Vol. 42 (2022)
No. 6

数 学 杂 志
J. of Math. (PRC)

P- 范分布尺度参数的似然比检验及其应用

胡宏昌, 任 欢

(湖北师范大学数学与统计学院, 湖北 黄石, 435002)

摘 要: P- 范分布包含拉普拉斯分布、正态分布、均匀分布、退化分布等在内的一类重要分布, 本文用似然比检验方法研究了该分布中未知尺度参数的检验问题. 首先, 介绍三个与 P- 范分布有密切关系的抽样分布; 然后, 用似然比检验方法研究了一个和两个总体为 P- 范分布的未知尺度参数的检验问题, 得到了似然比检验检验统计量的渐近分布及其 (近似) 密度函数, 并给出了拉普拉斯分布、正态分布等特殊情形下的 (近似) 密度函数; 最后, 给出了一些模拟算例来验证本文方法的有效性.

关 键 词: P- 范分布; 似然比检验; 尺度参数

MR(2010) **主题分类号:** 62F03 **中 图 分 类 号:** O212.2

文 献 标 识 码: A **文 章 编 号:** 0255-7797(2022)06-0493-10

1 引言

在研究线性回归模型时, 往往直接或间接地假设误差服从正态分布, 取得了大量令人满意的结果 (参见文献 [1-2]). 但在很多情况下, 误差并不服从正态分布, 由此将会产生一些用正态分布无法解决的问题. 由于 P- 范分布包含拉普拉斯分布、正态分布、均匀分布、退化分布等在内的一类重要分布, 在某些情况下 (如地图数字化误差) 它比正态分布更适合描述误差分布. 另外, 在常用的估计最小准则下, 参数的估值与总体服从 P- 范分布时的极大似然估值一致 (参见文献 [3]). 基于这些, 越来越多的人研究 P- 范分布及其相关问题, 如: 文献 [3] 研究了 P- 范分布及其抽样分布; 文献 [4] 系统地研究了 P- 范分布理论及其在现代测量数据处理中的应用; 文献 [5] 对 P - 范分布参数的估计方法进行了改进, 建立了 P- 范分布的参数估计的实数阶矩估计方法; 文献 [6] 用 P- 范分布对上证指数的日收益率进行拟合, 得到较好的结果. 似然比方法是一种应用很广泛的检验方法, 深入的结果有很多, 如: 文献 [7] 用似然比方法对线性混合模型的方差参数进行了检验; 文献 [8] 运用 Skovgaard's 修正似然比方法研究指数族非线性模型, 得到了修正似然比统计量近似服从分布; 文献 [9] 讨论了高维正态向量独立性的假设检验问题, 得到了似然比统计量的极限分布. 文献 [10] 用 Lq 似然比方法对广义极值分布的形状参数进行了检验; 文献 [11] 针对一般的污染分布提出了具有稳健性的 Lq 似然比检验, 得到了 Lq 似然比检验统计量的渐近分布; 文献 [12] 提出了基于似然比检验的二阶段多项式曲线控制图, 并通过平均运行长度来衡量控制图的性能表现.

*收稿日期: 2021-04-27 接收日期: 2021-12-23

基金项目: 国家自然科学基金资助项目 (11471105).

作者简介: 胡宏昌 (1971–), 男, 汉族, 湖北英山人, 教授, 主要研究方向: 回归模型的统计推断及其应用. E-mail: retutome@163.com.

通讯作者: 任欢 (1996–), 女, 汉族, 河南南阳人, 在读硕士研究生, 主要研究方向: 概率论与数理统计. E-mail: 2448144234@qq.com

尽管研究 P- 范分布和运用似然比检验方法的成果很多, 但是运用似然比方法对 P- 范分布中的参数进行检验的文献甚少. 为此, 本文用似然比方法对 P- 范分布中尺度参数的检验进行了研究.

2 P- 范分布及其抽样分布

本节简要介绍 P- 范分布及其抽样分布, 详细内容请参见文献 [3-4].

定义 2.1　若随机变量 X 的密度函数为

$$f(x)=\frac{p\lambda}{2\sigma\Gamma(1/p)}\exp\Big\{-\Big[\lambda\frac{|x-\mu|}{\sigma}\Big]^p\Big\}, \tag{2.1}$$

则称 X 服从一元 P- 范分布 (其中 $\lambda=(\Gamma(3/p)/\Gamma(1/p))^{1/2}$, 位置参数为 μ , 尺度参数为 $\sigma>0$).

注　易知拉普拉斯分布 (令 $p=1$)、正态分布 (令 $p=2$)、均匀分布 (令 $p\to\infty$) 与退化分布 (令 $p\to 0$) 均为一元 P- 范分布 (2.1) 的特例. 由此可知, P- 范分布揭示了表面上毫不相关的几个分布的内在本质联系.

定义 2.2　设随机变量 $X_1,X_2,\cdots,X_n$ 独立同分布, 服从 $\mu=0$、$\sigma=1$ 的一维 P- 范分布, 则称随机变量 $Y=\sum\limits_{i=1}^{n}|X_i|^p$ 服从自由度为 n 的 χ^p 分布, 记为 $Y\sim\chi^p(n)$, 其密度函数为

$$f(y)=\begin{cases}\dfrac{\lambda^n}{\Gamma(n/p)}y^{n/p-1}e^{-\lambda^p y}, & y>0\\ 0, & y\le 0\end{cases} \tag{2.2}$$

定义 2.3　设 X 服从 $\mu=0$、$\sigma=1$ 的一维 P- 范分布, Y 服从 $\chi^p(n)$ 分布, 且 X 与 Y 相互独立, 则

$$T=\frac{X}{(Y/n)^{1/p}}, \tag{2.3}$$

服从自由度为 n 的 t_p 分布, 记为 $T\sim t_p(n)$, 其密度函数为

$$f(t)=\frac{p\Gamma((n+1)/p)}{2\Gamma(1/p)\Gamma(n/p)n^{1/p}}\left(1+\frac{|t|^p}{n}\right)^{-(n+1)/p}. \tag{2.4}$$

定义 2.4　设 X、Y 分别服从自由度为 m,n 的 χ^p 分布, 且 X 与 Y 相互独立, 则

$$Z=\frac{X/m}{Y/n} \tag{2.5}$$

服从第一自由度为 m、第二自由度为 n 的 F_p 分布, 记为 $Z\sim F_p(m,n)$, 其密度函数为

$$f(z)=\begin{cases}\dfrac{(m/n)^{m/p}}{B(m/p,n/p)}z^{m/p-1}\left(1+\dfrac{m}{n}z\right)^{-\frac{m+n}{p}}, & z>0\\ 0, & z\le 0\end{cases} \tag{2.6}$$

式中 $B(x,y)$ 为贝塔函数.

3 似然比检验

本节假定 μ 已知, 不失一般性假定 $\mu = 0$. 当 μ 未知时, 由于此时一元 P- 范分布的极大似然估计 (参见文献 [5]) 不易求出, 所以在实际应用中用其矩估计 $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$ 替代, 仍可采用下面的密度函数进行近似计算.

3.1 单总体尺度参数的检验

本小节我们考虑单个总体的尺度参数的假设检验问题. 设随机样本 $x_1, x_2, \cdots, x_n$ 独立同分布, 取自参数为 μ 和 σ 的 P- 范分布, 检验问题为

$$H_0 : \sigma \in \Theta_0 \leftrightarrow H_1 : \sigma \in \Theta_1,$$

其中 Θ_0 和 Θ_1 分别为原假设和备择假设的参数空间. 如记

$$\begin{aligned} l_n(\sigma) &= \log L(\sigma; x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \log f(x_i; \sigma) \\ &= \sum_{i=1}^{n} \log \left\{ \frac{p\lambda}{2\sigma\Gamma(1/p)} \exp\left\{ -\left(\lambda \frac{|x_i|}{\sigma} \right)^p \right\} \right\} \\ &= n \log \left\{ \frac{p\lambda}{2\sigma\Gamma(1/p)} \right\} - \sum_{i=1}^{n} \left(\lambda \frac{|x_i|}{\sigma} \right)^p, \end{aligned} \tag{3.1}$$

则似然比检验统计量定义为

$$\begin{aligned} \lambda_n &= -2 \left\{ \sup_{\sigma\in\Theta_0} l_n(\sigma) - \sup_{\sigma\in\Theta_0\cup\Theta_1} l_n(\sigma) \right\} \\ &= 2 \sup_{\sigma\in\Theta_0} \left\{ n\log\sigma + \lambda^p/\sigma^p \sum_{i=1}^{n} |x_i|^p \right\} - 2 \sup_{\sigma\in\Theta_0\cup\Theta_1} \left\{ n\log\sigma + \lambda^p/\sigma^p \sum_{i=1}^{n} |x_i|^p \right\}. \end{aligned} \tag{3.2}$$

定理 3.1 记 σ^p 的极大似然估计为

$$\hat{\sigma}_n^p = \frac{p}{n}\lambda^p \sum_{i=1}^{n} |x_i|^p. \tag{3.3}$$

对于检验 $H_0 : \sigma = \sigma_0 \leftrightarrow H_1 : \sigma \neq \sigma_0$ 问题, 有

(1) 似然比检验统计量

$$\frac{\lambda_n}{n} = 2(\log\sigma_0 - \log\hat{\sigma}_n) + \frac{2\lambda^p}{n} \sum_{i=1}^{n} \left(\frac{|x_i|^p}{\sigma_0^p} - \frac{|x_i|^p}{\hat{\sigma}_n^p} \right) \xrightarrow{D} -\frac{2}{p}\log\xi, \tag{3.4}$$

其中 $\xi = \frac{p\lambda^p}{n} \sum_{i=1}^{n} \frac{|x_i|^p}{\sigma_0^p}$ 的分布为 $\chi^p(n)$.

(2) $-\frac{2n}{p}\log\xi$ 的密度函数为

$$f(y)=\frac{1}{2\Gamma(n/p)}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{-\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}-\frac{py}{2n}\right\}.$$

(3) 当 n 足够大时, λ_n 的近似密度函数为

$$g(y)=C^{-1}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{\frac{(n-p)y}{2n}\right\},-\infty<y<\infty.$$

其中 $C=\int_{-\infty}^{\infty}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{\frac{(n-p)y}{2n}\right\}dy.$

证 (1) 由 (3.2) 和 (3.3) 式得

$$\begin{aligned}\lambda_n&=2n\log\frac{\sigma_0}{\hat{\sigma}_n}+2\lambda^p\left\{\sum_{i=1}^{n}\left|\frac{x_i}{\sigma_0}\right|^p-\sum_{i=1}^{n}\left|\frac{x_i}{\hat{\sigma}_n}\right|^p\right\}\\&=2n\log\frac{\sigma_0}{\hat{\sigma}_n}+2\lambda^p\left\{\sum_{i=1}^{n}\left|\frac{x_i}{\sigma_0}\right|^p-\frac{n}{p}\left(\frac{\Gamma(3/p)}{\Gamma(1/p)}\right)^{-p/2}\right\}\\&=-2n\log\hat{\sigma}_n+2n\log\sigma_0+2\left(\frac{\Gamma(3/p)}{\Gamma(1/p)}\right)^{p/2}\sum_{i=1}^{n}\left|\frac{x_i}{\sigma_0}\right|^p-\frac{2n}{p}\\&=2\lambda^p\sum_{i=1}^{n}\left|\frac{x_i}{\sigma_0}\right|^p-\frac{2n}{p}-\frac{2n}{p}\log\left\{\sum_{i=1}^{n}\frac{|x_i|^p}{\sigma_0^p}\right\}-\frac{2n}{p}\log\left\{\frac{p}{n}\lambda^p\sigma_0^p\right\}+2n\log\sigma_0\\&=-\frac{2n}{p}\log\left\{\frac{p\lambda^p}{n}\sum_{i=1}^{n}\frac{|x_i|^p}{\sigma_0^p}\right\}+n\left(\frac{2\lambda^p}{n}\sum_{i=1}^{n}\frac{|x_i|^p}{\sigma_0^p}-\frac{2}{p}\right).\end{aligned}\tag{3.5}$$

注意到 $E\chi^p(n)=\frac{n}{p\lambda^p},Var\chi^p(n)=\frac{n}{p\lambda^{2p}}$(参见文献 [4]), 由大数定理容易证明:

$$\frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i}{\sigma_0}\right|^p\xrightarrow{P}E\left\{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i}{\sigma_0}\right|^p\right\}=\frac{1}{p\lambda^p},\tag{3.6}$$

由此得到

$$\frac{2\lambda^p}{n}\sum_{i=1}^{n}\left|\frac{x_i}{\sigma_0}\right|^p\xrightarrow{P}\frac{2}{p},\tag{3.7}$$

于是

$$\begin{aligned}\frac{\lambda_n}{n}&=-\frac{2}{p}\log\xi+2\lambda^p\left(\frac{1}{n}\sum_{i=1}^{n}\frac{|x_i|^p}{\sigma_0^p}-\frac{1}{p\lambda^p}\right)\\&=-\frac{2}{p}\log\xi+o_p(1)\\&\xrightarrow{D}-\frac{2}{p}\log\xi,\end{aligned}\tag{3.8}$$

其中 $\xi=\frac{p\lambda^p}{n}\sum\limits_{i=1}^{n}\frac{|x_i|^p}{\sigma_0^p}$ 的分布是 $\chi^p(n)$.

(2) 由 ξ 密度函数

$$h(x)=\frac{\lambda^n}{\Gamma(n/p)}\left(\frac{n}{p\lambda^p}\right)^{n/p}x^{n/p-1}\exp\left\{-\frac{nx}{p}\right\},\tag{3.9}$$

可得到 $-\frac{2n}{p}\log\xi$ 的密度函数为

$$\begin{aligned}f(y)&=\left(\frac{n}{p\lambda^p}\right)^{n/p}\frac{\lambda^n}{\Gamma(n/p)}\left(\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{-\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right\}\cdot\left(\frac{p}{2n}\right)\exp\left\{-\frac{py}{2n}\right\}\\&=\frac{1}{2\Gamma(n/p)}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{-\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}-\frac{py}{2n}\right\}.\end{aligned}$$

(3) 当 $n\to\infty$ 时, 有 $\frac{py}{2n}\to 0$ 和 $\exp\left\{-\frac{py}{2n}\right\}\approx 1-\frac{py}{2n}$, 从而有

$$\begin{aligned}\exp\left\{-\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}-\frac{py}{2n}\right\}&\approx\exp\left\{-\frac{n}{p}\left(1-\frac{py}{2n}\right)-\frac{py}{2n}\right\}\\&=\exp\left\{\frac{(n-p)y}{2n}-\frac{n}{p}\right\},\end{aligned}$$

于是可得

$$\begin{aligned}f(y)&=\frac{1}{2\Gamma(n/p)}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{-\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}-\frac{py}{2n}\right\}\\&\approx\frac{1}{2\Gamma(n/p)}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{\frac{(n-p)y}{2n}-\frac{n}{p}\right\}\\&=\frac{1}{2\Gamma(n/p)}\exp\left\{-\frac{n}{p}\right\}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{\frac{(n-p)y}{2n}\right\}\\&=C^{-1}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{\frac{(n-p)y}{2n}\right\},\end{aligned}\tag{3.10}$$

其中

$$C=\int_{-\infty}^{\infty}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{\frac{(n-p)y}{2n}\right\}dy,$$

因此, λ_n 的近似密度函数为

$$g(y)=C^{-1}\left(\frac{n}{p}\exp\left\{-\frac{py}{2n}\right\}\right)^{n/p-1}\exp\left\{\frac{(n-p)y}{2n}\right\},-\infty<y<\infty.$$

注 当 $|\lambda_n|$ 偏大时, 拒绝 H_0, 认为两总体方差有显著性差异; 当 $|\lambda_n|$ 接近 0 时, 接受 H_0, 认为两总体方差没有显著性差异.

注意到 $(\Gamma(3)/\Gamma(1))^{1/2}=\sqrt{2}$ 和 $(\Gamma(3/2)/\Gamma(1/2))^{1/2}=1/\sqrt{2}$, 由定理 3.1 容易得到下面的两个推论, 在此略去证明.

推论 3.1 (拉普拉斯分布) 令 $p=1$, $\hat{\sigma}_n=\frac{\sqrt{2}}{n}\sum\limits_{i=1}^{n}|x_i|$, 则 $\frac{\lambda_n}{n}\xrightarrow{D}-2\log\xi$, 其中 $\xi=\frac{\sqrt{2}}{n}\sum\limits_{i=1}^{n}\frac{|x_i|}{\sigma_0}$ 服从 $\chi(n)$. 且 λ_n 的近似密度函数为

$$f(y)=\frac{1}{2\Gamma(n)}\left(n\exp\left\{-\frac{y}{2n}\right\}\right)^{n-1}\exp\left\{-n\exp\left\{-\frac{y}{2n}\right\}-\frac{y}{2n}\right\}.$$

推论 3.2 (正态分布) 令 $p=2$, 则 $\frac{\lambda_n}{n}\xrightarrow{D}-\log\xi$, 其中 $\xi=\frac{1}{n}\sum\limits_{i=1}^{n}\frac{x_i^2}{\sigma_0^2}$ 服从的分布是 $\chi^2(n)$. 进而 λ_n 的近似密度函数为

$$f(y)=\frac{1}{2\Gamma(n/2)}\left(\frac{n}{2}\exp\left\{-\frac{y}{n}\right\}\right)^{n/2-1}\exp\left\{-\frac{n}{2}\exp\left\{-\frac{y}{n}\right\}-\frac{y}{n}\right\}.$$

3.2 双总体尺度参数的检验

本小节我们考虑两个总体的尺度参数的假设检验问题. 设随机样本 $x_1,x_2,\cdots,x_n$ 独立同分布, 取自参数为 μ_1,σ_1 的 P- 范分布; $y_1,y_2,\cdots,y_n$ 独立同分布, 均取自参数为 μ_1,σ_1 的 P- 范分布; 而且两个样本相互独立. 仍不妨假定 $\mu_1=0,\mu_0=0$. 检验如下问题

$$H_0:\sigma_1=\sigma_2\leftrightarrow H_1:\sigma_1\neq\sigma_2.$$

为了将 (3.5) 式适用于两个总体的尺度参数检验, 我们将它适当修改成如下似然比检验统计量

$$\begin{aligned}\lambda_n&=2\sup_{\sigma_1\in\tilde{\Theta}_1}\left\{n\log\sigma+\lambda^p/\sigma^p\sum_{i=1}^{n}|x_i|^p\right\}-2\sup_{\sigma_2\in\tilde{\Theta}_2}\left\{n\log\sigma+\lambda^p/\sigma^p\sum_{i=1}^{n}|x_i|^p\right\}\\&=2\left\{\log(\hat{\sigma}_{1n_1}^{n_1}/\hat{\sigma}_{2n_2}^{n_2})+\lambda^p\left[\sum_{i=1}^{n_1}\left|\frac{x_i}{\hat{\sigma}_{1n_1}}\right|^p-\sum_{i=1}^{n_2}\left|\frac{y_i}{\hat{\sigma}_{2n_2}}\right|^p\right]\right\}\\&=\frac{2}{p}\left\{\log(\hat{\sigma}_{1n_1}^{pn_1}/\hat{\sigma}_{2n_2}^{pn_2})+(n_1-n_2)\right\},\end{aligned}\tag{3.11}$$

其中 $\tilde{\Theta}_1$ 和 $\tilde{\Theta}_2$ 分别为第一、二个总体所对应的参数空间, σ_1^p 和 σ_2^p 对应的极大似然估计分别为

$$\hat{\sigma}_{1n_1}^p=\frac{p}{n_1}\lambda^p\sum_{i=1}^{n_1}|x_i|^p,\hat{\sigma}_{2n_2}^p=\frac{p}{n_2}\lambda^p\sum_{i=1}^{n_2}|y_i|^p,$$

定理 3.2 似然比统计量 $\lambda_n=\frac{2}{p}\left\{\log(\hat{\sigma}_{1n_1}^{pn_1}/\hat{\sigma}_{2n_2}^{pn_2})+(n_1-n_2)\right\}$ 的密度函数为

$$\begin{aligned}\tilde{f}(y)=&\frac{n_1^{n_1/p-1}n_2^{n_2/p-1}p^{1-(n_1+n_2)/p}}{2\Gamma(n_1/p)\Gamma(n_2/p)\sigma_1^p\sigma_2^p}\exp\left\{(1-p)\left[\frac{y}{2}-\frac{n_1-n_2}{p}\right]\right\}\\&\cdot\int_0^\infty z^{2/p-1}\exp\left\{-\frac{n_1\sqrt[n_1]{z}}{p\sigma_1^p}\exp\left\{\frac{py}{2n_1}-\left(1-\frac{n_2}{n_1}\right)\right\}-\frac{n_2\sqrt[n_2]{z}}{p\sigma_2^p}\right\}dz,-\infty<y<\infty.\end{aligned}\tag{3.12}$$

证 经计算容易得到 $\hat{\sigma}_n^p = \frac{p}{n}\left(\Gamma(3/p)/\Gamma(1/p)\right)^{p/2}\sigma^p\sum\limits_{i=1}^{n}\dfrac{|x_i|^p}{\sigma^p}$ 的密度函数如下:

$$\begin{aligned} f(y) &= \frac{n}{p}\frac{\lambda^{n-p}}{\Gamma(n/p)\sigma^p}\left(\frac{y}{\frac{p}{n}\lambda^p\sigma^p}\right)^{n/p-1}\exp\left\{-\lambda^p\frac{y}{\frac{p}{n}\lambda^p\sigma^p}\right\}, y>0 \\ &= \left(\frac{p}{n}\right)^{-n/p}\frac{1}{\Gamma(n/p)\sigma^p}y^{n/p-1}\exp\left\{-\frac{ny}{p\sigma^p}\right\}, y>0 \end{aligned} \tag{3.13}$$

由此可以得到 $\hat{\sigma}_n^{pn} = \left\{\frac{p}{n}\lambda^p\sigma^p\sum\limits_{i=1}^{n}\dfrac{|x_i|^p}{\sigma^p}\right\}^n$ 的密度函数为

$$\begin{aligned} h(y) &= \frac{1}{n}\left(\frac{p}{n}\right)^{-n/p}\frac{1}{\Gamma(n/p)\sigma^p}y^{n^{-1}(n/p-1)}y^{n^{-1}-1}\exp\left\{-\frac{ny^{n^{-1}}}{p\sigma^p}\right\} \\ &= \frac{1}{n}\left(\frac{p}{n}\right)^{-n/p}\frac{1}{\Gamma(n/p)\sigma^p}y^{1/p-1}\exp\left\{-\frac{ny^{n^{-1}}}{p\sigma^p}\right\}, y>0 \end{aligned} \tag{3.14}$$

由 (3.11) 和 (3.14) 式可得 $Z = \hat{\sigma}_{1n_1}^{pn_1}/\hat{\sigma}_{2n_2}^{pn_2}$ 的密度函数为

$$\begin{aligned} g(y) = &\int_{-\infty}^{\infty}|z|\frac{1}{n_2}\left(\frac{p}{n_2}\right)^{-n_2/p}\frac{1}{\Gamma(n_2/p)\sigma_2^p}z^{1/p-1}\exp\left\{-\frac{n_2z^{n_2^{-1}}}{p\sigma_2^p}\right\} \\ &\cdot\frac{1}{n_1}\left(\frac{p}{n_1}\right)^{-n_1/p}\frac{1}{\Gamma(n_1/p)\sigma_1^p}(zy)^{1/p-1}\exp\left\{-\frac{n_1(zy)^{n_1^{-1}}}{p\sigma_1^p}\right\}dz \\ = &\frac{n_1^{n_1/p-1}n_2^{n_2/p-1}p^{-(n_1+n_2)/p}}{\Gamma(n_1/p)\Gamma(n_2/p)\sigma_1^p\sigma_2^p}y^{1/p-1} \\ &\cdot\int_0^{\infty}z^{2/p-1}\exp\left\{-\frac{n_1(zy)^{n_1^{-1}}}{p\sigma_1^p}-\frac{n_2z^{n_2^{-1}}}{p\sigma_2^p}\right\}dz, y>0 \end{aligned} \tag{3.15}$$

由 (3.11) 和 (3.15) 式可得 $\lambda_n = \frac{2}{p}\left\{\log(\hat{\sigma}_{1n_1}^{pn_1}/\hat{\sigma}_{2n_2}^{pn_2})+(n_1-n_2)\right\}$ 的密度函数为

$$\begin{aligned} \tilde{f}(y) = &\frac{p}{2}\frac{n_1^{n_1/p-1}n_2^{n_2/p-1}p^{-(n_1+n_2)/p}}{\Gamma(n_1/p)\Gamma(n_2/p)\sigma_1^p\sigma_2^p}\exp\left\{\left(\frac{1}{p}-1\right)\left[\frac{py}{2}-(n_1-n_2)\right]\right\} \\ &\cdot\int_0^{\infty}z^{2/p-1}\exp\left\{-\frac{n_1z^{n_1^{-1}}}{p\sigma_1^p}\exp\left\{\frac{1}{n_1}(\frac{py}{2}-(n_1-n_2))\right\}-\frac{n_2z^{n_2^{-1}}}{p\sigma_2^p}\right\}dz \\ = &\frac{n_1^{n_1/p-1}n_2^{n_2/p-1}p^{1-(n_1+n_2)/p}}{2\Gamma(n_1/p)\Gamma(n_2/p)\sigma_1^p\sigma_2^p}\exp\left\{(1-p)\left[\frac{y}{2}-\frac{n_1-n_2}{p}\right]\right\} \\ &\cdot\int_0^{\infty}z^{2/p-1}\exp\left\{-\frac{n_1\sqrt[n_1]{z}}{p\sigma_1^p}\exp\left\{\frac{py}{2n_1}-\left(1-\frac{n_2}{n_1}\right)\right\}-\frac{n_2\sqrt[n_2]{z}}{p\sigma_2^p}\right\}dz. \end{aligned} \tag{3.16}$$

由于定理 3.2 的结论较复杂不便于应用, 所以下面我们考虑其特殊情形, 即 $n_1 = n_2 = n$. 利用定理 3.2 及其证明容易得到如下较简单的结论:

推论 3.3 若 H_0 成立, 且 $n_1=n_2=n$, 则 $\lambda_n=\frac{2n}{p}\log\left(\frac{\hat\sigma_{1n}}{\hat\sigma_{2n}}\right)^p$ 的密度函数为

$$\tilde f(y)=\frac{p}{2nB(n/p,n/p)}\exp\left\{\frac{y}{2}\left(1-\frac{n}{p}\right)\right\}\left[1+\exp\left(\frac{py}{2n}\right)\right]^{-2n/p},-\infty<y<\infty. \tag{3.17}$$

证 注意到

$$\left(\frac{\hat\sigma_{1n}}{\hat\sigma_{2n}}\right)^p=\frac{\lambda^p\frac{p}{n_1}\sigma_1^p\sum\limits_{i=1}^{n_1}\frac{|x_i-\mu_1|^p}{\sigma_1^p}}{\lambda^p\frac{p}{n_2}\sigma_2^p\sum\limits_{i=1}^{n_2}\frac{|y_i-\mu_2|^p}{\sigma_2^p}}=\frac{\sigma_1^p}{\sigma_2^p}\frac{n_2\sum\limits_{i=1}^{n_1}\frac{|x_i-\mu_1|^p}{\sigma_1^p}}{n_1\sum\limits_{i=1}^{n_2}\frac{|y_i-\mu_2|^p}{\sigma_2^p}}=\frac{n_2\sum\limits_{i=1}^{n_1}\frac{|x_i-\mu_1|^p}{\sigma_1^p}}{n_1\sum\limits_{i=1}^{n_2}\frac{|y_i-\mu_2|^p}{\sigma_2^p}}\sim F_p(n,n).$$

即 $(\hat\sigma_{1n}/\hat\sigma_{2n})^p$ 的密度函数为

$$f(z)=\frac{1}{B(n/p,n/p)}z^{n/p-1}(1+z)^{-2n/p},z>0 \tag{3.18}$$

于是由 (3.11) 和 (3.18) 式得 $\lambda_n=\frac{2n}{p}\log\left(\frac{\hat\sigma_{1n}}{\hat\sigma_{2n}}\right)^p$ 的密度函数为

$$\begin{aligned}\tilde f(y)&=\frac{1}{B(n/p,n/p)}z^{n/p-1}(1+z)^{-2n/p}\\&=\frac{p}{2nB(n/p,n/p)}\exp\left\{\frac{py}{2n}\left(\frac{n}{p}-1\right)\right\}\left[1+\exp\left(\frac{py}{2n}\right)\right]^{-2n/p}\\&=\frac{p}{2nB(n/p,n/p)}\exp\left\{\frac{y}{2}\left(1-\frac{n}{p}\right)\right\}\left[1+\exp\left(\frac{py}{2n}\right)\right]^{-2n/p},-\infty<y<\infty\end{aligned} \tag{3.19}$$

由推论 3.3 易得推论 3.4 和推论 3.5, 在此略去证明.

推论 3.4 令 $p=1$, 则在推论 3.3 的条件下, $\lambda_n=2n\log(\hat\sigma_{1n}/\hat\sigma_{2n})$ 的密度函数为

$$\tilde f(y)=\frac{1}{2nB(n,n)}\exp\left\{\frac{y}{2}\left(1-\frac{1}{n}\right)\right\}\left[1+\exp\left\{\frac{y}{2n}\right\}\right]^{-2n},-\infty<y<\infty$$

推论 3.5 令 $p=2$, 则在推论 3.3 的条件下, $\lambda_n=n\log(\hat\sigma_{1n}^2/\hat\sigma_{2n}^2)$ 的密度函数为

$$\tilde f(y)=\frac{2}{2nB(n/2,n/2)}\exp\left\{\frac{y}{2}\left(1-\frac{2}{n}\right)\right\}\left[1+\exp\left\{\frac{y}{n}\right\}\right]^{-n},-\infty<y<\infty$$

注 当 $|\lambda_n|$ 偏大时, 拒绝 H_0, 认为两总体方差有显著性差异; 当 $|\lambda_n|$ 接近 0 时, 接受 H_0, 认为两总体方差没有显著性差异.

4 应用举例

本节将选取不同的 p 值进行似然比检验. 由下面 3 个例子可以看出, 检验结果均与模拟数据相符, 从而说明了我们理论结果的可靠性.

例 1 在 (2.1) 式中令 $p=2,\mu=0,\sigma=2$, 随机产生 100 组样本, 其中样本容量为 100. 我们根据每组样本值求出统计量, 并将 100 个样本统计量求和取平均进行模拟实验 (显著性水平取为 $\alpha=0.05$).

(1) 检验 $H_0:\sigma=3\leftrightarrow H_1:\sigma\neq 3$.

根据样本可以计算得

$$\bar{\lambda}_n=2n(\log\sigma_0-\log\hat{\sigma}_n)+\lambda^p\sum_{i=1}^{n}\left(\frac{|x_i|^p}{\sigma_0^p}-\frac{|x_i|^p}{\hat{\sigma}_n^p}\right)=54.4694$$

由定理 3.1 中 (2) 式得

$$0.025=\int_{-\infty}^{q_1}f(y)dy,0.975=\int_{-\infty}^{q_2}f(y)dy$$

解得: $q_1=-25.8983,q_2=29.8111$. 由于 $\bar{\lambda}_n>q_2$, 并且在计算 100 个样本统计量时, 结果显示最小的统计量 λ_n 大于 q_2, 所以得到拒绝 H_0.

(2) 检验 $H_0:\sigma=2\leftrightarrow H_1:\sigma\neq 2$.

$$\bar{\lambda}_n=2n(\log\sigma_0-\log\hat{\sigma}_n)+\lambda^p\sum_{i=1}^{n}\left(\frac{|x_i|^p}{\sigma_0^p}-\frac{|x_i|^p}{\hat{\sigma}_n^p}\right)=0.23336$$

由于 $q_1<\bar{\lambda}_n<q_2$, 并且结果显示所求统计量 λ_n 的最大值和最小值在 (q_1,q_2), 所以接受 H_0.

例 2 在 (2.1) 式中令 $p=1,\mu_1=\mu_1=0,\sigma_1=2,\sigma_2=1$, 从两个总体中分别随机产生容量为 100 的样本, 检验 $H_0:\sigma_1=\sigma_2\leftrightarrow H_1:\sigma_1\neq\sigma_2$(显著性水平取为 $\alpha=0.05$).

根据样本计算得

$$\bar{\lambda}_n=\frac{2n}{p}\log\left(\frac{\hat{\sigma}_{1n}}{\hat{\sigma}_{2n}}\right)^p=69.6176$$

由推论 3.3 中 (3.17) 式得

$$0.025=\int_{-\infty}^{q_1}\tilde{f}(y)dy,0.975=\int_{-\infty}^{q_2}\tilde{f}(y)dy$$

其中 $q_1=-59.7589,q_2=47.4423$. 由于 $\bar{\lambda}_n>q_2$, 并且模拟结果显示, 统计量 λ_n 最小值大于 q_2, 所以拒绝 H_0, 认为两总体方差有显著性差异.

例 3 在 (2.1) 式中令 $p=4,\mu_1=\mu_1=0,\sigma_1=1,\sigma_2=1$, 随机产生容量为 100 的两个样本, 检验 $H_0:\sigma_1=\sigma_2\leftrightarrow H_1:\sigma_1\neq\sigma_2$(显著性水平取为 $\alpha=0.05$).

根据样本计算得

$$\bar{\lambda}_n=\frac{2n}{p}\log\left(\frac{\hat{\sigma}_{1n}}{\hat{\sigma}_{2n}}\right)^p=0.0503$$

由推论 3.3 中 (3.17) 式得

$$0.025=\int_{-\infty}^{q_3}\tilde{f}(y)dy,0.975=\int_{-\infty}^{q_4}\tilde{f}(y)dy$$

其中 $q_3=-32.4755,q_4=17.6155$. 由于 $q_3<\bar{\lambda}_n<q_4$, 并且结果显示所求统计量 λ_n 的最大值和最小值在 (q_1,q_2), 所以接受 H_0, 认为两总体方差没有显著性差异.

参 考 文 献

[1] Shah R D, B ü hlmann P. Goodness-of-fit tests for high dimensional linear models[J]. Journal of the Royal Statistical Society, Series B, 2018, 80: 113–135.

[2] Meinshausen N, Yu B. Lasso-type recovery of sparse representations for high-dimensional data[J]. The Annals of Statistics, 2009, 37(1): 246–270.

[3] 孙海燕, 胡宏昌. P- 范分布及其抽样分布 [J]. 应用概率统计, 2003,19(4): 424–428.

[4] 孙海燕. P- 范分布理论及其在现代测量数据处理中的应用 [D]. 武汉: 武汉测绘科技大学, 1995.

[5] 潘雄, 罗静, 汪耀. P 范分布的实数阶与对数矩估计法 [J]. 测绘学报, 2016, 45(3): 302–309.

[6] 童光荣, 李思维. 偏态 P 范分布刻画股指收益率的实证检验 [J]. 统计与决策, 2015, 8: 167–169.

[7] Giampaoli V, Singer J M. Likelihood ratio tests for variance components in linear mixed models[J]. Journal of Statistical Planning and Inference, 2009, 139: 1435–1448.

[8] Ferrari S L P, Cysneiros A H M A. Skovgaard's adjustment to likelihood ratio tests in exponential family nonlinear models[J]. Statistics and Probability Letters, 2008, 78: 3047–3055.

[9] Qi Y, Wang F, Zhang L. Limiting distributions of likelihood ratio test for independence of components for high-dimensional normal vectors[J]. Annals of the Institute of Statistical Mathematics, 2019, 71(4): 911–946.

[10] Huang C, Lin J, Ren Y Y. Testing for the shape parameter of generalized extreme value distribution based on the Lq-likelihood ratio statistic[J]. Metrika, 2013, 76: 641–671.

[11] Qin Y, Priebe C E. Robust hypothesis testing via Lq-likelihood[J]. Statistica Sinica, 2017, 27: 1793–1813.

[12] 张久军, 张慧, 杨瑞梅等. 基于似然比检验的监控多项式曲线的控制图 [J]. 数理统计与管理, 2018, 37(2): 289–297.

LIKELIHOOD RATIO TESTS FOR SCALE PARAMETER IN P-NORM DISTRIBUTION AND ITS APPLICATIONS

HU Hong-chang , REN Huan

(*School of Mathematics and Statistics, Hubei Normal University, Huangshi 435002, China*)

Abstract: P-norm distribution is a important distribution family, including Laplace distribution, normal distribution, uniform distribution, one-point distribution, etc. In this article, we consider the problem of testing scale parameter in P-norm distribution via likelihood ratio test method. Firstly, we introduce three sampling distributions, whose are closely related to P-norm distribution. Secondly, for a and two P-norm distributions, we investigate two hypothesis testing problems of unknown scale parameters by the likelihood ratio test method. We discuss the asymptotic distributions and (approximate) density functions of likelihood ratio test statistics, and also give the (approximate) density functions under some special cases such as Laplace distribution and normal distribution. At last, the validity of our method is illuminated by some numerical examples.

Keywords: P-norm distribution; Likelihood ratio tests; Scale parameter

2010 MR Subject Classification: 62F03