

第一章 随机事件及其概率

一、随机事件及其运算

1. 样本空间、随机事件

①样本点：随机试验的每一个可能结果，用 ω 表示；

②样本空间：样本点的全集，用 Ω 表示；

注：样本空间不唯一。

③随机事件：样本点的某个集合或样本空间的某个子集，用 A, B, C, \dots 表示；

④必然事件就等于样本空间；不可能事件 (\emptyset) 是不包含任何样本点的空集；

⑤基本事件就是仅包含单个样本点的子集。

2. 事件的四种关系

①包含关系： $A \subset B$ ，事件 A 发生必有事件 B 发生；

②等价关系： $A = B$ ，事件 A 发生必有事件 B 发生，且事件 B 发生必有事件 A 发生；

③互不相容（互斥）： $AB = \emptyset$ ，事件 A 与事件 B 一定不会同时发生。

④对立关系（互逆）： \bar{A} ，事件 \bar{A} 发生事件 A 必不发生，反之也成立；互逆满足
$$\begin{cases} \bar{\bar{A}} \cup A = \Omega \\ \bar{\bar{A}} A = \emptyset \end{cases}$$

注：互不相容和对立的关系（对立事件一定是互不相容事件，但互不相容事件不一定是对立事件。）

3. 事件的三大运算

①事件的并： $A \cup B$ ，事件 A 与事件 B 至少有一个发生。若 $AB = \emptyset$ ，则 $A \cup B = A + B$ ；

②事件的交： $A \cap B$ 或 AB ，事件 A 与事件 B 都发生；

③事件的差： $A - B$ ，事件 A 发生且事件 B 不发生。

4. 事件的运算规律

①交换律： $A \cup B = B \cup A, AB = BA$

②结合律： $(A \cup B) \cup C = A \cup (B \cup C), (A \cap B) \cap C = A \cap (B \cap C)$

③分配律： $A \cup (B \cap C) = (A \cup B) \cap (A \cup C), A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

④德摩根（De Morgan）定律：
$$\begin{aligned} \overline{A \cup B} &= \bar{A} \bar{B}, \\ \overline{AB} &= \bar{A} \cup \bar{B} \end{aligned}$$
 对于 n 个事件，有
$$\begin{aligned} \overline{\bigcup_{i=1}^n A_i} &= \bigcap_{i=1}^n \bar{A}_i, \\ \overline{\bigcap_{i=1}^n A_i} &= \bigcup_{i=1}^n \bar{A}_i \end{aligned}$$

二、随机事件的概率定义和性质

1. 公理化定义：设试验的样本空间为 Ω ，对于任一随机事件 $A (A \subset \Omega)$ ，

都有确定的实值 $P(A)$ ，满足下列性质：

(1) 非负性： $P(A) \geq 0$ ； (2) 规范性： $P(\Omega) = 1$ ；

(3) 有限可加性（概率加法公式）：对于 k 个互不相容事件 A_1, A_2, \dots, A_k ，有 $P(\sum_{i=1}^k A_i) = \sum_{i=1}^k P(A_i)$ 。

则称 $P(A)$ 为随机事件 A 的概率。

2. 概率的性质

① $P(\Omega) = 1, P(\emptyset) = 0$ ② $P(\bar{A}) = 1 - P(A)$

③若 $A \subset B$ ，则 $P(A) \leq P(B)$ ，且 $P(B - A) = P(B) - P(A)$

$$\textcircled{4} P(A \cup B) = P(A) + P(B) - P(AB)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC)$$

注：性质的逆命题不一定成立的。如若 $P(A) \leq P(B)$, 则 $A \subset B$ 。(×) 若 $P(A) = 0$, 则 $A = \phi$ 。(×)

三、古典概型的概率计算

古典概型：若随机试验满足两个条件：① 只有有限个样本点，

② 每个样本点发生的概率相同，则称该概率模型为古典概型， $P(A) = \frac{k}{n}$ 。

典型例题：设一批产品共 N 件，其中有 M 件次品，从这批产品中随机抽取 n 件样品，则

(1) 在放回抽样的方式下，取出的 n 件样品中恰好有 m 件次品（不妨设事件 A_1 ）的概率为

$$P(A_1) = \frac{C_n^m M^m (N-M)^{n-m}}{N^n}.$$

(2) 在不放回抽样的方式下，取出的 n 件样品中恰好有 m 件次品（不妨设事件 A_2 ）的概率为

$$P(A_2) = \frac{C_n^m A_M^m A_{N-M}^{n-m}}{A_N^n} = \frac{C_M^m \cdot C_{N-M}^{n-m}}{C_N^n}.$$

四、条件概率及其三大公式

1. **条件概率**： $P(B|A) = \frac{P(AB)}{P(A)}$, $P(A|B) = \frac{P(AB)}{P(B)}$

2. **乘法公式**：

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

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

3. **全概率公式**：若 B_1, B_2, \cdots, B_n 满足 $\bigcup_{i=1}^n B_i = \Omega$, $B_i B_j = \emptyset, i \neq j$, 则 $P(A) = \sum_{i=1}^n P(B_i)P(A|B_i)$ 。

4. **贝叶斯公式**：若事件 B_1, B_2, \cdots, B_n 和 A 如全概率公式所述，且 $P(A) > 0$, 则 $P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{i=1}^n P(B_i)P(A|B_i)}$ 。

五、事件的独立 1. 定义：若 $P(AB) = P(A)P(B)$, 则称 A, B 独立。

推广：若 A_1, A_2, \cdots, A_n 相互独立， $P(A_1 \cdots A_n) = P(A_1) \cdots P(A_n)$

2. 在 $\{A, B\}, \{A, \bar{B}\}, \{\bar{A}, B\}, \{\bar{A}, \bar{B}\}$ 四对事件中，只要有一对独立，则其余三对也独立。

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

3. 三个事件 A, B, C 两两独立： $P(BC) = P(B)P(C)$

$$P(AC) = P(A)P(C)$$

注： n 个事件的两两独立与相互独立的区别。（相互独立 \Rightarrow 两两独立，反之不成立。）

4. **伯努利概型**： $P_n(k) = C_n^k p^k q^{n-k}, k = 0, 1, 2, \cdots, n, q = 1 - p$ 。

1. 事件的对立与互不相容是等价的。(X)
2. 若 $P(A)=0$, 则 $A=\emptyset$ 。(X)
3. 若 $P(A)=0.1, P(B)=0.5$, 则 $P(AB)=0.05$ 。(X)
4. A, B, C 三个事件恰有一个发生可表示为 $\overline{A}\overline{B}C + \overline{A}B\overline{C} + A\overline{B}\overline{C}$ 。(V)
5. n 个事件若满足 $\forall i, j, P(A_i A_j) = P(A_i)P(A_j)$, 则 n 个事件相互独立。(X)
6. 当 $A \subset B$ 时, 有 $P(B-A) = P(B) - P(A)$ 。(V)

第二章 随机变量及其分布

一、随机变量的定义：设样本空间为 Ω , 变量 $X = X(\omega)$ 为定义在 Ω 上的单值实值函数, 则称 X 为随机变量, 通常用大写英文字母, 用小写英文字母表示其取值。

二、分布函数及其性质

1. 定义：设随机变量 X , 对于任意实数 $x \in R$, 函数 $F(x) = P\{X \leq x\}$ 称为随机变量 X 的概率分布函数, 简称分布函数。
注：当 $x_1 < x_2$ 时, $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$

(1) X 是离散随机变量, 并有概率函数 $p(x_i), i=1, 2, \dots$, 则有 $F(x) = \sum_{x_i \leq x} p(x_i)$.

(2) X 连续随机变量, 并有概率密度 $f(x)$, 则 $F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$.

2. 分布函数性质：

- (1) $F(x)$ 是单调非减函数, 即对于任意 $x_1 < x_2$, 有 $F(x_1) \leq F(x_2)$;
- (2) $0 \leq F(x) \leq 1$; 且 $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0$, $F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$;

(3) 离散随机变量 X , $F(x)$ 是右连续函数, 即 $F(x) = F(x+0)$; 连续随机变量 X , $F(x)$ 在 $(-\infty, +\infty)$ 上处处连续。

注：一个函数若满足上述 3 个条件, 则它必是某个随机变量的分布函数。

三、离散随机变量及其分布

1. 定义：设随机变量 X 只能取得有限个数值 x_1, x_2, \dots, x_n , 或可列无穷多个数值 $x_1, x_2, \dots, x_n, \dots$, 且 $P(X = x_i) = p_i (i=1, 2, \dots)$, 则称 X 为离散随机变量, $p_i (i=1, 2, \dots)$ 为 X 的概率分布, 或概率函数 (分布律)。

注：概率函数 p_i 的性质：(1) $p_i \geq 0, i=1, 2, \dots$; (2) $\sum_i p_i = 1$

2. 几种常见的离散随机变量的分布：

(1) 超几何分布, $X \sim H(N, M, n)$, $P\{X = k\} = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n} \quad k=0, 1, 2, \dots, n$

(2) 二项分布, $X \sim B(n, p)$, $P(X = k) = C_n^k p^k (1-p)^{n-k} \quad k=0, 1, \dots, n$

当 $n=1$ 时称 X 服从参数为 p 的两点分布 (或 0-1 分布)。

若 $X_i (i=1, 2, \dots, n)$ 服从同一两点分布且独立, 则 $X = \sum_{i=1}^n X_i$ 服从二项分布。

(3) 泊松(Poisson)分布, $X \sim P(\lambda)$, $P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!} (\lambda > 0), k = 0, 1, 2, \dots$

四、连续随机变量及其分布

1. 定义. 若随机变量 X 的取值范围是某个实数区间 I , 且存在非负函数 $f(x)$, 使得对于任意区间 $(a, b] \subset I$, 有 $P(a < X \leq b) = \int_a^b f(x)dx$, 则称 X 为连续随机变量; 函数 $f(x)$ 称为连续随机变量 X 的**概率密度函数**, 简称**概率密度**。

注 1: 连续随机变量 X 任取某一确定值的 x_0 概率等于 0, 即 $P(X = x_0) = 0$;

注 2: $P(x_1 < X < x_2) = P(x_1 \leq X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$

2. 概率密度 $f(x)$ 的性质: 性质 1: $f(x) \geq 0$; 性质 2: $\int_{-\infty}^{+\infty} f(x)dx = 1$.

注 1: 一个函数若满足上述 2 个条件, 则它必是某个随机变量的概率密度函数。

注 2: 当 $x_1 < x_2$ 时, $P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx$

且在 $f(x)$ 的连续点 x 处, 有 $F'(x) = f(x)$.

3. 几种常见的连续随机变量的分布:

(1) 均匀分布 $X \sim U(a, b)$,
$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{其它} \end{cases} \quad F(x) = \begin{cases} 0, & x < a; \\ \frac{x-a}{b-a}, & a \leq x < b; \\ 1, & x \geq b. \end{cases}$$

(2) 指数分布 $X \sim e(\lambda), \lambda > 0$
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

(3) 正态分布 $X \sim N(\mu, \sigma^2), \sigma > 0$
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \quad -\infty < x < +\infty$$

1. 概率函数与密度函数是同一个概念。(X)

2. 当 N 充分大时, 超几何分布 $H(n, M, N)$ 可近似成泊松分布。(X)

3. 设 X 是随机变量, 有 $P(a < X < b) = P(a \leq X \leq b)$ 。(X)

4. 若 X 的密度函数为 $f(x) = \cos x, x \in [0, \frac{\pi}{2}]$, 则 $P(0 < X < \pi) = \int_0^\pi \cos t dt$ 。(X)

第三章 随机变量的数字特征

一、期望(或均值)

1. 定义: $EX, EX = \begin{cases} \sum_{k=1}^{\infty} x_k p_k, & \text{离散型} \\ \int_{-\infty}^{+\infty} xf(x)dx, & \text{连续型} \end{cases}$

(1) $E(C) = C$, (C 为常数)

(2) $E(CX) = CE(X)$

2. 期望的性质:

(3) $E(X \pm Y) = E(X) \pm E(Y)$

(4) 若 X 与 Y 相互独立, 则 $E(XY) = E(X)E(Y)$, 反之结论不成立。

$$3. \text{ 随机变量函数的数学期望 } E[g(x)] = \begin{cases} \sum_{k=1}^{\infty} g(x_k) p_k, & X \text{ 离散型} \\ \int_{-\infty}^{+\infty} g(x) f(x) dx, & X \text{ 连续型} \end{cases}$$

4. 计算数学期望的方法

(1) 利用数学期望的定义; (2) 利用数学期望的性质;

常见的基本方法:

将一个比较复杂的随机变量 X 拆成有限多个比较简单的随机变量 X_i 之和, 再利用期望性质求得 X 的期望.

(3) 利用常见分布的期望;

$$1. \text{ 方差 } D(X) = E[X - E(X)]^2 = \begin{cases} \sum_i [x - E(X)]^2 p_i, & \text{离散型} \\ \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx, & \text{连续型} \end{cases}$$

注: $D(X) = E[X - E(X)]^2 \geq 0$; 它反映了随机变量 X 取值分散的程度, 如果 $D(X)$ 值越大(小), 表示 X 取值越分散(集中)。

2. 方差的性质

(1) $D(C) = 0$, (C 为常数)

(2) $D(CX) = C^2 D(X)$

(3) 若 X 与 Y 相互独立, 则 $D(X \pm Y) = D(X) + D(Y)$

(4) 对于任意实数 $C \in R$, 有 $E(X - C)^2 \geq D(X)$

当且仅当 $C = E(X)$ 时, $E(X - C)^2$ 取得最小值 $D(X)$.

(5) (切比雪夫不等式): 设 X 的数学期望 $E(X)$ 与方差 $D(X)$ 存在, 对于任意的正数 ε , 有

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

3. 计算

(1) 利用方差定义; (2) 常用计算公式 $D(X) = E(X^2) - [E(X)]^2$. (3) 方差的性质; (4) 常见分布的方差.

注: 常见分布的期望与方差

1. 若 $X \sim B(n, p)$, 则 $E(X) = np$, $D(X) = npq$;

2. 若 $X \sim P(\lambda)$, 则 $E(X) = D(X) = \lambda$;

3. 若 $X \sim U(a, b)$, 则 $E(X) = \frac{a+b}{2}$, $D(X) = \frac{(b-a)^2}{12}$;

4. 若 $X \sim e(\lambda)$, 则 $E(X) = \frac{1}{\lambda}$, $D(X) = \frac{1}{\lambda^2}$;

5. 若 $X \sim N(\mu, \sigma^2)$, 则 $E(X) = \mu$, $D(X) = \sigma^2$.

三、原点矩与中心矩

(总体) X 的 k 阶原点矩: $v_k(X) = E(X^k)$

(总体) X 的 k 阶中心矩: $u_k(X) = E[X - E(X)]^k$

1. 只要是随机变量, 都能计算期望和方差。(X)

2. 期望反映的是随机变量取值的中心位置, 方差反映的是随机变量取值的分散程度。(✓)

3. 方差越小, 随机变量取值越分散, 方差越大越集中。(X)

4. 方差的实质是随机变量函数的期望。(✓)

5. 对于任意的 X, Y , 都有 $D(X - Y) = DX + DY$ 成立。(X)

第四章 正态分布

一、正态分布的定义

1. 正态分布

(1) $X \sim N(\mu, \sigma^2)$ 概率密度为 $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty < x < +\infty$, 其分布函数为 $F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$

注: $F(\mu) = \frac{1}{2}$.

正态密度函数的几何特性:

(1) 曲线关于 $x = \mu$ 对称; (2) 当 $x = \mu$ 时, $f(x)$ 取得最大值 $\frac{1}{\sqrt{2\pi}\sigma}$;

(3) 当 $x \rightarrow \pm\infty$ 时, $f(x) \rightarrow 0$, 以 x 轴为渐近线; (4) $\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1 \Rightarrow \int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma$;

(5) 当固定 σ , 改变 μ 的大小时, $f(x)$ 的图形不变, 只是沿着 y 轴作平移变化.

(6) 当固定 μ , 改变 σ 的大小时, $f(x)$ 对称轴不变而形状在改变, σ 越小, 图形越高越瘦; σ 越大, 图形越矮越胖.

2. 标准正态分布

当 $\mu = 0, \sigma = 1$ 时, $X \sim N(0, 1)$, 其密度函数为 $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $-\infty < x < +\infty$. 且其分布函数为 $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$.

$\Phi(x)$ 的性质: (1) $\Phi(0) = \frac{1}{2}$;

(2) $\Phi(+\infty) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1 \Rightarrow \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$ (3) $\Phi(-x) = 1 - \Phi(x)$.

3. 正态分布与标准正态分布的关系

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

定理: 设 $X \sim N(\mu, \sigma^2)$, 则 $P(x_1 < X \leq x_2) = \Phi(\frac{x_2 - \mu}{\sigma}) - \Phi(\frac{x_1 - \mu}{\sigma})$.

二、正态分布的数字特征

设 $X \sim N(\mu, \sigma^2)$, 则 1. 期望 $E(X) = \mu$ $E(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu$

2. 方差 $D(X) = \sigma^2$ $D(X) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} (x - \mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2$

3. 标准差 $\sigma(X) = \sigma$

三、正态分布的性质

1. 线性性. 设 $X \sim N(\mu, \sigma^2)$, 则 $Y = a + bX \sim N(a + b\mu, b^2\sigma^2)$, ($b \neq 0$);

2. 可加性. 设 $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$, 且 X 和 Y 相互独立, 则 $Z = X + Y \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$;

3. 线性组合性. 设 $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2, \dots, n$, 且相互独立, 则 $\sum_{i=1}^n c_i X_i \sim N(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \sigma_i^2)$.

四、中心极限定理

1. 独立同分布的中心极限定理

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 服从相同的分布, 且 $E(X_i) = \mu, D(X_i) = \sigma^2, i = 1, 2, \dots, n, \dots$;

$$\text{则对于任何实数 } x, \text{ 有 } \lim_{n \rightarrow \infty} P\left(\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \leq x\right) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

定理解释: 若 X_1, X_2, \dots, X_n 满足上述条件, 当 n 充分大时, 有

$$(1) Y_n^* = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \sim AN(0,1); (2) Y_n^* = \sum_{i=1}^n X_i \sim AN(n\mu, n\sigma^2);$$

$$(3) \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim AN\left(\mu, \frac{\sigma^2}{n}\right)$$

2. 棣莫弗-拉普拉斯中心极限定理

$$\text{设 } Y_n \sim B(n, p), \text{ 则 } \lim_{n \rightarrow \infty} P\left(\frac{Y_n - np}{\sqrt{np(1-p)}} \leq x\right) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

定理解释: 若 $Y_n \sim B(n, p)$, 当 n 充分大时, 有

$$(1) \frac{Y_n - np}{\sqrt{np(1-p)}} \sim AN(0,1); (2) Y_n \sim AN(np, np(1-p))$$

1. 若 $X \sim N(0, 1), Y \sim N(2, 1)$, 则 $X - Y \sim N(-2, 2)$. (X)

2. 若 $X \sim N(\mu, \sigma^2)$, 则 $P\left(\frac{X - \mu}{\sigma} \leq 0\right) = \frac{1}{2}$. (√)

3. 设随机变量 X 与 Y 均服从正态分布: $X \sim N(\mu, 4^2), Y \sim N(\mu, 5^2)$

而 $p_1 = P(X \leq \mu - 4); p_2 = P(Y \geq \mu + 5)$, 则(B).

A. 对任何实数 μ , 都有 $p_1 < p_2$; B. 对任何实数 μ , 都有 $p_1 = p_2$

C. 只对 μ 的个别值, 才有 $p_1 = p_2$; D. 对任何实数 μ , 都有 $p_1 > p_2$.

4. 已知连续随机变量 X 的概率密度函数为 $f(x) = \frac{1}{\sqrt{\pi}} e^{-x^2 + 2x - 1}$ 则 X 的数学期望为 1; X 的方差为 1/2.

第五章 数理统计的基本知识

一、总体 个体 样本

1. 总体: 把研究对象的全体称为总体 (或母体). 它是一个随机变量, 记 X .

2. 个体: 总体中每个研究对象称为个体. 即每一个可能的观察值.

3. 样本: 从总体 X 中, 随机地抽取 n 个个体 X_1, X_2, \dots, X_n , 称为总体 X 的容量为 n 的样本。

注：(1) 样本 (X_1, X_2, \dots, X_n) 是一个 n 维的随机变量；(2) 本书中提到的样本都是指简单随机样本，其满足 2 个特性：

① 代表性： X_1, X_2, \dots, X_n 中每一个与总体 X 有相同的分布. ② 独立性： X_1, X_2, \dots, X_n 是相互独立的随机变量.

4. 样本 (X_1, X_2, \dots, X_n) 的联合分布

设总体 X 的分布函数为 $F(x)$ ，则样本 (X_1, X_2, \dots, X_n) 的联合分布函数为 $F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F(x_i)$;

(1) 设总体 X 的概率密度函数为 $f(x)$ ，则样本的联合密度函数为 $f(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i)$;

(2) 设总体 X 的概率函数为 $p(x)$, $(x=0,1,2,\dots)$ ，则样本的联合概率函数为 $p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i)$;

二、统计量

1. 定义

不含总体分布中任何未知参数的样本函数 $g(X_1, X_2, \dots, X_n)$ 称为统计量， $g(x_1, x_2, \dots, x_n)$ 是 $g(X_1, X_2, \dots, X_n)$ 的观测值.

注：(1) 统计量 $g(X_1, X_2, \dots, X_n)$ 是随机变量； (2) 统计量 $g(X_1, X_2, \dots, X_n)$ 不含总体分布中任何未知参数；

(3) 统计量的分布称为**抽样分布**.

2. 常用统计量

(1) 样本矩：①样本均值 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ ； 其观测值 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. 可用于推断：总体均值 $E(X)$.

②样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} (\sum_{i=1}^n X_i^2 - n\bar{X}^2)$;

其观测值 $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)$. 可用于推断：总体方差 $D(X)$.

③样本标准差 $S = \sqrt{S^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} = \sqrt{\frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right)}$.

其观测值 $s = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)}$.

④样本 k 阶原点矩 $V_k = \frac{1}{n} \sum_{i=1}^n X_i^k, (k=1,2,\dots)$ 其观测值 $v_k = \frac{1}{n} \sum_{i=1}^n x_i^k$

⑤样本 k 阶中心矩 $U_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, (k=1,2,\dots)$ 其观测值 $u_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$

注：比较样本矩与总体矩，如样本均值 \bar{X} 和总体均值 $E(X)$ ；样本方差 S^2 与总体方差 $D(X)$ ；

样本 k 阶原点矩 $V_k = \frac{1}{n} \sum_{i=1}^n X_i^k, (k=1,2,\dots)$ 与总体 k 阶原点矩 $E(X^k), (k=1,2,\dots)$ ；样本 k 阶中心矩

$U_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, (k=1,2,\dots)$ 与总体 k 阶原点矩 $E[X - E(X)]^k, (k=1,2,\dots)$. 前者是随机变量，后者是常数.

(2) 样本矩的性质:

设总体 X 的数学期望和方差分别为 $EX = \mu, DX = \sigma^2$, \bar{X}, S^2 为样本均值、样本方差, 则

$$1^\circ E(\bar{X}) = \mu; \quad 2^\circ D(\bar{X}) = \frac{1}{n}\sigma^2; \quad 3^\circ E(S^2) = \sigma^2.$$

3. 抽样分布: 统计量的分布称为抽样分布.

三、

3 大抽样分布

1. χ^2 分布: 定义. 设 X_1, X_2, \dots, X_k 相互独立, 且 $X_i \sim N(0,1), i=1,2,\dots,k$, 则 $\chi^2 = X_1^2 + X_2^2 + \dots + X_k^2 \sim \chi^2(k)$

注: 若 $X \sim N(0,1)$, 则 $X^2 \sim \chi^2(1)$.

(2) 性质 (可加性)

设 χ_1^2 和 χ_2^2 相互独立, 且 $\chi_1^2 \sim \chi^2(k_1), \chi_2^2 \sim \chi^2(k_2)$, 则 $\chi_1^2 + \chi_2^2 \sim \chi^2(k_1 + k_2)$.

2. t 分布: 设 X 与 Y 相互独立, 且 $X \sim N(0,1), Y \sim \chi^2(k)$, 则 $t = \frac{X}{\sqrt{Y/k}} \sim t(k)$.

注: t 分布的密度图像关于 $t=0$ 对称; 当 n 充分大时, t 分布趋向于标准正态分布 $N(0,1)$.

3. F 分布: 定义. 设 X 与 Y 相互独立, 且 $X \sim \chi^2(k_1), Y \sim \chi^2(k_2)$, 则 $F = \frac{X/k_1}{Y/k_2} \sim F(k_1, k_2)$.

(2) 性质. 设 $X \sim F(k_1, k_2)$, 则 $1/X \sim F(k_2, k_1)$.

四、分位点

定义: 对于总体 X 和给定的 $\alpha (0 < \alpha < 1)$, 若存在 x_α , 使得 $P(X \geq x_\alpha) = \alpha$ 则称 x_α 为 X 分布的 α 分位点.

注: 常见分布的分位点表示方法

(1) $\chi^2(k)$ 分布的 α 分位点 $\chi_\alpha^2(k)$; (2) $t(k)$ 分布的 α 分位点 $t_\alpha(k)$, 其性质: $t_{1-\alpha}(k) = -t_\alpha(k)$;

(3) $F_\alpha(k_1, k_2)$, 分布的 α 分位点 $F_\alpha(k_1, k_2)$, 其性质 $F_{1-\alpha}(k_1, k_2) = \frac{1}{F_\alpha(k_2, k_1)}$;

(4) $N(0,1)$ 分布的 α 分位点 u_α , 有 $P(X \geq u_\alpha) = 1 - P(X < u_\alpha) = 1 - \Phi(u_\alpha)$,

第六章 参数估计

一、点估计: 设 (X_1, X_2, \dots, X_n) 为来自总体 X 的样本, θ 为 X 中的未知参数, (x_1, x_2, \dots, x_n) 为样本值, 构造某个统计

量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 作为参数 θ 的估计, 则称 $\hat{\theta}(X_1, X_2, \dots, X_n)$ 为 θ 的点估计量, $\hat{\theta}(x_1, x_2, \dots, x_n)$ 为 θ 的估计值.

2. 常用点估计的方法: 矩估计法和最大似然估计法.

二、矩估计法

1. 基本思想: 用样本矩 (原点矩或中心矩) 代替相应的总体矩.

2. 求总体 X 的分布中包含的 m 个未知参数 $\theta_1, \theta_2, \dots, \theta_m$ 的矩估计步骤:

① 求出总体矩, 即 $E(X^k)$ 或 $E[X - E(X)]^k, k=1,2,\dots$; ② 用样本矩代替总体矩, 列出矩估计方程:

$$\frac{1}{n} \sum_{i=1}^n X_i^k = E(X^k) \text{ 或 } \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k = E[X - E(X)]^k, k=1,2,\dots$$

③ 解上述方程 (或方程组) 得到 $\theta_1, \theta_2, \dots, \theta_m$ 的矩估计量为: $\hat{\theta}_i = \hat{\theta}_i(X_1, X_2, \dots, X_n), i=1,2,\dots,m$

④ $\theta_1, \theta_2, \dots, \theta_m$ 的矩估计值为: $\hat{\theta}_i = \hat{\theta}_i(x_1, x_2, \dots, x_n), i=1,2,\dots,m$

3. 矩估计法的优缺点:

优点: 直观、简单; 只须知道总体的矩, 不须知道总体的分布形式.

缺点: 没有充分利用总体分布提供的信息; 矩估计量不具有唯一性; 可能估计结果的精度比其它估计法的低

三、最大似然估计法

1. 直观想法: 在试验中, 事件 A 的概率 $P(A)$ 最大, 则 A 出现的可能性就大; 如果事件 A 出现了, 我们认为事件 A 的概率最大.

2. 定义 设总体 X 的概率函数或密度函数为 $p(x, \theta)$ (或 $f(x, \theta)$), 其中参数 θ 未知, 则 X 的样本 (X_1, X_2, \dots, X_n) 的联合

$$\text{概率函数 (或联合密度函数)} \quad L(\theta) = \prod_{i=1}^n p(x_i, \theta) \quad (\text{或 } L(\theta) = \prod_{i=1}^n f(x_i, \theta))$$

称为似然函数.

3. 求最大似然估计的步骤:

$$(1) \text{ 求似然函数: } X \text{ 离散: } L(\theta) = \prod_{i=1}^n p(x_i, \theta) \quad X \text{ 连续: } L(\theta) = \prod_{i=1}^n f(x_i, \theta)$$

$$(2) \text{ 求 } \ln L(\theta) \text{ 和似然方程: } \frac{\partial \ln L(\theta)}{\partial \theta_i} = 0, i = 1, 2, \dots, m$$

$$(3) \text{ 解似然方程, 得到最大似然估计值: } \hat{\theta}_i = \hat{\theta}_i(x_1, x_2, \dots, x_n), i = 1, 2, \dots, m$$

$$(4) \text{ 最后得到最大似然估计量: } \hat{\theta}_i = \hat{\theta}_i(X_1, X_2, \dots, X_n), i = 1, 2, \dots, m$$

4. 最大似然估计法是在各种参数估计方法中比较优良的方法, 但是它需要知道总体 X 的分布形式.

四、估计量的评价标准

1. 无偏性: 设 $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ 是未知参数 θ 的估计量, 若 $E(\hat{\theta}) = \theta$, 则 $\hat{\theta} = \hat{\theta}(X_1, \dots, X_n)$ 是 θ 的无偏估计量,

$$\hat{\theta} = \hat{\theta}(x_1, \dots, x_n) \text{ 是 } \theta \text{ 的无偏估计值.}$$

有效性: 设 $\hat{\theta}_1 = \hat{\theta}_1(X_1, \dots, X_n)$ 和 $\hat{\theta}_2 = \hat{\theta}_2(X_1, \dots, X_n)$ 是未知参数 θ 的无偏估计量,

若 $D(\hat{\theta}_1) < D(\hat{\theta}_2) = \theta$, 则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效.

1. 若 (X_1, X_2, \dots, X_n) 是来自总体 X 的样本, 则 X_1, X_2, \dots, X_n 相互独立. (✓)

2. 不含总体 X 的任何未知参数的样本函数 $g(X_1, X_2, \dots, X_n)$ 就是统计量. (✓)

3. 样本矩与总体矩是等价的. (X)

4. 矩估计法的基本思想是用总体矩代替样本矩, 故矩估计量不唯一. (X)

设总体 $X \sim N(\mu, \sigma^2)$, 其中 μ, σ^2 未知, 则估计量 $\hat{\mu} = \bar{X}, \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ 分别是 μ, σ^2 的无偏估计

量. (X)