两个均值向量的比较

张伟平

zwp@ustc.edu.cn

Office: 东区管理科研楼 1006

Phone: 63600565

课件 http://staff.ustc.edu.cn/~zwp/

论坛 http://fisher.stat.ustc.edu.cn

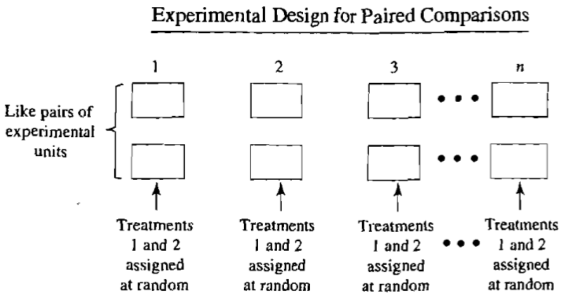
简介

1.1 成对比较问题

上一讲中对均值向量的 Hotelling's T^2 检验可以方便的推广到多个均值向量的比较问题中.

- **成对比较设计**: 每个样本单元使用两种不同的处理 (treatment), 来研究两种处理是否存在差异.
  - 投放广告前后某个市场某个产品的销售量变动来研究广告投放的效用
  - 服用某种降压药前后血压的变化, 来研究该药物的效用
  - 一些路口使用交通信号灯前后交通事故数的变化, 来研究交通信号等的效用
- 成对比较设计的优点是测量结果的差异仅仅是由于不同处理的效应造成的, 因为其他条件完全相同 (同一个体仅处理不同).

- 当然, 在一些问题里成对比较设计不会那么简单. 比如服药前后其他变量可能也会发生变化, 造成测量结果的差异除处理不同原因外, 还可能有其他条件发生变化的原因.



$p = 1$ 时的成对比较问题

- 在一元场合下, 以响应变量的某个连续取值的指标为例. 记 X_{1j} 和 X_{2j} 为第 j 个试验单元的响应变量分别在处理 1 和处理 2 下的测量值.

- 感兴趣的问题是处理 1 和处理 2 是否有差异.

- 设 $D_j = X_{1j} - X_{2j}, j = 1, \ldots, n$, 则 D_j 反应了两种处理的差异. 假设

  - $D_1, \ldots, D_n$ 相互独立同分布
  - $D_1 \sim N_1(\delta, \sigma_\delta^2)$

- 在上述假设条件下, 量

$$t(\delta) = \frac{\bar{D} - \delta}{s_d/\sqrt{n}} \sim t_{n-1}$$

其中 $\bar{D} = \frac{1}{n}\sum_j D_j, s_d^2 = \frac{1}{n-1}\sum_j (D_j - \bar{D})^2$.

- 因此假设 $H_0 : \delta = 0 \leftrightarrow H_1 : \delta \neq 0$ 的水平 α 检验拒绝域为 $|t(0)| > t_{n-1}(\alpha/2)$.

- 等价地, δ 的 $1-\alpha$ 置信区间为

$$\bar{D} - t_{n-1}(\alpha/2)\frac{s_d}{\sqrt{n}} \leq \delta \leq \bar{D} + t_{n-1}(\alpha/2)\frac{s_d}{\sqrt{n}}$$

- 当样本量 n 充分大时, $t(\delta)$ 有渐近正态分布 $N(0,1)$, 因此可以得到渐近水平 α 检验拒绝域

$$|t(0)| > z(\alpha/2), \text{等价的} \quad t^2(0) > \chi_1^2(\alpha)$$

类似可得渐近 $1-\alpha$ 置信区间

$$\bar{D} - z(\alpha/2)\frac{s_d}{\sqrt{n}} \leq \delta \leq \bar{D} + z(\alpha/2)\frac{s_d}{\sqrt{n}}$$

$p > 1$ 时的成对比较问题

- 当每个样本单元的 p 个变量被测量时候, 我们关心差异向量 (处理 1-处理 2):

$$\mathbf{D}_j = \begin{pmatrix} D_{j1} \\ \vdots \\ D_{jp} \end{pmatrix} = \begin{pmatrix} X_{1j1} \\ \vdots \\ X_{1jp} \end{pmatrix} - \begin{pmatrix} X_{2j1} \\ \vdots \\ X_{2jp} \end{pmatrix}$$

$j = 1, \ldots, n$. 其中 X_{1jk}, X_{2jk} 分别表示第 j 个样本单元在处理 1 和处理 2 下第 k 个变量的测量值.

- 因此, 假设 $\mathbf{D}_1, \ldots, \mathbf{D}_n$ $i.i.d \sim N_p(\delta, \Sigma)$ 时, 可以使用 Hotelling's T^2 统计量

$$T_\delta^2 = n(\bar{D} - \delta)' S_D^{-1} (\bar{D} - \delta) \sim \frac{p(n-1)}{n-p} F_{p,n-p}$$

其中 $\bar{D}$ 和 S_D 分别为基于 $\mathbf{D}_1, \ldots, \mathbf{D}_n$ 的样本均值和样本协方差.

- 实际中很多问题是研究假设 “两种处理没有平均差异” 是否成立, 这等价于 $H_0 : \delta = 0 \leftrightarrow H_1 : \delta \neq 0$. 由 Hotelling's T^2 检验, 其水平 α 检验的拒绝域为

$$T_0^2 = n\bar{D}'S_D^{-1}\bar{D} > \frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)$$

  当不能拒绝 H_0 时候, 我们得出结论: p 个变量的任何分量上不存在处理效应.

- δ 的 $1-\alpha$ 置信域为

$$\left\{\delta : n(\bar{D}-\delta)'S_D^{-1}(\bar{D}-\delta) \leq \frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)\right\}$$

- $\mathbf{a}'\delta$ 的 $1-\alpha$ 同时置信区间为

$$\mathbf{a}'\bar{D}_i \pm \sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)}\sqrt{\frac{\mathbf{a}'S_D^{-1}\mathbf{a}}{n}}, \forall \mathbf{a} \in R^p$$

- $\delta_1, \ldots, \delta_p$ 的 $1-\alpha$ 同时置信区间为

$$\bar{D}_i \pm \sqrt{\frac{p(n-1)}{n-p}F_{p,n-p}(\alpha)}\sqrt{\frac{s_{D,i}^2}{n}}, \quad i = 1, \ldots, p$$

其中 $s_{D,i}^2$ 表示矩阵 S_D 的 ii 对角元.

- $\delta_1, \ldots, \delta_p$ 的 $1-\alpha$ Bonferroni 同时置信区间为

$$\bar{D}_i \pm t_{n-1}(\frac{\alpha}{2p})\sqrt{\frac{s_{D,i}^2}{n}}, \quad i = 1, \ldots, p$$

- 当样本量 n 和 $n-p$ 都充分大时, 则

$$T_\delta^2 \rightsquigarrow \chi_p^2$$

  – 故渐近水平 α 检验的拒绝域为

$$T_0^2 = n\bar{D}'S_D^{-1}\bar{D} > \chi_p^2(\alpha)$$

  – δ 的渐近 $1-\alpha$ 置信域为

$$\left\{\delta : n(\bar{D}-\delta)'S_D^{-1}(\bar{D}-\delta) \le \chi_p^2(\alpha)\right\}$$

  – $\delta_1, \ldots, \delta_p$ 的渐近 $1-\alpha$ Bonferroni 同时置信区间为

$$\bar{D}_i \pm z(\frac{\alpha}{2p})\sqrt{\frac{s_{D,i}^2}{n}}, \quad i = 1, \ldots, p$$

重复测量下比较多个处理

- 一元成对 t 检验的另一个推广场合: **对一元响应变量的 q 种处理进行比较**.

- 在一段时间内, 每个个体或者试验单元被重复安排到 q 种处理中, 每次只接受一种处理. 因此, 第 j 个个体的观测记为

$$X_j = \begin{pmatrix} X_{j1} \\ \vdots \\ X_{jq} \end{pmatrix}, j = 1, \ldots, n.$$

  其中 X_{ji} 表示第 j 个个体的第 i 个处理下的值.

- 重复测量 —来源于是对同一个体施加全部 q 种处理.

- 实际中经常感兴趣的是 q 种处理的平均效应是否存在差异. 例

如考虑均值 $\mu = EX_j$ 的分量的对比 (contrast):

$$\begin{bmatrix} \mu_1 - \mu_2 \\ \mu_1 - \mu_3 \\ \vdots \\ \mu_1 - \mu_q \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_q \end{bmatrix} = C_1 \mu$$

或者

$$\begin{bmatrix} \mu_2 - \mu_1 \\ \mu_3 - \mu_2 \\ \vdots \\ \mu_q - \mu_{q-1} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_q \end{bmatrix} = C_2 \mu$$

$(q-1) \times q$ 矩阵 C_1 和 C_2 都称为对比矩阵 (行线性无关, 每个行都是一个对比向量).

- 假设$H_0: q$ **种处理的平均效应不存在差异**等价于 $H_0: C\mu = 0$ **对任何对比矩阵** C. 因此, 当 q 种处理的平均效应不存在差异时候有 $C_1\mu = C_2\mu = 0$.

- 若假设 $X_1, \ldots, X_n$ $i.i.d \sim N_q(\mu, \Sigma)$, 则对照 $CX_j \sim N_{q-1}(C\mu, C\Sigma C')$, 因此假设 $H_0: C\mu = 0 \leftrightarrow H_1: C\mu \neq 0$ 的 Hotelling T^2 检验拒绝域为

$$T^2 = n(C\bar{\mathbf{x}})'(CSC')^{-1}(C\bar{\mathbf{x}}) > \frac{(q-1)(n-1)}{n-q+1} F_{q-1,n-q+1}(\alpha)$$

其中 $\bar{\mathbf{x}}$ 和 S 为基于 $X_j's$ 的样本均值和样本协方差矩阵.

- $C\mu$ 的 $1-\alpha$ 置信域为

$$n(C\bar{\mathbf{x}} - C\mu)'(CSC')^{-1}(C\bar{\mathbf{x}} - C\mu) \leq \frac{(q-1)(n-1)}{n-q+1} F_{q-1,n-q+1}(\alpha)$$

1.2 两总体的均值比较

- 随机化试验: n_1 个个体随机地被分配到处理 1(或者是从总体 1 中随机抽取的个体), n_2 个个体被随机的分配到处理 2(或者是从总体 2 中随机抽取的个体)

- 每个个体的 p 个性状 (变量) 被测量

- 总体 k 的样本数据向量记为

$$X_{kj} = \begin{pmatrix} x_{kj1} \\ x_{kj2} \\ \vdots \\ x_{kjp} \end{pmatrix}, j = 1, \ldots, n_k, k = 1, 2.$$

- 记 μ_k 表示总体 k 的期望 $(k = 1, 2)$, 则感兴趣假设

$$H_0 : \mu_1 - \mu_2 = \delta_0 \leftrightarrow H_1 : \mu_1 - \mu_2 \neq \delta_0 \qquad \delta_0\text{已知}$$

- 记两组样本的样本均值和样本方差分别为

$$\bar{\mathbf{x}}_k = \frac{1}{n_k}\sum_{j=1}^{n_k} X_{1j}, \qquad S_k = \frac{1}{n_k - 1}\sum_{j=1}^{n_k}(X_{kj} - \bar{\mathbf{x}}_k)(X_{kj} - \bar{\mathbf{x}}_k)'$$

其中 $k = 1, 2.$

假设

- $X_{k1}, \ldots, X_{kn_k} i.i.d \sim (\mu_k, \Sigma_k)$, 其中 $k = 1, 2$. μ_k 为 p 维向量, Σ_k 为 $p \times p$ 正定矩阵.
- $X_{11}, \ldots, X_{1n_1}$ 和 $X_{21}, \ldots, X_{2n_2}$ 相互独立

大样本场合

对感兴趣的参数 $\theta = \mu_1 - \mu_2$, 其一个无偏估计为 $\hat{\theta} = \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2$, 显然 $\hat{\theta} \sim (\theta, \frac{1}{n_1}\Sigma_1 + \frac{1}{n_2}\Sigma_2)$. 而 S_1 和 S_2 分别为 Σ_1 和 Σ_2 的无偏估计, 于是由大样本理论知当 $n_1, n_2 \to \infty$ 时候,

$$(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \theta)' \left[\frac{1}{n_1}S_2 + \frac{1}{n_2}S_2\right]^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \theta) \to \chi_p^2$$

- 从而此时可以得到一个渐近水平 α 检验拒绝域:

$$(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \delta_0)' \left[\frac{1}{n_1}S_2 + \frac{1}{n_2}S_2\right]^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \delta_0) > \chi_p^2(\alpha)$$

- 等价地, $\theta = \mu_1 - \mu_2$ 的一个渐近 $1-\alpha$ 置信域为

$$\left\{\theta : (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \theta)' \left[\frac{1}{n_1}S_2 + \frac{1}{n_2}S_2\right]^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \theta) \le \chi_p^2(\alpha)\right\}$$

当样本量较小时候, 为得到合适的检验, 我们需要对分布作进一步的假设:

正态总体同方差 -协方差假设

- $X_{k1},\ldots,X_{kn_k} i.i.d \sim N_p(\mu_k,\Sigma)$, 其中 $k=1,2$.
- $X_{11},\ldots,X_{1n_1}$ 和 $X_{21},\ldots,X_{2n_2}$ 相互独立

容易得到

- $\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2$ 为 $\theta=\mu_1-\mu_2$ 的极大似然估计, Σ 的 (修正无偏的) 估计为

$$S_{pooled}=\frac{1}{n_1+n_2-2}\Big[(n_1-1)S_1+(n_2-1)S_2\Big]$$

- $\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2\sim N_p(\mu_1-\mu_2,(\frac{1}{n_1}+\frac{1}{n_2})\Sigma)$
- $\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2$ 和 S_{pooled} 相互独立.

因此

定理 1. $X_{k1},\ldots,X_{kn_k} i.i.d \sim N_p(\mu_k,\Sigma)$, 其中 $k=1,2.$, $\bar{\mathbf{x}}_1,\bar{\mathbf{x}}_2$ 和 S_{pooled} 如上定义, 则

$$(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\theta)'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_{pooled}\right]^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\theta)$$
$$\sim\frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{p,n_1+n_2-p-1}$$

证明. 记 $n=n_1+n_2$, 由前面的讨论知存在 $z,u_1,\ldots,u_{n-2}\ i.i.d\sim N_p(0,I_p)$, 使得

$$z\overset{d}{=}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\Sigma^{-1/2}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\theta)$$

$$(n-2)\Sigma^{-1/2}S_{pooled}\Sigma^{-1/2}\overset{d}{=}\sum_{i=1}^{n-2}u_iu_i'$$

因此

$$(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\theta)'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_{pooled}\right]^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\theta)\overset{d}{=}(n-2)z'\left(\sum_{i=1}^{n-2}u_iu_i'\right)^{-1}z$$

从而再类似于上一讲定理 1 的证明可证. □

由此定理, 对假设

$$H_0:\mu_1-\mu_2=\delta_0 \leftrightarrow H_1:\mu_1-\mu_2\neq\delta_0 \qquad \delta_0\text{已知}$$

- 在正态同方差 -协方差的假设下得到其一个精确的水平 α 检验拒绝域为

$$(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\delta_0)'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_{pooled}\right]^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\delta_0)$$
$$>\frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{p,n_1+n_2-p-1}(\alpha).$$

- 等价地, $\theta=\mu_1-\mu_2$ 的一个 $1-\alpha$ 置信域为

$$(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\theta)'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_{pooled}\right]^{-1}(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2-\theta) \le \frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{p,n_1+n_2-p-1}(\alpha).$$

同时置信区间

定理 2. 记 $c^2=\frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{p,n_1+n_2-p-1}(\alpha)$, 则

$$P\left(\mathbf{a}'(\mu_1-\mu_2)\in\mathbf{a}'(\bar{\mathbf{x}}_1-\bar{\mathbf{x}}_2)\pm c\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})\mathbf{a}'S_{pooled}\mathbf{a}}, \forall\mathbf{a}\ne 0\in R^p\right)=1-\alpha.$$

特别,

$$P\left(\mu_{1i}-\mu_{2i})\in(\bar{\mathbf{x}}_{1i}-\bar{\mathbf{x}}_{2i})\pm c\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})S_{ii,pooled}}, i=1,\ldots,p\right)=1-\alpha.$$

证明. 注意到 $\mathbf{a}'X_{kj} \sim N_1(\mathbf{a}'\mu_k, \mathbf{a}'\Sigma\mathbf{a}), j = 1, \ldots, n_k; k = 1, 2.$ 且相互独立, 故可利用一元两样本 t 统计量和 Cauchy-Schwarz 不等式 (2.7 节) 易证. $\square$

正态总体异方差假设

- $X_{k1}, \ldots, X_{kn_k} i.i.d \sim N_p(\mu_k, \Sigma_k)$, 其中 $k = 1, 2$. $\Sigma_1 \neq \Sigma_2$.
- $X_{11}, \ldots, X_{1n_1}$ 和 $X_{21}, \ldots, X_{2n_2}$ 相互独立

此时,$\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2$ 的协方差的自然估计为 $\frac{1}{n_1}S_1 + \frac{1}{n_2}S_2$, 统计量

$$\tilde{T}^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \theta)'\left[\frac{1}{n_1}S_2 + \frac{1}{n_2}S_2\right]^{-1}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2 - \theta)$$

的精确分布依赖于 Σ_1 和 Σ_2, 因此不能用来获得检验的临界值或者置信域的边界值.

解决方法

- 若样本量 n_1, n_2 都充分大, 则可以使用大样本理论方法.
- 若样本量不大, 但两总体均服从多元正态, 则 $\tilde{T}^2$ 的分布可以近似

$$\tilde{T}^2 \rightsquigarrow \frac{vp}{v-p+1} F_{p,v-p+1}$$

其中 v 通过下式估计

$$v = \frac{p+p^2}{\frac{a_1}{n_1} + \frac{a_2}{n_2}}$$

其中

$$a_i = tr\left[\left(\frac{1}{n_i} S_i \left(\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right)^{-1}\right)^2\right] + \left(tr\left[\frac{1}{n_i} S_i \left(\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right)^{-1}\right]\right)^2, i=1,2.$$

据此, 可得渐近水平 α 检验拒绝域

$$\tilde{T}^2 > \frac{vp}{v-p+1} F_{p,v-p+1}(\alpha)$$

1.3 方差 -协方差的检验

怎样判断多个总体的方差相等性假设是合理的?

- 对两总体, 一个粗放的经验法则: 如果存在 i 使得 $\sigma_{1,ii} > 4\sigma_{2,ii}$ 或者 $\sigma_{2,ii} > 4\sigma_{1,ii}$ 成立, 则很有可能 $\Sigma_1 \neq \Sigma_2$.

其他的定量检验方法多基于似然比检验方法.

单正态总体下检验 $\Sigma = \Sigma_0$

此时样本 $X_1, \ldots, X_n\ i.i.d \sim N_p(\mu, \Sigma)$, $\Sigma_0 > 0$ 已知, 考虑假设

$$H_0 : \Sigma = \Sigma_0 \leftrightarrow H_1 : \Sigma \neq \Sigma_0$$

则似然函数为

$$L(\theta, \Sigma) = (2\pi)^{-n/2}|\Sigma|^{-n/2} exp\Big[-\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)'\Sigma^{-1}(X_i-\mu)\Big]$$

从而似然比检验统计量为

$$\lambda_1 = \frac{\max\limits_{\mu} L(\mu, \Sigma_0)}{\max\limits_{\mu,\Sigma} L(\mu, \Sigma)} = (\frac{e}{n})^{np/2}|B\Sigma_0^{-1}|^{n/2}e^{-\frac{1}{2}tr(B\Sigma_0^{-1})}$$

其中 $B = \sum_{i=1}^{n}(X_i - \bar{\mathbf{x}})(X_i - \bar{\mathbf{x}})'$.

Sugiura and Nagao (1968) 证明了基于 λ_1 的检验是有偏的, 但是可以修正为无偏的:

$$\lambda_1^* = (\frac{e}{n-1})^{(n-1)p/2}|S\Sigma_0^{-1}|^{(n-1)/2}e^{-\frac{n-1}{2}tr(S\Sigma_0^{-1})}$$

注意到

$$-2log\lambda_1^* = (n-1)[tr(S\Sigma_0^{-1}) - log|S| + log|\Sigma_0| - p] := M$$

当 $n \to \infty$ 时, $M \to \chi^2_{p(p+1)/2}$. 从而可得渐近水平 α 检验拒绝域

$$M > \chi^2_{p(p+1)/2}(\alpha)$$

k 个正态总体的方差 -协方差齐性检验

假设样本 $X_1^{(j)}, \ldots, X_{n_j}^{(j)}\ i.i.d \sim N_p(\mu^{(j)}, \Sigma^{(j)}),\ j = 1, \ldots, k.$ $n = n_1 + \cdots + n_k$. 且 k 个总体的样本之间相互独立. 考虑假设

$$H_0 : \Sigma^{(1)} = \cdots = \Sigma^{(k)} \leftrightarrow H_1 : \exists j \neq l, s.t.\ \ \Sigma^{(j)} \neq \Sigma^{(l)}$$

使用似然比检验方法, 可以得到似然比检验统计量

$$\Lambda = \frac{\prod_{j=1}^k |B_j/n_j|^{n_j/2}}{|B_p/n|^{n/2}} = \prod_{j=1}^k \left(\frac{|B_j/n_j|}{|B_p/n|} \right)^{n_j/2}$$

其中 $B_j = \sum_{i=1}^{n_j} (X_i^{(j)} - \bar{\mathbf{x}}_i^{(j)})(X_i^{(j)} - \bar{\mathbf{x}}_i^{(j)})'$ 和 $B_p = \sum_{j=1}^k B_j$.

使用无偏估计 S_j 和 S_{pooled} 代替上式中的 B_j 和 B_p, 得到

$$\Lambda^* = \frac{\prod_{j=1}^k |S_j|^{n_j/2}}{|S_{pooled}|^{(n-k)/2}}$$

其中 $S_j = \frac{1}{n_j - 1} B_j$, $S_{pooled} = \frac{1}{n-k} \sum_{j=1}^k (n_j - 1) S_j$.

令

$$M = -2log\Lambda^* = (n-k)log|S_{pooled}| - \sum_{j=1}^{k}(n_j - 1)log|S_j|$$

由似然比检验方法的性质, 当样本量充分大时候 M 的分布在零假设下近似 $\chi^2_{p(p+1)(k-1)/2}$. 因此可得渐近水平 α 检验的拒绝域

$$M > \chi^2_{p(p+1)(k-1)/2}(\alpha).$$

$k=2$ 时候 Λ^* 在 H_0 下的精确分布可以得到, 但对 $k>2$ 则不可得.

对样本量中等或较小时候, 各种分布修正的方法被提出.

(1) **Bartlett test** (Bartlett, 1937) 在 $p=1$ 时, Bartlett 统计量定义为

$$X^2 = \frac{M}{d}$$

其中 $d = 1 + \frac{1}{3(k-1)}\left(\sum_j \frac{1}{n_j-1} - \frac{1}{n-k}\right)$, 在 H_0 下,

$$X^2 \rightsquigarrow \chi^2_{k-1}$$

从而水平 α 检验的拒绝域为 $X^2 > \chi^2_{k-1}(\alpha)$.

- Pitman (1939) 表明 Bartlett 检验是无偏的, 相合的.
- Box(1953) 表明 Bartlett 检验对正态性偏离非常敏感 (不稳健), 因此当零假设被拒绝时候, 我们不知道是因为异方差原因还是非正态的原因.

(2) **Box's M test/Bartlett-Box M test** (Box, 1949,1950) 考虑了多元场合下对 $-2log\Lambda^*$ 的分布进行卡方或 F 近似, 定义其 M 统计量为

$$M = -2log\Lambda^* = (n-k)log|S_{pooled}| - \sum_{j=1}^{k}(n_j-1)log|S_j|$$

然后证明了在 H_0 下,

$$C=(1-c)M \rightsquigarrow \chi_v^2, \quad v=p(p+1)(k-1)/2$$

其中

$$c=\left[\sum_j \frac{1}{n_j-1}-\frac{1}{n-k}\right]\frac{2p^2+3p-1}{6(p+1)(k-1)}$$

从而水平 α 检验的拒绝域为 $C>\chi_v^2(\alpha)$. 当 $n_j>20, k\le 5, p\le 5$ 时, C 逼近卡方分布的程度令人满意.

(3) **Box's M test/Bartlett-Box F test** 对较小的样本量, 使用 F 分布可以得到一个更好的逼近: 定义

$$c_2=\frac{(p-1)(p+2)}{6(k-1)}\left[\sum_j \frac{1}{(n_j-1)^2}-\frac{1}{(n-k)^2}\right]$$

$$v_2=\frac{v}{|c_2-c^2|}, a^+=\frac{v}{1-c-\frac{v}{v_2}}, a^-=\frac{v_2}{1-c+\frac{2}{v_2}}$$

定义 F 检验统计量

$$F = \begin{cases} \frac{M}{a^{+}}, & c_2 > c^2 \\ \frac{v_2 M}{v(a^{-} - M)}, & c_2 < c^2 \end{cases}$$

则 $F \rightsquigarrow F(v, v_2)$, 从而水平 α 检验的拒绝域为 $F > F_{v,v_2}(\alpha)$.

- **Box's M 检验对正态性偏离非常敏感!!** (Olsen, 1974)

- 一些稳健的方差 -协方差齐性检验方法已经被提出, Tiku and Singh (1985), O's Brien (1992), Anderson (2006) 等.