CH02 抽样分布及若干预备知识

更新：2026 年 1 月 9 日

目录

1 特征函数和矩母函数

1.1 特征函数

1.1.1 特征函数的定义

定义 1.1 (特征函数) 设 X 是一个随机变量，称

$$\varphi(t)=\mathbb{E}(e^{itX}),\quad t\in\mathbb{R}, \tag{1}$$

为 X 的**特征函数**。其中 i 为虚数单位，满足 $i^2=-1$。

当离散型随机变量 X 的分布列为 $p_k=P(X=x_k),k=1,2,\cdots$，则 X 的特征函数为

$$\varphi(t)=\sum_{k=1}^{\infty}e^{itx_k}p_k,\quad -\infty<t<\infty. \tag{2}$$

当连续随机变量 X 的密度函数为 $p(x)$，则 X 的特征函数为

$$\varphi(t)=\int_{-\infty}^{\infty}e^{itx}p(x)\,dx,\quad -\infty<t<\infty. \tag{3}$$

定义 1.2 特别地，对于随机向量 $X\in\mathbb{R}^p$，其中 p 为随机向量的分量数，或 X 的维度。则此时，特征函数定义为

$$\varphi(t)=\mathbb{E}(e^{it^TX}),\quad t\in\mathbb{R}^p \tag{4}$$

1.1.2 特征函数的性质

现在研究特征函数函数的性质，其中 $\varphi_X(t)$ 表示 X 的特征函数，其他类似。

命题 1.1 若 $Y=aX+b$，其中 a,b 为常数，则

$$\varphi_Y(t)=e^{ibt}\cdot\varphi_X(at) \tag{5}$$

证明 $\varphi_Y(t)=\mathbb{E}(e^{it(aX+b)})=e^{ibt}\cdot\mathbb{E}(e^{i(at)X})=e^{ibt}\cdot\varphi_X(at)$.

命题 1.2 独立随机变量和的特征函数为每个随机变量的特征函数的积，即设 X 与 Y 相互独立，则

$$\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t) \tag{6}$$

证明 因为 X 与 Y 相互独立，所以 e^{itX} 和 e^{itY} 也相互独立，从而有

$$\varphi_{X+Y}(t) = \mathbb{E}(e^{it(X+Y)}) = \mathbb{E}(e^{itX} \cdot e^{itY}) = \mathbb{E}(e^{itX}) \cdot \mathbb{E}(e^{itY}) = \varphi_X(t) \cdot \varphi_Y(t)$$

命题 1.3 特征函数唯一确定密度函数：若 X 为连续随机变量，其密度函数为 $p(x)$，特征函数为 $\varphi(t)$。若 $\int_{-\infty}^{\infty} |\varphi(t)|\, dt < \infty$，则

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t)\, dt \tag{7}$$

证明 详见《概率论与数理统计教程》（茆诗松等）定理 4.2.4（唯一性定理）。

1.1.3 正态分布的特征函数

定义 1.3 (正态分布) 若随机向量 X 满足如下特征函数

$$\varphi_X(t) = \mathbb{E}\left\{\exp\left(it^T X\right)\right\} = \exp\left(i\mu^T t - \frac{1}{2} \cdot t^T \Sigma t\right) \tag{8}$$

其中 $X, t, \mu \in \mathbb{R}^p$ 为 p 维向量，$\Sigma \in \mathbb{R}^{p\times p}$ 为 $p \times p$ 的矩阵。则称 X 服从 p **维正态分布**，期望为 μ，协方差矩阵为 Σ，记作 $X \sim N_p(\mu, \Sigma)$。

1.2 矩母函数

1.2.1 矩母函数的定义

定义 1.4 设随机变量 X 的分布为 $F(x)$，其**矩母函数**定义为

$$M_X(t) = \mathbb{E}(e^{tX}), \quad t \in \mathbb{R},$$

类似地，当 X 为随机向量时 $X \in \mathbb{R}^p$，矩母函数的 t 也是向量 $t \in \mathbb{R}^p$。

注 对于 r.v. X 的矩母函数，有

- $M_X(0) = 1$
- 不是所有随机变量都存在矩母函数。

1.2.2 矩母函数的性质

命题 1.4 对于随机变量 X，它的特征函数为 $\varphi_X(t)=\mathbb{E}(e^{itX})$，矩母函数为 $M_X(t)=\mathbb{E}(e^{tX})$，则有

$$M_X(t)=\varphi_X(-it)$$

若矩母函数存在时，上面的关系总成立。

命题 1.5 对于随机变量 X，它的矩母函数为 $M_X(t)$，若 $M_X(t)$ 在 $t=0$ 可导，则有

$$M_X^{(n)}(0)=\mathbb{E}(X^n)$$

即，矩母函数的第 n 阶导数在零点给出随机变量的 n 阶原点矩。

2 正态总体样本均值和样本方差分布

2.1 正态变量线性函数的分布

定理 2.1 设随机变量 $X_1,X_2,\cdots,X_n$ 相互独立且 $X_k\sim N(a_k,\sigma_k^2),k=1,2,\cdots,n$. 令 $c_1,c_2,\cdots,c_n$ 为常数，则有

$$T=\sum_{k=1}^{n}c_kX_k\sim N(\mu,\tau^2) \tag{9}$$

其中 $\mu=\sum_{k=1}^{n}c_ka_k$ 而 $\tau^2=\sum_{k=1}^{n}c_k^2\sigma_k^2$。

证明 由 $X_k\sim N(a_k,\sigma_k^2),k=1,2,\cdots,n$ 故其特征函数为

$$\varphi_k(t)=\mathbb{E}(e^{itX_k})=\exp\left(ia_kt-\frac{1}{2}t^2\sigma_k^2\right)$$

所以 T 的特征函数为

$$\begin{aligned}\varphi_T(t)=\mathbb{E}(e^{itT})=\mathbb{E}\left[\exp\left(it\sum_{k=1}^{n}c_kX_k\right)\right]&=\prod_{k=1}^{n}\mathbb{E}\left[\exp\left(i(c_kt)X_k\right)\right]\\&=\prod_{k=1}^{n}\exp\left(ic_ka_kt-\frac{1}{2}t^2c_k^2\sigma_k^2\right)=\exp\left[it\left(\sum_{k=1}^{n}c_ka_k\right)-\frac{1}{2}t^2\left(\sum_{k=1}^{n}c_k^2\sigma_k^2\right)\right]\\&=\exp\left(it\mu-\frac{1}{2}t^2\tau^2\right)\end{aligned}$$

可见 $T\sim N(\mu,\tau^2)$，定理证毕。

推论 1 在定理 2.1 中，若取 $a_k=a,\sigma_k^2=\sigma^2$，则有

$$T\sim N\left(a\sum_{k=1}^{n}c_k,\ \sigma^2\sum_{k=1}^{n}c_k^2\right)$$

推论 2 在推论 1 中，继续取 $c_k = 1/n$，即 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(a, \sigma^2)$，而 $T = \sum_{k=1}^{n} X_n/n = \bar{X}$，有

$$\bar{X} \sim N\left(a,\, \frac{\sigma^2}{n}\right)$$

下面，针对随机向量 X，我们推导更一般的定理。

定理 2.2 随机向量 $X \sim N_p(\mu, \Sigma)$，其中 $X, \mu \in \mathbb{R}^p$, $\Sigma \in \mathbb{R}^{p\times p}$。存在非奇异阵 $A \in \mathbb{R}^{p\times p}$，使得 $Y = AX \in \mathbb{R}^p$，则有

$$Y = AX \sim N_p(A\mu,\, A\Sigma A^T) \tag{10}$$

证明 由 $X \sim N_p(\mu, \Sigma)$，故其特征函数为

$$\varphi_X(t) = \mathbb{E}\left\{\exp\left(it^T X\right)\right\} = \exp\left(i\mu^T t - \frac{1}{2}\cdot t^T \Sigma t\right)$$

而 $Y = AX$，所以 Y 的特征函数为

$$\begin{aligned}
\varphi_Y(t) = \varphi_{AX}(t) &= \mathbb{E}\left\{\exp\left(it^T AX\right)\right\} = \mathbb{E}\left\{\exp\left(i(A^T t)^T X\right)\right\} \\
&= \exp\left(i\mu^T (A^T t) - \frac{1}{2}\cdot (A^T t)^T \Sigma (A^T t)\right) \\
&= \exp\left(i(A\mu)^T t - \frac{1}{2}\cdot t^T (A\Sigma A^T) t\right)
\end{aligned}$$

服从正态分布的特征函数，所以 $Y = AX \sim N_p(A\mu,\, A\Sigma A^T)$，定理证毕。

证明 使用密度函数证明。由 $X \sim N_p(\mu, \Sigma)$，故其概率密度函数为

$$f_X(x) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}$$

又 $\because A$ 是非奇异的，故 A^{-1} 存在，所以 $X = A^{-1}Y$，且 Jacobi 阵为

$$\frac{\partial X}{\partial Y} = A^{-1}$$

于是，根据卷积公式，可以得到 Y 的密度函数

$$\begin{aligned}
f_Y(y) = f_X(A^{-1}y)|A^{-1}| &= (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(A^{-1}y-\mu)^T \Sigma^{-1}(A^{-1}y-\mu)\right\} |A^{-1}| \\
&= (2\pi)^{-\frac{p}{2}} |A\Sigma A^T|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(y-A\mu)^T (A^{-1})^T \Sigma^{-1} A^{-1}(y-A\mu)\right\} \\
&= (2\pi)^{-\frac{p}{2}} |A\Sigma A^T|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(y-A\mu)^T (A\Sigma A^T)^{-1}(y-A\mu)\right\}
\end{aligned}$$

上面的变化用到了 $(AB)^T = B^T A^T$，$(AB)^{-1} = B^{-1}A^{-1}$，$|A| = |A^T|$ 和 $|A^{-1}| = |A|^{-1}$。

注意到 $f_Y(y)$ 也服从正态分布的密度函数，于是 $Y = AX \sim N_p(A\mu,\, A\Sigma A^T)$，定理证毕。

注 这里的 $|\Sigma|$ 指数为负数，所以要求 Σ 是正定的（$\Sigma > 0$），但我们的第一个证明可以发现，在特征函数定义的正态分布中，对 Σ 只有半正定的约束（$\Sigma \geq 0$）。

推论 3 特别地，在定理 2.2 中，若 $\mu = 0$ 为零向量，$\Sigma = I_p$ 为 p 阶单位阵，A 为正交阵（$AA^T = I_p$），则有 $Y = AX \sim N_p(0,\, I_p)$。

2.2 正态变量样本均值和样本方差的分布

定理 2.3 设 n 个服从正态分布的样本 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(a,\, \sigma^2)$ 记样本均值为 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 和样本方差为 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 则有:

(1) $\bar{X} \sim N(a,\, \sigma^2/n)$

(2) $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$

(3) $\bar{X}$ 和 S^2 相互独立。

此处的 χ^2_{n-1} 表示自由度为 $n-1$ 的卡方分布。

证明 (1) 由推论 2，易证 $\bar{X} \sim N(a,\, \sigma^2/n)$。

证明 (2) 构造一个正交阵 A 形如

$$A = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}_{n\times n}$$

这一正交阵的存在性可由 Schmidt 正交化方法保证。

作正交变换 $Y = AX$，其中 $X = (X_1, X_2, \cdots, X_n)^T$, $Y = (Y_1, Y_2, \cdots, Y_n)^T$，于是 $X, Y \in \mathbb{R}^n$。于是有

$$Y_1 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i = \sqrt{n}\bar{X}$$

由于正交变换的长度不变性，有

$$Y_1^2 + Y_2^2 + \cdots + Y_n^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

即

$$Y^TY = (AX)^T(AX) = X^TA^TAX = X^TX$$

所以，

$$\begin{aligned}(n-1)S^2 &= \sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i - n\bar{X}^2 \\ &= \sum_{i=1}^{n} Y_i - Y_1^2 = \sum_{i=2}^{n} Y_i\end{aligned}$$

而注意到由定理 2.2 可知，$Y = AX \sim N_n(A\mu,\, A\Sigma A^T)$，其中 $\mu = (a, a, \cdots, a)^T \in \mathbb{R}^n$ 为 X 的均值，而 $\Sigma = \sigma^2 I_n \in \mathbb{R}^{n\times n}$ 为 X 的协方差阵。

于是 Y 的协方差阵 $A\Sigma A^T = A\sigma^2 I_n A^T = \sigma^2 AA^T = \sigma^2 I$ 为对角阵。多元正态分布的协方差阵为对角阵，则随机向量的各分量相互独立。于是 $Y_i \sim N(\mu_i,\, \sigma^2)$，其中 μ_i 为 $A\mu$ 的第 i 个分量。

而由 A 的正交性，有

$$\mu_i = a\sum_{k=1}^{n} a_{ik} = \sqrt{n}a\sum_{k=1}^{n}\frac{1}{\sqrt{n}}\cdot a_{ik} = 0, \quad i \geq 2$$

因为倒数第二个式子即为 A 的第一行点乘第 i 行。或：正交阵有 $A_i \cdot A_1 = 0,\ i \geq 1$，其中 A_i 代表 A 的第 i 行，于是有 $\mathbb{E}(Y_i) = A_i\mathbb{E}(X) = A_i(a, a, \cdots, a) = A_i a\sqrt{n}A_1 = 0$。

于是，我们得到 $Y_2, Y_3, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$，故 $Y_2/\sigma, Y_3/\sigma, \cdots, Y_n/\sigma \overset{\text{i.i.d.}}{\sim} N(0, 1)$，所以由卡方分布的定义，可知

$$\frac{(n-1)S^2}{\sigma} = \sum_{i=2}^{n}\left(\frac{Y_i}{\sigma}\right)^2 \sim \chi^2_{n-1}$$

故 (2) 证毕。

证明 (3) 由 (2) 的证明过程可知，$Y_1, Y_2, \cdots, Y_n$ 相互独立，$\bar{X}$ 只与 Y_1 有关，S^2 只与 $Y_2, Y_3, \cdots, Y_n$ 有关。所以 $\bar{X}$ 和 S^2 相互独立，故 (3) 证毕。

3 χ^2 分布，t 分布和 F 分布

3.1 χ^2 分布

3.1.1 χ^2 分布的定义及 Γ 分布

定义 3.1 (χ^2 分布) 设 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(0, 1)$，则称

$$\xi = \sum_{i=1}^{n} X_i^2$$

是自由度为 n 的 χ^2 变量，其分布称为自由度为 n 的 χ^2 分布，记作 $\xi \sim \chi^2_n$。

χ^2 变量的概率密度函数由下面的定理给出。

定理 3.1 设随机变量 ξ 是自由度 n 的 χ^2 随机变量，则其概率密度函数为

$$g_n(x) = \begin{cases} \dfrac{1}{2^{n/2}\cdot\Gamma(n/2)}\cdot x^{n/2-1}e^{-x/2}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

证明 由于 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(0, 1)$，故其联合密度函数为

$$f(x_1, x_2, \cdots, x_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right\}$$

令 r.v. $\xi = \sum_{i=1}^{n} X_i^2$ 的分布函数为 $G_n(x)$ 则有

$$G_n(X) = P(\sum_{i=1}^{n} X_i^2 \leq x) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int_{\sum_{i=1}^{n} X_i^2 < x}\cdots\int \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right\} dx_1 dx_2\cdots dx_n$$

作 n 维球坐标变换

$$\begin{cases} x_1 = \rho\cos\theta_1\cos\theta_2\cdots\cos\theta_{n-2}\cos\theta_{n-1} \\ x_2 = \rho\cos\theta_1\cos\theta_2\cdots\cos\theta_{n-2}\sin\theta_{n-1} \\ x_3 = \rho\cos\theta_1\cos\theta_2\cdots\sin\theta_{n-2} \\ \quad\cdots\cdots \\ x_{n-1} = \rho\cos\theta_1\sin\theta_2 \\ x_n = \rho\sin\theta_1 \end{cases}$$

变换的 Jacobi 行列式的绝对值为

$$|\det J| = \left|\frac{\partial(x_1, x_2, \cdots, x_n)}{\partial(\rho, \theta_1, \cdots, \theta_{n-1})}\right| = \rho^{n-1}D(\theta_1, \theta_2, \cdots, \theta_{n-1})$$

其中 $D(\theta_1, \theta_2, \cdots, \theta_{n-1})$ 代表 $\theta_1, \theta_2, \cdots, \theta_{n-1}$ 的某个函数。且 $0 < \rho < \sqrt{x}$；$-\pi/2 < \theta_i < \pi/2,\ i \leq n-2$；$-\pi < \theta_{n-1} < \pi$ 因此有

$$\begin{aligned} G_n(x) &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \int_0^{\sqrt{x}} \rho^{n-1}e^{-\frac{\rho^2}{2}}\,d\rho \int_{-\pi}^{\pi}\int_{-\pi/2}^{\pi/2}\cdots\int_{-\pi/2}^{\pi/2} D(\theta_1, \theta_2, \cdots, \theta_{n-1})\,d\theta_1\cdots d\theta_{n-1} \\ &= C_n\int_0^{\sqrt{x}} \rho^{n-1}e^{-\frac{\rho^2}{2}}\,d\rho \end{aligned}$$

其中

$$C_n = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int_{-\pi}^{\pi}\int_{-\pi/2}^{\pi/2}\cdots\int_{-\pi/2}^{\pi/2} D(\theta_1, \theta_2, \cdots, \theta_{n-1})\,d\theta_1\cdots d\theta_{n-1}$$

令 $y = \rho^2$，则 $d\rho = 1/(2\sqrt{y})\,dy$，故有

$$G_n(x) = \frac{1}{2}C_n\int_0^x y^{\frac{n}{2}-1}e^{-\frac{y}{2}}\,dy \tag{11}$$

下面确定 C_n，由于

$$1 = G_n(+\infty) = \frac{1}{2}C_n\int_0^{+\infty} y^{\frac{n}{2}-1}e^{-\frac{y}{2}}\,dy = C_n 2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})$$

于是有 $C_n = 1/[2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})]$，代回式 11 有

$$G_n(x) = \frac{1}{2^{n/2}\Gamma(n/2)}\int_0^x y^{\frac{n}{2}-1}e^{-\frac{y}{2}}\,dy$$

因此 ξ 的密度函数为

$$g_n(x) = \frac{\partial G_n}{\partial x} = \begin{cases} \dfrac{1}{2^{n/2}\cdot\Gamma(n/2)}\cdot x^{n/2-1}e^{-x/2}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

（当 $n \leq 0$ 时，情况平凡）定理证毕。

3.1.2 Γ 函数和 Beta 函数

定义 3.2 (Γ 函数) 上面出现的 $\Gamma(\cdot)$ 为 **Gamma** 函数，其积分形式为

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t}\, dt$$

有关 Gamma 函数的性质：

(1) $\Gamma(1) = 1$，$\Gamma(1/2) = \sqrt{\pi}$

(2) $\Gamma(x+1) = x \cdot \Gamma(x)$

(3) $\Gamma(n) = (n-1)!$，其中 $n \in \mathbb{N}^{+}$

证明 使用数学归纳法证明。

(1) 当 $n = 1$ 时，$\xi = X_1^2$，于是 ξ 的分布函数为

$$G_1(x) = P(X_1^2 \le x) = P(-\sqrt{x} \le X_1 \le \sqrt{x}) = \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}}\, dx_1$$

由于被积函数 $\frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}}$ 为偶函数，故我们有 ξ 的密度函数为

$$g_1(x) = \frac{\partial G_1}{\partial x} = 2(\frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}})(\frac{1}{2} x^{-\frac{1}{2}}) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}} = \frac{1}{2^{1/2}\Gamma(1/2)} x^{-\frac{1}{2}} e^{-\frac{x}{2}}$$

所以 $n = 1$ 时结论成立。

(2) 假设 $n = k$ 时成立，即

$$g_k(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$$

(3) 当 $n = k+1$ 时，令 $\xi_k = \sum_{i=1}^{k} X_i^2$, $\eta = X_{k+1}^2$，则 $\xi_{k+1} = \xi_k + \eta$。由 $n = 1, k$ 时命题成立，所以 $\eta = X_{k+1}^2$ 的密度函数为 $g_1(x)$，ξ_k 的密度函数为 $g_k(x)$。于是有

$$\begin{aligned}
G_{k+1}(x) &= P(\xi_{k+1} \le x) = P(\xi_k + \eta \le x) \\
&= \iint_{\xi_k + \eta \le x} g_k(\xi_k) g_1(\eta)\, d\xi_k d\eta = \int_0^x \left(\int_0^{x-\eta} g_k(\xi_k)\, d\xi_k \right) g_1(\eta)\, d\eta \\
&= \int_0^x G_k(x-\eta) g_1(\eta)\, d\eta
\end{aligned}$$

于是有 ξ_{k+1} 的密度函数为

$$\begin{aligned}
g_{k+1}(x) &= \frac{\partial G_{k+1}}{\partial x} = \int_0^x g_k(x-\eta) g_1(\eta)\, d\eta \\
&= \int_0^x \frac{1}{2^{k/2}\Gamma(k/2)} (x-\eta)^{k/2-1} e^{-(x-\eta)/2} \frac{1}{2^{1/2}\Gamma(1/2)} \eta^{-1/2} e^{-\eta/2}\, d\eta \\
&= \frac{1}{2^{(k+1)/2}\Gamma(k/2)\Gamma(1/2)} e^{-x/2} \int_0^x (x-\eta)^{k/2-1} \eta^{-1/2}\, d\eta
\end{aligned}$$

对积分做换元 $\eta = xt,\ t \in [0,1]$，从而 $d\eta = xdt$，于是有

$$\int_0^x (x-\eta)^{k/2-1}\eta^{-1/2}\,d\eta = \int_0^1 x^{k/2-1}(1-t)^{k/2-1}(xt)^{-1/2}x\,dt = x^{(k+1)/2-1}\int_0^1 (1-t)^{k/2-1}t^{-1/2}\,dt$$

注意到 Beta 函数的定义，上式可写作

$$x^{(k+1)/2-1}\int_0^1 (1-t)^{k/2-1}t^{-1/2}\,dt = x^{(k+1)/2-1}B(\frac{1}{2},\frac{k}{2}) = x^{(k+1)/2-1}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{k}{2})}{\Gamma(\frac{k+1}{2})}$$

代回式 $g_{k+1}(x)$ 中得到

$$\begin{aligned} g_{k+1}(x) &= \frac{1}{2^{(k+1)/2}\Gamma(k/2)\Gamma(1/2)}e^{-x/2}\int_0^x (x-\eta)^{k/2-1}\eta^{-1/2}\,d\eta \\ &= \frac{1}{2^{(k+1)/2}\Gamma(k/2)\Gamma(1/2)}e^{-x/2}x^{(k+1)/2-1}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{k}{2})}{\Gamma(\frac{k+1}{2})} \\ &= \frac{1}{2^{(k+1)/2}\Gamma((k+1)/2)}x^{(k+1)/2-1}e^{-x/2} \end{aligned}$$

正是所需的形式。

由数学归纳法，命题对所有正整数 n 成立（当 $n \le 0$ 时，情况平凡），证毕。

定义 3.3 (Beta 函数) 上面出现的 $B(\cdot,\cdot)$ 为 **Beta 函数**，其积分形式为

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad t \in (0,1)$$

更多结论详见《概率论与数理统计教程》（茆诗松等）章节 2.5.5 贝塔分布。

定义 3.4 (Γ 分布) 若随机变量 X 满足如下概率密度函数

$$p(x;\alpha,\lambda) = \begin{cases} \dfrac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

则称 X 满足 Γ **分布**，记作 $X \sim \Gamma(\alpha,\lambda)$。

命题 3.2 Γ 分布的 2 个特例：

- 自由度为 n 的 χ^2 分布与 Γ 分布的关系为

$$\chi_n^2 = \Gamma\left(\frac{n}{2},\frac{1}{2}\right)$$

- 指数分布于 Γ 分布的关系为

$$\mathrm{Exp}(\lambda) = \Gamma\left(1,\lambda\right)$$

命题 3.3 若 r.v. X 服从 Γ 分布，即 $X \sim \Gamma(\alpha,\lambda)$，则 X 的期望和方差为

$$\mathbb{E}(X) = \frac{\alpha}{\lambda}, \quad var(X) = \frac{\alpha}{\lambda^2}$$

证明 期望为

$$\mathbb{E}(X) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\lambda x}\, dx = \frac{\lambda^\alpha}{\lambda^{\alpha+1}\Gamma(\alpha)} \int_0^\infty (\lambda x)^{(\alpha+1)-1} e^{-\lambda x}\, d\lambda x$$
$$= \frac{1}{\lambda}\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\alpha}{\lambda}$$

同理，二阶矩为

$$\mathbb{E}(X^2) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+1} e^{-\lambda x}\, dx = \frac{\lambda^\alpha}{\lambda^{\alpha+2}\Gamma(\alpha)} \int_0^\infty (\lambda x)^{(\alpha+2)-1} e^{-\lambda x}\, d\lambda x$$
$$= \frac{1}{\lambda^2}\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} = \frac{\alpha(\alpha+1)}{\lambda^2}$$

于是，方差为

$$var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{\alpha(\alpha+1)}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2 = \frac{\alpha}{\lambda^2}$$

综上，命题证毕。

命题 3.4 若 r.v. $Y \sim \Gamma(\alpha, \lambda)$，则 $Z = 2\lambda Y \sim \chi^2_{2\alpha}$

证明 若 $Y \sim \Gamma(\alpha, \lambda)$，其概率密度函数为

$$f_Y(y) = \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y}, \quad y > 0.$$

设 $Z = 2\lambda Y$，则有

$$y = \frac{z}{2\lambda}, \quad \frac{dy}{dz} = \frac{1}{2\lambda}$$

因此，Z 的密度函数为

$$f_Z(z) = f_Y(\frac{z}{2\lambda}) \cdot \frac{1}{2\lambda}, \quad z > 0.$$

代入有

$$f_Z(z) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \left(\frac{z}{2\lambda}\right)^{\alpha-1} e^{-\lambda\cdot\left(\frac{z}{2\lambda}\right)} \cdot \frac{1}{2\lambda}$$
$$= \frac{1}{2^\alpha \Gamma(\alpha)} z^{\alpha-1} e^{-z/2} = \frac{(\frac{1}{2})^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\frac{1}{2}z}$$
$$= f_{\Gamma(\alpha, \frac{1}{2})}(z)$$

综上 $Z = 2\lambda Y \sim \Gamma\left(\alpha, \frac{1}{2}\right) = \chi^2_{2\alpha}$，命题证毕。

3.1.3 χ^2 分布的性质

命题 3.5 χ^2_n 的密度函数 $g_n(x)$ 满足：

- 当 $n = 1, 2$ 时，曲线是单调下降趋于 0 的，即

$$g_n'(x) \leq 0, \quad n = 1, 2$$

- 当 $n \geq 3$ 时，曲线有单峰，从 0 开始先单调上升，达到峰值后，单调下降趋于 0。

命题 3.6 设 r.v. $\xi \sim \chi_n^2$，则 ξ 的特征函数为

$$\varphi_\xi(t) = (1-2it)^{-\frac{n}{2}}$$

证明 因为 $\xi \sim \chi_n^2$，则 $\xi \sim \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ 其密度函数为

$$f_\xi(x) = \frac{(1/2)^{n/2}}{\Gamma(n/2)} x^{n/2-1} e^{-x/2}, \qquad x > 0.$$

于是，根据特征函数的定义，有

$$\varphi_\xi(t) = \mathbb{E}(e^{it\xi}) = \int_0^\infty e^{itx} f_\xi(x)\, dx = \frac{(1/2)^{n/2}}{\Gamma(n/2)} \int_0^\infty x^{n/2-1} e^{-(1/2-it)x}\, dx.$$

注意积分为 Γ 函数的形式，即

$$\int_0^\infty x^{\alpha-1} e^{-\beta x}\, dx = \frac{\Gamma(\alpha)}{\beta^\alpha}$$

此处 $\alpha = n/2,\ \beta = 1/2 - it$，代入可得

$$\varphi_\xi(t) = \frac{(1/2)^{n/2}}{\Gamma(n/2)} \cdot \frac{\Gamma(n/2)}{(1/2-it)^{n/2}} = \left(\frac{1/2}{1/2-it}\right)^{\frac{n}{2}} = (1-2it)^{-\frac{n}{2}}$$

命题证毕。

注 或者可以采用 $\xi = \sum_{i=1}^n X_i^2$，其中 $X_i \overset{\text{i.i.d.}}{\sim} N(0,1)$。然后根据 X_i 标准正态分布的密度函数，计算 X_i^2 的特征函数。最后由命题 1.2 有关独立随机变量和的特征函数性质可证。

命题 3.7 设 r.v. $\xi \sim \chi_n^2$，则 ξ 的期望和方差分别为

$$\mathbb{E}(\xi) = n \quad var(\xi) = 2n$$

证明 由 Γ 分布的命题 3.2 可知，χ^2 分布是 Γ 分布的特例，即 $\xi \sim \chi_n^2 = \Gamma(n/2, 1/2)$。且对于 Γ 分布 $\Gamma(\alpha,\lambda)$，其期望和方差为（由命题 3.3）

$$\mathbb{E}(X) = \frac{\alpha}{\lambda}, \quad var(X) = \frac{\alpha}{\lambda^2}$$

于是 χ^2 分布的期望和方差为

$$\mathbb{E}(\xi) = \frac{n/2}{1/2} = n \quad var(\xi) = \frac{n/2}{(1/2)^2} = 2n$$

命题证毕。

命题 3.8 设 $Z_1 \sim \chi_{n_1}^2$，$Z_2 \sim \chi_{n_2}^2$，且 Z_1 和 Z_2 相互独立。则有 $Z_1 + Z_2 \sim \chi_{n_1+n_2}^2$。

证明 根据特征函数证明。由 χ^2 分布的特征函数可知

$$\varphi_{Z_1}(t) = (1-2it)^{-\frac{n_1}{2}} \quad \varphi_{Z_2}(t) = (1-2it)^{-\frac{n_2}{2}}$$

因为 $Z_1 \perp Z_2$，由命题 1.2 知

$$\varphi_{Z_1+Z_2}(t) = \varphi_{Z_1}(t)\cdot\varphi_{Z_2}(t) = (1-2it)^{-\frac{n_1}{2}} \cdot (1-2it)^{-\frac{n_2}{2}} = (1-2it)^{-\frac{n_1+n_2}{2}}$$

恰好为 $\chi_{n_1+n_2}^2$ 的特征函数。所以，$Z_1+Z_2 \sim \chi_{n_1+n_2}^2$，命题证毕。

证明 根据定义证明。根据 χ^2 分布的定义，可知 Z_1 和 Z_2 可以分别写作 n_1, n_2 个标准正态随机变量的平方和。不妨设

$$Z_1 = \sum_{i=1}^{n_1} X_i^2 \quad Z_2 = \sum_{j=1}^{n_2} Y_i^2$$

其中 $X_i, Y_j \overset{\text{i.i.d.}}{\sim} N(0,\, 1)$, $i = 1, 2, \cdots, n_1$, $j = 1, 2, \cdots, n_2$。于是有

$$Z_1 + Z_2 = \sum_{i=1}^{n_1} X_i^2 + \sum_{j=1}^{n_2} Y_i^2 \sim \chi^2_{n_1+n_2}$$

由 χ^2 分布的定义可知，证毕。

3.1.4 非中心 χ^2 分布

定义 3.5 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立，且 $X_i \sim N(a_i,\, 1)$ 其中 a_i, $i = 1, 2, \cdots, n$ 不全为 0。记

$$Y = \sum_{i=1}^{n} X_i^2$$

则称 Y 的分布是自由度为 n 和非中心参数为

$$\delta = \sqrt{\sum_{i=1}^{n} a_i^2}$$

的**非中心 χ^2 分布**，记作 $Y \sim \chi^2_{n,\delta}$。特别地，当 $\delta = 0$ 时称为中心的 χ^2 分布，即前面所述的 χ^2_n 分布。

命题 3.9 若 r.v. $Y \sim \chi^2_{n,\delta}$，则其概率密度函数为

$$\begin{aligned} g(x) &= \begin{cases} e^{-\delta^2/2} \displaystyle\sum_{i=1}^{\infty} \frac{1}{i!} \left(\frac{\delta^2}{2}\right)^i \frac{x^{i+n/2-1}}{2^{i+n/2}\Gamma(n/2+i)}, & x > 0 \\ 0, & x \leq 0 \end{cases} \\ &= \begin{cases} e^{-\delta^2/2} \displaystyle\sum_{i=1}^{\infty} \frac{(\delta^2/2)^i}{i!} \chi^2(x, 2i+n), & x > 0 \\ 0, & x \leq 0 \end{cases} \end{aligned}$$

此处 $\chi^2(x, 2i+n)$ 表示自由度为 $2i+n$ 的 χ^2 分布的概率密度函数。

证明 (1) 作正交变换使 $X_1 + X_2 + \cdots + X_n = Y_1^2 + Z$，其中 $Y_1 \sim N(\delta,\, 1)$，$Z \sim \chi^2_{n-1}$；(2) 再利用求 r.v. 和的分布公式求出 $Y_1^2 + Z$ 的密度函数，此即 $Y = \sum_{i=1}^{n} X_i$ 的密度函数。

命题 3.10 若 $Y \sim \chi^2_{n,\delta}$，则 Y 的特征函数为

$$\varphi_Y(t) = (1 - 2it)^{-\frac{n}{2}} \cdot \exp\left(\frac{i\delta^2 t}{1 - 2it}\right)$$

注 此处书本出现了笔误，书本在特征函数的指数部分多添加了负号。

证明 由非中心 χ^2 分布的定义，可将 Y 写作

$$Y = \sum_{i=1}^{n} X_i^2$$

其中 $X_i \sim N(a_i, 1)$，其 X_i, $i = 1, 2, \cdots, n$ 相互独立。由独立随机变量和的特征函数等于随机变量特征函数的积（命题 1.2），我们有

$$\varphi_Y(t) = \varphi_{\sum X_i^2}(t) = \prod_{i=1}^{n} \varphi_{X_i^2}(t) \tag{12}$$

下面，我们只需求每个 X_i^2 的特征函数 $\varphi_{X_i^2}(t)$ 即可，不妨去求 $X \sim N(\mu, 1)$ 的特征函数 $\varphi_{X^2}(t)$

$$\begin{aligned}
\varphi_{X^2}(t) = \mathbb{E}(e^{itX^2}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(itx^2 - \frac{1}{2}(x-\mu)^2\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-(\frac{1}{2} - it)x^2 + \mu x - \frac{\mu^2}{2}\right) dx
\end{aligned}$$

记 $a = 1/2 - it$，则有

$$-ax^2 + \mu x - \frac{\mu^2}{2} = -a(x - \frac{\mu}{2a})^2 + \frac{\mu^2}{4a} - \frac{\mu^2}{2}$$

代回得

$$\begin{aligned}
\varphi_{X^2}(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-a(x - \frac{\mu}{2a})^2 + \frac{\mu^2}{4a} - \frac{\mu^2}{2}\right) dx \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(\frac{\mu^2}{4a} - \frac{\mu^2}{2}\right) \int_{-\infty}^{\infty} \exp\left(-a(x - \frac{\mu}{2a})^2\right) dx
\end{aligned}$$

下面我们先证明一个引理：

引理 3.11 对任意正实数 $a \in \mathbb{R}^+$，我们有下面的积分结果

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \sqrt{\frac{\pi}{a}}$$

证明 设 $y = \sqrt{a}x$，则 $dx = 1/\sqrt{a}dy$，那么原式变成

$$\int_{-\infty}^{\infty} e^{-ax^2}\, dx = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} e^{-y^2}\, dy$$

而通过简单的极坐标变换即可得到 $\int_{-\infty}^{\infty} e^{-y^2}\, dy = \sqrt{\pi}$，代入即证。

回到本问题 3.10 的证明，代入有

$$
\begin{aligned}
\varphi_{X^2}(t) &= \frac{1}{\sqrt{2\pi}}\exp\left(\frac{\mu^2}{4a}-\frac{\mu^2}{2}\right)\sqrt{\frac{\pi}{a}} = \frac{1}{\sqrt{2a}}\exp\left(\frac{\mu^2}{4a}-\frac{\mu^2}{2}\right) \\
&= \frac{1}{\sqrt{2(1/2-it)}}\exp\left(\frac{\mu^2}{4(1/2-it)}-\frac{\mu^2}{2}\right) \\
&= (1-2it)^{-1/2}\exp\left(\frac{\mu^2}{2(1-2it)}-\frac{\mu^2}{2}\right) \\
&= (1-2it)^{-1/2}\exp\left(\frac{\mu^2/2(1-1-2it)}{1-2it}\right) \\
&= (1-2it)^{-1/2}\exp\left(\frac{i\mu^2 t}{1-2it}\right)
\end{aligned}
$$

于是，代入式 12 有

$$
\begin{aligned}
\varphi_Y(t) &= \prod_{i=1}^n \varphi_{X_i^2}(t) = \prod_{i=1}^n (1-2it)^{-1/2}\exp\left(\frac{ia_i^2 t}{1-2it}\right) \\
&= (1-2it)^{-n/2}\exp\left(\frac{it}{1-2it}\sum_{i=1}^n a_i^2\right) \\
&= (1-2it)^{-\frac{n}{2}}\cdot\exp\left(\frac{i\delta^2 t}{1-2it}\right)
\end{aligned}
$$

综上，命题证毕。

命题 3.12 *若 $Y \sim \chi^2_{n,\delta}$，则 Y 的期望和方差分别为*

$$
\mathbb{E}(Y) = n+\delta^2 \quad var(Y) = 2n+4\delta^2
$$

证明 已知特征函数 $\varphi_Y(t)$ 可以求得矩母函数为

$$
\begin{aligned}
M_Y(t) = \varphi_Y(-it) &= (1-2i(-it))^{-\frac{n}{2}}\cdot\exp\left(\frac{i\delta^2(-it)}{1-2i(-it)}\right) \\
&= (1-2t)^{-\frac{n}{2}}\cdot\exp\left(\frac{\delta^2 t}{1-2t}\right)
\end{aligned}
$$

先对 $M_Y(t)$ 作对数变换

$$
\log M_Y(t) = -\frac{n}{2}\log(1-2t)+\frac{\delta^2 t}{1-2t}
$$

然后求一阶导

$$
\frac{1}{M_Y(t)}M'_Y(t) = \frac{n}{1-2t}+\frac{\delta^2(1-2t)+2\delta^2 t}{(1-2t)^2} = \frac{n}{1-2t}+\frac{\delta^2}{(1-2t)^2}
$$

注意到 $M_Y(0)=1$，令 $t=0$，可求得期望等于 $M'_Y(0)$ 为

$$
\mathbb{E}(Y) = M'_Y(0) = M_Y(0)\cdot\left(\frac{n}{1-0}+\frac{\delta^2}{(1-0)^2}\right) = n+\delta^2
$$

类似地，求两次导为

$$
\frac{-M'_Y(t)}{M_Y^2(t)}M'_Y(t)+\frac{1}{M_Y(t)}M''_Y(t) = \frac{2n}{(1-2t)^2}+\frac{4\delta^2}{(1-2t)^3}
$$

令 $t=0$，可求得二阶矩 $\mathbb{E}(Y^2)=M_Y''(0)$ 为

$$\begin{aligned}\mathbb{E}(Y^2)=M_Y''(0)&=\left[\left(\frac{2n}{(1-0)^2}+\frac{4\delta^2}{(1-0)^3}\right)+\frac{(M_Y'(0))^2}{M_Y^2(0)}\right]M_Y(0)\\&=[(2n+4\delta^2)+(n+\delta^2)^2/1^2]\times 1=(2n+4\delta^2)+(n+\delta^2)^2\end{aligned}$$

所以，方差为

$$var(Y)=\mathbb{E}(Y^2)-(\mathbb{E}(Y))^2=(2n+4\delta^2)+(n+\delta^2)^2-(n+\delta^2)^2=2n+4\delta^2$$

综上，$\mathbb{E}(Y)=n+\delta^2\quad var(Y)=2n+4\delta^2$，命题证毕。

命题 3.13 若 $Y_1,Y_2,\cdots,Y_k$ 相互独立，且 $Y_i\sim\chi^2_{n_i,\delta_i}$，则有

$$\sum_{i=1}^{k}Y_i\sim\chi^2_{n,\delta}$$

其中

$$n=\sum_{i=1}^{k}n_i,\quad \delta=\sqrt{\sum_{i=1}^{k}\delta_i^2}.$$

证明 相互独立配合特征函数即可证毕。

3.2 t 分布

3.2.1 t 分布的定义

定义 3.6 (t 分布) 设 r.v. $X\sim N(0,1)$，$Y\sim\chi^2_n$，且 X 和 Y 相互独立，则称

$$T=\frac{X}{\sqrt{Y/n}}$$

是自由度为 n 的 t 变量，其分布称为自由度为 n 的 **t 分布**，记作 $T\sim t_n$。

命题 3.14 设随机变量 $T\sim t_n$，则其概率密度函数为

$$t_n(x)=\frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2},\quad x\in\mathbb{R}$$

证明 首先，根据 $Y\sim\chi^2_n$，我们先求 $U=\sqrt{Y/n}$ 的概率密度，由于，分布函数 $F_U(u)$ 为

$$F_U(u)=P(U=\sqrt{Y/n}\le u)=P(Y\le nu^2)=F_Y(nu^2)$$

对 u 求导得密度函数 $f_U(u)$

$$f_U(u)=\frac{d}{du}F_U(u)=f_Y(nu^2)\frac{d}{du}(nu^2)=2nu\cdot f_Y(nu^2)$$

而 $Y\sim\chi_n^2$，则可知 $f_Y(\cdot)$ 的表达式，代入则有

$$\begin{aligned}f_U(u)&=2nu\cdot f_Y(nu^2)=2nu\cdot\frac{1}{2^{n/2}\Gamma(n/2)}(nu^2)^{n/2-1}e^{-nu^2/2}\\&=\frac{2n^{n/2}}{2^{n/2}\Gamma(n/2)}u^{n-1}e^{-nu^2/2},\quad u\geq 0\end{aligned}$$

为方便记号，常数可写作

$$C:=\frac{2n^{n/2}}{2^{n/2}\Gamma(n/2)}$$

下面求解 $T=X/U,\ X\sim N(0,1),\ U\sim F_U$ 的概率密度。注意到 X 和 U 独立，故联合分布为

$$f_{X,U}(x,u)=f_X(x)\cdot f_U(u)$$

对于变换 $T=X/U,\ Z=U$ 有 $X=TZ,\ U=Z$，于是 Jacobi 行列式的绝对值为

$$|\det J|=\left|\det\begin{pmatrix}\partial x/\partial t & \partial x/\partial z\\ \partial u/\partial t & \partial u/\partial z\end{pmatrix}\right|=\left|\det\begin{pmatrix}z & t\\ 0 & 1\end{pmatrix}\right|=|z|$$

所以 T,Z 的联合密度函数为

$$f_{T,Z}(t,z)=f_X(tz)\cdot f_U(z)|z|=f_X(tz)\cdot f_U(z)\cdot z,\quad z\geq 0$$

于是 T 的边际分布概率密度为

$$\begin{aligned}f_T(t)&=\int_{-\infty}^{\infty}f_{T,Z}(t,z)\,dz=\int_0^{\infty}f_X(tz)\cdot f_U(z)\cdot z\,dz\\&=\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-t^2z^2/2}\cdot Cz^{n-1}e^{-nz^2/2}\,dz\\&=\frac{C}{\sqrt{2\pi}}\int_0^{\infty}z^n\exp\left(-\frac{t^2+n}{2}z^2\right)\,dz\end{aligned}$$

记 $(t^2+n)/2=\alpha$，则积分形式化为

$$\int_0^{\infty}z^ne^{-\alpha z^2}\,dz=\frac{1}{2\alpha^{(n+1)/2}}\int_0^{\infty}(\alpha z^2)^{(n+1)/2-1}e^{-\alpha z^2}\,d(\alpha z^2)=\frac{1}{2\alpha^{(n+1)/2}}\Gamma((n+1)/2)$$

代回原式，得到

$$\begin{aligned}f_T(t)&=\frac{C}{\sqrt{2\pi}}\cdot\frac{1}{2\alpha^{(n+1)/2}}\Gamma((n+1)/2)\\&=\frac{2n^{n/2}}{2^{n/2}\Gamma(n/2)}\cdot\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{2\alpha^{(n+1)/2}}\Gamma((n+1)/2)\\&=\frac{2n^{n/2}}{2^{n/2}\Gamma(n/2)}\cdot\frac{1}{\sqrt{2\pi}}\cdot\frac{1}{2((t^2+n)/2)^{(n+1)/2}}\Gamma((n+1)/2)\\&=\frac{1}{\sqrt{2\pi}}\cdot\frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)}\cdot\frac{2^{(n+1)/2}}{2^{n/2}}\cdot n^{n/2}(n+t^2)^{-(n+1)/2}\\&=\frac{1}{\sqrt{2\pi}}\frac{\sqrt{2}}{\sqrt{n}}\cdot\frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)}\cdot\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}\end{aligned}$$

所以 T 的密度函数为

$$f_T(t) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2} \quad \square$$

3.2.2 t 分布的性质

命题 3.15 t_n 的密度函数与标准正态分布 $N(0,1)$ 密度很相似，它们都是关于原点对称的单峰的偶函数，在 $x=0$ 处达到极大值。但 t_n 的峰值低于 $N(0,1)$ 的峰值，且 t_n 存在厚尾。

重要的是，t 分布的极限分布为标准正态分布。

$$\lim_{n\to\infty} t_n(x) = \phi(x)$$

其中 $\phi(x)$ 为 $N(0,1)$ 的概率密度函数，即 t_n **依分布收敛于** $N(0,1)$。

证明 通过函数极限可以证明 $\lim_{n\to\infty} t_n(x) = \phi(x)$，即可证明

$$t_n \xrightarrow{\mathscr{L}} N(0,1)$$

或者采用 **Slutsky 引理**证明：

由于 Y_n 可记作 $\sum_{i=1}^n X_i^2$，其中 $X_i \overset{\text{i.i.d.}}{\sim} N(0,1)$，所以根据 **Khinchin 大数定律**

$$Y_n/n = \bar{X_n^2} \xrightarrow{P} E(X_1^2) = 1$$

所以 $T = X/\sqrt{Y_n/n} \xrightarrow{\mathscr{L}} X/\sqrt{1} = X$，即 $N(0,1)$。$\square$

命题 3.16 若 r.v. $T\sim t_n$，则 $\mathbb{E}(T^r)$ 只有当 $r<n$ $(n>1)$ 时存在，且

$$\mathbb{E}(T^r) = \begin{cases} n^{r/2}\cdot\dfrac{\Gamma\left(\dfrac{r+1}{2}\right)\Gamma\left(\dfrac{n-r}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\Gamma\left(\dfrac{1}{2}\right)}, & r \text{ 为偶数} \\ 0, & r \text{ 为奇数} \end{cases}$$

特别地，当 $n\ge 2$ 时，$\mathbb{E}(T)=0$。当 $n\ge 3$ 时，$\mathbb{D}(T)=n/(n-2)$。

证明 由命题 3.14 知 T 的概率密度函数

$$t_n(x) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2}, \quad x\in\mathbb{R}$$

若矩存在，先讨论奇数阶矩。且由命题 3.15 可知 $t_n(x)$ 为偶函数，则 $x^r t_n(x)$ 在 $2\nmid r$ 时是奇函数。于是 $\mathbb{E}(T^r)=0$, $2\nmid r$。

下面具体讨论 r 阶矩的存在性和偶数阶矩的表达式。由于已知概率密度函数，实则探讨下面积分的存在性

$$\mathbb{E}(T^r) = \int_{-\infty}^{\infty} x^r \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\sqrt{n\pi}} \left(1+\frac{x^2}{n}\right)^{-(n+1)/2} dx$$

为方便化简，记常数

$$C := \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\sqrt{n\pi}}$$

而当 r 为偶数时，积分对偶函数积分，于是有

$$\mathbb{E}(T^r) = 2C\int_0^{\infty} x^r \left(1+\frac{x^2}{n}\right)^{-(n+1)/2} dx$$

作换元 $y = \left(1+\frac{x^2}{n}\right)^{-1}$，于是 $x = \left[n\left(\frac{1-y}{y}\right)\right]^{1/2}$，那么有

$$\begin{aligned}
\mathbb{E}(T^r) &= 2C\cdot\int_1^0 \left[n\left(\frac{1-y}{y}\right)\right]^{r/2} y^{\frac{n+1}{2}}\, d\left[n\left(\frac{1-y}{y}\right)\right]^{1/2} \\
&= 2C\cdot\int_1^0 n^{r/2}\frac{(1-y)^{r/2}}{y^{r/2}} y^{\frac{n+1}{2}} \frac{n^{1/2}}{2}\frac{(1-y)^{-1/2}}{y^{-1/2}}\frac{-1}{y^2}\, dy \\
&= C\cdot n^{(r+1)/2}\int_0^1 y^{-\frac{1}{r}\frac{n+1}{2}+\frac{1}{2}-2}\cdot(1-y)^{\frac{r}{2}-\frac{1}{2}}\, dy \\
&= C\cdot n^{(r+1)/2}\int_0^1 y^{\frac{n-r}{2}-1}\cdot(1-y)^{\frac{r+1}{2}-1}\, dy
\end{aligned}$$

注意到 Beta 函数 $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\, dt$，所以有

$$\mathbb{E}(T^r) = C\cdot n^{(r+1)/2}\cdot B\left(\frac{n-r}{2}, \frac{r+1}{2}\right)$$

由 Beta 函数的性质，存在性需要 $\frac{n-r}{2} > 0$ 即 $r < n$。此时，再由 $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$，以及 $\Gamma(1/2) = \sqrt{\pi}$，有

$$\begin{aligned}
\mathbb{E}(T^r) &= \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\sqrt{n\pi}}\cdot n^{(r+1)/2}\cdot\frac{\Gamma\left(\dfrac{n-r}{2}\right)\cdot\Gamma\left(\dfrac{r+1}{2}\right)}{\Gamma\left(\dfrac{n-r}{2}+\dfrac{r+1}{2}\right)} \\
&= n^{r/2}\cdot\frac{\Gamma\left(\dfrac{r+1}{2}\right)\Gamma\left(\dfrac{n-r}{2}\right)}{\Gamma\left(\dfrac{n}{2}\right)\Gamma\left(\dfrac{1}{2}\right)}, \quad r < n,\ 2 \mid r
\end{aligned}$$

于是，命题证毕。

命题 3.17 当 $n = 1$ 时，t 分布就是 Cauchy 分布，此时概率密度函数变为

$$t_1(x) = \frac{1}{\pi(1+x^2)}, \quad x\in\mathbb{R}$$

由命题 3.16 可知，Cauchy 分布不存在任何矩 $\mathbb{E}(T^r)$, $\forall\, r$。

3.2.3 非中心 t 分布

定义 3.7 设 r.v. $X \sim N(\delta, 1)$ 以及 r.v. $Y \sim \chi_n^2$ 且 X 和 Y 相互独立，则称

$$Z = \frac{X}{\sqrt{Y/n}}$$

的分布是自由度为 n 和非中心参数为 δ 的**非中心 t 分布**，记作 $Z \sim t_{n,\delta}$。特别地，当 $\delta = 0$ 时为中心的 t 分布，即定义 3.6 的 t_n。

命题 3.18 非中心 t 分布 $t_{n,\delta}$ 的密度函数为

$$t_{n,\delta}(x) = \frac{n^{n/2}}{\sqrt{\pi}\Gamma(n/2)} \cdot \frac{e^{-\delta^2/2}}{(n+x^2)^{\frac{n+1}{2}}} \sum_{i=0}^{\infty} \Gamma\left(\frac{n+i+1}{2}\right) \frac{(\delta x)^i}{i!} \left(\frac{2}{n+x^2}\right)^{i/2}, \quad x \in \mathbb{R}$$

命题 3.19 若 $Z_n \sim t_{n,\delta}$，则其依分布收敛于正态分布：

$$Z_n \xrightarrow{\mathscr{L}} N(\delta, 1)$$

证明 同样使用 Slutsky 引理，根据 $Y_n/n \xrightarrow{P} 1$ 可证。

命题 3.20 若 $Z_n \sim t_{n,\delta}$，则其期望和方差为

$$\mathbb{E}(Z_n) = \delta \left(\frac{n}{2}\right)^{\frac{1}{2}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}, \quad n \geq 2$$

$$\mathbb{D}(Z_n) = \frac{n(1+\delta^2)}{n-2} - \frac{\delta^2 n}{2} \left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right)^2, \quad n \geq 3$$

3.3 F 分布

3.3.1 F 分布的定义

定义 3.8 (F 分布) 设 r.v. $X \sim \chi_m^2$ 和 r.v. $Y \sim \chi_n^2$ 且 X, Y 相互独立，则称

$$F = \frac{X/m}{Y/n}$$

是自由度为 m 和 n 的 F 变量。其分布称为自由度为 m 和 n 的 **F 分布**，记作 $F \sim F_{m,n}$。

命题 3.21 设 r.v. $Z \sim F_{m,n}$，则其概率密度函数为

$$f_{m,n}(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma\left(\dfrac{m}{2}\right)\Gamma\left(\dfrac{n}{2}\right)} m^{\frac{m}{2}} n^{\frac{n}{2}} x^{\frac{m}{2}-1} (n+mx)^{-\frac{m+n}{2}}, & x > 0 \\ 0, & \text{其他} \end{cases}$$

3.3.2 F 分布的性质

命题 3.22 若 $Z \sim F_{m,n}$ 则 $1/Z \sim F_{n,m}$。

命题 3.23 若 $Z \sim F_{m,n}$ 则对 $r>0$ 有

$$\mathbb{E}(Z^r) = \left(\frac{n}{m}\right)^r \frac{\Gamma\left(\frac{m}{2}+r\right)\Gamma\left(\frac{n}{2}-r\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}, \quad 2r<n$$

特别地

$$\mathbb{E}(Z) = \frac{n}{n-2}, \quad n>2$$
$$\mathbb{D}(Z) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, \quad n>4$$

命题 3.24 若 $T \sim t_n$，则 $T^2 \sim F_{1,n}$

命题 3.25 若 $F_{m,n}(\alpha)$ 代表 $F_{m,n}$ 的**上 α 分位数点**，即对于 $Z \sim F_{m,n}$ 有 $P(Z \geq F_{m,n}(\alpha)) = \alpha$，于是有

$$F_{m,n}(\alpha) \cdot F_{n,m}(1-\alpha) = 1$$

证明 根据上分位数点的定义，对于 $Z \sim F_{m,n}$ 有

$$\alpha = P(Z \geq F_{m,n}(\alpha)) = P(1/Z \leq 1/F_{m,n}(\alpha)) = 1 - P(1/Z \geq 1/F_{m,n}(\alpha))$$

即有

$$P(1/Z \geq 1/F_{m,n}(\alpha)) = 1-\alpha$$

而 $1/Z \sim F_{n,m}$，于是 $1/F_{m,n}(\alpha)$ 是 $F_{n,m}$ 的上 $1-\alpha$ 分位数点，即

$$1/F_{m,n}(\alpha) = F_{n,m}(1-\alpha) \Rightarrow F_{m,n}(\alpha) \cdot F_{n,m}(1-\alpha) = 1$$

命题证毕。

3.3.3 非中心 F 分布

定义 3.9 设 r.v. $X \sim \chi^2_{m,\delta}$ 和 r.v. $Y \sim \chi^2_n$ 且 X, Y 相互独立，则称

$$Z = \frac{X/m}{Y/n}$$

的分布是自由度为 m,n 和非中心参数为 δ 的**非中心 F 分布**，记作 $Z \sim F_{m,n;\delta}$。当 $\delta = 0$ 时，称 Z 为中心的 F 分布，记作 $F_{m,n}$。

命题 3.26 若 $Z \sim F_{m,n;\delta}$，则 Z 的概率密度函数为

$$f_{m,n;\delta}(x) = \begin{cases} \dfrac{m^{\frac{n}{2}} n^{\frac{m}{2}}}{\Gamma\left(\frac{n}{2}\right)} e^{-\frac{\delta^2}{2}} x^{\frac{m}{2}-1} \displaystyle\sum_{k=0}^{\infty} \frac{\left(\frac{\delta^2 mx}{2}\right)^k \Gamma\left(\frac{m+n}{2}+k\right)}{k!\Gamma\left(\frac{m}{2}+k\right)(mx+n)^{\frac{m+n}{2}+k}}, & x>0 \\ 0, & \text{其他} \end{cases}$$

命题 3.27 若 $X \sim t_{n,\delta}$，则 $X^2 \sim F_{1,n;\delta}$。

命题 3.28 若 $Z_n \sim F_{m,n;\delta}$, $n = 1, 2, \cdots$ 且 δ 固定，则当 $n \to \infty$ 时

$$Z_n \xrightarrow{\mathscr{L}} \frac{\chi^2_{m,\delta}}{m}$$

命题 3.29 若 $Z \sim F_{m,n;\delta}$，则其的期望和方差为

$$\begin{aligned} \mathbb{E}(Z) &= \frac{m(m+\delta)}{m(n-2)}, \quad n > 2 \\ \mathbb{D}(Z) &= \frac{2n^2}{m(n-2)^2(n-4)}[(m+\delta^2)^2 + (n-2)(m+2\delta^2)], \quad n > 4 \end{aligned}$$

3.4 几个重要推论

回顾定理 2.3 以及上面的三大抽样分布（定义 3.1，3.6，3.8），我们得到下面几个重要的推理。

推论 4 设 $X_1, X_2, \cdots, X_n$ 相互独立，且 $X_i \sim N(a_i, \sigma_i^2)$, $i = 1, 2, \cdots, n$ 则

$$\sum_{i=1}^n \left(\frac{X_i - a_i}{\sigma_i}\right)^2 \sim \chi^2_n$$

证明 因为 $Y_i = (X_i - a_i)/\sigma_i \sim N(0,1)$，且 $Y_1, Y_2, \cdots, Y_n$ 相互独立，由定义 3.1 可知 $\sum_{i=1}^n (X_i - a_i)^2/\sigma_i^2 = \sum_{i=1}^n Y_i^2 \sim \chi^2_n$，命题证毕。

推论 5 设 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} N(a, \sigma^2)$，则

$$T = \frac{\sqrt{n}(\bar{X} - a)}{S} \sim t_{n-1}$$

其中 $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ 为样本均值，$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ 为样本无偏方差。

证明 由定理 2.3 可知 $\bar{X} \sim N(a, \sigma^2/n)$ 及 $(n-1)S/\sigma \sim \chi^2_{n-1}$，将 $\bar{X}$ 标准化得 $\sqrt{n}(\bar{X} - a)/\sigma \sim N(0,1)$，且 $\bar{X}$ 和 S 独立。于是根据定义 3.6 可知

$$T = \frac{\sqrt{n}(\bar{X} - a)}{\sqrt{S^2/\sigma^2}} = \frac{\sqrt{n}(\bar{X} - a)}{S} \sim t_{n-1}$$

命题证毕。

推论 6 设 $X_1, X_2, \cdots, X_m \overset{\text{i.i.d.}}{\sim} N(a_1, \sigma^2)$ 以及 $Y_1, Y_2, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} N(a_2, \sigma^2)$，且样本 $X_1, X_2, \cdots, X_m$ 和 $Y_1, Y_2, \cdots, Y_n$ 相互独立，则

$$T = \frac{(\bar{X} - \bar{Y}) - (a_1 - a_2)}{S_w} \cdot \sqrt{\frac{mn}{m+n}} \sim t_{m+n-2}$$

其中 $(m+n-2)S_w^2 = (m-1)S_1^2 + (n-1)S_2^2$ 此处

$$S_1^2 = \frac{1}{m-1}\sum_{i=1}^m (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n-1}\sum_{i=1}^m (Y_i - \bar{Y})^2$$

证明 由定理 2.3 可知 $\bar{X} \sim N(a_1, \sigma^2/m)$，$\bar{Y} \sim N(a_2, \sigma^2/n)$，故有 $\bar{X} - \bar{Y} \sim N(a_1 - a_2, (m+n)\sigma^2/mn)$ 将其标准化

$$\frac{(\bar{X} - \bar{Y}) - (a_1 - a_2)}{\sigma}\sqrt{\frac{mn}{m+n}} \sim N(0,1)$$

又 $(m-1)S_1^2/\sigma^2 \sim \chi^2_{m-1}$，$(n-1)S_2^2/\sigma^2 \sim \chi^2_{n-1}$，且二者独立，根据 χ^2 分布的性质

$$\frac{(m-1)S_1^2 + (n-1)S_2^2}{\sigma^2} \sim \chi^2_{m+n-2}$$

又 $(\bar{X}, \bar{Y})$ 和 (S_1^2, S_2^2) 独立，由定义 3.6 可知

$$\begin{aligned} T &= \frac{(\bar{X} - \bar{Y}) - (a_1 - a_2)}{\sigma}\sqrt{\frac{mn}{m+n}} \Big/ \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{\sigma^2(m+n-2)}} \\ &= \frac{(\bar{X} - \bar{Y}) - (a_1 - a_2)}{S_w} \cdot \sqrt{\frac{mn}{m+n}} \sim t_{m+n-2} \end{aligned}$$

命题证毕。

推论 7 设 $X_1, X_2, \cdots, X_m \overset{\text{i.i.d.}}{\sim} N(a_1, \sigma_1^2)$ 以及 $Y_1, Y_2, \cdots, Y_n \overset{\text{i.i.d.}}{\sim} N(a_2, \sigma_2^2)$，且样本 $X_1, X_2, \cdots, X_m$ 和 $Y_1, Y_2, \cdots, Y_n$ 相互独立，则

$$F = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{m-1,n-1}$$

此处 S_1^2, S_2^2 的定义与推论 6 相同。

证明 由定理 2.3 可知 $(m-1)S_1^2/\sigma_1^2 \sim \chi^2_{m-1}$，$(n-1)S_2^2/\sigma_2^2 \sim \chi^2_{n-1}$ 且二者独立，由定义 3.8 可知

$$F = \frac{\dfrac{(m-1)S_1^2}{\sigma_1^2} \Big/ (m-1)}{\dfrac{(n-1)S_2^2}{\sigma_2^2} \Big/ (n-1)} = \frac{S_1^2}{S_2^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F_{m-1,n-1}$$

命题证毕。

推论 8 设 $X_1, X_2, \cdots, X_n$ i.i.d. 服从指数分布 $f(x, \lambda) = \lambda e^{-\lambda x} I_{(0,\infty)}(x)$，则有

$$2\lambda n\bar{X} = 2\lambda\sum_{i=1}^{n} X_i \sim \chi^2_{2n}$$

证明 由命题 3.4 和命题 3.2 有 $\mathrm{Exp}(\lambda) = \Gamma(1, \lambda)$ 可知 $2\lambda X_i \sim \chi^2_2$。再由 χ^2 分布的性质，

$$2\lambda\sum_{i=1}^{n} X_i \sim \chi^2_{\sum_{i=1}^{n} 2} = \chi^2_{2n}$$

命题证毕。

4 次序统计量

定义 4.1 (次序统计量) 设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本，记

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$$

则称 $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ 分别为样本的第 $1, 2, \cdots, n$ 个次序统计量。

4.1 单个次序统计量的分布

命题 4.1 设 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$，$f$ 为 F 的密度函数，则第 $1 \le m \le n$ 个次序统计量 $X_{(m)}$ 的概率密度函数为

$$f_m(x) = m\binom{n}{m}[F(x)]^{m-1}[1-F(x)]^{n-m}f(x), \quad x \in \mathbb{R}$$

证明 因为

$$\begin{aligned} F_m(x) &= P(X_{(m)} < x) = P(X_1, X_2, \cdots, X_n \text{ 中至少有 } m \text{ 个 } X_i < x) \\ &= \sum_{i=m}^{n} P(X_1, X_2, \cdots, X_n \text{ 恰有 } i \text{ 个 } X_i < x) \end{aligned}$$

若记 $A_i = \{X_i < x\}$，$i = 1, 2, \cdots, n$ 则 $P(A_i) = P(X_i < x) = F(x)$，且 $A_1, A_2, \cdots, A_n$ 相互独立。于是

$$\{X_1, X_2, \cdots, X_n \text{ 中恰有 } i \text{ 个} < x\} = \{\text{事件 } A_1, A_2, \cdots, A_n \text{ 中恰有 } i \text{ 个发生}\}$$

而这个事件的概率可用二项分布 $b(n, F(x))$ 表示，即

$$P(\{\text{事件 } A_1, A_2, \cdots, A_n \text{ 中恰有 } i \text{ 个发生}\}) = \binom{n}{i}[F(x)]^i[1-F(x)]^{n-i}$$

故

$$F_m(x) = \sum_{i=m}^{n}\binom{n}{i}[F(x)]^i[1-F(x)]^{n-i}$$

利用恒等式（证明见作业 **HW 2**）

$$\sum_{i=m}^{n}\binom{n}{i}p^i(1-p)^{n-i} = m\binom{n}{m}\int_0^p t^{m-1}(1-t)^{n-m}\,dt$$

可知

$$\begin{aligned} F_m(x) &= m\binom{n}{m}\int_0^{F(x)} t^{m-1}(1-t)^{n-m}\,dt \\ \Rightarrow \quad f_m(x) &= F_m'(x) = m\binom{n}{m}[F(x)]^{m-1}[1-F(x)]^{n-m}f(x) \end{aligned}$$

即第 m 个次序统计量的概率密度函数为

$$f_m(x) = m\binom{n}{m}[F(x)]^{m-1}[1-F(x)]^{n-m}f(x), \quad x \in \mathbb{R}$$

命题证毕。

4.2 次序统计量的联合分布

命题 4.2 (n 个次序统计量的联合分布) 设 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$，$f$ 为 F 的密度函数，则次序统计量 $(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 的联合概率密度函数为

$$f(x_{(1)}, x_{(2)}, \cdots, x_{(n)}) = \begin{cases} n! f(x_{(1)}) f(x_{(2)}) \cdots f(x_{(n)}), & x_{(1)} < x_{(2)} < \cdots < x_{(n)} \\ 0, & \text{其他} \end{cases}$$

证明 令 $y_i = x_{(i)}$，$i = 1, 2, \cdots, n$，则

$$\begin{aligned} G(y_1, y_2, \cdots, y_n) &:= P(X_{(1)} < y_1,\ X_{(2)} < y_2,\ \cdots,\ X_{(n)} < y_n) \\ &= \begin{cases} n! P(X_{j_1} < y_1,\ X_{j_2} < y_2,\ \cdots,\ X_{j_n} < y_n), & y_1 < y_2 < \cdots < y_n \\ 0, & \text{其他} \end{cases} \end{aligned}$$

其中 $X_{j_1} < \cdots < X_{j_n}$ 且 $(j_1, j_2, \cdots, j_n)$ 是 $(1, 2, \cdots, n)$ 的任意排列。故 n 个次序统计量的联合概率密度函数为

$$g(y_1, y_2, \cdots, y_n) = \begin{cases} n! f(y_1) f(y_2) \cdots f(y_n), & y_1 < y_2 < \cdots < y_n \\ 0, & \text{其他} \end{cases}$$

命题证毕。

命题 4.3 (2 个次序统计量的联合分布) 设 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} F$，$f$ 为 F 的密度函数，则第 $1 \le i < j \le n$ 个次序统计量 $(X_{(i)}, X_{(j)})$ 的联合概率密度函数为

$$f_{i,j}(x, y) = \begin{cases} \dfrac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} [1 - F(y)]^{n-j} f(x) f(y), & x < y \\ 0, & \text{其他} \end{cases}$$

证明 作业 HW 3 证明了下面的结论：设 r.v. X 的分布函数为 $F(x)$，密度函数为 $f(x)$，且 $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} f$，则有

$$\int \cdots \int_{a < x_1 < \cdots < x_n < b} f(x_1) \cdots f(x_n)\, dx_1 \cdots dx_n = \frac{1}{n!} [F(b) - F(a)]^n \tag{13}$$

回到本题，由于 $f_{i,j}(x, y)$ 是 $(X_{(1)}, \cdots, X_{(n)})$ 联合密度 $g(y_1, \cdots, y_n)$，$y_i = x_{(i)}$ 的边缘密度，对 $x < y$ 有

$$f_{i,j}(x, y) = \int \cdots \int_{-\infty < x_1 < \cdots < x_n < \infty} n! \prod_{l \ne i,j} f(x_l) \cdot f(x) f(y)\, dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_{j-1} dx_{j+1} \cdots dx_n$$

由 $f \ge 0$ 且可积，根据 Fubini 定理，可将上式拆成 3 部分

$$f_{i,j}(x, y) = n! \cdot A(x) \cdot B(x, y) \cdot C(y)$$

根据式 13 可知 $A(\cdot), B(\cdot,\cdot), C(\cdot)$ 分别为

$$
\begin{aligned}
A(x) &= \int\cdots\int_{-\infty<x_1<\cdots<x_{i-1}<x} \prod_{l=1}^{i-1} f(x_l)\cdot f(x)\,dx_1\cdots dx_{i-1} \\
&= \frac{1}{(i-1)!}[F(x)-F(-\infty)]^{i-1}f(x) \\
B(x,y) &= \int\cdots\int_{x<x_{i+1}<\cdots<x_{j-1}<y} \prod_{l=i+1}^{j-1} f(x_l)\cdot f(y)\,dx_1\cdots dx_{i-1} \\
&= \frac{1}{(j-i-1)!}[F(y)-F(x)]^{j-i-1}f(y) \\
C(y) &= \int\cdots\int_{y<x_{j+1}<\cdots<x_n<+\infty} \prod_{l=j+1}^{n} f(x_l)\,dx_1\cdots dx_{i-1} \\
&= \frac{1}{(n-j)!}[F(+\infty)-F(y)]^{n-j}
\end{aligned}
$$

综上可知，当 $x<y$ 有概率密度

$$
\begin{aligned}
f_{i,j}(x,y) &= n!A(x)B(x,y)C(y) \\
&= \frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(x)]^{i-1}[F(y)-F(x)]^{j-i-1}[1-F(y)]^{n-j}f(x)f(y)
\end{aligned}
$$

命题证毕。

4.3 极差的分布

下面求 $V=X_{(j)}-X_{(i)}$，$i<j$ 的分布。

作变换 $V=X_{(j)}-X_{(i)}$，$Z=X_{(i)}$ 则有 $X_{(j)}=V+Z$，$X_{(i)}=Z$，于是 Jacobi 行列式为 $J=\left|\dfrac{\partial(X_{(j)},X_{(i)})}{\partial V\partial Z}\right|=1$。那么容易得到 (V,Z) 的联合概率密度函数（$v>0$）为

$$
g_{i,j}(v,z)=\frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(z)]^{i-1}[F(v+z)-F(z)]^{j-i-1}[1-F(v+z)]^{n-j}f(z)f(v+z)
$$

所以 V 的密度函数为

$$
g_V(v)=\int_{-\infty}^{\infty} g_{i,j}(v,z)\,dz
$$

5 统计量的极限分布

5.1 随机变量列的收敛

定义 5.1 (依概率收敛) 考虑定义在同一概率空间上的随机变量列 $\{X_n\}_{n\geq1}$ 以及随机变量 X。若对 $\forall\,\epsilon>0$ 都有

$$
\lim_{n\to\infty} P(|X_n-X|\leq\epsilon)=1
$$

则称 X_n **依概率收敛**到 X，记作 $X_n \xrightarrow{P} X$。

定义 5.2 (依分布收敛) 考虑定义在同一概率空间上的随机变量列 $\{X_n\}_{n\geq 1}$ 以及随机变量 X。X 的分布函数 $F(\cdot)$ 的连续点集合为 C。若 X_n 的分布函数 $F_n(\cdot)$ 对 $\forall\, x \in C$ 都满足

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

则称 X_n **依分布收敛**到 X，记作 $X_n \xrightarrow{\mathscr{L}} X$。

定义 5.3 (几乎处处收敛) 考虑定义在同一概率空间上的随机变量列 $\{X_n\}_{n\geq 1}$ 以及随机变量 X。若

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1,$$

则称 X_n **几乎处处收敛**到 X，记作 $X_n \xrightarrow{\text{a.s.}} X$。

命题 5.1 有关依概率收敛和依分布收敛的相关性质：

(1) 依概率收敛强于依分布收敛，即

$$X_n \xrightarrow{P} X \quad \Rightarrow \quad X_n \xrightarrow{\mathscr{L}} X$$

(2) 当极限是常数时，二者等价，即

$$X_n \xrightarrow{P} c \quad \Leftrightarrow \quad X_n \xrightarrow{\mathscr{L}} c$$

其中 $c \in \mathbb{R}$ 为常数。

(3) 随机向量列的收敛：

$$(X_n,\, Y_n)^T \xrightarrow{P} (X,\, Y)^T \quad \Leftrightarrow \quad X_n \xrightarrow{P} X,\, Y_n \xrightarrow{P} Y$$
$$(X_n,\, Y_n)^T \xrightarrow{\mathscr{L}} (X,\, Y)^T \quad \Rightarrow \quad X_n \xrightarrow{\mathscr{L}} X,\, Y_n \xrightarrow{\mathscr{L}} Y$$

即向量的依概率收敛等价于分量均依概率收敛，但分量的依分布收敛不能推出向量的依分布收敛。

引理 5.2 (Slutsky 引理) 令 $\{X_n\}$ 和 $\{Y_n\}$ 是两个随机变量的序列，满足当 $n \to \infty$ 时，

$$X_n \xrightarrow{\mathscr{L}} X, \quad Y_n \xrightarrow{P} c$$

其中 $c \in \mathbb{R}$ 为常数。则有：

(1) $X_n \pm Y_n \xrightarrow{\mathscr{L}} X \pm c$
(2) $X_n Y_n \xrightarrow{\mathscr{L}} cX$
(3) $X_n / Y_n \xrightarrow{\mathscr{L}} X/c$

5.2 大数定律

定理 5.3 (Khinchin 大数定律 / 弱大数定律) 考虑独立同分布的随机变量列 $(X_i)_{i\geq 1}$ 假设 $\mathbb{E}(X_1) < \infty$。记 $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ 且 $\mu = \mathbb{E}(X_1)$。则

$$\bar{X}_n \xrightarrow{P} \mu$$

定理 5.4 (Kolmogorov 大数定律 / 强大数定律) 考虑独立同分布的随机变量列 $(X_i)_{i\geq 1}$ 假设 $\mathbb{E}(X_1) < \infty$，且 $\mathrm{Var}(X_1^2) < \infty$，记 $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ 且 $\mu = \mathbb{E}(X_1)$。则

$$\bar{X}_n \xrightarrow{\text{a.s.}} \mu$$

定理 5.5 (Markov 大数定律) 设随机变量列 $\{X_n\}_{n\geq 1}$ 满足

$$\lim_{n\to\infty} \mathbb{D}(\frac{1}{n}\sum_{i=1}^n X_i) = 0$$

则有

$$\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{n}\sum_{i=1}^n \mathbb{E}(X_i) \xrightarrow{P} 0, \quad \text{as } n \to \infty$$

5.3 中心极限定理

定理 5.6 (Lindeberg-Lévy 中心极限定理) 考虑独立同分布的随机变量列 $(X_i)_{i\geq 1}$ 记 $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ 且 $\mu = \mathbb{E}(X_1)$。若 $\sigma^2 = \mathbb{D}(X_1) \in (0, +\infty)$，则当 $n \to \infty$ 时有

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{\mathscr{L}} N(0, 1)$$

6 指数族

6.1 定义和常见的指数族

定义 6.1 (指数型分布族) 设 $\mathscr{F} = \{f(x;\theta) : \theta \in \Theta\}$ 是定义在样本空间 $\mathscr{X}$ 上的分布族，其中 Θ 为参数空间。若概率密度函数 $f(x;\theta)$ 可表示成如下形式

$$f(x;\theta) = C(\theta)\exp\left\{\sum_{i=1}^k Q_i(\theta)T_i(x)\right\} h(x)$$

其中 $k \in \mathbb{N}^+$ 且 $C(\theta) > 0$ 和 $Q_i(\theta)$, $i = 1, 2, \cdots, k$ 都是定义在参数空间 Θ 上的函数，而 $h(x) > 0$ 和 $T_i(x)$, $i = 1, 2, \cdots, k$ 都是定义在样本空间 $\mathscr{X}$ 上的函数。则称此分布族为**指数型分布族**，简称指数族（exponential family）。

定义 6.2 (支撑集) 一个分布族 $\mathscr{F} = \{f(x;\theta) : \theta \in \Theta\}$ 的支撑集为

$$G(x) = \{x : f(x;\theta) > 0\}$$

命题 6.1 指数族中所有分布具有共同的支撑集，不难发现这个支撑集为

$$G(x) = \{x : f(x;\theta) > 0\} = \{x : h(x) > 0\}$$

且支撑集与参数 θ 无关。

命题 6.2 设 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T \in \mathbb{R}^n$ 是从总体 F 中抽取的简单随机样本，则

- 当 $F = N(\mu, \sigma^2)$ 时，样本分布族是指数族。
- 当 $F = \Gamma(\gamma, \lambda)$ 时，样本分布族是指数族。

设随机变量 X 来自总体 F 则

- 当 $F = b(n, p)$ 时，样本分布族是指数族。
- 当 $F = \text{Poi}(\lambda)$ 时，样本分布族是指数族。
- 当 $F = U(0, \theta)$ 时，样本分布族不是指数族。

6.2 指数族的自然形式及自然参数空间

定义 6.3 若指数族有下列形式

$$f(x; \boldsymbol{\theta}) = C^*(\boldsymbol{\theta}) \exp\left\{\sum_{i=1}^{k} \theta_i T_i(x)\right\} h(x)$$

则称它为**指数族的自然形式**（natural form）。此时，集合

$$\Theta^* = \left\{\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k) : \int_{\mathscr{X}} \exp\left\{\sum_{i=1}^{k} \theta_i T_i(x)\right\} h(x)\, dx < \infty\right\}$$

称为**自然参数空间**（natural parametric space）。

注 推导自然形式可设 $\varphi_i = Q_i(x)$ 而后将 $C(\theta)$ 写成 $\boldsymbol{\varphi}$ 的函数 $C^*(\varphi)$，其中 $\boldsymbol{\varphi} = (\varphi_1, \cdots, \varphi_k)$。

6.3 指数族的性质

定理 6.3 在指数族的自然形式下，自然参数空间为凸集。即对于自然参数空间

$$\Theta^* = \left\{\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_k) : \int_{\mathscr{X}} \exp\left(\sum_{i=1}^{k} \theta_i T_i(x)\right) h(x)\, dx < \infty\right\}$$

其中 $h(x) > 0$，而 $\mathscr{X}$ 为样本空间。结论是：对于任意 $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta^*$，设 $0 < \alpha < 1$，记 $\boldsymbol{\theta} = \alpha\boldsymbol{\theta}_1 + (1-\alpha)\boldsymbol{\theta}_2$，则 $\boldsymbol{\theta} \in \Theta^*$。

证明 记 $\boldsymbol{T}(x) = (T_1(x), T_2(x), \cdots, T_k(x))^T \in \mathbb{R}^k$，则我们需要证明：

$$\int_{\mathscr{X}} \exp\left(\boldsymbol{\theta}^T \boldsymbol{T}(x)\right) h(x)\, dx < \infty$$

其中 $\boldsymbol{\theta} = \alpha\boldsymbol{\theta}_1 + (1-\alpha)\boldsymbol{\theta}_2$ 代入即

$$\exp\left(\boldsymbol{\theta}^T \boldsymbol{T}(x)\right) h(x) = \left[e^{\alpha\boldsymbol{\theta}_1^T \boldsymbol{T}(x)} h(x)^{\alpha}\right] \cdot \left[e^{(1-\alpha)\boldsymbol{\theta}_2^T \boldsymbol{T}(x)} h(x)^{1-\alpha}\right]$$

记 $p = 1/\alpha$，$q = 1/(1-\alpha)$，同时记 $\exp\left(\boldsymbol{\theta}^T \boldsymbol{T}(x)\right) h(x) := f(x) g(x)$，其中

$$f(x) = e^{\frac{1}{p}\boldsymbol{\theta}_1^T \boldsymbol{T}(x)} h(x)^{\frac{1}{p}} > 0 \quad g(x) = e^{\frac{1}{q}\boldsymbol{\theta}_2^T \boldsymbol{T}(x)} h(x)^{\frac{1}{q}} > 0$$

由于 $\alpha \in (0,1) \Rightarrow p,q > 1$ 且 $\frac{1}{p} + \frac{1}{q} = \alpha + (1-\alpha) = 1$，所以由 Hölder 不等式 6.5 有

$$\int_{\mathscr{X}} f(x)g(x)\,dx \leq \left(\int_{\mathscr{X}} f(x)^p\,dx\right)^{1/p} \left(\int_{\mathscr{X}} g(x)^q\,dx\right)^{1/q} \tag{14}$$

而

$$f(x)^p = \left(e^{\frac{1}{p}\boldsymbol{\theta}_1^T\boldsymbol{T}(x)}h(x)^{\frac{1}{p}}\right)^p = e^{\boldsymbol{\theta}_1^T\boldsymbol{T}(x)}h(x)$$
$$g(x)^q = \left(e^{\frac{1}{q}\boldsymbol{\theta}_2^T\boldsymbol{T}(x)}h(x)^{\frac{1}{q}}\right)^q = e^{\boldsymbol{\theta}_2^T\boldsymbol{T}(x)}h(x)$$

又因为 $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Theta^*$，所以

$$\int_{\mathscr{X}} f(x)^p\,dx = \int_{\mathscr{X}} e^{\boldsymbol{\theta}_1^T\boldsymbol{T}(x)}h(x)\,dx < \infty$$
$$\int_{\mathscr{X}} g(x)^q\,dx = \int_{\mathscr{X}} e^{\boldsymbol{\theta}_2^T\boldsymbol{T}(x)}h(x)\,dx < \infty$$

由不等式 14 可知

$$\int_{\mathscr{X}} f(x)g(x)\,dx = \int_{\mathscr{X}} \exp\left(\boldsymbol{\theta}^T\boldsymbol{T}(x)\right)h(x)\,dx < \infty$$

即有 $\boldsymbol{\theta} \in \Theta^*$ 证毕。

定理 6.4 (Young 不等式) 设 $a,b \geq 0$ $p,q > 1$ 满足

$$\frac{1}{p} + \frac{1}{q} = 1$$

则有

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

当且仅当 $a^p = b^q$ 时，等号成立。

证明 由 $\varphi(x) = e^x$ 为凸函数，由凸函数的性质或 Jensen 不等式

$$\varphi(\lambda x_1 + (1-\lambda)x_2) \leq \lambda\varphi(x_1) + (1-\lambda)\varphi(x_2)$$

令 $\lambda = 1/p$ 由条件 $1-\lambda = 1/q$。下面，对任意 $a,b > 0$ 取 $x_1 = p\ln a$，$x_2 = q\ln b$，由上述的 Jensen 不等式

$$e^{\lambda p\ln a + (1-\lambda)q\ln b} \leq \lambda e^{p\ln a} + (1-\lambda)e^{q\ln b}$$

而左边

$$e^{\lambda p\ln a + (1-\lambda)q\ln b} = e^{\frac{1}{p}\cdot p\ln a} \cdot e^{\frac{1}{q}\cdot q\ln b} = ab$$

右边

$$\lambda e^{p\ln a} + (1-\lambda)e^{q\ln b} = \frac{1}{p}\cdot e^{\ln a^p} + \frac{1}{q}\cdot e^{\ln b^q} = \frac{a^p}{p} + \frac{b^q}{q}$$

综上

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

定理证毕。

定理 6.5 (Hölder 不等式) 设 $f(x), g(x)$ 为可测函数，$p, q > 1$ 满足

$$\frac{1}{p} + \frac{1}{q} = 1$$

则有

$$\int |f(x)g(x)|\, dx \leq \left(\int |f(x)|^p\, dx\right)^{1/p} \left(\int |g(x)|^q\, dx\right)^{1/q}$$

等号成立当且仅当存在常数 $C > 0$，使得对几乎处处的 x 有 $|f(x)|^p = C|g(x)|^q$ 成立。

证明 这里只证明不平凡的情形。记

$$A := \int |f(x)|^p\, dx \in (0,\ \infty), \quad B := \int |g(x)|^q\, dx \in (0,\ \infty)$$

设函数

$$u(x) = \frac{|f(x)|}{A^{1/p}} > 0, \quad v(x) = \frac{|g(x)|}{B^{1/q}} > 0$$

则容易得 $\int u(x)^p\, dx = \int v(x)^q\, dx = 1$。由 Young 不等式（定理 6.4），令 $a = u(x)$, $b = v(x)$，则

$$u(x)v(x) \leq \frac{u(x)^p}{p} + \frac{v(x)^q}{q}$$

由上式对任意的 x 都成立，且 $0 < u, v < \infty$，故积分不等式仍成立

$$\int u(x)v(x)\, dx \leq \int u(x)^p\, dx + \int v(x)^q\, dx = 1$$

即

$$\int \frac{|f(x)|}{A^{1/p}} \frac{|g(x)|}{B^{1/q}}\, dx = A^{-1/p}B^{-1/q} \int |f(x)g(x)|\, dx \leq 1$$

代入 A, B 即有

$$\int |f(x)g(x)|\, dx \leq \left(\int |f(x)|^p\, dx\right)^{1/p} \left(\int |g(x)|^q\, dx\right)^{1/q}$$

定理证毕。

定理 6.6 设指数族的自然形式中，自然参数空间有内点，其内点集为 Θ_0。设 $g(x)$ 为任一实函数，使得积分

$$G(\boldsymbol{\theta}) = \int_{\mathscr{X}} g(x) \exp\left\{\sum_{j=1}^{k} \theta_j T_j(x)\right\} h(x)\, dx$$

在 Θ_0 存在有限。则 $G(\boldsymbol{\theta})$ 的任意阶偏导数在 Θ_0 内存在且可在积分号下求得，即

$$\frac{\partial^m G(\boldsymbol{\theta})}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}} = \int_{\mathscr{X}} \frac{\partial^m}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}} \left[g(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\} h(x)\right] dx$$

其中 $m = \sum_{j=1}^{k} m_j$，即对 $G(\boldsymbol{\theta})$ 关于 $\boldsymbol{\theta}$ 的任意阶偏导数可在积分下求得。

定理 6.7 设指数族的自然形式中，自然参数空间有内点，其内点集为 Θ_0。对 $\boldsymbol{\theta} \in \Theta_0$，有 $T_i(x)$ 的各阶矩均存在有限，可通过在积分号下求导算出。例如：

$$\mathbb{E}(T_i(x)) = -\frac{\partial \log C^*(\boldsymbol{\theta})}{\partial\theta_i}$$

$$\mathrm{Cov}(T_i(x), T_j(x)) = -\frac{\partial^2 \log C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}$$

证明 由概率定义 $\int_{\mathscr{X}} f(x;\theta)=1$，有关系式

$$\frac{1}{C^*(\boldsymbol{\theta})}=\int_{\mathscr{X}}\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)\,dx \tag{15}$$

对式 15 两边对 θ_i 求偏导有

$$-\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}=\int_{\mathscr{X}}T_i(x)\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)\,dx \tag{16}$$

注意到

$$\mathbb{E}(T_i(x))=\int_{\mathscr{X}}T_i(x)C^*(\boldsymbol{\theta})\exp\left\{\sum_{j=1}^{k}\theta_j T_j(x)\right\}h(x)\,dx$$

所以

$$\mathbb{E}(T_i(x))=-\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\cdot C^*(\boldsymbol{\theta})=-\frac{\partial\log C^*(\boldsymbol{\theta})}{\partial\theta_i}$$

继续对式 16 关于 θ_j 求偏导有

$$\frac{2}{C^*(\boldsymbol{\theta})^3}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j}-\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial^2 C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}=\int_{\mathscr{X}}T_i(x)T_j(x)\exp\left\{\sum_{l=1}^{k}\theta_l T_l(x)\right\}h(x)\,dx \tag{17}$$

注意到

$$\begin{aligned}\mathrm{Cov}(T_i(x),T_j(x))&=\mathbb{E}(T_i(x)T_j(x))-\mathbb{E}(T_i(x))\mathbb{E}(T_j(x))\\&=\int_{\mathscr{X}}T_i(x)T_j(x)C^*(\boldsymbol{\theta})\exp\left\{\sum_{l=1}^{k}\theta_l T_l(x)\right\}h(x)\,dx-\mathbb{E}(T_i(x))\mathbb{E}(T_j(x))\end{aligned}$$

代入式 16 和 $\mathbb{E}(T_i(x)),\mathbb{E}(T_j(x))$ 的表达式得

$$\begin{aligned}\mathrm{Cov}(T_i(x),T_j(x))=&\left(\frac{2}{C^*(\boldsymbol{\theta})^3}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j}-\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial^2 C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right)\cdot C^*(\boldsymbol{\theta})\\&-\left(-\frac{1}{C^*(\boldsymbol{\theta})}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\right)\left(-\frac{1}{C^*(\boldsymbol{\theta})}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j}\right)\\=&\frac{1}{C^*(\boldsymbol{\theta})^2}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_i}\frac{\partial C^*(\boldsymbol{\theta})}{\partial\theta_j}-\frac{1}{C^*(\boldsymbol{\theta})}\frac{\partial^2 C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\\=&-\frac{\partial^2\log C^*(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\end{aligned}$$

命题证毕。

7 充分统计量

7.1 定义和示例

定义 7.1 (充分统计量) 设样本 $\boldsymbol{X}$ 的分布族为 $\{f(x;\boldsymbol{\theta}):\boldsymbol{\theta}\in\Theta\}$，其中 Θ 是参数空间。令 $\boldsymbol{T}=\boldsymbol{T}(\boldsymbol{X})$ 为一统计量，若在已知 $\boldsymbol{T}$ 的条件下，样本 $\boldsymbol{X}$ 的条件分布与参数 $\boldsymbol{\theta}$ 无关，则称 $\boldsymbol{T}(\boldsymbol{X})$ 为 $\boldsymbol{\theta}$ 的充分统计量（sufficient statistic），即

$$P_{\boldsymbol{\theta}}(\boldsymbol{X}|\boldsymbol{T}(\boldsymbol{X})=\boldsymbol{t})\text{ 与 }\boldsymbol{\theta}\text{ 无关}$$

例 7.1 设 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T \in \mathbb{R}^n$ 为从 $b(1,\theta)$ 分布中抽取的简单随机样本，则 $T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ 为充分统计量。

证明 由 $X_i \sim b(1,\theta)$ 可知 $T = \sum_{i=1}^n \sim B(n,\theta)$ 二项分布。由定义出发证明：

$$\begin{aligned}
P_{\boldsymbol{\theta}}(\boldsymbol{X}|\boldsymbol{T}(\boldsymbol{X})) = P(X_1 = x_1, \cdots, X_n = x_n|T = t_0) &= \frac{P(X_1 = x_1, \cdots, X_n = x_n, T = t_0)}{P(T = t_0)} \\
&= \frac{\theta^{t_0}(1-\theta)^{n-t_0}}{\binom{n}{t_0}\theta^{t_0}(1-\theta)^{n-t_0}} \cdot \mathbb{I}(\sum_{i=1}^n x_i = t_0) \\
&= \binom{n}{t_0}^{-1} \cdot \mathbb{I}(\sum_{i=1}^n x_i = t_0)
\end{aligned}$$

与 θ 无关，故 $T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ 为充分统计量。

例 7.2 设 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T \in \mathbb{R}^n$ 为从 $N(\theta, 1)$ 分布中抽取的简单随机样本，则 $T(\boldsymbol{X}) = \sum_{i=1}^n X_i/n = \bar{X}$ 为充分统计量。

证明 作正交变换 A 使得

$$\boldsymbol{Y} = (Y_1, \cdots, Y_n)^T = A\boldsymbol{X} = A(X_1, X_2, \cdots, X_n)^T$$

其中 A 为正交阵，且

$$A = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}_{n\times n}$$

由之前有关正态变量样本的性质推导可知 $Y_1, \cdots, Y_n$ 相互独立 $\sum_{i=1}^n Y_i^2 = \sum_{i=1}^n X_i^2$，

$$\begin{cases} Y_1 = \sum_{i=1}^n X_i/\sqrt{n} = \sqrt{n}\,\bar{X} \\ Y_j = \sum_{k=1}^n a_{jk}X_k, \quad j = 1, 2, \cdots, n \end{cases}$$

而 $\boldsymbol{Y} = A\boldsymbol{X} \sim N(A\theta, AIA^T)$，故易知 $Y_1 \sim N(\sqrt{n}\theta, 1)$ 以及 $Y_j \sim (0,1)$，$j = 1, 2, \cdots, n$。所以

$$\begin{aligned}
f(y_1, \cdots, y_n|y_1) = \frac{f(y_1, \cdots, y_n)}{f_{Y_1}(y_1)} &= \frac{(2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=2}^n y_i^2 - \frac{1}{2}(y_1 - \sqrt{n}\theta)^2\right\}}{(2\pi)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(y_1 - \sqrt{n}\theta)^2\right\}} \\
&= (2\pi)^{-\frac{n-1}{2}} \exp\left\{-\frac{1}{2}\sum_{i=2}^n y_i^2\right\}
\end{aligned}$$

与 θ 无关，即 $T = \bar{X} = Y_1/\sqrt{n}$ 为充分统计量。

7.2 因子分解定理

定理 7.1 (因子分解定理) 设样本 $\boldsymbol{X}=(X_1,\cdots,X_n)^T\in\mathbb{R}^n$ 的概率函数 $f(\boldsymbol{x};\theta)$ 依赖于参数 θ，$\boldsymbol{T}=T(\boldsymbol{X})$ 是一个统计量。

$$\boldsymbol{T}\text{ 为充分统计量}\Leftrightarrow f(\boldsymbol{x};\theta)\text{ 可分解为 }f(\boldsymbol{x};\theta)=g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x})$$

这里概率函数是指：若 $\boldsymbol{X}$ 连续，则 f 为概率密度函数；若 $\boldsymbol{X}$ 离散，则 $f(\boldsymbol{x},\theta)=P_\theta(X_1=x_1,\cdots,X_n=x_n)$ 为概率分布。

证明 (离散情形 $\Leftarrow$) 已知 $f(\boldsymbol{x},\theta)$ 可以分解，需证

$$P_\theta(\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})=\frac{P(\boldsymbol{X}=\boldsymbol{x},\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})}{P(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})}=\frac{P(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}|\boldsymbol{X}=\boldsymbol{x})P(\boldsymbol{X}=\boldsymbol{x})}{P(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})}\tag{18}$$

与 θ 无关（等号成立是因为 Bayes 变换 / 条件概率定义）。

记集合 $A_t=\{\boldsymbol{x}:\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}\}$，注意到当 $\boldsymbol{X}=\boldsymbol{x}$ 已知时，$\boldsymbol{T}(\boldsymbol{x})$ 也已知，从而对于 $P(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}|\boldsymbol{X}=\boldsymbol{x})$ 只能为 1 或 0。设示性函数 $\mathbb{I}_A(\omega)=1$ 当 $\omega\in A$ 否则 0。于是有

$$P(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}|\boldsymbol{X}=\boldsymbol{x})=\mathbb{I}_{\{\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}\}}(\boldsymbol{x})=\mathbb{I}_{A_t}(\boldsymbol{x})\in\{0,1\}\tag{19}$$

由全概率公式，以及 $f(\boldsymbol{x},\theta)$ 可以分解，即

$$P(\boldsymbol{X}=\boldsymbol{x})=g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x})\tag{20}$$

可知

$$\begin{aligned}P(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})&=\sum_{\boldsymbol{y}\in\mathscr{X}}P(\boldsymbol{T}(\boldsymbol{y})=\boldsymbol{t}|\boldsymbol{X}=\boldsymbol{y})P(\boldsymbol{X}=\boldsymbol{y})\\&=\sum_{\boldsymbol{y}\in\mathscr{X}}\mathbb{I}_{\{\boldsymbol{T}(\boldsymbol{y})=\boldsymbol{t}\}}(\boldsymbol{y})\cdot g(\boldsymbol{T}(\boldsymbol{y}),\theta)h(\boldsymbol{y})\\&=\sum_{\boldsymbol{y}\in\{\boldsymbol{x}:\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}\}}g(\boldsymbol{T}(\boldsymbol{y}),\theta)h(\boldsymbol{y})\\&=\sum_{\boldsymbol{y}\in A_t}g(\boldsymbol{t},\theta)h(\boldsymbol{y})\end{aligned}$$

将上式和式 19 和 20 代入式 18 可得

$$P_\theta(\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})=\frac{\mathbb{I}_{\{\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}\}}(\boldsymbol{x})\cdot g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x})}{\sum_{\boldsymbol{y}\in A_t}g(\boldsymbol{t},\theta)h(\boldsymbol{y})}=\frac{g(\boldsymbol{t},\theta)h(\boldsymbol{x})}{\sum_{\boldsymbol{y}\in A_t}g(\boldsymbol{t},\theta)h(\boldsymbol{y})}=\frac{h(\boldsymbol{x})}{\sum_{\boldsymbol{y}\in A_t}h(\boldsymbol{y})}$$

与 θ 无关，故 $\boldsymbol{T}$ 是充分统计量。

证明 (离散情形 $\Rightarrow$) 已知 $\boldsymbol{T}$ 为充分统计量，需证 $f(\boldsymbol{x},\theta)$ 可分解。由于 $\boldsymbol{T}$ 为充分统计量，所以

$$P_\theta(\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})=P(\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}):=h(\boldsymbol{x})$$

与 θ 无关。由条件概率定义有

$$f(\boldsymbol{x},\theta)=P_\theta(\boldsymbol{X}=\boldsymbol{x})=P(\boldsymbol{X}=\boldsymbol{x}|\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t})\cdot P_\theta(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}):=h(\boldsymbol{x})\cdot g(\boldsymbol{T}(\boldsymbol{x}),\theta)$$

其中 $P_\theta(\boldsymbol{T}(\boldsymbol{x})=\boldsymbol{t}):=g(\boldsymbol{T}(\boldsymbol{x}),\theta)$。综上，$f(\boldsymbol{x},\theta)$ 可分解为 $g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x})$。

证明 (连续情形 $\Leftarrow$) 作一一对应的变换

$$\boldsymbol{X}=(X_1,\cdots,X_n)^T\longrightarrow(\boldsymbol{T},\boldsymbol{W})=(T_1,\cdots,T_k,W_1,\cdots,W_{n-k})^T$$

于是有

$$\begin{aligned}&X_i=X_i(T_1,\cdots,T_k,W_1,\cdots,W_{n-k}),\quad i=1,2,\cdots,n\\&T_j=T_j(X_1,\cdots,X_n),\quad j=1,2,\cdots,k\\&W_l=W_l(X_1,\cdots,X_n),\quad l=1,2,\cdots,n-k\end{aligned}$$

于是变换的 Jacobi 行列式的绝对值为

$$|J|=\left|\det\left(\frac{\partial(x_1,\cdots,x_n)}{\partial(t_1,\cdots,t_k,w_1,\cdots,w_{n-k})}\right)\right|\triangleq J(\boldsymbol{t},\boldsymbol{w})$$

由于需证明 $\Leftarrow$，故已知 $f(\boldsymbol{x};\theta)$ 可分解为 $f(\boldsymbol{x};\theta)=g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x}):=g(\boldsymbol{t},\theta)h(\boldsymbol{x})$，于是，我们可以得到 $\boldsymbol{T},\boldsymbol{W}$ 的联合密度分布为

$$f_{\boldsymbol{T},\boldsymbol{W}}(\boldsymbol{t},\boldsymbol{w})=g(\boldsymbol{t},\theta)h(\boldsymbol{x}(\boldsymbol{t},\boldsymbol{w}))\cdot J(\boldsymbol{t},\boldsymbol{w}):=\tilde{g}(\boldsymbol{t},\theta)\tilde{h}(\boldsymbol{t},\boldsymbol{w})$$

其中 $\tilde{g}(\boldsymbol{t},\theta)=g(\boldsymbol{t},\theta)$ 而 $\tilde{h}(\boldsymbol{t},\boldsymbol{w})=h(\boldsymbol{x}(\boldsymbol{t},\boldsymbol{w}))\cdot J(\boldsymbol{t},\boldsymbol{w})$。于是有在 $\boldsymbol{T}$ 下的条件概率为

$$f_{\boldsymbol{W}|\boldsymbol{T}}(\boldsymbol{w}|\boldsymbol{t})=\frac{f_{\boldsymbol{T},\boldsymbol{W}}(\boldsymbol{t},\boldsymbol{w})}{f_{\boldsymbol{T}}(\boldsymbol{t})}=\frac{f_{\boldsymbol{T},\boldsymbol{W}}(\boldsymbol{t},\boldsymbol{w})}{\int_{\boldsymbol{y}}f_{\boldsymbol{T},\boldsymbol{W}}(\boldsymbol{t},\boldsymbol{y})\,d\boldsymbol{y}}=\frac{\tilde{g}(\boldsymbol{t},\theta)\tilde{h}(\boldsymbol{t},\boldsymbol{w})}{\int_{\boldsymbol{y}}\tilde{g}(\boldsymbol{t},\theta)\tilde{h}(\boldsymbol{t},\boldsymbol{y})\,d\boldsymbol{y}}=\frac{\tilde{h}(\boldsymbol{t},\boldsymbol{w})}{\int_{\boldsymbol{y}}\tilde{h}(\boldsymbol{t},\boldsymbol{y})\,d\boldsymbol{y}}$$

与 θ 无关，所以 $\boldsymbol{T}$ 为充分统计量。

证明 (连续情形 $\Rightarrow$) 已知 $\boldsymbol{T}$ 为充分统计量，则在 $\boldsymbol{T}=\boldsymbol{t}$ 条件下的概率函数与 θ 无关。由之前的推导，我们知 $\boldsymbol{T},\boldsymbol{W}$ 的联合密度分布为

$$f_{\boldsymbol{T},\boldsymbol{W}}(\boldsymbol{t},\boldsymbol{w})=f_{\boldsymbol{W}|\boldsymbol{T}}(\boldsymbol{w}|\boldsymbol{t})f_{\boldsymbol{T}}(\boldsymbol{t})\triangleq r(\boldsymbol{t},\boldsymbol{w})g(\boldsymbol{T}(\boldsymbol{x}),\theta)$$

其中 $r(\boldsymbol{t},\boldsymbol{w})=f_{\boldsymbol{W}|\boldsymbol{T}}(\boldsymbol{w}|\boldsymbol{t})$ 由 $\boldsymbol{T}$ 为充分统计量，故 $r(\boldsymbol{t},\boldsymbol{w})$ 与 θ 无关，而 $g(\boldsymbol{T}(\boldsymbol{x}),\theta)=f_{\boldsymbol{T}}(\boldsymbol{t})$。

根据变换 $T_j=T_j(X_1,\cdots,X_n)$ 和 $W_l=W_l(X_1,\cdots,X_n)$ 我们同样有变换的 Jacobi 行列式的绝对值为

$$|\tilde{J}|=\left|\det\left(\frac{\partial(t_1,\cdots,t_k,w_1,\cdots,w_{n-k})}{\partial(x_1,\cdots,x_n)}\right)\right|\triangleq\tilde{J}(\boldsymbol{x})$$

于是 $\boldsymbol{X}$ 的概率密度函数为

$$f(\boldsymbol{x},\theta)=f_{\boldsymbol{T},\boldsymbol{W}}(\boldsymbol{t}(\boldsymbol{x}),\boldsymbol{w}(\boldsymbol{x}))\cdot\tilde{J}(\boldsymbol{x})=r(\boldsymbol{t}(\boldsymbol{x}),\boldsymbol{w}(\boldsymbol{x}))g(\boldsymbol{T}(\boldsymbol{x}),\theta)\cdot\tilde{J}(\boldsymbol{x}):=g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x})$$

其中 $h(\boldsymbol{x})=r(\boldsymbol{t}(\boldsymbol{x}),\boldsymbol{w}(\boldsymbol{x}))\tilde{J}(\boldsymbol{x})$，即 $f(\boldsymbol{x},\theta)$ 可分解为 $g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x})$。命题证毕。

推论 9 设 $T=T(\boldsymbol{x})$ 是 θ 的充分统计量。$S=\varphi(T)$ 为单值可逆函数，则 S 也是 θ 的充分统计量。

证明 (充分统计量定义) 由于 $S=\varphi(T)$ 为单值可逆函数，故

$$\{\boldsymbol{X}:T(\boldsymbol{X})=t_0\}=\{\boldsymbol{X}:\varphi^{-1}(S(\boldsymbol{X}))=t_0\}=\{\boldsymbol{X}:S(\boldsymbol{X})=\varphi(t_0):=s_0\}$$

是相同事件，于是对于任何事件 A 有

$$P(A|T=t_0)=P(A|S=s_0)$$

而 T 是充分统计量，故 $P(A|S=s_0)$ 与 θ 无关，根据充分统计量的定义 7.1 知 S 也是充分统计量。

证明 (因子分解定理) 因为 $T=T(\boldsymbol{X})$ 为 θ 的充分统计量，所以概率函数 f 可分解

$$f(\boldsymbol{x},\theta)=g(T(\boldsymbol{x}),\theta)h(\boldsymbol{x})$$

又因为 φ 为单值可逆函数，故存在反函数 φ^{-1}，由 $S(\boldsymbol{X})=\varphi(T(\boldsymbol{X}))$ 有

$$T(\boldsymbol{x})=\varphi^{-1}(S(\boldsymbol{x}))$$

代入 $f(\boldsymbol{x},\theta)$ 有关系式

$$f(\boldsymbol{x},\theta)=g(\varphi^{-1}(S(\boldsymbol{x})),\theta)h(\boldsymbol{x}):=\tilde{g}(S(\boldsymbol{x}),\theta)h(\boldsymbol{x})$$

所以，由因子分解定理 7.1 知，$S(\boldsymbol{X})=\varphi(T(\boldsymbol{X}))$ 也是 θ 的充分统计量。

例 7.3 设 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 为从正态总体 $N(a,\sigma^2)$ 中抽取的简单随机样本，令 $\boldsymbol{\theta}=(a,\sigma^2)$，则 $(\bar{X},S^2)$ 为 $\boldsymbol{\theta}$ 的充分统计量。

证明 样本 $\boldsymbol{X}$ 的联合密度分布为

$$\begin{aligned}
f(\boldsymbol{x},\boldsymbol{\theta})&=(2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2\right\}\\
&=(2\pi)^{-n/2}\sigma^{-n}\exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n}x_i^2-2a\sum_{i=1}^{n}x_i+na^2\right)\right\}\\
&=g(\boldsymbol{T}(\boldsymbol{x}),\boldsymbol{\theta})h(\boldsymbol{x})
\end{aligned}$$

其中 $h(\boldsymbol{x})\equiv 1$，由因子分解定理 7.1 可知 $\boldsymbol{T}=\boldsymbol{T}(\boldsymbol{X})=(\sum_{i=1}^{n}X_i,\sum_{i=1}^{n}X_i^2)$ 为充分统计量。又因为存在变换

$$\begin{pmatrix}\bar{X}\\ S^2\end{pmatrix}=\begin{pmatrix}\frac{1}{n} & 0\\ -\frac{1}{n(n-1)} & \frac{1}{n-1}\end{pmatrix}\begin{pmatrix}\sum_{i=1}^{n}X_i\\ \sum_{i=1}^{n}X_i^2\end{pmatrix}\triangleq A\boldsymbol{T}$$

由于 $|A|\neq 0$ 可逆，故是一一变换。由推论 9 可知 $(\bar{X},S^2)$ 也是 $\boldsymbol{\theta}$ 的充分统计量。

7.3 极小充分统计量

定义 7.2 设 $T=T(\boldsymbol{X})$ 是分布族 $\mathscr{F}$ 的充分统计量，若对 $\mathcal{F}$ 的任一充分统计量 $S=S(\boldsymbol{X})$，存在一个函数 $q_S(\cdot)$ 使得

$$T(\boldsymbol{X})=q_S(S(\boldsymbol{X}))$$

则称 $T(\boldsymbol{X})$ 是此分布族的**极小充分统计量**。

例 7.4 设 $\boldsymbol{X}=(X_1,\cdots,X_n)^T$ 为从正态总体 $N(\mu,\sigma^2)$ 中抽取的简单随机样本，则

$$\boldsymbol{T}=\left(\sum\nolimits_{i=1}^n X_i,\sum\nolimits_{i=1}^n X_i^2\right)$$

为 $\boldsymbol{\theta}=(\mu,\sigma^2)$ 的极小充分统计量。

证明 对任一充分统计量 $S=S(\boldsymbol{X})$，由因子分解定理 7.1 可分解联合密度函数为

$$f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{\theta})=(2\pi\sigma^2)^{-n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right\}=g(S(\boldsymbol{x}),\boldsymbol{\theta})h(\boldsymbol{x}) \tag{21}$$

其中 $h(\boldsymbol{x})=1$。上式对任意 $\boldsymbol{\theta}$ 都成立。

于是，我们取 $\boldsymbol{\theta}_1=(0,1)$，$\boldsymbol{\theta}_2=(1,1)$ 于是有

$$\frac{f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{\theta}_1)}{f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{\theta}_2)}=\frac{(2\pi)^{-n/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^n x_i^2\right\}}{(2\pi)^{-n/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^n(x_i-1)^2\right\}}:=G(\sum\nolimits_{i=1}^n X_i)$$

而由式 21 可知

$$\frac{f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{\theta}_1)}{f_{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{\theta}_2)}=\frac{g(S(\boldsymbol{x}),\boldsymbol{\theta}_1)h(\boldsymbol{x})}{g(S(\boldsymbol{x}),\boldsymbol{\theta}_2)h(\boldsymbol{x})}:=H(S(\boldsymbol{X}))$$

所以

$$G(\sum\nolimits_{i=1}^n X_i)=H(S(\boldsymbol{X}))\Rightarrow\sum\nolimits_{i=1}^n X_i=G^{-1}(H(S(\boldsymbol{X})))\triangleq q_S(S(\boldsymbol{X}))$$

其中 $q_S(\cdot)=G^{-1}(H(\cdot))$。

类似地，取 $\boldsymbol{\theta}_3=(0,1)$，$\boldsymbol{\theta}_4=(0,2)$ 有

$$\sum\nolimits_{i=1}^n X_i^2=\tilde{G}^{-1}\left(\frac{g(S(\boldsymbol{x}),\boldsymbol{\theta}_3)h(\boldsymbol{x})}{g(S(\boldsymbol{x}),\boldsymbol{\theta}_4)h(\boldsymbol{x})}\right):=\tilde{G}^{-1}(\tilde{H}(S(\boldsymbol{X})))\triangleq\tilde{q}_S(S(\boldsymbol{X}))$$

其中 $\tilde{q}_S(\cdot)=\tilde{G}^{-1}(\tilde{H}(\cdot))$。

综上，根据极小充分统计量的定义 7.2 可知，$\boldsymbol{T}=(\sum_{i=1}^n X_i,\sum_{i=1}^n X_i^2)$ 为 $\boldsymbol{\theta}=(\mu,\sigma^2)$ 的极小充分统计量。

定理 7.2 (*) 设 T 为样本空间 $\mathscr{X}$ 上的充分统计量，则有以下两条陈述等价：

- T 为极小充分统计量。
- 对任一充分统计量 S 有 $S(x)=S(y)\Rightarrow T(x)=T(y)\quad x,y\in\mathscr{X}$

证明 (*) **(1)** 先证 $\Rightarrow$：已知 T 为极小充分统计量。对任一充分统计量 S 有

$$T(x)=q_S(S(x)),\quad \exists\, q_S(\cdot),\ \forall\, x\in\mathscr{X}$$

于是对 $\forall\, x,y\in\mathscr{X}$，若 $S(x)=S(y)$，则

$$T(x)=q_S(S(x))=q_S(S(y))=T(y)$$

即 $\Rightarrow$ 成立。

(2) 下面证 $\Leftarrow$：已知 $S(x)=S(y)\Rightarrow T(x)=T(y)$。样本空间为 $\mathscr{X}$，统计量为 $S:\mathscr{X}\to\mathscr{S}$ 其中 $\mathscr{S}:=\{S(x):x\in\mathscr{X}\}$。对任意 $s\in\mathscr{S}$ 定义原像

$$E_s=S^{-1}(\{s\})=\{x\in\mathscr{X}:S(x)=s\}$$

那么定义集族（集合的集合）$\mathcal{E}=\{E_s:s\in\mathscr{S}\}$，其中 $E_s\in\mathcal{S}$ 为集合。由于 S 为充分统计量，不平凡地有 $E_s\neq\phi$。于是我们构造选择映射 r 为

$$r:\mathscr{S}\to\mathscr{E}:=\cup_{s\in\mathscr{S}}E_s,\quad \text{s.t. } r(s)\in E_s,\quad \forall\, s\in\mathscr{S} \tag{22}$$

注意到 E_s 的定义，有对 $\forall\, x\in\cup_{s\in\mathscr{S}}E_s$，即 $\exists\, E_{s_0}$ 使得 $x\in E_{s_0}=\{x\in\mathscr{X}:S(x)=s_0\}\Rightarrow x\in\mathscr{X}$，所以 $\cup_{s\in\mathscr{S}}E_s\subseteq\mathscr{X}$。反过来，对 $\forall\, x\in\mathscr{X}$ 记 $s=S(x)$，则 $x\in E_s$，于是 $x\in\cup_{s\in\mathscr{S}}E_s$。所以 $\mathscr{X}\subseteq\cup_{s\in\mathscr{S}}E_s$。最终有 $\cup_{s\in\mathscr{S}}E_s=\mathscr{X}$。所以式 22 中映射 r 定义为

$$r:\mathscr{S}\to\mathscr{X},\quad \text{s.t. } r(s)\in E_s=\{x:S(x)=s\},\quad \forall\, s\in\mathscr{S}$$

即此时 $r(s),\ s\in\mathscr{S}$ 满足 $S(r(s))=s$。

回到原题，对任意 $s\in\mathscr{S}$，任意取 $x_s\in E_s$ 定义映射

$$g(s)=T(x_s)$$

即 $g:\mathscr{S}\to\mathscr{T}:=T(\mathscr{X})$。断言：$g$ 是定义良好的，即每一个 s 都有唯一的输出 $g(s)$。断言的证明：对于 $x_s,x_s'\in E_s$，有 $S(x_s)=S(x_s')=s$，根据条件有 $T(x_s)=T(x_s')$，即 $g(s)$ 的值唯一（即只要 s 的值确定了，对于任意 $x_s\in E_s$，都有 $T(x_s)=g(s)$ 唯一）。

那么，对于任一充分统计量 $S(x)$，有 $q_S(\cdot)=g(\cdot)$ 使得

$$T(x)=g(S(x)):=q_S(S(x))$$

即 T 为极小充分统计量。

定理 7.3 设 $X_1,\cdots,X_n\overset{\text{i.i.d.}}{\sim}f(x,\theta),\ \theta\in\Theta$，若下面等价条件成立：

$$\frac{f(x,\theta)}{f(y,\theta)}\ \text{不依赖于}\theta\Leftrightarrow T(x)=T(y)$$

则 T 为极小充分统计量。

证明 (*) 分 2 步证明，先证 T 为充分统计量，再证 T 极小。

(1) 先证 T 为充分统计量。设样本空间为 $\mathscr{X}$，统计量为 $T:\mathscr{X}\to\mathscr{T}$ 其中 $\mathscr{T}:=\{T(x):x\in\mathscr{X}\}$。对任意 $t\in\mathscr{T}$ 定义原像

$$F_t=T^{-1}(\{t\})=\{x\in\mathscr{X}:T(x)=t\}$$

那么定义集族（集合的集合）$\mathcal{F}=\{F_t:t\in\mathscr{T}\}$，其中 $F_t\in\mathcal{F}$ 为集合。由于 T 的定义，不平凡地有 $F_t\neq\phi$。于是我们构造选择映射 h 为

$$h:\mathscr{T}\to\mathscr{F}:=\cup_{t\in\mathscr{T}}F_t,\quad \text{s.t. } h(t)\in F_t,\quad \forall\, t\in\mathscr{T} \tag{23}$$

注意到 F_t 的定义，有对 $\forall\, x\in\cup_{t\in\mathscr{T}}F_t$，即 $\exists\, F_{t_0}$ 使得 $x\in F_{t_0}=\{x\in\mathscr{X}:T(x)=t_0\}\Rightarrow x\in\mathscr{X}$，所以 $\cup_{t\in\mathscr{T}}F_t\subseteq\mathscr{X}$。反过来，对 $\forall\, x\in\mathscr{X}$ 记 $t=T(x)$，则 $x\in F_t$，于是 $x\in\cup_{t\in\mathscr{T}}F_t$。所以 $\mathscr{X}\subseteq\cup_{t\in\mathscr{T}}F_t$。最终有 $\cup_{t\in\mathscr{T}}F_t=\mathscr{X}$。所以式 23 中映射 h 定义为

$$h:\mathscr{T}\to\mathscr{X},\quad \text{s.t. } h(t)\in F_t=\{x:T(x)=t\},\quad \forall\, t\in\mathscr{T}$$

即此时 $h(t),\, t\in\mathscr{T}$ 满足 $T(h(t))=t$。

回到原问题，取 $y=h(T(x))$，有 $T(y)=T(x)$ 根据题目条件知 $f(x,\theta)/f(y,\theta)$ 与 θ 无关，有

$$f(x,\theta)=\frac{f(x,\theta)}{f(y,\theta)}\cdot f(y,\theta)=\phi(x,h(T(x)))\cdot f(h(T(x)),\theta):=g(T(x),\theta)h(x)$$

其中 $g(T(x),\theta)=f(h(T(x)),\theta)$ 而 $h(x)=\phi(x,h(T(x)))$，由因子分解定理 7.1 可知，T 为充分统计量。

(2) 再证 T 为极小充分统计量。对任意 $S(x)$ 为充分统计量，可以分解 $f(x,\theta)=g(S(x),\theta)h(x)$，于是若有 $S(x)=S(y)$，有

$$\frac{f(x,\theta)}{f(y,\theta)}=\frac{g(S(x),\theta)h(x)}{g(S(y),\theta)h(x)}=\frac{h(x)}{h(y)}$$

与 θ 无关。那么根据条件有 $T(x)=T(y)$。即 $S(x)=S(y)\Rightarrow T(x)=T(y)$，其中 T,S 均为充分统计量。根据定理 7.2 可知 T 为极小充分统计量。

例 7.5 设 $X_1,\cdots,X_n\overset{\text{i.i.d.}}{\sim}U(0,\theta)$，试证：$T(\boldsymbol{X})=X_{(n)}$ 是极小充分统计量。其中 $X_{(n)}$ 代表第 n 个次序统计量。

证明 对任意 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 和 $\boldsymbol{Y}=(Y_1,\cdots,Y_n)$ 均是来自 $U(0,\theta)$ 的简单随机样本，则有它们的密度函数为

$$f_{\boldsymbol{X}}(\boldsymbol{x},\theta)=\frac{1}{\theta^n}\cdot\mathbb{I}(0<x_{(1)}\le x_{(n)}<\theta)$$
$$f_{\boldsymbol{Y}}(\boldsymbol{y},\theta)=\frac{1}{\theta^n}\cdot\mathbb{I}(0<y_{(1)}\le y_{(n)}<\theta)$$

于是有 $\frac{f_{\boldsymbol{X}}(\boldsymbol{x},\theta)}{f_{\boldsymbol{Y}}(\boldsymbol{y},\theta)}$ 与 θ 无关 $\Rightarrow X_{(n)}=Y_{(n)}$，否则取 θ 在 $X_{(n)}$ 和 $Y_{(n)}$ 之间。反过来，若 $X_{(n)}=Y_{(n)}$，则有 $\frac{f_{\boldsymbol{X}}(\boldsymbol{x},\theta)}{f_{\boldsymbol{Y}}(\boldsymbol{y},\theta)}$ 与 θ 无关。根据定理 7.3 可知，$T=X_{(n)}$ 为极小充分统计量。

定理 7.4 设样本空间 $\mathscr{X}$ 上的指数族为 $\mathscr{F}=\{f(x,\theta),\theta\in\Theta\}$，密度函数可写作

$$f(x,\theta)=C(\theta)\exp\left\{\sum_{i=1}^{k}Q_i(\theta)T_i(x)\right\}h(x)$$

若 $(T_1(x),T_2(x),\cdots,T_k(x))$ 以概率为 1 线性无关，$(1,Q_1(\theta),Q_2(\theta),\cdots,Q_k(\theta))$ 以概率为 1 线性无关，则 $(T_1(x),T_2(x),\cdots,T_k(x))$ 是 θ 的极小充分统计量。

证明 (*) 注意到

$$\frac{f(x,\theta)}{f(y,\theta)}=\exp\left\{\sum_{i=1}^{k}Q_i(\theta)\cdot(T_i(x)-T_i(y))\right\}\frac{h(x)}{h(y)}$$

(1) 若 $T(x)=T(y)$，又因为 $(T_1(x),T_2(x),\cdots,T_k(x))$ 以概率为 1 线性无关，故 $T(\cdot)\not\equiv c$ 常值。则有

$$\frac{f(x,\theta)}{f(y,\theta)}=\exp\left\{\sum_{i=1}^{k}Q_i(\theta)\cdot 0\right\}\frac{h(x)}{h(y)}=\frac{h(x)}{h(y)}$$

与 θ 无关。

(2) 若 $f(x,\theta)/f(y,\theta)$ 与 θ 无关，则有

$$\begin{aligned}
&\exp\left\{\sum_{i=1}^{k}Q_i(\theta)\cdot(T_i(x)-T_i(y))\right\}\frac{h(x)}{h(y)}=A(x,y)\\
\Rightarrow&\sum_{i=1}^{k}Q_i(\theta)\cdot(T_i(x)-T_i(y))+\log\frac{h(x)}{h(y)}=\log A(x,y)\\
\Rightarrow&\left(\log\frac{h(x)}{h(y)}-\log A(x,y)\right)\cdot 1+\sum_{i=1}^{k}Q_i(\theta)\cdot(T_i(x)-T_i(y))=0
\end{aligned}$$

因为 $(1,Q_1(\theta),Q_2(\theta),\cdots,Q_k(\theta))$ 以概率为 1 线性无关，即在 θ 的非零测集上，有

$$\begin{cases}\log\dfrac{h(x)}{h(y)}-\log A(x,y)=0\\ T_i(x)-T_i(y)=0\end{cases}$$

即 $T(x)=T(y)$。

综上，根据定理 7.3 可知 $(T_1,\cdots,T_k)$ 为极小充分统计量。

例 7.6 设 $X_1,\cdots,X_n\overset{\text{i.i.d.}}{\sim}N(\mu,\sigma^2)$。试证：$T(X)=(\sum_{i=1}^n X_i,\sum_{i=1}^n X_i^2)$ 为极小充分统计量。

证明 (*) 联合概率密度函数可写作

$$f(\boldsymbol{x},\boldsymbol{\theta})=(2\pi\sigma^2)^{-\frac{n}{2}}e^{\frac{-n\mu^2}{2\sigma^2}}\exp\left\{\frac{\mu}{\sigma^2}\sum_{i=1}^{n}x_i+(-\frac{1}{2\sigma^2})\sum_{i=1}^{n}x_i^2\right\}$$

(1) 下面证明 $T_1=\sum_{i=1}^n X_i$，$T_2=\sum_{i=1}^n X_i^2$ 几乎处处线性无关。设 λ_1,λ_2 使得

$$\lambda_1\sum_{i=1}^{n}X_i+\lambda_2\sum_{i=1}^{n}X_i^2=0$$

对几乎处处 $\boldsymbol{X}=(X_1,\cdots,X_n)\in\mathbb{R}^n$。若 $\lambda_1,\lambda_2\neq 0$，则这个条件一定无法对整个 $\mathbb{R}^n$ 满足，所以对于 $\boldsymbol{X}\in\mathbb{R}^n$ 无解，即 $P(\lambda_1 T_1(\boldsymbol{X})+\lambda_2 T_2(\boldsymbol{X})=0)=0\neq 1$。故 $\lambda_1=\lambda_2=0$，即 T_1,T_2 几乎处处线性无关。

(2) 下面证明 $(1,Q_1,Q_2)$ 几乎处处线性无关。这里

$$Q_1(\boldsymbol{\theta})=\frac{\mu}{\sigma^2}\quad Q_2(\boldsymbol{\theta})=-\frac{1}{2\sigma^2}$$

设 a_0,a_1,a_2 满足

$$\begin{aligned}&a_0+a_1\cdot\frac{\mu}{\sigma^2}+a_2\cdot(-\frac{1}{2\sigma^2})=0\\ \Rightarrow\ &a_0\cdot 2\sigma^2+a_1\cdot 2\mu-a_2=0\end{aligned}$$

对任意的 $\mu\in\mathbb{R}$, $\sigma^2\in\{0\}\cup\mathbb{R}^+$ 几乎处处成立。固定 σ^2 令 μ 任意取值，则 $a_1=0$；固定 μ 令 σ^2 任意取值，则 $a_0=0$；所以最后 $a_0=a_1=a_2=0$，即 $(1,Q_1,Q_2)$ 几乎处处线性无关。

综上，根据定理 7.4 可知，$T(\boldsymbol{X})$ 是极小充分统计量。

8 完全统计量

8.1 定义和示例

定义 8.1 (完全统计量) 设 $\mathscr{F}=\{f(x,\theta):\theta\in\Theta\}$ 为一分布族。Θ 为参数空间。设 $T=T(X)$ 为一统计量，若对任何满足条件

$$\mathbb{E}_\theta\{\varphi(T(X))\}=0,\quad\forall\,\theta\in\Theta \tag{24}$$

的 $\varphi(T(X))$ 都有

$$P_\theta\{\varphi(T(X))=0\}=1,\quad\forall\,\theta\in\Theta \tag{25}$$

则称 $T(X)$ 为一**完全统计量**（complete statistic）。

由定义 8.1 可知，若 $T(X)$ 是完全统计量，则它的任一可测函数 $\delta(T)$ 也是完全统计量。

注 为简单计，设统计量 $T(X)$ 有密度函数 $g_\theta(t)$ 则式 24 可写作

$$\int\varphi(t)g_\theta(t)\,dt=0,\quad\forall\,\theta\in\Theta \tag{26}$$

即可视作“φ 与 g_θ 正交”。于是式 26 $\Rightarrow$ 式 25 可视作“若 φ 与函数系 $\{g_\theta:\theta\in\Theta\}$ 正交，则 φ 必为 0”。

例 8.1 设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 为从总体 $b(1,\theta)$ 中抽取的简单随机样本，则 $T(\boldsymbol{X})=\sum_{i=1}^n X_i$ 是完全统计量。

证明 显然 $T(\boldsymbol{X}) \sim b(n,\theta)$，故有

$$P(T(\boldsymbol{X})=k)=\binom{n}{k}\theta^k(1-\theta)^k,\quad k=0,1,\cdots,n$$

设 $\varphi(t)$ 为任一实函数，满足 $\mathbb{E}_\theta\{\varphi(T)\}=0$, $\theta\in(0,1)$ 此即

$$\begin{aligned}&\sum_{k=0}^{n}\varphi(k)\binom{n}{k}\theta^k(1-\theta)^k=0\\ \Leftrightarrow&\sum_{k=0}^{n}\varphi(k)\binom{n}{k}\left(\frac{\theta}{1-\theta}\right)^k=0\end{aligned}$$

令 $\delta=\theta/(1-\theta)$ 则上式等价于

$$\sum_{k=0}^{n}\left[\varphi(k)\binom{n}{k}\right]\delta^k=0,\quad\forall\,\delta\in\mathbb{R}^+$$

左边是关于 δ 的多项式，对 $\forall\,\delta$ 均为 0，则系数为 0，即 $\varphi(k)\binom{n}{k}=0$。即有 $\varphi(k)=0$, $\forall\,k=0,1,\cdots,n$，而 $\{k:k=0,1,\cdots,n\}$ 为 T 的支撑集。所以 $P_\theta(\varphi(T)=0)=1$ 即 $\varphi(T)=0$, a.s. P_θ。由完全统计量的定义 8.1 可知 $T(\boldsymbol{X})$ 是完全统计量。

例 8.2 设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是从均匀分布 $U(0,\theta)$ 中抽取的简单随机样本，则 $T(\boldsymbol{X})=X_{(n)}=\max\{X_1,\cdots,X_n\}$ 是完全统计量。

证明 由次序统计量可知 $T(\boldsymbol{X})=X_{(n)}$ 的概率密度函数为

$$g_\theta(t)=\frac{nt^{n-1}}{\theta^n}\cdot\mathbb{I}_{(0,\theta)}(t)$$

设 $\varphi(t)$ 为任一实函数，满足 $\mathbb{E}_\theta\{\varphi(T)\}=0$ 此即

$$\frac{n}{\theta^n}\int_0^\theta\varphi(t)t^{n-1}\,dt=0,\quad\forall\,\theta>0$$

去除不相关项后，两边同时对 θ 求导得

$$\varphi(\theta)\theta^{n-1}=0\quad\Leftrightarrow\quad\varphi(\theta)=0,\quad\forall\,\theta>0$$

即 $\varphi(T)=0$ 对任意 $T>0$，故由定义可知 $T(\boldsymbol{X})=X_{(n)}$ 为完全统计量。

例 8.3 设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是从 $N(\theta,\theta^2)$ 中抽取的简单随机样本，则 $\boldsymbol{T}(\boldsymbol{X})=(\bar{X},S^2)$ 不是完全统计量。

证明 由正态样本的性质可知

$$\mathbb{E}_\theta\{\bar{X}\}=\theta\quad\mathbb{E}_\theta\{S^2\}=\theta^2\quad\mathbb{E}_\theta\{\bar{X}^2\}=\mathbb{D}_\theta\{\bar{X}\}+(\mathbb{E}_\theta\{\bar{X}\})^2=(1+1/n)\theta^2$$

于是，构造函数 $\varphi(\boldsymbol{T})=\varphi(T_1,T_2)=T_1^2-(1+1/n)T_2$。由上式，代入计算有

$$\mathbb{E}_\theta\{\varphi(\boldsymbol{T}(\boldsymbol{X}))\}=\mathbb{E}_\theta\{\varphi(\bar{X},S^2)\}=\mathbb{E}_\theta\{\bar{X}^2\}-(1+1/n)\mathbb{E}_\theta\{S^2\}=0$$

但是

$$P_\theta(\varphi(\boldsymbol{T}(\boldsymbol{X}))=0)=P_\theta(\varphi(\bar{X},S^2))=P_\theta(\bar{X}^2=(1+1/n)S^2)\neq1$$

因为 $\bar{X}^2$ 不可能以概率为 1 等于 $(1+1/n)S^2$。所以由定义知，$\boldsymbol{T}(\boldsymbol{X})=(\bar{X},S^2)$ 不是完全统计量。

8.2 指数族中统计量的完全性

定理 8.1 设样本 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 的概率函数

$$f(\boldsymbol{x},\boldsymbol{\theta})=C(\theta)\exp\left\{\sum_{i=1}^{k}\theta_i T_i(\boldsymbol{x})\right\}h(\boldsymbol{x}),\quad \boldsymbol{\theta}=(\theta_1,\cdots,\theta_k)\in\Theta^*$$

为指数族的自然形式。令 $\boldsymbol{T}(\boldsymbol{X})=(T_1(\boldsymbol{X}),\cdots,T_k(\boldsymbol{X}))$，若自然参数空间 Θ^* 作为 $\mathbb{R}^k$ 的子集有内点，则 $\boldsymbol{T}(\boldsymbol{X})$ 是完全统计量。

注 指数族有内点 $\Rightarrow$ 完全统计量；但没有内点 $\not\Rightarrow$ 不是完全统计量。

定义 8.2 (内点) 设 $A\subseteq\mathbb{R}^k$，若 $\exists r>0$ 使得

$$B(\boldsymbol{x},r)=\{\boldsymbol{y}\in\mathbb{R}^k:||\boldsymbol{y}-\boldsymbol{x}||<r\}$$

有 $B(\boldsymbol{x},r)\subseteq A$，则称 $\boldsymbol{x}$ 为 A 中一个内点。

例 8.4 设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是从 $N(\mu,\sigma^2)$ 中抽取的简单随机样本，参数空间 $\Theta=\{\boldsymbol{\theta}=(a,\sigma^2):a\in\mathbb{R},\sigma^2\in\mathbb{R}^+\}$，则 $(\bar{X},S^2)$ 是完全统计量。

证明 由例 7.6 可知 $\boldsymbol{X}$ 的分布属于指数族，其自然形式为

$$f(\boldsymbol{x},\boldsymbol{\varphi})=C^*(\boldsymbol{\varphi})\exp\{\varphi_1T_1(\boldsymbol{x})+\varphi_2T_2(\boldsymbol{x})\}h(\boldsymbol{x})$$

其中 $h(\boldsymbol{x})\equiv 1$，$\varphi_1=a/\sigma^2$，$\varphi_2=-1/2\sigma^2$，$\boldsymbol{\varphi}=(\varphi_1,\varphi_2)$，于是自然参数空间为

$$\Theta^*=\{(\varphi_1,\varphi_2):\varphi_1\in\mathbb{R},\ \varphi_2<0\}$$

故 Θ^* 作为 $\mathbb{R}^2$ 的子集有内点，由定理 8.1 可知 $(T_1(\boldsymbol{X}),T_2(\boldsymbol{X}))=(\sum_{i=1}^n X_i,\sum_{i=1}^n X_i^2)$ 是完全统计量，而 $\boldsymbol{T}^*(\boldsymbol{X})=(\bar{X},S^2)$ 作为 $(T_1(\boldsymbol{X}),T_2(\boldsymbol{X}))$ 的函数也易证为完全统计量。

例 8.5 设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是从均匀分布 $U(\theta-1/2,\theta+1/2)$ 中抽取的简单随机样本，则 $\boldsymbol{T}(\boldsymbol{X})=(X_{(1)},X_{(n)})$ 是充分统计量，但不是完全统计量。

证明 容易得到 $\boldsymbol{X}$ 的联合密度函数为

$$f(\boldsymbol{x},\theta)=\mathbb{I}(\theta-1/2<x_{(1)}\leq x_{(n)}<\theta+1/2):=g(\boldsymbol{T}(\boldsymbol{x}),\theta)h(\boldsymbol{x})$$

由因子分解定理 7.1 可知 $\boldsymbol{T}(\boldsymbol{X})=(X_{(1)},X_{(n)})$ 是充分统计量。

证明 $\boldsymbol{T}(\boldsymbol{X})$ 不是完全统计量，只需构造函数 $\varphi(\boldsymbol{t})$ 使得 $\mathbb{E}_\theta\{\varphi(\boldsymbol{t})\}=0$ 但"$\varphi(\boldsymbol{t})=0$ a.s. P_θ 不成立"即可。

令 $Y_i=X_i-(\theta-1/2)$ 则有 $Y_i\overset{\text{i.i.d.}}{\sim}U(0,1)$ 与 θ 无关。所以有 $Z=X_{(n)}-X_{(1)}=Y_{(n)}-Y_{(1)}$ 的分布也与 θ 无关。故此时 Z 的分布确定，一定能找到常数 $a<b$ 使得

$$P(Z<a)=P(Z>b)>0$$

成立。于是，我们构造函数

$$\varphi(Z)=\mathbb{I}_{(-\infty,a)}(Z)-\mathbb{I}_{(b,+\infty)}(Z)$$

其中 $Z=X_{(n)}-X_{(1)}=T_2-T_1$。此时满足

$$\mathbb{E}_\theta\{\varphi(\boldsymbol{T})\}=\mathbb{E}_\theta\{\varphi(\boldsymbol{Z})\}=P(Z<a)-P(Z>b)=1-1=0$$

但 $\varphi(\cdot)\not\equiv 0$ 即 $\varphi(\boldsymbol{t})=0$ a.s. P_θ 不成立，所以 $\boldsymbol{T}(\boldsymbol{X})$ 不是完全统计量。

8.3 有界完全统计量及 Basu 定理

定义 8.3 (有界完全统计量) 若对任何满足

$$\mathbb{E}_\theta\{\varphi(T(X))\}=0,\quad \forall\,\theta\in\Theta$$

的有界（或 a.s. 有界）的函数 $\varphi(\cdot)$ 都有

$$P_\theta\{\varphi(T(X))=0\}=1,\quad \forall\,\theta\in\Theta$$

则称 $T(X)$ 为**有界完全统计量**。

由有界完全统计量的定义 8.3 可知，一个“完全统计量”必为“有界完全统计量”，反之不必对。

定理 8.2 (Basu 定理) 设 $\mathscr{F}=\{f(x,\theta):\theta\in\Theta\}$ 为一分布族，Θ 是参数空间。样本 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是从 $\mathscr{F}$ 中抽取的简单随机样本，设 $T(\boldsymbol{X})$ 是一有界完全统计量，且是充分统计量。若随机变量 $V(\boldsymbol{X})$ 的分布与 θ 无关，则对任意 $\theta\in\Theta$，$V(\boldsymbol{X})$ 和 $T(\boldsymbol{X})$ 相互独立。

推论 10 (指数族中的 Basu 定理) 设样本 $\boldsymbol{X}$ 的分布族为指数族，即

$$f(\boldsymbol{x},\boldsymbol{\theta})=C(\boldsymbol{\theta})\exp\left\{\sum_{j=1}^{k}\theta_jT_j(\boldsymbol{x})\right\}h(\boldsymbol{x})$$

而自然参数空间 Θ^* 作为 $\mathbb{R}^k$ 的子集有内点。若随机变量 $\boldsymbol{V}(\boldsymbol{X})$ 的分布与 $\boldsymbol{\theta}$ 无关，对任何 $\boldsymbol{\theta}$，则 $\boldsymbol{V}(\boldsymbol{X})$ 与 $\boldsymbol{T}(\boldsymbol{X})=(T_1(\boldsymbol{X}),\cdots,T_k(\boldsymbol{X}))$ 相互独立。

例 8.6 设 $\boldsymbol{X}=(X_1,\cdots,X_n)$ 是从正态总体 $\{N(0,\sigma^2):\sigma>0\}$ 中抽取的简单随机样本，则 $T(\boldsymbol{X})=\sum_{i=1}^n X_i^2$ 与随机变量 $\sum_{i=1}^n\lambda_iX_i^2/\sum_{i=1}^n X_i^2$ 独立，此处 $\lambda_1,\cdots,\lambda_n$ 为任意实数。

证明 由正态分布 $N(0,\sigma^2)$ 为指数族，且自然参数空间 $\Theta^*=\{\theta=-1/2\sigma^2:\theta<0\}$ 作为 $\mathbb{R}^1$ 的子集有内点，故易证 $T(\boldsymbol{X})=\sum_{i=1}^n X_i^2$ 是充分完全统计量。

记 $Y_i=X_i/\sigma$ 则 $Y_i\overset{\text{i.i.d.}}{\sim}N(0,1)$ 与 θ 无关，故

$$V(\boldsymbol{X})=\frac{\sum_{i=1}^n\lambda_iX_i^2}{\sum_{i=1}^n X_i^2}=\frac{\sum_{i=1}^n\lambda_iY_i^2}{\sum_{i=1}^n Y_i}$$

的分布也与 θ 无关。由 Basu 定理 8.2 可知 $T(\boldsymbol{X})$ 与 $V(\boldsymbol{X})$ 独立。