# 随机变量的数字特征

## 数学期望与中位数

### 数学期望的定义

**定义 1.1** 设随机变量 $X$ 只取有限个可能值 $a_1, \cdots, a_m$. 其概率分布为 $P(X = a_i) = p_i, i = 1, \cdots, m$. 则 $X$ 的数学期望,记为 $E(X)^*$ 或 $EX$,定义为

$$E(X) = a_1 p_1 + a_2 p_2 + \cdots a_m p_m$$

**定义 1.2** 如果

$$\sum_{i=1}^{\infty} |a_i| p_i < \infty \tag{1.3}$$

则称(1.2)式右边的级数之和为 $X$ 的数学期望.

$$E(X) = \sum_{i=1}^{\infty} a_i p_i$$

**定义 1.3** 设 $X$ 有概率密度函数 $f(x)$. 如果

$$\int_{-\infty}^{\infty} |x| f(x) \mathrm{d}x < \infty \tag{1.4}$$

则称

$$E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x \tag{1.5}$$

为 $X$ 的数学期望.

**数学期望的性质**

**定理 1.1** 若干个随机变量之和的期望,等于各变量的期望之和,即

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n) \tag{1.12}$$

当然,这里要假定各变量 $X_i$ 的期望都存在.

**定理 1.2** 若干个独立随机变量之积的期望,等于各变量的期望之积:

$$E(X_1 X_2 \cdots X_n) = E(X_1) E(X_2) \cdots E(X_n)$$

当然,这里也要假定各变量 $X_i$ 的期望都存在.

(想想为什么后者要独立)

只有当$X_1, X_2$独立时,($X_1, X_2$)的联合密度才等于各分量的密度之积

(注意期望的概率密度推导)

3.设二维随机变量$(X,Y)$的概率密度函数为$f(x,y)$。

$$E(X+Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\,dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\,dxdy+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\,dxdy$$

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

4.若X和Y相互独立，其边缘概率密度函数为$f_X(x), f_Y(y)$，有$f(x,y)=f_X(x)f_Y(y)$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\,dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf_X(x)f_Y(y)\,dxdy$$

$$=\left[\int_{-\infty}^{\infty}xf_X(x)\,dx\right]\left[\int_{-\infty}^{\infty}yf_Y(y)\,dy\right]=E(X)E(Y)$$

## 条件数学期望

如果知道了$(X,Y)$的联合密度，则$E(Y|x)$的定义就可以具体化为：先定出在给定$X=x$之下，$Y$的条件密度函数$f(y|x)$，然后按定义1.3算出

$$E(Y|x)=\int_{-\infty}^{\infty}yf(y|x)\mathrm{d}y \tag{1.25}$$

公式(1.27)可给以另一种写法，记$g(x)=E(Y|x)$，它是$x$的函数，则(1.27)成为

$$E(Y)=\int_{-\infty}^{\infty}g(x)f_1(x)\mathrm{d}x \tag{1.28}$$

但据(1.18)，上式右边就是$E(g(X))$. 从$g(x)$的定义，$g(X)$是$E(Y|x)|_{x=X}$，可简写为$E(Y|X)$. 于是由(1.28)得

$$E(Y)=E[E(Y|X)] \tag{1.29}$$

公式(1.29)虽可算是概率论中一个比较高深的公式，它的实际含义其实很简单：它可理解为一个"分两步走"去计算期望的方法. 因为在不少情况下，迳直计算 $E(Y)$ 较难，而在限定某变量 $X$ 之值后，计算条件期望 $E(Y|x)$ 则较容易. 因此我们分两步走：第一步算出 $E(Y|x)$，再借助 $X$ 的概率分布，通过 $E(Y|x)$ 算出 $E(Y)$. 更直观一些，你可以把求 $E(Y)$ 看成为在一个很大的范围求平均. 限定 $X$ 之值从这个很大的范围内界定了一个较小的部分. 先对这较小的部分求平均，然后再对后者求平均. 比如要求全

推广：

有特殊的假设，故可适用于更为一般的情形. 例如，$X$ 不必是一维的，如果 $X$ 为 $n$ 维随机向量 $(X_1,\cdots,X_n)$，有概率密度 $f(x_1,\cdots,x_n)$，则公式(1.29)有形式

$$E(Y)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}E(Y|x_1,\cdots,x_n)f(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n \tag{1.30}$$

这里 $E(Y|x_1,\cdots,x_n)$ 就是在 $X_1=x_1,\cdots,X_n=x_n$ 的条件下，$Y$ 的条件期望，又 $X,Y$ 都可以是离散型的. 例如，设 $X$ 为一维离散型变量，有分布

$$P(X=a_i)=p_i,\quad i=1,2,\cdots$$

则公式(1.29)有形式

$$E(Y)=\sum_{i=1}^{\infty}p_iE(Y|a_i) \tag{1.31}$$

**中位数**

**定义 1.4** 设连续型随机变量 $X$ 的分布函数为 $F(x)$，则满足条件

$$P(X \leqslant m) = F(m) = 1/2 \tag{1.32}$$

的数 $m$ 称为 $X$ 或分布 $F$ 的中位数.

# 方差与矩

## 方差

**定义 2.1** 设 $X$ 为随机变量，分布为 $F$，则

$$\mathrm{Var}(X) = E(X - EX)^2 \tag{2.1}$$

称为 $X$（或分布 $F$）的方差*，其平方根 $\sqrt{\mathrm{Var}(X)}$（取正值）称为 $X$（或分布 $F$）的标准差.

暂记 $EX = a$. 由于 $(X-a)^2 = X^2 - 2aX + a^2$，按定理 1.1 得

$$\begin{aligned}\mathrm{Var}(X) &= E(X^2) - 2aE(X) + a^2 \\ &= E(X^2) - (EX)^2\end{aligned} \tag{2.2}$$

方差的这个形式在计算上往往较为方便.

**定理 2.1** 1° 常数的方差为 0. 2° 若 $C$ 为常数，则 $\mathrm{Var}(X+C) = \mathrm{Var}(X)$. 3° 若 $C$ 为常数，则 $\mathrm{Var}(CX) = C^2\mathrm{Var}(X)$.

**定理 2.2** 独立随机变量之和的方差，等于各变量的方差之和：

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) \tag{2.3}$$

## 矩

**定义 2.2** 设 $X$ 为随机变量，$c$ 为常数，$k$ 为正整数. 则量 $E[(X-c)^k]$称为$X$关于$c$点的$k$阶矩.

比较重要的有两个情况：

1. $c=0$. 这时 $a_k=E(X^k)$称为 $X$ 的$k$ 阶原点矩.

2. $c=E(X)$. 这时 $\mu_k=E[(X-EX)^k]$称为$X$ 的$k$ 阶中心矩.

一阶原点矩就是期望. 一阶中心矩 $\mu_1=0$，二阶中心矩 $\mu_2$ 就是 $X$ 的方差 $\mathrm{Var}(X)$. 在统计学上，高于4阶的矩极少使用. 三、四阶矩有些应用，但也不很多.

$$\beta_1=\mu_3/\mu_2^{3/2}$$

称为 $X$ 或其分布的“偏度系数”.

$$\beta_2=\mu_4/\mu_2^2$$

它称为 $X$ 或其分布的“峰度系数”.

## 协方差与相关系数

### 协方差

**定义 3.1** 称 $E[(X-m_1)(Y-m_2)]$为 $X,Y$ 的协方差，并记为 $\mathrm{Cov}(X,Y)^*$.

**定理 3.1** 1° 若 $X,Y$ 独立，则 $\mathrm{Cov}(X,Y)=0$.

2° $[\mathrm{Cov}(X,Y)]^2\leqslant\sigma_1^2\sigma_2^2$. 等号当且仅当 $X,Y$ 之间有严格线性关系(即存在常数 $a,b$ 使 $Y=a+bX$)时成立.

### 相关系数

**定义 3.2** 称 $\mathrm{Cov}(X,Y)/(\sigma_1\sigma_2)$ 为 $X,Y$ 的相关系数,并记为 * $\mathrm{Corr}(X,Y)$.

**定理 3.2** 1° 若 $X,Y$ 独立,则 $\mathrm{Corr}(X,Y)=0$. 2° $-1\leqslant\mathrm{Corr}(X,Y)\leqslant1$,或 $|\mathrm{Corr}(X,Y)|\leqslant1$,等号当且仅当 $X$ 和 $Y$ 有严格线性关系时达到.

## 大数定理和中心极限定理

### 大数定理

**定理 4.1** 设 $X_1,X_2,\cdots,X_n,\cdots$ 是独立同分布的随机变量,记它们的公共均值为 $a$. 又设它们的方差存在并记为 $\sigma^2$. 则对任意给定的 $\varepsilon>0$ 有

$$\lim_{n\to\infty}P(|\bar{X}_n-a|\geqslant\varepsilon)=0\quad(\bar{X}_n\text{ 见}(4.1))\qquad(4.2)$$

**马尔科夫不等式** 若 $Y$ 为只取非负值的随机变量,则对任给常数 $\varepsilon>0$ 有

$$P(Y\geqslant\varepsilon)\leqslant E(Y)/\varepsilon\qquad(4.3)$$

**契比雪夫不等式**. 若 $\mathrm{Var}(Y)$存在,则

$$P(|Y-EY|\geqslant\varepsilon)\leqslant\mathrm{Var}(Y)/\varepsilon^2\qquad(4.4)$$

### 中心极限定理

**定理 4.2** 设 $X_1, X_2, \cdots, X_n, \cdots$ 为独立同分布的随机变量，$E(X_i)=a$，$\mathrm{Var}(X_i)=\sigma^2$，$0<\sigma^2<\infty$. 则对任何实数 $x$，有

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{n}\sigma}(X_1+\cdots+X_n-na)\leqslant x\right)=\Phi(x) \quad (4.7)$$

**定理 4.3** 设 $X_1, X_2, \cdots, X_n, \cdots$ 独立同分布，$X_i$ 分布是

$$P(X_i=1)=p, P(X_i=0)=1-p, 0<p<1$$

则对任何实数 $x$，有

$$\lim_{n\to\infty} P\left(\frac{1}{\sqrt{np(1-p)}}(X_1+\cdots+X_n-np)\leqslant x\right)=\Phi(x) \quad (4.9)$$