文章编号: 1000-5641(2009)05-0127-11

对称性全局统计分析中的定理证明

潘长缘[1], 马海南[1,2], 陈雪平[1], 张应山[1]

(1. 华东师范大学 统计与精算学系, 上海 200241; 2. 浙江工业职业技术学院 基础部,
浙江 绍兴 312000)

摘要: 证明了对称性全局统计分析方法中的几个重要定理, 保证了任意系统函数能够进行正交对称分解, 确保了系统函数方差分解公式成立, 这些是对称性全局统计分析方法的核心基石. 通过考察对称函数在整个系统函数中所起作用的大小, 达到认识系统函数对称性的目的. 实例表明, 将贡献率的Monte-Carlo计算值作为全局分析中对称函数的敏感性度量指标, 可以较好地刻画系统函数的对称性.

关键词: 类对称算符; 正交幂等系统; 对称函数; 全局统计分析

中图分类号: O212. 6 **文献标识码**: A

Proof procedure of some theories in statistical analysis of global symmetry

Pan Changyuan[1], Ma Hainan[1,2], Chen Xueping[1], Zhang Yingshan[1]

(1. Deptment of Statistics and Actuarial Science, East China Normal University, Shanghai 200241, China; 2. Fundamentals Department, Polytechnic Zhejiang Industry College, Shaoxing Zhejiang 312000, China)

Abstract: This research guaranteed orthogonal symmetry demonstration to any system function, ensured the formula of demonstration to system function variance. They are the kernel and foundation stone of statistical analysis of global symmetry. By studying how these symmetry functions work in the whole system function, the symmetry of system function can be understood better. As illustrated by the examples, it showed the symmetry of system function clearly by using the Monte-Carlo calculated value of contribution rate as the global sensitivity index of symmetry function.

Key words: class idempotent function; systems of orthogonal idempotents; symmetry functions; global statistical analysis

收稿日期: 2008-12

基金项目: 国家自然科学基金 (10571045); 教育部高等学校博士点基金(44k55050)

第一作者: 潘长缘, 男, 硕士生, 研究方向为应用统计. E-mail: nicemice@126. com.

0 引 言

近年来, 全局分析在试验设计领域里得到了广泛重视[1−4], 主要原因是该方法可以在整体上考察系统函数的性质, 给工程人员带来许多有意义的参考信息. 在进行试验设计之前对系统函数有一个清晰的理解, 试验设计将变得更加具有针对性, 相应的试验结果也必定会更加有效.

本文重点讨论关于系统函数对称性方面的全局分析. 对称性概念由来已久, 主要有图像的对称和函数的对称两种. 关于对称性理论方面的研究, 国内外已经有许多专著. 例如关于对称图像方面, 唐有祺在1977年 [5]指出: 对称图像虽然千变万化, 但对称操作却只有七类. 其中的原理和性质是研究原子和分子空间分布规律的结构化学之基础.

对称函数一般指在坐标的变换中显示出不变的规律性质, 各个学科关于对称函数的定义不尽相同. 王伯英[6]利用全对称置换群的不可约特征标方法来对多元函数空间进行对称分解, 从而定义出对称函数. 为了寻找出更一般的对称函数概念, 文献[7, 8]利用类对称算符来定义对称函数, 并证明类对称算符虽然有别于特征标, 但却具有类似特征标的性质.

为了研究系统函数的对称性, 本文采用文献[7, 8]中定义的对称函数, 利用系统函数的对称分解技术, 提出一种敏感性度量指标-贡献率（对称函数方差在整个系统函数方差中所占的比率）, 重点考察各对称函数所起作用的大小.

本文由两大部分组成. 第一部分由类对称算符、正交分解、对称函数和全局统计分析等四个小节构成, 集中证明其中涉及的重要定理；第二部分对两个系统函数, 分别给出了对称性全局统计分析方法的应用实例.

1 类对称算符

定义 1 假设 n 阶有限群 $G = (\sigma_1, \cdots, \sigma_n)$ 能写成剖分形式: $G = [\theta_1] \cup \cdots \cup [\theta_k]$, 其中 $[\theta_i]$ 是 G 的包含 θ_i 的一个子集合, $[\theta_i] \cap [\theta_j] = \emptyset$, $\forall\, i \neq j$, 称 $G = [\theta_1] \cup \cdots \cup [\theta_k]$ 为 G 的一个**对称剖分**, 如果其满足如下两个条件.

(1) 若定义 $[\theta_i][\theta_j] = \{\theta\pi : \theta \in [\theta_i],\ \pi \in [\theta_j]\}$, 则有 $[\theta_i][\theta_j] = [\theta_j][\theta_i] \supset [\theta_i\theta_j], \forall i, j;$

(2) 若定义 $[\theta_i]^{-1} = \{\theta^{-1} : \theta \in [\theta_i]\}, i = 1, ..., k$, 则群 G 的剖分满足: $\{[\theta_1], \cdots, [\theta_k]\} = \{[\theta_1]^{-1}, \cdots, [\theta_k]^{-1}\}$.

如果单位元 e 自成一个集合, 这样的对称剖分称为**正规的**, 简称**正规对称剖分**.

定义示性函数: $I_i(\sigma) = \begin{cases} 1, & \sigma \in [\theta_i]; \\ 0, & \sigma \notin [\theta_i]. \end{cases}$ 称形如 $f(\sigma) = c_1 I_1(\sigma) + \cdots + c_k I_k(\sigma)$(其中 c_i 为常数, $i = 1, ..., k$) 的函数为**类函数**, 简记为 $f = C^{\mathrm{T}} I$, 其中 $C = (c_1, \cdots, c_k)^{\mathrm{T}}, I = (I_1, \cdots, I_k)^{\mathrm{T}}$.

定义 2 对 $\forall\, \pi \in G$, 定义运算:$(f + g)(\pi) = f(\pi) + g(\pi)$.

定义 3 对 $\forall\, \pi \in G$, 定义运算:$(f \circ g)(\pi) = \sum\limits_{\sigma \in G} f(\sigma)g(\sigma^{-1}\pi)$, 若函数 f 满足

(1) 幂等性：$f^2(\pi) = (f \circ f)(\pi) = |G|f(\pi)$,

(2) 共轭对称性：$\bar{f}(\pi) = f(\pi^{-1})$, $\forall\, \sigma \in G$,

则称 f 为**对称算符**($|G|$ 表示群 G 的元素个数).

显然, $f(\sigma) \equiv I_1(\sigma) + \cdots + I_k(\sigma)$ 和 $E = |G|I_{\{e\}}$ 一定是对称算符.

定义 4 如果 f 既是定义在有限群剖分 $G = [\theta_1] \cup \cdots \cup [\theta_k]$ 上的类函数, 又是群 G 上的对称算符, 那么称 f 为关于剖分 $G = [\theta_1] \cup \cdots \cup [\theta_k]$ 上的**类对称算符**.

很显然, $f(\sigma) \equiv I_1(\sigma) + \cdots + I_k(\sigma)$ 一定是该剖分上的类对称算符. 另外, 任意对称算符都可以看成是将每个有限群元素单独作为一个剖分情形下的类对称算符. 为了表述清楚, 我们在阐述类对称算符时, 会特别指明其对应的剖分.

类对称算符在代数上具有许多优良性质[9],

2　正交分解

定义 5　若定义在有限群 G 上的非零对称算符 $f_1, \cdots, f_{k'} (k' \leqslant |G|)$ 满足: $f_i \circ f_j = 0,\ i \neq j$, 则称 f_i, f_j 为**正交的**, 并且称 $\{f_1, \cdots, f_{k'}\}$ 为**正交幂等系统**.

定理 1　若 f, g 是有限群剖分 $G = [\theta_1] \cup \cdots \cup [\theta_k]$ 上的相互正交的非零类对称算符, 则 $f + g$ 也是该剖分下非零类对称算符.

证　明　按定义证明即可. 首先, 由于 f, g 是有限群剖分 $G = [\theta_1] \cup \cdots \cup [\theta_k]$ 上的类函数, 其和 $f + g$ 仍然是该剖分上的类函数. 其次, 由定义, 对 $\pi \in G$

$$(f \bar{+} g)(\pi) = (\bar{f} + \bar{g})(\pi) = (f + g)(\pi^{-1}).$$

$$(f+g) \circ (f+g)(\pi) = \sum_{\sigma \in G} (f+g)(\sigma)(f+g)(\sigma^{-1}\pi) = |G| f(\pi) + 0 + 0 + |G| g(\pi) = |G|(f+g).$$

此即证明 $f + g$ 也是对称算符.

而由 f, g 的正交性, $f + g$ 必定非零 (否则 $g = -f$, 这样 $f \circ g = -f \circ f = -|G| f \neq 0$).

由以上证明可得, $f + g$ 也是该剖分下非零类对称算符.

定理1中f, g 相互正交的条件可以改成f, g 正交, 即$f \circ g = 0 \Longrightarrow g \circ f = 0$ 成立. 这是因为若$f \circ g = 0$, 则对$\forall \pi \in G$,

$$0 = \overline{f \circ g(\pi)} = \sum_{\sigma \in G} \overline{f(\sigma) g(\sigma^{-1}\pi)} = \sum_{\pi^{-1}\sigma \in G} g(\pi^{-1}\sigma) f((\pi^{-1}\sigma)^{-1}\pi^{-1}) = g \circ (\pi^{-1}).$$

当π 遍历G 时, π^{-1} 也遍历G, 此即证明$g \circ f = 0$ 成立.

由于前文已提到任意对称算符都可以看成是将每个有限群元素单独作为一个剖分情形下的类对称算符, 因此该定理关于任意对称算符也是成立的.

定理 2　若 $f_1, \cdots, f_{k'}$ 是关于正规对称剖分 $G = [\theta_1] \cup \cdots \cup [\theta_k]$ 上的相互正交的非零类对称算符, 且其向量形式为 $f_1 = C_1^{\mathrm{T}} I, \cdots, f_{k'} = C_{k'}^{\mathrm{T}} I$, 其中 $C_1, \cdots, C_{k'}$ 为 k 维列向量, $I = (I_1, \cdots, I_k)^{\mathrm{T}}$, 则 $C_1, \cdots, C_{k'}$ 必定线性无关.

证　明　反证法证明该定理. 如果 $C_1, \cdots, C_{k'}$ 线性相关, 不妨设 $C_1 = l_2 C_2 + \cdots + l_{k'} C_{k'}$, 其中 $l_2, \cdots, l_{k'}$ 为一组不全为零的常数. 这样就得到 $f_1 = C_1^{\mathrm{T}} I = (l_2 C_2 + \cdots + l_{k'} C_{k'})^{\mathrm{T}} I = l_2 f_2 + \cdots + l_{k'} f_{k'}$. 再不妨假设其中的 $l_{k'} \neq 0$, 由于 $f_{k'}$ 与 $f_2, \cdots, f_{k'-1}$ 正交, 就会有

$$f_1 \circ f_{k'} = (l_2 f_2 + \cdots + l_{k'} f_{k'}) \circ f_{k'} = l_{k'} f_{k'} \circ f_{k'} = l_{k'} |G| f_{k'} \neq 0.$$

这与 $f_1, f_{k'}$ 相互正交的假设相矛盾. 因此可得, $C_1, \cdots, C_{k'}$ 必定线性无关.

定理 3　若 $\{f_1, \cdots, f_{k'}\}$ 为定义在有限群 G 上的正交幂等系统 ($k' \leqslant |G|$), 则对特殊的对称算符 $E = |G| I_{\{e\}}$, 可以进行如下正交分解.

$$E = f_1 + \cdots + f_{k'} + f_{\triangle},\ f_i \circ f_j = 0,\ i \neq j;\ f_{\triangle} \circ f_j = 0,\ \forall\ j, \tag{1}$$

证　明　证明分三步走.

(1) 证明对任何对称算符 f, $E-f$也是对称算符. 首先我们有

$$(\overline{E-f})(\pi)=\bar{E}(\pi)-\bar{f}(\pi)=E(\pi^{-1})-f(\pi^{-1})=(E-f)(\pi^{-1}),\forall\pi\in G.$$

其次有

$$\begin{aligned}(E-f)\circ(E-f)(\pi) &= \sum_{\sigma\in G}(E-f)(\sigma)(E-f)(\sigma^{-1}\pi)\\ &= E\circ E(\pi)-E\circ f(\pi)-f\circ E(\pi)+f\circ f(\pi)\\ &= |G|E(\pi)-E(e)f(e\pi)-f(\pi)E(\pi^{-1}\pi)+|G|f(\pi)\\ &= |G|E(\pi)-|G|f(\pi)-f(\pi)|G|+|G|f(\pi)\\ &= |G|(E-f)(\pi).\end{aligned}$$

综合上述即得 $E-f$ 也是对称算符.

(2) 证明 $E-f$ 与 f 正交.

$$\begin{aligned}(E-f)\circ f(\pi) &= \sum_{\sigma\in G}(E-f)(\sigma)f(\sigma^{-1}\pi)=E\circ f(\pi)-f\circ f(\pi)\\ &= E(e)f(e\pi)-|G|f(\pi)=|G|(f-f)(\pi)=0.\end{aligned}$$

(3) 证明题设中的 $f_\triangle\circ f_j=0$, $\forall\ j$.

前面已经证明 $f_\triangle=E-(f_1+\cdots+f_{k'})$ 与 $(f_1+\cdots+f_{k'})$正交, 即

$$f_\triangle\circ(f_1+\cdots+f_{k'})=0.$$

用 $f_j, j=1,\cdots,k'$ 右乘上式, 得

$$f_\triangle\circ(f_1+\cdots+f_{k'})\circ f_j=f_\triangle\circ(f_1+\cdots+f_{k'})\circ f_j=f_\triangle\circ|G|f_j=|G|f_\triangle\circ f_j=0.$$

此即 $f_\triangle\circ f_j=0$.

综上所述, 定理得证.

定义 6 取 $\{f_1,\cdots,f_{k'}\}$ 为定义在有限群 G 上的正交幂等系统 ($k'\leqslant|G|$). 当 $f_\triangle=0$ 时, $\{f_1,\cdots,f_{k'}\}$ 称为**饱和的正交幂等系统**.

定理 4 若 $f_1,\cdots,f_k$ 为关于正规对称剖分 $G=[\theta_1]\cup\cdots\cup[\theta_k]$ 上的相互正交的非零类对称算符, 则 $\{f_1,\cdots,f_k\}$ 构成一个饱和正交幂等系统.

证　明 此定理可以分两步来证明.

(1) 证明对正规对称剖分 $G=[\theta_1]\cup\cdots\cup[\theta_k]$ 来说, 其正交幂等系统中非零类对称算符的个数最多为 k个. 这点很容易证明.

如果假设 $f_1=C_1^{\mathrm{T}}I,\cdots,f_k=C_k^{\mathrm{T}}I$, 其中 $C_1,\cdots,C_k$ 为 k 维列向量, $I=(I_1,\cdots,I_k)^{\mathrm{T}}$. 由定理 2, $C_1,\cdots,C_k$ 必定线性无关. 而 k 维列向量的最大线性无关的列向量个数最多为 k, 此即证明了正交幂等系统中非零类对称算符的个数最多为 k 个.

(2) 证明 $f_\triangle=|G|I_{\{e\}}-(f_1+\cdots+f_k)=0$.

由于我们选取了正规对称剖分, 所以必有一个剖分就是 $\{e\}$. 这样就得到 $|G|I_{\{e\}}$ 其实也是该正规对称剖分上的类对称算符, 从而 $f_\triangle$ 也是. 如果 $f_\triangle\neq0$, 也就是说其正交幂等系统中非零类对称算符的个数为 $k+1$ 个, 这与第 1 步证明的结论相矛盾. 因此只可能 $f_\triangle=0$.

综上所述, $\{f_1,\cdots,f_k\}$ 构成一个饱和正交幂等系统.

定理 5　若 $f_1,\cdots,f_k$ 是关于正规对称剖分 $G=[\theta_1]\cup\cdots\cup[\theta_k]$ 上的相互正交的非零类对称算符, 则 $f_1,\cdots,f_k$ 除排列顺序外是唯一确定的.

证　明　由定理 4, 正交幂等系统 $\{f_1,\cdots,f_k\}$ 是饱和的.

假设 $f_1=C_1^{\mathrm{T}}I,\cdots,f_k=C_k^{\mathrm{T}}I$, 其中 $C_1,\cdots,C_k$ 为 k 维列向量, $I=(I_1,\cdots,I_k)^{\mathrm{T}}$. 由于 $E=|G|I_{\{e\}}=f_1+\cdots+f_k$ 且 $C_1,\cdots,C_k$ 线性无关, 因此对任意的类对称算符 $g=C^{\mathrm{T}}I$, 其中 C 为 k 维列向量. 由于 $C_1,\cdots,C_k$ 线性无关, 其可作为 k 维列向量空间的一组基, 因此 C 可以表示成 $C_1,\cdots,C_k$ 的线性组合, 即 $C=l_1C_1+\cdots+l_kC_k$, 也即有

$$g=l_1f_1+\cdots+l_kf_k.$$

又由于 g 为对称算符, 因此

$$g^2=g\circ g=(l_1f_1+\cdots+l_kf_k)\circ(l_1f_1+\cdots+l_kf_k)=|G|(l_1^2f_1+\cdots+l_k^2f_k).$$

这样就可以得到等式 $l_1^2=l_1,\cdots,l_k^2=l_k$, 也即 $l_1=0,1;\cdots;l_k=0,1$.

现在假设存在另一组非零类对称算符 $g_1,\cdots,g_k$. 由上面任意类对称算符 g 的讨论, $g_1,\cdots,g_k$ 必定可以表示成 $f_1,\cdots,f_k$ 的线性组合, 且系数只为 0 或 1. 不妨假设

$$g_1=l_{11}f_1+\cdots+l_{1k}f_k,\cdots,g_k=l_{k1}f_1+\cdots+l_{kk}f_k.$$

为保证 $g_1,\cdots,g_k$ 是对称算符, $l_{11},\cdots,l_{1k};\cdots;l_{k1},\cdots,l_{kk}$ 只能取 0 或 1. 不妨假设 $g_i,g_j,1\leqslant i,j\leqslant k,i\neq j$ 中的第 $s(1\leqslant s\leqslant k)$ 项系数均取为 1, 即 l_{is},l_{js} 都取 1. 则

$$g_i\circ g_j=(l_{i1}f_1+\cdots+l_{ik}f_k)\circ(l_{j1}f_1+\cdots+l_{jk}f_k)=|G|(l_{i1}l_{j1}f_1+\cdots+f_s+\cdots+l_{ik}l_{jk}f_k).$$

由于 $C_1,\cdots,C_k$ 是线性无关的, 因此 $f_1,\cdots,f_k$ 也不能相互线性表出, $l_{i1}l_{j1}f_1+\cdots+f_s+\cdots+l_{ik}l_{jk}f_k$ 中因为存在 f_s 非零所以肯定非零, 这与 g_i,g_j 相互正交矛盾.

由以上论述, 就得到了不存在 $g_i,g_j,1\leqslant i,j\leqslant k,i\neq j$ 中的第 $s(1\leqslant s\leqslant k)$ 项系数均取为 1. 换句话说, 非零类对称算符 $g_1,\cdots,g_k$ 中的每一个中只有一个系数是 1, 其余全取为 0. 这就表明 $g_1,\cdots,g_k$ 与 $f_1,\cdots,f_k$ 就是同一组类对称算符 (排列顺序可以不一致).

3　对称函数

定义 7　设 $F(x_1,\cdots,x_m)$ 是任一 m 元函数, 有限群 G 为 S_m 的一个子群, f 是关于群 G 的某个正规对称剖分上的一个类对称算符, 则 $F(x_1,\cdots,x_m)$ 关于 f 对称是指满足下式

$$F(x_1,\cdots,x_m)=\frac{1}{|G|}\sum_{\sigma\in G}f(\sigma)F(x_{\sigma(1)},\cdots,x_{\sigma(m)}).\tag{2}$$

定义 8　对称函数F_1,F_2的内积定义如下

$$\langle F_1,F_2\rangle=\int_{R^m}F_1\bar{F}_2p(x_1)\cdots p(x_m)\mathrm{d}x_1\cdots\mathrm{d}x_m.\tag{3}$$

且任意对称函数F_1,F_2,F满足:

(1) $\langle F,F\rangle\geqslant 0$; (2) $\overline{\langle F_1,F_2\rangle}=\langle F_2,F_1\rangle$; (3) $\langle\alpha F_1+\beta F_2,F\rangle=\alpha\langle F_1,F\rangle+\beta\langle F_2,F\rangle$.

如果对称函数 F_1, F_2 内积为零, 我们称它们**相互正交**.

4 全局统计分析

定义 9 假设随机变量 $x_1, \cdots, x_m$ 独立同分布, 其密度函数为 $p_t = p(x_t)$, $t = 1, \cdots, m$. 定义多维空间 $D = \prod_{t=1}^{m}[c, d]$ 上的平方可积函数 $H(x_1, \cdots, x_m)$ 的**期望**为

$$E(H(x_1, \cdots, x_m)) = \int_D H(x_1, \cdots, x_m)p(x_1)\cdots p(x_m)\mathrm{d}x_1\cdots\mathrm{d}x_m. \tag{4}$$

简记为

$$E(H) = \int_D Hp_1\cdots p_m\mathrm{d}x. \tag{5}$$

定义可测函数 H 的**方差**为

$$D(H) = \int_D [H - E(H)][H - \bar{E}(H)]p_1\cdots p_m\mathrm{d}x. \tag{6}$$

定理 6 任给一个 m 元多元函数 $H(x_1, \cdots, x_m)$ 和任意的关于有限群 $G \subset S_m$ 的正规对称剖分 $G = [\theta_1] \cup \cdots \cup [\theta_k]$ 上的类对称算符构成的正交幂等系统 $\{f_1, \cdots, f_{k'}\}, k' \leqslant k$, 并且 $f(\sigma) \equiv I_1(\sigma) + \cdots + I_k(\sigma)$必为其中某个类对称算符, 定义

$$\begin{aligned} H_j(x_1, \cdots, x_m) &= \frac{1}{|G|}\sum_{\sigma\in G} f_j(\sigma)H(x_{\sigma(1)}, \cdots, x_{\sigma(m)}),\ j = 1, \cdots, k^{'}. \\ H_\triangle(x_1, \cdots, x_m) &= \frac{1}{|G|}\sum_{\sigma\in G} f_\triangle(\sigma)H(x_{\sigma(1)}, \cdots, x_{\sigma(m)}),\ f_\triangle = E - \sum_{j=1}^{k^{'}} f_j. \end{aligned}$$

则内积 $\langle H_i, H_j\rangle = 0,\ \forall\ i \neq j,\ i, j = 1, \cdots, k^{'}, \triangle$, 且对应于分解式

$$E = |G|I_{\{e\}} = f_1 + \cdots + f_{k'} + f_\triangle,\ f_i \circ f_j = 0,\ i \neq j;\ f_\triangle \circ f_j = 0,\ \forall\ j,$$

有

$$H(x_1, \cdots, x_m) = H_1(x_1, \cdots, x_m) + \cdots + H_{k'}(x_1, \cdots, x_m) + H_\triangle(x_1, \cdots, x_m). \tag{7}$$

证 明 先证明等式成立, 再来证明各项之间相互正交.

由 $E = |G|I_{\{e\}} = f_1 + \cdots + f_{k'} + f_\triangle$ 代入可知

$$\begin{aligned} H(x_1, \cdots, x_m) &= \frac{1}{|G|}\sum_{\sigma\in G} E(\sigma)H(x_{\sigma(1)}, \cdots, x_{\sigma(m)}) \\ &= \frac{1}{|G|}\sum_{\sigma\in G} (f_1 + \cdots + f_{k'} + f_\triangle)(\sigma)H(x_{\sigma(1)}, \cdots, x_{\sigma(m)}) \\ &= H_1(x_1, \cdots, x_m) + \cdots + H_{k'}(x_1, \cdots, x_m) + H_\triangle(x_1, \cdots, x_m). \end{aligned}$$

即分解式是成立的. 下面证明各对称函数之间相互正交. 我们知道存在类对称算符 $f(\sigma) \equiv I_1(\sigma) + \cdots + I_k(\sigma)$. 不妨设 f_1就是这个类对称算符, 由期望定义

$$E(H_1(x_1, \cdots, x_m)) = \frac{1}{|G|}\sum_{\sigma\in G} f_1(\sigma)E(H(x_{\sigma(1)}, \cdots, x_{\sigma(m)})) = E(H(x_1, \cdots, x_m)).$$

上面等式中推导中用到了 $E(H(x_{\sigma(1)},\cdots,x_{\sigma(m)}))=E(H(x_1,\cdots,x_m))$ 的结论. 而对其他的对称函数 $H_j(x_1,\cdots,x_m), j=2,\cdots,k',\triangle$,

$$
\begin{aligned}
E(H_j(x_1,\cdots,x_m)) &= \frac{1}{|G|}\sum_{\sigma\in G} f_j(\sigma)E(H(x_{\sigma(1)},\cdots,x_{\sigma(m)}))\\
&= (\frac{1}{|G|}\sum_{\sigma\in G} f_j(\sigma)f_1(\sigma^{-1}\pi))E(H(x_1,\cdots,x_m))=0.
\end{aligned}
$$

上面等式中用到了 f_1, f_j 相互正交的假设. 下面将来证明对 $H_i, H_j, i\neq j$, 有 $<H_i,H_j>=0$.

$$
\begin{aligned}
<H_i,H_j> &= E(H_i\bar{H}_j)\\
&= \frac{1}{|G|}\sum_{\sigma\in G} f_i(\sigma)\frac{1}{|G|}\sum_{\pi\in G}\bar{f}_j(\pi)E(H(x_{\sigma(1)},\cdots,x_{\sigma(m)}))E(\bar{H}(x_{\pi(1)},\cdots,x_{\pi(m)}))\\
&= \frac{1}{|G|}\sum_{\sigma\in G} f_i(\sigma)\frac{1}{|G|}\sum_{\pi\in G}\bar{f}_j(\pi)E(H(z_1,\cdots,z_m))E(\bar{H}(z_{\sigma^{-1}\pi(1)},\cdots,z_{\sigma^{-1}\pi(m)}))\\
&= \frac{1}{|G|}\sum_{\sigma\in G} f_i(\sigma)\frac{1}{|G|}\sum_{u\in G}\bar{f}_j(\sigma u)E(H(z_1,\cdots,z_m))E(\bar{H}(z_{u(1)},\cdots,z_{u(m)}))\\
&= \frac{1}{|G|}\sum_{u\in G}(\frac{1}{|G|}\sum_{\sigma\in G} f_i(\sigma)\bar{f}_j(\sigma u))E(H(z_1,\cdots,z_m))E(\bar{H}(z_{u(1)},\cdots,z_{u(m)}))
\end{aligned}
$$

其中上式中做了两次变换：

(1) 令 $z_j=x_{\sigma(j)}, j=1,\cdots,m$, 从而 $z_{\sigma^{-1}i}=x_i, i=1,\cdots,m$, 进而 $x_{\pi(k)}=z_{\sigma^{-1}(\pi(k))}=z_{(\sigma^{-1}\pi)(k)}, k=1,\cdots,m$;

(2) $u=\sigma^{-1}\pi, \pi=\sigma u$.

从而如果令 $v=(\sigma u)^{-1}$, 那么 $v^{-1}=\sigma u, \sigma=v^{-1}u^{-1}$, 再由 $\bar{f}_j(\sigma u)=f_j((\sigma u)^{-1})$,

$$\frac{1}{|G|}\sum_{\sigma\in G} f_i(\sigma)\bar{f}_j(\sigma u)=\frac{1}{|G|}\sum_{v\in G} f_j(v)f_i(v^{-1}u^{-1})=0.$$

因此就有 $<H_i,H_j>=0$ 成立. 这样就证明了这些对称函数是相互正交的.

如果 $f_\triangle=E-\sum\limits_{j=1}^{k'} f_j=0$, 此时 $\{f_1,\cdots,f_{k'}\}$ 是饱和正交幂等系统, 那么

$$H(x_1,\cdots,x_m)=H_1(x_1,\cdots,x_m)+\cdots+H_{k'}(x_1,\cdots,x_m).$$

定理 6 表明任意系统函数 $H(x_1,\cdots,x_m)$ 都可以分解成若干个对称函数之和, 这些对称函数与类对称算符的选取有关.

定理 7　设 G 为置换群 S_m 的子群, 随机变量 $x_1,\cdots,x_m$ 独立同分布, 其密度函数为 $p_t=p(x_t)$, $t=1,\cdots,m$. $H(x_1,\cdots,x_m)$ 为多维空间 $D=\prod\limits_{t=1}^{m}[c,d]$ 上的平方可积函数, $\{f_1,\cdots,f_{k'}\}$ 为 G 的某正规对称剖分上的类对称算符, 构成一个正交幂等系统, 并且 $f(\sigma)\equiv I_1(\sigma)+\cdots+I_k(\sigma)$必为其中某个类对称算符. 对应于分解式:

$$E=|G|I_{\{e\}}=f_1+\cdots+f_{k'}+f_\triangle,\ f_i\circ f_j=0,\ i\neq j;\ f_\triangle\circ f_j=0,\ \forall\ j.$$

如下定义的对称函数

$$
\begin{aligned}
H_j(x_1,\cdots,x_m) &= \frac{1}{|G|}\sum_{\sigma\in G} f_j(\sigma)H(x_{\sigma(1)},\cdots,x_{\sigma(m)}),\ j=1,\cdots,k^{'}.\\
H_{\triangle}(x_1,\cdots,x_m) &= \frac{1}{|G|}\sum_{\sigma\in G} f_{\triangle}(\sigma)H(x_{\sigma(1)},\cdots,x_{\sigma(m)}),\ f_{\triangle}=E-\sum_{j=1}^{k^{'}} f_j.
\end{aligned}
$$

满足 $H=H_1+\cdots+H_{k'}+H_{\triangle}$. 且有如下方差分解公式成立.

$$
D(H)=D(H_1)+\cdots+D(H_{k'})+D(H_{\triangle}). \tag{8}
$$

证 明 令 $h=H-E(H), h_j=H_j-E(H_j), j=1,\cdots,k^{'},\triangle$, 由定理 6, 任意函数可以分解成若干个相互正交的对称函数的和. 因此 $H=H_1+\cdots+H_{k'}+H_{\triangle}$, 而 $E(H)=E(H_1)+\cdots+E(H_{k'})+E(H_{\triangle})$ 恒成立, 因此 $h=h_1+\cdots+h_{k'}+h_{\triangle}$ 也成立. 下面我们将证明 (8) 式成立.

由于存在类对称算符 $f(\sigma)\equiv I_1(\sigma)+\cdots+I_k(\sigma)$. 不妨假设 f_1 就是这个类对称算符, 由定理 6 的证明过程,

$$
\begin{aligned}
E(H_1(x_1,\cdots,x_m)) &= E(H(x_1,\cdots,x_m));\\
E(H_j(x_1,\cdots,x_m)) &= 0, j=2,\cdots,k^{'},\triangle.
\end{aligned}
$$

并且对常函数 $C=E(H(x_1,\cdots,x_m))$, 其关于对称算符 f_1 是对称的, 因此对应对称算符分解式也必有分解式 $C=C+0+\cdots+0+0$ 成立.

从而知道 $h_1,\cdots,h_{k'},h_{\triangle}$ 也是相互正交的对称函数. 由分解式 $h=h_1+\cdots+h_{k'}+h_{\triangle}$,

$$
\begin{aligned}
D(H) &= \int_D (H-E(H))(\overline{H-E(H)})p_1\cdots p_m \mathrm{d}x\\
&= \int_D h\bar{h}p_1\cdots p_m \mathrm{d}x\\
&= \int_D (h_1+\cdots+h_{k'}+h_{\triangle})(\overline{h_1+\cdots+h_{k'}+h_{\triangle}})p_1\cdots p_m \mathrm{d}x\\
&= D(H_1)+\cdots+D(H_{k'})+D(H_{\triangle}).
\end{aligned}
$$

在系统函数确定以及分布确定的情况下, (8) 式两边的各个项均是确定的值.

定义 10 定义 $S_j=\frac{D(H_j)}{D(H)}, j=1,\cdots,k'; S_{\triangle}=\frac{D(H_{\triangle})}{D(H)}$ 为各个对称函数对整个系统函数方差所做贡献的比率 (以后简称**贡献率**), 它是各个对称函数关于整个系统函数的敏感性指标. $S_1,\cdots,S_{k'},S_{\triangle}$ 是唯一确定的, 且有

$$
1=S_1+\cdots+S_{k'}+S_{\triangle}. \tag{9}
$$

全局分析中, 这些敏感性指标对认识系统函数的对称性具有重要的意义.

然而由于函数 $H(x_1,\cdots,x_m)$ 的形式通常都是比较复杂的, 虽然其在定义域上具有平方可积的性质, 但是该积分的计算相当复杂, 因此贡献率的计算也就变得非常困难, 目前国际上较多采用 Monte-Carlo 方法. 我们将利用统计手段估计各个对称函数贡献率的方法称为**对称性全局统计分析**.

5　应用举例

下面以6阶置换群$G=\{e,(23),(12),(123),(132),(13)\}=S_3$为例, 我们将群元素依次简记为$1,2,3,4,5,6$. 其对应的乘法表为:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 6 & 3 & 4 \\ 3 & 4 & 1 & 2 & 6 & 5 \\ 4 & 3 & 6 & 5 & 1 & 2 \\ 5 & 6 & 2 & 1 & 4 & 3 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{pmatrix}.$$

如果取群S_3的正规对称剖分为:$[\theta_1]=\{1\};[\theta_2]=\{2\};[\theta_3]=\{3,6\};[\theta_4]=\{4,5\}$. 找到相互正交的一组类对称算符为: $f_1(\sigma)=I_1(\sigma)+I_2(\sigma)+I_3(\sigma)+I_4(\sigma)$; $f_2(\sigma)=I_1(\sigma)-I_2(\sigma)-I_3(\sigma)+I_4(\sigma)$; $f_3(\sigma)=2I_1(\sigma)+2I_2(\sigma)-I_3(\sigma)-I_4(\sigma)$; $f_4(\sigma)=2I_1(\sigma)-2I_2(\sigma)+I_3(\sigma)-I_4(\sigma)$.

这四个相互正交的类对称算符构成了一个饱和正交幂等系统.

例 1　考虑三元二次函数$H(x_1,x_2,x_3)=4+2x_1+3x_2+4x_3+3x_1^2+2x_2^2+x_3^2+x_1x_2+3x_1x_3+2x_2x_3$, x_1,x_2,x_3服从$[0,1]$上的均匀分布.

运用上面四个类对称算符可以分别构造四个对称函数：

$$\begin{aligned} H_1(x_1,x_2,x_3) &= \frac{1}{6}[H(x_1,x_2,x_3)+H(x_1,x_3,x_2)+H(x_2,x_1,x_3) \\ &\quad +H(x_2,x_3,x_1)+H(x_3,x_1,x_2)+H(x_3,x_2,x_1)] \\ &= 4+3(x_1+x_2+x_3)+2(x_1^2+x_2^2+x_3^2)+2(x_1x_2+x_1x_3+x_2x_3); \\ H_2(x_1,x_2,x_3) &= \frac{1}{6}[H(x_1,x_2,x_3)-H(x_1,x_3,x_2)-H(x_2,x_1,x_3) \\ &\quad +H(x_2,x_3,x_1)+H(x_3,x_1,x_2)-H(x_3,x_2,x_1)] \\ &= 0; \\ H_3(x_1,x_2,x_3) &= \frac{1}{6}[2H(x_1,x_2,x_3)+2H(x_1,x_3,x_2)-H(x_2,x_1,x_3) \\ &\quad -H(x_2,x_3,x_1)-H(x_3,x_1,x_2)-H(x_3,x_2,x_1)] \\ &= \frac{1}{2}[(x_2+x_3-2x_1)-(x_2^2+x_3^2-2x_1^2)]; \\ H_4(x_1,x_2,x_3) &= \frac{1}{6}[2H(x_1,x_2,x_3)-2H(x_1,x_3,x_2)+H(x_2,x_1,x_3) \\ &\quad -H(x_2,x_3,x_1)-H(x_3,x_1,x_2)+H(x_3,x_2,x_1)] \\ &= \frac{1}{2}[-(x_2-x_3)+(x_2^2-x_3^2)-2(x_1x_2-x_1x_3)]. \end{aligned}$$

由于函数形式比较简单, 可以用定义的方法分别求出上述各函数的真实方差, 从而可以求出各对称函数的贡献率；同时利用Monte-Carlo(简记作M-C)得到各函数的方差(用D表示)及其贡献率(用S表示)的估计值, 见表1.

表 1 例1的对称性全局统计分析表

Tab. 1 Statistical analysis of global symmetry to Example 1.

度量指标	$D(S)$	$D_1(S_1)$	$D_2(S_2)$	$D_3(S_3)$	$D_4(S_4)$
真实值	12.4667(100%)	12.4000(99.47%)	0(0%)	0.0083(0.07%)	0.0583(0.47%)
M-C方法(1000点)	12.3219(100%)	12.3211(99.99%)	0(0%)	0.0082(0.07%)	0.0566(0.46%)
M-C方法(64000点)	12.5076(100%)	12.4434(99.49%)	0(0%)	0.0083(0.07%)	0.0580(0.46%)
M-C方法(1000000点)	12.4441(100%)	12.3800(99.48%)	0(0%)	0.0083(0.07%)	0.0583(0.47%)

例 2 考虑三元复杂函数$H(x_1,x_2,x_3)=\exp\{x_1^2+x_2\cos x_3\}$, x_1,x_2,x_3服从$[0,1]$上的均匀分布. 运用上面四个类对称算符可以分别构造四个对称函数：

$$\begin{aligned}
H_1(x_1,x_2,x_3) &= \frac{1}{6}[H(x_1,x_2,x_3)+H(x_1,x_3,x_2)+H(x_2,x_1,x_3)\\
&\quad +H(x_2,x_3,x_1)+H(x_3,x_1,x_2)+H(x_3,x_2,x_1)]\\
&= \frac{1}{6}(\exp\{x_1^2+x_2\cos x_3\}+\exp\{x_1^2+x_3\cos x_2\}+\exp\{x_2^2+x_1\cos x_3\} \qquad (10)\\
&\quad +\exp\{x_2^2+x_3\cos x_1\}+\exp\{x_3^2+x_1\cos x_2\}+\exp\{x_3^2+x_2\cos x_1\}); \qquad (11)\\
H_2(x_1,x_2,x_3) &= \frac{1}{6}[H(x_1,x_2,x_3)-H(x_1,x_3,x_2)-H(x_2,x_1,x_3)\\
&\quad +H(x_2,x_3,x_1)+H(x_3,x_1,x_2)-H(x_3,x_2,x_1)]\\
&= \frac{1}{6}(\exp\{x_1^2+x_2\cos x_3\}-\exp\{x_1^2+x_3\cos x_2\}-\exp\{x_2^2+x_1\cos x_3\}\\
&\quad +\exp\{x_2^2+x_3\cos x_1\}+\exp\{x_3^2+x_1\cos x_2\}-\exp\{x_3^2+x_2\cos x_1\});\\
H_3(x_1,x_2,x_3) &= \frac{1}{6}[2H(x_1,x_2,x_3)+2H(x_1,x_3,x_2)-H(x_2,x_1,x_3)\\
&\quad -H(x_2,x_3,x_1)-H(x_3,x_1,x_2)-H(x_3,x_2,x_1)]\\
&= \frac{1}{6}(2\exp\{x_1^2+x_2\cos x_3\}+2\exp\{x_1^2+x_3\cos x_2\}-\exp\{x_2^2+x_1\cos x_3\}\\
&\quad -\exp\{x_2^2+x_3\cos x_1\}-\exp\{x_3^2+x_1\cos x_2\}-\exp\{x_3^2+x_2\cos x_1\});\\
H_4(x_1,x_2,x_3) &= \frac{1}{6}[2H(x_1,x_2,x_3)-2H(x_1,x_3,x_2)+H(x_2,x_1,x_3)\\
&\quad -H(x_2,x_3,x_1)-H(x_3,x_1,x_2)+H(x_3,x_2,x_1)\\
&= \frac{1}{6}(2\exp\{x_1^2+x_2\cos x_3\}-2\exp\{x_1^2+x_3\cos x_2\}+\exp\{x_2^2+x_1\cos x_3\} \qquad (12)\\
&\quad -\exp\{x_2^2+x_3\cos x_1\}-\exp\{x_3^2+x_1\cos x_2\}+\exp\{x_3^2+x_2\cos x_1\}).
\end{aligned}$$

由于系统函数比较复杂, 要求出其真实方差比较困难, 表(2)给出利用M-C方法得到的各函数方差及其贡献率的估计值.

表 2 例2的对称性全局统计分析表

Tab. 2 Statistical analysis of global symmetry to Example 2.

度量指标	$D(S)$	$D_1(S_1)$	$D_2(S_2)$	$D_3(S_3)$	$D_4(S_4)$
真实值	$--(--\%)$	$--(--\%)$	$--(--\%)$	$--(--\%)$	$--(--\%)$
M-C方法(1000点)	0.9682(100%)	0.4148(42.85%)	0.0005(0.06%)	0.2439(25.19%)	0.2815(29.08%)
M-C方法(64000点)	0.9511(100%)	0.4220(44.37%)	0.0005(0.06%)	0.2374(24.96%)	0.3001(31.55%)
M-C方法(1000000点)	0.9620(100%)	0.4242(44.10%)	0.0005(0.06%)	0.2396(24.91%)	0.3012(31.31%)

结论：对比表1中贡献率的真实值和三个不同模拟点数的M-C计算值, 可以发现对称性全局统计分析方法能较好地估计贡献率, 并且随着模拟点数的增加而变得更加精确；表2中虽然贡

献率的真实值并不知道, 但是三个不同模拟点数的Monte-Carlo计算值比较稳定, 因此可以将贡献率的M-C计算值作为全局分析中对称函数的敏感性度量指标. 另外, 从以上两个表我们不难发现：系统函数方差确实等于分解得到的各个对称函数方差之和, 这正是本文定理 7所要求的. 因此, 选择文献[7, 8]中对称函数的定义来研究对称性也就合情合理.

[参　考　文　献]

[1] SOBOL I M, TARANTOLA S, GATELLI D, et al. Estimating the approximation error when fixing unessential facters in global sensitivity analysis[J]. Reliability Engineering and System Safety, 2007, 92:957-960.

[2] SOBOL I M. Theorems and examples on high dimensional model representation[J]. Reliability Engineering and System Safety, 2003, 79: 187-193.

[3] SOBOL I M, LEVITAN Yu L. On the use of variance reducing multipliers in Monte-Carlo computations of a global sensitivity index[J]. Computer Physics Communications, 1999, 117: 52-61.

[4] SOBOL I M. Global sensitivity indices for nonlinear mathematical models and their Monte-Carlo estimates[J]. Mathematics and Computers in Simulation, 2001, 55: 271-280.

[5] 唐有祺. 对称性原理[M]. 北京: 科学出版社, 1977.
TANG Y Q. Principle of Symmetry[M]. Beijing: Science Press, 1977.

[6] 王伯英. 多重线性代数[M]. 北京: 北京师范出版社, 1985.
WANG B Y. Multi-linear Algebra[M]. Beijing: Beijing Normal University Press, 1985.

[7] 张应山. 多边矩阵理论[M]. 北京: 中国统计出版社, 1993.
ZHANG Y S. Theory of Multilateral Matrix[M]. Beijing: China Statistics Press, 1993.

[8] 陈雪平, 潘长缘, 张应山. 多元函数空间的对称分解[J]. 数学的实践与认识, 39(2): 167-173.
CHEN X P, PAN C Y, ZHANG Y S. Partitioning the Multivariate Function Space into Symmetrical Classes[J]. Mathematics in Practice and Theory, 39(2): 167-173.

[9] 赵建立. 群的幂等正交类系统及应用[D]. 上海：华东师范大学, 2007.
ZHAO J L. Idempotent orthogonal class system of a group and its applications[D]. Shanghai：East China Normal University, 2007.

[10] ZHANG Y S , PANG S Q , JIAO Z M , ZHAO W Z. Group partition and systems of orthogonal idempotents[J]. Linear Algebra and its Application, 1998, 278: 249-262.

[11] 潘长缘, 陈雪平, 张应山. 正交幂等系统的构造[J]. 华东师范大学学报(自然科学版), 2008, 141: 51-58.
PAN C Y, CHEN X P, ZHANG Y S. Construct Systems of Orthogonal Idempotents[J]. Journal of East China Normal University(Natural Science), 2008, 141: 51-58.

(上接第 126 页)

[7] YEO K L, VALDEZ E A. Claim dependence with common effects in credibility models [J]. Insurance: Mathematics and Economics, 2006 (**38**): 609-629.

[8] WEN L, WU X, ZHOU X. The credibility premiums for models with dependence induced by common effects [J], Insurance: Mathematics and Economics, 2009 (44): 19-25.

[9] NIKOLAI K, DELHI P A. Random sums of exchangeable variables and actuarial applications [J]. Insurance: Mathematics and Economics, 2008 (42): 147-153.

[10] BÜHLMANN H. Experience rating and credibility [J]. Astin Bulletin, 1967, 4(1): 199-207.

[11] JEWELL W S. Credible means are exact Bayesian for exponential families [J]. Astin Bulletin 1974, 8(2): 77-90.

[12] RADHAKRISHNA R, HELGE T. Linear Models [M]. New York: Springer, 1995.

[13] BÜHLMANN H, GISLER A. A Course in Credibility Theory and its Applications [M]. Netherlands: Springer, 2005.

[14] BÜHLMANN H, STRAUB E. Glaubwüdigkeit für Schadensäze [J]. Bulletin of the Swiss Association of Actuaries, 1970, 70(1): 111-33.