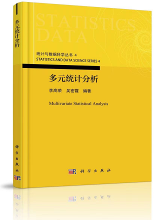
第4章
多元正态总体的抽样分布

李高荣

北京师范大学统计学院

E-mail: *ligaorong@bnu.edu.cn*



微信公众号：BNUlgr

- 扫二维码获取在线课件和相关教学电子资源
- 请遵守电子资源使用协议

- 统计量的分布称为抽样分布，在一元正态分布的研究中，导出了三个重要的抽样分布：χ^2分布、t分布和F分布。
- 这三个抽样分布已经被广泛用于正态总体的统计推断的问题中，为正态总体的统计推断提供了理论依据。
- 在多元正态分布统计推断的研究中，学者也导出了三个重要的抽样分布：
  1. Wishart分布
  2. Hotelling T^2分布
  3. Wilks分布

定义4.1：二次型

设$X_i \sim N_1(\mu_i, \sigma^2)(i = 1, \cdots, n)$的随机变量，并且相互独立。令$\boldsymbol{X} = (X_1, \cdots, X_n)'$，且$\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \sigma^2\mathbf{I}_n)$，其中$\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)'$。令

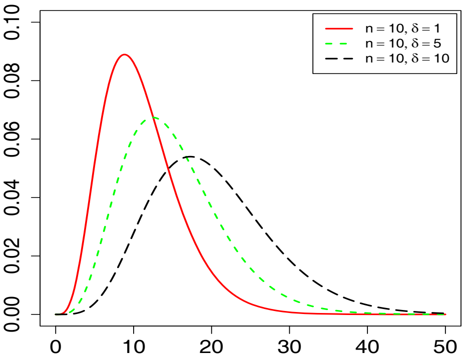
$$Y = \boldsymbol{X}'\boldsymbol{X} = \sum_{i=1}^{n} X_i^2$$

表示$\boldsymbol{X}$的二次型，则$Y/\sigma^2 = \boldsymbol{X}'\boldsymbol{X}/\sigma^2$服从自由度$n$，非中心参数$\delta = \boldsymbol{\mu}'\boldsymbol{\mu}/\sigma^2$的$\chi^2$分布，记为$Y/\sigma^2 \sim \chi_n^2(\delta)$或$Y \sim \sigma^2\chi_n^2(\delta)$。

$\chi_n^2(\delta)$分布具有概率密度函数：

$$f(x|n,\delta)=\begin{cases}\displaystyle\sum_{k=0}^{\infty}\frac{\exp\left(-\frac{\delta}{2}\right)\left(\frac{\delta}{2}\right)^k}{k!}\frac{\exp\left(-\frac{x}{2}\right)x^{\frac{n+2k}{2}-1}}{2^{\frac{n+2k}{2}}\Gamma\left(\frac{n+2k}{2}\right)}, & x>0;\\ 0, & x\le 0,\end{cases}$$

其中$\Gamma(\cdot)$为伽马(Gamma)函数。



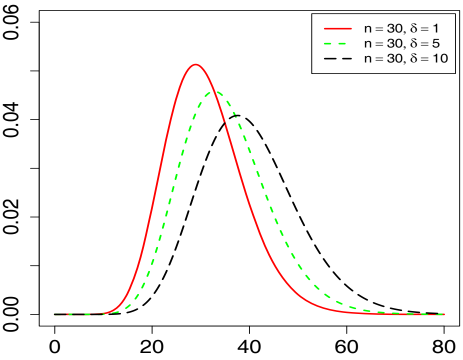


Figure 1: $\chi_n^2(\delta)$分布的概率密度函数。$E(Y)=n+\delta, \mathrm{Var}(Y)=2n+4\delta$。

下面讨论X的二次型Y的一些性质。

性质4.1.1

1. 当$\mu_i = 0, \sigma^2 = 1$，则$Y = X'X \sim \chi_n^2$;
2. 当$\mu_i = 0, \sigma^2 \neq 1$，则$Y/\sigma^2 = X'X/\sigma^2 \sim \chi_n^2$，或者记为$Y = X'X \sim \sigma^2\chi_n^2$;

   其中χ_n^2表示自由度是n的χ^2分布, $i = 1, \cdots, n$。

性质4.1.2

设$X \sim N_n(\mathbf{0}_n, \sigma^2\mathbf{I}_n)$，$\mathbf{A}$为对称矩阵，且$\text{rank}(\mathbf{A}) = r$，则二次型$X'\mathbf{A}X/\sigma^2 \sim \chi_r^2 \Longleftrightarrow \mathbf{A}^2 = \mathbf{A}$，即$\mathbf{A}$为对称幂等矩阵。

证明：“$\Rightarrow$:" 因为$\mathbf{A}$是对称矩阵，并且$\mathrm{rank}(\mathbf{A})=r$，则存在正交矩阵$\mathbf{\Gamma}$使得

$$\mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma}=\mathrm{diag}(\lambda_1,\cdots,\lambda_r,0,\cdots,0),$$

其中λ_i是$\mathbf{A}$的非零特征值，$i=1,\cdots,r$。

令$\mathbf{Z}=\mathbf{\Gamma}'\boldsymbol{X}\sim N_n(\mathbf{0}_n,\sigma^2\mathbf{I}_n)$，且$\boldsymbol{X}=\mathbf{\Gamma}\mathbf{Z}$，则

$$\xi=\boldsymbol{X}'\mathbf{A}\boldsymbol{X}/\sigma^2=\mathbf{Z}'\mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma}\mathbf{Z}/\sigma^2=\sum_{i=1}^{r}\lambda_i Z_i^2/\sigma^2,$$

又$Z_i\sim N(0,\sigma^2)(i=1,\cdots,r)$，并且相互独立。

对$i=1,\cdots,r$，故有Z_i^2/σ^2相互独立，且$Z_i^2/\sigma^2\sim\chi_1^2$。

$\sum_{i=1}^{r}\lambda_i Z_i^2/\sigma^2$的特征函数为

$$(1-2\mathrm{i}\lambda_1 t)^{-1/2}(1-2\mathrm{i}\lambda_2 t)^{-1/2}\cdots(1-2\mathrm{i}\lambda_r t)^{-1/2}.$$

由条件$\xi=\boldsymbol{X}'\mathbf{A}\boldsymbol{X}/\sigma^2\sim\chi_r^2$，故$\xi$的特征函数为$(1-2\mathrm{i}t)^{-r/2}$。由

$$(1-2\mathrm{i}\lambda_1 t)^{-1/2}(1-2\mathrm{i}\lambda_2 t)^{-1/2}\cdots(1-2\mathrm{i}\lambda_r t)^{-1/2}=(1-2\mathrm{i}t)^{-r/2}.$$

可得出：$\lambda_1=\lambda_2=\cdots=\lambda_r=1$。

因此，有

$$\mathrm{diag}(1,\cdots,1,0,\cdots,0)=\mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma}=\mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma}\cdot\mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma}=\mathbf{\Gamma}'\mathbf{A}^2\mathbf{\Gamma},$$

由上式证得$\mathbf{A}^2=\mathbf{A}$，即$\mathbf{A}$为对称幂等矩阵。

"$\Leftarrow$:" 既然$\mathbf{A}$为对称幂等矩阵，且$\mathrm{rank}(\mathbf{A})=r$，可知对称幂等矩阵$\mathbf{A}$的特征值非0即1，且只有$r$个为1的特征值，存在正交矩阵$\mathbf{\Gamma}$，使得

$$\mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma}=\begin{pmatrix}\mathbf{I}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}.$$

令$\mathbf{Z}=(Z_1,\cdots,Z_n)'=\mathbf{\Gamma}'\boldsymbol{X}$，则$\mathbf{Z}\sim N_n(\mathbf{0}_n,\sigma^2\mathbf{\Gamma}'\mathbf{I}_n\mathbf{\Gamma})=N_n(\mathbf{0}_n,\sigma^2\mathbf{I}_n)$，且

$$\frac{1}{\sigma^2}\boldsymbol{X}'\mathbf{A}\boldsymbol{X}=\frac{1}{\sigma^2}\mathbf{Z}'\mathbf{\Gamma}'\mathbf{A}\mathbf{\Gamma}\mathbf{Z}=\frac{1}{\sigma^2}\mathbf{Z}'\begin{pmatrix}\mathbf{I}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{Z}=\frac{1}{\sigma^2}\sum_{i=1}^{r}Z_i^2.$$

因为$Z_i\sim N(0,\sigma^2)(i=1,\cdots,r)$，且相互独立，故有

$$\xi=\frac{1}{\sigma^2}\boldsymbol{X}'\mathbf{A}\boldsymbol{X}=\frac{1}{\sigma^2}\sum_{i=1}^{r}Z_i^2\sim\chi_r^2.$$

性质4.1.3

设$\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \sigma^2\mathbf{I}_n)$，$\mathbf{A} = \mathbf{A}'$，则

$$\frac{1}{\sigma^2}\boldsymbol{X}'\mathbf{A}\boldsymbol{X} \sim \chi_r^2(\delta) \Longleftrightarrow \mathbf{A} = \mathbf{A}^2,$$

其中非中心参数为$\delta = \frac{1}{\sigma^2}\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$，$\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)'$，且$\text{rank}(\mathbf{A}) = r(r \leq n)$。

由定义4.1和性质4.1.2可得性质4.1.3的证明，留作习题课下练习。

性质4.1.4：二次型与线性函数的独立性

设$X \sim N_n(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_n)$, $\mathbf{A}$为n阶对称矩阵，$\mathbf{B}$为$m \times n$矩阵，令$\xi = X'\mathbf{A}X$, $\mathbf{Z} = \mathbf{B}X$ ($\mathbf{Z}$为m维的随机向量)，若$\mathbf{BA} = \mathbf{0}_{m\times n}$，则$\mathbf{Z}$与$\xi = X'\mathbf{A}X$相互独立。

性质4.1.4：二次型与线性函数的独立性

设$X \sim N_n(\mu, \sigma^2 \mathbf{I}_n)$, $\mathbf{A}$为n阶对称矩阵，$\mathbf{B}$为$m \times n$矩阵，令$\xi = X'\mathbf{A}X$, $\mathbf{Z} = \mathbf{B}X$ ($\mathbf{Z}$为m维的随机向量)，若$\mathbf{BA} = \mathbf{0}_{m\times n}$，则$\mathbf{Z}$与$\xi = X'\mathbf{A}X$相互独立。

- **课堂练习：** 利用性质4.1.4证明：$\bar{x} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n$ 和$\mathbf{V} = \mathbf{X}'\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\mathbf{X}$相互独立。
- **思考题：** 对多元线性模型$Y_{n\times 1} = \mathbf{X}_{n\times(p+1)}\beta_{(p+1)\times 1} + \varepsilon_{n\times 1}$中的应用，如$\beta$最小二乘估计和RSS的独立性。

证明： 设rank$(\mathbf{A})=r>0$ (当$r=0$时，$\mathbf{A}=0$，则结论显然成立), 存在正交矩阵$\boldsymbol{\Gamma}$使得

$$\boldsymbol{\Gamma}'\mathbf{A}\boldsymbol{\Gamma}=\mathrm{diag}(\lambda_1,\cdots,\lambda_r,0,\cdots,0),$$

其中λ_i是$\mathbf{A}$的非零特征值，$i=1,\cdots,r$。因为

$$\begin{aligned}\mathbf{BA} &= \mathbf{B}\boldsymbol{\Gamma}\mathrm{diag}(\lambda_1,\cdots,\lambda_r,0,\cdots,0)\boldsymbol{\Gamma}' \\ &= (\mathbf{C}_1,\mathbf{C}_2)\begin{pmatrix}\boldsymbol{\Lambda}_r & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(n-r)\times r} & \mathbf{0}_{(n-r)\times(n-r)}\end{pmatrix}\boldsymbol{\Gamma}' \\ &= (\mathbf{C}_1\boldsymbol{\Lambda}_r,\mathbf{0}_{m\times(n-r)})\boldsymbol{\Gamma}'=\mathbf{0}_{m\times n},\end{aligned}$$

其中$\boldsymbol{\Lambda}_r=\mathrm{diag}(\lambda_1,\cdots,\lambda_r)$，$\mathbf{C}_1$为$m\times r$矩阵，$\mathbf{C}_2$为$m\times(n-r)$ 矩阵。由条件$\mathbf{BA}=\mathbf{0}_{m\times n}$，且$\lambda_i\neq 0(i=1,\cdots,r)$，可得$\mathbf{C}_1=\mathbf{0}_{m\times r}$。

证明(续)： 令$\boldsymbol{Y} = \boldsymbol{\Gamma}'\boldsymbol{X}$，即$\boldsymbol{X} = \boldsymbol{\Gamma}\boldsymbol{Y}$。则

$$\boldsymbol{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \sim N_n(\boldsymbol{\Gamma}'\boldsymbol{\mu}, \sigma^2\mathbf{I}_n),$$

即$Y_1, \cdots, Y_n$相互独立。因为

$$\begin{aligned} \xi &= \boldsymbol{X}'\mathbf{A}\boldsymbol{X} = \boldsymbol{Y}'\boldsymbol{\Gamma}'\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{Y} = \sum_{i=1}^{r}\lambda_i Y_i^2, \\ \mathbf{Z} &= \mathbf{B}\boldsymbol{X} = \mathbf{B}\boldsymbol{\Gamma}\boldsymbol{Y} = (\mathbf{C}_1, \mathbf{C}_2)\begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} = \mathbf{C}_2\begin{pmatrix} Y_{r+1} \\ \vdots \\ Y_n \end{pmatrix}. \end{aligned}$$

由于$Y_1, \cdots, Y_r$和$Y_{r+1}, \cdots, Y_n$相互独立，$\mathbf{Z}$与$\xi = \boldsymbol{X}'\mathbf{A}\boldsymbol{X}$ 相互独立。

性质4.1.5: 两个二次型互相独立的条件

设$\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \sigma^2\mathbf{I}_n)$，且$\mathbf{A}$和$\mathbf{B}$是两个$n$阶对称矩阵，则

$$\mathbf{AB} = \mathbf{0}_{n\times n} \Longleftrightarrow \boldsymbol{X}'\mathbf{A}\boldsymbol{X}\text{与}\boldsymbol{X}'\mathbf{B}\boldsymbol{X}\text{相互独立}.$$

性质4.1.5的证明留作作业，课下完成。

下面讨论p元正态分布的二次型分布的性质：

性质4.1.6

设$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，$\boldsymbol{\Sigma} > 0$，则$\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} \sim \chi_p^2(\delta)$，其中$\delta = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$。

证明：因$\boldsymbol{\Sigma} > 0$，由正定矩阵的分解可得$\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}'$，其中$\mathbf{C}$为非退化的$p$维方阵。令$\boldsymbol{Y} = \mathbf{C}^{-1}\boldsymbol{X}$，即$\boldsymbol{X} = \mathbf{C}\boldsymbol{Y}$，则

$$\boldsymbol{Y} \sim N_p(\mathbf{C}^{-1}\boldsymbol{\mu}, \mathbf{C}^{-1}\boldsymbol{\Sigma}(\mathbf{C}^{-1})').$$

因为$\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}'$，所以$\boldsymbol{Y} \sim N_p(\mathbf{C}^{-1}\boldsymbol{\mu}, \mathbf{I}_p)$, 且有

$$\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} = \boldsymbol{Y}'\mathbf{C}'\boldsymbol{\Sigma}^{-1}\mathbf{C}\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{Y} \sim \chi_p^2(\delta),$$

其中$\delta = (\mathbf{C}^{-1}\boldsymbol{\mu})'(\mathbf{C}^{-1}\boldsymbol{\mu}) = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$。

性质4.1.7

设$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，$\boldsymbol{\Sigma} > 0$，$\mathbf{A}$为对称矩阵，$\mathrm{rank}(\mathbf{A}) = r$，则

$$(\boldsymbol{X} - \boldsymbol{\mu})'\mathbf{A}(\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi_r^2 \Leftrightarrow \boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}.$$

证明：因$\boldsymbol{\Sigma} > 0$，则有$\mathrm{rank}(\boldsymbol{\Sigma}) = p$。令存在正交矩阵$\boldsymbol{\Gamma}$和$\lambda_i > 0$, $i = 1, \cdots, p$，使得

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2},$$

其中$\boldsymbol{\Sigma}^{1/2} = \boldsymbol{\Gamma}\mathrm{diag}(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_p})\boldsymbol{\Gamma}'$ 为$\boldsymbol{\Sigma}$的平方根矩阵。

记
$$\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Gamma}\mathrm{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \cdots, \frac{1}{\sqrt{\lambda_p}}\right)\boldsymbol{\Gamma}'.$$

显然有$\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{-1/2} = \mathbf{I}_p$。令$\boldsymbol{Y} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu}) \sim N_p(\mathbf{0}, \mathbf{I}_p)$。我们有：

$$\begin{aligned}
\mathrm{Cov}(\boldsymbol{Y}) &= \mathrm{Cov}(\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu})) = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}(\boldsymbol{\Sigma}^{-1/2})' \\
&= \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2})\boldsymbol{\Sigma}^{-1/2} = \mathbf{I}_p,
\end{aligned}$$

$$(\boldsymbol{X} - \boldsymbol{\mu})'\mathbf{A}(\boldsymbol{X} - \boldsymbol{\mu}) = \boldsymbol{Y}'\boldsymbol{\Sigma}^{1/2}\mathbf{A}\boldsymbol{\Sigma}^{1/2}\boldsymbol{Y} =: \boldsymbol{Y}'\mathbf{C}\boldsymbol{Y}.$$

由性质4.1.2，有
$$\boldsymbol{Y}'\mathbf{C}\boldsymbol{Y} \sim \chi^2_r \Leftrightarrow \mathbf{C}^2 = \mathbf{C},$$

即$(\boldsymbol{\Sigma}^{1/2}\mathbf{A}\boldsymbol{\Sigma}^{1/2})(\boldsymbol{\Sigma}^{1/2}\mathbf{A}\boldsymbol{\Sigma}^{1/2}) = \boldsymbol{\Sigma}^{1/2}\mathbf{A}\boldsymbol{\Sigma}^{1/2}$，对其左右两边乘以$\boldsymbol{\Sigma}^{1/2}$，即得$\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{A}\boldsymbol{\Sigma}$。

性质4.1.8

设$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma), \Sigma > 0$，$\mathbf{A}$和$\mathbf{B}$为$p$阶对称矩阵，则

$$(\boldsymbol{X}-\boldsymbol{\mu})'\mathbf{A}(\boldsymbol{X}-\boldsymbol{\mu})\text{与}(\boldsymbol{X}-\boldsymbol{\mu})'\mathbf{B}(\boldsymbol{X}-\boldsymbol{\mu})\text{相互独立}$$

$$\Leftrightarrow \mathbf{A}\Sigma\mathbf{B} = \mathbf{0}_{p\times p}.$$

只证明充分性，必要性省略，留给大家！

证明：设$\Sigma^{1/2}\mathbf{A}\Sigma^{1/2}=\widetilde{\mathbf{A}}$, $\Sigma^{1/2}\mathbf{B}\Sigma^{1/2}=\widetilde{\mathbf{B}}$, $\boldsymbol{Y}=\Sigma^{-1/2}\boldsymbol{X}$，则$\boldsymbol{X}'\mathbf{A}\boldsymbol{X}=\boldsymbol{Y}'\widetilde{\mathbf{A}}\boldsymbol{Y}$, $\boldsymbol{X}'\mathbf{B}\boldsymbol{X}=\boldsymbol{Y}'\widetilde{\mathbf{B}}\boldsymbol{Y}$且$\boldsymbol{Y}\sim N_p(\Sigma^{-1/2}\boldsymbol{\mu},\mathbf{I}_p)$。
当$\mathbf{A}\Sigma\mathbf{B}=\mathbf{0}_{p\times p}$，有$\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}=\widetilde{\mathbf{B}}\widetilde{\mathbf{A}}=\mathbf{0}_{p\times p}$，故可以对 $\widetilde{\mathbf{A}}$ 和$\widetilde{\mathbf{B}}$ 同时进行对角化:

$$\Gamma'\widetilde{\mathbf{A}}\Gamma=\Lambda_1,\quad \Gamma'\widetilde{\mathbf{B}}\Gamma=\Lambda_2,$$

其中Γ为正交矩阵，Λ_i为对角阵，$i=1,2$。
由$\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}=\mathbf{0}_{p\times p}$，可知$\Lambda_1\Lambda_2=\mathbf{0}_{p\times p}$。因此

$$\boldsymbol{X}'\mathbf{A}\boldsymbol{X}=(\Gamma\boldsymbol{Y})'\Lambda_1(\Gamma\boldsymbol{Y}),\quad \boldsymbol{X}'\mathbf{B}\boldsymbol{X}=(\Gamma\boldsymbol{Y})'\Lambda_2(\Gamma\boldsymbol{Y}).$$

注意到：$\Gamma\boldsymbol{Y}\sim N_p(\Gamma\Sigma^{-1/2}\boldsymbol{\mu},\mathbf{I}_p)$，可知$\Gamma\boldsymbol{Y}$的每个分量相互独立，由$\Lambda_1\Lambda_2=\mathbf{0}_{p\times p}$，可知$\boldsymbol{X}'\mathbf{A}\boldsymbol{X}$和$\boldsymbol{X}'\mathbf{B}\boldsymbol{X}$ 依赖于$\Gamma\boldsymbol{Y}$的不同分量，因此相互独立。

- Wishart分布是一元统计中χ^2分布的推广，并且在样本协方差矩阵分析中具有非常重要的作用。

定义4.2: Wishart分布

设$\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$独立同$p$维正态分布$N_p(\mathbf{0}, \Sigma)$，记$\mathbf{X} = (\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n)'$为$n \times p$矩阵，则称$p$阶矩阵$\mathbf{W} = \mathbf{X}'\mathbf{X} = \sum_{i=1}^{n} \boldsymbol{X}_i\boldsymbol{X}_i'$的分布为$p$阶Wishart分布，简称Wishart分布，记为$\mathbf{W} \sim W_p(n, \Sigma)$，其中$n$称为它的自由度。

Wishart分布的定义

当$\boldsymbol{\Sigma} > 0$, $n \geq p$时，p阶Wishart分布的密度函数为：

$$\frac{1}{2^{np/2}\Gamma_p(n/2)|\boldsymbol{\Sigma}|^{n/2}}|\mathbf{W}|^{\frac{n-p-1}{2}}\exp\left\{-\frac{1}{2}\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W})\right\}, \quad \mathbf{W} > 0,$$

其中

- $\mathbf{W}$是对称矩阵
- $|\cdot|$表示矩阵的行列式
- $\Gamma_p(\cdot)$是p维Γ 函数，定义为

$$\Gamma_p(n/2) = \pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\left(\frac{n}{2} - \frac{i-1}{2}\right)$$

Wishart分布的性质

- 当$p=1$时，可知：$\mathbf{W} \sim \sigma^2 \chi_n^2$
- 当$p=1$，$\sigma^2=1$时，χ_n^2分布是Wishart分布的一种特殊情况

性质4.2.1：均值

若$\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$，则$E(\mathbf{W}) = n\mathbf{\Sigma}$。

性质4.2.2：变换

若$\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$，$\mathbf{C}$是$k \times p$阶矩阵，则$\mathbf{CWC}' \sim W_k(n, \mathbf{C\Sigma C}')$。

证明： 因$\mathbf{W} = \sum_{i=1}^{n} \boldsymbol{X}_i \boldsymbol{X}_i' \sim W_p(n, \boldsymbol{\Sigma})$，其中$\boldsymbol{X}_i \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$，$i = 1, \cdots, n$，相互独立。令$\boldsymbol{Y}_i = \mathbf{C}\boldsymbol{X}_i$，则$\boldsymbol{Y}_i \sim N_k(\mathbf{0}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}')$。因此，由Wishart分布的定义4.2可知

$$\mathbf{CWC}' = \sum_{i=1}^{n} \mathbf{C}\boldsymbol{X}_i \boldsymbol{X}_i' \mathbf{C}' = \sum_{i=1}^{n} \boldsymbol{Y}_i \boldsymbol{Y}_i' \sim W_k(n, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}').$$

性质4.2.3：特征函数

若$\mathbf{W} \sim W_p(n, \boldsymbol{\Sigma})$，则$\mathbf{W}$的特征函数为

$$E\Big(\exp(\mathrm{itr}(\mathbf{TW}))\Big) = |\mathbf{I}_p - 2\mathrm{i}\boldsymbol{\Sigma}\mathbf{T}|^{-n/2},$$

其中$\mathrm{i} = \sqrt{-1}$，$\mathbf{T}$为p阶实对称矩阵。

Wishart分布的性质

性质4.2.4：可加性

若$\mathbf{W}_i \sim W_p(n_i, \Sigma)(i = 1, \cdots, k)$相互独立，则

$$\sum_{i=1}^{k} \mathbf{W}_i \sim W_p(n, \Sigma), \qquad n = \sum_{i=1}^{k} n_i.$$

- 设$X_1, \cdots, X_n$独立同p维正态分布$N_p(\mathbf{0}, \Sigma)$，其中$\Sigma > 0$
- 记$\mathbf{X} = (X_1, \cdots, X_n)'$为$n \times p$矩阵，则称$\mathbf{Q} = \mathbf{X}'\mathbf{A}\mathbf{X}$为中心的矩阵二次型，简称为矩阵二次型，其中$n$阶方阵$\mathbf{A} \geq 0$
- 设$\mathbf{X} \sim N_{n\times p}(\mathbf{0}, \mathbf{I}_n \otimes \Sigma)$，则称$\mathbf{Q} = \mathbf{X}'\mathbf{A}\mathbf{X}$为矩阵二次型

Wishart分布的性质

性质4.2.5：矩阵二次型

1. 若$\mathbf{A}$为n阶幂等矩阵，则矩阵二次型$\mathbf{Q} = \mathbf{X}'\mathbf{A}\mathbf{X}$服从Wishart分布$W_p(m, \Sigma)$，其中$m = \mathrm{tr}(\mathbf{A})$。
2. 设$\mathbf{Q} = \mathbf{X}'\mathbf{A}\mathbf{X}$，$\mathbf{Q}_1 = \mathbf{X}'\mathbf{B}\mathbf{X}$是两个矩阵二次型，其中$\mathbf{A}$和$\mathbf{B}$都是$n$阶幂等矩阵。若$\mathbf{Q}_2 = \mathbf{Q} - \mathbf{Q}_1 \geq 0$，则$\mathbf{Q}_2$服从Wishart分布$W_p(m-r, \Sigma)$，其中$m = \mathrm{tr}(\mathbf{A})$，$r = \mathrm{tr}(\mathbf{B})$，且$\mathbf{Q}_1$与$\mathbf{Q}_2$相互独立。
3. 设$\mathbf{Q} = \mathbf{X}'\mathbf{A}\mathbf{X}$，$\mathbf{A}$为$n$阶对称幂等矩阵，则$\mathbf{P}'\mathbf{X}$与$\mathbf{Q}$独立的充分必要条件是$\mathbf{A}\mathbf{P} = \mathbf{0}$。

证明： (1) 当$\mathbf{A}$为n阶幂等矩阵时，存在正交矩阵$\mathbf{U}$，使得$\mathbf{A}=\mathbf{U}'\mathbf{\Gamma}\mathbf{U}$，其中

$$\mathbf{\Gamma}=\begin{array}{cc}\begin{pmatrix}\mathbf{I}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} & \begin{array}{c} m \\ n-m\end{array}\end{array}, \qquad m=\mathrm{tr}(\mathbf{A}).$$

由于$\mathbf{X}\sim N_{n\times p}(\mathbf{0},\mathbf{I}_n\otimes\mathbf{\Sigma})$，则根据矩阵的拉直运算和Kronecker积性质可得

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

这说明$\mathrm{Vec}(\mathbf{Y}')=\mathrm{Vec}(\mathbf{X}'\mathbf{U}')$的分布服从$N_{pn}(\mathbf{0},\mathbf{I}_n\otimes\mathbf{\Sigma})$。令$\mathbf{Y}'=(\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_n)$，则$\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_n$ 独立同p维正态分布$N_p(\mathbf{0},\mathbf{\Sigma})$。因此有

$$\mathbf{Q}=\mathbf{X}'\mathbf{A}\mathbf{X}=\mathbf{Y}'\mathbf{\Gamma}\mathbf{Y}=\sum_{i=1}^{m}\boldsymbol{Y}_i\boldsymbol{Y}_i'\sim W_p(m,\mathbf{\Sigma}).$$

证明(续)：对(2)的证明。由(1)的证明，以及$\mathbf{Q}=\mathbf{X}'\mathbf{A}\mathbf{X}$, $\mathbf{Q}_1=\mathbf{X}'\mathbf{B}\mathbf{X}$，容易得到$\mathbf{Q}=\mathbf{Y}'\mathbf{\Gamma}\mathbf{Y}$, $\mathbf{Q}_1=\mathbf{Y}'\mathbf{U}\mathbf{B}\mathbf{U}'\mathbf{Y}$。因此

$$\mathbf{Q}_2=\mathbf{Q}-\mathbf{Q}_1=\mathbf{Y}'\mathbf{\Gamma}\mathbf{Y}-\mathbf{Y}'\mathbf{U}\mathbf{B}\mathbf{U}'\mathbf{Y}=\mathbf{Y}'(\mathbf{\Gamma}-\mathbf{U}\mathbf{B}\mathbf{U}')\mathbf{Y}\geq 0,$$

则$\mathbf{\Gamma}-\mathbf{U}\mathbf{B}\mathbf{U}'\geq 0$。

进一步，由$\mathbf{\Gamma}$的定义可知，

$$\mathbf{U}\mathbf{B}\mathbf{U}'=\begin{array}{cc}\begin{pmatrix}\mathbf{E} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} & \begin{array}{c} m \\ n-m\end{array}\end{array},$$

其中$\mathbf{E}$是对角线元素不大于1的m阶矩阵。由于$\mathbf{B}$是n阶幂等矩阵，$\mathbf{U}$是正交矩阵，故$\mathbf{U}\mathbf{B}\mathbf{U}'$也是$n$阶幂等矩阵，进而$\mathbf{E}$是$m$阶幂等矩阵，并且可以计算$\mathbf{B}$的秩为

$$r=\mathrm{tr}(\mathbf{B})=\mathrm{tr}(\mathbf{U}\mathbf{B}\mathbf{U}')=\mathrm{tr}(\mathbf{E}).$$

Wishart分布的性质

证明(续)： 存在m阶正交矩阵$\boldsymbol{\Xi}$，使得$\mathbf{E}=\boldsymbol{\Xi}'\boldsymbol{\Phi}\boldsymbol{\Xi}$，其中$\boldsymbol{\Phi}=\mathrm{diag}(\mathbf{1}_r,\mathbf{0}_{m-r})$为$m\times m$的对角阵。这里，$\mathbf{1}_r$ 为所有元素为1的r维行向量，$\mathbf{0}_{m-r}$为所有元素为0的$m-r$维行向量。令

$$\boldsymbol{\Psi}=\begin{array}{cc}\begin{pmatrix}\boldsymbol{\Phi} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} & \begin{array}{c} m \\ n-m\end{array}\end{array}.$$

构造n阶正交矩阵

$$\mathbf{T}=\begin{pmatrix}\boldsymbol{\Xi} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{n-m}\end{pmatrix}\mathbf{U},$$

则$\mathbf{B}=\mathbf{T}'\boldsymbol{\Psi}\mathbf{T}$。并由$\boldsymbol{\Gamma}$的定义可知$\mathbf{A}=\mathbf{U}'\boldsymbol{\Gamma}\mathbf{U}=\mathbf{T}'\boldsymbol{\Gamma}\mathbf{T}$。

证明(续)：令$\mathbf{Z}' = \mathbf{X}'\mathbf{T}' = (\mathbf{Z}_1, \cdots, \mathbf{Z}_n)$，则$\text{Vec}(\mathbf{Z}')$服从$N_{pn}(\mathbf{0}, \mathbf{I}_n \otimes \mathbf{\Sigma})$，并且$\mathbf{Z}_1, \cdots, \mathbf{Z}_n$独立同服从$p$维正态分布$N_p(\mathbf{0}, \mathbf{\Sigma})$。

因此

$$\mathbf{Q} = \mathbf{X}'\mathbf{A}\mathbf{X} = \mathbf{Z}'\mathbf{\Gamma}\mathbf{Z} = \sum_{i=1}^{m} \mathbf{Z}_i\mathbf{Z}_i', \qquad \mathbf{Q}_1 = \mathbf{X}'\mathbf{B}\mathbf{X} = \mathbf{Z}'\mathbf{\Psi}\mathbf{Z} = \sum_{i=1}^{r} \mathbf{Z}_i\mathbf{Z}_i'.$$

因此有

$$\mathbf{Q}_2 = \mathbf{Q} - \mathbf{Q}_1 = \sum_{i=r+1}^{m} \mathbf{Z}_i\mathbf{Z}_i' \sim W_p(m-r, \mathbf{\Sigma}),$$

并且$\mathbf{Q}_1$与$\mathbf{Q}_2$相互独立。

证明(续)：现在证明(3)。

首先计算$\mathbf{AX}$与$\mathbf{P'X}$的协方差矩阵

$$\begin{aligned}\mathrm{Cov}(\mathbf{AX},\mathbf{P'X}) &= \mathrm{Cov}\Big((\mathbf{I}_p\otimes\mathbf{A})\mathrm{Vec}(\mathbf{X}),(\mathbf{I}_p\otimes\mathbf{P}')\mathrm{Vec}(\mathbf{X})\Big)\\ &= (\mathbf{I}_p\otimes\mathbf{A})\mathrm{Cov}(\mathrm{Vec}(\mathbf{X}))(\mathbf{I}_p\otimes\mathbf{P}')'\\ &= (\mathbf{I}_p\otimes\mathbf{A})(\boldsymbol{\Sigma}\otimes\mathbf{I}_n)(\mathbf{I}_p\otimes\mathbf{P}) = \boldsymbol{\Sigma}\otimes(\mathbf{AP}).\end{aligned}$$

由于$\boldsymbol{\Sigma} > 0$，所以$\mathbf{AX}$与$\mathbf{P'X}$相互独立的充分必要条件是$\mathbf{AP} = \mathbf{0}$。此外，由$\mathbf{Q} = \mathbf{X'AX} = (\mathbf{AX})'\mathbf{AX}$，故$\mathbf{Q}$与$\mathbf{P'X}$相互独立的充分必要条件是$\mathbf{AP} = \mathbf{0}$。

性质4.2.6：分解

设$\mathbf{W} \sim W_p(n, \Sigma), \Sigma > 0, n \geq p$。将$\mathbf{W}$与$\Sigma$作如下的剖分：

$$\mathbf{W} = \begin{array}{cc} \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix} & \begin{array}{c} q \\ p-q \end{array} \end{array}, \quad \Sigma = \begin{array}{cc} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} & \begin{array}{c} q \\ p-q \end{array} \end{array},$$

则有下面的结论：

1. $\mathbf{W}_{11} \sim W_q(n, \Sigma_{11})$，$\mathbf{W}_{22} \sim W_{p-q}(n, \Sigma_{22})$。进一步，当$\Sigma_{12} = \mathbf{0}$时，$\mathbf{W}_{11}$与$\mathbf{W}_{22}$相互独立。
2. 令$\mathbf{W}_{11.2} = \mathbf{W}_{11} - \mathbf{W}_{12}\mathbf{W}_{22}^{-1}\mathbf{W}_{21}$，则$\mathbf{W}_{11.2} \sim W_q(n-p+q, \Sigma_{11.2})$，并且$\mathbf{W}_{11.2}$与$\mathbf{W}_{22}$相互独立，其中$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$。
3. 在$\mathbf{W}_{22}$给定的条件下，有

$$\mathbf{W}_{12} \sim N_{q\times(p-q)}(\Sigma_{12}\Sigma_{22}^{-1}\mathbf{W}_{22}, \mathbf{W}_{22} \otimes \Sigma_{11.2}).$$

性质4.2.7：分解

设$\mathbf{W} \sim W_p(n, \Sigma), \Sigma > 0, n \geq p$。将$\mathbf{W}$与$\Sigma$作如下的剖分：

$$\mathbf{W} = \begin{array}{cc} \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix} & \begin{array}{c} q \\ p-q \end{array} \end{array}, \quad \Sigma = \begin{array}{cc} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} & \begin{array}{c} q \\ p-q \end{array} \end{array},$$

则有下面的结论:

1. $\mathbf{W}_{11} \sim W_q(n, \Sigma_{11})$，$\mathbf{W}_{22} \sim W_{p-q}(n, \Sigma_{22})$。进一步，当$\Sigma_{12} = \mathbf{0}$时，$\mathbf{W}_{11}$与$\mathbf{W}_{22}$相互独立。
2. 令$\mathbf{W}_{22.1} = \mathbf{W}_{22} - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\mathbf{W}_{12}$，则$\mathbf{W}_{22.1} \sim W_{p-q}(n-q, \Sigma_{22.1})$，并且$\mathbf{W}_{22.1}$与$\mathbf{W}_{11}$相互独立，其中$\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$。
3. 在$\mathbf{W}_{11}$给定的条件下，有

$$\mathbf{W}_{21} \sim N_{(p-q)\times q}(\Sigma_{21}\Sigma_{11}^{-1}\mathbf{W}_{11}, \mathbf{W}_{11} \otimes \Sigma_{22.1}).$$

性质4.2.8：行列式

设$\mathbf{W} \sim W_p(n, \boldsymbol{\Sigma})$，$\boldsymbol{\Sigma} > 0$，$n \geq p$，则

$$|\mathbf{W}| \stackrel{d}{=} |\boldsymbol{\Sigma}|\gamma_1\gamma_2\cdots\gamma_p,$$

其中$\gamma_1, \gamma_2, \cdots, \gamma_p$相互独立，并且$\gamma_i \sim \chi^2_{n-i+1}$，$i = 1, \cdots, p$。

- 注：性质4.2.8说明，$|\mathbf{W}|/|\boldsymbol{\Sigma}|$与相互独立的$p$个自由度分别为$n, n-1, \cdots, n-p+1$的$\chi^2$分布变量的乘积同分布。

证明：根据性质4.2.6，使用数学归纳法容易证明性质4.2.8。
令$\mathbf{M}=\mathbf{\Sigma}^{-1/2}\mathbf{W}\mathbf{\Sigma}^{-1/2}$，则$\mathbf{M}\sim W_p(n,\mathbf{I}_p)$。将$\mathbf{M}$剖分为

$$\mathbf{M}=\begin{pmatrix} m_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix} \begin{matrix} 1 \\ p-1 \end{matrix}.$$

容易看到$|\mathbf{M}|=m_{11}|\mathbf{M}_{22}-\mathbf{M}_{21}\mathbf{M}_{12}/m_{11}|$。由性质4.2.6知：当$\mathbf{\Sigma}_{12}=0$时，$m_{11}$与$\mathbf{M}_{22}-\mathbf{M}_{21}\mathbf{M}_{12}/m_{11}$相互独立，并$m_{11}\sim\chi^2_n$，$\mathbf{M}_{22}-\mathbf{M}_{21}\mathbf{M}_{12}/m_{11}\sim W_{p-1}(n-1,\mathbf{I}_{p-1})$。由于

$$|\mathbf{M}|=\frac{|\mathbf{W}|}{|\mathbf{\Sigma}|}=m_{11}\left|\mathbf{M}_{22}-\frac{\mathbf{M}_{21}\mathbf{M}_{12}}{m_{11}}\right|,\quad m_{11}\text{与}\mathbf{M}_{22}-\frac{\mathbf{M}_{21}\mathbf{M}_{12}}{m_{11}}\text{相互独立},$$

$$m_{11}\sim\chi^2_n,\quad \mathbf{M}_{22}-\frac{\mathbf{M}_{21}\mathbf{M}_{12}}{m_{11}}\sim W_{p-1}(n-1,\mathbf{I}_{p-1}),$$

所以由数学归纳法容易证明性质4.2.8。

性质4.2.9：逆矩阵期望

设$\mathbf{W} \sim W_p(n, \Sigma)$，$\Sigma > 0$，$n > p+1$，则

$$E(\mathbf{W}^{-1}) = \frac{1}{n-p-1}\Sigma^{-1}.$$

- 由性质4.2.9可知，Σ^{-1}的无偏估计为$(n-p-1)\mathbf{W}^{-1}$。如果$\mathbf{W}$为离差矩阵，进一步，Σ^{-1}的无偏估计为$[(n-p-1)/(n-1)]\mathbf{S}^{-1}$，其中$\mathbf{S}$为样本协方差矩阵。

证明：首先证明当$\mathbf{W} \sim W_p(n, \mathbf{I}_p)$时性质4.2.9成立。因为$\mathbf{W} \sim W_p(n, \mathbf{I}_p)$，则$E(\mathbf{W}^{-1})$一定具有下面的形式：

$$E(\mathbf{W}^{-1}) = d_0\mathbf{I}_p + d_1\mathbf{1}_p\mathbf{1}_p',$$

其中d_0和d_1是两个待定的常数。

因此，为了得到$E(\mathbf{W}^{-1})$，仅需计算常数d_0和d_1。

首先证明$d_1 = 0$。对任意的正交矩阵$\mathbf{U}$，$\mathbf{UWU}' \sim W_p(n, \mathbf{I}_p)$，所以

$$E((\mathbf{UWU}')^{-1}) = d_0\mathbf{I}_p + d_1\mathbf{1}_p\mathbf{1}_p'.$$

证明(续)：由于$E((\mathbf{UWU}')^{-1}) = \mathbf{U}E(\mathbf{W}^{-1})\mathbf{U}'$，因此对于任意的正交矩阵都有

$$\mathbf{U}(d_0\mathbf{I}_p + d_1\mathbf{1}_p\mathbf{1}_p')\mathbf{U}' = d_0\mathbf{I}_p + d_1\mathbf{1}_p\mathbf{1}_p'.$$

由上式容易看到$d_1(\mathbf{U1}_p)(\mathbf{U1}_p)' = d_1\mathbf{1}_p\mathbf{1}_p'$对于任意的正交矩阵$\mathbf{U}$都成立。要想该等式成立，必须要求$d_1 = 0$，则有

$$E(\mathbf{W}^{-1}) = d_0\mathbf{I}_p.$$

下面将证明$d_0 = 1/(n-p-1)$。令$\mathbf{W}^{-1} = (\omega^{ij})$，$i,j = 1,\cdots,p$，由$E(\mathbf{W}^{-1}) = d_0\mathbf{I}_p$知$d_0 = E(\omega^{11})$。根据分块矩阵逆矩阵的计算，若将$\mathbf{W}$剖分为

$$\mathbf{W} = \begin{pmatrix} \omega_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix} \begin{matrix} 1 \\ p-1 \end{matrix},$$

证明(续)：则有$\omega^{11}=(\omega_{11}-\mathbf{W}_{12}\mathbf{W}_{22}^{-1}\mathbf{W}_{21})^{-1}$。由性质4.2.6(2)，当$\mathbf{\Sigma}_{12}=\mathbf{0}$时，$\omega_{11}-\mathbf{W}_{12}\mathbf{W}_{22}^{-1}\mathbf{W}_{21}\sim\chi^2_{n-p+1}$，因此可得

$$d_0=E(\omega^{11})=E\left[(\omega_{11}-\mathbf{W}_{12}\mathbf{W}_{22}^{-1}\mathbf{W}_{21})^{-1}\right]=\frac{1}{n-p-1}.$$

(红颜色的部分留作作业)

则证明了当$\mathbf{W}\sim W_p(n,\mathbf{I}_p)$时性质4.2.9成立，即

$$E(\mathbf{W}^{-1})=\frac{1}{n-p-1}\mathbf{I}_p.$$

当$\mathbf{W}\sim W_p(n,\mathbf{\Sigma})$时，令$\mathbf{M}=\mathbf{\Sigma}^{-1/2}\mathbf{W}\mathbf{\Sigma}^{-1/2}$，则$\mathbf{M}\sim W_p(n,\mathbf{I}_p)$，因此容易计算

$$E(\mathbf{W}^{-1})=\mathbf{\Sigma}^{-1/2}E(\mathbf{M}^{-1})\mathbf{\Sigma}^{-1/2}=\frac{1}{n-p-1}\mathbf{\Sigma}^{-1}.$$

性质4.2.10：逆矩阵

设$\mathbf{W} \sim W_p(n, \mathbf{\Sigma}), \mathbf{\Sigma} > 0, n \geq p$，则对任意非零的$p$维向量$\boldsymbol{a}$，都有

$$\frac{\boldsymbol{a}'\mathbf{\Sigma}^{-1}\boldsymbol{a}}{\boldsymbol{a}'\mathbf{W}^{-1}\boldsymbol{a}} \sim \chi^2_{n-p+1}.$$

特别地，有下面三种特殊情况：

(1) 如果$\mathbf{\Sigma} = \mathbf{I}_p$，则对任意非零的$p$维向量$\boldsymbol{a}$，有

$$\frac{\boldsymbol{a}'\boldsymbol{a}}{\boldsymbol{a}'\mathbf{W}^{-1}\boldsymbol{a}} \sim \chi^2_{n-p+1}.$$

Wishart分布的性质

性质4.2.10(续)：逆矩阵

(2) 如果$\Sigma = \mathbf{I}_p$，对任意非零的p维向量$\boldsymbol{a}$，若$\boldsymbol{a}'\boldsymbol{a} = 1$，则有

$$\frac{1}{\boldsymbol{a}'\mathbf{W}^{-1}\boldsymbol{a}} \sim \chi^2_{n-p+1}.$$

(3) 如果$\Sigma = \mathbf{I}_p$，并令$\mathbf{W}^{-1} = (\omega^{ij}), i, j = 1, \cdots, p$，则

$$\frac{1}{\omega^{11}} \sim \chi^2_{n-p+1},$$

其中ω^{11}是$\mathbf{W}$逆矩阵的第1行第1列的元素。

- 所谓正定矩阵$\mathbf{W}$的Cholesky分解，就是存在一个下三角形矩阵$\mathbf{T}$，使得$\mathbf{W}=\mathbf{T}\mathbf{T}'$。
- Cholesky分解不是唯一存在的，但是如果给下三角矩阵一些约束条件，要求它的对角元素为正，则Cholesky分解就唯一存在了。
- 性质4.2.11回答了，当$\mathbf{W}\sim W_p(n,\mathbf{I}_p)$时，正定矩阵$\mathbf{W}$ 的Cholesky 分解具有下面的Bartlett分解性质。

性质4.2.11：Bartlett分解

设$\mathbf{W}\sim W_p(n,\mathbf{I}_p), n\geq p$。将$\mathbf{W}$作Bartlett分 解$\mathbf{W}=\mathbf{T}\mathbf{T}'$，其中$\mathbf{T}$是对角线元素为正的下三角形矩阵。令$\mathbf{T}=(t_{ij})_{p\times p}$，则$t_{11},t_{21},t_{22},\cdots,t_{p1},\cdots,t_{pp}$相互独立，在$i>j$时，$t_{ij}\sim N(0,1)$，当$i=j$时，$t_{ij}^2\sim\chi^2_{n-p+1}$。

定义4.3: 非中心Wishart分布

设$\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n$相互独立，$\boldsymbol{X}_i \sim N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}), \boldsymbol{\Sigma} > 0, i = 1,\cdots,n$，则称$\mathbf{W} = \sum_{i=1}^{n} \boldsymbol{X}_i \boldsymbol{X}_i'$服从非中心Wishart分布，记为$\mathbf{W} \sim W_p(n, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$，其中$\boldsymbol{\Delta}$为非中心参数，被定义为$\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{-1}\left(\sum_{i=1}^{n} \boldsymbol{\mu}_i \boldsymbol{\mu}_i'\right)$。

非中心Wishart分布

下面分两种特殊情况讨论非中心Wishart分布的密度函数：

- 当$\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_n = \boldsymbol{\mu}$，并且$\boldsymbol{\Sigma} = \mathbf{I}_p$时，非中心Wishart分布的密度函数为：

$$\exp\left(-\frac{n\boldsymbol{\mu}'\boldsymbol{\mu}}{2}\right)\frac{|\mathbf{W}|^{(n-p-1)/2}\exp\{-\mathrm{tr}(\mathbf{W})/2\}}{2^{np/2}\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma((n-i+1)/2)}$$
$$\sum_{k=0}^{\infty}\frac{1}{k!}\frac{\Gamma(n/2)}{\Gamma(n/2+k)}\left(\frac{n\boldsymbol{\mu}'\mathbf{W}\boldsymbol{\mu}}{4}\right)^k.$$

- 当$\boldsymbol{\Sigma} > 0$，$\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$相互独立，并同服从$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，非中心Wishart分布的密度函数为：

$$\exp\left(-\frac{n\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{2}\right)\frac{|\boldsymbol{\Sigma}|^{-n/2}|\mathbf{W}|^{(n-p-1)/2}\exp\{-\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W})/2\}}{2^{np/2}\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma((n-i+1)/2)}$$
$$\times\sum_{k=0}^{\infty}\frac{1}{k!}\frac{\Gamma(n/2)}{\Gamma(n/2+k)}\left(\frac{n\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{W}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{4}\right)^k.$$

- 假设随机变量$X \sim N(0,1)$，$Y \sim \chi^2_n$，并且X和Y相互独立，则

$$t = \frac{X}{\sqrt{Y/n}} \sim t_n,$$

称变量t服从自由度为n的t分布。

- 显然，若$X \sim N(0,\sigma^2)$，$Y \sim \sigma^2\chi^2_n$，并且X和Y相互独立，则变量$t \sim t_n$。
- 注意: $t^2 = nX^2/Y \sim F_{1,n}$。
- 在多元统计中，假设总体$\boldsymbol{X} \sim N_p(\boldsymbol{0},\boldsymbol{\Sigma})$，随机矩阵$\mathbf{W} \sim W_p(n,\boldsymbol{\Sigma})$，下面讨论$T^2 = n\boldsymbol{X}'\mathbf{W}^{-1}\boldsymbol{X}$的分布。

定义4.4: Hotelling T^2分布

设$\boldsymbol{X} \sim N_p(\mathbf{0}, \Sigma)$，随机矩阵$\mathbf{W} \sim W_p(n, \Sigma), \Sigma > 0, n \geq p$，且$\boldsymbol{X}$和$\mathbf{W}$相互独立，记$T^2 = n\boldsymbol{X}'\mathbf{W}^{-1}\boldsymbol{X}$，则称$T^2$的分布为Hotelling T^2分布，或者称为服从自由度为n的Hotelling T^2分布，记为$T^2 \sim T^2(p, n)$。

更一般地，如果$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$时，则称$T^2$的分布为非中心的Hotelling T^2分布，记为$T^2 \sim T^2(p, n, \boldsymbol{\mu})$。

定理5.1.1

设$\boldsymbol{x}_i = (x_{i1}, \cdots, x_{ip})'(i = 1, \cdots, n)$为来自$p$元正态总体$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$的一组随机样本，$\bar{\boldsymbol{x}}$为样本均值向量，$\mathbf{V}$为样本离差阵，则

1. $\bar{\boldsymbol{x}} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$；
2. $\mathbf{V} \sim W_p(n-1, \boldsymbol{\Sigma})$，其中$n > p$，并且$W_p(n-1, \boldsymbol{\Sigma})$是自由度是$n-1$的Wishart分布；
3. $\bar{\boldsymbol{x}}$与$\mathbf{V}$相互独立。
4. $\Pr(\mathbf{V} > 0) = 1$的充要条件是$n > p$。

详细的证明见第5章。

性质4.3.1

设$\boldsymbol{x}_i, i = 1, \cdots, n$，是来自$p$元正态总体$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$的随机样本，$\bar{\boldsymbol{x}}$，$\mathbf{V}$和$\mathbf{S}$分别是正态总体$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$的样本均值向量，样本离差阵和样本协方差阵，则变量

$$
\begin{aligned}
T^2 &= (n-1)[\sqrt{n}(\bar{\boldsymbol{x}} - \boldsymbol{\mu})]'\mathbf{V}^{-1}[\sqrt{n}(\bar{\boldsymbol{x}} - \boldsymbol{\mu})] \\
&= n(n-1)(\bar{\boldsymbol{x}} - \boldsymbol{\mu})'\mathbf{V}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \\
&= n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \sim T^2(p, n-1).
\end{aligned}
$$

证明： 因为$\bar{\boldsymbol{x}} \sim N_p\left(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma}\right)$，则

$$\sqrt{n}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}).$$

此外$\mathbf{V} \sim W_p(n-1, \boldsymbol{\Sigma})$，且$\bar{\boldsymbol{x}}$与$\mathbf{V}$相互独立，由定义4.4可知：

$$T^2 \sim T^2(p, n-1).$$

性质4.3.2

设$\boldsymbol{X} \sim N_p(\mathbf{0}, \Sigma)$，随机矩阵$\mathbf{W} \sim W_p(n, \Sigma)$，$\Sigma > 0$，$n \geq p$，且$\boldsymbol{X}$和$\mathbf{W}$相互独立，则

$$\boldsymbol{X}'\mathbf{W}^{-1}\boldsymbol{X} \stackrel{d}{=} \frac{\chi_p^2}{\chi_{n-p+1}^2}, \tag{1}$$

其中分子分母的两个χ^2分布相互独立。

证明： 由于$\boldsymbol{X}$和$\mathbf{W}$相互独立，在$\boldsymbol{X}$给定的条件下，$\mathbf{W}$的条件分布仍为$W_p(n,\Sigma)$，则由性质4.2.10可知

$$Y=\frac{\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}}{\boldsymbol{X}'\mathbf{W}^{-1}\boldsymbol{X}}\text{的条件分布为}\chi^2_{n-p+1}.$$

由于上面的条件分布与给定的条件$\boldsymbol{X}$没有关系，所以Y与$\boldsymbol{X}$相互独立，因此Y的无条件分布仍为χ^2_{n-p+1}。

由于$\boldsymbol{X}\sim N_p(\mathbf{0},\Sigma)$，由多元正态分布的二次型性质可知：

$$\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}\sim\chi^2_p.$$

由上面的讨论和下面式子可以证得：$\boldsymbol{X}'\mathbf{W}^{-1}\boldsymbol{X}=\dfrac{\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}}{\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}/\boldsymbol{X}'\mathbf{W}^{-1}\boldsymbol{X}}\overset{d}{=}\dfrac{\chi^2_p}{\chi^2_{n-p+1}}.$

性质4.3.3

T^2与F分布的关系：设$T^2 \sim T^2(p,n)$，则

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

注：

- t^2分布是性质4.3.3当$p=1$时的一种特殊情况，即当$p=1$时，则退化成$t^2 = nX^2/Y \sim F_{1,n}$。
- 性质4.3.3把Hotelling T^2的分布转化为了F分布。

证明： 由性质4.3.2有

$$\begin{aligned}
\frac{n-p+1}{np}T^2(p,n) &= \frac{n-p+1}{p}\boldsymbol{X}'\mathbf{W}^{-1}\boldsymbol{X} \\
&\overset{d}{=} \frac{\chi^2_p/p}{\chi^2_{n-p+1}/(n-p+1)} \\
&\sim F_{p,n-p+1}.
\end{aligned}$$

性质4.3.4

设$\boldsymbol{x}_i$，$i=1,\cdots,n$，是来自p元正态总体$N_p(\boldsymbol{\mu},\Sigma)$的随机样本，$\overline{\boldsymbol{x}}$，$\mathbf{V}$和$\mathbf{S}$分别是正态总体$N_p(\boldsymbol{\mu},\Sigma)$的样本均值向量，样本离差阵和样本协方差阵。记

$$T^2=n(n-1)\overline{\boldsymbol{x}}'\mathbf{V}^{-1}\overline{\boldsymbol{x}}=n\overline{\boldsymbol{x}}'\mathbf{S}^{-1}\overline{\boldsymbol{x}}.$$

则

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

其中$\delta=n\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}$。

性质4.3.5

T^2统计量对非退化变换保持不变。

- 假设随机变量$X \sim \chi_n^2$，$Y \sim \chi_m^2$，并且相互独立，则
$$F = \frac{X/n}{Y/m} \sim F_{n,m},$$
称变量F服从自由度为n和m的F分布。
- F分布和β分布可以相互转化。令
$$B = \frac{X}{X+Y} = \frac{F \cdot \frac{n}{m}}{1 + F \cdot \frac{n}{m}},$$
则$B \sim \beta(n/2, m/2)$。
- $\beta(a,b)$分布的密度函数为：
$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \qquad 0 \le x \le 1.$$

- F分布在一元统计中用于两个正态总体的方差齐性检验，即检验$H_0: \sigma_x^2 = \sigma_y^2$。
- 对p元正态总体$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，其协方差阵$\boldsymbol{\Sigma}$常用$\frac{1}{n-1}\mathbf{V}$或$\frac{1}{n}\mathbf{V}$来估计，前者为无偏估计，后者为极大似然估计，$\mathbf{V}$为样本离差阵。
- 尽管$\boldsymbol{\Sigma}$刻画了总体中数据间的离散程度，但它不是一个数值，所以不直观。
- 那么，如何用一个数值来描述总体中数据间的离散程度？

- F分布在一元统计中用于两个正态总体的方差齐性检验，即检验$H_0:\sigma_x^2=\sigma_y^2$。
- 对p元正态总体$N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$，其协方差阵$\boldsymbol{\Sigma}$常用$\frac{1}{n-1}\mathbf{V}$或$\frac{1}{n}\mathbf{V}$来估计，前者为无偏估计，后者为极大似然估计，$\mathbf{V}$为样本离差阵。
- 尽管$\boldsymbol{\Sigma}$刻画了总体中数据间的离散程度，但它不是一个数值，所以不直观。
- 那么，如何用一个数值来描述总体中数据间的离散程度？
- 最常用的就是$\boldsymbol{\Sigma}$的行列式$|\boldsymbol{\Sigma}|$和迹$\mathrm{tr}(\boldsymbol{\Sigma})$。

定义4.5: 广义方差

设随机向量$\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，则称$|\boldsymbol{\Sigma}|$为$\boldsymbol{X}$的广义方差。若$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$为$p$元正态总体$\boldsymbol{X}$的随机样本，$\mathbf{V}$为样本离差阵，则称$|\mathbf{V}/n|$或$|\mathbf{V}/(n-1)|$为样本广义方差。

- 由定义4.5中广义方差的概念，多元统计中两正态总体$N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$与$N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$协方差阵的齐性检验$H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$就变成了两总体广义方差之比的检验。

Wilks分布

- 两总体广义方差之比常用两个样本广义方差之比(Wilks统计量)来刻画。

定义4.6: Wilks分布

设$\mathbf{W}_1 \sim W_p(n, \Sigma)$，$\mathbf{W}_2 \sim W_p(m, \Sigma)$，$\Sigma > 0$，$n \geq p$，且$\mathbf{W}_1$与$\mathbf{W}_2$相互独立，则称

$$\Lambda = \frac{|\mathbf{W}_1|}{|\mathbf{W}_1 + \mathbf{W}_2|}$$

为Wilks统计量或Λ统计量，其分布为Wilks分布，记成$\Lambda \sim \Lambda(p, n, m)$。

注：当$p=1$时，Wilks分布正是一元统计中参数为$n/2, m/2$的β分布，记为$\beta(n/2, m/2)$。

性质4.4.1

$\Lambda(p,n,m) \overset{d}{=} B_1 B_2 \cdots B_p$，其中$B_i \sim \beta((n-i+1)/2, m/2), i=1,\cdots,p$，且$B_1, B_2, \cdots, B_p$相互独立。

- 注：性质4.4.1是Wilks分布最基本的一个性质，它表明$\Lambda(p,n,m)$和相互独立的p个参数分别为

$$(n/2, m/2), ((n-1)/2, m/2), \cdots, ((n-p+1)/2, m/2)$$

的β分布变量的乘积同分布。

性质4.4.2

$\Lambda(p,n,m) \stackrel{d}{=} \Lambda(m,n+m-p,p)$。

注：在$m<p$时，性质4.4.2表明：$\Lambda(p,n,m)$的分布计算可以转化为$\Lambda(m,n+m-p,p)$来计算。

性质4.4.3

(1) $\Lambda(2r,n,m) \stackrel{d}{=} B_1^2 B_2^2 \cdots B_r^2$，其中$B_i \sim \beta(n+1-2i,m)$，$i=1,\cdots,r$，且$B_1,B_2,\cdots,B_r$相互独立。

(2) $\Lambda(2r+1,n,m) \stackrel{d}{=} B_1^2 B_2^2 \cdots B_r^2 B_{r+1}$，其中$B_i \sim \beta(n+1-2i,m)$，$i=1,\cdots,r$；$B_{r+1} \sim \beta((n-2r)/2,m/2)$，且$B_1,B_2,\cdots,B_r,B_{r+1}$相互独立。

由性质4.4.1–4.4.3，在$p=1,2$或$m=1,2$时，为了计算简单，Wilks分布可转化为F分布。

(1) 当$p=1$时，由性质4.4.1知：$\Lambda(1,n,m)\overset{d}{=}\beta(n/2,m/2)$，所以

$$\frac{n}{m}\frac{1-\Lambda(1,n,m)}{\Lambda(1,n,m)}\overset{d}{=}F_{m,n}.$$

证明：由前面讨论可知：

$$B=\frac{F\cdot\frac{n}{m}}{1+F\cdot\frac{n}{m}}\overset{d}{=}\Lambda(1,n,m),$$

其中$F\sim F_{n,m}$。这时对上式求解可得结果。

(2) 当$m=1$时，由性质4.4.2知：$\Lambda(p,n,1)\stackrel{d}{=}\Lambda(1,n+1-p,p)\stackrel{d}{=}\beta((n+1-p)/2,p/2)$，所以

$$\frac{n+1-p}{p}\frac{1-\Lambda(p,n,1)}{\Lambda(p,n,1)}\stackrel{d}{=}F_{p,n+1-p}.$$

(3) 当$p=2$时，由性质4.4.3知：$\sqrt{\Lambda(2,n,m)}\stackrel{d}{=}\beta(n-1,m)$，所以

$$\frac{n-1}{m}\frac{1-\sqrt{\Lambda(2,n,m)}}{\sqrt{\Lambda(2,n,m)}}\stackrel{d}{=}F_{2m,2(n-1)}.$$

(4) 当$m=2$时，由性质4.4.2知：$\Lambda(p,n,2)\stackrel{d}{=}\Lambda(2,n+2-p,p)$，再由性质4.4.3知：$\sqrt{\Lambda(p,n,2)}\stackrel{d}{=}\beta(n+1-p,p)$，所以

$$\frac{n+1-p}{p}\frac{1-\sqrt{\Lambda(p,n,2)}}{\sqrt{\Lambda(p,n,2)}}\stackrel{d}{=}F_{2p,2(n+1-p)}.$$

Table 1: Λ与F统计量的关系

p	n	m	Λ与F统计量的关系
任意	任意	1	$\frac{n+1-p}{p}\frac{1-\Lambda(p,n,1)}{\Lambda(p,n,1)} \stackrel{d}{=} F_{p,n+1-p}$
任意	任意	2	$\frac{n+1-p}{p}\frac{1-\sqrt{\Lambda(p,n,2)}}{\sqrt{\Lambda(p,n,2)}} \stackrel{d}{=} F_{2p,2(n+1-p)}$
1	任意	任意	$\frac{n}{m}\frac{1-\Lambda(1,n,m)}{\Lambda(1,n,m)} \stackrel{d}{=} F_{m,n}$
2	任意	任意	$\frac{n-1}{m}\frac{1-\sqrt{\Lambda(2,n,m)}}{\sqrt{\Lambda(2,n,m)}} \stackrel{d}{=} F_{2m,2(n-1)}$

以上几个关系式说明对一些特殊的Λ统计量可以化为F统计量，而当$p \geq 3, m \geq 3$时，可用χ^2统计量或F统计量来近似表示。

- 当$p \geq 3$或$m \geq 3$时，Wilks分布的分布函数的精确计算很困难。Box (1949)提出可用χ^2统计量进行近似。

性质4.4.4

当$p \geq 3$或$m \geq 3$时，设$\Lambda \sim \Lambda(p, n, m)$，则当$n \to \infty$时，有

$$-r \ln \Lambda \sim \chi^2_{pm},$$

其中$r = n - \frac{1}{2}(p - m + 1)$。

- 注：详细的推导，可参考王静龙教授的《多元统计分析》教材中3.6节：Wilks分布的渐近展开。


多元统计分析
统计与数据科学丛书 4
STATISTICS AND DATA SCIENCE SERIES 4
多元统计分析
李高荣 吴密霞 编著
Multivariate Statistical Analysis
科学出版社

谢谢，请多提宝贵意见！