§2 特征函数

母函数有很好的性质，但是也有很大限制。考虑母函数的期望形式，我们定义下面的特征函数。

定义 对任一随机变量，称

$$f(t) = E(\mathrm{e}^{itX}), \quad -\infty < t < \infty$$

为 X 的特征函数。

对任意的随机变量 X 以及实数 t，都有

$$|f(t)| = |E(\mathrm{e}^{itX})| \leqslant E|\mathrm{e}^{itX}| = 1.$$

故任意随机变量的特征函数 $f(t)$ 都存在，是实变复值的。

特征函数有以下性质：

(1) $f(0) = 1$，$|f(t)| \leqslant 1$，$f(-t) = \overline{f(t)}$.

(2) $f(t)$ 一致连续。

(3) 若 X 与 Y 相互独立，则 $f_{X+Y}(t) = f_X(t) \cdot f_Y(t)$。

(4) 若 X 有 k 阶矩，则 $f_X(t)$ 是 k 阶可微的，且 $f_X^{(k)}(t) = E(i^k X^k e^{itX})$。

(5) 如果 X 的特征函数为 $f_X(t)$，则 $a + bX$ 的特征函数为

$$f_{a+bX}(t) = e^{itb} f_X(at).$$

(6) $f(t)$ 是非负定的，即对任意的正整数 n 及任意的实数 $t_1, \cdots, t_n$ 与复数 $\lambda_1, \cdots, \lambda_n$，总有

$$\sum_{k=1}^{n} \sum_{j=1}^{n} f(t_k - t_j) \lambda_k \overline{\lambda}_j \geqslant 0.$$

下面看几个常用函数的特征函数。

(1) 两点分布

根据定义，两点分布的特征函数为

$$f(t) = p\mathrm{e}^{it} + q.$$

于是，可以推出二项分布 $B(n,p)$ 的特征函数为

$$f(t)=(p\mathrm{e}^{it}+q)^n.$$

(2) $Poisson$ 分布

$$f(t)=\sum_{k=0}^{\infty}\mathrm{e}^{itk}\cdot\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=\sum_{k=0}^{\infty}\frac{(\lambda\mathrm{e}^{it})^k}{k!}\mathrm{e}^{-\lambda}=\mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}.$$

(3) 正态分布

先求标准正态分布 $N(0,1)$ 的特征函数。

$$\begin{aligned} f(t) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{itx}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-\frac{(x-it)^2}{x}}\mathrm{e}^{-\frac{t^2}{2}}\mathrm{d}x \\ &= \mathrm{e}^{-\frac{t^2}{2}}. \end{aligned}$$

于是，对一般的正态分布，$X\sim N(\mu,\sigma)$，由上面的性质 6 可知 X 的特征函数为

$$f(t)=\mathrm{e}^{i\mu t-\frac{1}{2}\sigma^2\mu^2}.$$

特征函数之所以称为特征函数，是因为从分布函数 $F(x)$ 到相应的特征函数 $f(t)$ 的变换是一一映射。也就是说，两个不同的分布函数对应的特征函数也是不同的。于是，给定了 $f(t)$ 之后，分布函数 $F(x)$ 也唯一确定了。从 $f(t)$ 到 $F(x)$ 的公式如下：

$$F(b)-F(a)=\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{\mathrm{e}^{-ita}-\mathrm{e}^{-itb}}{it}f(t)\mathrm{d}t,$$

其中 $a<b$ 都是 $F(x)$ 的连续点。

§3 几种收敛

对于一族分布或一族随机变量，容易想到它们是否有极限，有什么意义下的极限。如果有，极限是否还是一个分布函数。接下来就要讨论这个问题。

定义 对于分布函数列 $\{F_n\}$，如果存在一个函数 F，使得

$$\lim_{n\to\infty} F_n(x) = F(x)$$

对 F 的每一个连续点都成立，则称 $F_n(x)$ 依分布收敛（弱收敛）于 $F(x)$，记作 $F_n \xrightarrow{w} F$。

关于弱收敛，有下面的结论：

连续性定理：

设分布函数列 $F_n(x)$ 依分布收敛于某一分布函数 $F(x)$，则相应的特征函数列 $f_n(t)$ 收敛于 $f(t)$，且在 t 的一个有限区间内是一致收敛的。

设特征函数列 $f_n(t)$ 收敛于某一函数 $f(t)$，且 $f(t)$ 在 $t=0$ 处连续，则相应的分布函数列 $F_n(x)$ 依分布收敛于某一分布函数 $F(x)$，且 $f(t)$ 是 $F(x)$ 的特征函数。

对于随机变量，有以下几种收敛：

(1) 依分布收敛

称随机变量列 X_n 依分布收敛到 X，如果相应的分布函数 $F_n(x)$ 依分布收敛到 $F(x)$。记作 $X_n \xrightarrow{L} X$。

(2) 依概率收敛

称随机变量列 X_n 依概率收敛到 X，如果对任意的 $\varepsilon > 0$，都有

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0.$$

记作 $X_n \xrightarrow{P} X$。

(3) 几乎处处收敛（几乎必然收敛，概率 1 收敛）

称随机变量列 X_n 几乎处处收敛到 X，如果

$$P(\lim_{n\to\infty} X_n = X) = 1.$$

记作 $X_n \xrightarrow{a.s.} X$。

(4) r 阶矩收敛

称随机变量列 X_n 收敛到 X，如果

$$\lim_{n\to\infty} ||X_n - X||_{L^r} = 0.$$

记作 $X_n \xrightarrow{L^r} X$。

这几种收敛的强弱关系如下：

$$X_n \xrightarrow{a.s.} X \Longrightarrow X_n \xrightarrow{P} X \Longrightarrow X_n \xrightarrow{L} X,$$

$$X_n \xrightarrow{L^r} X \Longrightarrow X_n \xrightarrow{P} .X$$

依分布收敛不能推出依概率收敛，但有下面的特殊情形；
对任意的常数 c，有

$$X_n \xrightarrow{L} c \Longrightarrow X_n \xrightarrow{P} c.$$