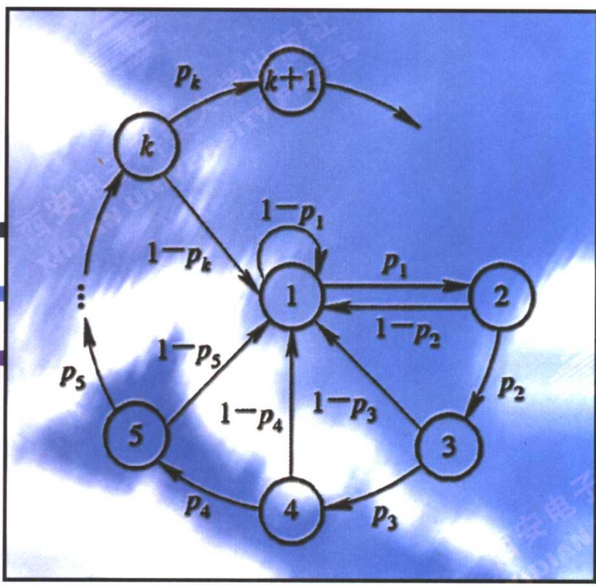
研究生
学习指导丛书

Postgraduate

《随机过程》同步学习指导



▪ 张卓奎 陈慧婵 编著

- 重点与难点
- 内容提要
- 例题解析
- 习题全解

西安电子科技大学出版社 http://www.xduph.com

研究生学习指导丛书

《随机过程》同步学习指导

张卓奎　陈慧婵　编著

西安电子科技大学出版社

2004

内 容 简 介

本书是《随机过程》一书的配套教材，也是《随机过程》一书的扩展，其内容包括概率论基础、随机过程的基本概念、随机分析、平稳过程、马尔可夫过程、排队和服务系统、更新过程、时间序列分析和鞅过程.

每章分为三部分：内容提要、例题解析和习题全解. 内容提要对每章的主要内容作了较为深入的概括；例题解析针对主要内容选择了适量的例题进行了解析；习题全解对《随机过程》一书每章后的所有习题作了较为详细的解答. 在每章的开头还指出了本章学习的重点和难点.

本书文字通俗易懂，概念清晰，可供工科研究生和本科高年级学生使用，也可作为有关科技人员的参考书.

研究生学习指导丛书
《随机过程》同步学习指导
张卓奎　陈慧婵　编著

策　　划　夏大平
责任编辑　夏大平
出版发行　西安电子科技大学出版社(西安市太白南路2号)
电　　话　(029)88242885　88201467　　邮　编　710071
http://www.xduph.com　　E-mail: xdupfxb@pub.xaonline.com
经　　销　新华书店
印刷单位　陕西光大印务有限责任公司
版　　次　2004年7月第1版　2004年7月第1次印刷
开　　本　787毫米×1092毫米　1/16　印张 14.375
字　　数　338千字
印　　数　1～4 000册
定　　价　18.00元
ISBN 7-5606-1282-2/O·0066(课)
XDUP 1553A01-1

前　　言

随机过程是研究随时间演变的随机现象的一门学科，已广泛地应用于通信、自动控制、生物物理、系统工程、空间技术、社会科学等领域，并且在这些领域中显示出十分重要的作用. 目前“随机过程”已经成为高等院校中工科研究生的一门重要的理论基础课.

在学习“随机过程”课程中，为了透彻理解随机过程的基本理论，掌握随机过程的基本方法，练习和巩固是不可缺少的. 为此，我们在编写《随机过程》一书的同时编写了本书，它是《随机过程》一书的配套教材，同时也是《随机过程》一书的扩展.

全书共分 9 章. 第 1 章为概率论基础；第 2 章为随机过程的基本概念；第 3 章为随机分析；第 4 章为平稳过程；第 5 章是马尔可夫过程；第 6 章是排队和服务系统；第 7 章是更新过程；第 8 章是时间序列分析；第 9 章是鞅过程.

本书每章由内容提要、例题解析、习题全解三部分组成. 内容提要对每章的基本内容以及读者应掌握的重点作了概括；为了使每章的例题不与课本及习题重复，例题解析部分特选编了适量的例题，用例题的形式来体现对每章的基本要求，通过例题解析，帮助读者理解基本概念，提高分析问题和解决问题的能力；针对读者(特别是初学者)在学习“随机过程”课程时往往感到内容看明白了，但题目不会做的现象，习题全解部分对《随机过程》一书每章的所有习题都作了比较详细的解答，同时习题全解也是例题解析的有益补充. 需要指出，这些题目的解法不是惟一的，也许读者会有更好的解法. 在每章的开头还指出了本章学习的重点和难点.

本书的编写得到了许多同行的热情支持和帮助；得益于许多专业课老师和学生的建议；西安电子科技大学出版社的领导和策划编辑夏大平也非常关心本书的出版，夏老师亲自担任此书的责任编辑，对该书的出版付出了辛勤的劳动. 编者在此一并致以诚挚的谢意！

由于编者水平有限，书中难免存在错误和不妥之处，恳请广大读者批评、指正.

编者

2004 年 2 月

目　　录

第 1 章　概率论基础

【重点与难点】 随机变量及其分布是本章的重点，要求既能正确理解随机变量的定义，熟练掌握随机变量的概率分布，包括分布函数、离散型随机变量的分布律、连续型随机变量的概率密度函数、边缘分布、条件分布，随机变量的数字特征，包括数学期望、方差、协方差、矩，随机变量的独立性、相关性，又能熟练地求出随机变量的分布、数字特征，判断随机变量是否独立、是否相关.

随机变量的特征函数、n 维正态随机变量和条件数学期望是本章的难点.

1.1　内容提要

1. 概率空间

定义

(1) 设 Ω 是样本空间，$\mathscr{F}$ 是 Ω 的某些子集构成的集合，如果满足：

① $\Omega\in\mathscr{F}$；

② 若 $A\in\mathscr{F}$，则 $\overline{A}\in\mathscr{F}$；

③ 若 $A_n\in\mathscr{F}$，$n=1, 2, \cdots$，则

$$\bigcup_{n=1}^{\infty}A_n\in\mathscr{F}$$

那么称 $\mathscr{F}$ 为一事件域，也称 $\mathscr{F}$ 为 σ 域.

(2) 设 Ω 是样本空间，$\mathscr{F}$ 是一事件域，定义在 $\mathscr{F}$ 上的实值函数 $P(\cdot)$ 如果满足：

① $\forall\ A\in\mathscr{F}$，$P(A)\geqslant 0$；

② $P(\Omega)=1$；

③ 若 $A_n\in\mathscr{F}$，$n=1, 2, \cdots$，且 $A_iA_j=\varnothing$，$i\neq j$，i，$j=1, 2, \cdots$，则

$$P\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}P(A_n)$$

那么称 P 是二元组 $(\Omega, \mathscr{F})$ 上的概率，称 $P(A)$ 为事件 A 的概率，称三元组 $(\Omega, \mathscr{F}, P)$ 为概率空间.

2. 随机变量及其分布

定义

(1) 设 $(\Omega, \mathscr{F}, P)$ 为一概率空间，定义在 Ω 上的实函数 $X(\cdot)$，如果 $\forall\ x\in\mathbf{R}$，$\{\omega|X(\omega)\leqslant x\}\in\mathscr{F}$，则称 X 是 $\mathscr{F}$ 的随机变量. 称

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

为随机变量 X 的分布函数.

(2) 设 $(\Omega, \mathscr{F}, P)$ 为一概率空间，定义在 Ω 上的 n 元实函数 $\boldsymbol{X}(\cdot)=(X_1(\cdot),$

$X_2(\cdot)$, $\cdots$, $X_n(\cdot)$)，如果$\forall\ \boldsymbol{x}=(x_1, x_2, \cdots, x_n)\in\mathbf{R}^n$，$\{\omega|X_1(\omega)\leqslant x_1, X_2(\omega)\leqslant x_2, \cdots, X_n(\omega)\leqslant x_n\}\in\mathscr{F}$，则称$\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$为$n$维随机变量或$n$维随机向量. 称

$$F(\boldsymbol{x}) = F(x_1, x_2, \cdots, x_n) = P(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n)$$
$$\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbf{R}^n$$

为$\boldsymbol{X}$的联合分布函数.

(3) 设$\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$是一$n$维随机变量，称$F_{X_i}(x_i)=P(X_i\leqslant x_i)$，$-\infty<x_i<+\infty$，为$\boldsymbol{X}$关于$X_i$的边缘分布函数，如果对于任意的$x_1, x_2, \cdots, x_n\in\mathbf{R}$，有

$$F(x_1, x_2, \cdots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$$

则称随机变量$X_1, X_2, \cdots, X_n$相互独立.

3. 随机变量的数字特征

1) 定义

(1) 设X是一个随机变量，$F(x)$是其分布函数，若$\int_{-\infty}^{+\infty}|x|\ \mathrm{d}F(x) <+\infty$，则称$EX \xlongequal{\text{def}} \int_{-\infty}^{+\infty} x\ \mathrm{d}F(x)$为随机变量$X$的数学期望或均值.

(2) 设X是随机变量，若$E|X|^2<+\infty$，则称

$$DX \xlongequal{\text{def}} E(X - EX)^2$$

为随机变量X的方差.

(3) 设X, Y是随机变量，若$E|X|^2<+\infty$，$E|Y|^2<+\infty$，则称

$$\mathrm{cov}(X, Y) \xlongequal{\text{def}} E[(X - EX)(Y - EY)]$$

为随机变量X, Y的协方差. 若$DX>0$，$DY>0$，则称

$$\rho_{XY} \xlongequal{\text{def}} \frac{\mathrm{cov}(X, Y)}{\sqrt{DX}\sqrt{DY}}$$

为随机变量X, Y的相关系数，若$\rho_{XY}=0$，则称X, Y不相关.

(4) 设$\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$是$n$维随机变量，则$E\boldsymbol{X} \xlongequal{\text{def}} (EX_1, EX_2, \cdots, EX_n)$为$n$维随机变量$\boldsymbol{X}$的均值向量. 称

$$\boldsymbol{B} \xlongequal{\text{def}} \begin{bmatrix} \mathrm{cov}(X_1, X_1) & \mathrm{cov}(X_1, X_2) & \cdots & \mathrm{cov}(X_1, X_n) \\ \mathrm{cov}(X_2, X_1) & \mathrm{cov}(X_2, X_2) & \cdots & \mathrm{cov}(X_2, X_n) \\ \vdots & \vdots & & \vdots \\ \mathrm{cov}(X_n, X_1) & \mathrm{cov}(X_n, X_2) & \cdots & \mathrm{cov}(X_n, X_n) \end{bmatrix}$$

为n维随机变量$\boldsymbol{X}$的协方差矩阵.

2) 主要的性质和结论

(1) 设X是一随机变量，其分布函数为$F(x)$，$y=g(x)$是连续函数，如果$\int_{-\infty}^{+\infty}g(x)\ \mathrm{d}F(x)$存在，则

$$EY = E[g(X)] = \int_{-\infty}^{+\infty} g(x)\ \mathrm{d}F(x)$$

(2) 设$\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$是$n$维随机变量，其联合分布函数为$F(x_1, x_2, \cdots, x_n)$，

$g(x_1, x_2, \cdots, x_n)$是连续函数，如果$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(x_1, x_2, \cdots, x_n)\,\mathrm{d}F(x_1, x_2, \cdots, x_n)$存在，则

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

(3) 设a，b是任意的常数，则$E(aX+bY)=aEX+bEY$.

(4) 设X，Y相互独立，a，b是常数，则$EXY=EXEY$，$D(aX+bY)=a^2DX+b^2DY$.

(5) 设X是随机变量，若$E|X|^r<+\infty$，$r>0$，则

$$P(|X| \geqslant \varepsilon) \leqslant \frac{E|X|^r}{\varepsilon^r}$$

特别地，有

$$P(|X - EX| \geqslant \varepsilon) \leqslant \frac{DX}{\varepsilon^2}$$

(6) 设X是随机变量，则$DX=0$的充要条件是$P(X=C)=1$(C是常数).

4. 随机变量的特征函数

1) 定义

(1) 设X是(实)随机变量，其分布函数为$F(x)$，则称

$$\varphi(t) \xlongequal{\text{def}} E[\mathrm{e}^{\mathrm{j}tX}] = \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{j}tx}\,\mathrm{d}F(x),\ -\infty < t < +\infty,\ \mathrm{j} = \sqrt{-1}$$

为随机变量X的特征函数.

(2) 设$\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$是$n$维随机变量，其联合分布函数为$F(\boldsymbol{x})=F(x_1, x_2, \cdots, x_n)$，则称

$$\begin{aligned}\varphi(\boldsymbol{t}) = \varphi(t_1, t_2, \cdots, t_n) &\xlongequal{\text{def}} E[\mathrm{e}^{\mathrm{j}\boldsymbol{t}\boldsymbol{X}^{\mathrm{T}}}] = E[\exp(\mathrm{j}\sum_{k=1}^{n} t_k X_k)] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \exp[\mathrm{j}(t_1x_1 + t_2x_2 + \cdots + t_nx_n)]\,\mathrm{d}F(x_1, x_2, \cdots, x_n) \\ \boldsymbol{t} = (t_1, t_2, \cdots, t_n) &\in \mathbf{R}^n\end{aligned}$$

为随机变量$\boldsymbol{X}$的特征函数.

(3) 设X是非负值随机变量，其分布函数为$F(x)$，则称

$$\widetilde{F}(s) \xlongequal{\text{def}} E[\mathrm{e}^{-sX}] = \int_{0}^{+\infty} \mathrm{e}^{-sx}\,\mathrm{d}F(x)$$

为$F(x)$的 Laplace 变换，其中$s=a+\mathrm{j}b$，$a>0$.

2) 主要的性质和结论

(1) 随机变量X的特征函数$\varphi(t)$具有如下性质：

① $|\varphi(t)| \leqslant \varphi(0)=1$；

② $\overline{\varphi(t)}=\varphi(-t)$，其中$\overline{\varphi(t)}$表示$\varphi(t)$的共轭；

③ 设随机变量$Y=aX+b$，其中a，b是常数，则$\varphi_Y(t)=\mathrm{e}^{\mathrm{j}bt}\varphi_X(at)$；

④ $\varphi(t)$在$(-\infty, +\infty)$上一致连续；

⑤ 设随机变量X，Y相互独立，$Z=X+Y$，则$\varphi_Z(t)=\varphi_X(t)\varphi_Y(t)$；

⑥ $\varphi(t)$是非负定的，即对于任意的正整数n，任意的复数z_1，z_2，$\cdots$，z_n和任意的实数

t_1, t_2, …, t_n, 有

$$\sum_{l=1}^{n}\sum_{k=1}^{n}\varphi(t_l-t_k)z_l\bar{z}_k\geqslant 0$$

⑦ 设随机变量 X 的 n 阶原点矩存在，则 $\varphi(t)$ 存在 $k(k\leqslant n)$ 阶导数，且

$$\varphi^{(k)}(0)=\mathrm{j}^k EX^k,\ k\leqslant n$$

(2) 随机变量的分布函数和特征函数相互惟一确定，是一一对应的.

(3) 设 $\varphi(t)$ 满足 $\varphi(0)=1$，且在 $-\infty<t<+\infty$ 上是连续的复值函数，则 $\varphi(t)$ 是特征函数的充要条件为它是非负定的.

(4) n 维随机变量 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ 的特征函数 $\varphi(\boldsymbol{t})=\varphi(t_1, t_2, \cdots, t_n)$ 具有如下性质：

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

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

③ 设 $Y=a_1X_1+a_2X_2+\cdots+a_nX_n$，则 $\varphi_Y(t)=\varphi(a_1t, a_2t, \cdots, a_nt)$；

④ $\varphi(t_1, t_2, \cdots, t_n)$ 在 $\mathbf{R}^n$ 上一致连续；

⑤ $X_1, X_2, \cdots, X_n$ 相互独立的充要条件是 $\varphi(t_1, t_2, \cdots, t_n)=\varphi_{X_1}(t_1)\varphi_{X_2}(t_2)\cdots\varphi_{X_n}(t_n)$；

⑥ $k(1\leqslant k<n)$ 维随机变量 $(X_1, X_2, \cdots, X_k)$ 的特征函数为

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

⑦ 如果 $E[X_1^{k_1}X_2^{k_2}\cdots X_n^{k_n}]$ 存在，则

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

(4) Laplace 变换具有如下性质：

① 设 $\widetilde{F}(s)=\mathscr{L}(f(x))$，则 $\mathscr{L}(f'(x))=s\widetilde{F}(s)-f(0)$；

② 设 $\widetilde{F}(s)=\mathscr{L}(f(x))$，则 $\mathscr{L}\left(\int_0^x f(t)\,\mathrm{d}t\right)=\dfrac{\widetilde{F}(s)}{s}$；

③ 设 $\widetilde{F}(s)=\mathscr{L}(f(x))$，则 $\mathscr{L}(\mathrm{e}^{-ax}f(x))=\widetilde{F}(s+a)$；

④ 设 $\widetilde{F}(s)=\mathscr{L}(f(x))$，则 $\mathscr{L}(f(x-a)h(x-a))=\mathrm{e}^{-as}\widetilde{F}(s)$，其中

$$h(x)=\begin{cases}0, & x<0\\ 1, & x>0\end{cases}$$

⑤ 设 $\widetilde{F}(s)=\mathscr{L}(f(x))$，$\widetilde{G}(s)=\mathscr{L}(g(x))$，则

$$\mathscr{L}(f(x)*g(x))=\widetilde{F}(s)\widetilde{G}(s)$$

其中

$$f(x)*g(x)=\int_0^x f(t)g(x-t)\,\mathrm{d}t=\int_0^x f(x-t)g(t)\,\mathrm{d}t$$

⑥ 设 $\widetilde{F}(s)=\mathscr{L}(f(x))$，则 Laplace 反变换为

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{e}^{sx}\widetilde{F}(s)\,\mathrm{d}s,\ \mathrm{Re}s>0$$

5. n 维正态随机变量

1) 定义

设 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ 是 n 维随机变量，如果其联合分布函数为

$$F(\boldsymbol{x}) = F(x_1, x_2, \cdots, x_n)$$
$$= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} \frac{1}{(2\pi)^{n/2}|\boldsymbol{B}|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})\boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\right] \mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n$$

其中

$$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$$
$$\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_n) \xlongequal{\text{def}} (EX_1, EX_2, \cdots, EX_n)$$
$$\boldsymbol{B} \xlongequal{\text{def}} \begin{bmatrix} \operatorname{cov}(X_1, X_1) & \operatorname{cov}(X_1, X_2) & \cdots & \operatorname{cov}(X_1, X_n) \\ \operatorname{cov}(X_2, X_1) & \operatorname{cov}(X_2, X_2) & \cdots & \operatorname{cov}(X_2, X_n) \\ \vdots & \vdots & & \vdots \\ \operatorname{cov}(X_n, X_1) & \operatorname{cov}(X_n, X_2) & \cdots & \operatorname{cov}(X_n, X_n) \end{bmatrix}$$

则称 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$服从 $\boldsymbol{\mu}$ 为均值向量，$\boldsymbol{B}$ 为协方差矩阵的 n 维正态分布，记为 $\boldsymbol{X}\sim N(\boldsymbol{\mu}, \boldsymbol{B})$.

2）主要的性质和结论

（1）设 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)\sim N(\boldsymbol{\mu}, \boldsymbol{B})$，$l_1, l_2, \cdots, l_n$ 是常数，则

$$Y = \sum_{k=1}^{n} l_k X_k \sim N\left(\sum_{k=1}^{n} l_k\mu_k, \sum_{i=1}^{n}\sum_{k=1}^{n} l_i l_k \operatorname{cov}(X_i, X_k)\right)$$

其中

$$\mu_k = EX_k, \ k=1, 2, \cdots, n$$

（2）设 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)\sim N(\boldsymbol{\mu}, \boldsymbol{B})$，$m<n$，则 $\boldsymbol{X}$ 的 m 个分量构成的 m 维随机变量 $\widetilde{\boldsymbol{X}}=(X_1, X_2, \cdots, X_m)$服从 m 维正态分布 $N(\tilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{B}})$，其中

$$\tilde{\boldsymbol{\mu}} = (\mu_1, \mu_2, \cdots, \mu_m)$$
$$\widetilde{\boldsymbol{B}} = \begin{bmatrix} \operatorname{cov}(X_1, X_1) & \operatorname{cov}(X_1, X_2) & \cdots & \operatorname{cov}(X_1, X_m) \\ \operatorname{cov}(X_2, X_1) & \operatorname{cov}(X_2, X_2) & \cdots & \operatorname{cov}(X_2, X_m) \\ \vdots & \vdots & & \vdots \\ \operatorname{cov}(X_m, X_1) & \operatorname{cov}(X_m, X_2) & \cdots & \operatorname{cov}(X_m, X_m) \end{bmatrix}$$

（3）设 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)\sim N(\boldsymbol{\mu}, \boldsymbol{B})$，且 m 维随机变量 $\boldsymbol{Y}$ 是 $\boldsymbol{X}$ 的线性变换，即 $\boldsymbol{Y}=\boldsymbol{XC}$，其中 $\boldsymbol{C}$ 是 $n\times m$ 阶矩阵，则 $\boldsymbol{Y}$ 服从 m 维正态分布 $N(\boldsymbol{\mu C}, \boldsymbol{C}^{\mathrm{T}}\boldsymbol{BC})$.

（4）设 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)\sim N(\boldsymbol{\mu}, \boldsymbol{B})$，则 $X_1, X_2, \cdots, X_n$ 相互独立的充要条件是 $X_1, X_2, \cdots, X_n$ 两两不相关.

6. 条件数学期望

1）定义

（1）设(X, Y)是二维随机变量，对于任意固定正数 ε 及给定 y，$P(y-\varepsilon<Y\leqslant y+\varepsilon)>0$，若对于任意的实数 x，极限

$$\lim_{\varepsilon\to 0^+} P(X\leqslant x | y-\varepsilon<Y\leqslant y+\varepsilon) = \lim_{\varepsilon\to 0^+} \frac{P(X\leqslant x, y-\varepsilon<Y\leqslant y+\varepsilon)}{P(y-\varepsilon<Y\leqslant y+\varepsilon)}$$

存在，则称此极限为(X, Y)关于 X 在条件 $Y=y$ 下的条件分布函数，记为 $P(X\leqslant x|Y=y)$ 或 $F_{X|Y}(x|y)$. 类似地，有 $P(Y\leqslant y|X=x)$或 $F_{Y|X}(y|x)$以及 n 维随机变量 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$关于 X_i 在 $X_1=x_1, \cdots, X_{i-1}=x_{i-1}, X_{i+1}=x_{i+1}, \cdots, X_n=x_n$ 下的条件分布函数

$F_{X_i|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n}(x_i|x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)$.

(2) 设(X,Y)是二维随机变量，$F_{X|Y}(x|y)$是 X 的条件分布函数，则称

$$E(X|y)=\int_{-\infty}^{+\infty}x\,\mathrm{d}F_{X|Y}(x|y)$$

为 X 在条件$Y=y$下的条件数学期望. 由于$E(X|y)$是 y 的函数，因此$E(X|Y)$是随机变量 Y 的函数，称为 X 在条件 Y 下的条件数学期望. 类似地，有 Y 在条件 X 下的条件数学期望 $E(Y|X)$.

(3) 设$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$是 n 维随机变量，$F_{X_i|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n}(x_i|x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)$为 X_i 的条件分布函数，则称

$$\begin{aligned}&E(X_i|x_1,\cdots,x_{i-1},x_{i-1},\cdots,x_n)\\&\xlongequal{\text{def}}\int_{-\infty}^{+\infty}x\,\mathrm{d}F_{X_i|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n}(x|x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)\end{aligned}$$

为 X_i 在条件 $X_1=x_1,\cdots,X_{i-1}=x_{i-1},X_{i+1}=x_{i+1},\cdots,X_n=x_n$ 下的条件数学期望. 称 $E(X_i|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n)$ 为 X_i 在条件 $X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n$ 下的条件数学期望.

2) 主要的性质和结论

(1) $E(E(X_i|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n))=EX_i$.

(2) 设 $X_1,X_2,\cdots,X_n$ 相互独立，则 $E(X_i|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n)=EX_i$.

(3) 设$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$是 n 维随机变量，$g(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)$是连续函数，则

$$\begin{aligned}&E(X_ig(X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n)|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n)\\&=g(X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n)E(X_i|X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_n)\end{aligned}$$

(4) 设$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$是 n 维随机变量，$k<n-1$，则

$$E(E(X_n|X_1,\cdots,X_{n-1})|X_1,\cdots,X_k)=E(X_n|X_1,\cdots,X_k)$$

7. 随机变量序列的收敛性

1) 定义

(1) (**弱收敛**)对于分布函数列$\{F_n(x)\}$，如果存在一个非降函数 $F(x)$，使

$$\lim_{n\to\infty}F_n(x)=F(x)$$

在 $F(x)$的每一连续点上都成立，则称$\{F_n(x)\}$弱收敛于 $F(x)$，并记为 $F_n(x)\xrightarrow{W}F(x)$.

(2) (**依分布收敛**)设随机变量 X_n，X 的分布函数分别为 $F_n(x)$和 $F(x)$，如果 $F_n(x)\xrightarrow{W}F(x)$，则称$\{X_n\}$依分布收敛于 X，并记为 $X_n\xrightarrow{L}X$.

(3) (**依概率收敛**)如果对于任意 $\varepsilon>0$，$\lim\limits_{n\to\infty}P(|X_n-X|\geqslant\varepsilon)=0$，则称$\{X_n\}$依概率收敛于 X，并记为 $X_n\xrightarrow{P}X$ 或 $p\lim\limits_{n\to\infty}X_n=X$.

(4) (**r 阶收敛**)设对随机变量 X_n 及 X 有 $E|X_n|^r<+\infty$，$E|X|^r<+\infty$，其中 $r(r>0)$ 为常数，如果 $\lim\limits_{n\to\infty}E|X_n-X|^r=0$，则称$\{X_n\}$$r$ 阶收敛于 X，并记为 $X_n\xrightarrow{r}X$.

(5) (**几乎处处收敛**)如果 $P(\lim\limits_{n\to\infty}X_n=X)=1$，则称$\{X_n\}$几乎处处收敛于 X，又称$\{X_n\}$以概率 1 收敛于 X，并记为 $X_n\xrightarrow{\text{a.s.}}X$.

2）主要的性质和结论

（1）随机变量序列的收敛性之间的关系为：

几乎处处收敛⇒依概率收敛⇒依分布收敛；

r 阶收敛⇒依概率收敛⇒依分布收敛.

（2）（弱）大数定律有：

① **Bernoulli 大数定律**：设 n_A 表示在 n 重 Bernoulli 试验中事件 A 出现的次数，而 p 表示每次试验中事件 A 出现的概率，则$\forall\ \varepsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{n_A}{n}-p\right|\geqslant\varepsilon\right)=0$$

② **Chebyshev 大数定律**：设 $X_1, X_2, \cdots, X_n, \cdots$是两两不相关的随机变量序列，每一随机变量都有有限的方差，并且它们有公共的上界，即存在 $M>0$，$DX_k\leqslant M$，$k=1, 2, \cdots$，则$\forall\ \varepsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}EX_k\right|\geqslant\varepsilon\right)=0$$

③ **Khintchine 大数定律**：设 $X_1, X_2, \cdots, X_n, \cdots$是相互独立同分布的随机变量序列，且有有限的数学期望，$\mu=EX_k<+\infty$，$k=1, 2, \cdots$，则$\forall\ \varepsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\mu\right|\geqslant\varepsilon\right)=0$$

④ **Markov 大数定律**：设 $X_1, X_2, \cdots, X_n, \cdots$是随机变量序列，且

$$\lim_{n\to\infty}\frac{1}{n^2}D\left(\sum_{k=1}^{n}X_k\right)=0$$

则$\forall\ \varepsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}EX_k\right|\geqslant\varepsilon\right)=0$$

（3）强大数定律有：

① **Borel 大数定律**：设 n_A 表示在 n 重 Bernoulli 试验中事件 A 出现的次数，而 p 表示每次试验中事件 A 出现的概率，则

$$P\left(\lim_{n\to\infty}\frac{n_A}{n}=p\right)=1$$

② **Kolmogorov 大数定律**：设 $X_1, X_2, \cdots, X_n, \cdots$是相互独立的随机变量序列，且

$$\sum_{n=1}^{\infty}\frac{DX_n}{n^2}<+\infty$$

则

$$P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}(X_k-EX_k)=0\right)=1$$

③ **Kolmogorov 大数定理**：设 $X_1, X_2, \cdots, X_n, \cdots$ 是相互独立同分布的随机变量序列，则 $\frac{1}{n}\sum_{k=1}^{n}X_k\xrightarrow{\text{a.s.}}\mu$ 的充要条件是 EX_k 存在且等于 μ.

1.2 例题解析

例 1.1 在一次选举中，候选人 A 得到 n 张选票，而候选人 B 得到 m 张选票，其中

$n>m$. 假定选票的一切排列次序是等可能的，证明在计票过程中，A 的票数始终领先的概率为$\frac{n-m}{n+m}$.

证明 设 p_{nm}为所求的概率，由全概率公式，得

$$p_{nm} = P(A\text{ 始终领先 }|A\text{ 得到最后一票}) \cdot \frac{n}{n+m}$$
$$+ P(A\text{ 始终领先 }|B\text{ 得到最后一票}) \cdot \frac{m}{n+m}$$

容易看出，在 A 得到最后一票的条件下，A 的票数始终领先的概率与 A 得到 $n-1$ 票而 B 得到 m 票的情形要计算的概率是一样的. 在 B 得到最后一票的条件下，可得类似的结果. 从而

$$p_{nm} = \frac{n}{n+m}p_{n-1\,m} + \frac{m}{n+m}p_{n\,m-1}$$

用数学归纳法证明如下：

当 $n+m=1$ 时，结论显然成立，即 $p_{10}=1$.

假设当 $n+m=k$ 时，结论成立，当 $n+m=k+1$ 时，有

$$p_{nm} = \frac{n}{n+m} \cdot \frac{(n-1)-m}{(n-1)+m} + \frac{m}{n+m} \cdot \frac{n-(m-1)}{n+(m-1)} = \frac{n-m}{n+m}$$

因此，A 的票数始终领先的概率为$\frac{n-m}{n+m}$.

例 1.2 n 个人将自己的帽子放在一起，充分混合后每人随机地取出一顶帽子，试求恰有 k 个人选中自己帽子的概率.

解 设 E 表示“n 个人全没有选中自己的帽子”这一事件及 $p_n=P(E)$，M 表示“第一个人选中自己的帽子”，由全概率公式，得

$$p_n = P(E) = P(M)P(E|M) + P(\overline{M})P(E|\overline{M})$$

显然 $P(E|M)=0$，从而

$$p_n = \frac{n-1}{n}P(E|\overline{M})$$

由于 $P(E|\overline{M})$是 $n-1$ 个人从 $n-1$ 顶帽子中各取一顶都未选中自己帽子的概率，因而其中有一个人的帽子不在这 $n-1$ 顶帽子中，此事件可以分成互不相容的两种情形：一种是都未选中且其帽子被第一个选取的人选中的那个人未能选中第一个选取的人的帽子；另一种是都未选中且其帽子被第一个选取的人选中的那个人选中了第一个选取的人的帽子. 由于第二种情形的概率为$\frac{1}{n-1}p_{n-2}$，因此

$$P(E|\overline{M}) = p_{n-1} + \frac{1}{n-1}p_{n-2}$$

于是

$$p_n = \frac{n-1}{n}p_{n-1} + \frac{1}{n}p_{n-2}$$

或
$$p_n - p_{n-1} = -\frac{1}{n}(p_{n-1} - p_{n-2})$$

显然
$$p_1 = 0,\ p_2 = \frac{1}{2}$$

从而

$$p_3 - p_2 = -\frac{1}{3}(p_2 - p_1) = -\frac{1}{3!} \quad 或 \quad p_3 = \frac{1}{2!} - \frac{1}{3!}$$

$$p_4 - p_3 = -\frac{1}{4}(p_3 - p_2) = \frac{1}{4!} \quad 或 \quad p_4 = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}$$

一般地，有

$$p_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + (-1)^n \frac{1}{n!}$$

由于任意固定的 k 个人，只有他们选中自己帽子的概率为

$$\frac{1}{n} \cdot \frac{1}{n-1} \cdots \frac{1}{n-(k-1)} p_{n-k} = \frac{(n-k)!}{n!} p_{n-k}$$

其中 p_{n-k}表示其余 $n-k$ 个人从他们自己的那些帽子中选取但都未选中的概率.

又 k 个人的选择法有 C_n^k 种，所以恰有 k 个人选中自己帽子的概率为

$$C_n^k \frac{(n-k)!}{n!} p_{n-k} = \frac{1}{k!}\left[\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-k} \frac{1}{(n-k)!}\right]$$

当 n 充分大时，上式近似地等于$\frac{1}{k!}e^{-1}$，即当 n 充分大时，选中自己帽子的人数近似地服从参数为 1 的 Poisson 分布.

例 1.3 设有 n 个点排列在直线上，随机地选取一对相邻的点，且选取点对$(i, i+1)$的概率为$\frac{1}{n-1}$，$i=1, 2, \cdots, n-1$. 连续随机地取这样的点对，但如果取到的点对中有前面已取到过的点，这一对就去掉不算. 这样一直取到只留下孤立点为止，试求孤立点个数的平均数.

解 令

$$I_{in} = \begin{cases} 1，若 i 是孤立点 \\ 0，其它 \end{cases}$$

则 $\sum_{i=1}^{n} I_{in}$ 表示孤立点的个数，因此

$$E[孤立点的个数] = \sum_{i=1}^{n} EI_{in} = \sum_{i=1}^{n} p_{in}$$

其中 p_{in}表示有 n 个点时，i 是孤立点的概率，令

$$p_n = p_{nn} = p_{1n}$$

p_n 表示端点 n(或 1)成孤立点的概率. 为了导出 p_{in}的表达式，可以认为这 n 个点由两个独立的片断组成，即

$$1, 2, \cdots, i 与 i, i+1, \cdots, n$$

因为点 i 未取到当且仅当第一段的右端点和第二段的左端点都未取到，所以

$$p_{in} = p_i p_{n-i+1}$$

因此如果我们能计算端点未取到的概率，则 p_{in}可确定. 为导出 p_n 的表达式，以初始点对(比如说$(i, i+1)$)为条件，且注意这个点对将这些点断成两个独立的片断：$1, 2, \cdots, i-1$ 与 $i+2, \cdots, n$，于是，如果初始点对是$(i, i+1)$，那么端点 n 是孤立点；如果 $n-i-1$ 个点的集合的端点是孤立点，那么

$$p_n=\sum_{i=1}^{n-1}\frac{p_{n-i-1}}{n-1}=\frac{p_1+p_2+\cdots+p_{n-2}}{n-1}$$

或

$$(n-1)p_n=p_1+p_2+\cdots+p_{n-2}$$

用 $n-1$ 代替 n，得

$$(n-2)p_{n-1}=p_1+p_2+\cdots+p_{n-3}$$

将两式相减，得

$$(n-1)p_n-(n-2)p_{n-1}=p_{n-2}$$

或

$$p_n-p_{n-1}=-\frac{1}{n-1}(p_{n-1}-p_{n-2})$$

由 $p_0=1$，$p_2=0$，得

$$p_3-p_2=-\frac{1}{2}(p_2-p_1)=\frac{1}{2!}\quad 或\quad p_3=\frac{1}{2!}$$

$$p_4-p_3=-\frac{1}{3}(p_3-p_2)=-\frac{1}{3!}\quad 或\quad p_4=\frac{1}{2!}-\frac{1}{3!}$$

一般地

$$p_n=\frac{1}{2!}-\frac{1}{3!}+\cdots+(-1)^{n-1}\frac{1}{(n-1)!}=\sum_{k=0}^{n-1}\frac{(-1)^k}{k!},\ n\geqslant 2$$

于是

$$p_{in}=\begin{cases}\displaystyle\sum_{k=0}^{n-1}\frac{(-1)^k}{k!}, & i=1,\ n\\ 0, & i=2,\ n-1\\ \displaystyle\sum_{k=0}^{i-1}\frac{(-1)^k}{k!}\sum_{k=0}^{n-i}\frac{(-1)^k}{k!}, & 2<i<n-1\end{cases}$$

对大的 i 与 $n-i$，由上面可见 $p_{in}\approx e^{-2}$，且事实上能从上式证明，当 n 充分大时，$\sum_{i=1}^{n}p_{in}$（孤立点个数的期望）渐近地为 $(n+2)e^{-2}$，即

$$\sum_{i=1}^{n}p_{in}\approx(n+2)e^{-2}$$

例 1.4 设随机变量 X 的特征函数为 $\varphi(t)=\dfrac{1}{1+t^2}$，试求 X 的概率密度函数 $f(x)$.

解 由于

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-jtx}\varphi(t)\,dt=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-jtx}\frac{1}{1+t^2}\,dt$$

$$=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{\cos tx}{1+t^2}\,dt$$

所以 $f(x)$ 是偶函数，且 $\int_{-\infty}^{+\infty}\frac{\cos tx}{1+t^2}\,dt$ 绝对可积. 不妨设 $x>0$，则

$$f'(x)=-\frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{t\sin tx}{1+t^2}\,dt$$

利用著名的 Dirichlet 积分 $\int_{-\infty}^{+\infty}\frac{\sin tx}{t}\,\mathrm{d}t=\frac{\pi}{2}$，得

$$2\pi f'(x)=-\int_{-\infty}^{+\infty}\frac{t\sin tx}{1+t^2}\,\mathrm{d}t=\int_{-\infty}^{+\infty}\sin tx\left(\frac{-t}{1+t^2}+\frac{1}{t}\right)\mathrm{d}t-\frac{\pi}{2}$$

$$=\int_{-\infty}^{+\infty}\frac{\sin tx}{t(1+t^2)}\,\mathrm{d}t-\frac{\pi}{2}$$

两边再对 x 求导得

$$2\pi f''(x)=\int_{-\infty}^{+\infty}\frac{\cos tx}{1+t^2}\,\mathrm{d}t=2\pi f(x)$$

求解 $f''(x)=f(x)$得

$$f(x)=C_1\mathrm{e}^{x}+C_2\mathrm{e}^{-x}$$

由于 $f(x)$为概率密度函数，因此 $\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x=1$，从而当 $x\to+\infty$ 时，$f(x)\to0$，所以 $C_1=0$，再由 $f(x)$ 的偶函数性得 $f(x)=C_2\mathrm{e}^{-|x|}$，再利用$\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x=1$ 得 $C_2=\frac{1}{2}$，故 $f(x)=\frac{1}{2}\mathrm{e}^{-|x|}$.

例 1.5 设 X 服从几何分布，即 $P(X=k)=pq^{k-1}$，$k=1, 2, \cdots$，$q=1-p$，试求 X 的特征函数.

解

$$\varphi(t)=E[\mathrm{e}^{\mathrm{j}tX}]=\sum_{k=1}^{\infty}\mathrm{e}^{\mathrm{j}tk}pq^{k-1}=p\mathrm{e}^{\mathrm{j}t}\sum_{k=1}^{\infty}(q\mathrm{e}^{\mathrm{j}t})^{k-1}$$

$$=\frac{p\mathrm{e}^{\mathrm{j}t}}{1-q\mathrm{e}^{\mathrm{j}t}}$$

例 1.6 设 X 服从 $\chi^2(n)$分布，即其概率密度为

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

试求 X 的特征函数及 EX 和 DX，若 X_1，X_2 相互独立并且分别服从 $\chi^2(n_1)$和 $\chi^2(n_2)$分布，试求 X_1+X_2 的分布.

解 X 的特征函数为

$$\varphi(t)=E[\mathrm{e}^{\mathrm{j}tX}]=\int_0^{+\infty}\mathrm{e}^{\mathrm{j}tx}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}\,\mathrm{d}x$$

$$=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}x^{\frac{n}{2}-1}\mathrm{e}^{-\left(\frac{1}{2}-\mathrm{j}t\right)x}\,\mathrm{d}x$$

$$=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\cdot\frac{\Gamma\left(\frac{n}{2}\right)}{\left(\frac{1}{2}-\mathrm{j}t\right)^{\frac{n}{2}}}$$

$$=(1-2\mathrm{j}t)^{-\frac{n}{2}}$$