荣腾中 孙荣恒 刘朝林 编著

随机过程及其应用

（第二版）

清华大学出版社

荣腾中 孙荣恒 刘朝林 编著

随机过程及其应用

（第二版）

清华大学出版社

北 京

内容简介

本书从创新型研究型专业人才培养的教学需要出发，结合多年来的教学基础编纂而成.本书内容从概率论的条件分布、极限理论开始，介绍了随机过程的概念与分类、马尔可夫过程、泊松过程与更新过程、连续参数马尔可夫链、随机分析、平稳过程和鞅论初步及其应用等.

本书可供高等院校本科生或研究生作为随机过程的教材使用，也可供相关专业科研工作者及工程技术人员参考.

图书在版编目(CIP)数据

随机过程及其应用/荣腾中，孙荣恒，刘朝林编著. —2版. —北京：清华大学出版社，2017
ISBN 978-7-302-48405-9

Ⅰ. ①随…　Ⅱ. ①荣…　②孙…　③刘…　Ⅲ. ①随机过程　Ⅳ. ①O211.6

中国版本图书馆CIP数据核字(2017)第216414号

责任编辑：汪　操
封面设计：傅瑞学
责任校对：赵丽敏
责任印制：沈　露

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn, http://www.wqbook.com
地　　址：北京清华大学学研大厦A座　　邮　　编：100084
社 总 机：010-62770175　　邮　　购：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：三河市金元印装有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　印　张：16.5　　字　　数：397千字
版　　次：2004年2月第1版　　2017年8月第2版　　印　　次：2017年8月第1次印刷
定　　价：49.00元

产品编号：075340-01

前　言

随机过程是统计学的一门分支，是动态地研究随机现象的统计规律性的数学学科.经过半个多世纪的发展，随机过程理论已经成为一个十分活跃的学科领域.它已广泛应用于通信、雷达、控制、生物、社会科学，以及其他工程科学技术领域中.人们已经认识到，在现代科学技术飞速发展的过程中，学习和掌握随机过程的基本理论日渐成为一种需要.现实科研活动启示我们，只有熟悉和掌握随机过程的基本理论和基本分析方法，才能更好地学习现代科学技术，探索新的科学领域.

随机过程是工科研究生的一门重要的公共数学课程，也是理科高年级本科生的基础课程.通过这一课程的学习，掌握对一连续(离散)变化的随机现象进行过程分析的方法，能对长期发展后的情形进行预测.随机过程既有严谨的理论基础知识，又重视实际应用和概念的直观背景，同时强调推断的思想和方法.因此，为营造良好的学习氛围，提高教学质量，使更多的同学主动地参与到随机过程的学习中来，编者从改编新版的随机过程教材入手，通过逐步的教学反馈不断修订完善，希望最终能出版一本针对性较强的随机过程教材.本书便是从创新型研究型专业人才培养的教学需要出发，结合多年来的教学基础编纂而成.本书内容从概率论的条件分布、极限理论开始，介绍了随机过程的概念与分类、马尔可夫过程、泊松过程与更新过程、连续参数马尔可夫链、随机分析、平稳过程和鞅论初步及其应用等.

本书在内容编排上力求由浅入深，加强背景概念的阐述，以最易接受的方式介绍随机过程的基本理论及各类常用的随机过程.书中列出较多例题，都是在教学过程中经过精心挑选的，以便对基本理论加深理解.为了提高分析问题和解决问题的能力，做习题是必不可少的，为此每章后面配有大量习题，书末部分附有较为详细的习题解答.

感谢重庆大学中央高校基本科研业务费(NO.106112017CDJXY100010)对本书编纂工作的资助!

自1986年起，本书一直以讲义的形式在重庆大学使用，2000年后被重庆多所高校选为教材，2004年经清华大学出版社成书出版.本书是在第一版的基础上经多次修改而成，然本书的结构设计还不完善，需要在教学实践中作进一步调整.同时，由于编者水平有限，书中难免存在错误和不妥之处，恳请读者批评指正.

编　者

2017年5月

目 录

第 1 章 预备知识 …… 1

1.1 随机过程介绍 …… 1

1.2 概率论基本概念 …… 2

1.3 多维随机变量 …… 4

1.4 随机变量间的关系 …… 6

1.5 条件数学期望 …… 7

1.6 全概率公式 …… 10

1.7 条件方差 …… 13

1.8 特征函数 …… 14

1.9 概率母函数 …… 20

1.10 极限定理 …… 21

习题 1 …… 23

第 2 章 随机过程的主要类型 …… 25

2.1 随机过程的定义 …… 25

2.2 随机过程的数字特征 …… 26

2.3 独立过程 …… 29

2.4 独立增量过程 …… 30

2.5 平稳增量过程 …… 32

2.6 马尔可夫过程 …… 35

2.7 鞅 …… 39

2.8 高斯过程 …… 41

2.9 维纳过程 …… 44

2.10 泊松过程 …… 48

2.11 平稳过程 …… 51

习题 2 …… 53

第 3 章 离散参数马氏链 …… 55

3.1 离散参数齐次马氏链 …… 55

3.2 离散分支过程 …… 65

3.3 状态的分类 …… 68

3.4 极限定理 …… 79

3.5 极限分布与平稳分布 …… 84

3.6 例子与应用 …… 91
习题3 …… 100
第4章 泊松过程 …… 104
4.1 泊松过程的性质与应用 …… 104
4.2 其他类型的泊松过程 …… 112
4.3 更新过程 …… 123
习题4 …… 132
第5章 连续参数齐次马尔可夫链 …… 134
5.1 柯尔莫戈洛夫方程 …… 134
5.2 特殊类型马尔可夫链 …… 141
5.3 随机服务系统(排队论)简介 …… 152
习题5 …… 161
第6章 随机分析 …… 163
6.1 随机序列的均方极限 …… 163
6.2 均方连续与均方导数 …… 167
6.3 均方积分 …… 169
习题6 …… 172
第7章 平稳过程 …… 175
7.1 例子与性质 …… 175
7.2 遍历性定理 …… 182
7.3 相关函数的谱分解 …… 187
7.4 线性系统中的平稳过程 …… 198
习题7 …… 208
第8章 鞅论初步及其应用 …… 211
8.1 σ代数下的条件数学期望 …… 211
8.2 离散参数鞅 …… 215
8.3 停时与任意停止定理 …… 222
8.4 停时的应用 …… 229
8.5 鞅的收敛定理及其应用 …… 235
8.6 连续参数鞅及其应用 …… 242
习题8 …… 247
部分习题参考答案 …… 250
参考文献 …… 255

第 1 章　预 备 知 识

在《概率论》与《数理统计》的学习中，讨论过初等随机现象，即一个随机变量描述一个随机现象，或少量随机变量描述简单的随机现象. 这时候讨论的随机现象都是静态的，而在应用多个随机变量的时候，更是作了独立性的假设.

事实上，很多的随机现象不可能仅用有限的随机变量就能来描述它们的丰富内涵；静态的随机变量也不足以反映随机现象的发展、延续、变化过程. 正因为如此，所以引发了动态的、无穷的随机变量的讨论.

1.1　随机过程介绍

例 1.1.1　设某个季节的天气可以用随机变量 $X(i)=1,2,3$ 来表示晴、阴、雨三种天气，i 表示日子，问随机变量 $X(i),i=1,2,\cdots$ 是否独立？对于 $X(i)$ 是一个取值为“晴、阴、雨”三状态的离散型随机变量，对如下两种情况，$X(i+1)$ 的分布是一样的吗？

$X(i-1)$	$X(i)$	$X(i+1)$
晴	阴	?
雨	阴	?

由生活常识可知，$X(i+1)$ 的分布既与 $X(i)$ 有关，也应与 $X(i-1)$ 有关. 这组随机变量列是不独立的，变量个数也无穷多(可数个).

例 1.1.2　某一商品每天的销售量，记为 X，若要研究其每天的变化，则需要研究依赖于时间 t 的随机变量列 $X(t),t=1,2,\cdots$.

这组随机变量列是不独立的，变量个数也无穷多(可数个).

例 1.1.3　某交换机在单位时间内收到的呼叫次数 X 为随机变量，$X\sim P(\lambda)$. 现在考虑在正常工作条件下 t 时刻以前收到的呼叫次数，它是一个随机变量，又与 t 有关，记为 $\{X(t),t\in[0,24]\}$.

这组随机变量列是不独立的，变量个数也无穷多(不可数个).

由此看出，要研究随机现象的发生发展过程，必须研究这一类与时间有关的随机变量簇.

(1) 对特定的时间 t_0，$X(t_0)$ 为一普通的随机变量；

(2) 对不同的 t_1,t_2，$X(t_1)$ 与 $X(t_2)$ 有不同的统计规律；

(3) 由于时间参数反映了随机现象发展变化的全过程，所以这类具有时间参数的随机变量簇称为随机过程——Stochastic Processes.

随机过程是指一簇随机变量，对随机过程的统计分析称为随机过程论，它是随机数学中的一个重要分支，产生于 20 世纪的初期，其研究对象与概率论一样是随机现象，而它特别研

究的是随“时间”变化的“动态”的随机现象,因此随机过程与概率论的关系类似于物理学中动力学与静力学的关系.

再来看几个例子,考虑是否独立,是否与时间参数有关?例如:

(1) 连续抛一枚骰子所得到的最大点数 $X(t),t=1,2,\cdots$;

(2) 某支股票的每日最高价 $Y(t),t=1,2,\cdots$;

(3) 布朗运动中一花粉在 t 时刻的坐标 $X(t),t\in[0,+\infty)$,其中 $X(t)=(x(t),y(t))$,这时称 $X(t)$ 为多维随机过程,或随机场;

(4) 生物种群模型:$X(t+1)=\sum\limits_{i=1}^{X(t)}Z_i^{(t)},t\in\mathbf{N}$,其中 $X(t)$ 表示 t 代个体数,$Z_i^{(t)}$ 是 t 代 i 个体的后代数.

总的来说,只要和时间有关的随机现象,都可以用随机过程来描述.随机过程是以概率论为理论基础,研究随机现象的动态特征(发生、发展过程)的一门数学分支.目前随机过程内容已经十分丰富,应用非常广泛.特别是近40年来,随着物理学、生物学、自动控制、无线电通信及管理科学等方面的需求的提出与解决,使随机过程逐步形成为一门独立的分支学科,在自然科学、工程技术及社会科学中日益呈现出广泛的应用前景和蓬勃的发展趋势.

1.2 概率论基本概念

本书的读者应具有初等的概率论知识,为了以后各章能顺利进行,下面先对本书有关的概率论基本内容作简要的回顾.

1. 概率论的基本概念

称随机试验 E 的最简单不能再分的每个结果为 E 的样本点,记为 ω 或 e.由所有样本点组成的集合 Ω 称为 E 的样本空间或必然事件.称不含样本点的空集 $\varnothing$ 为不可能事件.如果 Ω 中某些子集组成的集类 $\mathscr{F}$ 满足下列三个条件:

(1) $\Omega\in\mathscr{F}$;

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

(3) 如果 $A_i\in\mathscr{F},i=1,2,\cdots$,则 $\bigcup\limits_{i=1}^{\infty}A_i\in\mathscr{F}$,

则称 $\mathscr{F}$ 为 E 的事件域.称且仅称 $\mathscr{F}$ 中的元素为随机事件,简称为事件.

如果定义 $\mathscr{F}$ 上的实值集合函数 P 满足下列三个条件:

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

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

(3) 设 $A_i\in\mathscr{F},i=1,2,\cdots$,且当 $i\neq j$ 时,$A_iA_j=\varnothing$,则

$$P\Big(\bigcup_{i=1}^{\infty}A_i\Big)=\sum_{i=1}^{\infty}P(A_i),$$

则称 P 为概率测度,简称为概率.称 $(\Omega,\mathscr{F},P)$ 为概率空间.

如果定义于样本空间 Ω 上的单值实函数 ξ,对任意实数 x,$\{\omega:\xi(\omega)<x\}$(简记为 $\{\xi<x\}$)均为事件,即 $\{\xi<x\}\in\mathscr{F}$,则称 ξ 为随机变量.称概率

$$F_\xi(x) \triangleq P\{\xi < x\}, \quad x \in \mathbf{R}$$

为 ξ 的分布函数. 分布函数 $F_\xi(x)$ 有下列三个基本性质:

(1) 如果 $a<b$,则 $F_\xi(a) \leqslant F_\xi(b)$;

(2) $F_\xi(-\infty)=0, F_\xi(+\infty)=1$,其中 $F_\xi(-\infty)=\lim\limits_{x\to-\infty}F_\xi(x), F_\xi(+\infty)=\lim\limits_{x\to+\infty}F_\xi(x)$;

(3) $F_\xi(x-0)=F_\xi(x)$.

如果随机变量 ξ 只能取可数个不同的实数值,则称 ξ 为离散型随机变量. 如果存在非负函数 $f_\xi(x)$,使得对任意 $x\in\mathbf{R}$,有 $F_\xi(x)=\int_{-\infty}^{x} f_\xi(t)\mathrm{d}t$,则称 ξ 为连续型的随机变量,称 $f_\xi(x)$ 为 ξ 的密度函数.

2. 随机变量函数的分布

若 $\eta=g(\xi)$,如果 ξ 为离散型随机变量,可用列表法、公式法来求解 η 的离散型分布.

如果 ξ 为连续型,则分布函数法是最基本的方法.

$$\begin{cases} g\text{ 单调时}, & f_\eta(y)=f[g^{-1}(y)]\,|[g^{-1}(y)]'|, \\ g\text{ 不单调时}, & F_\eta(y)=P\{g(\xi)<y\}. \end{cases}$$

3. 数学期望

对离散型随机变量: 若有 $\sum\limits_{i\in S}|x_i|P\{\xi=i\}<+\infty$,则

$$E(\xi)=\sum_{i\in S}x_iP\{\xi=i\}.$$

对连续型随机变量: 若有 $\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x<+\infty$,则

$$E(\xi)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x.$$

数学期望的线性性质:

(1) $E(a\xi+b)=aE(\xi)+b$;

(2) $E(\xi+\eta)=E(\xi)+E(\eta)$.

4. 方差

对离散型随机变量: 若有 $\sum\limits_{i\in S}x_i^2P\{\xi=x_i\}<+\infty$,则

$$D(\xi)=\sum_{i\in S}(x_i-E(\xi))^2P\{\xi=x_i\}.$$

对连续型随机变量: 若有 $\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x<+\infty$,则

$$D(\xi)=\int_{-\infty}^{+\infty}(x-E(\xi))^2f(x)\mathrm{d}x.$$

方差的性质:

(1) $D(\xi)=E(\xi^2)-E^2(\xi)$;

(2) $D(a\xi+b)=a^2D(\xi)$.

5. k 阶矩

若有 $\int_{-\infty}^{+\infty}|x^k|\,\mathrm{d}F(x)<+\infty$,则

$$E(\xi^k)=\int_{-\infty}^{+\infty}x^k\mathrm{d}F(x).$$

6. 随机变量函数 $g(\xi)$ 的数学期望

如果存在,则

$$E(g(\xi))=\int_{-\infty}^{+\infty}g(x)\mathrm{d}F(x).$$

特别地,$g(\xi)=(\xi-E(\xi))^k$ 时,称

$$E(g(\xi))=\int_{-\infty}^{+\infty}(x-E(\xi))^k\mathrm{d}F(x)$$

为随机变量 ξ 的 k 阶中心矩. 其中一阶中心矩为 0,二阶中心矩为方差.

注意:数字特征是一个数值,$E(\xi)$ 与 $D(\xi)$ 都不再是随机变量了.

7. 数学期望的应用

例 1.2.1 按季节出售某种商品,每售 1kg 获利 5 元. 如果季节末有存货未售完,则每 1kg 净亏损 2 元. 设商品在季节内商品的销售量为 ξkg,$\xi\sim U(2000,5000)$,为使平均利润最大,问商店应进多少货?

解 设商店应进 $a$kg,由题知,$a\in(2000,5000)$,导出获利函数:

$$R(a,\xi)=\begin{cases}5\xi-2(a-\xi), & \xi\leqslant a,\\ 5a, & \xi>a.\end{cases}$$

$R(a,\xi)$ 是 ξ 的函数,仍为随机变量,则

$$\begin{aligned}E[R(a,\xi)]&=\int_{2000}^{5000}R(a,x)f_\xi(x)\mathrm{d}x\\ &=\int_{2000}^{a}(5x-2(a-x))\frac{1}{3000}\mathrm{d}x+\int_{a}^{5000}5a\frac{1}{3000}\mathrm{d}x\\ &=r(a),\end{aligned}$$

则最优的进货量 $a_0=\max\{r(a),a\in(2000,5000)\}$. 由 $r'(a)=0$,得 $a_0\approx4143$kg.

1.3 多维随机变量

在研究了概率论的基本概念之后,下面进一步讨论多维随机变量的分布描述,多维随机变量之间的关系,以及多维和一维之间的关系.

1. 二维联合分布 (ξ,η)

$$F_{\xi,\eta}(x,y)=P\{\xi<x,\eta<y\},\quad (x,y)\in\mathbf{R}^2.$$

二维联合分布函数的性质:

(1) 当 $x_1<x_2$ 时,$F(x_1,y)\leqslant F(x_2,y)$;

(2) 当 $y_1<y_2$ 时,$F(x,y_1)\leqslant F(x,y_2)$;

(3) $F(-\infty,-\infty)=0,F(-\infty,y)=0$,
$F(x,-\infty)=0,F(+\infty,+\infty)=1$;

(4) $F(x-0,y)=F(x,y),F(x,y-0)=F(x,y)$;

(5) 平行四边形法则:

$$P\{x_1\leqslant\xi<x_2,y_1\leqslant\eta<y_2\}=F(x_2,y_2)-F(x_1,y_2)-F(x_2,y_1)+F(x_1,y_1).$$

2. 二维离散型随机变量

$$P\{\xi=x_i,\eta=y_j\}=p_{ij},\quad \sum_i\sum_j p_{ij}=1.$$

称 p_{ij} 为二维联合分布律.

3. 二维连续型随机变量

$$F_{\xi,\eta}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}f(x,y)\mathrm{d}x\mathrm{d}y,\quad f(x,y)\geqslant 0.$$

称 $f_{\xi,\eta}(x,y)$ 为二维联合密度函数.

4. 边缘分布函数

$$F(x,+\infty)=P\{\xi<x,\eta<+\infty\}=P\{\xi<x\}=F_\xi(x).$$

5. 离散型边缘分布律

$\xi \backslash \eta$	y_1	y_2	$\cdots$	y_n	$\cdots$	
x_1	p_{11}	p_{12}	$\cdots$	p_{1n}	$\cdots$	$p_{1.}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2n}	$\cdots$	$p_{2.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nn}	$\cdots$	$p_{n.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
	$p_{.1}$	$p_{.2}$	$\cdots$	$p_{.n}$	$\cdots$	1

$$P\{\xi=x_i\}=\sum_j P\{\xi=x_i,\eta=y_j\}=\sum_j p_{ij}=p_{i.},$$
$$P\{\eta=y_j\}=\sum_i P\{\xi=x_i,\eta=y_j\}=\sum_i p_{ij}=p_{.j}.$$

6. 连续型边缘密度函数

(1) 由边缘分布求导:

$$f_\xi(x)=\frac{\partial F_\xi(x)}{\partial x}=\frac{\partial F_{\xi,\eta}(x,+\infty)}{\partial x}.$$

(2) 由联合密度求积分:

$$f_\xi(x)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y.$$

7. n 维联合分布$(\xi_1,\xi_2,\cdots,\xi_n)$

$$F_{(\xi_1,\xi_2,\cdots,\xi_n)}(x_1,x_2,\cdots,x_n)=P\{\xi_1<x_1,\xi_2<x_2,\cdots,\xi_n<x_n\},$$

其中$(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$.

1.4 随机变量间的关系

1. 随机变量的独立关系

两个事件 A,B 独立有 $P(AB)=P(A)P(B)$. 两个随机变量 (ξ,η) 独立有

$$P\{\xi<x,\eta<y\}=P\{\xi<x\}P\{\eta<y\},$$

$$F_{\xi,\eta}(x,y)=F_\xi(x)F_\eta(y).$$

具体来讲,(ξ,η) 同为离散型时,有

$$P\{\xi=x_i,\eta=y_j\}=P\{\xi=x_i\}P\{\eta=y_j\}.$$

(ξ,η) 同为连续型时,有

$$f_{\xi,\eta}(x,y)=f_\xi(x)f_\eta(y).$$

随机变量的独立是明确的,而相关关系则有相关程度的差异.

2. 随机变量的相关关系

(ξ,η) 的协方差:

$$\mathrm{Cov}(\xi,\eta)=E[(\xi-E(\xi))(\eta-E(\eta))]=E(\xi\eta)-E(\xi)E(\eta).$$

若两随机变量 (ξ,η) 相互独立,则 $E(\xi\eta)=E(\xi)E(\eta)$,即 $\mathrm{Cov}(\xi,\eta)=0$.

方差计算公式:

$$D(a\xi\pm b\eta)=a^2D(\xi)+b^2D(\eta)\pm 2ab\mathrm{Cov}(\xi,\eta).$$

相关系数:

$$\rho=\frac{\mathrm{Cov}(\xi,\eta)}{\sqrt{D(\xi)D(\eta)}},\quad D(\xi)D(\eta)>0.$$

例 1.4.1 试证明: $|\rho|\leqslant 1$.

证明 构造性证明,考虑 $t\in\mathbf{R}$ 的二次函数:

$$\begin{aligned}&E[t(\xi-E(\xi))+(\eta-E(\eta))]^2\geqslant 0\\ \Rightarrow\ &E[\xi-E(\xi)]^2t^2+E[\eta-E(\eta)]^2+2E[(\xi-E(\xi))(\eta-E(\eta))]t\geqslant 0\\ \Rightarrow\ &D(\xi)t^2+2\mathrm{Cov}(\xi,\eta)t+D(\eta)\geqslant 0\end{aligned}$$

由 $\Delta\leqslant 0$,有

$$\mathrm{Cov}^2(\xi,\eta)-D(\xi)D(\eta)\leqslant 0,$$

即 $|\rho|\leqslant 1$. □

这就是 Schwarz 不等式:

$$[E(xy)]^2\leqslant E(x^2)E(y^2).$$

3. 关于 ρ 取值的说明

$|\rho|=0$ 时,ξ 与 η 不相关;

$|\rho|\neq 0$ 时,ξ 与 η 相关;

$|\rho|$ 越接近于 1,ξ 与 η 线性相关程度就越高.

4. 独立和相关

(1) 独立则不相关：

$$E(\xi\eta)=E(\xi)E(\eta)\Rightarrow \mathrm{Cov}(\xi,\eta)=0\Rightarrow \rho=0;$$

(2) 不相关不一定独立；

(3) 正态型随机变量不相关与独立性等价.

例 1.4.2(不相关不独立反例) 设 $\xi\sim N(0,1)$，考虑随机变量 ξ 和 $\eta=\xi^2$.

$$\mathrm{Cov}(\xi,\eta)=E(\xi\eta)-E(\xi)E(\eta)=E(\xi^3)-E(\xi)E(\xi^2)=0,$$

所以 ξ 与 η 不相关，$P\{\xi<-1,\eta<1\}=P\{\xi<-1,\xi^2<1\}=P\{\xi<-1,-1<\xi<1\}=P\{\varnothing\}=0$.
另一方面

$$P\{\xi<-1\}P\{\eta<1\}>0,$$

则

$$P\{\xi<-1,\eta<1\}\neq P\{\xi<-1\}P\{\eta<1\},$$

所以随机变量 ξ 和 $\eta=\xi^2$ 虽然不相关，但不是独立的.

1.5 条件数学期望

为了研究复杂事件，而引入"条件概率"来描述某种条件下随机事件的发生概率：

$$P(A|B)=\frac{P(AB)}{P(B)},\quad P(B)>0.$$

如果 A,B 独立，则 $P(A|B)=P(A)$.

由条件概率的定义，若 ξ,η 是二维概率空间中的随机变量，由条件概率知，在 $\eta=y$ 的条件下，ξ 有条件分布 $F_{\xi|\eta}(x|y)$：

$$F_{\xi|\eta}(x\mid y)=\begin{cases}\dfrac{P\{\xi<x,\eta=y_j\}}{P\{\eta=y_j\}}, & \text{离散型.}\\[2ex] \displaystyle\int_{-\infty}^{x}\frac{f_{\xi,\eta}(x,y)}{f_\eta(y)}\mathrm{d}x, & \text{连续型.}\end{cases}$$

在离散型时，$P\{\eta=y_j\}>0$，有

$$P\{\xi=x_i|\eta=y_j\}=\frac{P\{\xi=x_i,\eta=y_j\}}{P\{\eta=y_j\}}.$$

条件分布律为

$$P\{\xi=x_i|\eta=y_j\}=\frac{p_{ij}}{p_{\cdot j}},\quad p_{\cdot j}>0.$$

在连续型时，$f_\eta(y)$ 为 η 的边缘密度函数，

$$f_\eta(y)=\int_{-\infty}^{+\infty}f_{\xi,\eta}(x,y)\mathrm{d}x.$$

条件密度为

$$f_{\xi|\eta}(x|y)=\frac{f_{\xi,\eta}(x,y)}{f_\eta(y)}.$$

1.5.1 条件期望的定义

条件分布是一个完整的概率分布,可以由此导出相应的数学期望,称为条件数学期望.

定义 1.5.1 如果 $\int_{-\infty}^{+\infty}|x|\,\mathrm{d}F_{\xi|\eta}(x\mid y)<+\infty$,则在 $\eta=y$ 的条件下,ξ 的条件数学期望为

$$E(\xi\mid\eta=y)=\int_{-\infty}^{+\infty}x\,\mathrm{d}F_{\xi|\eta}(x\mid y).$$

离散型:$E(\xi\mid\eta=y_j)=\sum\limits_{i\in S}x_i\dfrac{p_{ij}}{p_{\cdot j}}$.

连续型:$E(\xi\mid\eta=y)=\int_{-\infty}^{+\infty}xf_{\xi|\eta}(x\mid y)\mathrm{d}x$.

例 1.5.1 设 $\xi\sim P(\lambda_1)$,$\eta\sim P(\lambda_2)$,且 ξ 与 η 相互独立,求在 $\xi+\eta=k$(k 为非负整数)下 ξ 的条件数学期望 $E(\xi|\xi+\eta=k)$.

解 因为 ξ 与 η 相互独立,由泊松分布的和不变性,得

$$\xi+\eta\sim P(\lambda_1+\lambda_2).$$

从而

$$P\{\xi=i|\xi+\eta=k\}=\frac{P\{\xi=i,\xi+\eta=k\}}{P\{\xi+\eta=k\}}=\frac{P\{\xi=i,i+\eta=k\}}{P\{\xi+\eta=k\}}=\frac{P\{\xi=i\}P\{\eta=k-i\}}{P\{\xi+\eta=k\}}$$

$$=\frac{\mathrm{e}^{-\lambda_1}\dfrac{\lambda_1^i}{i!}\mathrm{e}^{-\lambda_2}\dfrac{\lambda_2^{k-i}}{(k-i)!}}{\mathrm{e}^{-(\lambda_1+\lambda_2)}\dfrac{(\lambda_1+\lambda_2)^k}{k!}}=\mathrm{C}_k^i\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^i\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{k-i}.$$

由 $i=0,1,\cdots,k$ 可知,当 $\xi+\eta=k$ 时,$\xi\sim B\left(k,\dfrac{\lambda_1}{\lambda_1+\lambda_2}\right)$,

$$E(\xi|\xi+\eta=k)=\frac{k\lambda_1}{\lambda_1+\lambda_2}.$$

由于数学期望 $E(\xi)$ 是一个数,那么条件期望 $E(\xi|\eta)$ 是否还是一个数呢?实际上 $E(\xi|\eta)$ 是关于 η 的函数.

1.5.2 条件期望的性质

性质 1 若随机变量 ξ 与 η 的期望均存在,则

$$E(\xi)=E[E(\xi|\eta)].\tag{1.5.1}$$

说明:数学期望 $E(\xi)$ 是一个实值,条件期望 $E(\xi|\eta)$ 是关于 η 的函数. 第二次期望是针对 η 来求的.

证明 仅考虑连续型随机变量的情形.

$$E(\xi\mid\eta=y)=\int_{-\infty}^{+\infty}xf_{\xi|\eta}(x\mid y)\mathrm{d}x=\int_{-\infty}^{+\infty}x\frac{f_{\xi,\eta}(x,y)}{f_\eta(y)}\mathrm{d}x=r(y),\tag{1.5.2}$$

$$E[E(\xi\mid\eta)]=\int_{-\infty}^{+\infty}r(y)f_\eta(y)\mathrm{d}y=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x\frac{f_{\xi,\eta}(x,y)}{f_\eta(y)}f_\eta(y)\mathrm{d}x\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf_{\xi,\eta}(x,y)\mathrm{d}x\mathrm{d}y=\int_{-\infty}^{+\infty}x\left[\int_{-\infty}^{+\infty}f_{\xi,\eta}(x,y)\mathrm{d}y\right]\mathrm{d}x$$

$$=\int_{-\infty}^{+\infty}xf_\xi(x)\mathrm{d}x=E(\xi).$$

□

由条件期望公式与式(1.5.2),得

$$E(\xi)=\int_{-\infty}^{+\infty}E(\xi\mid\eta=y)\mathrm{d}F_{\eta}(y).$$

此式也称为全期望公式,指求一个随机变量的期望有困难时,引入样本空间的一个剖分 η,先求在 $\eta=y$ 的条件下 ξ 的期望,再对这些条件期望进行加权和(积分):

$$E(\xi)=\begin{cases}\sum\limits_{j}E(\xi\mid\eta=y_j)P\{\eta=y_j\}, & \text{离散型},\\ \int_{-\infty}^{+\infty}E(\xi\mid\eta=y)f_{\eta}(y)\mathrm{d}y, & \text{连续型}.\end{cases}$$

例 1.5.2(矿工脱险) 一矿工在三个门的矿井中迷路,第一个门走 3 小时通向出口;第二个门走 7 小时回原地;第三个门走 5 小时回原地;如果他等可能的选择其中一个门,问他平均要花多少时间才能出井口?

解 设两个随机变量:ξ 表示走出井口所花的时间,η 表示选择的那个门.

由题知 η 的分布律:

$$P\{\eta=j\}=\frac{1}{3},\quad j=1,2,3.$$

$$\begin{aligned}E(\xi)&=E[E(\xi\mid\eta)]=\sum_{j=1}^{3}E(\xi\mid\eta=j)P\{\eta=j\}\\&=3\times\frac{1}{3}+[7+E(\xi)]\times\frac{1}{3}+[5+E(\xi)]\times\frac{1}{3},\end{aligned}$$

即

$$E(\xi)=(15+2E(\xi))\times\frac{1}{3}. \tag{1.5.3}$$

解方程(1.5.3),得 $E(\xi)=15$.

通过解方程来求数学期望的另类解法,这也体现了一个过程的概念.

性质 2 如果 ξ 与 η 独立,则 $E(\xi|\eta)=E(\xi)$.

证明 由 ξ 与 η 独立,得

$$f_{\xi,\eta}(x,y)=f_{\xi}(x)f_{\eta}(y),$$

$$f_{\xi|\eta}(x|y)=\frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}=f_{\xi}(x),$$

$$E(\xi\mid\eta=y)=\int_{-\infty}^{+\infty}xf_{\xi|\eta}(x\mid y)\mathrm{d}x=\int_{-\infty}^{+\infty}xf_{\xi}(x)\mathrm{d}x=E(\xi).\qquad\square$$

由于常数 c 与任意随机变量都是独立的,由性质 2 有

$$E(c|\eta)=c.$$

性质 3 $E(g(\eta)|\eta)=g(\eta)$.

证明
$$\begin{aligned}E(g(\eta)\mid\eta=y)&=\int_{-\infty}^{+\infty}g(y)f_{\xi|\eta}(x\mid y)\mathrm{d}x=g(y)\int_{-\infty}^{+\infty}\frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}\mathrm{d}x\\&=g(y)\frac{1}{f_{\eta}(y)}\int_{-\infty}^{+\infty}f_{\xi,\eta}(x,y)\,\mathrm{d}x=g(y)\frac{1}{f_{\eta}(y)}f_{\eta}(y)\\&=g(y).\end{aligned}\qquad\square$$

性质 4 $E(g(\eta)\xi|\eta)=g(\eta)E(\xi|\eta)$.

证明
$$E(g(\eta)\xi\mid\eta=y)=\int_{-\infty}^{+\infty}g(y)xf_{\xi|\eta}(x\mid y)\mathrm{d}x=g(y)\int_{-\infty}^{+\infty}x\frac{f_{\xi,\eta}(x,y)}{f_{\eta}(y)}\mathrm{d}x$$

$$= g(y)E(\xi \mid \eta = y).$$ □

性质 5 $E(g(\eta)\xi)=E[g(\eta)E(\xi|\eta)]$.

证明 由全期望公式有

$$E(g(\eta)\xi)=E[E(g(\eta)\xi|\eta)].$$

由性质 4 有

$$E(g(\eta)\xi)=E[g(\eta)E(\xi|\eta)] \tag{1.5.4}$$

□

性质 6 条件期望的线性性质:

$$E(a\xi+b\eta|\zeta)=aE(\xi|\zeta)+bE(\eta|\zeta).$$

证明 这要用到三维联合密度 $f_{\xi,\eta,\zeta}(x,y,z)$. 当 $\zeta=z$ 时,(ξ,η) 的条件联合密度为

$$f_{\xi,\eta|\zeta}(x,y|\zeta=z)=\frac{f_{\xi,\eta,\zeta}(x,y,z)}{f_\zeta(z)},$$

$$\begin{aligned}E(a\xi+b\eta \mid \zeta) &= \iint (ax+by)f_{\xi,\eta|\zeta}(x,y \mid z)\mathrm{d}x\mathrm{d}y \\ &= \iint ax f_{\xi,\eta|\zeta}(x,y \mid z)\mathrm{d}x\mathrm{d}y + \iint by f_{\xi,\eta|\zeta}(x,y \mid z)\mathrm{d}x\mathrm{d}y.\end{aligned} \tag{1.5.5}$$

因为

$$\begin{aligned}&\iint ax f_{\xi,\eta|\zeta}(x,y \mid z)\mathrm{d}x\mathrm{d}y = a\int_{-\infty}^{+\infty} x\left[\int_{-\infty}^{+\infty} f_{\xi,\eta|\zeta}(x,y \mid z)\mathrm{d}y\right]\mathrm{d}x \\ &= a\int_{-\infty}^{+\infty} x\left[\int_{-\infty}^{+\infty} \frac{f_{\xi,\eta,\zeta}(x,y,z)}{f_\zeta(z)}\mathrm{d}y\right]\mathrm{d}x = a\int_{-\infty}^{+\infty} x\left[\frac{f_{\xi,\zeta}(x,z)}{f_\zeta(z)}\right]\mathrm{d}x \\ &= a\int_{-\infty}^{+\infty} x f_{\xi|\zeta}(x \mid z)\mathrm{d}x = aE(\xi \mid \zeta),\end{aligned}$$

同理

$$\iint by f_{\xi,\eta|\zeta}(x,y \mid z)\mathrm{d}x\mathrm{d}y = bE(\eta \mid \zeta),$$

则由式(1.5.5),得

$$E(a\xi+b\eta|\zeta)=aE(\xi|\zeta)+bE(\eta|\zeta).$$ □

性质 7 $E[\xi-E(\xi|\eta)]^2 \leqslant E[\xi-g(\eta)]^2$.

说明:类似于一维分布里的 $E(\xi-E(\xi))^2 \leqslant E(\xi-x)^2, x\in \mathbf{R}$,与 ξ 均方差最小的 η 的函数是 $E(\xi|\eta)$.

证明 由全期望公式:

$$E[\xi-g(\eta)]^2=E\{E[[\xi-g(\eta)]^2|\eta]\},$$

$$E[[\xi-g(\eta)]^2 \mid \eta] = \int_{-\infty}^{+\infty} (x-g(y))^2 f_{\xi|\eta}(x \mid y)\mathrm{d}x.$$

由于条件分布是一个完整的分布,所以由一维分布的性质,可知当 $g(y)=E(\xi|\eta=y)$ 时,上式达到最小,进而对 η 求期望后,$E[\xi-E(\xi|\eta)]^2$ 达到最小. □

1.6 全概率公式

由全期望公式 $E(\xi)=E[E(\xi|\eta)]$,很容易推导出全概率公式.

设 A 为样本空间中一事件(见图 1.1),定义示性函数为

$$I_A(\omega)=\begin{cases}1, & \omega\in A,\\ 0, & \omega\notin A.\end{cases}$$

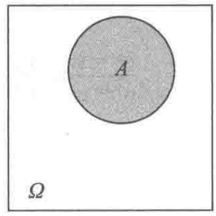

图 1.1

可知 $I_A(\omega)$ 为 $\Omega\to\mathbf{R}$ 的函数，是一个随机变量. 其分布为

$$P\{I_A(\omega)=1\}=P(A),\quad P\{I_A(\omega)=0\}=1-P(A),$$

于是

$$E[I_A(\omega)]=1\cdot P(A)=P(A).$$

另一方面，由全期望公式：

$$P(A)=E[I_A(\omega)]=E[E[I_A(\omega)\mid\eta]]=\int_{-\infty}^{+\infty}E[I_A(\omega)\mid\eta=y]\mathrm{d}F_\eta(y)$$
$$=\int_{-\infty}^{+\infty}[1\cdot P\{I_A(\omega)=1\mid\eta=y\}+0]\mathrm{d}F_\eta(y)=\int_{-\infty}^{+\infty}P\{A\mid\eta=y\}\mathrm{d}F_\eta(y),$$

这就是随机变量意义下的全概率公式.

离散型和连续型分别为

$$P(A)=\begin{cases}\sum\limits_{j\in S}P\{A\mid\eta=y_j\}P\{\eta=y_j\},\\ \int_{-\infty}^{+\infty}P\{A\mid\eta=y\}f_\eta(y)\mathrm{d}y.\end{cases}$$

例 1.6.1 设 ξ 与 η 独立且分布函数分别为 $F_\xi(x)$ 和 $F_\eta(y)$，求 $\xi+\eta$ 的分布函数.

解 由全概率公式：

$$F_{\xi+\eta}(a)=P\{\xi+\eta<a\}=\int_{-\infty}^{+\infty}P\{\xi+\eta<a\mid\eta=y\}\mathrm{d}F_\eta(y)$$
$$=\int_{-\infty}^{+\infty}P\{\xi<a-y\mid\eta=y\}\mathrm{d}F_\eta(y)=\int_{-\infty}^{+\infty}F_\xi(a-y)\mathrm{d}F_\eta(y).$$

例 1.6.2 设 ξ 与 η 为相互独立的随机变量，且分别服从参数 λ_1 与 λ_2 的指数分布，求 $P\{\xi<\eta\}$.

分析 参数为 λ_1 的指数分布，有

$$f_\xi(x)=\lambda_1\mathrm{e}^{-\lambda_1 x},\quad x\geqslant 0,$$
$$F_\xi(x)=1-\mathrm{e}^{-\lambda_1 x},\quad x\geqslant 0.$$

解

$$P\{\xi<\eta\}=\int_{-\infty}^{+\infty}P\{\xi<\eta\mid\eta=y\}f_\eta(y)\mathrm{d}y=\int_0^{+\infty}P\{\xi<y\mid\eta=y\}\lambda_2\mathrm{e}^{-\lambda_2 y}\mathrm{d}y$$
$$=\int_0^{+\infty}F_\xi(y)\lambda_2\mathrm{e}^{-\lambda_2 y}\mathrm{d}y=\int_0^{+\infty}[1-\mathrm{e}^{-\lambda_1 y}]\lambda_2\mathrm{e}^{-\lambda_2 y}\mathrm{d}y$$
$$=\int_0^{+\infty}[\lambda_2\mathrm{e}^{-\lambda_2 y}-\lambda_2\mathrm{e}^{-(\lambda_1+\lambda_2)y}]\mathrm{d}y=\int_0^{+\infty}\lambda_2\mathrm{e}^{-\lambda_2 y}\mathrm{d}y-\int_0^{+\infty}\lambda_2\mathrm{e}^{-(\lambda_1+\lambda_2)y}\mathrm{d}y$$
$$=1-\frac{\lambda_2}{\lambda_1+\lambda_2}\int_0^{+\infty}\mathrm{e}^{-(\lambda_1+\lambda_2)y}\mathrm{d}(\lambda_1+\lambda_2)y=1-\frac{\lambda_2}{\lambda_1+\lambda_2}$$
$$=\frac{\lambda_1}{\lambda_1+\lambda_2}.$$

例 1.6.3（抓阄问题） 一袋中 n 个白球、m 个黑球，不放回从袋中摸球，求第 k 次摸到黑球的概率.

解 设 A_k = "第 k 次摸到黑球"，$k=1,2,\cdots,m+n$. 显然