# Ch-02 随机变量与概率分布

## 2. 1 随机变量的概念

**定义 1. 1 随机变量的直观描述** 如果条件 $S$ 实现下的情况可以用一个数量 $X$ 来描述，$X$ 究竟等于多少不能预先确定，而随着条件 $S$ 下的结果不同而可能变化，但对任何实数 $c$，事件"$X$ 取值不超过 $c$"是有概率的，则把这样一种变量 $X$ 叫做**随机变量**。

**定义 1. 1' 随机变量的数学描述** 如果条件 $S$ 下所有可能的结果组成集合 $\Omega = \{\omega\}$，$X = X(\omega)$ 是 $\Omega$ 上有定义的实值函数，而且对任何实数 $c$，事件 $\{\omega : X(\omega) \le c\}$ 是有概率的，则称 $X$ 是**随机变量**。

如果随机变量 $X$ 所可能取的值只有有限个或者可排成一个无穷序列，则称 $X$ 为**离散型随机变量**。

## 2. 2 离散型随机变量

**定义 2. 1** 设 $X$ 的可能取值是 $x_1, x_2, \cdots$（有限个或可列无穷个），则称一列数

$$p_k = P(X = x_k)\ (k = 1, 2, \cdots)$$

为 $X$ 的**概率分布**（列），也称**概率函数**。关于 $\{p_k\}$，有以下性质

1. $p_k \ge 0\ (k = 1, 2, \cdots)$
2. $\sum_k p_k = 1$

### 1. 两点分布（伯努利分布）

**定义 2. 2** 如果随机变量 $X$ 的可能值是 0 和 1 且概率分布为

$$P(X = 1) = p,\ P(X = 0) = 1 - p \quad (0 \le p \le 1)$$

则称 $X$ 服从**两点分布**（也称**伯努利分布**），记作 $X \sim B(1, p)$.

示性函数

$$1_A(\omega) \begin{cases} 1, & \text{若 } \omega \in A \\ 0, & \text{否则} \end{cases}$$

**2.** 二项分布

**定义 2.3** 设随机变量 $X$ 的所有可能值是 $0,1,\cdots,n$，且

$$P(X=k)=C_n^k p^k(1-p)^{n-k}\ (k=0,1,\cdots,n)$$

这里 $n\geq 1$，$0\leq p\leq 1$，则称 $X$ 服从参数为 $n,p$ 的**二项分布**，记作 $X\sim B(n,p)$.

**定理 2.1** 设 $n\geq 2, 0<p<1, m=[(n+1)p]$ ( 不超过 $(n+1)p$ 的最大整数 )，

$$p_n(k)=C_n^k p^k(1-p)^{n-k}\ (k=0,1,\cdots,n)$$

则有以下结论：

1. 当 $(n+1)p$ 不是整数时，
$$\begin{aligned}&p_n(0)<p_n(1)<\cdots<p_n(m-1)<p_n(m)\\&>p_n(m+1)>\cdots>p_n(n)\end{aligned}$$
2. 当 $(n+1)p$ 是整数时，
$$\begin{aligned}&p_n(0)<p_n(1)<\cdots<p_n(m-1)=p_n(m)\\&>p_n(m+1)>\cdots>p_n(n)\end{aligned}$$

**3.** 泊松(Possion)分布

**定义 2.4** 设随机变量 $X$ 的所有可能值是全体非负整数，且

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

（其中 $\lambda$ 是正数），则称 $X$ 服从参数为 $\lambda$ 的**泊松分布**，记作 $X\sim P(\lambda)$

**定理 2.2** 设 $X$ 服从泊松分布，$p_k=P(X=k)=\dfrac{\lambda^k}{k!}e^{-\lambda}\ (\lambda>0;k=0,1,2,\cdots)$，则有下列结论：

1. 当 $\lambda$ 不是整数时，
$$p_0<p_1<\cdots<p_{[\lambda]}>p_{[\lambda]+1}>\cdots$$
（这里 $[\lambda]$ 是不超过 $\lambda$ 的最大整数）
2. 当 $\lambda$ 是整数时，
$$p_0<p_1<\cdots<p_{\lambda-1}=p_\lambda>p_{\lambda+1}>\cdots$$

**4. 超几何分布**

**定义 2.5** 若随机变量的概率分布是 $P(X=k)=\dfrac{C_D^k C_{N-D}^{n-k}}{C_N^n}\ (k=0,1,\cdots,l)$，其中 $N\geq D\geq 0$，$N\geq n\geq 1$，则称 $X$ 服从**超几何分布**。

**定理 2.3** 设超几何分布中 $D$ 是 $N$ 的函数，$D=D(N)$ 且 $\lim\limits_{N\to\infty}\dfrac{D(N)}{N}=p\ (0<p<1)$，则

$$\lim_{N\to\infty}\frac{C_{D(N)}^k C_{N-D(N)}^{n-k}}{C_N^n}=C_n^k p^k(1-p)^{n-k}$$

**5. 几何分布**

**定义 2.6** 设随机变量 $X$ 的所有可能值是全体正整数，且

$$P(X=k)=(1-p)^{k-1}p \quad (k=1,2,\cdots)$$

（其中 $0<p<1$），则称 $X$ 服从**几何分布**。

**6. 负二项分布**

**定义 2.7** 设随机变量 $X$ 的取值范围是 $\{r,r+1,\cdots\}$（$r$ 是正整数），且

$$P(X=k)=C_{k-1}^{r-1}p^r(1-p)^{k-r}\ (k=r,r+1,\cdots;\ 0<p<1)$$

则称 $X$ 服从**负二项分布**，记作 $X\sim NB(r,p)$.

**7. 离散均匀分布**

**定义 2.8** 设随机变量 $X$ 的取值范围是 $\{1,2,\cdots,N\}$（$N$ 是大于 1 的整数），且

$$P(X=k)=\frac{1}{N}\ (k=1,2,\cdots,N)$$

则称 $X$ 服从**离散均匀分布**。

## 2.3 连续型随机变量

**定义 3.1** 对于随机变量 $X$，如果存在非负函数 $p(x)$，使对任意 $a<b$，都有

$$P(a<X<b)=\int_a^b p(x)\mathrm{d}x$$

则称 $X$ 为**连续型随机变量**，并称 $p(x)$ 为 $X$ 的**概率密度函数**（简称**密度函数**或**分布密度**）。

作为密度函数的 $p(x)$ 满足：$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1$，对任何实数 $a$，$P(X=a)=0$. 注意，$p(x) \neq P(X=x)$.

### 1. 均匀分布

**定义 3.2** 如果随机变量 $X$ 的概率密度函数为

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{否则}\end{cases}$$

则称 $X$ 服从区间 $[a,b]$ 上的**均匀分布**，记作 $X \sim U(a,b)$.

其他写法为 $p(x)=\dfrac{1}{b-a}1_{\{a\leq x\leq b\}}$ 或 $p(x)=\dfrac{1}{b-a}$，其中 $a \leq x \leq b$.

### 2. 指数分布

**定义 3.3** 如果随机变量 $X$ 的概率密度函数为

$$p(x)=\begin{cases}\lambda e^{-\lambda x}. & x>0 \\ 0, & x \leq 0\end{cases} \quad (\lambda>0)$$

则称 $X$ 服从参数为 $\lambda$ 的**指数分布**，记为 $X \sim \operatorname{Exp}(\lambda)$。

**定理 3.1** 设 $X$ 是只取非负值的随机变量，则对任何 $s \geq 0, t \geq 0$，等式

$$P(X>s+t \mid X>s)=P(X>t)$$

恒成立的充要条件是 $X$ 服从指数分布。

### 3. 正态分布

**定义 3.4** 如果随机变量 $X$ 的概率密度函数为

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

（$\mu$ 是实数，$\sigma$ 是正数），则称 $X$ 服从参数为 $\mu, \sigma$ 的**正态分布**，也叫**高斯分布**，并记为 $X \sim N(\mu, \sigma^2)$

参数 $\mu=0, \sigma^2=1$ 时的正态分布叫做**标准正态分布**，密度函数 $p(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$，且 $\int_{-\infty}^{+\infty}\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x=1$

**定理 3.2** 设随机变量 $X \sim N(\mu, \sigma^2)$，则对一切 $a < b$，

$$P(a < X < b) = \phi(\frac{b-\mu}{\sigma}) - \phi(\frac{a-\mu}{\sigma})$$

**推论 3.1** 设随机变量 $X \sim N(\mu, \sigma^2)$，则对一切正数 $k$，有

$$P(\mu - k\sigma < X < \mu + k\sigma) = 2\Phi(k) - 1$$

**4. 威布尔（Weibull）分布**

**定义 3.5** 称随机变量 $X$ 服从**威布尔分布**，若它有密度函数

$$p(x) = \begin{cases} \dfrac{m}{\eta^m} x^{m-1} e^{-(\frac{x}{\eta})^m}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

其中 $m, \eta$ 是两个正参数，$m$ 叫做形状参数，$\eta$ 叫做刻度参数。

**5. 伽马分布（Γ 分布）**

**定义 3.6** 称随机变量 $X$ 服从**伽马分布（Γ 分布）**，若它有密度函数

$$p(x) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

其中 $\alpha$、$\beta$ 是两个正参数，$\Gamma(\alpha) = \int_0^{+\infty} u^{\alpha-1} e^{-u} \, \mathrm{d}u$ 是 $\Gamma$ 函数。

**6. 帕累托（Pareto）分布**

**定义 3.7** 称随机变量 $X$ 服从**帕累托分布**，若它有密度函数

$$p(x) = \begin{cases} (\alpha-1) x_0^{\alpha-1} x^{-\alpha}, & x \geq x_0 \\ 0, & x < x_0 \end{cases}$$

其中 $x_0$ 是正参数，$\alpha$ 是大于 1 的参数。

**7. 贝塔分布**

**定义 3.8** 称随机变量 $X$ 服从**贝塔分布**，若它有密度函数

$$p(x) = \begin{cases} \dfrac{1}{\mathrm{B}(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, & 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases}$$

其中 $\alpha,\beta$ 是正参数，$\mathrm{B}(\alpha,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,\mathrm{d}x$ 是贝塔函数。

## 2.4 随机变量的严格定义与分布函数

**定义 4.1** 设 $(\Omega,F,P)$ 是概率空间，$X=X(\omega)$ 是 $\Omega$ 上有定义的实值函数，如果对任何实数 $x$，集合 $\{\omega: X(\omega\le x)\}$ 属于 $F$，则称 $X$ 是 $(\Omega,F,P)$ 上的**随机变量**（简称 $X$ 是随机变量或条件 $S$ 下的随机变量）。

1. 随机变量 $X=X(\omega)$ 是基本事件 $\omega$ 的函数，它体现随机而变
2. 虽然 $X=X(\omega)$ 的值不能预先确定（因为无法预料将出现什么样的 $\omega$），但对给定的 $x$，事件 $\{\omega: X(\omega)\le x\}$ 是有确定概率的，这体现了随机变量的一种“规则性”，不是乱变到不可控制的程度。

**定理 4.1** 设 $X=X(\omega)$ 是 $(\Omega,F,P)$ 上的随机变量，则对直线上的任何 Borel 集 $B$，有

$$\{X\in B\}\in F$$

**定义 4.2** 设 $X=X(\omega)$ 是随机变量，则称函数

$$F(x)=P(X\le x)\quad(\text{一切实数 } x)$$

为 $X$ 的**分布函数**，有时记为 $F_X(x)$.

**定理 4.2** 分布函数 $F(x)$ 有下列三条性质：

1. 单调性：若 $a<b$，则 $F(a)\le F(b)$
2. $\lim\limits_{x\to-\infty}F(x)=0$，$\lim\limits_{x\to+\infty}F(x)=1$
3. 右连续性：$\lim\limits_{\delta\to 0+}F(x+\delta)=F(x)$

**定理 4.3** 设随机变量 $X$ 的分布函数 $F(x)$ 具有性质：$F'(x)$ 处处存在且是 $x$ 的连续函数，则 $X$ 是连续型的，且 $F'(x)$ 就是 $X$ 的概率密度函数。

**定理 4.4** 设随机变量 $X$ 的分布函数 $F(x)$ 连续且除有限个点 $c_1<c_2<\cdots<c_k$ 外 $F'(x)$ 存在且连续，则 $X$ 是连续型的且下列函数

$$p(x)=\begin{cases}F'(x), & \text{当 } x\ne c_1,c_2,\cdots,c_k \text{ 时}\\ a_i, & \text{当 } x=c_i\ (i=1,2,\cdots,k) \text{ 时}\end{cases}$$

（$a_1,a_2,\cdots,a_k$ 是任意的非负数）是 $X$ 的密度函数。

## 2. 5 随机变量的函数

**定理 5. 1** 设 $X = X(\omega)$ 是概率空间 $(\Omega, F, P)$ 上的随机变量，则对任何 Borel 函数 $f(x)$ ，$Y = f(X(\omega))$ 也是这个概率空间上的随机变量。

## 2. 6 随机变量的数学期望

### 1. 离散型随机变量的数学期望

**定义 6. 1** 设离散型随机变量的概率分布为 $p_k = P(X = x_k)$ $(k = 1, 2, \cdots)$，$X$ 的可能值为 $x_1, x_2, \cdots$（有限个或可列无穷个）。则称和数 $\sum_k x_k p_k$ 为 $X$ 的**数学期望**（简称**期望**或**均值**），记为 $E(X)$ 或 $EX$.

对常见离散型随机变量计算期望：

1. 两点分布。$E(X) = p$
2. 二项分布。$E(X) = np$
3. 泊松分布。$E(X) = \lambda$
4. 几何分布。$E(X) = \frac{1}{p}$
5. 负二项分布。$E(X) = \frac{r}{p}$
6. 离散均匀分布。$E(X) = \frac{N+1}{2}$
7. 超几何分布。$E(X) = \frac{n}{N}D$

### 2. 一般随机变量的数学期望

**定理 6. 1** 设随机变量 $X$ 有概率密度函数 $p(x)$，且积分

$$\int_{-\infty}^{+\infty} |x|\, p(x) \mathrm{d}x$$

收敛，则 $E(X)$ 存在且

$$E(X) = \int_{-\infty}^{+\infty} x p(x) \mathrm{d}x$$

计算常见离散型随机变量的数学期望：

1. 均匀分布。$E(X) = \frac{a+b}{2}$
2. 指数分布。$E(X) = \frac{1}{\lambda}$
3. 正太分布。$E(X) = \mu$
4. 伽马分布。$E(X) = \frac{\alpha}{\beta}$

**3. 数学期望的性质**

**定理 6.2** 设 $X, Y$ 是随机变量

1. 若 $X = a$（常量），则 $E(X) = a$
2. 若 $X \geq 0$（即对一切 $\omega$，$X(\omega) \geq 0$）且 $E(X)$ 存在，则 $E(X) \geq 0$
3. 若 $X$ 与 $Y$ 有相同的分布函数且 $E(X)$ 存在，则 $E(Y)$ 存在且 $E(X) = E(Y)$

**推论 6.1** 设 $P(X = Y) = 1$ 且 $E(X)$ 存在，则 $E(Y)$ 存在且 $E(X) = E(Y)$

**定理 6.3** 设 $X = X(\omega)$ 的期望 $E(X)$ 存在，$Y = Y(\omega)$ 的期望 $E(Y)$ 也存在，则

1. 对任何实数 $a$，$\xi = \xi(\omega) \triangleq aX(\omega)$ 的期望存在且 $E(\xi) = aE(X)$
2. $\eta = \eta(\omega) \triangleq X(\omega) + Y(\omega)$ 的期望存在且 $E(\eta) = E(X+Y) = E(X) + E(Y)$
3. 若 $X \geq Y$（即对一切 $\omega$，$X(\omega) \geq Y(\omega)$），则 $E(X) \geq E(Y)$

**推论 6.2** (1) 设 $X = X(\omega)$ 和 $Y = Y(\omega)$ 都是随机变量，$E(X)$ 和 $E(Y)$ 都存在，$a$ 和 $b$ 是实数，$\xi = \xi(\omega) = \triangleq aX(\omega) + bY(\omega)$，则

$$E(\xi) = aE(X) + bE(Y)$$

(2) 设 $X_i = X_i(\omega)$ $(i = 1, 2, \cdots, n)$ 都是随机变量，$E(X_i)$ $(i = 1, 2, \cdots, n)$ 都存在，$\eta = \eta(\omega) = \triangleq \sum_{i=1}^{n} X_i(\omega)$，则

$$E(\eta) = \sum_{i=1}^{n} E(X_i)$$

**定理 6.4 马尔可夫不等式** 设 $X = X(\omega)$ 是任何非负随机变量（即对一切 $\omega$，$X(\omega) \geq 0$）且 $E(X)$ 存在，则对任何 $C > 0$，有

$$P(X \geq C) \leq \frac{1}{C} E(X)$$

**4. 随机变量函数的数学期望**

**定理 6.5 均值公式** (1) 设 $X$ 是离散型随机变量，可能值是 $x_1, x_2, \cdots$（有限个或可列无穷个），概率分布是 $p_k = P(X = x_k)$ $(k = 1, 2, \cdots)$. 若 $f(x)$ 是任何函数，则

$$Ef(X) = \sum_{k} f(x_k) p_k$$

（当 $x_1, x_2, \cdots$ 是无穷序列时要求级数绝对收敛）

(2) 设 $X$ 是连续型随机变量，密度函数是 $p(x)$，若函数 $f(x)$ 使得积分 $\int_{-\infty}^{+\infty} \mid f(x) \mid p(x) \mathrm{d}x$ 收敛，则

$$Ef(X) = \int_{-\infty}^{+\infty} f(x) p(x) \, \mathrm{d}x$$

## 2.7 随机变量的方差及其他数字特征

**定义 7.1** 设 $X$ 是随机变量，$E(X)$ 存在且 $E(X-E(X))^2$ 也存在，则称 $E(X-E(X))^2$ 为 $X$ 的方差，记为 $var(X)$.（有时用 $D(X)$ 表示）

**定理 7.1 切比雪夫不等式** 设 $X$ 是随机变量，期望 $E(X)$ 和方差 $var(X)$ 都存在，则对任何 $\epsilon>0$ 成立：

$$P(|X-E(X)|\geq\epsilon)\leq\frac{1}{\epsilon^2}var(X)$$

**推论 7.1** 设随机变量 $X$ 的方差为 0，则 $P(X=E(X))=1$.

**定义 7.2** 设 $X$ 是随机变量，其方差 $var(X)$ 存在，则称 $\sqrt{var(X)}$ 为 $X$ 的**标准差**。

**定理 7.2** (1) 设 $X$ 是离散型随机变量，可能值是 $x_1,x_2,\cdots$（有限个或可列无穷个），$p_k=P(X=x_k)$ $(k=1,2,\cdots)$，则

$$var(X)=\sum_k x_k^2p_k-(E(X))^2$$

（当 $x_1,x_2,\cdots$ 是无穷序列时要求级数 $\sum\limits_{k=1}^{\infty}x_k^2p_k$ 收敛）

(2) 设 $X$ 是连续型随机变量，密度函数是 $p(x)$ 且积分 $\int_{-\infty}^{+\infty}x^2p(x)\mathrm{d}x$ 收敛，则

$$var(X)=\int_{-\infty}^{+\infty}x^2p(x)\mathrm{d}x-(E(X))^2$$

计算常见随机变量的方差：

1. 两点分布。$var(X)=p(1-p)$
2. 二项分布。$var(X)=np(1-p)$
3. 泊松分布。$var(X)=\lambda$
4. 均匀分布。$var(X)=\dfrac{(b-a)^2}{12}$
5. 指数分布。$var(X)=\dfrac{1}{\lambda^2}$
6. 正态分布。$var(X)=\sigma^2$
7. 伽马分布。$var(X)=\dfrac{\alpha}{\beta^2}$
8. 帕累托分布。$var(X)=\dfrac{(\alpha-1)x_0^2}{(\alpha-3)(\alpha-2)^2}$

**定义 7.3** 设 $X$ 是随机变量，如果 $E(X^k)$ 存在（$k$ 是正整数），则称 $E(X^k)$ 为 $X$ 的**$k$ 阶原点矩**，常把 $E(X^k)$ 记为 $\nu_k$.

**定义 7.4** 设 $X$ 是随机变量，如果 $E(X)$ 存在且 $E(X-E(X))^k$（$k$ 是正整数）存在，则称 $E(X-E(X))^k$ 为 $X$ 的**$k$ 阶中心矩**，常把 $E(X-E(X))^k$ 记为 $\mu_k$.

令 $\sigma=\sqrt{var(X)},\alpha=\dfrac{\mu_3}{\sigma^3},\gamma=\dfrac{\mu_4}{\sigma^4}$，通常称 $\alpha$ 为 $X$ 的**偏度系数**，$\gamma$ 为 $X$ 的**峰度系数**。$\alpha$ 刻画的是随机变量取值关于均值的对称程度；若 $X$ 有分布密度，$\gamma$ 用来刻画分布密度曲线的陡峭程度。

**定义 7.5** 设 $X$ 是随机变量，$0 < p < 1$，称 $a$ 是 $X$ 的 $p$ **分位数**（也称 $p$ 分位点），若成立

$$P(X \le a) \ge p \ge P(X > a)$$

常用 $a_p$ 表示 $p$ 分位数，0.5 分位数又叫**中位数**，记为 $\mathrm{med}(X)$.

**定理 7.3** 设 $F(x)$ 是 $X$ 的分布函数，给定 $p \in (0,1)$，则 $p$ 分位数唯一的充要条件是 方程 $F(x) = p$ 至多有一个根。

**定义 7.6** (1) 设 $X$ 是离散型随机变量，可能值为 $x_1, x_2, \cdots$（有限个或可列无穷个），称 $x_k$ 是 $X$ 的**众数**，若

$$P(X = x_k) \ge P(X = x_i) \quad (一切\ x_i)$$

(2) 设 $X$ 是连续型随机变量，密度函数是 $p(x)$，称 $p(x)$ 的最大值点 $x_0$ 为 $X$ 的**众数**。

## 2.8 补充知识

**定义 8.1** 设 $X$ 是非负随机变量，$E(X) > 0$ 且方差 $var(X)$ 存在，则称

$$v \triangleq \frac{\sqrt{var(X)}}{E(X)}$$

为 $X$ 的**变异系数**。

**定义 8.2** 称

$$L(u) = \frac{1}{u}\int_0^u F^{-1}(p)\mathrm{d}p \quad (0 \le u \le 1)$$

为**洛伦茨函数**，其图像叫做随机变量 $X$(或其分布函数 $F(x)$)的**洛伦茨曲线**。

**定义 8.3** 设 $L(u)(0 \le u \le 1)$ 是 $X$ 的洛伦茨函数，则称

$$G = 1 - 2\int_0^1 L(u)\mathrm{d}u$$

为 $X$（或其分布函数）的**基尼系数**。