随机过程

- 南京邮电大学
- 理学院
- 胡国雷

教材：随机过程，刘次华，华中理工大学出版社。
参考书：
1.应用随机过程，林元烈编著，清华大学出版社；
2.随机系统分析引论，盛昭瀚，东南大学出版社；
3.随机过程，伊曼纽尔、帕尔逊著，
 邓永录、杨振业译，高等教育出版社；
4.随机过程，Sheldon M[1].Ross著。

第一章 预备知识

简要回顾一下概率论中与本课程有关的基本概念：随机试验、样本空间、事件、概率、随机变量、概率分布、数字特征等。

§1.1 概率空间

一、基本概念

随机试验

- 试验结果事先不能准确预言，三个特征:
  - 可以在相同条件下重复进行；
  - 每次试验结果不止一个，可预先知道试验所有可能结果；
  - 每次试验前不能确定那个结果会出现。

样本空间

随机试验所有可能结果组成的集合，记为Ω

随机事件

样本空间Ω的子集A称为随机事件，用A、B、C表示

注：所谓某个事件在 试验中是否出现，当且仅当该事件所包含的某个样本点是否出现，因此一个事件实际上对应于的一个确定的子集。

事件的概率论运算 $\Longleftrightarrow$ Ω子集的集合论运算。

样本空间 Ω 也是一个事件，称 Ω 为必然事件，

空集 Φ 称为不可能事件。

注：由于事件是集合，故集合的运算（并、交、差、上极限、下极限、极限等）都适用于事件。

二、$\sigma-$代数（事件族）

在实际问题中，并不是对所有的事件：（样本空间Ω的所有子集）都感兴趣，而是关心某些事件（Ω的某些子集）及其发生的可能性大小（概率）。

为了数学上处理方便，我们常要求这些子集组成的类具有一些基本性质（即对事件需加一些约束）

定义1.1设样本空间 $\Omega=\{e\}$ 的某些子集构成的集合记为F，如果F满足下列性质：

(1). $\Omega \in F$；

(2).若 $A \in F$，则 $\overline{A}=\Omega-A\in F$

(3).若 $A_k\in F, k=1,2,\cdots$，则 $\bigcup_{k=1}^{\infty}A_k\in F$.

则称F为 σ —代数（*Bord*事件域），

(Ω,F) 称为可测空间

F中的元素称为事件。

对于某个事件A包含它的 $\sigma-$ 代数不是唯一的

例如，包含A的最大的 $\sigma-$ 代数是 Ω 的一切子集组成的集类

而包含*A*的最小的 $\sigma-$ 代数则是：$\{A,\overline{A},\Omega,\Phi\}$

注：F（Ω）表示由Ω的子集全体构成的集合类，显然满足上述定义的（1）~（3），但这个族常常显得太大以致对于某些样本空间而言不可以在这样的族上定义满足三条公理的概率函数 $P(\bullet)$

为了建立概率的数学理论通常只需把事件族取为具有定义（1）～（3）中并包含了我们感兴趣的所有集合的的最小子集族。

三、概率的公理化定义

为了完成随机现象的数学描述，还要规定随机事件族F上的概率函数 $P(\bullet)$ 即对F中的每个事件A要定义一个称作为的概率的数 $P(A)$，作为事件A的函数必须假定满足三条公理。

定义1. 2：设（Ω，F）是可测空间，$P(A)$ 是定义在F上的实值函数，如果 $P(A)$ 满足

(1) 对 $\forall A\in F$，有 $0\le P(A)\le 1$ 非负性；

（2） $P(\Omega)=1$ 规范性；

(3) 若 $A_1,A_2,\cdots,\in F$ 两两互不相容，即 $A_iA_j=\Phi\quad(i\ne j)$

有 $P(\bigcup_{k=1}^{\infty}A_k)=\sum_{k=1}^{\infty}P(A_k)$

则称P为（Ω，F）上的概率，（Ω，F，P）称为概率空间，P（A）为事件A的概率。

由此定义出发，可推出概率的其它一些性质：

(4) $P(\Phi)=0$;

(5) 若 $A,B\in F, A\subset B$，则 $P(B-A)=P(B)-P(A)$，且 $P(B)\ge P(A)$

即概率具有单调性；

(6) 设 $A_n\in F, n=1,2,\cdots$，则

$$\lim_{n\to\infty}P(A_n)=\begin{cases}P(\bigcup_{i=1}^{\infty}A_i) & 若\quad A_1\subset A_2\subset\cdots\\ P(\bigcap_{i=1}^{\infty}A_i) & 若\quad A_1\supset A_2\supset\cdots\end{cases}$$

连续性定理

当 $A_n\subset A_{n+1}, n\ge 1$ 新事件：$\lim_{n\to\infty}A_n=\bigcup_{i=1}^{\infty}A_i$

当 $A_n\supset A_{n+1}, n\ge 1$ $\lim_{n\to\infty}A_n=\bigcap_{i=1}^{\infty}A_i$

四、几个重要公式

加法公式

若 $A,B\in F$，则 $P(A\cup B)=P(A)+P(B)-P(AB)$

条件概率

- 在事件B已发生这一条件下，事件A发生的概率。

$$P(A\mid B)=\frac{P(A\cap B)}{P(B)}$$

全概率公式

- 若有N个互斥事件 B_n（n=1,2,…,N），它的并集等于整个样本空间，则

$$P(A)=\sum_{i=1}^{N}P(A\mid B_i)P(B_i)$$

贝叶斯公式

- 设事件 B_1，B_2，…，B_n 构成一个完备事件组，概率 $P(B_i)>0$, i=1,2,……,n,对于任何一个事件A，若P(A)>0, 有

$$P(B_i\mid A)=\frac{P(B_i)P(A\mid B_i)}{\sum_{i=1}^{N}P(B_i)P(A\mid B_i)}$$

独立事件 $P(A\cap B)=P(A)P(B)$

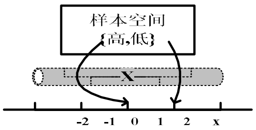
独立事件族：设（Ω，F，P）是概率空间，$Y\subset F$

如果对任意 $A_1,A_2,\cdots A_n\in Y, n=1,2,\cdots$ 有 $P(\bigcap_{i=1}^{n}A_i)=\prod_{i=1}^{n}P(A_i)$

则称Y为独立事件族。

§ 1.2　随机变量及其分布

一、一维随机变量及其分布函数



由于数学分析不能直接利用来研究集合函数，这样影响对随机现象的研究。解决这个问题的方法，主要是设法在集合函数与数学分析中所研究的点函数间建立某种联系，从而能用数学分析去研究随机现象。

X(e)就是一个函数，它把样本点映射到实数轴上，随机变量就是从原样本空间Ω到新样本空间的一种映射，我们通常把这样一种对应关系称之为在概率空间上的一个随机变量。下面我们给出随机变量的数学定义。

定义1.4：设（Ω，F，P）是概率空间，X=X(e)是定义在Ω上的实函数，如果对任意实数x，{e:X(e) ≤x} ∈F，则称X(e)是F上的随机变量。

事件 → 随机变量

离散型随机变量：

只取有限个数值或可列无穷多个值。

连续型随机变量：从原样本空间到新样本空间的映射是某一个范围，是一段（或几段）实线（也可能是整个坐标轴），随机变量可以取值于某一区间中的任一数。

分布函数（一个描述随机变量取值的概率分布情况的统一方法）

$$F(x)=P(e:X(e)\le x),\quad -\infty<x<\infty$$

分布函数$F(x)$具有下列性质：

(1) $F(x)$是非降函数：即当$x_1<x_2$时，有$F(x_1)\le F(x_2)$；

(2) $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0, F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$, $0\le F(x)\le 1$;

(3) $F(x)$右连续，即$F(x+0)=F(x)$。

离散型随机变量X的概率分布用分布律描述：

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

分布函数： $F(x)=\sum\limits_{x_k\le x}p_k$

连续型随机变量X的概率分布用密度函数$f(x)$描述：

分布函数为：$F(x)=\int_{-\infty}^{x}f(t)dt$

离散型随机变量的概率分布用分布列描述

0—1分布　$P(X=1)=p, P(X=0)=q$

二项分布　$P(X=k)=C_n^k p^k q^{n-k}, k=0,1,2\cdots n$

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

连续型随机变量的概率分布用概率密度描述

均匀分布　$f(x)=\begin{cases}\frac{1}{b-a}, & a<x<b\\ 0, & \text{其它}\end{cases}$

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

指数分布　$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$

随机变量函数的分布

在给定某任意的随机变量X，以及它的概率分布函数$F_X(x)$，希望进一步求出给定的随机变量的某些可测函数（如Y=g(X)）的概率分布函数。

X → 非线性放大器 → Y

Y的概率分布函数公式为

$$F_Y(y)=P(e:g(X)\le y,e\in\Omega_X)$$

如果上式右端概率的导数对于y处处存在，那么这个导数就给出了随机变量Y的概率密度

$$f_Y(y)=\frac{d}{dy}P(\{e:g(X)\le y,e\in\Omega_X\})$$

二、n维随机变量及其分布函数

定义1.5　设（Ω，F，P）是概率空间，**X=X**(e)＝（X_1(e),…,X_n(e)）是定义在Ω上的n维空间R^n中取值的向量函数。如果对于任意x=(x_1,…,x_n) ∈R^n，{e:X_1(e) ≤x_1,…,X_n(e) ≤x_n} ∈F，则称**X=X**(e)为n维随机变量。称

$$\begin{aligned}F(x)&=F(x_1,x_2\cdots x_n)\\&=P\{e:X_1(e)\le x_1,X_2(e)\le x_2,\cdots,X_n(e)\le x_n\}\\&=P\left\{e:\bigcap_{i=n}^{n}\{X_i\le x_i\}\right\}\end{aligned}$$

为X=(X_1,X_2,…,X_n)的联合分布函数

n维联合分布函数$F(x_1,x_2,\cdots x_n)$具有下列性质：

(1) 对于每个变量 $x_i(x_1,x_2,\cdots,x_n)$，$F(x_1,x_2,\cdots,x_n)$是非降函数；

(2) 对于每个变量 $x_i(x_1,x_2,\cdots,x_n)$，$F(x_1,x_2,\cdots,x_n)$是右连续的；

(3) 对于R^n中的任意区域，$(a_1,b_1;a_2,b_2;\cdots;a_n,b_n)$，其中 $a_i\le b_i,i=1,\cdots,n$

$$F(b_1,b_2,\cdots,b_n)-\sum_{i=1}^{n}F(b_1,\cdots,b_{i-1},a_i,b_{i+1},\cdots b_n)$$
$$+\sum_{\substack{i,j=1\\i<j}}F(b_1,\cdots,b_{i-1},a_i,b_{i+1},\cdots,b_{j-1},a_j,b_{j+1},\cdots,b_n),$$
$$+(-1)^nF(a_1,a_2,\cdots,a_n)\ge 0;$$ X落在R^n中任一超长方体中的概率

(4) $\lim\limits_{x_i\to-\infty}F(x_1,x_2,\cdots,x_i,\cdots,x_n)=0,i=1,2,\cdots,n$

$$\lim_{x_1,x_2,\cdots,x_n\to+\infty}F(x_1,x_2,\cdots,x_n)=1\qquad 0\le F(x_1,x_2,\cdots,x_n)\le 1$$

三、边缘分布

若二维联合分布函数中有一个变元趋于无穷，则其极限函数便是一维分布函数，对于这种特殊性质，我们称其为边缘分布。

对于任意两个随机变量X,Y，其联合分布函数为：

$$F(x,y)$$

则：

$$F_X(x)=P(X\le x)=P(X\le x,Y<\infty)=F(x,\infty)$$
$$F_Y(y)=P(Y\le y)=P(X<+\infty,Y\le y)=F(\infty,y)$$

分别称$F_X(x)$和$F_Y(y)$为$F(x,y)$ 关于X和关于Y的边缘分布函数。

离散型随机变量（X，Y）边缘分布律计算如下

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

连续型随机变量（X，Y）边缘概率密度计算如下

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

联合密度 → 边缘密度

联合密度 ← ? 边缘密度

相互独立的随机变量

设X，Y是两个随机变量，若对任意实数x，y有

$$P(X\le x,Y\le y)=P((X\le x)\cap(Y\le y))=P(X\le x)P(Y\le y)$$

则称X，Y为相互独立的随机变量。

若X，Y为相互独立随机变量，则有

$$F(x,y)=F_X(x)F_Y(y)$$
$$f(x,y)=f_X(x)f_Y(y)$$

四、条件分布

条件概率　　　　条件分布函数

$P(A|B)=\dfrac{P(A\cap B)}{P(B)} \Rightarrow F_{X|B}(x|B)=P(X\le x|B)=\dfrac{P(X\le x\cap B)}{P(B)}$

$$F_{X|Y}(x|Y=y)=\frac{\int_{-\infty}^{x} f(u,y)du}{f_Y(y)}$$

两边对x微分 $\Rightarrow f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}$

$\Rightarrow F_{X|Y}(x|Y=y)=\int_{-\infty}^{x} f_{X|Y}(u|y)du$

§ 1.3　随机变量的数字特征

- 随机变量的数学期望
- 随机变量函数的期望
- 方差
- 协方差
- 相关系数
- 独立与不相关

一、斯蒂尔吉斯积分（补充）

1.有限区间上的斯蒂尔吉斯积分

定义　设$f(x),g(x)$是定义在区间$[a,b]$上的两个有界函数，把区间$[a,b]$分成n个子区间，

分点为$a=x_0<x_1\cdots<x_n=b$

在每一个子区间$[x_{k-1},x_k]$上任意取一个点ξ_k作和式

$S=\sum_{k=1}^{n} f(\xi_k)[g(x_k)-g(x_{k-1})]$

令$\Delta=\max(x_k-x_{k-1}),1\le k\le n$

如果极限$\lim_{\Delta\to 0} S=\lim_{\Delta\to 0}\sum_{k=1}^{n} f(\xi_k)[g(x_k)-g(x_{k-1})]$存在，

且与子区间的分法和ξ_k的取法无关，

则称此极限为函数$f(x)$对函数$g(x)$在区间$[a,b]$上的斯蒂尔吉斯积分($Stieltjes$).

记为$\int_a^b f(x)dg(x)$

简称S积分，S积分是高等数学中，黎曼积分的推广，如果取$g(x)=x$，则S积分就变成黎曼积分。

2.无限区间上的S积分

定义　设$f(x),g(x)$是定义在无限区间$(-\infty,+\infty)$上的两个函数，若在任意有限区间$[a,b]$上，$f(x)$对$g(x)$是S可积的，

且极限$\lim_{\substack{a\to-\infty\\ b\to+\infty}}\int_a^b f(x)dg(x)$存在

则称此极限为$f(x)$对$g(x)$在无限区间$(-\infty,+\infty)$上的斯蒂吉斯积分，

记为$\int_{-\infty}^{+\infty} f(x)dg(x)$

3.在积分中，当$g(x)$取一些特殊形式时，积分可化为通常积分或级数

若$g(x)$是在$(-\infty,+\infty)$上的阶梯函数，它的跳跃点为$x_1,x_2,\cdots$(有限多个或无限可列多个),

则:$\int_{-\infty}^{+\infty} f(x)dg(x)=\sum_k f(x_k)[g(x_k+0)-g(x_k-0)]$

若$g(x)$是在$(-\infty,+\infty)$上的可微函数，它的导函数为$g'(x)$

则$\int_{-\infty}^{+\infty} f(x)dg(x)=\int_{-\infty}^{+\infty} f(x)g'(x)dx$

前者把S积分化为和式，后者把S积分化成黎曼积分。

二、数学期望

定义 设随机变量X的分布函数为$F(x)$，若$\int_{-\infty}^{+\infty}|x|dF(x)<\infty$

则称$EX=\int_{-\infty}^{+\infty}xdF(x)$为$X$的数学期望或均值。

左边的积分称为斯蒂尔吉斯积分

若X为离散型随机变量，分布律为

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

则$EX=\sum_{k=1}^{\infty}x_k p_k$

若X为连续型随机变量，概率密度为$f(x)$

则$EX=\int_{-\infty}^{+\infty}xf(x)dx$

随机变量函数的期望

已知随机变量X的数学期望，求随机变量函数Y=g(X)的数学期望，

$$E(Y)=E(g(X))=\int_{-\infty}^{\infty}yf_Y(y)dy=\int_{-\infty}^{\infty}g(x)f_X(x)dx$$

对于多维随机变量

若n维随机变量$(X_1,X_2\cdots,X_n)$的联合分布函数为$F(x_1,x_2\cdots,x_n)$，$g(X_1,X_2\cdots,X_n)$是n维连续函数

$$E[g(X_1,X_2\cdots,X_n)]$$
$$=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}g(x_1,x_2\cdots,x_n)dF(x_1,x_2\cdots,x_n)$$

设$X_1,X_2,\ldots,X_n$为随机变量，求随机变量函数$Y=a_1X_1+a_2X_2+\ldots+a_nX_n$的数学期望。

$$\begin{aligned}E(Y)&=E(a_1X_1+a_2X_2+\cdots+a_nX_n)\\&=E(a_1X_1)+E(a_2X_2)+\cdots+E(a_nX_n)\\&=a_1E(X_1)+a_2E(X_2)+\cdots+a_nE(X_n)\end{aligned}$$

已知随机变量X_1和X_2，求随机变量函数$Y=aX_1+bX_2$的数学期望

$$\begin{aligned}E(Y)&=\int_{-\infty}^{\infty}\int(ax_1+bx_2)f(x_1,x_2)dx_1dx_2\\&=a\int_{-\infty}^{\infty}\int x_1f(x_1,x_2)dx_1dx_2+b\int_{-\infty}^{\infty}\int x_2f(x_1,x_2)dx_1dx_2\\&=aE(X_1)+bE(X_2)\end{aligned}$$

加权和的期望等于加权期望的和

求数学期望是线性运算

数学期望的线性运算不受独立条件限制

已知随机变量X_1和X_2，求随机变量函数$Y=g_1(X_1)g_2(X_2)$的数学期望

$$E[Y]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g_1(x_1)g_2(x_2)f(x_1,x_2)dx_1dx_2$$

假设两个随机变量X_1和X_2相互独立，则有

$$f(x_1,x_2)=f_{X_1}(x_1)f_{X_2}(x_2)$$

因此，有

$$\begin{aligned}E[Y]&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g_1(x_1)g_2(x_2)f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2\\&=\int_{-\infty}^{\infty}g_1(x_1)f_{X_1}(x_1)dx_1\int_{-\infty}^{\infty}g_2(x_2)f_{X_2}(x_2)dx_2\\&=E[g_1(X_1)]E[g_2(X_2)]\end{aligned}$$

特别若X,Y相互独立，则$E(XY)=EX\cdot EY$

三、方差（随机变量取值的离散程度）

定义 设X是随机变量，若$EX^2<\infty$，

则称$DX=E(X-EX)^2$为X的方差

计算公式：$DX=EX^2-(EX)^2$

标准差 $\sigma(X)=\sqrt{D(X)}$

若X,Y相互独立，则

$D(aX+bY)=a^2DX+b^2DY$ a,b为常数

四、协方差与相关系数

定义1.9　设X,Y是随机变量,$EX^2<\infty, EY^2<\infty$

则称$B_{XY}=E[(X-EX)(Y-EY)]$为X,Y的协方差

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

引入一个描述两个随机变量相关程度的系数

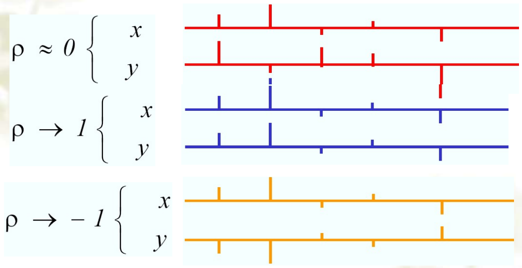
$$\rho_{XY}\overset{def}{=}\frac{Cov(X,Y)}{\sqrt{DX}\sqrt{DY}}$$

ρ_{XY}称为归一化的协方差系数或相关系数。

$$-1\le\rho_{XY}\le 1$$

相关系数ρ_{XY}表示X，Y之间的线性相关程度的大小

若$\rho_{XY}=0$，则称随机变量X和Y不相关。



若两个随机变量X和Y的联合矩满足　$E[X^iY^j]=E[X^i]E[Y^j]$

则称随机变量X和Y统计独立

五、K阶原点矩、k阶中心矩

随机变量X，若$E[|X|^k]<\infty$，称$E[X^k]$为k阶原点矩。

离散随机变量　连续随机变量

$$E[X^k]=\sum_{i=1}^{\infty}x_i^k p_i=\int_{-\infty}^{\infty}x^k f_X(x)dx$$

又若E[X]存在，且$E[|X-E[X]|^k]<\infty$，称

$$E[(X-E[X])^k]$$

为X的k阶中心矩。

离散随机变量　连续随机变量

$$E[(X-E[X])^k]=\sum_{i=1}^{\infty}(x_i-E[X])^k p_i=\int_{-\infty}^{\infty}(x-E[X])^k f_X(x)dx$$

一阶原点矩就是随机变量的数学期望，

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

数学期望大致的描述了概率分布的中心。

二阶中心矩就是随机变量的方差，

$$DX\overset{def}{=}E(X-EX)^2$$

方差反映随机变量取值的离散程度。

0—1分布
泊松分布
正态分布

常用分布的数学期望和方差（见表1—1）

中心化的两个随机变量X-E[X]，Y-E[Y]的互相关矩称为随机变量X和Y的协方差，

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

协方差是描述随机现象中，随机变量X和Y概率相关的程度。

相互独立 $\Rightarrow$ 不相关

$$Cov(X,Y)=E[(X-EX)(Y-EY)]$$
$$=E[XY]-E[X]E[Y]=0$$

相互独立 $\overset{?}{\Leftarrow}$ 不相关

设Z是一个随机变量，具有均匀概率密度

$$f_Z(z)=\begin{cases}\frac{1}{2\pi}, & 0\le z\le 2\pi\\ 0, & 其它\end{cases}$$

令X=sinZ，Y=cosZ，求随机变量X和Y是否相关，是否独立？

§1.4 特征函数、母函数

数字特征只反映了概率分布的某些侧面，一般并不能通过它们来确定分布函数，这里将要引进的特征函数，既能完全决定分布函数而又具有良好的分析性质。

一、复随机变量

如果X与Y都是概率空间(Ω, F, P)上的实值随机变量.

则称$Z = X + iY$为复随机变量

数学期望 $EZ = EX + iEY$

对复随机变量也可以平行于实随机变量建立起一系列结果。

例如 若Z_1，$Z_2, \cdots, Z_n$是相互独立的

则 $E(Z_1 Z_2 \cdots Z_n) = EZ_1 EZ_2 \cdots EZ_n$

二、特征函数

定义 设随机变量X的分布函数为$F(x)$

称$g(t) = E(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} dF(x) \quad -\infty < x < +\infty$

为X的特征函数

特征函数是一个实变量为t的复值函数，

由于$|e^{itx}| = 1$，故随机变量的特征函数必然存在。

对离散型随机变量，若其分布律为

$P\{X = x_k\} = p_k, k = 1,2\cdots$， 则 $g(t) = \sum_{k=1}^{\infty} e^{itx_k} \cdot p_k$

对于连续型随机变量，若其分布密度函数为$f(x)$

则 $g(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx$

这时，特征函数是密度函数$f(x)$的付里叶变换。

根据F积分理论，在$\int_{-\infty}^{+\infty} |g(t)| dt < \infty$的条件下

有反演公式 $f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} g(t) dt$

三、特征函数的性质

(1) $g(0) = 1, \quad |g(t)| \le 1, \quad g(-t) = \overline{g(t)}$

证：$g(0) = \int_{-\infty}^{+\infty} 1 \cdot dF(x) = 1$

$|g(t)| \le \int_{-\infty}^{+\infty} |e^{itx}| dF(x) = 1 = g(0)$

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

(2) 特征函数$g(t)$在$(-\infty, +\infty)$上一致连续

所谓$g(t)$在$(-\infty, +\infty)$上一致连续：$\forall \varepsilon > 0$，总$\exists$与t无关的$\delta > 0$，当$|h| < \delta$时

有$|g(t+h) - g(t)| < \varepsilon \quad (-\infty < t < +\infty)$

证：$|g(t+h) - g(t)| = \left| \int_{-\infty}^{+\infty} (e^{i(t+h)x} - e^{itx}) dF(x) \right|$

$\le \int_{-\infty}^{+\infty} |e^{ihx} - 1| dF(x) \le 2\int_{|x| \ge A} dF(x)$

$+ \int_{-A}^{A} |e^{ihx} - 1| dF(x)$

$= 2\int_{|x| \ge A} dF(x) + 2\int_{-A}^{A} \left| \sin \frac{hx}{2} \right| dF(x)$

上式右边已与t无关，可选足够大的A

使$\int_{|x| \ge A} dF(x)$任意小，然后选充分小的$|h|$可使

第二个积分也任意小，从而证明了结论。

(3) 若随机变量X的n阶矩EX^n存在，则X的特征函数$g(t)$可微分n次，且当$k \le n$时

$g^{(k)}(0) = i^k EX^k$

证：$\left| \frac{d^k}{dt^k} (e^{itx}) \right| = |i^k x^k e^{itx}| \le |x|^k$

由于X的k阶矩存在，故$\int_{-\infty}^{+\infty} |x|^k dF(x) < \infty$

因而可作下列积分号下的微分

$$g^{(k)}(t)=\int_{-\infty}^{+\infty}\frac{d^k}{dt^k}\left(e^{itx}\right)dF(x)=i^k\int_{-\infty}^{+\infty}x^k e^{itx}dF(x)$$

取$t=0$,即得 $g^{(k)}(0)=i^k EX^k$

此性质使我们可以方便地求得随机变量的各阶矩

(4) $g(t)$是非负定函数，即对任意正整数n
及任意实数$t_1,t_2,\cdots,t_n$和复数$Z_1,Z_2,\cdots,Z_n$有：

$$\sum_{k,l=1}^{n}g(t_k-t_l)Z_k\overline{Z}_l\geq 0$$

证：$$\sum_{k,l=1}^{n}g(t_k-t_l)Z_k\overline{Z}_l=\sum_{k=1}^{n}\sum_{l=1}^{n}E\left[e^{i(t_k-t_l)X}\right]Z_k\overline{Z}_l$$

$$=E\left[\sum_{k=1}^{n}\sum_{l=1}^{n}e^{it_kX}Z_k\cdot\overline{e^{it_lX}Z_l}\right]=E\left|\sum_{k=1}^{n}e^{it_kX}\cdot Z_k\right|^2\geq 0$$

(5) 若$X_1,X_2,\cdots,X_n$是相互独立的随机变量，
则$X=X_1+X_2+\cdots+X_n$的特征函数为：

$$g_X(t)=g_{X_1}(t)\cdot g_{X_2}(t)\cdots g_{X_n}(t)$$

其中$g_{X_i}(t)$是随机变量X_i的特征函数，$i=1,2,\cdots,n$

证：因为$X_1,X_2,\cdots$，X_n相互独立，
所以复随机变量$e^{itX_1},\cdots,e^{itX_n}$也相互独立
所以

$$g(t)=E\left(e^{itX}\right)=E\left[e^{it(X_1+X_2+\cdots+X_n)}\right]=E\left[e^{itX_1}\cdots e^{itX_n}\right]$$
$$=E\left(e^{itX_1}\right)\cdots E\left(e^{itX_n}\right)=g_{X_1}(t)\cdots g_{X_n}(t)$$

(6) 设$Y=aX+b$, a,b为常数
则$g_Y(t)=e^{ibt}g_X(at)$

证明：$g_Y(t)=E\left[e^{itY}\right]=E\left[e^{it(aX+b)}\right]$
$=e^{itb}E\left(e^{itaX}\right)=e^{itb}g_X(at)$

（7）特征函数与分布函数是相互唯一确定的

*逆转公式：设分布函数$F(x)$的特征函数
为$g(t)$,又x_1,x_2是$F(x)$的连续点，则

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

唯一性定理：
分布函数由其特征函数唯一决定

证：应用逆转公式，在$F(x)$的每一连续点
上，当y沿$F(x)$的连续点趋于$-\infty$时，有

$$F(x)=\lim_{y\to-\infty}[F(x)-F(y)]$$
$$=\frac{1}{2\pi}\lim_{y\to-\infty}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-ity}-e^{-itx}}{it}g(t)dt$$

而分布函数由其连续点上的值唯一决定
不连续点利用右连续性

特别当$g(t)$是绝可积函数时，有下列更强的结果

定理：若$\int_{-\infty}^{+\infty}|g(t)|dt<\infty$，则相应的分布函数$F(x)$
的导数存在连续，而且

$$F'(x)=f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}g(t)dt$$

$$g(t)=\int_{-\infty}^{+\infty}f(x)e^{itx}dx$$

即在特征函数绝对可积的条件下，概率密度
与特征函数构成一对付氏变换。

证明：由逆转公式，若$x-\delta$及$x+\delta$
是$F(x)$的连续点

则$F(x+\delta)-F(x-\delta)=\lim_{T\to\infty}\frac{1}{\pi}\int_{-T}^{T}\frac{\sin t\delta}{t}e^{-itx}g(t)dt$

因此

$$\frac{F(x+\delta)-F(x-\delta)}{2\delta}=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{\sin t\delta}{t\delta}e^{-itx}g(t)dt$$

由于 $\left|\frac{\sin t\delta}{t\delta}e^{-itx}g(t)\right|\leq|g(t)|$

因此用控制收敛定理知（极限号与积分号
交换的勒贝格控制收敛定理）

$$F'(x)=f(x)=\lim_{\delta\to 0}\frac{F(x+\delta)-F(x-\delta)}{2\delta}$$
$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itx}g(t)dt$$

因此，在$g(t)$是绝对可积的条件下，分布密度$f(x)$与特征函数$g(t)$可以通过付里叶变换来联系。

四、多元特征函数

1.定义 设$X=(X_1 X_2,\cdots,X_n)'$ 是n维随机向量，$t=(t_1,t_2,\cdots,t_n)'\in R^n$，

则称 $g_X(t)=g(t_1,t_2,\cdots,t_n)=Ee^{it'X}$
$$=E\left[\exp\left\{i\sum_{k=1}^{n}t_k X_k\right\}\right]$$
为X的特征函数

n维特征函数具有类似于一维随机变量的特征函数的性质。

2.性质

(1). $g(t_1,t_2,\cdots,t_n)$在R^n中一致连续，

且$|g(t_1,t_2,\cdots,t_n)|\le g(0,0,\cdots,0)=1$,

$g(-t_1,-t_2,\cdots,-t_n)=\overline{g(t_1,t_2,\cdots,t_n)}$

(2) 若$(X_1,X_2,\cdots, X_n)$相互独立，

$g_X(t_1,t_2,\cdots, t_n)=E\left[e^{i(t_1X_1+\cdots+t_nX_n)}\right]=g_{X_1}(t_1)\cdots g_{X_n}(t_n)$

反之也成立，可用此判别$X_1,\cdots, X_n$是否独立。

(3) 设$X_{j_1},X_{j_2},\cdots,X_{j_k}$是$X_1,X_2,\cdots,X_n$中任意$k$个

则$g(t_{j_1},\cdots,t_{j_k})=g_X(t_1,\cdots,t_n)\big|_{t_j=0},\ j\ne j_e\quad e=1,2,\cdots,k$

这是任意k个分量的边缘分布的特征函数

(4) $E\left[X_1^{k_1}X_2^{k_2}\cdots X_n^{k_n}\right]$
$$=\frac{\partial^{k_1+k_2+\cdots+k_n}}{\partial t_1^{k_1}\partial t_2^{k_2}\cdots\partial t_n^{k_n}}g(t_1,\cdots,t_n)i^{-\sum_{l=1}^{n}k_l}\Big|_{t_1=t_2=\cdots=t_n=0}$$

例1 设$X\sim B(n,p)$，求X的特征函数$g(t)$及EX,EX^2,DX

解：$\because$ X的分布律为：
$$P\{X=k\}=C_n^k p^k q^{n-k},$$
$$q=1-p,k=0,1,2,\cdots,n$$
$$g(t)=\sum_{k=0}^{n}e^{itk}C_n^k p^k q^{n-k}=\sum_{k=0}^{n}C_n^k\left(pe^{it}\right)^k q^{n-k}$$
$$=\left(pe^{it}+q\right)^n$$

由性质(3)知
$$EX=-ig'(0)=-i\frac{d}{dt}\left(pe^{it}+q\right)^n/_{t=0}=np$$
$$EX^2=(-i)^2g''(0)=(-i)^2\frac{d^2}{dt^2}\left(pe^{it}+q\right)^n/_{t=0}$$
$$=npq+n^2q^2$$

故 $DX=EX^2-(EX)^2=npq$

例2 设$X \sim N(0,1)$，求X的特征函数$g(t)$

解：$g(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{itx-\frac{x^2}{2}}dx$

由于$\left|ixe^{itx-\frac{x^2}{2}}\right|=|x|e^{-\frac{x^2}{2}}$ 且$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}|x|e^{-\frac{x^2}{2}}dx<\infty$

$$\therefore g'(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}ixe^{itx-\frac{x^2}{2}}dx=\frac{i}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{itx}\left(-de^{-\frac{x^2}{2}}\right)$$

$$=-\frac{i}{\sqrt{2\pi}}e^{itx-\frac{x^2}{2}}\Big|_{-\infty}^{+\infty}+\frac{i^2t}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{itx-\frac{x^2}{2}}dx=-tg(t)$$

即 $g'(t)+tg(t)=0$,

$$\frac{dg(t)}{g(t)}=-tdt \qquad \ln g(t)=-\frac{1}{2}t^2+c$$

$$g(t)=e^{-\frac{1}{2}t^2+c}$$

由$g(0)=1 \Rightarrow c=0$

$\therefore X$的特征函数为：$g(t)=e^{-\frac{t^2}{2}}$

例3 设随机变量$Y \sim N(a,\sigma^2)$，求Y的特征函数$g_Y(t)$

解 设$X \sim N(0,1)$,由例2知$g_X(t)=e^{-\frac{t^2}{2}}$

令$Y=\sigma X+a$,则$Y \sim N(a,\sigma^2)$,

由性质知Y的特征函数

$$g_Y(t)=e^{iat}g_X(\sigma t)=e^{iat}e^{-\frac{\sigma^2t^2}{2}}=e^{iat-\frac{\sigma^2t^2}{2}}$$

例4 若$X \sim B(n,p)$，$Y \sim B(m,p)$，且X与Y相互独立,则$Z=X+Y \sim B(m+n,p)$

证 $\because g_X(t)=(pe^{it}+q)^n$，$g_Y(t)=(pe^{it}+q)^m$

由性质(5)

$$g_z(t)=(pe^{it}+q)^{m+n}$$

由唯一性定理知 $Z \sim B(m+n,p)$

例5.设$X \sim U\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$,$Y=\cos X$,

利用特征函数求Y的概率密度。

解：因为X的概率密度为：

$$f(x)=\begin{cases}\frac{1}{\pi}, & x\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)\\ 0, & 其他\end{cases}$$

$$g_Y(t)=E[e^{itY}]=E[e^{it\cos X}]$$

$$=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}e^{it\cos x}\cdot\frac{1}{\pi}dx=\frac{2}{\pi}\int_0^{\frac{\pi}{2}}e^{it\cos x}dx$$

令 $\cos x=u$, $du=-\sin xdx=-\sqrt{1-u^2}dx$

$$\therefore \quad g_Y(t)=\frac{2}{\pi}\int_0^1 e^{itu}\cdot\frac{1}{\sqrt{1-u^2}}du$$

利用特征函数与分布一一对应的唯一性得

Y的概率密度为：$f_Y(y)=\frac{2}{\pi}\cdot\frac{1}{\sqrt{1-y^2}}$, $0<y<1$

注：求随机变量的特征函数的方法

（1）一般定义求解；

（2）对一些特殊分布可化为微分方程求解；

（3）用Fourier变换去求解。

要求：

（1）会求一些常用的随机变量的特征函数；

（2）记住一些重要分布的特征函数，如正态分布；

（3）利用特征函数求相应随机变量的各阶矩；

（4）利用特征函数求多个独立随机变量和的分布。

五、母函数

在离散型随机变量中，那些只取非负整数值$0,1,2,\cdots$的占有重要的地位，我们称取值非负整数的随机变量为整数随机变量。

对于整值随机变量，有一种处理方法很便于应用，这就是母函数法。

定义　设X是非负整数值随机变量

分布律为　$P\{X=k\}=p_k$，$k=0,1,2,\cdots$

则称　$$P(S)=E[S^X]=\sum_{k=0}^{\infty}s^k p_k$$

为X的母函数。

由于$\sum_{k=0}^{\infty}p_k=1$，由幂级数的收敛性知$P(s)$至少在$|s|\le 1$一致收敛且绝对收敛，因此母函数对任何整值随机变量都存在。

例、求二项分布、泊松分布、几何分布的母函数

二项分布　$P\{X=k\}=C_n^k p^k q^{n-k}$

$$P(s)=\sum_{k=0}^{n}C_n^k p^k q^{n-k}s^k=\sum_{k=0}^{n}C_n^k(ps)^k q^{n-k}=(ps+q)^n$$

泊松分布：$P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda},k=0,1,2\cdots$

$$P(s)=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}s^k=e^{-\lambda}\cdot e^{\lambda s}=e^{\lambda(s-1)}$$

几何分布　$P\{X=k\}=q^{k-1}p,k=1,2,\cdots$

$$P(s)=\sum_{k=1}^{\infty}q^{k-1}ps^k=ps\sum_{k=1}^{\infty}(qs)^{k-1}=\frac{ps}{1-qs}$$

六、母函数的性质

（1）唯一性，非负整数值随机变量的分布列由其母函数唯一确定

证：$P(s)=\sum_{k=0}^{\infty}p_k s^k=\sum_{k=0}^{n}p_k s^k+\sum_{k=n+1}^{\infty}p_k s^k$，$n=0,1,2,\cdots$

两边对s，求n阶导数得：

$$P^{(n)}(s)=n!p_n+\sum_{k=n+1}^{\infty}k(k-1)\cdots(k-n+1)p_k s^{k-n}$$

令$s=0$，则$P^{(n)}(0)=n!p_n\Rightarrow p_n=\dfrac{P^{(n)}(0)}{n!}$，

$n=0,1,\cdots$

(2)　设$P(s)$是X的母函数，若EX存在，则$EX=P'(1)$

若DX存在，则$DX=P''(1)+P'(1)-[P'(1)]^2$

证：由$P(s)=\sum_{k=0}^{\infty}p_k s^k\Rightarrow P'(s)=\sum_{k=1}^{\infty}kp_k s^{k-1}$，

$$P''(s)=\sum_{k=2}^{\infty}k(k-1)p_k s^{k-2}$$

令$s\uparrow 1$，得$EX=\sum_{k=1}^{\infty}kp_k=P'(1)$

$$E[X(X-1)]=\sum_{k=2}^{\infty}k(k-1)p_k=P''(1)$$

故　$DX=EX^2-(EX)^2=P''(1)+P'(1)-[P'(1)]^2$

例 二项分布，母函数为$P(s)=(q+ps)^n$

$$EX=P'(1)=n(q+ps)^{n-1}p\big|_{s=1}=np$$

$$P''(1)=n(n-1)(q+ps)^{n-2}p^2\big|_{s=1}=n(n-1)p^2$$

$$DX=n^2p^2-np^2+np-n^2p^2=npq$$

例 泊松分布：母函数为$P(s)=e^{\lambda(s-1)}$

$$EX=P'(1)=e^{\lambda(s-1)}\cdot\lambda\big|_{s=1}=\lambda$$

$$P''(1)=e^{\lambda(s-1)}\cdot\lambda^2\big|_{s=1}=\lambda^2$$

$$DX=\lambda^2+\lambda-\lambda^2=\lambda$$

3、独立随机变量之和的母函数等于母函数之积

设X,Y为相互独立的整值随机变量，概率分布律分别为$\{a_k\}$及$\{b_k\}$，相应的母函数为$A(s)$及$B(s)$。

下面首先计算$Z=X+Y$的概率分布，

显然Z也是整值随机变量

记$c_r=P\{Z=r\}=P\{X+Y=r\}=\sum_{i=0}^{r}P\{X=i,Y=r-i\}$

则Z的概率分布为：$c_r=a_0b_r+a_1b_{r-1}+\cdots+a_rb_0$

记$C(s)=\sum_{r=0}^{\infty}c_rs^r$

$$\because A(s)\cdot B(s)=\sum_{k=0}^{\infty}a_ks^k\cdot\sum_{l=0}^{\infty}b_ls^l=\sum_{k,l}a_kb_ls^{k+l}$$

$$=\sum_{r=0}^{\infty}\left(\sum_{k=0}^{r}a_kb_{r-k}\right)s^r=\sum_{r=0}^{\infty}c_rs^r$$

$$\therefore C(s)=A(s)\cdot B(s)$$

或$C(s)=E\left(s^{X+Y}\right)=E\left(s^X\cdot s^Y\right)=Es^X\cdot Es^Y$

$$=A(s)\cdot B(s)$$

在研究独立随机变量和的问题时，母函数很适用，此结论而推广到n个独立整值随机变量之和的场合

(4) 随机个随机变量之和的母函数

若$X_1,X_2,\cdots,X_n,\cdots$是相互独立具有相同分布的整值随机变量,$N$是与$X_1,X_2\cdots$独立的整值随机变量，则$Y=\sum_{j=1}^{N}X_j$的母函数：

$$H(s)=G[P(s)]$$

其中$G(s)$、$P(s)$分别是N、X_1的母函数

证 $H(s)=\sum_{k=0}^{\infty}P\{Y=k\}\cdot s^k$

$$=\sum_{k=0}^{\infty}P\{Y=k,\bigcup_{l=0}^{\infty}(N=l)\}s^k$$

$$=\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}P\{N=l\}\cdot P\{Y=k/N=l\}s^k$$

$$=\sum_{l=0}^{\infty}P\{N=l\}\cdot\sum_{k=0}^{\infty}P\{Y=k/N=l\}s^k$$

$$=\sum_{l=0}^{\infty}P\{N=l\}\cdot\sum_{k=0}^{\infty}P\left\{\sum_{j=1}^{l}X_j=k\right\}\cdot s^k$$

$$=\sum_{l=0}^{\infty}P\{N=l\}\cdot[P(s)]^l=G[P(s)]$$

$$EY=H'(s)\big|_{s\uparrow 1}=G'[P(s)]\cdot P'(s)_{s\uparrow 1}$$

$$=G'[1]\cdot P'[1]=EN\cdot EX_1$$

例 设商店在一天的顾客数N服从参数$\lambda=1000$人的泊松分布,又设每位顾客所化的钱$X_i\sim N(100,50^2)$,求商店的日销售额Z的平均值

解：$EN=1000$，$EX_1=100$

$$\therefore EZ=EN\cdot EX_1=1000\times 100=100000(\text{元})$$

§1.5 n维正态分布

一、密度函数与特征函数

定义 若n维随机变量$X=(X_1,X_2,\cdots X_n)'$的联合概率密度为：

$$f(x)=f_n(x_1,x_2,\cdots,x_n)$$
$$=\frac{1}{(2\pi)^{n/2}|B|^{1/2}}\exp\left\{-\frac{1}{2}(x-a)'B^{-1}(x-a)\right\}$$

79

式中$a=(a_1,a_2,\cdots,a_n)'$是常向量,$B=(b_{ij})_{n\times n}$是对称阵

则称X为n维正态随机变量或服从n维正态分布，记为$X\sim N(a,B)$

a,B分别是随机向量X的数学期望及协方差矩阵。

其中$a_j=EX_j\quad 1\le j\le n\quad b_{ij}=E[(X_i-a_i)(X_j-a_j)]$, $1\le i,j\le n$

其特征函数为：

$$g(t)=g(t_1,t_2,\cdots,t_n)=e^{ia't-\frac{1}{2}t'Bt}=\exp\left\{ia't-\frac{1}{2}t'Bt\right\}$$

80

二、几个常用结论

(1) $X=(X_1,X_2\cdots,X_n)'$服从n元正态分布$N(a,B)$的充要条件是它的任一个线性组合$Z=\sum_{j=1}^{n}l_jX_j$服从一元正态分布 $N\left(\sum_{j=1}^{n}l_ja_j,\sum_{j,k=1}^{n}l_jl_kb_{jk}\right)$

证:$\Rightarrow$若$X\sim N(a,B)$，则$g(t)=E[e^{it'X}]=\exp\{ia't-\frac{1}{2}t'Bt\}$

$$Z=\sum_{j=1}^{n}l_jX_j=l'X,l=(l_1,l_2,\cdots,l_n)',X=(X_1,X_2,\cdots,X_n)'$$

取$t=ul$，u为实数，则

81

$$g(u)=E[e^{iuZ}]=E[e^{i(ul)'X}]=\exp\{ia'(ul)-\frac{1}{2}(ul)'B(ul)\}$$
$$=\exp\{iu(a'l)-\frac{1}{2}u^2(l'Bl)\}$$

对u是任意实数都成立，所以$Z\sim N(l'a,l'Bl)$

$\Leftarrow$若$Z=l'X\sim N(l'a,l'Bl)$

在$g(u)$中取$u=1$，得：$E[e^{il'X}]=\exp\{ia'l-\frac{1}{2}l'Bl\}$

由于l的任意性，所以$X\sim N(a,B)$

82

(2) 若$X=(X_1，X_2,\cdots,X_n)'$服从n元正态分布$N(a,B)$，而A为任意$m\times n$阵

则$Y=AX$服从m元正态分布$N(Aa,ABA')$

(注：A的秩$m(\le n)$)

证：$a_Y=E[Y]=E[AX]=AEX=Aa_X$

$$B_Y=E[(Y-a_Y)(Y-a_Y)']$$
$$=E[A(X-a_X)(X-a_X)'A']=AB_XA'$$

83

$$g_Y(t)=E[e^{it'Y}]=E[e^{it'AX}]=E[e^{i(A't)'X}]$$
$$=\exp\left\{ia_X'(A't)-\frac{1}{2}(A't)'B_X(A't)\right\}$$
$$=\exp\left\{i(Aa_X)'t-\frac{1}{2}t'(AB_XA')t\right\}$$
$$=\exp\left\{ia_Y't-\frac{1}{2}t'B_Yt\right\}$$

$\therefore Y$为正态随机变量$Y\sim N(Aa_x,AB_xA')$

84

§1.6 条件期望

一、条件分布及条件期望

（1）随机变量关于事件的条件分布及条件期望

定义 设X是概率空间(Ω,F,P)上的一个随机变量，$A\in F$，且$P(A)>0$，

则称 $F(x/A)=P\{X\le x/A\}=\dfrac{P\{X\le x\bigcap A\}}{P(A)}$，$x\in k$

为X关于事件A的条件分布函数。

条件数学期望：

$$E(X/A)=\int_{-\infty}^{\infty}x\cdot dF(x/A)$$

（2）离散型随机变量的条件分布律及条件期望

定义 设X,Y是两个离散型随机变量

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

若对于给定的y_j，有$\sum_i p_{ij}>0$

则称 $p_{i/j}=P\{X=x_i/Y=y_j\}=\dfrac{p_{ij}}{p_{\cdot j}}$

为X关于$Y=y_j$的条件分布律；

则X关于 $Y=y_j$的条件期望为：

$$E(X/Y=y_j)=\sum_{i=1}^{\infty}x_ip_{i/j}=\frac{\sum_{i=1}^{\infty}x_ip_{ij}}{P\{Y=y_j\}}$$

(3) 连续型随机变量的条件密度及条件期望

定义 设X，Y是两个连续型随机变量，

联合密度函数为：

$f(x,y)\qquad(x,y)\in R^2$

对给定的y有 $\int_{-\infty}^{+\infty}f(x,y)dx>0$

则称 $f(x/y)=\dfrac{f(x,y)}{f_Y(y)}=\dfrac{f(x,y)}{\int_{-\infty}^{+\infty}f(x,y)dx}$

为X关于$Y=y$的条件密度函数

X关于$Y=y$的条件期望为：

$$E(X/Y=y)=\int_{-\infty}^{\infty}xdF(x/y)=\int_{-\infty}^{\infty}x\cdot f(x/y)dx$$

$$=\frac{\int_{-\infty}^{\infty}xf(x,y)dx}{f_Y(y)}$$

$E(X/Y=y)$是在给定$Y=y$的条件下X取值的加权平均

例如：假设Y为身高，X为体重，则$E(X)$为平均体重，而$E(X/Y=y)$则为身高为y的平均体重。

当$Y=y$固定，$E[X/Y=y]$是一个数

当$Y=y$不固定，则$E[X/Y=y]$是y的函数，

而y是随机变量Y的样本值.

$\therefore E(X/Y)$是随机变量Y的函数，也是随机变量，称为X在Y下的条件期望，它的数学期望为：

$$E[E(X/Y)]=\int_{-\infty}^{\infty}E(X/Y=y)dF_Y(y)$$

$$=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}x\cdot f(x/y)dx\right]\cdot f_Y(y)dy$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x\cdot f(x,y)dxdy=EX$$

（二）条件期望的性质

(1) 全期望公式，若随机变量X与Y的期望存在

则$EX=E[E(X/Y)]=\int_{-\infty}^{\infty}E(X/Y=y)dF_Y(y)$

(i)对离散型随机变量Y

$$EX=\sum_{j=1}^{\infty}E(X/Y=y_j)\cdot P\{Y=y_j\}$$

(ii)对连续型随机变量Y

$$EX=\int_{-\infty}^{\infty}E(X/Y=y)f(y)dy$$

(2) $E[g(X)]=E[E(g(X)/Y)]$

(3) 对任意常数c_1,c_2随机变量X_1,X_2期望存在，则有$E(c_1X_1+c_2X_2/Y)=c_1E(X_1/Y)+c_2E(X_2/Y)$

一般有 $E\left[\sum_{i=1}^{n}\alpha_iX_i/Y\right]=\sum_{i=1}^{n}\alpha_iE(X_i/Y)$

(4) 若X与Y相互独立，则

$E(X/Y)=EX$

此时条件期望与无条件期望一致

(5) $E(X\cdot g(Y)/Y]=E[X/Y]\cdot g(Y)$

证 $E[Xg(Y)/Y=y]=E[Xg(y)/Y=y]$

$=g(y)E[X/Y=y]$

将y以Y代之得：$E(X\cdot g(Y)/Y]=g(Y)\cdot E(X/Y)$

(6) $E\{E[g(X,Y)/Y]\}=E[g(X,Y)]$

以连续情况证：$E\{E[g(X,Y)/Y]\}$

$$=\int_{-\infty}^{+\infty}E[g(X,Y)/Y=y]\cdot f_Y(y)dy$$

$$=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}g(x,y)\cdot f(x/y)dx\right]\cdot f_Y(y)dy$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)\cdot f(x,y)dxdy=E[g(X,Y)]$$

(7) 连续型全概率公式

$$P(A)=\int_{-\infty}^{+\infty}P(A/Y=y)dF_Y(y)$$

$$=\int_{-\infty}^{+\infty}P(A/Y=y)f_Y(y)dy$$

证：定义 $I_A=\begin{cases}1, & A\text{发生}\\ 0, & \text{不发生}\end{cases}$

则I_A称为事件A的示性函数。

由全期望公式得

$$P(A)=E(I_A)=E[E(I_A/Y)]=\int_{-\infty}^{+\infty}E(I_A/Y=y)dF(y)$$

$$=\int_{-\infty}^{+\infty}P(A/Y=y)dF(y)=\int_{-\infty}^{+\infty}P(A/Y=y)f(y)dy$$

例 设X,Y相互独立,分布函数分别为$F_X(x),F_Y(y)$,记$X+Y$的分布函数为$F_Z(z)$.

$$F_Z(z)=P\{X+Y\le z\}$$

$$=\int_{-\infty}^{\infty}P\{X+Y\le z|Y=y\}dF_Y(y)$$

$$=\int_{-\infty}^{\infty}P\{X+y\le z/Y=y\}dF_Y(y)$$

$$=\int_{-\infty}^{\infty}P\{X\le z-y/Y=y\}dF_Y(y)$$

$$=\int_{-\infty}^{\infty}F_X(z-y)dF_Y(y)$$

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(z-y)\cdot f_Y(y)dy$$

例：已知$X\sim U(0,a)$，$Y\sim U(X,a)$,

试求：(1) $E(Y/X=x),0<x<a$;

(2) $E(Y)$

解 (1)：$f_{Y/X}(y/x)=\begin{cases}\dfrac{1}{a-x}, & 0<x<y<a\\ 0, & \text{其他}\end{cases}$

对任意的$0<x<a$有

$$E(Y/X=x)=\int_x^a\frac{y}{a-x}dy=\frac{a+x}{2}$$

(2) $E(Y)=E[E(Y/X)]=E\left(\dfrac{a+X}{2}\right)=\dfrac{3}{4}a$

例：假设$N,X_1,X_2,\cdots$,相互独立，N服从参数为$\lambda>0$的泊松分布，而$X_j,j=1,2,\cdots$的特征函数为$g(t)$，试求$Z=\sum_{j=1}^{N}X_j$的特征函数。

解：$g_Z(t)=E(e^{itz})=E\left(e^{it\sum_{j=1}^{N}X_j}\right)=E\left\{E\left(e^{it\sum_{j=1}^{N}X_j}/N\right)\right\}$

$$=\sum_{l=0}^{\infty}E\left(e^{it\sum_{j=1}^{l}X_j}\right)P\{N=l\}=\sum_{l=0}^{\infty}P\{N=l\}[g(t)]^l$$

$$=\sum_{l=0}^{\infty}\frac{\lambda^le^{-\lambda}}{l!}[g(t)]^l=e^{-\lambda}\sum_{l=0}^{\infty}\frac{[\lambda g(t)]^l}{l!}=e^{\lambda[g(t)-1]}$$

同理，母函数为

$$H(s)=E(s^{z})=E\left(s^{\sum_{j=1}^{N}X_j}\right)=E\left\{E\left(s^{\sum_{j=1}^{N}X_j}/N\right)\right\}$$

$$=\sum_{l=0}^{\infty}E\left(s^{\sum_{j=1}^{l}X_j}\right)P\{N=l\}$$

$$=\sum_{l=0}^{\infty}P\{N=l\}[P(s)]^{l}$$

$$=G[P(s)]$$

作业：设X和Y的联合概率密度为：

$$f(x,y)=\begin{cases}\frac{1}{2}ye^{-xy} & ,\quad 0<x<\infty,0<y<2\\ 0 & ,\quad 其他\end{cases}$$

求$E\{e^{\frac{X}{2}}/Y=1\}$

- 复习概率论与数理统计方面的知识；
- 掌握特征函数与母函数的性质和计算方法；
- 重点掌握条件分布与条件期望的性质和计算方法。

预备知识结束