1 习题解答

1 说明标准正态分布被其矩序列决定.

Proof. 我们可以直接计算标准正态分布的各阶矩，然后验证它满足 Risez 条件，从而说明矩序列决定标准正态分布。

$$\gamma_{2k+1}=\int_{-\infty}^{\infty}x^{2k+1}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}=0. \tag{1}$$

并且

$$\begin{aligned}\gamma_{2k}&=\int_{-\infty}^{\infty}x^{2k}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\\&=\int_{-\infty}^{\infty}(x^2)^{\frac{2k-1}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}d\frac{x^2}{2}\\&=2\frac{2^{\frac{2k-1}{2}}}{\sqrt{2\pi}}\int_0^{\infty}y^{\frac{2k-1}{2}}e^{-y}dy\\&=\frac{2^k}{\sqrt{\pi}}\Gamma(k+\frac{1}{2})=(2k-1)!!\le 2^kk!.\end{aligned} \tag{2}$$

Hence $\liminf(\frac{1}{k}\gamma_{2k}^{\frac{1}{2k}})<\infty$. 满足 Risez 条件，从而矩序列决定了分布 $\rho(x)$. □

2 求半圆律 $\rho(x)=\frac{1}{2\pi}\sqrt{4-x^2},\quad x\in[-2,2]$ 的 k 阶矩并验证其决定 $\rho(x)$。

Proof. (From 宋京倍的作业)

$k=2m+1$ 时, $E\left[x^k\right]=\int_{-2}^{2}x^k\cdot\frac{1}{2\pi}\sqrt{4-x^2}dx=0$.

当 $k=2m$ 时,

$$\begin{aligned}E\left[x^{2m}\right]&=\frac{1}{\pi}\int_0^2 x^{2m}\sqrt{4\pi^2}dx\\&=\frac{1}{\pi}\int_0^{\frac{\pi}{2}}2^{2m}\cdot\sin^{2m}x\cdot 2\cos x\cdot 2\cos xdx\\&=\frac{1}{\pi}\cdot 4^{m+1}\int_0^{\frac{2}{2}}\sin^{2m}x\cos^2xdx\\&=\frac{1}{\pi}4^{m+1}\cdot\frac{1}{2}B\left(\frac{2m+1}{2},\frac{3}{2}\right)\\&=\frac{1}{2\pi}4^{m+1}\frac{\Gamma\left(\frac{2m+m}{2}\right)\Gamma\left(\frac{3}{2}\right)}{\Gamma(m+2)}\\&=\frac{1}{2\pi}4^{m+!}\cdot\frac{1}{2^m}(2m-1)!!\cdot\frac{1}{2}\pi\cdot\frac{1}{(m+1)!}\\&=\frac{1}{m+1}C_{2m}^m.\end{aligned} \tag{3}$$

从而

$$E\left[X^k\right]=\begin{cases}0,\ k=2m+1\\ \frac{1}{m+1}C_{2m}^m,\ k=2m.\end{cases} \tag{4}$$

Riesz 条件:

$$\begin{aligned}\ln\left(\frac{1}{k}\left(\gamma_{2k}\right)^{\frac{1}{2k}}\right) &= \ln\left(\frac{1}{k}\left(\frac{1}{k+1}C_{2k}^{k}\right)^{\frac{1}{2k}}\right) \\ &= -\ln k - \frac{1}{2k}\ln(k+1) + \frac{1}{2k}\ln\frac{(2k)!}{k!k!},\end{aligned} \tag{5}$$

由

$$\begin{aligned}\lim_{k\to\infty}\frac{\ln\left(\frac{(2k)!}{k!k!}\right)}{2k} &= \lim_{k\to\infty}\frac{\ln((2k+2)!) - \ln(2k)! + 2\ln k! - 2\ln(k+1)!}{2k+2-2k} \quad (\text{Stolz 定理}) \\ &= \lim_{k\to\infty}\frac{\ln(2k+2) + \ln(2k+1) - 2\ln(k+1)}{2} = \ln 2,\end{aligned} \tag{6}$$

从而

$$\lim_{k\to\infty}\inf(\frac{1}{k}(\gamma_{2k}^{\frac{1}{2K}})) = 0 < \infty. \tag{7}$$

满足 Risez 条件，从而矩序列决定了分布 $\rho(x)$. □

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

Proof. 显然 $\{\gamma_k\}$ 满足 Risez 条件，我们尝试通过矩来计算特征函数。假设其为随机变量 X 的矩，则:

$$\begin{aligned}\mathbb{E}(e^{itX}) &= \sum_{n=0}^{\infty}\mathbb{E}(\frac{(it)^n X^n}{n!}) \\ &= \sum_{n=0}^{\infty}\frac{(it)^n\gamma_n}{n!} \\ &= \sum_{n=0}^{\infty}\frac{(it)^2 n}{(2n)!} \\ &= \frac{e^{it}+e^{-it}}{2}.\end{aligned} \tag{8}$$

为两点分布 $\mathbb{P}(X=1) = \mathbb{P}(X=-1) = \frac{1}{2}$ 的特征函数。易验证 X 的矩满足题目条件。 □

4 设 X_k 为独立同分布随机变量列，$E[X_1] = 0, Var(X_1) = 1, E[|X_1|^3] < \infty$，试用 Linderberg 替换法证明

$$|P(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}X_k \le t) - \Phi(t)| = \mathrm{O}(n^{-\frac{1}{8}}), \quad \forall t \in \mathbb{R} \tag{9}$$

Proof. (From 宋京倍的作业)

我们先证明 lindeberg 替换定理:

定理 1.1 (Lindeberg 替换)**.**

Lindeberg 替换定理证明. 令 $Y_1, Y_2..., Y_n$ 为服从正态分布 $N(0,1)$ 的独立同随机变量，设 X_k 为方差为 1 的独立同随机变量，记 $Z_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n} X_k$. 令

$$Z_{n,i} = \frac{1}{\sqrt{n}}\left(X_1 + \cdots + X_i + Y_{i+1} + \cdots + Y_n\right). \tag{10}$$

则

$$E\varphi\left(Z_n\right) - \varphi(Y) = -\sum_{i=0}^{n-1} E\left(\varphi\left(Z_{n,i}\right) - \varphi\left(Z_{n,i+1}\right)\right), \tag{11}$$

其中

$$Z_{n,i} = S_{n,i} + \frac{Y_{i+1}}{\sqrt{n}},\ Z_{n,i+1} = S_{n,i} + \frac{X_{i+1}}{\sqrt{n}}, \tag{12}$$

$$S_{n,i} = \frac{X_1 + \cdots + X_i + Y_{i+1} + Y_n}{\sqrt{n}}. \tag{13}$$

对 φ 做泰勒展开，

$$\varphi\left(Z_{n,i}\right) = \varphi\left(S_{n,i}\right) + \varphi'\left(S_{n,i}\right)\frac{Y_{i+1}}{\sqrt{n}} + \frac{1}{2}\varphi''\left(S_{n,i}\right)\frac{Y_{i+1}^2}{n} + \frac{Y_{i+1}^3}{6n^{\frac{3}{2}}}\varphi'''(\xi). \tag{14}$$

同样对 $\varphi(Z_{n,i+1})$ 有泰勒展开：

$$\varphi\left(Z_{n,i}\right) = \varphi\left(S_{n,i}\right) + \varphi'\left(S_{n,i}\right)\frac{X_{i+1}}{\sqrt{n}} + \frac{1}{2}\varphi''\left(S_{n,i}\right)\frac{X_{i+1}^2}{n} + \frac{X_{i+1}^3}{6n^{\frac{3}{2}}}\varphi'''(\xi). \tag{15}$$

取期望，由于 X 和 Y 的前二阶矩一致，$\varphi, \varphi', \varphi''$ 有一致上界

$$E\left[\varphi\left(Z_{n,i}\right) - \varphi\left(Z_{n,i+1}\right)\right] = O\left(\frac{E\left[|X|^3\right] + E\left[|Y|^3\right]}{n^{\frac{3}{2}}}\sup_{x\in R}|\varphi'''(x)|\right). \tag{16}$$

求和得到

$$\sum_{i=0}^{n-1} E\left[\varphi\left(z_{n,1}\right) - \varphi\left(z_{n,i+1}\right)\right] = O\left(n^{-\frac{1}{2}}E\left[|X|^3\right]\sup_{x\in R}|\varphi''|\left(x\right)|\right). \tag{17}$$

□

取 bump 函数

$$\varphi(t) = \begin{cases} 1, t\in(-\infty, 0] \\ 0, t\in[1, +\infty) \end{cases} \tag{18}$$

则 $\varphi_{\epsilon 1} = \varphi(\frac{t-x+\epsilon}{\epsilon})$,$\varphi_{\epsilon 2} = \varphi(\frac{t-x-\epsilon}{\epsilon})$ 满足：

$$\varphi_{\epsilon 1} \le 1_{t\le x} \le \varphi_{\epsilon 2},\quad |\varphi_{\epsilon i}'''| = O(\epsilon^{-3}) \tag{19}$$

从而对 $\varphi_{\epsilon i}$ 用 Lindeberg 替换

$$E\left[\varphi_{\epsilon i}\left(Z_n\right)\right] = E[\varphi_{\epsilon i}(Y)] + O\left(\frac{1}{\sqrt{n}}\left(E\left(|X^3|\right)\varepsilon^{-3}\right)\right. \tag{20}$$

$$\mathbb{P}\left(Z_n \leqslant x\right) \leqslant E\left[\varphi_{\epsilon 2}\left(Z_n\right)\right] \leq \Phi(x+\epsilon)+O(\epsilon^{-3}n^{-1/2}). \tag{21}$$

$$\mathbb{P}\left(Z_n \leqslant x\right) \geq E\left[\varphi_{\epsilon 1}\left(Z_n\right)\right] \geq \Phi(x-\epsilon)+O(\epsilon^{-3}n^{-1/2}). \tag{22}$$

从而

$$\mathbb{P}\left(Z_n \leqslant x\right) = \Phi(x)+O(\epsilon+\epsilon^{-3}n^{-1/2}) \tag{23}$$

取 $\epsilon = n^{-1/8}$ 即证。 □

2 Wick 公式 (Wick's probability theorem)

定理 2.1. ?? *如果 $(X_1, ..., X_n)$ 是均值为随机向量 0 的多元正态分布，那么*

$$\mathbb{E}[X_1X_2...X_n] = \sum_{p\in P_n^2} \Pi_{\{i,j\}\in p}\mathbb{E}[X_iX_j] = \sum_{p\in P_n^2} \Pi_{\{i,j\}\in p}Cov(X_i, X_j), \tag{24}$$

其中求和 p 是对 $\{1, ..., n\}$ 的所有配对 p 进行求和（即对将 $\{1, ..., n\}$ 划分成不相交 $\{i, j\}$ 的方式求和），后面的乘积对 p 中所有 $\{i, j\}$ 对做乘积。

更一般的如果 X_i 是多元复高斯的，那么这个公式同样成立。

注 2.1. *Wick 公式能够将多点关联函数（k 个乘积的期望）转换为两点关联函数（两个乘积的期望）简便计算。*

如果 k 是奇数，左边的期望自然是等于 0，右边不存在 k 的配对，等式当然成立。

当 $k = 2m$ 时，我们可以先看 X_i 都是同一个随机变量的情况，这时

$$\mathbb{E}(X_1...X_{2m}) = \mathbb{E}(\mathcal{N}(0,1)^{2m}) = \frac{(2m)!}{(2^m m!)} = (2m-1)!! \tag{25}$$

这正好就是 $2m$ 个元素的配对方法数 $\frac{2m!}{2^m m!}$. 也就是说当 X_i 都是同一个随机变量的时候是成立的。

对于一般的多元正态分布，我们总能找到一列独立的标准正态分布 $\{Y_i\}$，使得 $X_i = \sum_j \alpha_{ij}Y_j$. 此时我们可以把 $\prod_i X_i$ 拆成 Y_j 的乘积的线性组合，这时我们对 Y_j 乘积运用已经证明的结果，然后再做线性组合，就能得到 $\mathbb{E}\prod_i X_i$ 的结论。

3 谈概率论时是在谈论什么

回顾我们已经上过的概率论课程，主要内容就是以下这些。

- 概率论里的基本概念，随机变量、独立性、期望，以及一些具体的随机变量的计算.
- 大数定律，描述独立同分布的随机变量和 Sn/n 会集中到它的期望 $\mu = \mathbb{E}(X)$ 上.
- 中心极限定理，$(Sn - n\mu)/\sqrt{n\sigma^2} \to \mathcal{N}(0, 1)$. 描述随机变量集中的波动 (fluctuation) 情况。

- （大偏差，本课不涉及）对集中速度的估计，以及取条件在小概率事件时事情会是什么样。

对于更加一般的系统，我们也同样会关注它的“大数定律”（集中现象），大数定律之后下一阶的波动“中心极限定理”，以及它的大偏差。本科概率论课程考虑的是**独立同分布**随机变量的大数定律和中心极限定理。已经在足够大的范围中被运用，比如

$$\begin{cases} \text{统计} \\ \text{大物实验测量，不确定度分析} \\ \text{物理现象} \end{cases} \tag{26}$$

比如说我们可以从中心极限定理定理看出理想气体的 Boltzmann 分布。

例 1. *理想气体可以看成粒子之间没有相互作用，因此可以把他们的速度看成是独立的随机变量。*

他们的平均速度为 0，而且粒子数目充分大，所以可以认为粒子速度分布为正态分布。而且因为系统旋转不变，所以粒子速度分布正比于

$$p(\vec{v}) \propto e^{-a|\vec{v}|_2^2}.$$

也就是

$$p(\vec{v})dv_x dv_y dv_z \propto e^{-a|\vec{v}|_2^2} dv_x dv_y dv_z,$$

做个变量替换 $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$，*我们得到*

$$p(v) \propto v^2 e^{-av^2}. \tag{27}$$

也就是理想气体粒子速率的 Boltzmann 分布。

在统计力学中可以不通过中心极限定理得到 Boltzmann 分布，而是通过等概率原理以及其所对应的最大熵假设得到 Boltzmann 分布。这种”最大熵”的哲学不只可以用于理解独立同随机变量会出现中心极限定理，而且可以用来理解随机变量直接出现**耦合**、**非独立**的情况会出现什么对应的极限分布。依照这种方式去理解，很容易可以看出为什么有些分布会具有普适性 (universality)。但是数学上去证明这些却困难会非常困难，概率论进阶这门课会介绍几个随机变量间有相互作用的模型。

例 2.

- *Random Matrices Theory，随机矩阵特征值分布*
- *信息熵的定义*
- *Ising model 以及其他的物理模型等*

3.1 以随机矩阵为例

对于随机矩阵，我们主要关注它特征值和特征向量的分布。当随机矩阵的每个矩阵元是正态分布这一特殊的分布时 (课堂上提到的 GOE)，我们可以将特征值的联合密度分布函数精确地写出来

$$p(\lambda_1, ..., \lambda_n) \propto e^{-\frac{1}{2}\sum_i \lambda_i^2} \prod_{i<j} |\lambda_i - \lambda_j|. \tag{28}$$

此时我们可以运用解析工具去精确地把它特征值分布的极限现象描述清楚（我们猜测对矩阵元不是正态分布时这个也成立，但是我们可以先从特殊情况入手把现象研究清楚），此时依照我们之前提到的概率论关心问题的几个 level，我们有

- **大数定律 1** 特征值的整体分布, 所有的特征值 λ_i 放在一起会趋向于 Semicircle Law, 也就是当 $x \in [-2, 2]$

$$\frac{\#\{i|\lambda_i \le x\}}{n} \to \frac{1}{2\pi}\int_{-2}^{x}(4-x^2)^{1/2} \quad a.s. \tag{29}$$

- **大数定律 2**

$$\lambda_{\max} \to 2 \quad a.s. \tag{30}$$

- **波动，边界情况**

$$(\lambda_{\max} - 2)n^{3/2} \to \text{Tracy-Widom distribution.} \tag{31}$$

- **波动，内部情况**当 $x \in (-2, 2)$ 时，x 附近特征值的分布会收敛到 Sine process.

这些分布和中心极限定理一样，应该是具有普适性，不依赖与矩阵元的具体分布的。

我们可以再谈一下为什么会有这几道练习题。

- Lindeberg 替换，我们可以用 Lindeberg 替换将一般随机矩阵的矩阵元换成正态分布的情况，来证明一般矩阵的极限分布和 GOE 的一样。这个实际上就是 2010 年附近 Terence Tao 一系列工作的思想和方法。

- Wick formula，这个会告诉我们随机矩阵特征值分布是如何和其他数学（比如曲面技术）联系起来的。给我们对随机矩阵现象的更多理解。

References

[1] Isserlis’ theorem, Wikipedia, https://en.wikipedia.org/wiki/Isserlis’ theorem