

概率论进阶作业

林晓烁 2024 春

<https://xiaoshuo-lin.github.io>

习题 1 计算 Wigner 半圆律

$$w(x) = \frac{1}{2\pi} \sqrt{4-x^2}, \quad -2 \leq x \leq 2$$

的 k 阶矩.

解答 设 $X \sim w(x)$.

(1) 若 k 为奇数, $\mathbb{E}[X^k] = \int_{-2}^2 x^k w(x) dx = 0$.

(2) 若 k 为偶数, 设 $k = 2m$, 则

$$\begin{aligned} \mathbb{E}[X^k] &= \frac{1}{2\pi} \int_{-2}^2 x^{2m} \sqrt{4-x^2} dx = \frac{1}{2\pi} \int_0^2 (x^2)^{m-\frac{1}{2}} (4-x^2)^{\frac{1}{2}} dx^2 \\ &\stackrel{x^2=4t}{=} \frac{1}{2\pi} \int_0^1 (4t)^{m-\frac{1}{2}} (4-4t)^{\frac{1}{2}} d(4t) = \frac{2^{2m+1}}{\pi} \int_0^1 t^{m-\frac{1}{2}} (1-t)^{\frac{1}{2}} dt \\ &= \frac{2^{2m+1}}{\pi} \mathbf{B}\left(m+\frac{1}{2}, \frac{3}{2}\right) = \frac{2^{2m+1}}{\pi} \cdot \frac{\Gamma(m+\frac{1}{2})\Gamma(\frac{3}{2})}{\Gamma(m+2)} \\ &= \frac{2^{2m+1}}{\pi} \cdot \frac{\Gamma(2m)\Gamma(\frac{3}{2})}{\frac{2^{2m-1}}{\sqrt{\pi}}\Gamma(m)\Gamma(m+2)} = \frac{4}{\sqrt{\pi}} \cdot \frac{(2m-1)!\frac{\sqrt{\pi}}{2}}{(m-1)!(m+1)!} \\ &= \frac{1}{m+1} \binom{2m}{m}. \end{aligned} \quad \square$$

习题 2 (Gaussian Unitary Ensemble, GUE) 复矩阵 $X = (x_{ij})_{i,j=1}^n, \{\operatorname{Re} x_{ij}, \operatorname{Im} x_{ij}\}_{i,j=1}^n \stackrel{\text{iid}}{\sim} N(0, 1)$, 令

$$H = \frac{1}{2}(X + X^*).$$

证明:

(1) H 矩阵元联合密度为

$$f(H) = 2^{-\frac{n}{2}} \pi^{-\frac{n^2}{2}} e^{-\frac{1}{2} \operatorname{tr} H^2}.$$

(2) 酉群不变性, 即任给酉矩阵 $U \in \mathbf{U}(n)$ 有

$$UHU^{-1} \stackrel{D}{=} H.$$

证明 (1) 设 $H = (h_{ji})_{i,j=1}^n$, 则 $h_{ii} = \operatorname{Re} x_{ii} \sim N(0, 1)$, $h_{ij} = \frac{\operatorname{Re} x_{ij} + i \operatorname{Im} x_{ij} + \operatorname{Re} x_{ji} + i \operatorname{Im} x_{ji}}{2} \sim N_{\mathbb{C}}(0, 1)$. 故

$$f(H) dH = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{h_{ii}^2}{2}} dh_{ii} \prod_{1 \leq i < j \leq n} \frac{1}{\pi} e^{-|h_{ij}|^2} dh_{ij},$$

这里 $dH = \prod_{1 \leq i < j \leq n} dh_{ij}$. 因此 H 矩阵元联合密度为

$$f(H) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{h_{ii}^2}{2}} \prod_{1 \leq i < j \leq n} \frac{1}{\pi} e^{-|h_{ij}|^2} = 2^{-\frac{n}{2}} \pi^{-\frac{n^2}{2}} \prod_{i,j=1}^n e^{-\frac{|h_{ij}|^2}{2}} = 2^{-\frac{n}{2}} \pi^{-\frac{n^2}{2}} e^{-\frac{1}{2} \operatorname{tr} H^2}.$$

(2) 一方面, 对任意 $U \in \mathbf{U}(n)$, $\operatorname{tr}((UHU^{-1})^2) = \operatorname{tr}(UH^2U^{-1}) = \operatorname{tr}(H^2U^{-1}U) = \operatorname{tr} H^2$; 另一方面, 线性

变换 $H \mapsto UHU^{-1}$ 的 Jacobi 矩阵为 $U \otimes \bar{U}$, 设 $P, Q \in U(n)$ 分别将 U, \bar{U} 酉对角化:

$$PUP^{-1} = I_n, \quad Q\bar{U}Q^{-1} = I_n,$$

则

$$\begin{aligned} \det(U \otimes \bar{U}) &= \det\left((P \otimes Q)(U \otimes \bar{U})(P \otimes Q)^{-1}\right) = \det((PUP^{-1}) \otimes (Q\bar{U}Q^{-1})) \\ &= \det(I_n \otimes I_n) = 1. \end{aligned}$$

故 $UHU^{-1} \stackrel{D}{=} H$. □

习题 3 (Wishart 模型) $X = (x_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$, $\{x_{ij}\} \stackrel{\text{iid}}{\sim} N(0, 1)$, 假设 $\alpha = n - p \geq 0$, α 固定, 证明:

$$\frac{1}{p} \mathbb{E} \left[\text{tr} \left(\frac{XX^T}{p} \right)^m \right] \rightarrow C_m = \frac{1}{m+1} \binom{2m}{m}.$$

证明 首先

$$\frac{1}{p} \mathbb{E} \left[\text{tr} \left(\frac{XX^T}{p} \right)^m \right] = \frac{1}{p^{m+1}} \sum_{\substack{1 \leq i_1, \dots, i_m \leq p \\ 1 \leq j_1, \dots, j_m \leq n}} \mathbb{E}[x_{i_1 j_1} x_{i_2 j_1} x_{i_2 j_2} \cdots x_{i_m j_m} x_{i_1 j_m}].$$

由于 $\{x_{ij}\} \stackrel{\text{iid}}{\sim} N(0, 1)$, 只需考虑每条出现的“边”至少出现两次的求和项. 将每一个求和项视作顶点集 $\{i_1, \dots, i_m\}$ 与 $\{j_1, \dots, j_m\}$ 上的二部图, 总边数为 $2m$. 分别用 n_i, n_j 表示这两个顶点集中互异顶点的个数, 则 $n_i + n_j \leq m + 1$.

(1) 若 $n_i + n_j \leq m$.

- ① 对 $\mathbf{i} = (i_1, \dots, i_m)$ 定义用以描述 \mathbf{i} 中相同分量分布的 \mathbf{t}_i , 例如, 若 $\mathbf{i} = (2, 5, 2, 1, 1)$, 则 $\mathbf{t}_i = (1, 2, 1, 3, 3)$. 类似地, 对 \mathbf{j} 定义 \mathbf{t}_j . 由于 $\{x_{ij}\} \stackrel{\text{iid}}{\sim} N(0, 1)$, 具有相同 $\mathbf{t}_i, \mathbf{t}_j$ 的顶点集对应的求和项相同.
- ② 若某个 \mathbf{t}_i 对应 n_i 个不同的顶点, 具有这样的 \mathbf{t}_i 的 \mathbf{i} 的可能数 $p(p-1) \cdots (p-n_i+1) < p^{n_i}$. 同理, 若某个 \mathbf{t}_j 对应 n_j 个不同的顶点, 具有这样的 \mathbf{t}_j 的 \mathbf{j} 的可能数 $n(n-1) \cdots (n-n_j+1) < n^{n_j}$. 故 $n_i + n_j \leq m$ 对应的可能数不超过 $p^{n_i} n^{n_j} = p^{n_i} (p+\alpha)^{n_j} < Cp^{n_i+n_j} \leq Cp^m$, 这里 C 是一个依赖于 m, α 但与 p 无关的常数.
- ③ $n_i + n_j \leq m$ 的所有类型对题中 LHS 的贡献为 $\frac{1}{p^{m+1}} O(p^m) = O\left(\frac{1}{p}\right) \xrightarrow{p \rightarrow \infty} 0$.

(2) 若 $n_i + n_j = m + 1$.

- ① 此二部图共有 m 条边, 从而是树, 因此

$$\frac{1}{p^{m+1}} \mathbb{E}[x_{i_1 j_1} x_{i_2 j_1} x_{i_2 j_2} \cdots x_{i_m j_m} x_{i_1 j_m}] = \frac{1}{p^{m+1}} [\text{Var}(X_1)]^m = \frac{1}{p^{m+1}}.$$

- ② 若某个 \mathbf{t}_i 对应 n_i 个顶点, 某个 \mathbf{t}_j 具有 n_j 个顶点, 具有这样的 $\mathbf{t}_i, \mathbf{t}_j$ 的 \mathbf{i}, \mathbf{j} 的可能数为 $p(p-1) \cdots (p-n_i+1)n(n-1) \cdots (n-n_j+1)$. 由于 $\alpha = n - p \geq 0$ 为定值, 当 $p \rightarrow \infty$ 时 $n \approx p$, 因此 $n_i + n_j = m + 1$ 类型的可能数为 $p^{n_i} n^{n_j} + o(1) = p^{m+1} + o(1)$.
- ③ 由上述讨论知 LHS $\xrightarrow{p \rightarrow \infty} \#\{(\mathbf{t}_i, \mathbf{t}_j) : \text{对应的 } n_i + n_j = m + 1\}$. 每一闭路 $i_1 j_1 i_2 j_2 \cdots i_m j_m i_1$ 均对应于一个长度为 $2m$ 的类型序列, 其第 j 项为前 j 次移动中首次经过的边的数目与非首次经

过的边的数目, 例如 132524231 的类型序列为 12323210. 我们关注的情形对应的类型序列总是起始于 1 而终止于 0, 且相邻两项相差 ± 1 . 注意到类型序列中奇数项意味着经过这次移动后到达 j 顶点集, 偶数项则意味着到达 i 顶点集.

- ④ 注意到我们关注的 $(\mathbf{t}_i, \mathbf{t}_j)$ 一一对应于长度为 $2m$ 的类型序列, 记此数目为 β_m . 设一个长为 $2m$ 的类型序列在第 $2j$ 处首次变为 0, 则从第 $2j+1$ 到第 $2m$ 处为一个长为 $2m-2j$ 的类型序列, 而通过将前 $2j$ 的子序列去除首尾、余下每个数 -1 , 可将该长为 $2j$ 的 (有限制的) 类型序列对应于长为 $2j-2$ 的 (无限制的) 类型序列. 补充定义 $\beta_0 = 1$, 则有递推关系

$$\beta_m = \sum_{j=1}^m \beta_{j-1} \beta_{m-j}.$$

由此可得

$$\begin{aligned} \beta_1 &= \beta_0^2 = 1 = C_1, \\ \beta_2 &= 2\beta_0\beta_1 = 2 = C_2, \\ \beta_3 &= 2\beta_2 + \beta_1^2 = 5 = C_3, \\ &\dots \end{aligned}$$

一般地, 由 Catalan 数满足递推关系 $C_m = \sum_{j=1}^m C_{j-1}C_{m-j}$ 及 $C_0 = 1$ 可知 $\beta_m = C_m$.

综上所述, 我们证明了

$$\frac{1}{p} \mathbb{E} \left[\text{tr} \left(\frac{XX^T}{p} \right)^m \right] \rightarrow C_m = \frac{1}{m+1} \binom{2m}{m}. \quad \square$$

习题 4 实对称 Wigner 矩阵 $A_n = (a_{ij})_{i,j=1}^n$, $\{a_{ij} : i \leq j\}$ 独立, $\mathbb{E}[a_{ij}] = 0$, $\text{Var}(a_{ij}) = 1$, 高阶矩一致有界, 令

$$\|A_n\|_2 = \sup_{\|\mathbf{v}\|=1, \mathbf{v} \in \mathbb{R}^n} \|A_n \mathbf{v}\|.$$

证明: 对任意 $\delta > 0$ 有

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\|A_n\|_2 \geq n^{\frac{1}{2} + \delta} \right) = 0.$$

证明 利用

$$\max_i |\lambda_i(A_n)|^{2k} \leq \sum_i |\lambda_i(A_n)|^{2k} = \text{tr}(A_n^{2k})$$

与 $\|A_n\|_2 = \max_i \sigma_i(A_n) = \max_i |\lambda_i(A_n)|$ 即得

$$\begin{aligned} \mathbb{P} \left(\|A_n\|_2 \geq n^{\frac{1}{2} + \delta} \right) &= \mathbb{P} \left(\max_i |\lambda_i(A_n)| \geq n^{\frac{1}{2} + \delta} \right) = \mathbb{P} \left(\max_i |\lambda_i(A_n)|^{2k} \geq n^{2k(\frac{1}{2} + \delta)} \right) \\ &\leq \mathbb{P} \left(\text{tr}(A_n^{2k}) \geq n^{2k(\frac{1}{2} + \delta)} \right), \end{aligned}$$

再由 Markov 不等式,

$$\mathbb{P} \left(\|A_n\|_2 \geq n^{\frac{1}{2} + \delta} \right) \leq \mathbb{P} \left(\text{tr}(A_n^{2k}) \geq n^{2k(\frac{1}{2} + \delta)} \right) \leq \frac{\mathbb{E}[\text{tr}(A_n^{2k})]}{n^{2k(\frac{1}{2} + \delta)}}.$$

而

$$\frac{1}{n} \mathbb{E} \left[\operatorname{tr} \left(\frac{A_n}{\sqrt{n}} \right)^{2k} \right] = \frac{1}{n^{k+1}} \mathbb{E} [\operatorname{tr}(A_n^{2k})] \xrightarrow{n \rightarrow \infty} C_k = \frac{1}{k+1} \binom{2k}{k} \implies \mathbb{E} [\operatorname{tr}(A_n^{2k})] \sim C_k n^{k+1},$$

因此只需取满足 $2k\delta > 1$ 的正整数 k 即可得结论. \square

习题 5 (Hermite Wigner 矩阵)

$$A_n = (a_{ij})_{i,j=1}^n, \quad A_n = A_n^*,$$

- ▷ 实 $\{a_{ii}\} \stackrel{\text{iid}}{\sim} Y$.
- ▷ $\{\operatorname{Re} a_{ij}, \operatorname{Im} a_{ij}\}_{i < j} \stackrel{\text{iid}}{\sim} Z$.
- ▷ $\{a_{ii} : 1 \leq i \leq n\} \cup \{\operatorname{Re} a_{ij}, \operatorname{Im} a_{ij} : 1 \leq i < j \leq n\}$ 独立.
- ▷ $\mathbb{E}[Y] = 0, \mathbb{E}[Z] = 0, \operatorname{Var}(Y) < +\infty, \operatorname{Var}(Z) = \frac{1}{2}$.
- ▷ $\mathbb{E}[|Y|^k], \mathbb{E}[|Z|^k] < +\infty, \forall k \geq 3$.

证明: 类似于实 Wigner 矩阵情形, 我们有

$$\frac{1}{n} \mathbb{E} \left[\operatorname{tr} \left(\frac{A_n}{\sqrt{n}} \right)^k \right] \rightarrow \gamma_k = \int_{-2}^2 x^k \frac{1}{2\pi} \sqrt{4-x^2} dx.$$

证明 首先

$$\operatorname{tr}(A_n^k) = \sum_{i_1, i_2, \dots, i_k} a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_k i_1}.$$

由于 $\mathbb{E}[Y] = \mathbb{E}[Z] = 0$, 因此只需考虑每条出现的“边”至少出现两次的求和项.

- (1) 若 k 为奇数, 设 $k = 2m + 1$, 则每个求和项至多有 m 条“边”, 从而至多有 $m + 1$ 个“顶点”. 注意到 Y, Z 的各阶矩均有限, 因此

$$\text{LHS} = \frac{1}{n^{m+\frac{3}{2}}} O(n^{m+1}) = O\left(\frac{1}{\sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} 0 = \gamma_k.$$

- (2) 若 k 为偶数, 设 $k = 2m$, 则每个求和项中每条“边”恰出现两次, 且恰有 $m + 1$ 个“顶点”. 对这样的不自交路径 $(i_1, i_2, \dots, i_{2m-1}, i_{2m})$ 构造随机游走 $(b_1, b_2, \dots, b_{2m})$: 令

$$b_j = \begin{cases} 1, & \text{若 } i_j i_{j+1} \text{ 在游走中首次出现,} \\ -1, & \text{其他.} \end{cases}$$

令 $S_i = S_{i-1} + b_i, S_0 = 0$, 则对称随机游走 $\{S_n\}$ 从 0 出发经过 $2m$ 步回到 0, 且 $S_i \geq 0$ ($i = 1, \dots, 2m$), 这样的轨道数为 C_m . 因此

$$\text{LHS} = \frac{C_m}{n^{m+1}} n(n-1) \cdots (n-m) + o(1) \xrightarrow{n \rightarrow \infty} C_m = \gamma_k. \quad \square$$

习题 6 证明标准正态分布被其矩序列决定.

证明 标准正态分布的 k 阶矩为

$$\gamma_k = \begin{cases} 0, & k \text{ 为奇数,} \\ (k-1)!!, & k \text{ 为偶数.} \end{cases}$$

由 Wallis 公式与 Stirling 公式可知, 当 $k \rightarrow \infty$ 时,

$$\frac{1}{k} \gamma_{2k}^{\frac{1}{2k}} = \frac{1}{k} [(2k-1)!!]^{\frac{1}{2k}} \sim \frac{\left(\frac{2^k k!}{\sqrt{k\pi}}\right)^{\frac{1}{2k}}}{k} = \frac{\sqrt{2} \left(\frac{k!}{\sqrt{k\pi}}\right)^{\frac{1}{2k}}}{k} \sim \frac{\sqrt{2} \left[\frac{\sqrt{2k\pi} \left(\frac{k}{e}\right)^k}{\sqrt{k\pi}}\right]^{\frac{1}{2k}}}{k} = \frac{2^{\frac{1}{4k} + \frac{1}{2}}}{\sqrt{k}e} \rightarrow 0,$$

这说明标准正态分布的矩序列满足 Riesz 条件, 因此标准正态分布被其矩序列决定. \square

习题 7 求 Wigner 半圆律

$$\rho(x) = \frac{1}{2\pi} \sqrt{4-x^2}, \quad x \in [-2, 2]$$

的 k 阶矩, 并验证其决定 $\rho(x)$.

解答 习题 1 已求得 Wigner 半圆律的 k 阶矩为

$$\gamma_k = \begin{cases} 0, & k \text{ 为奇数,} \\ \frac{1}{m+1} \binom{2m}{m}, & k = 2m. \end{cases}$$

由 Stirling 公式, 当 $k \rightarrow \infty$ 时,

$$\frac{1}{k} \gamma_{2k}^{\frac{1}{2k}} \sim \frac{1}{k} \left\{ \frac{1}{k+1} \cdot \frac{\sqrt{4k\pi} \left(\frac{2k}{e}\right)^{2k}}{\left[\sqrt{2k\pi} \left(\frac{k}{e}\right)^k\right]^2} \right\}^{\frac{1}{2k}} \sim \frac{2}{k} (k\pi)^{-\frac{1}{4k}} \sim \frac{2}{k} \rightarrow 0,$$

这说明矩序列 $\{\gamma_k\}$ 满足 Riesz 条件, 因此其决定了 $\rho(x)$. \square

习题 8 序列 $\gamma_{2k+1} = 0, \gamma_{2k} = 1$ 是否对应随机变量矩序列?

解答 由于 $\limsup_{k \rightarrow \infty} \frac{1}{k} \gamma_{2k}^{\frac{1}{2k}} < +\infty$, 满足 Riesz 条件. 假设矩序列 $\{\gamma_k\}$ 对应于随机变量 X , 则其特征函数

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \sum_{k=0}^{\infty} \mathbb{E}\left[\frac{(itX)^k}{k!}\right] = \sum_{k=0}^{\infty} \frac{(it)^k \gamma_k}{k!} = \sum_{m=0}^{\infty} \frac{(-1)^m t^{2m}}{(2m)!} = \cos t,$$

这是两点分布 $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = \frac{1}{2}$ 的特征函数. \square

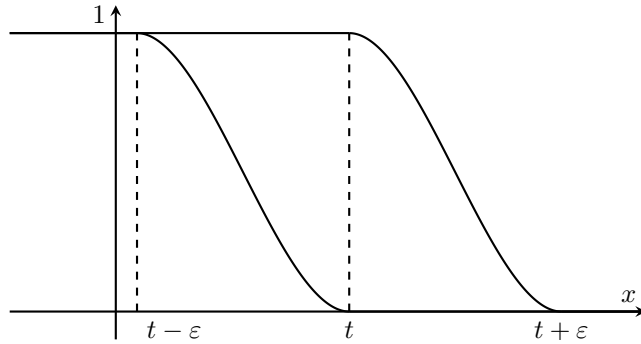
习题 9 设 $\{X_k\}$ 为独立同分布随机变量序列, $\mathbb{E}[X_1] = 0, \text{Var}(X_1) = 1, \mathbb{E}[|X_1|^3] < \infty$, 试用 Lindeberg 替换法证明 CLT 的收敛速度

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k \leq t\right) - \Phi(t) \right| = O\left(n^{-\frac{1}{8}}\right),$$

这里 $\Phi(t)$ 表示标准正态分布函数. 附注: 右边的指数 $\frac{1}{8}$ 可以改进到最优指数 $\frac{1}{2}$.

证明 取截断函数 $\varphi(x)$, 使得当 $x \leq 0$ 时 $\varphi(x) = 1$, 当 $x \geq 1$ 时 $\varphi(x) = 0$, 且 $\varphi(x) \in C^\infty(\mathbb{R})$. 令

$$\varphi_{t-\varepsilon, \varepsilon}(x) = \varphi\left(\frac{x-t+\varepsilon}{\varepsilon}\right), \quad \varphi_{t, \varepsilon}(x) = \varphi\left(\frac{x-t-\varepsilon}{\varepsilon}\right).$$



如上图, 有 $\mathbf{1}_{(-\infty, t-\varepsilon]}(x) \leq \varphi_{t-\varepsilon, \varepsilon}(x) \leq \mathbf{1}_{(-\infty, t]}(x) \leq \varphi_{t, \varepsilon}(x) \leq \mathbf{1}_{(-\infty, t+\varepsilon]}(x)$, 且

$$|\varphi_{t-\varepsilon, \varepsilon}'''(x)| = \frac{1}{\varepsilon^3} |\varphi'''(\frac{x-t+\varepsilon}{\varepsilon})| = O(\varepsilon^{-3}), \quad |\varphi_{t, \varepsilon}'''(x)| = \frac{1}{\varepsilon^3} |\varphi'''(\frac{x-t-\varepsilon}{\varepsilon})| = O(\varepsilon^{-3}).$$

取与 $\{X_k\}$ 独立的随机变量列 $Y, Y_1, Y_2, \dots \stackrel{\text{iid}}{\sim} N(0, 1)$, 令

$$Z_{n,k} = \sum_{1 \leq i < k} X_i + \sum_{k < i \leq n} Y_i,$$

则

$$\begin{aligned} Z_{n,n} + X_n &= \sum_{i=1}^n X_i, & Z_{n,1} + Y_1 &= \sum_{i=1}^n Y_i, \\ Z_{n,k} + X_k &= Z_{n,k+1} + Y_{k+1}, & k &= 1, 2, \dots, n-1. \end{aligned}$$

记 $S_n = \sum_{k=1}^n X_k$, 则

$$\mathbb{E} \left[\varphi_{t, \varepsilon} \left(\frac{S_n}{\sqrt{n}} \right) \right] - \mathbb{E}[\varphi_{t, \varepsilon}(Y)] = \sum_{k=1}^n \left\{ \mathbb{E} \left[\varphi_{t, \varepsilon} \left(\frac{Z_{n,k} + X_k}{\sqrt{n}} \right) \right] - \mathbb{E} \left[\varphi_{t, \varepsilon} \left(\frac{Z_{n,k} + Y_k}{\sqrt{n}} \right) \right] \right\},$$

对任意取定的 k , $Z_{n,k}$ 与 X_k, Y_k 独立, 从而

$$\mathbb{E} \left[\varphi'_{t, \varepsilon} \left(\frac{Z_{n,k}}{\sqrt{n}} \right) \frac{X_k - Y_k}{\sqrt{n}} \right] = \mathbb{E} \left[\varphi''_{t, \varepsilon} \left(\frac{Z_{n,k}}{\sqrt{n}} \right) \frac{X_k^2 - Y_k^2}{n} \right] = 0.$$

经 Taylor 展开可知

$$\mathbb{E} \left[\varphi_{t, \varepsilon} \left(\frac{S_n}{\sqrt{n}} \right) \right] - \mathbb{E}[\varphi_{t, \varepsilon}(Y)] = n \cdot O \left(\frac{\mathbb{E}[|X|^3] + \mathbb{E}[|Y|^3]}{n^{\frac{3}{2}}} \varepsilon^{-3} \right) = O(n^{-\frac{1}{2}} \varepsilon^{-3}).$$

同理

$$\mathbb{E} \left[\varphi_{t-\varepsilon, \varepsilon} \left(\frac{S_n}{\sqrt{n}} \right) \right] - \mathbb{E}[\varphi_{t-\varepsilon, \varepsilon}(Y)] = O(n^{-\frac{1}{2}} \varepsilon^{-3}).$$

故

$$\begin{aligned} \mathbb{P} \left(\frac{S_n}{\sqrt{n}} \leq t \right) &= \mathbb{E} \left[\mathbf{1}_{(-\infty, t]} \left(\frac{S_n}{\sqrt{n}} \right) \right] \leq \mathbb{E} \left[\varphi_{t, \varepsilon} \left(\frac{S_n}{\sqrt{n}} \right) \right] = \mathbb{E}[\varphi_{t, \varepsilon}(Y)] + O(n^{-\frac{1}{2}} \varepsilon^{-3}) \\ &\leq \mathbb{E}[\mathbf{1}_{(-\infty, t+\varepsilon]}(Y)] + O(n^{-\frac{1}{2}} \varepsilon^{-3}) = \Phi(t + \varepsilon) + O(n^{-\frac{1}{2}} \varepsilon^{-3}). \end{aligned}$$

同理可得

$$\begin{aligned} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \leq t\right) &= \mathbb{E}\left[\mathbf{1}_{(-\infty, t]}\left(\frac{S_n}{\sqrt{n}}\right)\right] \geq \mathbb{E}\left[\varphi_{t-\varepsilon, \varepsilon}\left(\frac{S_n}{\sqrt{n}}\right)\right] = \mathbb{E}[\varphi_{t-\varepsilon, \varepsilon}(Y)] + O\left(n^{-\frac{1}{2}}\varepsilon^{-3}\right) \\ &\geq \mathbb{E}[\mathbf{1}_{(-\infty, t-\varepsilon]}(Y)] + O\left(n^{-\frac{1}{2}}\varepsilon^{-3}\right) = \Phi(t-\varepsilon) + O\left(n^{-\frac{1}{2}}\varepsilon^{-3}\right). \end{aligned}$$

综上所述可得

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k \leq t\right) = \Phi(t) + O\left(\varepsilon + n^{-\frac{1}{2}}\varepsilon^{-3}\right),$$

取 $\varepsilon = n^{-\frac{1}{8}}$ 就有

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k \leq t\right) - \Phi(t) = O\left(n^{-\frac{1}{8}}\right) \implies \sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(\frac{1}{\sqrt{n}} \sum_{k=1}^n X_k \leq t\right) - \Phi(t) \right| = O\left(n^{-\frac{1}{8}}\right). \quad \square$$

习题 10 对取值 0 和 1 的 Bernoulli 分布求熵.

解答 记 $p = \mathbb{P}(X = 1)$, 则 $H(X) = -p \ln p - (1-p) \ln(1-p)$. □

习题 11 投掷两枚均匀的骰子, X 表示骰子点数之和, 求熵 $H(X)$.

解答 用 X_i 表示第 i 枚骰子的点数 ($i = 1, 2$), 则 $X = X_1 + X_2$ 的概率母函数

$$\begin{aligned} G_X(s) &= [G_{X_1}(s)]^2 = \left(\sum_{k=1}^6 \frac{s^k}{6}\right)^2 = \left[\frac{1}{6} \cdot \frac{s(1-s^6)}{1-s}\right]^2 \\ &= \frac{s^{12}}{36} + \frac{s^{11}}{18} + \frac{s^{10}}{12} + \frac{s^9}{9} + \frac{5s^8}{36} + \frac{s^7}{6} + \frac{5s^6}{36} + \frac{s^5}{9} + \frac{s^4}{12} + \frac{s^3}{18} + \frac{s^2}{36}. \end{aligned}$$

由此可得

$$\begin{aligned} H(X) &= -2 \left(\frac{1}{36} \ln \frac{1}{36} + \frac{1}{18} \ln \frac{1}{18} + \frac{1}{12} \ln \frac{1}{12} + \frac{1}{9} \ln \frac{1}{9} + \frac{5}{36} \ln \frac{5}{36} \right) - \frac{1}{6} \ln \frac{1}{6} \\ &= \frac{23}{18} \ln 2 + \frac{5}{3} \ln 3 - \frac{5}{18} \ln 5. \end{aligned} \quad \square$$

习题 12 离散型随机变量 X , 证明任给一元 Borel 可测函数 g , 均有 $H(g(X)) \leq H(X)$.

证明 我们有

$$\begin{aligned} H(g(X)) &= - \sum_i \mathbb{P}(g(X) = i) \ln \mathbb{P}(g(X) = i) \\ &= - \sum_i \left(\sum_{x: g(x)=i} \mathbb{P}(X = x) \right) \ln \left(\sum_{x: g(x)=i} \mathbb{P}(X = x) \right) \\ &\leq - \sum_i \sum_{x: g(x)=i} \mathbb{P}(X = x) \ln \mathbb{P}(X = x) \\ &= - \sum_x \mathbb{P}(X = x) \ln \mathbb{P}(X = x) \\ &= H(X). \end{aligned} \quad \square$$

习题 13 令 $D = \{x : f(x) > 0\}$, 证明结论 (2) 和 (3).

(1) $D = (-\infty, +\infty)$, $\mathbb{E}[X] = 0$, $\text{Var}(X) = 1$, 正态分布熵最大, 为 $\ln \sqrt{2\pi e}$.

(2) $D = (0, +\infty)$, $\mathbb{E}[X] = \frac{1}{\lambda}$, 指数分布 $\text{Exp}(\lambda)$ 熵最大, 为 $\ln \frac{e}{\lambda}$.

(3) $D = (0, a)$, 均匀分布熵最大, 为 $\ln a$.

证明 由 Gibbs 不等式, 对概率密度函数 $f(x), g(x)$, 有

$$\int_{\mathbb{R}} [f(x) - f(x) \ln f(x)] dx \leq \int_{\mathbb{R}} [g(x) - f(x) \ln g(x)] dx.$$

(2) 取 $g(x) = \lambda e^{-\lambda x}$, 若 $\int_{\mathbb{R}} x f(x) dx = \frac{1}{\lambda}$, 则有

$$-\int_{\mathbb{R}} f(x) \ln f(x) dx \leq -\int_{\mathbb{R}} f(x) \ln(\lambda e^{-\lambda x}) dx = -\ln \lambda + \lambda \int_{\mathbb{R}} x f(x) dx = \ln \frac{e}{\lambda},$$

后者恰为指数分布 $\text{Exp}(\lambda)$ 的熵.

(3) 取 $g(x) = \frac{1}{a} \mathbf{1}_{[0, a]}(x)$, 则有

$$-\int_0^a f(x) \ln f(x) dx \leq -\int_0^a f(x) \ln \frac{1}{a} dx = \ln a,$$

后者恰为均匀分布 $U(0, a)$ 的熵. □

习题 14 $\mu_1(\Omega)$ 表示有限样本空间 Ω 上的概率分布 (测度) 全体, 对 $\mu \in \mu_1(\Omega)$ 其熵

$$H(\mu) = -\sum_{\omega \in \Omega} \mu(\omega) \ln \mu(\omega).$$

试证对 $\mu, \nu \in \mu_1(\Omega)$, $\alpha \in [0, 1]$ 有

$$H(\alpha\mu + (1-\alpha)\nu) \geq \alpha H(\mu) + (1-\alpha)H(\nu).$$

证明 令 $f(x) = -x \ln x$, 则 $f'(x) = -\ln x - 1$, $f''(x) = -\frac{1}{x} < 0$ ($0 < x < 1$), 因此 $f(x)$ 为 $(0, 1)$ 上的凹函数. 对任意 $\omega \in \Omega$, 有

$$-[\alpha\mu + (1-\alpha)\nu](\omega) \ln[\alpha\mu + (1-\alpha)\nu](\omega) \geq -\alpha\mu(\omega) \ln \mu(\omega) - (1-\alpha)\nu(\omega) \ln \nu(\omega).$$

由此对 $\omega \in \Omega$ 求和即得证. □

习题 15 设维数 $d = 1$, 周期 Ising 模型对应的 Hamilton 量为

$$H(\sigma) = -J \sum_{k=1}^N \sigma_k \sigma_{k+1} - h \sum_{k=1}^N \sigma_k,$$

其中 $\sigma_{N+1} = \sigma_1$, 考虑 $M_N = \sum_{k=1}^N \sigma_k$.

(1) 证明：对一般的磁化强度 h , 我们有

$$\frac{M_N}{N} \xrightarrow{P} \frac{\sinh(\beta h)}{\sqrt{e^{-4\beta J} + \sinh^2(\beta h)}}.$$

(2) 对一般的磁化强度 h , $\frac{M_N}{N}$ 对应的中心极限定理定理是否仍然成立?

证明 考虑

$$\begin{aligned} \mathbb{E}\left[e^{z \frac{M_N}{N^\delta}}\right] &= \sum_{\sigma_1, \dots, \sigma_N} e^{z \frac{M_N}{N^\delta}} \frac{1}{Z_{N;\beta,h}} e^{-\beta H(\sigma)} \\ &= \frac{1}{Z_{N;\beta,h}} \sum_{\sigma_1, \dots, \sigma_N} \exp\left\{\frac{z}{N^\delta} \sum_{k=1}^N \sigma_k + \beta J \sum_{k=1}^N \sigma_k \sigma_{k+1} + \beta h \sum_{k=1}^N \sigma_k\right\} \\ &= \frac{1}{Z_{N;\beta,h}} \sum_{\sigma_1, \dots, \sigma_N} \exp\left\{\beta J \sum_{k=1}^N \sigma_k \sigma_{k+1} + \beta \left(h + \frac{z}{\beta N^\delta}\right) \sum_{k=1}^N \sigma_k\right\} \\ &= \frac{Z_{N;\beta,h+\frac{z}{\beta N^\delta}}}{Z_{N;\beta,h}}, \end{aligned}$$

我们有

$$Z_{N;\beta,h} = \text{tr} \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix}^N = (\lambda_h^+)^N + (\lambda_h^-)^N,$$

其中转移矩阵的两特征值

$$\begin{aligned} \lambda_h^+ &= e^{\beta J} \left(\cosh(\beta h) + \sqrt{\sinh^2(\beta h) + e^{-4\beta J}} \right), \\ \lambda_h^- &= e^{\beta J} \left(\cosh(\beta h) - \sqrt{\sinh^2(\beta h) + e^{-4\beta J}} \right). \end{aligned}$$

由于 $|\lambda_h^+| > |\lambda_h^-|$, 当 $N \rightarrow \infty$ 时,

$$\mathbb{E}\left[e^{z \frac{M_N}{N^\delta}}\right] = \frac{\left(\lambda_{h+\frac{z}{\beta N^\delta}}^+\right)^N + \left(\lambda_{h+\frac{z}{\beta N^\delta}}^-\right)^N}{\left(\lambda_h^+\right)^N + \left(\lambda_h^-\right)^N} \sim \left(\frac{\lambda_{h+\frac{z}{\beta N^\delta}}^+}{\lambda_h^+}\right)^N = \left(1 + \frac{\lambda_{h+\frac{z}{\beta N^\delta}}^+ - \lambda_h^+}{\lambda_h^+}\right)^N.$$

(1) 取 $\delta = 1$, 有

$$\begin{aligned} \mathbb{E}\left[e^{z \frac{M_N}{N}}\right] &\sim \left(1 + \frac{\lambda_{h+\frac{z}{\beta N}}^+ - \lambda_h^+}{\lambda_h^+}\right)^N \sim \left(1 + \frac{\frac{z}{\beta N} \cdot \frac{d\lambda_h^+}{dh}}{\lambda_h^+}\right)^N = \left(1 + \frac{\frac{z}{\beta} \cdot \frac{d \ln \lambda_h^+}{dh}}{N}\right)^N \\ &\xrightarrow{N \rightarrow \infty} \exp\left(\frac{z}{\beta} \cdot \frac{d \ln \lambda_h^+}{dh}\right), \end{aligned}$$

这说明

$$\frac{M_N}{N} \xrightarrow{P} \frac{1}{\beta} \cdot \frac{d \ln \lambda_h^+}{dh} = \frac{\sinh(\beta h)}{\sqrt{e^{-4\beta J} + \sinh^2(\beta h)}}.$$

(2) 记 $\mu = \frac{1}{\beta} \cdot \frac{d \ln \lambda_h^+}{dh}$. 取 $\delta = \frac{1}{2}$, 有

$$\begin{aligned} \mathbb{E} \left[e^{z \frac{M_N - N\mu}{\sqrt{N}}} \right] &\sim \left(1 + \frac{\lambda_h^+ + \frac{z}{\beta\sqrt{N}} - \lambda_h^+}{\lambda_h^+} \right)^N e^{-z\sqrt{N}\mu} \sim \left(1 + \frac{\frac{d\lambda_h^+}{dh} \cdot \frac{z}{\beta\sqrt{N}} + \frac{1}{2} \cdot \frac{d^2\lambda_h^+}{dh^2} \cdot \frac{z^2}{\beta^2 N}}{\lambda_h^+} \right)^N e^{-z\sqrt{N}\mu} \\ &= \exp \left\{ N \ln \left(1 + \frac{\frac{d\lambda_h^+}{dh} \cdot \frac{z}{\beta\sqrt{N}} + \frac{1}{2} \cdot \frac{d^2\lambda_h^+}{dh^2} \cdot \frac{z^2}{\beta^2 N}}{\lambda_h^+} \right) - z\sqrt{N}\mu \right\} \\ &\sim \exp \left(\frac{1}{2\lambda_h^+} \cdot \frac{d^2\lambda_h^+}{dh^2} \cdot \frac{z^2}{\beta^2} \right). \end{aligned}$$

因此

$$\frac{M_N - N\mu}{B\sqrt{N}} \xrightarrow{D} N(0, 1),$$

其中

$$B^2 = \frac{1}{\lambda_h^+ \beta^2} \cdot \frac{d^2\lambda_h^+}{dh^2} = \frac{\cosh(\beta h) + e^{4\beta J} \sinh^2(\beta h) \sqrt{\sinh^2(\beta h) + e^{-4\beta J}}}{\sqrt{\sinh^2(\beta h) + e^{-4\beta J}} [e^{4\beta J} \sinh^2(\beta h) + 1]} \quad \square$$

习题 16 证明: 存在 $c_-, c_+ > 0$, 使得对任意 $m \in A_N := \{-1 + \frac{2k}{N} : k = 0, 1, \dots, N\}$, 均有

$$c_- N^{-\frac{1}{2}} e^{NS(m)} \leq \binom{N}{\frac{1+m}{2}N} \leq c_+ N^{\frac{1}{2}} e^{NS(m)},$$

这里 $S(m) = -\frac{1-m}{2} \ln \frac{1-m}{2} - \frac{1+m}{2} \ln \frac{1+m}{2}$.

证明 当 $m = \pm 1$ 时, $S(m) = 0$, $\binom{N}{\frac{1+m}{2}N} = 0$, 结论成立. 当 $m \in A_N \setminus \{\pm 1\}$ 时, 由 Stirling 公式可得渐近展开

$$\begin{aligned} \binom{N}{\frac{1+m}{2}N} &= \frac{N!}{\left(\frac{1+m}{2}N\right)! \left(\frac{1-m}{2}N\right)!} \sim \frac{\sqrt{2\pi N} \left(\frac{N}{e}\right)^N}{\sqrt{2\pi \left(\frac{1+m}{2}N\right)} \left(\frac{1+m}{2} \frac{N}{e}\right)^{\frac{1+m}{2}N} \sqrt{2\pi \left(\frac{1-m}{2}N\right)} \left(\frac{1-m}{2} \frac{N}{e}\right)^{\frac{1-m}{2}N}} \\ &= \sqrt{\frac{2}{\pi N(1-m^2)}} e^{NS(m)}. \end{aligned}$$

由此及 $-1 + \frac{1}{N} < m < 1 - \frac{1}{N}$ 即知结论成立. □

习题 17 设 $\zeta : \mathbb{N} \rightarrow \mathbb{R}_{>0}$. 考虑如下对 Curie-Weiss 模型 Hamilton 量的修正 ($h = 0$):

$$H(\sigma) = -\frac{d}{\zeta(N)} \left(\sum_{i=1}^N \sigma_i \right)^2, \quad \mu_{N;\beta}(\sigma) = \frac{1}{Z_{N;\beta}} e^{-\beta H(\sigma)}.$$

证明:

(1) 若 $\lim_{N \rightarrow \infty} \frac{\zeta(N)}{N} = \infty$, 则对任意 $\beta \geq 0$, $\frac{1}{N} \sum_{i=1}^N \sigma_i \xrightarrow{P} 0$.

(2) 若 $\lim_{N \rightarrow \infty} \frac{\zeta(N)}{N} = 0$, 则对任意 $\beta > 0$, $\frac{1}{N} \left| \sum_{i=1}^N \sigma_i \right| \xrightarrow{P} 1$.

证明 首先

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N\sigma_i = m\right) = \mathbb{P}\left(\sum_{i=1}^N\sigma_i = Nm\right) = \frac{1}{Z_{N;\beta}} e^{\frac{d\beta}{\zeta(N)}(Nm)^2} \binom{N}{\frac{1+m}{2}N}.$$

由习题 16 中的渐近展开, 对上式取对数即得

$$\ln\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N\sigma_i = m\right) = N\left(\frac{d\beta N}{\zeta(N)}m^2 + S(m) + o(1)\right).$$

- (1) 若 $\lim_{N\rightarrow\infty}\frac{\zeta(N)}{N} = \infty$, 上式 RHS 以 $S(m)$ 为主项. 利用凸函数 $x\ln x$ 的 Jensen 不等式可得 $S(m)$ 在 $m=0$ 时取最大值 $\ln 2$, 因此当 $m \neq 0$ 时,

$$\ln\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N\sigma_i = m\right) - \ln\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N\sigma_i = 0\right) = N(-\lambda + o(1)), \quad \lambda > 0,$$

也即

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N\sigma_i = m\right) = e^{-\lambda N + o(N)}\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^N\sigma_i = 0\right), \quad \forall m \in A_N \setminus \{0\}.$$

故此时对任意 $\beta \geq 0$, 有 $\frac{1}{N}\sum_{i=1}^N\sigma_i \xrightarrow{P} 0$.

- (2) 若 $\lim_{N\rightarrow\infty}\frac{\zeta(N)}{N} = 0$, 且 $\beta > 0$, 上式 RHS 以 $\frac{d\beta N}{\zeta(N)}m^2$ 为主项, 由于 $m \in [-1, 1]$, 当 $m^2 = 1$ 时主项最大, 同 (1) 分析知此时 $\frac{1}{N}\sum_{i=1}^N\sigma_i \xrightarrow{P} 1$. □

习题 18 对 Curie–Weiss 模型证明: 极限 $\psi_\beta^{\text{CW}}(h) := \lim_{N\rightarrow\infty}\frac{1}{N}\ln Z_{N;\beta,h}^{\text{CW}}$ 存在, 且关于 h 为凸函数. 进一步证明: $\psi_\beta^{\text{CW}}(h) = \max_{-1 \leq m \leq 1} \{\beta hm + \beta dm^2 + S(m)\}$.

证明 此时配分函数为

$$Z_{N;\beta,h}^{\text{CW}} = \sum_{m \in A_N} \sum_{\substack{\sigma \in \Omega_N: \\ \frac{1}{N}\sum_{i=1}^N\sigma_i = m}} e^{-\beta H_{N;\beta,h}^{\text{CW}}(\sigma)} = \sum_{m \in A_N} \binom{N}{\frac{1+m}{2}N} e^{\beta dm^2 N + \beta hm N}.$$

由习题 16 中的渐近展开可得

$$\frac{1}{N}\ln\left\{\binom{N}{\frac{1+m}{2}N} e^{\beta dm^2 N + \beta hm N}\right\} \sim S(m) + \beta dm^2 + \beta hm.$$

取 $m_0 \in [-1, 1]$ 使此展开式 RHS 取最大值, 则其余 $m \neq m_0$ 对应项均以指数速度小于 m_0 对应项, 因此极限 $\psi_\beta^{\text{CW}}(h)$ 存在, 且 $\psi_\beta^{\text{CW}}(h) = \max_{-1 \leq m \leq 1} \{\beta hm + \beta dm^2 + S(m)\}$. 任取 h_1, h_2 , 有

$$\begin{aligned} \psi_\beta^{\text{CW}}\left(\frac{h_1+h_2}{2}\right) &= \max_{-1 \leq m \leq 1} \{\beta m \frac{h_1+h_2}{2} + \beta dm^2 + S(m)\} \\ &= \frac{1}{2} \max_{-1 \leq m \leq 1} \{[\beta h_1 m + \beta dm^2 + S(m)] + [\beta h_2 m + \beta dm^2 + S(m)]\} \\ &\leq \frac{1}{2} [\psi_\beta^{\text{CW}}(h_1) + \psi_\beta^{\text{CW}}(h_2)], \end{aligned}$$

故 $\psi_{\beta}^{\text{CW}}(h)$ 关于 h 为凸函数.

□