

SCNUSoS  
《概率论与数理统计》讲义



EVAN GAO

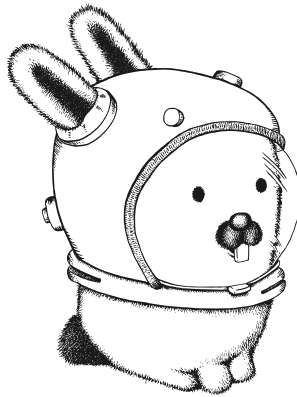
大二上



Version D:20231017

Made with X<sub>Y</sub>L<sup>A</sup>T<sub>E</sub>X  
(And Love)

evan@notchlp.xyz



# CONTENTS

## CHAPTER 1 随机事件及其概率 PAGE 3

1.1	随机事件	3
1.2	频率	3
1.3	概率	4
	概率的性质 — 4	
1.4	条件概率	4
	乘法公式 — 5 • 全概率公式 — 6 • Bayes 公式 — 7	
1.5	事件的独立性	8
1.6	Bernoulli 试验	9

## CHAPTER 2 随机变量及其分布 PAGE 10

2.1	随机变量	10
2.2	离散型随机变量及其分布	10
	常用离散分布 — 10	
2.3	累积分布函数	11
	随机变量的分布函数 — 11 • 离散型随机变量的分布函数 — 12	
2.4	连续型随机变量及其概率密度	12
	常用连续型分布 — 13	
2.5	随机变量函数的分布	16
	离散型随机变量函数的分布 — 16 • 连续型随机变量函数的分布 — 16	

## CHAPTER 3 多维随机变量及其分布 PAGE 18

3.1	二维随机变量及其分布	18
	二维随机变量 — 18 • 二维随机变量的分布函数 — 18	

# Chapter 1

## 随机事件及其概率

### 1.1 随机事件

事件是样本空间的子集

#### Definition 1.1.1: 样本空间

考虑样本空间集合  $S$ , 我们有  $S := \{\text{所有样本点}\}$ .

由定义, 我们可以得到几种特殊的样本空间:

- $\emptyset$  事件: 不可能发生的事件.
- $S - \emptyset$  发生的事件.
- 基本事件  $\omega : |\omega| = 1$  i.e. 基本事件只含有一个样本点.

#### Theorem 1.1.1 De Morgan 律

由于  $\emptyset$  事件和  $S - \emptyset$  事件是互为对偶的, 我们可以得到

$$\begin{aligned}\overline{A \cup B} &= \bar{A} \cap \bar{B} \\ \overline{A \cap B} &= \bar{A} \cup \bar{B}\end{aligned}\tag{1.1}$$

我们定义**和, 积事件**为

$$\begin{aligned}\bigcup_{i=1}^n A_i \text{ happen} &\iff \exists i \in [1, n] \text{ s.t. } A_i \text{ happens} \\ \bigcap_{i=1}^n A_i \text{ happen} &\iff \forall i \in [1, n] \text{ s.t. } A_i \text{ happens}\end{aligned}\tag{1.2}$$

### 1.2 频率

#### Definition 1.2.1: 频率

考虑事件  $A$ , 其发生的**频率**是

$$f_n(A) \stackrel{\text{def}}{=} \frac{r_n(A)}{n} \in [0, 1]$$

其中频数  $\frac{r_n(A)}{n} \in [0, n]$ . 显然有  $f_n(S) = 1$ .

### Corollary 1.2.1 有限可加性

若  $A_i \cap A_j = \emptyset$  且  $i \neq j$ ,  $i, j \in [1, k]$  i.e. 互斥事件, 则

$$f_n \left( \bigcap_{i=1}^k A_i \right) = \sum_{i=1}^k f_n(A_i)$$

## 1.3 概率

### Definition 1.3.1: 概率

(Kolmogorov 公理化定义) 设有随机试验  $E$  且与之对应的样本空间  $S$ , 考虑事件  $A$

for  $\forall A \in E$ , if

- ①  $0 \leq P(A) \leq 1$
- ②  $P(S) = 1$
- ③  $\Pr\{\bigcup_{i=1}^{\infty} A_i\} = \sum_{i=1}^{\infty} \Pr\{A_i\}$  i.e. 可列可加性

则称  $P$  为  $S$  上的**概率**.

### 1.3.1 概率的性质

由上面的定义我们能够得到概率的性质：

1.  $\Pr\{\emptyset\} = 0$
2.  $\Pr\{\bigcup_{i=1}^n A_i\} = \sum_{i=1}^n \Pr\{A_i\} \iff \forall i, j (i \neq j \rightarrow A_i A_j = \emptyset)$
3.  $\Pr\{\overline{A}\} + \Pr\{A\} = 1$
4.  $\Pr\{A - B\} = \Pr\{A\} - \Pr\{AB\} \Rightarrow (A - B) \cap B = \emptyset$
5. (单调性)  $B \subseteq A \Rightarrow \Pr\{B\} \leq \Pr\{A\}$
6. 若满足 5, 由 4 可得  $\Pr\{A - B\} = \Pr\{A\} - \Pr\{B\}$
7. (容斥原理)  $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{AB\}$

## 1.4 条件概率

### Definition 1.4.1: 条件概率

设  $AB$  是两个事件, 且  $P(A) \neq 0$ , 则称

$$\Pr\{A|B\} \stackrel{\text{def}}{=} \frac{\Pr\{AB\}}{\Pr\{B\}} \quad (1.3)$$

为在事件  $A$  发生的条件下, 事件  $B$  的**条件概率**.

### Corollary 1.4.1 条件概率之性质

概率满足的性质条件概率都满足.

### Theorem 1.4.1

设  $A_1, A_2, \dots, A_n$  是  $n$  个互斥事件, 则有

$$\Pr\left\{\bigcup_{i=1}^n A_i \mid A\right\} = \sum_{i=1}^n \Pr\{A_i \mid A\} \quad (1.4)$$

### Question 1: 证明

$$\Pr\{\bar{B} \mid A\} = 1 - \Pr\{B \mid A\} \quad (1.5)$$

Proof:

$$\begin{aligned} \Pr\{\bar{B} \mid A\} &= \frac{\Pr\{\bar{B}A\}}{\Pr\{A\}} \\ &= \frac{\Pr\{A\} - \Pr\{BA\}}{\Pr\{A\}} \\ &= 1 - \frac{\Pr\{BA\}}{\Pr\{A\}} \\ &= 1 - \Pr\{B \mid A\} \end{aligned}$$



## 1.4.1 乘法公式

由条件概率的定义, 我们可以得到**乘法公式** :

### Theorem 1.4.2 乘法公式

$$\begin{aligned} \Pr\{A_1 A_2 \cdots A_n\} &= \Pr\{A_1\} \Pr\{A_2 \mid A_1\} \Pr\{A_3 \mid A_1 A_2\} \cdots \Pr\{A_n \mid A_1 A_2 \cdots A_{n-1}\} \\ &= \prod_{i=1}^n \Pr\left\{A_i \mid \bigcup_{i=1}^{n-1} A_i\right\} \end{aligned} \quad (1.6)$$

### Example 1.4.1 (“买彩票”)

第一次买中的概率为  $\frac{1}{2}$ , 第二次买中而第一次未中的概率是  $\frac{7}{10}$ , 第三次买中而前两次未中的概率是  $\frac{9}{10}$ , 求三次都未中的概率.

**Solution:** 以  $A_i (i = 1, 2, 3)$  表示事件「第  $i$  次买中」, 以  $B$  表示事件「三次都未中」, 那么由(1.6)

$$\begin{aligned} \because B &= \overline{A_1 A_2 A_3} \\ \therefore \Pr\{B\} &= \Pr\{\overline{A_1 A_2 A_3}\} \\ &= \Pr\{\overline{A_1}\} \Pr\{\overline{A_2} | \overline{A_1}\} \Pr\{\overline{A_3} | \overline{A_1 A_2}\} \\ &= \left(1 - \frac{1}{2}\right) \left(1 - \frac{7}{10}\right) \left(1 - \frac{9}{10}\right) \\ &= \frac{3}{200} \end{aligned}$$

#### Question 2: P28-19

袋中装有  $a$  个红球,  $b$  个白球, 每次自袋中有放回地任取一球, 并同时再放入  $m$  个与之相同的球, 连续如此进行  $2n$  次, 求前  $n$  次为红球, 后  $n-1$  次为白球, 第  $2n$  次为红球的概率.

**Solution:** 设事件  $A_i$  表示第  $i$  次取出的球为红球,  $B_i$  表示白球, 那么可知所求概率为

$$\Pr\{A_1 A_2 A_3 \cdots A_n B_{n+1} B_{n+2} \cdots B_{2n-1} A_{2n}\}$$

利用(1.6)展开上式, 我们立即有

$$\Pr\{A_1\} \Pr\{A_2 | A_1\} \cdots \Pr\{A_n | A_1 A_2 \cdots A_{n-1}\} \times \Pr\{B_{n+1} | A_1 A_2 \cdots A_n\} \Pr\{B_{n+2} | \cdots B_{n+1}\} \\ \cdots \Pr\{A_{2n} | A_1 \cdots A_n B_{n+1} B_{n+2} \cdots B_{2n-1}\}$$

此式太长, 不妨将其分成三部分:

1.  $\Pr\{A_1\} \Pr\{A_2 | A_1\} \cdots \Pr\{A_n | A_1 A_2 \cdots A_{n-1}\}$  (前  $n$  个红球)
2.  $\Pr\{B_{n+1} | A_1 A_2 \cdots A_n\} \Pr\{B_{n+2} | \cdots B_{n+1}\} \cdots \Pr\{B_{2n-1} | \cdots B_{2n-2}\}$  (后  $n-1$  个白球)
3.  $\Pr\{A_{2n} | A_1 \cdots A_n B_{n+1} B_{n+2} \cdots B_{2n-1}\}$  (最后的红球)

三式相乘, 有

$$\prod_{i=1}^n \frac{a + (i-1)m}{a + b + (i-1)m} \prod_{i=n+1}^{2n-1} \frac{b + (i-n-1)m}{a + b + (i-1)m} \times \frac{a + (n-1)m}{a + b + (2n-1)m}$$

即为所求.

### 1.4.2 全概率公式

#### Lemma 1.4.1 完备事件组

设  $A_1, A_2, \cdots, A_n$  是有限或可数个事件, 若其满足 1. 两两互斥, 2.  $\bigcup_{i=1}^n A_i = S$ , 则称  $A_1, A_2, \cdots, A_n$  是一个**完备事件组**.

#### Theorem 1.4.3 全概公式

设  $A_1, A_2, \cdots, A_n$  是一个完备事件组, 且  $\Pr\{A_i\} > 0$  for  $i$  in  $1..n$ , 则对任一事件  $B$ , 有

$$\Pr\{B\} = \sum_{i=1}^n \Pr\{A_i\} \Pr\{B | A_i\} \quad (1.7)$$

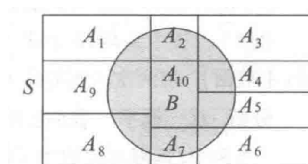


图 1-4-2

Figure 1.1: 全概公式示意图

### Question 3: P14-10

从 1 到 9 的整数中有放回地依次随机抽取 3 次, 求取出的 3 个数之积能被 10 整除的概率.

显然这三个数中必有两数为 5 和 {2, 4, 6, 8}, 因此

法一: 分情况讨论

1.  $A = \{\text{三个数里有两个 5 和一个偶数}\}$
2.  $B = \{\text{三个数里有一个 5 和两个偶数}\}$
3.  $C = \{\text{三个数里有一个 5 一个偶数和一个其他的奇数}\}$

那么有  $\Pr\{A\} = \frac{\binom{4}{1} \cdot 3}{9^3}$   $\Pr\{B\} = \frac{\binom{4}{1} \cdot 3 + \binom{4}{2} A_3^2}{9^3}$   $\Pr\{C\} = \frac{\binom{4}{1} \binom{4}{1} A_3}{9^3}$  相加得  $\frac{156}{729}$ .

法二: 对立事件

不妨设三个数中出现 5 的事件为  $A$ , 出现偶数的事件为  $B$ , 那么

$$\begin{aligned} \Pr\{AB\} &= 1 - \Pr\{\overline{AB}\} \\ &= 1 - (\Pr\{\overline{A}\} + \Pr\{\overline{B}\} - \Pr\{\overline{A}\overline{B}\}) \\ &= 1 - \left( \frac{8^3}{9^3} + \frac{5^3}{9^3} - \frac{4^3}{9^3} \right) \\ &= \frac{156}{729} \end{aligned}$$

法三: 另一种分情况讨论 (笔者的分法)

1.  $A = \{\text{三个数中有两个相同的}\}$
2.  $B = \{\text{三个数全不同}\}$ 
  - 2.1.  $B_1 = \{\text{三个数全不同且有两个偶数}\}$
  - 2.2.  $B_2 = \{\text{三个数全不同且有两个奇数}\}$

此项须要分成下述  
两项, 否则会重

那么有  $\Pr\{A\} = \frac{\binom{4}{1} \binom{3}{1} + \binom{4}{1} \binom{3}{1}}{9^3}$   $\Pr\{B_1\} = \frac{\binom{4}{1} \binom{4}{1} \binom{3}{1}}{9^3}$   $\Pr\{B_2\} = \frac{\binom{4}{1} \binom{4}{1} A_3}{9^3}$  相加得  $\frac{156}{729}$ .

#### Note

讨论各种情况的概率进而求得所求事件的概率的方法实际上是**全概率公式**的一种体现.

### 1.4.3 Bayes 公式

**全概公式**是通过计算某一事件会发生的所有原因和情况的可能性大小来计算该事件发生的概率, 而**Bayes 公式**则与之相反, 考察一件已经发生的事情的各种原因或或情况的可能性大小.



$\Pr\{A_i\}, \Pr\{A_i|B\}$   
分别称为原因的  
先验概率和后验概率

### Theorem 1.4.4 Bayes 公式

设  $A_1, A_2, \dots, A_n$  是一个完备事件组, 且  $\Pr\{A_i\} > 0$  for  $i$  in  $1..n$ , 则对任一可能发生的事件  $B$ , 有

$$\Pr\{A_i|B\} = \frac{\Pr\{A_i B\}}{\Pr\{B\}} = \frac{\Pr\{A_i\} \Pr\{B|A_i\}}{\sum_{j=1}^n \Pr\{A_j\} \Pr\{B|A_j\}} \quad (1.8)$$

特别地, 当  $n = 2$  时, Bayes 公式也可以写成

$$\Pr\{A|B\} = \frac{\Pr\{A\} \Pr\{B|A\}}{\Pr\{B\}} \quad (1.9)$$

## 1.5 事件的独立性

### Definition 1.5.1: 两事件独立性

设  $A, B$  是两个事件, 若

$$\Pr\{AB\} = \Pr\{A\} \Pr\{B\}$$

则称  $A, B$  相互独立.

### Theorem 1.5.1

设  $A, B$  两事件相互独立且  $\Pr\{B\} > 0$ , 则有

$$\Pr\{A|B\} = \Pr\{A\}$$

### Theorem 1.5.2

若  $A, B$  两事件相互独立, 则它们对立事件和其本身 (不同事件间) 的组合也相互独立.

根据三个事件 (略) 独立性的定义, 可以类推到  $n$  个事件的独立性: 设  $A_1, A_2, \dots, A_n$  是  $n$  个事件, 若对于其中任意  $k (2 \leq k \leq n)$  个事件  $A_{i_1}, A_{i_2}, \dots, A_{i_k}$ , 有

$$\Pr\{A_{i_1} A_{i_2} \dots A_{i_k}\} = \Pr\{A_{i_1}\} \Pr\{A_{i_2}\} \dots \Pr\{A_{i_k}\} \quad (1.10)$$

则称  $A_1, A_2, \dots, A_n$  相互独立.

### Definition 1.5.2

设  $A_1, A_2, \dots, A_n (n > 2)$  是  $n$  个事件, 若其中任意两个事件相互独立, 则称这  $n$  个事件两两独立.

由此可以得到多个独立事件所具备的性质

### Corollary 1.5.1

设  $A_1, A_2, \dots, A_n (n > 2)$  相互独立, 则其中任意  $k (2 \leq k \leq n)$  个 (它们的对立) 事件也相互独立.

## 1.6 Bernoulli 试验

即两点分布.

### Theorem 1.6.1 Bernoulli 定理

一次试验中, 事件  $A$  发生的概率为  $p$ , 进行这样的试验  $n$  次, 事件  $A$  发生  $k$  次的概率为

$$b(k; n; p) = \binom{n}{k} p^k (1-p)^{n-k}$$

这相当于在实验结果序列中任取  $k$  次发生  $A$  事件

#### Note

$n$  次 Bernoulli 试验的概率分布就是二项分布. 其概率密度函数为

$$f(k, n, p) = \Pr(k; n, p) = \Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

**Glossary:** 两点分布 = Bernoulli 试验; 二项分布 = 多次 Bernoulli 试验 = Bernoulli 概型.

**辨析:**  $\Pr\{A\}$  和  $\Pr(X)$  使用的括号不同. 大括号代表是事件, 即一系列样本点集合 (故用大括号), 小括号代表是随机变量 (因为此时是概率密度函数). ☺

# Chapter 2

## 随机变量及其分布

### 2.1 随机变量

#### Definition 2.1.1: 随机变量

设随机试验的样本空间为  $S$ , 称定义在  $S$  上的实值单值函数  $X = X(\omega)$  为**随机变量**. 亦即

$$X : \omega \mapsto x_i \in \mathbb{R} \text{ i.e. } S \rightarrow \mathbb{R}$$

实际上,  $\Pr(X = a)$  是  $A = \{\omega | X(\omega) = a\}$  的简记.

### 2.2 离散型随机变量及其分布

设  $X$  是一个随机变量, 如果  $X$  的所有可能取值为有限个或可列无限多个, 则称  $X$  为**离散型随机变量**.

#### Definition 2.2.1: 离散型随机变量分布律

设  $X$  是一个离散型随机变量, 如果

$$\Pr(X = x_k) = p_k, k = 1, 2, \dots$$

则称  $p_k$  为  $X$  的**分布律**或**概率分布**, 亦称**概率质量函数(PMF)**. 记为  $f_X(x)$ .

就是高中所学之分布列.

#### 2.2.1 常用离散分布

##### 两点分布

#### Definition 2.2.2: 两点分布

设  $X$  的分布律为

$$\Pr(X = k) = \begin{cases} p, & k = x_1 \\ 1 - p, & k = x_2 \end{cases} \quad (2.1)$$

其中  $0 < p < 1$ , 则称  $X$  服从以  $p$  为参数的**两点分布**.

若  $X$  服从  $x_1 = 1, x_2 = 0$  处参数为  $p$  的两点分布, 则称其服从参数为  $p$  的**0-1 分布**.

## 二项分布

当  $n = 1$  时, 二项分布退化为 0-1 分布.

### Definition 2.2.3: 二项分布

设  $X$  的分布律为

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n \quad (2.2)$$

其中  $n$  为正整数,  $0 < p < 1$ , 则称  $X$  服从参数为  $n, p$  的**二项分布**, 记为  $X \sim B(n, p)$ .

二项概率总存在一个最大值  $M$  s.t.  $(n+1)p - 1 \leq M < (n+1)p$ .

$\sim$  记号就是**服从**的意思

## Poisson 分布

### Definition 2.2.4: Poisson 分布

设  $X$  的分布律为

$$\Pr(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots \quad (2.3)$$

其中  $\lambda > 0$ , 则称  $X$  服从参数为  $\lambda$  的**Poisson 分布**, 记为  $X \sim P(\lambda)$ .

因为  $e^x = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!}$ , 将  $x$  换为  $\lambda$  可得泊松分布的概率和为 1.

## 2.3 累积分布函数

### 2.3.1 随机变量的分布函数

#### Definition 2.3.1: 随机变量 CDF

设  $X$  是一个随机变量, 对任意实数  $x$ , 定义

$$F(x) \stackrel{\text{def}}{=} \Pr(X \leq x) \in [0, 1]$$

称  $F(x)$  为  $X$  的**累积分布函数**(CDF). 有时记作  $F_X(x)$  或  $X \sim F(x)$ .

分布函数  $F(x)$  的值就表示  $X$  落在区间  $(-\infty, x]$  的概率.

#### 性质

1. 单调非减性:  $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$
2.  $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$ ,  $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$ .
3. 右连续性:  $\lim_{x \rightarrow x_0^+} F(x) = F(x_0)$

具有上述性质的函数一定是某个随机变量的 CDF.

### 2.3.2 离散型随机变量的分布函数

在上述基础上, 对于离散型随机变量  $X$  有 CDF:

$$F(x) = \Pr(X \leq x) = \sum_{x_i \leq x} \Pr(X = x_i)$$

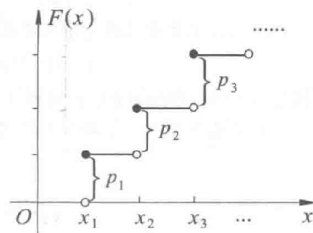


图 2-3-1

Figure 2.1: 阶梯型 CDF

$F(x)$  是一个**阶梯型函数**, 在点  $x_i$  处的跃度为  $\Pr(X = x_i)$ . 可以看出阶梯型函数符合 CDF 之性质: 总是右连续的; 而作为一个分段函数来说, 其各段区间总是左闭右开的.

#### Note

考题中由 CDF 反推分布列可以根据每一个左侧端点得到对应的随机变量  $x_i$ . 计算分布列时, 有公式  $\Pr\{a \leq X < b\} = F(b) - F(a)$ .

随机变量  $X$  的 CDF 为阶梯型函数  $\Leftrightarrow X$  是离散型随机变量.

## 2.4 连续型随机变量及其概率密度

### Definition 2.4.1: PDF

设  $X$  是一个随机变量, 如果存在非负可积函数  $f(x)$ , 使对任意实数  $x$  有

$$F(x) = \Pr\{X \leq x\} = \int_{-\infty}^x f(t)dt \quad (2.4)$$

则称  $X$  为**连续型随机变量**,  $f(x)$  为  $X$  的**概率密度函数(PDF)**.

立即可以得到 PDF 的性质

1.  $f(x) \geq 0$ ;
2.  $\int_{-\infty}^{+\infty} f(x)dx = 1$ .
3.  $F'(x) = f(x)$ . 若  $f(x)$  在  $x$  处连续

#### Note

若某个随机变量的 CDF 可以写成 PDF 积分函数的形式, 则称其为连续型随机变量.

证明  $X$  取任一特值的概率为 0:

$$\text{考虑极限 } \Pr\{X = a\} = \lim_{\Delta x \rightarrow 0^+} \Pr\{a - \Delta x \leq X \leq a\}$$

对 RHS 积分, 有

$$\begin{aligned} \Pr\{X = a\} &= \lim_{\Delta x \rightarrow 0^+} \int_{a-\Delta x}^a f(x)dx \\ &= 0. \end{aligned}$$

由上面的证明我们可以知道以下的性质

$$\Pr\{a < X < b\} = F(b) - F(a) = \int_a^b f(x)dx. \quad (2.5)$$

这个等式无论  $X$  两侧端点开闭都成立.

## 2.4.1 常用连续型分布

### 均匀分布

#### Definition 2.4.2: 均匀分布

设  $X$  的概率密度为

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{else} \end{cases} \quad (2.6)$$

U for Uniform

其中  $a < b$ , 则称  $X$  服从参数为  $a, b$  的**均匀分布**, 记为  $X \sim U(a, b)$ .

设  $[a, b]$  之间的任意子区间为  $[c, d]$ , 由上式可得  $\Pr\{c < X < d\} = \int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}$ . 因而我们可以得到

#### Corollary 2.4.1 均匀分布的 CDF

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases} \quad (2.7)$$

### 指数分布

#### Definition 2.4.3: 指数分布

若随机变量  $X$  的概率密度为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}, \lambda > 0 \quad (2.8)$$

则称  $X$  服从参数为  $\lambda$  的**指数分布**, 记为  $X \sim E(\lambda)$ .

积分得

#### Corollary 2.4.2 指数分布的 CDF

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2.9)$$

## 正态分布

### Definition 2.4.4: 正态分布

若随机变量  $X$  的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (-\infty < x < +\infty) \quad (2.10)$$

其中  $\mu, \sigma (\sigma > 0)$  为常数, 则称  $X$  服从参数为  $\mu, \sigma^2$  的**正态分布**, 记为  $X \sim N(\mu, \sigma^2)$ .

### Question 4: Poisson 积分

不使用 Wallis 公式, 求 Poisson 积分:  $\int_{-\infty}^{+\infty} e^{-t^2} dt$ .

**Solution:** 转化为求累次积分

要求原式, 不妨先求其平方, 亦即

$$\left( \int_{-\infty}^{+\infty} e^{-t^2} dt \right)^2 = \int_{-\infty}^{+\infty} e^{-x^2} dx \int_{-\infty}^{+\infty} e^{-y^2} dy$$

化为累次积分有

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy.$$

相当于在平面内求此积分, 化为极坐标有

$$\begin{aligned} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\rho^2} \rho d\rho d\theta &= \int_0^{2\pi} d\theta \int_0^{+\infty} e^{-\rho^2} \rho d\rho \\ &= \pi \end{aligned}$$

$$\therefore \int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi}. \quad (2.11)$$

通过证明 Poisson 积分, 我们可以得到正态分布的 PDF 重要性质

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1 \quad (2.12)$$

## 图形特征

$\sigma$  越小, 方差越大,  
曲线中峰越陡

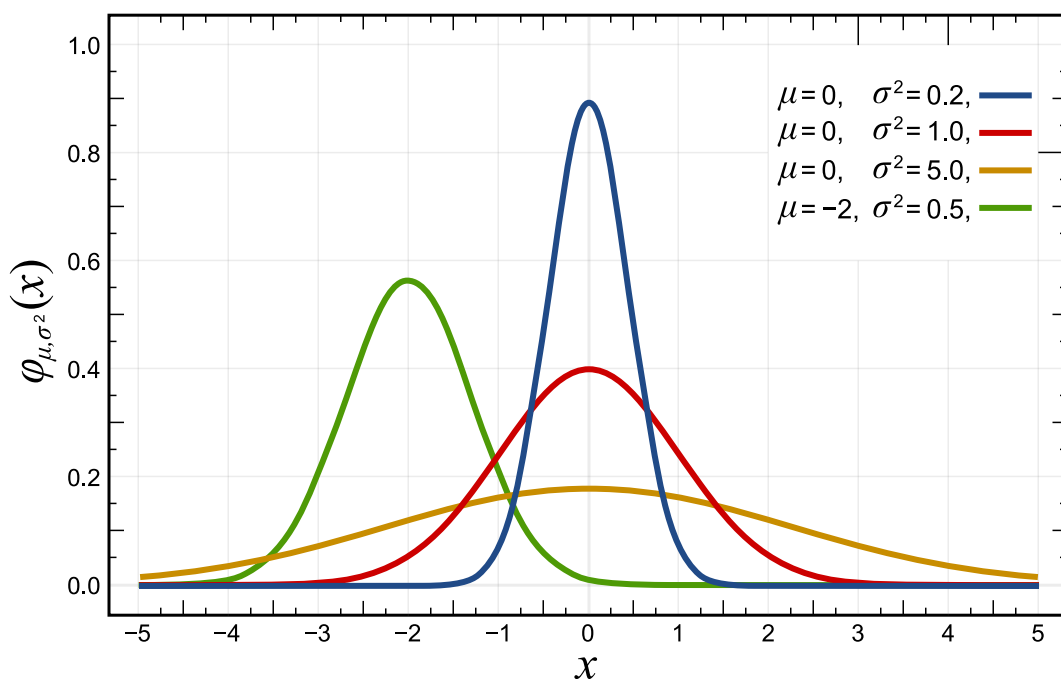


Figure 2.2: PDF

1.  $\mu$  确定曲线位置,  $\sigma$  确定曲线中峰的陡峭程度
2. 密度曲线关于  $x = \mu$  对称
3. 密度曲线在  $x = \mu$  处有最大值  $\frac{1}{\sqrt{2\pi}\sigma}$
4. 密度曲线在  $x = \mu \pm \sigma$  处有拐点且以  $x$  轴为渐近线

当  $\mu = 0, \sigma = 1$  时, 称为**标准正态分布**, 记为  $X \sim N(0, 1)$ , PDF 用  $\varphi(x)$  表示, CDF 用  $\Phi(x)$  表示:

$\varphi(x)$  为偶函数.

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \Phi(x) = \int_{-\infty}^x \varphi(t) dt \quad (2.13)$$

### Theorem 2.4.1

设  $X \sim N(\mu, \sigma^2)$ , 则  $Y = \frac{X-\mu}{\sigma} \sim N(0, 1)$ .

计算  $\Pr\{Y \leq x\}$  时, 可转化为计算  $\Pr\{X \leq \mu + \sigma x\}$ , 计算此时的  $\Phi(x)$  即得证.

这说明任意一般的正态分布都可以通过**线性变换**转化为标准正态分布. 因此  $X$  的分布函数也可以写成

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

$$\Pr\{a < X < b\} = \Pr\left\{\frac{a - \mu}{\sigma} < Y < \frac{b - \mu}{\sigma}\right\} = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$



## 2.5 随机变量函数的分布

前面提到 Theorem 2.4.1. 其中的  $Y = \frac{X-\mu}{\sigma}$  是随机变量  $X$  的函数, 因此  $Y$  也是随机变量.

### Definition 2.5.1: 随机变量函数

若存在一个函数  $g(x)$  s.t. 随机变量  $X, Y$  满足

$$Y = g(X)$$

称  $Y$  为  $X$  的一个**随机变量函数**.

### 2.5.1 离散型随机变量函数的分布

根据上述定义, 显然  $X$  的函数  $Y$  也是一个随机变量.

#### 导出 $Y$ 的分布律

首先确定  $Y$  的所有取值, 通过  $Y$  的每一个值  $y_i (i = 1, 2, \dots)$  确定相应的  
(注: 我们用  $C_i$  来表示  $X$  的取值为  $x_i$  时, 与对应  $Y$  匹配的  $X$  的取值的集合)

$$C_i = \{x_i | g(x_j) = y_i\},$$

$$\{Y = y_i\} = \{X \in C_i\},$$

那么有

$$\Pr\{Y = y_i\} = \Pr\{X \in C_i\} = \sum_{x_j \in C_i} \Pr\{X = x_j\} \quad (2.14)$$

为  $Y$  的分布律. 这表明:  $Y$  的分布律完全由  $X$  的分布律确定.

### 2.5.2 连续型随机变量函数的分布

分布律: 离散变量  
PDF: 连续变量

通过上述定义, 不难发现连续随机变量的函数不一定是连续型随机变量.

#### Note

考虑**符号函数**作为映射法则:

$$Y = \text{sign}(X) = \begin{cases} 1, & X > 0 \\ 0, & X = 0 \\ -1 & X < 0 \end{cases}$$

显然  $Y = 0, \pm 1$ , 是一个离散型随机变量. 因此, 在连续型的情况下,  $Y$  的连续性由其自身决定.

#### 导出 $Y$ 的 PDF

已知  $X$  的 CDF  $F_X(x)$  或 PDF  $f_X(x)$ , 则  $Y = g(x)$  的 CDF 可以通过以下公式求得:

$$F_Y(y) = \Pr\{Y \leq y\} = \Pr\{g(X) \leq y\} = \Pr\{X \in C_y\}, \text{ 其中 } C_y = \{x | g(x) \leq y\}. \quad (2.15)$$

之后通过积分  $\int_{C_y} f_X(x) dx$  求得 CDF.

下面两道例题来自第五次作业.

**Exercise 2.5.1** 离散变量函数导出 PMF

设随机变量  $X$  的分布律为

$X$	-2	-1	0	1	3
$p_i$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{11}{30}$

试求  $Y = X^2$  的分布律

**Solution:** 由题意得  $Y = 0, 1, 4, 9,$ , 因此利用(2.14)计算得

$$\Pr\{Y = 0\} = \Pr\{X = 0\} = \frac{1}{5}$$

$$\Pr\{Y = 1\} = \Pr\{X = \pm 1\} = \frac{7}{30}$$

$$\Pr\{Y = 4\} = \Pr\{X = -2\} = \frac{1}{5}$$

$$\Pr\{Y = 9\} = \Pr\{X = 3\} = \frac{11}{30}$$

$$\therefore \begin{array}{c|cccc} Y & 0 & 1 & 4 & 9 \\ \hline p_i & \frac{1}{5} & \frac{7}{30} & \frac{1}{5} & \frac{11}{30} \end{array}$$

**Exercise 2.5.2** 连续变量函数导出 PDF

设  $X$  服从标准正态分布, 求  $Y = 2X^2 + 1$  的 PDF

**Solution:** 要求 PDF, 不妨先求  $Y$  的 CDF  $F_Y(y)$  i.e.  $\Pr\{Y \leq y\}$ , 代入有

$$\begin{aligned} F_Y(y) &= \Pr\{2X^2 + 1 \leq y\} = \Pr\left\{\sqrt{-\frac{y-1}{2}} \leq X \leq \sqrt{\frac{y-1}{2}}\right\} \\ &= \begin{cases} \int_{-\sqrt{\frac{y-1}{2}}}^{\sqrt{\frac{y-1}{2}}} f_X(x) dx & y > 1 \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} F_X\left(\sqrt{\frac{y-1}{2}}\right) - F_X\left(-\sqrt{\frac{y-1}{2}}\right) & y > 1 \\ 0 & \text{else} \end{cases} \end{aligned}$$

$$\therefore f_Y(y) = F'_Y(y) = \begin{cases} \sqrt{2} \frac{f_X\left(\sqrt{\frac{y-1}{2}}\right)}{4\sqrt{y-1}} + \sqrt{2} \frac{f_X\left(-\sqrt{\frac{y-1}{2}}\right)}{4\sqrt{y-1}} & y > 1 \\ 0 & \text{else} \end{cases}$$

$$\therefore \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (2.14) \quad \text{注意到 } \varphi(x) \text{ 为偶函数}$$

Therefore we have

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{\pi(y-1)}} e^{-\frac{y-1}{4}} & y > 1 \\ 0 & \text{else} \end{cases}$$

# Chapter 3

## 多维随机变量及其分布

### 3.1 二维随机变量及其分布

#### 3.1.1 二维随机变量

##### Definition 3.1.1: 二维随机变量

设随机试验的样本空间为  $S$ , 定义在  $S$  上的实值单值函数  $X = X(\omega)$   $Y = Y(\omega)$  为两个随机变量. 称向量  $(X, Y)$  为定义在  $S$  上的**二维随机变量**

$$\begin{aligned} X &: S \rightarrow \mathbb{R} \\ \text{i.e. } Y &: S \rightarrow \mathbb{R} \\ (X, Y) &: S \rightarrow \mathbb{R}^2 \end{aligned}$$

类似地, 我们可以得到  $n$  维随机变量的定义.

#### 3.1.2 二维随机变量的分布函数

##### Definition 3.1.2: 二维随机变量的分布函数

设  $(X, Y)$  是二维随机变量, 对于任意实数  $x, y$ , 定义二元函数

$$F(x, y) = \Pr\{X \leq x \cap Y \leq y\} \stackrel{\text{def}}{=} \Pr\{X \leq x, Y \leq y\} \quad (3.1)$$

称为二维随机变量  $(X, Y)$  的**分布函数**, 或**联合分布函数 (Joint CDF)**.

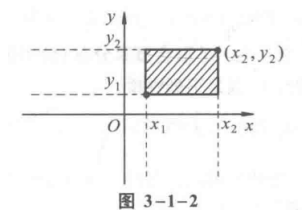


Figure 3.1: 分布函数示意图

将这个向量视作平面内随机点的座标, 分布函数就是随机点  $(X, Y)$  落入矩形域的概率. 通过容斥原理, 我们可以计算得到

$$\Pr\{x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2\} = F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1) \quad (3.2)$$

## 联合 CDF 的性质

右侧几个等式可通过矩形域的面积来理解.

1.  $0 \leq F(x, y) \leq 1$

(a) 固定  $y, F(-\infty, y) = 0$ ;

(b) 固定  $x, F(x, -\infty) = 0$ ;

(c)  $\lim_{y \rightarrow +\infty} F(x, y) = 1, \lim_{x \rightarrow -\infty} F(x, y) = 0$

2. (单调性)  $F(x, y)$  关于  $x$  和  $y$  都是单调不减的.

(a) 固定  $y, F(x, y)$  关于  $x$  单调不减;

(b) 固定  $x, F(x, y)$  关于  $y$  单调不减.

3. (连续性)  $F(x, y)$  关于  $x$  和  $y$  都是右连续的.

若已知  $(X, Y)$  的分布函数  $F(X, Y)$ , 则可由之导出各个参数 (在固定另一个参数的情况下) 各自的分布函数:

$$F_X(x) = \Pr\{X \leq x\} = \Pr\{X \leq x\} = \lim_{y \rightarrow +\infty} F(x, y), \quad (3.3)$$

$$F_Y(y) = \Pr\{Y \leq y\} = \Pr\{Y \leq y\} = \lim_{x \rightarrow +\infty} F(x, y), \quad (3.4)$$

如此导出的 CDF 称为**边缘分布函数**.