

# Probability Notes

颜成子游

2022 年 5 月 8 日

## 目录

<b>1 分布函数</b>	<b>2</b>
<b>2 随机变量的数字特征</b>	<b>2</b>
2.1 数学期望	2
2.1.1 数学期望的定义	2
2.1.2 数学期望的性质	3
2.1.3 常见分布的期望	3
2.2 其他数字特征	4
2.2.1 方差	5
2.2.2 协方差阵	6
2.2.3 相关系数	7
2.2.4 条件数学期望	8
2.2.5 矩	11
2.3 母函数	11
2.4 特征函数	13
2.4.1 特征函数的概念	13
2.4.2 反演公式与唯一性定理	15
2.4.3 多元特征函数	15
2.5 多元正态分布	16
2.5.1 密度函数与特征函数	16
2.5.2 正态分布的性质	17
2.5.3 线性变换	18
<b>3 极限定理</b>	<b>19</b>
3.1 随机变量列的收敛性	19
3.2 弱收敛	20
3.3 连续性定理：大的要来了	21

# 1 分布函数

## 2 随机变量的数字特征

### 2.1 数学期望

#### 2.1.1 数学期望的定义

引入质心坐标：

$$E_{\xi} = \int_{\Omega} \xi p(d\omega) = \int x d(F(x))$$

**定义 2.1** 设  $\xi$  是离散型随机变量，概率分布  $p_i = p(\xi = x_i), i = 1, 2, \dots$ ，如果：

$$\sum_i |x_i| p_i < +\infty$$

则  $E_{\xi} = \sum_i x_i p_i$  为  $\xi$  的数学期望。

**定义 2.2** 设  $\xi$  是连续型随机变量，密度函数为  $p(x)$ ，若：

$$\int_{-\infty}^{+\infty} |x| p(x) dx < +\infty$$

则称：

$$E_{\xi} = \int_{-\infty}^{+\infty} x p(x) dx$$

为  $\xi$  的数学期望。

显然，对于两个随机变量的和，差运算，取期望的操作是线性满足的。但是乘积并非如此：

**定理 2.1.1** 设  $\xi, \eta$  为  $(\Omega, \mathcal{F}, P)$  两个相互独立的随机变量，并且  $\xi, \eta$  均可积 ( $E_{\xi}, E_{\eta} < \infty$ )，那么  $\xi\eta$  也可积，并且：

$$E_{\xi\eta} = E_{\xi} E_{\eta}$$

**证明** (1) 假设  $\xi, \eta$  都是非负简单随机变量：

$$\xi = \sum_{i=1}^n a_i I_{A_i}, \quad \eta = \sum_{j=1}^m b_j I_{B_j}$$

由于  $\xi$  与  $\eta$  独立，可以推出  $A_i$  与  $B_j$  独立。(请自己思考为什么这两个事件是独立的。)

$$E_{\xi\eta} = \int \sum_{i=1}^n \sum_{j=1}^m a_i b_j I_{A_i} I_{B_j} p(d\omega) = \sum_i \sum_j a_i b_j p(A_i B_j) = \sum_i a_i p(A_i) \sum_j b_j p(B_j) = E_{\xi} E_{\eta}$$

(2)  $\xi, \eta$  是非负随机变量：

$$\xi_n = \sum_{k=0}^{n2^n} k/2^n I_{\{k/2^n \leq \xi < (k+1)/2^n\}}$$



$$\eta_n = \sum_{k=0}^{n2^n} k/2^n I_{\{k/2^n \leq \eta < (k+1)/2^n\}}$$

$\xi_n$  与  $\eta_n$  独立, 结合单调收敛定理可得结论。

(3)  $\xi, \eta$  是一般的随机变量:

$$E_{\xi\eta} = E[(\xi^+ - \xi^-)(\eta^+ - \eta^-)] = E\xi^+\eta^+ - E\xi^+\eta^- - E\xi^-\eta^+ + E\xi^-\eta^- = E\xi^+E\eta - E\xi^-E\eta = E\xi E\eta$$

证毕。 □

什么样的期望保持乘法时, 才会有随机变量独立呢?

**定理 2.1.2** 设  $\xi, \eta$  为  $(\Omega, \mathcal{F}, P)$  两个随机变量。那么:

$\xi, \eta$  独立  $\Leftrightarrow Ef(\xi)g(\eta) = Ef(\xi)Eg(\eta)$  对于任何 Borel 函数  $f, g$  使得  $f(\xi)g(\eta)$  可积。

**证明** 从左往右是显然的。

从右往左:

取  $f(x) = I_{B_1}(x)$ ,  $g(x) = I_{B_2}(x)$

$$Ef(\xi)g(\eta) = EI_{B_1}(\xi)I_{B_2}(\eta) = p(\xi \in B_1, \eta \in B_2)$$

$$Ef(\xi)Eg(\eta) = p(\xi \in B_1)p(\eta \in B_2)$$

从而知随机变量独立。 □

### 2.1.2 数学期望的性质

**定理 2.1.3** 1  $EC = C$

2  $EC\xi = CE\xi$

3  $E(\eta + \xi) = E(\eta) + E(\xi)$

4 若  $\xi \geq \eta$ , 则  $E\xi \geq E\eta$

5  $E(f(\xi)) = \int f(x)d(F(x))$

### 2.1.3 常见分布的期望

#### 1. Bernoulli 分布

$E\xi = \sum x_i p_i = p$  二项分布  $\xi = \xi_1 + \dots + \xi_n$  根据可加性知:  $E\xi = np$

#### 2. Poisson 分布

$$E\xi = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda$$

#### 3. 几何分布

$$E\xi = \sum_{k=1}^{\infty} k q^{k-1} p = 1/p$$

#### 4. 均匀分布

$$E\xi = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

#### 5. 正态分布

$$\xi \sim N(0, 1)$$

$$E\xi = \int_{\mathbb{R}} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

$$\eta \sim N(a, \sigma^2) \text{ 容易有 } E\eta = a$$

### 6. $\kappa^2$ 分布

$$\xi \sim N(0, 1) \quad \xi^2, n = 1$$

$$E\xi^2 = \int_{\mathbb{R}} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1$$

从而  $n = m$  时,  $\eta \sim \kappa^2(n)$ ,  $\eta = \sum_{k=1}^n \xi_k^2$   $\xi_k \sim N(0, 1)$

$$E\eta = m$$

### 7. Cauchy 分布

$$p(x) = \frac{1}{\pi(1+x^2)}$$

$$\int_{\mathbb{R}} |x|p(x)dx = \int_{\mathbb{R}} \frac{|x|}{\pi(1+x^2)} dx = +\infty$$

从而期望不存在!!!

## 2.2 其他数字特征

$$L^2 = \{\xi : E\xi^2 < +\infty\} \text{ 若 } \xi, \eta \in L^2, \text{ 则 } a\xi + b\eta \in L^2.$$

我们不加证明的指出以下的引理, 其证明是显然的。

**引理 2.2.1**  $E\xi^2 = 0 \Leftrightarrow \xi = 0, a.s.$

接下来给出的引理是方差概念的基石。

**引理 2.2.2** 设  $\xi, \eta \in L^2$  则

$$(E\xi\eta)^2 \leq E\xi^2 E\eta^2$$

等号成立的条件是存在实数  $t_0$ , 使得  $\eta = t_0\xi, a.s.$

**证明**  $\eta - t\xi \in L^2$  从而

$$u(t) = E(\eta - t\xi)^2 \geq 0$$

判别式即要证明的不等式。

要使不等式等号成立, 则上述不等式必须取等, 从而有唯一的  $t_0$ , 使得

$$\eta = t_0\xi, a.s.$$

我们也可以计算  $t_0$ :

$$t_0 = -\frac{-2E\xi\eta}{2E\xi^2} = \frac{E\xi\eta}{E\xi^2}$$



### 2.2.1 方差

**定义 2.3** 对于  $\xi \in L^2$ , 称:

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

为  $\xi$  的方差。

方差直接的会有如下性质:

**定理 2.2.3** 设  $\xi \in L^2$ , 则有:

1.  $D(\xi) \geq 0$
2.  $D\xi = E\xi^2 - (E\xi)^2$
3.  $D(c\xi) = c^2 D(\xi)$
4.  $f(C) = E(\xi - C)^2$  当且仅当  $C = E\xi$  时取到最小值。

**证明**

1. 利用引理2.2.1显然得到 1.
- 2.

$$D(\xi) = E[(\xi - E\xi)^2] = E[\xi^2 - 2\xi E\xi + (E\xi)^2] = E(\xi^2) - (E\xi)^2$$

3. 显然
- 4.

$$f(C) = E[(\xi - E\xi)^2] + 2E(\xi - E\xi)(E\xi - C) + (E\xi - C)^2 = E[(\xi - E\xi)^2] + (E\xi - C)^2 \geq D\xi$$

从而 4 也得证。 □

**例 2.1 (Poisson 分布的方差)**

$$E(\xi^2) = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} [k(k-1) + k] \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} + \lambda \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = \lambda^2 + \lambda$$

从而由上述定理的 2:

$$D\xi = \lambda$$

**例 2.2 (正态分布的方差)**

$$D\xi = \int_{\mathbb{R}} (x-a)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \sigma^2$$

具体计算过程 (换元, 利用已知积分等等) 见教材。

**例 2.3 ( $\kappa^2$  分布的方差)**

$$E(\xi^2) = \int_0^{\infty} \frac{1}{2^n \Gamma(n/2)} x^{n/2+1} e^{-x/2} dx = \frac{4}{\Gamma(n/2)} \Gamma(n/2 + 2) = n^2 + 2n$$

从而

$$D\xi = E\xi^2 - (E\xi)^2 = 2n$$



**定义 2.4 (一般随机变量的标准化)** 称  $\xi$  是非退化的, 如果其方差不为零。如果方差为零, 等价于  $\xi$  几乎处处为常数, 即单点分布。

$\xi \in L^2$  中的非退化随机变量可以标准化为均值为 0, 方差为 1 的随机变量  $\xi^*$ :

$$\xi^* = \frac{\xi - E\xi}{\sqrt{D(\xi)}}$$

很容易验证其的方差和期望满足我们“标准化”的要求。

### 2.2.2 协方差阵

协方差描述了多个随机变量之间相互依存的关系

**定义 2.5**  $\xi \in L^2, \eta \in L^2$ , 称:

$$Cov(\xi, \eta) = E[(\xi - E\xi)(\eta - E\eta)]$$

是  $\xi$  与  $\eta$  的协方差。

**命题 2.2.1** 协方差的等价公式:

$$Cov(\xi, \eta) = E[\xi\eta - \xi E\eta - \eta E\xi + E\xi E\eta] = E\xi\eta - E\xi E\eta$$

**命题 2.2.2** 方差与协方差的关系:

1.  $Cov(\xi, \xi) = E(\xi^2) - (E(\xi))^2 = D(\xi)$
2.  $D(\xi + \eta) = E[(\xi + \eta) - E(\xi + \eta)]^2 = E[(\xi - E\xi)^2] + E[(\eta - E\eta)^2] + 2E[(\xi - E\xi)(\eta - E\eta)] = D(\xi) + D(\eta) + 2Cov(\xi, \eta)$
3.  $D(\xi_1 + \xi_2 + \cdots + \xi_n) = \sum_{k=1}^n D(\xi_k) + 2 \sum_{1 \leq i < j \leq n} Cov(\xi_i, \xi_j)$

**定义 2.6 (协方差矩阵)**  $n$  维随机变量  $(\xi_1, \xi_2, \dots, \xi_n)$  的对称协方差矩阵定义为:

$$B = (b_{ij}), \quad b_{ij} = Cov(\xi_i, \xi_j)$$

**命题 2.2.3** 协方差矩阵是半正定的。

**证明** 对于任意  $n$  个实数  $a_k$ , 我们有:

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j b_{ij} = \sum_{i=1}^n \sum_{j=1}^n a_i a_j Cov(\xi_i, \xi_j) = E\left[\sum_{i=1}^n a_i (\xi_i - E\xi_i)\right]^2 \geq 0$$

**例 2.4 (离散型二维随机变量的协方差)**

$$E(g(\xi, \eta)) = \sum_{i,j} g(x_i, x_j) P\{\xi = x_i, \eta = x_j\}$$

$$Cov(\xi, \eta) = E(\xi\eta) - E(\xi)E(\eta)$$



### 2.2.3 相关系数

**定义 2.7** 若  $\xi, \eta$  是  $L^2$  中非退化 (方差不为零) 的随机变量, 称:

$$r = \frac{\text{cov}(\xi, \eta)}{\sqrt{D(\xi)}\sqrt{D(\eta)}}$$

是  $\xi$  与  $\eta$  的相关系数。若  $r = 0$ , 则  $\xi$  与  $\eta$  不相关。

由定义立刻得到:

$$r = E\left[\frac{\xi - E(\xi)}{\sqrt{D(\xi)}} \frac{\eta - E(\eta)}{\sqrt{D(\eta)}}\right]$$

即标准化后两个随机变量的协方差。

注明:

1. 常数与任何随机变量的协方差都是 0.
2. 常数与任何随机变量都是独立的。

**定理 2.2.4** 对于  $L^2$  空间中的非退化随机变量  $\xi, \eta$ , 相关系数取  $[-1, 1]$ 。且:

$$r = 1 \Leftrightarrow \xi^* = \eta^*$$

$$r = -1 \Leftrightarrow \xi^* = -\eta^*$$

**证明**

$$|r| = |E(\xi^* \eta^*)| \leq \sqrt{D(\xi^*)D(\eta^*)} = 1$$

根据 Cauchy-Schwarz 不等式取等条件, 立刻得到取等条件。 □

$r = 1$ , 称完全正线性相关;  $r = -1$ , 称完全负线性相关。

**定理 2.2.5** 对于空间  $L^2$  中的非退化随机变量  $\xi, \eta$  而言, 下列命题等价:

1.  $\xi, \eta$  等价。
2.  $\text{cov}(\xi, \eta) = 0$
3.  $E(\xi\eta) = E(\xi)E(\eta)$
4.  $D(\xi + \eta) = D(\xi) + D(\eta)$

**定理 2.2.6** 若  $\xi, \eta$  独立, 则不相关。

**证明**

$$\text{cov}(\xi, \eta) = E(\xi\eta) - E(\xi)E(\eta) = 0$$

(由独立立刻就有的) □

注意, 不相关不一定是独立的。我们将给出一个典型的例子:

**例 2.5**  $\theta$  是区间  $(0, 2\pi)$  上的均匀分布的随机变量。  $\xi = \cos \theta, \eta = \cos(\theta + a)$ , 不难验证, 当  $a = \pi/2$ ,  $\xi$  与  $\eta$  是不相关的, 但是不独立 ( $\xi^2 + \eta^2 = 1$ )

但是, 在二元正态分布下, 这就是等价的 (笑)

### 2.2.4 条件数学期望

1. 事件关于事件的条件概率:

$$P(A|B) = \frac{P(AB)}{P(B)}$$

或者写为:

$$E[I_A|B] = \frac{I_A I_B}{P(B)}$$

2. 随机变量关于事件的条件期望:

$$E[X|B] = \frac{E(XI_B)}{P(B)}$$

#### 命题 2.2.4

$$E[f(x)I_B]/P(B) = \int_{-\infty}^{+\infty} f(x)dF(X|B)$$

其中  $F(X|B) = P(X < x|B)$

**证明** 证明从简单函数开始, 然后是非负可测函数, 最后是可测函数。 □

3. 事件关于  $\sigma$ -代数的条件概率:

令  $\mathcal{G} = \sigma(B_j : 1 \leq j \leq m)$ ,  $B_j$  是  $\Omega$  上的一个有限划分

$$P(A|\mathcal{G}) = E[I_A|\mathcal{G}] = \sum_{j=1}^m \frac{E[I_A I_{B_j}]}{P(B_j)} I_{B_j}(\omega)$$

对任意的  $B_j$ , 不妨取  $B_1$

$$\int_{B_1} E[I_A|\mathcal{G}]p(d\omega) = \int_{B_1} \frac{E[I_A I_{B_1}]}{P(B_1)} I_{B_1}(\omega)p(d\omega) = E[I_A I_{B_1}] = \int_{B_1} I_A p(d\omega)$$

由测度论的知识, 得到:

$$\int_B E[I_A|\mathcal{G}]p(d\omega) = \int_B I_A p(d\omega)$$

4. 随机变量关于  $\sigma$ -代数的条件期望

$$E[X|\mathcal{G}] = \sum_{j=1}^m \frac{E[X I_{B_j}]}{P(B_j)} I_{B_j}$$

类似有:

$$\int_B E[X|\mathcal{G}]p(d\omega) = \int_B X p(d\omega)$$

5. 随机变量关于  $\sigma$ -代数的条件期望定义:

**定义 2.8**  $(\Omega, \mathcal{F}, P)$ , 令  $X$  是可积随机变量,  $\mathcal{G}$  是  $\mathcal{F}$  的子  $\sigma$ -代数,  $E[X|\mathcal{G}]$  是  $P$  等价意义下唯一的  $\mathcal{G}$  可测的, 满足:

$$\int_B E[X|\mathcal{G}]p(d\omega) = \int_B X p(d\omega)$$

的可积随机变量。



6. 随机变量关于随机变量的条件期望:

**定义 2.9**

$$\sigma(Y) = \sigma(\{\omega | Y(\omega) \in C, C \in \beta\})$$

被定义为  $Y$  生成的  $\sigma$ -代数。

**定义 2.10** 设  $(\Omega, \mathcal{F}, P)$  是概率空间,  $X, Y$  是可积随机变量,  $E[X|Y]$  是  $P$  等价意义下唯一的  $\sigma(Y)$  可测的, 满足:

$$\int_{Y \in C} E[X|Y] p(d\omega) = \int_{Y \in C} X p(d\omega)$$

的可积随机变量。

注:

$$\int_{Y \in C} E[X|Y] p(d\omega) = \int_C \phi(y) dF_Y(y)$$

7.  $X$  关于  $Y = y$  的条件期望:

**定义 2.11**  $(\Omega, \mathcal{F}, P)$  是概率空间,  $X$  是可积随机变量.  $E[X|Y = y]$  是  $\mathbf{F}_X()$  意义下唯一  $\beta$  可测的满足:

$$\int_C E[X|Y = y] dF_Y(y) = \int_{Y \in C} X p(d\omega)$$

的  $\mathbf{F}_X()$  下可积函数。

接下来给出一个公式:

**命题 2.2.5**

$$E[X] = E[E[X|\mathcal{G}]]$$

**证明** 根据定义:

$$\int_B E[X|\mathcal{G}] p(d\omega) = \int_B X p(d\omega)$$

令  $B = \Omega$ , 则有:

$$E[X] = E[E[X|\mathcal{G}]]$$

**命题 2.2.6** 如果  $X$  关于  $\mathcal{G}$  可测, 则:

$$E[XY|\mathcal{G}] = X E[Y|\mathcal{G}]$$

特别地,

$$E[X|\mathcal{G}] = X$$

**证明** 先验证  $I_A$ , 其中  $A$  是  $\mathcal{G}$  中的集合。再验证一般情形即可。 □

**定义 2.12** 称一个随机变量独立于  $\mathcal{G}$ , 若  $P(X \in C, B) = P(X \in C)P(B), C \in \beta, B \in \mathcal{G}$ .

等价于  $\sigma(X)$  独立于  $\mathcal{G}$ .

**命题 2.2.7** 如果  $X$  独立于  $\mathcal{G}$ , 则  $E[X|\mathcal{G}] = E(X)$

**证明** 先验证  $I_A$ , 其中  $A$  是独立于  $\mathcal{G}$  的集合。再验证一般情形即可。  $\square$

接着我们给出条件方差的概念:

**定义 2.13**  $D(\xi|\eta = y) = E[(\xi - E(\xi|\eta = y))^2|\eta = y]$

**例 2.6** 单位圆的均匀分布  $(\xi, \eta)$ :

$$\begin{aligned} E(\xi|\eta = y) &= 0 \\ D[\xi|\eta = y] &= \frac{1 - y^2}{3} \end{aligned}$$

**例 2.7** 二维正态分布:  $(\xi, \eta)$ :

$$\begin{aligned} E[\xi|\eta = y] &= a_1 + r \frac{\sigma_1}{\sigma_2} (y - a_2) \\ D(\xi|\eta = y) &= \sigma_1^2 (1 - r^2) \end{aligned}$$

**定理 2.2.7 (全期望公式)** 设  $\xi, \eta$  是联合分布随机变量,  $g(\xi)$  是可积随机变量, 则:

$$E[E[g(\xi)|\eta]] = E(g(\xi))$$

**证明** 根据 2.2.5 直接可得。  $\square$

**例 2.8** 见书上巴格达窃贼问题。

**例 2.9**  $\xi$  的密度函数为:

$$p_1(x) = \lambda^2 x^2 e^{-\lambda x}, x > 0$$

$\eta$  在区间  $[0, \xi]$  上均匀分布, 求  $E[\eta|\xi = x]$  和  $E[\eta]$

**证明** 显然  $E[\eta|\xi = x] = \frac{x}{2}, x > 0$ , 从而

$$E(\eta) = \int_0^\infty \frac{x}{2} p_1(x) dx = \frac{1}{\lambda}$$

**定理 2.2.8**  $\{\xi_k\}$  独立同分布,  $\xi_k$  与可积随机变量  $\xi$  同分布,  $\eta$  取自然数的随机变量,  $E(\eta)$  有限,  $\eta$  与  $\xi_k$  独立, 则:

$$E\left(\sum_{k=1}^{\eta} \xi_k\right) = E(\eta)E(\xi_k)$$

**例 2.10** 见书上例 3.2.9.

**定理 2.2.9**  $\xi, \eta \in L^2$ , 如果  $f(\eta)$  做  $\xi$  的预测, 那么使得预测误差:  $E[\xi - f(\eta)]^2$  最小的  $f$  应是  $E(\xi|\eta)$ .

**证明** 简便起见, 记  $E[\xi|\eta] = g(\eta)$

$$E[\xi - f(\eta)]^2 = E\{\xi - g(\eta) + g(\eta) - f(\eta)\}^2 = E[\xi - g(\eta)]^2 + E[g(\eta) - f(\eta)]^2 + I$$

其中, 有:

$$I = 2E[(\xi - g(\eta))(g(\eta) - f(\eta))] = 2E[E[(\xi - g(\eta))(g(\eta) - f(\eta))|\eta]]$$

注意到随机变量  $g(\eta) - f(\eta)$  是关于  $\sigma(\eta)$  可测的。从而有:

$$I = 2E[\{g(\eta) - f(\eta)\}E(\{\xi - g(\eta)\}|\eta)]$$

$$E(\xi - g(\eta)|\eta) = E(\xi|\eta) - E(E(\xi|\eta)|\eta) = 0$$

故  $I = 0$ . 证毕。  $\square$



### 2.2.5 矩

**定义 2.14** 如果  $E|\xi|^k < +\infty$ , 则称  $m_k = E\xi^k$  为  $\xi$  的  $k$  阶原点矩, 而称  $c_k = E[\xi - E\xi]^k$  为  $\xi$  的  $k$  阶中心矩。

**命题 2.2.8** 若  $\xi$  的高阶矩有限, 那么低阶矩有限。

**证明**  $|\xi|^k \leq |\xi|^{k+1} + 1$  □

下面的定理给出了两种矩的相互转化关系:

**定理 2.2.10** 若  $E|\xi|^k < +\infty$ , 有:

$$c_k = \sum_{i=0}^k \binom{k}{i} (-m_1)^{k-i} m_i$$

$$m_k = \sum_{i=0}^k \binom{k}{i} c_{k-i} m_1^i$$

**证明** 直接展开即可得到结果。 □

**例 2.11**  $\xi \sim N(0, 1)$ , 则容易验证  $E|\xi|^k < +\infty$  对于任何  $k$  都成立。从而:

$$m_{2k-1} = 0$$

$$m_{2k} = (2k-1)!!$$

## 2.3 母函数

母函数的出现弥补了部分分布函数在分析问题(分析性质, 独立和卷积计算复杂, 连续型和离散型泾渭分明)上的不足。

**定义 2.15** 对于任何实数列  $\{p_n\}$ , 如果幂级数:

$$G(s) = \sum_{n=0}^{+\infty} p_n s^n$$

的收敛半径  $s_0 > 0$ , 则称  $G(s)$  为数列  $\{p_n\}$  的母函数。特别的, 当  $\{p_n\}$  为某非负整值随机变量  $\xi$  的概率分布时, 上式至少在区间  $[-1, 1]$  绝对一致收敛, 此时有:

$$G(s) = E(s^\xi)$$

称这个  $G(s)$  为随机变量  $\xi$  或其概率分布  $\{p_n\}$  的母函数。

我们给出几个例子:

**例 2.12 (几何分布的母函数)**  $G(s) = \sum_{k=1}^{+\infty} s^k q^{k-1} p = \frac{sp}{1-qs}$   
级数在  $(-1/q, 1/q)$  收敛。

我们不加证明的指出, 母函数和非负整值随机变量  $\xi$  的概率分布有着——对应的关系。从而可以用母函数来刻画概率分布。

**定理 2.3.1** 非负整值随机变量  $\xi$  的母函数为  $G(s)$ , 如果  $E\xi$  与  $E\xi^2$  都收敛, 那么:

$$G'(1) = E(\xi)$$

$$G''(1) = E(\xi^2) - E(\xi)$$

**证明** 逐项求导即可。

级数的一致收敛性保证了逐项求导的可行性, 而两个特征收敛保证了在  $s = 1$  处也收敛。  $\square$

### 独立和的母函数

**定理 2.3.2** 设  $\xi, \eta$  相互独立的整值随机变量, 分别有概率分布  $\{a_n\}, \{b_n\}$ 。对应母函数  $A(s), B(s)$ , 则它们和  $\xi + \eta$  的母函数为:

$$C(s) = A(s)B(s)$$

**证明** 计算右边, 根据母函数的唯一性可以很容易得到答案。  $\square$

**推论 2.3.3** 如果  $\xi_1, \dots, \xi_n$  相互独立且同分布, 有母函数  $G(s)$ , 那么和  $\sum_{i=1}^n \xi_i$  的母函数为  $[G(s)]^n$ 。

这个推论应用在二项分布中, 可以很容易得到一些结论。

**例 2.13** 独立的  $\xi$  的  $\eta$  分别服从二项分布  $B(n, p), B(m, p)$ , 那么  $\xi + \eta$  服从二项分布  $B(n + m, p)$

因为计算  $\xi + \eta$  可以直接得到母函数, 从而可以根据唯一性立马得到  $\xi + \eta$  是  $B(m + n, p)$  的二项分布。

**例 2.14** 掷骰子问题。见书上 p137 例 3.3.4。

下面这个定理考虑随即多个非负整值随机变量之和的母函数。

**定理 2.3.4** 设  $\{\xi_k\}$  为相互独立的非负整值随机变量序列, 它们有共同的母函数  $G(s)$ 。如果  $\eta$  是另外一个非负整值随机变量, 与  $\xi_k$  均独立, 母函数  $F(s)$ 。那么  $\xi = \sum_{k=1}^{\eta} \xi_k$  的母函数为:

$$H(s) = F(G(s))$$

**证明** 使用全期望公式和  $H(s) = E(s^\xi)$  可以很容易得到结果:

$$H(s) = E(s^\xi) = E[E(s^\xi | \eta)] = \sum_{n=0}^{\infty} P\{\eta = n\} E(s^\xi | \eta = n) = \sum_{n=0}^{\infty} P\{\eta = n\} [G(s)]^n = F(G(s))$$

**推论 2.3.5** 可以通过上述定理证明在非负整值随机变量的情况下, 随机变量  $\xi$  的期望和方差的关系。

$$E(\xi) = E(\eta)E(\xi_1)$$

$$D(\xi) = E(\eta)D(\xi_1) + D(\eta)[E(\xi_1)]^2$$

## 2.4 特征函数

### 2.4.1 特征函数的概念

特征函数弥补了母函数只能考虑非负整值随机变量的缺点。

**定义 2.16** 设  $F(x)$  是  $\mathbb{R}$  上的一个分布函数, 称:

$$f(t) = \int_{\mathbb{R}} e^{itx} dF(x), t \in \mathbb{R}$$

是  $F(x)$  对应的特征函数。如果  $F(x)$  是  $\xi$  的分布函数, 那么  $f(t)$  也称为  $\xi$  的特征函数。此时:

$$f(t) = E(e^{it\xi})$$

#### 命题 2.4.1

1. 任意  $\xi$  的特征函数都存在, 因为  $|e^{itx}| \leq 1$ , 从而积分的模长小于 1.
2. Euler 公式:

$$f(t) = E(\cos t\xi) + iE(\sin t\xi)$$

**命题 2.4.2** 特征函数  $f(t)$  有如下性质:

1.  $|f(t)| \leq f(0) = 1$
2. 共轭对称性:  $f(-t) = \overline{f(t)}$
3.  $f(t)$  为  $t \in (-\infty, +\infty)$  上一致连续。
4. 半正定性: 任意  $n \geq 1$ , 任意  $n$  个实数  $t_1, \dots, t_n$  及  $n$  个复数  $a_1, \dots, a_n$ :

$$\sum_{j=1}^n \sum_{k=1}^n a_j \overline{a_k} f(t_j - t_k) \geq 0$$

5.  $f_{a+b\xi}(t) = e^{iab} f_{\xi}(bt)$

**证明** 1, 2 显然.

3. 相减后使用控制收敛定理显然。

4. 直接运算, 注意:

$$\sum_{j=1}^n \sum_{k=1}^n a_j e^{it_j x} \overline{a_k e^{it_k x}} = \left( \sum_{m=1}^n a_m e^{it_m x} \right)^2$$

5. 直接计算可得。 □

**命题 2.4.3** 对于  $f(t)$ , 称之为**特征函数**, 如果满足:

1.  $f(0) = 1$
2.  $f(t)$  在  $\mathbb{R}$  上连续。
3.  $f(t)$  满足半正定性。

对于任何一个特征函数  $f(t)$ , 都存在一个分布函数, 满足:

$$f(t) = \int_{\mathbb{R}} e^{itx} dF(x)$$

给出一些例子:

**例 2.15**  $\xi \sim N(0, 1) \Rightarrow f_{\xi} = e^{-t^2/2}$  则根据性质 5, 可以算出任何正态分布的特征函数。

进一步研究特征函数的性质前, 我们给出一个引理:

**引理 2.4.1** 对于  $x \in \mathbb{R}$  和自然数  $n$ , 有:

$$|e^{ix} - \sum_{k=1}^n \frac{(ix)^k}{k!}| \leq \left[ \frac{|x|^{n+1}}{(n+1)!} \right] \wedge \left[ \frac{2|x|^n}{n!} \right]$$

**证明** 对于第一个, 分部积分:

$$\int_0^x (x-s)^n e^{is} ds = \frac{x^{n+1}}{n+1} + \frac{i}{n+1} \int_0^x (x-s)^{n+1} e^{is} ds$$

递归即可。

对于第二个, 把上述积分中  $n$  换为  $n-1$ , 解出:

$$\int_0^x (x-s)^n e^{is} ds$$

代入即可。 □

这个引理给出了展开  $e^{ix}$  的尾项估计。较小的  $|x|$  用第一项, 较大的用第二项。

**定理 2.4.2** 如果随机变量  $\xi$  的各阶原点矩有限, 那么对一切满足:

$$\lim_n \frac{|t|^n E|\xi|^n}{n!} = 0$$

的  $t$ ,  $\xi$  的特征函数  $f(t)$  有展开式:

$$f(t) = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} E(\xi^k)$$

**证明** 展开显然。 □

**定理 2.4.3**  $\xi$  的  $k$  阶原点矩有限, 则其特征函数  $k$  阶可微。且有:

$$f^{(k)}(t) = E[(i\xi)^k e^{it\xi}]$$

特别的, 有  $f^{(k)}(0) = i^k E\xi^k$ 。

**证明** 显然。。。留做习题。 □

**命题 2.4.4** 独立的随机变量之和的特征函数是他们各自特征函数的积。

**证明** 运用独立的性质, 并且  $e^{it\xi}$  是 Borel 函数立刻就可以得到答案。 □

**定理 2.4.4** 设  $\{\xi_k\}$  为相互独立的随机变量序列, 它们有共同的特征函数  $f_{\xi_0}(t)$ 。如果  $\eta$  是另外一个非负整值随机变量, 与  $\xi_k$  均独立, 特征函数  $G(t)$ 。那么  $\xi = \sum_{k=1}^{\eta} \xi_k$  的特征函数为:

$$H(t) = G(f_{\xi_0}(s))$$

### 2.4.2 反演公式与唯一性定理

先给出一个引理。这是一个数分的结论：

#### 引理 2.4.5

$$\lim_{c \rightarrow +\infty} \int_0^c \frac{\sin ax}{x} dx = \frac{\pi}{2} \operatorname{sgn}|a|$$

**证明** 这是一个数分的结论。注意到换元  $y = ax$ ，此时只要考虑我们常用的迪利克雷积分：

$$\int_0^{+\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$$

即可证明结果。 □

这个公式给出了特征函数与分布函数的关系：

**定理 2.4.6 (反演公式)**  $f(t)$  是定义好的特征函数,  $F(x)$  是分布函数。那么：

$$1/2[F(b+0) - F(b)] - 1/2[F(a+0) - F(a)] = \lim_{c \rightarrow +\infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita} - e^{-itb}}{it} f(t) dt$$

**证明** 考虑式子的右边，把  $f(t)$  写为积分的形式。使用 Fubini 定理，交换积分的次序，从而得到关于  $t$  的积分。此时由控制收敛定理，交换极限和积分的次序，即可得到答案。 □

**定理 2.4.7 (唯一性定理)** 分布函数和特征函数是互相唯一决定的。

**证明** 假设  $F(x)$  有连续点集合  $C_F$ 。任一点  $x \in C_F$ ，由反演公式可得  $F(b)$ 。现在考虑分布函数左连续性，即可得到所有的  $F(x)$ 。

下面的定理也是分析性质的：

**定理 2.4.8** 如果特征函数  $f(x)$  的模可积，那么对应的分布函数  $F(x)$  为连续型，且密度函数为：

$$p(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} f(t) dt$$

**证明** 利用反演公式，取  $b_n > a, b_n \rightarrow a$ 。即可得到  $F(x)$  是连续的。

再求导就行。 □

以上的结果都可以在傅里叶分析的相关问题中得到答案。6

### 2.4.3 多元特征函数

**定义 2.17** 称：

$$f(t_1, t_2, \dots, t_n) = \int \dots \int_{\mathbb{R}^n} e^{i(t_1 x_1 + \dots + t_n x_n)} dF(x_1, \dots, x_n) = E[e^{i(t_1 \xi_1 + \dots + t_n \xi_n)}]$$

是一个多元函数的特征函数。



这是  $n$  个实变量  $t_1, \dots, t_n$  的复值函数。有性质：

$$|f(t_1, \dots, t_n)| \leq f(0, 0, \dots, 0) = 1$$

$$f(-t_1, \dots, -t_n) = \overline{f(t_1, \dots, t_n)}$$

类似于定理 2.4.3，我们也有以下结论：

**命题 2.4.5** 若混合矩  $E[\prod \xi_i^{k_i}]$  有限，那么就可以用其联合特征函数在原点的导数求矩：

$$E(\prod \xi_i^{k_i}) = i^{-\sum_{j=1}^n k_j} \frac{\partial^{k_1+\dots+k_n} f(0, 0, \dots, 0)}{\partial t_1^{k_1} \dots \partial t_n^{k_n}}$$

多元的特征函数也拥有反演公式。由于其复杂性，笔记中不再给出。读者可以在一般的概率论教材中找到。这同样说明特征函数和分布函数是一一对应的。

我们指出，如果想求出子向量  $(\xi_{s_1}, \dots, \xi_{s_k})$  的  $k$  维边缘特征函数，我们只需要把原有的特征函数中不相关的变量  $t_k$  取为 0。从其期望的结果来看，结果就是如此。

**定理 2.4.9**  $\xi_1, \dots, \xi_n$  独立的充分必要条件是他们的联合特征函数等于每个分量边缘特征函数的积。

$$f(t_1, t_2, \dots, t_n) = \prod_{k=1}^n f_k(t_k)$$

**证明** 如果独立，那么根据期望的可乘性质，即可得到：

$$E[e^{i(t_1\xi_1+t_2\xi_2)}] = E[e^{it_1\xi_1}e^{it_2\xi_2}] = E[e^{it_1\xi_1}]E[e^{it_2\xi_2}]$$

如果可分，那么独立。证明留给读者。 □

## 2.5 多元正态分布

学习了特征函数后，我们将尝试使用其来分析多元正态分布。这是强大的工具。

### 2.5.1 密度函数与特征函数

我们用列向量来记  $n$  维随机变量  $\vec{\xi}$  和其可能值  $\vec{x}$

**定理 2.5.1** 设

$$\vec{a} = \begin{pmatrix} a_1 \\ \dots \\ a_n \end{pmatrix} B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & & \dots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$

其中  $B$  是正定矩阵。那么

$$p(\vec{x}) = \frac{1}{(2\pi)^{n/2} |B|^{1/2}} \exp\{-1/2(\vec{x} - \vec{a})^T B^{-1}(\vec{x} - \vec{a})\}$$

满足  $p(\vec{x}) > 0$ ，并且积分为 1。从而其给出了一个连续型分布的密度函数，称为  $n$  维正态分布，记为  $N(\vec{a}, B)$ 。

**证明** 只需要根据  $B$  是正定矩阵，然后换元  $\vec{y} = L^{-1}(\vec{x} - \vec{a})$  即可得到答案 ( $B = LL^T$ )。 □





**定理 2.5.2**  $n$  维正态分布的特征函数为：

$$f(\vec{t}) = \exp\{i\vec{a}^T \vec{t} - \frac{1}{2} \vec{t}^T B \vec{t}\}$$

**证明** 实在是没有什么技术含量，直接计算就可以了 (?) □

我们注意到在特征函数中，并不需要  $B$  是正定的。只需要其是半正定的。事实上，根据特征函数与分布函数一一对应的性质（未证明），我们可以推断，对于非正定的  $B$ ，也会有一个特征函数，从而对应一个正态分布。

事实上，如果  $|B| = 0$ ，那么意味着正态分布发生了退化。比如一维退化为单点分布等。

### 2.5.2 正态分布的性质

由于正态分布的性质完全是由其协方差阵和期望所刻画的。从而我们可以猜想到，如果特征函数能保证期望，协方差阵有很好的（如线性）性质，那么这个，那么直接就会有正态分布的线性性质。

**定理 2.5.3** 设  $\vec{\xi}$  是多元正态分布，那么其子向量也是一个正态分布。其中，期望和协方差阵都是对应的子向量和子矩阵。

**证明** 根据之前的证明，直接让不相关变量取零，得到子向量的特征函数。其形式仍是正态分布。 □

下面的性质与  $n = 2$  时完全一致。不如说是，我们从  $n = 2$  的协方差阵和期望导出了正态分布的密度函数。当  $n$  一般化的时候，这个结论仍然成立。

#### 定理 2.5.4

$$a_k = E(\xi_k)$$

$$b_{jk} = \text{cov}(\xi_j, \xi_k)$$

**证明** 第一个根据上面的定理显然。第二个，根据矩的性质，可以计算  $E(\xi_j \xi_k) = b_{jk} + a_j a_k$ 。从而根据协方差：

$$\text{cov}(\xi_j, \xi_k) = E(\xi_j \xi_k) - E(\xi_j)E(\xi_k) = b_{jk}$$

之前我们也提到，正态分布的变量独立与不相关等价。 $n$  元情况是一样的。

**定理 2.5.5** 正态分布的各分量相互独立，等价于他们互不相关。

**证明** 如果互不相关，那么可以计算特征函数，发现其正好为各自特征函数的乘积。 □

注意，实际上这里还有，正态分布的变量相互独立，等价于它们两两独立。这是一个很好的结论。

上面的结论对分量形式也成立。即如果  $B$  是一个准对角矩阵，那么按照对角形式所产生的两个子向量，独立。

两个向量独立，当且仅当各取他们之中的两个分量都独立。



### 2.5.3 线性变换

类似于 ode 中, 我们对多元微分方程, 要使用矩阵和线性变换理论研究一样, 我们也研究正态分布的线性变换。

**定理 2.5.6**  $\vec{\eta} = C\vec{\xi}$  也是正态分布, 其中  $C$  是一个  $m \times n$  矩阵。如果  $\vec{\xi}$  是正态分布。并且:

$$\xi \sim N(\vec{a}, B)$$

$$\eta \sim N(C\vec{a}, CBC^T)$$

我们总可以通过正交变换, 把一个半正定矩阵变换为一个对角矩阵。于是这时候我们就得到了一个各分量独立的正态分布。

下面的这个定理很重要。给出了正态分布和一维正态分布之间的关系:

**定理 2.5.7**  $\vec{\xi}$  是正态分布  $N(\vec{a}, B)$ , 当且仅当对任何的  $n$  维向量  $\vec{s}$ , 有  $\vec{\eta} = \vec{s}^T \vec{\xi}$  满足  $N(\vec{s}^T \vec{a}, \vec{s}^T B \vec{s})$ 。

证明就留做习题。

注意! 如果只是  $n$  个分量都是正态分布, 是不能得出其为正态分布的。必须是分量的任一线性组合均服从一维正态分布, 才能有整个变量满足正态分布。

在上述定理中, 对于一维正态分布并不需要规定相应的形式。即下面加强的定理:

**定理 2.5.8**  $\vec{\xi}$  是正态分布  $N(\vec{a}, B)$ , 当且仅当对任何的  $n$  维向量  $\vec{s}$ , 有  $\vec{\eta} = \vec{s}^T \vec{\xi}$  也是一维正态分布

那什么时候, 分量是正态分布  $\Rightarrow$  正态分布呢?

**定理 2.5.9** 若  $\xi_1, \xi_2$  是正态分布, 那么当  $\xi_1, \xi_2$  独立的时候, 两个分量的联合分布函数也是正态分布,  $\xi_1 + \xi_2$  也是正态分布。

**证明**

$$E(e^{i(t_1\xi_1+t_2\xi_2)}) = E(e^{it_1\xi_1})E(e^{it_2\xi_2})$$

$$E(e^{it(\xi_1+\xi_2)}) = E(e^{it\xi_1})E(e^{it\xi_2})$$

**例 2.16 (分量为正态分布, 联合不一定正态, 和也不一定正态)**

我们假设随机变量  $\xi, \eta$  独立, 均满足  $N(0, 1)$ 。

令  $\zeta = |\eta|, \xi > 0; -|\eta|, \xi < 0$ 。

那么有:  $\zeta \sim N(0, 1)$ :

当  $x > 0$ ,

$$P(\zeta < x) = \frac{1}{2}P(|\eta| < x) + \frac{1}{2}P(-|\eta| < x) = \frac{1}{2}(\Phi(x) - \Phi(-x)) = \Phi(x)$$

同理也可以计算  $P(\zeta < x) = \Phi(x), x < 0$

但是我们断言,  $(\eta, \zeta)$  的联合分布不是正态分布, 和也不是。只需要看如下的等式就够了:

$$P(\eta + \zeta = 0) = \frac{1}{2}P(\eta + |\eta| = 0) + \frac{1}{2}P(\eta - |\eta| = 0) = \frac{1}{2}$$

### 3 极限定理

#### 3.1 随机变量列的收敛性

**定义 3.1** 称  $\xi_n$  依概率收敛到  $\xi$ , 如果对于任何的  $\epsilon > 0$ , 有:

$$\lim_{n \rightarrow +\infty} P(|\xi_n - \xi| \geq \epsilon) = 0$$

记为  $\xi_n \xrightarrow{P} \xi$

**命题 3.1.1** 1. 同一随机变量列概率测度收敛于不同的随机变量, 那么这两个随机变量几乎相同。

2.

$$\xi_n \xrightarrow{P} \xi, \eta_n \xrightarrow{P} \eta$$

那么有线性可加性质:

$$a\xi_n + b\eta_n \xrightarrow{P} a\xi + b\eta$$

3. 连续函数作用后, 测度收敛依然保持。

$$\xi_n \xrightarrow{P} \xi \Rightarrow f(\xi_n) \xrightarrow{P} f(\xi)$$

4. 随机变量列对应相乘后仍然概率收敛。

$$\xi_n \eta_n \xrightarrow{P} \xi \eta$$

**定义 3.2** 称  $\xi_n$  几乎必然收敛到  $\xi$ , 若只在零概率测度下  $\xi_n \rightarrow \xi$  不成立。

**命题 3.1.2** 几乎必然收敛的充分必要条件为:

$$\lim_{k \rightarrow +\infty} P\left\{\bigcup_{n=k}^{\infty} [|\xi_n - \xi| \geq \epsilon]\right\} = 0$$

**证明**

**定义 3.3** 设  $r > 0$  为常数, 若随机变量  $\xi$  与  $\xi_n$  的  $r$  阶矩都有限, 并且:

$$\lim_n E|\xi_n - \xi|^r = 0$$

则称  $\xi_n$  为  $r$  阶平均收敛到  $\xi$ :

$$\xi_n \xrightarrow{L_r} \xi$$

**定理 3.1.1 (Markov 不等式)** 设随机变量  $\xi$  的  $r$  阶矩有限, 则对与  $\epsilon > 0$ :

$$P(|\xi| \geq \epsilon) \leq \frac{E|\xi|^r}{\epsilon^r}$$



## 3.2 弱收敛

**定义 3.4**  $F_n(x)$  与  $F(x)$  均为实函数。若对于  $\forall x \in C_F$  均有：

$$\lim_n F_n(x) = F(x)$$

则称  $\{F_n\}$  弱收敛到  $F: F_n \xrightarrow{W} F$

任何函数列都弱收敛于一个处处不连续的函数。由此可见弱收敛的极限是不唯一的。

**定义 3.5** 若  $\xi_n$  与  $\xi$  对应的分布函数  $F_n(x)$  与  $F(x)$ ，有  $F_n(x)$  弱收敛到  $F(x)$ ，则称随机变量  $\{\xi_n\}$  依分布收敛到  $\xi$ ，记作  $\xi_n \xrightarrow{W} \xi$ 。

由于分布函数的不连续点是至多可列的。由左连续性，只需要在连续点的值给定，就能确定分布函数。于是极限唯一。

**定理 3.2.1**

$$\xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{W} \xi$$

**证明**  $P(\xi < y) = P(\xi < y, \xi_n < x) + P(\xi < y, \xi_n \geq x)$  □

弱收敛到底意味着什么？我们看下一个反例：

**例 3.1** 依分布收敛不一定是依概率收敛。

取  $\xi, \eta$  同分布于  $p = 1/2$  的伯努利分布。于是  $\{\xi_n\} = \{\eta\}$  依分布收敛到  $\xi$ 。但是显然不依概率收敛。

从这个例子我们看出，依分布收敛只描述了分布函数之间的关系，但不能说明随机变量之间取值的关系（即得不到联合分布）。从而依概率收敛是完全不能够保证的。

相反，依概率收敛给出了随机变量之间取值的关系，从而能够说明分布函数之间的关系。

**定理 3.2.2** 依分布收敛于一个常数的随机变量列，概率测度收敛于这个常数。

**证明**  $P(|\xi_n - C| \geq \epsilon) = 1 - F_n(c + \epsilon) + F_n((C - \epsilon) + 0) \rightarrow 1 - F(c + \epsilon) + F(c - \epsilon) = 0$  □

下面的定理给出了弱收敛的充分条件：

**定理 3.2.3** 如果在  $\mathbb{R}$  的稠子集  $D$  上有：

$$\lim F_n(x) = F(x) \quad , x \in D$$

则依分布收敛。

**证明**



### 3.3 连续性定理：大的要来了

我们先给出：

**定理 3.3.1** 分布函数列弱收敛等价于相应的特征函数列逐点收敛。

这意味着特征函数的定义与反演公式所建立的某种关系是有某种连续性的。得到这个结果的证明是分析的，可以参看傅里叶分析的教材。我们只简明扼要的给出一些结论和证明思路。

**定理 3.3.2 (Helly 第一定理)** 任何  $\mathbb{R}$  上的分布函数列都有一个弱收敛于某个有界非降且左连续函数的子序列。

不能断言得到的极限函数是一个分布函数。因为其不能保证值域是  $[0, 1]$ 。

**定理 3.3.3 (Helly 第二定理)** 分布函数列  $\{F_n\}$  弱收敛于分布函数  $F(x)$ ，则对任何有界连续函数  $\varphi$ ，有：

$$\int_{\mathbb{R}} \varphi dF_n \rightarrow \int_{\mathbb{R}} \varphi dF$$