第 4 章-随机变量的数字特征

李孟棠, 朱彬

中山大学-智能工程学院

{limt29,zhub26}@mail.sysu.edu.cn

December 18, 2020

章节概述

知识点:

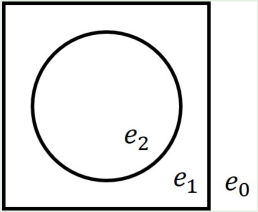
1. 数学期望
2. 方差
3. 协方差及相关系数
4. 矩、协方差矩阵

4.1 数学期望

例

一射手进行打靶，规定射入区域 e_2 得 2 分，e_1 得 1 分，e_0 得 0 分. 射手一次射击的分数 X 是一个随机变量，X 的分布律为

$$P\{X=k\}=p_k,\ \ k=0,1,2.$$



现射击 N 次，其中射入 e_0 次数为 a_0，e_1 次数为 a_1，e_2 次数为 a_2. 于是得分总和为 $a_0\times 0+a_1\times 1+a_2\times 2$，平均每次射击得分

$$\frac{a_0\times 0+a_1\times 1+a_2\times 2}{a_0+a_1+a_2}=\sum_{k=0}^{2}k\frac{a_k}{N}.$$

定义

(1) 设离散型随机变量 X 的分布律为

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

若级数 $\sum_{k=1}^{\infty} x_k p_k$ 绝对收敛，则称级数 $\sum_{k=1}^{\infty} x_k p_k$ 的值为随机变量 X 的数学期望，记为 $E(X)$. 即

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

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

$$E(X)=\int_{-\infty}^{\infty} xf(x)dx. \tag{2}$$

例 (1)

某医院当新生儿诞生时，医生要根据婴儿的皮肤颜色、肌肉弹性、反应敏感性、心脏搏动等方面进行评分，新生儿得分 X 是一个随机变量. 根据以下资料表明的 X 的分布律，求 X 的数学期望

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

解

$$E(X) = 0 \times 0.002 + 1 \times 0.001 + 2 \times 0.002 + \ldots + 9 \times .12 + 10 \times 0.01 = 7.15$$

例 (2)

有 2 个相互独立工作的电子装置，它们的寿命 (以小时计)$X_k(k=1,2)$ 服从同一指数分布，其概率密度为 ($\theta>0$)

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta} & x>0\\ 0 & x\le 0\end{cases}$$

若将这 2 个电子装置串联组成整机，求整机寿命 N 的数学期望.

解 X_k $(k=1,2)$ 的分布函数为

$$F(x)=\begin{cases}1-e^{-x/\theta} & x>0\\ 0 & x\le 0\end{cases}$$

由第三章可知 $N=\min\{X_1,X_2\}$ 的分布函数为

$$F_{\min}(x)=1-[1-F(x)]^2=\begin{cases}1-e^{-2x/\theta} & x>0\\ 0 & x\le 0\end{cases}$$

例 (2 续)

$$F_{\min}(x) = 1 - [1 - F(x)]^2 = \begin{cases} 1 - e^{-2x/\theta} & x > 0 \\ 0 & x \le 0 \end{cases}$$

因为 N 的概率密度为 $F_{\min}(x)$ 的导数

$$f_{\min}(x) = \begin{cases} \frac{2}{\theta} e^{-2x/\theta} & x > 0 \\ 0 & x \le 0 \end{cases}$$

所以 N 的数学期望为

$$E(N) = \int_{-\infty}^{\infty} x f_{\min}(x) dx = \int_0^{\infty} \frac{2x}{\theta} e^{-2x/\theta} dx = \frac{\theta}{2}$$

例 (3)

按规定，某车站每天 $8:00 \sim 9:00, 9:00 \sim 10:00$ 都有一辆客车到站，但到站的时刻是随机的，且两者到站的时间相互独立，其规律如下. 求一旅客 $8:20$ 到车站，他候车时间的数学期望.

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

解 设旅客候车的时间为 X(分钟)，则其分布律为

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

因此，候车时间 X 的数学期望为

$$E(X) = 10 \times \frac{1}{2} + 30 \times \frac{1}{3} + 50 \times \frac{1}{36} + 70 \times \frac{1}{12} + 90 \times \frac{1}{18} = 27.22$$

例 (4)

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(年), 求该商店一台这种家电付款 Y 的数学期望.

已知家电价格:

1. $X \leq 1$, 一台付款 1500 元;
2. $1 < X \leq 2$, 一台付款 2000 元;
3. $2 < X \leq 3$, 一台付款 2500 元;
4. $X > 3$, 一台付款 3000 元.

已知家电寿命 X 服从指数分布, 概率密度为

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

解 寿命 X 落在各个时间段的概率:

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

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

例 (4 续)

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

$$P\{1 < X \le 2\} = \int_1^2 \frac{1}{10} e^{-x/10} dx = e^{-0.1} - e^{-0.2} = 0.0861$$

$$P\{2 < X \le 3\} = \int_2^3 \frac{1}{10} e^{-x/10} dx = e^{-0.2} - e^{-0.3} = 0.0779$$

$$P\{X > 3\} = \int_3^\infty \frac{1}{10} e^{-x/10} dx = e^{-0.3} = 0.7408$$

因此家电收费 Y 的分布律为

Y	1500	2000	2500	3000
p_k	0.0952	0.0861	0.0779	0.7408

通过公式 $E(Y) = \sum_{k=1}^4 y_k p_k = 2732.15$ 元

例 (5)

对于某种传播性疾病，需要抽查 N 个人的血液，现有两种方法:(1) 将每个人的血分别化验，共需 N 次化验.(2) 按 k 个人一组，把从 k 个人的血混合化验，若呈阴性，则说明这 k 个人的血液都为阴性，只需 1 次化验；若为阳性，再分别对这 k 个人化验，则共需 $k+1$ 次化验. 假设每个人阳性概率为 p，且每个人互相独立，问当 p 较小时，如何选取 k 会使得第二种方法总化验次数更少?

解 每个人阳性概率 p，阴性概率 $q=1-p$. 则 k 个人混合呈阴性概率为 q^k，混合呈阳性概率为 $1-q^k$.
k 人一组，组内每人平均化验次数 X，其分布律为

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

则 X 的数学期望为

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

例 (5 续)

k 人一组, 组内每人平均化验次数 X, 其分布律为

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

则 X 的数学期望为

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

若方法二优于方法一, 则

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

例如,$p = 0.1$, $q = 0.9$, 当 $k = 4$, $1 - q^k + \frac{1}{k}$ 取得最小值. 若一共 1000 人, 按 $k = 4$ 分组, 则第一种方法需要化验 1000 次, 第二种方法需要化验 $1000(1-0.9^4+\frac{1}{4}) = 594$ 次, 减少工作量约 40%.

例 (6)

设 $X \sim \pi(\lambda)$, 求 $E(X)$.

解 X 的分布律为

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

X 的数学期望为

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

即泊松分布的数学期望 $E(X)$ 为其参数 λ.

例 (7)

设 $X \sim U(a, b)$, 求 $E(X)$.

解 X 的概率密度为

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

X 的数学期望为

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

即均匀分布的数学期望 $E(X)$ 为其区间中点.

定理 (Law of the unconscious statistician)

设 Y 是随机变量 X 的函数: $Y = g(X)$(g 是连续函数), 则对于

1. 离散型X, 分布律为 $P\{X = x_k\} = p_k,\ k = 1, 2, \ldots$, 若 $\sum_{k=1}^{\infty} g(x_k)p_k$ 绝对收敛, 有

$$E(Y) = E[g(X)] = \sum_{k=1}^{\infty} g(x_k)p_k; \tag{3}$$

2. 连续型X, 概率密度 $f(x)$, 若 $\int_{-\infty}^{\infty} g(x)f(x)dx$ 绝对收敛, 有

$$E(Y) = E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx. \tag{4}$$

证明见教科书 P95.

定理 (Law of the unconscious statistician, 2-d version)

设 $Z = g(X, Y)$ 是随机变量 X, Y 的函数，其中 g 为连续函数, 则 Z 是一维随机变量，对于

1. 离散型(X, Y), 分布律为 $P\{X = x_i, Y = y_j\} = p_{ij}$, $i, j = 1, 2, \ldots$, 有

$$E(Z) = E[g(X, Y)] = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} g(x_i, y_j)p_{ij}; \tag{5}$$

2. 连续型(X, Y), 概率密度为 $f(x, y)$, 有

$$E(Z) = E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f(x, y)dxdy. \tag{6}$$

例 (8)

风速 V 在区间 $(0, a)$ 上服从均匀分布, 即有概率密度 $f(v) = 1/a$, $0 < v < a$. 飞机机翼收到的正压力 W 是 V 的函数: $W = kV^2$, $k > 0$, 求 W 的数学期望.

解

$$E(W) = \int_{-\infty}^{\infty} kv^2 f(v) dv = \int_0^a kv^2 \frac{1}{a} dv = \frac{1}{3} ka^2.$$

例 (9)

设随机变量 (X,Y) 的概率密度为 $f(x,y)=\frac{3}{2x^3y^2}, \frac{1}{x}<y<x, x>1$, 求 $E(Y), E(\frac{1}{XY})$.

解 回顾：分部积分公式 $\int_a^b u\frac{dv}{dx}dx=\Big[uv\Big]_a^b-\int_a^b v\frac{du}{dx}dx$.

$$E(Y)=\int_{-\infty}^{\infty} yf(x,y)dydx=\int_1^{\infty}\int_{\frac{1}{x}}^{x}\frac{3}{2x^3y}dydx=\frac{3}{2}\int_1^{\infty}\frac{1}{x^3}\Big[\ln y\Big]_{\frac{1}{x}}^{x}dx$$

$$=3\int_1^{\infty}\frac{\ln x}{x^3}dx=\Big[-\frac{3}{2}\frac{\ln x}{x^2}\Big]_1^{\infty}+\frac{3}{2}\int_1^{\infty}\frac{1}{x^3}dx=\frac{3}{4}.$$

$$E(\frac{1}{XY})=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{xy}\cdot\frac{3}{2}\frac{1}{x^3y^2}dydx=\frac{3}{2}\int_1^{\infty}\frac{1}{x^4}dx\int_{\frac{1}{x}}^{x}\frac{1}{y^3}dy$$

$$=\frac{3}{2}\int_1^{\infty}\frac{1}{x^4}\Big[-\frac{1}{2}\cdot\frac{1}{y^2}\Big]_{\frac{1}{x}}^{x}dx=-\frac{3}{4}\int_1^{\infty}\frac{1}{x^4}(\frac{1}{x^2}-x^2)dx$$

$$=-\frac{3}{4}\int_1^{\infty}\frac{1}{x^6}-\frac{1}{x^2}dx=-\frac{3}{4}\Big[-\frac{1}{5}\frac{1}{x^5}+\frac{1}{x}\Big]_1^{\infty}=\frac{3}{5}.$$

例 (10)

某公司计划开发一种新产品，并确定产量. 估计售出一件可获利 m 元，而积压一件会损失 n 元. 估计销售量 Y(件) 服从指数分布，其概率密度为 $f_Y(y) = \frac{1}{\theta}e^{-y/\theta}, y > 0, \theta > 0$. 求要获利的数学期望最大，应生产多少件产品?

解 设生产 x 件，则获利 Q 是 x 的函数

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

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

令一阶导数 $\frac{d}{dx}E(Q) = 0$，即

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

令一阶导数 $\frac{d}{dx}E(Q)=0$, 即

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

得

$$x=-\theta\ln\frac{n}{m+n}$$

而二阶导数

$$\frac{d^2}{dx^2}E(Q)=\frac{-(m+n)}{\theta}e^{-x/\theta}<0 \quad \text{(全局凹函数)}$$

所以, $x=-\theta\ln\frac{n}{m+n}$ 时 $E(Q)$ 取得最大值.
例如若 $f_Y(y)=\frac{1}{10000}e^{-y/10000}$, $y>0$, $m=500$ 元, $n=2000$ 元, 则

$$x=-10000\ln\frac{2000}{500+2000}=2231.4.$$

也就是生产 2231.4 件时获利最大.

例 (11)

甲乙丙三人参加竞拍, 价高者获胜. 若甲获胜, 他就将此项目以 10 千美元转让给他人. 设三人竞拍价相互独立, 且都在 7∼11 千美元间均匀分布. 问甲如何报价才能使获益的数学期望最大?

解 甲乙丙三人报价为 X_1, X_2, X_3, 都在 $(7, 11)$ 上均匀分布, 其分布函数为

$$F(u) = \begin{cases} 0 & u < 7 \\ \frac{u-7}{4} & 7 \le u < 11 \\ 1 & u \ge 11 \end{cases}$$

以 Y 记为三人最大出价 $Y = \max\{X_1, X_2, X_3\}, Y$ 的分布函数为

$$F_Y(u) = \begin{cases} 0 & u < 7 \\ \left(\frac{u-7}{4}\right)^3 & 7 \le u < 11 \\ 1 & u \ge 11 \end{cases}$$

若甲报价 x $(7 \le x \le 10)$, 甲能赢下项目竞拍的概率为

$$p = P\{Y \le x\} = F_Y(x)\left(\frac{x-7}{4}\right)^3, \ 7 \le x \le 10$$

例 (11 续)

以 $G(X)$ 记甲赚钱数, 分布律为

$G(X)$	$10-x$	0
概率	$\left(\frac{x-7}{4}\right)^3$	$1-\left(\frac{x-7}{4}\right)^3$

所以甲赚钱数的数学期望

$$E[G(X)] = \left(\frac{x-7}{4}\right)^3 (10-x)$$

令

$$\frac{d}{dx}E[G(X)] = \frac{1}{4^3}[(x-7)^2(37-4x)] = 0$$

得 $x_1 = 37/4$, $x_2 = 7$(舍去). 又

$$\left.\frac{d^2}{dx^2}E[G(X)]\right|_{x=37/4} < 0,$$ （局部凹函数，Matlab 图示）

所以甲报价 $x = 37/4$ 时, 赚钱数的数学期望最大.

检查函数（多项式）

$$f(x) := E[G(X)] = \left(\frac{x-7}{4}\right)^3 (10-x)$$

在 $x = 7$ 的表现。利用 Taylor 展开

$$\begin{aligned} f(x) &= \left(\frac{x-7}{4}\right)^3 [3-(x-7)] \\ &= \frac{3}{4^3}(x-7)^3 - \frac{1}{4^3}(x-7)^4 \end{aligned}$$

当 $x = 7$ 时，f 的一阶导和二阶导均为零。

采用无穷小量分析 (infinitesimal analysis)，令 $x = 7 \pm \varepsilon$，其中 $\varepsilon > 0$ 是一个“无穷小量”，计算得到

$$f(7+\varepsilon) = \frac{\varepsilon^3}{4^3}(3-\varepsilon) > 0, \quad f(7-\varepsilon) = \frac{3}{4^3}(-\varepsilon)^3 - \frac{1}{4^3}\varepsilon^4 < 0.$$

因此 $x = 7$ 是 $f(x)$ 的非极值点（鞍点，saddle point）。

数学期望的性质

随机变量 (连续型、离散型) 具有以下重要性质:

1. 设 C 为常数, 则 $E(C) = C$.
2. 设 X 为随机变量,C 为常数, 则
$$E(CX) = CE(X)$$
3. 设 X, Y 为两个随机变量, 则
$$E(X+Y) = E(X) + E(Y)$$
这一性质可推广到任意有限个随机变量之和的情况.
4. 设 X, Y 是相互独立的随机变量, 则
$$E(XY) = E(X)E(Y)$$
这一性质可推广到任意有限个相互独立的随机变量之积的情况.

例 (12)

一公交车载有 20 位乘客, 路上共有 10 个站点可以下车. 如果没有乘客下车, 公交车就不停车. 以 X 表示停车次数, 求 $E(X)$(每位旅客在每个站点下车是等可能的, 且每位旅客是否下车相互独立).

解 设

$$X_i = \begin{cases} 0 & \text{在第 } i \text{ 站没人下车} \\ 1 & \text{在第 } i \text{ 站有人下车} \end{cases}$$

易知 $X = X_1 + X_2 + ... + X_{10}$ 按题意, 任一乘客在第 i 站下车概率为 $\frac{1}{10}$, 不下车概率 $\frac{9}{10}$, 因此 20 位乘客都不在第 i 站下车的概率为 $(\frac{9}{10})^{20}$, 在第 i 站有人下车的概率为 $1 - (\frac{9}{10})^{20}$, 即

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

$$E(X_i) = 0 \cdot (\frac{9}{10})^{20} + 1 \cdot [1 - (\frac{9}{10})^{20}] = 1 - (\frac{9}{10})^{20}$$

$$E(X) = E(X_1 + X_2 + ... + X_{10}) = 10[1 - (\frac{9}{10})^{20}] \approx 8.784$$

验算 $P\{X_2=0\}$。采用全概率公式，

$$\begin{aligned}P\{X_2=0\}=&P\{X_2=0|X_1=0\}P\{X_1=0\}\\&+P\{X_2=0|X_1=1\}P\{X_1=1\}\end{aligned}$$

不精确！准确来说，

$$\begin{aligned}P\{X_2=0\}=&P\{X_2=0|X_1=0\}P\{X_1=0\}\\&+\sum_{k=1}^{20}P\{X_2=0|\text{第一站下 } k \text{ 人}\}P\{\text{第一站下 } k \text{ 人}\}\\=&\left(\frac{9}{10}\right)^{20}\times\left(\frac{8}{9}\right)^{20}+\sum_{k=1}^{20}\left(\frac{8}{9}\right)^{20-k}\binom{20}{k}\left(\frac{1}{10}\right)^{k}\left(\frac{9}{10}\right)^{20-k}\\=&\left(\frac{8}{10}\right)^{20}+\left(\frac{1}{10}+\frac{8}{10}\right)^{20}-\left(\frac{8}{10}\right)^{20}=\left(\frac{9}{10}\right)^{20}.\end{aligned}$$

类似验证 $P\{X_3=0\}$（作为课后练习题）。

例 (13)

设一电路中电流 $I(A)$ 与电阻 $R(\Omega)$ 是两个**相互独立**的随机变量，其概率密度分别为

$$g(i) = 2i,\ 0 \le i \le 1 \qquad h(r) = r^2/9,\ 0 \le r \le 3$$

求电压 $V = IR$ 的均值.

解

$$E(V) = E(IR) = E(I)E(R)$$

$$= \left[\int_{-\infty}^{\infty} i \cdot g(i) di\right]\left[\int_{-\infty}^{\infty} r \cdot h(r) dr\right]$$

$$= \left[\int_0^1 i \cdot 2i di\right]\left[\int_0^3 r \cdot \frac{r^2}{9} dr\right] = 1.5V$$

期望不存在的例子：柯西分布，密度函数如下

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma} \cdot \frac{\gamma^2}{(x - x_0)^2 + \gamma^2}, \quad x \in \mathbb{R},$$

其中 $x_0 \in \mathbb{R}$ 为位置参数 (location parameter)，$\gamma > 0$ 为标度参数 (scale parameter)。按照期望的定义计算积分

$$\int_{-\infty}^{+\infty} x\, f(x; x_0, \gamma)\, \mathrm{d}x = \infty - \infty \quad \text{(不定式)}.$$

Ex. 验证上述结果。提示：取 $x_0 = 0$, $\gamma = 1$ 以简化计算。

4.2 方差

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 $\mathrm{Var}(X)$, 即

$$D(X)=\mathrm{Var}(X)=E\{[X-E(X)]^2\} \tag{7}$$

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

由定义可知, 方差实际上就是随机变量 X 的函数 $g(X)=[X-E(X)]^2$ 的数学期望 $E[g(X)]$.
对于离散型随机变量, X 的分布律为 $P\{X=x_k\}=p_k, k=1,2,\dots$ 由公式 (3) 可知

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2\cdot p_k. \tag{8}$$

对于连续型随机变量, X 的概率密度为 $f(x)$, 由公式 (4) 可知

$$D(X)=\int_{-\infty}^{\infty}[x-E(X)]^2\cdot f(x)dx. \tag{9}$$

Proposition

随机变量 X 的方差 $D(X)$ 可由下列公式计算

$$\begin{aligned} D(X) &= E\{[X - E(X)]^2\} \\ &= E\{X^2 - 2XE(X) + [E(X)]^2\} \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned} \tag{10}$$

例 (1)

随机变量 X 具有数学期望 $E(X)=\mu$, 方差 $D(X)=\sigma^2\neq 0$. 若

$$X^*=\frac{X-\mu}{\sigma}$$

则

$$E(X^*)=\frac{1}{\sigma}E(X-\mu)=\frac{1}{\sigma}[E(X)-\mu]=0$$

$$\begin{aligned} D(X^*)&=E(X^{*2})-[E(X^*)]^2=E[(\frac{X-\mu}{\sigma})^2] \\ &=\frac{1}{\sigma^2}E[(X-\mu)^2]=\frac{\sigma^2}{\sigma^2}=1 \end{aligned}$$

即 $X^*=(X-\mu)/\sigma$ 的数学期望为 0, 方差为 1. X^* 称为 X 的标准化变量.

例 (2)

随机变量 X 具有 $(0-1)$ 分布, 基于如下分布律, 求 $D(X)$.

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

解

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

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

由公式 (10) 可以计算 $D(X)$:

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

例 (3)

随机变量 $X \sim \pi(\lambda)$, 求 $D(X)$.

解 X 的分布律为

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

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

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

所以方差为

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

例 (4)

随机变量 $X \sim U(a,b)$, 求 $D(X)$.

解 X 的概率密度为

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

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

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_a^b \frac{x^2}{b-a}dx = \frac{b^3-a^3}{3(b-a)}$$

所以方差为

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

例 (5)

随机变量 X 服从指数分布，其概率密度如下 $(\theta > 0)$，求 $E(X), D(X)$.

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

解

$$E(X) = \int_{-\infty}^{\infty} x\frac{1}{\theta}e^{-x/\theta}dx = -xe^{-x/\theta}\Big|_0^{\infty} + \int_0^{\infty} e^{-x/\theta}dx = \theta$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2\frac{1}{\theta}e^{-x/\theta}dx = -x^2e^{-x/\theta}\Big|_0^{\infty} + \int_0^{\infty} 2xe^{-x/\theta}dx = 2\theta^2$$

所以方差为

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

方差的性质

随机变量 (连续型、离散型) 具有以下重要性质:

1. 设 C 为常数, 则 $D(C) = 0$.
2. 设 X 为随机变量,C 为常数, 则

$$D(CX) = C^2 D(X), D(X+C) = D(X)$$

3. 设 X, Y 为两个随机变量, 则

$$D(X+Y) = D(X) + D(Y) + 2E[(X-E(X))(Y-E(Y))] \quad (11)$$

若 X 与 Y**相互独立**, 则

$$D(X+Y) = D(X) + D(Y) \quad (12)$$

这一性质可推广到任意有限个**相互独立**的随机变量之和的情况.

4. $D(X) = 0$ 的**充要条件**是 X 以概率 1 取常数 $E(X)$, 即

$$P\{X = E(X)\} = 1$$

例 (6)

随机变量 $X \sim b(n,p)$, 求 $E(X), D(X)$.

解 引入新的随机变量

$$X_k = \begin{cases} 1 & A \text{ 在第 } k \text{ 次试验发生} \\ 0 & A \text{ 在第 } k \text{ 次试验不发生} \end{cases}, \quad k = 1, 2, ..., n$$

那么

$$X = X_1 + X_2 + ... + X_n \tag{13}$$

并且 X_k 只依赖第 k 次试验, 因此 $X_1, X_2, ..., X_n$相互独立. 并且我们知道 $E(X_k) = p, D(X_k) = p(1-p)$, 所以

$$E(X) = E(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} E(X_k) = np.$$

$$D(X) = D(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} D(X_k) = np(1-p).$$

例 (7)

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

解 化为标准正态变量 Z, 及其概率密度为

$$Z = \frac{X-\mu}{\sigma}, f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

于是

$$E(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} te^{-t^2/2} \mathrm{d}t = -\frac{1}{\sqrt{2\pi}} e^{-t^2/2} \Big|_{-\infty}^{\infty} = 0 \quad \text{（奇函数）}$$

$$\begin{aligned} D(Z) = E(Z^2) - E(Z)^2 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2/2} \mathrm{d}t \\ &= -\frac{1}{\sqrt{2\pi}} te^{-t^2/2} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} \mathrm{d}t = 1 \end{aligned}$$

所以

$$E(X) = E(\mu + \sigma Z) = \mu, D(X) = D(\mu + \sigma Z) = D(\sigma Z) = \sigma^2 D(Z) = \sigma^2$$

例 (7 续)

重要结论:

1. 由上章, 若多个正态分布 $X_i \sim N(\mu_i, \sigma_i^2)$ 相互独立, 则其线性组合 $X = C_1X_1 + C_2X_2 + ... + C_nX_n$ (C_i 不全为 0) 仍然服从正态分布
2. 具体为

$$X = C_1X_1 + C_2X_2 + ... + C_nX_n \sim N\left(\sum_{i=1}^{n} C_i\mu_i, \sum_{i=1}^{n} C_i^2\sigma_i^2\right) \quad (14)$$

   即数学期望为 $E(X) = \sum_{i=1}^{n} C_i\mu_i$, 方差为 $D(X) = \sum_{i=1}^{n} C_i^2\sigma_i^2$

3. 例如 $X \sim N(1,3)$, $Y \sim N(2,4)$,X, Y 相互独立, 则 $Z = 2X - 3Y$ 也服从正态分布, 且

$$E(Z) = 2 \times 1 - 3 \times 2 = -4, \quad D(Z) = 4D(X) + 9D(Y) = 48,$$
$$Z \sim N(-4, 48).$$

例 (8)

设活塞直径 (cm) $X \sim N(22.40, 0.03^2)$, 气缸直径 (cm) $Y \sim N(22.50, 0.04^2)$, X, Y 相互独立. 求任取一只活塞一只气缸, 活塞能放入气缸的概率.

解 题意即 $P\{X < Y\} = P\{X - Y < 0\}$. 而

$$X - Y \sim N(1 \times 22.40 - 1 \times 22.50, 1^2 \times 0.03^2 + 1^2 \times 0.04^2) = N(-0.10, 0.0025)$$

所以

$$P\{X < Y\} = P\{X - Y < 0\}$$

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

上面利用标准化操作, 把概率计算归结到 $N(0,1)$ 情形, 通过查标准正态分布表解决。

Proposition (切比雪夫 (Chebyshev) 不等式)

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, 有如下不等式成立

$$P\{|X-\mu|\geq\varepsilon\}\leq\frac{\sigma^2}{\varepsilon^2}. \tag{15}$$

证明.

$$P\{|X-\mu|\geq\varepsilon\}=\int_{|x-\mu|\geq\varepsilon}f(x)dx\leq\int_{|x-\mu|\geq\varepsilon}\frac{|x-\mu|^2}{\varepsilon^2}f(x)dx$$

$$\leq\frac{1}{\varepsilon^2}\int_{-\infty}^{\infty}(x-\mu)^2f(x)dx=\frac{\sigma^2}{\varepsilon^2}$$

其中 $\int_{-\infty}^{\infty}(x-\mu)^2f(x)dx$ 是 $D(X)$ 的定义, 因此等于 σ^2. □

Proposition (切比雪夫 (Chebyshev) 不等式)

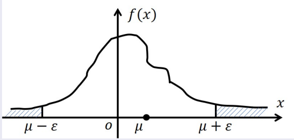
切比雪夫不等式

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

也可以写成

$$P\{|X-\mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2} \tag{16}$$

切比雪夫不等式给出了随机变量 X 在分布未知, 只知道 $E(X), D(X)$ 时, 估计 $P\{|X-E(X)| < \varepsilon\}$ 的界限的方法.



4.3 协方差及相关系数

定义

数字 $E\{[X-E(X)][Y-E(Y)]\}$ 称为随机变量 X, Y 的协方差 (covariance)，记为 $\mathrm{Cov}(X,Y)$，即

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

而称 ρ_{XY} 为随机变量 X, Y 的相关系数 (correlation coefficient)，

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \quad (\text{这里假设} D(X), D(Y)>0).$$

由定义可知如下性质

$$\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X), \mathrm{Cov}(X,X)=D(X)$$

$$\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y),\ a,b\in\mathbb{R}$$

$$\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$$

$$D(X+Y)=D(X)+D(Y)+2\mathrm{Cov}(X,Y) \tag{17}$$

$$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y) \tag{18}$$

ρ_{XY} 的重要性质

考虑: 以 X 的线性函数 $a+bX$ 来近似 Y. 以均方误差 e 来衡量近似好坏程度.

$$\begin{aligned} e &= E[(Y-(a+bX))^2] \\ &= E(Y^2)+b^2E(X^2)+a^2-2bE(XY)+2abE(X)-2aE(Y) \end{aligned} \tag{19}$$

问题: 求 a, b 使得 e 最小.

解 令偏导数分别为 0，即

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

解得

$$b_0 = \frac{\mathrm{Cov}(X,Y)}{D(X)}$$

$$a_0 = E(Y)-b_0E(X) = E(Y)-E(X)\frac{\mathrm{Cov}(X,Y)}{D(X)}$$

将 a_0, b_0 带入公式 (19) 得

$$\min_{a,b} e = E\{[Y-(a_0+b_0X)]^2\} = (1-\rho_{XY}^2)D(Y) \tag{20}$$

Proposition

1. $|\rho_{XY}| \le 1$
2. $|\rho_{XY}| = 1$ 的充要条件是, 存在常数 a, b 使得 $P\{Y = a + bX\} = 1$.

证明.

必要性: 若 $|\rho_{XY}| = 1$, 则由公式 (20) 知

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

从而由性质

$$0 = E([\sim]^2) = D[\sim] + (E[\sim])^2$$

得到

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

由方差的性质 4 可知

$$P\{Y - (a_0 + b_0X) = 0\} = 1 \implies P\{Y = (a_0 + b_0X)\} = 1$$

□

证明.

充分性: 若存在常数 a^*, b^* 使得

$$P\{Y = a^* + b^* X\} = 1$$

则

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

即

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

所以

$$0 = E\{[Y - (a^* + b^* X)]^2\} \geq E\{[Y - (a_0 + b_0 X)]^2\} = (1 - \rho_{XY}^2) D(Y)$$

即

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

□

说明: 均方误差 e 是 $|\rho_{XY}|$ 的严格单调减少函数.

1. 当 $|\rho_{XY}|$ 较大时, e 较小, 说明 X, Y线性关系紧密; 反之不紧密.
2. 当 $|\rho_{XY}| = 1$ 时, X, Y 之间以概率 1 存在线性关系.
3. 当 $|\rho_{XY}| = 0$ 时, X, Y (线性) 不相关. 注意此时 $\mathrm{Cov}(X, Y) = 0$, 线性近似的最优系数 $b_0 = 0$, $a_0 = E(Y)$.

注意: 当 X, Y互相独立时, 由数学期望的性质 4

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

以及公式 (18)

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

可知 $\mathrm{Cov}(X, Y) = 0$, 从而

$$\rho_{XY} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = 0$$

即 X, Y 相互独立 $\Longrightarrow$ X, Y 不相关. 但是反之不成立(见例 1).
也就是 X, Y 不相关 $\not\Longrightarrow$ X, Y 相互独立.

例 (1)

随机变量 (X, Y) 的分布律为

$Y \backslash X$	-2	-1	1	2	$P\{Y=i\}$
1	0	1/4	1/4	0	1/2
4	1/4	0	0	1/4	1/2
$P\{X=i\}$	1/4	1/4	1/4	1/4	1

可以看到,

$$E(X)=0, E(Y)=5/2, E(XY)=0, \rho_{XY}=0$$

所以 X, Y 不相关, 即不存在线性关系.
但是,

$$P\{X=-2, Y=1\}=0 \neq P\{X=-2\}P\{Y=1\}$$

所以 X, Y 不是相互独立的.
事实上, $Y=X^2$.

例 (2)

设随机变量 (X,Y) 服从二维正态分布，概率密度为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot$$

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

由上一章知道 (X,Y) 的边缘概率密度即为各自的正态分布

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}\exp\left\{\frac{-(x-\mu_1)^2}{\sigma_1^2}\right\},\ f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}\exp\left\{\frac{-(y-\mu_2)^2}{\sigma_2^2}\right\},\ x,y\in\mathbb{R}$$

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

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)dxdy=\rho\sigma_1\sigma_2$$

于是

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

对于二维正态分布随机变量:(X,Y) 不相关 $\Longleftrightarrow$ (X,Y) 相互独立.

其中积分

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)dxdy$$

的计算涉及二维变量替换（此处本质为线性变换）:

$$\begin{bmatrix} u \\ t \end{bmatrix}=\varphi(x,y)=\begin{bmatrix} \frac{x-\mu_1}{\sigma_1} \\ \frac{1}{\sqrt{1-\rho^2}}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right) \end{bmatrix}.$$

向量值函数 φ 的 Jacobi 矩阵为

$$D\varphi(x,y)=\begin{bmatrix} \frac{\partial\varphi_1}{\partial x}(x,y) & \frac{\partial\varphi_1}{\partial y}(x,y) \\ \frac{\partial\varphi_2}{\partial x}(x,y) & \frac{\partial\varphi_2}{\partial y}(x,y) \end{bmatrix}=\begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix},$$

所以（二维）面积微元变换为

$$\mathrm{d}u\mathrm{d}t=|\det(D\varphi(x,y))|\mathrm{d}x\mathrm{d}y=\frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}}\mathrm{d}x\mathrm{d}y.$$

4.4 矩、协方差矩阵

定义

给定随机变量 X, Y，定义如下量：

1. **k 阶原点矩**，简称 **k 阶矩**

$$E(X^k),\ k = 1, 2, \ldots$$

2. **k 阶中心矩**

$$E\{[X - E(X)]^k\},\ k = 2, 3, \ldots$$

3. **$k + \ell$ 阶混合矩**

$$E(X^k Y^\ell),\ k, \ell = 1, 2, \ldots$$

4. **$k + \ell$ 阶混合中心矩**

$$E\{[X - E(X)]^k [Y - E(Y)]^\ell\},\ k, \ell = 1, 2, \ldots$$

显然，

1. $E(X)$ 是 X 的一阶原点矩.
2. $D(X)$ 是 X 的二阶中心矩.
3. $\mathrm{Cov}(X, Y)$ 是 X, Y 的二阶混合中心矩.

定义

假设二维随机变量 (X_1, X_2) 的 4 个二阶中心矩都存在，记为

$$c_{11} = E\{[X_1 - E(X_1)]^2\}$$

$$c_{12} = E\{[X_1 - E(X_1)][X_2 - E(X_2)]\}$$
$$c_{21} = E\{[X_2 - E(X_2)][X_1 - E(X_1)]\}$$
$$c_{22} = E\{[X_2 - E(X_2)]^2\}$$

则称如下矩阵为 (X_1, X_2) 的协方差矩阵

$$\boldsymbol{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}.$$

显然 $c_{12} = c_{21}$，所以 $\boldsymbol{C}$ 为对称矩阵。

注意：随机变量的矩不一定存在，例如期望和方差不一定存在。

扩展到 n 维随机变量 $X_1, X_2, ..., X_n$, 设其二阶混合中心矩存在

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

则称如下矩阵为 $(X_1, X_2, ..., X_n)$ 的协方差矩阵

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

应用: 一般 n 维随机变量的分布未知, 或太复杂, 因此协方差矩阵可以简化其表达和计算 (近似).

二维正态分布随机变量 (X_1, X_2) 的概率密度

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot$$

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

引入 $\boldsymbol{X}, \boldsymbol{\mu}$

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

由本章前述知识可知, (X_1, X_2) 的协方差矩阵为

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

$$\boldsymbol{C}^{-1} = \frac{1}{\det \boldsymbol{C}} \cdot \mathrm{adj}(\boldsymbol{C}) = \frac{1}{\det \boldsymbol{C}} \cdot \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}$$

那么构造如下二次型

$$(\boldsymbol{X}-\boldsymbol{\mu})^\top \boldsymbol{C}^{-1}(\boldsymbol{X}-\boldsymbol{\mu}) = \frac{1}{\det \boldsymbol{C}} \begin{bmatrix} x_1-\mu_1 & x_2-\mu_2 \end{bmatrix} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix},$$

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

其中，$\det \boldsymbol{C} = \sigma_1^2\sigma_2^2(1-\rho^2)$。

所以，(X_1, X_2) 的概率密度可以写成矩阵形式

$$f(x_1, x_2) = \frac{1}{(2\pi)^{2/2}(\det \boldsymbol{C})^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^\top \boldsymbol{C}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right\}$$

从二维自然地推广到 n 维 $(X_1, X_2, ..., X_n)$

$$f(x_1, x_2, ..., x_n) = \frac{1}{(2\pi)^{n/2}(\det \boldsymbol{C})^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^\top \boldsymbol{C}^{-1}(\boldsymbol{X}-\boldsymbol{\mu})\right\}$$

1. 若 n 维随机变量 $(X_1, X_2, ..., X_n)$ 是 n 维正态分布，则每一个分量 $X_i, i = 1, 2, ..., n$ 都是正态随机变量；反之，若每一个分量 $X_i, i = 1, 2, ..., n$ 都是正态随机变量且相互独立，则 n 维随机变量 $(X_1, X_2, ..., X_n)$ 是 n 维正态分布.
2. n 维随机变量 $(X_1, X_2, ..., X_n)$ 是 n 维正态分布的充要条件是 $X_1, X_2, ..., X_n$ 的任意线性组合

   $$\ell_1 X_1 + \ell_2 X_2 + ... + \ell_n X_n$$

   服从一维正态分布.
3. 正态变量的线性变换不变性:n 维随机变量 $(X_1, X_2, ..., X_n)$ 是 n 维正态分布，设 $Y_1, Y_2, ..., Y_m$ 是 $X_j, j = 1, 2, ..., n$ 的线性函数，即 $Y_i = \sum_{j=1}^{n} a_{ij} X_j$，其中 $a_{ij} \in \mathbb{R}$，则 $(Y_1, Y_2, ..., Y_m)$ 是 m 维正态分布.
4. 若 n 维随机变量 $(X_1, X_2, ..., X_n)$ 是 n 维正态分布，则 $X_1, X_2, ..., X_n$ 相互独立 $\Longleftrightarrow$ $X_1, X_2, ..., X_n$ 两两不相关.

章节小结

知识点:

1. 数学期望

$$E(X)=\sum_{k=1}^{\infty}x_kp_k,\quad E(X)=\int_{-\infty}^{\infty}xf(x)dx$$

2. 方差

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2\cdot p_k,\ D(X)=\int_{-\infty}^{\infty}[x-E(X)]^2\cdot f(x)dx$$

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

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

3. 协方差及相关系数

$$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\},\ \rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

4. 矩、协方差矩阵
   k 阶原点矩, k 阶中心矩, $k+\ell$ 阶混合矩, $k+\ell$ 阶混合中心矩.

The End

本章作业：TBA...