第四章　随机变量的数字特征

§4.1　数学期望

知道了随机变量 X 的概率分布以后，X 的全部概率特征就都知道了.但在实际问题中，概率分布较难确定，而它的某些数字特征却比较容易估算出来；并且在不少问题中，只需知道它的某些数字特征就够了，而不必花大力气来细致的了解它的详细的概率特性.例如在水稻产区，水稻的产量是一个随机变量，如果要比较两个水稻产区的生产水平，通常只要比较两个产区的平均亩产量就可以了.平均亩产量越高，就意味着这个地区生产水平越高，如果不比较它们的平均亩产量，而只看它们的分布，虽然“全面”，却使人不得要领，既难以掌握，又难以迅速作出判断.这样的例子还可以举出很多.因此，我们经常需要研究随机变量的某些重要的特征，这些特征虽然不能全面描述随机变量，但对我们解决一些具体问题很有帮助.所以，对数字特征的研究就显得非常重要了.本节先讨论最常用的数字特征—数学期望. .

一、离散型随机变量的数学期望

设随机变量 X 的概率分布为

ξ	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

我们希望能找到这样一个数值，它能体现 X 取值的平均大小，就如同一组数字的平均数那样. 但是，对于一个随机变量 X 而言，X 的所有可能取值的和再除以总个数那种方式的平均数，并不能真正起到平均的作用.

例如，若 X 的分布列为

X	100	200
P	0.01	0.99

作为 X 的所有可能值的平均数是 150，但是凭直觉，150 并不能真正体现 X 的取值的平均，这是错误的对 X 的可能值 100 和 200 一视同仁的结果.事实上，从其分布可以看出，X 取 200 的机会远比取 100 的机会多得多.因此要真正体现 X 的取值的平均，不能只由它取什么值来决定，还必须考虑到它取那些值的相应的概率.

定义4.1　设随机变量 X 的分布列为　$P(X=x_k)=p_{k,}$　$k=1,2,\cdots$,若级数 $\sum\limits_{k=1}^{\infty}x_k p_k$ 绝对收敛，则称级数 $\sum\limits_{k=1}^{\infty}x_k p_k$ 为 X 的数学期望（均值），记为 $E(X)$.

即
$$E(X)=\sum_{k=1}^{\infty}x_k p_k \qquad (4.1)$$

在数学期望的定义中要求级数 $\sum\limits_{k=1}^{\infty}x_k p_k$ 绝对收敛，这是因为随机变量取值是随机的，并

不一定是按 $x_1, x_2, \cdots, x_k, \cdots$ 的顺序取值，而数学期望存在，则要求是惟一的.这就要求级数 $\sum_{k=1}^{\infty} x_k p_k$ 的项改变顺序而其和不变，要满足这一要求，级数必须绝对收敛.

不难看出，只取有限个值的随机变量的数学期望总是存在的..

对于上例，按照定义易得 $E(X)=100\times 0.01+200\times 0.99=199$.

显然，$E(X)$ 是一个实数，它在形式上是 X 的可能值关于概率的加权平均，实质上体现了随机变量 X 取值的真正平均

由于随机变量的数学期望由随机变量的概率分布惟一决定,所以有相同分布的随机变量必有相同的数学期望.

[例 1] 飞机在一次飞行后必须进行检修的概率是0.4，在以后的两次飞行中，每一次飞行后需检修的概率各增加0.1，假设每次飞行后是否需要检修与前次飞行后是否检修无关，求三次飞行后检修次数的数学期望.

解：设三次飞行后检修次数为随机变量 X，令 A_i 表示飞机在第 i 次飞行后需检修，则

$$P(A_1)=0.4,\ P(A_2)=0.5,\ P(A_3)=0.6$$

$$\begin{aligned}P\{X=1\}&=P(A_1\overline{A}_2\overline{A}_3)+P(\overline{A}_1A_2\overline{A}_3)+P(\overline{A}_1\overline{A}_2A_3)\\&=0.4\times0.5\times0.4+0.6\times0.5\times0.4+0.6\times0.5\times0.6\\&=0.38\end{aligned}$$

$$\begin{aligned}P\{X=2\}&=P(A_1A_2\overline{A}_3)+P(\overline{A}_1A_2A_3)+P(A_1\overline{A}_2A_3)\\&=0.4\times0.5\times0.4+0.6\times0.5\times0.6+0.4\times0.5\times0.6\\&=0.38\end{aligned}$$

$$\begin{aligned}P\{X=3\}&=P(A_1A_2A_3)\\&=0.4\times0.5\times0.6=0.12\end{aligned}$$

故 $E(X)=P\{X=1\}+2P\{X=2\}+3P\{X=3\}=1.5$

[例 2] 某种体育彩票，有顺序的 7 个数字组成一个号码，称为一注.7 个数字中的每个数字都选自 0，1，2，…，9，可以重复.如果彩票一元一张，且全体不同的彩票中只有一个大奖，中大奖可获得奖金 300 万元，上税 20%，甲购买一注时，期望盈利多少?

解：用 X 表示甲购买一注时的收益，则

$$P\{X=24\times10^5\}=10^{-7}\qquad P\{X=-1\}=1-10^{-7}$$

所以，X 的期望盈利为

$E(X)=(-1)\times(1-10^{-7})+24\times10^5\times10^{-7}=-0.76$.于是每购买一注，甲期望获得 -0.76 元.也就是说，每买一注，平均损失 0.76 元.

二、连续型随机变量的数学期望.

定义 4.2 设 X 具有密度函数 $p(x)$,如果积分 $\int_{-\infty}^{+\infty} xp(x)dx$ 绝对收敛，则称该积分值为 X 的数学期望.

记为 $$E(X) = \int_{-\infty}^{+\infty} xp(x)dx \quad (4.2)$$

[例 3] 设随机变量 X 的密度函数为

$$p(x) = \begin{cases} x, & 0 < x \le 1 \\ 2-x, & 1 < x \le 2 \\ 0, & \text{其他} \end{cases},$$

求 X 的数学期望.

解： $E(x) = \int_{-\infty}^{+\infty} xp(x)dx = \int_0^1 x^2 dx + \int_1^2 x(2-x)dx = 1$

[例 4] 设 X 的密度函数为

$$p(x) = \begin{cases} ax + bx^2 & 0 \le x \le 1 \\ 0 & \text{其他} \end{cases},$$

且 $E(X) = \frac{3}{5}$，求 a,b 的值.(a=18/5,b= -12/5)

解： 由 $E(x) = \int_{-\infty}^{+\infty} xp(x)dx = \frac{3}{5}$ 得

$$\int_0^1 x(ax+bx^2)dx = \frac{3}{5}$$

即 $$\frac{1}{3}a + \frac{1}{4}b = \frac{3}{5} \quad (1)$$

又由 $p(x)$ 是密度函数，所以有

$$\int_{-\infty}^{+\infty} p(x)dx = \int_0^1 (ax+bx^2)dx = 1$$

亦即 $$\frac{1}{2}a + \frac{1}{3}b = 1 \quad （2）$$

联立（1），（2）解得 $a = \frac{18}{5}, b = -\frac{12}{5}$

[例 5] 设 $$p(x) = \frac{1}{\pi(1+x^2)}, \quad x \in R$$

则（1） $p(x)$ 是一个密度函数；（2）该分布的期望不存在.

解：（1）显然，$p(x)$ 非负，并且

$$\int_{-\infty}^{+\infty} p(x)dx = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{1}{1+x^2}dx = \frac{1}{\pi}\arctan x \Big|_{-\infty}^{+\infty} = 1$$

所以 $p(x)$ 是一个密度函数（Cauchy 分布）.

（2）由于

$$\int_{-\infty}^{+\infty}|x|p(x)dx=\frac{2}{\pi}\int_{0}^{+\infty}\frac{x}{1+x^2}dx=\infty$$

所以 Cauchy 分布的期望不存在.

本例告诉我们，并不是任意随机变量的数学期望都存在.

三、常见分布的数学期望

（1） 两点分布 $P\{X=1\}=p, P\{X=0\}=1-p=q$

$$E(X)=1\cdot p+0.q=p.$$

（2） 二项分布 $X\sim B(n,p)$

$$E(X)=\sum_{k=0}^{n}kp_k=\sum_{k=1}^{n}kC_n^k p^k q^{n-k}$$

$$=\sum_{k=1}^{n}\frac{k\cdot n!}{k!(n-k)!}p^k q^{n-k}$$

$$=\sum_{k=1}^{n}\frac{np(n-1)!}{(k-1)![(n-1)-(k-1)]!}p^{k-1}q^{(n-1)-(k-1)}$$

$$\underline{\underline{k'=k-1}}\ np\sum_{k'=0}^{n-1}\frac{(n-1)!}{k'![(n-1)-k']!}p^{k'}q^{(n-1)-k'}$$

$$=np(p+q)^{n-1}=np$$

（3） Poisson 分布 $X\sim P(\lambda),\lambda>0$

$$E(X)=\sum_{k=0}^{\infty}k\cdot\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}\lambda$$

$$=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda$$

（4）几何分布 设随机变量 X 服从几何分布，则其分布列为

$$P\{X=k\}=pq^{k-1},k=1,2,\cdots$$

$$E(X)=\sum_{k=1}^{\infty}kpq^{k-1}=p(\sum_{k=0}^{\infty}q^k)'$$

$$=p(\frac{1}{1-q})'=\frac{1}{p}$$

（5）均匀分布 若 $X\sim U[a,b]$， 则

$$E(X)=\int_{-\infty}^{+\infty}xp(x)dx=\int_{a}^{b}x\cdot\frac{1}{b-a}dx=\frac{a+b}{2}$$

（6）正态分布　若 $X\sim N(\mu,\sigma^2)$

$$\begin{aligned}E(X)&=\int_{-\infty}^{+\infty}xp(x)dx\\&=\int_{-\infty}^{+\infty}\mu p(x)dx+\int_{-\infty}^{+\infty}(x-\mu)p(x)dx\\&=\mu+\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\\&=\mu.\end{aligned}$$

（7）　指数分布　　X 的概率密度　$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}\quad \lambda>0$

$$E(X)=\int_{-\infty}^{+\infty}xp(x)dx=\int_{0}^{+\infty}x\lambda e^{-\lambda x}dx=\frac{1}{\lambda}$$

四、随机变量函数的数学期望

定理 4.1　若随机变量 X 的分布列为 $p_k=P\{X=x_k\},k=1,2,\cdots$，则 X 的某一函数 $g(X)$ 的数学期望为

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

对连续型随机变量，完全平行地有

定理 4.2　若随机变量的密度函数为 $p(x)$，　则 X 的某一函数 $g(X)$ 的数学期望为

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

对随机向量函数的数学期望，我们这里只给出连续型的情况，离散型随机向量函数的期望读者可以自行推出.

定理 4.3　若随机向量 (X,Y) 的联合密度函数为 $p(x,y)$，则 (X,Y) 的某一函数 $g(X,Y)$ 的数学期望为

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

[例 6] 设 X 的密度函数为

$$p(x)=\frac{1}{2}e^{-|x|},-\infty<x<+\infty$$（称为拉普拉斯分布）

求 $E(\min(|X|,1))$.

解：由定理 4.2 得

$$E(\min(|X|,1))=\int_{-\infty}^{+\infty}\min(|x|,1)\,p(x)dx=\int_{|x|<1}|x|\,p(x)dx+\int_{|x|\ge 1}p(x)dx$$

$$=2\int_0^1 x\frac{1}{2}e^{-x}dx+2\int_1^{+\infty}\frac{1}{2}e^{-x}dx=1-e^{-1}$$

[例 7] 设 (X,Y) 服从二维正态分布，联合密度函数为

$$p(x,y)=\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)},\quad -\infty<x,y<+\infty$$

令 $Z=\sqrt{X^2+Y^2}$（这时 Z 的分布称为瑞利分布），求 Z 的数学期望.

解：由定理 4.3

$$E(Z)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)p(x,y)dxdy=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}dxdy$$

$$=\int_0^{2\pi}\int_0^{+\infty}r\frac{1}{2\pi}e^{-\frac{1}{2}r^2}rd\theta dr=\int_0^{+\infty}r^2e^{-\frac{1}{2}r^2}dr=\sqrt{\frac{\pi}{2}}$$

五、数学期望的性质：

（1） $E(C)=C$；

（2） $E(CX)=CE(X)$；

（3） $E(X+Y)=E(X)+E(Y)$；

（4） 设 X，Y 相互独立，则 $E(XY)=E(X)E(Y)$.

性质（1），2）显然成立，我们只证明（3）和（4）.

证明：（3） $E(X+Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)p(x,y)dxdy$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)dxdy+\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yp(x,y)dxdy$$

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

（4） $E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyp(x,y)dxdy$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyp_X(x)p_Y(y)dxdy$$

$$=(\int_{-\infty}^{+\infty}xp_X(x)dx)(\int_{-\infty}^{+\infty}yp_Y(y)dy)$$

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

[例 8] 有 n 把外型相同钥匙，其中只有一把能打开门上的锁，现随机的取出一把试开（显然，试开不成即除去），求试开次数 X 的数学期望.

解法一： X 的取值为$1,2,\cdots,n$，显然，第k次打开，必然是前$k-1$次试开不成功，

即 $$P\{X=k\}=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdots\cdot\frac{n-k+1}{n-k+2}\cdot\frac{1}{n-k+1}=\frac{1}{n}$$

$$E(X)=\sum_{k=1}^{n}kP\{X=k\}=\frac{1}{n}\sum_{k=1}^{n}k=\frac{n+1}{2}.$$

解法二： 利用数学期望的性质求解.

设 $X_i=\begin{cases}0, & 第i次试开不成功\\ i, & 第i次试开成功\end{cases}$ 则X_i的概率分布为$\begin{pmatrix}0 & i\\ 1-\frac{1}{n} & \frac{1}{n}\end{pmatrix}$,

所以，$E(X_i)=0\cdot(1-\frac{1}{n})+i\cdot\frac{1}{n}=\frac{i}{n}$，而$X=\sum_{i=1}^{n}X_i$，故

$$E(X)=E\sum_{i=1}^{n}X_i=\sum_{i=1}^{n}E(X_i)=\frac{1}{n}\sum_{i=1}^{n}i=\frac{n+1}{2}$$

§4.2 方差

一、方差的定义

数学期望反映了随机变量取值集中的位置，但有的随机变量取值相对集中，有的相对分散，因此，我们还要引入另外一个反映这种集中或分散程度的数字特征，这就是方差.

先看一个实例.设两名射手的射击成绩如下

甲 $\begin{pmatrix}1 & 2 & 3\\ 0.25 & 0.5 & 0.25\end{pmatrix}$ 乙 $\begin{pmatrix}1 & 2 & 3\\ 0.05 & 0.9 & 0.05\end{pmatrix}$

其中 1，2，3 为射击得分，为了比较甲乙两射手的成绩，先求数学期望

$E(X_1)=2$ $E(X_2)=2$

由于数学期望相同，故两名射手的平均成绩一样，是否凭此可以说明二人的水平一样呢？从直观上看，甲的成绩较分散，乙的成绩相对要集中一些，因此，有必要用“方差”这种数字特征作进一步的研究.

分散与集中，是一模糊概念，应当有一个共同的“参照物”作为比较的量，我们不妨选择$E(X)$，用$X-E(X)$表述分散与集中，称为离差，但是，离差的平均值为 0，事实上，$E(X-EX)=0$，为了解决第二次平均中正、负离差相互抵消的问题，我们可选用$E|X-EX|$来表示，但绝对值在运算中很不方便，故选用$E[X-EX]^2$，这就是方差.

定义 4.3 设X为一随机变量，若$E(X)$存在，$E[X-EX]^2$有意义，则称$E[X-EX]^2$

为 X 的方差，记为 $D(X)$，即 $\quad D(X)=E[X-EX]^2$.

$\sqrt{D(X)}$ 称为 X 的标准差（或均方差），记作 σ.

$D(X)=E[X-EX]^2$ 反映了 X 与 $E(X)$ 的偏离程度，易见，$D(X)$ 越小，则 X 取值越集中，反之，X 取值越分散.

下面讨论方差的计算公式.

设 X 为离散型的随机变量，设 X 的分布为 $p_k=P\{X=x_k\}$，k=1,2,… 则

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2 p_k$$

在前例中，$D(X_1)$=0.5，$D(X_2)$=0.1. $D(X_1)>D(X_2)$,所以乙的成绩较甲集中.

若 X 为连续型随机变量，其密度函数为 $p(x)$,则

$$D(X)=\int_{-\infty}^{+\infty}(x-E(X))^2 p(x)dx$$

在具体计算方差时，常用到下列计算公式

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

证明留给读者.

[例 1]

[例 2] 设随机变量 X 的密度函数为 $\quad p(x)=\frac{1}{2}e^{-|x|}$, $-\infty<x<+\infty$（拉普拉斯分布）,求 $E(X)$ 及 $D(X)$.

解：注意到密度函数 $p(x)=\frac{1}{2}e^{-|x|}$, $-\infty<x<+\infty$ 是一个偶函数，所以

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=0$$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2 f(x)dx=\int_{-\infty}^{0}x^2\frac{1}{2}e^{x}dx+\int_{0}^{\infty}x^2\frac{1}{2}e^{-x}dx$$

$$=\frac{1}{2}[e^{x}(x^2-2x+2]_{-\infty}^{0}+\frac{1}{2}[-e^{-x}(x^2+2x+2)]_{0}^{\infty}=2$$

所以 $\quad D(X)=E(X^2)-[E(X)]^2=2$

二、方差的性质

1）$D(C=0)$;

2）$D(CX)=C^2D(X)$;

3）设随机变量 X 与 Y 相互独立，则 $D(X+Y)=D(X)+D(Y)$.

证：$D(X+Y)=E[(X+Y)-E(X+Y)]^2=E\{[X-E(X)]+[Y-E(Y)]\}^2$

$$=E[X-E(X)]^2+2E\{[X-E(X)]\ [Y-E(Y)]\}+E[Y-E(Y)]^2$$

由于 X 与 Y 相互独立，所以，$X-E(X)$ 与 $Y-E(Y)$ 也相互独立，从而 $E\{[X-E(X)]\ [Y-E(Y)]\}=0$

故 $D(X+Y)=D(X)+D(Y)$

三、常见分布的方差

(1) 两点分布 $P\{X=1\}=p, P\{X=0\}=1-p=q$

由 $X^2=X$，$E(X)=p$ 立得

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

(2) 二项分布 $X\sim B(n,p)$

因为 $E(X^2)=E[X(X-1)]+E(X)$

$$=\sum_{k=0}^{n}C_n^k k(k-1)p^k q^{n-k}+np$$

$$=p^2\frac{d^2}{dx^2}\sum_{k=0}^{n}C_n^k x^k q^{n-k}\Big|_{x=p}+np$$

$$=p^2\frac{d^2}{dx^2}(x+q)^n\Big|_{x=p}+np$$

$$=n(n-1)p^2+np$$

故 $D(X)=E(X^2)-[E(X)]^2=n(n-1)p^2+np-(np)^2=npq$.

(3) Poisson 分布 $X\sim P(\lambda),\lambda>0$

$$E(X^2)=E[X(X-1)]+E(X)$$

$$=\sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k}{k!}e^{-\lambda}+\lambda$$

$$=\lambda^2\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}e^{-\lambda}+\lambda$$

$$=\lambda^2+\lambda$$

所以 $$D(X)=E(X^2)-[E(X)]^2=\lambda^2+\lambda-\lambda^2=\lambda.$$

（4）几何分布　设随机变量 X 服从几何分布，则其分布列为

$$P\{X=k\}=pq^{k-1},k=1,2,\cdots$$

$$E(X^2)=E[X(X-1)]+E(X)$$

$$=\sum_{k=1}^{\infty}k(k-1)pq^{k-1}+\frac{1}{p}$$

$$=pq(\sum_{k=1}^{\infty}q^k)''+\frac{1}{p}$$

$$=pq(\frac{1}{1-q})''+\frac{1}{p}$$

$$=\frac{2pq}{(1-q)^3}+\frac{1}{p}$$

$$=\frac{2q}{p^2}+\frac{1}{p}$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{2q}{p^2}+\frac{1}{p}-\frac{1}{p^2}=\frac{q}{p^2}$$

(5) 均匀分布　若 $X\sim U[a,b]$，　则

$$E(X^2)=\int_{-\infty}^{+\infty}x^2p(x)dx=\int_a^b x^2\cdot\frac{1}{b-a}dx=\frac{a^2+ab+b^2}{3}$$

$$D(X)=E(X^2)-[E(X)]^2=\frac{a^2+ab+b^2}{3}-(\frac{a+b}{2})^2=\frac{(b-a)^2}{12}$$

(6) 正态分布　　若 $X\sim N(\mu,\sigma^2)$

$$D(X)=\int_{-\infty}^{+\infty}(x-\mu)^2p(x)dx$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}(x-\mu)^2e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2e^{-\frac{t^2}{2}}dt$$　　（取 $x-\mu=\sigma t$）

$$=\sigma^2$$

(7) 指数分布　X 的概率密度　$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$　$\lambda>0$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2p(x)dx=\int_0^{+\infty}x^2\lambda e^{-\lambda x}dx$$

$$=\frac{1}{\lambda^2}\int_0^{+\infty}t^2e^{-t}dt$$

$$=\frac{2}{\lambda^2}$$

故 $D(X)=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2}$

[例 3] 将 n 张不同的信笺随机放入 n 个写好地址的信封，用 X 表示正确搭配的个数，求 X 的数学期望和方差.

解：定义随机变量

$$X_i=\begin{cases}1, & \text{第}i\text{封信正确搭配}\\ 0, & \text{第}i\text{封信没有正确搭配}\end{cases}$$

则 $E(X_i)=P\{X_i=1\}=\frac{1}{n}$，$X=X_1+X_2+\cdots X_n$ 是正确搭配的个数，，于是数学期望，即平均正确搭配的个数是

$$E(X)=E(X_1)+E(X_2)+\cdots E(X_n)=1.$$

对 $i<j$，有

$$E(X_iX_j)=P\{X_i=1,X_j=1\}=\frac{(n-2)!}{n!}$$

故 $D(X)=D(X_1+X_2+\cdots X_n)$

$$=E[(X_1+X_2+\cdots X_n)-E(X_1+X_2+\cdots X_n)]^2$$

$$=\sum_{j=1}^{n}\sum_{i=1}^{n}[E(X_iX_j)-E(X_i)E(X_j)]$$

$$=nE(X_1)+n(n-1)E(X_1X_2)-n^2(EX_1)^2 \qquad \text{（用 } {X_1}^2=X_1\text{）}$$

$$=n\cdot\frac{1}{n}+n(n-1)\frac{1}{n(n-1)}-n^2\frac{1}{n^2}$$

$$=1.$$

注意到这一配对问题的期望与方差都相等，且都等于 1，这是配对问题的一个共同特征.

四、切比雪夫不等式

我们知道，方差反映了随机变量 X 离开 EX 的平均偏离程度，设随机变量为 X 的数学期望为 EX，方差为 DX，那么，对任意的 $\varepsilon>0$，事件 $\{|X-EX|\ge\varepsilon\}$ 发生的概率应该

和 DX 有一定的关系，粗略地说，如果 DX 较大，那么，$P\{|X-EX|\geq\varepsilon\}$ 也会大一些，把这个直觉严格化，就是著名的切比雪夫不等式.

定理 4.4 对于任何具有有限方差的随机变量 X，对任意的 $\varepsilon>0$，有

$$P\{|X-EX|\geq\varepsilon\}\leq\frac{DX}{\varepsilon^2}$$

证明： 设 X 是一连续型随机变量，密度函数为 $p(x)$，则

$$P\{|X-EX|\geq\varepsilon\}=\int\limits_{|x-EX|\geq\varepsilon}p(x)dx\leq\int_{|x-EX|\geq\varepsilon}\frac{|x-EX|^2}{\varepsilon^2}p(x)dx$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{+\infty}|x-EX|^2\ p(x)dx=\frac{DX}{\varepsilon^2}$$

对离散型随机变量，类似可以证明.

不等式给出了在随机变量分布未知情况下对事件 $\{|X-EX|\geq\varepsilon\}$ 或 $\{|X-Ex|<\varepsilon\}$ 的概率的一种估计方法.

[例 4] 假设星期天上午来到某画展陈列室的顾客人数 X 是一个随机变量，其分布未知. 已知 X 的期望为 18（人），标准差为 2.5（人），试估计顾客人数 X 在 8 到 28 人之间的概率为多少?

解： 因为 $8<X<28=-10<X-18<10=\left|X-18\right|<10$，所以，利用切比雪夫不等式，得

$$P\{8<X<28\}=P\{\left|X-18\right|<10\}\geq1-\frac{(2.5)^2}{10^2}=0.9375.$$

即顾客人数 X 在 8 到 28 人之间的概率超过 93%.

§4.3 协方差与相关系数

对二维随机向量 (X,Y) 来说，EX，EY 只反映了 X,Y 各自的平均值，方差 DX,DY 也只反映了 (X,Y) 各自离开均值的偏离程度，它们对 X,Y 之间的相互联系不提供任何信息.然而随机变量之间可能存在着某种相关关系，如人的身高与体重，商品的广告费与销量等.如同数学期望与方差一样，我们当然也希望有一个数字特征能够在一定程度上反映这种联系.

容易证明，当 X,Y 相互独立时，$E[(X-EX)(Y-EY)]$=0.

也就是说，当 $E[(X-EX)(Y-EY)]\neq0$ 时，X,Y 肯定不独立，这说明 $E[(X-EX)(Y-EY)]$ 的数值在一定程度上反映了 X,Y 相互间的联系，因此，引入下列定义：

定义 4.4 若(X,Y)是二维随机向量，又$E|(X-EX)(Y-EY)|<\infty$，则称$E[(X-EX)(Y-EY)]$为X,Y的协方差，记为$Cov(X,Y)$.

即 $$Cov(X,Y)=E[(X-EX)(Y-EY)].$$

由协方差的定义可立得下述性质

（1）$Cov(X,Y)=Cov(Y,X)$

(2) 若 a,b是任意两个常数，则 $Cov(aX,bY)=abCov(X,Y)$

(3) $Cov(X_1+X_2,Y)=Cov(X_1,Y)+Cov(X_2,Y)$

协方差的计算常常使用下列公式

$$Cov(X,Y)=E(XY)-E(X)\cdot E(Y)$$

[例 1] 设二维随机向量(X,Y)的联合密度函数为

$$p(x,y)=\begin{cases} 4, & 0<x<\frac{1}{2},0<y<2x \\ 0, & \text{其他} \end{cases}$$

求$\mathrm{Cov}(X,Y)$.

解：由题设得

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xp(x,y)dxdy=\int_0^{\frac{1}{2}}dx\int_0^{2x}4xdy=\int_0^{\frac{1}{2}}8x^2dx=\frac{1}{3},$$

$$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yp(x,y)dxdy=\int_0^{\frac{1}{2}}dx\int_0^{2x}4ydy=\int_0^{\frac{1}{2}}8x^2dx=\frac{1}{3},$$

$$E(XY)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xyp(x,y)dxdy=\int_0^{\frac{1}{2}}xdx\int_0^{2x}4ydy=\int_0^{\frac{1}{2}}8x^3dx=\frac{1}{8},$$

所以 $$\mathrm{Cov}(X,Y)=E(XY)-E(X)\cdot E(Y)=\frac{1}{8}-\frac{1}{3}\cdot\frac{1}{3}=\frac{1}{72}$$

协方差的数值虽然在一定程度上反映了X,Y相互间的联系，但它还受X,Y本身数值大小的影响.比如说，令X,Y各自增大k倍，则$X_1=kX,Y_1=kY$，这时，X_1与Y_1的相互关系应与X,Y的相互关系是一样的，可是反映这种联系的协方差却增大了 k^2倍，即有$Cov(X_1,Y_1)=k^2Cov(X,Y)$.

为了克服这一缺点，在计算随机变量 X,Y 的协方差之前，先对 X 与 Y 进行“标准化”. 若 X 是一个服从 $N(\mu,\sigma^2)$ 分布的正态变量，前面已提到过 $X^*=\dfrac{X-\mu}{\sigma}$ 是服从 $N(0,1)$ 分布的标准正态变量.这里，所谓“标准”的意思无非是指 $E(X^*)=0,D(X^*)=1$ 而已.

现在对任意一个随机变量 X，若 $E(X)=a$, $D(X)=\sigma^2$,令 $X^*=\dfrac{X-a}{\sigma}$，由数学期望和方差的性质有 $E(X^*)=0$，$D(X^*)=1$，所以，X^* 就是对 X 标准化以后的随机变量. 引入下列定义

定义 4.5　若 (X,Y) 是一个二维随机向量，且 $E\left|\dfrac{X-EX}{\sqrt{DX}}\dfrac{Y-EY}{\sqrt{DY}}\right|<\infty$

则称　　$$\frac{Cov(X,Y)}{\sqrt{DXDY}}$$

为 X 与 Y 的相关系数，记为 ρ.

这是因为

$$Cov(X^*,Y^*)=E(X^*-EX^*)(Y^*-EY^*)=E[\frac{X-EX}{\sqrt{DX}}\frac{Y-EY}{\sqrt{DY}}]=\frac{Cov(X,Y)}{\sqrt{DXDY}}$$

即相关系数是标准化以后的随机变量的协方差.按照前面的讨论，协方差是一个有量纲的数，而相关系数是一个无量纲的数.所以，相关系数能够很好的反映随机变量 X 与 Y 之间的关系，不受量纲的影响.

顾名思义，相关系数反映了随机变量之间的相关——也就是它们相互之间的一种联系，到底是那一种联系呢？这是我们希望进一步弄清楚的问题.

定理 4.5　设二维随机向量 (X,Y) 的两个分量 X,Y 的相关系数为 ρ，则有

（1）$|\rho|\le 1$；

（2）$|\rho|=1$ 的充要条件为 X 与 Y 以概率 1 线性相关.即存在常数 a,b，使得 $P\{Y=aX+b\}=1$.

该定理告诉我们，相关系数只是随机变量间线性关系强弱的一个度量，因而说得更确切一些，应称为线性相关系数.

$|\rho|=1$, 表明 X 与 Y 之间存在完全线性关系；当 $|\rho|<1$ 时，这种线性相关程度随着 $|\rho|$ 的减小而减弱，当 $\rho=0$ 时，称 X 与 Y 不相关（零相关）.不相关，只说明 X 与 Y 不存在线性关系，但是它们之间还是有可能存在着别的函数关系（非线性关系）.

[例 2]　设二维随机向量 (X,Y) 的联合分布列为

$Y \backslash X$	0	1
-1	0	$1/3$
0	$1/3$	0
1	0	$1/3$

求 $\mathrm{Cov}(X,Y)$， $\rho(X,Y)$

解：根据联合分布列得 X,Y,XY 的分布列分别为

X	0	1
P	$1/3$	$2/3$

Y	-1	0	1
P	1/3	1/3	1/3

XY	-1	0	1
P	1/3	1/3	1/3

从而， $E(X)=\dfrac{2}{3}$， $E(Y)=0 \quad E(XY)=0$

故 $\quad Cov(X,Y)=E(XY)-E(X)\cdot E(Y)=0$

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{DXDY}}=0.$$

由此说明 X,Y 不相关.同时注意到 $P\{X=0,Y=0\}=\dfrac{1}{3}\neq P\{X=0\}P\{Y=0\}=\dfrac{1}{9}$，因此， X,Y 不独立.

作为本章的结束，最后我们介绍一下矩的概念.

前面讨论了随机变量的数学期望、方差及协方差等数字特征，这些计算式和物理学当中的静力矩、惯性矩的计算式非常相似，借用物理学中“矩”的名称，也把它们统称为随机变量的矩. $E(X)$， $E(X^2)$ 分别称为一阶矩和二阶矩.一般地，对任意正整数 k，可以自然定义

$$E(X^k)=\int_{-\infty}^{+\infty}x^k p(x)dx \quad 或 \qquad E(X^k)=\sum_{i=1}^{\infty}{x_i}^k p_i$$

如果它们存在的话，称为随机变量的 k 阶矩.

注意到方差$DX = E(X - EX)^2$，当然也可以对任意的正整数k ,考虑$E(X - EX)^k$，它也是一种矩，对应于前面的 k 阶矩，把它称为X的k阶中心矩，而$E\xi^k$是对原点的k阶矩，称为 k 阶原点矩.这样，期望是一阶原点矩，方差是二阶中心矩.