



温州大學
WENZHOU UNIVERSITY

机器学习-概率论回顾

黄海广 副教授

2021年07月

- 01** 随机事件和概率
- 02** 随机变量及其概率分布
- 03** 多维随机变量及其分布
- 04** 随机变量的数字特征
- 05** 数理统计的基本概念

1.随机事件和概率

3

01 随机事件和概率

02 随机变量及其概率分布

03 多维随机变量及其分布

04 随机变量的数字特征

05 数理统计的基本概念

1.随机事件和概率

4

1.事件的关系与运算

- (1) 子事件: $A \subset B$, 若 A 发生, 则 B 发生。
- (2) 相等事件: $A = B$, 即 $A \subset B$, 且 $B \subset A$ 。
- (3) 和事件: $A \cup B$ (或 $A + B$), A 与 B 中至少有一个发生。
- (4) 差事件: $A - B$, A 发生但 B 不发生。
- (5) 积事件: $A \cap B$ (或 AB), A 与 B 同时发生。
- (6) 互斥事件 (互不相容): $A \cap B = \emptyset$ 。
- (7) 互逆事件 (对立事件): $A \cap B = \emptyset, A \cup B = \Omega, A = \bar{B}, B = \bar{A}$ 。

1.随机事件和概率

5

2.运算律

(1) 交换律: $A \cup B = B \cup A, A \cap B = B \cap A$

(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: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

3.德.摩根律

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \quad \overline{A \cap B} = \bar{A} \cup \bar{B}$$

4.完全事件组

$A_1 A_2 \cdots A_n$ 两两互斥, 且和事件为必然事件, 即 $A_i \cap A_j = \emptyset, i \neq j, \bigcup_{i=1}^n A_i = \Omega$

1.随机事件和概率

6

5.概率的基本概念

(1) 概率：事件发生的可能性大小的度量，其严格定义如下：

概率 $P(g)$ 为定义在事件集合上的满足下面3个条件的函数：

1)对任何事件 A , $P(A) \geq 0$

2)对必然事件 Ω , $P(\Omega) = 1$

3)对 $A_1 A_2 \cdots A_n, \cdots$, 若 $A_i A_j = \emptyset (i \neq j)$, 则: $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

1.随机事件和概率

7

(2) 概率的基本性质

1) $P(\bar{A}) = 1 - P(A);$

2) $P(A - B) = P(A) - P(AB);$

3) $P(A \cup B) = P(A) + P(B) - P(AB)$

特别, 当 $B \subset A$ 时, $P(A - B) = P(A) - P(B)$ 且 $P(B) \leq P(A);$

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

4) 若 A_1, A_2, \dots, A_n 两两互斥, 则 $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n (P(A_i))$

1.随机事件和概率

8

(3) 古典型概率: 实验的所有结果只有有限个, 且每个结果发生的可能性相同, 其概率计算公式:

$$P(A) = \frac{\text{事件}A\text{发生的基本事件数}}{\text{基本事件总数}}$$

(4) 几何型概率: 样本空间 Ω 为欧氏空间中的一个区域, 且每个样本点的出现具有等可能性, 其概率计算公式:

$$P(A) = \frac{A\text{的度量(长度、面积、体积)}}{\Omega\text{的度量(长度、面积、体积)}}$$

1.随机事件和概率

9

6.概率的基本公式

(1) 条件概率: $P(B|A) = \frac{P(AB)}{P(A)}$, 表示A发生的条件下, B发生的概率

(2) 全概率公式: $P(A) = \sum_{i=1}^n P(A|B_i)P(B_i)$, $B_i B_j = \emptyset, i \neq j, \bigcup_{i=1}^n B_i = \Omega$.

(3) **Bayes**公式:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^n P(A|B_i)P(B_i)}, j = 1, 2, \dots, n$$

(4)乘法公式: $P(A_1 A_2) = P(A_1)P(A_2|A_1) = P(A_2)P(A_1|A_2)$ $P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \cdots P(A_n|A_1 A_2 \cdots A_{n-1})$

1.随机事件和概率

10

7.事件的独立性

(1) A与B相互独立 $\Leftrightarrow P(AB) = P(A)P(B)$

(2) A, B, C两两独立 $\Leftrightarrow P(AB) = P(A)P(B); P(BC) = P(B)P(C); P(AC) = P(A)P(C);$

(3) A, B, C相互独立 $\Leftrightarrow P(AB) = P(A)P(B); P(BC) = P(B)P(C); P(AC) = P(A)P(C);$
 $P(ABC) = P(A)P(B)P(C).$

1.随机事件和概率

11

8.独立重复试验

将某试验独立重复 n 次，若每次实验中事件 A 发生的概率为 p ，则 n 次试验中 A 发生 k 次的概率为：

$$P(X = k) = C_n^k p^k (1 - p)^{n-k}$$

9.重要公式与结论

$$(1) P(\overline{A}) = 1 - P(A)$$

$$(2) P(A \cup B) = P(A) + P(B) - P(AB)$$

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

$$(3) P(A - B) = P(A) - P(AB)$$

1.随机事件和概率

12

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

(5) 条件概率 $P(\cdot|B)$ 满足概率的所有性质,

$$\text{例如: } P(\bar{A}_1|B) = 1 - P(A_1|B) \quad P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1A_2|B) \\ P(A_1A_2|B) = P(A_1|B)P(A_2|A_1B)$$

$$(6) \text{ 若 } A_1, A_2, \dots, A_n \text{ 相互独立, 则 } P(\cap_{i=1}^n A_i) = \prod_{i=1}^n P(A_i), P(\cup_{i=1}^n A_i) = \prod_{i=1}^n (1 - P(A_i))$$

1.随机事件和概率

13

(7) 互斥、互逆与独立性之间的关系: A 与 B 互逆 $\Rightarrow A$ 与 B 互斥, 但反之不成立, A 与 B 互斥 (或互逆) 且均非零概率事件 $\Rightarrow A$ 与 B 不独立.

(8) 若 $A_1, A_2, \dots, A_m, B_1, B_2, \dots, B_n$ 相互独立, 则 $f(A_1, A_2, \dots, A_m)$ 与 $g(B_1, B_2, \dots, B_n)$ 也相互独立, 其中 $f(\cdot), g(\cdot)$ 分别表示对相应事件做任意事件运算后所得的事件, 另外, 概率为1 (或0) 的事件与任何事件相互独立.

2.随机变量及其概率分布

14

01 随机事件和概率

02 随机变量及其概率分布

03 多维随机变量及其分布

04 随机变量的数字特征

05 数理统计的基本概念

2.随机变量及其概率分布

15

1.随机变量及概率分布

取值带有随机性的变量，严格地说是定义在样本空间上，取值于实数的函数称为随机变量，概率分布通常指分布函数或分布律

2.分布函数的概念与性质

定义： $F(x) = P(X \leq x), -\infty < x < +\infty$

性质：(1) $0 \leq F(x) \leq 1$ (2) $F(x)$ 单调不减

(3) 右连续 $F(x+0) = F(x)$ (4) $F(-\infty) = 0, F(+\infty) = 1$

2. 随机变量及其概率分布

16

3. 离散型随机变量的概率分布

$$P(X = x_i) = p_i, i = 1, 2, \dots, n, \dots \quad p_i \geq 0, \sum_{i=1}^{\infty} p_i = 1$$

4. 连续型随机变量的概率密度

概率密度 $f(x)$;非负可积, 且:(1) $f(x) \geq 0$, (2) $\int_{-\infty}^{+\infty} f(x)dx = 1$ (3) x 为 $f(x)$ 的连续点, 则:

$$f(x) = F'(x) \text{ 分布函数 } F(x) = \int_{-\infty}^x f(t)dt$$

2.随机变量及其概率分布

17

5.常见分布

(1) 0-1分布: $P(X = k) = p^k(1 - p)^{1-k}, k = 0, 1$

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

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

(4) 均匀分布 $U(a, b): f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \end{cases}$

2. 随机变量及其概率分布

18

(5) 正态分布: $N(\mu, \sigma^2)$: $\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \sigma > 0, -\infty < x < +\infty$

(6) 指数分布: $E(\lambda)$: $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \lambda > 0 \\ 0, & \end{cases}$

(7) 几何分布: $G(p)$: $P(X = k) = (1 - p)^{k-1}p, 0 < p < 1, k = 1, 2, \dots$.

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

2. 随机变量及其概率分布

19

6. 随机变量函数的概率分布

(1) 离散型: $P(X = x_1) = p_i, Y = g(X)$

则: $P(Y = y_j) = \sum_{g(x_i)=y_j} P(X = x_i)$

(2) 连续型: $X \sim f_X(x), Y = g(x)$

则: $F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{g(x) \leq y} f_X(x) dx, \quad f_Y(y) = F'_Y(y)$

2.随机变量及其概率分布

20

7.重要公式与结论

$$(1) X \sim N(0,1) \Rightarrow \varphi(0) = \frac{1}{\sqrt{2\pi}}, \Phi(0) = \frac{1}{2}, \Phi(-a) = P(X \leq -a) = 1 - \Phi(a)$$

$$(2) X \sim N(\mu, \sigma^2) \Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1), P(X \leq a) = \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$(3) X \sim E(\lambda) \Rightarrow P(X > s + t | X > s) = P(X > t)$$

$$(4) X \sim G(p) \Rightarrow P(X = m + k | X > m) = P(X = k)$$

(5) 离散型随机变量的分布函数为阶梯间断函数；连续型随机变量的分布函数为连续函数，但不一定为处处可导函数。

(6) 存在既非离散也非连续型随机变量。

3.多维随机变量及其分布

21

01 随机事件和概率

02 随机变量及其概率分布

03 多维随机变量及其分布

04 随机变量的数字特征

05 数理统计的基本概念

3.多维随机变量及其分布

22

1.二维随机变量及其联合分布

由两个随机变量构成的随机向量 (X, Y) , 联合分布为 $F(x, y) = P(X \leq x, Y \leq y)$

2.二维离散型随机变量的分布

(1) 联合概率分布律 $P\{X = x_i, Y = y_j\} = p_{ij}; i, j = 1, 2, \dots$

(2) 边缘分布律 $p_{i\cdot} = \sum_{j=1}^{\infty} p_{ij}, i = 1, 2, \dots$ $p_{\cdot j} = \sum_{i=1}^{\infty} p_{ij}, j = 1, 2, \dots$

(3) 条件分布律 $P\{X = x_i | Y = y_j\} = \frac{p_{ij}}{p_{\cdot j}}$ $P\{Y = y_j | X = x_i\} = \frac{p_{ij}}{p_{i\cdot}}$

3. 多维随机变量及其分布

23

3. 二维连续性随机变量的密度

(1) 联合概率密度 $f(x, y)$:

$$1) f(x, y) \geq 0 \quad 2) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$

$$(2) \text{ 分布函数: } F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$$

$$(3) \text{ 边缘概率密度: } f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy \qquad f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$$

$$(4) \text{ 条件概率密度: } f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} \qquad f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

3.多维随机变量及其分布

24

4.常见二维随机变量的联合分布

(1) 二维均匀分布: $(x, y) \sim U(D)$, $f(x, y) = \begin{cases} \frac{1}{S(D)}, (x, y) \in D \\ 0, \text{其他} \end{cases}$

(2) 二维正态分布: $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

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

3.多维随机变量及其分布

25

5.随机变量的独立性和相关性

X 和 Y 的相互独立: $\Leftrightarrow F(x, y) = F_X(x)F_Y(y)$:

$$\Leftrightarrow p_{ij} = p_{i\cdot} \cdot p_{\cdot j} \quad (\text{离散型}) \quad \Leftrightarrow f(x, y) = f_X(x)f_Y(y) \quad (\text{连续型})$$

X 和 Y 的相关性:

相关系数 $\rho_{XY} = 0$ 时, 称 X 和 Y 不相关, 否则称 X 和 Y 相关

3.多维随机变量及其分布

26

6.两个随机变量简单函数的概率分布

离散型: $P(X = x_i, Y = y_i) = p_{ij}, Z = g(X, Y)$ 则:

$$P(Z = z_k) = P\{g(X, Y) = z_k\} = \sum_{g(x_i, y_i) = z_k} P(X = x_i, Y = y_j)$$

连续型: $(X, Y) \sim f(x, y), Z = g(X, Y)$ 则:

$$F_z(z) = P\{g(X, Y) \leq z\} = \iint_{g(x, y) \leq z} f(x, y) dx dy, \quad f_z(z) = F'_z(z)$$

3.多维随机变量及其分布

27

7.重要公式与结论

(1) 边缘密度公式: $f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy$, $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx$

(2) $P\{(X, Y) \in D\} = \iint_D f(x, y)dxdy$

3.多维随机变量及其分布

28

(3) 若 (X, Y) 服从二维正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 则有:

1) $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2).$

2) X 与 Y 相互独立 $\Leftrightarrow \rho = 0$, 即 X 与 Y 不相关。

3) $C_1X + C_2Y \sim N(C_1\mu_1 + C_2\mu_2, C_1^2\sigma_1^2 + C_2^2\sigma_2^2 + 2C_1C_2\sigma_1\sigma_2\rho)$

4) X 关于 $Y=y$ 的条件分布为: $N(\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(y - \mu_2), \sigma_1^2(1 - \rho^2))$

5) Y 关于 $X = x$ 的条件分布为: $N(\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1 - \rho^2))$

3.多维随机变量及其分布

29

(4) 若 X 与 Y 独立, 且分别服从 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$, 则:

$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, 0), C_1X + C_2Y \sim N(C_1\mu_1 + C_2\mu_2, C_1^2\sigma_1^2 + C_2^2\sigma_2^2).$

(5) 若 X 与 Y 相互独立, $f(x)$ 和 $g(x)$ 为连续函数, 则 $f(X)$ 和 $g(Y)$ 也相互独立。

4.随机变量的数字特征

30

- 01** 随机事件和概率
- 02** 随机变量及其概率分布
- 03** 多维随机变量及其分布
- 04** 随机变量的数字特征
- 05** 数理统计的基本概念

4. 随机变量的数字特征

31

1. 数学期望

离散型: $P\{X = x_i\} = p_i, E(X) = \sum_i x_i p_i$

连续型: $X \sim f(x), E(X) = \int_{-\infty}^{+\infty} x f(x) dx$

性质:

$$(1) E(C) = C, E[E(X)] = E(X)$$

$$(2) E(C_1X + C_2Y) = C_1E(X) + C_2E(Y)$$

$$(3) \text{若 } X \text{ 和 } Y \text{ 独立, 则 } E(XY) = E(X)E(Y) \quad (4) [E(XY)]^2 \leq E(X^2)E(Y^2)$$

4. 随机变量的数字特征

32

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

3. 标准差: $\sqrt{D(X)},$

4. 离散型: $D(X) = \sum_i [x_i - E(X)]^2 p_i$

4. 随机变量的数字特征

33

5. 连续型: $D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx$

性质:

(1) $D(C) = 0, D[E(X)] = 0, D[D(X)] = 0$

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

(3) $D(C_1X + C_2) = C_1^2 D(X)$

(4) 一般有 $D(X \pm Y) = D(X) + D(Y) \pm 2Cov(X, Y) = D(X) + D(Y) \pm 2\rho\sqrt{D(X)}\sqrt{D(Y)}$

(5) $D(X) < E(X - C)^2, C \neq E(X),$ (6) $D(X) = 0 \Leftrightarrow P\{X = C\} = 1$

4. 随机变量的数字特征

34

6. 随机变量函数的数学期望

(1) 对于函数 $Y = g(x)$

X 为离散型: $P\{X = x_i\} = p_i, E(Y) = \sum_i g(x_i)p_i;$

X 为连续型: $X \sim f(x), E(Y) = \int_{-\infty}^{+\infty} g(x)f(x)dx$

(2) $Z = g(X, Y); (X, Y) \sim P\{X = x_i, Y = y_j\} = p_{ij}$

$E(Z) = \sum_i \sum_j g(x_i, y_j)p_{ij} \quad (X, Y) \sim f(x, y)$

$E(Z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y)f(x, y)dxdy$

4. 随机变量的数字特征

35

7. 协方差 $Cov(X, Y) = E[(X - E(X))(Y - E(Y))]$

8. 相关系数 $\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$, k 阶原点矩 $E(X^k)$; k 阶中心矩 $E\{[X - E(X)]^k\}$

性质:

$$(1) Cov(X, Y) = Cov(Y, X) ,$$

$$(2) Cov(aX, bY) = abCov(X, Y)$$

$$(3) Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y) ,$$

$$(4) |\rho(X, Y)| \leq 1$$

$$(5) \rho(X, Y) = 1 \Leftrightarrow P(Y = aX + b) = 1 , \text{ 其中 } a > 0$$

$$\rho(X, Y) = -1 \Leftrightarrow P(Y = aX + b) = 1 , \text{ 其中 } a < 0$$

4. 随机变量的数字特征

36

9. 重要公式与结论

$$(1) D(X) = E(X^2) - E^2(X)$$

$$(2) Cov(X, Y) = E(XY) - E(X)E(Y)$$

$$(3) |\rho(X, Y)| \leq 1, \text{ 且 } \rho(X, Y) = 1 \Leftrightarrow P(Y = aX + b) = 1, \text{ 其中 } a > 0$$

$$\rho(X, Y) = -1 \Leftrightarrow P(Y = aX + b) = 1, \text{ 其中 } a < 0$$

4. 随机变量的数字特征

37

(4) 下面5个条件互为充要条件:

$$\rho(X, Y) = 0 \Leftrightarrow \text{Cov}(X, Y) = 0 \Leftrightarrow E(X, Y) = E(X)E(Y)$$

$$\Leftrightarrow D(X + Y) = D(X) + D(Y) \Leftrightarrow D(X - Y) = D(X) + D(Y)$$

注: X 与 Y 独立为上述5个条件中任何一个成立的充分条件, 但非必要条件。

5.数理统计的基本概念

38

01 随机事件和概率

02 随机变量及其概率分布

03 多维随机变量及其分布

04 随机变量的数字特征

05 数理统计的基本概念

5.数理统计的基本概念

39

1.基本概念

总体：研究对象的全体，它是一个随机变量，用 X 表示。

个体：组成总体的每个基本元素。

简单随机样本：来自总体 X 的 n 个相互独立且与总体同分布的随机变量 X_1, X_2, \dots, X_n ，称为容量为 n 的简单随机样本，简称样本。

统计量：设 X_1, X_2, \dots, X_n 是来自总体 X 的一个样本， $g(X_1, X_2, \dots, X_n)$ 是样本的连续函数，且 $g(\cdot)$ 中不含任何未知参数，则称 $g(X_1, X_2, \dots, X_n)$ 为统计量。

5.数理统计的基本概念

40

样本均值: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

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

样本矩: 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: $B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k, k = 1, 2, \dots$

5.数理统计的基本概念

41

2.分布

χ^2 分布: $\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$, 其中 X_1, X_2, \cdots, X_n 相互独立, 且同服从 $N(0,1)$

t 分布: $T = \frac{X}{\sqrt{Y/n}} \sim t(n)$, 其中 $X \sim N(0,1), Y \sim \chi^2(n)$, 且 X, Y 相互独立。

F 分布: $F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$, 其中 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 且 X, Y 相互独立。

分位数: 若 $P(X \leq x_\alpha) = \alpha$, 则称 x_α 为 X 的 α 分位数。

5.数理统计的基本概念

42

3.正态总体的常用样本分布

(1) 设 $X_1, X_2 \cdots, X_n$ 为来自正态总体 $N(\mu, \sigma^2)$ 的样本,

$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$, 则:

$$1) \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \text{ 或者 } \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

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

$$3) \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \mu)^2 \sim \chi^2(n)$$

$$4) \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

5.数理统计的基本概念

43

4.重要公式与结论

(1) 对于 $\chi^2 \sim \chi^2(n)$, 有 $E(\chi^2(n)) = n, D(\chi^2(n)) = 2n$;

(2) 对于 $T \sim t(n)$, 有 $E(T) = 0, D(T) = \frac{n}{n-2} (n > 2)$;

(3) 对于 $F \sim F(m, n)$, 有 $\frac{1}{F} \sim F(n, m), F_{\alpha/2}(m, n) = \frac{1}{F_{1-\alpha/2}(n, m)}$;

(4) 对于任意总体 X , 有 $E(\bar{X}) = E(X), E(S^2) = D(X), D(\bar{X}) = \frac{D(X)}{n}$

参考文献

44

1. <https://github.com/fengdu78>
2. 《概率论与数理统计》，同济大学

谢谢!

