



LATEX *Probability Theory*

Sun for morning, moon for night, and you forever.

作者：Lyshmily.Y & 木易

组织：Lyshmily.Y

时间：December 13, 2024

版本：V.2.0

邮箱：yjlpku.outlook.com & 845307723@qq.com



在没有结束前，总要做很多没有意义的事，这样才可以在未来某一天，用这些无
意义的事去堵住那些讨厌的缺口



目录

目录

A

1

第1部分 * 概率论

第1章 随机事件和概率	3
1.1 事件的关系和运算	4
1.2 概率定义	4
1.3 古典概率型和几何概率型	5
1.4 概率论基本公式	5
1.5 事件独立性和独立重复实验	6
第2章 一维随机变量及其分布	7
2.1 一维随机变量	7
2.2 一维离散型随机变量	8
2.3 一维连续型随机变量	9
2.4 一维随机变量函数的分布	11
第3章 多维随机变量及其分布	12
3.1 n 维随机变量及其分布函数	12
3.2 二维离散型随机变量	13
3.3 二维连续型随机变量	15
3.3.1 常见二维连续型随机变量分布	16
3.4 独立性	17
3.5 多维随机变量函数的分布	17
第4章 随机变量的数字特征	23
4.1 一维随机变量的数字特征	23
4.2 二维随机变量的数字特征	25

4.3 独立性与相关性判定、切比雪夫不等式.....	26
第 5 章 大数定理和中心极限定理	28
5.1 依概率收敛.....	28
5.2 大数定理.....	28
5.3 中心极限定理.....	29
第 6 章 数理统计	30
6.1 总体和样本.....	30
6.2 统计量及其分布.....	31
6.3 三大分布.....	32
6.4 参数的点估计.....	34
6.5 参数的区间估计.....	35
6.6 假设检验.....	36





第一部分

概率论



第1部分目录

第1章 随机事件和概率	3
1.1 事件的关系和运算	4
1.2 概率定义	4
1.3 古典概率型和几何概率型	5
1.4 概率论基本公式	5
1.5 事件独立性和独立重复实验	6
第2章 一维随机变量及其分布	7
2.1 一维随机变量	7
2.2 一维离散型随机变量	8
2.3 一维连续型随机变量	9
2.4 一维随机变量函数的分布	11
第3章 多维随机变量及其分布	12
3.1 n 维随机变量及其分布函数	12
3.2 二维离散型随机变量	13
3.3 二维连续型随机变量	15
3.4 独立性	17
3.5 多维随机变量函数的分布	17
第4章 随机变量的数字特征	23
4.1 一维随机变量的数字特征	23
4.2 二维随机变量的数字特征	25
4.3 独立性与相关性判定、切比雪夫不等式	26
第5章 大数定理和中心极限定理	28
5.1 依概率收敛	28
5.2 大数定理	28
5.3 中心极限定理	29
第6章 数理统计	30
6.1 总体和样本	30
6.2 统计量及其分布	31
6.3 三大分布	32
6.4 参数的点估计	34
6.5 参数的区间估计	35
6.6 假设检验	36





第 1 章 随机事件和概率

内容提要

- 随机试验
- 全概率公式
- 随机事件
- 贝叶斯公式
- 古典概率型和几何概率

定义 1.0.1 (随机试验)

- 试验可以在相同条件下重复进行
- 试验的所有结果是明确可知道的，并且不止一个
- 每一次试验出现哪一个结果，事先并不确定

定义 1.0.2 (随机事件)

- 每一次试验中可能出现也可能不出现的结果称为随机事件
- 在试验中一定发生的事件为必然事件，一定不发生的事件为不可能事件

定义 1.0.3 (样本空间)

- 随机试验的每一个可能的结果称为样本点，记作 ω ；样本点的全体组成的几何称为样本空间，记作 $\Omega \Rightarrow \Omega = \{\omega\}$
- 由一个样本点构成的事件为基本事件
- 随机事件 A 是由若干个基本事件组成 $\Rightarrow A \subset \Omega$



1.1 事件的关系和运算

定义 1.1.1 (事件间关系)

- 包含: 事件 A 发生, 事件 B 发生 $\Rightarrow A \subset B$
- 相等: $A \subset B, B \subset A \Rightarrow A = B$
- 相容: $AB \neq \emptyset$
- 互斥: $AB = \emptyset$
- 对立: $A \cup B = \Omega, A \cap B = \emptyset$

定义 1.1.2 (运算法则)

- 交换律: $A \cup B = B \cup A, A \cap B = B \cap A$
- 结合律: $(A \cup B) \cup C = A \cup (B \cup C), (A \cap B) \cap C = A \cap (B \cap C)$
- 分配律: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
- De Morgan's Laws: $\overline{A \cup B} = \overline{A} \cap \overline{B}, \overline{A \cap B} = \overline{A} \cup \overline{B}$

1.2 概率定义

定义 1.2.1 (概率定义)

(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, 记作 $P(A)$

(2). 统计性定义

在相同条件下做重复实验, 事件 A 出现的次数 k 和总的试验次数 n 的比 $\frac{k}{n}$ 称为事件 A 在这 n 次试验中出现的频率, 当试验次数 n 足够大时, 频率将稳定于某个常数 p , n 越大, 频率偏离 p 的可能性越小, 这个常数 p 称为事件 A 发生的概率

(3). 公理化定义

设随机事件的样本空间为 Ω , 对于每一个事件 A 都有一个确定的实数 $P(A)$, 且事件函数 $P(*)$ 满足

- 非负性: $P(A) \geq 0$
- 规范性: $P(\Omega) = 1$
- 可列可加性: 对于任意两个互不相容的事件 A_1, A_2, \dots, A_n

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$



1.3 古典概率型和几何概率型

定义 1.3.1 (古典概率型)

古典概率型样本空间

- 只有有限个基本事件
- 每个基本事件都是等可能发生

$$P(A) = \frac{k}{n} = \frac{N(A)}{N(\Omega)}$$



定义 1.3.2 (几何概率型)

几何概率型样本空间

- 无限个基本事件
- 每个基本事件都是等可能发生
- 样本空间是一个可以度量的有界区域

$$P(A) = \frac{S_A}{S_\Omega}$$



1.4 概率论基本公式

定义 1.4.1 (性质和基本公式)

- $0 \leq P(A) \leq 1, P(\emptyset) = 0, P(\Omega) = 1$
- $A \subset B \Rightarrow P(B - A) = P(B) - P(A), P(B) \geq P(A)$
- $P(\bar{A}) = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(AB)$



定理 1.4.1 (条件概率公式)

A, B 是两个任意事件, 如果 $P(A) > 0$, 称在 A 发生的条件下 B 发生的概率为条件概率, 记作 $P(B|A)$

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



定理 1.4.2 (乘法公式)

A_1, A_2, \dots, A_n 是 n 个事件, 且 $P(A_i) > 0 (i = 1, 2, \dots, n), P(A_1 A_2 \cdots A_{n-1}) > 0$

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



特别的, 当 $n = 2$ 时

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



定理 1.4.3 (全概率公式)

完备事件组 $A_i (i = 1, 2, \dots, n)$, 满足 $\bigcup_{i=1}^n A_i = 1, A_i A_j = \emptyset$, 对于任意事件 B

$$B = \bigcup_{i=1}^n A_i B, P(B) = \sum_{i=1}^n P(A_i)P(B|A_i)$$



定理 1.4.4 (贝叶斯公式)

完备事件组 $A_i (i = 1, 2, \dots, n)$, 满足 $\bigcup_{i=1}^n A_i = 1, A_i A_j = \emptyset$, 对于任意事件 B

$$P(A_j|B) = \frac{P(A_j B)}{P(B)} = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^n P(A_i)P(B|A_i)} (j = 1, 2, \dots, n)$$



1.5 事件独立性和独立重复实验

定义 1.5.1 (独立性)

(1). 事件的独立性

(i). 描述性定义

事件 A_1, A_2, \dots, A_n 中任意一个事件 A_i 发生的概率不受其他 $n - 1$ 个事件的影响, 称这 n 个事件相互独立

(ii). 数学定义

A, B 为两个事件, 如果 $P(AB) = P(A)P(B)$, 称事件 A, B 相互独立

(2). 试验的独立性

如果各个试验的结果是相互独立的, 称这些试验是相互独立的, 试验序列 $\{E_1, E_2, \dots, E_n\}$ 中任意两个试验 E_i, E_j , 在这两个试验中任意两个结果 A_{ip}, A_{iq} 满足 $P(A_{ip}A_{jq}) = P(A_{ip})P(A_{jq})$, 称试验序列相互独立





第 2 章 ➤ 一维随机变量及其分布

内容提要

- | | |
|--------------------|-------------|
| □ 分布函数 | 几何分布 |
| □ 概率密度 | □ 指数分布、正态分布 |
| □ 二项分布、泊松分布、几何分布、超 | |

2.1 一维随机变量

定义 2.1.1 (随机变量)

设随机试验 E 的样本空间 $\Omega = \{\omega\}$ 满足: $\forall \omega \in \Omega$ 都有唯一的实数 $X(\omega)$ 与之对应, 且对任意实数 x , 都有 $\{\omega | X(\omega) \leq x, \omega \in \Omega\}$ 是随机事件, 称定义在 Ω 上的单值函数 $X(\omega)$ 是随机变量

定义 2.1.2 (分布函数)

设 X 是随机变量, x 是任意实数, 称函数 $F(x) = P(X \leq x)$ 为随机变量 X 的分布函数, 或者称 X 服从 $F(x)$ 分布, 记作 $X \sim F(x)$

推论 2.1.1 (分布函数性质)

- $F(x)$ 是单调不减函数, $\forall x_1 \leq x_2, F(x_1) \leq F(x_2)$
- $F(x)$ 是右连续函数, $\lim_{x \rightarrow x_0^+} F(x) = F(x_0)$
- $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$
- $P\{X \leq a\} = F(a), P\{X < a\} = F(a^-), P\{X = a\} = F(a) - F(a^-)$



2.2 一维离散型随机变量

定义 2.2.1 (一维离散型随机变量)

随机变量 X 只能取有限个值 x_1, x_2, \dots , 称 X 为离散型随机变量

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

上面的式子称为随机变量 X 的分布列、分布律或者概率分布, 记作 $X \sim p_i$, 概率分布通常用表格或者矩阵形式表示

X	x_1	x_2	\dots
P	p_1	p_2	\dots

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$



推论 2.2.1 (离散型随机变量的性质)

- $P(X = x_i) \geq 0, \sum_{i=1}^n P(X = x_i) = 1$
- $P(X \leq x_i) = \sum_{x_j \leq x_i} P(X = x_j)$
- $P(X \in B) = \sum_{x_i \in B} P(X = x_i)$
- $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$



定义 2.2.2 (常见离散型随机变量分布)

(1). 0-1 分布

$$X \sim B(1, p)$$

$$P(X = 1) = p, P(X = 0) = 1 - p$$

(2). 二项分布

$X \sim B(n, p)$, 试验次数为 n , 成功概率为 p , 随机变量 X 为成功次数

$$P(X = k) = C_k^n p^k (1 - p)^{n-k} (k = 0, 1, 2, \dots, n), 0 < p < 1$$

(3). 泊松分布

$$X \sim P(\lambda)$$

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} (k = 0, 1, 2, \dots), 0 < p < 1$$

(4). 几何分布

$$X \sim G(p)$$

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



(5). 超几何分布

$$X \sim H(n, N, M)$$

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



2.3 一维连续型随机变量

定义 2.3.1 (连续型随机变量分布函数和密度函数)

随机变量 X 的分布函数可以表示为:

$$F(x) = \int_{-\infty}^x f(t)dt, t \in \mathbb{R}$$

其中 $f(x)$ 是非负可积函数, 且 $\int_{-\infty}^{+\infty} f(x)dx = 1$, 称 X 是连续型随机变量, $f(x)$ 是随机变量 X 的概率密度函数, 记作 $X \sim f(x)$



推论 2.3.1 (连续型随机变量分布函数性质)

- $\int_{-\infty}^{+\infty} f(x)dx = 1$
- $P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = \int_a^b f(x)dx$
- $P(X = c) = 0$

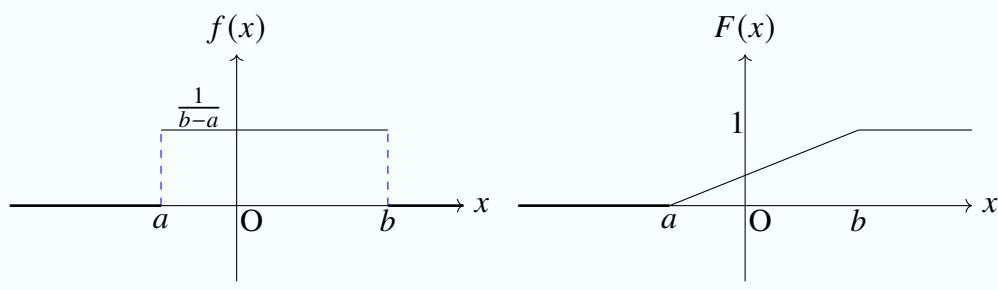


定义 2.3.2 (常见连续型随机变量分布)

(1). 均匀分布

$X \sim U(a, b)$, X 的概率密度函数 $f(x)$ 和分布函数 $F(x)$

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b) \\ 0 & x \in (-\infty, a] \cup [b, +\infty) \end{cases} \Rightarrow F(x) = \begin{cases} 0 & x \in (-\infty, a) \\ \frac{x-a}{b-a} & x \in [a, b) \\ 1 & x \in [b, +\infty) \end{cases}$$



(a) 概率密度

(b) 分布函数

图 2.1: 均匀分布概率密度和分布函数



(2). 指数分布

$X \sim E(\lambda)$, X 的概率密度函数 $f(x)$ 和分布函数 $F(x)$:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \in (0, +\infty) \\ 0 & x \in (-\infty, 0] \end{cases} \Rightarrow F(x) = \begin{cases} 1 - e^{-\lambda x} & x \in (0, +\infty) \\ 0 & x \in (-\infty, 0] \end{cases} (\lambda > 0)$$

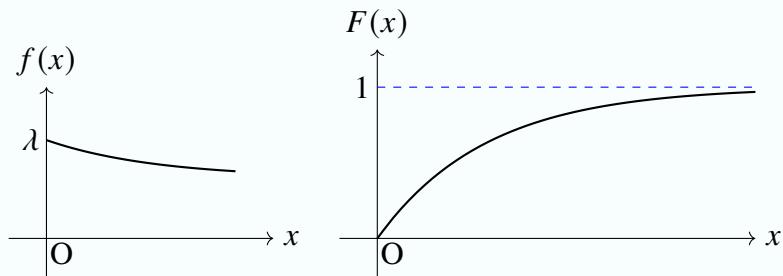


图 2.2: 指数分布概率密度和分布函数

(3). 正态分布

$X \sim N(\mu, \sigma^2)$, X 的概率密度 $f(x)$

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

特别的,当 $\mu = 0, \sigma = 1$ 时, $X \sim f(x)$ 是标准正态分布

$$\begin{cases} f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} (x \in \mathbb{R}) \\ \Phi(x) = \int_{-\infty}^x f(t) dt \end{cases}$$

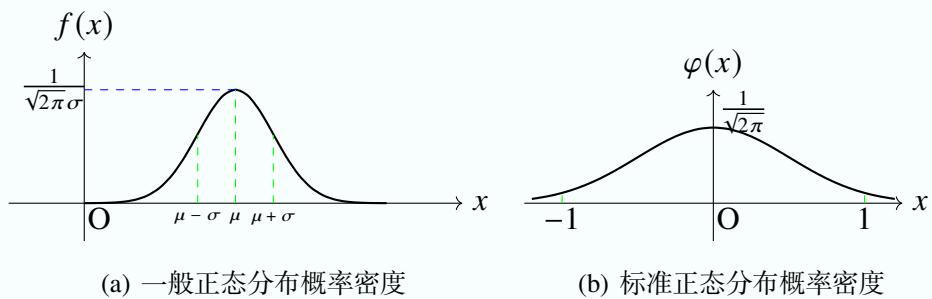


图 2.3: 一般正态分布和标准正态分布概率密度函数

推论 2.3.2 (正态分布函数性质)

- $F(x) = P(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$
- $P(a \leq X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$
- $F(\mu - x) + F(\mu + x) = 1, aX + b \sim N(a\mu + b, a^2\sigma^2)$



2.4 一维随机变量函数的分布

定义 2.4.1

设 X 是随机变量, 函数 $y = g(x)$, 以随机变量 X 作为自变量的函数 $Y = g(X)$ 也是随机变量, 称为随机变量函数

- 离散型 \rightarrow 离散型
- 连续型 \rightarrow 连续型
- 连续型 \rightarrow 离散型





第3章 多维随机变量及其分布

内容提要

- 多维随机变量
- 独立性
- 联合分布、边缘分布、条件分布
- 随机变量函数分布函数和概率密度

3.1 n 维随机变量及其分布函数

定义 3.1.1 (n 维随机变量)

如果 X_1, X_2, \dots, X_n 是定义在同一个样本空间 Ω 上的 n 个随机变量, 称 (X_1, X_2, \dots, X_n) 为 n 维随机向量

定义 3.1.2 (n 维随机变量分布函数)

对于任意的 n 个实数 x_1, x_2, \dots, x_n , n 元函数

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}$$

称为 n 维随机向量 (X_1, X_2, \dots, X_n) 的联合分布函数

特别的, 当 $n = 2$ 时, 记 (X, Y) 为 二维随机变量或者 二维随机向量, 称 $F(x, y) = P\{X \leq x, Y \leq y\}$ 为二维随机变量 (X, Y) 的联合分布函数

$$(X, Y) \sim F(x, y) \Leftrightarrow F(x, y) = P\{X \leq x, Y \leq y\}$$

推论 3.1.1 (二维随机变量联合分布性质)

(1). 单调性 (单调不减)

$$\forall x, y_1 < y_2, F(x, y_1) \leq F(x, y_2)$$



$$\forall y, x_1 < x_2, F(x_1, y) \leq F(x_2, y)$$

(2). 右连续性

$$\begin{cases} \lim_{x \rightarrow x_0^+} F(x, y) = F(x_0 + 0, y) = F(x_0, y) \\ \lim_{y \rightarrow y_0^+} F(x, y) = F(x, y_0 + 0) = F(x, y_0) \end{cases}$$

(3). 有界性

$$\begin{cases} F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0 \\ F(+\infty, +\infty) = 1 \end{cases}$$

(4). 非负性 $\forall x_1 < x_2, y_1 < y_2$

$$F(x_1 < x \leq x_2, y_1 < y \leq y_2) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$$

定义 3.1.3 (边缘分布函数)

设二维随机变量 (X, Y) 的分布函数 $F(X, Y)$, 随机变量 X 与 Y 的分布函数 $F_X(x)$ 和 $F_Y(y)$ 分别称为随机变量关于 (X, Y) 关于 X 和 Y 的边缘分布函数

$$\begin{cases} F_X(x) = P\{X \leq x\} = P\{X \leq x, Y < +\infty\} = F(x, +\infty) \\ F_Y(y) = P\{Y \leq y\} = P\{X < +\infty, Y \leq y\} = F(+\infty, y) \end{cases}$$

3.2 二维离散型随机变量

定义 3.2.1 (概率分布)

二维随机变量 (X, Y) 只能取有限对值或可列无限对值 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), \dots$, 称 (X, Y) 为二维离散型随机变量, (X, Y) 满足概率分布

$$p_{ij} = P\{X = x_i, Y = y_j\} (i, j = 1, 2, \dots)$$

上面的式子称为 (X, Y) 的联合分布律, 记作 $(X, Y) \sim p_{ij}$, 见 table : 3.1 所示

$$F(x, y) = P\{X \leq x, Y \leq y\} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij}$$



表 3.1: 离散型二维随机变量概率分布

$X \backslash Y$	y_1	\cdots	y_j	\cdots	$P\{X = x_i\}$
x_1	p_{11}	\cdots	p_{1j}	\cdots	p_{1*}
\vdots	\vdots		\vdots		\vdots
x_i	p_{i1}	\cdots	p_{ij}	\cdots	p_{i*}
\vdots	\vdots		\vdots		\vdots
$P\{Y = y_j\}$	p_{*1}	\cdots	p_{*j}	\cdots	1

定义 3.2.2 (边缘分布)

 X 边缘分布

$$p_{i*} = P\{X = x_i\} = \sum_{j=1}^{\infty} P\{X = x_i, Y = y_j\} = \sum_{j=1}^{\infty} p_{ij} (i = 1, 2, \dots)$$

 Y 边缘分布

$$p_{*j} = P\{Y = y_j\} = \sum_{i=1}^{\infty} P\{X = x_i, Y = y_j\} = \sum_{i=1}^{\infty} p_{ij} (j = 1, 2, \dots)$$

定义 3.2.3 (条件分布)

 X 在 $Y = y_j$ 下的条件分布

$$P\{X = x_i | Y = y_j\} = \frac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \frac{p_{ij}}{p_{*j}} (i = 1, 2, \dots)$$

 Y 在 $X = x_i$ 下的条件分布

$$P\{Y = y_j | X = x_i\} = \frac{P\{X = x_i, Y = y_j\}}{P\{X = x_i\}} = \frac{p_{ij}}{p_{i*}} (j = 1, 2, \dots)$$



3.3 二维连续型随机变量

定义 3.3.1 (分布函数和概率密度)

二维随机变量 (X, Y) 的分布函数 $F(X, Y)$ 可以表示为:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv, (x, y) \in \mathbb{R}^2$$

其中 $f(x, y)$ 是非负可积函数, 且 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$, 称 (X, Y) 是二维连续型随机变量, $f(x, y)$ 是随机变量 (X, Y) 的概率密度函数, 记作 $(X, Y) \sim f(x, y)$

推论 3.3.1 (分布函数和概率密度性质)

- $P\{(X, Y) \in G\} = \iint_G f(x, y) dx dy$
- $f(x, y)$ 在点 (x_0, y_0) 处连续 $\Rightarrow \frac{\partial^2 F(x, y)}{\partial x \partial y}|_{(x_0, y_0)} = f(x_0, y_0)$
- $F(x, y)$ 连续且可导, $\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$

定义 3.3.2 (边缘分布函数和边缘概率密度)

$(X, Y) \sim f(x, y)$, X, Y 的边缘分布函数和边缘概率密度

边缘分布函数

$$\begin{cases} F_X(x) = F(x, +\infty) = \int_{-\infty}^x \left[\int_{-\infty}^{+\infty} f(u, v) dv \right] du \\ F_Y(y) = F(+\infty, y) = \int_{-\infty}^y \left[\int_{-\infty}^{+\infty} f(u, v) du \right] dv \end{cases}$$

边缘概率密度

$$\begin{cases} f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy \\ f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx \end{cases}$$

定义 3.3.3 (条件分布函数和条件概率密度)

$(X, Y) \sim f(x, y)$, X 在 $Y = y$ 条件下的条件概率密度和 Y 在 $X = x$ 条件下的条件概率密度

$$\begin{cases} f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{f(x, y)}{\int_{-\infty}^{+\infty} f(x, y) dx} \\ f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{f(x, y)}{\int_{-\infty}^{+\infty} f(x, y) dy} \end{cases}$$



$$f(x, y) = \begin{cases} f_X(x)f_{Y|X}(y|x) \\ f_Y(y)f_{X|Y}(x|y) \end{cases}$$

$(X, Y) \sim f(x, y)$, X 在 $Y = y$ 条件下的条件分布函数和 Y 在 $X = x$ 条件下的条件分布函数

$$\begin{cases} F_{Y|X}(y|x) = \int_{-\infty}^y f_{Y|X}(y|x) dy = \int_{-\infty}^y \frac{f(x, y)}{f_X(x)} dy \\ F_{X|Y}(x|y) = \int_{-\infty}^x f_{X|Y}(x|y) dx = \int_{-\infty}^x \frac{f(x, y)}{f_Y(y)} dx \end{cases}$$

3.3.1 常见二维连续型随机变量分布

定义 3.3.4 (二维均匀分布)

(X, Y) 在有界区域 D 服从均匀分布, (X, Y) 的概率密度

$$f(x, y) = \begin{cases} \frac{1}{S_D} & (x, y) \in D \\ 0 & (x, y) \notin D \end{cases}$$

定义 3.3.5 (二维正态分布)

(X, Y) 概率密度

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right] \right\}$$

其中 $\mu_1, \mu_2 \in \mathbb{R}, \sigma_1, \sigma_2 > 0, -1 < \rho < 1$, 称 (X, Y) 服从参数为 $\mu_1, \mu_2, \sigma_1, \sigma_2, \rho$ 的二维正态分布, 记作 $(X, Y) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$

推论 3.3.2 (二维正态分布性质)

- 若 $(X_1, X_2) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho) \Rightarrow X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2)$
- $X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2)$, 且 X_1, X_2 相互独立, 则 $(X_1, X_2) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; 0)$
- $(X_1, X_2) \sim N \Rightarrow k_1 X_1 + k_2 X_2 \sim N$



3.4 独立性

定义 3.4.1 (二维随机变量独立性)

设二维随机变量 (X, Y) 的联合分布 $F(x, y)$, 边缘分布分别为 $F_X(x)$ 和 $F_Y(y)$, 如果任意实数对 (x, y)

$$F(x, y) = F_X(x)F_Y(y)$$

称 X 和 Y 相互独立

定义 3.4.2 (n 维随机变量独立性)

设 n 维随机变量 (X_1, X_2, \dots, X_n) 的联合分布 $F(x_1, x_2, \dots, x_n)$, 边缘分布分别为 $F_{X_1}(x_1), F_{X_2}(x_2), \dots, F_{X_n}(x_n)$, 如果任意实数对 (x_1, x_2, \dots, x_n)

$$F(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n)$$

称 X_1, X_2, \dots, X_n 相互独立

推论 3.4.1 (独立性性质)

- n 个随机变量 X_1, X_2, \dots, X_n 相互独立 $\Leftrightarrow \forall x_i \in \mathbb{R} (i = 1, 2, \dots, n)$, n 个事件 $\{X_1 \leq x_1\}, \{X_2 \leq x_2\}, \dots, \{X_n \leq x_n\}$ 相互独立
- 离散型二维随机变量 (X, Y) 相互独立 $\Leftrightarrow \forall x_i, y_j (i, j = 1, 2, \dots)$, $P\{X = x_i, Y = y_j\} = P\{X = x_i\} \cdot P\{Y = y_j\}$
- 连续型二维随机变量 (X, Y) 相互独立 $\Leftrightarrow \forall x, y, f(x, y) = f_X(x)f_Y(y)$
- (X_1, X_2, \dots, X_n) 相互独立, 其中任意 $k (2 \leq k \leq n)$ 个随机变量相互独立
- (X_1, X_2, \dots, X_n) 相互独立, $g_1(x), g_2(x), \dots, g_n(x)$ 是一元连续函数, $g_1(X_1), g_2(X_2), \dots, g_n(X_n)$ 相互独立
- 独立的二维随机变量, 边缘分布和条件分布相等, 边缘概率密度与条件概率密度相等

3.5 多维随机变量函数的分布

定义 3.5.1 (二维随机变量函数)

设 X, Y 为随机变量, $g(x, y)$ 是二元函数, 以随机变量 X, Y 作为自变量的函数 $Z = g(X, Y)$ 也是随机变量, 称为随机变量函数



定理 3.5.1 (随机变量函数分布函数和概率密度)

设 X, Y 是随机变量, $f(x, y)$ 是 (X, Y) 的概率密度, $\begin{cases} U = G(X, Y) \\ V = H(X, Y) \end{cases}$, 求 $U = g(X, Y)$ 和 $V = h(X, Y)$ 的分布函数和概率密度

$$\begin{cases} U = G(X, Y) \\ V = H(X, Y) \end{cases} \Rightarrow \begin{cases} X = P(U, V) \\ Y = Q(U, V) \end{cases} \Rightarrow |J| = \left\| \begin{array}{cc} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial Y}{\partial V} \end{array} \right\|$$

$$F(u, v) = P\{U(X, Y) \leq u, V(X, Y) \leq v\} = \iint_{D_{xy}} f(x, y) dx dy$$

$$\begin{aligned} F(u, v) &= \iint_{D_{xy}} f(x, y) dx dy \\ &= \iint_{D_{uv}} f(P(u, v), Q(u, v)) |J| du dv \end{aligned}$$

$$f(u, v) = \frac{\partial^2 F(u, v)}{\partial u \partial v} = f(P(u, v), Q(u, v)) |J|$$

$$\begin{cases} f_U(u) = \int_{-\infty}^{+\infty} f(u, v) dv = \int_{-\infty}^{+\infty} f(P(u, v), Q(u, v)) |J| dv \\ f_V(v) = \int_{-\infty}^{+\infty} f(u, v) du = \int_{-\infty}^{+\infty} f(P(u, v), Q(u, v)) |J| du \end{cases}$$

$$\begin{cases} F_U(u_0) = \int_{-\infty}^{u_0} du \int_{-\infty}^{+\infty} f(P(u, v), Q(u, v)) |J| dv \\ F_V(v_0) = \int_{-\infty}^{v_0} dv \int_{-\infty}^{+\infty} f(P(u, v), Q(u, v)) |J| du \end{cases}$$



命题 3.5.1 (和的分布)

$(X, Y) \sim f(x, y)$, $Z = X + Y$, 求 Z 的分布函数和概率密度

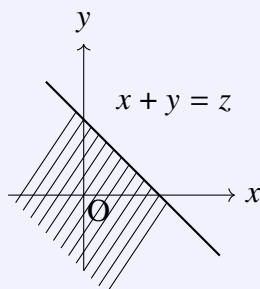
(a) $X + Y$

图 3.1: 和的分布



$$\begin{cases} Z = X + Y \\ W = Y \end{cases} \Rightarrow \begin{cases} X = Z - W \\ Y = W \end{cases} \Rightarrow |J| = 1$$

概率密度

$$f_Z(z) = \int_{-\infty}^{+\infty} f(z-w, w) dw$$

分布函数

$$F_Z(z_0) = \int_{-\infty}^{z_0} dz \int_{-\infty}^{+\infty} f(z-w, w) dw$$



命题 3.5.2 (差的分布)

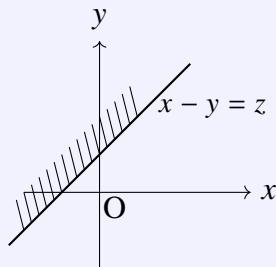
 $(X, Y) \sim f(x, y), Z = X - Y$, 求 Z 的分布函数和概率密度(a) $X - Y$

图 3.2: 差的分布

$$\begin{cases} Z = X - Y \\ W = Y \end{cases} \Rightarrow \begin{cases} X = Z + W \\ Y = W \end{cases} \Rightarrow |J| = 1$$

概率密度

$$f_Z(z) = \int_{-\infty}^{+\infty} f(z+w, w) dw$$

分布函数

$$F_Z(z_0) = \int_{-\infty}^{z_0} dz \int_{-\infty}^{+\infty} f(z+w, w) dw$$



命题 3.5.3 (积的分布)

 $(X, Y) \sim f(x, y), Z = XY$, 求 Z 的分布函数和概率密度

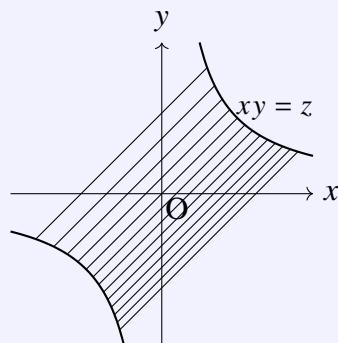
(a) XY

图 3.3: 积的分布

$$\begin{cases} Z = XY \\ W = Y \end{cases} \Rightarrow \begin{cases} X = \frac{Z}{W} \\ Y = W \end{cases} \Rightarrow |J| = \left| \frac{1}{w} \right|$$

概率密度

$$f_Z(z) = \int_{-\infty}^{+\infty} f\left(\frac{z}{w}, w\right) \left|\frac{1}{w}\right| dw$$

分布函数

$$F_Z(z_0) = \int_{-\infty}^{z_0} dz \int_{-\infty}^{+\infty} f(z-w, w) \left|\frac{1}{w}\right| dw$$

命题 3.5.4 (商的分布)

$(X, Y) \sim f(x, y)$, $Z = \frac{X}{Y}$, 求 Z 的分布函数和概率密度

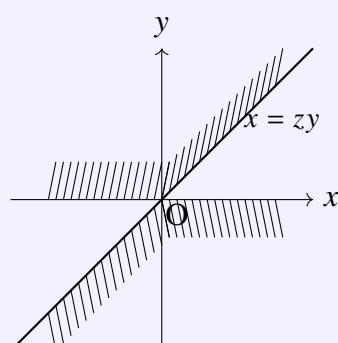
(a) $\frac{X}{Y}$

图 3.4: 商的分布

$$\begin{cases} Z = \frac{X}{Y} \\ W = Y \end{cases} \Rightarrow \begin{cases} X = ZW \\ Y = W \end{cases} \Rightarrow |J| = |w|$$

概率密度

$$f_Z(z) = \int_{-\infty}^{+\infty} f\left(\frac{z}{w}, w\right) |w| dw$$

分布函数

$$F_Z(z_0) = \int_{-\infty}^{z_0} dz \int_{-\infty}^{+\infty} f(z-w, w) |w| dw$$

命题 3.5.5 (最大值和最小值的分布)

$(X, Y) \sim f(x, y)$, $Z_1 = \max\{X, Y\}$, $Z_2 = \min\{X, Y\}$, 求 Z_1, Z_2 的分布函数和概率密度
 $\max\{X, Y\}$ 分布函数

$$F_{Z_1}(z) = P\{Z_1 \leq \max\{X, Y\}\} = P\{X \leq z, Y \leq z\} = F(z, z)$$

 $\max\{X, Y\}$ 概率密度

$$f_{Z_1}(z) = F'_{Z_1}(z)$$

 $\min\{X, Y\}$ 分布函数

$$F_{Z_2}(z) = P\{Z_2 \leq \min\{X, Y\}\} = P\{X \leq z\} \cup P\{Y \leq z\} = F_X(z) + F_Y(z) - F(z, z)$$

 $\min\{X, Y\}$ 概率密度

$$f_{Z_2}(z) = F'_{Z_2}(z)$$

推论 3.5.1 (独立同分布随机变量推论)

随机变量 X_1, X_2, \dots, X_n 相互独立, $Z_1 = \max\{X_1, X_2, \dots, X_n\}$, $Z_2 = \min\{X_1, X_2, \dots, X_n\}$,
 Z_1, Z_2 的分布函数

$$\begin{cases} F_{Z_1}(z) = F_{X_1}(z)F_{X_2}(z)\cdots F_{X_n}(z) \\ F_{Z_2}(z) = 1 - [1 - F_{X_1}(z)][1 - F_{X_2}(z)]\cdots [1 - F_{X_n}(z)] \end{cases}$$

假设随机变量 $X_i (i = 1, 2, \dots, n)$ 独立同分布

- $F_{Z_1}(x) = [F(x)]^n$, $f_{Z_1}(x) = n[F(x)]^{n-1}f(x)$
- $F_{Z_2}(x) = 1 - [1 - F(x)]^n$, $f_{Z_2}(x) = nf(x)[1 - F(x)]^{n-1}$



表 3.2: 常见独立分布可加性

X	Y	$X + Y$
$B(n, p)$	$B(m, p)$	$B(n + m, p)$
$P(\lambda_1)$	$P(\lambda_2)$	$P(\lambda_1 + \lambda_2)$
$N(\mu_1, \sigma_1^2)$	$N(\mu_2, \sigma_2^2)$	$N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$
$\chi^2(n)$	$\chi^2(m)$	$\chi^2(n + m)$





第4章 随机变量的数字特征

4.1 一维随机变量的数字特征

内容提要

- 数学期望
- 方差和标准差

- 相关系数
- 切比雪夫不等式

定义 4.1.1 (数学期望)

离散型随机变量

X 是离散型随机变量, X 的分布列为 $p_i = P\{X = x_i\} (i = 1, 2, \dots, n)$

如果级数 $\sum_{i=1}^{+\infty} x_i p_i$ 绝对收敛, 称随机变量 X 的数学期望存在, 将其记作 $E(X)$

$$E(X) = \sum_{i=1}^{+\infty} x_i p_i$$

连续型随机变量

X 是连续型随机变量, X 的概率密度为 $f(x)$

如果积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 绝对收敛, 称随机变量 X 的数学期望存在, 将其记作 $E(X)$

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

推论 4.1.1 (数学期望推论)

- $E(a) = a, E(E(X)) = E(X)$
- $E(aX \pm bY) = aE(X) \pm bE(Y)$



- $E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i)$
- X 与 Y 相互独立 $\Rightarrow E(XY) = E(X)E(Y)$
- X_1, X_2, \dots, X_n 相互独立

$$\begin{cases} E\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n E(X_i) \\ E\left[\prod_{i=1}^n g_i(X_i)\right] = \prod_{i=1}^n E[g_i(X_i)] \end{cases}$$

定义 4.1.2 (方差和标准差)

设 X 是随机变量, 如果 $E[(X - E(X))^2]$ 存在, 将 $E[(X - E(X))^2]$ 记作 X 的方差 $D(X)$

$$D(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2$$

将 $\sqrt{D(X)}$ 称为随机变量 X 的标准差或者均方差, 记作 $\sigma(X)$

随机变量 $X^* = \frac{X - E(X)}{\sqrt{D(X)}}$ 是 X 的标准化随机变量

$$\begin{cases} E(X^*) = 0 \\ D(X^*) = 1 \end{cases}$$

推论 4.1.2 (方差和标准差推论)

- $D(X) \geq 0, D(X) = E(X^2) - [E(X)]^2$
- $D(c) = 0, c \in \mathbb{R}$
- $D(aX + b) = a^2 D(X), D(X + b) = D(X)$
- $D(X \pm Y) = D(X) + D(Y) \pm 2Cov(X, Y)$
- $D\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 D(X_i) + 2 \sum_{1 \leq i < j \leq n} a_i a_j Cov(X_i, X_j)$
- X 和 Y 相互独立

$$\begin{cases} D(aX + bY) = a^2 D(X) + b^2 D(Y) \\ D(XY) = D(X)D(Y) + D(X)[E(Y)]^2 + D(Y)[E(X)]^2 \geq D(X)D(Y) \end{cases}$$
- X_1, X_2, \dots, X_n 相互独立

$$\begin{cases} D\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 D(X_i) \\ D\left(\sum_{i=1}^n g_i(X_i)\right) = \sum_{i=1}^n D[g_i(X_i)] \end{cases}$$
- $\forall c \in \mathbb{R}, D(X) \leq E[(X - c)^2]$



4.2 二维随机变量的数字特征

定义 4.2.1 (数学期望)

设 X, Y 为随机变量, $g(X, Y)$ 为 X, Y 的函数 (g 是连续函数)

离散型随机变量

$$p_{ij} = P\{X = x_i, Y = y_j\} (i, j = 1, 2, \dots)$$

级数 $\sum_i^m \sum_j^n g(x_i, y_j) p_{ij}$ 绝对收敛

$$E[g(X, Y)] = \sum_i^m \sum_j^n g(x_i, y_j) p_{ij}$$

连续型随机变量

(X, Y) 概率密度为 $f(x, y)$, 积分 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy$ 绝对收敛

$$E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy$$

定义 4.2.2 (协方差与相关系数)

随机变量 X 与 Y 的方差存在且 $D(X) > 0, D(Y) > 0$, 定义随机变量 X, Y 的协方差 $Cov(X, Y)$

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

$\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{DX}\sqrt{DY}}$ 定义为随机变量 X, Y 的相关系数

推论 4.2.1 (协方差和相关系数推论)

- $Cov(X, Y) = Cov(Y, X), Cov(X, X) = D(X), \rho_{XX} = 1$
- $Cov(aX + b, Y) = aCov(X, Y), Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$
- $\rho_{XY} = 0 \Rightarrow X, Y$ 不相关
- $\rho_{XY} \neq 0 \Rightarrow X, Y$ 相关



4.3 独立性与相关性判定、切比雪夫不等式

推论 4.3.1 (独立性与相关性判定)

随机变量 X, Y 相互独立充要条件

$$\begin{cases} f(x, y) = f_X(x) \cdot f_Y(y) \\ P\{X = x_i, Y = y_i\} = P\{X = x_i\} \cdot P\{Y = y_i\} \end{cases}$$

随机变量 X, Y 不相关充要条件 $\rho_{XY} = 0$

$$\rho_{XY} = 0 \Leftrightarrow Cov(X, Y) = 0 \Leftrightarrow E(XY) = E(X) \cdot E(Y) \Leftrightarrow D(X \pm Y) = D(X) + D(Y)$$



定义 4.3.1 (切比雪夫不等式)

随机变量 X 的期望 $E(X)$ 和方差 $D(X)$ 都存在

$$\begin{cases} \forall \varepsilon > 0, P\{|X - E(X)| \leq \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2} \\ \forall \varepsilon > 0, P\{|X - E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2} \end{cases}$$



表 4.1: 常用分布表

名称	概率分布	均值	方差	参数范围
两点分布	$P(X = k) = p^k q^{1-k}$ ($k = 0, 1$)	p	pq	$0 < p < 1$ $q = 1 - p$
二项分布 $B(n, p)$	$P(X = k) = C_n^k p^k q^{n-k}$ ($k = 0, 1, \dots, n$)	np	npq	$0 < p < 1$ $q = 1 - p$ $n \in N$
泊松分布 $P(\lambda)$	$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ ($k = 0, 1, 2, \dots$)	λ	λ	$\lambda > 0$
超几何分布 $H(n, N, M)$	$P(X = k) = \frac{C_{N-M}^{n-k} C_M^k}{C_N^n}$ ($k = 0, 1, \dots, \min\{M, n\}$)	$\frac{nM}{N}$	$\frac{n(N-n)(N-M)M}{N^2(N-1)}$	$n, N, M \in N$ $n \leq N, M \leq N$
几何分布 $G(p)$	$P(X = k) = (1-p)^{k-1} p$ ($k = 1, 2, \dots$)	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$0 < p < 1$ $q = 1 - p$
均匀分布 $U(a, b)$	$f(x) = \frac{1}{b-a}$ ($a \leq x \leq b$)	$\frac{a+b}{2}$	$\frac{(b-a)^3}{12}$	
指数分布 $E(\lambda)$	$f(x) = \lambda e^{-\lambda x}$ ($x > 0$)	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\lambda > 0$
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	$\sigma > 0$
Γ 分布 $\Gamma(\alpha, \beta)$	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ ($x > 0$)	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\alpha > 0$ $\beta > 0$



第 5 章 大数定理和中心极限定理

内容提要

- 依概率收敛
- 中心极限定理
- 大数定理

5.1 依概率收敛

定义 5.1.1 (依概率收敛)

设随机变量 X 与随机变量序列 $\{X_n\} (n = 1, 2, 3, \dots, n)$

$$\begin{cases} \forall \varepsilon > 0, \lim_{n \rightarrow \infty} P\{|X_n - X| \geq \varepsilon\} = 0 \\ \forall \varepsilon > 0, \lim_{n \rightarrow \infty} P\{|X_n - X| < \varepsilon\} = 1 \end{cases}$$

称随机变量序列 $\{X_n\}$ 依概率收敛于随机变量 X , 记作

$$\lim_{n \rightarrow \infty} X_n = X(P) \text{ or } X_n \xrightarrow{P} X(n \Rightarrow \infty)$$



5.2 大数定理

定理 5.2.1 (切比雪夫大数定理)

假设 $\{X_n\} (n = 1, 2, \dots, n)$ 是相互独立的随机变量序列, 如果方差 $D(X_i)$ 存在且一致有上



界, $\forall i \geq 1$, s.t. $D(X_i) \leq C$, $C \in \mathbb{R}$, $\{X_n\}$ 服从大数定理

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \frac{1}{n} \sum_{i=1}^n E(X_i)$$



定理 5.2.2 (伯努利大数定理)

假设 μ_n 是 n 重伯努利试验中事件 A 发生的次数, 在每次试验中事件 A 发生的概率为 $p(0 < p < 1)$, $\frac{\mu_n}{n} \xrightarrow{P} p$

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P\left\{ \left| \frac{\mu_n}{n} - p \right| < \varepsilon \right\} = 1$$



定理 5.2.3 (辛钦大数定理)

假设 $\{X_n\}$ 是独立同分布的随机变量序列, 如果 $E(X_i) = \mu(i = 1, 2, \dots, n)$ 存在,

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu$$

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P\left\{ \left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| < \varepsilon \right\} = 1$$



5.3 中心极限定理

中心极限定理

正态分布

定理 5.3.1 (列维-林德伯格定理)

假设 $\{X_n\}$ 是独立同分布的随机变量序列, 如果 $E(X_i) = \mu, D(X_i) = \sigma^2 > 0(i = 1, 2, \dots, n)$

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} P\left\{ \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \Phi(x)$$



推论 5.3.1 (中心极限定理推论)

- $\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$
- $P\{a < \sum_{i=1}^n X_i < b\} \approx \Phi\left(\frac{b - n\mu}{\sqrt{n}\sigma}\right) - \Phi\left(\frac{a - n\mu}{\sqrt{n}\sigma}\right)$



定理 5.3.2 (棣莫弗-拉普拉斯定理)

假设随机变量 $Y_n \sim B(n, p)(0 < p < 1, n \geq 1)$

$$\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} P\left\{ \frac{Y_n - np}{\sqrt{np(1-p)}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \Phi(x)$$





第 6 章 数理统计

内容提要

- | | |
|--|-------------------------------|
| <input type="checkbox"/> 统计量 | <input type="checkbox"/> 区间估计 |
| <input type="checkbox"/> χ^2 分布、 t 分布、 F 分布 | <input type="checkbox"/> 假设检验 |
| <input type="checkbox"/> 参数估计 | |

6.1 总体和样本

定义 6.1.1 (统计概念和统计量)

- **总体:** 研究对象的全体称为总体
- **样本:** n 个相互独立且与总体具有相同概率分布的随机变量 X_1, X_2, \dots, X_n 所组成的整体 (X_1, X_2, \dots, X_n) 称为来自总体 X , 容量为 n 的一个简单随机样本, 一次抽样结果的具体的 n 个数值 (x_1, x_2, \dots, x_n) 称为样本 (X_1, X_2, \dots, X_n) 的一个观测值

定义 6.1.2 (样本分布)

假设总体的分布函数 $F(x)$, 概率密度函数 $f(x)$, 样本 (X_1, X_2, \dots, X_n) 的分布函数和概率密度

- 离散型 $P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} = \prod_{i=1}^n P\{X_i = x_i\}$
- 连续型

$$\begin{cases} F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F(x_i) \\ f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i) \end{cases}$$

6.2 统计量及其分布

定义 6.2.1

设 X_1, X_2, \dots, X_n 是来自总体 X 的一个样本, $g(x_1, x_2, \dots, x_n)$ 是 n 元函数, 如果函数 g 中不含任何参数, 称 $g(X_1, X_2, \dots, X_n)$ 是样本 X_1, X_2, \dots, X_n 的一个统计量

定义 6.2.2 (样本数字特征)

- 样本均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
- 样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
- 样本标准差 $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$
- 样本 k 阶原点矩 $A_k = \frac{1}{n} \sum_{i=1}^n X_i^k (k = 1, 2, \dots)$
- 样本 k 阶中心矩 $B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k (k = 2, 3, \dots)$

定义 6.2.3 (顺序统计量)

将样本 X_1, X_2, \dots, X_n 的 n 个观测值从小到大顺序排序

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

随机变量 $X_{(k)}$ 称作第 k 顺序统计量, $X_{(1)}$ 为最小顺序统计量, $X_{(n)}$ 为最大顺序统计量

$$\begin{cases} X_{(1)} = \min\{X_1, X_2, \dots, X_n\} \\ X_{(n)} = \max\{X_1, X_2, \dots, X_n\} \end{cases}$$

推论 6.2.1 (统计量 (数字特征) 性质)

假设总体期望 $E(X) = \mu$, 总体方差为 $D(X) = \sigma^2$, X_1, X_2, \dots, X_n 是取自总体的一个样本, \bar{X}, S^2 分别为样本的均值和方差

- $E(X_i) = \mu$
- $D(X_i) = \sigma^2$
- $E(\bar{X}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = E(X) = \mu$
- $D(\bar{X}) = D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{1}{n}\sigma^2$
- $E(S^2) = D(X) = \sigma^2$



6.3 三大分布

定义 6.3.1 (χ^2 分布)

随机变量 X_1, X_2, \dots, X_n 相互独立, 且都服从标准正态分布, 随机变量 $X = \sum_{i=1}^n X_i^2$ 服从自由度为 n 的卡方分布 $\chi^2(n)$, 记作 $X \sim \chi^2(n)$

上 α 分位数: 对于给定的 $\alpha(0 < \alpha < 1)$, 称满足

$$P\{\chi^2 > \chi_{\alpha}^2(n)\} = \int_{\chi_{\alpha}^2(n)}^{+\infty} f(x)dx = \alpha$$

的 $\chi_{\alpha}^2(n)$ 为 $\chi^2(n)$ 分布的上 α 分位数

- $X_1 \sim \chi^2(n_1), X_2 \sim \chi^2(n_2), X_1, X_2$ 相互独立, $X_1 + X_2 \sim \chi^2(n_1 + n_2)$
- $X \sim \chi^2(n) \Rightarrow E(X) = n, D(X) = 2n$

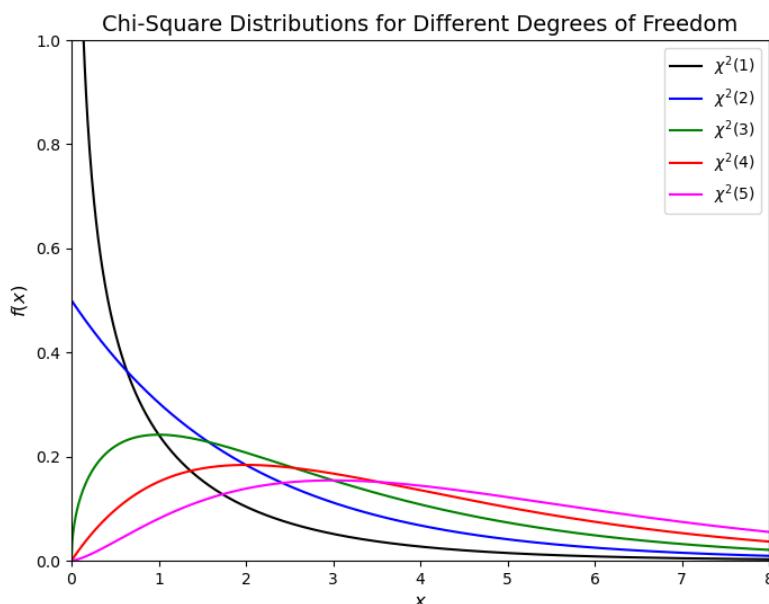


图 6.1: 卡方分布

定义 6.3.2 (t 分布)

随机变量 $X \sim N(0, 1), Y \sim \chi^2(n), X, Y$ 相互独立, 随机变量 $t = \frac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布, 记作 $t \sim t(n)$

上 α 分位数: 对于给定的 $\alpha(0 < \alpha < 1)$, 称满足

$$P\{t > t_{\alpha}(n)\} = \alpha$$



的 $t_\alpha(n)$ 为 t 分布的上 α 分位数

- t 分布概率密度关于 $x = 0$ 对称 $\Rightarrow E(t) = 0$
- $P\{t > -t_\alpha(n)\} = P\{t > t_{1-\alpha}(n)\} \Rightarrow t_{1-\alpha}(n) = -t_\alpha(n)$

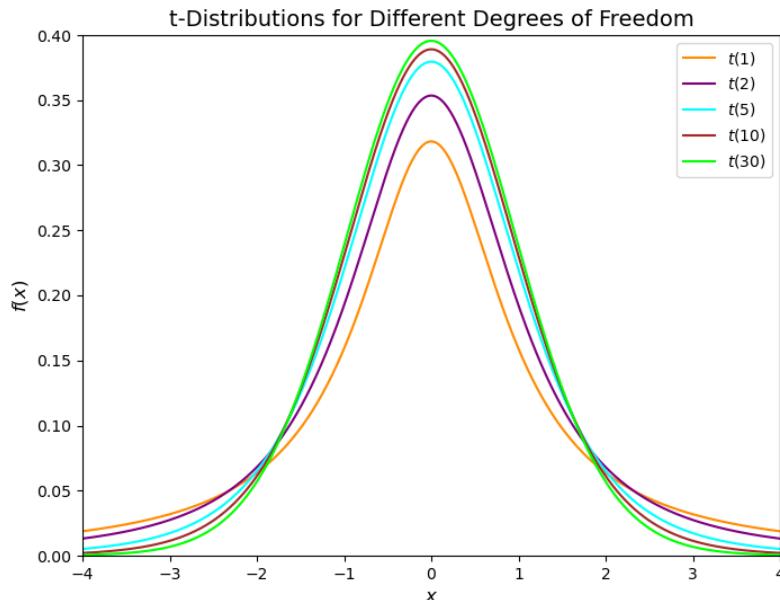


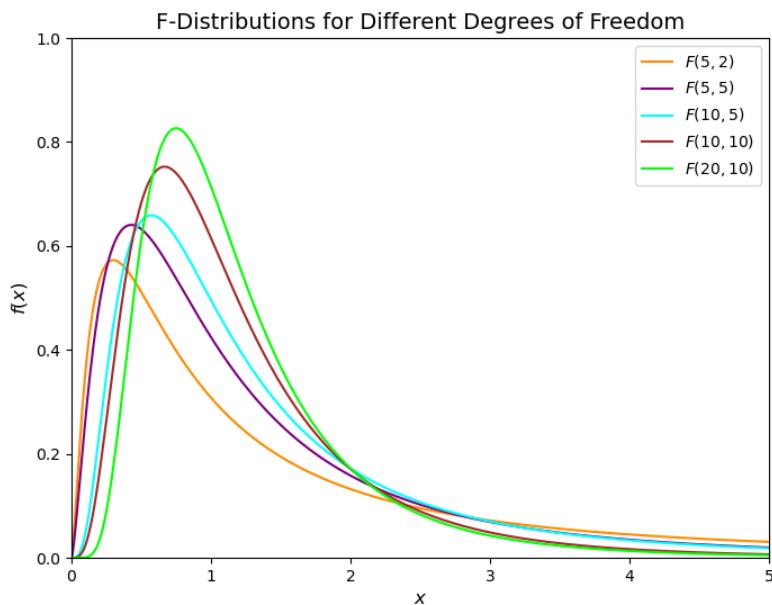
图 6.2: t 分布

定义 6.3.3 (F 分布)

随机变量 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 且 X, Y 相互独立, 随机变量 $F = \frac{X/n_1}{Y/n_2}$ 服从自由度为 (n_1, n_2) 的 F 分布, 记作 $F \sim F(n_1, n_2)$

- $F \sim F(n_1, n_2) \Rightarrow \frac{1}{F} \sim F(n_2, n_1)$
- $F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)}$
- $t \sim t(n) \Rightarrow t^2 \sim F(1, n)$



图 6.3: F 分布**推论 6.3.1 (正态总体推论)**

设 X_1, X_2, \dots, X_n 是来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, \bar{X}, S^2 分别是样本的均值和方差

- $\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$
- $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2(n)$
- $\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi^2(n-1)$
- \bar{X} 和 S^2 相互独立 $\Rightarrow \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1)$
- σ 未知时, $\frac{n(\bar{X} - \mu)^2}{S^2} \sim F(1, n-1)$

**6.4 参数的点估计****定义 6.4.1 (参数点估计)**

设总体 X 的分布函数 $F(x; \theta)$, 其中 θ 是一个未知参数, X_1, X_2, \dots, X_n 是来自总体的一个样本, 由样本构造一个适当的统计量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 作为参数 θ 的估计, 称统计量 $\hat{\theta}$ 为参数 θ 的估计量



定义 6.4.2 (矩估计)

设总体分布中有 k 个未知的参数 $\theta_1, \theta_2, \dots, \theta_k$, 来自总体 X 的一组样本 X_1, X_2, \dots, X_n , 如果 X 的原点矩 $E(X^l) (l = 1, 2, \dots, k)$ 存在, 样本原点矩 $\frac{1}{n} \sum_{i=1}^n X_i^l$ 可作为 $E(X^l)$ 的估计

定义 6.4.3 (最大似然估计)

对未知参数 θ 进行估计, 在该参数可能的取值范围 I 中选取, 使用使样本观测值 x_1, x_2, \dots, x_n 最大的参数 $\hat{\theta}$ 作为参数 θ 的估计值

似然函数

$$\begin{cases} L(\theta) = L(x_1, x_2, \dots, x_n; \theta) \\ L(\theta) = \prod_{i=1}^n f(x_i; \theta) \end{cases}$$

$$\exists \hat{\theta} \in I, \text{ s.t. } L(x_1, x_2, \dots, x_n, \hat{\theta}) = \max_{\theta \in I} L(x_1, x_2, \dots, x_n, \theta)$$

推论 6.4.1 (估计量评价标准)

- 无偏性: $E(\hat{\theta}) = \theta$
- 有效性: $D(\hat{\theta})$ 最小
- 一致性: $\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P\{|\hat{\theta} - \theta| < \varepsilon\} = 1$

6.5 参数的区间估计**定义 6.5.1 (区间估计和置信区间)**

设 θ 是总体 X 的一个未知参数, 对于给定的 $\alpha (0 < \alpha < 1)$, 如果由样本 X_1, X_2, \dots, X_n 确定的两个统计量 $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$ 和 $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$ 满足

$$P\{\hat{\theta}_1(X_1, X_2, \dots, X_n) < \theta < \hat{\theta}_2(X_1, X_2, \dots, X_n)\} = 1 - \alpha$$

则称随机区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 是 θ 的置信度为 $1 - \alpha$ 的置信区间, $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 分别称为置信上限和置信下限, $1 - \alpha$ 为置信水平, α 为显著性水平



表 6.1: 正态总体的置信区间

待估参数	其他参数	枢轴量分布	置信区间
μ	σ^2 已知	$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$	$\left(\bar{X} - \frac{\sigma}{\sqrt{n}}z_{\alpha/2}, \bar{X} + \frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right)$
μ	σ^2 未知	$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$	$\left(\bar{X} - \frac{S}{\sqrt{n}}t_{\alpha/2}(n-1), \bar{X} + \frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$
σ^2	μ 已知	$\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$	$\left(\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{\alpha/2}(n)}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{1-\alpha/2}(n)}\right)$
σ^2	μ 未知	$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\frac{(n-1)S^2}{\chi_{\alpha/2}^2(n)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n)}\right)$

6.6 假设检验

定义 6.6.1 (统计性检验)

- H_0 : 虚无假设
- H_1 : 备择假设

定义 6.6.2 (两类错误)

- 第一类错误: 虚无假设 H_0 为真, 但拒绝了 H_0 , 误认为备择假设 H_1 为真, 犯第一类错误概率 $\alpha = P\{R(H_0)|T(H_0)\}$
- 第二类错误: 备择假设 H_1 为真, 但接受了 H_0 , 误认为虚无假设 H_0 为真, 犯第二类错误概率 $\beta = P\{A(H_0)|T(H_1)\}$

推论 6.6.1 (正态总体下的六大检查和拒绝域)

- σ^2 已知, μ 未知, $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

$$\mu_r \in (-\infty, \mu_0 - \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}) \cup (\mu_0 + \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}, +\infty)$$

- σ^2 未知, μ 未知, $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

$$\mu_r \in (-\infty, \mu_0 - \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)) \cup (\mu_0 + \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1), +\infty)$$

- σ^2 已知, μ 未知, $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$

$$\mu_r \in (\mu_0 + \frac{\sigma}{\sqrt{n}}z_{\alpha}, +\infty)$$

- σ^2 已知, μ 未知, $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$

$$\mu_r \in (-\infty, \mu_0 - \frac{\sigma}{\sqrt{n}}z_{\alpha})$$



- σ^2 未知, μ 未知, $H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$
$$\mu_r \in (\mu_0 + \frac{S}{\sqrt{n}}t_\alpha(n-1), +\infty)$$

- σ^2 未知, μ 未知, $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$
$$\mu_r \in (-\infty, \mu_0 - \frac{S}{\sqrt{n}}t_\alpha(n-1))$$

