多元统计分析

主要内容

第一章 多元分布

基本概念:

随机向量、联合分布函数、联合概率密度函数、多元正态分布、边缘分布、边缘概率密度函数、条件密度、期望、协方差、特征函数、相关系数、偏相关系数、精度矩阵、矩阵正态分布

- 设 $X_1, X_2, \cdots, X_p$ 为*p*个随机变量，它们组成的向量 $X=(X_1, X_2, \cdots, X_p)'$ 称为**随机向量**。

- 随机向量的**联合分布函数** F 定义为

$$\begin{aligned} F(x_1, x_2, \cdots, x_p) &= P\{X_1 \leq x_1, X_2 \leq X_2, \cdots, X_p \leq x_p\} \\ &= P\{X \leq x\}, \end{aligned}$$

其中 $x=(x_1, x_2, \cdots, x_p)'$。

- **联合概率密度函数**：如果存在非负函数 $f(x_1, x_2, \cdots, x_p)$，使得对任意 $x_1, x_2, \cdots, x_p$ 有

$$F(x_1, x_2, \cdots, x_p)=\int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_p} f(t_1, t_2, \cdots, t_p)\, dt_1 dt_2 \cdots dt_p,$$

则称 $f(x_1, x_2, \cdots, x_p)$为 X 的联合概率密度函数。

- X的q个$(q<p)$分量$X^{(1)}=(X_1,X_2,\cdots,X_q)'$的分布称为**边缘分布**，即

$$\begin{aligned} P\{X^{(1)}\le u\} &= P\{X_1\le u_1,\ldots,X_q\le u_q\} \\ &= F(u_1,\cdots,u_q,+\infty,\cdots,+\infty). \end{aligned}$$

- $X^{(1)}$ 的**边缘概率密度函数**定义为：$g(u)=\int_{R^{p-q}} f(u,v)\,dv.$

- 若 $X=(X^{(1)\prime},X^{(2)\prime})'$ 有概率密度函数$f(x)=f(x^{(1)},x^{(2)})$，$X^{(1)}$有密度函数$g(u)$，则$X^{(2)}$在给定$X^{(1)}=x^{(1)}$的**条件密度**为

$$f(x^{(2)}|X^{(1)}=x^{(1)})=\frac{f(x^{(1)},x^{(2)})}{g(x^{(1)})}.$$

- $X_1,X_2,\cdots,X_p$ **相互独立**，当且仅当

$$F(x_1,\ldots,x_p)=\prod_{i=1}^{p}F_i(x_i),\forall(x_1,\ldots,x_p)'\in R^p,$$

其中 F_i 是 X_i的分布函数，$1\le i\le p$ 。

- 多元随机变量（随机向量）矩的性质：
- 期望 $E(X) = (E(X_1), E(X_2), \cdots, E(X_p))'$
- 协方差 $$\begin{aligned} Cov(X) &= \big(E[(X_i - E(X_i))(X_j - E(X_j))]\big)_{p\times p} \\ &= (E[(X - E(X))(X - E(X))']). \end{aligned}$$
- 若 $X_{p\times 1}, Y_{q\times 1}$ 为随机向量，则它们的协方差为

$$\begin{aligned} Cov(X, Y) &= \big(E[(X_i - E(X_i))(Y_j - E(X_j))]\big)_{p\times q} \\ &= (E[(X - E(X))(Y - E(Y))']). \end{aligned}$$

- 其它一些重要的运算

$$
\begin{aligned}
&E(tr(AXB)) = tr(A(E(X))B),\\
&Cov(AX) = ACov(X)A';\\
&E(X'AX) = (E(X))'A(E(X)) + tr(ACov(X));\\
&Cov(AX, BY) = ACov(X, Y)B'.
\end{aligned}
$$

- 多元特征函数

  随机向量 $X = (X_1, \cdots, X_p)'$ 的特征函数为：

$$\phi(t) = \phi(t_1, \cdots, t_p) = E[e^{\mathrm{i}(t_1X_1+\cdots+t_pX_p)}] = E[e^{\{\mathrm{i}t'X\}}],$$

  其中，$t = (t_1, \cdots, t_p)' \in R^p$， i 是虚数单位，$\mathrm{i}^2 = -1$.

- **特征函数的一些性质：**

性质1：对正整数 $k_1, \cdots, k_p$, 如果 $E(X_1^{k_1} \cdots X_p^{k_p})$ 存在，则

$$E(X_1^{k_1} \cdots X_p^{k_p}) = (-\mathrm{i})^{k_1+\cdots+k_p} \left[\frac{\partial^{k_1+\cdots+k_p}\phi(t_1, \cdots, t_p)}{\partial t_1^{k_1} \cdots \partial t_p^{k_p}}\right]_{t_1=\cdots=t_p=0}.$$

特别地，若期望 $E(X_j)$ 存在，则

$$E(X_j) = (-\mathrm{i}) \left[\frac{\partial\phi(t_1, \cdots, t_p)}{\partial t_j}\right]_{t_1=\cdots=t_p=0};$$

若二阶矩 $E(X_j^2)$存在，则

$$E(X_j^2) = -\left[\frac{\partial^2 \phi(t_1, \cdots, t_p)}{\partial t_j^2}\right]_{t_1=\cdots=t_p=0};$$

若二阶混合矩 $E(X_j X_k)$ 存在，则

$$E(X_j X_k) = -\left[\frac{\partial^2 \phi(t_1, \cdots, t_p)}{\partial t_j \partial t_k}\right]_{t_1=\cdots=t_p=0}.$$

性质2：对 $0 < k < p$，分量 $X^{(1)} = (X_1, \cdots, X_k)'$ 的特征函数为 $\phi(t_1, \cdots, t_k, 0, \cdots, 0)$.

性质3：记 $X_1, \cdots, X_p$ 的边缘特征函数分别为 $\phi_1(t_1), \cdots, \phi_p(t_p)$, 记 $X = (X_1, \cdots, X_p)'$ 的特征函数为 $\phi(t_1, \cdots, t_p)$, 则 $X_1, \cdots, X_p$ 相互独立的充分必要条件是：

$$\phi(t_1, \cdots, t_p) = \phi_1(t_1) \cdots \phi_p(t_p).$$

性质4：设 p 维随机向量 $Y_1, \cdots, Y_m$ 的特征函数分别为 $\phi^{(1)}(t), \cdots, \phi^{(m)}(t)$, 如果 $Y_1, \cdots, Y_m$ 相互独立，则随机向量和 $Y_1 + \cdots + Y_m$ 的特征函数为

$$\phi(t) = \phi^{(1)}(t) \cdots \phi^{(p)}(t).$$

- **分块矩阵的运算**

假设矩阵$\mathbf{A}$可以剖分为

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix},$$

记$\mathbf{A}_{2|1} = \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$,则有

$$|\mathbf{A}| = |\mathbf{A}_{11}||\mathbf{A}_{2|1}|,$$

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I} \end{pmatrix}.$$

如果记

$$\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix},$$

则有

$$\begin{aligned}
\mathbf{B}_{11} &= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}, \\
\mathbf{B}_{22} &= (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}, \\
\mathbf{B}_{12} &= -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}, \\
\mathbf{B}_{21} &= -(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1}.
\end{aligned}$$

矩阵拉直和Kronecker积

矩阵拉直： 记$X=(x_1,\cdots,x_p)$是$n\times p$的矩阵。矩阵拉直运算就是将矩阵按列拉直为向量，拉直后的向量记为$vec(X)$, 有

$$vec(X)=\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix},$$

即$vec(X)$是一个$(np)\times 1$的向量。

Kronecker积： 令$A=(a_{ij})_{n\times p}$和B分别是$n\times p$和$m\times q$的矩阵。矩阵A和B的Kronecker积记为$A\otimes B$, 有

$$A\otimes B=(a_{ij}B),$$

所以$A\otimes B$是$(nm)\times(pq)$的矩阵。

拉直运算和Kronecker积的性质

性质1: 对任意实数λ, 有$(\lambda A)\otimes B = A\otimes(\lambda B) = \lambda(A\otimes B)$.

性质2: $A\otimes(B+C) = A\otimes B + A\otimes C$, $(B+C)\otimes A = B\otimes A + C\otimes A$.

性质3: $(A\otimes B)\otimes C = A\otimes(B\otimes C)$.

性质4: $(A\otimes B)' = A'\otimes B'$.

性质5: $(A\otimes B)(C\otimes D) = (AC)\otimes(BD)$.

性质6: 若A和B都是非奇异的方阵，则$(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$.

性质7: $tr(A\otimes B) = tr(A)\cdot tr(B)$, $tr(C'D) = (vec(C))'(vec(D))$.

性质8: 若A和B分别是n和n阶方阵，则$|A\otimes B| = |A|^m\cdot|B|^n$.

性质9: 若A, Y和B分别是$n\times p, p\times q$和$q\times m$的矩阵，则

$$\mathbf{vec(AYB)} = (\mathbf{B}'\otimes\mathbf{A})\mathbf{vec(Y)}.$$

1.2 多元正态分布

1. 2. 1 多元正态分布密度

定义2: 称 p 元随机向量 X 服从参数为 μ 和 Σ 的多元正态分布，

如果其概率密度函数为

$$p(x) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\},$$

其中$\mu \in R^p$, Σ 为 p 阶正定矩阵. 记为 $X \overset{d}{\sim} N_p(\mu, \Sigma)$.

- **定理**1．设p元随机向量 $X=\mu+AY$，其中$\mu\in R^k, A$为 $k\times p$ 的行满秩矩阵, $k\leq p$，随机向量 $Y\overset{d}{\sim}N_p(0,I_p)$，则

$$X\overset{d}{\sim}N_k(\mu,\Sigma),$$

其中$\Sigma=AA'>0$.

1.2.2 多元正态分布的性质

性质1： 密度函数

$$p(x) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\}.$$

性质2：（特征函数）$X \overset{d}{\sim} N(\mu, \Sigma)$，则

$$E(\exp\{\mathrm{i}t'X\}) = \exp\left\{\mathrm{i}t'\mu - \frac{t'\Sigma t}{2}\right\}.$$

性质3： 若 $X \overset{d}{\sim} N_p(\mu, \Sigma)$, 则 $E(X) = \mu$, $Cov(X) = \Sigma$.

性质4： （线性变换）若 $X \overset{d}{\sim} N_p(\mu, \Sigma)$, $Y = \eta + AX$, $\eta \in R^k$, A 是 $k \times p$ 的矩阵, 则

$$Y \overset{d}{\sim} N_k(\eta + A\mu, A\Sigma A').$$

性质5：设 $X_1, \cdots, X_k$ 相互独立, $X_i \overset{d}{\sim} N_p(\mu_i, \Sigma_i)$, $1 \leq i \leq k$, 则

$$\sum_{i=1}^{k} a_i X_i \overset{d}{\sim} N_p\Big(\sum_{i=1}^{k} a_i \mu_i, \sum_{i=1}^{k} a_i^2 \Sigma_i\Big).$$

性质6： 若 $X \overset{d}{\sim} N_p(\mu, \Sigma)$, $\Sigma > 0$, 则

$$(X - \mu)' \Sigma^{-1} (X - \mu) \overset{d}{\sim} \chi_p^2,$$

其中 χ_p^2 是自由度为 p 的卡方分布.

性质7：若 $X \overset{d}{\sim} N_p(\mu, \Sigma)$, 有分解

$$X = \begin{pmatrix} X_1^{(q)} \\ X_2^{(p-q)} \end{pmatrix}, \ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

则 $X_1^{(q)} \overset{d}{\sim} N_q(\mu_1, \Sigma_{11})$, $X_2^{(p-q)} \overset{d}{\sim} N_{p-q}(\mu_2, \Sigma_{22})$.

性质8.（分量独立性） 设 $X \overset{d}{\sim} N_p(\mu, \Sigma)$, 有分解

$$X = \begin{pmatrix} X_1^{(q_1)} \\ \vdots \\ X_k^{(q_k)} \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & \Sigma_{1k} \\ \vdots & \ddots & \vdots \\ \Sigma_{k1} & \cdots & \Sigma_{kk} \end{pmatrix},$$

则 $X_i^{(q_i)}, X_j^{(q_j)}$ $(1 \le i < j \le k)$ 相互独立的充分必要条件是

$$Cov(X_i^{(q_i)}, X_j^{(q_j)}) = \Sigma_{ij} = 0.$$

性质9.（条件分布） 同上假设,

则 $(X_1|X_2=x_2) \overset{d}{\sim} N_q(\mu_{1|2}, \Sigma_{1|2})$, 其中

$$\mu_{1|2} = E(X_1|X_2=x_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2),$$
$$\Sigma_{1|2} = Cov(X_1|X_2=x_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

性质10.（变量的独立分解） 同上假设，令

$$Y_1 = X_1,\ Y_2 = X_2 - \Sigma_{21}\Sigma_{11}^{-1}X_1;$$
$$Z_2 = X_2,\ Z_1 = X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2,$$

则 Y_1 与 Y_2 相互独立, Z_1 与 Z_2 相互独立, 且

$$Y_1 \overset{d}{\sim} N_q(\mu_1, \Sigma_{11}),\ Y_2 \overset{d}{\sim} N_{p-q}(\mu_2 - \Sigma_{21}\Sigma_{11}^{-1}\mu_1, \Sigma_{2|1});$$
$$Z_2 \overset{d}{\sim} N_{p-q}(\mu_2, \Sigma_{22}),\ Z_1 \overset{d}{\sim} N_q(\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2, \Sigma_{1|2});$$
$$\Sigma_{2|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.$$

1.3 相关系数

1.3.1 相关系数

设随机向量 $X = (X_1, \cdots, X_p)'$, $Cov(X) = \Sigma = (\sigma_{ij})_{p\times p}$,

则 X_1 与 X_2 $(1 \le i < j \le p)$ 的相关系数 ρ_{ij} 为

$$\rho_{ij} = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)}\sqrt{Var(X_j)}} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}.$$

1.3.2 相关矩阵

$$R = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1p} \\ \vdots & & \vdots \\ \rho_{p1} & \cdots & \rho_{pp} \end{pmatrix} = diag(\sigma_{11}^{-1/2}, \cdots, \sigma_{pp}^{-1/2}) \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix} diag(\sigma_{11}^{-1/2}, \cdots, \sigma_{pp}^{-1/2}).$$

1.3.3 偏相关系数

设 $X \overset{d}{\sim} N_p(\mu, \Sigma), \Sigma > 0$, 有分解

$$X = \begin{pmatrix} X_1^{(q)} \\ X_2^{(p-q)} \end{pmatrix}, \ \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

则 $(X_1^{(q)} | X_2^{(p-q)}) \overset{d}{\sim} N_q(\mu_{1|2}, \Sigma_{1|2})$, 其中

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2),$$

$$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \overset{\text{记成}}{=} (\sigma_{(ij)|(1|2)})_{q\times q}.$$

在给定 $X_2^{(p-q)}$ 的条件下, X_i 与 X_j $(1 \le i < j \le q)$ 的条件相关系数为

$$\rho_{(ij)|(1|2)} = \frac{\sigma_{(ij)|(1|2)}}{\sqrt{\sigma_{(ii)|(1|2)}}\sqrt{\sigma_{(jj)|(1|2)}}}.$$

条件相关系数也称为偏相关系数。

1.3.4 精度矩阵

设随机向量 X, 有 $Cov(X)=\Sigma>0$, 那么称 $K=\Sigma^{-1}$ 为 X 的精度矩阵.

性质1: 若 $X=(X_1,\cdots,X_p)'\overset{d}{\sim}N_p(\mu,\Sigma)$, $\Sigma>0$, $K=\Sigma^{-1}=(k_{ij})_{p\times p}$, 则

$$k_{ii}=(Var(X_i|X_{(-i)}))^{-1},\ 1\le i\le p,$$

其中 $X_{(-i)}=(X_1,\cdots,X_{i-1},X_{i+1},\cdots,X_p)',\ 1\le i\le p$.

设 $X\overset{d}{\sim}N_p(\mu,\Sigma),\Sigma>0$, 有如下分解

$$X=\begin{pmatrix} X_1^{(p_1)} \\ X_2^{(p_2)} \\ X_3^{(p_3)} \end{pmatrix},\ \Sigma=\begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} \end{pmatrix},\ K=\begin{pmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{pmatrix}.$$

性质2: 在 X_3 给定的条件下, X_1 与 X_2 相互条件独立的充要条件是 $K_{12}=0$.

1.4 矩阵多元正态分布

设 $X_1, \cdots, X_n$ $i.i.d.$, $X_i = (x_{i1}, \cdots, x_{ip})' \overset{d}{\sim} N_p(\mu, \Sigma)$, 即

$X_1, \cdots, X_n$ 是来自 p 元正态总体 $N_p(\mu, \Sigma)$的独立样本.

记 $X = (X_1, \cdots, X_n)$, 则X是一个$p \times n$的随机矩阵.

随机矩阵的期望：$E(X) = (\mu, \cdots, \mu) = \mu \cdot \mathbf{1}_n'$,

其中 $\mathbf{1}_n = (1, \cdots, 1)'$.

矩阵的拉直运算：$vec(X) = vec((X_1, \cdots, X_n)) = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}_{(np)\times 1}$,

矩阵的拉直运算即是将矩阵依列拉直后形成一个向量。

随机矩阵的协方差阵：$Cov(X)=Cov(vec(X))$.

1.4.1 矩阵分布

随机矩阵的分布：随机矩阵拉直后的随机向量的分布。

矩阵 X 的运算：由于 $X=(X_1,\cdots,X_n)$，有

$$E(vec(X))=\begin{pmatrix}\mu\\ \vdots\\ \mu\end{pmatrix},Cov(vec(X))=\begin{pmatrix}\Sigma & 0 & \cdots & 0\\ 0 & \Sigma & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \Sigma\end{pmatrix},$$

即 $E(vec(X))=\mathbf{1}_n\otimes\mu$，$Cov(vec(X))=\mathbf{I}_n\otimes\Sigma$，其中 $\mathbf{I_n}$ 是 $\mathbf{n}$ 阶单位阵.

1.4.2 随机矩阵拉直运算的性质

性质1：对 $n \times m$ 的随机矩阵Y, 若有

$$E(vec(Y)) = \alpha \otimes \beta, \ Cov(vec(Y)) = A \otimes B,$$

其中, α, β分别是m和 n维列向量，A和B分别是m和n 阶方阵, 则

$$E(vec(Y')) = \beta \otimes \alpha, \ Cov(vec(Y')) = B \otimes A.$$

由性质1，对上述随机矩阵 X 有

$$E(vec(X')) = \mu \otimes \mathbf{1}_n, \ Cov(vec(X')) = \Sigma \otimes \mathbf{I}_n.$$

因此，对由 n个p 维正态总体的独立样本组成的随机矩阵X,

$$vec(X) \overset{d}{\sim} N_{np}(\mathbf{1}_n \otimes \mu, \ \mathbf{I}_n \otimes \Sigma),$$
$$vec(X') \overset{d}{\sim} N_{np}(\mu \otimes \mathbf{1}_n, \ \Sigma \otimes \mathbf{I}_n).$$

1.4.3 矩阵正态分布

若 $vec(X) \overset{d}{\sim} N_{np}(\mathbf{1}_n \otimes \mu,\ \mathbf{I}_n \otimes \Sigma)$ 或 $vec(X') \overset{d}{\sim} N_{np}(\mu \otimes \mathbf{1}_n,\ \Sigma \otimes \mathbf{I}_n)$，则称随机矩阵$X$和$X'$分别服从**矩阵正态分布**，记为

$$X \overset{d}{\sim} N_{p\times n}(\mu \cdot \mathbf{1}_n', \mathbf{I}_n \otimes \Sigma),$$
$$X' \overset{d}{\sim} N_{n\times p}(\mathbf{1}_n \cdot \mu', \Sigma \otimes \mathbf{I}_n).$$

一般地，记 $n \times p$ 的正态随机矩阵为 $X \overset{d}{\sim} N_{n\times p}(B, \Sigma \otimes V)$，其中

$$B = E(X),\ \Sigma \otimes V = Cov(vec(X)),$$

Σ 和 V分别是 p和 n阶方阵。

若 $X \overset{d}{\sim} N_{n\times p}(B, \Sigma \otimes V)$，则 $vec(X) \overset{d}{\sim} N_{np}(vec(B), \Sigma \otimes V)$.

1.4.4 矩阵正态分布的密度函数

若 $X \overset{d}{\sim} N_{p\times n}(B, V\otimes\Sigma)$, Σ和V均为正定的方阵,
则由 $vec(X) \overset{d}{\sim} N_{np}(vec(B), V\otimes\Sigma)$，可以导出矩阵 X 的密度函数如下：

$$\frac{1}{(2\pi)^{(np)/2}\sqrt{|V\otimes\Sigma|}}\exp\left\{-\frac{(vec(X-B))'(V\otimes\Sigma)^{-1}(vec(X-B))}{2}\right\},$$

上式等价于：

$$\frac{1}{(2\pi)^{(np)/2}|V|^{p/2}|\Sigma|^{n/2}}\exp\left\{-\frac{tr[(X-B)'\Sigma^{-1}(X-B)V^{-1}]}{2}\right\}.$$

1.4.5 矩阵正态分布的线性变换

性质2. 设 $p \times n$ 的矩阵 X 服从矩阵正态分布 $X \overset{d}{\sim} N_{p\times n}(B, V \otimes \Sigma)$, 有 $\Sigma_{p\times p} \geq 0, V_{n\times n} \geq 0$. 令

$$Y = C + AX\Gamma,$$

其中 $C_{q\times m}, A_{q\times p}$ 和 $\Gamma_{n\times m}$ 是常数矩阵，则

$$Y \overset{d}{\sim} N_{q\times m}((C + AB\Gamma), (\Gamma' V \Gamma) \otimes (A\Sigma A')).$$

第二章 其它重要多元分布

Wishart

Hotelling T^2

Wilks

2.1 Wishart分布

2. 1. 1 **Wishart 分布的定义**

设$X=(X_1,\cdots,X_n)$, 其中$X_1,\cdots,X_n$ $i.i.d.$, $X_i \overset{d}{\sim} N_p(0,\Sigma)$, $1\le i\le n$.

则称p阶随机矩阵$W=XX'=\sum_{i=1}^{n} X_iX_i'$的分布为$p$阶Wishart分布，记为

$$W \overset{d}{\sim} W_p(n,\Sigma),$$

其中n称为其自由度.

事实上，有$X \overset{d}{\sim} N_{p\times n}(0,\mathbf{I}_n\otimes\Sigma)$, $X' \overset{d}{\sim} N_{n\times p}(0,\Sigma\otimes\mathbf{I}_n)$是矩阵正态分布，则Wishart分布也可以定义为

$$W=XX' \overset{d}{\sim} W_p(n,\Sigma).$$

2.1.2 **Wishart 分布的密度函数**

当 $\Sigma > 0, n \geq p$时, p 阶Wishart分布有密度函数

$$f_p(W) = \frac{|W|^{(n-p-1)/2} \exp\left\{-\frac{1}{2} tr(\Sigma^{-1}W)\right\}}{2^{(np/2)}|\Sigma|^{n/2}\pi^{(p(p-1)/4)} \prod_{i=1}^{p} \Gamma(\frac{(n-i+1)}{2})}, \ W > 0.$$

当 $p = 1$时, $f_1(w) = 2^{-n/2}\Gamma^{-1}(n/2)\sigma^{-n}w^{(n-2)/2}\exp\{-w/(2\sigma^2)\}, \ w > 0.$

即 $W = X'X \overset{d}{\sim} \sigma^2\chi^2(n).$

2.1.3 Wishart分布的性质

性质1. 若$W \overset{d}{\sim} W_p(n,\Sigma)$, 则$E(W)=n\Sigma$.

性质2. 若$W \overset{d}{\sim} W_p(n,\Sigma)$, C是$k\times p$阶矩阵, 则$CWC' \overset{d}{\sim} W_k(n, C\Sigma C')$.

性质3. 若$W \overset{d}{\sim} W_p(n,\Sigma)$, 则$W$特征函数为

$$E(e^{\{\mathrm{i}tr(TW)\}}) = |I_p - 2\mathrm{i}\Sigma T|^{-n/2},$$

其中T为p阶实对称阵.

性质4. 若 $W_1,\cdots,W_k$ 相互独立, $W_i \overset{d}{\sim} W_p(n_i,\Sigma)$, $1\le i\le k$, 则

$$\sum_{i=1}^{k} W_i \overset{d}{\sim} W_p(\sum_{i=1}^{k} n_i, \Sigma).$$

矩阵二次型:

若随机矩阵$X \overset{d}{\sim} N_{p\times n}(0, I_n \otimes \Sigma)$, 或$X' \overset{d}{\sim} N_{n\times p}(0, \Sigma \otimes I_n)$, 则称

$$Q = XAX'$$

为矩阵二次型, 其中A是n阶方阵, $A \geq 0$.

若$X = (X_1, \cdots, X_n)$, 其中$X_1, \cdots, X_n$ $i.i.d.$, $X_i \overset{d}{\sim} N_p(0, \Sigma)$, $1 \leq i \leq n$, $A = (a_{ij})_{n\times n}$, 则

$$Q = XAX' = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} X_i X_j'.$$

特别地, 当$A = \mathbf{I}_n$时, $Q = XX' \overset{d}{\sim} W_p(n, \Sigma)$.

性质5. (矩阵二次型)

(1) 若A为幂等矩阵, 则矩阵二次型 $Q = XAX' \overset{d}{\sim} W_p(m, \Sigma)$,

其中, $m = Rank(A) = R(A) = tr(A)$.

(2) 设 $Q = XAX', Q_1 = XBX', A$ 和 B 都是幂等矩阵.

若 $Q_2 = Q - Q_1 \geq 0$, 则 $Q_2 \overset{d}{\sim} W_p(m - r, \Sigma)$,

其中, $m = R(A)$, $r = R(B)$, 且 Q_1 与 Q_2 相互独立.

(3) 设$Q = XAX', A$为幂等矩阵.

则 $P'X'$ 与 Q 独立的充要条件为 $AP = 0$, 其中 P 是 $n \times p$ 的矩阵.

性质6. (独立分解) 设$W \overset{d}{\sim} W_p(n, \Sigma), \Sigma > 0, n \geq p.$

将W和Σ作如下相同的q阶和$(p-q)$阶矩阵分块

$$W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix},\ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

则有:

(1) $W_{22} - W_{21}W_{11}^{-1}W_{12}$与$(W_{11}, W_{21})$相互独立;

(2) $W_{22} - W_{21}W_{11}^{-1}W_{12} \overset{d}{\sim} W_{p-q}((n-q), \Sigma_{2|1})$;

(3) $W_{11} \overset{d}{\sim} W_q(n, \Sigma_{11})$;

(4) 在W_{11}给定的条件下,

$$W_{21}W_{11}^{-1/2} \overset{d}{\sim} N_{(p-q)\times q}(\Sigma_{21}\Sigma_{11}^{-1}W_{11}^{1/2}, I_q \otimes \Sigma_{2|1}).$$

特别地, 当 $\Sigma_{21} = 0$ 时, 有

(1′) $W_{22} - W_{21}W_{11}^{-1}W_{12}$, W_{11} 与 $W_{21}W_{11}^{-1/2}$ 相互独立;

(2′) $W_{22} - W_{21}W_{11}^{-1}W_{12} \overset{d}{\sim} W_{p-q}((n-q), \Sigma_{22})$;

(3′) $W_{11} \overset{d}{\sim} W_q(n, \Sigma_{11})$;

(4′) $W_{21}W_{11}^{-1/2} \overset{d}{\sim} N_{(p-q)\times q}(0, I_q \otimes \Sigma_{22})$.

性质7. (行列式) 设 $W \overset{d}{\sim} W_p(n,\Sigma)$, $\Sigma>0$, $n\geq p$. 则

$$|W| \overset{d}{=} |\Sigma| \prod_{i=1}^{p} \gamma_i,$$

其中,$\gamma_1,\cdots,\gamma_p$ 相互独立, $\gamma_i \overset{d}{\sim} \chi^2(n-i+1)$, $1\leq i\leq p$.

性质8. (逆矩阵期望) 若 $W \overset{d}{\sim} W_p(n,\Sigma)$, $\Sigma>0$, $n>(p+1)$, 则

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

性质9. (逆矩阵的分布) 设 $W \overset{d}{\sim} W_p(n,\Sigma)$, $\Sigma>0$, $n\geq p$,

则对任意零的 p 维向量 a, 都有

$$\frac{a'\Sigma^{-1}a}{a'W^{-1}a} \overset{d}{\sim} \chi^2(n-p+1).$$

性质10. (Bartlett分解) 设$W \overset{d}{\sim} W_p(n, I_p), n \geq p$. 将 W 作分解

$W = TT'$, T是对角元为正的下三角矩阵.

令 $T = (t_{ij})_{p\times p}$, 则 $t_{11}, t_{21}, t_{22}, \cdots, t_{p1}, t_{p2}, \cdots, t_{pp}$ 相互独立, 且

$t_{ii}^2 \overset{d}{\sim} \chi^2(n-p+1)$,

$t_{ij} \overset{d}{\sim} N(0,1)$,

对 $1 \leq j < i \leq p$ 成立.

2.2 Hotelling T^2分布

定义: 设$X \overset{d}{\sim} N_p(0, \Sigma), W \overset{d}{\sim} W_p(n, \Sigma)$, 且$X$和$W$相互独立. 记

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

则称T^2的分布为**Hotelling** T^2分布.

特别地, 当$p = 1$, $\Sigma = 1$, 有

$$t^2 = nX'W^{-1}X = \frac{X^2}{(W/n)} \overset{d}{\sim} F(1, n).$$

2.2.2 **Hotelling** T^2分布的性质

性质1.

$$X'W^{-1}X \overset{d}{=} \frac{\chi^2(p)}{\chi^2(n-p+1)} \overset{d}{\sim} Z\left(\frac{p}{2}, \frac{n-p+1}{2}\right),$$

其中分子分母相互独立.

性质2.

$$\begin{aligned}\frac{n-p+1}{np}T_p^2(n) &\overset{d}{=} \frac{\chi^2(p)/p}{\chi^2(n-p+1)/(n-p+1)} \\ &\overset{d}{\sim} F(p,(n-p+1)).\end{aligned}$$

性质3. **(密度函数)** $T_p^2(n)$的密度函数为

$$p(t) = \frac{\Gamma((n+1)/2)}{\Gamma(p/2)\Gamma((n-p+1)/2)} \cdot \frac{(t/n)^{(p-2)/2}}{(1+t/n)^{(n+1)/2}}.$$

2.3 Wilks分布

2. 3. 1 **Wilks 分布的定义**

定义: 假设$W_1 \overset{d}{\sim} W_p(n,\Sigma), W_2 \overset{d}{\sim} W_p(m,\Sigma)$, $\Sigma > 0, n \geq p$, W_1和W_2相互独立.
记

$$\Lambda = \frac{|W_1|}{|W_1 + W_2|},$$

则称Λ的分布为**Wilks** 分布，记为 $\Lambda_{p,n,m}$.

2.3.2 Wilks 分布的性质

性质1. $\Lambda_{p,n,m} \overset{d}{=} B_1 B_2 \cdots B_p$，其中，$B_1, B_2, \cdots, B_p$ 相互独立,

$$B_i \overset{d}{\sim} B(\frac{n-i+1}{2}, \frac{m}{2}),\ 1 \le i \le p.$$

因此它是 $p = 1$ 时Beta分布的推广, 而不是F分布的直接推广.

性质2. $\Lambda_{p,n,m} \overset{d}{=} \Lambda_{m,(n+m-p),p}.$

性质3. (与F分布的关系)

1) $$\frac{n}{m}\cdot\frac{1-\Lambda_{1,n,m}}{\Lambda_{1,n,m}} \overset{d}{\sim} F(m,n);$$

2) $$\frac{n+1-p}{p}\cdot\frac{1-\Lambda_{p,n,1}}{\Lambda_{p,n,1}} \overset{d}{\sim} F(p,(n+1-p));$$

3) $$\frac{n-1}{m}\cdot\frac{1-\sqrt{\Lambda_{2,n,m}}}{\sqrt{\Lambda_{2,n,m}}} \overset{d}{\sim} F(2m,2(n-1));$$

4) $$\frac{n+1-p}{p}\cdot\frac{1-\sqrt{\Lambda_{p,n,2}}}{\sqrt{\Lambda_{p,n,2}}} \overset{d}{\sim} F(2p,2(n+1-p)).$$

几点总结：

Wishart分布

(1) Wishart分布是正态随机向量特殊二次型的分布;

(2) 样本离差阵是最常见的服从Wishart分布的随机矩阵;

(3) 一维情形下的Wishart分布就是卡方分布;

Hotelling T^2 分布

(1) 是一维情形下t分布平方的推广;

(2) Hotelling T^2 分布常见于检验统计量的分布;

(3) Hotelling T^2 分布的计算要转化为F分布.

Wilks 分布

(1) 是一维情形下Beta分布的推广;

(2) Wilks分布常见于似然比检验统计量的分布;

(3) Wilks分布的计算在很多情形下可以转化为F分布.