

前 言

党的二十大报告指出：“我们要坚持教育优先发展、科技自立自强、人才引领驱动，加快建设教育强国、科技强国、人才强国，坚持为党育人、为国育才，全面提高人才自主培养质量，着力造就拔尖创新人才，聚天下英才而用之。”^①教材是教学内容的主要载体，是教学的重要依据、培养人才的重要保障。在优秀教材的编写道路上，我们一直在努力。

在数字经济与智能技术深度融合的今天，全球产业链正在经历前所未有的重构，复杂系统的不确定性呈指数级增长。在这一背景下，如何在高波动、多维度随机环境中实现资源的最优配置与系统的韧性设计，已成为现代科学与工程领域亟待解决的核心挑战。随着现实问题的日益复杂化，运筹学作为一门融合数学建模与科学决策的交叉学科，面临着更加复杂的建模需求，迫切需要为其提供有效的解决方案。随机运筹学作为运筹学的重要分支，以概率论与随机过程为理论基础，专注于研究随机因素主导下的系统分析与优化问题，逐渐成为解决复杂建模难题的关键方法。

本书的编写源于四川大学在运筹学领域多年的积累与沉淀。自 20 世纪 80 年代起，四川大学便开设了“随机运筹学”课程并面向全校开放授课，至今已有 40 余年历史。在课程建设初期，赵玮和王荫清两位前辈于 1993 年编著的《随机运筹学》，成为国内首部系统的随机运筹学专业教材，在全国重点高校得到广泛应用。随着运筹学知识在工科类专业与经管类专业中的重要性日益凸显，教学实践中却暴露出一些问题：一方面，面向非数学专业的运筹学教材对于随机性问题的探讨往往停留在表面，缺乏深入分析，而在实际应用场景中，随机性恰恰是影响系统效能的关键因素；另一方面，面向数学专业的运筹学教材难度较高且缺乏足够的应用场景，让非数学专业学生既难以理解，也难以将理论与实际相结合。基于此，作者结合十余年经管类与工科类专业的教学经验，对随机运筹学的知识框架与教学难度进行了优化，注重提高理论的可接受性以及理论与实践的结合能力。本书适合工科类、经管类专业的本科生与研究生使用。

本书的出版得到了四川大学的资助，以及四川大学商学院与科学出版社的大力支持，在此深表谢意。由于作者水平有限，书中难免存在疏漏，恳请广大读者与学界同行不吝指正。若本书能为我国运筹学的教学研究贡献绵薄之力，我将倍感荣幸。

作 者

2026 年 2 月于成都

^① 《习近平：高举中国特色社会主义伟大旗帜为全面建设社会主义现代化国家而团结奋斗——在中国共产党第二十次全国代表大会上的报告》，https://www.gov.cn/xinwen/2022-10/25/content_5721685.htm，2022 年 10 月 25 日。

目 录

第 1 章 概率论预备知识	1
1.1 概率的元素	1
1.2 随机变量	2
1.3 二元随机变量	4
1.4 多元随机变量	8
1.5 常用分布	10
第 2 章 马尔可夫链	22
2.1 随机过程	22
2.2 马尔可夫链的定义	23
2.3 马尔可夫链的状态划分	33
2.4 马尔可夫链的长期特征	35
2.5 首达时间	42
2.6 吸收链	44
本章习题	47
第 3 章 排队论	53
3.1 排队模型的基本组成	53
3.2 到达和服务过程的刻画	56
3.3 生灭过程	63
3.4 稳态概率	66
3.5 $M/M/1/GD/\infty/\infty$ 模型	69
3.6 $M/M/1/GD/c/\infty$ 模型	74
3.7 $M/M/s/GD/\infty/\infty$ 模型	75
3.8 非指数分布的排队模型	82
3.9 有优先级的排队模型	85
本章习题	87
第 4 章 库存控制	98
4.1 库存模型的基础介绍	99

4.2 确定性库存模型	101
4.3 随机库存模型	115
本章习题	124
第 5 章 可靠性	133
5.1 系统的结构函数	133
5.2 系统可靠性	135
5.3 可靠性的计算	137
5.4 可靠性的边界	140
本章习题	144
参考文献	146

第 1 章 概率论预备知识

面对客观存在的不确定结果，概率论是目前对不确定性的规范化认知，同时也是处理不确定性问题的有效工具。通过本章的学习，我们将回顾概率论的关键概念。需要说明的是，概率的数学理论极为复杂，它的建立与发展以数学中的测度论为基础。对于非数学专业的学生，我们不强调背后的数学理论，而更关注有关概率的基本处理方法。

1.1 概率的元素

样本空间 Ω ：一个随机试验所有结果的集合。这里，每个结果 $\omega \in \Omega$ 可以被理解为对现实世界一个状态的完整描述。

事件空间 $\mathcal{F}$ ： $\mathcal{F}$ 是 Ω 上的一个 σ -代数^①，指样本空间中某些子集组成的集合类。

概率的公理化定义：一个函数 $P: \mathcal{F} \rightarrow \mathbb{R}$ ，如果满足下述三条性质，则被称作概率。

(1) $0 \leq P(A) \leq 1, \forall A \in \mathcal{F}$ 。

(2) $P(\Omega) = 1$ 。

(3) 如果 $A_1, A_2, \dots$ 不相交 (即 $A_i \cap A_j = \emptyset$, 对任意 $i \neq j$ 成立)，则 $P(\cup_i A_i) = \sum_i P(A_i)$ 。

例 1.1 对于一个六面骰子，它的样本空间为 $\Omega = \{1, 2, 3, 4, 5, 6\}$ 。我们可以在这个样本空间定义不同的事件空间，如事件空间 $\mathcal{F} = \{\emptyset, \Omega\}$ ，再如事件空间包含样本空间 Ω 的全部子集。对于前面的例子，两个事件的概率为 $P(\emptyset) = 0, P(\Omega) = 1$ ；对于后面的例子，每个含 i 个元素的事件空间的概率是 $\frac{i}{6}$ ，如 $P(\{1, 2, 3, 4\}) = \frac{4}{6}$ 、 $P(\{1, 2, 3\}) = \frac{3}{6}$ 。

性质：

(1) $A \subseteq B \implies P(A) \leq P(B)$ ；

(2) $P(A \cap B) \leq \min(P(A), P(B))$ ；

(3) $P(A \cup B) \leq P(A) + P(B)$ ；

(4) $P(\Omega \setminus A) = 1 - P(A)$ ；

(5) 如果 $A_1, \dots, A_k$ 是不相交集合且满足 $\cup_{i=1}^k A_i = \Omega$ ，则 $\sum_{i=1}^k P(A_i) = 1$ ^②。

令事件 B 的概率非零。在给定事件 B 下，任意事件 A 的条件概率定义为

$$P(A | B) \triangleq \frac{P(A \cap B)}{P(B)}$$

^① $\mathcal{F}$ 是 Ω 上的 σ -代数需要满足以下性质。包含全集： $\Omega \in \mathcal{F}$ 。补运算封闭： $A \in \mathcal{F} \implies \Omega \setminus A \in \mathcal{F}$ 。对可数并封闭：如果 $\{A_n\}_{n=1}^{\infty}$ 是 $\mathcal{F}$ 中的一列事件，则它们的并集也属于 $\mathcal{F}$ ，即 $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$ 。对可数交封闭：如果 $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{F}$ ，则 $\cap_{n=1}^{\infty} A_n \in \mathcal{F}$ 。

^② 该公式被称作全概率公式。

也就是说, $P(A|B)$ 是在事件 B 已发生的前提下, 对事件 A 发生概率的度量。两个事件是独立的, 当且仅当 $P(A \cap B) = P(A)P(B)$, 或 $P(A|B) = P(A)$ 。因此, 独立性意味着事件 B 是否发生不影响事件 A 发生的概率。

1.2 随机变量

考虑掷 10 个硬币的试验, 我们想知道出现正面的硬币数量。对于这个问题, 样本空间 Ω 的元素是长度为 10 并由正反面组成的所有可能序列。例如, $\omega_0 = \{H, H, T, H, T, H, H, T, T, T\} \in \Omega$ (H 表示正面, T 表示反面)。然而, 在现实中, 我们通常并不关注某一族序列 (特定结果) 的概率; 相反, 我们更关注某些结果的数值特征, 如正面出现的次数。这些函数被称为随机变量。

正式来说, 随机变量 X 是从样本空间到实数空间的映射, 即 $X: \Omega \rightarrow \mathbb{R}^{\text{①}}$ 。通常我们用大写字母 $X(\omega)$ 或更简单的 X 来表示随机变量, 用小写字母 x 表示这个随机变量的实现值。

例 1.2 在上面提到的试验中, 假设 $X(\omega)$ 表示在结果 ω 中硬币出现正面的次数。在掷 10 个硬币的设定下, $X(\omega)$ 的取值一定是有限的, 因此被称为离散随机变量。相应地, 一个随机变量 X 取特定值 k 的概率为

$$P(X = k) := P(\{\omega : X(\omega) = k\})$$

例 1.3 假设 $X(\omega)$ 作为一个随机变量, 刻画某个辐射物衰减的时间。在这种情形下, $X(\omega)$ 的取值是无限的, 因此被称为连续随机变量。一个随机变量的实现值落在两个实数 a 和 b (其中 $a < b$) 之间的概率为

$$P(a \leq X \leq b) := P(\{\omega : a \leq X(\omega) \leq b\})$$

1.2.1 累积分布函数

累积分布函数 (cumulative distribution function, CDF) $F_X: \mathbb{R} \rightarrow [0, 1]$ 是一种特定的概率度量, 即

$$F_X(x) \triangleq P(X \leq x)$$

利用该函数我们可以计算出事件空间 $\mathcal{F}$ 中任意事件发生的概率。

性质:

- (1) $0 \leq F_X(x) \leq 1$;
- (2) $\lim_{x \rightarrow -\infty} F_X(x) = 0$;
- (3) $\lim_{x \rightarrow \infty} F_X(x) = 1$;
- (4) $x \leq y \implies F_X(x) \leq F_X(y)$ 。

① 不是任意一个函数都可以被认为是随机变量。从测度论的角度来说, 随机变量必须是博雷尔可测 (Borel-measurable) 的函数。

1.2.2 概率质量函数

当一个随机变量的取值来自有限集 (即 X 是离散随机变量), 一种更简单地刻画概率度量的方式是明确随机变量每个取值对应的概率。具体来说, 概率质量函数 (probability mass function, PMF) $p_X : \Omega \rightarrow \mathbb{R}$ 定义为 $p_X(x) \triangleq P(X = x)$ 。

对于离散随机变量, 我们用符号 $\text{Val}(X)$ 表示随机变量 X 所有可能结果的集合。例如, 如果 $X(\omega)$ 是刻画掷 10 个硬币出现正面次数的随机变量, 则 $\text{Val}(X) = \{0, 1, 2, \dots, 10\}$ 。

性质:

- (1) $0 \leq p_X(x) \leq 1$;
- (2) $\sum_{x \in \text{Val}(X)} p_X(x) = 1$;
- (3) $\sum_{x \in A} p_X(x) = P(X \in A)$ 。

1.2.3 概率密度函数

对于某些连续随机变量, 若其累积分布函数 $F_X(x)$ 是处处可导的, 在这种情形下, 我们定义累积分布函数的导数为概率密度函数 (probability density function, PDF), 即

$$f_X(x) \triangleq \frac{dF_X(x)}{dx}$$

需要说明的是, 连续随机变量的概率密度函数并不总是存在, 因为 $F_X(x)$ 不一定处处可导。根据导数的定义, 对于充分小的 Δx , 我们有

$$P(x \leq X \leq x + \Delta x) \approx f_X(x) \Delta x$$

因此, 累积分布函数和概率密度函数可以用来计算不同事件发生的概率。但特别需要说明的是, 密度函数在任意一点 x 处的取值不是 $X = x$ 的概率, 即 $f_X(x) \neq P(X = x)$ 。另外, $f_X(x)$ 的取值非负且可以大于 1, 这从另一方面也说明了, $f_X(x)$ 的值不代表某个事件发生的概率。

性质:

- (1) $f_X(x) \geq 0$;
- (2) $\int_{-\infty}^{\infty} f_X(x) dx = 1$;
- (3) $\int_{x \in A} f_X(x) dx = P(X \in A)$ 。

1.2.4 期望

假设 X 是一个离散随机变量, 其概率质量函数为 $p_X(x)$ 。 $g : \mathbb{R} \rightarrow \mathbb{R}$ 是任意一个函数。为此, $g(X)$ 也是一个随机变量。我们定义 $g(X)$ 的期望如下:

$$E[g(X)] \triangleq \sum_{x \in \text{Val}(X)} g(x) p_X(x)$$

如果 X 是连续随机变量, 其概率密度函数为 $f_X(x)$, 则 $g(X)$ 的期望值定义为

$$E[g(X)] \triangleq \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

很显然，期望可以被理解为 X 所有取值的一种“加权平均”，其权重为 $p_X(x)$ 或 $f_X(x)$ 。作为特例，如果 $g(x) = x$ ，我们把 $E(X)$ 称为随机变量 X 的期望。

性质：

- (1) 对任意常数 $a \in \mathbb{R}$, $E(a) = a$;
- (2) 对任意常数 $a \in \mathbb{R}$, $E[af(X)] = aE[f(X)]$;
- (3) $E[f(X) + g(X)] = E[f(X)] + E[g(X)]$ ^①;
- (4) 对于离散随机变量 X , $E(1_{\{x=k\}}) = P(X = k)$ 。

例 1.4 假设 $g(x) = 1_{\{x \in A\}}$ ，其中 $1_{\{x \in A\}}$ 为示性函数且 $A \subseteq \Omega$ ，那么 $E[g(X)]$ 如何计算？

对于离散情形：

$$E[g(X)] = \sum_{x \in \text{Val}(X)} 1_{\{x \in A\}} P_X(x) = \sum_{x \in A} P_X(x) = P(x \in A)$$

对于连续情形：

$$E[g(X)] = \int_{-\infty}^{\infty} 1_{\{x \in A\}} f_X(x) dx = \int_{x \in A} f_X(x) dx = P(x \in A)$$

1.2.5 方差

随机变量的方差刻画均值中心化的程度。正式定义为 $\text{Var}(X) \triangleq E[(X - E(X))^2]$ 。利用基本的运算，我们可以将方差改写为

$$\begin{aligned} E[(X - E(X))^2] &= E[X^2 - 2E(X)X + E(X)^2] \\ &= E(X^2) - 2E(X)E(X) + E(X)^2 \\ &= E(X^2) - E(X)^2 \end{aligned}$$

性质：

- (1) 对任意常数 $a \in \mathbb{R}$, $\text{Var}(a) = 0$;
- (2) 对任意常数 $a \in \mathbb{R}$, $\text{Var}[af(X)] = a^2 \text{Var}[f(X)]$ 。

1.3 二元随机变量

在很多随机试验中，我们感兴趣的数值可能不止一个。例如，在连续掷 10 次硬币的试验中，我们可能关注正面出现的次数 $X(\omega)$ ，与此同时，也可能关注连续出现正面的次数 $Y(\omega)$ 。为此，我们需要考虑二元随机变量。

1.3.1 联合与边际分布

对于两个随机变量 X 和 Y ，尽管我们可以分开考虑这两个随机变量，但 X 和 Y 可以同时影响某一个随机试验的结果，因此需要更复杂的结构，即 X 和 Y 的联合累积分

^① 该性质表示期望的线性性。

布函数，其定义如下：

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

我们可以利用联合累积分布函数计算任意与 X 和 Y 相关事件的概率。此时，联合累积分布函数 $F_{XY}(x, y)$ 与每个随机变量的分布函数 $F_X(x)$ 和 $F_Y(y)$ 的关系如下：

$$F_X(x) = \lim_{y \rightarrow \infty} F_{XY}(x, y)$$

$$F_Y(y) = \lim_{x \rightarrow \infty} F_{XY}(x, y)$$

这里将 $F_X(x)$ 和 $F_Y(y)$ 称为 $F_{XY}(x, y)$ 的边际累积分布函数 (marginal cumulative distribution function)。

性质：

- (1) $0 \leq F_{XY}(x, y) \leq 1$;
- (2) $\lim_{x, y \rightarrow \infty} F_{XY}(x, y) = 1$;
- (3) $\lim_{x, y \rightarrow -\infty} F_{XY}(x, y) = 0$;
- (4) $F_X(x) = \lim_{y \rightarrow \infty} F_{XY}(x, y)$ 。

1.3.2 联合与边际概率质量函数

如果 X 和 Y 是离散随机变量，则联合概率质量函数 $p_{XY} : \mathbb{R} \times \mathbb{R} \rightarrow [0, 1]$ 定义如下：

$$p_{XY}(x, y) = P(X = x, Y = y)$$

其中， $0 \leq p_{XY}(x, y) \leq 1$ 对所有 x, y 成立，且 $\sum_{x \in \text{Val}(X)} \sum_{y \in \text{Val}(Y)} p_{XY}(x, y) = 1$ 。

联合概率质量函数和每个随机变量概率质量函数的关系为

$$p_X(x) = \sum_y p_{XY}(x, y), \quad p_Y(y) = \sum_x p_{XY}(x, y)$$

将 $p_X(x)$ 和 $p_Y(y)$ 称为边际概率质量函数。

1.3.3 联合与边际概率密度函数

如果 X 和 Y 是连续随机变量，则联合累积分布函数为 F_{XY} 。这意味着 $F_{XY}(x, y)$ 是关于 x 和 y 处处可微的。为此，可以定义联合概率密度函数：

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y}$$

与一元类似， $f_{XY}(x, y) \neq P(X = x, Y = y)$ ，但

$$\iint_{x \in A} f_{XY}(x, y) dx dy = P((X, Y) \in A)$$

特别需要说明的是 $f_{XY}(x, y)$ 的取值总是非负的，且可以高于 1。但无论如何，都满足 $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) dx dy = 1$ 。

与离散情形类似，我们定义

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$$

为 X 的边际概率密度函数。 $f_Y(y)$ 的定义与之类似。

1.3.4 条件概率

当一个随机变量 X 必须取特定值 x 时, 条件概率刻画的是 Y 的概率分布。在离散情形下, 给定 X 下的 Y 的概率质量函数为

$$p_{Y|X}(y | x) = \frac{p_{XY}(x, y)}{p_X(x)}$$

需要说明的是, 为了保证刻画有意义, 这里假定 $p_X(x) \neq 0$ 。

在连续情形下, 假设 $f_X(x) \neq 0$, 给定 $X = x$ 下的 Y 的概率密度函数为

$$f_{Y|X}(y | x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

需要说明的是, 连续情形的推导更为复杂, 因为 $P\{X = x\}$ 的取值永远为 0^①。

1.3.5 贝叶斯法则

贝叶斯公式在计算条件概率时非常有用, 其表现形式如下。

对于离散随机变量 X 和 Y ,

$$P_{Y|X}(y | x) = \frac{P_{XY}(x, y)}{P_X(x)} = \frac{P_{X|Y}(x | y)P_Y(y)}{\sum_{y' \in \text{Val}(Y)} P_{X|Y}(x | y')P_Y(y')}$$

对于连续随机变量 X 和 Y ,

$$f_{Y|X}(y | x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{f_{X|Y}(x | y)f_Y(y)}{\int_{-\infty}^{\infty} f_{X|Y}(x | y')f_Y(y') dy'}$$

1.3.6 独立性

如果 $F_{XY}(x, y) = F_X(x)F_Y(y)$ 对所有 x 和 y 都成立, 则随机变量 X 和 Y 相互独立。

与之等价, 我们有如下结论:

① 为了近似计算, 一个合理的方式是计算条件累积分布函数,

$$F_{Y|X}(y, x) = \lim_{\Delta x \rightarrow 0} P(Y \leq y | x \leq X \leq x + \Delta x)$$

因为 $F(x, y)$ 对于 x, y 处处可导, 可得

$$F_{Y|X}(y, x) = \int_{-\infty}^y \frac{f_{X,Y}(x, \alpha)}{f_X(x)} d\alpha$$

两边同时对 y 求导, 可得如下等式, 并将其作为条件概率密度函数的定义:

$$f_{Y|X}(y | x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

(1) 对于离散随机变量, $p_{XY}(x, y) = p_X(x)p_Y(y)$ 对所有 $x \in \text{Val}(X), y \in \text{Val}(Y)$ 成立;

(2) 对于离散随机变量, 当 $p_X(x) \neq 0$ 时, $p_{Y|X}(y | x) = p_Y(y)$ 对所有 $y \in \text{Val}(Y)$ 成立;

(3) 对于连续随机变量, $f_{XY}(x, y) = f_X(x)f_Y(y)$ 对所有 $x, y \in \mathbb{R}$ 成立;

(4) 对于离散随机变量, 当 $f_X(x) \neq 0$ 时, $f_{Y|X}(y | x) = f_Y(y)$ 对所有 $y \in \mathbb{R}$ 成立。
独立性最重要的结论如下。

引理 1.1 如果 X 和 Y 相互独立, 对任意子集 $A, B \subseteq \mathbb{R}$, 有

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

1.3.7 期望与协方差

假设 X, Y 是两个离散随机变量, 且 $g: \mathbf{R}^2 \rightarrow \mathbf{R}$ 是这两个随机变量的函数, 则 g 的期望定义为

$$E[g(X, Y)] \triangleq \sum_{x \in \text{Val}(X)} \sum_{y \in \text{Val}(Y)} g(x, y)p_{XY}(x, y)$$

如果 X, Y 是两个连续随机变量, 则 g 的期望定义为

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

我们利用期望的概念, 研究两个随机变量之间的关系, 尤其是两个随机变量的协方差, 其定义如下:

$$\text{Cov}(X, Y) \triangleq E[(X - E(X))(Y - E(Y))]$$

进一步可改写为

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E[XY - XE(Y) - YE(X) + E(X)E(Y)] \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

当 $\text{Cov}[X, Y] = 0$ 时, 我们说 X 和 Y 不相关^①。

性质:

- (1) 期望的线性性, $E[f(X, Y) + g(X, Y)] = E[f(X, Y)] + E[g(X, Y)]$;
- (2) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$;
- (3) 如果 X 与 Y 相互独立, 则 $\text{Cov}(X, Y) = 0$;
- (4) 如果 X 与 Y 相互独立, 则 $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$ 。

^① 相关性与独立性是两个不同的概念。例如, 如果 $X \sim \text{Uniform}(-1, 1)$, $Y = X^2$, 则可证明 X 与 Y 不相关, 尽管它们不独立。

1.4 多元随机变量

利用 1.3 节的记号和思路可以将二元随机变量推广至多元, 尤其是 n 元连续随机变量 $X_1(\omega), X_2(\omega), \dots, X_n(\omega)$ 。为了描述简便, 本节介绍只针对连续随机变量, 但相同的方法和思路可以适用于离散随机变量。

1.4.1 基础性质

我们定义 $X_1, X_2, \dots, X_n$ 的联合累积分布函数, $X_1, X_2, \dots, X_n$ 的联合概率密度函数, X_1 的边际概率密度函数和给定 $X_2, \dots, X_n$ 下 X_1 的条件概率密度函数分别为

$$\begin{aligned} F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) &= P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \\ f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) &= \frac{\partial^n F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n} \\ f_{X_1}(X_1) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) dx_2 \cdots dx_n \\ f_{X_1|X_2, \dots, X_n}(x_1 | x_2, \dots, x_n) &= \frac{f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)}{f_{X_2, \dots, X_n}(x_2, \dots, x_n)} \end{aligned}$$

由此, 计算任意事件 $A \subseteq \mathbb{R}^n$ 的概率的公式为

$$P((x_1, x_2, \dots, x_n) \in A) = \int_{(x_1, x_2, \dots, x_n) \in A} f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n$$

链式法则: 从多元随机变量条件概率的定义可以证明

$$\begin{aligned} f(x_1, x_2, \dots, x_n) &= f(x_n | x_1, x_2, \dots, x_{n-1}) f(x_1, x_2, \dots, x_{n-1}) \\ &= f(x_n | x_1, x_2, \dots, x_{n-1}) f(x_{n-1} | x_1, x_2, \dots, x_{n-2}) f(x_1, x_2, \dots, x_{n-2}) \\ &= \cdots = f(x_1) \prod_{i=2}^n f(x_i | x_1, \dots, x_{i-1}) \end{aligned}$$

独立性: 对于多个事件 $A_1, A_2, \dots, A_k$, 我们称 $A_1, A_2, \dots, A_k$ 是独立的, 对任意子集 $S \subseteq \{1, 2, \dots, k\}$, 下式成立:

$$P(\cap_{i \in S} A_i) = \prod_{i \in S} P(A_i)$$

类似于二元随机变量的独立性, 如果下式成立:

$$f(x_1, x_2, \dots, x_n) = f(x_1) f(x_2) \cdots f(x_n)$$

则称随机变量 $X_1, X_2, \dots, X_n$ 相互独立。

1.4.2 随机向量

对于 n 元随机变量, 向量化的表示更方便。因此, 随机向量可表示为 $\mathbf{X} = [X_1, X_2, \dots, X_n]^T$ 。随机向量的正式定义为, 它是一个从 Ω 到 $\mathbb{R}^n$ 的映射。基于向量化的刻

画方式, 原有 n 元随机变量的联合概率密度函数和联合累积分布函数都会相应变化。

期望: 对于任一函数 $g: \mathbb{R}^n \rightarrow \mathbb{R}$, 其期望值为

$$E[g(\mathbf{X})] = \int_{\mathbb{R}^n} g(x_1, x_2, \dots, x_n) f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n$$

如果函数 g 是从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的映射, 即

$$g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix}$$

则

$$E[g(\mathbf{X})] = \begin{pmatrix} E[g_1(\mathbf{X})] \\ E[g_2(\mathbf{X})] \\ \vdots \\ E[g_m(\mathbf{X})] \end{pmatrix}$$

协方差阵: 对于一个随机向量 $\mathbf{X}: \Omega \rightarrow \mathbb{R}^n$, 其协方差阵 Σ 是 $n \times n$ 方阵, 并且所有元素为 $\Sigma_{ij} = \text{Cov}(X_i, X_j)$ 。由此定义, 可得

$$\begin{aligned} \Sigma &= \begin{pmatrix} \text{Cov}(X_1, X_1) & \cdots & \text{Cov}(X_1, X_n) \\ \vdots & & \vdots \\ \text{Cov}(X_n, X_1) & \cdots & \text{Cov}(X_n, X_n) \end{pmatrix} \\ &= \begin{pmatrix} E(X_1^2) - E(X_1)E(X_1) & \cdots & E(X_1 X_n) - E(X_1)E(X_n) \\ \vdots & & \vdots \\ E(X_n X_1) - E(X_n)E(X_1) & \cdots & E(X_n^2) - E(X_n)E(X_n) \end{pmatrix} \\ &= \begin{pmatrix} E(X_1^2) & \cdots & E(X_1 X_n) \\ \vdots & & \vdots \\ E(X_n X_1) & \cdots & E(X_n^2) \end{pmatrix} - \begin{pmatrix} E(X_1)E(X_1) & \cdots & E(X_1)E(X_n) \\ \vdots & & \vdots \\ E(X_n)E(X_1) & \cdots & E(X_n)E(X_n) \end{pmatrix} \\ &= E(\mathbf{X}\mathbf{X}^T) - E(\mathbf{X})E(\mathbf{X})^T = \cdots = E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^T] \end{aligned}$$

协方差阵的性质:

- (1) $\Sigma \succeq 0$, 即 Σ 是半正定的;
- (2) $\Sigma = \Sigma^T$, 即 Σ 是对称的。

1.4.3 多元高斯分布

最为重要的一类多元概率分布是多元正态分布 (高斯分布)。一个随机向量 $\mathbf{X} \in \mathbb{R}^n$ 服从正态分布且均值为 $\mu \in \mathbb{R}^n$, 协方差阵为 $\Sigma \in \mathbb{S}_{++}^n$ (其中 $\mathbb{S}_{++}^n$ 表示对称正定的 $n \times n$ 矩阵的空间)。该分布的密度函数为

其记号为 $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 。值得注意的是, 当 $n = 1$ 时, 该密度函数退化为一元正态分布的定义, 其中均值为 μ_1 , 方差为 Σ_{11} 。

一元分布关系图如图 1.1 所示。

图 1.1 一元分布关系图

$$\text{标准差} = \sqrt{p(1-p)}$$