财经、统计、管理专业教材

多元统计分析及程序

于秀林 编著

中国统计出版社

财经、管理、统计等专业教材

多元统计分析及计算程序

于秀林　编著

（京）新登字041号

内 容 简 介

本书编写特点着重于实际应用，主要介绍多元统计分析中常用的各种方法，即多对多双重筛选逐步回归分析、判别分析、聚类分析、主成分分析、因子分析、对应分析、典型相关分析、路径分析、多维标度法等。为读者使用方便，对常用的方法给出BASIC语言程序，特别是双重筛选逐步回归分析程序，目前在其它书上很少见到。为加深对各种方法的理解，对理论部分作了适当的论证，但略去复杂的数学推导，本书可作为高等院校财经、管理、统计等专业的教材，也可作为在职科技人员处理和解决实际问题的参考书。

多元统计分析及计算程序

DUOYUAN TONGJI FENXI JI JISHUAN CHENGXU

于秀林　编著

中国统计出版社出版

（北京三里河月坛南街38号　100826）

北京市大兴县京南印刷厂印刷

787×1092毫米　16开本　15.5印张　30万字

1993年6月第1版　1993年6月北京第1次印刷

印数：1-2000

ISBN　7-5037-1235-X/C·772

定价：7.8元

前　　言

多元统计分析是数理统计中近二十余年发展最迅速的一个分支，随着电子计算机使用的日益广泛，使多元统计分析的方法很快应用到自然科学、社会科学以及经济领域中，实践证明它是一种很有用的处理数据的方法。它包括的主要内容：就其理论部分来说有统计量的分布、多元正态总体的参数估计和假设检验。就其方法部分来说有多对多双重筛选逐步回归分析、判别分析、聚类分析、主成分分析、因子分析、对应分析、典型相关分析、路径分析、多维标度法等。本书重点介绍多元统计分析中常用的各种方法，而对多元统计分析的理论内容，只是为了后面的需要做简单的概述。

本书特点：1.以结合经济实例为主，在一元统计分析的基础上，深入浅出地介绍多元统计分析的内容，并着重介绍多元统计分析中常用的各种方法，考虑到例题的典型性和多元统计分析应用的广泛性，也适当给出其它领域的例题。2.为了便于读者学完各种方法后，能用电子计算机进行计算，本书对主要方法附有BASIC语言程序。3.除介绍通常的几种方法（多元回归分析、判别分析、聚类分析、主成分分析、因子分析、对应分析、典型相关分析）外，还增加了目前在很多领域中已引起关注和应用的路径分析和多维标度法，目的使读者对多元统计分析的基本内容有个初步的较全面的了解。

为使本书能适合不同层次读者的需要和加深对各种方法的理解，以及期望读者能灵活地运用和推广这些方法，作者对各种方法的理论部分做了适当的论证和说明，但略去较深的数学推导，因此学习本书时只要具备初等数理统计和线性代数的基本知识即可阅读。

本书是在中国人民大学统计系开设多元统计分析课程以及科研工作的基础上整理编写而成的，由于作者水平有限，书中难免有错，敬请读者批评指正。

作者

1991年8月于北京

目　　录

第一章　预备知识--矩阵、行列式和消去变换

矩阵和行列式是研究多元统计分析的重要工具，虽然绝大多数读者已有矩阵代数的基础，但是为了后面学习的方便，本章对矩阵中必要的有关知识作简单的复习。

§1.　矩阵

1.1　矩阵的定义

将$n \times p$个实数a_{11}，a_{12}，…，a_{1p}，a_{21}，a_{22}，…，a_{2p}，…，a_{n1}，a_{n2}，…，a_{np}排成如下的形式，并且记为A：

$$A=\begin{pmatrix} a_{11} & a_{12}\cdots a_{1p} \\ a_{21} & a_{22}\cdots a_{2p} \\ \vdots & \vdots \quad\quad \vdots \\ a_{n1} & a_{n2}\cdots a_{np} \end{pmatrix}$$

则称A为$n \times p$阶矩阵，一般记为$A=(a_{ij})_{n \times p}$，其中a_{ij}是A中元素，它也可以是复数，但本书的a_{ij}均为实数。当$n=p$时，称A为n阶方阵；若$p=1$，A只有一列，称A为列向量，记为

$$a=\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix};$$

当$n=1$时，A只有一行，称A为行向量，记为

$$a'=(a_{11},\ a_{12},\ \cdots,\ a_{1n});$$

当A为n阶方阵，称a_{11}，a_{22}，…，a_{nn}为A的对角线元素，其它元素称为非对角元素，若方阵A的非对角元素全为零，称A为对角阵，记为

$$A=\begin{pmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}=\mathrm{diag}\ (a_{11},\ a_{22},\ \cdots,\ a_{nn});$$

进一步若$a_{11}=a_{22}=\cdots=a_{nn}=1$称$A=\begin{pmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & 1 \end{pmatrix}$

为n阶单位矩阵，记为I_n或$A=I$（I的阶数从上下文可明确）。

若A为n×p阶矩阵，它的转置A′是p×n阶矩阵，即

$$A'=\begin{pmatrix} a_{11} & a_{21}\cdots a_{n1} \\ a_{12} & a_{22}\cdots a_{n2} \\ \vdots & \vdots \qquad \vdots \\ a_{1p} & a_{2p} \qquad a_{np} \end{pmatrix};$$

若A是方阵，且$A'=A$，则称A为对称阵；若方阵$A=(a_{ij})$的元素$a_{ij}=0$对一切$i<j$成立，则称$A=\begin{pmatrix} a_{11} & & & 0 \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2}\cdots & & a_{nn} \end{pmatrix}$为下三角阵；若A′为下三角阵，则称A为上三角阵。

1.2 矩阵的运算

若A和B是n×m阵，则A和B的和定义为

$$A+B=(a_{ij}+b_{ij})$$

若α为一常数，它和A的积定义为

$$\alpha A=(\alpha a_{ij})$$

若A和B分别为p×q和q×r阵，则A和B的积定义为

$$AB=\left(\sum_{k=1}^{q} a_{ik}b_{kj}\right)$$

容易验证上述运算符合下面的规律：

$$A+(-1)A=0$$

$$(AB)'=B'A'$$

$$(A')'=A$$

$$(A+B)'=A'+B'$$

$$A(BC)=(AB)C$$

$$A(B+C)=AB+AC$$

$$(A+B)C=AC+BC$$

$$AI=IA=A$$

$$\alpha(A+B)=\alpha A+\alpha B$$

$$\alpha(AB)=(\alpha A)B=A(\alpha B)$$

若A为方阵满足$A'A=AA'=I$，则称A为正交阵

§2. 行列式、逆矩阵和矩阵的秩

2.1 行列式

若A为p阶方阵，记

$$|A|=\sum_{\pi}\varepsilon_{\pi}a_{1j_1},\ a_{2j_2}\cdots a_{pj_p}$$

其中 $(j_1, j_2, \cdots, j_p)$ 为 $(1, 2, \cdots, p)$ 的任一置换，$\sum_{\pi}$是对一切可能的$p!$ 个置换求和。$\varepsilon_\pi = 1$ 或 -1 取决于相应的是偶置换或奇置换。$|A|$ 称为A的行列式。

直接由行列式的定义来计算行列式是很麻烦的，通常利用行列式的如下的一些性质可以简化行列式的计算：

（1）若A的某行（或列）为零，则 $|A| = 0$；

（2）$|A| = |A'|$；

（3）将A的某行（或列）乘以数α所得矩阵的行列式等于$\alpha|A|$；

（4）若A的两行（或列）相同，则$|A| = 0$；

（5）若将A的两行（两列）互换，所得矩阵之行列式$= -|A|$；

（6）若将A的某一行（或列）乘上一个常数后加到另一行相应的元素上，所得矩阵的行列式不变仍等于$|A|$。

2.2 逆矩阵

设A为p阶方阵，若$|A| \neq 0$，则称A是非退化阵。

若A是p阶非退化阵，则存在唯一的矩阵B使得$AB = I_p$， B称为A的逆矩阵，并且记为 $B = A^{-1}$。B的元素b_{ij}可通过下式求得

$$b_{ij} = A_{ij}/|A|$$

其中A_{ij}为a_{ij}的代数余子式，即用 $(-1)^{i+j}$乘以a_{ij}的子式（将A的第i行和第j列去掉，得矩阵的行列式称为元素a_{ij}的子式）。

一般情况，这个求逆公式只有理论的价值，在多元分析中求逆矩阵是通过消去变换来实现，并且同时可求得该矩阵的行列式，消去变换将在后面介绍。

逆矩阵的基本性质如下：

（1）$AA^{-1} = A^{-1}A = I$；

（2）$(A')^{-1} = (A^{-1})'$；

（3）若A和C均为p阶非退化阵，则

$$(AC)^{-1} = C^{-1}A^{-1};$$

（4）设A为p阶非退化方阵，b和a为p维列向量，则方程

$$Ab = a$$

的解为

$$b = A^{-1}a;$$

（5）$|A^{-1}| = |A|^{-1}$，即逆矩阵行列式等于原矩阵行列式的逆；

（6）若A是正交阵，由于$AA' = I$推知$A^{-1} = A'$；若A是对角阵$A = \mathrm{diag}(a_{11}, a_{22}, \cdots, a_{pp})$且$a_{ii} \neq 0$，$i = 1, 2, \cdots, p$，则 $A^{-1} = \mathrm{diag}(a_{11}^{-1}, a_{22}^{-1}, \cdots, a_{pp}^{-1})$；

（7）上（下）三角阵的逆仍为上（下）三角阵。

2.3 矩阵的秩

设A为$p \times q$矩阵，若存在它的一个r阶子方阵的行列式不为零，而A的一切$(r+1)$阶子方阵的行列式均为零，则称A的秩为r，记作$rk(A) = r$。它有如下的基本性质：

（1）$rk(A) = 0$ 当且仅当$A = 0$；

（2）若A为$p \times q$阶阵，则$0 \leqslant rk(A) \leqslant \min(p, q)$；

（3） $rk(A)=rk(A')$；

（4） $rk(AB)\leqslant \min(rk(A),\ rk(B))$；

（5） $rk(A+B)\leqslant rk(A)+rk(B)$；

（6） 若A和C为非退化方阵，则

$$rk(ABC)=rk(B)。$$

§3. 特征根、特征向量和矩阵的迹

3.1 特征根和特征向量

设A为p阶方阵，则方程$|A-\lambda I_p|=0$是λ的p次多项式，由多项式理论知道必有p个根（可以有重根），记为λ_1，λ_2，…，λ_p，称为A的特征根或特征值。

若存在一个p维向量l_i使得$(A-\lambda_i I_p)l_i=0$，则称l_i为对应于λ_i的A的特征向量。今后总假设$l_i'l_i=1$。

特征根有如下一些性质：

（1） 若A为实对称阵，则A的特征根全为实数，故可按大小次序排成$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_p$。若$\lambda_i\neq\lambda_j$，则相应的特征向量$l_i$和$l_j$必正交，即$l_i'l_j=0$；

（2） A和A′有相同的特征根；

（3） 若A和B是$p\times q$和$q\times p$阶阵，则AB和BA有相同的非零特征根；

（4） 设A的特征根是λ_1，λ_2，…，λ_p，则CA的特征根为$c\lambda_1$，$c\lambda_2$，…，$c\lambda_p$；$A-cI_p$的特征根为$\lambda_1-c,\lambda_2-c,\cdots,\lambda_p-c$；$A^{-1}$的特征根为$\lambda_1^{-1}$，$\lambda_2^{-1}\cdots\lambda_p^{-1}$（若$A^{-1}$存在）；

（5） 设A为正交阵，则$|\lambda_i(A)|=1$，$i=1,2,\cdots,p$；

（6） 若$A=diag(a_1,\ a_2,\ \cdots,\ a_p)$，则$a_1$，$a_2$，…，$a_p$为A的特征根；$e_1=(1,\ 0,\ \cdots,\ 0)'$，$e_2=(0,\ 1,\ \cdots,\ 0)'$，…，$e_p=(0,\ \cdots,\ 0,\ 1)'$为相应的特征向量；

（7） 若A为三角阵（上三角或下三角），则A的特征根为其对角元素。

3.2 矩阵的迹

若A是p阶方阵，它的对角元素之和称为A的迹，记为$tr(A)=\sum\limits_{i=1}^{p}a_{ii}$。

若A是p阶方阵，它的特征根为λ_1，λ_2，…，λ_p则可以证明 $tr(A)=\sum\limits_{i=1}^{p}\lambda_i$。这表明特征根和迹有密切关系。

迹有如下性质：

（1） $tr(AB)=tr(BA)$；

（2） $tr(A)=tr(A')$；

（3） $tr(A+B)=tr(A)+tr(B)$。

§4. 正定阵和非负定阵

若A是p阶对称阵，且对p维空间R^p里的一切非零向量X都有$X'AX>0$，则称A是正定阵，记为$A>0$；若对一切$X\in R^p$有$X'AX\geqslant 0$，则称A是非负定阵，记为$A\geqslant 0$；记$A\geqslant B$表示$A-B\geqslant 0$；记$A>B$表示$A-B>0$。

正定阵和非负定阵有如下一些性质：

（1）一个对称阵是正（非负）定的当且仅当它的特征根为正（非负）；

（2）$BB' \geqslant 0$ 对一切矩阵B成立；

（3）若 $A>0$，则 $A^{-1}>0$；

（4）设 $A \geqslant 0$，则 $A>0$ 当且仅当 $|A| \neq 0$；

（5）设 $A>0$ 是 $p \times p$ 阶阵，B是 $q \times q$ 阶阵（$q \leqslant p$，）$rk(B)=r$，则 $BAB' \geqslant 0$，且 $BAB'>0$ 当且仅当 $r=q$；

（6）若 $A>0$，则 $cA>0$，其中c为正数；

（7）若 $A>0$，$B>0$，$A-B>0$，则 $B^{-1}-A^{-1}>0$ 且 $|A|>|B|$。

§5. 消去变换（矩阵变换）

在多元分析中经常要解线性方程组，求矩阵的逆和行列式，或进行某种递推运算，通过消去变换可以达到目的。

设 $A=(a_{ij})$ 是 $n \times m$ 阵，若 $a_{ij} \neq 0$，将A变换为 $A^*=(a^*_{ij})$ 使得

$$a^*_{\alpha\beta}=\begin{cases} a_{\alpha\beta}-a_{i\beta}a_{\alpha j}/a_{ij} & \text{当}\alpha\neq i,\ \beta\neq j \\ -a_{\alpha j}/a_{ij} & \text{当}\alpha\neq i,\ \beta=j \\ a_{i\beta}/a_{ij} & \text{当}\beta\neq j,\ \alpha=i \\ 1/a_{ij} & \text{当}\alpha=i,\ \beta=j \end{cases}$$

这个变换称为以（i，j）为枢轴的消去变换或称为矩阵变换。记作 $A^*=T_{ij}A$。

消去变换有如下性质：

（1）$T_{ij}(T_{ij}A)=A$，即对A连续施行两次（i，j）消去变换，其结果仍是A不变；

（2）若 $i\neq k$，$j\neq l$，则 $T_{ij}(T_{kl}A)=T_{kl}(T_{ij}A)$。

在后面介绍的回归分析和判别分析中将使用消去变换来筛选变量。

第二章　多元正态分布、参数估计和假设检验

多元正态分布在多元统计分析中占有相当重要的地位，这是因为迄今为止多元统计分析的主要理论都是建立在多元正态总体的基础上，多元正态分布是多元统计分析的基础。此外许多实际问题的分布常常是多元正态分布或近似正态分布，或虽然本身不是正态分布，但是它的样本均值近似于多元正态分布。基于此，介绍多元统计分析的种种具体方法之前，在本章里简单的介绍多元正态分布的定义及多元正态分布中参数的估计和假设检验。

若一个向量的分量都是随机变量，则称为随机向量；它是描述多个随机现象问题的基本工具。

本章在介绍多元正态分布之前，先回顾一下一元正态分布，然后推广到多元正态分布。

§1.复习一元正态分布

1.1　分布函数、数学期望和方差

设X是一个随机变量，称 $F(x) \triangleq P(X \leqslant x)$ 为X的概率分布函数，或简称为分布函数，记为 $X \sim F(x)$。

若随机变量以概率为1地在有限或可列个值 $\{x_K\}$ 上取值，记

$P(X=x_K)=P_K\ (k=1,2,\cdots)$ 且 $\sum\limits_K P_K=1$，则称X为离散型随机变量，并称 $P(X=x_K)=P_K,\ (k=1,2,\cdots)$ 为X的概率分布。

若X为离散型随机变量，且级数 $\sum\limits_{k=1}^{\infty} x_K P_K$ 绝对收敛，则称此级数为X的数字期望，简称期望或均值，记为 $E(X)$，即 $E(X)=\sum\limits_{k=1}^{\infty} x_K P_K$。

若级数 $\sum\limits_{k=1}^{\infty}(x_K-EX)^2 P_K$ 收敛，则称此级数为X的方差，记为 $D(X)$，即

$D(X)=\sum\limits_{k=1}^{\infty}(x_K-EX)^2 p_k$，显然 $D(X)=E(X-EX)^2$，在不引起混淆的情况下，将X的期望和方差简写为EX和DX。

设 $X \sim F(x)$，若存在一个非负可积函数 $f(x)$，使得对一切实数x有

$$F(x)=\int_{-\infty}^{x} f(t)\,dt$$

则称X为连续型随机变量，称 $f(x)$ 为X的分布密度函数，简称为密度。

一个函数f（x）能作为某个随机变量X的分布密度的充要条件是

（1）$f(x) \geqslant 0$ 对一切实数x，

（2）$\int_{-\infty}^{\infty} f(x)\,dx = 1$。

若连续型随机变量X的分布密度函数f（x）存在，且 $\int_{-\infty}^{\infty} xf(x)\,dx$ 绝对收敛，则称此积分为X的数学期望，即 $E(X) = \int_{-\infty}^{\infty} xf(x)\,dx$。

若 $\int_{-\infty}^{\infty} (x-EX)^2 f(x)\,dx$ 收敛，则称此积分为X的方差，记为D（X），

即 $D(x) = \int_{-\infty}^{\infty} (x-EX)^2 f(x)\,dx$

1.2 期望和方差基本性质

设X为随机变量，k、b为常数，则

（1）$E(kX+b) = kE(x) + b$

（2）$E(x+y) = E(x) + E(y)$

（3）$D(kx) = k^2 D(X)$

（4）$D(X) = E(x^2) - (Ex)^2$

1.3 一元正态分布

若随机变量X有分布密度函数：

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{\frac{1}{2\sigma^2}(x-\mu)^2\right\} \qquad -\infty < x < \infty$$

则称X服从正态分布，记为 $X \sim N(\mu, \sigma^2)$，其中 $-\infty < \mu < \infty$，$\sigma^2 > 0$。μ为总体X的期望，σ^2为总体X的方差。

当总体的期望和方差未知时，一般用样本均值 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 和样本方差

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2$ 分别作为总体期望和方差的估计值。

1.4 一元正态分布的性质

若总体 $X \sim N(\mu, \sigma^2)$，$x_1, \cdots, x_n$ 是从总体X中抽取的样本，则 $\sum_{k=1}^{n} a_K x_K$（a_K 为已知常数）也服从正态分布 $N(\mu \sum_{k=1}^{n} a_K, \sigma^2 \sum_{k=1}^{n} a^2_K)$，显然当 $a_K = \frac{1}{n}$ $(k=1,2,\cdots,n)$，

易得$\overline{X} \sim N\left(\mu, \frac{1}{n}\sigma^2\right)$。

§2. 多元正态分布

2.1 多元分布函数、边缘分布、均值和协差阵

设$X=(X_1, X_2, \cdots, X_p)'$是P维随机向量，它的分布函数（或联合分布函数）定义为

$F(x)=F(x_1, x_2, \cdots, x_p)=P(X_1 \leqslant x_1, \cdots, X_p \leqslant x_p)$ 式中$X=(X_1, X_2, \cdots, X_p)' \in R^p$（$R^p$表示P维欧氏空间），并记为$X \sim F(x)$。

若某个随机向量以概率为1地在有限个或可列个 向 量 上取值，则称为离散型的随机向量。

若$X \sim F(x)=F(x_1, \cdots, x_p)$且存在一个非负函数$f(x_1, \cdots, x_p)$对一切$x=(x_1, \cdots, x_p)' \in R^p$有

$$F(x)=F(x_1, \cdots, x_p)=\int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(t_1, \cdots, t_p)\, dt_1 \cdots dt_p$$

则称$f(x_1, \cdots, x_p)$为$X=(X_1, X\cdot\cdot, X_p)'$的联合分布密度函数，简称为联合密度，称$X=(X_1, \cdots, X_p)'$为连续型随机向量。

若$X=(X_1, \cdots, X_p)'$是P维随机向量，由它的q（<p）个分量组成的子向量$X^{(1)}=(X_{i1}, \cdots X_{iq})'$的分布称为X的边缘（或边际）分布；通过变换X中各 分量的次序，总可假定$X^{(1)}$正好是X的前q个分量，其余p－q个分量为$X^{(2)}$，即

$$X=\begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix}\begin{matrix} q \\ p-q \end{matrix}$$ 。相应的取值也剖分为两部分 $x=\begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix}$。

这时$X^{(1)}$的分布函数为$F(x_1, \cdots, x_q)=P(x_1 \leqslant x_1, \cdots, X_q \leqslant x_q)$，

若X有联合分布密度$f(x_1, \cdots, x_p)$，则$X^{(1)}$的边缘密度为

$$f_1(x_1, \cdots, x_q)=\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_q, x_{q+1}, \cdots, x_p)\, dt_{q+1} \cdots dt_p。$$

若P个随机变量$X_1, \cdots, X_p$的联合分布等于各自的边缘分布的乘积，则称$X_1, \cdots, X_p$是相互独立的。称给定$X^{(2)}$时$X^{(1)}$的分布为条件分布。当X的密度函数为$f(x^{(1)}, x^{(2)})$ $X^{(2)}$的密度函数为$f_2(x_2)$时，给定$X^{(2)}$时$X^{(1)}$的条件密度为

$$f_1(x^{(1)}|x^{(2)})=\frac{f(x^{(1)}, x^{(2)})}{f_2(x^{(2)})}$$ 。

设$X=(X_1, \cdots, X_p)'$，$Y=(Y_1, \cdots, Y_q)'$，若$E(X_i)=\mu_i\ (i=1, \cdots, P)$ 存

在，则称 $E(X)=\begin{pmatrix} EX_1 \\ \vdots \\ EX_p \end{pmatrix}=\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} \overset{\triangle}{=} \mu$为随机向量X的均值。

称

$$E(X-EX)(X-EX)'=\begin{pmatrix} COV(X_1,X_1) & COV(X_1,X_2) & \cdots & COV(X_1,X_p) \\ COV(X_2,X_1) & COV(X_2,X_2) & \cdots & COV(X_2,X_p) \\ \vdots & \vdots & & \vdots \\ COV(X_p,X_1) & COV(X_p,X_2) & \cdots & COV(X_p,X_p) \end{pmatrix}$$

$\triangleq(\sigma_{ij})_{p\times p}$为随机向量X的协差阵。

称随机向量X和Y的协差阵为

$$COV(X,Y)=E(X-EX)E(Y-EY)$$

$$=\begin{pmatrix} COV(X_1,Y_1) & COV(X_1,Y_2) & \cdots & COV(X_1,Y_q) \\ COV(X_2,Y_1) & COV(X_2,Y_2) & \cdots & COV(X_2,Y_q) \\ \vdots & \vdots & & \vdots \\ COV(X_p,Y_1) & COV(X_p,Y_2) & \cdots & COV(X_p,Y_q) \end{pmatrix}。$$

称随机变量X的相关阵为$R=(r_{ij})\ p\times p$，其中

$$r_{ij}=\frac{COV(X_i,X_j)}{\sqrt{D(X_i)}\sqrt{D(X_j)}}=\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}\ (i,\ j=1,\ \cdots,\ P)。$$

2.2　均值向量和协差阵的性质

设X、Y是随机向量，A、B是常数矩阵，则

(1) $E(AX)=AE(X)$

(2) $E(AXB)=AE(X)B$

(3) $D(AX)=A(DX)A'$

(4) $COV(AX,BY)=A\,COV(X,Y)B'$

(5) $D(X)=V$是对称非负定阵。

上述性质，根据定义不难验证。

2.3　多元正态分布

若p维随机向量$X=(X_1,\ \cdots,\ X_p)'$的密度函数为

$$f(x_1,\ \cdots,\ x_p)=\frac{1}{(\sqrt{2\pi})^p|V|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(X-\mu)'V^{-1}(X-\mu)\right\}$$

其中μ是p维向量，V是p阶正定阵，则称X服从P元正态分布，也称X为P维正态随机向量，简记为$X\sim N_p(\mu,\ V)$。

可以证明μ为X的均值向量，V为X的协差阵。

2.4　多元正态分布的性质

(1) 若$X=(X_1,\ \cdots,\ X_p)'\sim N_p(\mu,\ V)$，V是对角形阵，则$X_1,\ \cdots,\ X_p$相互独立。

(2) 若$X=(X_1,\ \cdots,\ X_p)'\sim N_p(\mu,\ V)$，令$Y=AX$，A为p×p阶阵，则$X\sim N_p(A\mu,\ AVA')$。

§3.　多元正态分布中参数的估计

在实际应用中，多元正态分布中的参数μ和V，通常是未知的，需由样本来估计；而参数估计的方法很多，这里用最常见的且具有很多优良性质的最大似然法来获得μ和V的估

计。

设$X_{(1)}$，$X_{(2)}$，$\cdots$，$X^{(n)}$为来自正态总体$N_p(\mu, V)$容量为n的样本，每一个$X_{(\alpha)}$为一个样品，记$X_{(\alpha)}=(x_{\alpha 1}, x_{\alpha 2}, \cdots, x_{\alpha p})'$，$\alpha=1, 2, \cdots, n$，则样本资料阵为

$$X=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}=\begin{pmatrix} X'_{(1)} \\ X'_{(2)} \\ \vdots \\ X'_{(n)} \end{pmatrix}$$

用最大似然法求出

μ的估计：

$$\hat{\mu}=\overline{X}=\frac{1}{n}\sum_{\alpha=1}^{n} X_{(\alpha)}=\left(\frac{1}{n}\sum_{\alpha=1}^{n} x_{\alpha 1}, \quad \cdots, \quad \frac{1}{n}\sum_{\alpha=1}^{n} x_{\alpha p}\right)'$$

V的估计：

$$\hat{V}=\frac{1}{n}S=\frac{1}{n}\sum_{\alpha=1}^{n}(X_{(\alpha)}-\overline{X})(X_{(\alpha)}-\overline{X})'$$

$$=\left(\frac{1}{n}\sum_{\alpha=1}^{n}(x_{\alpha i}-\overline{x}_i)(x_{\alpha j}-\overline{x}_j)\right)_{p\times p}$$

§4 三个重要的统计量的分布

为了进行假设检验，本节介绍三个重要的统计量的分布。它们分别是一元统计中x^2分布t分布及F分布的推广。而且这三个统计量的分布在多元统计分析的理论研究中也是相当重要的。

4.1 Wishart分布

前边已指出，在实际应用中，对正态总体 μ和V的估计，常采用$\overline{X}$和$\frac{1}{n-1}S$分别作为μ和V的估计，但$\overline{X}$的分布由正态分布的性质知仍为正态分布，自然我们也希望知道随机矩阵$S=\sum_{\alpha=1}^{n}(X_{(\alpha)}-\overline{X})(X_{(\alpha)}-\overline{X})'$的分布是什么，这一节给出它的分布，并将会看到S的分布也是构成其它重要分布的基础。

(1) Wishart分布的定义

Wishart分布是一元统计中x^2分布的推广，这个分布是Wishart在1928年推导出来的，有人就用这个时间作为多元统计分析的诞生时间，可见Wishart分布的重要性。

设 $X_{(\alpha)}=(X_{\alpha 1}, X_{\alpha 2}, \cdots, X_{\alpha p})'\sim N_p(\mu, V)$ $(\alpha=1, 2, \cdots n)$ 且相互独立，则由$X_{(\alpha)}$组成的随机矩阵$\underset{p\times p}{W}=\sum_{\alpha=1}^{n} X_{(\alpha)}X'_{(\alpha)}$的分布称为非中心Wishart分布，记为$W_p(n, V, Z)$，其中$Z=\mu'\mu$；当$\mu=0$时称W为（中心）Wishart分布，记为$W_p(n, V)$，当$n>p$，$V>0$，$W_p(n, V)$有密度存在，其表达式为

$$f(W)=\begin{cases}\dfrac{|W|^{\frac{1}{2}(n-p-1)}\exp\left\{-\frac{1}{2}\mathrm{tr}V^{-1}W\right\}}{2^{np/2}\pi^{p(p-1)/4}\ |V|^{n/2}\prod\limits_{i=1}^{p}\Gamma\left(\frac{n-i+1}{2}\right)} & |W|>0\\ 0 & 其它\end{cases}$$

当$p=1$时，$X_{(\alpha)}$为一维正态分布$N(0, \sigma^2)$，

则$W=\sum\limits_{\alpha=1}^{n}X_{(\alpha)}X_{(\alpha)}=\sum\limits_{\alpha=1}^{n}X_{(\alpha)}^2$，显然此时$\frac{1}{\sigma^2}\sum\limits_{\alpha=1}^{n}X_{(\alpha)}^2\sim x^2(n)$，因此

Wishart分布是x^2分布在P维正态情况下的推广。

这里要说明一下，一个随机矩阵的分布一般是指该矩阵的列向量一个接一个地组成一个长向量的分布，若随机矩阵

$$X=\begin{pmatrix}X_{11} & X_{12} & \cdots & X_{1p}\\ X_{21} & X_{22} & \cdots & X_{2p}\\ \vdots & \vdots & & \vdots\\ X_{n1} & X_{n2} & \cdots & x_{np}\end{pmatrix}$$

为对称阵时，由于$X_{ij}=X_{ji}$，$p=n$，故只取其上三角部分组成一个长向量，即

$X=(X_{11}, \cdots, X_{1n}, X_{22}, \cdots, X_{2n}, \cdots, X_{n-1\,n-1}X_{n-1\,n}X_{nn})$。

(2) 性质

i) 若$X_{(\alpha)}\sim N_p(\mu, V)$ $(\alpha=1, 2, \cdots, n)$且相互独立，则离差阵

$S=\sum\limits_{\alpha=1}^{n}(X_{(\alpha)}-\overline{X})(X_{(\alpha)}-\overline{X})'\sim W_p(n-1, V)$ 其中$\overline{X}=\frac{1}{n}\sum\limits_{\alpha=1}^{n}X_{(\alpha)}$。

ii) 若$S_i\sim W_p(n_i, V)$，$(i=1, 2, \cdots, k)$且相互独立，则

$S=S_1+S_2+\cdots+S_K\sim W_p(n_1+n_2+\cdots+n_K, V)$。

iii) 若$\underset{(p\times p)}{X}\sim W_p(n, V)$，$\underset{(p\times p)}{C}$为非奇异阵，则

$C\times C'\sim W_p(n, CVC')$。

4.2 Hotelling T^2分布

Hotelling T^2分布是一元统计中t分布的推广。

(1) Hotelling T^2分布的定义

设$X\sim W_p(\mu, V)$，$S\sim W_p(n, V)$且X与S相互独立、$n>p$，则称统计量$T^2=nX'S^{-1}X$的分布为非中心Hotelling T^2分布，记为$T^2\sim T^2(p, n, \mu)$。为了纪念多元统计分析的先驱者Harold Hotelling而命名的，因为他首先提出了这一统计量的分布当$\mu=0$时称T^2服从（中心）Hotelling T^2分布，记为$T^2(p, n)$。

在一元统计中，若$X_1, \cdots, x_n$是来自总体$N(\mu, \sigma^2)$的样本，则

$t=\dfrac{\sqrt{n}(\overline{X}-\mu)}{\hat{\sigma}}\sim t(n-1)$ 分布，其中$\hat{\sigma}^2=\frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\overline{X})^2$。

显然