重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

均值向量的推断

INFERENCES ABOUT MEAN VECTORS

数学与统计学院 杨炜明

Introduction 引言

χ^2-分布和Wishart分布

定义1 设 $X_1, X_2, \cdots, X_n$ 为 相互独立且同服从于 $N(0,1)$ 分布的随机变量。则

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

所服从的分布叫做χ^2-分布，n 称为自由度且记为 $\chi^2 \sim \chi^2(n)$ 。

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

Introduction 引言

定理．设 $X_1 \sim \chi^2(n_1)$ $X_2 \sim \chi^2(n_2)$，

且 X_1 与 X_2 相互独立，则

$$X_1 + X_2 \sim \chi^2(n_1 + n_2)$$

Introduction 引言

Wishart分布

定义1 设 $X_1, X_2, \cdots, X_n$ 为相互独立且同服从于分布 $N_p(0, \Sigma)$，令 $\boldsymbol{X} = [X_1, X_2, \cdots, X_n]'$ 则

$$W = XX' = \sum_{i=1}^{n} X_i X_i'$$

所服从的分布叫做 自由度为**n**的**p**维-维希特分布，记作 $W \sim W_p(n, \Sigma)$

Introduction 引言

$t-$分布与 T^2 分布

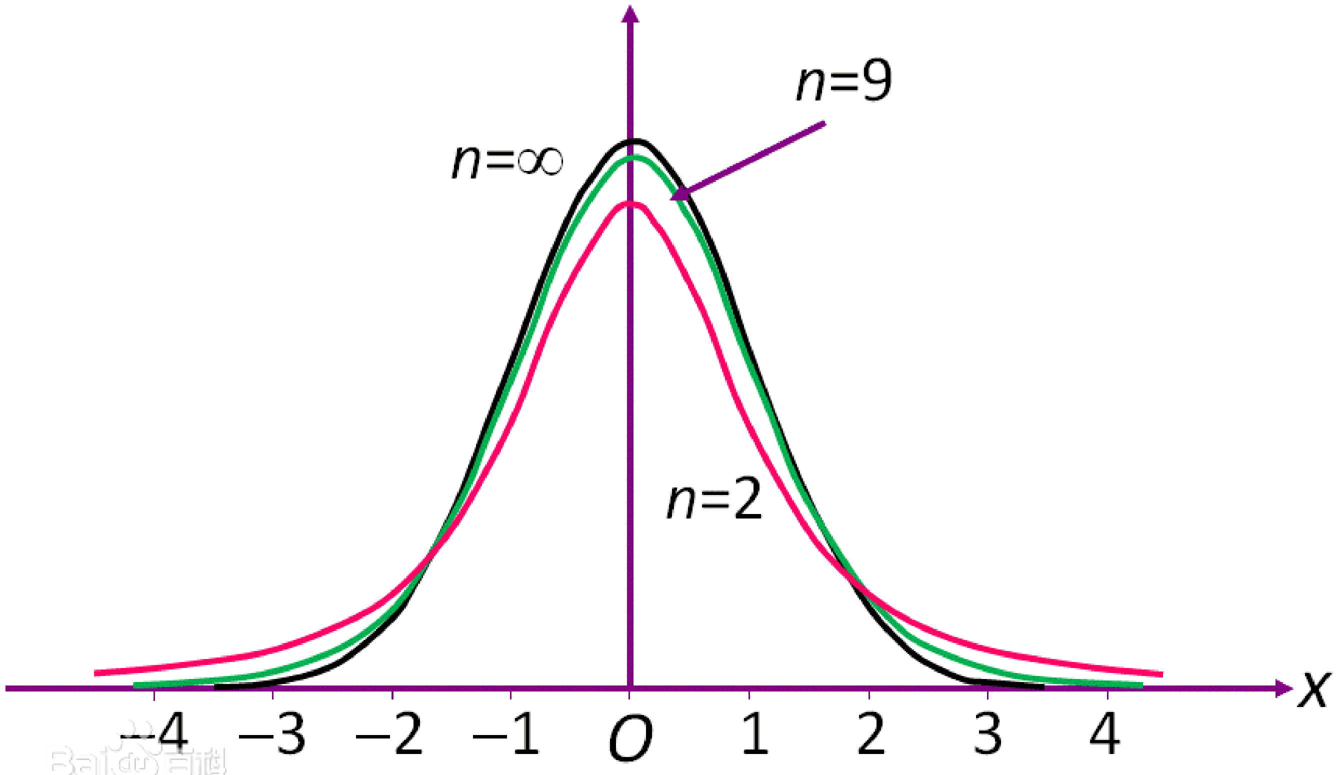
设 $X \sim N(0,1)$ $Y \sim \chi^2(n)$ ，且 X 与 Y 相互独立，则称随机变量

$T = \dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t- 分布，

记为 $T \sim t(n)$ 。

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

$$f(x)=\frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\,\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2}$$



重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

Introduction 引言

将*T*平方，即 $T^2 = n\frac{X^2}{Y} = nX'Y^{-1}X$

在多元统计中 T^2 分布是一元统计中*t*分布的推广

定理：若 $S \sim W_p(n,\Sigma)$，$X \sim N_p(0,\Sigma)$

S与X相互独立、称随机变量

$$T^2 = nX'S^{-1}X$$

是自由度为（**p**，**n**）的 Hotelling T^2 分布

T^2 可以转化为**F**分布

$$\frac{n-p+1}{np}T^2 \sim F(p, n-p+1)$$

重庆工商大学
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

Introduction 引言

F-分布

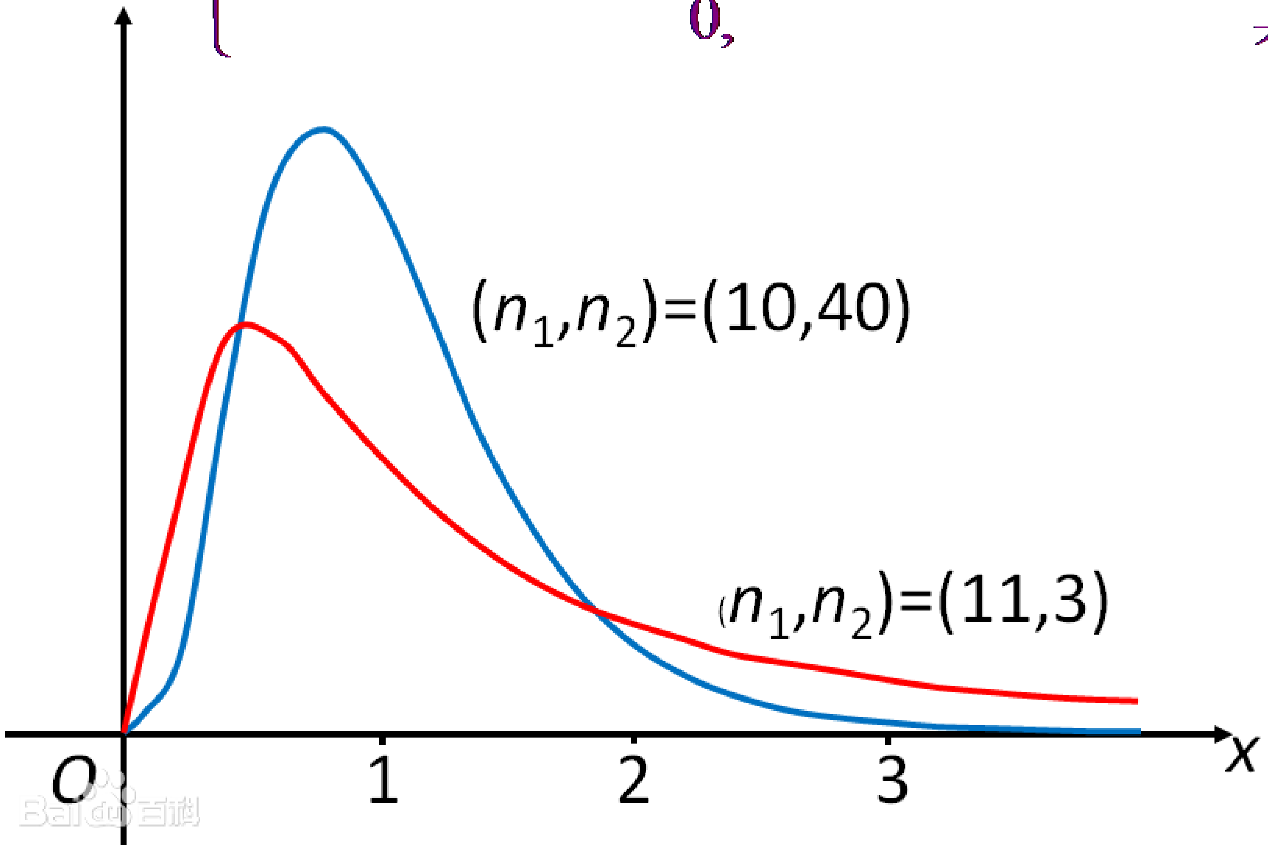
定义3　设 $X \sim \chi^2(m)$，$Y \sim \chi^2(n)$，且 X 与 Y 相互独立，则称随机变量

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

服从自由度为 (m,n) 的 F-分布，记为 $F \sim F(m,n)$。

重慶工商大學 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

$$f(x)=\begin{cases}\dfrac{\Gamma[(n_1+n_2)/2](n_1/n_2)^{n_1/2}x^{(n_1/2)-1}}{\Gamma(n_1/2)\Gamma(n_2/2)[1+(n_1x/n_2)]^{(n_1+n_2)/2}}, & x>0,\\ 0, & \text{其它}.\end{cases}$$



一个正态总体均值向量的检验

（1）协方差 / 协方差矩阵已知

- 单一变量情形

糖厂生产的袋装糖的重量是一个随机变量，当机器正常工作时，其均值为0.5kg，已知标准差为0.015kg。某日开工后，为检验包装机是否正常工作，随机地抽取它所包装的糖9袋，称得净重为(kg)：0.497, 0.506, 0.524, 0.518, 0.498, 0.511, 0.520, 0.515, 0.512，问机器是否正常工作?

一个正态总体均值向量的检验

$$H_0: \mu = 0.5 \leftrightarrow H_1: \mu \neq 0.5$$

构造检验统计量

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

决定 H_0 的拒绝域

$$|u| > Z_{\alpha/2}$$

$$|u| = \left|\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\right| = \left|\frac{0.511 - 0.5}{0.015/\sqrt{9}}\right| = 2.2 > Z_{0.975} = 1.96$$

拒绝 H_0

重庆工商大学
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

若$\boldsymbol{X} \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$，$\boldsymbol{\Sigma} > 0$，则 $\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} \sim \chi^2(p)$

证:因为$\Sigma > 0$,由于正定矩阵可以分解为$\Sigma = CC'$ $(C$为非退化矩阵$)$,令$Y = C^{-1}X$,即$X = CY$.则

$$Y \sim N_p\left(0, C^{-1}\Sigma\left(C^{-1}\right)'\right)$$

因$\Sigma = CC'$，所以$Y \sim N_p(0, I_p)$,且有

$$X'\Sigma^{-1}X = Y'C'\Sigma^{-1}CY = Y'Y \sim \chi^2(p)$$

重庆工商大学
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

一个正态总体均值向量的检验

- 多元情形 $H_0: \mu=\mu_0$ vs $H_1: \mu \neq \mu_0$

$$X \sim N(\mu_0,\sigma^2) \qquad \bar{X} \sim N(\mu_0,\frac{\sigma^2}{n})$$

$$\Rightarrow u^2 = n\left(\bar{X} - \mu_0\right)\sigma^{-1}\left(\bar{X} - \mu_0\right) \sim \chi^2(1)$$

$$X \sim N_p\left(\mu,\Sigma\right) \quad \bar{X} \sim N_p\left(\mu,\frac{1}{n}\Sigma\right)$$

$$\Rightarrow \chi_0^2 = n\left(\bar{X} - \mu\right)'\Sigma^{-1}\left(\bar{X} - \mu\right) \sim \chi^2\left(p\right)$$

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

一个正态总体均值向量的检验

设 $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_\mathrm{n}$ 是取自多元正态总体 $N_P(\mu, \Sigma)$ 的一个样本，现欲检验

$$H: \mu = \mu_0 \quad vs \quad H_1: \mu \neq \mu_0$$

总体协方差矩阵已知

$$\chi_0{}^2 = \left(\bar{\mathbf{x}} - \mu_0\right)^T \left(\frac{1}{n}\Sigma\right)^{-1} \left(\bar{\mathbf{x}} - \mu_0\right)$$

$$= n\left(\bar{\mathbf{x}} - \mu_0\right)^T \Sigma^{-1} \left(\bar{\mathbf{x}} - \mu_0\right)$$

重慶工商大學

CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

T_0^2 服从自由度为p的卡方分布，即 $T_0^2 \sim \chi^2(p)$。

可作为检验统计量，对给定的显著性水平 α ，

检验规则为

当 $T_0^2 \geq \chi_\alpha^2(p)$时，拒绝$H_0$;

当 $T_0^2 < \chi_\alpha^2(p)$时，接受H_0。

其中 $\chi_\alpha^2(p)$ 是 $\chi^2(p)$ 的上 α 分位点。

一个正态总体均值向量的检验

测定稻谷每亩穗数X1,每穗粒数X2,每亩稻谷产量X3,

$$\boldsymbol{X}=\begin{pmatrix}X_1\\X_2\\X_3\end{pmatrix}\sim N_3(\mu,\Sigma),\qquad \Sigma=\begin{bmatrix}6&-9&79\\-9&25&3\\79&3&2354\end{bmatrix}$$

X1	26.7	31.3	30.4	33.9	34.6	33.8	30.4	27	33.3	30.4	31.5	33	34
X2	73.4	59	65.9	58.2	64.6	64.6	62.1	71.4	64.5	64.1	61.1	56	59.8
X3	1008	959	1051	1022	1097	1103	992	945	1074	1029	1004	995	1045

检验均值是否等于 $\begin{pmatrix}32\\63\\1025\end{pmatrix}$

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

一个正态总体均值向量的检验

$$\bar{X}=\begin{pmatrix}31.5615\\63.4385\\1024.92\end{pmatrix},\quad \bar{X}-\mu_0=\begin{pmatrix}31.5615-32\\63.4385-63\\1024.92-1025\end{pmatrix}=\begin{pmatrix}-0.4385\\0.4385\\-0.0769\end{pmatrix}$$

$$\chi_0^2=n\left(\bar{X}-\mu\right)'\Sigma^{-1}\left(\bar{X}-\mu\right)$$

$$=13\times\begin{pmatrix}-0.4385 & 0.4385 & -0.0769\end{pmatrix}\begin{bmatrix}6 & -9 & 79\\-9 & 25 & 3\\79 & 3 & 2354\end{bmatrix}^{-1}\begin{pmatrix}-0.4385\\0.4385\\-0.0769\end{pmatrix}$$

$$=28.15$$

$$\chi_0^2>\chi_{0.05}^2(3)=7.81,\quad \text{reject}\ \ H_0$$

一个正态总体均值向量的检验

（2）协方差 / 协方差矩阵未知

例. 设某杨树品种在甲地区种植，今将此品种移植到乙地区20株，10年后测得这20株树的平均高为 $\bar{x}=15$ 标准差为 $s=2.4$。试以显著水平 $\alpha=0.05$ 检验移植到乙地区种植10年后树高总体的平均是否与17m显著不同。假定乙地区树高服从正态分布。

正态总体,方差未知 $H_0:\mu=\mu_0 \leftrightarrow H_1:\mu\neq\mu_0$

重庆工商大学
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

一个正态总体均值向量的检验

解： $H_0: \mu = 17 \leftrightarrow H_1: \mu \neq 17$

$n = 20$ 为小样本，σ未知，

计算检验统计量的值：

$$T = \frac{\bar{y} - \mu_0}{s/\sqrt{n}} = \frac{15 - 17}{2.4/\sqrt{20}} = -3.727$$

查临界值：$\alpha = 0.05,\ t_{1-\alpha/2}(n-1) = t_{0.975}(19) = 2.093$

$T = -3.727 < -T_{0.975}(19) = -2.093$ 所以拒绝 H_0

重慶工商大學 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

一个正态总体均值向量的检验

$$X \sim N_p(\mu,\Sigma) \quad \bar{X} \sim N_p\left(\mu,\frac{1}{n}\Sigma\right)$$

$$\Rightarrow T^2 = n\left(\bar{X}-\mu\right)'\Sigma^{-1}\left(\bar{X}-\mu\right) \sim \chi^2(p)$$

$$T^2 = n\left(\bar{X}-\mu\right)'\hat{\Sigma}^{-1}\left(\bar{X}-\mu\right)$$

$$= n\left(\bar{X}-\mu\right)'\left(\frac{L}{n-1}\right)^{-1}\left(\bar{X}-\mu\right)$$

$$= n\left(\bar{X}-\mu\right)'S^{-1}\left(\bar{X}-\mu\right) \sim ?$$

Hotelling T^2分布

$$\frac{n-p}{(n-1)p}T^2 \sim F(p,n-p)$$

重庆工商大学 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

当原假设 H_0 为真时，统计量 $\frac{n-p}{p(n-1)}T^2$ 服从自由度为 p 和 $n-p$ 的 F 分布。

对给定的显著性水平 α，检验规则为:

当 $\frac{n-p}{p(n-1)}T^2 \geq F_\alpha(p,n-p)$ 时，拒绝 H_0；

当 $\frac{n-p}{p(n-1)}T^2 < F_\alpha(p,n-p)$ 时，接受 H_0。

其中 $F_\alpha(p,n-p)$ 是 $F(p,n-p)$ 的上 α 分位点。

一个正态总体均值向量的检验

供试者	1	2	3	4	5	6	7	8
血糖变化	30	90	－10	10	30	60	0	40
收缩压变化	－8	7	－2	0	－2	0	－2	1
舒张压变化	－1	6	4	2	5	3	4	2

设总体$x \sim N_3(\mu, \Sigma)$,今欲检验假设

$$H_0: \mu = \mu_0 = 0, \quad H_1: \mu \neq 0$$

一个正态总体均值向量的检验

经计算得，$\bar{x}=\begin{bmatrix}31.25\\-0.75\\3.125\end{bmatrix}$， $S=\begin{bmatrix}1069.64 & 82.5 & 16.964\\82.5 & 17.357 & 6.393\\16.964 & 6.393 & 4.696\end{bmatrix}$

$$T^2=n\bar{x}'S^{-1}\bar{x}=79.064,$$

$$\frac{n-p}{(n-1)p}T^2=\frac{8-3}{(8-1)\times 3}\times 79.064=18.825>5.41=F_{0.05}(3,5)$$

检验表明，H_0被拒绝，即服药后，血糖和血压的变化是显著的。

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

一个正态总体均值向量的检验

测定稻谷每亩穗数X1,每穗粒数X2,每亩稻谷产量X3，

$$X=\begin{pmatrix}X_1\\X_2\\X_3\end{pmatrix}\sim N_3(\mu,\Sigma)$$

X1	26.7	31.3	30.4	33.9	34.6	33.8	30.4	27	33.3	30.4	31.5	33	34
X2	73.4	59	65.9	58.2	64.6	64.6	62.1	71.4	64.5	64.1	61.1	56	59.8
X3	1008	959	1051	1022	1097	1103	992	945	1074	1029	1004	995	1045

检验均值是否等于 $\begin{pmatrix}32\\63\\1025\end{pmatrix}$

一个正态总体均值向量的检验

$$\bar{X}=\begin{pmatrix}31.5615\\63.4385\\1024.92\end{pmatrix},\quad \bar{X}-\mu_0=\begin{pmatrix}31.5615-32\\63.4385-63\\1024.92-1025\end{pmatrix}=\begin{pmatrix}-0.4385\\0.4385\\-0.0769\end{pmatrix}$$

$$T^2=n\left(\bar{X}-\mu\right)'\hat{\Sigma}^{-1}\left(\bar{X}-\mu\right)$$

$$=13\times\left(-0.4385\quad 0.4385\quad -0.0769\right)\begin{pmatrix}6.6 & -9.11 & 78.90\\-9.11 & 24.66 & 3.24\\78.90 & 3.24 & 2353.74\end{pmatrix}^{-1}\begin{pmatrix}-0.4385\\0.4385\\-0.0769\end{pmatrix}$$

$$=1.97$$

$$\frac{n-p}{(n-1)p}T^2=\frac{13-3}{(13-1)3}\times 1.97=0.5463<F_{0.05}(3,10)=3.71$$

不拒绝 $\mathbf{H}_0$

置信域

$$T^2 = n(\bar{\mathbf{x}}-\mu)^T S^{-1}(\bar{\mathbf{x}}-\mu)$$

$\frac{(n-p)}{p(n-1)}T^2$ 服从自由度为 p 和 $n-p$ 的 F 分布，

即 $\frac{(n-p)}{p(n-1)}T^2 \sim F(p,n-p)$ ，从而

$$P\left(\frac{(n-p)}{p(n-1)}T^2 \le F_\alpha(p,n-p)\right) = 1-\alpha$$

重慶工商大學 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

即 $$P\left(n\left(\bar{\mathbf{x}}-\mu\right)^T S^{-1}\left(\bar{\mathbf{x}}-\mu\right)\le T_\alpha^{\ 2}\right)=1-\alpha$$

其中 $$T_\alpha^{\ 2}=\frac{p(n-1)}{n-p}F_\alpha(p,n-p)$$

由此得到 μ 的置信度为 $1-\alpha$ 的置信区间（confidence region）为

$$\left\{\mu : n\left(\bar{\mathbf{x}}-\mu\right)^T S^{-1}\left(\bar{\mathbf{x}}-\mu\right)\le T_\alpha^{\ 2}\right\}$$

当 $p=1$ 时，它是一个区间；当 $p=2$ 时，它是一个椭圆，这时可将其在坐标平面上画出；当 $p=3$ 时，它是一个椭球；当 $p>3$ 时，它是一个超椭球。

联合置信区间

- 联合 T^2 置信区间:

$$a'\bar{x}-T_\alpha\sqrt{a'Sa}/\sqrt{n}\leqslant a'\mu\leqslant a'\bar{x}+T_\alpha\sqrt{a'Sa}/\sqrt{n} \tag{4.2.13}$$

以 $1-\alpha$ 的概率对一切 $a\in R^p$ 成立，称它为一切线性组合 $\{a'\mu, a\in R^p\}$ 的置信度为 $1-\alpha$ 的**联合置信区间**(simultaneous confidence intervals)。

- 邦弗伦尼置信区间:

$$a_i'\bar{x}-t_{\alpha/2k}(n-1)\sqrt{a_i'Sa_i}/\sqrt{n}\leqslant a_i'\mu\leqslant a_i'\bar{x}+t_{\alpha/2k}(n-1)\sqrt{a_i'Sa_i}/\sqrt{n}$$

$$i=1,2,\cdots,k \tag{4.2.16}$$

两个区间的比较

- 若 $t_{\alpha/2k}(n-1)\le T_\alpha$ ，则邦弗伦尼区间比T^2区间要窄，这时宜采用前者作为联合置信区间；反之，若 $t_{\alpha/2k}(n-1)>T_\alpha$ ，则邦弗伦尼区间比T^2 区间宽，宜采用后者作为联合置信区间。
- 当$k=p$时，邦弗伦尼区间要比T^2 区间窄。故在求**μ**的所有p个分量 $\mu_1,\mu_2,\cdots,\mu_p$ 的联合置信区间时，应采用邦弗伦尼区间。

例 4.2.2 为评估某职业培训中心的教学效果，随机抽取 8 名受训者，进行甲和乙两个项目的测试，其数据列于表 4.2.2。假定 $\boldsymbol{x}=(x_1,x_2)'$ 服从二元正态分布。

表 4.2.2 两个项目的测试成绩

编号	1	2	3	4	5	6	7	8
甲项成绩 x_1	62	80	66	84	75	80	54	79
乙项成绩 x_2	70	77	75	87	87	91	61	84

该例中，$n=8$，$p=2$，取 $\alpha=0.1$，查 F 分布表得，$F_{0.1}(2,6)=3.46$，于是 $T_{0.1}^2=\frac{2\times 7}{6}F_{0.1}(2,6)=8.073$，$T_{0.1}=2.841$。经计算得

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

$$\bar{\boldsymbol{x}}=\begin{pmatrix}72.5\\79\end{pmatrix},\qquad \boldsymbol{S}=\begin{pmatrix}112.5714 & 96.1429\\96.1429 & 103.1429\end{pmatrix}$$

$$\boldsymbol{S}^{-1}=\begin{pmatrix}0.0436 & -0.0406\\-0.0406 & 0.0475\end{pmatrix}$$

由(4.2.11)式,$\boldsymbol{\mu}$ 的 0.90 置信区域为

$$8\times(72.5-\mu_1,79-\mu_2)\begin{pmatrix}0.0436 & -0.0406\\-0.0406 & 0.0475\end{pmatrix}\begin{pmatrix}72.5-\mu_1\\79-\mu_2\end{pmatrix}\leqslant 8.073$$

也就是

$$0.0436\times(\mu_1-72.5)^2-0.0812\times(\mu_1-72.5)(\mu_2-79)+0.0475\times(\mu_2-79)^2\leqslant 1.009$$

这是一个椭圆区域,如图 4.2.1 所示。μ_1 和 μ_2 的 0.90 联合 T^2 置信区间为

$$72.5-2.841\times\sqrt{112.5714/8}\leqslant\mu_1\leqslant 72.5+2.841\times\sqrt{112.5714/8}$$

$$79-2.841\times\sqrt{103.1429/8}\leqslant\mu_2\leqslant 79+2.841\times\sqrt{103.1429/8}$$

重慶工商大學 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

即

$$61.84 \leqslant \mu_1 \leqslant 83.16$$

$$68.80 \leqslant \mu_2 \leqslant 89.20$$

这两个区间分别正是椭圆在 μ_1 轴和 μ_2 轴上的投影。而 μ_1 和 μ_2 的

0.90 邦弗伦尼联合置信区间为(查表得,$t_{0.025}(7)=2.3646$)

$$72.5-2.3646\times\sqrt{112.5714/8} \leqslant \mu_1 \leqslant 72.5+2.3646\times\sqrt{112.5714/8}$$

$$79-2.3646\times\sqrt{103.1429/8} \leqslant \mu_2 \leqslant 79+2.3646\times\sqrt{103.1429/8}$$

即

$$63.63 \leqslant \mu_1 \leqslant 81.37$$

$$70.51 \leqslant \mu_2 \leqslant 87.49$$

显然,这个联合置信区间在精确度方面要好于 T^2 联合置信区间。由该联合置信区间可得到置信度至少为 0.90 的矩形置信区域(见图 4.2.1中的实践矩形),但其矩阵面积要大于椭圆面积。

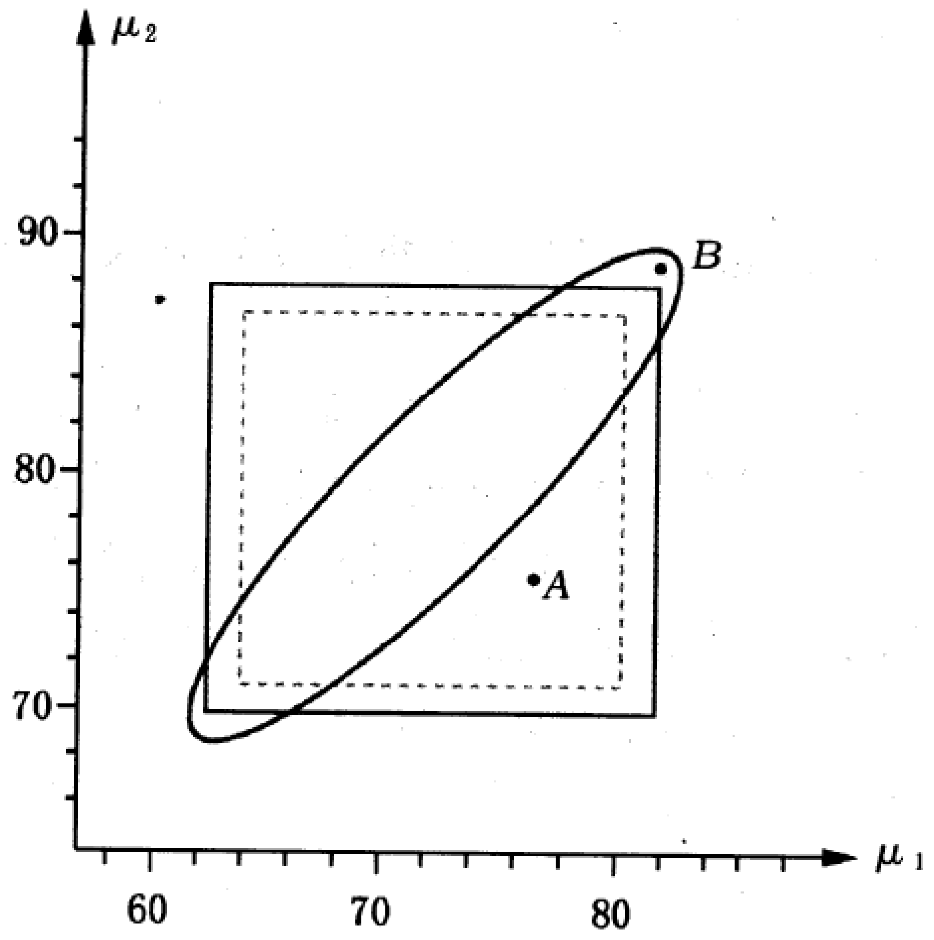


图 4.2.1　置信椭圆和联合置信区间

单个总体均值分量间结构关系的检验

$$H_0: \boldsymbol{C\mu}=\boldsymbol{\varphi}, \quad H_1: \boldsymbol{C\mu}\neq\boldsymbol{\varphi} \tag{4.3.1}$$

其中 $\boldsymbol{C}$ 为一已知的 $k\times p$ 矩阵，$k<p$，$\operatorname{rank}(\boldsymbol{C})=k$，$\boldsymbol{\varphi}$ 为已知的 k 维向量。

检验统计量为

$$T^2=n(\boldsymbol{C}\bar{\boldsymbol{x}}-\boldsymbol{\varphi})'(\boldsymbol{CSC}')^{-1}(\boldsymbol{C}\bar{\boldsymbol{x}}-\boldsymbol{\varphi})$$

拒绝规则为：

$$\text{若 } T^2\geq T_\alpha^2\text{，则拒绝 } H_0 \tag{4.3.5}$$

其中

$$T_\alpha^2=\frac{k(n-1)}{n-k}F_\alpha(k,n-k) \tag{4.3.6}$$

$F_\alpha(k,n-k)$ 是 $F(k,n-k)$ 的上 α 分位点。

特别地，若(4.3.1)式中的 $\boldsymbol{\varphi}=\mathbf{0}$，即欲检验

$$H_0: \boldsymbol{C\mu}=\mathbf{0}, \quad H_1: \boldsymbol{C\mu}\neq\mathbf{0} \tag{4.3.7}$$

则(4.3.3)式的 T^2 可简化为

$$T^2=n\bar{\boldsymbol{x}}'\boldsymbol{C}'(\boldsymbol{CSC}')^{-1}\boldsymbol{C}\bar{\boldsymbol{x}} \tag{4.3.8}$$

例 4.3.1 设 $\boldsymbol{x}\sim N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$，$\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_p)'$，$\boldsymbol{\Sigma}>0$，$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是取自该总体的一个样本，欲检验

$$H_0:\mu_1=\mu_2=\cdots=\mu_p,\quad H_1:\mu_i\neq\mu_j,\text{至少存在一对 } i\neq j$$

令

$$\boldsymbol{C}=\begin{pmatrix}1 & -1 & 0 & \cdots & 0\\ 1 & 0 & -1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & 0 & 0 & \cdots & -1\end{pmatrix}$$

则上面的假设可表达为

$$H_0:\boldsymbol{C\mu}=\boldsymbol{0},\quad H_1:\boldsymbol{C\mu}\neq\boldsymbol{0}$$

由于 $\boldsymbol{C}$ 为 $(p-1)\times p$ 矩阵，且 $\text{rank}(\boldsymbol{C})=p-1$，因此检验统计量为

$$T^2=n\bar{\boldsymbol{x}}'\boldsymbol{C}'(\boldsymbol{CSC}')^{-1}\boldsymbol{C}\bar{\boldsymbol{x}}$$

拒绝规则为：

$$\text{若 } T^2\geqslant T_\alpha^2,\text{ 则拒绝 } H_0$$

其中

$$T_\alpha^2=\frac{(p-1)(n-1)}{n-p+1}F_\alpha(p-1,n-p+1)$$

重慶工商大學 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

两个正态总体均值向量的检验

（1）协方差矩阵已知

研究株距对树高的影响。为此设计两种株距0.5米和1米，假设两种株距下的树高服从正态分布，标准差分别为2.3^2，1.8^2。7年后分别抽取9株和6株测量，测量结果为：$\bar{X}=24.02, \bar{Y}=28.93$

两种株距对树高有无显著影响？

重庆工商大学
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

两个正态总体均值向量的检验

解 假设检验类型为： $\mathbf{H_0}: \mu_X - \mu_Y = \mathbf{0} \leftrightarrow \mathbf{H_1}: \mu_X - \mu_Y \neq \mathbf{0}$

计算检验统计量的值

$$\mathrm{U} = \frac{24.02 - 28.93}{\sqrt{\frac{2.3^2}{9} + \frac{1.8^2}{6}}} = \frac{-4.91}{\sqrt{1.128}} = -4.6$$

$\mathrm{U} = -4.6 < -\mathrm{U}_{0.975} = -1.96$ 拒绝 H_0

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

两个正态总体均值向量的检验

当$X_{(\alpha)} \sim N_p(\mu_X, \Sigma_X)(\alpha = 1, 2, \cdots, n)$相互独立，则$\bar{X} \sim N_p\left(\mu_X, \frac{1}{n}\Sigma_X\right)$;

当$Y_{(\alpha)} \sim N_p(\mu_Y, \Sigma_Y)(\alpha = 1, 2, \cdots, m)$相互独立，则$\bar{Y} \sim N_p\left(\mu_Y, \frac{1}{m}\Sigma_Y\right)$;

$$\left(\bar{X} - \bar{Y}\right) \sim N_p\left(\mu_X - \mu_Y, \left(\frac{1}{n}\Sigma_X + \frac{1}{m}\Sigma_Y\right)\right)$$

$$\mathrm{H}_0 : \mu_X - \mu_Y = 0 \leftrightarrow \mathrm{H}_1 : \mu_X - \mu_Y \neq 0$$

$$T^2 = \left(\bar{X} - \bar{Y}\right)'\left(\frac{1}{n}\Sigma_X + \frac{1}{m}\Sigma_Y\right)^{-1}\left(\bar{X} - \bar{Y}\right) \sim \chi^2(p)$$

If $T^2 > \chi^2_\alpha(p)$, reject H_0

重庆工商大学
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

两个正态总体均值向量的检验

（2）协方差矩阵未知但相等

研究株距对树高的影响。为此设计两种株距0.5米和1米，假设两种株距下的树高服从正态分布,且方差相等。7年后分别抽取9株和6株测量树高数据,两种株距对树高有无显著影响?

$$\bar{X} = 24.02, \quad \bar{Y} = 28.93 \qquad S_X^2 = 2.342^2, \quad S_Y^2 = 1.825^2$$

X	22.1	24.4	24.3	28.8	23.3	22.0	21.0	25.8	24.5
Y	26.8	31.3	28.8	28.9	21.7	30.7			

两个正态总体均值向量的检验

解 假设检验类型为： $H_0: \mu_X - \mu_Y = 0 \leftrightarrow H_1: \mu_X - \mu_Y \neq 0$

计算检验统计量的值

$$T = \frac{24.02 - 28.93}{\sqrt{\frac{(9-1)\times 2.342^2 + (6-1)\times 1.825^2}{9+6-2}}\sqrt{\frac{1}{9}+\frac{1}{6}}} = -4.318$$

$T = -4.318 < -T_{0.975}(13) = -2.16$ 拒绝 H_0

重庆工商大学 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

$$L_X=\sum_{\alpha=1}^{n_1}\left(X_{(\alpha)}-\bar{X}\right)\left(X_{(\alpha)}-\bar{X}\right)'\sim W_p\left(n_1-1,\Sigma\right)$$

$$L_Y=\sum_{\alpha=1}^{n_2}\left(Y_{(\alpha)}-\bar{Y}\right)\left(Y_{(\alpha)}-\bar{Y}\right)'\sim W_p\left(n_2-1,\Sigma\right)$$

记$L=\sum_{\alpha=1}^{n_1}(x_\alpha-\bar{x})(x_\alpha-\bar{x})'+\sum_{\alpha=1}^{n_2}(y_\alpha-\bar{y})(y_\alpha-\bar{y})'$

$$\boldsymbol{L\sim W_p(n_1+n_2-2,\Sigma\ )}$$

重慶工商大學 CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

两个正态总体均值向量的检验

$$\bar{X} \sim N_p\left(\mu_X, \frac{1}{n_1}\Sigma\right), \ \bar{Y} \sim N_p\left(\mu_Y, \frac{1}{n_2}\Sigma\right),$$

$$\Rightarrow \left(\bar{X} - \bar{Y}\right) \sim N_p\left(0, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma\right)$$

$$\left(\bar{X} - \bar{Y}\right)'\left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right)S_{pooled}\right)^{-1}\left(\bar{X} - \bar{Y}\right) \sim T^2(p, n_1 + n_2 - 2)$$

其中 $S_{pooled} = \frac{n_1 - 1}{n_1 + n_2 - 2}S_1 + \frac{n_2 - 1}{n_1 + n_2 - 2}S_2$

$$\frac{\left(n_1 + n_2 - 2\right) - p + 1}{p\left(n_1 + n_2 - 2\right)}T^2 \sim F\left(p, \left(n_1 + n_2 - 2\right) - p + 1\right)$$

重慶工商大學
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

两个正态总体均值向量的检验

某品种小麦，在甲、乙两地各布30个点作试验，并取得了6个主要性状的数据，

$$\bar{X}=\begin{pmatrix}401.4\\37.5667\\27.6667\\38.3000\\32.8333\\32.1000\end{pmatrix},\quad L_X=\begin{bmatrix}789477.201 & & & & & \\ 41178.201 & 13679.367 & & & & \\ 38490.000 & 2004.666 & 7004.667 & & & \\ -14429.601 & -836.100 & 308.001 & 2570.301 & & \\ -8606.001 & -6576.168 & -1115.667 & 1848.501 & 6394.167 & \\ 5252.799 & 5544.300 & 2780.001 & 32.100 & -2773.500 & 5700.699\end{bmatrix}$$

$$\bar{Y}=\begin{pmatrix}355.1667\\62.7333\\12.4667\\21.1667\\8.2667\\22.3667\end{pmatrix},\quad L_Y=\begin{bmatrix}437592.168 & & & & & \\ 20580.333 & 11665.866 & & & & \\ 7875.666 & -676.266 & 1213.467 & & & \\ -6410.832 & -133.668 & -100.332 & 414.168 & & \\ -637.332 & 183.132 & -47.733 & 33.666 & 45.867 & \\ -3087.834 & 1698.933 & 702.867 & 190.167 & 144.066 & 2638.968\end{bmatrix}$$

重庆工商大学

CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY

两个正态总体均值向量的检验

$$\frac{L_X + L_Y}{30+30-2} = \begin{bmatrix} 21156.3684 & & & & & \\ 1064.8023 & 436.9868 & & & & \\ 799.4080 & 22.9035 & 136.5196 & & & \\ -359.3178 & -16.7201 & 3.5805 & 51.4564 & & \\ -159.3678 & -110.2248 & -20.0586 & 32.4512 & 111.0351 & \\ 37.3270 & 124.8823 & 60.0495 & 3.8322 & -45.3351 & 143.7874 \end{bmatrix}$$

$$F = \frac{(n_1+n_2-2)-p+1}{p(n_1+n_2-2)} T^2 = 20.63 > F_{0.01}(6,53) = 3.17$$

两个正态总体均值向量的检验

例：为研究某种疾病，对两组患者测量了 $x_1:\beta$ 脂蛋白，x_2：甘油三脂，x_3：α 脂蛋白，x_4：前 β 脂蛋白等四个指标的数据。这里第一组为20－25岁的女性患者，第二组为20－25岁男性患者，每一组样本数均为20人，试检验两组间有无显著差异？

设第一组总体为$x \sim N_4(\mu_1, \Sigma)$,第二组总体为$y \sim N_4(\mu_2, \Sigma)$

欲检验假设： $H_0: \mu_1 = \mu_2$， $H_1: \mu_1 \neq \mu_2$

两个正态总体均值向量的检验

$$\bar{x}=\begin{bmatrix}231.0\\89.6\\32.9\\17.1\end{bmatrix},\qquad \bar{y}=\begin{bmatrix}253.5\\72.55\\32.45\\17.9\end{bmatrix},$$

$$L_X=\begin{bmatrix}30530&&&\\6298&15736.8&&\\-1078&-796.8&955&\\198&1387.8&90.2&413.8\end{bmatrix},\qquad L_Y=\begin{bmatrix}51705&&&\\7021.5&12288.95&&\\-1571.5&-807.95&364.95&\\827&321.1&-5.1&133.8\end{bmatrix}$$

$$\bar{x}-\bar{y}=\begin{bmatrix}-22.5\\17.05\\0.45\\-0.8\end{bmatrix},\qquad S_p=\begin{bmatrix}2164.08&&&\\350.5&737.52&&\\-69.72&-42.23&34.74&\\26.97&44.97&2.24&14.41\end{bmatrix}$$

$$T^2=\frac{n_1n_2}{n_1+n_2}(\bar{x}-\bar{y})'S_p^{-1}(\bar{x}-\bar{y})=11.36248$$

两个正态总体均值向量的检验

H_0的拒绝域是：

$$T^2=\frac{n_1n_2}{n_1+n_2}(\bar{x}-\bar{y})'S_p^{-1}(\bar{x}-\bar{y})>\frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_\alpha(p,n_1+n_2-p-1)$$

$$F_{0.05}(4,35)=2.641464,\quad T_0^2=\frac{38\times4}{35}\times2.641464=11.4715$$

$$T^2=\frac{n_1n_2}{n_1+n_2}(\bar{x}-\bar{y})'S_p^{-1}(\bar{x}-\bar{y})=11.36248<11.4715=T_0^2$$

检验结果表明两组间从四个指标的整体看，在0.05的水平下无显著差异。

重庆工商大学
CHONGQING TECHNOLOGY AND BUSINESS UNIVERSITY