



LaTeX *main*

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

作者: Lyshmily.Y & 木易

组织: Lyshmily.Y

时间: November 21, 2024

版本: V.1.0

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	3
1.2 概率定义	4
1.3 古典概率型和几何概率型	4
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	7
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	10
第 3 章 多维随机变量及其分布	11
3.1 基本概念	11
3.2 二维离散型随机变量	12
3.3 二维连续型随机变量	13
3.4 独立性	15
第 4 章 随机变量的数字特征	20
4.1 一维随机变量的数字特征	20
4.2 二维随机变量的数字特征	21
第 5 章 大数定理和中心极限定理	24
5.1 依概率收敛	24

5.2 大数定理.....	24
5.3 中心极限定理.....	25
第 6 章 数理统计	
6.1 总体和样本.....	26
6.2 统计量及其分布.....	27
6.3 参数的点估计.....	28
6.4 参数的区间估计.....	29
6.5 假设检验.....	30



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	3
1.2 概率定义	4
1.3 古典概率型和几何概率型	4
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	7
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	10
第3章 多维随机变量及其分布	11
3.1 基本概念	11
3.2 二维离散型随机变量	12
3.3 二维连续型随机变量	13
3.4 独立性	15
第4章 随机变量的数字特征	20
4.1 一维随机变量的数字特征	20
4.2 二维随机变量的数字特征	21
第5章 大数定理和中心极限定理	24
5.1 依概率收敛	24
5.2 大数定理	24
5.3 中心极限定理	25
第6章 数理统计	26
6.1 总体和样本	26
6.2 统计量及其分布	27
6.3 参数的点估计	28
6.4 参数的区间估计	29
6.5 假设检验	30





第 1 章 随机事件和概率

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,我们把这个常数 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 ,如果对于每一个事件 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 , 我们有:

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

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满足:

- 只有有限个基本事件
- 每个基本事件都是等可能发生

$$P(A) = \frac{k}{n} = \frac{\text{事件 } A \text{ 所含的基本事件的个数}}{\text{样本空间内的基本事件个数}}$$



定义 1.3.2 (几何概率型)

几何概率型样本空间满足:

- 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

$$P(A) = \frac{S_A \text{的几何度量}}{\text{样本空间内的几何度量}}$$



1.4 概率论基本公式

定义 1.4.1 (性质和基本公式)

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



定义 1.4.2 (公式)

定理 1.4.1 (条件概率公式)

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

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



定理 1.4.2 (乘法公式)

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

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

特别的, 当 $n = 2$ 时, 我们有:

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



定理 1.4.3 (全概率公式)

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

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



定理 1.4.4 (贝叶斯公式)

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

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

**1.5 事件独立性和独立重复实验****定义 1.5.1 (定义)**

1. 事件的独立性

(1). 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, 我们称这 n 个事件相互独立.

(2). 数学定义

A, B 为两个事件, 如果我们有: $P(AB) = P(A)P(B)$, 则称事件 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, 我们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: $P(A_{ip}A_{jq}) = P(A_{ip})P(A_{jq})$, 我们称试验序列相互独立.

注

相互独立 \Rightarrow 两两独立, 两两独立 $\not\Rightarrow$ 相互独立





第 2 章 ➤ 一维随机变量及其分布

❖ 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, 我们则称定义在 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 $X \sim F(x)$

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

❖ 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 , 则称 X 为离散型随机变量, 我们有:

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

我们将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	...
P	p ₁	p ₂	...

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

数列 $\{p_i\}$ 是离散型随机变量的概率分布的充要条件为: $p_i \geq 0, \sum_{i=1}^n p_i = 1$

我们假设离散型随机变量的概率分布为: $P(X = x_i) = p_i$, 我们得到离散型随机变量 X 的分布函数:

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

对实轴上任意集合 B, 我们有:

$$P(X \in B) = \sum_{x_i \in B} P(X = x_i)$$

$$P(a < X \leq b) = P(x \leq b) - P(X \leq a) = F(b) - F(a)$$

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 $\int_{-\infty}^{+\infty} f(x)dx = 1$, 我们称 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 $X \sim f(x)$

对于连续型随机变量 X , 我们有:

$$P(X = c) = 0$$

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b)$$

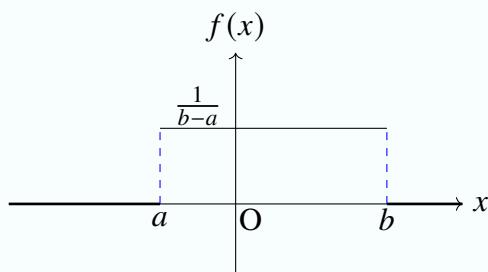
$$P(a < X < b) = \int_a^b f(x)dx$$

定义 2.3.2 (常见连续型随机变量分布)

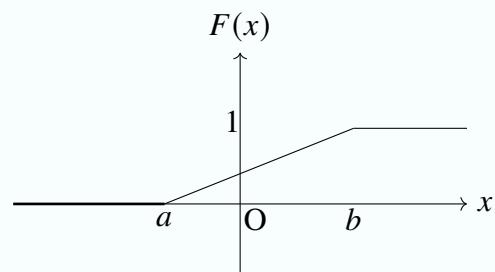
(1). 均匀分布

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

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



(a) 概率密度



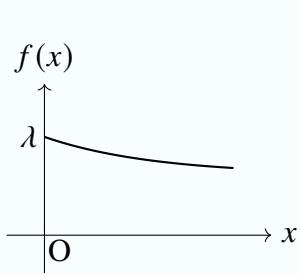
(b) 分布函数

(2). 指数分布

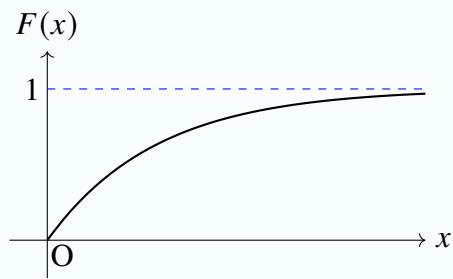
$X \sim E(\lambda)X$ 的概率密度函数 $f(x)$ 和分布函数 $F(x)$:

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





(c) 概率密度



(d) 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:

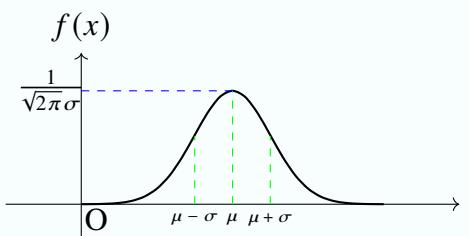
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, (-\infty < x < +\infty)$$

关于正态分布, 我们有:

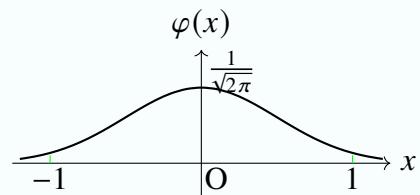
$$F(x) = P(X \leq x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$P(a \leq x \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$F(\mu - x) + F(\mu + x) = 1, aX + b \sim N(a\mu + b, a^2\sigma^2)$$



(e) 概率密度



(f) 分布函数

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.

我们一般有离散型 \rightarrow 离散型和连续型 \rightarrow 离散型两种随机变量函数.





第3章 多维随机变量及其分布

3.1 基本概念

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, 我们称 (X_1, X_2, \dots, X_n) 为 n 维随机向量.

对于任意的 n 个实数 x_1, x_2, \dots, x_n , 我们把 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, 我们称 $F(x, y) = P\{X \leq x, Y \leq y\}$ 为二维随机变量 (X, Y) 的联合分布函数, 记作:

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

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.

$$F_X(x) = P\{X \leq x\} = P\{X \leq x, Y < +\infty\} = F(x, +\infty)$$

$$F_Y(y) = P\{Y \leq y\} = P\{X < +\infty, Y \leq y\} = F(+\infty, y)$$



注

二维随机变量联合分布性质

(1). 单调性(单调不减)

$$\forall x, \text{ 当 } y_1 < y_2 \text{ 时, } F(x, y_1) \leq F(x, y_2)$$

$$\forall y, \text{ 当 } x_1 < x_2 \text{ 时, } F(x_1, y) \leq F(x_2, y)$$

(2). 右连续性

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

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

(3). 有界性

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

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

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



3.2 二维离散型随机变量

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 3.2.1 (二维离散型随机变量)

1. 概率分布

二维随机变量 (X, Y) 只能取有限对值或可列对值 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n), \dots$, 则称

(X, Y) 为离散型随机变量, (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 $(X, Y) \sim p_{ij}$, 如 **table : 3.1** 所示
数列 $\{p_{ij}\}, i, j = 1, 2, \dots$ 是某一二维随机变量的概率分布的充要条件:

$$p_{ij} \geq 0, \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} p_{ij} = 1$$

2. 联合分布函数、边缘分布、条件分布

(1). 联合分布函数

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

(2). 边缘分布

X 的边缘分布:

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

Y 的边缘分布:

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

(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 (二维连续型随机变量)

1. 联合分布函数、概率密度

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

$F(x, y)$ 是二维随机变量 (X, Y) 的概率分布函数, $f(x)$ 是二维随机变量 (X, Y) 的概率密度函数.



$$f(x, y) \geq 0, \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy$$

注

- $f(x, y)$ 在点 (x_0, y_0) 处连续, 我们有: $\frac{\partial^2 F(x, y)}{\partial x \partial y}|_{(x_0, y_0)} = f(x_0, y_0)$
- $F(x, y)$ 连续且可导, $\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$

2. 边缘分布函数、边缘概率密度

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

边缘分布函数

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

边缘概率密度

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

$(X, Y) \sim f(x, y)$, Y 的边缘分布函数和边缘密度函数:

边缘分布函数

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

边缘概率密度

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$$

3. 条件分布函数、条件概率密度

$(X, Y) \sim f(x, y)$, X 在 $Y = y$ 条件下的条件分布函数和条件概率密度:

条件概率密度

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, (f_Y(y) > 0)$$

条件分布函数

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

$(X, Y) \sim f(x, y)$, Y 在 $X = x$ 条件下的条件分布函数和条件概率密度:

条件概率密度

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}, (f_X(x) > 0)$$



条件分布函数

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

定义 3.3.2 (常见二维分布)

(1). 二维均匀分布

(X, Y) 在有界区域 D 服从均匀分布, (X, Y) 的概率密度为:

$$f(x, y) = \begin{cases} \frac{1}{S_D}, & (x, y) \in D \\ 0, & \text{其他} \end{cases} \quad S_D \text{是区域的面积}$$

(2). 二维正态分布

(X, Y) 概率密度为:

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

其中 $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, $-1 < \rho < 1$, 我们称 (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 $(X, Y) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$

注

- $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2), \rho$ 是 X 与 Y 的相关系数

$$\rho = \frac{Cov(X, Y)}{\sqrt{DX}\sqrt{DY}} = \frac{Cov(X, Y)}{\sigma_1\sigma_2}$$

- X 和 Y 的条件分布都是正态分布
- $aX + bY (a \neq 0 \text{ 或者 } b \neq 0)$ 服从正态分布
- X, Y 相互独立的充要条件为 X, Y 不相关 $\Leftrightarrow \rho = 0$

3.4 独立性

定义 3.4.1 (独立性)

1. 概念

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 $F_Y(y)$, 如果对任意实数 (x, y) , 我们都有:

$$F(x, y) = F_X(x)F_Y(y) \Leftrightarrow X \text{ 和 } Y \text{ 相互独立}$$

- 随机变量 X_1, X_2, \dots, X_n 相互独立 $\Leftrightarrow F(x_1, x_1, \dots, x_n) = F(x_1) \dots F(x_2) \dots F(x_n)$
- X_1, X_2, \dots, X_n 相互独立, 则其中任意 $k (2 \leq k \leq n)$ 个随机变量也相互独立



- 两个多维随机变量 (X_1, X_2, \dots, X_n) 与 (Y_1, Y_2, \dots, Y_m) 相互独立, 我们有:

$$F(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = F_1(x_1, x_2, \dots, x_n) \dots F_2(y_1, y_2, \dots, y_m)$$

- (X, Y) 为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. ($P\{Y = y_j\} > 0, P\{X = x_i\} > 0$)

$$P\{X = x_i | Y = y_j\} = P\{X = x_i\}, P\{Y = y_j | X = x_i\} = P\{Y = y_j\}$$

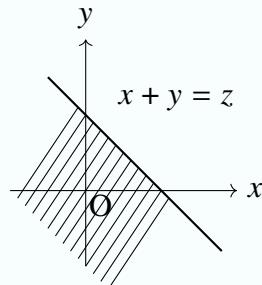
$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = f_X(x)(f_Y(y) > 0), f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = f_Y(y)(f_X(x) > 0)$$



定义 3.4.2 (连续型多维随机变量函数的分布)

$(X, Y) \sim f(x, y)$

(1). 和的分布



(a) $X + Y$

$Z = X + Y$

分布函数:

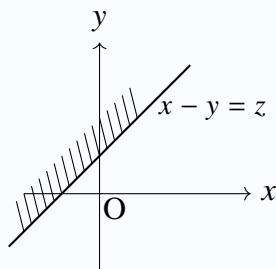
$$F_Z(z) = P\{X + Y \leq z\} = \iint_{D_z: x+y \leq z} f(x, y) d\sigma = \begin{cases} \int_{-\infty}^{+\infty} dx \int_{-\infty}^{z-x} f(x, y) dy \\ \int_{-\infty}^{+\infty} dy \int_{-\infty}^{z-y} f(x, y) dx \end{cases}$$

概率密度:

$$f_Z(z) = F'_Z(z) = \begin{cases} \int_{-\infty}^{+\infty} f(z - y, y) dy \\ \int_{-\infty}^{+\infty} f(x, z - x) dx \end{cases}$$

(2). 差的分布



(b) $X - Y$

$$Z = X - Y$$

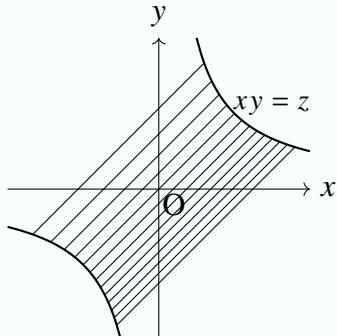
分布函数:

$$F_Z(z) = P\{X - Y \leq z\} = \iint_{D_z: x-y \leq z} f(x, y) d\sigma = \begin{cases} \int_{-\infty}^{+\infty} dx \int_{x-z}^{+\infty} f(x, y) dy \\ \int_{-\infty}^{+\infty} dy \int_{-\infty}^{y+z} f(x, y) dx \end{cases}$$

概率密度:

$$f_Z(z) = F'_Z(z) = \begin{cases} \int_{-\infty}^{+\infty} f(z+y, y) dy \\ \int_{-\infty}^{+\infty} f(x, x-z) dx \end{cases}$$

(3). 积的分布

(c) XY

$$Z = XY$$

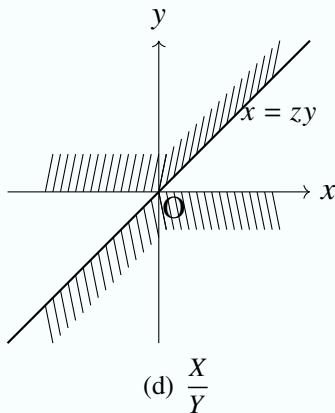
分布函数:

$$\begin{aligned} F_Z(z) &= P\{XY \leq z\} = \iint_{D_z: xy \leq z} f(x, y) d\sigma \\ &= \begin{cases} \int_{-\infty}^0 dy \int_{\frac{z}{y}}^{+\infty} f(x, y) dx + \int_0^{+\infty} dy \int_{-\infty}^{\frac{z}{y}} f(x, y) dx \\ \int_{-\infty}^0 dx \int_{\frac{z}{x}}^{+\infty} f(x, y) dy + \int_0^{+\infty} dx \int_{-\infty}^{\frac{z}{x}} f(x, y) dy \end{cases} \end{aligned}$$

概率密度:

$$f_Z(z) = F'_Z(z) = \begin{cases} \int_{-\infty}^0 \left(-\frac{1}{y}\right) f\left(\frac{z}{y}, y\right) dy + \int_0^{+\infty} \frac{1}{y} f\left(\frac{z}{y}, y\right) dy = \int_{-\infty}^{+\infty} \frac{1}{|y|} f\left(\frac{z}{y}, y\right) dy \\ \int_{-\infty}^0 \left(-\frac{1}{x}\right) f(x, \frac{z}{x}) dx + \int_0^{+\infty} \frac{1}{x} f(x, \frac{z}{x}) dx = \int_{-\infty}^{+\infty} \frac{1}{|x|} f(x, \frac{z}{x}) dx \end{cases}$$

(4). 商的分布



$$Z = \frac{X}{Y}$$

分布函数:

$$F_Z(z) = P\left\{\frac{X}{Y} \leq z\right\} = \iint_{D_z: \frac{X}{Y} \leq z} f(x, y) d\sigma = \int_{-\infty}^0 dy \int_{zy}^{+\infty} f(x, y) dx + \int_0^{+\infty} dy \int_{-\infty}^{zy} f(x, y) dx$$

概率密度:

$$f_Z(z) = F'_Z(z) = \int_{-\infty}^0 (-y) f(zy, y) dy + \int_0^{+\infty} y f(zy, y) dy = \int_{-\infty}^{+\infty} |y| f(zy, y) dy$$

(5). $\max\{X, Y\}$

$$Z = \max\{X, Y\}$$

分布函数:

$$\begin{aligned} F_Z(z) = P\{Z \leq \max\{X, Y\}\} &= P\{X \leq z\} \cup P\{Y \leq z\} \\ &= P\{X \leq z\} + P\{Y \leq z\} - P\{X \leq z, Y \leq z\} \\ &= F_X(z) + F_Y(z) - F(z, z) \end{aligned}$$

概率密度:

$$f_Z = F'_Z(z) = f(z, z)$$

(6). $\min\{X, Y\}$

$$Z = \min\{X, Y\}$$

分布函数:

$$F_Z(z) = P\{Z \leq \min\{X, Y\}\} = P\{X \leq z, Y \leq z\} = F(z, z) \Rightarrow F_{\max}(z) = F(z, z)$$



概率密度:

$$f_Z = F'_Z(z) = f_X(z) + f_Y(z) - f(z, z)$$

注

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

$$F_{\max}(z) = F_{X_1}(z)F_{X_2}(z) \cdots F_{X_n}(z)$$

$$F_{\min}(z) = 1 - [1 - F_{Z_1}(z)][1 - F_{Z_2}(z)] \cdots [1 - F_{Z_n}(z)]$$

- $X_i (i = 1, 2, \dots, n)$ 独立同分布, 我们有:

$$F_{\max}(x) = [F(x)]^n, f_{\max}(x) = n[F(x)]^{n-1}f(x)$$

$$F_{\min}(x) = 1 - [1 - F(x)]^n, f_{\min}(x) = nf(x)[1 - F(x)]^{n-1}$$



表 3.2: 常见分布可加性

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

定义 4.1.1 (数学期望)

1. 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, 我们称随机变量 X 的数学期望存在, 并将其记作 $E(X)$.

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

2. 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, 我们称随机变量 X 的数学期望存在, 并将其记作 $E(X)$.

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

- a_i 为常数, $E\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i E(X_i)$
- $E(X \pm Y) = E(X) \pm E(Y)$, $E(aX + bY) = aE(X) + bE(Y)$
- X 与 Y 相互独立 $\Rightarrow E(XY) = E(X)E(Y)$, $E[g_1(X) \dots g_2(Y)] = E[g_1(X)] \dots E[g_2(Y)]$
- X_1, X_2, \dots, X_n 相互独立, 我们有:

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



定义 4.1.2 (方差和标准差)

设 X 是随机变量, 如果 $E[(X - EX)^2]$ 存在, 我们将 $E[(X - EX)^2]$ 记作 X 的方差 $D(X)(DX)$:

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

我们将 $\sqrt{D(X)}$ 称为随机变量的**标准差**或者**均方差**, 记作 $\sigma(X)$.

- $DX \geq 0, E(X^2) = DX + (EX)^2$
- $D(c) = 0, c$ 为常数
- $D(aX + b) = a^2D(X), D(X + b) = D(X)$
- $D(X \pm Y) = D(X) + D(Y) \pm 2Cov(X, Y)$
- $X^* = \frac{X - E(X)}{\sigma(X)}$ 是 X 的**标准化随机变量**, $E(X^*) = 0, D(X^*) = 1$
- 如果 X 和 Y 相互独立, 我们得到

$$D(aX + bY) = a^2D(X) + b^2D(Y)$$

- X_1, X_2, \dots, X_n 是相互独立, 我们有:

$$D\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 D(X_i)$$

$$D\left(\sum_{i=1}^n g_i(X_i)\right) = \sum_{i=1}^n D[g_i(X_i)]$$

定义 4.1.3 (切比雪夫不等式)

如果随机变量 X 的期望 $E(X)$ 和方差 $D(X)$ 都存在, 对任意 $\varepsilon > 0$, 我们都有:

$$P\{|X - E(X)| \geq \varepsilon\} \leq \frac{D(X)}{\varepsilon^2} \text{ 或者 } P\{|X - E(X)| < \varepsilon\} \geq 1 - \frac{D(X)}{\varepsilon^2}$$

注

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)

1. (X, Y) 是离散型随机变量, 联合分布为:

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



级数 $\sum_i \sum_j g(x_i, y_j) p_{ij}$ 绝对收敛, 我们定义:

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

2. (X, Y) 是连续型随机变量, 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, 我们定义:

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

定义 4.2.2 (协方差与相关系数)

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

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

其中 $E(XY)$:

$$E(XY) = \begin{cases} \sum_i \sum_j x_i y_j p_{ij}, & (X, Y) \text{ 是离散型随机变量} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f(x, y) dx dy, & (X, Y) \text{ 是连续型随机变量} \end{cases}$$

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

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



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 $B(n, p)$	$P(X = k) = C_n^k p^k q^{n-k}$ ($x = 0, 1, \dots, n$)	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 $G(p)$	$P(X = k) = (1 - p)^{k-1} p$ ($k = 0, 1, \dots$)	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$0 < p < 1$ $q = 1 - p$
均匀分布 $U(a, b)$	$p(x) = \frac{1}{b-a}$ ($a \leq x \leq b$)	$\frac{a+b}{2}$	$\frac{(b-a)^3}{12}$	$0 < p < 1$ $q = 1 - p$ $r \in \mathbb{R}$
指数分布 $E(\lambda)$	$p(x) = \lambda e^{-\lambda x}$ ($x > 0$)	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\lambda > 0$
正态分布 $N(\mu, \sigma^2)$	$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$\sigma > 0$
Γ 分布 $\Gamma(\alpha, \beta)$	$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ ($x > 0$)	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\alpha > 0$ $\beta > 0$

第 5 章 大数定理和中心极限定理

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 $\{X_n\} (n = 1, 2, 3, \dots)$, 如果对于任意 $\varepsilon > 0$, 我们有:

$$\lim_{n \rightarrow \infty} P\{|X_n - X| \geq \varepsilon\} = 0 \text{ 或者 } \lim_{n \rightarrow \infty} P\{|X_n - X| < \varepsilon\} = 1$$

我们称随机变量序列 $\{X_n\}$ 依概率收敛于随机变量 X , 记作:

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



5.2 大数定理

定理 5.2.1 (切比雪夫大数定理)

假设 $\{X_n\} (n = 1, 2, \dots)$ 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, 即存在常数 $C, \forall i \geq 1, s.t. D(X_i) \leq C, \{X_n\}$ 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 $p (0 < p < 1)$, 则 $\frac{\mu_n}{n} \xrightarrow{P} p$.

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



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 $E(X_i) = \mu, (i = 1, 2, \dots)$ 存在, 则 $\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu$.

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



5.3 中心极限定理

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 $E(X_i) = \mu, D(X_i) = \sigma^2 > 0, (i = 1, 2, \dots)$ 存在, $\forall x \in \mathbb{R}$, 我们有:

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



注

- 定理满足 X_n 独立同分布、方差和期望存在
- $\sum_{i=1}^n X_i \sim N(n\mu, n\sigma^2)$
- $P\{a < \sum_{i=1}^n X_i < b\} \approx \Phi\left(\frac{b - n\mu}{\sqrt{n\sigma^2}}\right) - \Phi\left(\frac{a - n\mu}{\sqrt{n\sigma^2}}\right)$

定理 5.3.2 (棣莫弗-拉普拉斯定理)

假设随机变量 $Y_n \sim B(n, p), (0 < p < 1, n \geq 1), \forall x \in \mathbb{R}$, 我们有:

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



这一部分的内容我推荐一个 B 站视频观看.

- 中心极限定理
- 正态分布





第 6 章 数理统计

6.1 总体和样本

定义 6.1.1 (统计概念和统计量)

1. **总体**: 研究对象的全体称为总体
2. **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.
3. **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:

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

$$\text{连续型分布函数: } F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F(x_i)$$

$$\text{连续型概率密度: } f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i)$$



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, 我们称 $g(X_1, X_2, \dots, X_n)$ 是样本 X_1, X_2, \dots, X_n 的一个统计量.

1. 常用统计量

(1). 数字特征

- 样本均值 (一阶原点矩) $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
- 样本方差 (二阶中心矩) $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$
- 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 $B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, (k = 2, \dots)$

(2). 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得到:

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

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.

注

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, 我们有:

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

定义 6.2.2 (三大分布)

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, 则随机变量 $X = \sum_{i=1}^n X_i^2$ 服从自由度为 n 的卡方分布 $\chi^2(n)$, 记作 $X \sim \chi^2(n)$.

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

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

的 $\chi_{\alpha}^2(n)$ 为 $\chi^2(n)$ 分布上的 α 分位点.

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

2. t 分布

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

- $t_{1-\alpha}(n) = -t_\alpha(n)$
- t 分布概率密度关于 x 轴对称

3. F 分布

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

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



注

正态总体条件下常见结论:

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.

$$1. \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$$

$$2. \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2(n)$$

$$3. \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi^2(n)$$

$$4. \bar{X} \text{ 和 } S^2 \text{ 相互独立}, \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1)$$

$$5. \sigma \text{ 未知时: } \frac{n(\bar{X} - \mu)^2}{S^2} \sim F(1, n-1)$$

6.3 参数的点估计

定义 6.3.1 (矩估计和最大似然估计)

1. 概念

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, 称统计量为参数的估计量, $\theta = \hat{\theta}(X_1, X_2, \dots, X_n)$.

2. 矩估计法



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, 我们令样本的原点矩 = 总体原点矩:

$$\frac{1}{n} \sum_{i=1}^n X_i^l = E(X^l) = \begin{cases} \int_{-\infty}^{+\infty} x^l f(x; \theta_1, \theta_2, \dots, \theta_k) dx \\ \sum_{i=1}^n x_i^l P\{X = x_i; \theta_1, \theta_2, \dots, \theta_k\} \end{cases}$$

3. 最大似然估计法对未知参数 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.

(1). 总体 X 是离散型分布, 分布函数的参数为 $\theta_1, \theta_2, \dots, \theta_k$, 样本 X_1, X_2, \dots, X_n 出现取值 x_1, x_2, \dots, x_n 的概率为:

$$P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n p(x_i; \theta)$$

我们得到似然函数:

$$L(\theta) = L(x_1, x_2, \dots, x_n, \theta)$$

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

(2). 总体 X 是连续型分布, 概率密度为 $f(x; \theta), \theta \in I$, 样本 X_1, X_2, \dots, X_n 出现取值 x_1, x_2, \dots, x_n 的概率为:

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

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

我们得到了参数的最大似然估计 $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$

估计量的评判标准

1. 无偏性
2. 有效性 (最小方差性)
3. 一致性 (相和性)



6.4 参数的区间估计

定义 6.4.1 (概念)

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: 正态总体均值的置信区间

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	$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$	$\left(\bar{X} - \frac{S}{\sqrt{n}}t_{\alpha/2}(n-1), \bar{X} + \frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$

6.5 假设检验

定义 6.5.1 (统计性检验)

1. H_0 : 虚无假设, H_1 : 备择假设

2. 双边检验和单侧检验

3. 正态总体下的六大检查和拒绝域

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

$$\text{拒绝域: } (-\infty, \mu_0 - \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}) \cup (\mu_0 + \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}, +\infty)$$

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

$$\text{拒绝域: } (-\infty, \mu_0 - \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)) \cup (\mu_0 + \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1), +\infty)$$

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

$$\text{拒绝域: } (\mu_0 + \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}, +\infty)$$

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

$$\text{拒绝域: } (-\infty, \mu_0 - \frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}})$$

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

$$\text{拒绝域: } (\mu_0 + \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1), +\infty)$$

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

$$\text{拒绝域: } (-\infty, \mu_0 - \frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1))$$

