



清华大学
Tsinghua University



粒子物理与核物理实验中的 数据分析

第十章：特征函数

王喆
清华大学



回顾

1. 最小二乘法与最大似然法的联系

对于高斯变量 y_i , 二者相同

2. 线性的最小二乘法估计

通过求矩阵的逆完成估计, 估计量是测量量 y_i 的线性函数

3. 非线性的最小二乘法估计

通过迭代完成估计, 方差可采用线性情况 $\chi^2 = \chi_{min}^2 + 1$ 来估计

4. 约束条件下的最小二乘法拟合

在约束条件下引入拉格朗日乘子改进实验观测量的精度

5. 用最小二乘法检验拟合优度

用 χ_{min}^2 作拟合优度统计, 满足 $N - m$ 自由度下的卡方分布

6. 用最小二乘法处理分区数据

把 y_i 当作泊松变量, 不确定用 λ_i 估计, 或用 y_i 估计 (推广最小二乘法)

7. 不等精度关联实验结果的并合问题

对存在相关性的数据的处理, 不确定度的修正

本章要点

- 特征函数的定义
- 特征函数的性质
- 常用概率密度函数的特征函数
- 特征函数的应用
- 中心极限定理
- 利用特征函数求估计量的p.d.f.

特征函数的定义

设随机变量 x 的概率密度函数为 $f(x)$ ，则特征函数 $\phi_x(k)$ 定义为 e^{ikx} 的期望值，即

$$\phi_x(k) = E[e^{ikx}] = \int_{-\infty}^{+\infty} e^{ikx} f(x) dx$$

特征函数的本质是什么？

特征函数的本质是概率密度函数 $f(x)$ 的傅里叶变换。

任意概率密度函数都存在特征函数。

特征函数与概率密度函数的关系

已知概率密度 $f(x)$ ，我们往往关心其特征值(如均值、方差)。特征值提供了概率密度最重要的信息，但不能完全确定概率密度的所有性质。

特征函数则与概率密度函数——对应。
概率密度由特征函数的反傅里叶变换唯一确定

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} \phi_x(k) dk$$

即，概率密度函数 $f(x)$ 与其特征函数 $\phi_x(k)$ 是等价的。

为什么引入特征函数

问题：既然概率密度函数与特征函数一一对应，给出任意一个都可以完全确定概率密度函数的所有性质，为什么还需要引入特征函数？

很多问题直接用概率密度函数不易处理，但用特征函数处理则非常方便。比如，

- 1) 求独立随机变量之和的分布的卷积变为乘法运算；
 - 2) 求 n 阶原点矩变为求 n 阶微分
-

特征函数的性质

(1) $\phi_x(0) = 1$

(2) $|\phi_x(k)| \leq 1$

(3) $\phi_x(-k) = \phi_x^*(k)$

(4) 若 $y = ax + b$, 其中 a, b 为常数, 则

Eq. 1.32

$$\phi_y(t) = e^{ibt} \phi_x(ak)$$

(5) 独立随机变量之和的特征函数为特征函数的积, 即, 设 x, y 相互独立, 则

$$\phi_{x+y}(t) = \phi_x(k)\phi_y(k)$$

Eq. 1.36

(6) 若 $E[x^l]$ 存在, 则 $\phi_x(k)$ l 次可导, 且对 $1 \leq m \leq l$ 有

$$\phi_x^{(m)}(0) = i^m E[x^m]$$

特征函数的性质：证明(5)

(5) 独立随机变量之和的特征函数为特征函数的积，即，设 x, y 相互独立，则

$$\phi_{x+y}(t) = \phi_x(k)\phi_y(k)$$

证明：由于 x, y 相互独立，所以 e^{ikx} 和 e^{iky} 相互独立，从而

$$\phi_{x+y}(t) = E[e^{ik(x+y)}] = E[e^{ikx}e^{iky}] = E[e^{ikx}]E[e^{iky}] = \phi_x(k)\phi_y(k)$$

可推广到 n 个独立随机变量之和

$$z = x_1 + x_2 + \cdots + x_n$$
$$\phi_z(k) = \phi_{x_1}(k)\phi_{x_2}(k) \cdots \phi_{x_n}(k)$$

利用反傅立叶变换可求出 z 的概率密度函数

$$f(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikz} \phi_z(k) dk$$

特征函数的性质：证明(6)

(6) 若 $E[x^l]$ 存在, 则 $\phi_x(k)$ l 次可导, 且对 $1 \leq m \leq l$ 有

$$\phi_x^{(m)}(0) = i^m E[x^m]$$

证明: $E[x^l]$ 存在 $\rightarrow \int_{-\infty}^{+\infty} |x|^l f(x) dx < +\infty$, 于是含参变量 k 的广义积分 $\int_{-\infty}^{+\infty} e^{ikx} f(x) dx$ 可以对 k 求 l 次导。

所以, 对 $0 \leq m \leq l$, 有

$$\phi_x^{(m)}(k) = \frac{d^m}{dk^m} \int_{-\infty}^{+\infty} e^{ikx} f(x) dx = \int_{-\infty}^{+\infty} (ix)^m e^{ikx} f(x) dx = i^m E[x^m e^{ikx}]$$

令 $k = 0$, 即可以得到

$$\phi_x^{(m)}(0) = i^m E[x^m]$$

利用特征函数可以方便地求出各阶原点矩。

常用概率分布的特征函数

分布	p.d.f.	$\phi(k)$
二项分布	$f(n; N, p) = \frac{N!}{n! (N - n)!} p^n (1 - p)^{N - n}$	$[p(e^{ik} - 1) + 1]^N$
泊松分布	$f(n; \nu) = \frac{\nu^n}{n!} e^{-\nu}$	$e^{\nu(e^{ik} - 1)}$
均匀分布	$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta - \alpha}, & \alpha \leq x \leq \beta \\ 0, & \text{其他} \end{cases}$	$\frac{e^{i\beta k} - e^{i\alpha k}}{(\beta - \alpha)ik}$
指数分布	$f(x; \xi) = \frac{1}{\xi} e^{-\frac{x}{\xi}}$	$\frac{1}{1 - ik\xi}$
高斯分布	$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$e^{i\mu k - \frac{1}{2}\sigma^2 k^2}$
卡方分布	$f(z; n) = \frac{1}{2^{n/2}\Gamma(n/2)} z^{n/2-1} e^{-z/2}$	$(1 - 2ik)^{-n/2}$
柯西分布	$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2}$	$e^{- k }$

柯西分布在 $k=0$ 处不可导，即各阶矩都不存在。

特征函数的应用

问题：既然概率密度函数与特征函数一一对应，给出任意一个都可以完全确定概率密度函数的所有性质，为什么还需要引入特征函数？

很多问题直接用概率密度函数不易处理，但用特征函数处理则非常方便。比如，

- 1) 求独立随机变量之和的分布的卷积变为乘法运算；
- 2) 求 n 阶原点矩变为求 n 阶微分

.....

特征函数的应用：均值与方差

求均值和方差（以高斯分布为例）

特征函数为

$$E[x] = \frac{1}{i} \frac{d}{dk} e^{i\mu k - \frac{1}{2}\sigma^2 k^2} \Big|_{k=0} = \frac{1}{i} (i\mu - \sigma^2 k) e^{i\mu k - \frac{1}{2}\sigma^2 k^2} \Big|_{k=0} = \mu$$

$$\begin{aligned} V[x] &= E[x^2] - (E[x])^2 = \frac{1}{i^2} \frac{d^2}{dk^2} e^{i\mu k - \frac{1}{2}\sigma^2 k^2} \Big|_{k=0} - \mu^2 \\ &= \frac{1}{i^2} \left[-\sigma^2 + (i\mu - \sigma^2 k)^2 \right] e^{i\mu k - \frac{1}{2}\sigma^2 k^2} \Big|_{k=0} - \mu^2 \\ &= \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \end{aligned}$$

类似地，可以很容易求各阶中心矩。

特征函数的应用：p.d.f

求p.d.f的极限行为（以二项分布为例）

特征函数为 $\phi(k) = [p(e^{ik} - 1) + 1]^N$

取极限 $p \rightarrow 0, N \rightarrow \infty, \nu = pN$ 为常数

$$\phi(k) = [p(e^{ik} - 1) + 1]^N = \left[\frac{\nu}{N} (e^{ik} - 1) + 1 \right]^N$$
$$\xrightarrow{N \rightarrow \infty} e^{\nu(e^{ik} - 1)}$$

即二项分布在 N 很大且均值保持不变时，趋向于泊松分布。
同样可以证明 ν 很大时，泊松分布趋向于高斯分布。

特征函数的应用： p.d.f.

求独立随机变量之和的p.d.f.

两个独立的高斯随机变量 x 和 y ，均值为 μ_x 和 μ_y ，方差为 σ_x^2 和 σ_y^2 ，求 $z = x + y$ 的特征函数。

$$\begin{aligned}\phi(k) &= \phi_x(k)\phi_y(k) = e^{i\mu_x k - \frac{1}{2}\sigma_x^2 k^2} \cdot e^{i\mu_y k - \frac{1}{2}\sigma_y^2 k^2} \\ &= e^{i(\mu_x + \mu_y)k - \frac{1}{2}(\sigma_x^2 + \sigma_y^2)k^2}\end{aligned}$$

这正是均值 $\mu_z = \mu_x + \mu_y$ ，方差 $\sigma_z^2 = \sigma_x^2 + \sigma_y^2$ 的高斯分布的特征函数。

同样可证，泊松变量之和仍服从泊松分布。

特征函数的应用：高斯分布与卡方分布

n 个独立的高斯变量 $x_i \sim N(\mu_i, \sigma_i^2)$, 定义 $z = \sum_{i=1}^n \frac{(x_i - \mu_i)^2}{\sigma_i^2}$ 。

证明: $z \sim \chi^2(n)$ 。

证明: 首先易证 $y_i \equiv \frac{x_i - \mu_i}{\sigma_i} \sim N(0,1)$, 并且 $z = y_i^2$ 的概率密度函数为

$$g(z) = 2\varphi(y_i) \left| \frac{dy}{dz} \right| = \frac{1}{\sqrt{2\pi z}} e^{-z/2} = f(z; n=1)$$

则 $z \sim \chi^2(1)$, 其特征函数为 $\phi_z(k) = (1 - 2ik)^{-1/2}$ 。

若 $z = \sum_{i=1}^n y_i^2$, 显然特征函数为 $\phi_z(k) = (1 - 2ik)^{-n/2}$ 。

这正是自由度为 n 的卡方分布的特征函数, 即 $z \sim \chi^2(n)$ 。

中心极限定理

定理：假设 n 个独立随机变量 x_j ，均值和方差 μ_j, σ_j^2 ，则大 n 极限下， $z = \sum_j x_j$ 为高斯随机变量， $z \sim N(\sum_j \mu_j, \sum_j \sigma_j^2)$ 。

证明：定义 $y_j = \frac{x_j - \mu_j}{\sqrt{n}}$ ，则 $E[y_j] = 0$ ， $E[y_j^2] = \frac{\sigma_j^2}{n}$ 。对 y_j 的特征函数 $\phi_j(k)$ 作泰勒展开：

$$\phi_j(k) = \sum_{m=0}^{\infty} \phi_j^{(m)}(0) \frac{k^m}{m!} = \sum_{m=0}^{\infty} \frac{(ik)^m}{m!} E[y_j^m] = 1 - \frac{k^2}{2n} \sigma_j^2 - \frac{ik^3}{3!} \frac{E[(x_j - \mu_j)^3]}{n^{3/2}} + \dots$$

则在大 n 极限下，忽略高阶项， $\phi_j(k) \simeq \exp[-k^2 \sigma_j^2 / (2n)]$ 。

定义 $\sigma^2 = \sum_{j=1}^n \sigma_j^2$ ， $z' = \sum_{j=1}^n y_j$ ，则 z' 的特征函数为

$$\phi_{z'}(k) = \prod_{j=1}^n \phi_j(k) \simeq \exp[-k^2 \sigma^2 / (2n)]$$

即 $z' \sim N(0, \sigma^2/n)$ 。变换回 $z = \sum_j x_j$ ，则 $z \sim N(\sum_j \mu_j, \sum_j \sigma_j^2)$

中心极限定理（续）

n 有限时，中心极限定理成立的条件与程度：

大致说来，只要 z 的求和中，每个 x_j 的贡献都很小即可。

即 z 由大量微小贡献组合而成。

例如，早期有些程序经常用12个(0,1]均匀分布的随机变量之和近似高斯分布。

$$z = \frac{\sum_{i=1}^n x_i - \frac{n}{2}}{\sqrt{n/12}} \xrightarrow{n=12} \sum_{i=1}^{12} x_i - 6$$

如果某个或某几个 x_j 的贡献非常大，则求和的结果将明显偏离高斯分布。

求估计量的p.d.f.(1)

以指数分布为例: $f(x; \xi) = \frac{1}{\xi} e^{-\frac{x}{\xi}}$

参数 ξ 的最大似然估计量为 $\hat{\xi} = \frac{1}{n} \sum_{i=1}^n x_i$

其分布可以用特征函数法求得。

由于 $\phi_x(k) = 1/(1 - ik\xi)$, 所以 $z = \sum_{i=1}^n x_i = n\hat{\xi}$ 的特征函数为

$$\phi_z(k) = 1/(1 - ik\xi)^n$$

反傅立叶变换得到 z 的 p.d.f.:

$$g_z(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ikz}}{(1 - ik\xi)^n} dk = \frac{1}{(n-1)!} \frac{z^{n-1}}{\xi^n} e^{-z/\xi}$$

$$g(\hat{\xi}; n, \xi) = g_z(z) |dz/d\hat{\xi}| = \frac{(n/\xi)^n}{(n-1)!} \hat{\xi}^{n-1} e^{-(n/\xi)\hat{\xi}} = \Gamma(\hat{\xi}; n, n/\xi)$$

这是伽马分布, n 很大时趋于高斯分布。

求估计量的p.d.f.(2)

寿命 $\hat{\tau} = \frac{1}{n} \sum_{i=1}^n t_i$ 的平均值:

$$E[\hat{\tau}] = E\left[\frac{1}{n} \sum_{i=1}^n t_i\right] = \frac{1}{n} \sum_{i=1}^n E[t_i] = \frac{1}{n} \sum_{i=1}^n \tau = \tau$$

也可以用刚才的 p.d.f 进行积分求均值:

$$E[\hat{\tau}] = \int_0^{\infty} \hat{\tau} g(\hat{\tau}; n, \tau) d\hat{\tau} = \int_0^{\infty} \hat{\tau} \frac{(n/\tau)^n}{(n-1)!} \hat{\tau}^{n-1} e^{-(n/\tau)\hat{\tau}} d\hat{\tau} = \tau$$

有意思的是 $\lambda = 1/\tau$ 的最大似然估计量 $\hat{\lambda} = 1/\hat{\tau}$

如何求其期待值?

$$E[\hat{\lambda}] = E\left[\frac{n}{\sum_{i=1}^n t_i}\right] = ?$$

求估计量的p.d.f.(3)

可以先求 $\hat{\lambda}$ 的分布函数:

$$h(\hat{\lambda}; n, \lambda) = g(\hat{\tau}; n, \tau) \left| \frac{d\hat{\tau}}{d\hat{\lambda}} \right| = \frac{(n\lambda)^n}{(n-1)!} \frac{1}{\hat{\lambda}^{n+1}} e^{-(n\lambda)/\hat{\lambda}}$$

用这个 p.d.f 求解期待值:

$$E[\hat{\lambda}] = \int_0^{\infty} \hat{\lambda} h(\hat{\lambda}; n, \lambda) d\hat{\lambda} = \frac{n}{n-1} \lambda$$

可以看出 $\hat{\lambda}$ 不是无偏估计量。

求估计量期待值的置信区间

利用特征函数方法求出估计量(如 $\hat{\xi}$)的p.d.f., 有了估计量的p.d.f.(如 $g(\hat{\xi}; n, \xi)$), 可以方便地处理很多问题, 如置信区间。

对于给定 α, β 以及观测值 $\hat{\xi}_{obs}$, 通过

$$\begin{aligned}\alpha &= \int_{\hat{\xi}_{obs}}^{\infty} g(\hat{\xi}; a) d\hat{\xi} \\ \beta &= \int_{-\infty}^{\hat{\xi}_{obs}} g(\hat{\xi}; b) d\hat{\xi}\end{aligned}$$

求得置信区间 $[a, b]$ 。

小结

1. 特征函数的定义

$$\phi_x(k) = E[e^{ikx}] = \int_{-\infty}^{+\infty} e^{ikx} f(x) dx \rightarrow \text{傅立叶变换}$$

2. 特征函数的性质

$\phi_{x+y}(k) = \phi_x(k)\phi_y(k)$: 随机变量 x 和 y 相互独立

$\phi_x^{(m)}(0) = i^m E[x^m]$: 微分与原点矩的关系

3. 常用概率密度函数的特征函数

4. 特征函数的应用

5. 中心极限定理

6. 利用特征函数寻找估算子的p.d.f.