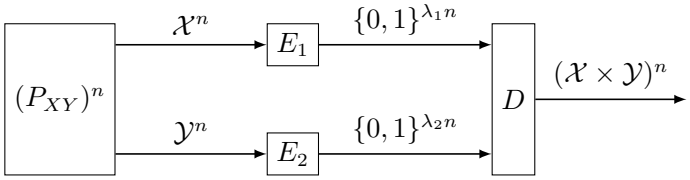
信息论

请在 12 月 12 日课前提交纸质作业.

1. (10 分) 本题中，信息熵 H 使用 2 作为底数. 令 P_{XY} 为一个支撑有限 $\mathcal{X}\times\mathcal{Y}$ 上的联合分布. 令常数 λ_1,λ_2 满足 $\lambda_1 > H(X|Y)$，$\lambda_2 > H(Y|X)$，$\lambda_1+\lambda_2 > H(X,Y)$.



   (1) 请构造两个压缩函数 $E_1 : \mathcal{X}^n \to \{0,1\}^{\lfloor\lambda_1 n\rfloor}$，$E_2 : \mathcal{X}^n \to \{0,1\}^{\lfloor\lambda_2 n\rfloor}$，和一个解压缩函数 $D : \{0,1\}^{\lfloor\lambda_1 n\rfloor+\lfloor\lambda_2 n\rfloor} \to (\mathcal{X}\times\mathcal{Y})^n$，并证明

$$\Pr_{((X_1,Y_1),\dots,(X_n,Y_n))\sim(P_{XY})^n}\Big[D(E_1(X_1,\dots,X_n),E_2(Y_1,\dots,Y_n)) \neq ((X_1,Y_1),\dots,(X_n,Y_n))\Big] \le 2^{-\Theta(n)}.$$

   (2) 如果改变参数，满足如下条件之一：a) $\lambda_1 < H(X|Y)$，b) $\lambda_2 < H(Y|X)$，c) $\lambda_1+\lambda_2 < H(X,Y)$. 说明这时不能构造满足前一问要求的压缩函数和解压缩函数.

2. (8 分) (n,m,k)-纠错码是一对映射 $E : \{0,1\}^n \to \{0,1\}^m$，$D : \{0,1\}^m \to \{0,1\}^n$ 满足

$$\forall x \in \{0,1\}^n, \forall c \in \{0,1\}^m, \Delta(c,E(x)) \le k \implies D(c) = x.$$

   这里 Δ 表示汉明距离（Hamming distance）.

   (1) 证明存在常数 α. 当 $m > 4k$ 且 $m \ge 2n + \alpha k\ln(m/2k)$ 时，存在 (n,m,k)-纠错码.
   提示：可以使用不等式 $\frac{\log p\cdot\log(1-p)}{\log e} \le h(p) \le \frac{\log p\cdot\log(1-p)}{\log 2}$. 其中 $h(p)$ 表示 $\mathrm{Bern}(p)$ 的熵.

   (2) 证明存在常数 α. 当 $m > 4k$ 且 $m \ge n + \alpha k\ln(m/2k)$ 时，存在 (n,m,k)-纠错码.

3. (10 分) 在有限空间 Ω 上有两个分布 P,Q. 区分器 $\mathcal{D}$ 是一个输入域为 Ω, 输出域为 $\{0,1\}$ 的算法 (更准确地说, 是 kernel). 我们希望让伪阳性概率 $\varepsilon_{\mathrm{FP}}$ 和伪阴性概率 $\varepsilon_{\mathrm{FN}}$ 尽量小.

$$\varepsilon_{\mathrm{FP}} = \Pr_{X\sim P}[\mathcal{D}(X)\to 1], \qquad \varepsilon_{\mathrm{FN}} = \Pr_{X\sim Q}[\mathcal{D}(X)\to 0].$$

   (1) 定义 likelihood ratio 为 $L : \Omega \to [-\infty,+\infty]$, $L(x) = \log(\frac{Q(x)}{P(x)})$.
   证明: 对任何区分器 $\mathcal{D}$, 存在算法 $\mathcal{D}' : [-\infty,+\infty] \to \{0,1\}$, 使得

$$\Pr_{X\sim P}[\mathcal{D}(X)\to 1] = \Pr_{X\sim P}[\mathcal{D}'(L(X))\to 1], \qquad \Pr_{X\sim Q}[\mathcal{D}(X)\to 0] = \Pr_{X\sim Q}[\mathcal{D}'(L(X))\to 0].$$

   (2) 证明: 为了最小化 $\varepsilon_{\mathrm{FP}},\varepsilon_{\mathrm{FN}}$, 只须考虑如下的 likelihood ratio test 区分器 $\mathcal{D}_{\tau,\theta}$

$$\mathcal{D}_{\tau,\theta}(x) = \begin{cases} 1, & \text{if } L(x) > \tau \\ \mathrm{Bern}(\theta), & \text{if } L(x) = \tau \\ 0, & \text{if } L(x) < \tau \end{cases}$$

(3) 改为区分 P^n 和 Q^n. 这时区分器是输入域为 Ω^n, 输出域为 $\{0,1\}$ 的算法. 随着 n 的增长, 是否可以让 $\varepsilon_{\mathrm{FP}}, \varepsilon_{\mathrm{FN}}$ 分别以 $\exp(-n\alpha), \exp(-n\beta)$ 的速度趋近于 0?

具体来说, 请确定以下区域的边界

$$\left\{ (\alpha,\beta) \in \mathbb{R}_+^2 \,\middle|\, \text{对任意充分大的 } n\text{, 存在区分器 } \mathcal{D}\text{, 同时满足} \begin{array}{r} \Pr_{X\sim P}[\mathcal{D}(X)\to 1] \le \exp(-n\alpha) \\ \Pr_{X\sim Q}[\mathcal{D}(X)\to 0] \le \exp(-n\beta) \end{array} \right\}$$

为了统一记号, 对任意 $\lambda \in [0,1]$, 定义分布 P_λ 为 $P_\lambda(x) \propto (P(x))^{1-\lambda}(Q(x))^\lambda$.

提示: 上次作业第 3 题.

4. (8 分) 有两个相互独立的秘密, 分别用随机变量 C_0, C_1 表示, 满足 $H[C_0] = H[C_1] = n > 0$. 根据 C_0, C_1, 用一个随机算法生成 A_0, A_1, B_0, B_1. Alice 选择 $\alpha \in \{0,1\}$, 并获得 A_α. Bob 选择 $\beta \in \{0,1\}$, 并获得 B_β. 我们要求, 无论 (α,β) 是多少:

- Alice 和 Bob 各自都没有得到 (C_0, C_1) 的任何信息;
- Alice 和 Bob 联合起来可以知道 $C_{\alpha\beta}$, 但没有得到 $C_{1-\alpha\beta}$ 的任何信息.

请问 A_0, A_1, B_0, B_1 可以有多短.

(1) 将两条要求用信息量（熵、条件熵、互信息等）表示.

(2) 证明 $\max(H[A_0], H[A_1], H[B_0], H[B_1]) > n$.

(3) 证明 $\max(H[A_0], H[A_1], H[B_0], H[B_1]) \ge 1.5n$.

(4) (0 分) 证明上一问的界是紧的. 构造 C_0, C_1 的分布, 以及生成 A_0, A_1, B_0, B_1 的随机算法.