第一章 事件、概率

§1.1 概率

例1. 掷硬币 $\Omega=\{H,T\}$.　$A=\{H\}$

电子自旋 $\Omega=\{\uparrow,\downarrow\}$　$B=\{\downarrow\}$

掷骰子 $\Omega=\{1,2,\cdots,6\}$　$A=\{1,3,5\}$

定义1. 样本点 一次试验中可能出现的结果 ω

样本空间 全体样本点 记号 Ω

例2. 道琼斯指数 $x(t)$ $t\in[0,T]$ 连续曲线作为样本点，函数空间为样本空间（随机过程）

事件 样本空间的某子集

事件运算 $\longleftrightarrow$ 集合运算

事件 A 发生指试验结果 $\omega\in A$

Ω 必然事件　$\varnothing$ 空集 不可能事件

事件交、并、余 $\longleftrightarrow$ $A\cap B$, $A\cup B$, A^c（对立事件）

$\omega\in A\cap B \Leftrightarrow \omega\in A \,\&\, \omega\in B \Leftrightarrow A, B$ 同时发生

$\omega\in A\cup B \Leftrightarrow \omega\in A$ or $\omega\in B \Leftrightarrow A$ 发生或 B 发生

$\omega\in A^c \Leftrightarrow \omega\notin A \Leftrightarrow A^c$ 发生 $\Leftrightarrow A$ 不发生

$A\subset B$ 表示 A 发生可推知 B 亦发生

若 $A\cap B=\varnothing$，称 A 与 B 互不相容。更一般，称 $A_1, A_2, \cdots, A_n$ 互不相容，若 $A_i\cap A_j=\varnothing$, $\forall i\neq j$

Question 是否所有子集均成为随机事件？

一般对交、并、余封闭，更多要求？

例3. $\Omega=\{1,2,3,\cdots\}$ $A=\{2,4,6,\cdots\}$

涉及到无限运算.

定义2. 称 $\mathcal{F}\subset 2^{\Omega}$ 为一个 σ-代数，若（事件域、σ域）

(i) $\Omega\in\mathcal{F}$

(ii) $A\in\mathcal{F}$，则 $A^c\in\mathcal{F}$

(iii) $A_n\in\mathcal{F}$, $n=1,2,\cdots$，则 $\bigcup_{n=1}^{\infty}A_n\in\mathcal{F}$

并称二元组 $(\Omega,\mathcal{F})$ 为一个可测空间.

例4. 最小 σ 代数 $\mathcal{F}=\{\Omega,\phi\}$

最大 σ 代数 $\mathcal{F}=2^{\Omega}$（有时太大）

$A\subset\Omega$　$\mathcal{F}=\{\Omega,\phi,A,A^c\}$

当 Ω 有限时，常取 $\mathcal{F}=2^{\Omega}$

概率：直观想法：频率稳定性. 重复试验 N 次，A 发生 N_A 次，经验表明 $N\to\infty$, $\frac{N_A}{N}\to$ 常数，记为 $P(A)$. 明显：$P(\phi)=0$　$P(\Omega)=1$

且对 $A\cap B=\varnothing$. 有 $N_{A\cup B}=N_A+N_B$，进而 $P(A\cup B)=P(A)+P(B)$.

更一般，$A_1, A_2, \cdots, A_n$ 互不相容时，

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$$

定义3. 称：$P:\mathcal{F}\to\mathbb{R}$ 为 $(\Omega,\mathcal{F})$ 上的一个 $\boxed{\text{概率测度}}$，若

(i) 非负性 $\forall A\in\mathcal{F}$, $P(A)\geq 0$

(ii) 规一性：$P(\Omega)=1$

(iii) 可列可加性：$A_n\in\mathcal{F}$, $n=1,2,\cdots$ 互不相容，$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$

并称三元组 $(\Omega,\mathcal{F},P)$ 为一个 $\boxed{\text{概率空间}}$

例5. 掷硬币: $\Omega=\{H,T\}$ $\mathcal{F}=2^{\Omega}$

$P(\{H\})=p$ $P(\{T\})=q$ $p\in[0,1]$ $p+q=1$.

掷均匀骰子 $\Omega=\{1,2,3,4,5,6\}$ $\mathcal{F}=2^{\Omega}$

$\forall A\subset\Omega$, $P(A)=\frac{\#A}{6}$

注: 有限等可能、古典概型

例6. $\Omega=\{1,2,3,\cdots\}$ $\mathcal{F}=2^{\Omega}$,

$\forall A\subset\Omega$. $P(A):=\sum_{i\in A}2^{-i}$

性质:

引理1. (i) $P(A^c)=1-P(A)$

(ii) 若 $A\subset B$, 则 $P(B)=P(A)+P(B\setminus A)$ $\geqslant P(A)$

(iii) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

(iv) Jordan 公式: $P(\bigcup_{i=1}^{n}A_i)$

$=\sum_{k=1}^{n}(-1)^{k-1}\sum_{i_1<\cdots<i_k}P(A_{i_1}\cap\cdots\cap A_{i_k})$.

证明: (i) $A\cup A^c=\Omega$ (ii) $B=A\cup(B\setminus A)$

(iii) $A\cup B=A\cup(B\setminus A)=A\cup(B\setminus(A\cap B))$

利用可列可加 + (ii)

引理2. (连续性)

(i) 若 $A_1\subset A_2\subset\cdots\subset A_n\subset\cdots$ (单调增), 则

则 $P(A)=\lim_{n\to\infty}P(A_n)$ $A=\bigcup_{i=1}^{\infty}A_i=\lim_{i\to\infty}A_i$

(ii) 若 $B_1\supset B_2\supset\cdots\supset B_n\supset\cdots$, 则

则 $P(B)=\lim_{n\to\infty}P(B_n)$ $B:=\bigcap_{i=1}^{\infty}B_i=\lim_{i\to\infty}B_i$

证明 (i) $A=A_1\cup(A_2\setminus A_1)\cup(A_3\setminus A_2)\cup\cdots$

用可列可加性, $P(A)=P(A_1)+\sum_{i=1}^{\infty}P(A_{i+1}\setminus A_i)$

$=P(A_1)+\lim_{n\to\infty}\sum_{i=1}^{n-1}P(A_{i+1}\setminus A_i)$

$=P(A_1)+\lim_{n\to\infty}\sum_{i=1}^{n-1}(P(A_{i+1})-P(A_i))$

$=\lim_{n\to\infty}P(A_n)$

(ii) 类似

例7. (配对问题) 思考题: 具体 $(\Omega,\mathcal{F},P)$

	1	2	3	$\cdots$	n
随机	S_1	S_2	S_3		S_n

问: $B=\{至少一个S_i=i\}$, $P(B)=?$

分析: $A_i=\{S_i=i\}$ 则 $B=\bigcup_{i=1}^{n}A_i$, 由Jordan公式

$P(B)=\sum_{k=1}^{n}(-1)^{k-1}\sum_{i_1<\cdots<i_k}P(A_{i_1}\cap\cdots\cap A_{i_k})$

又 $P(A_{i_1}\cap\cdots\cap A_{i_k})=\frac{(n-k)!}{n!}$

$P(B)=\sum_{k=1}^{n}(-1)^{k-1}\cdot\frac{1}{k!}\xrightarrow{n\to\infty}1-\frac{1}{e}$

§1.2. 条件概率、独立性

例1. Bridge $A=\{东家有3黑桃\}$

$B=\{西家有5黑桃\}$. 西家关心东家有3张黑桃概率, 即: B发生条件下A发生概率

直观想法: 重复试验N次, 观察事件A, B, B发生条件下A发生的概率.

$\frac{N_{AB}}{N_B}=\frac{N_{AB}/N}{N_B/N}\longrightarrow\frac{P(AB)}{P(B)}$

定义1. 对 $P(B)>0$, B发生条件下A的 $\boxed{\text{条件概率}}$

$P(A|B)=\frac{P(AB)}{P(B)}$

注: 乘法公式: $P(AB)=P(B)P(A|B)=P(A)P(B|A)$

划分: 称 $B_1,\cdots,B_n$ 为 Ω 的一个划分, 若 $\forall i\neq j$, $B_iB_j\neq\emptyset$, 且 $\Omega=\bigcup_{i=1}^{n}B_i$ (变复杂为简单)

引理1. 设 $B_1,\cdots,B_n$ 为 Ω 划分, 且 $\forall i$, $P(B_i)>0$, 则 $P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$ (全概公式)

证明: $A=A\cap\Omega=\bigcup_{i=1}^{n}AB_i$

则 $P(A)=\sum_{i=1}^{n}P(AB_i)=\sum_{i=1}^{n}P(B_i)P(A|B_i)$.

例2. 坛子里有5个球 (3绿2红), 无放回地每次取1个, 记 $A_i=\{$第i次取到绿球$\}$, $i=1,2$, 求 $P(A_2)$

解: $P(A_2)=P(A_1)P(A_2|A_1)+P(A_1^c)P(A_2|A_1^c)$

$=\frac{3}{5}\times\frac{2}{4}+\frac{2}{5}\times\frac{3}{4}=\frac{3}{5}$

引理2. (Bayes 公式). 设 $A_1,\cdots,A_n$ 为 Ω 的一个划分, $P(A_i)>0$, $\forall i$, 则当 $P(B)>0$ 时有:

$$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{P(B)}=\frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n}P(A_j)P(B|A_j)}$$

Bayes 1701-1761
1763《机会学说中一个问题的解》

注: 贝叶斯推断.

例3. 别墅里有只狗, 平均每周晚上叫3次, 一年里有2次盗窃, 且发生盗窃时狗叫概率为0.9, 问狗叫情况下发生盗窃概率?

解: $A=\{$狗晚上叫$\}$ $B=\{$盗窃发生$\}$

$P(A|B)=0.9$　$P(A)=\frac{3}{7}$　$P(B)=\frac{2}{365}$

$$P(B|A)=\frac{P(B)P(A|B)}{P(A)}=\frac{\frac{2}{365}\times0.9}{\frac{3}{7}}$$

独立性:

掷两次硬币, $A=\{$第2次H$\}$ $B=\{$第一次H$\}$, 则 $P(A|B)=\frac{1}{2}$　$P(A)=\frac{1}{2}$　$P(A|B)$ 不受B影响

$P(A|B)=P(A)=\frac{P(AB)}{P(B)}$

#. 定义2. 称A与B相互独立, 若

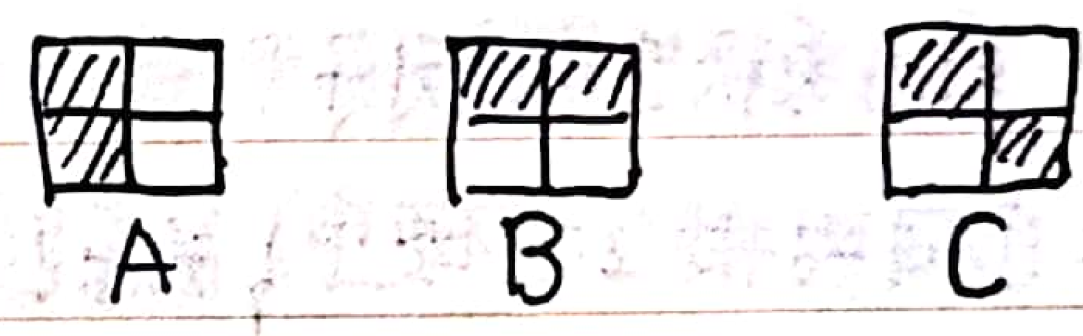
$P(AB)=P(A)P(B)$.

更一般, 称 $A_1,\cdots,A_n$ 相互独立, 若 $\forall k\geq2$, 有 $P(A_{i_1}\cdots A_{i_k})=P(A_{i_1})\cdots P(A_{i_k})$,

这里 $1\leq i_1<i_2<\cdots<i_k\leq n$

注: 两两独立指 $\forall i\neq j$, $P(A_iA_j)=P(A_i)P(A_j)$

例4.



将一个正方形分成四块, 随机取一块.

分析: $P(A)=P(B)=P(C)=\frac{1}{2}$

$P(AB)=\frac{1}{4}$　$P(BC)=\frac{1}{4}$　$P(CA)=\frac{1}{4}$

有A.B.C两两独立, 但 $P(ABC)=\frac{1}{4}\neq\frac{1}{8}$

表明A.B.C非相互独立.

引理3. 若A与B独立, 则A与 B^c、A^c 与B、A^c 与 B^c 亦独立.

分析: $P(AB^c)=P(A\setminus(A\cap B))=P(A)-P(A\cap B)$

$=P(A)-P(A)P(B)=P(A)(1-P(B))=P(A)P(B^c)$

例5. 三个小组独立破译某密码, 成功概率分别为0.4, 0.5, 0.7, 求成功破译概率

分析: $A_1, \cdots, A_n$ 相互独立, 则

$P(\bigcup_{i=1}^{n} A_i) = 1 - P(\bigcap_{i=1}^{n} A_i^c) = 1 - \prod_{i=1}^{n} P(A_i)$

具体所求 $= 1 - 0.4 \times 0.5 \times 0.7 = 0.91$.

例6. ~~独立~~"重复独立试验中, 小概率事件必然发生".

分析: $A_k = \{A$在第k次出现$\}$ $P(A_k) = \varepsilon \in (0,1)$

重复n次, 至少有1次发生:

$P(\bigcup_{i=1}^{n} A_k) = 1 - (1-\varepsilon)^n \xrightarrow{n \to \infty} 1$

§1.3 乘积空间. 例子

Question: $\left.\begin{array}{l}\text{①同时掷2枚硬币}\\ \text{②掷1枚硬币}\end{array}\right\}$ 关系?

Fact: $\forall i \in I$. $\mathcal{F}_i \subset 2^{\Omega}$ 为σ代数, 则 $\bigcap_{i \in I} \mathcal{F}_i$ 亦为σ代数

例子: $\Omega = \{1, 2, \cdots, n\}$ $A_i = \{i\}$ $(n > 2)$

$\mathcal{F}_i = \{\phi, \Omega, A_i, A_i^c\}$, 则

$\mathcal{F}_1 \cup \mathcal{F}_2 = \{\phi, \Omega, A_1, A_1^c, A_2, A_2^c\}$,

明显 $A_1 \cup A_2 = \{1, 2\} \notin \mathcal{F}_1 \cup \mathcal{F}_2$, 非σ代数

由小σ代数到大σ代数, 方式之一, 对 $\mathcal{F}, \mathcal{G} \subset 2^{\Omega}$ 称包含 $\mathcal{F}, \mathcal{G}$ 的最小σ代数为 $\mathcal{F}$ 与 $\mathcal{G}$ 生成的σ代数, 记 $\sigma(\mathcal{F} \cup \mathcal{G})$

典型例子: $\boxed{\text{1维Borel域}}$: $\mathbb{R}$上形如 $(a, b]$ 区间生成的σ代数, 记 $\mathcal{B}(\mathbb{R})$, 其中每个元素称为Borel集 $\{b\} = \bigcap_n (b - \frac{1}{n}, b]$

$(a, b) = (a, b] \setminus \{b\}$ $[a, b] = \{a\} \cup (a, b]$

$[a, b) = \{a\} \cup (a, b)$

即一切开区间、闭区间、半开半闭区间均为Borel集。

n维Borel域: $\mathbb{R}^n$ 上形如 $(a_1, b_1] \times \cdots \times (a_n, b_n]$ 生成的σ域, 记为 $\mathcal{B}(\mathbb{R}^n)$

乘积空间: $(\Omega_1, \mathcal{F}_1, P_1)$, $(\Omega_2, \mathcal{F}_2, P_2)$ 构造更大空间?

$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$

$\mathcal{F}_1 \times \mathcal{F}_2 = \{A_1 \times A_2 : A_i \in \mathcal{F}_i, i = 1, 2\}$

未必为σ代数

例2. 接例1. $\mathcal{F}_1 \times \mathcal{F}_2 = \{\phi, \Omega \times \Omega, A_1 \times A_2^c, A_1^c \times A_2, A_1^c \times A_2^c, \cdots\}$, 显见 $A_1 \times A_2 \in \mathcal{F}_1 \times \mathcal{F}_2$, 但

$(A_1 \times A_2)^c = \Omega \times \Omega \setminus \{1, 2\} \notin \mathcal{F}_1 \times \mathcal{F}_2$

$= \{2, 3, \cdots, n\} \times \Omega \cup \Omega \times \{1, 3, 4, \cdots, n\}$

记 $\mathcal{F} = \sigma(\mathcal{F}_1 \times \mathcal{F}_2)$ 看作 $\Omega = \Omega_1 \times \Omega_2$ 的σ代数, 概率测度? 引入 $\mathcal{F}_1 \times \mathcal{F}_2$ 上函数

$P_{12} : \mathcal{F}_1 \times \mathcal{F}_2 \to \mathbb{R}$ $P_{12}(A_1 \times A_2) := P_1(A_1) P_2(A_2)$

特别地, $\mathcal{F}_1 \times \mathcal{F}_2$ 中元素的不交并.

若 Ω_1, Ω_2 有限, $\mathcal{F}_i = 2^{\Omega_i}$, 这时 P_{12} 良定

对应: $A_i \in \mathcal{F}_i$, $A_1 \longleftrightarrow A_1 \times \Omega_2$

$A_2 \longleftrightarrow \Omega_1 \times A_2$

在$(\Omega_1\times\Omega_2, \sigma(F_1\times F_2), P_{12})$看$A_1\times A_2$

$P_{12}(A_1\times A_2)=P_1(A_1)P_2(A_2)=P_{12}(A_1\times\Omega_2)\cdot P_{12}(\Omega_1\times A_2)$

表明$A_1\times\Omega_2$与$\Omega_1\times A_2$独立

例3. 掷硬币2次

$\Omega=\{HH, HT, TH, TT\}=\{H,T\}\times\{H,T\}$

$P_{12}(HH)=P_1(H)=P_2(H)=\frac{1}{4}$, $P_{12}(HT)=\frac{1}{4}$

赌徒破产问题: 例4. Player 财富为k, Dealer 财富为$N-k$, 掷硬币, 出现正面H时 Player 赢1, 否则-1, 双方赌到一方输光概率?

解: A_k = Player 初始财富为k最后输光.

$B=\{$首局出现H$\}$

全概公式: $P(A_k)=P(B)P(A_k|B)+P(B^c)P(A_k|B^c)$

$=\frac{1}{2}P(A_{k+1})+\frac{1}{2}P(A_{k-1})$

记$P_k=P(A_k)$, 则 $P_k=\frac{1}{2}(P_{k+1}+P_{k-1})$ (*)

初始条件 $P_0=1$, $P_N=0$.

解之: $P_k=(P_1-P_0)k+P_0=1-\frac{k}{N}$

(*)别致: 首先观察 $P_k^{(0)}=a+\frac{k}{N}(b-a)$ 为(*)满足$P_0=a$, $P_N=b$的一个解.

又(*)表明 $P_k\le\max\{P_{k-1},P_{k+1}\}$. $\forall 0<k<N$

进而P_k关于k最大值在0或N处取得.

若P_k为满足$P_0=a$, $P_N=b$的解, 则

$P_k-P_k^{(0)}$为(*)满足边值为0的解

进而 $P_k-P_k^{(0)}\equiv 0$

计数: 古典概型: 样本点有限, 等可能发生

例5. 盒子里有4绿6红共10个球, 从中随机取4个, 试求2绿2红概率.

解: 样本点个数 C_{10}^4, 等可能, 利于事件发生样本点个数 $C_4^2C_6^2$ 所求 $\frac{C_4^2C_6^2}{C_{10}^4}=\frac{3}{7}$

古典概型, 关键在样本空间的选取, 排列组合: 从n个不同对象$a_1,a_2,\cdots,a_n$中任取m个, 有多少种方式?

		不可重复	可重复
排列	有序	A_n^m	n^m
组合	无序	C_n^m	C_{m+n-1}^{n-1}

可重复组合(无序, 可重复) $n=3$, $m=2$, 从 1 2 3 中取两个有6种

11; 22; 33; ~~12~~

12; 13; 23;

匣子表示: |00| | | | |00| | | | |00|

|0|0| | |0| |0| | |0|0|

看作3个有序匣子装入2个不可辨序小球, 忽略最外面两个板, 从2板2球共4个位置中选2板, 共有$C_{2+3-1}^{3-1}=6$

更一般地, 从$n-1$个板与m个小球共$m+n-1$位置中选$n-1$个板的位置, 共C_{m+n-1}^{n-1}

例6. 将n个小球投入$N\ge n$个盒子中, 每种

投法是等可能，记A={前n个盒中各有1个小球}，求P(A).

分析：小球是否可分辨，每个盒子是否限制容量。

(I) 球可辨，盒子无限制：(Maxwell-Boltzmann 统计)

样本点：N^n　A含有n!个点

$P(A)=\frac{n!}{N^n}$

(II) 不可辨，无限制：(Bose-Einstein 统计) 1924.

用匣子模型，样本点 C_{n+N-1}^{N-1}，$P(A)=\frac{1}{C_{n+N-1}^{N-1}}$

(III) 不可辨，每个盒子至多1个 (Fermi-Dirac 统计) 1925.

样本点：C_N^n　$P(A)=\frac{1}{C_N^n}$

例7.（Pólya 坛子模型）坛子里有b个黑球r个红球，从中取一个，取后放回并再放回c个同色球，记B_n={第n次取到黑球}，求$P(B_n)$

分析：① 抽取$\{B_1,B_2,R_3\}$，$\{B_1,R_2,B_3\}$，$\{R_1,B_2,B_3\}$

$$P(B_1B_2R_3)=P(B_1)P(B_2|B_1)P(R_3|B_1B_2)$$
$$=\frac{b}{b+r}\cdot\frac{b+c}{b+r+c}\cdot\frac{r}{b+r+2c}$$
$$P(B_1R_2B_3)=\frac{b}{b+r}\cdot\frac{r}{b+r+c}\cdot\frac{b+c}{b+r+2c}$$
$$P(R_1B_2B_3)=\frac{r}{b+r}\cdot\frac{b}{b+r+c}\cdot\frac{b+c}{b+r+2c}$$

② 在前n次抽取中含k个黑与(n-k)个红球，任意给定顺序概率为：

$$D_k(b)=\frac{b(b+c)\cdots(b+(k-1)c)\,r(r+c)\cdots(r+(n-k-1)c)}{(b+r)(b+r+c)\cdots(b+r+(n-1)c)}$$

③ $P(B_{n+1})=?$

记A_k={前n次中抽到k次黑球} $k=0,1,\cdots,n$

全概公式 $=\sum_{k=0}^{n}P(A_k)P(B_{n+1}|A_k)$

$$=\sum_{k=0}^{n}C_n^k D_k(b)\frac{b+kc}{b+r+nc}$$

$$\xlongequal{b'=b+c}\sum_{k=0}^{n}\underbrace{C_n^k D_k(b')}\cdot\frac{b}{b+r}$$

P(b'个黑球r个红球前n次取k黑球)

$=\frac{b}{b+r}$

注：c=0 有放回；c=-1 无放回，抽签次序无关

第二章 随机变量、分布函数

§2.1 随机变量

例1. 掷硬币.

$\Omega=\{H,T\}$　$X(H)=1$　$X(T)=-1$

$X:\Omega\to\mathbb{R}$　关心 $A(\omega):=\{\omega\in\Omega, X(\omega)\le x\}$

定义1. $(\Omega,\mathcal{F},P)$, $X:\Omega\to\mathbb{R}$, 若 $\forall x\in\mathbb{R}$ 有 $\{\omega\in\Omega: X(\omega)\le x\}\in\mathcal{F}$, 则称 X 为 $(\Omega,\mathcal{F},P)$ 上一个随机变量

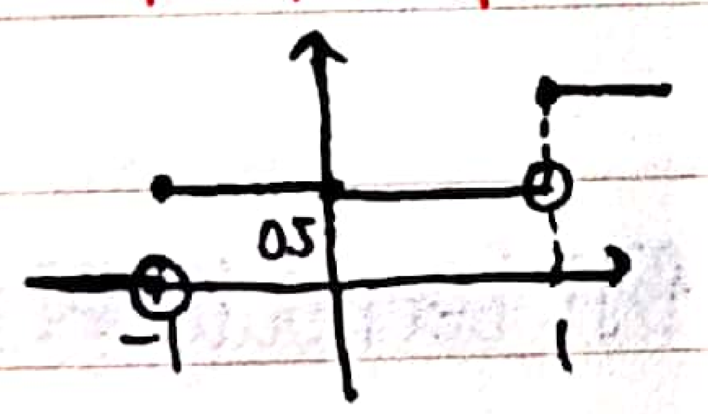
例2. 重复掷一枚硬币直到首次出现H, X表示投掷次数, 为随机变量.

注: $\{X\le x\}$ 表示 $\{\omega\in\Omega: X(\omega)\le x\}$

定义2. X 为 $(\Omega,\mathcal{F},P)$ 上随机变量, 则称函数 $F(x)=P(\{X\le x\})$ 为 X 的概率分布函数

例3. 接例1.

$$F(x)=\begin{cases}1 & x\ge 1\\ \frac{1}{2} & -1\le x<1\\ 0 & x<-1\end{cases}$$



引理1. 概率分布函数基本性质

(i) 单调递增: 若 $x<y$, 则 $F(x)\le F(y)$

(ii) $\lim_{x\to-\infty}F(x)=0$　$\lim_{x\to+\infty}F(x)=1$

(iii) 右连续: $F(x+0)=F(x)$

证明 (i) $F(y)-F(x)=P(\{x<X\le y\})\ge 0$

(ii) 令 $A_n=\{X\le -n\}$, 显然 $A_1\supset A_2\supset\cdots$, $\bigcap_{n=1}^{\infty}A_n=\emptyset$, 由P"连续性",

$\lim_{n\to\infty}F(-n)=\lim_{n\to\infty}P(A_n)=P(\bigcap_{n=1}^{\infty}A_n)=0$

类似有 $\lim_{n\to\infty}F(n)=1$, 再(i)即可

(iii) 令 $B_n=\{X\le x+\frac{1}{n}\}$　$B_n\downarrow$　$\bigcap_{n=1}^{\infty}B_n=\{X\le x\}$

$F(x+\frac{1}{n})\to F(x)$

注: (i) 满足引理1 (i)(ii)(iii) 一元函数 $F(x)$ 称为分布函数. 测度论知识表明: 任一分布函数必为某概率空间上某随机变量的概率分布函数.

(ii) 有的教材定义 $G(x)=P(X<x)$、$P(\{X<x\})$

这时 $G(x)$ 左连续, 保持 (i) (ii)

(iii) 随机变量的分布函数:

"忘记"样本空间

例4. 常值随机变量.

$X:\Omega\to\mathbb{R}$　$X(\omega)\equiv C$, $\forall\omega\in\Omega$

$$F(x)=\begin{cases}1 & x\ge C\\ 0 & x<C\end{cases}$$

几乎处处常值: $P(X=C)=1$

例5. Bernoulli两点分布.

$P(X=1)=p$　$P(X=0)=q$　$p+q=1$

分布函数 $F(x)=\begin{cases}1 & x\ge 1\\ q & 0\le x<1\\ 0 & x<0\end{cases}$

例6. 示性函数:

$A\in\mathcal{F}$　$I_A:\Omega\to\mathbb{R}$　$I_A(\omega)=1$, 若 $\omega\in A$, 否则 $I_A(\omega)=0$.

引理2. (i) $P(X>x)=1-F(x)$

(ii) $P(x<X\le y)=F(y)-F(x)$　　$x\le y$

(iii) $P(X=x)=F(x)-F(x-0)$

分析 (iii) 令 $A_n=\{x-\frac{1}{n}<X<x\}$, 则 $A_n\downarrow$

$\bigcap_{n=1}^{\infty}A_n=\{X=x\}$　　$P(A_n)=F(x)-F(x-\frac{1}{n})$

引理3. 设 X 为 (Ω,F,P) 上随机变量, 则

$\forall B\in\mathcal{B}(\mathbb{R})$ 有 $X^{-1}(B):=\{w\in\Omega: X(w)\in B\}\in\mathcal{F}$

分析: 记 $\mathcal{A}=\{A\subset\mathbb{R}: X^{-1}(A)\in\mathcal{F}\}$.

明显, $(-\infty,a]\in\mathcal{A}$, $(a,b]\in\mathcal{A}$　$\mathcal{F}$ ψ

因为 $X^{-1}((a,b])=X^{-1}((-\infty,b])\setminus X^{-1}((-\infty,a])$

断言: $\mathcal{A}$ 为 σ 代数, 若如此, 由 $\mathcal{B}(\mathbb{R})$ 最小性即可

(i) $X^{-1}(\mathbb{R})=\Omega\in\mathcal{F}$, 即 $\mathbb{R}\in\mathcal{A}$

(ii) $X^{-1}(A^c)=(X^{-1}(A))^c\in\mathcal{F}$, 当 $A\in\mathcal{A}$ 时

(iii) $X^{-1}(\bigcup_n A_n)=\bigcup_n X^{-1}(A_n)\in\mathcal{F}$, 当 $A_n\in\mathcal{A}$

引理4 X,Y 为 $(\Omega,\mathcal{F},P)$ 上随机变量, 则 $X+Y$ 亦为随机变量

证明: $\forall x\in\mathbb{R}$, $\{X+Y\le x\}\in\mathcal{F}$, 因为:

$\{X+Y\le x\}$

$=\bigcap_{n=1}^{\infty}(\{X\le r_n\}\cup\{Y\le x-r_n\})$

这里 r_n 取遍所有有理数

LHS ⊂ RHS ✓ 另一方面, 若 $w\notin$ LHS, 则

$X(w)+Y(w)>x$, 即 $X(w)>x-Y(w)$,

取 $r_m\in\mathbb{Q}$ s.t. $X(w)>r_m>x-Y(w)$, 表明:

$w\notin\{X\le r_m\}\cup\{Y\le x-r_m\}$, $w\notin$ RHS

例7. X,Y 为随机变量, 则 $\max\{X,Y\}$ 亦为随机变量.

分析: $\max\{X,Y\}=\frac{X+Y+|X-Y|}{2}$

Fact: (1) 若 X 为随机变量, 则 $|X|$ 亦是

(2) 若 X 为随机变量, 则 cX 亦是, $c\in\mathbb{R}$

§2.2. 离散与连续型随机变量

定义1. (离散型) 随机变量 X 至多取可数个实值 $x_1,x_2,\cdots,x_k,\cdots$, 记 $p_k=P(X=x_k)$, 并称 $p_1,p_2,\cdots,p_k,\cdots$ 为 X 的 $\boxed{\text{分布列}}$. 亦可用 $f(x)=P(X=x)$ 表示分布列

分布函数 $F(x)=\sum_{k:x_k\le x}p_k$

在 $x=x_1,x_2,\cdots,x_k,\cdots$ 有跳跃, F 亦称"原子分布"

例1. Bernoulli 两点分布

定义2. (连续型) 随机变量 X, 若存在非负可积函数 $f(x)$ s.t. X 的分布函数,

$F(x)=\int_{-\infty}^{x}f(u)du$

则称 X 为连续型, 且 $f(x)$ 为 X 的(概率)密度函数.

几点说明:

(i) $(F(x_0+\Delta x)-F(x_0))\frac{1}{\Delta x}=\frac{1}{\Delta x}\int_{x_0}^{x_0+\Delta x}f(u)du$

当 x_0 为 f 连续点时, $\frac{P(x_0<X\le x_0+\Delta x)}{\Delta x}\to f(x_0)$

密度函数+直观解释

(ii) 密度函数不唯一，改变有限多个点值亦为密度

(iii) $\int_{-\infty}^{+\infty} f(u)\,du = 1$

$P(X=a) < P(a-\frac{1}{n}<X\le a+\frac{1}{n})$

$=\int_{a-\frac{1}{n}}^{a+\frac{1}{n}} f(u)\,du \longrightarrow 0$

(iv) 若F连续，且除有限个点之外 $F'(x)$ 存在且连续，则微积分基本定理知 X 为连续型，且 $F'(x)$ 为密度

(v) X为连续型 $\Longleftrightarrow$ F绝对连续

区分"分布函数连续"与"连续型X的分布函数"

例2. $\Omega=[0,a]$ $a>0$, $\mathcal{F}=\{\Omega\cap A: A\in\mathcal{B}(\mathbb{R})\}$

$P(A)=\frac{|A|}{a}$ $\forall A\in\mathcal{F}$

随机变量 $X(\omega)=\omega$, $\omega\in\Omega$

$F(x)=P(X\le x)=\frac{x}{a}$ $x\in[0,a]$

密度函数 $f(x)=\frac{1}{a}$ $x\in[0,a]$

例3. X有密度 $f(x)=1$, $0\le x\le 1$, 掷硬币出现H时令 $Y=0$, 否则 $Y=X$

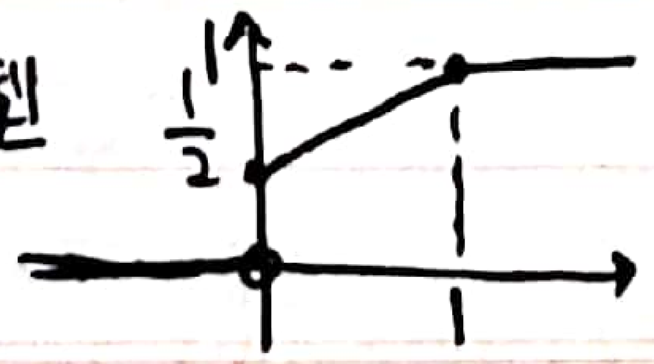
分析: $y\in[0,1]$. $P(Y\le y)$

$=P(H)P(Y\le y|H)+P(T)P(Y\le y|T)$

$=\frac{1}{2}+\frac{1}{2}y$

$y>1$, $P(Y\le y)=1$ $y<0$, $P(Y\le y)=0$

Y既非连续型亦非离散型



注：分布函数性质

(i) F的不连续点至多可数

(ii) F有勒贝格分解：$\cancel{F(x)=C_1F_1(x)+C_2F_2(x)+C_3}$

$F(x)=C_1F_1(x)+C_2F_2(x)+C_3F_3(x)$

F_1 离散型，F_2 连续型，F_3 奇异型

随机向量

定义3. $X_1,\cdots,X_n$ 为 $(\Omega,\mathcal{F},P)$ 上随机变量，称 $\vec{X}=(X_1,\cdots,X_n)$ 为一个n维随机向量，而n元函数 $F(x_1,\cdots,x_n)=P(X_1\le x_1,\cdots,X_n\le x_n)$ 为 $\vec{X}$ 的联合分布函数

离散型: $\vec{X}$ 与 $\mathbb{R}^n$ 中至多可数多个值，联合分布列

$f(x_1,\cdots,x_n)=P(X_1=x_1,\cdots,X_n=x_n)$

连续型 存在非负可积 $f(x_1,\cdots,x_n)$ s.t.

$F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f(u_1,\cdots,u_n)\,du_1\cdots du_n$

且称 $f(x_1,\cdots,x_n)$ 为 $\vec{X}$ 的联合密度函数

$n=2$ (X,Y) 联合分布函数 $F(x,y)$

引理1. (i) $F(x,y)$ 分别对 x, y 单调增

(ii) $F(x,y)$ 分别对 x, y 右连续

(iii) $\lim_{x\to-\infty}F(x,y)=\lim_{y\to-\infty}F(x,y)=0$

~~(iii)~~ $\lim_{x,y\to+\infty}F(x,y)=1$

(iv) $\forall x_1\le x_2$, $y_1\le y_2$

$F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\ge 0$

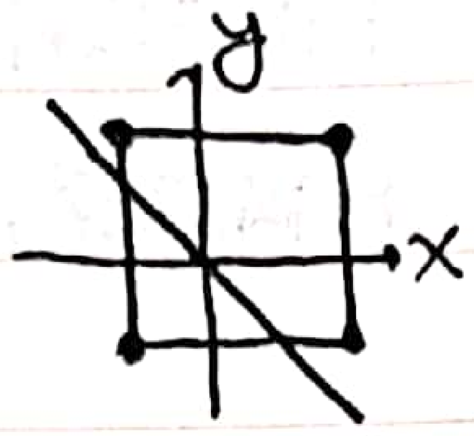
证明: (iv) $P(x_1<X<x_2,\ y_1<Y\le y_2)$

$=P(X\le x_2,\ y_1\le Y\le y_2)-P(X\le x_1,\ y_1<Y\le y_2)$

$=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\geq 0$

注：(iv)蕴含(i)，反之不然，满足(ii)(iii)(iv)的二元函数 $F(x,y)$ 必为某2维随机向量的联合分布函数。反例：

$$F(x,y)=\begin{cases}1 & x+y\geq 0\\ 0 & x+y<0\end{cases}$$



边际(缘)分布：$F(x,y)$：$F_X(x)$ 与 $F_Y(y)$ 称边缘分布，显见：$F(x,y)=P(X\le x, Y\le y)$

$F_X(x)=\lim_{y\to+\infty}F(x,y)$　$F_Y(y)=\lim_{x\to+\infty}F(x,y)$

对连续型密度 $f(x,y)$，边缘密度

$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)\,dy$

例4. 三项分布 掷均匀"三面"(H, T, E) n 次，分别出现 H_n, T_n, E_n，则 (H_n, T_n, E_n) 为随机向量，且 $H_n+T_n+E_n=n$.

$P((H_n,E_n,T_n)=(h,t,e))=\frac{n!}{h!\,t!\,e!}\cdot\left(\frac{1}{3}\right)^n$

更一般，$(X_1,\cdots,X_r)$ 服从参数为 $(n,p_1,\cdots,p_r)$ 的多项分布，若 $P((X_1,\cdots,X_r)=(k_1,\cdots,k_r))$
$=\frac{n!}{k_1!\cdots k_r!}p_1^{k_1}\cdots p_r^{k_r}$

例5. 均匀分布

G 为 R^n 中有限区域，$(X_1,\cdots,X_n)$ 联合密度.

$f(x_1,\cdots,x_n)=\frac{1}{|G|}$，$(x_1,\cdots,x_n)\in G$.

例6.(Buffon问题) 平面上有间距为2平行线，投长为1的针，试求针与线相交的概率？

分析：X 表示针与最近一条平行线距离，θ 表示针与线夹角，明显：$0\le X\le 1$，$0\le\theta\le\pi$，

$G=\{(\theta,x):\theta\in[0,\pi], x\in[0,1]\}$

A 表示针与线相交：$X\le\frac{1}{2}\sin\theta$

$P(A)=\frac{1}{\pi}\cdot\int_0^{\pi}\int_0^{\frac{1}{2}\sin\theta}1\,dx\,d\theta=\frac{1}{\pi}$

蒙特卡洛模拟 $g:[0,1]\to[0,1]$ 连续函数

算 $I=\int_0^1 g(x)\,dx$　(X,Y) 为 $[0,1]\times[0,1]$ 上均匀分布，$A=\{(x,y):0\le y\le g(x), x\in[0,1]\}$

落入A点数 N_A，则 $\frac{N_A}{N}\longrightarrow P(A)=I$

随机模拟中用到均匀分布

$X\sim U(0,1)$ (0,1)上均匀分布，当F严格增时，$Y=F^{-1}(X)$ 有分布函数 $F(y)$

$P(Y\le y)=P(F^{-1}(X)\le y)=P(X\le F(y))=F(y)$

一般地，对分布函数F，定义

$F^{-1}(y):=\sup\{x\mid F(x)<y\}$，$y\in(0,1)$

Fact：设F为分布函数，$X\sim U(0,1)$，则 $Y=F^{-1}(X)$ 的分布函数为F

分析：$P(Y\le y)=P(F^{-1}(X)\le y)\overset{?}{=}P(X\le F(y))$
$=F(y)$

另外Fact：$F^{-1}(y)\le x\Leftrightarrow y\le F(x)$ 思考

第三章 离散型随机变量

§3.1. 分布列.独立性

离散 X $p_k=P(X=x_k)$ $\sum_k p_k=1$

例1. 二项分布 $B(n,p)$

$P(X=k)=C_n^k p^k q^{n-k}$, $k=0,1,\cdots,n$

这里 $p\in(0,1)$, $q=1-p$, X服从参数为(n,p)二项分布. 记 $X\sim B(n,p)$

背景:掷硬币n次,每次正面出现概率为p, X为H出现次数,则 $X\sim B(n,p)$

例2. 几何分布

$P(X=k)=pq^{k-1}$ $k=1,2,\cdots$

这里 $p\in(0,1)$, $q=1-p$. 显然 $P(X>k)=q^k$

背景:独立重复掷硬币 X=首次出现H掷的次数/等待时间

无记忆性:前m次试验中A没有发生,此条件下A发生所需等待时间X'与m无关,且服从参数为p的几何分布

因为 $P(X'=k)=P(X=m+k\mid X>m)$

$=\frac{P(X=m+k)}{P(X>m)}=\frac{pq^{m+k-1}}{q^m}=pq^{k-1}$

Fact: 正整数值随机变量X. $\forall m=0,1,2,\cdots$ $P(X=m+1\mid X>m)$与m无关,则X服从几何分布。

分析:记 $p=P(X=k+1\mid X>k)$与k无关,

令 $r_k=P(X>k)$, 则 ($r_0=1$)

$p=\frac{P(X=k+1)}{P(X>k)}=\frac{r_k-r_{k+1}}{r_k}$, 解之 $r_k=(1-p)^k$

例3. 泊松(Poisson)分布

$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$ $k=0,1,2,\cdots$ $\lambda>0$

称X服从参数为λ的泊松分布 $X\sim P(\lambda)$

背景:网站访问量,保险中索赔数,物理学中放射粒子数:体积为V某块物质分割为n等份 $\Delta V=\frac{V}{n}$, 假设

(i)每一小块7.5S内放出一个α粒子概率为p, $p=\mu\cdot\Delta V$ $(\mu>0)$, 放出2个或更多概率为0

(ii)每一小块是否放出粒子相互独立.

分析: n块中恰好有k小块放出粒子概率用二项分布. $P(X=k)=C_n^k p^k q^{n-k}$

记$\lambda=\mu V$, 则 $p=\frac{\lambda}{n}$, 让$n\to\infty$

$C_n^k p^k q^{n-k}=\frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$

$=\frac{\lambda^k}{k!}\cdot\frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k}\to\frac{\lambda^k}{k!}e^{-\lambda}$

符合(i)(ii)条, Poisson出现,

二项分布逼近Poisson分布.

独立性:

定义1. 称离散型X与Y独立,若$\forall x,y\in\mathbb{R}$, $\{X=x\}$与$\{Y=y\}$独立. 更一般,称$X_1,\cdots,X_n$相互独立,若$\forall x_i\in\mathbb{R}$, 有$\{X_1=x_1\},\cdots,\{X_n=x_n\}$相互独立.

引理1. 离散型X, Y.

X和Y独立$\Leftrightarrow\forall x,y\in\mathbb{R}$, $P(x\le X;\ y\le Y)$,

$=P(X\le x)P(Y\le y)$

证明:"⇒" $P(X\le x, Y\le y)=\sum_{\substack{x_i\le x\\ y_j\le y}} P(X=x_i, Y=y_j)$

$=\sum_i P(X=x_i)\sum_j P(Y=y_j)=P(X\le x)P(Y\le y)$

"⇐"记 $f(x)=F(x)-F(x-0)$

$F(x,y)=F_X(x)F_Y(y)$ ①

$F(x-0,y)=F_X(x-0)F_Y(y)$ ②

$F(x,y-0)=F_X(x)F_Y(y-0)$ ③

$F(x-0,y-0)=F_X(x-0)F_Y(y-0)$ ④

①-②-(③-④); $f(x,y)=f_X(x)f_Y(y)$

分析: $\{g(X)=a, h(Y)=b\}=\bigcup_{\substack{g(x)=a\\ h(y)=b}}\{X=x, Y=y\}$ 不相容并

§3.2. 数学期望

分布列:

X	x_1	x_2	$\cdots$	x_n $\cdots$
P	p_1	p_2	$\cdots$	p_n $\cdots$

X ~~平均~~取值"平均"大小,用一个数来表示,如:

X	100	200	加权平均
P	0.01	0.99.	

例4. 泊松翻转:

掷硬币1次, $P(H)=p$, 记X、Y为H、T出现次数, 则X与Y不独立, 因为

$P(X=1, Y=1)=0$, $P(X=1)P(Y=1)=pq$

掷硬币N次, $N\sim P(\lambda)$, 则X与Y独立.

因为 $P(X=x, Y=y)=P(X=x, Y=y, N=x+y)$

$=P(X=x, Y=y\mid N=x+y)P(N=x+y)$

$=C_{x+y}^{x}p^x q^y\frac{\lambda^{x+y}}{(x+y)!}e^{-\lambda}$

$=\frac{(\lambda p)^x}{x!}e^{-\lambda p}\cdot\frac{(\lambda q)^y}{y!}e^{-\lambda q}$

$P(X=x)=\sum_{y=0}^{\infty}P(X=x, Y=y)=\frac{(\lambda p)^x}{x!}e^{-\lambda p}$

$P(Y=y)=\frac{(\lambda q)^y}{y!}e^{-\lambda q}$.

$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

$g:\mathbb{R}\to\mathbb{R}$ Borel可测, 随机变量函数

$Z=g(X)$, $Z(\omega)=g(X(\omega))$

引理2. 若X、Y独立, $g,h:\mathbb{R}\to\mathbb{R}$, 则: $g(X)$、$h(Y)$亦独立.

定义1. 离散型X分布列 $p_k=P(X=x_k)$ $k=1,2\cdots$ 称~~和~~和数: $\sum_k p_k x_k$ (绝对收敛时)为X的数学期望, 记 $E[X]$. 注: 想象某 $x_k=x_j$ 时, 对上面求和无影响。

例1(投资理财) 贵州茅台股票 或招商银行年利率5%

X	8	-2
P	0.3	0.7

如何选择?

分析: (I) $E[X]=0.3\times8+0.7\times(-2)=-1$

(II) $10\times5\%=0.5$

$X, g:\mathbb{R}\to\mathbb{R}, Y=g(X)$ (佚名统计学家公式)

引理1. X有分布列 $f(x)$. $g:\mathbb{R}\to\mathbb{R}$ Borel可测, 则 $E[g(X)]=\sum_{x:f(x)>0}g(x)f(x)$

这里求和绝对收敛时.

分析: 令 $Y=g(X)$, 则 $f_Y(y)=P(g(X)=y)$

$=P(\bigcup_{x:g(x)=y}\{X=x\})=\sum_{x:g(x)=y}P(X=x)$

$\mathbb{E}[Y]=\sum_y y f_Y(y)=\sum_y \sum_{x:g(x)=y} y f(x)$
$=\sum_x g(x) f(x)$

定义2. k阶矩 $m_k=\mathbb{E}[X^k]$，m_1为均值μ

k阶中心矩 $\sigma_k=\mathbb{E}[(X-m_1)^k]$

方差: $Var(X)=\mathbb{E}[(X-m_1)^2]=\mathbb{E}[(X-\mu)^2]$

标准差: $\sigma=\sqrt{Var(X)}$

$Var(X)=\sum_x (x-m_1)^2 f(x)$
$=\sum_x x^2 f(x)-2m_1\sum_x x f(x)+m_1^2\sum_x f(x)$
$=\mathbb{E}[X^2]-(\mathbb{E}[X])^2$

例2. Bernoulli分布: $P(X=1)=p$　$P(X=0)=q$

则 $\mathbb{E}[X]=p\cdot 1+q\cdot 0=p$

$\mathbb{E}[X^2]=p\cdot 1^2+q\cdot 0^2=p$

$Var(X)=pq$

例3. 二项分布 $B(n,p)$

$P(X=k)=C_n^k p^k q^{n-k}$　$k=0,1,\cdots,n.$

分析: $\mathbb{E}[X]=\sum_{k=0}^n k C_n^k p^k q^{n-k}$
$=\sum_{k=0}^n n C_{n-1}^{k-1} p^k q^{n-k}=np$

$\mathbb{E}[X(X-1)]=\sum_{k=0}^n k(k-1)C_n^k p^k q^{n-k}=n(n-1)p^2$

从而 $\mathbb{E}[X^2]=np(np+q)$　$Var(X)=npq$

类似, $X\sim P(\lambda)$, $\mathbb{E}[X]=\lambda$,　$Var(X)=\lambda$

数学期望可作为随机变量空间上线性"算子".

定理2. 期望$\mathbb{E}$基本性质

(i) 非负性: 若 $X\geq 0$, 则 $\mathbb{E}[X]\geq 0$

(ii) 归一性: $\mathbb{E}[C]=C$,　$C\in\mathbb{R}$

(iii) 线性性: $\forall a,b\in\mathbb{R}$有

$\mathbb{E}[aX+bY]=a\mathbb{E}[X]+b\mathbb{E}[Y]$

证明: (iii) 令 $A_x=\{X=x\}$　$B_y=\{Y=y\}$, 则

$X=\sum_x x I_{A_x}$　$Y=\sum_y y I_{A_y}=\sum_{x,y} y I_{B_y A_x}$

$aX+bY=\sum_x ax I_{A_x}+\sum_y by I_{B_y}$
$=\sum_{x,y}(ax+by) I_{A_x B_y}$

据定义 $\mathbb{E}[aX+bY]=\sum_{x,y}(ax+by)P(A_x B_y)$
$=\sum_{x,y} ax P(A_x B_y)+\sum_{x,y} by P(A_x B_y)$
$=\sum_x ax P(A_x)+\sum_y by P(B_y)=a\mathbb{E}[X]+b\mathbb{E}[Y]$

引理3. 若X, Y独立, 则 $\mathbb{E}[XY]=\mathbb{E}[X]\mathbb{E}[Y]$,
这里X与Y期望均存在

分析: $XY=\sum_{x,y} xy I_{A_x B_y}$, 则

$\mathbb{E}[XY]=\sum_{x,y} xy P(A_x B_y)$
$\overset{独立}{=}\sum_{x,y} xy P(A_x)P(B_y)=\mathbb{E}[X]\mathbb{E}[Y]$

定理4. (i) $\forall a,b\in\mathbb{R}$, $Var(aX+b)=a^2 Var(X)$

(ii) 若X与Y独立, 则 $Var(X+Y)=Var(X)+Var(Y)$

(iii) 若X与Y不相关, ($\mathbb{E}[XY]=\mathbb{E}[X]\mathbb{E}[Y]$), 则
$Var(X+Y)=Var(X)+Var(Y)$ 亦成立.

分析: $Var(X+Y)=\mathbb{E}[(X+Y-\mathbb{E}[X+Y])^2]$
$=\mathbb{E}[((X-\mathbb{E}[X])+(Y-\mathbb{E}[Y]))^2]$
$=Var(X)+Var(Y)+2\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$

数字特征

$= Var(X)+Var(Y)+2(E[XY]-E[X]E[Y])$

$\underline{\underline{\text{独立或不相关}}}\ Var(X)+Var(Y)$

例4. $P(X=x_k)=2^{-k}$ $x_k=(-1)^k\frac{2^k}{k}$ $k=1,2,\cdots$

则 $\sum_k p_k x_k = \sum_{k=1}^{\infty}2^{-k}(-1)^k\frac{2^k}{k}=-\ln 2.$

但 $\sum_k p_k|x_k|=\sum_{k=1}^{\infty}\frac{1}{k}=\infty$，期望不存在.

示性函数 $I_A(\omega)=\begin{cases}1 & \omega\in A\\ 0 & \omega\notin A\end{cases}$

注: $P(A):=E[I_A]$，期望算子E三条基本性质可作为概率公理化定义. 优点:

① 非~~线性~~交换概率、算子代数、量子物理、随机矩阵

② 非线性期望、金融风险度量

彭实戈(山东大学)

(i) 单调性: $X\geq Y\Rightarrow \hat{E}[X]\geq\hat{E}[Y]$

(ii) 归一性: $\hat{E}[C]=C$，$C\in\mathbb{R}$

(iii) 次可加: $\hat{E}[X+Y]\leq\hat{E}[X]+\hat{E}[Y]$

(iv) 正齐性: $\hat{E}[\lambda X]=\lambda\hat{E}[X]$，$\lambda\geq 0$

例5.(随机置换): $\begin{matrix}1 & 2 & 3 & \cdots & n\\ \sigma_1 & \sigma_2 & \sigma_3 & \cdots & \sigma_n\end{matrix}$

令 $N=\#\{i:\sigma_i=i\}$，不动点数，求N分布列:

$P(N=r)$，$r=0,1,\cdots,n$.

分析: 令 $A_i=\{\sigma_i=i\}$，$I_i=I_{A_i}$

特殊情形 $A_1\cdots A_r A_{r+1}^c\cdots A_n^c$

$=I_1\cdots I_r(1-I_{r+1})\cdots(1-I_n)$

构造 $X=\sum_{\substack{i_1<i_2<\cdots<i_r\\ i_{r+1}<i_{r+2}<\cdots<i_n}} I_{i_1}\cdots I_{i_r}(1-I_{r+1})\cdots(1-I_n)$

这里 $i_1,\cdots,i_n$ 为 $1,2,\cdots,n$ 的置换，则

$X=\begin{cases}1 & N=r\\ 0 & N\neq r\end{cases}$

取期望 $E[X]=C_n^r E[I_1\cdots I_r(1-I_{r+1})\cdots(1-I_n)]$

$=C_n^r\sum_{s=0}^{n-r}(-1)^s C_{n-r}^s E[I_1\cdots I_r I_{r+1}\cdots I_{r+s}]$

$=C_n^r\sum_{s=0}^{n-r}(-1)^s C_{n-r}^s\frac{(n-r-s)!}{n!}$

$=\frac{1}{r!}\sum_{s=0}^{n-r}\frac{(-1)^s}{s!}=P(N=r)$

例6. 接上例，$E[N]$，~~$Var[N]$~~ $Var[N]$

分析: $N=I_1+\cdots+I_n$

$E[N]=nE[I_1]=n\cdot\frac{(n-1)!}{n!}=1$

$Var[N]=E[N^2]-(E[N])^2$

$E[N^2]=\sum_{i=1}^{n}E[I_i^2]+\sum_{i\neq j}E[I_iI_j]$

$=\frac{1}{n}\cdot n+\frac{(n-2)!}{n!}2C_n^2=2$

$\Rightarrow Var[N]=1$

(Erdös 概率方法)

例7. 正17边形恰有5个红色顶点，证明存在7个相邻顶点，其中至少有3个红色顶点.

分析: 记 $\Omega=\{1,2,\cdots,17\}$ 古典概型，

令 $J_k=\begin{cases}1 & \text{第k个顶点为红色}\\ 0 & \text{否则}\end{cases}$

现随机取一个点k，令 $X(k):=(J_{k+1}+\cdots+J_{k+7})$ (mod 17)

则X为随机变量，则

$E[X]=\sum_{k=1}^{17}\frac{1}{17}(J_{k+1}+\cdots+J_{k+7})=\frac{35}{17}>2$

Claim: $P(X>2)>0$，否则 $P(X>2)=0$，

即 $P(X\leq 2)=1$，进而 $E[X]\leq E[2]=2$，矛盾

另一方面X取正整数则$P(X\geq 3)>0$，表明$\{X\geq 3\}$非空，即存在k s.t. $X(k)\geq 3$.

随机变量卷积 X与Y独立，称$Z=X+Y$为X与Y卷积，且 $f_Z(z)=P(Z=z)=P(X+Y=z)$
$=P(\bigcup_x\{X=x, Y=z-x\})$
$=\sum_x f_X(x)f_Y(z-x)=\sum_y f_Y(y)f_X(z-y)$

思考题：$a_1,\cdots,a_n\in\mathbb{R}$，试用概率方法证明：可选取$\varepsilon_1,\cdots,\varepsilon_n\in\{-1,1\}$, s.t.
$(\sum_{k=1}^n \varepsilon_k a_k)^2\leq \sum_{k=1}^n a_k^2$

§3.3. 条件期望

联合分布函数 $F(x,y)=P(X\leq x, Y\leq y)$

分布列 $f(x,y)=P(X=x, Y=y)$.

边缘分布列. $f_X(x)=\sum_y f(x,y)$

Fact: $\mathbb{E}[g(X,Y)]=\sum_{\substack{x,y:\\ f(x,y)>0}} g(x,y)f(x,y)$

定义1. 离散型(X,Y)

协方差: $Cov(X,Y)=\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$

相关系数: $\rho(X,Y)=\dfrac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

这里$Var(X),Var(Y)>0$.

注：协方差概念，从 $\mu_X=\mathbb{E}[X]$
$Var(X+Y)=Var(X)+Var(Y)+2\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$

更一般，$\vec{X}=(X_1,\cdots,X_n)$，协方差矩阵

$\Sigma=(\sigma_{ij})_{ij}^n$, $\sigma_{ij}=Cov(X_i,X_j)$
$=\mathbb{E}[X_iX_j]-\mathbb{E}[X_i]\mathbb{E}[X_j]$

Claim. Σ为非负定，因为

$\sum_{ij}\sigma_{ij}t_it_j=\sum_{ij}t_it_j\mathbb{E}[(X_i-\mu_{X_i})(X_j-\mu_{X_j})]$
$=\mathbb{E}[(\sum_i t_i(X_i-\mu_{X_i}))^2]\geq 0$

引理1. $\rho=\rho(X,Y)$相关系数，则

(i) $-1\leq\rho\leq 1$.

(ii) 当X与Y独立或不相关时，$\rho=0$

(iii) $|\rho|=1 \Leftrightarrow$ 存在$a,b\in\mathbb{R}$ s.t. $P(Y=aX+b)=1$

定理2. (Cauchy-Schwarz不等式):
$(\mathbb{E}[XY])^2\leq\mathbb{E}[X^2]\mathbb{E}[Y^2]$

等号成立$\Leftrightarrow$存在不全为零$a,b\in\mathbb{R}$ s.t.
$P(aX=bY)=1$.

Proof.

Case 1: $\mathbb{E}[X^2]>0$，则$\forall t\in\mathbb{R}$，有
$\mathbb{E}[(Y-tX)^2]=t^2\mathbb{E}[X^2]-2t\mathbb{E}[XY]+\mathbb{E}[Y^2]\geq 0$

t的二次函数，可知判别式：
$(\mathbb{E}[XY])^2-\mathbb{E}[X^2]\mathbb{E}[Y^2]\leq 0$.

等号成立$\Leftrightarrow t_0\in\mathbb{R}$ s.t. $\mathbb{E}[(Y-t_0X)^2]=0$
$\overset{\text{Case 2.}}{\Leftrightarrow} P(Y=t_0X)=1$

Case 2: $\mathbb{E}[X^2]=0$, Claim: $P(X=0)=1$.

$\mathbb{E}[X^2]=\sum_x x^2f(x)=0$，可知$x\neq 0$时$f(x)=0$

从而$\mathbb{E}[XY]=\sum_{x,y}xyf(x,y)$
$=\sum_{x\neq 0,y}xyf(x,y)+\sum_{x=0,y}xyf(x,y)$

由 $f_X(x)=\sum_y f(x,y)$，当 $x\neq 0$ 时 $f(x,y)=0$

$E[XY]=0+0=0$

定义2. (X,Y) 离散型，给定 $X=x$（假定 $P(X=x)>0$）下 Y 的 $\boxed{\text{条件分布列}}$，

$f_{Y|X}(y|x)=P(Y=y|X=x)$ 关于 y 构成分布列.

条件分布

$F_{Y|X}(y|x)=P(Y\le y|X=x)$

条件期望

$\psi(x)=E[Y|X=x]=\sum_y y f_{Y|X}(y|x)$

并称 $\psi(X)$ 为 Y 关于 X 的条件期望，记 $E[Y|X]$

注：~~$\psi(X)$~~ Y 关于 X 条件期望 $E[Y|X]$ 为一个与 X 有关的随机变量

注：$f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}$，$f_X(x)>0$，

当 $f_X(x)=0$ 时，必有 $f(x,y)=0$，$\forall y\in\mathbb{R}$

定理3. 记 $\psi(X)=E[Y|X]$，则

☆ $E[\psi(X)]=E[Y]=E[E[Y|X]]=\sum_x E[Y|X=x]\cdot f_X(x)$ 期望形式的全概公式

分析：佚名统计学家：

$$E[\psi(X)]=\sum_x \psi(x)f_X(x)=\cancel{\sum_x f_X(x)}$$
$$=\sum_x f_X(x)\sum_y y f_{Y|X}(y|x)$$
$$=\sum_{x,y} y f_X(x) f_{Y|X}(y|x)$$
$$=\sum_{x,y} y f(x,y)$$
$$=\sum_y y f_Y(y)=E[Y]$$

例1. 一只鸟下 N 枚蛋，$N\sim P(\lambda)$，每枚蛋独立地以概率 p 变为小鸟数，求 $E[K|N]$，$E[K]$，$E[N|K]$

分析：$f_N(n)=\dfrac{\lambda^n}{n!}e^{-\lambda}\quad n=0,1,\cdots$

$f_{K|N}(k|n)=C_n^k p^k q^{n-k}\quad k=0,1,\cdots,n.$

因此：$\psi(n)=E[K|N=n]=np$，$E[K|N]=pN$，

进而 $E[K]=E[E[K|N]]=E[pN]=p\lambda$

先算条件分布列：$f_{N|K}(n|k)=P(N=n|K=k)$

$$=\frac{P(N=n,K=k)}{P(K=k)}=\frac{P(N=n)P(K=k|N=n)}{\sum_m P(N=m)P(K=k|N=m)}$$

$$\overset{n\ge k}{=}\frac{\frac{\lambda^n}{n!}e^{-\lambda}C_n^k p^k q^{n-k}}{\sum_{m\ge k}\frac{\lambda^m}{m!}e^{-\lambda}C_m^k p^m q^{m-k}}=\frac{\frac{1}{k!(n-k)!}\lambda^n e^{-\lambda}p^k q^{n-k}}{\frac{\lambda^k}{k!}p^k\sum_{m\ge 0}q^m\frac{\lambda^m}{m!}e^{-\lambda}}$$

$$=\frac{(\lambda q)^{n-k}}{(n-k)!}e^{-\lambda q}$$

因此 $E[N|K=k]=\sum_{n\ge k} n\dfrac{(\lambda q)^{n-k}}{(n-k)!}e^{-\lambda q}$

$$\xlongequal{n\to n+k}\sum_{n\ge 0}(n+k)\frac{(\lambda q)^n}{n!}e^{-\lambda q}=k+\lambda q.$$

$E[N|K]=K+\lambda q$

定理4. 对 $\psi(X)=E[Y|X]$，有

$E[\psi(X)g(X)]=E[Yg(X)]$，　(※)

这里 $g:\mathbb{R}\to\mathbb{R}$ "好"函数

注：(※)可作为条件期望 $\psi(X)$ 的定义

§3.4. 随机游走 I

直线上随机游走是最简随机过程.

从 $t=0$ 时刻起粒子 $S_0=a\in\mathbb{Z}$，$t=n$ 时刻处于 x 轴上，$S_n=S_{n-1}+X_n=S_0+\sum_{i=1}^n X_i$

这里 $X_1,\cdots,X_n$ 相互独立，且 $P(X_i=1)=p$，

$P(X_i=-1)=q$. 注: $p=\frac{1}{2}$时, 对称简单随机游走

例1.(自由随机游走). $S_0=a$, 求$P(S_n=b)=?$

解: 令r向右移动次数, l向左移动次数.

则 $r+l=n$, $r-l=b-a$,

即 $r=\frac{1}{2}(n+b-a)$

因此 $P(S_n=b)=C_n^{\frac{1}{2}(n+b-a)}\,p^{\frac{1}{2}(n+b-a)}\,q^{\frac{1}{2}(n-b+a)}$

例2.(带两个吸收壁) 粒子在$t=0$时处于$x=k$. 在$x=0, N$处各有一个吸收壁(移动到0或N处时停止), 求粒子在$x=0, N$处吸收概率 $(0<k<N)$.

分析: 记P_k=初始处于$x=k$且最后在$x=0$吸收概率. 第1步式后全概公式:

$P_k=pP_{k+1}+qP_{k-1},\quad 0<k<N$

边界条件: $P_0=1$, $P_N=0$, 记$r=\frac{q}{p}$, 则

$P_{k+1}-P_k=r(P_k-P_{k-1})$

解之
$$P_k=\begin{cases}\dfrac{r^k-r^N}{1-r^N} & r\neq 1\\ 1-\dfrac{k}{N} & r=1.\end{cases}$$

$$P_k^*(\text{在}x=N\text{处吸收概率})=\begin{cases}\dfrac{1-r^k}{1-r^N} & r\neq 1\\ \dfrac{k}{N} & r=1\end{cases}$$

随机游走重要性质.

引理1. (i) 空齐性 $P(S_n=j|S_0=a)$

$=P(S_n=j+b|S_0=a+b)$

(ii) 时齐性: $P(S_n=j|S_0=a)=P(S_{n+m}=j|S_m=a)$

(iii) Markov性: (马氏性) 立足现在, 未来与过去无关.

$P(S_{m+n}=j|S_0=j_0, S_1=j_1,\cdots,S_m=j_m)$

$=P(S_{m+n}=j|S_m=j_m)$

这里, 条件概率指左右两边有意义

分析: (i) $\text{RHS}=\dfrac{P(S_n=j+b, S_0=a+b)}{P(S_0=a+b)}$

$=\dfrac{P(\sum_{i=1}^{n}X_i=j-a, S_0=a+b)}{P(S_0=a+b)}=P(\sum_{i=1}^{n}X_i=j-a)=\text{LHS}$

(ii) $\text{RHS}=\dfrac{P(S_{n+m}-S_m=j-a, S_m=a)}{P(S_m=a)}$

$=P(\sum_{i=m+1}^{n+m}X_i=j-a)$

$\text{LHS}=P(\sum_{i=1}^{n}X_i=j-a)$, 可知 RHS=LHS

(iii) $\text{LHS}=\dfrac{P(S_{m+n}=j, S_0=j_0,\cdots,S_m=j_m)}{P(S_0=j_0,\cdots,S_m=j_m)}$

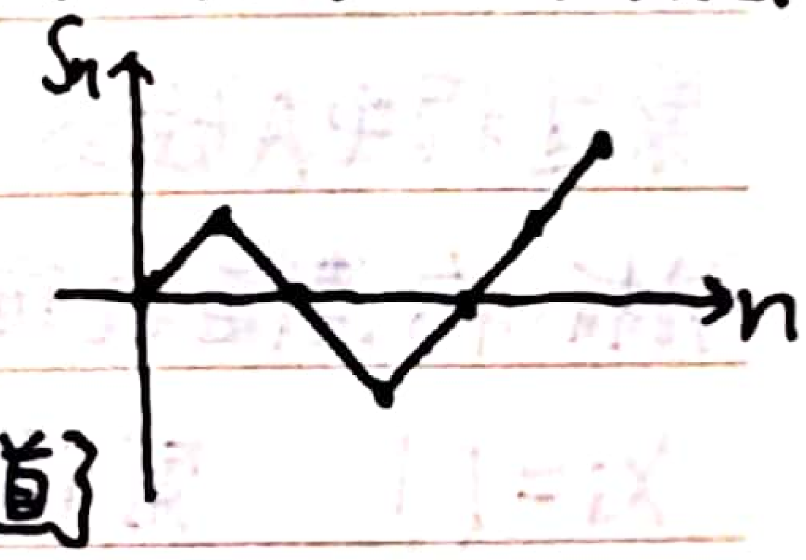
$=\dfrac{P(S_{m+n}-S_m=j-j_m, S_0=j_0,\cdots,S_m=j_m)}{P(S_0=j_0,\cdots,S_m=j_m)}$

$\overset{\text{独立}}{=}P(\sum_{i=m+1}^{m+n}X_i=j-j_m)=\text{RHS}$

轨道计数

粒子运动平面表示(n, S_n)

如从$S_0=0$开始, 相继访问x轴 1, 0, -1, 0, 1, 2.



记号: $S_0=a$, $S_n=b$,

$N_n(a,b)=$ #{(0,a)到(n,b)的轨道}

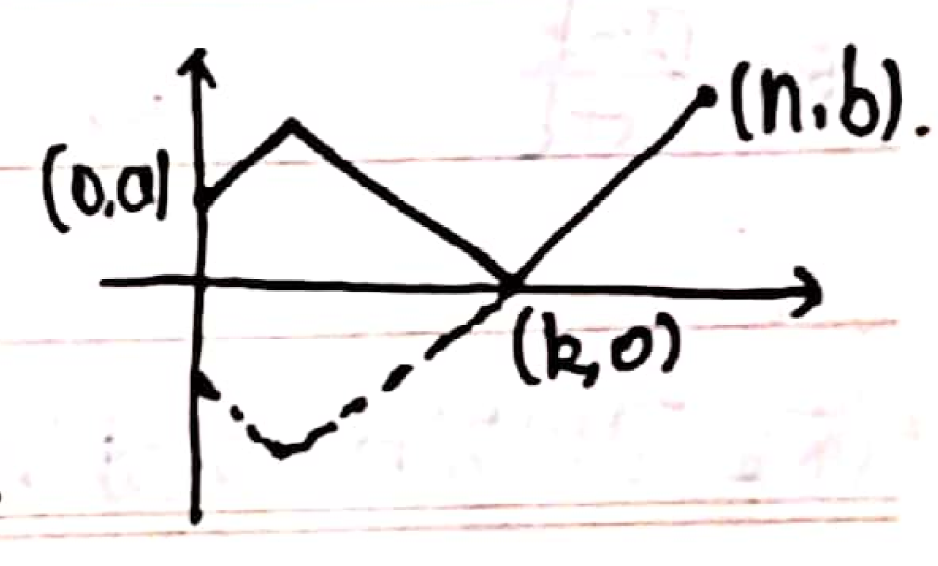
$N_n^0(a,b)=$ #{(0,a)到(n,b)且过x轴某点}

定理2. 若$a,b>0$, 则$N_n^0(a,b)=N_n(-a,b)$:

(反射原理)

分析: 构造1-1对应

在第1次和x轴交点



$(k,0)$处将前面轨道关于x轴反射，得到从$(0,-a)$到(n,b)的一条轨道. 易知这种对应为1-1对应

引理3. $N_n(a,b)=C_n^{\frac{1}{2}(n+b-a)}$

推论4. (投票定理): 若$b>0$, 则

$\#\{$从$(0,0)$到(n,b)且不再过x轴轨道$\}=\frac{b}{n}N_n(0,b)$.

分析: 第1步到$(1,1)$, 所求$(0,1)$到$(n-1,b)$且不再过x轴轨道数$=N_{n-1}(1,b)-N_{n-1}^0(1,b)$

$=N_{n-1}(1,b)-N_{n-1}(-1,b)$

$=\frac{(n-1)!}{(\frac{n-1+b-1}{2})!(\frac{n-1-b+1}{2})!}-\frac{(n-1)!}{(\frac{n-1+b+1}{2})!(\frac{n-1-b-1}{2})!}$

$=\frac{b}{n}C_n^{\frac{1}{2}(n+b)}=\frac{b}{n}N_n(0,b)$

例3. 选举中A得a票，B得b票，$a>b$，问投票过程中A始终领先B概率？

分析. 构造随机游走，令

$$X_i=\begin{cases}1 & 第i票给A\\ -1 & 否则\end{cases}$$

问题转化为从$(0,0)$到$(a+b,a-b)$的轨道中，不再过x轴数目与总数目之比，

即$\frac{a-b}{a+b}$

(不返回出发点) 定理5. $S_0=0$, $n\geq1$, 则.

$P(S_1S_2\cdots S_n\neq0,S_n=b)=\frac{|b|}{n}P(S_n=b)$

$P(S_1S_2\cdots S_n\neq0)=\frac{1}{n}E[|S_n|]$

证明: 不妨设$b>0$, 利用投票定理: 不再过x轴的轨道数. $\frac{b}{N}N_n(a,b)$. 因此

$P(S_1\cdots S_n\neq0,S_n=b)=\frac{b}{n}N_n(0,b)p^{\frac{n+b}{2}}q^{\frac{n-b}{2}}$

$=\frac{b}{n}P(S_n=b)$

(游走的最远距离) 令$M_n=\max\{S_i:0\leq i\leq n\}$

当$S_0=0$时 $M_n\geq0$

定理6. $S_0=0$, $r\geq1$, 则有:

$$P(M_n\geq r,S_n=b)=\begin{cases}P(S_n=b) & b\geq r\\ (\frac{q}{p})^{r-b}P(S_n=2r-b) & b<r\end{cases}$$

分析 设$b<r$, 记$Q=\{(0,0)\to(n,b)$且过某点$(i,r)\}$

$\forall\pi\in Q$, 记(i_π,r)为π与$y=r$的第一个交点，然后反射(i_π,r)后面部分得到新轨道π'，连接$(0,0)$与$(n,2r-b)$)，易知有1-1对应

另一方面: $\frac{P(\pi)}{P(\pi')}=\frac{p^{\frac{n-i_\pi+b-r}{2}}q^{\frac{n-i_\pi-b+r}{2}}}{p^{\frac{n-i_\pi-b+r}{2}}q^{\frac{n-i_\pi+b-r}{2}}}=(\frac{q}{p})^{r-b}$

因此: $\sum_{\pi\in Q}P(\pi)=(\frac{q}{p})^{r-b}\sum_{\pi'}P(\pi')$ ~~$=(\frac{q}{p})^{r-b}P(S_n\ldots)$~~

$=(\frac{q}{p})^{r-b}P(S_n=2r-b)$

当$p=\frac{1}{2}$时,

$P(M_n\geq r)=P(S_n\geq r)+\sum_{b=-\infty}^{r-1}(\frac{q}{p})^{r-b}P(S_n=2r-b)$

$\overset{c=2r-b}{=\!=}P(S_n\geq r)+\sum_{c=r+1}^{\infty}(\frac{q}{p})^{c-r}P(S_n=c)$

$=P(S_n=r)+\sum_{c=r+1}^{\infty}[1+(\frac{q}{p})^{c-r}]P(S_n=c)$

$\overset{p=\frac{1}{2}}{=\!=}P(S_n=r)+2P(S_n\geq r+1)$

§3.5 随机游走II

(首次到达某点) 定理7.(首中时定理).

$S_0=0$，在时刻 n 首次击中点 b 概率为

$$f_b(n)=\frac{|b|}{n}P(S_n=b),\qquad n\geq 1$$

分析: 不妨设 $b>0$，注意此时 $t=n$ 时达到新的最大值点，因此

$$f_b(n)=P(S_n=b,M_{n-1}=b-1,S_{n-1}=b-1)$$
$$=P(M_{n-1}=S_{n-1}=b-1,X_n=1)$$
$$=pP(M_{n-1}=b-1,S_{n-1}=b-1)$$
$$=p\big(P(M_{n-1}\geq b-1,S_{n-1}=b-1)-P(M_{n-1}\geq b,S_{n-1}=b-1)\big)$$
$$\overset{\text{定理6}}{=}pP(S_{n-1}=b-1)-p\cdot\frac{q}{p}P(S_{n-1}=b+1)$$
$$=\frac{n+b}{2n}P(S_n=b)-\frac{n-b}{2n}P(S_n=b)$$
$$=\frac{b}{n}P(S_n=b).$$

$$p\cdot P(S_{n-1}=b-1)=pC_{n-1}^{\frac12(n-1+b-1)}p^{\frac12(n-1+b-1)}\cdot q^{\frac12(n-1-b+1)}=\frac{n+b}{2n}P(S_n=b)$$

反正弦律

(最后一次访问原点) 定理8. $p=\frac12$，$S_0=0$，

记 $T_{2n}=\max\{0\leq i\leq 2n: S_i=0\}$，

则 $P(T_{2n}=2k)=P(S_{2k}=0)P(S_{2n-2k}=0)$

分析: $P(T_{2n}=2k)=P(S_{2k}=0,S_{2k+1}\cdots S_{2n}\neq 0)$

$$=P(S_{2k}=0)P(S_{2k+1}\cdots S_{2n}\neq0\mid S_{2k}=0)$$
$$\overset{\text{时齐}}{=}P(S_{2k}=0)P(S_1\cdots S_{2n-2k}\neq0\mid S_0=0)$$

令 $m=n-k$，求 $P(S_1S_2\cdots S_{2m}\neq 0)$

$$\overset{\text{定理5}}{=}\frac{1}{2m}E[|S_{2m}|]$$
$$=\frac{2}{2m}\sum_{i=1}^m 2iP(S_{2m}=2i)$$
$$=2\Big(\sum_{i=1}^m\frac{m+i-(m-i)}{2m}C_{2m}^{m+i}\Big)\cdot\Big(\frac12\Big)^{2m}$$
$$=2^{1-2m}\sum_{i=1}^m\big(C_{2m-1}^{m+i-1}-C_{2m-1}^{m+i}\big)$$
$$=2^{-2m}C_{2m}^m$$
$$=P(S_{2m}=0)$$

反正弦律 (Idea)

Stirling公式: $n!\sim\left(\frac{n}{e}\right)^n\sqrt{2\pi n}\qquad n\to\infty$

$P(S_{2k}=0)=\left(\frac12\right)^{2k}C_{2k}^k\sim\frac{1}{\sqrt{\pi k}}$, $k\to\infty$

$P(S_{2n-2k}=0)\sim\frac{1}{\sqrt{\pi(n-k)}}\qquad n-k\to\infty$

则 $P(T_{2n}\leq 2xn)=\sum_{k:k\leq xn}P(T_{2n}=2k)$

$$\sim\sum_{\frac kn\leq x}\frac{1}{\pi\sqrt{k(n-k)}}=\sum_{\frac kn\leq x}\frac1\pi\cdot\frac1n\cdot\frac{1}{\sqrt{\frac kn(1-\frac kn)}}$$
$$\sim\int_0^x\frac{1}{\pi\sqrt{u(1-u)}}du$$
$$=\frac2\pi\arcsin\sqrt x\qquad \frac{T_{2n}}{2n}\text{渐近分布}\frac2\pi\arcsin\sqrt x$$

$\mathbb{Z}^d$ 上随机游走

向量 $\{X_i\}$ 独立同分布，$P(X_i=\pm e_k)=\frac{1}{2d}$，

$k=1,\cdots,d$，$e_k=(0,\cdots,0,1,0,\cdots,0)$

$S_n=S_0+\sum_{i=1}^n X_i$.

$d=2$ 平面上随机游走，求 $P(S_{2n}=0)$. 这时 $S_0=0$

解: $P(S_{2n}=0)=\sum_{k=0}^n P(\text{向左右}k\text{次},\text{向上下}n-k\text{次})$

$$=\sum_{k=0}^n\frac{(2n)!}{k!k!(n-k)!(n-k)!}\left(\frac14\right)^{2n}$$
$$=C_{2n}^n\left(\frac14\right)^{2n}\sum_{k=0}^n(C_n^k)^2=\left(\frac14\right)^{2n}(C_{2n}^n)^2$$

§3.6. 母函数

A generating function is a device somewhat similar to a bag. Instead of carrying many little objects detatchedly, which could be embarrassing. We put them all in a bag, and we have only one object to carry, the bag. Pólya 1954

De Moivre 1730 引入

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k = \sum_{k=0}^{\infty}\binom{\alpha}{k}x^k$$

有 $G(s)=\frac{1}{2s}\left(1-\sum_{n=0}^{\infty}\binom{\frac12}{n}(-4s)^n\right)$

$=\frac{1}{2s}\left(-\sum_{n=1}^{\infty}\frac{\frac12(\frac12-1)\cdots(\frac12-n+1)}{n!}(-4s)^n\right)$

$=\sum_{n=1}^{\infty}\frac{2^{2n-1}(2n-3)!!}{n!}s^{n-1}$

$=\sum_{n=0}^{\infty}\frac{(2n)!}{(n+1)!n!}s^n$

可知: $C_n=\frac{(2n)!}{n!(n+1)!}$ $\boxed{\text{Catalan数}}$ 200多种场合出现

①数列的母函数

$\{a_i\}_{i=0}^{\infty}$ 母函数 $G_a(s)=\sum_{i=0}^{\infty}a_is^i$

例如: $C_n^0, C_n^1, \cdots, C_n^n$. $G(s)=\sum_{k=0}^{n}C_n^ks^k=(1+s)^n$

卷积: $\{a_i\}$与$\{b_i\}$的卷积$\{c_i\}$ $c_n=\sum_{k=0}^{n}a_kb_{n-k}$

记 $c=a*b$, 则 $G_c(s)=\sum_n c_ns^n$

$=\sum_{n=0}^{\infty}s^n\sum_{k=0}^{n}a_kb_{n-k}=\sum_{k=0}^{\infty}a_ks^k\sum_{n=k}^{\infty}s^{n-k}b_{n-k}$

$=G_a(s)G_b(s)$

即 $G_c(s)=G_a(s)G_b(s)$

例1. 对称随机游走 S_n, $S_0=0$, 求

$P(S_1\geq0,\cdots,S_{2n}\geq0, S_{2n}=0)$?

分析: 记 $C_n=\#\{(S_1,\cdots,S_{2n}): S_i\geq0,\forall i, S_{2n}=0\}$

则所求为 $(\frac12)^{2n}\cdot C_n$

设在 $t=2k$ 时刻首次返回原点, 则从 $t=0$ 到 $t=2k$ 轨道个数为 C_{k-1}, 则有

$C_n=\sum_{k=1}^{n}C_{k-1}C_{n-k}$, $C_0=1$

令 $G(s)=\sum_{i=0}^{\infty}C_is^i$, 则由 $C_n=\sum_{k=0}^{n-1}C_kC_{n-1-k}$

可知: $\frac{G(s)-1}{s}=G^2(s)$.

解之: $G(s)=\frac{1-\sqrt{1-4s}}{2s}$, $G(0)=1$

②非负整值随机变量.

定义1. X的概率母函数

$G_X(s)=E[s^X]$

显见, $G_X(s)=\sum_{i=0}^{\infty}s^iP(X=i)=\sum_{i=0}^{\infty}f(i)s^i$

$G(s)$有收敛半径 $R\geq1$, 在 $(-R,R)$ 内可微

$G(0)=P(X=0)$, $G(1)=1$ $f(i)=\frac{G^{(i)}(0)}{i!}$

· $G(s)$为概率母函数条件: $G(s)$系数非负, $G(1)=1$.

③几个分布的母函数.

(1) 二项分布: $B(n,p)$

$G(s)=\sum_{i=0}^{n}C_n^ip^iq^{n-i}s^i=(ps+q)^n$

(2) 几何分布

$P(X=i)=pq^{i-1}$, $i=1,2\cdots$

$G(s)=\sum_{i=1}^{\infty}pq^{i-1}s^i=\frac{ps}{1-qs}$

(3) Poisson分布

$P(X=i)=\frac{\lambda^i}{i!}e^{-\lambda}$ $i=0,1,2\cdots$

$G(s)=\sum_{i=0}^{\infty}s^i\frac{\lambda^i}{i!}e^{-\lambda}=e^{\lambda(s-1)}$

④母函数性质与数字特征矩

定理1. 记 $G(s)=E[s^X]$，则

(i) $E[X]=G'(1)$

(ii) $E[X(X-1)\cdots(X-k+1)]=G^{(k)}(1)$

(iii) $Var(X)=G''(1)+G'(1)(1-G'(1))$

这里 $G^{(k)}(1)=\lim_{s\uparrow 1}G^{(k)}(s)$ (Abel定理)

分析：当 $|s|<1$ 时，对 $G(s)=\sum_{i=0}^{\infty}s^i f(i)$

求导 $G^{(k)}(s)=\sum_i i(i-1)\cdots(i-k+1)s^{i-k}f(i)$

$=E[s^{X-k}X(X-1)\cdots(X-k+1)]$

让 $s\uparrow 1$ 即可

与随机变量卷积

定理2. 设 $X_1,\cdots,X_n$ 相互独立，~~$G_k(s)=E[X$~~

$G_k(s)=E[s^{X_k}]$，则 $Y=\sum_{k=1}^{n}X_k$，母函数为

$G_Y(s)=G_1(s)\cdots G_n(s)$

分析：$G_Y(s)=E[s^Y]=E[s^{X_1}\cdots s^{X_n}]$

$=G_1(s)\cdots G_n(s)$

定理3. 设 $\{X_k:k\geq 1\}$ 相互独立（独立同分布，指任意有限多个相互独立）且服从相同分布，共同母函数 G_X，又 N 与 $\{X_k:k\geq 1\}$ 独立，母函数为 G_N，则 $Y=\sum_{k=1}^{N}X_k$ 有母函数

$G_Y(s)=G_N(G_X(s))$

注：$P(N=n)=1$，此时定理3约化到定理2

$G_Y(s)=E[s^Y]=\sum_{n=0}^{\infty}E[s^Y|N=n]P(N=n)$

$=\sum_n E[s^{X_1+\cdots+X_n}|N=n]P(N=n)$

$\overset{独立}{=}\sum_n (G_X(s))^n P(N=n)=G_N(G_X(s))$

定义2. (X,Y) 非负整值随机向量，联合概率母函数 $G(s,t)=E[s^X t^Y]$

说明：$G(s,t)=\sum_{i,j=0}^{\infty}s^i t^j P(X=i,Y=j)$

定理4. X 与 Y 独立 $\Leftrightarrow G(s,t)=G_X(s)G_Y(t)$

证明："$\Rightarrow$"显见

"$\Leftarrow$"比较两边展开式 $s^i t^j$ 系数可知：

$P(X=i,Y=j)=P(X=i)P(Y=j)$，表明独立

例2. 掷3颗均匀骰子，求点数和为9概率？

分析：X_k 表示第 k 个骰子点数，则 $Y=X_1+X_2+X_3$

因为 $G_{X_k}(s)=\sum_{i=1}^{6}\frac{1}{6}s^i=\frac{1}{6}\cdot\frac{s(1-s^6)}{1-s}$

可知 $G_Y(s)=\frac{1}{6^3}\frac{s^3(1-s^6)^3}{(1-s)^3}$

$=\frac{1}{6^3}s^3(1-3s^6+3s^{12}-s^{18})\underbrace{\sum_{l=0}^{\infty}\binom{-3}{l}(-s)^l}_{\binom{l+2}{l}s^l}$

s^9 系数：$\frac{1}{6^3}(C_8^6-3)=\frac{25}{216}$

二项分布再生性.

$X_1\sim B(n_1,p)$，$X_2\sim B(n_2,p)$，X_1 与 X_2 独立

则 $G_{X_1+X_2}(s)=(ps+q)^{n_1+n_2}$，~~$X_1+X_2\sim B($~~

$X_1+X_2\sim B(n_1+n_2,p)$

Poisson 分布：

$X_1\sim P(\lambda_1)$ $X_2\sim P(\lambda_2)$，X_1 与 X_2 独立，

则 $G_{X_1+X_2}(s)=e^{(\lambda_1+\lambda_2)(s-1)}$

$X_1+X_2\sim P(\lambda_1+\lambda_2)$

$G(s)=\sum_i s^{x_i}P(X=x_i)$

如何定义"概率母函数"?(第五章)

习题.

(I) Feller: 抽查一个家庭,令

$A=\{$至多一个女孩$\}$ $B=\{$男女孩子都有$\}$

试验证:有3孩子之家A与B独立,但有4孩子之家A与B不独立.

解:3孩子之家:

$P(A)=P(3\text{boy})+P(2\text{boy }1\text{girl})=\frac{1}{8}+\frac{3}{8}=\frac{1}{2}$

$P(B)=P(2\text{boy}1\text{girl})+P(2\text{girl }1\text{boy})=\frac{3}{8}+\frac{3}{8}=\frac{3}{4}$.

$P(AB)=P(2\text{boy }1\text{girl})=\frac{3}{8}=P(A)P(B)$, 独立

4孩子之家:

$P(A)=P(4\text{boy})+P(3\text{boy }1\text{girl})=\frac{1}{16}+\frac{4}{16}=\frac{5}{16}$

$P(B)=P(3\text{boy}1\text{girl})+P(2\text{boy }2\text{girl})+P(3\text{girl }1\text{boy})=\frac{7}{8}$

$P(AB)=P(3\text{boy }1\text{girl})=\frac{4}{16}\neq P(A)P(B)$, 不独立

注:独立性在理论上按数学恒等式来定义,在应用时常用直观,但直观不可靠.

(II) $|P(AB)-P(A)P(B)|\leq\frac{1}{4}$

•分析:令 $p=P(A\cap B)$

$q=P(A\setminus B)$ $r=P(B\setminus A)$

$S=P(A^cB^c)$, 则 $p+q+r+s=1$.

因此 $|P(AB)-P(A)P(B)|=|p-(p+q)(p+r)|$

$=|p(1-p-q-r)-qr|=|ps-qr|$

$\leq\max\{ps,qr\}\leq\frac{1}{4}\max\{(p+s)^2,(q+r)^2\}\leq\frac{1}{4}$

"=" $\Leftrightarrow$ $ps=0$ 或 $qr=0$, 且 $p=s=\frac{1}{2}$ 或 $q=r=\frac{1}{2}$

•方法二,考察 I_A, I_B

则 $E[I_A]=P(A)$ $E[I_B]=P(B)$,

$E[I_AI_B]=E[I_{AB}]=P(AB)$,

因此 $E[I_AI_B]-E[I_A]E[I_B]=Cov(X,Y)$

当 $0<P(A), P(B)<1$时, $|\rho(I_A,I_B)|\leq1$,

有: $|Cov|\leq\sqrt{Var(I_A)Var(I_B)}$ (Bernoulli) $\leq\frac{1}{4}$

(III) 联合分布列

$y \backslash x$	-1	0	1	
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	

证明:X与Y不相关,也不独立.

解:$E[X]=0$. $E[Y]=\frac{2}{3}$. $E[XY]=0$

$Cov(X,Y)=0$, 不相关

$P(X=-1,Y=0)=0$

$P(X=-1)=\frac{1}{3}$ $P(Y=0)=\frac{1}{3}$ } 不独立

(IV) 对称随机游走 S_n, $S_0=0$, 记

$\tau_1=\inf\{n\geq0: S_n=1\}$,

证明: $P(\tau_1<\infty)=1$

证明:

$P(\tau_1=2r-1)\overset{\text{首中时}}{=\!=}\frac{1}{2r-1}P(S_{2r-1}=1)$

$=\frac{1}{2r-1}\left(\frac{1}{2}\right)^{2r-1}C_{2r-1}^{r}=\frac{1}{2r-1}P(S_{2r}=0)$

令 $u_r=P(S_{2r}=0)$, 则

$1-\sum_{r=1}^{m}\frac{u_r}{2r-1}=u_m$ 数学归纳法,

因此 $P(\tau_1>2m-1)=1-\sum_{r=1}^{m}P(\tau_1=2r-1)$

$=u_m=(\frac{1}{2})^{2m}C_{2m}^{m}=(\frac{1}{2})^{2m}\frac{(2m)!}{(m!)^2}\sim\frac{1}{\sqrt{\pi m}}$

$\to 0,\ n\to\infty$

于是 $P(\tau_1=\infty)=0$.

注:对称随机游走从原点出发迟早到达1

第四章 连续性随机变量

Born Rule (1926) 波函数统计诠释

The observation probability density = The absolute ~~value~~ square value of the wave function.

$$f(x)=|\psi(x,t)|^2$$

Schrödinger Equation (1925)

$$i\hbar\frac{\partial}{\partial t}\psi(x,t)=(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+V(x))\psi(x,t).$$

§4.1. 密度函数、独立性

X连续型、非负可积, $f:\mathbb{R}\to[0,\infty)$ 密度

分布函数 $F(x)=\int_{-\infty}^{x}f(u)du$

性质: (i) $\int_{-\infty}^{\infty}f(u)du=1$

(ii) $P(X=x)=0$

(iii) $P(a\le X\le b)=\int_a^b f(u)du$

注: 非负可积函数f为密度 $\Leftrightarrow\int_{\mathbb{R}}f(x)dx=1$

典型例子

(i) 均匀分布 $X\sim U[a,b]$ Uniform.

$$f(x)=\begin{cases}\frac{1}{b-a}, & x\in[a,b]\\ 0 & \text{否则}\end{cases}$$

$\forall[c,d]\subset[a,b]$ 有 $P(c<X\le d)=\frac{d-c}{b-a}$

背景: 均匀、"Lebesgue测度"

(ii) 指数分布 $Exp(\lambda),\lambda>0$

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x\ge 0\\ 0 & x<0\end{cases}$$

分布函数 $F(x)=\begin{cases}1-e^{-\lambda x} & x\geq 0\\ 0 & x<0\end{cases}$

背景:独立事件发生的时间间隔,如电子产品寿命, Wiki 百科新条目等待时间,分析如下:已使用 t 小时产品在以后 Δt 小时失效概率为 $\lambda\Delta t+o(\Delta t)$,则寿命 X 服从指数分布? $(\Delta t>0)$

$P(t<X\leq t+\Delta t \mid X>t)=\lambda\Delta t+o(\Delta t)$

因此 $LHS=\frac{P(t<X\leq t+\Delta t)}{P(X>t)}=\frac{F(t+\Delta t)-F(t)}{1-F(t)}$

进而 $\frac{F(t+\Delta t)-F(t)}{\Delta t}=(1-F(t))(\lambda+o(1))$

$\Delta t\to 0$ 有: $F'(t)=(1-F(t))\lambda$

初始条件 $F(0)=0$ (寿命为0概率为0)

解之 $F(t)=1-e^{-\lambda t}$, $t\geq 0$

反之,若 $X\sim Exp(\lambda)$, 则 $P(X>t+s \mid X>t)$
$=\frac{P(X>t+s)}{P(X>t)}=e^{-\lambda s}=P(X>s)$

(永远"年轻",几何分布的连续版本)

(iii) 正态分布(高斯分布) $N(\mu,\sigma^2)$

$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ $-\infty<x<\infty$

$\mu\in\mathbb{R}$, $\sigma>0$ (钟形曲线)

几何特征:

① $x=\mu$ 为对称轴, $f_{max}=\frac{1}{\sqrt{2\pi\sigma^2}}$

② $x=\mu\pm\sigma$ 为 f 的拐点

背景:测量误差、概率论成绩、二项分布极限方程、几何和数论等分支中重要。

如: $\frac{\partial u}{\partial t}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2}$, $u(0,x)=g(x)$,

Solution: $u(t,x)=\frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}}e^{-\frac{1}{2t}(x-y)^2}g(y)dy$

(iv) $\boxed{\text{Wigner 半圆律:}}$

$f(x)=\frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2-x^2}$ $-2\sigma\leq x\leq 2\sigma$

背景:随机矩阵和自由概率论(扮演正态分布角色)

例1. 验证半圆律为密度,并求 $P(0<X<\sigma)$?

解: $\int_{-2\sigma}^{2\sigma}\sqrt{4\sigma^2-x^2}\,dx \xlongequal{x=2\sigma\sin\theta}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}4\sigma^2\cos^2\theta\,d\theta$

$=4\sigma^2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{1}{2}(1+\cos 2\theta)d\theta$

$=4\sigma^2\cdot\frac{1}{2}(\theta+\frac{1}{2}\sin 2\theta)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}=2\pi\sigma^2$

$P(0<X<\sigma)=\frac{1}{2\pi\sigma^2}\cdot 4\sigma^2\cdot\frac{\theta+\frac{1}{2}\sin 2\theta}{2}\Big|_0^{\frac{\pi}{6}}=\frac{1}{6}+\frac{\sqrt{3}}{4\pi}$

注:正态分布通常用 $\boxed{\phi(x)}$ 表密度,
$\boxed{\Phi(x)}=\int_{-\infty}^{x}\phi(u)du$

$\mu=0$、$\sigma=1$ 时称为 $\boxed{\text{标准正态分布}}$

例2. $X\sim U(0,1)$, 则 $Y=\Phi^{-1}(X)\sim N(0,1)$

解: $P(Y\leq y)=P(\Phi^{-1}(X)\leq y)=P(X\leq\Phi(y))=\Phi(y)$

独立性 ← 一般随机变量独立性定义

定义1. $X_1,X_2,\cdots,X_n$ 为 $(\Omega,\mathcal{F},P)$ 上随机变量,若 $\forall x_1,\cdots,x_n\in\mathbb{R}$ 有:

$P(X_1\leq x_1,\cdots,X_n\leq x_n)=P(X_1\leq x_1)\cdots P(X_n\leq x_n)$

称 $X_1,\cdots,X_n$ 相互独立

一个等价刻划

定理1. $X_1,\cdots,X_n$ 相互独立 $\Leftrightarrow\forall B_1,\cdots,B_n\in\mathcal{B}(\mathbb{R})$

有: $P(X_1\in B_1,\cdots,X_n\in B_n)=P(X_1\in B_1)\cdots P(X_n\in B_n)$

证明: $\Leftarrow$ 显然

$\Rightarrow$ 不显然、用到更多抽象测度论,(见高概),略

定理2. 设 $g_1,\cdots,g_n$ 为1元Borel可测函数,若 $X_1,\cdots,X_n$ 相互独立,则 $g_1(X_1),\cdots,g_n(X_n)$ 亦相互独立

证明: 记 $B_i=\{x\in\mathbb{R}: g_i(x)\le x_i\}$, $i=1,\cdots,n$

则 $B_i\in\mathcal{B}(\mathbb{R})$, 且 $\{g_i(X_i)\le x_i\}=\{X_i\in B_i\}$

于是 $P(g_1(X_1)<x_1,\cdots,g_n(X_n)<x_n)$

$=P(X_1\in B_1,\cdots,X_n\in B_n)$

$\xlongequal[\text{定理1}]{\text{独立}} P(X_1\in B_1)\cdots P(X_n\in B_n)$

$=P(g_1(X_1)<x_1)\cdots P(g_n(X_n)<x_n)$

定理3. 设 $X_1,\cdots,X_n$ 分别有密度 $f_1(x)\cdots f_n(x)$,则 $X_1,\cdots,X_n$ 相互独立 $\Leftrightarrow$ $\prod_{i=1}^n f_i(x_i)$ 为 $(x_1,\cdots,x_n)$ 的联合密度

分析: $\Leftarrow$: 首先 f_i 为 X_i 密度,显然

$\Rightarrow$: $P(X_1\le x_1,\cdots,X_n\le x_n)$

$=P(X_1\le x_1)\cdots P(X_n\le x_n)$

$=\int_{-\infty}^{x_1}f_1(u_1)du_1\cdots\int_{-\infty}^{x_n}f_n(u_n)du_n$

$=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\underbrace{f_1(u_1)\cdots f_n(u_n)}_{\text{联合密度}}du_1\cdots du_n.$

卷积: 连续型X与Y独立时,称 $Z=X+Y$ 为X与Y卷积,记 $f_Z=f_X*f_Y$

定理4. (X,Y) 有密度 $f(x,y)$,则 $Z=X+Y$ 有密度 $f_Z(z)=\int_{-\infty}^{\infty}f(x,z-x)dx=\int_{-\infty}^{\infty}f(z-y,y)dy$

分析: $P(Z\le z)=P(X+Y\le z)$

$=\iint_{x+y\le z}f(x,y)dxdy$

$=\int_{-\infty}^{\infty}\int_{-\infty}^{z-x}f(x,y)dydx$

$\xlongequal{y\to y-x}\int_{-\infty}^{\infty}\int_{-\infty}^{z}f(x,y-x)dydx$

$=\int_{-\infty}^{z}\int_{-\infty}^{\infty}f(x,y-x)dxdy$

独立时, $f(x,y)=f_X(x)f_Y(y)$

$\frac{1}{\sqrt{2\pi}}e^{-\frac12x^2}$

例3. X与Y独立同 $N(0,1)$,求X+Y密度?

解: $f_{X+Y}(z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\frac12x^2-\frac12(z-x)^2}dx$

$=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-(x-\frac12z)^2-\frac14z^2}dx$

$=\frac{1}{2\pi}e^{-\frac14z^2}\int_{\mathbb{R}}e^{-x^2}dx$

$=\frac{1}{2\sqrt{\pi}}e^{-\frac14z^2}$, 即 $X+Y\sim N(0,2)$

§4.2 条件期望

定义1. X有密度f,当 $\int_{-\infty}^{\infty}|x|f(x)dx<\infty$ 时,称 $E[X]=\int_{-\infty}^{\infty}xf(x)dx$ 为X的(数学)期望

引理1. 连续型X分布为F,期望存在,则

$E[X]=\int_0^{\infty}(1-F(x))dx-\int_{-\infty}^0F(x)dx$

证明: $E[X]=\int_{-\infty}^{\infty}xf(x)dx$

$=\int_0^{\infty}tf(t)dt+\int_{-\infty}^0tf(t)dt$

$= \int_0^{\infty} (\int_0^t 1 dx) f(t) dt - \int_{-\infty}^0 (\int_t^0 1 dx) f(t) dt$

$\overset{Fubini}{=} \int_0^{\infty} (\int_x^{\infty} f(t) dt) dx - \int_{-\infty}^0 (\int_{-\infty}^x f(t) dt) dx$

$= \int_0^{\infty} (1-F(x)) dx - \int_{-\infty}^0 F(x) dx$

定理2. $g: \mathbb{R} \to \mathbb{R}$ Borel可测，X和$g(X)$均为连续型且期望均存在，则（何时$g(X)$连续型？）

$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$ （思考$g(x)=x^k$）

这里f为X的密度函数.

证明：记$Y=g(X)$由引理1，

$\mathbb{E}[Y] = \int_0^{\infty} P(Y>y) dy - \int_{-\infty}^0 P(Y \le y) dy$

$= \int_0^{\infty} (\int_{\{x: g(x)>y\}} f(x) dx) dy - \int_{-\infty}^0 (\int_{\{x: g(x)<y\}} f(x) dx) dy$

$\overset{Fubini}{=} \int_{\{x: g(x)>0\}} (\int_0^{g(x)} dy) f(x) dx$

$- \int_{\{x: g(x)<0\}} (\int_{g(x)}^0 dy) f(x) dx$

$= \int_{\{x: g(x)>0\}} g(x) f(x) dx - \int_{\{x: g(x)<0\}} (-g(x)) f(x) dx$

定理3. (X,Y)有联合密度$f(x,y)$，$g: \mathbb{R}^2 \to \mathbb{R}$ Borel可测函数且$g(X,Y)$为连续型，其期望存在，则$\mathbb{E}[g(X,Y)] = \iint g(x,y) f(x,y) dxdy$

特别地，若$g(x,y)=ax+by$时，

$\mathbb{E}[aX+bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$. 线性性

思考：$aX+bY$是否为连续型？$a^2+b^2 \neq 0$.

定理2和3可用来计算数字特征、相关系数等。

矩 $m_k = \mathbb{E}[X^k]$　中心矩 $\sigma_k = \mathbb{E}[(X-\mathbb{E}[X])^k]$

方差 $Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

标准差 $\sigma = \sqrt{Var(X)}$

当(X,Y)为连续型，定义协方差

$Cov(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$

相关系数 $\rho(X,Y) = \dfrac{Cov(X,Y)}{\sqrt{Var(X) Var(Y)}}$

$|\rho(X,Y)| \le 1$

定理4. (Cauchy-Schwarz不等式)

设(X,Y)为连续型，则 $(\mathbb{E}[XY])^2 \le (\mathbb{E}[X])^2 (\mathbb{E}[Y])^2$

注：当存在$a \in \mathbb{R}$, s.t. $P(Y=aX)=1$时等号成立。事实上，等号成立$\Leftrightarrow$存在不全为0的实数a,b s.t. $P(aX=bY)=1$ 但是这时(X,Y)不可能连续型.

（后面对一般随机变量定义期望）

例1. $X \sim N(\mu, \sigma^2)$，即X有密度.

$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty$

解：$\mathbb{E}[X] = \int_{\mathbb{R}} x f(x) dx$

$= \int_{\mathbb{R}} (x-\mu) f(x) dx + \int_{\mathbb{R}} \mu f(x) dx$

$\overset{x-\mu=y}{=} \int_{\mathbb{R}} y \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} y^2} dy + \mu = \mu$.

$Var[X] = \mathbb{E}[(X-\mu)^2] = \int_{\mathbb{R}} (x-\mu)^2 f(x) dx$

$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} y^2 e^{-\frac{1}{2\sigma^2} y^2} dy$

$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{1}{2} y^2} dy$

$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} y^2} dy = \sigma^2$

例2. 柯西分布：$x_0 \in \mathbb{R}$, $\gamma > 0$,

$f(x) = \frac{1}{\pi} \cdot \frac{\gamma}{\gamma^2 + (x-x_0)^2} \quad -\infty < x < \infty$

当 $x_0=0, \gamma=1$ 时, $\int_{-\infty}^{\infty}|x|f(x)dx$
$=\int_{-\infty}^{\infty}\frac{x}{\pi(1+x^2)}dx=\infty$

期望不存在。与正态分布比较:
柯西分布扁、宽;正态分布高、窄

例3. $\boxed{\text{二元正态}}$

$$f(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)}$$

$-\infty<x,y<\infty$, 这里 $-1<\rho<1$.

解: $f_X(x)=\int_{\mathbb{R}}\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}[(y-\rho x)^2+(1-\rho^2)x^2]}dy$
$=e^{-\frac{1}{2}x^2}\cdot\frac{1}{\sqrt{2\pi}}$. 即 $X\sim N(0,1)$, $Y\sim N(0,1)$,
表明 $E[X]=0$, $E[Y]=0$

$Cov(X,Y)=E[XY]=\iint xy\,f(x,y)\,dxdy$
$=\iint(x(y-\rho x)+\rho x^2)f(x,y)\,dxdy$
$=\iint x(y-\rho x)f(x,y)dxdy+\iint x^2 f(x,y)\,dxdy\cdot\rho$
$=\rho\int\frac{1}{\sqrt{2\pi}}x^2e^{-\frac{1}{2}x^2}dx=\rho$

ρ 为相关系数, 且 $\rho=0\Leftrightarrow X,Y$ 独立

条件期望:

直观上, $P(Y\le y\mid \cancel{X=x}\ x<X\le x+\Delta x)$
$=\frac{P(Y\le y,\ x<X\le x+\Delta x)}{P(x<X\le x+\Delta x)}$
$=\frac{\int_{-\infty}^{y}\int_{x}^{x+\Delta x}f(u,v)\,dudv}{\int_{x}^{x+\Delta x}f_X(u)\,du}$
$\xrightarrow{\Delta x\to 0}\frac{\int_{-\infty}^{y}f(x,v)\Delta x\,dv}{f_X(x)\Delta x}=\frac{\int_{-\infty}^{y}f(x,v)dv}{f_X(x)}$

定义2. (X,Y) 联合密度 $f(x,y)$ 给 $X=x$
$(f_X(x)>0)$ 下 Y 的条件密度:

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}$$ 关于 y 构成密度函数

条件分布:

$$F_{Y|X}(y|x)=\int_{-\infty}^{y}\frac{f(x,v)}{f_X(x)}dv$$

条件期望: (Y 关于 X 的条件期望 $\psi(X)$)

$E[Y|X]=\psi(X)$, $\psi(x)=E[Y|X=x]$
$=\int_{\mathbb{R}}y f_{Y|X}(y|x)\,dy$

定理5. $E[E[Y|X]]=E[Y]$
或 $E[Y]=\int_{\mathbb{R}}E[Y|X=x]f_X(x)dx$
期望形式的全概公式

例4. 见例3二元正态, 问 $f_{Y|X}(y|x)$
$=\frac{1}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{1}{2(1-\rho^2)}(y-\rho x)^2}$
关于 y: $N(\rho x, 1-\rho^2)$, 表明 $E[Y|X=x]=\rho x$
$E[Y|X]=\rho X$

例5. 设 (X,Y) 在 $D=\{(x,y)\in\mathbb{R}^2: x^2+y^2\le 4\sigma^2\}$,
上均匀分布, 即: $f(x,y)=\begin{cases}\frac{1}{4\pi\sigma^2} & (x,y)\in D\\ 0 & \text{否则}\end{cases}$

解: $f_X(x)=\frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2-x^2}$, $-2\sigma\le x\le 2\sigma$,
(Wigner半圆律)

$Cov(X,Y)=\iint_D xy\,f(x,y)\,dxdy=0$.
表明 X 与 Y 不相关, 又 $f(x,y)\ne f_X(x)f_Y(y)$,
表明 X 与 Y 不独立.

另外: $f_{Y|X}(y|x)=\frac{1}{2\sqrt{4\sigma^2-x^2}}$　$(x,y)\in D$

$E[Y|X=x]=0$, $E[Y|X]=0$

从随机向量到随机向量:

$(X_1,X_2)\longrightarrow(Y_1,Y_2)=T(X_1,X_2)$

$y_i=y_i(x_1,x_2)$ $i=1,2$,

$T: D\subset\mathbb{R}^2\mapsto R\subset\mathbb{R}^2$ 一一映射,

$(x_1,x_2)\longrightarrow(y_1,y_2)$

其逆 $x_i=x_i(y_1,y_2)$ 有连续偏导数

定理: (X_1,X_2) 有密度 $f(x_1,x_2)$, 则

$f_{Y_1,Y_2}(y_1,y_2)=f(x_1(y_1,y_2),x_2(y_1,y_2))\cdot|J(y_1,y_2)|\cdot I_R$

这里 $J=\begin{vmatrix}\frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1}\\ \frac{\partial x_1}{\partial y_2} & \frac{\partial x_2}{\partial y_2}\end{vmatrix}$

证明: 本质变量替换, ~~设A⊂D, B⊂?~~

设 $A\subset D$, $B=T(A)\subset T(D)=R$

$(X_1,X_2)\in A\Leftrightarrow(Y_1,Y_2)\in B$

则 $P((Y_1,Y_2)\in B)=P((X_1,X_2)\in A)$

$=\iint_A f(x_1,x_2)\,dx_1dx_2$

$=\iint_B f(x_1(y_1,y_2),x_2(y_1,y_2))|J|dy_1dy_2$

(取 $B=T(D)\cap(-\infty,y_1]\times(-\infty,y_2]$)

注: 若 $D_0\subset D$, $P((X_1,X_2)\in D_0)=1$, T 在 D_0 上单射

结论亦对

例: 设 X,Y 独立同 $N(0,1)$, 令

$X=R\cos\Theta$ $Y=R\sin\Theta$, $R\ge0$, $0\le\Theta\le2\pi$

求 (R,Θ) 联合密度

解: $f(x,y)=\frac{1}{2\pi}e^{-\frac12(x^2+y^2)}$, 则

$f_{R,\Theta}(r,\theta)=\frac{1}{2\pi}re^{-\frac12r^2}$, $r>0$, $\theta\in[0,2\pi)$

表明 R 与 Θ 独立, 且 $\Theta\sim U[0,2\pi)$,

$f_R(r)=re^{-\frac12r^2}$, $r>0$.

副产品: 产生独立正态随机数: U_1,U_2 独立同 $U[0,1]$, 令 $X=\sqrt{-2\log U_1}\cos(2\pi U_2)$

$Y=\sqrt{-2\log U_1}\sin(2\pi U_2)$

则 X 与 Y 独立同 N_1

§4.3. 多元正态分布

高斯过程.	高斯自由场
高斯滤波	高斯随机矩阵
高斯信道	布朗运动
高斯噪声	Wick Theorem. (量子场论)
高斯测度	

标准正态 $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac12x^2}$

二元正态 $f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac12Q(x_1,x_2)}$.

$-1<\rho<1$, $\sigma_1,\sigma_2>0$,

$Q=\frac{1}{1-\rho^2}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\frac{x_1-\mu_1}{\sigma_1}\frac{x_2-\mu_2}{\sigma_2}+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)$

更一般, 二次型

$Q(x_1,\cdots,x_n)=\sum_{i,j=1}^n a_{ij}x_ix_j=\vec{x}A\vec{x}'$

$\boxed{\text{定义1}}$ $\vec{X}=(X_1,\cdots,X_n)$ 服从n维正态分布, 若其密度 $f(\vec{x})=\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}\exp[-\frac12(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})']$

这里 $\vec{\mu}\in\mathbb{R}^n$, $\Sigma=(\sigma_{ij})_{i,j=1}^n$ 为正定矩阵,

记 $\vec{X}\sim N(\mu,\Sigma)$

$n=2$ 时, $\Sigma=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$

$$\Sigma^{-1}=\frac{1}{1-\rho^2}\begin{pmatrix}\frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2}\\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2}\end{pmatrix}$$

定理1. $\vec{X}\sim N(\vec{\mu},\Sigma)$, 则

(i) $E[\vec{X}]=\vec{\mu}$, 即 $E[X_i]=\mu_i$

(ii) Σ 为 $\vec{X}$ 的协方差矩阵, 即 $Cov(X_i,X_j)=\Sigma_{ij}$

证明: 存在正交阵 B, 即 $BB^T=I$, s.t.

$\Sigma=B^T\Lambda B$, $\Lambda=diag(\lambda_1,\cdots,\lambda_n)$

令 $\vec{Y}=(\vec{X}-\vec{\mu})\cdot B^{-1}$, 即 $\vec{X}=\vec{\mu}+\vec{Y}\cdot B$

①先CHECK: $f(x)$ 为密度: (注: $\Sigma^{-1}=B^T\Lambda^{-1}B$)

$$\int_{\mathbb{R}^n}f(x)d^nx\xlongequal{\vec{x}=\vec{\mu}+\vec{y}B}\int\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}e^{-\frac{1}{2}\vec{y}\Lambda^{-1}\vec{y}^T}d^ny$$

又 $|\Sigma|=\prod_{k=1}^n\lambda_k$, 则

$$\int_{\mathbb{R}^n}f(x)d^nx=\prod_{k=1}^n\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi\lambda_k}}e^{-\frac{1}{2\lambda_k}y_k^2}dy_k=1$$

② 由 $X_i=\mu_i+\sum_{j=1}^n y_jb_{ji}$ 知:

$$E[X_i]=\int_{\mathbb{R}^n}x_if(x)d^nx$$
$$=\int_{\mathbb{R}^n}(\mu_i+\sum_{j=1}^n y_jb_{ji})\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}e^{-\frac{1}{2}\vec{y}\Lambda^{-1}\vec{y}^T}d^ny$$
$$=\mu_i$$

$$Cov(X_i,X_j)=\int_{\mathbb{R}^n}(x_i-\mu_i)(x_j-\mu_j)f(x)\,d^nx$$
$$=\int_{\mathbb{R}^n}(\sum_{k=1}^n y_kb_{ki})(\sum_{l=1}^n y_lb_{lj})\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}e^{-\frac{1}{2}\vec{y}\Lambda^{-1}\vec{y}^T}d^ny$$
$$=\sum_{k,l}b_{ki}b_{lj}\int_{\mathbb{R}^n}y_ky_l\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}e^{-\frac{1}{2}\vec{y}\Lambda^{-1}\vec{y}^T}d^ny$$

($k=l$ 时非零, 且等于 λ_k)

$$=\sum_k b_{ki}b_{kj}\lambda_k=(B^T\Lambda B)_{ij}=\Sigma_{(ij)}.$$

线性变换之下:

$\vec{X}\sim N(\vec{\mu},\Sigma)$, A: $n\times m$ 矩阵, $\vec{Y}=\vec{X}A$ 分布?

定理2. $\vec{X}\sim N(\vec{\mu},\Sigma)$, n 阶方阵 D 非奇异, 则

$\vec{Y}=\vec{X}D\sim N(\vec{\mu}D,D^T\Sigma D)$

证明: 令 $B=\{\vec{y}\in\mathbb{R}^n: a_i<y_i\le b_i,\ i=1,2,\cdots,n\}$

$A=\{\vec{x}\in\mathbb{R}^n:\vec{x}D\in B\}$

则 $P(\vec{Y}\in B)=P(\vec{X}\in A)=\int_A f(x)d^nx$

$$\xlongequal{\vec{x}=\vec{y}D^{-1}}\int_B\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}e^{-\frac{1}{2}(\vec{y}D^{-1}-\vec{\mu})\Sigma^{-1}(\vec{y}D^{-1}-\vec{\mu})^T}|\det D^{-1}|d^ny$$
$$=\int_B\frac{1}{\sqrt{(2\pi)^n|D^T\Sigma D|}}e^{-\frac{1}{2}(\vec{y}-\vec{\mu}D)(D^T\Sigma D)^{-1}(\vec{y}-\vec{\mu}D)^T}d^ny$$

即 $\vec{Y}\sim N(\vec{\mu}D,D^T\Sigma D)$

定理3. $\vec{X}\sim N(\vec{\mu},\Sigma)$ $\Sigma=\begin{pmatrix}\Sigma_1 & 0\\ 0 & \Sigma_{22}\end{pmatrix}$, 则

$\vec{X}^{(i)}\sim N(\vec{\mu}^{(i)},\Sigma_{ii})$, $i=1,2$,

其中 Σ_{ii} 为 n_i 阶矩阵, $\vec{\mu}^{(i)}$、$\vec{X}^{(i)}$ 为相应分量

证明: 变量分离, 联合密度为两部分之积, 且 $\vec{X}^{(1)}$ 与 $\vec{X}^{(2)}$ 独立.

注: $\vec{X}$ 与 $\vec{Y}$ 独立, 指 $\forall A,B$ 有 $P(\vec{X}\in A,\vec{Y}\in B)=P(\vec{X}\in A)P(\vec{Y}\in B)$.

定理4. $\vec{X}\sim N(\vec{\mu},\Sigma)$, $\vec{X}=(\vec{X}^{(1)},\vec{X}^{(2)})$,

$\vec{\mu}=(\vec{\mu}^{(1)},\vec{\mu}^{(2)})$, $\Sigma=\begin{pmatrix}\Sigma_1 & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}$,

则 $\vec{X}^{(1)}\sim N(\vec{\mu}^{(1)},\Sigma_{11})$

证明: $\begin{pmatrix}I_{n_1} & 0\\ -\Sigma_{21}\Sigma_1^{-1} & I_{n_2}\end{pmatrix}\begin{pmatrix}\Sigma_1 & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\underbrace{\begin{pmatrix}I_{n_1} & -\Sigma_1^{-1}\Sigma_{12}\\ 0 & I_{n_2}\end{pmatrix}}_{=D}$

$$=\begin{pmatrix}\Sigma_{11} & 0\\ 0 & \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\end{pmatrix}$$

即 $D^T\Sigma D=\begin{pmatrix}\Sigma_{11} & 0\\ 0 & *\end{pmatrix}$

令 $(\vec{Y}^{(1)},\vec{Y}^{(2)})=(\vec{X}^{(1)},\vec{X}^{(2)})D$，有

$(\vec{Y}^{(1)},\vec{Y}^{(2)})=(\vec{X}^{(1)},*)$，

由定理2知 $\vec{Y}\sim N(\vec{\mu}D,D^T\Sigma D)$，但 $\vec{\mu}D=(\vec{\mu}^{(1)},*)$，$D^T\Sigma D$ 分块对角，再由定理3，

$\vec{X}^{(1)}=\vec{Y}^{(1)}\sim N(\vec{\mu}^{(1)},\Sigma_{11})$

定理5 $\vec{X}\sim N(\vec{\mu},\Sigma)$，$n\times m$ 矩阵 A 秩为 m $(m\le n)$，则 $\vec{Y}=\vec{X}A\sim N(\vec{\mu}A,A^T\Sigma A)$

证明 当 $m<n$ 时，取 $n\times(n-m)$ 矩阵 B，s.t. $D=(A\ B)$ 非奇异，则 $(\vec{X}A,\vec{X}B)=\vec{X}D$，

$\vec{X}D\sim N(\vec{\mu}D,D^T\Sigma D)$

因 $\vec{\mu}D=(\vec{\mu}A\ \ \vec{\mu}B)$

$D^T\Sigma D=\begin{pmatrix}A^T\Sigma A & A^T\Sigma B\\ B^T\Sigma A & B^T\Sigma B\end{pmatrix}$，定理4即可.

独立性、不相关性等价(协方差矩阵)

定理6. $\vec{X}\sim N(\vec{\mu},\Sigma)$，$\vec{X}=(X_1,\cdots,X_n)$，则 $X_1,\cdots,X_n$ 相互独立 $\Leftrightarrow$ Σ 为对角阵

Proof. $f(x)=\frac{1}{\sqrt{(2\pi)^n|\Sigma|}}e^{-\frac{1}{2}(\vec{x}-\vec{\mu})\Sigma^{-1}(\vec{x}-\vec{\mu})^T}$

复高斯(正态)分布

2维随机向量，复平面，

复随机变量 $Z=X+iY$，X、Y 实随机变量

$\mathbb{E}[Z]:=\mathbb{E}[X]+i\mathbb{E}[Y]$

复高斯分布 $N_{\mathbb{C}}(\mu,\sigma^2)$，$\mu\in\mathbb{C}$，$\sigma>0$，指其密度 $f(z)=\frac{1}{\pi\sigma^2}e^{-\frac{1}{\sigma^2}|z-\mu|^2}$，$z\in\mathbb{C}$

结论：若 $Z\sim N_{\mathbb{C}}(0,1)$，则

$$\mathbb{E}[Z^k\bar{Z}^l]=\begin{cases}k! & k=l\\ 0 & k\ne l\end{cases}$$

思考题：$\phi_n(x)=\frac{1}{(2\pi)^n}e^{-\frac{1}{2}\sum_{j=1}^n x_j^2}$，

$\Delta_n(x)=\prod_{1\le i<j\le n}(x_j-x_i)$

证明：当 $m=1$ 时，有

$$\int_{\mathbb{R}^n}|\Delta_n(x)|^{2m}\phi_n(x)\,dx=\prod_{j=1}^{m}\frac{\Gamma(1+jm)}{\Gamma(1+m)}.$$

注意：① $\Delta_n(x)$ 有行列式表示

② 结论对 $\forall m\ge 0$ 均成立，$m=\frac{1}{2},1,2$ 非常重要

来自正态分布.

两个统计量：从总体中抽取样本 $X_1,\cdots,X_n$，

样本均值 $\bar{X}=\frac{1}{n}\sum_{k=1}^n X_k$

样本方差 $S^2=\frac{1}{n-1}\sum_{k=1}^n(X_k-\bar{X})^2$

思考题：$X_1,\cdots,X_n$ 独立同分布 $N(\mu,\sigma^2)$，则 $\mathbb{E}[\bar{X}]=\mu$，$\mathbb{E}[S^2]=\sigma^2$

(用来估计 μ 和 σ^2)

定义2. 当 X 有密度

$$f(x)=\frac{1}{2^{\frac{d}{2}}\Gamma(\frac{d}{2})}x^{\frac{d}{2}-1}e^{-\frac{1}{2}x},\quad x>0$$

则 X 服从 d 个自由度的卡方分布，记 $X\sim\chi^2(d)$

引理7. $Y_1,\cdots,Y_d$ 独立同 $N(0,1)$,
$X=\sum_{i=1}^{d}Y_i^2$, 则 $X\sim\chi^2(d)$

证明: $(Y_1,\cdots,Y_d)$ 密度 $f(y)=\frac{1}{(\sqrt{2\pi})^d}e^{-\frac{1}{2}\sum_{i=1}^{d}y_i^2}$

因此 $F_X(x)=P(\sum_{i=1}^{d}Y_i^2\le x)=\int_{\sum y_i^2\le x}f(y)\,d^dy$

极坐标分解之后有:

$$F_X(x)=C_d\int_0^{\sqrt{x}}r^{d-1}e^{-\frac{1}{2}r^2}dr$$
$$=\frac{C_d}{2}\int_0^{x}r^{\frac{d}{2}-1}e^{-\frac{1}{2}r}dr$$

用 $F_X(\infty)=1$ 知: $\frac{C_d}{2}=\frac{1}{2^{\frac{d}{2}}\Gamma(\frac{d}{2})}$

定理8. $X_1,\cdots,X_n$ 独立同 $N(\mu,\sigma^2)$, 则

(i) $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\sim N(\mu,\frac{\sigma^2}{n})$

(ii) $\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1)$, $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$

(iii) $\overline{X}$ 与 S^2 独立.

Proof. 令 $Y_i=\frac{1}{\sigma}(X_i-\mu)$, 则 $Y_i\sim N(0,1)$, 且 $Y_1,\cdots,Y_n$ 相互独立, 显见

$$\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i=\frac{1}{\sigma}(\overline{X}-\mu)$$
$$\sum_{i=1}^{n}(Y_i-\overline{Y})^2=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{n-1}{\sigma^2}S^2$$

$(Y_1,\cdots,Y_n)$ 密度 $f(y)=\frac{1}{(\sqrt{2\pi})^n}e^{-\frac{1}{2}\sum_{i=1}^{n}y_i^2}$

取正交阵
$$A=\begin{pmatrix}\frac{1}{\sqrt{n}} & \\ \vdots & * \\ \frac{1}{\sqrt{n}} & \end{pmatrix}$$
令 $(Z_1,\cdots,Z_n)=(Y_1,\cdots,Y_n)A$,

则 $Z_1=\sqrt{n}\,\overline{Y}$, 且 $Z_1,\cdots,Z_n$ 独立同 $N(0,1)$

(因为 $Y\sim N(0,I_n)$ 知 $Z\sim N(0,\underbrace{A^TI_nA}_{I_n})$)

又可知 $Z_1^2+\cdots+Z_n^2=Y_1^2+\cdots+Y_n^2$,

有 $\sum_{i=2}^{n}Z_i^2=\sum_{i=1}^{n}Y_i^2-n\overline{Y}^2=\sum_{i=1}^{n}(Y_i-\overline{Y}+\overline{Y})^2-n\overline{Y}^2$
$=\sum_{i=1}^{n}(Y_i-\overline{Y})^2=\frac{n-1}{\sigma^2}S^2$

立即 $\overline{Y}\sim N(0,\frac{1}{n})$　$\frac{n-1}{\sigma^2}S^2\sim\chi^2(n-1)$

且 $\overline{X}$ 与 S^2 独立

第五章 特征函数·极限定理

§5.1. 再谈期望

1.记号

随机变量X、分布函数F.

$$E[X]=\begin{cases}\sum_x x f(x) & \text{离散型}\\ \int x f(x)\,dx & \text{连续型}\end{cases}$$

引入 $$dF(x)=\begin{cases}F(x)-F(x-0)=f(x)\\ \frac{dF}{dx}dx=f(x)dx\end{cases}$$

统一记为 $E[X]=\int x\,dF(x)$

性质 $E[g(x)]=\int g(x)\,dF(x)$

2.抽象积分

$(\Omega,\mathcal{F},P)$, X, F, 如何定义一般随机变量X期望$E[X]$? 三步走战略

Step1. (非负)简单随机变量

$X=\sum_{i=1}^{n} x_i I_{A_i}$, $A_i=\{X=x_i\}$, $\bigcup_{i=1}^{n} A_i=\Omega$

$E[X]=\sum_{i=1}^{n} x_i P(A_i)$

Step2. 非负随机变量:

存在单调增简单随机变量列 $X_n(w)\uparrow X(w)$

这里 $A_n=\{X\geq n\}$ ~~$A_{nj}=\{w\in\Omega:$~~

$A_{n,j}=\{w\in\Omega:\frac{j-1}{2^n}\leq X(w)<\frac{j}{2^n}\}$, $j=1,\cdots,n2^n$

$X_n=nI_{A_n}+\sum_{j=1}^{n\cdot 2^n}\frac{j-1}{2^n}I_{A_{nj}}$

定义 $E[X]:=\lim_{n\to\infty}E[X_n]$

要说明当 $X_n\uparrow X$, $Y_n\uparrow X$ 时有

$\lim_{n\to\infty}E[X_n]=\lim_{n\to\infty}E[Y_n]$?

见周民强《实变函数》

Step3. 一般随机变量:

$X=X^+-X^-$

这里 $X^+=\max\{X,0\}$ $X^-=\max\{-X,0\}$

当$E[X^+]<\infty$或$E[X^-]<\infty$时,

$E[X]=E[X^+]-E[X^-]$

特别,当$E[|X|]<\infty$时,称X的期望存在

统一记号:

$E[X]=\int_\Omega X\,dP$ 或 $E[X]=\int_\Omega X(w)\,P(dw)$

3.重要性质: 期望算子

(i) 非负性: 当 $X\geq 0$ 时, $E[X]\geq 0$

(ii) 线性性: $E[aX+bY]=aE[X]+bE[Y]$

(iii) 归一性: $E[C]=C$, $C\in\mathbb{R}$

E的连续性: $X_n(w)\to X(w)$, $\forall w\in\Omega$

($\exists\Omega_0$, $P(\Omega_0)=0$, $\forall w\in\Omega\backslash\Omega_0$)

(i) 单调收敛. $X_{n+1}(w)\geq X_n(w)\geq 0$, $\forall n,w$,

则 $E[X_n]\to E[X]$, $n\to\infty$

(ii) 控制收敛, $|X_n|\leq Y$, $\forall n$, 当$E[Y]<\infty$时,

有$E[X_n]\to E[X]$, $n\to\infty$

(iii) 有界收敛. $|X_n|\leq C$, $\forall n$,

则$E[X_n]\to E[X]$, $n\to\infty$

Fatou引理: $X_n\geq 0$, a.s (almost surely)

则 $E[\liminf_{n\to\infty}X_n]\leq\liminf_{n\to\infty}E[X_n]$

4. Lebesgue-Stieltjes 积分

X, F, 引入 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上测度 μ_F

$\mu_F((a,b]) = F(b) - F(a)$

有 $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu_F)$ 其上 $g: \mathbb{R} \to \mathbb{R}$ Borel可测

抽象积分: $\int g \, d\mu_F$, 有时也记 $\int g \, dF$

称为 Lebesgue-Stieltjes 积分

定理: $\mathbb{E}[g(X)] = \int_{\mathbb{R}} g \, dF$

CHECK分三步走: 示性→简单→非负→一般

(多元场合) $\mathbb{E}[g(X_1,\cdots,X_n)]$

$= \int_{\mathbb{R}^n} g(x_1,\cdots,x_n) \, dF(x_1,\cdots,x_n)$.

5. 独立随机变量之积期望.

定理: X与Y独立, $\mathbb{E}[|X|]<\infty$, $\mathbb{E}[|Y|]<\infty$,

则 $\mathbb{E}[|XY|]<\infty$, 且 $\mathbb{E}[XY]=\mathbb{E}[X]\mathbb{E}[Y]$.

证明:

Step1. 简单

$X=\sum_{i=1}^n x_i I_{A_i}$ $Y=\sum_{j=1}^n y_j I_{B_j}$, 这里 A_i B_j 独立

$XY = \sum_{i,j} x_i y_j I_{A_iB_j} \Rightarrow \mathbb{E}[XY] = \sum_{i,j} x_i y_j P(A_iB_j)$

$= \mathbb{E}[X]\mathbb{E}[Y]$

Step2. 非负.

取简单随机变量列 $X_n \uparrow X$, $Y_n \uparrow Y$, 且X_n与Y_n独立(可以做到),

$\mathbb{E}[XY] = \lim_{n\to\infty} \mathbb{E}[X_nY_n] \overset{\text{Step1}}{=\!=} \lim_{n\to\infty} \mathbb{E}[X_n]\mathbb{E}[Y_n]$

$= \mathbb{E}[X]\mathbb{E}[Y]$

Step3. 一般 $X=X^+-X^-$ $Y=X^+-X^-$

$\mathbb{E}[XY] = \mathbb{E}[X^+Y^+ + X^-Y^-] - \mathbb{E}[X^+Y^- + X^-Y^+]$

$= \mathbb{E}[X]\mathbb{E}[Y]$

6. 同分布随机变量刻划

X与Y同分布指 $\forall x \in \mathbb{R}$, 有 $P(X\le x)=P(Y\le x)$

定理: X与Y同分布 $\Longleftrightarrow$ $\forall g \in C_b(\mathbb{R})$: 有界连续函数, $\mathbb{E}[g(X)]=\mathbb{E}[g(Y)]$

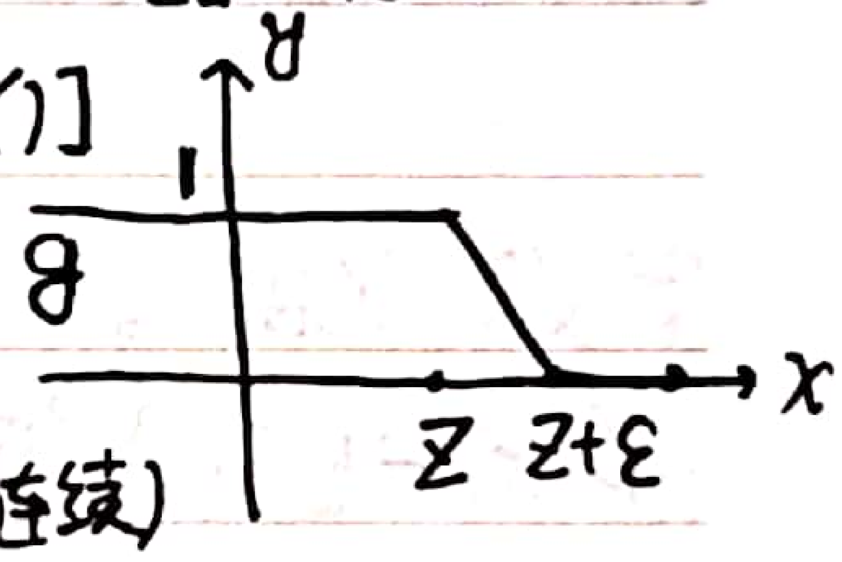
CHECK $\Rightarrow$ √

$\Leftarrow$ $F_X(z)=P(X\le z)=\mathbb{E}[I_{(-\infty,z]}(X)]$

(取g) $\le \mathbb{E}[g(X)]=\mathbb{E}[g(Y)]$

$\le \mathbb{E}[I_{(-\infty,z+\varepsilon]}(Y)]$

$= F_Y(z+\varepsilon)$.



(右连续)

有 $F_X(z) \le F_Y(z+\varepsilon)$, 令 $\varepsilon\to 0$, 有 $F_X(z)\le F_Y(z)$

同样 $F_Y(z)\le F_X(z)$, 即X与Y同分布

§5.2. 特征函数

非负整值X 母函数 $G(s)=\mathbb{E}[s^X]$.

一般随机变量 $s=e^t$, $t\in\mathbb{R}$, 矩母函数

$M_X(t)=\mathbb{E}[e^{tX}]$

若$\exists r>0$, s.t. 当$|t|<r$时, $M_X(t)$总存在,

但对Cauchy分布

$M(t)=\int_{-\infty}^{\infty} e^{tx}\frac{1}{\pi}\cdot\frac{1}{1+x^2}dx$, $t\ne 0$时不存在

定义1. X, Y为$(\Omega,\mathcal{F},P)$上随机变量, 复随机变量 $Z=X+iY$

注: (i) 复随机变量即2维随机向量

(ii) 当 (X_1,Y_1) 与 (X_2,Y_2) 独立即 $P(X_1\le x_1, Y_1\le y_1, X_2\le x_2, Y_2\le y_2) = P(X_1\le x_1, Y_1\le y_1)P(X_2\le x_2, Y_2\le y_2)$ 时，称 $Z_1=X_1+iY_1$ 与 $Z_2=X_2+iY_2$ 独立

此时可知 X_1 与 X_2 独立，Y_1 与 Y_2 独立

(iii) $E[Z]:=E[X]+iE[Y]$

(iv) $Z_1,\cdots,Z_n$ 相互独立时，

$E[Z_1\cdots Z_n]=E[Z_1]\cdots E[Z_n]$

$Z_1Z_2=(X_1X_2-Y_1Y_2)+i(X_1Y_2+X_2Y_1)$

定义2. X的特征函数:

$\phi(t)=E[e^{itX}]$，$-\infty<t<+\infty$

有时 $\phi_X(t)$ 表示

注: (i) $\phi(t)=E[\cos(tX)]+iE[\sin(tX)]$

(ii) $|e^{itx}|=1$，$\forall t\in\mathbb{R}$，表明 $\phi(t)$ 总存在

(iii) $\phi(t)=\int_{\mathbb{R}}e^{itx}dF(x)$

当F有密度时，$\phi(t)=\int_{\mathbb{R}}e^{itx}f(x)dx$

定理1. $\phi(t)=E[e^{itX}]=\int e^{itx}dF$

(i) $\phi(0)=1$、$|\phi(t)|\le 1$，$\overline{\phi(t)}=\phi(-t)$

(ii) ϕ 在 $(-\infty,\infty)$ 一致连续

(iii) ϕ 非负定，$\forall t_1,\cdots,t_n\in\mathbb{R}$，$Z_1\cdots Z_n\in\mathbb{C}$，

有 $\sum_{j,k=1}^{n}\phi(t_j-t_k)Z_j\overline{Z_k}\ge 0$.

证明: (i) $|\phi(t)|\le\int|e^{itx}|dF=1$

(ii) $|\phi(t+h)-\phi(t)|=\int|e^{itx}(e^{ihx}-1)|dF$

$\le\int\underbrace{|e^{ihx}-1|}_{有界}dF\xrightarrow[有界收敛定理]{h\to 0}0$

(iii) $\sum_{j,k}\phi(t_j-t_k)Z_j\overline{Z_k}=$ $\sum_{j,k}E[Z_je^{it_jX}\overline{Z_k}e^{-it_kX}]$

$=E[|\sum_j Z_je^{it_jX}|^2]\ge 0$

注: 定理1完全刻划了 $\phi:\mathbb{R}\to\mathbb{C}$ 是否为某随机变量的特征函数 (Bochner 1899-1982)

定理2. 若 $E[|X|^k]<\infty$，则 $\forall j\le k$ 有

$\phi^{(j)}(0)=i^jE[X^j]$，

进而 $\phi(t)=\sum_{j=0}^{k}\frac{(it)^j}{j!}E[X^j]+o(t^k)$，$t\to 0$

证明: $\phi(t)=\int e^{itx}dF(x)=E[e^{itX}]$

由题5.6.4 (见周民强《实变函数》§6.1 例)

知 $E[|X|^j]<\infty$，再由 $\frac{d^j}{dt^j}e^{itx}=|x|^j$，

$E[|X|^j]<\infty$ 可知积分与求导可交换

定理3.

(i) $Y=aX+b$，则 $\phi_Y(t)=e^{ibt}\phi_X(at)$

(ii) 当X与Y独立时，

$\phi_{X+Y}(t)=\phi_X(t)\phi_Y(t)$

证明: (i) $\phi_Y(t)=E[e^{it(aX+b)}]$

(ii) $\phi_{X+Y}(t)=E[e^{it(X+Y)}]=E[e^{itX}e^{itY}]$

多元情形

定义3. $\vec{X}=(X_1,\cdots,X_n)$ 特征函数

$\phi_{\vec{X}}(\vec{t})=E[e^{i\sum_j t_jX_j}]$

定理4. X与Y独立 $\Leftrightarrow\phi_{X,Y}(s,t)=\phi_X(s)\phi_Y(t)$

Proof. $\Rightarrow \surd$

$\Leftarrow$? 见反转公式.

例1. Bernoulli分布

$\phi(t)=\mathbb{E}[e^{itx}]=pe^{it}+q$

二项分布: $X\sim B(n,p)$, $\phi_X(t)=(pe^{it}+q)^n$

例2. 指数分布

$f(x)=\lambda e^{-\lambda x}$, $x>0$

$\phi(t)=\int_0^\infty e^{itx}\cdot\lambda e^{-\lambda x}dx=\lambda\int_0^\infty e^{-(\lambda-it)x}dx$

$=-\frac{\lambda}{\lambda-it}e^{-(\lambda-it)x}\Big|_0^\infty=\frac{\lambda}{\lambda-it}$

(复分析: $Re(\lambda)>0$, 另外分别算实部虚部)

例3. 标准正态分布 $N(0,1)$

$\phi(t)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{itx-\frac{1}{2}x^2}dx$

今 $S=it$, 当 $S\in\mathbb{R}$时,

$\phi(t)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{\frac{1}{2}s^2-\frac{1}{2}(x-s)^2}dx=e^{\frac{1}{2}s^2}$

$=e^{-\frac{1}{2}t^2}$ "Physics Proof"

再解析延拓.

$\phi(t)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\cos(tx)e^{-\frac{1}{2}x^2}dx$

$\phi'(t)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\sin(tx)(-xe^{-\frac{1}{2}x^2})dx$

$=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\sin(tx)\,de^{-\frac{1}{2}x^2}=-t\phi(t)$,

$\phi(0)=1$, 解之 $\phi(t)=e^{-\frac{1}{2}t^2}$

"Math Proof"

当$X\sim N(\mu,\sigma^2)$时, $\phi_X(t)=e^{i\mu t-\frac{1}{2}\sigma^2t^2}$

例4. $\vec{X}\sim N(\vec{\mu},\Sigma)$, 则 $Y=\vec{X}\cdot\vec{t}^T$

$\phi(\vec{t})=\mathbb{E}[e^{i\vec{X}\cdot\vec{t}^T}]=\mathbb{E}[e^{isY}]\big|_{s=1}$

$Y\sim N(\vec{\mu}\cdot\vec{t}^T,\vec{t}\Sigma\vec{t}^T)$ 当$\vec{t}\neq 0$时

$\phi(\vec{t})=e^{i\vec{\mu}\cdot\vec{t}^Ts-\frac{1}{2}\vec{t}\Sigma\vec{t}^Ts^2}\big|_{s=1}$

$=e^{i\vec{\mu}\cdot\vec{t}^T-\frac{1}{2}\vec{t}\Sigma\vec{t}^T}$

注: 多元正态密度与联合特征函数 Σ 出现形式不同!

例5. $X\sim U[-1,1]$, 则

$\phi(t)=\int_{-1}^1\frac{1}{2}e^{itx}dx=\frac{\sin t}{t}$

会用微积分或复分析求$\phi(t)$, 会运用$\phi(t)$性质

§5.3 反转与连续性定理

X与Y同分布 $\Leftrightarrow$ $\phi_X(t)=\phi_Y(t)$

定理1. (反转公式)

X的分布函数为F, $\phi(t)=\int e^{itx}dF$,

则$\forall a<b$有:

$\frac{F(b)+F(b-0)}{2}-\frac{F(a)+F(a-0)}{2}$

$=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^T\frac{e^{-iat}-e^{-ibt}}{it}\phi(t)\,dt$

推论2. (唯一性定理)

$\phi_X(t)=\phi_Y(t)\Longrightarrow X$与$Y$同分布

Proof. 记C_F为F连续点全体，则$\mathbb{R}\setminus C_F$至多可数，对$b\in C_F$，让a在C_F中趋于$-\infty$可知$F(b)$被$\phi(t)$确定.

当$b\notin C_F$时，取$b_n\in C_F$，$b_n\downarrow b$，由F右连续性 $F(b)=\lim_{n\to\infty}F(b_n)$

定理1证明:

记 $I_T = \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-iat}-e^{-ibt}}{it}\phi(t)dt$

$=\frac{1}{2\pi}\int_{-T}^T \frac{e^{-iat}-e^{-ibt}}{it}\int_{-\infty}^\infty e^{itx}dF(x)\,dt$

$=\int_{-T}^T\int_{-\infty}^\infty \frac{e^{-iat}-e^{-ibt}}{2\pi it}e^{itx}dF(x)\,dt$

由 $|e^{i\alpha}-1| = \left|\int_0^\alpha e^{ix}dx\right|\le \alpha$

利用Fubini Thm,

$I_T = \frac{1}{2\pi}\int_{-\infty}^\infty\int_{-T}^T \frac{e^{it(x-a)}-e^{it(x-b)}}{it}dt\,dF(x)$

$=\int_{-\infty}^\infty \underbrace{\int_0^T\frac{\sin t(x-a)-\sin t(x-b)}{\pi t}dt}_{g_T(x)}\,dF(x)$

利用 $\frac{1}{\pi}\int_0^\infty\frac{\sin(tx)}{t}dt = \begin{cases}\frac12 & x>0\\ 0 & x=0\\ -\frac12 & x<0\end{cases}$

知 $\int_0^T \frac{\sin(tx)}{t}dt$有界，进而$g_T(x)$有界，且

$\lim_{T\to\infty} g_T(x)=\begin{cases}1, & x\in(a,b)\\ \frac12, & x=a,b\\ 0, & 否则\end{cases}$

控制收敛定理:

$\lim_{T\to\infty} I_T = \frac12\mu_F(\{a\})+\frac12\mu_F(\{b\})+\mu_F((a,b))$

$=\frac12(F(a)-F(a-0))+\frac12(F(b)-F(b-0))$

$+(F(b-0)-F(a))$

$=\frac12(F(b)+F(b-0))-\frac12(F(a)+F(a-0))$

定理3 (多元场合反转公式)

记 $\phi(t_1,\cdots,t_n)=E[e^{i\sum_{j=1}^n t_jX_j}]$，则

$P(a_j<X_j\le b_j,\ j=1,2,\cdots,n) = \lim_{T_1,\cdots,T_n\to\infty}\frac{1}{(2\pi)^n}\int_{-T_1}^{T_1}\cdots\int_{-T_n}^{T_n}\prod_{j=1}^n\frac{e^{-it_ja_j}-e^{-it_jb_j}}{it_j}\phi(t_1,\cdots,t_n)\,dt_1\cdots dt_n$

这里假设$(X_1,\cdots,X_n)$落入平行体$[a_1,b_1]\times\cdots\times[a_n,b_n]$表面内概率为0，特别当$\phi_{X,Y}(s,t)=\phi_X(s)\phi_Y(t)$时，X与Y独立

例1. 求特征函数cost对应的随机变量

解: 利用反转公式计算，或构造X

$P(X=1)=\frac12=P(X=-1)=\frac12$

$\phi_X(t)=\cos t$

X_n, F_n, ϕ_n收敛？收敛性之间关系？

例2. $X_n(\omega)=\frac1n$ $X(\omega)=0$ $\forall\omega\in\Omega$

明显，$X_n(\omega)\to X(\omega)$, $n\to\infty$, $\forall \omega\in\Omega$

但 $F_n(x)=\begin{cases}1, & x\ge\frac1n\\0,&x<\frac1n\end{cases}$ $F(x)=\begin{cases}1,&x\ge0\\0,&x<0\end{cases}$

当$x\ne0$时，$F_n(x)\to F(x)$，当$x=0$时，$F_n(0)=0$，不收敛到$F(0)=1$

定义1. F、F_n为分布收敛，若对$\forall x\in C_F$(连续点全体)有$\lim_{n\to\infty}F_n(x)=F(x)$，则称$F_n$**弱收敛**到F，记$F_n\xrightarrow{w}F$

(连续性定理: Lévy-Cramér)

定理4. F_n为分布函数, $\phi_n(t)=\int e^{itx}dF_n$

(i) 若F为分布函数, 且 $F_n \xrightarrow{W} F$, 则

$\phi_n(t)\to\phi(t)$, $\forall t\in\mathbb{R}$ 且内闭一致收敛,

这里 $\phi(t)=\int e^{itx}dF$

(ii) 若 $\phi(t)=\lim_{n\to\infty}\phi_n(t)$ 存在, 且 $\phi(t)$ 在 $t=0$ 处连续, 则 ϕ 为某分布函数F的特征函数, 且 $F_n \xrightarrow{W} F$

例3、$X_n\sim U[-n,n]$

则 $\phi_n(t)=\int_{-n}^{n}\frac{1}{2n}e^{itx}dx=\begin{cases}\frac{\sin(nt)}{nt}, & t\neq 0\\ 1, & t=0\end{cases}$

$\lim_{n\to\infty}\phi_n(t)=\begin{cases}0, & t\neq 0\\ 1, & t=0\end{cases} := h(t)$

$h(t)$在$t=0$处不连续, 非特征函数

§5.4 极限定理I

Probability Theory Is Measure Theory With A Soul —M.Kac.

随机性 Randomness 关键词

复杂

▷噪 噪声, 图像去噪, 信号去噪, 数据去噪

▷大 大数据

▷高 高维

▷多 多变量、多体问题(量子物理)

▷强 强关联、拓扑材料 体系

1.问题

大量随机现象, 各自的偶然此消彼长, 混沌中涌现秩序, 如误差。

给定$\{X_k\}$, 实数列$\{a_k\}$ $\{b_k\}$, 目标

$T_n=\frac{1}{B_n}\sum_{k=1}^{\infty}(X_k-a_k)$

当$n\to\infty$时, T_n是否收敛? 极限是什么?

下面因素必须考虑:

(i) $\{X_k\}$有什么性质?

(ii) $\{a_k\}$和$\{b_k\}$选取?

(iii) 哪种收敛, 求极限?

回答: (i) $\{X_k\}$ 相互独立 (且同分布)

(ii) $a_k=\mathbb{E}[X_k]$ $B_k=C\sqrt{k}$

(iii) 几种收敛引入, 求极限

2. 大数律(LLN)和中心极限定理(CLT)

定义1. 称X_n依分布收敛到X, 若 $F_{X_n}\xrightarrow{W}F_X$, 记 $X_n \xrightarrow[\text{distribution}]{D} X$

也称弱收敛

定理1. $\boxed{\text{(LLN)}}$ 设$\{X_n\}$独立同分布随机变量列, 且 $\mu=\mathbb{E}[X_1]$存在, 令

$S_n=\sum_{k=1}^{n}X_k$

则 $\frac{1}{n}S_n\xrightarrow{D}\mu$

Proof. 由连续性定理只须

$\phi_n(t) := E[e^{it\frac{S_n}{n}}] \to e^{i\mu t}$

再令$\phi(t) = E[e^{itX_1}]$，有$\phi_n(t) = (\phi(\frac{t}{n}))^n$

又$\phi(\frac{t}{n}) = 1 + \frac{i\mu t}{n} + o(\frac{t}{n})$可知：

$\phi_n(t) = (1 + \frac{i\mu t}{n} + o(\frac{t}{n}))^n \to e^{i\mu t}$, $n \to \infty$

LLN表明：$\frac{S_n}{n} - \mu \xrightarrow{D} 0$

寻找$\sigma_n > 0$, s.t. $\sigma_n \frac{S_n - n\mu}{n} \xrightarrow{?} ?$　　普适性：Universality 大道至简，大美天成

取$\sigma_n = \sqrt{n}$　因为$Var(S_n) = n Var(X_1)$

定理2. $\boxed{\text{(CLT)}}$ 设$\{X_k\}$独立同，且$\mu = E[X_1]$，

$\sigma^2 = Var(X_1)$, $\sigma > 0$，则：

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{D} Z,\quad Z \sim N(0,1)$$

解：记$Y_k = \frac{1}{\sigma}(X_k - \mu)$，不妨设$\mu = 0$, $\sigma = 1$

$\phi_{\frac{S_n}{\sqrt{n}}}(t) = E[e^{it\frac{S_n}{\sqrt{n}}}] = (E[e^{it\frac{X_1}{\sqrt{n}}}])^n$

$= (\phi(\frac{t}{\sqrt{n}}))^n = (1 - \frac{t^2}{2n} + o(\frac{t^2}{n}))^n$

$\phi(t) := E[e^{itX_1}] = 1 + 0 \cdot \frac{it}{1!} + 1 \cdot \frac{(it)^2}{2!} + o(t^2)$

可知：$\phi_{\frac{S_n}{\sqrt{n}}}(t) \longrightarrow e^{-\frac{1}{2}t^2}$，恰为$N(0,1)$特征函数

连续性定理即知.　(有时Z写为$N(0,1)$)

例1：测量遥远恒星与地球间距离(光年)，若各次测量值为独立同分布$X_1, \cdots, X_n$, $E[X_1] = \mu$, $Var(X_1) = 4$，若欲以95%把握保证测量精度达到±0.5光年，问至少测量多少次？

解：$Z_n = \frac{\sum_{i=1}^n X_i - n\mu}{2\sqrt{n}}$ 逼近正态分布$N(0,1)$

$P(|\frac{1}{n}\sum_{i=1}^n X_i - \mu| \le 0.5)$

$= P(-0.5 \times \frac{\sqrt{n}}{2} \le Z_n \le 0.5 \times \frac{\sqrt{n}}{2})$

$\approx \Phi(\frac{\sqrt{n}}{4}) - \Phi(-\frac{\sqrt{n}}{4}) = 2\Phi(\frac{\sqrt{n}}{4}) - 1 \ge 0.95$

即$\Phi(\frac{\sqrt{n}}{4}) \ge 0.975$　CHECK正态分布表

$\frac{\sqrt{n}}{4} \ge 1.96 \Rightarrow n \approx 61.47$，取$n = 62$

3. Lindeberg条件

可以处理独立但未必同分布，对$X_1, \cdots, X_n$

记$a_k = E[X_k]$, $b_k^2 = Var[X_k]$,

$B_n^2 = \sum_{i=1}^n b_k^2$，$F_k$为$X_k$分布函数

Lindeberg条件：$\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x-a_k| > \varepsilon B_n} (x - a_k)^2 dF_k = 0 \quad \text{(L)}$$

定理3. (Lindeberg-Feller CLT)

设$X_1, \cdots, X_n$相互独立，满足Lindeberg条件，

则$\frac{1}{B_n} \sum_{k=1}^n (X_k - a_k) \longrightarrow N(0,1)$　(LF)

且$\frac{1}{B_n^2} \max_{1\le k\le n} b_k^2 \longrightarrow 0$　(F)

特别地，当$\{X_k\}$相互独立，$E[X_k] = 0$,

$b_k^2 = Var(X_k)$, $E[|X_k|^3] < \infty$，且

$\frac{1}{B_n^3} \sum_{k=1}^n E[|X_k|^3] \longrightarrow 0$, $n \to \infty$

则$\frac{1}{B_n} \sum_{i=1}^n X_k \longrightarrow N(0,1)$

CHECK：满足Lindeberg条件

$(\sum_{i=1}^n \int_{|x| > \varepsilon B_n} x^2 dF_k) \cdot \frac{1}{B_n^2}$

$\le \frac{1}{B_n^2} \sum_{i=1}^n \int_{|x| > \varepsilon B_n} \frac{|x|}{\varepsilon B_n} x^2 dF_k$

$\le \frac{1}{\varepsilon B_n^3} \sum_{i=1}^n \int_{-\infty}^{\infty} |x|^3 dF_k \longrightarrow 0$, $\forall \varepsilon > 0$

注：(1) 条件(L)概率意义

$\frac{1}{B_n^2} \sum_{i=1}^n \int_{|x-a_k| > \varepsilon B_n} (x - a_k)^2 dF_k$

$\geq \varepsilon^2 \sum_{k=1}^{n} P\left(\frac{1}{B_n}|X_k - a_k| > \varepsilon\right)$

$\geq \varepsilon^2 P\left(\bigcup_{k=1}^{n}\left\{\frac{1}{B_n}|X_k - a_k| > \varepsilon\right\}\right)$

$= \varepsilon^2 P\left(\max_{1\leq k\leq n} \frac{|X_k - a_k|}{B_n} > \varepsilon\right)$

(L)条件表明：$\forall \varepsilon > 0$,

$P\left(\max_{1\leq k\leq n} \frac{|X_k - a_k|}{B_n} > \varepsilon\right) \to 0,\ n\to\infty$

即相对偏差 $\frac{|X_k - a_k|}{B_n}$ 一致小概率接近1

(ii) 当 $X_1,\cdots,X_n$ 相互独立时，条件(L)不仅仅是(LF)成立的充分条件，差不多也是必要条件，只需增加两个普通条件：

$\lim_{n\to\infty} B_n = \infty \Big| \lim_{n\to\infty}\frac{1}{B_n^2}\max_{1\leq k\leq n} b_k^2 = 0$ (Feller条件)

(iii) Lyapunov条件：$\exists \delta > 0$, s.t.

$\frac{1}{B_n^{2+\delta}}\sum_{k=1}^{n} E\left[|X_k - a_k|^{2+\delta}\right] \to 0,\ n\to\infty$

可知 Lyapunov条件 $\Rightarrow$ Lindeberg条件 $\Rightarrow$ CLT (Lyapunov CLT, 1901)

(iv) $\frac{b_k^2}{B_n^2} = \frac{1}{B_n^2}\int_{\mathbb{R}} (x-a_k)^2\, dF_k$

$= \frac{1}{B_n^2}\left(\int_{|x-a_k|\leq\varepsilon B_n} + \int_{|x-a_k|>\varepsilon B_n}\right)(x-a_k)^2\, dF_k$

$\leq \varepsilon^2 + \frac{1}{B_n^2}\int_{|x-a_k|>\varepsilon B_n}(x-a_k)^2\, dF_k$

$\leq \varepsilon^2 + \underbrace{\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\varepsilon B_n}(x-a_k)^2\, dF_k}_{\to 0\ \text{(L)条件}}$

微积分：$\max \frac{b_k^2}{B_n^2} \xrightarrow{n\to\infty} 0$

CLT简史

1733. De Moivre，$p=\frac{1}{2}$ Bernoulli分布

Laplace, $p\in(0,1)$ Bernoulli

1901. Lyapunov Lyapunov CLT

1922. Lindeberg Lindeberg条件

1935. Feller (L)条件的必要性

现在. 统计力学、随机矩阵、自由概率论…

§5.5 极限定理 II

4. 二项分布的正态逼近

$S_n = X_1 + \cdots + X_n \quad X_k \sim B(1,p)$ 独立同

$P(S_n = k) = C_n^k p^k q^{n-k},\ 0\leq k\leq n$,

(局部极限定理)

定理4. 设 $p\in(0,1)$，记 $x_k = \frac{k-np}{\sqrt{npq}}$，$0\leq k\leq n$

则对所有满足 $|x_k|\leq A$，一致地有：

$C_n^k p^k q^{n-k} \sim \frac{1}{\sqrt{2\pi npq}} e^{-\frac{1}{2}x_k^2},\quad n\to\infty$

证明：对所有 k, s.t. $|x_k|\leq A$，有

$k = np + \sqrt{npq}\, x_k,\quad n-k = nq - \sqrt{npq}\, x_k$

一致地 $k\sim np \quad n-k\sim nq$

Stirling公式：$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n},\ n\to\infty$

则 $C_n^k p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$

$\sim \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}\ p^k q^{n-k}}{\left(\frac{k}{e}\right)^k \sqrt{2\pi k}\left(\frac{n-k}{e}\right)^{n-k}\sqrt{2\pi(n-k)}}$

$= \sqrt{\frac{n}{2\pi k(n-k)}}\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}$

$\sim \sqrt{\frac{1}{2\pi npq}}\underbrace{\left(\frac{np}{k}\right)^k\left(\frac{nq}{n-k}\right)^{n-k}}_{\varphi(n,k)}$

$\log\varphi(n,k) = k\log\left(1-\frac{\sqrt{npq}}{k}x_k\right)$

$+(n-k)\log\left(1+\frac{\sqrt{npq}}{n-k}x_k\right)$

$= k\left(-\frac{\sqrt{npq}}{k}x_k - \frac{1}{2}\frac{npq}{k^2}x_k^2 + O\left(\left|\frac{\sqrt{npq}}{k}x_k\right|^3\right)\right)$

$+(n-k)\left(\frac{\sqrt{npq}}{n-k}x_k+\frac{npq}{(n-k)^2}x_k^2+O\left(\left|\frac{\sqrt{npq}}{n-k}x_k\right|^3\right)\right)$

$=-\frac{n^2pq}{2k(n-k)}x_k^2+O\left(\frac{(\sqrt{npq}|x_k|)^3}{k^2}+\frac{(\sqrt{npq}|x_k|)^3}{(n-k)^2}\right)$

$=-\frac{1}{2}x_k^2+O\left(\frac{1}{\sqrt{n}}\right)$

Fact: $|x|\leq\frac{2}{3}$, 有 $\log(1+x)=x-\frac{1}{2}x^2+\theta(x)$, $|\theta(x)|\leq|x|^3$

定理5 (De Moivre-Laplace 积分形式LCT)

设 $S_n\sim B(n,p)$, 则

$$P\left(a<\frac{S_n-np}{\sqrt{npq}}\leq b\right)\longrightarrow\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{1}{2}x^2}dx$$

证明: 令 $x_k=\frac{k-np}{\sqrt{npq}}$, 则

$$\text{LHS}=\sum_{k:x_k\in(a,b]}P(S_n=k)$$

$$=\sum_{k:x_k\in(a,b]}C_n^k p^k q^{n-k}\quad\left(\begin{array}{l}\text{为什么还}\\\text{要可积?}\end{array}\right)$$

$$\underset{\text{一致地}}{\sim}\sum_{k:x_k\in(a,b]}\frac{1}{\sqrt{2\pi npq}}e^{-\frac{1}{2}x_k^2}$$

$$=\sum_{k:x_k\in(a,b]}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x_k^2}(x_{k+1}-x_k)\quad\text{黎曼和}$$

$$\longrightarrow\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx,\quad n\to\infty$$

注意: (i) n 固定时, k 与 x_k 1-1对应

(ii) $x_0,x_1,\cdots,x_n$ 在 $\left[-\sqrt{\frac{np}{q}},\sqrt{\frac{nq}{p}}\right]$ 等间隔 $\frac{1}{\sqrt{npq}}$, 且 $n\to\infty$ 时, 该区间总包含 $(a,b]$

5. 矩方法.

$$\gamma_k=\int_{\mathbb{R}}x^k\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx$$

$\gamma_{2m-1}=0$, $\gamma_{2m}=(2m-1)!!$

组合诠释:

将 $1,2,\cdots,2m$ 配成 m 对, 共有 $(2m-1)!!$ 种

定理6. $\{X_k\}$ 相互独立, 且满足

(i) $\mathbb{E}[X_k]=0$, $Var(X_k)=1$, $\forall k$

(ii) (一致有界高阶矩) $\forall m\geq3$,

$$C_m=\sup_k\mathbb{E}[|X_k|^m]<\infty$$

则对 $S_n=\sum_{i=1}^n X_i$, 有

$$\mathbb{E}\left[\left(\frac{S_n}{\sqrt{n}}\right)^k\right]\longrightarrow\gamma_k,\quad n\to\infty$$

(进而 $\frac{S_n}{\sqrt{n}}\xrightarrow{D}N(0,1)$)

Proof:

$$\mathbb{E}\left[\left(\frac{S_n}{\sqrt{n}}\right)^k\right]=n^{-\frac{k}{2}}\sum_{1\leq i_1,\cdots,i_k\leq n}\mathbb{E}[X_{i_1}X_{i_2}\cdots X_{i_k}]\quad(*)$$

(1) $k=0,1$, 显然

(2) $k=2$,

$$(*)\text{RHS}=\frac{1}{n}\sum_{i_1}\mathbb{E}[X_{i_1}^2]+\frac{1}{n}\sum_{i_1\neq i_2}\mathbb{E}[X_{i_1}X_{i_2}]$$

$$=1$$

(3) $k=3$,

$$(*)\text{RHS}=n^{-\frac{3}{2}}\sum_{i_1}\mathbb{E}[X_{i_1}^3]+n^{-\frac{3}{2}}\sum_{i_1\neq i_2}3\mathbb{E}[X_{i_1}^2X_{i_2}]$$

$$+n^{-\frac{3}{2}}\sum_{i_1\neq i_2\neq i_3}\mathbb{E}[X_{i_1}X_{i_2}X_{i_3}]=O(n^{-\frac{1}{2}})\to0$$

(4) $k=4$.

$$(*)\text{RHS}=\frac{1}{n^2}\sum_{i_1}\mathbb{E}[X_{i_1}^4]+\frac{1}{n^2}\sum_{i_1\neq i_2}4\mathbb{E}[X_{i_1}^3X_{i_2}]$$

$$+\frac{1}{n^2}\sum_{i_1\neq i_2}3\mathbb{E}[X_{i_1}^2X_{i_2}^2]+\frac{1}{n^2}\sum_{i_1\neq i_2\neq i_3\neq i_4}\mathbb{E}[X_{i_1}X_{i_2}X_{i_3}X_{i_4}]$$

$$=O\left(\frac{1}{n}\right)+0+3+0\longrightarrow3$$

(5) 一般情形

RHS中非零项 $\mathbb{E}[X_{i_1}\cdots X_{i_k}]$ 必可写为形如 $\mathbb{E}[X_{i_1}^{a_1}\cdots X_{i_m}^{a_m}]$, $i_1\neq\cdots\neq i_m$, 且 $a_1,\cdots,a_m\geq2$, 因 $a_1+\cdots+a_m=k$, 则 $m\leq[\frac{k}{2}]$, 表明 k 为奇数时, ~~$\text{RHS}=n^{-\frac{k}{2}}$~~

$$\text{RHS}=n^{-\frac{k}{2}}O(n^m)\to0$$

当 k 为偶数时, 主项必在 $m=\frac{k}{2}$ 时取到

这时 $a_1=a_2=\cdots=a_m=2$,

从 $i_1, i_2, \cdots, i_{2m-1}, i_{2m}$ 两两配对

每一种给定的配对方式,对应极限

$$n^{-\frac{k}{2}}\sum_{\substack{i_1\cdots i_k\\ \text{一种配对}}} E[X_{i_1}\cdots X_{i_k}]=\frac{n(n-1)\cdots(n-m+1)}{n^m}\longrightarrow 1$$

即证.

详细分类:

$\{1,2,\cdots,k\}$ 划分成 m 组,每组至少 2 个数字 记 $\Gamma_k^{(m)}$ 为所有划分方式,每种方式形如 $V=\{V_1,V_2,\cdots,V_m\}$,这里 $V_i\cap V_j=\phi$ $(\forall i\neq j)$, $\bigcup_j V_j=\{1,2,\cdots,k\}$

$\#V_j\geq 2$, 可定义等价类:

$\vec{i}=(i_1,i_2,\cdots,i_k)\in V$, 若 $i_p=i_q \Leftrightarrow p,q\in V_j$

定理 6 中依分布收敛可由矩收敛定理证出

定理 7. (矩收敛定理)

假设 (i) $\forall k\in\mathbb{N}$, $\gamma_{k,n}=\int x^k dF_n$ 存在

(ii) $\forall k$, $\lim_{n\to\infty}\gamma_{k,n}=\gamma_k$

(iii) $\gamma_k=\int x^k dF$, 且满足 Carleman 条件:

$\sum_{k=1}^{\infty}(\gamma_{2k})^{-\frac{1}{2k}}=\infty$ (思考:正态分布)

则 $F_n\xrightarrow{w}F$

科普小品: 靳志辉, 正态分布的前世今生

定理 8. (Wick 公式)

$E(X^{2m})=\sum_{\sigma\in P(2m)}E(X^2)\cdots E(X^2)$

非负定

假设 $(X_1,\cdots,X_{2n})\sim N(0,\Sigma)$, $\Sigma\geq 0$, 则:

$E(X_1X_2\cdots X_{2n})=\sum_{\sigma\in P(2m)}\prod_{i,j\in\sigma}E(X_iX_j)$

证明: 令 $Y=\sum_{j=1}^{2n}\lambda_jX_j$, 则 $Y\sim N(0,\sigma^2)$,

$\sigma^2=\sum_{i,j}\lambda_i\lambda_j E(X_iX_j)$

令 $g(\lambda)=E(e^Y)$, 则 $LHS=\frac{\partial^{2n}}{\partial\lambda_1\cdots\partial\lambda_{2n}}g(\lambda)\big|_{\lambda=0}$

另外 $g(\lambda)=\sum_{k=0}^{\infty}\frac{\sigma^{2k}(2k-1)!!}{(2k)!}$

只须看 $k=n$ 时, $LHS=\frac{(2n-1)!!}{(2n)!}$

$\frac{\partial^{2n}}{\partial\lambda_1\cdots\partial\lambda_{2n}}\left(\sum_{i,j}\lambda_i\lambda_j E(\lambda_i\lambda_j)\right)^n=\sum_{\sigma\in P(2m)}\prod_{i,j\in\sigma}E(X_iX_j)$

$X\sim N(0,1)$, $E(X^{2n})=(2n-1)!!$

定理 9. 多元 CLT.

设 $\vec{X_1},\cdots,\vec{X_n}$ 独立同分布 α 维向量, $E(\vec{X_1})=0$

$\Sigma=E(\vec{X_1}^{T}\vec{X_1})$ 正定, 则 $\frac{1}{\sqrt{n}}\sum\vec{X_k}\longrightarrow N(0,\Sigma)$

第六章 几种收敛

§6.1 四种收敛 I

回顾分析学函数列$\{f_n\}$收敛性概念、不妨$f、f_n:[0,1]\to\mathbb{R}$

逐点收敛:

$f_n(x)\to f(x)$, $\forall x\in[0,1]$

L_p范数: $\|g\|_p:=(\int_0^1|g|^p dx)^{\frac{1}{p}}$

$\|f_n-f\|_p\to 0$, $n\to\infty$

依测度: $\forall\varepsilon>0$,

$m(\{x:|f_n(x)-f(x)|>\varepsilon\})\to 0$, $n\to\infty$

定义1. X, X_n为$(\Omega,\mathcal{F},P)$上随机变量

(i) 几乎处处收敛(也称以概率1收敛)

$P(\{\omega\in\Omega: X_n(\omega)-X(\omega)\to 0, n\to\infty\})=1$

记 $X_n\xrightarrow{a.s.}X$

(ii) r阶收敛 $\forall n$, $\mathbb{E}[|X_n|^r]<\infty$, 且

$\mathbb{E}[|X_n-X|^r]\to 0$, $n\to\infty$

记$X_n\xrightarrow{r}X$. $r=1$时平均收敛

$r=2$ 均方收敛

(iii) 依概率收敛 $\forall\varepsilon>0$,

$P(|X_n-X|>\varepsilon)\to 0$, $n\to\infty$

记$X_n\xrightarrow{P}X$

(iv) 依分布收敛. 对$F(x)=P(X\le x)$的连续点x处, $P(X_n\le x)\to P(X\le x)$, $n\to\infty$

记$X_n\xrightarrow{D}X$

注:依分布收敛可与样本空间的选取无关,与其他三种收敛有明显不同.

例1. 设$P(X=0)=P(X=1)=\frac{1}{2}$, $\forall n$, $X_n=X$,

令$Y=1-X$, 则$X_n\xrightarrow{D}X$, $X_n\xrightarrow{D}Y$

但$|X_n-Y|=|2X-1|=1$. 表明在前三种收敛下X_n不收敛到Y

定理1.

$X_n\xrightarrow{a.s.}X\Rightarrow X_n\xrightarrow{P}X\Longrightarrow X_n\xrightarrow{D}X$

$X_n\xrightarrow{r}X\nearrow$

$r>s\ge 1$时, $X_n\xrightarrow{r}X\Rightarrow X_n\xrightarrow{s}X$

引理2. $X_n\xrightarrow{P}X\Rightarrow X_n\xrightarrow{D}X$

Proof. 记$F_n(x)=P(X_n\le x)$, $F(x)=P(X\le x)$

$\forall\varepsilon>0$, $F_n(x)=P(X_n\le x, X\le x+\varepsilon)$

$+P(X_n\le x, X>x+\varepsilon)$

$\le F(x+\varepsilon)+P(|X-X_n|>\varepsilon)$ (*)

X_n与X交换: $x\to x-\varepsilon$

~~$F(x)\le F_n(x+\varepsilon)+P(|X-X_n|>\varepsilon)$~~

$F(x-\varepsilon)\le F_n(x)+P(|X-X_n|>\varepsilon)$ (**)

(*)与(**)合并:

$F(x-\varepsilon)-P(|X_n-X|<\varepsilon)\le F_n(x)$

$\le F(x+\varepsilon)+P(|X_n-X|>\varepsilon)$

让$n\to\infty$, 依概率收敛知:

$F(x-\varepsilon)\le\liminf_{n\to\infty}F_n(x)\le\limsup_{n\to\infty}F_n(x)\le F(x+\varepsilon)$

$\forall x\in C_F$, 有$\varepsilon\to 0$时, $F_n(x)\to F(x)$

引理3.

(i) $r>s\geq 1$ 时，$X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{s} X$

(ii) $X_n \xrightarrow{1} X \Rightarrow X_n \xrightarrow{P} X$

Proof. (i) 由问题4.14.28

$(E[|Z|^s])^{\frac{1}{s}} \leq (E[|Z|^r])^{\frac{1}{r}}$

取 $Z=X_n-X$，即知.

或4.14.27. $\boxed{\text{Hölder不等式}}$

$E[|XY|] \leq (E[|X|^p])^{\frac{1}{p}}(E[|Y|^q])^{\frac{1}{q}}$

可推知4.14.28

(ii) Markov不等式:

$P(|X|\geq a) \leq \frac{E[|X|]}{a}$，$a>0$

分析: $|X| = |X|I_{\{|X|\geq a\}} + |X|I_{\{|X|<a\}}$

取期望: $E[|X|] \geq E[|X|I_{\{|X|\geq a\}}]$

$\geq aE[I_{\{|X|\geq a\}}] = aP(|X|\geq a)$

$\forall \varepsilon>0$，$P(|X_n-X|>\varepsilon) \leq \frac{E[|X_n-X|]}{\varepsilon} \to 0$，

$n\to\infty$

Chebyshev不等式:

$P(|X-E[X]|\geq a) \leq \frac{Var(X)}{a^2}$

例2. (上反例) $\Omega=(0,1]$，取P为其上Lebesgue测度，定义 $X(\omega)=0$，$\forall \omega\in\Omega$，

$X_n(\omega) = \begin{cases} n^{\frac{1}{r}} & 0<\omega\leq\frac{1}{n} \\ 0 & \text{否则} \end{cases}$

明显 $X_n(\omega)\longrightarrow X(\omega)$，$n\to\infty$，$\forall\omega$

又 $\forall\varepsilon>0$，$P(|X_n-X|>\varepsilon)\leq\frac{1}{n}\to 0$

因此 $X_n \xrightarrow{P} X$，

但 $E[|X_n-X|^r] = \frac{1}{n}\cdot(n^{\frac{1}{r}})^r = 1$

表明 X_n 非 r 阶收敛到 X

§6.2 四种收敛II

几乎处处收敛刻划

$\{\omega\in\Omega: \lim_{n\to\infty} X_n(\omega)=X(\omega)\} = \bigcap_{k=1}^{\infty}\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{\omega\in\Omega: |X_n(\omega)-X(\omega)|\leq\frac{1}{k}\}$

分析: 若 $\omega\in$ LHS，则 $\forall k\ \exists m$ s.t. $n>m$ 时，总有 $\omega\in\bigcap_{n=m}^{\infty}\{\omega\in\Omega: |X_n(\omega)-X(\omega)|\leq\frac{1}{k}\}$，进而 $\omega\in$ RHS.

反之，若 $\omega\in$ RHS，$\forall\varepsilon$ s.t. $k>\frac{1}{\varepsilon}$，则总存在m s.t. $n\geq m$ 时，$|X_n(\omega)-X(\omega)|\leq\frac{1}{k}<\varepsilon$

即 $\omega\in$ LHS

以概率1收敛等价刻划

$$P\left(\bigcap_{k=1}^{\infty}\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{|X_n-X|\leq\tfrac{1}{k}\}\right)=1 \quad (*)$$

$$P\left(\bigcup_{k=1}^{\infty}\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n-X|>\tfrac{1}{k}\}\right)=0 \quad (**)$$

引理4.

(i) $X_n \xrightarrow{a.s.} X \Leftrightarrow \forall\varepsilon>0$，

$P(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n-X|>\varepsilon\})=0 \Leftrightarrow \forall\varepsilon>0$，

$\lim_{m\to\infty} P(\bigcup_{n=m}^{\infty}\{|X_n-X|>\varepsilon\})=0$

(ii) $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X$

Proof. (ii) 可由(i)立即知，只看(i)第1个"$\Leftrightarrow$"

$\Rightarrow \bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n-X|>\varepsilon\} \subset \bigcup_{k=1}^{\infty}\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n-X|>\frac{1}{k}\}$

由(**)可知.

$\Leftarrow P(\bigcup_{k=1}^{\infty}\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n-X|>\frac{1}{k}\})$

$\leq \sum_{k=1}^{\infty} P(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n-X|>\frac{1}{k}\})=0$

例3(上反例). 令 $X_n=\begin{cases}1 & 以概率\frac{1}{n}\\ 0 & 以概率1-\frac{1}{n}\end{cases}$

假设 $\{X_n\}$ 相互独立，则 $\varepsilon\in(0,1)$

$P(|X_n-0|>\varepsilon)=\frac{1}{n}\longrightarrow 0$

另一方面，$P(\bigcup_{n=m}^{\infty}\{|X_n-0|>\varepsilon\})$

$=\lim_{r\to\infty}P(\bigcup_{n=m}^{r}\{|X_n-0|>\varepsilon\})$

$=1-\lim_{r\to\infty}P(\bigcap_{n=m}^{r}\{|X_n-0|\leq\varepsilon\})$

$=1-\lim_{r\to\infty}\prod_{n=m}^{r}(1-\frac{1}{n})=1-\lim_{r\to\infty}\frac{m-1}{r}=1$

表明非几乎处处收敛到0.

本例也说明：r 阶收敛推不出 a.s.

$E[|X_n-0|^r]=\frac{1}{n}\longrightarrow 0$. 即 $X_n\xrightarrow{r}0$,

但非几乎处处收敛

定理5. (加上合适条件后，定理1结论反命题√)

(i) 若 $C\in\mathbb{R}$，且 $X_n\xrightarrow{D}C$，则 $X_n\xrightarrow{P}C$

(ii) 若 $\exists K$, s.t. $P(|X_n|\leq K)=1$, $\forall n$, 且 $X_n\xrightarrow{P}X$，则 $X_n\xrightarrow{r}X$

(iii) 若 $\forall\varepsilon>0$ 有 $\sum_{n=1}^{\infty}P(|X_n-X|>\varepsilon)<\infty$，则 $X_n\xrightarrow{a.s.}X$

Proof. (i) $P(|X_n-C|>\varepsilon)$

$=P(X_n>C+\varepsilon)+P(X_n<C-\varepsilon)$

$=1-\underbrace{P(X_n\leq C+\varepsilon)}_{\downarrow\ 1}+\underbrace{P(X_n<C-\varepsilon)}_{\downarrow\ 0}\longrightarrow 0$

(ii) Claim: $P(|X|\leq K)=1$

因为 $\{|X|\leq K+\varepsilon\}\supset\{|X-X_n|\leq\varepsilon\}\cap\{|X_n|\leq K\}$，知：

$P(|X|\leq K+\varepsilon)\geq P(|X-X_n|\leq\varepsilon, |X_n|\leq K)\to 1$

即 $P(|X|\leq K+\varepsilon)=1$，再由分布函数右连续，让 $\varepsilon\to 0$ 即知 $P(|X|\leq K)=1$

$|X_n-X|^r=|X_n-X|^r I_{\{|X_n-X|\leq\varepsilon\}}+|X_n-X|^r I_{\{|X_n-X|>\varepsilon\}}$

$\leq\varepsilon^r I_{\{|X_n-X|\leq\varepsilon\}}+(2K)^r I_{\{|X_n-X|>\varepsilon\}}$ a.s

取期望：$E[|X_n-X|^r]\leq\varepsilon^r+(2K)^r\underbrace{P(\{|X_n-X|>\varepsilon\})}_{\downarrow n\to\infty\ 0}$，易知

(iii) $P(\bigcup_{n=m}^{\infty}\{|X_n-X|>\varepsilon\})$

$\leq\sum_{n=m}^{\infty}P(\{|X_n-X|>\varepsilon\})\xrightarrow{m\to\infty}0$

推论6.(弱大数律)

设 $X_1,\cdots,X_n$ 独立同，且 $\mu=E[X_i]$ 存在，则 $\frac{1}{n}\sum_{k=1}^{n}X_k\xrightarrow{P}\mu$

定理7. (Skorokhod 表示定理)

设 $X_n\xrightarrow{D}X$，则存在 $(\Omega,\mathcal{F},P)$ 其上 Y_n, Y 满足 (i) Y_n 与 X_n 同分布，Y 与 X 同分布

(ii) $Y_n\xrightarrow{a.s.}Y$

Proof.

Fact. F 为分布函数，定义"反函数"

$F^{-1}(y):=\sup\{x:F(x)<y\}$, $y\in(0,1)$

(或 $F^{-1}(y)>x \Longleftrightarrow y>F(x)$)

则 $F^{-1}(y)\le x \Longleftrightarrow y\le F(x)$,

且当U为(0,1)上均匀分布时, $F^{-1}(U)$的分布函数为F

CHECK: "$\Leftarrow$" 由 $y>F(x)$ 及右连续性可知: $\exists\,\delta>0$, s.t. $y>F(x+\delta)$,

再由 $F^{-1}(y)$ 定义知: $x+\delta\le F^{-1}(y)$, 从而 $x<F^{-1}(y)$

"$\Rightarrow$" 由定义知: $\exists\, x_*\in\{x: F(x)<y\}$, s.t. $x<x_*$, 从而 $F(x)\le F(x_*)<y$

令 $Y_n(u)=F_n^{-1}(u)$, $Y(u)=F^{-1}(u)$, $u\in(0,1)$,

则 $u\le F_n(x) \Longleftrightarrow Y_n(u)\le x$ (*)

$u\le F(x) \Longleftrightarrow Y(u)\le x$ (**)

进而 X_n 与 Y_n, X与Y同分布

Step1. $\forall\,\varepsilon>0$, $u\in(0,1)$, $x\in C_F$ s.t. $Y(u)-\varepsilon<x<Y(u)$

(**)表明 $F(x)<u$, 结合 $F_n(x)\longrightarrow F(x)$ 知充分大n, $F_n(x)<u$, 再由(*) $Y_n(u)>x$

从而 $Y(u)-\varepsilon<x<Y_n(u)$

于是 $\liminf\limits_{n\to\infty} Y_n(u)\ge Y(u)$

Step2: $u<u'<1$, $x\in C_F$, s.t. Y在u连续 $Y(u')<x<Y(u')+\varepsilon$

由(**)知: $u<u'\le F(x)\longleftarrow F_n(x)$

当n充分大时 $F_n(x)>u$, 结合(*),

$Y_n(u)\le x$, 进而 $Y_n(u)\le x<Y(u')+\varepsilon$

取上极限 $\limsup\limits_{n\to\infty} Y_n(u)\le Y(u')$

因Y单调增, 可取 $u'\in C_Y$ 且 $u'\downarrow u$, 于是

$\limsup\limits_{n\to\infty} Y_n(u)\le Y(u)$

从而 $\lim\limits_{n\to\infty} Y_n(u)=Y(u)$

定理8. $X_n\xrightarrow{D}X$ 有如下等价刻划 弱收敛名称来源

(i) $\forall\, g\in C_b(\mathbb{R})$, (有界连续函数), 有 $\mathbb{E}[g(X_n)]\longrightarrow\mathbb{E}[g(X)]$, $n\to\infty$

(ii) $\forall$有界且一致连续函数g, 有(i)

(iii) ~~$\forall g\in C_b(\mathbb{R})$~~ 给定m, $g, g', \dots, g^{(m)}\in C_b(\mathbb{R})$, 有(i)

(iv) $\phi_n(t)\to\phi(t)$, $\forall t$, $\phi(t)$在$t=0$处连续

Proof. ~~只看(i)~~ (ii)(iii) 因为其他部分

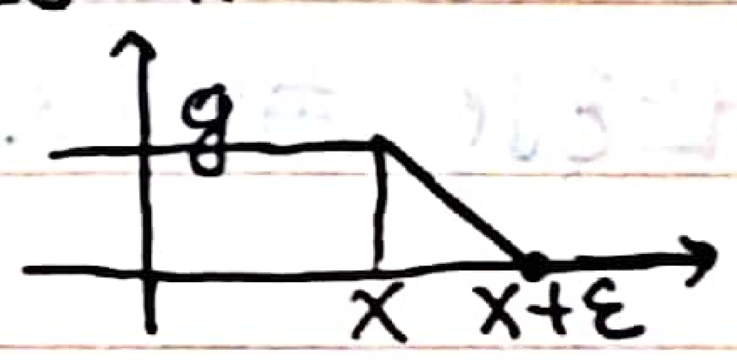
$\phi_n(t)=\mathbb{E}[\cos(tX_n)]+i\mathbb{E}[\sin(tX_n)]$

$\cos(tx)$和$\sin(tx)$关于x一致连续、有界, 为"好函数"

只看(i) "$\Rightarrow$" 利用表示定理, $\exists$ Y与X, Y_n与X_n同分布, 且 $Y_n\xrightarrow{a.s.}Y$, 又 $g\in C_b(\mathbb{R})$ 知 $g(Y_n)\xrightarrow{a.s.}g(Y)$, 有界控制收敛表明

$$\underset{\displaystyle\mathbb{E}[g(X_n)]}{\underset{\shortparallel}{\mathbb{E}[g(Y_n)]}}\longrightarrow\underset{\displaystyle\mathbb{E}[g(X)]}{\underset{\shortparallel}{\mathbb{E}[g(Y)]}}$$

"$\Leftarrow$" 构造 $g_{x,\varepsilon}(y)=\begin{cases}1 & y\le x\\ \frac{1}{\varepsilon}(x-y)+1 & x<y\le x+\varepsilon\\ 0 & y>x+\varepsilon\end{cases}$



$P(X_n\le x)=\mathbb{E}[I_{(-\infty,x]}(X_n)]\le\mathbb{E}[g_{x,\varepsilon}(X_n)]$

则 $\limsup\limits_{n\to\infty} P(X_n\le x)\le\mathbb{E}[g_{x,\varepsilon}(X)]$

$\le P(X\le x+\varepsilon)$

类似, $P(X_n\le x)\ge E[g_{x-\varepsilon,\varepsilon}(X_n)]$

$\liminf_{n\to\infty} P(X_n\le x)\ge E[g_{x-\varepsilon,\varepsilon}(X)]\ge P(X\le x-\varepsilon)$

因此当 $x\in C_F$ 时, 让 $\varepsilon\to 0$, 即可

§6.3 结论拾零

1. 几个不等式

Hölder 不等式: $p,q>1$, $\frac{1}{p}+\frac{1}{q}=1$,

$E[|XY|]\le (E[|X|^p])^{\frac{1}{p}}(E[|X|^q])^{\frac{1}{q}}$

Minkowski 不等式:

$(E[|X+Y|^p])^{\frac{1}{p}}\le (E[|X|^p])^{\frac{1}{p}}+(E[|X|^p])^{\frac{1}{p}}$

Markov 不等式: (尾概率)

$P(|X|\ge a)\le \frac{E[|X|^r]}{a^r}$, $a>0$

Chebyshev 不等式:

$P(|X-E[X]|\ge a)\le \frac{1}{a^2}Var(X)$

例1. 若存在 $r>0$, s.t. $E[|X|^r]=0$, 则 $P(X=0)=1$

分析: $\forall \varepsilon>0$, $P(|X|\ge\varepsilon)\le \frac{E[|X|^r]}{\varepsilon^r}=0$

表明: $P(|X|<\varepsilon)=1$, 进而 $P(|X|\le 2\varepsilon)=1$

让 $\varepsilon\downarrow 0$, 由分布函数右连续知 $P(|X|=0)=1$

2. 随机变量和的收敛

定理1. $\square=a.s, r, p$, 则

(i) 若 $X_n\xrightarrow{\square}X$, $X_n\xrightarrow{\square}Y$, 则 $P(X=Y)=1$

(ii) 若 $X_n\xrightarrow{\square}X$, $Y_n\xrightarrow{\square}Y$, 则 $X_n+Y_n\xrightarrow{\square}X+Y$

(iii) 对依分布收敛, (i) (ii) 中结论一般不对

证明 (i) 只看 r 阶收敛, $\|X\|_r=(E[|X|^r])^{\frac{1}{r}}$

$\|X-Y\|_r\le\|X-X_n\|_r+\|X_n-Y\|_r\to 0, n\to\infty$

表明 $E[|X-Y|^r]=0$, 再例1知 $P(X=Y)=1$

(ii) 只看 p 收敛,

$P(|X_n+Y_n-(X+Y)|>\varepsilon)$

$=P(|(X_n-X)+(Y_n-Y)|>\varepsilon)$

$\le P(\{|X_n-X|>\frac{\varepsilon}{2}\}\cup\{|Y_n-Y|>\frac{\varepsilon}{2}\})$

$\le P(|X_n-X|>\frac{\varepsilon}{2})+P(|Y_n-Y|>\frac{\varepsilon}{2})\to 0$

(iii) 取 $P(X=1)=P(X=-1)=\frac{1}{2}$, 设 X_n 与 X 同分布, $\forall n$, 再取 $Y=-X$, 则 Y 与 X 同分布, 且 $X_n\xrightarrow{D}X$, $X_n\xrightarrow{D}Y$, 但 X_n+X_n 非依分布收敛到 $X+Y$, $P(X=Y)\ne 1$

注: 当 $X_n\xrightarrow{D}X$, $X_n\xrightarrow{D}Y$, 则 X 与 Y 同分布, 因为 $\phi_n(t)=E[e^{itX_n}]\to\phi_X(t)$, 且 $\phi_n(t)\Rightarrow\phi_Y(t)$, 由反转公式知.

3. Borel-Cantelli 引理

刻划几乎处处收敛的一个充分条件: $\forall\varepsilon>0$, $\sum_{n=1}^{\infty}P(|X_n-X|<\varepsilon)<\infty$

先讨论事件列运算,

$(\Omega, \mathcal{F}, P)$, $A_m\in\mathcal{F}$, $m=1,2,\cdots$

$\bigcup_{m=n}^{\infty}A_m$ 表示 $A_n, A_{n+1},\cdots$ 至少有1个发生

$\bigcap_{m=n}^{\infty}A_m$ 表示 $A_n, A_{n+1},\cdots$ 同时发生

$\{A_n\}$ 的$\boxed{\text{上限事件}}$:

$\limsup_{n\to\infty}A_n=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m$

$\{A_n\}$的$\boxed{\text{下限事件}}$

$$\liminf_{n\to\infty} A_n := \bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m$$

上限: $\omega\in RHS \Leftrightarrow \omega$属于无穷多个$A_n$，表明$A_n$发生无穷多次

下限: $\omega\in RHS \Leftrightarrow \exists N$, s.t. $\omega\in\bigcap_{m=N}^{\infty}A_m$，表明$A_n$至多有限多个不发生

记号 $\{A_n\ i.o.\} = \limsup_{n\to\infty} A_n$ (infinitely often)

定理(Borel-Cantelli引理)

(i) 当 $\sum_{n=1}^{\infty}P(A_n)<\infty$ 时，有 0-1律

$P(A_n\ i.o.)=0$

(ii) ~~且~~假设$\{A_n\}$相互独立，当$\sum_{n=1}^{\infty}P(A_n)=\infty$时，有 $P(A_n\ i.o.)=1$

证明: 令 $A=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m$

(i) $P(A)\le P(\bigcup_{m=n}^{\infty}A_m)\le\sum_{m=n}^{\infty}P(A_m)\longrightarrow 0$, $n\to\infty$

(ii) $A^c=\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}A_m^c$，因为

$$P(\bigcap_{m=n}^{\infty}A_m^c)=\lim_{r\to\infty}P(\bigcap_{m=n}^{r}A_m^c)$$

$$\overset{\text{独立}}{=}\lim_{r\to\infty}\prod_{m=n}^{r}(1-P(A_m)) \qquad 1-x\le e^{-x}$$

$$\le\lim_{r\to\infty}e^{-\sum_{m=n}^{r}P(A_m)}=0$$

必有 ~~$P(A$~~ $P(A^c)=0$

例2. 设$\{X_n\}$相互独立，且服从参数为1的指数分布，试证

$$P(\limsup_{n\to\infty}\frac{X_n}{\log n}=1)=1$$

解: $P(X_n\ge x)=e^{-x}$, $x\ge 0$

令 $A_n=\{\frac{X_n}{\log n}\ge 1+a\}$, $|a|<1$

则 $P(A_n)=\frac{1}{n^{a+1}}$，且$\{A_n\}$相互独立

分情形讨论:

① 当$a\in(-1,0]$时，$\sum_{n=1}^{\infty}P(A_n)=\infty$,

Borel-Cantelli引理(ii) $P(A_n\ i.o.)=1$,

即A_n几乎处处发生无穷多次，亦即

$P(\limsup_{n\to\infty}\frac{X_n}{\log n}\ge 1)=1$

② 当$a\in(0,1)$时，$\sum_{n=1}^{\infty}P(A_n)<\infty$

Borel-Cantelli引理(i) $P(A_n\ i.o.)=0$,

即A_n几乎必然只发生有限多次，

亦即 $P(\limsup_{n\to\infty}\frac{X_n}{\log n}\le 1)=1$.

总之，即可

§6.4 强大数律

弱LLN: $\{X_k\}$独立同，$E[X_1]$存在 $S_n=\sum_{i=1}^{n}X_i$,

$\frac{S_n}{n}\xrightarrow{P}E[X_1]$.

更强版本:

定理1. 设$\{X_k\}$独立同，$E[X_1^2]<\infty$，$E[X_1]=\mu$，则

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{2}\mu \qquad (*)$$

$$\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{a.s.}\mu \qquad (**)$$

(*)证明: $S_n=\sum_{i=1}^{n}X_i$

$$E[(\frac{1}{n}S_n-\mu)^2]=\frac{1}{n^2}\sum_{i=1}^{n}Var(X_i)=\frac{1}{n}Var(X_1)$$

$\to 0$

(※※)证明: 寻找子列 $n_i = i^2$

$$P\left(\left|\frac{S_{n_i}}{n_i}-\mu\right|>\varepsilon\right) \le \frac{Var(S_{n_i})}{(n_i\varepsilon)^2} = \frac{Var(X_1)}{n_i\varepsilon^2} = \frac{Var(X_1)}{i^2\varepsilon}$$

即 $\sum_i P\left(\left|\frac{S_{n_i}}{n_i}-\mu\right|>\varepsilon\right)<\infty$

可知 $\frac{S_{i^2}}{i^2}\xrightarrow{a.s.}\mu$

·先设 $X_k\ge 0$, $\forall k$, 则当 $i^2\le n\le (i+1)^2$ 时,

有 $S_{i^2}\le S_n\le S_{(i+1)^2}$

$$\underbrace{\frac{i^2}{(i+1)^2}\frac{S_{i^2}}{i^2}}_{\downarrow a.s.\ \mu} \le \frac{S_n}{n} \le \underbrace{\frac{S_{(i+1)^2}}{(i+1)^2}\cdot\frac{(i+1)^2}{i^2}}_{\downarrow a.s.\ \mu}$$

从而 $\frac{1}{n}S_n\xrightarrow{a.s.}\mu$

一般, $X_k=X_k^+-X_k^-$, $X_k^+=\max\{X_k,0\}$,

因 X_k^+, X_k^- 数学期望均存在, 则

$\frac{1}{n}\sum_{i=1}^n X_i^+\xrightarrow{a.s.}E[X_1^+]$,

$\frac{1}{n}\sum_{i=1}^n X_i^-\xrightarrow{a.s.}E[X_1^-]$,

合在一起即可

定理2 (柯尔莫哥洛夫强大数律)

设 $\{X_k\}$ 独立同, 则

$$\frac{1}{n}\sum_{i=1}^n \xrightarrow{a.s.}\mu \iff E[|X_1|]<\infty \text{ 且 } \mu=E[X_1]$$

证明: $\Leftarrow$ 先设 $X_k\ge 0$, $\forall k$, 分三步

Step 1. 截尾术.

引入 $Y_n = X_n I_{\{X_n<n\}} = \begin{cases} X_n & X_n<n \\ 0 & X_n\ge n\end{cases}$

则 $P(Y_n\ne X_n)=P(X_n\ge n)$

$\sum_n P(Y_n\ne X_n)=\sum_n P(X_n\ge n)=\sum_n P(X_1\ge n)$

$\le E[X_1]<\infty$

Fact1. $X\ge 0$, $\sum_{m=1}^\infty P(X\ge m)\le E[X]\le 1+\sum_{m=1}^\infty P(X\ge m)$

因为 $\sum_{m=1}^\infty (m-1) I_{\{m-1\le X\le m\}}$

$\le X=\sum_{m=1}^\infty X I_{\{m-1\le X\le m\}} \le \sum_{m=1}^\infty m I_{\{m-1\le X\le m\}}$

取期望: $\sum_{m=1}^\infty (m-1)P(m-1\le X\le m)$

$\le E[X]\le \sum_{m=1}^\infty m P(m-1\le X\le m)$

故 $\sum_{m=1}^\infty P(X\ge m)\le E[X]\le \sum_{m=1}^\infty P(X\ge m)+1$

·Borel-Cantelli引理(i)表明:

$P(X_n\ne Y_n\ \ i.o.)=0$, 即几乎处处 $\{X_n\ne Y_n\}$

只发生有限次, 从而 $\frac{1}{n}\sum_{i=1}^n (X_i-Y_i)\xrightarrow{a.s.}0$

进而 $\frac{1}{n}\sum_{i=1}^n X_i\xrightarrow{a.s.}\mu \iff \frac{1}{n}\sum_{i=1}^n Y_i\xrightarrow{a.s.}\mu$

Step 2. 几乎处处收敛子列

对 $\alpha>1$, 令 $\beta_k=[\alpha^k]$, 则

$\alpha^k-1<\beta_k\le\alpha^k$, $\frac{\beta_{k+1}}{\beta_k}\longrightarrow\alpha, k\to\infty$,

且存在 $A=A(\alpha)$, s.t. $\forall m\ge 1$, 有

$\sum_{k=m}^\infty \frac{1}{\beta^k}\le\frac{A}{\beta_m^2}$

记 $S_n'=\sum_{i=1}^n Y_i$, 对 $\alpha>1$, $\varepsilon>0$, 用Chebyshov

不等式: $\sum_n P\left(\frac{1}{\beta_n}|S'_{\beta_n}-E[S'_{\beta_n}]|\ge\varepsilon\right)$

$\le\sum_n\frac{Var(S'_{\beta_n})}{\varepsilon^2\beta_n^2}$

$\overset{\text{独立性}}{=\!=}\frac{1}{\varepsilon^2}\sum_{n=1}^\infty\frac{1}{\beta_n^2}\sum_{i=1}^{\beta_n}Var(Y_i)$

$=\frac{1}{\varepsilon^2}\sum_{n=1}^\infty\left(\sum_{n:\beta_n\ge i}\frac{1}{\beta_n^2}\right)Var(Y_i)$

$\le\frac{A}{\varepsilon^2}\sum_{i=1}^\infty\frac{1}{i^2}Var(Y_i)$

$\le\frac{A}{\varepsilon^2}\sum_{i=1}^\infty\frac{1}{i^2}E[Y_i^2]<\infty\ \boxed{?}$

因此 $\frac{1}{\beta_n}(S'_{\beta_n}-E[S'_{\beta_n}])\xrightarrow{a.s.}0$

又 $E[Y_n]=E[X_nI_{\{X_n<n\}}]=E[X_1I_{\{X_1<n\}}]$

$\xrightarrow{\text{单调收敛}}E[X_1]=\mu$

从而 $\frac{1}{\beta_n}\mathbb{E}[S_{\beta_n}'] = \frac{1}{\beta_n}\sum_{i=1}^{n}\mathbb{E}[Y_i] \longrightarrow \mu$

则 $\frac{1}{\beta_n}S_{\beta_n}' \xrightarrow{a.s} \mu \Rightarrow \frac{1}{\beta_n}S_{\beta_n} \xrightarrow{a.s.} \mu$

⋮

$\boxed{?}$ Fact 2: $\sum_{i=1}^{\infty}\frac{1}{i^2}\mathbb{E}[Y_i^2] < \infty$

因为: 记 $B_{ij} = \{j-1 \le X_i \le j\}$, 则

$\sum_i \frac{1}{i^2}\mathbb{E}[Y_i^2] = \sum_{i=1}^{\infty}\frac{1}{i^2}\sum_{j=1}^{i}\mathbb{E}[Y_i^2 I_{B_{ij}}]$

$\le \sum_{i=1}^{\infty}\frac{1}{i^2}\sum_{j=1}^{i} j^2 P(B_{ij})$

$= \sum_{i=1}^{\infty}\frac{1}{i^2}\sum_{j=1}^{i} j^2 P(B_{1j})$

$= \sum_{j=1}^{\infty}\left(\sum_{i=j}^{\infty}\frac{1}{i^2}\right) j^2 P(B_{1j})$

$\le 2\sum_{j=1}^{\infty} j P(B_{1j}) \overset{Fact 1}{\le} 2(1+\mathbb{E}[X_1]) < \infty$

⋮

Step 3 几乎处处收敛

S_n 关于 n 单调增, 取 $\beta_m \le n < \beta_{m+1}$, 由

$\frac{\beta_m}{\beta_{m+1}}\cdot\frac{S_{\beta_m}}{\beta_m} \le \frac{S_n}{n} \le \frac{S_{\beta_{m+1}}}{\beta_{m+1}}\cdot\frac{\beta_{m+1}}{\beta_m}$

知: $\frac{1}{\alpha}\mu \le \liminf_{n\to\infty}\frac{S_n}{n} \le \limsup_{n\to\infty}\frac{S_n}{n} \le \alpha\mu$, a.s.

让 $\alpha \downarrow 1$ 知: $\lim_{n\to\infty}\frac{S_n}{n} = \mu$ a.s.

⋮

最后, $X_i = X_i^+ - X_i^-$ 即可。

⋮

$\Longrightarrow$ 由 $\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{a.s.} \mu \in \mathbb{R}$ 可知: $\frac{1}{n}X_n \xrightarrow{a.s.} 0$

立即有 $\sum_{n=1}^{\infty}P(|X_n| \ge n) < \infty$

(利用 Borel-Cantelli 引理 (ii))

否则 $P(\frac{1}{n}|X_n| \ge 1 \text{ i.o.}) = 1$, 与

$\frac{1}{n}X_n \xrightarrow{a.s.} 0$ 矛盾

即 $\sum_{n=1}^{\infty}P(|X_1| \ge n) < \infty$.

Fact 1 表明, $\mathbb{E}[|X_1|] \le 1 + \sum_{n=1}^{\infty}P(|X_1| \ge n)$

$< \infty$ 即 X_1 期望存在, $\Leftarrow$ 已证部分可知

$\mu = \mathbb{E}[X_1]$

注: 若 $\mathbb{E}[X_1^+] = \infty$, $\mathbb{E}[X_1^-] < \infty$, 则 $\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{a.s.} \infty$

定理1推论 (Borel 强 LLN):

试验中A发生概率为 p, S_n 为 n 次独立重复试验中A发生次数, 则

$\frac{1}{n}S_n \xrightarrow{a.s.} p$

独立和 $S_n = \sum_{i=1}^{n}X_i$, $0 < Var(X_i) < \infty$,

强 LLN $\quad \frac{S_n}{n} \xrightarrow{a.s.} \mu = \mathbb{E}[X_1]$

CLT $\quad \frac{S_n}{\sqrt{n\sigma^2}} \xrightarrow{D} N(0,1)$

定理3. (重对数律) 设 $\{X_k\}$ 独立同,

$\mathbb{E}[X_1] = 0$, $Var(X_1) = 1$, 则

$P\left(\limsup_{n\to\infty}\frac{S_n}{\sqrt{2n\log\log n}} = 1\right) = 1$

$P\left(\liminf_{n\to\infty}\frac{S_n}{\sqrt{2n\log\log n}} = -1\right) = 1$

习题四.

(I) 设 $\{X_k\}$ 独立同 $N(0,1)$, 令 $S_n = \sum_{i=1}^{n}X_i$, 对 $k < n$, 求:

(i) $S_k = y$ 下 S_n 的条件分布

(ii) 条件密度 $f_{S_k|S_n}(y|x)$ 与条件期望 $\mathbb{E}[S_k|S_n]$

解: (i) $S_n = S_k + \sum_{i=k+1}^{n}X_i \quad S_n - S_k = \sum_{i=k+1}^{n}X_i$

右边均值为0, 方差为 $n-k$, 因此给定

$S_k=y$下，S_n为均值为y，方差为$n-k$的正态分布

(ii) $f_{S_k|S_n}(y|x)=\dfrac{f_{S_k,S_n}(y,x)}{f_{S_n}(x)}$

$=\dfrac{f_{S_n|S_k}(x|y)f_{S_k}(y)}{f_{S_n}(x)}$

$=\dfrac{1}{\sqrt{2\pi(n-k)}}e^{-\frac{1}{2(n-k)}(x-y)^2}\dfrac{e^{-\frac{1}{2k}y^2}}{\sqrt{2\pi k}}\cdot\dfrac{\sqrt{2\pi n}}{e^{-\frac{1}{2n}x^2}}$

$=\sqrt{\dfrac{n}{2\pi k(n-k)}}e^{-\frac{n}{2k(n-k)}(y-\frac{kx}{n})^2}$

$E[S_k|S_n=x]=\dfrac{kx}{n}$

$E[S_k|S_n]=\dfrac{k}{n}S_n$

(II) 气体分子平衡态Maxwell分布.(1859)

解: 气体分速度分布 $U=(X,Y,Z)$ 三个假设:

(i) X,Y,Z相互独立且同分布

(ii) 具有连续可微的密度函数

(iii) 平衡态时速度分布在空间各向同性

分析: X,Y,Z共同密度为$p(x)$，则联合密度 $p(x)p(y)p(z)$，根据(iii)可知

$p(x)p(y)p(z)=q(r)$, $r=x^2+y^2+z^2$

取log后: $\log p(x)+\log p(y)+\log p(z)=\log q(r)$

对x求导: $\dfrac{p'(x)}{p(x)}=\dfrac{q'(r)}{q(r)}\cdot 2x$

即: $\dfrac{p'(x)}{xp(x)}=2\dfrac{q'(r)}{q(r)}$

同理: $\dfrac{p'(x)}{xp(x)}=\dfrac{p'(y)}{yp(y)}=\dfrac{p'(z)}{zp(z)}=2\dfrac{q'(r)}{q(r)}$

表明 $\dfrac{q'(r)}{q(r)}$ 与x,y,z无关，即$\exists a\in\mathbb{R}$, s.t.

$\dfrac{p'(x)}{xp(x)}=a$

解之 $p(x)=be^{\frac{a}{2}x^2}$

若记 $a=-\dfrac{m}{kT}$，利用规范条件 $\int p(x)dx=1$，

可知 $p(x)=\sqrt{\dfrac{m}{2\pi kT}}e^{-\frac{m}{2kT}x^2}$, $x\in\mathbb{R}$

(III) 量子力学Fermi子波函数

Schrödinger Equation

$i\hbar\dfrac{\partial\Psi}{\partial t}=\hat{H}\Psi$, $\hat{H}=-\dfrac{\hbar}{2m}\Delta+V$

Born 几率诠释:

$\int|\Psi|^2d\vec{r}=1$

一维谐振子 $V(x)=\frac{1}{2}mw^2x^2$

定态解: $\hat{H}\Psi=E\Psi$

取适当m与w，约化为

$-\Delta\Psi+x^2\Psi=E\Psi$

可解: $\varphi_n(x)=e^{-\frac{1}{2}x^2}H_n(x)$,

$H_n(x)$为Hermite多项式(在不同学科中均出现)

可构造多体波函数:

$$\Psi(x_1,\cdots,x_n)=\begin{vmatrix}\varphi_0(x_1) & \cdots & \varphi_0(x_n)\\ \varphi_1(x_1) & \cdots & \varphi_1(x_n)\\ \cdots & & \cdots\\ \varphi_{n-1}(x_1) & \cdots & \varphi_{n-1}(x_n)\end{vmatrix}$$

可具体计算 $\int\cdots\int|\Psi(x_1,\cdots,x_n)|^2dx_1\cdots dx_n$

(可参见第四章Hermite

习题五

(I) (1) 试证若 $X_n\xrightarrow{D}X$, $Y_n\xrightarrow{P}c$ ($c>0$为常数), 则 $X_nY_n\xrightarrow{D}cX$; (2) 若$X_1,\cdots,X_n$为独立同的非负随机变量，且$E(X_i)=1$, $Var(X_i)=\sigma^2$ $(0<\sigma<\infty)$, 令$S_n=\sum_{i=1}^nX_i$，则 $\dfrac{2}{\sigma}(\sqrt{S_n}-\sqrt{n})\xrightarrow{D}N(0,1)$

解: (1) 作业 7.2.5(a)

(2) $\dfrac{2}{\sigma}(\sqrt{S_n}-\sqrt{n})=\underbrace{\dfrac{S_n-n}{\sqrt{n\sigma^2}}}_{\text{CLT}\xrightarrow{D}N(0,1)}\cdot\dfrac{2}{1+\sqrt{\frac{S_n}{n}}}$ (强大数律)

(II) 利用概率方法证明：任给 $q>p>0$ 有：

$$\lim_{n\to\infty}\int_0^1\cdots\int_0^1 \frac{x_1^q+x_2^q+\cdots+x_n^q}{x_1^p+x_2^p+\cdots+x_{2n}^p}dx_1\cdots dx_{2n}=\frac{p+1}{2(q+1)}$$

证明：只需证

$$\int_0^1\cdots\int_0^1 \frac{x_1^q+x_2^q+\cdots+x_{2n}^q}{x_1^p+x_2^p+\cdots+x_{2n}^p}dx_1\cdots dx_{2n}\xrightarrow{n\to\infty}\frac{p+1}{q+1}$$

取独立同 $[0,1]$ 上均匀分布 $X_1,\cdots,X_{2n}$

则 $X_1^q,\cdots,X_{2n}^q$ 独立同，且

$$E[X_1^q]=\int_0^1 x^q dx=\frac{1}{q+1}$$

利用强 LLN，$\frac{1}{2n}\sum_{i=1}^{2n}X_i^q\xrightarrow{a.s.}\frac{1}{q+1}$，

$\frac{1}{2n}\sum_{i=1}^{n}X_i^p\xrightarrow{a.s.}\frac{1}{p+1}$，

另一方面，

$$LHS=E\left[\frac{X_1^q+\cdots+X_{2n}^q}{X_1^p+\cdots+X_{2n}^p}\right]$$

$$=E\left[\frac{(X_1^q+\cdots+X_{2n}^q)/(2n)}{(X_1^p+\cdots+X_{2n}^p)/(2n)}\right]$$

$$\xrightarrow{\text{控制收敛}}\frac{p+1}{q+1}$$

(III) 设 $\mu=E[X]$，$\sigma^2=Var(X)$，$\sigma>0$，试证：

$$P(X\geq\mu+a)\leq\frac{\sigma^2}{\sigma^2+a^2}\quad(a>0)$$

分析：令 $\mu=0$，

$$a=E[a-X]=E[(a-X)I_{\{X\geq a\}}]+E[(a-X)I_{\{X<a\}}]$$

$$\leq E[(a-X)I_{\{X<a\}}]$$

$$\leq\sqrt{E[(a-X)^2]E[I_{\{X<a\}}]}$$

$$=\sqrt{(a^2+\sigma^2)(1-P(X\geq a))}$$

(IV) 从对称随机游走到布朗运动

分析：$S_n=\sum_{i=1}^n X_i$，$\{X_i\}$ i.i.d.，$S_0=0$

$$P(X_i=1)=P(X_i=-1)=\frac{1}{2}$$

CLT：$\frac{S_n}{\sqrt{n}}\xrightarrow{D}N(0,1)$

$\forall t>0$，CHECK ~~$S_{[nt]}$~~ $\frac{S_{[nt]}}{\sqrt{n}}\xrightarrow{D}B_t$

$B_t\sim N(0,t)$

回顾：$Cov(S_n,S_m)=E[S_nS_m]=m\wedge n$（取小）

$S_{n+m}-S_n$ 与 S_n 独立

随机过程 B_t 有如下性质，称**布朗运动**：

① $\forall t_0=0<t_1<\cdots<t_n$ 有 $B_{t_1}, B_{t_1}-B_{t_0},\cdots,B_{t_n}-B_{t_{n-1}}$ 相互独立，且 $B_t\sim N(0,t)$

② $E[B_sB_t]=s\wedge t$，$s,t>0$

③ B_t 关于 t 连续

下阶段最重要研究对象！

作业题：$\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$

Hermite 多项式：$H_n(x)=(-1)^n\frac{1}{\phi(x)}\left(\frac{d^n}{dx^n}\phi(x)\right)$

参见习题 4.14, 37，利用

$$\sum_{n=0}^{\infty}H_n(x)\frac{t^n}{n!}=e^{xt-\frac{1}{2}t^2}$$

可证：① $H_n'(x)=nH_{n-1}(x)$

② $H_{n+1}(x)=xH_n(x)-nH_{n-1}(x)$

③ $H_n''(x)-xH_n'(x)+nH_n(x)=0$

引入 $\varphi_n(x)=\frac{1}{\sqrt{n!}}\sqrt{\phi(x)}H_n(x)$

$$\Psi(x_1,\cdots,x_N)=\det\left(\varphi_{i-1}(x_j)\right)_{i,j=1}^N$$

试证：

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}|\Psi(x_1,\cdots,x_N)|^2dx_1\cdots dx_N=N!$$

第七章 概率论外篇

内容:

- ▷信息熵
- ▷Lindeberg替换
- ▷矩问题
- ▷随机矩阵初步

§7.1 信息熵(entropy)

什么是信息?

对事件发生惊奇程度. 例如:

3个骰子点数之和 $\begin{cases} \text{偶数} & \frac{1}{2}\text{概率} \\ 18 & \frac{1}{6^3}\text{概率} \end{cases}$

概率为 p 事件发生后, 用 $S(p)$ 表示惊奇程度, $S(p)$, $0<p\le 1$ 有何要求?

公理1. $S(1)=0$

必然事件不带来惊奇

公理2. $S(p_1)>S(p_2)$, $p_1<p_2$ 严格减

越不可能发生, 越表达更多信息

公理3. $S(p)$ 为 p 的连续函数

p 微小变化引起 $S(p)$ 微小变化

公理4. $S(pq)=S(p)+S(q)$, $0<p,q\le 1$

E、F 独立, $P(E)=p$, $P(F)=q$, $E\cap F$ 惊奇数 $S(pq)$, 则 $S(pq)-S(p)$ 表示听到 E 发生后增加惊奇度, 但 E 与 F 独立, 故应该 $S(pq)-S(p)=S(q)$

定理1. 若 $S(p)$ 满足公理1-4, 则 $\exists C>0$, s.t.

$S(p)=-c\log p$

分析: $S(p^2)=2S(p)$, $\cdots$, $S(p^m)=mS(p)$

$S(p^{\frac{1}{n}})=\frac{1}{n}S(p)$, $S(p^{\frac{m}{n}})=\frac{m}{n}S(p)$

公理3表明 $S(p^x)=xS(p)$

取 $x=-\log p$, 则 $S(p)=S\left(\frac{1}{e^x}\right)=xS\left(\frac{1}{e}\right)$

$=-c\log p$, 这里 $C=S\left(\frac{1}{e}\right)>S(1)=0$,

(这里利用了公理1, 2)

离散型

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

定义1. Shannon 信息熵

$$H(X)=-\sum_{i=1}^{n}p_i\log p_i$$

联合熵

$$H(X,Y)=-\sum_i\sum_j p(x_i,y_j)\log p(x_i,y_j)$$

相对熵:

$$H_Y(X)=\sum_j H_{Y=y_j}(X)P_Y(y_j)$$

这里 $H_{Y=y_j}(X)=-\sum_i p(x_i|y_j)\log p(x_i|y_j)$

引理2. $H(X,Y)=H(Y)+H_Y(X)$

分析: $H_Y(X)=\sum_j\left(\sum_i p(x_i|y_j)\log p(x_i|y_j)\right)P_Y(y_j)$

$=-\sum_{i,j}p(x_i,y_j)\log\frac{p(x_i,y_j)}{p(y_j)}$

$=H(X,Y)-H(Y)$

注: 此定义可扩展到一般离散型

定理3. $H_Y(X) \le H(X)$,

取等号 $\Longleftrightarrow$ X与Y独立

分析: $H_Y(X) - H(X)$

$= -\sum_j \sum_i \left(p(x_i|y_j)\log p(x_i|y_j)\right) P_Y(y_j)$

$+ \sum_i \sum_j p(x_i, y_j) \log P_X(x_i)$

$= \sum_{i,j} p(x_i, y_j) \log \frac{P_X(x_i)}{p(x_i|y_j)}$　　$\log x \le x-1$

$\le \sum_{i,j} p(x_i, y_j) \left(\frac{P_X(x_i)}{p(x_i|y_j)} - 1\right)$

$= \sum_{i,j} P_X(x_i) P_Y(y_j) - \sum_{i,j} p(x_i, y_j)$

$= 0$

易知取等号条件✓

注: $P(X = x_i) = p_i$, $i = 1, \cdots, n$,

$H(X) = -\sum_{i=1}^n p_i \log p_i$　　$(\sum_i p_i = 1)$

$\max\{H(X)\} = H(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}) = \log n$.

还有 $H(\frac{1}{n-1}, \cdots, \frac{1}{n-1}) < H(\frac{1}{n}, \cdots, \frac{1}{n})$

熵表达一个系统混乱程度的一种度量

连续型X有密度f,

定义2. $H(X) = -\int_{-\infty}^{\infty} f(x) \log f(x)\, dx$

$H(X, Y) = -\int_{\mathbb{R}^2} f(x,y) \log f(x,y)\, dx dy$

引入 $\eta(u) = -u\log u$, $u \ge 0$, 则

$\eta(u) - \eta(v) = \eta'(v)(u-v) + \frac{1}{2}\eta''(\xi)(u-v)^2$

$\le \eta'(v)(u-v) \le -(\log v + 1)(u-v)$

Gibbs不等式 $u - u\log u \le v - u\log v$

立即有: $\int (f - f\log f)\, dx \le \int (g - f\log g) dx$

这里f, g均为密度, 进而

$-\int f \log f\, dx \le -\int f \log g\, dx$　　$(*)$

定理4. 记 $D = \{x \in \mathbb{R} : f(x) > 0\}$

(1) $D = (-\infty, \infty)$, $E[X] = 0$, $Var(X) = 1$,

正态分布时熵最大, 最大熵为 $\log\sqrt{2\pi e}$

(2) $D = (0, \infty)$, $E[X] = \frac{1}{\lambda}$, $\lambda > 0$,

指数分布 $Exp(\lambda)$ 熵最大 $\log\frac{e}{\lambda}$

(3) $D = [0, a]$, 均匀分布熵最大 $\log a$.

证明: (1) 取 $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, 则

$H(X) \le -\int f \log\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\right) dx$

$= -\int \left(\log\frac{1}{\sqrt{2\pi}} - \frac{1}{2}x^2\right) f\, dx$

$= \log\sqrt{2\pi} + \frac{1}{2}$　　标准正态分布熵

Boltzmann熵 $S = k_B \log W$

S: 宏观系统熵　　W: 微观状态数

k_B: Boltzmann 常数

$\frac{1}{W}$, W结果, 熵 $-\sum_{i=1}^W \frac{1}{W}\log\frac{1}{W} = \log W$

能级: $E_1, E_2, \cdots, E_n$

概率: $p_1, p_2, \cdots, p_n$

平均能量: $\sum_{i=1}^n p_i E_i = U$

最大熵时, $p_i = \frac{1}{Z} e^{-\beta E_i}$　　Gibbs分布

配分函数 $Z = \sum_i e^{-\beta E_i}$, β 由U决定

Lagrange乘子法:

$L = -\sum_i p_i \log p_i - \beta(\sum_i p_i E_i - U) - \lambda(\sum_i p_i - 1)$

$\frac{\partial L}{\partial p_i} = -(1 + \log p_i) - \beta E_i - \lambda = 0$

$\frac{\partial L}{\partial \beta} = -\sum_i p_i E_i + U = 0$

$\frac{\partial L}{\partial \lambda} = -\sum_i p_i + 1 = 0$

即 $p_i = e^{-1-\lambda-\beta E_i}$, $\sum_i p_i = 1$, $\sum_i p_i E_i = U$

f为密度，约束：$\int_{\mathbb{R}} h(x) f(x) dx = \bar{h}$ 能量

何时 $H(f)$ 最大？ (Gibbs分布)

猜：$f_0(x) = \frac{e^{-ch(x)}}{\int e^{-ch(x)} dx}$, $c \in \mathbb{R}$

类似正态分布情形，最大熵为

$H(f_0) = \ln Z + c\bar{h}$ 热力学熵

§7.2. Lindeberg Replacement Strategy

$X_1, \cdots, X_n$ 相互独立，$E[X_k] = 0$, ~~…~~

$b_k^2 = Var(X_k)$, $B_n = \sqrt{b_1^2 + \cdots + b_n^2}$

Lindeberg条件：(1922)

$\forall \varepsilon > 0$, $\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^n \int_{|x|>\varepsilon B_n} x^2 dF_k = 0$ (L)

CLT. 设$\{X_i\}$相互独立，满足(L)，则

$\frac{1}{B_n^2} \max_{1\le k\le n} b_k^2 \longrightarrow 0$, (Feller)

且 $\frac{1}{B_n} \sum_{k=1}^n X_k \xrightarrow{D} N(0,1)$

1922. Lindeberg

(i) 当$\{Y_k\}$为独立正态随机变量，$E[Y_k]=0$

$Var(Y_k) = b_k^2$ 时，

$E[g(\frac{1}{B_n} \sum_{k=1}^n Y_k)] = E[g(G)] + o(1)$,

$G \sim N(0,1)$

(ii) 三阶矩存在时~~…~~，对$\{X_k\}$，

$E[g(\frac{1}{B_n} \sum_{k=1}^n X_k)] = E[g(\frac{1}{B_n} \sum_{k=1}^n Y_k)] + o(1)$

(iii) 截断术去掉三阶假设

回顾：$X_n \xrightarrow{D} X$, $\Leftrightarrow \forall g \in C_b(\mathbb{R})$, (有界连续函数全体)

$E[g(X_n)] \longrightarrow E[g(X)]$

$\Leftrightarrow \forall g, g', g'', g''' \in C_b(\mathbb{R})$, $E[g(X_n)] \to E[g(X)]$

引入$\{Y_k\}$独立正态随机变量，且与$\{X_k\}$独立，$E[Y_k] = 0$, $b_k^2 = Var(Y_k)$

{思考：两个分布函数F, G, ~~是~~试CHECK存在独立随机变量X与Y, s.t. 对应分布函数为F, G)

令 $\mathfrak{Z}_k = X_1 + \cdots + X_{k-1} + Y_{k+1} + \cdots + Y_n$

则 $\mathfrak{Z}_n + X_n = \sum_{k=1}^n X_k := S_n$

$\mathfrak{Z}_1 + Y_1 = \sum_{k=1}^n Y_k \overset{D}{=} B_n Y$, $Y \sim N(0,1)$

更一般 $1 \le k < n$ 有

$\mathfrak{Z}_k + X_k = \mathfrak{Z}_{k+1} + Y_{k+1}$

$E[g(\frac{S_n}{B_n})] - E[g(Y)]$

$= \sum_{k=1}^n (E[g(\frac{\mathfrak{Z}_k + X_k}{B_n})] - E[g(\frac{\mathfrak{Z}_k + Y_k}{B_n})])$

每一步作替换 $X_n \to Y_n$, $X_{n-1} \to Y_{n-1}, \cdots, X_1 \to Y_1$

Observe：因$\mathfrak{Z}_k$, X_k, Y_k独立，有

$E[g'(\frac{\mathfrak{Z}_k}{B_n})(X_k - Y_k)] = 0$,

$E[g''(\frac{\mathfrak{Z}_k}{B_n})(X_k^2 - Y_k^2)] = 0$

引入 $h(t) = \sup_x \{g(x+t) - g(x) - g'(x)t - \frac{1}{2} g''(x) t^2\}$

则易知 $\exists K > 0$, s.t. $h(t) \le K\, t^2 \wedge |t|^3$

因此 $|\mathbb{E}[g(\frac{Z_k+X_k}{B_n})-g(\frac{Z_k+Y_k}{B_n})]|$

$\leq \mathbb{E}[h(\frac{X_k}{B_n})]+\mathbb{E}[h(\frac{Y_k}{B_n})]$

只须: $\sum_{k=1}^{n}\mathbb{E}[h(\frac{X_k}{B_n})]\longrightarrow 0$ ①

$\sum_{k=1}^{n}\mathbb{E}[h(\frac{Y_k}{B_n})]\longrightarrow 0$ ②

先看②: $h(t)\leq K|t|^3$.

LHS of ② $\leq K\sum_{k=1}^{n}[|\frac{Y_k}{B_n}|^3]$

$=K\sum_{k=1}^{n}\frac{b_k^3}{B_n^3}\mathbb{E}[|Y|^3]$

$\leq K\cdot\max_{1\leq k\leq n}b_k\cdot\frac{1}{B_n}\mathbb{E}[|Y|^3]\longrightarrow 0$ (Feller)

对①: 划分 $\{|X_k|\leq\varepsilon B_n\}$, $\{|X_k|>\varepsilon B_n\}$

LHS of ① $\leq K\cdot\sum_{k=1}^{n}$ ~~$\mathbb{E}[|X_k|^3]$~~ $\mathbb{E}[h(\frac{X_k}{B_n})]$

$\leq K\cdot\sum_{k=1}^{n}\frac{1}{B_n^3}\mathbb{E}[|X_k|^3 I_{\{|X_k|\leq\varepsilon B_n\}}]$

$+K\cdot\sum_{k=1}^{n}\frac{1}{B_n^2}\mathbb{E}[|X_k|^2 I_{\{|X_k|>\varepsilon B_n\}}]$

$\leq K\cdot\sum_{k=1}^{n}\frac{\varepsilon}{B_n^2}\mathbb{E}[|X_k|^2 I_{\{|X_k|\leq\varepsilon B_n\}}]$

$+K\cdot\sum_{k=1}^{n}\frac{1}{B_n^2}\mathbb{E}[|X_k|^2 I_{\{|X_k|>\varepsilon B_n\}}]$

$\leq\varepsilon K+K\cdot\underbrace{\sum_{k=1}^{n}\frac{1}{B_n^2}\int_{|x|>\varepsilon B_n}x^2dF_k}_{\text{Lindeberg条件}}\longrightarrow 0$

·何时矩唯一确定随机变量?

反例: 对数正态密度

$f_0(x)=\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{x}e^{-\frac{1}{2}(\log x)^2}$, $x>0$,

引入 $f_a(x)=f_0(x)(1+a\sin(2\pi\log x))$, $a\in[-1,1]$

Claim: $\int_0^\infty x^k f_a(x)dx=e^{\frac{1}{2}k^2}$, $k=0,1,\cdots$

一个充分条件

定理: 设 $\limsup_{k\to\infty}\frac{1}{2k}(\gamma_{2k})^{\frac{1}{2k}}=r<\infty$, 则

至多存在一个分布函数 F s.t.

$\gamma_k=\int x^k dF$, $k=0,1,\cdots$

证明: 令 $\mu_k=\int|x|^k dF(x)$, 则由 C-S 不等式 $\mu_{2k+1}\leq\sqrt{\mu_{2k}\mu_{2k+2}}$

因此 $\limsup_{k\to\infty}\frac{1}{k}(\mu_k)^{\frac{1}{k}}=r<\infty$

Fact:

$e^{it}-\sum_{m=0}^{n-1}\frac{(it)^m}{m!}=\frac{i^n}{(n-1)!}\int_0^t(t-s)^{n-1}e^{is}ds$

$|e^{it}-\sum_{m=0}^{n-1}\frac{(it)^m}{m!}|\leq\frac{|t|^n}{n!}$, $\forall t\in\mathbb{R}$

于是 $|e^{i\theta x}(e^{itx}-\sum_{m=0}^{n-1}\frac{(itx)^m}{m!})|\leq\frac{|tx|^n}{n!}$

取期望:

$|\phi(\theta+t)-\sum_{m=0}^{n-1}\frac{t^m}{m!}\phi^{(m)}(\theta)|\leq\frac{|t|^n}{n!}\mu_n$

又当 n 充分大时, 对 $\varepsilon>0$ 有

$\mu_n\leq(r+\varepsilon)^n n^n$

利用 $\frac{n^n}{n!}\leq e^n$ 知,

$\phi(\theta+t)=\phi(\theta)+\sum_{m=1}^{\infty}\frac{t^m}{m!}\phi^{(m)}(\theta)$ (*)

当 $|t|<\frac{1}{er}$ 时, $\frac{|t|^n}{n!}\mu_n\leq(\frac{|t|}{er})^n$, 故上级数可展开.

对 $\theta=0$, $\phi(0)=1$, $\phi^{(m)}(0)$ 由矩 γ_m 给出,

故 $\phi(t)$, $|t|<\frac{1}{er}$

由(*), 再次取 $\theta=\pm\frac{1}{er}$ 时决定

$\phi(t)$, $|t|<\frac{2}{er}$

如此下去, 矩决定 $\phi(t)$, 进而确定 F

§7.3 随机矩阵初步

When randomness meets matrix

$(\Omega, \mathcal{F}, P)$

随机变量 $X: \Omega \to \mathbb{R}$，可测

随机向量 $\vec{X}: \Omega \to \mathbb{R}^n$，可测映射

随机矩阵：矩阵值随机变量，或矩阵元为随机变量

随机矩阵 == 矩阵论 + 概率论

无处不在

1. 起源

(i) 统计 1928 Wishart

样本协方差阵

$X_k = (X_{1k}, \cdots, X_{pk})^T$

$W = \frac{1}{n}\sum_{k=1}^{n} X_k X_k^T = \frac{1}{n} X X^T$

这里 $X = (X_1, \cdots, X_n)$，$p \times n$ 矩阵，

$\{X_{ij}\}$ 为独立同 $N(0,1)$

矩阵 X 的矩阵之联合密度：$f(X) = \frac{1}{(\sqrt{2\pi})^{pn}} e^{-\frac{1}{2}\operatorname{tr} XX^T}$

(ii) 物理 1950S Wigner (1902-1995)

Schrödinger 方程：

$i\hbar \frac{\partial}{\partial t}\psi(t,\vec{x}) = H\psi(t,\vec{x})$

$H = -\frac{\hbar}{2m}\Delta + V(\vec{x})$ 哈密顿算子

$\psi(t,\vec{x})$ 波函数，$\int |\psi|^2 d\vec{x} = 1$

1950S Wigner

粒子多、强相互作用 ⟶ 复杂度增加

⟶ 随机性进入

H 截断 ⟶ $n \times n$ 矩阵 H_n，$H_n^* = H_n$，

矩阵元为随机变量

Wigner 观点：用随机矩阵模拟复杂体系哈密尔顿算子

最初模型：$H_{2n+1} = (V_{ij})_{i,j=-n}^{n}$，$V_{ii} = 0, \forall i$，

$\{V_{ij}\}$ 独立同 (Z)，$P(Z=1) = P(Z=-1) = \frac{1}{2}$

必须研究 $n \to \infty$ 时 H_n 谱的渐近性，

进而逼近 H 谱！ 大维现象

Connection with...

▷ MIMO 无线通信

▷ 高维数据

▷ 统计与量子物理

▷ 自由概率论

▷ 表示论 群 S_n

▷ 数论 Riemann ζ 函数

$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p (1 - \frac{1}{p^s})^{-1}$，$\operatorname{Re}(s) > 1$

Symmerty:

$\Gamma(\frac{s}{2})\pi^{-\frac{s}{2}}\zeta(s) = \Gamma(\frac{1-s}{2})\pi^{-\frac{1-s}{2}}\zeta(1-s)$

Riemann Hypothesis:

Non-trivial zeroes of $\zeta(s)$ lie on $\operatorname{Re}(s) = \frac{1}{2}$ (critical line)

传说：Hilbert-Polya Conjecture:

Non-trivial zeroes of $\zeta(s)$ are the spectrum of a self-disjoint

Hamiltonian
~~Hamitionian~~ operator.

2. 高斯正交系综 (Gaussian Orthogonal Ensemble, GOE)

$X_n : \Omega \longrightarrow M_{n\times n}(\mathbb{R})$

$X_n(\omega) \longrightarrow (x_{ij}(\omega))_{i,j=1}^n$

假设$\{x_{ij}\}$为iid $N(0,\sigma^2)$，构造对称矩阵 $A_n = \frac{1}{2}(X_n + X_n^T)$

明显：$a_{ii} \sim N(0,\sigma^2)$，$a_{ij} = a_{ji} \sim N(0, \frac{\sigma^2}{2})$

且$\{a_{ij}\}_{i\le j}$ 相互独立

A_n矩阵元联合密度：

$$f(A_n) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n a_{ii}^2} \cdot \frac{1}{(\sqrt{\pi\sigma^2})^{\frac{1}{2}n(n-1)}} e^{-\frac{1}{\sigma^2}\sum_{i<j} a_{ij}^2}$$
$$= \frac{1}{(\sqrt{2})^n} \cdot \frac{1}{(\sqrt{\pi\sigma^2})^{\frac{1}{2}n(n+1)}} e^{-\frac{1}{2\sigma^2}\operatorname{tr} A_n^2}$$

此时称A_n为一个GOE矩阵，记 $A_n \sim GOE_n(\sigma)$

GOE具有正交群作用下不变性，即：

$\forall Q \in O(n) := \{X \in M_{n\times n} : XX^T = I\}$，

令$B_n = QA_nQ^{-1}$，则B_n亦为对称阵，

Claim: $B_n \sim GOE_n(\sigma)$

因为$B_n = \frac{1}{2}(QX_nQ^T + (QX_nQ^T)^T)$

只须说明：QX_n ~~的~~ 矩阵元为iid $N(0,\sigma^2)$

另一方面，记$X_n = (\vec{x_{n1}}, \cdots, \vec{x_{nn}})$，而$\forall k$，

$\vec{x_k}$为$N(0,\sigma^2 I)$，可知$Q\vec{x_{nk}} \sim N(0,\sigma^2 I)$

定理1：设$A_n \sim GOE_n(\sigma)$，记$x_1,\cdots,x_n$为A_n特征值，则特征值联合密度为

$$\rho_n(x_1,\cdots,x_n) = \frac{1}{C_n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2} \cdot \prod_{1\le j<k\le n} |x_k - x_j|$$

这里$C_n = (2\pi)^{\frac{1}{2}n} \sigma^{\frac{1}{2}n(n+1)} \prod_{k=1}^n \frac{\Gamma(1+\frac{k}{2})}{\Gamma(1+\frac{1}{2})}$

来自Jacobian行列式

排斥效应，非独立，强相互作用

More is different 多则异

当$n=2$时，定理1易，大致想法：

$|\lambda I - A_n| = 0 \Rightarrow \lambda = \frac{1}{2}(a_{11} + a_{12} \pm \sqrt{\Delta})$

$\Delta = (a_{11} - a_{22})^2 + 4a_{12}$

明显$\Delta = (\lambda_+ - \lambda_-)^2$ ~~$|a_{11} - a_{12}| =$~~

$|a_{11} - a_{12}| = \sqrt{(\lambda_+ - \lambda_-)^2 - 4a_{12}^2}$

取$\lambda_0 = a_{12}$，算$(a_{11}, a_{12}, a_{22}) \longmapsto (\lambda_+, \lambda_-, \lambda_0)$

$\text{Jacobian} = \frac{\sqrt{(\lambda_+ - \lambda_-)^2 - 4\lambda_0^2}}{|\lambda_+ - \lambda_-|}$，

对λ_0积分后保留λ_+，λ_-即可

3. 半圆律

实Wigner矩阵：$A_n = (a_{ij})$，$A_n = A_n^T$

$\{a_{ii}\}$同分布，与Y

$\{a_{ij}\}_{i<j}$同分布，与Z

$\{a_{ij}\}_{i\le j}$相互独立

$\mathbb{E}[Y] = \mathbb{E}[Z] = 0$. $Var(Z) = 1$. $Var(Y) < \infty$,

$\forall k \ge 3$时，$\mathbb{E}[|Y|^k]$, $\mathbb{E}[|Z|^k] < \infty$

记$\gamma_k = \begin{cases} 0, & k = 2m+1 \\ \frac{1}{m+1}\binom{2m}{m} & k = 2m \end{cases}$

对象：$\frac{A_n}{\sqrt{n}}$

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

定理2: $\forall k=0,1,2,\cdots$, 有

$\frac{1}{n}\mathbb{E}[\mathrm{tr}(\frac{A_n}{\sqrt{n}})^k] \xrightarrow{n\to\infty} \gamma_k$

先算 $\int x^k w(x)dx = \int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \frac{x^k}{4\pi} I_{\{x^2+y^2\le 4\}}\,dydx$

$=\frac{1}{4\pi}\iint_{x^2+y^2\le 4} x^k\,dxdy$.

引入 $x=r\cos\theta$, $y=r\sin\theta$

$=\frac{1}{4\pi}\int_0^2 r^{k+1}dr\int_0^{2\pi}(\frac{e^{i\theta}+e^{-i\theta}}{2})^k d\theta$

$=\frac{2\pi}{4\pi}\cdot\frac{2^{k+2}}{k+2}\cdot\frac{1}{2^k}C_k^{\frac{1}{2}k}=\gamma_k$

直观上, x_j 为 A_n 特征值

$\sum_{j=1}^n x_j^2=\mathrm{tr}A_n^2$ $\quad \sum_{j=1}^n\mathbb{E}[x_j^2]=\sum_{i,j=1}^n\mathbb{E}[a_{ij}^2]$

$n\mathbb{E}[X^2]=n(n-1)+n\mathrm{Var}(Y)$

可知 $\mathbb{E}[X^2]=n-1+\mathrm{Var}(Y)$, $|X|\sim\sqrt{n}$

证明: $\frac{1}{n}\mathbb{E}[\mathrm{tr}(\frac{A_n}{\sqrt{n}})^k]$

$=n^{-1-\frac{1}{2}k}\sum_{i_1,\dots,i_k}^{n}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\cdots a_{i_ki_1}]$

(1) $k=1$, RHS $=\sum_{i_1}\mathbb{E}[a_{i_1i_1}]=0$

(2) $k=2$, RHS $=\sum_{i_1,i_2}\mathbb{E}[a_{i_1i_2}a_{i_2i_1}]=n\mathrm{Var}(Y)+n(n-1)$

(3) $k=3$, RHS $=\sum_{i_1,i_2,i_3}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}a_{i_3i_1}]$

$=\sum_{i_1,i_3}\mathbb{E}[a_{i_1i_1}a_{i_1i_3}^2]$: $O(n)$ $\quad i_1=i_2$

$+\sum_{i_1,i_2}\mathbb{E}[a_{i_1i_2}^2a_{i_1i_1}]$: 0 $\quad i_1\ne i_2, i_1=i_3$

$+$: 0 $\quad i_1\ne i_2, i_3=i_2$

$+\sum_{i_1,i_2,i_3}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}a_{i_3i_1}]$: 0 全不同

(4) $k=4$, RHS $=\sum_{i_1,\dots,i_4}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\cdots a_{i_4i_1}]$

四条边 $\{i_1,i_2\}$, $\{i_2,i_3\}$, $\{i_3,i_4\}$, $\{i_4,i_1\}$ 中若有1条与其他不同, 则因期望为0, 此项必为0. 每条边若出现, 则非零贡献时, 其出现次数至少为2. 分析非消失项具体形式。

若 $i_3\ne i_1$, 对 $a_{i_1i_2}a_{i_2i_3}a_{i_3i_4}a_{i_4i_1}$ 有 $a_{i_1i_2}$ 与 $a_{i_2i_3}$ 不同, 此时 $a_{i_3i_4}$ 必为两者之一, 可知 $i_4=i_2$ (否则 $i_1=i_4\ne i_3=i_2$, 此时 $a_{i_1i_2}a_{i_2i_2}a_{i_2i_1}a_{i_1i_1}$ 期望为0)

下面设 $i_3=i_1$, 非消失项典型形式

Type1: $i_3=i_1$, $i_4\ne i_2\ne i_1$,

$\sum_{i_1\ne i_2\ne i_4}\mathbb{E}[a_{i_1i_2}^2a_{i_1i_4}^2]=n(n-1)(n-2)$ 贡献!

Type2: $i_3=i_1$, $i_4=i_2\ne i_1$,

$\sum_{i_1,i_2}\mathbb{E}[a_{i_1i_2}^4]=O(n^2)$

Type3: $i_3=i_1=i_2$, $i_4\ne i_2$,

$\sum_{i_1,i_4}\mathbb{E}[a_{i_1i_1}^2a_{i_1i_4}^2]=O(n^2)$

Type4: $i_1=i_2=i_3=i_4$,

$\sum_{i_1}\mathbb{E}[a_{i_1i_1}^4]=O(n)$

当Type1, $i_4=i_2$ 但 $i_1\ne i_3\ne i_2$ 时,

$\sum_{i_1\ne i_2\ne i_3}\mathbb{E}[a_{i_1i_2}^2a_{i_2i_3}^2]=n(n-1)(n-2)$ 贡献!

(5) 一般 k

RHS $=\sum_{i_1,\dots,i_k}\mathbb{E}[a_{i_1i_2}a_{i_2i_3}\cdots a_{i_ki_1}]$

下标循环, 期望为0, 独立性, 高阶矩有限, 非消失项至多 $[\frac{k}{2}]$ 个不同边, 因此 $i_1,\cdots,i_k$ 中至多有 $1+[\frac{k}{2}]$ 个自由顶点(不相同的顶点数)

反证. 否则, 若至少有 $2+[\frac{k}{2}]$ 个不同顶点, 则除第一个外每一个自由顶点带来一条

新边，即至少有 $1+[\frac{k}{2}]$ 条不同边，矛盾！

立即知 RHS $= O(n^{1+[\frac{k}{2}]})$

进而 $k=2m+1$ 时结论成立。

当 $k=2m$ 时，RHS 求和式中主项由下面项构成：每条边恰好出现2次，且有 $(m+1)$ 个自由顶点，

这时称 $(i_1,i_2,\cdots,i_{2m})$ 为一个不相交的路径。

例：$k=6$，两个典型不相交路径

$(i_1,i_2,i_3,i_4,i_3,i_2)$，$i_1,i_2,i_3,i_4$ 互不相同

$a_{i_1i_2}^2 a_{i_2i_3}^2 a_{i_3i_4}^2$

$(i_1,i_2,i_1,i_3,i_1,i_4)$，$i_1,i_2,i_3,i_4$ 互不相同

$a_{i_1i_2}^2 a_{i_2i_3}^2 a_{i_3i_4}^2$

但是 $(i_1,i_2,i_3,i_1,i_2,i_3)$ i_1,i_2,i_3 互不相同

$a_{i_1i_2}^2 a_{i_2i_3}^2 a_{i_3i_1}^2$

计数问题：同构意义下，$i_1,i_2,\cdots,i_{2m}$ 从 $1,2,\cdots,m+1$ 中取。对不相交路径 $(i_1,i_2,\cdots,i_{2m})$，可构造一一对应：$(a_1,a_2,\cdots,a_{2m})$

定义 $a_j=\begin{cases}1, & i_j i_{j+1} \text{ 在 } i_1 i_2 \cdots i_{2m} i_1 \text{ 中首次出现} \\ -1, & \text{否则}\end{cases}$

这里先固定 $i_1=1$.

$k=6$，$i_1,i_2,i_3,i_4,i_3,i_2,i_1$

$1,1,1,-1,-1,-1$

$i_1,i_2,i_1,i_3,i_1,i_4,i_1$

$1\ -1\ 1\ -1\ 1\ -1$

CHECK：1-1对应？练习

构造一个随机游走：

令 $S_i=S_{i-1}+a_i$，$i=1,2,\cdots,2m$，$S_0=0$

则 $S_{2m}=0$，$S_i\geq 0$，$S_0=0$，$\forall i$

轨道数 Catalan数 $C_m=\frac{1}{m+1}C_{2m}^m$

参见母函数或反射原理.

总之，RHS $= n(n-1)\cdots(n-m)C_m+O(n^m)$.

Wigner半圆律偶阶矩 C_m，组合诠释

$1,2,\cdots,2m-1,2m$ 两两配对且不相交

1 2 3 4　　1 2 3 4　　1 2 3 4

Non-crossing　　相交　　不相交

N(0,1)四阶矩为3

半圆律四阶矩为2

练习：$1,2,\cdots,2m$ 不相交两两配对数为 C_m

4. Wishart矩阵模型

$X=(x_{ij})_{i=1,\cdots,p;j=1,\cdots,n}$　$\{x_{ij}\}$ iid $N(0,1)$

假设 $n-p=\alpha\geq 0$ 固定，试证

$\frac{1}{p}E[tr(\frac{1}{p}XX^T)^m]\xrightarrow{p\to+\infty}C_m$

$m=2$ 时作业