

第一节 数学期望

- 一、数学期望的概念
- 二、数学期望的性质
- 三、随机变量函数的数学期望
- 四、小结



一、数学期望的概念

引例1 分赌本问题(产生背景)

A, B 两人赌技相同, 各出赌金100元, 并约定先胜三局者为胜, 取得全部 200 元. 由于出现意外情况, 在 A 胜 2 局 B 胜1 局时, 不得不终止赌博, 如果要分赌金, 该如何分配才算公平?



分析 假设继续赌两局,则结果有以下四种情况:

AA	AB	BA	BB
A 胜 B 负	A 胜 B 负	B 胜 A 负	B 胜 A 负
A 胜 B 负	B 胜 A 负	A 胜 B 负	B 胜 A 负

把已赌过的三局(A 胜2局 B 胜1局)与上述结果相结合,即 A 、 B 赌完五局,

前三局: A 胜 2 局 B 胜 1 局

后二局:

AA	AB	BA	B
A 胜			B 胜



故有, 在赌技相同的情况下, A, B 最终获胜的可能性大小之比为 **3:1**,

即 A 应获得赌金的 $\frac{3}{4}$, 而 B 只能获得赌金的 $\frac{1}{4}$.

因此, A 能“**期望**”得到的数目应为

$$200 \times \frac{3}{4} + 0 \times \frac{1}{4} = 150(\text{元}),$$

而 B 能“**期望**”得到的数目, 则为

$$200 \times \frac{1}{4} + 0 \times \frac{3}{4} = 50(\text{元}).$$



若设随机变量 X 为:在 A 胜2局 B 胜1局的前提下,继续赌下去 A 最终所得的赌金.

则 X 所取可能值为: 200 0

其概率分别为: $\frac{3}{4}$ $\frac{1}{4}$

因而 A 期望所得的赌金即为 X 的 “期望” 值,

等于 $200 \times \frac{3}{4} + 0 \times \frac{1}{4} = 150(\text{元}).$

即为 X 的可能值与其概率之积的累加.



引例2 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 = $\frac{\text{射中靶的总环数}}{\text{射击次数}}$

$$= \frac{0 \times 2 + 1 \times 13 + 2 \times 15 + 3 \times 10 + 4 \times 20 + 5 \times 30}{90}$$

$$= 0 \times \frac{2}{90} + 1 \times \frac{13}{90} + 2 \times \frac{15}{90} + 3 \times \frac{10}{90} + 4 \times \frac{20}{90} + 5 \times \frac{30}{90}$$

$$= \sum_{k=0}^5 k \cdot \frac{n_k}{n} = 3.37.$$

设射手命中的环数为随机变量 Y .



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

平均射中环数
随机波动
频率随机波动

“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

→

↓

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



1. 离散型随机变量的数学期望

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

为随机变量 X 的数学期望, 记为 $E(X)$. 即

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



分赌本问题

A 期望所得的赌金即为 X 的数学期望

$$E(X) = 200 \times \frac{3}{4} + 0 \times \frac{1}{4} = 150(\text{元}).$$

射击问题

“平均射中环数” 应为随机变量 Y 的数学期望

$$E(Y) =$$

$$0 \times p_0 + 1 \times p_1 + 2 \times p_2 + 3 \times p_3 + 4 \times p_4 + 5 \times p_5.$$



关于定义的几点说明

(1) $E(X)$ 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.

(2) **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.

(3) 随机变量的数学期望与一般变量的算术平均值不同.

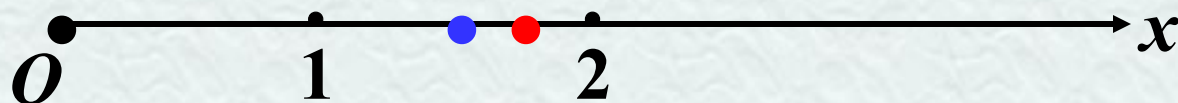


假设

X	1	2
p	0.02	0.98

随机变量 X 的算术平均值为 $\frac{1+2}{2} = 1.5$,

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



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。

实例1 谁的技术比较好？



甲、乙两个射手，他们射击的分布律分别为

甲射手

击中环数	8	9	10
概率	0.3	0.1	0.6

乙射手

击中环数	8	9	10
概率	0.2	0.5	0.3

试问哪个射手技术较好？



解 设甲、乙射手击中的环数分别为 X_1, X_2 .

$$E(X_1) = 8 \times 0.3 + 9 \times 0.1 + 10 \times 0.6 = 9.3(\text{环}),$$

$$E(X_2) = 8 \times 0.2 + 9 \times 0.5 + 10 \times 0.3 = 9.1(\text{环}),$$

故甲射手的技术比较好.



实例2 发行彩票的创收利润

某一彩票中心发行彩票 10万张, 每张2元. 设头等奖1个, 奖金 1万元, 二等奖2个, 奖金各 5 千元; 三等奖 10个, 奖金各1千元; 四等奖100个, 奖金各 100元; 五等奖1000个, 奖金各10 元. 每张彩票的成本费为 0.3 元, 请计算彩票发行单位的创收利润.

解 设每张彩票中奖的数额为随机变量 X , 则

X	10000	5000	1000	100	10	0
p	$1/10^5$	$2/10^5$	$10/10^5$	$100/10^5$	$1000/10^5$	p_0



每张彩票平均能得到奖金

$$\begin{aligned} E(X) &= 10000 \times \frac{1}{10^5} + 5000 \times \frac{2}{10^5} + \cdots + 0 \times p_0 \\ &= 0.5(\text{元}), \end{aligned}$$

每张彩票平均可赚

$$2 - 0.5 - 0.3 = 1.2(\text{元}),$$

因此彩票发行单位发行 10 万张彩票的创收利润为

$$100000 \times 1.2 = 120000(\text{元}).$$



实例3 如何确定投资决策方向？

某人有10万元现金，想投资于某项目，预估成功的机会为30%，可得利润8万元，失败的机会为70%，将损失2万元。若存入银行，同期间的利率为5%，问是否作此项投资？



解 设 X 为投资利润，则

X	8	-2
p	0.3	0.7

$E(X) = 8 \times 0.3 - 2 \times 0.7 = 1$ (万元), 存入银行的利息:
 $10 \times 5\% = 0.5$ (万元), 故应选择投资.



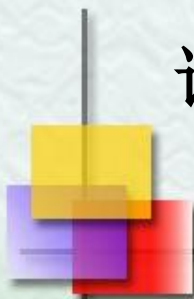
实例4 商店的销售策略

某商店对某种家用电器 的销售采用先使用后付款的方式 ,记使用寿命为 X (以年计),规定 :
 $X \leq 1$,一台付款 1500 元; $1 < X \leq 2$,一台付款 2000 元;
 $2 < X \leq 3$,一台付款 2500 元; $X > 3$,一台付款 3000 元.

设寿命 X 服从指数分布 ,概率密度为

$$f(x) = \begin{cases} \frac{1}{10} e^{-x/10}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

试求该商店一台家用电器收费 Y 的数学期望 .



解
$$P\{X \leq 1\} = \int_0^1 \frac{1}{10} e^{-x/10} dx = 1 - e^{-0.1} = 0.0952,$$

$$\begin{aligned} P\{1 < X \leq 2\} &= \int_1^2 \frac{1}{10} e^{-x/10} dx \\ &= e^{-0.1} - e^{-0.2} = 0.0861, \end{aligned}$$

$$\begin{aligned} P\{2 < X \leq 3\} &= \int_2^3 \frac{1}{10} e^{-x/10} dx \\ &= e^{-0.2} - e^{-0.3} = 0.0779, \end{aligned}$$



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

$$= e^{-0.3} = 0.7408.$$

因而一台收费 Y 的分布律为

Y	1500	2000	2500	3000
p_k	0.0952	0.0861	0.0779	0.7408

得 $E(Y) = 2732.15,$

即平均一台家用电器收费 2732.15 元.



实例5 分组验血

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,这样, k 个人的血共最多需化验 $k + 1$ 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 由于血液呈阳性反应的 概率为 p ,

所以血液呈阴性反应的 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, 且其分布律为



X	$\frac{1}{k}$	$\frac{k+1}{k}$
p_k	q^k	$1 - q^k$

X 的数学期望为

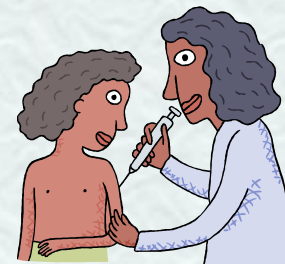
$$E(X) = \frac{1}{k}q^k + (1 + \frac{1}{k})(1 - q^k) = 1 - q^k + \frac{1}{k}.$$

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



因此,只要选择 k 使

$$1 - q^k + \frac{1}{k} < 1,$$



则 N 个人平均需化验的次数 $< N$.

当 p 固定时,选取 k 使得

$$L = 1 - q^k + \frac{1}{k} \text{ 小于1且取到最小值,}$$

此时可得到最好的分组 方法.



实例6 按规定,某车站每天 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

到站时刻	8:10	8:30	8:50
	9:10	9:30	9:50
概率	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$



- (i) 一旅客8:00到车站,求他候车时间的数学期望.
- (ii) 一旅客8:20到车站,求他候车时间的数学期望.



解 设旅客的候车时间为 X (以分计).

(i) X 的分布律为

X	10	30	50
p_k	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$

候车时间的数学期望为

$$\begin{aligned} E(X) &= 10 \times \frac{1}{6} + 30 \times \frac{3}{6} + 50 \times \frac{2}{6} \\ &= 33.33(\text{分}). \end{aligned}$$



(ii) X 的分布律为

X	10	30	50	70	90
p_k	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6} \times \frac{1}{6}$	$\frac{1}{6} \times \frac{3}{6}$	$\frac{1}{6} \times \frac{2}{6}$

候车时间的数学期望为

$$E(X) =$$

$$10 \times \frac{3}{6} + 30 \times \frac{2}{6} + 50 \times \frac{1}{6} \times \frac{1}{6} + 70 \times \frac{1}{6} \times \frac{3}{6} + 90 \times \frac{1}{6} \times \frac{2}{6}$$

$$= 27.22(\text{分}).$$



2.连续型随机变量数学期望的定义

设连续型随机变量 X 的概率密度为 $f(x)$,
若积分

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

绝对收敛, 则称积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 的值为随机
变量 X 的数学期望, 记为 $E(X)$. 即

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



实例7 顾客平均等待多长时间?

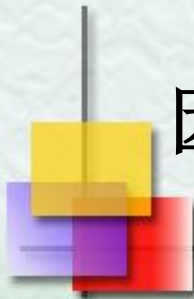
设顾客在某银行的窗口等待服务的时间 X (以分计)服从指数分布,其概率密度为

$$f(x) = \begin{cases} \frac{1}{5}e^{-x/5}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

试求顾客等待服务的平均时间?

解 $E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^{+\infty} x \cdot \frac{1}{5}e^{-x/5} dx = 5(\text{分钟}).$

因此, 顾客平均等待5分钟就可得到服务.



二、数学期望的性质

1. 设 C 是常数, 则有 $E(C) = C$.

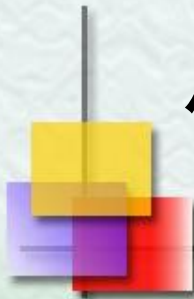
证明 $E(X) = E(C) = 1 \times C = C$.

2. 设 X 是一个随机变量, C 是常数, 则有

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

证明 $E(CX) = \sum_k Cx_k p_k = C \sum_k x_k p_k = CE(X)$.

例如 $E(X) = 5$, 则 $E(3X) = 3E(X) = 3 \times 5 = 15$.



3. 设 X, Y 是两个随机变量, 则有

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

证明

$$\begin{aligned} E(X + Y) &= \sum_k (x_k + y_k) p_k \\ &= \sum_k x_k p_k + \sum_k y_k p_k = E(X) + E(Y). \end{aligned}$$

4. 设 X, Y 是相互独立的随机变量, 则有

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

说明 连续型随机变量 X 的数学期望与离散型随机变量数学期望的性质类似.



实例8 一机场班车载有 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 $E(X)$ (设每位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 引入随机变量 X_i ,

$$X_i = \begin{cases} 0, & \text{在第 } i \text{ 站没有人下车,} \\ 1, & \text{在第 } i \text{ 站有人下车,} \end{cases} \quad i = 1, 2, \dots, 10.$$

则 $X = X_1 + X_2 + \dots + X_{10}$.



$$\text{则有 } P\{X_i = 0\} = \left(\frac{9}{10}\right)^{20}, \quad P\{X_i = 1\} = 1 - \left(\frac{9}{10}\right)^{20},$$
$$i = 1, 2, \dots, 10.$$

$$\text{由此 } E(X_i) = 1 - \left(\frac{9}{10}\right)^{20}, \quad i = 1, 2, \dots.$$

$$\begin{aligned} \text{得 } E(X) &= E(X_1 + X_2 + \dots + X_{10}) \\ &= E(X_1) + E(X_2) + \dots + E(X_{10}) \\ &= 10 \left[1 - \left(\frac{9}{10}\right)^{20} \right] = 8.784(\text{次}). \end{aligned}$$



三、随机变量函数的数学期望

1. 离散型随机变量函数的数学期望

设随机变量 X 的分布律为

$X = x_k$	-1	0	1	2
$P\{X=x_k\}=p_k$	p_1	p_2	p_3	p_4

若 $Y = g(X) = X^2$, 求 $E(Y)$.

解 先求 $Y = X^2$ 的分布律

$Y = X^2$	0	1	4
p	p_2	$p_1 + p_3$	p_4



$$\begin{aligned}
 \text{则有 } E(Y) &= E(g(X)) = E(X^2) \\
 &= 0 \cdot p_2 + 1 \cdot (p_1 + p_2) + 4 \cdot p_4 \\
 &= 0 \cdot p_2 + (-1)^2 \cdot p_1 + 1^2 \cdot p_2 + 2^2 \cdot p_4 \\
 &= \sum_{k=1}^4 g(x_k) P\{X = x_k\}.
 \end{aligned}$$

因此离散型随机变量函数的数学期望为

若 $Y=g(X)$, 且 $P\{X = x_k\} = p_k, \quad k = 1, 2, \dots,$

则有

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



2. 连续型随机变量函数的数学期望

若 X 是连续型的, 它的分布密度为 $f(x)$, 则

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

3. 二维随机变量函数的数学期望

(1) 设 X, Y 为离散型随机变量, $g(x, y)$ 为二元函数, 则 $E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j) p_{ij}.$

其中 (X, Y) 的联合概率分布为 $p_{ij}.$



(2) 设 X, Y 为连续型随机变量, $g(x, y)$ 为二元函数, 则

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

其中 (X, Y) 的联合概率密度为 $f(x, y)$.



实例9 设 (X, Y) 的分布律为

$Y \backslash X$	1	2	3
-1	0.2	0.1	0
0	0.1	0	0.3
1	0.1	0.1	0.1

求： $E(X)$, $E(Y)$, $E(Y/X)$, $E[(X - Y)^2]$.

解 X 的分布律为

X	1	2	3
p	0.4	0.2	0.4



得 $E(X) = 1 \times 0.4 + 2 \times 0.2 + 3 \times 0.4 = 2.$

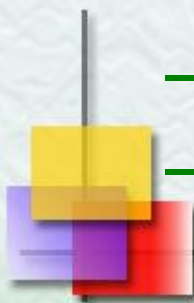
Y 的分布律为

Y	-1	0	1
p	0.3	0.4	0.3

得 $E(Y) = -1 \times 0.3 + 0 \times 0.4 + 1 \times 0.3 = 0.$

由于

p	0.2	0.1	0.1	0.1	0.1	0.3	0.1
(X, Y)	$(1, -1)$	$(1, 0)$	$(1, 1)$	$(2, -1)$	$(2, 1)$	$(3, 0)$	$(3, 1)$
Y/X	-1	0	1	$-1/2$	$1/2$	0	$1/3$



于是

$$E\left(\frac{Y}{X}\right) = -1 \times 0.2 + 0 \times 0.1 + 1 \times 0.1 - \frac{1}{2} \times 0.1 + \frac{1}{2} \times 0.1 + 0 \times 0.3 + \frac{1}{3} \times 0.1$$

$$= -\frac{1}{15}.$$

p	0.2	0.1	0.1	0.1	0.1	0.3	0.1
(X, Y)	(1, -1)	(1, 0)	(1, 1)	(2, -1)	(2, 1)	(3, 0)	(3, 1)
$(X - Y)^2$	4	1	0	9	1	9	4

得 $E[(X - Y)^2] = 4 \times 0.3 + 1 \times 0.2 + 0 \times 0.1 + 9 \times 0.4$

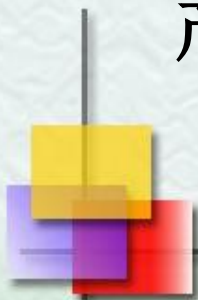
$$= 5.$$



实例10 某公司计划开发一种新 产品市场,并试图确定该产品的产量.他们估计出售一件产品 可获利 m 元,而积压一件产品导致 n 元的损失.再者,他们预测销售量 Y (件)服从指数分布其概率密 度为

$$f_Y(y) = \begin{cases} \frac{1}{\theta} e^{-y/\theta}, & y > 0, \\ 0, & y \leq 0. \end{cases} \quad \theta > 0,$$

问若要获得利润的数学 期望最大,应生产多少件产品(m, n, θ 均为已知)?



解 设生产 x 件, 则获利 Q 是 x 的函数:

$$Q = Q(x) = \begin{cases} mY - n(x - Y), & \text{若 } Y < x, \\ mx, & \text{若 } Y \geq x. \end{cases}$$

$$\begin{aligned} E(Q) &= \int_0^{+\infty} Q f_Y(y) \mathrm{d} y \\ &= \int_0^x [my - n(x - y)] \frac{1}{\theta} e^{-y/\theta} \mathrm{d} y + \int_x^{+\infty} mx \frac{1}{\theta} e^{-y/\theta} \mathrm{d} y \\ &= (m + n)\theta - (m + n)\theta e^{-x/\theta} - nx, \end{aligned}$$

令 $\frac{\mathrm{d}}{\mathrm{d} x} E(Q) = (m + n)e^{-x/\theta} - n = 0,$



得 $x = -\theta \ln\left(\frac{n}{m+n}\right).$

又 $\frac{d^2}{dx^2} E(Q) = \frac{-(m+n)}{\theta} e^{-x/\theta} < 0,$

因此, 当 $x = -\theta \ln\left(\frac{n}{m+n}\right)$ 时, $E(Q)$ 取得最大值.



实例11 (卖报问题) 设某卖报人每日的潜在卖报数 ξ 服从参数为 λ 的泊松分布. 如果每卖出一份报可得报酬 a , 卖不掉而退回则每份赔偿 b , 若某日报卖人买进 n 份报, 试求其期望所得. 进一步, 再求最佳的卖报份数.

解 若记其真正卖报数为 η , 则 η 与 ξ 的关系如下:

$$\eta = \begin{cases} \xi, & \xi < n \\ n, & \xi \geq n \end{cases},$$

则 η 的分布为



$$P\{\eta = k\} = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda}, & k < n, \\ \sum_{i=n}^{+\infty} \frac{\lambda^i}{i!} e^{-\lambda}, & k = n. \end{cases}$$

记所得为 ζ , 则 ζ 与 η 的关系如下:

$$\zeta = g(\eta) = \begin{cases} a\eta - b(n - \eta), & \eta < n, \\ an, & \eta = n. \end{cases}$$

因此期望所得为

$$M(n) = E[g(\eta)]$$



$$= \sum_{k=0}^{n-1} \frac{\lambda^k}{k!} e^{-\lambda} [ka - (n-k)b] + \left(\sum_{k=n}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \right) na$$

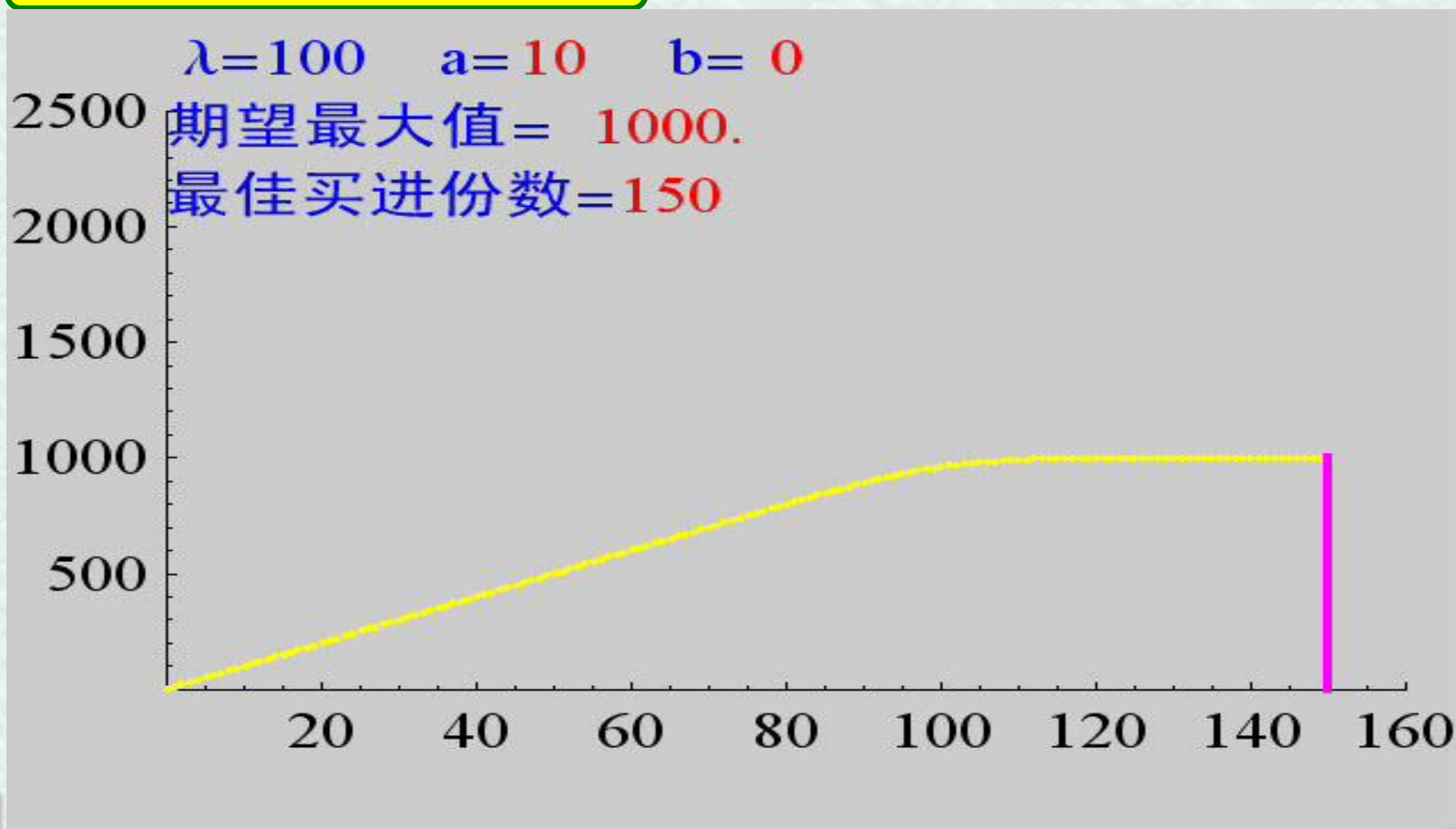
$$= (a+b)\lambda \sum_{k=0}^{n-2} \frac{\lambda^k}{k!} e^{-\lambda} - n(a+b) \sum_{k=0}^{n-1} \frac{\lambda^k}{k!} e^{-\lambda} + na$$

当 a, b, λ 给定后, 求 n 使 $M(n)$ 达到极大.



利用软件包求解,并演示计算结果.

单击图形播放/暂停 ESC键退出



四、小结

1. 数学期望是一个实数, 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**.

2. 数学期望的性质

$$1^{\circ} E(C) = C;$$

$$2^{\circ} E(CX) = CE(X);$$

$$3^{\circ} E(X + Y) = E(X) + E(Y);$$

$$4^{\circ} X, Y \text{ 独立} \Rightarrow E(XY) = E(X)E(Y).$$



第二节 方 差

- 一、随机变量方差的概念及性质
- 二、重要概率分布的方差
- 三、例题讲解
- 四、小结

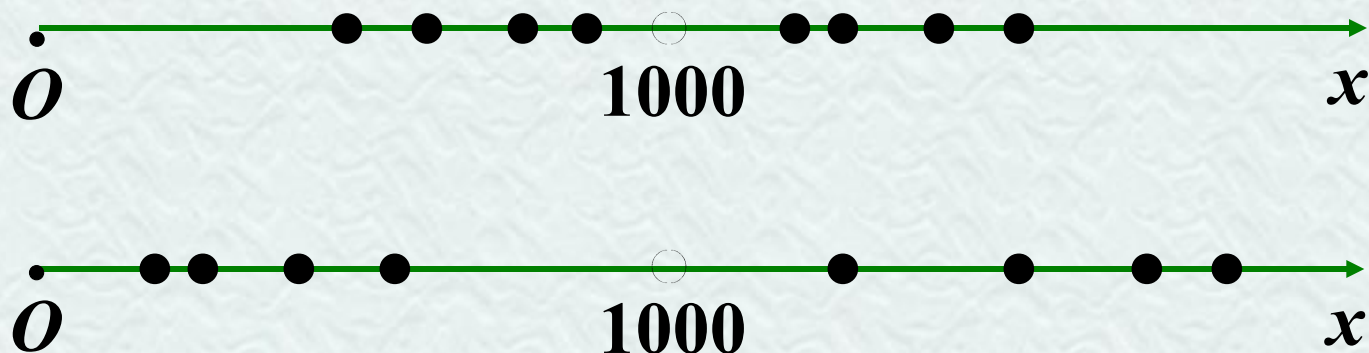


一、随机变量方差的概念及性质

1. 概念的引入

方差是一个常用来体现随机变量取值分散程度的量.

实例 有两批灯泡,其平均寿命都是 $E(X)=1000$ 小时.



2. 方差的定义

设 X 是一个随机变量, 若 $E\{[X - E(X)]^2\}$ 存在, 则称 $E\{[X - E(X)]^2\}$ 为 X 的方差, 记为 $D(X)$ 或 $\text{Var}(X)$, 即...

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

称 $\sqrt{D(X)}$ 为标准差或均方差, 记为 $\sigma(X)$.



3. 方差的意义

方差是一个常用来体现随机变量 X 取值分散程度的量. 如果 $D(X)$ 值大, 表示 X 取值分散程度大, $E(X)$ 的代表性差; 而如果 $D(X)$ 值小, 则表示 X 的取值比较集中, 以 $E(X)$ 作为随机变量的代表性好.



4. 随机变量方差的计算

(1) 利用定义计算

离散型随机变量的方差

$$D(X) = \sum_{k=1}^{+\infty} [x_k - E(X)]^2 p_k,$$

其中 $P\{X = x_k\} = p_k, k = 1, 2, \dots$ 是 X 的分布律.

连续型随机变量的方差

$$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx,$$

其中 $f(x)$ 为 X 的概率密度.



(2) 利用公式计算

$$D(X) = E(X^2) - [E(X)]^2.$$

证明

$$\begin{aligned} D(X) &= E\{[X - E(X)]^2\} \\ &= E\{X^2 - 2XE(X) + [E(X)]^2\} \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \\ &= E(X^2) - E^2(X). \end{aligned}$$



5. 方差的性质

(1) 设 C 是常数, 则有 $D(C) = 0$.

证明 $D(C) = E(C^2) - [E(C)]^2 = C^2 - C^2 = 0$.

(2) 设 X 是一个随机变量, C 是常数, 则有

$$D(CX) = C^2 D(X).$$

证明 $D(CX) = E\{[CX - E(CX)]^2\}$

$$= C^2 E\{[X - E(X)]^2\}$$

$$= C^2 D(X).$$



(3) 设 X, Y 相互独立, $D(X), D(Y)$ 存在, 则

$$D(X \pm Y) = D(X) + D(Y).$$

证明

$$\begin{aligned} D(X \pm Y) &= E\{[(X \pm Y) - E(X \pm Y)]^2\} \\ &= E\{[X - E(X)] \pm [Y - E(Y)]\}^2 \\ &= E[X - E(X)]^2 + E[Y - E(Y)]^2 \\ &\quad \pm 2E\{[X - E(X)][Y - E(Y)]\} \\ &= D(X) + D(Y). \end{aligned}$$



推广 若 $X_1, X_2, \dots, X_n$ 相互独立, 则有

$$D(X_1 \pm X_2 \pm \dots \pm X_m) = D(X_1) + D(X_2) + \dots + D(X_m).$$

(4) $D(X) = 0$ 的充要条件是 X 以概率 1 取常数 C , 即

$$P\{X = C\} = 1.$$



二、重要概率分布的方差

1. 两点分布

已知随机变量 X 的分布律为

X	1	0
p	p	$1-p$

则有 $E(X) = 1 \cdot p + 0 \cdot q = p,$

$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 \\ &= 1^2 \cdot p + 0^2 \cdot (1-p) - p^2 = pq. \end{aligned}$$



2. 二项分布

设随机变量 X 服从参数为 n, p 二项分布, 其分布律为

$$P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}, (k = 0, 1, 2, \dots, n),$$

则有

$$0 < p < 1.$$

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

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



$$= \sum_{k=0}^n \frac{kn!}{k!(n-k)!} p^k (1-p)^{n-k}$$

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

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

$$= np[p + (1-p)]^{n-1}$$

$$= np.$$



$$E(X^2) = E[X(X-1) + X]$$

$$= E[X(X-1)] + E(X)$$

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

$$= \sum_{k=0}^n \frac{k(k-1)n!}{k!(n-k)!} p^k (1-p)^{n-k} + np$$



$$= n(n-1)p^2 \sum_{k=2}^n \frac{(n-2)!}{(n-k)!(k-2)!} p^{k-2} (1-p)^{(n-2)-(k-2)} + np$$

$$= n(n-1)p^2 [p + (1-p)]^{n-2} + np$$

$$= (n^2 - n)p^2 + np.$$

$$D(X) = E(X^2) - [E(X)]^2$$

$$= (n^2 - n)p^2 + np - (np)^2$$

$$= np(1-p).$$



3. 泊松分布

设 $X \sim \pi(\lambda)$, 且分布律为

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots, \quad \lambda > 0.$$

则有

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \cdot \lambda \\ &= \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda. \end{aligned}$$



$$E(X^2) = E[X(X-1) + X]$$

$$= E[X(X-1)] + E(X)$$

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

$$= \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda.$$

所以 $D(X) = E(X^2) - [E(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$

泊松分布的期望和方差都等于参数 λ 。



4. 均匀分布

设 $X \sim U(a, b)$, 其概率密度为

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

$$\begin{aligned} \text{则有 } E(X) &= \int_{-\infty}^{\infty} xf(x)dx = \int_a^b \frac{1}{b-a} x dx \\ &= \frac{1}{2}(a+b). \end{aligned}$$



结论 均匀分布的数学期望位于区间的中点.

$$D(X) = E(X^2) - [E(X)]^2$$

$$= \int_a^b x^2 \frac{1}{b-a} dx - \left(\frac{a+b}{2} \right)^2$$

$$= \frac{(b-a)^2}{12}.$$



5. 指数分布

设随机变量 X 服从指数分布, 其概率密度为

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-x/\theta}, & x > 0, \\ 0, & x \leq 0. \end{cases} \quad \text{其中 } \theta > 0.$$

则有

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} xf(x) dx = \int_0^{+\infty} x \cdot \frac{1}{\theta} e^{-x/\theta} dx \\ &= -xe^{-x/\theta} \Big|_0^{+\infty} + \int_0^{+\infty} e^{-x/\theta} dx = \theta. \end{aligned}$$



$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 \\ &= \int_0^{+\infty} x^2 \cdot \frac{1}{\theta} e^{-x/\theta} dx - \theta^2 \\ &= 2\theta^2 - \theta^2 \\ &= \theta^2. \end{aligned}$$

指数分布的期望和方差分别为 θ 和 θ^2 。



6. 正态分布

设 $X \sim N(\mu, \sigma^2)$, 其概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \sigma > 0, \quad -\infty < x < +\infty.$$

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

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

$$\text{令 } \frac{x-\mu}{\sigma} = t \Rightarrow x = \mu + \sigma t,$$



$$\begin{aligned}
 \text{所以 } E(X) &= \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\mu + \sigma t) e^{-\frac{t^2}{2}} dt \\
 &= \mu \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt \\
 &= \mu.
 \end{aligned}$$



$$D(X) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$$

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

令 $\frac{x - \mu}{\sigma} = t$, 得

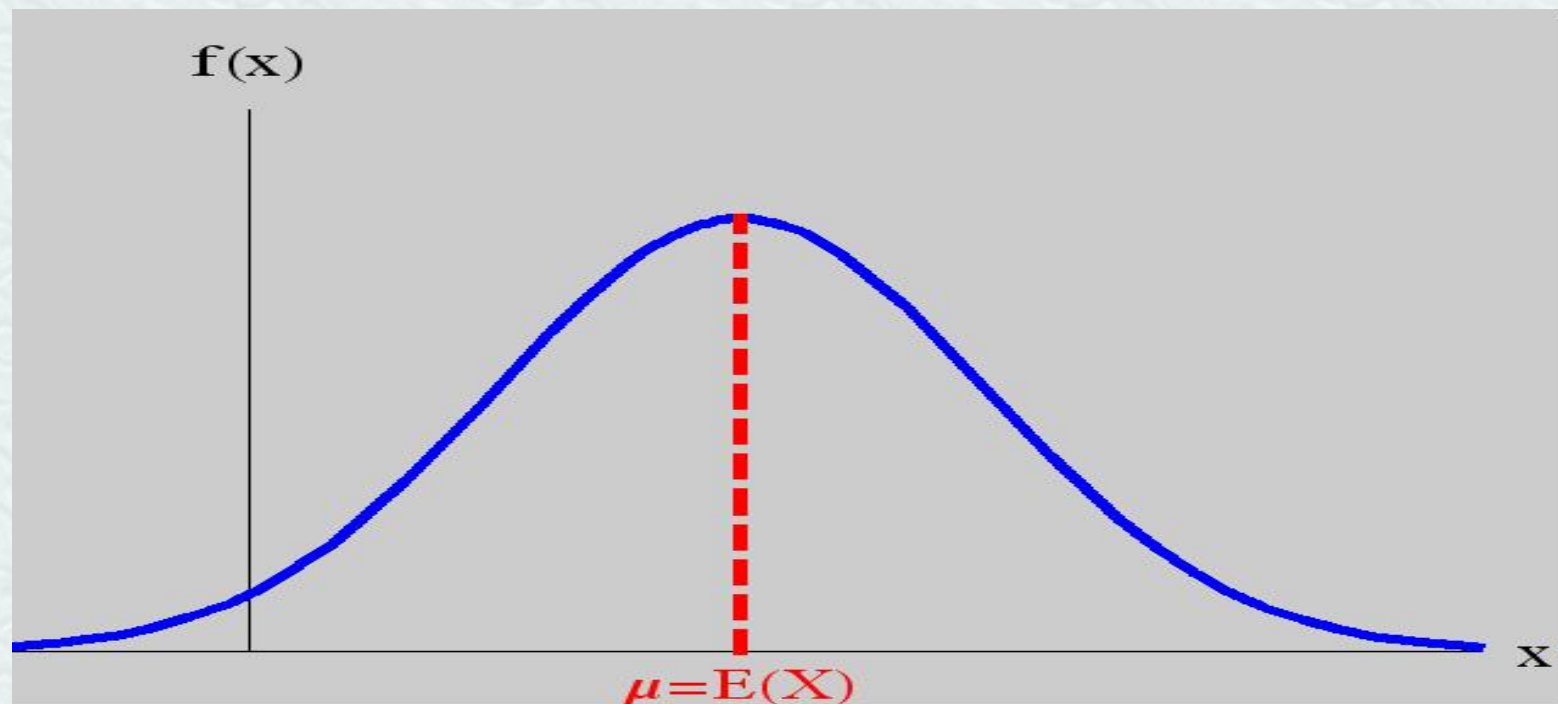
$$D(X) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} dt$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left(-te^{-\frac{t^2}{2}} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt \right)$$

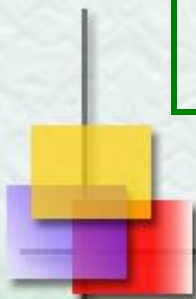


$$= 0 + \frac{\sigma^2}{\sqrt{2\pi}} \sqrt{2\pi} = \sigma^2.$$

正态分布的期望和方差分别为两个参数 μ 和 σ^2 。



分 布	参 数	数学期望	方差
两点分布	$0 < p < 1$	p	$p(1 - p)$
二项分布	$n \geq 1,$ $0 < p < 1$	np	$np(1 - p)$
泊松分布	$\lambda > 0$	λ	λ
均匀分布	$a < b$	$(a + b)/2$	$(b - a)^2/12$
指数分布	$\theta > 0$	θ	θ^2
正态分布	$\mu, \sigma > 0$	μ	σ^2



三、例题讲解

例1 设随机变量 X 具有概率密度

$$f(x) = \begin{cases} 1+x, & -1 \leq x < 0, \\ 1-x, & 0 \leq x < 1, \\ 0, & \text{其他.} \end{cases}$$

求 $D(X)$.

解
$$E(X) = \int_{-1}^0 x(1+x)dx + \int_0^1 x(1-x)dx$$
$$= 0,$$



$$\begin{aligned} E(X^2) &= \int_{-1}^0 x^2(1+x)dx + \int_0^1 x^2(1-x)dx \\ &= \frac{1}{6}, \end{aligned}$$

于是

$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 \\ &= \frac{1}{6} - 0^2 = \frac{1}{6}. \end{aligned}$$



例2 设活塞的直径 (以cm计) $X \sim N(22.40, 0.03^2)$, 气缸的直径 $Y \sim N(22.50, 0.04^2)$, X, Y 相互独立. 任取一只活塞, 任取一只气缸, 求活塞能装入气缸的概率.

解 因为 $X \sim N(22.40, 0.03^2)$, $Y \sim N(22.50, 0.04^2)$,
所以 $X - Y \sim N(-0.10, 0.0025)$,

故有 $P\{X < Y\} = P\{X - Y < 0\}$

$$= P\left\{\frac{(X - Y) - (-0.10)}{\sqrt{0.0025}} < \frac{0 - (-0.10)}{\sqrt{0.0025}}\right\} = \Phi(2) = 0.9772.$$



例3 设连续型随机变量 X 的概率密度为

$$f(x) = \begin{cases} \cos x, & 0 \leq x \leq \frac{\pi}{2}, \\ 0, & \text{其他.} \end{cases}$$

求随机变量 $Y = X^2$ 的方差 $D(Y)$.

解 $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$

$$= \int_0^{\frac{\pi}{2}} x^2 \cos x dx = \frac{\pi^2}{4} - 2,$$

$$E(X^4) = \int_{-\infty}^{+\infty} x^4 f(x) dx = \int_0^{\frac{\pi}{2}} x^4 \cos x dx$$



$$= \frac{\pi^4}{16} - 3\pi^2 + 24,$$

因为 $D(X) = E(X^2) - [E(X)]^2$,

所以 $D(X^2) = E(X^4) - [E(X^2)]^2$

$$\begin{aligned} &= \frac{\pi^4}{16} - 3\pi^2 + 24 - \left(\frac{\pi^2}{4} - 2 \right)^2 \\ &= 20 - 2\pi^2. \end{aligned}$$



例4 设 $X \sim \begin{pmatrix} -2 & 0 & 1 & 3 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{12} & \frac{1}{12} \end{pmatrix}$, 求 $D(2X^3 + 5)$.

$$\begin{aligned} \text{解} \quad D(2X^3 + 5) &= D(2X^3) + D(5) \\ &= 4D(X^3) \\ &= 4[E(X^6) - (E(X^3))^2] \end{aligned}$$

$$E(X^6) = (-2)^6 \times \frac{1}{3} + 0^6 \times \frac{1}{2} + 1^6 \times \frac{1}{12} + 3^6 \times \frac{1}{12} = \frac{493}{6},$$



$$[E(X^3)]^2 = \left[(-2)^3 \times \frac{1}{3} + 0^3 \times \frac{1}{2} + 1^3 \times \frac{1}{12} + 3^3 \times \frac{1}{12} \right]^2$$

$$= \frac{1}{9},$$

故 $D(2X^3 + 5) = 4[E(X^6) - (E(X^3))^2]$

$$= \frac{2954}{9}.$$



契比雪夫不等式

契比雪夫

定理 设随机变量 X 具有数学期望 $E(X) = \mu$, 方差 $D(X) = \sigma^2$, 则对于任意正数 ε , 不等式

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}$$

成立.

切比雪夫不等式

证明 取连续型随机变量的情况来证明.

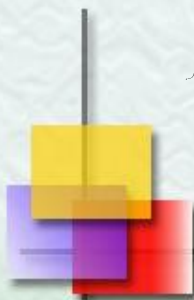
设 X 的概率密度为 $f(x)$, 则有



$$\begin{aligned}
 P\{|X - \mu| \geq \varepsilon\} &= \int_{|x - \mu| \geq \varepsilon} f(x) dx \\
 &\leq \int_{|x - \mu| \geq \varepsilon} \frac{|x - \mu|^2}{\varepsilon^2} f(x) dx \\
 &\leq \frac{1}{\varepsilon^2} \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \frac{1}{\varepsilon^2} \sigma^2.
 \end{aligned}$$

得 $P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2}.$

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} \Leftrightarrow P\{|X - \mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}.$$



四、小结

1. 方差是一个常用来体现随机变量 X 取值分散程度的量. 如果 $D(X)$ 值大, 表示 X 取值分散程度大, $E(X)$ 的代表性差; 而如果 $D(X)$ 值小, 则表示 X 的取值比较集中, 以 $E(X)$ 作为随机变量的代表性好.

2. 方差的计算公式

$$D(X) = E(X^2) - [E(X)]^2,$$

$$D(X) = \sum_{k=1}^{+\infty} [x_k - E(X)]^2 p_k,$$

$$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx.$$



3. 方差的性质

$$1^{\circ} D(C) = 0;$$

$$2^{\circ} D(CX) = C^2 D(X);$$

$$3^{\circ} D(X \pm Y) = D(X) + D(Y).$$

4. 契比雪夫不等式

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} \Leftrightarrow P\{|X - \mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}.$$



契比雪夫资料



Pafnuty Chebyshev

**Born: 16 May. 1821 in
Okatovo, Russia**

**Died: 8 Dec. 1894, in
St. Petersburg, Russia**



第三节 协方差及相关系数

- 一、协方差与相关系数的概念及性质
- 二、相关系数的意义
- 三、小结



一、协方差与相关系数的概念及性质

1. 问题的提出

若随机变量 X 和 Y 相互独立,那么

$$D(X + Y) = D(X) + D(Y).$$

若随机变量 X 和 Y 不相互独立

$$D(X + Y) = ?$$

$$\begin{aligned} D(X + Y) &= E(X + Y)^2 - [E(X + Y)]^2 \\ &= D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\}. \end{aligned}$$

协方差



2. 定义

量 $E\{[X - E(X)][Y - E(Y)]\}$ 称为随机变量 X 与 Y 的协方差. 记为 $\text{Cov}(X, Y)$, 即

$$\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}.$$

而

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}}$$

称为随机变量 X 与 Y 的相关系数.



3. 说明

(1) X 和 Y 的相关系数又称为标准协方差,它是一个无量纲的量.

(2) 若随机变量 X 和 Y 相互独立

$$\begin{aligned}\Rightarrow \text{Cov}(X, Y) &= E\{[X - E(X)][Y - E(Y)]\} \\ &= E[X - E(X)]E[Y - E(Y)] \\ &= 0.\end{aligned}$$

(3) 若随机变量 X 和 Y 相互独立

$$\begin{aligned}\Rightarrow D(X + Y) &= D(X) + D(Y) \\ &\quad + 2E\{[X - E(X)][Y - E(Y)]\} \\ &= D(X) + D(Y) + 2\text{Cov}(X, Y) = D(X) + D(Y).\end{aligned}$$



4. 协方差的计算公式

$$(1) \text{Cov}(X, Y) = E(XY) - E(X)E(Y);$$

$$(2) D(X + Y) = D(X) + D(Y) + 2\text{Cov}(X, Y).$$

证明 (1) $\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$

$$\begin{aligned} &= E[XY - YE(X) - XE(Y) + E(X)E(Y)] \\ &= E(XY) - 2E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y). \end{aligned}$$



$$\begin{aligned}
 (2) D(X + Y) &= E\{[(X + Y) - E(X + Y)]^2\} \\
 &= E\{[(X - E(X)) + (Y - E(Y))]^2\} \\
 &= E\{[X - E(X)]^2\} + E\{[Y - E(Y)]^2\} \\
 &\quad + 2E\{[X - E(X)][Y - E(Y)]\} \\
 &= D(X) + D(Y) + 2\text{Cov}(X, Y).
 \end{aligned}$$



5. 性质

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

$$(2) \operatorname{Cov}(aX, bY) = ab \operatorname{Cov}(X, Y), \quad a, b \text{ 为常数};$$

$$(3) \operatorname{Cov}(X_1 + X_2, Y) = \operatorname{Cov}(X_1, Y) + \operatorname{Cov}(X_2, Y).$$



例1 设 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 试求 X 与 Y 的相关系数.

解 由 $f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$

$$\Rightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}, -\infty < x < +\infty,$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}, -\infty < y < +\infty.$$



$$\Rightarrow E(X) = \mu_1, E(Y) = \mu_2, D(X) = \sigma_1^2, D(Y) = \sigma_2^2.$$

而

$$\begin{aligned} \text{Cov}(X, Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_1)(y - \mu_2) f(x, y) dx dy \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_1)(y - \mu_2) \\ &\quad \cdot e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} e^{-\frac{1}{2(1-\rho^2)} \left[\frac{y-\mu_2}{\sigma_2} - \rho \frac{x-\mu_1}{\sigma_1} \right]^2} dy dx. \end{aligned}$$

$$\text{令 } t = \frac{1}{\sqrt{1-\rho^2}} \left(\frac{y-\mu_2}{\sigma_2} - \rho \frac{x-\mu_1}{\sigma_1} \right), \quad u = \frac{x-\mu_1}{\sigma_1},$$



$\text{Cov}(X, Y)$

$$\begin{aligned}
 &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\sigma_1 \sigma_2 \sqrt{1-\rho^2} tu + \rho \sigma_1 \sigma_2 u^2) e^{-\frac{u^2}{2} - \frac{t^2}{2}} dt du \\
 &= \frac{\rho \sigma_1 \sigma_2}{2\pi} \left(\int_{-\infty}^{+\infty} u^2 e^{-\frac{u^2}{2}} du \right) \left(\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt \right) \\
 &\quad + \frac{\sigma_1 \sigma_2 \sqrt{1-\rho^2}}{2\pi} \left(\int_{-\infty}^{+\infty} u e^{-\frac{u^2}{2}} du \right) \left(\int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt \right) \\
 &= \frac{\rho \sigma_1 \sigma_2}{2\pi} \sqrt{2\pi} \cdot \sqrt{2\pi},
 \end{aligned}$$

故有 $\text{Cov}(X, Y) = \rho \sigma_1 \sigma_2$.



于是

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \rho.$$

结论

(1) 二维正态分布密度函数中, 参数 ρ 代表了 X 与 Y 的相关系数;

(2) 二维正态随机变量 X 与 Y 相关系数为零等价于 X 与 Y 相互独立.



例2 已知随机变量 X, Y 分别服从 $N(1, 3^2), N(0, 4^2)$,
 $\rho_{XY} = -1/2$, 设 $Z = X/3 + Y/2$.

(1) 求 Z 的数学期望和方差.

(2) 求 X 与 Z 的相关系数.

(3) 问 X 与 Z 是否相互独立? 为什么?

解 (1) 由 $E(X) = 1, D(X) = 9, E(Y) = 0, D(Y) = 16$.

$$\begin{aligned} \text{得 } E(Z) &= E\left(\frac{X}{3} + \frac{Y}{2}\right) = \frac{1}{3}E(X) + \frac{1}{2}E(Y) \\ &= \frac{1}{3}. \end{aligned}$$



$$D(Z) = D\left(\frac{X}{3}\right) + D\left(\frac{Y}{2}\right) + 2\text{Cov}\left(\frac{X}{3}, \frac{Y}{2}\right)$$

$$= \frac{1}{9}D(X) + \frac{1}{4}D(Y) + \frac{1}{3}\text{Cov}(X, Y)$$

$$= \frac{1}{9}D(X) + \frac{1}{4}D(Y) + \frac{1}{3}\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}$$

$$= 1 + 4 - 2 = 3.$$



$$(2) \text{Cov}(X, Z) = \text{Cov}\left(X, \frac{X}{3} + \frac{Y}{2}\right)$$

$$= \frac{1}{3} \text{Cov}(X, X) + \frac{1}{2} \text{Cov}(X, Y)$$

$$= \frac{1}{3} D(X) + \frac{1}{2} \rho_{XY} \sqrt{D(X)} \sqrt{D(Y)} = 3 - 3 = 0.$$

$$\text{故 } \rho_{XY} = \text{Cov}(X, Z) / (\sqrt{D(X)} \sqrt{D(Z)}) = 0.$$

(3) 由二维正态随机变量相关系数为零和相互独立两者是等价的结论, 可知: X 与 Z 是相互独立的.



二、相关系数的意义

1. 问题的提出

问 a, b 应如何选择, 可使 $aX + b$ 最接近 Y ?
接近的程度又应如何来 衡量?

$$\text{设 } e = E[(Y - (a + bX))^2]$$

则 e 可用来衡量 $a + bX$ 近似表达 Y 的好坏程度.
当 e 的值越小, 表示 $a + bX$ 与 Y 的近似程度越好.

确定 a, b 的值, 使 e 达到最小.



$$\begin{aligned} e &= E[(Y - (a + bX))^2] \\ &= E(Y^2) + b^2 E(X^2) + a^2 - 2bE(XY) + 2abE(X) \\ &\quad - 2aE(Y). \end{aligned}$$

将 e 分别关于 a, b 求偏导数, 并令它们等于零, 得

$$\begin{cases} \frac{\partial e}{\partial a} = 2a + 2bE(X) - 2E(Y) = 0, \\ \frac{\partial e}{\partial b} = 2bE(X^2) - 2E(XY) + 2aE(X) = 0. \end{cases}$$

解得 $b_0 = \frac{\text{Cov}(X, Y)}{D(X)}, a_0 = E(Y) - E(X) \frac{\text{Cov}(X, Y)}{D(X)}.$



将 a_0, b_0 代入 $e = E[(Y - (a + bX))^2]$ 中,得

$$\begin{aligned}\min_{a,b} e &= E[(Y - (a + bX))^2] \\ &= E[(Y - (a_0 + b_0X))^2] \\ &= (1 - \rho_{XY}^2)D(Y).\end{aligned}$$

2. 相关系数的意义

当 $|\rho_{XY}|$ 较大时 e 较小, 表明 X, Y 的线性关系联系较紧密.

当 $|\rho_{XY}|$ 较小时, X, Y 线性相关的程度较差.

当 $\rho_{XY} = 0$ 时, 称 X 和 Y 不相关.



例3 设 θ 服从 $[0, 2\pi]$ 的均匀分布, $\xi = \cos \theta$, $\eta = \cos(\theta + a)$, 这里 a 是常数, 求 ξ 和 η 的相关系数?

解
$$E(\xi) = \frac{1}{2\pi} \int_0^{2\pi} \cos x \, dx = 0,$$

$$E(\eta) = \frac{1}{2\pi} \int_0^{2\pi} \cos(x + a) \, dx = 0,$$

$$E(\xi^2) = \frac{1}{2\pi} \int_0^{2\pi} \cos^2 x \, dx = \frac{1}{2},$$

$$E(\eta^2) = \frac{1}{2\pi} \int_0^{2\pi} \cos^2(x + a) \, dx = \frac{1}{2},$$



$$E(\xi\eta) = \frac{1}{2\pi} \int_0^{2\pi} \cos x \cdot \cos(x+a) dx = \frac{1}{2} \cos a,$$

由以上数据可得相关系数为 $\rho = \cos a$.

当 $a = 0$ 时, $\rho = 1, \xi = \eta$,
 当 $a = \pi$ 时, $\rho = -1, \xi = -\eta$, } 存在线性关系.

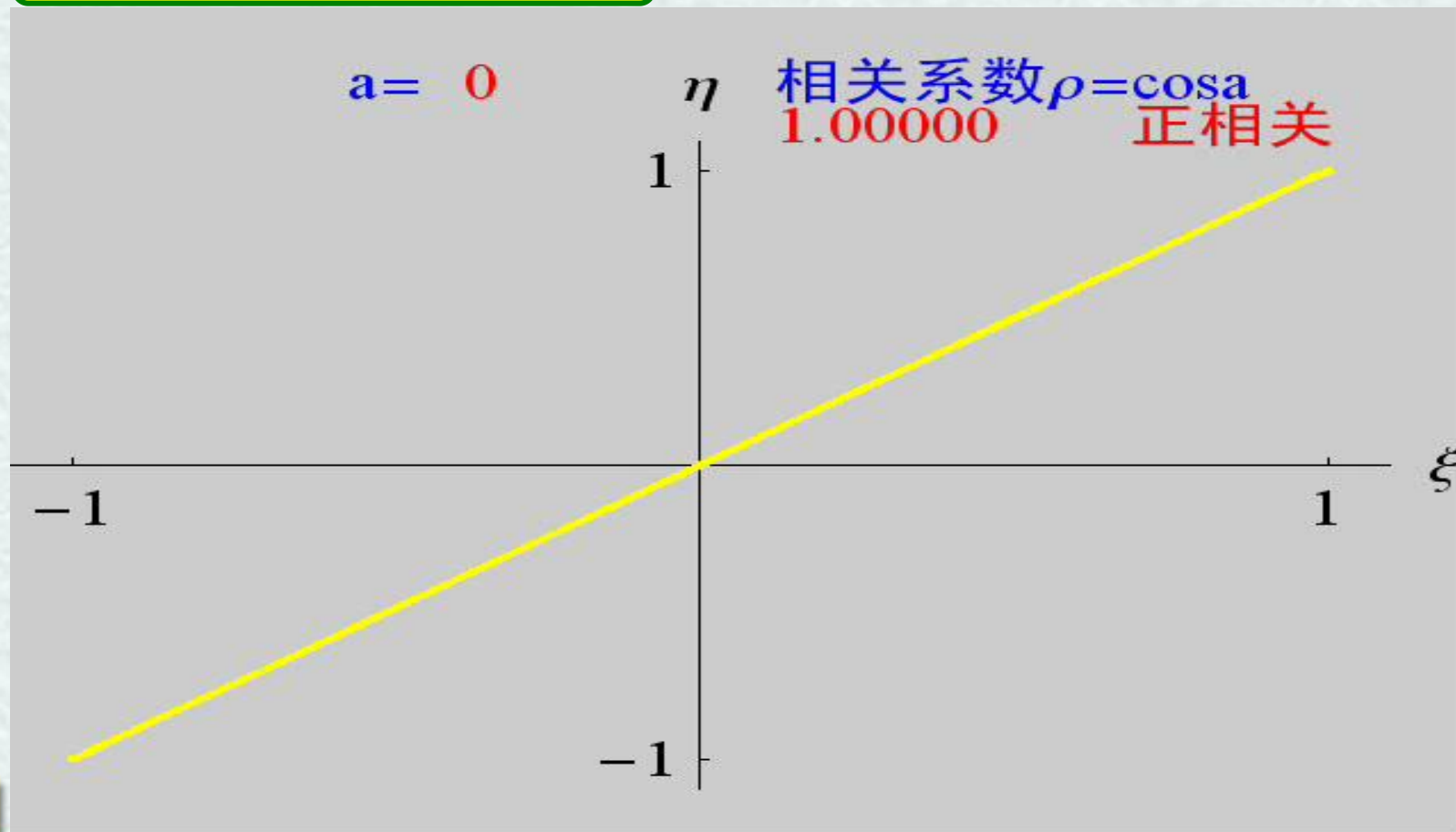
当 $a = \frac{\pi}{2}$ 或 $a = \frac{3\pi}{2}$ 时, $\rho = 0$, ξ 与 η 不相关.

但 $\xi^2 + \eta^2 = 1$, 因此 ξ 与 η 不独立.



动画演示 ξ 与 η 的相关关系.

单击图形播放/暂停 ESC键退出



3. 注意

(1) 不相关与相互独立的关系

相互独立 $\xrightarrow{\text{green}} \text{不相关}$
 $\xleftarrow{\text{red}}$

(2) 不相关的充要条件

1° X, Y 不相关 $\Leftrightarrow \rho_{XY} = 0$;

2° X, Y 不相关 $\Leftrightarrow \text{Cov}(X, Y) = 0$;

3° X, Y 不相关 $\Leftrightarrow E(XY) = E(X)E(Y)$.



4. 相关系数的性质

(1) $|\rho_{XY}| \leq 1.$

(2) $|\rho_{XY}| = 1$ 的充要条件是：存在常数 a, b 使

$$P\{Y = a + bX\} = 1.$$

证明 (1) $\min_{a,b} e = E[(Y - (a + bX))^2]$

$$= (1 - \rho_{XY}^2) D(Y) \geq 0$$

$$\Rightarrow 1 - \rho_{XY}^2 \geq 0$$

$$\Rightarrow |\rho_{XY}| \leq 1.$$



(2) $|\rho_{XY}| = 1$ 的充要条件是, 存在常数 a, b 使

$$P\{Y = a + bX\} = 1.$$

事实上, $|\rho_{XY}| = 1 \Rightarrow E[(Y - (a_0 + b_0X))^2] = 0$

$$\Rightarrow 0 = E[(Y - (a_0 + b_0X))^2]$$

$$= D[Y - (a_0 + b_0X)] + [E(Y - (a_0 + b_0X))]^2$$

$$\Rightarrow D[Y - (a_0 + b_0X)] = 0,$$

$$E[Y - (a_0 + b_0X)] = 0.$$

由方差性质知

$$P\{Y - (a_0 + b_0X) = 0\} = 1, \text{ 或 } P\{Y = a_0 + b_0X\} = 1.$$



反之,若存在常数 a^*, b^* 使

$$P\{Y = a^* + b^* X\} = 1 \Leftrightarrow P\{Y - (a^* + b^* X) = 0\} = 1,$$

$$\Rightarrow P\{[Y - (a^* + b^* X)]^2 = 0\} = 1,$$

$$\Rightarrow E\{[Y - (a^* + b^* X)]^2\} = 0.$$

故有

$$0 = E\{[Y - (a^* + b^* X)]^2\} \geq \min_{a,b} E[(Y - (a + bX))^2]$$

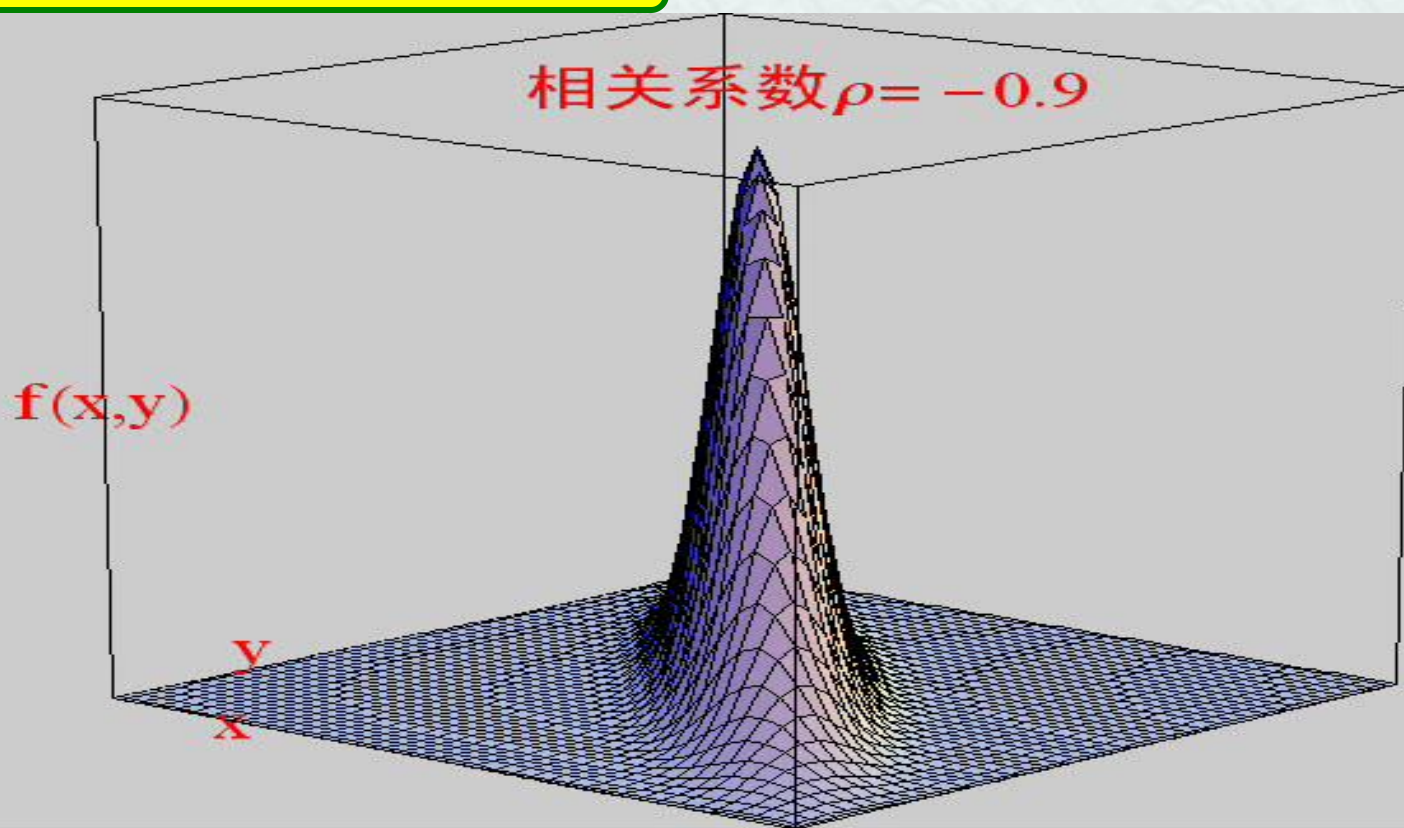
$$= E\{[Y - (a_0 + b_0 X)]^2\} = (1 - \rho_{XY}^2) D(Y)$$

$$\Rightarrow |\rho_{XY}| = 1.$$



二维正态随机变量 (X, Y) 的概率密度曲面与
相关系数 $\rho_{XY} = \rho$ 的关系.

单击图形播放/暂停 ESC键退出



三、小结

相关系数的意义

当 $|\rho_{XY}|$ 较大时, X, Y 的线性相关程度较高.

当 $|\rho_{XY}|$ 较小时, X, Y 的线性相关程度较差.

当 $\rho_{XY} = 0$ 时, X 和 Y 不相关.



第四节 矩、协方差矩阵

一、基本概念

二、 n 维正态变量的性质

三、小结



一、基本概念

1. 定义

设 X 和 Y 是随机变量, 若 $E(X^k)$, $k = 1, 2, \dots$ 存在, 称它为 X 的 k 阶原点矩, 简称 k 阶矩.

若 $E\{[X - E(X)]^k\}$, $k = 2, 3, \dots$ 存在, 称它为 X 的 k 阶中心矩.

若 $E(X^k Y^l)$, $k, l = 1, 2, \dots$ 存在, 称它为 X 和 Y 的 $k + l$ 阶混合矩.

若 $E\{[X - E(X)]^k [Y - E(Y)]^l\}$, $k, l = 1, 2, \dots$ 存在, 称它为 X 和 Y 的 $k + l$ 阶混合中心矩.



2. 说明

- (1) 以上数字特征都是随机变量函数的数学期望；
- (2) 随机变量 X 的数学期望 $E(X)$ 是 X 的一阶原点矩, 方差为二阶中心矩, 协方差 $\text{Cov}(X, Y)$ 是 X 与 Y 的二阶混合中心矩；
- (3) 在实际应用中, 高于 4 阶的矩很少使用.

三阶中心矩 $E\{[X - E(X)]^3\}$ 主要用来衡量随机变量的分布是否有偏。

四阶中心矩 $E\{[X - E(X)]^4\}$ 主要用来衡量随机变量的分布在均值附近的陡峭程度如何。



3. 协方差矩阵

设 n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的二阶混合中心矩

$$c_{ij} = \text{Cov}(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}$$
$$i, j = 1, 2, \dots, n$$

都存在, 则称矩阵

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为 n 维随机变量的协方差矩阵.



例如 二维随机变量 (X_1, X_2) 的协方差矩阵为

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

其中 $c_{11} = E\{[X_1 - E(X_1)]^2\},$

$$c_{12} = E\{[X_1 - E(X_1)][X_2 - E(X_2)]\},$$

$$c_{21} = E\{[X_2 - E(X_2)][X_1 - E(X_1)]\},$$

$$c_{22} = E\{[X_2 - E(X_2)]^2\}.$$



由于 $c_{ij} = c_{ji}$ ($i, j = 1, 2, \dots, n$), 所以协方差矩阵为对称的非负定矩阵 .

协方差矩阵的应用

协方差矩阵可用来表示多维随机变量的概率密度, 从而可通过协方差矩阵达到对多维随机变量的研究



以二维随机变量 (X_1, X_2) 为例.

由于

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}.$$

引入矩阵 $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$.

及 (X_1, X_2) 的协方差矩阵 $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$,



$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

由此可得

$$\begin{aligned} C^{-1} &= \frac{1}{\det C} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \\ &= \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}. \end{aligned}$$



由于

$$\begin{aligned}
 & (X - \mu)^T C^{-1} (X - \mu) \\
 &= \frac{1}{\det C} (x_1 - \mu_1, x_2 - \mu_2) \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} \\
 &= \frac{1}{1 - \rho^2} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right].
 \end{aligned}$$



于是 (X_1, X_2) 的概率密度可写成

$$f(x_1, x_2)$$

$$= \frac{1}{(2\pi)^{2/2} (\det C)^{1/2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\}.$$



推广

n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的概率密度可表示为 $f(x_1, x_2, \dots, x_n)$

$$= \frac{1}{(2\pi)^{n/2} (\det C)^{1/2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\}.$$

其中 $X = (x_1, x_2, \dots, x_n)^T$,

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{pmatrix}, \quad C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}.$$

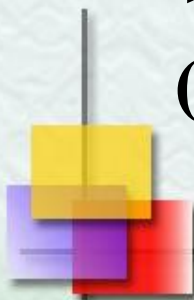


二、 n 维正态变量的性质

1. n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的每一个分量 $X_i, i = 1, 2, \dots, n$ 都是正态变量;

反之, 若 $X_1, X_2, \dots, X_n$ 都是正态变量, 且相互独立, 则 $(X_1, X_2, \dots, X_n)$ 是 n 维正态变量.

2. n 维随机变量 $(X_1, X_2, \dots, X_n)$ 服从 n 维正态分布的充要条件是 $X_1, X_2, \dots, X_n$ 的任意的线性组合 $l_1 X_1 + l_2 X_2 + \dots + l_n X_n$ 服从一维正态分布 (其中 $l_1, l_2, \dots, l_n$ 不全为零).



3.若 $(X_1, X_2, \dots, X_n)$ 服从 n 维正态分布, 设 $Y_1, \dots, Y_k$ 是 $X_j (j = 1, 2, \dots, n)$ 的线性函数, 则 $(Y_1, Y_2, \dots, Y_k)$ 也服从多维正态分布. **线性变换不变性**

4.设 $(X_1, \dots, X_n)$ 服从 n 维正态分布, 则“ $X_1, X_2, \dots, X_n$ 相互独立”与“ $X_1, X_2, \dots, X_n$ 两两不相关”是等价的.



三、小结

1. 矩是随机变量的数字特征.

随机变量 X 的

{ 数学期望 $E(X)$ 是 X 的一阶原点矩;
方差 $D(X)$ 是 X 的二阶中心矩;
协方差 $\text{Cov}(X, Y)$ 是 X 与 Y 的二阶混合中心矩.

2. 正态变量是最重要的随机变量, 其性质一定要熟练掌握.

