第三章：数字特征

1.1 数学期望的定义

回顾:

1. 随机现象可由概率空间$(\Omega, \mathcal{F}, P)$描述;

2. 随机事件的规律通过随机变量的分布来刻画：

包括取值和取每个值的概率大小

问题：

(1) 有时随机变量的分布很难获得，特别是取每个值的概率大小

(2) 即使知道了随机变量的分布，也不容易把握随机变量的特征。

- 需要研究随机变量的数字特征：平均值，方差，围绕平均值的波动等

- 平均值

(1) 算术平均

考虑n个实数$x_1, x_2, \cdots, x_n$。算术平均值为

$$\frac{1}{n}(x_1 + \cdots + x_n)$$

(2) 加权平均

考虑n个实数$x_1, x_2, \cdots, x_n$;

加权系数$\alpha_1, \alpha_2, \cdots, \alpha_n \geq 0 : \sum_{k=1}^{n} \alpha_k = 1$

加权平均值为

$$\sum_{k=1}^{n} x_k \alpha_k$$

应用：

体育比赛成绩计算

前3个是王宗源的，扣除一个最高分8.5，一个最低分8.0，剩余一个8.0

后3个是谢思埸的，扣除1个最高和1个最低，还剩一个9.0

Synchronisation那一行共5个分数：扣除1个最高分9.5，扣除1个最低分8.5，还剩3个分数：9.0, 9.5, 9.0.

上面剩下的5个分数：8.0，9.0，9.0，9.5，9.0。

相加得：44.5; 除以5乘以3得：26.7

Difficulty表示难度分，这里是3.4; 用$\underline{26.7 \times 3.4 = 90.78}$

- 大学期末成绩总平均分怎么计算?

公式:

总平均分=Σ（考分乘以学分）/ Σ学分

例：假如这个学期有5门课,学分分别是：2、3、3、2、1, 而期末考试的成绩分别为：75、80、78、86、90;

那么

$$\frac{75\times 2+80\times 3+78\times 3+86\times 2+90\times 1}{2+3+3+2+1}=88.36$$

- 公务员绩效工资计算

………

- 平均值

(3) 概率平均

$(\Omega, \mathcal{F}, P)$ 概率空间

$X: \Omega \to R$ 随机变量

X的平均值是随机变量的取值按<u>概率大小进行平均</u>

以下详细介绍。

- (1) 离散型随机变量

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_N \\ p_1 & p_2 & \cdots & p_N \end{pmatrix}$$

数学期望

$$EX = \sum_{k=1}^{N} x_k p_k$$

1.2 例子

1. 退化变量

$$X \sim \begin{pmatrix} c \\ 1 \end{pmatrix}$$

那么

$$EX = c \cdot P(X = c) = c$$

2. 两点分布

$$X \sim \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$$

那么

$$EX = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = p$$

3. 二项Bernoulli分布

$$X \sim B(n,p), \quad P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0,1,2,\cdots,n$$

那么

$$\begin{aligned} EX &= \sum_{k=0}^{n} kP(X=k) \\ &= \sum_{k=0}^{n} k\binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^{n} k\binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= np \end{aligned}$$

4. Poisson 分布

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

那么

$$\begin{aligned} EX &= \sum_{k=0}^{\infty} kP(X=k) \\ &= \sum_{k=0}^{\infty} k\frac{\lambda^k e^{-\lambda}}{k!} \\ &= \sum_{k=1}^{\infty} k\frac{\lambda^k e^{-\lambda}}{k!} \\ &= \lambda \end{aligned}$$

平均值等于参数

$$EX = \lambda$$

5. 几何分布

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

那么

$$\begin{aligned} EX &= \sum_{k=1}^{\infty} kP(X=k) \\ &= \sum_{k=1}^{\infty} kp(1-p)^{k-1} \\ &= p\sum_{k=1}^{\infty} k(1-p)^{k-1} \\ &= \frac{1}{p} \end{aligned}$$

- 绝对可求和

考虑随机变量

$$P(X=(-1)^{k+1}k)=\frac{c}{k^2},\qquad k=1,2,\cdots,\qquad c=\frac{6}{\pi^2}$$

编号(1). 令$x_k=(-1)^{k+1}k,\quad p_k=\frac{c}{k^2}$。那么

$$\begin{aligned}EX &= \sum_{k=1}^{\infty}x_kp_k\\ &= c\sum_{k=1}^{\infty}(-1)^{k+1}k\cdot\frac{1}{k^2}\\ &= c\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}\\ &= c\ln 2\end{aligned}$$

- 莱布尼茨判别法

定理：如果交错级数$\sum_{k=1}^{\infty}(-1)^{k+1}u_k$ 满足以下两个条件：

（1）$u_k \geq 0$; 数列$\{u_k\}$单调递减；

（2）$\lim_{k\to\infty} u_k = 0$.

那么该交错级数收敛，且其和满足$\sum_{k=1}^{\infty}(-1)^{k+1}u_k \leq u_1$

绝对收敛：如果$\sum_{k=1}^{\infty}|u_k| < \infty$, 那么称$\sum_{k=1}^{\infty}u_k$绝对收敛。

条件收敛：$\sum_{k=1}^{\infty}u_k$收敛，但不是绝对收敛，则称$\sum_{k=1}^{\infty}u_k$ 条件收敛。

条件收敛和绝对收敛的区别：<u>重排不同</u>

1、条件收敛：条件收敛任意重排后所得的级数非条件收敛，且有不相同的和数。

2、绝对收敛：绝对收敛任意重排后所得的级数也绝对收敛，且有相同的和数。

注：$\sum_{k=1}^{\infty} x_k p_k = c\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}$不是绝对收敛的。

编号(2). 令

$$x_1 = 1, x_2 = 3, x_3 = -2, x_4 = -4, x_5 = 5, x_6 = 7,$$

$$x_7 = -6, x_8 = -8, \cdots$$

$$p_1 = c, p_2 = \frac{c}{3^2}, p_3 = \frac{c}{2^2}, p_4 = \frac{c}{4^2}, p_5 = \frac{c}{5^2}, p_6 = \frac{c}{7^2},$$

$$p_7 = \frac{c}{6^2}, p_8 = \frac{c}{8^2}, \cdots$$

那么

$$\begin{aligned} EX &= \sum_{k=1}^{\infty} x_k p_k \\ &= c\left[1 + \frac{1}{3} - \frac{1}{2} - \frac{1}{4} + \frac{1}{5} + \frac{1}{7} - \frac{1}{6} - \frac{1}{8} + \cdots\right] \\ &= ? \end{aligned}$$

- 同一个随机变量取值如果编号不同，会引起平均值不同。
- 为避免不唯一或混乱，要求

$$\sum_{k=1}^{\infty} x_k p_k$$

绝对可和，即$\sum_{k=1}^{\infty} |x_k| p_k < \infty$

- (2) 连续型随机变量

假设$X \sim p(x)$，

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

正如离散型一样，需要绝对可积性

绝对可积:

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

- 1. 均匀分布

$$X \sim U(a,b), \quad a < b$$

那么

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} xp(x)dx \\ &= \int_a^b x\frac{1}{b-a}dx \\ &= \frac{a+b}{2} \end{aligned}$$

平均值等于区间中点

- 2. 正态分布

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

那么

$$\begin{aligned} \int_{-\infty}^{\infty} |x| p(x) dx &= 2\int_0^{\infty} x p(x) dx \\ &= 2\int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \frac{\sqrt{2}}{\sqrt{\pi}} \end{aligned}$$

X的密度函数绝对可积. 因此,

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} x p(x) dx \\ &= \int_{-\infty}^{\infty} x p(x) dx = 0 \end{aligned}$$

- 3. 指数分布

$$X \sim \exp(\lambda), \quad p(x) = \lambda e^{-\lambda x}, \quad x > 0$$

那么

$$\begin{aligned} \int_{-\infty}^{\infty} |x| p(x) dx &= \int_0^{\infty} x \lambda e^{-\lambda x} dx \\ &= \frac{1}{\lambda} \int_0^{\infty} x e^{-x} dx \\ &= \frac{1}{\lambda} \end{aligned}$$

X的密度函数绝对可积。因此，

$$EX = \frac{1}{\lambda}$$

1.3 数学期望的性质

1.

$$a \le X \le b \quad \Rightarrow \quad a \le EX \le b$$

2. 线性运算

$$E(a+bX) = a + bEX$$

从求和或求积分的运算性质容易得到。

3. 加法定理

$$\underbrace{E(X+Y)=EX+EY}$$

左右两边同时存在

证明：以连续型为例。

假设$(X,Y)\sim p(x,y)$，那么$Z=:X+Y$具有密度

$$p_Z(z)=\int_{-\infty}^{\infty}p(x,z-x)dx$$

所以

$$E(X+Y)=\int_{-\infty}^{\infty}zp_Z(z)dz$$

$$
\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty} z p_Z(z) dz \\
&= \int_{-\infty}^{\infty} z \int_{-\infty}^{\infty} p(x, z-x) dx dz \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) p(x,y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x p(x,y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y p(x,y) dx dy \\
&= \int_{-\infty}^{\infty} x p_X(x) dx + \int_{-\infty}^{\infty} y p_Y(y) dx \\
&= EX + EY
\end{aligned}
$$

推广：

$$
E(a_1 X_1 + a_2 X_2 + \cdots + a_m X_m) = a_1 E X_1 + a_2 E X_2 + \cdots + a_m E X_m
$$

- 加法定理的应用：

例1. 假设$X_1, X_2, \cdots, X_m$是非负、独立同分布的随机变量，求

$$E\frac{X_1+\cdots+X_k}{X_1+\cdots+X_m}$$

解：

$$E\frac{X_1}{X_1+\cdots+X_m}=\cdots=E\frac{X_m}{X_1+\cdots+X_m}$$

存在，有限。另外，

$$\begin{aligned}
1 &= E\frac{X_1+\cdots+X_m}{X_1+\cdots+X_m}\\
&= E\frac{X_1}{X_1+\cdots+X_m}+\cdots+E\frac{X_m}{X_1+\cdots+X_m}\\
&= m\cdot E\frac{X_1}{X_1+\cdots+X_m}
\end{aligned}$$

所以，

$$E\frac{X_1}{X_1+\cdots+X_m}=\frac{1}{m}$$

$$\Downarrow$$

$$E\frac{X_1+\cdots+X_k}{X_1+\cdots+X_m}=\frac{k}{m}$$

例2. 假设$S_n \sim B(n,p)$。令

$$X_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases} \qquad i = 1, 2, \cdots, n, \ \underline{\text{i.i.d.}}$$

那么

$$S_n = \sum_{i=1}^{n} X_i \quad \text{独立随机变量和}$$

$$E(S_n) = \sum_{i=1}^{n} E(X_i) = np$$

例3. 考虑超几何分布随机变量。

令N是产品总数，M是次品数，现抽取n件产品检查，其中$n \le M$。

令S_n表示n件抽查产品中次品的个数。那么

$$P(S_n = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

下面给出X的另一种表示：令ξ_i表示第i-次抽检时次品个数，

$$X_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases} \qquad i = 1, 2, \cdots, n, \underline{\text{同分布，但是不独立}}$$

那么

$$S_n = \sum_{i=1}^{n} X_i \qquad \text{随机变量和}$$

$$ES_n = \sum_{i=1}^{n} EX_i = n\frac{M}{N}$$

1.4 随机变量函数的数学期望

假设$(\Omega, \mathcal{F}, P)$是一个概率空间，$X : \Omega \to \mathbb{R}$ —r.v.

$$f : \mathbb{R} \to \mathbb{R} \qquad \text{实值可测函数}$$

求$Ef(X) =?$

(1) 假设X 是离散型随机变量，

$$P(X = x_k) = p_k, k = 1, 2, \cdots, N,$$

$N < \infty$ 或$N = \infty$

那么

$$Ef(X) = \sum_{k=1}^{N} f(x_k)p_k$$

证明：令$Y=f(X)$，按定义计算EY。
Y的取值为$\{f(x_k),k=1,2,\cdots,N\}$。
重新编号记为$\{y_1,y_2,\cdots,y_M\}$。令

$$A_l =: \{k : f(x_k)=y_l\}, \quad 1\le l\le M$$

$$P(Y=y_l)=\sum_{k\in A_l} p_k$$

$$\begin{aligned} EY &= \sum_{l=1}^{M} y_l P(Y=y_l) = \sum_{l=1}^{M} y_l \sum_{k\in A_l} p_k \\ &= \sum_{l=1}^{M}\sum_{k\in A_l} f(x_k)p_k = \sum_{k=1}^{N} f(x_k)p_k \end{aligned}$$

(2) 假设X 是连续型随机变量，

$$X \sim p(x), \quad -\infty < x < \infty$$

那么

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

(3) 假设X具有分布函数$F(x)$，那么

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

2. 方差

2.1 方差的定义

考察两个学生的成绩：

(1)

$$75, 80, 80, 85, 80$$

(2)

$$55, 90, 60, 100, 85$$

注:

(1) 两个学生的<u>平均成绩都是80</u>

(2) 第一个学生成绩波动很小；第二个学生成绩波动很大

波动大不一定是“坏事”。

除退化变量外，一般随机变量都取多个不同的值
用平均值单个数字去刻画随机变量，肯定会有偏差。
如何衡量偏差的大小？

自然地考虑

$$|X - EX| \qquad \text{绝对偏差}$$

这仍是一个随机变量。因此进一步考虑

$$E|X - EX| \qquad \text{平均偏差}$$

然而，绝对值通常很难计算。

普遍采用

$$\left(E(X-EX)^2\right)^{1/2}$$

称为X的标准差。

记

$$Var(X) =: E(X-EX)^2$$

称为X的方差(Variance)。

方差(或者标准差)反映随机变量取值偏离平均值的平均程度。

2.2 方差的计算

方差公示：

$$Var(X) = EX^2 - (EX)^2$$

$$EX^2 = Var(X) + (EX)^2$$

- 例子

1. 退化分布

$$X \sim \begin{pmatrix} c \\ 1 \end{pmatrix}$$

那么

$$EX = c \cdot P(X = c) = 1$$

$$EX^2 = c^2$$

$$Var(X) = EX^2 - (EX)^2 = 0$$

2. 二项Bernoulli分布: $X \sim B(n,p)$

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

$$\begin{aligned}
EX^2 &= \sum_{k=0}^{n} k^2 P(X=k) = \sum_{k=1}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k(k-1)\binom{n}{k} p^k(1-p)^{n-k} + \sum_{k=1}^{n} k\binom{n}{k}p^k(1-p)^{n-k} \\
&= \sum_{k=2}^{n} \frac{n!}{(k-2)!(n-k)}p^k(1-p)^{n-k} + np \\
&= n(n-1)p^2 + np
\end{aligned}$$

$$\begin{aligned}
Var(X) &= EX^2 - (EX)^2 \\
&= n(n-1)p^2 + np - (np)^2 = np(1-p)
\end{aligned}$$

- 3. Poisson 分布: $X \sim \mathcal{P}(\lambda)$

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

$$\begin{aligned}
EX^2 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k e^{-\lambda}}{k!} \\
&= \sum_{k=1}^{\infty} k(k-1)\frac{\lambda^k e^{-\lambda}}{k!} + \sum_{k=1}^{\infty} k\frac{\lambda^k e^{-\lambda}}{k!} \\
&= \sum_{k=2}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-2)!} + \lambda
\end{aligned}$$

$$EX^2=\lambda^2+\lambda$$

$$Var(X)=EX^2-(EX)^2=\lambda^2+\lambda-(\lambda)^2=\lambda$$

- 4. 均匀分布: $X \sim U(a,b)$

$$p(x) = \frac{1}{b-a}, \quad a<x<b$$

$$\begin{aligned} EX^2 &= \int_{-\infty}^{\infty} x^2 p(x)dx \\ &= \int_a^b x^2 \frac{1}{b-a} dx \\ &= \frac{1}{3}(a^2+ab+b^2) \end{aligned}$$

$$\begin{aligned} Var(X) &= EX^2 - (EX)^2 \\ &= \frac{1}{3}(a^2+ab+b^2) - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{(b-a)^2}{12} \end{aligned}$$

- 5. 指数分布: $X \sim exp(\lambda)$

$$p(x) = \lambda e^{-\lambda x}, \quad x > 0$$

$$\begin{aligned} EX^2 &= \int_{-\infty}^{\infty} x^2 p(x) dx \\ &= \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx \\ &= \frac{1}{\lambda^2} \int_0^{\infty} x^2 e^{-x} dx \\ &= \frac{2}{\lambda^2} \end{aligned}$$

$$Var(X) = EX^2 - (EX)^2 = \frac{1}{\lambda^2}$$

- 6. 正态分布: $X \sim N(0,1)$

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

$$\begin{aligned} EX^2 &= \int_{-\infty}^{\infty} x^2 p(x) dx \\ &= \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^2 e^{-\frac{x^2}{2}} dx \\ &= 1 \end{aligned}$$

$$\underline{Var(X) = 1}$$

可以类似证明：当$X \sim N(\mu, \sigma^2)$

$$Var(X) = \sigma^2$$

注：σ越小，方差越小，取值集中在μ附近；

σ越大，方差越大，取值围绕μ散开

2.3 方差的性质

1.

$$Var(X+a) = Var(X), \quad a \text{ 为常数}$$

2.

$$Var(bX) = b^2 Var(X), \quad b \text{ 为常数}$$

3.

$$Var(a+bX) = b^2 Var(X), \quad a, b \text{ 为常数}$$

4.

$$Var(X+Y) = Var(X) + Var(Y) + 2E(X-EX)(Y-EY)$$

5. 当X, Y相互独立时，

$$Var(X+Y) = Var(X) + Var(Y)$$

推广：当$X_1, \cdots, X_m$相互独立时，

$$Var(\sum_{k=1}^{m} X_k) = \sum_{k=1}^{m} Var(X_k)$$

独立和的方差等于方差的和

- 应用：

（1）假设$S_n \sim B(n,p)$。令

$$X_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases} \qquad i = 1, 2, \cdots, n, \underline{\text{ i.i.d.}}$$

那么

$$S_n = \sum_{i=1}^{n} X_i \quad \text{独立随机变量和}$$

$$Var(S_n) = \sum_{i=1}^{n} Var(X_i) = np(1-p)$$

（2）考虑超几何分布随机变量。

令N是产品总数，M是次品数，现抽取n件产品检查，其中$n \le M$。令S_n表示n件抽查产品中次品的个数。那么

$$P(S_n = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

下面给出X的另一种表示：令X_i表示第i-次抽检时次品个数，

$$X_i = \begin{cases} 1, & p \\ 0, & 1-p \end{cases} \qquad i = 1, 2, \cdots, n, \underline{\text{同分布，但是不独立}}$$

那么

$$S_n = \sum_{i=1}^{n} X_i \qquad \text{随机变量和}$$

$$Var(S_n) = Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{N} Var(X_i) + 2\sum_{i<j} \underbrace{E(X_i - EX_i)(X_j - EX_j)}$$

6. 当$c \neq EX$时，

$$Var(X) < E(X-c)^2$$

证明：

$$\begin{aligned}
Var(X) &= E(X-EX)^2 \\
&= E(X-c-(EX-c))^2 \\
&= E(X-c)^2+(EX-c)^2-2E(X-c)(EX-c) \\
&= E(X-c)^2-(EX-c)^2 \\
&< E(X-c)^2, \qquad c \neq EX
\end{aligned}$$

3. Chebyschev不等式

3.1 Chebyschev 切比雪夫（1821 1894）

俄罗斯数学家、力学家。他一生发表了70多篇科学论文，内容涉及数论、概率论、函数逼近论、积分学等方面。他不仅重视纯数学，而且十分重视数学的应用。

生平简介

- 少年时期。切比雪夫出身于贵族家庭. 1832年，全家迁往莫斯科, 父母请了一位当时莫斯科最有名的私人教师。切比雪夫从家庭教师那里学到了很多东西，并对数学产生了强烈的兴趣。他对欧几里得(Euclid)《几何原本》(Elements)当中关于没有最大素数的证明留下了极深刻的印象。
- 大学时代。1837年，进入莫斯科大学物理数学专业。1865年，切比雪夫递交“方程根的计算”的论文，提出一种建立在反函数的级数展开式基础之上的方程近似解法，获得系里设立的银质奖章。毕业后莫斯科大学当助教，同时攻读硕士学位。
- 1843年，切比雪夫通过硕士课程的考试，并在刘维尔的《纯粹与应用数学杂志》(Journal des mathématiques pures et appliquées)上发表了一篇关于多重积分的文章。

1844年，他又在格列尔的同名杂志(Journal für die reine und angewandte Mathematik) 上发表了一篇讨论泰勒级数收敛性的文章。

1845年，他完成了硕士论文“试论概率论的基础分析”，于次年夏天通过了答辩。

- 执教圣彼得堡:

1846年，切比雪夫任圣彼得堡大学助教。他的数学才干得到布尼亚科夫斯基和奥斯特罗格拉茨基这两位数学前辈的赏识。

1847年春天，在题为“关于用对数积分”的晋职报告中，彻底解决了奥斯特罗格拉茨基不久前才提出的一类代数无理函数的积分问题。

1849年，他的博士论文“论同余式”在圣彼得堡大学通过答辩。荣获圣彼得堡科学院的最高数学荣誉奖。

1850年升为副教授，1860年升为教授。

1872年，圣彼得堡堡大学任教25周年之际，学校授予他功勋教授的称号。

- 35年间，切比雪夫教过数论、高等代数、积分运算、椭圆函数、有限差分、概率论、分析力学、傅里叶级数、函数逼近论、工程机械学等十余门课程。

他讲课深受学生们欢迎，李雅普诺夫评论道：“他的课程是精练的，他不注重知识的数量，而是热衷于向学生阐明一些最重要的观念。

他讲解是生动的、富有吸引力的，总是充满了对问题和科学方法之重要意义的奇妙评论。”

1853年，切比雪夫被选为彼得堡科学院候补院士，同时兼任应用数学部主席。

1856年成为副院士。1859年成为院士。

- 切比雪夫曾先后六次出国考察或进行学术交流。

与法国数学界联系甚为密切，曾三次赴巴黎出席法国科学院的年会。

1860年、1871年与1873年分别当选为法兰西科学院、柏林皇家科学院的通讯院士与意大利波隆那科学院的院士。

1877年、1880 年、1893年分别成为伦敦皇家学会、意大利皇家科学院与瑞典皇家科学院的外籍成员。

他是全俄罗斯所有大学的荣誉成员、全俄中等教育改革委员会的成员和圣彼得堡炮兵科学院的荣誉院士。

他圣彼得堡和莫斯科两地数学会的热心支持者。

他发起召开的全俄自然科学家和医生代表大会对于科学界之间的相互了解与科学在民众中的影响起到了很大的作用。

切比雪夫是圣彼得堡数学学派的奠基人和领袖。

3.2 Chebyschev不等式:

$(\Omega, \mathcal{A}, P)$概率空间, $X : \Omega \to \mathbb{R}$随机变量

那么$\varepsilon > 0$,

$$P(|X - EX| > \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

证明：仅就$X \sim p(x)$加以证明。

$$\begin{aligned}
P(|X - EX| > \varepsilon) &= \int_{x:|x-EX|>\varepsilon} p(x)dx \\
&\leq \int_{x:|x-EX|>\varepsilon} \frac{|x - EX|^2}{\varepsilon^2} p(x)dx \\
&\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} |x - EX|^2 p(x)dx \\
&= \frac{Var(X)}{\varepsilon^2}
\end{aligned}$$

推广：f是单调不减严格正函数，那么

$$P(X > \varepsilon) \leq \frac{Ef(X)}{f(\varepsilon)}$$

事实上，

$$P(X > \varepsilon) \leq P(f(X) \geq f(\varepsilon)) \leq \frac{Ef(X)}{f(\varepsilon)}$$

3.3 Chebyschev不等式的应用

(1)假设$X \sim N(\mu, \sigma^2)$，那么

$$P(|X - \mu| > 3\sigma) \leq \frac{\sigma^2}{9\sigma^2} = \frac{1}{9}$$

注：上述误差上界与真实值的大小比较相差甚远。

(2) 假设$S_n \sim B(n, p)$，那么

$$\begin{aligned} P(|\frac{S_n}{n} - p| > \varepsilon) &\leq \frac{Var(\frac{S_n}{n})}{\varepsilon^2} \\ &= \frac{p(1-p)}{n\varepsilon} \end{aligned}$$

(3) 如果$Var(X)=0$，那么

$$P(X=EX)=1$$

证明：不妨设$EX=0$, 需要证明：$P(X=0)=1$.
往证$P(|X|>0)=0$
根据Chebyschev不等式，对任何$\varepsilon>0$

$$P(|X|>\varepsilon)\leq\frac{Var(X)}{\varepsilon^2}=0$$

所以，

$$\begin{aligned}P(|X|>0) &= P(\cup_{n=1}^{\infty}\{|X|>\frac{1}{n}\}) \\ &\leq \sum_{n=1}^{\infty}P(|X|>\frac{1}{n})=0\end{aligned}$$

3. 均值向量和协方差矩阵

在实际问题中，我们经常需要研究随机向量。

假设(X, Y)为2-维随机向量，联合分布函数为$F(x, y)$

作为1-维随机变量，我们可以分别讨论X和Y的期望和方差以及其它矩。

但我们更关心X, Y之间的内在联系。

除了联合分布外，也可以通过数字特征来反映它们之间的联系。

- 均值向量

假设X和Y的数学期望存在，那么令

$$\vec{\mu} = (EX, EY)$$

表示向量均值

- 协方差

假设X和Y的方差存在，即$EX^2 < \infty$，$EY^2 < \infty$。令

$$Cov(X, Y) = E(X - EX)(Y - EY)$$

表示X和Y之间的协方差。

一个方便的计算公式：$Cov(X, Y) = EXY - EXEY$

- Cauchy-Schwarz不等式

$$E|X-EX||Y-EY| \leq \left(E(X-EX)^2E(Y-EY)^2\right)^{1/2}$$

证明：不妨设，$EX=0, EY=0$, 需要证明：

$$E|XY| \leq (EX^2EY^2)^{1/2}$$

注意到，对任何实数t，

$$\begin{aligned} 0 &\leq E(|X|+t|Y|)^2 \\ &= EX^2+2tE(|X|\cdot|Y|)+t^2EY^2, \quad \forall t \in R \end{aligned} \tag{1}$$

所以$EX^2+2tE(|X|\cdot|Y|)+t^2EY^2 \geq 0, \quad \forall t \in R$

$$4[E(|X|\cdot|Y|)]^2 \leq 4EX^2EY^2$$

即

$$[E(|X|\cdot|Y|)]^2 \leq EX^2EY^2$$

$$E(|X|\cdot|Y|) \leq (EX^2EY^2)^{1/2}$$

- 协方差阵

假设X和Y的方差存在，令

$$\Sigma = \begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix}$$

称Σ为协方差阵。

- 1：协方差阵是非负定的，即对任意实数x, y，

$$(x,y)\begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \geq 0$$

事实上，进行展开可得

$$x^2E(X-EX)^2 + 2xyE(X-EX)(Y-EY) + y^2E(Y-EY)^2 \geq 0$$

$$x^2Var(X) + 2xyCov(X,Y) + y^2Var(Y) \geq 0$$

- 2. 如果X, Y独立，那么

$$Cov(X, Y) = 0$$

反过来，$Cov(X, Y) = 0$ 并不意味着X, Y独立.

- 定义: 如果

$$Cov(X, Y) = 0$$

称X, Y是不相关的

显然，

如果X, Y独立，那么X, Y 不相关。

- 3. X, Y 不相关并不意味X, Y独立。

例1. 假设$\theta \sim U(0, 2\pi)$，$X = \sin\theta, Y = \cos\theta$，那么

$$EX = 0, EY = 0, \quad Var(X) = \frac{1}{2}, Var(Y) = \frac{1}{2}$$

并且

$$Cov(X, Y) = 0 \quad \Rightarrow \gamma = 0$$

但明显，

$$X^2 + Y^2 \equiv 1$$

例2. 二元联合正态分布

假设$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$

注意，$X\sim N(\mu_1,\sigma_1^2),\quad Y\sim N(\mu_2,\sigma_2^2)$

为简单起见，先假设$\mu_1=\mu_2=0,\sigma_1^2=\sigma_2^2=1$

$$
\begin{aligned}
Cov(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\frac{1}{2\pi\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)}dxdy\\
&= \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\Big(\int_{-\infty}^{\infty} y\frac{1}{\sqrt{2\pi(1-\rho^2)}}e^{-\frac{1}{2(1-\rho^2)}(y-\rho x)^2}dy\Big)dx\\
&= \rho\int_{-\infty}^{\infty} x^2\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\\
&= \rho
\end{aligned}
$$

一般情况下，

$$Cov(X,Y)=\rho\sigma_1\sigma_2$$

- 相关系数（Correlation coefficient）

假设X, Y为二元随机向量，

$$\vec{\mu} = (EX, EY), \quad \Sigma = \begin{pmatrix} Var(X) & Cov(X,Y) \\ Cov(X,Y) & Var(Y) \end{pmatrix}$$

定义

$$\gamma = \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}}$$

并称作X和Y的相关系数。γ 的大小反映X和Y相关程度。

上述γ也称为皮尔逊相关系数，它是由Pearson从Galton在19世纪80年代提出的一个相似却又稍有不同的想法演变而来的。

性质：

(1)根据Cauchy-Schwarz 不等式，

$$|\gamma| \leq 1$$

(2) 如果$\gamma = 1$，那么存在$t_0 = \left(\frac{Var(X)}{Var(Y)}\right)^{1/2}$

$$P(X = t_0(Y - EY) + EX) = 1$$

即，X和Y是正线性相关的：

$$X = t_0 Y + EX - t_0 EY$$

类似地，如果$\gamma = -1$，那么存在$t_0 = -\left(\frac{Var(X)}{Var(Y)}\right)^{1/2}$

$$P(X = t_0(Y - EY) + EX) = 1$$

即，X和Y是负线性相关的：

$$X = t_0 Y + EX - t_0 EY$$

证明：假设$\gamma = 1$，那么

$$Cov(X,Y) = [Var(X) \cdot Var(Y)]^{1/2}$$

取$t_0 = \left(\frac{Var(X)}{Var(Y)}\right)^{1/2}$，则

$$\begin{aligned}
& E((X-EX)-t_0(Y-EY))^2 \\
&= E(X-EX)^2 - 2t_0 E(X-EX)(Y-EY) + t_0^2 E(Y-EY)^2 \\
&= Var(X) - 2t_0 Cov(X,Y) + t_0^2 Var(Y) = 0
\end{aligned}$$

所以，

$$X - EX = t_0(Y - EY)$$

即

$$\begin{aligned}
X &= t_0(Y-EY) + EX \\
&= t_0 Y + EX - t_0 EY
\end{aligned}$$

(3)

$$\gamma = 0 \Leftrightarrow \text{不相关}$$

事实上，“不相关”解释为“不线性相关”似乎更合理些。

例1（续）. 假设$\theta \sim U(0, 2\pi)$，$X = \sin\theta, Y = \cos\theta$，那么

$$EX = 0, EY = 0, \quad Var(X) = \frac{1}{2}, Var(Y) = \frac{1}{2}$$

并且

$$Cov(X, Y) = 0 \quad \Rightarrow \gamma = 0$$

但明显，

$$X^2 + Y^2 \equiv 1, \qquad (X, Y)\text{在圆周上}$$

- 条件期望和全期望公式

回忆条件分布

(1) 离散型随机变量的条件分布

假设

$$P(X = x_i, Y = y_j) = p_{ij}, \quad i, j = 1, 2, \cdots$$

那么边际分布为

$$X \sim p_{i,\cdot}, \qquad Y \sim p_{\cdot,j}$$

给定$Y = y_j$，X的条件分布为

$$P(X = x_i | Y = y_j) = \frac{p_{ij}}{p_{\cdot,j}}, i = 1, 2, 3, \cdots$$

给定$Y=y_j$，X的条件期望定义为

$$E(X|Y=y_j)=\sum_{i=1}^{\infty}x_iP(X=x_i|Y=y_j)$$

当然要求：

$$\sum_{i=1}^{\infty}|x_i|P(X=x_i|Y=y_j)<\infty$$

否则，认为条件期望不存在。

类似地，给定$X=x_j$，Y的条件期望定义为

$$E(Y|X=x_i)=\sum_{j=1}^{\infty}y_jP(Y=y_j|X=x_i)$$

当然要求：

$$\sum_{j=1}^{\infty}|y_j|P(Y=y_j|X=x_i)<\infty$$

否则，认为条件期望不存在。

例1. 假设$\xi_1, \xi_2, \cdots, \xi_n$独立同分布，

$$P(\xi_k = 1) = p, \quad P(\xi_k = 0) = 1 - p$$

记$S_n = \sum_{k=1}^n \xi_k$.

求$E(\xi_k | S_n = m), \quad 0 \le m \le n$?

$$\begin{aligned} P(\xi_k = 0 | S_n = m) &= \frac{P(\xi_k = 0, S_n = m)}{P(S_n = m)} \\ &= \frac{P(\xi_k = 0) P(S_{n-1} = m)}{P(S_n = m)} \end{aligned} \tag{2}$$

$$\begin{aligned} P(\xi_k = 1 | S_n = m) &= \frac{P(\xi_k = 1, S_n = m)}{P(S_n = m)} \\ &= \frac{P(\xi_k = 1) P(S_{n-1} = m - 1)}{P(S_n = m)} \end{aligned} \tag{3}$$

$$E(\xi_k | S_n = m) = 0 \cdot P(\xi_k = 0 | S_n = m) + 1 \cdot P(\xi_k = 1 | S_n = m)$$

(2) 连续型随机变量的条件分布

假设$(X, Y) \sim p(x, y)$，那么

$$X \sim p_X(x) = \int_{-\infty}^{\infty} p(x, y) dy$$

$$Y \sim p_Y(y) = \int_{-\infty}^{\infty} p(x, y) dx$$

给定$Y = y$，X的条件分布密度

$$P(X = x | Y = y) = \frac{p(x, y)}{p_Y(y)}$$

给定$Y=y$，X的条件期望

$$E(X|Y=y)=\int_{-\infty}^{\infty} xP(X=x|Y=y)dx$$

当然要求

$$\int_{-\infty}^{\infty} |x|P(X=x|Y=y)dx<\infty$$

类似地，给定$X=x$，Y的条件期望

$$E(Y|X=x)=\int_{-\infty}^{\infty} yP(Y=y|X=x)dy$$

例2. 假设$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$。那么，

$$X\sim N(\mu_1,\sigma_1^2),\quad Y\sim N(\mu_2,\sigma_2^2)$$

给定$Y=y$下，

$$X\sim N\Big(\mu_1+\frac{\rho\sigma_1}{\sigma_2}(y-\mu_2),(1-\rho^2)\sigma_1^2\Big)$$

那么

$$E(X|Y=y)=\mu_1+\frac{\rho\sigma_1}{\sigma_2}(y-\mu_2)$$

类似地，给定$X=x$下，

$$Y\sim N\Big(\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1),(1-\rho^2)\sigma_2^2\Big)$$

$$E(Y|X=x)=\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)$$

- 全期望公式

回忆条件期望：

$$P(X=x_i,Y=y_j)=p_{ij},\quad i,j=1,2,\cdots$$

每一个y_j，对应一个条件期望$E(X|Y=y_j)$，即

$$y_j\to E(X|Y=y_j)$$

定义

$$g(y_j)=E(X|Y=y_j)$$

即

$$g(Y)=E(X|Y)$$

它是Y的函数，所以是随机变量。求$Eg(Y)=?$

$$
\begin{aligned}
Eg(Y) &= \sum_{j=1}^{\infty} g(y_j) P(Y = y_j) \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} x_i P(X = x_i | Y = y_j) P(Y = y_j) \\
&= \sum_{i=1}^{\infty} x_i \underbrace{\sum_{j=1}^{\infty} P(X = x_i | Y = y_j) P(Y = y_j)} \\
&= \sum_{i=1}^{\infty} x_i \underbrace{P(X = x_i)} \\
&= EX
\end{aligned}
$$

结论：

$$E(E(X|Y)) = EX$$

类似地，

$$E(E(Y|X)) = EY$$

以上结论对离散, 连续或一般随机变量都成立。

例2(续). $(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$
给定$Y=y$，X的条件期望

$$g(y)=E(X|Y=y)=\mu_1+\frac{\rho\sigma_1}{\sigma_2}(y-\mu_2)$$

因此，

$$g(Y)=\mu_1+\frac{\rho\sigma_1}{\sigma_2}(Y-\mu_2)$$

所以，

$$\begin{aligned}Eg(Y) &= E\Big(\mu_1+\frac{\rho\sigma_1}{\sigma_2}(Y-\mu_2)\Big)\\ &= \mu_1+\frac{\rho\sigma_1}{\sigma_2}(EY-\mu_2)\\ &= EX\end{aligned}$$

例. Suppose the number of children in a family is a random variable X with mean μ, and given $X = n$ for $n \geq 1$, each of the n children in the family is a girl with probability p and a boy with probability $1 - p$. What is the expected number of girls in a family?

解：令G是该家庭中女孩的个数，求EG

$$E(G|X = n) = np$$

$$\begin{aligned} EG &= \sum_{n=1}^{\infty} E(G|X = n)P(X = n) \\ &= \sum_{n=1}^{\infty} npP(X = n) = \mu p \end{aligned}$$

4. 矩

- k 阶矩，k阶中心矩:

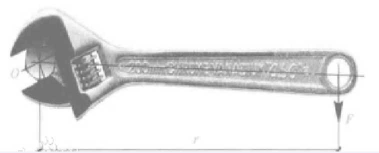
假设$(\Omega, \mathcal{A}, P)$概率空间，$X : \Omega \to \mathbb{R}$ 随机变量

如果$E|X|^k < \infty, k \geq 1$，那么称

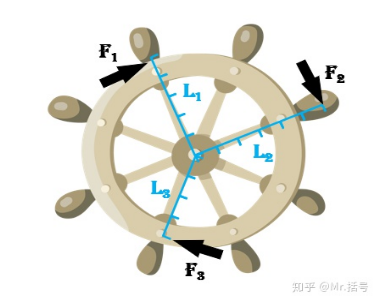
$$EX^k, \qquad k \text{ 阶矩} \quad k \geq 1$$

$$E(X - EX)^k, \qquad k \text{ 阶中心矩} \quad k \geq 1$$

- 力矩(Moment of force)(百度百科)：力对物体产生转动作用的物理量, $M = L \times F$，其中L是从转动轴到着力点的距离矢量，F是矢量力，力矩也是矢量。



“给我一个支点，我就能撬动地球” — 古希腊物理学家阿基米德



$$\text{一阶原点矩} = F_1L_1 + F_2L_2 + F_3L_3$$

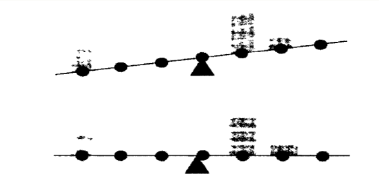
- 如果有一个力，比如F_2，方向相反，可以用二阶矩衡量总能量：

$$\text{二阶原点矩} = (F_1L_1)^2 + (-F_2L_2)^2 + (F_3L_3)^2$$

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_i & \cdots & x_N \\ p_1 & p_2 & \cdots & p_i & \cdots & p_N \end{pmatrix}$$

$$EX = \sum_{i=1}^{N} x_i p_i$$

$$0 = E(X - EX) = \sum_{i=1}^{N} (x_i - EX) p_i$$



$$\sum_{x_i > EX} \underbrace{(x_i - EX) p_i} = \sum_{x_i < EX} \underbrace{(-(x_i - EX)) p_i}$$

例1. $X \sim N(0, \sigma^2)$，那么

$$E|X|^k < \infty, \quad k \geq 1$$

并且

$$EX^{2k} = (2k-1)!!\sigma^{2k}, \qquad EX^{2k+1} = 0, \quad k \geq 1$$

$$\begin{aligned} EX^{2k} &= \int_{-\infty}^{\infty} x^{2k} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}} dx \\ &= -\int_{-\infty}^{\infty} x^{2k-1} \frac{1}{\sqrt{2\pi}} de^{-\frac{x^2}{2}} \\ &= -x^{2k-1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Big|_{-\infty}^{\infty} + (2k-1) \int_{-\infty}^{\infty} x^{2k-2} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}} dx \\ &= (2k-1) EX^{2k-2} \end{aligned}$$

$$EX^{2k} = (2k-1)EX^{2k-2} = \cdots = (2k-1)(2k-3)\cdots 1 = (2k-1)!!$$

例2. $X \sim \mathcal{P}(\lambda)$，那么

$$E|X|^k < \infty, \quad k \geq 1$$

并且

$$E\,\underbrace{X(X-1)\cdots(X-(k-1))} = \lambda^k, \quad k \geq 1$$

$$\Updownarrow$$

$$EX^2 = EX(X-1) + EX$$

$$EX^3 = EX(X-1)(X-2) + 3EX(X-1) + EX$$

$$EX^4 = EX(X-1)(X-2)(X-3) + 6EX(X-1)(X-2) + 7EX(X-1) + EX$$

$$EX^k = EX(X-1)\cdots(X-(k-1)) + \cdots$$

- 问题：随机变量X的分布是否由它的矩所唯一确定？

假设X,Y为随机变量，并且

$$EX^k = EY^k, \qquad k \geq 1$$

问：

$$X \stackrel{d}{=} Y?$$

一般情况下，不能断定$X \stackrel{d}{=} Y$。

例. 假设$\xi \sim N(0,1)$，并且$X = e^{\xi}$. 那么$X \sim p_X(x)$:

$$p_X(x) = \frac{1}{\sqrt{2\pi}x} e^{-\frac{1}{2}(\log x)^2}, \quad x > 0$$

$$p_X(x) = 0, \quad x \leq 0$$

定义$Y \sim p_Y(y)$:

$$p_Y(y) = p_X(y)(1 + \sin(2\pi \log y)), \quad y > 0$$

试证（作为课外练习）

(1) $p_Y(y)$确实是概率密度函数

(2) $EX^k = EY^k, \quad k \geq 0$

定理：假设X, Y是两个随机变量，并且对任意$k \geq 1$，

$$EX^k = EY^k = m_k < \infty$$

如果下列三个条件之一成立：

(i)

$$\sum_{k=1}^{\infty} \frac{m_{2k} t^{2k}}{(2k)!} < \infty, \quad \text{对某个} t > 0$$

(ii)

$$\sum_{k=1}^{\infty} m_{2k}^{-1/2k} = \infty$$

(iii)

$$\limsup_{k \to \infty} |m_k|^{1/k} < \infty$$

那么

$$X \stackrel{d}{=} Y$$

5. 特征函数

$(\Omega, \mathcal{A}, P)$, $X : \Omega \mapsto \mathbb{R}$, $X \sim F_X(x)$. 定义

$$\varphi(t) = Ee^{itX}, \qquad t \in \mathbb{R}$$

其中

$$Ee^{itX} = E\cos tX + iE\sin tX$$

一定存在有限。

$$\varphi(t) : \mathbb{R} \to \mathbb{C}$$

实变量复值函数

目的：利用复分析研究随机变量的分布性质。

意义：对概率论的发展起着重要作用

- 典型分布的特征函数

假设$X \sim F(x)$，那么

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$$

(1) 退化分布

$$P(X = c) = 1$$

那么

$$\varphi(t) = e^{ict}$$

(2) 两点分布

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

那么

$$\varphi(t) = pe^{it} + 1 - p$$

(3) n-Bernoulli 分布

$$X \sim B(n,p)$$

那么

$$\begin{aligned}
\varphi(t) &= Ee^{itX} = \sum_{k=0}^{n} e^{itk} \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} (pe^{it})^k (1-p)^{n-k} \\
&= (1-p+pe^{it})^n
\end{aligned}$$

(4) Poisson 分布

$$X \sim \mathcal{P}(\lambda)$$

那么

$$\begin{aligned}
\varphi(t) &= Ee^{itX} = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k e^{-\lambda}}{k!} \\
&= e^{\lambda e^{it} - \lambda} = e^{\lambda(e^{it}-1)}
\end{aligned}$$

(5) 均匀分布

$$X \sim U(a,b)$$

那么

$$\begin{aligned}\varphi(t) &= Ee^{itX} = \int_a^b e^{itx}\frac{1}{b-a}dx \\ &= \frac{e^{itb}-e^{ita}}{it(b-a)}\end{aligned}$$

(6) 指数分布

$$X \sim exp(\lambda)$$

那么

$$\begin{aligned}\varphi(t) &= Ee^{itX} = \int_0^\infty e^{itx}\lambda e^{-\lambda x}dx \\ &= \int_0^\infty \lambda e^{-(\lambda-it)x}dx = \frac{\lambda}{\lambda-it}\end{aligned}$$

(6) 正态分布

$$X \sim N(0,1)$$

那么

$$\begin{aligned}
&\varphi(t) = Ee^{itX} \\
&= \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
&= \int_{-\infty}^{\infty} \cos tx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx + i \int_{-\infty}^{\infty} \sin tx \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{2k} e^{-\frac{x^2}{2}}\, dx} \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \underbrace{(2k-1)!!} = \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{2^k k!} = e^{-\frac{t^2}{2}}
\end{aligned}$$

5.2 特征函数的分析性质

(1)

$$\varphi(0) = 1$$

(2)

$$|\varphi(t)| \leq 1 = \varphi(0)$$

(3)

$$\varphi(-t) = \overline{\varphi(t)}$$

(4) $\varphi(t)$在$\mathbb{R}$上一致连续。

证明：对任何$\varepsilon > 0$，存在$\delta > 0$，只要$|h| < \delta$，那么对任何$t \in \mathbb{R}$，

$$|\varphi(t+h) - \varphi(t)| < \varepsilon$$

$$\begin{aligned}|\varphi(t+h)-\varphi(t)| &= |Ee^{i(t+h)X}-Ee^{itX}| \\ &= |E(e^{i(t+h)X}-e^{itX})| \\ &= |Ee^{itX}(e^{ihX}-1)| \\ &\leq E|(e^{ihX}-1)|\end{aligned}$$

现在固定$\varepsilon > 0$. 假设$X \sim F(x)$，那么存在$M > 0$

$$P(|X|>M)=1-F(M)+F(-M)\leq \frac{\varepsilon}{4}$$

所以

$$\begin{aligned}|\varphi(t+h)-\varphi(t)| &\leq E|(e^{ihX}-1)|=\int_{-\infty}^{\infty}|e^{ihx}-1|dF(x) \\ &= \int_{|x|>M}|e^{ihx}-1|dF(x)+\int_{|x|\leq M}|e^{ihx}-1|dF(x) \\ &< 2\frac{\varepsilon}{4}+hM\end{aligned}$$

选取$\delta > 0$足够小，使得$\delta M < \frac{\varepsilon}{2}$。这样，当$|h| \leq \delta$时，

$$|\varphi(t+h) - \varphi(t)| < 2\frac{\varepsilon}{4} + hM < \varepsilon$$

(5) Bochner 非负定性

对任何实数$t_1, t_2, \cdots, t_n$，任何复数$a_1, a_2, \cdots, a_n$

$$\sum_{k,l=1}^{n} a_k \bar{a}_l \varphi(t_k - t_l) \geq 0$$

证明：

$$\begin{aligned}\sum_{k,l=1}^{n} a_k \bar{a}_l \varphi(t_k - t_l) &= \sum_{k,l=1}^{n} a_k \bar{a}_l E e^{i(t_k - t_l)X} \\ &= E \sum_{k,l=1}^{n} a_k \bar{a}_l e^{it_k X} \overline{e^{it_l X}} \\ &= E \left| \sum_{k=1}^{n} a_k e^{it_k X} \right|^2 \geq 0\end{aligned}$$

(6) 可微性

引理：令$f(x,t):\mathbb{R}\times[a,b]\to\mathbb{C}$，并且对每一个$t$，$f(\cdot,t):\mathbb{R}\to\mathbb{C}$ 关于函数$p(x)$ 可积. 记

$$F(t)=\int_{\mathbb{R}} f(x,t)p(x)dx$$

(1) 假设存在$g\geq 0$使得$\int_{\mathbb{R}} g(x)p(x)dx<\infty$，并且对所有$x,t$都有$|f(x,t)|<g(x)$. 如果对一个$x$，

$$\lim_{t\to t_0} f(x,t)=f(x,t_0),$$

那么$\lim_{t\to t_0} F(t)=F(t_0)$;

(2) 假设$\frac{\partial f}{\partial t}$存在，存在$g\geq 0$使得$\int_{\mathbb{R}} g(x)p(x)dx<\infty$，并且对所有$x,t$都有$|\frac{\partial f(x,t)}{\partial t}|<g(x)$. 那么$F$ 是可微的，并且

$$F'(t)=\int_{\mathbb{R}} \frac{\partial f(x,t)}{\partial t}p(x)dx.$$

假设$E|X| < \infty$, $EX = \mu$, 那么

$$\varphi(t) \quad \text{可微}$$

并且

$$\varphi'(0) = i\mu$$

事实上，

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x).$$

因为

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

所以

$$\begin{aligned} \varphi'(t) &= \int_{-\infty}^{\infty} ixe^{itx} dF(x) \\ &= i \int_{-\infty}^{\infty} xe^{itx} dF(x) \end{aligned}$$

类似地，如果$E|X|^k<\infty$, 那么

$$\varphi^{(k)}(t)=i^k\int_{-\infty}^{\infty}x^k e^{itx}dF(x)$$

特别，如果$E|X|^k<\infty$，那么$\varphi(t)$ 在0处可以进行k次展开：

$$\begin{aligned}\varphi(t) &= \varphi(0)+\varphi'(0)t+\frac{\varphi''(0)}{2!}t^2+\cdots+\frac{\phi^{(k)}(0)}{k!}+o(t^k)\\ &= 1+iEXt-\frac{EX^2}{2}t^2+\cdots+i^k\frac{EX^k}{k!}t^k+o(t^k), \qquad t\to 0.\end{aligned}$$

- 特征函数的运算性质

(1)令X的特征函数为$\varphi_X(t)$，那么

$$Ee^{it(aX+c)} = e^{itc}\varphi_X(at)$$

作为应用：

如果$Y \sim N(\mu, \sigma^2)$，那么可写成

$$Y = \sigma X + \mu, \qquad X \sim N(0,1)$$

因此，

$$\varphi_Y(t) = e^{i\mu t - \frac{\sigma^2 t^2}{2}}$$

(2) 乘法公式：

如果X, Y相互独立，那么$Z =: X + Y$的特征函数为

$$\varphi_Z(t) = \varphi_X(t) \cdot \varphi_Y(t)$$

证明：

$$\begin{aligned} \varphi_Z(t) &= Ee^{itZ} = Ee^{it(X+Y)} \\ &= Ee^{itX} \cdot e^{itY} \\ &= Ee^{itX} \cdot Ee^{itY} \\ &= \varphi_X(t) \cdot \varphi_Y(t) \end{aligned}$$

推广：

假设$X_1, X_2, \cdots, X_n$相互独立，令

$$Z = X_1 + X_2 + \cdots + X_n$$

那么

$$\varphi_Z(t) = \prod_{k=1}^{n} \varphi_{X_k}(t)$$

应用：

$$S_n \sim B(n,p), \qquad S_n = \sum_{k=1}^{n} \xi_k$$

所以，

$$Ee^{itS_n} = \prod_{k=1}^{n} Ee^{it\xi_k} = (1-p+pe^{it})^n$$

- 唯一性问题：分布函数和特征函数相互唯一确定吗？

更具体地说，假设X, Y是两个随机变量

(1) 显然，如果$X \stackrel{d}{=} Y$ (分布相同)，那么

$$\varphi_X(t) \equiv \varphi_Y(t)$$

(2) 反过来，如果

$$\varphi_X(t) \equiv \varphi_Y(t)$$

那么

$$X \stackrel{d}{=} Y, \qquad F_X(x) \equiv F_Y(x)?$$

- 唯一性定理：如果

$$\varphi_X(t) \equiv \varphi_Y(t)$$

那么

$$X \stackrel{d}{=} Y, \qquad F_X(x) \equiv F_Y(x)$$

事实上，

$$F_X(x_2) - F_X(x_1) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx_2} - e^{-itx_1}}{it} \cdot \varphi_X(t)dt$$

证明：略

推论：

(1) 如果X的特征函数$\varphi(t)$绝对可积，即

$$\int_{-\infty}^{\infty} |\varphi(t)| dt$$

那么X具有密度函数$p(x)$，并且

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt$$

例. $\varphi(t)=e^{-|t|}$是一个特征函数，求相应的随机变量的分布？

显然，$\varphi(t)$绝对可积，即

$$\int_{-\infty}^{\infty} e^{-|t|}dt < \infty$$

所以

$$\begin{aligned} p(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{-|t|}dt \\ &= \frac{1}{2\pi}[\int_{0}^{\infty} e^{-itx}e^{-t}dt + \int_{-\infty}^{0} e^{-itx}e^{t}dt] \end{aligned}$$

得

$$X \sim p(x) = \frac{1}{\pi}\cdot\frac{1}{1+x^2} \qquad \text{Cauchy r.v.}$$

(2) 假设$\varphi(t)$是一个特征函数，如果

$$\varphi(t) = \sum_{k=-\infty}^{\infty} a_k e^{ikt}$$

并且

$$a_k \geq 0, \quad \sum_{k=-\infty}^{\infty} a_k = 1$$

那么

$$P(X = k) = a_k, \quad k = \cdots, -2, -1, 0, 1, 2, \cdots,$$

注意，某些a_k可能为0.

例.

$$\varphi(t) = \cos t$$

那么

$$\varphi(t) = \frac{1}{2}e^{-it} + \frac{1}{2}e^{it}$$

所以,

$$P(X = -1) = \frac{1}{2}, \qquad P(X = 1) = \frac{1}{2}$$

- 利用唯一性可以计算随机变量的分布

例. 假设$X_k \sim N(\mu_k, \sigma_k^2)$, $k = 1, 2, \cdots, n$

令

$$Z = \sum_{k=1}^{n} X_k$$

求Z的分布？

(a) 先求Z的特征函数$\varphi_Z(t)$

$$\begin{aligned}\varphi_Z(t) &= Ee^{itZ} \\ &= \prod_{k=1}^{n} Ee^{itX_k}\end{aligned}$$

$$
\begin{aligned}
\varphi_Z(t) &= \prod_{k=1}^{n} e^{it\mu_k - \frac{\sigma_k^2 t^2}{2}} \\
&= e^{it\sum_{k=1}^{n}\mu_k - \frac{\sum_{k=1}^{n}\sigma_k^2 t^2}{2}}
\end{aligned}
$$

(b) 然后，将所得到的特征函数和熟知的特征函数进行对比，可确定其分布

$$
Z \sim N(\mu, \sigma^2)
$$

其中

$$
\mu = \sum_{k=1}^{n}\mu_k, \quad \sigma^2 = \sum_{k=1}^{n}\sigma_k^2
$$

- 二元随机向量的特征函数

假设(X,Y)是二元随机向量，那么它的特征函数是一个二元函数：

$$\phi(t_1,t_2)=Ee^{i(t_1X+t_2Y)}, \qquad t_1,t_2\in R$$

特别，当X,Y相互独立时，

$$\begin{aligned}\phi(t_1,t_2) &= Ee^{i(t_1X+t_2Y)}\\ &= Ee^{it_1X}Ee^{it_2Y}\\ &= \phi_X(t_1)\phi_Y(t_2)\end{aligned}$$

- 例：假设(X,Y)是二元联合正态随机变量

$$(X,Y)\sim N(\mu_1,\sigma_1^2;\mu_2,\sigma_2^2;\rho)$$

求：$\phi(t_1,t_2)=?$
为简单起见，假设$\mu_1=0,\sigma_1^2=1;\mu_2=0,\sigma_2^2=1$，即

$$(X,Y)\sim N(0,1;0,1;\rho)$$

令

$$\Sigma=\begin{pmatrix}1&\rho\\\rho&1\end{pmatrix}$$

作线性变换：

$$
\begin{aligned}
\begin{pmatrix} U \\ V \end{pmatrix} &= \Sigma^{-1/2} \begin{pmatrix} X \\ Y \end{pmatrix} \\
&= \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}^{-1/2} \begin{pmatrix} X \\ Y \end{pmatrix}
\end{aligned}
$$

这样，$(U, V) \sim N(0, 1; 01; 0)$，即$U, V$相互独立。
所以，

$$
\begin{aligned}
\phi_{U,V}(t_1, t_2) &= e^{-\frac{1}{2}(t_1^2+t_2^2)} \\
&= e^{-\frac{1}{2}(t_1,t_2)\cdot(t_1,t_2)'}
\end{aligned}
$$

并由此得到

$$
\begin{aligned}
\phi_{X,Y}(t_1,t_2) &= Ee^{i(t_1,t_2)(X,Y)'} \\
&= Ee^{i(t_1,t_2)\Sigma^{1/2}(U,V)'} \\
&= e^{-\frac{1}{2}(t_1,t_2)\Sigma(t_1,t_2)'}
\end{aligned}
$$

Fourier 及其Fourier 变换简介

约瑟夫·傅里叶（Joseph Fourier，1768 -1830 ）
法国欧塞尔人，著名数学家、物理学家。

1780年，就读于地方军校。

1795年，任巴黎综合工科大学助教，跟随拿破仑军队远征埃及，成为伊泽尔省格伦诺布尔地方长官。

1817年，当选法国科学院院士。

1822年，担任该院终身秘书，后又任法兰西学院终身秘书和理工科大学校务委员会主席，敕（Chi）封为男爵。

主要贡献是在研究《热的传播》和《热的分析理论》，创立一套数学理论，对19 世纪的数学和物理学的发展都产生了深远影响。

1830年5月16日，在巴黎去世，时年六十三岁。其墓地现位于拉雪兹神父公墓。

小行星10101号傅里叶星

他是名字被刻在埃菲尔铁塔的七十二位法国科学家与工程师其中一位

约瑟夫.傅里叶大学

傅里叶变换

1、傅里叶变换是线性算子，若赋予适当的范数，它还是酉算子。

2、傅里叶变换的逆变换容易求出，而且形式与正变换非常类似。

3、正弦基函数是微分运算的本征函数，从而使得线性微分方程的求解可以转化为常系数的代数方程的傅里叶求解。在线性时不变的物理系统内，频率是个不变的性质，从而系统对于复杂激励的响应可以通过组合其对不同频率正弦信号的响应来获取。

4、著名的卷积定理指出：傅里叶变换可以化复杂的卷积运算为简单的乘积运算，从而提供了计算卷积的一种简单手段。

5、离散形式的傅里叶变换可以利用数字计算机快速的算出(其算法称为快速傅里叶变换算法(FFT))。

正是由于上述的良好性质，傅里叶变换在物理学、数论、组合数学、信号处理、概率、统计、密码学、声学、光学等领域都有着广泛的应用。

傅立叶变换

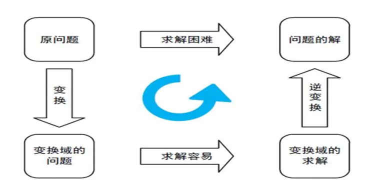
傅立叶变换，表示能将满足一定条件的某个函数表示成三角函数（正弦和/或余弦函数）或者它们的积分的线性组合。

在不同的研究领域，傅立叶变换具有多种不同的变体形式，如连续傅立叶变换和离散傅立叶变换。

最初傅立叶分析是作为热过程的解析分析的工具被提出的。

从数学的角度理解积分变换就是通过积分运算，把一个函数变成另一个函数。也可以理解成是算内积，然后就变成一个函数向另一个函数的投影：

$$F(s)=\int_a^b f(t)K(s,t)dt$$



$K(s,t)$积分变换的核(Kernel)。
当选取不同的积分域和变换核时，就得到不同名称的积分变换。
向核空间投影，将原问题转化到核空间。
下表列出常见的变换及其核函数：

变换名称	核
傅里叶(Fourier)变换	$K(t,\omega)=\frac{1}{\sqrt{2\pi}}e^{-i\omega t}$
拉普拉斯(Laplace)变换	$K(t,s)=e^{-st}$
梅林(Mellin)变换	$K(t,s)=t^{s-1}$
汉科尔(Hankel）变换	$K(t,s)=t\cdot J_v(st)$
魏尔斯特拉斯(Weierstrass) 变换	$K(t,s)=\frac{1}{\sqrt{4\pi}}e^{-\frac{(s-t)^2}{4}}$
阿贝尔(Abel)变换	$K(t,s)=\frac{2t}{\sqrt{t^2-s^2}}$
希尔伯特(Hilbert)变换	$K(t,s)=\frac{1}{\pi}\frac{1}{s-t}$