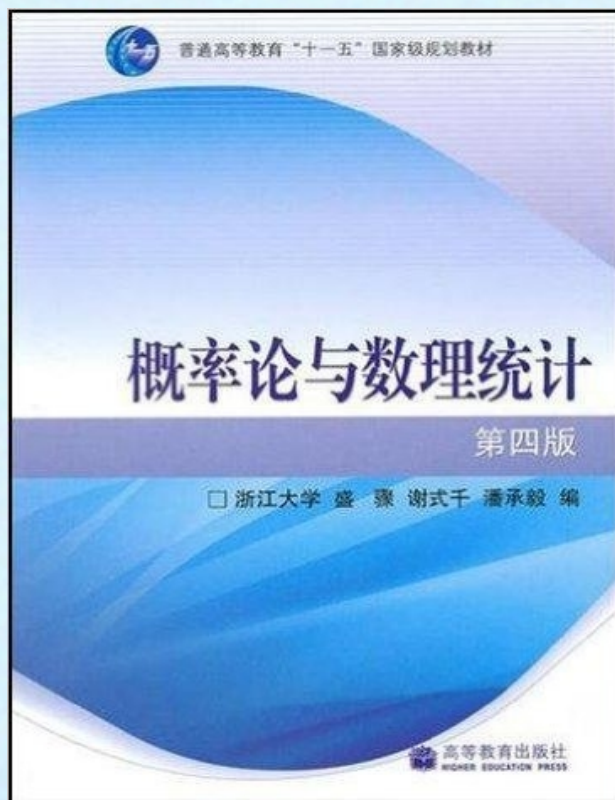


# 目 录



- 第一章 概率论的基本概念
- 第二章 随机变量及其分布
- 第三章 多维随机变量及其分布
- 第四章 随机变量的数字特征
- 第五章 大数定律及中心极限定理
- 第六章 样本及抽样分布
- 第七章 参数估计
- 第八章 假设检验
- 第九章 方差分析及回归分析

## 第四章 随机变量的数字特征

### 第一节 数学期望

### 第二节 方差

### 第三节 协方差及相关系数

### 第四节 矩、协方差矩阵



## 第一节 数学期望

### (一) 离散型随机变量的数学期望

一、问题引入

二、离散型数学期望的定义

三、例题

四、小结、思考

# 一、问题引入

前面讨论了随机变量的分布函数，从中知道随机变量的分布函数能完整地描述随机变量的统计规律。

但在许多实际问题中，人们并不需要去全面考察随机变量的变化情况，而只需要知道它的数字特征即可。

例如：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；

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.

做某项投资时,希望预期收益越高越好,同时又要尽可能地降低投资风险;

由上面例子看到,与  $r.v.$  有关的某些数值,虽不能完整地描述  $r.v.$  但能清晰地描述  $r.v.$  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.



## 引例 射击问题

设某射击手在同样的条件下,瞄准靶子相继射击90次(命中的环数是一个随机变量).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 平均射中环数 =  $\frac{\text{射中靶的总环数}}{\text{射击次数}}$

$$\begin{aligned} &= \frac{0 \times 2 + 1 \times 13 + 2 \times 15 + 3 \times 10 + 4 \times 20 + 5 \times 30}{90} \\ &= 0 \times \frac{2}{90} + 1 \times \frac{13}{90} + 2 \times \frac{15}{90} + 3 \times \frac{10}{90} + 4 \times \frac{20}{90} \\ &\quad + 5 \times \frac{30}{90} \\ &= \sum_{k=0}^5 k \cdot \frac{n_k}{n} = 3.37. \end{aligned}$$

设射手命中的环数为随机变量.

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

随机波动

“平均射中环数”的稳定值 = ?

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

随机波动  $\longrightarrow$  稳定值

“平均射中环数”等于

射中环数的可能值与其概率之积的累加



## 二、离散型数学期望的定义

定义 设离散型随机变量  $X$  的分布律为

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

若级数  $\sum_{k=1}^{\infty} x_k p_k$  绝对收敛, 则称级数的和  $\sum_{k=1}^{\infty} x_k p_k$

为随机变量  $X$  的数学期望, 记为  $E(X)$ .

即

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

$x$ 的概率  
平均值
  $\leftarrow$ 
 $E(X) = \sum_{k=1}^{\infty} x_k p_k$ 
 $\rightarrow$ 
 权重

## 射击问题

“平均射中环数”应为随机变量 $X$ 的数学期望

$$E(X) = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + 4 \cdot p_4 + 5 \cdot p_5.$$

## 关于定义的几点说明

(1) 级数的绝对收敛性保证了级数的和不随级数各项次序的改变而改变,之所以这样要求是因为数学期望是反映随机变量 $X$ 取可能值的平均值,它不应随可能值的排列次序而改变.

(2)  $E(X)$  是一个实数, 而非变量, 它是一种**加权平均**, 与一般的平均值不同, 它从本质上体现了随机变量  $X$  取**可能值**的**真正的平均值**, 也称均值.

假设 

|     |      |      |
|-----|------|------|
| $X$ | 1    | 2    |
| $p$ | 0.02 | 0.98 |

 $x$  算术平均值为  $\frac{1+2}{2} = 1.5$ ,

$$E(X) = 1 \times 0.02 + 2 \times 0.98 = 1.98.$$



它从本质上体现了随机变量  $X$  取可能值的平均值.

当随机变量  $X$  取各个可能值是等概率分布时,  $X$  的期望值与算术平均值相等.



### (3) 不是所有离散型随机变量的数学期望都存在

例如：设随机变量 $X$ 的分布律为

$$P\{X = (-1)^{j+1} \frac{3^j}{j}\} = \frac{2}{3^j}, \quad j = 1, 2, \dots,$$

因为

$$\sum_{j=1}^{\infty} \left| (-1)^{j+1} \frac{3^j}{j} \right| P\{X = (-1)^{j+1} \frac{3^j}{j}\} = \sum_{j=1}^{\infty} \frac{3^j}{j} \frac{2}{3^j} = 2 \sum_{j=1}^{\infty} \frac{1}{j} \quad (\text{发散})$$

所以 $X$ 的数学期望不存在.

### 三、例题

例1 某医院当新生儿诞生时,医生要根据婴儿的皮肤颜色、肌肉弹性、反应的敏感性、心脏的搏动等方面的情况进行评分,新生儿的得分  $X$  是一个随机变量.据以往的资料表明  $X$  的分布律为

| $X$   | 0     | 1     | 2     | 3     | 4    | 5    | 6    | 7    |
|-------|-------|-------|-------|-------|------|------|------|------|
| $p_k$ | 0.002 | 0.001 | 0.002 | 0.005 | 0.02 | 0.04 | 0.18 | 0.37 |
|       | 8     | 9     | 10    |       |      |      |      |      |
|       | 0.25  | 0.12  | 0.01  |       |      |      |      |      |

试求  $X$  的数学期望  $E(X)$ .

**解**

$$\begin{aligned} E(X) = & 0 \times 0.002 + 1 \times 0.001 + 2 \times 0.002 \\ & + 3 \times 0.005 + 4 \times 0.02 + 5 \times 0.04 \\ & + 6 \times 0.18 + 7 \times 0.37 + 8 \times 0.25 \\ & + 9 \times 0.12 + 10 \times 0.01 = 7.15 \text{ (分)} \end{aligned}$$

结果表明,若考察医院出生的很多新生儿,例如1000个,那么一个新生儿的平均得分约7.15分,1000个新生儿共得分7150分.



**例 2** 某省的体育彩票中,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, ..., 8, 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,中大奖的奖金300万元,交20%的税,某人购买一注时期望盈利多少?

**解** 用  $X$  表示其购买一注时的收益, 中大奖获利

$300 \times 80\% = 240$  万元, 于是

$$P(X = 240 \text{ 万元}) = \frac{1}{10^7} \quad P(X = -1 \text{ 元}) = 1 - \frac{1}{10^7}$$

$$EX = -1 \times (1 - 10^{-7}) + 24 \times 10^5 \times 10^{-7} = -0.76.$$

此说明此人每买一注, 平均要损失 0.76 元。

**例3** 在一个人数很多的团体中普查某种疾病, 为此要抽验  $N$  个人的血, 可以用两种方法进行:

- (i) 将每个人的血分别去化验, 这就需化验  $N$  次.
- (ii) 按  $k$  个人一组进行分组, 把从  $k$  个人抽来的血混合在一起进行化验, 如果这混合血液呈阴性反应, 就说明  $k$  个人的血都呈阴性反应, 这样, 这  $k$  个人的血就只需验一次. 若呈阳性, 则再对这  $k$  个人的血液分别进行化验, 此时,  $k$  个人的血总共最多化验  $k + 1$  次. 假设每个人化验呈阳性的概率为  $p$ , 且

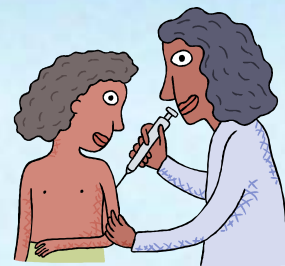
这些人的化验反应是相互独立的. 试说明当  $p$  较小时, 选取适当的  $k$ , 按第二种方法可以减少化验的次数. 并说明  $k$  取什么值时最适宜.

**解** 各人的血呈阴性反应的 概率为  $q = 1 - p$ ,  
因而  $k$  个人的混合血呈阴性反 应的概率为  $q^k$ ,  $k$  个人的混合血呈阳性反应 的概率为  $1 - q^k$ .

设以  $k$  个人为一组时, 组内每人的血化验的次数为  $X$ , 则  $X$  为一随机变量, 其分布律为



| $X$   | $1/k$ | $(k+1)/k$ |
|-------|-------|-----------|
| $p_k$ | $q^k$ | $1 - q^k$ |



$X$  的数学期望为  $E(X) = \frac{1}{k}q^k + (1 + \frac{1}{k})(1 - q^k) = 1 - q^k + \frac{1}{k}$ .

$N$  个人平均需化验的次数为  $N(1 - q^k + \frac{1}{k})$ .

由此可知, 只要选择  $k$  使  $1 - q^k + \frac{1}{k} < 1$ ,

且达到最小值, 这时就能得到最好的分组方法.

例如,  $p = 0.1$ , 则  $q = 0.9$ ,

当  $k = 4$  时,  $L = 1 - q^k + \frac{1}{k}$  取到最小值, 此时得到最

好的分组方法. 若  $N = 1000$ , 此时以  $k = 4$  分组, 则

按第二种方法平均只需化验

$$1000(1 - 0.9^4 + \frac{1}{4}) = 594 \text{ (次)}.$$

这样平均来说, 可以减少 40% 的工作量.

## 四、小结、思考

1. 离散型随机变量的数学期望反映了随机变量取值的概率平均值，其中概率就是对应取值的权重。
2. 当可能取值为无限个值时，其均值不一定存在，若存在，等于无穷级数的收敛值。

### 思考：

1. 连续型随机变量数学期望的计算公式是怎么样的？
2. 数学期望有哪些性质？



谢 谢！