备选题目

1. (10 分) 使用自然演绎 (Natural deduction) [1]推出 $(((p \to q) \to p) \to p)$. 需要写明每步使用的推导规则.

2. (5 分) 考虑命题逻辑的自然演绎系统, 证明 $r \to p \vdash ((p \to q) \to r) \to p$.

3. (10 分) (1) 用自然演绎推导下面两个命题时, 哪个需要使用 RAA 规则, 哪个不需要使用 RAA 规则?

   - $((p \to q) \to (\neg q \to \neg p))$
   - $((\neg q \to \neg p) \to (p \to q))$

   (2) 现在我们考虑一个修改过的自然演绎系统, 去掉 RAA 规则, 但是增加一个公理模式: 对于任何命题 ϕ, 可以把 $(\neg\neg\phi \to \phi)$ 直接作为公理使用. 证明: $(((p \to q) \to p) \to p)$ 可以被推出.

   (3) 现在我们考虑一个修改过的自然演绎系统, 去掉 RAA 规则, 但是增加一个公理模式: 对于任何命题 ϕ, ψ, 可以把 $(((\phi \to \psi) \to \phi) \to \phi)$ 直接作为公理使用. 证明: 对于任何命题 ϕ, $(\neg\neg\phi \to \phi)$ 可以被推出.

4. (5 分) 用 System K[2] 证明 $((p \to q) \to (\neg q \to \neg p))$, 也就是要推出 $\Rightarrow ((p \to q) \to (\neg q \to \neg p))$.

5. (5 分) 考虑两个定义了良序关系的集合 $(S, \leq_S), (T, \leq_T)$, 证明: 要么 $|S| \leq |T|$, 要么 $|T| \leq |S|$.

   提示: 对于两个定义了序关系的集合 $(S, \leq_S), (T, \leq_T)$, 如果有一个双射 $\pi : S \to T$ 满足

   $$\forall a, b \in S, a \leq_S b \iff \pi(a) \leq_T \pi(b)$$

   那么我们称 π 是 $(S, \leq_S), (T, \leq_T)$ 之间的序同构（order isomorphism）.

   提示: 对于任意 $s \in S$, $t \in T$, 定义集合 S_s, T_t 为

   $$S_s = \{a \in S | a <_S s\}, \qquad T_t = \{b \in T | b <_T t\}.$$

   考虑 $(S_s, \leq_S), (T_t, \leq_T)$ 之间是否存在序同构.

   注意到, 假如我们使用选择公理, 那么任何集合都有良序. 这题说明选择公理蕴含了集合之间的势可以比较.

6. (15 分) 考虑只包含连接词 $\to$ 和 $\wedge$ 的命题逻辑, 它的自然演绎系统（natural deduction system）包括如下内容:

   - 命题逻辑的公式, 包括永假常元 $\perp$.
   - 没有公理.

[1] 使用没有 $\neg$ 和 $\vee$ 的简化版. 参见 Logic and Structure 第 2.4 节.
[2] 参见 Sequents and Trees 第 1.2.2 节.

- 推导规则如下:

$$\frac{\phi \quad \psi}{\phi \wedge \psi} \wedge I \qquad \frac{\phi \wedge \psi}{\phi} \wedge E_1 \qquad \frac{\phi \wedge \psi}{\psi} \wedge E_2 \qquad \frac{\bot}{\phi} \bot$$

$$\begin{array}{c} [\phi] \\ \mathcal{D} \\ \dfrac{\psi}{\phi \to \psi} \to I \end{array} \qquad \frac{\phi \quad \phi \to \psi}{\psi} \to E \qquad \begin{array}{c} [\neg\phi] \\ \mathcal{D} \\ \dfrac{\bot}{\phi} \text{RAA} \end{array}$$

  其中, 横线上面是前提, 下面是推导的结果（结论）, 横线右边是规则的名字（例如 $\wedge I$）. $\mathcal{D}$ 表示省略的推导步骤；方括号表示假设该公式已经推出, 在此基础上进行推理.

我们将 $\phi \to \bot$ 缩写为 $\neg\phi$（注意, $\neg$ 不属于字符集）.

(1) 证明: 对任意公式 ψ 和 ϕ, $\psi \vdash \phi$ 当且仅当 $\vdash \psi \to \phi$.

(2) 利用自然演绎系统证明 $(\phi \to \psi) \wedge (\phi \to \neg\psi) \vdash \neg\phi$.

(3) 利用完全性定理（completeness theorem）证明 $(\phi \to \psi) \wedge (\phi \to \neg\psi) \vdash \neg\phi$.

7. (5 分) 考虑通常的命题逻辑自然演绎系统, 作为简化, 连接词只有 $\to$, 命题字母的集合为 P. 将推导规则除去 RAA, 我们就得到了直觉主义命题逻辑, 推理符号为 $\vdash_i$. 本题证明在经典的语义 $\vDash$ 下, 直觉主义命题逻辑是不完全的（incomplete）, 即存在公式 ϕ, $\vDash \phi$ 但 $\nvdash_i \phi$.

为此, 我们将会给一种新的语义 $\Vdash$. 考虑一个偏序集 $(W, \geq)$, 每一个 $w \in W$ 上都被赋予一些命题字母 $p, q, \ldots$, 由映射 $V : W \to 2^P$ 给出, 满足: 如果 $v \geq w$, 那么 $V(v) \supseteq V(w)$. 语义 $\Vdash$ 可以被归纳定义为:

- $(W, V, w) \Vdash p$ 当且仅当 $p \in V(w)$.
- $(W, V, w) \Vdash \bot$ 永远不成立.
- $(W, V, w) \Vdash \psi \to \phi$ 当且仅当对所有 $v \geq w$, 如果 $(W, V, v) \Vdash \psi$, 那么 $(W, V, v) \Vdash \phi$.

符号 $\vDash_i \phi$ 表示 $(W, V, w) \Vdash \phi$ 对所有 W, V, w 成立.

(1) 证明: 如果 $(W, V, w) \Vdash \phi$ 且 $v \geq w$, 那么 $(W, V, v) \Vdash \phi$.

(2) 证明一致性（soundness）定理: 对任意公式 ϕ, 如果 $\vdash_i \phi$, 那么 $\vDash_i \phi$.

(3) 证明: 公式 $\phi = ((p \to q) \to p) \to p$ 同时满足 $\vDash \phi$ 和 $\nvdash_i \phi$. 因而在经典语义 $\vDash$ 下, 直觉主义命题逻辑是不完全的.

注. *语义 $\vDash_i$ 和语法 $\vdash_i$ 实际上满足完全性定理: 对任意公式 ϕ, 如果 $\vDash_i \phi$, 那么 $\vdash_i \phi$.*

8. (15 分) 考虑一阶逻辑, 它的变元是 $x, y, \ldots$, 至少包括一个二元谓词 P. 考虑公式 $\phi := \exists x \forall y P(x, y)$, $\psi := \forall y \exists x P(x, y)$.

(1) 用一阶逻辑的自然演绎（有 $\forall I, \forall E$ 规则, “$\exists$”是“$\neg\forall\neg$”的简记）推导出 $\vdash \phi \to \psi$.

(2) 证明: $\phi \to \psi$ 是有效的, 即对于任意解释 I, 都有 $I \vDash \phi \to \psi$.

(3) 给一个语义解释 I, 说明 $\psi \to \phi$ 不是有效的.

9. (10 分) 令 Γ 表示如下公理集合, 不难看出它们是在 Peano 公理中不涉及乘法的公理.

$$\begin{aligned}
&\forall x \neg(0 = S(x)) \\
&\forall xy(S(x) = S(y) \to x = y) \\
&\forall x(x + 0 = x) \\
&\forall xy(x + S(y) = S(x + y)) \\
&(\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(S(x))) \to \forall x\varphi(x) \qquad \text{对任意 } \varphi
\end{aligned}$$

为了简化证明, 我们使用几个稍作修改的相等规则, 不再局限于变量符号. 它们都很容易从原有的相等公理使用一次 $\forall E$ 规则得出.

$$\frac{}{t = t}\,\mathrm{RI}_1^* \qquad \frac{t_1 = t_2}{t_2 = t_1}\,\mathrm{RI}_2^* \qquad \frac{t_1 = t_2 \qquad t_2 = t_3}{t_1 = t_3}\,\mathrm{RI}_3^* \qquad \frac{t_1 = t_2}{t(t_1) = t(t_2)}\,\mathrm{RI}_4^* \quad \frac{t_1 = t_2}{\varphi(t_1) \to \varphi(t_2)}\,\mathrm{RI}_4^*$$

作为例子, 我们给出 $\Gamma \vdash \forall x(0 + x = x)$ 的证明. 记 $\varphi(x) := 0 + x = x$. 也就是要证明 $\forall x\varphi(x)$.

$$\dfrac{\dfrac{\dfrac{\forall x(x+0=x)}{\substack{0+0=0 \\ \text{i.e. } \varphi(0)}}\,\forall E \qquad \dfrac{\dfrac{\dfrac{\dfrac{\forall xy(x+S(y)=S(x+y))}{0+S(x)=S(0+x)}\,\forall E \qquad \dfrac{[0+x=x]}{S(0+x)=S(x)}\,\mathrm{RI}_4^*}{0+S(x)=S(x)}\,\mathrm{RI}_3^*}{\substack{0+x=x \to 0+S(x)=S(x) \\ \text{i.e. } \varphi(x)\to\varphi(S(x))}}\,\to I}{\forall x(\varphi(x)\to\varphi(S(x))}\,\forall \mathrm{I}}{\varphi(0)\wedge\forall x(\varphi(x)\to\varphi(S(x))}\,\wedge I \qquad (\varphi(0)\wedge\forall x(\varphi(x)\to\varphi(S(x)))\to\forall x\varphi(x)}{\forall x\varphi(x)}\,\to E$$

作为例子, 我们给出 $\Gamma \vdash \forall xy(S(x) + y = S(x + y))$ 的证明. 记 $\psi(y) := \forall x(S(x) + y = S(x + y))$. 也就是要证明 $\forall y\psi(y)$.

$$\dfrac{\dfrac{\dfrac{\forall x(x+0=x)}{S(x)+0=S(x)}\,\forall E \qquad \dfrac{\dfrac{\dfrac{\forall x(x+0=x)}{x+0=x}\,\forall E}{x=x+0}\,\mathrm{RI}_2^*}{S(x)=S(x+0)}\,\mathrm{RI}_4^*}{S(x)+0=S(x+0)}\,\mathrm{RI}_3^*}{\forall x S(x)+0=S(x+0) \text{ i.e. } \psi(0)}\,\forall I$$

$$
\dfrac{\dfrac{\psi(0)\qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\forall xy(x+S(y)=S(x+y))}{S(x)+S(y)=S(S(x)+y)}\,\forall\mathrm{E}\qquad \dfrac{\dfrac{[\forall x(S(x)+y=S(x+y))]\text{ i.e. }\psi(y)}{S(x)+y=S(x+y)}\,\forall E}{S(S(x)+y)=S(S(x+y))}\,\mathrm{RI}_4^*}{S(x)+S(y)=S(S(x+y))}\,\mathrm{RI}_3^*\qquad \dfrac{\dfrac{\dfrac{\forall xy(x+S(y)=S(x+y))}{x+S(y)=S(x+y)}\,\forall\mathrm{E}}{S(x+S(y))=S(S(x+y))}\,\mathrm{RI}_4^*}{S(S(x+y))=S(x+S(y))}\,\mathrm{RI}_2^*}{S(x)+S(y)=S(x+S(y))\text{ i.e. }\psi(S(y))}\,\mathrm{RI}_3^*}{\psi(y)\to\psi(S(y))}\,{\to}\mathrm{I}}{\forall y(\psi(y)\to\psi(S(y)))}\,\forall\mathrm{I}}{\psi(0)\land\forall y(\psi(y)\to\psi(S(y))}\,\land I\qquad (\psi(0)\land\forall y(\psi(y)\to\psi(S(y)))\to\forall y\psi(y)}{\forall y\psi(y)}\,\to E
$$

(1) 使用自然演绎法, 证明 $\Gamma\vdash\forall xy(x+y=y+x)$.

(2) 使用自然演绎法, 证明 $\Gamma\vdash\forall x(\exists z(0=x+z)\to x=0)$. 人们常用记号 $y\geq x$ 表示 $\exists z(y=x+z)$, 这时待证命题可以写为 $\forall x(0\geq x\to x=0)$.

可以直接使用例子中已经证明的命题. 后一问可以直接使用前一问的结论.

10. (10 分) 考虑 ZFC 集合论, 它的变元是 $x,y,\ldots$, 谓词包括 $\in,=$.

(1) "存在且唯一"的符号是 $\exists!$, 请用 ZFC 公式给出它的定义. 也就是说, 给一个公式 ϕ_A, 使得 ϕ_A 表示 $\exists!xA(x)$, 读作"存在唯一的 x 使 $A(x)$ 成立".

(2) 利用第一问的记号, 给出二元关系 R 是从集合 X 到集合 Y 的函数关系的 ZFC 公式定义.

提示: 第二问的公式中允许使用集合论的常用符号, 例如交 $\cap$、并 $\cup$、包含 $\subseteq$、笛卡尔积 $\times$、序对 (x,y)、子集符号 $\{x\in X:\phi(x)\}$, 幂集符号 2^X 等.

11. (15 分) 考虑 ZFC 集合论, 本题讨论正则公理.

(1) 证明, 不存在一对互相包含的集合 x,y. 即证明 $\forall x\forall y\neg(x\in y\land y\in x)$.

(2) 证明, 不存在一列集合 $x_0,x_1,\ldots,x_n,\ldots$ 满足 $\forall i\in\mathbb{N},x_{i+1}\in x_i$.

(3) 在保留 ZFC 中其它公理的前提下, 证明前一问的命题可以推出正则公理.

12. (5 分) 在这个问题中, 我们试图构造一个比 ZFC 更近完备的形式系统.

令 Ω 表示 ZFC 使用的字母表允许出现的所有不含自由变量的合式公式（closed well-formed formula）, 可以写为 $\Omega=\{\varphi_1,\varphi_2,\ldots\}$.

用以下方式递归定义 T_i,F_i,Γ_i:

- 令 T_0 表示所有 ZFC 可以推理演绎得到的命题. 令 $F_0:=\{\text{“}\varphi\text{”}\mid\text{“}\neg\varphi\text{”}\in T_0\}$ 表示 T_0 中命题的否命题, 也就是 ZFC 可以"否定"的命题. 根据 Gödel 的不完备定理, 我们知道或者 $T_0\cap F_0\neq\varnothing$ (不一致), 或者 $T_0\cup F_0\not\supseteq\Omega$ (不完备). 我们暂且假设 ZFC 是一致的.
- 令 $\Gamma_0=\varnothing$.
- 如果 $\varphi_i\in T_{i-1}\cup F_{i-1}$, 那么定义 $T_i:=T_{i-1},F_i:=F_{i-1},\Gamma_i:=\Gamma_{i-1}$.
- 如果 $\varphi_i\notin T_{i-1}\cup F_{i-1}$, 那么定义 $\Gamma_i:=\Gamma_{i-1}\cup\{\varphi_i\}$. 令 T_i 表示所有 ZFC + Γ_i 可以推理演绎得到的命题. 令 F_i 表示所有 ZFC + Γ_i 可以"否定"的命题.

不难看出, $T_i \supseteq T_{i-1}, F_i \supseteq F_{i-1}, \Gamma_i \supseteq \Gamma_{i-1}$ 并且 $T_{i-1} \cap F_{i-1} = \varnothing \implies T_i \cap F_i = \varnothing$. 令 $\Gamma = \bigcup_i \Gamma_i$. 那么 $T = \bigcup_i T_i$ 是所有 ZFC + Γ 可以推理演绎得到的命题, $F = \bigcup_i F_i$ 是所有 ZFC + Γ 可以"否定"的命题. 且 $T \cup F \supseteq \Omega$. 也就是说, 如果把 Γ 中的命题都作为公理加入 ZFC, 可以在不破坏一致性的同时获得完备性.

判断上述结论是否违反了 Gödel 不完备定理. 如果是, 请指出证明中的错误. 如果否, 请解释.

13. (10 分) (1) 求 $\gcd(10^6 - 1, 10^{15} - 1)$.

(2) 设自然数 $n, m \geq 1$, 证明: $\gcd(2^m - 1, 2^n - 1) = 2^{\gcd(m,n)} - 1$.

14. (10 分) 对实数 $x \in \mathbb{R}$, 定义

$$\mu(x) = \inf\left\{\alpha \in \mathbb{R} : \left|x - \frac{p}{q}\right| \leq \frac{1}{q^\alpha} \text{ 仅有有限组互素整数解 } (p,q), q > 0\right\}.$$

证明: 对 $x \in \mathbb{Q}$, $\mu(x) = 1$.

提示: 首先证明 $\mu(x) \geq 1$, 然后考虑 $|x - p/q| \leq 1/q^{1+\epsilon}$ 的解个数, 进而证明 $\mu(x) < 1 + \epsilon$.

15. (5 分) 已知群 G 满足 $\forall g \in G, g^2 = e$. 证明 G 是阿贝尔群.

16. (10 分) 设 G 是一个有限阿贝尔群, 证明以下命题

(1) $\prod_{g \in G} g$ 的平方等于单位元 e.

(2) 如果 G 中没有阶（order）为 2 的元素, 或 G 中有超过一个阶为 2 的元素, 那么 $\prod_{g \in G} g = e$.

(3) 如果 G 中唯一的阶（order）为 2 的元素 y, 那么 $\prod_{g \in G} g = y$.

(4) (Wilson's theorem) 如果 p 是素数, $(p-1)! \equiv -1 \mod p$.

17. (10 分) 形如 $aba^{-1}b^{-1}$ 被称作 a, b 的换位子. 给定群 G, 定义它的换位子群（commutator subgroup）$G' = \langle aba^{-1}b^{-1} : a, b \in G \rangle$ 为所有换位子生成的群.

(1) 证明: $G' \trianglelefteq G$.

(2) 考虑 $N \trianglelefteq G$, 证明: G/N 是 Abel 群当且仅当 $G' \leq N$.

(3) 考虑 Sym(4), 即 $\{1,2,3,4\}$ 上的对称群, 请给出一个序列:

$$S_4 = G^0 \triangleright G^1 \triangleright \cdots \triangleright G^n = \{1\},$$

满足 G^i/G^{i+1}（$i = 0, \ldots, n-1$）是 Abel 群.

18. (5 分) 考虑集合 S 与 S 上的二元运算 $\cdot$. 如果 $\cdot$ 满足结合律, 那么 $(S, \cdot)$ 构成半群 (semigroup). 如果还存在单位元, 那么称作幺半群 (monid). 如果还存在逆元, 那么就是群.

假设半群 $(S, \cdot)$ 额外满足

- 对称性 $\forall a \forall b\ a \cdot b = b \cdot a$
- 消去律 $\forall a \forall b \forall c\ a \cdot b = a \cdot c \to b = c$

证明, 可以将 S 嵌入一个群 G 中. 也就是存在群 $(G, \star)$, 满足 $S \subseteq G$ 且 $\forall a \forall b\ a \cdot b = a \star b$.

19. (5 分) 证明或证伪以下命题:

(1) 单同态 $\varphi : G \to G$ 一定是自同构.

(2) 满同态 $\varphi : G \to G$ 一定是自同构.

20. (5 分) 对任意群 G, 定义 $\operatorname{Aut} G$ 是所有 G 的自同构 (automorphism), 定义 $\operatorname{Inn} G$ 为所有 G 的内自同构 (inner automorphism).

$$\operatorname{Aut} G := \{\text{同构 } \sigma : G \to G\},$$
$$\operatorname{Inn} G := \{\phi_g : h \mapsto ghg^{-1} | g \in G\},$$

证明, $\operatorname{Inn} G \trianglelefteq \operatorname{Aut} G$.

21. (10 分) 如果 $p = 2p' + 1$, 其中 p, p' 都是素数, 那么 p 被称作"安全素数". 考虑两个安全素数 $p = 2p' + 1, q = 2q' + 1$, 其中 p, p', q, q' 两两不同且均大于 2. 记 $n = pq$.

证明: $\mathbb{Z}_{n^2}^* \cong \mathbb{Z}_n^* \times \mathbb{Z}_n$.

提示: 可以考虑如下三个映射, $\pi_1 : \mathbb{Z}_n^* \to \mathbb{Z}_{n^2}^*$, $\pi_2 : \mathbb{Z}_n \to \mathbb{Z}_{n^2}^*$ 和 $\pi : \mathbb{Z}_n^* \times \mathbb{Z}_n \to \mathbb{Z}_{n^2}^*$

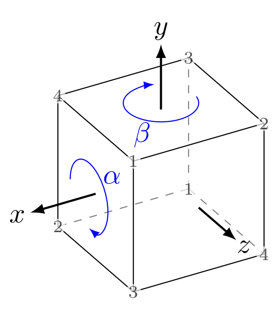
$$\pi_1(a) = a^n, \qquad \pi_2(t) = (1+n)^t, \qquad \pi(a,t) = a^n(1+n)^t.$$

22. (10 分) 给定一个正方体, 按照某种特定方式对它整体旋转（即特殊正交变换, 可以保角度的旋转, 但没有镜面操作）的时候, 它会与原来的正方体重合, 尽管点和面可能换了位置. 以正方体的中心为原点, 沿着正方体的边建立 x 轴（左右方向）、y 轴（上下方向）和 z 轴（前后方向）, 正方体的"基本旋转"恰好就是顺时针沿着 x 轴或 y 轴转九十度, 记为 α, β. 可以证明, 保持正方体占位不变的旋转都是由这两种旋转生成的, 因此正方体的旋转构成了一个群, 记为 R.

(1) 证明: $R \cong \mathrm{Sym}(4)$, 因此 $\mathrm{Sym}(4)$ 可以被视为正方体的旋转群.

提示: 对正方体的顶点编号 $1, 2, 3, 4$, 并且对径点编上相同的号, 这样一来, 每一个面的顶点都恰好具有四个编号, 考虑底面的编号, 给出 α, β 所对应的 $\mathrm{Sym}(4)$ 中的元素, 证明他们生成了 $\mathrm{Sym}(4)$.

(2) 写出 $\mathrm{Sym}(4)$ 的类方程（class equation）, 并解释它的几何意义（即每个共轭类对应的旋转类型）.



23. (10 分) 若 $N \triangleleft G$, 考虑商群 G/N, 并考虑自然映射 $\eta : G \to G/N$, $\eta(g) = gN$. 若 $K' \triangleleft H' \triangleleft G/N$. 定义

$$H = \eta^{-1}(H') = \{g \in G \mid \eta(g) \in H'\},$$
$$K = \eta^{-1}(K') = \{g \in G \mid \eta(g) \in K'\}.$$

证明:

(1) H, K 是群且 $N \triangleleft K \triangleleft H \triangleleft G$.

(2) $H/K \cong H'/K'$.

24. (10 分) 设 $N \trianglelefteq G$, $|N| = n$, $[G : N] = m$. 这里记号 $[G : N]$ 表示 G 中 N 的左陪集的数目, 被称作指数 (index).

(1) 设 $g \in G$ 且 $\gcd(\mathrm{order}(g), m) = 1$. 证明 $g \in N$.

(2) 设 m 和 n 互素. 证明, N 是 G 的唯一的大小为 n 的正规子群.

25. (10 分) 设 G 是一个 15 阶群.

(1) 证明 G 有阶分别为 3 和 5 的正规子群.

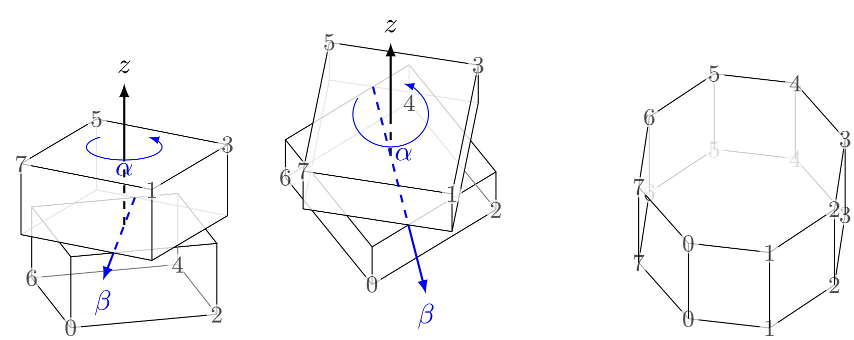
(2) 证明 G 是循环群.

提示: 需要使用 Sylow 第三定理: 若 $|G| = p^k m$ (p 为素数, $m > 0$, $\gcd(p, m) = 1$), 记 Sylow p-子群的数目为 n_p, 那么 $n_p \mid m$ 且 $n_p \equiv 1 \pmod p$.

26. (10 分) 群 G 是有限生成群(finitely generated group)当且仅当存在有限集合 $F \subseteq G$ 使得 $G = \langle F \rangle$. 设 G 是有限生成群, $\{g_1, \ldots, g_n\}$ 是 G 的一个生成集. 用 $\mathsf{free}(S)$ 表示由一个给定符号集合 $S = \{s_1, \ldots, s_n\}$ 生成的自由群. 证明, 存在 $\mathsf{free}(S)$ 的一个正规子群 H, 满足 $G \cong \mathsf{free}(S)/H$.

27. (5 分) 设素数 $p > 2$ 且, 我们知道二面体群 D_p 是一个 $2p$ 的非循环群. 证明, 在同构意义下, 阶为 $2p$ 的非循环群只有 D_p.

28. (5 分) 给定一个正方体, 垂直于 z 轴将其平分为两半, 并将其下半部分绕 z 轴旋转 45 度.



按照某种特定方式对它整体旋转（即特殊正交变换, 可以保角度的旋转, 但没有镜面操作）的时候, 它会与原来的占位重合, 尽管点和面可能换了位置. 占位不变的旋转构成一个群 R. 记 α 为逆时针绕着 z 轴转 90 度. 记 β 为绕着图中标出的轴转 180 度. 可以证明, α, β 生成了 R.

(1) 证明: R 同构于 D_8 的一个子群.

提示: 对正方体的顶点编号 $0,1,2,3,4,5,6,7$.

(2) 写出 R 的类方程 (class equation), 并解释它的几何意义 (即每个共轭类对应的旋转类型).

29. (10 分) 考虑环的两个性质

- ACC 指不存在理想的无穷严格上升列. 即不存在无穷个理想 $I_1, I_2, \dots$ 满足 $I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \dots$
- DCC 指不存在理想的无穷严格下降列. 即不存在无穷个理想 $I_1, I_2, \dots$ 满足 $I_1 \supsetneq I_2 \supsetneq I_3 \supsetneq \dots$

课上证明 PID 是 UFD 时, 实际就是证明 PID 环必满足 ACC.

(1) 给出一个满足 ACC 但不满足 DCC 的整环的例子.

(2) 证明满足 DCC 的整环（无零因子含单位元的交换环）一定满足 ACC. *Remark:* 实际上该环一定是域, 所以满足 ACC.

30. (10 分) 证明环版本的中国剩余定理. 给定一个含幺交换环 R 以及若干真理想 (proper ideals) $I_1, \dots, I_k$. 这些理想之间两两互素, 也就是 $R = I_a + I_b$. 证明

$$R/I_1 I_2 \dots I_k \cong (R/I_1) \times (R/I_2) \times \cdots \times (R/I_k).$$

为此, 我们先证明几个小结论

(1) 如果理想 I, J 互素, 那么 $IJ = I \cap J$.

(2) 如果理想 I, J 互素, 那么 $R/(I \cap J) \cong (R/I) \times (R/J)$.

(3) 如果理想 I, J, K 两两互素, 那么 IJ, K 互素.

(4) 证明环版本的中国剩余定理.

31. (5 分) 给定一个域 $\mathbb{F}$, 证明 $M_n(\mathbb{F})$ 是单环（即不存在非平凡理想）.

32. (10 分) 给定一个环 R 以及它的一个理想 I. R/I 是域可以推出 I 是极大理想. 这里我们证明相反方向.

(1) 如果额外知道 R 是（含有单位元的）交换环, 那么 I 是极大理想可以推出 R/I 是域.

Remark: 作为推论, 如果交换环 R 没有非平凡理想, 那么 R 一定是域.

(2) 一个理想 I 被称作 R 的素理想, 当且仅当 $\forall a \in R, \forall b \in R, ab \in I \implies a \in I \vee b \in I$. 证明（含有单位元的）交换环 R 的任何一个极大理想都是素理想.

33. (10 分) 对任意集合 S, 用 $M_n(S)$ 表示 S 中元素构成的 $n \times n$ 矩阵的集合.

(1) 给定一个域 $\mathbb{F}$, 证明 $M_n(\mathbb{F})$ 是单环, 即不存在非平凡理想.

(2) 给定一个环 R, 证明 $M_n(R)$ 的理想一定形如 $M_n(I)$, 其中 I 是 R 的理想.

(3) 证明 $M_n(R)/M_n(I) \cong M_n(R/I)$.

34. (5 分) 给定一个环 R 以及它的两个理想 I, J, 满足 $R = I + J$. 证明

$$R/(I \cap J) \cong (R/I) \times (R/J)$$

Remark: 这是中国剩余定理的推广. 当 $R = \mathbb{Z}$, $I = p\mathbb{Z}$, $J = q\mathbb{Z}$ 便转化为中国剩余定理.

35. (5 分) 考虑 $\mathbb{F}_{103}[x]$ 中的多项式 $f(x) = x^3 - 2$, $\mathbb{F}_{103}[x]/(f(x))$ 是否是域? 请计算说明.

36. (10 分) 设 $k \geq 1$ 是整数, 找到最小的 d, 使得对任意 n, 都存在一个 $\mathbb{F}_2$ 上的次数不超过 d 的多项式 $f(x_1, \dots, x_n)$

$$\forall x_1, \dots, x_n \in \{0, 1\},\ f(x_1, \dots, x_n) = \begin{cases} 1, & \text{if } x_1, \dots, x_n \text{ 中 } 1 \text{ 的个数} \equiv -1 \pmod{2^k} \\ 0, & \text{otherwise} \end{cases}$$

提示: 先考虑 $n = 2^k - 1$ 的情况.

37. (4 分) $\mathbb{F}$ 是有限域, 证明 $\mathbb{F}^* = \mathbb{F} \setminus \{0\}$ 是循环群.

提示: 可以利用 $n = \sum_{d|n} \varphi(d)$.

38. (5 分) 定义 $R_n = \sum_{k \geq 0} \binom{n+k}{k} 2^{-k}$. 化简 R_n 的表达式.

39. (5 分) 定义 $R_n = \sum_{k \leq n} \binom{n-k}{k} (-1)^k$. 化简 R_n 的表达式.

40. (10 分) 每一个置换 $g \in \mathrm{Sym}(n)$ 都可以写成若干不交的轮换.
    (1) 对任意 $k > n/2$, 问有多少个置换包含一个长度恰好为 k 的轮换.
    (2) 对任意 $\alpha > 1/2$, 对一个随机的置换 $g \in \mathrm{Sym}(n)$ 包含一个长度至少为 αn 的轮换的概率大概是多少. 这里假设 n 充分大.

注. *(这往往会被包装为以下问题.)* 监狱中有 n 位囚犯. 现在让他们进行如下游戏. 在房间里布置标号 $1, \dots, n$ 的 n 个柜子, 其中分别写有 n 个囚犯的名字. 每个囚犯分别被带到房间中, 允许打开其中至多 αn 个柜子. 如果所有囚犯都找到了包含自己名字的柜子, 那么所有囚犯都被释放. 否则, 所有囚犯都被处死. 游戏开始后囚犯之间不能交流. 请问囚犯应该采取怎样的策略.

41. (10 分) 给定一个函数 $f : 2^{[n]} \to \mathbb{R}$. 证明如果定义 $\tilde{f}(S) = \sum_{T \supseteq S} f(T)$, 那么

$$f(S) = \sum_{T \supseteq S} (-1)^{|T \setminus S|} \tilde{f}(T).$$

Remark: 对于一组有限集 $A_1, \dots, A_n$ 和 $\Omega = A_1 \cup A_2 \cup \cdots \cup A_n$. 如果定义

$$f(S) = |\{x \in \Omega | \forall i \in [n], x \in A_i \iff i \in S\}|,$$

那么题目结论可以推出容斥原理.

Remark: 对称地, 如果定义 $\hat{f}(S) = \sum_{T \subseteq S} f(T)$, 那么

$$f(S) = \sum_{T \subseteq S} (-1)^{|S \setminus T|} \hat{f}(T).$$

42. (10 分) 令 $M \in \mathrm{Mat}_{n,n}(\mathbb{F})$ 是一个有限域 $\mathbb{F}$ 上的 $n \times n$ 矩阵. 定义 M 是一个 MDS（maximum distance separable）矩阵, 当且仅当对任何不同的 $x, x' \in \mathbb{F}^n$, (Mx, x) 和 (Mx', x') 至少在 $n+1$ 个位置不同. 不难证明以下命题等价,

   a. M 是 MDS 矩阵;
   b. M 可逆, 且 M^{-1} 是 MDS 矩阵;
   c. M 的任何子矩阵满秩;
   d. 对任何非零 $x \in \mathbb{F}^n$, (Mx, x) 至少在 $n+1$ 个位置非零.
   e. 考虑方程 $(y_1, \dots, y_n) = M(x_1, \dots, x_n)$, 任意固定 $x_1, \dots, x_n, y_1, \dots, y_n$ 中的 n 个变量, 方程仍有解;
   f. 考虑方程 $(y_1, \dots, y_n) = M(x_1, \dots, x_n)$, 任意固定 $x_1, \dots, x_n, y_1, \dots, y_n$ 中的 n 个变量, 方程有唯一解.

   由等价命题 c 可以看出, 当 $|\mathbb{F}|$ 足够大时, 大部分矩阵都是 MDS 矩阵. 因此, 可以说 MDS 刻画了"一般的"矩阵.

   (1) 若 $M \in \mathrm{Mat}_{n,n}(\mathbb{F})$ 是一个 MDS 矩阵, 求出满足 $(x_{n+1}, \dots, x_{2n}) = M(x_1, \dots, x_n)$ 且 $x_1, \dots, x_{2n}$ 均不为 0 的解的个数.

   提示: 对每个集合 $S \subseteq [2n]$, 计算满足 $(x_{n+1}, \dots, x_{2n}) = M(x_1, \dots, x_n)$ 且 $x_i = 0 \iff i \in S$ 的解的个数.

   (2) 记上问求出的解的个数为 L. 证明

$$\left| L - \frac{(|\mathbb{F}| - 1)^{2n}}{|\mathbb{F}|^n} \right| \le 2^{2n}.$$

43. (10 分) 令 t_n 表示 n 个带标号的点组成的有根树的个数. 不妨令 $t_0 = 0$. 其初始几项为

$$t_0 = 0, \quad t_1 = 1, \quad t_2 = 2, \quad t_3 = 9, \quad t_4 = 64, \dots$$

   (1) 求出 t_n 的通项公式.

   提示:

   (2) 考虑 t_n 的 EGF $\tilde{T}(x) = \sum_i \frac{1}{n!} t_n x^n$. $\tilde{T}(x)$ 是一个简洁的方程的解, 请找到这个 (超越) 方程.

44. (5 分) 设 G 是点集 $V = [n]$ 上的一个简单无向图. 图中的点形成了 k 个联通子块 $S_1, \dots, S_k \subseteq V$. 向 G 添加 $k-1$ 条边, 使得图联通. 问有多少种不同的添加边的方法.

   提示: 如果每个联通子块都是单点集, 那么题目就是在问有标号的 k 个点组成的无根树的个数.

45. (5 分) 用 $\mathsf{M}_n(\mathbb{F})$ 表示所有 $\mathbb{F}$ 上的 $n \times n$ 矩阵. 用 $\mathsf{GL}_n(\mathbb{F})$ 表示所有 $\mathbb{F}$ 上的 $n \times n$ 可逆矩阵. 我们称两个矩阵 $A, B \in \mathsf{M}_n(\mathbb{F})$ 相似当且仅当 $\exists G \in \mathsf{GL}_n(\mathbb{F})$, $GAG^{-1} = B$. 相似是一个等价关系. 求在相似关系下, $\mathsf{M}_2(\mathbb{F}_q)$ 有多少个等价类. 证明你的结果.

   提示: 可以使用 Burnside 引理.

46. (5 分) 用 $C \geq 6$ 种颜色对立方体进行面染色, 要求相邻面的颜色不能相同. 求有多少种不同的染色方案. 两个染色方案等价, 当且仅当其中一种方案可以经过旋转 (不包括镜像) 转化为另一种方案.

47. (8 分) 考虑从 $(0,0)$ 到 (n,n) 的长度为 $2n$ 的不降路径 (每次向上或向右走一步)

$$\vec{p} = [(x_0, y_0), (x_1, y_1), \cdots, (x_{2n}, y_{2n})]$$

其中 $(x_0, y_0) = (0,0), (x_{2n}, y_{2n}) = (n,n), \forall 0 < i \leq 2n\ (x_i - x_{i-1}, y_i - y_{i-1}) \in \{(0,1),(1,0)\}$. 称路径 $\vec{p}$ 不越过对角线, 如果 $\forall 0 \leq i \leq 2n,\ y_i \geq x_i$. 将一条不越过对角线的不降路径沿着 $(0,n)(n,0)$ 的连线翻转, 可以得到一条不越过对角线的不降路径. 计算考虑这种翻转后, 仍然不等价的不越过对角线的不降路径数目. 严格来说, 称路径 $\vec{p} = [(x_0, y_0), \cdots, (x_{2n}, y_{2n})]$ 和路径 $\vec{q} = [(x'_0, y'_0), \cdots, (x'_{2n}, y'_{2n})]$ 等价, 如果

$$\forall 0 \leq i \leq 2n,\ x_i + y'_{2n-i} = y_i + x'_{2n-i} = n.$$

问有多少个等价类?

48. (10 分) 令 $\mathbb{F}$ 是一个有限域. 考虑对域中的每个数用 C 种颜色之一染色. 每个染色方案可以表示为一个映射 $f : \mathbb{F} \to C$. (这里 $C := \{0, 1, \ldots, C-1\}$.) 考虑在仿射变换下仍然不同的染色方法数. 严格来说, 我们说两个染色方案 f, f' 是等价的, 当且仅当存在一个 $\mathbb{F}$ 上的可逆仿射映射 $g : x \mapsto ax + b$ (这里 $a \in \mathbb{F} \setminus \{0\}, b \in \mathbb{F}$), 使得 $f' = f \circ g$.

当 $|\mathbb{F}| = 7^5 = 16807 = 2 \times 3 \times 2801 + 1$, 请计算在仿射变换下仍然不同的染色方法数?

提示: 不妨先考虑一般的有限域 $|\mathbb{F}| = p^k$. 对任何有限域 $\mathbb{F}$, 其乘法群 $\mathbb{F} \setminus \{0\}$ 是循环群.

49. (10 分) 考虑 n 个无差异点构成的圈, 每个点上可以标记 $C := \{0, 1, \ldots, C-1\}$ 中的一个整数, 经过旋转 (不包括镜面) 可重合的标号方式视为同一种. 选以下问题中的 2 个作答即可.

(1) 如果要求相邻两个点的奇偶性不同, 有多少种不同的标号方案? ($n = 30$, $C = 3$)

(2) 如果要求所有点标的和为偶数, 有多少种不同的标号方案? ($C = 3$, $n = p^2$ 为奇素数平方)

(3) 如果要求相邻两个点的标号不同, 有多少种不同的标号方案? ($n = 30$)

(4) 如果要求所有点的标号和为 $C - 1$ 的倍数, 有多少种不同的标号方案? (n 为素数)

50. (6 分) 用 $\text{Poisson}(\lambda)$ 表示期望为 λ 的泊松分布. 随机变量 X 服从柏松分布, 记为 $X \sim \text{Poisson}(\lambda)$, 当且仅当 $\forall k \in \mathbb{N}, \Pr[X = k] = \frac{e^{-\lambda}\lambda^k}{k!}$.

(1) 如果独立的随机变量 X, Y 分别服从 $X \sim \text{Poisson}(\lambda_1)$, $Y \sim \text{Poisson}(\lambda_2)$, 定义另一个随机变量 $Z = X + Y$. 证明 $Z \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

(2) 随机变量 X, Y, Z 的定义和上一文相同. 求已知 Z 时, X 的条件分布. 具体来说, 对任意 $n, k \in \mathbb{N}$, 计算 $\Pr[X = k | Z = n]$.

51. (4 分) 有一族相关事件 $E_1, \ldots, E_{2024}$, 满足 $\Pr[E_i] = \frac{1}{2}, \Pr[E_i \wedge E_j] = \frac{1}{3}, \ldots$. 更一般地, 对任意 $S \subseteq \{1, 2, \ldots, 2024\}$, $\Pr[\bigwedge_{i \in S} E_i] = \frac{1}{|S|+1}$. 计算 $\Pr[\bigvee_{i=1}^{2024} E_i]$.

52. (10 分) 令 $X_1, X_2, \cdots \in \{H, T\}$ 表示反复独立的投掷一枚均匀硬币的结果. 用 T_{HHH} (resp. $T_{HHT}, T_{HTH}, \dots$) 表示首次出现连续的 HHH (resp. $HHT, HTH, \dots$) 的投掷次数. 求出 $\mathbb{E}[T_{HHH}], \mathbb{E}[T_{HHT}], \mathbb{E}[T_{HTH}]$.

Remark: 对于计算密集的步骤, 不妨使用计算机辅助.

53. (5 分) 证明信息熵是凹函数. 也就是对任意分布 P, Q 和 $\lambda \in (0,1)$

$$H[\lambda P + (1-\lambda)Q] \geq \lambda H[P] + (1-\lambda)H[Q].$$

54. (5 分) 对于随机变量 X, Y, 我们定义了信息熵、条件 (信息) 熵、互信息

$$\begin{aligned}
H[X] &= \sum_x \Pr[X=x] \log \frac{1}{\Pr[X=x]} = \mathbb{E}\Big[\log \frac{1}{P_X(X)}\Big], \\
H[X|Y] &= \sum_{x,y} \Pr[X=x, Y=y] \log \frac{1}{\Pr[X=x|Y=y]} = \mathbb{E}\Big[\log \frac{1}{P_{X|Y}(X|Y)}\Big], \\
I(X;Y) &= H[X] - H[X|Y].
\end{aligned}$$

考虑事件 E 或另一随机变量 Z, 还定义了条件互信息

$$\begin{aligned}
I(X;Y|E) &= H[X|E] + H[Y|E] - H[X,Y|E] \\
I(X;Y|Z) &= \sum_z \Pr[Z=z] I(X;Y|Z=z).
\end{aligned}$$

如果我们定义三个随机变量之间的互信息为

$$I(X;Y;Z) = H[X] + H[Y] + H[Z] - H[X,Y] - H[X,Z] - H[Y,Z] + H[X,Y,Z].$$

(1) 证明 $I(X;Y;Z) = I(X;Y) - I(X;Y|Z)$.

(2) 举例说明 $I(X;Y;Z)$ 可以大于零, 也可以小于零. 也就是说, 泄露额外信息 Z 后, X, Y 之间的互信息可能增加, 也可能减少.

55. (5 分) 令 P_{XY} 表示 X, Y 的联合分布. 证明

$$I(X;Y) = D(P_{XY} \| P_X P_Y).$$

56. (5 分) 证明 $d(p\|q) = D(\mathrm{Bern}(p) \| \mathrm{Bern}(q)) \geq 2 \log e \cdot (p-q)^2$.

Remark: 不妨两边都除以 $\log e$, 这等价于使用 e 作底数.

57. (5 分) 证明散度的 data-processing 不等式. 对任意 P_X, Q_X 和 kernel $P_{Y|X}$, 令 $P_Y = P_X \circ P_{Y|X}$, $Q_Y = Q_X \circ P_{Y|X}$ (也就是说, P_Y, Q_Y 分别是 $P_{XY} = P_X P_{Y|X}, Q_{XY} = Q_X P_{Y|X}$ 的边缘分布). 证明

$$D(P_X \| Q_X) \geq D(P_Y \| Q_Y).$$

Remark: 互信息的 data-processing 不等式可以由散度的 data-processing 不等式推出. 如果 X, Y, Z 的依赖关系可以用有向图 $X \to Y \to Z$ 表示 (即 $P_{XYZ} = P_X P_{Y|X} P_{Z|Y}$), 注意到

$$I(X;Y) = D(P_{XY} \| P_X P_Y), \qquad I(X;Z) = D(P_{XZ} \| P_X P_Z).$$

只需定义一个合适的 kernel 便证明了互信息的 data-processing 不等式 $I(X;Y) \geq I(X;Z)$.

58. (5 分) 使用 data-processing 不等式, 证明

$$\sqrt{\frac{1}{2\log e}D(P\|Q)} \geq \Delta(P,Q)$$

这里 $\Delta(P,Q)$ 表示 P,Q 之间的统计距离（statistical distance, 也可以更精确地称为 total variation distance）

$$\Delta(P,Q) = \frac{1}{2}\sum_x |P(x)-Q(x)| = \max_{\text{事件 } E}(P(E)-Q(E)).$$

59. (5 分) 证明对于服从任意联合分布的随机变量 X,Y,Z,

$$2H[X,Y,Z] \leq H[X,Y]+H[X,Z]+H[Y,Z].$$

据此证明 Shearer 引理: 令 Ω 是 $\mathbb{R}^3$ 上 n 个点组成的集合, Ω 向三个坐标平面投影分别有 n_1,n_2,n_3 个像, 那么 $n^2 \leq n_1n_2n_3$. 并说明何时可以取到等号.

60. (5 分) 考虑 Markov kernel $P_{Y|X}:\{0,1\}\times\{0,1\}\to\mathbb{R}$, $P_{Y|X}(0|0)=1$, $P_{Y|X}(0|1)=P_{Y|X}(1|1)=\frac{1}{2}$.
   (1) 找到 P_X^* 使得 $I(X;Y)$ 最大, 其中 $(X,Y)\sim P_X^*P_{Y|X}$. 这个最大值被称作 $P_{Y|X}$ 的容量.
   (2) 令 P_X^* 是前一问找到的分布. 定义 P_Y^* 为 $P_X^*P_{Y|X}$ 的边缘分布. 请计算 $I(X;Y)$, $D(P_{Y|X=0}\|P_Y^*)$ 和 $D(P_{Y|X=1}\|P_Y^*)$ 的值.
   (3) (0 分) 现在考虑任意 Markov kernel $P_{Y|X}:\mathcal{Y}\times\mathcal{X}\to\mathbb{R}$. 为了保证最值存在, 我们要求 $\mathcal{X},\mathcal{Y}$ 都是有限集合. 证明

$$\max_{P_X} I(X;Y) = \max_{P_X}\min_{Q_Y} D(P_{Y|X}\|Q_Y|P_X) = \min_{Q_Y}\max_{P_X} D(P_{Y|X}\|Q_Y|P_X).$$

   课上我们定义了

$$D(P_{Y|X}\|Q_{Y|X}|P_X) = \sum_x P_X(x)D(P_{Y|X=x}\|Q_{Y|X=x}) = D(P_XP_{Y|X}\|P_XQ_{Y|X}).$$

   类似地, 可以自然地定义 $D(P_{Y|X}\|Q_Y|P_X)$, 只需将 Q_Y 视作一个退化的 kernel

$$D(P_{Y|X}\|Q_Y|P_X) = \sum_x P_X(x)D(P_{Y|X=x}\|Q_Y) = D(P_XP_{Y|X}\|P_XQ_Y).$$

61. (10 分) 完成以下关于 Chernoff bound 的证明.
   (1) (0 分) 设随机变量 $(X_1,\ldots,X_n)\sim(\mathrm{Bern}(p))^n$, 即它们独立地服从 $\mathrm{Bern}(p)$. 对任意 $t>0$,

$$\Pr\Big[\frac{X_1+\cdots+X_n}{n}\geq q\Big] = \Pr[e^{t(X_1+\cdots+X_n)}\geq e^{tqn}] \underset{\text{Markov's bound}}{\leq} \frac{\mathbb{E}[e^{t(X_1+\cdots+X_n)}]}{e^{tqn}} = \Big(\frac{\mathbb{E}[e^{tX_1}]}{e^{tq}}\Big)^n.$$

   当 $0\leq p\leq q\leq 1$ 时, 请选取合适的 t 使得上式最紧. 得到的结果应为

$$\Pr\Big[\frac{X_1+\cdots+X_n}{n}\geq q\Big] \leq \exp(-n\cdot d(q\|p)).$$

   Remark: 对称地, 当 $0\leq q\leq p\leq 1$ 时, 可以证明

$$\Pr\Big[\frac{X_1+\cdots+X_n}{n}\leq q\Big] \leq \exp(-n\cdot d(q\|p)).$$

(2) 设 $(X_1,\dots,X_n)\sim P_1P_2\dots P_n$, 即它们相互独立. 每个 P_i 都是 $[0,1]$ 上的期望等于 p 的分布. 证明当 $0\le p\le q\le 1$ 时,

$$\Pr\Big[\frac{X_1+\cdots+X_n}{n}\ge q\Big]\le \exp(-n\cdot d(q\|p)).$$

提示: 比较 $\mathbb{E}_{X\sim P_i}[e^{tX}]$ 和 $\mathbb{E}_{X\sim \mathrm{Bern}(p)}[e^{tX}]$ 的大小.

(3) 有 $m>n$ 个球, 其中 pm 个是白球. 从中无放回的随机选取 n 个球. 用随机变量 $(X_1,\dots,X_n)$ 表示这 n 次选取的结果. $X_i=1$ 表示第 i 个球是白球, $X_i=0$ 表示第 i 个球不是白球. 显然 $\mathbb{E}[X_i]=p$. 证明当 $0\le p\le q\le 1$ 时,

$$\Pr\Big[\frac{X_1+\cdots+X_n}{n}\ge q\Big]\le \exp(-n\cdot d(q\|p)).$$

62. (10 分) 根据 Sanov's Theorem 我们可以看出, Chernoff bound 对于

$$\Pr_{(X_1,\dots,X_n)\sim(\mathrm{Bern}(p))^n}\Big[\frac{X_1+\cdots+X_n}{n}\ge q\Big]$$

的估计已经很精确, 指数上的系数是紧的. 这个估计对非 Bernoulli 分布是否也同样精确?

考虑有限个正实数上的分布 P. 记 $\mathrm{Supp}(P)=\{v_1,\dots,v_T\}\subseteq\mathbb{R}^+$. 记 $p_i:=P(v_i)>0$. 这个分布的期望是 $\bar v=\sum p_iv_i$. 考虑任意 $b\in(\bar v,\max_i v_i)$, 定义

$$Q^*=\arg\min_{\substack{\text{分布 } Q\\ \mathbb{E}_{X\sim Q}[X]\ge b}} D(Q\|P).$$

根据 Sanov's Theorem,

$$\Pr_{(X_1,\dots,X_n)\sim P^n}\Big[\frac{X_1+\cdots+X_n}{n}\ge b\Big]\le (n+1)^T\cdot\exp(-n\cdot D(Q^*\|P)).$$

而根据 Chernoff bound,

$$\Pr_{(X_1,\dots,X_n)\sim P^n}\Big[\frac{X_1+\cdots+X_n}{n}\ge b\Big]\le \min_{t>0}\Big(\frac{\mathbb{E}_{X\sim P}[e^{tX}]}{e^{tb}}\Big)^n.$$

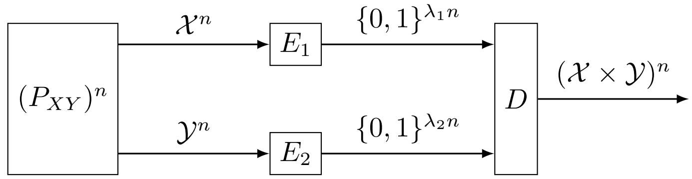
请问是否存在 P 和 $b\in(\bar v,\max_i v_i)$ 使得 Chernoff bound 的估计要弱于 Sanov's Theorem?

提示: 拉格朗日乘数.

63. (5 分) 设 $X_1,\dots,X_n,Y_1,\dots,Y_n$ 是 $2n$ 个独立的随机变量. $X_i\sim\mathrm{Bern}(4/5)$, $Y_i\sim\mathrm{Bern}(1/2)$.

求最小的不依赖于 n 的常数 c 使得 $\Pr[\sum_{i=1}^n X_i\le\sum_{i=1}^n Y_i]\le O(c^n)$.

64. (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.

65. (10 分) $[\ell, n, d]$-纠错码可以由其编码函数 $E : \{0,1\}^n \to \{0,1\}^\ell$ 定义, 满足

$$\forall \text{distinct } x, y \in \{0,1\}^n,\ \Delta(E(x), E(y)) \ge d.$$

这里 Δ 表示汉明距离 (Hamming distance).

(1) 证明存在常数 α. 当 $\ell > 2d$ 且 $\ell \ge 2n + \alpha d \ln(\ell/d)$ 时, 存在 $[\ell, n, d]$-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 $\ell > 2d$ 且 $\ell \ge n + \alpha d \ln(\ell/d)$ 时, 存在 $[\ell, n, d]$-纠错码.

66. (5 分) $[\ell, n, d]_p$-纠错码可以由其编码函数 $E : \mathbb{Z}_p^n \to \mathbb{Z}_p^\ell$ 定义, 满足

$$\forall \text{distinct } x, y \in \mathbb{Z}_p^n,\ \Delta(E(x), E(y)) \ge d.$$

这里 Δ 表示汉明距离 (Hamming distance).

(1) (0 分) 证明当 p 为素数幂且 $p \ge \ell$ 时, 存在 $[\ell, n, \ell - n + 1]_p$-纠错码.

(2) 证明当 $n + d > \ell + 1$ 时, 不存在 $[\ell, n, d]_p$-纠错码.

67. (15 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

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

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 2 题.

68. (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.

69. (6 分) 对于一个布尔函数 $f:\{0,1\}^n\to\{0,1\}$, 函数第 i 位的影响被定义为

$$\mathrm{Influence}_i(f) := \Pr_{x\leftarrow\{0,1\}^n}\big[f(x)\ne f(x\oplus e_i)\big],$$

其中 $e_i=(0,\dots,0,1,0,\dots,0)$ 只在第 i 位等于 1.

(1) 布尔函数 f 被称为**单调**（monotone）, 如果 $x\ge y \implies f(x)\ge f(y)$. （这里 $x\ge y$ 表示 $\forall i, x_i\ge y_i$.）请寻找一个单调的布尔函数 f, 使得 $\sum_i \mathrm{Influence}_i(f)$ 最大, 并证明.

(2) 布尔函数 f 被称为**平衡**（balanced）, 如果 $\Pr_{x\leftarrow\{0,1\}^n}[f(x)=1]=\frac{1}{2}$. 请寻找一个平衡的布尔函数 f, 使得 $\max_i \mathrm{Influence}_i(f)$ 尽量小, 可以忽略常数系数.

Remark: 证明 $\max_i \mathrm{Influence}_i(f)$ 的下界需要非常有技巧地使用傅里叶变换. 本题不需要证明结果最优.

70. (12 分) 对于函数 $f:\{0,1\}^n\to\mathbb{R}$, 用 f 的傅里叶系数表示如下量

(1) $\widehat{g_s}(x)$, 其中 $g_s(x):=f(x\oplus s)$.

(2) $\widehat{g_y}(x)$, 其中 $g_y(x):=(-1)^{\langle x,y\rangle}f(x)$.

(3) $\widehat{f_i}(x)$, 其中 $f_i(x):=f(x)-f(x\oplus e_i)$.

(4) $\widehat{g_k}(x)$, 其中 $g_k(x_1,\dots,x_k):=f(x_1,\dots,x_k,0,\dots,0)$.

(5) $\widehat{g_a}(x)$, 其中 $g_a(x_1,\dots,x_k):=\mathbb{E}_{y\leftarrow\{0,1\}^{n-k}}[f(x,y)\chi_a(y)]$.

(6) $\mathrm{Var}[f(X)]$, 其中 X 服从均匀分布.

这里约定 $\hat{f}(a) = \mathbb{E}_x[f(x)\overline{\chi_a(x)}]$, $f(x) = \sum_a \hat{f}(a)\chi_a(x)$.

71. (4 分) 给定一个函数 $f : \{0,1\}^n \to [-1,1]$, 你能看到若干独立采样 $(x, f(x))$（其中 x 在 $\{0,1\}^n$ 中均匀分布）.

(1) 需要多少个采样才能以 $1-\delta$ 的信心和不错过 ε 的绝对误差估计 $\hat{f}(a)$. 换言之, 需要设计一个估计方法使得 $\Pr[|(\text{估计值}) - \hat{f}(a)| > \varepsilon] \le \delta$.

(2) 问题改为估计 $\sum_{a\in\{0,1\}^n} \hat{f}(a)$, 需要多少个采样.

这里约定 $\hat{f}(a) = \mathbb{E}_x[f(x)\overline{\chi_a(x)}]$, $f(x) = \sum_a \hat{f}(a)\chi_a(x)$.

提示: 可以使用作业中 Chernoff bound 的变种.

72. (10 分) 令 $X_1, \ldots, X_{2n}$ i.i.d. 服从 Bern(1/2) 分布. 我们要选取一组系数 $c_1, \ldots, c_{2n} \in \mathbb{Z}$, 使得 $\sum_i c_i X_i$ 接近均匀分布. 当然, 并不存在 $\mathbb{Z}$ 上的均匀分布, 我们实际的要求是统计距离

$$\Delta\Big(\sum_i c_i X_i, \sum_i c_i X_i + 1\Big) \le 2^{-\lambda}. \tag{*}$$

一种显然的做法, 是令 $n = \lambda/2$, 令 $c_i = 2^{i-1}$; 这样 $\sum_i c_i X_i$ 服从 $\{0, 1, \ldots, 2^\lambda - 1\}$ 上的均匀分布, 而 $\sum_i c_i X_i + 1$ 服从 $\{1, 2, \ldots, 2^\lambda\}$ 上的均匀分布, 满足我们对统计距离的要求.

进一步, 假设 $X_1, \ldots, X_n$ 中有一半值被泄漏. 要求即使已经看到泄露值, (*) 仍然成立.

为了方便分析, 我们令 $c_1, \ldots, c_n$ 是 i.i.d. 从某个分布 P_C 中选取的. 这样不管哪部分值泄露, 分析都相同. 不失一般性, 可以假设前一半值没有泄露. 定义函数

$$\mathrm{Err}(c_1, \ldots, c_n) = \Delta\Big(\sum_i c_i X_i, \sum_i c_i X_i + 1\Big)$$

我们要求当 $c_1, \ldots, c_n$ 是从 $(P_C)^n$ 选取时, 以 $1 - 2^{-\lambda}$ 的概率（这个随机性只依赖于 $c_1, \ldots, c_n$）

$$\mathrm{Err}(c_1, \ldots, c_n) \le 2^{-\lambda}.$$

请根据 λ, 选取合适的 n 以及分布 P_C, 使得要求被满足. 请让 n 的取值尽量小, 可以忽略常数系数. 建议选取 P_C 为 $\{1, 2, 3, \ldots, B\}$ 上的均匀分布, 其中 $B = 2^{O(\lambda)}$ 根据 λ 选取.

73. (5 分) 对于函数 $f : \{0,1\}^n \to \mathbb{R}$, 它的傅里叶系数 $\hat{f} : \{0,1\}^n \to \mathbb{R}$ 满足

$$\hat{f}(y) = \frac{1}{2^n} \sum_x f(x)\chi_y(x), \qquad f(x) = \sum_y \hat{f}(y)\chi_y(x).$$

考虑一个“噪音算子”$T_\rho : (\{0,1\}^n \to \mathbb{R}) \to (\{0,1\}^n \to \mathbb{R})$, 其中 $\rho \in [0,1]$.

$$T_\rho(f)(x) = \mathop{\mathbb{E}}_{y\sim(\mathrm{Bern}(\rho))^n}\Big[f(x \oplus y)\Big].$$

求 $\widehat{T_\rho(f)}(y)$. 化简后的表达式不应该出现 $\sum$ 或 $\mathbb{E}$.

74. (6 分) 考虑一个有限状态空间 Ω 上不可约的（irreducible）马尔可夫核 P. 我们知道 P 存在唯一的稳态分布（stationary distribution）π. 证明对于任意初始分布 μ

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\mu P^j=\pi.$$

75. (4 分) 考虑大小为 n 的有限状态空间 Ω 上的一个不可约（irreducible）马尔可夫核 P, 令 π 是稳态分布.

    (1) 证明 P 只有一个特征值等于 1.

    (2) 假设 P 有周期 $T>1$. 这里 $T=\gcd\{t:\exists x, P^t(x|x)>0\}$. 不难证明, 周期性说明状态空间可以划分为 T 个非空子集 $\Omega_0,\Omega_1,\dots,\Omega_{T-1}$ 满足

    $$\forall x,y\in\Omega,\ \forall j\in\mathbb{Z}_T,\ P(y|x)>0\wedge x\in\Omega_j \implies y\in\Omega_{j+1}.$$

    证明 1 的所有 T 次根 $e^{2\pi i\frac{k}{T}}$（for $k\in\mathbb{Z}_T$）都是 P 的特征值.

76. (6 分) 考虑 $\mathbb{Z}$ 上的随机游走. 马尔可夫核是

$$P(x+1|x)=p,\quad P(x-1|x)=1-p$$

其中 $p\in(0,1)$ 是参数. 请计算这个马尔可夫链返回初始点的概率.

$$\Pr[\exists i>0 \text{ such that } X_i=X_0].$$

77. (6 分) 考虑 $\mathbb{Z}^d$ 上的随机游走. 马尔可夫核是

$$P(y_1,\dots,y_d|x_1,\dots,x_d)=\begin{cases}1/3^d, & \text{if } \forall i, |y_i-x_i|\le 1\\ 0, & \text{otherwise}\end{cases}$$

这个马尔可夫核在各个维度上独立, 便于分析. 证明

$$\Pr[\exists i>0 \text{ such that } X_i=X_0]=\begin{cases}1, & \text{if } d=2\\ 1-\Omega(1), & \text{if } d>2\end{cases}$$

提示：考虑

$$\mathbb{E}[\text{number of } i>0 \text{ such that } X_i=X_0].$$

78. (6 分) 有 n 种不同卡片. 可以从一个抽卡机中, 每次独立地获得一张随机卡片.

    (1) 期望需要抽多少次卡, 才能收集到每种卡片至少一张.

    (2) 需要抽多少次卡, 才能以至少 $1-1/n$ 的概率收集到每种卡片至少一张.（给出一个尽量紧的上界即可. 可以有常数倍的放松.）

79. (10 分) 简单图 G 中有 n 个点, 最大度数记为 Δ. 用 $C>5\Delta$ 种颜色对 G 随机点染色, 要求任意一对相邻点的染色不同. 为了均匀采样一个随机染色, 我们使用 MCMC 方法. 马尔可夫核是:

- 假设当前染色为 $f: V \to C$.
- 随机选取一个点 $v \in C$, 随机选取一个颜色 $c \in C$. (TODO 明年改成随机选取一个邻居没有的颜色)
- 如果 v 的邻居的颜色都不是 c, 就将 v 的染色修改为 c; 否则保持染色不变.

请估算混合时间 $\tau(\varepsilon)$, 给出一个尽量好的上届.

$$\tau(\varepsilon) = \text{smallest } t \text{ s.t. } d(t) \le \varepsilon$$
$$d(t) = \max_x \Delta_{\mathrm{TV}}(P^t(x,\cdot),\pi)$$

注. *如果想用 coupling 分析 $C > 2\Delta$ 的情形, 建议用 $S_t \subseteq V$ 表示 t 时刻 coupling 中两个染色一致的点集. 考虑被 S_t 切的边 (一个端点在 S_t 中, 另一个端点在 S_t 外) 有怎样的影响.*

80. (5 分) 对一个马尔可夫链 P, 用 π 表示它的一个稳态分布, 用 $\tau(\varepsilon)$ 表示它的混合时间.

$$\tau(\varepsilon) = \text{smallest } t \text{ s.t. } d(t) \le \varepsilon$$
$$d(t) = \max_x \Delta_{\mathrm{TV}}(P^t(x,\cdot),\pi)$$

证明, 对任意 $\varepsilon > 0$, $\tau(2\varepsilon^2) \le 2\tau(\varepsilon)$.

81. (10 分) 随机图 $G(n,p)$ 是连通的概率是多少?
    (1) 证明存在常数 $\alpha > 0$, 使得对所有充分大的 n, $G(n, \alpha \ln n/n)$ 是连通图的概率不高于 $1/n$.
    (2) 证明存在常数 $\alpha > 0$, 使得对所有充分大的 n, $G(n, \alpha \ln n/n)$ 是连通图的概率不低于 $1 - 1/n$.
    提示: $G(n,p)$ 可以如此采样: 先从二项分布 $\mathrm{Binom}(\binom{n}{2}, p)$ 中采样 m, 再从 $G(n,m)$ 中采样. $G(n,m)$ 可以如此采样: 从没有边的图开始, 每次随机添加一条新边, 重复 m 次.

82. (8 分) 用 $G(2n,p,q)$ 表示如下的随机图的分布: 点集为 $V = \{1,2,\ldots,2n\}$. 选取随机的 $S^* \subseteq V$ 且 $|S^*| = n$. 这样点集被划分为了两个大小相同的块 S 和 $V \setminus S$. 对于任意两个点 (u,v), 如果 u,v 在同一个块中（$u,v \in S^*$ 或者 $u,v \notin S^*$）, 那么 u,v 之间以概率 p 有边相连; 如果 u,v 在不同的块中, 那么 u,v 之间以概率 q 有边相连.

考虑高度稀疏的场景. 令 $p = \alpha/n$, $q = \beta/n$, 其中 $\alpha,\beta \in \mathbb{R}$ 是常数. 请问是否可以依据图本身的信息对 S^* 进行一个非平凡的估计. 具体来说, 是否存在一个算法 $\mathcal{M}$ 和常数 $c > 0$ 使得

$$\Pr\left[|\hat{S}| = n \wedge \frac{|\hat{S} \cap S^*|}{n} \notin (1/2 - c, 1/2 + c) \middle| \begin{array}{c} G \leftarrow G(2n, \alpha/n, \beta/n) \\ \hat{S} \leftarrow \mathcal{M}(G) \end{array}\right] = 1 - O(1/n).$$

答案显然依赖于 (α,β) 的取值. 请找到一个尽量大的 (α,β) 的范围使得可以对 S^* 进行上述的非平凡估计.

提示: 当 $(\alpha-\beta)^2 < \alpha+\beta$ 时, 不存在任何算法可以非平凡地估计 S^* [Mossel-Neeman-Sly 2012].

83. (8 分) 有若干集合 $A_1,\ldots,A_n,B_1,\ldots,B_n$, 满足 $\forall i,j \;\; A_i \cap B_j = \varnothing \iff i = j$. 证明

$$\sum_i \frac{1}{\binom{|A_i|+|B_i|}{|A_i|}} \le 1.$$

84. (5 分) 考虑如下生成社交网络过程: 初始时只有一人; 每当一个新人加入时, 会随机关注一个人, 被关注概率正比于当前被关注次数加上 β. 这里 $\beta > 0$ 是一个常数. 请问当网络中有 n 人时, 第 k 个加入网络的人的被关注次数的期望约是多少.

85. (8 分) 如果集合 $S \subseteq \{0,1\}^n$ 满足

$$\forall \text{distinct } a,b,c \in S,\ \Delta(a,b) + \Delta(b,c) > \Delta(a,c),$$

我们称 S 是不共线的. 这里 Δ 表示汉明距离（Hamming distance）.

(1) 证明, 对所有足够大的 n, 存在不共线的 $S \subseteq \{0,1\}^n$ 满足 $|S| \geq 1.01^n$.

(2) 证明, 对所有足够大的 n, 任何不共线的 $S \subseteq \{0,1\}^n$ 都满足 $|S| \leq 1.99^n$.

86. (5 分) 假设 n 足够大. 证明存在不依赖 n 的常数 $\alpha > 0$, 使得

采样有 αn 条边的随机图 $G \sim G(n, \alpha n)$, 并采样两个不同的随机点 u, v. 那么 u, v 在 G 上联通的概率不超过 $1/n$.