第四章 随机变量的数字特征

随机变量的概率特性
- 分布函数CDF
- 密度函数PDF
- 频率函数PMF

特点 全面、详细、完整

不足 复杂、重点不突出

问题

怎样粗线条地描述随机变量的特性?

简单明了、特征鲜明、直观实用

第四章 随机变量的数字特征

问题 怎样粗线条地描述随机变量的特性？

数字特征

例 甲、乙两射手进行打靶训练，每人各打了100发子弹，成绩如下：

甲：

环数	8	9	10
次数	15	40	45

乙：

环数	8	9	10
次数	35	10	55

怎样评估两人的射击水平？

分析 两人的总环数分别为

甲：$8\times 15+9\times 40+10\times 45=930$ （环）

乙：$8\times 35+9\times 10+10\times 55=920$ （环）

每枪平均环数为

甲：$8\times\dfrac{15}{100}+9\times\dfrac{40}{100}+10\times\dfrac{45}{100}=9.3$ （环）

乙：$8\times\dfrac{35}{100}+9\times\dfrac{10}{100}+10\times\dfrac{55}{100}=9.2$ （环）

◆ 平均值的概念广泛存在

例如 某班级某课程考试的平均成绩

电子产品的平均无故障时间

某地区的日平均气温和日平均降水量

某地区水稻的平均亩产量

某地区的家庭平均年收入

某国家国民的平均寿命

问题 question 怎样定义随机变量的平均值概念？

例 甲、乙两射手进行打靶训练，每人各打了100发子弹，成绩如下：

甲：

环数	8	9	10
次数	15	40	45

乙：

环数	8	9	10
次数	35	10	55

怎样评估两人的射击水平？

进一步分析 记甲每枪击中的环数为 X ,因为射击次数较多, 故可认为 X 的频率函数为

X	8	9	10
p_k	0.15	0.40	0.45

则甲射手每枪平均环数为

$$8\times 0.15+9\times 0.40+10\times 0.45 = 9.3$$

即平均环数为

$$E(X)\triangleq\sum_{k=1}^{3}x_k p_k$$

(一) 离散型 r.v 的数学期望

定义 设 r.v X的频率函数为

X	x_1	x_2	$\cdots$	x_k	$\cdots$
$P\{X=x_k\}$	p_1	p_2	$\cdots$	p_k	$\cdots$

若级数 $\sum_{k=1}^{\infty}|x_k|p_k<+\infty$，则称

$$E(X)\triangleq\sum_{k=1}^{\infty}x_kp_k=\sum_{k=1}^{\infty}x_kP\{X=x_k\}$$

为r.v X的**数学期望(期望、均值).**

“数学期望”(Expectation, Expected value)的由来

“数学期望”是历史上沿用下来的一个名词，可理解为在数学上对r.v进行计算期望得到的值，即平均值

问题 在数学期望的定义中，为什么要求

$$\sum_{k=1}^{\infty}|x_k|p_k<+\infty\ ?$$

分析 由高等数学知

$$\sum_{k=1}^{\infty}|x_k|p_k<+\infty \Longrightarrow E(X)=\sum_{k=1}^{\infty}x_k p_k \text{收敛}$$

且 $E(X)$与 $x_k p_k$ 出现的先后位置无关！

注 若 $\sum_{k=1}^{\infty}|x_k|p_k=+\infty$, 则称 $E(X)$不存在.

例 某商店对某种家用电器的销售采用先使用后付款的方式. 付款额根据使用寿命 X 来确定：

寿命(年)	$X \le 1$	$1 < X \le 2$	$2 < X \le 3$	$X > 3$
付款(元)	1500	2000	2500	3000

假设$X \sim EXP(0.1)$,试求该商店出售一台电器的平均收费额.

解 设出售一台电器的收费额为 Y, 频率函数为

$$P\{Y=1500\}=P\{X\le 1\}=\int_0^1 \frac{1}{10}e^{-x/10}dx=0.0952$$

$$P\{Y=2000\}=P\{1<X\le 2\}=\int_1^2 \frac{1}{10}e^{-x/10}dx=0.0861$$

$$P\{Y=2500\}=P\{2<X\le 3\}=\int_2^3 \frac{1}{10}e^{-x/10}dx=0.0779$$

$$P\{Y=3000\}=P\{X>3\}=\int_3^\infty \frac{1}{10}e^{-x/10}dx=0.7408$$

即

Y	1500	2000	2500	3000
p_k	0.0952	0.0861	0.0779	0.7408

例 某商店对某种家用电器的销售采用先使用后付款的方式. 付款额根据使用寿命 X 来确定:

寿命(年)	$X \le 1$	$1 < X \le 2$	$2 < X \le 3$	$X > 3$
付款(元)	1500	2000	2500	3000

假设$X \sim EXP(0.1)$,试求该商店出售一台电器的平均收费额.

解 设出售一台电器的收费额为 Y, 频率函数为

Y	1500	2000	2500	3000
p_k	0.0952	0.0861	0.0779	0.7408

$$\therefore \quad E(Y) = \sum_{k=1}^{4} y_k P\{Y = y_k\}$$

$$= 1500 \times 0.0952 + 2000 \times 0.0861 + 2500 \times 0.0779 + 3000 \times 0.7408$$

$$= 2732.17$$

即商店出售一台电器平均收费额为 2732.17 元.

例 设 $X \sim P(\lambda)$, 求 $E(X)$.

解 X 的频率函数为

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

$\therefore$ X 的均值为

$$E(X)=\sum_{k=0}^{\infty}k\cdot P\{X=k\}$$

$$=\sum_{k=0}^{\infty}k\cdot\frac{\lambda^k}{k!}e^{-\lambda}$$

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

令 $k-1=i$

$$=\lambda e^{-\lambda}\cdot\sum_{i=0}^{\infty}\frac{\lambda^i}{i!}=\lambda e^{-\lambda}\cdot e^{\lambda}=\lambda$$

例 设 $X \sim b(n,p)$, 求 $E(X)$.

解 $E(X)=\sum_{k=0}^{n} k\cdot C_n^k p^k (1-p)^{n-k}$

$\forall\ p,q\in(0,1)$, 令

$$\varphi(p,q)\overset{\Delta}{=}\sum_{k=0}^{n} C_n^k p^k q^{n-k} \equiv (p+q)^n$$

$$\therefore \quad \frac{\partial\varphi}{\partial p}=\sum_{k=1}^{n} kC_n^k p^{k-1}q^{n-k} \equiv n(p+q)^{n-1}$$

特别令 $q=1-p$,则有

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

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

(二) 连续型 r.v 的数学期望

定义 设 r.v X 的概率密度函数为 $f(x)$, 若

$$\int_{-\infty}^{\infty} |x| f(x)dx < +\infty$$

则称

$$E(X) \triangleq \int_{-\infty}^{\infty} x f(x)dx$$

为 r.v X 的**数学期望(期望、均值)**.

注意离散型和连续型情形的形式一致性

注 若 $\int_{-\infty}^{\infty} |x| f(x)dx = +\infty$, 则称 $E(X)$ 不存在.

例 设 $X \sim U(a,b)$, 求 $E(X)$.

解 X的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其它}\end{cases}$$

$$\therefore \quad E(X)=\int_{-\infty}^{\infty} x f(x) dx$$

$$=\int_a^b \frac{x}{b-a} dx$$

$$=\frac{1}{b-a}\cdot\frac{b^2-a^2}{2}$$

$$=\frac{a+b}{2}$$

例 设$X \sim N(\mu,\sigma^2)$, 求 $E(X)$.

解 $E(X)=\int_{-\infty}^{\infty} x f(x)dx$

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

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

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

$$=\frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt+\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t e^{-\frac{t^2}{2\sigma^2}} dt$$

$$=\frac{\mu}{\sqrt{2\pi}}\sqrt{2\pi}+0$$

$$=\mu$$

例 设某元器件的寿命 X 服从指数分布,其密度为

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta}, & x>0\\ 0, & x\le 0\end{cases}$$

求 X 的数学期望 $E(X)$.

解

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

$$=\int_{0}^{\infty}x\frac{1}{\theta}e^{-\frac{x}{\theta}}dx$$

$$=\theta\cdot\int_{0}^{\infty}\frac{x}{\theta}e^{-\frac{x}{\theta}}d(\frac{x}{\theta})$$

$$=\theta\cdot\int_{0}^{\infty}te^{-t}dt$$

$$=\theta$$

即该元器件的平均寿命为 θ.

工程背景

如果某产品的平均寿命为

$$\theta = 10^k \quad （小时）$$

则称该产品为”k级”产品 $(k = 1, 2, \cdots)$.k越大(级别越高),失效率 10^{-k}越低,则产品的平均寿命越长,可靠性越高.

在航空、航天、军事、医疗等领域，通常要求元器件达9级以上，这意味着该元器件的平均寿命至少为

$$\frac{10^9}{365 \times 24} \approx 114160 \text{ (年)}$$

问题

由一万个9级元器件组成的电子设备的平均寿命为多少年?

$$\frac{10^9}{365 \times 24 \times 10^4} \approx 11.4 \text{ (年)}$$

例 设r.v X服从标准Cauchy分布, 其概率密度为

$$f(x)=\frac{1}{\pi}\cdot\frac{1}{1+x^2},\quad -\infty<x<\infty$$

计算$E(X)$.

解 因为

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

$$=\frac{2}{\pi}\int_{0}^{\infty}\frac{x}{1+x^2}dx$$

$$=\frac{1}{\pi}\int_{0}^{\infty}\frac{dx^2}{1+x^2}$$

$$=\frac{1}{\pi}\ln(1+x^2)\Big|_0^{\infty}=\infty$$

故 $E(X)$不存在.

定理（马尔可夫Markov不等式）

设 r.v. X 满足 $P\{X \geq 0\} = 1$，且 $E(X)$ 存在，则

$$P\{X \geq t\} \leq \frac{E(X)}{t}.$$

证 只证离散型情形，连续情形完全类似.

$$E(X) = \sum_{x} xp(x) = \sum_{x<t} xp(x) + \sum_{x \geq t} xp(x)$$

因 X 只取非负值，故和式中每一项都是非负的.

因此，

$$E(X) \geq \sum_{x \geq t} xp(x) \geq \sum_{x \geq t} tp(x) = tP\{X \geq t\}.$$

即

$$P\{X \geq t\} \leq \frac{E(X)}{t}.$$

对于一般的非负r.v，
无论其概率分布，该结论都成立

(三) r.v 的函数的数学期望

实际背景

飞机机翼受到的压力为

$$W = kV^2$$

其中 V 是风速, $k(>0)$是常数,问机翼受到的平均压力多大?

一般地 设已知

$$X \sim f(x) \quad (\text{概率函数})$$

$$y = g(x) \quad (\text{普通函数})$$

则要求

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

一般的思路

1. $X \sim f(x), Y = g(X) \Rightarrow Y \sim f_Y(y)$
2. $E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy$

一般的思路 1. $X \sim f(x), Y = g(X) \Rightarrow Y \sim f_Y(y)$

2. $E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy$

分析 ① 设 $y = g(x)$ 单调增，其反函数为 $x = h(y)$，则

$$f_Y(y) = h'(y) f(h(y))$$

②
$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y f_Y(y) dy \\ &= \int_{-\infty}^{\infty} y h'(y) f(h(y)) dy \\ &= \int_{-\infty}^{\infty} y f(h(y)) dh(y) \\ &= \int_{-\infty}^{\infty} g(x) f(x) dx \end{aligned}$$

有意思的结果 $E(X) = \int_{-\infty}^{\infty} x f(x) dx$

$$E(Y) = E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$$

定理 设 $y = g(x)$ 为普通函数, 则

① 设 X 为离散型r.v, 其频率函数为

$$P\{X = x_k\} = p_k, k = 1, 2, \cdots$$

若 $\sum_{k=1}^{\infty} |g(x_k)| \cdot p_k < +\infty$, 则

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

② 设 X 为连续型r.v, 其概率密度为 $f(x)$,

若 $\int_{-\infty}^{\infty} |g(x)| f(x) dx < \infty$, 则

$$E(Y) = E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

注意二者的形式一致性

例　设风速$V \sim U(0,a)$,设飞机机翼受到的正压力W与风速的关系是$W = kV^2$ $(k > 0$常数$)$.求$E(W)$.

解　V的密度函数为

$$f(v) = \begin{cases} \frac{1}{a}, & 0 < v < a \\ 0, & \text{其它} \end{cases}$$

$$\therefore \quad E(W) = \int_{-\infty}^{\infty} kv^2 f(v)dv$$

$$= \frac{k}{a}\int_0^a v^2 dv$$

$$= \frac{1}{3}ka^2$$

即飞机机翼受到的平均正压力为$\frac{1}{3}ka^2$.

推广的定理 设 $z = g(x, y)$ 为二元函数，则

① 设 X, Y 的联合频率函数为

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

若 $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |g(x_i, y_j)| p_{ij} < +\infty$，则

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

$$= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g(x_i, y_j) p_{ij}$$

② 设 X, Y 的联合密度为 $f(x, y)$，若

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |g(x, y)| f(x, y) dxdy < \infty,$$

则

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

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y) dxdy$$

注：公式可推广到一般的高维随机变量

例 设(X,Y)的密度函数为 $f(x,y)=\begin{cases}15xy^2, 0\le y\le x\le 1,\\ 0\quad, \quad 其它,\end{cases}$

试求$E(X)$，$E(Y)$，$E(XY)$．

一般的思路

$$f(x,y)$$

$$\Downarrow$$

$$f_X(x),\ f_Y(y)$$

$$\Downarrow$$

$$E(X)=\int_{-\infty}^{\infty} x f_X(x)dx$$

$$E(Y)=\int_{-\infty}^{\infty} y f_Y(y)dy$$

另一种方法

$$\boxed{g(x,y)=x}$$

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

例　设(X,Y)的密度函数为 $f(x,y)=\begin{cases}15xy^2, 0\le y\le x\le 1,\\ 0\quad,\quad \text{其它},\end{cases}$

试求 $E(X)$，$E(Y)$，$E(XY)$.

解

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy=\iint_{0\le y\le x\le 1} x\cdot 15xy^2dxdy$$

$$=\int_0^1 dx\int_0^x 15x^2y^2dy=\frac{15}{3}\int_0^1 x^5dx=\frac{5}{6}$$

$$E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy=\frac{5}{8}$$

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

$$=\iint_{0\le y\le x\le 1} xy\cdot 15xy^2dxdy=\frac{15}{28}$$

(四) 数学期望的基本性质

① 设 $a \le X \le b \quad (a.e)$, 则 $a \le E(X) \le b$

② 设 c 为常数, 则 $E(cX) = cE(X)$

③ 设 X、Y 为r.v, 则有 $E(X+Y)=E(X)+E(Y)$

④ 设 X, Y 相互独立,则有 $E(XY)=E(X)E(Y)$

几个推论

1. 若 $X = c \quad (a.e)$, 则 $E(X) = c$

2. 设 $a_1, a_2, \cdots, a_n$ 为常数, $X_1, X_2, \cdots, X_n$ 为r.v, 则

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

3. 设 $X_1, X_2, \cdots, X_n$ 相互独立, 则

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

例 一民航客车载有20位旅客自机场开出，沿途有十个停靠站，如到达一个车站时没有乘客下车就不停车．以 X 表示停车的次数，求 $E(X)$．（假定每位旅客在任一车站下车是等可能的，且各旅客是否下车相互独立）．

解 引入 r.v

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 站有人下车} \\ 0, & \text{第 } i \text{ 站没人下车} \end{cases} \qquad i = 1, 2, \cdots, 10$$

$\because \quad X_i = 0 \Leftrightarrow$ 20位乘客在第 i 站都不下车

$$\therefore \quad P\{X_i = 0\} = \left(\frac{9}{10}\right)^{20}, P\{X_i = 1\} = 1 - \left(\frac{9}{10}\right)^{20}, i = 1, 2, \cdots, 10$$

易知 $X = X_1 + X_2 + \cdots + X_{10}$，从而

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_{10})$$

$$= 10E(X_1) = 10\left[1 - \left(\frac{9}{10}\right)^{20}\right] = 8.784 \text{（次）}$$

例 旅游团的 N 个游客出酒店时都将自己房间的钥匙交给了导游. 回到酒店后, 每人从导游处任取一把钥匙去开自己房间的门. 试问平均有多少人能打开房门.

解 令

$$X_i = \begin{cases} 1, & \text{第}\, i \,\text{人能打开房门} \\ 0, & \text{第}\, i \,\text{人不能打开房门} \end{cases} \quad i = 1, 2, \cdots, N$$

则能打开房门的人数为

$$X = X_1 + X_2 + \cdots + X_N$$

且

$$P\{X_i = 1\} = \frac{1}{N}, P\{X_i = 0\} = 1 - \frac{1}{N}, i = 1, 2, \cdots, N$$

故能打开房门的平均人数为

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_N)$$

$$= N \cdot E(X_1) = N \cdot \frac{1}{N} = 1 \quad (\text{人})$$

例 设N件产品中有M件次品,在该批产品中任意取 n 件,记 X 表示取出的次品个数, 求$E(X)$.

解 X的频率函数为

$$p_k = P\{X = k\} = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k = 0,1,2,\cdots,n$$

称 X 服从**超几何分布**. 故

$$E(X) = \sum_{k=0}^{n} k p_k$$

$$= \sum_{k=0}^{n} k \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

直接求和很难

例 设N件产品中有M件次品,在该批产品中任意取 n 件,记 X 表示取出的次品个数, 求$E(X)$.

解二 令

$$X_k=\begin{cases}1, & \text{第}k\text{件取出次品}\\ 0, & \text{第}k\text{件取出正品}\end{cases}\qquad k=1,2,\cdots,n$$

则

$$P\{X_k=1\}=\frac{M}{N},P\{X_k=0\}=\frac{N-M}{N},k=1,2,\cdots,n$$

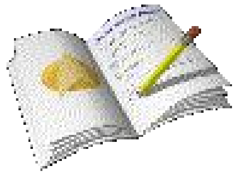
因为$X=\sum_{k=1}^{n}X_k$, 故

$$E(X)=\sum_{k=1}^{n}E(X_k)$$

$$=\sum_{k=1}^{n}(1\cdot\frac{M}{N}+0\cdot\frac{N-M}{N})=\frac{nM}{N}$$

从而求得公式

$$\sum_{k=0}^{n}k\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}=\frac{nM}{N}$$



课后作业

P116：6、15、20、21、31，补充题1，2

1. 设随机变量X的概率密度为

$$f(x)=\begin{cases} e^{-x}, & x>0, \\ 0, & x\leqslant 0. \end{cases}$$

求 (1) $Y=2X$; (2) $Y=e^{-2X}$ 的数学期望.

2. 设随机变量(X,Y)的概率密度为

$$f(x,y)=\begin{cases} 12y^2, & 0\leqslant y\leqslant x\leqslant 1, \\ 0, & \text{其他}. \end{cases}$$

求 $E(X)$, $E(Y)$, $E(XY)$, $E(X^2+Y^2)$.