

多个正态总体均值相等的 VDR 检验

戴家佳^{①*}, 程维虎^②, 张忠占^②, 杨振海^②

① 贵州大学理学院, 贵阳 550025;

② 北京工业大学应用数理学院, 北京 100124

E-mail: sci.jjdai@gzu.edu.cn, chengweihu@bjut.edu.cn, zzhang@bjut.edu.cn, zhyang@bjut.edu.cn

收稿日期: 2013-05-27; 接受日期: 2014-06-25; * 通信作者
国家自然科学基金 (批准号: 11361015) 资助项目

摘要 本文将随机估计由一维参数扩展至多维参数, 基于随机估计的密度函数提出 VDR 检验. 在总体方差已知和未知的两种情形下, 本文讨论多个正态总体均值是否相同的 VDR 检验过程, 而且得到精确的检验. 单因素方差分析是 VDR 检验的特例. 模拟研究表明, VDR 检验是一个普遍适用的方法.

关键词 随机估计 置信分布 垂直密度表示

MSC (2010) 主题分类 62H12, 62H15

1 引言

参数推断是统计学的基本问题, 包含 3 种推断方式, 即经典频率统计推断、Bayes 统计推断和信仰推断 (fiducial inference). 信仰推断是基于似然函数的, 观测到样本后, 还不能准确确定未知参数 θ 的值, 而只能确定参数 θ 落在一个区间的可能性, 即概率. 若认为参数 θ 是随机变量, 则可以确定 θ 的分布或密度函数. Fisher^[1] 提出了信仰分布的初衷是在经典统计体系下挑战后验分布的概念. 信仰分布的定义如下:

设 $F(\cdot; \theta)$ 是随机变量 X 的分布函数, 其概率密度函数是 $f(\cdot; \theta)$. 给定 x , 信仰分布密度函数为

$$\text{fid}(\theta | x) = -\frac{\partial F}{\partial \theta}. \quad (1.1)$$

我们可以解释 (1.1) 如下:

- (1) 给定参数 θ , $V = F(X, \theta) \sim U(0, 1)$, $\forall \theta \in \Theta$;
- (2) 如果 X 用它的观察值 x 代替, 即固定 X , 则

$$V = F(x, \theta), \quad \text{且} \quad V \sim U(0, 1), \quad (1.2)$$

此时, θ 是 V 的函数, 而且也是随机变量, 其对应的密度函数是 $f(x; \theta) = \left| \frac{\partial F(x; \theta)}{\partial \theta} \right|$, 对常见分布族有 $\left| \frac{\partial F(x; \theta)}{\partial \theta} \right| = -\frac{\partial F(x; \theta)}{\partial \theta}$, 即 (1.1). 不过这里存在不足, 确定 V 的分布时, 参数 θ 是常数, 样本是随机的; 而计算信仰分布密度函数时, 样本是常数, θ 是随机的. 在同一过程中 θ 的解释不同, 这就是所谓 Fisher 疑惑.

英文引用格式: Dai J J, Cheng W H, Zhang Z Z, et al. VDR test for equality of means of multiple normally distributed populations (in Chinese). Sci Sin Math, 2014, 44: 1203–1224, doi: 10.1360/N012013-00091

事实上, 解决不一致性最简单的方法是在 (1.2) 中用随机变量 W 代替参数 θ , 用随机变量 W 估计或推断参数 θ , 既坚持了参数是未知常数的基本观点, 又可得出信仰推断. 设 $\mathbf{X} = (X_1, \dots, X_n)$ 是来自 $F(\cdot; \theta)$ 的样本, 可用参数 θ 的枢轴量 $h(\mathbf{X}; \theta)$ 代替 $F(\cdot; \theta)$, 且 W 代替 θ , (1.2) 成为

$$V = h(\mathbf{X}, W), \quad (1.3)$$

若 $P(a < W \leq b) \geq 1 - \alpha$, 则 $(a, b]$ 就是在经典统计意义下参数 θ 的一个置信水平为 $1 - \alpha$ 的置信区间.

考虑更一般的模型如下:

$$X \sim F(\cdot; \theta, \lambda), \quad \theta \in \Theta, \quad \lambda \in \Lambda,$$

其中 $F(\cdot; \theta, \lambda)$ 是分布函数, θ 和 λ 是未知参数, Θ 和 Λ 是参数空间, $\theta \in \Theta$ 是感兴趣的参数, 而 $\lambda \in \Lambda$ 是冗余参数. 设 $\mathbf{X} = (X_1, \dots, X_n)$ 为来自 $F(\cdot; \theta, \lambda)$ 的随机样本, $\mathbf{X} \in \mathcal{X}$, 其中 $\mathcal{X}$ 是样本空间. 基本问题是检验假设 $H_0: \theta = \theta_0$ 或者构造参数 θ 的置信区间. 解决上述问题的经典基本方法是枢轴量法.

设 $h(\mathbf{X}; \theta)$ 是参数 θ 的枢轴量, 即

$$h(\mathbf{X}; \theta) \sim F_h(\cdot), \quad \forall \theta \in \Theta,$$

其中 $F_h(\cdot)$ 是不依赖于任何未知参数的已知分布函数.

记 $Q_h(\alpha)$ 是 $F_h(\cdot)$ 的 α 分位点, 则

$$1 - \alpha_2 - \alpha_1 = P(Q_h(\alpha_1) < h(\mathbf{X}; \theta) \leq Q_h(1 - \alpha_2)) = P(\underline{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X}) < \theta \leq \bar{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})),$$

其中 $\underline{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})$ 和 $\bar{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})$ 满足:

(1) 若 $h(\mathbf{X}; \cdot)$ 是单调增的,

$$h(\mathbf{X}; \underline{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})) = Q_h(\alpha_1), \quad h(\mathbf{X}; \bar{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})) = Q_h(1 - \alpha_2); \quad (1.4a)$$

(2) 若 $h(\mathbf{X}; \cdot)$ 是单调减的,

$$h(\mathbf{X}; \underline{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})) = Q_h(1 - \alpha_2), \quad h(\mathbf{X}; \bar{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})) = Q_h(\alpha_1). \quad (1.4b)$$

$[\underline{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X}), \bar{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})]$ 是参数 θ 的置信度为 $1 - \alpha_1 - \alpha_2$ 的置信区间, $h(\mathbf{X}; \theta)$ 是随机区间 $[\underline{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X}), \bar{\theta}(\alpha_1, 1 - \alpha_2; \mathbf{X})]$ 的观测值.

引入置信分布 (confidence distribution, CD) 函数, 可以简化上述公式. 置信分布函数定义^[2,3] 为

$$H(\cdot; \mathbf{X}) = \begin{cases} F_h(h(\mathbf{X}; \cdot)), & h(\mathbf{X}, \cdot) \text{ 是单调增的,} \\ 1 - F_h(h(\mathbf{X}; \cdot)), & h(\mathbf{X}, \cdot) \text{ 是单调减的.} \end{cases} \quad (1.5)$$

$H(\cdot; \mathbf{X})$ 是以 $\mathbf{X}$ 为分布参数的参数空间 Θ 上的分布函数. 记 $Q_H(\alpha; \mathbf{X})$ 为 $H(\cdot; \mathbf{X})$ 的 α 分位点, 则 (1.4) 可简洁地表示为

$$\underline{\theta}(\alpha_1, \alpha_2; \mathbf{X}) = \underline{\theta}(\alpha_1; \mathbf{X}) = Q_H(\alpha_1; \mathbf{X}), \quad \bar{\theta}(\alpha_1, \alpha_2; \mathbf{X}) = \bar{\theta}(\alpha_2; \mathbf{X}) = Q_H(1 - \alpha_2; \mathbf{X}).$$

设随机变量 ξ 的分布函数是 $H(\cdot; \mathbf{X})$, 若

$$P(a < \xi \leq b) = H(b; \mathbf{X}) - H(a; \mathbf{X}) = \gamma,$$

则 $(a, b]$ 是参数 θ 的置信度为 γ 的置信区间. 称 ξ 为 θ 的随机化估计^[2], 简称随机估计. ξ 可由下式定义:

$$U = H(\xi; \mathbf{X}), \quad U \sim U(0, 1). \quad (1.6)$$

易见,

$$P(\xi \leq w) = P(U \leq H(w; \mathbf{X})) = H(w; \mathbf{X}).$$

若 $H(\cdot; \mathbf{X})$ 由 (1.5) 定义, 则由 (1.3) 定义的 W 和 ξ 有相同的分布函数. 因此, 由 W 确定的置信区间就是频率意义下的置信区间. 用随机估计 W 推断参数是频率统计的方法, 也容易推广到多维参数的情形.

本文通过推广多维随机估计, 提出 VDR 检验, 并讨论 VDR 检验的应用. 本文结构如下: 第 2 节基于随机估计的密度函数提出 VDR 检验; 第 3 节讨论多元正态分布参数的 VDR 检验; 第 4 节运用 VDR 检验讨论多个正态总体均值相等的检验; 第 5 节通过模拟研究, 指出所提出的 VDR 检验是可用的; 第 6 节总结本文内容.

2 多维参数的随机估计和 VDR 检验

随机估计也适用于多维参数, 其基本思想描述如下. 设总体分布族为

$$f(\mathbf{x}; \boldsymbol{\eta}, \boldsymbol{\lambda}), \quad \mathbf{x} \in \mathbb{R}^p, \quad \boldsymbol{\eta} = (\eta_1, \dots, \eta_l)^T \in \mathcal{N} \subseteq \mathbb{R}^l, \quad \boldsymbol{\lambda} \in \Lambda \subseteq \mathbb{R}^{s-l},$$

其中 $\boldsymbol{\eta} \in \mathcal{N}$ 是感兴趣的参数, $\boldsymbol{\lambda}$ 是冗余参数, $\mathbb{R}^p$ 为实数域 $\mathbb{R}$ 上 p 维向量空间, T 表示转置. 设抽自 $f(\cdot; \boldsymbol{\eta}, \boldsymbol{\lambda})$ 的随机样本和具体样本分别记作

$$\mathbb{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n), \quad \mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n),$$

其中 n 是样本容量, 样本空间记作 $\mathcal{X} = \prod_{i=1}^n \mathbb{R}^p$.

考虑假设

$$H_0: \boldsymbol{\eta} = \boldsymbol{\eta}_0 \leftrightarrow H_1: \boldsymbol{\eta} \neq \boldsymbol{\eta}_0. \quad (2.1)$$

用枢轴量法实现假设 (2.1) 的检验.

定义 1 设 $\mathbf{h}(\mathcal{X}; \boldsymbol{\eta})$ 是 $\mathcal{X} \times \mathcal{N} \rightarrow \mathcal{N}_h$ 的 l 维向量函数, 其中 $\mathcal{N}_h$ 是 $\mathbf{h}(\cdot; \cdot)$ 的值域空间. 若满足:

- (1) 给定样本 $\mathcal{X}$, $\mathbf{h}(\mathcal{X}; \cdot)$ 是 $\mathcal{N} \rightarrow \mathcal{N}_h$ 的 1-1 映射;
- (2) 给定参数 $\boldsymbol{\eta}$, $\mathbf{h}(\mathbb{X}; \boldsymbol{\eta})$ 的分布与参数 $\boldsymbol{\eta}$ 和 $\boldsymbol{\lambda}$ 无关,

则称 $\mathbf{h}(\cdot, \cdot)$ 是参数 $\boldsymbol{\eta}$ 的枢轴量.

设 $\mathbf{h}(\mathbb{X}; \boldsymbol{\eta}) \sim F_h(\cdot)$, $\forall \boldsymbol{\eta} \in \mathcal{N}$, $\boldsymbol{\lambda} \in \Lambda$, 其中 $F_h(\cdot)$ 是参数空间 $\mathcal{N}_h$ 上的分布函数. 事实上, 在不依赖于参数的意义下, $F_h(\cdot)$ 不含参数信息, 但是 $\mathbf{h}(\cdot; \cdot)$ 的自变量是样本和参数, 含参数的充分信息. $F_h(\cdot)$ 包含了分布族的结构信息, 是推断参数的依据.

设 $\mathcal{B}_h$ 是 $\mathcal{N}_h$ 的 Borel 域, P_h 是 $F_h(\cdot)$ 确定的可测空间 $(\mathcal{N}, \mathcal{B}_h)$ 上的概率测度, $(\mathcal{N}, \mathcal{B}_h, P_h)$ 是概率空间, 参数 $\boldsymbol{\eta}$ 的随机估计 $\mathbf{W}$ 可定义为

$$\mathbf{V} = \mathbf{h}(\mathcal{X}; \mathbf{W}), \quad (2.2)$$

$\mathbf{V}$ 是枢轴量在值域空间 $\mathcal{N}_h$ 上的随机向量, 其分布函数是 $F_h(\cdot)$. 称 $\mathbf{V}$ 为参数 $\boldsymbol{\eta}$ 的枢轴随机向量, 简称为枢轴向量, 一维参数称为枢轴变量. 显然, $\mathbf{W}$ 是参数空间 $\mathcal{N}$ 上的随机向量. 若 $\mathbf{W} = (W_1, \dots, W_l)^T$, 则 W_i 是 η_i 的随机估计, $i = 1, \dots, l$. $\mathbf{W}$ 的概率密度函数是

$$f_{\boldsymbol{\eta}}(\mathbf{v}; \mathcal{X}) = f_h(\mathbf{h}(\mathcal{X}; \mathbf{v})) \left| \det \left(\frac{\partial \mathbf{h}(\mathcal{X}; \mathbf{v})}{\partial \mathbf{v}} \right) \right|,$$

其中 $f_h(\cdot)$ 是 $F_h(\cdot)$ 的概率密度函数, 也是枢轴向量 $\mathbf{V}$ 的密度函数. $f_h(\cdot; \mathcal{X})$ 是值域空间 $\mathcal{N}_h$ 上的概率密度函数. Schweder 和 Hjort^[4] 将 $f_{\boldsymbol{\eta}}(\mathbf{v}; \mathcal{X})$ 称为收缩似然函数 (reduced likelihoods). 而且, $f_{\boldsymbol{\eta}}(\mathbf{v}; \mathcal{X})$ 是感兴趣的参数 $\boldsymbol{\eta}$ 的似然函数, 同时滤除了冗余参数 $\boldsymbol{\lambda}$ 的信息, 有类似于枢轴量的作用. 因此可用 $f_{\boldsymbol{\eta}}(\mathbf{v}; \mathcal{X})$ 估计参数和构造置信区间. 但是, 文献 [4] 仅论述了一维参数的收缩似然函数. 本文将推广至多维参数的情形.

为检验假设 (2.1), 令 $Z = f_{\boldsymbol{\eta}}(\mathbf{W}; \mathcal{X})$, 且称为检验随机变量. 由垂直密度表示 (vertical density representation, VDR) 理论^[5-9] 知, Z 的概率密度函数是

$$f_Z(v; \mathcal{X}) = -v \frac{dL_l(D_{[f_{\boldsymbol{\eta}}(\cdot; \mathcal{X})]}(v))}{dv},$$

其中 $D_{[f]}(v) = \{\mathbf{x} : f(\mathbf{x}) > v, \mathbf{x} \in \mathbb{R}^p\}$, $f(\cdot)$ 为定义 $\mathbb{R}^p$ 上的非负函数, 且 $L_l(\cdot)$ 是 $\mathbb{R}^p$ 上的 Lebesgue 测度.

进一步, 可以计算 Z 的 γ 分位点 $Q_Z(\gamma; \mathcal{X})$,

$$\begin{aligned} \gamma &= P(Z \leq Q_Z(\gamma; \mathcal{X})) = \int_0^{Q_Z(\gamma; \mathcal{X})} f_Z(v; \mathcal{X}) dv \\ &= \int_0^{Q_Z(\gamma; \mathcal{X})} L_l(D_{[f_{\boldsymbol{\eta}}(\cdot; \mathcal{X})]}(v)) dv - L_l(D_{[f_{\boldsymbol{\eta}}(\cdot; \mathcal{X})]}(Q_Z(\gamma; \mathcal{X}))) Q_Z(\gamma; \mathcal{X}). \end{aligned} \quad (2.3)$$

在显著水平 α 下, 检验假设 (2.1) 的 VDR 检验规则如下:

规则 1 若 $f_{\boldsymbol{\eta}}(\boldsymbol{\eta}_0; \mathcal{X}) < Q_Z(\alpha; \mathcal{X})$, 接受 H_0 ; 否则拒绝 H_0 .

事实上, 上述的 VDR 检验规则是极大似然原则的应用, 似然函数在 $\boldsymbol{\eta}_0$ 处的值越大越支持原假设 $H_0 : \boldsymbol{\eta} = \boldsymbol{\eta}_0$. 注意在很多情形下, $f_{\boldsymbol{\eta}}(\boldsymbol{\eta}_0; \mathcal{X})$ 就是经典意义下的检验统计量.

综合上述讨论, 构建 VDR 检验的步骤如下:

第 1 步 基于总体模型构造枢轴量 $\mathbf{h}(\mathcal{X}; \boldsymbol{\eta})$, 且 $\mathbf{h}(\mathbb{X}; \boldsymbol{\eta}) \sim F_h(\cdot)$, 其中 $F_h(\cdot)$ 是与未知参数无关的 l 维分布函数;

第 2 步 确定随机估计, 参数 $\boldsymbol{\eta}$ 的随机估计 $\mathbf{W}$ 由方程 $\mathbf{h}(\mathcal{X}; \mathbf{W}) = \mathbf{V}$, $\mathbf{V} \sim F_h(\cdot)$ 的解给出;

第 3 步 计算随机估计密度函数 $f_{\boldsymbol{\eta}}(\cdot; \mathcal{X})$,

$$f_{\boldsymbol{\eta}}(\mathbf{v}; \mathcal{X}) = |J| f_h(\mathbf{h}(\mathcal{X}; \mathbf{v})),$$

其中

$$J = \det \left(\frac{\partial \mathbf{h}(\mathcal{X}; \mathbf{v})}{\partial \mathbf{v}} \right), \quad f_h(\mathbf{v}) = f_h(v_1, \dots, v_l) = \frac{\partial F_h(\mathbf{v})}{\partial v_l \cdots \partial v_1};$$

第 4 步 按公式 (2.3) 计算临界值 $Q_Z(\alpha; \mathcal{X})$, 即 $P(Z \leq Q_Z(\alpha; \mathcal{X})) = \alpha$;

第 5 步 按检验规则确定检验结果: 若 $f_{\boldsymbol{\eta}}(\boldsymbol{\eta}_0; \mathcal{X}) \leq Q_Z(\alpha; \mathcal{X})$, 则拒绝接受原假设 $H_0 : \boldsymbol{\eta} = \boldsymbol{\eta}_0$.

下面考虑多元正态分布的参数 VDR 检验.

3 多元正态分布参数的 VDR 检验

设 $\mathbf{X}$ 是来自 p 元正态分布 $N_p(\boldsymbol{\mu}, \Sigma)$ 的随机样本, $\mathcal{X}$ 是具体样本, 样本容量为 n . 记

$$\bar{\mathbf{X}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i, \quad \hat{\Sigma} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^T.$$

考虑假设

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0 \leftrightarrow H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0. \quad (3.1)$$

假设 (3.1) 的检验, 分总体协方差阵已知和未知两种情形讨论.

3.1 协方差阵已知时的均值检验

当 Σ 已知时, 参数 $\boldsymbol{\mu}$ 的枢轴量是

$$h_1(\bar{\mathbf{X}}; \boldsymbol{\mu}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \Sigma^{-\frac{1}{2}} (\mathbf{X}_i - \boldsymbol{\mu}) = \sqrt{n} \Sigma^{-\frac{1}{2}} (\bar{\mathbf{X}} - \boldsymbol{\mu}).$$

$\boldsymbol{\mu}$ 的随机估计 $W_{\boldsymbol{\mu}}$ 由下式确定,

$$\sqrt{n} \Sigma^{-\frac{1}{2}} (\bar{\mathbf{x}} - W_{\boldsymbol{\mu}}) = \mathbf{V}, \quad \mathbf{V} \sim N_p(\mathbf{0}_p, I_p),$$

即

$$W_{\boldsymbol{\mu}} = \bar{\mathbf{x}} - \frac{1}{\sqrt{n}} \Sigma^{\frac{1}{2}} \mathbf{Z}, \quad \mathbf{Z} \sim N_p(\mathbf{0}_p, I_p). \quad (3.2)$$

可知 $W_{\boldsymbol{\mu}}$ 的密度函数为

$$f_{\boldsymbol{\mu}}(\mathbf{u}; \bar{\mathbf{x}}) = \frac{\sqrt{n}}{\sqrt{|\Sigma|}} \varphi_p(\sqrt{n} \Sigma^{-\frac{1}{2}} (\bar{\mathbf{x}} - \mathbf{u})), \quad \mathbf{u} \in \mathbb{R}^p,$$

其中 $\varphi_p(\cdot)$ 是 $N_p(\mathbf{0}_p, I_p)$ 的密度函数.

设 (3.1) 的检验随机变量是 $\mathbf{Z} = f_{\boldsymbol{\mu}}(W_{\boldsymbol{\mu}}; \bar{\mathbf{x}})$, 其 α 分位点记作 $Q_Z(\alpha, \bar{\mathbf{x}})$, 且

$$0 < Q_Z(\alpha, \bar{\mathbf{x}}) < \frac{\sqrt{n}}{\pi^{\frac{p}{2}} \sqrt{|\Sigma|}}.$$

于是有

$$\begin{aligned} \alpha &= P(\mathbf{Z} \leq Q_Z(\alpha, \bar{\mathbf{x}})) = P(f_{\boldsymbol{\mu}}(W_{\boldsymbol{\mu}}; \bar{\mathbf{x}}) \leq Q_Z(\alpha, \bar{\mathbf{x}})) \\ &= P\left(\varphi_p(\sqrt{n} \Sigma^{-\frac{1}{2}} (\bar{\mathbf{x}} - W_{\boldsymbol{\mu}})) \leq \frac{\sqrt{|\Sigma|}}{\sqrt{n}} Q_Z(\alpha, \bar{\mathbf{x}})\right) \\ &= P\left(\varphi_p(\mathbf{V}) \leq \frac{\sqrt{|\Sigma|}}{\sqrt{n}} Q_Z(\alpha, \bar{\mathbf{x}})\right) \\ &= P\left(\mathbf{V}^T \mathbf{V} > -2 \ln \left(Q_Z(\alpha, \bar{\mathbf{x}}) \frac{\pi^{\frac{p}{2}} \sqrt{|\Sigma|}}{\sqrt{n}} \right)\right) \\ &= P(\mathbf{V}^T \mathbf{V} > \chi_p^2(1 - \alpha)), \end{aligned}$$

$$\chi_p^2(1-\alpha) = -2 \ln \left(Q_Z(\alpha, \bar{\mathbf{x}}) \frac{\pi^{\frac{p}{2}} \sqrt{|\hat{\Sigma}|}}{\sqrt{n}} \right),$$

其中 $\chi_p^2(\cdot)$ 是自由度为 p 的 χ^2 分布. 进而得到检验假设 (3.1) 的 VDR 检验规则如下:

规则 2 给定显著水平 α , 当 $f_{\mu}(\mu_0; \bar{\mathbf{x}}) \leq Q_Z(\alpha, \bar{\mathbf{x}})$ 或者 $n(\bar{\mathbf{X}} - \mu_0)^T \Sigma^{-1}(\bar{\mathbf{X}} - \mu_0) > \chi_p^2(1-\alpha)$ 时, 拒绝接受 (3.1) 的原假设.

上述结果与似然比检验结果一致, 即在协方差阵已知的条件下, 检验假设 (3.1) 的 VDR 检验与经典结果是一致的.

接下来, 讨论总体协方差阵未知的情形下, 如何构造假设 (3.1) 的 VDR 检验.

3.2 协方差阵未知时的均值检验

当 Σ 未知时, 利用多元 t 分布构造均值参数 μ 的检验. 参数 μ 的枢轴量是

$$\mathbf{h}_1(\mathcal{X}; \mu) = \hat{\Sigma}^{-\frac{1}{2}} \sqrt{n}(\bar{\mathbf{x}} - \mu), \quad (3.3)$$

且

$$\mathbf{h}_1(\mathbb{X}; \mu) = \hat{\Sigma}^{-\frac{1}{2}} \sqrt{n}(\bar{\mathbf{X}} - \mu) \stackrel{d}{=} S^{-\frac{1}{2}}(\mathbb{Z}) \sqrt{n} \bar{\mathbf{Z}} \sim t_{(p, n-1)},$$

其中

$$\mathbb{Z} = (\mathbf{Z}_1, \dots, \mathbf{Z}_n); \quad \bar{\mathbf{Z}} = (\bar{\mathbf{X}} - \mu), \quad \text{且 } \mathbf{Z}_i \sim N_p(\mathbf{0}_p, I_p), \quad i = 1, \dots, n.$$

$\hat{\Sigma}$ 为 Σ 的估计, $\stackrel{d}{=}$ 表示两端变量有相同的分布.

令 $\mathbf{T}_{p, n-1} = S^{-\frac{1}{2}}(\mathbb{Z}) \sqrt{n} \bar{\mathbf{Z}}$ 是枢轴向量, 其概率密度函数为

$$f_{\mathbf{T}}(\mathbf{t}; n-1, p) = \frac{\Gamma(\frac{n+p-1}{2})}{((n-1)\pi)^{\frac{p}{2}} \Gamma(\frac{n-1}{2})} \left(1 + \frac{\sum_{i=1}^p t_i^2}{n-1} \right)^{-\frac{n-1+p}{2}},$$

设 $\mathcal{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ 是观测样本, 则 μ 的随机估计 $\mathbf{W}_{\mu}$ 由

$$\hat{\Sigma}^{-\frac{1}{2}} \sqrt{n}(\bar{\mathbf{x}} - \mathbf{W}_{\mu}) = \mathbf{T}_{p, n-1}$$

定义, 即

$$\mathbf{W}_{\mu} = \bar{\mathbf{x}} - \frac{\hat{\Sigma}^{\frac{1}{2}}}{\sqrt{n}} \mathbf{T}_{p, n-1}.$$

$\mathbf{W}_{\mu}$ 的密度函数为

$$f_{\mu}(\mathbf{u}; \bar{\mathbf{x}}, \hat{\Sigma}) = \sqrt{n} |\hat{\Sigma}|^{-\frac{1}{2}} f_{\mathbf{T}}(\hat{\Sigma}^{-\frac{1}{2}} \sqrt{n}(\bar{\mathbf{x}} - \mathbf{u}), n-1, p) = \sqrt{n} |\hat{\Sigma}|^{-\frac{1}{2}} f_{\mathbf{T}}(\mathbf{h}_1(\mathcal{X}, \mathbf{u}), n-1, p).$$

检验假设 (3.1) 的检验随机变量是

$$Z = \sqrt{n} |\hat{\Sigma}|^{-\frac{1}{2}} f_{\mathbf{T}}(\hat{\Sigma}^{-\frac{1}{2}} \sqrt{n}(\bar{\mathbf{x}} - \mathbf{W}_{\mu}), n-1, p).$$

当 Z 的值过小时, 拒绝原假设.

设 $Q_V(\alpha)$ 是 $V = f_{\mathbf{T}}(\mathbf{T}_{p, n-1}, n-1, p)$ 的 α 分位点. 因此, 检验假设 (3.1) 的规则如下:

规则 3 给定显著水平 α , 若 $f_{\mathbf{T}}(\mathbf{h}_1(\mathcal{X}, \mu_0), n-1, p) \leq Q_V(\alpha)$, 拒绝接受 (3.1) 的原假设; 否则接受原假设.

进一步计算 $Q_V(\alpha)$ 可知 (参见附录 A1), 在协方差阵未知的条件下, 检验假设 (3.1) 的 VDR 检验就是 Hotelling 检验.

综上所述, 对多元正态总体检验均值是否为已知向量的 VDR 检验与经典结果一致. 特别地, 当 $p = 1$ 时就是 t 检验.

4 多个正态总体期望相等的 VDR 检验

设 $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{in_i})^T$, 是来自 $N(\mu_i, \sigma_i^2)$ 的样本容量为 n_i 的样本, 其中 μ_i 和 σ_i^2 均未知, 且 $\mathbf{X}_i$ 相互独立, $i = 1, 2, \dots, p$, 记 $n = \sum_{i=1}^p n_i$.

考虑假设

$$H_0: \mu_1 = \mu_2 = \dots = \mu_p \leftrightarrow H_1: \mu_1, \dots, \mu_p \text{ 不全相等.} \quad (4.1)$$

问题 (4.1) 是 Behrens-Fisher 问题的直接推广, 将检验两个正态总体均值是否相等扩展为检验多个正态总体期望是否相等. 单因素方差分析是在各正态总体方差相等的前提下, 检验期望是否相等. 在不假定各总体方差相等条件下, 如何利用 VDR 检验来讨论假设 (4.1) 的检验是本节讨论的问题. 基于随机估计和 VDR 检验实现检验假设 (4.1) 的方法很多, 这里仅讨论有普遍意义的方法. 首先给出各总体方差相等时, 均值是否相等的 VDR 检验就是单因素方差分析, 然后讨论总体方差无约束时均值是否相等时的 VDR 检验. 两者处理问题的思路是一致的, 这种方法可以应用于解决其他问题.

设

$$\begin{aligned} \bar{x}_i &= \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}, \quad s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2, \quad i = 1, \dots, p, \\ \bar{\mathbf{x}} &= (\bar{x}_1, \dots, \bar{x}_p)^T, \quad \mathbf{s}_p^2 = ((n_1 - 1)s_1^2, \dots, (n_p - 1)s_p^2)^T, \\ \mathbf{n} &= (n_1, \dots, n_p)^T, \quad s_n^2 = \frac{1}{n - p} \sum_{i=1}^p (n_i - 1)s_i^2. \end{aligned}$$

4.1 等方差总体均值相等的 VDR 检验

假设 (4.1) 可以写为

$$H_0: \boldsymbol{\mu} \in \Theta_{\mathbf{1}_p} \leftrightarrow H_1: \boldsymbol{\mu} \notin \Theta_{\mathbf{1}_p}, \quad (4.2)$$

其中 $\boldsymbol{\mu} \in \Theta_{\mathbf{1}_p} = \{\mathbf{v}: \mathbf{v} = \mu \mathbf{1}_p, \mu \in \mathbb{R}\}$, $\mathbf{1}_p = (1, \dots, 1) \in \mathbb{R}^p$. 这意味着 $\boldsymbol{\mu}$ 受到参数 μ 约束, 称 μ 为约束参数, 即只要存在实数 μ 使得 $\boldsymbol{\mu} = \mu \mathbf{1}_p$, 假设 (4.1) 的原假设就成立, 这时约束参数 μ 就是各总体的共享均值.

设 μ 的随机估计为 W_μ , 且均值向量 $\boldsymbol{\mu}$ 的随机估计为 $\mathbf{W}_\mu$, 则有

$$\mathbf{W}_\delta = \mathbf{W}_\mu - W_\mu \mathbf{1}_p$$

是参数 $\boldsymbol{\delta} = \boldsymbol{\mu} - \mu \mathbf{1}_p$ 的随机估计. 当 (4.2) 的原假设成立时, $\boldsymbol{\delta} = \mathbf{0}_p$. 因此, 检验假设 (4.2) 等价于基于随机估计 $\mathbf{W}_\delta$ 检验假设

$$H_0: \boldsymbol{\delta} = \mathbf{0}_p \leftrightarrow H_1: \boldsymbol{\delta} \neq \mathbf{0}_p. \quad (4.3)$$

事实上, 基于随机估计 $\mathbf{W}_\mu$ 可以检验假设

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0 \leftrightarrow H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0. \quad (4.4)$$

假设 (4.3) 和 (4.4) 是同类假设, 后者更普遍些. 这里只需考虑后者的检验方法和理论, 但是两者的差异体现在随机估计对应的枢轴向量的自由度上. $\mathbf{W}_\mu$ 对应的枢轴向量的各分量是独立的, 或者说独立变化的分量个数是 p , 即枢轴向量的自由度是 p . 而 $\mathbf{W}_\delta$ 对应的枢轴向量的自由度是 $p-1$, 因为随机估计 W_μ 对应的枢轴变量是 $\mathbf{W}_\mu$ 对应的枢轴向量的函数, 形成了约束, 因此, 自由度减少 1.

4.1.1 假设 (4.4) 的 VDR 检验

当各总体方差相等时, 未知参数为 $(\boldsymbol{\mu}^T, \sigma^2)^T = (\mu_1, \dots, \mu_p, \sigma^2)^T$, 它的枢轴量为

$$\mathbf{h}(\bar{\mathbf{x}}, \mathbf{S}_p^2; \boldsymbol{\mu}, \sigma^2) = \begin{pmatrix} h(\bar{x}_1, s_1^2; \mu_1, \sigma^2) \\ \vdots \\ h(\bar{x}_p, s_p^2; \mu_p, \sigma^2) \\ h_1(s_n^2; \sigma^2) \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{n_1}(\bar{x}_1 - \mu_1)}{\sigma} \\ \vdots \\ \frac{\sqrt{n_p}(\bar{x}_p - \mu_p)}{\sigma} \\ \frac{(n-p)s_n^2}{\sigma^2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{\mathbf{n}}(\bar{\mathbf{x}} - \boldsymbol{\mu})}{\sigma} \\ h_1(s_n^2; \sigma^2) \end{pmatrix}. \quad (4.5)$$

显然,

$$\mathbf{h}(\bar{\mathbf{X}}, \mathbf{S}_p^2; \boldsymbol{\mu}, \sigma^2) \stackrel{d}{=} \mathbf{h}(\bar{\mathbf{X}}, \mathbf{S}_p^2; \mathbf{0}_p, 1) \stackrel{d}{=} \begin{pmatrix} \mathbf{Z} \\ \chi_{n-p}^2 \end{pmatrix} \sim \begin{pmatrix} N_p(\mathbf{0}_p, I_p) \\ \chi_{n-p}^2 \end{pmatrix},$$

其中 $\mathbf{Z} = (Z_1, \dots, Z_p)^T$ 与 χ_{n-p}^2 相互独立. 枢轴向量 $(\mathbf{Z}, \chi_{n-p}^2)$ 的联合密度函数为

$$f_{\mathbf{Z}, \chi_{n-p}^2}(\mathbf{u}, v) = \frac{1}{(\sqrt{2\pi})^p} \prod_{i=1}^p e^{-\frac{1}{2}u_i^2} \frac{1}{2^{\frac{n-p}{2}} \Gamma(\frac{n-p}{2})} v^{\frac{n-p}{2}-1} e^{-\frac{v}{2}} = \phi_p(\mathbf{u}) d\chi(v, n-p),$$

其中 $d\chi(v, n-p)$ 是自由度为 $n-p$ 的 χ^2 分布的密度函数.

设 $\boldsymbol{\mu}$ 和 σ^2 的随机估计分别为 $\mathbf{W}_\mu$ 和 W_{σ^2} , 则满足

$$\begin{pmatrix} \frac{\sqrt{\mathbf{n}}(\bar{\mathbf{x}} - \mathbf{W}_\mu)}{\sqrt{W_{\sigma^2}}} \\ \frac{(n-p)s_n^2}{W_{\sigma^2}} \end{pmatrix} = \begin{pmatrix} \mathbf{Z} \\ \chi_{n-p}^2 \end{pmatrix},$$

即

$$\begin{pmatrix} \mathbf{W}_\mu \\ W_{\sigma^2} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{x}} - \frac{s_n}{\sqrt{\mathbf{n}}} \frac{\mathbf{Z}}{\sqrt{\frac{\chi_{n-p}^2}{n-p}}} \\ \frac{(n-p)s_n^2}{\chi_{n-p}^2} \end{pmatrix}. \quad (4.6)$$

计算 $\mathbf{W}_\mu$ 的密度函数 (见附录 A2), 在各总体方差相等的条件下, 假设 (4.1) 的检验随机变量 Z 为

$$\begin{aligned} Z &= f_\mu(\mathbf{W}_\mu, \mathcal{X}) \\ &= \frac{\Gamma(\frac{n}{2})}{\pi^{\frac{p}{2}} \Gamma(\frac{n-p}{2})} \frac{\prod_{i=1}^p \sqrt{n_i}}{[(n-p)s_n^2]^{\frac{p}{2}}} \left(1 + \frac{\sum_{i=1}^p n_i (\bar{x}_i - W_i)^2}{(n-p)s_n^2} \right)^{-\frac{n}{2}} \\ &= C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2)) (1 + (\bar{\mathbf{x}} - \mathbf{W}_\mu)^T \Lambda(\mathbf{n}, s_n^2) (\bar{\mathbf{x}} - \mathbf{W}_\mu))^{-\frac{n}{2}}, \end{aligned}$$

其中

$$C(p, n) = \frac{\Gamma(\frac{n}{2})}{\pi^{\frac{p}{2}} \Gamma(\frac{n-p}{2})}, \quad \Lambda(\mathbf{n}, s_n^2) = \text{diag}\left(\frac{\mathbf{n}}{(n-p)s_n^2}\right).$$

记 Z 的 α 分位点为 $Q_Z(\alpha, s_n^2)$, 有

$$Q_V(\alpha) = \frac{Q_Z(\alpha, s_n^2)}{C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2))}, \quad 0 < Q_V(\alpha) < 1.$$

于是, 检验假设 (4.4) 的 VDR 检验规则如下:

规则 4 给定显著水平 α , 若

$$(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^T \Lambda(\mathbf{n}, s_n^2) (\bar{\mathbf{x}} - \boldsymbol{\mu}_0) > (Q_V(\alpha))^{-\frac{2}{n}} - 1,$$

则拒绝原假设.

接下来的问题就是关于 $Q_V(\alpha)$ 的计算, 附录 A3 中分别运用经典方法和 VDR 检验临界值的方法计算 $Q_V(\alpha)$, 结果显示, 两种方法得到的计算结果是一致的.

4.1.2 假设 (4.3) 的 VDR 检验

假设 (4.3) 为

$$H_0: \boldsymbol{\delta} = \mathbf{0}_p \leftrightarrow H_1: \boldsymbol{\delta} \neq \mathbf{0}_p.$$

先确定共享均值 μ 的随机估计. 已知 μ 的枢轴量为

$$\tilde{h}(\bar{\mathbf{x}}; \mu, \sigma^2) = \sqrt{n} \frac{\bar{x} - \mu}{\sigma} = \sum_{i=1}^p \sqrt{\frac{n_i}{n}} \frac{\sqrt{n_i}(\bar{x}_i - \mu)}{\sigma} = \sum_{i=1}^p \sqrt{\frac{n_i}{n}} \frac{\sqrt{n_i}(\bar{x}_i - \mu_i)}{\sigma},$$

且

$$\tilde{h}(\bar{\mathbf{X}}, \mu, \sigma^2) \stackrel{d}{=} \sqrt{n} \frac{\bar{X} - \mu}{\sigma} \stackrel{d}{=} \sum_{i=1}^p \sqrt{\frac{n_i}{n}} Z_i.$$

当 (4.1) 的原假设成立时, μ 的随机估计 W_μ 满足方程

$$\frac{\sqrt{n}(\bar{x} - W_\mu)}{\sqrt{W_{\sigma^2}}} = \sum_{i=1}^p \frac{\sqrt{n_i}}{\sqrt{n}} \frac{\sqrt{n_i}(\bar{x}_i - W_\mu)}{\sqrt{W_{\sigma^2}}} = \sum_{i=1}^p \sqrt{\frac{n_i}{n}} Z_i = \frac{\sqrt{\mathbf{n}^T}}{\sqrt{n}} \mathbf{Z}. \quad (4.7)$$

计算 (4.6) - $\frac{\sqrt{n}}{\sqrt{n}}$ (4.7) 可得

$$\frac{\sqrt{n}((\bar{\mathbf{x}} - \bar{x}\mathbf{1}_p) - (\mathbf{W}_\mu - W_\mu\mathbf{1}_p))}{\sqrt{W_{\sigma^2}}} = \mathbf{Z} - \frac{\sqrt{n}}{\sqrt{n}} \frac{\sqrt{\mathbf{n}^T}}{\sqrt{n}} \mathbf{Z}. \quad (4.8)$$

设 $\mathbf{A}$ 是 $p \times p$ 正交阵, 其中第 p 行元素为 $\frac{\sqrt{\mathbf{n}^T}}{\sqrt{n}}$, 记前 $p-1$ 行为 $\mathbf{A}_{p-1}$. $\mathbf{A}$ 左乘 (4.8) 两端得

$$\begin{pmatrix} \mathbf{A}_{p-1}(\bar{\mathbf{x}} - \bar{x}\mathbf{1}_p) - \mathbf{A}_{p-1}\mathbf{W}_\delta \\ 0 - \frac{\sqrt{\mathbf{n}^T}}{\sqrt{n}}\mathbf{W}_\delta \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{p-1}\mathbf{Z} \\ 0 \end{pmatrix},$$

则 $\mathbf{A}_{p-1}\mathbf{W}_\delta$ 是参数 $\mathbf{A}_{p-1}\boldsymbol{\delta}$ 的随机估计, 对应的枢轴向量是 $\mathbf{A}_{p-1}\mathbf{Z} \sim N_{p-1}(\mathbf{0}_{p-1}, I_{p-1})$. 假设 (4.3) 等价于

$$H_0: \mathbf{A}_{p-1}\boldsymbol{\delta} = \mathbf{0}_p \leftrightarrow H_1: \mathbf{A}_{p-1}\boldsymbol{\delta} \neq \mathbf{0}_p.$$

于是假设 (4.3) 归结为假设 (4.4) 的检验. 应用 4.1.1 小节的结果且注意到

$$(\bar{\mathbf{x}} - \bar{x}\mathbf{1}_p)^T (\bar{\mathbf{x}} - \bar{x}\mathbf{1}_p) = (\mathbf{A}_{p-1}(\bar{\mathbf{x}} - \bar{x}\mathbf{1}_p))^T \mathbf{A}_{p-1}(\bar{\mathbf{x}} - \bar{x}\mathbf{1}_p),$$

则等方差均值相等的 VDR 检验规则如下:

规则 5 给定显著水平 α , 若

$$\frac{\sum_{i=1}^p n_i (\bar{x}_i - \bar{x})^2}{(n-p)s_n^2} > \frac{p-1}{n-p} F_{p-1, n-p}(\alpha),$$

则拒绝接受假设 $H_0: \boldsymbol{\mu} \in \Theta_{\bar{\mathbf{1}}_p}$.

这也证明了当总体方差相等时, 均值相等的 VDR 检验就是单因素方差分析.

4.2 异方差多总体均值相等的 VDR 检验

本节讨论当总体方差未知且不相等时, 多总体均值检验的 VDR 检验的构造. 参数向量 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)^T$ 的枢轴量为

$$\mathbf{h}(\bar{\mathbf{x}}, \mathbf{s}_p^2; \boldsymbol{\mu}) = (h(\bar{x}_1, s_1^2; \mu_1), \dots, h(\bar{x}_p, s_p^2; \mu_p))^T = \frac{\sqrt{n}(\bar{\mathbf{x}} - \boldsymbol{\mu})}{\mathbf{s}_p^2},$$

且

$$\mathbf{h}(\bar{\mathbf{X}}, \mathbf{S}_p^2; \boldsymbol{\mu}) \stackrel{d}{=} \mathbf{h}(\bar{\mathbf{X}}, \mathbf{S}_p^2; \mathbf{0}_p) \stackrel{d}{=} \mathbf{T} \sim f_T(\mathbf{t}, \mathbf{n}),$$

其中 $\mathbf{T} = (T_1, \dots, T_p)$ 是枢轴向量, 其概率密度函数为

$$f_T(\mathbf{t}, \mathbf{n}) = \prod_{i=1}^p d_t(t_i, n_i - 1),$$

这里 $d_t(\cdot, p)$ 是自由度为 p 的 t 分布的密度函数.

设 $\boldsymbol{\mu}$ 的随机估计为 $\mathbf{W}_\mu = (W_{\mu_1}, \dots, W_{\mu_p})^T$, 且满足

$$\mathbf{h}(\bar{\mathbf{x}}, \mathbf{s}_p^2; \mathbf{W}_\mu) = \frac{\sqrt{n}(\bar{\mathbf{x}} - \mathbf{W}_\mu)}{\mathbf{s}_p} = \mathbf{T}, \quad (4.9)$$

即

$$\mathbf{W}_\mu = \bar{\mathbf{x}} - \frac{\mathbf{s}_p}{\sqrt{n}} \mathbf{T}.$$

可知共享均值 μ 的对数似然函数和似然方程分别为

$$\begin{aligned} l(\mu, \sigma_1^2, \dots, \sigma_p^2) &= \sum_{i=1}^p \sum_{j=1}^{n_i} \left(-\ln(\sigma_i^2) - \frac{1}{2\sigma_i^2} (x_{ij} - \mu)^2 \right), \\ \sum_{i=1}^p \sum_{j=1}^{n_i} \frac{1}{2\sigma_i^2} (x_{ij} - \mu) &= \sum_{i=1}^p \frac{n_i}{\sigma_i^2} (\bar{x}_i - \mu) = 0, \end{aligned}$$

同时参数 μ 的枢轴量为

$$\tilde{h}(\bar{\mathbf{x}}, \mathbf{s}_p^2; \mu) = \sum_{i=1}^p \frac{n_i}{s_i^2} (\bar{x}_i - \mu) = \sum_{i=1}^p \sqrt{\frac{n_i}{s_i^2}} \frac{\sqrt{n_i}(\bar{x}_i - \mu)}{s_i},$$

且

$$\tilde{h}(\bar{\mathbf{X}}, \mathbf{S}_p^2; \mu) \stackrel{d}{=} \tilde{h}(\bar{\mathbf{X}}, \mathbf{S}_p^2; \mathbf{0}_p) \stackrel{d}{=} \sum_{i=1}^p \sqrt{\frac{n_i}{s_i^2}} T_i = \left(\sqrt{\frac{\mathbf{n}}{\mathbf{s}_p^2}} \right)^T \mathbf{T},$$

其中 $(\sqrt{\frac{n}{s_p^2}})^T \mathbf{T}$ 是参数 μ 的枢轴变量. μ 的随机估计 W_μ 满足

$$\sum_{i=1}^p \frac{n_i}{s_i^2} (\bar{x}_i - W_\mu) = \left(\sqrt{\frac{n}{s_p^2}} \right)^T \mathbf{T},$$

即

$$\bar{x} - W_\mu = \frac{1}{\sqrt{\sum_{i=1}^p \frac{n_i}{s_i^2}}} \left(\frac{\sqrt{n/s_p^2}}{\sqrt{\sum_{i=1}^p \frac{n_i}{s_i^2}}} \right)^T \mathbf{T}, \quad (4.10)$$

其中

$$\tilde{x} = \sum_{i=1}^p \frac{\frac{n_i}{s_i^2}}{\sum_{i=1}^p \frac{n_i}{s_i^2}} \bar{x}_i.$$

计算 (4.9) 的第 p 行 $-\sqrt{\frac{n}{s_p^2}}$ (4.10) 得

$$\frac{\sqrt{n}((\bar{x} - \tilde{x}\mathbf{1}_p) - (\mathbf{W}_\mu - W_\mu\mathbf{1}_p))}{\sqrt{s_p^2}} = \begin{pmatrix} \frac{\sqrt{n_1}((\bar{x}_1 - \tilde{x}) - (W_{\mu_1} - W_\mu))}{s_1} \\ \vdots \\ \frac{\sqrt{n_p}((\bar{x}_p - \tilde{x}) - (W_{\mu_p} - W_\mu))}{s_p} \end{pmatrix} = (I_p - \mathbf{n}_1 \mathbf{n}_1^T) \mathbf{T} = \mathbf{T}^*, \quad (4.11)$$

其中

$$\mathbf{n}_1 = \left(\frac{\sqrt{\frac{n}{s_p^2}}}{\sqrt{\sum_{i=1}^p \frac{n_i}{s_i^2}}} \right).$$

由于 $\mathbf{W}_\delta = \mathbf{W}_\mu - W_\mu \mathbf{1}_p$ 是参数 $\delta = \mu - \mu \mathbf{1}_p$ 的随机估计, 而且假设 (4.1) 等价于

$$H_0: \delta = \mathbf{0}_p \leftrightarrow H_1: \delta \neq \mathbf{0}_p, \quad (4.12)$$

则 $\mathbf{W}_\delta$ 对应的枢轴向量 $\mathbf{T}^* = (I_p - \mathbf{n}_1 \mathbf{n}_1^T) \mathbf{T}$ 与 $\mathbf{n}_1$ 正交, 这蕴含着, $\mathbf{T}^* = (T_1^*, \dots, T_p^*)^T$ 是退化的, 取值于超平面

$$\mathbb{R}(\mathbf{n}_1) = \{\mathbf{t} : \mathbf{n}_1^T \mathbf{t} = 0, \mathbf{t} \in \mathbb{R}^p\}.$$

若 $\mathbf{t} \in \mathbb{R}(\mathbf{n}_1)$, 则

$$\mathbf{t} = (I_p - \mathbf{n}_1 \mathbf{n}_1^T)(\mathbf{t} + c\mathbf{n}_1), \quad \forall c \in \mathbb{R}.$$

故 $\mathbf{T}^*$ 的概率密度函数是

$$f_{\mathbf{T}^*}(\mathbf{t}, \mathbf{n}) = \int_{-\infty}^{\infty} f_{\mathbf{T}}(\mathbf{t} + c\mathbf{n}_1, \mathbf{n}) dc = \int_{-\infty}^{\infty} \prod_{i=1}^p d_t \left(t_i + c \sqrt{\frac{\frac{n_i}{s_i^2}}{\sum_{j=1}^p \frac{n_j}{s_j^2}}}, n_i - 1 \right) dc, \quad \forall \mathbf{t} \in \mathbb{R}(\mathbf{n}_1).$$

随机估计 $\mathbf{W}_\delta = \bar{x} - \tilde{x}\mathbf{1}_p - \frac{s_p}{\sqrt{n_p}} \mathbf{T}^*$ 的概率密度函数为

$$f_\delta(\mathbf{t}; \bar{x}, s_p^2) = \int_{-\infty}^{\infty} \prod_{i=1}^p \frac{\sqrt{n_i}}{s_i} d_t \left(\frac{\sqrt{n_i}}{s_i} ((\bar{x}_i - \tilde{x}) - t_i) + c \sqrt{\frac{\frac{n_i}{s_i^2}}{\sum_{j=1}^p \frac{n_j}{s_j^2}}}, n_i - 1 \right) dc, \quad \forall \mathbf{t} \in \mathbb{R}(\mathbf{n}_1). \quad (4.13)$$

记 VDR 检验随机变量 $Z_1 = f_\delta(\mathbf{W}_\delta; \bar{x}, s_p^2)$ 的 α 分位点为 $Q_{Z_1}(\alpha, \bar{x}, s_p^2)$, 则假设 (4.1) 的 VDR 检验规则如下:

规则 6 给定显著水平 α , 若

$$f_{T^*}(\mathbf{h}(\bar{\mathbf{x}}, \mathbf{s}_p^2; \mathbf{0}_p), \mathbf{n}) = \int_{-\infty}^{\infty} d_t \left(\frac{\sqrt{n_i}}{s_i} (\bar{x}_i - \tilde{x}) + c \sqrt{\frac{\frac{n_i}{s_i^2}}{\sum_{j=1}^p \frac{n_j}{s_j^2}}}, n_i - 1 \right) dc \leq Q_V(\alpha),$$

则拒绝接受各总体均值相等的原假设, 其中

$$Q_V(\alpha) = \frac{1}{\prod_{i=1}^p \frac{\sqrt{n_i}}{s_i}} Q_{Z_1}(\alpha, \bar{\mathbf{x}}, \mathbf{s}_p^2).$$

直接计算 $Q_{Z_1}(\alpha, \bar{\mathbf{x}}, \mathbf{s}_p^2)$ 较困难, 但是可以计算模拟值. 事实上, 只需得到 Z_1 的 N 个观测值便可得到其经验分布 $F_N(\cdot; \bar{\mathbf{x}}, \mathbf{s}_p^2)$, 而 Z_1 的一个观测值 z_1 获得过程如下:

- (1) 生成服从自由度为 $n_i - 1$ 的 t 分布随机数 $t_i, i = 1, \dots, p$, 记为 $\mathbf{t} = (t_1, \dots, t_p)^T$;
- (2) 计算

$$\mathbf{w} = \bar{\mathbf{x}} - \sqrt{\frac{\mathbf{s}_p^2}{n}} \mathbf{t}, \quad \mathbf{u} = \tilde{\mathbf{x}} - \frac{1}{\sqrt{\sum_{i=1}^n \frac{n_i}{s_i^2}}} \left(\frac{\sqrt{\frac{n}{\mathbf{s}_p^2}}}{\sqrt{\sum_{i=1}^n \frac{n_i}{s_i^2}}} \right)^T \mathbf{t}, \quad \mathbf{d} = \mathbf{w} - \mathbf{u} \mathbf{1}_p;$$

- (3) 按 (4.13) 计算 $z_1 = f_{\delta}(\mathbf{d}; \bar{\mathbf{x}}, \mathbf{s}_p^2)$.

4.3 近似 VDR 检验

在前面的讨论中, VDR 检验都是基于 t 分布来构造的. 而当自由度趋于无穷时, t 分布的极限分布是标准正态分布, 我们自然就想到了基于标准正态分布来构造 VDR 检验.

令 $\mathbf{Z} = (Z_1, \dots, Z_p)^T \sim N_p(\mathbf{0}_p, I_p)$, 随机向量 $\mathbf{W}^*$ 满足 $\mathbf{h}(\bar{\mathbf{x}}, \mathbf{s}; \mathbf{W}^*) = \mathbf{Z}$, 写成分量形式为

$$\begin{pmatrix} \frac{\sqrt{n_1}(\bar{x}_1 - W_1^*)}{s_1} \\ \vdots \\ \frac{\sqrt{n_p}(\bar{x}_p - W_p^*)}{s_p} \end{pmatrix} = \begin{pmatrix} Z_1 \\ \vdots \\ Z_p \end{pmatrix}. \quad (4.14)$$

称 $\mathbf{W}^*$ 为参数 $\boldsymbol{\mu}$ 的极限随机估计, 事实上, $\mathbf{W}^*$ 是确定随机估计的方程右端取极限而左端不动的条件下确定的随机估计. 显然, $\mathbf{W}$ 和 $\mathbf{W}^*$ 的分布接近, 或密度函数近似相等.

$\mathbf{W}^*$ 的密度函数为

$$f_{\boldsymbol{\mu}}^*(\mathbf{u}) = \phi_p(\mathbf{h}(\bar{\mathbf{x}}, \mathbf{s}^2; \mathbf{u})) = \prod_{i=1}^p \frac{\sqrt{n_i}}{s_i \sqrt{2\pi}} e^{-\frac{1}{2} \frac{n_i(\bar{x}_i - u_i)^2}{s_i^2}},$$

其中 $\phi_p(\cdot)$ 是 p 维标准正态分布密度函数.

为检验假设 (4.1), 令 $Z^* = f_{\boldsymbol{\mu}}^*(\mathbf{W}^*)$, 则称 Z^* 为近似检验随机变量, 其对应的 α 分位点记作 $Q_{Z^*}(\alpha; \mathbf{s}_p^2)$, 则

$$\begin{aligned} \alpha &= P(Z^* \leq Q_{Z^*}(\alpha; \mathbf{s}_p^2)) = P\left(\phi_p(\mathbf{h}(\bar{\mathbf{x}}, \mathbf{s}^2, \mathbf{W}^*)) \leq Q_{Z^*}(\alpha; \mathbf{s}_p^2) \prod_{i=1}^p \frac{s_i}{\sqrt{n_i}}\right) \\ &= P(\phi_p(\mathbf{Z}) \leq Q_{V^*}(\alpha)) \end{aligned}$$

$$\begin{aligned}
&= P\left(\sum_{i=1}^p Z_i^2 > -\ln((\sqrt{2\pi})^p Q_{V^*}(\alpha))\right) \\
&= P(\chi_p^2 > \chi_p^2(1-\alpha)),
\end{aligned}$$

因此,

$$-\ln((\sqrt{2\pi})^p Q_{V^*}(\alpha)) = \chi_p^2(1-\alpha), \quad Q_{Z^*}(\alpha; \mathbf{s}_p^2) = \prod_{i=1}^p \frac{\sqrt{n_i}}{s_i} Q_{V^*}(\alpha).$$

于是, 假设 (4.1) 的近似 VDR 检验规则如下:

规则 7 给定显著水平 α , 若 $\sum_{i=1}^p Z_i^2 = \sum_{i=1}^p \frac{n_i(\bar{x}_i - \mu_{0i})^2}{s_i^2} > \chi_p^2(1-\alpha)$, 则拒绝原假设 $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$.

直观地看, 近似 VDR 检验规则恰是 $\lim_{n \rightarrow \infty} P(\frac{\sqrt{n_i}(\bar{x}_i - \mu_{0i})}{s_i} \leq a) = \Phi(a)$ 的直接结果, 这从一个侧面说明 VDR 的实际意义.

另一方面, 考虑共享均值 μ 的随机估计. μ 的极限随机估计 W_μ^* 满足

$$\sum_{i=1}^p \frac{n_i}{s_i^2} (\bar{x}_i - W_\mu^*) = \left(\sqrt{\frac{\mathbf{n}}{\mathbf{s}_p^2}}\right)^T \mathbf{Z}. \quad (4.15)$$

计算 (4.14) 的第 i 行 $-\sqrt{\frac{n_i}{s_i^2}}(4.15)$, $i = 1, \dots, p$, 并写成向量形式有

$$\begin{pmatrix} \frac{\sqrt{n_1}[(\bar{x}_1 - \tilde{x}) - (W_1^* - W_\mu^*)]}{s_1} \\ \vdots \\ \frac{\sqrt{n_p}[(\bar{x}_p - \tilde{x}) - (W_p^* - W_\mu^*)]}{s_p} \end{pmatrix} = \left(I_p - \frac{\sqrt{\frac{\mathbf{n}}{\mathbf{s}_n^2}}}{\sqrt{\sum_{i=1}^p \frac{n_i}{s_i^2}}} \left(\frac{\sqrt{\frac{\mathbf{n}}{\mathbf{s}_n^2}}}{\sqrt{\sum_{i=1}^p \frac{n_i}{s_i^2}}}\right)^T\right) \mathbf{Z}. \quad (4.16)$$

归结为 (4.14) 的形式, 不难得到检验假设 (4.1) 的近似 VDR 检验规则如下:

规则 8 给定显著水平 α , 若 $\sum_{i=1}^p \frac{n_i}{s_i^2} (\bar{x}_i - \tilde{x})^2 \leq \chi_{p-1}^2(1-\alpha)$, 则接受假设 (4.1) 的原假设; 否则拒绝接受原假设.

事实上, $\sum_{i=1}^p \frac{n_i}{s_i^2} (\bar{x}_i - \tilde{x})^2$ 是经典意义下假设 (4.1) 的检验统计量, 其渐近分布为自由度为 $p-1$ 的 χ^2 分布. 换言之, 近似 VDR 检验随机变量渐近于自由度为 $p-1$ 的 χ^2 分布.

5 模拟研究

在第 4 节中, 理论上, 我们构造了多个正态总体均值相等的 VDR 检验, 并指出 VDR 检验是一种普遍适用的方法, 其特例是经典的检验方法. 本节将通过模拟研究, 进一步说明 VDR 检验的实用性.

首先给出三个正态总体均值是否相等的 VDR 检验. 模拟计算过程如下:

第 1 步 给定样本容量 $\mathbf{n} = (n_1, n_2, n_3)^T$, 均值和标准差分别为 $D = (d_1, d_2, d_3)^T$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)$.

第 2 步 (a) 从正态分布 $N(\mu_i, d_i^2)$, $i = 1, 2, 3$ 中各自抽取一个样本, 记为

$$\mathbf{x}_1 = (x_{11}, x_{12}, \dots, x_{1n_1}), \quad \mathbf{x}_2 = (x_{11}, x_{12}, \dots, x_{1n_2}), \quad \mathbf{x}_3 = (x_{11}, x_{12}, \dots, x_{1n_3});$$

(b) 计算 $z_1 = f_{T^*}(\mathbf{h}(\bar{\mathbf{x}}, \mathbf{s}_p^2; \mathbf{0}_p), \mathbf{n})$;

(c) 基于 $\mathbf{n}_1$ 和 $\mathbf{n}$ 模拟出分位点 $Q_V(\alpha)$;

(d) 计算 $r = z_1 - Q_V(\alpha)$.

第 3 步 重复第 2 步的过程 1000 次, 将第 i 次重复计算出的 r 记作 r_i , 1000 次重复结果记为 $\boldsymbol{r} = (r_1, \dots, r_{1000})^T$.

第 4 步 改变初值 $\boldsymbol{n}, D$ 和 $\boldsymbol{\mu}$ 的值, 重复第 2 和 3 步.

表 1 列出了三个正态总体均值是否相等的 VDR 检验与单因素方差分析的功效比较. 模拟重复 1000 次. 模拟结果表明, 在正态总体方差相等的前提下, VDR 检验的功效比单因素方差分析的模拟功效略低, 而且随着样本容量的增加, VDR 检验的功效与单因素方差分析的功效差距减小. 当原假设不真时, 两个检验的模拟功效几乎相等. 例如, 正态总体样本容量分别为 100、150 和 300, 标准差为 4, 均值向量为 $\boldsymbol{\mu} = (0, 0.5\mu_3, \mu_3)$, 其中 μ_3 分别取 0.0、0.25、0.5、0.75、1.0、1.5 和 2.0, 显著水平 $\alpha = 0.05$ 时, VDR 检验的模拟功效分别为 0.053、0.063、0.138、0.290、0.480、0.847 和 0.982, 单因素方差分析的模拟功效分别为 0.053、0.064、0.158、0.301、0.511、0.859 和 0.985, 两者的模拟功效差别很小. 但是 VDR 检验的优点在于不需要进行方差齐性检验, 可以直接使用, 而且功效与单因素方差分析相差不多.

进一步, 我们在正态总体方差不等的情形下, 我们将 VDR 检验分别与文献 [10, 11] 中的方法进行比较. 文献 [10] 中的检验方法 (简记为 ZHONG 方法) 要求各总体的样本容量相等. 而文献 [11] 提出了总体样本容量不等时均值检验的方法 (简记为 TAO 方法). 本文分别在总体样本容量相等和不等的情形下, 进行模拟研究, 结果见表 2. 模拟结果表明, 当总体样本容量相等时, VDR 检验的功效比文献 [10] 中的方法的功效低, 但是, 随着总体个数和样本容量的增大, VDR 检验与文献 [10] 中的方法的检验功效差距减少. 例如, 考虑 3 个总体, 样本容量都为 20, 总体均值分别为 0.1、0.5、1、2.3 和 0, 标准差为 1、5、7、3 和 6, 检验水平 $\alpha = 0.05$ 时, VDR 检验和文献 [10] 中的方法功效分别为 0.581 和 0.717. 考虑 8 个正态总体, 其样本容量都为 10, 总体均值分别为 0、1、0.5、0.7、2、2.5、3.2 和 -0.4 , 标准差为 1、2、4、3、5、1、7 和 9, 检验水平 $\alpha = 1.0$ 时, VDR 检验和文献 [10] 中的方法功效分别为

表 1 VDR 检验和单因素方差分析 (ANOVA) 模拟功效比较

样本容量 (n_1, n_2, n_3)	标准差 $\sigma_1, \sigma_2, \sigma_3$	显著水平 α	检验方法	最大均值 μ_3 (均值向量 $\boldsymbol{\mu} = (0, 0.5\mu_3, \mu_3)$)						
				0.0	0.25	0.5	0.75	1.0	1.5	2.0
(10,15,22)	(0.5,1,1.5)	0.10	VDR	0.090	0.166	0.301	0.581	0.802	0.987	1.000
		0.05		0.0417	0.085	0.208	0.433	0.687	0.961	0.999
		0.01		0.0717	0.022	0.066	0.205	0.418	0.867	0.992
(10,15,20)	(1,1,1)	0.10	VDR	0.088	0.143	0.276	0.452	0.700	0.950	1.000
			ANOVA	0.1003	0.150	0.296	0.490	0.746	0.965	1.000
		0.05	VDR	0.0415	0.066	0.173	0.311	0.567	0.897	0.999
			ANOVA	0.0503	0.080	0.210	0.365	0.626	0.921	1.000
		0.01	VDR	0.0083	0.015	0.055	0.109	0.295	0.715	0.963
			ANOVA	0.0095	0.021	0.064	0.143	0.367	0.788	0.983
(100,150,300)	(4,4,4)	0.10	VDR	0.102	0.116	0.255	0.410	0.622	0.914	0.991
			ANOVA	0.105	0.127	0.271	0.436	0.640	0.931	0.991
		0.05	VDR	0.053	0.063	0.138	0.290	0.480	0.847	0.982
			ANOVA	0.053	0.064	0.158	0.301	0.511	0.859	0.985
		0.01	VDR	0.012	0.016	0.040	0.135	0.240	0.645	0.944
			ANOVA	0.012	0.019	0.047	0.138	0.263	0.698	0.951

0.984 和 1.00. 当各个总体方差不等, 且样本容量不等时, VDR 检验的功效比文献 [11] 中的方法的功效高. 例如, 当总体样本容量分别为 5、15、20、30 和 40, 总体均值分别为 0.1、2、3、0.5 和 1, 标准差分别为 1、2、3、4 和 5, 检验水平分别为 0.01、0.05 和 0.1 时, VDR 检验功效分别为 0.564、0.831 和 0.907, 文献 [11] 中的方法的模拟功效分别为 0.381、0.668 和 0.784. 这表明, 在总体方差不等的情形下, 无论样本容量是否相等, VDR 检验都是有效的, 而且检验功效不会低. 这也是 VDR 检验的优势. 但是 VDR 检验的缺点是计算相对复杂. 另一方面, 我们也考虑了多个正态总体均值检验经典的两两比较 (pairwise comparison) 方法, 并利用 Bonferroni 不等式得出结论, 将两两比较的方法和 VDR 检验进行模拟功效的比较, 结果显示, VDR 检验比两两比较的 t 检验有更高的功效见表 2.

最后比较三个正态总体均值是否相等的 VDR 检验和单因素方差分析的显著水平. 对每一个显著水平, 共做 6 次模拟, 每一次模拟重复 1000 次, 模拟结果见表 3. 模拟结果显示, VDR 检验的显著水平比单因素方差分析更接近于真实值. 例如, 当样本容量分别取 10、15 和 20, 标准差分别为 1、2 和 4, 真实的显著水平为 0.10 时, VDR 检验 6 次模拟显著水平的均值是 0.094, 而单因素方差分析 6 次模拟显著水平的均值是 0.063, 当真实的显著水平取为 0.05 时, VDR 检验和单因素方差分析 6 次模拟显著水平的均值分别是 0.045 和 0.034.

通过模拟研究说明, 对于多个正态总体均值的检验, 无论总体方差是否相同, 各样本容量是否相等, 都可以使用 VDR 检验, 且检验功效与已有的检验方法的功效接近, 甚至在一些的特定的假设下, 例如, 当总体方差不等和各个样本容量不等时, VDR 检验功效比其他方法的功效高. 这些都说明了 VDR 检验的适用性. 但是 VDR 检验计算相对复杂. 因此, 我们给出了多个总体均值检验的 VDR 检

表 2 VDR 检验、ZHONG 方法、TAO 方法和两两比较 (PAIRWISE) 功效比较

样本容量 ($n_1, \dots, n_p$)	均值 ($\mu_1, \dots, \mu_p$)	标准差 ($\sigma_1, \dots, \sigma_p$)	检验方法	显著水平		
				0.01	0.05	0.1
(40,40,40)	(0,0,0)	(3,7,10)	VDR	0.005	0.035	0.074
			ZHONG	0.006	0.045	0.094
(20,20,20,20,20)	(0.1,0.5,1,2.3,0)	(1.5,7.3,6)	VDR	0.312	0.581	0.696
			ZHONG	0.497	0.717	0.812
(10,10,10,10,10,10,10,10)	(0.1,0.5,0.7,2,2.5,3.2,-0.4)	(1.2,4.3,5.1,7.9)	VDR	0.808	0.954	0.984
			ZHONG	0.980	0.997	1.00
(100,150,300)	(0.1,0.5,0.8)	(2,4,8)	VDR	0.102	0.246	0.356
			TAO	0.117	0.266	0.368
(10,15,50,35,45)	(0,0,0,0,0)	(1,2,3,4,5)	VDR	0.009	0.039	0.080
			TAO	0.011	0.044	0.083
(5,15,20,30,40)	(0.1,2,3,0.5,1)	(1,2,3,4,5)	VDR	0.564	0.831	0.907
			TAO	0.381	0.668	0.784
(5,15,20,30,40,50,55,60)	(0.1,2,3,0.5,1,0.02,-1,1.5)	(1,2,3,4,5,6,7,8)	VDR	0.767	0.919	0.955
			TAO	0.511	0.736	0.838
(10,30,20)	(0.3,0.5,0))	(1,1,1)	VDR	0.124	0.293	0.420
			PAIRWISE	0.011	0.085	0.155
(10,30,20,15,40)	(0.3,0.5,0.1,1,-0.2))	(2,2,2,2,2)	VDR	0.127	0.0321	0.459
			PAIRWISE	0.007	0.039	0.077

表 3 VDR 检验与 ANOVA 方法模拟显著水平比较

样本容量	标准差	显著水平	检验方法	模拟次数						
(n_1, n_2, n_3)	$(\sigma_1, \sigma_2, \sigma_3)$	α		1	2	3	4	5	6	平均值
(10, 15, 22)	(0.5, 1, 1.5)	0.10	VDR	0.088	0.097	0.093	0.070	0.106	0.085	0.0900
		0.05		0.038	0.059	0.046	0.029	0.045	0.033	0.0417
		0.01		0.005	0.008	0.007	0.004	0.013	0.006	0.0072
(10, 15, 20)	(1, 2, 4)	0.10	VDR	0.086	0.106	0.089	0.098	0.095	0.087	0.0940
			ANOVA	0.056	0.077	0.058	0.066	0.067	0.053	0.0630
		0.05	VDR	0.034	0.053	0.042	0.053	0.046	0.041	0.0450
			ANOVA	0.025	0.039	0.034	0.040	0.033	0.030	0.0340
		0.01	VDR	0.007	0.011	0.007	0.011	0.011	0.005	0.0087
			ANOVA	0.006	0.010	0.008	0.014	0.010	0.004	0.0090
(10, 15, 20)	(1, 2, 3)	0.10	VDR	0.086	0.080	0.096	0.089	0.089	0.084	0.0870
			ANOVA	0.070	0.050	0.062	0.047	0.063	0.066	0.0600
		0.05	VDR	0.044	0.035	0.048	0.045	0.054	0.046	0.0450
			ANOVA	0.030	0.025	0.037	0.023	0.039	0.031	0.0310
		0.01	VDR	0.004	0.006	0.012	0.004	0.008	0.009	0.0072
			ANOVA	0.003	0.003	0.008	0.004	0.011	0.010	0.0065

验的 R 软件主要程序, 见附录 A4. 在具体检验时, 我们只需要根据构造的 VDR 检验统计量, 修改对应的函数表达即可.

6 结论

我们将一维随机估计的概念推广至多维的情形, 结合多维枢轴量和置信分布, 提出了 VDR 检验. VDR 检验的思想可简述如下: 首先构造枢轴量, 根据枢轴量得到随机估计的方程; 其次计算出随机估计的密度函数, 将密度函数中的自变量换成随机估计, 定义检验随机变量; 最后利用垂直密度表示计算检验随机变量的分位点, 给出检验规则.

运用 VDR 检验, 给出了多个正态总体均值相等的检验, 并进行模拟研究, 模拟结果显示, 利用 VDR 作检验, 无论样本方差是否相等, 各总体的样本容量是否相等, VDR 检验都是有效的, 而且功效比某个特定的检验的功效低一点, 但是随着样本容量的增加, 两者之间的差异变小, 甚至相等. 从整个过程来看, VDR 检验的计算相对复杂, 这是它的弱点.

对于 VDR 检验, 我们可以考虑如下的问题:

(1) VDR 检验理论是围绕枢轴量展开的, 利用枢轴量定义随机估计, 如果当枢轴量没法找到时, 可以考虑是否可用其他方式定义随机估计, 且具有与枢轴量定义的随机估计类似的基本特性. 而且如果存在两个以上的枢轴量, 应该选择哪个来构造随机估计? 如果两个枢轴量都用于构造随机估计, 所得到的 VDR 检验原则会导致统计推断结论的变化吗? 这些 VDR 检验中谁更优, 评价的标准是什么?

(2) 在模拟研究中, 实现模拟推断较复杂, 当针对高维参数和高维总体时, 模拟推断会遇见计算难点. 如何实现快速模拟推断是需要研究的问题.

致谢 作者衷心感谢审稿专家提出的宝贵意见.

参考文献

- 1 Fisher R A. Inverse probability. Proc Cambridge Philos Soc, 1930, 26: 528–535
- 2 Singh K, Xie M G, Strawderman W E. Confidence Distribution (CD)-distribution estimator of a parameter. IMS Lecture Notes Monogr Ser, 2007, 54: 132–150
- 3 Xie M G, Singh K, Strawderman W E. Confidence distribution and Unifying framework for meta-analysis. J Amer Statist Assoc, 2011, 106: 320–333
- 4 Schwder T, Hjort N L. Confidence and likelihood. Scand J Statist, 2002, 29: 309–332
- 5 杨振海. 随机估计和 VDR 检验. 北京: 高等教育出版社, 2014
- 6 Fang K T, Yang Z H, Kotz S. Generation of Multivariate Distributions by Vertical Density Representation. Statistics, 2001, 35: 281–293
- 7 Troutt M D. A theorem on the density of the density ordinate and an alternative derivation of the Box-muller method. Statistics, 1991, 22: 436–466
- 8 Troutt M D, Pang W K, Hou S H. Vertical density representation and Its Application. Singapore: World Scientific Publishing, 2004
- 9 Yang Z H, Samuel K. Center-similar distribution with applications in multivariate analysis. Statist Probab Lett, 2003, 64: 335–345
- 10 仲崇新. 方差不等时的多总体均值的相等性检验. 扬州师院学报 (自然科学版), 1994, 14: 28–31
- 11 陶熙, 郭民之. 不等重复试验下的单因素方差分析. 云南民族大学学报 (自然科学版), 2008, 17: 116–118

附录 A1 假设 (3.1) 的 VDR 检验为 Hoteling 检验

设 Z 的 α 分位点记作 $Q_Z(\alpha; \mathcal{X})$, 则

$$\begin{aligned} \alpha &= P(Z \leq Q_Z(\alpha; \mathcal{X})) \\ &= P\left(f_T(\mathbf{h}_1(\mathcal{X}, \mathbf{W}_\mu), n-1, p) \leq \frac{|\hat{\Sigma}|^{\frac{1}{2}}}{\sqrt{n}} Q_Z(\alpha; \mathcal{X})\right) \quad (0 < Q_Z(\alpha; \mathcal{X}) \leq \sqrt{n} |\hat{\Sigma}|^{-\frac{1}{2}} f_0) \\ &= P(f_T(\mathbf{h}_1(\mathcal{X}, \mathbf{W}_\mu), n-1, p) \leq Q_V(\alpha)) \quad (0 < Q_V(\alpha) \leq f_0), \end{aligned}$$

其中

$$\frac{|\hat{\Sigma}|^{\frac{1}{2}}}{\sqrt{n}} Q_Z(\alpha; \mathcal{X}) = Q_V(\alpha), \quad f_0 = f_T(\mathbf{0}_p, n-1, p) = (k\pi)^{\frac{p}{2}} \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})},$$

且 $Q_V(\alpha)$ 是 $V = f_T(\mathbf{T}_{p,n-1}, n-1, p)$ 的 α 分位点. 因此, 检验假设 (3.1) 的规则如下:

规则 A1 给定显著水平 α , 若 $f_T(\mathbf{h}(\mathcal{X}, \boldsymbol{\mu}_0), n-1, p) \leq Q_V(\alpha)$, 则拒绝接受 (3.1) 的原假设; 否则接受原假设.

进一步, 计算 $Q_V(\alpha)$. 由于

$$f_T(\mathbf{t}_p, n-1) \leq v \Leftrightarrow \left(1 + \frac{\sum_{i=1}^p t_i^2}{n-1}\right)^{-\frac{n}{2}} \leq \frac{v}{f_0} \Leftrightarrow \mathbf{t}^T \mathbf{t} = \sum_{i=1}^p t_i^2 \geq (n-1) \left(\left(\frac{f_0}{v} \right)^{\frac{2}{n}} - 1 \right),$$

故

$$f_T(\mathbf{T}; \boldsymbol{\mu}_0), n-1, p \leq Q_V(\alpha) \Leftrightarrow \mathbf{T}^T \mathbf{T} \geq (n-1) \left(\left(\frac{f_0}{Q_V(\alpha)} \right)^{\frac{2}{n}} - 1 \right).$$

又

$$\mathbf{T}^T \mathbf{T} = n \bar{\mathbf{Z}}^T \mathbf{S}(\mathbb{Z}) \bar{\mathbf{Z}} \sim T^2(p, n-1),$$

$$\alpha = P(f_T(\mathbf{T}, n-1, p) \leq Q_V(\alpha)) \Leftrightarrow \left(\left(\frac{f_0}{Q_V(\alpha)} \right)^{\frac{2}{n}} - 1 \right) = T_{p,n-1}(1-\alpha),$$

其中 $T_{p,n-1}(\gamma)$ 是 Hoteling 统计量 $T(p, n)$ 的 γ 分位点. 这意味着, 在协方差阵未知的条件下, 检验假设 (3.1) 的 VDR 检验就是 Hoteling 检验.

附录 A2 计算随机估计 $\mathbf{W}_\mu$ 的密度函数

注意到

$$M = \frac{\partial(\mathbf{z}^T, \chi_{n-p}^2)^T}{\partial(W_\mu^T, W_{\sigma^2})^T} \Big|_{W_\mu=\mathbf{t}, W_{\sigma^2}=v} = \begin{pmatrix} \text{diag}(\sqrt{\frac{n}{v}}) & -\frac{\sqrt{n}(\bar{\mathbf{x}}-\mathbf{t})}{2v^{\frac{3}{2}}} \\ \mathbf{0}_p^T & -\frac{(n-p)s_n^2}{v^2} \end{pmatrix},$$

$$J(v; s_n^2) = |\det(M)| = \frac{(n-p)s_n^2}{v^{\frac{p}{2}+2}} \prod_{i=1}^p \sqrt{n_i}.$$

于是有 $\mathbf{W}_\mu$ 和 W_{σ^2} 的联合概率密度函数为

$$\begin{aligned} f_{\mu, \sigma^2}(\mathbf{t}, v; \bar{\mathbf{x}}, s_p^2) &= J(v; s_n^2) f_{\mathbf{Z}, \chi_{n-p}^2}(\mathbf{h}(\bar{\mathbf{x}}, s_p^2; \mathbf{t}, v)) \\ &= \frac{\prod_{i=1}^p \sqrt{n_i}}{(\sqrt{2\pi})^p 2^{\frac{n-p}{2}} \Gamma(\frac{n-p}{2})} \left(\frac{(n-p)s_n^2}{v} \right)^{\frac{n-p}{2}-1} \frac{(n-p)s_n^2}{v^{\frac{p}{2}+2}} e^{-\frac{1}{2} \sum_{i=1}^p \frac{n_i(\bar{x}_i-t_i)^2}{v} - \frac{1}{2} \frac{(n-p)s_n^2}{v}} \\ &= \frac{\prod_{i=1}^p \sqrt{n_i}}{(\sqrt{2\pi})^p 2^{\frac{n-p}{2}} \Gamma(\frac{n-p}{2})} \frac{1}{[(n-p)s_n^2]^{\frac{p}{2}+1}} \left(\frac{(n-p)s_n^2}{v} \right)^{\frac{n}{2}+1} e^{-\frac{(n-p)s_n^2}{2v} \left(1 + \frac{\sum_{i=1}^p n_i(\bar{x}_i-t_i)^2}{(n-p)s_n^2} \right)}, \end{aligned}$$

同时, μ 的随机估计 $\mathbf{W}_\mu = (W_{\mu_1}, \dots, W_{\mu_p})^T$ 的密度函数为

$$\begin{aligned} f_\mu(\mathbf{t}; \bar{\mathbf{x}}, s_p^2) &= \int_0^\infty f_{\mu, \sigma^2}(\mathbf{t}, v; \bar{\mathbf{x}}, s_p^2) dv \\ &= \frac{1}{(\sqrt{2\pi})^p 2^{\frac{n-p}{2}} \Gamma(\frac{n-p}{2})} \frac{\prod_{i=1}^p \sqrt{n_i}}{[(n-p)s_n^2]^{\frac{p}{2}}} \int_0^\infty \left(\frac{1}{v} \right)^{\frac{n}{2}+1} e^{-\frac{1}{2v} (1 + \frac{\sum_{i=1}^p n_i(\bar{x}_i-t_i)^2}{(n-p)s_n^2})} dv \\ &= \frac{1}{(\sqrt{2\pi})^p 2^{\frac{n-p}{2}} \Gamma(\frac{n-p}{2})} \frac{\prod_{i=1}^p \sqrt{n_i}}{[(n-p)s_n^2]^{\frac{p}{2}}} \int_0^\infty s^{\frac{n}{2}-1} e^{-\frac{s}{2} (1 + \frac{\sum_{i=1}^p n_i(\bar{x}_i-t_i)^2}{(n-p)s_n^2})} ds \\ &= \frac{1}{(\sqrt{2\pi})^p 2^{\frac{n-p}{2}} \Gamma(\frac{n-p}{2})} \frac{\prod_{i=1}^p \sqrt{n_i}}{[(n-p)s_n^2]^{\frac{p}{2}}} \left(1 + \frac{\sum_{i=1}^p n_i(\bar{x}_i-t_i)^2}{(n-p)s_n^2} \right)^{-\frac{n}{2}} \int_0^\infty s^{\frac{n}{2}-1} e^{-\frac{s}{2}} ds \\ &= \frac{\Gamma(\frac{n}{2})}{\sqrt{\pi^p} \Gamma(\frac{n-p}{2})} \frac{\prod_{i=1}^p \sqrt{n_i}}{[(n-p)s_n^2]^{\frac{p}{2}}} \left(1 + \frac{\sum_{i=1}^p n_i(\bar{x}_i-t_i)^2}{(n-p)s_n^2} \right)^{-\frac{n}{2}} \\ &= C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2)) (1 + (\bar{\mathbf{x}} - \mathbf{t})^T \Lambda(\mathbf{n}, s_n^2) (\bar{\mathbf{x}} - \mathbf{t}))^{-\frac{n}{2}}, \end{aligned}$$

其中

$$\Lambda(\mathbf{n}, s_n^2) = \text{diag} \left(\frac{\mathbf{n}}{(n-p)s_n^2} \right).$$

在各总体方差相等的条件下, 假设 (4.1) 的检验变量 Z 为

$$Z = f_\mu(\mathbf{W}_\mu, \mathcal{X}) = C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2)) (1 + (\bar{\mathbf{x}} - \mathbf{W}_\mu)^T \Lambda(\mathbf{n}, s_n^2) (\bar{\mathbf{x}} - \mathbf{W}_\mu))^{-\frac{n}{2}}.$$

记 Z 的 α 分位点为 $Q_Z(\alpha, s_n^2)$, 有

$$\alpha = P(Z \leq Q_Z(\alpha, s_n^2))$$

$$\begin{aligned}
&= P(f_{\mu}(\mathbf{W}_{\mu}; \bar{\mathbf{x}}, s_p^2) \leq Q_Z(\alpha, s_n^2)), \quad (0 < Q_Z(\alpha, s_n^2) \leq C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2))) \\
&= P((1 + (\bar{\mathbf{x}} - \mathbf{W}_{\mu})^T \Lambda(\mathbf{n}, s_n^2)(\bar{\mathbf{x}} - \mathbf{W}_{\mu}))^{-\frac{n}{2}} \leq Q_V(\alpha)) \\
&= P((\bar{\mathbf{x}} - \mathbf{W}_{\mu})^T \Lambda(\mathbf{n}, s_n^2)(\bar{\mathbf{x}} - \mathbf{W}_{\mu}) > (Q_V(\alpha))^{-\frac{2}{n}} - 1),
\end{aligned}$$

其中

$$Q_V(\alpha) = \frac{Q_Z(\alpha, s_n^2)}{C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2))}, \quad 0 < Q_V(\alpha) < 1.$$

于是, 检验假设 (4.1) 的 VDR 检验规则如下:

规则 A2 若

$$(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^T \Lambda(\mathbf{n}, s_n^2)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) > (Q_V(\alpha))^{-\frac{2}{n}} - 1,$$

则在显著水平 α 拒绝接受假设 $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$; 否则接受假设 H_0 .

附录 A3 $Q_V(\alpha)$ 的两种计算方法

首先按经典方法求 $Q_V(\alpha)$. 将上式的二次型写成显式形式, 有

$$K(\bar{\mathbf{x}}, \mathbf{s}^2; \boldsymbol{\mu}_0) = (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^T \Lambda(\mathbf{n}, s_n^2)(\bar{\mathbf{x}} - \boldsymbol{\mu}_0) = \frac{\sum_{i=1}^p n_i (\bar{x}_i - \mu_{0i})^2}{(n-p)s_n^2}.$$

上式是经典意义下的检验统计量. 因为有 $(n-p)S_n^2 \sim p_{\chi}(\cdot, n-p)$, 其中 $p_{\chi}(\cdot, n-p)$ 是自由度为 $n-p$ 的 χ^2 分布函数, 同时, 在各总体方差相等的条件下,

$$\sum_{i=1}^p n_i (\bar{x}_i - \mu_{0i})^2 \sim p_{\chi}(\cdot, p).$$

所以,

$$\frac{n-p}{p} K(\bar{\mathbf{X}}, \mathbf{S}^2; \boldsymbol{\mu}_0) \sim p_F(\cdot, p, n-p),$$

其中 $p_F(\cdot, p, n-p)$ 是自由度为 p 和 $n-p$ 的 F 分布函数. 故

$$(Q_V(\alpha))^{-\frac{2}{n}} - 1 = \frac{p}{n-p} F_{p, n-p}(\alpha).$$

是否按 VDR 检验临界值计算方法计算的 $Q_V(\alpha)$ 仍满足上式? 回答是肯定的.

其次应用临界值计算通用公式 (2.3) 计算检验随机变量 Z 的分位点, 即计算检验假设 (4.1) 的 VDR 检验临界值 $Q_Z(\alpha, s_n^2) = C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2)) Q_V(\alpha)$,

$$\begin{aligned}
\alpha &= \int_0^{Q_Z(\alpha, s_n^2)} f_Z(v; \bar{\mathbf{x}}, \mathbf{s}_p^2) dv = \int_0^{Q_Z(\alpha, s_n^2)} -v \frac{dL_p(D_{[f_{\mu}(\cdot; \bar{\mathbf{x}}, \mathbf{s}_p^2)]}(v))}{dv} dv \\
&\equiv \int_0^{Q_Z(\alpha, s_n^2)} -v \frac{dL_p(D_{[f_{\mu}]}(v))}{dv} dv \\
D_{[f_{\mu}]}(v) &= \{\mathbf{t}: f_{\mu}(\mathbf{t}; \bar{\mathbf{x}}, \mathbf{s}_p^2) \geq v\} \\
&= \left\{ \mathbf{t}: (\bar{\mathbf{x}} - \mathbf{t})^T \Lambda(\mathbf{n}, s_n^2)(\bar{\mathbf{x}} - \mathbf{t}) \leq \left(\frac{v}{C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2))} \right)^{-\frac{2}{n}} - 1 \right\}
\end{aligned}$$

$$\begin{aligned}
&= \{\mathbf{t} : (\bar{x} - \mathbf{t})^T \Lambda(\mathbf{n}, s_n^2)(\bar{x} - \mathbf{t}) \leq (v^*)^{-\frac{2}{n}} - 1\}, \\
&= S_p^2(\sqrt{(v^*)^{-\frac{2}{n}} - 1}, \bar{\mathbf{x}}, \Lambda(\mathbf{n}, s_n^2)), \\
L_p(D_{[f_\mu]}(v)) &= \frac{1}{\det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2))} L_p(S_p^2(\sqrt{(v^*)^{-\frac{2}{n}} - 1})) \\
&= \frac{\sqrt{\pi^p}}{\det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2)) \Gamma(\frac{p}{2} + 1)} ((v^*)^{-\frac{2}{n}} - 1)^{\frac{p}{2}},
\end{aligned}$$

其中

$$S_p^2(r, \mathbf{a}, A) = \{\mathbf{x} : (\mathbf{x} - \mathbf{a})^T A(\mathbf{x} - \mathbf{a}) \leq r^2\}, \quad v^* = \frac{v}{C(p, n) \det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2))}, \quad 0 < v^* \leq 1,$$

于是,

$$\begin{aligned}
\alpha &= \int_0^{Q_Z(\alpha, s_n^2)} -v \frac{dL_p(D_{[f_\mu]}(v))}{dv} dv \\
&= \frac{\sqrt{\pi^p}}{\det(\Lambda^{\frac{1}{2}}(\mathbf{n}, s_n^2)) \Gamma(\frac{p}{2} + 1)} \int_0^{Q_Z(\alpha, s_n^2)} \frac{p}{n} ((v^*)^{-\frac{2}{n}} - 1)^{\frac{p}{2}-1} (v^*)^{-\frac{2}{n}} dv \\
&= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-p}{2}) \Gamma(\frac{p}{2} + 1)} \int_0^{Q_V(\alpha)} \frac{p}{n} (v^{-\frac{2}{n}} - 1)^{\frac{p}{2}-1} v^{-\frac{2}{n}} dv.
\end{aligned}$$

作变量代换 $v^{-\frac{2}{n}} - 1 = s$, 则

$$v = \frac{1}{(1+s)^{\frac{n}{2}}}, \quad dv = \frac{-n}{2(1+s)^{\frac{n}{2}+1}} ds.$$

于是,

$$\begin{aligned}
\alpha &= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-p}{2}) \Gamma(\frac{p}{2} + 1)} \int_0^{Q_V(\alpha)} \frac{p}{n} (v^{-\frac{2}{n}} - 1)^{\frac{p}{2}-1} v^{-\frac{2}{n}} dv \\
&= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-p}{2}) \Gamma(\frac{p}{2} + 1)} \frac{p}{2} \int_{(Q_V(\alpha))^{-\frac{2}{n}} - 1}^{\infty} s^{\frac{p}{2}-1} (1+s)^{-\frac{n}{2}} ds \\
&= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-p}{2}) \Gamma(\frac{p}{2})} \int_{(Q_V(\alpha))^{-\frac{2}{n}} - 1}^{\infty} s^{\frac{p}{2}-1} (1+s)^{-\frac{n}{2}} ds \quad \left(\text{作变换 } s = \frac{p}{n-p} r \right) \\
&= \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-p}{2}) \Gamma(\frac{p}{2})} \int_{\frac{n-p}{p}((Q_V(\alpha))^{-\frac{2}{n}} - 1)}^{\infty} \left(\frac{p}{n-p} \right)^{\frac{p}{2}-1} r^{\frac{p}{2}-1} \left(1 + \frac{p}{n-p} r \right)^{-\frac{n}{2}} dr.
\end{aligned}$$

将上式与 F 分布密度函数比较知,

$$(Q_V(\alpha))^{-\frac{2}{n}} - 1 = \frac{p}{n-p} F_{p, n-p}(1 - \alpha).$$

这说明了, 用 VDR 检验计算的临界值与经典方法计算的结果是一致的. 至此解决了假设 (4.4) 的 VDR 检验和临界值计算.

附录 A4 VDR 检验的主要 R 程序

Midu < - function(c, n1, n) {

```

if (!(length(sum(n1^2)) == 1)) return "模长不等 1, 请重新输入"
l = length(n1)
if(!(l == length(c)&l == length(n))) return "数据不匹配, 请重新输入"
c = c - sum(c * n1) * n1
tlianhe < -function(x){
  a = 1;
  for (i in 1:l)a = a * dt(c[i] + x * n1[i], n[i] - 1)
  return(a)
}
b=integrate (tlianhe, -Inf, Inf)[[1]]
return(c(c, b))
}
Qv< -function(n1, n, N = 10000){
  l=length(n)
  F=numeric()
  for (i in 1:N){
    t=numeric()
    for(j in 1:l)t = c(t, rt(1, n[j] - 1))
    F = c(F, Midu(t, n1, n)[l + 1])
  }
  F = sort(F)
  return(F)
}

```

说明:

(1) Midu (c, n, s) 功能: 计算差参数密度函数值, 其中 n 为样本容量向量, s 为样本标准差向量, c 为密度函数自变量.

(2) Qv ($\alpha, n, s, N = 10000$) 功能: 模拟枢轴向量的 VDR 变量的分位点, 其中 α 为显著水平, N 为模拟重复次数.

VDR test for equality of means of multiple normally distributed populations

DAI JiaJia, CHENG WeiHu, ZHANG ZhongZhan & YANG ZhenHai

Abstract This paper develops the concept on randomized estimation of one-dimensional parameter to that of vector parameter. The randomized estimator of vector parameter is defined by using both pivotal quantity

of the parameters and pivotal random vector. Based on test variable which is defined through both randomized estimation and its probability density function, the VDR test is constructed. As an application, the proposed method is used to test whether the means of several normal distribution populations are equal or not. One-way analysis of variance is a special case of VDR test. The simulation study reveals that VDR test is a valid and unified approach. Some illustrations show that some classical results can be derived by VDR test.

Keywords randomized estimation, confidence distribution, vertical density representation

MSC(2010) 62H12, 62H15

doi: 10.1360/N012013-00091