第 1 章　引　论

学习目标与要求

1. 回顾概率论的基本概念以及相关性质.
2. 理解随机过程的概念、平稳性、独立增量性和平稳增量性.
3. 理解泊松过程的有关概念, 并掌握它的重要性质.
4. 理解布朗运动的有关概念, 并掌握它的重要性质.

1.1 预备知识

1.1.1 概率空间

在概率论中, 通常把按照一定的想法去做的事情称为试验. 一个试验, 若它的结果预先无法确定, 则称之为**随机试验**, 简称为**试验**. 试验的每个可能结果称为**样本点**, 样本点的集合称为**样本空间**. 在本书中, 用 Ω 表示样本空间, 用 ω 表示样本点, 于是

$$\Omega=\{\omega:\omega \text{ 是试验的样本点}\}.$$

样本空间 Ω 中的样本点 ω 也称为**基本事件**. 样本空间的子集, 也即由基本事件构成的集合称为**事件**, 通常用大写英文字母 A,B,C 等表示. 样本空间 Ω 称为**必然事件**, 空集 $\varnothing$ 称为**不可能事件**. 由样本空间 Ω 中的若干子集构成的集合称为 Ω 的**集类**, 用花写的字母 $\mathcal{A},\mathcal{B},\mathcal{C}$ 等表示. 显然, 样本空间 Ω 的集类就是由一些事件构成的集合.

在实际问题中, 人们通常不是对样本空间的所有子集都感兴趣, 而是只关心某些事件及其发生的可能性大小. 为了方便地在人们感兴趣的事件上定义概率, 我们引入如下概念.

定义 1.1　设 $\mathcal{F}$ 是样本空间 Ω 的集类. 如果满足:

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

(2) 若 $A\in\mathcal{F}$, 则 $A^{\mathrm{c}}=\Omega-A\in\mathcal{F}$;

(3) 若 $A_n\in\mathcal{F}, n\in\mathbb{N}$, 则 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathcal{F}$,

则称 $\mathcal{F}$ 为样本空间 Ω 的一个 **σ 域** (或 **σ 代数**). 将 $(\Omega,\ \mathcal{F})$ 称为**可测空间**. $\mathcal{F}$ 中的元素便是人们感兴趣的事件.

容易验证, 若 $\mathcal{F}$ 是样本空间 Ω 的一个 σ 域, 则有: (1) $\varnothing\in\mathcal{F}$; (2) 若 $A,B\in\mathcal{F}$, 则 $A-B\in\mathcal{F}$; (3) 若 $A_n\in\mathcal{F}, n\in\mathbb{N}$, 则 $\bigcap\limits_{n=1}^{\infty}A_n\in\mathcal{F}$.

定义 1.2 设 $\mathcal{A}$ 为样本空间 Ω 的集类, 称一切包含 $\mathcal{A}$ 的 σ 域的交集为**由 $\mathcal{A}$ 生成的 σ 域**, 或称为**包含 $\mathcal{A}$ 的最小 σ 域**, 记为 $\sigma(\mathcal{A})$.

例 1.1 容易看出, 样本空间 Ω 的集类 $\mathcal{A}_0=\{\varnothing,\Omega\}$, $\mathcal{A}_1=\{\varnothing,A,A^{\mathrm{c}},\Omega\}$ 和 $\mathcal{A}_2=\{A:\ A\subseteq\Omega\}$ 都是 σ 域, 而集类 $\mathcal{C}=\{\varnothing,A,\Omega\}$ 不是 σ 域. 由 $\mathcal{C}$ 生成的 σ 域为 $\sigma(\mathcal{C})=\{\varnothing,A,A^{\mathrm{c}},\Omega\}$.

定义 1.3 设 $\Omega=\mathbb{R}$, 集类 $\mathcal{A}=\{(-\infty,x):x\in\mathbb{R}\}$, 则称由 $\mathcal{A}$ 生成的最小 σ 域 $\sigma(\mathcal{A})$ 为 $\mathbb{R}$ 上的 **Borel σ 域**, 记为 $\mathcal{B}(\mathbb{R})$. $\mathcal{B}(\mathbb{R})$ 中的元素称为 **Borel 集合**. 类似地, 可定义 $\mathbb{R}^n$ 上的 Borel σ 域 $\mathcal{B}(\mathbb{R}^n)$.

定义 1.4 设 $(\Omega,\mathcal{F})$ 是一个可测空间, $P(\cdot)$ 是一个定义在 $\mathcal{F}$ 上的集函数. 如果 $P(\cdot)$ 满足以下条件:

(1) 非负性: $P(A)\geqslant 0,\forall A\in\mathcal{F}$;

(2) 完全性: $P(\Omega)=1$;

(3) 可列可加性: 对两两互不相容的事件 $A_i\in\mathcal{F}, i=1,2,\cdots$ (即当 $i\neq j$ 时, $A_i\cap A_j=\varnothing$), 有

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

那么称 P 是 $(\Omega,\mathcal{F})$ 上的一个**概率测度**, 简称**概率**. 称 $(\Omega,\mathcal{F},P)$ 为**概率空间**. 称 $P(A)$ 为事件 A 的概率. 本书约定概率为零的事件的任何子集都属于 $\mathcal{F}$. 满足这样约定的概率空间通常称为**完备的概率空间**.

概率具有如下基本性质:

(1) $P(\varnothing)=0, P(A^{\mathrm{c}})=1-P(A),\forall A\in\mathcal{F}$.

(2) 有限可加性: 如果 $A_i\in\mathcal{F}, i=1,2,\cdots,n$, 且两两互不相容, 那么

$$P\Big(\bigcup_{i=1}^{n}A_i\Big)=\sum_{i=1}^{n}P(A_i).$$

(3) 单调性: 如果 $A,B\in\mathcal{F}$, 且 $A\subseteq B$, 那么 $P(A)\leqslant P(B)$.

(4) 进出公式: 如果 $A_i\in\mathcal{F},i=1,2,\cdots,n$, 那么

$$P\Big(\bigcup_{i=1}^{n}A_i\Big)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n+1}P(A_1A_2\cdots A_n).$$

(5) 次可加性: 如果 $A_i\in\mathcal{F},i=1,2,\cdots,n$, 那么

$$P\Big(\bigcup_{i=1}^{n}A_i\Big)\leqslant\sum_{i=1}^{n}P(A_i).$$

下面介绍概率的连续性. 为此引入事件序列的极限.

定义 1.5 若一个事件序列 $\{A_n,n\geqslant1\}\subseteq\mathcal{F}$ 满足 $A_n\subseteq A_{n+1}$ (或 $A_n\supseteq A_{n+1}$), $n\geqslant1$, 则称该事件序列 $\{A_n,n\geqslant1\}$ 为**单调递增事件序列**(或**单调递减事件序列**). 若 $\{A_n,n\geqslant1\}$ 是一个单调递增事件序列, 则定义 $\lim\limits_{n\to\infty}A_n=\bigcup\limits_{n=1}^{\infty}A_n$. 若 $\{A_n,n\geqslant1\}$ 是一个单调递减事件序列, 则定义 $\lim\limits_{n\to\infty}A_n=\bigcap\limits_{n=1}^{\infty}A_n$. 若 $\{A_n,n\geqslant1\}$ 是一个事件序列, 则定义 $\limsup\limits_{n\to\infty}A_n=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{i=n}^{\infty}A_i$.

(6) 概率的连续性: 如果$\{A_n,n\geqslant1\}$是一个单调递增(或单调递减)事件序列, 那么

$$\lim_{n\to\infty}P(A_n)=P\big(\lim_{n\to\infty}A_n\big).$$

证明 设 $\{A_n,n\geqslant1\}$ 是一个单调递增事件序列, 并令

$$B_1=A_1,\quad B_n=A_n-A_{n-1},\quad n\geqslant2.$$

显然 $\{B_n,n\geqslant1\}$ 两两互不相容, 而且对于每个 $n\geqslant1$, 有 $\bigcup\limits_{k=1}^{n}A_k=\bigcup\limits_{k=1}^{n}B_k$, 从而 $\bigcup\limits_{n=1}^{\infty}A_n=\bigcup\limits_{n=1}^{\infty}B_n$. 于是得到

$$\begin{aligned}P\big(\lim_{n\to\infty}A_n\big)&=P\Big(\bigcup_{n=1}^{\infty}A_n\Big)=P\Big(\bigcup_{n=1}^{\infty}B_n\Big)=\sum_{n=1}^{\infty}P(B_n)\\&=\lim_{n\to\infty}\sum_{k=1}^{n}P(B_k)=\lim_{n\to\infty}P\Big(\bigcup_{k=1}^{n}B_k\Big)\\&=\lim_{n\to\infty}P\Big(\bigcup_{k=1}^{n}A_k\Big)=\lim_{n\to\infty}P(A_n).\end{aligned}$$

另一种情况的证明, 留给读者练习. 证毕

(7) Borel-Cantelli 引理: 如果 $\{A_n, n \geqslant 1\}$ 是一个事件序列, 且满足 $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$, 那么

$$P\Big(\limsup_{n\to\infty} A_n\Big) = 0.$$

证明 显然 $\bigcup\limits_{k=n}^{\infty} A_k (n = 1, 2, \cdots)$, 是关于 n 的单调递减事件序列. 根据性质 (5) 和性质 (6) 得

$$0 \leqslant P\Big(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\Big) = P\Big(\lim_{n\to\infty}\bigcup_{k=n}^{\infty} A_k\Big) = \lim_{n\to\infty} P\Big(\bigcup_{k=n}^{\infty} A_k\Big) \leqslant \lim_{n\to\infty}\sum_{k=n}^{\infty} P(A_k) = 0.$$

证毕

事件的独立性是事件间的一种重要关系. 下面介绍事件的独立性概念.

定义 1.6 称两个事件 $A, B \in \mathcal{F}$ **相互独立**, 如果满足

$$P(AB) = P(A)P(B).$$

称三个事件 $A, B, C \in \mathcal{F}$ **相互独立**, 如果满足

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(BC) = P(B)P(C)$$

和

$$P(ABC) = P(A)P(B)P(C).$$

一般地, 称 n 个事件 $A_1, A_2, \cdots, A_n \in \mathcal{F}$ **相互独立**, 如果对于其中任意 k 个事件 $A_{i_1}, A_{i_2}, \cdots, A_{i_k}$ $(1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n)$, 满足

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

称事件序列 $A_1, A_2, \cdots, A_n, \cdots \in \mathcal{F}$ **相互独立**, 如果任取其中有限个均相互独立.

容易证明, 如果事件 $A_1, A_2, \cdots, A_n$ 相互独立, 且令 $\mathcal{F}_1 = \sigma(A_k, 1 \leqslant k \leqslant m)$, $\mathcal{F}_2 = \sigma(A_k, m+1 \leqslant k \leqslant n)$, 那么对于 $B_1 \in \mathcal{F}_1$ 和 $B_2 \in \mathcal{F}_2$, 有 B_1 和 B_2 相互独立.

(8) 如果事件序列 $A_1, A_2, \cdots, A_n, \cdots$ 相互独立, 且 $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$, 那么

$$P\Big(\limsup_{n\to\infty} A_n\Big) = 1.$$

性质 (8) 的证明留给读者.

条件概率是概率论中的重要概念, 用途广泛. 下面介绍一下这个概念.

定义 1.7　设 A 是一个事件, 且 $P(A)>0$, 则称 $P(AB)/P(A)$ 为事件 A 发生下事件 B 的**条件概率**, 记为 $P(B|A)$, 即

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

由条件概率的定义, 易得 $P(AB)=P(B|A)P(A)$, 这称为**乘法公式**. 进一步, 可得更一般的乘法公式

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

借助于条件概率的概念, 容易得到下列全概率公式和贝叶斯公式. 设 $A_1,A_2,\cdots$ 两两互不相容, 且 $\bigcup\limits_n A_n=\Omega$, 则称 $\{A_n\}$ 为 Ω 的一个**分割**.

(9) 全概率公式: 如果事件 $\{A_n\}$ 为 Ω 的一个分割, 且 $P(A_n)>0$, 那么 $\forall B\in\mathcal{F}$,

$$P(B)=\sum_n P(A_n)P(B|A_n).$$

(10) 贝叶斯公式: 如果事件 $\{A_n\}$ 为 Ω 的一个分割, 且 $P(A_n)>0$, 那么对于事件 $P(A)>0$,

$$P(A_n|A)=\frac{P(A_n)P(A|A_n)}{\sum\limits_k P(A_k)P(A|A_k)},\quad n\geqslant 1.$$

容易看出, 如果记 $P_A(\cdot)=P(\cdot|A)$, 那么 $P_A(\cdot)$ 也是可测空间 $(\Omega,\mathcal{F})$ 上的一个概率测度, 从而 $(\Omega,\mathcal{F},P_A(\cdot))$ 也构成概率空间, 而且对于任意的 $A,B,C\in\mathcal{F}$, 当 $P(AB)>0$ 时, 有 $P_A(C|B)=P(C|AB)$; 并且 $P(C|AB)=P(C|A)$ 与 $P(BC|A)=P(B|A)P(C|A)$ 等价. 同样全概率公式也成立, 即如果 $\{B_n\}$ 是 Ω 的一个分割, 那么对于 $C\subseteq\bigcup\limits_n B_n$, 有

$$P_A(C)=\sum_n P_A(B_n)P_A(C|B_n),$$

也即

$$P(C|A)=\sum_n P(B_n|A)P(C|AB_n).$$

1.1.2　随机变量

定义 1.8　设 $(\Omega,\mathcal{F},P)$ 是一个概率空间, X 是定义在 Ω 上, 取值于实数集 $\mathbb{R}$ 的函数. 如果对任意实数 $x\in\mathbb{R}$, $\{\omega:X(\omega)\leqslant x\}\in\mathcal{F}$, 那么称 $X(\omega)$ 是概率空间 $(\Omega,\mathcal{F},P)$ 上的**随机变量**, 简称为随机变量. 函数

$$F(x)=P(\{\omega:X(\omega)\leqslant x\}),\quad -\infty<x<\infty$$

称为随机变量 X 的**分布函数**.

需要说明的是, 定义中样本点 ω 的集合 $\{\omega : X(\omega) \leqslant x\}$ 是一个事件, 一般简记为 $\{X \leqslant x\}$ 或 $\{X \in (-\infty, x]\}$. 容易验证, 对于任意实数 x, $\{X \leqslant x\}, \{X \geqslant x\}, \{X < x\}$ 和 $\{X > x\}$ 中只要有一个属于 σ 域 $\mathcal{F}$, 那么其余的就都属于 $\mathcal{F}$; 而且对于两个随机变量 X 和 Y 而言, $\{X < Y\}, \{X \leqslant Y\}, \{X = Y\}$ 和 $\{X \neq Y\}$ 也都属于 $\mathcal{F}$. 给定随机变量 X, 称包含所有形如 $\{X \leqslant x\}(x \in \mathbb{R})$ 的最小 σ 域为**由随机变量 X 生成的 σ 域** (**或代数**), 记为 $\sigma(X)$. 类似地, 可以定义由随机变量 $X_1, X_2, \cdots, X_n$ 生成的 σ 域 (或代数).

设 X 和 Y 是同一个概率空间 $(\Omega, \mathcal{F}, P)$ 上的两个随机变量, 若 $P(\{X \neq Y\}) = 0$ 成立, 则称它们是**几乎必然相等**, 记为 $X = Y$, a.s.. 对于随机变量 X 和 Y 而言, $X \pm Y$ 和 XY 也都是随机变量. 不难证明, 如果 $\{X_n\}$ 是一个随机变量序列, 那么 $\sup_n X_n, \inf_n X_n, \limsup\limits_{n\to\infty} X_n$ 和 $\liminf\limits_{n\to\infty} X_n$ 都是随机变量.

一般地, 常用的随机变量有两种类型: 离散型随机变量和连续型随机变量.

定义 1.9 若随机变量 X 的可能取值构成一个有限集或可列集, 则称 X 是一个**离散型随机变量**. 若随机变量 X 的可能取值充满数轴上的一个区间 (a, b), 则称其为**连续型随机变量**, 其中 a 可以是 $-\infty$, b 可以是 ∞.

对于一个离散型随机变量 X 而言, 如果 X 的所有可能取值是 $x_1, x_2, \cdots, x_n, \cdots$, 那么 X 取 x_i 的概率

$$p_i = P(\{X = x_i\}), \quad i = 1, 2, \cdots, n, \cdots$$

称为**概率分布列**或简称**分布列**, 其分布函数为

$$F(x) = \sum_{x_k \leqslant x} p_k.$$

对于一个连续型随机变量 X 而言, 存在实数轴上的一个非负可积函数 $f(x)$, 使得其分布函数 $F(x)$ 可表示为

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

此时, 称非负可积函数 $f(x)$ 为 X 的**概率密度函数**, 简称为**密度函数**或**密度**. 密度函数 $f(x)$ 的值虽然不是概率, 但是乘以微分元 $\mathrm{d}x$ 就可得小区间 $(x, x + \mathrm{d}x)$ 上概率的近似值, 即

$$f(x)\mathrm{d}x \approx P(x < X < x + \mathrm{d}x).$$

将相邻微分元累积起来就得到 X 在 (a,b) 上取值的积分, 此积分值就是 X 在 (a,b) 上取值的概率, 即

$$\int_a^b f(x)\mathrm{d}x = P(a < X < b).$$

更一般地, 对于 $(-\infty,\infty)$ 的子集 A, 有

$$P(X \in A) = \int_A f(x)\mathrm{d}x.$$

定义 1.10 如果 $X_1, X_2, \cdots, X_n$ 都是随机变量, 那么称 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 是 n **维随机向 (变) 量**. 称 $\mathbb{R}^n$ 上的 n 元函数

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

为 $\boldsymbol{X}$ 的 n **维联合分布函数**或 n **维分布函数**. 如果有 $\mathbb{R}^n$ 上的非负函数 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$, 使得对于 $\mathbb{R}^n$ 的任何子立方体

$$D = \{(x_1, x_2, \cdots, x_n) : a_i < x_i \leqslant b_i,\ 1 \leqslant i \leqslant n\},$$

有

$$P(\boldsymbol{X} \in D) = \int_D f(\boldsymbol{x})\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n, \tag{1.1}$$

那么称 $\boldsymbol{X}$ 是连续型随机向量, 称函数 $f(\boldsymbol{x})$ 为 $\boldsymbol{X}$ 的**概率密度函数**, 简称为**密度函数**或**密度**.

事实上, 可以证明对于 $\mathbb{R}^n$ 的任何子区域 D, (1.1) 式仍然成立. 而且显然

$$F(\boldsymbol{x}) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f(s_1, s_2, \cdots, s_n)\mathrm{d}s_n\cdots\mathrm{d}s_1.$$

若 $F(\boldsymbol{x})$ 在 $\mathbb{R}^n$ 的开区域 D 中有连续的 n 阶混合偏导数, 且 $P(\boldsymbol{X} \in D) = 1$, 则

$$f(\boldsymbol{x}) = \begin{cases} \dfrac{\partial^n F(\boldsymbol{x})}{\partial x_n \cdots \partial x_1}, & \boldsymbol{x} \in D, \\ 0, & \text{其他}. \end{cases}$$

在后面章节里, 要经常用到一个随机向量变换的重要公式, 这就是下面的定理.

定理 1.1 设随机向量 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 的概率密度函数是 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$, $Y_i = g_i(\boldsymbol{X})(i = 1, 2, \cdots, n)$ 是 $\boldsymbol{X}$ 的函数, 且存在唯一的反函数

$X_i = h_i(\boldsymbol{Y}) = h_i(Y_1, Y_2, \cdots, Y_n)$, $i = 1, 2, \cdots, n$. 如果 g_i, h_i 有连续偏导数, 那么 $\boldsymbol{Y} = (Y_1, Y_2, \cdots, Y_n)$ 的概率密度函数为

$$g(\boldsymbol{y}) = \begin{cases} f(\boldsymbol{x})|\boldsymbol{J}|, & \text{若 } \boldsymbol{y} \text{ 属于 } (g_1, g_2, \cdots, g_n) \text{ 的值域, 且 } |\boldsymbol{J}| \neq 0, \\ 0, & \text{其他.} \end{cases}$$

其中, $\boldsymbol{y} = (y_1, y_2, \cdots, y_n), \boldsymbol{x} = (x_1, x_2, \cdots, x_n), x_i = h_i(\boldsymbol{y})$, 而 $\boldsymbol{J}$ 为坐标变换的雅可比矩阵, 即

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix}.$$

由随机向量的联合分布函数, 可以得到如下边际 (或边缘) 分布函数的概念.

定义 1.11 设 $F(\boldsymbol{x})$ 为随机向量 $\boldsymbol{X}$ 的联合分布函数. 对于 $1 \leqslant i_1 < i_2 < \cdots < i_m \leqslant n$, 我们称如下函数

$$F_{i_1, \cdots, i_m}(x_{i_1}, \cdots, x_{i_m}) = F(\infty, \cdots, \infty, i_1, \infty, \cdots, \infty, i_2, \infty, \cdots, \infty, i_m, \infty, \cdots, \infty)$$

为随机向量 $\boldsymbol{X}$ 的 m **维边际 (或边缘) 分布函数**.

定义 1.12 设随机向量 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 的分布函数为 $F(\boldsymbol{x})$. 若 $\forall \boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$, 有

$$F(\boldsymbol{x}) = F_1(x_1) F_2(x_2) \cdots F_n(x_n), \quad \text{其中} \quad F_i(x_i) = P(X_i \leqslant x_i),$$

则称 $X_1, X_2, \cdots, X_n$ **相互独立**.

可以证明, 若随机变量 X, Y, Z 相互独立, 则 $X \pm Y$, $X \cdot Y$, X/Y $(Y \neq 0)$ 均与 Z 相互独立. 更一般地, $g(X, Y)$ 与 $h(Z)$ 相互独立.

下面回顾一下常用随机变量的分布.

(1) 离散型随机变量的典型分布

二项分布 设 $0 \leqslant p < 1, n \geqslant 1$. 若随机变量 X 的分布列为

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

则称 X 的分布为参数是 (n, p) 的**二项分布**, 简记为 $X \sim B(n, p)$.

泊松分布　设 $\lambda > 0$. 若随机变量 X 的分布列为

$$P(X=k)=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\quad k=0,1,2,\cdots,$$

则称 X 的分布为参数是 λ 的**泊松分布**, 简记为 $X \sim P(\lambda)$.

几何分布　设 $0 < p < 1$. 若随机变量 X 的分布列为

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

则称 X 的分布为参数是 p 的**几何分布**, 简记为 $X \sim G(p)$.

(2) 连续型随机变量的典型分布

均匀分布　设 $a < b$. 若随机变量 X 的概率密度函数为

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

则称 X 的分布为 (a,b) 上的**均匀分布**, 简记为 $X \sim U(a,b)$.

正态分布　设 $\mu \in \mathbb{R}, \sigma > 0$. 若随机变量 X 的概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\{-(x-\mu)^2/2\sigma^2\},$$

则称 X 的分布为参数是 (μ,σ^2) 的**正态分布**, 简记为 $X \sim N(\mu,\sigma^2)$.

指数分布　设 $\lambda > 0$. 若随机变量 X 的概率密度函数为

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

则称 X 的分布为参数是 λ 的**指数分布**, 简记为 $X \sim \mathrm{Exp}(\lambda)$.

n **维正态分布**　设 $\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_n)$, $\boldsymbol{\Sigma}$ 是 n 阶正定对称阵, 并且其行列式为 $|\boldsymbol{\Sigma}|$. 如果 n 维随机向量 $\boldsymbol{X}$ 的概率密度函数为

$$f(\boldsymbol{x})=(2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\right],$$

其中 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$, 那么称 $\boldsymbol{X}$ 的分布为 n **维正态分布**, 简记为 $\boldsymbol{X} \sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$.

例 1.2 设 $Z_1, Z_2, \cdots, Z_n$ 是独立同分布、非负的随机变量, 密度函数为 $f(z)$, 它们相应的次序统计量为 $Z_{(1)} \leqslant Z_{(2)} \leqslant \cdots \leqslant Z_{(n)}$, 请写出次序统计量 $Z_{(1)}, Z_{(2)}, \cdots, Z_{(n)}$ 的联合概率密度.

解 对于任意给定的正数 $z_1 < z_2 < \cdots < z_n$, 取充分小的 $h > 0$, 使得

$$0 < z_1 < z_1 + h < z_2 < z_2 + h < \cdots < z_{n-1} + h < z_n < z_n + h.$$

考虑到

$$\{z_k < Z_{(k)} \leqslant z_k + h,\ 1 \leqslant k \leqslant n\} = \bigcup_{(i_1, i_2, \cdots, i_n)} \{z_j < Z_{i_j} \leqslant z_j + h,\ 1 \leqslant j \leqslant n\},$$

且上式右边事件互不相容, 于是

$$\begin{aligned}
&\lim_{h \to 0} \frac{P(z_1 < Z_{(1)} \leqslant z_1 + h, z_2 < Z_{(2)} \leqslant z_2 + h, \cdots, z_n < Z_{(n)} \leqslant z_n + h)}{h^n} \\
= &\lim_{h \to 0} n! \frac{P(z_1 < Z_{i_1} \leqslant z_1 + h, z_2 < Z_{i_2} \leqslant z_2 + h, \cdots, z_n < Z_{i_n} \leqslant z_n + h)}{h^n} \\
= &\lim_{h \to 0} n! \prod_{j=1}^{n} \frac{P(z_j < Z_{i_j} \leqslant z_j + h)}{h} = n! \prod_{j=1}^{n} f(z_j).
\end{aligned}$$

故而 $Z_{(1)}, Z_{(2)}, \cdots, Z_{(n)}$ 的联合概率密度为

$$f(z_1, z_2, \cdots, z_n) = \begin{cases} n! \displaystyle\prod_{j=1}^{n} f(z_j), & 0 < z_1 < z_2 < \cdots < z_n, \\ 0, & \text{其他}. \end{cases} \tag{1.2}$$

由 (1.2) 式 知, 若 $Z_i (1 \leqslant i \leqslant n)$ 是独立同分布的随机变量, 而且共同的分布为在 $[0,\ t]$ 上的均匀分布, 则其顺序统计量 $Z_{(1)}, Z_{(2)}, \cdots, Z_{(n)}$ 的联合概率密度为

$$f(z_1, z_2, \cdots, z_n) = \begin{cases} \dfrac{n!}{t^n}, & 0 < z_1 < z_2 < \cdots < z_n \leqslant t, \\ 0, & \text{其他}. \end{cases} \tag{1.3}$$

1.1.3 黎曼-斯蒂尔切斯积分

为了统一表述随机变量的数字特征, 我们引入黎曼-斯蒂尔切斯积分. 设 $G(x)$ 为 $(-\infty, \infty)$ 上单调不减右连续函数, $h(x)$ 为 $(-\infty, \infty)$ 上的单值实函数. 对于 $\forall a < b$, 我们可以如微积分中定义定积分那样定义关于 $G(x)$ 的黎曼-斯蒂尔切斯积分.