

## 第三章、数学期望与特征函数

### §3.1 数学期望的定义

#### 一、数学期望的含义

例3.1.1 (空间平均).  $N$  位学生, 身高分别为  $i_1, \dots, i_N$ .

- 平均身高:  $\bar{i} = (i_1 + \dots + i_N)/N$ .
- 古典概型:  $\Omega = \{1, \dots, N\}$ ;  $X = \text{身高}$ ,  $X(r) = i_r$ .
- 分布列记为  $P(X = x_k) = p_k$ ,  $k = 1, \dots, K$ .  
设恰有  $N_k$  为同学身高为  $x_k$ , 则  $p_k = N_k/N$ .
- $i_1 + \dots + i_N = N_1 x_1 + \dots + N_K x_K$ , 故

$$\bar{i} = p_1 x_1 + \dots + p_K x_K.$$

- RHS: 所有可能值的加权平均值.

### 例3.1.2 (时间平均).

- 放回抽样 $n$ 次, 得 $X$ 的观测值 $a_1, \dots, a_n$  和时间平均值 $\bar{a}$ .
- 设其中恰有 $n_k$ 个为 $x_k$ , 则

$$i_1 + \dots + i_N = n_1 x_1 + \dots + n_K x_K.$$

- 令 $n \rightarrow \infty$ .  $n_k/n \rightarrow p_k$ . 故

$$\bar{a} \approx p_1 x_1 + \dots + p_K x_K.$$

## 二、离散型随机变量的数学期望

### 1. 非负情形

- 分布列:  $P(X = x_k) = p_k, k \in I.$
- $EX := \sum_{k \in I} p_k x_k$ , 良定.  
称为 $X$  的期望, 或 $\mathcal{L}(X)$  的期望.
- 例3.1.3.  $X \sim B(1, p),$

$$EX = 1 \cdot p + 0 \cdot (1 - p) = p.$$

- 特别地,  $E\mathbf{I}_A = P(A).$

例3.1.4.  $X \sim B(n, p)$ ,  $EX = np$ .

- $b(n, p; k) = C_n^k p^k (1-p)^{n-k}.$

$$k \cdot \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = np \cdot \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k}.$$

- $\underline{k \cdot b(n, p; k) = np \cdot b(n-1, p; k-1)}.$

- 记  $k' = k - 1$ , 则

$$EX = \sum_{k=1}^n k \cdot b(n, p; k) = np \sum_{k'=0}^{n-1} b(n-1, p; k') = np.$$

- 类似地, 超几何分布  $H(N, M, n)$  的期望为  $n \frac{M}{N}$  (例3.1.5);  
泊松分布  $P(\lambda)$  的期望为  $\lambda$  (例3.1.6).

命题3.1.7. 设 $X$  取非负整数, 则

$$EX = \sum_{n=1}^{\infty} P(X \geq n) = \sum_{n=0}^{\infty} P(X > n).$$

•  $EX = p_1 + (p_2 + p_2) + (p_3 + p_3 + p_3) + (p_4 + p_4 + p_4 + p_4) + \cdots,$

$$EX = \sum_{k=1}^{\infty} k p_k = \sum_{k=1}^{\infty} \sum_{n=1}^k p_k = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} p_k.$$

• 一般地, 若 $X$  为非负“格点型”: 取值 $\delta\mathbb{Z}_+$ , 则

$$EX = \delta \sum_{m=1}^{\infty} P(X \geq m\delta) = \delta \sum_{n=0}^{\infty} P(X > n\delta).$$

• 例3.1.8.  $X \sim G(p), EX = \sum_{n=0}^{\infty} q^n = \frac{1}{p}.$

## 2. 一般情形

- 若两个级数  $x_{\pm} := \sum_{k \in I} x_k^{\pm} p_k$  中至少有一个收敛,  
则级数  $\sum_{k \in I} x_k p_k$  良定, 且它等于  $x_+ - x_-$ .  
此时称此级数为  $X$  (或  $\mathcal{L}(X)$ ) 的期望.
- 例3.1.11.  $EX^{\pm} = x_{\pm}$ ,  $E|X| = x_+ + x_-$ .
- 若  $x_+ = x_- = \infty$ , 则级数不是良定的. 此时称  $X$  (或  $\mathcal{L}(X)$ ) 的期望不存在.

### 三、连续型随机变量的数学期望

- 设  $X$  的密度函数为  $p(x)$ . 记

$$x_+ := \int_0^{\infty} x \cdot p(x) dx = \lim_{M \rightarrow \infty} \int_0^M x \cdot p(x) dx,$$
$$x_- := \int_{-\infty}^0 |x| \cdot p(x) dx = \lim_{M \rightarrow \infty} \int_{-M}^0 |x| \cdot p(x) dx.$$

- 当  $x_+$  与  $x_-$  不同时为无穷时,

$$EX := \int_{-\infty}^{\infty} x \cdot p(x) dx = \lim_{M, N \rightarrow \infty} \int_{-M}^N x \cdot p(x) dx = x_+ - x_-.$$

是良定的, 称为  $X$  的期望.

- 若  $x_+ = x_- = \infty$ , 称  $X$  (或  $\mathcal{L}(X)$ ) 的期望不存在.

● 例3.1.12.  $U(a, b)$ .

●  $x_+, x_- < \infty$ , 故 $EX$  存在.

●  $EX = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{(b^2 - a^2)/2}{b-a} = \frac{a+b}{2}.$

● 例3.1.13.  $N(0, 1)$ .

●  $x_+, x_- < \infty$ , 故 $EX$  存在.

●  $x_+ = x_-$ , 故 $EX = 0$ .

● 总结1:  $|X| \leq M$  a.s., 则 $EX$  存在且 $|EX| \leq M$ .

总结2:  $p_X(x)$  关于 $\mu$  对称, 且 $EX$  存在, 则 $EX = \mu$ .

● 例3.1.14.  $C(0, 1), p(x) = \frac{1}{\pi(1+x^2)}.$

●  $x_+ = x_- = \infty$ , 故 $EX$  不存在.

命题3.1.15. 设 $X$  为连续型随机变量, 取非负实数, 则

$$\underbrace{EX = \int_0^{\infty} x \cdot p(x) dx}_{\text{wavy line}} = \underbrace{\int_0^{\infty} P(X > x) dx}_{\text{underline}} \triangleq c.$$

- 若 $EX = c = \infty$ , 则 $\checkmark$ .
- 往证: 若 $EX < \infty$  或  $c < \infty$ , 则  $a \cdot G(a) \xrightarrow{a \rightarrow \infty} 0$ .

于是由下式可推出 $EX = c$ :

$$\int_0^a x \cdot p_X(x) dx = -a \cdot G(a) + \int_0^a G(x) dx.$$

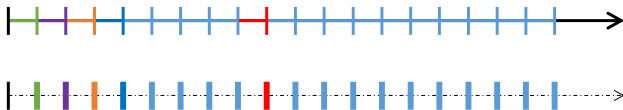
- $a \cdot G(a) = a \cdot \int_a^{\infty} p(x) dx \leq \int_a^{\infty} x \cdot p(x) dx \xrightarrow{a \rightarrow \infty} 0$ .
- $a \cdot G(a) = 2 \int_{a/2}^a G(x) dx \leq 2 \int_{a/2}^a G(x) dx \xrightarrow{a \rightarrow \infty} 0$ .
- 例3.1.16.  $\text{Exp}(\lambda)$ .  $EX = \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}$ .

## 四、数学期望的一般定义

### 1. 非负随机变量的数学期望

- 取定  $\delta > 0$ .  $X \in ((k-1)\delta, k\delta]$  时, 令  $\tilde{X}_\delta = k\delta$ .

$|X - \tilde{X}_\delta| \leq \delta$ , 故  $|EX - E\tilde{X}_\delta| \leq \delta$ .



- $$\begin{aligned} \star &= \sum_{k=0}^{\infty} (k+1)\delta \cdot P(k\delta < \tilde{X}_\delta \leq (k+1)\delta) \\ &= \delta + \sum_{k=0}^{\infty} x_k \cdot \Delta F(x_k) \approx \int_0^{\infty} x dF(x). \end{aligned}$$
- $$\star = \delta \sum_{n=0}^{\infty} P(\tilde{X}_\delta > n\delta) \approx \delta \sum_{n=1}^{\infty} P(X > n\delta) \approx \int_0^{\infty} G(x) dx.$$

- $EX := \int_0^{\infty} x dF(x)$ , 称为  $X$  (或  $\mathcal{L}(X)$ ) 的期望).
- 命题3.1.17 & 注3.1.18.

$$EX = \int_0^{\infty} P(X > x) dx = \int_0^{\infty} P(X \geq x) dx.$$

- 推论3.1.19.

$$EX < \infty \Leftrightarrow \sum_{n=0}^{\infty} P(X > n\delta) < \infty, \forall \delta > 0 \Leftrightarrow \exists \delta > 0 \text{ 使得 } \star < \infty.$$

$$\bullet \int_0^{\infty} P(X > x) dx = \sum_{n=0}^{\infty} \int_{n\delta}^{(n+1)\delta} P(X > x) dx.$$

$$\bullet P(X > (n+1)\delta) \leq \star \leq P(X > n\delta).$$

$$\bullet \star\delta - \delta \leq \sum_{n=0}^{\infty} \star \leq \star\delta.$$

$$\bullet \text{ 特别地, } EX < \infty \text{ 当且仅当 } \sum_{n=0}^{\infty} P(X > n) < \infty.$$

## 2. 一般随机变量的数学期望

- 当  $EX^+$  与  $EX^-$  不同时为无穷时,  $EX := \int_{-\infty}^{\infty} x dF(x)$  良定, 称为  $X$  (或  $\mathcal{L}(X)$ ) 的期望.
- 定义与离散型、连续型的一致.
- 重心:



- $E\vec{X} = (EX_1, \dots, EX_n).$

- 计算公式:

$$\begin{aligned} EX^+ &= \int_0^{\infty} P(X^+ \geq x) dx = \int_0^{\infty} P(X \geq x) dx, \\ EX^- &= \int_0^{\infty} P(X^- \geq x) dx = \int_0^{\infty} P(X \leq -x) dx. \end{aligned}$$

$$EX = EX^+ - EX^- = \int_0^{\infty} G(x) dx - \int_{-\infty}^0 F(x) dx.$$

- 例3.1.21.  $E|X| = \int_0^{\infty} P(|X| \geq x) dx = EX^+ + EX^-.$
- 推论3.1.22.  $|X| \leq M$  a.s., 则  $E|X| \leq M.$

## 五、随机变量函数的数学期望

离散型(随机向量类似):  $E f(X) = \sum_{k \in I} f(x_k) p_k$ .

- 例3.1.24.  $X \sim P(\lambda)$ .  $E X(X-1) \cdots (X-r+1) = \lambda^r$ .

- $f(k) = 0, k = 0, 1, \cdots, r-1$ ;

$$f(k)p_k = \frac{\lambda^k}{(k-r)!} e^{-\lambda} = \lambda^r \cdot p_{k-r}, \quad k \geq r$$

$$\Rightarrow E f(X) = \sum_{k=r}^{\infty} f(k)p_k = \lambda^r \sum_{k=r}^{\infty} p_{k-r} = \lambda^r.$$

- 例3.1.25.  $X \sim B(n, p)$ , 类似.
- 例3.1.10 (古典概型).  $\Omega = \{1, 2, \cdots, n\}$ , 则

$$EX = \frac{1}{n} (X(1) + \cdots + X(n)).$$

- 令  $Y: \Omega \rightarrow \mathbb{R}, i \mapsto i; f(i) := X(i)$ , 则  $X = f(Y): i \rightarrow X(i)$ .

连续型(随机向量类似):

$$Ef(X) = \int_{-\infty}^{\infty} f(x) \cdot p_X(x) dx.$$

● 命题3.1.26.  $f \geq 0$  且严格 $\nearrow$ , 则上式成立.

● 证:  $Ef(X) = \int_0^{\infty} P(f(X) > y) dy$ :

$$\begin{aligned} &= \int_0^{\infty} P(X > f^{-1}(y)) dy = \int_0^{\infty} \int_{f^{-1}(y)}^{\infty} p_X(x) dx dy \\ &= \int_{-\infty}^{\infty} \int_0^{f(x)} p_X(x) dy dx = \int_{-\infty}^{\infty} f(x) \cdot p_X(x) dx. \end{aligned}$$

或者, 先求 $f(X)$  的密度, 再求期望.

●  $k$  阶矩:  $EX^k$ ;  $k$  阶中心矩:  $E(X - EX)^k$ .

例3.1.29. 设 $r > s > 0$ . 若 $E|X|^r < \infty$ , 则 $E|X|^s < \infty$ .

- 当 $x \geq 1$  时,  $x^{1/r} < x^{1/s}$ . 因此

$$P(|X| > x^{1/s}) \leq P(|X| > x^{1/r}).$$

- 于是

$$\begin{aligned} E|X|^s &= \int_0^1 P(|X| > x^{1/s})dx + \int_1^\infty P(|X| > x^{1/s})dx \\ &\leq 1 + \int_1^\infty P(|X| > x^{1/r})dx \leq 1 + E|X|^r. \end{aligned}$$

- 特别地, 又设 $s \geq 1$ , 则 $(E|X|^s)^{1/s} \leq (E|X|^r)^{1/r}$ .

例3.1.30.  $\forall r > 0$  和  $a > 0$ ,

$$\mathbb{E}|X|^r \cdot \mathbf{I}_{\{|X| \leq a\}} = \int_0^a r x^{r-1} \mathbb{P}(x < |X| \leq a) dx \leq \int_0^a r x^{r-1} \mathbb{P}(|X| > x) dx.$$

- $Y \geq 0$  a.s., 则

$$\mathbb{E}Y \cdot \mathbf{I}_A = \int_0^\infty \mathbb{P}(Y \cdot \mathbf{I}_A > y) dy = \int_0^\infty \mathbb{P}(A \cap \{Y > y\}) dy.$$

- $\{|X| \leq a, |X|^r > y\} = \{y^{1/r} < |X| \leq a\}.$
- 记  $x = y^{1/r}$ .  $x > a$  时以上集合为空集, 故结论成立.

## §3.2 数学期望的性质

### 一、数学期望的线性性质

- 命题3.2.1.  $EX$  存在, 则  $E(aX)$  存在且

$$E(aX) = aEX.$$

- 命题3.2.2. 设  $EX$  和  $EY$  都存在, 若不是  $\pm\infty$  情形, 则  $E(X + Y)$  存在且

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

- 由归纳法,

$$E(X_1 + \cdots + X_n) = EX_1 + \cdots + EX_n.$$

## 1. 随机数目的数学期望

- $X = X_1 + \cdots + X_n$ , 其中  $X_i = \mathbf{I}_{A_i}$ , 则

$$EX = P(A_1) + \cdots + P(A_n).$$

- 例3.2.3.  $X \sim B(n, p)$  或  $X \sim H(N, M, n)$ , 则

$$EX = np, \quad p = M/N.$$

- 例3.2.4.  $n$  双不同颜色的鞋, 随机匹配成功  $X$  双, 则  $EX = 1$ .
- 例2.5.6(小数定律).  $EX = p_1 + \cdots + p_n = \lambda$ .

## 2. 具有对称性的模型

- 例3.2.6. 设 $U_1, \dots, U_n$  独立,  $\sim U(0, 1)$ , 则 $EU_{(i)} = \frac{i}{n+1}$ .
  - 记 $X_i = U_{(i)} - U_{(i-1)}$ , ( $U_{(0)} = 0, U_{(n+1)} = 1$ ).
  - $X_1, \dots, X_{n+1}$  可交换, 故 $EX_i \equiv c$ .
  - $X_1 + \dots + X_{n+1} \equiv 1$ , 故 $c = 1/(n+1)$ .
- 例3.2.7. 设 $X_1, \dots, X_n$  独立同分布, 且 $X_1 > 0$  a.s., 则

$$E \frac{X_1 + \dots + X_k}{X_1 + \dots + X_n} = \frac{k}{n}, \quad 1 \leq k \leq n.$$

- 记 $Y_i = \frac{X_i}{X_1 + \dots + X_n}$ .
- $X_1, \dots, X_n$  可交换, 故 $Y_1, \dots, Y_n$  可交换, 故 $EY_i \equiv 1/n$ .

例2.6.4 & 3.2.8.  $\mathbb{R}^3$  的单位球面  $S$  上,  $P(\text{四点不共半球}) = 1/8$ .

- 设  $\xi_i$ 's i.i.d.,  $\sim U(S)$ .  $A = \{\xi_4 \in \triangle_{-\xi_1, -\xi_2, -\xi_3}\}$ .  $P(A) = E\Delta$ .
- 独立地取独立同分布的  $W_1, W_2, W_3$ , 等可能取  $\pm 1$ .
- $\eta_i = W_i \xi_i$ , 则根据对称性  $E\Delta = 1/8$ .
- $B = \text{“}\eta_1, \eta_2, \eta_3, \xi_4 \text{不共半球”} = \{\xi_4 \in \triangle_{-\eta_1, -\eta_2, -\eta_3}\}$ .
- $\forall \alpha_1, \alpha_2, \alpha_3$ ,  $P(B | \xi_1 = \alpha_1, \xi_2 = \alpha_2, \xi_3 = \alpha_3, \xi_4 = \alpha_4) \equiv 1/8$ .
- 故  $P(B) = 1/8$ . ( $B$  与  $\xi_i$ 's 独立.)

## 二、数学期望的其他性质

命题3.2.11. 若 $X$ 与 $Y$ 相互独立且它们的期望都存在且有限, 则 $E(XY)$ 存在且

$$E(XY) = (EX)(EY).$$

- 连续型:

$$EXY = \iint_{\mathbb{R}^2} xy \underbrace{p_X(x)p_Y(y)}_{\text{联合密度}} dx dy = \int_{-\infty}^{\infty} xp_X(x) dx \int_{-\infty}^{\infty} yp_Y(y) dy.$$

- 一般情形:  $\mu_{X,Y}(B \times D) = \mu_X(B)\mu_Y(D)$ ,

$$EXY = \iint_{\mathbb{R}^2} xy \underbrace{d\mu_X d\mu_Y}_{\text{乘积测度}} = \int_{-\infty}^{\infty} \text{id}_X d\mu_X \int_{-\infty}^{\infty} \text{id}_Y d\mu_Y.$$

命题3.2.13. 设 $EX$  与 $EY$  都存在且有限.

若 $X \geq Y$  a.s., 则 $EX \geq EY$ .

- 若 $X, Y$  都非负, 则

$$EY = \int_0^{\infty} P(Y > x)dx \leq \int_0^{\infty} P(X > x)dx = EX.$$

- 一般情形,  $X^+ \geq Y^+, X^- \leq Y^-$ , 故

$$EX = EX^+ - EX^- \geq EY^+ - EY^- = EY.$$

定理3.2.14 (Markov 不等式). 设  $X \geq 0$  a.s. (且  $EX < \infty$ ), 则

$$P(X \geq \varepsilon) \leq \frac{EX}{\varepsilon}, \quad \forall \varepsilon > 0.$$

- 记  $A = \{X \geq \varepsilon\}$ . 目标:  $P(A)$  的上估计.
- $P(A) = E\mathbf{I}_A$ .
- 找  $Y \geq \mathbf{I}_A$ , 满足

$$Y \geq 0; \quad Y|_A \geq 1.$$

- 于是  $P(A) = E\mathbf{I}_A \leq EY$ .
- 取  $Y = X/\varepsilon$  即可.
- 例. 一般情形,

$$P(X \geq C) = P(e^{a(X-C)} \geq 1) \leq Ee^{a(X-C)}, \quad \forall a > 0.$$

- 推论3.2.15. 设 $E|X - EX|^r$  存在(且有限), 其中 $r > 0$ , 则

$$P(|X - EX| \geq \varepsilon) \leq \frac{E|X - EX|^r}{\varepsilon^r}, \quad \forall \varepsilon > 0.$$

- 取 $Y = |X - EX|^r/\varepsilon^r$  即可.
- 特别地,  $r = 2$  时称为Chebyshev 不等式.
- 推论3.2.16. 若 $E|X| = 0$ , 则 $P(X = 0) = 1$ .
  - $P(|X| > 1/n) \leq n \cdot E|X| = 0$ .
  - $P(|X| > 0) = \downarrow \lim_{n \rightarrow \infty} P(|X| > 1/n) = 0$ .
- 命题3.2.17 (Jensen 不等式). 设 $X$  是随机变量,  $EX$  存在且有限,  $f(\cdot)$  是凸函数且 $Ef(X)$  存在且有限, 则 $f(EX) \leq Ef(X)$ . (注: 随机向量类似.)

### 三、数学期望与尾分布函数的收敛性

命题3.2.18. 设  $X \geq 0$  a.s. 且  $EX < \infty$ , 则当  $x \rightarrow \infty$  时,

$$xP(X > x) \rightarrow 0, \quad EX \cdot \mathbf{I}_{\{X > x\}} \rightarrow 0.$$

●  $xG(x) \rightarrow 0$ :

$$xG(x) = 2 \int_{x/2}^x G(y) dy \leq 2 \int_{x/2}^x G(y) dy \leq 2 \int_{x/2}^{\infty} G(y) dy \rightarrow 0.$$

●  $EX \cdot \mathbf{I}_{\{X > x\}} \rightarrow 0$ :

$$\begin{aligned} EX \cdot \mathbf{I}_{\{X > x\}} &= \int_0^{\infty} P(X > x, X > y) dy \\ &= \int_0^x P(X > x) dy + \int_x^{\infty} P(X > y) dy \rightarrow 0. \end{aligned}$$

命题3.2.19. 设 $EX$  存在, 则当 $x \rightarrow \infty$  时,

$$EX \cdot \mathbf{I}_{\{|X| \leq x\}} \rightarrow EX.$$

- $(X \cdot \mathbf{I}_{\{|X| \leq x\}})^- = X^- \cdot \mathbf{I}_{\{|X| \leq x\}}$ . 故  $E\star = E\star\star \leq EX^-$ .
- $\forall y > 0, \{\star > y\} = \{\underbrace{X^-}_{\text{blue}} > y, \underbrace{|X| \leq x}_{\text{red}}\} = \{\underbrace{-X > y}_{\text{blue}}, |X| \leq x\}$ .

$$\begin{aligned} E\star\star &= \int_0^\infty P(y < -X \leq x) dy \geq \int_0^M P(y < -X \leq x) dy \\ &= \int_0^M P(-X > y) dy - M \cdot P(-X > x) \\ &\xrightarrow{x \rightarrow \infty} \int_0^M P(X^- > y) dy \xrightarrow{M \rightarrow \infty} EX^-. \end{aligned}$$

- $\liminf_{x \rightarrow \infty} E\star \geq EX^-$ , 从而  $E\star \xrightarrow{x \rightarrow \infty} EX^-$ .
- 类似地,  $\lim_{x \rightarrow \infty} E(X \cdot \mathbf{I}_{\{|X| \leq x\}})^+ \xrightarrow{x \rightarrow \infty} EX^+$ . 故结论成立.

### §3.3 条件期望

#### 一、在事件 $A$ 发生的条件下的数学期望

- 设 $P(A) > 0$ . 用 $P_A(\cdot)$  代替 $P(\cdot)$  计算期望.
- 非负情形:  $Y \geq 0$  a.s.,

$$\begin{aligned} E_A Y &= \int_0^{\infty} P_A(Y > y) dy = \int_0^{\infty} \frac{P(\{Y > y\} \cap A)}{P(A)} dy \\ &= \frac{1}{P(A)} \int_0^{\infty} P(Y \cdot \mathbf{I}_A > y) dy = \frac{E(Y \cdot \mathbf{I}_A)}{P(A)}. \end{aligned}$$

- 一般情形: 若 $E(Y^+ \cdot \mathbf{I}_A)$  与 $E(Y^- \cdot \mathbf{I}_A)$  不同时发散, 则

$$E_A Y = E_A Y^+ - E_A Y^- = \frac{E(Y^+ \cdot \mathbf{I}_A) - E(Y^- \cdot \mathbf{I}_A)}{P(A)} = \frac{E(Y \cdot \mathbf{I}_A)}{P(A)}.$$

- $EY$  存在/有限  $\Rightarrow E_A Y$  存在/有限.
- $Y$  在 $A$  中的加权总和与加权平均值:  $P(A) \cdot \star = \star$ .
- 命题3.3.1. 设 $Y$  与 $A$  独立, 则 $E_A Y = EY$ .

## 二、离散型

- 称 $\varphi(X)$  为 $Y$  关于 $X$  的条件期望, 记为 $\boxed{E(Y|X)}$ , 其中

$$\varphi(x_i) := E(Y|X = x_i) = \sum_{j \in J} y_j P(Y = y_j | X = x_i), \quad \forall i \in I.$$

- 命题3.3.2. (1)  $E(c|X) = c$ ; (2)  $E(X|X) = X$ ;  
(3) 若 $X, Y$  独立, 则 $E(Y|X) = EY$ .
- 命题3.3.3. (1)  $E(Y + Z|X) = E(Y|X) + E(Z|X)$ .  
(2)  $E(f(X) \cdot Y|X) = f(X) \cdot E(Y|X)$ .
  - 证(2):  $\psi(x_i) = E(W|A_i) = E(f(x_i) \cdot Y|A_i)$   
 $= f(x_i)E(Y|A_i) = f(x_i)\varphi(x_i)$ .
- 注: 随机向量类似.

命题3.3.4 (重期望公式). 设 $EY$  存在且有限, 则 $EY = EE(Y|X)$ .

- $\varphi(x_i) = E(Y|X = x_i),$

$$P(X = x_i)\varphi(x_i) = E(Y \cdot \mathbf{I}_{\{X=x_i\}}).$$

- $E\varphi(X) = \sum_{i \in I} P(X = x_i)\varphi(x_i) = \sum_{i \in I} E(Y \cdot \mathbf{I}_{\{X=x_i\}}) = EY.$

- 以下级数绝对收敛:

$$\begin{aligned} E\varphi(X) &= \sum_i \varphi(x_i)P(X = x_i) \\ &= \sum_i \sum_j y_j P(X = x_i, Y = y_j) = EY. \end{aligned}$$

例3.3.5 (Polya 坛子).  $X_n$  和  $Y_n$ :  $n$  次后的黑球数和黑球比例.  
求  $E(Y_{n+1}|Y_n)$  和  $EY_{n+1}$ .

- $X_n, Y_n$  的可能值:  $x_i = b + ic, y_i = \frac{x_i}{b + r + nc}, i = 0, \dots, n$ .
- $X_{n+1} = X_n + c \cdot \mathbf{I}_{B_{n+1}}$ , 其中  $B_{n+1} =$  “第  $n+1$  次摸到黑球”.
- $P(B_{n+1}|Y_n = y_i) = y_i$ , 故

$$P(B_{n+1}|Y_n) := E(\mathbf{I}_{B_{n+1}}|Y_n) = Y_n.$$

从而  $E(X_{n+1}|Y_n = y_i) = x_i + c \cdot y_i = y_i(b + r + (n+1)c)$ .

- $E(Y_{n+1}|Y_n = y_i) = y_i$ , 故

$$E(Y_{n+1}|Y_n) = Y_n.$$

- $EY_{n+1} = EY_n = \dots = \frac{b}{b+r}, P(B_{n+1}) = EY_n = \frac{b}{b+r}$ .

命题3.3.7. 设 $Y_1, Y_2, \dots$  独立同分布;  $X$  与之独立, 取值于 $\mathbb{Z}_+$ ; 期望均存在且有限. 令  $Z = Y_1 + \dots + Y_X$ , 则 $EZ = (EX)(EY_1)$ .

- 记 $S_n = Y_1 + \dots + Y_n$ , 则

$$\mathcal{L}(Z|X = n) = \mathcal{L}(S_n|X = n).$$

- $\varphi(n) = E(Z|X = n) = E(S_n|X = n) = ES_n = n \cdot \mu_Y$ .
- $E(Z|X) = \varphi(X) = X \cdot \mu_Y$ .
- $EZ = EE(Z|X) = E(X \cdot \mu_Y) = \mu_X \cdot \mu_Y$ .

例3.3.6 (Bernoulli 试验).  $T_k$ : 首次连续成功 $k$ 次的时刻. 求 $ET_k$ .

- 记 $t_k = ET_k$ . 先求 $E(T_{k+1}|T_k)$ .
- 方法一、若 $A_{k,n} = \{T_k = n\}$  发生, 则

$$T_{k+1} = n + 1 + T_k^{(n+1)} \cdot \mathbf{I}_{\{\xi_{n+1}=0\}},$$

其中 $T_k^{(n+1)}$  = 第 $n+2$ 次起连续成功 $k$ 次所需的试验次数.

- $E(T_{k+1}|T_k = n) = n + 1 + (1-p)t_{k+1}$ .
- $E(T_{k+1}|T_k) = T_k + 1 + (1-p)t_{k+1} \Rightarrow t_{k+1} = t_k + 1 + (1-p)t_{k+1}$ .
- 方法二、 $T_{1,k} = T_k, T_{2,k}, \dots$  i.i.d.,  $X$  与之独立,  $X \sim G(p)$ ,

$$T_{k+1} = (T_{1,k} + 1) + \dots + (T_{X,k} + 1).$$

- $t_{k+1} = (t_k + 1) \cdot \frac{1}{p}$ . 由 $t_1 = \frac{1}{p}$  解得 $t_k = \frac{1 - p^k}{p^k(1-p)}$ .

### 三、连续型

- 称 $\varphi(X)$  为 $Y$  关于 $X$  的条件期望, 记为 $E(Y|X)$ , 其中

$$\varphi(x) := \int_{-\infty}^{\infty} yp_{Y|X}(y|x)dy.$$

- 一般情形:  $\varphi(x)$  为 $\mathcal{L}(Y|X=x)$  的期望,  $E(Y|X) := \varphi(X)$ .
- 注: 随机向量类似, 性质都与离散型类似.

例3.1.9 & 3.3.8. 设 $U_1, U_2, \dots$  独立同分布,  $\sim U(0, 1)$ . 求 $EX$ .

$$X = \inf\{n \geq 1 : S_n > 1\}, \quad S_n = U_1 + \dots + U_n.$$

- 方法一、 $X$  取正整数且 $X \geq 2$ .  $G(1) = 1$ ; 对 $n \geq 2$ ,

$$G(n) = P(S_n \leq 1) = \int_{0 < s_1 < \dots < s_n < 1} 1 \, ds_1 \cdots ds_n = \frac{1}{n!}.$$

其中, 由 $s_n < 1$  可推出 $s_r - s_{r-1} < 1, r = 1, \dots, n$ .

- $EX = \sum_{n=0}^{\infty} G(n) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} = e$ .
- 类似地, 记 $X_a = \inf\{n \geq 1 : S_n > a\}, a \leq 1$ .

$$G_a(n) = P(X_a > n) = P(S_n \leq a) = \frac{a^n}{n!}.$$

- $EX_a = 1 + \sum_{n=1}^{\infty} \frac{a^n}{n!} = e^a$ .

例3.1.9 & 3.3.8(续). 记  $X_a := \inf\{n \geq 1 : S_n > a\}$ .

- 方法二、  $P(U_1 + U_2 > 1) > 0$ , 故  $EX < \infty$ .
- $X_a = 1 + \mathbf{I}_{\{U_1 \leq a\}} \cdot \hat{X}_{a-U_1}$ , 其中

$$\hat{X}_b = \inf\{n \geq 1 : U_2 + \cdots + U_{n+1} > b\}.$$

- $E(\star\star|U_1) = \underbrace{\mathbf{I}_{\{U_1 \leq a\}} \cdot f(a - U_1)}_{\text{wavy line}}:$

$$\varphi(x) = E(\star\star|U_1 = x) = \mathbf{I}_{\{x \leq a\}} \cdot E(\hat{X}_{a-x}|U_1 = x) = \mathbf{I}_{\{x \leq a\}} \cdot f(a-x).$$

- $f(a) : EX_a = 1 + E\star\star$ , 其中

$$E\star\star = E\underset{\text{wavy line}}{\star \cdot \star} = \int_0^a f(a-x)dx = \int_0^a f(y)dy.$$

- $f'(a) = f(a) \Rightarrow f(a) = Ce^a$ . 由  $f(0+) = 1 \Rightarrow C = 1$ .

## §3.4 方差

### 一、方差的定义

- 定义3.4.1. 设 $E|X| < \infty$ .

称 $E(X - EX)^2$  为 $X$  的方差, 记为 $\text{Var}(X)$  或 $D(X)$ .

称 $\sqrt{\text{Var}(X)}$  为 $X$  的标准差, 记为 $\sigma_X$ .

- $\mathcal{L}(X)$  的方差/标准差, 计算公式: (记 $\mu = EX$ ),

$$(X - \mu)^2 = X^2 - 2\mu \cdot X + \mu^2 \Rightarrow \underbrace{\text{Var}(X) = EX^2 - \mu^2}.$$

- $EX^2 < \infty \Rightarrow E|X| < \infty$ . 设 $E|X| < \infty$ , 则

$$EX^2 < \infty \Leftrightarrow \text{Var}(X) < \infty, \quad (\text{方差有限}),$$

$$EX^2 = \infty \Leftrightarrow \text{Var}(X) = \infty, \quad (\text{方差发散}).$$

- 例3.4.2.  $X \sim B(1, p)$ ,  $X^k = X$ ,  $\text{Var}(X) = p - p^2 = pq$ .

## 二、线性变换的方差

● 命题3.4.8.  $\text{Var}(aX + b) = \text{E}a^2(X - a\text{E}X)^2 = a^2\text{Var}(X)$ .

● 例3.4.9.  $X \sim N(\mu, \sigma^2)$ ,  $\text{Var}(X) = \sigma^2$ .

●  $Z := X^* = (X - \mu)/\sigma \sim N(0, 1)$ ,  $X = \mu + \sigma Z$ .

●  $\text{Var}(Z) = \text{E}Z^2 = 1$ :

$$\frac{2}{\sqrt{2\pi}} \int_0^\infty x^2 e^{-\frac{x^2}{2}} dx = -\frac{2}{\sqrt{2\pi}} \int_0^\infty x de^{-\frac{x^2}{2}} = \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2}} dx = 1.$$

● 称

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

为 $X$  的标准化.

### 三、不等式

- 定理3.4.14 (Chebyshev 不等式).  $EX = \mu$ ,  $\text{Var}(X) = \sigma^2$ , 则

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}, \quad \forall \varepsilon > 0.$$

- 例3.4.16. 设  $EX = 0$ ,  $\text{Var}(X) = \sigma^2$ , 则

$$P(X \geq \varepsilon) \leq \frac{\sigma^2}{\sigma^2 + \varepsilon^2}, \quad \forall \varepsilon > 0.$$

- 推论3.4.15. 若  $\text{Var}(X) = 0$ , 则  $X$  退化.
- 方差刻画权重的分散程度:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}, \quad \forall k > 0.$$

- $\mu \gg \sigma$  (即  $\mu^2 \gg \sigma^2$ ) 时,  $X \approx \mu$ :

$$P\left(\left|\frac{X}{\mu} - 1\right| \geq a\right) \leq \frac{1}{a^2} \cdot \frac{\sigma^2}{\mu^2}, \quad \forall a > 0.$$

## 随机图:

例3.2.5 & 3.4.13 & 3.6.4. 完全图 $K_n$  中的每条边, 独立地以概率 $p$  保留, 以概率 $1 - p$  删除. 得到随机(子)图 $G(n, p)$ .

- 记 $X = G(n, p)$  中的三角形 $\triangle$  的数目, 则

$$X = \mathbf{I}_{A_1} + \cdots + \mathbf{I}_{A_{C_n^3}},$$

其中,  $A_i = “\triangle_i \text{ 在 } G(n, p) \text{ 中}”$ .

- 计算期望:  $P(A_i) = p^3$ ,  $EX = C_n^3 p^3$ .

随机图(续).

- 计算方差:  $X = \mathbf{I}_{A_1} + \cdots + \mathbf{I}_{A_{C_n^3}} = \sum_{i=1}^{C_n^3} \mathbf{I}_{A_i},$

$$E(X^2) - (EX)^2 = E\left(\sum_{i,j} \mathbf{I}_{A_i} \mathbf{I}_{A_j}\right) - \sum_{i,j} (E\mathbf{I}_{A_i})(E\mathbf{I}_{A_j}),$$

- $\text{Var}(X) = \sum_i \sigma_i^2 + \sum_{i \neq j} \sigma_{ij},$  其中

$$\sigma_i^2 = P(A_i) - P(A_i)^2, \quad \sigma_{ij} = P(A_i A_j) - P(A_i)P(A_j).$$

- $\text{Var}(X) = C_n^3 \left( \sigma_1^2 + \sum_{j=2}^{C_n^3} \sigma_{1j} \right).$  (命题3.4.11)

- $\sigma_1^2 = p^3(1-p^3).$  下面研究  $\sum_{j=2}^{C_n^3} \sigma_{1j}.$

随机图(续).

- $3(n-3)$  个  $\triangle_j$  与  $\triangle_1$  有一条共同的边,  $\sigma_{1j} = p^5 - p^6$ .
- 0 个  $\triangle_j$  与  $\triangle_i$  有两条或三条共同的边,
- 其他  $\triangle_j$  与  $\triangle_i$  无共同的边,  $\sigma_{1j} = 0$ .
- 综上,  $\sum_{j=2}^{C_n^3} \sigma_{1j} = 3(n-3)p^5(1-p)$ .
- 因此

$$\text{Var}(X) = C_n^3 p^3 (1-p^3) + C_n^3 3(n-3)p^5(1-p).$$

随机图(续).

- $\mu := EX = C_n^3 p^3,$   
 $\sigma^2 := \text{Var}(X) = C_n^3 p^3 (1 - p^3) + C_n^3 3(n-3)p^5(1-p).$
- 令  $n \rightarrow \infty, p \rightarrow 0.$

$$\mu^2 = O((np)^6), \quad \sigma^2 = O((np)^3) + o((np)^4).$$

- 若  $np \rightarrow 0$ , 则

$$P(\text{存在} \triangle) = P(X \geq 1) \leq EX \leq \frac{1}{6}(np)^3 \rightarrow 0.$$

- 若  $np \rightarrow \infty$ , 则

$$P(\text{不存在} \triangle) = P(X = 0) \leq P(|X - \mu| \geq \mu) \leq \frac{\sigma^2}{\mu^2} \rightarrow 0.$$

## §3.5 协方差和相关系数

### 协方差:

- 设  $EX^2 < \infty, EY^2 < \infty$ .  
 $|XY| \leq (X^2 + Y^2)/2 \Rightarrow E|XY| < \infty$ .
- 定义3.5.5. 称  $\underline{E(X - EX)(Y - EY)}$  为  $X, Y$  的协方差, 记为  $\text{Cov}(X, Y)$ , 或  $\sigma_{X,Y}$ , 或  $\sigma_{XY}$ .
- 计算公式:  $\text{Cov}(X, Y) = E(XY) - (EX)(EY)$ .
- 和的方差:  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$ .
- 若  $\underline{\star\star} = (\geq, \leq) 0$ , 则称  $X$  与  $Y$  不(正, 负)相关.
- 命题3.5.6. 设  $X_1, \dots, X_n$  两两不相关, 则

$$\text{Var}(X_1 + \dots + X_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n).$$

## 和的方差vs协方差:

例3.4.7 & 3.5.13. 设 $U_1, \dots, U_n$  独立同分布,  $\sim U(0, 1)$ .

记 $X_i = U_{(i)} - U_{(i-1)}$ . 求 $\text{Cov}(X_i, X_j)$ .

- $EX_1 = \frac{1}{n+1}, \text{Var}(X_1) = \frac{n}{(n+1)^2(n+2)} \triangleq \sigma^2:$

$$EX_1^2 = \int_0^1 P(X_1^2 > y) dy = \int_0^1 P(U_1 > x)^n dx^2 = \int_0^1 (1-x)^n 2x dx.$$

- 由可交换性,  $\text{Var}(X_i) = \sigma^2; \forall i \neq j, \text{Cov}(X_i, X_j) \triangleq c.$

- $X_1 + \dots + X_{n+1} \equiv 1$ , 故 $\text{Var}(\star) = 0$ .

- $\text{Var}(\star) = (n+1)\sigma^2 + n(n+1)c$ . 故 $c = -\frac{1}{n}\sigma^2$ .

- 进一步,  $U_{(i)} = X_1 + \dots + X_i$ , 故

$$EU_{(i)} = i/(n+1), \quad \text{Var}(U_{(i)}) = i\sigma^2 + i(i-1)c.$$

## 不相关vs独立:

- 命题3.5.7.  $X, Y$  相互独立, 则  $X, Y$  不相关.
- “不相关 $\nRightarrow$ 独立”: 例3.5.8.  $X \sim N(0, 1)$ .
  - $\text{Cov}(X, X^2) = EX^3 - (EX)(EX^2) = 0$ .
  - $\mathcal{L}(X^2|X=x)$  是退化的,  $\mathcal{L}(X|X^2=a)$  是两点分布( $a > 0$ ).
- “不相关 $\Leftrightarrow$ 独立”的特例1: 例3.5.9.  $X = \mathbf{I}_A, Y = \mathbf{I}_B$ .
  - $X, Y$  不相关 $\Leftrightarrow E\mathbf{I}_A\mathbf{I}_B = (E\mathbf{I}_A)(E\mathbf{I}_B)$

$$\Leftrightarrow P(AB) = P(A)P(B) \Leftrightarrow A, B \text{ 独立} \Leftrightarrow X, Y \text{ 独立}.$$

- 例3.5.10. 若  $X, Y$  正相关(即  $P(AB) \geq P(A)P(B)$ ),  
则称  $A$  与  $B$  正相关. (注: 负相关类似, 例3.5.11, 不放回抽样.)
- “不相关 $\Leftrightarrow$ 独立”的特例2: 二维正态的两个边缘.

正相关:

例3.5.14. 设 $f, g$  有界、 $\nearrow$ . 证明:  $\forall X, f(X)$  与 $g(X)$  正相关.

- $Ef(X)^2, Eg(X)^2 < \infty$ .

- 取与 $X$  独立同分布的随机变量 $Y$ , 则

$$Z := (f(X) - f(Y))(g(X) - g(Y)) \geq 0. \text{ 故 } EZ \geq 0.$$

- $EZ = 2\text{Cov}(f(X), g(X))$ :

$$\begin{aligned} EZ &= E(f(X)g(X) + f(Y)g(Y) - f(X)g(Y) - f(Y)g(X)) \\ &= Ef(X)g(X) + Ef(Y)g(Y) - Ef(X)g(Y) - Ef(Y)g(X) \\ &= 2Ef(X)g(X) - 2(Ef(X))(Eg(X)). \end{aligned}$$

- FKG 不等式. 设 $X_1, \dots, X_n$  独立;  $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ , “ $\nearrow$ ”, 且 $f(X_1, \dots, X_n), g(X_1, \dots, X_n) \in L^2$ , 则 $\star$  与 $\star$  正相关.

协方差矩阵:

- $\vec{X} = (X_1, \dots, X_n)$  的协方差矩阵:

$$\Sigma = (\sigma_{ij})_{n \times n}, \quad \sigma_{ij} = \text{Cov}(X_i, X_j).$$

- 命题3.5.16.  $\Sigma$  半正定.  $\Sigma$  正定当且仅当

$$a_0 + a_1 X_1 + \dots + a_n X_n = 0 \text{ a.s.} \quad \Rightarrow \quad a_0 = a_1 = \dots = a_n = 0.$$

- $\Sigma$  对称.  $\Sigma$  半正定:

$$\begin{aligned} \vec{a}^T \Sigma \vec{a} &= \sum_{i,j=1}^n a_i \sigma_{ij} a_j = \sum_{i,j=1}^n a_i a_j \mathbb{E}(X_i - \mathbb{E}X_i)(X_j - \mathbb{E}X_j) \\ &= \mathbb{E} \left( \sum_{i=1}^n a_i (X_i - \mathbb{E}X_i) \right)^2 \geq 0. \end{aligned}$$

- $\Sigma$  正定指:

$$\sum_{i=1}^n a_i (X_i - \mathbb{E}X_i) = 0 \text{ a.s.} \quad \Rightarrow \quad a_1 = \dots = a_n = 0.$$

- $a_0 := -a_1 \mathbb{E}X_1 - \dots - a_n \mathbb{E}X_n.$

## 相关系数:

- 定义3.5.17. 设  $0 < \text{Var}(X) < \infty, 0 < \text{Var}(Y) < \infty$ . 称

$$\frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

为  $X$  与  $Y$  的相关系数, 记为  $\rho_{XY}$ .

- 不相关:  $\rho_{X,Y} = 0$ .
- $\rho_{XY} = \rho_{X^*Y} = \rho_{XY^*} = \rho_{X^*Y^*}$ :

$$\rho_{XY} = \frac{\text{E}(X - \text{E}X)(Y - \text{E}Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}} = \text{E}X^*Y^*.$$

例3.5.19. 设  $(X, Y) \sim N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2; \rho)$ , 则  $\rho_{XY} = \rho$ .

$$\bullet u = \frac{x - \mu_1}{\sigma_1}, v = \frac{y - \mu_2}{\sigma_2}, C = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

$$p(x, y) = C \exp \left\{ -\frac{u^2 - 2\rho uv + v^2}{2(1-\rho^2)} \right\} = C \exp \left\{ -\frac{(v - \rho u)^2}{2(1-\rho^2)} - \frac{u^2}{2} \right\}.$$

$$\bullet \rho_{XY} = \rho: X^* = \frac{X - \mu_1}{\sigma_1}, Y^* = \frac{Y - \mu_2}{\sigma_2},$$

$$\begin{aligned} EX^*Y^* &= \iint_{\mathbb{R}^2} uv \cdot C \exp \left\{ -\frac{(v - \rho u)^2}{2(1-\rho^2)} - \frac{u^2}{2} \right\} \underbrace{\sigma_1\sigma_2 du dv}_{\text{differential}} \\ &= C\sigma_1\sigma_2 \int_{-\infty}^{\infty} u \exp \left\{ -\frac{u^2}{2} \right\} \left( \int_{-\infty}^{\infty} v \exp \left\{ -\frac{(v - \rho u)^2}{2(1-\rho^2)} \right\} dv \right) du \\ &= C\sigma_1\sigma_2 \int_{-\infty}^{\infty} u \exp \left\{ -\frac{u^2}{2} \right\} \left( \sqrt{2\pi(1-\rho^2)} \rho u \right) du. \\ &= C\sigma_1\sigma_2 \sqrt{2\pi(1-\rho^2)} \rho \sqrt{2\pi} = \rho. \end{aligned}$$

例3.5.20. 构造 $X, Y \sim N(0, 1)$ , 不相关( $EXY = 0$ ), 但不独立.

- 找不恒为0 的 $f$  使得

$$p(x, y) = \phi(x)\phi(y) + f(x)f(y), \quad \text{其中 } \phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

- 需满足 $p(x, y) \geq 0$ . 要求 $|f(x)| \leq \phi(x)$  即可.
- 需满足 $\iint_{\mathbb{R}^2} p(x, y) dx dy = 0$ , 故要求 $\int_{-\infty}^{\infty} f(x) dx = 0$ .  
要求 $f$  为奇函数 即可. 此时,
- $EXY = 0$ , 要求 $\int_{-\infty}^{\infty} xf(x) dx = 0$ .  $xf(x)$  为偶函数,  
要求 $\int_0^{\infty} xf(x) dx = 0$  即可.
- 不独立.

## Cauchy-Schwarz不等式与Hölder 不等式:

引理3.5.1. 设 $EX^2 < \infty, EY^2 < \infty$ , 则 $(EXY)^2 \leq (EX^2) \cdot (EY^2)$ .

等号成立 $\Leftrightarrow$  存在常数 $c$  使得 $Y = cX$  a.s. 或 $X = cY$  a.s..

- 下设 $EX^2 > 0$ .
- $f(t) := E(tX - Y)^2 = (EX^2) \cdot t^2 - 2(EXY) \cdot t + EY^2$ .
- 最小值点记为 $c$ , 则

$$c = \frac{EXY}{EX^2}, \quad f(c) = \frac{(EX^2) \cdot (EY^2) - (EXY)^2}{EX^2}.$$

- $f(c) \geq 0$ , 故 $\star\star$ . 等号成立 $\Leftrightarrow f(c) = 0 \Leftrightarrow Y = cX$  a.s..
- 命题3.5.4. 设 $p, q > 1$  且 $1/p + 1/q = 1$ , 则

$$|EXY| \leq (E|X|^p)^{\frac{1}{p}} (E|Y|^q)^{\frac{1}{q}}.$$

- $|\rho_{XY}| \leq 1$ .
- 正相关:  $\rho_{X,Y} \geq 0$ ; 负相关:  $\rho_{X,Y} \leq 0$ .
- 完全正相关:  $\rho_{X,Y} = 1, Y^* = X^*$ ; 完全负相关:  
 $\rho_{X,Y} = -1, Y^* = -X^*$ .
- 例3.5.18.  $U \sim U(0, 2\pi), X = \cos U, Y = \cos(U + a)$ .
  - $Y = \cos \hat{U} \stackrel{d}{=} X, \hat{U} \sim U(0, 2\pi)$ .
  - 期望、方差:  $EX = 0, EX^2 = \frac{1}{2}$ .

$$EX^2 = \frac{1}{2\pi} \int_0^{2\pi} \cos^2 u du = \frac{1}{2}.$$

- 协方差、相关系数.  $\rho_{XY} = \cos a$ :

$$\text{Cov}(X, Y) = EXY = \frac{1}{2\pi} \int_0^{2\pi} \cos u \cos(u + a) du = \frac{1}{2} \cos a.$$

- $a = 0$ :  $Y = X, \rho_{XY} = 1$ , 完全正相关.
- $a = \pi$ :  $Y = -X, \rho_{XY} = -1$ , 完全负相关.
- $a = \frac{\pi}{2}$  或  $\frac{3\pi}{2}$ :  $\rho_{XY} = 0$ , 不相关, 但是不独立.

## 几何含义:

例3.5.24 (古典概型).  $\Omega = \{1, 2, \dots, n\}$ .

- 所有随机变量  $= \mathbb{R}^n$ :

$$\varphi: X: \Omega \rightarrow \mathbb{R} \mapsto \vec{x} = (x_1, \dots, x_n): x_i = X(i).$$

- $\varphi(Y) = \vec{y} = (y_1, \dots, y_n): y_i = Y(i).$

- $\mathbb{R}^n$  中的标准内积  $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i$ , 则

$$EXY = \frac{1}{n} \sum_{i=1}^n X(i)Y(i) = \frac{1}{n} \sum_{i=1}^n x_i y_i = \frac{1}{n} \langle \vec{x}, \vec{y} \rangle.$$

- Cauchy-Schwarz不等式:

$$(EXY)^2 \leq (EX^2) \cdot (EY^2) \iff |\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\|.$$

一系列线性空间:

- $\Omega$  为非空集合,

$\mathbf{X}_1 = \Omega$ 上的所有函数(组合的集合).

- $\mathcal{F}$  是 $\Omega$  上的 $\sigma$  代数,

$\mathbf{X}_2 =$  所有满足可测性要求的函数 = 所有随机变量.

- $P$  是 $(\Omega, \mathcal{F})$  上的概率, 等价关系 $X \sim Y \Leftrightarrow X = Y$  a.s.,

$\mathbf{X}_3 = \mathbf{X}_2 / \sim$ , 几乎必然相等的随机变量视为同一元素.

- 记号上不区分 $[X]$  和 $X$ ,  $\mathbf{X}_3 \triangleq \mathbf{X}$ .

- $L^2$  是线性空间:

$$L^2 := \{X \in \mathbf{X} : EX^2 < \infty\}.$$

- $\langle X, Y \rangle := EXY$  是  $L^2$  上的内积:

(1) 对称:  $\langle X, Y \rangle = \langle Y, X \rangle$ ;

(2) (双)线性:  $\langle aX + Y, Z \rangle = a\langle X, Z \rangle + \langle Y, Z \rangle$ ;

(3) 正定:  $\langle X, X \rangle \geq 0$ , 且  $\langle X, X \rangle = 0$  当且仅当  $X = 0$  a.s..

- 长度:  $\|X\| := \sqrt{\langle X, X \rangle} = \sqrt{EX^2}$ ;

$$L^2 \text{ 距离: } d(X, Y) := \|X - Y\| = \sqrt{E(X - Y)^2}.$$

夹角:  $\langle X, Y \rangle = \|X\| \cdot \|Y\| \cdot \cos \theta$ ,  $0 \leq \theta \leq \pi$ .

- 正交,  $X \perp Y$ :  $EXY = 0$ .

- 设  $\hat{L}$  是  $L^2$  的子空间. 设  $\forall Y \in L^2$ , 存在正交分解:

$$Y = \hat{Y} + Y - \hat{Y}, \quad \text{其中 } \hat{Y} \in \hat{L}, \quad Y - \hat{Y} \perp \hat{L}.$$

- 垂足  $\hat{Y}$ :

$$\|Y - \hat{Y}\| = d(Y, \hat{L}) := \inf_{V \in \hat{L}} d(Y, V).$$

- 勾股定理:

$$\|Y\|^2 = \|\hat{Y}\|^2 + \|Y - \hat{Y}\|^2, \quad d(Y, \hat{L}) = \|Y\|^2 - \|\hat{Y}\|^2.$$

- $\hat{L}$  的正交补空间存在:

$$\hat{L}^\perp = \{W : W \perp V, \forall V \in \hat{L}\} = \{Y - \hat{Y} : Y \in L\}.$$

- 记  $X_0 = X - EX$ .

方差:  $\|X_0\|^2$ , 协方差:  $\langle X_0, Y_0 \rangle$ , 相关系数:  $\cos \theta_{X_0, Y_0}$ .

## 最优预测:

例3.5.21. 设  $EY^2 < \infty$ . 求  $f(a) := E(Y - a)^2$  的最小值/最小值点.

- $\hat{L} = \mathbb{R}$  所有常值函数  $= \text{span}\{\mathbf{1}\}$ , 是  $L^2$  的一维子空间.  
求  $d(Y, \hat{L})^2$  和垂足  $\hat{Y}$ .
- $L_*^\perp = L_0^2$ :

$$L_0^2 = \{X \in L^2 : EX = \langle X, \mathbf{1} \rangle = 0\}.$$

- 最小值点为  $\hat{Y} = EY$ :

$$Y - a_0 \perp \mathbf{1}, \quad \text{即 } E(Y - a_0) = 0, \quad \text{即 } a_0 = EY.$$

- 最小值为  $d(Y, \hat{L})^2 = \text{Var}(Y)$ :

$$\|Y\|^2 - \|\hat{Y}\|^2 = EY^2 - (EY)^2 = \text{Var}(Y).$$

例3.5.22. 设  $EX^2 < \infty, EY^2 < \infty$  且  $0 < \text{Var}(X) < \infty$ .

求  $Q(a, b) = E(Y - (a + bX))^2$  的最小值/最小值点.

- $\hat{L} = \text{span}\{1, X\}$ , 是  $L^2$  的2 维子空间. 求  $d(Y, \hat{L})^2$  和垂足  $\hat{Y}$ .
- $\hat{L}$  的标准正交基:  $1, X^*$ .
- 最小值点为  $\hat{Y} = a_0 \cdot 1 + b_0 X^*$ , 其中

$$a_0 = E(Y \cdot 1) = EY, \quad b_0 = EY X^* = \sigma_Y \cdot \rho_{XY}.$$

- 最小值为  $d(Y, \hat{L})^2 = \text{Var}(Y)(1 - \rho_{XY}^2)$ :

$$d(Y, \hat{L})^2 = \underbrace{EY^2 - a_0^2}_{\text{Var}(Y)} - b_0^2 = \underbrace{\text{Var}(Y)} - \sigma_Y^2 \rho_{XY}^2 = \sigma_Y^2 (1 - \rho_{XY}^2).$$

例3.5.23. 设 $EY^2 < \infty$ .  $X$  是任一随机变量.

求 $Q(\psi) = E(Y - \psi(X))^2$  的最小值/最小值点.

- $\hat{L} = \{\psi(X) : E\psi(X)^2 < \infty\}$ , 是 $L^2$  的子空间.

- 最小值点为 $\hat{Y} = E(Y|X)$ :

- $Y - \hat{\psi}(X) \perp \hat{L}$ , 即

$$E(\underbrace{Y - \hat{\psi}(X)}_{\perp \hat{L}} \psi(X)) = 0, \quad \forall \psi(X) \in \hat{L}.$$

- 记 $\varphi(x) = E(Y|X = x)$ , 则

$$E(\star\star|X = x) = (\varphi(x) - \hat{\psi}(x))\psi(x).$$

- 故 $E\star\star = E(\varphi(X) - \hat{\psi}(X))\psi(X)$ . 取 $\psi = \varphi - \hat{\psi}$  得证.

- 最小值为

$$d(Y, \hat{L})^2 = EY^2 - E(E(Y|X)^2) = \text{Var}(Y) - \text{Var}(E(Y|X)).$$

## §3.7 母函数

注：在本节中, 随机变量都取非负整数值.

### 一、定义

- 定义3.7.1. 设 $P(X = k) = p_k, k = 0, 1, 2, \dots$ . 称

$$g_X(s) = E s^X = \sum_{k=0}^{\infty} p_k s^k, \quad s \in [-1, 1]$$

为 $X$  (或 $\mathcal{L}(X)$ ) 的母函数(generating function).

- 命题3.7.5.  $X \stackrel{d}{=} Y$  当且仅当 $\exists \delta$  使 $g_X(s) = g_Y(s), \forall s \in [0, \delta]$ .
  - $g^{(k)}(0) = p_k \cdot k!$ .
- 例3.7.2 ~ 3.7.4.  $B(n, p), P(\lambda), G(p)$  的母函数分别为

$$(q + ps)^n = (1 + p(s - 1))^n, \quad e^{\lambda(s-1)}, \quad \frac{ps}{1 - qs}.$$

## 二、性质

- 性质1. 设  $X_1, \dots, X_n$  相互独立, 则

$$g_{X_1+\dots+X_n}(s) = g_{X_1}(s) \cdots g_{X_n}(s).$$

- $E s^{X_1+\dots+X_n} = E s^{X_1} \cdots s^{X_n} = (E s^{X_1}) \cdots (E s^{X_n}).$
- 例3.7.7.  $B(n, p)$  的母函数为  $(q + ps)^n$ .
- 性质2. 设  $P(A) = p$ ;  $\mathbf{I}_A$  与  $X$  (或  $Y$ ) 独立, 则

$$g_Z(s) = p \cdot g_X(s) + (1-p) \cdot g_Y(s), \quad \text{其中 } Z = \mathbf{I}_A \cdot X + \mathbf{I}_{A^c} \cdot Y.$$

- $E s^Z = E \mathbf{I}_A \cdot s^X + E \mathbf{I}_{A^c} \cdot s^Y = (E \mathbf{I}_A)(E s^X) + (E \mathbf{I}_{A^c})(E s^Y).$

- 性质3. 设  $X$  与  $Y = Y_1, Y_2, \dots$  (i.i.d.) 独立, 则

$$g_Z(s) = g_X(g_Y(s)), \quad \text{其中 } \boxed{Z = Y_1 + \dots + Y_X}.$$

- $E(s^Z | X = n) = E(s^{Y_1 + \dots + Y_n} | X = n) = E_\star = (Es^Y)^n.$
- 例3.7.8.  $X \sim P(\lambda)$  时称  $\boxed{Z}$  服从复合泊松分布.
  - $g_X(t) = e^{\lambda(t-1)}, g_Z(s) = e^{\lambda(g_Y(s)-1)}.$
  - 若  $Y \sim B(1, p)$ , 则  $e^{\lambda((1-p+ps)-1)} = e^{\lambda p(s-1)}$ , 即  $Z \sim P(\lambda p).$

- 性质4.  $EX(X-1)\cdots(X-n+1) = g_X^{(n)}(1-)$ .
  - $g_X^{(n)}(s) = \sum_{k=n}^{\infty} p_k k(k-1)\cdots(k-n+1)s^{k-n}$ .
- 推论3.7.9.  $EX = g'_X(1-)$ .
- 例3.7.10.  $X \sim P(\lambda)$ , 则  $EX = g'_X(1-) = \lim_{s \rightarrow 1-} \lambda e^{\lambda(s-1)} = \lambda$ .
- 推论3.7.11.  $Z = Y_1 + \cdots + Y_X$ , 则  $EZ = (EX)(EY)$ .
  - $EZ = \lim_{s \rightarrow 1-} g'_Z(s) = \lim_{s \rightarrow 1-} g'_X(g_{Y_1}(s)) \cdot g'_{Y_1}(s) = (EX)(EY_1)$ .

### 例3.7.12 (分支过程).

- $\xi_{n,i}$  = 第 $n$ 代第 $i$ 个男丁的儿子数. 设 $\xi_{n,i}$ 's 独立同分布,
- 子代分布列:  $p_k = P(\xi = k)$ . 设 $p_1 \neq 1$  (否则平凡).
- $X_n$  = 第 $n$ 代的男丁总数, 设 $X_0 = 1$ . 求 $X_n$  的母函数与期望.
- 解:  $g_{X_1}(s) = g_\xi(s)$ , 记为  $g(s)$ .
- $X_{n+1} := \sum_{i=1}^{X_n} \xi_{n,i}$ , 故  $g_{X_{n+1}}(s) = g_{X_n}(g(s))$ .
- 由归纳法,  $g_{X_n}(s) = g^{(n)}(s)$ , ( $g(\cdot)$  迭代 $n$  次).
- 记  $m = \sum_{k=0}^{\infty} k \cdot p_k$ ,  
则  $EX_{n+1} = (EX_n) \cdot (E\xi) = mEX_n = \cdots = m^n$ .

### 三、随机向量的母函数

- $g_{\vec{X}}(\vec{s}) := \mathbb{E}s_1^{X_1} \cdots s_n^{X_n}$ .
- $\vec{X} \stackrel{d}{=} \vec{Y}$  当且仅当  $g_{\vec{X}}(\cdot) \equiv g_{\vec{Y}}(\cdot)$ .
- 边缘: 例如,  $g_{X_1}(s) = g_{\vec{X}}(s, 1, \cdots, 1)$ .
- 命题3.7.13.  $X_1, \cdots, X_n$  相互独立当且仅当  $\forall \vec{s} \in [-1, 1]^n$ ,

$$g_{\vec{X}}(\vec{s}) = g_{X_1}(s_1) \cdots g_{X_n}(s_n).$$

例3.7.14 & 3.7.15. 设 $X, Y_1, Y_2, \dots$  相互独立;  $Y_n \sim B(1, p)$ . 令

$$Z = Y_1 + \dots + Y_X, W = X - Z,$$

则 $X$  服从泊松分布当且仅当 $Z$  与 $W$  相互独立.

- $g_{Z,W}(s, t) := \mathbb{E}s^Z t^W = \mathbb{E}\mathbb{E}(s^Z t^W | X)$ .
- $s^Z t^W = (s^{Y_1} t^{1-Y_1}) \dots (s^{Y_X} t^{1-Y_X})$ .
- $\star = (\mathbb{E}s^{Y_1} t^{1-Y_1})^X$ , 因此 $g_{Z,W}(s, t) = g_X(ps + (1-p)t)$ .
- $\star\star$ 当且仅当

$$g(ps + (1-p)t) \stackrel{(*)}{=} g(ps + (1-p)) \cdot g(p + (1-p)t).$$

- $\Rightarrow$ : 直接验证 $g(s) = e^{\lambda(s-1)}$  满足 $(*)$ .
- $\Leftarrow$ : 令 $f(s) = -\ln g(1-s)$ ,  $x = p(1-s)$ ,  $y = (1-p)(1-t)$ .  
将 $(*)$  改写为:  $f(x+y) = f(x) + f(y)$ , 推出 $f(x) = \lambda x$ .

## §3.8 特征函数

### 一、复数运算与复值随机变量

- 复值随机变量:  $X + iY$ , 等价于二维随机向量  $(X, Y)$ .
- 命题3.8.1.  $\|E(X + iY)\| \leq E\|X + iY\|$ .
- 命题3.8.2. 设  $(X_1, Y_1), \dots, (X_n, Y_n)$  相互独立, 则

$$E(X_1 + iY_1) \cdots (X_n + iY_n) = \left(E(X_1 + iY_1)\right) \cdots \left(E(X_n + iY_n)\right).$$

- 引理3.8.3.  $\left\| \int_a^b z(t) dt \right\| \leq \int_a^b \|z(t)\| dt.$

引理3.8.4.  $\forall n \geq 0, y \in \mathbb{R},$

$$\left\| e^{iy} - \sum_{m=0}^n \frac{(iy)^m}{m!} \right\| \leq \min \left\{ \frac{2|y|^n}{n!}, \frac{|y|^{n+1}}{(n+1)!} \right\}.$$

$$\bullet e^{ib} - e^{ia} = i \int_a^b e^{it} dt.$$

$$\bullet e^{iy} - 1 = i \int_0^y e^{iu_1} du_1,$$

$$e^{iy} - 1 - iy = i \int_0^y (e^{iu_1} - 1) du_1 = i^2 \int_0^y \int_0^{u_1} e^{iu_2} du_2 du_1.$$

$$\bullet e^{iy} - \sum_{m=0}^n \frac{(iy)^m}{m!}$$

$$= i^n \int_0^y \int_0^{u_1} \cdots \int_0^{u_{n-1}} (e^{iu_n} - 1) du_n \cdots du_2 du_1$$

$$= i^{n+1} \int_0^y \int_0^{u_1} \cdots \int_0^{u_{n-1}} \int_0^{u_n} e^{iu_{n+1}} du_{n+1} du_n \cdots du_2 du_1.$$

## 二、特征函数的定义

- 定义3.8.5. 称 $f(\cdot)$  为 $X$  (或 $\mathcal{L}(X)$ )的特征函数, 其中

$$f_X(t) := \mathbb{E}e^{itX} = \mathbb{E} \cos(tX) + i\mathbb{E} \sin(tX), \quad t \in \mathbb{R}.$$

- 例3.8.6.  $X \sim P(\lambda)$ ,  $f(t) = e^{\lambda(e^{it}-1)}$ .
- 例3.8.7.  $X \sim \text{Exp}(1)$ , 则

$$\int_0^\infty e^{ity} e^{-y} dy = \frac{1}{it-1} e^{(it-1)y} \Big|_0^\infty = \frac{1}{1-it} = \frac{1+it}{1+t^2}.$$

- 例3.8.8.  $X \sim N(0, 1)$ , 则

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{x^2}{2} + itx} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2}(x-it)^2 + \frac{1}{2}(it)^2} dx = e^{-\frac{t^2}{2}}.$$

- 命题3.8.9 (逆转公式).  $\forall a < b$ ,

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} f(t) dt \\ &= P(a < X < b) + \frac{1}{2}P(X = b) + \frac{1}{2}P(X = a). \end{aligned}$$

- 推论3.8.10. 若  $\int_{-\infty}^{\infty} \|f_X(t)\| dt < \infty$ , 则  $X$  为连续型且

$$p_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f_X(t) dt.$$

- 推论3.8.11. 若  $f_X \equiv f_Y$ , 则  $X \stackrel{d}{=} Y$ .

### 三、特征函数的性质

- 性质1.  $f(0) = 1$ , 且  $\|f(t)\| \leq 1$ .
  - $\|f(t)\|^2 = (\mathbb{E} \cos(tX))^2 + (\mathbb{E} \sin(tX))^2$   
 $\leq \mathbb{E} \cos^2(tX) + \mathbb{E} \sin^2(tX) = 1$ .
- 性质2.  $\forall \varepsilon > 0, \exists \delta > 0$ , s.t.  $|t - s| < \delta \Rightarrow \|f(t) - f(s)\| \leq \varepsilon$ .
  - $e^{itX} - e^{isX} = e^{isX}(e^{iuX} - 1)$ .
  - $\|f(t) - f(s)\| \leq \mathbb{E} \|e^{iuX} - 1\|$ .
  - $\star \leq \min\{|uX|, 2\}$ ,

$$\mathbb{E} \star = \underbrace{\mathbb{E} \star \cdot \mathbf{I}_{\{|X| \leq M\}}}_{\text{}} + \underbrace{\mathbb{E} \star \cdot \mathbf{I}_{\{|X| > M\}}}_{\text{}}.$$

- $\exists M > 0$  s.t.  $P(|X| > M) < \varepsilon/4$ . 于是  $\underbrace{\star \star \star}_{\text{}} \leq \varepsilon/2$ .
- 取  $\delta = \varepsilon/(2M)$ , 则  $\underbrace{\star \star \star}_{\text{}} \leq \delta M \leq \varepsilon/2$ .

- 性质3.  $\forall t_1, \dots, t_n \in \mathbb{R}$  和  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ ,

$$\sum_{j,k=1}^n f(t_j - t_k) \lambda_j \overline{\lambda_k} \geq 0.$$

- $\sum_{j,k=1}^n \mathbb{E} e^{i(t_j - t_k)X} \lambda_j \overline{\lambda_k} = \mathbb{E} \left\| \sum_{j=1}^n e^{it_j X} \lambda_j \right\|^2.$
- 注:  $f_X(-t) = \mathbb{E} e^{-itX} = \mathbb{E}(\overline{e^{itX}}) = \overline{\mathbb{E} e^{itX}} = \overline{f_X(t)}.$

- 定理3.8.12 (博赫纳-辛钦定理). 若  $f: \mathbb{R} \rightarrow \mathbb{C}$  连续, 且满足性质1 & 3, 则  $f$  是(某随机变量的)特征函数.

- 性质4.  $f_{a+bX}(t) = e^{iat} f_X(bt)$ . 特别地,  $f_{-X}(t) = \overline{f_X(t)}$ .
  - $Ee^{it(a+bX)} = e^{iat} Ee^{i(bt)X} = e^{iat} f_X(bt)$ .
  - 例.  $X \stackrel{d}{=} -X$  当且仅当  $f_X(t) \in \mathbb{R}, \forall t$ .
- 性质5. 若  $X_1, \dots, X_n$  相互独立, 则  $f_{X_1+\dots+X_n} = f_{X_1} \cdots f_{X_n}$ .
  - $Ee^{it(X_1+\dots+X_n)} = Ee^{itX_1} \cdots Ee^{itX_n} = (Ee^{itX_1}) \cdots (Ee^{itX_n})$ .  $\square$
- 性质6. 设  $P(A) = p$ ;  $\mathbf{I}_A$  与  $X$  (或  $Y$ ) 独立, 则

$$f_Z(t) = pf_X(t) + (1-p)f_Y(t), \quad \text{其中 } Z = \mathbf{I}_A \cdot X + \mathbf{I}_{A^c} \cdot Y.$$

- $e^{itZ} = \mathbf{I}_A \cdot e^{itX} + \mathbf{I}_{A^c} \cdot e^{itY}$ .
- $Ee^{itZ} = (E\mathbf{I}_A)(Ee^{itX}) + (E\mathbf{I}_{A^c})(Ee^{itY})$ .

例3.8.14 求柯西分布 $C(0, 1)$ 与拉普拉斯分布的特征函数.

$$p_X(x) = \frac{1}{2}e^{-|x|}, \quad p_Y(y) = \frac{1}{\pi(1+y^2)}.$$

- $\text{Exp}(1)$  的特征函数为  $f(t) = \frac{1+it}{1+t^2}$ , 故

$$f_X(t) = \frac{1}{2}f(t) + \frac{1}{2}\overline{f(t)} = \frac{1}{1+t^2}.$$

- 由反转公式,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \frac{1}{1+t^2} dt = p_X(x) = \frac{1}{2}e^{-|x|}.$$

- 改写:  $t \rightarrow y, -x \rightarrow t$  知  $f_Y(t) = e^{-|t|}$ .

例3.8.15. 设  $Y = Y_1, \dots, Y_n$  独立同分布,  $\sim C(0, 1)$ , 则

$$Y_1 + \dots + Y_n \stackrel{d}{=} nY, \quad \text{i.e. } \frac{1}{n} S_n \stackrel{d}{=} Y.$$

- $f(t) = e^{-|t|}$ ,  $f_{S_n}(t) = f(t)^n = e^{-n|t|}$ .

- $S_n \stackrel{d}{=} nY$ :

$$f_{nY}(t) = f_Y(nt) = e^{-n|t|}.$$

- 注:  $nY = Y + \dots + Y$ , ( $n$  个  $Y$  之和).

- 柯西分布族具有可加性.

- $Z \stackrel{d}{=} \mu + \alpha Y_1 \sim C(\mu, \alpha) = C(\mu, -\alpha)$ .

- 柯西分布族:  $\{C(\mu, \alpha) : \mu \in \mathbb{R}, \alpha > 0\}$ .

特征函数:  $e^{-i\mu t} e^{-\alpha \cdot |t|}$ .

- $C(\mu_1, \alpha_1) * C(\mu_2, \alpha_2) = C(\mu_1 + \mu_2, \alpha_1 + \alpha_2)$ . ( $\alpha_1, \alpha_2 > 0$ .)

- 性质7. 设  $Y_1, Y_2, \dots$  i.i.d.;  $X$  与它们独立且取非负整数, 则

$$f_Z(t) = g_X(f_{Y_1}(t)),$$

其中  $Z = Y_1 + \dots + Y_X$ ,  $g_X(z) := \sum_{k=0}^{\infty} p_k z^k$ .

- $g_X(z)$ ,  $\|z\| \leq 1$  良定:

$$\sup_{m \geq 1} \left\| \sum_{k=n}^{n+m} p_k z^k \right\| \leq \sum_{k=n}^{\infty} p_k \rightarrow 0.$$

- $E(e^{itZ} | X = k) = E(e^{it(Y_1 + \dots + Y_k)} | X = k) = (f_{Y_1}(t))^k$ .

- 性质8. 设  $E|X|^n < \infty$ , 则  $f_X^{(m)}(0) = i^m EX^m$ ,  $m = 1, \dots, n$ .

- $\|e^{itX} - \sum_{m=0}^n \frac{(itX)^m}{m!}\| \leq \min \left\{ \frac{2|tX|^n}{n!}, \frac{|tX|^{n+1}}{(n+1)!} \right\},$

$$\begin{aligned} \|f_X(t) - \sum_{m=0}^n \frac{i^m EX^m}{m!} t^m\| &\leq E_{\star}(\mathbf{I}_{\{|X|>M\}} + \mathbf{I}_{\{|X|\leq M\}}) \\ &\leq \frac{2}{n!} \cdot |t|^n \underbrace{E|X|^n \cdot \mathbf{I}_{\{|X|>M\}}}_{\text{...}} + |t|^{n+1} \frac{M^{n+1}}{(n+1)!}. \end{aligned}$$

- 推论3.8.16. 若  $EX^2 < \infty$ , 则当  $t \rightarrow 0$  时,

$$\varphi(t) = f_X(t) - 1 - itEX + \frac{EX^2}{2!}t^2 = o(t^2).$$

- $\vec{X}$  的特征函数:  $f_{\vec{X}}(\vec{t}) := \mathbb{E} e^{i\vec{t} \cdot \vec{X}} = \mathbb{E} \prod_{k=1}^n e^{it_k X_k}$ .
- 反转公式:  $f_{\vec{X}} \rightarrow F_{\vec{X}}$ .
- $\vec{X} \stackrel{d}{=} \vec{Y}$  当且仅当  $f_{\vec{X}}(\cdot) \equiv f_{\vec{Y}}(\cdot)$ .
- 性质9.  $X_1, \dots, X_n$  相互独立当且仅当

$$f_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = f_{X_1}(t_1) \cdots f_{X_n}(t_n).$$

- 注: ★★ 可推出下式, 反之不然!

$$f_{X_1 + \dots + X_n}(t) = f_{X_1}(t) \cdots f_{X_n}(t).$$

- 性质10.  $f_{(X_1, \dots, X_m)}(t_1, \dots, t_m) = f_{\vec{X}}(t_1, \dots, t_m, 0, \dots, 0)$ .

### §3.9 正态分布与高斯分布

- 设 $\Sigma_{n \times n}$  为正定矩阵. 若

$$p_{\vec{X}}(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp \left\{ -\frac{1}{2}(\vec{x} - \vec{\mu})^T \Sigma^{-1}(\vec{x} - \vec{\mu}) \right\},$$

则称 $\vec{X}$  服从( $n$  维/元)正态分布, 记为 $\vec{X} \sim N(\vec{\mu}, \Sigma)$ .

- 称 $N(\vec{0}, \mathbf{I})$  为( $n$  维)标准正态分布.
- $\vec{X} \sim N(\vec{0}, \mathbf{I})$  当且仅当 $X_1, \dots, X_n$  独立同分布,  $\sim N(0, 1)$ .  
联合密度函数:

$$\frac{1}{\sqrt{(2\pi)^n}} e^{-\|\vec{x}\|^2/2} = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-x_i^2/2}.$$

- 约定:  $\vec{Z}$  表示服从标准正态分布的随机向量.

## 一、正态分布的非退化线性变换

- 考虑  $n$  维列向量和  $n$  阶非退化方阵,  $\Sigma$  正定.
- 结论1. 设  $\mathbf{A}_{n \times n}$  非退化, 则  $\mathbf{A}\vec{Z} + \vec{\mu} \sim N(\vec{\mu}, \Sigma)$ .

特别地, 正交变换  $\mathbf{O}\vec{Z} \sim N(\vec{0}, \mathbf{I})$ .

- 注:  $\Sigma = \mathbf{A}\mathbf{A}^T$  为正定矩阵(线性代数).
- 记  $\vec{x} = \mathbf{A}\vec{z} + \vec{\mu}$ ,

$$p_{\vec{X}}(\vec{x}) = p_{\vec{Z}}(\vec{z}) \cdot \frac{1}{|\det \mathbf{A}|} = C e^{-\|\vec{z}\|^2/2}.$$

- 其中,

$$\begin{aligned}\|\vec{z}\|^2 &= \vec{z}^T \vec{z} = (\mathbf{A}^{-1}(\vec{x} - \vec{\mu}))^T (\mathbf{A}^{-1}(\vec{x} - \vec{\mu})) \\ &= (\vec{x} - \vec{\mu})^T (\mathbf{A}^T)^{-1} \mathbf{A}^{-1} (\vec{x} - \vec{\mu}) = (\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu}).\end{aligned}$$

- 特别地, 若  $\mathbf{O}$  为正交矩阵, 则  $\mathbf{O}\mathbf{O}^T = \mathbf{I}$ .

例3.9.1. 设  $X_1, \dots, X_n$  独立同分布,  $\sim N(0, 1)$ , 则

$$\bar{X} \sim N(0, \frac{1}{n}); \quad \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1); \quad \bar{X} \text{ 与 } \star \text{ 相互独立.}$$

- 取正交矩阵  $\mathbf{O}_{n \times n}$  满足  $o_{1j} \equiv \frac{1}{\sqrt{n}}$ .  $\vec{Y} := \mathbf{O} \vec{X} \sim N(\vec{0}, \mathbf{I})$ .
- $Y_1 \sim N(0, 1)$ , 因此  $\bar{X} = \frac{1}{\sqrt{n}} Y_1 \sim N\left(0, \frac{1}{n}\right)$ .
- $\frac{1}{n} \star = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X})^2 = \frac{1}{n} \sum_{i=1}^n Y_i^2 - \frac{1}{n} Y_1^2 = \frac{1}{n} \sum_{i=2}^n Y_i^2$ .
- $\bar{X}$  称为样本均值;  $S^2 = \frac{1}{n-1} \star$  称为样本方差.

- 结论2. 设  $\vec{X} \sim N(\vec{\mu}, \Sigma)$ .

若  $\Sigma = \mathbf{A}\mathbf{A}^T$ , 则  $\mathbf{A}^{-1}(\vec{X} - \vec{\mu}) \sim N(\vec{0}, \mathbf{I})$ .

- $\vec{X} \stackrel{d}{=} \mathbf{A}\vec{Z} + \vec{\mu}$ , 故  $\star \stackrel{d}{=} \vec{Z}$ .

- 例3.9.2.  $\mathbf{A} = \sqrt{\Sigma}$ :

取正交矩阵  $\mathbf{O}$  使得  $\Sigma = \mathbf{O}\mathbf{D}\mathbf{O}^T$ , 其中  $\mathbf{D} = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ .

令  $\mathbf{A} = \mathbf{O}\sqrt{\mathbf{D}}\mathbf{O}^T$ , 其中  $\sqrt{\mathbf{D}} = \text{diag}\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\}$ .

- 结论3. 设  $\star$ , 则  $\mathbf{A}\vec{X} + \vec{b} \sim N(\mathbf{A}\vec{\mu} + \vec{b}, \mathbf{A}\Sigma\mathbf{A}^T)$ .

- $\vec{Z} := \mathbf{M}^{-1}(\vec{X} - \vec{\mu}) \sim N(\vec{0}, \mathbf{I})$ , 其中  $\mathbf{M} = \sqrt{\Sigma}$ .

- $\star = \mathbf{A}(\mathbf{M}\vec{Z} + \vec{\mu}) + \vec{b} = \mathbf{A}\mathbf{M}\vec{Z} + \mathbf{A}\vec{\mu} + \vec{b} \sim N(\star, \tilde{\Sigma})$ ,

其中  $\tilde{\Sigma} = \mathbf{A}\mathbf{M}(\mathbf{A}\mathbf{M})^T = \mathbf{A}\Sigma\mathbf{A}^T$ .

## 二、高斯分布及其线性变化

- 向量维数不固定, 矩阵可以退化甚至不是方阵.
- 结论4  $N(\vec{0}, \mathbf{I})$  的特征函数为  $e^{-\|\vec{t}\|^2/2}$ .

$$\bullet f_{\vec{Z}}(\vec{t}) = \mathbb{E} \prod_{k=1}^n e^{it_k Z_k} = \prod_{k=1}^n \mathbb{E} e^{it_k Z_k} = \prod_{k=1}^n e^{-t_k^2/2} = e^{-\|\vec{t}\|^2/2}.$$

- 结论5.  $N(\vec{\mu}, \Sigma)$  的特征函数为  $e^{i\vec{t} \cdot \vec{\mu} - \frac{1}{2}\vec{t}^T \Sigma \vec{t}}$ , (其中  $\Sigma$  正定.)

- $\vec{X} := \mathbf{A}\vec{Z} + \vec{\mu} \sim N(\vec{\mu}, \Sigma)$ , 其中  $\Sigma = \mathbf{A}\mathbf{A}^T$ ,

$$\begin{aligned} f_{\vec{X}}(\vec{t}) &= \mathbb{E} e^{i\vec{t} \cdot (\mathbf{A}\vec{Z} + \vec{\mu})} = e^{i\vec{t} \cdot \vec{\mu}} \mathbb{E} e^{i\vec{t}^T \mathbf{A} \vec{Z}} = e^{i\vec{t} \cdot \vec{\mu}} \mathbb{E} e^{i(\mathbf{A}^T \vec{t})^T \vec{Z}} \\ &= e^{i\vec{t} \cdot \vec{\mu}} e^{-\|\mathbf{A}^T \vec{t}\|^2 / 2} = e^{i\vec{t} \cdot \vec{\mu}} e^{-\frac{1}{2}\vec{t}^T \mathbf{A}\mathbf{A}^T \vec{t}}. \end{aligned}$$

- A 可以退化甚至不是方阵, 此时  $\Sigma$  为半正定(非负定)矩阵.
- 特征函数形如  $\star\star$  的分布称为( $n$  维)高斯分布, 记为  $N(\vec{\mu}, \Sigma)$ .  
其中  $\vec{\mu}$  为( $n$  维)实向量,  $\Sigma$  为( $n$  阶)半正定矩阵.
- 当  $\Sigma$  退化时, 也称  $N(\vec{\mu}, \Sigma)$  为退化的正态分布.
- 结论6.  $\mathbf{A}\vec{Z} + \vec{\mu} \sim N(\vec{\mu}, \Sigma)$ , 其中  $\Sigma = \mathbf{A}\mathbf{A}^T$ .
  - 注:  $\vec{Z}$  服从  $m$  维标准正态,  $\mathbf{A}$  是  $n \times m$  矩阵,  $\vec{\mu}$  是  $n$  维.
- 例3.9.3.  $\mathbf{A} = \sqrt{\Sigma}$ .

- 结论7. 设  $\vec{X} \sim N(\vec{\mu}, \Sigma)$ , 则  $\mathbf{A}\vec{X} + \vec{b} \sim N(\mathbf{A}\vec{\mu} + \vec{b}, \mathbf{A}\Sigma\mathbf{A}^T)$ .
  - 注:  $\vec{X}$  是  $n$  维,  $\mathbf{A}$  是  $m \times n$  矩阵,  $\vec{b}$  是  $m$  维.
  - $\vec{X} \stackrel{d}{=} \mathbf{M}\vec{Z} + \vec{\mu}$ , 其中  $\mathbf{M} = \sqrt{\Sigma}$ ,  $\vec{Z}$  是  $n$  维.
  - $\star \stackrel{d}{=} \mathbf{A}(\mathbf{M}\vec{Z} + \vec{\mu}) + \vec{b} = \mathbf{AM}\vec{Z} + \mathbf{A}\vec{\mu} + \vec{b}$ ,  $\hat{\mathbf{A}}\hat{\mathbf{A}}^T = \star$ .
  - 注: 若  $\Sigma$  退化, 没有  $\vec{Z} := \mathbf{M}^{-1}(\vec{X} - \vec{\mu})$ .
- 结论8. 设  $\vec{X} \sim N(\vec{\mu}, \Sigma)$ , 则  $\vec{\mu} = E\vec{X}$ ,  $\Sigma$  为  $\vec{X}$  的协方差矩阵.
  - $\vec{X} \stackrel{d}{=} \vec{Y} := \mathbf{A}\vec{Z} + \vec{\mu}$ , 其中  $\mathbf{A} = \sqrt{\Sigma}$ .
  - $E\vec{X} = E\vec{Y} = \vec{\mu}$ ,

$$\begin{aligned}
 \text{Cov}(X_i, X_j) &= \text{Cov}(Y_i, Y_j) = E \left( \sum_{r=1}^m a_{ir} Z_r \right) \cdot \left( \sum_{s=1}^n a_{js} Z_s \right) \\
 &= \sum_{r,s=1}^n a_{ir} a_{js} E Z_r Z_s = \sum_{r=1}^n a_{ir} a_{jr} = (\mathbf{A}\mathbf{A}^T)(i, j).
 \end{aligned}$$

### 三、高斯分布的边缘分布

- $\vec{X} = \begin{pmatrix} \vec{V} \\ \vec{W} \end{pmatrix}$ , 分别是  $n$ ,  $m$  和  $k$  维的,  $n = m + k$ .
- $\vec{\mu} = \begin{pmatrix} \vec{\varphi} \\ \vec{\psi} \end{pmatrix}$ , 分别是  $***$  维的.
- $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ , 分别是  $***$  阶的.
- $V_i = X_i$ ,  $\varphi_i = \mu_i$ ,  $i = 1, \dots, m$ .
- $W_r := X_{m+r}$ ,  $\psi_r = \mu_{m+r}$ ,  $r = 1, \dots, k$ .

- 结论9. 若  $\vec{X} \sim N(\vec{\mu}, \Sigma)$ , 则  $\vec{V} \sim N(\vec{\varphi}, \Sigma_{11})$ ,  $\vec{W} \sim N(\vec{\psi}, \Sigma_{22})$ .

- 方法一、边缘特征函数.

$$\mathbb{E} e^{i\vec{v} \cdot \vec{V}} = \mathbb{E} e^{i\vec{t} \cdot \vec{X}} = e^{i\vec{t} \cdot \vec{\mu} - \frac{1}{2} \vec{t}^T \Sigma \vec{t}}, \quad \text{其中 } \vec{t} = \begin{pmatrix} \vec{v} \\ \vec{0} \end{pmatrix}.$$

- $\vec{t} \cdot \vec{\mu} = \vec{v} \cdot \vec{\varphi}, \quad \vec{t}^T \Sigma \vec{t} = \vec{v}^T \Sigma_{11} \vec{v}.$
- 方法二、线性变换:  $\vec{V} = \begin{pmatrix} \mathbf{I}_{m \times m}, & \mathbf{0} \end{pmatrix} \vec{X}.$
- $\vec{V}$  的期望为  $\vec{\varphi}$ , 协方差矩阵为  $\Sigma_{11}$ .

- 结论10. 设  $\vec{X} \sim N(\vec{\mu}, \Sigma)$ .

$\vec{V}$  与  $\vec{W}$  相互独立当且仅当  $\Sigma_{12} = \mathbf{0}$ , (等价地,  $\Sigma_{21} = \mathbf{0}$ .)

- 联合特征函数:

$$f_{\vec{V}, \vec{W}}(\vec{v}, \vec{w}) = \mathbb{E} e^{i\vec{v} \cdot \vec{V} + i\vec{w} \cdot \vec{W}} = \mathbb{E} e^{i\vec{t} \cdot \vec{X}} = e^{i\vec{t} \cdot \vec{\mu} - \frac{1}{2} \vec{t}^T \Sigma \vec{t}}, \quad \vec{t} = \begin{pmatrix} \vec{v} \\ \vec{w} \end{pmatrix}.$$

- $\vec{t} \cdot \vec{\mu} = \vec{v} \cdot \vec{\varphi} + \vec{w} \cdot \vec{\psi},$

$$\vec{t}^T \Sigma \vec{t} = \vec{v}^T \Sigma_{11} \vec{v} + \vec{w}^T \Sigma_{22} \vec{w} + 2\vec{v}^T \Sigma_{12} \vec{w}.$$

- $f_{\vec{V}, \vec{W}}(\vec{v}, \vec{w}) = f_{\vec{V}}(\vec{v}) f_{\vec{W}}(\vec{w}) e^{-\star}.$

## 四、高斯分布的结构

- 结论11. 若  $\vec{X} \sim N(\vec{\mu}, \Sigma)$ , 则  $\exists \vec{Z}$  和  $\mathbf{A}$  使得  $\vec{X} = \mathbf{A}\vec{Z} + \vec{\mu}$  a.s..

- 注: 都是  $n$  维的. 只需处理  $\Sigma$  退化的情形. 不妨设  $\vec{\mu} = \vec{0}$ .

- $\vec{X} \stackrel{d}{=} \mathbf{M}\vec{Y}$ , 其中  $\mathbf{M} = \sqrt{\Sigma}$ ,  $\vec{Y} \sim N(\vec{0}, \mathbf{I})$ .

- 设  $\text{rank}(\Sigma) = m$ , 等价地,  $\text{rank}(\mathbf{M}) = m$ .

取  $\mathbf{M}$  的行向量的极大线性无关组, 不妨设为前  $m$  个:  $\vec{\alpha}^{(i)}$ 's.

- 对  $r = 1, \dots, \underbrace{k = n - m}_{\text{wavy line}}, \exists b_{ri}$ 's 使得

$$\vec{\alpha}^{(m+r)} = b_{r1}\vec{\alpha}^{(1)} + \dots + b_{rm}\vec{\alpha}^{(m)}.$$

- 令  $V_i := \vec{\alpha}^{(i)}\vec{Y}$ ,  $W_r := \vec{\alpha}^{(m+r)}\vec{Y}$ , 则

$$W_r = b_{r1}V_1 + \dots + b_{rm}V_m, \quad r = 1, \dots, k.$$

- 结论11(续).  $\vec{X} \sim N(\vec{\mu}, \Sigma) \Rightarrow \vec{X} = \mathbf{A}\vec{Z} + \vec{\mu}$  a.s..

- $\vec{X} = \begin{pmatrix} \vec{X}^{(1)} \\ \vec{X}^{(2)} \end{pmatrix}$  与  $\mathbf{M}\vec{Y} = \begin{pmatrix} \vec{V} \\ \vec{W} \end{pmatrix}$  同分布.

- $\vec{Y} \sim N(\vec{0}, \mathbf{I}_n)$ ,  $V_i := \vec{\alpha}^{(i)}\vec{Y}$ ,  $\vec{\alpha}^{(i)}$ 's 线性无关,  $\boxed{\vec{W} = \mathbf{B}\vec{V}}$ .

- $\vec{V}$  服从正态分布:

由 $\star$ 和 $\star$ 可推出 $V_i$ 's 线性无关, 故 $\vec{V}$  的协方差矩阵非退化.

- 另证 $\star$ :

$\vec{V} = \tilde{\mathbf{M}}_{m \times n} \vec{Y}$ , 其协方差矩阵为  $\tilde{\mathbf{M}}\mathbf{I}_{n \times n}\tilde{\mathbf{M}}^T$ , 秩为  $m$ .

- $\vec{X}^{(1)} \stackrel{d}{=} \vec{V}$ , 故服从正态分布. 于是

$$\vec{Z}^{(1)} = \mathbf{C}^{-1}\vec{X}^{(1)} \sim N(\vec{0}, \mathbf{I}_m), \quad \mathbf{C} = \sqrt{\Sigma_{11}}.$$

- 由 $\boxed{\vec{W} = \mathbf{B}\vec{V}}$  知  $\vec{X}^{(2)} = \mathbf{B}\vec{X}^{(1)}$  a.s..

- 取 $\vec{Z}^{(2)}$  得  $\vec{Z} \sim N(\vec{0}, \mathbf{I}_n)$ .

$$\vec{X} = \mathbf{A}\vec{Z}, \quad \text{其中 } \mathbf{A} = \begin{pmatrix} \mathbf{C} & \mathbf{0}_{m \times k} \\ \mathbf{BC} & \mathbf{0}_{k \times k} \end{pmatrix}.$$

## 五、高斯分布的条件分布

例3.9.5. 设 $(X, Y)$  服从正态分布,  $EX = EY = 0$ , 则

$$\mathcal{L}(Y|X = x) = N\left(\frac{EXY}{EX^2} \cdot x, EY^2 - \frac{(EXY)^2}{EX^2}\right).$$

- $Y = aX + (Y - aX)$ . 目标: 取 $a$  使得 $EWX = 0$ .
- $(X, W)$  服从高斯分布, 于是 $X$  与 $W$  相互独立,

$$\mathcal{L}(Y|X = x) = \mathcal{L}(ax+W|X = x) = \mathcal{L}(ax+W) \sim N(ax, \sigma_W^2).$$

- $E(Y - aX)X = 0$  当且仅当 $a = \frac{EXY}{EX^2}$ .
- 此时,

$$EW^2 = EY^2 - E(aX)^2 = EY^2 - \frac{(EXY)^2}{EX^2}.$$

例. 设  $\vec{X} = \begin{pmatrix} \vec{V} \\ \vec{W} \end{pmatrix} \sim N(\vec{0}, \Sigma)$ ,  $\Sigma_{11}$  非退化, 则

$\mathcal{L}(\vec{W} | \vec{V} = \vec{v}) = N(\mathbf{A}\vec{v}, \hat{\Sigma}_{22})$ , 其中  $\mathbf{A} = \Sigma_{21}\Sigma_{11}^{-1}$ ,  $\hat{\Sigma}_{22} = \Sigma_{22} - \mathbf{A}\Sigma_{12}$ .

- $\vec{W} = \mathbf{A}\vec{V} + (\vec{W} - \mathbf{A}\vec{V})$ . 目标: 取  $\mathbf{A}$  使得  $\mathbb{E}T_r V_i \stackrel{(*)}{=} 0, \forall i, r$ .
- $(\vec{V}, \vec{T})$  服从高斯分布, 于是  $\vec{V}$  与  $\vec{T}$  相互独立,

$$\mathcal{L}(\vec{W} | \vec{V} = \vec{v}) = \mathcal{L}(\mathbf{A}\vec{v} + \vec{T} | \vec{V} = \vec{v}) = \mathcal{L}(\mathbf{A}\vec{v} + \vec{T}) \sim N(\mathbf{A}\vec{v}, \Sigma_{\vec{T}}).$$

- 取  $\mathbf{A} = \Sigma_{21}\Sigma_{11}^{-1}$  以使得  $\Sigma_{21} - \mathbf{A}\Sigma_{11} = 0$ :

$$\mathbb{E}T_r V_i = \mathbb{E}(W_r - \sum_{j=1}^m a_{rj} V_j) V_i = \sigma_{ri} - \sum_{j=1}^m a_{rj} \sigma_{ij}.$$

- $\Sigma_{\vec{T}} = \hat{\Sigma}_{22}$ :  $\mathbb{E}T_s T_r = \mathbb{E}T_s (W_r - \sum_{i=1}^m a_{ri} V_i) \stackrel{(*)}{=} \mathbb{E}T_s W_r,$

$$\mathbb{E}T_s T_r = \mathbb{E}(W_s - \sum_{j=1}^m a_{sj} V_j) W_r = \sigma_{m+s, s+r} - \sum_{j=1}^m a_{sj} \sigma_{j, m+r}.$$

- 结论12 设  $\vec{X} = \begin{pmatrix} \vec{V} \\ \vec{W} \end{pmatrix} \sim N(\vec{\mu}, \Sigma)$ . 设  $\Sigma_{11}$  非退化, 则

$$\mathcal{L}(\vec{W} | \vec{V} = \vec{v}) = N(\vec{\theta}_{\vec{v}}, \hat{\Sigma}_{22}),$$

其中  $\vec{\theta}_{\vec{v}} = \vec{\psi} + \Sigma_{21}\Sigma_{11}^{-1}(\vec{v} - \vec{\varphi})$ ,  $\hat{\Sigma}_{22} = \Sigma_{22} - \mathbf{A}\Sigma_{12}$ .

- 先考虑  $\vec{W} - \vec{\psi}$  和  $\vec{V} - \vec{\varphi}$ ,

$$\mathcal{L}(\vec{W} - \vec{\psi} | \vec{V} - \vec{\varphi} = \vec{\theta}) = N(\mathbf{A}\vec{\theta}, \hat{\Sigma}_{22}).$$

- 取  $\theta = \vec{v} - \vec{\varphi}$ , 推出

$$\mathcal{L}(\vec{W} - \vec{\psi} | \vec{V} = \vec{v}) = N(\mathbf{A}\vec{\theta}, \hat{\Sigma}_{22}).$$

## 六、高斯分布的其他性质

命题3.9.6.  $\vec{X} = (X_1, \dots, X_n)^T$  服从高斯分布当且仅当对任意  $a_1, \dots, a_n \in \mathbb{R}$ ,  $\sum_{k=1}^n a_k X_k$  服从一维高斯分布.

- $\Leftarrow$ : 取  $a_k = 1$ ;  $a_j = 0, \forall j \neq k$  知  $X_k$  服从高斯分布.
- $\vec{X}$  的期望  $\vec{\mu}$  和协方差矩阵  $\Sigma$  都存在.
- $f_{\vec{X}}(\vec{t}) = \mathbb{E}e^{i\vec{t} \cdot \vec{X}}$ ,  $Y = \sum_{k=1}^n t_k X_k$  服从高斯分布.
- $\mu := \mathbb{E}Y = \sum_{k=1}^n t_k \mu_k = \vec{t} \cdot \vec{\mu}$ ;

$$\sigma^2 := \text{Var}(Y) = \sum_{k,j=1}^n t_k t_j \text{Cov}(X_k, X_j) = \sum_{k,j=1}^n t_k t_j \sigma_{kj} = \vec{t}^T \Sigma \vec{t}.$$

- $f_{\vec{X}}(\vec{t}) = f_Y(1) = e^{i\mu - \sigma^2/2} = e^{i\vec{t} \cdot \vec{\mu} - \frac{1}{2} \vec{t}^T \Sigma \vec{t}}.$