arXiv:2602.04318v1 [stat.ME] 4 Feb 2026

Rao 间距检验零分布的精确高效近似

Yoshiki Kinoshita[1], **AyaShinozaki**[2] **和 Toshinari Kamakura**[3]

yoshikik.92g@g.chuo-u.ac.jp; [2]shinozaki.223@g.chuo-u.ac.jp; [3]
ktoshinari001f@g.chuo-u.ac.jp[1,2,3]日本东京文京区春日 1-13-27，中央大学，邮编 112-8551

摘要

Rao 间距检验是一种广泛使用的非参数方法，用于评估圆周上的均匀性。然而，由于零分布难以计算，该方法在实际应用中的适用范围受到限制。因此，研究者通常依赖预先制表的临界值，而这些临界值仅针对有限的样本量计算，限制了该方法的灵活性和通用性。本文通过递归计算 Rao 间距检验统计量的高阶矩，并采用格拉姆-沙利叶展开，推导出其零分布的精确近似，从而解决这一局限。该方法能够高效、直接地计算任意样本量下的 P 值，无需依赖现有的临界值表。此外，我们证实，即使对于现有表格未涵盖的大样本量，我们的方法仍然准确有效，从而克服了一个重要的实际局限。与已发表的临界值和鞍点近似进行比较评估后发现，我们的方法在广泛的样本量范围内均具有很高的准确性。这些发现大幅提升了 Rao 间距检验在理论研究和应用统计分析中的实用性。

1 引言

圆周上的均匀性检验是方向数据分析中的一个基本问题。对于单峰分布，文献中广泛使用冯·米塞斯分布 Mardiaand Jupp (2009)；Jammalamadaka 和 Sengupta (2001)。由似然比检验推导出的均匀性雷利检验，已知在冯·米塞斯分布下具有较高检验效能。然而，当数据呈多峰分布时，Rao 间距检验等非参数方法已被证明比雷利检验等参数方法更有效。

由于难以通过数值方法计算原假设下 Rao 间距检验统计量的分布，该检验在实际中的应用有限。因此，使用者必须查阅由Russell and Levitin (1995)提供的表格临界值，而这些临界值仅适用于特定样本量。这一限制严重制约了该检验的灵活性。

本文递归计算 Rao 间距检验统计量的高阶矩，并应用格拉姆-沙利叶展开来近似其零分布。借此可以直接计算 P 值，从而使 Rao 间距检验能够适用于任意样本量。我们通过将推导出的近似值与Russelland Levitin (1995)的现有表格以及Gatto 和 Jammalamadaka (1999)的表格进行比较，验证了我们的方法。比较结果表明，我们提出的方法达到了与他们的方法相当或更高的准确度。此外，现有表格的样本量上限为 $n = 1000$，对于更大的样本，通常使用 $n = 1000$ 的临界值作为近似值。然而，这种替代的影响尚未得到充分研究。本研究通过比较我们提出的方法与现有方法在较大样本下的表现，具体考察 $n = 10000$时这种近似的影响。

2 Rao 间距检验的矩

Rao 间距检验统计量在检验圆周上的均匀性时发挥着核心作用。尽管其概率密度函数（PDF）已知，但在实际应用中，概率密度函数和累积分布函数（CDF）都难以计算。因此，Russell and Levitin (1995)针对特定样本量和角度设置制成了临界值表，并通过 R 和 MATLAB 实现提供这些数值。该方法基于数值积分，在样本量较小时效率较高。然而，随着样本量增加，递归项的计算成本变得高昂，使得实时计算不切实际。因此，用户通常只能使用预计算表中包含的离散设置。为解决这一局限，我们推导了原假设下 Rao 间距检验统计量的高阶矩

，并利用格拉姆-沙利叶展开构造近似分布。这种方法无需依赖表格值，即可针对任意样本量进行检验。

This 本节首先介绍 Rao 间距检验统计量。随后，我们给出其一般表达式 ，并推导其高阶矩。最后，我们推导累积分布函数的格拉姆-沙利叶展开。

2.1 Rao 间距检验统计量

设 α_1、.....、 α_n 为圆周上的观测值，并按以下方式排序：

$$0 \leq \alpha_{(1)} \leq \cdots \leq \alpha_{(n)} < 2\pi.$$

Rao 间距检验统计量定义如下：

$$U_n = \frac{1}{2} \sum_{i=1}^{n} \left| T_i - \frac{2\pi}{n} \right|, \tag{1}$$

其中 $T_i = \alpha_{(i+1)} - \alpha_{(i)} (i = 1$、.....、 $n-1)$ 和 $T_n = \alpha_{(1)} - \alpha_{(n)} + 2\pi$。

W虽然 U_n 的精确值概率密度函数已知，但计算起来非常耗时。其密度函数 ion U_n 的 已知为：

$$f_n(u) = (n-1)! \sum_{j=1}^{n-1} \binom{n}{j} \left(\frac{u}{2\pi}\right)^{n-j-1} \frac{\phi_j(nu)}{(n-j-1)! n^{j-1}} \quad \left(0 \leq u \leq 2\pi \left(1 - \frac{1}{n}\right)\right),$$

其中

$$\phi_j(x) = \frac{1}{2\pi(j-1)!} \sum_{k=0}^{j-1} (-1)^k \binom{j}{k} \left[\max\left(0, \frac{x}{2\pi} - k\right)\right]^{j-1}.$$

这里， $\phi_j(x)$ 表示厄温-霍尔分布，即在 $[0, 2\pi]$ 上独立均匀随机变量之和的分布。

2.2 高阶矩的推导

Rao (1969) 通过修改 Sherman (1950) 最初报告的数值，给出了 Rao 间距检验统计量临界点表。然而，该表仅涵盖样本量从 $n = 2$ 到 $n = 20$的范围，且仅包含三个显著性水平（$\alpha = 0.01$，0.05 和 0.10）。这一局限源于检验统计量的精确密度函数涉及复杂表达式，因此难以通过解析方法计算临界值。

为克服这些局限，Russell and Levitin (1995) 使用梯形法则，以数值方法计算样本量从 $n = 4$ 到 $n = 1000$范围内的临界值。他们提出了一种递归方法，以更高效地计算梯形积分中的一部分。尽管如此，计算临界点仍然耗时，因此无法实际即时计算，必须使用预计算表。

对此，我们推导出一种递归方法，用于计算 Rao 间距检验统计量的高阶矩，并应用格拉姆-沙利叶展开这一渐近级数来近似其分布。这种方法使得快速计算 P 值成为可能，而此前由于分布的复杂性，这一计算难以实现。因此，我们的方法带来了显著的计算优势，尤其适用于大规模应用。

在本小节中，我们推导以下定理，以递归方式计算 Rao 间距检验统计量的 r 阶矩。

定理 1。

$$E(U_n^r) = \frac{(2\pi)^r}{n^{n+r-1}(n)_{\overline{r}}} \sum_{j=1}^{r} a_j^{(r)} (n)_{\underline{j}} (n-j)^{n+r-1}$$

where $a_j^{(r)}$ *satisfies:*

$$\begin{aligned} a_r^{(r)} &= 1, \\ a_1^{(r+1)} &= (r+1) a_1^{(r)}, \\ a_j^{(r+1)} &= (r+j) a_j^{(r)} + a_{j-1}^{(r)} \qquad (j = 2, \ldots, r), \end{aligned}$$

and for all $n \in \mathbb{N}_{\geq 2}$ *and* $r \in \mathbb{N}_{\geq 1}$.

注 1. *Theorem 1 is expressed as a linear combination of* $(n)_j(n-j)^{n+r-1}_{_}, (j=1,\ldots,r)$ *and the coefficients are as follows:*

$$\frac{(2\pi)^r}{n^{n+r-1}(n)_{\overline{r}}} a_j^{(r)} (n)_{\underline{j}} \quad (j=1,\ldots,r).$$

定理1的证明见附录A。

通过应用定理1中的通用公式，并使用表5中列出的系数，可计算出一阶矩和二阶矩，如下所示。

$$\begin{aligned} E(U_n) &= 2\pi \left(1-\frac{1}{n}\right)^n, \\ E(U_n^2) &= \frac{(2\pi)^2}{n^{n+1}(n+1)} \left(2(n-1)^{n+1} + (n-1)(n-2)^{n+1}\right). \end{aligned}$$

2.3 累积分布函数的格拉姆-沙利叶展开

Us给定原始矩 $\mu^r = E(U_n^r)$，我们计算累积量并构造格拉姆-沙利叶展开 ：

$$\begin{aligned} f_n(u) &= \sum_{j=3}^{\infty} B_j(0,0,\kappa'_{n,3},\ldots,\kappa'_{n,j}) \frac{H_j(u)}{j!}\phi(u), \\ F_n(u) &= \Phi(u) - \phi(u) \sum_{j=3}^{\infty} B_j(0,0,\kappa'_{n,3},\ldots,\kappa'_{n,j}) \frac{H_{j-1}(u)}{j!}, \end{aligned}$$

其中， ϕ 和 Φ 分别是标准正态概率密度函数和累积分布函数，而 B_j 是 j阶完全指数贝尔多项式。此外， $H_j(j=0,$ 1，.....)表示厄米特多项式，定义如下：

$$\begin{aligned} H_0(x) &= 1, \\ H_{(j+1)}(x) &= xH_j(x) - \frac{d}{dx}H_j(x) \qquad (j=0,1,\ldots). \end{aligned}$$

该展开无需依赖列表中的临界值，即可准确计算P 值。当 n 较小时（例如6），这种近似可能不够准确；在这种情况下，可以通过直接对 $f_n(u)$的概率密度函数积分得到精确结果。

3 与现有方法的比较

本节通过与现有方法进行比较，评估所提出的 Gram–Charlier 近似的性能。具体而言，我们将我们的方法与 Russell andLevitin (1995) 提供的表列值，以及 Gatto 和 Jammalamadaka (1999) 提出的鞍点近似进行比较。

表 1 列出了样本量固定为 $n=10$ 时，Rao 间距检验统计量在若干取值下的累积分布函数值。$\Pr_E(T<t)$ 列给出Russelland Levitin (1995) 的精确值，而 $\Pr_{LR}(T<t)$、$\Pr_{BN}(T<t)$ 和 $\Pr_{ID}(T<t)$ 则表示Gatto 和 Jammalamadaka (1999) 给出的鞍点近似。$\Pr_{G_{10}}(T<t)$ 列给出使用最高至 10 阶矩计算的 Gram–Charlier 近似。结果表明，使用最高至 10 阶矩的 Gram–Charlier 展开在整个取值范围内都与精确概率吻合。此外，我们的方法在所有情况下的准确性均优于鞍点近似。

4 个示例

我们采用推导出的方法，对Russell and Levitin (1995)中的示例进行均匀性检验。该数据集包含一名癫痫患者的假想数据，记录了患者在多次发作时的时间。角度值如下：

$$5^\circ, 10^\circ, 10^\circ, 12^\circ, 17^\circ, 85^\circ, 90^\circ, 99^\circ, 100^\circ, 110^\circ, 153^\circ, 233^\circ, 235^\circ, 296^\circ, 331^\circ.$$

在我们的方法中，使用10阶矩（$r=10$时，该示例的样本量为 $n=15$ 且$U_{15}=3.089233(=177^\circ)$，如公式（1）所示。相应的P值为0.0174，因此拒绝均匀性原假设。

我们还采用推导出的方法，对Mardia and Jupp (2009)中的示例6.9进行均匀性检验。该示例涉及一项鸽子归巢实验：研究人员放飞了13只鸟，并记录了它们的消失角，如下所示：

$$20^\circ, 135^\circ, 145^\circ, 165^\circ, 170^\circ, 200^\circ, 300^\circ, 325^\circ, 335^\circ, 350^\circ, 350^\circ, 350^\circ, 355^\circ.$$

表 1：Gram–Charlier 近似与鞍点近似的准确度比较。($n = 10$)

角度	精确值	鞍点近似			我们的方法
t	$\Pr_E(T < t)$	$\Pr_{LR}(T < t)$	$\Pr_{BN}(T < t)$	$\Pr_{ID}(T < t)$	$\Pr_{G_{10}}(T < t)$
50°	.001	.001	.001	0	.001
60°	.004	.004	.004	.004	.004
70°	.015	.015	.015	.015	.015
80°	.042	.041	.041	.043	.042
90°	.093	.093	.092	.096	.093
100°	.178	.176	.175	.182	.178
110°	.294	.292	.291	.300	.294
120°	.433	.430	.429	.439	.433
130°	.577	.580	.578	.583	.577
140°	.708	.706	.705	.713	.708
150°	.815	.813	.812	.819	.815
160°	.892	.891	.890	.895	.892
170°	.943	.942	.941	.945	.943
180°	.972	.972	.971	.973	.972
190°	.988	.987	.987	.989	.988
200°	.995	.995	.995	.996	.995
210°	.998	.998	.998	.999	.998
220°	.999	.999	.999	1.000	.999

在我们的方法中，使用 10阶矩（$r = 10$）时，该示例的样本量为 $n = 13$ 和 $U_{13} = 2.826091 (= 161.9231°)$，如方程（1）所示。对应的 P值为 0.0786，因此不会拒绝均匀性原假设。我们的方法可以计算 P值，从而无需借助统计表即可进行均匀性检验。相比之下，Mardia and Jupp (2009) 参考 Rao (1969) 提供的表格来获得均匀性检验的结果。他们得出结论：在 5%显著性水平下，均匀性假设未被拒绝，因为检验统计量 $U_{13} = 162°$ 介于 Rao（1969）表中的 5%临界值（167.76°）和 10% 临界值（158.40°）之间。

我们还对来自Kitamura （1989）的一个例子进行均匀性检验，该检验基于我们所推导的方法。该例子涉及Kamiokande中微子数据，数据集如下：

$$30°, 36°, 60°, 64°, 76°, 98°, 136°, 140°, 182°, 216°, 244°, 270° (n = 12),$$
$$30°, 36°, 60°, 64°, 76°, 98°, 140°, 182°, 216°, 244°, 270° (n = 11).$$

在我们的方法中，当使用1阶至10阶矩（$r = 10$）时，该例子得到的样本量为 $n = 12$ 和 $U_{12} = 1.989675$（$= 114°$），以及样本量为 $n = 11$ 和 $U_{11} = 1.869089$（$= 107.0909°$），如式（1）所示。相应的P 值为 0.685（$n = 12$）和0.762（$n = 11$），这并未导致拒绝均匀性原假设。我们的方法能够计算P值，从而使均匀性检验无需借助统计表即可进行。相比之下，Kitamura （1989）参考了Rao （1969）所提供的表来获得均匀性检验的结果。他指出，Kamiokande数据的检验统计量均低于临界值 $\alpha = 0.10$，因此无法揭示偏离随机性的显著性。他还声称，Rao的表中没有与该统计量对应的P 值。因此，能够计算P 值的我们的方法，被认为有助于判断在Kamiokande数据中是否应拒绝原假设。就Kamiokande数据而言，所得到的P 值都非常大，从而使我们得出不应拒绝该假设的结论。

5 模拟研究

在 R 的 circular软件包中，Rao 均匀性检验采用 Russell and Levitin (1995) 表格，该表仅提供样本量不超过 $n = 1000$的临界值。当样本量大于 1000 时进行检验，会使用 $n = 1000$ 的临界值作为替代。为评估这一做法的影响，我们以样本量 $n = 10000$ 进行模拟，并将我们提出的方法与 R 的 circular软件包进行比较。

对于 $n = 1000$，Russell and Levitin (1995) 表中的拒绝临界值与我们提出的方法所得结果一致。对于 $n = 10000$，$d = 10$，临界值列于表 2。

表2：临界值。

显著性水平	0.001	0.01	0.05	0.10
Russell and Levitin (1995) ($n = 1000$)	140.99	138.84	136.94	135.92
我们提出的方法（$n = 1000$）	140.99	138.84	136.94	135.92
Russell and Levitin (1995) ($n = 2000$)	NA	NA	NA	NA
我们提出的方法（$n = 2000$）	138.49	136.97	135.63	134.91
Russell and Levitin (1995) ($n = 10000$)	NA	NA	NA	NA
我们提出的方法（$n = 10000$）	135.14	134.47	133.87	133.55

对于 $n = 10000$，临界值小于 $n = 1000$，这表明随着样本量增大，拒绝域也会扩大。如果样本量为 10000 时仍用 $n = 1000$的表格代替，则在本应拒绝原假设的情况下，可能会导致无法拒绝原假设。

为比较拒绝率，我们在单位圆上生成一个均匀分布，其中 $n = 10000$ 并以 $\alpha = 0.05$为显著性水平进行检验。该过程重复 10000 次，拒绝次数记录于表 3。

表 3：接受次数和拒绝次数。

	接受	拒绝	重复次数
circular软件包	10000	0	10000
我们提出的方法	9482	518	10000

结果表明，我们提出的方法的拒绝率通常与设定的显著性水平（$\alpha = 0.05$）一致，而 circular软件包则完全没有拒绝。由于该检验针对均匀分布数据进行，因此没有拒绝并不是问题。不过，这表明使用 circular软件包时，对于本应被拒绝的分布，其检验效能预计会较低。

此处，我们比较样本量为10000、参数为（均值参数 $\theta = 0$、集中参数 $\kappa = 0.3$）的冯·米塞斯分布的检验效能，结果见表 4。

表4：接受次数和拒绝次数。

	接受	拒绝	重复次数
circular软件包	9624	376	10000
我们提出的方法	514	9486	10000

结果表明，我们提出的方法的拒绝率较高，而 circular软件包的拒绝率较低，这说明我们提出的方法具有更高的检验效能。

6 讨论与结论

我们的研究使得此前无法实现的、在圆周均匀性检验中计算 Rao 检验统计量 P 值成为可能。迄今为止，使用该检验的唯一可行方法是查阅表格值($n \leq 1000$)，这些表格值见于基于 Russell and Levitin (1995)编制的 R circular 软件包。然而，有了我们推导出的 P 值，就不再需要查阅这些表格。因此，与使用现有表格的传统方法相比，我们的新方法有望推动多个应用领域取得新的进展。

此外，Rao 间距检验向球面的扩展尚未实现。涉及球面角度的问题，例如岩石中古地磁的方向、地球指向恒星的方向，以及石英晶体中光轴的方向，都是此类问题的例子（Mardia and Jupp，2009）。因此，将 Rao 间距检验扩展至球面是未来研究的一项重要挑战。

附录A 定理证明1

为计算矩，我们引入波赫哈默符号以及以下与第二类斯特林数相关的引理。

首先，波赫哈默符号定义如下 ows:

$$(n)_{\overline{r}} = n(n+1)\cdots(n+r-1),$$
$$(n)_{\underline{r}} = n(n-1)\cdots(n-r+1),$$

其中 $(n)_0 \equiv (n)_{\underline{0}} = 1$。

接下来，我们引入一个引理，用于表达多项式与斯特林数之间的关系。**引理 1**。 *(Aigner and Axler, 2007)*

$$x^n = \sum_{k=0}^{n} (x)_{\underline{k}} S_2(n,k).$$

Note that we can stop the summation at n, since $S_2(n,k) = 0$ *for* $k = n+1, n+2, \ldots$.

引理 2。 *(Aigner and Axler, 2007) The recurrence relation for the Stirling numbers of the second kind is given as follows:*

$$S_2(n+1,j) = S_2(n,j-1) + jS_2(n,j).$$

这里，第二类斯特林数记作 $S_2(n,k)$。 $S_2(n,k)$ 表示一个 n 元集合的 k 划分数。根据定义， $S_2(0,0) = 1$, $S_2(0,k) = 0 (k = 1,2,\ldots)$。

这里，我们给出以下涉及 N¨orlund 多项式的展开式，以推导 U_n 的矩。**引理 3**。 *The* r*th moment of* U_n *is given by*

$$E(U_n^r) = \frac{(2\pi)^r}{n^{n+r-1}} \sum_{j=1}^{n-1} (n)_{\underline{j}} \binom{n-1}{j} B_{n+r-j-1}^{(-j)},$$

where $B_n^{(z)}$ *is the Nörlund polynomial, defined as the coefficient in the expansion of the following expression (Carlitz, 1960).*

$$\left(\frac{x}{e^x - 1}\right)^z = \sum_{n=0}^{\infty} B_n^{(z)} \frac{x^n}{n!}.$$

在引理3的方程中，N¨orlund多项式 $B_n^{(z)}$可以用第二类斯特林数 S_2 表示如下（Quaintance和Gould，2015）：

$$S_2(n+z,z) = \binom{n+z}{n} B_n^{(-z)}.$$

Then

$$\begin{aligned}
E(U_n^r) &= \frac{(2\pi)^r}{n^{n+r-1}} \sum_{j=1}^{n-1} (n)_{\underline{j}} \binom{n-1}{j} B_{n+r-j-1}^{(-j)} \\
&= \frac{(2\pi)^r}{n^{n+r-1}} \sum_{j=1}^{n-1} (n)_{\underline{j}} \binom{n-1}{j} \frac{S_2(n+r-1,j)}{\binom{n+r-1}{j}} \\
&= \frac{(2\pi)^r}{n^{n+r-1}} \sum_{j=1}^{n-1} \frac{(n)_{\underline{j}} (n-1)_{\underline{j}}}{(n+r-1)_{\underline{j}}} S_2(n+r-1,j) \\
&= \frac{(2\pi)^r}{n^{n+r-1}} I_{n,r,0},
\end{aligned}$$

其中

$$I_{n,r,u} = \sum_{j=1}^{n-1} A_{n,j,r} (n)_{\underline{j}} S_2(n-1+r+u,j),$$

with

$$A_{n,j,r} = \frac{(n-1)!}{(n-1-j)!} \frac{(n-1+r-j)!}{(n-1+r)!}.$$

为了递归计算矩，我们引入以下定理。利用递推关系和斯特林数的性质，我们建立了以下定理：

定理 2。

$$I_{n,r,u} = \frac{1}{(n)_{\overline{r}}} \sum_{j=1}^{r} a_j^{(r)} (n)_{\underline{j}} (n-j)^{n+r-1+u} \tag{2}$$

where $a_j^{(r)}$ *satisfies:*

$$\begin{aligned} a_r^{(r)} &= 1, \\ a_1^{(r+1)} &= (r+1) a_1^{(r)}, \\ a_j^{(r+1)} &= (r+j) a_j^{(r)} + a_{j-1}^{(r)} \qquad (j = 2, \ldots, r). \end{aligned}$$

Proof. 该证明通过对 r 使用归纳法完成。

对于 $r = 1$，由引理 1可得以下结果，该引理刻画了第二类斯特林数与某些多项式之间的关系。

$$I_{n,1,u} = \frac{1}{(n)_{\overline{1}}} \sum_{j=1}^{1} (n)_{\underline{j}} (n-j)^{n+1-1+u}.$$

假设 (2) 对 r 成立，下面证明它对 $r+1$ 也成立。利用引理 2 以及 $I_{n、r、u}$ 和$A_{n、j、r}$的定义，可得:

$$I_{n,r+1,u} = I_{n,r,u+1} - D_1 + D_2,$$

其中

$$\begin{aligned} D_1 &= \frac{1}{n+r} \sum_{j=1}^{n-1} A_{n,j,r} (n)_{\underline{j}} S_2(n+r+u+1, j), \\ D_2 &= \frac{1}{n+r} \sum_{j=1}^{n-1} A_{n,j,r} (n)_{\underline{j}} S_2(n+r+u, j-1). \end{aligned}$$

此外，我们有

$$D_1 = \frac{1}{n+r} I_{n,r,u+2} \tag{3}$$

and

$$\begin{aligned} D_2 &= \frac{n}{n+r} \sum_{j=1}^{n-1} A_{n,j,r} (n-1)_{\underline{j-1}} S_2(n+r+u, j-1) \\ &= \frac{n}{n+r} \frac{n-1}{n+r-1} I_{n-1,r,u+2}. \end{aligned} \tag{4}$$

对于 $n = 2$，求和指标并不恰当；但这并非关键，因为 $I_{1、r、u+2} = 0$。综合以上结果并进行必要的计算，可得如下结果。

$$\begin{aligned} I_{n,r+1,u} &= I_{n,r,u+1} - \frac{1}{n+r} I_{n,r,u+2} + \frac{n}{n+r} \frac{n-1}{n+r-1} I_{n-1,r,u+2} \\ &= \frac{1}{(n)_{\overline{r+1}}} \sum_{j=1}^{r+1} a_j^{(r+1)} (n)_{\underline{j}} (n-j)^{n+r+u}, \end{aligned} \tag{5}$$

其中

$$\begin{aligned} a_{r+1}^{(r+1)} &= 1, \\ a_1^{(r+1)} &= (r+1) a_1^{(r)}, \\ a_j^{(r+1)} &= (r+j) a_j^{(r)} + a_{j-1}^{(r)} \qquad (j = 2, \ldots, r). \end{aligned}$$

由上可知，通过数学归纳法，等式（2）对任意 r 均成立。 □

附录B 格拉姆-沙利叶展开

在定理 1中，我们递归地计算了原始矩 $\mu'_{n,\,r} := E\left(U_n^r\right)$ ，对所有 $n = 2,\ 3,\ \ldots\ldots$以及 $r = 1,\ 2,\ \ldots\ldots$。此外，原始累积量 $\kappa'_{n,\,r}$ 可由原始矩（Smith，1995）得到：

$$\kappa'_{n,r} = \mu'_{n,r} - \sum_{k=1}^{r-1} \binom{r-1}{k-1} \kappa'_{n,k} \mu'_{n,r-k},$$

对所有 $n = 2,\ 3,\ \ldots$ 和 $r = 1,\ 2,\ \ldots$。在本小节中，我们利用累积量 $\kappa'_{n,\,r}$，根据格拉姆-沙利叶展开计算 U_n 的分布函数。令 ϕ 为标准正态分布的概率密度函数， D 为微分算子。则概率函数 $f_n(u)$ 可利用累积量 $\kappa'_{n,\,r}$ (Stuart和Ord，2010) 展开如下：

$$f_n(u) = \exp\left(\sum_{j=3}^{\infty} \kappa'_{n,j} \frac{(-D)^j}{j!}\right) \phi(u).$$

此外，右侧还可以用第 j 个指数型完全贝尔多项式 B_j 表示。

$$f_n(u) = \sum_{j=3}^{\infty} B_j(0, 0, \kappa'_{n,3}, \ldots, \kappa'_{n,j}) \frac{(-D)^j}{j!} \phi(u),$$

其中， B_j 表示定义1和定义2中定义的贝尔多项式。

定义1 （指数型部分贝尔多项式）。

$$\begin{aligned} B_{j,k}(x_1, \ldots, x_{j-k+1}) &= \sum \frac{j!}{c_1! \cdots c_{j-k+1}! (1!)^{c_1} \cdots ((j-k+1)!)^{j-k+1}} x_1^{c_1} \cdots x_{j-k+1}^{c_{j-k+1}} \\ &= j! \sum \prod_{i=1}^{j-k+1} \frac{x_i^{c_i}}{(i!)^{c_i} c_i!}, \end{aligned}$$

where the summation takes place over all integers c_1， ...， $c_{j-k+1} \geq 0$*(partitions of* j *into* k *summands), suchthat:*

$$\begin{aligned} c_1 + \cdots + c_{j-k+1} &= k, \\ c_1 + \cdots + (j-k+1) c_{j-k+1} &= j. \end{aligned}$$

定义2 （指数型完全贝尔多项式）。

$$B_j = \sum_{k=1}^{j} B_{j,k} \ .$$

我们使用指数型完全贝尔多项式 B_n （Comtet，2012）。

令 $H_j(j = 0,\ 1,\ \ldots\ldots)$ 为 Hermite 多项式

$$\begin{aligned} H_0(x) &= 1, \\ H_{(j+1)}(x) &= x H_j(x) - \frac{d}{dx} H_j(x) \qquad (j = 0, 1, \ldots). \end{aligned}$$

则，概率密度函数 $f_n(u)$ 可重写为

$$f_n(u) = \sum_{j=3}^{\infty} B_j(0, 0, \kappa'_{n,3}, \ldots, \kappa'_{n,j}) \frac{H_j(u)}{j!} \phi(u),$$

因为

$$D^j \phi(u) = (-1)^j H_j(u) \phi(u).$$

对两边积分，得到如下结果：

$$F_n(u) = \Phi(u) - \phi(u) \sum_{j=3}^{\infty} B_j(0, 0, \kappa'_{n,3}, \ldots, \kappa'_{n,j}) \frac{H_{j-1}(u)}{j!},$$

其中， Φ 表示标准正态分布的分布函数， F 表示 U_n的分布函数。概率 $\Pr(U_n > U_n^*)$可由该无穷级数截至第 d 项的部分和近似得到。

$$\Pr(U_n > U_n^*) \simeq 1 - \Phi(u) + \phi(u) \sum_{j=3}^{d} B_j(0, 0, \kappa'_{n,3}, \ldots, \kappa'_{n,j}) \frac{H_{j-1}(u)}{j!}.$$

附录 C U_n的高阶矩

利用第 2.2 节推导出的递推关系，我们计算了截至 10 阶的 r阶 Rao 间距检验统计量的矩的系数。表 5 展示了系数矩阵 $a_j^{(r)}$，其中每一行对应一个矩阶 r，各列表示系数 $a_j^{(r)}$。

表 5: r阶矩的系数。 $\left(a_j^{(r)}\right)$

r	$a_1^{(r)}$	$a_2^{(r)}$	$a_3^{(r)}$	$a_4^{(r)}$	$a_5^{(r)}$	$a_6^{(r)}$	$a_7^{(r)}$	$a_8^{(r)}$	$a_9^{(r)}$	$a_{10}^{(r)}$
1	1	0	0	0	0	0	0	0	0	0
2	2	1	0	0	0	0	0	0	0	0
3	6	6	1	0	0	0	0	0	0	0
4	24	36	12	1	0	0	0	0	0	0
5	120	240	120	20	1	0	0	0	0	0
6	720	1800	1200	300	30	1	0	0	0	0
7	5040	15120	12600	4200	630	42	1	0	0	0
8	40320	141120	141120	58800	11760	1176	56	1	0	0
9	362880	1451520	1693440	846720	211680	28224	2016	72	1	0
10	3628800	16329600	21772800	12700800	3810240	635040	60480	3240	90	1

T这些系数可用于根据给定的一般公式计算 U_n 的 r阶原始矩 in 理论 em 1.

参考文献

Aigner , M. 和 Axler , S. (2007). *A course in enumeration*, 第 238 卷。Springer , 柏林。

Carlitz, L. (1960). 关于 N¨orlund 多项式 $B_n^{(z)}$ 的注记。 *Proceedings of the American Mathematical Society*, 11 3 452 455( ): – . Comtet, L. (2012). *Advanced Combinatorics: The art of finite and infinite expansions*. 施普林格科学与商业媒体。Gatto, R. 和 Jammalamadaka, S. R. (1999). 用于检验问题的条件鞍点近似。 *Journal of the American Statistical Association*, 94(446):533–541. Jammalamadaka, S. R. 和 Sengupta, A. (2001). *Topics in circular statistics*, 第 5 卷。世界科学出版社，新加坡。Kitamura, T. (1989). 对 SN 1987A 中微子暴发角分布的圆形统计检验。*Kinki University Research Institute for Science and Technology*, 1:49–57. Mardia, K. V. 和 Jupp, P. E. (2009). *Directional statistics*. 约翰·威立父子出版公司，英国。 Quaintance, J. 和 Gould, H. W. (

2015 *Combinatorial identities for Stirling numbers: the unpublished*). . *notes of HW Gould*

世界科学出版社。Rao, J. S. (1969). *Some Contributions to the Analysis of Circular Data*. 博士论文，印度统计研究所 - 印度加尔各答。Russell, G. S. 和 Levitin, D. J. (1995). Rao 间距检验概率值的扩展表。 *Communications in Statistics-Simulation and Computation*, 24(4):879–888. Sherman, B. (

1950 *The Annals of Mathematical*). *Statistics* 21 3 339 361

一个与样本值间距有关的随机变量。, ( ): – . Smith, P. J. (1995). 从累积量求矩及反向求解这一经典问题的递归公式。 *The American Statistician*, 49(2):217–218. Stuart, A. 和 Ord, K. (2010). *Kendall's advanced theory of statistics, distribution theory*, 第 1 卷。约翰·威立父子出版公司，英国。