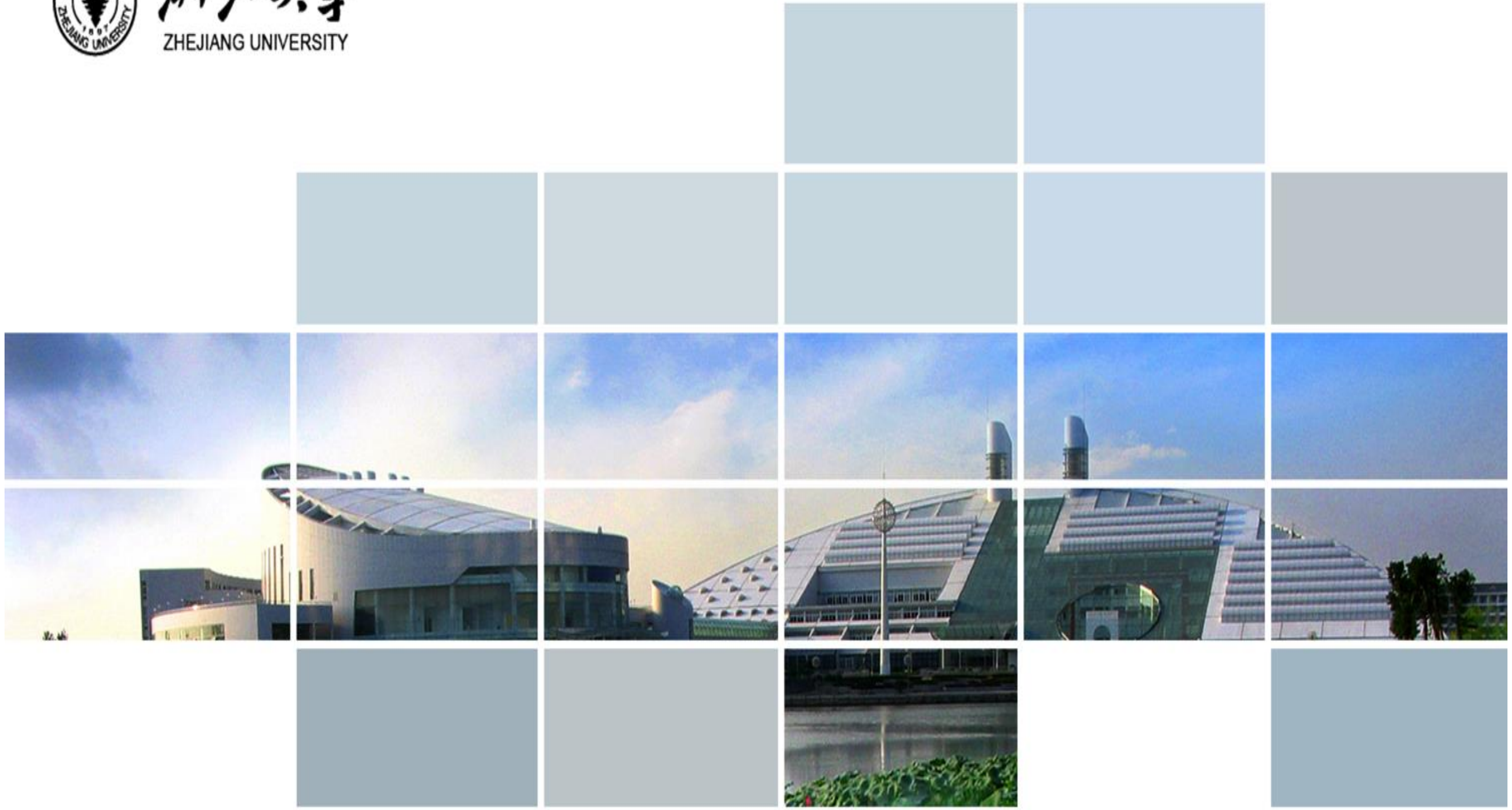


概率复习

§1 随机变量

定义1.1 设Ω 是一个非空集合，$\mathcal{F}$是Ω 中的子集类，如果满足：

(1) $\emptyset, \Omega \in \mathcal{F}$;

(2)(关于取补封闭)：若$A \in \mathcal{F}$, 则$A^c \in \mathcal{F}$;

(3)(关于可列并封闭):若$A_1, A_2, \cdots \in \mathcal{F}$, 则$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$,

则称$\mathcal{F}$是Ω 上σ-**代数**, 称$(\Omega, \mathcal{F})$ 为 **可测空间**。

设$\mathbb{P}:\mathcal{F}\to[0,1]$, 如果满足：

(a) $\mathbb{P}(\Omega)=1$,

(b)(可列可加性：) 若$A_1,A_2,\cdots\in\mathcal{F}$且两两互斥，则

$$\mathbb{P}(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}\mathbb{P}(A_i),$$

则称$\mathbb{P}$ 是$(\Omega,\mathcal{F})$ 上的**概率测度**。

此时称$(\Omega,\mathcal{F},\mathbb{P})$ 为 **概率空间**。

性质：

(1)Ω上任意多个σ-代数的交仍然是σ-代数.

(2)若$\mathcal{H}$是Ω中的子集类，则Ω上所有包含$\mathcal{H}$的σ-代数的交是包含$\mathcal{H}$的最小σ-代数,

称为由$\mathcal{H}$生成的σ-代数, 记为$\sigma(\mathcal{H})$.

A是Ω中非空真子集

例1. 令Ω为样本空间，$A \subset \Omega$。定义

$$\mathcal{F} = \{\emptyset, \Omega, A, \bar{A}\}$$

$$P(\emptyset) = 0, \quad P(\Omega) = 1$$

$$P(A) = p, \quad P(\bar{A}) = 1 - p, \quad 0 < p < 1$$

那么，$(\Omega, \mathcal{F}, P)$是概率空间，最简单的非平凡概率空间。

例2. 古典概率模型

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_N\}, \quad N < \infty$$

$$\mathcal{F} = 2^{\Omega}$$

所有子集构成的σ-域

$$P(A) = \frac{|A|}{N}$$

那么，$(\Omega, \mathcal{F}, P)$是一个古典概率空间。

等价地，每个基本结果等可能地发生。

例3： 设$\Omega=(a,b)$,

$\mathcal{F}$是由Ω中所有开集生成的σ-代数

令m是$(\Omega,\mathcal{F})$ 上的Lebesgue测度

特别地，对$(c,d)\subseteq(a,b), m(c,d)=d-c$

对$A\in\mathcal{F}$，令：

$$P(A)=\frac{m(A)}{b-a}$$

那么，$(\Omega,\mathcal{F},P)$是一个几何概率空间。

定义1.2 设$(\Omega, \mathcal{F})$ 和$(\Omega', \mathcal{F}')$ 是给定的两个可测空间，ξ 是Ω 到Ω'上的映射。如果对任何$A \in \mathcal{F}'$ 有$\xi^{-1}(A) \in \mathcal{F}$, 则称$\xi$ 为**可测映射**。

特别地，当$(\Omega', \mathcal{F}') = (\mathbb{R}, \mathcal{B})$ 时，

称ξ 为**随机变量**(或可测函数), 这里$\mathcal{B}$是$\mathbb{R}$上Borel σ-代数。

定义1.3 设$\xi:\Omega\to\mathbb{R}$，记$\sigma(\xi)$**是Ω上使得**

ξ成为随机变量的最小σ-代数。

称为由ξ **生成的σ-代数**.

更一般地，设T是一个指标集，对$i\in T$, 有$\xi_i:\Omega\to\mathbb{R}$。记

$\sigma(\xi_i;i\in T)$

是使得所有ξ_i $(i\in T)$ 成为随机变量的最小σ-代数。

称为由$\{\xi_i:i\in T\}$ **生成的σ-代数**,

例1.1 班上有6个学生, 编号分别为1 ~ 6 号, 其中1 ~ 4号是男生, 5 ~ 6号是女 生, 1,2,6号戴眼镜,3,4,5号不带眼镜.

任取一名学生, 用X 表示这个学生的性别(0表示女生，1表示男生), 用Y 表示这个学生戴眼镜情况(0表示不戴眼镜,1表示带眼镜).写出概率空间$(\Omega, \mathcal{F}, \mathbb{P})$以及$X, Y, \sigma(X), \sigma(Y), \sigma(X, Y)$.

解： $\Omega = \{1,2,3,4,5,6\},\ \mathcal{F} = 2^{\Omega} = \{A : A \subseteq \Omega\},$

$$\mathbb{P}(A) = |A|/6$$

$$X(\omega) = \begin{cases} 1, & \omega = 1,2,3,4, \\ 0, & \omega = 5,6. \end{cases} \quad Y(\omega) = \begin{cases} 1, & \omega = 1,2,6, \\ 0, & \omega = 3,4,5. \end{cases}$$

$$\sigma(X) = \{\{1,2,3,4\},\{5,6\},\emptyset,\Omega\};$$

$$\sigma(Y) = \{\{1,2,6\},\{3,4,5\},\emptyset,\Omega\};$$

$$\sigma(X,Y) = \sigma\{\{1,2\},\{3,4\},\{5\},\{6\}\} =$$

$$\{\emptyset,\Omega,\{1,2\},\{1,2,3,4\},\{1,2,5\},\{1,2,6\},\{1,2,3,4,5\},\{1,2,3,4,6\},$$
$$\{1,2,5,6\},\{3,4\},\{3,4,5\},\{3,4,6\},\{3,4,5,6\},\{5\},\{5,6\},\{6\}\}.$$

- 1. 数学期望

假设X是一个离散型随机变量，其分布列为

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n \cdots \\ p_1 & p_2 & \cdots & p_n \cdots \end{pmatrix}$$

如果

$$\sum_{n=1}^{\infty} |x_n| p_n < \infty$$

那么称X的数学期望存在，记为

$$EX = \sum_{n=1}^{\infty} x_n p_n$$

假设X是一个连续型随机变量，其密度函数为$p(x)$。如果

$$\int_{-\infty}^{\infty} |x|p(x)dx < \infty$$

那么称X的数学期望存在，记为

$$EX = \int_{-\infty}^{\infty} xp(x)dx$$

假设X是一个随机变量，其分布函数为$F(x)$。如果

$$\int_{-\infty}^{\infty}|x|dF(x)<\infty$$

其中积分为Riemann-Stieltjes积分，那么称X的数学期望存在，记为

$$EX=\int_{-\infty}^{\infty}xdF(x)$$

Riemann-Stieltjes 积分

回忆f关于$G(x)$的积分

$$\int_a^b f(x)dG(x)$$

(1) 分割

在$[a,b]$上插入$n-1$个分点：$a=x_0<x_1<\cdots<x_n=b$

(2) 取点

在每个$[x_{i-1},x_i]$任取一点ξ_i

(3) 作和(Riemann-Stieltjes 和)

$$S_n=\sum_{i=1}^{n} f(\xi_i)(G(x_i)-G(x_{i-1}))$$

(4) 求极限

如果存在一个S，使得对任意划分，任意取点，

$$\lim_{\max_i(x_i - x_{i-1})\to 0} S_n = S$$

那么称S为函数f在$[a,b]$上的Riemann-Stieltjes 积分，记作

$$\int_a^b f(x)dG(x) = S$$

- 什么条件下$\int_a^b f(x)dG(x)$存在呢？

- 有界变差函数

假设G是$[a,b]$上的实值函数。对$[a,b]$进行分割

$$\Delta : a = x_0 < x_1 < \cdots < x_n = b$$

如果

$$\sup_{\Delta} \sum_{i=1}^{n} |G(x_i) - G(x_{i-1})| < \infty$$

那么称G在$[a,b]$上具有有界变差。

假设G是$[0,\infty)$上的实值函数。如果

$$\sup_{t>0} \sup_{\Delta} \sum_{i=1}^{n} |G(x_i) - G(x_{i-1})| < \infty$$

其中Δ是$[0,t]$上的分割，

那么称G在$[0,\infty)$上具有有界变差。

- 有界变差函数的性质

(1) 任何有界变差函数都可以写成两个单调增函数的差;

(2) 任何有界变差函数都几乎处处可微;

(3) $[a,b]$上具有有界导函数的函数是有界变差函数;

(4) $[a,b]$上的单调函数是有界变差函数

- 假设f是$[a,b]$上的连续函数，G是$[a,b]$上的有界变差函数，那么

$$\int_a^b f(x)dG(x) \quad 存在$$

数学期望具有以下基本性质:

(1) 如果$P(X=c)=1$,那么$E(X)=c$

(2) 如果$a \le X \le b$，那么$a \le EX \le b$

(3) (线性性质) $E(aX+b)=aEX+b$

(4) (加法定理) $E(X+Y)=EX+EY$

(5) 如果$X \sim F(x)$，$f:\mathbb{R} \mapsto \mathbb{R}$是Borel 可测函数，那么

$$Ef(X)=\int_{-\infty}^{\infty} f(x)dF(x)$$

(假定上述等式两边存在)

数学期望的另一种定义：（两个定义等价）

$$1.\ E(1_A(\omega)) = P(A),\text{这里}1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

$$2.\ \text{如}\ \eta=\sum_{i=1}^{n} a_i 1_{A_i}\ \text{是简单随机变量，则令}\ E(\eta) = \sum_{i=1}^{n} a_i P(A_i)$$

3. 如果X是非负随机变量，则令

$$E(X) = \sup\{E\eta : 0 \le \eta \le X, \eta\text{是简单随机变量}\}$$

4. 如果X是随机变量，则$X=X^{+}-X^{-}$,

如果$E(X^{+})<\infty, E(X^{-})<\infty$,则称$X$的数学期望存在,

且$E(X)=E(X^{+})-E(X^{-})$, 这里 $X^{+}=\max(X,0)$, $X^{-}=\max(-X,0)$.

$$EX=\int_{\Omega} X(\omega)\,dP(\omega)=\int_{\Omega} X\,dP$$

假设$\mathbb{E}(X^2) < \infty$。如果用一个常数c来估计X，那么当c为何值时，均方误差$\mathbb{E}[(X-c)^2]$最小呢？

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

所以当且仅当$c = \mathbb{E}(X)$时，$E[(X-c)^2]$最小。

定理： 设$\{X_n\}$是随机变量序列.

1.（单调收敛定理）如果X_n非负且递增收敛于X,

那么$\lim\limits_{n\to\infty} E(X_n)=E(X)$

2. (Fatou引理）如$\{X_n\}$非负，那么

$$E(\liminf_{n\to\infty} X_n)\le \liminf_{n\to\infty} E(X_n)$$

3.（Lebesgue控制收敛定理）如果$\lim\limits_{n\to\infty} X_n=X$,且存在

非负随机变量η使得$E(\eta)<\infty$,且对所有n,$|X_n|\le\eta$,

则有$\lim\limits_{n\to\infty} E(X_n)=E(X)$

例: 设$\Omega=(0,1)$, $\mathcal{F}$是由Ω中开集生成的σ-代数

$$P(A)=m(A)$$

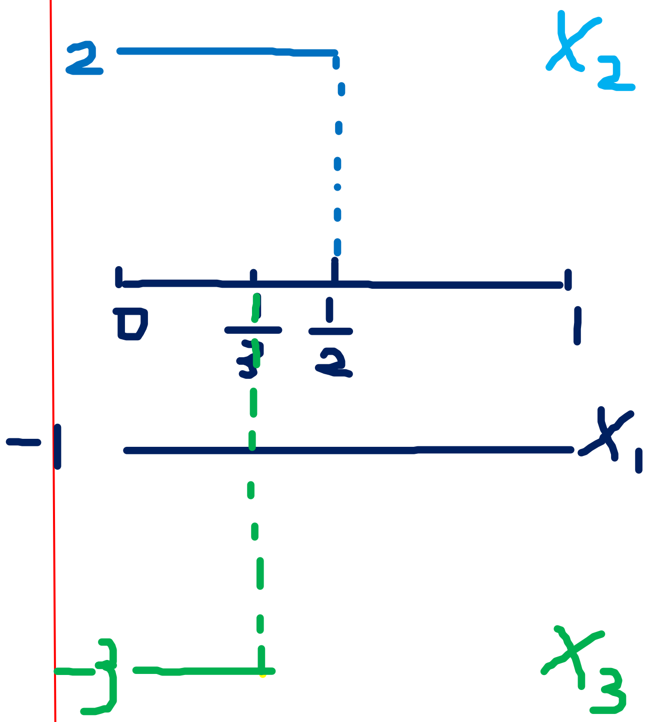
令$X_n(\omega)=(-1)^n n 1_{(0,1/n)}(\omega)$

$$=\begin{cases}(-1)^n n, & \text{当}\ \omega\in(0,1/n)\text{时},\\ 0, & \text{当}\ \omega\in[1/n,1)\text{时},\end{cases}$$

则对所有$\omega\in\Omega$, $\lim\limits_{n\to\infty}X_n(\omega)=X(\omega)=0$,

$E(X)=0$, 但$E(X_n)=(-1)^n$极限不存在

即$E(X_n)$并不收敛于$E(X)$



Fubini定理： 设对$a \le t \le b$，$X(t)$是随机变量.

如果$\int_a^b E(|X(t)|)dt < \infty$

或对所有$a \le t \le b, X(t)$都非负,

则 $E[\int_a^b X(t)dt] = \int_a^b E(X(t))dt.$

例： 设$\xi \sim U(0,1)$, 令$X(t)=\dfrac{t^2-\xi^2}{(t^2+\xi^2)^2}, t\in[0,1]$

$$\text{则}\int_0^1 X(t)dt=\int_0^1\frac{t^2-\xi^2}{(t^2+\xi^2)^2}dt=-\frac{t}{t^2+\xi^2}\Bigg|_0^1=-\frac{1}{1+\xi^2}$$

$$E\int_0^1 X(t)dt=\int_0^1-\frac{1}{1+x^2}dx=-\frac{\pi}{4}$$

$$EX(t)=\int_0^1\frac{t^2-x^2}{(t^2+x^2)^2}dx=\frac{x}{x^2+t^2}\Bigg|_0^1=\frac{1}{1+t^2}$$

$$\int_0^1 EX(t)dt=\int_0^1\frac{1}{1+t^2}dt=\frac{\pi}{4}\quad \text{则}E[\int_0^1 X(t)dt]\neq\int_0^1 E(X(t))dt$$

例： 设X是一非负随机变量，α为一正数，

证明 $E(X^{\alpha})=\int_0^{\infty}\alpha t^{\alpha-1}P(X>t)dt.$

证明： $E(X^{\alpha})=E(\int_0^{X}\alpha t^{\alpha-1}dt)$

$$=E(\int_0^{\infty}\alpha t^{\alpha-1}1_{\{X>t\}}dt)$$

由于$\alpha t^{\alpha-1}1_{\{X>t\}}$非负，由Fubini定理，

$$E(X^{\alpha})=\int_0^{\infty}E(\alpha t^{\alpha-1}1_{\{X>t\}})dt=\int_0^{\infty}\alpha t^{\alpha-1}P(X>t)dt$$

对于数列，也有类似结论. 设$\{a_{nk};k\geq 1\}$是数列.

定理:1.(单调收敛定理) 如果a_{nk}非负且递增收敛于a_k,

$$\text{那么}\lim_{n\to\infty}\sum_{k=1}^{\infty}a_{nk}=\sum_{k=1}^{\infty}a_k$$

2. (Fatou引理) 如果a_{nk}非负，那么

$$\sum_{k=1}^{\infty}\liminf_{n\to\infty}a_{nk}\leq\liminf_{n\to\infty}\sum_{k=1}^{\infty}a_{nk}$$

3.(Lebesgue控制收敛定理)如果$\lim\limits_{n\to\infty} a_{nk} = a_k$,且存在

非负数列$\{b_k;\ k \geq 1\}$使得$\sum\limits_{k=1}^{\infty} b_k < \infty$,

且对所有n,k,有$|a_{nk}| \leq b_k$,

则有 $$\lim_{n\to\infty}\sum_{k=1}^{\infty} a_{nk} = \sum_{k=1}^{\infty} a_k$$

$$\begin{array}{cccc|c}
a_{11} & a_{12} & a_{13} & \cdots & \sum_{k=1}^{\infty} a_{1k} \\
a_{21} & a_{22} & a_{23} & \cdots & \sum_{k=1}^{\infty} a_{2k} \\
a_{31} & a_{32} & a_{33} & \cdots & \sum_{k=1}^{\infty} a_{3k} \\
\downarrow & \downarrow & \downarrow & \vdots & \downarrow \\
a_1 & a_2 & a_3 & \cdots & \sum_{k=1}^{\infty} a_k
\end{array}$$

4.（Fubini定理）：

如果$\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}|a_{nk}|<\infty$,

或对所有n,k,有$a_{nk}\geq 0$,

则有 $$\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}a_{nk}=\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}a_{nk}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} & \cdots \\ a_{21} & a_{22} & a_{23} & \cdots \\ a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{matrix}$$

例： 设$a_{nk}=\begin{cases}1/n, & 若 1\le k\le n,\\ 0, & 若 k>n,\end{cases}$

则$\lim\limits_{n\to\infty} a_{nk}=0=a_k$,

但$\lim\limits_{n\to\infty}\sum\limits_{k=1}^{\infty} a_{nk}=1\ne 0=\sum\limits_{k=1}^{\infty} a_k$

$$\begin{array}{cccccc}
1 & 0 & 0 & 0 & 0 & \cdots\\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \cdots\\
\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \cdots\\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow\\
0 & 0 & 0 & 0 & 0 & 0\cdots
\end{array}$$

例： 设 $a_{nk}=\begin{cases}1, & 若n=k,\\ -1, & 若k=n-1,\\ 0, & 其他,\end{cases}$

则 $\sum\limits_{n=1}^{\infty}\sum\limits_{k=1}^{\infty}a_{nk}=1+0+0+\cdots=1$

$$\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}a_{nk}=0+0+0+\cdots=0$$

$$\therefore\sum_{n=1}^{\infty}\sum_{k=1}^{\infty}a_{nk}\neq\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}a_{nk}$$

$$\begin{matrix}1 & 0 & 0 & 0 & 0 & \cdots\\ -1 & 1 & 0 & 0 & 0 & \cdots\\ 0 & -1 & 1 & 0 & 0 & \cdots\\ 0 & 0 & -1 & 1 & 0 & \cdots\\ 0 & 0 & 0 & -1 & 0 & \cdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots\end{matrix}$$

- 特征函数

特征函数对于研究随机变量的分布函数起着重要作用。假设X是一个随机变量，具有分布函数$F(x)$。定义

$$\phi(t) = Ee^{itX} = \int_{-\infty}^{\infty} e^{itx} dF(x), \quad t \in \mathbb{R}$$

称$\phi(t)$为X的特征函数。这里 $e^{itx} = \cos tx + i\sin tx$.

任何随机变量的特征函数总是存在的。

特征函数具有下列基本性质：

(1) $\phi(0)=1$;

(2) $|\phi(t)|\leq 1$;

(3) $\phi(t)$ 在$\mathbb{R}$上一致连续；

(4) $\phi(t)$是非负定的，即对任意实数$t_1,t_2,\cdots,t_m$和复数$z_1,z_2,\cdots,z_m$,

$$\sum_{i,j=1}^{n} z_i\bar{z}_j\phi(t_i-t_j)\geq 0$$

(5) 如果X,Y相互独立，那么

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

唯一性定理：分布函数和特征函数相互唯一确定。

2. 随机变量序列收敛

依概率收敛

依分布收敛

几乎处处收敛

均方收敛

1. 依概率收敛

假设$(\Omega, \mathcal{F}, P)$是概率空间，$X, X_n, n \geq 1$是一列随机变量。

如果对任意$\varepsilon > 0$

$$P(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) \to 0$$

那么称X_n依概率收敛到X，记作$X_n \xrightarrow{P} X$。

依概率收敛的基本性质：

(1) 如果$X_n \xrightarrow{P} X$，$X_n \xrightarrow{P} Y$，那么

$$P(X = Y) = 1$$

(2) 如果$X_n \xrightarrow{P} X$，$X_n - Y_n \xrightarrow{P} 0$，那么

$$Y_n \xrightarrow{P} X$$

(3) 如果$X_n \xrightarrow{P} X$，$Y_n \xrightarrow{P} Y$，那么

$$X_n \pm Y_n \xrightarrow{P} X \pm Y$$

并且

$$X_n \cdot Y_n \xrightarrow{P} X \cdot Y$$

(4) 如果$X_n \xrightarrow{P} X$，$Y_n \xrightarrow{P} Y$，$Y \neq 0$，那么

$$\frac{X_n}{Y_n} \xrightarrow{P} \frac{X}{Y}$$

(5) 如果$X_n \xrightarrow{P} X$，f是连续函数，那么

$$f(X_n) \xrightarrow{P} f(X)$$

2. 依分布收敛

假设$F(x), F_n(x), n\geq 1$是一列分布函数，如果对每一个F的连续点x，都有

$$F_n(x)\to F(x)$$

那么称F_n弱收敛到F，记作$F_n \xrightarrow{w} F$。

假设随机变量X, X_n分别定义在概率空间$(\Omega,\mathcal{F},P)$，$(\Omega_n,\mathcal{F}_n,P_n)$上，$n\geq 1$。如果相应的分布函数弱收敛，那么称$X_n$依分布收敛到$X$，记作$X_n \xrightarrow{d} X$。

依分布收敛的基本性质

(1) 如果$X_n \xrightarrow{P} X$，那么$X_n \xrightarrow{d} X$

(2) 如果$X_n \xrightarrow{d} c$，那么$X_n \xrightarrow{P} c$

(3) 如果$X_n - Y_n \xrightarrow{P} 0$，并且$X_n \xrightarrow{d} X$，那么$Y_n \xrightarrow{d} X$

(4) 如果$X_n \xrightarrow{d} X$，f是连续函数，那么$f(X_n) \xrightarrow{d} f(X)$

Levy连续性定理

假设$X, X_n, n \geq 1$是一列随机变量，特征函数分别为$\phi(t)$，$\phi_n(t)$。那么$X_n \xrightarrow{d} X$ 当且仅当对每个t，$\phi_n(t) \longrightarrow \phi(t)$。

假设$X_n, n \geq 1$是一列随机变量，特征函数为$\phi_n(t)$。如果对每个t，$\phi_n(t)$ 收敛到某函数$\phi(t)$，并且ϕ在$t = 0$处连续，那么$\phi(t)$一定是某随机变量X的特征函数，进而$X_n \xrightarrow{d} X$ 。

3.几乎处处收敛(以概率1收敛)

假设$X, X_n, n \geq 1$是定义在概率空间$(\Omega, \mathcal{F}, P)$上得一列随机变量，如果存在$\Omega_0 \in \mathcal{F}$，使得

$$P(\Omega_0) = 0$$

并且对任意$\omega \in \Omega \setminus \Omega_0$,

$$X_n(\omega) \longrightarrow X(\omega)$$

那么称X_n几乎处处收敛到X，记作$X_n \longrightarrow X$ a.s.

几乎处处收敛的基本性质：

1. 如果$X_n \longrightarrow X$ a.s.，那么$X_n \xrightarrow{P} X$。

2. 如果对任意$\varepsilon > 0$，

$$\sum_{n=1}^{\infty} P(\omega : |X_n(\omega) - X(\omega)| > \varepsilon) < \infty$$

那么$X_n \longrightarrow X$ a.s.

4. 均方收敛

假设$X, X_n, n \geq 1$是一列定义在概率空间$(\Omega, \mathcal{F}, P)$上的随机变量，具有有限r阶矩$(r > 0)$。如果

$$E|X_n - X|^r \longrightarrow 0$$

那么称X_n r阶均方收敛到X，记作$X_n \xrightarrow{L^r} X$。

如果$X_n \xrightarrow{L^r} X$，那么$X_n \xrightarrow{P} X$。

各种收敛关系：

$$X_n \xrightarrow{L^r} X \Longrightarrow X_n \xrightarrow{P} X \Longrightarrow X_n \xrightarrow{d} X$$

$$X_n \longrightarrow X \ \ a.s \Longrightarrow X_n \xrightarrow{P} X$$

$$X_n \xrightarrow{d} X \Longleftrightarrow \phi_{X_n}(t) \longrightarrow \phi_X(t)$$

3. 条件概率和条件期望

设$(\Omega,\mathcal{F},\mathbb{P})$是一个概率空间，$A\in\mathcal{F}$,且$\mathbb{P}(A)>0$.

当A发生时事件B发生的条件概率定义为：

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

此时$(\Omega,\mathcal{F},\mathbb{P}(\cdot|A))$也是一个概率空间。

若X是$(\Omega,\mathcal{F},\mathbb{P})$上可积的随机变量，定义$X$在$A$发生时的条件期望

为X关于条件概率$\mathbb{P}(\cdot|A)$的期望，计为$\mathbb{E}(X|A)$.

注意到$\mathbb{E}(X)=\int_\Omega X(\omega)\,d\mathbb{P}(\omega)$是$X$在$\Omega$上的加权平均. 而

$$\mathbb{E}(X|A)=\int_\Omega X(\omega)\,d\mathbb{P}(\omega|A)=\frac{1}{\mathbb{P}(A)}\int_A X(\omega)\,d\mathbb{P}(\omega)$$

是X在A上的加权平均.

性质：(1)

$$\mathbb{E}(X1_A)=\mathbb{E}(X|A)\mathbb{P}(A);$$

(2)若$A_1,A_2,\cdots$是Ω的划分，且$\mathbb{P}(A_i)>0$, 则以下全期望公式成立：

$$\mathbb{E}(X)=\sum_i \mathbb{E}(X|A_i)\mathbb{P}(A_i).$$

例：独立重复地抛一枚硬币直至有k次相继出现正面，计算所抛次数的均值。（设每次出现正面的概率为p）

解：令N_k：所抛次数，T：首次出现反面时抛的次数

$$\text{则} E(N_k \mid T=i)=\begin{cases} i+E(N_k) & \text{如果} i\le k, \\ k & \text{如果} i>k, \end{cases}$$

$$\therefore E(N_k)=\sum_{i\ge 1} E(N_k \mid T=i)P(T=i)$$

$$=\sum_{i=1}^{k}[i+E(N_k)]p^{i-1}q+kp^k \quad \text{这里，} q=1-p$$

$$\text{解得} E(N_k)=\frac{1}{p}+\frac{1}{p^2}\cdots+\frac{1}{p^k}$$

设Y是一取值为$y_1, y_2, \cdots$的离散型随机变量，由于$E(X|Y=y)$是y的函数，所以定义给定Y的条件下X的条件期望$\mathbb{E}(X|Y)$为Y的一个函数，它在$Y=y$时的取值是$E(X|Y=y)$，即

$$\mathbb{E}(X|Y) = \sum_i \mathbb{E}(X|Y=y_i)1_{\{Y=y_i\}} = \begin{cases} a_1, & 若\ Y=y_1 \\ a_2, & 若\ Y=y_2 \\ a_3, & 若\ Y=y_3 \\ \vdots & \vdots \end{cases}$$

这里 $a_i = E(X|Y=y_i)$

它是关于$\sigma(Y)$可测的随机变量。

此时全期望公式$\mathbb{E}(X) = \sum \mathbb{E}(X|Y=y_i)\mathbb{P}(Y=y_i)$即

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

例1.3 设班级有3名男生，数学成绩分别为$70,80,90$分，有两名女生成绩分别为$80,100$分。从班级中任取一人，以X表示该生成绩，Y表示该生性别(0表示男，1表示女). 计算 $\mathbb{E}(X),\mathbb{E}(X|Y)$.

解：

$$\begin{aligned}
\mathbb{E}(X) &= (70+80+90+80+100)/5=84,\\
\mathbb{E}(X|Y=0) &= (70+80+90)/3=80,\\
\mathbb{E}(X|Y=1) &= (80+100)/2=90,\\
\mathbb{E}(X|Y) &= \begin{cases} 80, & \text{如果}Y=0;\\ 90, & \text{如果}Y=1.\end{cases}
\end{aligned}$$

确实有$\mathbb{E}[\mathbb{E}(X|Y)]=80\times\frac{3}{5}+90\times\frac{2}{5}=84=\mathbb{E}(X)$.

定义: 如果X, Y都是随机变量，且X可积，

(1)如果Y离散型, 所有取值为$y_1, y_2, \cdots$, 则 **定义**

$$\mathbb{E}(X|Y)(\omega) = \sum_{n=1}^{\infty} \mathbb{E}(X|Y = y_n) 1_{\{Y=y_n\}}(\omega).$$

(2)如果(X, Y)是二维连续型随机变量具有概率密度$f(x, y)$,

并且$g(X)$是可积的随 机变量, **则定义** $\mathbb{E}(g(X)|Y) = h(Y)$, 这里

$$h(y) = \begin{cases} \int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)\, dx, & f_Y(y) > 0; \\ 0, & f_Y(y) = 0. \end{cases}$$

条件期望的性质：（假设下面的随机变量都是可积的。）

(1) 线性性：对常数$a,b,c,$

$$E(aX+bY+c\,|\,Z)=aE(X\,|\,Z)+bE(Y\,|\,Z)+c$$

(2) 单调性：如果$X\le Y,$那么 $E(X\,|\,Z)\le E(Y\,|\,Z)$

(3) 如果$Y=h(X)$是X的可测函数，则 $E(Y\,|\,X)=Y.$

(4) 如果$Y=h(X)$是X的可测函数，则 $E(YZ\,|\,X)=YE(Z\,|\,X).$

(5) 如果X与Y相互独立，则 $E(Y\,|\,X)=E(Y)$

(6) 全期望公式：$E[E(Y\,|\,X)]=E(Y).$

例1.4 设$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 计算$\mathbb{E}(X|Y)$.

解： 方法一：

$Y=y$时,X 服从$N(\mu_1+\rho\sigma_1(y-\mu_2)/\sigma_2,\sigma_1^2(1-\rho^2))$,

所以 $\mathbb{E}(X|Y)=\mu_1+\rho\sigma_1(Y-\mu_2)/\sigma_2$

方法二：令$Z=X-aY$, a为一个待定常数，使得Z与Y独立.

由正态分布性质知, Z与Y独立当且仅当$\mathrm{Cov}(Z,Y)=0$,

即$a=\rho\sigma_1/\sigma_2$. 此时,

$$\mathbb{E}(X|Y)=\mathbb{E}(Z+aY|Y)=\mathbb{E}(Z)+aY$$

$$=\mu_1+\rho\sigma_1(Y-\mu_2)/\sigma_2.$$

例：设$N \sim B(n,p)$，在$N=k$的条件下X服从$B(k,q)$，

这里$0 \le p,q \le 1$.求X的分布律.

解：对$0 \le m \le n$，由全概率公式，

$$P(X=m)=\sum_{k=0}^{n} P(X=m \mid N=k)P(N=k)$$

$$=\sum_{k=m}^{n} C_k^m q^m (1-q)^{k-m} C_n^k p^k (1-p)^{n-k}$$

$$=C_n^m (pq)^m \underline{\sum_{k=m}^{n} C_{n-m}^{k-m} [p(1-q)]^{k-m} (1-p)^{n-k}} = \sum_{i=0}^{n-m} C_{n-m}^{i} [p(1-q)]^i (1-p)^{n-m-i}$$

$$=C_n^m (pq)^m (1-pq)^{n-m}$$

这说明$X \sim B(n,pq)$

另解：计算X的特征函数：

$$f(t)=E(e^{itX})=E[E(e^{itX}\mid N)]$$

$$E(e^{itX}\mid N=k)=(qe^{it}+1-q)^k$$

$$\therefore E(e^{itX}\mid N)=(qe^{it}+1-q)^N$$

$$\therefore f(t)=E[(qe^{it}+1-q)^N]$$

$$=[p(qe^{it}+1-q)+1-p]^n=(pqe^{it}+1-pq)^n$$

这说明$X\sim B(n,pq)$

Gamma分布

X服从参数为$\alpha > 0$，$\beta > 0$的Gamma分布，

$$X \sim \Gamma(\alpha, \beta) \equiv \text{Gamma}(\alpha, \beta)$$

是指具有下面的概率密度函数：

$$f(x; \alpha, \beta) = \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad \text{for } x > 0 \text{ and } \alpha, \beta > 0$$

这里$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$

$Gamma$分布的性质：

1. 如果 $X \sim \Gamma(\alpha, \beta)$ ，那么$E(X) = \dfrac{\alpha}{\beta}$

2. $\Gamma(1, \beta)$就是参数为β的指数分布

3. 如果$X_1, \ldots, X_n$相互独立，且$X_i \sim \Gamma(\alpha_i, \beta)$，
那么$X_1 + \ldots + X_n \sim \Gamma(\alpha_1 + \ldots + \alpha_n, \beta)$

全概率公式：

1. X 离散型

$$P(A)=\sum_i P(A\mid X=x_i)\,P(X=x_i)$$

2. X 连续型

$$P(A)=\int_{-\infty}^{+\infty}P(A\mid X=x)\,f_X(x)\,dx$$

例： 设$\lambda \sim Gamma(\alpha,1)$,在$\lambda = x$的条件下$X$服从$Poi(x)$,

求X的分布律和EX.

解： 对非负整数k,由全概率公式,

$$P(X=k)=\int_0^{\infty} P(X=k\mid\lambda=x)f_\lambda(x)dx$$

$$=\int_0^{\infty}(e^{-x}\frac{x^k}{k!})(\frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)})dx$$

$$=\frac{1}{k!\Gamma(\alpha)}\int_0^{\infty}x^{k+\alpha-1}e^{-2x}dx$$

$$= \frac{1}{k!\Gamma(\alpha)} \frac{\Gamma(k+\alpha)}{2^{k+\alpha}} \int_0^{\infty} \frac{2^{k+\alpha}}{\Gamma(k+\alpha)} x^{k+\alpha-1} e^{-2x} dx$$

$$= \frac{1}{k!\Gamma(\alpha)} \frac{\Gamma(k+\alpha)}{2^{k+\alpha}}$$

$$E(X \mid \lambda = x) = x \qquad \Rightarrow E(X \mid \lambda) = \lambda$$

$$\therefore E(X) = E[E(X \mid \lambda)] = E(\lambda) = \alpha$$

例1.5 设$Y_1, Y_2, \cdots$独立同分布，均值为μ，方差为σ^2.

设N是取非负整数值的随机变量,与$Y_1, Y_2, \cdots$独立,

且$\mathbb{E}(N^2) < \infty$. 令$X = Y_1 + \cdots + Y_N$，计算$\mathbb{E}(X)$和Var(X).

解： 对非负整数n,由于N与$Y_1, Y_2, \cdots, Y_n$独立,所以

$$\mathbb{E}(X|N = n) = \mathbb{E}(Y_1 + \cdots + Y_N|N = n)$$

$$= \mathbb{E}(Y_1 + \cdots + Y_n|N = n) = \mathbb{E}(Y_1 + \cdots + Y_n) = n\mu.$$

因此$\mathbb{E}(X|N) = N\mu$,

$$\mathbb{E}(X) = \mathbb{E}\left(\mathbb{E}(X|N)\right) = \mathbb{E}(N\mu) = \mu\mathbb{E}(N)$$

$$E(X^2 \mid N=n)=E((Y_1+\ldots+Y_N)^2 \mid N=n)$$

$$=E((Y_1+\ldots+Y_n)^2 \mid N=n)=E((Y_1+\ldots+Y_n)^2)$$

$$=Var(Y_1+\ldots+Y_n)+|[E(Y_1+\ldots+Y_n)]^2=n\sigma^2+n^2\mu^2$$

$$\therefore E(X^2 \mid N)=N\sigma^2+N^2\mu^2$$

$$\therefore E(X^2)=E[E(X^2 \mid N)]=E[N\sigma^2+N^2\mu^2]$$

$$=\sigma^2 E(N)+\mu^2 E(N^2)$$

$$\therefore Var(X)=E(X^2)-[E(X)]^2=\sigma^2 E(N)+\mu^2 Var(N)$$

例 1.6. 设$N, X_1, X_2, \cdots$相互独立，$N \sim Poi(\lambda)$,

$X_i \sim B(1, p), i \geq 1$.令$Y = \sum_{i=1}^{N} X_i$. 求$Y$的分布.

解法1: $P(Y = m) = \sum_{n=m}^{\infty} P(Y = m \mid N = n) P(N = n)$

$$= \sum_{n=m}^{\infty} \binom{n}{m} p^m (1-p)^{n-m} e^{-\lambda} \frac{\lambda^n}{n!} = \frac{e^{-\lambda}(p\lambda)^m}{m!} \sum_{n=m}^{\infty} \frac{(\lambda(1-p))^{n-m}}{(n-m)!}$$

$$= \frac{e^{-\lambda p}(p\lambda)^m}{m!}$$ 所以$Y \sim Poi(\lambda p)$

解法2：令$\phi_Y(t)=E(e^{itY}),\phi_{X_1}(t)=E(e^{itX_1})$,

$$E(e^{itY}\mid N=n)=E(e^{it(X_1+\cdots X_n)}\mid N=n)=E(e^{it(X_1+\cdots X_n)})$$

$$=[\phi_{X_1}(t)]^n=(1-p+pe^{it})^n$$

$$E(e^{itY}\mid N)=(1-p+pe^{it})^N$$

$$\phi_Y(t)=E(1-p+pe^{it})^N=\sum_{n=0}^{\infty}(1-p+pe^{it})^n e^{-\lambda}\frac{\lambda^n}{n!}$$

$$=e^{\lambda(1-p+pe^{it})}e^{-\lambda}=e^{\lambda p(e^{it}-1)}$$

所以$Y\sim Poi(\lambda p)$

例1.7： 设$\xi \sim U(0,1)$,在$\lambda = x$的条件下X服从$B(n,x)$,

求X的分布律.

解法1: 对整数$0 \le k \le n$,由全概率公式,

$$P(X=k)=\int_{-\infty}^{\infty} P(X=k \mid \xi = x) f(x)dx$$

$$=\int_0^1 \binom{n}{k} x^k (1-x)^{n-k} dx = \binom{n}{k} \frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1}$$

解法2: 令$\phi(t)=E(e^{itX}),\quad E(e^{itX}\mid\xi=x)=(1+x(e^{it}-1))^n$

$$E(e^{itX}\mid\xi)=(1+\xi(e^{it}-1))^n$$

$$\phi(t)=E(1+\xi(e^{it}-1))^n=\int_0^1(1+x(e^{it}-1))^n dx$$

$$=\frac{(1+x(e^{it}-1))^{n+1}}{(n+1)(e^{it}-1)}\Bigg|_0^1=\frac{e^{it(n+1)}-1}{(n+1)(e^{it}-1)}$$

$$=\frac{1}{n+1}(1+e^{it}+e^{2it}+\cdots e^{nit})\ \therefore P(X=k)=\frac{1}{n+1},0\le k\le n$$