

第五章 随机变量的数字特征

分布函数能够完整地描述随机变量的统计特性，但在一些实际问题中，只需知道随机变量的某些特征，因而不需要求出它的分布函数。

例如：

评定某企业的经营能力时，只要知道该企业**人均赢利水平**；

研究水稻品种优劣时，我们关心的是稻穗的**平均粒数**及每粒的**平均重量**；

检验棉花的质量时，既要注意纤维的**平均长度**，又要注意**纤维长度与平均长度的偏离程度**，平均长度越长、偏离程度越小，质量就越好；

考察一射手的水平，既要看他的**平均环数**是否高，还要看他弹着点的范围是否小，即**数据的波动**是否小。

由上面例子看到，与随机变量有关的某些数值，虽不能完整地描述随机变量，但能清晰地描述随机变量在某些方面的重要特征，这些数字特征在理论和实践上都具有重要意义。

随机变量某一方面的概率特性
可用**数字**来描写

本章内容

- 随机变量的平均取值 —— 数学期望
- 随机变量取值平均偏离平均值的情况 —— 方差
- 描述两个随机变量之间的某种关系的数 —— 协方差与相关系数

§5.1 随机变量的数学期望

引例1 测量 50 个圆柱形零件直径（见下表）

尺寸 (cm)	8	9	10	11	12	
数量 (个)	8	7	15	10	10	50

则这 50 个零件的平均直径为

$$\frac{8 \times 8 + 9 \times 7 + 10 \times 15 + 11 \times 10 + 12 \times 10}{50}$$

$$= 8 \times \frac{8}{50} + 9 \times \frac{7}{50} + 10 \times \frac{15}{50} + 11 \times \frac{10}{50} + 12 \times \frac{10}{50} = 10.14 \text{ cm}$$

换一个角度看，从一堆零件中任取一个零件，它的尺寸为随机变量 X ，且 X 的概率分布为

X	8	9	10	11	12
P	$\frac{8}{50}$	$\frac{7}{50}$	$\frac{15}{50}$	$\frac{10}{50}$	$\frac{10}{50}$

则加权平均

$$\begin{aligned} & 8 \times \frac{8}{50} + 9 \times \frac{7}{50} + 10 \times \frac{15}{50} + 11 \times \frac{10}{50} + 12 \times \frac{10}{50} \\ &= \sum_{k=8}^{12} k \times P(X = k) = \sum_{k=8}^{12} k p_k = 10.14 \end{aligned}$$

可以表示这堆零件的平均直径，称之为数学期望

数学期望的定义

定义1 设 X 为离散型随机变量，其概率分布为

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

若无穷级数

$$\sum_{k=1}^{+\infty} x_k p_k$$

绝对收敛，则称其和为随机变量 X 的**数学期望**，
记作 $E(X)$

$$E(X) = \sum_{k=1}^{+\infty} x_k p_k$$

定义2 设 X 为连续型随机变量, 其密度函数为 $f(x)$

若积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛,

则称此积分为随机变量 X 的**数学期望**,
记作 $E(X)$

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

随机变量的**数学期望的本质** —— 加权平均,
它是一个数不再是随机变量

例1 $X \sim B(n, p)$, 求 $E(X)$.

解

$$\begin{aligned} E(X) &= \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{k=0}^{n-1} C_{n-1}^k p^k (1-p)^{(n-1)-k} \\ &= np \end{aligned}$$

例2 $X \sim N(\mu, \sigma^2)$, 求 $E(X)$.

解

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &\stackrel{\text{令 } \frac{x-\mu}{\sigma}=y}{=} \int_{-\infty}^{+\infty} (y\sigma + \mu) \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2}} \sigma dy \\ &= \int_{-\infty}^{+\infty} y\sigma \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &\quad + \int_{-\infty}^{+\infty} \mu \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= \mu \end{aligned}$$

常见随机变量的数学期望

分布	概率分布	期望
参数为 p 的 0-1分布	$P(X = 1) = p$ $P(X = 0) = 1 - p$	p
$B(n, p)$	$P(X = k) = C_n^k p^k (1 - p)^{n-k}$ $k = 0, 1, 2, \dots, n$	np
$P(\lambda)$	$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ $k = 0, 1, 2, \dots$	λ

分布	概率密度	期望
区间 (a,b) 上的 均匀分布	$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & \text{其它} \end{cases}$	$\frac{a+b}{2}$
$E(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{其它} \end{cases}$	$\frac{1}{\lambda}$
$N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ

注意：不是所有的随机变量都有数学期望

例如：Cauchy分布的密度函数为

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

但 $\int_{-\infty}^{+\infty} |x| f(x) dx = \int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)} dx$ 发散

它的数学期望不存在

随机变量函数的数学期望

□ 设 X 为离散型随机变量，概率分布为

$$P(X = x_i) = p_i, \quad i = 1, 2, \dots$$

$$Y = g(X),$$

若级数 $\sum_{i=1}^{\infty} g(x_i) p_i$ 绝对收敛，则

$$E(Y) = \sum_{i=1}^{\infty} g(x_i) p_i$$

$$E(g(X)) = \sum_{i=1}^{\infty} g(x_i) p_i \quad E(Y) = \sum_{i=1}^{\infty} y_i p(Y = y_i)$$

$$\begin{aligned}
E(Y) &= \sum_{i=1}^{\infty} y_i p(Y = y_i) \\
&= \sum_{i=1}^{\infty} \left\{ g(x_j) \sum_{g(x_j)=y_i} p(x_j) \right\} \\
&= \sum_{i=1}^{\infty} \left\{ \sum_{g(x_j)=y_i} g(x_j) p(x_j) \right\} \\
&= \sum_{k=1}^{\infty} g(x_k) p(x_k) \\
&= \sum_{i=1}^{\infty} g(x_i) p_i
\end{aligned}$$

□ 设 X 为连续型随机变量，密度函数为 $f(x)$

$$Y = g(X),$$

若积分 $\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛，则

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

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

□ 设 (X, Y) 为二维离散型随机变量，概率分布为

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

$$Z = g(X, Y),$$

若级数 $\sum_{i,j=1}^{\infty} g(x_i, y_j) p_{ij}$ 绝对收敛，则

$$E(Z) = \sum_{i,j=1}^{\infty} g(x_i, y_j) p_{ij}$$

$$E(g(X, Y)) = \sum_{i,j=1}^{\infty} g(x_i, y_j) p_{ij}$$

□ 设 (X, Y) 为二维连续型随机变量，
密度函数为 $f(x, y)$

$$Z = g(X, Y),$$

若积分 $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$ 绝对收敛，则

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

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

几个重要的随机变量函数的数学期望

$E(X^k)$ —— X 的 k 阶原点矩

$E(X)$ —— X 的 数学期望

$E(|X|^k)$ —— X 的 k 阶绝对原点矩

$E((X - E(X))^k)$ —— X 的 k 阶中心矩

$E((X - E(X))^2) = D(X)$ —— X 的 方差

$E(X^k Y^l)$ —— X, Y 的 $k + l$ 阶混合原点矩

$E((X - E(X))^k (Y - E(Y))^l)$
—— X, Y 的 $k + l$ 阶混合中心矩

$E(XY)$ —— X, Y 的二阶原点矩

$E((X - E(X))(Y - E(Y)))$
—— X, Y 的二阶混合中心矩
 X, Y 的协方差

$$E\left(\frac{(X - E(X))(Y - E(Y))}{\sqrt{D(X)}\sqrt{D(Y)}}\right) = \rho_{XY}$$

—— X, Y 的相关系数

例4 设二维连续型随机变量 (X, Y) 的密度函数为

$$f(x, y) = \begin{cases} \frac{1}{4}x(1+3y^2), & 0 < x < 2, 0 < y < 1, \\ 0, & \text{其它} \end{cases}$$

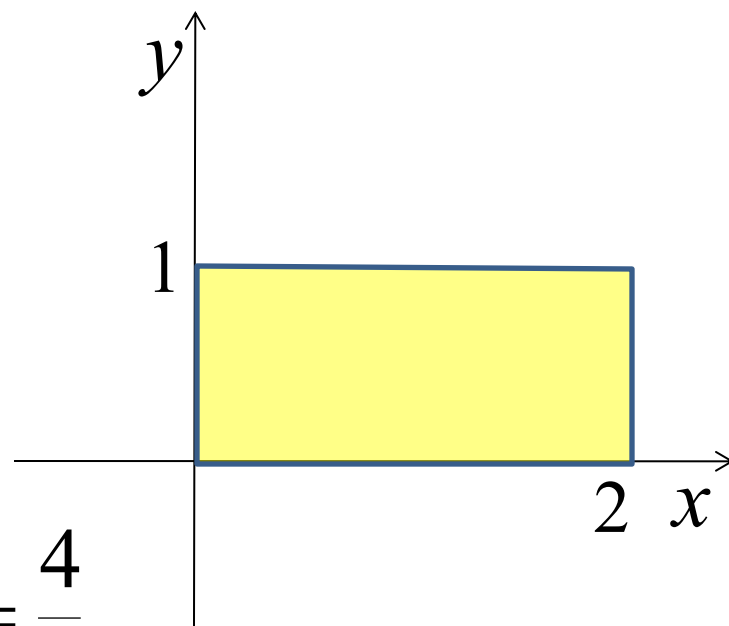
求 $E(X)$, $E(Y)$, $E(X+Y)$, $E(XY)$, $E(Y/X)$

解

$$E(X) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y)dx dy$$

$$= \int_0^2 \int_0^1 x \cdot \frac{1}{4}x(1+3y^2)dy dx$$

$$= \int_0^2 x \cdot \frac{1}{4}x dx \int_0^1 (1+3y^2)dy = \frac{4}{3}$$



$$\begin{aligned}
 E(Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y) dx dy \\
 &= \int_0^2 \int_0^1 y \cdot \frac{1}{4} x (1 + 3y^2) dy dx \\
 &= \int_0^2 \frac{1}{4} x dx \int_0^1 y (1 + 3y^2) dy = \frac{5}{8}
 \end{aligned}$$

$$\begin{aligned}
 E(X + Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) f(x, y) dx dy \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dx dy + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y) dx dy \\
 &= E(X) + E(Y) \quad \text{—— 数学期望的性质} \\
 &= \frac{4}{3} + \frac{5}{8} = \frac{47}{24}
 \end{aligned}$$

$$\begin{aligned}
E(XY) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (xy) f(x, y) dx dy \\
&= \int_0^2 \int_0^1 xy \cdot \frac{1}{4} x(1+3y^2) dy dx \\
&= \int_0^2 x \cdot \frac{1}{2} x dx \cdot \int_0^1 y \cdot \frac{1}{2} (1+3y^2) dy \\
&= \frac{4}{3} \cdot \frac{5}{8} = \frac{5}{6}
\end{aligned}$$

$$= E(X) \cdot E(Y) \quad \text{—— 数学期望的性质}$$

注意：X, Y 相互独立

$$f(x, y) = \begin{cases} \frac{1}{4}x(1+3y^2), & 0 < x < 2, 0 < y < 1, \\ 0, & \text{其它} \end{cases}$$

$$E\left(\frac{Y}{X}\right) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(\frac{y}{x}\right) f(x, y) dx dy$$

$$= \int_0^2 \int_0^1 \left(\frac{y}{x}\right) \cdot \frac{1}{4} x(1+3y^2) dy dx$$

$$= \int_0^2 \frac{1}{2} dx \int_0^1 y \cdot \frac{1}{2} (1+3y^2) dy$$

$$= \frac{5}{8}$$

