第四章　随机变量的数字特征

在前面的课程中, 我们讨论了随机变量及其分布, 如果知道了随机变量X的概率分布, 那么X的全部概率特性也就知道了. 然而，在实际问题中，概率分布一般是较难确定的, 而且在一些实际应用中, 人们并不需要知道随机变量的一切概率性质，只要知道它的某些数字特征就够了.

例如　考察一射手的水平：既要看他的平均环数是否高，还要看他弹着点的范围是否小, 即数据的波动是否小.

判断棉花质量时，既看纤维的平均长度，又要看纤维长度与平均长度的偏离程度，平均长度越长，偏离程度越小，质量就越好.

由上面例子看到，与随机变量有关的某些数值，虽不能完整地描述随机变量但能清晰地描述随机变量在某些方面的重要特征，这些数字特征在理论和实践上都具有重要意义.

本章内容

- 数学期望 — —随机变量的平均取值；
- 方差 — — — 取值平均偏离均值的情况；
- 协方差、相关系数 — —随机变量相互关系各种矩等；
- 极限定理 — —大数定律，中心极限定理.

§4.1　随机变量的数学期望

一　数学期望的概念

引例　设某射击手在同样的条件下，瞄准靶子相继射击90次，射中次数记录如下：

命中环数 k	0	1	2	3	4	5
命中次数 n_k	2	13	15	10	20	30
频率 $\frac{n_k}{n}$	$\frac{2}{90}$	$\frac{13}{90}$	$\frac{15}{90}$	$\frac{10}{90}$	$\frac{20}{90}$	$\frac{30}{90}$

试问：该射手每次射击平均命中靶多少环?

解：$\text{平均射中环数}=\dfrac{\text{射中靶的总环数}}{\text{射击次数}}$

$$=\frac{0\times 2+1\times 13+2\times 15+3\times 10+4\times 20+5\times 30}{90}$$

$$=0\times\frac{2}{90}+1\times\frac{13}{90}+2\times\frac{15}{90}+3\times\frac{10}{90}+4\times\frac{20}{90}+5\times\frac{30}{90}$$

$$=\sum_{k=0}^{5}k\cdot(n_k/n)=3.37.$$

设射手命中的环数为随机变量X.

$$\text{平均射中环数}=\sum_{k=0}^{5}k\cdot\boxed{\frac{n_k}{n}}\to\text{频率随机波动}$$

“平均射中环数”的稳定值$=?$

$$\sum_{k=0}^{5} k\cdot\frac{n_k}{n} \xrightarrow{n\to\infty} \sum_{k=0}^{5} k\cdot p_k$$

以频率为权的加权平均

以概率为权的加权平均

随机波动 → 稳定值

平均射中环数的稳定值=$\sum\limits_{k=0}^{5} k\cdot p_k$.

数学期望的概念源于此.

二　离散型场合

将上述公式推广到更一般的离散型随机变量，我们引进如下定义.

定义1 设X是一离散型随机变量, 其概率分布为

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

如果级数$\sum_{i=1}^{\infty}|x_i|p_i < +\infty$，则称级数$\sum_{i=1}^{\infty}x_i p_i$为随机变量$X$的数学期望,记为$EX$；如果级数$\sum_{i=1}^{\infty}|x_i|p_i = +\infty$，则称随机变量$X$的数学期望不存在.

注：(1) 要求级数$\sum_{i=1}^{\infty} x_i p_i$绝对收敛，是为了保证级数$\sum_{i=1}^{\infty} x_i p_i$的和不随$x_i$位置的变化而改变.

(2) 一般利用逐项求导或逐项积分求级数$\sum_{i=1}^{\infty} x_i p_i$的和. 例如：

$$\sum_{n=1}^{\infty} n \cdot \frac{1}{2^n} = \frac{1}{2}\sum_{n=1}^{\infty}(x^n)'\Big|_{x=\frac{1}{2}} = \frac{1}{2}(\sum_{n=1}^{\infty} x^n)'\Big|_{x=\frac{1}{2}}$$

$$= \frac{1}{2}(\frac{x}{1-x})'\Big|_{x=\frac{1}{2}} = \frac{1}{2} \cdot \frac{1}{(1-x)^2}\Big|_{x=\frac{1}{2}} = 2$$

例如：$$\sum_{n=1}^{\infty}\frac{1}{n\cdot 2^n}=\sum_{n=1}^{\infty}\int_0^{\frac{1}{2}}x^{n-1}\mathrm{d}x=\int_0^{\frac{1}{2}}(\sum_{n=1}^{\infty}x^{n-1})\mathrm{d}x$$

$$=\int_0^{\frac{1}{2}}\frac{1}{1-x}\mathrm{d}x=-\ln(1-x)\Big|_0^{\frac{1}{2}}=\ln 2$$

显然数学期望由概率分布列唯一确定，因此也称为某概率分布的数学期望.

三　连续型场合

连续型随机变量X可以用：以概率$f(x_i)(x_i-x_{i-1})$取值为x_i的离散型随机变量近似，而其数学期望为

$$\sum_i x_i f(x_i)(x_i - x_{i-1})$$

它是积分$\int_{-\infty}^{+\infty} xf(x)dx$的渐进和式． 因而有如下定义.

定义4.1.2 设X是连续型随机变量，其密度函数为$f(x)$，当$\int_{-\infty}^{+\infty} |x| f(x)dx < +\infty$时，称$\int_{-\infty}^{+\infty} xf(x)dx$为$X$的数学期望(或均值),记为$EX$；即

$$EX = \int_{-\infty}^{+\infty} xf(x)dx.$$

同离散一样，数学期望只与分布有关.

！！
数学期望是一个数，不再是随机变量.

四　例子

一些重要分布的数学期望：

【两点分布】

$$P\{X=1\}=p,\quad P\{X=0\}=q.$$

$$E(X)=p$$

【二项分布】

$$P\{X=k\}=C_n^k p^k q^{n-k},\qquad k=0,1,2,\cdots,n.$$

$$EX=\sum_{k=0}^{n} kC_n^k p^k q^{n-k}=\sum_{k=1}^{n} kC_n^k p^k q^{n-k}$$

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

【泊松分布】

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}, \qquad k=0,1,2,\cdots$$

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

【几何分布】

$$P\{X=k\}=q^{k-1}p, \qquad k=1,2,\cdots.$$

$$EX=\sum_{k=1}^{\infty}kq^{k-1}p=p\sum_{k=1}^{\infty}(q^k)'_q=p(\sum_{k=1}^{\infty}q^k)'_q$$

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

【超几何分布】【帕斯卡分布】

后面利用数学期望的性质求.

【均匀分布】 $X \sim U[a,b]$.

$$EX=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x=\int_a^b x\cdot\frac{1}{b-a}\mathrm{d}x=\frac{a+b}{2}.$$

【正态分布】 $X \sim N(\mu,\sigma^2)$.

$$EX=\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x=\int_{-\infty}^{+\infty}\frac{x}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\mathrm{d}x$$

$$\overset{x=\sigma z+\mu}{=\!=\!=\!=\!=}\int_{-\infty}^{+\infty}(\sigma z+\mu)\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\mathrm{d}z=\mu.$$

练习题　【指数分布】　$X \sim \exp(\lambda)$. 试求EX.

【柯西分布】　柯西分布的密度函数为

$$f(x)=\frac{1}{\pi(1+x^2)}.$$

由于

$$\int_{-\infty}^{+\infty}|x|\,f(x)dx=\int_{-\infty}^{+\infty}|x|\cdot\frac{1}{\pi(1+x^2)}dx=+\infty$$

因而柯西分布不存在数学期望.

练习题　给出离散型数学期望不存在的例子.

【例1】（一种验血新技术）假定有N个人需要验血，可有两种验血方式：（1）每个人的血分别化验，这时需要化验N次；（2）把k个人的血混在一起进行化验，如果结果显阴性，那么对这k 个人只做一次化验就可以了；如果结果显阳性，那么必须对这k个人再逐个进行化验，这时对这个人共需要$k+1$次化验. 假定对所有的人来说，化验是阳性的概率都是p，而且他们的反应是独立的. 设每个k 人小组的检查次数是X. 则

$$P\{X=1\}=q^k, \qquad P\{X=k+1\}=1-q^k.$$

N个人需要的化验次数的期望值是

$$\frac{N}{k}\left[1\cdot q^{k}+(k+1)\cdot(1-q^{k})\right]=N(1-q^{k}+\frac{1}{k}).$$

当$q^{k}-\frac{1}{k}>0$时，就能减少化验次数.

例如当$p=0.1$，取$k=4$，则$q^{k}-\frac{1}{k}=0.4$. 即方法2平均能减少40%的工作量. p越小, 这种方法减少的工作量越大.

五　随机变量函数的数学期望

已知随机变量X的分布，需要计算X的函数$Y=g(X)$的期望. 那么应该如何计算呢？

方法1　由随机变量X的分布，求出随机变量$Y=g(X)$的分布，再根据$g(X)$的分布，按照期望的定义把$E[g(X)]$计算出来.

缺点：求$g(X)$的概率分布比较复杂.

方法2　直接由随机变量X的分布，按照下面定理给出的计算公式把$E[g(X)]$计算出来.

优点：不用求解$g(X)$的概率分布.

定理1 设Y是随机变量X的函数：$Y = g(X)$.

(1) 若X为离散型随机变量，其概率分布为

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

若无穷级数$\sum_{k=1}^{+\infty} g(x_k) p_k$绝对收敛，则有

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

(2) 若X为连续型时，其密度函数为$f(x)$.

若广义积分$\int_{-\infty}^{+\infty} g(x) f(x) dx$绝对收敛，则有

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

定理2 设(X,Y)是二维随机变量：$Z=g(X,Y)$.

(1) 若(X,Y)为离散型随机变量，其概率分布为

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

若无穷级数$\sum\limits_{i,j=1}^{\infty} g(x_i, y_j)p_{ij}$绝对收敛，则有

$$E(Z)=E[g(X,Y)]=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} g(x_i, y_j)p_{ij}.$$

(2) 若(X,Y)为连续型时，其密度函数为$f(x,y)$.

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

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

$$E(Y)=E[g(X)]=\begin{cases}\sum_{k=1}^{\infty}g(x_k)p_k, & X\text{离散型},\\ \int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x, & X\text{连续型}.\end{cases}$$

★

$$E(Z)=E[g(X,Y)]$$

★

$$=\begin{cases}\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}g(x_i,y_j)p_{ij}, & (X,Y)\text{是离散型},\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y, & (X,Y)\text{是连续型}.\end{cases}$$

上述定理还可以推广多维随 机变量的函数的情况.

(练习写出来)

例2　设随机变量X的概率分布如下：

X	-2	-1	0	1
P	0.1	0.2	0.3	0.4

试求$E(3X+1)$，EX^2

解：$E(3X+1)=\sum_{i=1}^{4}(3x_i+1)p_i$

$$=-5\times 0.1-2\times 0.2+1\times 0.3+4\times 0.4=1,$$

$$EX^2=\sum_{i=1}^{4}x_i^2 p_i=4\times 0.1+1\times 0.2+0\times 0.3+1\times 0.4=1.$$

例3 设随机变量X的概率密度为拉普拉斯分布

$$f(x)=\frac{1}{2}\mathrm{e}^{-|x|},\quad -\infty<x<\infty$$

试求$E(X)$，EX^2.

解： $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x=\int_{-\infty}^{+\infty}x\cdot\frac{1}{2}\mathrm{e}^{-|x|}\mathrm{d}x$

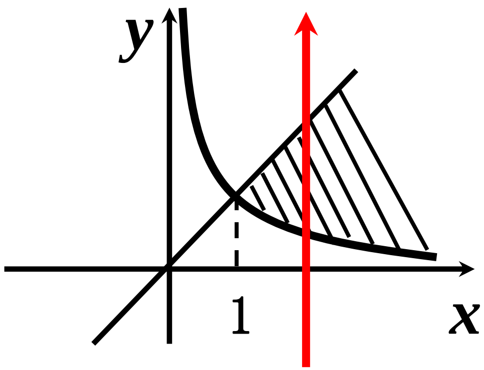
$=0.$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2f(x)\mathrm{d}x=\int_{-\infty}^{+\infty}x^2\cdot\frac{1}{2}\mathrm{e}^{-|x|}\mathrm{d}x$$

$$=\int_0^{+\infty}x^2\cdot\mathrm{e}^{-x}\mathrm{d}x=2.$$

例4　设随机变量(X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\dfrac{3}{2x^3y^2}, & \dfrac{1}{x}<y<x,\ x>1,\\ 0, & \text{其它}.\end{cases}$$



求 $E(\dfrac{1}{XY}),E(Y)$.

解：$E(\dfrac{1}{XY})=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\dfrac{1}{xy}f(x,y)\mathrm{d}x\mathrm{d}y$

$$=\int_{1}^{+\infty}\mathrm{d}x\int_{\frac{1}{x}}^{x}\frac{3}{2x^4y^3}\mathrm{d}y=\frac{3}{2}\cdot\frac{1}{-2}\int_{1}^{+\infty}\frac{1}{x^4}(\frac{1}{x^2}-x^2)\mathrm{d}x$$

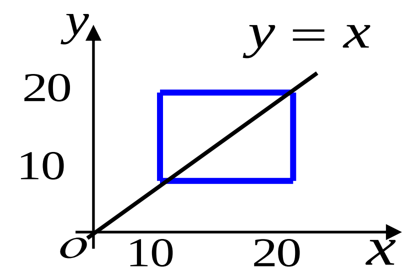
$$=-\frac{3}{4}\int_{1}^{+\infty}(\frac{1}{x^6}-\frac{1}{x^2})\mathrm{d}x=\frac{3}{5}.$$

例5 某商店经销某种商品，每周进货量X与需求量Y是相互独立的随机变量，且都在区间[10, 20]上均匀分布. 商店每售出一单位商品可获利1000元；若需求量超过进货量，商店可从它处调剂供应，这时每单位商品可获利500元；试计算此商店经销该种商品每周所获得利润的数学期望.

解： 设Z表示该种商品每周所得的利润，则

$$Z = g(X,Y) = \begin{cases} 1000Y, & \text{若} Y \le X, \\ 1000X + 500(Y - X), & \text{若} Y > X. \end{cases}$$

X和Y相互独立，因此(X,Y)的概率密度为



$$f(x,y) = \begin{cases} 1/100, & 10 \le x \le 20, 10 \le y \le 20. \\ 0, & \text{其他}. \end{cases}$$

$$E(Z)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_{10}^{20}\mathrm{d}x\int_{10}^{x}1000y\times 1/100\,\mathrm{d}y$$

$$+\int_{10}^{20}\mathrm{d}x\int_{x}^{20}500(x+y)\times 1/100\,\mathrm{d}y$$

$$\approx 14166.7(\text{元}).$$

练习题:

设$(X,Y)\sim N\left(0,0,1,1,0\right), Z=\sqrt{X^2+Y^2}$.

求Z的数学期望$E(Z)$.

练习题解答：

$$E(Z)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}\,f(x,y)\mathrm{d}x\mathrm{d}y$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\sqrt{x^2+y^2}\,\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}\mathrm{d}x\mathrm{d}y$$

$$=\int_{0}^{2\pi}\left[\int_{0}^{+\infty}r\frac{1}{2\pi}e^{-\frac{r^2}{2}}r\mathrm{d}r\right]\mathrm{d}\theta=\sqrt{\frac{\pi}{2}}.$$

六 数学期望的性质

性质1 $E(C)=C$.

性质2 $E(aX)=aE(X)$.

性质3 $E(X+Y)=E(X)+E(Y)$.

更一般地

$$E\left(\sum_{i=1}^{n} a_i X_i + C\right) = \sum_{i=1}^{n} a_i E(X_i) + C,$$

称为数学期望的线性性.

性质4 当X,Y独立时，$E(XY)=E(X)E(Y)$.

证明：设$(X,Y) \sim f(x,y)$.

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

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf_X(x)f_Y(y)dxdy$$

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

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

若 $E(XY)=E(X)E(Y)$，　X,Y不一定独立.

反例　$X \sim N(0,1)$，　$Y=X^2$.

例6 试求二项分布的数学期望.

解: 设$X \sim B(n, p)$, 则X表示n重贝努里试验中的"成功"次数. 令

$$X_i = \begin{cases} 1 & \text{如第}i\text{次试验成功,} \\ 0 & \text{如第}i\text{次试验失败.} \end{cases} \quad i = 1, 2, \cdots, n.$$

则 $$X = X_1 + X_2 + \cdots + X_n.$$

由于 $E(X_i) = p$, 所以

$$E(X) = \sum_{i=1}^{n} E(X_i) = np.$$

练习题

求超几何分布的数学期望.

例6 把数字$1,2,\cdots,n$任意地排成一列，如果数字k恰好出现在第k个位置上，则称为一个巧合，求巧合个数的数学期望.

解: 设巧合个数为X,引入

$$X_k=\begin{cases}1, & \text{数字}k\text{恰好出现在第}k\text{个位置上},\\ 0, & \text{否则}.\end{cases}$$

则 $X=\sum_{k=1}^{n}X_k$, 由

$$E(X_k)=P(X_k=1)=\frac{(n-1)!}{n!}=\frac{1}{n}.$$

得 $E(X)=\sum_{k=1}^{n}E(X_k)=n\cdot\frac{1}{n}=1.$

七　数学期望的其他应用

案例1　发行彩票的创收利润

案例2　如何确定投资决策方向?

案例3　保险费＝? 赔偿金＝?

案例4　报童问题——批发多少报纸?

……………………….