

# 概率论与随机过程

## 一、概率与分布

### 1.1 条件概率与独立事件

1. 条件概率：已知  $A$  事件发生的条件下  $B$  发生的概率，记作  $P(B | A)$ ，它等于事件  $AB$  的概率相对于事件  $A$  的概率，即：
$$P(B | A) = \frac{P(AB)}{P(A)}$$
。其中必须有  $P(A) > 0$ 。

2. 条件概率分布的链式法则：对于  $n$  个随机变量  $X_1, X_2, \dots, X_n$ ，有：

$$P(X_1, X_2, \dots, X_n) = P(X_1) \prod_{i=2}^n P(X_i | X_1, \dots, X_{i-1})$$

3. 两个随机变量  $X, Y$  相互独立的数学描述： $P(X, Y) = P(X)P(Y)$ 。记作： $X \perp Y$ 。

4. 两个随机变量  $X, Y$  关于随机变量  $Z$  条件独立的数学描述： $P(X, Y | Z) = P(X | Z)P(Y | Z)$ 。  
记作： $X \perp Y | Z$ 。

### 1.2 联合概率分布

1. 定义  $X$  和  $Y$  的联合分布为： $P(a, b) = P\{X \leq a, Y \leq b\}$ ， $-\infty < a, b < +\infty$ 。

◦  $X$  的分布可以从联合分布中得到：

$$P_X(a) = P\{X \leq a\} = P\{X \leq a, Y \leq \infty\} = P(a, \infty), \quad -\infty < a < +\infty$$

◦  $Y$  的分布可以从联合分布中得到：

$$P_Y(b) = P\{Y \leq b\} = P\{X \leq \infty, Y \leq b\} = P(\infty, b), \quad -\infty < b < +\infty$$

2. 当  $X$  和  $Y$  都是离散随机变量时，定义  $X$  和  $Y$  的联合概率质量函数为： $p(x, y) = P\{X = x, Y = y\}$   
则  $X$  和  $Y$  的概率质量函数分布为：

$$p_X(x) = \sum_y p(x, y) \quad p_Y(y) = \sum_x p(x, y)$$

3. 当  $X$  和  $Y$  联合地连续时，即存在函数  $p(x, y)$ ，使得对于所有的实数集合  $\mathbb{A}$  和  $\mathbb{B}$  满足：

$$P\{X \in \mathbb{A}, Y \in \mathbb{B}\} = \int_{\mathbb{B}} \int_{\mathbb{A}} p(x, y) dx dy$$

则函数  $p(x, y)$  称为  $X$  和  $Y$  的概率密度函数。

◦ 联合分布为： $P(a, b) = P\{X \leq a, Y \leq b\} = \int_{-\infty}^a \int_{-\infty}^b p(x, y) dx dy$ 。

◦  $X$  和  $Y$  的分布函数以及概率密度函数分别为：

$$\begin{aligned}
P_X(a) &= \int_{-\infty}^a \int_{-\infty}^{\infty} p(x, y) dx dy = \int_{-\infty}^a p_X(x) dx \\
P_Y(b) &= \int_{-\infty}^{\infty} \int_{-\infty}^b p(x, y) dx dy = \int_{-\infty}^b p_Y(y) dy \\
p_X(x) &= \int_{-\infty}^{\infty} p(x, y) dy \\
p_Y(y) &= \int_{-\infty}^{\infty} p(x, y) dx
\end{aligned}$$

## 二、期望和方差

### 2.1 期望

1. 期望描述了随机变量的平均情况，衡量了随机变量  $X$  的均值。它是概率分布的泛函（函数的函数）。

- 离散型随机变量  $X$  的期望： $\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i p_i$ 。

若右侧级数不收敛，则期望不存在。

- 连续性随机变量  $X$  的期望： $\mathbb{E}[X] = \int_{-\infty}^{\infty} x p(x) dx$ 。

若右侧极限不收敛，则期望不存在。

2. 定理：对于随机变量  $X$ ，设  $Y = g(X)$  也为随机变量， $g(\cdot)$  是连续函数。

- 若  $X$  为离散型随机变量，若  $Y$  的期望存在，则： $\mathbb{E}[Y] = \mathbb{E}[g(X)] = \sum_{i=1}^{\infty} g(x_i) p_i$ 。

也记做： $\mathbb{E}_{X \sim P(X)}[g(X)] = \sum_x g(x) p(x)$ 。

- 若  $X$  为连续型随机变量，若  $Y$  的期望存在，则： $\mathbb{E}[Y] = \mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) p(x) dx$ 。

也记做： $\mathbb{E}_{X \sim P(X)}[g(X)] = \int g(x) p(x) dx$ 。

该定理的意义在于：当求  $\mathbb{E}(Y)$  时，不必计算出  $Y$  的分布，只需要利用  $X$  的分布即可。

该定理可以推广至两个或两个以上随机变量的情况。对于随机变量  $X, Y$ ，假设  $Z = g(X, Y)$  也是随机变量， $g(\cdot)$  为连续函数，则有： $\mathbb{E}[Z] = \mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) p(x, y) dx dy$ 。也记做：

$$\mathbb{E}_{X, Y \sim P(X, Y)}[g(X, Y)] = \int g(x, y) p(x, y) dx dy。$$

3. 期望性质：

- 常数的期望就是常数本身。
- 对常数  $C$  有： $\mathbb{E}[CX] = C\mathbb{E}[X]$ 。
- 对两个随机变量  $X, Y$ ，有： $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ 。

该结论可以推广到任意有限个随机变量之和的情况。

- 对两个相互独立的随机变量，有： $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ 。

该结论可以推广到任意有限个相互独立的随机变量之积的情况。

### 2.2 方差

1. 对随机变量  $X$ ，若  $\mathbb{E}[(X - \mathbb{E}[X])^2]$  存在，则称它为  $X$  的方差，记作  $Var[X]$ 。

$X$  的标准差为方差的开平方。即：

$$\begin{aligned} \text{Var}[X] &= \mathbb{E}[(X - \mathbb{E}[X])^2] \\ \sigma &= \sqrt{\text{Var}[X]} \end{aligned}$$

- 方差度量了随机变量  $X$  与期望值偏离的程度，衡量了  $X$  取值分散程度的一个尺度。
  - 由于绝对值  $|X - \mathbb{E}[X]|$  带有绝对值，不方便运算，因此采用平方来计算。
- 又因为  $|X - \mathbb{E}[X]|^2$  是一个随机变量，因此对它取期望，即得  $X$  与期望值偏离的均值。

2. 根据定义可知：

$$\begin{aligned} \text{Var}[X] &= \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ \text{Var}[f(X)] &= \mathbb{E}[(f(X) - \mathbb{E}[f(X)])^2] \end{aligned}$$

3. 对于一个期望为  $\mu$ ，方差为  $\sigma^2, \sigma \neq 0$  的随机变量  $X$ ，随机变量  $X^* = \frac{X - \mu}{\sigma}$  的数学期望为0，方差为1。称  $X^*$  为  $X$  的标准化变量。

4. 方差的性质：

- 常数的方差恒为 0。
- 对常数  $C$ ，有  $\text{Var}[CX] = C^2 \text{Var}[X]$ 。
- 对两个随机变量  $X, Y$ ，有： $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + 2\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$   
当  $X$  和  $Y$  相互独立时，有  $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$ 。这可以推广至任意有限多个相互独立的随机变量之和的情况。
- $\text{Var}[X] = 0$  的充要条件是  $X$  以概率1取常数。

## 2.3 协方差与相关系数

1. 对于二维随机变量  $(X, Y)$ ，可以讨论描述  $X$  与  $Y$  之间相互关系的数字特征。

- 定义  $\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$  为随机变量  $X$  与  $Y$  的协方差，记作  $\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$ 。
- 定义  $\rho_{XY} = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]}\sqrt{\text{Var}[Y]}}$  为随机变量  $X$  与  $Y$  的相关系数，它是协方差的归一化。

2. 由定义可知：

$$\begin{aligned} \text{Cov}[X, Y] &= \text{Cov}[Y, X] \\ \text{Cov}[X, X] &= \text{Var}[X] \\ \text{Var}[X + Y] &= \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y] \end{aligned}$$

3. 协方差的性质：

- $\text{Cov}[aX, bY] = ab\text{Cov}[X, Y]$ ， $a, b$  为常数。
- $\text{Cov}[X_1 + X_2, Y] = \text{Cov}[X_1, Y] + \text{Cov}[X_2, Y]$
- $\text{Cov}[f(X), g(Y)] = \mathbb{E}[(f(X) - \mathbb{E}[f(X)])(g(Y) - \mathbb{E}[g(Y)])]$
- $\rho[f(X), g(Y)] = \frac{\text{Cov}[f(X), g(Y)]}{\sqrt{\text{Var}[f(X)]}\sqrt{\text{Var}[g(Y)]}}$

4. 协方差的物理意义：

- 协方差的绝对值越大，说明两个随机变量都远离它们的均值。
- 协方差如果为正，则说明两个随机变量同时趋向于取较大的值或者同时趋向于取较小的值；如果为负，则说明一个随变量趋向于取较大的值，另一个随机变量趋向于取较小的值。
- 两个随机变量的独立性可以导出协方差为零。但是两个随机变量的协方差为零无法导出独立性。

因为独立性也包括：没有非线性关系。有可能两个随机变量是非独立的，但是协方差为零。如：假设随机变量  $X \sim U[-1, 1]$ 。定义随机变量  $S$  的概率分布函数为：

$$P(S = 1) = \frac{1}{2}P(S = -1) = \frac{1}{2}$$

定义随机变量  $Y = SX$ ，则随机变量  $X, Y$  是非独立的，但是有： $Cov[X, Y] = 0$ 。

5. 相关系数的物理意义：考虑以随机变量  $X$  的线性函数  $a + bX$  来近似表示  $Y$ 。以均方误差

$$e = \mathbb{E}[(Y - (a + bX))^2] = \mathbb{E}[Y^2] + b^2\mathbb{E}[X^2] + a^2 - 2b\mathbb{E}[XY] + 2ab\mathbb{E}[X] - 2a\mathbb{E}[Y]$$

来衡量以  $a + bX$  近似表达  $Y$  的好坏程度。 $e$  越小表示近似程度越高。

为求得最好的近似，则对  $a, b$  分别取偏导数，得到：

$$a_0 = \mathbb{E}[Y] - b_0\mathbb{E}[X] = \mathbb{E}[Y] - \mathbb{E}[X] \frac{Cov[X, Y]}{Var[X]}$$

$$b_0 = \frac{Cov[X, Y]}{Var[X]}$$

$$\min(e) = \mathbb{E}[(Y - (a_0 + b_0X))^2] = (1 - \rho_{XY}^2)Var[Y]$$

因此有以下定理：

- $|\rho_{XY}| \leq 1$  ( $|\cdot|$  是绝对值)。
- $|\rho_{XY}| = 1$  的充要条件是：存在常数  $a, b$  使得  $P\{Y = a + bX\} = 1$ 。
- 6. 当  $|\rho_{XY}|$  较大时， $e$  较小，意味着随机变量  $X$  和  $Y$  联系较紧密。于是  $\rho_{XY}$  是一个表征  $X, Y$  之间线性关系紧密程度的量。
- 7. 当  $\rho_{XY} = 0$  时，称  $X$  和  $Y$  不相关。
  - 不相关是就线性关系来讲的，而相互独立是一般关系而言的。
  - 相互独立一定不相关；不相关则未必独立。

## 2.4 协方差矩阵

1. 设  $X$  和  $Y$  是随机变量。

- 若  $\mathbb{E}[X^k], k = 1, 2, \dots$  存在，则称它为  $X$  的  $k$  阶原点矩，简称  $k$  阶矩。
- 若  $\mathbb{E}[(X - \mathbb{E}[X])^k], k = 2, 3, \dots$  存在，则称它为  $X$  的  $k$  阶中心矩。
- 若  $\mathbb{E}[X^k Y^l], k, l = 1, 2, \dots$  存在，则称它为  $X$  和  $Y$  的  $k + l$  阶混合矩。
- 若  $\mathbb{E}[(X - \mathbb{E}[X])^k (Y - \mathbb{E}[Y])^l], k, l = 1, 2, \dots$  存在，则称它为  $X$  和  $Y$  的  $k + l$  阶混合中心矩。

因此：期望是一阶原点矩，方差是二阶中心矩，协方差是二阶混合中心矩。

2. 协方差矩阵：

- 二维随机变量  $(X_1, X_2)$  有四个二阶中心矩（假设他们都存在），记作：

$$c_{11} = \mathbb{E}[(X_1 - \mathbb{E}[X_1])^2]$$

$$c_{12} = \mathbb{E}[(X_1 - \mathbb{E}[X_1])(X_2 - \mathbb{E}[X_2])]$$

$$c_{21} = \mathbb{E}[(X_2 - \mathbb{E}[X_2])(X_1 - \mathbb{E}[X_1])]$$

$$c_{22} = \mathbb{E}[(X_2 - \mathbb{E}[X_2])^2]$$

称矩阵

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

为随机变量  $(X_1, X_2)$  的协方差矩阵。

- 设  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  的二阶混合中心矩

$c_{ij} = Cov[X_i, X_j] = E[(X_i - E[X_i])(X_j - E[X_j])]$  都存在, 则称矩阵

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

为  $n$  维随机变量  $(X_1, X_2, \dots, X_n)$  的协方差矩阵。

由于  $c_{ij} = c_{ji}, i \neq j, i, j = 1, 2, \dots, n$  因此协方差矩阵是个对称阵。

- 通常  $n$  维随机变量的分布是不知道的, 或者太复杂以致数学上不容易处理。因此实际中协方差矩阵非常重要。

## 三、大数定律及中心极限定理

### 3.1 切比雪夫不等式

- 切比雪夫不等式: 假设随机变量  $X$  具有期望  $E[X] = \mu$ , 方差  $Var(X) = \sigma^2$ , 则对于任意正数  $\varepsilon$ , 下面的不等式成立:

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

- 其意义是: 对于距离  $E[X]$  足够远的地方 (距离大于等于  $\varepsilon$ ), 事件出现的概率是小于等于  $\frac{\sigma^2}{\varepsilon^2}$ 。即事件出现在区间  $[\mu - \varepsilon, \mu + \varepsilon]$  的概率大于  $1 - \frac{\sigma^2}{\varepsilon^2}$ 。

该不等式给出了随机变量  $X$  在分布未知的情况下, 事件  $\{|X - \mu| \leq \varepsilon\}$  的下限估计。如:

$$P\{|X - \mu| < 3\sigma\} \geq 0.8889。$$

- 证明:

$$\begin{aligned} P\{|X - \mu| \geq \varepsilon\} &= \int_{|x-\mu| \geq \varepsilon} p(x)dx \leq \int_{|x-\mu| \geq \varepsilon} \frac{|x - \mu|^2}{\varepsilon^2} p(x)dx \\ &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{\infty} (x - \mu)^2 p(x)dx = \frac{\sigma^2}{\varepsilon^2} \end{aligned}$$

- 切比雪夫不等式的特殊情况: 设随机变量  $X_1, X_2, \dots, X_n, \dots$  相互独立, 且具有相同的数学期望和方差:  $E[X_k] = \mu, Var[X_k] = \sigma^2$ 。作前  $n$  个随机变量的算术平均:  $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ , 则对于任意正数  $\varepsilon$  有:

$$\lim_{n \rightarrow \infty} P\{|\bar{X} - \mu| < \varepsilon\} = \lim_{n \rightarrow \infty} P\{|\frac{1}{n} \sum_{k=1}^n X_k - \mu| < \varepsilon\} = 1$$

证明: 根据期望和方差的性质有:  $E[\bar{X}] = \mu, Var[\bar{X}] = \frac{\sigma^2}{n}$ 。根据切比雪夫不等式有:

$$P\{|\bar{X} - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{n\varepsilon^2}$$

则有  $\lim_{n \rightarrow \infty} P\{|\bar{X} - \mu| \geq \varepsilon\} = 0$ , 因此有:  $\lim_{n \rightarrow \infty} P\{|\bar{X} - \mu| < \varepsilon\} = 1$ 。

### 3.2 大数定理

1. 依概率收敛：设  $Y_1, Y_2, \dots, Y_n, \dots$  是一个随机变量序列， $a$  是一个常数。

若对于任意正数  $\varepsilon$  有： $\lim_{n \rightarrow \infty} P\{|Y_n - a| \leq \varepsilon\} = 1$ ，则称序列  $Y_1, Y_2, \dots, Y_n, \dots$  依概率收敛于  $a$ 。

记作： $Y_n \xrightarrow{P} a$

2. 依概率收敛的两个含义：

- 收敛：表明这是一个随机变量序列，而不是某个随机变量；且序列是无限长，而不是有限长。
- 依概率：表明序列无穷远处的随机变量  $Y_\infty$  的分布规律为：绝大部分分布于点  $a$ ，极少数位于  $a$  之外。且分布于  $a$  之外的事件发生的概率之和为0。

3. 大数定理一：设随机变量  $X_1, X_2, \dots, X_n, \dots$  相互独立，且具有相同的数学期望和方差：

$E[X_k] = \mu, Var[X_k] = \sigma^2$ 。则序列： $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$  依概率收敛于  $\mu$ ，即  $\bar{X} \xrightarrow{P} \mu$ 。

注意：这里并没有要求随机变量  $X_1, X_2, \dots, X_n, \dots$  同分布。

4. 伯努利大数定理：设  $n_A$  为  $n$  次独立重复实验中事件  $A$  发生的次数， $p$  是事件  $A$  在每次试验中发生的概率。则对于任意正数  $\varepsilon$  有：

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1$$
$$or: \lim_{n \rightarrow \infty} P\left\{\left|\frac{n_A}{n} - p\right| \geq \varepsilon\right\} = 0$$

即：当独立重复实验执行非常大的次数时，事件  $A$  发生的频率逼近于它的概率。

5. 辛钦定理：设随机变量  $X_1, X_2, \dots, X_n, \dots$  相互独立，服从同一分布，且具有相同的数学期望：

$E[X_k] = \mu$ 。则对于任意正数  $\varepsilon$  有：

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} \sum_{k=1}^n X_k - \mu\right| < \varepsilon\right\} = 1$$

- 注意：这里并没有要求随机变量  $X_1, X_2, \dots, X_n, \dots$  的方差存在。
- 伯努利大数定理是辛钦定理的特殊情况。

### 3.3 中心极限定理

1. 独立同分布的中心极限定理：设随机变量  $X_1, X_2, \dots, X_n$  独立同分布，且具有数学期望和方差：

$E[X_k] = \mu, Var[X_k] = \sigma^2$ ，则随机变量之和  $SX_n = \sum_{k=1}^n X_k$  的标准变化量：

$$Y_n = \frac{SX_n - E[SX_n]}{\sqrt{Var[SX_n]}} = \frac{SX_n - n\mu}{\sqrt{n}\sigma}$$

的概率分布函数  $F_n(x)$  对于任意  $x$  满足：

$$\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P\{Y_n \leq x\} = \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} \leq x\right\}$$
$$= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x)$$

- 其物理意义为：均值方差为  $\mu, \sigma^2$  的独立同分布的随机变量  $X_1, X_2, \dots, X_n$  之和  $SX_n = \sum_{k=1}^n X_k$  的标准变化量  $Y_n$ ，当  $n$  充分大时，其分布近似于标准正态分布。

即： $SX_n = \sum_{k=1}^n X_k$  在  $n$  充分大时，其分布近似于  $N(n\mu, n\sigma^2)$ 。

- 一般情况下，很难求出  $n$  个随机变量之和的分布函数。因此当  $n$  充分大时，可以通过正态分布来做理论上的分析或者计算。

2. **Liapunov** 定理: 设随机变量  $X_1, X_2, \dots, X_n, \dots$  相互独立, 具有数学期望和方差:

$\mathbb{E}[X_k] = \mu_k, \text{Var}[X_k] = \sigma_k^2$ 。记:  $B_n^2 = \sum_{k=1}^n \sigma_k^2$ 。若存在正数  $\delta$ , 使得当  $n \rightarrow \infty$  时,  $\frac{1}{B_n^{2+\delta}} \sum_{k=1}^n \mathbb{E}[|X_k - \mu_k|^{2+\delta}] \rightarrow 0$ 。则随机变量之和  $\overline{SX_n} = \sum_{k=1}^n X_k$  的标准变化量:

$$Z_n = \frac{\overline{SX_n} - \mathbb{E}[\overline{SX_n}]}{\sqrt{\text{Var}[\overline{SX_n}]}} = \frac{\overline{SX_n} - \sum_{k=1}^n \mu_k}{B_n}$$

的概率分布函数  $F_n(x)$  对于任意  $x$  满足:

$$\begin{aligned} \lim_{n \rightarrow \infty} F_n(x) &= \lim_{n \rightarrow \infty} P\{Z_n \leq x\} = \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{k=1}^n X_k - \sum_{k=1}^n \mu_k}{B_n} \leq x\right\} \\ &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt = \Phi(x) \end{aligned}$$

- 其物理意义为: 相互独立的随机变量  $X_1, X_2, \dots, X_n, \dots$  之和  $\overline{SX_n} = \sum_{k=1}^n X_k$  的衍生随机变量序列  $Z_n = \frac{\overline{SX_n} - \sum_{k=1}^n \mu_k}{B_n}$ , 当  $n$  充分大时, 其分布近似与标准正态分布。
- 这里并不要求  $X_1, X_2, \dots, X_n, \dots$  同分布。

9. **Demoiver-Laplace** 定理: 设随机变量序列  $\eta_n, n = 1, 2, \dots$  服从参数为  $(n, p)$  的二项分布, 其中  $0 < p < 1$ 。则对于任意  $x$ , 有:

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

该定理表明, 正态分布是二项分布的极限分布。当  $n$  充分大时, 可以利用正态分布来计算二项分布的概率。

## 五、常见概率分布

### 5.1 均匀分布

1. 离散随机变量的均匀分布: 假设  $X$  有  $k$  个取值:  $x_1, x_2, \dots$ ,