

目录

第四章 随机变量的数字特征	3
4.1 数学期望	3
4.2 方差	18
4.3 协方差与相关系数	26
4.4 矩 协方差矩阵	32
4.5 二维正态分布	33

第四章 随机变量的数字特征

4.1 数学期望

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

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

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 $E(X)$ ，即

$$E(X) = \sum_k x_k p_k.$$

注记。数学期望简称期望，又称均值

离散型随机变量的期望

例 1. $X \sim b(1, p)$, 求 $E(X)$.

解. 易知 X 的分布律为

X	0	1
p_k	$1-p$	p

因此 X 的数学期望为

$$E(X) = 0 \times (1-p) + 1 \times p = p.$$

离散型随机变量的期望

例 2. 设 $X \sim \pi(\lambda)$, 求 $E(X)$

解. 易知泊松分布的分布律为

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

因此 X 的数学期望为

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

离散型随机变量的期望

例 3. 甲、乙两工人每天生产出相同数量同种类型的产品, 用 X_1, X_2 分别表示甲、乙两人某天生产的次品数, 经统计得以下数据:

次品数 X_1	0	1	2	3
p_k	0.3	0.3	0.2	0.2
次品数 X_2	0	1	2	3
p_k	0.2	0.5	0.3	0

试比较他们的技术水平的高低.

解. 根据定义, X_1 和 X_2 的数学期望分别为

$$E(X_1) = 0 \times 0.3 + 1 \times 0.3 + 2 \times 0.2 + 3 \times 0.2 = 1.3,$$

$$E(X_2) = 0 \times 0.2 + 1 \times 0.5 + 2 \times 0.3 + 3 \times 0 = 1.1,$$

显然甲的技术水平比乙低.

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

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

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

每注彩票所得奖金 X 的数学期望

$$\begin{aligned}
 E(X) &= 2p_2 + 50p_3 + 1500p_4 + 40000p_5 + 2000000p_6 \\
 &= 0.292 + 1.055 + 1.8735 + 1.0248 + 0.2134 \\
 &= 4.4587.
 \end{aligned}$$

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 $E(X)$, 即

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

连续型随机变量的期望

例 4. 设随机变量 X 在区间 (a, b) 内服从均匀分布, 求 $E(X)$.

解. 由题意知, X 的概率密度为

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

于是有

$$E(X) = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}.$$

连续型随机变量的期望

例 5. 设随机变量 X 服从参数为 $\lambda (\lambda > 0)$ 的指数分布, 求 $E(X)$.

解. 由题意知, X 的概率密度为

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

于是有

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

连续型随机变量的期望

例 6. 设随机变量 X 的概率密度为

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

求 X 的数学期望 $E(X)$.

解. 因为 $xf(x)$ 是奇函数, 所以 $E(X) = 0$.

连续型随机变量的期望

例 7. 设随机变量 X 服从柯西 (Cauchy) 分布, 其概率密度为

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

求 $E(X)$.

解. 因为反常积分 $\int_{-\infty}^{+\infty} \frac{|x|}{(x^2 + 1)} dx$ 不收敛, 所以 $E(X)$ 不存在.

二维随机变量的数学期望

对二维随机变量 (X, Y) , 定义它们的数学期望为

$$E(X, Y) = (E(X), E(Y))$$

设二维离散型随机变量 (X, Y) 的联合分布律为

$$P(X = x, Y = y) = p_{ij}, \quad i, j = 1, 2, \dots,$$

则

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

$$E(Y) = \sum_{j=1}^{+\infty} y_j p_{.j} = \sum_{i=1}^{+\infty} \sum_{j=1}^{+\infty} y_j p_{ij}.$$

二维随机变量的数学期望

设二维连续型随机变量 (X, Y) 的联合概率密度为 $f(x, y)$, 则

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

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

例 8. 设二维连续型随机变量 (X, Y) 的概率密度为

$$f(x, y) = \begin{cases} 12y^2, & 0 \leq y \leq x \leq 1, \\ 0, & \text{其他.} \end{cases}$$

求 $E(X), E(Y)$.

解. 由条件易得

$$E(X) = \iint_D x f(x, y) d\sigma = \int_0^1 x dx \int_0^x 12y^2 dy = \frac{4}{5}.$$

$$E(Y) = \iint_D y \cdot 12y^2 d\sigma = \int_0^1 dx \int_0^x 12y^3 dy = \frac{3}{5}.$$

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 $E(Y)$?

直观思路: 根据 X 的分布算出 Y 的分布, 然后利用定义计算 $E(Y)$. 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

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

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

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

随机变量函数的数学期望

例 9. 设随机变量 X 的分布律为

X	-2	-1	0	1	2	3
p_k	0.10	0.20	0.25	0.20	0.15	0.10

求随机变量函数 $Y = X^2$ 的数学期望.

解 (方法 1). 先求 Y 的分布律

Y	0	1	4	9
p_k	0.25	0.40	0.25	0.10

因此 $E(Y) = 0 \times 0.25 + 1 \times 0.40 + 4 \times 0.25 + 9 \times 0.10 = 2.30$.

解 (方法 2). 由条件易得,

$$\begin{aligned} E(Y) &= (-2)^2 \times 0.10 + (-1)^2 \times 0.20 + 0^2 \times 0.25 + 1^2 \times 0.20 \\ &\quad + 2^2 \times 0.15 + 3^2 \times 0.10 \\ &= 2.30 \end{aligned}$$

随机变量函数的数学期望

例 10. 设随机变量 X 在区间 $(0, \pi)$ 内服从均匀分布, 求随机变量函数 $Y = \sin X$ 的数学期望.

解 (方法 1). 先利用分布函数法求得 Y 的概率密度为

$$f_Y(y) = \begin{cases} \frac{2}{\pi\sqrt{1-y^2}}, & 0 < y < 1, \\ 0, & \text{其他.} \end{cases}$$

因此

$$E(Y) = \int_0^1 y \cdot \frac{2}{\pi\sqrt{1-y^2}} dy = \frac{2}{\pi}.$$

解 (方法 2). 由题意知, X 的概率密度为

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

因此

$$E(Y) = \int_0^{\pi} \sin x \cdot \frac{1}{\pi} dx = \frac{2}{\pi}.$$

随机变量函数的数学期望

例 11. 设随机变量 $X \sim N(0, 1)$, 求 $E(X)$ 和 $E(X^2)$.

解. X 为连续型随机变量, 其概率密度为

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

所以

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0,$$

随机变量函数的数学期望

解(续).

$$\begin{aligned}
 E(X^2) &= \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \\
 &= -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x d\left(e^{-\frac{x^2}{2}}\right) \\
 &= -\frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} \Big|_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \\
 &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 1.
 \end{aligned}$$

随机变量函数的数学期望

例 12. 按季节出售的某种应时商品, 每售出 1 kg 获利润 6 元, 如到季末尚有剩余商品, 则每千克净亏损 2 元, 设某商店在季节内这种商品的销售量 X (以 kg 计) 是一随机变量, X 在区间 $(8, 16)$ 内服从均匀分布, 为使商店所获得利润最大, 问商店应进多少货?

随机变量函数的数学期望

解. 设 t 表示进货量, 易知应取 $8 < t < 16$, 进货 t 所得利润记为 $W_t(X)$, 且有

$$W_t(X) = \begin{cases} 6X - 2(t - X), & 8 < X < t \text{ (有积压)} \\ 6t, & t < X < 16 \text{ (无积压)} \end{cases}$$

利润 $W_t(X)$ 是随机变量, 如何获得最大利润? 自然是取“平均利润”的最大值, 即求 t 使 $E[W_t(X)]$ 最大. X 的概率密度为

$$f(x) = \begin{cases} \frac{1}{8}, & 8 < x < 16 \\ 0, & \text{其他.} \end{cases}$$

随机变量函数的数学期望

解. 因此

$$\begin{aligned} E[W_t(X)] &= \int_{-\infty}^{+\infty} W_t(x)f(x) dx = \frac{1}{8} \int_8^{16} W_t(x) dx \\ &= \frac{1}{8} \int_8^t [6x - 2(t-x)] dx + \frac{1}{8} \int_t^{16} 6t dx \\ &= \frac{1}{8} \left[\int_8^t (8x - 2t) dx + 6t(16-t) \right] = 14t - \frac{t^2}{2} - 32. \end{aligned}$$

令

$$\frac{dE[W_t(X)]}{dt} = 14 - t = 0$$

得 $t = 14$. 而

$$\frac{d^2E[W_t(X)]}{dt^2} = -1 < 0$$

故知当 $t = 14$ 时, $E[W_t(X)]$ 取极大值, 且可知这也是最大值.

二维随机变量函数的数学期望

定理. 设 (X, Y) 为二维随变量, $Z = g(X, Y)$, 则

1. 若 (X, Y) 为二维离散型随机变量, 概率分布为

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

则

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

2. 若 (X, Y) 为二维连续型随机变量, 概率密度为 $f(x, y)$, 则

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

二维随机变量函数的数学期望

例 13. 设随机变量 X 与 Y 相互独立, 概率密度分别是

$$f_X(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & x \leq 0, \end{cases} \quad f_Y(y) = \begin{cases} e^{-y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

求随机变量函数 $Z = X + Y$ 的数学期望.

解 (方法 1). 首先求出 Z 的概率密度为

$$f_Z(z) = \begin{cases} ze^{-z}, & z > 0, \\ 0, & \text{其他.} \end{cases}$$

因此

$$E(Z) = \int_0^{+\infty} z \cdot ze^{-z} dz = 2.$$

二维随机变量函数的数学期望

解 (方法 2). 因为随机变量 X 与 Y 是相互独立的, 所以二维随机变量 (X, Y) 的联合概率密度为

$$f(x, y) = f_X(x)f_Y(y) = \begin{cases} e^{-x-y}, & x > 0, y > 0, \\ 0, & \text{其他.} \end{cases}$$

因此

$$\begin{aligned} E(Z) &= E(X + Y) = \int_0^{+\infty} \int_0^{+\infty} (x + y)e^{-x-y} dx dy \\ &= \int_0^{+\infty} xe^{-x} dx \int_0^{+\infty} e^{-y} dy + \int_0^{+\infty} e^{-x} dx \int_0^{+\infty} ye^{-y} dy \\ &= 1 + 1 = 2. \end{aligned}$$

数学期望的性质

设 X, X_1, X_2, \dots, 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

4. 若 X_1, X_2 相互独立, 则有

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

推论: 若 X_1, X_2, \dots, 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!

数学期望的性质

例 14. 设随机变量 $X \sim N(\mu, \sigma^2)$, 求 $E(X)$.

解. 令 $Y = \frac{X - \mu}{\sigma}$, 则 $Y \sim N(0, 1)$, 于是

$$E(Y) = 0$$

而

$$X = \sigma Y + \mu$$

所以

$$E(X) = E(\sigma Y + \mu) = \sigma E(Y) + E(\mu) = \mu$$

数学期望的性质

例 15. 设随机变量 $X \sim b(n, p)$, 求 $E(X)$.

解. 引入随机变量

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 次试验中事件 } A \text{ 发生,} \\ 0, & \text{第 } i \text{ 次试验中事件 } A \text{ 不发生,} \end{cases} \quad i = 1, 2, \dots, n,$$

其中, $P(A) = p$, 则 X_i 服从 $(0-1)$ 分布, 于是 $E(X_i) = p$, 又 $X = \sum_{i=1}^n X_i$, 所以

$$\begin{aligned} E(X) &= E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) \\ &= p + p + \dots + p = np \end{aligned}$$

即 X 的数学期望为 np .

数学期望的性质

例 16. 一民航公司的客车载有 20 位旅客自机场开出, 旅客有 10 个车站可以下车, 若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.)

数学期望的性质

解. 引随机变量 $X_i = \begin{cases} 0, & \text{在第 } i \text{ 站没有人下车,} \\ 1, & \text{在第 } i \text{ 站有人下车} \end{cases}, i = 1, 2, \dots, 10$. 易知 $X = X_1 + X_2 + \dots + X_{10}$, 现求 $E(X)$. 依题意

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

因此

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

从而

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

数学期望的性质

练习 1. 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.

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

连续型随机变量的期望

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

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	$X \sim \pi(\lambda)$	$E(X) = \lambda$

4.2 方差

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

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

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

方差

方差刻画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, $D(X) = 0$ 当且仅当 X 取某个常数的概率为 1.

方差

方差的常用计算公式： $D(X) = E(X^2) - [E(X)]^2$.

证明. 由期望的性质易得

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

方差

例 1. 设随机变量 $X \sim (0-1)$ 分布, 求 $D(X)$.

解. X 的分布律为

$$P\{X=0\} = 1-p, P\{X=1\} = p.$$

显然

$$E(X) = p,$$

又 $E(X^2) = 0^2 \cdot (1-p) + 1^2 \cdot p = p$, 故

$$D(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p).$$

方差

例 2. 设随机变量 $X \sim \pi(\lambda), \lambda > 0$, 求 $D(X)$.

解. X 的分布律为

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

于是 $E(X) = \lambda$, 且

$$\begin{aligned} E(X^2) &= E[X(X-1) + X] = E[X(X-1)] + E(X) \\ &= \sum_{k=0}^{\infty} k(k-1) \cdot \frac{\lambda^k e^{-\lambda}}{k!} + \lambda = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda \\ &= \lambda^2 e^{\lambda} e^{-\lambda} + \lambda = \lambda^2 + \lambda, \end{aligned}$$

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

方差

例 3. 设随机变量 $X \sim U(a, b)$, 求 $D(X)$.

解. X 的概率密度为

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

则 $E(X) = \frac{a+b}{2}$, 而

$$E(X^2) = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{1}{3}(a^2 + ab + b^2),$$

所以

$$D(X) = E(X^2) - [E(X)]^2 = \frac{1}{3}(a^2 + ab + b^2) - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

方差

例 4. 设随机变量 $X \sim E(\lambda), \lambda > 0$, 求 $D(X)$.

解. 由条件, X 的概率密度为

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{其他.} \end{cases}$$

则 $E(X) = \frac{1}{\lambda}$, 而

$$\begin{aligned} E(X^2) &= \int_0^{+\infty} x^2 \cdot \lambda e^{-\lambda x} dx = - \int_0^{+\infty} x^2 d(e^{-\lambda x}) \\ &= -x^2 e^{-\lambda x} \Big|_0^{+\infty} + 2 \int_0^{+\infty} x e^{-\lambda x} dx = \frac{2}{\lambda^2} \end{aligned}$$

所以, $D(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$.

方差的性质

设 X, Y 为随机变量, a, b, c 为常数, 则有

1. $D(c) = 0$

2. $D(aX + b) = a^2 D(X)$

3. $D(X \pm Y) = D(X) + D(Y) \pm 2E\{[X - E(X)][Y - E(Y)]\}$ 特别地, 若 X 和 Y 相互独立, 则有

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

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

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

方差的性质

例 5. 设随机变量 $X \sim b(n, p)$, 求 $D(X)$.

解. 设 X 表示 n 重伯努利试验中事件 A 发生的次数, 且在每次试验中 A 发生的概率为 p . 设 X_i 表示第 i 次试验中事件 A 发生的次数, 则 X_i 服从 $(0-1)$ 分布, 于是

$$E(X_i) = p, D(X_i) = p(1-p), i = 1, 2, \dots, n,$$

而 X_1, X_2, \dots, X_n 相互独立, 且

$$X = X_1 + X_2 + \dots + X_n,$$

所以

$$\begin{aligned} D(X) &= D(X_1 + X_2 + \dots + X_n) = D(X_1) + D(X_2) + \dots + D(X_n) \\ &= p(1-p) + p(1-p) + \dots + p(1-p) = np(1-p). \end{aligned}$$

方差的性质

例 6. 设随机变量 $X \sim N(\mu, \sigma^2)$, 求 $D(X)$.

解. 由于 $\frac{X-\mu}{\sigma} \sim N(0, 1)$, 令 $Y = \frac{X-\mu}{\sigma}$, 则有

$$Y \sim N(0, 1).$$

易知 (前面结论) $E(Y) = 0, E(Y^2) = 1$, 由方差的性质可得

$$D(X) = D(\sigma Y + \mu) = \sigma^2 D(Y) = \sigma^2.$$

注记. 正态分布的概率密度中的两个参数 μ 和 σ 分别就是该分布的数学期望和均方差, 因而正态分布完全可由它的数学期望和方差确定.

方差的性质

若 $X_i \sim N(\mu_i, \sigma_i^2), i = 1, 2, \dots, n$, 且它们相互独立, 则它们的线性组合: $c_1X_1 + c_2X_2 + \dots + c_nX_n$ (c_1, c_2, \dots, c_n 是不全为 0 的常数) 仍然服从正态分布, 于是由数学期望和方差的性质知道:

$$c_1X_1 + c_2X_2 + \dots + c_nX_n \sim N\left(\sum_{i=1}^n c_i\mu_i, \sum_{i=1}^n c_i^2\sigma_i^2\right)$$

方差的性质

例 7. 设活塞的直径 (以 cm 计) $X \sim N(22.40, 0.03^2)$, 汽缸的直径 $Y \sim N(22.50, 0.04^2)$, X 与 Y 相互独立, 任取一只活塞, 任取一只汽缸, 求活塞能装入汽缸的概率.

解. 由题意, 只要求 $P\{X < Y\} = P\{X - Y < 0\}$ 即可. 由于

$$X - Y \sim N(-0.10, 0.0025),$$

故有

$$\begin{aligned} P\{X < Y\} &= P\{X - Y < 0\} \\ &= P\left\{\frac{(X - Y) - (-0.10)}{\sqrt{0.0025}} < \frac{0 - (-0.10)}{\sqrt{0.0025}}\right\} \\ &= \Phi\left(\frac{0.1}{0.05}\right) = \Phi(2) = 0.9772. \end{aligned}$$

方差的性质

例 8. 设随机变量 X_1, X_2, \dots, X_n 相互独立, 且具有相同的数学期望 μ 和方差 σ^2 , 令

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

求 $E(\bar{X}), D(\bar{X}), E(S^2)$.

解. 由条件易知 $E(X_i) = \mu, D(X_i) = \sigma^2, i = 1, 2, \dots, n$,

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu \\ D(\bar{X}) &= D\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} D\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}. \end{aligned}$$

方差的性质

由于

$$\begin{aligned}
 \sum_{i=1}^n (X_i - \bar{X})^2 &= \sum_{i=1}^n (X_i^2 - 2X_i\bar{X} + \bar{X}^2) \\
 &= \sum_{i=1}^n X_i^2 - 2\bar{X} \left(\sum_{i=1}^n X_i \right) + \sum_{i=1}^n \bar{X}^2 \\
 &= \sum_{i=1}^n X_i^2 - 2\bar{X}(n\bar{X}) + n\bar{X}^2 \\
 &= \sum_{i=1}^n X_i^2 - n\bar{X}^2
 \end{aligned}$$

方差的性质

而

$$E(X_i^2) = D(X_i) + [E(X_i)]^2 = \sigma^2 + \mu^2,$$

$$E(\bar{X}^2) = D(\bar{X}) + [E(\bar{X})]^2 = \frac{\sigma^2}{n} + \mu^2$$

所以

$$\begin{aligned}
 E(S^2) &= E \left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{1}{n-1} E \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right) \\
 &= \frac{1}{n-1} \left[\sum_{i=1}^n E(X_i^2) - nE(\bar{X}^2) \right] \\
 &= \frac{1}{n-1} \left[n(\sigma^2 + \mu^2) - n \left(\frac{\sigma^2}{n} + \mu^2 \right) \right] = \sigma^2
 \end{aligned}$$

常用随机变量的方差

随机变量	X	$E(X)$	$D(X)$
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	$\pi(\lambda)$	λ	λ
均匀分布	$U(a, b)$	$(a+b)/2$	$(b-a)^2/12$
指数分布	$E(\lambda)$	$1/\lambda$	$1/\lambda^2$
正态分布	$N(\mu, \sigma^2)$	μ	σ^2

常用随机变量的方差

练习 1. 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.

答案. $X = X_1 + X_2 + X_3$, 其中 $X_i \sim B(1, p_i)$, 求得 $E(X) = 0.06$, $D(X) = 0.0586$.

切比雪夫不等式

定理. 设随机变量 X 有期望和方差, 则对于任给的 $\varepsilon > 0$, 有

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2},$$

即

$$P(|X - E(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}.$$

证明. 设 X 是一个连续随机变量, 其概率密度函数为 $f(x)$, 则

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

切比雪夫不等式

例 9. 设电站供电网有 10000 盏电灯, 夜晚每一盏灯开灯的概率都是 0.7, 假定开、关时间彼此独立, 估计夜间同时使用的灯的盏数在 6800 与 7200 之间的概率.

解. 令 X 表示在夜间同时使用的灯的盏数, 它服从参数 $n = 10000$, $p = 0.7$ 的二项分布, 若要准确计算, 应该用如下公式来计算:

$$P\{6800 < X < 7200\} = \sum_{k=6801}^{7199} C_{10000}^k \times 0.7^k \times 0.3^{10000-k},$$

显然这是比较困难的.

解. 利用切比雪夫不等式估计:

$$E(X) = np = 10000 \times 0.7 = 7000,$$

$$D(X) = np(1-p) = 10000 \times 0.7 \times 0.3 = 2100,$$

$$P\{6800 < X < 7200\} = P\{|X - 7000| < 200\}$$

$$\geq 1 - \frac{2100}{200^2} \approx 0.95.$$

4.3 协方差与相关系数

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. 对于二维随机向量 (X, Y) , 称

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

为 X 与 Y 的协方差(Covariance).

由定义可得: 对任意两个随机变量 X 和 Y , 有

$$\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y).$$

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

设 X, Y, Z 为随机变量, a, b, c, d 为常数, 则有

1. $\text{Cov}(X, X) = D(X)$;
2. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
3. $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$;
4. $\text{Cov}(X, C) = 0$, C 为任意常数;
5. $\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$;
6. 如果 X 和 Y 相互独立, 则 $\text{Cov}(X, Y) = 0$;
7. $D(X \pm Y) = D(X) + D(Y) \pm 2 \text{Cov}(X, Y)$.

协方差的性质

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

证明. 等式右边各项为

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

$$D(Y) = E(Y^2) - E(Y)^2$$

$$2 \operatorname{Cov}(X, Y) = 2E(XY) - 2E(X) \cdot E(Y)$$

而等式左边为

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

比较等式两边可知等式成立.

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 $\operatorname{Cov}(X - Y, Y)$.

协方差

答案.

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{2}{3} - \frac{2}{3} \cdot 1 = 0,$$

$$D(Y) = E(Y^2) - E(Y)^2 = \frac{5}{3} - 1 = \frac{2}{3},$$

$$\text{Cov}(X - Y, Y) = \text{Cov}(X, Y) - D(Y) = -\frac{2}{3}.$$

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

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

为 X 与 Y 的**相关系数**(Correlation), 也可以记为 $\rho(X, Y)$.

令

$$X^* = \frac{X - E(X)}{\sqrt{D(X)}}, \quad Y^* = \frac{Y - E(Y)}{\sqrt{D(Y)}}$$

称 X^*, Y^* 分别为 X, Y 的**标准化随机变量**, 易知

$$E(X^*) = 0, D(X^*) = 1, E(Y^*) = 0, D(Y^*) = 1,$$

$$\rho_{XY} = \text{Cov}(X^*, Y^*) = E(X^*Y^*)$$

相关系数

性质. 相关系数表示随机变量之间的**线性相关程度**:

1. $|\rho_{XY}| \leq 1$.
2. $|\rho_{XY}| = 1$ 当且仅当存在常数 a, b , 使得 $P\{Y = aX + b\} = 1$.

定义 2. 当 $\rho_{XY} = 0$ 时, 称 X 与 Y 不相关; 当 $|\rho_{XY}| = 1$ 时, 称 X 与 Y 完全相关

性质. 相互独立 \implies 不相关; 反之未必成立.

例 1. 设二维随机变量 (X, Y) 的联合概率分布如右表所示, 证明: X 与 Y 不相关, 但不相互独立.

$Y \backslash X$	-1	0	1
0	0	1/3	0
1	1/3	0	1/3

证明. 易知 X 与 Y 的边缘概率分布分别是

X	-1	0	1
p_i	1/3	1/3	1/3

Y	0	1
p_j	1/3	2/3

$$\begin{aligned} \text{Cov}(X, Y) &= (-1) \times 1 \times \frac{1}{3} + 0 \times 0 \times \frac{1}{3} + 1 \times 1 \times \frac{1}{3} \\ &\quad - \left[(-1) \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} \right] \left[0 \times \frac{1}{3} + 1 \times \frac{2}{3} \right] = 0 \end{aligned}$$

所以 X 与 Y 不相关. 由 $p_{00} = \frac{1}{3} \neq p_{0 \cdot} p_{\cdot 0} = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$ 知 X 与 Y 不是相互独立的

相关系数

例 2. 设 X 服从 $(-\pi, \pi)$ 上的均匀分布, $X_1 = \sin X, X_2 = \cos X$, 求 $\rho_{X_1 X_2}$.

相关系数

解. 随机变量 X 的概率密度为 $f(x) = \begin{cases} \frac{1}{2\pi}, & x \in (-\pi, \pi), \\ 0, & \text{其他}. \end{cases}$ 因此

$$E(X_1) = E(\sin X) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin x \, dx = 0,$$

$$E(X_2) = E(\cos X) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos x \, dx = 0,$$

$$E(X_1 X_2) = E(\sin X \cos X) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sin x \cos x \, dx = 0.$$

所以 $\text{Cov}(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2) = 0$, 故 $\rho_{X_1 X_2} = 0$.

注记. X_1, X_2 不相关, 但 $X_1^2 + X_2^2 = 1$.

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：

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

例子. 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 $D(X + Y)$ 和 $D(X - Y)$.

解. $D(X + Y) = 3$, $D(X - Y) = 7$.

例子 (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 $\sigma_1^2 = 0.25$ 、 $\sigma_2^2 = 0.49$ 、 $\rho = 0.6$.

投资风险组合

投资组合的收益 $Z = tX + (1 - t)Y$, 则平均收益为

$$E(Z) = t\mu_1 + (1 - t)\mu_2.$$

投资风险为

$$\begin{aligned} D(Z) &= D(tX + (1 - t)Y) \\ &= t^2D(X) + (1 - t)^2D(Y) + 2t(1 - t)\text{Cov}(X, Y) \\ &= t^2\sigma_1^2 + (1 - t)^2\sigma_2^2 + 2t(1 - t)\rho\sigma_1\sigma_2 \end{aligned}$$

当 $t = 87.5\%$ 时, 函数有最小值, 此时风险最小.

4.4 矩 协方差矩阵

矩与中心矩

定义：设 X 和 Y 是随机变量, k, l 为正整数

1. 称 $E(X^k)$ 为 X 的 k 阶原点矩,
2. 称 $E[(X - EX)^k]$ 为 X 的 k 阶中心矩,
3. 称 $E(X^k Y^l)$ 为 X 和 Y 的 $k + l$ 阶混合原点矩,
4. 称 $E\{[X - E(X)]^k [Y - E(Y)]^l\}$ 为 X 和 Y 的 $k + l$ 阶混合中心矩.

例:

1. 期望 $E(X)$ 为 X 的一阶原点矩,
2. 方差 $D(X)$ 为 X 的二阶中心矩.

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

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

都存在, 则称矩阵

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

为 n 维随机变量 (X_1, X_2, \dots, X_n) 的协方差矩阵.

易知上述矩阵对称.

4.5 二维正态分布

二维正态分布

定义. 设二维随机变量 (X, Y) 的联合概率密度为

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

其中 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$ 都是常数, 且 $\sigma_1 > 0, \sigma_2 > 0, -\infty < \mu_1, \mu_2 < +\infty, -1 < \rho < 1$, 我们称 (X, Y) 服从参数为 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$ 的二维正态分布, 记为:

$$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho).$$

二维正态分布

可以证明: 参数 μ_1, μ_2 是随机变量 X 和 Y 的数学期望, 参数 σ_1, σ_2 分别是它们的标准差, ρ 是它们的相关系数.

证明. 首先计算随机变量 X 的边缘概率密度:

$$f_X(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} e^{-u(x,y)} dy$$

其中

$$\begin{aligned} u(x, y) &= \frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \\ &= \frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{1}{2(1-\rho^2)} \left[\frac{y-\mu_2}{\sigma_2} - \frac{\rho(x-\mu_1)}{\sigma_1} \right]^2 \end{aligned}$$

二维正态分布

置换积分变量 $\frac{1}{\sqrt{1-\rho^2}} \left[\frac{y-\mu_2}{\sigma_2} - \frac{\rho(x-\mu_1)}{\sigma_1} \right] = t$ 得到

$$f_X(x) = \frac{1}{2\pi\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

由对称性得 $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$

故二维正态分布的两个边缘分布都是正态分布, 且有

$$\mu_1 = E(X), \mu_2 = E(Y), \sigma_1 = \sqrt{D(X)}, \sigma_2 = \sqrt{D(Y)}.$$

二维正态分布

可以证明, 参数 ρ 是随机变量 X 与 Y 的相关系数

$$\rho_{XY} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{x-\mu_1}{\sigma_1} \cdot \frac{y-\mu_2}{\sigma_2} e^{-u(x,y)} dx dy$$

化为二次积分, 得

$$\rho_{XY} = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \frac{x-\mu_1}{\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} I(x) dx$$

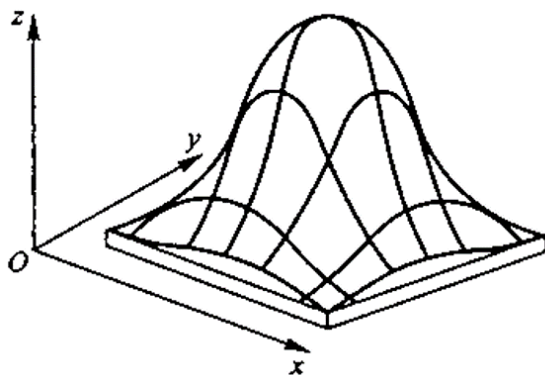
其中

$$I(x) = \int_{-\infty}^{+\infty} \frac{y-\mu_2}{\sigma_2} e^{-\frac{1}{2(1-\rho^2)} \left[\frac{y-\mu_2}{\sigma_2} - \frac{\rho(x-\mu_1)}{\sigma_1} \right]^2} dy$$

二维正态分布

设 $\frac{1}{\sqrt{1-\rho^2}} \left[\frac{y-\mu_2}{\sigma_2} - \frac{\rho(x-\mu_1)}{\sigma_1} \right] = t$, 则得

$$\begin{aligned} I(x) &= \sigma_2 \sqrt{1-\rho^2} \int_{-\infty}^{+\infty} \left[t\sqrt{1-\rho^2} + \frac{\rho(x-\mu_1)}{\sigma_1} \right] e^{-\frac{t^2}{2}} dt \\ &= \sigma_2 (1-\rho^2) \int_{-\infty}^{+\infty} t e^{-\frac{t^2}{2}} dt + \frac{\rho\sigma_2(x-\mu_1)}{\sigma_1} \sqrt{1-\rho^2} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt \\ &= \frac{\rho\sigma_2(x-\mu_1)}{\sigma_1} \sqrt{2\pi(1-\rho^2)} \end{aligned}$$



$$\rho_{XY} = \frac{\rho}{\sqrt{2\pi}\sigma_1} \int_{-\infty}^{+\infty} \left(\frac{x-\mu_1}{\sigma_1}\right)^2 e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} dx$$

二维正态分布

设 $\frac{x-\mu_1}{\sigma_1} = t$, 则得 $\rho_{XY} = \frac{\rho}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-\frac{t^2}{2}} dt = \rho$

二维正态分布曲面图所示

二维正态分布

当相关系数 $\rho = 0$ 时

$$\begin{aligned} f(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} = f_X(x)f_Y(y) \end{aligned}$$

注记. 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 $\rho_{X, Y} = 0$.

n 维正态分布

若二维随机变量 $(X_1, X_2) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 则

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

引入下面的列向量和矩阵

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

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

其中

$$c_{11} = \text{COV}(X_1, X_1), c_{12} = c_{21} = \text{COV}(X_1, X_2), c_{22} = \text{COV}(X_2, X_2)$$

称矩阵 C 为随机变量 (X_1, X_2) 的协方差矩阵

$$|C| = \sigma_1^2\sigma_2^2(1-\rho^2), C^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}$$

经过计算得

$$(X - \mu)^T C^{-1} (X - \mu) = \frac{1}{(1 - \rho^2)} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1 \sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right]$$

于是随机变量 (X_1, X_2) 的联合概率密度可写成

$$f(x_1, x_2) = (2\pi)^{-\frac{2}{2}} |C|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\}$$

将公式

$$f(x_1, x_2) = (2\pi)^{-\frac{2}{2}} |C|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\}$$

推广到 n 维随机变量 (X_1, X_2, \dots, X_n) 的情形引入列向量和矩阵:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

其中 C 为对称正定矩阵

n 维正态分布

如 n 维随机变量 (X_1, X_2, \dots, X_n) 的联合概率密度为

$$f(x_1, x_2, \dots, x_n) = (2\pi)^{-\frac{n}{2}} |C|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\}$$

则称 (X_1, X_2, \dots, X_n) 服从 n 维正态分布, C 为协方差矩阵, 其中

$$c_{ij} = \text{cov}(X_i, X_j), (i, j = 1, \dots, n), \mu_i = E(X_i) (i = 1, \dots, n)$$

n 维正态随机变量的重要性质

- 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
- 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 0)

- 若 (X_1, X_2, \dots, X_n) 服从 n 维正态分布, 设 Y_1, Y_2, \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. 这一性质称为正态变量的线性变换不变性
- 设 (X_1, X_2, \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

例 1. 已知 (X, Y) 的联合概率密度为

$$f(x, y) = \frac{1}{2\pi \cdot 5^2} e^{-\frac{1}{2} \left(\frac{x^2}{5^2} + \frac{y^2}{5^2} \right)} \quad (-\infty < x, y < +\infty)$$

问 X, Y 是否独立?

解. 由于 e 的指数式中不含 x 与 y 相乘的交叉项, 故 $\rho = 0$ 又显然 $\sigma_1 = \sigma_2 = 5, \mu_1 = \mu_2 = 0$, 故 (X, Y) 服从二维正态分布. 因 $\rho = 0$, 故 X 与 Y 相互独立.

例 2. 设随机变量 X 与 Y 相互独立, 且都服从正态分布 $N(0, \sigma^2)$ 记 $U = \alpha X + \beta Y, V = \alpha X - \beta Y$ (α, β 为不相等的常数) 求

1. U 与 V 的相关系数 ρ_{UV}

2. U 与 V 的相互独立的条件

解. (1) 由条件易知:

$$E(X) = E(Y) = 0, D(X) = D(Y) = \sigma^2$$

故

$$E(U) = E(\alpha X + \beta Y) = 0, E(V) = 0$$

$$D(U) = D(\alpha X + \beta Y) = \alpha^2 D(X) + \beta^2 D(Y) = (\alpha^2 + \beta^2) \sigma^2$$

$$D(V) = D(\alpha X - \beta Y) = \alpha^2 D(X) + (-\beta)^2 D(Y) = (\alpha^2 + \beta^2) \sigma^2$$

解. 因此

$$\begin{aligned} \text{Cov}(U, V) &= E(UV) - E(U)E(V) \\ &= E(UV) = E[(\alpha X + \beta Y)(\alpha X - \beta Y)] \\ &= E(\alpha^2 X^2 - \beta^2 Y^2) = \alpha^2 E(X^2) - \beta^2 E(Y^2) \\ &= \alpha^2 D(X) - \beta^2 D(Y) = (\alpha^2 - \beta^2) \sigma^2 \end{aligned}$$

最后得 $\rho_{UV} = (\alpha^2 - \beta^2)/(\alpha^2 + \beta^2)$.

(2) 因为 X 与 Y 是相互独立的随机变量, 故 X 与 Y 的线性组合的随机变量 U, V 也是正态随机变量. 由于 U 与 V 相互独立的充要条件是 $\rho_{UV} = 0$ 所以 $\frac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2} = 0$, 即 $\alpha = -\beta$

设二维正态随机变量 $X = (X_1, X_2)$ 的均值为 $E(X) = (0, 1)$, 协方差矩阵为 $C = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$

试计算

1. $D(2X_1 - X_2)$

2. $E(X_1^2 - X_1X_2 + X_2^2)$

解. 由条件易知

$$D(X_1) = D(X_2) = 1, \quad \text{Cov}(X_1, X_2) = 0.5$$

故

$$\begin{aligned} D(2X_1 - X_2) &= D(2X_1) + D(X_2) - 2 \text{cov}(2X_1, X_2) \\ &= 4D(X_1) + D(X_2) - 4 \text{cov}(X_1, X_2) = 3. \end{aligned}$$

$$E(X_1^2) = D(X_1) + (E(X_1))^2 = 1$$

$$E(X_2^2) = D(X_2) + (E(X_2))^2 = 2$$

$$E(X_1X_2) = \text{cov}(X_1, X_2) + E(X_1)E(X_2) = 0.5$$

所以 $E(X_1^2 - X_1X_2 + X_2^2) = 1 - 0.5 + 2 = 2.5$.