第 1 章 概率论的基本知识

为了能够顺利阅读后面章节，本章对与本书有关的概率论基本内容作简要介绍.

1.1 概率论的基本概念

定义 1.1 设随机试验 E 的样本空间为 Ω，Ω 的某些子集组成集类$\mathfrak{I}$，称$\mathfrak{I}$为随机试验 E 的事件域，如果它满足下列条件：

(1) $\Omega\in\mathfrak{I}$；

(2) 若 $A\in\mathfrak{I}$，则 $A^c\in\mathfrak{I}$；

(3) 若 $A_j\in\mathfrak{I}$，$j=1,2,\cdots$，则 $\bigcup\limits_{j=1}^{\infty}A_j\in\mathfrak{I}$.

称$\mathfrak{I}$中的元素为随机事件，简称事件.

样本空间 Ω 的全体子集构成的集合就是随机试验 E 的一个事件域.

定义 1.2 对事件域$\mathfrak{I}$中的每个事件 A 赋予一个实数，记为 $P(A)$，称为事件 A 的**概率**，如果集合函数 $P(\cdot)$满足下列条件：

(1) 对任一 $A\in\mathfrak{I}$，$P(A)\geqslant 0$；

(2) $P(\Omega)=1$；

(3) 对两两互不相容事件 $A_1,A_2,\cdots$(即当 $j\neq k$ 时，$A_j\cap A_k=\varnothing$) 有

$$P\left(\bigcup_{j=1}^{\infty}A_j\right)=\sum_{j=1}^{\infty}P(A_j).$$

三元组$(\Omega,\mathfrak{I},P)$称为**概率空间**.

如果定义在样本空间 Ω 上的单实值函数 ξ，对任意实函数 x，$\{\omega\mid\xi(\omega)\leqslant x\}$(简记为$\{\xi\leqslant x\}$)均为事件，即$\{\xi\leqslant x\}\in\mathfrak{I}$，则称 ξ 为随机变量. 称 $F(x)=P\{\xi\leqslant x\}$，$x\in\mathbb{R}$ 为随机变量 ξ 的分布函数.

设 X,Y 为随机变量，施瓦茨(Schwarz)不等式 $E(|XY|)\leqslant\sqrt{E(X^2)}\sqrt{E(Y^2)}$成立.

1.2 随机变量的特征函数

1.2.1 复随机变量

定义 1.3 如果 X 与 Y 都是实随机变量，称 $Z=X+\mathrm{i}Y$ 为复随机变量，其中 $\mathrm{i}=\sqrt{-1}$.

复随机变量 Z 的数学期望定义为$E(Z)=E(X)+\mathrm{i}E(Y)$，其中 $E(X)$，$E(Y)$分别是实随机变量 X,Y 的数学期望.

若 X 是实随机变量,对任意的实数 t,显然 e^{itX} 是复随机变量.

1.2.2 特征函数的定义

定义 1.4 设 X 是实随机变量,则对任意实数 t,有

$\varphi(t)=E(e^{itX})=E(\cos tX+i\sin tX)=E(\cos tX)+iE(\sin tX)$.

称 $\varphi(t)$ 为随机变量 X 的**特征函数**.

设离散型随机变量 X 的分布律为

$$p_k=P\{X=x_k\},\quad k=1,2,\cdots,$$

则 X 的特征函数可表示成

$$\varphi(t)=E(e^{itX})=\sum_{k=1}^{\infty}e^{itx_k}p_k. \tag{1.1}$$

设连续型随机变量 X 的概率密度为 $f(x)$,则 X 的特征函数可表示为

$$\varphi(t)=E(e^{itX})=\int_{-\infty}^{+\infty}e^{itx}f(x)\,dx. \tag{1.2}$$

一般地,设随机变量 X 的分布函数为 $F(x)$,则 X 的特征函数可表示为

$$\varphi(t)=E(e^{itX})=\int_{-\infty}^{+\infty}e^{itx}\,dF(x). \tag{1.3}$$

由(1.2)式可见,连续型随机变量的特征函数 $\varphi(t)$ 是概率密度 $f(x)$ 的傅里叶积分,简称 F 积分.(1.3)式表明随机变量的特征函数 $\varphi(t)$ 是分布函数 $F(x)$ 的傅里叶-斯蒂尔吉斯积分或 F-S 积分.

例 1.1 设离散型随机变量 X 的分布律为

X	-1	0	1
p_k	0.25	0.5	0.25

求 X 的特征函数.

解 $\varphi(t)=E(e^{itX})=0.25e^{-it}+0.5e^{i0t}+0.25e^{it}=0.5\cos t+0.5$.

下面计算一些重要概率分布的特征函数.

1.2.3 重要概率分布的特征函数

例 1.2(二项分布的特征函数) 设随机变量 X 的分布律为 $P\{X=k\}=C_n^k p^k q^{n-k}$,其中 $0<p<1,q=1-p,k=0,1,2,\cdots,n$,求随机变量 X 的特征函数.

解 $\varphi(t)=\sum_{k=0}^{n}e^{itk}C_n^k p^k q^{n-k}=(pe^{it}+q)^n$.

当 $n=1$ 时,X 服从(0-1)分布,其特征函数为

$$\varphi(t)=pe^{it}+q,\quad t\in\mathbb{R}.$$

例 1.3(泊松分布的特征函数) 设随机变量 X 的分布律为 $P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda}$,其中 $\lambda>0,k=0,1,2,\cdots$,求随机变量 X 的特征函数.

解 $\varphi(t)=\sum\limits_{k=0}^{\infty}\mathrm{e}^{\mathrm{i}tk}\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\mathrm{e}^{-\lambda}\sum\limits_{k=0}^{\infty}\dfrac{(\lambda\mathrm{e}^{\mathrm{i}t})^k}{k!}=\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda\mathrm{e}^{\mathrm{i}t}}=\mathrm{e}^{\lambda(\mathrm{e}^{\mathrm{i}t}-1)},\quad t\in\mathbb{R}.$

例 1.4(指数分布的特征函数) 设 X 服从参数为 λ 的指数分布,求 X 的特征函数.

解 X 的概率密度函数为

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

根据特征函数的定义有

$$\begin{aligned}\varphi(t)&=E(\mathrm{e}^{\mathrm{i}tX})=\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x\\&=\int_{0}^{+\infty}\mathrm{e}^{\mathrm{i}tx}\lambda\mathrm{e}^{-\lambda x}\mathrm{d}x=\lambda\int_{0}^{+\infty}\mathrm{e}^{(\mathrm{i}t-\lambda)x}\mathrm{d}x=\frac{\lambda}{\lambda-\mathrm{i}t}=\left(1-\frac{\mathrm{i}t}{\lambda}\right)^{-1},\quad t\in\mathbb{R}.\end{aligned}$$

例 1.5(正态分布的特征函数) 设随机变量 X 服从正态分布 $N(a,\sigma^2)$,其中 $-\infty<a<+\infty,\sigma>0$,求 X 的特征函数.

解 X 的概率密度是

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

故得

$$\begin{aligned}\varphi(t)&=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}\mathrm{e}^{-\frac{(x-a)^2}{2\sigma^2}}\mathrm{d}x\xlongequal{\text{令 }u=\frac{x-a}{\sigma}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}t(a+\sigma u)}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u\\&=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}ta-\frac{\sigma^2t^2}{2}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{(u-\mathrm{i}\sigma t)^2}{2}}\mathrm{d}u\\&=\mathrm{e}^{\mathrm{i}at-\frac{1}{2}\sigma^2t^2}.\end{aligned}$$

特殊地,标准正态分布 $N(0,1)$的特征函数为

$$\varphi(t)=\mathrm{e}^{-\frac{1}{2}t^2},\quad t\in\mathbb{R}.$$

唯一性定理 分布函数 $F(x)$与其特征函数 $\varphi(t)$是一一对应的.

例如,例 1.1 中的特征函数 $\varphi(t)=0.5\cos t+0.5$ 对应的分布律为

X	-1	0	1
p_k	0.25	0.5	0.25

1.2.4 特征函数的性质

性质 1 和性质 6 仅对连续概率分布的情形证明特征函数的性质.

性质 1 $|\varphi(t)|\leqslant\varphi(0)=1$.

证明 $\varphi(0)=E(\mathrm{e}^{\mathrm{i}0X})=1$.

$$|\varphi(t)|=|E(\mathrm{e}^{\mathrm{i}tX})|=\left|\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x\right|\leqslant\int_{-\infty}^{+\infty}|\mathrm{e}^{\mathrm{i}tx}|f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1.$$

性质2 共轭对称性 $\varphi(-t)=\overline{\varphi(t)}$.

证明 $\varphi(-t)=E(\mathrm{e}^{-\mathrm{i}tX})=E[\cos(-tX)]+\mathrm{i}E[\sin(-tX)]$

$$=E(\cos tX)-\mathrm{i}E(\sin tX)=\overline{E(\cos tX)+\mathrm{i}E(\sin tX)}=\overline{\varphi(t)}.$$

性质3 特征函数 $\varphi(t)$ 在区间 $(-\infty,+\infty)$ 上一致连续.

证明 略.

性质4 设随机变量 $Y=aX+b$,其中 a,b 是常数,则

$$\varphi_Y(t)=\mathrm{e}^{\mathrm{i}bt}\varphi_X(at),$$

其中 $\varphi_X(t),\varphi_Y(t)$ 分别表示随机变量 X,Y 的特征函数.

证明 $\varphi_Y(t)=E(\mathrm{e}^{\mathrm{i}tY})=E(\mathrm{e}^{\mathrm{i}t(aX+b)})=\mathrm{e}^{\mathrm{i}bt}E(\mathrm{e}^{\mathrm{i}(at)X})=\mathrm{e}^{\mathrm{i}bt}\varphi_X(at)$.

性质5 设随机变量 X,Y 相互独立,又 $Z=X+Y$,则

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

证明 $\varphi_Z(t)=E(\mathrm{e}^{\mathrm{i}tZ})=E[\mathrm{e}^{\mathrm{i}t(X+Y)}]=E(\mathrm{e}^{\mathrm{i}tX}\cdot\mathrm{e}^{\mathrm{i}tY})=E(\mathrm{e}^{\mathrm{i}tX})E(\mathrm{e}^{\mathrm{i}tY})=\varphi_X(t)\varphi_Y(t)$.

性质6 设随机变量 X 的 n 阶原点矩存在,则它的特征函数 $\varphi(t)$ 的 k 阶导数 $\varphi^{(k)}(t)$ 存在,且有

$$\varphi^{(k)}(0)=\mathrm{i}^kE(X^k)\text{或}E(X^k)=\mathrm{i}^{-k}\varphi^{(k)}(0),\quad k=1,2,\cdots,n.$$

证明 因 $\varphi(t)=E(\mathrm{e}^{\mathrm{i}tX})=\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x$,故

$$\varphi^{(k)}(t)=\frac{\mathrm{d}^k}{\mathrm{d}t^k}\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x$$

$$=\int_{-\infty}^{+\infty}\mathrm{i}^kx^k\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x=\mathrm{i}^k\int_{-\infty}^{+\infty}x^k\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x.$$

因而

$$\varphi^{(k)}(0)=\mathrm{i}^k\int_{-\infty}^{+\infty}x^kf(x)\mathrm{d}x=\mathrm{i}^kE(X^k),\quad k=1,2,\cdots,n.$$

上述推导过程中 $\frac{\mathrm{d}^k}{\mathrm{d}t^k}\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathrm{e}^{\mathrm{i}tx}f(x)\mathrm{d}x$ 成立,需要满足条件

$$\int_{-\infty}^{+\infty}\left|\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathrm{e}^{\mathrm{i}tx}\right|f(x)\mathrm{d}x<+\infty,\quad \text{即}\int_{-\infty}^{+\infty}|\mathrm{i}^kx^k\mathrm{e}^{\mathrm{i}tx}|f(x)\mathrm{d}x<+\infty.$$

事实上,由于 X 的 n 阶原点矩存在,所以 $E(|X|^k)<+\infty,k=1,2,\cdots,n$,从而

$$\int_{-\infty}^{+\infty}\left|\frac{\mathrm{d}^k}{\mathrm{d}t^k}\mathrm{e}^{\mathrm{i}tx}\right|f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}|\mathrm{i}^kx^k\mathrm{e}^{\mathrm{i}tx}|f(x)\mathrm{d}x\leqslant\int_{-\infty}^{+\infty}|x|^kf(x)\mathrm{d}x<+\infty.$$

此性质表明,随机变量的各阶原点矩可由其特征函数在原点的相应阶导数得到.

例1.6 若随机变量 X 和 Y 相互独立,且分别服从参数为 λ_1 和 λ_2 的泊松分布.试用特征函数求随机变量 $Z=X+Y$ 的概率分布.

解 由泊松分布特征函数有 $\varphi_X(t)=\mathrm{e}^{\lambda_1(\mathrm{e}^{\mathrm{i}t}-1)},\varphi_Y(t)=\mathrm{e}^{\lambda_2(\mathrm{e}^{\mathrm{i}t}-1)}$.

利用特征函数性质(5),有

$$\varphi_Z(t)=\mathrm{e}^{\lambda_1(\mathrm{e}^{\mathrm{i}t}-1)}\mathrm{e}^{\lambda_2(\mathrm{e}^{\mathrm{i}t}-1)}=\mathrm{e}^{(\lambda_1+\lambda_2)(\mathrm{e}^{\mathrm{i}t}-1)}.$$

由唯一性定理得随机变量 Z 为具有参数 $\lambda_1+\lambda_2$ 的泊松分布.

例1.7 设 $X_1,X_2,\cdots,X_n$ 相互独立,且 X_j 服从参数为 m_j 和 p 的二项分布,$j=1,2,\cdots,n$. 证明 $Y=\sum_{j=1}^{n}X_j$ 服从参数为 $\sum_{j=1}^{n}m_j$ 和 p 的二项分布.

证明　因为 X_j 服从参数为 m_j 和 p 的二项分布，故其特征函数分别为

$$\varphi_{X_j}(t)=(p\mathrm{e}^{\mathrm{i}t}+q)^{m_j},\quad j=1,2,\cdots,n,$$

由 $X_1,X_2,\cdots,X_n$ 相互独立得 $Y=\sum_{j=1}^{n}X_j$ 的特征函数为

$$\varphi_Y(t)=\prod_{j=1}^{n}\varphi_{X_j}(t)=\prod_{j=1}^{n}(p\mathrm{e}^{\mathrm{i}t}+q)^{m_j}=(p\mathrm{e}^{\mathrm{i}t}+q)^{\sum_{j=1}^{n}m_j}.$$

可见，随机变量 $Y=\sum_{j=1}^{n}X_j$ 服从参数为 $\sum_{j=1}^{n}m_j$ 和 p 的二项分布.

例 1.8　利用特征函数求例 1.1 中随机变量的数学期望、方差和三阶原点矩.

解　因为 $\varphi(t)=E(\mathrm{e}^{\mathrm{i}tX})=0.5\cos t+0.5$，所以

$$\varphi'(t)=-0.5\sin t,\quad \varphi''(t)=-0.5\cos t,\quad \varphi'''(t)=0.5\sin t.$$

从而 $\varphi'(0)=0,\varphi''(0)=-0.5,\varphi'''(0)=0$. 因此，由性质(6)知

$$E(X)=\frac{1}{\mathrm{i}}\varphi'(0)=0,\quad D(X)=\mathrm{i}^{-2}\varphi''(0)-(E(X))^2=0.5,\quad E(X^3)=\frac{1}{\mathrm{i}^3}\varphi'''(0)=0.$$

1.3　多元特征函数和多维正态分布

1.3.1　多维概率分布及其数字特征

我们学过二维随机变量，下面把二维随机变量的一些概念推广到多维随机变量.

1. 分布函数

设 n 维随机变量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$，称

$$F(x_1,x_2,\cdots,x_n)=P\{X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n\},\quad x_1,x_2,\cdots,x_n\in\mathbb{R}$$

为 n 维随机变量 $(X_1,X_2,\cdots,X_n)$ 的 **n 维分布函数**，用向量的形式可表示为 $F(\boldsymbol{x})=P\{\boldsymbol{X}\leqslant\boldsymbol{x}\}$，其中 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$，而 $\boldsymbol{X}\leqslant\boldsymbol{x}$ 理解为 $\boldsymbol{X}$ 对每一个分量都有 $X_j\leqslant x_j$.

2. 数字特征

n 维随机变量 $(X_1,X_2,\cdots,X_n)$ 的**数学期望** $E(\boldsymbol{X})$ 定义为

$$E(\boldsymbol{X})=(E(X_1),E(X_2),\cdots,E(X_n)).$$

$E(\boldsymbol{X})$ 的分量是 $\boldsymbol{X}$ 各分量的数学期望.

n 维随机变量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的**协方差(矩)阵**定义为

$$\boldsymbol{B}=\begin{pmatrix}\mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n)\\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_n)\\ \vdots & \vdots & & \vdots\\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Cov}(X_n,X_n)\end{pmatrix},$$

其中，$\mathrm{Cov}(X_j,X_k)=E[(X_j-E(X_j))(X_k-E(X_k))],\quad j,k=1,2,\cdots,n.$

当 $j\neq k$ 时，$\mathrm{Cov}(X_j,X_k)$ 是随机变量 X_j 和 X_k 的协方差.

当 $j=k$ 时，$\mathrm{Cov}(X_j,X_j)$ 是随机变量 X_j 的方差.

n 维随机向量 $\boldsymbol{X}$ 的协方差阵刻画了它的各个分量概率分布的分散程度，以及各分量之

间线性联系的密切程度,它的主对角线是 $\boldsymbol{X}$ 的各分量的方差.

协方差阵也可以表示为

$$\boldsymbol{B}=E\begin{pmatrix}[X_1-E(X_1)]^2 & [X_1-E(X_1)][X_2-E(X_2)] & \cdots & [X_1-E(X_1)][X_n-E(X_n)]\\ [X_2-E(X_2)][X_1-E(X_1)] & [X_2-E(X_2)]^2 & \cdots & [X_2-E(X_2)][X_n-E(X_n)]\\ \vdots & \vdots & & \vdots\\ [X_n-E(X_n)][X_1-E(X_1)] & [X_n-E(X_n)][X_2-E(X_2)] & \cdots & [X_n-E(X_n)]^2\end{pmatrix}$$

$$=E[\boldsymbol{X}-E(\boldsymbol{X})][\boldsymbol{X}-E(\boldsymbol{X})]^{\mathrm{T}}.$$

定理 1.1　协方差(矩)阵是对称的非负定矩阵.

证明　由协方差的性质 $\mathrm{Cov}(X_j,X_k)=\mathrm{Cov}(X_k,X_j)$ 知,协方差(矩)阵为对称矩阵.

再证协方差矩阵为非负定矩阵.事实上,对于任意 n 个实数 $t_1,t_2,\cdots,t_n$,有

$$\begin{aligned}\sum_{k=1}^{n}\sum_{j=1}^{n}\mathrm{Cov}(X_j,X_k)t_jt_k&=\sum_{k=1}^{n}\sum_{j=1}^{n}E[X_j-E(X_j)][X_k-E(X_k)]t_jt_k\\&=E\left\{\sum_{k=1}^{n}\sum_{j=1}^{n}t_j[X_j-E(X_j)]\cdot t_k[X_k-E(X_k)]\right\}\\&=E\left\{\sum_{j=1}^{n}t_j[X_j-E(X_j)]\right\}^2\geqslant 0,\end{aligned}$$

所以协方差矩阵为非负定矩阵.

1.3.2　多元特征函数及其性质

$n(n>1)$维随机变量的特征函数称为多元特征函数,它是一维随机变量的特征函数的推广.多元特征函数的某些性质与一元特征函数的性质类似.

定义 1.5　设 n 维随机变量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$,称

$$\varphi(t_1,t_2,\cdots,t_n)=E(\mathrm{e}^{\mathrm{i}(t_1X_1+t_2X_2+\cdots+t_nX_n)})$$

为 n 维随机变量 $\boldsymbol{X}$ 的特征函数,其中 $\mathrm{i}=\sqrt{-1}$,$t_1,t_2,\cdots,t_n\in\mathbb{R}$.

记 $\boldsymbol{t}=(t_1,t_2,\cdots,t_n)$,则 n 元特征函数可以简单地表示为

$$\varphi(\boldsymbol{t})=E(\mathrm{e}^{\mathrm{i}\boldsymbol{t}\boldsymbol{X}^{\mathrm{T}}})=E(\mathrm{e}^{\mathrm{i}\boldsymbol{X}\boldsymbol{t}^{\mathrm{T}}}).$$

当 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$是离散型随机变量,$\boldsymbol{X}$ 的特征函数表示为

$$\varphi(t_1,t_2,\cdots,t_n)=\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n}\mathrm{e}^{\mathrm{i}(t_1x_1+t_2x_2+\cdots+t_nx_n)}P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}.$$

当 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$具有概率密度 $f(x_1,x_2,\cdots,x_n)$时,$\boldsymbol{X}$ 的特征函数表示为

$$\varphi(t_1,t_2,\cdots,t_n)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}(t_1x_1+t_2x_2+\cdots+t_nx_n)}f(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n.$$

唯一性定理　n 维随机变量的分布函数与其特征函数是一一对应的.

n 维随机变量的特征函数有以下性质:

(1) $|\varphi(t_1,t_2,\cdots,t_n)|\leqslant\varphi(0,0,\cdots,0)=1$.

(2) $\varphi(-t_1,-t_2,\cdots,-t_n)=\overline{\varphi(t_1,t_2,\cdots,t_n)}$.

(3) $\varphi(t_1,t_2,\cdots,t_n)$在 n 维欧几里得空间$\mathbb{R}^n$上一致连续.

(4) 若 $\varphi(t_1,t_2,\cdots,t_n)$是 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$的特征函数,则 $k(k\in\{1,2,\cdots,n\})$维随机变量$(X_1,X_2,\cdots,X_k)$的特征函数为

$$\varphi_{X_1,X_2,\cdots,X_k}(t_1,t_2,\cdots,t_k)=\varphi(t_1,t_2,\cdots,t_k,0,\cdots,0).$$

此性质表明,如果要获得 n 维随机变量 $\boldsymbol{X}$ 的 k 维边缘概率分布的特征函数,只要在原来的特征函数中保留自变量 $t_1,t_2,\cdots,t_k$,其他自变量置零即可.

(5) 若 $\varphi(t_1,t_2,\cdots,t_n)$ 是 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的特征函数,则随机变量

$$Y=a_1X_1+a_2X_2+\cdots+a_nX_n$$

的特征函数为

$$\varphi_Y(t)=\varphi(a_1t,a_2t,\cdots,a_nt).$$

(6) 若 $\varphi(t_1,t_2,\cdots,t_n)$ 是 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的特征函数,而随机变量 X_j 的特征函数是 $\varphi_{X_j}(t)$,$j=1,2,\cdots,n$,则随机变量 $X_1,X_2,\cdots,X_n$ 相互独立的充分必要条件是

$$\varphi(t_1,t_2,\cdots,t_n)=\varphi_{X_1}(t_1)\varphi_{X_2}(t_2)\cdots\varphi_{X_n}(t_n).$$

(7) 如果矩 $E(X_1^{k_1}X_2^{k_2}\cdots X_n^{k_n})$ 存在,则

$$E(X_1^{k_1}X_2^{k_2}\cdots X_n^{k_n})=\mathrm{i}^{-\sum_{j=1}^{n}k_j}\left[\frac{\partial^{k_1+k_2+\cdots+k_n}\varphi(t_1,t_2,\cdots,t_n)}{\partial t_1^{k_1}\partial t_2^{k_2}\cdots\partial t_n^{k_n}}\right]_{t_1=t_2=\cdots=t_n=0}.$$

1.3.3 n 维正态随机变量及其性质

n 维正态分布在概率论、数理统计和随机过程中占有重要的位置.

若二维随机变量 (X_1,X_2) 的概率密度为

$$f(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-a_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2}+\frac{(x_2-a_2)^2}{\sigma_2^2}\right]\right\},$$
$$-\infty<x_1<+\infty,\ -\infty<x_2<+\infty$$

则称随机变量 (X_1,X_2) 为**二维正态随机变量**,该概率密度称为**二维正态概率密度**,其中

$$a_1=E(X_1),\quad a_2=E(X_2),\quad \sigma_1^2=D(X_1),\quad \sigma_2^2=D(X_2),$$

ρ 是随机变量 X_1 与 X_2 的相关系数.

令 $\boldsymbol{B}=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$,$\boldsymbol{a}=(a_1,a_2)$,$\boldsymbol{x}=(x_1,x_2)$,则有 $|\boldsymbol{B}|=(1-\rho^2)\sigma_1^2\sigma_2^2$,

$$\boldsymbol{B}^{-1}=\frac{1}{(1-\rho^2)\sigma_1^2\sigma_2^2}\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}=\frac{1}{1-\rho^2}\begin{pmatrix}\dfrac{1}{\sigma_1^2} & -\rho\dfrac{1}{\sigma_1\sigma_2}\\ -\rho\dfrac{1}{\sigma_1\sigma_2} & \dfrac{1}{\sigma_2^2}\end{pmatrix},$$

从而

$$(\boldsymbol{x}-\boldsymbol{a})\boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{a})^{\mathrm{T}}=\frac{1}{1-\rho^2}\left[\frac{(x_1-a_1)^2}{\sigma_1^2}-2\rho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2}+\frac{(x_2-a_2)^2}{\sigma_2^2}\right],$$

由此可知二维正态概率密度可以表示为

$$f(\boldsymbol{x})=\frac{1}{2\pi|\boldsymbol{B}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})\boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{a})^{\mathrm{T}}\right].$$

可以将二维正态概率密度推广到 n 维正态概率密度.

定义 1.6 如果 n 维随机变量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的分布密度为

$$f(\boldsymbol{x})=\frac{1}{(2\pi)^{\frac{n}{2}}|\boldsymbol{B}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})\boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{a})^{\mathrm{T}}\right],$$

其中 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$，$\boldsymbol{a}=E(\boldsymbol{X})=(E(X_1),E(X_2),\cdots,E(X_n))$，

$$\boldsymbol{B}=\begin{pmatrix}\mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n)\\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_n)\\ \vdots & \vdots & & \vdots\\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Cov}(X_n,X_n)\end{pmatrix},$$

且矩阵 $\boldsymbol{B}$ 是正定的，称 $\boldsymbol{X}$ 为 n 维正态随机变量. $f(\boldsymbol{x})$ 称为 **n 维正态概率密度**. n 维正态分布记为 $N(\boldsymbol{a},\boldsymbol{B})$.

可以证明，n 维正态分布 $N(\boldsymbol{a},\boldsymbol{B})$ 的特征函数是

$$\varphi(\boldsymbol{t})=\exp\left(\mathrm{i}\boldsymbol{a}\boldsymbol{t}^{\mathrm{T}}-\frac{1}{2}\boldsymbol{t}\boldsymbol{B}\boldsymbol{t}^{\mathrm{T}}\right).$$

定义 1.7 如果 n 维随机变量 $\boldsymbol{X}$ 的特征函数是

$$\varphi(\boldsymbol{t})=\exp\left(\mathrm{i}\boldsymbol{a}\boldsymbol{t}^{\mathrm{T}}-\frac{1}{2}\boldsymbol{t}\boldsymbol{B}\boldsymbol{t}^{\mathrm{T}}\right),$$

其中 $\boldsymbol{a}$ 是 n 维行向量，$\boldsymbol{B}$ 是 n 阶非负定矩阵，那么称 $\boldsymbol{X}$ 的概率分布是 **n 维正态分布** $N(\boldsymbol{a},\boldsymbol{B})$.

当 $|\boldsymbol{B}|=0$，即 $\boldsymbol{B}$ 不可逆时，定义 1.7 也是有意义的，而用定义 1.6 定义 n 维正态分布则要求 $|\boldsymbol{B}|>0$. 所以，用多维特征函数定义 n 维正态分布更为一般. 显然，在 $|\boldsymbol{B}|>0$ 的情况下两种定义方式是等价的. $|\boldsymbol{B}|>0$ 时的 n 维正态分布也称为**非退化的 n 维正态分布**.

下面介绍 n 维正态分布的性质.

性质 1 若 n 维正态随机变量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的 $m(m<n)$ 个分量构成 m 维随机变量 $\widetilde{\boldsymbol{X}}=(X_1,X_2,\cdots,X_m)$，则它是 m 维正态随机变量，且其数学期望为 $\widetilde{\boldsymbol{a}}=(a_1,a_2,\cdots,a_m)$，协方差阵为

$$\widetilde{\boldsymbol{B}}=\begin{pmatrix}\mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_m)\\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_m)\\ \vdots & \vdots & & \vdots\\ \mathrm{Cov}(X_m,X_1) & \mathrm{Cov}(X_m,X_2) & \cdots & \mathrm{Cov}(X_m,X_m)\end{pmatrix}.$$

特别地，当 $m=1$ 时，可得随机变量 X_1 服从正态分布 $N(a_1,D(X_1))$. 一般有随机变量 X_j 服从正态分布 $N(a_j,D(X_j))$，$j=1,2,\cdots,n$.

证明 记 $\widetilde{\boldsymbol{t}}=(t_1,t_2,\cdots,t_m)$，利用特征函数的性质(4)，有

$$f_{X_1,X_2,\cdots,X_m}(t_1,t_2,\cdots,t_m)=f(t_1,t_2,\cdots,t_m,0,\cdots,0)=\mathrm{e}^{\mathrm{i}\boldsymbol{a}\boldsymbol{t}^{\mathrm{T}}-\frac{1}{2}\boldsymbol{t}\boldsymbol{B}\boldsymbol{t}^{\mathrm{T}}}\Big|_{t_{m+1}=\cdots=t_n=0},$$

其中 $\boldsymbol{a}\boldsymbol{t}^{\mathrm{T}}\big|_{t_{m+1}=\cdots=t_n=0}=a_1t_1+a_2t_2+\cdots+a_mt_m=(a_1,a_2,\cdots,a_m)(t_1,t_2,\cdots,t_m)^{\mathrm{T}}=\widetilde{\boldsymbol{a}}\widetilde{\boldsymbol{t}}^{\mathrm{T}}$，

$$\boldsymbol{tBt}^{\mathrm{T}}\big|_{t_{m+1}=\cdots=t_n=0}$$

$$=(t_1,t_2,\cdots,t_m,0,\cdots,0)\begin{pmatrix}\mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n)\\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_n)\\ \vdots & \vdots & & \vdots\\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Cov}(X_n,X_n)\end{pmatrix}\begin{pmatrix}t_1\\ \vdots\\ t_m\\ 0\\ \vdots\\ 0\end{pmatrix}$$

$$=\sum_{k=1}^{m}\sum_{j=1}^{m}\mathrm{Cov}(X_j,X_k)t_jt_k$$

$$=(t_1,t_2,\cdots,t_m)\begin{pmatrix}\mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_m)\\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_m)\\ \vdots & \vdots & & \vdots\\ \mathrm{Cov}(X_m,X_1) & \mathrm{Cov}(X_m,X_2) & \cdots & \mathrm{Cov}(X_m,X_m)\end{pmatrix}\begin{pmatrix}t_1\\ t_2\\ \vdots\\ t_m\end{pmatrix}=\widetilde{\boldsymbol{t}}\widetilde{\boldsymbol{B}}\widetilde{\boldsymbol{t}}^{\mathrm{T}},$$

所以

$$f_{X_1,X_2,\cdots,X_m}(t_1,t_2,\cdots,t_m)=f(t_1,t_2,\cdots,t_m,0,\cdots,0)=\exp\left(\mathrm{i}\widetilde{\boldsymbol{a}}\widetilde{\boldsymbol{t}}^{\mathrm{T}}-\frac{1}{2}\widetilde{\boldsymbol{t}}\widetilde{\boldsymbol{B}}\widetilde{\boldsymbol{t}}^{\mathrm{T}}\right),$$

这是 m 维正态分布的特征函数.

由唯一性定理,m 维正态随机变量 $\widetilde{\boldsymbol{X}}=(X_1,X_2,\cdots,X_m)$服从正态分布 $N(\widetilde{\boldsymbol{a}},\widetilde{\boldsymbol{B}})$.

性质 2 设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$是 n 维正态随机变量,则随机变量 $X_1,X_2,\cdots,X_n$ 相互独立的充分必要条件是它们两两不相关.

证明 必要性. 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立,则其中任意两个随机变量 X_j 与 $X_k(j\neq k)$相互独立,进而 X_j 与 X_k 不相关,所以 $X_1,X_2,\cdots,X_n$ 两两不相关.

充分性. 因为 $\mathrm{Cov}(X_j,X_k)=0,j\neq k$,故 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$的协方差阵为

$$\boldsymbol{B}=\begin{pmatrix}D(X_1) & 0 & \cdots & 0\\ 0 & D(X_2) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & D(X_n)\end{pmatrix},$$

因此 $\boldsymbol{X}$ 的特征函数

$$\varphi(t_1,t_2,\cdots,t_n)=\exp\left(\mathrm{i}\sum_{j=1}^{n}a_jt_j-\frac{1}{2}\sum_{j=1}^{n}D(X_j)t_j^2\right)=\prod_{j=1}^{n}\exp\left(\mathrm{i}a_jt_j-\frac{1}{2}D(X_j)t_j^2\right)=\prod_{j=1}^{n}f_{X_j}(t_j).$$

由多元特征函数的性质(6)得随机变量 $X_1,X_2,\cdots,X_n$ 相互独立.

性质 3 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$服从 n 维正态分布 $N(\boldsymbol{a},\boldsymbol{B})$当且仅当对任意常数 $l_1,l_2,\cdots,l_n$,随机变量 $Y=\sum_{j=1}^{n}l_jX_j$ 服从一维正态分布

$$N\left(\sum_{j=1}^{n}l_ja_j,\sum_{j=1}^{n}\sum_{k=1}^{n}l_jl_k\mathrm{Cov}(X_j,X_k)\right).$$

证明 充分性. 令 $\boldsymbol{l}=(l_1,l_2,\cdots,l_n)$,$\boldsymbol{a}=(a_1,a_2,\cdots,a_n)$,若

$$Y=\sum_{j=1}^{n}l_jX_j\sim N\left(\sum_{j=1}^{n}l_ja_j,\sum_{j=1}^{n}\sum_{k=1}^{n}l_jl_k\mathrm{Cov}(X_j,X_k)\right),$$

则有

$$E(e^{itY})=E(e^{it\boldsymbol{lX}^{\mathrm{T}}})=\exp\left(i\boldsymbol{al}^{\mathrm{T}}t-\frac{1}{2}\boldsymbol{lBl}^{\mathrm{T}}t^{2}\right),$$

其中 $\boldsymbol{B}=(\mathrm{Cov}(X_j,X_k))$ 为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的协方差矩阵.

令 $t=1$,得 $E(e^{i\boldsymbol{lX}^{\mathrm{T}}})=\exp\left(i\boldsymbol{al}^{\mathrm{T}}-\frac{1}{2}\boldsymbol{lBl}^{\mathrm{T}}\right)$.

由 $\boldsymbol{l}$ 的任意性,知 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的特征函数

$$\varphi(\boldsymbol{l})=\exp\left(i\boldsymbol{al}^{\mathrm{T}}-\frac{1}{2}\boldsymbol{lBl}^{\mathrm{T}}\right).$$

由特征函数的唯一性定理知,$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)\sim N(\boldsymbol{a},\boldsymbol{B})$.

必要性.若 $\boldsymbol{X}$ 服从 n 维正态分布 $N(\boldsymbol{a},\boldsymbol{B})$,则其特征函数 $E(e^{i\boldsymbol{t}^{\mathrm{T}}\boldsymbol{X}})=\exp\left(i\boldsymbol{at}^{\mathrm{T}}-\frac{1}{2}\boldsymbol{tBt}^{\mathrm{T}}\right)$,由多元特征函数性质(5),$Y$ 的特征函数

$$\varphi_Y(t)=\varphi(l_1t,l_2t,\cdots,l_nt)=\exp\left[it\sum_{j=1}^{n}l_ja_j-\frac{1}{2}t^2\sum_{j=1}^{n}\sum_{k=1}^{n}\mathrm{Cov}(X_j,X_k)l_jl_k\right],$$

这是正态分布 $N\left(\sum_{j=1}^{n}l_ja_j,\sum_{j=1}^{n}\sum_{k=1}^{n}l_jl_k\mathrm{Cov}(X_j,X_k)\right)$ 的特征函数.

由唯一性定理,$Y\sim N\left(\sum_{j=1}^{n}l_ja_j,\sum_{j=1}^{n}\sum_{k=1}^{n}l_jl_k\mathrm{Cov}(X_j,X_k)\right)$.

性质 4 若 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 服从 n 维正态分布 $N(\boldsymbol{a},\boldsymbol{B})$,又 m 维随机变量 $\boldsymbol{Y}=\boldsymbol{XC}$,其中 $\boldsymbol{C}$ 是 $n\times m$ 矩阵,则 $\boldsymbol{Y}$ 服从 m 维正态分布 $N(\boldsymbol{aC},\boldsymbol{C}^{\mathrm{T}}\boldsymbol{BC})$.

证明 $\boldsymbol{Y}$ 的特征函数为

$$\begin{aligned}\varphi(\boldsymbol{t})&=E(e^{i\boldsymbol{tY}^{\mathrm{T}}})=E(e^{i\boldsymbol{t}(\boldsymbol{XC})^{\mathrm{T}}})=E(e^{i(\boldsymbol{Ct}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}})\\&=\exp\left(i\boldsymbol{a}(\boldsymbol{Ct}^{\mathrm{T}})-\frac{1}{2}(\boldsymbol{Ct}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{BCt}^{\mathrm{T}}\right)\\&=\exp\left(i(\boldsymbol{aC})\boldsymbol{t}^{\mathrm{T}}-\frac{1}{2}\boldsymbol{t}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{BC})\boldsymbol{t}^{\mathrm{T}}\right).\end{aligned}$$

这是 m 维正态分布 $N(\boldsymbol{aC},\boldsymbol{C}^{\mathrm{T}}\boldsymbol{BC})$ 的特征函数,由唯一性定理得证.

性质 4 表明,正态随机变量经过线性变换后仍为正态随机变量.

例 1.9 设 $\boldsymbol{X}=(X_1,X_2)$,$\boldsymbol{X}\sim N(\boldsymbol{0},\boldsymbol{B})$,$\boldsymbol{B}=\begin{pmatrix}1&2\\2&1\end{pmatrix}$,$Y_1=2X_1+3X_2$,$Y_2=X_2$,求:(1) $\boldsymbol{Y}=(Y_1,Y_2)$ 的特征函数;(2) $Y_3=X_1+2X_2$ 的分布.

解 $\boldsymbol{A}=\begin{pmatrix}2&0\\3&1\end{pmatrix}$,$\boldsymbol{Y}=\boldsymbol{XA}$,$E(\boldsymbol{Y})=(0,0)\begin{pmatrix}2&0\\3&1\end{pmatrix}=(0,0)$,

$$\boldsymbol{B_Y}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{BA}=\begin{pmatrix}2&0\\3&1\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}1&2\\2&1\end{pmatrix}\begin{pmatrix}2&0\\3&1\end{pmatrix}=\begin{pmatrix}37&7\\7&1\end{pmatrix}.$$

故 $\boldsymbol{Y}$ 的特征函数为

$$\varphi_{\boldsymbol{Y}}(\boldsymbol{t})=\exp\left(-\frac{1}{2}\boldsymbol{tB_Yt}^{\mathrm{T}}\right)=\exp\left[-\frac{1}{2}(37t_1^2+t_2^2+14t_1t_2)\right].$$