高等数理统计

张鑫航

国防科技大学

版本：1.0

更新：2023 年 12 月 24 日

1 基本概念

1.1 统计结构

例 1.1 对一物理量进行测量，其真值 μ 未知，测量值为 x，但测量有误差，故可认为

$$x = \mu + \varepsilon$$

$\{x_1, x_2 \cdots, x_n\}$ 是测量值。

可加上一个假设，进一步设 $\varepsilon \sim N(0, \sigma^2)$，这就建立了一个统计结构。

定义 1.1 (统计结构) 设 $(\mathscr{X}, \mathscr{B})$ 为一可测空间，$\mathscr{P}$ 为其上的一族概率分布，则成三元组 $(\mathscr{X}, \mathscr{B}, \mathscr{P})$ 为统计结构（模型）。若 $\mathscr{P}$ 仅依赖于某参数（向量）θ，即 $\mathscr{P} = \{\mathscr{P}_\theta : \theta \in \Theta\}$，就称为参数结构，否则成为非参数结构。

定义 1.2 (乘积结构与重复抽样结构) 设 $(\mathscr{X}, \mathscr{B}, \mathscr{P})$ 与 $(\mathscr{X}', \mathscr{B}', \mathscr{P}')$ 为两个统计结构，则称 $(\mathscr{X} \otimes \mathscr{X}', \mathscr{B} \otimes \mathscr{B}', \mathscr{P} \otimes \mathscr{P}')$ 为二者的乘积结构，记为 $(\mathscr{X}, \mathscr{B}, \mathscr{P}) \otimes (\mathscr{X}', \mathscr{B}', \mathscr{P}')$，特别的，$n$ 个同样的统计结构的乘积称为重复抽样结构，记为 $(\mathscr{X}, \mathscr{B}, \mathcal{P})^n$ 或 $(\mathscr{X}^n, \mathscr{B}^n, \mathcal{P}^n)$。

定义 1.3 (样本分布函数) 设 $(\mathscr{X}, \mathscr{B}, \mathscr{P})^n$ 为一重复抽样结构，$\forall (X_1, X_2, \cdots, X_n) \in X^n$，定义

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(X_i \leqslant x)}$$

为经验分布函数（样本分布函数）。

定义 1.4 (所控) 设 u 和 v 是可测空间上 $(\mathscr{X},\mathscr{B})$ 上的两个σ-有限测度。若 $N\in\mathscr{B}$，$u(N)=0\Rightarrow v(N)=0$，则称v 关于 u 绝对连续，或 v 被 u 所控，记为 $v<<u$。

定义 1.5 (可控结构) 设 $(\mathscr{X},\mathscr{B},\mathscr{P})$ 为一统计结构，若存在 u 为可测空间 $(\mathscr{X},\mathscr{B})$ 上的 $\sigma-$ 有限测度，使得 $\forall P\in\mathscr{P}$，$P<<u$，则称该结构为可控结构。进而 $p(x)=\dfrac{dP(x)}{du}$ 成为 p.r 密度。

1.2 常用分布族

1.2.1 Gamma 分布族

定义 1.6 (Gamma 分布族) 在 $(R^+,\mathscr{B}_{R^+})$ 上的密度函数形如

$$p(x;\alpha,\lambda)=\frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}I_{(0,+\infty)}(x),\ (\alpha>0,\lambda>0)$$

的分布称为参数为 α，λ 的Gamma 分布族，记为 $Ga(\alpha,\lambda)$。其中，$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}e^{-x}\,dx$ 为 Gamma 函数。

注 因为 $\int_0^{+\infty}\lambda^\alpha x^{\alpha-1}e^{-\lambda x}\,dx=\int_0^{+\infty}(\lambda x)^{\alpha-1}e^{-\lambda x}\,d(\lambda x)=\Gamma(\alpha)$，所以，

$$\int_{-\infty}^{+\infty}p(x;\alpha,\lambda)=\frac{\Gamma(\alpha)}{\Gamma(\alpha)}=1$$

注
- 由图 1，α 影响 $Ga(\alpha,\lambda)$ 的形状，λ 影响 $Ga(\alpha,\lambda)$ 的尺寸

- $\alpha \leqslant 1$ 时，严减；$1 < \alpha \leqslant 2$ 时，先上凸，后下凸；$\alpha > 2$ 时，先下凸，再上凸，最后下凸，两个拐点
- λ 影响密度函数的胖瘦

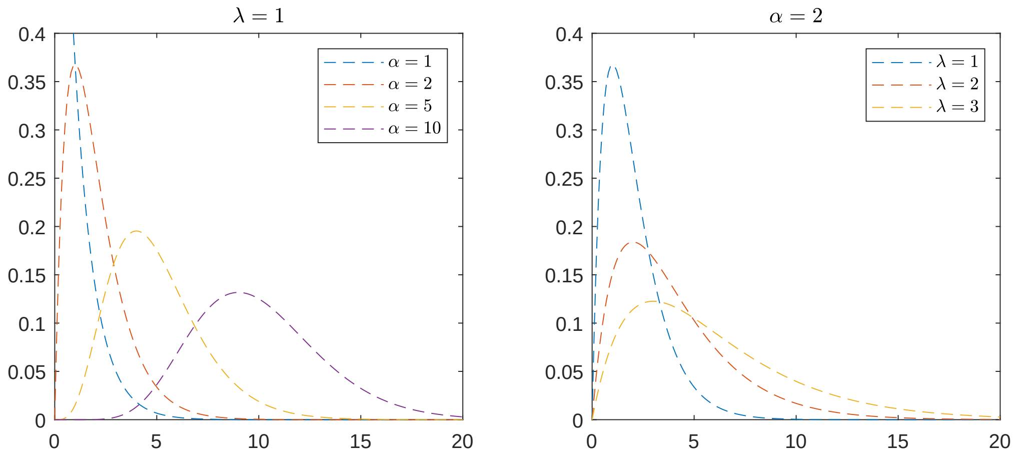

图 1: α 和 λ 对 $Ga(\alpha, \lambda)$ 的影响

设 $Z \sim Ga(\alpha, \lambda)$，则其 k 阶矩

$$
\begin{aligned}
EZ^k &= \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha+k-1} e^{-\lambda x} \, dx \\
&= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\lambda^k} \int_0^{+\infty} \frac{\lambda^{\alpha+k}}{\Gamma(\alpha+k)} x^{\alpha+k-1} e^{-\lambda x} \, dx \\
&= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\lambda^k} = \frac{(\alpha+k-1)(\alpha+k-2)\cdots\alpha}{\lambda^k}
\end{aligned}
$$

$Ga(\alpha,\lambda)$ 的特征函数

$$f(t) = Ee^{ixt} = \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda(1-\frac{it}{\lambda})x} dx$$
$$= (1-\frac{it}{\lambda})^{-\alpha} \int_0^{+\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} [(1-\frac{it}{\lambda})x]^{\alpha-1} e^{-\lambda(1-\frac{it}{\lambda})x} d(1-\frac{it}{\lambda})x$$
$$= (1-\frac{it}{\lambda})^{-\alpha}$$

于是，设 $Z_1, Z_2, \cdots, Z_n \sim Ga(\alpha,\lambda)$，且 Z_i 相互独立，则

$$Z_1 + Z_2 + \cdots + Z_n \sim Ga(\alpha_1 + \alpha_2 + \cdots + \alpha_n, \lambda)$$

两个特殊的 Gamma 分布

- 1) $Ga(1,\lambda) = Exp(\lambda),\ p(x,\lambda) = \lambda e^{-\lambda x}$
- 2) $Ga(\frac{n}{2},\frac{1}{2}) = \chi^2(n), \quad p(x,n) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}, x > 0$

注 $Z \sim \chi^2(n)$，则

$$EZ = \frac{n/2}{1/2} = n$$
$$EZ^2 = \frac{(n/2+1)(n/2)}{(1/2)^2} = n^2 + 2n$$
$$VarZ = EZ^2 - (EZ)^2 = 2n$$

例 1.2 电子产品的失效常常是由于外界的“冲击”引起。若在 $(0,t)$ 内发生冲击的次数 $N(t)$ 服从参数为 λt 的泊松分布，试证明第 n 冲击到来的时间服从伽马分布 $Ga(n,\lambda)$

证明.

$$\{S_n \leqslant t\} = \{N(t) \geqslant n\}$$

$$F_{S_n(t)} = P\{S_n \leqslant t\} = P\{N(t) \geqslant n\} = \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

$$F_{Ga(n,\lambda)}(t) = \int_0^t \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x}\, dx \overset{?}{=} 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t} = \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t}$$

1.2.2 Beta 分布族

定义 1.7 (Beta 分布) 设 $D = (0,1)$，定义在 $(D, \mathscr{B}_D)$，密度函数形如 $p(x; a, b) = \frac{1}{B(a,b)} x^{\alpha-1}(1-x)^{b-1} I_{(0,1)}(x)$，$(a > 0, b > 0)$ 的分布成为参数为 a，b 的 Beta 分布，记为 $Be(a,b)$。其中 $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\, dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

- $a > 1$ 和 $b > 1$，$p(x)$ 单峰状，在 $x = (a-1)/(a+b-2)$ 处达到最大值
- $a < 1$ 和 $b < 1$，$p(x)$U 形，在 $x = (a-1)/(a+b-2)$ 处达到最小值
- 当 $a = b + 1/2$ 时，Beta 分布为反正弦分布
- $a < 1$ 和 $b > 1$，$p(x)$ 严减
- $a > 1$ 和 $b < 1$，$p(x)$ 严增

注 设 $Z \sim Be(a,b)$，则 Z 的 k 阶矩

$$\begin{aligned} EZ^k &= \int_0^1 \frac{1}{B(a,b)} x^{a+k-1}(1-x)^{b-1}\ dx \\ &= \frac{B(a+k,b)}{B(a,b)} \int_0^1 \frac{1}{B(a+k,1)} x^{a+k-1}(1-x)^{b-1}\ dx \\ &= \frac{B(a+k,b)}{B(a,b)} = \frac{\Gamma(a+k)\Gamma(b)}{\Gamma(a+k+b)} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{\Gamma(a+k)\Gamma(a+b)}{\Gamma(a)\Gamma(a+k+b)} \\ &= \frac{(a+k-1)(a+k-2)\cdots(a)}{(a+b+k-1)(a+b+k-2)\cdots(a+b)} \end{aligned}$$

特别的，

$$\mathrm{E}\,Z = \frac{a}{a+b}, \quad \mathrm{E}\,Z^2 = \frac{(a+1)a}{(a+b+1)(a+b)}, \quad \mathrm{Var}\,Z = \mathrm{E}\,Z^2 - (\mathrm{E}\,Z)^2 = \left[(\frac{a+1}{a+b+1})^2 - 1 \right] (\frac{a}{a+b})^2$$

Beta 分布与 Gamma 分布的关系

设 $X_1 \sim \Gamma(\alpha_1, \lambda)$，$X_2 \sim \Gamma(\alpha, \lambda)$，且相互独立，则 $Y = \dfrac{X_1}{X_1 + X_2} \sim Be(\alpha_1, \alpha_2)$

证明. X_1 和 X_2 的联合分布为

$$p_{X_1,X_2}(x_1,x_2) = \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x_1^{\alpha_1-1} e^{-\lambda x_1} x_2^{\alpha_2-1} e^{-\lambda x_2}$$

令 $U = X_1$, $V = \dfrac{X_1}{X_1+X_2}$, 则 $\begin{cases} X_1 = U \\ X_2 = U/V - U \end{cases}$ ，且变换的行列式为

$$\begin{vmatrix} 1 & 0 \\ 1/v - 1 & -u/v^2 \end{vmatrix}$$

U, V 的联合分布为

$$\begin{aligned} p_{U,V}(u,v) &= p_{X_1,X_2}(u,v)|J| \\ &= \frac{\lambda^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} u^{\alpha_1-1} e^{-\lambda u} (\tfrac{u}{v} - u)^{\alpha_2-1} e^{-\lambda(u/v-u)} \tfrac{u}{v^2} \end{aligned}$$

则 V 的边缘分布为

$$p_V(v) = \int_0^\infty p_{U,V}(u,v)\, du = \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} v^{\alpha_1-1}(1-v)^{\alpha_2-1}$$

即 $Y_2 \sim Be(\alpha_1, \alpha_2)$

Beta-Binomial 共轭性

证明. 假定二项分布 $b(n,p)$ 的参数 p 服从 $Be(a,b)$ 的先验分布。然后又做了 $n_1 + n_2$ 次伯努利试验

（记为 W）成功 n_1 次，失败 n_2 次，于是后验分布

$$P(p|W)=\frac{P(p,W)}{P(W)}=\frac{P(W|p)P(p)}{\int_0^1 P(W|p)P(p)dp}$$

$$\frac{C_{n_1+n_2}^{n_1}p^{n_1}(1-p)^{n_2}\frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}}{\int_0^1 C_{n_1+n_2}^{n_1}p^{n_1}(1-p)^{n_2}\frac{1}{B(a,b)}p^{a-1}(1-p)^{b-1}\ dp}=\frac{p^{n_1+a-1}(1-p)^{n_2+b-1}}{\int_0^1 p^{n_1+a-1}(1-p)^{n_2+b-1}\ dp}$$

$$=\frac{p^{n_1+a-1}(1-p)^{n_2+b-1}}{B(n_1+a-1,n_2+b-1)}$$

p 的后验分布为 $Be(n_1+a,n_2+b)$，$Be(a,b)+\mathrm{BinomCount}(n_1,n_2)=Be(n_1+a,n_2+b)$。

1.2.3 Fisher 分布族

定义 1.8 (Fisher Z 分布) 定义在 $(R^+,\mathscr{B}_{R^+})$ 上的密度函数形如

$$p(x;a,b)=\frac{1}{B(a,b)}\frac{x^{a-1}}{(1+x)^{a+b}}I_{(0,+\infty)}(x),\ (a>0,b>0)$$

的分布承诺为参数为 a,b 的 Fisher Z 分布，记作 $Z(a,b)$。

设 $Z\sim Z(a,b)$，则 Z 的 k 阶矩

$$\begin{aligned}
EZ^k &= \int_0^{\infty}\frac{1}{B(a,b)}\frac{x^{a+k-1}}{(1+x)^{a+b}}\,dx\\
&= \frac{B(a+k,b-k)}{B(a+b)}\int_0^{\infty}\frac{1}{B(a+k,b-k)}\frac{x^{a+k-1}}{(1+x)^{a+b}}\,dx\\
&= \frac{B(a+k,b-k)}{B(a+b)}
\end{aligned}$$

特别的

$$EZ=\frac{a}{b-1},b>1\ ;\quad EZ^2=\frac{(a+1)a}{(b-1)(b-2)},b>2$$

Fisher 分布与 Beta 分布的关系

若 $Z\sim Be(a,b)$，则 $Y=\dfrac{Z}{1-Z}\sim Z(a,b)$

证明. 有 $Z=\dfrac{Y}{1+Y}$，那么，$|\frac{dz}{dy}|=(\frac{1}{1+y})^2$ 因此

$$\begin{aligned}
&f_Y(y)=f_Z(\tfrac{y}{1+y})|\tfrac{dz}{dy}|\\
&=\frac{1}{B(a,b)}(\frac{y}{1+y})^{a-1}(\frac{1}{1+y})^{b-1}(\frac{1}{1+y})^2\\
&=\frac{1}{B(a,b)}\frac{y^{a-1}}{(1+y)^{a+b}}
\end{aligned}$$

若 $Z \sim Z(a,b)$，则 $Y = \dfrac{Z}{1+Z} \sim Be(a,b)$

证明. 有 $Z = \dfrac{Y}{1-Y}$，那么，$|\frac{dz}{dy}| = (\frac{1}{1-y})^2$ 因此

$$
\begin{aligned}
f_Y(y) &= f_Z(\tfrac{y}{1-y})|\tfrac{dz}{dy}| \\
&= \frac{1}{B(a,b)} \frac{(\frac{y}{1-y})^{a-1}}{(1+\frac{y}{1-y})^{a+b}} (\frac{1}{1-y})^2 \\
&= \frac{1}{B(a,b)} y^{a-1}(1-y)^{b-1}
\end{aligned}
$$

Fisher 分布与 Gamma 分布的关系

设 $X_1 \sim \Gamma(\alpha_1, \lambda),\quad , X_2 \sim (\alpha_2, \lambda)$，且相互独立，则

$$Y = X_1/X_2 \sim Z(\alpha_1, \alpha_2)$$

证明. 设 $\begin{cases} U = X_1 \\ V = X_1/X_2 \end{cases}$，有 $\begin{cases} X_1 = U \\ X_2 = U/V \end{cases}$ 且变换的行列式为

$$J = \begin{vmatrix} 1 & 0 \\ 1/V & -U/V^2 \end{vmatrix}$$

U,V 的联合分布为

$$\begin{aligned} p_{U,V}(u,v) &= p_{X,Y}(u,v)|J| \\ &= \frac{\frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)}u^{\alpha_1-1}e^{-\lambda u}}{\frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)}(u/v)^{\alpha_2-1}e^{-\lambda u/v}}\frac{u}{v^2} \end{aligned}$$

V 的边缘分布为

$$\begin{aligned} p_V(v) &= \int_0^{+\infty} p_{U,V}(u,v)\ du \\ &= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\frac{v^{\alpha_1-1}}{(1+v)^{\alpha_1+\alpha_2}}\int_0^{+\infty}\frac{1}{\Gamma(\alpha_1+\alpha_2)}\left(\frac{\lambda(1+v)}{v}\right)^{\alpha_1+\alpha_2} u^{\alpha_1+\alpha_2-1}e^{-\lambda\frac{1+v}{v}u}\ du \\ &= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\frac{v^{\alpha_1-1}}{(1+v)^{\alpha_1+\alpha_2}} \end{aligned}$$

1.2.4 t 分布族

定义 1.9 (t 分布族) 在 $(R,\mathscr{B}_R)$ 上的密度函数形如

$$p(x;\alpha) = \frac{\Gamma(\frac{\alpha+1}{2})}{\sqrt{\alpha\pi}\Gamma(\frac{\alpha}{2})}(1+\frac{x^2}{\alpha})^{-\frac{\alpha+1}{2}}$$

的分布族称作自由度为 α 的 t 分布族，记为 $t(\alpha)$

注　1° 设 $X \sim t(\alpha)$，则由于其分布函数为偶函数，则 Z 的 k 阶矩为

$$\begin{aligned} E^{2k+1} &= 0, & \alpha > 2k+1 \\ E^{2k} &= \frac{\alpha^k}{\sqrt{\pi}} \frac{\Gamma(\frac{\alpha}{2}-k)\Gamma(k+\frac{1}{2})}{\Gamma(\frac{\alpha}{2})} & \alpha > 2k \end{aligned}$$

2° t 分布于标准正态分布的关系

$$\lim_{\alpha\to\infty} \frac{\Gamma(\frac{\alpha+1}{2})}{\sqrt{\alpha\pi}\Gamma(\frac{\alpha}{2})} = \frac{1}{2}$$
$$\lim_{\alpha\to\infty} (1+\frac{x^2}{\alpha})^{-\frac{\alpha+1}{2}} = \lim_{\alpha\to\infty} [(1+\frac{x^2}{\alpha})^{\frac{\alpha}{x^2}}]^{-\frac{x^2}{\alpha}\cdot\frac{\alpha+1}{2}} = e^{-\frac{x^2}{2}}$$

3° 考虑 Cauchy 分布 $X \sim t(1)$ 的 k 阶矩 $k \geqslant 1$

$$\begin{aligned} E\,|X|^k &= \int_{-\infty}^{\infty} |\,x\,|^k \cdot \frac{1}{\pi(1+x^2)} dx = \int_0^{\infty} x^{k-1} \cdot \frac{2x}{\pi(1+x^2)} dx \\ &= \int_0^{\infty} \frac{x^{k-1}}{\pi} d[\ln(1+x^2)] = \frac{x^{k-1}\ln(1+x^2)}{\pi}\Bigg|_0^{\infty} - \frac{k-1}{\pi}\int_0^{\infty} \ln(1+x^2)x^{k-2}dx \end{aligned}$$

$k \geqslant 2,\ \lim\limits_{x\to\infty} \dfrac{x^k}{\pi(1+x^2)} = \infty$，不存在。

1.2.5　多元正态分布族

1° 已知一元标准正态分布 $N(0,1)$ 的随机变量为 U，也即

$$U \sim \phi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2},\ u \in R$$

对于任意的 $\mu \in R, \sigma > 0$，易有

$$\begin{aligned} &X = \mu + \sigma U \sim p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &X \sim N(\mu, \sigma^2) \end{aligned}$$

2° 设 $U = (u_1, u_2, \cdots, u_n)$ 满足 $u_i \sim N(0,1), i = 1, 2, \cdots, n$，且相互独立，则

$$U \sim p(u) = \frac{1}{(\sqrt{2\pi})^2} e^{\frac{1}{2}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}}$$

且 $E(\boldsymbol{U}) = 0, \boldsymbol{U} = \boldsymbol{I}_n$，称$U$ 为 n 元标准正态分布，记为 $U \sim N_n(0, \boldsymbol{I}_n)$。

3° 设 $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_n)^{\mathrm{T}}$ 是常数向量，$\boldsymbol{\Sigma}$ 是一个 n 阶正定矩阵，则经特征值分解有 $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$，于是 $\boldsymbol{\Sigma}^{-1} = (\boldsymbol{A}^{\mathrm{T}})^{-1}\boldsymbol{A}^{-1}$，$|\boldsymbol{\Sigma}|^{-\frac{1}{2}} = |\boldsymbol{A}|^{-1}$。令 $\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{U}$ 则可以算出

$$\boldsymbol{X} \sim p(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi})^n} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})},\ \boldsymbol{X} \in \boldsymbol{R}^n$$

此时，$E(\boldsymbol{X}) = \boldsymbol{\mu}$，$\mathrm{Var}(\boldsymbol{X}) = \boldsymbol{\Sigma}$，称 $\boldsymbol{X}$ 服从多元正态分布，记为 $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$。

若 $\boldsymbol{A}$ 不是满秩的，定义多元正态分布$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{U}$

4° 考虑 $\boldsymbol{X}$ 的特征函数

$$f_{\boldsymbol{X}}(t) = Ee^{i\boldsymbol{t}^{\mathrm{T}}\boldsymbol{x}} = e^{i\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu}} Ee^{i\boldsymbol{t}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{U}}$$

令 $\boldsymbol{t}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{a}^{\mathrm{T}}=\begin{pmatrix}a_1 & a_2 & \cdots & a_n\end{pmatrix}$ 于是后验分布

$$\begin{aligned} f_{AU}(\boldsymbol{t}) &= Ee^{i\boldsymbol{t}^{\mathrm{T}}\boldsymbol{AU}}=\prod_{i=1}^{n}Ee^{ia_iU_i}=\prod_{i=1}^{n}f_{U_i}(a_i)=\prod_{i=1}^{n}e^{-\frac{1}{2}a_i^2} \\ &= e^{-\frac{1}{2}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}=e^{-\frac{1}{2}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{t}}=e^{-\frac{1}{2}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{t}} \end{aligned}$$

所以 $f_X(t)=\exp(i\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu}-\frac{1}{2}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{t})$，与 $\boldsymbol{A}$ 满秩时一致。

定义 1.10 (n 元正态分布) 设 $\boldsymbol{X}=\begin{pmatrix}X_1 & X_2 & \cdots & X_n\end{pmatrix}$ 是一个 n 维随机向量，且 $E\boldsymbol{X}=\boldsymbol{\mu}$, $\mathrm{Var}\,\boldsymbol{X}=\boldsymbol{\Sigma}$ (非负定)，若其特征函数为 $f_{\boldsymbol{X}}(t)=\exp(i\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu}-\frac{1}{2}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{t})$，则称 $\boldsymbol{X}$ 为 n 元正态分布，记为 $X\sim N_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$，而矩阵 $\boldsymbol{\Sigma}$ 的秩 $\mathrm{rank}(\boldsymbol{\Sigma})=r$ 称为这个分布的秩。

注 若 $\mathrm{rank}(\boldsymbol{\Sigma})=n$，$\boldsymbol{\Sigma}^{-1}$ 存在，$\boldsymbol{X}$ 具有非奇的 n 元正态分布，密度函数为

$$\boldsymbol{X}\sim p(\boldsymbol{x})=\frac{1}{(\sqrt{2\pi})^n}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})},\quad \boldsymbol{X}\in\boldsymbol{R}^n$$

若 $\mathrm{rank}(\boldsymbol{\Sigma})=r<n$，$\Sigma^{-1}$ 不存在，则其密度函数形式又该如何？

定义 1.11 (多元正态分布) 设 $\boldsymbol{X}=\begin{pmatrix}X_1 & X_2 & \cdots & X_n\end{pmatrix}^{\mathrm{T}}$ 是一个 n 维随机向量，若 $\boldsymbol{X}$ 与 $\boldsymbol{\mu}+\boldsymbol{BU}$ 具有相同的分布，其中 $\boldsymbol{\mu}$ 为 n 维向量，$\boldsymbol{B}$ 是一个秩为 r 的 $n\times r$ 阶矩阵，$\boldsymbol{U}\sim N_r(0,\boldsymbol{I}_r)$，那么称 $\boldsymbol{X}\sim N_n(\boldsymbol{\mu},\boldsymbol{B}\boldsymbol{B}^{\mathrm{T}})$

1.3 统计量及其分布

1.3.1 统计量的概念

定义 1.12 (统计量) 设 $(\mathscr{X},\mathscr{B},\mathscr{P})$ 是一个统计结构，$(\mathscr{T},\mathscr{C})$ 是一个可测空间，若 $T:\mathscr{X}\to\mathscr{T}$ 是一个可测映射，且与 P 无关，则称 T 是此结构的统计量。

对定义的理解：

注 1° T 是可测映射，即 σ 代数 C 中任一元素（集合）C 的原像 $T^{-1}(C)=\{x:T(x)\in C\}$ 是 σ 代数 B 中的元素（集合）。

2° T 与 P 无关，即不含未知参数

3° T 可以是向量，$T(X)=\begin{pmatrix}T(X_1) & T(X_2) & \cdots & T(X_k)\end{pmatrix}$，也可以是一维的。统计量 T 的值域 T 一般常用 R 或者 R^k。

1.3.2 统计量的分布（抽样分布）

定义 1.13 (抽样分布) 统计量的概率分布，称为抽样分布，也成为诱导分布。

设 $T:(\boldsymbol{X},\boldsymbol{B})\to(\boldsymbol{T},\boldsymbol{C})$，对任意 $C\in\boldsymbol{C}$，概率

$$P^{\mathrm{T}}=P(T(x)\in C)=\int_{x:T(x)\in C} dP=\int_{T^{-1}(C)} dP=P(T^{-1}(C))$$

其中 $P\in\boldsymbol{P}$，称 P^{T} 是 P 的诱导测度，$P^{\mathrm{T}}=\{P^{\mathrm{T}}:P\in\boldsymbol{P}\}$ 成为诱导分布族，而 (Y,C,P^{T}) 称为 T 的诱导结构。

注 若 $(\boldsymbol{X},\boldsymbol{B},\boldsymbol{P})$ 是可控结构，则诱导结构 $(\boldsymbol{T},\boldsymbol{C},\boldsymbol{P}^{\mathrm{T}})$ 也是可控结构。

证明. 因为 $(\boldsymbol{X},\boldsymbol{B},\boldsymbol{P})$ 是可控结构，则存在 σ 有限测度 μ，使 $P << \mu \quad (\forall P \ in \boldsymbol{P})$，令 $\mu^{\mathrm{T}}(C) = \mu(T^{-1}(C))$，$\forall C \in \boldsymbol{C}$。

若有 $\boldsymbol{\mu}^{\mathrm{T}}$ 的零测集 N，$\mu^{\mathrm{T}}(C) = \mu(T^{-1}(C)) = 0$，因为 $\forall P \in \boldsymbol{P}, \boldsymbol{P} << \mu$，则 $P^{\mathrm{T}}(N) = P(T^{-1}(N)) = 0$，从而 $P^{\mathrm{T}} << \mu^{\mathrm{T}}$

1° 设 $T = T(X_1, X_2, \cdots, X_n)$ 且为可微函数，其梯度的模为正，即

$$\|\mathrm{grad}\ \ T(x_1, x_2 \cdots, x_n)\| = \sqrt{\sum_{i=1}^{n}\left[\frac{\partial}{\partial x_i}T(x_1, x_2 \cdots, x_n)\right]^2} > 0$$

则此时 T 的分布函数为

$$F_T(t) = P\{T(x_1, \cdots, x_n) \leqslant t\} = \int_D^{n\text{重}} \cdots \int p(x_1, x_2 \cdots, x_n) dx_1 \cdots dx_n$$

其中 $D = \{(x_1, x_2 \cdots, x_n) : T(x_1, x_2 \cdots, x_n) \leqslant t\}$。可以计算出其密度函数为

$$p_T(t) = \int_{S_{n-1}}^{n-1\text{重}} \cdots \int p(x_1, x_2 \cdots, x_n) \frac{dS_{n-1}}{\|\mathrm{grad}\ \ T(x_1, x_2 \cdots, x_n)\|}$$

其中积分域是由方程 $T(x_1, x_2 \cdots, x_n) = t$ 所决定的 $n-1$ 维曲面 S_{n-1}

2° T 是 k 维统计量 $(k < n)$

$$p_T(t) = \int_{S_{n-k}}^{n-k\text{重}} \cdots \int \frac{p(x_1,\cdots,x_n)}{\left(\sum\limits_{i_1<\cdots<i_n}\left[\dfrac{D(T_1,\cdots,T_k)}{D(x_{i_1},\cdots,x_{i_k})}\right]^2\right)^{\frac{1}{2}}} dS_{n-k}$$

其中积分域是由 k 个方程 $T_j(x_1,\cdots,x_n) = t_j$，$j = 1,\cdots,k$ 所决定的 $n-k$ 维曲面 S_{n-k}，而 $\dfrac{D(T_1,\cdots,xT_k)}{D(x_{i_1},\cdots,x_{i_k})} = \left|\begin{pmatrix} \dfrac{\partial T_1}{\partial x_{i_1}} & \cdots & \dfrac{\partial T_1}{\partial x_{i_k}} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial T_k}{\partial x_{i_1}} & \cdots & \dfrac{\partial T_k}{\partial x_{i_k}} \end{pmatrix}\right|$ 是函数 $T_1,\cdots,T_k$ 对变量 $x_{i_1},\cdots,x_{i_k}$ 的雅可比行列式。

1.3.3 次序统计量及其分布

定义 1.14 (次序统计量) 设 $X_1,\cdots,X_n$ 是来自某总体的一个样本，将其按从小到大的次序排列成 $X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$，称 $\begin{pmatrix} X_{(1)} & X_{(2)} & \cdots & X_{(n)} \end{pmatrix}$ 为该样本的次序统计量。$X_{(1)}$ 称为该样本的最小次序统计量，$X_{(n)}$ 称为样本的最大次序统计量。

“概率元方法”（p.r.）元法的引入

把 $X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$ 的观察值记为 $y_1 \leqslant y_2 \leqslant \cdots \leqslant y_n$，设总体 X 的密度函数为 $p(x)$，则连续随机变量落在很小区间 $(x, x+dx)$ 的概率为

$$P(x < X < x + dx) = p(x)dx + x(dx)$$

$p(x)dx$ 称为 X 的概率元；反正，若存在函数 $p(x)$ 使上式成立，则 $p(x)$ 就是 X 的密度函数。

次序统计量的分布

区间 $(-\infty, y_k)$，$[y_k, y_k + dy_k)$，$[y_k, +\infty)$

$X_{(k)}$ 的概率函数 $g(y_k)$，其中 $1 \leqslant k \leqslant n$

$$g(y_k)dy_k = \frac{n!}{(k-1)!(n-k)!}[F(y_k)]^{k-1}p(y_k)dy_k[1 - F(y_k + dy_k)]^{n-k}$$

两边约去 dy_k 后，再让 $dy_k \to 0$

$$g(y_k) = \frac{n!}{(k-1)!(n-k)!}[F(y_k)]^{k-1}[1 - F(y_k)]^{n-k}p(y_k)$$

最小次序统计量　$g(y_1) = n[1 - F(y_1)]^{n-1}p(y_1)$

最大次序统计量　$g(y_n) = n[F(y_n)]^{n-1}p(y_n)$

例 1.3 【猜数游戏】一个魔盒上面有一个按钮，每按下按钮，就均匀地输出一个 $[0, 1]$ 之间的随机数，甲按 10 下得到 10 个数，要乙猜第 7 大的数是什么，偏离不超过 0.01 就算对。乙应该怎么猜呢?

- $X_1, X_2, \cdots, X_n \overset{\text{iid}}{\sim} U(0, 1)$
- 把这 n 个随机变量排序后得到顺序统计量 $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$
- $F(y) = y$，$p(y) = 1$

$X_{(7)}$ 的分布 $g(y_7)$ 为

$$\begin{aligned} g(y_7) &= \frac{10!}{(7-1)!(10-7)!} y_7^{7-1}(1-y_7)^{10-7} \times 1 \\ &= \frac{10!}{6!3!} y_7^6 (1-y_7)^3 \end{aligned}$$

$X_{(7)} \sim Be(7,4)$，在 $y = \dfrac{7-1}{7+4-2} = \dfrac{2}{3}$ 时，概率最大。

例 1.4 设 $X_1, \cdots, X_n$ 是来自均匀分布 $U(0,1)$ 的一个样本，要求该样本第 k 个次序统计量 $X_(k)$ 的分布与期望

$$g(y_k) = \frac{n!}{(k-1)!(n-k)!} y_k^{k-1}(1-y_k)^{n-k},$$

可以知道 $X_{(k)} \sim Be(k, n-k+1)$，期望为 $E(X_{(k)}) = \dfrac{k}{n+1}$

$X_{(k)}$ 与 $X_{(j)}$ 的联合密度函数 $g(y_k, y_j)$，其中 $1 \leqslant k < j \leqslant n$

区间 $(-\infty, y_k)$，$[y_k, y_k + dy_k)$，$[y_k + dy_k, y_j)$，$[y_j, y_j + dy_j)$，$[y_j + dy_j, +\infty)$

$$\begin{aligned} g(y_k, y_j) dy_k dy_j &= \frac{n!}{(k-1)!(j-1-k)(n-j)!} [F(y_k)]^{k-1} p(y_k) dy_k \times \\ & [F(y_j) - F(y_k + dy_k)]^{j-1-k} p(y_j) dy_j [1 - F(y_j + dy_j)]^{n-j} \end{aligned}$$

两边约去 dy_k，dy_j 后，再让 $dy_k \to 0$，$dy_j \to 0$

$$\begin{aligned} g(y_k, y_j) &= \frac{n!}{(k-1)!(j-1-k)!(n-j)!}[F(y_k)]^{k-1} \times \\ & [F(y_j) - F(y_k)]^{j-1-k}[1 - F(y_j)]^{n-j} p(y_k) p(y_j) \end{aligned}$$

最小次序统计量与最大次序统计量的联合密度

$$g(y_1, y_n) = n(n-1)[F(y_n) - F(y_1)]^{n-2} p(y_1) p(y_n),\ y_1 \leqslant y_n$$

前 r 个次序统计量 $X_{(1)}, X_{(2)}, \cdots, X_{(r)}$ 的联合密度函数

$$g(y_1, \cdots, y_r) = \frac{n!}{(n-r)!}[1 - F(y_r)]^{n-r} \cdot p(y_1) \cdots p(y_2), \quad y_1 \leqslant \cdots y_n$$

次序统计量的矩的存在性问题

定理 1.1 设 $X_1, \cdots, X_n$ 是来自于某总体 X 的一个样本，且对某个 $a > 0$ 有 $E|X|^a < \infty$。则当 n, k 和 r 满足

$$r \leqslant a \cdot \min(k, n-k+1)$$

有 $E|X_{(k)}|^r < \infty$，其中 $X_{(k)}$ 为该样本的第 k 个次序统计量。

证明. 要证明 $E|X_{(k)}|^r < \infty$，根据定义，会用到 $\int_{-\infty}^{0} |x|^r dG_k(x) + \int_0^{\infty} x^r dG_k(x)$ 的值必须有界。而 $G_k(x)$ 与总体分布 $F(x)$ 有关，故要转化

证明分三步:

1° 证明：$|x|^a[1-F(x)]$ 有界

2° 证明：$|x|^a[F(x)]$ 有界

3° $X_{(k)}$ 代入计算 $E|X_{(k)}|^r$ 积分的值

$$\begin{aligned}\infty > E|X|^a &= \int_{-\infty}^{+\infty} |x|^a \, dP\{|X| \leqslant x\} \geqslant \int_{|x|}^{+\infty} |x|^a \, dP\{|X| \leqslant x\} \\ &\geqslant |x|^a \int_{|x|}^{+\infty} \, dP\{|X| \leqslant t\} = |x|^\alpha P\{|X| \geqslant t\} \\ &= |x|^a\{[1-F(x)] + [F(-x)]\}\end{aligned}$$

对于

$$\int_{|x|}^{+\infty} |x|^a \, dP\{|X| \leqslant x\} \geqslant |x|^a\{[1-F(x)] + [F(-x)]\}$$

$=$ 两边让 $x \to +\infty$ 有

$$0 = \lim_{x\to\infty} \int_{|x|}^{+\infty} |x|^a \, dP\{|X| \leqslant x\} \geqslant \lim_{x\to+\infty} |x|^a[1-F(x)] + \lim_{x\to+\infty} |x|^a[F(-x)] \geqslant 0$$

综上所述 $\lim\limits_{x\to+\infty} |x|^a[1-F(x)] = 0,\ \lim\limits_{x\to+\infty} |x|^a[F(-x)] = 0$

$$\begin{aligned} E|X_{(k)}|^r &= \int_{-\infty}^{0} |x|^r \frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1}[1-F(x)]^{n-k}p(x)\ dx + \\ & \int_{0}^{+\infty} |x|^r \frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1}[1-F(x)]^{n-k}p(x)\ dx \\ &= \frac{n!}{(k-1)!(n-k)!}[(1)+(2)] \end{aligned}$$

考虑 $(1) < \infty$

$$\begin{aligned} (1) &= \int_{-\infty}^{0} |x|^r [F(x)]^{k-1}[1-F(x)]^{n-k}p(x)\ dx \\ &= \int_{-\infty}^{0} \{|x|^a [F(x)]\}^{\frac{r-a}{a}}[F(x)]^{k-1-\frac{r-a}{a}}[1-F(x)]^{n-k}x^a p(x)\ dx \end{aligned}$$

需要

$$k-1-\frac{r-a}{a} > 0 \to r < ak$$

考虑 $(2) < \infty$

$$\begin{aligned} (2) &= \int_{0}^{+\infty} |x|^r [F(x)]^{k-1}[1-F(x)]^{n-k}p(x)\ dx \\ &= \int_{0}^{+\infty} \{|x|^a [1-F(x)]\}^{\frac{r-a}{a}}[1-F(x)]^{n-k-\frac{r-a}{a}}[F(x)]^{k-1}x^a p(x)\ dx \end{aligned}$$

需要

$$n-k-\frac{r-a}{a} > 0 \to r < a(n-k+1)$$

综上，$r \leqslant a \cdot \min(k, n-k+1)$

例 1.5 设随机变量 X 服从 Cauchy 分布，其密度函数为

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

请考虑总体的 k 阶矩和次序统计量的 k 阶矩为：

因为 X 服从 Cauchy 分布，则 $E|X|^r=\infty,\ \forall r\leqslant 1$，然而对于任意小 $\varepsilon>0$，有

$$\begin{aligned}
E\,|X|^{1-\varepsilon}&=\int_{-\infty}^{\infty}|\,x\,|^{1-\varepsilon}\cdot\frac{1}{\pi(1+x^2)}dx=\int_0^{\infty}x^{-\varepsilon}\cdot\frac{2x}{\pi(1+x^2)}dx\\
&=\int_0^{\infty}\frac{x^{-\varepsilon}}{\pi}d[\ln(1+x^2)]=\frac{\ln(1+x^2)}{\pi x^{\varepsilon}}\bigg|_0^{\infty}+\frac{\varepsilon}{\pi}\int_0^{\infty}\ln(1+x^2)x^{-\varepsilon-1}dx\\
&=0+\frac{\varepsilon}{\pi}\int_0^{\infty}x^{-1-\varepsilon/2}\frac{\ln(1+x^2)}{x^{\varepsilon/2}}dx
\end{aligned}$$

由于 $\lim\limits_{x\to\infty}\dfrac{\ln(1+x^2)}{x^{\varepsilon/2}}=0$，故存在 $x_0>0,M>0$，使得当 $x>x_0$ 时，$\dfrac{\ln(1+x^2)}{x^{\varepsilon/2}}<M$，从而

$$\begin{aligned}
E|X|^{1-\varepsilon}&=\frac{\varepsilon}{\pi}\int_0^{\infty}x^{-1-\varepsilon/2}\frac{\ln(1+x^2)}{x^{\varepsilon/2}}dx\\
&\leq\frac{\varepsilon}{\pi}[\int_0^{x_0}x^{-1-\varepsilon}\ln(1+x^2)dx+M\int_{x_0}^{\infty}x^{-1-\varepsilon/2}dx]<\infty
\end{aligned}$$

这表明：Cauchy 随机变量的 $1-\varepsilon$ 阶矩存在，其中 $\varepsilon>0$ 可以任意小。

评论

1.4 统计量的近似分布

1.4.1 随机变量序列的两种收敛性

定义 1.15 (依概率收敛) 设 $\{Z_n\}$ 为一随机变量序列，Z 为另一随机变量，若 $\forall \varepsilon > 0$，有 $P\{|Z_n - Z| > \varepsilon\} \to 0$, $n \to \infty$，则称随机变量序列 $\{Z_n\}$依概率收敛于 Z，记为 $Z_n \xrightarrow{P} Z$。

定义 1.16 (依分布收敛) 设 $\{Z_n\}$ 为一随机变量序列，Z 为另一随机变量，又 $F_n(x)$ 与 $F(x)$ 分别是 $\{Z_n\}$ 和 Z 的分布函数，若对 $F(x)$ 的每个连续点 x 有 $F_n(x) \to F(x)$, $n \to \infty$，则称随机变量序列 $\{Z_n\}$依分布收敛于 Z，记为 $Z_n \xrightarrow{L} Z$。

定理 1.2 设 $Z_n \xrightarrow{P} Z$，则 $Z_n \xrightarrow{L} Z$.

证明. 对于 $\forall x, x'\,(x' < x)$，由于

$$\begin{aligned}\{Z < x'\} &= \{Z_n < x, Z < x'\} \cup \{Z_n \geq x, Z < x'\}\\ &\subset \{Z_n < x\} \cup \{Z_n \geq x, Z < x'\}\end{aligned}$$

因而有

$$F(x') \leqslant F_n(x) + P\{Z_n \geqslant x, Z < x'\}$$

另一方面，有 $Z_n \xrightarrow{p} Z$，当 $n \to \infty$

$$P\{Z_n \geqslant x, Z < x'\} \leqslant P\{|Z_n - Z| \geqslant x - x'\} \to 0$$

故而 $F(x') \leqslant \lim_{n\to\infty}\inf F_n(x)$。

同理，$\forall x, x''\,(x'' > x)$，由于

$$\begin{aligned}\{Z_n < x\} &= \{Z < x'', Z_n < x\} \cup \{Z \geqslant x'', Z_n < x\}\\ &\subset \{Z < x''\} \cup \{Z_n \geq x, Z < x'\}\end{aligned}$$

因而有

$$F(x'') \geqslant F_n(x) - P\{Z_n \geqslant x, Z < x''\}$$

另一方面，有 $Z_n \xrightarrow{p} Z$，当 $n \to \infty$

$$P\{Z < x'', Z_n \geqslant x\} \geqslant P\{|Z - Z_n| \geqslant x'' - x\} \to 0$$

故而

$$F(x'') \geq \lim_{n\to\infty}\sup F_n(x),$$

所以，$\forall x' < x < x''$，有

$$F(x') \leq \lim_{n\to\infty}\inf F_n(x) \leq \lim_{n\to\infty}\sup F_n(x) \leq F(x'')$$

当 x 时 $F(x)$ 的连续点，令 $x' \to x, x'' \to x$，则有

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

即 $Z_n \xrightarrow{L} Z$

定理 1.3 设 c 为常数，若 $Z_n \xrightarrow{L} c$，则 $Z_n \xrightarrow{P} c$

证明. 对 $\forall \varepsilon > 0$，有

$$
\begin{aligned}
&P\{\left|Z_n - c\right| \geqslant \varepsilon\} = P\{Z_n - c \geq \varepsilon\} + P\{Z_n - c \leq -\varepsilon\} \\
&= 1 - P\{Z_n < c + \varepsilon\} + P\{Z_n \leq c - \varepsilon\} \leq 1 - F_n(c + \frac{\varepsilon}{2}) + F_n(c - \varepsilon) \\
&\to 1 - F(c + \frac{\varepsilon}{2}) + F(c - \varepsilon) = 1 - 1 + 0 = 0 \quad (n \to \infty)
\end{aligned}
$$

定理 1.4 (Slutsky 定理) 设 $\{Z_n\}$ 和 $\{U_n\}$ 是两个随机变量序列，若 $Z_n \xrightarrow{L} Z$, $U_n \xrightarrow{P} c$(常数)，则有

$$
\begin{aligned}
&\text{a)} \quad Z_n + U_n \xrightarrow{L} Z + c \\
&\text{b)} \quad Z_n U_n \xrightarrow{L} Zc \\
&\textbf{c)} \quad Z_n / U_n \xrightarrow{L} Z/c \quad (c \neq 0)
\end{aligned}
$$

证明. a) 一方面

$$
\begin{aligned}
&P\{Z_n + U_n \leq x\} \\
&= P\{Z_n + U_n \leq x, |U_n - c| \leq \varepsilon\} + P\{Z_n + U_n \leq x, |U_n - c| > \varepsilon\} \\
&\leq P\{Z_n \leq x - c + \varepsilon, |U_n - c| \leq \varepsilon\} + P\{|U_n - c| > \varepsilon\} \\
&\leq P\{Z_n \leq x - c + \varepsilon\} + P\{|U_n - \varepsilon| > \varepsilon\}
\end{aligned}
$$

因此

$$
\begin{aligned}
&\lim_{n\to\infty}\sup P\{Z_n+U_n\le x\}\\
&\le \lim_{n\to\infty} P\{Z_n\le x-c+\varepsilon\}+\lim_{n\to\infty} P\{|U_n-c|>\varepsilon\}\\
&=P\{Z\le x-c+\varepsilon\}=F(x-c+\varepsilon)
\end{aligned}
$$

定理 1.5 设 $\{Z_n\}$ 为一随机变量序列，且 $Z_n \xrightarrow{P} c$(常数)，又函数 $g(\cdot)$ 在点 c 处连续，则 $g(Z_n)\xrightarrow{P} g(c)$

例 1.6 设 $\{X_n\}$ 是独立同分布随机变量序列，其均值为 μ，方差为 $\sigma^2<\infty$。模仿正态总体下的 t 统计量，构造

$$t_n=\sqrt{n}\left(\bar{X}-\mu\right)S_n$$

常称 t_n 为 t 化统计量，现求 t_n 的渐进分布。

由辛钦大数定律可知

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\xrightarrow{P}\sigma^2,\quad \bar{X}_n=\sum_{i=1}^{n}X_i\xrightarrow{P}\mu$$

因此由 Slutsky 定理可知，

$$\frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)-(\bar{X}-\mu)^2\right]\xrightarrow{P}\sigma^2$$

另一方面，由中心极限定理

$$Z_n=\sqrt{n}\left(\bar{X}_n-\mu\right)/\sigma\xrightarrow{L}N(0,1)$$

最后，由 Slutsky 定理可知 $t_n\xrightarrow{L}N(0,1)$

定理 1.6 设 $\{a_n\}$ 为一趋于 ∞ 的序列，b 为常数，并且对随机变量序列 $\{Z_n\}$ 有

$$a_n\left(Z_n-b\right) \xrightarrow{L} Z$$

有设 $g(\cdot)$ 为可微函数，且 g' 在点 b 处连续，则有

$$a_n\left[g(Z_n)-g(b)\right] \xrightarrow{L} g'(b)Z$$

例 1.7

1.5 充分统计量

2 点估计

2.1 基本概念

定义 2.1 (点估计) 设参数 θ，X 为样本，用统计量 $\hat{\theta}(X)$ 作为未知参数 θ 的“猜测”称为点估计。

该估计的均方误差 MSE 为

$$\begin{aligned}
\mathrm{MSE}(\hat{\theta}) &= \mathrm{E}(\hat{\theta}(X)-\theta)^2 = \mathrm{E}\{[\hat{\theta}(X)-\mathrm{E}(\hat{\theta}(X))]-[\theta-\mathrm{E}(\hat{\theta}(X))]\}^2 \\
&= \mathrm{E}[\hat{\theta}(X)-\mathrm{E}(\hat{\theta}(X))]^2+\mathrm{E}[\theta-\mathrm{E}(\hat{\theta}(X))]^2 \\
&= \mathrm{var}(\hat{\theta})+\mathrm{bias}(\hat{\theta})
\end{aligned}$$

2.2 无偏估计及 UMVUE

2.2.1 无偏估计

定义 2.2 (无偏估计) 设 $(\mathscr{X},\mathscr{B},\{P_\theta,\theta\in\Theta\})$ 为可控参数统计结构,$g(\theta)$ 是未知参数,$X=(X_1,\cdots,X_n)$ 是来自该统计结构的一个样本，若用 $\hat{g}(X)$ 估计 $g(\theta)$，且

$$\mathrm{E}_\theta(\hat{g}(X))=g(\theta),\forall\theta\in\Theta$$

则称 $\hat{g}(X)$ 为 $g(\theta)$ 的无偏估计。

评论 关于无偏估计，我们有三点需要说明

1. 无偏估计不一定存在
2. 对可估参数，无偏估计一般不唯一
3. 无偏估计并不一定是一个好估计

2.2.2 UMVUE

定义 2.3 设 $g(\theta)$ 是可估参数，$T(X)\in U_g$，若对于 $\forall\phi(X)\in U_g$ 有

$$\mathrm{var}_\theta T(X)\leq\mathrm{var}_\theta\,\phi(X),\forall\theta\in\Theta$$

则称 $T(X)$ 是 $g(\theta)$ 的一致最小方差无偏估计（UMVUE）。

引理 2.1 设 $S(X)$ 是分布族 $\{p_\theta, \theta \in \Theta\}$ 的充分统计量，$\phi(X)$ 是 $g(\theta)$ 的无偏估计，令 $T(X) = \mathrm{E}(\phi(X)|S(X))$，则 $T(X)$ 也是 $g(\theta)$ 的无偏估计，且 $\operatorname{var} T(X) \le \phi(X)$

评论 求 UMVUE 的方法

1. 寻找充分完备统计量的函数 $S(X)$，使其属于 U_g
2. 任取 $\hat{g}(X) \in U_g$，令 $T(X) = \mathrm{E}(\hat{g}(X)|S(X))$

例 2.1 设 $X = (X_1, \cdot, X_n)$ 是来自 $b(1,\theta), \theta \in (0,1)$ 的一个样本，由指数组性质知，$S(X) = \sum X_i$ 是充分完备统计量。

1. 对 θ，因为 $E(S(X)) = n\theta$，则 $\bar{X} = S(X)/n$ 是 θ 的无偏估计，从而也 θ 的 UMVUE。
2. 对于 $g(\theta) = \theta^k + (1-\theta)^{n-k}, 0 \le k \le n$，如何找 $g(\theta)$ 的 UMVUE?

对于 $g(\theta) = \theta^k + (1-\theta)^{n-k}, 0 \le k \le n$，先找一个 $g(\theta)$ 的无偏估计 $\phi(X)$。令

$$\phi_1(X) = \begin{cases} 1, & \sum\limits_{i=1}^{k} X_i = k \\ 0, & others \end{cases}, \quad \phi_2(X) = \begin{cases} 1, & \sum\limits_{i=k+1}^{n} X_i = 0 \\ 0, & others \end{cases}$$

又令 $\phi(X) = \phi_1(X) + \phi_2(X)$，则

$$\mathrm{E}(\phi(X)) = \mathrm{E}(\phi_1(X)) + \mathrm{E}(\phi_2(X)) = \theta^k + (1-\theta)^{n-k} = g(\theta)$$

所以 $\phi(X)\in U_g$。故而 $T(X)=\mathrm{E}(\phi(X)|S(X))$ 是 $g(\theta)$ 的 UMVUE。当 $k\le s$ 时，

$$\begin{aligned}
&\mathrm{E}(\phi_1(X)|S(X))=P(\phi_1(X)=1|S(X)=s)\\
&=\frac{P(\phi_1(X)=1,S(X)=s)}{P(S(X)=s)}=\frac{P(\sum\limits_{i=1}^{k}X_i=k,\sum\limits_{i=k+1}^{n}X_i=s-k)}{P(\sum\limits_{i=1}^{n}X_i=s)}\\
&=\frac{\theta^k\cdot C_{n-k}^{s-k}\theta^{s-k}(1-\theta)^{n-s}}{C_n^s\theta^k(1-\theta)^{n-k}}=\frac{C_{n-k}^{s-k}}{C_n^s}
\end{aligned}$$

当 $k\ge s$，$\mathrm{E}(\phi_1(X)|S(X))=0$。

同理，当 $k\ge s$ 时，$\mathrm{E}(\phi_2(X)|S(X))=\frac{C_k^s}{C_n^s}$

当 $k<s$ 时，$\mathrm{E}(\phi_2(X)|S(X))=0$。

因此，$g(\theta)$ 的 UMVUE 为

$$T(X)=\begin{cases}\frac{C_{n-k}^{s-k}}{C_n^s} & k\le s\\ \frac{C_k^s}{C_n^s} & k>s\end{cases}$$

例 2.2 设 $X_1,\cdots,X_n$ 是来自总体 X 的几何分布，其概率分布为 $P_\theta(X=x)=(1-\theta)^x\theta,x=0,1,\cdots,$。试求 θ 的 UMVUE。

对于 $\phi(X)$

$$\phi(X)=\begin{cases}1, & X_1=0\\ 0, & others\end{cases}$$

有 $\mathrm{E}(\phi(X))=P_\theta(X_1=1)=(1-\theta)^0\theta=\theta$。且 $\sum_{i=1}^n X_i\sim Nb(n;s,\theta)$ 服从负二项分布。因此，

$$\begin{aligned}\mathrm{E}(\phi(X)|S(X))&=P(\phi(X)=1|S(X)=s)\\&=\frac{P(\phi_1(X)=1,S(X)=s)}{P(S(X)=s)}=\frac{P(X_1=0,\sum\limits_{i=2}^n X_i=s)}{P(\sum\limits_{i=1}^n X_i=s)}\\&=\frac{\theta C_{s+n-2}^{n-2}\theta^{n-2}(1-\theta)^s}{C_{s+n-1}^{n-1}\theta^{n-1}(1-p)^s}=\frac{n-1}{s+n-1}\end{aligned}$$

例 2.3 设 $X_1,\cdot,X_n$ 独立同分布，均服从区间 $(0,\theta),\theta>0$ 上的均匀分布，样本为 $X=(X_1,\cdots,X_n)$。求参数 $g(\theta)=\theta^2$ 的 UMVUE。

首先，我们知道 $X_{(n)}$ 是为一个完全充分统计量。

直接方法：找到一个合适的 $X_{(n)}$ 的函数 $\phi(X_{(n)})$，使得 $\phi(X_n)$ 称为 $g(\theta)=\theta^2$ 的无偏估计，即 $E_\theta(\phi(X_n))=\theta^2$。为此，首先注意到 $X_{(n)}$ 的概率密度函数为

$$p(t;\theta)=\begin{cases}n\dfrac{t^{n-1}}{\theta^n}, & 0<t<\theta\\ 0, others\end{cases}$$

我们看一下其期望

$$E_\theta X_{(n)} = n\int_0^\theta n\frac{t^{n-1}}{\theta^n}t\,dx = \frac{n}{n+1}\theta$$

$$E_\theta X_{(n)}^2 = n\int_0^\theta n\frac{t^{n-1}}{\theta^n}t^2\,dt = \frac{n}{n+2}\theta^2$$

于是我们就有 $E_\theta\left(\frac{n+2}{n}X_{(n)}^2\right)=\theta^2$ 就得到 $\frac{n+2}{n}X_{(n)}^2$ 就是 θ^2 的无偏估计，进而是 UMVUE。

条件期望法：我们先找到 θ^2 的一个无偏估计，$E_\theta X_1^2=\int_0^\theta t^2\frac{1}{\theta}\,dt=\frac{\theta^2}{3}$，课件 $3X_1^2$ 是 θ^2 的一个无偏估计，进而求期望 $E_\theta\left[3X_1^2|X_{(n)}\right]$.

$$\begin{aligned} E_\theta\left[3X_1^2|X_{(n)}=t\right] &= \frac{1}{n}\cdot 3t^2+(1-\frac{1}{n})\int_0^t 3u^2\frac{1}{t}\,du \\ &= \frac{3t^2}{n}+\frac{n-1}{n}t^2 \\ &= \frac{n+2}{n}t^2 \end{aligned}$$

可见，$\frac{n+2}{n}X_{(n)}^2$ 为 θ^2 的 UMVUE

2.3 极大似然估计（MLE）

极大似然估计在直观上可以这样解释：使得出现所选样本最大概率的分布参数的估计。

定义 2.4 设 $X \sim f(x;\theta), \theta \in \Theta$，把 $f(x;\theta)$ 是为 θ 的哈桑农户，则称它为 X 关于 θ 的似然函数，$L(\theta, X) = \ln f(x;\theta) = L(\theta)$ 称为对数似然函数，若 $\hat{\theta}$ 满足

$$f(x, \hat{\theta}(x)) = \max_{\theta \in \Theta} f(x;\theta)$$

称 $\hat{\theta}$ 为 θ 的极大似然估计。

例 2.4 设 $X_1, \cdots, X_n$ 是来自 $b(1,\theta)$ 的一个样本，$0 < \theta < 1$

$$l(\theta; x) = \left(\sum x_i\right) \ln \theta + \left(n - \sum x_i\right) \ln(1-\theta)$$

由 $\dfrac{\partial l}{\partial \theta} = 0$ 知

$$\hat{\theta} = \bar{x}$$

评论 MLE 还有一个非常有吸引力的性质：如果 $\hat{\theta}$ 是 θ 的 MLE，$g(\cdot)$ 为可测函数，那么 $g(\hat{\theta})$ 也是 $g(\theta)$ 的 MLE。

对于多参数指数族 $p(x;\theta) = \exp\{\sum_{j=1}^k \theta_j T_j(x) + c(\theta) + d(x)\}$，似然方程化为

$$\sum T_j(x_i) = -n \frac{\partial c(\theta)}{\partial \theta_j}$$

定理 2.2 (MLE 的相合性) 设 $\{p_\theta(x;\theta) : \theta \in \Theta\}$ 是可识别的，且 $p_\theta(x;\theta)$ 关于 θ 可微，则似然方程在 $n \to \infty$ 时，以概率 1 有解，且此解关于 θ 时相合的。

定理 2.3 (MLE 的渐进正态性) 在一系列正则条件下，对于 $p(x;\theta)$ 的相合解 $\hat{\theta}_n$，有

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{L} N(0, I^{-1}(\theta_0))$$

其中 θ_0 为真值。

3 假设检验

3.1 N-P 基本引理

定义 3.1 (最优势检验) 在检验问题 (θ_0, θ_1) 中，若 ϕ 是一个 α 水平的检验，若对任意一切 α 水平的检验 ϕ'，均有 $\mathrm{E}_{\theta_1}\phi(X) \geq \mathrm{E}_{\theta_1}\phi'(X)$，则称 ϕ 时最优势检验 MPT。

$$\mathrm{E}_\theta\phi(X) = g_\phi(\theta) = \begin{cases} \alpha(\theta) & \theta \in \Theta_0 \\ 1 - \beta(\theta) & \theta \in \Theta_1 \end{cases}$$

定理 3.1 (N-P 引理) 对于问题 (θ_0, θ_1)，设分布 $p_{\theta_0}, p_{\theta_1}$ 密度函数存在，则 $\forall \alpha \in (0, 1)$

1. 存在一个 α 水平的检验 ϕ 及 $\lambda_0 > 0$ 使得 ϕ 形如

$$\phi(X) = \begin{cases} 1 & p(x;\theta_1) \geq \lambda_0 p(x;\theta_0) \\ 0 & p(x;\theta_1) < \lambda_0 p(x;\theta_0) \end{cases}$$

且 $\mathrm{E}_{\theta_0}\phi(X) = \alpha$

2. 此 $\phi(X)$ 是一个 MPT
3. 反之，如果 $\phi(X)$ 是水平为 α 的 MPT，则一定存在常数 $\lambda_0 \geq 0$，使得 $\phi(X)$ 满足 1

例 3.1 设 $X_1, X_2, \cdots, X_n$ 来自两点分布 $b(1,\theta)$ 的样本，对于假设问题 $H_0: \theta = \dfrac{1}{2}, H_1: \theta = \theta_1(\theta_1 > \frac{1}{2})$，试求其最优势检验

似然比：

$$\lambda(X) = \frac{\prod_{i=1}^n p(x_i;\theta_1)}{\prod_{i=1}^n p(x_i;\theta_0)} = 2^n \left(\frac{\theta_1}{1-\theta_1}\right)^{\sum_{i=1}^n x_i} (1-\theta_1)^n$$

原假设成立条件下，$T = \sum_{i=1}^n X_i \sim b(n, \dfrac{1}{2})$，且 $\lambda(X)$ 是关于 T 递增的，那么 $\{\lambda \geq k\} = \{T \geq c\}$

$$\phi(X) = \begin{cases} 1 & T > c \\ r & T = c \\ 0 & T < c \end{cases}$$

$$\alpha = \mathrm{E}_{\theta_0}\phi(X) = P\{T > c\} + rP(T = c)$$

得到 $c = u_{1-\alpha_1}, r = \dfrac{\alpha - \alpha_1}{G(c-0) - G(c)} = \dfrac{\alpha - \alpha_1}{F(c) - F(c-0)}$

例 3.2 设 $X_1, X_2, \cdots, X_n$ 来自两点分布 $N(\mu, 1)$ 的样本，对于假设问题 $H_0: \mu = 0, H_1: \mu = \mu_1(\mu_1 > 0)$，试求其最优势检验

似然比：

$$\lambda(X) = \frac{\prod_{i=1}^n p(x_i;\theta_1)}{\prod_{i=1}^n p(x_i;\theta_0)} = \exp\{-\frac{1}{2}n\mu_1^2 + \sum_{i=1}^n x_i\mu\}$$

原假设成立条件下，$T=\bar{X}\sim N(0,\dfrac{1}{n})$，且 $\lambda(X)$ 是关于 T 递增的，那么 $\{\lambda\geq k\}=\{T\geq c\}$

$$\phi(X)=\begin{cases}1 & T\geq c\\ 0 & T<c\end{cases}$$

$$\alpha=\mathrm{E}_{\theta_0}\phi(X)=P\{T\geq c\}$$

得到 $c=\frac{u_{1-\alpha}}{\sqrt{n}}$

例 3.3 设 $X=(X_1,\cdots,X_n)$ 是来自均匀分布组 $\{R(0,\theta):\theta>0\}$ 的样本，考虑如下检验问题：$H_0:\theta=1,H_1:\theta_1=\theta_1(\theta_1>1)$，取水平为 $\alpha(0<\alpha<1)$。构造似然比统计量

$$\lambda(X)=\frac{\prod_{i=1}^n p(x_i;\theta_1)}{\prod_{i=1}^n p(x_i;\theta_0)}=\begin{cases}\theta_1^{-n} & 0<x_{(n)}<1\\ \infty & 1\leq x_{(n)}\leq\theta_1\end{cases}$$

$\lambda(X)$ 为退化分布

取非随机化检验：

$$\phi(x)=\begin{cases}1 & c\leq x_{(n)}<\theta_1\\ 0 & 0<x_{(n)}<c\end{cases}$$

原假设 H_0 成立时，$T=X_{(n)}$ 的密度函数为 $nt^{n-1},(1<t<1)$，故由 $\mathrm{E}_{\theta_0}\pi(X)=\alpha$ 得 $c=\sqrt[n]{1-\alpha}$

例 3.4 设 $X=(X_1,\cdots,X_n)$ 是来自均匀分布组 $\{R(0,\theta):\theta>0\}$ 的样本，考虑如下检验问题：$H_0:\theta=1,H_1:\theta_1=\theta_1(\theta_1<1)$，取水平为 $\alpha(0<\alpha<1)$。构造似然比统计量

$$\lambda(X)=\frac{\prod_{i=1}^n p(x_i;\theta_1)}{\prod_{i=1}^n p(x_i;\theta_0)}=\begin{cases}\theta_1^{-n} & 0<x_{(n)}<\theta_1\\ 0 & \theta_1\leq x_{(n)}<1\end{cases}$$

$\lambda(X)$ 为退化分布

取非随机化检验:

$$\phi(x)=\begin{cases}1 & 0\leq x_{(n)}\leq c\\ 0 & c<x_{(n)}<\theta_1\end{cases}$$

原假设 H_0 成立时，$T=X_{(n)}$ 的密度函数为 $nt^{n-1},(1<t<1)$，故由 $\mathrm{E}_{\theta_0}\phi(X)=\alpha$ 得 $c=\sqrt[n]{\alpha}$

例 3.5 设 $X_1,\cdots,X_n\sim U(\theta,1)$ 的样本，考虑如下检验问题：$H_0:\theta=0,H_1:\theta_1=\theta_1(\theta_1<0)$，取水平为 $\alpha(0<\alpha<1)$，试求其 MPT。

构造似然比统计量

$$\lambda(X)=\frac{\prod_{i=1}^n p(x_i;\theta_1)}{\prod_{i=1}^n p(x_i;\theta_0)}=\begin{cases}(1-\theta_1)^{-n} & 0<x_{(1)}<1\\ \infty & \theta_1\leq x_{(1)}\leq 0\end{cases}$$

$\lambda(X)$ 为退化分布

取非随机化检验:

$$\phi(x)=\begin{cases}1 & 0\leq x_{(1)}\leq c\\ 0 & c<x_{(n)}<1\end{cases}$$

原假设 H_0 成立时,$T=X_{(1)}$ 的密度函数为 $n(1-t)^{n-1},(1<t<1)$,故由 $\mathrm{E}_{\theta_0}\phi(X)=\alpha$ 得 $c=1-\sqrt[n]{1-\alpha}$

例 3.6 电话交换台单位时间内接到的呼唤次数服从 Poisson 分布 $P(\lambda),\lambda>0$. λ 为单位时间内接到的平均呼唤次数. 设 $x=(x_1,\cdots,x_{10})$ 是该电话交换台的 10 次记录. 考虑假设检验问题：原假设 $H_0:\lambda\geqslant 1$ 对备择假设 $H_1:\lambda<1$. 取水平为 $\alpha=0.05$.

解：取检验统计量为 λ 的完备充分统计量 $T(x)=\sum_{i=1}^{n}x_i$.

对于 $H_0:\lambda\geqslant 1$ 和 $H_1:\lambda<1$ ，其拒绝域为 $W=\{x:\sum_{i=1}^{n}x_i\leqslant c\}$，

检验函数为：

$$\Phi(x)=\begin{cases}1, & T(x)<c,\\ r, & T(x)=c,\\ 0, & T(x)>c.\end{cases}$$

势函数为：

$$g(\lambda)=\mathrm{P}_{\lambda}(x\in W)=\sum_{k=0}^{c-1}\frac{(n\lambda)^k}{k!}e^{-n\lambda}+r\frac{(n\lambda)^c}{c!}e^{-n\lambda}$$

，

当 $n=10$ ，$\lambda=1$ 时，由

$$\begin{cases}\sum_{k=0}^{4}\frac{(n\lambda)^k}{k!}e^{(-n\lambda)}=0.02921\\ \sum_{k=0}^{5}\frac{(n\lambda)^k}{k!}e^{(-n\lambda)}=0.06704\end{cases}$$

得 c=5 .

即 $\sum_{k=0}^{4}\frac{(n\lambda)^k e^{(-n\lambda)}}{k!}+r\frac{(n\lambda)^c}{c!}e^{-n\lambda}=0.05$ ，解得 $r=0.5496$. 故检验函数为：

$$\Phi(x)=\begin{cases}1, & T(x)<c,\\ 0.5496, & T(x)=c,\\ 0, & T(x)>c.\end{cases}$$

例 3.7 设 $X=(X_1,\cdots,X_n)$ 是来自正态分布族 $\{N(0,\sigma^2):0<\sigma^2<\infty\}$ 的样本，考虑原假设 $H_0:\sigma^2=1$ 对备择假设 $H_1:\sigma^2=\sigma_1^2(\sigma_1^2>1)$ 的检验问题，取水平为 $\alpha(0<\alpha<1)$ ，试求其 MPT.

解：

密度函数：$p(x;\sigma^2)=(2\pi)^{-1/2}\sigma^{-1}\exp\{x^2/(2\sigma^2)\}$ ，

似然函数：$L(x;\sigma^2)=(2\pi)^{-n/2}\sigma^{-n}\exp\{-\sum_{i=1}^n x_i^2/(2\sigma^2)\}$ ，

由因子分解定理知，$T(x)=\sum_{i=1}^n x_i^2$ 为该分布的完备充分统计量.

构造似然比统计量：

$$\begin{aligned}\lambda(x)=\frac{\prod_{i=1}^n p\left(x_i;\sigma_1^2\right)}{\prod_{i=1}^n p\left(x_i;\sigma_0^2\right)}&=\frac{\sigma_0^n}{\sigma_1^n}\exp\left\{\sum_{i=1}^n x_i^2\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\right\}\\&=\frac{\sigma_0^n}{\sigma_1^n}\exp\left\{T(x)\left(\frac{1}{2\sigma_0^2}-\frac{1}{2\sigma_1^2}\right)\right\},\end{aligned}$$

$\lambda(x)$ 关于 $T(x)$ 严格单调上升, 根据 N-P 基本引理, MPT 的拒绝域形式为 $W=\{x:T(x)=\sum_{i=1}^n x_i^2\geqslant c\}$.

当 H_0 成立时，$T\sim\chi^2(n)$, 所以对给定的水平 α ，$c=\chi^2_{1-\alpha}(n)$.

MPT 检验仅与水平 α 有关，而与 σ_1^2 的具体数值无关，只要求 $\sigma_1^2>1$ 就行了.

故 MPT 为：

$$\phi(x)=\begin{cases}1, & T\geqslant\chi^2_{1-\alpha}(n),\\ 0, & T<\chi^2_{1-\alpha}(n).\end{cases}$$