1. 期望值

在研究期望值之前，我们看看数据：(3,3,3,4,5)。以前，我们计算均值的方法是：

$$\frac{3+3+3+4+5}{5}=3.6$$

我们可以换一种思路：

$$\begin{aligned}&\frac{3\times 3+1\times 4+1\times 5}{5}\\&=0.6\times 3+0.2\times 4+0.2\times 5\\&=60\%\times 3+20\%\times 4+20\%\times 5\end{aligned}$$

也就是说，我们运算转换为了将数据的概率与数据相乘再求和的过程。这种通过数据出现的概率来计算均值的方式就是期望值的运算方式。

随机变量的期望值E(X)其实也就是总体的均值，只是此时总体是无穷的,所以无法用全部求和然后除以数目的方式来求均值。但是，如果我们知道值的频率，计算均值就成为可能，这就是计算随机变量期望值的方式。

如何知道值的频率呢？那就是参照概率分布。我们用频率作为权重，然后计算出所有结果的加权平均值。

例如，投篮命中率为50%，投掷6次，统计投中的命中率：

命中数(K)	概率(P)	N个选择K		P(X=K)
0	0.015625	1		0.015625
1	0.015625	6		0.09375
2	0.015625	15		0.234375
3	0.015625	20		0.3125
4	0.015625	15		0.234375
5	0.015625	6		0.09375
6	0.015625	1		0.015625

期望值为：

$$\begin{aligned}E(X)&=0\times 1.563\%+1\times 0.938\%+2\times 23.438\%+3\times 31.25\%+4\times 23.438\%+5\times 0.938\%+6\times 1.563\%\\&=3\end{aligned}$$

也就是说投6次，投中3次的概率最大。

> 期望值利用频率进行加权平均

2. 期望值的一般公式

$$\begin{aligned}&X=\#\text{ of sucesses with probablility p after n trails}\\&E(x)=n\cdot p\end{aligned}$$

例如：X = # of baskets i make after 10 shots with p = 40%. 那么E(x) = 10×40% =4。

接下来证明一下这个公式。我们知道二项分布的概率公式为：

$$P(X=k)=\binom{n}{k}p^k\cdot(1-p)^{n-k}$$

则：

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

提出一个np来

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

我们希望期望值为np，也就是说让后面的求和式等于1即可。我们使用换元法：

$$\begin{aligned}&a=k-1,b=n-1\\&a+1=k,b+1=n\\&n-k=b-a\end{aligned}$$

所以：

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

np后面的求和公式是对二项分布的每一项求和，让后将所有的概率加到一起，所以就是1。最终推导出：

$$E(X)=np$$