应用随机过程

Applied Stochastic Processes

- 概率论与数理统计
- 王英:1181955855@qq.com

前　言

随机过程是研究随机现象统计规律性的一门数学分支，属于随机数学的范畴，是现代科学的一门数学基础。随机过程在理论和应用上有着蓬勃的发展，在通信、电子、工程技术、控制、生物科学、环境科学、经济和社会科学等领域都有着广泛的应用。

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

引入

在自然界和人类的活动中经常遇到各种各样的现象，大体上分为两类：必然现象和随机现象 。

具有随机性的现象—随机现象

对随机现象的观察或为观察而进行的试验—随机试验　（有3个特征）

§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。

样本点(ω)： & 随机试验的基本结果

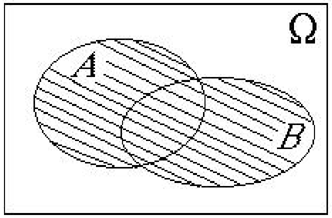
样本空间Ω： & 随机试验所有可能结果组成的集合

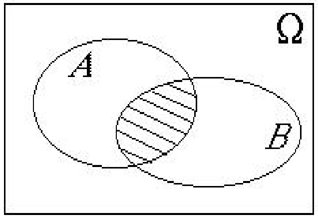
基本事件： & Ω中的样本点 ω

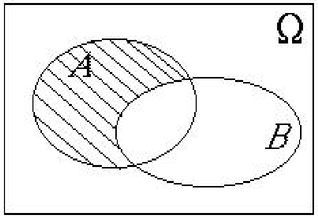
不可能事件： & 空集 ϕ

事件： &由基本事件组成的Ω中的子集A

事件的表示和运算







$A \cup B$

$A \cap B$ 或 AB

$A - B$ 或 $A\overline{B}$

事件运算的性质

假设A,B,C是任意事件，则他们满足：

(1)交换律 $A\cup B=B\cup A$

(2)结合律 $A\cup(B\cup C)=(A\cup B)\cup C$

$$A\cap(B\cap C)=(A\cap B)\cap C$$

(3)分配律 $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$

$$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$$

(4)对偶律 (De Morgan律)

$$\overline{A\cup B}=\overline{A}\cap\overline{B}\qquad\overline{A\cap B}=\overline{A}\cup\overline{B}$$

例1

计算 $(A \cup B) - B$ $= (A \cup B) \cap \overline{B}$

$$= (A \cap \overline{B}) \cup (B \cap \overline{B})$$

$$= (A \cap \overline{B}) \cup \Phi$$

$$= A\overline{B}$$

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）

事件域的定义

设Ω是样本空间，$\mathscr{F}$是Ω中随机事件组成的集合族，若满足

（1）$\Omega \in \mathscr{F}$;

（2）如果$A \in \mathscr{F}$，则 $\overline{A} \in \mathscr{F}$;

（3）如果 $A_i \in \mathscr{F}$，则 $\bigcup\limits_{i=1}^{+\infty} A_i \in \mathscr{F}$.

那么$\mathscr{F}$称为事件域或者 σ-代数。

$\mathscr{F}$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$P(\bullet)$

概率的定义

设$\mathscr{F}$是定义在样本空间Ω上的事件域，$P(A)$ 是定义在$\mathscr{F}$上的非负集函数（$A \in \mathscr{F}$为事件），且满足：

（1）$0 \le P(A) \le 1$；

（2）$P(\Omega) = 1$；

（3）对任意事件A_i，$i = 1,2,\cdots,$满足$A_i \cap A_j = \Phi, i \ne j$

$$P\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} P(A_i)$$

则称P是定义在$\mathscr{F}$上的概率，$(\Omega,\mathscr{F},P)$为概率空间，$P(A)$ 是事件A的概率。

概率的基本性质

(1) $P(\Phi)=0$,

(2) $P(\overline{A})=1-P(A)$,

(3)若$A\subset B$, 则$P(A)\le P(B)$, —单调性

(4)若$A\subset B$, 则$P(B-A)=P(B)-P(A)$,

(5)推论：$P(B-A)=P(B)-P(AB)$,

(6)推论：$P(A\cup B)=P(A)+P(B)-P(AB)$.

例2

$P(A)=0.4$，$P(B)=0.3$，$P(A\cup B)=0.6$， 求 $P(A-B)$.

解：因为 $P(A-B) = P(A)-P(AB)$，所以先求 $P(AB)$

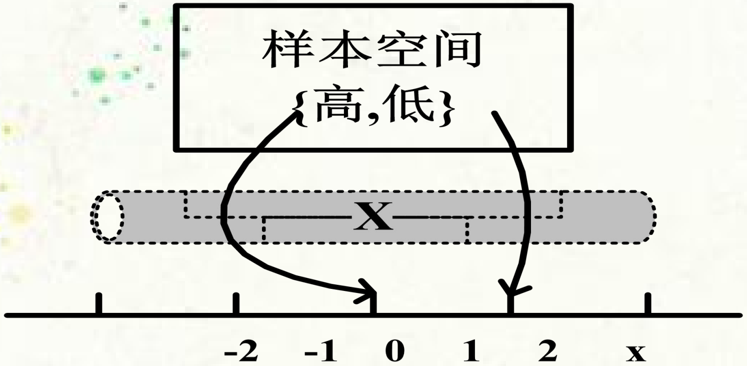
由加法公式得 $P(AB) = P(A)+P(B)-P(A\cup B)$

$= 0.4+0.3-0.6=0.1$

所以 $P(A-B) = P(A)-P(AB) = 0.3$

§1.2　随机变量与分布函数

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。

定义1.2.1 设 $(\Omega,\mathscr{F},P)$为概率空间，X为定义在Ω上取值于实数集R上的函数$(\omega \to X(\omega))$，且对$\forall x \in R$，$\{\omega : X(\omega) \le x\} \in \mathscr{F}$，则称X为定义在Ω上的随机变量

关于随机变量的两点说明：

(1)定义中ω为自变量，为了书写方便，简记

$$\{\omega : X(\omega) \le a\} \triangleq \{X \le a\}，其中a是实数。$$

一般随机变量符号常用大写字母X,Y,Z等表示。

(2)$X(\omega)$满足$\{\omega : X(\omega) \le a\} \in \mathscr{F}$,则易证：

$$\forall\ a,b \in R, \{X > a\}, \{X < a\}, \{X = a\}, \{a \le X < b\},$$
$$\{a < X < b\}, \{a \le X \le b\} \in \mathscr{F}.$$

定义1.2.2 设$X(\omega)$是随机变量，函数

$$F(x)=P(X\le x),\qquad -\infty<x<+\infty$$

称为随机变量X的分布函数。

分布函数的含义： 分布函数$F(x)$表示随机变量X取值不超过x的概率 (x为任意实数).

分布函数 $F(x)$ 的性质：

(1) $0\le F(x)\le 1$;

(2) $F(x)$是非降函数，即$x_1<x_2 \Longrightarrow F(x_1)\le F(x_2)$;

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

(4)$F(x)$是右连续的, 即$F(x+0)=\lim\limits_{t\to x^+}F(t)=F(x)$.

随机变量的类型：

离散型：$P(X = x_k) = p_k \quad \sum_{k=1}^{\infty} p_k = 1$

$$F(x) = P(X \le x) = \sum_{x_k \le x} p_k$$

连续型：$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$

（其中$f(x)$为概率密度函数，$\int_{-\infty}^{+\infty} f(x)dx = 1$）

$$f(x) = \frac{dF(x)}{dx}$$

常用离散型随机变量的概率分布用分布列

0－1分布 $P(X=1)=p, P(X=0)=q$

伯努利分布 $P(X=k)=p^k(1-p)^{1-k}, k=0,1$

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

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

离散均匀分布 $P(X=x_k)=\dfrac{1}{n}$

常用连续型随机变量的概率分布用概率密度

均匀分布

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

指数分布

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

正态分布

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

三、二维随机变量及其分布函数

定义1.4.2 任对实数 x 和 y，称

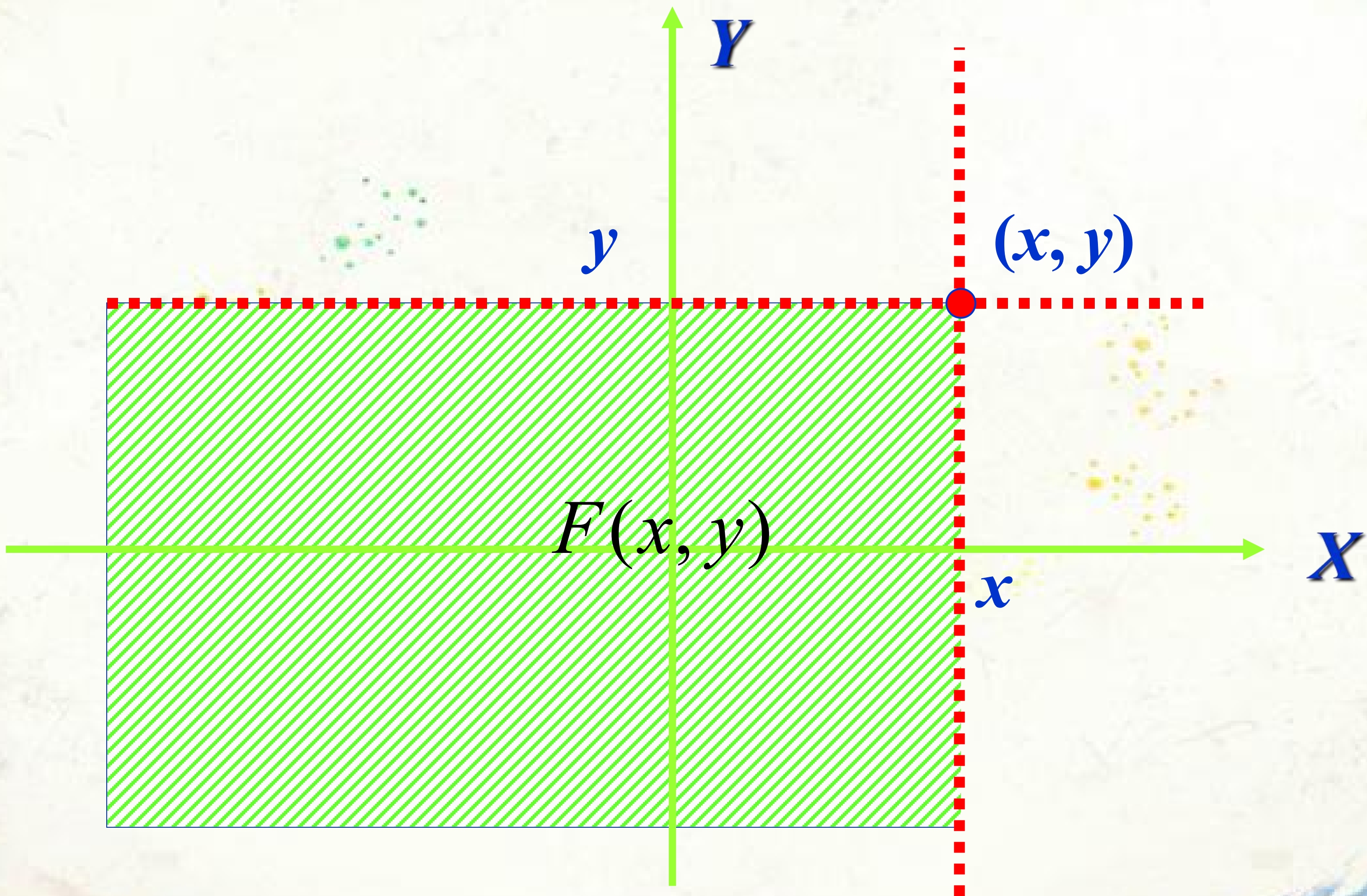
$$F(x, y) = P(X \le x, Y \le y)$$

为 (X, Y) 的联合分布函数.

注意：(几何意义）

F(x, y)为(X, Y)落在点(x, y)的左下区域的概率.

二维联合分布函数



在概率统计这门课中我们学习了多维随机变量的概念（以二维随机变量为主）。二维联合分布函数（分布列、密度函数）含有丰富的信息，主要有以下三个方面信息：

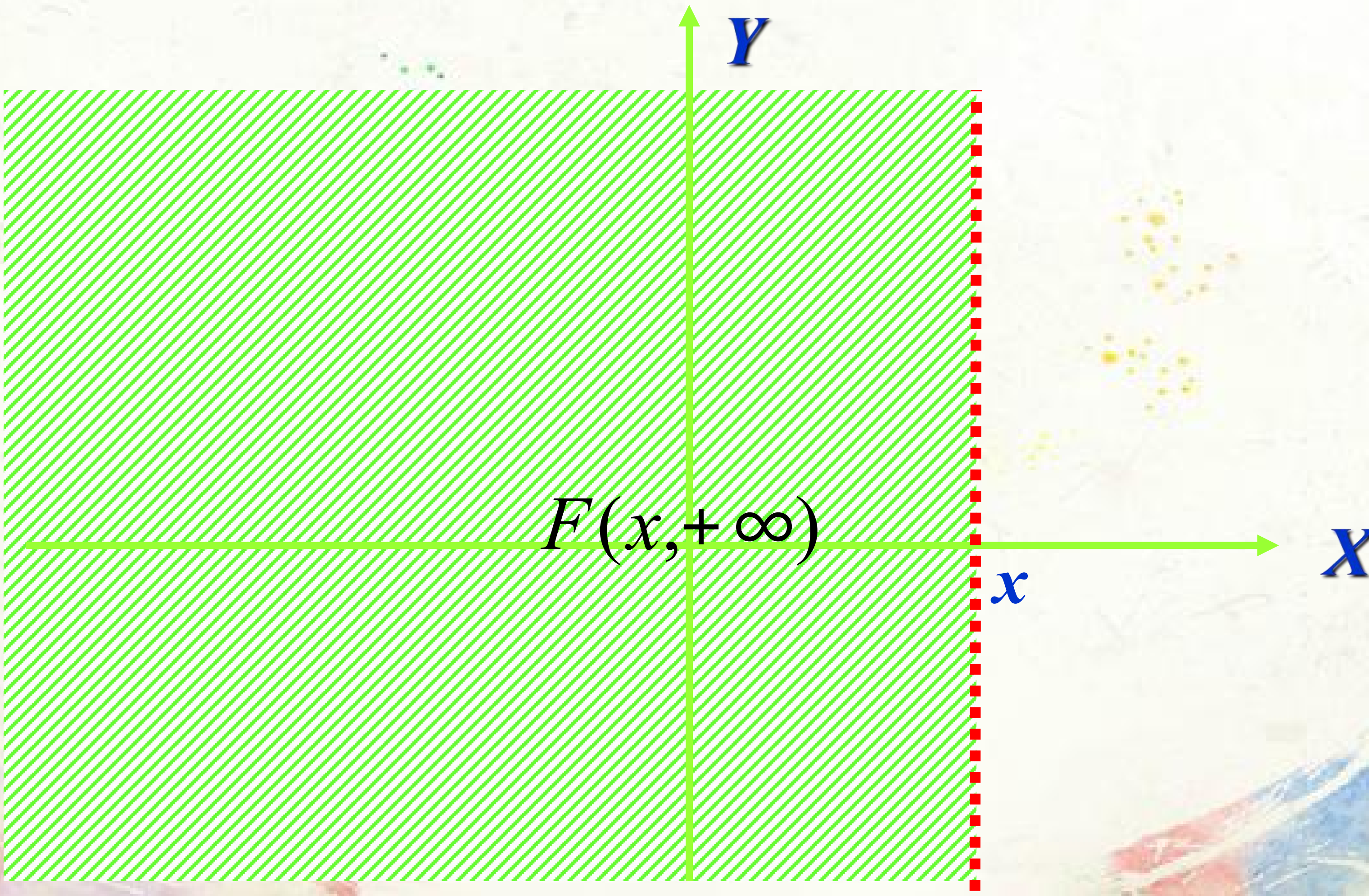
1. 每个分量的分布（边缘分布）。

2. 两个分量之间的关联程度（相关系数）。

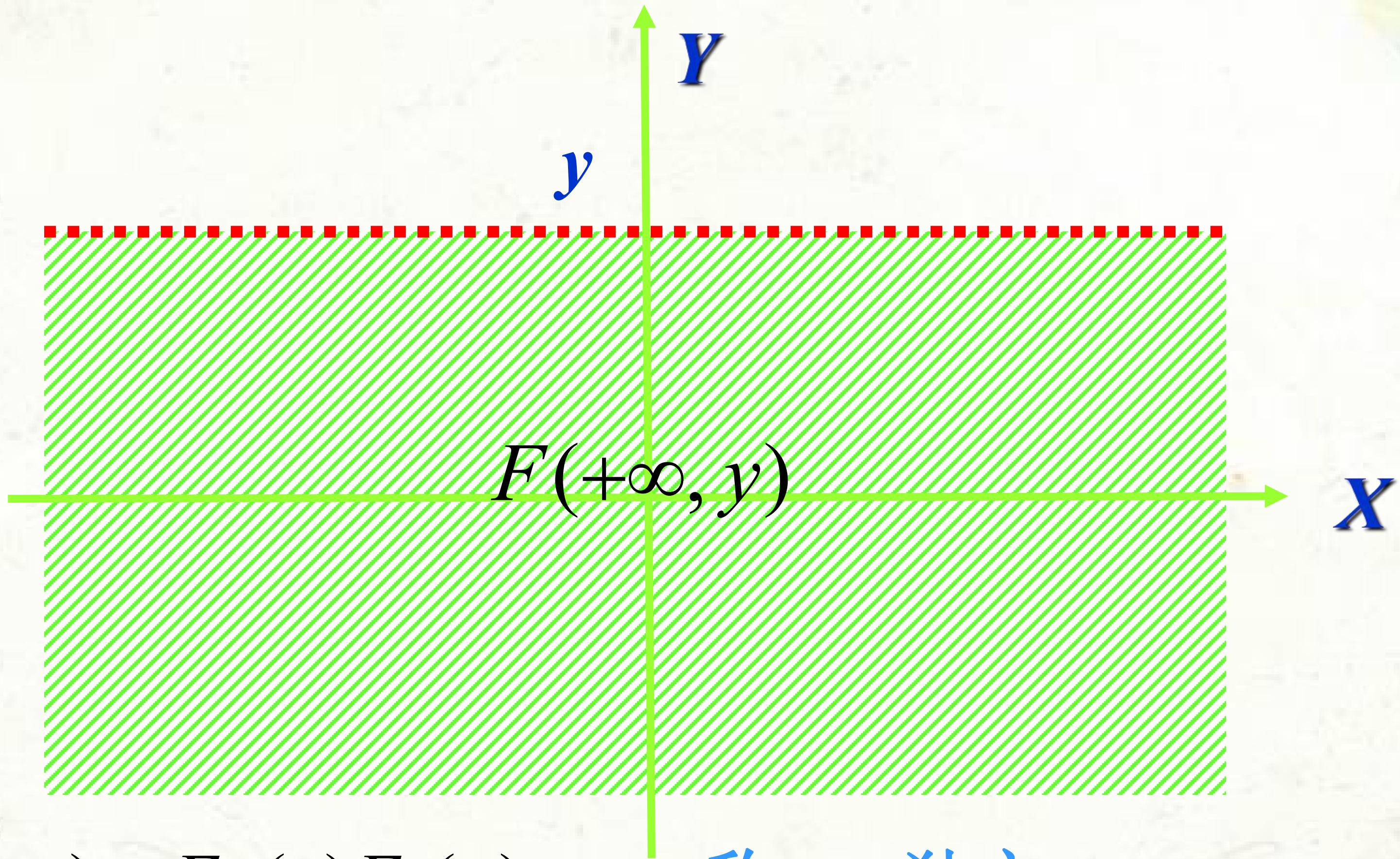
3. 给定一个分量时，另一个分量的分布（条件分布）。

设(X,Y)的联合分布函数为$F(x,y)$，则X,Y的分布函数为$F_X(x),F_Y(y)$，依次称为关于X和Y的边缘分布函数，且有

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



同理 $F_Y(y) \triangleq F(+\infty, y)$



若 $F(x,y) = F_X(x)F_Y(y)$ ——称X,Y独立.

离散的情形

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

为(X,Y) 的联合分布列，其表格形式如下:

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

若(X,Y)的联合分布列 $P(X=x_i, Y=y_j)=p_{ij}$

则 $$P_{i\cdot}=P(X=x_i)=\sum_{j=1}^{\infty} p_{ij} \qquad (i=1,2,\cdots)$$

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

$P_{i\cdot}$ 及$P_{\cdot j}$分别称为(X,Y)关于X和Y的边缘分布列.

X和Y相互独立的充要条件是：$P_{ij}=P_{i\cdot}P_{\cdot j}$

连续的情形

若随机变量(X,Y)的概率密度函数为$f(x,y)$,

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

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

$f_X(x)$及$f_Y(y)$分别称为(X,Y)关于X和Y的边缘概率密度。

X和Y相互独立的充要条件是:

$$f(x,y)=f_X(x)\,f_Y(y)$$

例3

(X, Y) 的联合分布列为:

$X \backslash Y$	0	1
0	0.3	0.4
1	0.2	0.1

问 X与Y是否独立?

解: 边缘分布列分别为:

X	0	1
P	0.7	0.3

Y	0	1
P	0.5	0.5

因为 $P(X=0,\ Y=0)=0.3$

$P(X=0)P(Y=0)=0.7\times 0.5=0.35$ 所以不独立

一、随机变量的数学期望

定义3.1：设随机变量X的分布函数为F(x)，若

$\int_{-\infty}^{+\infty}|x|dF(x)<+\infty$,则 $E(X)=\int_{-\infty}^{+\infty}xdF(x)$ 称为随机变量X的数学期望或均值。

若X是连续型随机变量，则

$$E(X)=\int_{-\infty}^{+\infty}xF'(x)dx=\int_{-\infty}^{+\infty}xf(x)dx$$

若X是离散型随机变量，则

$$E(X)=\sum_{i=1}^{\infty}x_i p(X=x_i)$$

注：随机变量X的数学期望是对所有可能取值的加权平均

二、方差

定义3.2：设随机变量X的分布函数为F(x)，则称

$$D(X)=E\{[X-E(X)]^2\}=E(X^2)-[E(X)]^2$$

为随机变量X的方差，有时记$D(X)=VarX=\sigma_X{}^2$,称

$\sigma_X=\sqrt{D(X)}$为X的标准差。

注：随机变量X的方差反应的是其偏离程度

定理3.3：(切比雪夫不等式)设随机变量X的期望和方差分别为$E(X),D(X)$,则对任给的$\varepsilon>0$,有

$$P\{|X-E(X)|\geq\varepsilon\}\leq\frac{D(X)}{\varepsilon^2}$$

或

$$P\{|X-E(X)|<\varepsilon\}\geq 1-\frac{D(X)}{\varepsilon^2}$$

注：切比雪夫不等式描述了不管X的分布是什么，X落在$(E(X)-\varepsilon,E(X)+\varepsilon)$中的概率不小于$1-\dfrac{D(X)}{\varepsilon^2}$。

三、协方差与相关系数

定义3.4 设X,Y是随机变量, $EX^2<\infty, EY^2<\infty$则称

$$\mathrm{Cov}(X,Y)=E\left[\left(X-EX\right)\left(Y-EY\right)\right]=E(XY)-E(X)E(Y)$$

为(X,Y)的协方差。

相互独立 ⟹ 不相关

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

相互独立 ⟸ ? 不相关

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。

定义3.5 设X,Y是随机变量,$0<D(X)<+\infty$, $0<D(Y)<+\infty$,称

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

ρ_{XY}称为归一化的协方差系数或相关系数。

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。

四、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，

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

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。

常用分布的数学期望和方差

常用离散型随机变量的概率分布用分布列

0—1分布

伯努利分布

$$\left.\begin{array}{l}P(X=1)=p, P(X=0)=q \\ X\sim B(1,p)\quad E(X)=p, D(X)=p(1-p)\end{array}\right.$$

二项分布

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

$$X\sim B(n,p)\quad E(X)=np, D(X)=np(1-p)$$

泊松分布

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

$$X\sim P(\lambda)\quad E(X)=\lambda, D(X)=\lambda$$

几何分布

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

$$X\sim Ge(p)\quad E(X)=\frac{1}{p}, D(X)=\frac{1-p}{p^2}$$

离散均匀分布

$$P(X=x_k)=\frac{1}{n}$$

$$X\sim DU(a,b)\quad E(X)=\frac{a+b}{2}, D(X)=\frac{(b-a+1)^2-1}{12}$$

连续型随机变量的概率分布用概率密度描述

均匀分布

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & 其它\end{cases}$$

$$X\sim U(a,b) \quad E(X)=\frac{a+b}{2}, D(X)=\frac{(b-a)^2}{12}$$

指数分布

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

$$X\sim Exp(\lambda) \qquad E(X)=\frac{1}{\lambda}, D(X)=\frac{1}{\lambda^2}$$

正态分布

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

$$X\sim N(\mu,\sigma^2) \qquad E(X)=\mu, D(X)=\sigma^2$$

§1.5 独立性与条件期望

一、独立性

定义5.1：(1)A,B为两个事件，若

$$P(A\cap B)=P(A)P(B)$$

则称 A,B相互独立。

特别的 设$A_1,A_2,\cdots,A_n$为n个事件，若对任意$m\le n$，$1\le k_1<k_2<\cdots<k_m\le n$，有

$$\mathrm{P}\left(\bigcap_{j=1}^{m}A_{k_j}\right)=\prod_{j=1}^{m}P(A_{k_j})$$

则称$A_1,A_2,\cdots,A_n$相互独立。注：$A_1,A_2,\cdots,A_n$两两独立不一定相互独立。

（2）设$\{A_i, i \in I\}$是一簇事件，若对I的任意有限子集$\{i_1, i_2, \cdots, i_k\} \neq \phi$,有

$$\mathrm{P}(\bigcap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j})$$

则称$\{A_i, i \in I\}$相互独立。

（3）设$\{X_i, i \in I\}$是Ω上的一簇随机变量，若σ 代数族$\{\sigma(X_i), i \in I\}$是独立事件类,则称$\{X_i, i \in I\}$相互独立。

注：设随机变量$X_1, X_2, \cdots, X_n$独立$\Leftrightarrow$其联合分布函数可分解为边际分布函数的乘积，即

$$\mathrm{F}(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$$

二、条件期望

回忆：条件概率

在事件B已发生这一条件下，事件A发生的概率。

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

显然，给定事件B，条件概率有如下性质：

(1) $\forall A \in F, 0 \le P(A|B) \le 1;$

(2) $P(\Omega|B) = 1;$

(3) $A_i \in F(i = 1, 2, \cdots)$, 且 $A_i \cap A_j = \phi(i \neq j)$, 则

$$P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) = \sum_{i=1}^{\infty} P(A_i|B)$$

定义5.2：记 $P_B = P(\cdot|B)$, 则 P_B 是可测空间 (Ω, F) 上的概率, 称 (Ω, F, P_B) 是条件概率空间。

二. 条件分布

离散型：　若 $P(Y=y_j)>0$ 则称

$$P(X=x_i \mid Y=y_j)=\frac{P(X=x_i,\ Y=y_j)}{P(Y=y_j)}=\frac{P_{ij}}{P_{\cdot j}}$$

$i=1,2,3,\cdots$

为条件 $Y=y_j$ 下，随机变量X的分布列.

同理 $$P(Y=y_j \mid X=x_i)=\frac{P(X=x_i,\ Y=y_j)}{P(X=x_i)}=\frac{P_{ij}}{P_{i\cdot}}$$

$j=1,2,3,\cdots$

称为条件 $X=x_i$ 下，随机变量Y的分布列.

连续型：$f(x|y)=\dfrac{f(x,y)}{f_Y(y)}, \quad -\infty < x < +\infty$

称为在条件$Y=y$下随机变量X的条件密度函数.

$$f(y|x)=\frac{f(x,y)}{f_X(x)} \quad -\infty < y < +\infty$$

称为在条件$X=x$下随机变量Y的条件密度函数.

三、条件数学期望

条件数学期望是随机过 程中最基本最重要的概 念之一.

X的条件数学期望

$$E(X \mid Y = y) = \begin{cases} \sum_i x_i P(X = x_i \mid Y = y) \\ \int_{-\infty}^{\infty} x f(x \mid y) \mathrm{d}x \end{cases}$$

实用案例

- 若用X表示中国成年人的身高，则E(X)表示中国成年人的平均身高。若用Y表示中国成年人的足长，则E(X|Y=y)表示足长为y的中国成年人的平均身高，我国公安部门研究获得
- $$E(X|Y=y)=6.876y.$$
- 这个公式很有用，比如测的案犯留下的足印25.3厘米，则可推算此案犯身高174厘米。

引进$E(X|Y)$是随机变量Y的函数, 当$Y=y_j$时，

$E(X|Y)$取值为$E(X|Y=y_j)$,

由于$E(X|Y)$是随机变量 Y的函数,也是随机变量,

故它的数学期望为:

$$E(E(X|Y))=\sum_j E(X|Y=y_j)\ P(Y=y_j)$$

$$=E(X).$$

这就是著名的重期望公式。

重期望公式具体使用如下：

（i）如果Y是一个离散型随机变量，则有

$$E(X)=\sum_{j}E(X\mid Y=y_j)P(Y=y_j).$$

（ii）如果Y是一个连续型随机变量，则有

$$E(X)=\int_{-\infty}^{+\infty}E(X\mid Y=y)\,f(y)dy.$$

（2）$E\left[g(X)\right]=E\left[E\left(g(X)\middle|Y\right)\right]$

⑶对任意常数c_1,c_2随机变量X_1,X_2期望存在

则有$E\left(c_1X_1+c_2X_2\middle|Y\right)=c_1E\left(X_1\middle|Y\right)+c_2E\left(X_2\middle|Y\right)$

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

（4） 若 X 与 Y 相 互 独 立 则

$E\left(X\middle|Y\right)=EX$

此 时 条 件 期 望 与 无 条 件 期 望 一 致

（5）$E\left(X\cdot g(Y)\middle|Y\right]=E\left[X\middle|Y\right]\cdot g(Y)$

证 $E\left[Xg(Y)\middle|Y=y\right]=E\left[Xg(y)\middle|Y=y\right]$

$$=g(y)E\left[X\middle|Y=y\right]$$

将y以Y代之得：$E\left(X\cdot g(Y)\middle|Y\right]=g(Y)\cdot E\left(X\middle|Y\right)$

（6）$E\left\{E\left[g(X,Y)\middle|Y\right]\right\}=E\left[g(X,Y)\right]$

以连续情况证：$E\left\{E\left[g(X,Y)\middle|Y\right]\right\}$

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

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

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

例4：一矿工被困在有三个门的矿井里。第一个门通一坑道，沿此坑道走3个小时可以到达安全区；第二个门通一坑道，沿此坑道走5个小时又回到原处；第三个门通一坑道，沿此坑道走7个小时也回到原处。假定此矿工总是等可能地在三个门中中选择一个，试求他平均要用多少时间才能到达安全区？

解：设该矿工需要X小时到达安全区，直接写出X的分布列时困难的，所以无法直接求出E(X).为此记Y表示第一次所选的门，{Y=i}就是选择第i个门。由题设知

$$P(Y=1)=P(Y=2)=P(Y=3)=\frac{1}{3}.$$

因为选择第一个门后3个小时可以到达安全区，所以有

$$E(X\mid Y=1)=3.$$

因为选择第二个门后5个小时回到原处，所以有

$$E(X \mid Y=2) = 5 + E(X).$$

因为选择第三个门后7个小时回到原处，所以有

$$E(X \mid Y=3) = 7 + E(X).$$

综上所述，由重期望公式

$$E(X) = \quad {}_{j} E(X \mid Y = y_j) P(Y = y_j)$$

$$= 3 \times \tfrac{1}{3} + (5 + E(X)) \times \tfrac{1}{3} + (7 + E(X)) \times \tfrac{1}{3}$$

解之可得

$$E(X) = 15.$$

即该矿工平均需要15小时才能到达安全区。

例5

设 X 和 Y 的联合概率密度为：

$$f(x,y)=\begin{cases}\dfrac{1}{2}e^{-y}, & y>|x|,-\infty<x<+\infty\\ 0, & \text{其他}\end{cases}$$

求 $E\{Y|X=x\}$

作业:

1.设随机变量X的概率分布函数

$$F(x)=\begin{cases} A+Be^{-\frac{x^2}{2}}, & x>0, \\ 0, & x\le 0, \end{cases}$$ 求 **A和B.**

2.设随机变量X的密度函数为

$$p(x)=\begin{cases} ke^{-3x} & x>0 \\ 0 & x\le 0 \end{cases}$$

求 (1) 常数 k, (2) $F(x)$.

3. 设(X, Y) 的联合分布列为:

$X \backslash Y$	0	1
0	0.2	0.3
1	0.2	0.3

问 X与Y 是否独立?

4. 一只老鼠被困在一个有三个通道的隧道之中，从第一个通道行进2个小时后将到达安全地带；从第二个通道行进3个小时后将绕回原地；从第三个通道行进5个小时后将绕回隧道原地。假定该老鼠对此隧道的通道完全未知，那么该老鼠到达安全地带所需要的平均时间是多少?

5. 设X和Y的联合概率密度为：

$$f(x,y)=\begin{cases}\frac{1}{y}e^{-\frac{x}{y}}e^{-y}, & 0<y<+\infty, 0<x<+\infty\\ 0, & \text{其他}\end{cases}$$

求$E\{X|Y=y\}$

6. 已知 $X \sim U(0,a)$， $Y \sim U(X,a)$，

试求：（1） $E(Y|X=x), 0<x<a$;

（2） $E(Y)$

复习概率论与数理统计方面的知识；
重点掌握条件分布与条件期望的性质和计算方法。