随机变量及其分布

王友乐

电话：18980082143

邮箱：**Eleconnection@126.com**

2015 年 5 月 7 日

随机变量的数字特征

1. 数学期望
   - 数学期望的定义及计算
   - 随机变量函数的数学期望
   - 数学期望的性质
2. 方差
   - 方差的定义及计算
   - 方差的性质
   - 变异系数、矩及中心矩
3. 协方差系数和相关系数
   - 协方差
   - 相关系数

数学期望

定义

设离散型随机变量 X 的概率分布为

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

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛, 则称这个级数为随机变量 X 的**数学期望**, 简称为**期望**, 记为

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

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 不绝对收敛, 称 X 的数学期望不存在.

数学期望

定义

设连续型随机变量 X 的密度为 $f(x)$, 若反常积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛, 则称这个积分为 X 的数学期望, 记为

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx. \tag{2}$$

$X \sim B(n,p)$, 那么 $E(X) = np$.

$X \sim P(\lambda)$, 那么 $E(X) = \lambda$.

$X \sim e(\lambda)$, 那么 $E(X) = \frac{1}{\lambda}$.

定理

设 X 为随机变量,$y=g(x)$ 是 x 的 (分段) 连续函数或单调函数, 则对 $Y=g(X)$,

(1) 若 X 是离散型的, 其分布律为 $p_k=P(X=x_k)$,$k=1,2,\cdots$, 且级数 $\sum_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛, 则有

$$E(Y)=E[g(X)]=\sum_{k=1}^{\infty} g(x_k)p_k.$$

(2) 若 X 为连续型的, 其密度函数为 $f(x)$, 且反常积分 $\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛, 则有

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

数学期望

定理

设 (X,Y) 为二维随机变量,$z=g(x,y)$ 是 (x,y) 的 *(*分区域*)* 连续函数, 则对 $Z=g(X,Y)$,

(1) 若 (X,Y) 为离散型, 其二维概率分布为

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

$\sum_i\sum_j g(x_i,y_j)p_{ij}$ 绝对收敛, 则有

$$E(Z)=E[g(X,Y)]=\sum_i\sum_j g(x_i,y_j)p_{ij};$$

(2) 若 (X,Y) 为连续型, 其二维密度函数为 $f(x,y)$, 且反常积分

$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)dxdy$ 绝对收敛, 则有

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

数学期望的性质

(1) C 是常数, 则 $E(C)=C$.

(2) $E(CX)=CE(X)$.

(3) $E(X+Y)=E(X)+E(Y)$.

(4) 设 $X_1,X_2,\cdots,X_n$ 是 n 个随机变量,$C_1,C_2,\cdots,C_n,b$ 是 $n+1$ 个常量, 则

$$E(\sum_{i=1}^{n}C_iX_i+b)=\sum_{i=1}^{n}C_iE(X_i)+b.$$

(5) 若 X 与 Y 相互独立, 则 $E(XY)=E(X)E(Y)$. 一般地, 若 $X_1,X_2,\cdots,X_n$ 相互独立, 则

$$E(\Pi_{i=1}^{n}X_i)=\Pi_{i=1}^{n}E(X_i).$$

方差

定义

若期望 $E[X-E(X)]^2$ 存在, 称为 X 的**方差**, 记为

$$D(X)=E[X-E(X)]^2. \tag{3}$$

而 $\sqrt{D(X)}$ 称为 X 的**均方差或标准差**.

计算方差的公式

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

$X\sim B(n,p)$, 那么 $D(X)=np(1-p)$.

$X\sim P(\lambda)$, 那么 $D(X)=\lambda$.

$X\sim e(\lambda)$, 那么 $D(X)=\frac{1}{\lambda^2}$.

方差

方差的性质

(1) C 为常量, 则

$$D(C) = 0. \tag{4}$$

(2) a, b 为常量, 且 $a \neq 0$, 则

$$D(aX + b) = a^2 E(X). \tag{5}$$

(3) 若 X 与 Y 独立, 则

$$D(X \pm Y) = D(X) + D(Y). \tag{6}$$

方差的性质

(4) 随机变量 $X_1, X_2, \cdots, X_n$ 相互独立,$C_1, C_2, \cdots, C_n$ 是 n 个常数, 则

$$D(\sum_{i=1}^{n} C_i X_i) = \sum_{i=1}^{n} C_i^2 D(X_i). \tag{7}$$

(5) $D(X) = 0$ 其充要条件为存在常数 C, 使得 $P(X = C) = 1$, 且 $C = E(X)$.

定义

若随机变量 X 的期望, 方差均存在, 且 $E(X) \neq 0$, 称

$$C_v = \frac{\sqrt{D(X)}}{|E(X)|}. \tag{8}$$

为 X 的变异系数.

方差

定义

若随机变量 X 对非负整数 k 有下列期望存在,

$$m_k = E(X^k), \tag{9}$$

称 m_k 为 X 的 k 阶**原点矩**;

$$\mu_k = E[X - E(X)]^k, \tag{10}$$

称 μ_k 为 X 的 k 阶**中心矩**.

显然,$E(X) = m_1, D(X) = \mu_2$.

协方差和相关系数

定义

对二维随机变量 (X,Y), 若 $E[[X-E(X)][Y-E(Y)]]$ 存在, 称为 X 与 Y 的**协方差**, 记为

$$Cov(X,Y)=E[[X-E(X)][Y-E(Y)]]. \tag{11}$$

特别地, 有 $Cov(X,X)=D(X)$.

协方差和相关系数

协方差的性质

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

(2)$Cov(X,a)=0$;

(3)$Cov(aX,bY)=abCov(X,Y)$;

(4)$Cov(X+Y,Z)=Cov(X,Y)+Cov(Y,Z)$;

(5)$Cov(\sum_{i=1}^{n}a_iX_i,\sum_{j=1}^{m}b_jY_j)=\sum_{i=1}^{n}\sum_{j=1}^{m}a_ib_jCov(X_i,Y_j)$;

(6)$D(X\pm Y)=D(X)+D(Y)\pm 2Cov(X,Y)$;

(7) 若 X 与 Y 相互独立, 则 $Cov(X,Y)=0$.

对二维随机变量 (X, Y), 称向量 $(E(X), E(Y))$ 为 (X, Y) 的数学期望或**均值向量**, 称矩阵

$$V = \begin{pmatrix} D(X) & Cov(X, Y) \\ Cov(Y, X) & D(Y) \end{pmatrix}$$

为 (X, Y) 的**协方差矩阵**.

协方差和相关系数

一般, 对 n 维随机变量 $(X_1, X_2, \cdots, X_n)$, 称向量 $(E(X_1), E(X_2), \cdots, E(X_n))$ 为 $(X_1, X_2, \cdots, X_n)$ 的**均值向量**, 记

$$\sigma_{ij} = Cov(X_i, Y_j), \quad i, j = 1, 2, \cdots, n,$$

称矩阵

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

为 $(X_1, X_2, \cdots, X_n)$ 的**协方差矩阵**. 显然 V 是一个对称阵.

定义

设随机变量 X 与 Y 的期望, 方差都存在, 且 $D(X)>0, D(Y)>0$, 称

$$X^* = \frac{X-E(X)}{\sqrt{D(X)}}, \quad Y^* = \frac{Y-E(Y)}{\sqrt{D(Y)}}. \tag{12}$$

分别为 X 与 Y 的**标准化随机变量**; 称

$$Cov(X^*, Y^*) = R(X, Y) \tag{13}$$

为 X 与 Y 的**相关系数**.

协方差和相关系数

定理

$$(1)E(X^*)=0; D(X^*)=D(Y^*)=1;$$
$$(2)R(X,Y)=\frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}.$$

定理

$(1)R(X,Y)=R(Y,X);$

$(2)|R(X,Y)|\leq 1;$

$(3)|R(X,Y)|=1$ 的充要条件为; 存在常数 a,b 且 $a\neq 0$, 使得 $P(aX+b=Y)=1.$

$aX+b=Y$, 若 $a>0$, 可得 $R(X,Y)=1$, 这时称 X 与 Y 完全正相关; 若 $a<0$, 可得 $R(X,Y)=-1$, 这时称 X 与 Y 完全负相关.

一般地,$R(X,Y)>0$, 称 X 与 Y 正线性相关;$R(X,Y)<0$, 称 X 与 Y 负线性相关. 当 $R(X,Y)=0$ 时, 称 X 与 Y **不相关**.

当 X 与 Y 独立时, 则有 X 与 Y 不相关. 但当 X 与 Y 不相关时, 不能得出 X 与 Y 不相关.