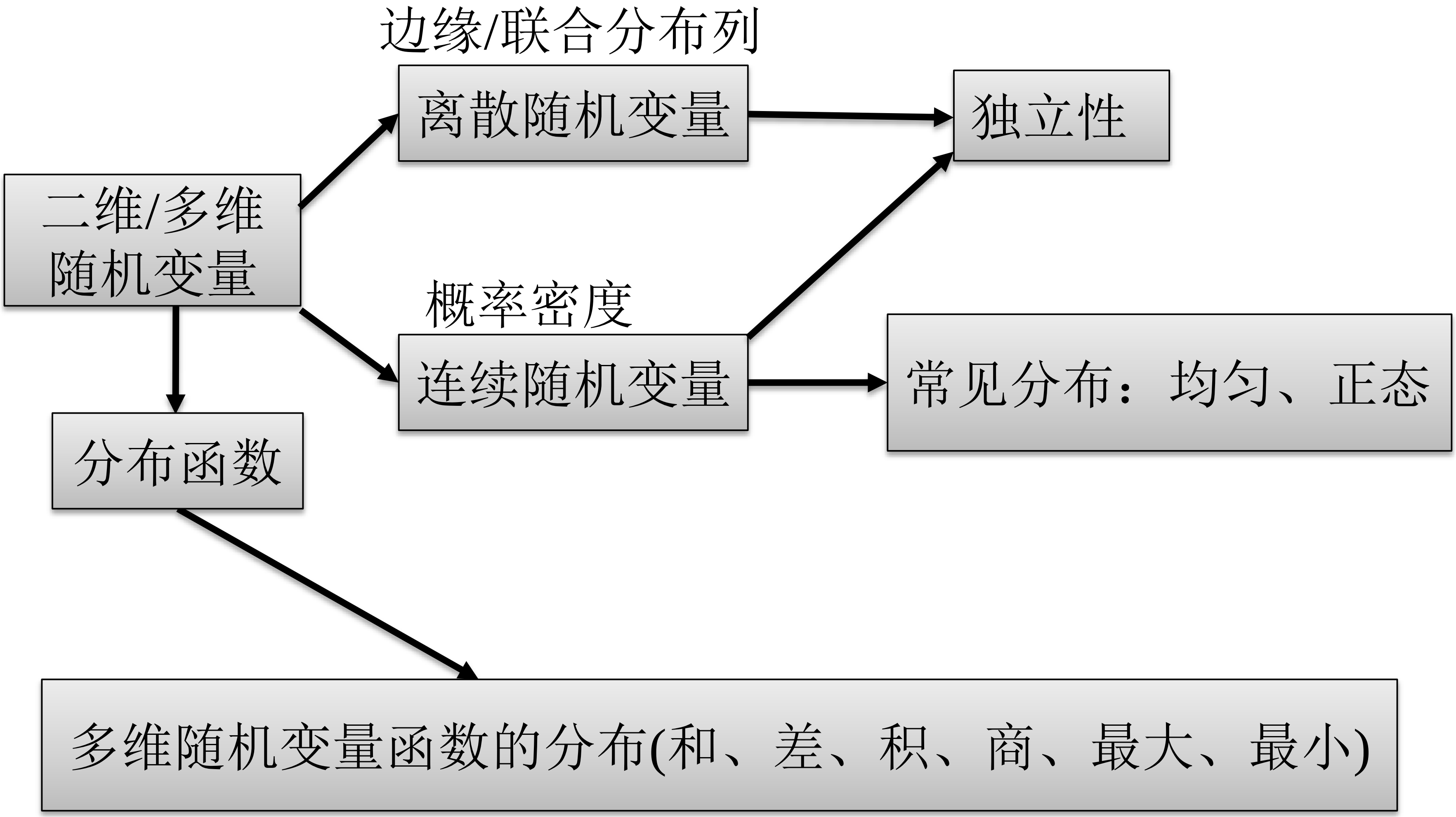
第三章 多维随机变量

知识结构


边缘/联合分布列
离散随机变量
独立性
二维/多维
随机变量
概率密度
连续随机变量
常见分布：均匀、正态
分布函数
多维随机变量函数的分布(和、差、积、商、最大、最小)

二维连续型随机变量

二维随机变量的分布函数为$F(x,y)$，如存在二元非负可积函数 $f(x,y)$ 使得对 (x,y) 有 $F(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)dudv$ 则称 **$f(x,y)$称为随机变量(X,Y)的概率密度**, 或称**联合概率密度**

- 非负性：$f(x,y)\geq 0$;
- 规范性：$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(u,v)dudv$
- 若$f(x,y)$ 在(x,y) 连续, 则$f(x,y)=\partial^2F(x,y)/\partial x\partial y$
- 若G为平面上的一个区域, 则点(X,Y)落入G的概率为

$$P((X,Y)\in G)=\iint_{(x,y)\in G}f(x,y)dxdy$$

例题

设X和Y的分布函数为

$$F(x,y)=\begin{cases}(a-e^{-x})(b-e^{-2y}) & x,y\in[0,+\infty)\\ 0 & \text{其他}\end{cases}$$

其中a,b 非负，求$F(1,1)$

例题

设X和Y的概率密度函数为

$$f(x,y)=\begin{cases}A(x+y) & x,y\in[0,1]\\ 0 & \text{其他}\end{cases}$$

求

- $P(X+Y\leq 1)$
- $P(\max(X,Y)\leq 1/2)$
- $P(\min(X,Y)\leq 1/2)$
- $E[\max(X,Y)+\min(X,Y)]$
- $E[\max(X,Y)\times\min(X,Y)]$

边缘概率密度

设二维随机变量(X,Y)的概率密度为$f(x,y)$，则随机变量X和Y的边缘概率密度为

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

随机变量X和Y相互独立的等价条件：

- 分布函数$F(x,y) = F_X(x)F_Y(y)$
- 概率密度$f(x,y) = f_X(x)f_Y(y)$
- 条件概率$f_{Y|X}(y|x) = f_Y(y)$

例题

设二维随机变量(X,Y)的概率密度

$$f(x,y)=\begin{cases} c & 0<x<1, x^2<y<x \\ 0 & \text{其它} \end{cases}$$

求边缘概率密度

练习

1）设X和Y的分布函数为

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

求X和Y的边缘分布函数

2）设(X,Y)的分布函数为$F(x,y)$，边缘分布函数分别为$F_X(x)$和$F_Y(y)$, 则Z=max(X,Y)的分布函数为(　)

A) $F_X(x)F_Y(y)$　B) $F_X(x)F_Y(x)$　C) $F(x,x)$　D) $F(x,y)$

二维正态分布

设$|\rho| < 1$, 令$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$ $\quad \Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$

若随机变量X和Y的联合概率密度函数为

$$f(x,y) = (2\pi)^{-2/2}|\Sigma|^{-1/2}e^{-\frac{1}{2}(\xi-\mu)^{\mathsf{T}}\Sigma^{-1}(\xi-\mu)} \qquad \xi = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2}-2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}+\frac{(y-\mu_y)^2}{\sigma_y^2}\right]}$$

称$\boldsymbol{X}$**和**$\boldsymbol{Y}$**服从参数为**$\boldsymbol{\mu}$**和**$\boldsymbol{\Sigma}$**的正太分布**, 记 $(\boldsymbol{X},\boldsymbol{Y}) \sim \boldsymbol{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$

- 参数的含义： μ, Σ, ρ
- X和Y的条件概率
- X与Y的独立性
- 联合分布可以推出边缘分布，反只不成立

条件概率

条件分布列：二维离散型随机变量(X,Y)的分布列为$\{p_{ij}\}$，称

$P\left(X=x_i \middle| Y=y_j\right)=\frac{P\left(X=x_i,Y=y_j\right)}{P\left(Y=y_j\right)}=\frac{p_{ij}}{p_{\cdot j}}$ 在$Y=y_j$条件下随机变量X的**条件分布列**

条件概率密度：随机变量(X,Y)的联合概率密度为$f(x,y)$，以及Y的边缘概率密度为$f_Y(y)>0$，称$f_{X|Y}(x|y)=f(x,y)/f_Y(y)$在**$Y=y$条件下随机变量$X$的条件概率密度**

性质与例题

性质：非负性、规范性

条件概率的计算

$$P(a < X < b|Y = y_0) = \int_a^b f_{X|Y}(x|y_0)dx = \int_a^b f(x,y)/f_Y(y_0)dx$$

例题：已知随机变量$X \sim U(0,1)$，当观察到$X = x$的条件下随机变量$Y \sim U(x,1)$，求Y的概率密度以及概率$P(X + Y \leq 1)$

随机变量函数的分布

离散：$p_{ij} = P(X = x_i, Y = y_j), Z = g(X, Y)$ 将相同值合并

连续：i) 给出分布函数$f(x, y)$，求和、差、积

ii) 两个独立随机变量的和、差、积

方法：公式讨论独立随机变量的和、差、积、商、最大、最小

分布函数法

- Step 1: (X, Y) 的联合概率密度$f(x, y)$，则Z的分布函数为

$$F_Z(z) = P(g(X, Y) \le z) = \int_{g(x,y)\le z} f(x, y)dxdy$$

- Step 2: 求概率密度$f_Z(z) = F_Z'(z)$

难点：分类讨论

随机变量函数的结论

多维随机变量函数$\max(X,Y)$ 和 $\min(X,Y)$

$$f_{X+Y}(z)=\int_{-\infty}^{+\infty}f(x,z-x)dx$$

$$f_{XY}(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)dx$$

$$f_{Y/X}(z)=\int_{-\infty}^{+\infty}|x|f(x,xz)dx$$

$X\sim B(n_1,p)$和$Y\sim B(n_2,p)$独立, $X+Y\sim B(n_1+n_2,p)$

$X\sim P(\lambda_1)$ 和$Y\sim P(\lambda_2)$ 独立, 则$X+Y\sim P(\lambda_1+\lambda_2)$

$X\sim N(\mu_1,\sigma_1^2)$和$Y\sim N(\mu_2,\sigma_2^2)$ 独立, 则

$$X+Y\sim \mathrm{N}(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$$

例题

设(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}1 & x,y\in[0,1]\\ 0 & \text{其他}\end{cases}$$

求$Z=|X-Y|$的概率密度

综合题

设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases}Ae^{-y} & y>x>0\\ 0 & \text{其它}\end{cases}$$

求1) 常数A

2) 求X与Y的边缘概率密度

3) 求X与Y是否独立?

4) 求$Z=X+Y$的概率密度

5) 求条件概率密度$f_{X|Y}(x|y)$, 或条件概率$P(1<X<2|Y=1)$

第四章 随机变量的数字特征

知识结构

定义、性质

期望

常见分布期望

随机变量函数的期望

定义、性质

方差

常见分布方差

定义、性质

协方差

正态分布协方差

相关系数

独立与不相关

期望及其性质

离散随机变量**期望** $\boldsymbol{E(X)=\sum_k p_k x_k}$

连续随机变量$E(X)=\int_{-\infty}^{+\infty} xf(x)dx$

若随机变量$X\equiv c$, 则$E(c)=c$

对随机变量X和常数$a,b\in \mathrm{R}$, 有$E(aX+b)=aE(X)+b$

对随机变量X,Y和常数$a,b\in \mathrm{R}$, 有$E(aX+bY)=aE(X)+bE(Y)$

对随机变量X和连续凸函数$g:[a,b]\to \mathrm{R}$, 有$g\big(E(X)\big)\le\ E\big(g(X)\big)$

离散变量X, 及连续函数$g: R\to R$, 有$E[g(X)]=\sum_{k\ge 1} g(x_k)p_k$

连续变量X, 及连续函数$g: R\to R$, 有$E\big(g(X)\big)=\int_{-\infty}^{+\infty} g(t)f(t)dt$

例题

- 设随机变量$X\sim \mathrm{E}(\lambda)$, 则$P\left(X>\sqrt{\mathrm{Var}(X)}\right)=$

- X与Y相互独立且期望存在，记$U=\max(X,Y)$, $V=\min(X,Y)$, 则$E(UV)=$

- 设X与Y服从正态分布，则下列服从正太分布的有( )

  A) $X+Y$　　B) $X-Y$　　C) (X,Y)　　D) $2X+1$

- 设随机变量X的分布函数$F(x)=0.3\Phi(x)+0.7\Phi((x-1)/2)$，其中$\Phi(x)$为标准正态分布的分布函数，则$E(X)=$

  A) 0　　B) 0.3　　C) 0.7　　D) 1

方差的定义及其性质

随机变量方差：$Var\ (X) = E\big(X - E(X)\big)^2 = E(X)^2 - \big(E(X)\big)^2$

- 若随机变量$X \equiv c$, 则$\mathrm{Var}(X) = 0$
- 对随机变量X和常数$a, b \in \ R$, 有

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}\ (X)$$

- 一般情况下方差不具有线性性, 即

$$\mathrm{Var}\big(f(X) + g(X)\big) \neq \ \mathrm{Var}\big(f(X)\big) + \mathrm{Var}\big(g(X)\big)$$

常见分布的期望和方差

- 0/1分布：$X \sim Ber(p)$，$E(X) = p \qquad Var(X) = p(1-p)$
- 二项分布：$X \sim B(n,p)$，$E(X) = np \qquad Var(X) = np(1-p)$
- 几何分布：$X \sim G(p)$，$\mathrm{E}(X) = \frac{1}{p} \qquad \mathrm{Var}(X) = \frac{1-p}{p^2}$
- 负二项分布：X服从参数为r和p的负二项分布

$$\mathrm{E}(X) = \frac{r}{p} \qquad 和 \qquad \mathrm{Var}(X) = \frac{r(1-p)}{p^2}$$

- 泊松分布：$X \sim P(\lambda)$, $E(X) = \lambda \qquad \mathrm{Var}\,(X) = \lambda$

均匀分布 $X \sim U(a,b)$: $E(X) = \frac{a+b}{2} \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$

指数分布 $X \sim e(\lambda)$: $E(X) = \frac{1}{\lambda} \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2}$

正态分布 $X \sim N(\mu,\sigma^2)$：$E(X) = \mu \qquad \mathrm{Var}(X) = \sigma^2$

多元随机变量的期望

离散随机变量 $E[Z] = E[g(X,Y)] = \sum_{i,j} g\left(x_i, y_j\right) p_{ij}$

连续随机变量$E[Z] = E[g(X,Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) f(x,y) dxdy$

- 对任意随机变量X, Y和常数a, b有$E[aX + bY] = aE[X] + bE[Y]$
- 对**独立**随机变量X和Y, 以及任意函数$h(x)$和$g(y)$, 有

$$E[XY] = E[X]E[Y] \text{ 和 } E[h(X)g(Y)] = E[h(X)]E[g(Y)]$$

- 对任意随机变量X和Y, 有Cauchy-Schwartz不等式

$$E[XY] \leq \sqrt{E[X^2]E[Y^2]}$$

协方差及其性质

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

对随机变量X和常数c, 有$\mathrm{Cov}(X,c)=0$.

交换律 $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$

对任意常数a和b, 随机变量X和Y, 有

$$\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y),$$

$$\mathrm{Cov}(X+a,Y+b)=\mathrm{Cov}(X,Y)$$

$$\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$$

$$Cov\left(\sum_{i=1}^{n}X_i,\sum_{j}^{m}Y_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}Cov(X_i,Y_j)$$

$$Var\left(\sum_{i}^{n}X_i\right)=\sum_{i}^{n}Var(X_i)+2\sum_{i<j}Cov(X_i,X_j)$$

相关系数的定义

X与Y的相关系数：$\rho_{XY} = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$

- $|\rho_{XY}| = 1$的充要条件为X与Y有线性关系$Y = aX + b$

- **本质上$\boldsymbol{\rho_{XY}}$刻画了$\boldsymbol{X}$与$\boldsymbol{Y}$的线性相关程度, 又称为“线性相关系数”**

若相关系数$|\rho_{XY}| = 0$，则随机变量***X*和*Y*不相关**

X和Y不相关　　　　X和Y独立

例题

随机变量X和Y在以点$(0,1),(1,0),(1,1)$为顶点的三角形区域服从均匀分布，求$U=X-Y$的方差

第五章 大数定律与中心极限定理

问题

什么是大数定律？

什么是中心极限定理？

其局限是什么，可以采用什么方式弥补

知识结构

大数定律 → **依概率收敛** → 四个大数定律

中心极限定理 → **依分布收敛** → 两个中心极限定理

大数定律

设$X_1, X_2, \cdots, X_n, \cdots$是一随机变量序列，$a$是一常数，如果对任意$\epsilon > 0$有$\lim_{n\to\infty} P[|X_n - a| < \epsilon] = 1$, or $\lim_{n\to\infty} P[|X_n - a| > \epsilon] = 0$

则称随机变量序列$X_1, X_2, \cdots, X_n, \cdots$依概率收敛于$a$, 记$X_n \xrightarrow{P} a$

若随机变量序列$X_1, X_2, \cdots, X_n, \cdots$满足

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \frac{1}{n}\sum_{i=1}^{n} E[X_i],$$

则称$\{X_n\}$服从大数定律

大数定理刻画了随机变量的均值 (算术平均值) 依概率收敛于期望的均值 (算术平均值)

大数定律总结

Markov大数定律: 若随机变量序列$\{X_i\}$满足$\mathrm{Var}(\sum_{i=1}^{n} X_n)/n^2 \to 0$,
则满足大数定律

Chebyshev大数定律: 若独立随机变量序列$\{X_i\}$满足$\mathrm{Var}(X_i) \le c$,
则满足大数定律

Khintchine大数定律: 若独立同分布随机变量序列$\{X_i\}$期望存在,
则满足大数定律

Bernoulli大数定律: 对二项分布$X_n \sim \mathrm{B}(n,p)$, 有$X_n/n \xrightarrow{P} p$

依分布收敛

设随机变量Y的分布函数为$F_Y(y)=P(Y\le y)$，以及随机变量序列$Y_1,Y_2,\cdots,Y_n,\cdots$的分布函数分别为$F_{Y_n}(y)=P(Y_n\le y)$, 如果

$$\lim_{n\to\infty}P[Y_n\le y]=P[Y\le y] \text{ 即 } \lim_{n\to\infty}F_{Y_n}(y)=F_Y(y)$$

则称随机变量序列$\boldsymbol{Y_1,Y_2,\cdots,Y_n,\cdots}$**依分布收敛于**$\boldsymbol{Y}$, 记$\boldsymbol{Y_n\xrightarrow{d}Y}$.

- 林德贝格-勒维中心极限定理：独立同分布随机变量，若$E[X_k]=\mu$和$\mathrm{Var}(X_k)=\sigma^2$, 则

$$\frac{\sum_{k=1}^{n}X_k-n\mu}{\sigma\sqrt{n}}\xrightarrow{d}N(0,1)$$

- 李雅普诺夫定理: 独立不同分布中心极限定理