第五周　随机变量函数的分布及随机变量的数字特征

5.2 随机变量的数学期望

例 5.2.1

项目 1：投资 10 万元

60%可能回收 10 万元保本；40%可能回收 15 万元，盈利 5 万元

平均收益为 $0\times\frac{3}{5}+5\times\frac{2}{5}=2$ 万元

项目 2：投资 10 万元

60%可能回收 0 万元，亏损 10 万元；40%可能回收 30 万元，盈利 20 万元

平均收益为 $-10\times\frac{3}{5}+20\times\frac{2}{5}=2$ 万元

两项投资预期的平均收益都是 2 万元，若要从中做出决策，如何决定？

可能不少人会选择第一种最差也能保本的项目，但也有人更愿意尝试第二个看似风险更大的项目。所谓的风险大就是不同可能性对应的收益差别大，收益的波动大。一般而言，人们会依据平均收益和风险程度两个方面进行判断。

上述例子反映了随机或不确定情况下，平均值是人们常用的参考量，分散程度或波动程度的不同又会带来差异。在概率论中，描述随机变量这两方面特征的标准概念是期望和方差，下面分别给出它们在数学上的定义。

****

随机变量 X 的数学期望：（加权）平均值

$$E(X)=\begin{cases}\sum_{i=1}^{\infty}x_i\cdot P(X=x_i), & X\text{为离散随机变量}\\ \int_{-\infty}^{+\infty}x\cdot f(x)dx, & X\text{为连续随机变量}\end{cases}$$

备注：离散型随机变量需要满足 $\sum_{i=1}^{\infty}|x_i|P(X=x_i)<\infty$，

连续型随机变量需要满足$\int_{-\infty}^{+\infty}|x|f(x)dx<\infty$ 才称其数学期望存在。

****

例 5.2.2 投掷一颗均匀的色子，求掷出点数的数学期望。

解：设投出的点数为随机变量X，则X服从下面分布

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix},$$

$$E(X)=\sum_{i=1}^{6}x_i\cdot P(X=x_i)=\sum_{i=1}^{6}i\cdot\frac{1}{6}=\frac{1}{6}\frac{6\times7}{2}=3.5$$

****

随机变量X函数的数学期望

$$E(g(X))=\begin{cases}\sum_{k=1}^{\infty}g(x_k)\cdot P(X=x_k), & X\text{为离散随机变量}\\ \int_{-\infty}^{+\infty}g(x)f(x)dx, & X\text{为连续随机变量}\end{cases}$$

例 5.2.3 已知随机变量$X\sim\begin{pmatrix}-2 & -1 & 0 & 1 & 2\\ 0.2 & 0.1 & 0.1 & 0.3 & 0.3\end{pmatrix}$，求$Y=X^2+X$的期望。

$$E(Y)=\sum_{k=1}^{5}(x_k^2+x_k)\cdot P(X=x_k)$$

$$=\left[(-2)^2+(-2)\right]\times0.2+\left[(-1)^2+(-1)\right]\times0.1+(0^2+0)\times0.1+(1^2+1)\times0.1+(2^2+2)\times0.1$$

$$=1.2$$

****

数学期望的几个基本性质

$E(c)=c$ （c为常数，常值分布）　　　$E(cX)=cE(X)$

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

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

$$E(g_1(X)+g_2(X))=E(g_1(X))+E(g_2(X))$$

前两个结论很容易验证，后面几个关于随机变量求和的期望的结论，证明有些麻烦，我们暂且先不加证明地引入这几个结论。

****

例 5.2.4（匹配问题） n 封写给不同人的信随机放入 n 个写好收信人姓名的信封，求平均有几封信会装对信封？

解：将 n 封不同的信分别编号 $1,2,\cdots,n$， n 个对应的信封同样编号 $1,2,\cdots,n$，

定义随机变量 $X_k=\begin{cases}1, & \text{编号为}k\text{的信件装入了编号为}k\text{的信封}\\ 0, & \text{其他}\end{cases}$， $k=1,2,\cdots,n$，

$P(X_k=1)=\dfrac{1}{n}$， $P(X_k=0)=1-\dfrac{1}{n}$， 所以 $E(X_k)=1\cdot\dfrac{1}{n}+0\cdot\left(1-\dfrac{1}{n}\right)=\dfrac{1}{n}$。

装对了信封的信件总数为： $X_1+X_2+\cdots+X_n$，

所以，装对信封的信件的平均数为

$$E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n)=n\cdot\frac{1}{n}=1。$$

****