概率论与数理统计

第四章 · 数字特征

2020 年 2 月 12 日

■ 暨南大学数学系 ■ 吕荐瑞

主要内容

数学期望 随机变量的平均取值

方差 随机变量取值偏离平均值的程度

协方差 两个随机变量的总体偏差

相关系数 两个随机变量之间的线性相关程度

第一节 期望

第二节 方差

第三节 协方差与相关系数

第一节	期望
A	离散型随机变量的期望
B	连续型随机变量的期望
C	随机变量函数的期望
D	数学期望的性质

离散型随机变量的期望

概念引入：某服装公司生产两种套装，一种是大众装，每件价格 200 元，每月生产 1 万件；另一种是高档装，每件 1800 元，每月生产 100 件．现在问该公司生产的套装平均价格是多少？

离散型随机变量的期望

概念引入：某服装公司生产两种套装，一种是大众装，每件价格 200 元，每月生产 1 万件；另一种是高档装，每件 1800 元，每月生产 100 件．现在问该公司生产的套装平均价格是多少？.....216

离散型随机变量的期望

定义 1 设离散型随机变量 X 的分布律为

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

若级数

$$\sum_k x_k p_k$$

绝对收敛，则称其和为随机变量 X 的数学期望，记为 $E(X)$.

离散型随机变量的期望

例 2　设甲、乙两射手在各次射击中得分 X 和 Y 的分布律如下：

X	1	2	3	4
P	0.1	0.2	0.5	0.2

Y	1	2	3	4
P	0.2	0.4	0.1	0.3

试比较甲、乙两个射手的技术.

离散型随机变量的期望

两点分布	$X \sim B(1,p)$	$E(X)=p$
二项分布	$X \sim B(n,p)$	$E(X)=np$
泊松分布	$X \sim P(\lambda)$	$E(X)=\lambda$

离散型随机变量的期望

例 3　某种产品次品率为 0.1．检验员每天检验 4 次，每次随机抽取 10 件产品进行检验，如发现次品数大于 1，就调整设备．若各件产品是否为次品是相互独立的，求一天中调整设备次数的期望．

彩票派奖问题

例子　美国波士顿的“Cash WinFall”彩票，每注价格为 2 美元，从 1 – 46 中选择 6 个不重复号码．在连续多期无人中头奖时的派奖规则如下：

- 2 个号码和开奖号码相同，奖金 2 美元.
- 3 个号码和开奖号码相同，奖金 50 美元.
- 4 个号码和开奖号码相同，奖金 1500 美元.
- 5 个号码和开奖号码相同，奖金 40000 美元.
- 6 个号码和开奖号码相同，奖金 200 万美元.

求购买每注彩票所得奖金 X 的数学期望 $E(X)$.

彩票号码与开奖号码相同的个数为 k 的概率等于

$$p_k = \frac{C_6^k C_{40}^{6-k}}{C_{46}^6}, \quad k = 0, 1, \cdots, 6.$$

彩票号码与开奖号码相同的个数为 k 的概率等于

$$p_k = \frac{C_6^k C_{40}^{6-k}}{C_{46}^6}, \quad k = 0, 1, \cdots, 6.$$

k	X	P
0 or 1	0	0.831
2	2	0.146
3	50	0.0211
4	1500	0.001249
5	40000	0.00002562
6	2000000	0.0000001068

k	X	P
0 or 1	0	0.831
2	2	0.146
3	50	0.0211
4	1500	0.001249
5	40000	0.00002562
6	2000000	0.0000001068

k	X	P
0 or 1	0	0.831
2	2	0.146
3	50	0.0211
4	1500	0.001249
5	40000	0.00002562
6	2000000	0.0000001068

每注彩票所得奖金 X 的数学期望

$$E(X) = 2p_2 + 50p_3 + 1500p_4 + 40000p_5 + 2000000p_6$$

$$= 0.292 + 1.055 + 1.8735 + 1.0248 + 0.2134$$

$$= 4.4587.$$

第一节	期望
A	离散型随机变量的期望
B	连续型随机变量的期望
C	随机变量函数的期望
D	数学期望的性质

连续型随机变量的期望

定义 2 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，则称此积分为随机变量 X 的数学期望，记为 $E(X)$.

连续型随机变量的期望

例 4　设随机变量 X 的概率密度为

$$f(x) = \frac{1}{2}\mathrm{e}^{-|x|}, \qquad x \in \mathbb{R},$$

求 X 的数学期望 $E(X)$.

连续型随机变量的期望

均匀分布	$X \sim U[a,b]$	$E(X)=\dfrac{a+b}{2}$
指数分布	$X \sim EP(\lambda)$	$E(X)=\dfrac{1}{\lambda}$
正态分布	$X \sim N(\mu,\sigma^2)$	$E(X)=\mu$

连续型随机变量的期望

例 5　设某型号电子管的寿命 X 服从指数分布，其平均寿命为 1000 小时，计算 $P\{1000 < X \leqslant 1200\}$.

连续型随机变量的期望

练习 2　已知随机变量 X 的概率密度为

$$f(x) = \begin{cases} 2x, & 0 < x < 1; \\ 0, & \text{otherwise}. \end{cases}$$

求 EX.

连续型随机变量的期望

练习 2 已知随机变量 X 的概率密度为

$$f(x)=\begin{cases} 2x, & 0<x<1; \\ 0, & \text{otherwise}. \end{cases}$$

求 EX.

练习 3 已知随机变量 X 的概率密度为

$$f(x)=\begin{cases} a+bx^2, & 0<x<1; \\ 0, & \text{otherwise}. \end{cases}$$

且 $EX=\frac{3}{5}$，求未知系数 a, b.

第一节	期望
A	离散型随机变量的期望
B	连续型随机变量的期望
C	随机变量函数的期望
D	数学期望的性质

随机变量函数的数学期望

问题：设随机变量 X 的分布已知，在实际问题中有时需要计算的量并非 X 的期望，而是 X 的某个函数 $Y = g(X)$ 的期望．如何根据 X 的分布计算 EY？

随机变量函数的数学期望

问题：设随机变量 X 的分布已知，在实际问题中有时需要计算的量并非 X 的期望，而是 X 的某个函数 $Y = g(X)$ 的期望．如何根据 X 的分布计算 EY？

直观思路：根据 X 的分布算出 Y 的分布，然后利用定义计算 EY．但求 Y 的分布的计算一般很麻烦．

随机变量函数的数学期望

定理 设 X 为随机变量，$Y = g(X)$，则

1. 若 X 为离散型随机变量，分布律为

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

则

$$E(Y) = \sum_k g(x_k) p_k.$$

随机变量函数的数学期望

定理 设 X 为随机变量，$Y = g(X)$，则

1. 若 X 为离散型随机变量，分布律为

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

则

$$E(Y) = \sum_k g(x_k) p_k.$$

2. 若 X 为连续型随机变量，概率密度为 $f(x)$，则

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

随机变量函数的数学期望

例 6　设随机变量 X 的概率分布为

X	1	2	3
P	0.2	0.1	0.7

求随机变量函数 $Y = X^2$ 的期望.

随机变量函数的数学期望

例 6　设随机变量 X 的概率分布为

X	1	2	3
P	0.2	0.1	0.7

求随机变量函数 $Y = X^2$ 的期望.

例 7　设随机变量 X 在区间 $[0, \pi]$ 上服从均匀分布, 求随机变量函数 $Y = \sin X$ 的数学期望.

随机变量函数的数学期望

例 8　设随机变量 $X \sim N(0,1)$，求 $E(X^2)$.

随机向量函数的数学期望

定理 1 设 (X,Y) 为随机向量，$Z=g(X,Y)$，则

1. 若 (X,Y) 为离散型随机向量，概率分布为

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

则

$$EZ=\sum_i\sum_j g(x_i,y_j)p_{ij}.$$

随机向量函数的数学期望

定理 1 设 (X,Y) 为随机向量，$Z=g(X,Y)$，则

1. 若 (X,Y) 为离散型随机向量，概率分布为
$$P\{X=x_i,Y=y_j\}=p_{ij},\quad i,j=1,2,\cdots.$$
则
$$EZ=\sum_i\sum_j g(x_i,y_j)p_{ij}.$$

2. 若 (X,Y) 为连续型随机向量，概率密度为 $f(x,y)$，则
$$EZ=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

随机向量函数的数学期望

例 9　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.

$X \backslash Y$	1	2
1	$\frac{1}{8}$	$\frac{1}{4}$
2	$\frac{1}{2}$	$\frac{1}{8}$

随机向量函数的数学期望

练习 4　假设 (X,Y) 的联合分布为

(X,Y)	$(0,0)$	$(0,2)$	$(1,1)$	$(1,2)$	$(2,1)$
P	0.15	0.25	0.1	0.2	0.3

求 $2X-Y$ 的期望.

随机向量函数的数学期望

例 10　设 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases} x+y, & 0\leqslant x\leqslant 1, 0\leqslant y\leqslant 1 \\ 0, & \text{otherwise} \end{cases}$$

求随机变量函数 $Z=XY$ 的数学期望.

第一节	期望
A	离散型随机变量的期望
B	连续型随机变量的期望
C	随机变量函数的期望
D	数学期望的性质

数学期望的性质

设 $X, X_1, X_2, \cdots, X_n$ 为随机变量，c, k 为常数，则有

1. $E(c) = c$;
2. $E(kX) = kE(X)$;
3. $E(X_1 + X_2) = E(X_1) + E(X_2)$;

数学期望的性质

设 $X, X_1, X_2, \cdots, X_n$ 为随机变量，c, k 为常数，则有

1. $E(c) = c$;
2. $E(kX) = kE(X)$;
3. $E(X_1 + X_2) = E(X_1) + E(X_2)$;

推论：$E\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} E(X_k)$.

数学期望的性质

例子　设在某试验中事件 A 的概率为 p，将该试验独立地进行 n 次．记 X 为 n 次试验中事件 A 发生的总次数，X_i 为第 i 次试验中事件 A 发生的次数，则 $X \sim B(n,p)$，$X_i \sim B(1,p), i = 1,2,\cdots,n$，且

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

故

$$EX = EX_1 + EX_2 + \cdots + EX_n = np.$$

例 11　将 n 个球放入 M 个盒子中，设每个球落入各个盒子是等可能的，求有球的盒子数 X 的期望.

例 11 将 n 个球放入 M 个盒子中，设每个球落入各个盒子是等可能的，求有球的盒子数 X 的期望.

解答 令 $X_i = \begin{cases} 1, & \text{第 } i \text{ 个盒子中有球;} \\ 0, & \text{第 } i \text{ 个盒子中无球.} \end{cases}$

例 11　将 n 个球放入 M 个盒子中，设每个球落入各个盒子是等可能的，求有球的盒子数 X 的期望.

解答　令 $X_i = \begin{cases} 1, & \text{第 } i \text{ 个盒子中有球;} \\ 0, & \text{第 } i \text{ 个盒子中无球.} \end{cases}$ 则有 $X = X_1 + \cdots + X_M$.

例 11 将 n 个球放入 M 个盒子中，设每个球落入各个盒子是等可能的，求有球的盒子数 X 的期望.

解答 令 $X_i = \begin{cases} 1, & \text{第 } i \text{ 个盒子中有球;} \\ 0, & \text{第 } i \text{ 个盒子中无球.} \end{cases}$ 则有 $X = X_1 + \cdots + X_M$. 于是 $E(X) = E(X_1) + \cdots + E(X_M)$.

例 11 将 n 个球放入 M 个盒子中，设每个球落入各个盒子是等可能的，求有球的盒子数 X 的期望.

解答 令 $X_i = \begin{cases} 1, & \text{第 } i \text{ 个盒子中有球;} \\ 0, & \text{第 } i \text{ 个盒子中无球.} \end{cases}$ 则有 $X = X_1 + \cdots + X_M$. 于是 $E(X) = E(X_1) + \cdots + E(X_M)$.

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

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

例 11 将 n 个球放入 M 个盒子中，设每个球落入各个盒子是等可能的，求有球的盒子数 X 的期望.

解答 令 $X_i = \begin{cases} 1, & \text{第 } i \text{ 个盒子中有球;} \\ 0, & \text{第 } i \text{ 个盒子中无球.} \end{cases}$ 则有 $X = X_1 + \cdots + X_M$. 于是 $E(X) = E(X_1) + \cdots + E(X_M)$.

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

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

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

数学期望的性质

练习 5　从一副标准扑克牌（52 张，不包含大小王）中任意抽取 10 张，求其中红桃张数的期望、脸牌（J,Q,K）张数的期望.（提示：利用抽签的公平性）

数学期望的性质

4 若 X_1, X_2 相互独立，则有

$$E(X_1X_2) = E(X_1)E(X_2).$$

数学期望的性质

4 若 X_1, X_2 相互独立，则有

$$E(X_1X_2) = E(X_1)E(X_2).$$

推论：若 $X_1, X_2, \cdots, X_n$ 相互独立，则有

$$E\left(\prod_{k=1}^{n} X_k\right) = \prod_{k=1}^{n} E(X_k)$$

数学期望的性质

4 若 X_1, X_2 相互独立，则有

$$E(X_1X_2) = E(X_1)E(X_2).$$

推论：若 $X_1, X_2, \cdots, X_n$ 相互独立，则有

$$E\left(\prod_{k=1}^{n} X_k\right) = \prod_{k=1}^{n} E(X_k)$$

注：如果没有相互独立这一条件，上式一般不成立！

随机变量函数的数学期望

例 12　设随机变量 X 和 Y 相互独立，且概率密度分别为

$$f_X(x) = \begin{cases} 4e^{-4x}, & x \geq 0 \\ 0, & \text{otherwise;} \end{cases}$$

$$f_Y(y) = \begin{cases} 2e^{-2y}, & y \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

求 $E(X+Y)$ 和 $E(XY)$.

随机变量函数的数学期望

练习 6　设二维离散型随机向量 (X,Y) 的概率分布如下表所示：

$X \backslash Y$	1	2
1	$\frac{1}{9}$	$\frac{2}{9}$
2	$\frac{2}{9}$	$\frac{4}{9}$

求 $E(X+Y)$ 和 $E(XY)$.

复习与提高

选择 (2011)　设随机变量 X 和 Y 相互独立，且 $E(X)$ 与 $E(Y)$ 存在，记 $U=\max\{X,Y\}$，$V=\min\{X,Y\}$，则 $E(UV)=$ ……（　　）

(A) $E(U)E(V)$　　(B) $E(X)E(Y)$

(C) $E(U)E(Y)$　　(D) $E(X)E(V)$

第一节 期望

第二节 方差

第三节 协方差与相关系数

第二节	方差
A	方差的定义
B	方差的性质
C	常用随机变量的方差

方差的概念

例 1　设甲、乙两射手在各次射击中得分 X 和 Y 的分布律如下：

X	1	2	3	4
P	0.1	0.2	0.5	0.2

Y	1	2	3	4
P	0.2	0.1	0.4	0.3

试比较甲、乙两个射手的技术.

方差的概念

在实际问题中，仅靠期望值不能完善地说明随机变量的分布特征．我们常常需要知道分布相对于期望值的离散程度．

方差的概念

在实际问题中，仅靠期望值不能完善地说明随机变量的分布特征．我们常常需要知道分布相对于期望值的离散程度．

定义 1　设 X 是一随机变量，若 $[X-E(X)]^2$ 的期望存在，则称该期望为 X 的方差(Variance)，记为 $Var(X)$（或 $D(X)$），即

$$\mathrm{Var}(X) := E[X-E(X)]^2.$$

称 $\sqrt{\mathrm{Var}(X)}$ 为 X 的标准差(Standard deviation)，记为 $\sigma(X)$.

方差

方差刻划了随机变量的取值相对于其数学期望的偏离程度.

1. 若 X 的取值比较分散，则方差较大；
2. 若 X 的取值比较集中，则方差较小；

特别地，$\mathrm{Var}(X) = 0$ 当且仅当 X 取某个常数的概率为 1.

方差的常用计算公式：$\mathrm{Var}(X) = E(X^2) - [E(X)]^2$.

方差的常用计算公式：$\mathrm{Var}(X) = E(X^2) - [E(X)]^2$.

例 2　已知随机变量 X 的概率分布律为

X	0	1	2
P	0.2	0.5	0.3

求 $\mathrm{Var}(X)$.

方差的常用计算公式：$\mathrm{Var}(X) = E(X^2) - [E(X)]^2$.

例 2　已知随机变量 X 的概率分布律为

X	0	1	2
P	0.2	0.5	0.3

求 $\mathrm{Var}(X)$.

例 3　设连续型随机变量 X 的密度函数为

$$f(x) = \begin{cases} 2x, & 0 \leqslant x \leqslant 1 \\ 0, & \text{otherwise} \end{cases},$$

求 $\mathrm{Var}(X)$.

练习 1　已知随机变量 X 的概率分布律为

X	0	1	2	3
P	0.4	0.3	0.2	0.1

求 $\mathrm{Var}(X)$.

练习 1 已知随机变量 X 的概率分布律为

X	0	1	2	3
P	0.4	0.3	0.2	0.1

求 $\mathrm{Var}(X)$.

练习 2 设连续型随机变量 X 的密度函数为

$$f(x) = \begin{cases} |x|, & -1 \leqslant x \leqslant 1 \\ 0, & \text{otherwise} \end{cases},$$

求 $\mathrm{Var}(X)$.

第二节	方差
A	方差的定义
B	方差的性质
C	常用随机变量的方差

方差的性质

设 X、Y 为随机变量，c、k 为常数，则有

1. $\mathrm{Var}(c) = 0$

方差的性质

设 X、Y 为随机变量，c、k 为常数，则有

1. $\mathrm{Var}(c) = 0$
2. $\mathrm{Var}(X + c) = \mathrm{Var}(X)$

方差的性质

设 X、Y 为随机变量，c、k 为常数，则有

1. $\mathrm{Var}(c) = 0$
2. $\mathrm{Var}(X + c) = \mathrm{Var}(X)$
3. $\mathrm{Var}(kX) = k^2\,\mathrm{Var}(X)$

方差的性质

设 X、Y 为随机变量，c、k 为常数，则有

1. $\mathrm{Var}(c) = 0$
2. $\mathrm{Var}(X + c) = \mathrm{Var}(X)$
3. $\mathrm{Var}(kX) = k^2\,\mathrm{Var}(X) \implies \mathrm{Var}(-X) = \mathrm{Var}(X)$

方差的性质

设 X、Y 为随机变量，c、k 为常数，则有

1. $\mathrm{Var}(c) = 0$
2. $\mathrm{Var}(X + c) = \mathrm{Var}(X)$
3. $\mathrm{Var}(kX) = k^2\,\mathrm{Var}(X) \implies \mathrm{Var}(-X) = \mathrm{Var}(X)$
4. 若 X 和 Y 相互独立，则有

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

方差的性质

设 X、Y 为随机变量，c、k 为常数，则有

1. $\mathrm{Var}(c) = 0$
2. $\mathrm{Var}(X + c) = \mathrm{Var}(X)$
3. $\mathrm{Var}(kX) = k^2\,\mathrm{Var}(X) \implies \mathrm{Var}(-X) = \mathrm{Var}(X)$
4. 若 X 和 Y 相互独立，则有

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

注记　性质 4 可以推广到多个相互独立的随机变量.

方差的性质

例 4　连续掷两次骰子，用随机变量 X 表示两次的点数之和，求 $\mathrm{Var}(X)$.

例 5　设随机变量 X 的期望和方差分别为 $E(X)$ 和 $\mathrm{Var}(X)$，且 $\mathrm{Var}(X) > 0$，求

$$X^* = \frac{X - E(X)}{\sqrt{\mathrm{Var}(X)}}$$

的期望和方差.

例 5　设随机变量 X 的期望和方差分别为 $E(X)$ 和 $\mathrm{Var}(X)$，且 $\mathrm{Var}(X) > 0$，求

$$X^* = \frac{X - E(X)}{\sqrt{\mathrm{Var}(X)}}$$

的期望和方差.

通常将 $X^* = \dfrac{X - E(X)}{\sqrt{\mathrm{Var}(X)}}$ 称为 X 的标准化的随机变量.

第二节	方差
A	方差的定义
B	方差的性质
C	常用随机变量的方差

常用随机变量的方差

随机变量	X	$E(X)$	$\mathrm{Var}(X)$
两点分布	$B(1,p)$	p	$p(1-p)$
二项分布	$B(n,p)$	np	$np(1-p)$
泊松分布	$P(\lambda)$	λ	λ
均匀分布	$U[a,b]$	$(a+b)/2$	$(b-a)^2/12$
指数分布	$EP(\lambda)$	$1/\lambda$	$1/\lambda^2$
正态分布	$N(\mu,\sigma^2)$	μ	σ^2

常用随机变量的方差

例 6　一台设备由三个部件构成，在设备运转中各部件需要调整的概率分别为 0.01, 0.02, 0.03. 设各部件的状态相互独立，用 X 表示同时需要调整的部件数，求 X 的期望和方差.

常用随机变量的方差

练习 3　已知随机变量 X 服从二项分布 $B(n,p)$,

$$EX = 12, \qquad \mathrm{Var}(X) = 8$$

求 n 和 p.

复习与提高

复习 1　设 (X,Y) 的联合概率密度为

$$f(x,y)=\begin{cases} 3x, & x\in[0,1], y\in[0,x]; \\ 0, & \text{otherwise.} \end{cases}$$

求 $E(Y)$ 和 $\mathrm{Var}(Y)$.

复习与提高

复习 2　假设随机变量 X 的概率密度为

$$f(x)=\begin{cases} ax^2+bx+c, & x\in(0,1); \\ 0, & \text{otherwise.} \end{cases}$$

已知 $EX=0.5, \mathrm{Var}(X)=0.15$，求 a,b,c.

复习与提高

复习 3　设 (X,Y) 的联合概率分布如下表所示：

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

求 $\mathrm{Var}(X), \mathrm{Var}(Y)$ 和 $\mathrm{Var}(X+Y)$.

第一节 期望

第二节 方差

第三节 协方差与相关系数

第三节	协方差与相关系数
A	协方差
B	相关系数

协方差

对于二维随机向量 (X, Y)，除了其分量 X 和 Y 的期望与方差外，还有一些数字特征，用以刻画 X 与 Y 之间的相关程度，其中最主要的就是下面要讨论的协方差和相关系数.

协方差

定义 1 定义：对于二维随机向量 (X,Y)，称

$$\mathrm{Cov}(X,Y) := E[(X-EX)(Y-EY)]$$

为 X 与 Y 的协方差(Covariance).

协方差

定义 1 定义：对于二维随机向量 (X,Y)，称

$$\mathrm{Cov}(X,Y) := E[(X-EX)(Y-EY)]$$

为 X 与 Y 的协方差(Covariance).

由定义直接可得：任意随机变量与其自身的协方差就是该随机变量的方差，即

$$\mathrm{Cov}(X,X) = \mathrm{Var}(X).$$

协方差的性质

设 X, Y, Z 为随机变量，a, b, c, d 为常数，则有

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$；

协方差的性质

设 X, Y, Z 为随机变量，a, b, c, d 为常数，则有

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$；
2. $\mathrm{Cov}(aX + b, cY + d) = ac\,\mathrm{Cov}(X, Y)$；

协方差的性质

设 X, Y, Z 为随机变量，a, b, c, d 为常数，则有

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$；
2. $\mathrm{Cov}(aX + b, cY + d) = ac\,\mathrm{Cov}(X, Y)$；
3. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$；

协方差的性质

设 X, Y, Z 为随机变量，a, b, c, d 为常数，则有

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$；
2. $\mathrm{Cov}(aX + b, cY + d) = ac\,\mathrm{Cov}(X, Y)$；
3. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$；
4. $\mathrm{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$；

协方差的性质

设 X, Y, Z 为随机变量，a, b, c, d 为常数，则有

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;
2. $\mathrm{Cov}(aX + b, cY + d) = ac\,\mathrm{Cov}(X, Y)$;
3. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$;
4. $\mathrm{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$;
5. $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y)$.

协方差的性质

设 X, Y, Z 为随机变量，a, b, c, d 为常数，则有

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;
2. $\mathrm{Cov}(aX + b, cY + d) = ac\,\mathrm{Cov}(X, Y)$;
3. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$;
4. $\mathrm{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$;
5. $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y)$.

协方差的性质

设 X, Y, Z 为随机变量，a, b, c, d 为常数，则有

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$;
2. $\mathrm{Cov}(aX + b, cY + d) = ac\,\mathrm{Cov}(X, Y)$;
3. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$;
4. $\mathrm{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$;
5. $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y)$.

推论 两随机变量相互独立，则协方差等于零；反之未必成立.

协方差的性质

例 1　设二维离散型随机向量 (X,Y) 的概率分布如下表所示：

$X \backslash Y$	0	1
0	1/8	1/4
1	1/2	1/8

求 $\mathrm{Cov}(X,Y)$.

协方差

练习 1 假设二维随机变量 (X, Y) 的联合分布为

$X \backslash Y$	0	1	2
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{3}$	0
2	$\frac{1}{12}$	0	$\frac{1}{12}$

求 $\mathrm{Cov}(X, Y)$.

协方差

练习 1 假设二维随机变量 (X, Y) 的联合分布为

$X \backslash Y$	0	1	2
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	0	$\frac{1}{3}$	0
2	$\frac{1}{12}$	0	$\frac{1}{12}$

求 $\mathrm{Cov}(X - Y, Y)$.

第三节	协方差与相关系数
A	协方差
B	相关系数

定义 对于二维随机变量 (X,Y)，如果两个变量的方差都不为零，称

$$\rho_{X,Y} := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}}$$

为 X 与 Y 的**相关系数**，也可以记为 $\rho(X,Y)$.

定义 对于二维随机变量 (X,Y)，如果两个变量的方差都不为零，称

$$\rho_{X,Y} := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}} = \mathrm{Cov}(X^*,Y^*)$$

为 X 与 Y 的**相关系数**，也可以记为 $\rho(X,Y)$.

定义 对于二维随机变量 (X,Y)，如果两个变量的方差都不为零，称

$$\rho_{X,Y} := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}} = \mathrm{Cov}(X^*,Y^*)$$

为 X 与 Y 的相关系数，也可以记为 $\rho(X,Y)$.

性质 相关系数表示随机变量之间的线性相关程度：

1. $|\rho_{X,Y}| \leqslant 1$.
2. $\rho_{X,Y} = -1$ 当且仅当 $Y = aX + b$，$a < 0$.
3. $\rho_{X,Y} = 1$ 当且仅当 $Y = aX + b$，$a > 0$.

相关系数

定义 若随机变量 X 与 Y 的相关系数 $\rho_{X,Y}=0$，则称 X 与 Y 线性互不相关，简称不相关.

相关系数

定义 若随机变量 X 与 Y 的相关系数 $\rho_{X,Y}=0$，则称 X 与 Y 线性互不相关，简称不相关.

完全负相关	负相关	不相关	正相关	完全正相关
$\rho=-1$	$\rho<0$	$\rho=0$	$\rho>0$	$\rho=1$

- $\rho_{X,Y}=-1$ 当且仅当 $Y=aX+b$，$a<0$；
- $\rho_{X,Y}=1$ 当且仅当 $Y=aX+b$，$a>0$.

相关系数

定义 若随机变量 X 与 Y 的相关系数 $\rho_{X,Y}=0$，则称 X 与 Y 线性互不相关，简称不相关.

完全负相关	负相关	不相关	正相关	完全正相关
$\rho=-1$	$\rho<0$	$\rho=0$	$\rho>0$	$\rho=1$

- $\rho_{X,Y}=-1$ 当且仅当 $Y=aX+b$，$a<0$；
- $\rho_{X,Y}=1$ 当且仅当 $Y=aX+b$，$a>0$.

性质 相互独立 $\Longrightarrow$ 不相关；反之未必成立.

相关系数

$X \backslash Y$	1	2	3	4
4	a	0	0	a
3	0	a	a	0
2	0	a	a	0
1	a	0	0	a

$\rho = 0$

$X \backslash Y$	1	2	3	4
4	0	0	0	c
3	0	c	0	0
2	0	0	c	0
1	c	0	0	0

$\rho = 0.8$

$X \backslash Y$	1	2	3	4
4	0	0	0	c
3	0	0	c	0
2	0	c	0	0
1	c	0	0	0

$\rho = 1$

$X \backslash Y$	1	2	3	4
4	c	0	0	0
3	0	0	c	0
2	0	c	0	0
1	0	0	0	c

$\rho = -0.8$

$X \backslash Y$	1	2	3	4
4	c	0	0	0
3	0	c	0	0
2	0	0	c	0
1	0	0	0	c

$\rho = -1$

独立性和相关性

例 2　二维随机向量 (X,Y) 的服从以下分布律

$X \backslash Y$	-1	0	1
-1	1/8	1/8	1/8
0	1/8	0	1/8
1	1/8	1/8	1/8

判断 X,Y 的独立性和相关性.

独立性和相关性

例 3　设 (X,Y) 服从 $D=\{(x,y):x^2+y^2\leqslant 1\}$ 上的均匀分布，即概率密度函数为

$$f(x,y)=\begin{cases}\frac{1}{\pi}, & x^2+y^2\leqslant 1\\ 0, & \text{otherwise}\end{cases}$$

求 $\rho_{X,Y}$.

独立性和相关性

例 3　设 (X,Y) 服从 $D=\{(x,y):x^2+y^2\leqslant 1\}$ 上的均匀分布，即概率密度函数为

$$f(x,y)=\begin{cases}\frac{1}{\pi}, & x^2+y^2\leqslant 1\\ 0, & \text{otherwise}\end{cases}$$

求 $\rho_{X,Y}$.

注记　在这个例子中，X 和 Y 不相关，但是两者不是相互独立的.

相关系数

例 4　设 $(X,Y) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求 $\rho_{X,Y}$.

相关系数

例 4　设 $(X,Y) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求 $\rho_{X,Y}$.

答案：$\rho_{X,Y} = \rho$.

相关系数

例 4　设 $(X,Y) \sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，求 $\rho_{X,Y}$.

答案：$\rho_{X,Y} = \rho$.

结论：对于二维正态随机向量 (X,Y)，X 与 Y 相互独立的充分必要条件是 $\rho_{X,Y} = 0$.

相关系数

练习 2 假设随机变量 X 的概率分布为

X	-1	0	1
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

求 X 与 X^2 的相关系数以及 X 与 X^3 的相关系数.

相关系数

由协方差的性质及相关系数与协方差的关系可得：

$$\mathrm{Var}(X \pm Y)$$

$$= \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\rho_{X,Y}\sqrt{\mathrm{Var}(X)} \cdot \sqrt{\mathrm{Var}(Y)}.$$

相关系数

由协方差的性质及相关系数与协方差的关系可得：

$$\begin{aligned}&\mathrm{Var}(X \pm Y)\\ =&\mathrm{Var}(X)+\mathrm{Var}(Y) \pm 2\rho_{X,Y}\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}.\end{aligned}$$

例 5　已知随机变量 X 和 Y 的方差分别为 1 和 4，相关系数为 -0.5．求 $\mathrm{Var}(X+Y)$ 和 $\mathrm{Var}(X-Y)$.

例 6 (投资风险组合)　设有 1 百万用于投资甲、乙两种证券：若将资金 t 投资于甲证券，将资金 $1-t$ 投资于乙证券，则称 $(t, 1-t)$ 为一个投资组合.

例 6 (投资风险组合)　设有 1 百万用于投资甲、乙两种证券：若将资金 t 投资于甲证券，将资金 $1-t$ 投资于乙证券，则称 $(t, 1-t)$ 为一个投资组合.

用随机变量 X 和 Y 分别表示投资甲、乙证券的收益率. 已知 X 和 Y 的期望（代表平均收益）分别为 μ_1 和 μ_2，方差（代表风险）分别为 σ_1^2 和 σ_2^2，相关系数为 ρ.

例 6 (投资风险组合)　设有 1 百万用于投资甲、乙两种证券：若将资金 t 投资于甲证券，将资金 $1-t$ 投资于乙证券，则称 $(t, 1-t)$ 为一个投资组合.

用随机变量 X 和 Y 分别表示投资甲、乙证券的收益率. 已知 X 和 Y 的期望（代表平均收益）分别为 μ_1 和 μ_2，方差（代表风险）分别为 σ_1^2 和 σ_2^2，相关系数为 ρ.

1. 求投资组合的平均收益和风险.
2. 求投资风险最小的投资组合.

例 6 (投资风险组合) 设有 1 百万用于投资甲、乙两种证券：若将资金 t 投资于甲证券，将资金 $1-t$ 投资于乙证券，则称 $(t, 1-t)$ 为一个投资组合.

用随机变量 X 和 Y 分别表示投资甲、乙证券的收益率. 已知 X 和 Y 的期望（代表平均收益）分别为 μ_1 和 μ_2，方差（代表风险）分别为 σ_1^2 和 σ_2^2，相关系数为 ρ.

1. 求投资组合的平均收益和风险.
2. 求投资风险最小的投资组合.

在第二问中假设 $\sigma_1^2 = 0.25$、$\sigma_2^2 = 0.49$、$\rho = 0.6$.

复习与提高

选择 (2001)　将一枚硬币重复掷 n 次，以 X 和 Y 分别表示正面向上和反面向上的次数，则 X 和 Y 的相关系数等于 . . . . . (　　)

(A) -1　　(B) 0　　(C) $\frac{1}{2}$　　(D) 1