



第四章 多元统计基础

§ 4.1 多元分布概念

§ 4.2 多元正态分布

§ 4.3 均值向量与协方差阵的估计

§ 4.4 均值向量的检验





§ 4.1 多元分布概念

§ 4.1.1 随机向量

§ 4.1.2 分布函数与密度函数

§ 4.1.3 多元变量的独立性

§ 4.1.3 随机向量的数字特征





§ 4.1.1 随机向量

讨论多变量总体，数据是同时观测 p 个变量，进行 n 次观测得到的，把这 p 个指标表示为 X_1, X_2, \dots, X_p ，用向量 $X = (X_1, X_2, \dots, X_p)^T$ 表示对同一个体观测的 p 个变量。

观测 n 个个体，数据如下表所示，称每一个个体的 p 个变量为一个样品，而全体 n 个样品形成一个样本。



变量	X_1	X_2	\dots	X_p
1	x_{11}	x_{12}	\dots	x_{1p}
2	x_{21}	x_{21}	\dots	x_{2p}
\dots	\dots	\dots		\dots
n	x_{n1}	x_{n2}	\dots	x_{np}

记 $X_{(a)} = (x_{a1}, x_{a2}, \dots, x_{ap})^T$, $a = 1, 2, \dots, n$

表示第 a 个样品的观测值。

第 j 列的元素 $X_j = (x_{1j}, x_{2j}, \dots, x_{nj})^T$, $j = 1, 2, \dots, p$

表示对第 j 个变量 X_j 的 n 次观测数值。



因此，样本资料矩阵可用矩阵语言表示为

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = (X_1, X_2, \cdots, X_p) = \begin{bmatrix} X_{(1)}^T \\ X_{(2)}^T \\ \vdots \\ X_{(n)}^T \end{bmatrix}$$

定义4.1.1 设 X_1, X_2, \cdots, X_p 为 p 个随机变量，由它们组成的向量 $X = (X_1, X_2, \cdots, X_p)^T$ 称为随机向量。



§ 4.1.2 分布函数与密度函数

定义4.1.2 设随机向量 $X = (X_1, X_2, \dots, X_p)^T$, 多元分布函数为

$$F(x) = F(x_1, x_2, \dots, x_p) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_p \leq x_p)$$

式中, $x = (x_1, x_2, \dots, x_p) \in R^p$, 并记成 $X \sim F$ 。





定义4.1.3 $X \sim F(x) = F(x_1, x_2, \dots, x_p)$, 若存在一个非负的函数 $f(\cdot)$, 使得

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

对一切 $x \in R^p$ 成立, 则称 X (或 $F(x)$) 有分布密度 $f(\cdot)$, 并称 X 为连续型随机向量。

p 维变量的函数 $f(\cdot)$ 能作为 R^p 中某个随机向量的分布密度, 当且仅当 $f(x) \geq 0, \forall x \in R^p$, 且 $\int_{R^p} f(x) dx = 1$ 。





§ 4.1.3 多元变量的独立性

定义4.1.4 两个随机向量 X 和 Y 称为互相独立的, 若
 $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ 对一切 x, y 均成立。

若 $F(x, y)$ 为 (X, Y) 的联合分布函数, $G(x)$ 和 $H(y)$ 分别为 X 和 Y 的分布函数, 则 X 和 Y 独立当且仅当

$$F(x, y) = G(x)H(y)$$





若 (X, Y) 有密度 $f(x, y)$ ，用 $g(x)$ 和 $h(y)$ 分别表示 X 和 Y 的分布密度，则 X 和 Y 独立当且仅当

$$f(x, y) = g(x)h(y)$$

在上述定义中， X 和 Y 的维数一般是不同的。

类似地，若联合分布等于各自分布的乘积，则称 p 个随机向量 X_1, X_2, \dots, X_p 相互独立。可以推知任何 X_i 与 $X_j (i \neq j)$ 独立，但若已知任何 X_i 与 $X_j (i \neq j)$ 独立，并不能推出 X_1, X_2, \dots, X_p 相互独立。





§ 4.1.4 随机向量的数字特征

随机向量 X 的均值

设 $X = (X_1, X_2, \dots, X_p)^T$ 有 p 个分量。

若 $E(X_i) = \mu_i (i = 1, 2, \dots, p)$ 存在，定义随机向量 X 的均值为

$$E(X) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \mu$$





μ 是一个 p 维向量，称为均值向量。

当 A 和 B 为常数矩阵时，由定义可立即推出如下性质

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

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





随机向量 X 的协方差阵

$$\begin{aligned}\Sigma &= cov(X, X) = E(X - EX)(X - EX)^T = D(X) \\ &= \begin{bmatrix} D(X_1) & cov(X_1, X_2) & \cdots & cov(X_1, X_p) \\ cov(X_2, X_1) & D(X_2) & \cdots & cov(X_2, X_p) \\ \vdots & \vdots & & \vdots \\ cov(X_p, X_1) & cov(X_p, X_2) & \cdots & D(X_p) \end{bmatrix} \\ &= (\sigma_{ij})\end{aligned}$$

称它为 p 维随机向量 X 的协方差阵。称 $|cov(X, X)|$ 为 X 的广义方差，是协方差阵的行列式的值。



随机向量 X 和 Y 的协方差阵

设 $X = (X_1, X_2, \dots, X_p)^T$ 和 $Y = (Y_1, Y_2, \dots, Y_q)^T$ 分别为 p 维和 q 维随机向量，它们之间的协方差阵定义为一个 $p \times q$ 矩阵，其元素是 $cov(X_i, Y_j)$ ，即

$$cov(X, Y) = \left(cov(X_i, Y_j) \right), i = 1, 2, \dots, p; j = 1, 2, \dots, q$$

若 $cov(X, Y) = 0$ ，称 X 和 Y 是不相关的。



当 A 和 B 为常数矩阵时，由定义可推出如下性质

$$(1) \quad D(AX) = AD(X)A^T = A\Sigma A^T$$

$$(2) \quad cov(AX, BY) = Acov(X, Y)B^T$$

(3) 设 X 为 p 维随机向量，期望和协方差存在，记
 $\mu = E(X)$, $\Sigma = D(X)$, A 为 $p \times p$ 常数阵，则

$$E(X^T AX) = tr(A\Sigma) + \mu^T A\mu$$

随机向量 $X = (X_1, X_2, \dots, X_p)^T$ ，其协方差阵 Σ 都是对称阵，同时总是半正定的。大多数情形下是正定的。



随机向量 X 的相关阵

若随机向量 $X = (X_1, X_2, \dots, X_p)^T$ 的协方差阵存在, 且每个分量的方差大于零, 则 X 的相关阵定义为:

$$R = \left(\text{corr}(X_i, X_j) \right) = (r_{ij})_{p \times p}$$

$$r_{ij} = \frac{\text{cov}(X_i, X_j)}{\sqrt{D(X_i)}\sqrt{D(X_j)}}, \quad i, j = 1, 2, \dots, p$$

r_{ij} 也称分量 X_i 和 X_j 之间的（线性）相关系数。

在数据处理时，为克服指标量纲不同的影响，使用统计分析方法前，将每个指标“标准化”：

$$X_j^* = \frac{X_j - E(X_j)}{[var(X_j)]^{1/2}}, j = 1, 2, \dots, p$$

$$X^* = (X_1^*, X_2^*, \dots, X_p^*)$$

于是 $E(X^*) = \mathbf{0}$, $D(X^*) = corr(X) = R$

即标准化数据的协方差矩阵正好是原指标的相关阵：

$$R = \frac{1}{n-1} X^{*T} X^*$$

§ 4.2 多元正态分布

§ 4.2.1 多元正态分布的定义

§ 4.2.2 多元正态分布的基本性质

§ 4.2.3 多元正态条件分布

§ 4.2.1 多元正态分布的定义

回顾二元正态分布

$x = (x_1, x_2)^T \sim N(\mu_1, \sigma_1^2; \mu_2, \sigma_2^2; \rho)$, 其密度函数为

$p(x_1, x_2)$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right] \right\}$$

其中 $\sigma_1^2 > 0, \sigma_2^2 > 0, |\rho| < 1$ 。特征函数为

$$\varphi(t_1, t_2) = E[e^{i(t_1x_1 + t_2x_2)}]$$

$$= \exp \left\{ i(t_1\mu_1 + t_2\mu_2) - \frac{1}{2}(\sigma_1^2 t_1^2 + 2\rho\sigma_1\sigma_2 t_1 t_2 + \sigma_2^2 t_2^2) \right\}$$

若令 $\mu = (\mu_1, \mu_2)^T, t = (t_1, t_2)^T$, 以及

$$V = \text{Var}(x) = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$



则分别有

$$p(x) = (2\pi)^{-\frac{n}{2}} |V|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (x - \mu)^T V^{-1} (x - \mu) \right\}$$

$$\varphi(t) = E(e^{it^T x}) = \exp \left\{ it^T \mu - \frac{1}{2} t^T V t \right\}$$

上述两个表达式主要由二元正态向量 x 的数学期望 $E(x) = \mu$ 和协方差阵 $Var(x) = V = (\sigma_{ij})$ 决定, 易推广到多维随机向量的场合, 可引出下述多元正态分布的密度函数定义和特征函数定义。



定义4.2.1（密度函数定义） 若 p 维随机向量（随机变量） $X = (X_1, X_2, \dots, X_p)^T$ （联合）概率密度函数为

$$f(x_1, x_2, \dots, x_p) = \frac{1}{(2\pi)^{\frac{p}{2}} |V|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (X - \mu)^T V^{-1} (X - \mu) \right\}$$

则称随机向量 X 为 p 维正态随机向量，其中 μ 称为均值向量， V 为协方差矩阵(协差阵)，且 $V > 0$ 。



对于一般情形 $V \geq 0$ ，仍可定义多维正态随机向量，记为 $X \sim N_p(\mu, V)$ 。当 $V > 0$ 时， X 由前面密度表示，而当 $|V| = 0$ 时， X 的分布是退化的正态分布。



定义4.2.2 (特征函数定义) 若 p 维随机向量 (随机变量) $X = (X_1, X_2, \dots, X_p)^T$ 特征函数为

$$\varphi(t) = E(e^{it^T x}) = \exp \left\{ it^T \mu - \frac{1}{2} t^T V t \right\}$$

则称随机向量 X 为 p 维正态随机向量, 其中 μ 称为均值向量, V 为协方差矩阵(协差阵), 且 $V \geq 0$ 。

定义4.2.3 若对任何非零向量 $a \in R^p$, X 的线性组合 $a^T X$ 服从一元正态分布 $N(a^T \mu, a^T V a)$, 则称 X 服从 p 元正态分布 $N_p(\mu, V)$ 。

利用特征函数的性质可以证明定义4.2.2与定义4.2.3是等价的。

§ 4.2.2 多元正态分布的基本性质

- (1) p 维正态分布由其均值向量和协方差阵唯一确定。
- (2) 对于任一 p 维向量 μ 及 p 阶非负定矩阵 V ，必存在 p 维正态随机向量 $X \sim N_p(\mu, V)$ 。
- (3) 设 $X \sim N_p(\mu, V)$ ， A 是 $m \times p$ 常数矩阵， b 是 m 维向量，若令 $Y = AX + b$ ，则 $Y \sim N_p(A\mu + b, AVA^T)$ 。

证明 首先计算的 Y 特征函数

$$\begin{aligned}\varphi_Y(t) &= E(e^{it^T Y}) = E[e^{it^T (AX+b)}] = e^{it^T b} E(e^{it^T AX}) \\ &= e^{it^T b} \varphi_X(A^T t) = e^{it^T b} \exp \left\{ it^T A\mu - \frac{1}{2} t^T AVA^T t \right\} \\ &= \exp \left\{ it^T (A\mu + b) - \frac{1}{2} t^T (AVA^T) t \right\}\end{aligned}$$

因此由多元正态分布的特征函数定义**4.2.2**，有

$$Y \sim N_p(A\mu + b, AVA^T)$$



(4) X 为 p 维正态随机向量的充要条件为对任一 p 维向量 c , $c^T X$ 是一维正态随机变量。

(5) 设 $X = (X_1^T, X_2^T)^T$ 为多维正态随机向量, 则 X_1 与 X_2 互不相关的充要条件是 X_1 与 X_2 相互独立。

注: 若 $Cov(X, Y) = 0$, 则称 X 与 Y 互不相关。

(6) 设 $X \sim N_p(\mu, V)$, 则 $rank(V) = m$ 的充要条件是存在 $m \times p$ 的矩阵 $B(BB^T = V)$ 使得 $X = BY + \mu$, 其中 $Y \sim N_m(0, I_m)$ 。



证明 充分性由性质3立得。下证必要性。

由于 V 是秩为 m 的非负定阵，则必存在正交矩阵 U 使得

$$U^T V U = \begin{pmatrix} \lambda_1 & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \lambda_m & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}$$

其中 $\lambda_i > 0, i = 1, \cdots, m$ 。



$$\text{令 } \Lambda = \begin{pmatrix} \lambda_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \lambda_m \end{pmatrix}, \text{ 则有}$$

$$\begin{pmatrix} \Lambda^{-\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & I_{p-m} \end{pmatrix} U^T V U \begin{pmatrix} \Lambda^{-\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & I_{p-m} \end{pmatrix} = \begin{pmatrix} I_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \triangleq D$$

$$\text{令 } \mathbf{Z} = \begin{pmatrix} Y \\ W \end{pmatrix} = \begin{pmatrix} \Lambda^{-\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & I_{p-m} \end{pmatrix} U^T (X - \mu)$$

则由性质3知 $\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{D})$, 且 $\mathbf{Y} \sim N_m(\mathbf{0}, \mathbf{I}_m)$, 由上式

可得 $\mathbf{X} - \boldsymbol{\mu} = \mathbf{U} \begin{pmatrix} \Lambda^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{p-m} \end{pmatrix} \mathbf{Z} = \mathbf{U} \begin{pmatrix} \Lambda^{\frac{1}{2}} \\ \mathbf{0} \end{pmatrix} \mathbf{Y}$, 若记

$\mathbf{B} = \mathbf{U} \begin{pmatrix} \Lambda^{\frac{1}{2}} \\ \mathbf{0} \end{pmatrix}$, 它是 $p \times m$ 矩阵, 即有 $\mathbf{X} = \mathbf{B}\mathbf{Y} + \boldsymbol{\mu}$



(7) 若 $X \sim N_p(\mu, V)$, 且 $|V| \neq 0$, 则

$$\eta \triangleq (X - \mu)^T V^{-1} (X - \mu) \sim \chi^2(p).$$

证明 由 $|V| \neq 0$ 可知 V 是正定矩阵, 所以 $V^{-\frac{1}{2}}$ 存在且为对称矩阵, 这样

$$\eta = (X - \mu)^T V^{-\frac{1}{2}} V^{-\frac{1}{2}} (X - \mu) = \left(V^{-\frac{1}{2}} (X - \mu) \right)^T \left(V^{-\frac{1}{2}} (X - \mu) \right)$$

令 $Y = V^{-\frac{1}{2}} (X - \mu)$, 则 $\eta = Y^T Y$, 且 $Y \sim N_p(0, I_p)$.



由性质3知 Y 的每个分量 Y_i 服从标准正态分布，且相互独立，故由 χ^2 -分布的定义知 $Y^T Y \sim \chi^2(p)$.



例4.2.1 设 $X \sim N_p(\mu, V)$, $V > 0$, 将 X 分块为 $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ $\begin{matrix} r \\ p-r \end{matrix}$,

$$\text{相应 } \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},$$

证明 $X_1 - V_{12}V_{22}^{-1}X_2$ 与 X_2 相互独立, $X_2 - V_{21}V_{11}^{-1}X_1$ 与 X_1 也相互独立。

证明 令 $A = (I_r, -V_{12}V_{22}^{-1})$, $B = (0, I_{p-r})$, 则

$$\text{rank}(A) = r, \text{rank}(B) = p - r$$





$$X_1 - V_{12}V_{22}^{-1}X_2 = AX, X_2 = BX$$

且由性质 (3) 知, $\begin{pmatrix} A \\ B \end{pmatrix} X$ 服从 p 维正态分布。由于

$$\text{Cov}(AX, BX) = A\text{Cov}(X, X)B^T = \begin{pmatrix} I_r & -V_{12}V_{22}^{-1} \end{pmatrix} V \begin{pmatrix} 0 \\ I_{p-r} \end{pmatrix}$$

所以由性质 (5) 知, AX 与 BX 相互独立, 即 $X_1 -$

$V_{12}V_{22}^{-1}X_2$ 与 X_2 相互独立。同理可证 $X_2 - V_{21}V_{11}^{-1}X_1$ 与 X_1 也相互独立。



§ 4.2.3 多元正态条件分布

设 $X \sim N_p(\mu, V)$, $V > 0$, 将 X 分块为 $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}_{p-k}^k$, 同时对 μ 和 V 相应分块

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

本小节主要介绍在给定 X_2 时 X_1 的条件分布。

定理4.2.1 设 $X \sim N_p(\mu, V)$, $V > 0$, 则给定 X_2 时 X_1 的条件分布为 $N_k(\mu_{1 \cdot 2}, V_{11 \cdot 2})$, 其中

$$\mu_{1 \cdot 2} = \mu_1 + V_{12}V_{22}^{-1}(X_2 - \mu_2)$$

$$V_{11 \cdot 2} = V_{11} - V_{12}V_{22}^{-1}V_{21}$$

同样, 给定 X_1 时 X_2 的条件分布为 $N_{p-k}(\mu_{2 \cdot 1}, V_{22 \cdot 1})$, 其中

$$\mu_{2 \cdot 1} = \mu_2 + V_{21}V_{11}^{-1}(X_1 - \mu_1)$$

$$V_{22 \cdot 1} = V_{22} - V_{21}V_{11}^{-1}V_{12}$$

证明 由 $V > 0$, 可知 $V_{11} > 0, V_{22} > 0, V_{11 \cdot 2} > 0$ 。令

$$A = \begin{pmatrix} I_k & -V_{12}V_{22}^{-1} \\ \mathbf{0} & I_{p-k} \end{pmatrix},$$

$$Y = AX = \begin{pmatrix} X_1 - V_{12}V_{22}^{-1}X_2 \\ X_2 \end{pmatrix} \triangleq \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

则由4.2.2性质 (3) 知, $Y_1 \sim N_k(\mu_1 - V_{12}V_{22}^{-1}\mu_2, V_{11 \cdot 2})$

由例4.2.1可知 Y_1 与 $Y_2 = X_2$ 相互独立, 因此

$$f_Y(Y) = f_{Y_1}(Y_1) f_{Y_2}(Y_2)$$

注意到变换雅可比行列式 $J(Y, X) = |A| = 1$ ，则 X 的联合密度可以表示为

$$\begin{aligned} f_X(X) &= f_Y(Y) |J(Y, X)| = f_Y(Y) \\ &= f_{Y_1}(Y_1) f_{Y_2}(Y_2) = f_{Y_1}(Y_1) f_{X_2}(X_2) \end{aligned}$$

故给定 X_2 时 X_1 的条件密度函数为

$$f(X_1|X_2) = \frac{f_X(X)}{f_{X_2}(X_2)} = f_{Y_1}(Y_1)$$



$$= (2\pi)^{-\frac{k}{2}} |V_{11.2}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (Y_1 - \mu_1 + V_{12}V_{22}^{-1}\mu_2)^T V_{11.2}^{-1} (Y_1 - \mu_1 + V_{12}V_{22}^{-1}\mu_2) \right\}$$

将 $Y_1 = X_1 - V_{12}V_{22}^{-1}X_2$ 代入, 可得

$$f(X_1|X_2) = (2\pi)^{-\frac{k}{2}} |V_{11.2}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (X_1 - \mu_{1.2})^T V_{11.2}^{-1} (X_1 - \mu_{1.2}) \right\}$$

即给定 X_2 时, $X_1 \sim N_k(\mu_{1.2}, V_{11.2})$ 。





例4.2.2 设 $X \sim N_3(\mu, V)$, 其中 $\mu = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$, $V = \begin{pmatrix} 16 & -4 & 2 \\ -4 & 4 & -1 \\ 2 & -1 & 4 \end{pmatrix}$

在给定 $X_1 + 2X_3$ 时, 求 $\begin{pmatrix} X_2 - X_3 \\ X_1 \end{pmatrix}$ 的条件分布。

解 令 $Y_1 = \begin{pmatrix} X_2 - X_3 \\ X_1 \end{pmatrix}$, $Y_2 = X_1 + 2X_3$, 则

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$$



由4.2.2性质 (3) , $Y \sim N_3(E(Y), Var(Y))$, 其中

$$E(Y) = E \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}$$

$$Var(Y) = Var \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 16 & -4 & 2 \\ -4 & 4 & -1 \\ 2 & -1 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix}^T$$

$$\text{Var}(Y) = \begin{pmatrix} 10 & -6 & -16 \\ -6 & 16 & 20 \\ -16 & 20 & 40 \end{pmatrix}$$

所以, 由定理4.2.1, 在给定 $X_1 + 2X_3$ 时,

$$\begin{pmatrix} X_2 - X_3 \\ X_1 \end{pmatrix} \sim N_3(E(Y_1|Y_2), \text{Var}(Y_1|Y_2)), \text{ 其中}$$

$$E(Y_1|Y_2) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} -16 \\ 20 \end{pmatrix} \times \frac{1}{40} \times (Y_2 + 3) = \begin{pmatrix} -\frac{2}{5}X_1 - \frac{4}{5}X_3 + \frac{4}{5} \\ \frac{1}{2}X_1 + X_3 + \frac{5}{2} \end{pmatrix}$$

$$\text{Var}(Y_1|Y_2) = \begin{pmatrix} 10 & -6 \\ -6 & 16 \end{pmatrix} - \begin{pmatrix} -16 \\ 20 \end{pmatrix} \times \frac{1}{40} \times (-16 \quad 20) = \begin{pmatrix} 3.6 & 2 \\ 2 & 6 \end{pmatrix}$$

当 X_1 与 X_2 存在相关关系时，条件数学期望

$E(X_1|X_2) = \mu_{1\cdot 2} = \mu_1 + V_{12}V_{22}^{-1}(X_2 - \mu_2)$ 可以被视为 X_1 对 X_2 的多元回归函数， $V_{12}V_{22}^{-1}$ 可以称为 X_1 对 X_2 的回归系数矩阵。

条件协方差矩阵 $V_{11\cdot 2} = V_{11} - V_{12}V_{22}^{-1}V_{21}$ 又称为协方差矩阵。它是 X_1 对 X_2 回归后的剩余方差，是从 X_1 的方差中扣除能被线性表出的部分（即

$\mu_1 + V_{12}V_{22}^{-1}(X_2 - \mu_2)$ ）的方差之后的剩余部分。

例4.2.3 设 $X = (X_1, X_2, X_3)^T \sim N_3(\mu, V)$, 其中 $\mu =$

$\begin{pmatrix} 10 \\ 4 \\ 7 \end{pmatrix}$, $V = \begin{pmatrix} 9 & -3 & -3 \\ -3 & 5 & 1 \\ -3 & 1 & 5 \end{pmatrix}$ 。求 X_1 对 $(X_1, X_2)^T$ 的回归函数。

解 X_1 对 $(X_1, X_2)^T$ 的回归函数即 $E(X_1 | (X_1, X_2)^T)$ 。根据条件期望计算公式可得

$$\begin{aligned} E(X_1 | (X_1, X_2)^T) &= \mathbf{10} + (-3 \quad -3) \begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}^{-1} \begin{pmatrix} X_2 - 4 \\ X_3 - 7 \end{pmatrix} \\ &= 15.5 - 0.5X_2 - 0.5X_3 \end{aligned}$$

因此, X_1 对 $(X_1, X_2)^T$ 的回归函数为

$$X_1 = 15.5 - 0.5X_2 - 0.5X_3$$

相关性是指两个随机变量之间存在线性关联，可以用

简单相关系数 $\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$ 来进行度量。

但两个统计相关的随机变量不一定有因果关系， X_1 与 X_2 关系有可能会受到第三个因素的影响 X_3 。

把给定 X_3 的条件下 X_1 与 X_2 的相关系数称为的偏相关系数。偏相关系数也可以解释成在扣除了 X_3 的影响之后 X_1 与 X_2 的相关系数。

定义4.2.4 设 $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}_{p-k}^k \sim N_p(\mu, V), V > 0$ 。在给定 X_2 条件下, X_1 的偏协方差矩阵为

$$\text{Var}(X_1|X_2) = V_{11.2} = (\sigma_{ij \cdot k+1 \dots p})_{k \times k}$$

在给定 X_2 条件下, X_1 与 X_2 的偏相关系数定义为

$$\rho_{ij \cdot k+1 \dots p} = \frac{\sigma_{ij \cdot k+1 \dots p}}{\sqrt{\sigma_{ii \cdot k+1 \dots p} \sigma_{jj \cdot k+1 \dots p}}}$$

由偏相关系数组成的矩阵

$$R(X_1|X_2) = R_{11 \cdot 2} = (\rho_{ij \cdot k+1 \cdots p})_{k \times k}$$

称为给定 X_2 条件下 X_1 的偏相关矩阵。

设总体 $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}_{p-k}^k \sim N_p(\mu, V), V > 0$ 。在给定 X_2 条件下, X_1 的偏协方差矩阵为

$$\text{Var}(X_1|X_2) = V_{11.2} = (\sigma_{ij \cdot k+1 \dots p})_{k \times k}$$

在给定 X_2 条件下, X_1 与 X_2 的偏相关系数定义为

$$\rho_{ij \cdot k+1 \dots p} = \frac{\sigma_{ij \cdot k+1 \dots p}}{\sqrt{\sigma_{ii \cdot k+1 \dots p} \sigma_{jj \cdot k+1 \dots p}}}$$

§ 4.3 均值向量与协方差阵的估计



在实际问题中，多元正态分布的均值 μ 和协方差阵 V 往往未知，需要利用样本对它们进行估计和检验，这些内容可以看成是一元正态分布参数 μ 和 σ^2 的估计和检验问题在多元情况下的推广。

下面介绍多元正态总体参数 μ 和 V 的估计。



设 x_1, x_2, \dots, x_n 是来自多元正态总体 $N_p(\mu, V)$ 的简单样本，样本容量 $n > p$ 。

为简单起见，在此仅考虑 $V > \mathbf{0}$ ，即非退化情形。令

$$X_{n \times p} = (x_1, x_2, \dots, x_n)^T = \begin{pmatrix} x_1^T \\ \vdots \\ x_n^T \end{pmatrix}$$

则样本的联合分布密度函数为

$$p(X; \mu, V) = \prod_{i=1}^n p(x_i; \mu, V)$$

化简后

$$p(X; \mu, V)$$

$$= (2\pi)^{-\frac{np}{2}} |V|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \{ V^{-1} [S + n(\bar{x} - \mu)(\bar{x} - \mu)^T] \} \right\}$$

其中,

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k \quad \text{——样本均值向量}$$

$$S = \sum_{k=1}^n (X_k - \bar{X})(X_k - \bar{X})^T \quad \text{——样本离差阵}$$

为求出 μ 和 V 的极大似然估计, 先介绍下列引理。

引理4.3.1 设 A 是 $m \times m$ 阶正定矩阵, 则当 $A = nI_m$ 的

时候, 函数 $f(A) = |A|^{\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(A) \right\}$, 达最大值

$$n^{\frac{mn}{2}} e^{-\frac{mn}{2}}。$$

定理4.3.1 设 $X_1, X_2, \dots, X_n (n > p)$ 是多元正态总体 $N_p(\mu, V)$ 的简单样本, 且 $V > 0$, 则 \bar{X} 是 μ 的极大似然估计, $\frac{1}{n}S$ 是 V 的极大似然估计。

多元正态总体 $N_p(\mu, V)$ 参数 μ 和 V 的估计量有如下性质

性质4.3.1 设 $X_1, X_2, \dots, X_n (n > p)$ 是多元正态总体 $N_p(\mu, V)$ 的简单样本, 且 $V > 0$, 则 \bar{X} 是 μ 的一致最小方差无偏估计, $\frac{1}{n-1}S$ 是 V 的一致最小方差无偏估计。

性质4.3.2 $\bar{X} \sim N_p\left(\mu, \frac{1}{n}V\right)$, 且 \bar{X} 与 S 相互独立。

性质4.3.3 样本离差阵 S 服从自由度为 $n - 1$ 的Wishart分布, 即 $S \sim W_p(n - 1, V)$ 。

定义4.3.1 设 $X = (X_1, X_2, \dots, X_n)^T$, $X_i \sim N_p(\mu_i, V)$ ($i = 1, 2, \dots, n$), 且 X_1, X_2, \dots, X_n 相互独立, 记 $M = E(X) = (\mu_1, \mu_2, \dots, \mu_n)^T$, 称 $A = X^T X = \sum_{i=1}^n X_i X_i^T$ 所服从的分布是自由度为 n 的非中心Wishart分布, 记为 $A \sim W_p(n, V, \tau)$, 其中 $\tau = M^T M$ 为非中心参数。当 $\mu_1 = \mu_2 = \dots = \mu_n = 0$ 时有 $\tau = 0$, 此时称 A 服从中心Wishart分布, 记为 $W_p(n, V)$ 。



§ 4.4 均值向量的检验

§ 4.4.1 协差阵 \mathbf{V} 已知时, 均值 μ 的检验

§ 4.4.2 协差阵 \mathbf{V} 未知时, 均值 μ 的检验

§ 4.4.3 两个正态总体均值相等的检验



§ 4.4.1 协差阵 V 已知时, 均值 μ 检验

设 $X_1, X_2, \dots, X_n (n > p)$ 是多元正态总体 $N_p(\mu, V)$ 的简单样本, 其中 V 已知。考虑假设检验问题

$$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$$

首先导出检验似然比。当 V 已知时, 参数空间 θ 为 p 维欧式空间 R^p , 原假设 H_0 的参数空间为 $\theta_0 = \{\mu_0\}$ 。



由于样本的联合密度函数为

$$p(\mathbf{X}; \boldsymbol{\mu}) = (2\pi)^{-\frac{np}{2}} |\mathbf{V}|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \{ \mathbf{V}^{-1} [\mathbf{S} + n(\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})^T] \} \right\}$$

从而

$$\sup_{\boldsymbol{\mu} \in \boldsymbol{\Theta}} \{p(\mathbf{X}; \boldsymbol{\mu})\} = p(\mathbf{X}; \bar{\mathbf{X}})$$

$$= (2\pi)^{-\frac{np}{2}} |\mathbf{V}|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{V}^{-1} \mathbf{S}) \right\}$$



$$\sup_{\mu \in \Theta_0} \{p(X; \mu)\} = p(X; \mu_0)$$

$$= (2\pi)^{-\frac{np}{2}} |V|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \{ V^{-1} [S + n(\bar{X} - \mu_0)(\bar{X} - \mu_0)^T] \} \right\}$$

因此检验假设检验问题的似然比为

$$\lambda(X) = \frac{\sup_{\mu \in \Theta} \{p(X; \mu)\}}{\sup_{\mu \in \Theta_0} \{p(X; \mu)\}} = n(\bar{X} - \mu_0)^T V^{-1} (\bar{X} - \mu_0)$$

令 $D = n(\bar{X} - \mu_0)^T V^{-1}(\bar{X} - \mu_0)$, 则可以证明当 H_0 成立时, 即 $\mu = \mu_0$ 时, $D \sim \chi^2(p)$

而当 H_0 不成立时, D 有偏大的趋势。因此, 对给定的显著性水平 α , 当

$$D = n(\bar{X} - \mu_0)^T V^{-1}(\bar{X} - \mu_0) \geq \chi_{1-\alpha}^2(p)$$

时拒绝 H_0 , 否则接受 H_0 , 即拒绝域为

$$W = \{D: D \geq \chi_{1-\alpha}^2(p)\}$$

实际上，检验统计量 D 可以看成是相应一维情形的推广。在一维情形中，当 σ^2 已知，使用统计量 $z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ 作为检验统计量。当 H_0 成立时， $z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$ ，平方有 $z^2 = n(\bar{X} - \mu_0)^T V^{-1}(\bar{X} - \mu_0) \sim \chi^2(1)$ 。在多元情形， $\bar{X} - \mu_0$ 变成列向量，而总体方差 σ^2 变为协方差矩阵 V ，就可得到检验统计量 D 。

§ 4.4.2 协差阵 V 未知时, 均值 μ 检验

设 $X_1, X_2, \dots, X_n (n > p)$ 是多元正态总体 $N_p(\mu, V)$ 的简单样本, 其中 V 未知。考虑假设检验问题

$$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$$

仍根据似然比原则导出检验统计量及检验的拒绝域。

当 V 未知时, 样本的联合密度函数为

$$p(X; \mu, V) = (2\pi)^{-\frac{np}{2}} |V|^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \{ V^{-1} [S + n(\bar{X} - \mu)(\bar{X} - \mu)^T] \} \right\}$$

参数空间为 $\Theta = \{(\mu, V): \mu \in R^p, V > \mathbf{0}\}$, 由定理4.3.1

知, μ 和 V 的极大似然估计分别为 \bar{X} 和 $\frac{1}{n}S$ 。

$$\sup_{\theta \in \Theta} \{p(X; \mu, V)\} = p\left(X; \bar{X}, \frac{1}{n}S\right) = (2\pi)^{-\frac{np}{2}} |S|^{-\frac{n}{2}} n^{\frac{np}{2}} \exp \left\{ -\frac{np}{2} \right\}$$

原假设 H_0 的参数空间为 $\Theta_0 = \{(\mu_0, V): V > \mathbf{0}\}$, 与定理4.3.1证明过程类似, 可得 V 的极大似然估计为

$$\begin{aligned}\hat{V}_0 &= \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)(X_i - \mu_0)^T \\ &= \frac{1}{n} [S + n(\bar{X} - \mu_0)(\bar{X} - \mu_0)^T]\end{aligned}$$



$$\sup_{\theta \in \Theta_0} \{p(X; \mu, V)\} = p(X; \mu_0, \hat{V}_0)$$

$$= (2\pi)^{-\frac{np}{2}} |S + n(\bar{X} - \mu_0)(\bar{X} - \mu_0)^T|^{-\frac{n}{2}} n^{\frac{np}{2}} \exp\left\{-\frac{np}{2}\right\}$$

进而检验假设 H_0 的似然比为

$$\begin{aligned} \lambda(X) &= \frac{\sup_{\theta \in \Theta} \{p(X; \mu, V)\}}{\sup_{\theta \in \Theta_0} \{p(X; \mu, V)\}} = \frac{|S|^{-\frac{n}{2}}}{|S + n(\bar{X} - \mu_0)(\bar{X} - \mu_0)^T|^{-\frac{n}{2}}} \\ &= [1 + (\bar{X} - \mu_0)^T S^{-1} (\bar{X} - \mu_0)]^{\frac{n}{2}} \end{aligned}$$



令 $F = \frac{n}{p}(n - p)(\bar{X} - \mu_0)^T S^{-1}(\bar{X} - \mu_0)$ ，则可以证明

当 H_0 成立时，即 $\mu = \mu_0$ 时， $F \sim F(p, n - p)$

而当 H_0 不成立时， F 有偏大的趋势。因此，对给定的显著性水平 α ，当

$$F = \frac{n}{p}(n - p)(\bar{X} - \mu_0)^T S^{-1}(\bar{X} - \mu_0) \geq F_{1-\alpha}(p, n - p)$$

时拒绝 H_0 ，否则接受 H_0 ，即拒绝域为

$$W = \{F: F \geq F_{1-\alpha}(p, n - p)\}$$

§ 4.4.3 两个正态总体均值相等的检验

设 $X_1, X_2, \dots, X_{n_1} (n_1 > p)$ 是多元正态总体 $N_p(\mu_1, V)$ 的简单样本, $Y_1, Y_2, \dots, Y_{n_2} (n_2 > p)$ 是多元正态总体 $N_p(\mu_2, V)$ 的简单样本, 且两个样本相互独立, 协方差阵 $V > 0$ 。考虑假设检验问题

$$H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$$

根据协方差阵 V 已知和未知分两种情形:

(1) V 已知

在一维情形，当 σ^2 已知时，检验两个正态总体均值相等

的检验统计量为 $z = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma^2}}$ ，平方并重新整理有

$$z^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X} - \bar{Y}) (\sigma^2)^{-1} (\bar{X} - \bar{Y})$$

推广到 p 维情形， $\bar{X} - \bar{Y}$ 变为向量， σ^2 变为协方差阵 V 。

$$\text{检验统计量 } D^* = \frac{n_1 n_2}{n_1 + n_2} (\bar{X} - \bar{Y})^T V^{-1} (\bar{X} - \bar{Y})$$



可以证明当 H_0 成立时, 即 $\mu_1 = \mu_2$ 时,

$$D^* = \frac{n_1 n_2}{n_1 + n_2} (\bar{X} - \bar{Y})^T V^{-1} (\bar{X} - \bar{Y}) \sim \chi^2(p)$$

而当 H_0 不成立时, D^* 有偏大的趋势。因此, 对给定的显著性水平 α , 当

$$D^* = \frac{n_1 n_2}{n_1 + n_2} (\bar{X} - \bar{Y})^T V^{-1} (\bar{X} - \bar{Y}) \geq \chi_{1-\alpha}^2(p)$$

时拒绝 H_0 , 否则接受 H_0 , 即拒绝域为

$$W = \{D^*: D^* \geq \chi_{1-\alpha}^2(p)\}$$



(2) V 未知

在一维情形，当 σ^2 未知时，检验两个正态总体均值相

等的检验统计量为 $t = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_w^2}}$ ，其中

$$S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2,$$

$$S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

即用作估计 σ^2 的 S_w^2 是样本方差 S_1^2 和 S_2^2 的加权平均。对 t

平方并重新整理有 $t^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X} - \bar{Y})(S_w^2)^{-1}(\bar{X} - \bar{Y})$

推广到 p 维情形, $\bar{X} - \bar{Y}$ 变为向量, 用样本协方差阵



$$\begin{aligned}\hat{V}_1 &= \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})(X_i - \bar{X})^T, \hat{V}_2 \\ &= \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})(Y_i - \bar{Y})^T\end{aligned}$$

分别替代一维时的样本方差 S_1^2 和 S_2^2 , S_w^2 变为样本联合

$$\text{协方差阵 } \hat{V} = \frac{(n_1 - 1)\hat{V}_1 + (n_2 - 1)\hat{V}_2}{n_1 + n_2 - 2}$$



检验统计量 $F^* = \frac{n_1 n_2 (n_1 + n_2 - p - 1)}{p (n_1 + n_2) (n_1 + n_2 - 2)} (\bar{X} - \bar{Y})^T \hat{V}^{-1} (\bar{X} - \bar{Y})$

其中 $\hat{V} = \frac{1}{n_1 + n_2 - 2} (S_1 + S_2)$ 是协方差阵 V 的估计量。

可以证明当 H_0 成立时, 即 $\mu_1 = \mu_2$ 时,

$$F^* \sim F(p, n_1 + n_2 - p - 1)$$

而当 H_0 不成立时, F^* 有偏大的趋势。因此, 对给定的显著性水平 α , 拒绝域为

$$W = \{F^*: F^* \geq F_{1-\alpha}(p, n_1 + n_2 - p - 1)\}$$