应用随机过程

刘秀芹 李 娜 赵金玲 编

科学出版社

应用随机过程

刘秀芹　李　娜　赵金玲　编

科 学 出 版 社

北　京

内 容 简 介

本书主要介绍随机过程的基础理论及其实际应用. 全书共 6 章, 内容包括概率论基础知识、随机过程的基本概念及其分类、泊松过程及其推广、马尔可夫过程、平稳过程及其谱分析. 各章配有练习题和相关的科学家简介.

本书以高等数学、概率论为基础, 可作为高等院校理工科等相关专业本科生与研究生的教学用书.

图书在版编目(CIP)数据

应用随机过程/刘秀芹, 李娜, 赵金玲编. —北京: 科学出版社, 2015.4

ISBN 978-7-03-044058-7

Ⅰ. ①应… Ⅱ. ①刘… ②李… ③赵… Ⅲ. ①随机过程-高等学校-教材 Ⅳ. ①O211.6

中国版本图书馆 CIP 数据核字(2015) 第 072279 号

责任编辑: 昌 盛 周金权 / 责任校对: 张凤琴
责任印制: 霍 兵 / 封面设计: 陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

北京市文林印务有限公司 印刷
科学出版社发行 各地新华书店经销
*
2015 年 5 月第 一 版 开本: 720 × 1000 1/16
2015 年 5 月第一次印刷 印张: 10 1/4
字数: 210 000

定价: 25.00 元

(如有印装质量问题, 我社负责调换)

前　言

随机过程是描述自然界中随机现象随着时间变化的统计规律的一门学科. 它在金融、通信、生物、控制等科学技术领域都有广泛的实际应用.

本书是由北京科技大学从事本科及研究生 “随机过程” 课程教学的主讲教师在教学与研究的基础上, 经过对所用讲义进行反复修改与补充编写而成的. 编者在编写过程中注重由浅入深, 表述简洁, 强调可读性, 以最朴素的语言介绍随机过程的基本理论和分析方法, 通过大量精选例题使读者能够比较容易地理解随机过程的基本概念; 采用形象生动的图形展现随机过程的一些关键知识点 (例如, 随机过程的样本曲线、柯尔莫哥洛夫前进后退方程的时间状态对应示意图等); 强调随机过程在实际中的应用, 结合编者的教学研究情况, 把随机过程在科技前沿中的应用引入到本书中 (例如, 平稳分布在 Google 搜索中的应用等).

本书主要介绍随机过程的基本理论和基本方法, 重点介绍实际应用中常见的几类随机过程. 全书共分为 6 章: 第 1 章, 预备知识; 第 2 章, 随机过程的概念与基本类型; 第 3 章, 泊松过程; 第 4 章, 马尔可夫链; 第 5 章, 连续时间的马尔可夫链; 第 6 章, 平稳过程.

本书的编写过程中, 李娜编写了第 1 章和第 4 章; 赵金玲编写了第 2 章和第 6 章; 刘秀芹编写了第 3 章和第 5 章并负责全书统稿. 中国科学院数学与系统科学院巩馥洲研究员、北京科技大学数理学院汪飞星教授、廖福成教授、张志刚副教授等多位老师为本书提出了丰富而宝贵的意见和建议, 编者在此对诸位老师表示衷心的感谢!

随机过程课程目前在我校多个专业的本科生和研究生中开设, 根据不同专业的专业设置和学时安排调整教学内容, 其中带有 $*$ 部分作为选讲内容. 本书的编写与出版得到了北京科技大学教材建设基金的资助.

限于作者水平, 书中难免有不足之处, 敬请读者批评指正.

编　者

2014 年 12 月于北京科技大学

目录

第 1 章　预 备 知 识

1.1　随机事件及概率

1.1.1　随机事件

实际生活中我们遇到过各种各样的试验, 如果一个试验具有以下三个特性:

(1) 可以在相同的条件下重复进行;

(2) 每次试验的结果不止一个, 但预先知道试验的所有可能结果;

(3) 每次试验前不能确定哪个结果会出现.

则称这样的试验为**随机试验**. 随机试验是概率论的基本概念, 试验的结果事先不能准确地预言.

随机试验所有可能结果组成的集合称为**样本空间**, 记作 Ω; Ω 的子集称为**随机事件**, 简称事件, 常用大写字母 A, B 等表示; 随机试验的每一个可能结果, 即 Ω 中的每一个元素称为**样本点**或**基本事件**, 记作 ω; 样本空间 Ω 称为**必然事件**; 空集 $\varnothing$ 称为**不可能事件**.

在一个样本空间中往往有很多的事件, 因此需要研究事件之间的关系和事件之间的运算. 由于事件是集合, 事件之间的关系和事件之间的运算可以按照集合论中集合的关系和运算来处理. 此时要求事件是 Ω 的子集, 同时满足以下三个条件:

(1) Ω 是事件;

(2) 若 A 是事件, 则 $\bar{A}$ 是事件;

(3) 若 A_i 是事件, 则 $\bigcup\limits_{i=1}^{\infty} A_i$ 是事件.

由此可知, 如果 A, B 是事件, 则 $A \cup B$, $A \cap B$ 等都是事件, 即事件经过有限次集合运算所得的都还是事件.

1.1.2　随机事件的概率

对于 Ω 的一个事件 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 $P(\cdot)$ 满足以下三个条件:

(1) 非负性: 对任意事件 A, 有 $0 \leqslant P(A) \leqslant 1$;

(2) 规范性: 对必然事件 Ω, 有 $P(\Omega) = 1$;

(3) 可列可加性: 对两两互不相容的事件$A_1, A_2, \cdots$, 即当 $i \neq j$ 时, $A_i \cap A_j = \varnothing$,

$i, j=1,2,\cdots$, 有 $P\left(\bigcup\limits_{i=1}^{\infty} A_i\right)=\sum\limits_{i=1}^{\infty} P(A_i)$.

下面给出概率的一些基本性质：

(1) $P(\varnothing)=0$;

(2) 有限可加性：对两两互不相容的事件 $A_1, A_2, \cdots, A_n$, 有 $P\left(\bigcup\limits_{i=1}^{n} A_i\right)=\sum\limits_{i=1}^{n} P(A_i)$;

(3) 对于任意事件 A, 有 $P(\bar{A})=1-P(A)$;

(4) 单调性：若 $A \subset B$, 则有 $P(A) \leqslant P(B)$;

(5) 加法公式：$P(A \cup B)=P(A)+P(B)-P(AB)$;

(6) 条件概率：当 $P(A)>0$ 时, $P(B|A)=\dfrac{P(AB)}{P(A)}$;

(7) 乘法公式：当 $P(A)>0$ 时, $P(AB)=P(A) \cdot P(B|A)$;

(8) 全概率公式: 若事件 $A_1, A_2, \cdots, A_n$ 是一组划分, 即 $\bigcup\limits_{i=1}^{n} A_i=\Omega$ 且 $A_i \cap A_j=\varnothing(i \neq j)$, $P(A_j)>0$, 有 $P(B)=\sum\limits_{i=1}^{n} P(A_i) P(B|A_i)$;

(9) 贝叶斯公式: 若事件 $A_1, A_2, \cdots, A_n$ 是一组划分, 即 $\bigcup\limits_{i=1}^{n} A_i=\Omega$ 且 $A_i \cap A_j=\varnothing(i \neq j), P(A_i)>0$, $P(B)>0$, 有 $P(A_i|B)=\dfrac{P(A_i) P(B|A_i)}{\sum\limits_{j=1}^{n} P(A_j) P(B|A_j)}$;

(10) 独立性：事件 A, B 相互独立, 即 $P(AB)=P(A) \cdot P(B)$.

1.2 随机变量及其分布

1.2.1 随机变量

设随机试验的样本空间为 $\Omega=\{\omega\}$, $X=X(\omega)$ 是定义在样本空间 Ω 上的实值单值函数, 且对任意给定的 $x \in \mathbf{R}$, $\{\omega | X(\omega) \leqslant x\}$ 的概率存在, 称 X 为**随机变量**. 一般地, 随机变量常用大写字母 X, Y, Z 等来表示, 其取值用小写字母 x, y, z 等来表示.

定义随机变量 X 的分布函数为：$F(x)=P\{X \leqslant x\}$, $x \in \mathbf{R}$. 分布函数具有以下三条性质：

(1) 单调性：$F(x)$ 在整个实数域内单调非减;

(2) 有界性：对任意的 $x, 0 \leqslant F(x) \leqslant 1$, 且 $\lim\limits_{x\to-\infty} F(x)=0$, $\lim\limits_{x\to+\infty} F(x)=1$;

(3) 右连续性：$F(x)$ 是 x 的右连续函数.

1.2.2 离散型随机变量

若随机变量 X 的可能取值仅有有限个或可列个, 则称此随机变量为**离散型随机变量**. 设 X 的所有可能取值为 $x_k, k=1,2,\cdots,X$ 取各个可能值的**概率**, 即事件 $\{X=x_k\}$ 的概率为 $P\{X=x_k\}=p_k$, $k=1,2,\cdots$, 则称 $P\{X=x_k\}=p_k$ 为离散型随机变量 X 的**概率分布律**或简称为**分布律**.

由定义可知, 分布律满足如下两条性质:

(1) 非负性：$p_k \geqslant 0, k=1,2,\cdots$;

(2) 完备性：$\sum\limits_{k=1}^{\infty} p_k=1$.

1.2.3 连续型随机变量

若对于随机变量 X 的分布函数, 存在非负函数 $f(x)$, 使得对于任意实数 x 有 $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$, 则称 X 为**连续型随机变量**, 其中 $f(x)$ 称为 X 的**概率密度函数**.

由分布函数的性质可知, 任意连续型随机变量的概率密度函数 $f(x)$ 都具有如下性质:

(1) 非负性：$f(x) \geqslant 0$;

(2) 正则性：$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$;

(3) 若 $f(x)$ 在点 x 处连续, 则有 $F'(x)=f(x)$;

(4) 对任意实数 $x_1, x_2 (x_1 \leqslant x_2)$ 有

$$P\{x_1<X<x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)\mathrm{d}x.$$

1.2.4 常见随机变量的分布

(1) 两点分布：若随机变量 X 的分布律为：$P\{X=k\}=p^k(1-p)^{1-k}$, $k=0,1$, 这时称随机变量服从两点分布, 又称伯努利分布, 记为 $X\sim(0,1)$.

(2) 二项分布：若随机变量 X 的分布律为：$P\{X=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k}$, $k=0,1,\cdots,n$, 这时称随机变量服从以 n 和 p 为参数的二项分布, 记为 $X\sim B(n,p)$.

(3) 泊松分布：若随机变量 X 的分布律为：$P\{X=k\}=\dfrac{\lambda^k \mathrm{e}^{-\lambda}}{k!}$, $k=0,1,\cdots$,

这时称随机变量服从以 λ 为参数的泊松分布, 记为 $X \sim P(\lambda)$ 或 $X \sim \pi(\lambda)$.

(4) 几何分布: 若随机变量 X 的分布律为: $P\{X=k\}=p(1-p)^{k-1}, k=1,2,\cdots$, 则称随机变量服从几何分布.

(5) 均匀分布: 若随机变量 X 的概率密度函数为: $f(x)=\begin{cases} \dfrac{1}{b-a}, & x\in[a,b], \\ 0, & x\notin[a,b], \end{cases}$ 这时称随机变量服从区间 $[a,b]$ 上的均匀分布, 记为 $X \sim U(a,b)$.

(6) 指数分布: 若随机变量 X 的概率密度函数为: $f(x)=\begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x\geqslant 0, \\ 0, & x<0, \end{cases}$ 这时称随机变量服从参数为 λ 的指数分布, 记为 $X \sim Z(\lambda)$.

(7) 正态分布: 若随机变量 X 的概率密度函数为: $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, x\in\mathbf{R}$, 这时称随机变量服从参数为 μ, σ^2 的正态分布, 记为 $X \sim N(\mu,\sigma^2)$.

(8) 伽马分布: 若随机变量 X 的概率密度函数为: $f(x)=\begin{cases} \dfrac{\lambda^s}{\Gamma(s)}x^{s-1}\mathrm{e}^{-\lambda x}, & x\geqslant 0, \\ 0, & x<0, \end{cases}$ 其中 $\Gamma(s)=\displaystyle\int_0^{+\infty}\mathrm{e}^{-x}x^{s-1}\mathrm{d}x, s>0$, 这时称随机变量服从参数为 $s>0, \lambda>0$ 的伽马分布, 记为 $X \sim \Gamma(s,\lambda)$.

1.3 随机变量的数字特征

1.3.1 数学期望

设离散型随机变量 X 的分布律为: $P\{X=x_k\}=p_k, k=1,2,\cdots$, 若级数 $\displaystyle\sum_{k=1}^{\infty}x_kp_k$ 绝对收敛, 则称级数 $\displaystyle\sum_{k=1}^{\infty}x_kp_k$ 的和为随机变量 X 的**数学期望**, 记为 $E(X)$, 即为 $E(X)=\displaystyle\sum_{k=1}^{\infty}x_kp_k$, 设连续型随机变量 X 的概率密度函数为 $f(x)$, 若积分 $\displaystyle\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 绝对收敛, 则称积分 $\displaystyle\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$ 的值为随机变量 X 的**数学期望**, 记为 $E(X)$, 即为 $E(X)=\displaystyle\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$.

1.3.2 方差与矩

设 X 为随机变量, 若 $E(X-E(X))^2$ 存在, 则称它为随机变量 X 的**方差**, 记为 $D(X)$, 即 $D(X)=E(X-E(X))^2=E(X^2)-(E(X))^2$, 称 $\sqrt{D(X)}$ 为随机变量 X 的**标准差**, **或均方差**.

设 X 为随机变量, 若 $E(X^k)$ 存在 (k 为正整数), 则称它为随机变量 X 的 k**阶原点矩**, 记为 μ_k, 即 $\mu_k=E(X^k)$. 数学期望是一阶原点矩.

设 X 为随机变量, 若 $E(X-E(X))^k$ 存在 (k 为正整数), 则称它为随机变量 X 的 k**阶中心矩**, 记为 ν_k, 即 $\nu_k=E(X-E(X))^k$. 方差是二阶中心矩.

1.3.3 协方差

设 (X,Y) 为二维随机变量, 若 $E[(X-E(X))(Y-E(Y))]$ 存在, 则称它为随机变量 X,Y 的**协方差**, 记为 $\mathrm{cov}(X,Y)$, 即

$$\mathrm{cov}(X,Y)=E[(X-E(X))(Y-E(Y))]=E(XY)-E(X)E(Y).$$

设 X,Y 为两个随机变量, 若 $E((X-E(X))^k(Y-E(Y))^l)$ 存在 (k,l 为正整数), 则称它为随机变量 X,Y 的 $k+l$**阶混合中心矩**, 协方差是二阶混合中心矩.

1.4 条件期望

1.4.1 条件期望的定义

设 X,Y 为离散型随机变量, 对一切使 $P\{Y=y_j\}\neq 0$ 的 y_j, 定义给定 $Y=y_j$ 时, X 的条件分布律为

$$P\{X=x_i|Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}.$$

给定 $Y=y$ 时, X 的条件分布函数为

$$F_{X|Y}(x|y)=P\{X\leqslant x|Y=y\}.$$

而给定 $Y=y_j$ 时, X 的条件期望为

$$E(X|Y=y_j)=\sum_i x_iP\{X=x_i|Y=y_j\}.$$

例 1.4.1 一个射手进行射击, 命中目标的概率为 p, $0<p<1$, 射击到命中两次目标为止, 设 X 表示首次命中目标所进行的射击次数, Y 表示总共进行的射击次数, 试求 X 和 Y 的条件期望.

解 根据题意可知, (X,Y) 的联合分布律为

$$P\{X=i,Y=j\}=p^2(1-p)^{j-2},\quad i=1,2,\cdots,\quad j=2,3,\cdots,$$

则 X 与 Y 的边缘分布律为

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

$$P\{Y=j\}=(j-1)p^2(1-p)^{j-2},\quad j=2,3,\cdots,$$

当 $Y=j(j=2,3,\cdots)$ 时, X 的条件分布律为

$$P\{X=i|Y=j\}=\frac{1}{j-1},\quad i=1,2,\cdots,j-1,$$

则

$$E\{X|Y=j\}=\sum_{i=1}^{j-1}\frac{i}{j-1}=\frac{j}{2},\quad j=2,3,\cdots.$$

同理, 当 $X=i(i=1,2,\cdots)$ 时, Y 的条件分布律为

$$P\{Y=j|X=i\}=p(1-p)^{j-i-1},\quad j=i+1,i+2,\cdots,$$

则

$$E\{Y|X=i\}=\sum_{j=i+1}^{\infty}jp(1-p)^{j-i-1}=\sum_{k=0}^{\infty}(k+i+1)p(1-p)^k$$
$$=\sum_{k=0}^{\infty}(k+1)p(1-p)^k+i\sum_{k=0}^{\infty}p(1-p)^k=\frac{1}{p}+i,\quad i=1,2.\cdots.$$

设 X,Y 为连续型随机变量, 联合概率密度函数为 $f(x,y)$, 则对一切使 $f_Y(y)>0$ 的 y, 定义给定 $Y=y,X$ 的条件概率密度为

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}.$$

给定 $Y=y$ 时, X 的条件分布函数为

$$F_{X|Y}(x|y)=P\{X\leqslant x|Y=y\}=\int_{-\infty}^{x}f_{X|Y}(x|y)\mathrm{d}x.$$

而给定 $Y=y$ 时, X 的条件期望为

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

因此, 除了概率是关于事件 $Y=y$ 的条件概率外, 一切定义均与无条件的情形一样.

例 1.4.2 设在 $Y=y,0<y<1$ 条件下 X 的条件概率密度为

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{3x^2}{y^3}, & 0<x<y,\\ 0, & \text{其他},\end{cases}$$

试求条件期望 $E(X|Y=y)$.

解 根据条件期望的定义, 可得

$$E(X|Y=y)=\int_0^y x\frac{3x^2}{y^3}\mathrm{d}x=\frac{1}{y^3}\cdot\frac{1}{4}x^4\Big|_0^y=\frac{3y}{4}.$$

例 1.4.3 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 求 $E(X|Y=y)$.

解 由条件概率密度的定义: $f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}$, 将已知条件代入即得

$$f_{XY}(x|y)=\frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}\right)^2}.$$

显然, 它也是正态分布的密度函数, 服从 $N\left(\mu_1+\rho\dfrac{\sigma_1}{\sigma_2}(y-\mu_2),(1-\rho^2)\sigma_1^2\right)$. 因此, $E(X|Y=y)=\mu_1+\rho\dfrac{\sigma_1}{\sigma_2}(y-\mu_2)$.

例 1.4.3 中的条件期望 $E(X|Y=y)$ 是 y 的线性函数. 一般地, $E(X|Y=y)$ 只是 y 的函数, 以 $E(X|Y)$ 表示随机变量 Y 的函数, 在 $Y=y$ 时的取值为 $E(X|Y=y)$, 因此它也是一个随机变量, 称为 X 在 Y 条件下的条件期望.

1.4.2 条件期望的性质

条件期望在概率论、数理统计和随机过程中是一个十分重要的概念, 下面介绍条件期望的一些性质.

(1) 全期望公式: 若随机变量 X 与 Y 的期望存在, 则

$$E(X)=E[E(X|Y)]=\int E(X|Y=y)\mathrm{d}F_Y(y).$$

如果 Y 是离散型随机变量, 则 $E(X)=\sum\limits_j E(X|Y=y_j)P\{Y=y_j\}$.

如果 Y 是连续型随机变量, 具有概率密度 $f_Y(x)$, 则

$$E(X)=\int_{-\infty}^{\infty}E(X|Y=y)f_Y(y)\mathrm{d}y.$$

现在给出 X 与 Y 都是离散型随机变量时的证明.

$$\begin{aligned}&\sum_j E(X|Y=y_j)P\{Y=y_j\}\\&=\sum_i\sum_j x_iP\{X=x_i|Y=y_j\}P\{Y=y_j\}\\&=\sum_i\sum_j x_iP\{X=x_i,Y=y_j\}\end{aligned}$$

$$=\sum_i x_i \sum_j P\{X=x_i, Y=y_j\}$$
$$=\sum_i x_i P\{X=x_i\}=E(X).$$

连续型的情形可类似证明.

(2) 设 X,Y 相互独立, 则 $E(X|Y)=E(X), E(Y|X)=E(Y)$.

(3) $E(g(Y)X|Y)=g(Y)E(X|Y)$.

实际上, 对于任意固定的 y, 有等式 $E(g(y)X|Y=y)=g(y)E(X|Y=y)$ 成立, 可知结论为真. 由条件期望的定义和数学期望的性质还可以得到以下性质:

(4) $E(g(Y)X)=E(g(Y)E(X|Y))$.

(5) $E(c|Y)=c$, c 为常数.

(6) $E(g(Y)|Y)=g(Y)$.

(7) 线性可加性 $E((aX+bY)|Z)=aE(X|Z)+bE(Y|Z)$.

例 1.4.4　设在某一天走进商店的人数是期望为 1000 的随机变量, 又设这些顾客在该商店所花钱数都为期望为 100 元的相互独立的随机变量, 并设单个顾客的花钱数和进入该商店的总人数独立. 问在给定的一天内, 顾客在该商店所花钱数的期望是多少?

解　设 N 表示这天进入该商店的总人数, X_i 表示第 i 个顾客所花的钱数, 则 N 个顾客所花的钱数总和为 $\sum\limits_{i=1}^{N} X_i$. 由于

$$E\left(\sum_{i=1}^{N} X_i\right)=E\left[E\left(\sum_{i=1}^{N} X_i|N\right)\right],$$

又

$$E\left(\sum_{i=1}^{N} X_i|N=n\right)=E\left(\sum_{i=1}^{n} X_i|N=n\right)=E\left(\sum_{i=1}^{n} X_i\right)=nE(X_1).$$

则

$$E\left(\sum_{i-1}^{N} X_i\right)=E\left[E\left(\sum_{i=1}^{N} X_i|N\right)\right]=E(NE(X_1))=E(N)E(X_1).$$

根据题目已知 $E(N)=1000, E(X_1)=100$, 于是

$$E\left(\sum_{i=1}^{N} X_i\right)=1000\times 100=100000.$$

因此该天顾客花费在该商店的钱数的期望为 100000 元.

例 1.4.5(巴格达窃贼问题)　一窃贼被关在有 3 个门的地牢中, 其中第一个门通向自由, 出这个门后 3 个小时便回到地面; 第二个门通向一个地道, 在此地道中

走 5 个小时后将返回地牢; 第三个门通向一个更长的地道, 沿这个地道走 7 个小时也回到地牢, 如果窃贼每次选择 3 个门的可能性总相等, 试求他为获得自由而奔走的平均时间.

解 设窃贼需要走 X 个小时到达地面, 并设 Y 为窃贼每次对 3 个门的选择, 则 Y 均以 $\frac{1}{3}$ 的概率取值为 1, 2, 3, 可利用全期望公式得

$$E(X) = E[E(X|Y)] = \sum_{j=1}^{3} E(X|Y=j)P\{Y=j\},$$

而由题意可知, $E(X|Y=1) = 3, E(X|Y=2) = 5 + E(X), E(X|Y=1) = 7 + E(X)$, 因此,

$$E(X) = \frac{1}{3}[3 + 5 + E(X) + 7 + E(X)],$$

求解得 $E(X) = 15$, 即窃贼为获得自由而奔走的平均时间是 15 小时.

1.5 特征函数

通过前面的讨论, 我们已经知道如何去计算随机变量的数字特征, 数字特征一般由各阶矩决定, 随着阶数的增高, 矩的计算总是比较麻烦的. 而且在一般情况下, 随机变量的数学期望和方差只能粗略地反映分布函数的某些特征性质, 不能完整地刻画分布函数; 另一方面, 从前面的讨论我们就看到, 只利用分布函数和概率密度函数求独立随机变量的和的分布比较麻烦, 需要用到卷积, 要解决复杂的问题没有更优越的数学工具是不行的. 我们在学习数学分析时知道傅里叶变换可以把卷积运算变成乘法运算, 它在数学中是非常重要而有效的工具, 将傅里叶变换引入到概率论中, 就产生了"特征函数"的概念. 可以证明, 不同的分布函数对应着不同的特征函数, 而特征函数具有简单实用的特点. 因此在研究随机变量的分布特性时, 特征函数起着重要的作用. 毫不夸张地说, 概率统计自从引入特征函数以后, 就把理论的研究推上了一个新的台阶.

1.5.1 复随机变量

设 X 与 Y 是两个实随机变量, 则称 $Z = X + \mathrm{i}Y$ 为**复随机变量**, 其中 $\mathrm{i} = \sqrt{-1}$. 如果二维随机变量 $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ 相互独立, 则称复随机变量 $Z_1 = X_1 + \mathrm{i}Y_1, Z_2 = X_2 + \mathrm{i}Y_2, \cdots, Z_n = X_n + \mathrm{i}Y_n$ **是相互独立的**.

如果随机变量 X 和 Y 的数学期望存在, 记为 $E(X)$ 和 $E(Y)$, 则称 $E(Z) = E(X) + \mathrm{i}E(Y)$ 为复随机变量 Z 的**数学期望**.

显然可知, 关于实随机变量数学期望的性质对于复随机变量也同样成立.

1.5.2　特征函数的定义

设 X 为随机变量, 则称复随机变量 $\mathrm{e}^{\mathrm{i}uX}$ 的数学期望为 X 的**特征函数**, 记为 $\varphi_X(u)$, 即

$$\varphi_X(u)=E(\mathrm{e}^{\mathrm{i}uX}),\quad u\in(-\infty,+\infty).$$

由于对于任意的实数 u, 总有 $|\mathrm{e}^{\mathrm{i}uX}|=1$, 所以 $E(\mathrm{e}^{\mathrm{i}uX})$ 总是存在的, 即对一切随机变量其特征函数都存在.

1.5.3　常见分布的特征函数

1. *两点分布*

设 X 服从两点分布, 其分布律为 $P\{X=1\}=p, P\{X=0\}=1-p, 0<p<1$, 则其特征函数为

$$\varphi_X(u)=E(\mathrm{e}^{\mathrm{i}uX})=\mathrm{e}^{\mathrm{i}u\cdot 1}\cdot p+\mathrm{e}^{\mathrm{i}u\cdot 0}\cdot(1-p)=1-p+p\cdot\mathrm{e}^{\mathrm{i}u}.$$

2. *二项分布*

设 X 服从二项分布 $B(n,p)$, 其分布律为

$$P\{X=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k},\quad k=0,1,\cdots,n,0<p<1,$$

则其特征函数为

$$\begin{aligned}\varphi_X(u)=&E(\mathrm{e}^{\mathrm{i}uX})=\sum_{k=0}^{n}\mathrm{e}^{\mathrm{i}uk}\cdot\mathrm{C}_n^k p^k(1-p)^{n-k}\\=&\sum_{k=0}^{n}\mathrm{C}_n^k(p\mathrm{e}^{\mathrm{i}u})^k(1-p)^{n-k}=(1-p+p\mathrm{e}^{\mathrm{i}u})^n.\end{aligned}$$

3. *泊松分布*

设 X 服从泊松分布 $X\sim\pi(\lambda)$, 其分布律为

$$P\{X=k\}=\lambda^k\frac{\mathrm{e}^{-\lambda}}{k!},\quad k=0,1,\cdots,\lambda>0,$$

则其特征函数为

$$\varphi_X(u)=E(\mathrm{e}^{\mathrm{i}uX})=\sum_{k=0}^{+\infty}\frac{\mathrm{e}^{\mathrm{i}uk}\cdot\lambda^k\mathrm{e}^{-\lambda}}{k!}=\sum_{k=0}^{+\infty}\frac{(\lambda\mathrm{e}^{\mathrm{i}u})^k\mathrm{e}^{-\lambda}}{k!}=\mathrm{e}^{\lambda\mathrm{e}^{\mathrm{i}u}-\lambda}.$$

4. 均匀分布

设 X 服从均匀分布, 其概率密度函数为

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

则其特征函数为

$$\begin{aligned}\varphi_X(u)=&E(\mathrm{e}^{\mathrm{i}uX})=\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}ux}f(x)\mathrm{d}x=\int_a^b\mathrm{e}^{\mathrm{i}ux}\frac{1}{b-a}\mathrm{d}x\\ =&\frac{1}{b-a}\frac{1}{\mathrm{i}u}\mathrm{e}^{\mathrm{i}ux}\Big|_a^b=\frac{-\mathrm{i}}{(b-a)u}(\mathrm{e}^{\mathrm{i}ub}-\mathrm{e}^{\mathrm{i}ua}).\end{aligned}$$

5. 指数分布

设 X 服从指数分布, 其概率密度函数为

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

则其特征函数为

$$\begin{aligned}\varphi_X(u)=&E(\mathrm{e}^{\mathrm{i}uX})=\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}ux}\cdot f(x)\mathrm{d}x=\int_0^{+\infty}\lambda\mathrm{e}^{\mathrm{i}ux}\mathrm{e}^{-\lambda x}\mathrm{d}x=\lambda\int_0^{+\infty}\mathrm{e}^{(\mathrm{i}u-\lambda)}\mathrm{d}x\\ =&\frac{\lambda}{\mathrm{i}u-\lambda}\int_0^{+\infty}\mathrm{e}^{(\mathrm{i}u-\lambda)x}\mathrm{d}(\mathrm{i}u-\lambda)x=\frac{\lambda}{\mathrm{i}t-\lambda}\mathrm{e}^{(\mathrm{i}u-\lambda)x}\Big|_0^{+\infty}\\ =&\frac{\lambda}{\mathrm{i}u-\lambda}(0-1)=\frac{\lambda}{\lambda-\mathrm{i}u}=\frac{\lambda^2+\mathrm{i}\lambda u}{\lambda^2+u^2}.\end{aligned}$$

6. 标准正态分布

设 X 服从标准正态分布, 其概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}},$$

则其特征函数为

$$\begin{aligned}\varphi_X(u)=E(\mathrm{e}^{\mathrm{i}uX})=&\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}ux}\cdot\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}ux-\frac{x^2}{2}}\mathrm{d}x\\ =&\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2}[(x-\mathrm{i}u)^2-(\mathrm{i}u)^2]}\mathrm{d}x\\ =&\mathrm{e}^{-\frac{u^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-\frac{1}{2}(x-\mathrm{i}u)^2}\mathrm{d}x,\end{aligned}$$