第四章　随机变量的数字特征

§1 数学期望

- 数学期望的定义
- 随机变量函数的数学期望
- 数学期望的性质
- 应用问题中数学期望的求解举例

一、数学期望定义

1）　离散型

设离散型随机变量X的分布律为：

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

若级数 $\sum_{k=1}^{+\infty} x_k p_k$ 绝对收敛，则称该级数的和为随机变量 X 的数学期望，记作 $E(X)$，

即

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

数学期望简称期望，也称为均值。

2）连续型

设连续型随机变量X的概率密度为 $f(x)$，

若积分 $\int_{-\infty}^{+\infty} xf(x)dx$ 绝对收敛，则称积分$\int_{-\infty}^{+\infty} xf(x)dx$

的值为 X 的数学期望。

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

说 明

(1) $E(X)$是一个常数，它刻划了 X变化的平均值，完全由 X的概率分布所决定。

(2)随机变量的数学期望存在和不存在的问题:

$$\left.\begin{array}{l}\text{对离散型,}\quad \text{只有当}\sum_k |x_k|\cdot p_k\text{收敛时}\\ \text{对连续型,}\quad \text{只有当}\int_{-\infty}^{+\infty}|x|f(x)dx\text{收敛时}\end{array}\right\}\text{, 称}X\text{的数学期望存在}$$

当条件收敛时，仍然称 X的数学期望不存在。

(3)按定义求数学期望:

a)对取无穷多个值的离散型随机变量，是无穷级数求和问题;

b)对连续型随机变量，是无穷积分问题.

例1

设随机变量 X 服从$Cauchy$分布，其密度函数为

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

由于

$$\int_{-\infty}^{+\infty}|x|\cdot f(x)\ dx=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{|x|}{1+x^2}dx=\frac{2}{\pi}\int_0^{+\infty}\frac{x}{1+x^2}dx$$

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

这表明积分$\int_{-\infty}^{+\infty}xf(x)dx$不绝对收敛，因而$E(X)$不存在。

例 2

某商店对某种家电的销售采用先使用后付款的方式.记使用寿命为 X (以年记), 规定

$$X \le 1,\quad \text{一台付款1500元;}$$

$$1 < X \le 2,\quad \text{一台付款2000元;}$$

$$2 < X \le 3,\quad \text{一台付款2500元;}$$

$$X > 3,\quad \text{一台付款3000元;}$$

设寿命X服从指数分布,概率密度为$f(x)=\begin{cases}\dfrac{1}{10}e^{-\frac{x}{10}} & x>0\\ 0 & x\le 0\end{cases}$

试求该商店一台家电收费 Y 元的数学期望。

解:

先求寿命X落在各个时间区间的概率.即有

$$P\{X \le 1\} = \int_0^1 \frac{1}{10} e^{-\frac{x}{10}} dx = 1 - e^{-0.1} = 0.0952$$

$$P\{1 < X \le 2\} = \int_1^2 \frac{1}{10} e^{-\frac{x}{10}} dx = e^{-0.1} - e^{-0.2} = 0.0861$$

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

$$P\{X > 3\} = \int_3^{+\infty} \frac{1}{10} e^{-\frac{x}{10}} dx = e^{-0.3} = 0.7408$$

Y的分布律

Y	1500	2000	2500	3000
P	0.0952	0.0861	0.0779	0.7408

$E(Y) = 2732.15$, 即平均一台家电收费2732.15元。

二、随机变量函数的数学期望

定理 1： 设 Y=g(X), g(x) 是连续函数，

（1）若 X 的分布律为 $p_k = P\{X = x_k\}，\ k = 1,2,\cdots$

且 $\sum_{k=1}^{+\infty} p_k g(x_k)$ 绝对收敛， 则 $E(Y) = \sum_{k=1}^{+\infty} p_k g(x_k)$

(2)若 X 的概率密度为 $f(x)$，且 $\int_{-\infty}^{+\infty} g(x)f(x)dx$ 绝对收敛，

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

定理 2： 若(X, Y)是二维随机变量，$g(x, y)$是二元连续函数，$Z = g(X,Y)$

(1) 若(X, Y)的分布律为

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

且$\sum\limits_{i,j=1}^{+\infty} g(x_i, y_j)p_{ij}$绝对收敛，则 $E(Z) = \sum\limits_{i,j=1}^{+\infty} g(x_i, y_j)p_{ij}$

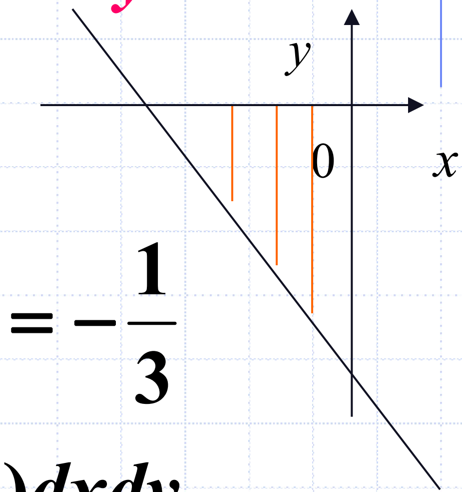
(2) 若(X,Y)的概率密度为$f(x, y)$，且

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

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

例 3 设(X, Y)在区域 A 上服从均匀分布，其中A为x轴，y 轴和直线 $x+y+1=0$ 所围成的区域。求 $E(X)$，$E(-3X+2Y)$，$E(XY)$。

$x + y + 1 = 0$



解：

$$f(x,y)=\begin{cases}2,(x,y)\in A\\0,\text{其它};\end{cases}$$

$$E(X)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dxdy=\int_{-1}^{0}dx\int_{-1-x}^{0}x\cdot 2dy=-\frac{1}{3}$$

$$E(-3X+4Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(-3x+4y)f(x,y)dxdy$$

$$=\int_{-1}^{0}dx\int_{-x-1}^{0}2(-3x+2y)dy=\frac{1}{3}$$

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)dxdy=\int_{-1}^{0}dx\int_{-1-x}^{0}x\cdot 2ydy=\frac{1}{12}$$

例4　国际市场上每年对我国某种出口商品的需求量是随机变量 X（吨），$X \sim U(2000,4000)$，每售出这种商品一吨，可为国家挣得外汇3万元，但销售不出而囤积在仓库，则每吨需浪费保养费1万元。问需要组织多少货源，才能使国家收益最大。

解:

设 y 为预备出口的该商品的数量，则 $2000 < y < 4000.$

用 Z 表示国家的收益（万元）

$$Z = \begin{cases} 3y, & X \geq y \\ 3X - (y - X), & X < y \end{cases}$$

$$z = g(x) = \begin{cases} 3y, & x \ge y \\ 3x-(y-x), & x < y \end{cases}, \quad 2000 < y < 4000$$

下面求 $E(Z)$，并求 y 使 $E(Z)$ 达到最大值，

$$E(Z) = \int_{-\infty}^{\infty} g(x)f(x)dx = \int_{2000}^{y} \frac{3x-(y-x)}{2000}dx + \int_{y}^{4000} \frac{3y}{2000}dx$$

$$= -\frac{1}{1000}[y^2 - 7000y + 4\times10^6]$$

$$= -\frac{1}{1000}[(y-3500)^2 - 3500^2 + 4\times10^6]$$

$$= -\frac{1}{1000}(y-3500)^2 + 8250$$

即，组织3500吨此种商品是最佳的决策。

三、数学期望的性质

1) $E(c)=c$,　c是常数.

若　$a\le X\le b$，则　$a\le E(X)\le b$.

2) $E(cX)=cE(X)$,　c是常数.

3) $E(aX+bY)=aE(X)+bE(Y$ 常数.

有重要应用!

推广：$E(\sum_{i=1}^{n}a_iX_i)=\sum_{i=1}^{n}a_iE(X_i)$,　其中$a_i$是常数.

4) 若X、Y 相互独立，则$E(XY)=E(X)\cdot E(Y)$.

如$X_1,\cdots,X_n$相互独立,则$E(X_1\cdots X_n)=E(X_1)\cdots E(X_n)$

例 5　对N个人进行验血，有两种方案：

（1）对每人的血液逐个化验，共需 N 次化验；

（2）将采集的每个人的血分成两份，然后取其中的一份，按k个人一组混合后进行化验（设N是k的倍数），若呈阴性反应，则认为k个人的血都是阴性反应，这时k个人的血只要化验一次；如果混合血液呈阳性反应，则需对k个人的另一份血液逐一进行化验，这时k个人的血要化验$k+1$次；

假设每个人的血液呈阳性反应的概率都是 p，且各次化验结果是相互独立的。

试说明适当选取 k 可使第二个方案减少化验次数。

解： 设 X 表示第二个方案下的总化验次数，

X_i 表示第 i 个组的化验次数，则

$$X=\sum_{i=1}^{N/k} X_i\,,\ \text{且}\ E(X)=\sum_{i=1}^{N/k} E(X_i)$$

$E(X)$表示第二种方案下总的平均化验次数，$E(X_i)$表示第i个组的平均化验次数。

X_i 只可能取两个值 1 或 $k+1$,

$$P\{X_i=1\}=q^k,\quad P\{X_i=k+1\}=1-q^k,$$

$$E(X_i)=q^k+(k+1)(1-q^k)=k+1-kq^k \qquad (q=1-p)$$

$$E(X_i) = q^k + (k+1)(1-q^k) = k+1-kq^k \quad (q = 1-p)$$

$$i = 1,2,\cdots,N/k.$$

所以 $E(X) = \dfrac{N}{k}(k+1-kq^k) = N(1+\dfrac{1}{k}-q^k)$

只要选 k 使 $1+1/k-q^k < 1$, 即

$$1/k < q^k$$

就可使第二个方案减少化验次数;

当 q 已知,

若选 k 使 $f(k) = 1+1/k-q^k$ 取最小值,

就可使化验次数最少。

例如：当p=0.1，q=0.9时，可证明 k=4 可使最小；这时，

$$E(X) = N(1+1/4-0.9^4) = 0.5939N$$

工作量将减少40%.

例6 将n只球放入到M只盒子中去,设每只球落入各个盒子是等可能的,求有球的盒子数X的均值$E(X)$.

解: 设 $X_i=\begin{cases}1,\text{第}i\text{个盒子有球},\\0,\text{第}i\text{个盒子无球}.\end{cases}\quad i=1,2,\cdots,M.$

则 $X=X_1+\cdots+X_M,\quad E(X)=\sum_{i=1}^{M}E(X_i),$

$$P\{X_i=0\}=(1-\frac{1}{M})^n,\ P\{X_i=1\}=1-(1-\frac{1}{M})^n,$$

$$E(X_i)=1-(1-\frac{1}{M})^n,\quad i=1,\cdots,M,$$

$$E(X)=\sum_{i=1}^{M}E(X_i)=M\cdot[1-(1-\frac{1}{M})^n].$$

典型方法:

将随机变量 X 分解成几个分布相对简单的随机变量之和,

然后利用

随机变量和的数学期望等于随机变量数学期望之和

来求 X 的数学期望。

课堂练习

1 某人的一串钥匙上有n把钥匙,其中只有一把能打开自己的家门,他随意地试用这串钥匙中的某一把去开门,若每把钥匙试开一次后除去,求打开门时试开次数的数学期望.

2 设随机变量X的概率密度为

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

求$Y=e^{-2X}$的数学期望。

1. 解： 设试开次数为X,

X是离散型随机变量，其分布律为：

$P(X=k)=1/n$, $k=1, 2, \ldots, n$

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

2 .解： Y是随机变量X的函数,

$$E(Y)=\int_{-\infty}^{+\infty} e^{-2x} f(x)dx=\int_{0}^{+\infty} e^{-2x} e^{-x} dx=\frac{1}{3}$$

练习题　用某台机器生产某种产品，已知正品率随着该机器所用次数的增加而指数下降，即

$$P\{\text{第}k\text{次生产出的产品是正品}\} = e^{-\lambda k}, k = 1,2,\cdots,\lambda > 0.$$

假设每次生产100件产品，试求这台机器前10次生产中平均生产的正品总数。

解： 设X是前10次生产的产品中的正品数，并设

$$X_{ki} = \begin{cases} 1, \text{第}k\text{次生产的第}i\text{件产品是正品;} \\ 0, \text{否则.} \end{cases}$$

$$k = 1,2,\cdots,10, i = 1,2,\cdots,100,$$

则

$$X = \sum_{k=1}^{10}\sum_{i=1}^{100} X_{ki}.$$

而X_{ki}服从 $p=e^{-\lambda k}$ 的(0—1)分布，

$$E(X_{ki})=e^{-\lambda k},\quad i=1,2,\cdots,100,$$

所以

$$E(X)=\sum_{k=1}^{10}\sum_{i=1}^{100}E(X_{ki})=\sum_{k=1}^{10}100e^{-k\lambda}$$

$$=100\sum_{k=1}^{10}e^{-k\lambda}=\frac{100e^{-\lambda}(1-e^{-10\lambda})}{1-e^{-\lambda}}$$

本节小结：

1）数学期望的定义。

2）随机变量函数的数学期望。

3）数学期望的性质及典型应用。

作业:

P87-88:　24，　26, 27, 29