第二章 由多元正态分布导出的分布

2.0 由一元正态分布导出的分布

卡方分布：

设$X = (x_1, \cdots, x_n)$, 其中$x_1, \cdots, x_n$ $i.i.d.$, $x_i \overset{d}{\sim} N_1(0,1), 1 \le i \le n.$

则称随机变量$Y = XX' = \sum_{i=1}^{n} x_i^2$为服从自由度为$n$的 **卡方分布**, 记为

$$Y = \sum_{i=1}^{n} x_i^2 \overset{d}{\sim} \chi^2(n).$$

t分布：

假设随机变量X和Y相互独立, 且$X \overset{d}{\sim} N_1(0,1)$, $Y \overset{d}{\sim} \chi^2(n)$,
则称随机变量 $t = X/(\sqrt{Y/n})$为服从自由度为 n 的 **t分布**, 记为

$$t = \frac{X}{\sqrt{Y/n}} \overset{d}{\sim} t(n).$$

F分布

假设随机变量X和Y相互独立, 且$X \overset{d}{\sim} \chi^2(n)$, $Y \overset{d}{\sim} \chi^2(m)$,
则称随机变量 $F = (X/n)/(Y/m)$为服从自由度为 n 和 m 的 **F分布**, 记为

$$F = \frac{X/n}{Y/m} \overset{d}{\sim} F(n, m).$$

应用： 1）构造参数的置信区间； 2）假设检验

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

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

若记$\Gamma_p(x) = \pi^{(p(p-1)/4)}\prod\limits_{i=1}^{p}\Gamma(x - \dfrac{i-1}{2})$, 并称之为$p$维$\Gamma$函数, 则有

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

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

说明1： 记$W=XX'=(w_{ij})_{p\times p}$是对称矩阵, 实际上$W$的分布是随机向量$(w_{11},\cdots,w_{1p},w_{22},\cdots,w_{2p},\cdots,w_{pp})'$的分布.

说明2： $n\geq p$是为了保证$W>0$成立的概率为1.

Wishart分布的推导：

1) 从2阶到高阶；

2) 对2阶，从独立到相关，从特殊到一般.

2.1.3 Wishart分布的性质

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

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

性质3. 若$W \stackrel{d}{\sim} W_p(n,\Sigma)$, 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, $W_i \stackrel{d}{\sim} W_p(n_i,\Sigma)$, $1\le i\le k$, 则

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

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5.(1)的证明:

由A幂等, 知存在正交阵U, 使得$A = UBU'$, 其中

$$B = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}, \; m = R(A).$$

下面考虑矩阵正态分布的正交变换$Y = XU$的分布.

矩阵拉直的性质: 对矩阵$C_{n\times p}, Z_{p\times q}, D_{q\times m}$, 有

$$vec(CZD) = (D' \otimes C)vec(Z).$$

计算

$$E(Y) = E(XU) = E(X)U = 0,$$

$$\begin{aligned} Cov[vec(Y)] &= Cov[vec(XU)] = Cov[vec(I_p XU)] \\ &= Cov[(U' \otimes I_p)vec(X)] \\ &= (U' \otimes I_p)Cov[vec(X)](U \otimes I_p) \\ &= (U' \otimes I_p)(I_n \otimes \Sigma)(U \otimes I_p) \\ &= I_n \otimes \Sigma. \end{aligned}$$

因此有 $Y \overset{d}{\sim} N_{p\times n}(0, I_n \otimes \Sigma)$.

令$Y = (Y_1, \cdots, Y_n)$, 可知 $Y_1, \cdots, Y_n$是独立同分布的p维正态随机向量, 均值为0, 协方差为Σ. 进而有,

$$Q = XAX' = Y{\color{red}B}Y' = \sum_{i=1}^{m} Y_i Y_i' \overset{d}{\sim} W_p(m, \Sigma).$$

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 $W \overset{d}{\sim} W_p(n,\Sigma)$, $\Sigma > 0$, $n \geq p$,

则对任意非零的 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

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

假设 $\Sigma > 0$, 有

$$
\begin{aligned}
T^2 &= nX'W^{-1}X \\
&= n(\Sigma^{-1/2}X)'(\Sigma^{-1/2}W\Sigma^{-1/2}))^{-1}(\Sigma^{-1/2}X),
\end{aligned}
$$

而

$$
\Sigma^{-1/2}X \overset{d}{\sim} N_p(0, I_p),
$$

$$
\Sigma^{-1/2}W\Sigma^{-1/2} \overset{d}{\sim} W_p(n, I_p),
$$

因此**Hotelling** T^2分布与 Σ 无关, 记为 $T_p^2(n)$.

2.2.2 **Hotelling** T^2分布的性质

性质1.

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

其中分子分母相互独立.

性质2.

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

性质3. (密度函数) $T_p^2(n)$的密度函数为

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

2.2.3 非中心Hotelling T^2分布

定义： 设$X \overset{d}{\sim} N_p(\mu, \Sigma), W \overset{d}{\sim} W_p(n, \Sigma)$, 且 X 和 W 相互独立.
则$T^2 = nX'W^{-1}X$的分布为非中心的Hotelling T^2分布,
记为$T_p^2(n, a)$, 其中 $a = \mu'\Sigma^{-1}\mu$ 是非中心参数.

性质3.

1) $X'W^{-1}X \overset{d}{=} \dfrac{\chi^2(p, a)}{\chi^2(n-p+1)}$,

2) $\dfrac{n-p+1}{np} T_p^2(n, a) \overset{d}{=} \dfrac{\chi^2(p, a)/p}{\chi^2(n-p+1)/(n-p+1)} \overset{d}{\sim} F(p, (n-p+1), a).$

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$\Lambda_{p,n,m}$.

由于

$$\Lambda = \frac{|\Sigma^{-1/2}W_1\Sigma^{-1/2}|}{|\Sigma^{-1/2}W_1\Sigma^{-1/2} + \Sigma^{-1/2}W_2\Sigma^{-1/2}|},$$

且 $\Sigma^{-1/2}W_1\Sigma^{-1/2} \overset{d}{\sim} W_p(n, I_p)$, $\Sigma^{-1/2}W_2\Sigma^{-1/2} \overset{d}{\sim} W_p(m, I_p)$,

故**Wilks**分布$\Lambda_{p,n,m}$与Σ无关.

F分布与Beta分布的关系：

设随机变量$F \overset{d}{\sim} F(n,m)$，则

$$\frac{\frac{n}{m}F(n,m)}{1+\frac{n}{m}F(n,m)} \overset{d}{=} B(\frac{n}{2},\frac{m}{2}),$$

其中$B(\frac{n}{2},\frac{m}{2})$是自由度为$(\frac{n}{2},\frac{m}{2})$的$Beta$分布.

同理有

$$\frac{1-B(\frac{m}{2},\frac{n}{2})}{B(\frac{m}{2},\frac{n}{2})}\cdot\frac{m}{n} \overset{d}{=} F(n,m).$$

2.3.2 Wilks 分布的性质

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

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

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

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.

作业： 若$W \overset{d}{\sim} W_p(n, \Sigma), \Sigma > 0$. A是 p 阶常数方阵，试求$E(|AW|)$.

拍照发邮件给助教：

- 樊瑜：
- Email： fanyu16@mails.ucas.ac.cn