

1 随机变量 X 的数学期望、方差

(1) X 为标量

对随机变量 $X \in \mathbb{R}$ ，分两种情况讨论：

(1.1) X 为离散型随机变量

令 X 的取值为 $\{x_1, x_2, \dots, x_n\}$ ，这些取值点上对应的概率为 $\{p_1, p_2, \dots, p_n\}$ 。

X 的数学期望为：

$$\begin{aligned} E(X) &= x_1 p_1 + x_2 p_2 + \dots + x_n p_n \\ &= \sum_{i=1}^n x_i p_i \end{aligned} \quad (1)$$

X 的方差为：

$$\begin{aligned} \text{Var}(X) &= E\{(X - E(X))^2\} \\ &= (x_1 - E(X))^2 p_1 + (x_2 - E(X))^2 p_2 + \dots + (x_n - E(X))^2 p_n \\ &= E(X^2) - E^2(X) \end{aligned} \quad (2)$$

(1.2) X 为连续型随机变量

令 X 的概率密度函数为 $f(x)$ ， X 的定义域为 $[a, b]$ 。则 X 的数学期望为：

$$E(X) = \int_a^b f(x) dx \quad (3)$$

X 的方差为：

$$\begin{aligned} \text{Var}(X) &= \int_a^b (x - E(X))^2 dx \\ &= E(X^2) - E^2(X) \end{aligned} \quad (4)$$

(2) X 为向量

记 $X \in \mathbb{R}^n$ ， $X = [X_1, X_2, \dots, X_n]^T$ ，其中 $X_i, (i = 1, 2, \dots, n)$ ，均为随机变量。

X 的数学期望为各个分量的数学期望：

$$\begin{aligned}
 E(X) &= E \left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \right) \\
 &= \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{bmatrix}
 \end{aligned} \tag{5}$$

X 的方差为各个分量的方差：

$$\begin{aligned}
 Var(X) &= Var \left(\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \right) \\
 &= \begin{bmatrix} Var(X_1) \\ Var(X_2) \\ \vdots \\ Var(X_n) \end{bmatrix}
 \end{aligned} \tag{6}$$

(3) X 为矩阵

记随机变量 $X \in \mathbb{R}^{m \times n}$ 为：

$$X = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ X_{m,1} & X_{m,2} & \cdots & X_{m,n} \end{bmatrix} \tag{7}$$

X 的数学期望定义为：

$$\begin{aligned}
E(X) &= E \left(\begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m,1} & X_{m,2} & \cdots & X_{m,n} \end{bmatrix} \right) \\
&= \begin{bmatrix} E(X_{1,1}) & E(X_{1,2}) & \cdots & E(X_{1,n}) \\ E(X_{2,1}) & E(X_{2,2}) & \cdots & E(X_{2,n}) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{m,1}) & E(X_{m,2}) & \cdots & E(X_{m,n}) \end{bmatrix}
\end{aligned} \tag{8}$$

X 的方差定义为：

$$\begin{aligned}
Var(X) &= Var \left(\begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m,1} & X_{m,2} & \cdots & X_{m,n} \end{bmatrix} \right) \\
&= \begin{bmatrix} Var(X_{1,1}) & Var(X_{1,2}) & \cdots & Var(X_{1,n}) \\ Var(X_{2,1}) & Var(X_{2,2}) & \cdots & Var(X_{2,n}) \\ \vdots & \vdots & \ddots & \vdots \\ Var(X_{m,1}) & Var(X_{m,2}) & \cdots & Var(X_{m,n}) \end{bmatrix}
\end{aligned} \tag{9}$$

随机变量的协方差

1. X, Y 为标量

- 协方差的定义：(维基百科关于协方差的定义)

In probability theory and statistics, covariance is a measure of the joint variability of two random variables. If the greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the lesser values (that is, the variables tend to show similar behavior), the covariance is positive. In the opposite case, when the greater values of one variable mainly correspond to the lesser values of the other, (that is, the variables tend to show opposite behavior), the covariance is negative. The sign of the covariance therefore shows the tendency in the linear relationship between the variables. The magnitude of the covariance is not easy to interpret because it is not normalized and hence depends on the magnitudes of the variables. The normalized version of the covariance, the correlation coefficient, however, shows by its magnitude the strength of the linear relation.

- 协方差的前提要求

X 和 Y 之间存在联合概率分布 $f(x, y)$

- 协方差的作用

- 度量两个随机变量 X 和 Y 之间的联合变异性
- 协方差的符号
 - 为正： X 与 Y 正相关
 - 为负： X 与 Y 负相关
- 协方差的绝对值：没有实际意义
- 相关系数：是协方差的归一化形式，表明 X 和 Y 之间的线性相关性

- 协方差公式

$$\begin{aligned} Cov(X, Y) &= E((X - E(X))(Y - E(Y))) \\ &= E(XY) - E(X)E(Y) \end{aligned} \quad (10)$$

- 当 $f(x, y)$ 为离散函数时

- X 的概率质量函数为：

$$f_X(x_i) = \sum_j f(x_i, y_j)$$

- Y 的概率质量函数为：

$$f_Y(y_i) = \sum_j f(x_j, y_i)$$

- X, Y 的协方差为

$$Cov(X, Y) = \sum_{i,j} x_i y_j f(x_i, y_j) - \sum_i x_i \left(\sum_j f(x_i, y_j) \right)$$

- 当 $f(x, y)$ 为连续函数时

- X 的概率密度函数为：

$$f_X(x) = \int f(x, y) dy$$

- Y 的概率密度函数为：

$$f_Y(y) = \int f(x, y) dx$$

- X, Y 的协方差为

$$Cov(X, Y) = \iint f(x, y) dx dy -$$

2. X, Y 为向量

3. X, Y 为矩阵