三. 连续型随机变量

1. 典型密度函数

(1) 均匀分布 $X \sim U[a,b]$ $f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$ ~~F(x)~~ $P(c \le X \le d) = \frac{d-c}{b-a}$

(2) 指数分布 $X \sim \exp(\lambda)$ $f(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & \text{otherwise} \end{cases}$ $F(x) = 1 - e^{-\lambda x}$ $x \ge 0$

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

$$LHS = \frac{P(t<X<t+\Delta t)}{P(t<X)} = \frac{F(t+\Delta t) - F(t)}{1-F(t)} = \lambda \Delta t + o(t)$$

$$\Rightarrow F'(t) = \lambda(1 - F(t))$$

解 ODE $\Rightarrow F(t) = 1 - e^{-\lambda t}$.

$P(X > t+s \mid X > t) = e^{-\lambda s} = P(X > s)$ 几何分布的连续版本

(指数分布 $\Leftrightarrow$ "永远年轻")

(3) 正态分布 $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ 记为 $N(\mu, \sigma)$

分布函数为 $\Phi(x)$

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

check $\frac{1}{2\pi\sigma^2}\int_{-2\sigma}^{2\sigma}\sqrt{4\sigma^2 - x^2}\,dx \xlongequal{x=2\sigma\sin\theta} \frac{1}{2\pi\sigma^2}\cdot(2\sigma)^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\theta\, d\theta$

$$= \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\theta\, d\theta = 1$$

$$P(0 < X < \sigma) = \frac{\sqrt{3}}{4\pi} + \frac{1}{6}.$$

$\boxed{EX\,3.2}$ 二元正态 $f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}$ $\begin{cases} -\infty < x, y < \infty \\ -1 < \rho < 1 \end{cases}$

$$f_X(x) = \int_{\mathbb{R}} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}[(y-\rho x)^2 + (1-\rho^2)x^2]}\,dy$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \int \frac{1}{\sqrt{2\pi(\sqrt{1-\rho^2})^2}} e^{-\frac{1}{2(1-\rho^2)}(y-\rho x)^2}\,dy$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \sim N(0,1).$$

$\boxed{\text{EX 3.3}}$ (X,Y) 有密度 $f(x,y)$ ~~即~~则 $Z=X+Y$ 有密度

$$\cancel{f_Z(z)=(f_X * f_Y)(z)} \quad f_Z(z)=\int_{\mathbb{R}} f(x,z-x)\,dx=\int f(z-y,y)\,dy$$

X Y 独立时变成卷积

$$P(X+Y\le z)=\iint_{x+y\le z} f(x,y)\,dx\,dy$$

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

$$\overset{y\to y-x}{=}\int_{-\infty}^{\infty}\int_{-\infty}^{z} f(x,y-x)\,dy\,dx$$

$$\overset{\text{Fubini}}{=}\int_{-\infty}^{z}\Big(\underbrace{\int_{-\infty}^{\infty} f(x,y-x)\,dx}_{\substack{\text{并且独立} \\ f_X(x)f_Y(y-x)}}\Big)\,dy$$

$\square$

$\boxed{\text{EX 3.4}}$ 对 $X\sim N(0,1)$ $Y\sim N(0,1)$

$Z=X+Y$ 的密度：
$$f_Z(z)=\frac{1}{2\pi}\int_{\mathbb{R}} e^{-\frac12 x^2-\frac12(z-x)^2}\,dx$$

$$=\frac{1}{2\pi}\int_{\mathbb{R}} e^{-\frac{1}{2(\frac12)}(x-\frac12 z)^2-\frac14 z^2}\,dx$$

$$=\frac{1}{\sqrt{2\pi}}\Big(\frac{1}{\sqrt{2\pi\frac12}}\int_{\mathbb{R}}\ \ dx\Big)e^{-\frac14 z^2}$$

$$=\frac{1}{2\sqrt{\pi}}e^{-\frac14 z^2}$$

$\underline{Z\sim N(0,2)}$ $\square$

2、期望与条件期望

$\boxed{\text{Def 3.1}}$ X 有密度 f，若 $\int_{\mathbb{R}}|x|f<\infty$，称 $\mathbb{E}[X]=\int_{-\infty}^{\infty}x f(x)$ 为 X 的数学期望.

$\boxed{\text{Lem 3.2}}$ 连续型 r.v. X 的期望存在，则

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

proof:
$$\mathbb{E}[X]=\int_0^{\infty}t f(t)\,dt+\int_{-\infty}^{0}t f(t)\,dt$$

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

$$0\le x\le t\le\infty \quad \overset{\text{Fubini}}{=} \cancel{\int_0^{\infty}\int_0^{t} f(t)\,dx\,dt}$$

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

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

$\square$

26

Thm 3.3 设 $g \in \mathcal{B}$ X, & $g(X)$ 均为连续型 且 $g(X)$ 期望存在

则 $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$.

proof $\mathbb{E}[g(X)] = \mathbb{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy$

$= \cancel{\int_{-\infty}^{\infty} y P(} \int_0^{\infty} (1 - F_Y(y)) dy - \int_{-\infty}^{0} F_Y(y) dy$

$= \int_0^{\infty} P(g(X) > y) dy - \int_{-\infty}^{0} P(g(X) \le y) dy$

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

$= \int_{\{g(x) > 0\}} \left(\int_0^{g(x)} f(x) dy \right) dx - \int_{\{g(x) \le 0\}} \left(\int_{g(x)}^{0} f(x) dy \right) dx$

$= \int_{\{g(x) > 0\}} g(x) f(x) dx + \int_{\{g(x) \le 0\}} g(x) f(x) dx$

$= \int_{-\infty}^{\infty} g(x) f(x) dx$ □

类似地

Thm 3.4 X, Y 有联合密度 $f(x,y)$, $g(X,Y)$ 为连续型随机变量且期望存在, 则

$\mathbb{E}[g(X,Y)] = \iint g(x,y) f(x,y) dx dy$

特别地 $g(x,y) = ax + by$ 时 $\Rightarrow \mathbb{E}[aX+bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

这样后面能定义更多内容

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

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

(3) 协方差 $Cov(X_i, X_j) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$ 内积!

(4) 相关系数 $\rho = \dfrac{Cov(X,Y)}{\sqrt{Var(X) Var(Y)}}$ $|\rho| \le 1$ By C-S inequality

Thm 3.5 (Cauchy - Schwarz 不等式) $(\mathbb{E}[XY])^2 \le \mathbb{E}[X^2]\mathbb{E}[Y^2]$ ...~~subtle~~

"=" 成立 $\Leftrightarrow$ a, b 不全为0, $P(aX = bY) = 1$ 但 (X,Y) 并非联合连续型.

EX 3.5 $X \sim N(\mu, \sigma^2)$ $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ $e^{-\frac{y^2}{2\sigma^2}}$, $-\frac{y}{\sigma^2}$

$\mathbb{E}[X] = \underbrace{\int (x-\mu) f(x) dx}_{奇} + \mu = \mu$

$Var(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} y^2 e^{-\frac{1}{2\sigma^2} y^2} dy$ $\cancel{= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \frac{-2\sigma^2 y}{2} d e^{-\frac{y^2}{2\sigma^2}}}$ $= \frac{\sigma^2}{\sqrt{2\pi}} \int y^2 e^{-\frac{1}{2}y^2} dy$

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

$\boxed{\text{EX 3.6}}$ 二元正态

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

$$\Rightarrow f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\qquad f_Y(y)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$$

$$\begin{aligned}Cov(X,Y)&=\iint xy\, f(x,y)\,dxdy\\&=\iint (x(y-\rho x)+\rho x^2)\,f(x,y)\,dxdy\\&=\iint x(y-\rho x)\,f(x,y)\,dxdy+\int \rho x^2 f(x,y)\,dxdy\\&=\rho\end{aligned}$$

这里 $\rho=0 \Leftrightarrow X, Y$ 独立.

相关系数

条件期望 $P(Y\le y \mid x<X<x+\Delta x)\sim\frac{\int_{-\infty}^{y}f(u,v)\Delta x\,dv}{f_X(x)\Delta x}$

$$=\int_{-\infty}^{y}\frac{f(x,v)}{f_X(x)}dv\qquad f_X(x)>0.$$

$\boxed{\text{Def 3.6}}$ "连续型" $f_X(x)>0$ 下 Y 的条件密度为

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}$$

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

条件期望 $\mathbb{E}[Y|X]$: $\mathbb{E}[Y|X=x]=\int y\, f_{Y|X}(y|x)\,dy$.

其实是个随机变量.

$\boxed{\text{Thm 3.7}}$ $\mathbb{E}[\mathbb{E}[Y|X]]=\mathbb{E}[Y]$

$$\begin{aligned}LHS&=\int \mathbb{E}[Y|X=x]\,f_X(x)\,dx\\&=\int\left(\int y f_{Y|X}(y|x)\,dy\right)f_X(x)\,dx\\&=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}y\frac{f(x,y)}{f_X(x)}dy\right)f_X(x)\,dx\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}y\, f(x,y)\,dy\,dx\\&=\int_{-\infty}^{\infty}y\, f_Y(y)\,dy=\mathbb{E}[Y].=RHS\qquad\square\end{aligned}$$

$\boxed{EX\ 3.7}$ 二元正态

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

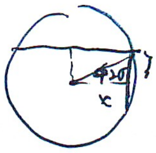
$$= \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^2)}(y-\rho x)^2} \sim N(\rho x, 1-\rho^2)$$

$\mathbb{E}[Y|X] = \rho x$. $\mathbb{E}[Y|X] = \rho X$. $\square$

$\boxed{EX\ 3.8}$ (X,Y) 在 $D = \{x^2+y^2 \le 4\sigma^2\}$ 上均匀分布

$$f(x,y) = \begin{cases} \frac{1}{4\pi\sigma^2} & (x,y)\in D \\ 0 & \text{otherwise.} \end{cases}$$

$$f_X(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2} \quad |x| \le 2\sigma$$



可算 $Cov(X,Y)=0$

但 $f(x,y) \neq f_X(x) f_Y(y)$

$X \quad f$

$\mathbb{E}[X] = \int x f(x)\,dx$

矩 $\mathbb{E}[X^k] = \int x^k f(x)\,dx$

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

$Cov(X,Y) = \iint (x-\mu_x)(y-\mu_y) f(x,y)\,dx\,dy$

条件密度

$f_{Y|X}(y|x) = \dfrac{f(x,y)}{f_X(x)} \quad (f_X(x) > 0)$ 本身也是密度.

条件期望 $\mathbb{E}[Y|X]$: $\mathbb{E}[Y|X=x] = \int y f_{Y|X}(y|x)\,dy$

$\Downarrow$

r.v. $\psi(X)$ 称为 Y 关于 X 的条件期望.

$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$ 全期望公式

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

可以与 Cov 联系

正态分布 $X \sim N(\mu, \sigma^2)$　　$f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$

二元正态　ρ 为相关系数

3. 多元正态分布

一元　$\phi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

二元　$f(x_1, x_2) = \dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x_1,x_2)} = \dfrac{1}{2\pi(1-\rho^2)} e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)}$

$Q(x_1,x_2) = \left[\left(\dfrac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\dfrac{x_1-\mu_1}{\sigma_1}\cdot\dfrac{x-\mu_2}{\sigma_2} + \left(\dfrac{x-\mu_2}{\sigma_2}\right)^2\right]$

$\boxed{\text{Def 3.8}}$ (多元正态分布) 若 $X = (X_1, \cdots X_n)$ 密度为

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \exp\left\{-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T\right\}$$ ，Σ 正定.

则称 X 服从 n 元正态分布.　$\vec{X} \sim N(\vec{\mu}, \Sigma)$

$\boxed{\text{Rmk}}$ 二元标准正态 $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$　$\Sigma^{-1} = \dfrac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$

$\boxed{\text{Thm 3.9}}$ $X \sim N(\vec{\mu}, \Sigma)$，则

(1) $\mathbb{E}[X] = \vec{\mu}$ $\quad \mathbb{E}[X_i] = \mu_i$

(2) Σ 为协方差矩阵，$\sigma_{ij} = Cov(X_i, X_j)$

proof (Idea：尝试变量分离)．取正交阵 B.

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

$\Sigma^{-1} = B^T \Lambda^{-1} B \qquad$ 令 $y = (x-\mu)B^T \Rightarrow x = \vec{\mu} + yB \quad dx = \overset{|\det B|}{dy}$

$$\int f(\vec{x}) d\vec{x} = \int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \cdot \exp\left(-\frac{1}{2} y \overset{-1}{\Lambda} \vec{y}^T\right) dx$$

$$= \int_{\mathbb{R}^n} \frac{1}{\sqrt{(2\pi)^n \det\Sigma}} \exp\left(-\frac{1}{2} y \overset{-1}{\Lambda} y^T\right) |\det B|\, dy$$

$$= \prod_{i=1}^n \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi \cdot \lambda_i}} e^{-\frac{1}{2}\frac{y^2}{\lambda_i}} dy$$

$$= 1.$$

$$\mathbb{E}[X_i] = \int_{\mathbb{R}^n} x_i f(\vec{x}) d\vec{x} = \int \underbrace{(\mu_i + \sum_{j=1}^n y_i b_{ji})}_{\text{奇}} \frac{1}{\sqrt{(2\pi)^n |\Lambda|}} e^{-\sum_i \frac{1}{2\lambda_i} y_i^2} dy$$

$$= \mu_i$$

(ii) $$Cov(X_i, X_j) = \int_{\mathbb{R}^n} (x_i - \mu_i)(x_j - \mu_j) f(\vec{x}) d\vec{x}$$

$$= \int_{\mathbb{R}^n} \left(\sum_k y_k b_{ki}\right)\left(\sum_l y_l b_{lj}\right) f(\vec{y})\, dy$$

$$= \sum_k \sum_l b_{ki} b_{lj} \int y_l y_k$$

$$= \sum_{k=l} b_{ki} b_{kj} \int y_k^2 \frac{1}{\sqrt{2\pi \lambda_k}} e^{-\frac{1}{2}\frac{y_k^2}{\lambda_k}} dy_k$$

$$= \sum_{k=1}^n b_{ki} b_{kj} \lambda_k$$

$$= (B^T \Lambda B)_{ij} = \sigma_{ij} \qquad \square$$

$\boxed{\text{Thm 3.10}}$ $X \sim N(\vec{\mu}, \Sigma)$，若 $\operatorname{rank} A = m \le n \quad A \in \mathbb{F}^{n\times m}$ $\quad Y = XA \sim N(\mu A, A^T \Sigma A)$

$\boxed{\text{Thm 3.11}}$ (特殊情形)．A 为 n 阶方阵，$\det A \neq 0 \quad \vec{X}A \sim N(\vec{\mu}A, A^T \Sigma A)$

proof 令 $Y = XA$．求 Y 分布函数．记 $B = [a_1, b_1] \times \cdots [a_n, b_n]$.

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

$$P(Y\in B) = P(X\in D) = \int_D \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^T\right) dx$$

$Y = XA \quad X = YA^{-1}$

$dx = |\det A^{-1}| \cdot dy$

$$= \int_B \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} \exp\left(-\frac{1}{2}(YA^{-1}-\mu)\Sigma^{-1}(YA^{-1}-\mu)^T\right) |\det A^{-1}|\, dy$$

$$= \int_B \frac{1}{\sqrt{(2\pi)^n |A^T\Sigma A^{-1}|}} \exp\left(-\frac{1}{2}(Y-\mu A)(A^T\Sigma A)^{-1}(Y-\mu A)^T\right) dy$$

$\Rightarrow Y \sim N(\mu A, A^T\Sigma A)$. $\square$

复习. 三元组 $(\Omega, \mathcal{F}, P)$

样本空间　事件域　概率测度

事件域：σ-代数 { φ；补；可数并 }

概率测度 { 非负；规范；可加 }

事件独立　$P(AB) = P(A)P(B)$

随机变量　$X: \Omega \to (\mathbb{R}, \mathcal{B})$　$\forall x\in\mathbb{R}$, $\{X\le x\}\in\mathcal{F}$

$\vec{X}$ 独立　$P(X_1\le x_1, \cdots X_n\le x_n) = \prod_{i=1}^n P(X_i\le x_i)$

$\mathbb{E}[X] = \sum x f(x)$　$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \ge 0 \Rightarrow \mathbb{E}[X^2] \ge (\mathbb{E}[X])^2$　$Y=1$

推广(仅有用)　$\mathbb{E}[X^2]\cdot\mathbb{E}[Y^2] \ge \cancel{\mathbb{E}[X]\cdot\mathbb{E}[Y]}\ (\mathbb{E}[XY])^2$

计算 $P(A) = \underbrace{\sum_i P(B_i)P(A|B_i)}_{\text{全概率}}$　B_i 是 Ω 的划分. $P(B_i)>0$

$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$　全期望公式

$= \sum_x \mathbb{E}[Y|X=x]P(X=x)$.

$\mathbb{E}[Y|X=x] = \sum_y y\, f_{Y|X}(y|x)$　　$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$

概率模型：

概率方法：

随机游走. $S_n = X_1 + \cdots + X_n$　　母函数（离散 非负整数）$G(z) = \mathbb{E}[z^X]$

多项式 & Taylor 级数

$\boxed{\text{Thm 3.12}}$ $X \sim N(\mu, \Sigma)$ $\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$

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

$\boxed{\text{Thm 3.13}}$ $X \sim N(\mu, \Sigma)$ $X = (X^{(1)}, X^{(2)})$ $\mu = (\mu^{(1)}, \mu^{(2)})$

$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \Rightarrow X^{(1)} \sim N(\mu^{(1)}, \Sigma_{11})$

$\begin{pmatrix} I & \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix} \Sigma \begin{pmatrix} I & -\Sigma_{11}^{-1}\Sigma_{12} \\ & I \end{pmatrix} = D^T \Sigma D$

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

$Y \sim N(\mu D, D^T \Sigma D) = N\left(\mu D, \begin{matrix} \Sigma_{11} & \\ & * \end{matrix}\right)$

$\Rightarrow X^{(1)} \sim N(\mu^{(1)}, \Sigma_{11})$.

~~Thm 3.14~~

$\boxed{\text{Thm 3.10 's proof}}$ $m<n$ 时 $D = A_{n\times m}, B_{n\times(n-m)}$ B 使其非奇异

$(XA \;\; XB) = XD \sim N(\mu D, D^T \Sigma D)$

$D^T \Sigma D = \begin{pmatrix} A^T \Sigma A & \cdot \\ \cdot & \cdot \end{pmatrix}$

$XA \sim N(\mu A, A^T \Sigma A)$. $\square$

$\boxed{\text{Thm 3.14}}$ $X \sim N(\mu, \Sigma)$. $X = (X_1, \cdots X_n)$ 相互独立 $\Leftrightarrow$ Σ 对角

$\boxed{\text{EX 3.9}}$ $X, Y \sim N(0,1)$ 且独立.

令 $\begin{cases} X = R\cos\Theta \\ Y = R\cos\Theta \end{cases}$ $R \geq 0$, $0 \leq \Theta < 2\pi$

求 R, Θ 的联合密度

$\Rightarrow f(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}x^2 - \frac{1}{2}y^2}$

$f_{R,\Theta}(r,\theta) = \frac{1}{2\pi} r e^{-\frac{1}{2}r^2}$ $\Rightarrow$ $\Theta \sim U[0, 2\pi]$

↑

雅可比

~~$R \sim N$~~

$f_R(r) = r e^{-\frac{1}{2}r^2}$ $\square$

副产品 令 $X = \sqrt{-2\log U_1} \cos 2\pi U_2$ $\Rightarrow$ X, Y 独立. $X, Y \sim N(0,1)$

$Y = \sqrt{-2\log U_1} \sin 2\pi U_2$ $R = \sqrt{X^2+Y^2}$ $f_R(r) = r e^{-\frac{1}{2}r^2}$

check!

$\boxed{\text{EX 3.10}}$ X, Y 独立 $\sim N(0,1)$

$\begin{cases} U = \sigma_1 X \\ V = \sigma_2 \rho X + \sigma_2\sqrt{1-\rho^2}\, Y \end{cases}$ 代入有 $f(u,v) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(u,v)}$ check!

$\Rightarrow \mathbb{E}[UV] = \sigma_1\sigma_2\rho$ $\mathbb{E}[V|U] = \frac{\sigma_2}{\sigma_1}\rho U$.

4、Wick公式与GOE模型

$\boxed{\text{Thm 3.13}}$ (Wick公式) $(X_1, \cdots X_n) \sim N(0, \Sigma)$

则 $\mathbb{E}[X_1 \cdots X_n] = \sum_{p \in P_2(n)} \prod_{(i,j) \in p} \mathbb{E}[X_i X_j]$

$P_2(n)$ 为 $\{1, \cdots n\}$ 所有可能的两两配对 $\prod_{(i,j)\in p}$ 表示对每个配对的乘积

$\boxed{\text{Idea}}$ 用 $\mathbb{E}[e^{\theta X}]$ 算 $\mathbb{E}[X^{2k}]$

展开 $\mathbb{E}\left[\sum_{n=0}^{\infty} \frac{(\theta X)^n}{n!}\right]$ 视 $\theta \in \mathbb{R}$.

proof $2k+1$ 时结论为 0

$2k = 2n$ 为偶数时

先算 $\mathbb{E}[e^{\lambda_1 X_1 + \cdots + \lambda_n X_n}]$

$$X \sim N(0, \Sigma) \qquad X\begin{pmatrix}\lambda_1 \\ \vdots \\ \lambda_n\end{pmatrix} = XD \sim N(\mu D, D^T \Sigma D)$$

$$D^T\Sigma D = \sigma^2 = \sum_{ij} \lambda_i \lambda_j \sigma_{ij}$$

$\Rightarrow \mathbb{E}[e^{\tilde{X}}]$ $\tilde{X} = XD$ $D \to 0 \Rightarrow$ ~~$\mathbb{E}[e$~~

$M(\tilde{X}) = 1$

$M(\lambda) = e^{\frac{1}{2}\sigma^2}$

$$\frac{\partial^n}{\partial \lambda_1 \cdots \partial \lambda_n} M(\lambda) \Big|_{\lambda=0} = \mathbb{E}[X_1 \cdots X_n]$$

~~$M(\lambda) = \sum_{n=0}^{\infty}$~~

$$\sigma^2 = Var(\lambda_1 X_1 + \cdots + \lambda_n X_n)$$
$$= \mathbb{E}[(\lambda_1 X_1 + \cdots + \lambda_n X_n)^2] - \left(\sum \lambda_i \mathbb{E}[X_i]\right)^2$$
$$= \sum_{ij} \lambda_i \lambda_j \mathbb{E}[X_i X_j]$$

因为 $M(\lambda) = \mathbb{E}\left[\sum_k \frac{(XD)^k}{k!}\right]$

$$= \sum_{k=0}^{\infty} \frac{1}{(2k)!} \sigma^{2k} (2k-1)!!$$

算

复随机变量. $Z = X + iY$ $\quad \mathbb{E}[Z] = \mathbb{E}[X] + i\,\mathbb{E}[Y]$

可证. 独立. $\mathbb{E}[Z_1 Z_2] = \mathbb{E}[Z_1]\,\mathbb{E}[Z_2]$

复正态 $Z \sim N_{\mathbb{C}}(\mu, \sigma^2)$ $\quad \mu \in \mathbb{C}$ $\quad \sigma^2 > 0$

$$f(z) = \frac{1}{\pi\sigma^2} e^{-\frac{1}{\sigma^2}|z-\mu|^2} \quad z \in \mathbb{C}$$

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

$\boxed{\text{Thm 3.16}}$ $X_j = \sum_i a_{ij} Z_i$ $\quad Z_i$ 独立同分布 $N(0,1)$

$$\mathbb{E}[X_1 \cdots X_n] = \sum_{p \in P_2(n)} \prod_{(i,j) \in p} \mathbb{E}[X_i X_j].$$

$\boxed{\text{Def 3.17}}$ n 阶实对称 H 为高斯正交阵. 若矩阵元联合密度函数

$$f(H) = 2^{-\frac{n}{2}} (2\pi)^{-\frac{1}{4}n(n+1)} e^{-\frac{1}{4}\operatorname{tr}(H^2)}$$

四、大数定律

引入 $dF = \begin{cases} f(x) & \text{discrete} \\ f(x)dx & \text{cts} \end{cases}$

$\mathbb{E}[X] := \int x \, dF(x)$.

自然有

$\boxed{\text{Thm 4.1}}$ $\mathbb{E}[g(X)] = \int g(x)\,dF(x)$ where g has fine properties.

统一记号 $\mathbb{E}[X] = \int_\Omega X dP$

以某分布的标准操作，可构造 $\mu_F((a,b]) = F(b) - F(a)$，扩张为概率测度

proof by integral three steps

$\boxed{\text{Thm 4.2}}$ 期望算子的性质

(1) 非负，若 $X \geq 0$ $\mathbb{E}[X] \geq 0$

(2) 归1 $\mathbb{E}[1] = 1$

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

$\boxed{\text{Thm 4.3}}$ 设 $X_n(\omega) \to X(\omega)$ for $\forall \omega \in \Omega$ or $\omega \in \Omega \setminus \Omega_0$, $P(\Omega_0) = 0$.

(1) MCT $X_{n+1} \geq X_n \geq 0$, $\mathbb{E}[X_n] \to \mathbb{E}[X]$

(2) DCT $|X_n| \leq Y$ $\mathbb{E}[Y] < \infty$, $\mathbb{E}[X_n] \to \mathbb{E}[X]$

$\boxed{\text{Thm 4.4}}$ (Fatou) $X_n \geq 0$ a.s. $\mathbb{E}[\liminf X_n] \leq \liminf \mathbb{E}[X_n]$.

almost surely.

$\boxed{\text{Thm 4.5}}$ X, Y 相互独立. $\mathbb{E}|X|, \mathbb{E}[|Y|] < \infty$ 则 ~~$\mathbb{E}[|XY|] \mathbb{E}[|X|]\mathbb{E}[|Y|]$~~ $\mathbb{E}[|XY|] < \infty$ ① & ② $\mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[XY]$.

proof STEP 1. $X = \sum_i^n x_i I_{A_i}$ $Y = \sum_j^m y_j I_{B_j}$

$XY = \sum_{ij} x_i y_j I_{A_i} I_{B_j}$

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

STEP 2. $X_n \nearrow X$, $Y_n \nearrow Y$ $\Rightarrow X_n Y_n \nearrow XY$

$\mathbb{E}[XY] \overset{MCT}{=} \lim \mathbb{E}[X_n Y_n] = \lim \mathbb{E}[X_n]\mathbb{E}[Y_n] = \mathbb{E}[X]\mathbb{E}[Y]$.

STEP 3. $X = X^+ - X^-$ $Y = Y^+ - Y^-$ $\square$

记 $C_b(\mathbb{R})$ 为 $\mathbb{R}$ 上有界连续函数全体 ← 测试函数

$\boxed{\text{Thm 4.6}}$ X, Y 同分布 $\Leftrightarrow \mathbb{E}[g(X)] = \mathbb{E}[g(Y)]$ $\forall g \in C_b$

proof ($\Rightarrow$) ✓

($\Leftarrow$) 令 $g = \begin{cases} 1 & x \leq z \\ 1 - \frac{1}{\varepsilon}(x-z) & z < x \leq z+\varepsilon \\ 0 & x > z+\varepsilon \end{cases}$ $F_X(z) = \mathbb{E}[I_{(-\infty, z]}(X)] \leq \mathbb{E}[g(X)]$

$F_Y(z+\varepsilon) = \mathbb{E}[I_{(-\infty, z+\varepsilon]}(Y)] \geq \mathbb{E}[g(Y)]$ （$\mathbb{E}[g(X)] = \mathbb{E}[g(Y)]$）

令 $\varepsilon \to 0^+$ $F_Y(z) \geq F_X(z)$. $\square$

四种收敛

$\boxed{\text{Def 4.7}}$ X_n, X 是 $(\Omega, \mathscr{F}, P)$ 上的随机变量

(1) 依分布收敛 $\underbrace{F(x)=P(X\le x)}_{\text{的连续点}}$ $P(X_n\le x)\to P(X\le x)$ $n\to\infty$

记 $X_n \xrightarrow{D} X$ 和分析中不太一样.

$(F_{X_n}\to F_X)$

(2) L^p 收敛 $\mathbb{E}[|X_n|^p]<\infty$, $\mathbb{E}[|X|^p]<\infty$ $\forall n$

$\mathbb{E}[|X_n-X|^p]\to 0$ as $n\to\infty$

记 $X_n \xrightarrow{L^p} X$

(3) 依概率收敛 (测度) $\forall \varepsilon>0$.

记 $X_n \xrightarrow{P} X$. $P(|X_n-X|>\varepsilon)\to 0$ as $n\to\infty$

(4) 几乎处处收敛

$P(\{X_n\to X \text{ as } n\to\infty\})=1$.

记 $X_n \xrightarrow{a.s.} X$ a.s. = almost surely.

$\boxed{\text{Rmk}}$ 依分布收敛与样本空间选取无关，下面的例子表明依分布收敛是不能比另外几个强的.

$\boxed{\text{EX 4.1}}$ 设 $P(X=1)=P(X=0)=\frac{1}{2}$, $X_n=X$, $\forall n\ge 1$. 令 $Y=1-X$ 则

$$X_n\xrightarrow{D}X, \quad X_n\to Y$$

但 $|X_n-Y|=|2X-1|=1 \Rightarrow X_n \not\to Y$ $\underline{\text{a.s.}, L^p, P}$ □

$\boxed{\text{Thm 4.8}}$

(1)
$$\begin{array}{ccc} X_n\xrightarrow{a.s.}X & & \\ & \searrow & \\ & & X_n\xrightarrow{P}X \Rightarrow X_n\xrightarrow{D}X \\ & \nearrow & \\ X_n\xrightarrow{L^p}X & & \end{array}$$

(2) $r>s\ge 1$

$L^r\Rightarrow L^s$

$\boxed{\text{Lem 4.9}}$ $X_n\xrightarrow{P}X \Rightarrow X_n\xrightarrow{D}X$

令 $F_n(x)=P(X_n\le x)$ $F(x)=P(X\le x)$

$F_n(x)=P(X_n\le x, X\le x+\varepsilon)+P(X_n\le x, X>x+\varepsilon)$

$\le F(x+\varepsilon) + P(|X-X_n|>\varepsilon) \Rightarrow F(x-\varepsilon)+P\cdots \le F_n(x)\le F(x+\varepsilon)+P\cdots$

对称为 $F(x-\varepsilon)\le F_n(x)+P(|X-X_n|>\varepsilon)$

取极限 $\Rightarrow F(x-\varepsilon)\le \varliminf_n F_n(x)\le \varlimsup_n F_n(x)\le F(x+\varepsilon)$ $\varepsilon\to 0$

$\Rightarrow \lim_n F_n(x)=F(x)$ 在 F 的连续点处

$\boxed{\text{Lem 4.10}}$ Markov $\forall p>0$ $\forall a>0$

$$P(|X|\geq a) \leq \frac{1}{a^p}\mathbb{E}[|X|^p]$$

证明: $\mathbb{E}[|X|^p] = \int |X|^p \, dP \geq \int_{|X|>a} |X|^p \, dP \geq a^p P(|X|>a)$. □

尾概率被矩控制

$\boxed{\text{Rmk}}$ (Chebyshev 不等式)

$p=2$ 时 $P(|X-\mathbb{E}[X]| \geq a) \leq \frac{1}{a^2}\mathbb{E}[(X-\mathbb{E}[X])^2]$

$= \frac{1}{a^2}\mathrm{Var}(X)$

$\boxed{\text{Lem 4.11}}$ $p,q>1$. $\frac{1}{p}+\frac{1}{q}=1 \Rightarrow \mathbb{E}[|XY|] \leq (\mathbb{E}[|X|^p])^{\frac{1}{p}}(\mathbb{E}[|Y|^q])^{\frac{1}{q}}$

(1) Hölder

$\Downarrow$ def $\|X\|_p := (\mathbb{E}[|X|^p])^{\frac{1}{p}}$ $\quad$ i.e. $\|XY\|_1 \leq \|X\|_p \|Y\|_q$

(2) Minkowski $p\geq 1$ $\|X+Y\|_p \leq \|X\|_p + \|Y\|_p$.

(3) Lyapunov $\|X\|_p \geq \|X\|_q$ $(p\geq q\geq 1)$

proof of $\boxed{4.8}$: $L^p \Rightarrow P$ $\quad$ Markov

$L^r \Rightarrow L^s$ $\quad$ Lyapunov $\quad$ □

有时,加条件能 $D \Rightarrow P$

$\boxed{\text{Thm 4.12}}$ 若对常数 c 有 $X_n \xrightarrow{D} c$,则 $X_n \xrightarrow{P} c$

Proof: $P(|X_n - c| > \varepsilon) = P(X_n > c+\varepsilon) + P(X_n < c-\varepsilon)$

$= 1 - \underset{1}{P(X_n \leq c+\varepsilon)} + P(X_n < c-\varepsilon)$

$\leq P(X_n \leq c-\varepsilon/2)$ $\downarrow 0$ □

$\longrightarrow 0$.

$\boxed{\text{EX 4.2}}$ 若 $r>0$. $\mathbb{E}[|X|^r]=0 \Rightarrow P(X=0)=1$.

proof. $\forall \varepsilon>0$, $P(|X|>\varepsilon) \leq \frac{1}{\varepsilon^r}\mathbb{E}[|X|^r] = 0$

$P(|X|\leq\varepsilon) = 1$ $\quad \varepsilon\to 0^+$ 由分布函数的右连续性. □

$\boxed{\text{Thm 4.13}}$ (1) $X_n \xrightarrow{□} X$, $X_n \xrightarrow{□} Y \Rightarrow P(X=Y)=1$ $\quad$ □ = a.s. P, L^p

(2) $X_n \xrightarrow{□} X$. $Y_n \xrightarrow{□} Y \Rightarrow X_n+Y_n \xrightarrow{□} X_n+Y_n$

(3) □ = D 时 一般不对

proof: (1) $\|X-Y\|_p \leq \|X_n - X\|_p + \|X_n - Y\|_p \longrightarrow 0$

$P(X-Y=0)=1$ $\quad$ 依概率类似

(2) "概率三角不等式" $\quad$ (3) $X = U(1,-1)$ $\quad Y_n = -X$.

一些注记. P阶矩收敛 似乎仅要求 $p>0$. 因为和 L^p 空间关心的内容不同

Rmk 依分布收敛仅要求在关于 $F_X(x)$ 的连续点处收敛.

不等式中比较重要的是 Markov 不等式 $P(|X|>a) \le \frac{\mathbb{E}[|X|^p]}{a^p}$.

本节考虑 a.s收敛和 P收敛. (其实有 a.s $\xrightarrow{\text{egoroff}}$ a.un $\Rightarrow$ P)

由分析知识. 熟知 $\{w\in\Omega : \lim_{n\to\infty} X_n(w) = X(w)\} = \bigcap_{k=1}^{\infty}\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{w\in\Omega : |X_n(w)-X(w)|\le\frac{1}{k}\}$

$$X_n \xrightarrow{a.s} X \iff P\Big(\bigcup_{k=1}^{\infty}\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n(w)-X(w)|>\tfrac{1}{k}\}\Big)=0$$

$$\iff P\Big(\bigcap_{k=1}^{\infty}\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{|X_n(w)-X(w)|\le\tfrac{1}{k}\}\Big)=1$$

$\boxed{\text{lem 4.14}}$ (1) $X_n \xrightarrow{a.s} X \iff \forall\varepsilon>0.\ P\Big(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\{|X_n-X|>\varepsilon\}\Big)=0$

$$\iff \forall\varepsilon>0\ \lim_{m\to\infty} P\Big(\bigcup_{n=m}^{\infty}\{|X_n-X|>\varepsilon\}\Big)=0$$

(2) $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X$.

proof (1). 第1个 $\iff$ 是显然的. 第2个 $\iff$ 只要看到 $\bigcup_{n=m}^{\infty}$ $\downarrow$, 由测度的连续性.

(2) 由第2个 $\iff$, 左边去掉 $\bigcup$ 即可. □

$\boxed{\text{EX 4.3}}$ 这个例子说明 以概率1收敛严格弱于 a.s.

$(\Omega=[0,1), \mathcal{L}, m)$ 随机变量 $h_k^{(i)}(x) = \chi_{[\frac{i-1}{k},\frac{i}{k})}$ $\quad i=1,\cdots k$, $k=1,\cdots$

这个例子是经典的. 有 $X_n \xrightarrow{P} 0$ (这里 $X_1=h_1^{(1)}$ $X_2=h_2^{(1)}$. $X_3=h_2^{(2)},\cdots$)

但 $X_n \stackrel{a.s}{\not\to} 0$. 这个例子也有 $X_n \xrightarrow{L^p} 0$ □

Recall 分析知识 $\limsup_{n\to\infty} A_n = \bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty} A_n$ 上限集 无穷多次

$\liminf_{n\to\infty} A_n = \bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty} A_n$ 下限集. 有限多次外.

i.o. = infinitely often $\quad \{A_n\ \text{i.o.}\} = \limsup_{n\to\infty} A_n$

$\boxed{\text{Thm 4.15}}$ (Borel - Catelli) (1) 当 $\sum_{n=1}^{\infty} P(A_n) < \infty$ 时. 有 $P(A_n\ \text{i.o.})=0$

(2) 当 $A_1,\cdots A_n,\cdots$ 相互独立时. $\sum_{n=1}^{\infty} P(A_n)=\infty \iff P(A_n\ \text{i.o.})=1$

proof. (1) $P(A_n\ \text{i.o.}) \le \sum_{n=m}^{\infty} P(A_n) \to 0$ as $m\to 0$

(2) ($\Leftarrow$) 由(1)知

($\Rightarrow$) $A^c = \bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty} A_m^c$

$$P\Big(\bigcap_{m=n}^{\infty} A_m^c\Big) = \lim_{r\to\infty} P\Big(\bigcap_{m=n}^{r} A_m^c\Big) = \lim_{r\to\infty}\prod_{m=n}^{r}(1-P(A_m)) \le e^{-\sum_{m=n}^{r} P(A_m)} = 0$$

□

Borel-Catelli 引理能用来处理 a.s. 收敛？

$\boxed{\text{EX4.4}}$ 设 $\{X_n\}$ 相互独立且服从参数为 1 的指数分布，证明.

$$P\left(\limsup_{n\to\infty}\frac{X_n}{\log n}=1\right)=1$$

$P(X\geq x)=e^{-x}$ 设 $A_n=\left\{\frac{X_n}{\log n}\geq 1+a\right\}$ $|a|<1$

A_n 相互独立. $P(A_n)=e^{-(\log n)(1+a)}=\frac{1}{n^{1+a}}$

(1) $-1<a\leq 0$ $\Rightarrow \sum_n P(A_n)=\infty$

By B-C 引理. $P(A_n \ i.o)=1$

$$\Rightarrow P\left(\limsup\frac{X_n}{\log n}\geq 1\right)=1$$

(2) 当 $0<a<1$ $\sum P(A_n)<\infty \Rightarrow P(A_n\ i.o)=0$

$$\Rightarrow P\left(\limsup\frac{X_n}{\log n}\leq 1\right)=1 \qquad \square$$

$\boxed{\text{EX4.5}}$ 非负 r.v. 列 $\{X_n\}$ 与 X 同分布. $\mathbb{E}[X]<\infty$. 截尾

$$Y_n=X_n I_{X_n<n}$$

证明 $\frac{1}{a_n}\sum_{i=1}^{n}(X_i-Y_i)\xrightarrow{a.s.}0$

积分面积转置原理.

$\sum_n P(X_n\neq Y_n)=\sum_n P(X_n\geq n)=\sum_n P(X\geq n)\leq \mathbb{E}[X]<\infty$

$\Rightarrow P(X_n\neq Y_n\ i.o)=0 \Rightarrow \frac{1}{a_n}\sum_{i=1}^{n}(X_i-Y_i)\xrightarrow{a.s.}0$

几乎必然只发生有限次.

承认如下 Skorokhod 表示定理.

$X_n\xrightarrow{D}X$ $\exists\ Y_n, Y$ s.t. (1) Y_n, Y 与 X_n, X 同分布
(2) $Y_n\xrightarrow{a.s.}Y$

那么有

$\boxed{\text{Thm 4.16}}$ $X_n\xrightarrow{D}X \Leftrightarrow \forall g\in C_b(\mathbb{R})$. $\mathbb{E}[g(X_n)]\to\mathbb{E}[g(X)]$

(与之前"同分布 $\Leftrightarrow$ 测试函数期望相同).

Proof $(\Rightarrow)$ 表示定理. 只要证 $\mathbb{E}[g(Y_n)]\to\mathbb{E}[g(Y)]$. 而 $Y_n\xrightarrow{a.s.}Y\Rightarrow g(Y_n)\xrightarrow{a.s.}g(Y)$

用控制收敛定理

$(\Leftarrow)$ 用之前的想法造测试函数. $g_{x,\varepsilon}(y)=\begin{cases}1 & y\leq x\\ -\frac{y+x}{\varepsilon} & x\leq y\leq x+\varepsilon\\ 0 & \cdots\end{cases}$

$P(X_n\leq x)=\mathbb{E}[I_{\{X_n\leq x\}}]\leq \mathbb{E}[g_{x,\varepsilon}(X_n)]$

连续点

$P(X\leq x)$

$\overline{\lim}P(X_n\leq x)\leq \mathbb{E}[g_{x,\varepsilon}(X)]\leq P(X\leq x+\varepsilon)$ 再造下界

$\Rightarrow P(X\leq x-\varepsilon)\leq \underline{\lim}P(X_n\leq x)\leq\overline{\lim}P(X_n\leq x)\leq P(X\leq x+\varepsilon)$. $\varepsilon\to 0$ $\square$

40.

Borel - Catelli：$\begin{cases} P(A_n \ i.o) = 0 & \text{when } \sum P(A_n) < \infty \Rightarrow \text{用来处理 a.s. 收敛} \\ P(A_n \ i.o) = 0 \Leftrightarrow \sum P(A_n) = \infty & \text{when } A_n \text{ 相互独立} \end{cases}$

$\boxed{\text{Thm 4.17}}$ (辛钦弱大数定律) 设 $\{X_i\}$ 独立同分布，$\mathbb{E}|X_i| < \infty$ $\mathbb{E}[X_i] = \mu$ 则

$$\underbrace{\frac{1}{n}\sum_{1} X_i}_{\text{独立和的算术平均}} \xrightarrow{P} \mu$$

proof.（技术是截尾法）

引入数列 $a_i = i^\delta$ δ 待定.

截断！$X_i^{(1)} = X_i I_{\{|X_i| < a_i\}}$ $X_i^{(2)} = X_i I_{\{|X_i| \geq a_i\}}$.

令 $S_n^{(1)} = \sum_1^n X_i^{(1)}$ $S_n^{(2)} = \sum_1^n X_i^{(2)}$

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq P\left(\left|S_n^{(1)} - \mathbb{E}[S_n^{(1)}]\right| > \frac{\varepsilon n}{2}\right) + P\left(\left|S_n^{(2)} - \mathbb{E}[S_n^{(2)}]\right| > \frac{\varepsilon n}{2}\right).$$

由 Chebyshev 不等式 $\left(P(|X - \mathbb{E}[X]| \geq a) \leq \frac{1}{a^2}Var(X).\right)$

$$P\left(\left|S_n^{(1)} - \mathbb{E}[S_n^{(1)}]\right| > \frac{\varepsilon n}{2}\right) = \frac{4}{\varepsilon^2 n^2} Var(S_n^{(1)}) \overset{\text{独立}}{=} \frac{4}{\varepsilon^2 n^2}\sum_1^n Var(X_i^{(1)})$$

$$\overset{\text{方差不等式}}{\leq} \frac{4}{\varepsilon^2 n^2}\sum_1^n \mathbb{E}[X_i^{(1)2}] \overset{\text{截断}}{\leq} \frac{4}{\varepsilon^2 n^2}\sum_1^n \mathbb{E}[a_i|X_i|]$$

$X_i^{(1)2} \leq a_i X_i$?

$$= \frac{4}{\varepsilon^2 n^2}\mathbb{E}|X_1| \sum_1^n a_i \approx \frac{4}{\varepsilon^2}\frac{(n+1)^{1+\delta}}{n^2} \xrightarrow{n\to\infty} 0$$

令 $\delta < 1$.

另一方面

$$P\left(\left|S_n^{(2)} - \mathbb{E}[S_n^{(2)}]\right| > \frac{\varepsilon n}{2}\right) \overset{\text{Markov}}{\leq} \frac{2}{\varepsilon n}\mathbb{E}\left[\left|S_n^{(2)} - \mathbb{E}[S_n^{(2)}]\right|\right]$$

$$\overset{\text{三角}}{\leq} \frac{4}{\varepsilon n}\sum_1^n \mathbb{E}[|X_i^{(2)}|]$$

$$\leq \frac{4}{\varepsilon n}\sum_1^n \mathbb{E}\left[|X_1 I_{\{|X_1| \geq a_i\}}|\right] \xrightarrow{\text{Stolz}} 0 \quad \text{as } n \to \infty.$$

因此当 $\delta < 1$. 完成了证明！ □

这里我们只做到了 P 收敛. 能改进！

$\boxed{\text{Thm 4.18}}$ (二阶矩版本) $\{X_i\}$ 独立同分布，$\mathbb{E}[X_i^2] < \infty$ $\mathbb{E}[X_i] = \mu$，则

$$\begin{cases} \frac{1}{n}\sum X_i \xrightarrow{L^2} \mu \\ \frac{1}{n}\sum X_i \xrightarrow{a.s.} \mu. \end{cases}$$

Proof $\mathbb{E}\left[\left(\frac{1}{n}S_n - \mu\right)^2\right] = Var\left(\frac{1}{n}S_n - \mu\right) = \frac{1}{n^2}Var(S_n) = \frac{1}{n^2}\sum_1^n Var(X_i)$

$= \frac{1}{n}Var(X_1) \to \infty$ i.e. $\frac{1}{n}S_n \xrightarrow{L^2} \mu$.

为了证明 $\frac{1}{n}\sum X_i \xrightarrow{a.s.} \mu$. 采用子列方法(?)

$n_i = i^2$. 用B-C引理. 这里可以先假设 X_i 非负.

$$P\left(\left|\frac{S_{n_i}}{n_i}-\mu\right|>\varepsilon\right) \le \frac{Var(S_{n_i})}{(n_i\varepsilon)^2} = \frac{Var(X_1)}{n_i\varepsilon^2}$$

由B-C. $\frac{S_{i^2}}{i^2} \xrightarrow{a.s.} \mu$.

有 $S_{i^2} \le S_n \le S_{(i+1)^2}$ $i^2 \le n^2 \le (i+1)^2$

$$\frac{S_{i^2}}{i^2}\cdot\frac{i^2}{n} \le \frac{S_n}{n} \le \frac{S_{(i+1)^2}}{(i+1)^2}\frac{(i+1)^2}{n}$$

两边取极限 $\Rightarrow \frac{S_n}{n} \xrightarrow{a.s} \mu$. □

$\boxed{\text{Thm 4.19}}$ (柯尔莫洛夫, 强大数定律) 设 $\{X_i\}$ 独立同分布, 则

$$\frac{1}{n}\sum_1^n X_i \xrightarrow{a.s} \mu \iff \mathbb{E}[|X_1|]<\infty, \mu=\mathbb{E}[X_1]$$

Proof. 三步走

STEP 1. 截尾术 $Y_n = X_n I_{\{X_n<n\}}$. 由 $\boxed{\text{EX 4.5}}$ 有

$a_n=n$. $\Rightarrow \frac{1}{n}\sum_1^n(X_n-Y_n) \xrightarrow{a.s} 0$. 那么

$$\frac{1}{n}\sum_1^n X_i \xrightarrow{a.s} \mu \iff \frac{1}{n}\sum_1^n Y_i \xrightarrow{a.s} \mu.$$

STEP 2 a.s.收敛子列 for $\alpha>1$

令 $\beta_k=[\alpha^k]$ $\Rightarrow \frac{\beta_{k+1}}{\beta_k}\to\alpha$.

$\exists A=A(\alpha)$ $\sum_{k=m}^\infty \frac{1}{\beta_k^2} \le \frac{A}{\beta_m^2}$ (这在干嘛...) $\forall m$.
不易过应该就是几何级数的估计

记 $S_n'=\sum_{i=1}^n Y_i$ $\alpha>1$ $\varepsilon>0$.

$$\sum_n P\left(\frac{1}{\beta_n}\left|S_{\beta_n}'-\mathbb{E}[S_{\beta_n}']\right|>\varepsilon\right) \overset{\text{chebyshev}}{\le} \sum \frac{1}{\varepsilon^2\beta_n^2}Var(S_{\beta_n}')$$

$$=\sum_n \frac{1}{\varepsilon^2\beta_n^2}\sum_{i=1}^{\beta_n} Var\, Y_i \overset{\text{Fubini}}{=} \frac{1}{\varepsilon^2}\sum_{i=1}^\infty\left(\sum_{\beta_n\ge i}\frac{1}{\beta_n^2}\right)Var(Y_i)$$

$$\le \frac{A}{\varepsilon^2}\sum\frac{1}{i^2}\mathbb{E}[Y_i^2] \quad <\infty$$

$B_{ij}=\{j-1\le Y_i<j\}$ $\Rightarrow \sum_1^\infty \frac{1}{i^2}\mathbb{E}[Y_i^2] = \sum_1^\infty \frac{1}{i^2}\sum_j^i \mathbb{E}[Y_i^2 I_{B_{ij}}]$

$$\le \sum_1^\infty \frac{1}{i^2}\sum_j^i j^2 P(B_{ij})$$

$$=\sum_{j=1}^\infty\left(\sum_{i=j}^\infty\frac{1}{i^2}\right)j^2P(B_{ij})$$

$$\approx \sum_j jP(B_{ij}) \approx \mathbb{E}[X] < \infty.$$

由BC. $\cancel{\frac{1}{\beta_n}S_{\beta_n}' \to \mu}$ $\frac{1}{\beta_n}\left(S_{\beta_n}'-\mathbb{E}[S_{\beta_n}']\right) \xrightarrow{a.s} 0$.

和. 这里. $\mathbb{E}[Y_n] = \mathbb{E}[X_1 I\{X_1<n\}] \longrightarrow \mathbb{E}[X_1] = \mu.$

由Stolz. $\frac{1}{\beta_n} S'_{\beta_n} \xrightarrow{a.s.} \mu.$

STEP 3. a.s.收敛.

$\beta_m \leq n \leq \beta_{m+1}$ 仍设 $X_i \geq 0$

$S_{\beta_m} \leq S_n \leq S_{\beta_{m+1}}$

$$\frac{\beta_m}{\beta_{m+1}} \frac{S_{\beta_m}}{\beta_m} \leq \frac{\beta_m}{n} \frac{S_{\beta_m}}{\beta_m} \leq \frac{S_n}{n} \leq \frac{S_{\beta_{m+1}}}{\beta_{m+1}} \frac{\beta_{m+1}}{n} \leq \frac{S_{\beta_{m+1}}}{\beta_{m+1}} \frac{\beta_{m+1}}{\beta_m}$$

$$\Rightarrow \quad \frac{1}{\alpha}\mu \leq \underline{\lim} \frac{S_n}{n} \leq \overline{\lim} \frac{S_n}{n} \leq \alpha \mu.$$

$\alpha \downarrow 1^+ \quad \Rightarrow \lim \frac{S_n}{n} = \mu$ a.s.

$\Rightarrow \frac{S_n}{n} \xrightarrow{a.s.} \mu$

另一方面 $\frac{1}{n}\sum_1^n X_i \longrightarrow \mu$ a.s. $\Rightarrow \frac{1}{n} X_i \longrightarrow 0$

$(\Rightarrow)$ 断言 $\sum_{n=1}^{\infty} P(|X_n| \geq n) < \infty$

若不然. $\sum_{n=1}^{\infty} P(|X_n| \geq n) = \infty \overset{独立}{\Longrightarrow} P(\frac{1}{n}|X_n| \geq 1 \text{ i.o.})$. 这与 $\frac{1}{n} X_i \xrightarrow{a.s.} 0$ 矛

又 $\mathbb{E}[X_1] \leq 1 + \sum_1^{\infty} P(|X_n| \geq n) < \infty$

"蛋糕"

且由 $(\Leftarrow)$ 知 $\mathbb{E}[X_1] = \mu$. $\square$

五. 中心极限定理

1. 特征函数

矩母函数有缺点：不存在

总

Def 5.1 X, Y 为 $(\Omega, \mathcal{F}, P)$ 上的随机变量，称 $Z = X + iY$ 为一个复随机变量

规定其期望 $\mathbb{E}[Z] = \mathbb{E}[X] + i\mathbb{E}[Y]$.

Rmk 复随机变量视为二维随机向量.

独立定义为：(X_1, Y_1) 和 (X_2, Y_2) 独立，若 $\forall x_1, x_2, y_1, y_2 \in \mathbb{R}$

$P(X_1 \le x_1, Y_1 \le y_1, X_2 \le x_2, Y_2 \le y_2) = P(X_1 \le x_1, Y_1 \le y_1) \cdot P(X_2 \le x_2, Y_2 \le y_2)$.

$Z_1 \cdots Z_n$ 独立时有 $\mathbb{E}[Z_1 \cdots Z_n] = \mathbb{E}[Z_1] \cdots \mathbb{E}[Z_n]$.

随机变量 X 的特征函数

Def 5.2 $\phi(t) = \mathbb{E}[e^{itX}] \quad t \in \mathbb{R}$.

Rmk (1) $\mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)]$

(2) $|e^{itX}| \le 1 \Rightarrow \phi(t)$ 总存在

(3) $\phi(t) = \int_X e^{itX} \, dF(x) \overset{\text{若F有密度}}{=} \underbrace{\int_X e^{itX} f(x) \, dx}_{\text{Fourier 变换}}$

Thm 5.3 (1) $\phi(0) = 1$, $|\phi(t)| \le 1$, $\phi(-t) = \overline{\phi(t)}$

(2) $\phi(t)$ 在 $(-\infty, \infty)$ 上一致连续

(3) $\phi(t)$ 非负定, i.e. $\forall t_1, \cdots t_n \in \mathbb{R}$, $z_1, \cdots z_n \in \mathbb{C}$

$\sum_{j,k} \phi(t_j - t_k) z_j \bar{z}_k \ge 0$

Proof. (1) $|\phi(t)| \le \int_X |e^{itX}| \, dF = 1$, $\int e^{-itX} dF(x) = \overline{\phi(t)}$

(2) $|\phi(t+h) - \phi(t)| = \left| \int_X e^{itx}(e^{ihx} - 1) \, dF(x) \right|$

$\le 2 \underbrace{\int_{A^c} dF}_{\varepsilon/4} + \underbrace{\int_A |e^{ihx} - 1| \, dF}_{h \to 0} < \varepsilon$. ✓

(3) $\sum_{j,k} \mathbb{E}[e^{i(t_j - t_k)X} z_j \bar{z}_k] = \sum_{j,k} \mathbb{E}[z_j e^{it_jX} \overline{z_k e^{it_kX}}]$

$= \sum_j \mathbb{E}[|\sum_j z_j e^{it_jX}|^2] \ge 0$. □

Rmk Bochner：该性质完全刻画了特征函数.

矩与可微性有关

$\boxed{\text{Thm 5.4}}$ 若 $\mathbb{E}[X^k]<\infty$，则 $\forall j\le k$，有

$$\phi^{(j)}(0)=i^j\,\mathbb{E}[X^j].$$

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

Proof. $\phi(t)=\int_X e^{itX}\,dF$

求导 $\Rightarrow$ $\phi^{(j)}(t)=\int (iX)^j e^{itX}\,dF$ （低阶矩的存在由 Lyapunov 不等式给出 $\|X\|_r\le\|X\|_s$ $r<s$）.

$\Rightarrow \phi^{(j)}(0)=i^j\int x^j\,dF=i^j\,\mathbb{E}[X^j]$.

Taylor 展开 $\Rightarrow$ $\phi(t)=\sum_{j=0}^{k}\frac{(it)^j}{j!}\mathbb{E}[X^j]+o(t^k)$. □

$\boxed{\text{Thm 5.5}}$ (1) $Y=aX+b$. $\phi_Y(t)=e^{itb}\phi_X(at)$

(2) X,Y 独立时 $\phi_{X+Y}(t)=\phi_X(t)\phi_Y(t)$.

$\boxed{\text{Def 5.6}}$ (多元) $\vec{X}=(X_1,\cdots,X_n)$ $\phi_{\vec{X}}(\vec{t})=\mathbb{E}[e^{i\sum_j t_jX_j}]$

$\boxed{\text{Thm 5.7}}$ $\vec{X},\vec{Y}$ 独立 $\Leftrightarrow$ $\phi_{X,Y}(\vec{s},\vec{t})=\phi_{\vec{X}}(\vec{s})\cdot\phi_{\vec{Y}}(\vec{t})$.

Proof. $\Rightarrow$ 显然

$\Leftarrow$ 反演公式

$\boxed{\text{EX 5.1}}$ 参数 p 的 Bernoulli 分布 $\phi(t)=pe^{it}+q$

$X\sim B(n,p)$ $\phi(t)=(pe^{it}+q)^n$.

算 $\phi(t)=\mathbb{E}[e^{itX}]=e^{it0}\cdot\underbrace{(1-p)}_{q}+e^{it1}\cdot p$

$=pe^{it}+q$

$\phi(t)=\mathbb{E}[e^{itX}]=\sum e^{itk}\binom{n}{k}p^kq^{n-k}$

$=(e^{it}\cdot p+q)^n$. □.

$\boxed{\text{EX 5.2}}$ $X\sim\exp(\lambda)$. $f(x)=\lambda e^{-\lambda x}$ $x\ge0$

$\phi(t)=\int_0^\infty \lambda e^{-\lambda x}\cdot e^{itx}\,dx=\frac{\lambda}{\lambda-it}$.

$\boxed{\text{EX 5.3}}$ $X\sim N(0,1)$

$\phi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{itx-\frac{1}{2}x^2}\,dx=e^{-\frac{1}{2}t^2}$

$X\sim N(\mu,\sigma^2)\Rightarrow\phi(t)=e^{i\mu t-\frac{1}{2}\sigma^2t^2}$

$\boxed{\text{EX 5.4}}$ X 多元正态 $X \sim N(\vec{\mu}, \Sigma)$ $Y = Xt' = \sum t_n X_n$

$$\phi(t) = \mathbb{E}[e^{i t X t'}] = \mathbb{E}[e^{isY}]\Big|_{s=1} \quad Y \sim N(\vec{\mu} t', t^{T}\Sigma t).$$

$$= \underline{e^{-\frac{1}{2} t\Sigma t' + i\mu t'}} \quad \cancel{Y \sim N(\mu}$$

$\boxed{\text{Rmk.}}$ 多元正态分布的另一观点.

$\boxed{\text{EX 5.5}}$ $X \sim U[0,1]$. $\phi(t) = \int_0^1 e^{itx} dx = \dfrac{e^{it}-1}{it}$

$X \sim U[-1,1]$ $\phi(t) = \dfrac{\sin t}{t}$.

两个问题 $\begin{cases} \phi_X(t) = \phi_Y(t) \overset{?}{\Rightarrow} X, Y \text{同分布} \\ \phi_{X,Y}(s,t) = \phi_X(s)\phi_Y(t) \overset{?}{\Rightarrow} X, Y \text{独立} \end{cases}$

$\boxed{\text{Thm 5.8}}$ (反演公式) 设 X 的分布函数为 F. $\phi(t) = \int e^{itX} dF$, 则对 $-\infty < a < b < \infty$, 有

$$\frac{F(b)+F(b-0)}{2} - \frac{F(a)+F(a-0)}{2} = \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-iat}-e^{-ibt}}{it}\phi(t)\,dt.$$

Proof: $I_T = \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-iat}-e^{-ibt}}{it} \cancel{\phi(t)dt} = \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-iat}-e^{-ibt}}{it}dt \int_{-\infty}^{\infty} e^{itX} dF$

主值积分

$$\overset{\text{交换}}{=} \frac{1}{2\pi}\int_{\mathbb{R}}\int_{-T}^{T} \frac{e^{-iat}-e^{-ibt}}{it}\,dt\,dF = \int_{\mathbb{R}} g_T\,dF$$

$$g_T(x) = \frac{1}{\pi}\int_0^T \left(\frac{\sin t(x-a)}{t} - \frac{\sin t(x-b)}{t}\right)dt \xrightarrow{T\to\infty} \begin{cases} 1 & x \in (a,b) \\ \frac{1}{2} & x = a \text{ or } b \\ 0 & \text{otherwise} \end{cases}$$

$$\lim_{T\to\infty} I_T = \lim \int_{\mathbb{R}} \overset{\text{有界}}{g_T}\, dF \overset{\text{控制}}{=} \int g_\infty\, dF$$

$$= F(a,b) + \tfrac{1}{2}F(a) + \tfrac{1}{2}F(b)$$

$$= \frac{F(b)+F(b-0)}{2} - \frac{F(a)+F(a-0)}{2} \qquad \square$$

$\boxed{\text{Cor 5.9}}$ (唯一性定理) $\phi_X(t) = \phi_Y(t) \Rightarrow F_X(x) = F_Y(x)$.

Proof: 记 C_F 为 F 的连续点全体. $\mathbb{R}\setminus C_F$ 至多可数. 设 $b \in C_F$.

令 $a \to -\infty \Rightarrow F(b) = \lim_{a\to-\infty} \frac{1}{2\pi}\int \frac{e^{-iat}-e^{-ibt}}{it}\phi(t)\,dt$

当 $x \notin C_F$. $b_n \downarrow x$ $b_n \in C_F \Rightarrow F(x) = \lim_{n\to\infty} F(b_n)$. $\square$

~~若 $b \notin C_F$. $\frac{F(b)+F(b-0)}{2} = \lim_{a\to-\infty}\frac{1}{2\pi}\int\frac{e^{-iat}-e^{-ibt}}{it}\phi(t)dt$~~

46.

Thm 5.10（多元反演公式）记 $\phi(t_1,\cdots t_n)=\mathbb{E}[e^{i\sum_{j=1}^n t_jX_j}]$

则 $P(a_j \le X_j \le b_j, j=1,\cdots n) = \lim_{T_1,\cdots T_n\to\infty}\frac{1}{(2\pi)^n}\int_{-T_1}^{T_1}\cdots\int_{-T_n}^{T_n}\prod_{j=1}^n \frac{e^{-it_ja_j}-e^{-it_jb_j}}{it_j}$.

Rmk 这里假设边界面上的概率为0. 特别地. $\phi_{X,Y}(s,t)=\phi_X(s)\phi_Y(t) \Rightarrow X,Y$ 独立. $\phi(t_1,\cdots t_n) d\cdot$

用反演公式找特征函数~~的分~~对应的分布函数.

EX 5.6 $\cos t$ 对应什么. 用唯一性定理. 可以猜.

$\mathbb{E}[e^{itX}] = \frac{1}{2}e^{it1}+\frac{1}{2}e^{it(-1)} = \cos t$.

Def 5.11 F, F_n 分布函数. 若对 F 的任意连续点 x 都有 $\lim_n F_n(x)=F(x)$

则称 F_n 弱收敛于 F. 记 $F_n \xrightarrow{w} F$

Thm 5.12 设分布函数为 $F_n(x)$. 特征函数 $\phi_n(t)=\int e^{itX}dF_n$

(1) 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 $\phi_n(t)$ 内闭一致收敛

(2) 若 $\phi(t)=\lim_{n\to\infty}\phi_n(t)$ 且 $\phi(t)$ 在 $t=0$ 处连续, 则 $\phi(t)$ 为某分布函数 F

的特征函数. 且 $F_n \xrightarrow{w} F$.

EX 5.7 $X_n \sim U[-n,n]$

$\phi_n(t) = \frac{\sin nt}{nt}$ $t=0 \mapsto 1$

$h(t) \to \delta$ 在 t 处不连续. 不是特征函数

这里有唯一性

Cor 5.13 $X_n \xrightarrow{D} X$, $X_n \to Y$ $X \overset{D}{=} Y \Leftrightarrow \phi_X=\phi_Y \Leftrightarrow \phi_n(t) \xrightarrow{p.w.}$

中心极限定理. $S_n = X_1+\cdots+X_n$

大数: $\frac{S_n-n\mu}{n} \xrightarrow{a.s.} 0$. 现在找 $\sigma_n\to\infty$ 使 $\sigma_n\cdot\frac{S_n-n\mu}{n}$.

取 $\sigma_n=\sqrt{n}$. 因为在 i.i.d 条件下 $Var(S_n)=n^2\sigma^2$

$\boxed{\text{Thm 5.14}}$ (中心极限定理) X_i 独立同. $\mu = \mathbb{E}[X_i]$. $\sigma^2 = Var(X)$. $\sigma \in (0,\infty)$.

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{D} z \sim N(0,1)$$

Proof. $Y_k = \frac{X_k - \mu}{\sigma}$ $U_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{1}{\sqrt{n}}(Y_1 + \cdots Y_n)$

$\phi(t) = \mathbb{E}[e^{itY_1}]$ $\phi_n(t) = \mathbb{E}[e^{itU_n}]$

$$\phi(t) = 1 + \frac{it}{1!} 0 + \frac{(it)^2}{2!} + o(t^2)$$

$$\phi_n(t) = \left(\phi\left(\frac{t}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{1}{2n} t^2 + o\left(\frac{t^2}{n}\right)\right)^n \longrightarrow e^{-\frac{1}{2}t^2}$$ □

$\boxed{\text{Thm 5.15}}$ (多元) $\vec{X}_i$ n 维独立同 r.v. $0 = \mathbb{E}[\vec{X}_i]$ $\Sigma \mathbb{E}[X_i X_i^T] > 0$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \vec{X}_i \xrightarrow{P} N(0, \Sigma).$$

$\boxed{\text{rmk}}$ 二阶矩是重要的.

如果~~要~~独立呢? (不同分布)

$\mu_k = \mathbb{E}[X_k]$. $\sigma_k^2 = Var(X_k)$ ~~则~~ $B_n^2 = \sum_{k=1}^n \sigma_k^2$

Lindeberg 条件: $\forall \varepsilon > 0$, $\lim \frac{1}{B_n^2} \sum_{k=1}^n \mathbb{E}[(X_k - \mu_k)^2 1_{\{|X_k - \mu_k| > \varepsilon B_n\}}] = 0$ (L)

$\boxed{\text{Thm 5.16}}$ X_k ~~相~~ 独立 L.

$$\frac{1}{B_n} \sum_{k=1}^n (X_k - \mu_k) \xrightarrow{D} Z \sim N(0,1). \quad (LF)$$

且 Feller 条件. $\boxed{\frac{1}{B_n^2} \max \sigma_k^2 \to 0.}$ (F).

没有说同分布.
先不证.

推论(?). 当 $\{X_k\}$ 相互独立. $\mathbb{E}[X_k] = 0$, $\sigma_k^2 = Var(X_k)$. $\mathbb{E}[|X_k|^3] < \infty$.

且 $\frac{1}{B_n^3} \sum_{k=1}^n \mathbb{E}[|X_k|^3] \longrightarrow 0$ $n \to \infty$

则 $\frac{1}{B_n} \sum_{k=1}^n X_k \xrightarrow{D} N(0,1)$.

验证 L 条件: $\frac{1}{B_n^2} \sum_{k=1}^n \int_{|x| > \varepsilon B_n} x^2 \, dF_k \le \frac{1}{B_n^2} \sum_{k=1} \int_{|x| > \varepsilon B_n} x^2 \cdot \frac{|x|}{\varepsilon B_n} dF_k$.

$$\le \frac{1}{\varepsilon B_n^3} \int |x|^3 dF_k \longrightarrow 0 \quad n \to \infty.$$

$\boxed{\text{Rmk}}$ (1) (L) 是什么? $\frac{1}{B_n^2} \sum_{k=1}^n \int_{|x-\mu_k| > \varepsilon B_n} (x - \mu_k)^2 dF_k \overset{\text{Markov}}{\ge} \varepsilon^2 \sum_{k=1} P\left(\frac{|X_k - \mu_k|}{B_n} > \varepsilon\right)$

$$\ge ~~\Sigma~~ \varepsilon^2 P\left(\bigcup_{k=1}^n \{\tfrac{1}{B_n}|X_k - \mu_k| > \varepsilon\}\right)$$

$$= \varepsilon^2 P\left(\max \frac{|X_k - \mu_k|}{B_n} > \varepsilon\right).$$

也就是说，"相对偏差" $\frac{1}{B_n}|X_k - \mu_k|$ 一致地依概率收敛.

(2) (L) $\Rightarrow$ (F) 不妨让期望为0.

$$\frac{\sigma_k^2}{B_n^2} = \frac{\mathbb{E}[X_k^2]}{B_n^2} = \frac{\mathbb{E}\left[X_k^2\left(I_{\{|X_k| > \varepsilon B_n\}} + I_{\{|X_k| \le \varepsilon B_n\}}\right)\right]}{B_n^2}$$

$$\le \underbrace{\frac{1}{B_n^2}\sum_{k=1}^{n}\mathbb{E}\left[X_k^2 I_{\{|X_k| > \varepsilon B_n\}}\right]}_{\to 0 \text{ as } n\to\infty \quad (L).} + \varepsilon^2.$$

$$\Rightarrow \limsup \frac{\max \sigma_k^2}{B_n^2} \le \varepsilon^2 \quad \text{这个不等式中 } \varepsilon \text{ 任意.}$$

$$\Rightarrow \frac{\max \sigma_k^2}{B_n} \to 0. \quad \text{as } n\to\infty.$$

(3) (L) 是 (LF) 的充分条件 (这是 CLT 的结论). 且几乎是必要条件. ((F) + $\overset{LF}{\cancel{(CLT)}}$ $\Rightarrow$ (L)).

Feller条件：没有某个随机变量比较突出，这实际上也很符合我们的需求.

证明 CLT 需要更多铺垫. 先看一下下面的结果.

$\boxed{\text{Thm 5.1}}$ (DL 定理) $S_n = \sum_{k=1}^{n} X_k$ $X_k \sim B(1,p)$. $P(S_n = k) = \binom{n}{k} p^k q^{n-k}$.

引入. $0 < p < 1$. $x_k = \frac{k - np}{\sqrt{npq}}$ $0 \le k \le n$.

设 $p \in (0,1)$. 则对所有 k s.t. $|x_k| \le c$. 一致有 $P(S_n = k) \sim \frac{e^{-\frac{1}{2}x_k^2}}{\sqrt{2\pi npq}}$.

proof: $k = \boxed{np} + \sqrt{npq}\, x_k$ $\quad k \sim np$

$n-k = \boxed{nq} + \sqrt{npq}\, x_k$ 有 $\quad n-k \sim nq$

主项.

组合数就用 Stirling 公式. $\Rightarrow P(S_n = k) \sim \left(1 - \frac{\sqrt{npq}}{k}x_k\right)^k \left(1 + \frac{\sqrt{npq}}{n-k}x_k\right)^{n-k} \frac{1}{\sqrt{2\pi npq}}$

令 $\varphi_k = k\log\left(1 - \frac{\sqrt{npq}}{k}x_k\right) + (n-k)\log\left(1 + \frac{\sqrt{npq}}{n-k}x_k\right)$

$$\overset{\text{展开}}{=} -\sqrt{npq}\,x_k - \frac{npq}{k}x_k^2 + o\left(\frac{1}{\sqrt{n}}\right)$$

$$\sqrt{npq}\,x_k - \frac{npq}{n-k}x_k^2 + o\left(\frac{1}{\sqrt{n}}\right)$$

$$= -\frac{n^2pq}{2k(n-k)}x_k^2 \sim \underline{-\frac{1}{2}x_k^2}. \qquad \square$$

Thm 5.18 (积分形式 CLT)

$$P\left(\frac{S_n-np}{\sqrt{npq}}\in(a,b]\right)\longrightarrow\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx.$$

Proof: $LHS=\sum_{k:x_k\in(a,b]}P(S_n=k)\sim\sum_{k:x_k\in(a,b]}\frac{e^{-\frac{1}{2}x_k^2}}{\sqrt{2\pi npq}}$

$$=\sum\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x_k^2}\cdot(x_{k+1}-x_k)\sim\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx$$

$\boxed{x_k=\frac{k-np}{\sqrt{npq}}}$ □

矩方法

$$\underbrace{\int_{\mathbb{R}}x^n\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx}_{\gamma_n}=\begin{cases}(n-1)!! & n\text{ 偶}\\ 0 & n\text{ 奇}\end{cases}$$

组合意义、$\{1,\cdots,\underset{n}{2m}\}$ 两两配对

$(2m-1)!!$ 种.

Thm 5.19 设 $\{X_n\}$ 独立.满足

(1) $\mathbb{E}[X_k]=0$ $Var(X_k)=1$.

(2) (一致有界高阶矩) $C_m=\sup_k\mathbb{E}[|X_k|^m]<\infty$ ($m\geq 3$).

则 $\mathbb{E}\left[\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}x_i\right)^k\right]\longrightarrow\gamma_k$

proof. 直接有计算一般情况

项形如 $x_{i_1}^{k_1}\cdots x_{i_m}^{k_m}$ $k_1+\cdots+k_m=k$. 有是偶数才行

若 $k_j>2$、其余 $k_i\geq 2$. 这类项的贡献大 $\frac{E[x_{i_1}^{k_1}\cdots x_{i_m}^{k_m}]}{n^{k/2}}$ $(m<\frac{k}{2})$.

只能 $k_i=2$. ⟶ 配对数. $k=2m$ $(2m-1)!!\frac{n(n-1)\cdots(n-m+1)}{n^m}\to 0$. □

$\downarrow$

$(2m-1)!!$

Thm 5.20 矩收敛定理.

(1) $\forall k\in\mathbb{N}$, $\gamma_{k,n}=\int x^k dF_n$

(2) $\gamma_{k,n}\longrightarrow\gamma_k$ $n\to\infty$

(3) ∃ 分布 F, $\gamma_k=\int x^k dF$ 且 矩满足 Carleman $\sum_{k=1}^{\infty}(\gamma_{2k})^{-\frac{1}{2k}}<\infty$

或 Riesz $\liminf\frac{1}{k}\cdot\gamma_{2k}^{\frac{1}{2k}}<\infty$ 则

$F_n\xrightarrow{w}F$. $n\to\infty$.

50

$\boxed{\text{Thm 5.27}}$ (依分布收敛刻画) $X_n \xrightarrow{D} X \iff$ 下列之一

(1) $\forall g \in C_b(\mathbb{R})$ $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$

(2) 给定 k. $\forall g.\ g' \cdots g^{(k)} \in C_b(\mathbb{R})$ $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$.

(3) $\phi_n(t) \to \phi(t)$ $\forall t \in \mathbb{R}$.

The Proof of CLT: 思路 矩替换, 不想看

复习课

1. $\{X_k\}$ iid $X_k \sim N(0,1)$ $S_n = \sum_n X_k$, $k<n$, 求 $f_{S_n|S_k}(x|y)$, $f_{S_k|S_n}(y|x)$ $\mathbb{E}(S_k|S_n)$ $\boxed{\text{条件密度, 期望.}}$

2. Maxwell 分布率 $\boxed{\text{3维正态}}$

3. $X_n \xrightarrow{D} X$, $Y_n \xrightarrow{P} c > 0 \Rightarrow X_n Y_n \xrightarrow{D} cX$; X_k iid 非负 $\mathbb{E}[X_i]=1$ $Var(X_i)=\sigma^2$

$\Rightarrow \frac{2}{\sigma}(\sqrt{S_n}-\sqrt{n}) \xrightarrow{D} N(0,1)$

$\frac{2}{\sigma}\sqrt{S_n}-\sqrt{n} = \frac{2}{\sigma}\frac{\frac{S_n-n}{\sqrt{n}}}{\frac{\sqrt{S_n}+\sqrt{n}}{\sqrt{n}}} = \frac{\frac{S_n-n}{\sqrt{n}\sigma}}{\frac{1}{2}\sqrt{\frac{S_n}{n}}+1}$ $\frac{S_n-n}{\sqrt{n}\sigma} \xrightarrow{D} N(0,1)$, $\frac{1}{2}\sqrt{\frac{S_n}{n}}+1 \xrightarrow{a.s.} 1$

4. $q > p > 0$

$\lim_n \int_0^1 \cdots \int_0^1 \frac{x_1^q + \cdots + x_n^q}{x_1^p + \cdots + x_n^p} dx_1 \cdots dx_n = \frac{p+1}{q+1}$

设 X_k iid $U[0,1]$ $LHS = \mathbb{E}\left[\frac{\sum X_k^q}{\sum X_k^p}\right] = \mathbb{E}\left[\frac{\frac{1}{n}\sum X_k^q}{\frac{1}{n}\sum X_k^p}\right] \xrightarrow{\text{交换}} \frac{p+1}{q+1}$ ($\frac{1}{n}\sum X_k^q \xrightarrow{LLN} \frac{1}{q+1}$, $\frac{1}{n}\sum X_k^p \xrightarrow{LLN} \frac{1}{p+1}$)

5. X, $\mathbb{E}[X]=\mu$. $Var(X)=\sigma^2$ $\forall a>0 \Rightarrow P(X \geq \mu+a) \leq \frac{\sigma^2}{\sigma^2+a^2}$

不妨 $\mu=0$

$a = \mathbb{E}[a-X] = \mathbb{E}[(a-X)I_{(a-X)>0} + (a-X)I_{(a-X)\leq 0}]$

$\leq \mathbb{E}[(a-X)I_{(a-X)>0}]$

$\leq \mathbb{E}[(a-X)^2]\cdot P(X<a)$

$= \sqrt{(a^2+\sigma^2)[1-P(X\geq a)]}$

$\Rightarrow P(X \leq a) \leq \frac{\sigma^2}{a^2+\sigma^2}$

6. Brown 运动

$S_n = \sum X_n$ 两点

$\mathbb{E}[S_n S_m] = n \wedge m$

$t \in [0,1]$

$\frac{S_{[nt]}}{\sqrt{n}} \to B_t \sim N(0,t)$.

$B_t - B_s \sim N(0, t-s)$

真复习. X_i iid

(I) $S_n = \sum_{i=1}^n X_i$ $\frac{S_n}{\sqrt{n}} \xrightarrow{a.s.} \mu$ LLN ~~充要~~条件: 期望存在 为 μ

(II) 独立情形 Lindeberg $\mu_{n,k} = \mathbb{E}[X_k]$. $\sigma_{n,k}^2 = Var(X_k)$

$$\lim_n \frac{1}{\sum_k^n \sigma_{n,k}^2} \sum_k^n \mathbb{E}[(X_k-\mu_{n,k})^2 I_{|X_k-\mu_{n,k}|>\varepsilon B_n}] = 0 \quad (L)$$

$$\Rightarrow \frac{1}{B_n}\sum (X_k-\mu_{n,k}) \xrightarrow{D} N(0,1)$$