

# 第三章 多元正态分布的估计与检验

## 3.1 多元正态分布样本统计量

设  $X_1, X_2, \dots, X_n$  为来自多元正态总体  $N_p(\mu, \Sigma)$  的独立样本, 其中  $\mu \in R^p, \Sigma > 0, n > p$ . 记

$$\bar{X} = n^{-1} \sum_{i=1}^n X_i, \text{ 称为样本均值};$$

$$V = \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})', \text{ 称为样本离差阵}.$$

事实:  $(\bar{X}, V)$  是  $(\mu, \Sigma)$  的完全充分统计量.

3.1.1  $(\bar{X}, V)$  的分布性质:

- 1)  $\bar{X} \stackrel{d}{\sim} N_p(\mu, \Sigma/n)$ ;
- 2)  $V \stackrel{d}{\sim} W_p(n-1, \Sigma)$ ;
- 3)  $\bar{X}$  与  $V$  相互独立.

证明: 记  $X = (X_1, \dots, X_n)$ , 则有  $E(X) = \mu 1'_n$ ,  $Cov(vec(X)) = I_n \otimes \Sigma$ .

令  $U = (U_1, U_2, \dots, U_n)$  为  $n$  阶正交矩阵, 其中

$$U_1 = \left( \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \dots, \frac{1}{\sqrt{n}} \right)' = \frac{1}{\sqrt{n}} 1_n.$$

再令  $Y = XU \stackrel{\text{记为}}{=} (Y_1, Y_2, \dots, Y_n)$ , 则

$$E(Y) = E(X)U = \mu 1'_n U = \mu 1'_n (1_n/\sqrt{n}, U_2, \dots, U_n) = (\sqrt{n}\mu, 0, \dots, 0),$$

$$\begin{aligned}
Cov(vec(Y)) &= Cov(vec(\mathbf{I}_p \mathbf{X} \mathbf{U})) \\
&= Cov((\mathbf{U}' \otimes \mathbf{I}_p) vec(\mathbf{X})) \\
&= (\mathbf{U}' \otimes \mathbf{I}_p) Cov(vec(\mathbf{X})) (\mathbf{U} \otimes \mathbf{I}_p) \\
&= (\mathbf{U}' \otimes \mathbf{I}_p) (\mathbf{I}_n \otimes \Sigma) (\mathbf{U} \otimes \mathbf{I}_p) \\
&= \mathbf{I}_n \otimes \Sigma.
\end{aligned}$$

即知  $Y_1, Y_2, \dots, Y_n$  相互独立, 且  $Y_1 = \sqrt{n} \bar{X} \stackrel{d}{\sim} N_p(\sqrt{n}\mu, \Sigma)$ ,  
 $Y_2, \dots, Y_n \stackrel{d}{\sim} N_p(0, \Sigma)$ .

因而有  $\bar{X} \stackrel{d}{\sim} N_p(\mu, \Sigma/n)$ , 即1)成立.

由于  $YY' = (XU)(U'X') = XX'$ , 即  $\sum_{i=1}^n X_i X'_i = \sum_{i=1}^n Y_i Y'_i$ , 因而有

$$\begin{aligned}
 V &= \sum_{i=1}^n X_i X'_i - n\bar{X}\bar{X}' = \sum_{i=1}^n X_i X'_i - Y_1 Y'_1 \\
 &= \sum_{i=1}^n Y_i Y'_i - Y_1 Y'_1 \\
 &= \sum_{i=2}^n Y_i Y'_i \stackrel{d}{\sim} W_p(n-1, \Sigma).
 \end{aligned}$$

所以2)成立.

又由于  $\sqrt{n}\bar{X} = Y_1$ ,  $V = \sum_{i=2}^n Y_i Y'_i$ , 因此  $\bar{X}$  与  $V$  独立, 即3)成立. #

## 3.2 多元正态分布的参数估计

### 3.2.1 极大似然估计:

观测样本  $X = (X_1, X_2, \dots, X_n)$  的联合密度,

$$f(X) = (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} [\Sigma^{-1} (V + n(\bar{X} - \mu)(\bar{X} - \mu)')] \right\}.$$

首先求  $\mu$  的极大似然估计.

$$\begin{aligned} f(X) &= (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} [\Sigma^{-1} (V + n(\bar{X} - \mu)(\bar{X} - \mu)')] \right\} \\ &= (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} (\Sigma^{-1} V) - \frac{n}{2} \text{tr} [\Sigma^{-1} (\bar{X} - \mu)(\bar{X} - \mu)'] \right\} \\ &= (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} (\Sigma^{-1} V) - \frac{n}{2} (\bar{X} - \mu)' \Sigma^{-1} (\bar{X} - \mu) \right\}, \end{aligned}$$

易知  $\mu$  的极大似然估计为:  $\hat{\mu} = \bar{X}$ .

即正态总体均值的极大似然估计是样本均值.

将 $\mu = \bar{X}$ 代入似然函数并去掉与参数无关的项, 有

$$f(X) \propto |\Sigma|^{-n/2} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1} V) \right\}. \quad (1)$$

考虑正交分解 $\Sigma^{-1/2} V \Sigma^{-1/2} = U \Lambda U'$ , 其中 $U$ 是正交矩阵,

$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$ . 那么有

$$|\Sigma|^{-1} = |V|^{-1} \prod_{i=1}^p \lambda_i,$$

$$\text{tr}(\Sigma^{-1} V) = \text{tr}(\Sigma^{-1/2} V \Sigma^{-1/2}) = \text{tr}(U \Lambda U') = \text{tr}(\Lambda U' U) = \text{tr}(\Lambda) = \sum_{i=1}^p \lambda_i.$$

则(1)式化为:

$$f(X) \propto |V|^{-n/2} \prod_{i=1}^p \left[ \lambda_i^{n/2} \exp \left\{ -\frac{\lambda_i}{2} \right\} \right]. \quad (2)$$

易知(2)式在  $\lambda_1 = \cdots = \lambda_p = n$  时取最大值, 故  $\Sigma$  的极大似然估计  $\hat{\Sigma}$  满足:

$$\hat{\Sigma}^{-1/2} V \hat{\Sigma}^{-1/2} = n I_p \implies \hat{\Sigma} = V/n.$$

因此, 正态总体参数  $(\mu, \Sigma)$  的极大似然估计为  $(\hat{\mu}, \hat{\Sigma}) = (\bar{X}, V/n)$ .

记

$$\begin{aligned} S &= \frac{V}{n-1}, & \text{为样本协方差阵,} \\ K &= S^{-1}, & \text{为样本精度矩阵,} \end{aligned}$$

有

$$\begin{aligned} E(\bar{X}) &= \mu; \\ E(S) &= E(V)/(n-1) = \Sigma. \end{aligned}$$

因此,  $(\bar{X}, S)$  是  $(\mu, \Sigma)$  的无偏估计(一致最小).

### 3.2.2 样本相关系数

记  $\Upsilon = (\rho_{ij})_{p \times p}$  为正态总体的相关系数矩阵,

并记  $V = (v_{ij})_{p \times p}$ ,  $S = (s_{ij})_{p \times p}$ ,

则  $\rho_{ij}$  的极大似然估计为

$$\hat{\rho}_{ij} = r_{ij} = \frac{v_{ij}}{\sqrt{v_{ii}}\sqrt{v_{jj}}} = \frac{s_{ij}}{\sqrt{s_{ii}}\sqrt{s_{jj}}}, \quad 1 \leq i, j \leq p.$$

记  $R = (r_{ij})_{p \times p}$  为样本相关系数矩阵.



## 样本相关系数的精确分布与渐近分布

### 1) 精确分布

考虑二元正态分布的情形, 即 $(X, Y)'$ 的 $n$ 个观测

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \dots, \begin{pmatrix} x_n \\ y_n \end{pmatrix} \text{ i.i.d. } \stackrel{d}{\sim} N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).$$

由3.1.1的性质2), 知样本离差阵  $V \stackrel{d}{\sim} W_2((n-1), \Sigma)$ , 其中

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

再由Wishart分布的定义知, 随机向量 $(v_{xx}, v_{xy}, v_{yy})$ 的密度函数为

$$\frac{(v_{xx}v_{yy} - v_{xy}^2)^{(n-4)/2} \exp \left\{ -\frac{v_{xx}-2\rho v_{xy}+v_{yy}}{2(1-\rho^2)} \right\}}{2^{n-1}\pi^{1/2}\Gamma(\frac{n-1}{2})\Gamma(\frac{n-2}{2})(1-\rho^2)^{(n-1)/2}},$$

其中,  $v_{xx} > 0$ ,  $v_{yy} > 0$ ,  $v_{xx}v_{yy} - v_{xy}^2 > 0$ .

由变换 $(v_{xx}, v_{xy}, v_{yy}) \rightarrow (v_{xx}, v_{yy}, r)$ 导出 $(v_{xx}, v_{yy}, r)$ 的密度函数, 再对 $v_{xx}$ 和 $v_{yy}$ 积分后, 可得 $r$ 的密度函数为

$$f(r|\rho, n) = \frac{2^{n-3}(1-\rho^2)^{(n-1)/2}(1-r^2)^{(n-4)/2}}{\pi(n-3)!} \sum_{k=0}^{\infty} \frac{(2\rho r)^k}{k!} \Gamma^2\left(\frac{n+k-1}{2}\right), \quad |r| < 1.$$

当  $\rho = 0$  时, 即  $X$  与  $Y$  独立, 样本相关系数  $r$  的密度函数为

$$f(r|n) = \frac{2^{n-3}(1-r^2)^{(n-4)/2}}{\pi(n-3)!} \Gamma^2\left(\frac{n-1}{2}\right), \quad |r| < 1.$$

因此, 可以用此分布作为零假设为  $X$  与  $Y$  独立时的统计量  $r$  的零分布来检验零假设是否成立.

检验问题:  $X$  与  $Y$  独立

等价于:

$$\mathbf{H}_0 : \rho = 0, \quad v.s. \quad \mathbf{H}_1 : \rho \neq 0.$$

检验方案一:

1) 给定显著性水平  $0 < \alpha < 0.5$  (通常取为 0.05);

2) 计算临界值  $C > 0$ , 满足

$$\int_{-C}^C f(r|n) dr = 1 - \alpha;$$

3) 如果  $|r| = \frac{|v_{xy}|}{\sqrt{v_{xx}}\sqrt{v_{yy}}} > C$ , 则拒绝零假设, 即认为  $X$  与  $Y$  不独立.

检验方案二:

通常采用统计量

$$T = \sqrt{n-2} \cdot \frac{r}{\sqrt{1-r^2}}$$

作为检验 $X$ 与 $Y$ 是否独立的统计量.

不难推导得

$$T = \sqrt{n-2} \cdot \frac{v_{xy}v_{xx}^{-1/2}}{\sqrt{v_{yy} - v_{xy}^2v_{xx}^{-1}}},$$

由Wishart分布的**性质6**, 即**独立分解性质**知, 当  $X$  与  $Y$  独立, 即  $\rho = 0$  时, 有

(1)  $v_{yy} - v_{xy}^2 v_{xx}^{-1}$  与  $v_{xy} v_{xx}^{-1/2}$  **相互独立**;

(2)  $v_{yy} - v_{xy}^2 v_{xx}^{-1} \stackrel{d}{\sim} \chi^2(\mathbf{n} - \mathbf{2})$ ;

(3)  $v_{xy} v_{xx}^{-1/2} \stackrel{d}{\sim} N_1(0, 1)$ ,

因此在零假设  $\rho = 0$ , 即  $X$  与  $Y$  独立的假设下, 有

$$\begin{aligned} T &= \sqrt{n-2} \cdot \frac{r}{\sqrt{1-r^2}} = \frac{\mathbf{v_{xy} v_{xx}^{-1/2}}}{\sqrt{\frac{\mathbf{v_{yy} - v_{xy}^2 v_{xx}^{-1}}}{\mathbf{n-2}}}} \\ &\stackrel{d}{=} \frac{N_1(0, 1)}{\sqrt{\frac{\chi^2(n-2)}{n-2}}} \\ &\stackrel{d}{\sim} t(n-2). \end{aligned}$$

### 检验方案二(续):

1) 给定显著性水平  $0 < \alpha < 0.5$ ;

2) 如果  $|T| > t_{1-\frac{\alpha}{2}}(n-2)$ , 则拒绝零假设, 即认为  $X$  与  $Y$  不独立,

其中  $t_{1-\frac{\alpha}{2}}(n-2)$  是自由度为  $(n-2)$  的标准t分布的  $(1 - \frac{\alpha}{2})$  分位点.

### 检验方案三:

对独立 *v.s.* 正相关的检验问题:

$$\mathbf{H}_0 : \rho = 0, \quad v.s. \quad \mathbf{H}_1 : \rho > 0,$$

计算检验的p值:  $p = P\{t(n-2) > T\}$ ,

如果p值小于显著性水平  $\alpha$ , 则拒绝零假设.

正态独立性检验的例子:

水泥在凝固时释放的热量与水泥中化学成分的关系案例(续)

样本相关系数矩阵为

	$x_1$	$x_2$	$x_3$	$x_4$	$y$
$x_1$	1.000				
$x_2$	0.229	1.000			
$x_3$	-0.824	-0.139	1.000		
$x_4$	-0.245	-0.973	0.030	1.000	
$y$	0.731	0.816	-0.535	-0.821	1.000

设定显著性水平为0.05.



(1) 变量  $x_3$  与  $x_4$  独立与正相关检验问题:

检验统计量为

$$\begin{aligned} T &= \sqrt{n-2} \cdot \frac{r}{\sqrt{1-r^2}} \\ &= \sqrt{13-2} \cdot \frac{0.030}{\sqrt{1-0.030^2}} \\ &= 0.0995. \end{aligned}$$

因此, 变量  $x_3$  与  $x_4$  独立与正相关检验问题的p值为

$$p = P\{t(11) \geq 0.0995\} = 0.4612 > 0.05,$$

故而不能拒绝  $x_3$  与  $x_4$  之间的相互独立性.

(2) 变量  $x_2$  与  $x_3$  独立与负相关检验问题的p值为

$$p = P\{t(11) \leq -0.4655\} = 0.3253 > 0.05,$$

则不能拒绝  $x_2$  与  $x_3$  之间的相互独立性.

(3) 变量  $y$  与  $x_3$  独立与负相关检验问题的p值为

$$p = P\{t(11) \leq -2.1\} = 0.0298 < 0.05,$$

则拒绝  $y$  与  $x_3$  之间的相互独立性, 认为它们负相关.

## 2) 渐近分布 (总体不服从正态分布时)

注意到, 样本相关系数  $r$  可以改写成

$$r = \frac{t_5 - t_1 t_2}{\sqrt{t_3 - t_1^2} \sqrt{t_4 - t_2^2}} = g(t_1, t_2, t_3, t_4, t_5),$$

其中,

$$\begin{aligned} t_1 &= n^{-1} \sum_{i=1}^n x_i, \\ t_2 &= n^{-1} \sum_{i=1}^n y_i, \\ t_3 &= n^{-1} \sum_{i=1}^n x_i^2, \\ t_4 &= n^{-1} \sum_{i=1}^n y_i^2, \\ t_5 &= n^{-1} \sum_{i=1}^n x_i y_i. \end{aligned}$$

记随机向量序列  $z_i = (x_i, y_i, x_i^2, y_i^2, x_i y_i)'$ ,  $1 \leq i \leq n$ ,

并令

$$\bar{Z} = n^{-1} \sum_{i=1}^n z_i = (t_1, t_2, t_3, t_4, t_5)'.$$

由于  $z_1, \dots, z_n$  *i.i.d.*, 因此由独立同分布随机向量和的中心极限定理知

$$\sqrt{n}(\bar{Z} - \mu_Z) \xrightarrow{d} N_5(0, \Sigma_Z),$$

其中,

$$\mu_Z = E(z_1) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ \rho \end{pmatrix}, \quad \Sigma_Z = Cov(z_1) = \begin{pmatrix} 1 & \rho & 0 & 0 & 3\rho \\ \rho & 1 & 0 & 0 & 3\rho \\ 0 & 0 & 2 & 2\rho^2 & 2\rho \\ 0 & 0 & 2\rho^2 & 2 & 2\rho \\ 3\rho & 3\rho & 2\rho & 2\rho & 1 + \rho^2 \end{pmatrix}.$$

此时有

$$g(\bar{Z}) = g(t_1, t_2, t_3, t_4, t_5) = r,$$

$$g(\mu_Z) = \rho,$$

再由Cramér定理, 可得样本相关系数  $r$  的渐近分布:

$$\begin{aligned} \sqrt{n}(r - \rho) &= \sqrt{n}(g(\bar{Z}) - g(\mu_Z)) \\ &\xrightarrow{d} N(0, (\dot{g}(\mu_Z))' \Sigma_Z \dot{g}(\mu_Z)) \\ &\xrightarrow{d} N(0, (1 - \rho^2)^2), \end{aligned} \tag{3}$$

其中  $\dot{g}$  是  $g$  的一阶偏导.

**注意:** 上面的结论与  $\rho$  是否为0无关.

### 3) 置信区间 (区间估计)

由于样本相关系数的精确分布不是分布自由的, 即与未知的总体相关系数  $\rho$  有关, 因此无法用于置信区间的构造.

可以用其渐近分布. 由(3)知

$$\frac{\sqrt{n}(r - \rho)}{1 - r^2} \xrightarrow{d} N(0, 1),$$

因此,  $\rho$  的一个水平为  $(1 - \alpha)$  的区间估计为

$$\left[ r - n^{-1/2}(1 - r^2)Z_{1-\alpha/2}, r + n^{-1/2}(1 - r^2)Z_{1-\alpha/2} \right],$$

其中  $Z_{1-\alpha/2}$  是标准正态分布的  $(1 - \alpha/2)$  分位点.

## 置信区间的统计意义

记

$$\begin{aligned}C_L &= r - n^{-1/2}(1 - r^2)Z_{1-\alpha/2}, \\C_U &= r + n^{-1/2}(1 - r^2)Z_{1-\alpha/2},\end{aligned}$$

有

$$\begin{aligned}Pr\{C_L \leq \rho \leq C_U\} &= Pr\left\{\left|\frac{\sqrt{n}(r - \rho)}{1 - r^2}\right| \leq Z_{1-\alpha/2}\right\} \\&\approx Pr\{|N(0, 1)| \leq Z_{1-\alpha/2}\} \\&= 1 - \alpha.\end{aligned}$$

相关系数置信区间的例子:

水泥在凝固时释放的热量与水泥中化学成分的关系案例(续)

样本相关系数矩阵为

	$x_1$	$x_2$	$x_3$	$x_4$	$y$
$x_1$	1.000				
$x_2$	0.229	1.000			
$x_3$	-0.824	-0.139	1.000		
$x_4$	-0.245	-0.973	0.030	1.000	
$y$	0.731	0.816	-0.535	-0.821	1.000

设定置信水平  $1 - \alpha = 0.95$ .



(1) 变量  $x_3$  与  $x_4$  的相关系数  $\rho_{34}$  的置信区间

$$\begin{aligned}C_L &= r - n^{-1/2}(1 - r^2)Z_{1-\alpha/2}, \\&= 0.030 - 13^{-1/2}(1 - 0.030^2) * 1.96 \\&= -0.513,\end{aligned}$$

$$\begin{aligned}C_U &= r + n^{-1/2}(1 - r^2)Z_{1-\alpha/2} \\&= 0.030 + 13^{-1/2}(1 - 0.030^2) * 1.96 \\&= 0.573,\end{aligned}$$

故变量  $x_3$  与  $x_4$  的相关系数  $\rho_{34}$  的置信区间为  $[-0.513, 0.573]$ .

0 在此区间里, 故不能拒绝  $x_3$  与  $x_4$  之间的相互独立性.

(2) 变量  $x_2$  与  $x_3$  的相关系数  $\rho_{23}$  的置信区间为  $[-0.672, 0.394]$ .

0 在此区间里，故也不能拒绝  $x_2$  与  $x_3$  之间的相互独立性.

(3) 变量  $x_1$  与  $x_3$  的相关系数  $\rho_{13}$  的置信区间为  $[-0.999, -0.650]$ .

0 不在此区间里，故不能认为  $x_1$  与  $x_3$  之间是相互独立的.

#### 4) 方差齐性变换

考虑一个单调变换  $h$ , 使得

$$\sqrt{n}(h(r) - h(\rho)) \xrightarrow{d} N(0, 1).$$

由(3)和Cramér定理知

$$\sqrt{n}(h(r) - h(\rho)) \xrightarrow{d} N(0, (h'(\rho))^2(1 - \rho^2)^2),$$

求解  $(h'(\rho))^2(1 - \rho^2)^2 = 1$  知

$$h(x) = \frac{1}{2} \ln \frac{1+x}{1-x}.$$

因此有

$$\sqrt{n} \left( \frac{1}{2} \ln \frac{1+r}{1-r} - \frac{1}{2} \ln \frac{1+\rho}{1-\rho} \right) \xrightarrow{d} N(0, 1).$$

则可以导出  $\rho$  的另一个区间估计

$$\left[ \frac{\frac{1+r}{1-r}e^{-a} - 1}{\frac{1+r}{1-r}e^{-a} + 1}, \frac{\frac{1+r}{1-r}e^a - 1}{\frac{1+r}{1-r}e^a + 1} \right],$$

其中  $a = \frac{2}{\sqrt{n}}Z_{1-\alpha/2}$ .

称  $Z = \frac{1}{2} \ln \frac{1+r}{1-r}$  为Fisher的 **Z变换**.

水泥在凝固时释放的热量与水泥中化学成分的关系案例(续)

设定置信水平  $1 - \alpha = 0.95$ .

(1) 变量  $x_3$  与  $x_4$  的相关系数  $\rho_{34}$  的置信区间为

$$\left[ \frac{\frac{1+0.030}{1-0.030}e^{-a} - 1}{\frac{1+0.030}{1-0.030}e^{-a} + 1}, \frac{\frac{1+0.030}{1-0.030}e^a - 1}{\frac{1+0.030}{1-0.030}e^a + 1} \right] = [-0.237, 0.293],$$

$$\text{其中 } a = \frac{2}{\sqrt{n}} Z_{1-\alpha/2} = \frac{2}{\sqrt{13}} \cdot 1.96 = 1.0872.$$

(2) 变量  $x_2$  与  $x_3$  的相关系数  $\rho_{23}$  的置信区间为  $[-0.390, 0.131]$ ,

(3) 变量  $x_1$  与  $x_3$  的相关系数  $\rho_{13}$  的置信区间为  $[-0.894, -0.716]$ .

注意: 置信区间变窄.

### 3.2.3 正态总体均值的置信域估计

#### 1) 单总体的情形

##### (1) 总体协方差 $\Sigma$ 已知的情形

由

$$n(\bar{X} - \mu)' \Sigma^{-1} (\bar{X} - \mu) \stackrel{d}{\sim} \chi^2(p),$$

则  $\mu$  的水平为  $(1 - \alpha)$  的置信域估计为

$$D = \{\mu \in R^p : n(\bar{X} - \mu)' \Sigma^{-1} (\bar{X} - \mu) \leq \chi_{1-\alpha}^2(p)\},$$

即有

$$Pr\{\mu \in D\} = 1 - \alpha.$$

## (2) $\Sigma$ 未知的情形

令

$$T^2 = n(n-1)(\bar{X} - \mu)'V^{-1}(\bar{X} - \mu),$$

由正态样本统计量的性质知

$$\sqrt{n}(\bar{X} - \mu) \stackrel{d}{\sim} N(0, \Sigma), \quad V \stackrel{d}{\sim} W_p(n-1, \Sigma), \quad \text{且 } \bar{X} \text{ 与 } V \text{ 独立},$$

因此有

$$T^2 = (n-1)(\sqrt{n}(\bar{X} - \mu))'V^{-1}(\sqrt{n}(\bar{X} - \mu)) \stackrel{d}{\sim} T_p^2(n-1),$$

$$\frac{1}{n-1}T^2 = n(\bar{X} - \mu)'V^{-1}(\bar{X} - \mu) \stackrel{d}{=} \frac{\chi^2(p)}{\chi^2(n-p)},$$

$$\frac{n-p}{(n-1)p}T^2 \stackrel{d}{=} \frac{\chi^2(p)/p}{\chi^2(n-p)/(n-p)} \stackrel{d}{\sim} F(p, n-p).$$

则当  $\Sigma$  未知时,  $\mu$  的水平为  $(1 - \alpha)$  的置信域估计为

$$D = \left\{ \mu \in R^p : \frac{n(n-p)}{p} (\bar{X} - \mu)' V^{-1} (\bar{X} - \mu) \leq F_{1-\alpha}(p, n-p) \right\},$$

即有

$$Pr\{\mu \in D\} = 1 - \alpha.$$



## 2) 两总体的情形

设独立总体  $X \stackrel{d}{\sim} N_p(\mu_1, \Sigma)$ ,  $Y \stackrel{d}{\sim} N_p(\mu_2, \Sigma)$ ,  $\mu_1, \mu_2 \in R^p$ ,  $\Sigma > 0$ .

记  $\mathbf{X} = (X_1, \dots, X_m)$ ,  $\mathbf{Y} = (Y_1, \dots, Y_n)$  分别为来自总体  $X$  和  $Y$  的样本,  $\min\{n, m\} > p$ .

我们要构造总体均值差  $\delta = \mu_1 - \mu_2$  的置信域估计.

### (1) $\Sigma$ 已知的情形

由  $\bar{X} \stackrel{d}{\sim} N_p(\mu_1, m^{-1}\Sigma)$ ,  $\bar{Y} \stackrel{d}{\sim} N_p(\mu_2, n^{-1}\Sigma)$ , 有

$$(\bar{X} - \bar{Y}) \stackrel{d}{\sim} N_p(\delta, (m^{-1} + n^{-1})\Sigma),$$

则

$$\frac{mn}{m+n} ((\bar{X} - \bar{Y}) - \delta)' \Sigma^{-1} ((\bar{X} - \bar{Y}) - \delta) \stackrel{d}{\sim} \chi^2(p).$$

由此得到  $\delta$  的水平为  $(1 - \alpha)$  的置信域估计为

$$D = \left\{ \delta \in R^p : \frac{mn}{m+n} ((\bar{X} - \bar{Y}) - \delta)' \Sigma^{-1} ((\bar{X} - \bar{Y}) - \delta) \leq \chi_{1-\alpha}^2(p) \right\}.$$

## (2) $\Sigma$ 未知的情形

记 $V_X$ 和 $V_Y$ 分别为总体 $X$ 和 $Y$ 的样本离差阵.

由 $X$ 和 $Y$ 的联合密度函数

$$(2\pi)^{-(m+n)p/2} |\Sigma|^{-(m+n)/2} \cdot \exp \left\{ -\frac{1}{2} \text{tr} [\Sigma^{-1} (V_X + V_Y + m(\bar{X} - \mu_1)(\bar{X} - \mu_1)' + n(\bar{Y} - \mu_2)(\bar{Y} - \mu_2)')] \right\},$$

知 $(\mu_1, \mu_2, \Sigma)$ 的极大似然估计为 $(\bar{X}, \bar{Y}, (V_X + V_Y)/(m+n))$ .

记 $V = V_X + V_Y$ , 并令

$$T^2 = \frac{mn(m+n-2)}{m+n} ((\bar{X} - \bar{Y}) - \delta)' V^{-1} ((\bar{X} - \bar{Y}) - \delta).$$

由于  $(\bar{X} - \bar{Y})$  与  $V$  相互独立, 且

$$\sqrt{\frac{mn}{m+n}}((\bar{X} - \bar{Y}) - \delta) \stackrel{d}{\sim} N_p(0, \Sigma), \quad V \stackrel{d}{\sim} W_p(m+n-2, \Sigma),$$

有  $T^2 \stackrel{d}{\sim} T_p^2(m+n-2)$ . 进而可知

$$\begin{aligned} & \frac{m+n-p-1}{(m+n-2)p} T^2 \\ &= \frac{(m+n-p-1)mn}{(m+n)p} ((\bar{X} - \bar{Y}) - \delta)' V^{-1} ((\bar{X} - \bar{Y}) - \delta) \\ &\stackrel{d}{\sim} F(p, m+n-p-1). \end{aligned}$$

由此得到  $\delta$  的水平为  $(1 - \alpha)$  的置信域估计为

$$D = \left\{ \delta \in R^p : \begin{aligned} & \frac{(m+n-p-1)mn}{(m+n)p} ((\bar{X} - \bar{Y}) - \delta)' V^{-1} ((\bar{X} - \bar{Y}) - \delta) \\ & \leq F_{1-\alpha}(p, m+n-p-1) \end{aligned} \right\}.$$

### 3.2.4 正态总体均值的Bayes估计

Bayes方法两要素:

- 1) 样本的密度函数形式已知, 只是参数未知
- 2) 参数也被视为随机变量, 其密度函数完全确定

称参数的密度为先验密度, 它是Bayes方法的核心.

Bayes先验密度(分布)的确定

- 1) 专业知识和专家经验
- 2) 数学方法, 如: 共轭先验、Jeffery先验、无信息先验等

## Bayes推断

基于后验密度(分布)

$$\text{后验密度} \propto \text{似然函数} \times \text{先验密度}$$

Bayes认为数据(样本)的似然函数为给定参数下的条件密度:  $f(X|\theta)$

数据(样本)和参数的联合密度:  $h(X, \theta) = f(X|\theta)\pi(\theta)$

数据(样本)的边缘密度:  $H(X) = \int f(X|\theta)\pi(\theta) d\theta$

Bayes后验密度为给定数据(样本)下参数的条件密度:

$$p(\theta|X) = \frac{h(X, \theta)}{H(X)} = \frac{f(X|\theta)\pi(\theta)}{H(X)} \propto f(X|\theta)\pi(\theta).$$

## 1) 逆Wishart分布

若  $X^{-1} \stackrel{d}{\sim} W_p(n, \Sigma)$ , 则称随机矩阵  $X$  服从逆Wishart分布, 记为  $X \stackrel{d}{\sim} IW_p(n, \Sigma^{-1})$ ,  $\Sigma > 0$ .

### 逆Wishart分布的密度函数

$$f_p(X) = \frac{|\Sigma|^{-n/2} |X|^{-(n+p+1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(\Sigma^{-1} X^{-1}) \right\}}{2^{(np/2)} \Gamma_p\left(\frac{n}{2}\right)}, \quad X > 0.$$

## 2) 多元正态分布参数 $(\mu, \Sigma)$ 的先验分布

$$\begin{cases} \Sigma \stackrel{d}{\sim} IW_p(\alpha, T), \\ \Sigma \text{ 给定的条件下, } \mu \stackrel{d}{\sim} N_p(\theta, k^{-1} \Sigma), \end{cases}$$

其中,  $\alpha, T, \theta$  和  $k$  为超参数.



### 3) 后验密度

$$\pi_1(\Sigma|X) = \frac{|T_n|^{\alpha_n/2} |\Sigma|^{-(\alpha_n+p+1)/2} \exp \left\{ -\frac{1}{2} \text{tr}(T_n \Sigma^{-1}) \right\}}{2^{(\alpha_n p/2)} \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma((\alpha_n - i + 1)/2)} = IW_p(\alpha_n, T_n),$$

$$\pi_2(\mu|\Sigma, X) = \frac{k_n^{p/2}}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{k_n}{2} (\mu - \theta_n)' \Sigma^{-1} (\mu - \theta_n) \right\} = N_p(\theta_n, k_n^{-1} \Sigma),$$

其中,

$$\alpha_n = n + \alpha, \quad T_n = T + V + \frac{kn(\bar{X} - \theta)(\bar{X} - \theta)'}{k + n},$$

$$\theta_n = \frac{n\bar{X} + k\theta}{n + k}, \quad k_n = n + k.$$

因此, 该先验分布是**共轭先验分布**.

注: 后验密度(分布)的形式一般都比较复杂

通常采用Markov chain Monte Carlo方法来计算.

#### 4) Bayes估计

Bayes估计 = 参数的后验均值

$\mu$  的Bayes估计:  $\hat{\mu}_B = \theta_n$ ,

$\Sigma$  的Bayes估计:  $\hat{\Sigma}_B = \frac{T_n}{\alpha_n - p - 1}$ . (性质8)

作业:

多元正态独立同分布样本下, 证明协方差矩阵的极大似然估计是强相合的.

(矩阵估计的相合性即每个元素对应相合)

拍照发到 fanyu16@mails.ucas.ac.cn