Table 1: $X+Y$ 的分布列

2	3	4	5	6	7	8	9	10	11	12
$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

1 第三次习题解答

1. 设 X 为两个骰子的点数之和所对应的随机变量，求 X 对应的熵 $H(X)$.

Proof. 直接写出骰子点数和的分布列，从而可以直接计算熵 $H(X)$. □

2. 设 X 为离散型随机变量，证明对于任意的函数 g，$H(g(X)) \leq H(X)$.

Proof.

$$\begin{aligned} H(g(X)) &= -\sum_{i \in I} \mathbb{P}(g(X)=i) \log \mathbb{P}(g(X)) \\ &= -\sum_{i} \sum_{X:g(X)=i} \mathbb{P}(X) \log(\sum_{X:g(X)=i} \mathbb{P}(X)) \\ &\leq -\sum_{i} \sum_{X:g(X)=i} \mathbb{P}(X) (\sum_{X:g(X)=i} \log \mathbb{P}(X)) \\ &= \sum_{X} \mathbb{P}(X) \log \mathbb{P}(X). \end{aligned} \tag{1}$$

□

3. 设 X 是取值在 $[0,\infty)$ 上的随机变量，满足 $\mathbb{E}(X)=\frac{1}{\lambda}$，求证指数分布在这些限制条件下取得熵的最大值，为 $\ln\frac{e}{\lambda}$.

Proof. 用 Gibbs 不等式，若 $f(x)$，$g(x) \geq 0$，$\int f(x)=1, \int g(x)=1$ 则

$$-\int f(x) \log f(x) \leq -\int f(x) \log g(x).$$

取 $g(x)=\lambda e^{-\lambda x}$. 则有

$$-\int f(x) \log f(x) \leq -\int f(x) \log(\lambda e^{-\lambda x}) = \log\frac{e}{\lambda}.$$

□

4. 设 X 是取值在 $[0,a]$ 上的随机变量，求证 X 为均匀分布时熵最大，为 $\ln a$.

Proof. 用 Gibbs 不等式，取 $g(x)=\frac{1}{a}1_{\{x\in[0,a]\}}$, 我们有：

$$-\int f(x) \log f(x) \leq \int f(x) \log a = \log a.$$

□

2 第四次习题解答

1. 设维数 $d=1$, 考虑此时的周期 Ising model, $H(\sigma)=-J\sum_{k=1}^{N}\sigma_k\sigma_{k+1}-h\sum_{k=1}^{N}\sigma_k$, 探索 $M_N=\sum_{k=1}^{N}\sigma$ 对应的大数律和中心极限定理。

Proof. 我们考虑 M_N 的矩母函数:

$$< e^{t\frac{M_N}{N^\delta}} > = \frac{Z_{N;\beta,h+\frac{t}{\beta N^\delta}}}{Z_{N;\beta,h}}. \tag{2}$$

于是我们只需要考虑 $Z_{N;\beta,h}$ 的计算。类似 $h=0$ 时的 Ising model 的计算，我们知道其为:

$$Z_{N;\beta,h}=TrA^N \tag{3}$$

其中

$$A=\begin{pmatrix} e^{\beta J+\beta h} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J-\beta h}\end{pmatrix} \tag{4}$$

计算其特征值为

$$\lambda_h=\frac{e^{\beta J+\beta h}+e^{\beta J-\beta h}\pm\sqrt{(e^{\beta J+\beta h}-e^{\beta J-\beta h})^2+4e^{-2\beta J}}}{2}$$

由于 N 会很大，我们只要考虑 $\frac{\lambda^N_{h+\frac{t}{\beta N^\delta}}}{\lambda_h^N}$ 的渐近行为即可。如取 $\delta=1$，我们可以得到 M_N 的大数定律现象

$$\log < e^{tM_N/N} > \to \frac{d\log\lambda_h}{dh}\frac{t}{\beta}.$$

也就是 $\frac{M_N}{N}\to\frac{d\log\lambda_h}{\beta dh}$.

我们再考虑其波动，考虑

$$\log < e^{\frac{M_N}{N^\delta}} > -N^{1-\delta}\frac{d\log\lambda_h}{\beta dh}$$

用 $\log\lambda_h$ 二阶可导，将其展开到第二阶，我们知道 δ 要取 $\frac{1}{2}$，并且此时有中心极限定理。 □

2. 设 $\zeta:\mathbb{N}\to\mathbb{R}^+$，考虑如下修改后的 Curie-Weiss 哈密顿量 ($h=0$):

$$H_{N;\beta,0}=-\frac{\beta}{\zeta(N)}\sum_{i,j=1}^{N}\omega_i\omega_j.$$

设 $m_N=\frac{M_N}{N}$，说明有如下结论:

a. 如果 $\lim_{N\to\infty}\frac{\zeta(N)}{N}=\infty$, 则 $m_N\to 0$ 依概率收敛对于所有 $\beta\geq 0$.

b. 如果 $\lim_{N\to\infty}\frac{\zeta(N)}{N}=0$, 则 $|m_N|\to 1$ 依概率收敛对于所有 $\beta>0$.

这说明只有当 $\zeta(N)$ 和 N 同阶时 m_N 才会有与 β 有关的非平凡表现。

Proof. 由于

$$H_{N;\beta,0} = -\frac{\beta}{\zeta(N)} M_N^2 = -\frac{\beta}{\zeta(N)} N^2 m_N^2.$$

我们可以得到

$$\mathbb{P}(m_N = m) \propto e^{\frac{\beta}{\zeta(N)} N^2 m^2} \binom{N}{\frac{1+m}{2}N}.$$

我们要考虑 $\mathbb{P}(m_N)$ 的渐近行为可以从 $e^{\frac{\beta}{\zeta(N)} N^2 m_N^2} \binom{N}{\frac{1+m}{2}N}$ 的渐近行为考虑。由 Stirling 公式

$$\begin{aligned}\binom{N}{\frac{1+m}{2}N} = \frac{N!}{(\frac{1+m}{2}N)!(\frac{1-m}{2}N)!} &\sim \frac{\sqrt{2\pi N}(\frac{N}{e})^N}{2\pi N\sqrt{\frac{1-m^2}{4}}(\frac{(1+m)N}{2e})^{\frac{1+m}{2}N}(\frac{(1-m)N}{2e})^{\frac{1-m}{2}N}} \\ &\frac{1}{\sqrt{2\pi N}\sqrt{\frac{1-m^2}{4}}}((\frac{1+m}{2})^{\frac{1+m}{2}}(\frac{1-m}{2})^{\frac{1-m}{2}})^{-N}.\end{aligned} \tag{5}$$

可以看到上面这个是个指数级增加的量（注意到 $(\frac{1+m}{2})^{\frac{1+m}{2}}(\frac{1-m}{2})^{\frac{1-m}{2}} < 1$，可以取对数证明)。

从而根据 $\frac{N}{\zeta(N)}$ 的不同关系，我们可以得到

$$\log \mathbb{P}(m_N = m) \sim \frac{N\beta}{\zeta(N)} N m_N^2 - N(\frac{1+m}{2}\log(\frac{1+m}{2}) + \frac{1-m}{2}\log(\frac{1-m}{2})). \tag{6}$$

当 $\frac{N}{\zeta(N)} \to 0$ 时，右边第二项起到主要作用，从而在使 $-(\frac{1+m}{2}\log(\frac{1+m}{2}) + \frac{1-m}{2}\log(\frac{1-m}{2}))$ 最大的 $m = 0$ 外，其他部分相对于此都会有个指数衰减的速度，从而此时 m_N 会依概率收敛到 0。

而当 $\frac{N}{\zeta(N)} \to \infty$ 时，(8) 的第一项是主项，从而 m_N 会以指数速度往使得 $\frac{N\beta}{\zeta(N)} N m^2$ 最大的 m 集中，也就是 $|m_N| \to 1$ 依概率收敛。 □

注 2.1. 这种方法在物理里面叫鞍点法 *(Saddle point method)*，在概率论中叫大偏差 *(Large Deviation Principle)*。都是用来刻画随机变量朝某一事件的集中速度的。一般来说在随机变量比较好的时候都可以有指数的集中速度。可以证明对类似于 m_N 一般的随机变量 $J(m) = \lim \frac{1}{N}\log P(m_N = m)$ 也存在并且成立，这样的 $J(m)$ 如果是物理模型，可能就是对应到他的熵或者自由能等物理参数。从某种程度上这也是用概率说明物理系统往特定状态演化的方法。

3. 说明极限

$$\psi_\beta^{CW}(h) = \lim_{N\to\infty} \frac{1}{N} \log Z_{N;\beta,h}^{CW}$$

存在，并且对于 h 是凹 (凸) 函数的，进一步其等于自由能的 Legendre 变换：

$$\psi_\beta^{CW}(h) = \max_{m\in[-1,1]} \{hm - f_\beta^{CW}(m)\}.$$

Proof. 我们可以计算其对应的配分函数。

$$Z_{N;\beta,h}^{CW} = \sum_{i=1}^{N} \binom{N}{i} e^{d\beta(\frac{N-2i}{N})^2 N + h\frac{N-2i}{N}N}. \tag{7}$$

我们找到使得 $\frac{1}{N}\log\binom{N}{i}e^{d\beta(\frac{N-2i}{N})^2N+h\frac{N-2i}{N}N}$ 最大的那个 i，则其他的项会以指数的速度小于那一项。取对数之后主要项由那个 i 所给出。设 $\frac{N-2i}{N} = m$ 则我们考虑下式的渐近行为（如前一题中考虑的）

$$e^{(d\beta m^2+hm)N}\binom{N}{\frac{1-m}{2}N}.$$

用 Stirling 公式或者式 (8) 其渐近行为为:

$$\frac{1}{N}\log\mathbb{P}(m_N = m) \sim hm + d\beta m^2 - (\frac{1+m}{2}\log(\frac{1+m}{2}) + \frac{1-m}{2}\log(\frac{1-m}{2})). \tag{8}$$

可以将其他项都放成这个，只不过是多了个 $\frac{\log N}{N}$ 这个无穷小量，所以 $\psi_\beta^{CW}(h) = \frac{1}{N}\log Z_{N;\beta,h}^{CW}$ 是存在的，并且

$$\psi_\beta^{CW}(h) = \max_{m\in[-1,1]}\{hm - f_\beta^{CW}(m)\}.$$

□

*

References

[1] S. FRIEDLI, Y. VELENIK, Statistical Mechanics of Lattice Systems A Concrete Mathematical Introduction(群里发的那本书的第二章.