

第四章 数字特征



讨论原因：

(1) 实际中，有的随机变量的概率分布难确定，有的不可能知道，而它的一些**数字特征较易确定**。

(2) 实际应用中，人们更**关心概率分布的数字特征**。

(3) 一些常用的重要分布，如二项分布、泊松分布、指数分布、正态分布等，只要知道了它们的某些数字特征，就能**完全确定其具体的分布**。



本章内容

- 随机变量的平均取值 —— 数学期望
- 随机变量取值平均偏离平均值的情况 —— 方差
- 描述两个随机变量之间的某种关系的数 —— 协方差与相关系数

§ 1 数学期望

- ❖ 离散型随机变量的数学期望
- ❖ 连续型随机变量的数学期望
- ❖ 随机变量函数的数学期望
- ❖ 数学期望的性质 应用

1.离散性随机变量的数学期望

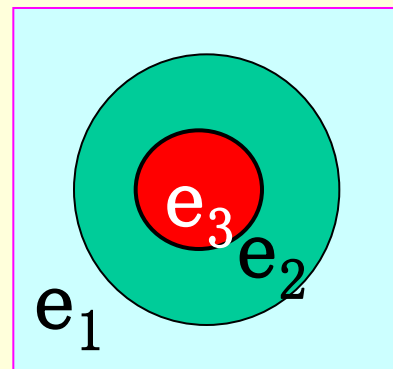
引例：射手打靶练习

射手每次射击得分数 X 是随机变量，射击 N 次，得1分 a_1 次，得2分 a_2 次，得3分 a_3 次。

总得分 $\text{sum}=1*a_1+2*a_2+3*a_3$,

每次平均得分

$$=\text{sum}/N=1*a_1/N+2*a_2/N+3*a_3/N.$$



每次平均得分

$$= \text{sum}/N = 1*a_1/N + 2*a_2/N + 3*a_3/N$$

$$= \sum k * a_k / N.$$

{X=k}的
频率

N充分大

$$\sum k * P\{X=k\}$$

X取值的平均



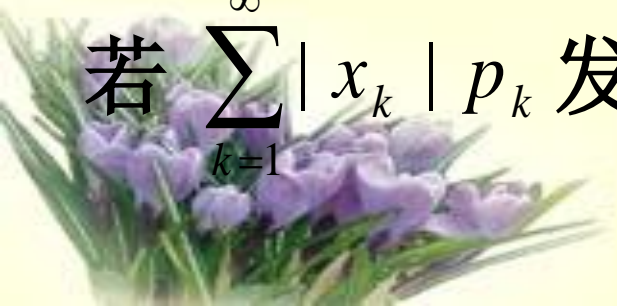
定义 1 设离散型随机变量 X 的分布律为

$$P(X=x_k)=p_k, k=1,2,\dots$$

若级数 $\sum_{k=1}^{\infty} x_k p_k$ **绝对收敛**，则称它为 X 的**数学期望或均值**，记作 $E(X)$ ，

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

若 $\sum_{k=1}^{\infty} |x_k| p_k$ 发散，则称 X 的数学期望不存在。



说明：

- (1) 随机变量的数学期望是一个**实数**，它体现了**随机变量取值的平均**；
- (2) 要注意数学期望存在的条件：**绝对收敛**；
- (3) 当 X 服从某一分布时，也称某**分布的数学期望**为 $E(X)$ 。



例1: 设X服从参数为p的0-1分布, 求E(X)

例2: 设 $X \sim B(n, p)$, 求E(X)

解: $E(X) = \sum_{k=0}^n k P(x = k)$

$$= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k}$$

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

$$n \binom{n-1}{k-1}$$

$$= np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{n-1-(k-1)}$$

$$= np$$



例3: 设X服从参数为 λ 的泊松分布, 求EX。

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

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

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

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



注意：不是所有的随机变量都有数学期望

例如 $P(X = (-1)^k \frac{2^k}{k}) = \frac{1}{2^k}, \quad k = 1, 2, \dots,$

$E(X)$ 不存在



例 4

按规定，火车站每天 8:00~9:00, 9:00~10:00 都恰有一辆客车到站，但到站的时刻是随机的，且两者到站的时间相互独立，其规律为：

到站时间	8:10,9:10	8:30,9:30	8:50,9:50
概率	1/6	3/6	2/6

- (1) 旅客 8:00 到站，求他候车时间的数学期望。
- (2) 旅客 8:20 到站，求他候车时间的数学期望。

解：设旅客的候车时间为 X （以分记）

(1) X 的分布律：

X	10	30	50
P	1/6	3/6	2/6

$$EX=10*(1/6)+30*(3/6)+50*(2/6)=33.33(\text{分})$$



(2) 旅客8: 20分到达

X的分布率为

X	10	30	50	70	90
P	3/6	2/6	$(1/6)*(1/6)$	$(3/6)*(1/6)$	$(2/6)*(1/6)$

$$EX=10*(3/6)+30*(2/6)+50*(1/36)+70*(3/36)+90*(2/36) \\ =27.22(\text{分})$$

到站时间	8:10,9:10	8:30,9:30	8:50,9:50
概率	1/6	3/6	2/6

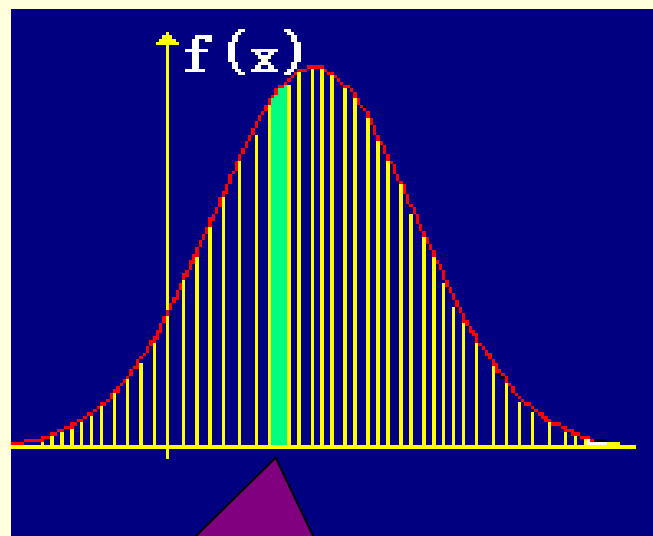
2. 连续型随机变量的数学期望

设 X 是连续型随机变量，其密度函数为 $f(x)$ ，在数轴上取很密的分点 $x_0 < x_1 < x_2 < \dots$ ，则 X 落在小区间 $[x_i, x_{i+1})$ 的概率是

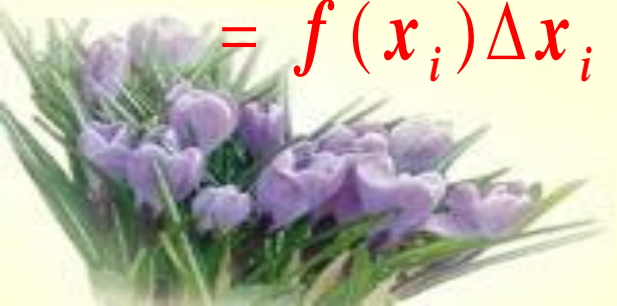
$$\int_{x_i}^{x_{i+1}} f(x) dx$$

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

$$= f(x_i) \Delta x_i$$



小区间 $[x_i, x_{i+1})$



由于 x_i 与 x_{i+1} 很接近, 所以区间 $[x_i, x_{i+1})$ 中的值可以用 x_i 来近似代替.

因此 X 与以概率 $f(x_i)\Delta x_i$ 取值 x_i 的离散型r.v近似, 该离散型r.v的数学期望是

$$\sum_i x_i f(x_i) \Delta x_i$$

这正是 $\int_{-\infty}^{\infty} x f(x) dx$ 的渐近和式.



定义2 设 X 是连续型随机变量, 其密度函数为 $f(x)$, 如果

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

绝对收敛, 定义 X 的数学期望为

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



例5: 设X服从 $U[a, b]$, 求 $E(X)$ 。

例6: 设X服从参数为 λ 的指数分布, 求 $E(X)$

例7: $X \sim N(\mu, \sigma^2)$, 求 $E(X)$ 。

解: $E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$

设 $\frac{x-\mu}{\sigma} = y$

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} (\mu + \sigma y) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \\ &= \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \end{aligned}$$

$$\begin{aligned} &+ \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y \exp\left(-\frac{y^2}{2}\right) dy \\ &= \mu \end{aligned}$$



注意：不是所有的随机变量都有数学期望

例如：Cauchy分布的密度函数为

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

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

它的数学期望不存在



常见随机变量的数学期望

分布	概率分布	期望
参数为 p 的 0-1分布	$P(X = 1) = p$ $P(X = 0) = 1 - p$	p
$B(n, p)$	$P(X = k) = C_n^k p^k (1 - p)^{n-k}$ $k = 0, 1, 2, \dots, n$	np
$\pi(\lambda)$	$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ $k = 0, 1, 2, \dots$	λ



分布	概率密度	期望
区间 (a,b) 上的均匀分布	$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & \text{其它} \end{cases}$	$\frac{a+b}{2}$
$E(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{其它} \end{cases}$	$1/\lambda$
$N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ



3 随机变量函数的数学期望

问题： 求 $Y=g(X)$ 的数学期望

方法一： 根据定义

根据 $g(X)$ 的分布， 求出对应的期望 $E[g(X)]$.



方法二： 直接求

设 $Y = g(X)$, 若离散型随机变量 X 的分布律为

$P(X=x_k)=p_k$, $k=1,2,\dots$, 级数

$$\sum_{k=1}^{\infty} g(x_k) p_k$$

绝对收敛, 则 Y 的数学期望存在, 且为此级数.



例8. 设离散型随机向量 X 的概率分布如下表所示, 求: $Z=X^2$ 的期望.

X	0	-1	1
P	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

解:
$$E(Z) = g(0) \times 0.5 + g(-1) \times 0.25 + g(1) \times 0.25$$
$$= 0.5$$

注: 这里的 $g(x) = x^2$.



设 $Y = g(X)$, 若连续型随机变量 X 的概率密度为 $f(x)$,

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

绝对收敛, 则 Y 的数学期望存在, 且为此积分。



例9. 设 $X \sim U[0, \pi]$, $Y = \sin X$, 求 $E(Y)$ 。

解: X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{\pi}, & 0 \leq x \leq \pi; \\ 0, & \text{其他。} \end{cases}$$

所以

$$EY = E[\sin X]$$

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

$$= \frac{-1}{\pi} \cos x \Big|_0^{\pi} = \frac{2}{\pi}.$$



例10.在随机服务系统中，等待时间可以认为服从指数分布。现已知人们在某银行办理业务的等待时间 X （单位：分钟）服从期望为15的指数分布。某人到银行办理业务，由于办理银行业务后还需要去办另外一件事情，故此人先等待，如果20分钟后没有等到自己办理业务就离开银行。设此人在银行的实际等待时间为 Y ，求此人实际等待时间的平均值。

解：易知 Y 和 X 的关系为 $Y=\min(X,20)$ 。

X 的密度函数为

$$f(x) = \begin{cases} \frac{1}{15} e^{-\frac{1}{15}x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$



$$f(x) = \begin{cases} \frac{1}{15} e^{-\frac{1}{15}x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad Y = \min(X, 20).$$

由定理2知

$$\begin{aligned} E(Y) &= E(\min(X, 20)) = \int_{-\infty}^{+\infty} \min(x, 20) f(x) dx = \int_0^{+\infty} \min(x, 20) \frac{1}{15} e^{-\frac{1}{15}x} dx \\ &= \int_0^{20} x \frac{1}{15} e^{-\frac{1}{15}x} dx + \int_{20}^{+\infty} 20 \times \frac{1}{15} e^{-\frac{1}{15}x} dx = 15(1 - e^{-\frac{4}{3}}) \approx 11.04 \end{aligned}$$

注意：可以证明实际等待时间 Y 不是连续型随机变量，从而 Y 没有密度函数，故不能先求分布函数，然后用分布函数的导数的积分的方法求 EY 。



设二维随机变量 (X, Y) 的函数 $Z=g(X, Y)$, 若二维离散型随机变量 (X, Y) 的联合分布律为

$P(X=x_i, Y=y_j)=P_{ij}$, $i, j=1, 2, \dots$ 且有级数

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

绝对收敛, 则 Z 的数学期望存在, 且为此级数.



注意:

$$E(X) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_i p_{ij}$$

例11. 设二维离散型随机向量 (X, Y) 的概率分布如下表所示, 求: $Z=X^2+Y$ 的期望.

X \ Y	1	2
1	1/8	1/4
2	1/2	1/8

解:
$$E(Z) = g(1, 1) \times 0.125 + g(1, 2) \times 0.25 \\ + g(2, 1) \times 0.5 + g(2, 2) \times 0.125 = 4.25$$

注: 这里的

$$g(x, y) = x^2 + y.$$



若二维连续型随机变量 (X, Y) 的联合概率密度为 $f(x, y)$ ，并且

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy$$

绝对收敛，则Z的数学期望存在，且为此积分

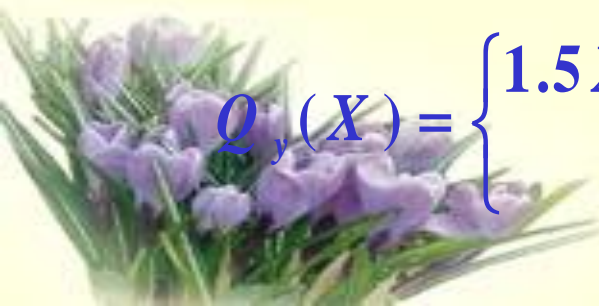
注意： $E(X) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dx dy$



例12 某水果商店，冬季每周购进一批苹果。
已知该店一周苹果销售量 X (单位:kg)服从
 $U[1000,2000]$ 。购进的苹果在一周内售出，
1kg获纯利1.5元；一周内没售出，1kg需付
耗损、储藏等费用0.3元。问一周应购进多
少千克苹果，商店才能获得最大的平均利润。

设购进 y (kg)苹果，获利数为 $Q_y(X)$

$$Q_y(X) = \begin{cases} 1.5X - 0.3(y - X) & y > X \\ 1.5y & y \leq X \end{cases}$$



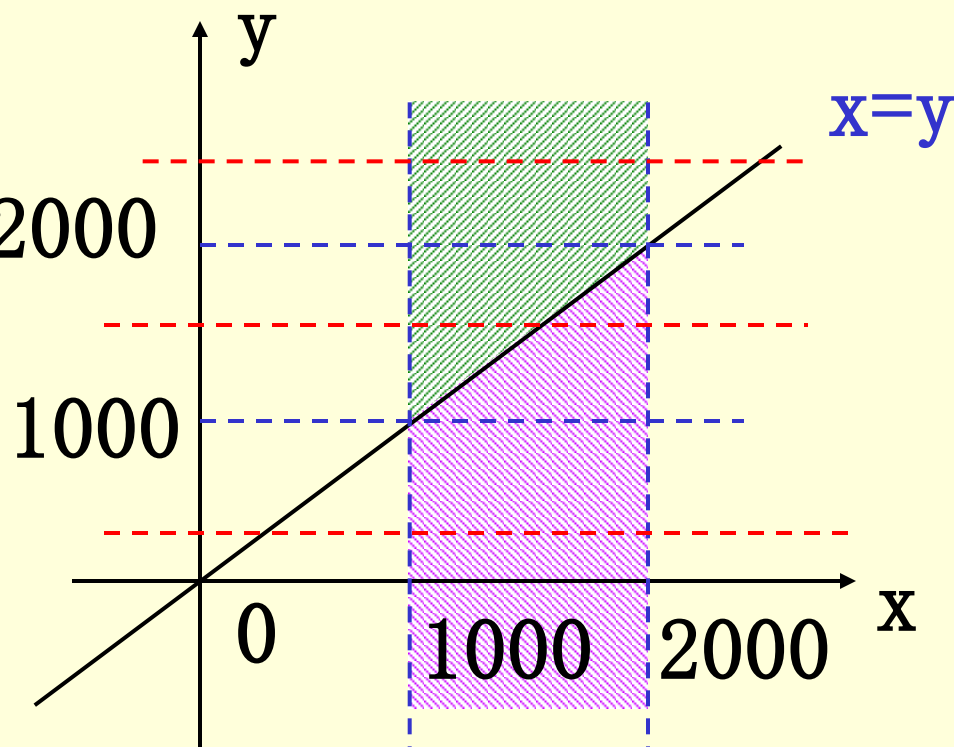
解 设购进 y (kg)苹果, 获利数为 $Q_y(X)$,

$$Q_y(X) = \begin{cases} 1.5X - 0.3(y - X) & y > X \\ 1.5y & y \leq X \end{cases} \quad Q_y(x) = \begin{cases} 1.8x - 0.3y & y > x \\ 1.5y & y \leq x \end{cases}$$

$$EQ_y(X) = \int_{-\infty}^{\infty} Q_y(x) f(x) dx$$

$Q_y(x)f(x)$

$$= \begin{cases} \frac{1}{1000} [1.8x - 0.3y] & y > x \quad 1000 < x < 2000 \\ \frac{1.5y}{1000} & y \leq x \quad 1000 < x < 2000 \\ 0 & \text{其他} \end{cases}$$



$$= \begin{cases} \int_{1000}^{2000} 1.5y \cdot \frac{1}{1000} dx = 1.5y & y \leq 1000 \\ \int_{1000}^y [1.5x - 0.3(y-x)] \frac{1}{1000} dx + \int_y^{2000} 1.5y \cdot \frac{1}{1000} dx \\ = \frac{-0.9y^2 + 3300y - 900000}{1000} & 1000 < y < 2000 \\ \int_{1000}^{2000} [1.5x - 0.3(y-x)] \frac{1}{1000} dx = 2700 - 0.3y & y \geq 2000 \end{cases}$$

$$y = -\frac{3300}{2 \cdot (-0.9)} \approx 1833$$



例13 设二维连续随机变量 (X, Y) 的密度函数为

$$f(x, y) = \begin{cases} \frac{1}{4}x(1 + 3y^2), & 0 < x < 2, 0 < y < 1, \\ 0, & \text{其它} \end{cases}$$

求 $E(X)$, $E(Y)$, $E(X + Y)$, $E(XY)$

解 $E(X) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xf(x, y)dx dy$

$$= \int_0^2 x \cdot \frac{1}{4} x dx \int_0^1 (1 + 3y^2) dy = \frac{4}{3}$$

$$E(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} yf(x, y)dx dy$$

$$= \int_0^2 \frac{1}{4} x dx \int_0^1 y(1 + 3y^2) dy = \frac{5}{8}$$



$$\begin{aligned}
 E(X + Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + y) f(x, y) dx dy \\
 &= \int_0^2 \left[\int_0^1 (x + y) \frac{1}{4} x(1 + 3y^2) dy \right] dx \\
 &= \int_0^2 \left[\int_0^1 x \cdot \frac{1}{4} x(1 + 3y^2) dy \right] dx \\
 &\quad + \int_0^2 \left[\int_0^1 y \cdot \frac{1}{4} x(1 + 3y^2) dy \right] dx \\
 &= \frac{4}{3} + \frac{5}{8} = \frac{47}{24} \\
 &= E(X) + E(Y)
 \end{aligned}$$



$$E(XY) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (xy) f(x, y) dx dy$$

$$= \int_0^2 x \cdot \frac{1}{2} x dx \int_0^1 y \cdot \frac{1}{2} (1 + 3y^2) dy$$

$$= \frac{4}{3} \cdot \frac{5}{8} = \frac{5}{6}$$

$$= E(X) \cdot E(Y)$$



第四章

作业1： 2,8,12,18,21,26





4 数学期望的性质

(1) 设C为常数，则有 $E(C)=C$

(2) 设C为常数，X为随机变量，则有

$$E(CX)=CE(X)$$

(3) 设X, Y为任意两个随机变量，则有

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

推广

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n).$$

$$E(a_1 X_1 + a_2 X_2 + \dots + a_n X_n)$$

$$= a_1 E(X_1) + a_2 E(X_2) + \dots + a_n E(X_n).$$



(4) 设 X, Y 为相互独立的随机变量，则有

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

推广

设 X_1, \dots, X_n 为相互独立的随机变量，则有

$$E(X_1 X_2 \dots X_n) = E(X_1)E(X_2) \dots E(X_n).$$



(5) 若随机变量几乎处处只取非负值，即

$$X \geq 0, \text{ a.e.}$$

又 $E(X)$ 存在，则 $E(X) \geq 0$ 。

推论：若 $X \leq Y$, a.e., $E(X)$, $E(Y)$ 都存在，则

$$E(X) \leq E(Y)$$

特别地，若 $a \leq X \leq b$, a.e., a , b 为常数，则 $E(X)$ 存在，且

$$a \leq E(X) \leq b$$



注 性质 4 的逆命题不成立, 即

若 $E(XY) = E(X)E(Y)$, X, Y 不一定相互独立

反例

p_{ij} $Y \backslash X$	-1	0	1	$p_{\cdot j}$
-1	1/8	1/8	1/8	3/8
0	1/8	0	1/8	2/8
1	1/8	1/8	1/8	3/8
$p_{i \cdot}$	3/8	2/8	3/8	

$X \ Y$	-1	0	1
P	$2/8$	$4/8$	$2/8$

$$E(X) = E(Y) = 0; \quad E(XY) = 0;$$

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

但 $P(X = -1, Y = -1) = \frac{1}{8}$

$$\neq P(X = -1)P(Y = -1) = \left(\frac{3}{8}\right)^2$$



5 数学期望性质的应用

例11 设X的概率密度为

$$f(x) = \begin{cases} a + bx^2 & 0 \leq x \leq 1 \\ 0 & \text{其它} \end{cases}$$

其中a, b为常数, 且 $E(X) = 3/5$ 。求a, b的值。



例14 求二项分布的数学期望

若 $X \sim B(n, p)$,

则 X 表示 n 重贝努里试验中的“成功”次数.

若设

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



则 $X = X_1 + X_2 + \dots + X_n$

因为 $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p$

$$E(X_i) = 1 \cdot p + 0 \cdot (1 - p) = p$$

所以 $E(X) = \sum_{i=1}^n E(X_i) = np$



例 15

一民航送客载有 20 位旅客自机场开出，旅客有 10 个车站可以下车，如到达一个车站没有旅客下车就不停车。以 X 表示停车的次数。

求 EX （设每个旅客在各个车站下车是等可能的，并设各旅客是否下车相互独立）。

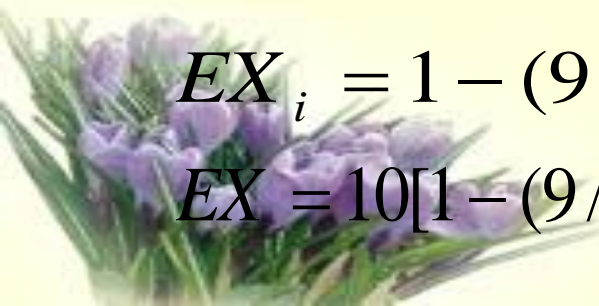
解： 设 $X_i = \begin{cases} 0, & \text{第} i \text{站没人下车} \\ 1, & \text{第} i \text{站有人下车} \end{cases}, \quad i = 1, 2, \dots, 10,$

易见 $X = X_1 + \dots + X_{10}, \quad EX = \sum_{i=1}^{10} EX_i,$

$P\{X_i = 0\} = (9/10)^{20}, \quad P\{X_i = 1\} = 1 - (9/10)^{20}, \quad i = 1, \dots, 10,$

$EX_i = 1 - (9/10)^{20}, \quad i = 1, \dots, 10,$

$EX = 10[1 - (9/10)^{20}] = 8.784(\text{次})。$



例16 设一批同类型的产品共有 N 件，其中次品有 M 件。今从中任取 n (假定 $n \leq N-M$)件，记这 n 件中所含的次品数为 X ，求 $E(X)$

解：若设

$$X_i = \begin{cases} 1 & \text{如第 } i \text{ 次抽到次品} \\ 0 & \text{如第 } i \text{ 次抽到合格品} \end{cases} \quad i=1, 2, \dots, n$$

则 $X = X_1 + X_2 + \dots + X_n$

$$P(X_i = 1) = \frac{M}{N} \quad P(X_i = 0) = 1 - \frac{M}{N}$$

$$E(X_i) = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = \frac{M}{N}$$

$$\text{所以 } E(X) = \sum_{i=1}^n E(X_i) = \frac{nM}{N}$$



例17. 验血方案的选择

为普查某种疾病， n 个人需验血。有如下两种验血方案：

- (1) 分别化验每个人的血，共需化验 n 次；
- (2) 分组化验。每 k 个人分为1组， k 个人的血混在一起化验，若结果为阴性，则只需化验一次；若为阳性，则对 k 个人的血逐个化验，找出有病者，此时 k 个人的血需化验 $k+1$ 次。

设：每个人血液化验呈阳性的概率为 p ，且每个人化验结果是相互独立的。试说明选择哪一方案较经济。



解：只需计算方案（2）所需化验次数 X 的期望。

为简单计，不妨设 n 是 k 的倍数，共分成 $j=n/k$ 组。

设第 i 组需化验的次数为 X_i ，则其分布律为

X_i	1	$k+1$
P	$(1-p)^k$	$1 - (1-p)^k$

$$E(X_i) = 1 \times (1-p)^k + (k+1) \times [1 - (1-p)^k] = (k+1) - k(1-p)^k$$



方案2需要化验的总次数为 $X = X_1 + X_2 + \cdots + X_j$

$$E(X) = \sum_{i=1}^j E(X_i) = \frac{n}{k} [(k+1) - k(1-p)^k] = n[1 - ((1-p)^k - \frac{1}{k})]$$

若 $(1-p)^k - \frac{1}{k} > 0$, 则 $E(X) < n$, 即方案2优于方案1

如: $n=1000$, $p=0.001$, $k=10$

$$E(X) = 1000[1 - (0.999^{10} - \frac{1}{10})] \approx 110 \ll 1000.$$

