高等工程數學（6）

厦門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黄 煒
博士，副教授
Dr. Wei Huang

实二次型

矩阵论（6）

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院）School of Informatics
黄 炜 博士·副教授 Dr. Wei Huang

1.6 实二次型

二次型的研究起源于化二次曲线和二次曲面的方程为标准形式的问题. 其理论在数学、力学和工程技术中都有应用.

一般的一个实二次型是指 n 个变量 $x_1,\cdots,x_n$ 的二次齐次多项式

$$\begin{aligned} f(x_1,x_2,\cdots,x_n) &= a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots \\ &\quad + 2a_{1n}x_1x_n + a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots \\ &\quad + 2a_{2n}x_2x_n + \cdots + a_{nn}x_n^2, \end{aligned} \tag{1.6-1}$$

其中 $x_i \in \mathbf{R}, i = 1,\cdots,n, a_{ij} \in \mathbf{R}, i,j = 1,\cdots,n$.

取 $a_{ji} = a_{ij}$，则 $2a_{ij}x_ix_j = a_{ij}x_ix_j + a_{ji}x_jx_i$，从而(1.6-1)式可写成

$$f(x_1,x_2,\cdots,x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j. \tag{1.6-2}$$

(1.6-2)式的系数确定了一个 n 阶实对称矩阵

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \text{其中 } a_{ij} = a_{ji},$$

称 $\mathbf{A}$ 为实二次型(1.6-2)的矩阵. 反之，任给一个 n 阶实对称矩阵 $\mathbf{A} = [a_{ij}]$，由(1.6-2)式也唯一确定了一个实二次型.

1.6 实二次型

记 $\boldsymbol{x}=[x_1,x_2,\cdots,x_n]^{\mathrm{T}}$，则二次型(1.6-2)可表示成

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}. \tag{1.6-2$'$}$$

例 1 将二次型

$$f(x_1,x_2,x_3)=3x_1^2+2x_2^2-x_3^2-2x_1x_3+4x_2x_3$$

表示成形如(1.6-2′)式.

解 令

$$\boldsymbol{A}=\begin{bmatrix}3 & 0 & -1\\ 0 & 2 & 2\\ -1 & 2 & -1\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix},$$

则

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}.$$ ∎

1.6 实二次型

数域 F 上之向量空间 V 的一个变换 $\mathscr{T}$ 称为**线性变换**,如果满足

(1) $\forall \boldsymbol{\alpha},\boldsymbol{\beta}\in V$,有 $\mathscr{T}(\boldsymbol{\alpha}+\boldsymbol{\beta})=\mathscr{T}\boldsymbol{\alpha}+\mathscr{T}\boldsymbol{\beta}$;

(2) $\forall \boldsymbol{\alpha}\in V$, $\forall k\in F$,有 $\mathscr{T}(k\boldsymbol{\alpha})=k\mathscr{T}\boldsymbol{\alpha}$.

设有 n 阶实方阵 $\boldsymbol{C}=[c_{ij}]$,则容易验证 $\mathbf{R}^n$ 上的变换 $\mathscr{T}$:

$$\mathscr{T}\boldsymbol{\alpha}=\boldsymbol{C}\boldsymbol{\alpha},\quad \boldsymbol{\alpha}\in\mathbf{R}^n$$

是线性变换.若记 $\boldsymbol{\beta}=\mathscr{T}\boldsymbol{\alpha}$,那么,当 $\boldsymbol{C}$ 可逆时,对任一 $\boldsymbol{\beta}\in\mathbf{R}^n$,有

$$\boldsymbol{\alpha}=\boldsymbol{C}^{-1}\boldsymbol{\beta}.$$

显然这是 $\boldsymbol{\beta}$ 到 $\boldsymbol{\alpha}$ 的线性变换,称其为线性变换 $\mathscr{T}$ 的**逆变换**,记为 $\mathscr{T}^{-1}$,即 $\boldsymbol{\alpha}=\mathscr{T}^{-1}\boldsymbol{\beta}$,并称 $\mathscr{T}$ 为**可逆线性变换**.

二次型主要的问题是,寻找适当的可逆线性变换

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}, \tag{1.6-3}$$

使二次型(1.6-2′)化成只含平方项的二次型,即用(1.6-3)式代入(1.6-2′)式能成为

$$f=d_1y_1^2+\cdots+d_ny_n^2=[y_1\quad\cdots\quad y_n]\begin{bmatrix}d_1&&\\&\ddots&\\&&d_n\end{bmatrix}\begin{bmatrix}y_1\\\vdots\\y_n\end{bmatrix}=\boldsymbol{y}^{\mathrm{T}}\begin{bmatrix}d_1&&\\&\ddots&\\&&d_n\end{bmatrix}\boldsymbol{y}. \tag{1.6-4}$$

这种只含平方项的二次型(1.6-4)叫做二次型 f 的**标准形(或法式)**.

1.6 实二次型

将(1.6-3)式代入(1.6-2′)式，得

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{C}\boldsymbol{y})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{C}\boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})\boldsymbol{y}.$$

由 $\boldsymbol{A}$ 是实对称矩阵知，$(\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C})^{\mathrm{T}}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$，即 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 也是实对称矩阵，故 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 是经过可逆线性变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 后，f 关于新变量 $y_1,\cdots,y_n$ 的矩阵. 因此，化二次型为标准形的问题也就是，对实对称矩阵 $\boldsymbol{A}$ 寻找可逆的实矩阵 $\boldsymbol{C}$，使 $\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 为对角矩阵.

定义 1.6-1 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 n 阶方阵，若存在可逆矩阵 $\boldsymbol{P}$，使

$$\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}, \tag{1.6-5}$$

则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **合同**，记为 $\boldsymbol{A}\simeq\boldsymbol{B}$. 这种对 $\boldsymbol{A}$ 所作的运算称为对 $\boldsymbol{A}$ 的**合同变换**，称可逆矩阵 $\boldsymbol{P}$ 为**合同变换矩阵**.

方阵之间的合同具有下列性质：

(1) 自反性 $\boldsymbol{A}\simeq\boldsymbol{A}$；

(2) 对称性 若 $\boldsymbol{A}\simeq\boldsymbol{B}$，则 $\boldsymbol{B}\simeq\boldsymbol{A}$；

(3) 传递性 若 $\boldsymbol{A}\simeq\boldsymbol{B}$，$\boldsymbol{B}\simeq\boldsymbol{C}$，则 $\boldsymbol{A}\simeq\boldsymbol{C}$.

特别地，当合同变换矩阵 $\boldsymbol{Q}$ 为正交矩阵时，由于 $\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{Q}^{-1}$，所以合同变换此时也是相似变换. 因此，有下述定理.

1.6 实二次型

定理 1.6-1 对任意 n 个变量的实二次型

$$f = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}, \quad \text{其中 } \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A},$$

存在正交线性变换(即合同变换矩阵是正交矩阵,一般也简称为正交变换)

$$\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$$

使二次型 f 化为标准形

$$f = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2, \tag{1.6-6}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值.

证 因为 $\boldsymbol{A}$ 是实对称矩阵,所以由定理 1.5-9 的推论知,存在正交矩阵 $\boldsymbol{Q}$ 使

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix},$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值. 于是,正交线性变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$ 就使二次型 f 化成标准形(1.6-6). ∎

1.6 实二次型

由代数基本定理知，特征方程(1.5-3)在复数域中必有解，且解的个数等于方程的次数(重根按重数计)，因而 n 阶方阵 $\boldsymbol{A}$ 有 n 个特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$(有重根时重复出现)，于是

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda). \tag{1.5-6}$$

定义 1.5-3 若 λ_i 是方阵 $\boldsymbol{A}$ 的特征方程(1.5-3)的 k 重根，则称特征值 λ_i 的代数重数是 k.

由于 $$(\boldsymbol{A}-\lambda\boldsymbol{I})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}-\lambda\boldsymbol{I},\ (\boldsymbol{A}-\bar{\lambda}\boldsymbol{I})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}-\lambda\boldsymbol{I},$$

故由行列式的性质得 $$\det(\boldsymbol{A}^{\mathrm{T}}-\lambda\boldsymbol{I})=\det(\boldsymbol{A}-\lambda\boldsymbol{I}),$$

$$\det(\boldsymbol{A}^{\mathrm{H}}-\lambda\boldsymbol{I})=\overline{\det(\boldsymbol{A}-\bar{\lambda}\boldsymbol{I})}.$$

这就是说，$\boldsymbol{A}^{\mathrm{T}}$ 与 $\boldsymbol{A}$ 有相同的特征多项式，故它们有相同的特征值；$\boldsymbol{A}^{\mathrm{H}}$ 的特征多项式与 $\boldsymbol{A}$ 的特征多项式之差别仅在于对应的各项系数互为共轭复数，故 $\boldsymbol{A}^{\mathrm{H}}$ 的特征值是 $\boldsymbol{A}$ 的特征值的共轭复数.

1.6 实二次型

例 2 求一个正交线性变换，把二次型

$$f = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4$$

化为标准形.

解 f 的矩阵是
$$A = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{bmatrix}.$$

由

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 1-\lambda & 1-\lambda & 1-\lambda \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix}$$

$$= (1-\lambda)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & -\lambda-1 & -2 & 0 \\ 0 & -2 & -\lambda-1 & 0 \\ 0 & 2 & 2 & -\lambda+1 \end{vmatrix} = (1-\lambda)^2\begin{vmatrix} -\lambda-1 & -2 \\ -2 & -\lambda-1 \end{vmatrix}$$

$$= (\lambda+3)(\lambda-1)^3,$$

知 A 的特征值是 $\lambda_1 = -3, \lambda_2 = \lambda_3 = \lambda_4 = 1$.

1.6 实二次型

对 $\lambda_1=-3$，解线性方程组 $(\boldsymbol{A}-(-3)\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix} 3 & 1 & 1 & -1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{bmatrix}\boldsymbol{x}=\boldsymbol{0},$$

得其一个基础解系 $\boldsymbol{x}_1=[1,-1,-1,1]^{\mathrm{T}}$，单位化后得

$$\boldsymbol{\varepsilon}_1=\left[\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right]^{\mathrm{T}}.$$

对 $\lambda_2=\lambda_3=\lambda_4=1$，解线性方程组 $(\boldsymbol{A}-1\cdot\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}\boldsymbol{x}=\boldsymbol{0},$$

1.6 实二次型

亦即 $$(-1)\cdot x_1+1\cdot x_2+1\cdot x_3+(-1)\cdot x_4=0.$$

不难得到它的三个相互正交的解(向量)：

$$\boldsymbol{x}_2=[1,1,0,0]^{\mathrm{T}},\ \boldsymbol{x}_3=[0,0,1,1]^{\mathrm{T}},\ \boldsymbol{x}_4=[1,-1,1,-1]^{\mathrm{T}}.$$

将它们单位化后分别得

$$\boldsymbol{\varepsilon}_2=\left[\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0\right]^{\mathrm{T}},\quad \boldsymbol{\varepsilon}_3=\left[0,0,\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right]^{\mathrm{T}},$$

$$\boldsymbol{\varepsilon}_4=\left[\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right]^{\mathrm{T}}.$$

于是，令 $$\boldsymbol{Q}=[\boldsymbol{\varepsilon}_1\quad \boldsymbol{\varepsilon}_2\quad \boldsymbol{\varepsilon}_3\quad \boldsymbol{\varepsilon}_4]=\begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} \end{bmatrix},$$

则 $\boldsymbol{Q}$ 为正交矩阵，且经正交线性变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$，二次型 f 化为

$$f=-3y_1^2+y_2^2+y_3^2+y_4^2\quad(\text{为什么？}).$$

1.6 实二次型

定义 1.6-2 设有实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，如果对任意的实向量$\boldsymbol{x}\neq\boldsymbol{0}$，都有 $f>0(f<0)$，则称 f 为**正定(负定)二次型**，相应的实对称矩阵 $\boldsymbol{A}$ 称为**正定(负定)矩阵**，记为 $\boldsymbol{A}>\boldsymbol{0}(\boldsymbol{A}<\boldsymbol{0})$；如果对任意的实向量 $\boldsymbol{x}\neq\boldsymbol{0}$，都有 $f\geqslant0(f\leqslant0)$，则称 f 为**半正定(半负定)二次型**，相应的实对称矩阵 $\boldsymbol{A}$ 称为**半正定(半负定)矩阵**，记为 $\boldsymbol{A}\geqslant\boldsymbol{0}$ $(\boldsymbol{A}\leqslant\boldsymbol{0})$.

定理 1.6-2 实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定(负定)二次型的充要条件是，f 的矩阵 $\boldsymbol{A}$ 的特征值全都大于(小于)零.

证 设 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值，由定理 1.6-1 知，存在正交线性变换 $\boldsymbol{x}=\boldsymbol{Q}\boldsymbol{y}$ 使

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q})\boldsymbol{y}=\lambda_1y_1^2+\cdots+\lambda_ny_n^2.$$

若 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 都大于 0，则只要 $\boldsymbol{y}\neq\boldsymbol{0}$ 就有 $f>0$，从而只要 $\boldsymbol{x}\neq\boldsymbol{0}$就有 $f>0$，即 f 是正定二次型.

反之，若有一个特征值不大于零，不妨设为 $\lambda_1\leqslant0$. 则取 $\boldsymbol{y}_0=[1,0,\cdots,0]^{\mathrm{T}}$ 就使 $f\leqslant0$，从而存在 $\boldsymbol{x}_0=\boldsymbol{Q}\boldsymbol{y}_0\neq\boldsymbol{0}$使 $f\leqslant0$. 这与 f 是正定的矛盾. ∎

1.6 实二次型

推论 1 实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定(负定)二次型的充要条件是，存在可逆的实矩阵 $\boldsymbol{P}$ 使

$$\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P} \quad (\boldsymbol{A}=-\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}).$$

证 **充分性**：对任意非零实向量 $\boldsymbol{x}$，由于 $\boldsymbol{P}$ 可逆，所以 $\boldsymbol{P}\boldsymbol{x}\neq\boldsymbol{0}$. 因而

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}=(\boldsymbol{P}\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{P}\boldsymbol{x})=|\boldsymbol{P}\boldsymbol{x}|^{2}>0.$$

必要性：设 f 是正定二次型，则存在正交矩阵 $\boldsymbol{Q}$ 使

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix},$$

其中 $\lambda_i>0(i=1,\cdots,n)$ 为 $\boldsymbol{A}$ 的 n 个特征值. 于是

$$\boldsymbol{A}=\boldsymbol{Q}\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}\boldsymbol{Q}^{\mathrm{T}}.$$

令

$$\boldsymbol{P}=\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}\boldsymbol{Q}^{\mathrm{T}},$$

显然 $\boldsymbol{P}$ 是可逆的，且有 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$. ▌

1.6 实二次型

推论 2 若 $\boldsymbol{A}$ 是正定矩阵，则 $\boldsymbol{A}$ 的行列式大于零.

证 由推论 1 知，存在可逆的实矩阵 $\boldsymbol{P}$ 使

$$\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P},$$

因而由 $|\boldsymbol{P}|\neq 0$ 推出

$$|\boldsymbol{A}|=|\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}|=|\boldsymbol{P}^{\mathrm{T}}||\boldsymbol{P}|=|\boldsymbol{P}|^2>0. \quad \blacksquare$$

定理 1.6-3 实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定(负定)二次型的充要条件是，$\boldsymbol{A}$ 的各阶顺序主子式都大于零($\boldsymbol{A}$ 的奇数阶顺序主子式小于零，而偶数阶顺序主子式大于零).

证 **必要性**：设 f 正定，对 $\boldsymbol{A}$ 作如下分块

$$\boldsymbol{A}=\left[\begin{array}{ccc:ccc} a_{11} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \cdots & & \cdots & \cdots & & \cdots \\ a_{k1} & \cdots & a_{kk} & a_{k,k+1} & \cdots & a_{kn} \\ \hdashline a_{k+1,1} & \cdots & a_{k+1,k} & a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \cdots & & \cdots & \cdots & & \cdots \\ a_{n1} & \cdots & a_{nk} & a_{n,k+1} & \cdots & a_{nn} \end{array}\right]=\left[\begin{array}{c:c} \boldsymbol{A}_k & \boldsymbol{B} \\ \hdashline \boldsymbol{C} & \boldsymbol{D} \end{array}\right],$$

取非零向量 $\boldsymbol{x}=\left[\begin{array}{c} \boldsymbol{\alpha} \\ \hdashline \boldsymbol{0} \end{array}\right]\begin{array}{l} \}k \\ \}n-k \end{array}$，其中 $\boldsymbol{\alpha}=[x_1,\cdots,x_k]^{\mathrm{T}}\neq\boldsymbol{0}$，则有

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=[\boldsymbol{\alpha}^{\mathrm{T}} \; \vdots \; \boldsymbol{0}^{\mathrm{T}}]\left[\begin{array}{c:c} \boldsymbol{A}_k & \boldsymbol{B} \\ \hdashline \boldsymbol{C} & \boldsymbol{D} \end{array}\right]\left[\begin{array}{c} \boldsymbol{\alpha} \\ \hdashline \boldsymbol{0} \end{array}\right]=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{\alpha}.$$

1.6 实二次型

显然，$\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{\alpha}$ 是 k 个变量 $x_1,\cdots,x_k$ 的二次型，由于 f 正定，所以 $\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}_k\boldsymbol{\alpha}$ 也是正定二次型，故 $\boldsymbol{A}_k$ 是正定矩阵. 因此由推论 2 知

$$|\boldsymbol{A}_k| = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} > 0.$$

因为 k 可取不大于 n 的任何正整数，所以 $\boldsymbol{A}$ 的各阶顺序主子式都大于零.

充分性：对变量的个数 n 用归纳法.

当 $n=1$ 时，$f(x_1)=a_{11}x_1^2$. 由 $a_{11}>0$，故 $f(x_1)$ 是正定的.

现假设对 $n-1$ 个变量的二次型结论成立，要证对 n 个变量的二次型结论也成立.

对二次型的 n 阶实对称矩阵 $\boldsymbol{A}$ 作如下分块

$$\boldsymbol{A} = \left[\begin{array}{c:c} \boldsymbol{A}_{n-1} & \boldsymbol{\beta} \\ \hdashline \boldsymbol{\beta}^{\mathrm{T}} & a_{nn} \end{array}\right],$$

其中 $\boldsymbol{\beta}=[a_{1n},\cdots,a_{n-1,n}]^{\mathrm{T}}=[a_{n1},\cdots,a_{n,n-1}]^{\mathrm{T}}$.

由归纳假设知 $\boldsymbol{A}_{n-1}$ 是正定矩阵，故 $\boldsymbol{A}_{n-1}$ 可逆，所以有

$$\left[\begin{array}{c:c} \boldsymbol{I}_{n-1} & \\ \hdashline -\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1} & 1 \end{array}\right]\left[\begin{array}{c:c} \boldsymbol{A}_{n-1} & \boldsymbol{\beta} \\ \hdashline \boldsymbol{\beta}^{\mathrm{T}} & a_{nn} \end{array}\right]\left[\begin{array}{c:c} \boldsymbol{I}_{n-1} & -\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta} \\ \hdashline & 1 \end{array}\right] = \left[\begin{array}{c:c} \boldsymbol{A}_{n-1} & \\ \hdashline & a_{nn}-\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta} \end{array}\right]. \quad (1.6\text{-}7)$$

1.6 实二次型

于是
$$|\boldsymbol{A}|=(a_{nn}-\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta})|\boldsymbol{A}_{n-1}|.$$

由于$|\boldsymbol{A}|>0$ 和$|\boldsymbol{A}_{n-1}|>0$,故 $a_{nn}-\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}>0$.

令 $b_{nn}=a_{nn}-\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\beta}$, $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}_{n-1}\\ \hdashline x_n\end{bmatrix}$, $\boldsymbol{B}=\left[\begin{array}{c:c}\boldsymbol{A}_{n-1} & \\ \hdashline & b_{nn}\end{array}\right]$,

其中 $\boldsymbol{x}_{n-1}=[x_1,\cdots,x_{n-1}]^{\mathrm{T}}$,则有
$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}=\left[\begin{array}{c:c}\boldsymbol{x}_{n-1}^{\mathrm{T}} & x_n\end{array}\right]\left[\begin{array}{c:c}\boldsymbol{A}_{n-1} & \\ \hdashline & b_{nn}\end{array}\right]\begin{bmatrix}\boldsymbol{x}_{n-1}\\ \hdashline \boldsymbol{x}_n\end{bmatrix}=\boldsymbol{x}_{n-1}^{\mathrm{T}}\boldsymbol{A}_{n-1}\boldsymbol{x}_{n-1}+b_{nn}x_n^2.$$

由归纳假设知,$n-1$ 个变量的实二次型 $\boldsymbol{x}_{n-1}^{\mathrm{T}}\boldsymbol{A}_{n-1}\boldsymbol{x}_{n-1}$是正定的,故 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{x}$ 也是正定的.而由(1.6-7)式知 $\boldsymbol{A}\simeq\boldsymbol{B}$,故 $\boldsymbol{A}$ 是正定矩阵(为什么?)亦即 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是正定二次型.

对于负定二次型的情况,只需注意 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 负定的充要条件是 $-f=-\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}(-\boldsymbol{A})\boldsymbol{x}$为正定. ■

类似地可证下述定理及其推论.

1.6 实二次型

定理 1.6-4 实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为半正定(半负定)二次型的充要条件是，f 的矩阵 $\boldsymbol{A}$ 的特征值全都非负(非正)。

推论 1 实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为半正定(半负定)二次型的充要条件是，存在实矩阵 $\boldsymbol{P}$ 使

$$\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\quad(\boldsymbol{A}=-\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}).$$

推论 2 若 $\boldsymbol{A}$ 是半正定矩阵，则 $\boldsymbol{A}$ 的行列式非负.

例 3 设 $\boldsymbol{A}$ 是正定矩阵，证明 $\boldsymbol{A}^{-1}$，$\mathrm{adj}\boldsymbol{A}$ 是正定矩阵.

证 由 $\boldsymbol{A}$ 是正定矩阵知，$\boldsymbol{A}$ 的特征值 $\lambda_1,\cdots,\lambda_n$ 全都大于零，故由 1.5 节例 2 知，$\lambda_1^{-1},\cdots,\lambda_n^{-1}$ 是 $\boldsymbol{A}^{-1}$ 的 n 个特征值，而它们是全大于零的. 又 $(\boldsymbol{A}^{-1})^{\mathrm{T}}=(\boldsymbol{A}^{\mathrm{T}})^{-1}=\boldsymbol{A}^{-1}$，故 $\boldsymbol{A}^{-1}$ 是正定矩阵.

容易验证，大于零的数乘正定矩阵仍是正定矩阵，故由 $\mathrm{adj}\boldsymbol{A}=\dfrac{\boldsymbol{A}^{-1}}{|\boldsymbol{A}|}$，$|\boldsymbol{A}|>0$ 和

$$(\mathrm{adj}\boldsymbol{A})^{\mathrm{T}}=\left(\frac{\boldsymbol{A}^{-1}}{|\boldsymbol{A}|}\right)^{\mathrm{T}}=\frac{1}{|\boldsymbol{A}|}(\boldsymbol{A}^{-1})^{\mathrm{T}}=\frac{\boldsymbol{A}^{-1}}{|\boldsymbol{A}|}=\mathrm{adj}\boldsymbol{A}$$

知，$\mathrm{adj}\boldsymbol{A}$ 也是正定矩阵. ∎

1.6 实二次型

例 4 设 $\boldsymbol{A}$、$\boldsymbol{B}$ 都是 n 阶正定矩阵，证明：$\boldsymbol{AB}$ 的所有特征值全都为正，但 $\boldsymbol{AB}$ 未必是正定矩阵；若还有 $\boldsymbol{AB}=\boldsymbol{BA}$，则 $\boldsymbol{AB}$ 是正定矩阵.

证 因 $\boldsymbol{A}$ 正定，由定理 1.6-2 的推论 1 知 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$，这里 $\boldsymbol{P}$ 是可逆的. 又因 $\boldsymbol{B}$ 正定，故 $\boldsymbol{PBP}^{\mathrm{T}}$ 是正定矩阵，从而它的特征值全都为正.

由于
$$\boldsymbol{AB}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{PB}=\boldsymbol{P}^{\mathrm{T}}(\boldsymbol{PBP}^{\mathrm{T}})(\boldsymbol{P}^{\mathrm{T}})^{-1},$$
所以 $\boldsymbol{AB}$ 与 $\boldsymbol{PBP}^{\mathrm{T}}$ 相似，故它们有相同的特征值，因此 $\boldsymbol{AB}$ 的特征值全都为正.

因 $(\boldsymbol{AB})^{\mathrm{T}}=\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{BA}\neq\boldsymbol{AB}$，故 $\boldsymbol{AB}$ 未必正定，仅当 $\boldsymbol{BA}=\boldsymbol{AB}$ 时，$\boldsymbol{AB}$ 才是正定矩阵. ∎

1.6 实二次型

例 5 设有二次型

$$f = x_1^2 + x_2^2 + 5x_3^2 + 2tx_1x_2 - 2x_1x_3 + 4x_2x_3,$$

问 t 取何值时，它是正定二次型？

解 f 的矩阵为

$$\boldsymbol{A} = \begin{bmatrix} 1 & t & -1 \\ t & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix},$$

f 为正定二次型的充要条件是

$$1 > 0, \quad \begin{vmatrix} 1 & t \\ t & 1 \end{vmatrix} = 1 - t^2 > 0, \quad \begin{vmatrix} 1 & t & -1 \\ t & 1 & 2 \\ -1 & 2 & 5 \end{vmatrix} = -5t^2 - 4t > 0,$$

即

$$\begin{cases} t^2 - 1 < 0, \\ t(5t + 4) < 0. \end{cases}$$

解之得 $-\dfrac{4}{5} < t < 0$，故当 $-\dfrac{4}{5} < t < 0$ 时，f 正定. ∎

作业（第1部分）

习题 1.6

1. 用正交线性变换将二次型

$$f = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_{12} - 4x_2x_3$$

化为标准形，并给出所用的正交线性变换.

2. 问 t 取何值时，对称矩阵 $\mathbf{A}$ 是正定的，其中

$$\mathbf{A} = \begin{bmatrix} 1 & t & 1 \\ t & 2 & 0 \\ 1 & 0 & 1-t \end{bmatrix}.$$

区间估计

数理统计（6）

厦門大學 XIAMEN UNIVERSITY 信息学院（国家示范性软件学院）School of Informatics 黄 炜 博士·副教授 Dr. Wei Huang

2.2.3 相和估计与渐进正态性

由于未知参数 θ 的估计量 $\hat{\theta}(X_1, X_2, \cdots, X_n)$ 是与样本容量 n 有关的，所以自然希望当 n 无限增大时，$\hat{\theta}(X_1, X_2, \cdots, X_n)$ 应该趋向于 θ.

定义 2.2-8 设 $\hat{\theta}(X_1, X_2, \cdots, X_n)$ 是未知参数 θ 的估计量，如果对任意 $\varepsilon>0$，都有

$$\lim_{n\to+\infty} P\{|\hat{\theta}(X_1, X_2, \cdots, X_n) - \theta| \geqslant \varepsilon\} = 0, \tag{2.2-19}$$

则称 $\hat{\theta}(X_1, X_2, \cdots, X_n)$ 为 θ 的**相合估计量**.

验证一个估计量是否为相合估计量，常常要用到所谓 Чебыщев 型不等式. 现用一个引理给出其证明.

引理 设 $g(X)$ 是随机变量 X 的非负连续函数，如果 $E[g(X)]$ 存在，则对于任一正常数 c，均有

$$P\{g(X) \geqslant c\} \leqslant \frac{E[g(X)]}{c}. \tag{2.2-20}$$

证 由于 g 是非负连续函数，所以

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x \geqslant \int_{\{x|g(x)\geqslant c\}} g(x)f(x)\mathrm{d}x$$

$$\geqslant c \int_{\{x|g(x)\geqslant c\}} f(x)\mathrm{d}x = cP\{g(X) \geqslant c\},$$

从而由 $c>0$ 知，(2.2-10)式成立. ∎

2.2.3 相和估计与渐进正态性

由于样本方差 S^2 是总体方差 σ^2 的无偏估计，所以有

$$P\{|S^2-\sigma^2|\geqslant\varepsilon\}\leqslant\frac{D(S^2)}{\varepsilon^2}.$$

而由 $\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$ 及(1.3-7)式知

$$D(S^2)=\frac{2\sigma^4}{n-1},$$

故 $\lim\limits_{n\to+\infty}P\{|S^2-\sigma^2|\geqslant\varepsilon\}=0$，从而 S^2 是 σ^2 的相合估计量.

请读者自证 $\lim\limits_{n\to\infty}P\{|\tilde{S}^2-\sigma^2|>\varepsilon\}=0$.

2.3 区间估计

以上讨论的是参数的点估计问题，由样本构造一个估计量去估计未知参数，用样本观测值所得的相应估计量的观测值作为参数的估计值. 对于未知参数 θ，除了求出它的点估计 $\hat{\theta}$ 外，我们还希望估计出一个范围，并希望知道这个范围包含参数 θ 真值的可靠程度. 这样的范围通常用区间的形式给出，区间的两个端点是样本的函数，记为 $\underline{\theta}(X_1, X_2, \cdots, X_n)$（或 $\underline{\theta}$）与 $\bar{\theta}(X_1, X_2, \cdots, X_n)$（或 $\bar{\theta}$）. 由于样本 $(X_1, X_2, \cdots, X_n)$ 是 n 维随机变量，所以区间 $[\underline{\theta}, \bar{\theta}]$ 是一个随机区间，它是否包含了未知参数 θ 是一个随机事件，这个随机事件的概率的大小反映了可靠程度，这种带有一定概率的区间 $[\underline{\theta}, \bar{\theta}]$ 称为**置信区间**，以置信区间的形式对未知参数 θ 进行估计，称为**区间估计**.

定义 2.3-1 设总体分布的未知参数为 θ. 对于给定的 α，$0<\alpha<1$，如果存在两个统计量 $\underline{\theta}(X_1, X_2, \cdots, X_n)$ 与 $\bar{\theta}(X_1, X_2, \cdots, X_n)$，使得

$$P\{\underline{\theta}(X_1, X_2, \cdots, X_n) \leqslant \theta \leqslant \bar{\theta}(X_1, X_2, \cdots, X_n)\} = 1-\alpha, \tag{2.3-1}$$

则称 $[\underline{\theta}, \bar{\theta}]$ 为 θ 的**双侧 $1-\alpha$ 置信区间**；$1-\alpha$ **称为置信水平**；$\underline{\theta}$（$\bar{\theta}$）称为 θ 的**双侧 $1-\alpha$ 置信区间的下(上)限**，简称为**双侧置信下(上)限**.

2.3 区间估计

(2.3-1)式的意义是什么呢？若每次抽取一个大小为 n 的样本，则相应的样本观测值确定一个具体区间 $[\underline{\theta},\overline{\theta}]$，这个区间要么包含 θ 的真值，要么不包含 θ 的真值．当反复抽取很多次样本时，根据 Bernoulli 大数定律，在这很多个具体区间中大约有 $100(1-\alpha)\%$ 的区间包含 θ 真值，不包含 θ 真值的仅有 $100\alpha\%$ 左右．例如，若 $\alpha=0.05$，反复抽取 1000 次，则所得到的 1000 个区间中不包含 θ 真值的仅 50 个左右，而其余的 950 个都包含 θ 真值．这就是说，如果认为"区间 $[\underline{\theta},\overline{\theta}]$ 包含未知参数 θ 真值"，那么犯错误的概率约为 α．显然，α 越小，$1-\alpha$ 越大，从而随机区间 $[\underline{\theta},\overline{\theta}]$ 包含 θ 真值的可靠度越高．那么是否 α 取得越小越好呢？不是的！因为 α 越小，区间 $[\underline{\theta},\overline{\theta}]$ 的长度就会越大，如果大得过分了，区间估计就失去了它的意义．反过来，是否区间越小越好呢？也不是．区间太小，置信水平 $1-\alpha$ 就会变小，从而 $[\underline{\theta},\overline{\theta}]$ 不包含 θ 真值的概率变大；如果估计的区间经常不包含 θ 真值，这也失去了区间估计的意义．因此，选取置信区间应该遵循这样的原则：在给定的较大的置信水平 $1-\alpha$ 下，尽量使 $[\underline{\theta},\overline{\theta}]$ 的平均长度最短．

2.3 区间估计

例 1 已知某炼铁厂的铁水含碳量(%)在正常情况下服从正态分布，且标准差 $\sigma=0.12$. 现测量四炉铁水，其含碳量分别是

$$4.28,\quad 4.40,\quad 4.42,\quad 4.36$$

求该厂铁水平均含碳量 μ 的双侧 95% 置信区间.

解 由于样本均值 $\overline{X}$ 是 μ 的一个较优点估计，且 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}=\dfrac{2}{0.12}(\overline{X}-\mu)$ 服从 $N(0,1)$，不依赖于 μ；又考虑到

$$P\{\underline{\mu}\leqslant\mu\leqslant\bar{\mu}\}=1-\alpha=0.95$$

等价于

$$P\{-\underline{\mu}\geqslant-\mu\geqslant-\bar{\mu}\}=P\{\overline{X}-\underline{\mu}\geqslant\overline{X}-\mu\geqslant\overline{X}-\bar{\mu}\}$$

$$=P\left\{\frac{2}{0.12}(\overline{X}-\underline{\mu})\geqslant\frac{2}{0.12}(\overline{X}-\mu)\geqslant\frac{2}{0.12}(\overline{X}-\bar{\mu})\right\}$$

$$=0.95,$$

所以借助于标准正态分布的分位数，有

$$P\left\{u_{0.025}\leqslant\frac{2}{0.12}(\overline{X}-\mu)\leqslant u_{0.975}\right\}=0.95,$$

即

$$P\left\{-1.96\leqslant\frac{2}{0.12}(\overline{X}-\mu)\leqslant1.96\right\}=0.95.$$

2.3 区间估计

因此,若令 $\frac{2}{0.12}(\overline{X}-\overline{\mu})=-1.96$, $\frac{2}{0.12}(\overline{X}-\underline{\mu})=1.96$,则分别得到 $\underline{\mu}=\overline{X}-0.118$, $\overline{\mu}=\overline{X}+0.118$.

由样本观测值得 $\overline{x}=\frac{1}{4}(4.28+4.40+4.42+4.36)=4.365$,故 μ 的双侧 95%的置信区间为

$$[4.365-0.118, 4.365+0.118]=[4.247, 4.483].$$ ■

值得注意的是,置信区间不是唯一的.例如,由于

$$P\left\{u_{0.04}\leqslant\frac{2}{0.12}(\overline{X}-\mu)\leqslant u_{0.99}\right\}=0.95,$$

即

$$P\left\{-1.75\leqslant\frac{2}{0.12}(\overline{X}-\mu)\leqslant 2.33\right\}=0.95,$$

所以 μ 的双侧 95%的置信区间也可以取为

$$[4.365-0.140, 4.365+0.105]=[4.225, 4.470].$$

不过,这个区间的长度为 0.245,比上述区间的长度 0.236 要长,因而不采用这个区间.

2.3 区间估计

根据上述应使未知参数的双侧 $1-\alpha$ 置信区间的长度尽可能小的原则，一般取双侧 α 分位数来确定置信区间.

依照例 1 的求解过程，求双侧 $1-\alpha$ 置信区间的步骤可归纳如下。

（1）求出未知参数 θ 的一个较优的点估计 $\hat{\theta}$，一般用 θ 的极大似然估计.

（2）以 $\hat{\theta}$ 为基础寻找一个随机变量

$$J=J(X_1,X_2,\cdots,X_n;\theta),$$

J 的分布要已知，从而可以通过查表或直接计算得到 J 的分位数.

（3）对于给定的置信水平 $1-\alpha$，确定常数 a 和 b 使

$$P\{a\leqslant J\leqslant b\}=1-\alpha. \tag{2.3-2}$$

一般取 a 为 J 的 $\frac{\alpha}{2}$ 分位数，取 b 为 J 的 $1-\frac{\alpha}{2}$ 分位数.

（4）将不等式“$a\leqslant J\leqslant b$”进行等价变形，使它具有“$\underline{\theta}(X_1,X_2,\cdots,X_n)\leqslant\theta\leqslant\overline{\theta}(X_1,X_2,\cdots,X_n)$”的形式，从而区间$[\underline{\theta},\overline{\theta}]$即为所求的置信区间.

在有些问题中，所关心的仅是未知参数 θ 至少有多大（例如灯泡的寿命问题），或者不超过多大（例如某批产品的不合格率问题），这时需要进行单侧区间估计.

2.3 区间估计

定义 2.3-2 设总体分布的未知参数为 θ，对于给定的 α，$0<\alpha<1$，如果存在统计量 $\underline{\theta}(X_1,X_2,\cdots,X_n)$，使得

$$P(\underline{\theta}(X_1,X_2,\cdots,X_n)\leqslant\theta)=1-\alpha, \qquad (2.3\text{-}3)$$

则称 $[\underline{\theta},+\infty]$ 为 θ 的**单侧 $1-\alpha$ 置信区间**，$\underline{\theta}$ 称为 θ 的**单侧 $1-\alpha$ 置信区间的下限**。类似地，如果存在统计量 $\overline{\theta}(X_1,X_2,\cdots,X_n)$，使得

$$P(\theta\leqslant\overline{\theta}(X_1,X_2,\cdots,X_n))=1-\alpha, \qquad (2.3\text{-}4)$$

则称 $(-\infty,\overline{\theta}]$ 为 θ 的单侧 $1-\alpha$ 置信区间，$\overline{\theta}$ 称为 θ 的**单侧 $1-\alpha$ 置信区间的上限**。

上述求置信区间的一般步骤对于单侧置信区间依然有效，只不过步骤 3 要稍作修改。将(2.3-2)式改为求常数 a，使

$$P(J\geqslant a)=1-\alpha,$$

或者改为求常数 b，使
$$P(J\leqslant b)=1-\alpha.$$

这时，一般取 a 为 J 的 α 分位数，取 b 为 J 的 $1-\alpha$ 分位数。

2.3 区间估计

下面就一个正态总体，二个正态总体及非正态总体的情况分别讨论区间估计问题.

1. 一个正态总体的情况

设$(X_1,X_2,\cdots,X_n)$是取自正态总体$N(\mu;\sigma^2)$的样本，具体讨论如何求得μ、σ^2的置信区间，置信水平为$1-\alpha$.

(1) μ未知但σ^2已知，求μ的置信区间.

这时，μ的极大似然估计量为$\overline{X}$，且取

$$J=\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma},$$

则$J\sim N(0;1)$. 由于σ已知，故$J=J(X_1,\cdots,X_n;\mu)$，所以(2.3-2)式为

$$P\{u_{\alpha/2}\leqslant J\leqslant u_{1-\alpha/2}\}=1-\alpha. \tag{2.3-5}$$

注意到$u_{\alpha/2}=-u_{1-\alpha/2}$，从而(2.3-5)式可改写为

$$P\left\{-u_{1-\alpha/2}\leqslant\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\leqslant u_{1-\alpha/2}\right\}=1-\alpha,$$

亦即

$$P\left\{\overline{X}-\frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}\leqslant\mu\leqslant\overline{X}+\frac{\sigma}{\sqrt{n}}u_{1-\alpha/2}\right\}=1-\alpha. \tag{2.3-6}$$

2.3 区间估计

因此，μ 的双侧 $1-\alpha$ 置信区间是

$$\left[\overline{X}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\overline{X}+u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

如果要求单侧置信下限，则把(2.3-5)式改为

$$P\{J\leqslant u_{1-\alpha}\}=1-\alpha,$$

并将它等价地表示为

$$P\left\{\overline{X}-u_{1-\alpha}\frac{\sigma}{\sqrt{n}}\leqslant\mu\right\}=1-\alpha,$$

便得到 μ 的单侧 $1-\alpha$ 置信区间的下限 $\overline{X}-u_{1-\alpha}\frac{\sigma}{\sqrt{n}}$.

类似地，只要把(2.3-5)式改为

$$P\{J\geqslant -u_{1-\alpha}\}=1-\alpha,$$

则可得到 μ 的单侧 $1-\alpha$ 置信区间的上限 $\overline{X}+u_{1-\alpha}\frac{\sigma}{\sqrt{n}}$.

2.3 区间估计

(2) μ 与 σ^2 均未知，求 μ 的置信区间.

μ 的极大似然估计量仍为 $\overline{X}$，但上述的 J 中含有未知参数 σ，所以不符合要求. 这时一个自然的想法是用样本标准差 S 来替代 σ，并注意到(1.4-3)式，则取

$$J=\sqrt{n}\frac{\overline{X}-\mu}{S},$$

从而 $J\sim t(n-1)$，且有 $J=J(X_1,X_2,\cdots,X_n;\mu)$. 于是，(2.3-2)式为

$$P\{t_{\alpha/2}(n-1)\leqslant J\leqslant t_{1-\alpha/2}(n-1)\}=1-\alpha.$$

因为 $t_{\alpha/2}(n-1)=-t_{1-\alpha/2}(n-1)$，所以上式又可等价地表示成

$$P\left\{\overline{X}-t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\leqslant\mu\leqslant\overline{X}+t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right\}=1-\alpha. \qquad (2.3\text{-}7)$$

这样便得到 μ 的双侧 $1-\alpha$ 置信区间为

$$\left[\overline{X}-t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}},\overline{X}+t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right],$$

与(1)类似地可得到 μ 的单侧 $1-\alpha$ 置信区间的下限或上限，这里不再详述，其结果见表 2.3-1.

2.3 区间估计

(3) μ 已知但 σ^2 未知，求 σ^2 的置信区间.

这时，σ^2 的极大似然估计量为 $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$，取

$$J = \frac{n\widehat{\sigma}^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2,$$

则 $J \sim \chi^2(n)$，且 $J = J(X_1, X_2, \cdots, X_n; \sigma^2)$. 因此，(2.3-2)式为

$$P\{\chi^2_{\alpha/2}(n) \leqslant J \leqslant \chi^2_{1-\alpha/2}(n)\} = 1-\alpha.$$

将上式等价地表示为

$$P\left\{\frac{n\widehat{\sigma}^2}{\chi^2_{1-\alpha/2}(n)} \leqslant \sigma^2 \leqslant \frac{n\widehat{\sigma}^2}{\chi^2_{\alpha/2}(n)}\right\} = 1-\alpha, \tag{2.3-8}$$

从而得到 σ^2 的双侧 $1-\alpha$ 置信区间为

$$\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}, \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}\right].$$

2.3 区间估计

(4) μ 与 σ^2 均未知，求 σ^2 的置信区间

这时，σ^2 的极大似然估计量是 $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$，即 $\hat{\sigma}^2=\tilde{S}^2$，而由(1.4-2)式知 $\frac{(n-1)S^2}{\sigma^2}=\sum_{i=1}^{n}(X_i-\overline{X})^2/\sigma^2\sim\chi^2(n-1)$，故取 J 为

$$J=\frac{(n-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2,$$

则 J 的分布是已知的，且 $J=J(X_1,X_2,\cdots,X_n;\sigma^2)$. 从而可得到 σ^2 的双侧 $1-\alpha$ 置信区间为

$$\left[\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{1-\alpha/2}(n-1)},\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\chi^2_{\alpha/2}(n-1)}\right].$$

上式也可写为

$$\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)},\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right].$$

至于 σ^2 的单侧 $1-\alpha$ 置信区间的下限与上限见表 2.3-1.

2.3 区间估计

表 2.3-1 一个正态总体下未知参数的 $1-\alpha$ 置信区间

未知参数		随机变量 J	J 的分布	双侧置信区间下，上限	单侧置信下限	单侧置信上限
μ	σ^2 已知	$\sqrt{n}\dfrac{\bar{X}-\mu}{\sigma}$	$N(0,1)$	$\bar{X} \mp u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}$	$\bar{X} - u_{1-\alpha}\dfrac{\sigma}{\sqrt{n}}$	$\bar{X} + u_{1-\alpha}\dfrac{\sigma}{\sqrt{n}}$
	σ^2 未知	$\sqrt{n}\dfrac{\bar{X}-\mu}{S}$	$t(n-1)$	$\bar{X} \mp t_{1-\alpha/2}(n-1)\dfrac{S}{\sqrt{n}}$	$\bar{X} - t_{1-\alpha}(n-1)\dfrac{S}{\sqrt{n}}$	$\bar{X} + t_{1-\alpha}(n-1)\dfrac{S}{\sqrt{n}}$
σ^2	μ 已知	$\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}$	$\chi^2(n)$	$\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}, \dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}$	$\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha}(n)}$	$\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha}(n)}$
	μ 未知	$\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2}$	$\chi^2(n-1)$	$\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\chi^2_{1-\alpha/2}(n-1)}, \dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\chi^2_{\alpha/2}(n-1)}$	$\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\chi^2_{1-\alpha}(n-1)}$	$\dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\chi^2_{\alpha}(n-1)}$

2.3 区间估计

例 2 为确定某种溶液中的甲醛浓度，随机地抽取容量为 4 的样本，得到样本均值的观测值 $\overline{x}=8.34(\%)$，样本标准差的观测值 $s=0.26(\%)$，假设被测总体近似地服从正态总体 $N(\mu;\sigma^2)$，求 μ 的双侧 95% 置信区间.

解 由于 σ^2 未知而要求 μ 的双侧置信区间，所以按表 2.3-1 知，μ 的双侧 95% 置信区间为

$$\left[\overline{x}-t_{0.975}(3)\frac{s}{\sqrt{4}},\overline{x}+t_{0.975}(3)\frac{s}{\sqrt{4}}\right].$$

将 $\overline{x}=8.34$，$s=0.26$ 及 $t_{0.975}(3)=3.1824$ 代入上式，得 μ 的双侧 95% 置信区间是 $[7.926\%,8.754\%]$. ∎

2.3 区间估计

例 3　从自动机床加工的同类产品中随机地抽取 16 件，测得它们的长度（单位：厘米）为

12.15，12.12，12.01，12.28，12.09，12.16，12.03，12.01，
12.06，12.13，12.07，12.11，12.08，12.01，12.03，12.06

假设产品的长度服从正态分布 $N(\mu;\sigma^2)$，求 μ 的双侧 90% 置信区间和 σ 的双侧 95% 置信区间.

解　根据所测得的数据可得

$$\bar{x}=12.0875,\quad s=0.0712.$$

由表 2.3-1，得到

$$\bar{x}\pm t_{0.95}(15)\frac{s}{\sqrt{16}}=12.0875\pm 1.7531\times\frac{0.0712}{4}=12.0875\pm 0.0312.$$

因此，μ 的双侧 90% 置信区间为[12.0563，12.1187].

再则，

$$\frac{(n-1)s^2}{\chi^2_{0.975}(n-1)}=\frac{15\times(0.0712)^2}{27.488}=0.0028$$

$$\frac{(n-1)s^2}{\chi^2_{0.025}(n-1)}=\frac{15\times(0.0712)^2}{6.262}=0.0121.$$

于是，σ^2 的双侧 95% 置信区间为[0.0028，0.0121]. 由此得到 σ 的双侧 95% 置信区间为

$$[\sqrt{0.0028},\sqrt{0.0121}]=[0.0529,0.1100].$$ ∎

2.3 区间估计

例 4 对某型号飞机的飞行速度进行了 15 次试验，测得其最大飞行速度(单位：米/秒)为

422.2，417.2，425.6，420.3，425.8，423.1，418.7

428.2，438.3，434.0，412.3，431.5，413.5，441.3，423.0

根据长期经验可以认为最大飞行速度服从正态分布 $N(\mu;\sigma^2)$，求 μ 的单侧 95%置信区间下限和 σ 的单侧 90%置信区间上限.

解 根据所测得的数据，计算可得

$$\bar{x}=\frac{1}{15}\sum_{i=1}^{15}x_i=424.99,\quad s^2=\frac{1}{14}\sum_{i=1}^{15}(x_i-\bar{x})^2=71.859.$$

由表 2.3-1 分别得到

$$\bar{x}-t_{0.95}(n-1)\frac{s}{\sqrt{n}}=424.99-1.7613\times\frac{8.4770}{\sqrt{15}}=421.135,$$

$$\frac{(n-1)s^2}{\chi^2_{0.10}(n-1)}=\frac{14\times 71.859}{7.790}=129.14.$$

因此，μ 的单侧 95%的置信下限为 421.135 米/秒；σ^2 的单侧 90%的置信上限为 129.14 米^2/秒^2，从而 σ 的单侧 90%的置信上限为 11.36 米/秒. ▌

2.3 区间估计

2. 两个正态总体的情况

在实际问题中经常会遇到需要同时处理两个正态总体的情况. 例如, 比较两种不同型号的灯泡寿命时就有两个总体, 同一种型号的灯泡是一个总体.

对于两个正态总体, 我们用 $\overline{X}$ 与 S_1^2 分别表示总体 $N(\mu_1;\sigma_1^2)$ 的容量为 m 的样本 $(X_1, X_2, \cdots, X_m)$ 的样本均值与方差, 而用 $\overline{Y}$ 与 S_2^2 分别表示总体 $N(\mu_2;\sigma_2^2)$ 的容量为 n 的样本 $(Y_1, Y_2, \cdots, Y_n)$ 的样本均值与方差, 并且假定这两个样本相互独立.

(1) μ_1 与 μ_2 均未知, 但 σ_1^2 与 σ_2^2 都已知, 求未知参数 $\mu_1-\mu_2$ 的置信区间

$\mu_1-\mu_2$ 的极大似然估计量是 $\overline{X}-\overline{Y}$, 且

$$J=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}\sim N(0;1),$$

因而 $J=J(X_1,\cdots,X_m,Y_1,\cdots,Y_n;\mu_1-\mu_2)$. 于是, 将

$$P\{-u_{1-\alpha/2}\leqslant J\leqslant u_{1-\alpha/2}\}=1-\alpha$$

等价地表示为

$$P\left\{(\overline{X}-\overline{Y})-u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\leqslant\mu_1-\mu_2\leqslant(\overline{X}-\overline{Y})+u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\right\}=1-\alpha, \tag{2.3-9}$$

2.3 区间估计

即可得到 $\mu_1-\mu_2$ 的双侧 $1-\alpha$ 置信区间为

$$\left[(\overline{X}-\overline{Y})-u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}},(\overline{X}-\overline{Y})+u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\right].$$

(2) μ_1 与 μ_2 均未知，且 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 σ^2 未知，求未知参数 $\mu_1-\mu_2$ 的置信区间

这时，由于 J 中包含 σ_1^2、σ_2^2，因而不符合要求. 考虑到(1.4-4)式，取

$$J=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{m}+\frac{1}{n}}},$$

则有 $J\sim t(m+n-2)$，且 $J=J(X_1,\cdots,X_m,Y_1,\cdots,Y_n;\mu_1-\mu_2)$，从而由

$$P\{-t_{1-\alpha/2}(m+n-2)\leqslant J\leqslant t_{1-\alpha/2}(m+n-2)\}=1-\alpha$$

可得 $\mu_1-\mu_2$ 的双侧 $1-\alpha$ 置信区间为

$$\left[(\overline{X}-\overline{Y})-t_{1-\alpha/2}(m+n-2)S_w\sqrt{\frac{1}{m}+\frac{1}{n}},(\overline{X}-\overline{Y})+t_{1-\alpha/2}(m+n-2)S_w\sqrt{\frac{1}{m}+\frac{1}{n}}\right].$$

2.3 区间估计

(3) μ_1、μ_2，σ_1^2、σ_2^2 均未知，求$\dfrac{\sigma_1^2}{\sigma_2^2}$的置信区间

这种统计问题实际上是要求比较两个正态总体方差 σ_1^2 与 σ_2^2 之间的差异.

考虑到(1.4-5)式，取 $$J=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2},$$

则 $J\sim F(m-1,n-1)$，且 $J=J(X_1,\cdots,X_m,Y_1,\cdots,Y_n;\dfrac{\sigma_1^2}{\sigma_2^2})$，从而将

$$P\{F_{\alpha/2}(m-1,n-1)\leqslant J\leqslant F_{1-\alpha/2}(m-1,n-1)\}=1-\alpha$$

等价地表示为

$$P\left\{\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\alpha/2}(m-1,n-1)}\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{S_1^2}{S_2^2}\frac{1}{F_{\alpha/2}(m-1,n-1)}\right\}=1-\alpha,\quad(2.3\text{-}10)$$

即可得到$\dfrac{\sigma_1^2}{\sigma_2^2}$的双侧 $1-\alpha$ 置信区间为

$$\left[\frac{S_1^2}{S_2^2}\frac{1}{F_{1-\alpha/2}(m-1,n-1)},\frac{S_1^2}{S_2^2}\frac{1}{F_{\alpha/2}(m-1,n-1)}\right].$$

由此也可推出$\dfrac{\sigma_1}{\sigma_2}$的双侧 $1-\alpha$ 置信区间

$$\left[\frac{S_1}{S_2}\frac{1}{\sqrt{F_{1-\alpha/2}(m-1,n-1)}},\frac{S_1}{S_2}\frac{1}{\sqrt{F_{\alpha/2}(m-1,n-1)}}\right].$$

2.3 区间估计

类似地可得到 $\mu_1-\mu_2$ 及 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的单侧置信区间下限或上限. 我们将有关的结果列成表 2.3-2.

表 2.3-2　两个正态总体下未知参数的 $1-\alpha$ 置信区间

未知参数		随机变量 J	J 的分布	双侧置信区间下、上限	单侧置信下限	单侧置信上限
$\mu_1-\mu_2$	σ_1^2,σ_2^2 均已知	$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}}$	$N(0,1)$	$(\bar{X}-\bar{Y})\mp u_{1-\alpha/2}\cdot\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}$	$(\bar{X}-\bar{Y})-u_{1-\alpha}\cdot\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}$	$(\bar{X}-\bar{Y})+u_{1-\alpha}\cdot\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}$
	$\sigma_1^2=\sigma_2^2=\sigma^2$ 但未知	$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}$	$t(m+n-2)$	$(\bar{X}-\bar{Y})\mp t_{1-\alpha/2}(m+n-2)\cdot S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}$	$(\bar{X}-\bar{Y})-t_{1-\alpha}(m+n-2)\cdot S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}$	$(\bar{X}-\bar{Y})+t_{1-\alpha}(m+n-2)\cdot S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}$
$\dfrac{\sigma_1^2}{\sigma_2^2}$	μ_1,μ_2 均已知	$\dfrac{\sum_{i=1}^{m}(X_i-\mu_1)^2/m\sigma_1^2}{\sum_{i=1}^{n}(Y_i-\mu_2)^2/n\sigma_2^2}$	$F(m,n)$	$\dfrac{n\sum_{i=1}^{m}(X_i-\mu_1)^2}{m\sum_{i=1}^{n}(Y_i-\mu_2)^2}\cdot\dfrac{1}{F_{1-\alpha/2}(m,n)}$, $\dfrac{n\sum_{i=1}^{m}(X_i-\mu_1)^2}{m\sum_{i=1}^{n}(Y_i-\mu_2)^2}\cdot\dfrac{1}{F_{\alpha/2}(m,n)}$	$\dfrac{n\sum_{i=1}^{m}(X_i-\mu_1)^2}{m\sum_{i=1}^{n}(Y_i-\mu_2)^2}\cdot\dfrac{1}{F_{1-\alpha}(m,n)}$	$\dfrac{n\sum_{i=1}^{m}(X_i-\mu_1)^2}{m\sum_{i=1}^{n}(Y_i-\mu_2)^2}\cdot\dfrac{1}{F_{\alpha}(m,n)}$
	μ_1,μ_2 均未知	$\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$	$F(m-1,n-1)$	$\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{1-\alpha/2}(m-1,n-1)}$, $\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{\alpha/2}(m-1,n-1)}$	$\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{1-\alpha}(m-1,n-1)}$	$\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{\alpha}(m-1,n-1)}$

2.3 区间估计

例 5 设自总体 $N(\mu_1;25)$ 得到一个大小为 10 的样本，其样本均值 $\bar{x}=19.8$；而自总体 $N(\mu_2;36)$ 得到一个大小为 12 的样本，其样本均值 $\bar{y}=24.0$，并且两者独立，求 $\mu_1-\mu_2$ 的双侧 90%置信区间.

解 由于 $\sigma_1^2=25$、$\sigma_2^2=36$ 均已知，且 $m=10$，$n=12$，查标准正态分布表得 $u_{1-\alpha/2}=u_{0.95}=1.645$，所以从表 2.3-2 得到 $\mu_1-\mu_2$ 的双侧 90%置信区间上、下限为

$$(\bar{x}-\bar{y})\pm u_{1-\alpha/2}\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}=(19.8-24.0)\pm 1.645\times\sqrt{\frac{25}{10}+\frac{36}{12}}=-4.200\pm 3.858.$$

■

2.3 区间估计

例 6 为比较甲、乙两种型号灯泡的寿命，随机地抽取甲种型号灯泡 5 只做试验，测得平均寿命 $\bar{x}=1000$ 小时，样本标准差 $s_1=28$ 小时；又随机地抽取乙种型号灯泡 7 只做试验，测得平均寿命 $\bar{y}=980$ 小时，样本标准差 $s_2=32$ 小时. 假定两种型号灯泡的寿命服从各自的正态分布，且方差相同，求甲、乙两种型号灯泡的平均寿命之差的双侧 95% 置信区间.

解 因为 $m=5$，$n=7$，$\alpha=0.05$，查 t 分布表可得 $t_{0.975}(5+7-2)=t_{0.975}(10)=2.2281$. 又

$$s_w=\sqrt{\frac{(m-1)s_1^2+(n-1)s_2^2}{m+n-2}}=\sqrt{\frac{4\times 28^2+6\times 32^2}{10}}=30.46,$$

$\bar{x}-\bar{y}=20$，故由表 2.3-2 得到 $\mu_1-\mu_2$ 的双侧 95% 置信区间上、下限为

$$(\bar{x}-\bar{y})\pm t_{0.975}(10)s_w\sqrt{\frac{1}{m}+\frac{1}{n}}=20\pm 2.2281\times 30.46\times\sqrt{\frac{1}{5}+\frac{1}{7}}$$

$$=20\pm 39.74.$$ ∎

2.3 区间估计

例 7 两个正态总体 $N(\mu_1;\sigma_1^2)$、$N(\mu_2;\sigma_2^2)$的参数 $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ 均未知，今依次从这两个正态总体中随机抽取容量为 15 与 21 的两个独立样本，算得样本方差分别为 $s_1^2=6.38$ 与 $s_2^2=5.15$，求这两个总体标准差之比 $\frac{\sigma_1}{\sigma_2}$ 的双侧 90%置信区间.

解 因为 $m=15,n=21,\alpha=0.10$，查 F 分布表得 $F_{0.95}(14,20)=2.22$，$F_{0.05}(14,20)=1/F_{0.95}(20,14)=1/2.39=0.42$；又因 $s_1^2/s_2^2=6.38/5.15=1.24$，所以从表 2.3-2 可得 σ_1^2/σ_2^2 的双侧 90%置信区间下、上限分别为

$$\frac{s_1^2}{s_2^2}\,\frac{1}{F_{1-\alpha/2}(m-1,n-1)}=1.24\times\frac{1}{2.22}=0.56,$$

$$\frac{s_1^2}{s_2^2}\,\frac{1}{F_{\alpha/2}(m-1,n-1)}=1.24\times 2.39=2.96.$$

于是，σ_1/σ_2 的双侧 90%置信区间是

$$[\sqrt{0.56},\sqrt{2.96}]=[0.75,1.72].$$ ■

2.3 区间估计

3. 非正态总体情况

对于非正态总体，上述求置信区间的方法仍然有效，只不过要利用中心极限定理，对大样本（样本容量 $n>30$）作近似估计，这时点估计的渐近正态性起着重要的作用。例如，设 $(X_1,X_2,\cdots,X_n)$ 是取自总体 X 的样本，要求未知参数 $\mu=E(X)$ 的双侧 $1-\alpha$ 置信区间。尽管不知道总体 X 究竟服从什么分布，由于当 n 充分大时，$J=\sqrt{n}\dfrac{\overline{X}-\mu}{S}$ 近似地服从标准正态分布 $N(0;1)$，所以在大样本时，μ 的双侧 $1-\alpha$ 置信区间近似地为

$$\left[\overline{X}-\mu_{1-\alpha/2}\frac{S}{\sqrt{n}},\overline{X}+u_{1-\alpha/2}\frac{S}{\sqrt{n}}\right]. \tag{2.3-11}$$

类似地可得到 μ 的单侧 $1-\alpha$ 置信区间下限或上限。

2.3 区间估计

例 8 从一大批产品中随机地抽取 100 件进行检查，发现有 4 件是不合格品，求不合格率 p 的单侧 95%置信区间上限的近似值.

解 由于总体 X 服从 0-1 分布 $B(1;p)$，其中 $0\leqslant p\leqslant 1$，但 p 未知，所以 p 的极大似然估计量是 $\overline{X}$. 又因当 n 充分大时，$\sqrt{n}\dfrac{\overline{X}-p}{\sqrt{\overline{X}(1-\overline{X})}}$ 近似地服从标准正态分布 $N(0;1)$，故 p 的单侧 95%置信区间上限的近似值为 $\overline{x}+u_{0.95}\sqrt{\dfrac{1}{n}\overline{x}(1-\overline{x})}$.

现 $n=100$，$\overline{x}=0.04$，因而所要求的上限的近似值为

$$0.04+1.645\times\sqrt{\frac{1}{100}\times 0.04\times 0.96}=0.072.$$

■

2.3 区间估计

在讨论两个正态总体均值之间的差异时，我们假定要么 σ_1^2 与 σ_2^2 均已知，要么 $\sigma_1^2=\sigma_2^2=\sigma^2$ 但 σ^2 未知. 若取消两个总体方差齐性的假定，即 $\sigma_1^2\neq\sigma_2^2$，且 σ_1^2、σ_2^2 均未知时，如何求 $\mu_1-\mu_2$ 的置信区间呢？这时，$\mu_1-\mu_2$ 的极大似然估计仍是 $\overline{X}-\overline{Y}$，而用样本方差 S_1^2、S_2^2 分别代替 $\sigma^{2\,1}$、$\sigma^{2\,2}$，令

$$J=\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{S_1^2}{m}+\dfrac{S_2^2}{n}}},$$

可以证明，当 m、n 都很大时，近似地有 $J\sim N(0;1)$. 从而 $\mu_1-\mu_2$ 的双侧（近似）$1-\alpha$ 置信区间为

$$\left[(\overline{X}-\overline{Y})-u_{1-\alpha/2}\sqrt{\frac{S_1^2}{m}+\frac{S_2^2}{n}},(\overline{X}-\overline{Y})+u_{1-\alpha/2}\sqrt{\frac{S_1^2}{m}+\frac{S_2^2}{n}}\right]. \tag{2.3-13}$$

2.3 区间估计

例 9 从某地区随机地选取男、女各 100 名测量身高. 测得男子高度的平均值为 171 厘米，标准差为 3.5 厘米；女子高度的平均值为 161 厘米，标准差为 3.8 厘米. 假定身高服从正态分布，且男女身高相互独立，求该地区男女平均身高之差的双侧 95% 置信区间.

解 因为 $m=n=100$，$\alpha=0.05$，$\bar{x}=171$，$s_1^2=(3.5)^2=12.25$，$\bar{y}=161$，$s_2^2=(3.8)^2=14.44$，查标准正态分布表得 $u_{0.975}=1.96$，所以由(2.3-13)式得到该地区男女平均身高之差的双侧 95% 置信区间为

$$\left[(171-161)-1.96\times\sqrt{\frac{12.25}{100}+\frac{14.44}{100}},\right.$$

$$\left.(171-161)+1.96\times\sqrt{\frac{12.25}{100}+\frac{14.44}{100}}\right]=[8.99,11.01].$$

■

作业（第2部分）

习题 2

12. 随机地从一批钉子中抽取 16 只，测得其长度(单位：厘米)为

$$2.14, 2.10, 2.13, 2.15, 2.13, 2.12, 2.13, 2.10,$$
$$2.15, 2.12, 2.14, 2.10, 2.13, 2.11, 2.14, 2.11$$

假定钉长分布是正态的，求总体均值 μ 的双侧 90% 置信区间：(1)若已知 $\sigma=0.01$ 厘米；(2)若 σ 未知.

17. 甲、乙两位化验员独立地对一种聚合物的含氯量用相同的方法各作了 10 次测定，得 $s_1^2=0.5419$，$s_2^2=0.6050$，求他们测定值的方差比的双侧 90% 置信区间，假定测定值服从正态分布.

20. 设从某种型号的一大批晶体管中随机抽取 100 只样品，测得其寿命标准差 $s=45$ 小时，求这批晶体管寿命标准差 σ 的双侧 95% 置信区间.

謝謝觀看！

厦門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黄 煒
博士，副教授
Dr. Wei Huang