概率论与数理统计

学习笔记

作者：夏同

组织：华南理工大学

时间：April 2, 2026

版本：1.0

自定义：信息

ElegantLaTeX Program

居敬持志，守正出奇。——张红兵

目录

第 1 章　随机事件与概率 **1**

1.1　基本概念 2

1.2　事件的关系与运算 3

1.2.1　事件的关系 3

1.2.2　事件的运算 3

1.2.3　事件的运算律 3

1.3　概率的定义 6

1.3.1　描述性定义 6

1.3.2　统计性定义 6

1.4　古典概型 9

1.4.1　几何概型 13

1.5　条件概率 14

1.5.1　条件概率的定义 14

1.5.2　乘法公式 14

1.5.3　全概率公式与贝叶斯公式 15

1.6　事件的独立性 21

1.6.1　独立性的定义 21

1.6.2　独立性的判定 22

1.7　独立重复试验与伯努利概型 24

第 2 章　一维随机变量及其分布 **26**

2.1　随机变量与分布函数 27

2.1.1　随机变量 27

2.1.2　分布函数 27

2.1.3　分布函数的性质 27

2.1.4　分布函数与事件概率 28

2.2　离散型随机变量 29

2.3　连续型随机变量 30

2.4　常见离散型随机变量及分布 34

2.4.1　(0-1) 分布 34

2.4.2　二项分布 35

2.4.3　泊松分布 36

2.4.4　几何分布 37

2.4.5　超几何分布 38

2.4.6　极值分布 38

2.5　常见连续型随机变量及分布 39

2.5.1 均匀分布 39
2.5.2 指数分布 40
2.5.3 正态分布 42
2.6 常见分布的期望与方差 44
2.7 典型例题 44
2.8 一维随机变量函数的分布 49
2.8.1 基本概念 49
2.8.2 离散型随机变量函数的分布 49
2.8.3 连续型随机变量函数的分布 50
2.8.3.1 当 Y 为离散型时 50
2.8.3.2 当 Y 为连续型时 50
2.8.4 一般情形 51
2.9 典型例题 52

第 3 章 二维随机变量及分布 58
3.1 二维随机变量 59
3.1.1 联合分布函数 59
3.2 二维离散型随机变量 62
3.2.1 联合分布律 62
3.2.2 边缘分布律 63
3.2.3 条件分布律 63
3.3 二维连续型随机变量 70
3.3.1 边缘概率密度 71
3.3.2 条件概率密度 71
3.4 常见的二维分布 78
3.4.1 二维均匀分布 78
3.4.2 二维正态分布 79
3.5 n 维随机变量 79
3.6 随机变量的独立性 80
3.6.1 独立性的定义 80
3.6.2 独立的充要条件 81
3.6.2.1 离散型随机变量的独立性 81
3.6.2.2 连续型随机变量的独立性 82
3.6.3 独立性的性质 82
3.6.4 不独立的判断与证明 83
3.7 二维随机变量函数的分布 84
3.7.1 离散型随机变量函数的分布 84
3.7.2 连续型随机变量函数的分布 85
3.7.2.1 分布函数法 85
3.7.2.2 卷积公式与常用分布函数公式 86

3.7.2.3 max 和 min 的分布 89

第 4 章 随机变量的数字特征 96

4.1 数学期望 97

4.1.1 数学期望的定义 97

4.1.2 二维随机变量函数的数学期望 101

4.1.3 数学期望的性质 101

4.2 方差 104

4.2.1 方差的定义 104

4.2.2 方差的性质 105

4.3 常见分布的数学期望与方差 107

4.4 协方差与相关系数 109

4.4.1 协方差 109

4.4.2 相关系数 112

4.4.3 不相关性的判定 115

4.4.4 正态分布的数字特征 117

第 5 章 大数定律和中心极限定理 119

5.1 依概率收敛 119

5.2 切比雪夫不等式 120

5.3 大数定律 120

5.3.1 切比雪夫大数定律 122

5.3.2 伯努利大数定律 122

5.3.3 辛钦大数定律 123

5.4 中心极限定理 124

5.4.1 列维–林德伯格中心极限定理 124

5.4.2 棣莫弗–拉普拉斯中心极限定理 127

第 6 章 样本与抽样分布 130

6.1 基本概念 131

6.1.1 总体与样本 131

6.1.2 统计量 131

6.2 三大抽样分布 133

6.2.1 χ^2 分布 133

6.2.2 t 分布 133

6.2.3 F 分布 134

6.3 正态总体下的抽样分布 135

6.3.1 单个正态总体 135

6.3.2 两个正态总体 138

6.4 分布的分位数 138

6.5 综合练习 139

第 7 章 点估计 **141**
7.1 点估计的基本概念 141
7.2 矩估计法（MME） 142
7.3 最大似然估计法（MLE） 142
7.4 估计量的评价标准 147

第 8 章 区间估计 **156**
8.1 区间估计的基本概念 156
8.2 构造置信区间的基本原理 157
8.3 常见的参数置信区间总结 159
8.4 典型例题 160

第 9 章 假设检验 **163**
9.1 基本概念 163
9.2 假设检验的步骤 164
9.3 正态总体参数的假设检验 164
9.3.1 单个正态总体均值的检验 164
9.3.2 两个正态总体均值差的检验 165
9.3.3 单个正态总体方差的检验 166
9.3.4 两个正态总体方差比的检验 166
9.4 置信区间与假设检验的关系 167
9.5 两类错误 168

第 1 章 随机事件与概率

- 基本概念
  - 随机试验 E : 可重复性、结果不唯一、所有结果已知、不确定性
  - 样本空间 S : 所有可能结果构成的集合
  - 随机事件 A : 样本空间的子集
    - 关系：包含、相等、互斥、对立
    - 表示：和事件、积事件、差事件、逆 (对立) 事件
    - 运算：交换律、结合律、分配律、德摩根根律
- 概率概念
  - 统计概念：频率
  - 公理化概念：非负性、规范性和可列可加性
  - 古典模型 (等可能概型)：有限元素、可能性相同
  - 计算公式
    - 加法公式
      - $P(A\cup B)=P(A)+P(B)-P(AB)$
      - $P(A\cup B)=P(A)+P(B)$ (若事件 A 与 B 互斥)
    - 减法公式
      - $P(A-B)=P(A)-P(AB)=P(A\overline{B})$
      - $P(A-B)=P(A)-P(B)$ (若事件 $A\supset B$)
    - 求逆公式： $P(\overline{A})=1-P(A)$
- 条件概率概念
  - 定义：当 $P(A)>0$ 时, $P(B|A)=\dfrac{P(AB)}{P(A)}$, 且满足概率性质与上述公式 (非负性、规范性和可列可加性)
  - 计算: 定义法，缩小样本空间方法
  - 应用
    - 乘法公式：当 $P(A)>0$ 时, $P(AB)=P(B|A)P(A)$
    - 全概率公式 (由因求果)： $P(A)=\sum_{i=1}^{n}P(A|B_i)P(B_i)$
    - 贝叶斯公式 (执果寻因)： $P(B_i|A)=\dfrac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n}P(A|B_j)P(B_j)}$ (其中 $B_1,B_2,\ldots,B_n$ 是一组完备事件组, $P(B_i)>0$)
- 独立性
  - 定义
    - 两个事件: $P(AB)=P(A)P(B)\Leftrightarrow$ 当 $0<P(A)<1, P(B|A)=P(B|\overline{A})$
    - 多个事件：相互独立 $\Rightarrow$ 两两独立, 两两独立 $\not\Rightarrow$ 相互独立
  - 性质
    - 相互独立的事件中部分事件间相互独立
    - 相互独立的事件中部分事件的逆事件与其余事件仍相互独立
  - 应用
    - 计算多个事件的和事件概率转化为逆事件概率乘积
    - 伯努利试验：只有两个结果 A 与 $\overline{A}$, n 次试验相互独立

内容提要

- 随机试验与样本空间 1.1
- 随机事件及其运算 1.3
- 完备事件组 1.4
- 概率的公理化定义 1.5
- 概率的基本性质 1.3.2
- 古典概型 1.6
- 几何概型 1.7
- 条件概率 1.8
- 乘法公式 1.1
- 全概率公式与贝叶斯公式 1.2
- 事件的独立性 1.9
- 伯努利概型 1.11

1.1 基本概念

定义 1.1 (随机试验)

如果一个试验满足以下三个条件：

1. 试验可以在相同的条件下重复进行；
2. 试验所有可能结果明确可知，且不止一个；
3. 每一次试验会出现哪一个结果，事先并不能确定，

则称此试验为随机试验。随机试验简称试验，用字母 E 或 $E_1, E_2, \ldots$ 表示。 ♣

我们是通过研究随机试验来研究随机现象的。

笔记 在不少情况下，不能确切知道某一随机试验的全部可能结果，但可以知道它不超出某个范围。这时，也可以用这个范围来作为该试验的全部可能结果的集合。例如，需要记录某个城市一天的交通事故数量，则试验结果将是非负整数 x。无法确定 x 的可能取值的确切范围，但可以把这个范围取为 $[0, +\infty)$，它总能包含一切可能的试验结果，尽管明知某些结果，如 $x > 10000$ 是不会出现的。甚至把这个范围取为 $(-\infty, +\infty)$ 也无妨。这里体现了一定的数学抽象，它可以带来很大的方便。

定义 1.2 (样本空间)

随机试验的每一个可能结果称为样本点，记为 ω。样本点的全体组成的集合称为样本空间（或基本事件空间），记为 Ω（或 S），即 $\Omega = \{\omega\}$。 ♣

由一个样本点构成的事件称为**基本事件**。

定义 1.3 (随机事件)

在一次试验中可能出现，也可能不出现的结果称为随机事件，简称事件，用大写字母 A, B, C 等表示。随机事件是样本空间 Ω 的子集。在每次试验中，当且仅当这一子集中的一个样本点出现时，称这一事件发生。 ♣

1. 对于仅由一个基本结果构成的单点集，称为**基本事件**；
2. 将每次试验中一定发生的事件称为**必然事件**，记为 Ω；
3. 每次试验中一定不发生的事件称为**不可能事件**，记为 $\varnothing$。

笔记 概念之间的联系：随机试验产生基本结果，基本结果 $\in$ 随机事件 $\subset$ 样本空间。随机事件在一次试验中是否发生虽然不能事先确定，但是在大量重复试验的情况下，它的发生呈现出一定的规律性，这门课程正是要研究这种规律性。

1.2 事件的关系与运算

事件是一个集合，因此事件之间的关系和运算与集合论中集合之间的关系和运算完全类似，可以利用集合论的知识来理解事件的关系与运算。

1.2.1 事件的关系

1. **包含**——如果事件 A 发生必导致事件 B 发生，则称事件 B 包含事件 A（或 A 被 B 包含），记为 $A \subset B$。也称 A 是 B 的子事件。
2. **相等**——如果 $A \subset B$ 且 $B \subset A$，则称事件 A 与 B 相等，记为 $A = B$。A 与 B 相等，事实上也就是说，A 与 B 由一些完全相同的试验结果构成，它不过是同一事件表面上看起来不同的两个说法而已。
3. **相容**——若 $AB \neq \varnothing$，则称事件 A 和 B 相容。
4. **互斥（互不相容）**——若 $AB = \varnothing$，则称事件 A 与 B 互不相容，也叫互斥。如果一些事件中任意两个事件都互斥，则称这些事件是两两互斥的。任意一对基本事件都是互斥的。
5. **对立**——事件 A, B 不能同时发生，且至少一个发生，称事件 A, B 互为对立事件，或互为逆事件，即 $A \cap B = \varnothing$，$A \cup B = \Omega$。

定义 1.4 (完备事件组)

如果 $\bigcup_{i=1}^{n} A_i = \Omega$（或 $\bigcup_{i=1}^{\infty} A_i = \Omega$），且 $A_i A_j = \varnothing$（对一切 $i \neq j$；$i, j = 1, 2, \ldots, n(\ldots)$），称有限个（或可列个）事件 $A_1, A_2, \ldots, A_n(\ldots)$ 构成一个完备事件组。 ♣

完备事件组在概率论中具有重要作用，全概率公式和贝叶斯公式都需要完备事件组作为前提条件。

笔记 互斥与对立的区别：互斥只要求 $AB = \varnothing$，而对立还要求 $A \cup B = \Omega$。因此对立一定是互斥的，但互斥不一定是对立的。

1.2.2 事件的运算

1. **积事件（交）**——称" 事件 A 与 B 同时发生" 的事件为事件 A 与 B 的**积事件**（或交事件），记为 $A \cap B$ 或 AB。称" 有限个（或可列个）事件 $A_1, A_2, \ldots, A_n(\ldots)$ 同时发生" 的事件为这些事件的积事件，记为 $\bigcap_{i=1}^{n} A_i$（或 $\bigcap_{i=1}^{\infty} A_i$）。
2. **和事件（并）**——称" 事件 A 与 B 至少有一个发生" 的事件为事件 A 与 B 的**和事件**（或并事件），记为 $A \cup B$。称" 有限个（或可列个）事件 $A_1, A_2, \ldots, A_n(\ldots)$ 至少有一个发生" 的事件为这些事件的和事件，记为 $\bigcup_{i=1}^{n} A_i$（或 $\bigcup_{i=1}^{\infty} A_i$）。
3. **差事件**——称" 事件 A 发生而事件 B 不发生" 的事件为事件 A 与 B 的**差事件**，记为 $A - B$。
4. **逆事件（对立事件）**——称" 事件 A 不发生" 的事件为事件 A 的**逆事件**或**对立事件**，记为 $\overline{A}$。

   由定义易知

$$A - B = A - AB = A\overline{B}, \quad B = \overline{A} \Leftrightarrow AB = \varnothing \text{ 且} A \cup B = \Omega.$$

1.2.3 事件的运算律

1. **吸收律**——若 $A \subset B$，则 $A \cup B = B$，$A \cap B = A$。

2. **交换律**——$A\cup B=B\cup A$，$A\cap B=B\cap A$。
3. **结合律**——$(A\cup B)\cup C=A\cup(B\cup C)$，$(A\cap B)\cap C=A\cap(B\cap C)$。
4. **分配律**——$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$，$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$。特别地，

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

5. **对偶律（德·摩根律）**——$\overline{A\cup B}=\overline{A}\cap\overline{B}$，$\overline{A\cap B}=\overline{A}\cup\overline{B}$。一般地，对于 $n\geq 1$，

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i};\quad \overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}.$$

笔记 德摩根律记忆技巧：长线变短线，开口换方向。事件运算顺序约定为先进行逆运算，然后进行交运算，最后进行并或差运算。事件的关系、运算与集合的关系、运算相当，且具有相同的运算法则，所以我们可以对比着理解记忆，并要学会用集合关系去考虑事件关系。

练习 1.1 已知 A、B、C 为三个随机事件，则 A、B、C 不都发生的事件为（　）。

A. $\overline{ABC}$　B. $\overline{ABC}$　C. $A\cup B\cup C$　D. ABC

解 "A、B、C 不都发生" 的对立事件为"A、B、C 都发生"，即 ABC，故"A、B、C 不都发生" 为 $\overline{ABC}$。答案选 B。

练习 1.2 设有随机事件 $A_i\ (i=1,2,\ldots,n)$，$A=\bigcap_{i=1}^{n}A_i$，则 A 的对立事件 $\overline{A}=$ _____。

解 根据德摩根律，

$$\overline{A}=\overline{\bigcap_{i=1}^{n}A_i}=\overline{A_1\cap A_2\cap\cdots\cap A_n}=\overline{A_1}\cup\overline{A_2}\cup\cdots\cup\overline{A_n}=\bigcup_{i=1}^{n}\overline{A_i}.$$

练习 1.3 设 A,B,C 是任意三个事件，则下列选项中正确的是（　）。

A. 若 $A\cup C=B\cup C$，则 $A=B$　B. 若 $A-C=B-C$，则 $A=B$　C. 若 $AC=BC$，则 $A=B$　D. 若 $AB=\varnothing$ 且 $\overline{A}\,\overline{B}=\varnothing$，则 $\overline{A}=B$

解 应选 D。

方法一（直接法）：$\overline{A}\,\overline{B}=\overline{A\cup B}=\varnothing$，故 $A\cup B=\Omega$。又 $AB=\varnothing$，因此 A,B 互为对立事件，即 $\overline{A}=B$。

方法二（排除法）：取 $C=\Omega$，则 $A\cup C=B\cup C=\Omega$ 恒成立，但 $A\neq B$ 仍可成立，故 (A) 错。取 $C=\varnothing$，则 $AC=BC=\varnothing$ 恒成立，故 (C) 错。(B) 类似取 $C=\Omega$ 即可排除。

笔记 本题反映了事件运算与数的运算的不同之处。事件的并、交、差运算不满足消去律。

练习 1.4 用事件 A,B,C 的运算关系表示下列事件：

(1) A,B,C 都不发生；

(2) A,B,C 不都发生；

(3) A,B,C 不多于一个发生。

解 (1) A,B,C 都不发生 = 三者均不发生：

$$\overline{A}\,\overline{B}\,\overline{C}=\overline{A\cup B\cup C}.$$

(2) A,B,C 不都发生 = " 都发生" 的逆事件：

$$\overline{ABC}=\overline{A}\cup\overline{B}\cup\overline{C}.$$

(3) A,B,C 不多于一个发生 = 至多一个发生 = 至少有两个不发生：

$$\overline{A}\,\overline{B}\cup\overline{A}\,\overline{C}\cup\overline{B}\,\overline{C}=A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C\cup\overline{A}\,\overline{B}\,\overline{C}.$$

练习 1.5 设随机变量 X, Y，事件 $A = \{X \geq c\}$，事件 $B = \{Y \geq c\}$。用 A, B 表示下列事件：

$D = \{\min(X, Y) < c\}$； $E = \{\max(X, Y) \geq c\}$； $F = \{\min(X, Y) \geq c\}$； $G = \{\max(X, Y) < c\}$。

解

- $D = \overline{A} \cup \overline{B}$（$X, Y$ 中至少有一个 $< c$）；
- $E = A \cup B$（X, Y 中至少有一个 $\geq c$）；
- $F = A \cap B$（X, Y 均需 $\geq c$）；
- $G = \overline{A} \cap \overline{B} = \overline{A \cup B}$（$X, Y$ 均需 $< c$）。

练习 1.6 设 A 和 B 是任意两个事件，则下列两个命题（　）。

(1) 若 $P(A) = P(B)$，则 $A = B$；

(2) 若 $P(AB) = 0$，则 $AB = \varnothing$。

A. (1) 正确，(2) 不正确　B. (1) 不正确，(2) 正确　C. (1)、(2) 均正确　D. (1)、(2) 均不正确

解 应选 D。

反例：向 $[0, 2]$ 上随机投点。令 $A = \{点落入[0, 1]\}$，$B = \{点落入[1, 2]\}$。则 $P(A) = P(B) = 0.5$，但 $A \neq B$；$P(AB) = P(\{1\}) = 0$，但 $AB = \{1\} \neq \varnothing$。故两个命题均不正确。

笔记 $P(A) = 0$ 不能推出 $A = \varnothing$，$P(A) = 1$ 不能推出 $A = \Omega$。概率只反映可能性大小，不反映事件的集合构成。

练习 1.7 已知事件 A, B 满足 $P(AB) = P(\overline{A}\,\overline{B})$，记 $P(A) = p$，则 $P(B) =$ _____。

解 应填 $1 - p$。

由德摩根律和加法公式：

$$P(\overline{A}\,\overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - P(A) - P(B) + P(AB),$$

代入 $P(AB) = P(\overline{A}\,\overline{B})$，得

$$1 - P(A) - P(B) = 0 \implies P(B) = 1 - P(A) = 1 - p.$$

练习 1.8 已知 $P(A) = 0.7$，$P(A - B) = 0.3$，则 $P(\overline{AB}) =$ _____。

解 应填 0.6。

由减法公式 $P(A - B) = P(A) - P(AB)$，得 $P(AB) = 0.7 - 0.3 = 0.4$，故

$$P(\overline{AB}) = 1 - P(AB) = 0.6.$$

练习 1.9 设 A, B 为两个随机事件，且 $P(A) = 0.4$，$P(A \cup B) = 0.7$。若 A, B 互不相容，则 $P(B) =$ _____；若 A, B 相互独立，则 $P(B) =$ _____；若 A 发生则 B 必发生，则 $P(B) =$ _____。

解 应填 0.3；0.5；0.7。

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

- 互不相容时 $P(AB) = 0$：$P(B) = 0.7 - 0.4 = 0.3$；
- 相互独立时 $P(AB) = P(A)\,P(B)$：$0.7 = 0.4 + P(B) - 0.4P(B)$，$P(B) = \dfrac{0.3}{0.6} = 0.5$；
- A 发生 B 必发生即 $A \subset B$：$P(B) = P(A \cup B) = 0.7$。

1.3 概率的定义

1.3.1 描述性定义

通常将随机事件 A 发生的可能性大小的度量（非负值）称为事件 A 发生的**概率**，记为 $P(A)$。

1.3.2 统计性定义

在相同条件下做重复试验，事件 A 出现的次数 k 和总的试验次数 n 之比 $\frac{k}{n}$ 称为事件 A 在这 n 次试验中出现的**频率**。当试验次数 n 充分大时，频率将" 稳定" 于某常数 p。n 越大，频率偏离这个常数 p 的可能性越小。这个常数 p 就称为事件 A 的**概率**。

笔记 概率的统计性定义实质上是说，用频率 $\frac{k}{n}$ 作为事件 A 的概率 $P(A)$ 的估计。其直观理解为某事件出现的可能性大小，可由其在多次重复试验中出现的频率去刻画。从上述可以看出，频率只是概率的估计，而非概率本身。概率的统计性定义的重要性主要基于以下两点：

1. 它提供了估计概率的方法。比如在一批产品中抽取样品，来估计该批产品的合格率（合格率是客观的数据，抽取样品计算出来的合格率只是一种估计）。
2. 它提供了一种检验某结论是否正确的准则。比如，你说某批产品的合格率是 95%，我们做试验，抽取样品进行计算，得出的结果是合格率为 20%，远远低于你所说的 95%，于是毫不犹豫地拒绝你的结论。

公理化定义

定义 1.5 (概率的公理化定义)

设随机试验的样本空间为 Ω，如果对每一个事件 A 都有一个确定的实数 $P(A)$，且事件函数 $P(\cdot)$ 满足：

1. 非负性：$P(A) \geq 0$；
2. 规范性：$P(\Omega) = 1$；
3. 可列可加性：对任意可列个两两互不相容事件 $A_1, A_2, \ldots, A_n, \ldots$（即 $A_iA_j = \varnothing$，$i \neq j$；$i, j = 1, 2, \ldots$），有
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),$$

则称 $P(\cdot)$ 为概率，$P(A)$ 为事件 A 的概率。 ♣

注 (1) 数学上所说的" 公理"，就是一些不加证明而承认的前提，上述公理化定义只是界定了概率这个概念所必须满足的一般性质，它不解决具体场合下的概率计算。

(2) 概率 $P(\cdot)$ 是事件的函数。

(3) 虽然公理化定义不解决具体场合下的概率计算，但是我们却经常用它来判断事件函数 $P(\cdot)$ 是否是概率，这种题型在考研试题中也会经常遇到。

由概率的公理化定义，可以推导出以下基本性质：

性质 概率的基本性质：

1. $P(\varnothing) = 0$；

2. 有限可加性：若 $A_1, A_2, \ldots, A_n$ 两两互不相容，则
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i);$$
3. 减法公式：若 $B \subset A$，则 $P(A-B) = P(A) - P(B)$。更一般地，
$$P(A-B) = P(A) - P(AB);$$
4. 单调性：若 $B \subset A$，则 $P(B) \le P(A)$；
5. 有界性：对任意事件 A，$0 \le P(A) \le 1$；
6. 求逆公式：$P(\overline{A}) = 1 - P(A)$；
7. 加法公式：$P(A \cup B) = P(A) + P(B) - P(AB)$。当 A、B 互斥时，$P(A \cup B) = P(A) + P(B)$。
8. 设 A_1, A_2, A_3 为任意三个事件，则有
$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3) + P(A_1A_2A_3).$$
9. 若 $A_1, A_2, \ldots, A_n$ 是两两互不相容的事件，则有
$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

笔记 概率为 0 并不意味着为空集。即 $P(A) = 0 \not\Rightarrow A = \varnothing$，而 $A = \varnothing \Rightarrow P(A) = 0$。同样的，$P(A) = 0$ 不能断言 $A = \varnothing$，而 $P(A) = 1$ 不能断言 $A = \Omega$.

练习 1.10 设 A, B 为两个随机事件，$P(A) = 0.4$，$P(A \cup B) = 0.7$。

(1) 求 $P(B - A)$。

(2) 若事件 A 和事件 B 互不相容，求 $P(B)$。

解 (1) 由加法公式
$$P(A \cup B) = P(A) + P(B) - P(AB) = P(A) + P(B - A),$$
故
$$P(B - A) = P(A \cup B) - P(A) = 0.7 - 0.4 = 0.3.$$

(2) 若 A 与 B 互不相容，则 $P(AB) = 0$，此时加法公式简化为
$$P(A \cup B) = P(A) + P(B),$$
故
$$P(B) = P(A \cup B) - P(A) = 0.7 - 0.4 = 0.3.$$

练习 1.11 设 A 和 B 是任意两个概率不为 0 的互不相容事件，则下列结论中肯定正确的是（　）。

A. $\overline{A}$ 与 $\overline{B}$ 不相容　B. $\overline{A}$ 与 $\overline{B}$ 相容　C. $P(AB) = P(A)P(B)$　D. $P(A - B) = P(A)$

解 应选 D。

由 A 与 B 互不相容，$P(AB) = 0$，故
$$P(A - B) = P(A) - P(AB) = P(A).$$

当 A 与 B 互为对立事件时，$\overline{A}$ 与 $\overline{B}$ 不相容；反之则相容。故选项 A、B 不一定正确。选项 C：$P(AB) = 0 \ne P(A)P(B)$（因为 $P(A) > 0, P(B) > 0$），故错误。

练习 1.12 考虑一元二次方程 $x^2 + Bx + C = 0$，其中 B, C 分别是将一枚骰子接连抛两次先后出现的点数。则该方程有实根的概率为（　）。

A. $\dfrac{19}{36}$　B. $\dfrac{17}{36}$　C. $\dfrac{15}{36}$　D. $\dfrac{13}{36}$

解 应选 A。

方程有实根的条件为 $B^2-4C\geq 0$。总样本数 $6^2=36$。列表统计满足条件的样本点数：

$B\setminus C$	1	2	3	4	5	6
1	−	−	−	−	−	−
2	0	−	−	−	−	−
3	+	+	−	−	−	−
4	+	+	+	0	−	−
5	+	+	+	+	+	+
6	+	+	+	+	+	+

"+" 和 "0" 对应的样本点共 $1+2+4+6+6=19$ 个，故 $P=\dfrac{19}{36}$。

练习 1.13 甲口袋有 5 个白球、3 个黑球，乙口袋有 4 个白球、6 个黑球。从两个口袋中各任取一个球，则取到的两个球颜色相同的概率为（　）。

A. $\dfrac{19}{40}$　B. $\dfrac{21}{40}$　C. $\dfrac{1}{4}$　D. $\dfrac{9}{40}$

解 应选 A。

两个球颜色相同分两种情况：都白或都黑。总取法 $8\times 10=80$ 种，有利取法 $5\times 4+3\times 6=38$ 种，故

$$P=\frac{5\times 4+3\times 6}{8\times 10}=\frac{38}{80}=\frac{19}{40}.$$

练习 1.14 已知 $P(B)=0.5$，$P(A\cup B)=0.7$，则 $P(A\overline{B})=$ _____，$P(\overline{AB})=$ _____。

解 应填 0.2；0.3。

由加法公式：

$$P(A\overline{B})=P(A)-P(AB)=P(A\cup B)-P(B)=0.7-0.5=0.2,$$
$$P(\overline{AB})=P(\overline{A\cup B})=1-P(A\cup B)=1-0.7=0.3.$$

练习 1.15 对任意事件 A,B，下面结论正确的是（　）。

A. 若 $P(AB)=0$，则 $\overline{A}\,\overline{B}=\Omega$　B. $P(A-B)=P(A)-P(B)$　C. $P(A\overline{B})=P(A)-P(B)+P(\overline{A}B)$　D. $P(A\cup B)=P(A)+P(B)-P(A)P(B)$

解 应选 C。

A 错误：$P(AB)=0$ 不能推出 $AB=\varnothing$，从而不能推出 $\overline{A}\,\overline{B}=\Omega$。B 错误：$P(A-B)=P(A)-P(AB)$，一般不等于 $P(A)-P(B)$。D 错误：题目未说明 A,B 独立，故 $P(AB)\neq P(A)P(B)$。

C 正确：$P(A\overline{B})=P(A)-P(AB)$，$P(\overline{A}B)=P(B)-P(AB)$，两式消去 $P(AB)$ 得

$$P(A\overline{B})=P(A)-P(B)+P(\overline{A}B).$$

练习 1.16 设 A 和 B 是任意两个随机事件，则（　）。

A. $P(A+B)\leq P(A\cup B)\leq P(A)+P(B)$　B. $P(A)+P(B)-1\leq P(AB)\leq P(A\cup B)\leq P(A)+P(B)$　C. $P(A)+P(B)-1\leq P(A(A\cup B))\leq P(AB)\leq P(A\cup(A\cup B))$　D. $1-P(A)-P(B)\leq P(AB)\leq P(A)+P(B)\leq P(A\cup B)$

解 应选 B。

A 错误：$P(A+B)=P(A\cup B)$（概率中"+" 即并）。C 错误：$P(A(A\cup B))=P(A)\geq P(AB)$。D 错误：$P(A)+P(B)\geq P(A\cup B)$，故 $P(A)+P(B)\leq P(A\cup B)$ 不成立。

B 正确：由 $P(AB)=P(A)+P(B)-P(A\cup B)$：

- 左边：$P(A\cup B)\le 1$，故 $P(AB)\ge P(A)+P(B)-1$；
- 中间：$AB\subset A\cup B$，故 $P(AB)\le P(A\cup B)$；
- 右边：$P(A\cup B)\le P(A)+P(B)$（加法公式的直接推论）。

练习 1.17 设 A 与 B 是两个随机事件，且 $P(AB)=0$，则（　）。

A. $P(A-B)=P(A)$　B. A 与 B 相互独立　C. $P(A)=0$ 或 $P(B)=0$　D. A 与 B 互不相容

解 应选 A。$P(A-B)=P(A)-P(AB)=P(A)$，故 A 正确。B 与条件无必然联系；C 错误（例如连续型随机变量取特定值的概率为零，但事件本身可能发生）；D 错误（$P(AB)=0$ 是互不相容的必要条件而非充分条件）。

练习 1.18 提高 设事件 A,B 相互独立，事件 B,C 互不相容，事件 A 与 C 不能同时发生，且 $P(A)=P(B)=0.5$，$P(C)=0.2$。则事件 A,B 和 C 中仅 C 发生或仅 C 不发生的概率为 _____。

解 应填 0.45。“仅 C 发生”即 $C\overline{A}\,\overline{B}$。由于 $AC=\varnothing$ 且 $BC=\varnothing$，故 $C\subset\overline{A}$ 且 $C\subset\overline{B}$，即 $C\subset\overline{A}\,\overline{B}$，从而 $C\overline{A}\,\overline{B}=C$，$P(C)=0.2$。

“仅 C 不发生”即 $\overline{C}AB$。由独立性 $P(AB)=P(A)P(B)=0.25$，且 $AB\subset\overline{C}$（因为 $AC=\varnothing$），故 $\overline{C}AB=AB$，$P(\overline{C}AB)=0.25$。

又 C 与 AB 互斥（$C\subset\overline{A}$，$AB\subset A$），故

$$P(C\overline{A}\,\overline{B}\cup\overline{C}AB)=P(C)+P(AB)=0.2+0.25=0.45.$$

练习 1.19 提高 设 A,B 为任意两个随机事件，则（　）。

A. $P(AB)\le P(A)P(B)$　B. $P(AB)\ge P(A)P(B)$

C. $P(AB)\le\dfrac{P(A)+P(B)}{2}$　D. $P(AB)\ge\dfrac{P(A)+P(B)}{2}$

解 应选 C。由 $AB\subset A$ 和 $AB\subset B$，得 $P(AB)\le P(A)$ 且 $P(AB)\le P(B)$，因此

$$2P(AB)\le P(A)+P(B)\implies P(AB)\le\frac{P(A)+P(B)}{2}.$$

A、B 的反例：当 $A\subset B$ 时 $P(AB)=P(A)$，与 $P(A)P(B)$ 无确定大小关系；当 $AB=\varnothing$ 时 $P(AB)=0$。D 的反例同 A。

1.4 古典概型

定义 1.6 (古典概型)

如果随机试验的样本空间满足：

1. 只有有限个样本点（基本事件）；
2. 每个样本点（基本事件）发生的可能性都一样，

则称随机试验的概率模型为古典概型（等可能概型）。 ♣

如果古典概型的基本事件总数为 n，事件 A 包含 k 个基本事件，则 A 的概率为

$$P(A)=\frac{k}{n}=\frac{\text{事件 }A\text{ 所含基本事件的个数}}{\text{基本事件总数}}.$$

由上式计算得出的概率称为 A 的**古典概率**。

笔记 计算古典概率的关键是基本事件和样本空间的选定以及基本事件数的计算：

1. 列举法（直接数数法）：基本事件数不多时常用。

2. 集合对应法：
   (a). 加法原理——完成一件事有 n 类办法，第一类办法中有 m_1 种方法，第二类办法中有 m_2 种方法，…，第 n 类办法中有 m_n 种方法，则完成此事共有 $\sum_{i=1}^{n} m_i$ 种方法。
   (b). 乘法原理——完成一件事有 n 个步骤，第一步有 m_1 种方法，第二步有 m_2 种方法，…，第 n 步有 m_n 种方法，则完成此事共有 $\prod_{i=1}^{n} m_i$ 种方法。
   (c). 排列——从 n 个不同的元素中取出 $m(\leq n)$ 个元素，并按照一定顺序排成一列，叫作排列。所有排列的个数叫作排列数，记作
   $$\mathrm{P}_n^m = n(n-1)(n-2)\cdots(n-m+1) = \frac{n!}{(n-m)!}.$$
   当 $m=n$ 时，$\mathrm{P}_n^n = \frac{n!}{0!} = n!$，叫作全排列。
   (d). 组合——从 n 个不同的元素中取出 $m(\leq n)$ 个元素，并成一组，叫作组合。所有组合的个数叫作组合数，记作
   $$\mathrm{C}_n^m = \frac{\mathrm{P}_n^m}{m!} = \frac{n!}{m!(n-m)!}, \quad \mathrm{C}_n^m = \mathrm{C}_n^{n-m}, \quad 0! = 1, \quad \mathrm{C}_n^0 = 1.$$
3. 逆数法：先求 $\overline{A}$ 中的基本事件数 $n_{\overline{A}}$，用基本事件总数 n 减去 $n_{\overline{A}}$ 便得 A 中的基本事件数，这种方法常用于计算含有”至少”字样的事件的概率。

随机分配问题

随机分配也叫随机占位，突出一个”放”字，即将 n 个可辨质点随机地分配到 N 个盒子中，区分每盒可以容纳任意多个质点（不同分法的总数 N^n）和最多可以容纳一个质点（不同分法的总数 $\mathrm{P}_N^n = N(N-1)\cdots(N-n+1)$）。

笔记 每盒容纳任意多个质点：每个质点均可放到 N 个盒子中的任何一个，即有 N 种放法，于是 n 个可辨质点放到 N 个盒子中共有 N^n 种不同放法。

每盒容纳至多一个质点：质点可辨，且一个盒子至多容纳一个质点，故 n 个质点放到 $N(N \geq n)$ 个盒子中的所有不同放法即从 N 个元素中选取 n 个元素的排列数 P_N^n。

练习 1.20 将 n 个球随机放入 $N(n \leq N)$ 个盒子中，每个盒子可以放任意多个球。求下列事件的概率：

(1) $A = \{$某指定 n 个盒子各有一球$\}$；

(2) $B = \{$恰有 n 个盒子各有一球$\}$；

(3) $C = \{$指定 $k(k \leq n)$ 个盒子各有一球$\}$。

解 设 n 个球、N 个盒子是可分辨的（例如编号），由于每个盒子可以放任意多个球，因此每个球都有 N 种不同的放置方法。将 n 个球随机放入 N 个盒子的一种放法作为基本事件，则基本事件总数为 N^n。

(1) 事件 A 所含的基本事件是 n 个不同球的一种排列，故
$$n_A = n!, \quad P(A) = \frac{n!}{N^n}.$$

(2) 事件 B 中的基本事件可以设想为先从 N 个盒子中选出 n 个（共有 C_N^n 种不同选法），而后把 n 个球随机放入这 n 个盒子中，每盒一球（共有 $n!$ 种不同放法），因此
$$n_B = \mathrm{C}_N^n \cdot n!, \quad P(B) = \frac{\mathrm{C}_N^n \cdot n!}{N^n}.$$

(3) 事件 C 的基本事件可以设想为先从 n 个球中选出 k 个球（有 C_n^k 种不同选法），再将这 k 个球随机放入指定的 k 个盒子中，每盒一球（有 $k!$ 种不同放法），最后将余下的 $n-k$ 个球随机放入其余的

$N-k$ 个盒子中，每个球都有 $N-k$ 种放置方法，因此共有 $(N-k)^{n-k}$ 种不同放法，故

$$n_C = \mathrm{C}_n^k \cdot k! \cdot (N-k)^{n-k}, \quad P(C) = \frac{\mathrm{C}_n^k k!(N-k)^{n-k}}{N^n} = \frac{n!(N-k)^{n-k}}{(n-k)!N^n}.$$

笔记 许多问题的结构形式与分球入盒问题相同，都属于随机占位问题。例如生日问题（n 个人生日，相当于 n 个球随机放入 365 个盒子中，每盒可以放多个球）；住房分配问题（n 个人被分配到 N 个房间中去，每个房间可住多个人）；乘客下车问题（n 个乘客在 N 个车站下车的各种可能情况）；等等。对这些问题的求解都可以用”将 n 个球等可能地投放到 N 个盒子中”的思路来考虑。

简单随机抽样问题

设 $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ 含 N 个元素，称 Ω 为总体。如果各元素被抽到的可能性相同，则总体 Ω 的抽样称作**简单随机抽样**，突出一个”取”字。

简单随机抽样分为先后有放回、先后无放回及任取这三种不同的方式。

1. 先后有放回取 n 次，抽取法总数 N^n
2. 先后无放回取 n 次，抽取法总数 $\mathrm{P}_N^n = N(N-1)\cdots(N-n+1)$
3. 任取 n 个，抽取法总数 C_N^n

笔记 先后有放回取 n 次：既考虑抽到何元素，又考虑各元素出现的顺序，每次从 Ω 中随意抽取一个元素，并在抽下一元素前将其放回 Ω，于是每次都有 N 个元素可被抽取，即有 N 种抽取方法，抽取 n 次，即 N^n。

先后无放回取 n 次：既考虑抽到何元素，又考虑各元素出现的顺序，只是抽出的元素均不再放回 Ω，于是每次抽取时都比上一次少了一个元素，抽 $n(n \le N)$ 次，即 $\mathrm{P}_N^n = N(N-1)\cdots(N-n+1)$。

任取 n 个：一次性取 n 个元素，相当于将 n 个元素无序且无放回地取走，共抽取法总数为 $\mathrm{C}_N^n = \dfrac{\mathrm{P}_N^n}{n!}$。

练习 1.21 袋中有 5 个球，3 个白球，2 个黑球。

(1) 先后有放回取 2 个球；

(2) 先后无放回取 2 个球；

(3) 任取 2 个球。

求取的 2 个球中至少 1 个是白球的概率。

解 用对立事件思想，计算”两球全黑”的种数，再用总数减去它。

1. 先后有放回取 2 个球：总数 $5^2 = 25$，两球全黑 $2^2 = 4$，故
$$P = \frac{5^2 - 2^2}{5^2} = \frac{21}{25}.$$
2. 先后无放回取 2 个球：总数 $5 \cdot 4 = 20$，两球全黑 $2 \cdot 1 = 2$，故
$$P = \frac{5 \cdot 4 - 2 \cdot 1}{5 \cdot 4} = \frac{18}{20} = \frac{9}{10}.$$
3. 任取 2 个球：总数 $\mathrm{C}_5^2 = 10$，两球全黑 $\mathrm{C}_2^2 = 1$，故
$$P = \frac{\mathrm{C}_5^2 - \mathrm{C}_2^2}{\mathrm{C}_5^2} = \frac{9}{10}.$$

笔记 上述 (2) 与 (3) 概率相同，这是因为 $5 \cdot 4 = \mathrm{C}_5^2 \cdot 2!$，左边上下有序，右边上下无序，相当于把顺序”消掉”了，故”先后无放回取 k 个球”与”任取 k 个球”的概率相同。由于 (3) 较方便，因此计算 (2) 时可按 (3) 来计算。

练习 1.22 袋中有 100 个球，40 个白球，60 个黑球。

(1) 先后有放回取 20 个球，求取出 15 个白球、5 个黑球的概率；

(2) 先后无放回取 20 个球，求取出 15 个白球、5 个黑球的概率；

(3) 先后有放回取 20 个球，求第 20 次取到白球的概率；

(4) 先后无放回取 20 个球，求第 20 次取到白球的概率。

解

1. $$p_1 = \frac{C_{20}^{15} \cdot 40^{15} \cdot 60^5}{100^{20}}.$$

2. 按照”任取 20 个球”来计算，即
$$p_2 = \frac{C_{40}^{15} C_{60}^{5}}{C_{100}^{20}}.$$

3. 有放回取球，每次抽取的样本空间没有变化，故每次取到白球的概率始终为 $\frac{40}{100} = \frac{2}{5}$，即
$$p_3 = \frac{40}{100} = \frac{2}{5}.$$

4. 看作随机占位问题：设有 100 个盒子，每个盒子中放 1 个球，则要求第 20 个盒子中放入白球即可。故
$$p_4 = \frac{C_{40}^{1} \cdot 99!}{100!} = \frac{40}{100} = \frac{2}{5}.$$

笔记 $p_4 = p_3$。无放回抽取时，每次取到白球的概率也不会变，这可理解成依概率摸球（抓阄模型）。类似地，设有 100 个灰球，白的成分与黑的成分之比为 $40:60$，每次取到球中白的成分始终是 40%，不受抽取顺序的影响。

练习 1.23 从一副扑克牌（52 张）中任取 3 张（不重复），则取出 3 张牌中至少有 2 张花色相同的概率为 _____。

解 应填 $\frac{256}{425}$。利用对立事件：所求概率的对立事件为”3 张牌花色各不相同”。从 4 种花色中选 3 种，每种花色取 1 张：
$$P(\text{花色各不相同}) = \frac{C_4^3 \times 13^3}{C_{52}^3} = \frac{4 \times 2197}{22100} = \frac{8788}{22100}.$$
因此
$$P(\text{至少 2 张同花色}) = 1 - \frac{8788}{22100} = \frac{13312}{22100} = \frac{256}{425}.$$

练习 1.24 提高 从 1 到 9 的 9 个整数中有放回地随机取 3 次，每次取一个数，则取出的 3 个数之积能被 10 整除的概率为 _____。

解 应填 $\frac{52}{243}$。积能被 10 整除当且仅当积同时含有因子 2 和因子 5。从 $\{1,\ldots,9\}$ 中，5 的倍数仅有 $\{5\}$，2 的倍数为 $\{2,4,6,8\}$。

设 A 为”至少取出一个 5 的倍数”，B 为”至少取出一个 2 的倍数”。由容斥原理：
$$\begin{aligned} P(AB) &= 1 - P(\overline{A}) - P(\overline{B}) + P(\overline{A}\,\overline{B}) = 1 - \frac{8^3}{9^3} - \frac{5^3}{9^3} + \frac{4^3}{9^3} \\ &= 1 - \frac{512}{729} - \frac{125}{729} + \frac{64}{729} = \frac{156}{729} = \frac{52}{243}. \end{aligned}$$

1.4.1 几何概型

定义 1.7 (几何概型)

如果随机试验的样本空间满足：

1. 样本空间（基本事件空间）Ω 是一个可度量的有界区域；
2. 每个样本点（基本事件）发生的可能性都一样，即样本点落入 Ω 的某一可度量的子区域 S 的可能性大小与 S 的几何度量成正比，而与 S 的位置及形状无关，

则称随机试验的概率模型为几何概型。 ♣

如果 S_A 是样本空间 Ω 的一个可度量的子区域，则事件 $A=\{\text{样本点落入区域}S_A\}$ 的概率为

$$P(A)=\frac{S_A\text{ 的几何度量}}{\Omega\text{ 的几何度量}}.$$

笔记 古典概型与几何概型的区别：基本事件有限、等可能发生的随机试验为古典概型；基本事件无限且具有几何度量、等可能发生的随机试验为几何概型。

练习 1.25 在区间 $(0,1)$ 中随机地取两个数，则这两个数之差的绝对值小于 0.5 的概率为 _____。

解 设两个数分别为 x,y，则 (x,y) 的取值范围为正方形区域 $D=\{(x,y)\mid 0<x<1,0<y<1\}$。又 $|x-y|<0.5$，则所在区域如图中阴影部分所示，于是所求概率为

$$p=\frac{1^2-0.5^2}{1^2}=0.75.$$

练习 1.26 长方形 $ABCD$ 的长 $AB=10\,\text{cm}$，宽 $BC=5\,\text{cm}$，在长方形 $ABCD$ 内任取一点 E，则 $P\left(\angle AEB\geq\frac{\pi}{2}\right)=$ _____。

解 应填 $\frac{\pi}{4}$。$\angle AEB\geq 90^\circ$ 当且仅当点 E 位于以 AB 为直径的半圆内。该圆半径 $r=\frac{AB}{2}=5\,\text{cm}$，半圆面积 $S_{\text{半圆}}=\frac{1}{2}\pi r^2=\frac{25\pi}{2}$。由于半圆半径恰好等于矩形宽 $BC=5\,\text{cm}$，半圆完全包含在矩形内，故

$$P=\frac{S_{\text{半圆}}}{S_{\text{矩形}}}=\frac{25\pi/2}{50}=\frac{\pi}{4}.$$

练习 1.27 从 $[0,1]$ 中随机地取两个数，其积大于 $\frac{1}{4}$，其和小于 $\frac{5}{4}$ 的概率为 _____。

解 应填 $\frac{15}{32}-\frac{1}{2}\ln 2$。设两个数分别为 x,y，则 (x,y) 在正方形 $\Omega=[0,1]^2$ 内均匀分布。所求事件为

$$A=\left\{(x,y)\,\middle|\,x+y<\frac{5}{4},\ xy>\frac{1}{4}\right\}.$$

曲线 $xy=\frac{1}{4}$ 与 $x+y=\frac{5}{4}$ 交于 $x=\frac{1}{4}$ 和 $x=1$。对 $x\in\left[\frac{1}{4},1\right]$，$y$ 的范围从 $\frac{1}{4x}$ 到 $\frac{5}{4}-x$：

$$\begin{aligned}P(A)&=\int_{1/4}^{1}\left(\frac{5}{4}-x-\frac{1}{4x}\right)\mathrm{d}x=\left[\frac{5}{4}x-\frac{1}{2}x^2-\frac{1}{4}\ln x\right]_{1/4}^{1}\\&=\frac{3}{4}-\left(\frac{9}{32}+\frac{1}{2}\ln 4\right)=\frac{15}{32}-\frac{1}{2}\ln 2.\end{aligned}$$

练习 1.28 随机地向半圆 $0<y<\sqrt{2ax-x^2}$（$a>0$）内投掷一点，求该点和原点连线与 x 轴的夹角 $\theta\leq\frac{\pi}{4}$ 的概率。

解 应填 $\frac{1}{2}+\frac{1}{\pi}$。半圆方程 $y=\sqrt{2ax-x^2}$ 即 $(x-a)^2+y^2=a^2$（$y>0$），是以 $(a,0)$ 为圆心、a 为半径的上半圆。直线 $y=x$ 与半圆交于 $(0,0)$ 和 (a,a)。$\theta\leq\pi/4$ 对应 $y\leq x$。在半圆内，该区域由两部分

组成：

- 三角形区域 $(0,0),(a,0),(a,a)$：面积为 $\dfrac{a^2}{2}$；
- 半圆从 $x=a$ 到 $x=2a$ 的部分：面积为 $\dfrac{1}{4}\pi a^2$。

半圆总面积 $S_\Omega=\dfrac{1}{2}\pi a^2$，故

$$P=\frac{a^2/2+\pi a^2/4}{\pi a^2/2}=\frac{1}{2}+\frac{1}{\pi}.$$

1.5 条件概率

1.5.1 条件概率的定义

定义 1.8 (条件概率)

设 A,B 是两个事件，且 $P(A)>0$，称

$$P(B|A)=\frac{P(AB)}{P(A)}$$

为在事件 A 发生的条件下事件 B 发生的条件概率。♣

性质 条件概率 $P(\cdot|A)$ 满足概率的三条公理：

1. 非负性：对于每一事件 B，有 $P(B|A)\geq 0$；
2. 规范性：$P(\Omega|A)=1$；
3. 可列可加性：设 $B_1,B_2,\ldots,B_n,\ldots$ 是两两互斥的事件，则

$$P\left(\bigcup_{i=1}^{\infty}B_i\middle|A\right)=\sum_{i=1}^{\infty}P(B_i|A).$$

因此，概率的一切性质和公式对条件概率都适用。

笔记 对于条件概率 $P(B|A)$，可以理解为样本空间由原来的 Ω 缩小为 A，即

$$P(B)=P(B|\Omega)\xrightarrow{\Omega\to A}P(B|A).$$

从这个角度理解，样本空间的缩小使不确定性减小了，即当 $A\subset B\subset\Omega$ 时，有 $P(A|B)\geq P(A)$ 成立。条件概率就是在”已知某事件发生了”这一附加条件下所计算的概率。例如，

$$P(\overline{B}|A)=1-P(B|A),\quad P[(B-C)|A]=P(B|A)-P(BC|A).$$

1.5.2 乘法公式

定理 1.1 (乘法公式)

设 $P(A)>0$，则

$$P(AB)=P(B|A)\,P(A).$$

一般地，设 $A_1,A_2,\ldots,A_n$ 为 n 个事件（$n\geq 2$），且 $P(A_1A_2\cdots A_{n-1})>0$，则

$$P(A_1A_2\cdots A_n)=P(A_1)\,P(A_2|A_1)\,P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

♡

1.5.3 全概率公式与贝叶斯公式

前面已经介绍了完备事件组的概念（见定义 1.4），它是全概率公式和贝叶斯公式的前提条件。

定理 1.2 (全概率公式)

设 $B_1, B_2, \ldots, B_n$ 是一组完备事件组，A 为任一事件，且 $P(B_i) > 0\ (i = 1, 2, \ldots, n)$，则

$$P(A) = \sum_{i=1}^{n} P(B_i)\, P(A|B_i).$$

全概率公式的实质是**全集分解法**：如果事件 A 的发生总是与多个" 原因" B_i 相联系，则

$$A = A\Omega = A\left(\bigcup_{i=1}^{n} B_i\right) = \bigcup_{i=1}^{n} B_i A,$$

然后应用有限可加性和条件概率即可得到。

定理 1.3 (贝叶斯公式)

设 $B_1, B_2, \ldots, B_n$ 是一组完备事件组，A 为任一事件，且 $P(A) > 0$，$P(B_k) > 0\ (k = 1, 2, \ldots, n)$，则

$$P(B_k|A) = \frac{P(B_k)\, P(A|B_k)}{P(A)} = \frac{P(B_k)\, P(A|B_k)}{\sum_{i=1}^{n} P(B_i)\, P(A|B_i)}, \quad k = 1, 2, \ldots, n.$$

笔记 全概率公式与贝叶斯公式的记忆技巧：

- 全概率公式——由因求果。将 A 视为结果，B_i 视为导致 A 发生的原因，已知各原因的概率 $P(B_i)$ 及各原因导致结果 A 的概率 $P(A|B_i)$，求结果 A 发生的概率 $P(A)$。
- 贝叶斯公式——执果寻因。已知结果 A 已经发生，反推是由第 k 个原因 B_k 导致的概率 $P(B_k|A)$。

练习 1.29 某人外出可以乘坐飞机、火车、轮船、汽车四种交通工具，其概率依次为 $0.05, 0.15, 0.30, 0.50$，而乘坐这几种交通工具能如期到达的概率依次为 $0.80, 0.70, 0.60, 0.90$。求：

(1) 该人如期到达的概率；

(2) 已知该人误期到达，求他是乘坐火车的概率。

解 设 $A_i\ (i = 1, 2, 3, 4)$ 分别表示乘坐飞机、火车、轮船、汽车，B 表示如期到达。

(1) 利用全概率公式：

$$P(B) = \sum_{i=1}^{4} P(A_i)\, P(B|A_i) = 0.05 \times 0.80 + 0.15 \times 0.70 + 0.30 \times 0.60 + 0.50 \times 0.90 = 0.775.$$

(2) 利用贝叶斯公式：

$$P(A_2|\overline{B}) = \frac{P(A_2)\, P(\overline{B}|A_2)}{P(\overline{B})} = \frac{0.15 \times (1 - 0.70)}{1 - 0.775} = \frac{0.045}{0.225} = 0.2.$$

练习 1.30 设 A, B 为两个随机事件，则（　）。

A. $P(AB) + P(\overline{A}\,\overline{B}) \le 1$　　B. $P(AB) + P(\overline{A}\,\overline{B}) \ge 1$

C. $P(A - B) \le P(A) - P(B)$　　D. $P(A \cup B) + P(\overline{A} \cup \overline{B}) \le 1$

解 应选 A。由于 $AB \subset A \cup B$，即 $P(AB) \le P(A \cup B)$，从而

$$P(AB) \le P(A \cup B) = 1 - P(\overline{A \cup B}) = 1 - P(\overline{A}\,\overline{B}),$$

即 $P(AB) + P(\overline{A}\,\overline{B}) \le 1$，故选 A。

同时有

$$P(A\cup B)+P(\overline{A}\cup\overline{B})-1=P(A\cup B)+P(\overline{AB})-1\geq 0,$$

故选项 B、D 不正确。又 $AB\subset B$，有 $P(AB)\leq P(B)$，从而

$$P(A-B)=P(A)-P(AB)\geq P(A)-P(B),$$

知选项 C 不正确。

练习 1.31 设随机事件 A,B 满足 $P(A)=P(B)=\dfrac{1}{2}$ 和 $P(A\cup B)=1$，则 ()。

A. $A\cup B=\Omega$　B. $AB=\varnothing$　C. $P(\overline{A}\cup\overline{B})=1$　D. $P(A-B)=0$

解 应选 C。$P(A\cup B)=1$ 不能推出 $A\cup B=\Omega$，且 $P(AB)=P(A\cup B)-P(A)-P(B)=0$ 不能推出 $AB=\varnothing$。故选项 A、B 不正确。

又 $P(A-B)=P(A)-P(AB)=\dfrac{1}{2}$，选项 D 不正确。

从而 $P(\overline{A}\cup\overline{B})=P(\overline{AB})=1-P(AB)=1$，故选 C。

练习 1.32 设 A,B,C 是三个随机事件，A 与 C 互斥，$P(AB)=\dfrac{1}{2}$，$P(C)=\dfrac{1}{3}$，则 $P(AB|\overline{C})=$ _____。

解 应填 $\dfrac{3}{4}$。因为 A 与 C 互斥，所以 $P(AC)=0$。又 $ABC\subset AC$，所以 $P(ABC)=0$，故

$$P(AB|\overline{C})=\frac{P(AB\overline{C})}{P(\overline{C})}=\frac{P(AB)-P(ABC)}{1-P(C)}=\frac{\frac{1}{2}}{1-\frac{1}{3}}=\frac{3}{4}.$$

笔记 由于 $AC=\varnothing$，则 $ABC=\varnothing\cdot B=\varnothing$，得到 $P(ABC)=0$。

练习 1.33 已知 $P(A)=\dfrac{1}{4}$，$P(B|A)=\dfrac{1}{3}$，$P(A|B)=\dfrac{1}{2}$，则 $P(A\cup B)=$ _____。

解 应填 $\dfrac{1}{3}$。由乘法公式知

$$P(AB)=P(A)\,P(B|A)=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12},\quad P(B)=\frac{P(AB)}{P(A|B)}=\frac{1/12}{1/2}=\frac{1}{6},$$

所以

$$P(A\cup B)=P(A)+P(B)-P(AB)=\frac{1}{4}+\frac{1}{6}-\frac{1}{12}=\frac{1}{3}.$$

练习 1.34 某人给四位亲友各写一封信，然后随机地装入 4 个写好地址的信封中，且每个信封装一封信。求：

(1) 4 封信都装对的概率；

(2) 4 封信都装错的概率。

解 设事件 $A_i\ (i=1,2,3,4)$ 为"第 i 封信装对了"，A 为"4 封信都装对了"，B 为"4 封信都装错了"。

(1) 利用乘法公式：

$$\begin{aligned}P(A)&=P(A_1A_2A_3A_4)=P(A_1)\,P(A_2|A_1)\,P(A_3|A_1A_2)\,P(A_4|A_1A_2A_3)\\&=\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}\times\frac{1}{1}=\frac{1}{24}.\end{aligned}$$

(2) 利用逆向思维和对偶律：

$$\begin{aligned}P(B)&=P(\overline{A_1}\,\overline{A_2}\,\overline{A_3}\,\overline{A_4})=1-P(A_1\cup A_2\cup A_3\cup A_4)\\&=1-\sum_{i=1}^{4}P(A_i)+\sum_{1\leq i<j\leq 4}P(A_iA_j)-\sum_{1\leq i<j<k\leq 4}P(A_iA_jA_k)+P(A_1A_2A_3A_4),\end{aligned}$$

其中 $P(A_i)=\frac{1}{4}$，$P(A_iA_j)=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12}$，$P(A_iA_jA_k)=\frac{1}{4}\times\frac{1}{3}\times\frac{1}{2}=\frac{1}{24}$，于是

$$P(B)=1-4\times\frac{1}{4}+6\times\frac{1}{12}-4\times\frac{1}{24}+\frac{1}{24}=\frac{3}{8}.$$

笔记 装信过程可看作无放回抽取信件，依次装入信封。第 (2) 问直接计算较为困难，采用逆向思维方法更为简捷。

练习 1.35 从 $1,2,3,4$ 中任取一个数，记为 X，再从 $1,\ldots,X$ 中任取一个数，记为 Y，则 $P\{Y=2\}=$ _____。

解 应填 $\frac{13}{48}$。

由全概率公式：

$$\begin{aligned}P\{Y=2\}&=\sum_{x=1}^{4}P\{X=x\}\,P\{Y=2|X=x\}\\&=\frac{1}{4}\times 0+\frac{1}{4}\times\frac{1}{2}+\frac{1}{4}\times\frac{1}{3}+\frac{1}{4}\times\frac{1}{4}=\frac{13}{48}.\end{aligned}$$

练习 1.36 一批产品有 10 个正品和 2 个次品，任意抽取两次，每次抽一个，抽出后不再放回，则第二次抽出的是次品的概率为 _____。

解 应填 $\frac{1}{6}$。记 $A_i=\{第 i 次取得次品\}\ (i=1,2)$，由全概率公式得

$$P(A_2)=P(A_1)\,P(A_2|A_1)+P(\overline{A_1})\,P(A_2|\overline{A_1})=\frac{2}{12}\times\frac{1}{11}+\frac{10}{12}\times\frac{2}{11}=\frac{1}{6}.$$

笔记 此例为抓阄模型：无放回抽取时，每次取到次品的概率不变，均为 $\frac{2}{12}=\frac{1}{6}$。从第 1 次到第 12 次取到次品的概率均为 $\frac{1}{6}$。

练习 1.37 设有两批数量相同的零件，已知有一批产品全部合格，另一批产品有 25% 不合格。从两批产品中任取 1 只，经检验是正品，放回原处，并在原所在批次再取 1 只，求这只产品是次品的概率。

解 设 $H_i\ (i=1,2)$ 为"第一次从第 i 批产品中抽取"，A 为"所取产品为正品"。

第一次抽取：

$$P(H_1)=P(H_2)=\frac{1}{2},\quad P(A|H_1)=1,\quad P(A|H_2)=\frac{3}{4},$$

$$P(A)=P(H_1)\,P(A|H_1)+P(H_2)\,P(A|H_2)=\frac{1}{2}\times 1+\frac{1}{2}\times\frac{3}{4}=\frac{7}{8}.$$

由贝叶斯公式更新概率：

$$P(H_1|A)=\frac{P(H_1)\,P(A|H_1)}{P(A)}=\frac{4}{7},\quad P(H_2|A)=\frac{3}{7}.$$

第二次抽取：设所取为次品的概率为

$$P(\overline{A})=P(H_1|A)\cdot P(\overline{A}|H_1)+P(H_2|A)\cdot P(\overline{A}|H_2)=\frac{4}{7}\times 0+\frac{3}{7}\times\frac{1}{4}=\frac{3}{28}.$$

笔记 本题是全概率公式与贝叶斯公式的复合应用：第一次用全概率公式计算 $P(A)$，再用贝叶斯公式更新各原因的概率，最后第二次再用全概率公式计算结果。两次抽取的关键区别在于：第一次在完全不知情的情况下等可能地抽取；第二次在已知第一次取正品的条件下，各批次被抽到的概率已经发生了变化。

练习 1.38 设 $P(\overline{A})=0.3$，$P(B)=0.4$，$P(A\overline{B})=0.5$，则 $P(B|A\cup\overline{B})=$ _____。

解 应填 0.25。

首先，$P(A)=1-P(\overline{A})=0.7$。根据减法公式：

$$P(AB)=P(A)-P(A\overline{B})=0.7-0.5=0.2.$$

根据加法公式：

$$P(A\cup\overline{B})=P(A)+P(\overline{B})-P(A\overline{B})=0.7+0.6-0.5=0.8.$$

利用分配律化简分子：

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

故

$$P(B|A\cup\overline{B})=\frac{P(AB)}{P(A\cup\overline{B})}=\frac{0.2}{0.8}=0.25.$$

练习 1.39 设 A,B 为任意两事件且 $A\subset B$，$P(B)>0$，则下列结论一定正确的是（　）。

A. $P(A)\geq P(A|B)$　B. $P(A)>P(A|B)$　C. $P(A)<P(A|B)$　D. $P(A)\leq P(A|B)$

解 应选 D。

由 $A\subset B$，有 $AB=A$，故

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(A)}{P(B)}\geq P(A).$$

从样本空间更新的角度理解：$P(A)=P(A|\Omega)$，而 $A\subset B\subset\Omega$，样本空间缩小为 B 后，A 在新样本空间中的”占比”不变，但不确定性减小，概率增大。

练习 1.40 假如在数字通信中传送信号 0 与 1 的概率分别为 0.7 和 0.3。由于随机干扰，当传送信号 0 时接收到信号 0 的概率是 0.8；当传送信号 1 时接收到信号 1 的概率是 0.9。求：

(1) 接收到信号 0 的概率；

(2) 当接收到信号 0 时，传送的信号是 0 的概率。

解 设事件 A：传送信号 0；事件 B：接收到信号 0。已知 $P(A)=0.7$，$P(\overline{A})=0.3$，$P(B|A)=0.8$，$P(\overline{B}|\overline{A})=0.9$，故 $P(B|\overline{A})=0.1$。

(1) 由全概率公式：

$$P(B)=P(B|A)\,P(A)+P(B|\overline{A})\,P(\overline{A})=0.8\times0.7+0.1\times0.3=0.59.$$

(2) 由贝叶斯公式：

$$P(A|B)=\frac{P(B|A)\,P(A)}{P(B)}=\frac{0.8\times0.7}{0.59}=\frac{56}{59}.$$

笔记 本题是全概率公式与贝叶斯公式的典型应用：传送信号为”因”，接收信号为”果”。第 (1) 问”由因求果”用全概率公式，第 (2) 问”执果寻因”用贝叶斯公式。

练习 1.41 甲、乙两人独立地对同一目标射击一次，其命中率分别为 0.6 和 0.7。已知目标被击中，则它是甲射中的概率为 _____。

解 应填 $\frac{15}{22}$（约 0.682）。

记 A 为”目标被击中”，B_1 为”甲射中”，B_2 为”乙射中”。$A=B_1\cup B_2$，由独立性：

$$P(A)=P(B_1)+P(B_2)-P(B_1B_2)=0.6+0.7-0.6\times0.7=0.88.$$

由于 $B_1\subset A$，

$$P(B_1|A)=\frac{P(B_1)}{P(A)}=\frac{0.6}{0.88}=\frac{15}{22}.$$

练习 1.42 口袋中有 1 个球，不知其颜色是黑还是白。现再往口袋中放入 1 个白球，然后从中任意取出 1 个，发现取出的是白球。求口袋中原来那个球是白球的概率。

解 应填 $\frac{2}{3}$。

记 A 为"取出的是白球"，B 为"原来那个球是白球"。由于不知道原球颜色，取等可能先验概率 $P(B)=P(\overline{B})=\frac{1}{2}$。

$P(A|B)=1$（原来若为白球，袋中有 2 白，必取白），$P(A|\overline{B})=\frac{1}{2}$（原来若为黑球，袋中 1 白 1 黑）。

由贝叶斯公式：

$$P(B|A)=\frac{P(B)\,P(A|B)}{P(B)\,P(A|B)+P(\overline{B})\,P(A|\overline{B})}=\frac{\frac{1}{2}\times 1}{\frac{1}{2}\times 1+\frac{1}{2}\times\frac{1}{2}}=\frac{2}{3}.$$

练习 1.43 已知装有同种零件的产品两箱，第一箱内装 50 件，其中一等品 10 件；第二箱内装 30 件，其中一等品 18 件。现从两箱中任意挑选一箱，从中先后取出两件产品（取后不放回）。求：先取出的产品是一等品的概率 p；已知先取出的产品是一等品，则第二次取出的产品仍然是一等品的概率 q。

解 设 A_i 为"第 i 次取出一等品"（$i=1,2$），C_j 为"挑选第 j 箱"（$j=1,2$）。$C_1\cup C_2=\Omega$，$C_1C_2=\varnothing$，$P(C_1)=P(C_2)=\frac{1}{2}$。

由全概率公式：

$$p=P(A_1)=P(C_1)\,P(A_1|C_1)+P(C_2)\,P(A_1|C_2)=\frac{1}{2}\times\frac{10}{50}+\frac{1}{2}\times\frac{18}{30}=\frac{2}{5},$$

$$\begin{aligned}P(A_1A_2)&=P(C_1)\,P(A_1A_2|C_1)+P(C_2)\,P(A_1A_2|C_2)\\&=\frac{1}{2}\times\frac{10\times 9}{50\times 49}+\frac{1}{2}\times\frac{18\times 17}{30\times 29}=\frac{276}{1421}.\end{aligned}$$

故 $q=P(A_2|A_1)=\dfrac{P(A_1A_2)}{P(A_1)}=\dfrac{276}{1421}\times\dfrac{5}{2}=\dfrac{690}{1421}$。

练习 1.44 每箱产品有 10 件，其次品数从 0 到 2 是等可能的。开箱检验时从中任取 1 件，检验出次品则拒收。由于检验有误，将正品误认为次品的概率为 2%，次品被漏查的概率为 5%。求该箱产品通过验收的概率。

解 应填 0.887。

设 A_i（$i=0,1,2$）为"箱中有 i 件次品"，B 为"通过验收"。$P(A_i)=\frac{1}{3}$。

取到正品时通过验收的概率为 0.98，取到次品时通过验收的概率为 0.05。由全概率公式（两层）：

$$\begin{aligned}P(B|A_0)&=1\times 0.98=0.98,\\P(B|A_1)&=\frac{9}{10}\times 0.98+\frac{1}{10}\times 0.05=0.887,\\P(B|A_2)&=\frac{8}{10}\times 0.98+\frac{2}{10}\times 0.05=0.794,\\P(B)&=\frac{1}{3}(0.98+0.887+0.794)=\frac{2.661}{3}=0.887.\end{aligned}$$

练习 1.45 设水杯成箱出售，每箱 12 个，每箱中含 0 个、1 个、2 个残次品的概率分别为 0.7、0.19、0.11。某人欲购一箱，售货员随机取一箱，从中任取 3 只查看，若无残次品则购买。

(1) 此人购买一箱的概率；

(2) 在此人购买一箱的条件下，该箱无残次品的概率。

解 设 A_i（$i=0,1,2$）为"箱中有 i 件残次品"，B 为"购买此箱"（即取出的 3 只均无残次品）。

(1) 由全概率公式：

$$P(B)=\sum_{i=0}^{2}P(A_i)\,P(B|A_i),$$

$$P(B|A_0)=1,\quad P(B|A_1)=\frac{C_{11}^3}{C_{12}^3}=\frac{3}{4},\quad P(B|A_2)=\frac{C_{10}^3}{C_{12}^3}=\frac{6}{11},$$

$$P(B)=0.7\times 1+0.19\times\frac{3}{4}+0.11\times\frac{6}{11}=0.9025.$$

(2) 由贝叶斯公式：

$$P(A_0|B)=\frac{P(A_0)\,P(B|A_0)}{P(B)}=\frac{0.7\times 1}{0.9025}=\frac{280}{361}\approx 0.7756.$$

练习 1.46 统计资料显示，在出口罐头导致的索赔事件中，有 50% 是质量问题，30% 是数量问题，20% 是包装问题。在这三类问题各引起的索赔事件中，不经过法律诉讼而通过协商解决的概率分别为 40%、60% 和 75%。

(1) 索赔事件通过协商解决的概率；

(2) 已知一索赔事件通过协商解决，求该事件不是包装问题的概率。

解 设 A_1,A_2,A_3 分别为质量、数量、包装问题引起的索赔，B 为"通过协商解决"。

(1) 由全概率公式：

$$P(B)=\sum_{i=1}^{3}P(A_i)\,P(B|A_i)=0.5\times 0.40+0.3\times 0.60+0.2\times 0.75=0.53.$$

(2) 由贝叶斯公式：

$$P(A_3|B)=\frac{P(A_3)\,P(B|A_3)}{P(B)}=\frac{0.2\times 0.75}{0.53}=\frac{15}{53},$$

故所求概率为 $1-P(A_3|B)=\dfrac{38}{53}$。

练习 1.47 在某城市中，下雨和晴天的天数各占一半，而天气无论在雨天还是晴天都有 $\dfrac{2}{3}$ 的概率预报准确。如果预报下雨，王明同学就一定带雨伞；如果预报无雨，则不带伞。设 $A=\{$天下雨$\}$，$B=\{$预报有雨$\}$，$C=\{$王明带雨伞$\}$。

(1) 事件 $\overline{A}\cap\overline{B}\cap C$、$\overline{A}\cap B\cap C$ 的含义是什么？哪个为不可能事件？

(2) 求他带雨伞而没有下雨的概率。

解 由题意：$P(A)=\dfrac{1}{2}$，$P(B|A)=\dfrac{2}{3}$（预报准确），$P(\overline{B}|\overline{A})=\dfrac{2}{3}$（预报准确），故 $P(B|\overline{A})=\dfrac{1}{3}$。又 $P(C|B)=1$（预报有雨必带伞），$P(C|\overline{B})=0$（预报无雨不带伞）。

(1) $\overline{A}\cap\overline{B}\cap C$：天没下雨、预报无雨、却带了雨伞。由于 $P(C|\overline{B})=0$，这是一个不可能事件。

$\overline{A}\cap B\cap C$：天没下雨、预报有雨、带了雨伞。这是可能事件。

(2)"带雨伞而没有下雨"即 $\overline{A}\cap C$。由于 C 完全由 B 决定，$\overline{A}\cap\overline{B}\cap C$ 为不可能事件，故

$$P(\overline{A}\cap C)=P(\overline{A}\cap B\cap C)=P(\overline{A})\,P(B|\overline{A})\,P(C|\overline{A}B)=\frac{1}{2}\times\frac{1}{3}\times 1=\frac{1}{6}.$$

练习 1.48 提高 设某同学的手机在一天内收到短信数服从泊松分布 $P(\lambda)$，每个短信是否为垃圾短信与其到达时间独立，也与其他短信是否为垃圾短信相互独立。假设每个短信是垃圾短信的概率为 p。

(1) 已知一天内收到了 n 条短信，求其中恰有 k 条垃圾短信的概率（$0\le k\le n$）。

(2) 求一天内收到 k 条垃圾短信的概率（$k=0,1,2,\ldots$）。

解 (1) 记 X 为一天内收到短信的总数，Y 为其中垃圾短信的条数。已知 $X=n$ 时，每条短信独立地以

概率 p 为垃圾短信，故

$$P(Y=k \mid X=n)=\mathrm{C}_n^k\, p^k\,(1-p)^{n-k}.$$

(2) 由全概率公式：

$$\begin{aligned}P(Y=k)&=\sum_{n=k}^{\infty} P(Y=k \mid X=n)\,P(X=n)=\sum_{n=k}^{\infty}\mathrm{C}_n^k\, p^k\,(1-p)^{n-k}\,\frac{\lambda^n}{n!}\,e^{-\lambda}\\&=\frac{p^k\,e^{-\lambda}\,\lambda^k}{k!}\sum_{n=k}^{\infty}\frac{[\lambda(1-p)]^{n-k}}{(n-k)!}=\frac{(\lambda p)^k}{k!}\,e^{-\lambda}\,e^{\lambda(1-p)}=\frac{(\lambda p)^k}{k!}\,e^{-\lambda p}.\end{aligned}$$

即 $Y\sim P(\lambda p)$。

笔记 泊松分布的分解定理（稀疏性）：以概率 p 独立地对泊松过程 $P(\lambda)$ 进行"稀疏"，得到的新过程仍为泊松分布，参数变为 λp。本题是全概率公式与泊松分布结合的典型应用。

1.6 事件的独立性

1.6.1 独立性的定义

定义 1.9 (事件的独立性)

设 A,B 为两个事件，如果

$$P(AB)=P(A)\,P(B),$$

则称事件 A 与 B 相互独立，简称 A 与 B 独立。 ♣

顺带剧透一下，我们后面会讨论多个事件的独立性概念。

定义 1.10 (n 个事件的相互独立)

设 $A_1,A_2,\ldots,A_n$ 为 $n\,(n\geq 2)$ 个事件，如果对其中任意有限个事件 $A_{i_1},A_{i_2},\ldots,A_{i_k}\,(2\leq k\leq n)$，都有

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})\,P(A_{i_2})\cdots P(A_{i_k}),$$

则称这 n 个事件相互独立。 ♣

特别地，对于三个事件 A,B,C，相互独立要求同时满足以下四个等式：

$$\begin{cases}P(AB)=P(A)\,P(B),\\P(AC)=P(A)\,P(C),\\P(BC)=P(B)\,P(C),\\P(ABC)=P(A)\,P(B)\,P(C).\end{cases}$$

若仅满足前三个等式（即任意两个事件独立），则称 A,B,C **两两独立**。

笔记 从条件概率的角度理解独立性：若 A,B 相互独立且 $P(B)>0$，则 $P(A|B)=P(A)$。这意味着样本空间缩小为 B 后，事件 A 发生的概率并没有改变——B 的发生对 A 没有任何影响，这正是"独立"一词的含义。

1.6.2 独立性的判定

性质 A 与 B 独立（$P(AB)=P(A)P(B)$）的等价条件（设 $0<P(A)<1$）：

1. A 与 B 独立 $\Leftrightarrow$ A 与 $\overline{B}$ 独立 $\Leftrightarrow$ $\overline{A}$ 与 B 独立 $\Leftrightarrow$ $\overline{A}$ 与 $\overline{B}$ 独立；
2. A 与 B 独立 $\Leftrightarrow P(B|A)=P(B|\overline{A})=P(B)$；
3. A 与 B 独立 $\Leftrightarrow P(B|A)+P(\overline{B}|\overline{A})=1$。

注 (1) 关于等价条件 (1) 的证明：由 A,B 独立，有 $P(AB)=P(A)P(B)$，于是

$$P(A\overline{B})=P(A)-P(AB)=P(A)-P(A)P(B)=P(A)[1-P(B)]=P(A)P(\overline{B}),$$

故 $A,\overline{B}$ 独立，其余同理。

(2) 推广：将相互独立的事件组中的任何几个事件换成各自的对立事件，所得的新事件组仍相互独立。

性质 独立性的判定与运算：

1. 若 $P(A)=0$ 或 $P(A)=1$，则 A 与任意事件 B 相互独立。
2. 若 $P(A)>0,P(B)>0$，且 A,B 互斥，则 A,B 一定不独立；反之，若 A,B 独立，则 A,B 一定不互斥（除非 $P(A)=0$ 或 $P(B)=0$）。
3. 若一组事件相互独立，则其中部分事件也相互独立；对不含相同事件的独立组分别作运算，所得新事件仍独立。例如 A,B,C,D 相互独立，则 AB 与 CD 独立，A 与 $BC-D$ 独立。

笔记 常见误区：

1. ”独立性”是概率意义上的数学定义，与逻辑因果上的”独立”无关，只看 $P(AB)=P(A)P(B)$ 是否成立。
2. 两两独立 $\not\Rightarrow$ 相互独立。三个事件两两独立只保证 $P(AB)=P(A)P(B)$、$P(AC)=P(A)P(C)$、$P(BC)=P(B)P(C)$，但不保证 $P(ABC)=P(A)P(B)P(C)$。

练习 1.49 将一枚硬币独立地掷两次，设 $A_1=\{$掷第一次出现正面$\}$，$A_2=\{$掷第二次出现正面$\}$，$A_3=\{$正反面各出现一次$\}$，$A_4=\{$正面出现两次$\}$，则事件（　）。

A. A_1,A_2,A_3 相互独立　B. A_2,A_3,A_4 相互独立

C. A_1,A_2,A_3 两两独立　D. A_2,A_3,A_4 两两独立

解 应选 C。

根据古典概型：

$$P(A_1)=P(A_2)=P(A_3)=\frac{1}{2},\quad P(A_1A_2)=P(A_1A_3)=P(A_2A_3)=\frac{1}{4},$$

满足两两独立条件。但

$$P(A_1A_2A_3)=0\neq P(A_1)P(A_2)P(A_3)=\frac{1}{8},$$

故 A_1,A_2,A_3 两两独立但不相互独立。

又 $A_3\cap A_4=\varnothing$，$P(A_3A_4)=0\neq P(A_3)P(A_4)=\frac{1}{8}$，故 A_2,A_3,A_4 不两两独立。

练习 1.50 设 A,B,C 为三个两两独立的随机事件，$P(A)=P(B)=P(C)$，$P(A\cup B\cup C)=\dfrac{23}{25}$，$P(\overline{A}\cup\overline{B}\cup\overline{C})=\dfrac{14}{25}$，求 $P(A)$。

解 由两两独立性，

$$\begin{aligned}P(\overline{A}\cup\overline{B}\cup\overline{C})&=P(\overline{A})+P(\overline{B})+P(\overline{C})-P(\overline{A}\,\overline{B})-P(\overline{B}\,\overline{C})-P(\overline{A}\,\overline{C})+P(\overline{A}\,\overline{B}\,\overline{C})\\&=3P(\overline{A})-3[P(\overline{A})]^2+P(A\cup B\cup C).\end{aligned}$$

代入数据：

$$3P(\overline{A})-3[P(\overline{A})]^2+\frac{2}{25}=\frac{14}{25}\implies 3P(\overline{A})-3[P(\overline{A})]^2=\frac{12}{25}.$$

令 $x=P(\overline{A})$，则 $3x-3x^2=\frac{12}{25}$，即 $25x^2-25x+4=0$，解得 $x=\frac{1}{5}$ 或 $\frac{4}{5}$。

又 $P(\overline{A})\le P(\overline{A}\cup\overline{B}\cup\overline{C})=\frac{14}{25}$，故 $P(\overline{A})=\frac{1}{5}$，即 $P(A)=\frac{4}{5}$。

练习 1.51 设 A,B,C 为三个随机事件，且 A 与 C 相互独立，B 与 C 相互独立，则 $A\cup B$ 与 C 相互独立的充分必要条件是（　）[1]。

A. A 与 B 相互独立　B. A 与 B 互不相容　C. AB 与 C 相互独立　D. AB 与 C 互不相容

解 应选 C。

$A\cup B$ 与 C 相互独立，即 $P[(A\cup B)C]=P(A\cup B)\,P(C)$。

$$\begin{aligned}P[(A\cup B)C]&=P(AC\cup BC)=P(AC)+P(BC)-P(ABC)\\&=P(A)\,P(C)+P(B)\,P(C)-P(ABC),\\P(A\cup B)\,P(C)&=[P(A)+P(B)-P(AB)]\,P(C)\\&=P(A)\,P(C)+P(B)\,P(C)-P(AB)\,P(C),\end{aligned}$$

因此充要条件为 $P(ABC)=P(AB)\,P(C)$，即 AB 与 C 相互独立。

练习 1.52 已知 A,B 相互独立，$P(A)=0.5$，$P(B)=0.4$，$C=\overline{A}\cup B$。求 $P(\overline{C})$，$P(C\mid(A\cup B))$。

解 $\overline{C}=\overline{\overline{A}\cup B}=A\overline{B}$，由独立性：

$$P(\overline{C})=P(A\overline{B})=P(A)(1-P(B))=0.5\times0.6=0.3.$$

又 $P(A\cup B)=P(A)+P(B)-P(A)P(B)=0.5+0.4-0.2=0.7$。注意到

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

故

$$P(C\mid(A\cup B))=\frac{P(B)}{P(A\cup B)}=\frac{0.4}{0.7}=\frac{4}{7}.$$

练习 1.53 下列命题中，正确的是（　）。

A. 若 A,B 互不相容，则 $\overline{A},\overline{B}$ 也互不相容　B. 若 A,B 相容，则 $\overline{A},\overline{B}$ 也相容　C. 若 A,B 独立，则 $\overline{A},\overline{B}$ 也独立　D. 若 A,B 独立，则 $\overline{A},\overline{B}$ 不相容

解 应选 C。

由独立性 $P(AB)=P(A)P(B)$：

$$P(\overline{A}\,\overline{B})=P(\overline{A\cup B})=1-P(A)-P(B)+P(AB)=(1-P(A))(1-P(B))=P(\overline{A})P(\overline{B}),$$

故 $\overline{A},\overline{B}$ 也独立。

A 错误：例如 $A=\varnothing$ 时 $\overline{A}=\Omega$，$\overline{A}\,\overline{B}=\overline{B}$ 一般不为空。B 错误：类似地，$\overline{A},\overline{B}$ 可相容也可不相容。D 错误：独立与相容没有必然关系。

练习 1.54 提高 设 A,B,C 相互独立，且 $P(A)\ne0$，$0<P(C)<1$。则下面四个事件中不独立的是（　）。

A. $\overline{A}\cup B$ 与 C　B. $\overline{AC}$ 与 $\overline{C}$　C. $\overline{A-B}$ 与 $\overline{C}$　D. $\overline{AB}$ 与 $\overline{C}$

解 应选 B。

[1]验证独立性不要看直觉，会被骗，还是要验证定义

A、C、D 中左边的事件均由 A,B 构成，而 A,B 各自与 C 独立，因此由 A,B 经过并、交、差运算构成的新事件也与 C 独立。

B 中，$\overline{AC}\cap\overline{C}=\overline{C}$（因为 $\overline{AC}\supset\overline{C}$）。若两者独立，需 $P(\overline{AC})\,P(\overline{C})=P(\overline{C})$，即 $P(\overline{AC})=1$。但

$$P(\overline{AC})=P(\overline{A})+P(\overline{C})-P(\overline{A})P(\overline{C})<1\quad(\text{因为}P(\overline{A})<1,P(\overline{C})<1),$$

矛盾。故 $\overline{AC}$ 与 $\overline{C}$ 不独立。

等价地，若独立则需 $P(\overline{C})=P(\overline{AC}\,\overline{C})=P(\overline{C})$ 恒成立，但由独立性 $P(AC)=P(A)P(C)>0$，可得 $\overline{AC}\neq\Omega$，从而 $P(\overline{AC})<1$。

1.7 独立重复试验与伯努利概型

定义 1.11 (n 重伯努利概型)

如果随机试验满足以下条件：

1. 每次试验只有两个可能结果，记为 A（成功）和 $\overline{A}$（失败）；
2. 每次试验中 $P(A)=p$（$0<p<1$）保持不变；
3. 各次试验相互独立，

则将这种试验独立重复进行 n 次的概率模型称为 $\boldsymbol{n}$ 重伯努利概型。 ♣

在 n 重伯努利概型中，事件 A 恰好发生 k 次的概率为

$$P\{A\text{ 恰好发生}k\text{ 次}\}=\mathrm{C}_n^k\,p^k\,(1-p)^{n-k},\quad k=0,1,\ldots,n.$$

如果用 X 表示 n 重伯努利概型中事件 A 发生的次数，则 X 服从参数为 n,p 的**二项分布**，记为 $X\sim B(n,p)$。

笔记 要善于判定 n 重伯努利概型：题目中出现“将……独立重复进行 n 次”“对……重复观察 n 次”等字样，或可以转化为 n 次独立重复试验的问题，都应考虑使用二项分布计算概率。例如射击问题、抛硬币问题、产品抽样检验问题等。

练习 1.55 对同一目标接连进行 3 次独立重复射击，假设至少命中目标一次的概率为 $\dfrac{7}{8}$，则每次射击命中目标的概率 $p=$ _____。

解 应填 $\dfrac{1}{2}$。

记 $A_i=\{\text{第}i\text{ 次命中}\}\ (i=1,2,3)$，$A_1,A_2,A_3$ 相互独立且 $P(A_i)=p$。由

$$P(A_1\cup A_2\cup A_3)=1-P(\overline{A_1}\,\overline{A_2}\,\overline{A_3})=1-(1-p)^3=\frac{7}{8},$$

得 $(1-p)^3=\dfrac{1}{8}$，故 $1-p=\dfrac{1}{2}$，$p=\dfrac{1}{2}$。

练习 1.56 某人向同一目标独立重复射击，每次命中目标的概率为 $p\ (0<p<1)$，则此人第 4 次射击恰好第 2 次命中目标的概率为（　）。

A. $3p(1-p)^2$　B. $6p(1-p)^2$　C. $3p^2(1-p)^2$　D. $6p^2(1-p)^2$

解 应选 C。

”第 4 次恰好第 2 次命中”意味着：前 3 次中有 1 次命中，第 4 次命中。设前 3 次命中次数为 $X\sim B(3,p)$，则

$$P\{X=1\}=\mathrm{C}_3^1\,p\,(1-p)^2=3p(1-p)^2.$$

第 4 次命中的概率为 p，故所求概率为

$$3p(1-p)^2\cdot p=3p^2(1-p)^2.$$

笔记 一般地，在独立重复试验中，第 n 次试验恰好第 k 次成功的概率为

$$\mathrm{C}_{n-1}^{k-1}p^k(1-p)^{n-k},$$

其中前 $n-1$ 次中有 $k-1$ 次成功，第 n 次成功。这实际上是负二项分布（帕斯卡分布）的一个特例。

练习 1.57 在四次独立试验中，事件 A 至少发生一次的概率为 $\dfrac{80}{81}$，则事件 A 在每次试验中发生的概率为（　）。　A. $\dfrac{1}{3}$　B. $\dfrac{2}{3}$　C. $\dfrac{2}{3}\sqrt[4]{5}$　D. $1-\dfrac{2}{3}\sqrt[4]{5}$

解 应选 B。设 X 为事件 A 在 4 次试验中发生的次数。

$$P\{X\geq 1\}=1-P\{X=0\}=1-(1-p)^4=\frac{80}{81},$$

故 $(1-p)^4=\dfrac{1}{81}$，即 $1-p=\dfrac{1}{3}$，$p=\dfrac{2}{3}$。

练习 1.58 三人独立地破译一个密码，他们能单独译出的概率分别为 $\dfrac{1}{5},\dfrac{1}{3},\dfrac{1}{4}$，则此密码被译出的概率为 _____。

解 应填 $\dfrac{3}{5}$。

记事件 A_i 为"第 i 个人译出密码"（$i=1,2,3$），B 为"密码被译出"。由独立性：

$$P(B)=P(A_1\cup A_2\cup A_3)=1-P(\overline{A_1}\,\overline{A_2}\,\overline{A_3})=1-\frac{4}{5}\times\frac{2}{3}\times\frac{3}{4}=\frac{3}{5}.$$

笔记 互不相容可简化并事件的概率计算，相互独立可简化交事件的概率计算。这里利用对偶律将事件的并转化为事件的交，从而利用独立性进行计算，这一方法经常用到。

练习 1.59 设甲、乙两人轮流射击同一目标，直到射中为止，每次命中率分别为 0.3 和 0.4，甲先射击。

(1) 目标是由甲射中的概率 p_1；

(2) 击中目标者射击了 k 次的概率 p_2。

解 (1) 甲在第一轮命中的概率为 0.3；若第一轮都未命中（概率 $0.7\times 0.6=0.42$），则回到初始状态。故

$$p_1=0.3+0.42\times p_1\implies p_1=\frac{0.3}{1-0.42}=\frac{15}{29}\approx 0.5172.$$

(2)"击中者射击了 k 次"意味着该射手在前 $k-1$ 轮均未命中，第 k 次命中。若甲击中：前 $k-1$ 轮甲乙均未命中（概率 0.42^{k-1}），甲第 k 次命中（概率 0.3）。若乙击中：前 $k-1$ 轮均未命中（概率 0.42^{k-1}），甲第 k 次未命中（概率 0.7），乙第 k 次命中（概率 0.4）。故

$$p_2=0.42^{k-1}\times 0.3+0.42^{k-1}\times 0.7\times 0.4=0.42^{k-1}\times 0.58\quad(k=1,2,\ldots).$$

第 2 章 一维随机变量及其分布

- 随机变量分布函数$F(x)=P(X\le x)$ 的性质
  - (1) 单调不减
  - (2) $0\le F(x)\le 1,\ F(-\infty)=0,\ F(+\infty)=1$
  - (3) 右连续: $F(x+0)=F(x)$
- 随机变量分类及概率分布
  - 离散型: 分布律
  - 连续型: 概率密度　　满足概率分布的性质
  - 混合型: 分布函数
- 常见随机变量及其分布
  - 离散型
    - $(0-1)$分布: $P(X=k)=p^k(1-p)^{1-k}\ (k=0,1)$
    - 二项分布: $P(X=k)=C_n^k p^k(1-p)^{n-k}\ (k=0,1,2,\ldots,n)$
    - 泊松分布: $P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}\ (k=0,1,2\cdots)$
  - 连续型
    - 均匀分布
      - $X\sim U(a,b)$
      - 性质: 区间概率等于该区间占整个区间的比例
    - 指数分布
      - $X\sim E(\lambda)$
      - 性质: 无记忆性
    - 正态分布
      - $X\sim N(\mu,\sigma^2)$
      - 性质
        - 参数含义: μ 为位置参数 (注意对称性), σ^2 为形状参数
        - 利用标准化求解概率
- 随机变量函数的概率分布
  - 离散型: 关键取值及对应概率
  - 连续型
    - 定义法: 求解函数的分布函数
    - 公式法: 适用单调的随机变量

内容提要

- 随机变量的定义与分类
- 分布函数的概念与性质
- 离散型随机变量及其分布律
- 连续型随机变量及其概率密度
- 常见离散分布：(0-1) 分布、二项分布、泊松分布
- 常见连续分布：均匀分布、指数分布、正态分布
- 随机变量函数的分布求法

2.1 随机变量与分布函数

2.1.1 随机变量

定义 2.1 (随机变量)

设随机试验 E 的样本空间为 $\Omega = \{\omega\}$。如果对每一个 $\omega \in \Omega$，都有唯一的实数 $X(\omega)$ 与之对应，并且对任意实数 x，集合

$$\{\omega \mid X(\omega) \le x,\ \omega \in \Omega\}$$

都是随机事件，则称定义在 Ω 上的实值单值函数 $X(\omega)$ 为随机变量，简记为 X。♣

一般用大写字母 $X, Y, Z, \ldots$ 或希腊字母 $\xi, \eta, \zeta, \ldots$ 来表示随机变量。

笔记 随机变量是"其值随试验结果而定"的变量，其定义有三个要点：

1. 它是定义在样本空间 Ω 上的实值单值函数——定义域是样本空间，而非实数集；
2. 对任意实数 x，$\{X \le x\}$ 必须是随机事件（即可定义概率）；
3. 取值在试验之前不能预知，且取各值有一定的概率。

随机事件是从静态的观点研究随机现象，而随机变量则引入了动态的观点（类似于高等数学中常量与变量的区别），使概率论的研究从对单个事件的孤立分析，上升到对随机变量整体规律的系统研究。

笔记 随机变量一定范围内的取值本质上与随机事件等价。例如，"掷骰子出现的点数"是一个随机变量 X，则 $X = 3$、$X \le 4$、X 为偶数等都是随机事件。

2.1.2 分布函数

定义 2.2 (分布函数)

设 X 是随机变量，x 是任意实数，称函数

$$F(x) = P\{X \le x\} \quad (-\infty < x < +\infty)$$

为随机变量 X 的分布函数，也称 X 服从 $F(x)$ 分布，记为 $X \sim F(x)$。♣

当 x 从 $-\infty$ 到 $+\infty$ 取遍所有实数时，$\{X \le x\}$ 就相应从不可能事件 $\varnothing$ 取到必然事件 Ω，因此分布函数**完整地描述了随机变量的概率规律**。

2.1.3 分布函数的性质

性质 设 $F(x)$ 为随机变量 X 的分布函数，则它满足以下三条性质：

1. 单调不减：对任意实数 $x_1 < x_2$，有 $F(x_1) \le F(x_2)$；
2. 有界性：$0 \le F(x) \le 1$，且
$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \to +\infty} F(x) = 1;$$
3. 右连续：对任意 $x_0 \in \mathbf{R}$，有 $\lim_{x \to x_0^+} F(x) = F(x_0)$。

注 上述三条性质也是**充要条件**：若函数 $F(x)$ 满足性质 (1)–(3)，则它一定是某个随机变量的分布函数。因此，这三条性质常用于判断一个函数能否作为分布函数。

2.1.4 分布函数与事件概率

分布函数最重要的应用是通过它来计算各种事件的概率。记 $F(a^-) = \lim_{x\to a^-} F(x)$ 为 F 在点 a 处的左极限值，则：

性质 设随机变量 X 的分布函数为 $F(x)$，则：

1. $P\{X \le a\} = F(a)$；
2. $P\{X < a\} = F(a^-)$；
3. $P\{X = a\} = F(a) - F(a^-)$；
4. $P\{a < X < b\} = F(b^-) - F(a)$；
5. $P\{a < X \le b\} = F(b) - F(a)$；
6. $P\{a \le X < b\} = F(b^-) - F(a^-)$；
7. $P\{a \le X \le b\} = F(b) - F(a^-)$；
8. $P\{X > a\} = 1 - F(a)$。

笔记 记忆技巧：事件中包含等号 $\le$ 或 $=$ 的一端取函数值 $F(\cdot)$，不包含等号的一端（开区间端）取左极限 $F(\cdot^-)$。例如：

- $P\{a < X \le b\}$ 中，b 端有等号取 $F(b)$，a 端无等号取 $F(a^-)$；
- $P\{X = a\} = F(a) - F(a^-)$ 即函数值减去左极限。

特别地，当 $F(x)$ 在点 a 处连续时，$F(a^-) = F(a)$，从而 $P\{X = a\} = 0$。

练习 2.1 设随机变量 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < 0, \\ \dfrac{1}{2}, & 0 \le x < 1, \\ 1 - e^{-x}, & x \ge 1. \end{cases}$$

求 $P\{X = 0\}$、$P\{0 < X \le 2\}$ 和 $P\{X > 1\}$。

解

1. $P\{X = 0\} = F(0) - F(0^-) = \dfrac{1}{2} - 0 = \dfrac{1}{2}$；
2. $P\{0 < X \le 2\} = F(2) - F(0) = (1 - e^{-2}) - \dfrac{1}{2} = \dfrac{1}{2} - e^{-2}$；
3. $P\{X > 1\} = 1 - F(1) = 1 - (1 - e^{-1}) = e^{-1}$。

练习 2.2 下列函数中，可以作为某个随机变量的分布函数的是（　）。

A. $F(x) = \dfrac{1}{\pi}\arctan x + \dfrac{1}{2}$, $x \in \mathbf{R}$　B. $F(x) = \dfrac{1}{1+x^2}$, $x \in \mathbf{R}$

C. $F(x) = \begin{cases} 0, & x < 0 \\ 1 - x, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$　D. $F(x) = \begin{cases} \sin x, & 0 \le x \le \pi \\ 0, & \text{其他} \end{cases}$

解 应选 A。

逐条验证分布函数的三条充要条件：

- A：$\arctan x$ 单调递增，故 $F(x)$ 单调不减；$F(-\infty) = 0$，$F(+\infty) = 1$；$\arctan x$ 连续，故右连续。三条均满足。
- B：$F(x)$ 在 $(-\infty, +\infty)$ 上不是单调不减的（例如 $F(0) = 1$，$F(1) = \frac{1}{2}$），排除。
- C：在 $[0, 1)$ 上 $F(x) = 1 - x$ 单调递减，不满足单调不减，排除。

- D：$\sin x$ 在 $[0,\pi]$ 上不是单调函数，不满足单调不减，排除。

练习 2.3 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}a+be^{-x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

求常数 a,b 及概率 $P\{|X|<2\}$。

解 $F(x)$ 中含有 2 个未知参数，依据分布函数的性质建立两个独立方程解之。

由分布函数的有界性，有

$$F(+\infty)=\lim_{x\to+\infty}F(x)=\lim_{x\to+\infty}(a+be^{-x})=a=1,$$

得 $a=1$。又由分布函数的右连续性，在 $x=0$ 处有

$$F(0)=\lim_{x\to 0^+}F(x)=a+b=0,$$

得 $b=-1$。所以

$$F(x)=\begin{cases}1-e^{-x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

从而

$$P\{|X|<2\}=P\{-2<X<2\}=F(2^-)-F(-2)=1-e^{-2}-0=1-e^{-2}.$$

2.2 离散型随机变量

定义 2.3 (离散型随机变量)

如果随机变量 X 只可能取有限个或可列无限个值 $x_1,x_2,\dots$，则称 X 为离散型随机变量。称

$$P\{X=x_i\}=p_i,\quad i=1,2,\dots$$

为 X 的分布律（或分布列、概率分布），记为 $X\sim p_i$。 ♣

分布律常用表格形式表示：

X	x_1	x_2	$\cdots$
P	p_1	p_2	$\cdots$

性质 数列 $\{p_i\}$ $(i=1,2,\dots)$ 是某离散型随机变量的分布律的充要条件：

1. $p_i\ge 0$ $(i=1,2,\dots)$；
2. $\sum\limits_{i=1}^{\infty}p_i=1$。

性质 设离散型随机变量 X 的分布律为 $P\{X=x_i\}=p_i$ $(i=1,2,\dots)$，则：

1. X 的分布函数为 $F(x)=\sum\limits_{x_i\le x}p_i$，它是阶梯函数，在 $x=x_i$ 处有跳跃，跳跃高度为 p_i；
2. $P\{X=x_i\}=F(x_i)-F(x_i^-)$；
3. 对实数轴上的任一集合 B，有 $P\{X\in B\}=\sum\limits_{x_i\in B}p_i$。

笔记 离散型随机变量的分布函数是阶梯函数，从图形上看，在每个取值点处有一个跳跃，跳跃的高度恰好等于该点的概率值。这意味着离散型随机变量的分布函数不是连续函数。

2.3 连续型随机变量

定义 2.4 (连续型随机变量)

如果随机变量 X 的分布函数 $F(x)$ 可以表示为

$$F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t \quad (x\in\mathbf{R}),$$

其中 $f(x)$ 是非负可积函数，则称 X 为连续型随机变量，称 $f(x)$ 为 X 的概率密度函数（简称概率密度），记为 $X\sim f(x)$。 ♣

性质 函数 $f(x)$ 为某一随机变量的概率密度的充要条件：$f(x)\geq 0$ 且 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x=1$。

性质 设连续型随机变量 $X\sim f(x)$，则：

1. 对任意实数 $a<b$，$P\{a<X\leq b\}=\int_a^b f(x)\,\mathrm{d}x=F(b)-F(a)$；
2. 连续型随机变量取一点的概率为零：$P\{X=c\}=0$，从而
$$P\{a<X<b\}=P\{a\leq X<b\}=P\{a<X\leq b\}=P\{a\leq X\leq b\};$$
3. 若 $f(x)$ 在点 x 处连续，则 $F'(x)=f(x)$。更一般地，
$$f(x)=\begin{cases}0, & x\text{ 为}F(x)\text{ 的不可导点},\\ F'(x), & x\text{ 为}F(x)\text{ 的可导点}.\end{cases}$$

注 (1) 概率密度没有 $0\leq f(x)\leq 1$ 的限制。例如 $X\sim U[0,0.1]$，则 $f(x)=10\ (x\in[0,0.1])$，$f(x)$ 可以大于 1。

(2) 概率密度反映的是概率在点 x 处的”密集程度”而非概率值本身。$P\{x<X\leq x+h\}\approx f(x)\cdot h$（当 h 很小时），因此 $f(x)$ 类似于”单位长度上的概率”。从几何上看，X 落入某区间的概率等于该区间上概率密度曲线下的曲边梯形面积。

(3) 连续型随机变量的分布函数一定是**连续**函数（但反之不成立：$F(x)$ 连续的随机变量不一定是连续型的）。

(4) 同一个分布函数的概率密度函数并不唯一——改变有限个点的值不影响积分结果。

笔记 除离散型和连续型外，还存在其他类型的随机变量。例如

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^3}{2}, & 0\leq x<1,\\ 1, & x\geq 1,\end{cases}$$

该分布函数连续但不绝对连续（因 $\int_0^1 F'(x)\,\mathrm{d}x=\dfrac{1}{2}\neq 1$），故对应的随机变量既非离散型又非连续型。

练习 2.4 分布函数性质判断 已知 $F(x)$ 是分布函数，下列函数：

（1）$aF(x)$ （$a>0,a\neq 1$）；　（2）$F(x)+F(-x)$；　（3）$F(x)-F(-x)$；　（4）$F(x)\cdot F(-x)$。

其中**不能作为**分布函数的个数是（　）。 A. 1　B. 2　C. 3　D. 4

解 应用分布函数的充要条件求解。由于

- (1) $aF(+\infty)=a\neq 1$；
- (2) $F(-\infty)+F(+\infty)=1\neq 0$；
- (3) $F(-\infty)-F(+\infty)=-1\neq 0$；

- (4) $F(+\infty)\cdot F(-\infty)=0\neq 1$;

故 (1)、(2)、(3)、(4) 都不能作为分布函数，答案选 **(D)**。

✍ **练习 2.5** 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ Ax^2+\dfrac{2}{3}x, & 0\le x<1,\\ 1, & x\ge 1.\end{cases}$$

求常数 A、X 的概率密度函数 $f(x)$ 以及 $P\{|X|<0.5\}$。

解 由连续型随机变量分布函数的连续性，有

$$\lim_{x\to 1^-}F(x)=A+\frac{2}{3}=1\implies A=\frac{1}{3}.$$

在 $[0,1)$ 上，$F(x)=\dfrac{x^2}{3}+\dfrac{2x}{3}$，故

$$f(x)=F'(x)=\frac{2x}{3}+\frac{2}{3},\quad x\in[0,1),$$

即

$$f(x)=\begin{cases}\dfrac{2x}{3}+\dfrac{2}{3}, & 0\le x<1,\\ 0, & \text{其他}.\end{cases}$$

$$P\{|X|<0.5\}=P\{-0.5<X<0.5\}=F(0.5)-F(-0.5)=\left(\frac{1}{12}+\frac{1}{3}\right)-0=\frac{5}{12}.$$

✍ **练习 2.6** 已知随机变量 X 的概率分布为

X	1	2	3
$P\{X=k\}$	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

且 $P\{X\ge 2\}=\dfrac{3}{4}$。求未知参数 θ 及 X 的分布函数 $F(x)$。

解 由 $P\{X\ge 2\}=P\{X=2\}+P\{X=3\}=2\theta(1-\theta)+(1-\theta)^2=1-\theta^2=\dfrac{3}{4}$，解得 $\theta=\pm\dfrac{1}{2}$。

又 $P\{X=2\}=2\theta(1-\theta)\ge 0$，当 $\theta=-\dfrac{1}{2}$ 时 $P\{X=2\}=-\dfrac{3}{2}<0$，故 $\theta=\dfrac{1}{2}$。从而得 X 的概率分布为

X	1	2	3
$P\{X=k\}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

X 的分布函数为

$$F(x)=P\{X\le x\}=\sum_{x_i\le x}P\{X=x_i\}=\begin{cases}0, & x<1,\\ \dfrac{1}{4}, & 1\le x<2,\\ \dfrac{3}{4}, & 2\le x<3,\\ 1, & x\ge 3.\end{cases}$$

✍ **练习 2.7** 已知离散型随机变量 X 的分布律为 $P\{X=k\}=c(0.75)^k$，$k=1,2,3,\ldots$，则常数 c 的值为_____。

解 应填 $\frac{1}{3}$。由离散型随机变量分布律的归一性：

$$\sum_{k=1}^{+\infty} P\{X=k\} = \sum_{k=1}^{+\infty} c\cdot(0.75)^k = c\cdot\frac{0.75}{1-0.75} = 3c = 1,\ \ c=\frac{1}{3}$$

练习 2.8 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ 0.4, & -1\le x<1,\\ 0.8, & 1\le x<3,\\ 1, & x\ge 3.\end{cases}$$

求：(1) X 的分布律；(2) $P\{X<2\mid X\neq 1\}$。

解 (1) 根据离散型随机变量的性质：$P\{X=k\}=F(k)-F(k^-)$。

$$P\{X=-1\}=F(-1)-F(-1^-)=0.4-0=0.4,$$
$$P\{X=1\}=F(1)-F(1^-)=0.8-0.4=0.4,$$
$$P\{X=3\}=F(3)-F(3^-)=1-0.8=0.2.$$

故 X 的分布律为：

X	-1	1	3
P	0.4	0.4	0.2

(2) 由条件概率公式：

$$P\{X<2\mid X\neq 1\}=\frac{P\{X<2,\,X\neq 1\}}{P\{X\neq 1\}}=\frac{P\{X=-1\}}{1-P\{X=1\}}=\frac{0.4}{0.6}=\frac{2}{3}.$$

练习 2.9 连续型分布函数性质 设 $F(x)$ 为某连续型随机变量的分布函数，$f(x)$ 为相应的密度函数，则（　）。

(A) $F(x)$ 连续但不一定可导；

(B) $F(x)$ 可导；

(C) $f(x)$ 连续；

(D) $0\leqslant f(x)\leqslant 1$。

解 答案选 **(A)**。分析：

- 连续型随机变量的分布函数 $F(x)$ 一定是连续的；
- 但 $F(x)$ 仅在密度 $f(x)$ 的连续点处可导；
- 密度函数 $f(x)$ 可以不连续（例如分段均匀分布）；
- 密度函数 $f(x)$ 可以大于 1（例如 $U(0,0.5)$ 的密度为 2）。

练习 2.10 已知 X 的分布函数为 $F(x)$，概率密度为 $f(x)$。当 $x\le 0$ 时，$f(x)$ 连续且 $f(x)=F(x)$。若 $F(0)=1$，求 $f(x)$。

解 由于 $F(x)$ 单调不减且 $F(0)=1$，故对任意 $x>0$，$F(x)=1$。又当 $x\le 0$ 时，$f(x)$ 连续，故 $f(x)=F'(x)$。已知 $f(x)=F(x)$，由微分方程 $F'(x)=F(x)$ 及初始条件 $F(0)=1$，解得 $F(x)=e^x\ (x\le 0)$，从而

$$F(x)=\begin{cases}e^x, & x<0,\\ 1, & x\ge 0,\end{cases}\qquad f(x)=\begin{cases}e^x, & x\le 0,\\ 0, & x>0.\end{cases}$$

笔记 当分布函数是不连续的分段函数时，为保证右连续性，分段小区间除第一个外都应是左闭右开的（等号跟在”大于等于”端）。

练习 2.11 已知随机变量 X 的概率密度为

$$f(x)=\begin{cases}2\left(1-\dfrac{1}{x^2}\right), & 1\le x\le 2,\\ 0, & \text{其他}.\end{cases}$$

求 X 的分布函数 $F(x)$。

解 分段积分：

- 当 $x<1$ 时，$F(x)=\displaystyle\int_{-\infty}^{x}0\,\mathrm{d}t=0$；
- 当 $1\le x<2$ 时，
$$F(x)=\int_1^x 2\left(1-\frac{1}{t^2}\right)\mathrm{d}t=2\left[t+\frac{1}{t}\right]_1^x=2\left(x+\frac{1}{x}-2\right);$$
- 当 $x\ge 2$ 时，$F(x)=\displaystyle\int_1^2 2\left(1-\frac{1}{t^2}\right)\mathrm{d}t=2\left[\frac{3}{2}-1\right]=1$。

故

$$F(x)=\begin{cases}0, & x<1,\\ 2\left(x+\dfrac{1}{x}-2\right), & 1\le x<2,\\ 1, & x\ge 2.\end{cases}$$

练习 2.12 下列函数中，可以作为某个随机变量的概率密度函数的是（　）。

A. $f(x)=\begin{cases}2(1-|x|), & |x|\le 1\\ 0, & |x|>1\end{cases}$　　B. $f(x)=\begin{cases}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, & x\ge 0\\ 0, & x<0\end{cases}$

C. $f(x)=\begin{cases}x, & -1<x<0\\ \frac{3}{4}x, & 0\le x<2\\ 0, & \text{其他}\end{cases}$　　D. $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}\ (\lambda>0)$

解 应选 D。

逐条验证概率密度函数的充要条件（非负性和归一性）：

- A：$\displaystyle\int_{-\infty}^{\infty}f(x)\,\mathrm{d}x=2\int_0^1 2(1-x)\,\mathrm{d}x=4\left[x-\frac{x^2}{2}\right]_0^1=2\ne 1$，错误；
- B：正态分布的概率密度在 $x<0$ 时截断为 0，但 $\displaystyle\int_0^{+\infty}f(x)\,\mathrm{d}x=\frac{1}{2}\ne 1$，错误；
- C：当 $-1<x<0$ 时，$x<0$，故 $f(x)=x<0$，不满足非负性，错误；
- D：指数分布的标准形式，满足 $\displaystyle\int_0^{+\infty}\lambda e^{-\lambda x}\,\mathrm{d}x=1$，正确。

练习 2.13 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}A\sin x, & 0\le x\le\pi\\ 0, & \text{其他}\end{cases}$$

求：(1) 系数 A；(2) X 的分布函数 $F(x)$；(3) $P\{0<X<1\}$。

解 (1) 由归一性条件：

$$\int_0^{\pi} A\sin x\,\mathrm{d}x = A[-\cos x]_0^{\pi} = A(1+1) = 2A = 1,\quad A = \frac{1}{2}$$

(2) 当 $x<0$ 时，$F(x)=0$；当 $0\le x<\pi$ 时，

$$F(x) = \int_0^x \frac{1}{2}\sin t\,\mathrm{d}t = \frac{1}{2}[-\cos t]_0^x = \frac{1}{2}(1-\cos x);$$

当 $x\ge\pi$ 时，$F(x)=1$。故

$$F(x) = \begin{cases} 0, & x<0, \\ \frac{1}{2}(1-\cos x), & 0\le x<\pi, \\ 1, & x\ge\pi. \end{cases}$$

(3)

$$P\{0<X<1\} = \int_0^1 \frac{1}{2}\sin t\,\mathrm{d}t = \frac{1}{2}(1-\cos 1).$$

练习 2.14 设 $F_1(x), F_2(x)$ 是随机变量的分布函数，$f_1(x), f_2(x)$ 是相应的概率密度，则下列选项中正确的是（　）[1]。

A. $F_1(x)+F_2(x)$ 是分布函数　B. $F_1(x)\cdot F_2(x)$ 是分布函数　C. $f_1(x)+f_2(x)$ 是概率密度　D. $f_1(x)\cdot f_2(x)$ 是概率密度

解 应选 B。逐条验证：

- A：$F_1(+\infty)+F_2(+\infty)=2\ne 1$，不满足边界条件，排除；
- B：$F_1(x)\cdot F_2(x)$ 单调不减（两个非负单调不减函数之积仍单调不减），右连续，且 $F_1(-\infty)F_2(-\infty)=0$，$F_1(+\infty)F_2(+\infty)=1$。三条均满足；
- C：$\displaystyle\int_{-\infty}^{+\infty}[f_1(x)+f_2(x)]\,\mathrm{d}x = 2\ne 1$，排除；
- D：$\displaystyle\int_{-\infty}^{+\infty} f_1(x)f_2(x)\,\mathrm{d}x \ne 1$（一般不等于两个积分的乘积），排除。

2.4 常见离散型随机变量及分布

2.4.1 (0-1) 分布

定义 2.5 ((0-1) 分布)

随机变量 X 只可能取 0 与 1 两个值，它的分布律为

$$P\{X=k\} = p^k(1-p)^{1-k},\quad k=0,1\ (0<p<1).$$

则称 X 服从参数为 p 的 (0-1) 分布，记为 $X\sim B(1,p)$。 ♣

笔记 (0-1) 分布是描述伯努利试验结果的计数变量——事件发生计为 1，不发生计为 0。它是所有离散型分布的基础。

练习 2.15（0-1）分布的分布函数 设 X 服从 0-1 分布，其分布律为 $P\{X=k\}=p^k(1-p)^{1-k}, k=0,1, 0<p<1$，求 X 的分布函数。

[1] 验证概率密度和分布函数的定义即可

解 0-1 分布（伯努利分布）的分布函数为：

$$F(x)=\begin{cases}0, & x<0,\\ 1-p, & 0\leqslant x<1,\\ 1, & x\geqslant 1.\end{cases}$$

解释：

- 当 $x<0$ 时，$F(x)=P\{X\leqslant x\}=0$；
- 当 $0\leqslant x<1$ 时，$F(x)=P\{X=0\}=1-p$；
- 当 $x\geqslant 1$ 时，$F(x)=P\{X=0\}+P\{X=1\}=1-p+p=1$。

2.4.2 二项分布

定义 2.6 (二项分布)

若随机变量 X 的分布律为

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

其中 $0<p<1$，则称 X 服从参数为 (n,p) 的二项分布，记为 $X\sim B(n,p)$。 ♣

性质 若 $X\sim B(n,p)$，则令 $Y=n-X$，有 $Y\sim B(n,1-p)$。

笔记 二项分布描述了 n 重独立伯努利试验中事件 A 发生的次数。当 $n=1$ 时，二项分布退化为 (0-1) 分布。

练习 2.16 二项分布参数求解 设 $X\sim B(2,p)$，$Y\sim B(4,p)$，且 $P(X\geqslant 1)=\dfrac{5}{9}$，则 $P(Y\geqslant 1)=$（　）。

A. $\dfrac{65}{81}$　B. $\dfrac{16}{81}$　C. 1　D. $\dfrac{4}{7}$

解 由题设 $P(X\geqslant 1)=\dfrac{5}{9}=1-P(X=0)=1-\mathrm{C}_2^0p^0(1-p)^2$，即 $(1-p)^2=\dfrac{4}{9}$，解得 $p=\dfrac{1}{3}$，从而有

$$P(Y\geqslant 1)=1-P(Y=0)=1-\mathrm{C}_4^0p^0(1-p)^4=1-\left(\frac{2}{3}\right)^4=\frac{65}{81},$$

答案选 **(A)**。

练习 2.17 概率密度与二项分布结合 设随机变量 X 的概率密度为 $f(x)=\begin{cases}2x, & 0<x<1,\\ 0, & \text{其他},\end{cases}$，$Y$ 表示对 X 的 3 次独立重复试验中事件 $\left\{X\leqslant\dfrac{1}{2}\right\}$ 出现的次数，则 $P\{Y=2\}=$ _____。

解 首先计算单次试验中事件发生的概率：

$$p=P\left\{X\leqslant\frac{1}{2}\right\}=\int_0^{\frac{1}{2}}2x\,\mathrm{d}x=\left[x^2\right]_0^{\frac{1}{2}}=\frac{1}{4}.$$

Y 服从二项分布 $B\left(3,\dfrac{1}{4}\right)$，故

$$P\{Y=2\}=\mathrm{C}_3^2\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)=3\cdot\frac{1}{16}\cdot\frac{3}{4}=\frac{9}{64}.$$

答案：$\mathbf{\dfrac{9}{64}}$。

2.4.3 泊松分布

定义 2.7 (泊松分布)

若随机变量 X 的分布律为

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda},\quad \lambda>0;\ k=0,1,2,\dots,$$

则称 X 服从参数为 λ 的泊松分布，记为 $X\sim P(\lambda)$。 ♣

定理 2.1 (泊松定理)

设 $\lambda>0$ 是一个常数，n 是任意正整数，设 $\lim_{n\to\infty}np_n=\lambda$，则对于任一固定的非负整数 k，有

$$\lim_{n\to\infty}\mathrm{C}_n^k p_n^k(1-p_n)^{n-k}=\frac{\lambda^k e^{-\lambda}}{k!}.$$

♡

笔记 泊松定理表明：当 n 很大，p 很小，且 $\lambda=np$ 适中时，二项分布可用泊松分布近似，即

$$\mathrm{C}_n^k p^k(1-p)^{n-k}\approx\frac{\lambda^k}{k!}e^{-\lambda}.$$

一般地，当 $n\geq 20$, $p\leq 0.05$ 时，用泊松近似公式逼近二项分布效果比较好，特别当 $n\geq 100$, $np\leq 10$ 时，逼近效果更佳。这一结论在计算机出现前极大简化了二项分布的计算。

练习 2.18 设 X 服从泊松分布，且 $P\{X=1\}=P\{X=2\}$，则 $P\{X=4\}=$ _____。

解 由泊松分布的分布律 $P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}$，由条件 $P\{X=1\}=P\{X=2\}$ 得

$$\lambda e^{-\lambda}=\frac{\lambda^2e^{-\lambda}}{2}\implies\lambda=2\implies P\{X=4\}=\frac{2^4e^{-2}}{4!}=\frac{2}{3}e^{-2}.$$

练习 2.19 泊松分布 假设某段时间内来百货公司的顾客数服从参数为 λ 的泊松分布，而每位顾客购买电视机的概率均为 p，且顾客之间是否购买电视机相互独立。求该段时间内百货公司售出 k 台电视机的概率（假设每位顾客至多买一台电视机）。

解 设 A_i 表示这段时间内到达百货公司的顾客数为 i $(i=0,1,2,\dots)$，则 $P(A_i)=\dfrac{\lambda^i}{i!}e^{-\lambda}$。当顾客数为 i 时，售出的电视机数 $Y|A_i\sim B(i,p)$，故

$$P\{Y=k\mid A_i\}=\mathrm{C}_i^k\,p^k(1-p)^{i-k},\quad k\leq i.$$

由全概率公式：

$$\begin{aligned}P\{Y=k\}&=\sum_{i=0}^{\infty}P(A_i)\,P\{Y=k\mid A_i\}\\&=\sum_{i=k}^{\infty}\frac{\lambda^i}{i!}e^{-\lambda}\cdot\frac{i!}{(i-k)!k!}\,p^k(1-p)^{i-k}\\&=\frac{e^{-\lambda}(\lambda p)^k}{k!}\sum_{i=k}^{\infty}\frac{[\lambda(1-p)]^{i-k}}{(i-k)!}.\end{aligned}$$

令 $m=i-k$，则

$$\sum_{i=k}^{\infty}\frac{[\lambda(1-p)]^{i-k}}{(i-k)!}=\sum_{m=0}^{\infty}\frac{[\lambda(1-p)]^m}{m!}=e^{\lambda(1-p)}.$$

因此

$$P\{Y=k\}=\frac{e^{-\lambda}(\lambda p)^k}{k!}\cdot e^{\lambda(1-p)}=\frac{(\lambda p)^k}{k!}e^{-\lambda p},\quad k=0,1,2,\dots.$$

即 $Y \sim P(\lambda p)$。这表明：若顾客数服从泊松分布，每位顾客独立地以概率 p 购买，则购买总数仍服从泊松分布，参数变为 λp。

练习 2.20 超几何分布 20 个产品中有 5 个不合格品，若从中随机取出 8 个，求其中不合格品数 X 的概率分布。

解 X 服从超几何分布 $H(N=20, M=5, n=8)$，其分布律为

$$P\{X=k\} = \frac{\mathrm{C}_5^k \mathrm{C}_{15}^{8-k}}{\mathrm{C}_{20}^8}, \quad k=0,1,2,3,4,5.$$

具体计算：

X	0	1	2	3	4	5
P	$\frac{\mathrm{C}_5^0\mathrm{C}_{15}^8}{\mathrm{C}_{20}^8}$	$\frac{\mathrm{C}_5^1\mathrm{C}_{15}^7}{\mathrm{C}_{20}^8}$	$\frac{\mathrm{C}_5^2\mathrm{C}_{15}^6}{\mathrm{C}_{20}^8}$	$\frac{\mathrm{C}_5^3\mathrm{C}_{15}^5}{\mathrm{C}_{20}^8}$	$\frac{\mathrm{C}_5^4\mathrm{C}_{15}^4}{\mathrm{C}_{20}^8}$	$\frac{\mathrm{C}_5^5\mathrm{C}_{15}^3}{\mathrm{C}_{20}^8}$
数值	0.0511	0.2554	0.3973	0.2384	0.0542	0.0036

其中 $\mathrm{C}_{20}^8 = 125970$。答案：$X$ 的超几何分布如上表所示。

2.4.4 几何分布

定义 2.8 (几何分布)

若随机变量 X 的分布律为

$$P\{X=k\} = (1-p)^{k-1}p, \quad k=1,2,\ldots;\ 0<p<1,$$

则称 X 服从参数为 p 的几何分布，记为 $X \sim G(p)$。 ♣

笔记 几何分布描述了在 n 重伯努利试验中，事件 A 首次出现所需的试验次数。它具有与指数分布类似的无记忆性。

练习 2.21 几何分布的无记忆性 设 X 为取正整数值的离散型随机变量，证明：X 是服从参数为 p 的几何分布的充要条件是 X 具有无记忆性，即 $P(X=k+n \mid X>k)$ 与 k 无关（k,n 为任意正整数）。

解 必要性证明：若 X 服从几何分布，则 $P(X=k)=p(1-p)^{k-1}$, $k=1,2,\ldots$。对任意正整数 k,n：

$$\begin{aligned} P(X=k+n \mid X>k) &= \frac{P(X=k+n, X>k)}{P(X>k)} = \frac{P(X=k+n)}{P(X>k)} \\ &= \frac{p(1-p)^{k+n-1}}{\sum_{j=k+1}^{\infty} p(1-p)^{j-1}} = \frac{p(1-p)^{k+n-1}}{(1-p)^k} = p(1-p)^{n-1}. \end{aligned}$$

结果与 k 无关，故几何分布具有无记忆性。

充分性证明：设 $P(X=k+1 \mid X>k)=p$ 与 k 无关。记 $q=1-p$，则

$$P(X>k) = q \cdot P(X>k-1) = q^k \cdot P(X>0).$$

由 $P(X>0)=1$ 可得

$$P(X=k) = P(X>k-1) - P(X>k) = q^{k-1} - q^k = p(1-p)^{k-1}.$$

这正是参数为 p 的几何分布。

2.4.5 超几何分布

定义 2.9 (超几何分布)

若随机变量 X 的分布律为

$$P\{X=k\}=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n},\quad \max\{0,n-N+M\}\le k\le \min\{M,n\},$$

其中 M,N,n 为正整数且 $M\le N$, $n\le N$，则称 X 服从参数为 (n,N,M) 的超几何分布，记为 $X\sim H(n,N,M)$。♣

笔记 超几何分布描述的是不放回抽样场景：设总体中有 N 个产品，其中 M 个不合格品。若从中随机不放回地抽取 n 个，则其中不合格品的个数 X 服从超几何分布。当 $n\ll N$ 时（抽取个数远小于总体），每次抽取后总体不合格品率 $p=\frac{M}{N}$ 改变甚微，此时不放回抽样可近似看作放回抽样，超几何分布可用二项分布 $H(n,N,M)\approx B(n,\frac{M}{N})$ 近似。

2.4.6 极值分布

性质 对于 n 个相互独立的随机变量 $\xi_1,\xi_2,\dots,\xi_n$，其分布函数分别为 $F_1(x),F_2(x),\dots,F_n(x)$，则：

- 最大值 $\eta_1=\max\{\xi_1,\xi_2,\dots,\xi_n\}$ 的分布函数为

$$F_{\eta_1}(x)=P\{\eta_1\le x\}=P\left(\bigcap_{k=1}^n \xi_k\le x\right)=\prod_{k=1}^n F_k(x);$$

- 最小值 $\eta_2=\min\{\xi_1,\xi_2,\dots,\xi_n\}$ 的分布函数为

$$F_{\eta_2}(x)=P\{\eta_2\le x\}=1-P\{\eta_2>x\}=1-P\left(\bigcap_{k=1}^n \xi_k>x\right)=1-\prod_{k=1}^n(1-F_k(x)).$$

练习 2.22 最小值分布（系统寿命） 一辆新车装有 4 个轮胎并配有 1 个备胎，当使用的轮胎损坏时就用备胎更换。设所有轮胎的寿命均服从参数为 0.2 的指数分布且相互独立，记 T 为开始使用备胎的时间，求 T 的分布函数。

解 设四个在用轮胎的寿命分别为 X_1,X_2,X_3,X_4，则 $X_i\overset{\text{i.i.d.}}{\sim}E(0.2)$。

开始使用备胎的时间为

$$T=\min\{X_1,X_2,X_3,X_4\}.$$

单个轮胎的分布函数：

$$F_X(x)=1-e^{-0.2x},\quad x\geqslant 0.$$

T 的分布函数：

$$\begin{aligned}F_T(t)&=P\{T\leqslant t\}=1-P\{T>t\}=1-P\{X_1>t,X_2>t,X_3>t,X_4>t\}\\&=1-[P\{X_1>t\}]^4=1-[1-F_X(t)]^4=1-\left[e^{-0.2t}\right]^4=1-e^{-0.8t},\quad t\geqslant 0.\end{aligned}$$

当 $t<0$ 时，$F_T(t)=0$。故

$$F_T(t)=\begin{cases}0, & t<0,\\ 1-e^{-0.8t}, & t\geqslant 0.\end{cases}$$

即 $T\sim E(0.8)$。

2.5 常见连续型随机变量及分布

练习 2.23 偶函数概率密度的性质证明 设连续型随机变量 X 的概率密度 $f(x)$ 是偶函数,证明:(1)$F(-a)=1-F(a)$; (2) $P(|X|<a)=2F(a)-1$; (3)$P(|X|>a)=2[1-F(a)]$

证明 由于 $f(x)$ 是偶函数，即 $f(-x)=f(x)$。

(1) 由分布函数定义，令 $t=-x$:

$$F(-a)=\int_{-\infty}^{-a} f(x)\,\mathrm{d}x=\int_{+\infty}^{a} f(-t)\,(-dt)=\int_{a}^{+\infty} f(t)\,dt=1-\int_{-\infty}^{a} f(t)\,dt=1-F(a).$$

(2) 利用 (1) 的结果:

$$P(|X|<a)=P(-a<X<a)=F(a)-F(-a)=F(a)-(1-F(a))=2F(a)-1.$$

(3) 由概率的互补性:

$$P(|X|>a)=1-P(|X|\leqslant a)=1-[F(a)-F(-a)]=1-[F(a)-(1-F(a))]=2[1-F(a)].$$

2.5.1 均匀分布

定义 2.10 (均匀分布)

若随机变量 X 的概率密度为

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

则称 X 在区间 (a,b) 内服从均匀分布，记为 $X\sim U(a,b)$。其分布函数为

$$F(x)=\begin{cases}0, & x<a,\\ \frac{x-a}{b-a}, & a\leq x<b,\\ 1, & x\geq b.\end{cases}$$

♣

笔记 均匀分布描述了随机变量 X 在区间 (a,b) 内任意子区间取值的概率只与子区间长度成正比，与其位置无关。几何概型是均匀分布的实际背景。

练习 2.24 设随机变量 X 满足 $1\leq X\leq 4$，且 $P\{X=1\}=\frac{1}{4}$，$P\{X=4\}=\frac{1}{4}$，当 X 在区间 $(1,4)$ 内取值时，X 服从均匀分布，求 X 的分布函数 $F(x)$。

解 当 $x<1$ 时，$F(x)=0$；当 $1\leq x<4$ 时，

$$F(x)=P\{X\leq 1\}+P\{1<X\leq x\}=\frac{1}{4}+\frac{x-1}{4-1}\cdot\frac{1}{2}=\frac{5+2x}{12};$$

当 $x\geq 4$ 时，$F(x)=1$。故

$$F(x)=\begin{cases}0, & x<1,\\ \frac{5+2x}{12}, & 1\leq x<4,\\ 1, & x\geq 4.\end{cases}$$

练习 2.25 设 $X\sim U(a,b)$（$a>0$），且 $P\{X>4\}=\frac{1}{2}$，$P\{0<X<3\}=\frac{1}{4}$，求 a,b 及 $P\{1<X<5\}$。

解 由 $P\{X>4\}=\frac{1}{2}$ 及均匀分布的对称性，得 $\frac{a+b}{2}=4$，即 $a+b=8$。

又 $P\{3<X<4\}=1-P\{X\geq 4\}-P\{X\leq 3\}=1-\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$，而 $P\{3<X<4\}=\frac{4-3}{b-a}$，故 $\frac{1}{b-a}=\frac{1}{4}$，得 $b-a=4$。

联立 $\begin{cases} a+b=8, \\ b-a=4, \end{cases}$ 解得 $\begin{cases} a=2, \\ b=6. \end{cases}$

因此 $P\{1<X<5\}=P\{2<X<5\}=\frac{5-2}{6-2}=\frac{3}{4}$。

练习 2.26 设随机变量 X 在区间 $(0,8)$ 上服从均匀分布。

(1) 求方程 $y^2+2y+X=0$ 有实根的概率；

(2) 若对随机变量 X 进行 4 次独立观察，记 Y 为上述方程有实根的次数，求 Y 的数学期望和方差。

解 (1) 方程有实根的条件为判别式 $\Delta=2^2-4X\geq 0$，即 $X\leq 1$。由均匀分布性质：

$$P\{X\leq 1\}=\frac{1-0}{8}=\frac{1}{8}.$$

(2) 易知 $Y\sim B\left(4,\frac{1}{8}\right)$，故

$$E(Y)=np=4\times\frac{1}{8}=\frac{1}{2},\quad D(Y)=np(1-p)=4\times\frac{1}{8}\times\frac{7}{8}=\frac{7}{16}.$$

练习 2.27 均匀分布与二项分布结合 设随机变量 X 服从区间 $(2,5)$ 上的均匀分布，则对 X 进行 3 次独立观测中，至少有 2 次的观测值大于 3 的概率为 _____。

解 首先计算单次观测值大于 3 的概率：

$$P\{X>3\}=\frac{5-3}{5-2}=\frac{2}{3}.$$

设 Y 表示 3 次独立观测中观测值大于 3 的次数，则 $Y\sim B\left(3,\frac{2}{3}\right)$。所求概率：

$$P\{Y\geqslant 2\}=P\{Y=2\}+P\{Y=3\}=\mathrm{C}_3^2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)+\mathrm{C}_3^3\left(\frac{2}{3}\right)^3=3\cdot\frac{4}{9}\cdot\frac{1}{3}+\frac{8}{27}=\frac{12}{27}+\frac{8}{27}=\frac{20}{27}.$$

答案：$\mathbf{\frac{20}{27}}$。

2.5.2 指数分布

定义 2.11 (指数分布)

若随机变量 X 的概率密度为

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

其中 $\lambda>0$，则称 X 服从参数为 λ 的指数分布，记为 $X\sim E(\lambda)$。其分布函数为

$$F(x)=\begin{cases} 1-e^{-\lambda x}, & x\geq 0, \\ 0, & x<0. \end{cases}$$

♣

性质 [无记忆性] 指数分布具有无记忆性：对于任意 $s,t>0$，有

$$P\{X>s+t\mid X>s\}=P\{X>t\}.$$

这意味着，对于一个已经工作了 s 小时的元件，其再工作至少 t 小时的概率，与从最开始起就至少工作 t 小时的概率相同。

练习 2.28 设某医院门诊的内科看病窗口有 5 个，患者平均就诊时间为 10 分钟，假定每个患者的就诊时间服从指数分布。求：

(1) 患者就诊时间超过 10 分钟的概率；

(2) 若某时刻 5 个窗口都有病人就诊，则恰有 2 人的看病时间超过 10 分钟的概率。

解 由题意，平均就诊时间为 10 分钟，故 $\dfrac{1}{\lambda}=10$，即 $\lambda=0.1$。记患者就诊时间为 X，则 $X\sim E(0.1)$。

(1)

$$P\{X\geq 10\}=\int_{10}^{\infty}0.1\,e^{-0.1x}\,\mathrm{d}x=e^{-1}.$$

(2) 设 5 个窗口都有病人就诊时，恰有 2 人看病时间超过 10 分钟为事件 B。每位患者就诊时间超过 10 分钟的概率为 e^{-1}，则

$$P(B)=\mathrm{C}_5^2\,(e^{-1})^2\,(1-e^{-1})^3=10\,e^{-2}\,(1-e^{-1})^3.$$

练习 2.29 设随机变量 X 服从参数为 $\lambda(\lambda>0)$ 的指数分布，且 $P\{X\leq 1\}=\frac{1}{2}$。

(1) 求参数 λ；

(2) 求 $P\{X>2\mid X>1\}$。

解 (1) 根据指数分布的分布函数，$P\{X\leq 1\}=1-e^{-\lambda}=\frac{1}{2}$，解得 $\lambda=\ln 2$。

(2) 由无记忆性，

$$P\{X>2\mid X>1\}=P\{X>1\}=1-P\{X\leq 1\}=\frac{1}{2}.$$

或者直接计算：

$$P\{X>2\mid X>1\}=\frac{P\{X>2\}}{P\{X>1\}}=\frac{e^{-2\ln 2}}{e^{-\ln 2}}=e^{-\ln 2}=\frac{1}{2}.$$

练习 2.30（指数分布无记忆性应用） 设某电子元件的寿命服从参数 $\lambda=\dfrac{1}{4}$ 的指数分布（单位：万小时），某系统并联了两个这种电子元件，计算：

(1) 已知一个元件已经工作了三万小时，求再正常工作四万小时的概率；

(2) 求系统寿命大于五万小时的概率。

解 设元件寿命为 X，则 $X\sim E(\lambda)$，其中 $\lambda=\dfrac{1}{4}$。

(1) 由指数分布的无记忆性：

$$P\{X>3+4\mid X>3\}=P\{X>4\}=e^{-\lambda\cdot 4}=e^{-1}.$$

(2) 并联系统寿命大于 5 万小时，即至少一个元件正常工作。设两个元件寿命为 X_1,X_2，则系统寿命为 $\max\{X_1,X_2\}$：

$$\begin{aligned}P\{\max\{X_1,X_2\}>5\}&=1-P\{X_1\leqslant 5,X_2\leqslant 5\}\\&=1-\left(1-e^{-\lambda\cdot 5}\right)^2\\&=1-\left(1-e^{-\frac{5}{4}}\right)^2.\end{aligned}$$

练习 2.31 指数分布与系统更换 设一批元件的寿命独立同分布，其概率密度函数为

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

系统开始工作时只有一个元件工作，在它损坏后立即更换一个新元件接替工作，求系统从开始工作到时刻 $T(T>0)$ 为止，恰更换了一个元件的概率。

解 设两个元件的寿命分别为 X_1 和 X_2，则 $X_1, X_2 \overset{\text{i.i.d.}}{\sim} E(1)$，"恰更换了一个元件"等价于事件：

$$\{X_1 < T,\ X_1 + X_2 > T\}.$$

所求概率：

$$\begin{aligned} P\{X_1 < T,\ X_1 + X_2 > T\} &= \int_0^T f_{X_1}(x) \cdot P\{X_2 > T - x\}\, dx = \int_0^T e^{-x} \cdot e^{-(T-x)}\, dx \\ &= \int_0^T e^{-T}\, dx = Te^{-T}. \end{aligned}$$

✍ **练习 2.32** 某元件的工作寿命 X（小时）服从参数为 $\lambda(\lambda > 0)$ 的指数分布。

(1) 求该元件正常工作 t 小时的概率；

(2) 已知该元件已正常工作 10 小时，求在此基础上再工作 10 小时的概率（$\lambda = 0.01$）。

解 由题意 $X \sim F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & \text{其他}, \end{cases}$

(1) $P\{X > t\} = 1 - F(t) = e^{-\lambda t}$。

(2) 由无记忆性，

$$P\{X > 20 \mid X > 10\} = P\{X > 10\} = e^{-0.01 \times 10} = e^{-0.1}.$$

或者直接计算：

$$P\{X > 20 \mid X > 10\} = \frac{P\{X > 20\}}{P\{X > 10\}} = \frac{e^{-0.2}}{e^{-0.1}} = e^{-0.1}.$$

2.5.3 正态分布

定义 2.12 (正态分布)

若随机变量 X 的概率密度为

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

其中 μ 为常数，$\sigma > 0$ 为常数，则称 X 服从参数为 (μ, σ^2) 的正态分布，记为 $X \sim N(\mu, \sigma^2)$。 ♣

定义 2.13 (标准正态分布)

当 $\mu = 0,\ \sigma = 1$ 时，称 X 服从标准正态分布，记为 $X \sim N(0, 1)$。其概率密度和分布函数分别记为

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Phi(x) = \int_{-\infty}^{x} \varphi(t) dt.$$

♣

性质 标准正态分布具有以下性质：

1. $\Phi(0) = \frac{1}{2},\ \Phi(-a) = 1 - \Phi(a)$；
2. 若 $X \sim N(\mu, \sigma^2)$，则 $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$；若 $X \sim N(\mu, \sigma^2)$，则 $\frac{X-\mu}{\sigma} \sim N(0, 1)$；
3. 若 $X \sim N(\mu, \sigma^2)$，则 $P\{X \leq \mu\} = P\{X > \mu\} = \frac{1}{2}$；
4. $P\{a < X < b\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$。

✍ **练习 2.33 正态分布对称点** 设随机变量 $X \sim N(3, 4)$，且 $P(X < C) = P(X > C)$，则常数 $C =$ _____。

解 对于正态分布 $N(\mu,\sigma^2)$，其概率密度函数关于均值 μ 对称，因此

$$P(X<\mu)=P(X>\mu)=\frac{1}{2}.$$

由题设条件 $P(X<C)=P(X>C)$ 可知，C 必为分布的对称中心，即

$$C=\mu=3.$$

✍ **练习 2.34** 设随机变量 $X\sim N(10,\sigma^2)$，且 $P\{10<X<20\}=0.35$，则 $P\{X<0\}=$ _____。

解 由正态分布的对称性，

$$P\{10<X<20\}=\Phi\left(\frac{20-10}{\sigma}\right)-\Phi(0)=\Phi\left(\frac{10}{\sigma}\right)-0.5=0.35,$$

得 $\Phi\left(\frac{10}{\sigma}\right)=0.85$。因此

$$P\{X<0\}=\Phi\left(\frac{0-10}{\sigma}\right)=1-\Phi\left(\frac{10}{\sigma}\right)=0.15.$$

✍ **练习 2.35** 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)$，则随着 σ 的增大，$P\{|X-\mu|<1\}$（　）。

(A) 单调增加　(B) 单调减少　(C) 保持不变　(D) 先增后减

解 应选 (B)。讨论 $P\{|X-\mu|<1\}$ 的单调性应将其标准化后进行。由

$$P\{|X-\mu|<1\}=P\left\{\left|\frac{X-\mu}{\sigma}\right|<\frac{1}{\sigma}\right\}=2\Phi\left(\frac{1}{\sigma}\right)-1,$$

其中 $\Phi(u)$ 为单调增加函数，而 $u=\frac{1}{\sigma}$ 随 σ 增大而单调减少，故 $\Phi\left(\frac{1}{\sigma}\right)$ 单调减少，从而 $2\Phi\left(\frac{1}{\sigma}\right)-1$ 单调减少，故选择 (B)。

✍ **练习 2.36** 设随机变量 X 服从正态分布 $N(0,\sigma^2)$，对给定的 $\alpha(0<\alpha<1)$，数 u_α 满足 $P\{X>u_\alpha\}=\alpha$，若 $P\{|X|<x\}=\alpha$，则 x 等于（　）。(A) $u_{\frac{\alpha}{2}}$　(B) $u_{\frac{1-\alpha}{2}}$　(C) $u_{1-\frac{\alpha}{2}}$　(D) $u_{1-\alpha}$

解 应选 (B)。由 $P\{|X|<x\}=\alpha$ 可得

$$P\{-x<X<x\}=\Phi(x/\sigma)-\Phi(-x/\sigma)=2\Phi(x/\sigma)-1=\alpha,$$

即 $\Phi(x/\sigma)=\frac{1+\alpha}{2}$。而 $P\{X>u_{(1-\alpha)/2}\}=\frac{1-\alpha}{2}$，故 $x=u_{\frac{1-\alpha}{2}}$。

✍ **练习 2.37 正态分布标准差比较** 设 $X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)$，且 $P(|X-\mu_1|<1)>P(|Y-\mu_2|<1)$，则（　）。(A)$\sigma_1<\sigma_2$　(B)$\sigma_1>\sigma_2$　(C) $\mu_1<\mu_2$　(D) $\mu_1>\mu_2$。

解 答案选 **(A)**。标准化得：

$$P\left\{\frac{|X-\mu_1|}{\sigma_1}<\frac{1}{\sigma_1}\right\}>P\left\{\frac{|Y-\mu_2|}{\sigma_2}<\frac{1}{\sigma_2}\right\}\Rightarrow 2\Phi\left(\frac{1}{\sigma_1}\right)-1>2\Phi\left(\frac{1}{\sigma_2}\right)-1$$

$$\Rightarrow\Phi\left(\frac{1}{\sigma_1}\right)>\Phi\left(\frac{1}{\sigma_2}\right)\implies\frac{1}{\sigma_1}>\frac{1}{\sigma_2}\Rightarrow\sigma_1<\sigma_2.$$

✍ **练习 2.38 正态分布标准差比较** 设随机变量 X 服从正态分布 $N(\mu_1,\sigma_1^2)$，随机变量 Y 服从正态分布 $N(\mu_2,\sigma_2^2)$，$\sigma_1,\sigma_2>0$，且 $P\{|X-\mu_1|<1\}>P\{|Y-\mu_2|<1\}$，则必有（　）。

A. $\sigma_1<\sigma_2$　B. $\sigma_1>\sigma_2$
C. $\mu_1<\mu_2$　D. $\mu_1>\mu_2$

解 正态分布标准化：

$$P\left\{\frac{|X-\mu_1|}{\sigma_1}<\frac{1}{\sigma_1}\right\}>P\left\{\frac{|Y-\mu_2|}{\sigma_2}<\frac{1}{\sigma_2}\right\},$$

$$2\Phi\left(\frac{1}{\sigma_1}\right)-1>2\Phi\left(\frac{1}{\sigma_2}\right)-1\implies\Phi\left(\frac{1}{\sigma_1}\right)>\Phi\left(\frac{1}{\sigma_2}\right),$$

由 $\Phi(x)$ 单调递增，得 $\dfrac{1}{\sigma_1}>\dfrac{1}{\sigma_2}$，即 $\sigma_1<\sigma_2$，答案选 **(A)**。

✍ **练习 2.39 正态分布参数求解** 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)(\sigma>0)$，且二次方程 $y^2+4y+X=0$

无实根的概率为 $\frac{1}{2}$，则 $\mu =$ _____。

解 二次方程无实根的充要条件是判别式 $\Delta = 16 - 4X < 0$，即 $X > 4$。由题意 $P\{X > 4\} = \frac{1}{2}$。对于正态分布 $N(\mu, \sigma^2)$，有

$$P\{X > \mu\} = \frac{1}{2}.$$

因此 $\mu = 4$（因为 $P\{X > 4\} = 1/2$ 说明 4 是正态分布的中位数，对于正态分布中位数等于均值）。

2.6 常见分布的期望与方差

表 2.1: 常见分布的数学期望与方差

名称与记号	分布列或概率密度	数学期望	方差
(0-1) 分布 $B(1,p)$	$P\{X=1\}=p,\ P\{X=0\}=1-p$	p	$p(1-p)$
二项分布 $B(n,p)$	$P\{X=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k}$	np	$np(1-p)$
泊松分布 $P(\lambda)$	$P\{X=k\}=e^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ
几何分布 $G(p)$	$P\{X=k\}=(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布 $H(n,N,M)$	$P\{X=k\}=\frac{\mathrm{C}_M^k\mathrm{C}_{N-M}^{n-k}}{\mathrm{C}_N^n}$	$\frac{nM}{N}$	$\frac{nM}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$
均匀分布 $U(a,b)$	$f(x)=\frac{1}{b-a},\ a<x<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $E(\lambda)$	$f(x)=\lambda e^{-\lambda x},\ x>0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
正态分布 $N(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2

2.7 典型例题

练习 2.40 拉普拉斯分布变换 设随机变量 X 的概率密度为

$$f(x) = \frac{1}{2\lambda}e^{-\frac{|x|}{\lambda}}, \quad -\infty < x < +\infty,\ \lambda > 0.$$

(1) 求方程 $y^2 + 2Xy + \lambda^2 = 0$ 有实根的概率；

(2) 求 $Z = e^{-|X|}$ 的概率密度。

解 (1) 方程有实根的条件是判别式

$$\Delta = (2X)^2 - 4\lambda^2 = 4(X^2 - \lambda^2) \geqslant 0 \Leftrightarrow |X| \geqslant \lambda.$$

所求概率：

$$\begin{aligned} P\{|X| \geqslant \lambda\} &= \int_{-\infty}^{-\lambda} f(x)\,dx + \int_{\lambda}^{\infty} f(x)\,dx = 2\int_{\lambda}^{\infty} \frac{1}{2\lambda}e^{-\frac{x}{\lambda}}\,dx \\ &= \frac{1}{\lambda}\int_{\lambda}^{\infty} e^{-\frac{x}{\lambda}}\,dx = e^{-1}. \end{aligned}$$

(2) $Z = e^{-|X|}$ 的取值范围是 $(0,1]$。当 $0 < z < 1$ 时：

$$\begin{aligned} F_Z(z) &= P\{Z \leqslant z\} = P\{e^{-|X|} \leqslant z\} = P\{|X| \geqslant -\ln z\} \\ &= 2\int_{-\ln z}^{\infty} \frac{1}{2\lambda}e^{-\frac{x}{\lambda}}\,dx = z^{\frac{1}{\lambda}}. \end{aligned}$$

当 $z \leqslant 0$ 时，$F_Z(z) = 0$；当 $z \geqslant 1$ 时，$F_Z(z) = 1$。

求导得密度函数：

$$f_Z(z)=\begin{cases}\dfrac{1}{\lambda}z^{\frac{1}{\lambda}-1}, & 0<z<1,\\ 0, & \text{其他}.\end{cases}$$

练习 2.41 设随机变量 X 的概率密度为 $f(x)=\begin{cases}2^{-x}\ln 2, & x>0,\\ 0, & \text{其他},\end{cases}$ 对 X 进行独立重复的观测，直到第 2 个大于 3 的观测值出现时停止，记 Y 为观测次数，求 Y 的概率分布。

解 记 p 为" 观测值大于 3" 的概率，则

$$p=P\{X>3\}=\int_3^{+\infty}2^{-x}\ln 2\,\mathrm{d}x=\frac{1}{8}.$$

依题意，Y 为离散型随机变量，取值为 $2,3,\ldots$，则 Y 的概率分布为

$$P\{Y=k\}=\mathrm{C}_{k-1}^{1}p(1-p)^{k-2}\cdot p=(k-1)p^2(1-p)^{k-2}=(k-1)\left(\frac{1}{8}\right)^2\left(\frac{7}{8}\right)^{k-2},\quad k=2,3,\ldots.$$

练习 2.42 分布函数计算 设随机变量 X 的分布函数

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{1}{2}, & 0\leqslant x<1,\\ 1-e^{-x}, & x\geqslant 1,\end{cases}$$

则 $P\{0\leqslant X\leqslant 1\}=$ _____，$P\{0<X<1\}=$ _____。

解

$$P\{0\leqslant X\leqslant 1\}=F(1)-F(0^-)=1-e^{-1};$$

$$P\{0<X<1\}=F(1^-)-F(0)=\frac{1}{2}-\frac{1}{2}=0.$$

练习 2.43 离散型分布求解 离散型随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ 0.4, & -1\leqslant x<2,\\ 0.8, & 2\leqslant x<3,\\ 1, & x\geqslant 3,\end{cases}$$

则 X 的分布列为 _____，$P\{X\leqslant 3.2\}=$ _____，方程 $x^2+Xx+1=0$ 有实根的概率为 _____。

解 由分布函数可得分布列：

$$P\{X=-1\}=F(-1)-F(-1^-)=0.4-0=0.4,$$

$$P\{X=2\}=F(2)-F(2^-)=0.8-0.4=0.4,$$

$$P\{X=3\}=F(3)-F(3^-)=1-0.8=0.2.$$

故分布列：$X\sim\begin{pmatrix}-1 & 2 & 3\\ 0.4 & 0.4 & 0.2\end{pmatrix}$。

$P\{X\leqslant 3.2\}=1$（因为最大取值为 3）。

方程有实根 $\Leftrightarrow \Delta = X^2 - 4 \geqslant 0 \Leftrightarrow X \geqslant 2$ 或 $X \leqslant -2$（后者不可能），

$$P(X \geqslant 2) = P\{X = 2\} + P\{X = 3\} = 0.4 + 0.2 = 0.6.$$

答案：分布列如上；**1**；**0.6**。

练习 2.44 连续型概率密度参数求解 设连续型随机变量 X 的概率密度为

$$f(x) = \begin{cases} \dfrac{A}{x^2}, & x > 100, \\ 0, & x \leqslant 100, \end{cases}$$

则 $A =$ _____，$P\{100 < X \leqslant 150\} =$ _____，$P\{X \geqslant 1000\} =$ _____，$P\{X = 1000\} =$ _____，X 的分布函数 $F(x) =$ _____。

解 由归一性条件 $\displaystyle\int_{100}^{+\infty} \frac{A}{x^2}\,\mathrm{d}x = 1$ 得

$$A\left[-\frac{1}{x}\right]_{100}^{+\infty} = \frac{A}{100} = 1 \Rightarrow A = 100.$$

相关概率计算：

$$P\{100 < X \leqslant 150\} = \int_{100}^{150} \frac{100}{x^2}\,\mathrm{d}x = 100\left(\frac{1}{100} - \frac{1}{150}\right) = \frac{1}{3},$$

$$P\{X \geqslant 1000\} = \int_{1000}^{+\infty} \frac{100}{x^2}\,\mathrm{d}x = 100 \cdot \frac{1}{1000} = \frac{1}{10},$$

$$P\{X = 1000\} = 0 \quad (\text{连续型随机变量单点概率为 } 0).$$

分布函数：

$$F(x) = \int_{100}^{x} \frac{100}{t^2}\,dt = 1 - \frac{100}{x}, \quad x \geqslant 100,$$

故

$$F(x) = \begin{cases} 0, & x < 100, \\ 1 - \dfrac{100}{x}, & x \geqslant 100. \end{cases}$$

答案：**100**；$\mathbf{\frac{1}{3}}$；$\mathbf{\frac{1}{10}}$；**0**；分布函数如上。

练习 2.45 反正弦分布 设连续型随机变量 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < -a, \\ A + B\arcsin\dfrac{x}{a}, & -a \leqslant x \leqslant a, \\ 1, & x \geqslant a, \end{cases}$$

则 $A =$ _____，$B =$ _____，X 的概率密度 $f(x) =$ _____，$P\left\{-\dfrac{a}{2} < X < \dfrac{a}{2}\right\} =$ _____。

解 由分布函数的连续性：

$$\begin{cases} \displaystyle\lim_{x \to (-a)^+} F(x) = A + B\arcsin(-1) = A - \frac{\pi}{2}B = 0, \\ \displaystyle\lim_{x \to a^-} F(x) = A + B\arcsin(1) = A + \frac{\pi}{2}B = 1. \end{cases}$$

解得 $A = \dfrac{1}{2}$，$B = \dfrac{1}{\pi}$。

概率密度：

$$f(x)=F'(x)=\begin{cases}\dfrac{1}{\pi\sqrt{a^2-x^2}}, & -a<x<a,\\ 0, & \text{其他}.\end{cases}$$

概率计算：

$$P\left\{-\frac{a}{2}<X<\frac{a}{2}\right\}=F\left(\frac{a}{2}\right)-F\left(-\frac{a}{2}\right)=\frac{1}{2}+\frac{1}{\pi}\arcsin\frac{1}{2}-\left[\frac{1}{2}+\frac{1}{\pi}\arcsin\left(-\frac{1}{2}\right)\right]=\frac{2}{\pi}\cdot\frac{\pi}{6}=\frac{1}{3}.$$

练习 2.46 同分布随机变量参数求解 设随机变量 X 和 Y 同分布，X 的概率密度为

$$f(x)=\begin{cases}\dfrac{3}{8}x^2, & 0<x<2,\\ 0, & \text{其他}.\end{cases}$$

已知事件 $A=\{X>a\}$ 和 $B=\{Y>a\}$ 独立，且 $P(A\cup B)=\dfrac{3}{4}$，则常数 $a=$ _____。

解 首先计算 $P(A)=P\{X>a\}$：

$$P\{X>a\}=\int_a^2\frac{3}{8}x^2\,\mathrm{d}x=\left[\frac{x^3}{8}\right]_a^2=1-\frac{a^3}{8}.$$

由于 X 和 Y 同分布，故 $P(A)=P(B)$。由独立性及加法公式：

$$P(A\cup B)=P(A)+P(B)-P(AB)=2P(A)-[P(A)]^2=\frac{3}{4}.$$

设 $p=P(A)$，解方程 $2p-p^2=\dfrac{3}{4}$ 得 $p=\dfrac{1}{2}$（舍去 $p=\dfrac{3}{2}$）。于是

$$1-\frac{a^3}{8}=\frac{1}{2}\implies\frac{a^3}{8}=\frac{1}{2}\implies a^3=4\implies a=\sqrt[3]{4}.$$

练习 2.47 分段概率密度的分布函数 随机变量 X 的概率密度为

$$f(x)=\begin{cases}x, & 0\leqslant x<1,\\ 2-x, & 1\leqslant x<2,\\ 0, & \text{其他}.\end{cases}$$

求 X 的分布函数 $F(x)$。

解 分段计算分布函数 $F(x)=\int_{-\infty}^x f(t)\,dt$：

1. 当 $x<0$ 时：$F(x)=0$；
2. 当 $0\leqslant x<1$ 时：
$$F(x)=\int_0^x t\,dt=\frac{x^2}{2};$$
3. 当 $1\leqslant x<2$ 时：
$$F(x)=\int_0^1 t\,dt+\int_1^x(2-t)\,dt=\frac{1}{2}+\left[2t-\frac{t^2}{2}\right]_1^x=-\frac{x^2}{2}+2x-1;$$
4. 当 $x\geqslant 2$ 时：$F(x)=1$。

综上

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^2}{2}, & 0\leqslant x<1,\\ -\dfrac{x^2}{2}+2x-1, & 1\leqslant x<2,\\ 1, & x\geqslant 2.\end{cases}$$

练习 2.48 分布函数性质判断 下列函数可作为随机变量的分布函数的是（　）。

A. $F(x)=\begin{cases}0, & x<-1,\\ \dfrac{1}{3}, & -1\leqslant x\leqslant 2,\\ 1, & x>2,\end{cases}$　B. $F(x)=\begin{cases}0, & x<0,\\ \sin x, & 0\leqslant x<\pi,\\ 1, & x\geqslant \pi,\end{cases}$

C. $F(x)=\begin{cases}0, & x<0,\\ \dfrac{x+2}{5}, & 0\leqslant x<2,\\ 1, & x\geqslant 2,\end{cases}$　D. $F(x)=\dfrac{3}{4}+\dfrac{1}{2\pi}\arctan x.$

解 选项 C 正确。分析：

- 选项 A：在 $x=2$ 处，$F(2)=\dfrac{1}{3}$，但 $\lim\limits_{x\to 2^+}F(x)=1$，不满足右连续性；
- 选项 B：$\sin x$ 在 $[0,\pi)$ 上不单调（先增后减），违反分布函数的单调不减性；
- 选项 C：满足分布函数四条性质：单调不减、右连续、$F(-\infty)=0$、$F(+\infty)=1$；
- 选项 D：$F(-\infty)=\dfrac{3}{4}-\dfrac{1}{4}=\dfrac{1}{2}\neq 0$，不满足极限性质。

练习 2.49 反正切分布函数 已知随机变量 X 的分布函数为

$$F(x)=\begin{cases}A_1, & x<-a,\\ A_2+A_3\arctan\dfrac{x}{a}, & x\in[-a,a],\\ A_4, & x>a,\end{cases}$$

a 为已知常数，求：

(1) 使 $F(x)$ 连续的常数 A_i $(i=1,2,3,4)$；

(2) 随机变量 X 的概率密度函数；

(3) 二次方程 $t^2+Xt+\dfrac{a^2}{16}=0$ 的两个根均为正根的概率。

解 (1) 由分布函数极限性质：

$$\lim_{x\to-\infty}F(x)=0,\quad \lim_{x\to+\infty}F(x)=1\Rightarrow A_1=0,\ A_4=1.$$

由连续性条件：

$$\begin{cases}F(-a)=A_2+A_3\arctan(-1)=A_2-\dfrac{\pi}{4}A_3=0,\\ F(a)=A_2+A_3\arctan(1)=A_2+\dfrac{\pi}{4}A_3=1.\end{cases}$$

解得 $A_2=\dfrac{1}{2}$，$A_3=\dfrac{2}{\pi}$。

(2) 概率密度为分布函数的导数：

$$f(x)=F'(x)=\begin{cases}\dfrac{2a}{\pi(x^2+a^2)}, & -a\leqslant x\leqslant a,\\ 0, & \text{其他}.\end{cases}$$

(3) 方程有两个正根的条件：

$$\begin{cases}\Delta=X^2-\dfrac{a^2}{4}\geqslant 0,\\ X<0 \quad (\text{由韦达定理，根的和}=-X>0).\end{cases}$$

即 $X\leqslant -\dfrac{a}{2}$。所求概率：

$$P\left\{X\leqslant -\frac{a}{2}\right\}=F\left(-\frac{a}{2}\right)=\frac{1}{2}-\frac{2}{\pi}\arctan\frac{1}{2}.$$

练习 2.50 分布函数线性组合 设 $F_1(x),F_2(x)$ 分别为随机变量 X_1,X_2 的分布函数，为使 $aF_1(x)-bF_2(x)$ 是某一随机变量的分布函数，在下列给定的各组数值中应取（　）。
(A)$a=\dfrac{3}{5},b=-\dfrac{2}{5}$　(B)$a=\dfrac{2}{3},b=\dfrac{2}{3}$　(C)$a=-\dfrac{1}{2},b=\dfrac{3}{2}$　(D)$a=\dfrac{1}{2},b=-\dfrac{3}{2}$。

解 答案选 **(A)**。由分布函数性质，必须满足 $F(+\infty)=1$，即

$$F(+\infty)=aF_1(+\infty)-bF_2(+\infty)=a-b=1.$$

只有选项 (A) 满足 $a-b=\dfrac{3}{5}-\left(-\dfrac{2}{5}\right)=1$。其他选项：

- (B): $a-b=0\neq 1$；
- (C): $a-b=-2\neq 1$；
- (D): $a-b=2\neq 1$。

2.8 一维随机变量函数的分布

2.8.1 基本概念

设 X 为随机变量，函数 $y=g(x)$ 在随机变量 X 的取值范围内有定义，则以随机变量 X 作为自变量的函数 $Y=g(X)$ 也是一个随机变量，称为随机变量 X 的**函数**。例如：

$$Y=aX^2+bX+c,\quad Y=|X-a|,\quad Y=\begin{cases}X, & X\leq 1,\\ 1, & X>1\end{cases}$$

等都是随机变量的函数。本节主要研究如何由已知随机变量 X 的分布来求其函数 $Y=g(X)$ 的分布。

笔记 随机变量的函数在概率论与数理统计中具有重要应用。例如，在物理实验中，我们可能直接测量的是电压 X，但关心的却是功率 $Y=X^2/R$；在金融领域，资产的收益率经过某种变换后得到新的随机变量。因此，掌握随机变量函数的分布求法是解决实际问题的关键。

2.8.2 离散型随机变量函数的分布

设 X 为离散型随机变量，其概率分布为

$$P\{X=x_i\}=p_i,\quad i=1,2,\ldots$$

则 X 的函数 $Y=g(X)$ 也是离散型随机变量。若 $g(x_i)$ 的值互不相同，则 Y 的分布律为

$$P\{Y=g(x_i)\}=p_i,\quad i=1,2,\ldots$$

即

$$Y\sim\begin{pmatrix} g(x_1) & g(x_2) & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}.$$

如果有若干个 $g(x_i)$ 的值相同，则需要将这些相同值对应的概率相加合并。具体来说，若 $g(x_{i_1})=g(x_{i_2})=\cdots=g(x_{i_k})=y_j$，则

$$P\{Y=y_j\}=\sum_{t=1}^{k}P\{X=x_{i_t}\}=\sum_{t=1}^{k}p_{i_t}.$$

性质 离散型随机变量函数的分布求解步骤：

1. 确定 $Y=g(X)$ 的所有可能取值 y_j；
2. 对每个 y_j，找出所有满足 $g(x_i)=y_j$ 的 x_i；
3. 将相应的概率 p_i 相加，得到 $P\{Y=y_j\}$。

2.8.3 连续型随机变量函数的分布

设 X 为连续型随机变量，其概率密度函数为 $f_X(x)$，分布函数为 $F_X(x)$。对于连续型随机变量 X 的函数 $Y=g(X)$，需要根据 Y 的类型（离散型或连续型）采用不同的方法。

2.8.3.1 当 Y 为离散型时

若 $Y=g(X)$ 只能取有限个或可列个值，则 Y 是离散型随机变量。此时应先确定 Y 的所有可能取值，然后通过 X 的概率分布计算 Y 取各个值的概率。

具体来说，若 Y 的可能取值为 $y_1,y_2,\ldots,y_n,\ldots$，则

$$P\{Y=y_j\}=P\{g(X)=y_j\}=\int_{\{x:g(x)=y_j\}}f_X(x)\,\mathrm{d}x.$$

这通常涉及到对 X 的取值空间按 $g(x)$ 的取值进行划分。

2.8.3.2 当 Y 为连续型时

若 $Y=g(X)$ 的取值充满一个区间（或若干区间的并），则 Y 是连续型随机变量。此时主要有两种求解方法：分布函数法（定义法）和公式法。

2.8.3.2.1 方法一：分布函数法（通用方法） 根据分布函数的定义，先求 Y 的分布函数 $F_Y(y)$，再对 y 求导得到概率密度 $f_Y(y)$。

具体步骤如下：

1. 对任意实数 y，求 Y 的分布函数：

$$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=\int_{\{x:g(x)\le y\}}f_X(x)\,\mathrm{d}x.$$

这一步的关键是解不等式 $g(x)\le y$，确定积分区域 $\{x:g(x)\le y\}$。

2. 对 $F_Y(y)$ 关于 y 求导（在可导点处）：

$$f_Y(y) = F_Y'(y).$$

3. 检查 $f_Y(y)$ 是否满足概率密度函数的基本条件：非负性和归一性。

笔记 分布函数法是求连续型随机变量函数分布的最通用方法，无论 $g(x)$ 是单调还是非单调、是连续还是分段连续，都适用。虽然计算过程可能相对复杂，但其思想直观、逻辑严密，是理解随机变量函数分布的基础。

2.8.3.2.2 方法二：公式法（单调函数情形） 当 $y = g(x)$ 是严格单调函数时，可以使用公式法更快捷地求得 Y 的概率密度。

定理 2.2 (随机变量函数的密度公式)

设连续型随机变量 X 的概率密度为 $f_X(x)$，函数 $y = g(x)$ 在 X 的取值区间 (a, b) 上严格单调且可导，其反函数为 $x = h(y)$，且 $h(y)$ 连续可导。则 $Y = g(X)$ 的概率密度为

$$f_Y(y) = \begin{cases} f_X[h(y)] \cdot |h'(y)|, & \alpha < y < \beta, \\ 0, & \text{其他}, \end{cases}$$

其中 (α, β) 是函数 $y = g(x)$ 的值域，即

$$\alpha = \min\left\{\lim_{x\to a^+} g(x), \lim_{x\to b^-} g(x)\right\}, \quad \beta = \max\left\{\lim_{x\to a^+} g(x), \lim_{x\to b^-} g(x)\right\}.$$

♡

证明 考虑 $g(x)$ 严格单调递增的情形（单调递减类似）：

$$F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\} = P\{X \le h(y)\} = F_X[h(y)].$$

两边对 y 求导，由链式法则得

$$f_Y(y) = F_Y'(y) = f_X[h(y)] \cdot h'(y).$$

由于 $h'(y) > 0$（递增情形），故 $f_Y(y) = f_X[h(y)] \cdot |h'(y)|$。对于单调递减情形，$h'(y) < 0$，此时

$$F_Y(y) = P\{g(X) \le y\} = P\{X \ge h(y)\} = 1 - F_X[h(y)],$$

求导后得 $f_Y(y) = -f_X[h(y)] \cdot h'(y) = f_X[h(y)] \cdot |h'(y)|$。综上，公式成立。

注 公式法的适用条件：

1. $g(x)$ 必须是严格单调函数（递增或递减）；
2. $g(x)$ 在 X 的取值区间内可导；
3. 若 $g(x)$ 在若干区间上分段单调，则需要在每个单调区间上分别应用公式法，然后将结果相加（这实际上等价于分布函数法）。

当 $g(x)$ 不满足严格单调条件时，必须使用分布函数法。

2.8.4 一般情形

若 X 既非离散型也非连续型（或形式未知），则只能使用分布函数的定义来求 Y 的分布：

$$F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\}.$$

这种方法不依赖于 X 的具体类型，具有完全的通用性。

2.9 典型例题

练习 2.51 线性变换与平方变换 设连续型随机变量 X 的概率密度函数为

$$f_X(x)=\begin{cases}1, & -\dfrac{1}{2}<x<\dfrac{1}{2},\\ 0, & \text{其他},\end{cases}$$

求下列随机变量 Y 的概率密度函数：

(1) $Y=a+bX$, 其中 $b\neq 0$；

(2) $Y=X^2$；

(3) $Y=|X|$。

解 (1) 线性变换：使用公式法最为简便。函数 $y=a+bx$ 是严格单调函数，其反函数为 $x=h(y)=\dfrac{y-a}{b}$，导数为 $h'(y)=\dfrac{1}{b}$。

当 $b>0$ 时：

$$f_Y(y)=f_X\left(\frac{y-a}{b}\right)\cdot\frac{1}{b},\quad \text{当}-\frac{1}{2}<\frac{y-a}{b}<\frac{1}{2}.$$

解得 $a-\dfrac{b}{2}<y<a+\dfrac{b}{2}$。因此

$$f_Y(y)=\begin{cases}\dfrac{1}{b}, & a-\dfrac{b}{2}<y<a+\dfrac{b}{2},\\ 0, & \text{其他}.\end{cases}$$

当 $b<0$ 时，$|h'(y)|=-\dfrac{1}{b}$，得到

$$f_Y(y)=\begin{cases}-\dfrac{1}{b}, & a+\dfrac{b}{2}<y<a-\dfrac{b}{2},\\ 0, & \text{其他}.\end{cases}$$

可以统一表示为 $f_Y(y)=\dfrac{1}{|b|}$，在区间 $\left(a-\dfrac{|b|}{2},a+\dfrac{|b|}{2}\right)$ 上。

(2) 平方变换：函数 $y=x^2$ 在 $(-\dfrac{1}{2},\dfrac{1}{2})$ 上不是单调函数，必须使用分布函数法。

由于 $Y=X^2\geq 0$，当 $y\leq 0$ 时，$F_Y(y)=0$。

当 $0<y<\dfrac{1}{4}$ 时：

$$F_Y(y)=P\{X^2\leq y\}=P\{-\sqrt{y}\leq X\leq\sqrt{y}\}=\int_{-\sqrt{y}}^{\sqrt{y}}1\,\mathrm{d}x=2\sqrt{y}.$$

当 $y\geq\dfrac{1}{4}$ 时，$F_Y(y)=1$。

求导得概率密度：

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{y}}, & 0<y<\dfrac{1}{4},\\ 0, & \text{其他}.\end{cases}$$

(3) 绝对值变换：同样使用分布函数法。由于 $Y=|X|\geq 0$，当 $y<0$ 时，$F_Y(y)=0$。

当 $0 \le y < \frac{1}{2}$ 时：

$$F_Y(y) = P\{|X| \le y\} = P\{-y \le X \le y\} = \int_{-y}^{y} 1\,\mathrm{d}x = 2y.$$

当 $y \ge \frac{1}{2}$ 时，$F_Y(y) = 1$。

求导得：

$$f_Y(y) = \begin{cases} 2, & 0 \le y < \frac{1}{2}, \\ 0, & \text{其他}. \end{cases}$$

练习 2.52 离散型随机变量的函数 设随机变量 X 的分布律为

$$P\{X=-1\} = \frac{1}{4}, \quad P\{X=0\} = \frac{1}{2}, \quad P\{X=1\} = \frac{1}{4},$$

求 $Y = X^2$ 的分布。

解 $Y = X^2$ 的可能取值为 0 和 1。

当 $Y = 0$ 时：

$$P\{Y=0\} = P\{X^2=0\} = P\{X=0\} = \frac{1}{2}.$$

当 $Y = 1$ 时：

$$P\{Y=1\} = P\{X^2=1\} = P\{X=1\text{或}X=-1\} = P\{X=1\} + P\{X=-1\} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

因此，Y 的分布律为

$$Y \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

练习 2.53 概率积分变换 设随机变量 X 的分布函数 $F_X(x)$ 是严格单调增加函数，其反函数 $F_X^{-1}(y)$ 存在，$Y = F_X(X)$。证明：Y 服从区间 $(0,1)$ 上的均匀分布。

证明 由于 $F_X(x) \in [0,1]$，故 Y 的取值范围是 $(0,1)$。

当 $y < 0$ 时，$F_Y(y) = P\{Y \le y\} = 0$。

当 $0 \le y < 1$ 时：

$$F_Y(y) = P\{F_X(X) \le y\} = P\{X \le F_X^{-1}(y)\} = F_X[F_X^{-1}(y)] = y.$$

当 $y \ge 1$ 时，$F_Y(y) = 1$。

综上，Y 的分布函数为

$$F_Y(y) = \begin{cases} 0, & y < 0, \\ y, & 0 \le y < 1, \\ 1, & y \ge 1, \end{cases}$$

这正是均匀分布 $U(0,1)$ 的分布函数，故 $Y \sim U(0,1)$。

注 概率积分变换是统计模拟和假设检验中的重要工具。它表明，任何连续型随机变量都可以通过其分布函数变换为标准均匀分布。反之，若 $U \sim U(0,1)$，则 $X = F_X^{-1}(U)$ 服从分布 $F_X(x)$。这为随机数生成提供了理论基础。

练习 2.54 分段随机变量的函数 在区间 $(0,2)$ 上随机取一点，将该区间分成两段，较短一段的长度记为 X，较长一段的长度记为 Y。令 $Z = \frac{Y}{X}$。

(1) 求 X 的概率密度；(2) 求 Z 的概率密度。

解 (1) 设随机点的坐标为 $V \sim U(0,2)$，则 $X = \min\{V, 2-V\}$。显然 $0 \le X \le 1$。

当 $x < 0$ 时，$F_X(x) = 0$；当 $x \ge 1$ 时，$F_X(x) = 1$。

当 $0 \le x < 1$ 时：

$$\begin{aligned} F_X(x) &= P\{\min\{V, 2-V\} \le x\} = 1 - P\{\min\{V, 2-V\} > x\} \\ &= 1 - P\{x < V < 2-x\} = 1 - \frac{2-x-x}{2} = x. \end{aligned}$$

因此

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{其他}. \end{cases}$$

(2) 由于 $Y = 2 - X$，故 $Z = \dfrac{Y}{X} = \dfrac{2-X}{X} = \dfrac{2}{X} - 1$。当 $X \in (0,1)$ 时，$Z \in (1, +\infty)$。

函数 $z = \dfrac{2}{x} - 1$ 在 $(0,1)$ 上严格单调递减，反函数为 $x = \dfrac{2}{1+z}$，导数为 $\dfrac{\mathrm{d}x}{\mathrm{d}z} = -\dfrac{2}{(1+z)^2}$。

由公式法：

$$f_Z(z) = f_X\left(\frac{2}{1+z}\right) \cdot \left| -\frac{2}{(1+z)^2} \right| = \frac{2}{(1+z)^2}, \quad z > 1.$$

故

$$f_Z(z) = \begin{cases} \dfrac{2}{(1+z)^2}, & z > 1, \\ 0, & \text{其他}. \end{cases}$$

练习 2.55 分段密度函数的变换 设随机变量 X 的概率密度为

$$f_X(x) = \begin{cases} \dfrac{1}{2}, & -1 < x < 0, \\ \dfrac{1}{4}, & 0 \le x < 2, \\ 0, & \text{其他}, \end{cases}$$

令 $Y = X^2$，求 Y 的概率密度 $f_Y(y)$。

解 使用分布函数法。由于 $Y = X^2 \ge 0$，当 $y < 0$ 时，$F_Y(y) = 0$。

当 $0 \le y < 1$ 时：

$$\begin{aligned} F_Y(y) &= P\{-\sqrt{y} \le X \le \sqrt{y}\} \\ &= P\{-\sqrt{y} \le X < 0\} + P\{0 \le X \le \sqrt{y}\} \\ &= \frac{1}{2}\sqrt{y} + \frac{1}{4}\sqrt{y} = \frac{3}{4}\sqrt{y}. \end{aligned}$$

当 $1 \le y < 4$ 时（注意 $-1 < x < 0$ 对应 $0 < y < 1$，而 $0 \le x < 2$ 对应 $0 \le y < 4$）：

$$F_Y(y) = P\{-1 < X < 0\} + P\{0 \le X \le \sqrt{y}\} = \frac{1}{2} + \frac{1}{4}\sqrt{y}.$$

当 $y \ge 4$ 时，$F_Y(y) = 1$。

求导得概率密度：

$$f_Y(y)=\begin{cases}\dfrac{3}{8\sqrt{y}}, & 0<y<1,\\ \dfrac{1}{8\sqrt{y}}, & 1\le y<4,\\ 0, & \text{其他}.\end{cases}$$

此题体现了当 X 的密度函数分段定义时，$Y=g(X)$ 的分布需要分段讨论，特别注意变换后不同区间的合并与拆分。

练习 2.56 离散型随机变量函数分布 已知 X 为随机变量，$Y=X^2+1$，已知 X 的概率分布为 $P\{X=-1\}=P\{X=0\}=P\{X=1\}=\dfrac{1}{3}$，则 Y 的分布函数 $F_Y(y)=$ _____。

解 Y 的可能取值为：

- 当 $X=-1$ 或 $X=1$ 时，$Y=(-1)^2+1=1^2+1=2$
- 当 $X=0$ 时，$Y=0^2+1=1$

因此

$$P\{Y=1\}=P\{X=0\}=\frac{1}{3},\quad P\{Y=2\}=P\{X=-1\}+P\{X=1\}=\frac{2}{3}.$$

分布函数：

$$F_Y(y)=\begin{cases}0, & y<1,\\ \dfrac{1}{3}, & 1\leqslant y<2,\\ 1, & y\geqslant 2.\end{cases}$$

练习 2.57 设随机变量 X 服从参数为 2 的指数分布，求 $Y=1-e^{-2X}$ 的概率密度 $f_Y(y)$。

解 由题意，$X\sim E(2)$，概率密度为 $f_X(x)=\begin{cases}2e^{-2x}, & x>0\\ 0, & x\le 0\end{cases}$。

函数 $y=1-e^{-2x}$ 在 $x>0$ 上严格单调递增，值域为 $(0,1)$。反函数为 $x=-\dfrac{1}{2}\ln(1-y)$，导数为

$$\frac{\mathrm{d}x}{\mathrm{d}y}=\frac{1}{2(1-y)}.$$

由公式法：

$$\begin{aligned}f_Y(y)&=f_X\left(-\frac{1}{2}\ln(1-y)\right)\cdot\left|\frac{1}{2(1-y)}\right|\\&=2e^{-2\cdot\left[-\frac{1}{2}\ln(1-y)\right]}\cdot\frac{1}{2(1-y)}\\&=2(1-y)\cdot\frac{1}{2(1-y)}=1,\quad 0<y<1.\end{aligned}$$

故

$$f_Y(y)=\begin{cases}1, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

即 $Y\sim U(0,1)$。

练习 2.58 设随机变量 $X\sim N(4,16)$，求 $Y=|X-4|$ 的概率密度 $f_Y(y)$。

解 令 $Z=X-4$，则 $Z\sim N(0,16)$，概率密度为 $f_Z(z)=\dfrac{1}{\sqrt{32\pi}}e^{-\frac{z^2}{32}}$。

函数 $y=|z|$ 在 $(-\infty,+\infty)$ 上不是严格单调函数，使用分布函数法。

由于 $Y=|Z|\geq 0$，当 $y\leq 0$ 时，$F_Y(y)=0$。

当 $y>0$ 时：

$$F_Y(y)=P\{|Z|\leq y\}=P\{-y\leq Z\leq y\}=F_Z(y)-F_Z(-y).$$

求导得：

$$f_Y(y)=F_Z'(y)-F_Z'(-y)\cdot(-1)=f_Z(y)+f_Z(-y)=\frac{2}{\sqrt{32\pi}}e^{-\frac{y^2}{32}}.$$

故

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{8\pi}}e^{-\frac{y^2}{32}}, & y>0,\\ 0, & y\leq 0.\end{cases}$$

练习 2.59 柯西分布变换 设 $f(x)=\dfrac{k}{1+x^2}$，求：(1) k (2)$Y=1-\sqrt[3]{X}$ 的概率密度。

解 (1) 由归一性条件：

$$\int_{-\infty}^{+\infty}\frac{k}{1+x^2}\,\mathrm{d}x=k\left[\arctan x\right]_{-\infty}^{+\infty}=k\pi=1\implies k=\frac{1}{\pi}.$$

即 X 服从标准柯西分布。

(2) 函数 $y=1-\sqrt[3]{x}$ 在 $(-\infty,+\infty)$ 上严格单调递减，反函数为 $x=(1-y)^3$，导数为

$$\frac{\mathrm{d}x}{dy}=-3(1-y)^2.$$

由随机变量函数的密度公式（公式法）：

$$f_Y(y)=f_X[(1-y)^3]\cdot\left|\frac{\mathrm{d}x}{dy}\right|=\frac{1}{\pi}\cdot\frac{1}{1+(1-y)^6}\cdot 3(1-y)^2=\frac{3(1-y)^2}{\pi\left[1+(1-y)^6\right]},\quad y\in\mathbb{R}.$$

答案：(1) $\boldsymbol{k=\dfrac{1}{\pi}}$；(2) $\boldsymbol{f_Y(y)=\dfrac{3(1-y)^2}{\pi[1+(1-y)^6]}}$。

练习 2.60 柯西分布尺度变换 设随机变量 X 的概率密度 $f(x)=\dfrac{1}{\pi(1+x^2)}$，则 $Y=2X$ 的分布密度为（　）。　（A）$\dfrac{1}{\pi(1+4y^2)}$；　（B）$\dfrac{2}{\pi(4+y^2)}$；　（C）$\dfrac{1}{\pi(1+y^2)}$；　（D）$\dfrac{1}{\pi}\arctan y$。

解 答案选 **(B)**。由单调函数密度变换公式：

$$f_Y(y)=f_X\left(\frac{y}{2}\right)\cdot\left|\frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{y}{2}\right)\right|=\frac{1}{\pi\left[1+\left(\frac{y}{2}\right)^2\right]}\cdot\frac{1}{2}=\frac{2}{\pi(4+y^2)}.$$

练习 2.61 正态分布平方变换 设随机变量 $X\sim N(0,1)$，求 $Y=2X^2+1$ 的概率密度函数。

解 当 $y\leqslant 1$ 时，$F_Y(y)=0$，故 $f_Y(y)=0$。

当 $y>1$ 时：

$$\begin{aligned}F_Y(y)&=P\{2X^2+1\leqslant y\}=P\left\{X^2\leqslant\frac{y-1}{2}\right\}=P\left\{-\sqrt{\frac{y-1}{2}}\leqslant X\leqslant\sqrt{\frac{y-1}{2}}\right\}\\&=2\Phi\left(\sqrt{\frac{y-1}{2}}\right)-1,\end{aligned}$$

其中 Φ 为标准正态分布函数。求导得：

$$f_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=2\varphi\left(\sqrt{\frac{y-1}{2}}\right)\cdot\frac{1}{2\sqrt{2(y-1)}},$$

其中 $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$ 为标准正态密度。化简得：

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{\pi(y-1)}}\exp\left\{-\dfrac{y-1}{4}\right\}, & y>1,\\ 0, & y\leqslant 1.\end{cases}$$

练习 2.62 逻辑分布变换 设 $f(x)=\dfrac{e^x}{(1+e^x)^2}$，$Y=e^X$，求 X 的分布函数和 Y 的概率密度。

解 (1) 分布函数：

$$F(x)=\int_{-\infty}^{x}\frac{e^t}{(1+e^t)^2}\,dt=\left[-\frac{1}{1+e^t}\right]_{-\infty}^{x}=\frac{e^x}{1+e^x},\quad x\in\mathbb{R}.$$

这是标准的逻辑分布（Logistic distribution）。(2) 函数 $y=e^x$ 在 $(-\infty,+\infty)$ 上严格单调递增，值域为 $(0,+\infty)$，反函数为 $x=\ln y$，导数为 $\dfrac{\mathrm{d}x}{\mathrm{d}y}=\dfrac{1}{y}$。

由公式法：

$$f_Y(y)=f_X(\ln y)\cdot\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right|=\frac{e^{\ln y}}{(1+e^{\ln y})^2}\cdot\frac{1}{y}=\frac{y}{(1+y)^2}\cdot\frac{1}{y}=\frac{1}{(1+y)^2},\quad y>0.$$

故

$$f_Y(y)=\begin{cases}\dfrac{1}{(1+y)^2}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

答案：(1) $\boldsymbol{F(x)=\dfrac{e^x}{1+e^x}}$；(2) 密度函数如上。

练习 2.63 均匀分布绝对值变换 设 $X\sim U(-1,2)$，求 $Y=|X|$ 的概率密度。

解 X 的概率密度为

$$f_X(x)=\begin{cases}\dfrac{1}{3}, & -1<x<2,\\ 0, & \text{其他}.\end{cases}$$

函数 $y=|x|$ 在区间 $(-1,2)$ 上不是单调函数，使用分布函数法。由于 $Y=|X|\in[0,2)$：

- 当 $0<y<1$ 时：

$$F_Y(y)=P\{|X|\leqslant y\}=P\{-y\leqslant X\leqslant y\}=\frac{y-(-y)}{3}=\frac{2y}{3},$$

  故 $f_Y(y)=\dfrac{2}{3}$；
- 当 $1\leqslant y<2$ 时：

$$F_Y(y)=P\{-1\leqslant X\leqslant y\}=\frac{y-(-1)}{3}=\frac{y+1}{3},$$

  故 $f_Y(y)=\dfrac{1}{3}$；
- 当 $y\leqslant 0$ 时，$F_Y(y)=0$；当 $y\geqslant 2$ 时，$F_Y(y)=1$。

综上，

$$f_Y(y)=\begin{cases}\dfrac{2}{3}, & 0<y<1,\\ \dfrac{1}{3}, & 1\leqslant y<2,\\ 0, & \text{其他}.\end{cases}$$

第 3 章 二维随机变量及分布

$$
\begin{cases}
\text{二维随机变量的概念：二维平面中随机点的坐标} \\
\text{二维随机变量的概率分布定义及性质}\begin{cases}\text{(1) 单调不减性}\\ \text{(2) 规范性}\\ \text{(3) 右连续性}\end{cases}\quad \text{及}\begin{cases}\text{边缘分布}\\ \text{条件分布}\end{cases} \\
\text{二维离散型}\begin{cases}
\text{联合分布律：} P\{X=x_i,Y=y_j\}=p_{ij};\ p_{ij}\geqslant 0,\ \sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty}p_{ij}=1 \\
\text{边缘分布律：}\begin{cases}P\{X=x_i\}=\sum\limits_{j=1}^{\infty}p_{ij}=p_{i\cdot}\\ P\{Y=y_j\}=\sum\limits_{i=1}^{\infty}p_{ij}=p_{\cdot j}\end{cases} \\
\text{条件分布律：}\begin{cases}P\{X=x_i|Y=y_j\}=\dfrac{p_{ij}}{p_{\cdot j}}\\ P\{Y=y_j|X=x_i\}=\dfrac{p_{ij}}{p_{i\cdot}}\end{cases}
\end{cases} \\
\text{二维连续型}\begin{cases}
\text{联合概率密度：} f(x,y)\geqslant 0,\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(u,v)dudv=1 \\
\text{边缘分布：}\begin{cases}f_X(x)=\int_{-\infty}^{+\infty}f(x,y)dy,\ F_X(x)=\int_{-\infty}^{x}f_X(x)dx\\ f_Y(y)=\int_{-\infty}^{+\infty}f(x,y)dx,\ F_Y(y)=\int_{-\infty}^{y}f_Y(y)dy\end{cases} \\
\text{条件概率密度：}\begin{cases}f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}\\ f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}\end{cases}
\end{cases} \\
\text{随机变量的独立性：}\begin{cases}\text{离散型：} p_{ij}=p_{i\cdot}p_{\cdot j}\\ \text{连续型：} f(x,y)=f_X(x)\cdot f_Y(y)\end{cases}\quad\Longleftrightarrow\ F(x,y)=F_X(x)\cdot F_Y(y) \\
\text{二维随机变量函数的概率分布}\begin{cases}\text{离散型：逐点法或表格法}\\ \text{连续型：卷积公式法，分布函数法}\end{cases}
\end{cases}
$$

内容提要

- 二维随机变量与联合分布函数 3.1 3.2 3.1.1
- 边缘分布函数 3.3
- 二维离散型随机变量：联合分布律 3.4 3.5、边缘分布律 3.6、条件分布律 3.7
- 二维连续型随机变量：联合概率密度 3.8 3.3、边缘概率密度 3.9、条件概率密度 3.10
- 常见二维分布：均匀分布 3.12、正态分布 3.13 3.4.2
- n 维随机变量 3.14 3.15
- 随机变量的独立性：定义 3.17 3.18、充要条件与性质 3.6.3
- 二维随机变量函数的分布：分布函数法、卷积公式 3.3、max 和 min 的分布、常见分布的可加性

3.1 二维随机变量

在许多实际问题中，随机试验的结果往往需要用两个或更多个随机变量来描述。例如，研究某地区的气象状况，需要同时考虑温度和湿度；检查产品的质量，需要同时考察长度和直径等。这时，我们不仅要研究每个随机变量的统计规律，还要研究它们之间的相互关系。

定义 3.1 (二维随机变量)

设 E 是一个随机试验，其样本空间为 $S=\{e\}$，设 $X=X(e)$ 和 $Y=Y(e)$ 是定义在 S 上的两个随机变量，由它们构成的向量 (X,Y) 称为二维随机向量（或二维随机变量）。 ♣

笔记 二维随机变量 (X,Y) 的性质不仅与 X 及 Y 各自的性质有关，还依赖于这两个随机变量之间的相互关系。因此，逐个研究 X 和 Y 的性质是不够的，还需要将 (X,Y) 作为一个整体来研究。

类似地，如果 $X_1,X_2,\ldots,X_n$ 是定义在同一个样本空间 Ω 上的 n 个随机变量，则称 $(X_1,X_2,\ldots,X_n)$ 为 n **维随机变量**或 n **维随机向量**，$X_i\ (i=1,2,\ldots,n)$ 称为第 i 个**分量**。

3.1.1 联合分布函数

定义 3.2 (联合分布函数)

设 (X,Y) 是二维随机变量，对于任意实数 x,y，称二元函数

$$F(x,y)=P\{X\leqslant x,Y\leqslant y\}$$

为二维随机变量 (X,Y) 的分布函数，或称为随机变量 X 和 Y 的联合分布函数，记为 $(X,Y)\sim F(x,y)$。 ♣

几何上，如果把二维随机变量 (X,Y) 看成平面上随机点的坐标，那么分布函数 $F(x_0,y_0)$ 就表示随机点 (X,Y) 落在以 (x_0,y_0) 为顶点且位于该点左下方的无穷矩形区域内的概率（见图 **??**）。

性质 联合分布函数 $F(x,y)$ 具有以下基本性质：

1. 单调性：$F(x,y)$ 关于每个变量单调不减。即对任意固定的 y，当 $x_1<x_2$ 时，$F(x_1,y)\leqslant F(x_2,y)$；对任意固定的 x，当 $y_1<y_2$ 时，$F(x,y_1)\leqslant F(x,y_2)$。
2. 有界性：$0\leqslant F(x,y)\leqslant 1$，且

$$F(-\infty,y)=0,\quad F(x,-\infty)=0,\quad F(-\infty,-\infty)=0,\quad F(+\infty,+\infty)=1.$$

3. 右连续性：$F(x,y)$ 关于每个变量右连续，即

$$F(x+0,y)=F(x,y),\quad F(x,y+0)=F(x,y).$$

4. 非负性：对于任意的 $x_1<x_2$，$y_1<y_2$，有

$$P\{x_1<X\leqslant x_2,\ y_1<Y\leqslant y_2\}=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\geqslant 0.$$

注 性质 (4) 是二维分布函数特有的性质，也是区别于一维情形的关键。仅满足前三个性质的函数不一定能成为某个二维随机变量的联合分布函数，必须同时满足性质 (4)。例如，设

$$F(x,y)=\begin{cases}1, & x+y\geqslant 0,\\ 0, & x+y<0,\end{cases}$$

满足性质 (1)–(3)，但不满足性质 (4)，因此不是联合分布函数。

定义 3.3 (边缘分布函数)

设二维随机变量 (X,Y) 的联合分布函数为 $F(x,y)$，分别称

$$F_X(x)=P\{X\leqslant x\}=F(x,+\infty)=\lim_{y\to+\infty}F(x,y)$$

$$F_Y(y)=P\{Y\leqslant y\}=F(+\infty,y)=\lim_{x\to+\infty}F(x,y)$$

为 (X,Y) 关于 X 和关于 Y 的边缘分布函数。 ♣

练习 3.1 设 (X,Y) 的联合分布函数为 $F(x,y)=a\left(b+\arctan\dfrac{x}{2}\right)\left(c+\arctan\dfrac{y}{2}\right)$，求常数 a,b,c 及 (X,Y) 的概率密度。

解 由联合分布函数的性质：

(1) $F(-\infty,y)=a\left(b-\dfrac{\pi}{2}\right)\left(c+\arctan\dfrac{y}{2}\right)=0$，由于此式对所有 y 成立，故 $b=\dfrac{\pi}{2}$。

(2) $F(x,-\infty)=a\left(b+\arctan\dfrac{x}{2}\right)\left(c-\dfrac{\pi}{2}\right)=0$，同理得 $c=\dfrac{\pi}{2}$。

(3) $F(+\infty,+\infty)=a\left(\dfrac{\pi}{2}+\dfrac{\pi}{2}\right)\left(\dfrac{\pi}{2}+\dfrac{\pi}{2}\right)=a\pi^2=1$，故 $a=\dfrac{1}{\pi^2}$。

概率密度为联合分布函数的二阶混合偏导数：

$$f(x,y)=\frac{\partial^2F(x,y)}{\partial x\,\partial y}=\frac{1}{\pi^2}\cdot\frac{1/2}{1+(x/2)^2}\cdot\frac{1/2}{1+(y/2)^2}=\frac{4}{\pi^2(4+x^2)(4+y^2)}.$$

练习 3.2 密度函数线性组合 设二维随机变量 (X_1,Y_1) 和 (X_2,Y_2) 的概率密度分别为 $f_1(x,y)$ 与 $f_2(x,y)$，令 $f(x,y)=af_1(x,y)+bf_2(x,y)$。若 $f(x,y)$ 是某二维连续型随机变量的概率密度，则 a,b 满足条件（　）。

(A) $a+b=1$　　(B) $a>0$ 且 $b>0$

(C) $0\leqslant a\leqslant 1,\ 0\leqslant b\leqslant 1$　　(D) $a\geqslant 0,\ b\geqslant 0$，且 $a+b=1$

解 $f(x,y)$ 为概率密度的充要条件：$f(x,y)\geqslant 0$ 且 $\iint f(x,y)\,\mathrm{d}x\mathrm{d}y=1$。由于 f_1,f_2 本身是概率密度（非负、归一），有

$$\iint f\,\mathrm{d}x\mathrm{d}y=a\cdot 1+b\cdot 1=a+b=1.$$

又 $f_1,f_2\geqslant 0$，故 $f\geqslant 0$ 要求 $a\geqslant 0$ 且 $b\geqslant 0$。答案选 **D**。

练习 3.3 标准正态独立随机变量 设随机变量 X,Y 相互独立，且均服从标准正态分布 $N(0,1)$，则（　）。

(A) $P\{X+Y\geqslant 1\}=1/2$　　(B) $P\{X-Y\geqslant 1\}=1/2$

(C) $P\{\max(X,Y)\geqslant 0\}=1/4$　　(D) $P\{\min(X,Y)\geqslant 0\}=1/4$

解 $X+Y\sim N(0,2)$，$X-Y\sim N(0,2)$，均值为 0 但对称点不在 1，故 $P(\cdot\geqslant 1)\neq 1/2$，排除 A、B。

$$P\{\max(X,Y)\geqslant 0\}=1-P\{X<0,Y<0\}=1-\frac{1}{2}\cdot\frac{1}{2}=\frac{3}{4}\neq\frac{1}{4}$$

排除 C。

$$P\{\min(X,Y)\geqslant 0\}=P\{X\geqslant 0,Y\geqslant 0\}=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

答案选 **D**。

练习 3.4max 的分布函数 设随机变量 X 和 Y 独立同分布，且 X 的分布函数为 $F(x)$，则 $Z=\max\{X,Y\}$ 的分布函数为（　）。

(A) $F^2(x)$　　(B) $F(x)F(y)$

(C) $1-[1-F(x)]^2$　　(D) $[1-F(x)][1-F(y)]$

解

$$F_Z(z)=P\{\max(X,Y)\leqslant z\}=P\{X\leqslant z,\ Y\leqslant z\}=F(z)\cdot F(z)=F^2(z).$$

答案选 **A**。（选项 C 为 min 的分布函数。）

练习 3.5 边缘分布函数 如果二维随机变量 (X,Y) 的分布函数为

$$F(x,y)=\begin{cases}(1-e^{-\lambda_1 x})(1-e^{-\lambda_2 y}+e^{-\lambda_3 x-\lambda_2 y}), & x>0,\ y>0,\\ 0, & \text{其他},\end{cases}$$

其中 $\lambda_1,\lambda_2,\lambda_3>0$，求 X 和 Y 的边缘分布函数。

解

$$F_X(x)=\lim_{y\to+\infty}F(x,y)=\begin{cases}(1-e^{-\lambda_1 x})\cdot 1=1-e^{-\lambda_1 x}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

$$F_Y(y)=\lim_{x\to+\infty}F(x,y)=\begin{cases}1\cdot(1-e^{-\lambda_2 y})=1-e^{-\lambda_2 y}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

笔记 $X\sim E(\lambda_1)$，$Y\sim E(\lambda_2)$。注意 $F(x,y)\neq F_X(x)F_Y(y)$，故 X 与 Y 不独立（除非 $\lambda_3=0$）。

练习 3.6 边缘密度不保证联合为正态 随机变量 (X,Y) 的概率密度为

$$f(x,y)=\frac{1+\sin x\sin y}{2\pi}e^{-\frac{x^2+y^2}{2}},\quad -\infty<x,y<+\infty,$$

求边缘概率密度 $f_X(x)$ 和 $f_Y(y)$。

解 $\sin y\cdot e^{-y^2/2}$ 是奇函数，在 $\mathbb{R}$ 上积分为 0，故

$$f_X(x)=\int_{-\infty}^{+\infty}\frac{1+\sin x\sin y}{2\pi}e^{-\frac{x^2+y^2}{2}}\,\mathrm{d}y=\frac{e^{-x^2/2}}{2\pi}\left[\int_{-\infty}^{+\infty}e^{-y^2/2}\,\mathrm{d}y+\sin x\cdot 0\right]=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

同理 $f_Y(y)=\dfrac{1}{\sqrt{2\pi}}e^{-y^2/2}$。

笔记 $X\sim N(0,1)$，$Y\sim N(0,1)$，但 (X,Y) 不服从二维正态分布（密度中含 $\sin x\sin y$ 项）。这说明：边缘分布为正态不能推出联合分布为二维正态。

3.2 二维离散型随机变量

定义 3.4 (二维离散型随机变量)

如果二维随机变量 (X,Y) 全部可能取到的值是有限对或可列无限对，则称 (X,Y) 为二维离散型随机变量。 ♣

3.2.1 联合分布律

定义 3.5 (联合分布律)

设二维离散型随机变量 (X,Y) 所有可能取的值为 (x_i,y_j) $(i,j=1,2,\dots)$，则称

$$p_{ij}=P\{X=x_i,Y=y_j\},\quad i,j=1,2,\dots$$

为 (X,Y) 的联合分布律，或称为随机变量 X 和 Y 的联合概率分布。 ♣

联合分布律也可用表格形式表示：

$X\backslash Y$	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}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	

性质 数列 $\{p_{ij}\}$ $(i,j=1,2,\dots)$ 是某一二维离散型随机变量的联合分布律的充分必要条件为：

1. 非负性：$p_{ij}\geqslant 0$；
2. 归一性：$\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty}p_{ij}=1$。

由联合分布律可以求得联合分布函数：

$$F(x,y)=\sum_{x_i\leqslant x}\sum_{y_j\leqslant y}p_{ij}.$$

更一般地，设 G 为平面上的某个区域，则

$$P\{(X,Y)\in G\}=\sum_{(x_i,y_j)\in G}p_{ij},$$

即 (X,Y) 落入区域 G 的概率等于所有落在 G 内的可能值 (x_i,y_j) 对应的概率之和。

3.2.2 边缘分布律

定义 3.6 (边缘分布律)

设 (X,Y) 的联合分布律为 $P\{X=x_i,Y=y_j\}=p_{ij}$，则分别称

$$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{\infty}p_{ij},\quad i=1,2,\ldots$$

$$p_{\cdot j}=P\{Y=y_j\}=\sum_{i=1}^{\infty}p_{ij},\quad j=1,2,\ldots$$

为 (X,Y) 关于 X 和关于 Y 的边缘分布律。 ♣

笔记 边缘分布律写在联合分布律表格的最后一列和最后一行。需要注意的是，联合分布律可以完全确定边缘分布律，但反过来由边缘分布律一般不能确定联合分布律，因为边缘分布律不包含两个随机变量之间相互关系的信息。

3.2.3 条件分布律

定义 3.7 (条件分布律)

设 (X,Y) 为二维离散型随机变量，其联合分布律为 $P\{X=x_i,Y=y_j\}=p_{ij}\ (i,j=1,2,\ldots)$。若对固定的 j，$p_{\cdot j}=P\{Y=y_j\}>0$，则称

$$P\{X=x_i\mid Y=y_j\}=\frac{P\{X=x_i,Y=y_j\}}{P\{Y=y_j\}}=\frac{p_{ij}}{p_{\cdot j}},\quad i=1,2,\ldots$$

为在 $\{Y=y_j\}$ 条件下 X 的条件分布律。

同理，若对固定的 i，$p_{i\cdot}=P\{X=x_i\}>0$，则称

$$P\{Y=y_j\mid X=x_i\}=\frac{p_{ij}}{p_{i\cdot}},\quad j=1,2,\ldots$$

为在 $\{X=x_i\}$ 条件下 Y 的条件分布律。 ♣

练习 3.7 将两封信投入 3 个信箱，设 X_1,X_2 分别表示第一个和第二个信箱投进的信的数量。求：

(1) (X_1,X_2) 的联合分布律及边缘分布律；

(2) 在条件 $X_2=1$ 下，X_1 的条件分布。

解 (1) 每封信有 3 种投法，两封信共有 $3^2=9$ 种等可能结果。X_1 与 X_2 的可能取值均为 $0,1,2$。

$$\begin{aligned}
p_{00}&=P\{X_1=0,X_2=0\}=\frac{1}{9}, & p_{01}&=P\{X_1=0,X_2=1\}=\frac{2}{9},\\
p_{02}&=P\{X_1=0,X_2=2\}=\frac{1}{9}, & p_{10}&=P\{X_1=1,X_2=0\}=\frac{2}{9},\\
p_{11}&=P\{X_1=1,X_2=1\}=\frac{2}{9}, & p_{12}&=P\{X_1=1,X_2=2\}=0,\\
p_{20}&=P\{X_1=2,X_2=0\}=\frac{1}{9}, & p_{21}&=p_{22}=0.
\end{aligned}$$

联合分布律与边缘分布律如下表所示：

$X_1 \backslash X_2$	0	1	2	$p_{i\cdot}$
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{4}{9}$
1	$\frac{2}{9}$	$\frac{2}{9}$	0	$\frac{4}{9}$
2	$\frac{1}{9}$	0	0	$\frac{1}{9}$
$p_{\cdot j}$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$	1

(2) $P\{X_2=1\}=p_{\cdot 1}=\frac{4}{9}$。在 $X_2=1$ 的条件下，X_1 的条件分布为：

$$P\{X_1=0\mid X_2=1\}=\frac{p_{01}}{p_{\cdot 1}}=\frac{2/9}{4/9}=\frac{1}{2}$$
$$P\{X_1=1\mid X_2=1\}=\frac{p_{11}}{p_{\cdot 1}}=\frac{2/9}{4/9}=\frac{1}{2}$$
$$P\{X_1=2\mid X_2=1\}=0$$

即 $X_1\mid(X_2=1)\sim\begin{pmatrix}0 & 1\\ 1/2 & 1/2\end{pmatrix}$。

练习 3.8 已知二维随机变量 (X,Y) 的联合概率分布如下表，若随机事件 $\{X=0\}$ 与 $\{X+Y=1\}$ 相互独立，则（　）。

$X\backslash Y$	0	1
0	0.4	a
1	b	0.1

A. $a=0.2,b=0.3$　B. $a=0.1,b=0.4$　C. $a=0.3,b=0.2$　D. $a=0.4,b=0.1$

解 答案选 **D**。由联合分布的性质，先由归一性得：

$$0.4+a+b+0.1=1\implies a+b=0.5.$$

分别计算三个事件的概率：

$$P\{X=0\}=0.4+a,\quad P\{X+Y=1\}=a+b=0.5,\quad P\{X=0,X+Y=1\}=P\{X=0,Y=1\}=a.$$

由 $\{X=0\}$ 与 $\{X+Y=1\}$ 相互独立，有 $P\{X=0\}\cdot P\{X+Y=1\}=P\{X=0,X+Y=1\}$，即

$$(0.4+a)\times 0.5=a\implies 0.2+0.5a=a\implies 0.5a=0.2\implies a=0.4.$$

从而 $b=0.5-a=0.1$。

练习 3.9 联合分布的确定与独立性 已知 X_1 和 X_2 的分布律分别为

X_1	-1	0	1
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

X_2	0	1
P	$\frac{1}{2}$	$\frac{1}{2}$

且 $P\{X_1X_2=0\}=1$。求：

(1) X_1 和 X_2 的联合分布律；

(2) X_1 和 X_2 是否独立？为什么？

(3) $Z=X_1-X_2$ 的分布律。

解 (1) 由 $P\{X_1X_2=0\}=1$，得 $P\{X_1X_2\neq 0\}=0$，即

$$P\{X_1=-1,X_2=1\}=0,\quad P\{X_1=1,X_2=1\}=0.$$

由边缘分布：

$$\begin{aligned}P\{X_2=1\}=\frac{1}{2}&=P\{X_1=-1,X_2=1\}+P\{X_1=0,X_2=1\}+P\{X_1=1,X_2=1\}\\&=0+P\{X_1=0,X_2=1\}+0\end{aligned}$$

故 $P\{X_1=0,X_2=1\}=1/2$。同理 $P\{X_2=0,X_1=-1\}=P\{X_1=-1\}-P\{X_1=-1,X_2=1\}=1/4$，$P\{X_2=0,X_1=1\}=1/4$，$P\{X_2=0,X_1=0\}=P\{X_1=0\}-P\{X_1=0,X_2=1\}=1/2-1/2=0$。联合分布律如下：

$X_2\backslash X_1$	-1	0	1	$p_{\cdot j}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
1	0	$\frac{1}{2}$	0	$\frac{1}{2}$
$p_{i\cdot}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

(2) 不独立。例如 $P\{X_2=0,X_1=-1\}=1/4$，但 $P\{X_2=0\}\cdot P\{X_1=-1\}=\frac{1}{2}\times\frac{1}{4}=\frac{1}{8}\neq\frac{1}{4}$。

(3) $Z=X_1-X_2$ 的可能取值为 $-2,-1,0,1$：

$$\begin{aligned}P\{Z=-2\}&=P\{X_1=-1,X_2=1\}=0,\\P\{Z=-1\}&=P\{X_1=-1,X_2=0\}+P\{X_1=0,X_2=1\}=\frac{1}{4}+\frac{1}{2}=\frac{3}{4},\\P\{Z=0\}&=P\{X_1=0,X_2=0\}+P\{X_1=1,X_2=1\}=0+0=0,\\P\{Z=1\}&=P\{X_1=1,X_2=0\}=\frac{1}{4}.\end{aligned}$$

$$Z\sim\begin{pmatrix}-2&-1&0&1\\0&3/4&0&1/4\end{pmatrix}.$$

练习 3.10 由独立性求未知参数 设二维随机变量 (X,Y) 的联合分布律为

(X,Y)	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$
P	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{18}$	$\frac{1}{3}$	α	β

若 X 与 Y 相互独立，求 α 和 β。

解 设 $a_1=P\{X=1\}$，$a_2=P\{X=2\}$，$b_1=P\{Y=1\}$，$b_2=P\{Y=2\}$，$b_3=P\{Y=3\}$。由独立性 $P\{X=i,Y=j\}=P\{X=i\}\cdot P\{Y=j\}$：

$$a_1b_1=\frac{1}{6},\quad a_2b_1=\frac{1}{3}\implies\frac{a_1}{a_2}=\frac{1}{2}.$$

同理：

$$a_1b_2=\frac{1}{9}\implies a_2b_2=\frac{2}{9}\cdot\frac{a_2}{a_1}=\alpha,\quad \alpha=2\times\frac{1}{9}=\frac{2}{9}.$$

$$a_1b_3=\frac{1}{18}\implies a_2b_3=2\times\frac{1}{18}=\frac{1}{9},\quad \beta=\frac{1}{9}.$$

验证归一性：

$$\frac{1}{6}+\frac{1}{9}+\frac{1}{18}+\frac{1}{3}+\frac{2}{9}+\frac{1}{9}=\frac{3+2+1+6+4+2}{18}=\frac{18}{18}=1$$

笔记 本题的解题技巧是利用独立性的比例关系 $\frac{a_1}{a_2}=\frac{a_1b_j}{a_2b_j}$（对任意 j），由已知项快速求出未知项。

练习 3.11 行列式分布 设随机变量 X_1,X_2,X_3,X_4 相互独立且同分布，$P\{X_i=0\}=0.6$，$P\{X_i=1\}=0.4$（$i=1,2,3,4$），求行列式 $Z=\begin{vmatrix}X_1 & X_2\\ X_3 & X_4\end{vmatrix}$ 的概率分布。

解 $Z=X_1X_4-X_2X_3$，可能取值为 $\{-1,0,1\}$。

$$\begin{aligned}P\{Z=1\}&=P\{X_1X_4=1\}\cdot P\{X_2X_3=0\}\\&=P\{X_1=1,X_4=1\}\cdot P\{(X_2=0)\cup(X_3=0)\}\\&=0.4^2\cdot(1-0.4^2)=0.16\times0.84=0.1344.\\P\{Z=-1\}&=P\{X_1X_4=0\}\cdot P\{X_2X_3=1\}\\&=(1-0.4^2)\cdot0.4^2=0.84\times0.16=0.1344\\P\{Z=0\}&=1-0.1344-0.1344=0.7312.\end{aligned}$$

所以

$$Z\sim\begin{pmatrix}-1 & 0 & 1\\ 0.1344 & 0.7312 & 0.1344\end{pmatrix}.$$

练习 3.12 独立性填表 设 X 与 Y 相互独立，下表是 (X,Y) 的联合分布律以及边缘分布律中的部分数值，试将其余数值填入空白处：

$X\backslash Y$	y_1	y_2	y_3	$P\{X=x_i\}$
x_1		1/8		
x_2	1/8			
$P\{Y=y_j\}$	1/6			1

解 由独立性 $p_{ij}=p_{i\cdot}\cdot p_{\cdot j}$：$p_{21}=p_{2\cdot}\cdot p_{\cdot1}=1/8$，故 $p_{2\cdot}=(1/8)/(1/6)=3/4$，$p_{1\cdot}=1-3/4=1/4$。$p_{12}=p_{1\cdot}\cdot p_{\cdot2}=1/8$，故 $p_{\cdot2}=(1/8)/(1/4)=1/2$。$p_{\cdot3}=1-1/6-1/2=1/3$。填满后：

$X\backslash Y$	y_1	y_2	y_3	$P\{X=x_i\}$
x_1	1/24	1/8	1/12	1/4
x_2	1/8	3/8	1/4	3/4
$P\{Y=y_j\}$	1/6	1/2	1/3	1

验证：行和 $1/24+1/8+1/12=1/4$，$1/8+3/8+1/4=3/4$；列和 $1/24+1/8=1/6$，$1/8+3/8=1/2$，$1/12+1/4=1/3$。✓

练习 3.13 无放回取球 袋中有编号为 $1,1,2,3$ 的四个球，从中无放回地取两次，每次任取一个。设 X_1,X_2 分别为第一次、第二次取得的球的号码。求：

(1) (X_1,X_2) 的联合分布律，并判断 X_1 与 X_2 的独立性；

(2) 在 $X_2=2$ 的条件下，X_1 的条件分布律；

(3) $Y = X_1X_2$ 的分布。

解 (1) 四个球编号 $\{1,1,2,3\}$，无放回取两次，共 $4\times 3 = 12$ 种等可能结果。

$X_1\backslash X_2$	1	2	3	$p_{i\cdot}$
1	1/6	1/6	1/6	1/2
2	1/6	0	1/12	1/4
3	1/6	1/12	0	1/4
$p_{\cdot j}$	1/2	1/4	1/4	1

独立性：$p_{11} = 1/6$，$p_{1\cdot}\cdot p_{\cdot 1} = (1/2)(1/2) = 1/4 \neq 1/6$，故 X_1 与 X_2 不独立。

(2) $P\{X_2 = 2\} = p_{\cdot 2} = 1/4$。

$$P\{X_1 = 1 \mid X_2 = 2\} = \frac{p_{12}}{p_{\cdot 2}} = \frac{1/6}{1/4} = \frac{2}{3}, \quad P\{X_1 = 3 \mid X_2 = 2\} = \frac{p_{32}}{p_{\cdot 2}} = \frac{1/12}{1/4} = \frac{1}{3}.$$

(3) $Y = X_1X_2$ 的可能取值及概率：

$$\begin{aligned}
P\{Y=1\} &= P\{X_1=1, X_2=1\} = 1/6,\\
P\{Y=2\} &= P\{X_1=1, X_2=2\} + P\{X_1=2, X_2=1\} = 1/6+1/6 = 1/3,\\
P\{Y=3\} &= P\{X_1=1, X_2=3\} + P\{X_1=3, X_2=1\} = 1/6+1/6 = 1/3,\\
P\{Y=6\} &= P\{X_1=2, X_2=3\} + P\{X_1=3, X_2=2\} = 1/12+1/12 = 1/6.
\end{aligned}$$

$$Y \sim \begin{pmatrix} 1 & 2 & 3 & 6 \\ 1/6 & 1/3 & 1/3 & 1/6 \end{pmatrix}.$$

✍ **练习 3.14 离散型函数的分布** 已知 (X,Y) 的概率分布为

$X\backslash Y$	-1	1
-1	1/3	1/6
0	1/3	1/6

求：(1) $Z = X - Y$ 的概率分布；(2) 设 $U = \max\{X,Y\}$，$V = \min\{X,Y\}$，求 (U,V) 及 UV 的分布。

解 (1) Z 的可能取值：$Z(-1,-1) = 0$，$Z(-1,1) = -2$，$Z(0,-1) = 1$，$Z(0,1) = -1$。

$$P\{Z=-2\} = 1/6, \quad P\{Z=-1\} = 1/6, \quad P\{Z=0\} = 1/3, \quad P\{Z=1\} = 1/3.$$

$$Z \sim \begin{pmatrix} -2 & -1 & 0 & 1 \\ 1/6 & 1/6 & 1/3 & 1/3 \end{pmatrix}.$$

(2) (U,V) 的可能取值：

$$\begin{aligned}
(X,Y) = (-1,-1) &: U=-1,\ V=-1, \quad P = 1/3.\\
(X,Y) = (-1,1) &: U=1,\ V=-1, \quad P = 1/6.\\
(X,Y) = (0,-1) &: U=0,\ V=-1, \quad P = 1/3.\\
(X,Y) = (0,1) &: U=1,\ V=0, \quad P = 1/6.
\end{aligned}$$

(U,V) 联合分布律：

$V\backslash U$	-1	0	1
-1	$1/3$	$1/3$	$1/6$
0	0	0	$1/6$

UV 的分布：$UV \in \{-1,0,1\}$。

$$UV \sim \begin{pmatrix} -1 & 0 & 1 \\ 1/6 & 1/2 & 1/3 \end{pmatrix}.$$

练习 3.15 事件的示性函数 设 A,B 为两个随机事件，$P(A)=1/4$，$P(B)=1/6$，$P(AB)=1/12$。令

$$X=\begin{cases}1, & A\text{发生}, \\ 0, & A\text{不发生}.\end{cases} \qquad Y=\begin{cases}1, & B\text{发生}, \\ 0, & B\text{不发生}.\end{cases}$$

求：(1) (X,Y) 的联合分布律；(2) $Z=X^2+Y^2$ 的分布律；(3) $W=\min(X,Y)$ 的分布律。

解 (1) 由事件的概率关系：

$$\begin{aligned}
P(X=0,Y=0)&=P(\bar{A}\bar{B})=1-P(A\cup B)=1-(P(A)+P(B)-P(AB))=\frac{2}{3},\\
P(X=1,Y=0)&=P(A\bar{B})=P(A)-P(AB)=\frac{1}{4}-\frac{1}{12}=\frac{1}{6},\\
P(X=0,Y=1)&=P(\bar{A}B)=P(B)-P(AB)=\frac{1}{6}-\frac{1}{12}=\frac{1}{12},\\
P(X=1,Y=1)&=P(AB)=\frac{1}{12}.
\end{aligned}$$

$X\backslash Y$	0	1	$P\{X=x_i\}$
0	$2/3$	$1/12$	$3/4$
1	$1/6$	$1/12$	$1/4$
$P\{Y=y_j\}$	$5/6$	$1/6$	1

(2) $Z=X^2+Y^2$：$(0,0)\to 0$，$(1,0)$ 或 $(0,1)\to 1$，$(1,1)\to 2$。

$$P(Z=0)=2/3,\quad P(Z=1)=1/6+1/12=1/4,\quad P(Z=2)=1/12.$$

$$Z\sim\begin{pmatrix} 0 & 1 & 2 \\ 2/3 & 1/4 & 1/12 \end{pmatrix}.$$

(3) $W=\min(X,Y)$：$W=0$ 当至少一个为 0，$W=1$ 仅当 $(1,1)$。

$$P(W=0)=2/3+1/6+1/12=11/12,\quad P(W=1)=1/12.$$

练习 3.16 独立性条件求参数 设 X,Y 相互独立，$P(X=1)=P(Y=1)=p$，$P(X=0)=P(Y=0)=1-p$（$0<p<1$）。令

$$Z=\begin{cases}1, & X+Y\text{为偶数}, \\ 0, & X+Y\text{为奇数}.\end{cases}$$

要使 X 与 Z 独立，则 $p=$（　）

A. $1/2$　B. $1/4$
C. $1/3$　D. $2/3$

解 $X+Y$ 为偶数当且仅当 $(X,Y)=(0,0)$ 或 $(1,1)$，故

$$P(Z=1)=(1-p)^2+p^2,\quad P(Z=0)=2p(1-p).$$

又 $P(X=1,Z=1)=P(X=1,Y=1)=p^2$。由独立性 $P(X=1,Z=1)=P(X=1)\,P(Z=1)$：

$$p^2=p\left[(1-p)^2+p^2\right].$$

若 $p=0$ 或 $p=1$，Z 退化为常数，平凡独立。对 $0<p<1$，两边除以 p：

$$p=1-2p+2p^2 \implies 2p^2-3p+1=0 \implies (2p-1)(p-1)=0 \implies p=1/2.$$

答案选 **A**。

笔记 验证 $p=1/2$：$P(Z=1)=1/2$，$P(X=1,Z=1)=1/4=P(X=1)P(Z=1)$；$P(X=1,Z=0)=1/4=P(X=1)P(Z=0)$ 。

练习 3.17 离散型函数分布：和、max、min X,Y 独立同分布，$P(X=0)=1/2$，$P(X=1)=1/4$，$P(X=2)=1/4$。求：(1) $Z=X+Y$；(2) $U=\max(X,Y)$；(3) $W=\min(X,Y)$ 的分布律。

解 (1) $Z=X+Y$ 的可能取值为 $0,1,2,3,4$。由独立性逐项计算：

$$\begin{aligned}
P(Z=0)&=P(X=0)P(Y=0)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4},\\
P(Z=1)&=P(X=0)P(Y=1)+P(X=1)P(Y=0)=2\cdot\frac{1}{2}\cdot\frac{1}{4}=\frac{1}{4},\\
P(Z=2)&=P(0,2)+P(2,0)+P(1,1)=2\cdot\frac{1}{2}\cdot\frac{1}{4}+\frac{1}{4}\cdot\frac{1}{4}=\frac{5}{16},\\
P(Z=3)&=P(1,2)+P(2,1)=2\cdot\frac{1}{4}\cdot\frac{1}{4}=\frac{1}{8},\\
P(Z=4)&=P(2,2)=\frac{1}{16}.
\end{aligned}$$

(2) $U=\max(X,Y)$：利用 $P(U\leqslant k)=P(X\leqslant k)^2$：

$$P(U=0)=\frac{1}{4},\quad P(U=1)=\left(\frac{3}{4}\right)^2-\left(\frac{1}{2}\right)^2=\frac{5}{16},\quad P(U=2)=1-\frac{1}{4}-\frac{5}{16}=\frac{7}{16}.$$

(3) $W=\min(X,Y)$：利用 $P(W\geqslant k)=P(X\geqslant k)^2$：

$$P(W=0)=1-\left(\frac{1}{2}\right)^2=\frac{3}{4},\quad P(W=1)=\left(\frac{1}{2}\right)^2-\left(\frac{1}{4}\right)^2=\frac{3}{16},\quad P(W=2)=\frac{1}{16}.$$

Z	0	1	2	3	4
P	1/4	1/4	5/16	1/8	1/16

U	0	1	2
P	1/4	5/16	7/16

W	0	1	2
P	3/4	3/16	1/16

练习 3.18 泊松变量的复合函数 设 X_1,X_2,X_3 相互独立且同服从 $P(1)$，记

$$X=\begin{cases}1, & X_1+X_2=1,\\ 0, & X_1+X_2\neq 1,\end{cases}\qquad Y=\begin{cases}1, & X_2+X_3=1,\\ 0, & X_2+X_3\neq 1.\end{cases}$$

求：(1) (X,Y) 的联合分布律；(2) 边缘分布律；(3) $P(X+Y\leqslant 1)$。

解 (1) $\{X=1,Y=1\}$ 要求 $X_1+X_2=1$ 且 $X_2+X_3=1$。分两种情况：

情形 $X_2=0$：需 $X_1=1$，$X_3=1$。$P=e^{-1}\cdot e^{-1}\cdot e^{-1}=e^{-3}$。

情形 $X_2=1$：需 $X_1=0$，$X_3=0$。$P=e^{-1}\cdot e^{-1}\cdot e^{-1}=e^{-3}$。

当 $X_2\geqslant 2$ 时 $X_1+X_2\geqslant 2\neq 1$，不贡献。

$$P(X=1,Y=1)=2e^{-3}.$$

边缘概率：

$$P(X=1)=P(X_1+X_2=1)=P(1,0)+P(0,1)=2e^{-2},\qquad P(Y=1)=P(X_2+X_3=1)=2e^{-2}.$$

$$P(X=0,Y=1)=P(Y=1)-P(X=1,Y=1)=2e^{-2}-2e^{-3},$$
$$P(X=1,Y=0)=P(X=1)-P(X=1,Y=1)=2e^{-2}-2e^{-3},$$
$$P(X=0,Y=0)=1-2e^{-3}-(2e^{-2}-2e^{-3})-(2e^{-2}-2e^{-3})=1-4e^{-2}+2e^{-3}.$$

$X\backslash Y$	1	0	$P\{X=x_i\}$
1	$2e^{-3}$	$2e^{-2}-2e^{-3}$	$2e^{-2}$
0	$2e^{-2}-2e^{-3}$	$1-4e^{-2}+2e^{-3}$	$1-2e^{-2}$
$P\{Y=y_j\}$	$2e^{-2}$	$1-2e^{-2}$	1

(2) 边缘分布律见上表右侧和下方。

(3) $P(X+Y\leqslant 1)=1-P(X=1,Y=1)=1-2e^{-3}$。

笔记 X 与 Y 不独立：$P(X=1,Y=1)=2e^{-3}$，而 $P(X=1)P(Y=1)=4e^{-4}\neq 2e^{-3}$。这是因为 X 和 Y 共同依赖 X_2。

练习 3.19 离散条件均匀分布 从 $\{1,2,3\}$ 中等可能地取一个数记为 X，再从 $\{1,\ldots,X\}$ 中等可能地取一个整数记为 Y。求 (X,Y) 的联合分布律和边缘分布律，并计算 $P(X+Y=4)$。

解 $P(X=a,Y=b)=P(X=a)\,P(Y=b|X=a)=\dfrac{1}{3}\cdot\dfrac{1}{a}$（$1\leqslant b\leqslant a$）。逐项列出联合分布律：

$X\backslash Y$	1	2	3	$P\{X=x_i\}$
1	1/3	0	0	1/3
2	1/6	1/6	0	1/3
3	1/9	1/9	1/9	1/3
$P\{Y=y_j\}$	$1/3+1/6+1/9$ $=11/18$	$1/6+1/9$ $=5/18$	$1/9$ $=1/9$	1

$P(X+Y=4)$：满足 $a+b=4$ 且 $1\leqslant b\leqslant a\leqslant 3$ 的对为 $(2,2)$ 和 $(3,1)$。

$$P(X+Y=4)=P(X=2,Y=2)+P(X=3,Y=1)=\frac{1}{6}+\frac{1}{9}=\frac{5}{18}.$$

3.3 二维连续型随机变量

定义 3.8 (二维连续型随机变量)

设 (X,Y) 为二维随机变量，若存在非负可积函数 $f(x,y)$，使得对于任意实数 x,y，都有

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)\,\mathrm{d}u\,\mathrm{d}v,$$

则称 (X,Y) 为二维连续型随机变量，称 $f(x,y)$ 为 (X,Y) 的联合概率密度函数（简称概率密度），记为 $(X,Y)\sim f(x,y)$。♣

性质 联合概率密度 $f(x,y)$ 具有以下性质：

1. 非负性：$f(x,y) \geqslant 0$；
2. 归一性：$\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$；
3. 设 G 为 xOy 平面上的某个区域，则
$$P\{(X,Y)\in G\} = \iint_G f(x,y)\,\mathrm{d}x\,\mathrm{d}y;$$
4. 若 $f(x,y)$ 在点 (x,y) 处连续，则
$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x\,\partial y}.$$

笔记 几何意义：$z = f(x,y)$ 表示空间中的一个曲面，该曲面与 xOy 平面之间所围成的空间区域的体积为 1。$P\{(X,Y)\in G\}$ 的值等于以 G 为底、以曲面 $z = f(x,y)$ 为顶面的曲顶柱体的体积。

改变 $f(x,y)$ 在有限个点（甚至一条曲线上）的值（仍取非负值），$f(x,y)$ 仍然是 (X,Y) 的概率密度。

3.3.1 边缘概率密度

定义 3.9 (边缘概率密度)

设 $(X,Y)\sim f(x,y)$，则分别称
$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y,\quad f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x$$
为 (X,Y) 关于 X 和关于 Y 的边缘概率密度。 ♣

相应的边缘分布函数分别为
$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = \int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty} f(t,y)\,\mathrm{d}y\right]\mathrm{d}t,$$
$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)\,\mathrm{d}t = \int_{-\infty}^{y}\left[\int_{-\infty}^{+\infty} f(x,t)\,\mathrm{d}x\right]\mathrm{d}t.$$

3.3.2 条件概率密度

定义 3.10 (条件概率密度)

设 $(X,Y)\sim f(x,y)$，若边缘概率密度 $f_X(x)>0$，则称
$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$
为在 $\{X=x\}$ 条件下 Y 的条件概率密度。

同理，若 $f_Y(y)>0$，则称
$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$
为在 $\{Y=y\}$ 条件下 X 的条件概率密度。 ♣

性质 [概率密度乘法公式] 若 $f_X(x)>0$，$f_Y(y)>0$，则有
$$f(x,y) = f_X(x)\,f_{Y|X}(y|x) = f_Y(y)\,f_{X|Y}(x|y).$$

定义 3.11 (条件分布函数)

在 $\{X=x\}$ 条件下 Y 的条件分布函数为

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

同理，在 $\{Y=y\}$ 条件下 X 的条件分布函数为

$$F_{X|Y}(x|y)=\int_{-\infty}^{x} f_{X|Y}(u|y)\,\mathrm{d}u=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}\,\mathrm{d}u.$$

♣

笔记 在研究二维连续型随机变量时，通常需要先确定联合密度函数 $f(x,y)$，一般有三种情形：

1. 题目中直接给出 $f(x,y)$；
2. 通过边缘密度函数及随机变量的独立性来确定；
3. 通过条件密度及边缘密度利用乘法公式求出。

练习 3.20 设 (X,Y) 是二维随机变量，X 的边缘概率密度为

$$f_X(x)=\begin{cases}3x^2, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

在给定 $X=x$（$0<x<1$）的条件下，Y 的条件概率密度为

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

(1) 求 (X,Y) 的联合概率密度 $f(x,y)$；

(2) 求 Y 的边缘概率密度 $f_Y(y)$；

(3) 求 $P\{X>2Y\}$。

解 (1) 由概率密度乘法公式：

$$f(x,y)=f_{Y|X}(y|x)\,f_X(x)=\begin{cases}\dfrac{3y^2}{x^3}\cdot 3x^2=\dfrac{9y^2}{x}, & 0<y<x<1,\\ 0, & \text{其他}.\end{cases}$$

(2) 当 $0<y\leqslant 1$ 时，$f_Y(y)=\displaystyle\int_y^1\frac{9y^2}{x}\,\mathrm{d}x=9y^2\big[-\ln y\big]_y^1=-9y^2\ln y$；其余范围 $f_Y(y)=0$。

(3) $\{X>2Y\}$ 等价于 $\{Y<X/2\}$，积分区域为 $\{(x,y):0<x<1,\ 0<y<x/2\}$：

$$P\{X>2Y\}=\int_0^1\mathrm{d}x\int_0^{x/2}\frac{9y^2}{x}\,\mathrm{d}y=\int_0^1\frac{3x^2}{8}\,\mathrm{d}x=\left[\frac{x^3}{8}\right]_0^1=\frac{1}{8}.$$

练习 3.21 求系数、边缘密度与条件密度 二维随机变量 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}A(6-x-y), & 0\leqslant x\leqslant 2,\ 2\leqslant y\leqslant 4,\\ 0, & \text{其他}.\end{cases}$$

(1) 求系数 A；(2) 求 X,Y 的边缘概率密度；(3) X,Y 是否独立？(4) 求条件概率密度 $f_{X|Y}(x|y)$ 和 $f_{Y|X}(y|x)$。

解 (1) 由归一性：

$$\int_2^4\mathrm{d}y\int_0^2 A(6-x-y)\,\mathrm{d}x=A\int_2^4\big[6x-\tfrac{x^2}{2}-xy\big]_0^2\,\mathrm{d}y=A\int_2^4(10-2y)\,\mathrm{d}y=A\cdot 8=1,\quad A=\frac{1}{8}$$

(2) 边缘概率密度：

$$f_X(x)=\int_2^4\frac{1}{8}(6-x-y)\,\mathrm{d}y=\frac{3-x}{4},\quad 0<x<2.$$

$$f_Y(y)=\int_0^2\frac{1}{8}(6-x-y)\,\mathrm{d}x=\frac{5-y}{4},\quad 2<y<4.$$

(3) 取 $x=1,y=3$：$f(1,3)=\frac{1}{8}(6-1-3)=\frac{1}{4}$，而 $f_X(1)\cdot f_Y(3)=\frac{2}{4}\cdot\frac{2}{4}=\frac{1}{4}$。看似相等，但一般点检查不充分。实际上 $f(x,y)=\dfrac{1}{8}(6-x-y)$，$f_X(x)\,f_Y(y)=\dfrac{(3-x)(5-y)}{16}=\dfrac{15-8x-3y+xy}{16}$，而 $\dfrac{6-x-y}{8}=\dfrac{12-2x-2y}{16}$。两者不相等（例如取 $x=0,y=2$：前者 $=12/16=3/4$，后者 $=15/16$），故 X 与 Y 不独立。

(4)

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{(6-x-y)/8}{(5-y)/4}=\frac{6-x-y}{2(5-y)},\quad 0<x<2.$$

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{(6-x-y)/8}{(3-x)/4}=\frac{6-x-y}{2(3-x)},\quad 2<y<4.$$

性质 [区域概率公式] 对于联合分布函数 $F(x,y)$ 及边缘分布函数 $F_X(x),F_Y(y)$，有

$$P\{X>x_0,Y>y_0\}=1-F_X(x_0)-F_Y(y_0)+F(x_0,y_0).$$

推导：将全空间分为四个区域，

$$P\{X>x_0,Y>y_0\}=1-P\{X\leqslant x_0\}-P\{Y\leqslant y_0\}+P\{X\leqslant x_0,Y\leqslant y_0\}.$$

注意：只有当 X 与 Y 独立时，$P\{X>x_0,Y>y_0\}=[1-F_X(x_0)][1-F_Y(y_0)]$。

练习 3.22 二维分布的性质判断 下列叙述中正确的是（　）

(1) 二维正态分布的边缘分布是正态分布，条件分布也是正态分布；

(2) 二维正态分布的边缘分布是正态分布，但条件分布不是正态分布；

(3) 二维均匀分布的边缘分布是均匀分布，条件分布是均匀分布；

(4) 二维均匀分布的边缘分布不一定是均匀分布，条件分布也不一定为均匀分布。

A. (1)(3)　B. (2)(3)　C. (1)(4)　D. (2)(4)

解 答案选 **C**。(1) 正确：二维正态分布的边缘分布和条件分布均为正态分布。

(4) 正确：二维均匀分布的边缘不一定均匀。例如

$$f(x,y)=\begin{cases}2, & 0<x<1,\ 0<y<1-x,\\ 0, & \text{其他},\end{cases}$$

则 $f_X(x)=2(1-x)$（$0<x<1$），$f_Y(y)=2(1-y)$（$0<y<1$），均非均匀分布；条件密度 $f_{Y|X}(y|x)=1/(1-x)$（$0<y<1-x$），在 $[0,1]$ 上也不均匀。

练习 3.23 抛物线区域上的均匀分布 设 (X,Y) 服从区域 $G=\{(x,y):0\leqslant y\leqslant 1-x^2\}$ 上的均匀分布。求：

(1) 联合概率密度 $f(x,y)$；

(2) 边缘密度 $f_X(x)$，$f_Y(y)$ 及条件密度 $f_{Y|X}(y|x)$；

(3) X 与 Y 是否独立？

(4) $P\{Y\geqslant X^2\}$。

解 (1) 区域 G 的面积：

$$S_G=\int_{-1}^{1}(1-x^2)\,\mathrm{d}x=\left[x-\frac{x^3}{3}\right]_{-1}^{1}=\frac{4}{3}.$$

故

$$f(x,y)=\begin{cases}3/4, & -1<x<1,\ 0\leqslant y\leqslant 1-x^2,\\ 0, & \text{其他}.\end{cases}$$

(2) 边缘密度：

$$f_X(x)=\int_0^{1-x^2}\frac{3}{4}\,\mathrm{d}y=\frac{3}{4}(1-x^2),\quad |x|<1.$$

$$f_Y(y)=\int_{-\sqrt{1-y}}^{\sqrt{1-y}}\frac{3}{4}\,\mathrm{d}x=\frac{3}{2}\sqrt{1-y},\quad 0\leqslant y\leqslant 1.$$

条件密度（$|x|<1$ 时）：

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\begin{cases}\dfrac{1}{1-x^2}, & 0\leqslant y\leqslant 1-x^2,\\ 0, & \text{其他}.\end{cases}$$

(3) $f_X(x)\,f_Y(y)=\dfrac{3}{4}(1-x^2)\cdot\dfrac{3}{2}\sqrt{1-y}=\dfrac{9}{8}(1-x^2)\sqrt{1-y}\neq\dfrac{3}{4}=f(x,y)$，故不独立。

(4) $\{Y\geqslant X^2\}$ 要求 $x^2\leqslant y\leqslant 1-x^2$，即 $2x^2\leqslant 1$，$|x|\leqslant 1/\sqrt{2}$：

$$P\{Y\geqslant X^2\}=\int_{-1/\sqrt{2}}^{1/\sqrt{2}}\int_{x^2}^{1-x^2}\frac{3}{4}\,\mathrm{d}y\,\mathrm{d}x=\frac{3}{2}\int_0^{1/\sqrt{2}}(1-2x^2)\,\mathrm{d}x$$

$$=\frac{3}{2}\left[x-\frac{2x^3}{3}\right]_0^{1/\sqrt{2}}=\frac{3}{2}\left(\frac{1}{\sqrt{2}}-\frac{2}{3\sqrt{8}}\right)=\frac{3}{2}\cdot\frac{2}{3\sqrt{2}}=\frac{\sqrt{2}}{2}.$$

练习 3.24Gamma 型联合密度 (X,Y) 的概率密度为

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

求：(1) 边缘密度；(2) 独立性；(3) $Z=X+Y$ 的密度。

解 (1) 利用 $\displaystyle\int_0^{+\infty}ye^{-y}\,\mathrm{d}y=1$ 和 $\displaystyle\int_0^{+\infty}y^2e^{-y}\,\mathrm{d}y=2$：

$$f_X(x)=\frac{e^{-x}}{2}\int_0^{+\infty}(x+y)e^{-y}\,\mathrm{d}y=\frac{e^{-x}}{2}\left(x\int_0^{+\infty}e^{-y}\,\mathrm{d}y+\int_0^{+\infty}ye^{-y}\,\mathrm{d}y\right)=\frac{x+1}{2}e^{-x},\quad x>0.$$

(2) $f_X(x)\,f_Y(y)=\dfrac{(x+1)(y+1)}{4}e^{-(x+y)}\neq\dfrac{x+y}{2}e^{-(x+y)}=f(x,y)$，不独立。

(3) 用卷积公式，$z>0$ 时：

$$f_Z(z)=\int_0^z\frac{1}{2}(x+(z-x))e^{-(x+(z-x))}\,\mathrm{d}x=\frac{z}{2}e^{-z}\int_0^z\mathrm{d}x=\frac{z^2}{2}e^{-z}.$$

笔记 验证：$\displaystyle\int_0^{+\infty}\frac{z^2}{2}e^{-z}\,\mathrm{d}z=\frac{1}{2}\Gamma(3)=1$。即 $Z\sim\Gamma(3,1)$。

练习 3.25 联合均匀分布的判断 X,Y 独立且均服从 $U[0,1]$，则服从均匀分布的随机变量是（　）

A. (X,Y)　B. $X+Y$　C. $X-Y$　D. X^2

解 (X,Y) 的联合密度 $f(x,y)=1$（$0\leqslant x,y\leqslant 1$），为矩形上的均匀分布。

$X+Y$ 的密度为三角形分布 $f_Z(z)=z$（$0<z<1$）或 $2-z$（$1\leqslant z<2$），非均匀。

$X-Y$ 的密度为 $f_W(w)=1-|w|$（$|w|<1$），非均匀。

X^2 的密度为 $f(u)=1/(2\sqrt{u})$（$0<u<1$），即 Beta$(1/2,1)$ 分布，非均匀。

答案选 **A**。

练习 3.26 两种边缘密度计算 求下列联合密度的边缘概率密度：

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

(2) $f_2(x,y)=\begin{cases} 5(x^2+y), & 0<y<1-x^2, \\ 0, & \text{其他}. \end{cases}$

解 (1)

$$f_X(x)=\int_x^{+\infty} e^{-y}\,\mathrm{d}y=e^{-x}\ (x>0),\qquad f_Y(y)=\int_0^y e^{-y}\,\mathrm{d}x=ye^{-y}\ (y>0).$$

即 $X\sim E(1)$，$Y\sim\Gamma(2,1)$。

(2) $f_2(x,y)\neq 0$ 要求 $0<y<1-x^2$，即 $|x|<1$。

$$f_X(x)=\int_0^{1-x^2} 5(x^2+y)\,\mathrm{d}y=5\left[x^2y+\frac{y^2}{2}\right]_0^{1-x^2}=5\left[x^2(1-x^2)+\frac{(1-x^2)^2}{2}\right]=\frac{5}{2}(1-x^4),\quad |x|<1.$$

$$f_Y(y)=\int_{-\sqrt{1-y}}^{\sqrt{1-y}} 5(x^2+y)\,\mathrm{d}x=10\sqrt{1-y}\cdot\frac{1+2y}{3}=\frac{10}{3}(1+2y)\sqrt{1-y},\quad 0<y<1.$$

练习 3.27 抛物线区域上的均匀分布 $f(x,y)=\begin{cases} kx^2, & 0<x<1,\ x^2<y<1, \\ 0, & \text{其他}. \end{cases}$ 求：(1) k；(2) $P\{X>1/2\}$，$P\{Y<1/2\}$。

解 (1) 由归一性：

$$\int_0^1\int_{x^2}^1 kx^2\,\mathrm{d}y\,\mathrm{d}x=k\int_0^1 x^2(1-x^2)\,\mathrm{d}x=k\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_0^1=k\cdot\frac{2}{15}=1,$$

故 $\boldsymbol{k=15/2}$。

(2)

$$P\left\{X>\frac{1}{2}\right\}=\frac{15}{2}\int_{1/2}^1\int_{x^2}^1 x^2\,\mathrm{d}y\,\mathrm{d}x=\frac{15}{2}\int_{1/2}^1 x^2(1-x^2)\,\mathrm{d}x$$

$$=\frac{15}{2}\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_{1/2}^1=\frac{15}{2}\cdot\frac{47}{480}=\frac{47}{64}.$$

$$P\left\{Y<\frac{1}{2}\right\}=\frac{15}{2}\int_0^{1/\sqrt{2}}\int_{x^2}^{1/2} x^2\,\mathrm{d}y\,\mathrm{d}x=\frac{15}{2}\int_0^{1/\sqrt{2}} x^2\left(\frac{1}{2}-x^2\right)\mathrm{d}x$$

$$=\frac{15}{2}\left[\frac{x^3}{6}-\frac{x^5}{5}\right]_0^{1/\sqrt{2}}=\frac{15}{2}\cdot\frac{1}{15\sqrt{2}}=\frac{1}{2\sqrt{2}}=\frac{\sqrt{2}}{4}.$$

练习 3.28 概率计算 $f(x,y)=\begin{cases} Ae^{-y}, & 0<x<y, \\ 0, & \text{其他}. \end{cases}$

求 A，$P\{X+Y\geqslant 1\}$，$P\{X/Y\leqslant 1/2\}$。

解 (1) 归一性：$\displaystyle\int_0^{+\infty}\int_0^y Ae^{-y}\,\mathrm{d}x\,\mathrm{d}y=A\int_0^{+\infty} ye^{-y}\,\mathrm{d}y=A=1$，故 $A=1$。

(2) 利用 $Z=X+Y$ 的分布函数（$f(x,y)=e^{-y}$ 时 $F_Z(z)=1-2e^{-z/2}+e^{-z}$，$z>0$）：

$$P\{X+Y\geqslant 1\}=1-F_Z(1)=1-(1-2e^{-1/2}+e^{-1})=2e^{-1/2}-e^{-1}.$$

(3) $P\{X/Y\leqslant 1/2\}=P\{X\leqslant Y/2\}$。在 $\{0<x<y\}$ 中附加 $x\leqslant y/2$，即 $0<x\leqslant y/2$：

$$P\{X\leqslant Y/2\}=\int_0^{+\infty}\int_0^{y/2}e^{-y}\,\mathrm{d}x\,\mathrm{d}y=\int_0^{+\infty}\frac{y}{2}e^{-y}\,\mathrm{d}y=\frac{1}{2}.$$

练习 3.29 混合型：连续 + 离散 $X\sim E(4)$，Y 为离散型随机变量：$P\{Y=-1\}=1/4$，$P\{Y=1\}=3/4$，且 X 与 Y 相互独立。求 $P\{X-Y\leqslant 1\}$ 和 $P\{XY\leqslant 2\}$。

解 利用全概率公式（对 Y 的取值分解）。

$$\begin{aligned}P\{X-Y\leqslant 1\}&=P\{Y=-1\}P\{X+1\leqslant 1\}+P\{Y=1\}P\{X-1\leqslant 1\}\\&=\frac{1}{4}P\{X\leqslant 0\}+\frac{3}{4}P\{X\leqslant 2\}=\frac{1}{4}\cdot 0+\frac{3}{4}(1-e^{-8})=\frac{3}{4}(1-e^{-8}).\\P\{XY\leqslant 2\}&=P\{Y=-1\}P\{-X\leqslant 2\}+P\{Y=1\}P\{X\leqslant 2\}\\&=\frac{1}{4}\cdot 1+\frac{3}{4}(1-e^{-8})=1-\frac{3}{4}e^{-8}.\end{aligned}$$

（由于 $X>0$ 恒成立，$P\{-X\leqslant 2\}=1$。）

练习 3.30 三角形均匀分布求概率与卷积

$$f(x,y)=\begin{cases}2-x-y, & 0<x<1,\ 0<y<1,\\0, & \text{其他}.\end{cases}$$

求：(1) $P\{X>2Y\}$；(2) $Z=X+Y$ 的密度。

解 验证归一性：$\int_0^1\int_0^1(2-x-y)\,\mathrm{d}y\,\mathrm{d}x=1$。

(1) $\{X>2Y\}$ 等价于 $\{0<y<x/2\}$（$0<x<1$）：

$$\begin{aligned}P\{X>2Y\}&=\int_0^1\int_0^{x/2}(2-x-y)\,\mathrm{d}y\,\mathrm{d}x=\int_0^1\left[2y-xy-\frac{y^2}{2}\right]_0^{x/2}\mathrm{d}x\\&=\int_0^1\left(x-\frac{x^2}{2}-\frac{x^2}{8}\right)\mathrm{d}x=\int_0^1\left(x-\frac{5x^2}{8}\right)\mathrm{d}x=\left[\frac{x^2}{2}-\frac{5x^3}{24}\right]_0^1=\frac{7}{24}.\end{aligned}$$

(2) 当 $0<z<1$ 时（$x\in(0,z)$，$y\in(0,z-x)$）：

$$f_Z(z)=\int_0^z(2-z)\,\mathrm{d}x=z(2-z).$$

当 $1\leqslant z<2$ 时（$x\in(z-1,1)$，$y\in(0,z-x)$）：

$$f_Z(z)=\int_{z-1}^1(2-z)\,\mathrm{d}x=(2-z)(2-z)=(2-z)^2.$$

综上：

$$f_Z(z)=\begin{cases}z(2-z), & 0<z<1,\\(2-z)^2, & 1\leqslant z<2,\\0, & \text{其他}.\end{cases}$$

练习 3.31 三角形区域上的均匀分布 已知二维随机变量 (X,Y) 在区域 G 上服从均匀分布，G 由直线 $x-y=0$、$x+y=2$ 与 $y=0$ 围成。求：

(1) 边缘概率密度 $f_X(x)$ 和 $f_Y(y)$；

(2) 条件概率密度 $f_{X|Y}(x|y)$。

解 (X,Y) 的正概率密度区域 G 是以 $(0,0)$、$(2,0)$、$(1,1)$ 为顶点的三角形，面积 $S_G=\frac{1}{2}\times 2\times 1=1$。由于面积恰好为 1，联合概率密度为

$$f(x,y)=\begin{cases}1, & 0\leqslant y\leqslant 1,\ y\leqslant x\leqslant 2-y,\\ 0, & \text{其他}.\end{cases}$$

(1) 边缘概率密度。$f_X(x)$：固定 x，对 y 积分。

- 当 $0<x<1$ 时，y 从 0 到 x：$f_X(x)=\int_0^x 1\,\mathrm{d}y=x$；
- 当 $1<x<2$ 时，y 从 0 到 $2-x$：$f_X(x)=\int_0^{2-x} 1\,\mathrm{d}y=2-x$。

$$f_X(x)=\begin{cases}x, & 0<x<1,\\ 2-x, & 1<x<2,\\ 0, & \text{其他}.\end{cases}$$

$f_Y(y)$：固定 y，对 x 积分。

- 当 $0<y<1$ 时，x 从 y 到 $2-y$：$f_Y(y)=\int_y^{2-y} 1\,\mathrm{d}x=2-2y$。

$$f_Y(y)=\begin{cases}2-2y, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

(2) 在 $0<y<1$ 的条件下：

$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=\frac{1}{2-2y},\quad y\leqslant x\leqslant 2-y.$$

即

$$f_{X|Y}(x|y)=\begin{cases}\dfrac{1}{2-2y}, & y\leqslant x\leqslant 2-y,\\ 0, & \text{其他}.\end{cases}$$

笔记 求边缘概率密度的技巧（口诀）：求谁不积谁，不积先定限，限内画条线，先交写下限，后交写上限。即对 $f(x,y)$ 求关于 y 的积分得 $f_X(x)$，对 $f(x,y)$ 求关于 x 的积分得 $f_Y(y)$。定积分限时，画出正概率密度区域的图形，用平行于积分轴的直线穿过区域，先交到的位置为下限，后交到的位置为上限。

练习 3.32 条件均匀分布 已知随机变量 X 在区间 $[0,1]$ 上服从均匀分布，在 $X=x$（$0<x<1$）的条件下，随机变量 Y 在区间 $[0,x]$ 上服从均匀分布。求：

(1) (X,Y) 的联合概率密度；

(2) Y 的概率密度；

(3) 概率 $P\{X+Y>1\}$。

解 (1) 由题设，X 的边缘概率密度为

$$f_X(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他}.\end{cases}$$

在 $X=x$（$0<x<1$）的条件下，Y 在 $[0,x]$ 上的条件概率密度为

$$f_{Y|X}(y|x)=\begin{cases}\dfrac{1}{x}, & 0<y<x,\\ 0, & \text{其他}.\end{cases}$$

由乘法公式 $f(x,y)=f_X(x)\,f_{Y|X}(y|x)$，当 $0<y<x<1$ 时 $f(x,y)=1\cdot\dfrac{1}{x}=\dfrac{1}{x}$。故联合概率密度为

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

(2) 当 $0<y<1$ 时，固定 y 对 x 积分（x 从 y 到 1）：

$$f_Y(y)=\int_y^1\frac{1}{x}\,\mathrm{d}x=\big[-\ln y\big]_y^1=-\ln y.$$

当 $y\leqslant 0$ 或 $y\geqslant 1$ 时，$f_Y(y)=0$。因此

$$f_Y(y)=\begin{cases}-\ln y, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

(3) $\{X+Y>1\}$ 的积分区域为 $\{(x,y):0<y<x<1,\ x+y>1\}$，即 $x>\dfrac{1}{2}$ 且 $1-x<y<x$：

$$\begin{aligned}P\{X+Y>1\}&=\int_{1/2}^1\mathrm{d}x\int_{1-x}^x\frac{1}{x}\,\mathrm{d}y=\int_{1/2}^1\frac{x-(1-x)}{x}\,\mathrm{d}x=\int_{1/2}^1\frac{2x-1}{x}\,\mathrm{d}x=\int_{1/2}^1\left(2-\frac{1}{x}\right)\mathrm{d}x\\&=\big[2x-\ln x\big]_{1/2}^1=(2-0)-(1-\ln 2)=1-\ln 2.\end{aligned}$$

笔记 第 (1) 问中，乘法公式 $f(x,y)=f_X(x)\,f_{Y|X}(y|x)$ 仅在 $f_X(x)>0$ 时成立，因此最终结果需要加上 $f_X(x)=0$ 的部分，在分段表达式的“其他”中自然已包含。

3.4 常见的二维分布

3.4.1 二维均匀分布

定义 3.12 (二维均匀分布)

设 D 为 xOy 平面上的有界区域，其面积为 S_D。如果 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{1}{S_D}, & (x,y)\in D,\\ 0, & \text{其他},\end{cases}$$

则称 (X,Y) 在区域 D 上服从二维均匀分布，记为 $(X,Y)\sim U(D)$。♣

二维均匀分布是区间上均匀分布在二维情形的推广。若 $(X,Y)\sim U(D)$，则对 D 的任何子区域 G（其面积为 S_G），有

$$P\{(X,Y)\in G\}=\frac{S_G}{S_D}.$$

3.4.2 二维正态分布

定义 3.13 (二维正态分布)

如果 (X,Y) 的联合概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\},$$

其中 $\mu_1,\mu_2\in\mathbb{R}$，$\sigma_1>0$，$\sigma_2>0$，$-1<\rho<1$，则称 (X,Y) 服从参数为 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$ 的二维正态分布，记为 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2;\rho)$。 ♣

性质 二维正态分布具有以下重要性质：

1. 边缘分布仍为正态分布：若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2;\rho)$，则
$$X\sim N(\mu_1,\sigma_1^2),\quad Y\sim N(\mu_2,\sigma_2^2).$$
2. 独立性的充要条件：若 (X,Y) 服从二维正态分布，则 X 与 Y 相互独立的充分必要条件是 $\rho=0$。
3. 独立正态变量联合分布：若 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，且 X 与 Y 相互独立，则
$$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2;0).$$

笔记 需要注意以下几点：

1. 边缘分布为正态分布不能推出联合分布为二维正态分布。例如，设 (X,Y) 的联合密度为
$$f(x,y)=\frac{1}{2\pi}e^{-(x^2+y^2)/2}\left[1+\sin x\sin y\right],$$
则 X 和 Y 都服从标准正态分布，但 (X,Y) 并不服从二维正态分布。
2. 性质 (2) 仅对二维正态分布成立。一般地，由 X 与 Y 的边缘分布和 $\rho_{XY}=0$ 不能推出 X 与 Y 独立。
3. 二维正态分布的条件分布仍为正态分布：在 $X=x$ 的条件下，Y 的条件分布为
$$Y\mid X=x\sim N\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\ \sigma_2^2(1-\rho^2)\right).$$

3.5 n 维随机变量

前面主要讨论了二维随机变量，本节将相关概念推广到 n 维情形。在实际问题中，有时需要同时研究多个随机变量之间的关系。

定义 3.14 (n 维随机变量)

如果 $X_1,X_2,\ldots,X_n$ 是定义在同一个样本空间 Ω 上的 n 个随机变量，则称 $(X_1,X_2,\ldots,X_n)$ 为 n 维随机变量（或 n 维随机向量），X_i（$i=1,2,\ldots,n$）称为第 i 个分量。 ♣

定义 3.15 (n 维联合分布函数)

对任意的 n 个实数 $x_1,x_2,\ldots,x_n$，称 n 元函数

$$F(x_1,x_2,\ldots,x_n)=P\{X_1\leqslant x_1,X_2\leqslant x_2,\ldots,X_n\leqslant x_n\}$$

为 n 维随机变量 $(X_1,X_2,\ldots,X_n)$ 的联合分布函数（或分布函数）。 ♣

性质 [n 维联合分布函数的性质] n 维联合分布函数 $F(x_1,x_2,\ldots,x_n)$ 具有以下性质：

1. 单调性：F 关于每个变量 x_i 单调不减；
2. 右连续性：F 关于每个变量 x_i 右连续；
3. 有界性：

$$F(\underbrace{-\infty,\ldots,-\infty}_{k\text{个}},x_{k+1},\ldots,x_n)=0\quad(k=1,2,\ldots,n-1),\quad F(+\infty,+\infty,\ldots,+\infty)=1;$$

4. 非负性：对任意的 $a_i<b_i$（$i=1,2,\ldots,n$），以 F 计算的 n 维矩形区域的概率

$$P\{a_1<X_1\leqslant b_1,\ldots,a_n<X_n\leqslant b_n\}\geqslant 0.$$

定义 3.16 (n 维边缘分布函数)

n 维随机变量 $(X_1,X_2,\ldots,X_n)$ 中任意 k 个（$1\leqslant k<n$）分量所构成的 k 维随机变量的分布函数，称为原 n 维随机变量的 k 维边缘分布函数。特别地，关于第 i 个分量 X_i 的边缘分布函数为

$$F_{X_i}(x_i)=F(+\infty,\ldots,+\infty,x_i,+\infty,\ldots,+\infty).$$

♣

类似地，对于 n 维离散型随机变量，可以定义**联合分布律**

$$p_{i_1i_2\cdots i_n}=P\{X_1=x_{i_1},X_2=x_{i_2},\ldots,X_n=x_{i_n}\}$$

及相应的边缘分布律和条件分布律。

对于 n 维连续型随机变量，可以定义**联合概率密度** $f(x_1,x_2,\ldots,x_n)$，满足

$$F(x_1,x_2,\ldots,x_n)=\int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_1}f(t_1,t_2,\ldots,t_n)\,\mathrm{d}t_1\cdots\mathrm{d}t_n,$$

以及边缘概率密度、条件概率密度和概率密度乘法公式等，这些概念与二维情形完全类似。

3.6 随机变量的独立性

在第一章中我们讨论了事件的独立性，现在将这一概念推广到随机变量的情形。独立性是概率论中最重要的概念之一，它能够大大简化联合分布的研究。

3.6.1 独立性的定义

定义 3.17 (两个随机变量的独立性)

设 $F(x,y)$ 及 $F_X(x)$、$F_Y(y)$ 分别是二维随机变量 (X,Y) 的联合分布函数及边缘分布函数。若对任意实数 x,y，都有

$$P\{X\leqslant x,Y\leqslant y\}=P\{X\leqslant x\}\cdot P\{Y\leqslant y\},$$

即

$$F(x,y)=F_X(x)\cdot F_Y(y),$$

则称随机变量 X 和 Y 是相互独立的，否则称 X 与 Y 不相互独立。

♣

将独立性推广到 n 个随机变量的情形：

定义 3.18 (n 个随机变量的相互独立性)

设 n 维随机变量 $(X_1, X_2, \ldots, X_n)$ 的联合分布函数为 $F(x_1, x_2, \ldots, x_n)$，X_i 的边缘分布函数为 $F_{X_i}(x_i)$。若对任意实数 $x_1, x_2, \ldots, x_n$，都有

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

则称 $X_1, X_2, \ldots, X_n$ 是相互独立的。

进一步地，两个多维随机变量 $(X_1, \ldots, X_n)$ 与 $(Y_1, \ldots, Y_m)$ 相互独立，是指对任意实数 x_i（$i = 1, \ldots, n$）与 y_j（$j = 1, \ldots, m$），都有

$$F(x_1, \ldots, x_n, y_1, \ldots, y_m) = F_X(x_1, \ldots, x_n) \cdot F_Y(y_1, \ldots, y_m),$$

其中 F_X 和 F_Y 分别是 $(X_1, \ldots, X_n)$ 和 $(Y_1, \ldots, Y_m)$ 的分布函数。 ♣

定义 3.19 (随机变量序列的独立性)

若对任意正整数 k 和任意 $1 \leqslant i_1 < i_2 < \cdots < i_k$，随机变量 $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ 都相互独立，则称随机变量序列 $X_1, X_2, \ldots$ 是相互独立的。 ♣

3.6.2 独立的充要条件

n 个随机变量 $X_1, X_2, \ldots, X_n$ 相互独立 $\Longleftrightarrow$ 对任意的 n 个实数 x_i（$i = 1, 2, \ldots, n$），n 个事件 $\{X_1 \leqslant x_1\}, \{X_2 \leqslant x_2\}, \ldots, \{X_n \leqslant x_n\}$ 相互独立。

3.6.2.1 离散型随机变量的独立性

定理 3.1 (离散型随机变量独立的判定)

设 (X, Y) 为二维离散型随机变量，其联合分布律为 $P\{X = x_i, Y = y_j\} = p_{ij}$，则 X 与 Y 相互独立的充分必要条件是对一切 i, j，都有

$$p_{ij} = p_{i\cdot} \cdot p_{\cdot j}.$$

更一般地，n 个离散型随机变量 $X_1, X_2, \ldots, X_n$ 相互独立的充分必要条件是对任意的 $x_i \in D_i$（X_i 的一切可能取值），有

$$P\{X_1 = x_1, \ldots, X_n = x_n\} = \prod_{i=1}^{n} P\{X_i = x_i\}.$$

♡

证明 必要性显然。充分性：若 $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$ 对一切 i, j 成立，则

$$F(x, y) = \sum_{x_i \leqslant x} \sum_{y_j \leqslant y} p_{ij} = \sum_{x_i \leqslant x} \sum_{y_j \leqslant y} p_{i\cdot} p_{\cdot j} = \left(\sum_{x_i \leqslant x} p_{i\cdot}\right) \left(\sum_{y_j \leqslant y} p_{\cdot j}\right) = F_X(x) F_Y(y).$$

3.6.2.2 连续型随机变量的独立性

定理 3.2 (连续型随机变量独立的判定)

设 (X,Y) 为二维连续型随机变量，联合概率密度为 $f(x,y)$，边缘概率密度为 $f_X(x)$ 和 $f_Y(y)$，则 X 与 Y 相互独立的充分必要条件是

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

几乎处处成立（即除了面积为零的集合外处处成立）。

更一般地，n 维连续型随机变量 $(X_1,\ldots,X_n)$ 相互独立的充分必要条件是

$$f(x_1,x_2,\ldots,x_n)=\prod_{i=1}^{n}f_i(x_i),$$

其中 $f_i(x_i)$ 为 X_i 的边缘概率密度。

证明 充分性：若 $f(x,y)=f_X(x)\,f_Y(y)$，则

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}f(u,v)\,\mathrm{d}u\,\mathrm{d}v=\int_{-\infty}^{x}f_X(u)\,\mathrm{d}u\cdot\int_{-\infty}^{y}f_Y(v)\,\mathrm{d}v=F_X(x)\,F_Y(y).$$

必要性：若 $F(x,y)=F_X(x)\,F_Y(y)$，求二阶混合偏导数即得 $f(x,y)=f_X(x)\,f_Y(y)$。

笔记 利用独立性，可以大大简化联合分布的确定和概率计算。例如，当 X 与 Y 独立时，

$$P\{a<X\leqslant b,\ c<Y\leqslant d\}=P\{a<X\leqslant b\}\cdot P\{c<Y\leqslant d\}.$$

此外，常数与任何随机变量相互独立。

3.6.3 独立性的性质

性质 随机变量的独立性具有以下重要性质：

1. 若 $X_1,X_2,\ldots,X_n$ 相互独立，则其中任意 k 个（$2\leqslant k\leqslant n$）随机变量也相互独立。
2. 设 (X,Y) 为二维离散型随机变量，X 与 Y 独立，则条件分布等于边缘分布：
$$P\{X=x_i\mid Y=y_j\}=P\{X=x_i\}\ (P\{Y=y_j\}>0)$$
$$P\{Y=y_j\mid X=x_i\}=P\{Y=y_j\}\ (P\{X=x_i\}>0)$$
3. 设 (X,Y) 为二维连续型随机变量，X 与 Y 独立，则条件概率密度等于边缘概率密度：
$$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}=f_X(x)\ (f_Y(y)>0),\quad f_{Y|X}(y|x)=f_Y(y)\ (f_X(x)>0).$$
4. 若 $X_1,X_2,\ldots,X_n$ 相互独立，$g_1,g_2,\ldots,g_n$ 为一元连续函数，则 $g_1(X_1),g_2(X_2),\ldots,g_n(X_n)$ 也相互独立。更一般地，若 $X_{11},\ldots,X_{1i_1},X_{21},\ldots,X_{2i_2},\ldots,X_{n1},\ldots,X_{ni_n}$ 相互独立，g_k 是 i_k（$k=1,2,\ldots,n$）元连续函数，则 $g_1(X_{11},\ldots,X_{1i_1}),g_2(X_{21},\ldots,X_{2i_2}),\ldots,g_n(X_{n1},\ldots,X_{ni_n})$ 也相互独立。
5. 若 (X,Y) 服从二维正态分布，则 X 与 Y 相互独立当且仅当 $\rho=0$。但一般而言，X 与 Y 不相关（$\mathrm{Cov}(X,Y)=0$）不能推出 X 与 Y 独立。
6. 独立一定不相关，但不相关不一定独立（除非联合分布为正态分布）。

3.6.4 不独立的判断与证明

性质 X 与 Y 不独立 $\Longleftrightarrow$ 存在 x_0, y_0，使得

$$F(x_0, y_0) \neq F_X(x_0) \cdot F_Y(y_0).$$

也可通过取特定值使 $P\{X \leqslant x_0, Y \leqslant y_0\} \neq P\{X \leqslant x_0\} \cdot P\{Y \leqslant y_0\}$ 来证明不独立。对于离散型，寻找一对 (i,j) 使 $p_{ij} \neq p_{i\cdot} \cdot p_{\cdot j}$；对于连续型，验证 $f(x,y) \neq f_X(x)\, f_Y(y)$。

练习 3.33 设随机变量 X 与 Y 相互独立，且均服从参数为 1 的指数分布，即 $X, Y \sim E(1)$，求 $P\{X < Y\}$。

解 由独立性，联合概率密度为

$$f(x,y) = f_X(x)\, f_Y(y) = \begin{cases} e^{-(x+y)}, & x > 0,\ y > 0, \\ 0, & \text{其他}. \end{cases}$$

$$P\{X < Y\} = \iint_{x<y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{+\infty} \mathrm{d}y \int_0^y e^{-(x+y)}\,\mathrm{d}x = \int_0^{+\infty} e^{-y}(1-e^{-y})\,\mathrm{d}y$$

$$= \int_0^{+\infty} (e^{-y} - e^{-2y})\,\mathrm{d}y = \left[-e^{-y} + \frac{1}{2}e^{-2y}\right]_0^{+\infty} = 1 - \frac{1}{2} = \frac{1}{2}.$$

练习 3.34 条件正态联合密度 设随机变量 $X \sim N(0,1)$，在 $X = x$ 的条件下，随机变量 $Y \sim N(x,1)$，求 (X,Y) 的概率密度 $f(x,y)$。

解 由题意知

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}} e^{-(y-x)^2/2},$$

故由乘法公式

$$f(x,y) = f_X(x)\, f_{Y|X}(y|x) = \frac{1}{2\pi} e^{-\frac{x^2+(y-x)^2}{2}} = \frac{1}{2\pi} e^{-\frac{2x^2-2xy+y^2}{2}}.$$

练习 3.35 证明 X 与 $|X|$ 不独立 设随机变量 X 的概率密度为

$$f(x) = \frac{1}{2} e^{-|x|}, \quad x \in (-\infty, +\infty),$$

证明：X 与 $|X|$ 不独立。

证明 取 $x_0 = 1$，则

$$P\{X \leqslant 1\} = \int_{-\infty}^1 \frac{1}{2} e^{-|x|}\,\mathrm{d}x = \frac{1}{2}\int_{-\infty}^0 e^x\,\mathrm{d}x + \frac{1}{2}\int_0^1 e^{-x}\,\mathrm{d}x = \frac{1}{2} + \frac{1}{2}(1-e^{-1}) = 1 - \frac{1}{2}e^{-1},$$

$$P\{|X| \leqslant 1\} = \frac{1}{2}\int_{-1}^1 e^{-|x|}\,\mathrm{d}x = \int_0^1 e^{-x}\,\mathrm{d}x = 1 - e^{-1}.$$

又事件 $\{|X| \leqslant 1\} \subset \{X \leqslant 1\}$，即 $\{|X| \leqslant 1\} \cap \{X \leqslant 1\} = \{|X| \leqslant 1\}$，故

$$P\{X \leqslant 1, |X| \leqslant 1\} = P\{|X| \leqslant 1\} = 1 - e^{-1}.$$

而

$$P\{X \leqslant 1\} \cdot P\{|X| \leqslant 1\} = \left(1 - \frac{1}{2}e^{-1}\right)(1 - e^{-1}) = 1 - \frac{3}{2}e^{-1} + \frac{1}{2}e^{-2} \neq 1 - e^{-1},$$

因此 $P\{X \leqslant 1, |X| \leqslant 1\} \neq P\{X \leqslant 1\} \cdot P\{|X| \leqslant 1\}$，$X$ 与 $|X|$ 不相互独立。

笔记 本题中 x_0 可取任意正数均可证明不独立。关键在于 $|X| \leqslant x_0$ 隐含了 X 的信息，因此 X 与 $|X|$

存在函数关系，不可能独立。

练习 3.36 判断独立性与条件概率 设随机变量 X, Y 的联合概率密度函数为

$$f(x,y)=\begin{cases} e^{-y}, & x>0,\ y>x, \\ 0, & \text{其他}, \end{cases}$$

(1) X, Y 是否相互独立？(2) 求 $P\{X>2 \mid Y<4\}$。

解 (1) 求边缘概率密度：

$$f_X(x)=\int_x^{+\infty} e^{-y}\,\mathrm{d}y=e^{-x}\ (x>0), \qquad f_Y(y)=\int_0^y e^{-y}\,\mathrm{d}x=ye^{-y}\ (y>0).$$

由于 $f_X(x)\,f_Y(y)=xe^{-x-y}\neq e^{-y}=f(x,y)$（当 $y>x>0$ 时），故 X 与 Y 不相互独立。

(2)

$$P\{X>2,\ Y<4\}=\int_2^4 \mathrm{d}y\int_2^y e^{-y}\,\mathrm{d}x=\int_2^4 (y-2)e^{-y}\,\mathrm{d}y=e^{-2}-3e^{-4},$$

$$P\{Y<4\}=\int_0^4 ye^{-y}\,\mathrm{d}y=1-5e^{-4},$$

故

$$P\{X>2 \mid Y<4\}=\frac{P\{X>2,\ Y<4\}}{P\{Y<4\}}=\frac{e^{-2}-3e^{-4}}{1-5e^{-4}}.$$

3.7 二维随机变量函数的分布

设 X, Y 为随机变量，$g(x,y)$ 是二元函数，则 $Z=g(X,Y)$ 也是随机变量，称为随机变量 X, Y 的函数。本节讨论已知 (X,Y) 的分布，求 $Z=g(X,Y)$ 的分布的方法。

常考问题包括：已知 (X,Y) 的分布，求 $Z=X+Y$、$Z=XY$、$Z=X/Y$、$Z=\max\{X,Y\}$、$Z=\min\{X,Y\}$ 等的分布。

3.7.1 离散型随机变量函数的分布

对于离散型随机变量 (X,Y)，函数 $Z=g(X,Y)$ 的分布可以通过**逐点法**（也称**表格法**）来确定。

性质 [逐点法] 设 (X,Y) 的联合分布律为 $P\{X=x_i, Y=y_j\}=p_{ij}$，则 $Z=g(X,Y)$ 的分布律为

$$P\{Z=z_k\}=\sum_{g(x_i,y_j)=z_k} p_{ij},$$

即将使得 $g(x_i,y_j)=z_k$ 的所有 (i,j) 对应的 p_{ij} 求和。若不同的 (x_i,y_j) 对应相同的 z_k 值，则需要合并概率。

练习 3.37 设随机变量 X 和 Y 相互独立，且 $X\sim B(2,p)$，$Y\sim B(4,p)$，求 $Z=X+Y$ 的分布律。

解 由于 X 和 Y 相互独立，利用二项分布的可加性（独立二项分布之和仍为二项分布）：

$$Z=X+Y\sim B(2+4,p)=B(6,p).$$

即 $P\{Z=k\}=\mathrm{C}_6^k p^k(1-p)^{6-k}$，$k=0,1,\ldots,6$。

笔记 更一般地，若 $X\sim B(n_1,p)$，$Y\sim B(n_2,p)$，且 X 与 Y 独立，则 $X+Y\sim B(n_1+n_2,p)$。类似地，若 $X\sim P(\lambda_1)$，$Y\sim P(\lambda_2)$，且 X 与 Y 独立，则 $X+Y\sim P(\lambda_1+\lambda_2)$。

练习 3.38 和与差的分布 将两封信投入 3 个信箱，设 X_1, X_2 分别表示第一个和第二个信箱投进的信的数量，求随机变量 $Y_1=X_1+X_2$ 与 $Y_2=X_1-X_2$ 的分布。

解 由 3.7 知 (X_1, X_2) 的联合分布律。$Y_1 = X_1 + X_2$ 的可能取值为 $0, 1, 2$：

$$P\{Y_1 = 0\} = p_{00} = \frac{1}{9}, \quad P\{Y_1 = 1\} = p_{01} + p_{10} = \frac{4}{9}, \quad P\{Y_1 = 2\} = p_{02} + p_{11} + p_{20} = \frac{4}{9}.$$

$Y_2 = X_1 - X_2$ 的可能取值为 $-2, -1, 0, 1, 2$：

$$P\{Y_2 = -2\} = p_{02} = \frac{1}{9}, \quad P\{Y_2 = -1\} = p_{01} = \frac{2}{9}, \quad P\{Y_2 = 0\} = p_{00} + p_{11} = \frac{3}{9} = \frac{1}{3},$$
$$P\{Y_2 = 1\} = p_{10} = \frac{2}{9}, \quad P\{Y_2 = 2\} = p_{20} = \frac{1}{9}.$$

综上：

$$Y_1 \sim \begin{pmatrix} 0 & 1 & 2 \\ 1/9 & 4/9 & 4/9 \end{pmatrix}, \qquad Y_2 \sim \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 1/9 & 2/9 & 1/3 & 2/9 & 1/9 \end{pmatrix}.$$

3.7.2 连续型随机变量函数的分布

对于连续型随机变量 (X, Y)，求 $Z = g(X, Y)$ 的分布主要有**分布函数法**和**卷积公式法**两种方法。

如果 (X, Y) 中一个是离散型、一个是连续型，可以将事件对离散型的一切可能值进行**全概率公式分解**，逐项求分布（见 3.43）。

3.7.2.1 分布函数法

分布函数法是最基本的方法，适用于任何函数 $g(X, Y)$。

性质 [分布函数法] 求 $Z = g(X, Y)$ 的分布，步骤如下：

1. 先求 Z 的分布函数：
$$F_Z(z) = P\{Z \leqslant z\} = P\{g(X, Y) \leqslant z\} = \iint_{\{(x,y):g(x,y)\leqslant z\}} f(x, y)\,\mathrm{d}x\,\mathrm{d}y;$$
2. 再对 $F_Z(z)$ 求导得概率密度：$f_Z(z) = F_Z'(z)$。

练习 3.39 设随机变量 X 和 Y 相互独立，$X \sim U(0, 1)$，$Y \sim E(1)$，求 $Z = \max\{X, Y\}$ 的分布函数 $F_Z(z)$。

解 X 的分布函数为 $F_X(x) = x$（$0 < x < 1$），Y 的分布函数为 $F_Y(y) = 1 - e^{-y}$（$y > 0$）。

对于独立的 X 和 Y：

$$\max\{X, Y\} \leqslant z \iff X \leqslant z \text{ 且} Y \leqslant z,$$

故

$$F_Z(z) = P\{Z \leqslant z\} = P\{X \leqslant z\}\,P\{Y \leqslant z\} = F_X(z)\,F_Y(z).$$

当 $z \leqslant 0$ 时，$F_Z(z) = 0$；

当 $0 < z < 1$ 时，$F_Z(z) = z(1 - e^{-z})$；

当 $z \geqslant 1$ 时，$F_Z(z) = 1 \cdot (1 - e^{-z}) = 1 - e^{-z}$。

综上：

$$F_Z(z) = \begin{cases} 0, & z \leqslant 0, \\ z(1 - e^{-z}), & 0 < z < 1, \\ 1 - e^{-z}, & z \geqslant 1. \end{cases}$$

✍ **练习 3.40** 设 (X,Y) 的联合概率密度为

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

求 $Z=X+Y$ 的概率密度。

解 用分布函数法。当 $z>0$ 时：

$$F_Z(z)=P\{X+Y\leqslant z\}=\iint_{x+y\leqslant z}e^{-(x+y)}\,\mathrm{d}x\,\mathrm{d}y=\int_0^z\mathrm{d}x\int_0^{z-x}e^{-(x+y)}\,\mathrm{d}y.$$

内层积分：

$$\int_0^{z-x}e^{-(x+y)}\,\mathrm{d}y=e^{-x}(1-e^{-(z-x)})=e^{-x}-e^{-z}.$$

外层积分：

$$F_Z(z)=\int_0^z(e^{-x}-e^{-z})\,\mathrm{d}x=1-e^{-z}-ze^{-z}=1-(1+z)e^{-z}.$$

当 $z\leqslant 0$ 时，$F_Z(z)=0$。求导得：

$$f_Z(z)=\begin{cases}ze^{-z}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

即 $Z\sim\Gamma(2,1)$（形状参数为 2、尺度参数为 1 的 Gamma 分布）。

3.7.2.2 卷积公式与常用分布函数公式

定理 3.3 (卷积公式)

设 $(X,Y)\sim f(x,y)$，则 $Z=X+Y$ 的概率密度为

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

当 X 与 Y 相互独立时，

$$f_Z(z)=\int_{-\infty}^{+\infty}f_X(x)\,f_Y(z-x)\,\mathrm{d}x=\int_{-\infty}^{+\infty}f_X(z-y)\,f_Y(y)\,\mathrm{d}y.$$

♡

类似地，对于其他形式的函数，有如下公式：

性质 [差、积、商的分布] 设 $(X,Y)\sim f(x,y)$，则

1. 差的分布（$Z=X-Y$）：

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

当 X 与 Y 独立时，$f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}f_X(x)\,f_Y(x-z)\,\mathrm{d}x$。

2. 积的分布（$Z=XY$）：

$$f_Z(z)=\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)\,\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{1}{|y|}f\left(\frac{z}{y},y\right)\,\mathrm{d}y.$$

当 X 与 Y 独立时，$f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}\frac{1}{|x|}f_X(x)\,f_Y\left(\frac{z}{x}\right)\,\mathrm{d}x$。

3. 商的分布（$Z = X/Y$）：

$$f_Z(z) = \int_{-\infty}^{+\infty} |y|\, f(yz, y)\, \mathrm{d}y.$$

当 X 与 Y 独立时，$f_Z(z) = \displaystyle\int_{-\infty}^{+\infty} |y|\, f_X(yz)\, f_Y(y)\, \mathrm{d}y$。

笔记 上述卷积公式可用口诀记忆：积谁不换谁，换完求偏导。

例如，$Z = X - Y$ 中对 x 积分（“积谁”= 积 x），则 x 不换，y 换为 $y = x - z$，写成 $f(x, x - z)$；然后“换完求偏导”：$\dfrac{\partial(x-z)}{\partial z} = -1$，取绝对值为 $|-1| = 1$，即得公式 $f_Z(z) = \displaystyle\int_{-\infty}^{+\infty} f(x, x-z)\, \mathrm{d}x$。

更一般地，若 $Z = g(X, Y)$，由 $u = g(x, y)$ 解出 $y = h(x, u)$（要求存在且可导），则

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, h(x, z)) \left| \frac{\partial h(x, z)}{\partial z} \right| \mathrm{d}x.$$

练习 3.41 卷积公式分段积分 设随机变量 X, Y 相互独立，其概率密度分别为

$$f_X(x) = \begin{cases} 2e^{-2x}, & x > 0, \\ 0, & \text{其他}, \end{cases} \qquad f_Y(y) = \begin{cases} 3y^2, & 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

(1) 求 (X, Y) 的概率密度；(2) 求 $Z = X + Y$ 的概率密度。

解 (1) 由独立性，

$$f(x, y) = f_X(x)\, f_Y(y) = \begin{cases} 6y^2 e^{-2x}, & x > 0,\ 0 < y < 1, \\ 0, & \text{其他}. \end{cases}$$

(2) 用卷积公式 $f_Z(z) = \displaystyle\int_{-\infty}^{+\infty} f_X(x)\, f_Y(z - x)\, \mathrm{d}x$。

当 $z \leqslant 0$ 时，$f_Z(z) = 0$。

当 $0 < z < 1$ 时，$0 < z - x < 1$ 要求 $0 < x < z$：

$$f_Z(z) = \int_0^z 2e^{-2x} \cdot 3(z - x)^2\, \mathrm{d}x = 2z + 2e^{-2z} - 2.$$

当 $z \geqslant 1$ 时，$0 < z - x < 1$ 要求 $z - 1 < x < z$（但 $x > 0$，实际为 $z - 1 < x < z$）：

$$f_Z(z) = \int_{z-1}^z 2e^{-2x} \cdot 3(z - x)^2\, \mathrm{d}x = 2e^{-z}.$$

综上：

$$f_Z(z) = \begin{cases} 2z + 2e^{-2z} - 2, & 0 < z < 1, \\ 2e^{-z}, & z \geqslant 1, \\ 0, & \text{其他}. \end{cases}$$

笔记 本题的关键在于卷积公式中积分区间的确定，需根据 $f_Y(z - x) \neq 0$ 即 $0 < z - x < 1$ 的条件来确定 x 的范围，这取决于 z 的取值。当 $0 < z < 1$ 和 $z \geqslant 1$ 时，x 的范围不同，因此需要分段积分。本题亦可用分布函数法验证。

练习 3.42 积的分布：矩形面积 设二维随机变量 (X, Y) 在矩形区域 $D = \{(x, y) : 0 \leqslant x \leqslant 2,\ 0 \leqslant y \leqslant 1\}$ 上服从均匀分布，求边长为 X 和 Y 的矩形面积 $Z = XY$ 的概率密度。

解 由题设，

$$f(x,y)=\begin{cases}\dfrac{1}{2}, & 0\leqslant x\leqslant 2,\ 0\leqslant y\leqslant 1,\\ 0, & \text{其他}.\end{cases}$$

用公式 $f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}\frac{1}{|x|}f\left(x,\frac{z}{x}\right)\mathrm{d}x$。

$f(x,z/x)\neq 0$ 要求 $0\leqslant x\leqslant 2$ 且 $0\leqslant z/x\leqslant 1$，即 $x\geqslant z$ 且 $x\leqslant 2$。

当 $z\leqslant 0$ 或 $z\geqslant 2$ 时，$f_Z(z)=0$。

当 $0<z<2$ 时：

$$f_Z(z)=\int_z^2\frac{1}{2}\cdot\frac{1}{x}\,\mathrm{d}x=\frac{1}{2}(\ln 2-\ln z).$$

综上：

$$f_Z(z)=\begin{cases}\dfrac{1}{2}(\ln 2-\ln z), & 0<z<2,\\ 0, & \text{其他}.\end{cases}$$

笔记 本题亦可用分布函数法验证。$F_Z(z)=P\{XY\leqslant z\}$ 的积分区域取决于 z 与矩形区域的关系：当 $0<z<2$ 时，$F_Z(z)=\dfrac{z}{2}(1+\ln 2-\ln z)$；当 $z\geqslant 2$ 时，$F_Z(z)=1$。求导后结果一致。

练习 3.43 混合型：全概率公式 设随机变量 X_1,X_2,X_3 相互独立，其中 X_1 与 X_2 均服从标准正态分布，X_3 的概率分布为 $P\{X_3=0\}=P\{X_3=1\}=1/2$。令 $Y=X_1X_3^{X_2}$。(1) 求 (X_1,Y) 的分布函数（结果用 $\Phi(x)$ 表示）；(2) 证明 Y 服从标准正态分布。

解 (1) 记 (X_1,Y) 的分布函数为 $F(x,y)$，由全概率公式：

$$\begin{aligned}F(x,y)&=P\{X_1\leqslant x,X_1X_3^{X_2}\leqslant y\}\\&=P\{X_3=0\}\,P\{X_1\leqslant x,0\leqslant y\mid X_3=0\}+P\{X_3=1\}\,P\{X_1\leqslant x,X_1\leqslant y\mid X_3=1\}\\&=\frac{1}{2}P\{X_1\leqslant x\}+\frac{1}{2}P\{X_1\leqslant\min(x,y)\}=\frac{1}{2}\Phi(x)+\frac{1}{2}\Phi(\min(x,y)).\end{aligned}$$

(2) Y 的分布函数为

$$F_Y(y)=\lim_{x\to+\infty}F(x,y)=\lim_{x\to+\infty}\left[\frac{1}{2}\Phi(x)+\frac{1}{2}\Phi(\min(x,y))\right]=\frac{1}{2}+\frac{1}{2}\Phi(y)=\Phi(y).$$

因此 Y 服从标准正态分布 $N(0,1)$。

笔记 本题的特点是 X_3 为离散型而 X_1,X_2 为连续型，求分布函数时需对离散型变量的一切可能值进行全概率分解，分别在 $X_3=0$ 和 $X_3=1$ 的条件下计算。这是处理混合型随机变量问题的标准方法。

性质 [常见分布的可加性] 有些相互独立且服从同类型分布的随机变量，其和的分布也是同类型的。设 X 与 Y 相互独立，则

1. 二项分布：若 $X\sim B(n,p)$，$Y\sim B(m,p)$，则 $X+Y\sim B(n+m,p)$（注意 p 必须相同）。
2. 泊松分布：若 $X\sim P(\lambda_1)$，$Y\sim P(\lambda_2)$，则 $X+Y\sim P(\lambda_1+\lambda_2)$。
3. 正态分布：若 $X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，则 $X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$。更一般地，独立正态变量的线性组合
$$\sum_{i=1}^n c_iX_i\sim N\left(\sum_{i=1}^n c_i\mu_i,\ \sum_{i=1}^n c_i^2\sigma_i^2\right).$$
4. χ^2 分布：若 $X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，则 $X+Y\sim\chi^2(n_1+n_2)$（χ^2 分布详见第六章）。

5. Gamma 分布：若 $X \sim \Gamma(\alpha_1, \beta)$，$Y \sim \Gamma(\alpha_2, \beta)$，则 $X+Y \sim \Gamma(\alpha_1+\alpha_2, \beta)$。特别地，$X_i \sim E(\lambda)$ 独立时，$\sum_{i=1}^n X_i \sim \Gamma(n, \lambda)$。

以上结果对 n 个相互独立的随机变量也成立。

3.7.2.3 max 和 min 的分布

在可靠性理论和排序问题中，$\max\{X,Y\}$ 和 $\min\{X,Y\}$ 的分布具有重要的应用价值。

性质 [max 和 min 的分布] 设 $(X,Y) \sim f(x,y)$，分布函数为 $F(x,y)$。

(1) 一般情形：

1. $Z = \max\{X,Y\}$ 的分布函数为
$$F_{\max}(z) = P\{X \leqslant z, Y \leqslant z\} = F(z,z).$$
2. $Z = \min\{X,Y\}$ 的分布函数为
$$F_{\min}(z) = P\{X \leqslant z\} + P\{Y \leqslant z\} - P\{X \leqslant z, Y \leqslant z\} = F_X(z) + F_Y(z) - F(z,z).$$

(2) 独立情形：

1. 当 X 与 Y 独立时，$F_{\max}(z) = F_X(z) \cdot F_Y(z)$，$F_{\min}(z) = 1 - [1-F_X(z)][1-F_Y(z)]$。
2. 推广到 n 个相互独立的随机变量 $X_1, X_2, \ldots, X_n$（分布函数为 $F_{X_i}(x)$）：
$$F_{\max}(z) = \prod_{i=1}^n F_{X_i}(z), \qquad F_{\min}(z) = 1 - \prod_{i=1}^n [1 - F_{X_i}(z)].$$
3. 当 $X_1, \ldots, X_n$ 独立同分布（分布函数为 $F(x)$，概率密度为 $f(x)$）时：
$$F_{\max}(z) = [F(z)]^n, \quad f_{\max}(z) = n[F(z)]^{n-1} f(z),$$
$$F_{\min}(z) = 1 - [1-F(z)]^n, \quad f_{\min}(z) = n[1-F(z)]^{n-1} f(z).$$

这些结果在数理统计部分有重要应用。

注

1. "$\max\{X_1, \ldots, X_n\} \leqslant z$" 等价于" 所有 $X_i \leqslant z$"，因此取概率时用乘积。
2. "$\min\{X_1, \ldots, X_n\} > z$" 等价于" 所有 $X_i > z$"，因此取逆事件后用乘积。
3. 推导 $F_{\min}(z)$ 的关键步骤：
$$P\{\min(X,Y) \leqslant z\} = 1 - P\{\min(X,Y) > z\} = 1 - P\{X > z, Y > z\} = 1 - [1-F_X(z)][1-F_Y(z)].$$

练习 3.44 一辆新车装有 4 个轮胎并配有 1 个备胎，当使用的轮胎损坏时就用备胎更换。设所有轮胎的寿命均服从参数为 $\lambda = 0.2$（单位：万小时）的指数分布且相互独立，记 T 为开始使用备胎的时间，求 T 的分布函数。

解 开始使用备胎的时刻 T 等于 4 个轮胎中寿命最短的那个的寿命，即 $T = \min\{X_1, X_2, X_3, X_4\}$。

每个轮胎的寿命 $X_i \sim E(0.2)$，分布函数为 $F(x) = 1 - e^{-0.2x}$（$x \geqslant 0$）。

由最小值的分布公式：
$$F_T(t) = 1 - [1-F(t)]^4 = 1 - [e^{-0.2t}]^4 = 1 - e^{-0.8t}, \quad t \geqslant 0.$$
即 $T \sim E(0.8)$，备胎在参数为 0.8 的指数分布时刻开始被使用。

笔记 更一般地，若 $X_1, X_2, \ldots, X_n$ 独立同分布于 $E(\lambda)$，则 $\min\{X_1, \ldots, X_n\} \sim E(n\lambda)$。

练习 3.45 并联系统最大值分布 设系统由两个元件并联组成，两个元件的寿命分别为 X 和 Y，且 X, Y

相互独立，均服从参数为 $\lambda = 1/4$（万小时）的指数分布。求系统寿命 $S = \max\{X, Y\}$ 大于 5 万小时的概率。

解 $X, Y \sim E(1/4)$，分布函数 $F(x) = 1 - e^{-x/4}$（$x \geqslant 0$）。

$S = \max\{X, Y\}$ 的分布函数为 $F_S(s) = [F(s)]^2 = (1 - e^{-s/4})^2$。

所求概率：

$$P\{S > 5\} = 1 - F_S(5) = 1 - (1 - e^{-5/4})^2 = 2e^{-5/4} - e^{-5/2}.$$

练习 3.46 均匀分布最大值密度 设随机变量 X, Y 相互独立，且都在区间 $[a, b]$ 上服从均匀分布，求 $Z = \max\{X, Y\}$ 的概率密度。

解 由独立性，(X, Y) 的概率密度为

$$f(x, y) = \begin{cases} \dfrac{1}{(b-a)^2}, & a \leqslant x \leqslant b,\ a \leqslant y \leqslant b, \\ 0, & \text{其他}. \end{cases}$$

$Z = \max\{X, Y\}$ 的取值范围为 $[a, b]$。当 $a \leqslant z \leqslant b$ 时，

$$F_Z(z) = P\{Z \leqslant z\} = P\{X \leqslant z,\ Y \leqslant z\} = \int_a^z \mathrm{d}x \int_a^z \frac{1}{(b-a)^2}\,\mathrm{d}y = \frac{(z-a)^2}{(b-a)^2}.$$

求导得概率密度：

$$f_Z(z) = F_Z'(z) = \begin{cases} \dfrac{2(z-a)}{(b-a)^2}, & a \leqslant z \leqslant b, \\ 0, & \text{其他}. \end{cases}$$

笔记 这与独立同分布时 $f_{\max}(z) = n[F(z)]^{n-1} f(z)$ 的公式一致：$n = 2$，$F(z) = \dfrac{z-a}{b-a}$，$f(z) = \dfrac{1}{b-a}$，故 $f_{\max}(z) = 2 \cdot \dfrac{z-a}{b-a} \cdot \dfrac{1}{b-a} = \dfrac{2(z-a)}{(b-a)^2}$。

练习 3.47 柯西分布线性变换 设随机变量 X 的概率密度为 $f(x) = \dfrac{1}{\pi(1+x^2)}$（柯西分布），$Y = 2X$，求 Y 的概率密度。

解 $y = g(x) = 2x$ 是严格单调函数，反函数 $x = h(y) = y/2$，$h'(y) = 1/2$。

由公式法：

$$f_Y(y) = f_X\left(\frac{y}{2}\right) \cdot \left|\frac{1}{2}\right| = \frac{1}{\pi\left[1 + (y/2)^2\right]} \cdot \frac{1}{2} = \frac{2}{\pi(4+y^2)}.$$

练习 3.48 正态平方变换 设 $X \sim N(0, 1)$，求 $Y = 2X^2 + 1$ 的概率密度函数。

解 当 $y \leqslant 1$ 时，$2X^2 + 1 \leqslant y$ 无解（$X^2 \geqslant 0$），故 $F_Y(y) = 0$，$f_Y(y) = 0$。

当 $y > 1$ 时：

$$F_Y(y) = P\{2X^2 + 1 \leqslant y\} = P\left\{-\sqrt{\frac{y-1}{2}} \leqslant X \leqslant \sqrt{\frac{y-1}{2}}\right\} = 2\Phi\left(\sqrt{\frac{y-1}{2}}\right) - 1.$$

求导，令 $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$：

$$f_Y(y) = 2\varphi\left(\sqrt{\frac{y-1}{2}}\right) \cdot \frac{1}{2\sqrt{2(y-1)}} = \frac{1}{2\sqrt{\pi(y-1)}}\, e^{-(y-1)/4}.$$

综上：

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{\pi(y-1)}}\exp\left\{-\dfrac{y-1}{4}\right\}, & y>1,\\ 0, & y\leqslant 1.\end{cases}$$

练习 3.49 指数分布条件区域 设二维随机变量 (X,Y) 的概率密度函数为

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

其中 $\lambda>0$。求：(1) $Z=X+Y$ 的概率密度；(2) $N=\min(X,Y)$ 的概率密度。

解 验证归一性：$\displaystyle\int_0^{+\infty}\mathrm{d}y\int_0^y\lambda e^{-\lambda y}\,\mathrm{d}x=\int_0^{+\infty}y\lambda e^{-\lambda y}\,\mathrm{d}y=1$（$Y\sim\Gamma(2,\lambda)$）。

边缘密度：

$$f_X(x)=\int_x^{+\infty}\lambda e^{-\lambda y}\,\mathrm{d}y=e^{-\lambda x},\quad x>0.\qquad \text{即}X\sim E(\lambda).$$

$$f_Y(y)=\int_0^y\lambda e^{-\lambda y}\,\mathrm{d}x=\lambda y\,e^{-\lambda y},\quad y>0.\qquad \text{即}Y\sim\Gamma(2,\lambda).$$

(1) 当 $z>0$ 时，$\{X+Y\leqslant z\}$ 等价于 $\{0<x<y,\ x+y\leqslant z\}$，即 $0<x<z/2$，$x<y\leqslant z-x$：

$$\begin{aligned}F_Z(z)&=\int_0^{z/2}\mathrm{d}x\int_x^{z-x}\lambda e^{-\lambda y}\,\mathrm{d}y=\int_0^{z/2}(e^{-\lambda x}-e^{-\lambda(z-x)})\,\mathrm{d}x\\&=\left[-\frac{1}{\lambda}e^{-\lambda x}-\frac{1}{\lambda}e^{-\lambda z}e^{\lambda x}\right]_0^{z/2}=\frac{1}{\lambda}(1-e^{-\lambda z/2})-\frac{1}{\lambda}e^{-\lambda z}(e^{\lambda z/2}-1)\\&=\frac{1}{\lambda}(1-e^{-\lambda z/2}-e^{-\lambda z/2}+e^{-\lambda z})=\frac{1}{\lambda}(1-2e^{-\lambda z/2}+e^{-\lambda z}).\end{aligned}$$

当 $z\leqslant 0$ 时，$F_Z(z)=0$。求导：

$$f_Z(z)=\frac{1}{\lambda}(\lambda e^{-\lambda z/2}-\lambda e^{-\lambda z})=e^{-\lambda z/2}-e^{-\lambda z},\quad z>0.$$

所以

$$f_Z(z)=\begin{cases}e^{-\lambda z/2}-e^{-\lambda z}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

(2) 在区域 $\{0<x<y\}$ 中，$\min(X,Y)=X$。因此 $N=X$，其密度为

$$f_N(n)=f_X(n)=\begin{cases}e^{-\lambda n}, & n>0,\\ 0, & n\leqslant 0.\end{cases}$$

即 $N\sim E(\lambda)$。

练习 3.50 卷积公式：公式法与分布函数法对照 随机变量 X,Y 相互独立，概率密度分别为

$$f_X(x)=\begin{cases}e^{-x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\qquad f_Y(y)=\begin{cases}\dfrac{1}{2}e^{-y/2}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

求 $Z=X+Y$ 的概率密度。

解 方法一（公式法）：由 $Z=X+Y$，故 $Z>0$。$f_Z(z)=\displaystyle\int_{-\infty}^{+\infty}f_X(x)\,f_Y(z-x)\,\mathrm{d}x$。$f_Y(z-x)\neq 0$ 要求 $z-x>0$，即 $x<z$；$f_X(x)\neq 0$ 要求 $x>0$。当 $z>0$ 时：

$$f_Z(z)=\int_0^z e^{-x}\cdot\frac{1}{2}e^{-(z-x)/2}\,\mathrm{d}x=\frac{1}{2}e^{-z/2}\int_0^z e^{-x/2}\,\mathrm{d}x=\frac{1}{2}e^{-z/2}\left[-2e^{-x/2}\right]_0^z=e^{-z/2}-e^{-z}.$$

方法二（分布函数法）：当 $z>0$ 时，

$$\begin{aligned}F_Z(z)=P\{X+Y\leqslant z\}&=\int_0^z e^{-x}\,\mathrm{d}x\int_0^{z-x}\frac{1}{2}e^{-y/2}\,\mathrm{d}y=\int_0^z e^{-x}\big(1-e^{-(z-x)/2}\big)\,\mathrm{d}x\\&=\int_0^z(e^{-x}-e^{-z/2}e^{-x/2})\,\mathrm{d}x=\big[-e^{-x}+2e^{-z/2}e^{-x/2}\big]_0^z\\&=1-e^{-z}-2e^{-z/2}+2e^{-z}=1-2e^{-z/2}+e^{-z}.\end{aligned}$$

求导：$f_Z(z)=e^{-z/2}-e^{-z}$（$z>0$）。综上：

$$f_Z(z)=\begin{cases}e^{-z/2}-e^{-z}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

笔记 本题用两种方法得到一致结果。当 $X\sim E(1)$，$Y\sim E(1/2)$ 时，$Z=X+Y$ 的密度为 $e^{-z/2}-e^{-z}$。这恰好是两个不同参数指数分布之和，不服从简单的指数分布（因为参数不同，不可直接用 Gamma 可加性公式 $\Gamma(\alpha_1,\beta)+\Gamma(\alpha_2,\beta)=\Gamma(\alpha_1+\alpha_2,\beta)$，这里 β 不同）。

练习 3.51 独立性选择题 设随机变量 X,Y 相互独立，分布函数分别为 $F_X(x),F_Y(y)$，则对任意实数 z，$P\{\max(X,Y)\leqslant z\}$ 等于（　）。

A. $F_X(z)F_Y(z)$　　B. $(1-F_X(z))F_Y(z)$

C. $1-(1-F_X(z))(1-F_Y(z))$　　D. $(1-F_X(z))F_Y(z)$

解 由独立性：

$$P\{\max(X,Y)\leqslant z\}=P\{X\leqslant z,\,Y\leqslant z\}=P\{X\leqslant z\}\,P\{Y\leqslant z\}=F_X(z)\,F_Y(z).$$

答案选 **A**。

笔记 选项 C 为 $P\{\min(X,Y)\leqslant z\}$ 的表达式。不要混淆 max 和 min 的公式。

练习 3.52 综合性：常数、独立性、和的分布、min 的概率 设二维随机变量 (X,Y) 具有联合密度函数

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

求：

(1) 常数 c；

(2) X 与 Y 是否独立？

(3) $Z=X+Y$ 的密度函数；

(4) $P\{\min(X,Y)<1\}$。

解 (1) 由概率密度的归一性：

$$1=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int_0^{+\infty}\int_0^y cxe^{-y}\,\mathrm{d}x\,\mathrm{d}y.$$

内层积分 $\int_0^y x\,\mathrm{d}x=\dfrac{y^2}{2}$，因此

$$\int_0^{+\infty}c\cdot\frac{y^2}{2}e^{-y}\,\mathrm{d}y=\frac{c}{2}\Gamma(3)=\frac{c}{2}\cdot 2!=c=1.$$

(2) 求边缘概率密度：

$$f_X(x)=\int_x^{+\infty}xe^{-y}\,\mathrm{d}y=xe^{-x},\quad x>0.\qquad f_Y(y)=\int_0^y xe^{-y}\,\mathrm{d}x=\frac{y^2}{2}e^{-y},\quad y>0.$$

由于 $f_X(x)\,f_Y(y)=\dfrac{xy^2}{2}e^{-(x+y)}\neq xe^{-y}=f(x,y)$（取 $x=1,y=2$ 可直接验证），故 X 与 Y 不独立。

(3) 方法一（卷积公式）：$f_Z(z)=\int_{-\infty}^{+\infty} f(x,z-x)\,\mathrm{d}x$。$f(x,z-x)\neq 0$ 要求 $0<x<z-x$，即 $0<x<\dfrac{z}{2}$。当 $z>0$ 时：

$$f_Z(z)=\int_0^{z/2} xe^{-(z-x)}\,\mathrm{d}x=e^{-z}\int_0^{z/2} xe^{x}\,\mathrm{d}x=e^{-z}\left[\left(\frac{z}{2}-1\right)e^{z/2}+1\right]=\left(\frac{z}{2}-1\right)e^{-z/2}+e^{-z}.$$

方法二（分布函数法）：当 $z<0$ 时，$F_Z(z)=0$；当 $z\geqslant 0$ 时，积分区域为 $\{0<x<y,\ x+y\leqslant z\}$，即 $0<x<z/2$，$x<y\leqslant z-x$：

$$\begin{aligned}F_Z(z)&=\int_0^{z/2}\mathrm{d}x\int_x^{z-x} xe^{-y}\,\mathrm{d}y=\int_0^{z/2} x(e^{-x}-e^{-(z-x)})\,\mathrm{d}x\\&=\left[1-\left(\frac{z}{2}+1\right)e^{-z/2}\right]-e^{-z}\left[\left(\frac{z}{2}-1\right)e^{z/2}+1\right]=1-ze^{-z/2}-e^{-z}.\end{aligned}$$

求导验证：

$$f_Z(z)=F_Z'(z)=\left(\frac{z}{2}-1\right)e^{-z/2}+e^{-z},\quad z\geqslant 0,$$

与方法一结果一致。综上：

$$f_Z(z)=\begin{cases}\left(\dfrac{z}{2}-1\right)e^{-z/2}+e^{-z}, & z\geqslant 0,\\ 0, & z<0.\end{cases}$$

(4) 在区域 $\{0<x<y\}$ 中，$\min(X,Y)=X$。利用逆事件：

$$\begin{aligned}P\{\min(X,Y)<1\}&=1-P\{\min(X,Y)\geqslant 1\}=1-P\{X\geqslant 1,Y\geqslant 1\}\\&=1-\int_1^{+\infty}\mathrm{d}y\int_1^{y} xe^{-y}\,\mathrm{d}x=1-\int_1^{+\infty}\frac{y^2-1}{2}e^{-y}\,\mathrm{d}y\\&=1-\frac{1}{2}\left[\int_1^{+\infty} y^2e^{-y}\,\mathrm{d}y-\int_1^{+\infty} e^{-y}\,\mathrm{d}y\right]=1-\frac{1}{2}(5e^{-1}-e^{-1})=1-2e^{-1}.\end{aligned}$$

笔记

1. 第 (3) 问分布函数法中，$F_Z(z)=1-ze^{-z/2}-e^{-z}$（注意 e^{-z} 前为负号）。可通过取 $z=2$ 验证：直接计算得 $1-2e^{-1}-e^{-2}$，而 $1-2e^{-1}+e^{-2}$ 与之不符。
2. 第 (4) 问的技巧：在 $\{0<x<y\}$ 的条件下 $\min(X,Y)=X$，因此 $P\{X\geqslant 1,Y\geqslant 1\}$ 的积分区域为 $\{1\leqslant x<y\}$。

练习 3.53 商的分布 X,Y 独立同服从指数分布 $E(\lambda)$，求 $Z=X/Y$ 的概率密度。

解 由商的分布公式，$Z>0$ 时：

$$\begin{aligned}f_Z(z)&=\int_0^{+\infty}|y|\,f_X(zy)\,f_Y(y)\,\mathrm{d}y=\int_0^{+\infty} y\cdot\lambda e^{-\lambda zy}\cdot\lambda e^{-\lambda y}\,\mathrm{d}y\\&=\lambda^2\int_0^{+\infty} ye^{-\lambda(z+1)y}\,\mathrm{d}y=\lambda^2\cdot\frac{1}{\lambda^2(z+1)^2}=\frac{1}{(z+1)^2}.\end{aligned}$$

$$f_Z(z)=\begin{cases}\dfrac{1}{(z+1)^2}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

笔记 $Z=X/Y$ 服从 **Pareto** 型分布，其期望 $E[Z]=\displaystyle\int_0^{+\infty}\frac{z}{(z+1)^2}\,\mathrm{d}z$ 发散（令 $u=z+1$ 可验证）。

练习 3.54 均匀分布 min 的密度 X,Y 独立同服从均匀分布 $U[a,b]$，求 $Z=\min(X,Y)$ 的概率密度。

解 X,Y 的分布函数 $F(x)=\dfrac{x-a}{b-a}$（$a\leqslant x\leqslant b$），密度 $f(x)=\dfrac{1}{b-a}$。由 min 分布公式：

$$F_Z(z)=1-[1-F(z)]^2=1-\left(1-\frac{z-a}{b-a}\right)^2=1-\left(\frac{b-z}{b-a}\right)^2.$$

求导得密度：

$$f_Z(z)=\frac{2(b-z)}{(b-a)^2}\cdot\frac{1}{b-a}=\frac{2(b-z)}{(b-a)^2},\quad a\leqslant z\leqslant b.$$

所以

$$f_Z(z)=\begin{cases}\dfrac{2(b-z)}{(b-a)^2}, & a\leqslant z\leqslant b,\\ 0, & \text{其他}.\end{cases}$$

笔记 对照 3.46 中 max 的密度 $\dfrac{2(z-a)}{(b-a)^2}$，可见 min 和 max 的密度关于区间中点对称。

练习 3.55 泊松可加性 $X\sim P(5)$，$Y\sim P(3)$，且 X 与 Y 独立，则（　）

A. $X+Y\sim P(2)$　B. $X-Y\sim P(2)$　C. $X+Y\sim P(8)$　D. $X-Y\sim P(8)$

解 泊松分布具有可加性：若 $X\sim P(\lambda_1)$，$Y\sim P(\lambda_2)$ 且独立，则 $X+Y\sim P(\lambda_1+\lambda_2)$。故 $X+Y\sim P(8)$。泊松分布不具有"可减性"：$X-Y$ 可取负值，不服从泊松分布。答案选 **C**。

练习 3.56 次序统计量的分布 $X_1,X_2,\ldots,X_5$ 为总体 $N(1,4)$ 的简单随机样本，$X_{(5)}=\max(X_1,\ldots,X_5)$，求 $P(X_{(5)}<3)$。

解 $P(X<3)=\Phi\left(\dfrac{3-1}{2}\right)=\Phi(1)\approx 0.8413$。

由独立性，$X_{(5)}=\max$ 的分布函数 $F_{(5)}(z)=[F(z)]^5$，故

$$P(X_{(5)}<3)=\Phi^5(1)=0.8413^5\approx 0.4214.$$

练习 3.57min+max=X+Y (X,Y) 服从区域 $G=\{(x,y):0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 2\}$ 上的均匀分布，令 $Z=\min(X,Y)$，$W=\max(X,Y)$，求 $P(Z+W\geqslant 1)$。

解 由于 $Z+W=X+Y$ 恒成立，$P(Z+W\geqslant 1)=P(X+Y\geqslant 1)$。$(X,Y)$ 的密度为 $f(x,y)=1/2$（$0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 2$）。

$$P(X+Y\geqslant 1)=1-P(X+Y<1)=1-\int_0^1\mathrm{d}x\int_0^{1-x}\frac{1}{2}\,\mathrm{d}y=1-\frac{1}{2}\int_0^1(1-x)\,\mathrm{d}x=1-\frac{1}{4}=\frac{3}{4}.$$

练习 3.58 半正态联合密度与极坐标变换 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{2}{\pi}e^{-\frac{x^2+y^2}{2}}, & x>0,\ y>0,\\ 0, & \text{其他}.\end{cases}$$

求：(1) 边缘密度；(2) 独立性；(3) $Z=X^2+Y^2$ 的密度。

解 (1) 利用 $\displaystyle\int_0^{+\infty}e^{-y^2/2}\,\mathrm{d}y=\sqrt{\frac{\pi}{2}}$：

$$f_X(x)=\frac{2}{\pi}e^{-x^2/2}\int_0^{+\infty}e^{-y^2/2}\,\mathrm{d}y=\frac{2}{\pi}e^{-x^2/2}\cdot\sqrt{\frac{\pi}{2}}=\sqrt{\frac{2}{\pi}}e^{-x^2/2},\quad x>0.$$

同理 $f_Y(y)=\sqrt{2/\pi}\,e^{-y^2/2}$（$y>0$）。此为半正态分布（折叠正态分布）。

(2) $f(x,y)=\dfrac{2}{\pi}e^{-(x^2+y^2)/2}=f_X(x)\,f_Y(y)$，故 X 与 Y 独立。

(3) $Z=X^2+Y^2>0$。用极坐标变换 $x=r\cos\theta$，$y=r\sin\theta$（$0<\theta<\pi/2$，$r>0$）：

$$F_Z(z)=\iint_{x^2+y^2\leqslant z,\ x>0,\ y>0}\frac{2}{\pi}e^{-(x^2+y^2)/2}\,\mathrm{d}x\,\mathrm{d}y=\frac{2}{\pi}\int_0^{\pi/2}\mathrm{d}\theta\int_0^{\sqrt{z}}r\,e^{-r^2/2}\,\mathrm{d}r$$

$$= \frac{2}{\pi} \cdot \frac{\pi}{2} \left[-e^{-r^2/2} \right]_0^{\sqrt{z}} = 1 - e^{-z/2}, \quad z > 0.$$

求导得 $f_Z(z) = \frac{1}{2} e^{-z/2}$（$z > 0$），即 $Z \sim E(1/2)$（或等价地 $Z \sim \chi^2(2)$）。

笔记 本题说明：两个独立半正态变量的平方和服从指数分布（自由度为 **2** 的 χ^2 分布）。更一般地，若 X_i 独立同分布于 $N(0, \sigma^2)$，则 $\sum_{i=1}^n X_i^2 \sim \sigma^2 \chi^2(n)$。

练习 3.59 混合条件分布求 F_Y $P\{X = 1\} = P\{X = 2\} = 1/2$，在 $X = i$ 的条件下 $Y \sim U(0, i)$。求 Y 的分布函数 $F_Y(y)$ 和密度 $f_Y(y)$。

解 由全概率公式：$F_Y(y) = P\{X = 1\} F_{Y|X=1}(y) + P\{X = 2\} F_{Y|X=2}(y) = \frac{1}{2} F_1(y) + \frac{1}{2} F_2(y)$。

$$Y|X = 1 \sim U(0,1)\text{：} F_1(y) = \begin{cases} 0, & y \leqslant 0, \\ y, & 0 < y < 1, \\ 1, & y \geqslant 1. \end{cases}$$

$$Y|X = 2 \sim U(0,2)\text{：} F_2(y) = \begin{cases} 0, & y \leqslant 0, \\ y/2, & 0 < y < 2, \\ 1, & y \geqslant 2. \end{cases}$$ 分区间计算 $F_Y(y)$：

- $y \leqslant 0$：$F_Y(y) = 0$。
- $0 < y \leqslant 1$：$F_Y(y) = \frac{1}{2} y + \frac{1}{2} \cdot \frac{y}{2} = \frac{3y}{4}$。
- $1 < y \leqslant 2$：$F_Y(y) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{y}{2} = \frac{1}{2} + \frac{y}{4}$。
- $y \geqslant 2$：$F_Y(y) = 1$。

求导得密度：

$$F_Y(y) = \begin{cases} 0, & y \leqslant 0, \\ \frac{3y}{4}, & 0 < y \leqslant 1, \\ \frac{1}{2} + \frac{y}{4}, & 1 < y \leqslant 2, \\ 1, & y \geqslant 2. \end{cases} \qquad f_Y(y) = \begin{cases} \frac{3}{4}, & 0 < y \leqslant 1, \\ \frac{1}{4}, & 1 < y \leqslant 2, \\ 0, & \text{其他}. \end{cases}$$

第 4 章 随机变量的数字特征

$$
\begin{cases}
\text{数学期望}\begin{cases}
\text{计算公式}\begin{cases}
\text{离散型}\begin{cases}
\text{一维:}\begin{cases} EX=\sum\limits_i x_ip_i \\ E\varphi(X)=\sum\limits_i \varphi(x_i)p_i \end{cases} \\
\text{二维: } E[\varphi(X,Y)]=\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}\varphi(x_i,y_j)p_{ij}
\end{cases} \\
\text{连续型}\begin{cases}
\text{一维:}\begin{cases} EX=\int_{-\infty}^{+\infty} xf(x)dx \\ E\varphi(X)=\int_{-\infty}^{+\infty}\varphi(x)f(x)dx \end{cases} \\
\text{二维: } E[\varphi(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\varphi(x,y)f(x,y)dxdy
\end{cases}
\end{cases} \\
\text{性质}\begin{cases}
(1)\ \text{常数的期望不变: } EC=C,\ \text{特别地: } E(EX)=EX \\
(2)\ \text{线性性质: 如果}EX\ \text{和}EY\ \text{都存在, 则}E(aX+bY)=aEX+bEY \\
(3)\ \text{如果}EX\ \text{存在, 则}E(kX)=kEX \\
(4)\ \text{若}X,Y\ \text{相互独立, 且期望都存在, 则}E(XY)=EXEY
\end{cases}
\end{cases} \\
\text{方差}\begin{cases}
\text{计算公式: } DX=E(X-EX)^2=E(X^2)-(EX)^2 \\
\text{性质}\begin{cases}
(1)\ \text{常数的方差为零: } DC=0,\quad (2)DX\geqslant 0\Rightarrow E(X^2)\geqslant (EX)^2 \\
(3)\ D(aX+b)=a^2DX,\quad (4)\ D(X\pm Y)=DX+DY\pm 2\mathrm{Cov}(X,Y) \\
(5)\ \text{若}X,Y\ \text{相互独立, 则}D(X\pm Y)=DX+DY \\
(6)\ DX=E(X-EX)^2\leqslant E(X-c)^2
\end{cases}
\end{cases} \\
\text{协方差}\begin{cases}
\text{计算公式: } \mathrm{Cov}(X,Y)=E(X-EX)(Y-EY)=E(XY)-EXEY \\
\text{其中}E(XY)=\begin{cases}
\sum\limits_i\sum\limits_j x_iy_jP\{X=x_i,Y=y_j\}\quad (\text{离散型}) \\
\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy\quad (\text{连续型})
\end{cases} \\
\text{性质}\begin{cases}
(1)\ \mathrm{Cov}(X,X)=DX,\quad (2)\ \mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X) \\
(3)\ \mathrm{Cov}(X,C)=0(C\text{是常数}),(4)\ \mathrm{Cov}(aX+bY,Z)=a\mathrm{Cov}(X,Z)+b\mathrm{Cov}(Y,Z)
\end{cases}
\end{cases} \\
\text{相关系数}\begin{cases}
\text{计算公式: } \rho_{XY}=\dfrac{\mathrm{Cov}(X,Y)}{\sqrt{DX}\sqrt{DY}},\ |\rho_{XY}|\leqslant 1 \\
\text{性质}\begin{cases}
(1)\ |\rho_{XY}|=1\Leftrightarrow P\{Y=aX+b\}=1 \\
(2)\ \rho_{XY}=0\Leftrightarrow X,Y\ \text{不相关}\Leftrightarrow\begin{cases}
(1)\ \mathrm{Cov}(X,Y)=0 \\
(2)\ E(XY)=EXEY \\
(3)\ D(X\pm Y)=DX+DY
\end{cases} \\
(3)\ \text{若}(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho),\ \text{则}X,Y\ \text{独立与}X,Y\ \text{不相关等价}
\end{cases}
\end{cases}
\end{cases}
$$

内容提要

- 数学期望的定义与性质 4.1 4.1.3
- 方差的定义与性质 4.4 4.2.2
- 常见分布的数学期望与方差 4.1
- 协方差 4.5 4.4.1
- 相关系数 4.6 4.4.2
- 独立性与不相关性 4.4.2
- 矩（原点矩、中心矩、混合矩） 7.2
- 切比雪夫不等式 5.1

4.1 数学期望

4.1.1 数学期望的定义

定义 4.1 (数学期望)

设 X 是一个随机变量。

1. 离散型：若 X 的分布律为 $P\{X=x_i\}=p_i\ (i=1,2,\dots)$，且级数 $\sum\limits_{i=1}^{\infty} x_ip_i$ 绝对收敛，则称 X 的数学期望存在，其值为
$$EX=\sum_{i=1}^{\infty} x_ip_i.$$
若级数不绝对收敛，则称 X 的数学期望不存在。
2. 连续型：若 X 的概率密度为 $f(x)$，且积分 $\int_{-\infty}^{+\infty} xf(x)\,\mathrm{d}x$ 绝对收敛，则称 X 的数学期望存在，其值为
$$EX=\int_{-\infty}^{+\infty} xf(x)\,\mathrm{d}x.$$

♣

练习 4.1 设连续型随机变量 X 的概率密度为
$$f(x)=\frac{1}{\pi(1+x^2)},\quad -\infty<x<+\infty,$$
则 EX（　）。

A. 等于 0　B. 等于 1　C. 等于 π　D. 不存在

解 应选 D。由于
$$\int_{-\infty}^{+\infty}\frac{x}{\pi(1+x^2)}\,\mathrm{d}x=\frac{1}{2\pi}\ln(1+x^2)\Big|_{-\infty}^{+\infty}$$
该积分发散，因此 EX 不存在。

练习 4.2 已知连续型随机变量 X 与 Y 有相同的概率密度，且
$$X\sim f(x)=\begin{cases}2x\theta^2, & 0<x<\dfrac{1}{\theta}\ (\theta>0),\\ 0, & \text{其他},\end{cases}\qquad E[a(X+2Y)]=\frac{1}{\theta},$$
则 $a=$ _____。　A. $\dfrac{2}{3}$　B. $\dfrac{1}{2}$　C. $\dfrac{1}{3}$　D. $\dfrac{1}{6}$

解 应选 B。由于 X 与 Y 同分布，$EX=EY$，故
$$E[a(X+2Y)]=a(3EX)=3a\int_0^{1/\theta}2x^2\theta^2\,\mathrm{d}x=3a\cdot\frac{2}{3\theta}=\frac{2a}{\theta}=\frac{1}{\theta},$$
解得 $a=\dfrac{1}{2}$。

练习 4.3 设随机变量 X 服从参数为 λ 的泊松分布，且已知 $E[(X-1)(X-2)]=1$，则 $\lambda=$ _____。

解 由 $EX=DX=\lambda$，得

$$E[(X-1)(X-2)]=E(X^2)-3EX+2=DX+(EX)^2-3EX+2=\lambda^2-2\lambda+2=1,$$

即 $(\lambda-1)^2=0$，解得 $\lambda=1$。

练习 4.4 设随机变量 X 的概率分布为 $P\{X=k\}=\dfrac{C}{k!}$, $k=0,1,2,\ldots$，则 $E(|X-EX|)=$ _____。

解 应填 2。由 $\sum\limits_{k=0}^{\infty}P\{X=k\}=\sum\limits_{k=0}^{\infty}\dfrac{C}{k!}=Ce=1$，得 $C=e^{-1}$，故 $X\sim P(1)$，$EX=1$。

$$\begin{aligned}E(|X-EX|)=E(|X-1|)&=\sum_{k=0}^{\infty}P(X=k)|k-1|\\&=P(X=0)\cdot 1+\sum_{k=2}^{\infty}P(X=k)(k-1)\\&=e^{-1}+\left[\sum_{k=0}^{\infty}P(X=k)(k-1)-P(X=0)(0-1)-P(X=1)(1-1)\right]\\&=e^{-1}+\left[(1\cdot 1-1)+e^{-1}\right]=e^{-1}+(0+e^{-1})=2.\end{aligned}$$

笔记 由泊松分布的分布列 $P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!}$ 与本题的 $P\{X=k\}=\dfrac{C}{k!}$ 比较可知，X 服从参数 $\lambda=1$ 的泊松分布。直接利用 $EX=\lambda=1$ 可简化计算。

练习 4.5 一部机器一天内发生故障的概率为 0.2，机器发生故障时全天停产。若一周 5 个工作日里：无故障获利 10 万元，发生一次故障获利 5 万元，发生两次故障获利 0 元，发生三次及以上故障亏损 2 万元。求一周内利润的期望。

解 设一周内故障次数为 X，则 $X\sim B(5,0.2)$。

$$P\{X=0\}=0.8^5=0.32768,\quad P\{X=1\}=\mathrm{C}_5^1\times 0.2\times 0.8^4=0.4096,$$

$$P\{X=2\}=\mathrm{C}_5^2\times 0.04\times 0.512=0.2048,\quad P\{X\geqslant 3\}=1-\sum_{k=0}^{2}P\{X=k\}=0.05792.$$

设利润为 Y（万元），则

$$EY=10\times 0.32768+5\times 0.4096+0\times 0.2048+(-2)\times 0.05792=5.20896(\text{万元}).$$

练习 4.6 某商店销售某商品，销量 X 在 $(50,100)$ 上服从均匀分布。若每销售一单位获利 500 元；如果需求量大于进货量 n，可以从其他部门调剂，此时每单位获利 300 元；如果有积压，则每单位亏损 200 元。进货量 n 为多少时平均获利最大？最大为多少？

解 设进货量为 n（$n\in[50,100]$），获利为

$$Y=\begin{cases}500n+300(X-n)=300X+200n, & X\geqslant n,\\ 500X-200(n-X)=700X-200n, & X<n.\end{cases}$$

$$
\begin{aligned}
EY &= \frac{1}{50}\left[\int_{50}^{n}(700x-200n)\,\mathrm{d}x+\int_{n}^{100}(300x+200n)\,\mathrm{d}x\right]\\
&= \frac{1}{50}\left[350(n^2-2500)-200n(n-50)+150(10000-n^2)+200n(100-n)\right]\\
&= \frac{1}{50}\left[-200n^2+30000n+625000\right]\\
&= -4n^2+600n+12500.
\end{aligned}
$$

令 $(EY)'=-8n+600=0$，得 $n=75$。此时

$$EY_{\max}=-4\times 75^2+600\times 75+12500=35000.$$

练习 4.7 假定国际市场每年对我国某种商品的需求量 X（单位：吨）服从 $[2000,4000]$ 上的均匀分布。每售出一吨获利 3 万美元，若销售不出则每吨需仓储费 1 万美元。问：每年组织多少资源，才能使期望收益最大？

解 设进货量为 y，收益为

$$L(y,X)=\begin{cases}3y, & X\geqslant y,\\ 3X-(y-X)=4X-y, & X<y.\end{cases}$$

$$E[L]=\frac{1}{2000}\left[\int_{2000}^{y}(4x-y)\,\mathrm{d}x+\int_{y}^{4000}3y\,\mathrm{d}x\right]=\frac{1}{2000}(-2y^2+14000y-8000000).$$

令 $\dfrac{\mathrm{d}E[L]}{\mathrm{d}y}=\dfrac{1}{2000}(-4y+14000)=0$，得 $y=3500$ 吨。

练习 4.8 某决赛采取五局三胜制（三局先胜后比赛终止），双方每局胜率相同。目前比分 1:0（对方领先）。若每场比赛总收入 160 万元，胜方分得 120 万，败方分得 40 万，求对方收入的期望。

解 对方需在余下最多 4 局中至少胜 3 局：

$$P\{\text{对方夺冠}\}=\mathrm{C}_4^3\left(\frac{1}{2}\right)^4+\mathrm{C}_4^4\left(\frac{1}{2}\right)^4=\frac{5}{16}.$$

夺冠收入 120 万元，否则收入 40 万元：

$$E=120\times\frac{5}{16}+40\times\frac{11}{16}=65(\text{万元}).$$

定义 4.2 (随机变量函数的期望)

设 $Y=g(X)$（g 为连续函数）。

1. 离散型：若 $\sum\limits_{i=1}^{\infty}g(x_i)p_i$ 绝对收敛，则

$$E[g(X)]=\sum_{i=1}^{\infty}g(x_i)p_i.$$

2. 连续型：若 $\int_{-\infty}^{+\infty}g(x)f(x)\,\mathrm{d}x$ 绝对收敛，则

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

♣

笔记 (1) 数学期望又称为概率平均值，常简称期望或均值。它是描述随机变量平均取值状况的数字特征，刻画了随机变量一切可能值的集中位置。

(2) 定义中要求级数（或积分）绝对收敛，这是为了保证期望的值不依赖于求和（或积分）的顺序。

(3) 并非所有随机变量都存在数学期望。例如柯西分布 $X \sim f(x) = \dfrac{1}{\pi(1+x^2)}$，由于

$$\int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)}\,\mathrm{d}x = +\infty,$$

其数学期望不存在。

练习 4.9 设随机变量 $X \sim B(2, 0.1)$，则 $E(2^X) =$ _____。

解

$$E(2^X) = 0.9^2 \times 2^0 + \mathrm{C}_2^1 \times 0.1 \times 0.9 \times 2^1 + 0.1^2 \times 2^2 = 1.21.$$

练习 4.10 设连续型随机变量 X 的概率密度为 $f(x) = \dfrac{1}{\pi(1+x^2)}\ (-\infty < x < +\infty)$，求 $E(\min(|X|, 1))$。

解 由对称性，

$$\begin{aligned}
E(\min(|X|,1)) &= 2\int_0^{+\infty} \min(x,1)\, f(x)\,\mathrm{d}x = 2\left[\int_0^1 x f(x)\,\mathrm{d}x + \int_1^{+\infty} f(x)\,\mathrm{d}x\right] \\
&= 2\left(\int_0^1 \frac{x}{\pi(1+x^2)}\,\mathrm{d}x + \int_1^{+\infty} \frac{1}{\pi(1+x^2)}\,\mathrm{d}x\right) \\
&= \frac{1}{\pi}\ln(1+x^2)\Big|_0^1 + \frac{2}{\pi}\arctan x\Big|_1^{+\infty} = \frac{\ln 2}{\pi} + \frac{1}{2}.
\end{aligned}$$

练习 4.11 已知连续型随机变量 Y 的概率密度为

$$f(y) = \begin{cases} \dfrac{y}{a^2}\, e^{-\frac{y^2}{2a^2}}, & y > 0,\ a > 0, \\ 0, & y \leqslant 0, \end{cases}$$

求 $Z = \dfrac{1}{Y}$ 的数学期望 EZ。

解

$$EZ = E\left(\frac{1}{Y}\right) = 0 + \int_0^{+\infty} \frac{1}{y} \cdot \frac{y}{a^2}\, e^{-\frac{y^2}{2a^2}}\,\mathrm{d}y = \frac{1}{a^2}\int_0^{+\infty} e^{-\frac{y^2}{2a^2}}\,\mathrm{d}y.$$

利用正态分布密度的归一性：$N(0, a^2)$ 的密度 $\varphi(y) = \dfrac{1}{\sqrt{2\pi}\,a}\, e^{-\frac{y^2}{2a^2}}$ 满足 $\displaystyle\int_{-\infty}^{+\infty} \varphi(y)\,\mathrm{d}y = 1$，由对称性

$$\int_0^{+\infty} \frac{1}{\sqrt{2\pi}\,a}\, e^{-\frac{y^2}{2a^2}}\,\mathrm{d}y = \frac{1}{2},$$

故 $\displaystyle\int_0^{+\infty} e^{-\frac{y^2}{2a^2}}\,\mathrm{d}y = \frac{\sqrt{2\pi}\,a}{2}$，于是

$$EZ = \frac{1}{a^2} \cdot \frac{\sqrt{2\pi}\,a}{2} = \frac{\sqrt{2\pi}}{2a}.$$

练习 4.12 设随机变量 X 在 $[-1, 2]$ 上服从均匀分布，令

$$Y = \begin{cases} 1, & X > 0, \\ 0, & X = 0, \\ -1, & X < 0, \end{cases}$$

则 $DY =$ _____。

解 $X \sim U(-1,2)$，$f(x)=\dfrac{1}{3}\,(-1<x<2)$。

$$P\{Y=1\}=P\{X>0\}=\int_0^2 \frac{1}{3}\,\mathrm{d}x=\frac{2}{3}$$
$$P\{Y=0\}=P\{X=0\}=0$$
$$P\{Y=-1\}=P\{X<0\}=\int_{-1}^0 \frac{1}{3}\,\mathrm{d}x=\frac{1}{3}$$
$$EY=1\times\frac{2}{3}+(-1)\times\frac{1}{3}=\frac{1}{3},\quad E(Y^2)=1^2\times\frac{2}{3}+(-1)^2\times\frac{1}{3}=1$$
$$DY=E(Y^2)-(EY)^2=1-\frac{1}{9}=\frac{8}{9}.$$

练习 4.13 已知随机变量 $(X,Y)\sim N(0,0;\,1,1;\,0.5)$，$U=\max\{X,Y\}$，$V=\min\{X,Y\}$，求 $E(U+V)$ 和 $E(UV)$。

解 利用恒等式

$$U+V=X+Y,\quad UV=XY.$$

故

$$E(U+V)=EX+EY=0,\quad E(UV)=E(XY)=\mathrm{Cov}(X,Y)+EX\cdot EY=0.5.$$

4.1.2 二维随机变量函数的数学期望

定义 4.3 (二维随机变量函数的期望)

设 (X,Y) 为二维随机变量，$g(X,Y)$ 为连续函数。

1. 离散型：若 (X,Y) 的联合分布律为 $p_{ij}=P\{X=x_i,Y=y_j\}$ $(i,j=1,2,\dots)$，且级数 $\sum\limits_i\sum\limits_j |g(x_i,y_j)|\,p_{ij}$ 收敛，则
$$E[g(X,Y)]=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i,y_j)\,p_{ij}.$$
2. 连续型：若 (X,Y) 的联合概率密度为 $f(x,y)$，且积分 $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}|g(x,y)|\,f(x,y)\,\mathrm{d}x\mathrm{d}y$ 收敛，则
$$E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,f(x,y)\,\mathrm{d}x\mathrm{d}y.$$

♣

4.1.3 数学期望的性质

性质 数学期望具有以下性质（假设所涉及的期望均存在）：

1. $E(C)=C$，其中 C 为任意常数。特别地，$E(EX)=EX$。
2. $E(kX)=kEX$，其中 k 为任意常数。
3. 线性性质：$E(aX+bY)=aEX+bEY$。更一般地，
$$E\left(\sum_{i=1}^{n} a_iX_i\right)=\sum_{i=1}^{n} a_iEX_i.$$
4. 若 X,Y 相互独立，则 $E(XY)=EX\cdot EY$。

笔记 (1) 性质 4 的逆命题不成立，即由 $E(XY)=EX\cdot EY$ 不能推出 X 与 Y 独立。

(2) 在计算期望时，如果积分区间关于原点对称，应当考虑被积函数的奇偶性，这往往能简化计算。

(3) 实数域上的 Gamma 函数定义为 $\Gamma(\alpha)=\int_0^{+\infty}t^{\alpha-1}e^{-t}\,\mathrm{d}t$，其有以下性质：

$$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi},\quad \Gamma(x+1)=x\Gamma(x),\quad \Gamma(n)=(n-1)!.$$

在计算正态分布的各阶矩时，利用 Gamma 函数能够降低计算量。

练习 4.14 设随机变量 X 的概率密度为 $f(x)=\begin{cases}\frac{1}{2}x, & 0<x<2,\\ 0, & 其他,\end{cases}$ $F(x)$ 为 X 的分布函数，则 $P\{F(X)>EX-1\}=$ _____。

解 应填 $\dfrac{2}{3}$。令 $Y=F(X)$，因为 $f(x)$ 严格单调增加，所以 $Y\sim U(0,1)$。又 $EX=\frac{1}{2}\int_0^2 x^2\,\mathrm{d}x=\dfrac{4}{3}$，故

$$P\{F(X)>EX-1\}=P\left(Y>\frac{1}{3}\right)=\frac{2}{3}.$$

笔记 本题若不知 $Y=F(X)\sim U(0,1)$ 的结论，可直接计算：

$$F(x)=\begin{cases}0, & x\leqslant 0,\\ \dfrac{x^2}{4}, & 0<x<2,\\ 1, & x\geqslant 2.\end{cases}$$

于是 $P\left(F(X)>\dfrac{1}{3}\right)=P\left(\dfrac{X^2}{4}>\dfrac{1}{3},\ 0\leqslant X<2\right)=P\left(X>\dfrac{2}{\sqrt{3}}\right)=\displaystyle\int_{\frac{2}{\sqrt{3}}}^{2}x\,\mathrm{d}x=\dfrac{2}{3}$。

练习 4.15 设随机变量 X,Y 相互独立，且 X 服从区间 $[0,2]$ 上的均匀分布，Y 服从参数为 3 的指数分布，则 $E(XY)=$ _____。

解 X,Y 相互独立，故

$$E(XY)=EX\cdot EY=1\times\frac{1}{3}=\frac{1}{3}.$$

练习 4.16 甲、乙两人相约于 12:00~13:00 会面。设 X,Y 分别是甲、乙到达超过 12:00 的时间（单位：小时），且相互独立，已知

$$f_X(x)=\begin{cases}3x^2, & 0<x<1,\\ 0, & 其他,\end{cases}\qquad f_Y(y)=\begin{cases}2y, & 0<y<1,\\ 0, & 其他,\end{cases}$$

求先到达者需要等待的时间的数学期望。

解 所求为 $E(|X-Y|)$。由独立性，$f(x,y)=6x^2y\ (0<x,y<1)$，故

$$\begin{aligned}E(|X-Y|)&=\int_0^1\int_0^1|x-y|\cdot 6x^2y\,\mathrm{d}x\mathrm{d}y\\&=\iint_{x>y}(x-y)\cdot 6x^2y\,\mathrm{d}x\mathrm{d}y+\iint_{x<y}(y-x)\cdot 6x^2y\,\mathrm{d}x\mathrm{d}y\\&=\frac{1}{12}+\frac{1}{6}=\frac{1}{4}.\end{aligned}$$

练习 4.17 从长度为 3 的线段上任取两个点，求两点间距离 D 的数学期望和方差。

解 设两个点的位置分别为 X_1,X_2，独立且均服从 $U(0,3)$，则距离 $D=|X_1-X_2|$。单个均匀分布 $U(0,3)$ 的概率密度为：

$$f_{X_i}(x_i)=\begin{cases}\dfrac{1}{3}, & 0<x_i<3\\ 0, & 其他\end{cases}\qquad (i=1,2)$$

由独立性，联合密度 = 边缘密度的乘积：

$$f(x_1,x_2)=f_{X_1}(x_1)\cdot f_{X_2}(x_2)=\begin{cases}\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{9}, & 0<x_1<3,\ 0<x_2<3\\ 0, & \text{其他}\end{cases}$$

由对称性，只需计算 $x_1>x_2$ 的区域再乘以 2：

$$ED=2\int_0^3\int_0^{x_1}(x_1-x_2)\cdot\frac{1}{9}\,\mathrm{d}x_2\,\mathrm{d}x_1=\frac{2}{9}\int_0^3\frac{x_1^2}{2}\,\mathrm{d}x_1=1.$$

$$E(X_i)=\frac{3}{2},\quad E(X_i^2)=DX_i+(EX_i)^2=\frac{3}{4}+\frac{9}{4}=3.$$

$$E(D^2)=E[(X_1-X_2)^2]=E(X_1^2)-2EX_1EX_2+E(X_2^2)=3-2\times\frac{9}{4}+3=\frac{3}{2},$$

$$DD=ED^2-(ED)^2=\frac{3}{2}-1=\frac{1}{2}.$$

命题 4.1 (均匀分布次序统计量的通用结论)

对于 $U(0,1)$ 上的第 k 个次序统计量 $X_{(k)}$，其期望为：

$$\boldsymbol{E(X_{(k)})=\frac{k}{n+1}}$$

当 $k=1$（最小值）：$E(X_{(1)})=\frac{1}{n+1}$；当 $k=n$（最大值）：$E(X_{(n)})=\frac{n}{n+1}$

证明 对于 $U(0,1)$，单个变量的分布函数 $F(x)=x\ (0\le x\le 1)$，密度 $f(x)=1$。第 k 个次序统计量 $X_{(k)}$ 的密度公式为：

$$f_{X_{(k)}}(x)=\frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1}[1-F(x)]^{n-k}f(x)$$

代入 $F(x)=x, f(x)=1$，得：

$$f_{X_{(k)}}(x)=\frac{n!}{(k-1)!(n-k)!}x^{k-1}(1-x)^{n-k},\quad 0<x<1$$

这正是 Beta 分布 Beta$(k,n-k+1)$ 的密度。Beta 分布 Beta(α,β) 的期望为 $\frac{\alpha}{\alpha+\beta}$，这里 $\alpha=k,\beta=n-k+1$，因此：

$$E(X_{(k)})=\frac{k}{k+(n-k+1)}=\frac{k}{n+1}$$

练习 4.18 设相互独立的随机变量 $X_1,X_2,\ldots,X_n$ 均服从 $(0,\theta)$ 上的均匀分布，求 $M=\max\{X_1,\ldots,X_n\}$ 和 $N=\min\{X_1,\ldots,X_n\}$ 的数学期望。

解 $F(x)=\dfrac{x}{\theta}$，$f(x)=\dfrac{1}{\theta}\ (0<x<\theta)$。

最大值：$F_M(x)=[F(x)]^n=\left(\dfrac{x}{\theta}\right)^n$，$f_M(x)=\dfrac{n\,x^{n-1}}{\theta^n}$，

$$EM=\int_0^\theta x\cdot\frac{n\,x^{n-1}}{\theta^n}\,\mathrm{d}x=\frac{n}{n+1}\theta.$$

最小值：$1-F_N(x)=[1-F(x)]^n=\left(1-\dfrac{x}{\theta}\right)^n$，$f_N(x)=\dfrac{n}{\theta}\left(1-\dfrac{x}{\theta}\right)^{n-1}$，

$$EN=\int_0^\theta x\cdot\frac{n}{\theta}\left(1-\frac{x}{\theta}\right)^{n-1}\mathrm{d}x=\frac{1}{n+1}\theta.$$

练习 4.19 在区间 $[0,1]$ 上任取 n 个点 $X_1,X_2,\ldots,X_n$，记

$$X_{(1)}=\min\{X_1,\ldots,X_n\},\ X_{(n)}=\max\{X_1,\ldots,X_n\},\ R=X_{(n)}-X_{(1)}\text{（极差）}$$

求 ER。

解 设 $X_1, X_2, \ldots, X_n$ 是独立同分布的随机变量，都服从区间 $[0,1]$ 上的均匀分布 $U(0,1)$。定义次序统计量：

- 最小值 $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$（第 1 个次序统计量）
- 最大值 $X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$（第 n 个次序统计量）

极差 $R = X_{(n)} - X_{(1)}$，表示 n 个点覆盖的区间长度，我们要求它的数学期望 ER。由次序统计量的期望公式（$\theta = 1$）：

$$E(X_{(n)}) = \frac{n}{n+1}, \quad E(X_{(1)}) = \frac{1}{n+1}, \quad ER = \frac{n}{n+1} - \frac{1}{n+1} = \frac{n-1}{n+1}.$$

4.2 方差

4.2.1 方差的定义

定义 4.4 (方差)

设 X 是随机变量，若 $E[(X-EX)^2]$ 存在，则称其为 X 的方差，记为 DX（或 $\mathrm{Var}(X)$），即

$$DX = E[(X-EX)^2] = E(X^2) - (EX)^2.$$

称 $\sqrt{DX}$ 为 X 的标准差（或均方差），记为 $\sigma(X)$。

若 $DX > 0$，称 $X^* = \dfrac{X - EX}{\sqrt{DX}}$ 为 X 的标准化随机变量，此时 $EX^* = 0$，$DX^* = 1$。 ♣

笔记 从方差的定义可以看出，只有当随机变量的期望存在时，其方差才可能存在。但期望存在的随机变量，其方差也不一定存在。

练习 4.20 设随机变量 X 的分布函数为

$$F(x) = \begin{cases} 0, & x < -1, \\ 0.2, & -1 \leqslant x < 0, \\ 0.8, & 0 \leqslant x < 1, \\ 1, & x \geqslant 1, \end{cases}$$

则 $D(X^2) =$ _____。

解 由分布函数知 X 的分布列为

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.2 & 0.6 & 0.2 \end{pmatrix}.$$

于是

$$E(X^2) = (-1)^2 \times 0.2 + 0^2 \times 0.6 + 1^2 \times 0.2 = 0.4,$$

$$E(X^4) = (-1)^4 \times 0.2 + 0^4 \times 0.6 + 1^4 \times 0.2 = 0.4,$$

$$D(X^2) = E(X^4) - [E(X^2)]^2 = 0.4 - 0.16 = 0.24.$$

练习 4.21 设连续型随机变量 X 服从参数为 1 的指数分布，则 $D(e^{-2X}) =$ _____。

解

$$E(e^{-2X}) = \int_0^{+\infty} e^{-2x}\, e^{-x}\, \mathrm{d}x = \int_0^{+\infty} e^{-3x}\, \mathrm{d}x = \frac{1}{3},$$

$$E(e^{-4X}) = \int_0^{+\infty} e^{-4x}\, e^{-x}\, \mathrm{d}x = \int_0^{+\infty} e^{-5x}\, \mathrm{d}x = \frac{1}{5},$$

$$D(e^{-2X}) = E(e^{-4X}) - [E(e^{-2X})]^2 = \frac{1}{5} - \frac{1}{9} = \frac{4}{45}.$$

练习 4.22 设一次试验成功的概率为 p，进行 100 次独立重复试验，则成功次数的标准差的最大值为_____。

解 设成功次数为 X，则 $X \sim B(100, p)$，标准差为 $\sqrt{DX} = \sqrt{100p(1-p)}$。由于 $p(1-p) \leqslant \frac{1}{4}$，当且仅当 $p = \frac{1}{2}$ 时等号成立，故标准差的最大值为

$$\sqrt{100 \times \frac{1}{4}} = 5.$$

练习 4.23 设随机变量 X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{2}\cos x, & 0 \leqslant x \leqslant \pi, \\ 0, & \text{其他}, \end{cases}$$

对 X 独立观察 4 次，用 Y 表示观察值大于 $\frac{\pi}{3}$ 的次数，求 $E(Y^2)$。

解

$$P\left(X > \frac{\pi}{3}\right) = \int_{\frac{\pi}{3}}^{\pi} \frac{1}{2}\cos x\, \mathrm{d}x = \frac{1}{2},$$

所以 $Y \sim B\left(4, \frac{1}{2}\right)$，于是

$$EY = 2, \quad DY = 1, \quad E(Y^2) = DY + (EY)^2 = 1 + 4 = 5.$$

4.2.2 方差的性质

性质 方差具有以下性质（假设所涉及的方差均存在）：

1. $DX \geqslant 0$，从而 $E(X^2) = DX + (EX)^2 \geqslant (EX)^2$。
2. $D(C) = 0$，其中 C 为常数。特别地，$D(EX) = 0$。
3. $DX = 0 \Leftrightarrow$ 存在常数 a，使得 $P\{X = a\} = 1$。
4. $D(aX + b) = a^2 DX$，其中 a, b 为常数。
5. 对任意两个随机变量 X, Y，有

$$D(X \pm Y) = DX + DY \pm 2\,\mathrm{Cov}(X, Y).$$

更一般地，

$$D\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 DX_i + 2 \sum_{1 \leqslant i < j \leqslant n} a_i a_j\, \mathrm{Cov}(X_i, X_j).$$

6. 若 X, Y 相互独立，则 $D(X \pm Y) = DX + DY$，且

$$D(XY) = DX \cdot DY + DX \cdot (EY)^2 + DY \cdot (EX)^2 \geqslant DX \cdot DY.$$

一般地，若 $X_1, X_2, \ldots, X_n$ 相互独立，g_i 为连续函数，则

$$D\left(\sum_{i=1}^{n} g_i(X_i)\right) = \sum_{i=1}^{n} D[g_i(X_i)].$$

7. 对任意常数 c，有 $DX = E[(X-EX)^2] \leqslant E[(X-c)^2]$。

证明 [性质 6 中 $D(XY)$ 的证明] 因为 X, Y 相互独立，故 X^2, Y^2 也相互独立，即有

$$E(XY) = EX \cdot EY, \quad E(X^2Y^2) = E(X^2) \cdot E(Y^2),$$

于是

$$\begin{aligned} D(XY) &= E(X^2Y^2) - [E(XY)]^2 = E(X^2)E(Y^2) - (EX)^2(EY)^2 \\ &= [DX + (EX)^2][DY + (EY)^2] - (EX)^2(EY)^2 \\ &= DX \cdot DY + (EX)^2 \cdot DY + (EY)^2 \cdot DX. \end{aligned}$$

又 $(EX)^2 \cdot DY + (EY)^2 \cdot DX \geqslant 0$，所以 $D(XY) \geqslant DX \cdot DY$。

证明 [性质 7 的证明] 令 $g(c) = E[(X-c)^2] = E(X^2) - 2cEX + c^2$。由

$$g'(c) = -2EX + 2c \overset{令}{=} 0,$$

得 $c = EX$。又 $g''(c) = 2 > 0$，故 $c = EX$ 是 $g(c)$ 的最小值点。

命题 4.2 (均值的期望与方差)

[] 若总体 X 的期望为 μ，方差为 σ^2，$X_1, \ldots, X_n$ 是简单随机样本，样本均值为 $\overline{X}$，则：

$$\boldsymbol{E(\overline{X}) = \mu, \quad D(\overline{X}) = \frac{\sigma^2}{n}}$$

证明

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} EX_i = \frac{1}{n} \cdot n\mu = \mu.$$

由于简单随机样本中 $X_1, \ldots, X_n$ 相互独立，故

$$D(\overline{X}) = D\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} DX_i = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}.$$

练习 4.24 对于任意两个随机变量 X 和 Y，若 $E(XY) = EXEY$，则（　）。

A. X, Y 相互独立　B. X, Y 不相互独立　C. $D(XY) = DX \cdot DY$　D. $D(X+Y) = DX + DY$

解 应选 D。$E(XY) = EXEY$ 即 $\mathrm{Cov}(X, Y) = 0$，故 $D(X+Y) = DX + DY$。

练习 4.25 已知 $D(X) = 4$，$D(Y) = 1$，$D(X+Y) = 4$，则 $D(X-Y) =$ _____。

解 应填 6。由 $D(X+Y) = DX + DY + 2\mathrm{Cov}(X, Y) = 5 + 2\mathrm{Cov}(X, Y) = 4$，得 $\mathrm{Cov}(X, Y) = 0$，故

$$D(X-Y) = DX + DY - 2\mathrm{Cov}(X, Y) = 5 + 1 = 6.$$

练习 4.26 设随机变量 X 和 Y 的方差存在且不等于 0，则 $D(X+Y) = D(X) + D(Y)$ 是 X 与 Y（　）。

A. 不相关的充分不必要条件　B. 不相关的充分必要条件

C. 独立的充分不必要条件　D. 独立的充分必要条件

解 应选 B。充分性：$D(X+Y) = DX + DY + 2\mathrm{Cov}(X, Y) = DX + DY \Rightarrow \mathrm{Cov}(X, Y) = 0 \Rightarrow X$ 与 Y 不相关。

必要性：X 与 Y 不相关 $\Rightarrow \mathrm{Cov}(X, Y) = 0 \Rightarrow D(X+Y) = DX + DY$。

练习 4.27 设连续型随机变量 X 的概率密度为

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

则 $EX\cdot\sqrt{DX}=$ _____。

解 应填 $\frac{1}{\sqrt{2}}$。由

$$f(x)=\frac{1}{\sqrt{\pi}}e^{-x^2+2x-1}=\frac{1}{\sqrt{2\pi}\cdot\frac{1}{\sqrt{2}}}\exp\left\{-\frac{(x-1)^2}{2\cdot\left(\frac{1}{\sqrt{2}}\right)^2}\right\},$$

知 $X\sim N\left(1,\frac{1}{2}\right)$，故 $EX=1$，$\sqrt{DX}=\frac{1}{\sqrt{2}}$，于是 $EX\cdot\sqrt{DX}=\frac{1}{\sqrt{2}}$。

练习 4.28 设 X,Y 是随机变量，相关系数为 0.4，$EX=2$，$DX=25$，$EY=1$，$DY=16$，则 $E(2X-3Y+4)^2=$ _____。

解 先计算各分量：

$$EX^2=DX+(EX)^2=29,\quad EY^2=DY+(EY)^2=17,$$

$$EXY=EX\cdot EY+\mathrm{Cov}(X,Y)=2+0.4\times5\times4=10.$$

于是

$$\begin{aligned}E(2X-3Y+4)^2&=E(4X^2+9Y^2-12XY+16+16X-24Y)\\&=4\times29+9\times17-12\times10+16+16\times2-24\times1\\&=116+153-120+16+32-24=173.\end{aligned}$$

练习 4.29 设随机变量 (X,Y) 服从区域 $G=\{(x,y)\mid 0\leqslant x\leqslant 2,\ 0\leqslant y\leqslant 2\}$ 上的均匀分布，$U=|X-Y|$。

(1) 求 U 的概率密度 $f_U(u)$；

(2) 求 EU 和 DU。

解 (1) $F_U(u)=P(|X-Y|\leqslant u)$。当 $u\leqslant 0$ 时，$F_U(u)=0$；当 $u\geqslant 2$ 时，$F_U(u)=1$。当 $0<u<2$ 时，区域 $|x-y|\leqslant u$ 在 $[0,2]^2$ 内的面积为 $4-2\times\frac{1}{2}(2-u)^2=4u-u^2$，故

$$F_U(u)=\frac{4u-u^2}{4}=\frac{(4-u)u}{4}.$$

因此

$$f_U(u)=F_U'(u)=\begin{cases}\dfrac{2-u}{2}, & 0<u<2,\\ 0, & \text{其他}.\end{cases}$$

(2) $EU=\int_0^2 u\cdot\frac{2-u}{2}\,\mathrm{d}u=\frac{2}{3}$；$EU^2=\int_0^2 u^2\cdot\frac{2-u}{2}\,\mathrm{d}u=\frac{2}{3}$；$DU=EU^2-(EU)^2=\frac{2}{3}-\frac{4}{9}=\frac{2}{9}$。

4.3 常见分布的数学期望与方差

笔记 表中仅列出各分布概率密度（或分布律）的非零区域。

证明 [0-1 分布] $P\{X=1\}=p$, $P\{X=0\}=1-p$。

$$EX=1\cdot p+0\cdot(1-p)=p.$$

$$EX^2=1^2\cdot p+0^2\cdot(1-p)=p.$$

表 4.1: 常见分布的数学期望与方差

分布	分布律或概率密度	数学期望	方差
0-1 分布	$P\{X=k\}=p^k(1-p)^{1-k},\ k=0,1$	p	$p(1-p)$
$B(n,p)$	$P\{X=k\}=\mathrm{C}_n^k p^k(1-p)^{n-k},\ k=0,1,\ldots,n$	np	$np(1-p)$
$P(\lambda)$	$P\{X=k\}=\dfrac{\lambda^k e^{-\lambda}}{k!},\ k=0,1,\ldots$	λ	λ
$G(p)$	$P\{X=k\}=(1-p)^{k-1}p,\ k=1,2,\ldots$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
$U(a,b)$	$f(x)=\dfrac{1}{b-a},\ a<x<b$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$E(\lambda)$	$f(x)=\lambda e^{-\lambda x},\ x>0$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
$N(\mu,\sigma^2)$	$f(x)=\dfrac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\dfrac{(x-\mu)^2}{2\sigma^2}\right\}$	μ	σ^2
$\chi^2(n)$	（概率密度不作要求）	n	$2n$

$$DX=EX^2-(EX)^2=p-p^2=p(1-p).$$

证明 $[B(n,p)]$ 设 $X_i\ (i=1,\ldots,n)$ 相互独立，均服从 0-1 分布，则 $X=X_1+X_2+\cdots+X_n\sim B(n,p)$。

$$EX=\sum_{i=1}^{n}EX_i=np,\qquad DX=\sum_{i=1}^{n}DX_i=np(1-p).$$

证明 $[P(\lambda)]$

$$EX=\sum_{k=0}^{\infty}k\cdot\frac{\lambda^k e^{-\lambda}}{k!}=\sum_{k=1}^{\infty}\frac{\lambda^k e^{-\lambda}}{(k-1)!}=\lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda.$$

$$EX^2=\sum_{k=0}^{\infty}k^2\cdot\frac{\lambda^k e^{-\lambda}}{k!}=\sum_{k=1}^{\infty}k\cdot\frac{\lambda^k e^{-\lambda}}{(k-1)!}=\sum_{k=1}^{\infty}[(k-1)+1]\frac{\lambda^k e^{-\lambda}}{(k-1)!}$$

$$=\sum_{k=2}^{\infty}(k-1)\frac{\lambda^k e^{-\lambda}}{(k-1)!}+\sum_{k=1}^{\infty}\frac{\lambda^k e^{-\lambda}}{(k-1)!}=\lambda^2 e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda=\lambda^2+\lambda.$$

$$DX=EX^2-(EX)^2=\lambda^2+\lambda-\lambda^2=\lambda.$$

证明 $[G(p)]$ 令 $q=1-p$，利用幂级数求导技巧：

$$\sum_{k=1}^{\infty}kq^{k-1}=\frac{\mathrm{d}}{\mathrm{d}q}\left(\sum_{k=0}^{\infty}q^k\right)=\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{1}{1-q}\right)=\frac{1}{(1-q)^2}=\frac{1}{p^2},$$

$$\sum_{k=1}^{\infty}k^2q^{k-1}=\frac{\mathrm{d}}{\mathrm{d}q}\left(\sum_{k=1}^{\infty}kq^k\right)=\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{q}{(1-q)^2}\right)=\frac{1+q}{(1-q)^3}=\frac{2-p}{p^3}.$$

$$EX=\sum_{k=1}^{\infty}k\,p\,q^{k-1}=p\cdot\frac{1}{p^2}=\frac{1}{p}.$$

$$EX^2=\sum_{k=1}^{\infty}k^2\,p\,q^{k-1}=p\cdot\frac{2-p}{p^3}=\frac{2-p}{p^2}.$$

$$DX=EX^2-(EX)^2=\frac{2-p}{p^2}-\frac{1}{p^2}=\frac{1-p}{p^2}.$$

证明 $[U(a,b)]$

$$EX=\int_a^b x\cdot\frac{1}{b-a}\,\mathrm{d}x=\frac{1}{b-a}\cdot\frac{b^2-a^2}{2}=\frac{a+b}{2}.$$

$$EX^2=\int_a^b x^2\cdot\frac{1}{b-a}\,\mathrm{d}x=\frac{1}{b-a}\cdot\frac{b^3-a^3}{3}=\frac{a^2+ab+b^2}{3}.$$

$$DX=EX^2-(EX)^2=\frac{a^2+ab+b^2}{3}-\frac{(a+b)^2}{4}=\frac{4(a^2+ab+b^2)-3(a^2+2ab+b^2)}{12}=\frac{(b-a)^2}{12}.$$

证明 $[E(\lambda)]$ 利用分部积分（$\int_0^\infty x^n e^{-\lambda x}\,\mathrm{d}x=n!/\lambda^{n+1}$）：

$$EX=\int_0^\infty x\cdot\lambda e^{-\lambda x}\,\mathrm{d}x=\lambda\cdot\frac{1}{\lambda^2}=\frac{1}{\lambda}.$$

$$EX^2=\int_0^\infty x^2\cdot\lambda e^{-\lambda x}\,\mathrm{d}x=\lambda\cdot\frac{2}{\lambda^3}=\frac{2}{\lambda^2}.$$

$$DX=EX^2-(EX)^2=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}.$$

证明 $[\chi^2(n)]$ 由定义，$\chi^2=X_1^2+X_2^2+\cdots+X_n^2$，其中 $X_i\sim N(0,1)$ 且相互独立。

$$E(\chi^2)=\sum_{i=1}^n E(X_i^2)=n\cdot 1=n.$$

$$D(\chi^2)=\sum_{i=1}^n D(X_i^2).$$

利用 $X_i\sim N(0,1)$：$E(X_i^2)=1$，$E(X_i^4)=3$（由正态分布的四阶矩公式），故

$$D(X_i^2)=E(X_i^4)-[E(X_i^2)]^2=3-1=2.$$

$$D(\chi^2)=2n.$$

4.4 协方差与相关系数

4.4.1 协方差

定义 4.5 (协方差)

设随机变量 X 与 Y 的方差都存在，称

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

为 X 与 Y 的协方差。其中

$$E(XY)=\begin{cases}\sum\limits_i\sum\limits_j x_iy_jP\{X=x_i,Y=y_j\} & \text{（离散型）},\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\,f(x,y)\,\mathrm{d}x\mathrm{d}y & \text{（连续型）}.\end{cases}$$

♣

性质 协方差具有以下性质：

1. $\mathrm{Cov}(X,X)=DX$。
2. $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$（对称性）。
3. $\mathrm{Cov}(X,C)=0$，其中 C 为常数。

4. $\mathrm{Cov}(aX, bY) = ab\,\mathrm{Cov}(X, Y)$。
5. $\mathrm{Cov}(X_1 + X_2, Y) = \mathrm{Cov}(X_1, Y) + \mathrm{Cov}(X_2, Y)$（双线性）。更一般地，
$$\mathrm{Cov}\left(\sum_{i=1}^{n} a_i X_i,\ \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j\,\mathrm{Cov}(X_i, Y_j).$$
6. 若 $\mathrm{Cov}(X, Y) = 0$，则 $D(aX+bY) = a^2DX + b^2DY$。更一般地，$D(aX+bY) = a^2DX + b^2DY + 2ab\mathrm{Cov}(X, Y)$

笔记 协方差 $\mathrm{Cov}(X, Y)$ 是随机变量 X 与 Y 之间偏差关联程度的度量，研究的是 X 与 Y 之间的线性相依性。

练习 4.30 二维随机变量 $(X, Y) \sim N(1, 1; 4, 9; -0.5)$，设 $Z = \frac{1}{2}X - \frac{1}{3}Y$，则 $DZ =$ _____。

解 应填 3。已知 $DX = 4$，$DY = 9$，$\rho = -0.5$，则
$$\mathrm{Cov}(X, Y) = \rho\sqrt{DX}\sqrt{DY} = -0.5 \times 2 \times 3 = -3.$$
$$DZ = \frac{1}{4}DX + \frac{1}{9}DY + 2 \times \frac{1}{2} \times \left(-\frac{1}{3}\right)\mathrm{Cov}(X, Y) = 1 + 1 + 1 = 3.$$

练习 4.31 设 $X_1, X_2, \ldots, X_5$ 为总体 $N(1, 4)$ 的简单随机样本，$X_{(5)} = \max\{X_1, X_2, \ldots, X_5\}$。求：(1) $P(X_{(5)} < 3)$；(2) $\mathrm{Cov}(X_1, \bar{X})$，其中 $\bar{X} = \frac{1}{5}\sum_{i=1}^{5} X_i$。

解 (1) $P(X_i < 3) = \Phi\left(\frac{3-1}{2}\right) = \Phi(1)$，由样本独立性
$$P(X_{(5)} < 3) = [P(X_1 < 3)]^5 = [\Phi(1)]^5.$$
(2) $\mathrm{Cov}(X_1, \bar{X}) = \mathrm{Cov}\left(X_1,\ \frac{1}{5}\sum_{i=1}^{5} X_i\right) = \frac{1}{5}\mathrm{Cov}(X_1, X_1) = \frac{DX_1}{5} = \frac{4}{5}$。

练习 4.32 设二维随机变量 (X, Y) 的概率密度函数为
$$f(x, y) = \begin{cases} xe^{-y}, & 0 < x < y, \\ 0, & \text{其他}, \end{cases}$$
求：(1) $Z = X + Y$ 的概率密度；(2) $N = \min(X, Y)$ 的概率密度；(3) EZ 和 EN。

解 (1) 由卷积公式，约束 $0 < x < y$ 且 $x + y = z$，即 $0 < x < z - x$，故 $0 < x < z/2$（$z > 0$）：
$$\begin{aligned} f_Z(z) &= \int_0^{z/2} x\,e^{-(z-x)}\,\mathrm{d}x = e^{-z}\int_0^{z/2} x\,e^{x}\,\mathrm{d}x \\ &= e^{-z}\left[e^{z/2}(z/2 - 1) + 1\right] = e^{-z/2}\left(\frac{z}{2} - 1\right) + e^{-z} \quad (z > 0). \end{aligned}$$
(2) 在定义域内始终有 $0 < X < Y$，故 $\min(X, Y) \equiv X$。
$$f_X(x) = \int_x^{+\infty} x\,e^{-y}\,\mathrm{d}y = x\,e^{-x} \quad (x > 0)$$
所以 $f_N(n) = n\,e^{-n}$ $(n > 0)$

(3) $EN = EX = \int_0^{+\infty} x^2\,e^{-x}\,\mathrm{d}x = \Gamma(3) = 2$。

$f_Y(y) = \int_0^{y} x\,e^{-y}\,\mathrm{d}x = \frac{y^2}{2}e^{-y}$ $(y > 0)$，$EY = \frac{1}{2}\Gamma(4) = 3$。

$EZ = EX + EY = 2 + 3 = 5$。

练习 4.33 设二维随机变量 (X,Y) 的概率分布为

$X\backslash Y$	0	1	2
-1	0.1	0.1	b
1	a	0.1	0.1

若事件 $\{\max(X,Y)=2\}$ 与事件 $\{\min(X,Y)=1\}$ 相互独立，则 $\mathrm{Cov}(X,Y)=$ _____。

解 首先确定未知参数：

$$P\{\max(X,Y)=2\}=P\{Y=2\}=b+0.1,$$
$$P\{\min(X,Y)=1\}=P\{X=1,Y=1\}+P\{X=1,Y=2\}=0.2.$$

由独立性：$P\{X=1,Y=2\}=P\{\max=2\}\,P\{\min=1\}$，即 $0.1=(b+0.1)\times 0.2$，解得 $b=0.4$。

由分布律之和为 1：$0.1+0.1+0.4+a+0.1+0.1=1$，解得 $a=0.2$。

计算协方差：

$$EX=-1\times 0.6+1\times 0.4=-0.2,\quad EY=0\times 0.3+1\times 0.2+2\times 0.5=1.2,$$
$$E(XY)=(-1)(2)(0.4)+1(2)(0.1)=-0.6,$$
$$\mathrm{Cov}(X,Y)=-0.6-(-0.2)(1.2)=-0.36.$$

练习 4.34 设随机变量

$$X\sim\begin{pmatrix}0 & 1\\ \frac{1}{4} & \frac{3}{4}\end{pmatrix},\quad Y\sim\begin{pmatrix}0 & 1\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix},$$

且 $\mathrm{Cov}(X,Y)=\dfrac{1}{8}$，求 X 与 Y 的联合分布律。

解 $EX=\dfrac{3}{4}$，$EY=\dfrac{1}{2}$，由

$$\mathrm{Cov}(X,Y)=E(XY)-EX\cdot EY=E(XY)-\frac{3}{8}=\frac{1}{8},$$

得 $E(XY)=\dfrac{1}{2}$。而 $E(XY)=P\{X=1,Y=1\}$，故 $P\{X=1,Y=1\}=\dfrac{1}{2}$。由此：

$$P\{X=1,Y=0\}=P\{X=1\}-P\{X=1,Y=1\}=\frac{1}{4},$$
$$P\{X=0,Y=1\}=P\{Y=1\}-P\{X=1,Y=1\}=0,$$
$$P\{X=0,Y=0\}=P\{X=0\}=\frac{1}{4}.$$

联合分布律为

$X\backslash Y$	0	1	P_i
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
P_j	$\frac{1}{2}$	$\frac{1}{2}$	1

4.4.2 相关系数

定义 4.6 (相关系数)

设 $DX > 0$，$DY > 0$，称

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

为 X 与 Y 的相关系数。

若 $\rho_{XY} = 0$，则称 X 与 Y 不相关；若 $\rho_{XY} > 0$，称 X 与 Y 正相关；若 $\rho_{XY} < 0$，称 X 与 Y 负相关。 ♣

性质 相关系数具有以下性质：

1. $|\rho_{XY}| \leqslant 1$。
2. $|\rho_{XY}| = 1 \Leftrightarrow$ 存在常数 $a \neq 0, b$，使得 $P\{Y = aX + b\} = 1$。具体地：
   - $\rho_{XY} = 1 \Leftrightarrow P\{Y = aX + b\} = 1\ (a > 0)$；
   - $\rho_{XY} = -1 \Leftrightarrow P\{Y = aX + b\} = 1\ (a < 0)$。
3. $\rho_{XY} = 0$（即不相关）的等价条件：
   - $\mathrm{Cov}(X,Y) = 0$；
   - $E(XY) = EX \cdot EY$；
   - $D(X+Y) = DX + DY$。
4. 若 $(X,Y) \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$，则 X 与 Y 独立等价于 X 与 Y 不相关。

笔记 相关系数 ρ_{XY} 是刻画 X 与 Y 之间线性相关程度的度量。$\rho_{XY} = 0$ 只表示不存在线性关系，但它们之间仍可能存在非线性关系。因此，不相关不等于独立。[1]当 X 与 Y 相互独立时，可推出 X 与 Y 不相关；但当 X 与 Y 不相关时，X 与 Y 却不一定相互独立。

性质 独立性与不相关性的关系：

1. X 与 Y 相互独立 $\Rightarrow$ X 与 Y 不相关，但反之不一定成立。
2. 在以下特殊情形中，独立性与不相关性等价：
   - (X,Y) 服从二维正态分布；
   - X 与 Y 均服从两点分布。

练习 4.35 设二维随机变量 (X,Y) 的概率密度函数为

$$f(x,y) = \begin{cases} 24y(1-x), & 0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant x, \\ 0, & \text{其他}, \end{cases}$$

求：(1) X 的边缘密度 $f_X(x)$；(2) $P(X+Y<1)$；(3) $\mathrm{Cov}(X,Y)$。

解 (1) $f_X(x) = \displaystyle\int_0^x 24y(1-x)\,\mathrm{d}y = 12x^2(1-x) \quad (0 \leqslant x \leqslant 1)$。

(2) 积分区域为 $\{0 \leqslant y \leqslant x \leqslant 1,\ x+y<1\}$。以 $x = 1/2$ 为分界：

- 当 $0 \leqslant x \leqslant \dfrac{1}{2}$ 时，y 从 0 到 x（此时 $x+y \leqslant 2x \leqslant 1$ 自动满足）；
- 当 $\dfrac{1}{2} < x \leqslant 1$ 时，y 从 0 到 $1-x$。

$$P = \int_0^{1/2} 12x^2(1-x)\,\mathrm{d}x + \int_{1/2}^1 12(1-x)^3\,\mathrm{d}x = \frac{5}{16} + \frac{3}{16} = \frac{1}{2}.$$

[1]不相关：没有直线关系。独立：什么关系都没有。

(3)

$$EX=\int_0^1 12x^3(1-x)\,\mathrm{d}x=\frac{3}{5},\quad EY=\int_0^1\int_0^x 24y^2(1-x)\,\mathrm{d}y\,\mathrm{d}x=\frac{2}{5},$$

$$EXY=\int_0^1\int_0^x 24xy^2(1-x)\,\mathrm{d}y\,\mathrm{d}x=\frac{4}{15},$$

$$\mathrm{Cov}(X,Y)=\frac{4}{15}-\frac{3}{5}\times\frac{2}{5}=\frac{2}{75}.$$

✍ **练习 4.36** 将长度为 1 m 的木棒随机地截成两段，则两段长度的相关系数为（　）。

解 设一段长为 X，另一段为 Y，则 $X+Y=1$，即 $Y=-X+1$（$a=-1<0$），故 $\rho_{XY}=-1$。

✍ **练习 4.37** 掷硬币 2 次，正面出现的次数记为 X，反面出现的次数记为 Y，则 X 与 Y 的相关系数 $\rho_{XY}=$ _____。

解 应填 -1。因为总次数固定为 2，故 $X+Y=2$，即 $Y=-X+2$（斜率 $a=-1<0$），所以 $\rho_{XY}=-1$。

✍ **练习 4.38** 如果存在常数 $a,b\,(a\neq 0)$，使 $P(Y=aX+b)=1$，且 $0<DX<+\infty$，则 $\rho_{XY}=$ _____。

解 若存在 $Y=aX+b$ 的线性关系，则

$$\rho_{XY}=\begin{cases}1, & a>0,\\ -1, & a<0.\end{cases}$$

✍ **练习 4.39** 设 A,B 为两个随机事件，且 $P(A)=\frac{1}{4}$，$P(B|A)=\frac{1}{3}$，$P(A|B)=\frac{1}{2}$。令

$$X=\begin{cases}1, & A\text{ 发生},\\ 0, & A\text{ 不发生},\end{cases}\qquad Y=\begin{cases}1, & B\text{ 发生},\\ 0, & B\text{ 不发生}.\end{cases}$$

求：(1) (X,Y) 的概率分布；(2) ρ_{XY}；(3) $Z=X^2+Y^2$ 的概率分布。

解 (1) 先计算基本概率：

$$P(AB)=P(A)\,P(B|A)=\frac{1}{4}\times\frac{1}{3}=\frac{1}{12},\quad P(B)=\frac{P(AB)}{P(A|B)}=\frac{1/12}{1/2}=\frac{1}{6}.$$

于是

$$\begin{aligned}
P\{X=0,Y=0\}&=P(\overline{A}\,\overline{B})=1-P(A)-P(B)+P(AB)=1-\frac{1}{4}-\frac{1}{6}+\frac{1}{12}=\frac{2}{3},\\
P\{X=0,Y=1\}&=P(\overline{A}\,B)=P(B)-P(AB)=\frac{1}{6}-\frac{1}{12}=\frac{1}{12},\\
P\{X=1,Y=0\}&=P(A\,\overline{B})=P(A)-P(AB)=\frac{1}{4}-\frac{1}{12}=\frac{1}{6},\\
P\{X=1,Y=1\}&=P(AB)=\frac{1}{12}.
\end{aligned}$$

(2) 边缘分布：$X\sim(0:\frac{3}{4},\ 1:\frac{1}{4})$，$Y\sim(0:\frac{5}{6},\ 1:\frac{1}{6})$。

$$DX=\frac{3}{16},\quad DY=\frac{5}{36},\quad E(XY)=\frac{1}{12},$$

$$\mathrm{Cov}(X,Y)=\frac{1}{12}-\frac{1}{4}\times\frac{1}{6}=\frac{1}{24},\quad \rho_{XY}=\frac{1/24}{\sqrt{3/16}\sqrt{5/36}}=\frac{1}{\sqrt{15}}=\frac{\sqrt{15}}{15}.$$

(3) Z 的可能取值为 $0,1,2$：

$$P\{Z=0\}=\frac{2}{3},\quad P\{Z=1\}=\frac{1}{12}+\frac{1}{6}=\frac{1}{4},\quad P\{Z=2\}=\frac{1}{12}.$$

练习 4.40 设随机变量 X 与 Y 的概率分布分别为

X	0	1	2
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

,

Y	-1	0	1
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

,

且 $P\{X^2=Y^2\}=1$。

(1) 求 (X,Y) 的概率分布；

(2) 求 $Z=XY$ 的概率分布；

(3) 求 ρ_{XY}。

解 (1) 由 $P\{X^2=Y^2\}=1$ 知

$$P\{X=0,Y=-1\}=P\{X=0,Y=1\}=P\{X=1,Y=0\}=0,$$

于是

$$P\{X=0,Y=0\}=\frac{1}{3},\quad P\{X=1,Y=1\}=\frac{1}{3},\quad P\{X=1,Y=-1\}=\frac{1}{3}.$$

(X,Y) 的联合分布为

$X\backslash Y$	-1	0	1	P_i
0	0	$\frac{1}{3}$	0	$\frac{1}{3}$
1	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
P_j	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

(2) $Z=XY$ 的可能取值为 $-1,0,1$：

$$P\{Z=-1\}=P\{X=1,Y=-1\}=\frac{1}{3},$$
$$P\{Z=0\}=P\{X=0,Y=0\}=\frac{1}{3},$$
$$P\{Z=1\}=P\{X=1,Y=1\}=\frac{1}{3}.$$

(3) $EX=\dfrac{2}{3}$，$DX=\dfrac{2}{9}$；$EY=0$，$DY=\dfrac{2}{3}$；$E(XY)=0$。故 $\mathrm{Cov}(X,Y)=0$，$\rho_{XY}=0$。此例说明 X 与 Y 不相关，但由联合分布可知 X 与 Y 并不独立。

练习 4.41 设离散型随机变量 X_1,X_2 都只取 1 和 -1，已知 $P(X_1=-1)=P(X_2=-1)=\dfrac{1}{2}$，且 $P(X_2=-1\mid X_1=-1)=\dfrac{1}{3}$。

(1) 求 (X_1,X_2) 的联合概率函数；

(2) 求 $P(X_1+X_2=0)$；

(3) 分别求 $\mathrm{Cov}(X_1,X_2)$ 和 $\rho_{X_1X_2}$。

解 (1) $P(X_2=-1,X_1=-1)=P(X_2=-1\mid X_1=-1)\,P(X_1=-1)=\dfrac{1}{3}\times\dfrac{1}{2}=\dfrac{1}{6}$，

$P(X_2=1,X_1=-1)=P(X_1=-1)-P(X_2=-1,X_1=-1)=\dfrac{1}{3}$。

由对称性：$P(X_2=1,X_1=1)=\dfrac{1}{6}$，$P(X_2=-1,X_1=1)=\dfrac{1}{3}$。

$X_1\backslash X_2$	-1	1
-1	$\frac{1}{6}$	$\frac{1}{3}$
1	$\frac{1}{3}$	$\frac{1}{6}$

(2) $P(X_1+X_2=0)=\dfrac{1}{3}+\dfrac{1}{3}=\dfrac{2}{3}$。

(3) $EX_1=0$，$EX_2=0$，$DX_1=DX_2=1$。$E(X_1X_2)=1\times\left(\frac{1}{6}+\frac{1}{6}\right)+(-1)\times\left(\frac{1}{3}+\frac{1}{3}\right)=-\frac{1}{3}$。

$\mathrm{Cov}(X_1,X_2)=-\frac{1}{3}$，$\rho_{X_1X_2}=-\frac{1}{3}$。

练习 4.42 设随机变量 X 的概率密度为 $f(x)$，$DX=2$，而随机变量 Y 的概率密度为 $f(-y)$，它们的相关系数 $\rho_{XY}=\frac{1}{4}$，记 $Z=X+Y$，求 EZ，DZ。

解 $EY=\int_{-\infty}^{+\infty}y\,f(-y)\,\mathrm{d}y\overset{t=-y}{=}-\int_{+\infty}^{-\infty}(-t)\,f(t)\,\mathrm{d}t=-EX$，故 $EZ=EX+EY=0$。

$EY^2=\int_{-\infty}^{+\infty}y^2\,f(-y)\,\mathrm{d}y=\int_{-\infty}^{+\infty}x^2\,f(x)\,\mathrm{d}x=EX^2$，故 $DY=EX^2-(EY)^2=DX=2$。

$$DZ=DX+DY+2\,\mathrm{Cov}(X,Y)=2+2+2\times\frac{1}{4}\times\sqrt{2}\times\sqrt{2}=6.$$

练习 4.43 随机变量 X,Y 相互独立，均服从参数为 λ 的指数分布。

(1) 求 $Z=X+Y$ 的概率密度函数；

(2) 求 EZ，DZ；

(3) 求 ρ_{XZ}。

解 (1) $F_Z(z)=P(X+Y\leqslant z)=\int_0^z\mathrm{d}x\int_0^{z-x}\lambda^2e^{-\lambda x}e^{-\lambda y}\,\mathrm{d}y=1-(1+\lambda z)e^{-\lambda z}$，故

$$f_Z(z)=\begin{cases}\lambda^2z\,e^{-\lambda z}, & z>0,\\ 0, & z\leqslant 0.\end{cases}$$

(2) $EZ=EX+EY=\frac{2}{\lambda}$，$DZ=DX+DY=\frac{2}{\lambda^2}$。

(3) $\mathrm{Cov}(X,Z)=\mathrm{Cov}(X,X+Y)=DX+\mathrm{Cov}(X,Y)=DX=\frac{1}{\lambda^2}$，故

$$\rho_{XZ}=\frac{\mathrm{Cov}(X,Z)}{\sqrt{DX}\sqrt{DZ}}=\frac{1/\lambda^2}{\sqrt{1/\lambda^2}\sqrt{2/\lambda^2}}=\frac{1}{\sqrt{2}}.$$

4.4.3 不相关性的判定

练习 4.44 设随机变量 X 在 $[-1,1]$ 上服从均匀分布，$Y=X^3$，则 X 与 Y（　）。

A. 不相关且相互独立　B. 不相关且不相互独立　C. 相关且相互独立　D. 相关且不相互独立

解 应选 D。X 与 Y 之间存在确定性函数关系 $Y=X^3$，故不可能独立，排除 A、C。计算协方差：$EX=0$，

$$E(XY)=E(X^4)=\int_{-1}^{1}x^4\cdot\frac{1}{2}\,\mathrm{d}x=\frac{1}{5}\neq 0,$$

故 $\mathrm{Cov}(X,Y)=E(XY)-EX\cdot EY=\frac{1}{5}\neq 0$，$X$ 与 Y 相关。

笔记 本题常见误区：有人误将 $E(XY)$ 写成 $\int x^3\cdot(1/2)\,\mathrm{d}x=0$（奇函数），但这计算的是 $E(X^3)$ 而非 $E(XY)=E(X^4)$。注意 $XY=X\cdot X^3=X^4$，是偶函数。

练习 4.45 设 $X\sim N(0,\sigma^2)\ (\sigma>0)$，$Y=X^3$。

(1) 求 EY；

(2) 判断 X 与 Y 是否相关？是否独立？说明理由。

解 (1) X 的密度 $f(x)$ 为偶函数，x^3 为奇函数，故

$$EY=E(X^3)=\int_{-\infty}^{+\infty}x^3f(x)\,\mathrm{d}x=0.$$

(2) $E(XY)=E(X^4)$。对 $X\sim N(0,\sigma^2)$，利用 Gamma 函数计算：

$$E(X^4)=\int_{-\infty}^{+\infty}x^4\cdot\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x=\frac{2}{\sqrt{2\pi}\,\sigma}\int_0^{+\infty}x^4\,e^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x=3\sigma^4\neq 0.$$

故 $\mathrm{Cov}(X,Y)=E(XY)-EX\cdot EY=3\sigma^4\neq 0$，$X$ 与 Y 相关，从而不独立。

练习 4.46 设随机变量 X 的密度函数 $f(x)$ 是偶函数，令 $Y=X^2$，假定 EX、EY 都存在，求证 $\mathrm{Cov}(X,Y)=0$。

解 由于 $f(x)$ 为偶函数，x^3 为奇函数，故

$$\int_{-\infty}^{+\infty}x^3f(x)\,\mathrm{d}x=0.$$

于是

$$\mathrm{Cov}(X,Y)=E(XY)-EX\cdot EY=E(X^3)-EX\cdot E(X^2)=\int_{-\infty}^{+\infty}x^3f(x)\,\mathrm{d}x-EX\cdot E(X^2)=0.$$

练习 4.47 设 θ 服从 $[-\pi,\pi]$ 上的均匀分布，又 $X=\sin\theta$，$Y=\cos\theta$，则 $\rho_{XY}=$ _____。

解 应填 0。

$$EX=\int_{-\pi}^{\pi}\frac{\sin\theta}{2\pi}\,\mathrm{d}\theta=0,\quad EY=\int_{-\pi}^{\pi}\frac{\cos\theta}{2\pi}\,\mathrm{d}\theta=0,$$

故

$$\rho_{XY}=\frac{EXY-EX\cdot EY}{\sqrt{DX}\sqrt{DY}}=\frac{\frac{1}{2\pi}\int_{-\pi}^{\pi}\sin\theta\cos\theta\,\mathrm{d}\theta}{\sqrt{DX}\sqrt{DY}}=0.$$

笔记 此例中 $X^2+Y^2=\sin^2\theta+\cos^2\theta=1$，说明 X 与 Y 之间存在非线性关系（不独立），但 $\rho_{XY}=0$（不相关）。这是”不相关不等于独立”的典型例子。

练习 4.48 已知随机变量 X,Y 独立同分布，记 $U=X-Y$，$V=X+Y$，则 U,V（　）。

A. 不独立　B. 独立　C. 相关系数不为 0　D. 相关系数为 0

解 应选 D。$EU=EX-EY=0$，$EV=EX+EY=2EX$。

$E(UV)=E(X^2-Y^2)=EX^2-EY^2=0$（因为 X,Y 同分布），故

$$\rho_{UV}=\frac{E(UV)-EU\cdot EV}{\sqrt{DU}\sqrt{DV}}=0.$$

由于未给出具体分布，无法判断是否独立，故只能选 D。

笔记 若 X,Y 同为正态分布，则 U,V 服从联合正态分布，不相关即独立，此时应选 B。但本题未说明分布类型，故只能选 D。

练习 4.49 在区间 $[0.5,1]$ 中任取一个值 x，再在区间 $[-x,x]$ 中任取一个值 y，构成二维随机变量 (X,Y)。

(1) 求 (X,Y) 的联合密度函数 $f(x,y)$，关于 Y 的边缘密度函数 $f_Y(y)$，并判断 X,Y 是否独立；

(2) 求 $\mathrm{Cov}(X,Y)$，并据此判断 X,Y 是否相关。

解 (1) (X,Y) 的取值区域为 $0.5<x<1$，$-x<y<x$。

$X\sim U(0.5,1)$，故 $f_X(x)=\dfrac{1}{0.5}=2\ (0.5<x<1)$。给定 $X=x$，$Y\sim U(-x,x)$，故 $f(y|x)=\dfrac{1}{2x}\ (-x<y<x)$。

联合密度函数：

$$f(x,y)=f_X(x)\,f(y|x)=2\times\frac{1}{2x}=\frac{1}{x},\quad 0.5<x<1,\ -x<y<x.$$

关于 Y 的边缘密度函数。对固定的 y，x 的范围由 $\max(0.5,|y|)<x<1$ 确定：

$$f_Y(y)=\begin{cases}\displaystyle\int_y^1\frac{1}{x}\,\mathrm{d}x=-\ln y, & 0.5\leqslant y<1,\\ \displaystyle\int_{0.5}^1\frac{1}{x}\,\mathrm{d}x=\ln 2, & -0.5\leqslant y<0.5,\\ \displaystyle\int_{-y}^1\frac{1}{x}\,\mathrm{d}x=-\ln(-y), & -1\leqslant y<-0.5.\end{cases}$$

$f_X(x)=\displaystyle\int_{-x}^x\frac{1}{x}\,\mathrm{d}y=2$。由于 $f(x,y)=\dfrac{1}{x}\neq 2\,f_Y(y)=f_X(x)\,f_Y(y)$，故 X 与 Y 不独立。

(2) $EX=\displaystyle\int_{0.5}^1 2x\,\mathrm{d}x=\frac{3}{4}$。由对称性（$f_Y(y)$ 关于 $y=0$ 对称），$EY=0$。

$$\mathrm{Cov}(X,Y)=EXY-EX\cdot EY=EXY=\int_{0.5}^1\int_{-x}^x xy\cdot\frac{1}{x}\,\mathrm{d}y\,\mathrm{d}x=\int_{0.5}^1 x\left[\int_{-x}^x y\,\mathrm{d}y\right]\mathrm{d}x=0.$$

所以 X,Y 不相关（但由 (1) 知二者不独立）。

4.4.4 正态分布的数字特征

练习 4.50 设随机变量 (X,Y) 服从二维正态分布，X 与 Y 的期望均为 μ，方差均为 σ^2，相关系数 $\rho_{XY}=0$。记 $Z_1=2X+Y$，$Z_2=2X-Y$，则 $\rho_{Z_1Z_2}=$ _____。

解 由 $\rho_{XY}=0$ 知 X 与 Y 相互独立。

$$DZ_1=4DX+DY=5\sigma^2,\quad DZ_2=4DX+DY=5\sigma^2,$$

$$\mathrm{Cov}(Z_1,Z_2)=\mathrm{Cov}(2X+Y,\,2X-Y)=4DX-DY=3\sigma^2,$$

$$\rho_{Z_1Z_2}=\frac{3\sigma^2}{\sqrt{5\sigma^2}\sqrt{5\sigma^2}}=\frac{3}{5}.$$

练习 4.51 设两个随机变量 X,Y 相互独立，且都服从均值为 0、方差为 0.5 的正态分布，求 $|X-Y|$ 的数学期望与方差。

解 由条件知 $X-Y\sim N(0,1)$，故

$$E(|X-Y|)=\int_{-\infty}^{+\infty}|x|\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}\,\mathrm{d}x=\frac{2}{\sqrt{2\pi}}\int_0^{+\infty}x\,e^{-\frac{x^2}{2}}\,\mathrm{d}x=\sqrt{\frac{2}{\pi}}.$$

又 $E(|X-Y|^2)=D(X-Y)+[E(X-Y)]^2=1$，故

$$D(|X-Y|)=1-\frac{2}{\pi}.$$

练习 4.52 已知随机变量 $X\sim N(0,1)$，$Y\sim N(1,4)$，且 X 与 Y 的相关系数为 $\dfrac{1}{3}$，设 $Z=2X+3Y$，求 ρ_{XZ}。

解 由 $\rho_{XY}=\dfrac{\mathrm{Cov}(X,Y)}{1\times 2}=\dfrac{1}{3}$，得 $\mathrm{Cov}(X,Y)=\dfrac{2}{3}$。

$$\mathrm{Cov}(X,Z)=\mathrm{Cov}(X,2X+3Y)=2DX+3\,\mathrm{Cov}(X,Y)=2+3\times\frac{2}{3}=4,$$

$$DZ=4DX+9DY+12\,\mathrm{Cov}(X,Y)=4+36+12\times\frac{2}{3}=48,$$

故

$$\rho_{XZ}=\frac{\mathrm{Cov}(X,Z)}{\sqrt{DX}\sqrt{DZ}}=\frac{4}{\sqrt{48}}=\frac{4}{4\sqrt{3}}=\frac{1}{\sqrt{3}}.$$

✍ **练习 4.53** 已知随机变量 $X \sim N(1, 3^2)$，$Y \sim N(0, 4^2)$，且 $\rho_{XY} = \dfrac{1}{2}$。设 $Z = \dfrac{X}{3} + \dfrac{Y}{2}$，求：(1) EZ 和 DZ；(2) ρ_{XZ}。

解 (1) $\mathrm{Cov}(X, Y) = \dfrac{1}{2} \times 3 \times 4 = 6$。

$$EZ = \frac{1}{3} \times 1 + \frac{1}{2} \times 0 = \frac{1}{3}, \quad DZ = \frac{1}{9} \times 9 + \frac{1}{4} \times 16 + 2 \times \frac{1}{3} \times \frac{1}{2} \times 6 = 1 + 4 + 2 = 7.$$

(2) $\mathrm{Cov}(X, Z) = \dfrac{1}{3} DX + \dfrac{1}{2} \mathrm{Cov}(X, Y) = 3 + 3 = 6$，故

$$\rho_{XZ} = \frac{\mathrm{Cov}(X, Z)}{\sqrt{DX}\sqrt{DZ}} = \frac{6}{\sqrt{9 \times 7}} = \frac{2\sqrt{7}}{7}.$$

✍ **练习 4.54** 设二维随机变量 $(X, Y) \sim N(1, 3^2; 0, 4^2, 0)$，则 $E(XY) =$ _____。

解 应填 0。由 $\rho = 0$ 得 $\mathrm{Cov}(X, Y) = 0$，由二维正态分布的性质知 X 与 Y 相互独立，故

$$E(XY) = EX \cdot EY = 1 \times 0 = 0.$$

第 5 章 大数定律和中心极限定理

$$\begin{cases} \text{切比雪夫不等式：} P\{|X-EX| \geqslant \varepsilon\} \leq \dfrac{DX}{\varepsilon^2} \\ \text{大数定理} \begin{cases} \text{切比雪夫大数定律} \\ \text{独立同分布的大数定理} \\ \text{辛钦大数定理} \end{cases} \\ \text{中心极限定理} \begin{cases} \text{列维林德伯格中心极限定理} \\ \text{拉普拉斯中心极限定理} \end{cases} \end{cases}$$

内容提要

- 依概率收敛的定义与性质 5.1
- 切比雪夫不等式的应用 5.1
- 切比雪夫大数定律 5.2
- 伯努利大数定律 5.3
- 辛钦大数定律 5.4
- 列维–林德伯格中心极限定理 5.5
- 棣莫弗–拉普拉斯中心极限定理 5.6

5.1 依概率收敛

大数定律和中心极限定理研究的是大量随机变量之和在极限状态下的行为规律，而”依概率收敛”是描述这一极限行为的基本工具。

定义 5.1 (依概率收敛)

设 X 为随机变量，$\{X_n\}$ $(n=1,2,\dots)$ 为随机变量序列。若对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty} P\{|X_n-X| \geqslant \varepsilon\}=0 \quad \text{或等价地} \quad \lim_{n\to\infty} P\{|X_n-X|<\varepsilon\}=1,$$

则称序列 $\{X_n\}$ 依概率收敛于 X，记为

$$X_n \xrightarrow{P} X \ (n\to\infty).$$

特别地，若 $X=a$ 为常数，则称 $\{X_n\}$ 依概率收敛于常数 a。 ♣

性质 依概率收敛具有以下运算性质。设 $X_n \xrightarrow{P} a$，$Y_n \xrightarrow{P} b$，$g(x,y)$ 为连续函数，则

$$g(X_n,Y_n) \xrightarrow{P} g(a,b).$$

更一般地，对 m 元连续函数 $g(x_1,x_2,\dots,x_m)$，若 $X_n^{(i)} \xrightarrow{P} a_i$ $(i=1,2,\dots,m)$，则

$$g\left(X_n^{(1)},X_n^{(2)},\dots,X_n^{(m)}\right) \xrightarrow{P} g(a_1,a_2,\dots,a_m).$$

笔记 依概率收敛是最弱的收敛方式之一（弱于几乎必然收敛和依分布收敛，但强于依概率为 1 收敛的概念框架）。在本章中，大数定律所描述的正是随机变量序列的算术平均值依概率收敛于其期望值的平均值这一事实。

5.2 切比雪夫不等式

定理 5.1 (切比雪夫不等式)

设随机变量 X 的数学期望 EX 和方差 DX 都存在，则对任意 $\varepsilon > 0$，有

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

或等价地，

$$P\{|X - EX| < \varepsilon\} \geqslant 1 - \frac{DX}{\varepsilon^2}.$$

笔记 切比雪夫不等式说明了方差越小，X 的取值越集中在 EX 附近。它在理论上具有重要价值，是大数定律的理论基础，但在具体估计概率时往往比较粗糙。其中 $\geqslant \varepsilon \leqslant$ 刚好是一个颜文字，方便读者进行记忆。

练习 5.1 设 X 为随机变量，$P(|X - EX| < \varepsilon) \geqslant 0.9$，$DX = 0.009$，用切比雪夫不等式估计 ε 的取值范围。

解 由切比雪夫不等式：

$$P(|X - EX| < \varepsilon) \geqslant 1 - \frac{DX}{\varepsilon^2} = 1 - \frac{0.009}{\varepsilon^2} \geqslant 0.9,$$

即 $\dfrac{0.009}{\varepsilon^2} \leqslant 0.1$，解得 $\varepsilon \geqslant 0.3$。

5.3 大数定律

大数定律是概率论中最重要的极限定理之一，它从理论上严格证明了”频率稳定于概率”这一经验事实。切比雪夫不等式（定理 5.1）是大数定律的理论基础——正是利用这一不等式对概率进行粗略估计，才能证明各种形式的大数定律。

在应用切比雪夫不等式时，需要特别注意不等号的方向：原式为

$$P\{|X - EX| \geq \varepsilon\} \leq \frac{DX}{\varepsilon^2},$$

其等价形式为

$$P\{|X - EX| < \varepsilon\} \geq 1 - \frac{DX}{\varepsilon^2}.$$

公式中的随机变量 X 可以替换为任何满足方差存在条件的随机变量。

练习 5.2 设随机变量 $X \sim U(-1, 2)$，由切比雪夫不等式得 $P(|X - 0.5| < 1) >$ _____。

解 应填 $\dfrac{1}{4}$。$EX = \dfrac{-1+2}{2} = 0.5$，$DX = \dfrac{9}{12} = \dfrac{3}{4}$。由切比雪夫不等式（$\varepsilon = 1$）：

$$P(|X - 0.5| < 1) \geq 1 - \frac{DX}{\varepsilon^2} = 1 - \frac{3}{4} = \frac{1}{4}.$$

练习 5.3 设随机变量 X 的方差 $DX = 2$，则由切比雪夫不等式，有 $P(|X - EX| \geq 2) \leq$ _____。

解 直接代入切比雪夫不等式（$\varepsilon = 2$）：

$$P(|X - EX| \geq 2) \leq \frac{DX}{\varepsilon^2} = \frac{2}{4} = \frac{1}{2}.$$

练习 5.4 设 $EX = -2$，$EY = 2$，$DX = 1$，$DY = 4$，$\rho(X, Y) = 0.5$。求 $D(X + Y) =$ _____，并根据切比雪夫不等式估计 $P\{|X + Y| \geq 6\}$ _____。

解 $\mathrm{Cov}(X,Y)=0.5\times\sqrt{1\times 4}=1$。

$$D(X+Y)=DX+DY+2\,\mathrm{Cov}(X,Y)=1+4+2=7.$$

$E(X+Y)=EX+EY=0$，由切比雪夫不等式（$\varepsilon=6$）：

$$P\{|X+Y|\geq 6\}\leq\frac{D(X+Y)}{36}=\frac{7}{36}.$$

练习 5.5 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{2}x^2e^{-x}, & x>0,\\ 0, & x\leq 0,\end{cases}$$

利用切比雪夫不等式估计 $P(1<X<5)\geq$ _____。

解 利用公式 $\displaystyle\int_0^{+\infty}e^{-x}x^k\,\mathrm{d}x=k!$（即 Gamma 函数）计算期望和方差：

$$EX=\int_0^{+\infty}\frac{1}{2}x^3e^{-x}\,\mathrm{d}x=\frac{1}{2}\Gamma(4)=3,\quad EX^2=\int_0^{+\infty}\frac{1}{2}x^4e^{-x}\,\mathrm{d}x=\frac{1}{2}\Gamma(5)=12,$$

$$DX=EX^2-(EX)^2=12-9=3.$$

注意到 $P(1<X<5)=P(|X-3|<2)$，由切比雪夫不等式：

$$P(|X-3|<2)\geq 1-\frac{DX}{2^2}=1-\frac{3}{4}=\frac{1}{4}.$$

笔记 本题没有直接给出切比雪夫不等式的标准形式，需要先计算 EX 和 DX 来识别 $|X-EX|<\varepsilon$ 的结构。此外，题目中不等号方向为"$\geq$"（求概率下界），因此应使用等价形式 $P\{|X-EX|<\varepsilon\}\geq 1-DX/\varepsilon^2$。

练习 5.6 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}2x, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

则 $P(|X-EX|\geq 2\sqrt{DX})$ 等于（　）。

A. $\dfrac{9-8\sqrt{2}}{9}$　B. $\dfrac{6+4\sqrt{2}}{9}$　C. $\dfrac{6-8\sqrt{2}}{9}$　D. $\dfrac{6-4\sqrt{2}}{9}$

解 应选 D。

$$EX=\int_0^1 2x^2\,\mathrm{d}x=\frac{2}{3},\quad EX^2=\int_0^1 2x^3\,\mathrm{d}x=\frac{1}{2},\quad DX=\frac{1}{2}-\frac{4}{9}=\frac{1}{18}.$$

$2\sqrt{DX}=2\times\dfrac{\sqrt{2}}{6}=\dfrac{\sqrt{2}}{3}$，不等式 $|X-2/3|\geq\sqrt{2}/3$ 的解为 $X\geq\dfrac{2+\sqrt{2}}{3}$ 或 $X\leq\dfrac{2-\sqrt{2}}{3}$。由于 $\dfrac{2+\sqrt{2}}{3}>1$（超出取值范围），第一部分概率为 0。只需计算：

$$P\left(X\leq\frac{2-\sqrt{2}}{3}\right)=\int_0^{\frac{2-\sqrt{2}}{3}}2x\,\mathrm{d}x=\left(\frac{2-\sqrt{2}}{3}\right)^2=\frac{6-4\sqrt{2}}{9}.$$

笔记 作为对比，切比雪夫不等式仅给出粗略估计 $P(|X-EX|\geq 2\sqrt{DX})\leq 1/4=0.25$，而精确值为 $\dfrac{6-4\sqrt{2}}{9}\approx 0.139$。

5.3.1 切比雪夫大数定律

定理 5.2 (切比雪夫大数定律)

设 $X_1, X_2, \ldots, X_n, \ldots$ 是相互独立的随机变量序列，若方差 $DX_i\ (i \geq 1)$ 存在且一致有界，即存在常数 $C > 0$，使得

$$DX_i \leq C \quad (\text{对一切} i \geq 1),$$

则对任意 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} EX_i\right| < \varepsilon\right\} = 1.$$

笔记 切比雪夫大数定律的意义在于：当独立随机变量的方差一致有界时，n 个随机变量的算术平均值依概率收敛于其期望的算术平均值。特别地，若 $EX_i = \mu$（常数），则 $\frac{1}{n}\sum X_i \xrightarrow{P} \mu$。

练习 5.7 设 $X_1, X_2, \ldots, X_n, \ldots$ 是相互独立的随机变量序列，X_n 服从参数为 $n\ (n \geq 1)$ 的指数分布，则下列随机变量序列中不满足切比雪夫大数定律条件的是（　）。

A. $X_1, \frac{1}{2}X_2, \ldots, \frac{1}{n}X_n, \ldots$　B. $X_1, X_2, \ldots, X_n, \ldots$

C. $X_1, 2X_2, \ldots, nX_n, \ldots$　D. $X_1, 2^2X_2, \ldots, n^2X_n, \ldots$

解 应选 D。

切比雪夫大数定律要求：$\{X_n\}$ 相互独立，方差存在且一致有界（$DX_n \leq C$）。

由题设，$DX_n = \dfrac{1}{n^2}$（指数分布 $E(\lambda)$ 的方差为 $\frac{1}{\lambda^2}$，此处 $\lambda = n$）。

逐一检验：

- A: $D\left(\dfrac{1}{n}X_n\right) = \dfrac{1}{n^2}DX_n = \dfrac{1}{n^4} \leq 1$ ✓
- B: $DX_n = \dfrac{1}{n^2} \leq 1$ ✓
- C: $D(nX_n) = n^2\,DX_n = 1 \leq 1$ ✓
- D: $D(n^2X_n) = n^4\,DX_n = n^2$，当 $n \to \infty$ 时无上界 ×

故选 D。

5.3.2 伯努利大数定律

定理 5.3 (伯努利大数定律)

设 μ_n 是 n 重伯努利试验中事件 A 发生的次数，p 是事件 A 在每次试验中发生的概率 $(0 < p < 1)$，则对任意 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right\} = 1,$$

即频率 $\frac{\mu_n}{n}$ 依概率收敛于概率 p。

笔记 伯努利大数定律是切比雪夫大数定律的特例（取 $X_i \sim B(1,p)$），它为"频率稳定于概率"这一经验事实提供了严格的数学证明，也是用频率定义概率的理论依据。

在数理统计中，经验分布函数

$$F_n(x) = \frac{\text{样本值中} \leq x\text{的个数}}{n}$$

正是事件 $\{X \le x\}$ 在 n 次试验中出现的频率。由伯努利大数定律，当 n 充分大时，$F_n(x)$ 可作为总体分布函数 $F(x)$ 的近似，且 n 越大近似越好。

练习 5.8 设 $(2,1,5,2,1,3,1)$ 是来自总体 X 的简单随机样本值，求总体 X 的经验分布函数 $F_7(x)$。

解 将样本值按从小到大排列：$1,1,1,2,2,3,5$。

$$F_7(x) = \begin{cases} 0, & x < 1, \\ \dfrac{3}{7}, & 1 \le x < 2, \\ \dfrac{5}{7}, & 2 \le x < 3, \\ \dfrac{6}{7}, & 3 \le x < 5, \\ 1, & x \ge 5. \end{cases}$$

5.3.3 辛钦大数定律

定理 5.4 (辛钦大数定律)

设 $X_1, X_2, \ldots, X_n, \ldots$ 是独立同分布的随机变量序列，若数学期望 $EX_i = \mu\ (i = 1, 2, \ldots)$ 存在，则对任意 $\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| < \varepsilon\right\} = 1,$$

即 $\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i \xrightarrow{P} \mu$。

证明 [证明思路] 在方差 $DX_k = \sigma^2\ (k = 1, 2, \ldots)$ 存在的条件下给出证明。因为

$$E\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\sum_{k=1}^{n} EX_k = \mu,$$

又由独立性得

$$D\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} DX_k = \frac{\sigma^2}{n}.$$

由切比雪夫不等式：

$$P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| \ge \varepsilon\right\} \le \frac{\sigma^2/n}{\varepsilon^2} \xrightarrow{n\to\infty} 0.$$

笔记 辛钦大数定律不要求方差存在（上述证明是方差存在时的特例），仅要求期望存在，适用范围更广。

辛钦大数定律的更一般形式：若 $X_1, X_2, \ldots$ 独立同分布且 $E(X_1^k)$ 存在，则

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{P} E(X_1^k).$$

在数理统计中，判断估计量是否具有一致性（相合性）时，常常需要利用依概率收敛和大数定律。

练习 5.9 设随机变量序列 $X_1, X_2, \ldots, X_n, \ldots$ 独立同分布，X_1 的概率密度为

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

则当 $n \to \infty$ 时，$\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2$ 依概率收敛于（　）。

A. $\dfrac{1}{8}$　B. $\dfrac{1}{6}$　C. $\dfrac{1}{3}$　D. $\dfrac{1}{2}$

解 应选 B。由辛钦大数定律，$\frac{1}{n}\sum X_i^2 \xrightarrow{P} E(X_1^2)$。由偶函数对称性：

$$E(X_1^2) = \int_{-1}^{1} x^2(1-|x|)\,\mathrm{d}x = 2\int_0^1 x^2(1-x)\,\mathrm{d}x = 2\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^1 = 2\left(\frac{1}{3}-\frac{1}{4}\right) = \frac{1}{6}.$$

练习 5.10 将一枚骰子重复掷 n 次，则当 $n \to \infty$ 时，n 次掷出点数的算术平均值 $\overline{X}_n$ 依概率收敛于_____。

解 应填 $\dfrac{7}{2}$。设第 i 次掷出的点数为 X_i，则 $X_1, X_2, \ldots, X_n$ 独立同分布，且

$$EX_i = \frac{1+2+3+4+5+6}{6} = \frac{7}{2}.$$

由辛钦大数定律，$\overline{X}_n = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i \xrightarrow{P} EX_i = \dfrac{7}{2}$。

5.4 中心极限定理

中心极限定理研究的是：在什么条件下，大量独立随机变量之和的分布趋近于正态分布。它是正态分布在概率论中占据核心地位的根本原因，也是大样本统计推断的理论基础。

5.4.1 列维–林德伯格中心极限定理

定理 5.5 (列维–林德伯格中心极限定理)

设 $X_1, X_2, \ldots, X_n, \ldots$ 是独立同分布的随机变量序列，且数学期望 $EX_i = \mu$ 和方差 $DX_i = \sigma^2 > 0$ 存在 $(i = 1, 2, \ldots)$，则对任意实数 x，有

$$\lim_{n\to\infty} P\left\{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} \le x\right\} = \Phi(x).$$

笔记 应用列维–林德伯格中心极限定理时，需满足三个条件：独立、同分布、期望和方差存在，缺一不可。

当 n 足够大时，$\sum\limits_{i=1}^{n} X_i$ 近似服从正态分布 $N(n\mu, n\sigma^2)$，因此

$$P\left\{a < \sum_{i=1}^{n} X_i < b\right\} \approx \Phi\left(\frac{b-n\mu}{\sqrt{n}\,\sigma}\right) - \Phi\left(\frac{a-n\mu}{\sqrt{n}\,\sigma}\right).$$

只要题目涉及独立同分布随机变量之和的概率计算，且 n 较大，就应考虑使用中心极限定理。

练习 5.11 设随机变量 $X_1, X_2, \ldots, X_n$ 相互独立，$S_n = X_1 + X_2 + \cdots + X_n$。根据列维–林德伯格定理，当 n 充分大时，S_n 近似服从正态分布，只要（　）。

A. 有相同的期望和方差　B. 服从同一离散型分布　C. 服从同一指数分布　D. 服从同一连续型分布

解 应选 C。列维–林德伯格定理要求：独立、同分布、期望和方差均存在。

- A：有相同的期望和方差 $\Rightarrow$ 未必同分布，不满足条件。
- B：服从同一离散型分布 $\Rightarrow$ 方差未必存在（如取值为所有正整数的几何分布，期望和方差存在；但若取值范围无穷且概率下降不够快，方差可能不存在）。
- C：指数分布的期望和方差都存在，满足全部三个条件 ✓
- D：服从同一连续型分布 $\Rightarrow$ 方差未必存在（如柯西分布）。

练习 5.12 设随机变量 $X_1, X_2, \ldots, X_{100}$ 相互独立，均服从参数为 $\frac{1}{2}$ 的指数分布，利用中心极限定理计算

$$P\left(\sum_{i=1}^{100} X_i < 240\right) = \underline{\qquad\qquad}.$$

解 $X_i \sim E\left(\frac{1}{2}\right)$，故 $\mu = EX_i = 2$，$\sigma^2 = DX_i = 4$。

设 $S = \sum_{i=1}^{100} X_i$，则 $ES = 200$，$DS = 400$。

由列维–林德伯格定理，S 近似服从 $N(200, 400)$，故

$$P(S < 240) = P\left(\frac{S-200}{20} < \frac{240-200}{20}\right) \approx \Phi(2).$$

练习 5.13 设某种元件的寿命（小时）服从指数分布，其概率密度函数为

$$f(x) = \begin{cases} \frac{1}{100} e^{-x/100}, & x > 0, \\ 0, & x \le 0, \end{cases}$$

各元件的寿命相互独立。随机取 36 个元件，求寿命总和大于 4200 小时的概率近似值。

解 由密度函数知 $\lambda = \frac{1}{100}$。设第 i 只元件的寿命为 X_i，则 $X_i \sim E\left(\frac{1}{100}\right)$。

$EX_i = 100$，$DX_i = 10000$。令 $Y = \sum_{i=1}^{36} X_i$，则

$$EY = 3600, \quad DY = 36 \times 10000 = 360000.$$

由列维–林德伯格定理，Y 近似服从 $N(3600, 360000)$，故

$$P(Y > 4200) = 1 - P(Y \le 4200) \approx 1 - \Phi\left(\frac{4200-3600}{600}\right) = 1 - \Phi(1) \approx 0.1587.$$

练习 5.14 生产线生产的产品成箱包装，每箱质量是随机的。假设平均每箱重 50 千克，标准差为 5 千克。若用载重为 5 吨的汽车承运，试用中心极限定理说明每辆车最多可以装多少箱，才能保证不超载的概率大于 0.977。（$\Phi(2) = 0.977$）

解 设 $X_i\ (i = 1, 2, \ldots, n)$ 为第 i 箱的质量（千克），X_i 独立同分布，$EX_i = 50$，$\sqrt{DX_i} = 5$。

n 箱总质量 $T_n = \sum_{i=1}^{n} X_i$，$ET_n = 50n$，$\sqrt{DT_n} = 5\sqrt{n}$。

由列维–林德伯格定理，T_n 近似服从 $N(50n, 25n)$：

$$P(T_n \le 5000) = P\left(\frac{T_n - 50n}{5\sqrt{n}} \le \frac{5000-50n}{5\sqrt{n}}\right) \approx \Phi\left(\frac{1000-10n}{\sqrt{n}}\right) > 0.977 = \Phi(2),$$

即 $\dfrac{1000-10n}{\sqrt{n}}>2$。

令 $t=\sqrt{n}$，则 $1000-10t^2>2t$，即 $10t^2+2t-1000<0$。解得 $t<\dfrac{-2+\sqrt{4+40000}}{20}=\dfrac{-2+2\sqrt{10001}}{20}\approx 9.950$，故 $n<99.0$。

因此每辆车最多装 **98** 箱。

练习 5.15 设 $X_1,X_2,\ldots,X_n,\ldots$ 为独立同分布随机变量序列，且均服从参数为 $\lambda\ (\lambda>1)$ 的指数分布，记 $\Phi(x)$ 为标准正态分布函数，则（　）。

A. $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum_{i=1}^n X_i-n\lambda}{\lambda\sqrt{n}}\le x\right\}=\Phi(x)$　　B. $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum_{i=1}^n X_i-n\lambda}{\sqrt{\lambda n}}\le x\right\}=\Phi(x)$

C. $\lim\limits_{n\to\infty}P\left\{\dfrac{\lambda\sum_{i=1}^n X_i-n}{\sqrt{n}}\le x\right\}=\Phi(x)$　　D. $\lim\limits_{n\to\infty}P\left\{\dfrac{\sum_{i=1}^n X_i-\lambda}{\sqrt{n\lambda}}\le x\right\}=\Phi(x)$

解 应选 C。$X_i\sim E(\lambda)$，故 $EX_i=\dfrac{1}{\lambda}$，$DX_i=\dfrac{1}{\lambda^2}$。设 $S_n=\sum\limits_{i=1}^n X_i$，则 $ES_n=\dfrac{n}{\lambda}$，$DS_n=\dfrac{n}{\lambda^2}$。由列维–林德伯格中心极限定理：

$$\frac{S_n-ES_n}{\sqrt{DS_n}}=\frac{S_n-n/\lambda}{\sqrt{n}/\lambda}=\frac{\lambda S_n-n}{\sqrt{n}}\to N(0,1),$$

故选 C。

练习 5.16 设 $X_1,X_2,\ldots,X_n$ 相互独立，且均服从泊松分布 $P(2)$，则由列维–林德伯格中心极限定理

$$\lim_{n\to\infty}P\left(\frac{\sum_{i=1}^n X_i-2n}{\sqrt{n}}<2\right)=(\quad).$$

A. $\Phi(\sqrt{2})$　B. $\Phi(2)$　C. $\Phi(1)$　D. 1

解 应选 A。$X_i\sim P(2)$，故 $EX_i=2$，$DX_i=2$。由中心极限定理：

$$\frac{\sum X_i-n\cdot 2}{\sqrt{n\cdot 2}}\to N(0,1).$$

题目中式子变形：

$$P\left(\frac{\sum X_i-2n}{\sqrt{n}}<2\right)=P\left(\frac{\sum X_i-2n}{\sqrt{2n}}<\frac{2}{\sqrt{2}}\right)=P(Z<\sqrt{2})=\Phi(\sqrt{2}).$$

练习 5.17 假设 $X_1,X_2,\ldots,X_n$ 是来自总体 X 的简单随机样本，已知 $E(X^k)=\alpha_k\ (k=1,2,3,4)$。证明：当 n 充分大时，随机变量

$$Z_n=\frac{1}{n}\sum_{i=1}^n X_i^2$$

近似服从正态分布，并指出其分布参数。

解 由题设，$X_1,X_2,\ldots,X_n$ 独立同分布，故 $X_1^2,X_2^2,\ldots,X_n^2$ 也独立同分布，且

$$E(X_i^2)=\alpha_2,\quad D(X_i^2)=E(X_i^4)-[E(X_i^2)]^2=\alpha_4-\alpha_2^2.$$

由列维–林德伯格中心极限定理，

$$\frac{\sum_{i=1}^n X_i^2-n\alpha_2}{\sqrt{n(\alpha_4-\alpha_2^2)}}\to N(0,1),$$

等价于

$$\frac{Z_n-\alpha_2}{\sqrt{(\alpha_4-\alpha_2^2)/n}}\to N(0,1),$$

因此当 n 充分大时，Z_n 近似服从正态分布，参数为

$$\mu = \alpha_2, \quad \sigma^2 = \frac{\alpha_4 - \alpha_2^2}{n}.$$

5.4.2 棣莫弗–拉普拉斯中心极限定理

定理 5.6 (棣莫弗–拉普拉斯中心极限定理)

设随机变量 $Y_n \sim B(n,p)$ $(0 < p < 1,\ n \geq 1)$，则对任意实数 x，有

$$\lim_{n\to\infty} P\left\{\frac{Y_n - np}{\sqrt{np(1-p)}} \leq x\right\} = \Phi(x).$$

♡

笔记 棣莫弗–拉普拉斯定理是列维–林德伯格定理的特例。若将 Y_n 看作 n 个独立同分布的 $B(1,p)$ 随机变量之和，即 $Y_n = \sum_{i=1}^n X_i$，其中 $X_i \sim \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$，则 $\mu = p$，$\sigma^2 = p(1-p)$，代入列维–林德伯格定理即得。

二项分布概率计算的三种方法：设 $X \sim B(n,p)$。

1. n 较小（$n \leq 10$）：直接用公式 $P\{X = k\} = \mathrm{C}_n^k p^k (1-p)^{n-k}$。
2. n 较大且 p 较小（$n > 10$, $p < 0.1$, np 适中）：泊松近似 $P\{X = k\} \approx \dfrac{\lambda^k}{k!} e^{-\lambda}$，其中 $\lambda = np$。
3. n 较大且 p 不太大（$np \geq 10$）：正态近似（中心极限定理）

$$P\{a < X < b\} \approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right).$$

✍ **练习 5.18** 从废品率为 0.03 的大量产品中随机抽取 1000 个，根据中心极限定理，废品数 X 近似服从的分布为_____（要求写出分布的参数）。

解 $X \sim B(1000, 0.03)$。$EX = 30$，$DX = 1000 \times 0.03 \times 0.97 = 29.1$。由棣莫弗–拉普拉斯中心极限定理，$X$ 近似服从 $N(30,\ 29.1)$。

✍ **练习 5.19** 某宿舍楼按 400 名学生的规模建造，每人每天有约 10% 的时间占用一个水龙头（各人独立）。为能以 95% 的概率保证用水需要，至少应安装多少个水龙头？ $(\Phi(1.645) = 0.95)$

解 设同一时刻用水人数 $X \sim B(400, 0.1)$，$EX = 40$，$DX = 36$。由中心极限定理，X 近似服从 $N(40, 36)$。设安装 N 个水龙头：

$$P(X \leq N) \approx \Phi\left(\frac{N - 40}{6}\right) \geq 0.95 = \Phi(1.645),$$

故 $\dfrac{N-40}{6} \geq 1.645$，$N \geq 49.87$，取 $N = 50$。

✍ **练习 5.20** 某生产线产品合格率为 0.8。要使一批产品的合格率落在 76% 与 84% 之间的概率不小于 90%，问这批产品至少要生产多少件？ $(\Phi(1.64) = 0.95)$

解 设生产 n 件，合格数 $X \sim B(n, 0.8)$。要求

$$P\left(0.76 \leq \frac{X}{n} \leq 0.84\right) \geq 0.90.$$

标准化：$E(X/n) = 0.8$，$\mathrm{Var}(X/n) = 0.16/n$。

$$P\left(\left|\frac{X}{n} - 0.8\right| \leq 0.04\right) = P\left(|Z| \leq \frac{0.04}{\sqrt{0.16/n}}\right) = P(|Z| \leq 0.1\sqrt{n}) = 2\Phi(0.1\sqrt{n}) - 1 \geq 0.90.$$

$\Phi(0.1\sqrt{n}) \geq 0.95$，$0.1\sqrt{n} \geq 1.64$，$n \geq 268.96$，取 $n = 269$。

✍ **练习 5.21** 设 X_n 表示将一枚硬币随意投掷 n 次出现正面的次数，则（　）。

A. $\lim\limits_{n\to\infty} P\left\{\dfrac{X_n-n}{\sqrt{n}}\le x\right\}=\Phi(x)$　B. $\lim\limits_{n\to\infty} P\left\{\dfrac{2X_n-n}{\sqrt{n}}\le x\right\}=\Phi(x)$

C. $\lim\limits_{n\to\infty} P\left\{\dfrac{X_n-2n}{\sqrt{n}}\le x\right\}=\Phi(x)$　D. $\lim\limits_{n\to\infty} P\left\{\dfrac{2X_n-2n}{\sqrt{n}}\le x\right\}=\Phi(x)$

解 应选 B。

由题设 $X_n\sim B(n,1/2)$，$np=n/2$，$np(1-p)=n/4$。由棣莫弗–拉普拉斯定理：

$$\frac{X_n-n/2}{\sqrt{n/4}}=\frac{X_n-n/2}{\sqrt{n}/2}=\frac{2X_n-n}{\sqrt{n}}\to N(0,1).$$

✍ **练习 5.22** 设某种电气元件不能承受超负荷试验的概率为 0.05。现对 100 个这样的元件进行超负荷试验，以 X 表示不能承受试验而烧毁的元件数。根据中心极限定理，求 $P\{5\le X\le 10\}$ 的近似值。($\Phi(2.29)=0.989$)

解 $X\sim B(100,0.05)$，$EX=np=5$，$DX=np(1-p)=4.75$。由棣莫弗–拉普拉斯中心极限定理：

$$\begin{aligned}P\{5\le X\le 10\}&=P\left\{\frac{5-5}{\sqrt{4.75}}\le\frac{X-5}{\sqrt{4.75}}\le\frac{10-5}{\sqrt{4.75}}\right\}\\&=P\left\{0\le\frac{X-5}{\sqrt{4.75}}\le 2.29\right\}\approx\Phi(2.29)-\Phi(0)=0.989-0.5=0.489.\end{aligned}$$

✍ **练习 5.23** 设 $X_1,X_2,\ldots,X_{100}$ 为来自总体 X 的简单随机样本，其中 $P\{X=0\}=P\{X=1\}=\dfrac{1}{2}$。利用中心极限定理求 $P\left\{\sum\limits_{i=1}^{100}X_i\le 55\right\}$ 的近似值（　）。

A. $1-\Phi(1)$　B. $\Phi(1)$　C. $1-\Phi(0.2)$　D. $\Phi(0.2)$

解 应选 B。$X_i\sim B(1,1/2)$，$EX_i=\dfrac{1}{2}$，$DX_i=\dfrac{1}{4}$。设 $S=\sum\limits_{i=1}^{100}X_i$，则 $ES=50$，$DS=25$。由中心极限定理，S 近似服从 $N(50,25)$：

$$P(S\le 55)=P\left(\frac{S-50}{5}\le\frac{55-50}{5}\right)\approx\Phi(1).$$

✍ **练习 5.24** 在一次集体登山活动中，每个人意外受伤的概率为 1%，各人是否受伤相互独立。

(1) 为保证没有人受伤的概率大于 0.90，应如何控制参加人数？

(2) 若有 100 人参加，求受伤人数不超过 2 人的概率的近似值（用中心极限定理，结果可用 Φ 表示）。

解 (1) 设人数为 n，受伤人数 $X\sim B(n,0.01)$。要求 $P(X=0)=0.99^n>0.9$。取对数：

$$n\ln 0.99>\ln 0.9\implies n<\frac{\ln 0.9}{\ln 0.99}\approx 10.48,$$

故人数控制在 **10** 人及以内。

(2) $n=100$，$X\sim B(100,0.01)$，$EX=1$，$DX=0.99$。

$$P(X\le 2)=P\left(\frac{X-1}{\sqrt{0.99}}\le\frac{1}{\sqrt{0.99}}\right)\approx\Phi\left(\frac{1}{\sqrt{0.99}}\right)=\Phi(\sqrt{100/99}).$$

✍ **练习 5.25** 从甲地前往乙地有两条路线：第一条穿过市区，用时 $X_1\sim N(50,100)$（分钟）；第二条走环城公路，用时 $X_2\sim N(60,16)$（分钟）。若只有 65 分钟可用，应走哪条路线？

解

$$P(X_1\le 65)=\Phi\left(\frac{65-50}{10}\right)=\Phi(1.5)=0.9332,$$

$$P(X_2 \le 65) = \Phi\left(\frac{65-60}{4}\right) = \Phi(1.25) = 0.8944.$$

$0.9332 > 0.8944$，故走第一条路线较好。

笔记 本题本质是比较两个正态分布在同一阈值下的累积概率，不涉及中心极限定理。但作为正态分布在实际决策中的应用，与 CLT 应用题有很好的互补性。

练习 5.26 某地区某种疾病的患病率为 0.1，现对 10000 人进行血检普查（各人独立）。

(1) 用中心极限定理估计患病人数在 $[970, 1030]$ 内的概率（用 Φ 表示）；

(2) 有两套化验方案：方案一逐一化验（10000 次）；方案二将每 k 人分为一组，先混合化验，阴性则只需 1 次，阳性则再逐一化验另半血样（共 $k+1$ 次）。问 k 满足什么条件时方案二更优？$k=4$ 时方案二的平均化验次数是多少？

解 (1) 患病人数 $X \sim B(10000, 0.1)$，$EX = 1000$，$DX = 900$，$\sigma = 30$。

$$P(970 \le X \le 1030) \approx \Phi\left(\frac{30}{30}\right) - \Phi\left(\frac{-30}{30}\right) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1.$$

(2) 每组 k 人全为阴性的概率为 0.9^k，每组化验次数 Y_i 满足：

$$E(Y_i) = 1 \cdot 0.9^k + (k+1)(1-0.9^k) = k + 1 - k \cdot 0.9^k.$$

总期望 $E(T_2) = \dfrac{10000}{k}(k+1-k\cdot 0.9^k) = 10000\left(1 + \dfrac{1}{k} - 0.9^k\right)$。

方案二更优的条件为 $E(T_2) < 10000$，即

$$k \cdot 0.9^k > 1.$$

当 $k=4$ 时：

$$E(Y_i) = 5 - 4 \times 0.9^4 = 5 - 4 \times 0.6561 = 2.3756,$$

$$E(T_2) = \frac{10000}{4} \times 2.3756 = 5939.$$

第 6 章 样本与抽样分布

- 基本概念
  - 总体与个体
  - 样本
    - 代表性：与总体具有相同概率分布
    - 独立性：个体之间相互独立
- 统计量
  - 概念：不含其它未知参数的样本函数
  - 常见统计量：
    - 样本均值：$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$
    - 样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\bar{X}^2\right)$
    - 样本k阶原点矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$
    - 样本k阶中心距：$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k, k=2,3\cdots$
  - 常用统计量的数字特征
    - $E\bar{X}=EX, D\bar{X}=\frac{DX}{n}$
    - $E(S^2)=DX$
- 抽样分布
  - χ^2分布
    - 形式：n个独立标准正态分布的平方和，自由度为n
    - 性质：(1) 独立可加性；(2) 若$X\sim\chi^2(n)$, 则$EX=n, DX=2n$
  - t分布
    - 形式：$Z=\frac{X}{\sqrt{Y/n}}$, 其中X,Y独立且$X\sim N(0,1), Y\sim\chi^2(n)$
    - 性质：若$Z\sim t(n)$, 则$EZ=0, DZ=\frac{n}{n-2}$
  - F分布
    - 形式：$F=\frac{X/n_1}{Y/n_2}\sim F(n_1,n_2)$, 其中$X,Y$独立且$X\sim\chi^2(n_1), Y\sim\chi^2(n_2)$
    - 性质：若$X\sim t(n)$, 则$X^2\sim F(1,n)$
  - 单正态总体$X\sim N(\mu,\sigma^2)$抽样
    - $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1), \frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1), \bar{X}$与$S^2$独立
    - $\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n), \frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$
  - 两个正态总体$X\sim N(\mu_1,\sigma_1^2), Y\sim N(\mu_2,\sigma_2^2)$的抽样分布

内容提要

- 总体、个体与样本的概念 6.1
- 统计量与常用统计量 6.3
- 三大抽样分布 6.5, 6.6, 6.7
- 正态总体下的抽样分布定理 6.1, 6.2
- 分布的分位数 6.8

6.1 基本概念

6.1.1 总体与样本

定义 6.1 (总体与个体)

研究对象的全体称为总体（population），组成总体的每一个元素称为个体（individual）。在对总体进行统计研究时，通常将总体与某个随机变量 X 等同起来，称"总体 X"，总体的分布即指随机变量 X 的分布。 ♣

定义 6.2 (简单随机样本)

设 X 是具有分布函数 F 的随机变量，若 $X_1, X_2, \ldots, X_n$ 是与 X 具有同一分布函数 F 且相互独立的随机变量，则称 $X_1, X_2, \ldots, X_n$ 为从总体 X（或分布函数 F）得到的容量为 n 的简单随机样本，简称样本。它们的观察值 $x_1, x_2, \ldots, x_n$ 称为样本值。 ♣

性质 [样本的分布] 设总体 X 的分布函数为 $F(x)$，则样本 $X_1, X_2, \ldots, X_n$ 的联合分布函数为

$$F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F(x_i).$$

对于离散型总体，联合分布为 $P\{X_1 = x_1, \ldots, X_n = x_n\} = \prod_{i=1}^{n} P\{X_i = x_i\}$；对于连续型总体，联合密度为 $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i)$。

6.1.2 统计量

定义 6.3 (统计量)

设 $X_1, X_2, \ldots, X_n$ 为来自总体 X 的一个样本，$g(X_1, X_2, \ldots, X_n)$ 是 $X_1, X_2, \ldots, X_n$ 的函数。若 g 中不含任何未知参数，则称 $g(X_1, X_2, \ldots, X_n)$ 为一个统计量。 ♣

笔记 统计量是样本的函数，也是随机变量。统计量中绝不能出现未知参数，这是判断一个表达式是否为统计量的核心依据。

练习 6.1 设总体 $X \sim N(\mu, \sigma^2)$，其中 μ 未知，σ^2 已知，X_1, X_2, X_3 是简单随机样本，则下列表达式中不是统计量的是（　）。

A. $X_1 + X_2 + X_3$　　B. $\min(X_1, X_2, X_3)$　　C. $\sum_{i=1}^{3} \frac{X_i^2}{\sigma^2}$　　D. $\overline{X} + 2\mu$

解 应选 D。选项 D 中含有未知参数 μ，因此不是统计量。

练习 6.2 设 (X_1, X_2, X_3) 为取自总体 $N(0, \sigma^2)$ 的样本，σ^2 为未知参数。$T_1 = X_1 + X_2 + X_3$，$T_2 = \frac{X_3 - X_1}{2}$，$T_3 = \frac{1}{\sigma}\sum_{i=1}^{3} X_i$，$T_4 = \max\{X_1, X_2, X_3\}$，其中不是统计量的是（　）。

A. T_1　　B. T_2　　C. T_3　　D. T_4

解 应选 C。T_3 中含有未知参数 σ，因此不是统计量。

练习 6.3 设 (X_1, X_2, X_3) 为取自总体的样本，σ^2 已知，μ 未知，$T_1 = X_1 + X_2 + X_3$，$T_2 = (X_3 - \mu)/2$，$T_3 = \sum_{i=1}^{3} X_i$，$T_4 = \max\{X_1, X_2, X_3\}$，其中不是统计量的是（　）。

A. T_1　B. T_2　C. T_3　D. T_4

解 应选 B。统计量是不含任何未知参数的样本函数。A、C、D 均由样本构成且不含未知参数。$T_2 = (X_3 - \mu)/2$ 含有未知参数 μ，不是统计量。

定义 6.4 (常用统计量)

设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的样本，$\overline{X}$ 和 S^2 分别为样本均值与样本方差。

1. 样本均值：$\overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$；
2. 样本方差：$S^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2 = \dfrac{1}{n-1}\left(\sum\limits_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$；
3. 样本标准差：$S = \sqrt{S^2}$；
4. 样本 k 阶原点矩：$A_k = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^k \ (k = 1, 2, \ldots)$；
5. 样本 k 阶中心矩：$B_k = \dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^k \ (k = 2, 3, \ldots)$。

♣

性质 [统计量的数字特征] 设总体 X 的期望为 $EX = \mu$，方差为 $DX = \sigma^2$，$X_1, X_2, \ldots, X_n$ 为容量为 n 的样本，则

$$EX_i = \mu, \quad DX_i = \sigma^2, \quad E\overline{X} = \mu, \quad D\overline{X} = \frac{\sigma^2}{n}, \quad E(S^2) = \sigma^2.$$

笔记 由大数定律，样本均值和样本方差依概率收敛于总体均值和总体方差：$\overline{X} \xrightarrow{P} \mu$，$S^2 \xrightarrow{P} \sigma^2$，$A_k \xrightarrow{P} EX^k$。

练习 6.4 设 $X_1, X_2, \ldots, X_n \ (n > 2)$ 为独立同分布的随机变量，均服从 $N(0,1)$，记 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，$Y_i = X_i - \overline{X} \ (i = 1, 2, \ldots, n)$。则 $DY_i =$ _____。

解 应填 $\dfrac{n-1}{n}$。

$$DY_i = D\left[\left(1 - \frac{1}{n}\right)X_i - \frac{1}{n}\sum_{k \neq i} X_k\right] = \left(1 - \frac{1}{n}\right)^2 + \frac{n-1}{n^2} = \frac{n-1}{n}.$$

练习 6.5 已知总体 X 的期望 $EX = 0$，方差 $DX = \sigma^2$，从总体 X 中抽取容量为 n 的简单随机样本，其样本均值和样本方差分别为 $\overline{X}, S^2$。记 $S_k^2 = k\overline{X}^2 + kS^2 \ (k = 1, 2, 3, 4)$，则（　）。

A. $E(S_1^2) = \sigma^2$　B. $E(S_2^2) = \sigma^2$　C. $E(S_3^2) = \sigma^2$　D. $E(S_4^2) = \sigma^2$

解 应选 B。

$E(\overline{X}^2) = D\overline{X} + (E\overline{X})^2 = \dfrac{\sigma^2}{n} + 0 = \dfrac{\sigma^2}{n}$，$E(S^2) = \sigma^2$。

$$E(S_k^2) = k\,E(\overline{X}^2) + k\,E(S^2) = k \cdot \frac{\sigma^2}{n} + k\sigma^2 = k\sigma^2\left(\frac{1}{n} + 1\right) = k\sigma^2 \cdot \frac{n+1}{n}.$$

令 $E(S_k^2) = \sigma^2$，得 $k(n+1) = n$，即 $k = \dfrac{n}{n+1}$。当 $n \to \infty$ 时 $k \to 1$，但精确值取决于 n。

另由 $E(S_k^2) = k \cdot \frac{\sigma^2}{n} + k\sigma^2 = k\sigma^2\left(1 + \frac{1}{n}\right)$，需要 $k\left(1 + \frac{1}{n}\right) = 1$，即 $k = \dfrac{n}{n+1}$。

对于一般的 n，四个选项都不精确成立。但当 n 足够大时，$k \approx 1$（即 A 选项）最接近。但若题目隐含要求精确成立，则无正确选项。

笔记 此题需注意 $E(\overline{X}^2) = D\overline{X} = \sigma^2/n$（因为 $E\overline{X} = 0$），而非 $(EX)^2 = 0$。

6.2 三大抽样分布

6.2.1 χ^2 分布

定义 6.5 (χ^2 分布)

设 $X_1, X_2, \ldots, X_n$ 相互独立且都服从标准正态分布 $N(0,1)$，令

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2,$$

则称 χ^2 服从自由度为 n 的 χ^2 分布，记为 $\chi^2 \sim \chi^2(n)$。♣

性质

1. $E[\chi^2(n)] = n$，$D[\chi^2(n)] = 2n$。
2. 可加性：若 $X \sim \chi^2(m)$，$Y \sim \chi^2(n)$，且 X, Y 相互独立，则 $X + Y \sim \chi^2(m+n)$。

✍ **练习 6.6** 设 $(X_1, X_2, \ldots, X_{10})$ 与 $(Y_1, Y_2, \ldots, Y_{10})$ 分别是来自正态总体 $N(0,1)$ 的两组独立样本，$Z = \sum_{i=1}^{10}(X_i - \overline{X})^2 + \sum_{i=1}^{10}(Y_i - \overline{Y})^2$，则 $EZ =$ _____。

解 应填 18。$\sum_{i=1}^{10}(X_i - \overline{X})^2 \sim \chi^2(9)$，$\sum_{i=1}^{10}(Y_i - \overline{Y})^2 \sim \chi^2(9)$，且两组独立。由 χ^2 分布可加性，$Z \sim \chi^2(18)$，故 $EZ = 18$。

✍ **练习 6.7** 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, 2^2)$ 的样本，令 $Y = (X_1 + X_2)^2 + (X_3 - X_4)^2$。当 $C =$ _____ 时，$CY \sim \chi^2(2)$。

解 应填 $\dfrac{1}{8}$。

$X_1 + X_2 \sim N(0,8)$，$X_3 - X_4 \sim N(0,8)$。标准化：$\dfrac{X_1 + X_2}{\sqrt{8}} \sim N(0,1)$，$\dfrac{X_3 - X_4}{\sqrt{8}} \sim N(0,1)$。

$$\left(\frac{X_1 + X_2}{\sqrt{8}}\right)^2 + \left(\frac{X_3 - X_4}{\sqrt{8}}\right)^2 \sim \chi^2(2) \implies C = \frac{1}{8}.$$

✍ **练习 6.8** 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, 2^2)$ 的简单随机样本，记

$$X = a(X_1 - 2X_2)^2 + b(3X_3 - 4X_4)^2,$$

若统计量 X 服从自由度为 2 的 χ^2 分布，则 $ab =$ _____。

解 应填 $\dfrac{1}{2000}$。

$X_1 - 2X_2 \sim N(0,\ 4 \times 1 + 4 \times 4) = N(0, 20)$，$3X_3 - 4X_4 \sim N(0,\ 9 \times 4 + 16 \times 4) = N(0, 100)$。

由 χ^2 分布的定义：

$$\frac{(X_1 - 2X_2)^2}{20} + \frac{(3X_3 - 4X_4)^2}{100} \sim \chi^2(2),$$

故 $a = \dfrac{1}{20}$，$b = \dfrac{1}{100}$，$ab = \dfrac{1}{2000}$。

6.2.2 t 分布

定义 6.6 (t 分布)

设 $X \sim N(0,1)$，$Y \sim \chi^2(n)$，且 X, Y 相互独立，令

$$t = \frac{X}{\sqrt{Y/n}},$$

则称 t 服从自由度为 n 的 t 分布，记为 $t \sim t(n)$。♣

性质

1. t 分布的概率密度关于 $x=0$ 对称，$Et=0\ (n\geq 2)$，$Dt=\dfrac{n}{n-2}\ (n>2)$。
2. 当 n 足够大时，t 分布近似于标准正态分布；n 越大，曲线越"瘦"越"高"。
3. 由对称性：$t_{1-\alpha}(n)=-t_\alpha(n)$。
4. 若 $t\sim t(n)$，则 $t^2\sim F(1,n)$。

练习 6.9 设 $X_1,X_2,\ldots,X_n$ 为来自正态分布 $N(\mu,\sigma^2)$ 的简单随机样本，$\overline{X}$ 和 S^2 分别为样本均值和样本方差。又设 $X_{n+1}\sim N(\mu,\sigma^2)$ 且与 $X_1,\ldots,X_n$ 相互独立，则统计量

$$\frac{X_{n+1}-\overline{X}}{S}\sqrt{\frac{n}{n+1}}$$

服从_____分布（注明自由度）。

解 应填 $t(n-1)$。

$\overline{X}\sim N(\mu,\sigma^2/n)$，$X_{n+1}\sim N(\mu,\sigma^2)$，且相互独立，故

$$X_{n+1}-\overline{X}\sim N\left(0,\ \left(1+\frac{1}{n}\right)\sigma^2\right),\quad \frac{X_{n+1}-\overline{X}}{\sigma\sqrt{1+1/n}}\sim N(0,1).$$

又 $\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，且与 $\overline{X}$ 独立，故

$$\frac{X_{n+1}-\overline{X}}{S}\sqrt{\frac{n}{n+1}}=\frac{(X_{n+1}-\overline{X})/\left(\sigma\sqrt{1+\frac{1}{n}}\right)}{\sqrt{(n-1)S^2/[\sigma^2(n-1)]}}\sim t(n-1).$$

练习 6.10 设随机变量 $X\sim t(n)$，$Y\sim F(1,n)$，给定 $\alpha\ (0<\alpha<0.5)$，常数 c 满足 $P\{X>c\}=\alpha$，则 $P\{Y>c^2\}=$_____。

A. α　B. $1-\alpha$　C. 2α　D. $1-2\alpha$

解 应选 C。由 $X\sim t(n)$ 知 $X^2\sim F(1,n)$，故 Y 与 X^2 同分布：

$$P\{Y>c^2\}=P\{X^2>c^2\}=P\{X>c\}+P\{X<-c\}=\alpha+\alpha=2\alpha.$$

笔记 此题只能断言 Y 与 X^2 同分布（即 $P(Y>c^2)=P(X^2>c^2)$），不能说 $Y=X^2$——服从相同分布的随机变量不一定相同。

6.2.3 F 分布

定义 6.7 (F 分布)

设 $X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，且 X,Y 相互独立，令

$$F=\frac{X/n_1}{Y/n_2},$$

则称 F 服从自由度为 (n_1,n_2) 的 F 分布，记为 $F\sim F(n_1,n_2)$，其中 n_1 为第一自由度，n_2 为第二自由度。 ♣

性质

1. 若 $F\sim F(n_1,n_2)$，则 $\dfrac{1}{F}\sim F(n_2,n_1)$。
2. $F_{1-\alpha}(n_1,n_2)=\dfrac{1}{F_\alpha(n_2,n_1)}$，常用于求 F 分布表中未列出的分位数。
3. 若 $t\sim t(n)$，则 $t^2\sim F(1,n)$。

练习 6.11 设 X_1, X_2, X_3, X_4 相互独立，均服从 $N(0,1)$。$Z = c \times \dfrac{X_1^2/1}{X_2^2+X_3^2+X_4^2}$ 服从 F 分布，则常数 $c =$ _____。

A. 3　B. $\dfrac{1}{3}$　C. $\dfrac{2}{3}$　D. 1

解 应选 A。$X_1^2 \sim \chi^2(1)$，$X_2^2+X_3^2+X_4^2 \sim \chi^2(3)$，且相互独立。由 F 分布的定义：

$$\frac{X_1^2/1}{(X_2^2+X_3^2+X_4^2)/3} \sim F(1,3) \implies c = 3.$$

练习 6.12 设随机变量 $T \sim t(15)$，则 T^2 服从的分布为_____。

解 应填 $F(1,15)$。自由度为 n 的 t 分布的平方服从 $F(1,n)$。

练习 6.13 设随机变量 $X \sim t(n)\ (n>1)$，则 $Y = \dfrac{1}{X^2}$ 服从的分布为（　）。

A. $\chi^2(n)$　B. $\chi^2(n-1)$　C. $F(n,1)$　D. $F(1,n)$

解 应选 C。设 $X = \dfrac{U}{\sqrt{V/n}}$，其中 $U \sim N(0,1)$，$V \sim \chi^2(n)$ 且相互独立。$X^2 = \dfrac{U^2/1}{V/n} \sim F(1,n)$，取倒数得

$$Y = \frac{1}{X^2} = \frac{V/n}{U^2/1} \sim F(n,1).$$

6.3 正态总体下的抽样分布

6.3.1 单个正态总体

定理 6.1 (单正态总体抽样分布)

设总体 $X \sim N(\mu,\sigma^2)$，$X_1, X_2, \ldots, X_n$ 为来自 X 的样本，$\overline{X}, S^2$ 分别为样本均值和样本方差，则

1. $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，即 $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$；
2. $\dfrac{(n-1)S^2}{\sigma^2} = \displaystyle\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$，且 $\overline{X}$ 与 S^2 相互独立；
3. $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$（当 σ 未知时，用 S 替代 σ）。

笔记 定理中结论 (2) 的证明涉及矩阵知识，不要求掌握，但其结论必须牢记。注意自由度：$\sum(X_i-\overline{X})^2/\sigma^2$ 的自由度为 $n-1$（而非 n），因为 $\overline{X}$ 消耗了一个自由度。

进一步有 $\dfrac{n(\overline{X}-\mu)^2}{S^2} \sim F(1,n-1)$。

练习 6.14 设 $X_1, X_2, \ldots, X_n$ 是总体 $X \sim N(\mu,\sigma^2)$ 的简单随机样本，$\overline{X}$ 为样本均值，则下列结论正确的是（　）。

A. $\dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2 \sim \chi^2(n-1)$　　B. $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2 \sim \chi^2(n-1)$

C. $\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2 \sim \chi^2(n-1)$　　D. $\dfrac{1}{\sigma^2}\displaystyle\sum_{i=1}^{n}(X_i-\overline{X})^2 \sim \chi^2(n)$

解 应选 A。由定理，$\dfrac{(n-1)S^2}{\sigma^2} = \displaystyle\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$，这正是选项 A。

练习 6.15 设 $X_1, X_2, \ldots, X_n$ 是来自总体 $X \sim N(\mu,\sigma^2)$ 的简单随机样本，$\overline{X}$ 与 S^2 分别为其样本均值和样本方差，则下列结论正确的是（　）。

A. $2X_2 - X_1 \sim N(\mu, \sigma^2)$　B. $\dfrac{n(\overline{X}-\mu)^2}{S^2} \sim F(1, n-1)$

C. $\dfrac{S^2}{\sigma^2} \sim \chi^2(n-1)$　D. $\dfrac{\overline{X}-\mu}{S}\sqrt{n-1} \sim t(n-1)$

解 应选 B。

A: $D(2X_2 - X_1) = 4\sigma^2 + \sigma^2 = 5\sigma^2 \neq \sigma^2$ ×。

B: $\dfrac{\sqrt{n}(\overline{X}-\mu)}{S} \sim t(n-1)$，其平方 $\dfrac{n(\overline{X}-\mu)^2}{S^2} \sim F(1, n-1)$ ✓。

C: 应为 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ ×。

D: 应为 $\dfrac{\sqrt{n}(\overline{X}-\mu)}{S} \sim t(n-1)$，系数为 $\sqrt{n}$ 而非 $\sqrt{n-1}$ ×。

练习 6.16 设总体 $X \sim N(1,4)$,$X_1, X_2, \ldots, X_n$ 为来自总体 X 的样本,$\overline{X}$ 为样本均值,则 $E\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right) =$ _____。

A. $4n$　B. $4(n-1)$　C. $2n$　D. $2(n-1)$

解 应选 B。

$\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2 \sim \chi^2(n-1)$，自由度为 $n-1$ 的 χ^2 分布期望为 $n-1$，故

$$E\left(\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = \sigma^2(n-1) = 4(n-1).$$

练习 6.17 设总体 $X \sim N(0, 3^2)$，$X_1, X_2, \ldots, X_8$ 是来自总体 X 的简单随机样本，则统计量

$$Y = \frac{X_1 + X_2 + X_3 + X_4}{\sqrt{X_5^2 + X_6^2 + X_7^2 + X_8^2}}$$

服从（　）。

A. $\chi^2(4)$　B. $t(4)$　C. $F(4,4)$　D. $N(0, 4^2)$

解 应选 B。

$$U = \frac{X_1 + X_2 + X_3 + X_4}{\sqrt{4 \times 9}} = \frac{X_1 + X_2 + X_3 + X_4}{6} \sim N(0,1),$$

$$V = \frac{X_5^2 + X_6^2 + X_7^2 + X_8^2}{9} \sim \chi^2(4),$$

且 U, V 相互独立。故

$$Y = \frac{(X_1 + X_2 + X_3 + X_4)/6}{\sqrt{(X_5^2 + \cdots + X_8^2)/4}} = \frac{U}{\sqrt{V/4}} \sim t(4).$$

练习 6.18 设 X_1, X_2 是取自正态总体 $N(\mu, \sigma^2)$ 的简单随机样本，$Y_1 = X_1 + X_2$，$Y_2 = X_1 - X_2$。则 $\mathrm{Cov}(Y_1, Y_2) =$ _____；已知 (Y_1, Y_2) 服从二维正态分布，若 c 为非零常数，则当 $c =$ _____ 时，$\dfrac{c(Y_1 - 2\mu)}{|Y_2|}$ 服从自由度为 _____ 的 _____ 分布。

解 第一空填 0：

$$\mathrm{Cov}(Y_1, Y_2) = \mathrm{Cov}(X_1 + X_2,\ X_1 - X_2) = DX_1 - DX_2 = \sigma^2 - \sigma^2 = 0.$$

第二空填 1，第三空填 1，第四空填 t。由 (Y_1, Y_2) 服从二维正态分布且协方差为 0，知 Y_1, Y_2 相互独立。

$Y_1 \sim N(2\mu, 2\sigma^2)$，故 $\dfrac{Y_1 - 2\mu}{\sqrt{2}\,\sigma} \sim N(0,1)$；$Y_2 \sim N(0, 2\sigma^2)$，故 $\left(\dfrac{Y_2}{\sqrt{2}\,\sigma}\right)^2 \sim \chi^2(1)$。构造 t 分布：

$$t = \frac{(Y_1 - 2\mu)/(\sqrt{2}\,\sigma)}{\sqrt{(Y_2/(\sqrt{2}\,\sigma))^2}} = \frac{Y_1 - 2\mu}{|Y_2|} \sim t(1),$$

故 $c = 1$。

练习 6.19 设 $(X_1, X_2, \ldots, X_{10})$ 是取自正态总体 $N(0,4)$ 的一个样本，$\overline{X}$ 为样本均值，S^2 为样本方差。当 $c =$ _____ 时，$c\dfrac{\overline{X}^2}{S^2}$ 服从 F 分布。

解 应填 10。$n = 10$，$\overline{X} \sim N(0, 4/10)$，$\dfrac{\sqrt{10}\,\overline{X}}{2} \sim N(0,1)$，故 $\dfrac{10\overline{X}^2}{4} \sim \chi^2(1)$。$\dfrac{9S^2}{4} \sim \chi^2(9)$，且 $\overline{X}$ 与 S^2 独立，故

$$F = \frac{\chi^2(1)/1}{\chi^2(9)/9} = \frac{10\overline{X}^2/4}{S^2/4} = 10\,\frac{\overline{X}^2}{S^2} \sim F(1,9).$$

练习 6.20 设 $X_1, X_2, \ldots, X_8$ 是来自正态总体 $N(0,4)$ 的简单随机样本。常数 $c =$ _____ 时，统计量 $Y = \dfrac{c(X_1 - X_2)}{\sqrt{\sum_{i=3}^8 X_i^2}} \sim$ _____；$P\left\{|\overline{X}| \le \dfrac{\sqrt{2}}{2}\right\} =$ _____。

解 第一空填 $\sqrt{3}$，第二空填 $t(6)$，第三空填 $2\Phi(1) - 1$。

$X_1 - X_2 \sim N(0,8)$，$\dfrac{X_1 - X_2}{\sqrt{8}} \sim N(0,1)$；$\sum_{i=3}^8 X_i^2 = 4\sum_{i=3}^8 \left(\dfrac{X_i}{2}\right)^2$，其中 $\dfrac{X_i}{2} \sim N(0,1)$，故 $\sum_{i=3}^8 X_i^2 \sim 4\chi^2(6)$。

由 t 分布定义：

$$t = \frac{(X_1 - X_2)/\sqrt{8}}{\sqrt{\sum_{i=3}^8 X_i^2/(4\times 6)}} = \sqrt{\frac{24}{8}}\,\frac{X_1 - X_2}{\sqrt{\sum_{i=3}^8 X_i^2}} = \sqrt{3}\,\frac{X_1 - X_2}{\sqrt{\sum_{i=3}^8 X_i^2}} \sim t(6).$$

故 $c = \sqrt{3}$。

概率计算：$\overline{X} \sim N(0, 4/8) = N(0, 1/2)$，$\sqrt{2}\,\overline{X} \sim N(0,1)$。

$$P\left\{|\overline{X}| \le \frac{\sqrt{2}}{2}\right\} = P\left\{-1 \le \sqrt{2}\,\overline{X} \le 1\right\} = 2\Phi(1) - 1.$$

练习 6.21 设总体 $X \sim N(1, \sigma^2)$，$X_1, X_2, \ldots, X_6$ 是取自 X 的简单随机样本，$\overline{X}, S$ 分别为样本均值与样本标准差。求 $P\{\overline{X} > 1\}$ 和 $P\{\overline{X} < 1,\ S^2 < 1.8472\sigma^2\}$。$(\chi^2_{0.05}(5) = 9.236)$

解 $\overline{X} \sim N(1, \sigma^2/6)$。由正态分布关于均值对称：$P\{\overline{X} > 1\} = 0.5$。在正态总体下 $\overline{X}$ 与 S^2 相互独立，故

$$P\{\overline{X} < 1,\ S^2 < 1.8472\sigma^2\} = P\{\overline{X} < 1\}\,P\{S^2 < 1.8472\sigma^2\}.$$

$\dfrac{5S^2}{\sigma^2} \sim \chi^2(5)$，$\chi^2_{0.05}(5) = 9.236$ 表示 $P\{\chi^2(5) > 9.236\} = 0.05$，故

$$P\{S^2 < 1.8472\sigma^2\} = P\left\{\frac{5S^2}{\sigma^2} < 9.236\right\} = 1 - 0.05 = 0.95.$$

结果：$0.5 \times 0.95 = 0.475$。

6.3.2 两个正态总体

定理 6.2 (双正态总体抽样分布)

设总体 $X \sim N(\mu_1, \sigma_1^2)$，$Y \sim N(\mu_2, \sigma_2^2)$，$X_1, \ldots, X_{n_1}$ 和 $Y_1, \ldots, Y_{n_2}$ 分别为来自 X 和 Y 的样本，两组样本相互独立。记 $\overline{X}, S_1^2$ 和 $\overline{Y}, S_2^2$ 分别为两组的样本均值与样本方差，则

1. $\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$；
2. 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时，
$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2),$$
其中 $S_w^2 = \dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ 称为合并样本方差。

练习 6.22 设 $X_1, X_2, \ldots, X_{15}$ 是来自正态总体 $N(0,9)$ 的简单随机样本，则统计量
$$Y = \frac{2X_1^2+X_2^2+\cdots+X_{10}^2}{X_{11}^2+X_{12}^2+\cdots+X_{15}^2}$$
服从（　）。

A. $\chi^2(10)$　B. $t(15)$　C. $F(10,5)$　D. $F(5,10)$

解 应选 C。

$$\chi_1^2 = \frac{2X_1^2+X_2^2+\cdots+X_{10}^2}{9} = \frac{X_1^2}{9}+\frac{X_1^2}{9}+\frac{X_2^2}{9}+\cdots+\frac{X_{10}^2}{9}$$

注意到 $2X_1^2/9 = X_1^2/9 + X_1^2/9$，但实际上更直接的方法是：
$$Y = \frac{(2X_1^2+X_2^2+\cdots+X_{10}^2)/9}{(X_{11}^2+\cdots+X_{15}^2)/9} = \frac{U/10}{V/5},$$
其中 $U = \dfrac{2X_1^2+\cdots+X_{10}^2}{9} \sim \chi^2(10)$（等价于 10 个独立 $N(0,9)$ 变量的平方和除以方差），$V = \dfrac{X_{11}^2+\cdots+X_{15}^2}{9} \sim \chi^2(5)$。

故 $Y = \dfrac{U/10}{V/5} \sim F(10,5)$。

笔记 严格地说，$2X_1^2$ 等价于两个独立的 X_1^2（因为 $N(0,9)$ 变量除以 3 后平方服从 $\chi^2(1)$，$X_1^2/9 \sim \chi^2(1)$，所以 $2X_1^2/9 = 2\chi^2(1)$），因此分子共 10 个自由度。

6.4 分布的分位数

定义 6.8 (上 α 分位数)

设 X 为随机变量，$0<\alpha<1$。若实数 a_α 满足 $P\{X > a_\alpha\} = \alpha$（即上侧尾部面积为 α），则称 a_α 为 X 的上 α 分位数。

三大抽样分布的上 α 分位数分别为：

1. χ^2 分布：$P\{\chi^2 > \chi_\alpha^2(n)\} = \alpha$；
2. t 分布：$P\{t > t_\alpha(n)\} = \alpha$，由对称性 $t_{1-\alpha}(n) = -t_\alpha(n)$；

3. F 分布：$P\{F > F_\alpha(m,n)\} = \alpha$，且 $F_{1-\alpha}(m,n) = \dfrac{1}{F_\alpha(n,m)}$。

笔记 不同教材对分位数的定义可能不同：有的用上侧（$P\{X > a\} = \alpha$），有的用下侧（$P\{X \le a\} = \alpha$）。使用时需注意对应关系：

- 上侧 α 分位数 a_α 对应下侧 $1-\alpha$ 分位数；
- 两种定义下，χ^2 分布满足 $\chi^2_{\alpha,\text{上侧}}(n) = \chi^2_{1-\alpha,\text{下侧}}(n)$。

练习 6.23 设总体 $X \sim N(\mu, 16)$，$X_1, X_2, \ldots, X_{10}$ 是取自 X 的样本，S^2 为样本方差。求常数 $a =$ _____，使 $P(S^2 \ge a) = 0.1$ 成立。

A. $\dfrac{16}{9}\chi^2_{0.1}(9)$　B. $\dfrac{16}{9}\chi^2_{0.05}(10)$　C. $\chi^2_{0.1}(10)$　D. $\chi^2_{0.05}(9)$

解 应选 A。

$\dfrac{9S^2}{16} \sim \chi^2(9)$，故

$$P(S^2 \ge a) = P\left(\frac{9}{16}S^2 \ge \frac{9}{16}a\right) = 0.1 \implies \frac{9}{16}a = \chi^2_{0.1}(9) \implies a = \frac{16}{9}\chi^2_{0.1}(9).$$

练习 6.24 设随机变量 $X \sim \chi^2(2)$，用 $\chi^2_\alpha(2)$ 表示自由度为 2 的 χ^2 分布的上 α 分位数。已知 $P(x < X < y) = 0.95$，$P(X > y) = 0.02$，请用分位数表示 x 和 y。

解 由 $P(X > y) = 0.02$，得 $y = \chi^2_{0.02}(2)$。

由 $P(x < X < y) = 0.95$ 且 $P(X > y) = 0.02$，知 $P(X \le x) = 1 - 0.95 - 0.02 = 0.03$，即 $P(X > x) = 0.97$。

故 $x = \chi^2_{0.97}(2)$。

6.5 综合练习

练习 6.25 设 $X \sim N(0,1)$，$X_1, X_2, \ldots, X_n$ 是取自 X 的样本，$\overline{X}$ 为样本均值，S^2 为样本方差，则（　）。

A. $\overline{X} \sim N(0,1)$　B. $n\overline{X} \sim N(0,1)$

C. $\displaystyle\sum_{i=1}^n X_i^2 \sim \chi^2(n)$　D. $\dfrac{\overline{X}}{S} \sim t(n-1)$

解 应选 C。

A: $\overline{X} \sim N(0, 1/n)$，非 $N(0,1)$ ×。B: $n\overline{X} = \sum X_i \sim N(0,n)$，非 $N(0,1)$ ×。C: 由 χ^2 分布定义直接可得 ✓。D: $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$，此处分子缺少 μ 和 $\sqrt{n}$ ×。

练习 6.26 设 $X_1, X_2, \ldots, X_n$ 是来自总体 $X \sim N(0,1)$ 的简单随机样本，则下列统计量的分布不正确的是（　）。

A. $\displaystyle\sum_{i=1}^n X_i^2 \sim \chi^2(n)$　B. $\dfrac{\sqrt{n-1}\,X_n}{\sqrt{\sum_{i=1}^{n-1} X_i^2}} \sim t(n-1)$

C. $\displaystyle\frac{1}{n}\sum_{i=1}^n X_i \sim N(0,1)$　D. $\dfrac{(n-1)\sum_{i=1}^2 X_i^2}{2\sum_{i=3}^n X_i^2} \sim F(2, n-2)$

解 应选 C。

A: χ^2 分布定义，正确。B: $X_n \sim N(0,1)$，$\sum_{i=1}^{n-1} X_i^2 \sim \chi^2(n-1)$，故 $\dfrac{X_n}{\sqrt{\sum_{i=1}^{n-1} X_i^2/(n-1)}} \sim t(n-1)$，即 B 正确。C: $\dfrac{1}{n}\sum X_i = \overline{X} \sim N(0, 1/n)$，非 $N(0,1)$，错误。D: $\dfrac{\sum_{i=1}^2 X_i^2/2}{\sum_{i=3}^n X_i^2/(n-2)} \sim F(2, n-2)$，即 D 正确。

练习 6.27 设 $X_1, X_2, \ldots, X_n$ 是总体 $N(0,1)$ 的简单随机样本，则 $a =$ _____ 时 $a(X_1 - X_2)^2 \sim$ _____；$\dfrac{2(X_1 - X_2)}{\sqrt{\sum_{i=3}^n X_i^2}} \sim$ _____。

解 第一空填 $\dfrac{1}{2}$，第二空填 $\chi^2(1)$，第三空填 $t(8)$（此时 $n = 10$）。

前两空：$X_1 - X_2 \sim N(0,2)$，$\dfrac{X_1 - X_2}{\sqrt{2}} \sim N(0,1)$，平方后 $\dfrac{(X_1 - X_2)^2}{2} \sim \chi^2(1)$，故 $a = 1/2$。

第三空：$\dfrac{(X_1 - X_2)/\sqrt{2}}{\sqrt{\sum_{i=3}^n X_i^2/(n-2)}} = \sqrt{\dfrac{n-2}{2}}\dfrac{X_1 - X_2}{\sqrt{\sum_{i=3}^n X_i^2}} \sim t(n-2)$。

令 $\sqrt{\dfrac{n-2}{2}} = 2$，得 $n = 10$，自由度为 $10 - 2 = 8$。故 $\dfrac{2(X_1 - X_2)}{\sqrt{\sum_{i=3}^n X_i^2}} \sim t(8)$。

练习 6.28 设总体 $X \sim N(\mu, \sigma^2)$，$X_1, X_2, \ldots, X_n$ 是取自 X 的简单随机样本，$\overline{X}, S^2$ 分别为样本均值和样本方差。记 $T = \overline{X}^2 - \dfrac{1}{n}S^2$。

(1) 证明 $ET = \mu^2$；

(2) 当 $\mu = 0$，$\sigma = 1$ 时，求 DT。

解 (1) $E(\overline{X}) = \mu$，$D(\overline{X}) = \sigma^2/n$，$E(S^2) = \sigma^2$。

$$E(\overline{X}^2) = D(\overline{X}) + [E(\overline{X})]^2 = \frac{\sigma^2}{n} + \mu^2.$$

$$ET = E\left(\overline{X}^2 - \frac{1}{n}S^2\right) = \left(\frac{\sigma^2}{n} + \mu^2\right) - \frac{\sigma^2}{n} = \mu^2.$$

(2) 当 $\mu = 0$，$\sigma = 1$ 时，$X \sim N(0,1)$。$\sqrt{n}\,\overline{X} \sim N(0,1)$，故 $n\overline{X}^2 \sim \chi^2(1)$。$(n-1)S^2 \sim \chi^2(n-1)$。由 $\overline{X}$ 与 S^2 独立：

$$D(\overline{X}^2) = D\left(\frac{\chi^2(1)}{n}\right) = \frac{2}{n^2}, \quad D\left(\frac{S^2}{n}\right) = \frac{1}{n^2}\cdot\frac{2(n-1)}{(n-1)^2} = \frac{2}{n^2(n-1)}.$$

$$DT = D(\overline{X}^2) + D\left(\frac{S^2}{n}\right) = \frac{2}{n^2} + \frac{2}{n^2(n-1)} = \frac{2n}{n^2(n-1)} = \frac{2}{n(n-1)}.$$

第 7 章 点估计

- 点估计
  - 具体内容: 总体分布未知, 寻求适当的统计量估计分布中的未知参数
  - 估计方法
    - 矩估计
      - 思想: 样本矩替换总体矩
      - 掌握: 一阶矩、二阶矩
    - 最大似然估计
      - 思想: 样本似然函数取最大值时对应的未知参数
      - 掌握: 最大似然估计步骤
  - 估计量的评选标准
    - 无偏性: $E(\hat{\theta}) = \theta$
    - 有效性: 若$\hat{\theta}_1, \hat{\theta}_2$ 为θ 的无偏估计量, 且$D(\hat{\theta}_1) \leq D(\hat{\theta}_2)$, 则$\hat{\theta}_1$ 比$\hat{\theta}_2$ 有效
    - 一致性 (相合性): $\lim\limits_{n\to\infty} P\left\{|\hat{\theta} - \theta| < \varepsilon\right\} = 1$, 也可记为$\hat{\theta} \xrightarrow{P} \theta$
    - 均方误差: $\hat{\theta}$ 为θ 的估计量, 称$M(\hat{\theta}) = E(\hat{\theta} - \theta)^2$ 为$\hat{\theta}$ 关于θ 的均方误差,
      若$\hat{\theta}_1, \hat{\theta}_2$ 为θ 的无偏估计量, 且$M(\hat{\theta}_1) \leq M(\hat{\theta}_2)$,
      则$\hat{\theta}_1$ 在均方误差下比$\hat{\theta}_2$ 有效

内容提要

- 点估计的基本概念 7.1
- 矩估计法 7.2
- 最大似然估计法 7.3
- 估计量的评价标准 7.4

7.1 点估计的基本概念

定义 7.1 (点估计)

设总体 X 的分布函数形式已知，但含有一个或多个未知参数 θ。利用总体 X 的一个样本

$$X_1, X_2, \ldots, X_n$$

构造一个适当的统计量

$$\hat{\theta}(X_1, X_2, \ldots, X_n)$$

来估计未知参数 θ，则称其为 θ 的估计量。将样本观测值 $x_1, x_2, \ldots, x_n$ 代入估计量，得到的数值

$$\hat{\theta}(x_1, x_2, \ldots, x_n)$$

称为 θ 的估计值。 ♣

笔记 估计量是随机变量（样本的函数），估计值是具体的数值。常用的点估计方法有矩估计法和最大似然估计法。

7.2 矩估计法（MME）

定义 7.2 (矩)

设 X, Y 为随机变量，k, l 为正整数。

1. 称 $E(X^k)$ 为 X 的k 阶原点矩。
2. 称 $E[(X-EX)^k]$ 为 X 的k 阶中心矩。
3. 称 $E(X^kY^l)$ 为 X 与 Y 的$k+l$ 阶混合原点矩。
4. 称 $E[(X-EX)^k(Y-EY)^l]$ 为 X 与 Y 的$k+l$ 阶混合中心矩。

♣

笔记 从矩的定义可以看出：数学期望 EX 是 X 的一阶原点矩，方差 DX 是 X 的二阶中心矩，协方差 $\mathrm{Cov}(X,Y)$ 是 X 与 Y 的二阶混合中心矩。

矩估计法的核心思想是**用样本矩替换总体矩**：由大数定律，样本矩依概率收敛于总体矩，因此可以用样本矩来估计总体矩，从而解出未知参数。

性质 [矩估计的步骤] 设总体 X 含有 k 个未知参数 $\theta_1, \theta_2, \ldots, \theta_k$，$X_1, X_2, \ldots, X_n$ 为样本。

1. 计算总体的前 k 阶矩 $\mu_j = E(X^j)$ $(j=1,2,\ldots,k)$，它们是 $\theta_1, \ldots, \theta_k$ 的函数。
2. 计算对应的样本矩 $A_j = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^j$ $(j=1,2,\ldots,k)$。
3. 令 $\mu_j = A_j$ $(j=1,2,\ldots,k)$，建立 k 个方程。
4. 解方程组，得到 $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ 作为矩估计量。

笔记 求矩估计时，以低阶为原则——取能解出参数的最小 j 值。例如，若一个参数可用一阶矩解出，就不必使用二阶矩。对于两个参数，通常用 EX 和 EX^2（或等价地 EX 和 DX）建立两个方程。

7.3 最大似然估计法（MLE）

最大似然估计的基本思想：在参数可能的取值范围内，选取使”样本获得当前观测值的概率最大”的参数值作为估计值——”参数 θ 为何值时，观测值出现的概率最大”。

性质 [最大似然估计的步骤] 设总体 X 的分布（离散型为 $p(x;\theta)$，连续型为 $f(x;\theta)$）含有未知参数 θ，样本观测值为 $x_1, x_2, \ldots, x_n$。

1. 写出似然函数：
$$L(\theta) = \prod_{i=1}^{n} p(x_i;\theta) \quad \text{（离散型）} \qquad \text{或} \qquad L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) \quad \text{（连续型）}.$$
2. 取对数：$\ln L(\theta)$（将乘积化为求和，便于求导）。
3. 对 θ 求导并令导数为零，解出 $\hat{\theta}$：
   - 有驻点：令 $\dfrac{\mathrm{d}\ln L}{\mathrm{d}\theta} = 0$，解出 $\hat{\theta}$；
   - 单调型（无驻点）：$\hat{\theta}$ 在参数范围的边界取得（通常是 $\max\{X_i\}$ 或 $\min\{X_i\}$）；
   - 常数型：$\hat{\theta}$ 不唯一。

性质 [最大似然估计的不变性] 设 $\hat{\theta}$ 是 θ 的最大似然估计，$u = u(\theta)$ 具有反函数，则 $\hat{u} = u(\hat{\theta})$ 是 $u(\theta)$ 的最大似然估计。例如，若 $\hat{\theta}$ 是 θ 的 MLE，则 $e^{\hat{\theta}}$ 是 e^{θ} 的 MLE。

✍ **练习 7.1** 为检验某种自然水消毒设备的效果，从消毒后的水中随机抽取 50 升，化验每升水中大肠杆菌的个数（服从泊松分布），结果如下：

大肠杆菌数/升	0	1	2	3	4	5	6
升数	17	20	10	2	1	0	0

问：平均每升水中大肠杆菌的个数为多少时，出现上述情况的概率最大？

解 此题等价于求泊松分布参数 λ 的最大似然估计。由泊松分布的 MLE 为样本均值：

$$\overline{x} = \frac{0\times 17 + 1\times 20 + 2\times 10 + 3\times 2 + 4\times 1}{50} = \frac{50}{50} = 1.$$

故平均每升 1 个时出现上述情况的概率最大。

练习 7.2 设 $(X_1, X_2, \ldots, X_n)$ 为取自总体 $N(0, \sigma^2)$ 的样本，则 σ^2 的最大似然估计是（　）。

A. S^2　　B. $\dfrac{n-1}{n}S^2$　　C. $\overline{X}^2$　　D. $\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2$

解 应选 D。注意此题中 $\mu = 0$ 为已知。似然函数为

$$L(\sigma^2) = \frac{1}{(\sqrt{2\pi\sigma^2})^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right) \qquad \ln L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2.$$

令 $\dfrac{\partial \ln L}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\sum\limits_{i=1}^{n} x_i^2 = 0$，得

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

命题 7.1

在数理统计中，如果一个概率分布（离散或连续均可）的概率质量函数或密度函数可以写成如下形式，它就属于单参数指数族分布：

$$P(X = x; \theta) = h(x)c(\theta)\exp\{w(\theta)t(x)\}$$

这里最关键的是 $t(x)$，它被称为充分统计量（Sufficient Statistic）。在几何分布和上述的二项分布中，我们总能把公式化简，使得 $t(x) = x$。而 $t(x) = x$ 会导致矩估计和最大似然估计相等。

证明

1. 写出对数似然函数：

$$\ln L(\theta) = \sum_{i=1}^{n}\ln P(X_i;\theta) = \sum_{i=1}^{n}\ln h(X_i) + n\ln c(\theta) + w(\theta)\sum_{i=1}^{n} X_i$$

2. 求导寻找最大似然估计（MLE）：将对数似然函数对 θ 求导并令其等于 0：

$$\frac{\mathrm{d}\ln L}{\mathrm{d}\theta} = n\frac{c'(\theta)}{c(\theta)} + w'(\theta)\sum_{i=1}^{n} X_i = 0$$

等式两边同除以 $n\cdot w'(\theta)$，我们得到 MLE 的核心方程：

$$-\frac{c'(\theta)}{c(\theta)w'(\theta)} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

3. 再来看看理论期望（总体一阶矩）：在数理统计中有一个基础定理（正则条件下对数似然导数的期望为 0），即 $E\left[\frac{\mathrm{d}\ln P(X;\theta)}{\mathrm{d}\theta}\right] = 0$。代入单样本的导数 $\frac{c'(\theta)}{c(\theta)} + w'(\theta)X$，取期望得到：

$$\frac{c'(\theta)}{c(\theta)} + w'(\theta)E(X) = 0 \implies E(X) = -\frac{c'(\theta)}{c(\theta)w'(\theta)}$$

4. 见证奇迹的时刻：把期望 $E(X)$ 代回第 2 步的 MLE 核心方程中，MLE 的求解方程竟然变成了：

$$E(X) = \overline{X}$$

而这，恰恰就是一阶矩估计的定义！

结论：只要总体分布属于指数族分布，并且其自然充分统计量就是 X 本身，那么通过最大似然估计求导得到的方程，在数学恒等变换后，一定会变成 $E(X) = \overline{X}$。所以，MLE 和 MME 必然殊途同归，结果完全一致。

练习 7.3 设总体 X 的分布律为 $P(X = x) = p(1-p)^{x-1}$ $(x = 1, 2, \ldots)$，其中 p $(0 < p < 1)$ 未知。$X_1, \ldots, X_n$ 为样本，求 p 的矩估计量和最大似然估计量。

解 几何分布的数学期望

$$E(X) = \sum_{x=1}^{\infty} x \cdot p(1-p)^{x-1} = p\sum_{x=1}^{\infty} x(1-p)^{x-1} = p \cdot \frac{1}{[1-(1-p)]^2} = p \cdot \frac{1}{p^2} = \frac{1}{p}$$

矩估计的核心思想是用样本矩代替总体矩，建立方程求解未知参数。对于单参数问题，我们用一阶原点矩（期望）来求解即：

$$E(X) = \overline{X}, \quad \text{其中}\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \implies \frac{1}{p} = \overline{X} \implies \hat{p}_{\text{矩}} = \frac{1}{\overline{X}}$$

最大似然估计是想找到一个参数值 $\hat{p}$，使得当前样本出现的概率最大（即似然函数取最大值），似然函数是样本的联合概率（离散型），即所有样本观测值概率的乘积：

$$L(p) = \prod_{i=1}^{n} P(X = x_i) = \prod_{i=1}^{n} \left[p(1-p)^{x_i - 1}\right] = p^n \cdot (1-p)^{\left(\sum_{i=1}^{n} x_i\right) - n}$$

$$\ln L(p) = n \ln p + \left(\sum_{i=1}^{n} X_i - n\right) \ln(1-p)$$

$$\frac{\mathrm{d}}{\mathrm{d}p} \ln L(p) = \frac{n}{p} - \frac{\sum_{i=1}^{n} X_i - n}{1-p} = 0 \implies n(1-p) = p\left(\sum_{i=1}^{n} X_i - n\right), \; n = p\sum_{i=1}^{n} X_i$$

故 $\hat{p} = \frac{1}{\overline{X}}$

练习 7.4 设总体 X 的概率分布为 $P\{X = 0\} = \theta^2$，$P\{X = 1\} = 2\theta(1-\theta)$，$P\{X = 2\} = (1-\theta)^2$，其中 θ $(0 < \theta < 1/2)$ 是未知参数。利用样本值 2, 0, 2, 2, 1, 0, 2, 2，求 θ 的矩估计值和最大似然估计值。

解 矩估计：

$$EX = 2\theta(1-\theta) + 2(1-\theta)^2 = 2 - 2\theta.$$

$\overline{x} = \frac{11}{8}$，令 $2 - 2\theta = \frac{11}{8}$，解得 $\hat{\theta}_{\text{矩}} = \frac{5}{16}$。

最大似然估计：样本值中 $X = 0$ 出现 2 次，$X = 1$ 出现 1 次，$X = 2$ 出现 5 次。

$$L(\theta) = \theta^4 \cdot 2\theta(1-\theta) \cdot (1-\theta)^{10} = 2\theta^5(1-\theta)^{11}.$$

$$\ln L = \ln 2 + 5\ln\theta + 11\ln(1-\theta), \quad \frac{\mathrm{d}\ln L}{\mathrm{d}\theta} = \frac{5}{\theta} - \frac{11}{1-\theta} = 0.$$

解得 $\hat{\theta}_{\text{MLE}} = \frac{5}{16}$。

练习 7.5 设总体 X 的分布列如下：

X	1	2	3
P	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

其中 θ 为未知参数，样本观测值为 $X_1 = 1, X_2 = 2, X_3 = 1$。求 θ 的矩估计值和最大似然估计值。

解 矩估计：

$$EX = 1\cdot\theta^2 + 2\cdot 2\theta(1-\theta) + 3(1-\theta)^2 = 3-2\theta.$$

$\overline{x} = \frac{4}{3}$，令 $3-2\theta = \frac{4}{3}$，得 $\hat{\theta}_{矩} = \frac{5}{6}$。

最大似然估计：

$$L(\theta) = \theta^2\cdot 2\theta(1-\theta)\cdot\theta^2 = 2\theta^5(1-\theta).$$

$\ln L = \ln 2 + 5\ln\theta + \ln(1-\theta)$，令 $\dfrac{5}{\theta} - \dfrac{1}{1-\theta} = 0$，得 $\hat{\theta}_{\text{MLE}} = \frac{5}{6}$。

笔记 几何分布和二项分布都属于指数族分布。下面给出一些别的例子。

练习 7.6 设总体 X 的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 θ $(0<\theta<1/2)$ 是未知参数。从总体 X 中抽取容量为 8 的样本，其样本值为 3, 1, 3, 0, 3, 1, 2, 3。求 θ 的矩估计值和最大似然估计值。

解 矩估计：总体仅含一个参数，用一阶矩。

$$EX = 0\cdot\theta^2 + 1\cdot 2\theta(1-\theta) + 2\cdot\theta^2 + 3(1-2\theta) = 3-4\theta.$$

令 $3-4\theta = \overline{X}$，$\overline{x} = \dfrac{3+1+3+0+3+1+2+3}{8} = 2$，故 $\hat{\theta}_{矩} = \dfrac{3-2}{4} = \dfrac{1}{4}$.

最大似然估计：样本值中 $X=0$ 出现 1 次，$X=1$ 出现 2 次，$X=2$ 出现 1 次，$X=3$ 出现 4 次。

$$L(\theta) = \theta^2\cdot[2\theta(1-\theta)]^2\cdot\theta^2\cdot(1-2\theta)^4 = 4\theta^6(1-\theta)^2(1-2\theta)^4.$$

$$\ln L(\theta) = \ln 4 + 6\ln\theta + 2\ln(1-\theta) + 4\ln(1-2\theta).$$

令 $\dfrac{\mathrm{d}\ln L}{\mathrm{d}\theta} = \dfrac{6}{\theta} - \dfrac{2}{1-\theta} - \dfrac{8}{1-2\theta} = 0$，化简得 $12\theta^2 - 14\theta + 3 = 0$。解得 $\theta = \dfrac{14\pm\sqrt{196-144}}{24} = \dfrac{7\pm\sqrt{13}}{12}$

因为 $\dfrac{7+\sqrt{13}}{12} > \dfrac{1}{2}$（不合题意），故

$$\hat{\theta}_{\text{MLE}} = \frac{7-\sqrt{13}}{12}.$$

练习 7.7 设总体 X 的分布列为

X	0	1	2	3
P	$\dfrac{\theta(1-\theta)}{2}$	θ^2	$\dfrac{\theta(1-\theta)}{2}$	$1-\theta$

其中 θ $(0<\theta<1)$ 未知。样本观测值为 0, 0, 1, 2, 2, 3, 3, 3, 0, 2。求 θ 的矩估计值与最大似然估计值。

解 频数统计：$X=0$ 出现 3 次，$X=1$ 出现 1 次，$X=2$ 出现 3 次，$X=3$ 出现 3 次。$\overline{x} = 1.6$。矩估计：

$$EX = \theta^2 + \theta(1-\theta) + 3(1-\theta) = 3-2\theta.$$

令 $3-2\theta = 1.6$，得 $\hat{\theta}_{矩} = 0.7$。

最大似然估计：

$$L(\theta) = \left[\frac{\theta(1-\theta)}{2}\right]^3\theta^2\left[\frac{\theta(1-\theta)}{2}\right]^3(1-\theta)^3 = \frac{1}{64}\theta^8(1-\theta)^9.$$

$\ln L = 8\ln\theta + 9\ln(1-\theta) - \ln 64$，令 $\dfrac{8}{\theta} - \dfrac{9}{1-\theta} = 0$，得 $\hat{\theta}_{\text{MLE}} = \dfrac{8}{17}$。

注 以上是离散型的，我们再来看一下连续型的。

练习 7.8 设总体 X 服从泊松分布 $P(\theta)$，总体 Y 具有密度函数 $f(y,\lambda)=\dfrac{6y}{\lambda^3}(\lambda-y)\ (0<y<\lambda)$。$X_1,\ldots,X_n$ 及 $Y_1,\ldots,Y_n$ 分别为来自 X 和 Y 的样本。

(1) 求 θ 的最大似然估计 $\hat\theta$； (2) 求 λ 的矩估计 $\hat\lambda$； (3) 求 $\hat\lambda$ 的方差。

解 (1) 即 $P(X=x)=e^{-\theta}\dfrac{\theta^x}{x!}$，$\theta>0$，泊松分布的似然函数：

$$L(\theta)=\prod_{i=1}^n\frac{\theta^{x_i}e^{-\theta}}{x_i!}=e^{-n\theta}\,\theta^{\sum x_i}\prod_{i=1}^n\frac{1}{x_i!}.$$

$$\ln L=-n\theta+\sum x_i\ln\theta+C,\quad \frac{\partial\ln L}{\partial\theta}=-n+\frac{\sum x_i}{\theta}=0\implies\hat\theta=\overline{X}$$

(2)

$$EY=\int_0^\lambda y\cdot\frac{6y}{\lambda^3}(\lambda-y)\,\mathrm{d}y=\frac{6}{\lambda^3}\left[\lambda\int_0^\lambda y^2\,\mathrm{d}y-\int_0^\lambda y^3\,\mathrm{d}y\right]=\frac{6}{\lambda^3}\left(\frac{\lambda^4}{3}-\frac{\lambda^4}{4}\right)=\frac{\lambda}{2}$$

令 $\dfrac{\lambda}{2}=\overline{Y}$，得 $\hat\lambda=2\overline{Y}$。

(3)

$$EY^2=\int_0^\lambda\frac{6y^3}{\lambda^3}(\lambda-y)\,\mathrm{d}y=\frac{6}{\lambda^3}\left(\frac{\lambda^5}{4}-\frac{\lambda^5}{5}\right)=\frac{3\lambda^2}{10}$$

$DY=EY^2-(EY)^2=\dfrac{3\lambda^2}{10}-\dfrac{\lambda^2}{4}=\dfrac{\lambda^2}{20}$。$D(\hat\lambda)=D(2\overline{Y})=\dfrac{4}{n}DY=\dfrac{\lambda^2}{5n}$。

笔记 对于单调型似然函数（$\ln L$ 关于 θ 单调），$\hat\theta$ 通常在参数范围的边界取得。常见模式：若 $L(\theta)$ 关于 θ 单调递减，则 $\hat\theta=\max\{X_i\}$；若单调递增，则 $\hat\theta=\min\{X_i\}$。

练习 7.9 单减取最大，单增取最小 设总体 X 的概率密度为

$$f(x;\theta)=\begin{cases}\dfrac{1}{1-\theta}, & \theta\le x\le 1,\\ 0, & \text{其他},\end{cases}$$

其中 θ 为未知参数，$X_1,X_2,\ldots,X_n$ 为简单随机样本。

(1) 求 θ 的矩估计量；

(2) 求 θ 的最大似然估计量。

解 (1) 总体 $X\sim U(\theta,1)$，$EX=\dfrac{1+\theta}{2}$。令 $\dfrac{1+\theta}{2}=\overline{X}$，得 $\hat\theta_{矩}=2\overline{X}-1$。

(2) 似然函数为[1]

$$L(\theta)=\frac{1}{(1-\theta)^n},\quad \text{约束条件：}\theta\le\min\{x_1,\ldots,x_n\},\ x_i\le 1.$$

$L(\theta)$ 关于 θ 单调递增（分母 $1-\theta$ 递减），故当 $\theta=\min\{X_1,\ldots,X_n\}$ 时 L 最大：

$$\hat\theta_{\mathrm{MLE}}=\min\{X_1,X_2,\ldots,X_n\}.$$

练习 7.10 设某种产品的使用寿命 T 的分布函数 $F(t)$ 满足方程

$$F'(t)+\frac{2t}{\theta^2}[F(t)-1]=0,\quad t\ge 0,\quad F(0)=0,$$

其中 $\theta>1$ 为未知参数。产品性能 $Q(\theta)=\dfrac{\theta^2}{2}\ln\theta-\dfrac{3}{4}\theta^2+\theta$。

(1) 求 $P\{T>t\}$ 与 $P\{T>s+t\mid T>s\}$（$s,t>0$）；

(2) 任取 n 个产品做寿命试验，测得寿命 $t_1,\ldots,t_n$，求 θ 的最大似然估计值 $\hat\theta$；

[1]密度为 0 的区间直接被过滤了，否则整个乘积为 0，似然函数取最小值，不可能对应最大似然估计。

(3) 求 Q 的最大似然估计值 $\hat{Q}$。

解 先求 $F(t)$。微分方程 $F'(t)+\dfrac{2t}{\theta^2}F(t)=\dfrac{2t}{\theta^2}$ 是一阶线性 ODE，积分因子为 e^{t^2/θ^2}：

$$F(t)\,e^{t^2/\theta^2}=\int e^{t^2/\theta^2}\,\mathrm{d}\left(\frac{t^2}{\theta^2}\right)+C=e^{t^2/\theta^2}+C.$$

由 $F(0)=0$ 得 $C=-1$，故

$$F(t)=1-e^{-t^2/\theta^2}\quad(t\geq 0).$$

(1) $P\{T>t\}=e^{-t^2/\theta^2}$。

$$P\{T>s+t\mid T>s\}=\frac{P\{T>s+t\}}{P\{T>s\}}=\frac{e^{-(s+t)^2/\theta^2}}{e^{-s^2/\theta^2}}=e^{-(t^2+2st)/\theta^2}.$$

(2) 密度 $f(t;\theta)=F'(t)=\dfrac{2t}{\theta^2}\,e^{-t^2/\theta^2}\ (t>0)$。

$$L(\theta)=\theta^{-2n}\cdot 2^n\left(\prod_{i=1}^n t_i\right)e^{-\sum t_i^2/\theta^2},\quad \ln L=n\ln 2-2n\ln\theta+\sum\ln t_i-\frac{1}{\theta^2}\sum t_i^2.$$

令 $\dfrac{\mathrm{d}\ln L}{\mathrm{d}\theta}=-\dfrac{2n}{\theta}+\dfrac{2}{\theta^3}\sum t_i^2=0$，得 $\hat{\theta}=\left(\dfrac{1}{n}\sum\limits_{i=1}^n t_i^2\right)^{1/2}$。

(3) 由不变性原则，$Q'(\theta)=\theta\ln\theta-\theta+1>0$，故 $Q(\theta)$ 有反函数。

$$\hat{Q}=Q(\hat{\theta})=\frac{\hat{\theta}^2}{2}\ln\hat{\theta}-\frac{3}{4}\hat{\theta}^2+\hat{\theta}.$$

7.4 估计量的评价标准

定义 7.3 (无偏性)

若 $E\hat{\theta}=\theta$，则称 $\hat{\theta}$ 为 θ 的无偏估计量。 ♣

命题 7.2 (系数和为 1)

若估计量形式为：

$$\hat{\mu}=a_1X_1+a_2X_2+\cdots+a_nX_n$$

则

$$\hat{\mu}\text{ 是}\mu\text{ 的无偏估计}\iff a_1+a_2+\cdots+a_n=1$$

♠

证明

$$E(\hat{\mu})=E\left(\sum_{i=1}^n a_iX_i\right)=\sum_{i=1}^n a_i\,EX_i=\sum_{i=1}^n a_i\,\mu=\left(\sum_{i=1}^n a_i\right)\mu.$$

故 $E(\hat{\mu})=\mu\iff\sum\limits_{i=1}^n a_i=1$。

✍ **练习 7.11** 设 X_1,X_2,X_3 是总体的样本，则下列统计量不是总体均值 μ 的无偏估计量的是（　）。

A. $\dfrac{1}{3}(X_1+X_2+X_3)$　B. $\dfrac{1}{3}X_1+\dfrac{1}{6}X_2+\dfrac{1}{2}X_3$　C. $\dfrac{1}{6}X_1+\dfrac{1}{3}X_2+\dfrac{1}{2}X_3$　D. $\dfrac{1}{6}X_1+\dfrac{1}{6}X_2+\dfrac{1}{3}X_3$

解 应选 D。无偏估计要求系数之和为 1。

✍ **练习 7.12** 设 $X_1,X_2,\ldots,X_n$ 是来自分布 $f(x;\theta)=\dfrac{3x^2}{\theta^3},0\leq x\leq\theta$ 的简单随机样本。若统计量 $g(X_1,\ldots,X_n)=c\overline{X}$ 为 θ 的无偏估计，则常数 c 为？

解 应填 $\frac{4}{3}$。$EX=\frac{3\theta}{4}$，$E(c\overline{X})=c\cdot\frac{3\theta}{4}=\theta$，故 $c=\frac{4}{3}$。

练习 7.13 设总体 X 具有概率密度

$$f(x,\theta)=\begin{cases}\frac{1}{\theta}\,e^{-x/\theta}, & x>0\ (\theta>0),\\ 0, & x\le 0,\end{cases}$$

其中 θ 未知，$X_1,\ldots,X_n$ 为样本。求 θ 的最大似然估计 $\hat\theta$，并证明 $\hat\theta$ 是 θ 的无偏估计。

解 似然函数 $L(\theta)=\theta^{-n}e^{-\sum X_i/\theta}$，$\ln L=-n\ln\theta-\sum X_i/\theta$。令 $-n/\theta+\sum X_i/\theta^2=0$，得 $\hat\theta=\overline{X}$。指数分布 $EX=\theta$，故 $E(\hat\theta)=E(\overline{X})=\theta$，$\hat\theta$ 是 θ 的无偏估计。

练习 7.14 设 $X_1,X_2,\ldots,X_n$ 是来自正态总体 $X\sim N(\mu,\sigma^2)$ 的简单随机样本，为使

$$D=k\sum_{i=1}^{n-1}(X_{i+1}-X_i)^2$$

成为总体方差 σ^2 的无偏估计量，则 $k=$ _____。

解 $E(X_{i+1}-X_i)^2=E(X_{i+1}^2)+E(X_i^2)-2E(X_i)E(X_{i+1})=2(\sigma^2+\mu^2)-2\mu^2=2\sigma^2$。

$ED=k\cdot(n-1)\cdot 2\sigma^2=2k(n-1)\sigma^2$。令 $ED=\sigma^2$，得 $k=\frac{1}{2(n-1)}$。

练习 7.15 设 $X_1,\ldots,X_m$ 是取自总体 $N(1,\sigma^2)$ 的简单随机样本，$Y_1,\ldots,Y_n$ 是取自总体 $N(2,\sigma^2)$ 的简单随机样本。

(1) 求联合密度函数；

(2) 求 σ^2 的最大似然估计 $\hat\sigma^2$；

(3) $\hat\sigma^2$ 是否为 σ^2 的无偏估计？

解 (1) 设 $T=\sum_{i=1}^{m}(X_i-1)^2+\sum_{j=1}^{n}(Y_j-2)^2$，则

$$f=(2\pi\sigma^2)^{-(m+n)/2}\exp\left\{-\frac{T}{2\sigma^2}\right\}.$$

(2) $\ln f=-\frac{m+n}{2}\ln(2\pi\sigma^2)-\frac{T}{2\sigma^2}$，令 $\frac{\partial\ln f}{\partial\sigma^2}=-\frac{m+n}{2\sigma^2}+\frac{T}{2(\sigma^2)^2}=0$，得

$$\hat\sigma^2=\frac{1}{m+n}\left[\sum_{i=1}^{m}(X_i-1)^2+\sum_{j=1}^{n}(Y_j-2)^2\right].$$

(3) 是。$E\left[\sum(X_i-1)^2\right]=m\sigma^2$，$E\left[\sum(Y_j-2)^2\right]=n\sigma^2$，故

$$E(\hat\sigma^2)=\frac{(m+n)\sigma^2}{m+n}=\sigma^2.$$

练习 7.16 设总体 X 的分布函数为

$$F(x,\alpha,\beta)=\begin{cases}1-\left(\frac{\alpha}{x}\right)^{\beta}, & x>\alpha,\\ 0, & x\le\alpha,\end{cases}$$

其中 $\alpha>0$，$\beta>1$，$X_1,\ldots,X_n$ 为样本。

(1) 当 $\alpha=1$ 时，求 β 的矩估计和最大似然估计；

(2) 当 $\beta=2$ 时，求 α 的最大似然估计并讨论它的无偏性。

解 密度 $f(x)=\beta\,\alpha^{\beta}\,x^{-(\beta+1)}\ (x>\alpha)$。

(1) $\alpha=1$ 时，$f(x)=\beta\,x^{-(\beta+1)}\ (x>1)$。

$$EX=\int_1^{\infty}x\cdot\beta\,x^{-(\beta+1)}\,\mathrm{d}x=\frac{\beta}{\beta-1}.$$

矩估计：令 $\dfrac{\beta}{\beta-1}=\overline{X}$，得 $\hat{\beta}_{矩}=\dfrac{\overline{X}}{\overline{X}-1}$。

似然函数

$$L(\beta)=\beta^n\left(\prod X_i\right)^{-(\beta+1)},\quad \ln L=n\ln\beta-(\beta+1)\sum\ln X_i$$

令 $n/\beta-\sum\ln X_i=0$，得 $\hat{\beta}_{\text{MLE}}=\dfrac{n}{\sum\ln X_i}$。

(2) $\beta=2$ 时，$f(x)=2\alpha^2x^{-3}\ (x>\alpha)$。约束 $\alpha\le X_{(1)}$。

$\ln L=n\ln 2+2n\ln\alpha-3\sum\ln X_i$，$\dfrac{\partial\ln L}{\partial\alpha}=\dfrac{2n}{\alpha}>0$，故 L 关于 α 递增，$\hat{\alpha}=X_{(1)}$。

对于 Pareto(α,2)，$E(X_{(1)})=\dfrac{2n\alpha}{2n-1}\ne\alpha$，故 $\hat{\alpha}$ 是有偏估计。

练习 7.17 设总体 X 的概率密度为 $f(x,\theta)=\dfrac{3x^2}{\theta^3}\ (0\le x\le\theta)$，其中 $\theta>0$ 未知，$X_1,\ldots,X_n$ 为简单随机样本。

(1) 求 θ 的矩估计量 $\hat{\theta}_1$ 和最大似然估计量 $\hat{\theta}_2$；

(2) 判断 $\hat{\theta}_2$ 是否是 θ 的无偏估计量。

解 (1) $EX=\displaystyle\int_0^\theta\frac{3x^3}{\theta^3}\,\mathrm{d}x=\frac{3\theta}{4}$。令 $\dfrac{3\theta}{4}=\overline{X}$，得 $\hat{\theta}_1=\dfrac{4}{3}\overline{X}$。似然函数 $L(\theta)=\dfrac{3^n\prod x_i^2}{\theta^{3n}}$ 关于 θ 单减，故

$$\hat{\theta}_2=\max\{X_1,X_2,\ldots,X_n\}.$$

(2) 由于样本同分布，$P(X_i\le x)=P(X\le x)=F_X(x)$，对于样本最大值，分布函数公式为：

$$\begin{aligned}F_{\max}(x)&=P\left(\max\{X_1,X_2,\ldots,X_n\}\le x\right)=P\left(X_1\le x,\ X_2\le x,\ \ldots,\ X_n\le x\right)\\&=\prod_{i=1}^nP(X_i\le x)=\prod_{i=1}^nF_X(x)=[F_X(x)]^n\end{aligned}$$

求总体 X 的分布函数 $F_X(x)$

$$F_X(x)=\begin{cases}0, & x<0\\ \displaystyle\int_0^x\frac{3t^2}{\theta^3}dt=\frac{x^3}{\theta^3}, & 0\le x\le\theta\\ 1, & x>\theta\end{cases}$$

代入最大值分布公式，当 $0\le x\le\theta$ 时：

$$F_{\hat{\theta}_2}(x)=(F_X(x))^n=\left(\frac{x^3}{\theta^3}\right)^n=\left(\frac{x}{\theta}\right)^{3n}$$

当 $x<0$ 时 $F=0$，$x>\theta$ 时 $F=1$。密度是分布函数的导数：

$$f_{\hat{\theta}_2}(x)=F'_{\hat{\theta}_2}(x)=\begin{cases}\dfrac{3nx^{3n-1}}{\theta^{3n}}, & 0\le x\le\theta\\ 0, & 其他\end{cases}$$

根据期望公式 $E(\hat{\theta}_2)=\int_{-\infty}^{+\infty}xf_{\hat{\theta}_2}(x)dx$：

$$E(\hat{\theta}_2)=\int_0^\theta x\cdot\frac{3nx^{3n-1}}{\theta^{3n}}dx=\frac{3n}{\theta^{3n}}\int_0^\theta x^{3n}dx=\frac{3n}{\theta^{3n}}\cdot\frac{x^{3n+1}}{3n+1}\bigg|_0^\theta=\frac{3n}{\theta^{3n}}\cdot\frac{\theta^{3n+1}}{3n+1}=\frac{3n}{3n+1}\theta$$

故 $\hat{\theta}_2$ 是有偏估计量。

笔记 作为对比，$\hat{\theta}_1=\dfrac{4}{3}\overline{X}$ 满足 $E(\hat{\theta}_1)=\dfrac{4}{3}\cdot\dfrac{3\theta}{4}=\theta$，是无偏的。

练习 7.18 设总体 X 的密度函数为

$$f_X(x,\theta)=\begin{cases}4\,e^{-4(x-\theta)}, & x\geq\theta,\\ 0, & \text{其他},\end{cases}$$

$X_1,\ldots,X_n$ 为来自总体 X 的样本。

(1) 求 θ 的最大似然估计 $\hat{\theta}$;

(2) $\hat{\theta}$ 是否为无偏估计？如果不是，修正为无偏估计。

解 (1) 构造似然函数：

$$L(\theta)=\prod_{i=1}^{n}4e^{-4(X_i-\theta)}=4^n e^{-4\sum_{i=1}^{n}X_i+4n\theta},\quad \theta\leq X_{(1)}=\min\{X_i\}$$

$L(\theta)$ 关于 θ 严格单调递增，因此在定义域右端点取最大值，故最大似然估计量：

$$\hat{\theta}=X_{(1)}=\min\{X_1,X_2,\ldots,X_n\}$$

(2) 我们先做一个变量平移：令 $Y_i=X_i-\theta$，则当 $X_i\geq\theta$ 时，$Y_i\geq 0$，代入密度函数：

$$f_{Y_i}(y)=f_X(y+\theta,\theta)=4e^{-4y},\quad y\geq 0$$

这正是参数为 $\lambda=4$ 的指数分布：$Y_i\sim\text{Exp}(4)$，且 $Y_1,Y_2,\ldots,Y_n$ 独立同分布。样本最小值 $X_{(1)}=\min\{X_i\}=\min\{Y_i+\theta\}=\theta+\min\{Y_i\}$，因此：

$$X_{(1)}-\theta=\min\{Y_1,Y_2,\ldots,Y_n\}$$

即：$\boldsymbol{X_{(1)}-\theta}$ 是 n 个独立 Exp(4) 变量的最小值。单个 Exp(4) 的分布函数：$F_Y(y)=1-e^{-4y}$, $y\geq 0$，n 个独立变量的最小值的分布函数：

$$F_{\min Y}(y)=P(\min Y_i\leq y)=1-P(\min Y_i>y)=1-[1-F_Y(y)]^n=1-e^{-4ny}$$

这正是参数为 $\lambda=4n$ 的指数分布：$\min Y_i\sim\text{Exp}(4n)$。根据指数分布的期望公式：若 $Z\sim\text{Exp}(\lambda)$，则 $E(Z)=\frac{1}{\lambda}$。因此：

$$E\left(\min Y_i\right)=\frac{1}{4n}$$

由 $X_{(1)}=\theta+\min Y_i$，两边取期望：

$$E(\hat{\theta})=E(X_{(1)})=E\left(\theta+\min Y_i\right)=\theta+E\left(\min Y_i\right)=\theta+\frac{1}{4n}\neq\theta$$

因此 $\hat{\theta}=X_{(1)}$ 是有偏估计量，偏差为 $\frac{1}{4n}$，我们只需要把偏差减掉，构造新的估计量：

$$\hat{\theta}_{\text{unbiased}}=X_{(1)}-\frac{1}{4n}$$

验证无偏性：

$$E\left(\hat{\theta}_{\text{unbiased}}\right)=E\left(X_{(1)}-\frac{1}{4n}\right)=\left(\theta+\frac{1}{4n}\right)-\frac{1}{4n}=\theta$$

修正后的估计量是 θ 的无偏估计。

笔记 题目中的 $f_X(x,\theta)$ 是位置参数型指数分布（也叫平移指数分布）：标准指数分布 Exp(λ) 是 $f(x)=\lambda e^{-\lambda x}$, $x\geq 0$；平移 θ 后，分布变为 $f(x)=\lambda e^{-\lambda(x-\theta)}$, $x\geq\theta$，θ 是位置参数（分布的左端点）。对于这类位置参数、支撑集为 $[\theta,+\infty)$ 的分布，极大似然估计一定是样本最小值，这是通用结论。

命题 7.3 (样本方差的无偏性)

样本方差 S^2 是总体方差 σ^2 的无偏估计，这个结论对“任意”具有有限方差的总体分布都成立

证明 首先给出样本方差的定义

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

令 $Y_i=X_i-\mu$，则总体满足 $Y_i\sim N(0,\sigma^2)$，样本均值 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，偏差平方和为：

$$\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=1}^{n}[(X_i-\mu)-(\overline{X}-\mu)]^2=\sum_{i=1}^{n}(Y_i-\overline{Y})^2$$

其中 $\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i$，展开：

$$\begin{aligned}\sum_{i=1}^{n}(Y_i-\overline{Y})^2&=\sum_{i=1}^{n}(Y_i^2-2Y_i\overline{Y}+\overline{Y}^2)=\sum Y_i^2-2\overline{Y}\sum Y_i+n\overline{Y}^2\\&=\sum Y_i^2-2\overline{Y}(n\overline{Y})+n\overline{Y}^2=\sum Y_i^2-2n\overline{Y}^2+n\overline{Y}^2=\sum Y_i^2-n\overline{Y}^2\end{aligned}$$

所以

$$\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=1}^{n}Y_i^2-n\overline{Y}^2$$

求期望：

$$\begin{aligned}E\left(\sum(X_i-\overline{X})^2\right)&=E\left(\sum Y_i^2-n\overline{Y}^2\right)=E(\sum Y_i^2)-E(n\overline{Y}^2)\\&=\sum_{i=1}^{n}E(Y_i^2)-E(n\overline{Y}^2)=nE(Y_i^2)-E(n\overline{Y}^2)\end{aligned}$$

由于

$$\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i\sim N\left(0,\frac{\sigma^2}{n}\right)$$

所以

$$E(\overline{Y}^2)=D(\overline{Y})+[E(\overline{Y})]^2=\frac{\sigma^2}{n}+0=\frac{\sigma^2}{n}\implies E(n\overline{Y}^2)=n\cdot\frac{\sigma^2}{n}=\sigma^2$$

所以

$$nE(Y_i^2)-E(n\overline{Y}^2)=n\sigma^2-\sigma^2=(n-1)\sigma^2$$

又

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

两边取期望：

$$E(S^2)=E\left(\frac{1}{n-1}\sum(X_i-\overline{X})^2\right)=\frac{1}{n-1}E\left(\sum(X_i-\overline{X})^2\right)=\frac{1}{n-1}\cdot(n-1)\sigma^2=\sigma^2$$

练习 7.19 设 $X_1,\ldots,X_n$ 和 $Y_1,\ldots,Y_m$ 分别是来自总体 $X\sim N(\mu_1,\sigma^2)$ 和 $Y\sim N(\mu_2,\sigma^2)$ 的简单随机样本，样本方差分别为 S_X^2,S_Y^2。则 σ^2 的无偏估计是（　）。

A. $S_X^2+S_Y^2$　B. $(m-1)S_X^2+(n-1)S_Y^2$　C. $\dfrac{S_X^2+S_Y^2}{m+n-2}$　D. $\dfrac{(m-1)S_X^2+(n-1)S_Y^2}{m+n-2}$

解 应选 D。$E(S_X^2)=E(S_Y^2)=\sigma^2$。选项 D 中：

$$E\left(\frac{(m-1)S_X^2+(n-1)S_Y^2}{m+n-2}\right)=\frac{(m-1)\sigma^2+(n-1)\sigma^2}{m+n-2}=\sigma^2.$$

命题 7.4 (泊松分布无偏估计的不唯一性)

设 $X_1, X_2, \ldots, X_n$ 是来自泊松总体 $P(\lambda)$ 的简单随机样本，那么样本均值 $\overline{X}$ 和样本方差 S^2 都是未知参数 λ 的无偏估计。

证明 根据泊松分布的性质，总体期望和总体方差相等，均等于参数 λ。即：

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

对于样本均值 $\overline{X}$，利用期望的线性性质：

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\cdot n\lambda = \lambda$$

因此，$\overline{X}$ 是 λ 的无偏估计。而样本方差 S^2 总是总体方差的无偏估计：

$$E(S^2) = D(X) = \lambda$$

因此，S^2 也是 λ 的无偏估计。（注：虽然两者都是无偏估计，但在实际应用中我们通常使用 $\overline{X}$，因为可以进一步证明 $\overline{X}$ 的方差比 S^2 的方差更小，即 $\overline{X}$ 是更有效的估计量）。

练习 7.20 设 $X_1, X_2, \ldots, X_n$ 是取自总体 $P(\lambda)$ 的样本，$\overline{X}$ 为样本均值，则 λ^2 的一个无偏估计为（　）。

A. $\overline{X}^2 - \overline{X}$　　B. $\overline{X}^2 + \overline{X}$　　C. $\overline{X}^2 - \frac{1}{n}\overline{X}$　　D. $\overline{X}^2 + \frac{1}{n}\overline{X}$

解 应选 C。泊松分布 $EX = \lambda$，$DX = \lambda$。

$$E(\overline{X}^2) = D(\overline{X}) + (E\overline{X})^2 = \frac{\lambda}{n} + \lambda^2, \quad E\left(\overline{X}^2 - \frac{1}{n}\overline{X}\right) = \left(\frac{\lambda}{n} + \lambda^2\right) - \frac{\lambda}{n} = \lambda^2$$

命题 7.5 (均匀分布参数的无偏估计)

设 $X_1, X_2, \ldots, X_n$ 是来自均匀总体 $U(0, \theta)$ 的简单随机样本，次序统计量最大值

$$X_{(n)} = \max(X_1, \ldots, X_n)$$

是 θ 的极大似然估计量，但它是有偏的。经过修正后的统计量 $T = \frac{n+1}{n}X_{(n)}$ 才是 θ 的无偏估计。

证明 首先求出最大值 $X_{(n)}$ 的分布。因为样本相互独立，所以 $X_{(n)} \le x$ 等价于所有的 $X_i \le x$：

$$F_{X_{(n)}}(x) = P(X_{(n)} \le x) = \prod_{i=1}^{n} P(X_i \le x) = \left(\frac{x}{\theta}\right)^n \quad (0 < x < \theta)$$

对其求导，得到 $X_{(n)}$ 的概率密度函数：

$$f_{X_{(n)}}(x) = \frac{n}{\theta^n}x^{n-1} \quad (0 < x < \theta)$$

接下来求 $X_{(n)}$ 的数学期望：

$$E(X_{(n)}) = \int_0^\theta x \cdot \frac{n}{\theta^n}x^{n-1}dx = \frac{n}{\theta^n}\int_0^\theta x^n dx = \frac{n}{\theta^n}\cdot\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta$$

由于 $E(X_{(n)}) \ne \theta$，所以极大似然估计量 $X_{(n)}$ 是有偏的（它总是倾向于低估总体的上限 θ）。为了消除偏差，我们在等式两边同乘 $\frac{n+1}{n}$：

$$E\left(\frac{n+1}{n}X_{(n)}\right) = \frac{n+1}{n}\cdot\frac{n}{n+1}\theta = \theta$$

因此，修正后的估计量 $T = \frac{n+1}{n}X_{(n)}$ 是总体参数 θ 的无偏估计量。

定义 7.4 (有效性)

若 $\hat{\theta}_1, \hat{\theta}_2$ 都是 θ 的无偏估计量，且 $D\hat{\theta}_1 < D\hat{\theta}_2$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。 ♣

练习 7.21 设总体 $X \sim N(\mu, \sigma^2)$，X_1, X_2, X_3 是来自总体 X 的简单随机样本，记

$$\hat{\mu}_1 = \frac{1}{5}X_1 + \frac{3}{10}X_2 + \frac{1}{2}X_3, \quad \hat{\mu}_2 = \frac{1}{3}X_1 + \frac{1}{4}X_2 + \frac{5}{12}X_3, \quad \hat{\mu}_3 = \frac{1}{3}X_1 + \frac{1}{6}X_2 + \frac{1}{2}X_3.$$

证明 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 都是 μ 的无偏估计，并指出哪一个最有效。

解 验证无偏性（系数之和均为 1）：

$$E\hat{\mu}_1 = \left(\frac{1}{5} + \frac{3}{10} + \frac{1}{2}\right)\mu = E\hat{\mu}_2 = \left(\frac{1}{3} + \frac{1}{4} + \frac{5}{12}\right)\mu = E\hat{\mu}_3 = \left(\frac{1}{3} + \frac{1}{6} + \frac{1}{2}\right)\mu = \mu$$

比较方差：

$$D\hat{\mu}_1 = \left(\frac{1}{25} + \frac{9}{100} + \frac{1}{4}\right)\sigma^2 = \frac{38}{100}\sigma^2, \quad D\hat{\mu}_2 = \left(\frac{1}{9} + \frac{1}{16} + \frac{25}{144}\right)\sigma^2 = \frac{50}{144}\sigma^2,$$

$$D\hat{\mu}_3 = \left(\frac{1}{9} + \frac{1}{36} + \frac{1}{4}\right)\sigma^2 = \frac{14}{36}\sigma^2.$$

换算：$\frac{38}{100} = 0.380$，$\frac{50}{144} \approx 0.347$，$\frac{14}{36} \approx 0.389$。$\frac{50}{144} < \frac{38}{100} < \frac{14}{36}$，故 $\hat{\mu}_2$ 最有效。

练习 7.22 设总体 $X \sim N(0, 1)$，X_1, X_2 是从总体抽取的一个样本，下列四个无偏估计量中最有效的是（　）。

A. $\mu_1 = X_1$　　B. $\mu_2 = \frac{2}{3}X_1 + \frac{1}{3}X_2$　　C. $\mu_3 = \frac{1}{4}X_1 + \frac{3}{4}X_2$　　D. $\mu_4 = \frac{1}{2}X_1 + \frac{1}{2}X_2$

解 应选 D。$D\mu_1 = 1$，$D\mu_2 = \frac{4}{9} + \frac{1}{9} = \frac{5}{9}$，$D\mu_3 = \frac{1}{16} + \frac{9}{16} = \frac{5}{8}$，$D\mu_4 = \frac{1}{2}$。

$\frac{1}{2} < \frac{5}{9} < \frac{5}{8} < 1$，故 μ_4 最有效。

练习 7.23 设样本 X_1, X_2, X_3 来自总体 $U(0, \theta)$，则估计量 $\hat{\theta}_1 = 2\overline{X}$，$\hat{\theta}_2 = X_1 + X_2$，$\hat{\theta}_3 = \frac{2}{5}X_1 + \frac{4}{5}(X_2 + X_3)$ 中，哪一个较优？

解 三个估计量均无偏（系数之和均为 1）。$DX = \theta^2/12$。

$$D\hat{\theta}_1 = 4D(\overline{X}) = \frac{4\theta^2}{12 \cdot 3} = \frac{\theta^2}{9}, \quad D\hat{\theta}_2 = 2 \cdot \frac{\theta^2}{12} = \frac{\theta^2}{6},$$

$$D\hat{\theta}_3 = \left(\frac{4}{25} + \frac{16}{25} + \frac{16}{25}\right)\frac{\theta^2}{12} = \frac{36\theta^2}{300} = \frac{3\theta^2}{25}.$$

$\frac{\theta^2}{9} < \frac{3\theta^2}{25} < \frac{\theta^2}{6}$，故 $\hat{\theta}_1 = 2\overline{X}$ 最优（方差最小）。

定义 7.5 (一致性（相合性）)

若 $\hat{\theta} \xrightarrow{P} \theta$（即对任意 $\varepsilon > 0$，$\lim_{n\to\infty} P\{|\hat{\theta} - \theta| < \varepsilon\} = 1$），则称 $\hat{\theta}$ 为 θ 的一致估计量（或相合估计量）。 ♣

笔记 验证一致性常用的两种方法：

1. 切比雪夫不等式：若 $D\hat{\theta} \to 0$（当 $n \to \infty$），则 $P\{|\hat{\theta} - \theta| \geq \varepsilon\} \leq D\hat{\theta}/\varepsilon^2 \to 0$。
2. 辛钦大数定律：若 $\hat{\theta}$ 是样本值的函数且可表示为独立同分布随机变量的均值，则可直接得出依概率收敛。

练习 7.24 设总体 X 的概率密度为 $f(x) = \frac{1}{2\theta}e^{-|x|/\theta}$ $(-\infty < x < \infty)$，其中 $\theta > 0$ 未知，$X_1, \ldots, X_n$ 为样本。

(1) 求 θ 的最大似然估计量 $\hat{\theta}$；

(2) 证明 $\hat{\theta}$ 是 θ 的无偏估计量；

(3) 证明 $\hat{\theta}$ 是 θ 的一致估计量。

解 (1) $L(\theta) = (2\theta)^{-n} e^{-\sum |X_i|/\theta}$，$\ln L = -n\ln(2\theta) - \sum |X_i|/\theta$。

令 $-n/\theta + \sum |X_i|/\theta^2 = 0$，得 $\hat{\theta} = \dfrac{1}{n}\sum_{i=1}^n |X_i|$。

(2) Laplace 分布的 $E|X| = \theta$，故 $E(\hat{\theta}) = \theta$ ✓。

(3) $|X_1|, |X_2|, \ldots, |X_n|$ 独立同分布，$E|X_i| = \theta$。由辛钦大数定律：

$$\hat{\theta} = \frac{1}{n}\sum |X_i| \xrightarrow{P} E|X| = \theta.$$

练习 7.25 设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的简单随机样本，$EX = \mu$，$DX = \sigma^2$。证明统计量

$$Y = \frac{2}{n(n+1)}\sum_{i=1}^n iX_i$$

是 μ 的无偏相合估计量。

解 无偏性：

$$EY = \frac{2}{n(n+1)}\sum_{i=1}^n i\,EX_i = \frac{2\mu}{n(n+1)}\cdot\frac{n(n+1)}{2} = \mu.$$

一致性：由 $\sum i^2 = \dfrac{n(n+1)(2n+1)}{6}$，

$$DY = \left[\frac{2}{n(n+1)}\right]^2 \sigma^2 \sum_{i=1}^n i^2 = \frac{4\sigma^2}{n^2(n+1)^2}\cdot\frac{n(n+1)(2n+1)}{6} = \frac{2(2n+1)}{3n(n+1)}\sigma^2.$$

由切比雪夫不等式，对任意 $\varepsilon > 0$：

$$P\{|Y-\mu| < \varepsilon\} \geq 1 - \frac{DY}{\varepsilon^2} = 1 - \frac{2(2n+1)\sigma^2}{3n(n+1)\varepsilon^2}.$$

令 $n \to \infty$，右端 $\to 1$，故 $Y \xrightarrow{P} \mu$，Y 是 μ 的相合估计量。

定义 7.6 (均方误差)

设 $\hat{\theta}$ 为参数 θ 的估计量，称

$$\mathrm{MSE}(\hat{\theta}) = E(\hat{\theta} - \theta)^2$$

为 $\hat{\theta}$ 关于 θ 的均方误差（Mean Squared Error, MSE）。 ♣

性质 [均方误差的分解] 均方误差可分解为偏差的平方与方差之和：

$$\mathrm{MSE}(\hat{\theta}) = [E(\hat{\theta}) - \theta]^2 + D(\hat{\theta}) = \mathrm{Bias}^2(\hat{\theta}) + D(\hat{\theta}).$$

证明

$$E(\hat{\theta} - \theta)^2 = E\Big[(\hat{\theta} - E\hat{\theta}) + (E\hat{\theta} - \theta)\Big]^2 = D(\hat{\theta}) + [E(\hat{\theta}) - \theta]^2.$$

交叉项 $2E[(\hat{\theta} - E\hat{\theta})(E\hat{\theta} - \theta)] = 0$。

笔记 均方误差综合考虑了估计量的偏差和方差：

- 若 $\hat{\theta}$ 是无偏估计（$E\hat{\theta} = \theta$），则 $\mathrm{MSE}(\hat{\theta}) = D(\hat{\theta})$，此时均方误差最小等价于方差最小（即有效性）。
- 若 $\hat{\theta}_1, \hat{\theta}_2$ 为 θ 的估计量（不一定无偏），且 $\mathrm{MSE}(\hat{\theta}_1) \leq \mathrm{MSE}(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 在均方误差意义下比 $\hat{\theta}_2$ 有效。
- 一个有偏但方差很小的估计量，其均方误差可能优于无偏但方差较大的估计量。

练习 7.26 设总体 X 的概率密度为

$$f(x)=\begin{cases} e^{\theta-x}, & x>\theta, \\ 0, & x\le\theta, \end{cases}$$

θ 未知，$X_1,\ldots,X_n$ 为样本。

(1) 求 θ 的矩估计量 $\hat{\theta}_M$ 和最大似然估计量 $\hat{\theta}_L$；

(2) 比较 $\hat{\theta}_M$ 与 $\hat{\theta}_L$ 的优劣。

解 (1) $EX=\int_\theta^\infty x\,e^{\theta-x}\,\mathrm{d}x=\theta+1$。令 $\theta+1=\overline{X}$，得 $\hat{\theta}_M=\overline{X}-1$。$L(\theta)=e^{n\theta-\sum X_i}$，关于 θ 递增（$\theta\le X_{(1)}$），故 $\hat{\theta}_L=X_{(1)}$。

(2) $\hat{\theta}_M$ 无偏，$D(\hat{\theta}_M)=D(\overline{X})=1/n$（因为 $DX=1$），故 $\mathrm{MSE}_M=1/n$。

$X_{(1)}-\theta$ 为 n 个 $\mathrm{Exp}(1)$ 的最小值，$E(X_{(1)}-\theta)=1/n$，$D(X_{(1)}-\theta)=1/n^2$。故 $E(\hat{\theta}_L)=\theta+1/n$，$D(\hat{\theta}_L)=1/n^2$，

$$\mathrm{MSE}_L=\left(\frac{1}{n}\right)^2+\frac{1}{n^2}=\frac{2}{n^2}.$$

当 $n>2$ 时，$2/n^2<1/n$，$\hat{\theta}_L$ 的均方误差更小，更优。

第 8 章 区间估计

$$\begin{cases}\text{单个正态}X\sim N(\mu,\sigma^2)\\ \text{置信水平为}1-\alpha\end{cases}\Rightarrow\begin{cases}\text{总体均值}\mu:\begin{cases}\sigma^2\text{ 已知}: & \left(\overline{X}-\dfrac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}},\overline{X}+\dfrac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)\\ \sigma^2\text{ 未知}: & \left(\overline{X}-\dfrac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1),\overline{X}+\dfrac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)\end{cases}\\ \text{总体方差}\sigma^2:\begin{cases}\mu\text{ 已知}: & \left(\dfrac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)},\dfrac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right)\\ \mu\text{ 未知}: & \left(\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)\end{cases}\end{cases}$$

$$\begin{array}{l}\text{两个正态}\\ X_1\sim N(\mu_1,\sigma_1^2);\\ X_2\sim N(\mu_2,\sigma_2^2)\\ \text{置信水平为}1-\alpha\end{array}\begin{cases}\text{总体均值差}(\mu_1-\mu_2):\\ \quad\begin{cases}\sigma_1^2,\sigma_2^2\text{ 已知}:\\ \quad(\bar{X}_1-\bar{X}_2)\pm z_{\alpha/2}\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2};\\ \sigma_1^2=\sigma_2^2=\sigma^2,\ \sigma^2\text{ 未知}:\\ \quad(\bar{X}_1-\bar{X}_2)\pm t_{\alpha/2}(n_1+n_2-2)\,S_w\sqrt{1/n_1+1/n_2};\end{cases}\\ \text{总体方差比}\left(\sigma_1^2/\sigma_2^2\right):\\ \quad\begin{cases}\mu_1,\mu_2\text{ 已知}:\\ \quad\left(\dfrac{\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{\frac{1}{n_2}\sum_{j=1}^{n_2}(X_j-\mu_2)^2}\cdot\dfrac{1}{F_{\alpha/2}(n_1,n_2)},\ \dfrac{\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{\frac{1}{n_2}\sum_{j=1}^{n_2}(X_j-\mu_2)^2}\cdot\dfrac{1}{F_{1-\alpha/2}(n_1,n_2)}\right)\\ \mu_1,\mu_2\text{ 未知}:\\ \quad\left(\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\ \dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right)\end{cases}\end{cases}$$

内容提要

- 区间估计的基本概念：双侧置信区间与单侧置信区间
- 构造置信区间的基本原理与枢轴量法
- 正态总体参数（均值、方差）的区间估计

8.1 区间估计的基本概念

对于未知参数 θ，除了求出它的点估计外，我们往往还希望给出一个合理的范围，并知道这个范围包含参数 θ 真值的可信程度。这样的范围通常以区间的形式给出，这种估计形式被称为**区间估计**。

定义 8.1 (置信区间与置信水平)

设 θ 是总体 X 的分布函数中的一个未知参数，对于给定的显著性水平 α $(0<\alpha<1)$，如果由样本 $X_1,X_2,\ldots,X_n$ 确定的两个统计量 $\underline{\theta}=\underline{\theta}(X_1,X_2,\ldots,X_n)$ 和 $\overline{\theta}=\overline{\theta}(X_1,X_2,\ldots,X_n)$，满足

$\underline{\theta} < \overline{\theta}$，且对于任意的 $\theta \in \Theta$ 均有：

$$P\{\underline{\theta} < \theta < \overline{\theta}\} \geq 1 - \alpha \tag{8.1}$$

则称随机区间 $(\underline{\theta}, \overline{\theta})$ 是 θ 的置信度（或置信水平）为 $1-\alpha$ 的双侧置信区间。其中，$\underline{\theta}$ 和 $\overline{\theta}$ 分别称为置信度为 $1-\alpha$ 的置信下限和置信上限。

特别地，如果满足 $P\{\theta < \underline{\theta}\} = P\{\theta > \overline{\theta}\} = \frac{\alpha}{2}$，则称这种置信区间为等尾置信区间。 ♣

在实际应用中，有时我们只关心参数的下界或上界（例如零件的寿命越长越好，废品率越低越好），此时就需要使用单侧置信区间。

定义 8.2 (单侧置信区间)

对于给定值 α $(0 < \alpha < 1)$，若由样本确定的统计量 $\underline{\theta}$ 满足 $P\{\theta > \underline{\theta}\} \geq 1 - \alpha$，则称 $(\underline{\theta}, +\infty)$ 是 θ 的置信水平为 $1-\alpha$ 的单侧置信区间，$\underline{\theta}$ 称为单侧置信下限。

同理，若由样本确定的统计量 $\overline{\theta}$ 满足 $P\{\theta < \overline{\theta}\} \geq 1 - \alpha$，则称 $(-\infty, \overline{\theta})$ 是 θ 的置信水平为 $1-\alpha$ 的单侧置信区间，$\overline{\theta}$ 称为单侧置信上限。 ♣

笔记 若总体 X 为连续型随机变量，我们通常要求上述概率严格等于 $1-\alpha$；若 X 为离散型随机变量，则使其尽可能地接近（且不小于）$1-\alpha$。

8.2 构造置信区间的基本原理

构造置信区间最常用的方法是**枢轴量法**。枢轴量是一个由样本和未知参数共同构成的函数，其最大的特点是**它的分布完全已知，不依赖于任何未知参数**。下面我们以正态分布均值的区间估计为例，说明枢轴量法的基本步骤。

练习 8.1 正态总体均值的区间估计 设总体 $X \sim N(\mu, \sigma^2)$，其中方差 σ^2 已知，均值 μ 未知。现有一组来自 X 的样本 $X_1, X_2, \ldots, X_n$，求 μ 的置信水平为 $1-\alpha$ 的置信区间。

证明 首先，寻找一个包含未知参数 μ 的枢轴量 $W = W(X_1, X_2, \ldots, X_n; \mu)$。由于样本均值 $\overline{X} \sim N(\mu, \sigma^2/n)$，我们可以对其进行标准化，得到：

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

由于标准正态分布 $N(0,1)$ 不依赖于任何未知参数，因此 W 就是一个理想的枢轴量。

其次，根据枢轴量的已知分布和给定的置信水平 $1-\alpha$，利用分位数的定义，寻找常数 a 和 b，使得 $P\{a < W < b\} = 1-\alpha$。通常我们取等尾情况，即利用标准正态分布的上 $\alpha/2$ 分位点 $z_{\alpha/2}$：

$$P\left\{-z_{\alpha/2} < \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right\} = 1 - \alpha$$

最后，从不等式中解出未知参数 μ，即得到：

$$P\left\{\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu < \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right\} = 1 - \alpha$$

由此，我们得到了 μ 的置信水平为 $1-\alpha$ 的置信区间为：

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$$

除了上述 σ^2 已知时 μ 的区间估计外，我们还可以利用类似的方法推导其他常见参数的置信区间。

练习 8.2 正态总体均值的区间估计（方差未知） 设总体 $X \sim N(\mu, \sigma^2)$，其中均值 μ 和方差 σ^2 均未知。现有一组来自 X 的独立样本 $X_1, X_2, \ldots, X_n$，样本均值为 $\overline{X}$，样本方差为 S^2。求 μ 的置信水平为 $1-\alpha$ 的置信区间。

证明 由于 σ^2 未知，我们不能再使用包含 σ 的枢轴量。考虑到样本方差 S^2 是 σ^2 的无偏估计，我们构造如下枢轴量：

$$W = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

该统计量服从自由度为 $n-1$ 的 t 分布，完全不依赖于未知参数。

根据 t 分布的对称性，利用上 $\alpha/2$ 分位点 $t_{\alpha/2}(n-1)$，我们有：

$$P\left\{-t_{\alpha/2}(n-1) < \frac{\overline{X} - \mu}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\right\} = 1-\alpha$$

从中解出 μ，即得到：

$$P\left\{\overline{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1) < \mu < \overline{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right\} = 1-\alpha$$

由此得到 μ 的置信水平为 $1-\alpha$ 的置信区间为：

$$\left(\overline{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1),\ \overline{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right)$$

练习 8.3 正态总体方差的区间估计（均值未知） 设总体 $X \sim N(\mu, \sigma^2)$，其中均值 μ 和方差 σ^2 均未知。现有一组来自 X 的独立样本 $X_1, X_2, \ldots, X_n$，样本方差为 S^2。求方差 σ^2 的置信水平为 $1-\alpha$ 的置信区间。

证明 此时我们要估计的参数是 σ^2。根据正态总体样本方差的抽样分布性质，我们知道：

$$W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

由于 χ^2 分布不依赖于未知参数 μ 和 σ^2，因此 W 就是一个理想的枢轴量。

由于 χ^2 分布是不对称的，我们分别在其两端各取 $\alpha/2$ 的概率面积。利用 χ^2 分布的上 $\alpha/2$ 和 $1-\alpha/2$ 分位点，我们有：

$$P\left\{\chi^2_{1-\alpha/2}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}(n-1)\right\} = 1-\alpha$$

对不等式进行变形，取倒数并同乘 $(n-1)S^2$，解出 σ^2：

$$P\left\{\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right\} = 1-\alpha$$

由此得到 σ^2 的置信区间为：

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$$

对于两个正态总体 $X \sim N(\mu_1, \sigma_1^2)$ 和 $Y \sim N(\mu_2, \sigma_2^2)$，我们同样可以利用类似的方法构造枢轴量，来估计它们的均值差或方差比。

练习 8.4 两正态总体均值差的区间估计（方差相等且未知） 设有两个独立的正态总体 $X \sim N(\mu_1, \sigma^2)$ 和 $Y \sim N(\mu_2, \sigma^2)$，方差相等但具体数值未知。分别抽取容量为 n_1 和 n_2 的独立样本，求均值差 $\mu_1 - \mu_2$ 的置信水平为 $1-\alpha$ 的置信区间。

证明 要估计 $\mu_1 - \mu_2$，一个自然的无偏估计量是 $\overline{X} - \overline{Y}$。由于两总体方差相同，我们可以利用两个样

本的方差 S_1^2 和 S_2^2 构造联合样本方差 S_w^2：

$$S_w^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$

由此，我们构造出不含未知参数的枢轴量：

$$W = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

利用 t 分布的对称性和上 $\alpha/2$ 分位点，我们有：

$$P\left\{-t_{\alpha/2} < \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} < t_{\alpha/2}\right\} = 1-\alpha$$

（这里简记 $t_{\alpha/2}$ 为 $t_{\alpha/2}(n_1+n_2-2)$）。

从中解出 $\mu_1-\mu_2$，得到置信区间为：

$$\left((\overline{X}-\overline{Y})-t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ (\overline{X}-\overline{Y})+t_{\alpha/2}(n_1+n_2-2)S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

练习 8.5 两正态总体方差比的区间估计（均值未知） 设有两个独立的正态总体 $X \sim N(\mu_1,\sigma_1^2)$ 和 $Y \sim N(\mu_2,\sigma_2^2)$，其均值均未知。求方差比 σ_1^2/σ_2^2 的置信水平为 $1-\alpha$ 的置信区间。

证明 根据两个样本方差的比值，可以构造服从 F 分布的统计量。为了让其成为枢轴量，我们将包含未知参数的部分分离出来：

$$W = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1-1,n_2-1)$$

利用 F 分布的上 $\alpha/2$ 和 $1-\alpha/2$ 分位点，我们有：

$$P\left\{F_{1-\alpha/2}(n_1-1,n_2-1) < \frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2} < F_{\alpha/2}(n_1-1,n_2-1)\right\} = 1-\alpha$$

将其变形，提取出 σ_1^2/σ_2^2。取倒数时不等号反向：

$$P\left\{\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right\} = 1-\alpha$$

由于 F 分布的性质有 $F_{1-\alpha/2}(n_1-1,n_2-1) = 1/F_{\alpha/2}(n_2-1,n_1-1)$，因此上界可以进一步化简为 $\frac{S_1^2}{S_2^2}\cdot F_{\alpha/2}(n_2-1,n_1-1)$。

由此得到 σ_1^2/σ_2^2 的置信区间为：

$$\left(\frac{S_1^2}{S_2^2}\cdot\frac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\ \frac{S_1^2}{S_2^2}\cdot F_{\alpha/2}(n_2-1,n_1-1)\right)$$

8.3 常见的参数置信区间总结

利用上述枢轴量法，我们推导出了现实中最常见的正态总体（包括单个和两个正态总体）在不同情况下的参数置信区间。为了便于查阅，我们将相关公式归纳为表 8.1。

注 上述表格中公式的推导核心在于“寻找正确的枢轴量”，因此读者不应当采用死记硬背的方式来记忆公式，而是应当通过枢轴量和分位数的定义，自行推导得出。

表 8.1: 常见的正态总体参数置信区间（置信度为 $1-\alpha$）

总体条件	待估参数	枢轴量及分布	置信区间
		单个正态总体 $X \sim N(\mu, \sigma^2)$	
σ^2 已知	μ	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2},\ \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$
σ^2 未知	μ	$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$\left(\overline{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1),\ \overline{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n-1)\right)$
μ 未知	σ^2	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\ \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$
μ 已知	σ^2	$\frac{\sum(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$	$\left(\frac{\sum(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)},\ \frac{\sum(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right)$
		两个正态总体 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$	
σ_1^2, σ_2^2 已知	$\mu_1 - \mu_2$	正态分布 $N(0,1)$	$\left((\overline{X} - \overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$
$\sigma_1^2 = \sigma_2^2$ 未知	$\mu_1 - \mu_2$	t 分布 $t(n_1+n_2-2)$	$\left((\overline{X} - \overline{Y}) \pm t_{\alpha/2}(n_1+n_2-2) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$
μ_1, μ_2 已知	$\frac{\sigma_1^2}{\sigma_2^2}$	F 分布 $F(n_1, n_2)$	$\left(\frac{\sum(X_i-\mu_1)^2/n_1}{\sum(Y_j-\mu_2)^2/n_2} \frac{1}{F_{\alpha/2}(n_1,n_2)},\ \frac{\sum(X_i-\mu_1)^2/n_1}{\sum(Y_j-\mu_2)^2/n_2} F_{\alpha/2}(n_2, n_1)\right)$
μ_1, μ_2 未知	$\frac{\sigma_1^2}{\sigma_2^2}$	F 分布 $F(n_1-1, n_2-1)$	$\left(\frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(n_1-1,n_2-1)},\ \frac{S_1^2}{S_2^2} F_{\alpha/2}(n_2-1, n_1-1)\right)$

8.4 典型例题

练习 8.6 设总体 $X \sim N(\mu, \sigma^2)$（σ^2 已知），则在给定样本容量 n 及置信度 $1-\alpha$ 的情况下，未知参数 μ 的置信区间长度随着样本均值 $\overline{X}$ 的增加而（　　）。

(A) 增加

(B) 减少

(C) 不变

(D) 不能确定增或减

解 应选 (C)。

在给定条件下，由表 8.1 可知，未知参数 μ 的置信区间为：

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2},\ \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$$

该置信区间的长度为 $L = 2\frac{\sigma}{\sqrt{n}} z_{\alpha/2}$。显然，区间长度 L 只与总体标准差 σ、样本容量 n 以及对应置信水平的分位点 $z_{\alpha/2}$ 有关，而与样本均值 $\overline{X}$ 毫无关系。因此，随着 $\overline{X}$ 的变化，置信区间的长度保持不变。

练习 8.7 设总体 X 的方差为 1，根据来自 X 的容量为 100 的简单随机样本，测得样本均值为 5。已知标准正态分布函数值 $\Phi(1.96) = 0.975$。求 X 的数学期望 μ 的置信度近似等于 0.95 的置信区间。

解 题目中并未说明总体 X 是否为正态分布，但是由于样本容量 $n = 100 \geq 30$，属于大样本。根据中心极限定理，样本均值 $\overline{X}$ 近似服从正态分布，即：

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - \mu}{1/\sqrt{100}} \overset{\text{近似}}{\sim} N(0,1)$$

又因为 $\Phi(1.96) = 0.975$，可知 $z_{0.025} = 1.96$，故 $P\{-1.96 < Z < 1.96\} \approx 0.95$。

由此可得 μ 的近似置信区间为：

$$\left(\overline{X}-\frac{1}{10}\times 1.96,\ \overline{X}+\frac{1}{10}\times 1.96\right)$$

代入已知数据 $\overline{X}=5$，计算出置信区间的下限为 4.804，上限为 5.196。

所以 μ 的置信区间近似为 $(4.804,\ 5.196)$。

练习 8.8 对于正态总体 $X\sim N(\mu,\sigma^2)$，σ^2 已知，记 L 是 μ 的置信度为 0.95 的置信区间的长度，那么当样本容量 n _____ 时，$L\le\sigma$。

（参考数据：$z_{0.05}=1.645$，$z_{0.025}=1.96$，$t_{0.025}(15)=1.753$，$t_{0.025}(16)=2.12$）

解 因为 σ^2 已知，使用正态分布作为枢轴量。置信区间长度为：

$$L=2\cdot\frac{\sigma}{\sqrt{n}}z_{0.025}$$

代入 $z_{0.025}=1.96$，并令 $L\le\sigma$：

$$2\cdot 1.96\cdot\frac{\sigma}{\sqrt{n}}\le\sigma\implies\frac{3.92}{\sqrt{n}}\le 1\implies\sqrt{n}\ge 3.92\implies n\ge 15.3664$$

由于样本容量 n 必须为正整数，故 $n\ge 16$。

练习 8.9 生产一个零件所需时间 $X\sim N(\mu,\sigma^2)$。现观察了 25 个零件，测得样本均值 $\overline{X}=5.5$，样本标准差 $S=1.73$。求 μ 的置信度为 0.95 的置信区间 _____。

（参考数据：$t_{0.05}(24)=1.7109$, $t_{0.025}(24)=2.0639$, $\Phi(1.96)=0.975$）

解 由于总体方差 σ^2 未知，需要使用 t 分布。自由度 $df=n-1=24$，查表得 $t_{0.025}(24)=2.0639$。置信区间公式为：

$$\left(\overline{X}-\frac{S}{\sqrt{n}}t_{0.025}(24),\ \overline{X}+\frac{S}{\sqrt{n}}t_{0.025}(24)\right)$$

计算边际误差：

$$\frac{1.73}{5}\times 2.0639\approx 0.7141$$

下限为 $5.5-0.7141=4.7859$，上限为 $5.5+0.7141=6.2141$。故置信区间为 $(4.7859,\ 6.2141)$。

练习 8.10 设总体 $X\sim N(\mu,\sigma^2)$，有一个容量为 10 的样本，测得样本方差 $S^2=2$。求 σ^2 的置信水平为 0.95 的置信区间 _____。

（参考数据：$\chi^2_{0.975}(9)=2.70$，$\chi^2_{0.025}(9)=19.023$）

解 由于 μ 未知，求 σ^2 的置信区间使用 χ^2 分布，自由度 $df=10-1=9$。方差的置信区间公式为：

$$\left(\frac{(n-1)S^2}{\chi^2_{0.025}(9)},\ \frac{(n-1)S^2}{\chi^2_{0.975}(9)}\right)$$

计算分子：$(n-1)S^2=9\times 2=18$。代入分母得到：

$$\frac{18}{19.023}\approx 0.9462,\quad \frac{18}{2.70}\approx 6.6667$$

故置信区间为 $(0.9462,\ 6.6667)$。

练习 8.11 已知 $X\sim N(\mu,\sigma^2)$，样本容量 $n=9$，样本均值 $\overline{X}=6$，求 μ 在置信水平 0.95 下的单侧置信下限 _____。

（参考数据：$\Phi(1.96)=0.975$, $\Phi(1.645)=0.95$）

解 由于需要用到 σ，我们将其视为已知。单侧置信下限的要求是 $P\{\mu\ge\underline{\mu}\}=0.95$。利用标准正态分布的枢轴量，有：

$$P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\le z_{0.05}\right)=0.95$$

解得 $\mu \geq \overline{X} - z_{0.05}\frac{\sigma}{\sqrt{n}}$。代入数据 $\overline{X} = 6, n = 9, z_{0.05} = 1.645$：

$$\underline{\mu} = 6 - 1.645\frac{\sigma}{3} \approx 6 - 0.5483\sigma$$

练习 8.12 对 $\alpha \in (0,1)$，由样本 $X_1, \ldots, X_n$ 得到统计量 $\underline{\theta}$，若对一切参数 θ 均有 _____，则称 $(\underline{\theta}, +\infty)$ 为 θ 的置信水平为 $1-\alpha$ 的单侧置信区间，$\underline{\theta}$ 称为单侧置信下限。

解 根据定义，单侧置信下限意味着参数真值大于等于该下限的概率不低于要求的置信水平。由于是针对连续情况下的理论定义，满足 $P(\theta \geq \underline{\theta}) \geq 1-\alpha$ 或 $P(\theta \geq \underline{\theta}) = 1-\alpha$。

练习 8.13 设总体 $X \sim N(\mu, \sigma^2)$，有一个容量 $n = 9$ 的样本，已知 $\sum_{i=1}^{9} X_i = 198$，$\sum_{i=1}^{9} X_i^2 = 4372$。求 μ 与 σ 的置信水平为 0.95 的双侧置信区间（结果保留四位小数）。

（参考数据：$t_{0.025}(8) = 2.306$，$\chi^2_{0.975}(8) = 2.180$，$\chi^2_{0.025}(8) = 17.535$）

解 首先计算样本均值和样本方差：

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{9} X_i = \frac{198}{9} = 22$$

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{9} X_i^2 - n\overline{X}^2\right) = \frac{1}{8}\left(4372 - 9 \times 22^2\right) = \frac{16}{8} = 2$$

于是样本标准差 $S = \sqrt{2} \approx 1.4142$。

求 μ 的置信区间（方差未知，用 t 分布）：自由度 $df = 8$，分位点 $t_{0.025}(8) = 2.306$。边际误差为 $\frac{\sqrt{2}}{3} \times 2.306 \approx 1.0871$。下限：$22 - 1.0871 = 20.9129$，上限：$22 + 1.0871 = 23.0871$。即 μ 的置信区间为 $(20.9129,\ 23.0871)$。

求 σ 的置信区间：先求出 σ^2 的置信区间：

$$\left(\frac{(n-1)S^2}{\chi^2_{0.025}(8)}, \frac{(n-1)S^2}{\chi^2_{0.975}(8)}\right) = \left(\frac{16}{17.535}, \frac{16}{2.180}\right) \approx (0.9125,\ 7.3395)$$

对上下限开平方，得到标准差 σ 的置信区间：

$$(\sqrt{0.9125},\ \sqrt{7.3395}) \approx (0.9552,\ 2.7091)$$

综上，μ 的置信区间为 $(20.9129,\ 23.0871)$，σ 的置信区间为 $(0.9552,\ 2.7091)$。

第 9 章 假设检验

$$
\begin{cases}
\text{假设检验}\begin{cases}
\text{基本思想: 随机试验中小概率事件一般不会发生} \\
\text{基本步骤: 提出假设} \Rightarrow \text{构建统计量} \Rightarrow \text{写出拒绝域} \Rightarrow \text{检验} \\
\text{主要检验: 双边检验或单边检验} \\
\text{显著性检验: 仅考虑控制第一类错误的假设检验，限定第一类错误发生的最大概率为}\alpha
\end{cases} \\
\text{单个正态总体}\begin{cases}
\text{总体均值}\begin{cases}
\text{总体方差已知, } Z\text{检验，也称}U\text{检验} \\
\text{总体方差未知, } t\text{检验}
\end{cases} \\
\text{总体方差}\begin{cases}
\text{总体均值已知, } \chi^2(n)\text{检验} \\
\text{总体均值未知, } \chi^2(n-1)\text{检验}
\end{cases}
\end{cases} \\
\text{两个正态总体区间估计} \\
(X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2))\begin{cases}
\text{总体均值差}(\mu_1 - \mu_2)\begin{cases}
\sigma_1^2 = \sigma_2^2 = \sigma^2\text{已知, } Z\text{检验} \\
\sigma_1^2 = \sigma_2^2 = \sigma^2\text{未知, } t(n_1 + n_2 - 2)\text{检验}
\end{cases} \\
\text{总体方差比}\left(\frac{\sigma_1^2}{\sigma_2^2}\right)\begin{cases}
\mu_1 = \mu_2 = \mu\text{已知, } F(n_1, n_2)\text{检验} \\
\mu_1 = \mu_2 = \mu\text{未知, } F(n_1 - 1, n_2 - 1)\text{检验}
\end{cases}
\end{cases}
\end{cases}
$$

内容提要

- 原假设与备择假设 9.1
- 两类错误与显著性水平 9.2
- 假设检验的基本步骤 9.2
- 正态总体参数的假设检验 9.3
- 置信区间与假设检验的关系 9.4

9.1 基本概念

定义 9.1 (假设检验)

关于总体分布中的未知参数或分布类型的每一种论断称为统计假设。根据样本信息去推断该假设是否成立的统计推断问题称为假设检验。其中需要检验的假设称为原假设（或零假设），记为 H_0；与 H_0 对立的假设称为备择假设，记为 H_1。 ♣

定义 9.2 (两类错误)

1. 第一类错误（弃真）：H_0 为真时，错误地拒绝了 H_0。

$$\alpha = P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\}.$$

2. 第二类错误（取伪）：H_0 不真时，错误地接受了 H_0。

$$\beta = P\{\text{接受 } H_0 \mid H_1 \text{ 为真}\}.$$

♣

定义 9.3 (显著性水平与显著性检验)

在假设检验中，只控制犯第一类错误的概率不超过 α（α 通常取 0.1, 0.05, 0.01），而不考虑第二类错误的概率，这种检验称为显著性检验，α 称为显著性水平。 ♣

笔记 小概率原理是假设检验的理论基础：概率很小的事件在一次试验中认为不会发生。若在 H_0 成立的条件下，样本观测值导致某个概率 $\leq \alpha$ 的事件发生，则拒绝 H_0。

定义 9.4 (双边检验与单边检验)

- 双边检验：$H_0: \mu = \mu_0$，$H_1: \mu \neq \mu_0$。
- 右边检验：$H_0: \mu \leq \mu_0$，$H_1: \mu > \mu_0$。
- 左边检验：$H_0: \mu \geq \mu_0$，$H_1: \mu < \mu_0$。 ♣

笔记 原假设 H_0 与备择假设 H_1 的地位不对等：H_0 是受保护的（控制弃真概率），因此选择 H_0 需要谨慎，一般遵循以下原则：

1. 将后果较严重的错误设为第一类错误。例如检验药品真假，H_0：药品为假，H_1：药品为真（误判假药为真的后果更严重）。
2. 若无严重后果，常选 H_0 为"维持现状"。例如检验新技术是否提高收益，H_0：新技术未提高收益。
3. 用反证法思想：若要验证 x 显著高于 y，则设 $H_0: x \leq y$，拒绝 H_0 后便有充分理由认为 $x > y$。

9.2 假设检验的步骤

性质 [假设检验的四步流程]

1. 根据问题提出 H_0 和 H_1，确定检验类型（双边/单边）。
2. 给定显著性水平 α 和样本容量 n。
3. 选择检验统计量，确定拒绝域。
4. 由样本计算统计量的观测值，判断是否落入拒绝域，作出结论。

笔记 拒绝域的形式与备择假设 H_1 一致，便于记忆：

- $H_1: \theta > \theta_0$（右边检验）$\Longrightarrow$ 拒绝域在右尾；
- $H_1: \theta < \theta_0$（左边检验）$\Longrightarrow$ 拒绝域在左尾；
- $H_1: \theta \neq \theta_0$（双边检验）$\Longrightarrow$ 拒绝域在双尾。

接受 H_0 并不意味着 H_0 一定为真，只是没有足够证据拒绝它。

9.3 正态总体参数的假设检验

9.3.1 单个正态总体均值的检验

性质 [Z 检验（σ^2 已知）] 检验统计量 $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$。

H_1	拒绝域
$\mu > \mu_0$	$z \geq z_\alpha$
$\mu < \mu_0$	$z \leq -z_\alpha$
$\mu \neq \mu_0$	$\lvert z\rvert \geq z_{\alpha/2}$

练习 9.1 某工厂生产灯泡的寿命 $X \sim N(\mu, 200^2)$。过去灯泡平均寿命为 1500 小时，采用新工艺后抽取 25 只，测得平均寿命为 1675 小时。问在显著性水平 $\alpha = 0.05$ 下，新工艺是否显著提高了灯泡寿命？（$z_{0.05} = 1.65$）

解 H_0: $\mu \leq 1500$，H_1: $\mu > 1500$（右边检验，$\alpha = 0.05$）。检验统计量 $Z = \dfrac{\overline{X} - 1500}{200/\sqrt{25}} \sim N(0,1)$，拒绝域 $z \geq z_{0.05} = 1.65$。

$$z = \frac{1675 - 1500}{200/5} = \frac{175}{40} = 4.375 > 1.65.$$

z 落入拒绝域，拒绝 H_0，认为新工艺显著提高了灯泡寿命。

性质 [t 检验（σ^2 未知）] 检验统计量 $t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$。

H_1	拒绝域
$\mu > \mu_0$	$t \geq t_\alpha(n-1)$
$\mu < \mu_0$	$t \leq -t_\alpha(n-1)$
$\mu \neq \mu_0$	$\lvert t\rvert \geq t_{\alpha/2}(n-1)$

练习 9.2 某厂生产灯泡寿命 $X \sim N(\mu, \sigma^2)$（μ, σ^2 均未知），抽取 25 只测得 $\overline{x} = 1700$，$s = 490$。在 $\alpha = 0.01$ 下，能否认为平均寿命为 2000 小时？（$t_{0.005}(24) = 2.7969$）

解 H_0: $\mu = 2000$，H_1: $\mu \neq 2000$（双边检验，$\alpha = 0.01$）。$t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t(24)$，拒绝域 $\lvert t\rvert \geq t_{0.005}(24) = 2.7969$。

$$t = \frac{1700 - 2000}{490/5} = \frac{-300}{98} = -3.061, \quad \lvert t\rvert = 3.061 > 2.7969.$$

落入拒绝域，拒绝 H_0，不能认为平均寿命为 2000 小时。

练习 9.3 某高校某专业 25 名学生高数成绩服从正态分布，样本均值 $\overline{x} = 69.5$，标准差 $s = 15$。在 $\alpha = 0.05$ 下，检验该专业平均成绩是否**显著高于** 70 分。（$t_{0.05}(24) = 1.711$）

解 H_0: $\mu \leq 70$，H_1: $\mu > 70$（右边检验，$\alpha = 0.05$）。拒绝域 $t \geq t_{0.05}(24) = 1.711$。

$$t = \frac{69.5 - 70}{15/\sqrt{25}} = \frac{-0.5}{3} = -0.167 < 1.711.$$

未落入拒绝域，接受 H_0，认为平均成绩没有显著高于 70 分。

9.3.2 两个正态总体均值差的检验

性质 [两个正态总体均值差的检验] 设 $X_1, \ldots, X_{n_1}$ 来自 $N(\mu_1, \sigma_1^2)$，$Y_1, \ldots, Y_{n_2}$ 来自 $N(\mu_2, \sigma_2^2)$，两组独立。

情形 **1**：σ_1^2, σ_2^2 已知。检验统计量

$$Z = \frac{(\overline{X} - \overline{Y}) - \delta}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0,1).$$

情形 **2**：$\sigma_1^2=\sigma_2^2=\sigma^2$ 未知。检验统计量

$$t=\frac{(\overline{X}-\overline{Y})-\delta}{S_w\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2),$$

其中 $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ 为合并样本方差。

两种情形下，$H_1\colon \mu_1-\mu_2>\delta$ 对应右尾拒绝域，$H_1\colon \mu_1-\mu_2<\delta$ 对应左尾，$H_1\colon \mu_1-\mu_2\neq\delta$ 对应双尾。

9.3.3 单个正态总体方差的检验

性质 [χ^2 检验] μ 未知时，检验统计量

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1).$$

μ 已知时，检验统计量

$$\chi^2=\frac{\sum_{i=1}^n(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n).$$

H_1	拒绝域（μ 未知）
$\sigma^2>\sigma_0^2$	$\chi^2\geq\chi_\alpha^2(n-1)$
$\sigma^2<\sigma_0^2$	$\chi^2\leq\chi_{1-\alpha}^2(n-1)$
$\sigma^2\neq\sigma_0^2$	$\chi^2\geq\chi_{\alpha/2}^2(n-1)$ 或 $\chi^2\leq\chi_{1-\alpha/2}^2(n-1)$

✍ **练习 9.4** 某厂铜线折断力 $X\sim N(\mu,\sigma^2)$，抽查 10 根，$\overline{x}=575.2$，$s^2=68.16$。在 $\alpha=0.05$ 下检验 $H_0\colon\sigma^2=64$，$H_1\colon\sigma^2\neq64$。（$\chi_{0.025}^2(9)=19.0$，$\chi_{0.975}^2(9)=2.7$）

解 $\chi^2=\dfrac{(10-1)\times68.16}{64}=\dfrac{613.44}{64}=9.585$。

接受域：$\chi_{0.975}^2(9)=2.7\leq\chi^2\leq\chi_{0.025}^2(9)=19.0$。

$2.7<9.585<19.0$，落入接受域，接受 H_0，认为方差为 64。

✍ **练习 9.5** 测定溶液水分，10 个测定值 $s=0.037\%$，总体服从正态分布。在 $\alpha=0.05$ 下检验 $H_0\colon\sigma\geq0.04\%$，$H_1\colon\sigma<0.04\%$。（$\chi_{0.95}^2(9)=3.325$）

解 等价于检验 $H_0\colon\sigma^2\geq(0.04\%)^2$，$H_1\colon\sigma^2<(0.04\%)^2$（左边检验）。拒绝域：$\chi^2\leq\chi_{0.95}^2(9)=3.325$。

$$\chi^2=\frac{(10-1)\times(0.037\%)^2}{(0.04\%)^2}=\frac{9\times0.001369}{0.0016}=7.701>3.325.$$

未落入拒绝域，接受 H_0。

9.3.4 两个正态总体方差比的检验

性质 [F 检验] μ_1,μ_2 未知时，检验统计量

$$F=\frac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1).$$

H_1	拒绝域
$\sigma_1^2 > \sigma_2^2$	$F \geq F_\alpha(n_1-1, n_2-1)$
$\sigma_1^2 < \sigma_2^2$	$F \leq F_{1-\alpha}(n_1-1, n_2-1)$
$\sigma_1^2 \neq \sigma_2^2$	$F \geq F_{\alpha/2}$ 或 $F \leq F_{1-\alpha/2}$

其中 $F_{1-\alpha}(n_1-1, n_2-1) = \dfrac{1}{F_\alpha(n_2-1, n_1-1)}$。

练习 9.6 两个小麦品种从播种到抽穗天数如下：

品种 1	101	100	99	99	98	100	98	99	99	99	行和 992
品种 2	100	98	100	99	98	99	98	98	99	100	行和 989

在 $\alpha = 0.05$ 下：(1) 检验方差是否有显著差异；(2) 检验均值是否有显著差异。（$F_{0.025}(9,9) = 4.03$，$F_{0.975}(9,9) = 0.248$，$t_{0.025}(18) = 2.101$）

解 $\overline{x}_1 = 99.2$, $\overline{x}_2 = 98.9$, $n_1 = n_2 = 10$。

$$\sum(x_{1i} - \overline{x}_1)^2 = 7.6, \quad \sum(x_{2j} - \overline{x}_2)^2 = 6.9.$$

$$S_1^2 = \frac{7.6}{9} \approx 0.844, \quad S_2^2 = \frac{6.9}{9} \approx 0.767.$$

(1) $H_0\colon \sigma_1^2 = \sigma_2^2$，$H_1\colon \sigma_1^2 \neq \sigma_2^2$。

$F = S_1^2/S_2^2 = 0.844/0.767 \approx 1.100$。拒绝域：$F \geq 4.03$ 或 $F \leq 0.248$。

$0.248 < 1.100 < 4.03$，未落入拒绝域，接受 H_0，方差无显著差异。

(2) $H_0\colon \mu_1 = \mu_2$，$H_1\colon \mu_1 \neq \mu_2$（方差已认为相等，用 pooled t 检验）。

$$S_w^2 = \frac{9 \times 0.844 + 9 \times 0.767}{18} = \frac{14.499}{18} \approx 0.806, \quad S_w\sqrt{\frac{2}{10}} \approx 0.898 \times 0.447 \approx 0.401.$$

$$t = \frac{99.2 - 98.9}{0.401} \approx 0.747.$$

拒绝域 $|t| \geq t_{0.025}(18) = 2.101$。$0.747 < 2.101$，未落入拒绝域，接受 H_0，均值无显著差异。

9.4 置信区间与假设检验的关系

性质 [置信区间与假设检验的等价性] 设 (θ_L, θ_U) 为参数 θ 的置信水平为 $1-\alpha$ 的置信区间，则对于双边检验

$$H_0\colon \theta = \theta_0, \quad H_1\colon \theta \neq \theta_0,$$

“接受 H_0”等价于“θ_0 落在置信区间 (θ_L, θ_U) 内”。对于单边检验也有类似的对应关系：$H_0\colon \theta \leq \theta_0$ 的拒绝域 $z \geq z_\alpha$ 等价于 $\theta_0 < \theta_U$（上置信限）。

练习 9.7 零件长度 $X \sim N(\mu, \sigma^2)$（μ, σ^2 均未知），$n = 16$，$\overline{x} = 10$，$s^2 = 0.16$。

(1) 求 μ 的置信水平为 0.95 的置信区间；

(2) 在 $\alpha = 0.05$ 下检验 $H_0\colon \mu = 9.7$，$H_1\colon \mu \neq 9.7$。（$t_{0.025}(15) = 2.132$）

解 (1) 置信区间：

$$\left(10 - \frac{0.4}{4} \times 2.132,\ 10 + \frac{0.4}{4} \times 2.132\right) = (9.7868,\ 10.2132).$$

(2) $9.7 \notin (9.7868,\ 10.2132)$，等价于落入拒绝域，故拒绝 H_0，认为 $\mu \neq 9.7$。验证：$\overline{x} = 10 > 9.7 + 0.4 \times 2.132/4 = 9.9132$，确实落入拒绝域。

9.5 两类错误

笔记 在固定样本容量 n 的条件下，α 与 β 存在此消彼长的关系：α 小则 β 大，β 小则 α 大。它们不满足 $\beta = 1 - \alpha$。实际中，总是在控制 α（使弃真概率较小）的条件下尽量使 β 小，因为人们通常认为错误地拒绝 H_0 比错误地接受 H_0 后果更严重。

练习 9.8 设连续型随机变量 U 的密度分别为：

$$H_0\colon f(x) = \begin{cases} 1/2, & 0 \le x \le 2, \\ 0, & \text{其他}, \end{cases} \qquad H_1\colon f(x) = \begin{cases} x/2, & 0 \le x \le 2, \\ 0, & \text{其他}. \end{cases}$$

检验规则：当 $U > 3/2$ 时拒绝 H_0。求犯第一类错误的概率 α 和犯第二类错误的概率 β。

A. $3/4,\ 7/16$　　B. $7/16,\ 3/4$　　C. $9/16,\ 1/4$　　D. $1/4,\ 9/16$

解 应选 D。

$$\alpha = P\{U > \tfrac{3}{2} \mid H_0\} = \int_{3/2}^{2} \frac{1}{2}\,\mathrm{d}x = \frac{1}{4}.$$

$$\beta = P\{U \le \tfrac{3}{2} \mid H_1\} = \int_{0}^{3/2} \frac{x}{2}\,\mathrm{d}x = \frac{(3/2)^2}{4} = \frac{9}{16}.$$

练习 9.9 在假设检验中，H_0 不成立时，检验统计量值未落入拒绝域从而接受了 H_0，这种错误称为第二类错误。

解 第二类错误（取伪）：H_0 不真时接受 H_0。

练习 9.10 显著性水平 α 的含义是（　）。

A. H_0 成立，经检验被拒绝的概率　　B. H_0 成立，经检验不能拒绝的概率

C. H_0 不成立，经检验被拒绝的概率　　D. H_0 不成立，经检验不能拒绝的概率

解 应选 A。$\alpha = P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\}$。