arXiv:2303.04402v1 [stat.ME] 8 Mar 2023

Goodness-of-fit tests for multivariate skewed distributions based on the characteristic function

Maicon J. Karling[a1], Marc G. Genton[a], Simos G. Meintanis[b, c2]

[a]沙特阿拉伯图瓦勒阿卜杜拉国王科技大学统计学项目[b]希腊雅典雅典国立卡波季斯特里亚大学经济学系[c]南非波切夫斯特鲁姆西北大学纯数学与应用分析学部

2023年3月9日

摘要。 我们采用一种通用的蒙特卡罗方法，对若干能够兼顾不对称性和重尾特征的常见多元模型进行拟合优度复合假设检验。具体而言，我们考虑基于差异度量的加权L2型检验，该度量涉及经验特征函数之间的距离，因此无需使用相应的总体量；这些总体量可能未知，也可能难以处理。我们的检验仅要求能够从待检验分布中抽取样本，并拥有一种合理的方法来估计未知的分布参数。我们通过蒙特卡罗研究，考察了检验准则在有限样本下对若干类偏斜分布的表现。此外，我们还提供了真实数据示例，以说明所用方法。

Keywords: 经验特征函数，拟合优度检验，重尾，偏态分布，偏斜正态分布，Tukey g-and-h 分布。

AMS 2020 classification numbers: 62F03，62H12，62H15。

1 引言

自 20 世纪 80 年代末以来，基于特征函数（CF）的 L2型拟合优度检验日益受到关注。主要原因在于，特征函数能够唯一确定其对应的分布，并且可以通过经验特征函数进行一致估计。对于多元分布而言，另一个优势是多元特征函数和经验特征函数定义良好且光滑；而累积分布函数及其经验对应量则不具备这些特点。因此，即使已知总体分布函数，使用前者也更为简便。

毫不意外，多元正态性的检验在这一领域占据重要地位（例如，参见 Chen 和 Genton [14], Ebner 等人 [15], Henze [23], Henze 等人 [24], Henze 和 Wagner [25],

[1]通讯作者：maicon.karling@kaust.edu.sa[2]目前在雅典大学休学术假。

和 Pude-lko [43]），因为有多种检验程序可供选择，其中包括基于特征函数的检验。然而，在多元高斯情形之外，基于特征函数的拟合优度程序仅适用于少数几种分布，其中大多数属于椭圆类（例如，参见 Fragiadakis 和 Meintanis [19], Meintanis 等人 [36], 以及 Sz´ekely 和 Rizzo [48]）。可用程序匮乏的主要原因之一是，尽管特征函数和经验特征函数都是平滑的，但通常仍需在 L^2情形下对它们进行数值积分；在高维情形下，这项任务可能颇具挑战。此外，对于大多数接受检验的多元分布，总体特征函数的解析形式可能完全未知。

近期，Chen 等人 [13] 提出了一种检验程序，该程序基于对待检验特征函数的蒙特卡洛近似，从而避免使用相应的总体量。然而，Chen 等人 [13], 所考察的椭圆分布族尽管可能十分重要，但其可呈现的形状范围有限，因此潜在应用受到限制；这是因为在经济学、金融学及大多数其他学科的实际数据分析中，通常除了超额峰度之外，还预期存在不对称性。

本文遵循 Chen 等人提出的方法 [13], ，但同时摒弃了其中采用的椭圆性设定，转而考虑更一般的形状。具体而言，我们针对若干常见的多元偏态分布族构造拟合优度检验。与 Chen 等人相比，实施这些 [13] 检验时还需处理一个额外问题：与该文提出的无参数检验不同，在当前设定下，形状参数的存在使得有必要增加一个重抽样循环，以便在给定参数配置的情况下复现检验统计量的经验分布。为此，只要待检验分布族的典范形式可用，我们便加以利用。在第2节中，我们将重新审视 Chen 等人工作的若干主要思想，并介绍本文检验的背景。 [13]

本文其余部分安排如下。第 3 节介绍如何通过自助重抽样实际实施新的检验程序。第4 节介绍并简要回顾将用于模拟和拟合优度检验的一系列分布族，并研究它们各自的典范形式。第5节通过一项大规模蒙特卡罗研究展示检验的有限样本性质。论文最后在第 6节给出若干实际数据应用，并在第 7 节讨论总体结果。在线补充材料包含一些额外的蒙特卡罗结果。

2 基于特征函数的检验

设 $\boldsymbol{X} \in \mathbb{R}^p$ $(p \geq 1)$ 是一个随机向量，其分布函数 $F_{\boldsymbol{X}}$ 绝对连续。我们关注由原假设表示的复合拟合优度检验问题

$\mathcal{H}_0$: $\boldsymbol{X} \in \mathcal{F}_{\vartheta}$ 的分布律，对于某个 $\vartheta \in \Theta$, (2.1)

其中 $\mathcal{F}_{\boldsymbol{\vartheta}} = \{F_{\boldsymbol{\vartheta}}, \boldsymbol{\vartheta} \in \Theta\}$ 表示一个特定的参数分布族，可用参数向量 $\boldsymbol{\vartheta}$进行参数化。相应的参数空间 Θ 将取为 $\mathbb{R}^q$ $(q \geq 1)$ 的一个开子集。鉴于特征函数的唯一性，我们也可以将式（2.1）中的原假设等价地表述为

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \varphi_{\boldsymbol{\vartheta}}(\boldsymbol{t}) \quad \forall\, \boldsymbol{t} \in \mathbb{R}^p, \text{ 对于某个 } \boldsymbol{\vartheta} \in \Theta, \tag{2.2}$$

其中，$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \mathbb{E}(e^{\mathrm{i}\boldsymbol{t}^\top \boldsymbol{X}})$表示 $\boldsymbol{X}$ 的特征函数，$\varphi_{\boldsymbol{\vartheta}}(\boldsymbol{t})$对应于族 $\mathcal{F}_{\boldsymbol{\vartheta}}$ 中某个随机向量的特征函数。这里，$\mathrm{i} = \sqrt{-1}$ 和 $\top$ 表示向量和矩阵的转置。

基于特征函数的 拟合优度统计量通常用 $\|\varphi_n - \varphi_{\widehat{\boldsymbol{\vartheta}}_n}\|_w^2$ 来表示，其中

$$\|f - g\|_w^2 := \int_{\mathbb{R}^p} |f(\boldsymbol{t}) - g(\boldsymbol{t})|^2 \, w(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} \tag{2.3}$$

是复值函数对 (f, g) 之间的一种 L2 型加权距离，

$$\varphi_n(\boldsymbol{t}) = \frac{1}{n} \sum_{j=1}^{n} e^{\mathrm{i}\boldsymbol{t}^\top \boldsymbol{X}_j} \tag{2.4}$$

是根据 $\boldsymbol{X}$ 的一组独立同分布（i.i.d.）副本 $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ 计算得到的经验特征函数，而 $\varphi_{\widehat{\boldsymbol{\vartheta}}_n}$ 是对应于原假设 $\mathcal{H}_0$ 的特征函数，其中$\widehat{\quad}$ $\widehat{\quad}$参数 $\boldsymbol{\vartheta}$ 被估计量 $\boldsymbol{\vartheta}_n$ 替代：$= \boldsymbol{\vartheta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$。权函数 $w > 0$ 将在下文进一步说明。

不过，确实存在一些分布情形（本文将讨论其中的一些），其原假设特征函数 $\varphi_{\boldsymbol{\vartheta}}(\cdot)$ 要么完全未知，要么过于复杂而难以处理。在这种情况下，我们建议以类似的方式构造检验统计量，但不直接引用待检验分布的特征函数。具体而言，按照 Chen 等人 [13], 的做法，我们建议使用以下检验统计量

$$T_{n,m}^{(w)} = \|\varphi_n - \widehat{\varphi}_{0,m}\|_w^2, \tag{2.5}$$

其中 $\varphi_n(\cdot)$ 与 (2.4) 中的定义相同，而

$$\widehat{\varphi}_{0,m}(\boldsymbol{t}) = \frac{1}{m} \sum_{j=1}^{m} e^{\mathrm{i}\boldsymbol{t}^\top \boldsymbol{X}_{0,j}} \tag{2.6}$$

是根据样本 $(\boldsymbol{X}_{0,1}, \ldots\ldots, \boldsymbol{X}_{0,m})$ 计算得到的经验特征函数，该样本取自 $F_{\widehat{\boldsymbol{\vartheta}}_n}$，即从待检验分布中抽取的样本量为 m $(m \geq n)$ 的样本，其参数由相合估计量 $\boldsymbol{\vartheta}_n := \boldsymbol{\vartheta}_n(\boldsymbol{X}_1, \ldots\ldots, \boldsymbol{X}_n)$ 估计。换言之，$\widehat{\varphi}_{0,m}$ 是原假设特征函数 $\varphi_{\boldsymbol{\vartheta}}(\cdot)$ 的蒙特卡洛近似。对于较大的 $T_{n,m}^{(w)}$，我们拒绝原假设。

使用检验统计量 $T_{n,m}^{(w)}$ 的一个明显优点是计算简便。为此，记 $|\cdot|$ 为复数的模，随后通过标准代数运算可得

$$\begin{aligned} |\varphi_n(\boldsymbol{t}) - \widehat{\varphi}_{0,m}(\boldsymbol{t})|^2 \quad = \quad & \frac{1}{n^2} \sum_{j,k=1}^{n} \cos \boldsymbol{t}^\top (\boldsymbol{X}_j - \boldsymbol{X}_k) + \frac{1}{m^2} \sum_{j,k=1}^{m} \cos \boldsymbol{t}^\top (\boldsymbol{X}_{0,j} - \boldsymbol{X}_{0,k}) \\ & - \frac{2}{nm} \sum_{j=1}^{n} \sum_{k=1}^{m} \cos \boldsymbol{t}^\top (\boldsymbol{X}_j - \boldsymbol{X}_{0,k}). \end{aligned}$$

现在假设，式 (2.3) 和 (2.5) 中的权函数 $w(\cdot)$ 取为服从某种球对称分布的随机向量 $\boldsymbol{W} \in \mathbb{R}^p$ 的密度。众所周知，$\boldsymbol{W}$的特征函数可简化为 $\varphi_{\boldsymbol{W}}(\boldsymbol{t}) = \mathbb{E}(\cos \boldsymbol{t}^\top \boldsymbol{W})$，最终可表示为 $\Psi(\|\boldsymbol{t}\|^2)$，其中 $\Psi(\cdot)$ 称为与 $\boldsymbol{W}$ 相关联的“核”，而 $\|\cdot\|$ 表示 $\mathbb{R}^p$ 中的标准欧几里得范数（见 Fang 等人 [16]）。将最后一个等式以及式 (2.5) 中这样的权函数 $w(\cdot)$ 代入，可得检验统计量

$$
\begin{aligned}
(2.7) \quad T_{n,m}^{(\Psi)} \;=\; & \frac{1}{n^2}\sum_{j,k=1}^{n}\Psi(\|\boldsymbol{X}_j-\boldsymbol{X}_k\|^2)+\frac{1}{m^2}\sum_{j,k=1}^{m}\Psi(\|\boldsymbol{X}_{0,j}-\boldsymbol{X}_{0,k}\|^2) \\
& -\frac{2}{nm}\sum_{j=1}^{n}\sum_{k=1}^{m}\Psi(\|\boldsymbol{X}_j-\boldsymbol{X}_{0,k}\|^2),
\end{aligned}
$$

其中，我们已明确表示检验统计量对核的依赖关系 Ψ。只要核 $\Psi(\cdot)$足够简单，便可直接计算（2.7）。我们可以使用的一些典型简单核包括：

- 标准正态核 $\Psi(\xi) = e^{-\xi/2}$;
- 源自稳定分布的核 $\Psi(\xi) = e^{-\xi^{b/2}}$, $b \in (0,2]$（参见 Nolan[40]）；
- 以及广义拉普拉斯核 $\Psi(\xi) = (1+\xi)^{-b}, \quad b > 0$ （参见 Kozubowski 等人 [32]）。

在本文中，我们将仅使用标准正态核来进行检验。有关核的更深入讨论，请参见 Micchelli 等人 [37]。

3 重抽样检验实施

当 $\boldsymbol{\vartheta}$ 中出现的某些分量参数可以通过标准化消除时，式 (2.5) 中所提出的检验统计量 $\widehat{T}_n^{(\Psi)},_m$ 的渐近零分布不依赖于这些参数。这类参数通常是位置参数和散布参数；而其他参数（称为形状参数），例如偏度和峰度，通常无法通过标准化消除，因此最终会影响检验统计量的渐近零分布（见 Meintanis 和 Swanepoel [34]）。在这种情况下，我们可以将参数向量分解为 $\boldsymbol{\vartheta} = (\boldsymbol{\theta}, \boldsymbol{\lambda})$，其中 $\boldsymbol{\theta}$ 表示非形状参数，而 $\boldsymbol{\lambda}$ 表示 $\boldsymbol{\vartheta}$ 中仅包含形状参数的部分。若待检验分布具有典范形式（见第 4 节），则可通过设定 $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ 来模拟检验统计量的渐近零分布，其中 $\boldsymbol{\theta}_0$ 是 $\boldsymbol{\theta}$ 的某个标准值，而 $\boldsymbol{\lambda}$ 设为其在典范形式中的值 $\sim \quad \sim \boldsymbol{\lambda} := \boldsymbol{\lambda}(\boldsymbol{\lambda}, \boldsymbol{\theta})$。

下面，我们概述本文所用的重抽样过程，以近似原假设下检验统计量的渐近分布，并说明如何在实践中进行检验。为明确起见，对于固定的 (n, m, Ψ)，我们将 (2.7) 中的检验统计量写作 $T(\mathcal{X}_n; \mathcal{X}_{0,m})$，其中 $\mathcal{X}_n = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ 表示观测数据，而 $\mathcal{X}_{0,m} = (\boldsymbol{X}_{0,1}, \ldots, \boldsymbol{X}_{0,m})$

表示由 $\mathcal{F}_{\widehat{\vartheta}_n}$ 生成的数据，即来自带有估计参数的零分布的数据。我们考虑两种原假设情形：一种是所有参数均需估计的“复合”情形，另一种称为“简单”情形，尽管在后一种情形中，也有部分参数需要估计，但并非所有参数都需要估计。

3.1 简单原假设

在此，我们关注与简单原假设相关的拟合优度检验问题

$$\mathcal{H}_0^{\mathrm{s}}: \boldsymbol{X} \in \mathcal{F}_{\boldsymbol{\theta}} \text{ 的分布律, } \boldsymbol{\lambda}, \text{ 对于固定的 } \boldsymbol{\lambda}=\boldsymbol{\lambda}_0, \text{ 以及某个 } \boldsymbol{\theta}; \quad (3.1)$$

以及备择假设

$$\mathcal{H}_1^{\mathrm{s}}: \boldsymbol{X} \notin \mathcal{F}_{\boldsymbol{\theta}} \text{ 的分布律, } \boldsymbol{\lambda}, \text{ 对于固定的 } \boldsymbol{\lambda}=\boldsymbol{\lambda}_0, \text{ 以及任意 } \boldsymbol{\theta}. \quad (3.2)$$

尽管我们将 (3.1) 标记为简单原假设，但需要指出的是，参数$\boldsymbol{\theta}$ 在 $\mathcal{H}_0^{\mathrm{s}}$中未作具体说明，而且我们的检验程序包含对该参数的估计步骤。对于这种情况，临界值的计算基于从原假设中所涉及的分布律进行的简单蒙特卡洛抽样。该次蒙特卡罗运行的步骤如下：

步骤 1 - 从 $\mathcal{F}_{\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0}$ 生成一个随机样本 $\mathcal{X}_n=\{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n\}$，计算估计值 $\widehat{\boldsymbol{\theta}}_n$，并得到标准化样本 $\widehat{\mathcal{X}}_n=(\widehat{\boldsymbol{X}}_1,\ldots,\widehat{\boldsymbol{X}}_n)$，其中 $\widehat{\boldsymbol{X}}_j=\boldsymbol{X}_j(\boldsymbol{X}_j,\widehat{\boldsymbol{\theta}}_n)$ $j\in\{1,\ldots,n\}$。

步骤 2 - 从 $\mathcal{F}_{\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0}$ 生成一个随机样本 $\mathcal{X}_{0,m}=\{\boldsymbol{X}_{0,1},\ldots,\boldsymbol{X}_{0,m}\}$。

步骤 3 - 根据（ $T:=T^{\Psi}(\widehat{\mathcal{X}}_n,\mathcal{X}_{0,m})$ ，按照（2.7）计算检验统计量。

步骤 4 - 重复步骤 1–3 若干次，设为 M，并得到检验统计量集合 $\{T_1,\ldots,T_M\}$。然后，临界值设为 $\widehat{c}_\delta$，定义为（$T_m,\ m=1,\ldots,M$）的（$1-\delta$）% 分位数。

获得经验临界值后， $\widehat{c}_\delta$ 用于计算检验的经验检验功效。在此过程中，我们从备择假设 $\mathcal{H}_1^{\mathrm{s}}$中的备择分布集合内任意一个分布生成随机样本 $\mathcal{X}_n=\{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n\}$，并执行上述步骤 1–3，从而计算检验统计量 T。若 $T>\widehat{c}_\delta$，则拒绝原假设 $\mathcal{H}_0^{\mathrm{s}}$。我们重复此过程若干次，例如 L次，并将经验检验功效率计算为 $L^{-1}\sum_{\ell=1}^{L}\mathbb{1}_{T_\ell>\widehat{c}_\delta}$，其中 T_ℓ 表示第 ℓ^{th} 个样本对应的检验统计量， $\ell\in\{1,\ldots,L\}$。

3.2 复合原假设

这里我们关注（完全）复合拟合优度检验问题，即所有分布参数均由观测数据估计。为明确起见，我们将原假设表述为

$$\mathcal{H}_0^{\mathrm{c}}: \boldsymbol{X} \in \mathcal{F}_{\boldsymbol{\theta}} \text{的分布律, } \boldsymbol{\lambda}, \text{ 对某个（} \boldsymbol{\theta}, \ \boldsymbol{\lambda}\text{）；} \quad (3.3)$$

以及备择假设

$$\mathcal{H}_1^c: \boldsymbol{X} \notin \mathcal{F}_{\boldsymbol{\theta}} \text{的分布律，} \lambda \text{，对于任意的} (\boldsymbol{\theta}, \boldsymbol{\lambda}) \text{。} \quad (3.4)$$

对于这种情况，重新抽样方案如下。根据 $\mathcal{X}_n$，计算 $\boldsymbol{\vartheta}$ 的估计量 $\widehat{\boldsymbol{\vartheta}}_n = (\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\lambda}}_n)$，并将观测值标准化为 $\widehat{\boldsymbol{X}}_j = \widehat{\boldsymbol{X}}_j(\boldsymbol{X}_j, \widehat{\boldsymbol{\theta}}_n)$，$j \in \{1, \ldots, n\}$。现在在原假设下生成一个随机样本 $\widetilde{\mathcal{X}}_{0,m} := (\widetilde{\boldsymbol{X}}_{0,1}, \ldots, \widetilde{\boldsymbol{X}}_{0,m})$，其中 $(\widetilde{\boldsymbol{\theta}}_0, \widehat{\boldsymbol{\lambda}}_n)$，而 $\widehat{\boldsymbol{\lambda}}_n = \widehat{\boldsymbol{\lambda}}_n(\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\lambda}}_n)$是由参数化所导出的 $\boldsymbol{\lambda}$ 的参数估计值。随后，根据（2.7）计算原始检验统计量的值：$T = T(\mathcal{X}_n; \widetilde{\mathcal{X}}_{0,m})$，其中 $\mathcal{X}_n := (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$。接着，使用参数自助法程序计算用于与 T 的值进行比较的临界值，具体步骤如下：

步骤1：在原假设下生成随机样本 $\mathcal{X}_n^*$：=（$\boldsymbol{X}_1^*$，……，$\boldsymbol{X}_n^*$），其中（$\boldsymbol{\theta}$，$\boldsymbol{\lambda}$）设为 $\widetilde{}$（$\boldsymbol{\theta}_0$，$\boldsymbol{\lambda}_n$）。

步骤2：根据 $\mathcal{X}_n^*$，计算估计量 $\widehat{\boldsymbol{\vartheta}}_n^* =$（$\widehat{\boldsymbol{\theta}}_n^*$，$\widehat{\boldsymbol{\lambda}}_n^*$）。

步骤3：将 $\mathcal{X}_n^*$ 的各分量标准化为 $\widehat{\boldsymbol{X}}_j^* = \widehat{\boldsymbol{X}}_j^*$（$\boldsymbol{X}_j^*$，$\widehat{\boldsymbol{\theta}}_n^*$），$j \in \{1, \ldots, n\}$。

步骤4：在原假设下生成随机样本 $\widehat{\mathcal{X}}_{0,m}^*$：=（$\widehat{\boldsymbol{X}}_{0,1}^*$，……，$\widehat{\boldsymbol{X}}_{0,m}^*$），其中（$\boldsymbol{\theta}$，$\boldsymbol{\lambda}$）设为（$\boldsymbol{\theta}_0$，$\widehat{\widehat{\boldsymbol{\lambda}}}_n^*$），这里 $\widehat{\widehat{\boldsymbol{\lambda}}}_n^* = \widehat{\widehat{\boldsymbol{\lambda}}}^*$（$\widehat{\boldsymbol{\theta}}_n^*$，$\widehat{\boldsymbol{\lambda}}_n^*$）是 $\widehat{\widehat{\boldsymbol{\lambda}}}_n$ 的自助法参数估计值。

步骤 5：根据 (2.7)，计算自助检验统计量的值，记为 $T^* = T(\mathcal{X}_n^*; \mathcal{X}_{0,m}^*)$，其中 $\mathcal{X}_n^*$:=($\widehat{\boldsymbol{X}}_1^*$，…，$\widehat{\boldsymbol{X}}_n^*$)。

步骤 6：重复步骤 1–5 若干次，设为 B，从而计算 $(1-\delta)$% 分位数 c_δ，其中 $\delta \in (0, 1)$，作为检验统计量的样本量为 δ 时的临界值；该分位数来自 $(T_b^*, b = 1, \ldots, B)$ 的经验分布。

步骤 7：重复步骤 1–6 若干次，设为 M，从而获得检验统计量及其对应的自助法临界值组成的配对 $(T_m, c_{\delta,m}), m \in \{1, \ldots, M\}$。

步骤 8：计算经验拒绝率：$M^{-1}\sum_{m=1}^{M} \mathbb{1}_{T_m > c_{\delta,m}}$。

由于上述参数自助法程序耗时较长，我们采用 Giacomini 等人提出的快速自助法 。因此，我们不再为每个 M 蒙特卡罗样本计算临界值$c_{\delta,m}$ 而是计算一个适用于所有蒙特卡罗样本的临界值。为此，对于每个 M蒙特卡罗样本，我们只生成一个自助样本，即在步骤 6 中令 $B = 1$，并根据该自助样本计算相应的自助检验统计量，记为 T_m^*。然后，按照与上述步骤 6 类似的方式，根据 $(T_m^*, m \in \{1, \ldots, M\})$ 计算快速自助法临界值，记为 $\widetilde{c}_\delta$，经验拒绝率则为 $M^{-1}\sum_{m=1}^{M} \mathbb{1}_{T_m > \widetilde{c}_\delta}$。

4 偏态分布族

本节将考察五类偏斜分布族，并举例说明如何应用第 2 和3节所述的方法，对这些分布进行拟合优度检验。在接下来的第 5 和6节中，我们将分别在模拟研究和真实数据集应用中使用这五类分布族。

4.1 多元偏态正态分布

多元偏斜正态分布（SN）可通过特征函数（CF）方便地定义（参见 Azzalini 和 Dalla Valle [6]）。

$$\varphi_{\boldsymbol{\vartheta}}(\boldsymbol{t}) = 2e^{\mathrm{i}\boldsymbol{t}^\top\boldsymbol{\xi}-\frac{1}{2}\boldsymbol{t}^\top\boldsymbol{\Omega}\boldsymbol{t}}\,\Phi\left(\mathrm{i}\frac{\boldsymbol{\alpha}^\top\boldsymbol{\Omega}\boldsymbol{t}}{\sqrt{1+\boldsymbol{\alpha}^\top\boldsymbol{\Omega}\boldsymbol{\alpha}}}\right), \tag{4.1}$$

其中，$\Phi(\cdot)$ 是标准正态累积分布函数，$\boldsymbol{\vartheta} := (\boldsymbol{\xi}, \Omega\ \boldsymbol{\alpha})$ 是相应的参数向量，其中 $(\boldsymbol{\xi},\ \boldsymbol{\alpha}) \in \mathbb{R}^p \times \mathbb{R}^p$ 分别为位置参数和偏度参数，且 $\Omega \in \mathbb{R}^{p\times p}$ 是一个对称正定矩阵。我们将该分布记为 SN$_p(\boldsymbol{\xi},\ \Omega\ \boldsymbol{\alpha})$，其中 $\boldsymbol{\alpha} = \mathbf{0}$ 表示均值为 $\boldsymbol{\xi}$、协方差矩阵为 Ω 的 p元正态分布。应用我们的检验需要 $\boldsymbol{\vartheta}$ 中参数的相合估计量。估计这些参数的方法有多种，包括极大似然估计（MLE）和基于矩的方法（例如，参见 Azzalini and Capitanio [4]、Azzalini 等人 [8]，和 Flecher 等人 [18]），也有可用于此目的的软件包（参见 Azzalini [3]）。此外，针对这种情形也有一些拟合优度检验，即 Balakrishnan 等人 [9]、Gonz´alez-Estrada 等人[21]、Jim´enez-Gamero 和 Kim [27]，以及 Meintanis 和 Hl´avka [35] 提出的检验；本文稍后将进一步讨论这些检验（见第5.4小节）。

可以证明，若 $\boldsymbol{X} \sim$ SN$_p(\boldsymbol{\xi}, \Omega\ \boldsymbol{\alpha})$，则存在矩阵 $\mathrm{H} \in \mathbb{R}^{p\times p}$，使得 $\mathrm{H}^\top(\boldsymbol{X}-\boldsymbol{\xi}) \sim$ SN$_p(0, \mathrm{I}\ \boldsymbol{\alpha}^*)$，其中 0 和 I 分别表示相应维数下的零向量和单位矩阵，且 $\boldsymbol{\alpha}^* = (\alpha^*, 0, \ldots, 0)^\top$，其中 $\alpha^* = (\boldsymbol{\alpha}^\top \bar{\Omega}\boldsymbol{\alpha})^{1/2}$，而 $\Omega = \boldsymbol{\omega}\bar{\Omega}\boldsymbol{\omega}$，其中$\boldsymbol{\omega} = \mathrm{diag}(\omega_1, \ldots, \omega_p)$ 是正定尺度矩阵（见 Azzalini and Capitanio [5] 的公式 (5.2)）。在文献中（例如，见 Azzalini and Capitanio [5] 和 Capitanio [12]），SN$_p(0, \mathrm{I}\ \boldsymbol{\alpha}^*)$ 分布也称为典范形式。此外，Capitanio [12] 证明，选择

$$\mathbf{H} = \boldsymbol{\Omega}^{-1/2}\mathbf{Q}, \tag{4.2}$$

其中 $\Omega^{-1/2}$ 是 Ω 的正定对称平方根矩阵的唯一逆矩阵，而 Q 通过谱分解 $\mathrm{Q}\Lambda\mathrm{Q}^\top = \Omega^{-1/2}\Sigma\Omega^{-1/2}$得到，其中 $\boldsymbol{\Sigma}$ 为 $\boldsymbol{X}$的协方差矩阵，由此可得 $\mathrm{H}^\top(\boldsymbol{X}-\boldsymbol{\xi})$服从典范偏斜正态$\widehat{\ }$ $\widehat{\ }$ $\widehat{\ }$分布。在此背景下，用 $\boldsymbol{\vartheta}_n = (\boldsymbol{\xi}_n, \Omega_n, \widehat{\boldsymbol{\alpha}}_n)$表示 $\boldsymbol{\vartheta}$的一个估计量，并考虑标准化观测值 $\widehat{\boldsymbol{X}}_j = \widehat{\ }\mathrm{H}_n^\top(\boldsymbol{X}_j - \widehat{\boldsymbol{\xi}}_n)$ $j \in \{1, \ldots, n\}$，其中 $\widehat{\ }\mathrm{H}_n = \widehat{\ }\Omega_n^{-1/2}\widehat{\ }\mathrm{Q}_n$。然后，只需将 $\boldsymbol{X}_j$ 替换为 $\boldsymbol{X}_j$，即可直接应用式 (2.7) 中的检验$\widehat{\ }$；其中 $\boldsymbol{X}_{0,j}$ 取自参数为 $(\boldsymbol{\xi}, \Omega) = (0, \mathrm{I})$ 的 SN 分布，而 $\boldsymbol{\alpha}^*$则替换为 $\widehat{\boldsymbol{\alpha}}_n^* = (\widehat{\alpha}_n^*, 0, \ldots, 0)^\top$，

其中 $\widehat{\alpha}_n^* = (\widehat{\boldsymbol{\alpha}}_n^\top \widehat{\boldsymbol{\Omega}}_n \widehat{\boldsymbol{\alpha}}_n)^{1/2}$。为了获得 $\widehat{\boldsymbol{\xi}}_n$， $\widehat{\mathbf{H}}_n$和 $\widehat{\alpha}_n^*$的估计值，我们建议使用`sn` [3] 软件包，它运行于R [44] 软件环境中。

4.2 多变量偏斜 t 分布

多变量偏斜 t 分布（ST）与多元偏斜正态分布通过随机方程 $\boldsymbol{Y} = \boldsymbol{\xi} + \sqrt{\eta}\boldsymbol{X}$ 相关，其中 $\boldsymbol{X}$ 服从多元偏斜正态分布，$\boldsymbol{X} \sim \mathrm{SN}_p(\mathbf{0}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$，而$^{-}\eta$ 服从形状参数和尺度参数均为 $\nu/2$ 的逆伽马分布，即 $\eta \sim \mathrm{IG}(\nu/2, \nu/2)$。Kim 和 Genton [30]，定理 7 证明，$\boldsymbol{Y}$ 的特征函数为

$$\varphi_{\boldsymbol{\vartheta}}(\boldsymbol{t}) = \exp(\mathrm{i}\boldsymbol{t}^\top \boldsymbol{\xi})[\psi_{T_p}(\boldsymbol{\Omega}^{1/2}\boldsymbol{t}) + \mathrm{i}\tau^+(\boldsymbol{\rho}, \omega\boldsymbol{t})], \tag{4.3}$$

其中

$$\begin{aligned}
\boldsymbol{\rho} &= \boldsymbol{\Omega}\boldsymbol{\alpha}/(1 + \boldsymbol{\alpha}^\top \boldsymbol{\Omega}\boldsymbol{\alpha})^{1/2}, \\
\psi_{T_p}(\boldsymbol{t}) &= \frac{\|\sqrt{\nu}\boldsymbol{t}\|^{\nu/2}}{\Gamma(\nu/2)\, 2^{\nu/2-1}}\, K_{\nu/2}(\|\sqrt{\nu}\boldsymbol{t}\|), \quad \text{for } \boldsymbol{t} \in \mathbb{R}^p,\ \nu > 0, \\
\tau^+(\boldsymbol{\rho}, \omega\boldsymbol{t}) &= \int_0^\infty \exp(-x\boldsymbol{t}^\top \boldsymbol{\Omega}\boldsymbol{t}/2)\, \tau(\sqrt{x}\boldsymbol{\rho}^\top \omega\boldsymbol{t})\, \mathrm{d}H(x), \quad \text{for } \boldsymbol{\rho}^\top \omega\boldsymbol{t} > 0,
\end{aligned}$$

其中 $\tau^+(\boldsymbol{\rho}, -\omega\boldsymbol{t}) = -\tau^+(\boldsymbol{\rho}, \omega\boldsymbol{t})$， $\tau(x) = \int_0^x \sqrt{2/\pi}\exp(u^2/2)\,\mathrm{d}u$，对于 $x > 0$， $\tau(-x) = -\tau(x)$，$H(x) = \Gamma(\nu/2, \nu/(2x))/\Gamma(\nu/2)$，对于 $x > 0$，表示 η 的累积分布函数；$\Gamma(a) = \Gamma(a, 0)$，且其中 $\Gamma(a, b) = \int_b^\infty t^{a-1}e^{-t}\mathrm{d}t$，对于 $b \geq 0$，表示上不完全伽马函数； $K_\lambda(\cdot)$ 是第三类修正贝塞尔函数的积分表示，其定义为 $K_\lambda(w) = \frac{1}{2}\int_0^\infty x^{\lambda-1}\exp\left\{-\frac{w}{2}\left(x + \frac{1}{x}\right)\right\}\mathrm{d}x$，对于 $w > 0$ 和 $\lambda \in \mathbb{R}$。这里 $\boldsymbol{\vartheta} = (\boldsymbol{\xi}, \Omega\ \boldsymbol{\alpha}, \nu)$，并记作 $\boldsymbol{Y} \sim \mathrm{ST}_p(\boldsymbol{\xi}, \Omega\ \boldsymbol{\alpha}, \nu)$。

由于多元偏斜 t 分布可以表示为偏态正态分布的尺度混合，Capitanio [12] 中的命题 2保证，再次采用式 (4.2) 中定义的矩阵 H，便可将任意随机向量 $\boldsymbol{Y} \sim \mathrm{ST}_p(\boldsymbol{\xi}, \Omega\ \boldsymbol{\alpha}, \nu)$转换为典范偏斜 t 分布 $\mathbf{H}^\top \boldsymbol{Y} \sim \mathrm{ST}_p(0, \mathrm{I}\ \boldsymbol{\alpha}_{\boldsymbol{Y}}^*, \nu)$。特别地，与偏斜正态情形一样， $\boldsymbol{\xi}$ 和 Ω 是干扰参数，因此在完成标准化后，可在假设检验中将其忽略。此外，对于第 5 节的模拟研究，只需针对不同的 p、 ν 和 α^*（ $\boldsymbol{\alpha}_{\boldsymbol{Y}}^* = (\alpha^*, 0, \ldots, 0)^\top$ 中唯一非零的分量）实施检验，从而大幅减少需要详细研究的情形。这里也可以使用 R [44] 中的 sn [3] 包来获取所需的多元偏斜 t 分布参数估计值，因为该包的设计也考虑了这一用途。随后即可按照前一小节讨论的偏斜正态情形进行假设检验。

4.3 多元偏斜拉普拉斯分布

多元偏斜拉普拉斯（SL）分布可通过特征函数方便地定义（参见 Arslan [2]）

$$\varphi_{\boldsymbol{\vartheta}}(\boldsymbol{t}) = \frac{e^{\mathrm{i}\boldsymbol{t}^\top \boldsymbol{\xi}}}{(1+\boldsymbol{t}^\top \boldsymbol{\Omega}\,\boldsymbol{t} - 2\mathrm{i}\boldsymbol{t}^\top \boldsymbol{\alpha})^{(p+1)/2}}, \tag{4.4}$$

其参数为 $\boldsymbol{\vartheta} = (\boldsymbol{\xi},\ \boldsymbol{\Omega}\ \boldsymbol{\alpha})$，其中（$\boldsymbol{\xi}$，$\boldsymbol{\alpha}$）$\in \mathbb{R}^p \times \mathbb{R}^p$ 分别为位置参数和偏度参数，而$\boldsymbol{\Omega}$ 为对称正定矩阵。我们将使用记号 SL$_p(\boldsymbol{\xi},\ \boldsymbol{\Omega}\ \boldsymbol{\alpha})$表示该分布。尽管 Arslan [2] 提出的多元偏斜拉普拉斯分布与 Kotz 等人 [31], 引入的另一种版本具有非常相似的性质，但 Arslan 的替代版本具有更简单的概率密度函数，因此在多元情形下可以采用简便的参数估计方法。此外，Arslan [2] 提出了一种高效的 EM 算法，可用于估计$\boldsymbol{\xi}$，$\boldsymbol{\Omega}$，和 $\boldsymbol{\alpha}$。

与偏斜正态分布和偏斜 t 分布类似，在满足 $\boldsymbol{\alpha} \neq \mathbf{0}$的条件下，可以证明 $\boldsymbol{X} \sim$ SL$_p(\boldsymbol{\xi},\ \boldsymbol{\Omega}\ \boldsymbol{\alpha})$可化为典范形式。这一新颖结果减轻了对干扰参数进行检验的负担，同时也将偏度简化为一个单维分量。

命题 4.1（典范形式 - SL$_p$ 分布）。设 $\boldsymbol{X} \sim$ SL$_p(\boldsymbol{\xi},\boldsymbol{\Omega}\ \boldsymbol{\alpha})$，且 $\boldsymbol{\alpha} \neq \mathbf{0}$，考虑仿射非奇异变换

$$\boldsymbol{X}^* = \mathbf{H}^\top(\boldsymbol{X} - \boldsymbol{\xi}), \tag{4.5}$$

其中 $\mathbf{H}$ $=\boldsymbol{\Omega}^{-1/2}\mathbf{Q}$，这里 $\boldsymbol{\Omega}^{-1/2}$ 表示 $\boldsymbol{\Omega}$ 的唯一正定对称平方根矩阵的逆，而 $\mathbf{Q}$ $= [\boldsymbol{v}_1 \dots \boldsymbol{v}_p]$ 是一个正交矩阵，其第一列向量为 $\boldsymbol{v}_1 = \boldsymbol{\Omega}^{-1/2}\boldsymbol{\alpha}/\|\boldsymbol{\Omega}^{-1/2}\boldsymbol{\alpha}\|$，其余各列 $\boldsymbol{v}_2, \dots, \boldsymbol{v}_p$属于 $\boldsymbol{v}_1$ 的正交补。于是 $\boldsymbol{X}^* \sim$ SL$_p(\mathbf{0}, \mathbf{I}\ \boldsymbol{\alpha}^*_{\boldsymbol{X}})$，且 $\boldsymbol{\alpha}^*_{\boldsymbol{X}} = (\alpha^*, 0, \dots, 0)^\top$，以及 $\alpha^* = \|\boldsymbol{\Omega}^{-1/2}\boldsymbol{\alpha}\|$。

证明。根据 Arslan 命题 3，若 A $\in \mathbb{R}^{p\times p}$ 是任意满秩矩阵，则 A$(\boldsymbol{X} - \boldsymbol{\xi})$ $\sim$SL$_p(0, \mathrm{A}\Omega\mathrm{A}^\top, \mathrm{A}\boldsymbol{\alpha})$。由于 $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_p)^\top \neq \mathbf{0}$ 且 $\Omega^{-1/2}$ 非奇异，因此 $\Omega^{-1/2}\boldsymbol{\alpha} \neq \mathbf{0}$。因此，在 $\boldsymbol{u} = (u_1, \dots, u_p)^\top = \Omega^{-1/2}\boldsymbol{\alpha}$ 中，至少有一个分量 u_k 非零。此外，若 $\boldsymbol{e}_j$ 表示 $\mathbb{R}^p$中的第 j 个标准基向量（ $j \in \{1, \dots, p\}$），则 $\{\boldsymbol{u}\} \cup \{\boldsymbol{e}_j \in \mathbb{R}^p\colon\ j \neq k\}$ 构成 $\mathbb{R}^p$ 的一组基。为便于阅读，将向量重新记为 $\boldsymbol{u}_1 = \boldsymbol{u}, \boldsymbol{u}_{j+1} = \boldsymbol{e}_j$（ $1 \leq j < k$），以及$\boldsymbol{u}_j = \boldsymbol{e}_j$（ $k < j \leq p$）。接着，我们可以应用 Gram-Schmidt过程，求得 $\mathbb{R}^p$ 的一组标准正交基$\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_p\}$。为此，取 $\tilde{\boldsymbol{v}}_1 = \boldsymbol{u}_1 = \Omega^{-1/2}\boldsymbol{\alpha}$ 和 $\tilde{\boldsymbol{v}}_j = \boldsymbol{u}_j - \sum_{i=1}^{j-1} \frac{\tilde{\boldsymbol{v}}_i^\top \boldsymbol{u}_j}{\tilde{\boldsymbol{v}}_i^\top \tilde{\boldsymbol{v}}_i} \tilde{\boldsymbol{v}}_i$（ $2 \leq j \leq p$）。所需的基由归一化向量 $\boldsymbol{v}_j = \tilde{\boldsymbol{v}}_j/\|\tilde{\boldsymbol{v}}_j\|$（ $j \in \{1, \dots, p\}$）得到。因此，取 Q $= [\boldsymbol{v}_1 \dots \boldsymbol{v}_p]$ 且 H $= \Omega^{-1/2}$Q，可得 H$^\top(\boldsymbol{X} - \boldsymbol{\xi})$ $\sim$SL$_p(0, \mathrm{H}^\top\Omega\mathrm{H}, \mathrm{H}^\top\boldsymbol{\alpha})$，其中$\overline{\mathrm{H}^\top}\Omega\mathrm{H}$ $=\mathrm{Q}^\top\Omega^{-1/2}$ $\Omega\,\Omega^{-1/2}$Q $=$I，且 $\mathrm{H}^\top\boldsymbol{\alpha} = \mathrm{Q}^\top\Omega^{-1/2}\boldsymbol{\alpha} = (\|\Omega^{-1/2}\boldsymbol{\alpha}\|, 0, \dots, 0)^\top$。

□

与偏斜正态分布和偏斜 t 分布的情形类似，偏斜拉普拉斯分布存在典范形式，这一点有助于实施拟合优度检验，因为需要考虑的参数较少

。命题 4.1 的证明同时给出了一种求取多元 SL^p 分布典范形式的算法。考虑到这一点，令 $\widehat{\boldsymbol{X}}_j = \widehat{\mathbf{H}}_n^{\top}(\boldsymbol{X}_j - \widehat{\boldsymbol{\xi}}_n)$、 $j \in \{1,\ldots,n\}$为标准化观测值。那么，将 $\boldsymbol{X}_j$ 替换为 $\widehat{\boldsymbol{X}}_j$，即可应用(2.7) 中的检验，其中 $\boldsymbol{X}_{0,j}$ 取自参数为 $(\boldsymbol{\xi},\boldsymbol{\Omega}) = (\mathbf{0},\mathbf{I})$ 的 SL 分布，且令 $\boldsymbol{\alpha}$ 等于 $(\|\widehat{\boldsymbol{\Omega}}_n^{-1/2}\widehat{\boldsymbol{\alpha}}_n\|,0,\ldots,0)^{\top}$。为估计 $\widehat{\boldsymbol{\xi}}_n$、 $\widehat{\boldsymbol{\Omega}}_n$ 和$\widehat{\boldsymbol{\alpha}}_n$，我们建议采用 Arslan [2]提出的 EM 算法。

4.4 多元 Tukey g 与 h 分布

多元 Tukey g 与 h 分布（GH）最早由 Field 和 Genton[17] 提出，是对 Tukey 于 1977年提出的单变量对应分布的推广。由于其边缘分布具有灵活性，该分布日益受到关注，可用于拟合来自气候与环境问题的偏斜重尾数据集（例如，参见 Jeong 等人 [26], Yan 和 Genton [51],以及 Yan 等人 [52]）。给定两个参数向量 $\boldsymbol{g} = (g_1、\ldots\ldots、 g_p)^{\top} \in \mathbb{R}^p$ 和 $\boldsymbol{h} = (h_1、\ldots\ldots、 h_p)^{\top} \in \mathbb{R}_+^p$，若随机向量 $\boldsymbol{Y} \in \mathbb{R}^p$ 可表示为

$$\boldsymbol{Y} = \boldsymbol{\tau}_{\boldsymbol{g},\boldsymbol{h}}(\boldsymbol{Z}) := (\tau_{g_1,h_1}(Z_1),\ldots,\tau_{g_p,h_p}(Z_p))^{\top}, \tag{4.6}$$

其中， $\boldsymbol{Z} = (Z_1、\ldots\ldots、 Z_p)^{\top} \sim N_p(\mathbf{0}, \mathbf{I})$ 服从标准多元正态分布，并且对于给定的 $g \in \mathbb{R}$ 和 $h \in \mathbb{R}_+$，单变量函数 $\tau_{g,\,h}$ 定义为

$$\tau_{g,h}(z) = \begin{cases} \left(\dfrac{\exp(gz)-1}{g}\right)\exp\left(\dfrac{hz^2}{2}\right), & g \in \mathbb{R}\setminus\{0\}, \\ z\exp\left(\dfrac{hz^2}{2}\right), & g = 0, \end{cases} \tag{4.7}$$

对于任意 $z \in \mathbb{R}$。一般的多元 Tukey g-and-h 分布定义如下:

$$\boldsymbol{Y} = \boldsymbol{\Omega}\,\boldsymbol{\tau}_{\boldsymbol{g},\boldsymbol{h}}(\boldsymbol{Z}) + \boldsymbol{\xi}, \tag{4.8}$$

其中$\boldsymbol{\Omega} \in \mathbb{R}^{p\times p}$ 和 $\boldsymbol{\xi} \in \mathbb{R}^p$ 分别是正定矩阵和位置向量。这里$\boldsymbol{\vartheta} = (\boldsymbol{\xi},\boldsymbol{\Omega}\ \boldsymbol{g},\boldsymbol{h})$，干扰参数 $\boldsymbol{\xi}$和$\boldsymbol{\Omega}$ 可以被标准化消去。我们将采用记号 GH^p($\boldsymbol{\xi}$,$\boldsymbol{\Omega}$ $\boldsymbol{g},\boldsymbol{h}$)。

与偏斜正态分布、偏斜 t 分布和 SL分布不同，Tukey g-and-h 分布没有已知的变换可使其表示为可约规范形式。此外，由于τ_g 的逆函数h(·)没有闭式表达式，经典估计方法（例如最大似然估计法）依赖数值近似。同样，特征函数、概率密度函数和累积分布函数也只能通过数值方法计算。已有多种方法被提出，用于估计和拟合单变量 Tukey g-and-h 分布（参见 M¨ostel 等人 [39]的综述论文）。然而，对于多元情形，可用的方法寥寥无几。例如，Field 和 Genton [17] 使用多元分位数进行数据拟合。He 和 Raghunathan

[22] 假设$\boldsymbol{\Omega}$ 为对角矩阵，因此提出了一种利用单变量 Tukey g-and-h分位数的算法。

在这项研究中，我们选择了最大似然估计法。为此，我们需要找到使对数似然函数达到最大值的参数 $\boldsymbol{\vartheta}$

$$\ell(\boldsymbol{\vartheta}|\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_n) = -n\ln|\boldsymbol{\Omega}| + \sum_{i=1}^{n}\sum_{j=1}^{p}\left[\ln\left\{\phi(u_i^{(j)})\right\} - \ln\left\{\tau'_{g_j,h_j}(u_i^{(j)})\right\}\right], \tag{4.9}$$

其中 $u_i^{(j)} = \tau^{-1}_{g_j,h_j}\left(\{\boldsymbol{\Omega}^{-1}(\boldsymbol{Y}_i - \boldsymbol{\xi})\}^\top \boldsymbol{e}_j\right)$，对于 $i \in \{1,\ldots,n\}$ 和 $j \in \{1,\ldots,p\}$，$\boldsymbol{e}_j$表示 $\mathbb{R}^p$ 中的第 j 个标准基向量，$\phi(\cdot)$ 是标准单变量正态分布的概率密度函数，而 $\tau'_{g_j,h_j}(z)$ 表示 $\tau_{g_j,h_j}(z)$ 关于其自变量 z 的导数。为了得到量 $u_i^{(j)}$ 的近似值，我们使用 R 中 stats [45] 包提供的 uniroot 函数。对于最大化过程，我们使用 R 中 MASS [46] 包提供的 optim 函数，并采用 “Nelder-Mead” 方法。

4.5 多变量 α-稳定分布

α-稳定分布（AS）与多变量偏斜 t 分布以及 Tukey g–h 分布类似，是多变量高斯分布的另一种可行扩展，能够表现偏度。然而，多变量偏斜 t 分布族具有有限的二阶矩，而 α-稳定分布由参数 $\alpha \in (0,2]$, 称为尾指数）控制；仅当 $\alpha = 2$时，它才具有有限的二阶矩，此时退化为多变量高斯情形。特别地，如果 $\boldsymbol{X}$ 服从多变量α-稳定分布，则当 $0 < s < \alpha$时，$\mathbb{E}(\|\boldsymbol{X}\|^s) < \infty$；当 $s \geq \alpha$时，$\mathbb{E}(\|\boldsymbol{X}\|^s) = \infty$。文献中已经提出了若干种特征函数的参数化形式。出于数值方面的考虑，本文采用 Abdul-Hamid 和 Nolan [1], 提出的 S^0 参数化形式，其特征函数为

$$\varphi_{\boldsymbol{\vartheta}}(\boldsymbol{t}) = \exp\left(-\int_{\mathcal{S}_p}\psi_\alpha(\boldsymbol{t}^\top\boldsymbol{s})\,\Gamma(d\boldsymbol{s}) + \mathrm{i}\boldsymbol{t}^\top\boldsymbol{\xi}\right), \quad \boldsymbol{t}\in\mathbb{R}^p, \tag{4.10}$$

其中

$$\psi_\alpha(u) = \begin{cases} |u|^\alpha\left(1 + \mathrm{i}\,\mathrm{sign}(u)\tan\left(\frac{\pi\alpha}{2}\right)\left(|u|^{1-\alpha}-1\right)\right), & \alpha \neq 1, \\ |u|\left(1 + \mathrm{i}\frac{2}{\pi}\mathrm{sign}(u)\ln|u|\right), & \alpha = 1. \end{cases} \tag{4.11}$$

在这种情况下，我们用 $\boldsymbol{X} \sim S^0_\alpha(\Gamma,\boldsymbol{\xi})$ 表示随机向量 $\boldsymbol{X}$ 具有 α-稳定分布，其谱测度 $\Gamma(\cdot)$ 有限，定义在单位球面 $\mathcal{S}_p := \{\boldsymbol{s}\in\mathbb{R}^p : \|\boldsymbol{s}\| = 1\}$ 上，位移向量为 $\boldsymbol{\xi}\in\mathbb{R}^p$。由式 (4.10) 可知，多元 α-稳定分布属于半参数族，由三元组 $(\alpha,\Gamma,\boldsymbol{\xi})$ 完全确定，并且属于更一般的无限可分分布类。此外，$\boldsymbol{X} - \boldsymbol{\xi} \sim S^0_\alpha(\Gamma, 0)$，因此 $\boldsymbol{\xi}$ 是干扰参数向量；当 $1 < \alpha \leq 2$ 时，它与均值向量一致。

Byczkowski 等人 [11] 证明，可以对谱测度 Γ给出一种近似，这对于数值计算和模拟很有用。该近似描述如下：

考虑 $\mathcal{S}_p$ 的一个有限划分 $P=\{A_1,\ldots,A_m\}$ 以及一个点集 $S=\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_m\}\subseteq\mathcal{S}_p$。然后，通过设定 $\gamma_i=\Gamma(A_i)$，与（P、S、Γ）相关联的离散谱测度定义为

$$\Gamma^*(\cdot)=\sum_{i=1}^{m}\Gamma(A_i)\,\mathbb{1}_{\boldsymbol{s}_i}(\cdot)=\sum_{i=1}^{m}\gamma_i\,\mathbb{1}_{\boldsymbol{s}_i}(\cdot), \tag{4.12}$$

对于 中的任意博雷尔集 $\mathcal{S}_p$。这里 Γ^* 是通过将 $\Gamma(A_i)$ 的质量集中到每个点$\boldsymbol{s}_i$而隐式定义的，唯一的要求是，对于每个 $_{\boldsymbol{s}\in A_i}|\boldsymbol{s}-\boldsymbol{s}_i|$，sup 足够小 $i=1$，……，m。因此，实际应用中可以使用离散谱测度 Γ^* 来代替其连续对应物 Γ。为了模拟多元 αα-稳定随机向量，我们将采用 Modarres 和 Nolan 的结果[38]，该结果指出，若 $\boldsymbol{X}\sim \mathrm{S}_\alpha^0(\Gamma^*,\ \boldsymbol{\xi})$，且 $\Gamma^*(\cdot)$ 在 (4.12) 中定义，则

$$\boldsymbol{X}=\begin{cases}\sum_{i=1}^{m}\gamma_i^{1/\alpha}Z_i\boldsymbol{s}_i+\tilde{\boldsymbol{\xi}}, & \alpha\neq 1,\\ \sum_{i=1}^{m}\gamma_i(Z_i+\frac{2}{\pi}\ln(\gamma_i))\boldsymbol{s}_i+\tilde{\boldsymbol{\xi}}, & \alpha=1,\end{cases} \tag{4.13}$$

其中 $\tilde{\boldsymbol{\xi}}=\boldsymbol{\xi}$–tan $\left(\frac{\pi\alpha}{2}\right)\sum_{i=1}^{m}\gamma_i\boldsymbol{s}_i$ 和 Z_1，……，Z_m 是独立同分布的一维 α-稳定随机变量，其 $Z_i\sim \mathrm{S}\alpha(1,\ 1,\ 0)$（即尺度 $=$、偏度 $=1$ 和位置参数 $=0$）。至于估计部分，我们采用 Nolan 等人 [41], 提出的投影法，该方法依赖于将多元样本投影到从单位球面中特别选取的一组网格值上。

众所周知，$\boldsymbol{X}\sim \mathrm{S}_\alpha^0(\Gamma,\ \boldsymbol{\xi})$ 对称，当且仅当 $\Gamma(\cdot)$ 在 $\mathcal{S}_p$ 上对称。迄今为止，文献中尚无专门针对一般非对称多元 α-稳定分布设计的检验，Meintanis 等人（2015）提出的检验除外，但该检验仅适用于对称情形。本文提出一种适用于对称和非对称情形的检验。尽管文献中并不常用这种记号，我们使用 $\boldsymbol{X}\sim \mathrm{AS}p(\boldsymbol{\xi},\ \Gamma,\ \alpha)$ 表示 $\boldsymbol{X}$ 服从非对称 α-稳定分布。若想进一步了解这些内容以及多元 α-稳定分布相关的技术细节，建议阅读 Karling 等人 [29] 以及 Samorodnitsky 和 Taqqu [47]。

5 模拟研究

本节介绍使用第2 节和第3 节所述检验，并采用第4节介绍的五个偏斜分布族进行模拟研究所得的结果。在这些模拟中，我们将 $M=L=1000$ 用作第3.1节和第3.2节所述步骤中的标准值。首先，我们计算在预设的 $\delta=0.05$ 名义水平下，各分布族对应检验的经验估计检验尺寸。随后，在第二轮和第三轮模拟中，我们分别考察简单假设情形和复合假设情形下检验的功效。在后一种情形中，我们以双曲正弦-反双曲正弦分布族为参照，对五个分布族进行检验（见 Jones 和 Pewsey [28]）。值得指出的是，随着样本量 n 的增加，检验运行所需的计算时间自然也会增加。此外，检验的效率容易受到各分布族所含参数数量以及估计方法的影响

。最后，本节将我们的检验与文献中已有的几种偏斜正态族检验方法进行比较。

5.1 估计检验尺寸

表 1 中给出的数值由第3 节所述的检验生成，对应于复合原假设情形，其中 $\delta = 0.05$ 和 $M = 1000$ 为重复试验次数。检验中样本所考虑的维数为 $p = 2$。关于 $p = 3$ 的类似表格可见于补充材料。为简化模拟，在每种情况下，我们将 m 固定为等于 n（样本量），其中 $n \in \{100, 250, 500, 750, 1000\}$。在后文中，为了更好地理解人工样本尺寸 m的影响，我们给出固定样本量 n 并令 m 变化的模拟结果；参见第12个模拟，该模拟位于第5.4小节。

模拟1。我们从一个 SN n 分布中模拟了 $p(\mathbf{0}, \mathbf{I}\ \boldsymbol{\alpha})$ 的观测值，其中 $\boldsymbol{\alpha} = (3, 0, \ldots\ldots, 0)^\top$。我们计算了检验的经验尺寸，结果见表 1。从该表可以看出，随着样本量增加，检验的经验尺寸稳定在 。05 附近，这与设定的名义水平相对应。

模拟2。接下来，我们模拟了 n 个来自偏斜 t 分布的观测值$p(\mathbf{0}, \mathbf{I}\ \boldsymbol{\alpha}, \nu)$ 分布，其中 $\boldsymbol{\alpha} = (3, 0, \ldots, 0)^\top$ 且 $\nu = 5$。检验的估计检验尺寸列于表1。将此情形与模拟1中的情形进行比较，我们注意到，当 n (100,250,500)取较小值时，估计检验尺寸不像偏斜正态分布情形中观察到的那样接近0。05，但随着 n的值增大，它们开始趋近于设定的名义水平，表明具有一致性。

模拟3。这里，我们模拟了 n 个来自 SL 分布的观测值$p(\mathbf{0}, \mathbf{I}\ \boldsymbol{\alpha})$ 分布，其中 $\boldsymbol{\alpha} = (3, 0, \ldots, 0)^\top$。检验的估计检验尺寸列于表1。这里可以观察到与前两次模拟相似的结果；随着样本量增加，检验的估计检验尺寸趋近于设定的名义水平。

模拟 4。接下来，我们模拟了 n 个来自 GH 分布的观测值$p(\mathbf{0}, \mathbf{I}\ \boldsymbol{g}, \boldsymbol{h})$ 分布，其中 $\boldsymbol{g} = (1, \ldots, 1)^\top$ 和 $\boldsymbol{h} = (0.5, \ldots, 0.5)^\top$。检验的估计检验尺寸见表 1。由于 $\boldsymbol{h}$ 的所有边缘分量均等于 0。5，因此此时方差不存在，且与偏斜正态分布、偏斜 t 分布和 SL 情形相比，尾部更重。这一点尤其反映在检验的估计检验尺寸上。我们观察到，对于 $n = 1000$，拒绝率为 0.071，这相对较高。这可能是由于观测数据集在尾部的分散范围较广。下一次模拟中展示了另一个与此类似的情形。

模拟 5。最后，我们从一个 $\mathrm{AS}_2(0, \Gamma, \alpha)$ 分布中模拟了 n 个观测值，该分布具有离散谱测度 $\Gamma(\cdot) = (1/3) \sum_{k=1}^{3} \mathbb{1}_{\boldsymbol{s}_k}(\cdot)$，其中 $\boldsymbol{s}_k = (\cos(2\pi k/3), \sin(2\pi k/3))^\top$，$k \in \{1, 2, 3\}$，且稳定指数为 $\alpha = 1.5$。在估计程序中，我们使用了一个包含 $N = 24$ 个投影的网格（参见 Nolan 等人 [41]）。检验的估计检验尺寸见表1。这里，与模拟 4 类似，

观测值来自一个具有无限方差和重尾的分布。然而，在这种情况下，由于我们发现，当将样本量 n 提高到 1000 时，检验的估计尺寸也随之增加，因此我们又进行了两轮模拟，分别使用 $n = 2500$ 和 $n = 5000$，以确保其没有偏离预设的名义水平。这些情况下得到的拒绝率分别为 0.066 和 0.070。

表1：模拟1 - 5中检验对应的估计检验尺寸。

	$n = 100$	$n = 250$	$n = 500$	$n = 750$	$n = 1000$
SN	0.045	0.047	0.049	0.045	0.053
ST	0.063	0.054	0.062	0.052	0.057
SL	0.063	0.045	0.048	0.055	0.046
GH	0.063	0.046	0.057	0.063	0.071
AS	0.051	0.064	0.063	0.076	0.080

5.2 简单原假设情形的估计功效函数

考虑简单原假设下的拟合优度问题$\mathcal{H}_0^s$，如式（3.1）所示；备择假设为 $\mathcal{H}_1^s$，如式（3.2）所示。在接下来的五组模拟（6-9）中，我们针对 GH、SL、偏斜正态分布、偏斜 t 分布和 α-稳定分布的若干情形，计算了检验的经验功效函数。前四组模拟中，我们考虑了样本量 $n \in \{100, 250, 500, 750, 1000\}$ 和 $p \in \{2, 3\}$，并考察生成观测值的维数。结果汇总并展示于图1中。

模拟6。我们从 $\mathrm{GH}_p(\mathbf{0}, \mathrm{I}\ \boldsymbol{g}, \boldsymbol{h})$ 分布生成了 n 个观测值。对于原假设，我们取 $\boldsymbol{\lambda}_0 = (\boldsymbol{g}_0, \boldsymbol{h}_0)$，其中 $\boldsymbol{g}_0 = (2, \ldots, 2)^\top$ 和 $\boldsymbol{h}_0 = (1, \ldots, 1)^\top$。我们针对 $\boldsymbol{\lambda} = (\boldsymbol{g}, \boldsymbol{h})$ 计算了功效函数，其中 $\boldsymbol{h} = (h, \ldots, h)^\top$，且 $h \in \{0.2, 0.4, 0.6, 0.8\}$ 和 $\boldsymbol{g} = 2\boldsymbol{h}$，因此它们仅取决于 h 的选择。对于 $p = 2$ 和 $p = 3$，结果分别绘于图1的 (a) 和 (b) 项中。快速浏览绘制的函数可见，样本量增加时，检验功效也随之提高。此外，当 h 趋近于 1 时，功效函数如理论预期般收敛于检验尺寸。

模拟7。接下来，我们从 $\mathrm{SL}_p(\mathbf{0}, \mathrm{I}\ \boldsymbol{\alpha})$ 分布中生成了 n 个观测值。对于原假设，我们取 $\boldsymbol{\lambda}_0 = \boldsymbol{\alpha}_0 = (\alpha_0^*, 0, \ldots, 0)^\top$，其中 $\alpha_0^* = 3$。随后，我们计算了 $\boldsymbol{\lambda} = \boldsymbol{\alpha} = (\alpha^*, 0, \ldots, 0)^\top$ 在 $\alpha^* \in \{0, 1, 2, 5, 8, 13\}$ 时的功效函数，结果分别针对 $p = 2$ 和 $p = 3$绘于图1的 (c) 和 (d) 项中。从图中的陡峭程度可以看出，在两种维数情况下，功效函数对 α^* 的微小变化都非常敏感，其参数增大或减小时，检验功效都会提高。此外，可以观察到图形在 $\alpha^* = 3$ 附近存在轻微的不对称性，检验的大小恰好位于此处。

模拟 8。这里，我们从 $\mathrm{SN}_p(0, \mathrm{I}\ \boldsymbol{\alpha})$ 分布中生成了 n 个观测值。对于原假设，我们考虑 $\boldsymbol{\lambda}_0 = \boldsymbol{\alpha}_0 = (\alpha_0^*, 0, \ldots, 0)^\top$，其中 $\alpha_0^* = 3$；并在 $\alpha^* \in \{0, 1, 2, 5, 8, 13\}$ 时考虑 $\boldsymbol{\lambda} = \boldsymbol{\alpha} = (\alpha^*, 0, \ldots, 0)^\top$，以计算经验功效函数。所得函数分别针对 $p = 2$ 和 $p = 3$绘于图1的 (e) 和 (f) 项中。从图中可以看出，检验的经验估计大小充当了拐点；当 $\alpha^* < 3$ 时，功效函数在较高的一侧呈不对称分布。这种现象可能是由分布的不对称性造成的。α^* 越接近 0，分布就越接近多元正态分布。当 α^* 增大时，检验功效不高，表明该方向上的曲线较为平缓。

模拟 9。最后，我们生成了 n 个来自 $\mathrm{ST}_p(\mathbf{0}, \mathbf{I}\ \boldsymbol{\alpha}, \nu)$ 分布的观测值，其中 $\boldsymbol{\alpha} = (3, 0, \ldots\ldots, 0)^\top$ 固定。我们考虑了 $\nu_0 = 5$ 作为原假设，并将 $\nu \in \{3, 8, 13, 21, 34, \infty\}$ 用于计算经验功效函数，这些函数绘于图 1中的项目（g）和（h），分别对应$p = 2$ 和 $p = 3$。通过观察这些图可以注意到，随着 ν 增加，对于较大的样本量，功效函数会迅速增加至 1，而检验的大小在 $\nu = 5$处达到。这里，$\nu = \infty$ 被解释为当 $\nu \to \infty$ 时的渐近分布。

对于下一次模拟，考虑 $n \in \{250, 500, 750, 1000, 2500\}$ 和 $p = 2$。

模拟 10。我们从 $\mathrm{AS}_2(0, \Gamma, \alpha)$ 分布中生成了 n 个观测值。对于原假设，我们取谱测度 $\Gamma_0(\cdot) = (1/3)\sum_{k=1}^3 \mathbb{1}_{\boldsymbol{s}_k}(\cdot)$，其中 $\boldsymbol{s}_k = (\cos(2\pi k/3), \sin(2\pi k/3))^\top$，$k \in \{1, 2, 3\}$，以及 $\alpha_0 = 1.5$。随后，我们为备择假设取两组不同的集合。在第一组中，我们固定稳定指数 $\alpha = \alpha_0$，并计算备择谱测度 $\Gamma_q(\cdot) = (1/q)\sum_{k=1}^q \mathbb{1}_{\boldsymbol{s}_k}(\cdot)$ 的功效函数，其中 $\boldsymbol{s}_k = (\cos(2\pi k/q), \sin(2\pi k/q))^\top$，$k \in \{1, \ldots, q\}$ 和 $q \in \{34, 21, 13, 8, 5\}$。所得的估计功效函数绘于图 2 (a) 中。在第二组中，我们将谱测度固定为 $\Gamma(\cdot) = (1/3)\sum_{k=1}^3 \mathbb{1}_{\boldsymbol{s}_k}(\cdot)$，转而改变参数 α，$\alpha \in \{1.1, 1.3, 1.7, 1.9\}$。该情形下所得的估计功效函数绘于图 2 (b) 中。在两幅图中，我们还绘制了 $\alpha = \alpha_0$ 以及 $\Gamma(\cdot) \equiv \Gamma_3(\cdot)$ 时检验的估计大小。

5.3 复合原假设情形下的估计功效函数 e

对于接下来的模拟，考虑具有复合原假设 $\mathcal{H}_0^c$ （如3.3）和备择假设 $\mathcal{H}_1^c$ （如3.4）的拟合优度检验问题。

模拟 11。我们从一个 $\mathrm{ST}_p(0, \mathrm{I}\ \boldsymbol{\alpha}, \nu)$ 分布中生成了 n 个观测值，其中 $p \in \{2, 3\}$，$\boldsymbol{\alpha} = (3, 0, \ldots, 0)^\top$ 为固定值，且 $\nu \in \{1, 2, 3, 5, 8, 13, \infty\}$。对于原假设，我们考虑偏斜正态分布族。因此，生成的观测值属于备择假设集合。随后，我们计算了经验功效函数，结果绘于图 3。我们注意到，随着 $\nu \to \infty$，$\mathrm{ST}_p(0, \mathrm{I}\ \boldsymbol{\alpha}, \nu)$ 分布收敛于 $\mathrm{SN}_p(0, \mathrm{I}\ \boldsymbol{\alpha})$ 分布。特别是，这种效应在功效函数中也有所体现，其收敛至显著性水平，此处设为 5%。

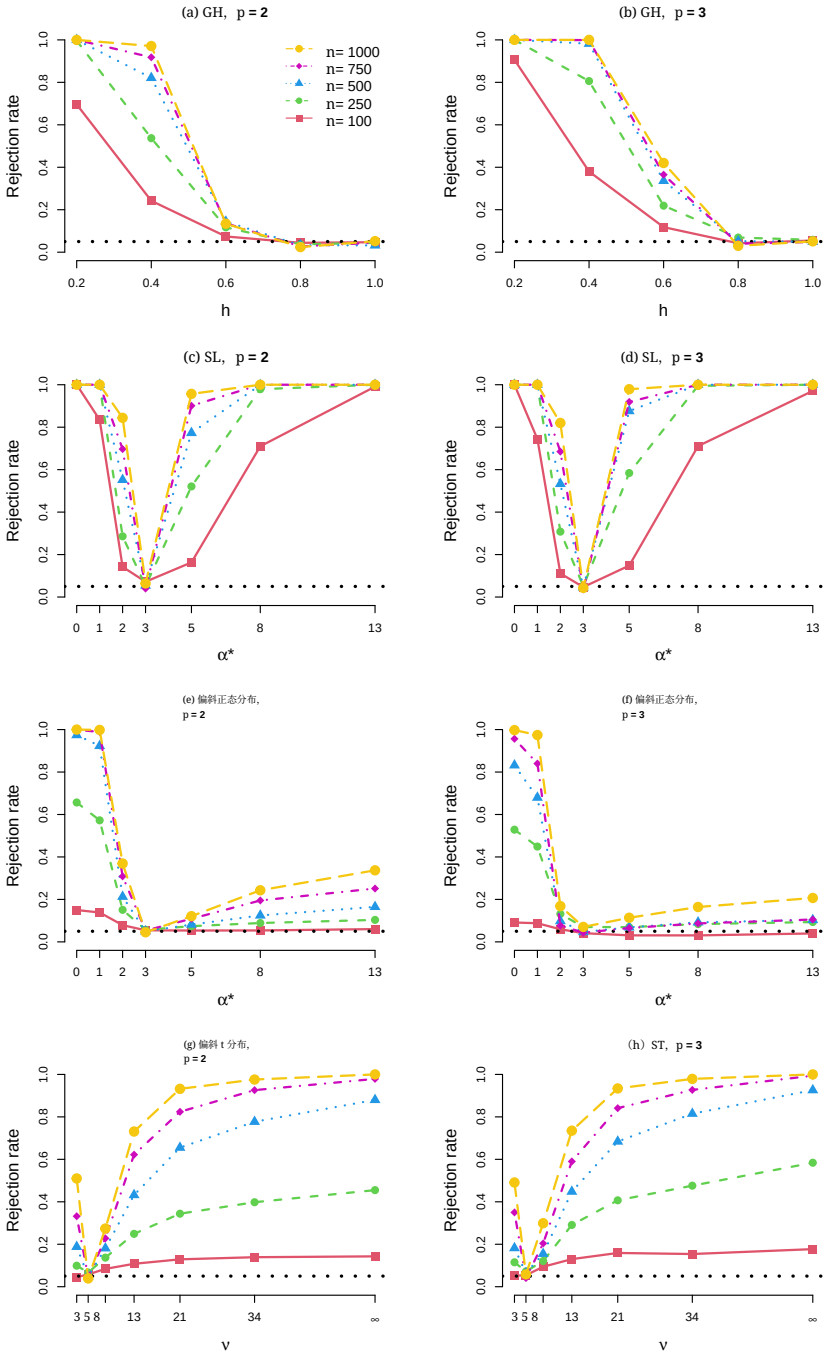


图1：模拟6-9中生成的功效函数：(a)、(b) 模拟 6 $\to \mathrm{GH}_p(0, \mathrm{I}\ \boldsymbol{g}, \boldsymbol{h})$；(c)、(d) 模拟 7 $\to \mathrm{SL}_p(0, \mathrm{I}\ \boldsymbol{\alpha})$；(e)、(f) 模拟 8 $\to \mathrm{SN}_p(0, \mathrm{I}\ \boldsymbol{\alpha})$；(g)、(h) 模拟 9 $\to \mathrm{ST}_p(0, \mathrm{I}\ \boldsymbol{\alpha}, \nu)$。水平点线对应于5%显著性水平。

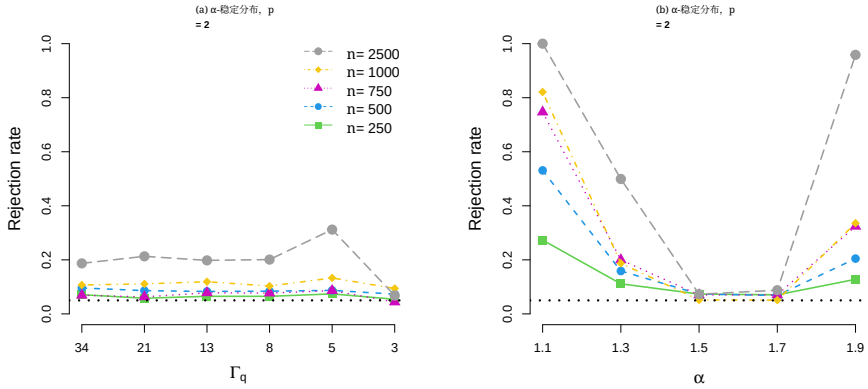


图 2：当原假设为 $\mathrm{AS}_2(\mathbf{0}, \Gamma_0, \alpha_0)$ 分布时检验的功效函数，其中$\alpha_0 = 1.5$和 $\Gamma_0(\cdot) = (1/3)\sum_{k=1}^{3} \mathbb{1}_{\boldsymbol{s}_k}(\cdot)$，$\boldsymbol{s}_k = (\cos(2\pi k/3), \sin(2\pi k/3))^\top$，对应于 $k \in \{1,2,3\}$，即模拟 10。左侧图中，横坐标上的索引对应于备择谱测度 $\Gamma_q(\cdot) = (1/q)\sum_{k=1}^{q} \mathbb{1}_{\boldsymbol{s}_k}(\cdot)$，其中 $\boldsymbol{s}_k = (\cos(2\pi k/q), \sin(2\pi k/q))^\top$，对应于 $k \in \{1,\ldots,q\}$和 $q \in \{34,21,13,8,5\}$。右侧图中，我们绘制了检验的功效，此时固定 $\Gamma(\cdot)$，而备择假设对应于 $\alpha \in \{1.1, 1.3, 1.7, 1.9\}$。

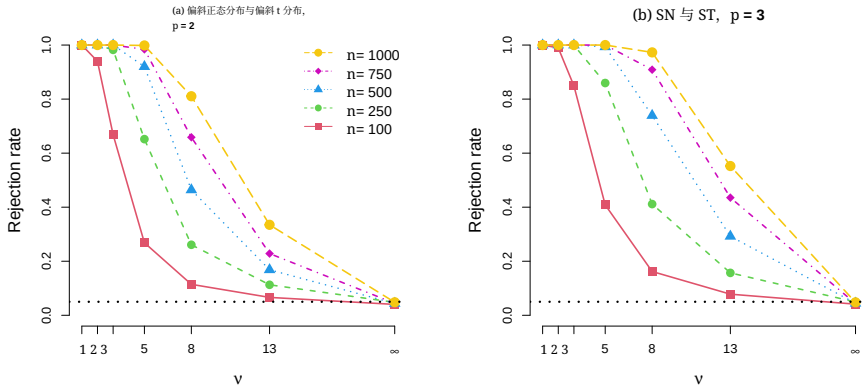


图 3：当原假设为 $\mathrm{SN}_p(\mathbf{0}, \mathbf{I}\ \boldsymbol{\alpha})$ 分布（某个 $\boldsymbol{\alpha} = (\alpha^*, 0, \ldots\ldots, 0)^\top$ 且 $\alpha^* \in [0,\infty)$）时，针对 $\mathrm{ST}_p(\mathbf{0}, \mathbf{I}\ \boldsymbol{\alpha}, \nu)$ 备择分布族的检验功效，其中 $p \in \{2,3\}$、$\boldsymbol{\alpha} = (3, 0, \ldots\ldots, 0)^\top$ 固定，且 $\nu \in \{1,2,3,5,8,13,\infty\}$。水平虚线对应 5%显著性水平。

模拟12。现在考虑 Jones和Pewsey [28]提出的多元双曲正弦-反双曲正弦分布族。本次模拟旨在将双曲正弦-反双曲正弦分布视为检验的备择假设，并将第4节介绍的五类分布视为原假设，从而计算功效函数。为定义多元双曲正弦-反双曲正弦

分布，让我们考虑单变量变换

$$\mathrm{S}_{a,b}(z) = \sinh\{b\sinh^{-1}(z) - a\}, \quad a \in \mathbb{R},\ b \in \mathbb{R}_+. \tag{5.1}$$

此外，设（$\boldsymbol{e}$，$\boldsymbol{f}$）$\in \mathbb{R}^p \times \mathbb{R}^p_+$，其中 $\boldsymbol{e} =$（e_1，……，e_p）$^\top$ 和 $\boldsymbol{f} =$（f_1，……，f_p）$^\top$，且 $\boldsymbol{Z} \sim$ N$_p$（$\mathbf{0}$，$\mathbf{I}$）服从标准 p元高斯分布。然后，对各分量应用（5.1），我们称

$$\boldsymbol{Y}_{\boldsymbol{e},\boldsymbol{f}} = \mathbf{S}_{-\boldsymbol{e}/\boldsymbol{f},1/\boldsymbol{f}}(\boldsymbol{Z}) := \left(\mathrm{S}_{-\frac{e_1}{f_1},\frac{1}{f_1}}(Z_1), \ldots, \mathrm{S}_{-\frac{e_p}{f_p},\frac{1}{f_p}}(Z_p)\right)^\top \tag{5.2}$$

服从参数为（$\boldsymbol{e}$，$\boldsymbol{f}$）的双曲正弦-反双曲正弦分布。当参数 $\boldsymbol{e} =$（e，……，e）$^\top$、$e \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$，以及 $\boldsymbol{f} =$（f，……，f）$^\top$、$f =$（$e+1$）$^{-1}$，从而满足 $1/f \in \{1, 1.1, 1.2, 1.3, 1.4, 1.5\}$ 和 $-e/f \in \{0, -1/11, -1/6, -3/13, -2/7, -1/3\}$时，我们从 $\boldsymbol{Y}_{\boldsymbol{e},\boldsymbol{f}}$ 中模拟生成了 n 个随机样本。特别地，当 $e = 0$、$\boldsymbol{Y}_{\boldsymbol{e},\boldsymbol{f}}$ 时，其服从标准高斯分布。此外，随着 e 增大，其分布会迅速偏离标准高斯分布，并且相对于原点，在每个轴上都呈现正偏。为生成这些随机样本，我们考虑了两种设置。在第一种设置中，我们在检验中取 $n \in \{50, 100, 250, 500\}$ 和 $m = n$ 。而在第二种设置中，我们固定生成样本的大小，令 $n = 100$，并沿集合$\{100, 250, 500, 1000\}$移动 m ，即测试所需的新生成数据的样本量，具体讨论见第2节。随后，对于每个检验，我们以第4节中考虑的五个不同分布族作为设计的原假设，计算了经验功效函数。两种不同设置下各检验的结果分别绘于图 4 和图 5 中。后者表明，如果数据的样本量较小（在本次特定模拟中为 $n = 100$），则将 m 的值从100逐步增加到1000，也会使检验功效略有提高；这对于小样本量的情形而言是一个有用的结论；例如，参见第 6.1 小节中的 AIS 数据集。

5.4 与竞争检验的比较

尽管文献中提出的主要检验仅适用于偏正态族，但本小节将讨论并比较我们提出的检验与这些替代方法。回顾相关研究，Meintanis 和 Hl´avka [35]提出了首批适用于多元偏态正态分布族的拟合优度检验之一，该检验利用经验矩母函数。然而，该检验的计算公式实际上仅适用于二维情形，因为在情形 $p > 2$ 下，需要求解一个难以进行数值计算的微分方程。随后，Balakrishnan 等人 [9] 提出了一种基于偏斜正态分布典范形式的检验。其主要优点在于无需进行重采样步骤，从而节省计算时间。然而，该检验仅适用于偏态正态分布，因为它基于这样一个事实：若 $\boldsymbol{X}_i$'s 服从典范偏态正态分布，则观测值 $\boldsymbol{X}_i = (X_{i,1}, \ldots, X_{i,p})^\top$ 的比值 $X_{i,j}/|X_{i,1}|$（$j \in \{2, \ldots, p\}$）服从柯西随机变量。最近，Gonz´alez-Estrada 等人 [21] 提出了两种随机化检验，与 Balakr- ishnan 等人 [9]的检验类似，它们均基于 SN 分布的估计典范形式。第一种检验（W）在对样本进行变换后，应用 Shapiro-Wilk 检验推广形式的原理

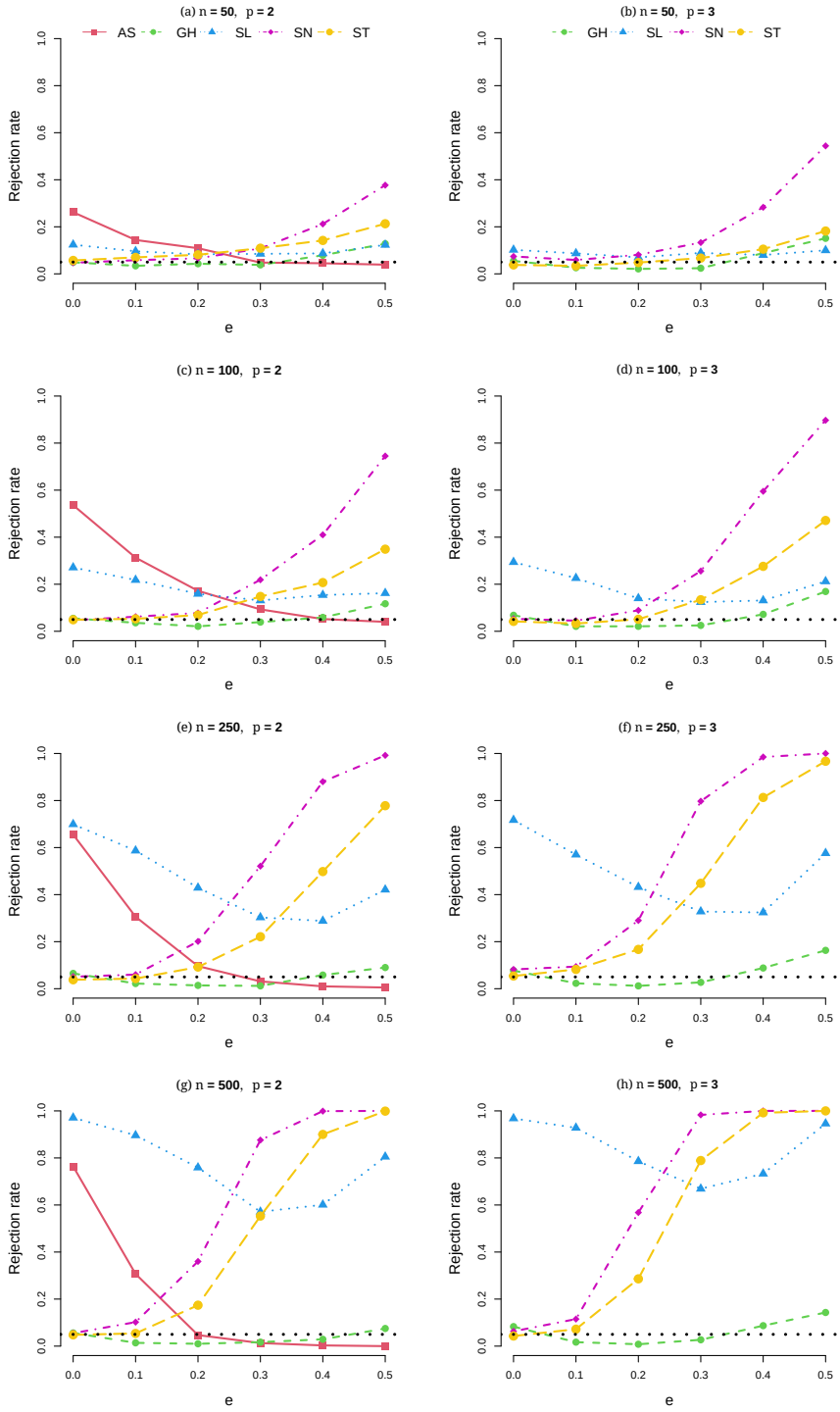


图 4：检验对双曲正弦-反双曲正弦分布备择分布族的检验功效，对应模拟 12。此处考虑的样本量为 $n \in \{50, 100, 250, 500\}$ 和 $m = n$。

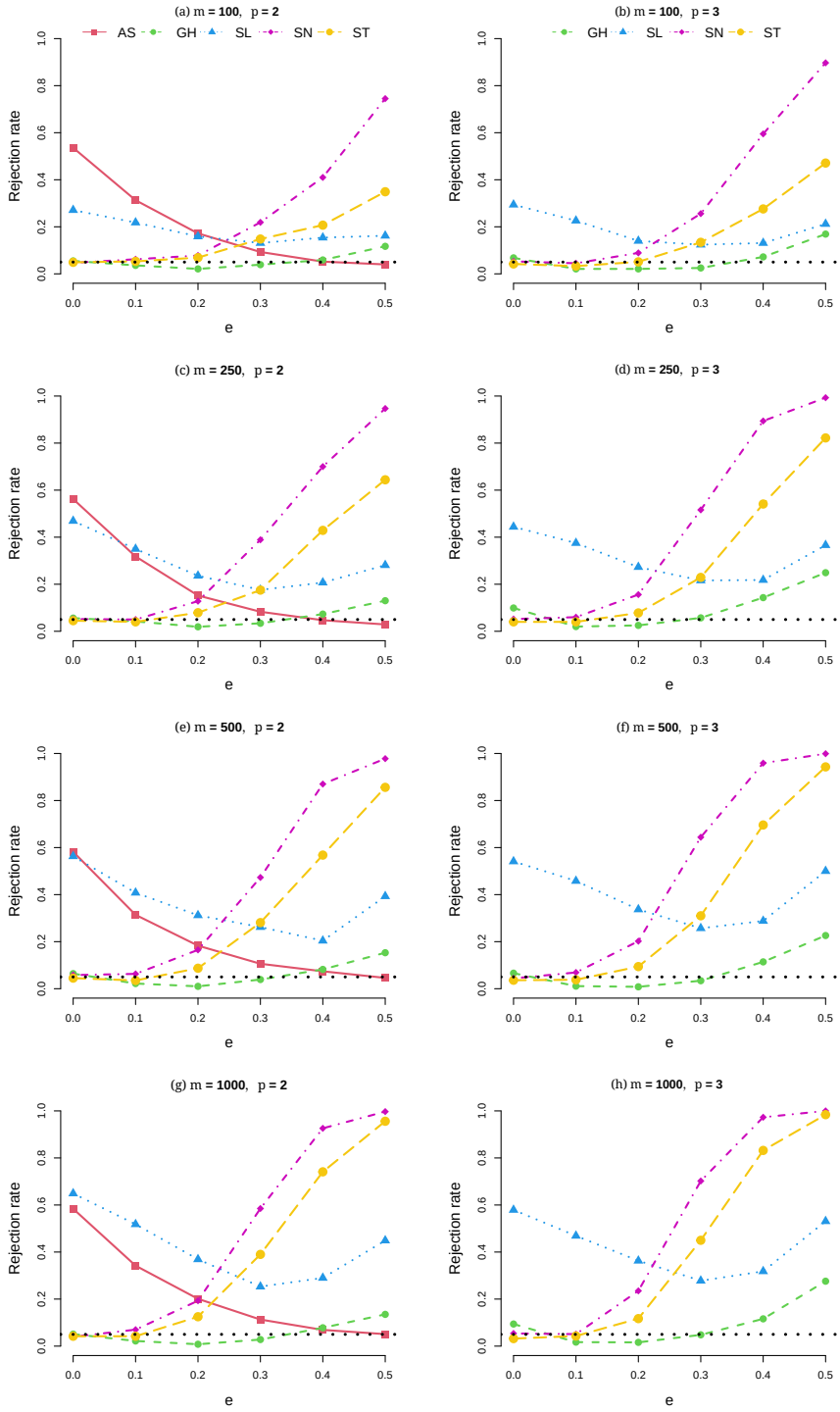


图5：针对与模拟对应的双曲正弦-反双曲正弦分布备择分布族的检验功效 12 ，样本量为 $n = 100$ 且固定 $m \in \{100, 250, 500, 1000\}$。

转化为近似的多元标准正态观测值，而第二个检验（S）则依赖于单变量独立偏态正态和正态随机变量之和的闭包性质。Jim´enez-Gamero 和 Kim [27] 提出了一对重抽样方案，其中之一是参数自助法（论文中称为 PB），我们在自己的检验中也采用了这种方法。

在下一项模拟中，我们重复 Balakrishnan 等人 [9] 的一项原始模拟研究，并将我们检验的估计检验功效与这些作者报告的结果以及 Gonz´alez-Estrada 等人 [21]给出的结果进行比较。

模拟 13。 这项研究针对复合原假设采用三元偏态正态分布，并在备择集合中采用 $ST_3(\boldsymbol{\xi},\ \boldsymbol{\Omega}\ \boldsymbol{\alpha},\ \nu)$ 分布，其中 $\boldsymbol{\xi}$、$\boldsymbol{\Omega}$ 和 $\boldsymbol{\alpha}$定义如下

$$\boldsymbol{\xi} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad \boldsymbol{\Omega} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2.5 & 1 \\ 1 & 1 & 5 \end{pmatrix}, \qquad \boldsymbol{\alpha} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}. \tag{5.3}$$

自由度 ν取自集合 $\{1, 2, 3, 5, 10\}$。当样本量为 $n = 100$ 和 $m = 1000$时，我们使用不同的随机种子对检验进行了十轮重复试验。拒绝率列于表 2中。为便于比较，我们转录了 Balakrishnan 等人报告的该情形下的检验功效值（见表 2 和表 5） [9] 以及 Gonz´alez-Estrada 等人 [21]的结果。可以看出，十轮检验的检验功效均高于 Balakrishnan 等人 [9]的检验。与 Gonz´alez-Estrada 等人 [21]的检验相比，除 $\nu = 10$外，我们的检验也表现出更高的检验功效。这表明，尤其是在以 ST 分布为备择来检验 SN 分布时，与文献中介绍的竞争检验相比，我们的检验在检验功效方面不逊色，甚至更优。

6 数据应用

本节将考察文献中此前介绍的一些真实数据样本示例。我们将应用并讨论本文提出的拟合优度检验的结果。我们对每项检验均采用 5% 的置信水平。下表中以粗体显示的 p值表示拒绝原假设。此外，为了提高每项检验的检验功效，我们设定 $m = \max\{n, 1000\}$ 与模拟 12 中呈现的结果一致。

6.1 AIS 数据集

澳大利亚体育学院（AIS）数据集是 Azzalini and Capitanio [4] 用于展示偏态正态分布拟合的经典示例之一。该数据集包含在澳大利亚体育学院采集的 100 名女性和 102 名男性运动员的生物医学测量数据，包括体质指数（BMI）、体脂百分比（BFP）、皮褶厚度总和（SSF）以及

表 2：样本量为 $n = 100$ 且数据来自偏斜 t 分布$_3(\boldsymbol{\xi}, \boldsymbol{\Omega}\ \boldsymbol{\alpha})$ 分布，其中 $\boldsymbol{\xi}$，$\boldsymbol{\Omega}$，且 $\boldsymbol{\alpha}$ 的定义见（5.3），并假设复合原假设属于偏正态族时，基于 1000 次重复试验模拟得到的检验功效。粗体数值表示估计的检验功效低于 Gonz´alez-Estrada 等人 [21]的检验。

ν （自由度）	1	2	3	5	10
Balakrishnan 等人 [9] 的检验	0.864	0.475	0.277	0.193	0.165
Gonz´alez-Estrada 等人 [21]的检验	1.000	-	0.979	0.758	0.299
我们的检验 - 第 1 轮	1.000	1.000	0.993	0.838	0.308
我们的检验 - 第 2 轮	1.000	1.000	0.988	0.797	**0.247**
我们的检验 - 第 3 轮	1.000	1.000	0.992	0.808	**0.273**
我们的检验 - 第 4 轮	1.000	0.999	0.989	0.801	**0.257**
我们的检验 - 第 5 轮	1.000	1.000	0.994	0.785	**0.281**
我们的检验 - 第 6 轮	1.000	1.000	0.990	0.816	0.321
我们的检验 - 第 7 轮	1.000	1.000	0.994	0.791	**0.282**
我们的检验 - 第 8 轮	1.000	1.000	0.991	0.778	**0.265**
我们的检验 - 第9轮	1.000	1.000	0.995	0.791	**0.238**
我们的检验 - 第10轮	1.000	1.000	0.993	0.806	**0.269**

瘦体重（LBM）等数据，可通过 R [44]中的 sn [3] 程序包获取。Balakrishnan 等人 [9] 和 Gonz´alez-Estrada 等人 [21] 最近也使用了上述四项指标来检验偏态正态分布的拟合优度。除了偏正态族的检验之外，我们还纳入了非对称 α-稳定分布（仅限二维情形）、Tukey g-and-h、偏斜拉普拉斯分布和偏斜 t 分布的检验。

我们将检验应用于按性别（女性和男性）分组的运动员二维和四维数据，并得到了表 3所示的估计 p 值。我们观察到，在四维情形下，无论是女性还是男性运动员，唯一未能拒绝原假设的检验是基于 GH 分布的检验。对于 SL 和 ST 分布，检验结果仅表明应拒绝女性运动员数据的原假设。此外，检验结果支持女性运动员数据服从 SN 分布，而对于男性运动员数据，检验则表明应拒绝 SN 分布。这两项检验因此与 Balakrishnan 等人 [9] 和 Gonz´alez-Estrada 等人[21]提出的结论一致。对于成对二维情形，我们观察到，大多数基于 AS 分布的检验都表明应拒绝原假设；12 对数据中有 10 对的 p 值低于 0.05。

表3：四维 AIS 数据集和成对二维 AIS 数据的检验所得 p值估计值。

四维 AIS 数据				
	GH	SL	SN	ST
女性运动员	0.715	**0.004**	0.068	**0.022**
男性运动员	0.175	0.503	**0.013**	0.247

成对二维 AIS 数据						
		AS	GH	SL	SN	ST
Female athletes	体质指数 & BFP	**0.002**	0.074	**0.003**	0.234	0.201
	体质指数 & SSF	**0.000**	**0.000**	**0.008**	**0.044**	**0.010**
	体质指数 & 瘦体重	**0.000**	0.618	0.203	0.102	0.236
	BFP & SSF	**0.000**	0.287	**0.001**	**0.049**	**0.037**
	BFP & 瘦体重	**0.000**	**0.000**	**0.010**	0.115	0.087
	SSF & 瘦体重	0.271	0.105	**0.012**	0.162	0.123
Male athletes	体质指数 & BFP	**0.000**	0.258	**0.033**	**0.001**	0.648
	体质指数 & SSF	0.257	0.058	0.209	**0.003**	0.778
	体质指数 & 瘦体重	**0.001**	0.157	**0.038**	0.300	0.486
	BFP & SSF	**0.022**	**0.018**	0.213	**0.002**	0.204
	BFP & 瘦体重	**0.000**	0.281	0.052	**0.032**	0.243
	SSF & 瘦体重	**0.041**	**0.001**	0.118	0.199	0.472

6.2 澳大利亚双胞胎样本生物统计数据的体质指数

从统计学角度来看，如今已明确，体质指数的总体分布并非对称，通常呈右偏，即体重与身高之比偏向较高值（参见 Nuttall [42]）。Marchenko 和 Genton [33]通过考察第 6.1 小节讨论的 AIS 数据中的体质指数观测值，提出了有力证据，表明偏度参数不为零。文献中也有研究指出（例如，Tran 等人 [49] 和 Tsang 等人 [50]），在描述单峰体重指数数据时，偏斜t分布具有相当强的竞争力。因此，作为第二个应用，我们考察了从`twinData` 数据集中获取的同卵双胞胎体质指数观测值。该数据集可通过 `OpenMx` [10] 软件包在 R [44]中获取。我们选择使用该数据集而非 AIS 数据，是因为它包含更多观测值，且数据更加同质。在分析中，我们纳入所有年龄段的个体，并按性别分组，其中包括 1171 对女性双胞胎和 532 对男性双胞胎；仅剔除了体质指数数据缺失的双胞胎对。

我们将观测到的体质指数二维向量拟合到第4节介绍的五种分布，并应用了拟合优度检验。估计得到的p值列于表4中。在5%的置信水平下，检验拒绝了女性和男性

单卵双胞胎数据中的偏斜正态分布和ST分布。不过，值得注意的是，ST检验的p值显著高于SN检验所得的p值，这印证了文献中的观点，即ST分布的拟合效果相当更好。至于α-稳定分布和SL分布，检验对两种性别得出了不明确的结果：在男性样本中拒绝了这两种分布，而在女性样本中则得到相对较高的p值。最后，对于女性和男性数据，针对GH分布的检验均未提供足够证据来拒绝原假设。

表4：澳大利亚单卵双胞胎样本数据集中二维体质指数对应检验的估计 p值。

	AS	GH	SL	SN	ST
单卵女性双胞胎	0.458	0.290	0.288	**0.000**	**0.042**
单卵男性双胞胎	**0.000**	0.216	**0.000**	**0.012**	**0.046**

6.3 风速数据

作为第三个也是最后一个示例，我们检验了风速数据集，该数据集由 Azzalini 和 Genton 提出[7]），包含 278 个观测值，即 2003年2月25日至11月30日记录并收集的每小时平均风速测量值，数据在午夜时分采集自三座气象塔：Goodnoe Hills（gh）、Kennewick（kw）和 Vansycle（vs）。这些气象塔位于哥伦比亚峡谷沿线及美国太平洋西北地区俄勒冈州与华盛顿州交界处。Azzalini 和 Genton [7] 提出使用独立同分布偏斜t三维模型拟合该数据，并声称该模型“显著优于正态分布”。Arslan 也使用了同一数据集 [2] 来展示偏斜拉普拉斯分布的拟合效果。作者考察了在气象塔 (gh, kw) 和 (vs, gh) 记录的风速所构成的二维向量，并指出，偏斜拉普拉斯分布能够令人满意地拟合散点图数据，同时捕捉到偏度和明显的重尾性。

我们对三维风速数据集进行拟合优度检验，以验证第 4 节介绍的偏斜模型中是否有不适用的模型。为获取更多信息，我们还对成对二维数据集应用了相同的检验，此时还包括 α-稳定分布。估计得到的 p 值列于表 5中。以 5%显著性水平为准，在三维情况下，只有针对 GH 分布的检验未提供足够证据拒绝原假设；而其他所有检验，即针对 SL、SN 和 ST 分布的检验，所得 p 值均低于 0.05，因此表明应拒绝这三种分布。在成对二维情况下，大多数检验都表明应拒绝原假设。这一结论可能是因为数据呈现出双峰性的迹象，采用分布混合或许更适合对该数据集建模。

表5：与三维和成对二维风速数据集对应的拟合优度检验的估计p值。

三维数据			
GH	SL	SN	ST
0.938	**0.000**	**0.000**	**0.000**

成对二维数据					
	AS	GH	SL	SN	ST
gh & kw	**0.006**	**0.000**	**0.000**	**0.002**	**0.006**
gh & vs	**0.001**	0.067	**0.000**	**0.000**	**0.004**
kw & vs	0.397	**0.001**	**0.000**	**0.000**	**0.000**

7 结论

本文提出了一种针对多种多元偏斜分布的拟合优度检验。一方面，本研究所述方法的主要优势在于其灵活性，可应用于任何多元参数分布族，前提是能够以合理的方法估计其参数，并且能够生成新的重复样本。另一方面，就计算成本而言，该方法的实现开销很大，因为参数自助法步骤要求在每次蒙特卡罗运行中额外进行一轮重抽样。

尽管文献中大多数拟合优度检验都需要进行此类嵌套重抽样，但这一缺点很容易通过采用并行算法来规避，因为参数自助法无需任何顺序处理程序；也可以采用 Giacomini 等人的扭曲速度自助法。[20]需要指出的是，所有检验均借助 Intel Xeon Gold 6230R CPU 运行，其中 104 个线程中的 100 个被密集使用，以进一步加快模拟的完成。我们通过五类多元分布展示了该方法的有效性，即多元偏斜正态分布、偏斜 t 分布、非对称偏斜拉普拉斯分布、偏斜 α-稳定分布和 Tukey g-与-h 分布（其中大多数没有可用的检验），并尽可能采用相应的典范形式。

第5.1 小节中的模拟结果表明（见表1），对于全部五类分布，检验的估计检验尺寸均合理且一致。同样，5.2和5.3小节中的模拟表明，该检验具有足够的检验功效，能够检测并拒绝备择假设。与文献中介绍的其他检验方案相比（见5.4小节），对于在 SN 分布这一复合原假设下进行检验的特定情形，我们的检验在针对备择 ST分布进行检验时也表现出更高的检验功效。

我们检验的有效性还通过第 6 节中的真实数据示例得到了说明，展示了将其应用于分别在日常生活和环境中观察到的生物现象和自然事件时的适用性与实用价值。

最后，我们想提醒读者，我们的检验在实际使用的核方面具有一定的灵活性 Ψ；请参阅第 2 节的最后一段。关于这一点，研究这种选择对我们检验的有限样本性质有何影响将会很有意思。这方面还需要进一步研究。

致谢

本研究得到了阿卜杜拉国王科技大学（KAUST）的资助。

利益冲突

作者　对本文讨论的任何材料均无经济或专有利益。

参考文献

[1] Abdul-Hamid, H., and Nolan, J. P. 多维稳定密度作为一维投影的函数。*Journal of Multivariate Analysis 67*, 1 (1998), 80–89.[2] Arslan, O. 一种替代的多元偏斜拉普拉斯分布：性质与估计。 *Statistical Papers*51, 4 (2010), 865–887.[3] Azzalini, A. `sn: the skew-normal and related distributions such as the skew-t and the SUN`, 2022. R 软件包版本 2.1.0. https://cran.r-project.org/web/packages/sn/.[4] Azzalini, A., and Capitanio, A. 多元偏斜正态分布的统计应用。 *Journal ofthe Royal Statistical Society: Series B 61*, 3 (1999), 579–602.[5] Azzalini, A., and Capitanio, A. *The Skew-Normal and Related Families*。剑桥大学出版社，纽约，2014.[6] Azzalini, A., and Dalla Valle, A. 多元偏斜正态分布。 *Biometrika 83*, 4 (1996),715–726.[7] Azzalini, A., and Genton, M. G. 基于偏斜 t 分布及相关分布的稳健似然方法。*International Statistical Review 76*, 1 (2008), 106–129.[8] Azzalini, A., Genton, M. G., and Scarpa, B. 针对偏斜对称分布的基于不变性的估计方程。 *Metron 68*, 3 (2010), 275–298.[9] Balakrishnan, N., Capitanio, A., and Scarpa, B. 基于典范形式的多元偏斜正态性检验。 *Journal of Multivariate Analysis 128* (2014), 19–32.[10] Boker, S. M., Neale, M. C., Maes, H. H., Spiegel, M., Brick, T. R., Estabrook, R., Bates, T. C., Gore,R. J., Hunter, M. D., Pritikin, J. N., Zahery, M., and Kirkpatrick, R. M. `OpenMx: extended structuralequation modelling`, 2022. R 软件包版本 2.20.7. https://cran.r-project.org/package=OpenMx.[11] Byczkowski, T., Nolan, J. P., and Rajput, B. 多维稳定密度的近似。 *Journal ofMultivariate Analysis 46*, 1 (1993), 13–31.[12] Capitanio, A. 关于偏态正态分布尺度混合的典范形式。 *Statistica 80*, 2 (2020), 145–160.[13] Chen, F., Jiménez-Gamero, M. D., Meintanis, S., and Zhu, L. 一种用于多元拟合优度检验的一般蒙特卡罗方法及其在椭圆分布族中的应用。 *Computational Statistics & Data Analysis 175*, 107548 (2022)。

[14] Chen, W., and Genton, M. G. 你们都服从正态分布吗? 这要视情况而定! *International Statistical Review* (2022), doi: 10.1111/insr.12512.[15] Ebner, B., Henze, N., and Strieder, D. 通过多元 Stein 方程中的傅里叶方法检验任意维数下的正态性。 *Canadian Journal of Statistics 50*, 3 (2021), 992–1033.[16] Fang, K. T., Kotz, S., and Ng, K. W. *Symmetric Multivariate and Related Distributions*。Chapman & Hall/CRC, 博卡拉顿, 1990.[17] Field, C., and Genton, M. G. 多元 g-and-h 分布。 *Technometrics 48*, 1 (2006), 104–111.[18] Flecher, C., Naveau, P., and Allard, D. 使用加权矩估计闭斜正态分布的参数。 *Statistics & Probability Letters 79*, 19 (2009), 1977–1984.[19] Fragiadakis, K., and Meintanis, S. G. 多元拉普拉斯分布的拟合优度检验。 *Mathematical and Computer Modelling 53*, 5–6 (2011), 769–779.[20] Giacomini, R., Politis, D. N., and White, H. 一种快速方法, 用于开展涉及自助法估计量的蒙特卡罗实验。 *Econometric Theory 29*, 3 (2013), 567–589.[21] González-Estrada, E., Villaseñor, J. A., and Acosta-Pech, R. 多元偏斜正态性的 Shapiro-Wilk 检验。 *Computational Statistics 37*, 4 (2022), 1985–2001.[22] He, Y., and Raghunathan, T. E. 使用多元 gh 变换进行多重插补。 *Journal of Applied Statistics 39*, 10 (2012), 2177–2198.[23] Henze, N. 多元正态性的不变检验: 一篇批判性综述。 *Statistical Papers 43*, 4 (2002), 467–506.[24] Henze, N., Jiménez-Gamero, M. D., and Meintanis, S. G. 多元正态性的刻画及相应的拟合检验, 包括针对 GARCH 模型的检验。 *Econometric Theory 35*, 3 (2019), 510–546.[25] Henze, N., and Wagner, T. 一种用于多元正态性 BHEP 检验的新方法。 *Journal of MultivariateAnalysis 62*, 1 (1997), 1–23.[26] Jeong, J., Yan, Y., Castruccio, S., and Genton, M. G. 基于 Tukey g-and-h 自回归过程的全球月度风能随机生成器。 *Statistica Sinica 29*, 3 (2019), 1105–1126.[27] Jiménez-Gamero, M. D., and Kim, H. 基于特征函数的快速拟合优度检验。 *Computational Statistics & Data Analysis 89* (2015), 172–191.[28] Jones, M. C., and Pewsey, A. 双曲正弦-反双曲正弦分布。 *Biometrika 96*, 4 (2009), 761–780.[29] Karling, M. J., Lopes, S. R. C., and de Souza, R. M. 多元 α-稳定分布: VAR(1) 过程、依赖性度量及其估计。 *Journal of Multivariate Analysis 195*, 105153 (2023).[30] Kim, H., and Genton, M. G. 偏态正态分布尺度混合的特征函数。*Journal of Multivariate Analysis 102*, 7 (2011), 1105–1117.[31] Kotz, S., Kozubowski, T. J., and Podgórski, K. *The Laplace Distribution and Generalizations*。Birkh¨auser, 波士顿, 2001.[32] Kozubowski, T. J., Podgórski, K., and Rychlik, I. 多元广义拉普拉斯分布及相关随机场。 *Journal of Multivariate Analysis 113* (2013), 59–72.[33] Marchenko, Y. V., and Genton, M. G. 一套用于拟合偏斜正态与偏斜 t 模型的命令。*The Stata Journal 10*, 4 (2010), 507–539.[34] Meintanis, S., and Swanepoel, J. 基于经验变换且含估计参数的自助法拟合优度检验。 *Statistics & Probability Letters 77*, 10 (2007), 1004–1013.[35] Meintanis, S. G., and Hlávka, Z. 二元与多元斜正态分布的拟合优度检验。*Scandinavian Journal of Statistics 37*, 4 (2010), 701–714.

[36] Meintanis, S. G., Ngatchou-Wandji, J., and Taufer, E. 基于经验特征函数的多元稳定分布拟合优度检验。 *Journal of Multivariate Analysis 140* （2015），171–192。[37] Micchelli, C. A., Xu, Y., and Zhang, H. 通用核。 *Journal of Machine Learning Research 7*，95（2006），2651–2667。[38] Modarres, R., and Nolan, J. P. 稳定随机向量的模拟方法。 *Computational Statistics 9*，1（1994），11–19。[39] Möstel, L., Fischer, M., Pfälzner, F., and Pfeuffer, M. Tukey 型分布的参数估计：比较分析。 *Communications in Statistics - Simulation and Computation 50*，4（2021），957–992。[40] Nolan, J. P. 多元椭圆等高稳定分布：理论与估计。 *ComputationalStatistics 28*，5（2013），2067–2089。[41] Nolan, J. P., Panorska, A. K., and McCulloch, J. H. 稳定谱测度的估计。 *Mathematicaland Computer Modelling 34*，9-11（2001），1113–1122。[42] Nuttall, F. Q. 体质指数：肥胖、BMI 与健康。 *Nutrition Today 50*，3（2015），117–128。[43] Pudełko, J. 关于一种新的仿射不变量多元正态性相合检验。 *Probability and MathematicalStatistics 25*，1（2005），43–54。[44] R Core Team。 `R: a language and environment for statistical computing`，2022。R 统计计算基金会，维也纳，奥地利。https://www.R-project.org/.[45] R Core Team and contributors worldwide。 `stats: the R stats package`，2022。R 程序包版本 4.3.0。 https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html.[46] Ripley, B. `MASS: support functions and datasets for venables and Ripley's MASS`，2022。R 程序包版本 7.3-58.1。 https://cran.r-project.org/package=MASS。[47] Samorodnitsky, G., and Taqqu, M. S. *Stable Non-Gaussian Random Processes: Stochastic models with infinitevariance*。Chapman & Hall/CRC，博卡拉顿，2000。[48] Székely, G. J., and Rizzo, M. L. 能量统计：一类基于距离的统计量。 *Journal of Statistical Planning and Inference 143*，8（2013），1249–1272。[49] Tran, T., Wiskow, C., and Aziz, M. A. 偏态和灵活偏斜分布：从现代视角看体质指数分布。 *American Journal of Undergraduate Research 14*，2（2017），45–63。[50] Tsang, S., Duncan, G. E., Dinescu, D., and Turkheimer, E. 体质指数（BMI）偏度下的双生子相关微分模型。 *PLoS ONE 13*，3（2018），e0194968。[51] Yan, Y., and Genton, M. G. 采用 Tukey g-and-h 变换的非高斯自回归过程。 *Envi-ronmetrics 30*，2（2019），e2503。[52] Yan, Y., Jeong, J., and Genton, M. G. 多元变换高斯过程。 *Japanese Journal ofStatistics and Data Science 3*，1（2020），129–152。