为非负随机变量

2. 切比雪夫不等式 (Chebyshev's Inequality)

   $P(|X - E[X]| \geq a) \leq \frac{E[(X-E[X])^2]}{a^2})$

3. 柯西-施瓦茨不等式 (Cauchy-Schwarz inequality)

   $|E[XY]|^2 \leq E[X^2]E[Y^2]$

   ($var(X) \geq 0 \Rightarrow E[X^2] \geq E[X]^2$，即为柯西不等式的特例)

证明

1. 马氏不等式 (Markov Inequality)

   $E[X] = \sum_{a_i} a_i P(X = a_i) \geq \sum_{a_i \geq a} a_i P(X = a_i) \geq aP(X \geq a)$

2. 切比雪夫不等式 (Chebyshev's Inequality)

   令马氏不等式中的 $X = |X - E[X]|, a = a^2$，则 $P(|X - E[X]| \geq a) = P(|X - E[X]|^2 \geq a^2) \leq \frac{E[(X-E[X])^2]}{a^2})$

3. 柯西-施瓦茨不等式 (Cauchy-Schwarz inequality)

   $0 \leq E[(X-tY)^2] = E[X^2] - 2tE[XY] + t^2 E[Y^2], \forall t, \Delta = 4E[XY]^2 - 4E[X^2]E[Y^2] \leq 0$

11 例 X, Y 独立同分布，$E[X|X+Y] = ?$

证明 $2E[X|X+Y]=E[X|X+Y]+E[Y|X+Y]=E[X+Y|X+Y]=X+Y$

12 问题 (向量期望长度) 令 $v_1,\ldots,v_n\in S^{n-1}$ (即 v_i 落在 $n-1$ 维球面上，或 $v_i\in\mathbb{R}^n$ and $|v_i|=1$)，求 $|\sum\varepsilon_i v_i|$ 的性质，其中 $\varepsilon_i\in\{\pm1\}$

解答 随机变量 $X=|\sum\varepsilon_i v_i|$，$E[X^2]=E[\sum_{i\neq j}\varepsilon_i\varepsilon_j v_i v_j]+E[\sum_i\varepsilon^2 v_i v_i]$

由于 $\varepsilon_i,\varepsilon_j$ 独立，$\Rightarrow E[\varepsilon_i\varepsilon_j]=E[\varepsilon_i]E[\varepsilon_j]=0$

因此，$E[X^2]=n,\Rightarrow\begin{cases}\exists\varepsilon_i,\ \text{s.t.}|\sum\varepsilon_i v_i|\geq\sqrt{n}\\ \exists\varepsilon_i,\ \text{s.t.}|\sum\varepsilon_i v_i|\leq\sqrt{n}\end{cases}$

13 定义 (独立性)

$X_1,\ldots,X_k$ 为随机变量，满足

$$\forall x_1,\ldots,x_k,P(X_1=x_1,\ldots,X_k=x_k)=\prod_i P(X_i=x_i)$$

，则称这 k 个随机变量相互独立

思考题

14 问题 $A_1,\ldots,A_k$ 为 k 个事件,满足 $\forall S\subset[k],\prod_{i\in S}P(A_i)=P(\bigcap_{i\in S}A_i)$,试证明 $1_{A_1},\ldots,1_{A_k}$ 独立

15 问题 $(X_i)_{i\in I}$ 为独立的随机变量。$J\subset I,K\subset,J\bigcap K=\varnothing$,对于随机变量 $(X_i)_{i\in J},(X_i)_{i\in K}$，及函数 $f\in\mathbb{R}^J$, $g\in\mathbb{R}^K$，试问随机变量 $Y=f((X_j)_{j\in J})$ 与 $Z=g((X_k)_{k\in K})$ 是否独立

16 问题 如果 X 和 Y 是相互独立的随机变量，那么 $E[XY]=E[X]E[Y]$, $var(X+Y)=var(X)+var(Y)$

17 问题 若 $P(X_i=1)=P(x_i=-1)=\frac{1}{2}$, $i=1,2,,\ldots$

1. 若 $Y_i=X_iX_{i+1}X_{i+2}$，问 $Y_1,Y_2,\ldots$ 是否相互独立？

2. 若 $S\subset\{1,2,\ldots\}|S|<\infty$, $Y_S=\prod_{i\in S}X_i$ 是否相互独立