基于傅里叶分析提取的特征，使用多元方差分析进行手写体证据的贝叶斯评估

Lampis Tzai洛桑大学刑事司法学院，瑞士；雅典经济与商业大学统计系，希腊 电子邮件：lampis.tzai@unil.ch

Ioannis Ntzoufras雅典经济与商业大学统计系，希腊 电子邮件：ntzoufras@aueb.gr

Silvia Bozza威尼斯卡福斯卡里大学经济系，意大利；洛桑大学刑事司法学院，瑞士 电子邮件：silvia.bozza@unive.it

arXiv:2601.07534v1 [stat.AP] 12 Jan 2026

摘要

本文提出一种新颖的统计方法，旨在通过贝叶斯建模，在笔迹检验中识别有效且有用的模式。研究从13位法语母语书写者中选取一组字符样本，并通过傅里叶分析实现精确的环形重建。手写字符的轮廓形状可由前四对傅里叶系数及表面大小描述。本文针对这些手写特征考察了六种贝叶斯模型。这些模型源自两种似然结构：(a) 多元正态模型；(b) 能够考虑字符层级变异的多元方差分析模型。对于每种似然，研究考察了三种不同的先验形式，由此得到不同的贝叶斯模型：(i) 共轭正态-逆威沙特先验；(ii) Bozza 等人提出的层次正态-逆威沙特先验Bozza 等人(2008年)；(iii) 正态-对数正态-LKJ先验设定。层次先验形式尤为重要，因为它们能够纳入书写者之间的变异性，这是区分不同书写者的关键因素。这些方法无法以闭式表达式计算边际似然，因此采用桥接抽样对其进行估计。通过估计贝叶斯因子来比较所提出模型的表现，并评估其判别效能。采用正态-对数正态-LKJ先验的贝叶斯多元方差分析，在判别能力和模型拟合方面总体表现更优。最后，对先验分布参数的设定进行了敏感性分析。

关键词：1. 笔迹证据 2. 法庭科学 3. 傅里叶分析 4. 贝叶斯层次建模 5. 共轭分析 6. 桥接抽样 7. 贝叶斯因子 8. 敏感性

1 引言

笔迹检验是法庭科学的一个领域，文件鉴定专家经常受托为法律程序中的相关人员提供意见；这些人员面对的是来源存在争议或不明的手写文件。笔迹证据的评估仍是法庭科学中的一个未决问题。尽管已有大量研究致力于支持个体识别过程，但该过程在很大程度上仍取决于专家对笔迹特征进行定性和主观评估。即便许多审慎的司法文献已指出，仅依据所谓证据材料、且不参考背景信息而得出的当前类别性结论（确定性陈述），在逻辑上也无法成立（Aitken 等人 2021年，Izenman 2020年）。

近期有多项研究探索了检查手写文件、评估书写者认定问题相关证据的新方法。Johnson 和 Ommen(2022年) 开发了一种方法，利用机器学习和统计技术量化手写文件之间的相似性。他们采用基于分数的似然比（SLRs）评估两种情形下证据的价值：共同来源和特定来源。Wydra 和 Matuszewski (2022年) 提出了一种运用似然比方法评估笔迹证据的方法。他们通过杰卡德指数测量笔迹样本之间的相似性；杰卡德指数是一种用于衡量两个特征集合重叠程度的统计量。Crawford 等人 (2023) 提出了一种统计模型，用于在检验文件的书写者属于封闭书写者集合时计算作者身份概率。他们利用基于旋转的特征从手写文件中提取测量值，并采用贝叶斯层级模型估计作者身份的后验预测概率。

本文研究并扩展了 Bozza 等人(2008年) 提出的贝叶斯概率方法。该文使用了 Marquis 等人 (2005年) 提出的临时图像分析流程所获得的特定手写特征。该技术基于傅里叶分析，能够精确重建字符环的轮廓形状。按照这一方法，每个字符环都可以用傅里叶系数描述，这些系数可用于表征形状复杂度及其他几何属性。Marquis 等人 (2006年) 的初步研究表明，这种特征表征具有良好的区分能力。随后，通过贝叶斯因子（简称 BF）评估证据的价值。贝叶斯因子可解释为一种衡量指标，用于衡量证据对法律诉讼程序中对立双方提出的竞争假设的支持力度。不同法庭科学学科的操作标准和建议广泛支持将贝叶斯因子用作评估法庭科学证据证明价值的指标（例如，欧洲法庭科学研究所网络的指南（Willis 等人 2015年））。Bozza 等人 (2022年) 从操作视角对贝叶斯因子在法庭科学决策分析中的应用作了综述。

特别是，Bozza 等人（2008年）提出了一种分层双层随机效应模型，其优点是能够同时对书写者内变异和书写者间变异进行建模。然而，正如作者本人所指出的，这一方案的主要局限在于，它没有对每种字符类型各自的变异进行建模。因此，该模型要么不区分不同字符类型，因而会丢失一个重要的信息来源；要么针对每种字符类型分别应用。这种方法可能会带来问题，尤其是在尝 H_1时，另一字符则支持 H_2）。例如，与一种字符（如a）相关的测量值可能支持这样的假设：所比较的材料出自同一书写者；而另一种字符（如d）的测量值则可能支持相反的假设，即这些材料出自不同的书写者。

为减轻这一模型限制的影响，本文提出了一种贝叶斯多变量方差分析（MANOVA）模型，并采用环形字符作为预测变量。环形字符指标被转换为虚拟变量（角点表示法，见附录C），从而能够对描述不同类型手写字符的变量进行联合建模，同时考虑字符间变异、每种字符对应的书写者间变异以及书写者内变异。

作者指出的另一个尚待解决的问题，是贝叶斯因子对先验分布设定的敏感性，以及对用于描述书写者内手写变异的先验分布进行设定时所产生的敏感性。

本文结构如下。第2节介绍可用数据，尤其是基于傅里叶的图像分析程序。第3节介绍关注的情境、法庭科学证据评估的概念、贝叶斯因子的定义，特别是其在法庭科学中的应用。第4.1 节和第4.2节分别阐述所比较的概率模型。本文考察了三种先验设定方案，首先介绍共轭方法（见第4.1.1 节和第4.2.1节）。这种方法的主要局限在于无法分别描述书写者内变异和书写者间变异。但其关键优势是，贝叶斯因子计算所需的边际分布可以通过解析方法获得。本文还提出了两种层次方法，可分别描述这两种变异来源；见第4.1.2节、第4.1.3 节和第4.2.2节、第4.2.3节。由于与共轭方法不同，这些方法无法通过解析方式获得边际似然，因此采用桥接抽样估计所需的边际似然。第5节呈现实验结果，重点讨论模型比较（第5.1节），以及评估各模型在支持正确假设方面的效能（见第5.2节）。最后，第6节针对先验分布参数的设定进行了敏感性分析。分析还考察了逆威沙特自由度的先验参数以及 η LKJ分布参数的设定敏感性；这两个参数均用于描述书写者内手写变异。第7 节总结全文，阐述主要结论，讨论本文的主要发现，并指出未来研究方向。

2 手写体数据的工程与结构

本研究所使用的数据来自一项先前研究（Marquis 等人 2006年）收集的13名书写者样本。这些书写者来自洛桑大学刑事司法学院（瑞士）的法语母语书写者群体，因其习惯闭合环形字符而被选中。字符 a、d、o 和 q 的轮廓形状根据 Schmittbuhl 等人 （1998年）提出的图像分析程序进行处理。该程序最初用于分析某些灵长类动物骨骼（即下颌骨）梨状孔的多态性，后经 Marquis 等人 （2005年）改用于笔迹检验。

这种方法存在实际局限性，因为它无法直接应用于不闭合或未完整形成环形字符的书写者。不过，插值和滤波等技术可以帮助重建开放或不完整的环形字符（Thiel 等人 2011）。然而，还需要进一步研究以确定这些技术的有效性，并优化其应用。另一个问题是，所提出的方法依赖于语言，因为不同语言中的环形字符存在显著差异——例如，罗马字母、中文或斯拉夫字母书写系统之间便是如此。最后，该数据集中环形字符的识别和标注由人工根据笔迹专家的判断完成。因此，数据中可能存在一些评估者引入的错误；不过，鉴于评估者的专业水平，这些错误被认为可以忽略不计；详情参见 Marquis 等人 （2006年）。

2.1 傅里叶系数作为环形字符特征

图像分析程序如图 1所示，可概括如下：

(a) 将字符数字化为图像；(b) 对图像进行二值化；(c) 获取图像的骨架以提取轮廓；(d) 用极坐标表示轮廓。

按照这一流程，通过 128 对极坐标重构每个手写字符的轮廓形状。请注意，为了消除尺寸对形状分析的影响，对轮廓进行了归一化。调整所有轮廓坐标，使其围成的面积等于 $1\,cm^2$。因此，在归一化之前提取表面尺寸测量值 (S)，作为一项额外的一般特征，用于测量每个环形字符的近似面积（例如，使用梯形法则或辛普森法则）。

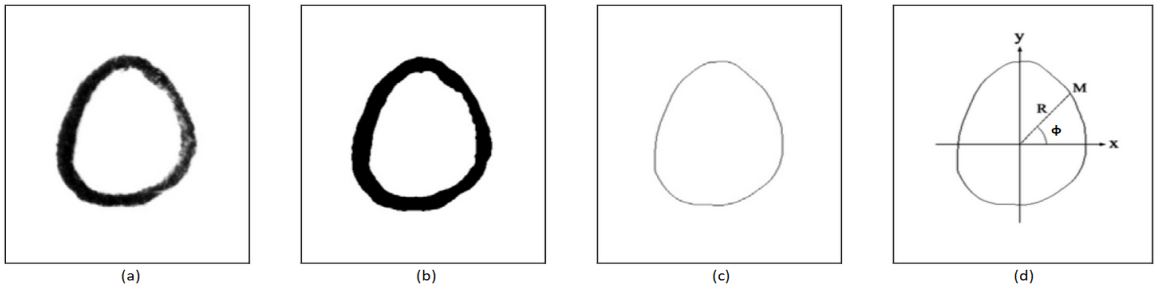


图1：图像分析程序：从原始字符环到极坐标。改编自Marquis 等人 (2005年)。

从极坐标出发，可以通过傅里叶级数重构每个环形字符：

$$R(\phi) = a_0 + \sum_{h=1}^{H}[a_h cos(h\phi) + b_h sin(h\phi)], \qquad \phi \in [0, 2\pi). \tag{1}$$

具体而言，每个字符环都可以用一个离散函数 R(ϕ来描述，该函数表示连接轮廓上的一点与质心的线段长度，其中 ϕ 是该线段与水平轴所成的角度。因此，轮廓形状可以用一系列谐波（$h = 1, \ldots\ldots, H$）来描述，每个谐波刻画对形状的一种特定贡献。每个谐波由一对傅里叶系数组成，即 a_h 和 b_h ($\in R$)。请注意，Marquis 等人 (2005年)最初提出使用傅里叶级数的振幅-相位形式：

$$R(\phi) = A_0 + \sum_{h=1}^{H}[A_h cos(h\phi - \phi_h)], \qquad \phi \in [0, 2\pi) \tag{2}$$

其中，振幅 A_h ($\in R^+$)和相位 ϕ_h (度或弧度)分别是系数 a_h 和 b_h，以及 $a_h = A_h cos(\phi_h)$、$b_h = A_h sin(\phi_h)$和 $A_0 = a_0$的极坐标。然而，傅里叶描述子（A_h 和 ϕ_h）取非负值，无法直接用多元高斯分布建模，而本研究选择多元高斯分布作为概率模型。此外，与傅里叶描述子（a_h 和 b_h）相比，傅里叶系数（Mardia（1970））不仅定义域是全体实数，而且在每位书写者上的 Mardia 偏度和峰度值也更优。即便对傅里叶描述子进行对数或平方根变换，这一结论仍然成立。因此，我们接下来使用傅里叶系数（a_h 和 b_h）进行分析。

字符轮廓形状中的谐波贡献如图2所示：第一谐波（$h = 1$）反映形状中的卵形贡献，第二谐波（$h = 2$）反映形状的椭圆度，第三谐波（$h = 3$）反映三角度，第四谐波（$h = 4$）反映四边形度，第五谐波（$h = 5$）反映五边形度。有关此图像分析程序的更详细说明，参见Schmittbuhl 等人 (1998年）和Marquis 等人(2005年）。

使用此类基于傅里叶的程序重建原始轮廓的结果如图3所示，涵盖每位书写者和每个字符。对于每个分析的环形字符，取傅里叶参数的平均值，谐波总数分别为 $H = 4$ 和 $H = 10$ 。最终图像呈现的是环形字符的平均轮廓。可以看出，有些书写者的环形结构具有明显的独特特征，而另一些则更为相似。此图未体现书写者内变异，这将在数据建模部分讨论（见第4节）。图中未展示尺寸，因为原始图像已归一化。

本文仅保留了前四个谐波，这是根据Marquis 等人（2006年的建议进行的。事实上，无需其他谐波的贡献即可重构每个字符轮廓的整体形状，如图3所示。此外，必须强调的是，常数项α_0 （或振幅-相位形式中的 A_0 ）不会影响所研究字符形状的形态特征，正如 Marquis 等人 （2006年）所讨论的那样。因此，该项被视为干扰参数，在后续分析中不再考虑。这样，每个字符环都可以通过 $p = 9$ 个变量进行描述，这些变量表示字符的表面大小（S）

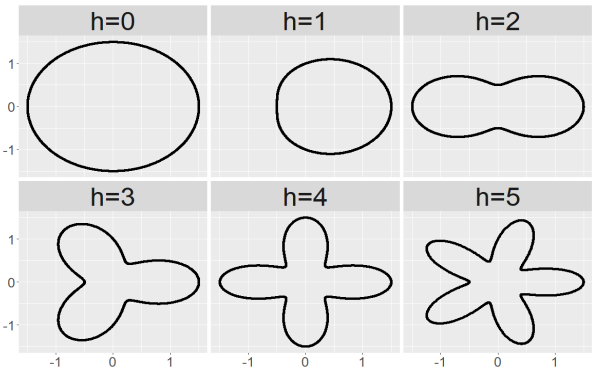
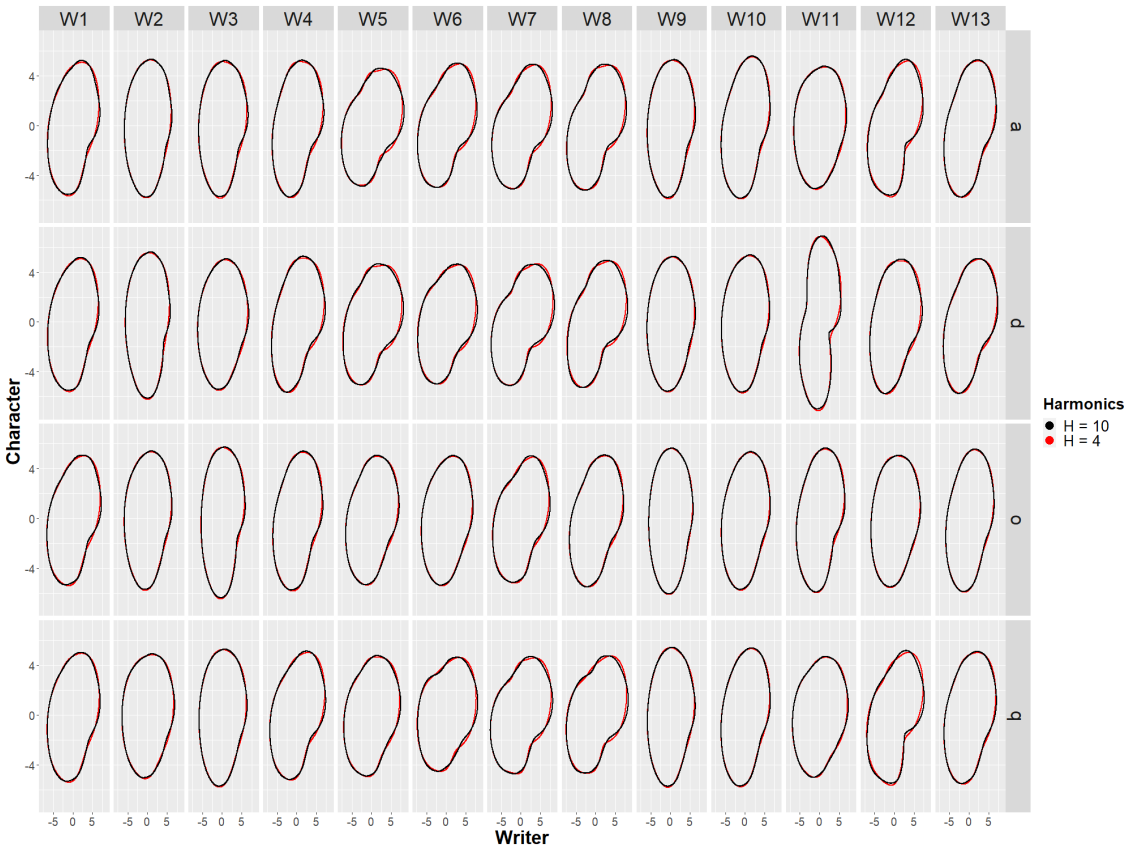


图2：单位圆（h=0）与关注的特定谐波（ $\alpha_h = 0.5$ 和 $b_h = 0$）之和所得到的谐波贡献

.



All characters were normalized to 1 cm^2 of surface size. H refers to the number of harmonics.

图3：展示了使用每位书写者和每个字符的平均傅里叶系数重建的环形字符：采用前四对傅里叶系数（H $= 4$；红色曲线）和前十对傅里叶系数（H= 10；黑色曲线）。

、前四个谐波的 α_h 和 b_h、 $h = 1, \ldots\ldots, 4$（$H = 4$）。因此，可用数据包含在对应于 $m = 13$名书写者处收集的 $L = 4$ 个字符（即 a、d、o 和 q）的测量值，并可表示为一个四维数组 D，其元素为 $\mathcal{D}_{i\ell j\kappa}$，其中 $i = 1, \ldots\ldots, m$（书写者）、 $\ell = 1, \ldots\ldots, L$（字符）、 $j = 1, \ldots\ldots, n_{i\ell}$（重复），以及 $\kappa = 1, \ldots\ldots, p$（傅里叶系数和表面大小）。请注意，傅里叶系数和表面大小均已通过将每个数值除以相应傅里叶系数的总体标准差进行标准化。有关可用字符特征的详细、全面分析，请参阅附录A。

3 笔迹证据评估的贝叶斯概率方法

法证文件检验员经常需要处理涉及书写者身份存疑的手写文件的情况。设想以下涉及一份来源存在争议的手写文件的情形。为进行比较分析，收集并检验被怀疑是争议材料实际来源的个人所书写的材料。应在审判中由对立双方（即检察官和辩护律师）分别提出的一组命题及给定信息框架下，对现有证据进行评估。因此，相关命题可表述如下：

H_1：相关人员（PoI）是检验文件的书写者；

H_2：相关人员（PoI）不是检验文件的书写者。

首先，从匿名手稿文件中选取 $n_1 > 0$ 个字符；这些字符称为检验材料。然后，从相关人员所书写的手写材料中选取 $n_2 > 0$ 个字符；这些字符称为对照材料或参考材料。对检验材料和对照材料进行的测量（即傅里叶系数和表面大小）表示为：$\boldsymbol{y}_{w\bullet\bullet\bullet} = (\boldsymbol{y}_{w\ell j\bullet},\ w = 1,\ 2,\ \ell = 1,\ \ldots\ldots,\ L,\ j = 1,\ \ldots\ldots,\ n_{w\ell})$，其中 $\boldsymbol{y}_1$ 表示对检验材料进行的测量，$\boldsymbol{y}_2$ 表示对对照材料进行的测量。因此，提供有关证据数值（$\boldsymbol{y}_1$，$\boldsymbol{y}_2$）的信息，以支持竞争命题，至关重要。这一信息由竞争命题下边际似然之比给出：

$$BF = \frac{m(\boldsymbol{y}_1, \boldsymbol{y}_2|H_1)}{m(\boldsymbol{y}_1, \boldsymbol{y}_2|H_2)} = \frac{m(\boldsymbol{y}_1, \boldsymbol{y}_2|H_1)}{m(\boldsymbol{y}_1|H_2)m(\boldsymbol{y}_2|H_2)} \tag{3}$$

式（3）中的边际似然可采用同一概率模型计算，该模型将用于描述手写体数据（并将在第4节中加以说明）。具体而言，分子中的边际似然 $m(\boldsymbol{y}_1, \boldsymbol{y}_2|H_1)$ 的计算假设合并数据集 $\{\boldsymbol{y}_1, \boldsymbol{y}_2\}$ 中的所有观测值共享同一个参数向量，因为假定 H_1 成立，且数据来源于同一位书写者。另一方面，若 H_2 成立，且相互竞争的材料来源于不同的书写者，则测量值 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$ 可视为相互独立。因此，边际似然 $m(\boldsymbol{y}_1, \boldsymbol{y}_2|H_2)$ 可表示为两个独立边际似然 $m(\boldsymbol{y}_1|H_2)$ 和 $m(\boldsymbol{y}_2|H_2)$ 的乘积。后两个边际似然采用同一概率模型计算，但分别对数据集 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$ 拟合该模型，并使用不同的参数向量。请注意，所采用的贝叶斯因子假设在质疑材料与对照材料之间满足 H_2条件下的独立性。这一假设意味着不考虑可能存在的伪装行为（有关所采用的贝叶斯因子假设，请参见附录B）。关于法庭科学中鉴定问题的进一步定义，请参见 Ommen等人（2017）。

本文分析并比较了六种不同的贝叶斯建模方法。第一组模型假设可用数据服从多元正态分布（第4.1节），并采用三种不同的先验设定：(a) 标准共轭先验（第4.1.1节）；(b) 一种层次先验结构，其中方差-协方差矩阵的建模独立于均值向量的先验分布（第4.1.2节）；以及 (c) 一种协方差分解方法，其中书写者内变异的先验服从对数正态-LKJ分布（第4.1.3节）。在此模型下，不对字符层级变异进行建模，因为所有观测字符共用一个方差-协方差矩阵。因此，所有测量值均被联合评估，不区分字符类型。因此，为了考虑字符类型的影响，将针对每种字符分别拟合模型。

为解决这一局限，第 4.2 节提出了一种贝叶斯多元方差分析模型。在该方法中，字符类型由一组虚拟变量（指示变量）表示（详见附录 C），从而使模型能够考虑字符间变异。与第一种建模方法类似，我们考虑三种先验设定方案：共轭先验（第 4.2.1 节）、层次正态-逆 Wishart 先验（第4.2.2 节）以及正态-对数正态-LKJ先验（第 4.2.3 节）。对于层次模型，我们使用桥接抽样估计计算贝叶斯因子（3）所需的边际似然；详见第 4.3 节。

4 傅里叶系数与表面大小建模

让我们考虑第 3 节所述的疑问材料和对照材料的可用测量值，分别记为 $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$，$\boldsymbol{y}_{w\bullet\bullet\bullet} = (\boldsymbol{y}_{w\ell j\bullet}$, $w = 1,\ 2,\ \ell = 1,\ \ldots,\ L,\ j = 1,\ \ldots,\ n_{w\ell})$。根据第 2 节所述的基于傅里叶的特征，这些数据构成一个长度为 p 的向量，对应于表面大小，并包含四对傅里叶系数，记为 $\mathrm{y}_{w\ell j\bullet}$。我们用 $\boldsymbol{\theta}_w \in \mathbb{R}^p$ 表示材料 w 内的均值向量，用 $\boldsymbol{W}_w \in \mathbb{R}^{p\times p}$ 表示材料 w 内的方差 - 协方差矩阵，其元素为 $W_{w\kappa_1\kappa_2}$，其中 $(\kappa_1,\ \kappa_2) \in \{1, 2, \ldots, p\}^2$。那么，给定 $\boldsymbol{\theta}_w$ 和 $\boldsymbol{W}_w$，$\mathrm{y}_{w\ell j\bullet}$ 的分布被视为 p 维正态分布 N_p，其中

$$\boldsymbol{y}_{w\ell j\bullet} \sim N_p\left(\boldsymbol{\theta}_w, \boldsymbol{W}_w\right). \tag{4}$$

进一步假设，我们有一个来自不同书写者、且与本案无关的手稿数据集（或数据库），记为 $\boldsymbol{X}$。我们将其称为背景数据，之后将在需要时用于指定模型参数（即，推导先验分布）。我们还假设，背景数据采用与 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$相同的特征工程流程。因此，我们的建模框架涉及三个不同的数据集：(i) 检材数据 $\boldsymbol{y}_1$，(ii) 对照数据 $\boldsymbol{y}_2$，以及 (iii) 背景数据 $\boldsymbol{X}$。检材数据和对照数据 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$直接用于通过贝叶斯因子（3）评估假设 H_1 和 H_2；而背景数据 $\boldsymbol{X}$ 则通过先验间接用于辅助各模型中的参数估计。

4.1 贝叶斯正态模型

在第一种建模方法中，为参数（$\boldsymbol{\theta}_w$，$\boldsymbol{W}_w$）选取正态-逆威沙特（NIW）先验分布。在这一通用先验设定下，并考虑4的抽样模型分布 $\boldsymbol{y}_{w\ell j\bullet}$，可指定如下贝叶斯模型：

$$\begin{aligned} \boldsymbol{y}_{w\ell j\bullet} &\sim N_p(\boldsymbol{\theta}_w, \boldsymbol{W}_w) \\ \boldsymbol{\theta}_w|\boldsymbol{X} &\sim N_p(\boldsymbol{\mu}, \boldsymbol{G}) \\ \boldsymbol{W}_w|\boldsymbol{X} &\sim IW(\boldsymbol{U}, \nu). \end{aligned} \tag{5}$$

在此模型形式中，IW 表示逆威沙特分布，$\boldsymbol{\mu}$ 是 $\boldsymbol{\theta}_w$的先验均值向量，$\boldsymbol{W}_w$ 是书写者内协方差矩阵，$\boldsymbol{G}$ 是 $\boldsymbol{\theta}_w$的协方差矩阵，而 $\boldsymbol{U}$ 和 ν 分别表示用于描述书写者内变异的逆威沙特分布的尺度矩阵和自由度。

根据先验方差-协方差矩阵 $\boldsymbol{G}$的设定，式（5）中的贝叶斯模型有两种不同的变体。首先，令 $\boldsymbol{G} = \boldsymbol{W}_w k_0^{-1}$，即可得到共轭先验（第 4.1.1节），先验均值向量记为 $\boldsymbol{\theta}_w|\boldsymbol{X}$，$\boldsymbol{W}_w$。相反，若将参数 $\boldsymbol{G}$ 设为书写者间协方差矩阵 $\boldsymbol{B}$，则可得到采用层次先验设置的模型（第 4.1.2节）。为了改进后一种先验方法，我们对书写者内协方差矩阵 $\boldsymbol{W}_w$进行协方差分解，并采用第 4.1.3节所述的对数正态-LKJ先验方法。

4.1.1 共轭先验：正态-逆威沙特

假设数据服从多元正态分布 $N_p(\boldsymbol{\theta}_w,\ \boldsymbol{W}_w)$，则 $(\boldsymbol{\theta}_w,\ \boldsymbol{W}_w)$ 对应的共轭先验分布为正态-逆威沙特（NIW）分布，

$$(\boldsymbol{\theta}_w, W_w)|\boldsymbol{X} \sim NIW(\boldsymbol{\mu}, k_0, \boldsymbol{U}, \nu),$$

其中，先验参数 k_0 是一个控制先验精度的大小以及 $\boldsymbol{\theta}_w$方差的参数。k_0越高，$\boldsymbol{\theta}_w$的先验分布所包含的信息就越多。此外，$\boldsymbol{U}$ 和 ν分别表示逆威沙特分布的尺度矩阵和自由度；更多详情请参见 Gelman 等人(2013年，第 3.6 章)。因此，将 $\boldsymbol{y}_{w\ell j}$ 的抽样模型分布（5）与 NIW 先验分布结合后，即可得到贝叶斯模型（$\boldsymbol{G} = \boldsymbol{W}_w k_0^{-1}$。

由于正态-逆威沙特分布先验是共轭的，因此可用测量值 $\boldsymbol{y}$ 的边际似然可以直接写成闭式，形式如下：

$$m(\boldsymbol{y}) = \frac{1}{\pi^{np/2}} \frac{\Gamma_p(\nu_n/2)}{\Gamma_p(\nu/2)} \frac{|\boldsymbol{U}|^{\nu/2}}{|\boldsymbol{U}_n|^{\nu_n/2}} \left(\frac{k_0}{k_n}\right)^{p/2},$$

其中， n 为样本量， p 为变量数， Γ_p 为多元伽马函数，且

$$k_n = k_0 + n,\ \nu_n = \nu + n,\ \boldsymbol{U}_n = \boldsymbol{U} + \boldsymbol{S} + \frac{k_0 n}{k_0 + n}(\bar{\boldsymbol{y}} - \boldsymbol{\mu})(\bar{\boldsymbol{y}} - \boldsymbol{\mu})^T,\ \boldsymbol{S} = \sum_i (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T\ ;$$

其中， $\bar{\boldsymbol{y}}$ 待检材料与对照材料的平均测量值，更多详情请参见 Murphy (2007年)。在本研究中，先验参数 $\boldsymbol{\mu}$ 和 $\boldsymbol{U}$ 根据易于获取的背景数据 $\boldsymbol{X}$中的信息确定；参见附录 D.1。

自由度 ν 设为 $p+2$，这是逆威沙特分布的均值存在的最小取值（Press 2005年）。最后，先验参数 k_0 通过对背景数据中所有可用书写者的 (0,1) 区间进行网格搜索来选取。具体而言，选取能够使背景数据的边际似然最大化的数值 k_0 。

尽管共轭先验在计算上十分便利，因而是自然的选择，但它有一个缺点：无法分别对同一书写者内部和不同书写者之间的变异进行建模。在这一特定情境下，这可能构成重大限制，因为书写的基本规律指出：(1) 没有人会两次以完全相同的方式书写同一个词（即书写者内部的变异）；(2) 也没有两个人会以完全相同的方式书写（即书写者间变异）。因此，我们考虑了另一个对先验进行独立设定的模型。

4.1.2 层次扩展：正态分布-逆威沙特

Bozza 等人（2008年）为模型参数（$\boldsymbol{\theta}$， $\boldsymbol{W}$）$|\boldsymbol{X}$ 设定了如下形式的层次先验

$$\boldsymbol{\theta}_w|\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{B}) \text{ and } \boldsymbol{W}_w|\boldsymbol{X} \sim IW(\boldsymbol{U}, \nu). \tag{6}$$

Th所提出的模型由（5）给出，其中 $\boldsymbol{B} = \boldsymbol{G}$ 固定，且独立于 W_w .

具体而言，先验参数 $\boldsymbol{B}$ 表示书写者间的方差-协方差矩阵。先验参数 $\boldsymbol{\mu}$、 $\boldsymbol{B}$和 $\boldsymbol{U}$ 可根据背景书写者的信息，以相对直接的方式估计（参见附录 D.1）。实证证据表明，纳入书写者间变异的模型表现优于不具备这种灵活性的模型（Bozza等（2008））。必须指出，缺少背景信息会大幅削弱有效区分书写者的能力。

遗憾的是，该模型的边际似然无法通过解析方法求得。因此，按照第 4.3 节所述并讨论的方法，基于 MCMC算法的输出应用桥接抽样进行估计。具体而言，MCMC算法 No-U-Turn 采样器（NUTS）通过*Stan* （参见Carpenter等， 2017）2.26.1 版实现，并通过*RStan*接口和*R* 4.2.1 版运行。

4.1.3 协方差分解：正态-对数正态-LKJ

在贝叶斯层级模型中，对协方差矩阵建模的一种广泛使用且灵活的方法，是将协方差矩阵分解为标准差和相关矩阵。具体而言，协方差矩阵 $\boldsymbol{W}_w$ 可以表示为：

$$\boldsymbol{W}_w = \boldsymbol{D}_w \boldsymbol{R}_w \boldsymbol{D}_w \text{ and } \boldsymbol{D}_w = Diag\left(W_{w11}^{1/2}, W_{w22}^{1/2}, \ldots, W_{wpp}^{1/2}\right); \tag{7}$$

其中， $\boldsymbol{D}_w$ 是一个对角矩阵，包含材料 w 的协变量的标准差（对于$w = \{1, 2\}$），而 $\boldsymbol{R}_w$ 是相应的相关矩阵。 $\boldsymbol{W}_w$ 的对角元素指定了 $\boldsymbol{D}_w$ 的相应元素，而 $\boldsymbol{R}_w$ 则通过对 $\boldsymbol{W}_w$进行标准化得到，即只需令 $\boldsymbol{R}_w = \boldsymbol{D}_w^{-1} \boldsymbol{W}_w \boldsymbol{D}_w^{-1}$。

在这种分解中，标准差 $\boldsymbol{D}_w$ 通常采用独立先验，例如对数正态分布，或者采用半柯西分布作为信息较弱的先验设定。因此，这种替代参数化方式可以更灵活地对边际方差进行建模。相关矩阵 $\boldsymbol{R}_w$ 则单独采用 Lewandowski-Kurowicka-Joe（LKJ）分布建模（Lewandowski 等人 2009年），该分布提供了一种专为相关性设计的非共轭先验

矩阵。LKJ分布具有一个形状参数 $\eta > 0$，用于控制先验向单位矩阵（独立性）集中的程度，其概率函数为：

$$\begin{aligned} f(\boldsymbol{R};\eta) &= \mathcal{C} \times |\boldsymbol{R}|^{\eta-1} \text{ with} \\ \log(\mathcal{C}) &= \sum_{\kappa=1}^{p}(2\eta-2+p-\kappa)\log 2 + \sum_{\kappa=1}^{p-1}(p-\kappa)\log \mathcal{B}(\eta+\tfrac{p-\kappa-1}{2},\ \eta+\tfrac{p-\kappa-1}{2}); \end{aligned}$$

其中， $\boldsymbol{R}$ 是一个对角线为单位值的 $p \times p$ 正定矩阵， $|\boldsymbol{R}|$ 表示矩阵 $\boldsymbol{R}$的行列式，归一化常数 $\mathcal{C}$确保 $f(\boldsymbol{R};\ \eta)$在 $p \times p$相关矩阵空间上的积分为 1， p 是相关矩阵 $\boldsymbol{R}$的维数， κ 是从 1 到p （或在第二个求和式中从 1 到 $p-1$ ）的索引，而 $\mathcal{B}()$是 Beta函数。

因此，贝叶斯模型设定如下：

$$\begin{aligned} \boldsymbol{y}_{w\ell j\bullet} &\sim N_p(\boldsymbol{\theta}_w, \boldsymbol{W}_w), \quad \boldsymbol{W}_w = \boldsymbol{D}_w \boldsymbol{R}_w \boldsymbol{D}_w \\ \boldsymbol{\theta}_w|\boldsymbol{X} &\sim N_p(\boldsymbol{\mu}, \boldsymbol{B}) \\ D_{w\kappa\kappa}|\boldsymbol{X} &\sim \text{LogNormal}(\upsilon, \sigma) \text{ for } \kappa \in \{1,2,\ldots,p\}, \\ \boldsymbol{R}_w &\sim LKJ(\eta) \end{aligned} \tag{8}$$

对数正态分布（υ， σ）由位置参数 υ 和尺度参数 σ参数化；这两个参数根据背景数据所推导出的标准差先验知识进行估计，从而实现先验 elicitation（见附录 D.1），而 LKJ 参数 η 设为 1，作为相关矩阵的无信息先验。选择对数正态先验，是基于背景数据所推导出的先验知识，以便确定先验参数 υ 和σ。其他模型参数也采用相同的先验设定方法，详见第 4.1.2 节。正态-对数正态-LKJ 先验设置的边际似然无法通过解析方法以闭式形式求得。因此，采用桥接抽样来估计边际似然，详见第 4.3 节。为提取参数的后验样本，使用 *Stan*（见 Carpenter等，2017）2.26.1 版，通过 *RStan* 接口在 *R* 4.2.1 版中运行 MCMC算法 No-U-Turn 采样器（NUTS）。

根据 Huang & Wand (2013年)，与经典的逆Wishart先验方法相比，LKJ先验方法的表现可能更好，尤其是在需要对相关矩阵采用更灵活的先验，或关于协方差结构的先验信息有限时。此外，在稀疏协方差矩阵估计领域，LKJ先验是最常用的方法。

第 4.1 节所讨论的所有建模方法都未考虑字符变异性，因为不同类型的字符要么不加区分地处理，要么分别分析。上述两种方法——分析全部数据但不区分字符，或分别分析每个字符——都有明显的缺点。具体而言，第一种方法不区分字符，因此既忽略了每位书写者手写字符的变异性，也忽略了不同书写者之间字符的差异。另一方面，分别分析每个字符会带来一个难题：如何将个人分析结果汇总为针对 H_1 或 H_2的单一证据度量。针对不同字符的贝叶斯因子无法直接合并；当出现冲突证据时，情况会更加复杂，也就是说，不同类型的字符支持不同的假设。因此，在第 4.2 节中，我们将提出考虑字符内部变异性的模型。

4.2 贝叶斯多元方差分析

本节将介绍第二种建模方法，以尝试解决上述关键问题。下面将要描述的贝叶斯多元方差分析模型具有一项显著优势，即还可以对字符层级变异进行建模。具体而言，通过选择字符 a 作为参照组，将字符的指标转换为虚拟变量（见附录C）。因此，使用字符虚拟变量作为预测变量构建回归模型。

该模型的一般形式可以写成

$$\begin{aligned} \boldsymbol{y}_{w\ell j\bullet} &\sim N_p(\boldsymbol{\Theta}_w^T \boldsymbol{C}_{w\bullet j}, \boldsymbol{W}_w) \\ \boldsymbol{\Theta}_w|\boldsymbol{W}_w, \boldsymbol{X} &\sim MN_{L,p}(\boldsymbol{\mathcal{M}}, \boldsymbol{K}_0^{-1}, \boldsymbol{W}_w) \\ \boldsymbol{W}_w|\boldsymbol{X} &\sim IW(\boldsymbol{U}, \nu), \end{aligned} \tag{9}$$

其中，$MN_{L,\ p}$ 表示具有 L=4 个字符和 $p=9$（即 4 对傅里叶系数及表面大小）的矩阵正态分布；$\boldsymbol{C}$是一个三维数组，包含针对 $\boldsymbol{y}$的解释变量虚拟变量。子向量 $\boldsymbol{C}_{w\bullet j}$ 的长度为 L ，包含每个书写者以及每个保留字符重复 L 的虚拟变量。 $\boldsymbol{C}$ 中的每个元素 $C_{w\ell j}$ 定义为 $C_{wlj}=1$，其中 $l=1$ 且对所有 w，j均成立；而对于 $l=2$，……，L，其定义为

$$\begin{aligned} C_{w\ell j} &= 1, \quad \text{if } \ell = \ell_{wj}^{obs}; \\ C_{w\ell j} &= 0, \quad \text{otherwise}, \end{aligned}$$

其中，ℓ_{wj}^{obs} 表示疑问材料或对照材料 w 在重复 j 下的观测字符（即 a对应 (1,0,0,0)，d对应 (1,1,0,0)，o 对应 (1,0,1,0)，等等）；更详细的说明请参见附录 C 。此外， $\boldsymbol{\Theta}_w$ 是 $L\times p$ 系数矩阵，可改写如下

$$\boldsymbol{\Theta}_w = \begin{pmatrix} \boldsymbol{\vartheta}_{w1}^T \\ \boldsymbol{\vartheta}_{w2}^T \\ \vdots \\ \boldsymbol{\vartheta}_{wL}^T \end{pmatrix}, \tag{10}$$

其中每个 $\boldsymbol{\vartheta}_{wl}$，$l=1$，……，$L$，是一个长度为 p的向量。 $\boldsymbol{W}_w$是书写者内协方差矩阵， ν表示自由度， $\boldsymbol{U}$ 是逆威沙特分布的尺度矩阵，$\boldsymbol{\mathcal{M}}$ 是先验均值矩阵（维度为 $L\times p$的先验均值矩阵），且 $\boldsymbol{K}_0$是一个 $L\times L$用于调整 $\boldsymbol{\Theta}_w$方差的尺度矩阵。（9）中的矩阵正态分布可以改写为

$$\text{vec}(\boldsymbol{\Theta}_w)|\boldsymbol{X} \sim N_{L\times p}(\text{vec}(\boldsymbol{\mathcal{M}}), \boldsymbol{W}_w \otimes \boldsymbol{K}_0^{-1})$$

其中，vec() 表示对矩阵进行向量化， $\otimes$ 表示克罗内克积。因此，vec($\boldsymbol{\Theta}_w$) 的总体方差可直接表示为 $\boldsymbol{W}_w \otimes \boldsymbol{K}_0^{-1}$。然而，由于此处关注的是每种不同类型字符的分布（即 $\boldsymbol{\Theta}_w$ 的每一行）， $\boldsymbol{W}_w$ 将作为主要方差分量，而 $\boldsymbol{K}_0^{-1}$将作为方差乘数，并利用背景数据进行估计。

4.2.1 共轭先验

关于第4.1.1节中的共轭正态-逆 Wishart先验设置，出于计算便利性的考虑，首先为（9）中的多元方差分析模型考虑共轭方法。在此方法下，使用向量化参数的自然共轭先验 $\boldsymbol{\Theta}_w$ 形式如下：

$$\begin{aligned} \text{vec}(\boldsymbol{\Theta}_w)|\boldsymbol{W}_w, \boldsymbol{X}, \boldsymbol{K}_0 &\sim N_{L\times p}(\text{vec}(\boldsymbol{\mathcal{M}}), \boldsymbol{W}_w \otimes \boldsymbol{K}_0^{-1})) \\ \boldsymbol{W}_w|\boldsymbol{X} &\sim IW(\boldsymbol{U}, \nu), \end{aligned}$$

其中，此处将 $\boldsymbol{K}_0$ 视为从背景数据中提取的固定参数。在上述贝叶斯表述下，后验分布仍属于同一分布族。因此，边际似然可以写成闭式，具体如下：

$$m(\boldsymbol{y}) = \frac{1}{2\pi^{np/2}} \frac{\Gamma_p(\nu_n/2)}{\Gamma_p(\nu/2)} \frac{|\boldsymbol{U}/2|^{\nu/2}}{|\boldsymbol{U}_n/2|^{\nu_n/2}} \left(\frac{|K_0|}{|K_n|}\right)^{p/2},$$

其中， n 为样本量， p 为变量的数量（即傅里叶系数和表面大小），

$$\begin{aligned} \nu_n &= \nu + n \\ K_n &= C^T C + K_0 \\ \boldsymbol{U}_n &= \boldsymbol{U} + \boldsymbol{y}^T\boldsymbol{y} + \boldsymbol{\mathcal{M}}^T K_0 \boldsymbol{\mathcal{M}} - \boldsymbol{\mathcal{M}}_n^T K_n \boldsymbol{\mathcal{M}}_n, \text{ and} \\ \boldsymbol{\mathcal{M}}_n &= K_n^{-1}\left(C^T\boldsymbol{y} + K_0\boldsymbol{\mathcal{M}}\right); \end{aligned}$$

参见 Rowe （2002年，第 8.4 章）以及 Soch （2019，第 2.4 章）。关于先验参数设定的详细说明，请参见附录 D.2。

然而，正如前文所指出的，尽管共轭方法在计算上可能颇具吸引力，但它有一个重要局限：无法对书写者间异质性（即书写者间变异）建模。因此，为了克服这一问题，下一节将介绍多元方差分析模型的一个非共轭版本。

4.2.2 层次扩展：多元方差分析正态分布-逆威沙特

我们考虑模型（9）的一个简化版本，其中 $\boldsymbol{K}_0^{-1}$设为单位矩阵，且$\boldsymbol{W}_w$ 在 $\boldsymbol{\Theta}_w$的抽样分布中被替换为书写者间协方差矩阵 $\boldsymbol{B}$ 。这样可以写成

$$\boldsymbol{\Theta}_w|\boldsymbol{W}_w, \boldsymbol{X} \sim MN_{L,p}(\boldsymbol{\mathcal{M}}, \boldsymbol{I}_L, \boldsymbol{B}),$$

化简后得到

$$\boldsymbol{\theta}_{w\ell}|\boldsymbol{X} \quad \sim \quad N_p(\boldsymbol{\mu}_\ell, \boldsymbol{B}), \qquad \text{for} \quad \ell \in \{1, \ldots, L\}$$

其中， $\boldsymbol{\mu}_\ell$ 是 $\boldsymbol{\theta}_{w\ell}$的先验均值向量。此外，将共同的方差-协方差矩阵 $\boldsymbol{B}$替换为 $\boldsymbol{B}_l, l = 1$、.....、 L，以考虑上述模型的异方差版本；在该版本中，不同字符层级 l的书写者间变异有所不同。在这些设定下，最终模型为

$$\begin{aligned} \boldsymbol{y}_{w\ell j\bullet} &\sim N_p(\boldsymbol{\Theta}_w^T \boldsymbol{C}_{w\bullet j}, \boldsymbol{W}_w) \\ \boldsymbol{\theta}_{w\ell}|\boldsymbol{X} &\sim N_p(\boldsymbol{\mu}_\ell, \boldsymbol{B}_\ell) \quad \text{for} \quad \ell \in \{1, \ldots, L\} \\ \boldsymbol{W}_w|\boldsymbol{X} &\sim IW(\boldsymbol{U}, \nu). \end{aligned} \tag{11}$$

第4.1节所述的 elicitation 程序经过了少量调整，以便根据现有背景数据指定均值向量 $\boldsymbol{\theta}_{w\ell}$ 和协方差矩阵 $\boldsymbol{W}_w$的先验分布参数。具体而言，通过计算估计的 $\widehat{\boldsymbol{\mu}}_a$（即字符 a 的背景数据样本均值）与 $\widehat{\boldsymbol{\mu}}_l$之间的差值，得出了表征 $\boldsymbol{\theta}_\ell$ 针对 $\ell = 2$、.....、 L 的先验分布的超参数数值；对于 $l = 2$、.....、 L， L是字符 l 相对于字符 a 的样本均值差。这种方法有效地捕捉了各字符 l 与参考字符 a 之间的均值差异。更多详情见附录 D.2。至于表征 $\boldsymbol{W}_w$先验分布的尺度矩阵 $\boldsymbol{U}$，则采用背景数据中每位书写者的方差-协方差矩阵的平均值。因此，这是对书写者内变异的初始先验估计，其后验估计则通过 $\boldsymbol{W}_w$ ，利用材料 w的手写体数据得到。同样， $\boldsymbol{B}_\ell$ 是字符 ℓ的方差-协方差矩阵，而 $\boldsymbol{B}_\ell$ 则根据背景数据中所有书写者对字符 ℓ 测量值所得的样本方差-协方差矩阵确定。类似地，正如第4.1.2节（NIW 的模型独立先验设置）所述，如果先验中未纳入背景信息，那么有效区分不同书写者的能力将显著下降。

与第4.2.1节的共轭方法不同，在这种情况下，边际似然无法解析地写成闭式表达式。因此，使用 MCMC 样本通过桥接抽样来估计边际似然，具体方法将在第4.3节介绍。具体而言，通过*Stan*(参见 Carpenter等，2017) 2.26.1 版实现 No-U-Turn 采样器（NUTS），并通过*RStan*接口和*R* 4.2.1 版运行。

4.2.3 协方差分解：多元方差分析正态-对数正态-LKJ分布

与第 4.1.3 节类似，我们通过将协方差结构表示为标准差和相关矩阵的函数，对模型（11）中的书写者内变异进行分解。具体而言，我们为标准差指定独立的对数正态先验，并将 LKJ分布用作相关矩阵的先验。所得模型可表示如下：

$$
\begin{aligned}
\boldsymbol{y}_{w\ell j\bullet} &\sim N_p(\boldsymbol{\Theta}_w^T \boldsymbol{C}_{w\bullet j}, \boldsymbol{W}_w), \quad \boldsymbol{W}_w = \boldsymbol{D}_w \boldsymbol{R}_w \boldsymbol{D}_w \\
\boldsymbol{\theta}_{w\ell}|\boldsymbol{X} &\sim N_p(\boldsymbol{\mu}_\ell, \boldsymbol{B}_\ell) \text{ for } \ell \in \{1, \ldots, L\} \\
D_{w\kappa\kappa}|\boldsymbol{X} &\sim \text{LogNormal}(\upsilon, \sigma) \text{ for } \kappa \in \{1, 2, \ldots, p\}, \\
\boldsymbol{R}_w &\sim LKJ(\eta)
\end{aligned} \tag{12}
$$

对于正态–对数正态–LKJ 先验设置，其边际似然没有闭式表达式，与第 4.3 节类似。因此，我们按照第 4.3 节所述，使用桥接抽样对其进行估计。

4.3 边际似然估计量

本节介绍数据分析中所考虑的两种主要情形下的边际似然及其估计方法。第一种情形下，我们使用完整数据集$\boldsymbol{\mathcal{D}}$ 以及与单个作者对应的数据子集，记为$\boldsymbol{\mathcal{D}}_i$，其中 $i = 1$、.....、13，用于评估和比较第 $M_1 - M_6$ 4节中描述的模型；参见表1。

第二种情形下，我们通过使用贝叶斯因子方程（ H_1 与 H_2之间的比较）来评估我们关注的主要比较，具体采用贝叶斯因子方程（3）及其对应的边际似然。此处使用了三个不同的数据集： $\boldsymbol{y}_1$、$\boldsymbol{y}_2$和 $\boldsymbol{X}$，其定义见第4节开头。这些数据集被视为完整数据集$\boldsymbol{\mathcal{D}}$的子集。关于如何从 $\boldsymbol{y}_1$、 $\boldsymbol{y}_2$和 $\boldsymbol{X}$ 构建这些数据集，取决于比较情形；具体说明见第$\boldsymbol{\mathcal{D}}$ 5节的实验结果。

值得注意的是，在模型比较和证据评估中，从完整数据集 $\boldsymbol{\mathcal{D}}$ 中选取子集，是基于这样一项原则：数据不应被重复使用（一次用于先验，另一次用于似然），以避免结果产生偏差。

模型	描述（模型和先验设定）	先验缩写	章节
M_1	采用共轭正态-逆 Wishart 先验的正态模型	NIW 共轭模型	4.1.1
M_2	采用分层正态-逆 Wishart 的正态模型	NIW 分层模型	4.1.2
M_3	采用正态-对数正态-LKJ 的正态模型		4.1.3
M_4	采用共轭正态-逆 Wishart 的多元方差分析模型	NIW 共轭模型	4.2.1
M_5	采用分层正态-逆 Wishart 的多元方差分析模型	NIW 分层模型	4.2.2
M_6	正态-对数正态-LKJ 多元方差分析模型		4.2.3

表1：研究中考虑的模型和先验设定概览。

4.3.1 使用完整数据集进行模型比较的边际似然

为了比较所提出的模型，我们针对每位书写者，将任意一对模型 M_l和 M_ξ 的边际似然之比作为贝叶斯因子进行计算，其中 l, $\xi \in \{1, \ldots, 6\}$。贝叶斯因子量化了基于每位书写者可用数据 $\boldsymbol{\mathcal{D}}_i$ 的被比较模型的相对证据， $i = 1$, .. . , m，先验参数则利用其余书写者的数据 $(m - i)$ 进行设定。贝叶斯因子越大，表明对其中一个模型相较于另一个模型的支持越强。

使用每位书写者的数据$\boldsymbol{\mathcal{D}}_i$ 得到的各模型边际似然为 $i = 1, \ldots, m$,

$$
m(\boldsymbol{\mathcal{D}}_i|M_l) = \int f(\boldsymbol{\mathcal{D}}_i|\boldsymbol{\Theta}, \boldsymbol{W}, M_l)\pi(\boldsymbol{\Theta}, \boldsymbol{W}|M_l)d(\boldsymbol{\Theta}, \boldsymbol{W}). \tag{13}
$$

对于任意模型 M_l（针对 l = 1、2、3、4、5、6）以及模型参数 $\boldsymbol{\Theta}$ 和 $\boldsymbol{W}$。模型 M_1 和 M_4 基于共轭先验设置，这意味着边际似然可以用解析形式表示。然而，模型（M_2, M_3）和（M_5, M_6）并非如此，其边际似然无法用闭式表达式表示。在这种情况下， M_2、 M_3、 M_5 和 M_6 的边际似然通过桥接抽样近似进行估计；桥接抽样近似是一种广泛用于边际似然估计的蒙特卡罗方法。该过程通过 R 中的 bridgesampling 软件包实现（Gronau、Singmann 与 Wagenmakers（2020）），该软件包为计算贝叶斯建模中的归一化常数提供了一个高效框架。值得注意的是，Gronau、Heathcote 和 Matzke（2020）提出的“Warp-III”变体已在 bridgesampling 软件包中实现，指的是一种基于变换的策略，旨在进一步提高边际似然估计的效率和准确性。通过对后验样本应用非线性变换，“Warp-III”旨在

使后验分布与提议分布的形状更好地匹配。有关桥接抽样近似的更详细说明，请参阅附录 E。最后，在第 E.1节中，我们展示了每位书写者所考虑的各个模型的边际似然估计的蒙特卡罗误差。

4.3.2 笔迹证据评估的边际似然

用于评估式（3）中贝叶斯因子的边际似然（ML），表示在竞争命题 $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2) = (\boldsymbol{y}_{w\ell j\bullet}, w = 1{,}2, \ell = 1, ..L, j = 1, ..., n_{w\ell})$下响应数据对应的证据度量 H_1 和 H_2，针对特定概率模型 M。此处，M 指第4节介绍的六种贝叶斯模型中的任意一种 $\{M_1, M_2, M_3, M_4, M_5, M_6\}$；详见表 41。

对于 式（3）的分子，假设 H_1 下的边际似然表示为 ：

$$m(\boldsymbol{y}_1, \boldsymbol{y}_2|H_1, M_l) = \int f(\boldsymbol{y}_1, \boldsymbol{y}_2|\boldsymbol{\Theta}, \boldsymbol{W}, H_1, M_l)\pi(\boldsymbol{\Theta}, \boldsymbol{W}|H_1, M_l)d(\boldsymbol{\Theta}, \boldsymbol{W}). \tag{14}$$

对于式（3）的分母，在假设 H_2 下：

$$m(\boldsymbol{y}_w|H_2, M_l) = \int f(\boldsymbol{y}_w|\boldsymbol{\Theta}, \boldsymbol{W}, H_2, M_l)\pi(\boldsymbol{\Theta}, \boldsymbol{W}|H_2, M_l)d(\boldsymbol{\Theta}, \boldsymbol{W}), \text{ for } w \in \{1, 2\}, \tag{15}$$

以及任意模型 M_l（对于 $l = 1$、2、3、4、5、6）和模型参数 $\boldsymbol{\Theta}$ 与 $\boldsymbol{W}$。若假设 H_1 成立（即相关人员是待检验手稿的书写者），则假定模型参数相等，即 $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 = \boldsymbol{\Theta}$ 与 $\boldsymbol{W}_1 = \boldsymbol{W}_2 = \boldsymbol{W}$。另一方面，若假设 H_2 成立（即相关人员不是检验文件的书写者），则假定模型参数（$\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ 与 $\boldsymbol{W}_1, \boldsymbol{W}_2$）在两个不同来源 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$之间不同。

在第4.1节和第4.2节中介绍的模型表述里，对于模型 M_1, M_2, M_3，在 H_2 下的参数向量 $\boldsymbol{\Theta}$记为 $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$；对于模型 M_4, M_5, M_6，则记为 $\boldsymbol{\Theta} = (\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2)$，其中 $\boldsymbol{\Theta}_w$ 按照（10）、 $w = 1{,}2$ 中的方式指定。此外，模型 M_1 和 M_4 基于共轭先验设定，因此可以解析形式获得其边际似然。然而，对于模型 M_2, M_3, M_5和 M_6，边际似然无法用闭式表达式表示。因此，我们使用通过 R 中的 bridgesampling 软件包实现的“Warp-III”方法进行桥接抽样近似来估计这些边际似然（Gronau、Singmann 与 Wagenmakers（2020））。

5 实验结果

本节分析了第 3 节介绍的笔迹检验问题，案例情境涉及一份来源存在争议的文件。第 4 节所述的两种似然结构和三种不同先验设定所形成的模型（共六种模型；见表 1）使用第 2 节所述的数据集进行了比较。所考察的模型方法包括贝叶斯正态模型的共轭和非共轭版本，分别采用正态-逆威沙特先验设置或正态-对数正态-LKJ先验设置（M_1、 M_2 和 M_3），以及贝叶斯多元方差分析模型的共轭和非共轭版本，分别采用正态-逆威沙特先验设置或正态-对数正态-LKJ先验设置（M_4、 M_5、 M_6）。

首先，在第 5.1 节中，我们针对完整可用数据中的每位个人书写者，比较 M_1、.....、 M_6。$\boldsymbol{\mathcal{D}}_i$ 对于每位书写者 $i = 1$、.....、 m。对于每位书写者 m，该过程支持在各模型之间进行成对比较。此外，在为所考虑的模型进行先验参数 elicitation 时，会使用所有其他书写者的数据（即排除正在评估的书写者）。具体而言，在比较书写者 i的模型时，会利用其余 $m - i$ 名书写者的数据进行先验 elicitation。

随后，在第 5.2 节中，我们按照第 3 节所述，在贝叶斯因子框架下评估所考察模型的性能。具体而言，我们开展了一系列旨在模拟真实案例研究的实验。在同一书写者实验中，从完整数据集 i 中选取每位书写者$\boldsymbol{\mathcal{D}}$，并将其数据随机划分为待检数据和对照集。随后评估式3 中的贝叶斯因子。所有其他书写者的数据（$m - i$）用作背景数据 $\boldsymbol{X}$，以确定所考察模型的先验参数。在不同书写者实验中，从完整数据集$\boldsymbol{\mathcal{D}}$中随机选取两位不同书写者的数据，并按照同一书写者实验中所述的相同流程进行。

5.1 按书写者进行模型比较

本节将第1 表中所考虑的六个模型进行比较，这些模型已在第 4 节中介绍。对于每位书写者 i，将模型 M_l 拟合到表征该书写者的数据$\boldsymbol{\mathcal{D}}_i$ 上，并使用其余书写者的数据拟合先验参数（在确定先验参数时排除i 该书写者）。随后估计每位书写者 i对应的所有待比较模型的边际似然。对于书写者 M_l 和 M_ξ 的两个模型 i，用于比较二者的贝叶斯因子定义如下：

$$\mathbf{BF}_{l,\xi}(\boldsymbol{\mathcal{D}}_i) = \frac{m(\boldsymbol{\mathcal{D}}_i \mid M_l)}{m(\boldsymbol{\mathcal{D}}_i \mid M_\xi)} \text{ for } l, \xi \in \{1, 2, 3, 4, 5, 6\}.$$

贝叶斯多元方差分析与正态模型的比较。首先，将贝叶斯多元方差分析模型与贝叶斯正态模型进行比较，分别采用共轭（M_4 与 M_1）和非共轭方法（M_5 与 M_2）、（M_6 与 M_3）版本。因此，对每位书写者分别计算贝叶斯因子 $\text{BF}_{4,1}$，$\text{BF}_{5,2}$和 $\text{BF}_{6,3}$ 。需要注意的是，计算 $\text{BF}_{4,1}$ 所需的边际似然可以通过解析方法获得，而第4.3 节所述的桥接抽样则用于估计 $\text{BF}_{5,2}$ 和 $\text{BF}_{6,3}$。表2 给出了 $\text{BF}_{4,1}$ （第二列）、$\text{BF}_{5,2}$ （第三列）和 $\text{BF}_{6,3}$ （第四列）的结果（以对数尺度表示）。

书写者	正态-逆威沙特		正态-对数正态-LKJ
	共轭方法 对数 $BF^{*}_{4,1}$	层次方法 对数 $BF^{**}_{5,2}$	协方差分解 对数 $BF^{**}_{6,3}$
1	20.6	19.1	16.6
2	174.7	173.4	164.5
3	113.5	115.7	89.7
4	92.8	93.1	78.8
5	205.9	203.8	202.9
6	116.3	111.7	114.2
7	156.8	157.3	152.1
8	132.8	131.8	128.8
9	157.4	152.6	144.9
10	53.6	45.3	42.3
11	215.6	184.9	181.6
12	196.7	193.0	198.3
13	155.4	159.0	154.6

**Closed form expression;* ***Bridge Sampling estimate*

表2：多元方差分析模型（M_3 或 M_4 或 M_6）相对于正态分布模型（M_1 或 M_2 或 M_5）的对数贝叶斯因子。

从这些结果可以看出，对数贝叶斯因子$_{4,1}$，对数贝叶斯因子$_{5,2}$ 和对数贝叶斯因子$_{6,3}$ 均明显高于 5。这些较高的贝叶斯因子数值表明，有非常强的证据（Kass & Raftery 1995年）支持 MANOVA 构式。因此，结合现有数据集可以得出结论：对于任何一种先验设置，以字符为预测变量的贝叶斯多元方差分析在书写者层面的数据拟合效果都更好。

共轭与非共轭。接下来，我们比较各模型的共轭版本与非共轭版本，具体考察正态-逆威沙特层次先验，以及正态-对数正态-LKJ先验与共轭版本之间的差异。随后计算贝叶斯因子 $\text{BF}_{1,2}$，$\text{BF}_{1,3}$，$\text{BF}_{2,3}$，$\text{BF}_{4,5}$，$\text{BF}_{4,6}$ 和 $\text{BF}_{5,6}$ ，以比较采用不同先验设置（共轭、正态-逆威沙特层次先验、正态-对数正态-LKJ先验）的两层随机效应模型与多元方差分析模型。如第 4节所述，手写文献强调，对书写者内部和书写者间变异进行建模十分重要。因此，将使用现有数据集评估结果是否与这一理论一致。为此，我们比较采用正态-逆威沙特先验的正态模型和多元方差分析模型的共轭版本与非共轭版本。表 3列出了按书写者评估的贝叶斯因子。结果显示，对于正态模型，9 名书写者中的 13 名的贝叶斯因子支持非共轭先验（因此也说明有必要考虑书写者间变异）

在 13 名书写者中，有 9 名的结果支持非共轭先验。另一方面，对于采用正态-逆威沙特先验设置（共轭和非共轭）的多元方差分析模型，贝叶斯因子并未明显倾向于非共轭先验；仅约半数书写者（13 名中的 7 名）支持该先验。因此，这一问题可能需要进一步研究。

逆威沙特与对数正态-LKJ先验的比较。最后，我们比较模型 M_1、 M_2、 M_4、 M_5（基于正态-逆威沙特先验）与模型 M_3 和 M_6（其对应的正态-对数正态-LKJ先验模型）。表4 列出了每位书写者的贝叶斯因子 $BF_{2,3}$ 和 $BF_{5,6}$。根据表中结果，在贝叶斯多元方差分析下，13位书写者中有9位的贝叶斯因子支持正态-对数正态-LKJ先验，而非对应的正态-逆威沙特先验。对于贝叶斯正态层次模型，证据呈现出更明显的一致性：几乎所有书写者（13位中有12位）都倾向于正态-对数正态-LKJ先验。附录 F中的表F.1 也列出了共轭正态-逆威沙特先验与正态-对数正态-LKJ先验的类似结果，包括各书写者对应的估计贝叶斯因子 $BF_{1,3}$ 和 $BF_{4,6}$。

书写者	正态-逆威沙特					
	正态分布 对数 $BF^*_{1,2}$			多元方差分析 对数 $BF^*_{4,5}$		
	数值	Sign	解释	数值	Sign	解释
1	0.14	+	单纯提及	1.60	+	显著
2	-4.92	-	强	-3.58	-	强
3	-3.77	-	强	-6.02	-	极端
4	-2.44	-	显著	-2.72	-	显著
5	1.85	+	显著	3.99	+	强
6	-2.13	-	显著	2.44	+	显著
7	-2.52	-	显著	-3.05	-	强
8	-3.68	-	强	-2.66	-	显著
9	-0.17	-	单纯提及	4.69	+	强
10	6.30	+	极端	14.62	+	极端
11	3.33	+	强	34.05	+	极端
12	-5.45	-	极端	-1.72	-	显著
13	-12.54	-	极端	-16.09	-	极端

** Bridge Sampling estimate*

表3：比较贝叶斯正态模型和多元方差分析模型中正态-逆威沙特先验的共轭方法与非共轭方法的对数贝叶斯因子（按书写者划分）。

书写者	正态分布 逆威沙特与对数正态-LKJ 对数 $BF^*_{2,3}$			多元方差分析 逆威沙特与对数正态-LKJ 对数 $BF^*_{5,6}$		
	数值	Sign	解释	数值	Sign	解释
1	-17.65	-	极端	-15.10	-	极端
2	-28.13	-	极端	-19.47	-	极端
3	-34.46	-	极端	-31.75	-	极端
4	-48.50	-	极端	-45.21	-	极端
5	5.39	+	极端	6.34	+	极端
6	-17.67	-	极端	-20.14	-	极端
7	-3.80	-	强	1.47	+	显著
8	-2.00	-	显著	1.00	+	单纯提及
9	-2.94	-	显著	4.69	+	强
10	-23.91	-	极端	-20.98	-	极端
11	-37.50	-	极端	-34.19	-	极端
12	-13.76	-	极端	-19.12	-	极端
13	-15.49	-	极端	-11.08	-	极端

** Bridge Sampling estimate*

表4：比较贝叶斯正态模型和多元方差分析模型的作者内协方差先验设置（a）逆威沙特或（b）对数正态-LKJ先验的对数贝叶斯因子（按书写者划分）。

5.2 模型准确性

本节通过模拟研究来评估所考察模型的性能。研究关注的命题是：某位个人是否为待检验手稿的书写者；与之相对的备择命题是，书写者为未知个人（参见第3节）。为此，我们将比较这些模型，考察它们能否给出支持正确命题的贝叶斯因子值。

为评估所提出模型在命题 H_1 （同一书写者情形）成立时的性能，我们生成了多种模拟案例研究，分别将同一书写者的字符测量值作为检材 $\boldsymbol{y}_1$ 和对照 $\boldsymbol{y}_2$ 数据。因此，将可用数据库中每位书写者的数据分为两部分，一部分用作检材数据，另一部分用作对照数据。用作检验材料的数据比例 π_{split} 在 0.35 和 0.65 之间随机取值，即 $\pi_{split} \in (0.35,0.65)$，其余数据则作为对照材料。对于 13 位可用书写者中的每一位，该过程重复 100 次（使用不同的随机划分），共得到 1300 个同一书写者比较。其余书写者的数据作为背景数据$\boldsymbol{X}$ ，用于为模型提取先验参数；具体而言，我们使用除待分析书写者之外的所有书写者的数据。由此可以评估假阴性率，即在贝叶斯因子应大于 1 时却小于 1 的案例所占的百分比。

比较不同书写者的材料时，也采用了类似的程序，即 H_2 为真。对于每一对书写者，将第一位书写者的测量值视为检材数据 $\boldsymbol{y}_1$ ，将第二位书写者的测量值视为对照数据 $\boldsymbol{y}_2$，由此生成案例研究。为考虑样本量的影响，采用了与针对 H_1 的模拟程序相同的方法。因此，对于每一对书写者，只考虑从每位书写者的数据中随机抽取的一个子集。具体而言，从每对中的第一位个人（即检验文件的书写者）随机抽取比例为$\pi_{split} \in (0.35,0.65)$的测量值；而第二位个人的测量值中，比例为 $(1-\pi_{split})$ 的部分则作为对照数据。对 78 对书写者中的每一对重复这一过程 100 次，共得到 7800 次不同的比较。其余书写者的数据作为背景数据 $\boldsymbol{X}$ ，用于确定模型的先验参数；也就是说，我们使用除正在分析的两位书写者之外所有书写者的数据。在这些比较中，目标是评估假阳性率，其定义为贝叶斯因子本应小于 1、但实际大于 1 的案例所占百分比。

在所有模拟情境中，为检验竞争假设 3，所得到的贝叶斯因子对数（ H_1 和 H_2 ；见第4.3.2节）分别列于图4a 和4b 中，对应同一书写者和不同书写者的比较。这些图中，从左数的前五个面板给出了字符 a、d、o和q 以及所有字符整体的贝叶斯正态模型（M_1、 M_2、 M_3）的贝叶斯因子。最后一个面板给出了使用所有字符的多元方差分析模型（M_4、 M_5、 M_6）的贝叶斯因子。在每个面板中，最左侧的第一个箱线图展示了使用共轭先验解析计算所得的贝叶斯因子；箱线图 2 和 3 则分别展示了使用分层正态-逆 Wishart 先验和正态-对数正态-LKJ先验所得的贝叶斯因子，其中边际似然通过附录E中所述的桥接抽样进行估计。

从这些箱线图可以明显看出，无论采用哪种统计建模方法，几乎所有贝叶斯因子都支持正确的假设。具体而言，在第一种情境（同一书写者）中，几乎所有贝叶斯因子（对数尺度）都明显大于零；而在第二种情境（不同书写者）的大多数比较中，贝叶斯因子都小于零。尤其是在同一书写者的情况下（即在 H_1 下），对于字符 q ，NIW 共轭模型和 NIW 分层方法均仅出现一个假阴性案例；对于字符 d ，在正态-对数正态-LKJ先验下也仅出现一个假阴性案例；其余字符和方法均未出现假阴性。

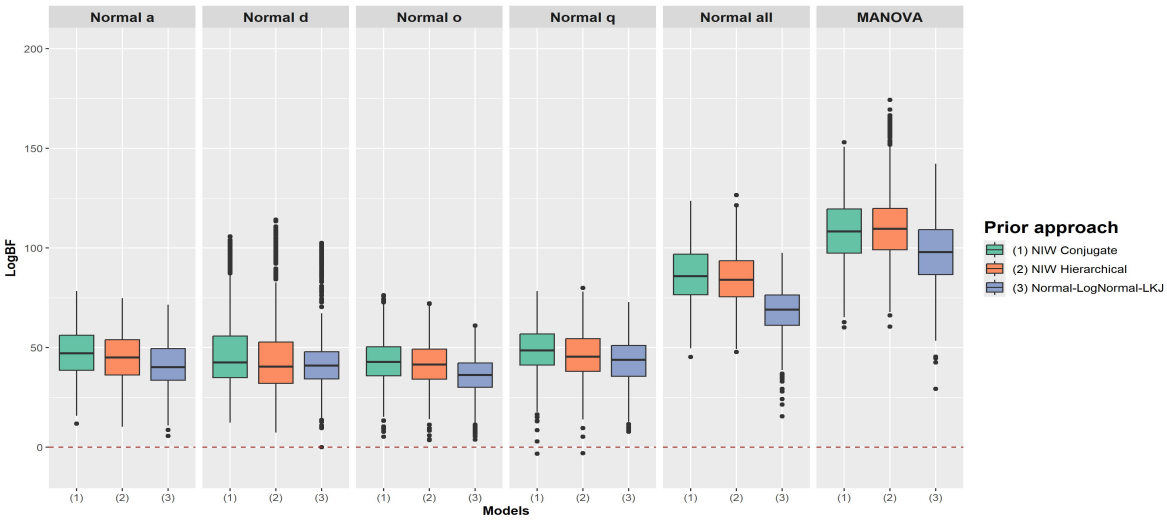
关于假设 H_2 下不同方法性能的评估，表 5列出了两种建模方法及不同先验设置下，在总计 7800 次不同书写者比较中出现的假阳性数量。具体而言，在总计 7800 次不同书写者比较中，假阳性总数最少为 47 次，最多为 589 次（即 0.6%–7.6%）。

对于单个字符，字符 a的假阳性率最高（NIW 共轭方法下高达 7.6%），而字符 q 的假阳性率始终最低。考虑所有字符时，假阳性率进一步下降：从字符 q（表现最佳的字符）的最低假阳性率 2.2%，降至正态-对数正态-LKJ 先验设置下的 0.9%。

关于先验设置，NIW 共轭模型和 NIW 分层模型对大多数字符产生了相近的假阳性率，不过层次版本的表现通常略好。与 NIW 方法相比，正态-对数正态-LKJ 设置使字符 a 和o的假阳性率更低，但使字符 d 和 q 的假阳性率更高。

在所有方法中，多元方差分析模型的假阳性比例最低，且始终低于 0.7%。这表明，在将所有字符综合考虑时，多元方差分析在跨书写者判别方面表现更优。此外，多元方差分析构式中不同先验设置之间仅观察到细微差异。

(a) 同一书写者比较



(b) 不同书写者比较

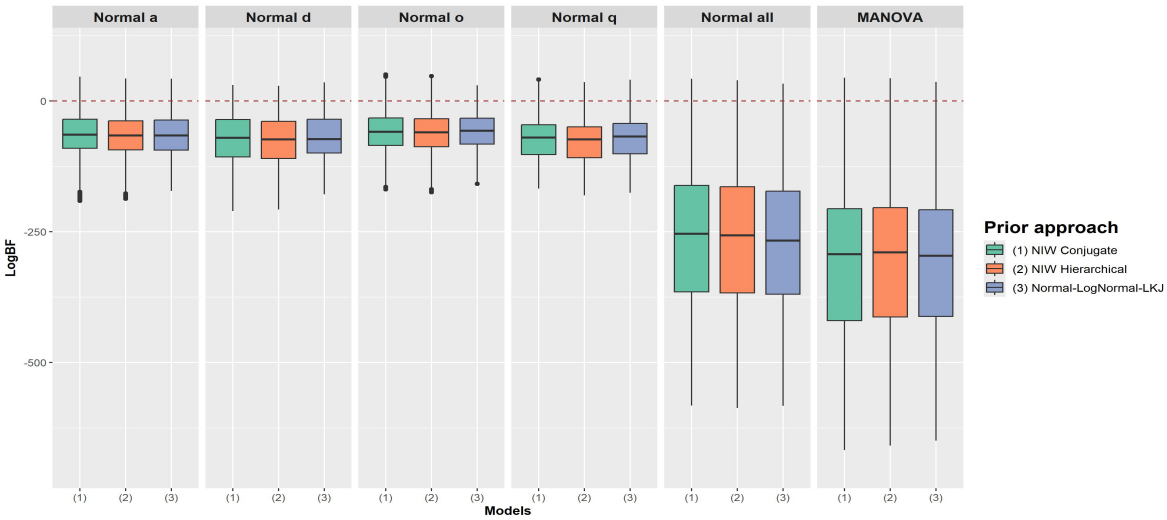


图4：采用第 4.1 节和 4.2节所述的数据建模方法，对同一书写者（a）和不同书写者（b）进行笔迹评估所得的对数贝叶斯因子（logBF）。

模型	字符	正态-逆威沙特		正态-对数正态-LKJ
		共轭	层次	
正态分布	a	589 (7.6%)	560 (7.2%)	545 (7.0%)
	d	491 (6.3%)	308 (3.9%)	589 (7.6%)
	o	410（5.3%）	363（4.7%）	301（3.9%）
	q	213（2.7%）	168（2.2%）	265 (3.4%)
	all	113 (1.4%)	86 (1.1%)	68 (0.9%)
多元方差分析	all	53 (0.7%)	51（0.7%）	47（0.6%）

表5：假阳性数 在7800组不同的书写者比较中（括号内为百分比）。

6 先验 elicitation 的敏感性分析

在贝叶斯层次建模中，先验 elicitation 至关重要，因为它会显著影响最终的贝叶斯因子数值。在本研究中，我们采用了一种主观先验方法，根据一组书写者的笔迹样本所产生的经验数据来确定先验分布；这组书写者与被比较的书写者不同。该方法明确纳入了有关研究中基于傅里叶分析所用笔迹特征属性的先验知识。当我们希望判定现有笔迹证据是否源自同一位书写者时，这种策略尤其有利。

为探究这一主观先验框架的影响，我们在第6.1节对先验 elicitation 框架进行敏感性分析。按照第5.2节所述的评估流程，从完整数据集$\boldsymbol{\mathcal{D}}$ 中随机选取一名书写者或一对书写者的数据，作为检材数据 $\boldsymbol{y}_1$ 和对照数据 $\boldsymbol{y}_2$，其余书写者则视为背景数据 $\boldsymbol{X}$。敏感性分析通过从每位背景书写者的数据中随机抽取子样本来构建先验分布，而不是使用整个背景数据集$\boldsymbol{X}$。具体而言，对每位背景书写者，均考虑从其数据中进行有放回抽样，抽取占其数据量50%的随机子样本。

此外，对于 M_{1-2} 和 M_{4-5}中 Wishart先验的自由度，我们将 $\nu = p + 2$设为默认的低信息量值；对于 M_3 和 M_6中的 LKJ先验，我们使用 $\eta = 1$。因此，在第6.2节敏感性分析的第二部分中，我们考察这些方差—协方差参数的不同先验数值会如何影响后验贝叶斯证据，并以贝叶斯因子进行衡量。

6.1 背景数据的子抽样

本节中，我们通过对用于先验 elicitation 的背景数据进行子抽样，研究先验设定的影响。用于评估两个假设（H_1 对比 H_2）的随机选定案例，其处理流程与第5.2节所述相同。具体而言，为了隔离先验 elicitation 的影响，我们在同一书写者实验和不同作者实验中均采用同一次随机数据划分，从而尽量减少数据分区可能带来的混杂效应。本分析聚焦于贝叶斯多元方差分析模型，该模型在第5.2节中被确定为效果最佳且推荐使用的方法。为评估贝叶斯因子及其所导出的决策的稳健性和稳定性，我们进行了 30 次有放回随机子抽样；每个评估案例中，每位书写者均使用 50% 的背景数据。进一步说明，本分析考虑三个不同的数据集：(i) 检材数据$\boldsymbol{y}_1$，(ii) 对照数据 $\boldsymbol{y}_2$，以及 (iii) 背景数据 $\boldsymbol{X}$；本节旨在评估改变第三个数据集所产生的影响。

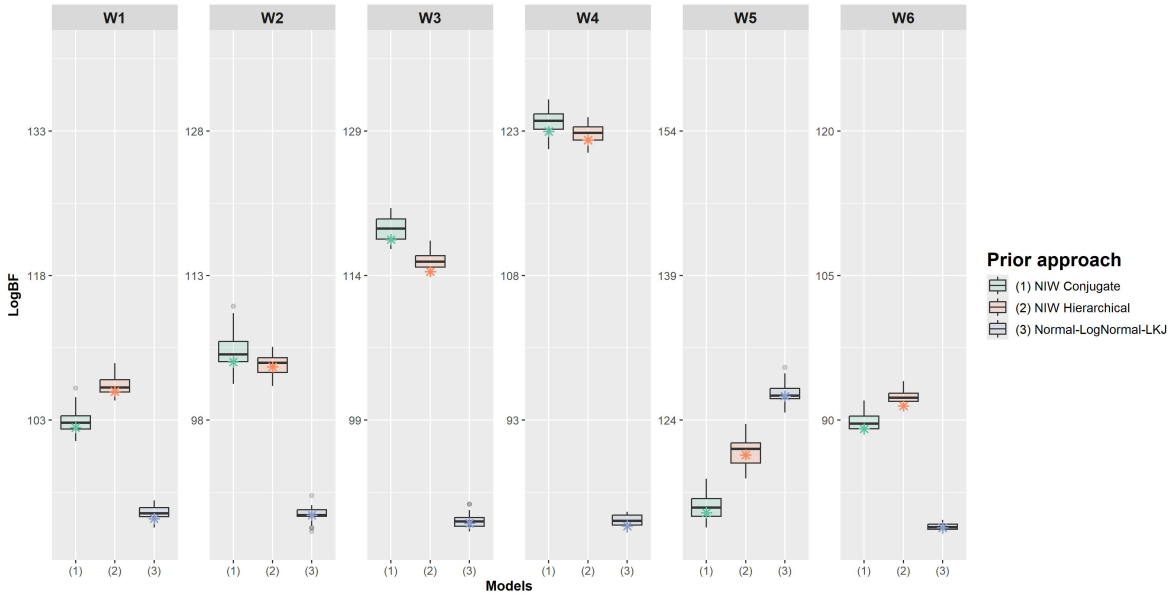
图 5 和 6 展示了在 30 种不同的 elicited先验下，手写体评估的对数贝叶斯因子（LogBF）数值，并比较了贝叶斯多元方差分析模型在三种先验设置（共轭、分层正态-逆 Wishart、正态-对数正态-LKJ）下的表现。图5 展示了同一书写者比较的结果，而图 6 展示了不同书写者之间的比较。图 6 中的星号表示使用基于完整背景数据得到的先验计算出的 LogBF 数值，可作为子抽样分布的参考点。这些参考点通常位于箱线图附近，表明我们提出的方法对于用于先验 elicitation 的

背景数据的变化相对稳健（无论是贝叶斯因子数值还是由此得出的判断），并且与基于完整数据集得到的贝叶斯因子接近。

总体而言，结果表明，LogBF 提供的支持不受所构建先验 elicitation 的影响，并且在几乎所有情况下都能正确识别真实假设。然而，仍可观察到一定程度的变异性，这在预期之中。例如，涉及不同书写者的比较（如 W8–W13、W7–W8）显示，共轭方法的 LogBF 数值存在显著变异性，尽管这些变异并未改变对 H_2的总体支持。这一模式表明，贝叶斯因子会随背景数据的选择而产生一定变异性；相比之下，同一书写者比较（如 W3、W4）表现出显著更高的稳定性。更多示例见附录 G。

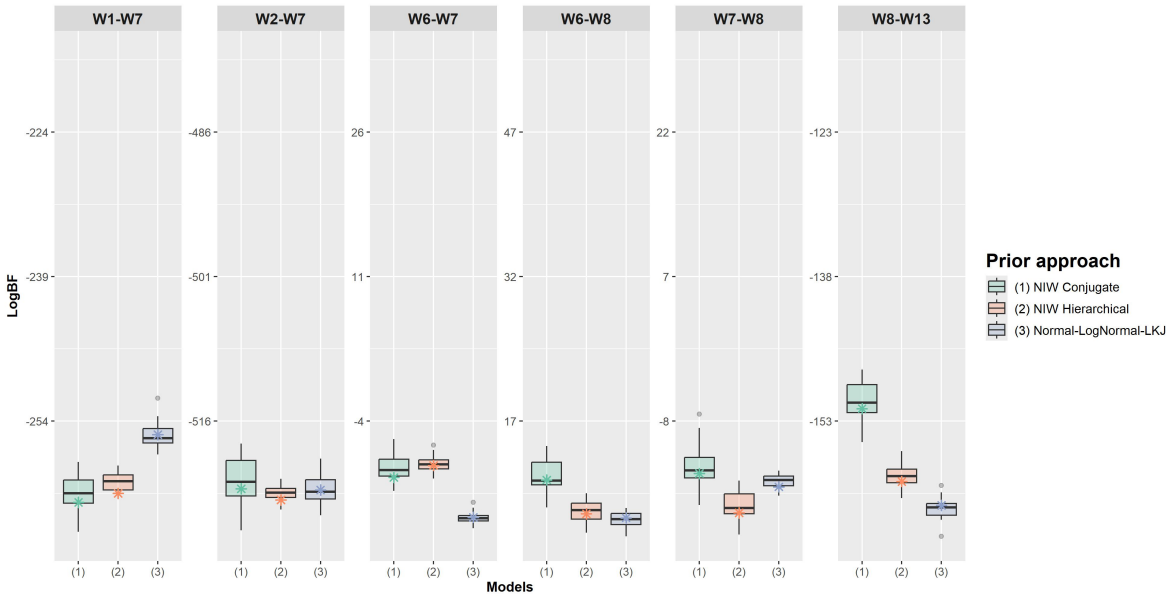
为评估这种变异性的影响，我们考察了四组相似度最高的书写者对（W5–W6、W6–W7、W6–W8 和 W7–W8）；这些书写者对是根据包含 $|\log BF| < 10$ 的案例选出的。对于其余书写者对，观察到的变异性似乎并未影响总体支持程度。为进一步说明这一点，我们按照第 5.2 节所述的步骤，从每一对书写者中随机选取数据 30 次，即更换受质疑数据集和对照数据集。随后，对于每一对书写者，从其余书写者的数据中进行有放回抽样，抽取其中 50% 的数据作为背景数据集。总计，该过程产生了 3600 次比较（案例研究）供分析。

表6 总结了在每种多元方差分析先验设定下，支持竞争假设的敏感性。在从不同背景数据子集得到的30个先验中，NIW 共轭先验方法出现支持不一致（即支持度随背景数据子样本而变化）的案例最多，共9例，占120例的7.5%。这9例的平均 logBF 范围为8.51个单位，而最宽的 logBF 子抽样区间（最差情境）的取值范围为（-6.24，4.79）。相比之下，正态-对数正态-LKJ 先验设置中的此类案例较少（2例，占1.67%），平均差异也较小，表明其推断在先验变化下相对更稳定。NIW 层级先验的表现介于这两种设置之间。



*indicates the 对数贝叶斯因子 using the complete background dataset for each case.

图 5：贝叶斯多元方差分析方法下，同一书写者情景的笔迹评估在不同背景数据子样本中的对数贝叶斯因子（logBF）箱线图。



*indicates the对数 贝叶斯因子 using the complete background dataset for each case.

图6：贝叶斯多元方差分析方法下，不同书写者情景在不同背景数据子样本上的手写体评估对数贝叶斯因子（对数贝叶斯因子）的箱线图。

Inconsistent Cases*			
		LogBF Summaries	
先验设置	**支持度偏移** (%)**	**平均范围†**	**最坏情况区间‡**
NIW 共轭模型	9 (7.50)	8.51	(−6.24, 4.79)
NIW 分层模型	5 (4.17)	3.88	(−2.69, 1。48) (−1.49<风格 id='2'>, </风格>3<风格 id='4'>。</风格>05)
正态-对数正态-LKJ	2 (1.67)	3.42	（最小值 最大值）

* *Inconsistent cases based on 30 random subsets of four compared writer pairs and 30 random 50% subsamples of background writers' data (4 × 30 × 30 = 3600).*
** *Percentage of inconsistent-support cases across 30 random subsets of compared writers (4 × 30=120).*
† *Average logBF range across inconsistent cases, range is computed over 30 background-data subsamples.*
‡ *LogBF interval , of the case with highest range of logBFs.*

表 6：针对最相似书写者对的敏感性分析（W5–W6、W6–W7、W6–W8 和 W7–W8），比较不同先验设置。

6.2 逆Wishart先验自由度及 LKJ分布参数的设定

在贝叶斯层次建模中，协方差矩阵先验的选择至关重要，因为它会直接影响最终的贝叶斯因子数值，正如第5节所示。

逆Wishart先验是对协方差矩阵进行建模时的传统选择。它是多元正态分布的共轭先验，能够简化贝叶斯推断的计算过程。逆Wishart先验由尺度矩阵和自由度两个参数确定，它们控制先验分布在尺度矩阵周围的集中程度。

另一方面，采用 LKJ 方法的协方差分解为相关矩阵建模提供了更灵活的替代方案。LKJ先验由形状参数η进行参数化，该参数控制先验在单位矩阵附近的集中程度。当 $\eta = 1$时，LKJ先验在相关矩阵空间上服从均匀分布，因此属于无信息先验。 η 的数值越高，先验在单位矩阵附近的集中程度越强，反映出对较弱相关性的先验信念。在本节中，我们将针对使用逆Wishart先验的方法中不同自由度的数值，以及 LKJ先验方法中η的不同数值，进行敏感性分析

采用逆Wishart先验的方法中的自由度，以及采用 LKJ先验方法时 η 的不同数值。

在使用逆Wishart先验的实验中，自由度 ν 从允许的最小值 11* 变化至 20 到 50，步长为 10，即 $\nu \in \{11, 20, 30, 40, 50\}$。因此，针对 ν的每个取值，均计算了贝叶斯因子，用于比较来自不同书写者的所有字符，并考虑采用正态分布和多元方差分析模型的正态分布-逆Wishart先验设置（共轭和非共轭）。

结果表明，贝叶斯因子对自由度的选择相当敏感，自由度会逐步影响其数值；详见附录G.2。这是因为自由度越高， $\boldsymbol{W}$ 的先验就越集中在 $\mathbb{R}^p$ (Gaborini 2021年)中较小的区域内。因此，随着 ν增大，方差-协方差矩阵的变异性会降低。因此，我们预计，当 ν 变得很大时，所得的对数贝叶斯因子将趋近于一个上限，并始终支持测量值来自同一书写者这一假设（即使测量值实际上来自不同的书写者）；更多详情请见附录G.2 。

现在，我们来考察采用 LKJ先验 方法的实验；参数 η 设定为从非信息性的数值 1 到信息性逐渐增强的数值 2、5、10 和 20。 η 数值越高，先验分布越集中于单位矩阵周围，这反映了傅里叶系数彼此不相关的假设（该假设由傅里叶分析所启发）。对于每个 η数值，我们使用所考虑的两种模型——贝叶斯正态模型和采用正态-对数正态-LKJ 先验设置的多元方差分析模型——计算不同书写者之间所有成对比较的贝叶斯因子。结果表明，贝叶斯因子对 η的选择相当敏感，尤其是当其数值大于 5 时，影响会愈发显著；相关讨论见附录G.3 。这种敏感性源于 η数值越高，意味着对变量之间不存在相关性的先验信念越强。

对两种方差-协方差矩阵建模方法（逆Wishart先验和 LKJ先验）的实证比较表明，对数正态-LKJ模型通常能提供更稳健的结果，尤其适用于贝叶斯多元方差分析模型。这体现了该模型因采用协方差分解而具有的灵活性。对数正态先验为方差-协方差矩阵对角线上的标准差提供了恰当的信息，而 LKJ 先验则控制有关相关结构的信息。根据我们的分析，我们建议使用对数正态 -LKJ 先验，并采用根据背景数据测量值推导出的信息性对数正态先验。采用对数正态-LKJ 先验设置的这一方法，为实施所提出的模型奠定了坚实基础。

7 结论与未来工作

本研究探讨了如何应对法证笔迹鉴定中的不确定性管理这一挑战，该不确定性会影响法律情境下的推断和决策过程。本研究开展了一项实验研究，旨在促进：（a）复杂多变量数据的管理；（b）整合多个信息来源（在这一特定情境中，是不同类型的字符），以评估联合概率值；以及（c）确定最优建模方法。所提出的贝叶斯统计框架提供了一种全面的定量工具，可用于揭示笔迹数据中有效且有信息价值的模式，有望对手写文档分析的多个领域产生显著影响。

所提出的贝叶斯概率框架已应用于一系列基于真实笔迹数据构建的法证案例研究。具体而言，本研究将Bozza 等人(2008年)提出的用于笔迹检验的两层随机效应模型与贝叶斯分层多元方差分析模型进行了比较；后者在这一情境中的应用尚属首次。本研究探讨了以下五个问题：

（i）哪个模型对数据的拟合更好？（ii）哪个模型具有更好的判别性能（假阳性率和假阴性率更低）？（iii）对书写者间变异进行建模是否会对最终结果产生影响?

*也就是说， $p + 2 = 11$ 适用于 $p = 9$。

（iv）就识别涉案手写稿件书写者这一最终目标而言，哪个字符提供的信息最多？（v）所实施模型的结果在不同背景数据和先验设定下是否仍然稳健？

为回答第一个研究问题，我们考虑并比较了六种不同的贝叶斯模型。这些模型基于两种似然结构：(a) 多元正态模型（参见第 4.1 节），以及 (b) 考虑字符层级变异的多元方差分析模型（参见第 4.2 节）。针对每种似然结构，我们考察了三种不同的先验形式，从而得到不同的贝叶斯模型：(i) 共轭正态-逆威沙特先验，(ii) 层次正态-逆威沙特先验，以及 (iii) 正态-对数正态-LKJ先验设定。表 2 所示结果表明，证据强烈支持贝叶斯多元方差分析模型。将书写者间变异纳入正态-逆威沙特（NIW）先验形式似乎产生了积极影响：如表 3 所示，对于 13 位书写者中的 9 位，贝叶斯因子支持这种先验设置（因而不支持共轭模型）。对于采用正态-逆威沙特先验的多元方差分析模型，共轭先验设置与层次先验设置之间没有明显优势，因为结果同时支持这两种方法（见表 3）。最后，我们比较了正态-对数正态-LKJ先验与正态-逆威沙特先验这两种先验设置。表 4 所示贝叶斯因子表明，在两种统计模型中，对于大多数书写者，正态-对数正态-LKJ先验优于正态-逆威沙特先验。

为回答研究问题 (ii)–(iv)，我们通过从现有数据中随机选取测量值，分别作为待检材料和对照材料，模拟了一系列法证案例场景。关于第二个研究问题，当联合分析所有字符时，贝叶斯多元方差分析模型的表现优于贝叶斯正态模型，其假阳性率约为 0.6%，而正态模型为 0.9%。这意味着，与贝叶斯正态模型相比，假阳性率降低了约 33.3%。这并不令人意外，因为后者模型未考虑字符层级变异。此外，在贝叶斯多元方差分析框架中采用正态-对数正态-LKJ先验，与采用正态-逆威沙特先验相比，假阳性减少了 0.1%。这一发现与模型拟合的结果一致。然而，使用贝叶斯正态模型分别分析每个字符时，对于所考察的四个字符中的两个，正态-逆威沙特先验表现更优。对于手写材料出自同一书写者时（即在H_1条件下）能否支持正确假设，两种模型均始终正确分类所有相关案例；唯一例外是字符 q 的 NIW 模型，其误分类比例极低（$< 0.1\%$；在生成的 1200 个情境中仅有一个）。

关于研究问题（iii），第 5.2 节呈现的结果表明，与共轭方法相比，能够考虑书写者间变异的层次先验设置，采用正态-对数正态-LKJ先验或正态-逆威沙特先验时，整体表现均更好（见表 5）。相比之下，多元方差分析模型在正态-对数正态-LKJ先验设置下表现更好。这一发现与第 5.2节讨论的其他实验结果一致；这些实验检验了所提出模型在手写体判别中的准确性，也与针对第一个研究问题开展的模型比较结果相符（见第 5.1节）。同样，采用正态-逆威沙特先验的贝叶斯正态模型，其共轭版本与非共轭版本之间也呈现出一致的性能下降。不过，这些结果仍需进一步研究和验证；理想情况下，应在多种不同的手写体数据集上实施所提出的方法，并纳入更多书写者。

为回答研究问题（iv）所进行的分析揭示了，在所分析的四个字符中，哪个字符对于区分正在考察的两个假设最具信息量。研究始终发现，字符q 产生误导性证据的情况较少，假阳性率为 2.2%。另一方面，字符a的区分能力最弱，使用层级正态-逆威沙特先验时，假阳性率高达 7.2%。最后，字符d 和o的假阳性率分别为 3.9% 和 4.7%。

我们进行了敏感性分析，以评估结果对先验 elicitation 和设定的稳健性，从而回答研究问题（v）。通过对每个评估案例中 50% 的背景作者数据进行子抽样，我们发现贝叶斯因子的变化程度不足以决定性地改变模型支持度。特别是在采用

正态-对数正态-LKJ先验时，在 120 次比较中，待检验假设的支持度仅在两种情况下发生逆转。这两种情况涉及笔迹文本最相似的书写者。这一发现凸显了该模型在对给定假设提供一致的支持或反对方面的稳健性（见第 6节）。此外，逆威沙特分布的自由度参数以及 LKJ 分布的 η 参数，对所得贝叶斯因子的表现起着至关重要的作用。贝叶斯因子对这些先验参数的设定尤为敏感。具体而言，较高的自由度数值和较大的 η 会使先验变得越来越有信息性，从而产生更大的贝叶斯因子，并因此对（错误的）假设 H_1提供更强的支持。鉴于这种敏感性，在设定先验自由度（用于逆威沙特分布）以及先验参数 η（用于 LKJ 分布）时，必须慎重考虑。建议采用弱信息先验。具体做法是，将逆威沙特分布的自由度设为其最小值 $p+2$，并为 LKJ 分布指定 $\eta=1$。

最后，还有一个未解决问题，即所得贝叶斯因子的尺度。本研究中的贝叶斯因子大小与此前相关研究（Bozza等（2008））的结果一致，但其稳健性仍可能受到质疑，而这种质疑无法通过实证加以证明。遗传（DNA）证据领域也讨论过类似情况（例如参见Hopwood 等人（2012年））。不过，这种量级的数值可能难以论证，尤其是在数据库如此有限的情况下。

未来的研究可考虑将环形字符的相邻字符纳入分析，这或许能提供更多信息并提高结果的准确性。此类数据可以通过双向多变量方差分析模型进行建模，其中相邻字符和环形字符作为虚拟变量。

8 致谢

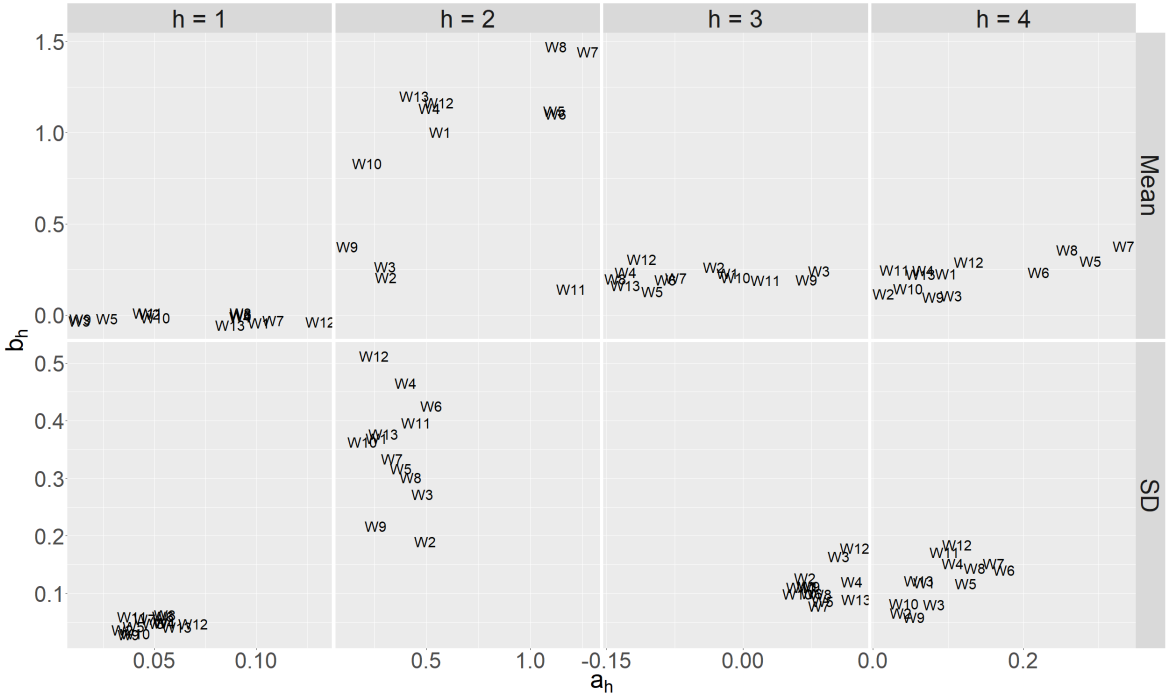
我们感谢瑞士国家科学基金会通过编号为 100011 204554/1（《法证推断与决策的剖析》）的项目慷慨资助，使我们得以开展这项研究。我们也感谢 Raymond Marquis博士（洛桑大学）为我们提供数据，并在这项 omfatt的笔迹分析中给予支持和咨询。

附录

傅里叶系数与表面大小的描述性分析

本节将对用于实施所提方法的样本主要特征进行基本的描述性分析。如图所示，傅里叶系数可用笛卡尔坐标表示A.1，图中展示了每个谐波下，各书写者以及所有字符的系数对（a_h，b_h）的均值和标准差。这种图形表示有助于了解各谐波的区分能力。具体而言，用于测量椭圆度的第二谐波似乎最适合用于区分。关于现有字符特征的分析性描述，读者可参阅图A.3。

最后，我们通过量化马氏距离来衡量所有字符中字书写者两两之间的变异性。两位书写者之间的马氏距离，是通过测量一位书写者样本的特征向量与另一位书写者样本的均值和协方差结构之间的距离来计算的。从图 A.2（其中以平方根表示马氏距离）可以清楚地看出，虽然某些书写者组合的距离很小（例如书写者6、7和8），但其他组合的距离较大（例如书写者10和13），这证实了这些组合具有更显著的个体特征，因此更容易区分。



图A.1：每个谐波、每位书写者以及所有字符中，每对傅里叶系数（a_h，b_h）的均值和标准差。

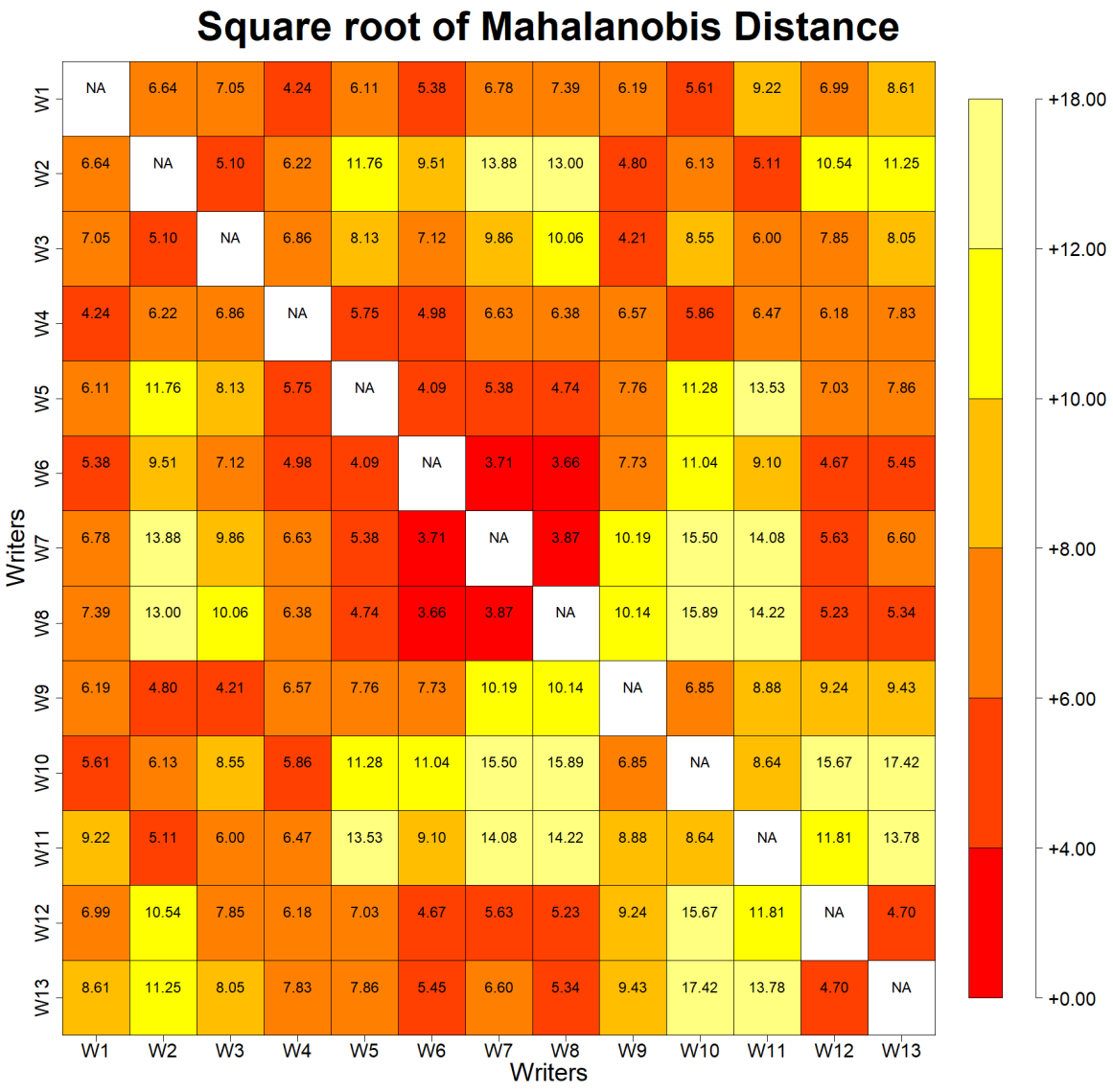
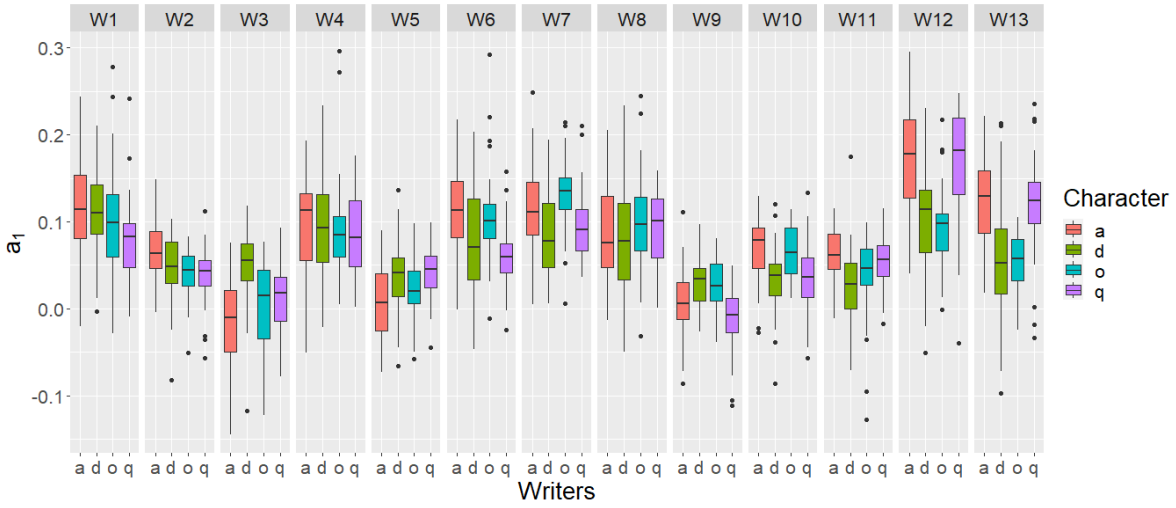


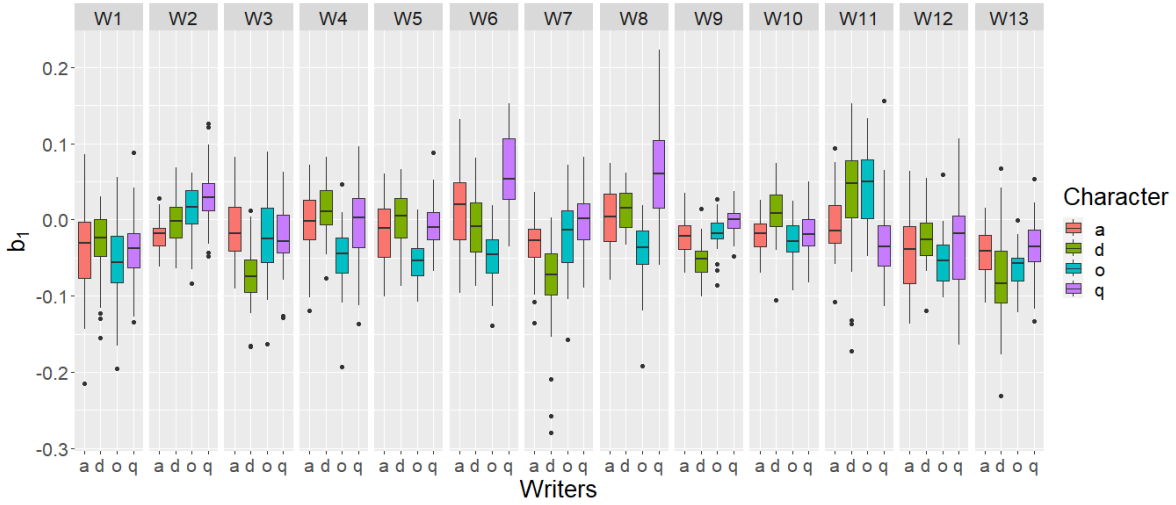
图 A.2：如第2节所述，使用所有字符的傅里叶系数和表面大小计算的第2.1节书写者之间马氏距离的平方根。



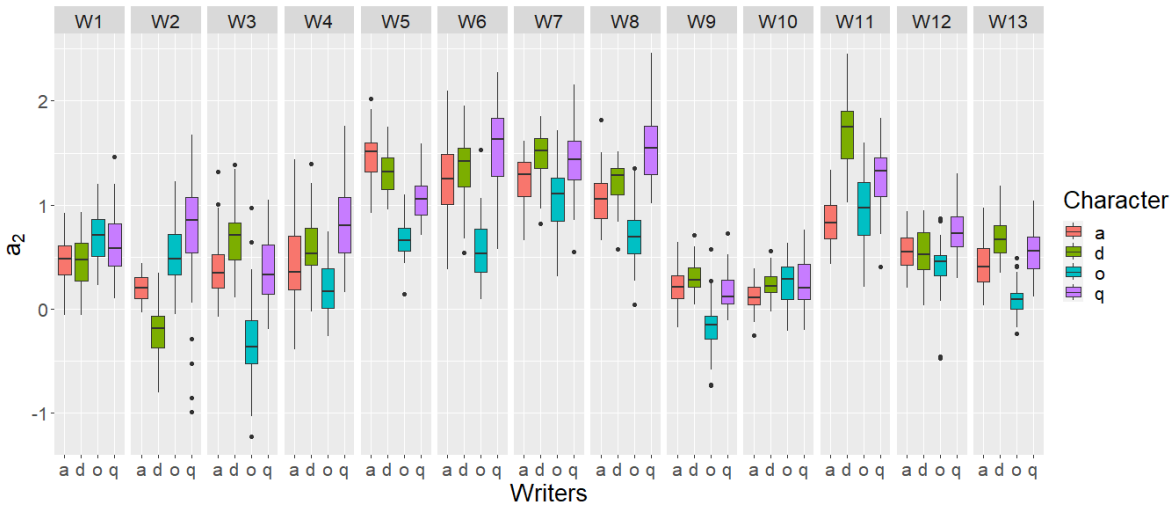
(a) 每个书写者各环形字符表面大小（以 cm^2 为单位）的箱线图。



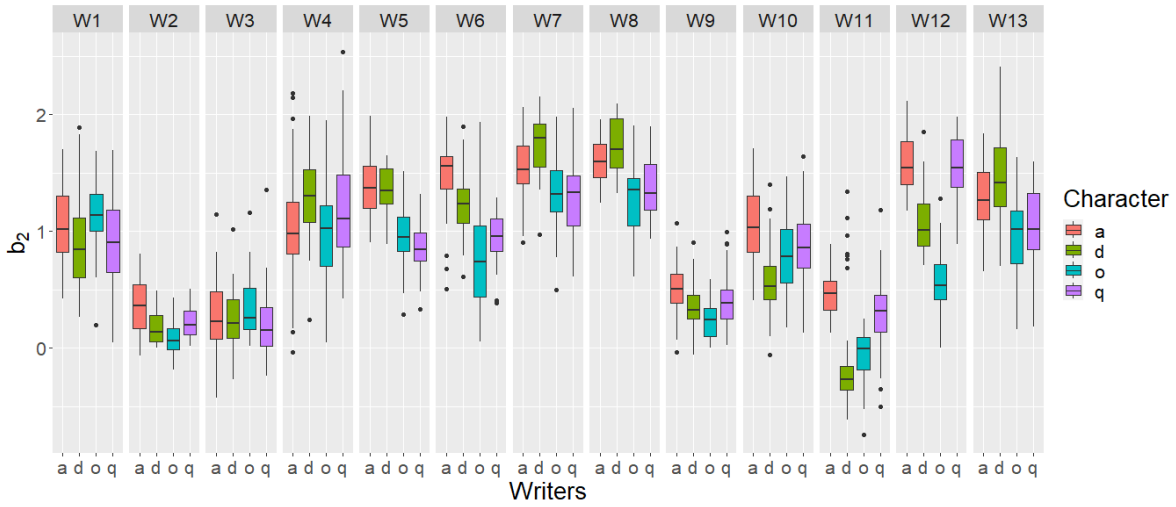
(b) 每个环形字符和书写者的傅里叶系数 a_1 箱线图。



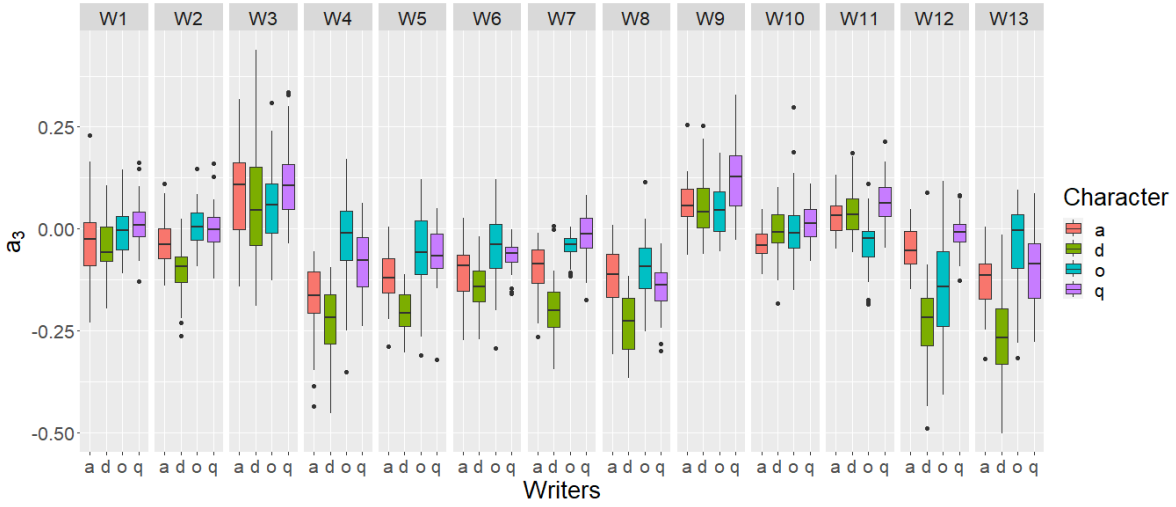
(c) 每个环形字符和书写者的傅里叶系数 b_1 箱线图。



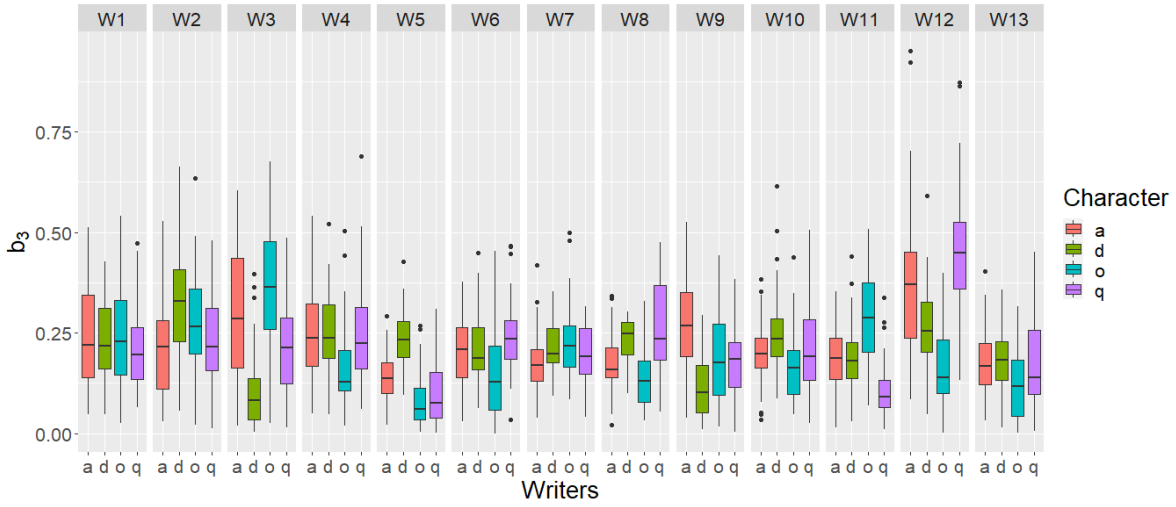
(d) 每个环形字符和书写者的傅里叶系数 a_2 箱线图。



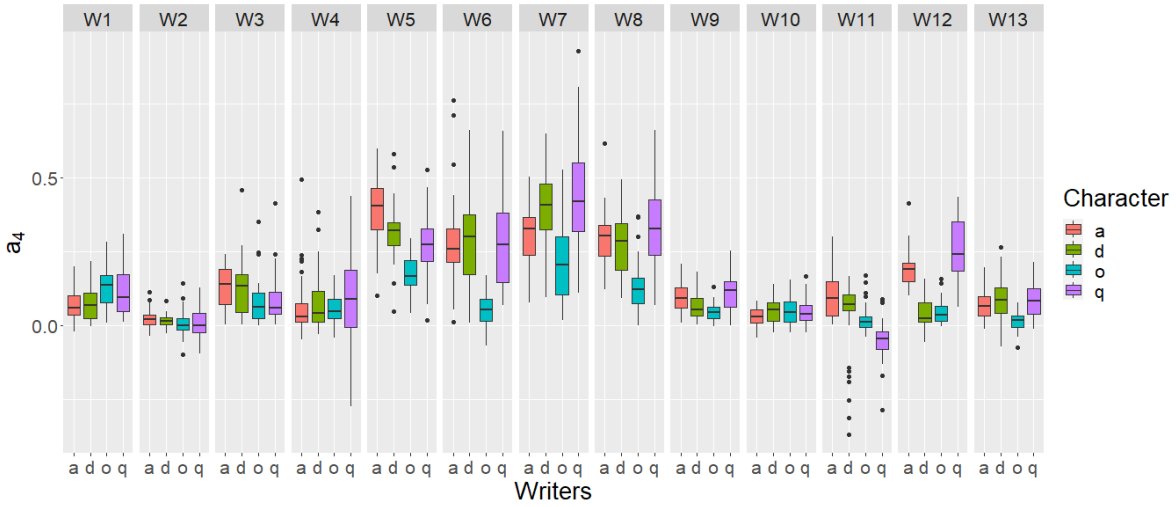
(e) 每个环形字符和书写者的傅里叶系数 b_2 箱线图。



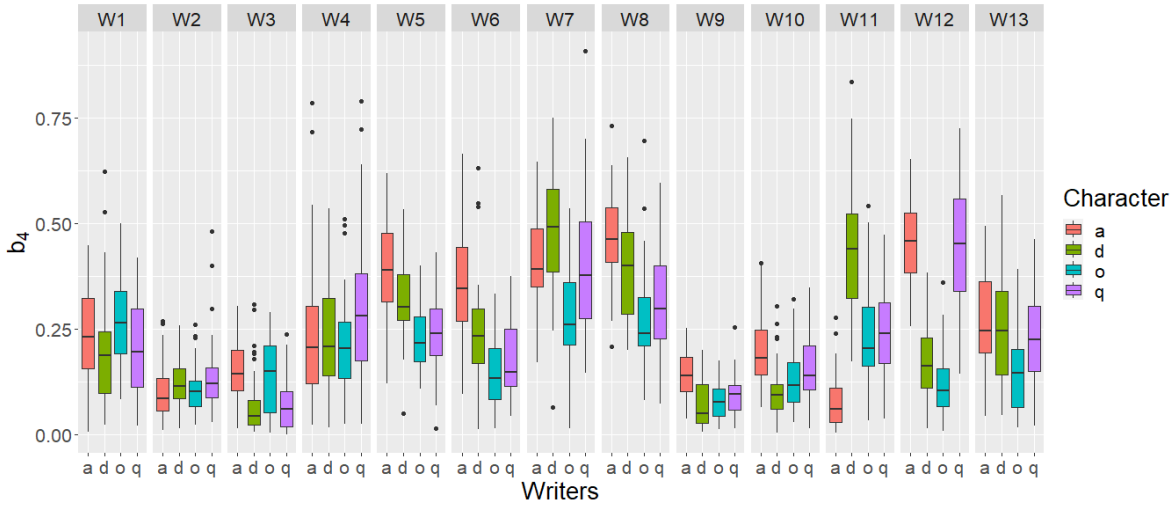
(f) 每个环形字符和书写者的傅里叶系数 a_3 箱线图。



(g) 每个环形字符和每位书写者的 b_3 傅里叶系数分组箱线图。



(h) 每个环形字符和书写者的傅里叶系数箱线图 a_4 。



(i) 每个环形字符和书写者的傅里叶系数箱线图 b_4 。

图 A.3：每个环形字符和每位书写者的表面大小及前四对傅里叶系数的分组箱线图。表面大小为 cm^2

B 贝叶斯因子的假设

本研究用于笔迹评估的贝叶斯因子（BF）定义如下：

$$BF = \frac{m(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid H_1)}{m(\boldsymbol{y}_1, \boldsymbol{y}_2 \mid H_2)} = \frac{m(\boldsymbol{y}_1 \mid \boldsymbol{y}_2, H_1) m(\boldsymbol{y}_2 \mid H_1)}{m(\boldsymbol{y}_1 \mid \boldsymbol{y}_2, H_2) m(\boldsymbol{y}_2 \mid H_2)} \tag{16}$$

该表达式还可以进一步分解为：

$$BF = \frac{m(\boldsymbol{y}_1 \mid \boldsymbol{y}_2, H_1)}{m(\boldsymbol{y}_1 \mid H_2)} \times \frac{m(\boldsymbol{y}_2 \mid H_1)}{m(\boldsymbol{y}_2 \mid H_2)} \tag{17}$$

假设当 H_2 成立时， $\boldsymbol{y}_1$ 的似然不依赖于 $\boldsymbol{y}_2$ ，则贝叶斯因子可简化为：

$$BF = \frac{m(\boldsymbol{y}_1 \mid \boldsymbol{y}_2, H_1)}{m(\boldsymbol{y}_1 \mid H_2)} \tag{18}$$

此外，如果 $\boldsymbol{y}_2$ 的似然（即对照测量）独立于 H_1 或 H_2 是否成立，则可得：

$$m(\boldsymbol{y}_2 \mid H_1) = m(\boldsymbol{y}_2 \mid H_2) \tag{19}$$

边际似然 m（$\boldsymbol{y}_1 \mid \boldsymbol{y}_2$， H_1）可表示为：

$$m(\boldsymbol{y}_1 \mid \boldsymbol{y}_2, H_1) = \int_\theta f(\boldsymbol{y}_1 \mid \theta) f(\theta \mid \boldsymbol{y}_2, H_1) d\theta \tag{20}$$

其中 $f(\theta \mid \boldsymbol{y}_2, H_1)$ 表示在 $\boldsymbol{y}_2$ 条件下、基于 H_1 的 θ 后验分布（为简便起见，本节中将模型参数记作 θ ；本研究中所有考虑的模型参数见第4节）。

或者，利用贝叶斯定理，可将其改写为：

$$m(\boldsymbol{y}_1 \mid \boldsymbol{y}_2, H_1) = \frac{\int_\theta f(\boldsymbol{y}_1 \mid \theta) f(\boldsymbol{y}_2 \mid \theta) f(\theta \mid H_1) d\theta}{m(\boldsymbol{y}_2 \mid H_1)} \tag{21}$$

H_2 下的边际似然类似地表示为：

$$m(\boldsymbol{y}_1 \mid H_2) = \int_\theta f(\boldsymbol{y}_1 \mid \theta) f(\theta \mid H_2) d\theta \tag{22}$$

结合这些结果，并注意到 $m(\boldsymbol{y}_2 \mid H_1) = m(\boldsymbol{y}_2 \mid H_2)$，贝叶斯因子可表示为：

$$BF = \frac{m(\boldsymbol{y}_1 \mid \boldsymbol{y}_2, H_1)}{m(\boldsymbol{y}_1 \mid H_2)} = \frac{\int_\theta f(\boldsymbol{y}_1 \mid \theta) f(\boldsymbol{y}_2 \mid \theta) f(\theta \mid H_1) d\theta}{\int_\theta f(\boldsymbol{y}_1 \mid \theta) f(\theta \mid H_2) d\theta \int_\theta f(\boldsymbol{y}_2 \mid \theta) f(\theta \mid H_2) d\theta} \tag{23}$$

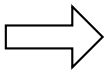
请注意，所实现的贝叶斯因子假设在 H_2 条件下，待检材料与对照材料相互独立。此假设意味着未考虑可能存在的伪装行为。本文正文中给出了这一表述。

C 虚拟变量

对于每位书写者，字符指示变量都会转换为一组虚拟变量。出于回归模型可识别性的考虑，通常会从方程中省略一个虚拟变量，这一做法最早由Suits (1957)提出。此外，与被省略的虚拟变量对应的类别将作为参照组，用于比较其他所有水平。在我们的情形中，字符a被选为参照类别；见表C.1和C.2。

指标	傅里叶系数
a	···
d	···
o	···
q	···

表 C.1：处理后的数据



a	d	o	q	傅里叶系数
1	0	0	0	···
1	1	0	0	···
1	0	1	0	···
1	0	0	1	···

表 C.2：回归数据

D 先验的确定

D.1 贝叶斯正态模型的先验参数确定

对于正态-逆威沙特模型，可以根据背景数据集确定先验参数。书写者 i 对字符 ℓ 的先验均值（记为 $\boldsymbol{\theta}_{i\ell}$）由相应背景数据在所有重复次数上的样本均值给出，即

$$\widehat{\boldsymbol{\theta}}_{i\ell} = \frac{\sum_{j=1}^{n_{i\ell}} \boldsymbol{X}_{i\ell j}}{n_{i\ell}}, \tag{24}$$

其中 $i = 1、\ldots、m$, $\ell = 1、\ldots、L$, $\boldsymbol{\theta}_{i\ell}$和 $\boldsymbol{X}_{i\ell j}$是长度为 p 的向量；后一个向量的元素为背景傅里叶系数以及书写者 i 在所有重复次数 $n_{i\ell}$中书写字符 ℓ 时的表面大小。环形字符 ℓ 的总体均值记为 $\boldsymbol{\mu}_\ell$，其估计值为

$$\widehat{\boldsymbol{\mu}}_\ell = \sum_{i=1}^{m} \widehat{\boldsymbol{\theta}}_{i\ell} \frac{n_{i\ell}}{n_\ell} \tag{25}$$

其中 $n_\ell = \sum_{i=1}^{m} n_{i\ell}$。等价地，所有字符和书写者的总体均值可简单表示为 $\widehat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{m}\sum_{\ell=1}^{L} n_{i\ell}\widehat{\boldsymbol{\theta}}_{i\ell}$；其中 $n = \sum_{\ell=1}^{L} n_{\ell}$。

环形字符 ℓ 的书写者间协方差矩阵 $\boldsymbol{B}_\ell$ 通过将其设为相应背景数据的样本协方差来确定，即

$$\widehat{\boldsymbol{B}}_\ell = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{X}_{ij\ell} - \widehat{\boldsymbol{\mu}}_\ell)(\boldsymbol{X}_{ij\ell} - \widehat{\boldsymbol{\mu}}_\ell)'. \tag{26}$$

按照 R 和 Stan 中逆威沙特分布的参数化方式，其均值为 $E[\boldsymbol{W}_{i\ell}] = \boldsymbol{U}_\ell/(\nu - p - 1)$，其中 p是表示响应数据的每个环形字符所保留的傅里叶系数和表面大小的数量。因此，通过将 $\boldsymbol{U}_\ell$ 设为以下数值，可以确定参数 $\boldsymbol{U}_\ell$ （用于描述书写者内变异）：

$$\widehat{\boldsymbol{U}}_\ell = \widehat{\boldsymbol{W}}_{i\ell}(\nu - p - 1), \tag{27}$$

其中 $\boldsymbol{W}_{i\ell}$ 可通过以下方式估计：

$$\widehat{\boldsymbol{W}}_{i\ell} = \frac{1}{n-m}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\boldsymbol{X}_{ij\ell} - \widehat{\boldsymbol{\theta}}_{i\ell})(\boldsymbol{X}_{ij} - \widehat{\boldsymbol{\theta}}_{i\ell})' \ . \tag{28}$$

我们考虑以下数值： ν 使得 $\nu \geq p + 2$ 从而保证先验均值有定义；更多细节请参阅Press (2005年)。最后，先验参数 k_0 根据对所有书写者的背景数据在 (0,1) 区间内进行的网格搜索来选取。具体而言，我们选择使背景数据的边际似然最大的k_0 值。

对于正态-对数正态-LKJ 先验方法，正态先验的参数根据公式（25）和（26）中给出的表达式进行一致估计。随后，通过提取$\widehat{\ }$估计协方差矩阵的对角元素 $\boldsymbol{W}_{i\ell}$，我们得出对数正态分布的先验参数如下：

$$\hat{\upsilon}_\ell = \frac{1}{p}\sum_{\kappa=1}^{p}\log(\widehat{\boldsymbol{W}}_{i\ell\kappa\kappa}) \tag{29}$$

$$\hat{\sigma}_\ell = \frac{1}{p-1}\sum_{\kappa=1}^{p}\left(\log(\widehat{\boldsymbol{W}}_{i\ell\kappa\kappa}) - \hat{\upsilon}_\ell\right) \tag{30}$$

对于 η，LKJ分布的参数设为 1。

D.2 贝叶斯多元方差分析的先验参数估计

同样，对于贝叶斯多元方差分析模型，也使用现有的背景数据来确定先验参数。参考字符‘a’ $\boldsymbol{\mu}_1$ 和 $\boldsymbol{B}_1$ 的 $\boldsymbol{\theta}_1$ 先验参数通过以下方程确定：25–26 适用于 $\ell = 1$。对于先验参数 $\boldsymbol{\mu}_\ell$ 和 $\boldsymbol{B}_\ell$，即针对 $\ell = 2$、3、4，我们使用相同的方程，但不采用原始数据 $\boldsymbol{X}_{i\ell j}$，而是考虑字符a的差值，即 $d_{i\ell j} = \boldsymbol{X}_{i\ell j} - \widehat{\boldsymbol{\mu}}_1$。

书写者内变异的逆威沙特分布 $\boldsymbol{U}$ 的迹矩阵通过方程27 和28 确定，方法与正态模型相同，并将所有字符一并纳入考虑；详情请参见Press(1980年)。

最后，共轭方法的 $\boldsymbol{K}_0$ 通过在 $(0,1)^L$ 内进行网格搜索来确定，搜索范围涵盖背景书写者的所有字符：选取使边际似然最大的数值。

对于正态-对数正态-LKJ先验方法，正态先验的参数通过计算本节开头所述的差值，并采用公式（25）和（26）中给出的表达式进行一致估计。此外，通过提取$\widehat{\ }$估计协方差矩阵 $\boldsymbol{W}$的对角元素，我们依据公式（29）、（30）推导出对数正态分布的参数，并将 η LKJ分布的参数设为 1。

E 桥接抽样

最著名且有效的边际似然蒙特卡洛估计量由Meng & Wong （1996年）提出。该方法基于以下简单恒等式：

$$m(\boldsymbol{y}) = \frac{\int h(\boldsymbol{\Theta},\boldsymbol{W})g(\boldsymbol{\Theta},\boldsymbol{W})f(\boldsymbol{y}|\boldsymbol{\Theta},\boldsymbol{W})\pi(\boldsymbol{\Theta},\boldsymbol{W})d(\boldsymbol{\Theta},\boldsymbol{W})}{\int h(\boldsymbol{\Theta},\boldsymbol{W})g(\boldsymbol{\Theta},\boldsymbol{W})\frac{f(\boldsymbol{y}|\boldsymbol{\Theta},\boldsymbol{W})\pi(\boldsymbol{\Theta},\boldsymbol{W})}{m(\boldsymbol{y})}d(\boldsymbol{\Theta},\boldsymbol{W})}, \tag{31}$$

对于任意函数 g（$\boldsymbol{\Theta}$，$\boldsymbol{W}$）和 h（$\boldsymbol{\Theta}$，$\boldsymbol{W}$）。这些函数分别称为提议分布和桥函数。提议分布应近似后验分布，并确保与其充分重叠。此外，函数 h（·）充当连接这两个分布的桥梁。因此，它必须与两个分布都相容，也就是说，所选函数应能针对提议分布和后验分布进行有意义的求值，从而有效地将二者“桥接”起来。对（31）进行简单变换，可得以下恒等式：

$$\begin{aligned} m(\boldsymbol{y}) &= \frac{\int h(\boldsymbol{\Theta},\boldsymbol{W})g(\boldsymbol{\Theta},\boldsymbol{W})f(\boldsymbol{y}|\boldsymbol{\Theta},\boldsymbol{W})\pi(\boldsymbol{\Theta},\boldsymbol{W})d(\boldsymbol{\Theta},\boldsymbol{W})}{\int h(\boldsymbol{\Theta},\boldsymbol{W})g(\boldsymbol{\Theta},\boldsymbol{W})\pi(\boldsymbol{\Theta},\boldsymbol{W}|\boldsymbol{y})d(\boldsymbol{\Theta},\boldsymbol{W})} \\ &= \frac{E_{g(\boldsymbol{\Theta},\boldsymbol{W})}\left[h(\boldsymbol{\Theta},\boldsymbol{W})f(\boldsymbol{y}|\boldsymbol{\Theta},\boldsymbol{W})\pi(\boldsymbol{\Theta},\boldsymbol{W})\right]}{E_{\pi(\boldsymbol{\Theta},\boldsymbol{W}|\boldsymbol{y})}\left[h(\boldsymbol{\Theta},\boldsymbol{W})g(\boldsymbol{\Theta},\boldsymbol{W})\right]}. \end{aligned} \tag{32}$$

因此，可以根据式（32）中的恒等式，使用蒙特卡洛估计量来估计边际似然。该蒙特卡洛边际似然估计量为

$$\widehat{m}(\boldsymbol{y}) = \frac{\frac{1}{T_2}\sum_{t=1}^{T_2} h(\boldsymbol{\Theta}^{*(t)},\boldsymbol{W}^{*(t)})f(\boldsymbol{y}|\boldsymbol{\Theta}^{*(t)},\boldsymbol{W}^{*(t)})\pi(\boldsymbol{\Theta}^{*(t)},\boldsymbol{W}^{*(t)})}{\frac{1}{T_1}\sum_{t=1}^{T_1} h(\widetilde{\boldsymbol{\Theta}}^{(t)},\widetilde{\boldsymbol{W}}^{(t)})g(\widetilde{\boldsymbol{\Theta}}^{(t)},\widetilde{\boldsymbol{W}}^{(t)})}, \tag{33}$$

其中（$\widetilde{\boldsymbol{\Theta}}^{(t)}$，$\widetilde{\boldsymbol{W}}^{(t)}$），对于 $t = 1, \ldots\ldots, T_1$，以及（$\boldsymbol{\Theta}^{*(t)}$，$\boldsymbol{W}^{*(t)}$），对于 $t = 1, \ldots\ldots, T_2$，分别是来自后验分布 $\pi(\boldsymbol{\Theta}, \boldsymbol{W}|\boldsymbol{y})$ 和提议分布 $g(\boldsymbol{\Theta}, \boldsymbol{W})$ 的样本。

提议分布 $g(\cdot)$ 必须接近目标后验分布。函数 $h(\cdot)$ 起到连接这两个分布的桥梁作用。根据 Meng & Wong (1996年，均方误差下的最优函数为

$$h(\boldsymbol{\Theta}, \boldsymbol{W}) = \frac{T_2}{T_2 g(\boldsymbol{\Theta}, \boldsymbol{W}) + T_1 \pi(\boldsymbol{\Theta}, \boldsymbol{W} | \boldsymbol{y})}.$$

将这个最优桥函数代入（32）后，经过一些运算，桥采样指标可简化为

$$\widehat{m}_o(\boldsymbol{y}) = \frac{\frac{1}{T_2} \sum_{t=1}^{T_2} \frac{f(\boldsymbol{y}|\boldsymbol{\Theta}^{*(t)}, \boldsymbol{W}^{*(t)}) \pi(\boldsymbol{\Theta}^{*(t)}, \boldsymbol{W}^{*(t)})}{T_2 g(\boldsymbol{\Theta}^{*(t)}, \boldsymbol{W}^{*(t)}) + T_1 \pi(\boldsymbol{\Theta}^{*(t)}, \boldsymbol{W}^{*(t)} | \boldsymbol{y})}}{\frac{1}{T_1} \sum_{t=1}^{T_1} \frac{g(\widetilde{\boldsymbol{\Theta}}^{(t)}, \widetilde{\boldsymbol{W}}^{(t)})}{T_2 g(\widetilde{\boldsymbol{\Theta}}^{(t)}, \widetilde{\boldsymbol{W}}^{(t)}) + T_1 \pi(\widetilde{\boldsymbol{\Theta}}^{(t)}, \widetilde{\boldsymbol{W}}^{(t)} | \boldsymbol{y})}}. \tag{34}$$

有关此方法的更多详情和全面解释，请参见 Gronau 等人 （2017年）。

按照 Meng & Wong （1996年，第 837 页）的描述，实现式（34）的迭代表达式，以获得最优桥采样估计量：根据预先设定的容差水平，不断更新边际似然的初始估计，直至收敛。随后应用贝叶斯定理得到联合后验分布，并经过一些代数运算，得到第 $z+1$ 次迭代时边际似然 $\widehat{m}_o(\boldsymbol{y})^{(z+1)}$ 的表达式：

$$\widehat{m}_o(\boldsymbol{y})^{(z+1)} = \frac{\frac{1}{T_2} \sum_{t=1}^{T_2} \frac{S(\boldsymbol{\Theta}^{*(t)}, \boldsymbol{W}^{*(t)})}{T_2 \widehat{m}_o(\boldsymbol{y})^{(z)} + T_1 S(\boldsymbol{\Theta}^{*(t)}, \boldsymbol{W}^{*(t)})}}{\frac{1}{T_1} \sum_{t=1}^{T_1} \frac{1}{T_2 \widehat{m}_o(\boldsymbol{y})^{(z)} + T_1 S(\widetilde{\boldsymbol{\Theta}}^{(t)}, \widetilde{\boldsymbol{W}}^{(t)})}}, \tag{35}$$

其中 $S(\boldsymbol{\Theta}, \boldsymbol{W}) = \frac{f(\boldsymbol{y}|\boldsymbol{\Theta}, \boldsymbol{W}) \pi(\boldsymbol{\Theta}, \boldsymbol{W})}{g(\boldsymbol{\Theta}, \boldsymbol{W})}$。

此外，为了获得可靠的边际似然估计，Overstall & Forster （2010年）提出将马尔可夫链蒙特卡罗程序（例如本研究采用的吉布斯抽样）得到的后验样本分为两部分：（a）第一部分用于确定提议分布的参数；（b）第二部分用于式（35）中，以计算桥采样估计量的分子。

E.1 各书写者模型的蒙特卡洛误差

本节比较所考虑模型的边际似然估计量的蒙特卡洛误差（MCE）。为此，采用所实现模型的先验设置得到的边际似然估计，M_2、M_3以及 M_5，M_6对整个数据集 $i = 1$中的每位书写者 m 进行估计$\boldsymbol{\mathcal{D}}_i$。具体而言，对于每位书写者，使用总计 10 次马尔可夫链蒙特卡罗运行来估计对数边际似然，每次运行迭代 2000 次，并额外丢弃 1000 次迭代作为预热期。最后，使用其余书写者的数据来确定先验参数（$m - i$）。表E.1 和E.2分别报告了正态模型在先验设置 M_2、M_3下，以及多元方差分析模型在先验设置 M_5、M_6下的对数边际似然估计均值和蒙特卡洛误差的标准差。

从表格E.1 和E.2中可以观察到两个主要结果。首先，正如科学文献所预期的那样，桥采样估计量系统地产生较小的蒙特卡洛误差（Sinharay & Stern 2005年，Ardia 等人 2012年）。其次，比较所考虑模型的蒙特卡洛误差时，贝叶斯正态模型采用正态-对数正态-LKJ先验所得的蒙特卡洛误差较小。然而，对于贝叶斯多元方差分析，并未呈现出明显的规律；结果取决于书写者，且差异相当均衡。

书写者	正态分布			
	逆威沙特		对数正态-LKJ	
	对数边际似然	MCE	对数边际似然	MCE
1	-1380.3	0.109	-1362.8	0.065
2	-1298.4	0.086	-1270.4	0.078
3	-1639.4	0.084	-1604.6	0.094
4	-1869.8	0.082	-1821.1	0.081
5	-1021.8	0.075	-1027.5	0.073
6	-1378.2	0.094	-1359.7	0.070
7	-1120.7	0.075	-1116.6	0.080
8	-1010.6	0.087	-1007.8	0.098
9	-969.4	0.129	-966.4	0.114
10	-1319.9	0.105	-1296.2	0.077
11	-1851.4	0.120	-1813.8	0.109
12	-1501.5	0.122	-1485.9	0.055
13	-1366.0	0.096	-1349.6	0.092

表 E.1：贝叶斯正态模型在两种不同先验设置（M_2、 M_3）下的对数边际似然（log $-ML$）估计值（均值）和蒙特卡洛误差（标准差）。

书写者	多元方差分析			
	逆威沙特		对数正态-LKJ	
	对数边际似然	MCE	对数边际似然	MCE
1	-1361.3	0.142	-1346.2	0.152
2	-1125.1	0.155	-1105.7	0.104
3	-1523.9	0.126	-1491.8	0.150
4	-1776.7	0.267	-1731.5	0.163
5	-818.2	0.111	-824.0	0.162
6	-1266.4	0.234	-1246.5	0.134
7	-963.3	0.201	-964.3	0.112
8	-878.8	0.116	-879.1	0.153
9	-817.0	0.175	-821.2	0.168
10	-1274.7	0.116	-1254.1	0.164
11	-1666.5	0.182	-1632.8	0.144
12	-1308.4	0.184	-1288.4	0.207
13	-1207.1	0.178	-1195.3	0.085

表 E.2：采用两种不同先验设置（M_5、 M_6）的贝叶斯多元方差分析模型的对数边际似然（log $-ML$）估计值（均值）和蒙特卡洛误差（标准差）。

F 模型比较

表 F.1 展示了每位书写者估计得到的贝叶斯因子 $BF^{1,3}(\mathbf{X}_i)$ 和 $BF^{4,6}(\mathbf{X}_i)$（以对数尺度表示）。具体而言，这涉及比较共轭正态-逆 Wishart 与正态分布-对数正态分布-LKJ分布先验结构。

书写者	正态分布 逆威沙特 vs 对数正态-LKJ log BF^{*}_{13}			多元方差分析 逆威沙特 vs 对数正态-LKJ log BF^{*}_{46}		
	数值	Sign	解释	数值	Sign	解释
1	-17.51	-	极端	-13.51	-	极端
2	-33.05	-	极端	-23.05	-	极端
3	-38.23	-	极端	-37.76	-	极端
4	-50.95	-	极端	-47.93	-	极端
5	7.24	+	极端	10.33	+	极端
6	-19.80	-	极端	-17.71	-	极端
7	-6.31	-	极端	-1.58	-	显著
8	-5.68	-	极端	-1.66	-	显著
9	-3.11	-	强	9.39	+	极端
10	-17.61	-	极端	-6.36	-	极端
11	-34.16	-	极端	-0.15	-	单纯提及
12	-19.21	-	强	-20.85	-	极端
13	-28.03	-	极端	-27.17	-	极端

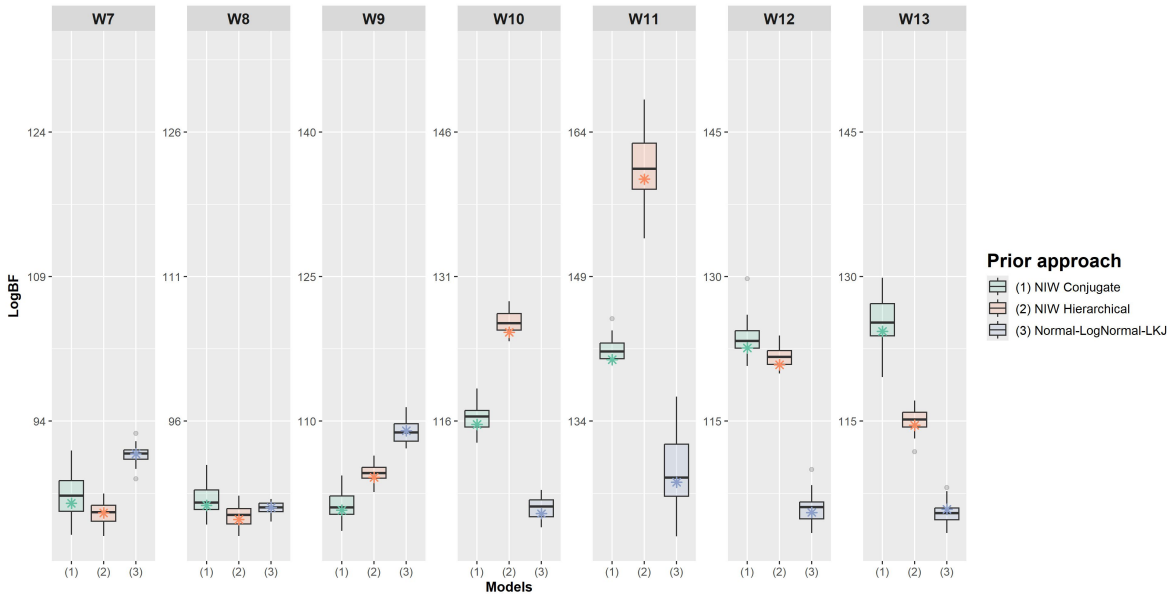
Bridge Sampling estimate

表 F.1：比较贝叶斯正态模型和多元方差分析模型中共轭正态-逆威沙特先验设置与正态-对数正态-LKJ 先验设置的对数贝叶斯因子（按书写者统计）。

G 先验设定的敏感性分析

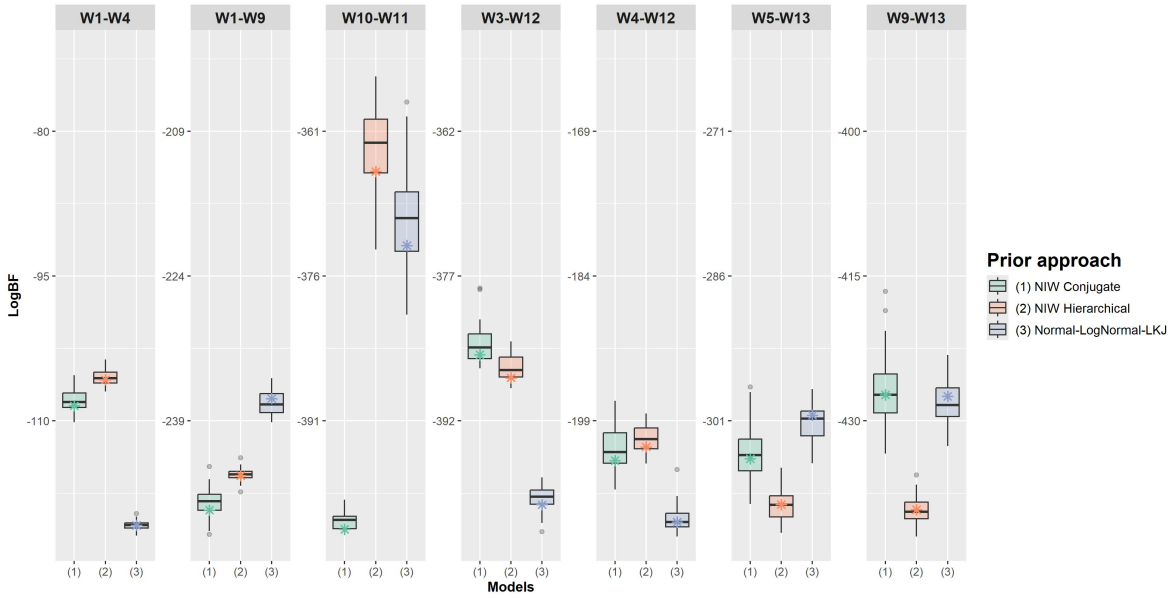
G.1 背景数据子抽样示例

本节将进一步展示如何对与第5.2节中随机选定案例相关的背景数据实施子抽样。为隔离先验 elicitation 的影响，我们对同一书写者实验和不同作者实验均采用同一个随机数据划分，从而尽量减少数据分区可能产生的混杂效应。本分析聚焦于贝叶斯多元方差分析模型，该模型已被确定为适用于此情境的最有效且推荐的方法。对于每个案例，我们在背景作者数据的 50% 中进行 30 次有放回抽样迭代，以确保结果的稳健性。



*indicates the 对数贝叶斯因子 using the complete background dataset for each case.

图 G.1：采用贝叶斯多元方差分析方法，对不同背景数据子样本下同一书写者情景的笔迹评估所得对数贝叶斯因子（对数贝叶斯因子）的箱线图。



*indicates the 对数贝叶斯因子 using the complete background dataset for each case.

图 G.2：采用贝叶斯多元方差分析方法，对不同背景数据子样本下不同书写者情景的笔迹评估所得对数贝叶斯因子（对数贝叶斯因子）的箱线图。

G.2 逆威沙特自由度的敏感性 m

本节进行了敏感性分析，以考察逆威沙特分布自由度的选择对结果的影响；该分布用于建模书写者内变异。先验参数的设定尤为重要，因为它会对最终得到的贝叶斯因子数值产生重大影响。

敏感性分析考察的自由度数值范围为 11（最小值（$p + 2 = 11$））至 50，步长为 10，即 $\nu \in \{11, 20, 30, 40, 50\}$。因此，对于 ν的每个数值，均计算了贝叶斯因子，涵盖同一书写者或不同书写者的字符之间的所有比较，并使用所有模型和边际似然估计方法。按照第5.2节所述的流程，为每对书写者抽取了 10 个子样本。

图 G.1 展示了不同自由度取值下贝叶斯因子（对数尺度）的平均值。正如预期，贝叶斯因子对自由度的选择相当敏感，自由度会对其数值产生递增影响。对于不同书写者比较，逆威沙特分布的自由度每增加 1，对数贝叶斯因子就会增加约 2.2 个单位（其中 $R^2 = 0.97$）。因此，证据对“比较的材料出自同一书写者”这一假设的支持会（错误地）增强。这并不令人意外，因为自由度在逆威沙特分布的先验 elicitation 过程中发挥着重要作用。随着自由度参数 ν 增大， $\boldsymbol{W}$ 的先验分布会更集中于 $\mathbb{R}^p$ 较小的区域，正如 Gaborini (2021年所讨论的那样。由于自由度取值越大，方差-协方差矩阵的变异性越小，因此可以合理地预期，当 ν取值较大时，得到的对数贝叶斯因子将受到这样一种情况的限制：方差-协方差矩阵为常数，且固定为逆威沙特分布的先验均值。图 G.2为此提供了佐证，其中展示了书写者 7 和 8 比较所得的平均对数贝叶斯因子。从图 A.2中的马氏距离可以轻易看出，这两位书写者确实非常相似，其书写风格难以区分。对于这对书写者，错误假设（即 H_1）所获得的证据仍随自由度增加，但超过 40 后，似乎已接近最大值。此外还可以看到，当自由度较小（低于 30）时，贝叶斯因子正确支持假设 H_2 （即比较的手写材料出自不同书写者）；而当自由度较大时，证据则支持错误假设 H_1 （即比较的手写材料出自同一书写者）。

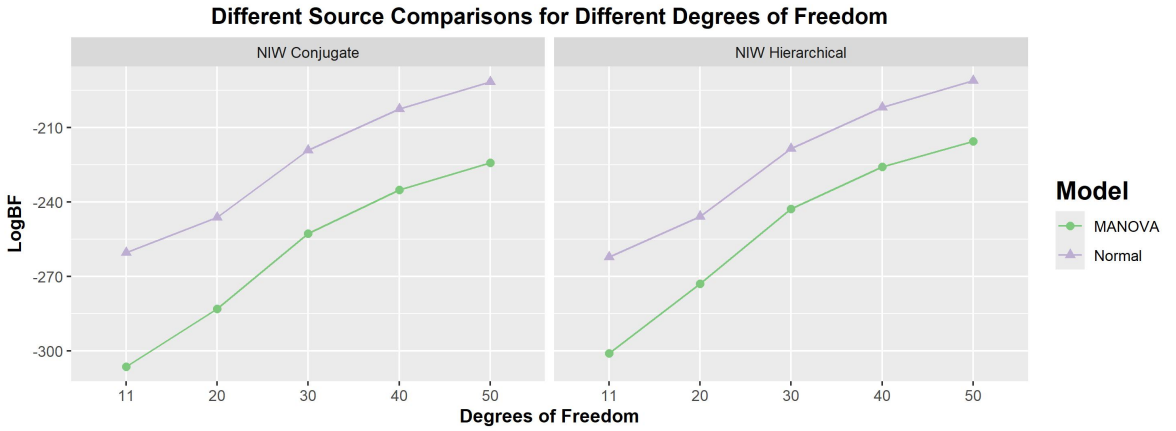


图 G.1：不同自由度下，对不同书写者进行比较时笔迹检验的平均对数贝叶斯因子（对数贝叶斯因子）。正态模型采用所有字符类型联合实现。

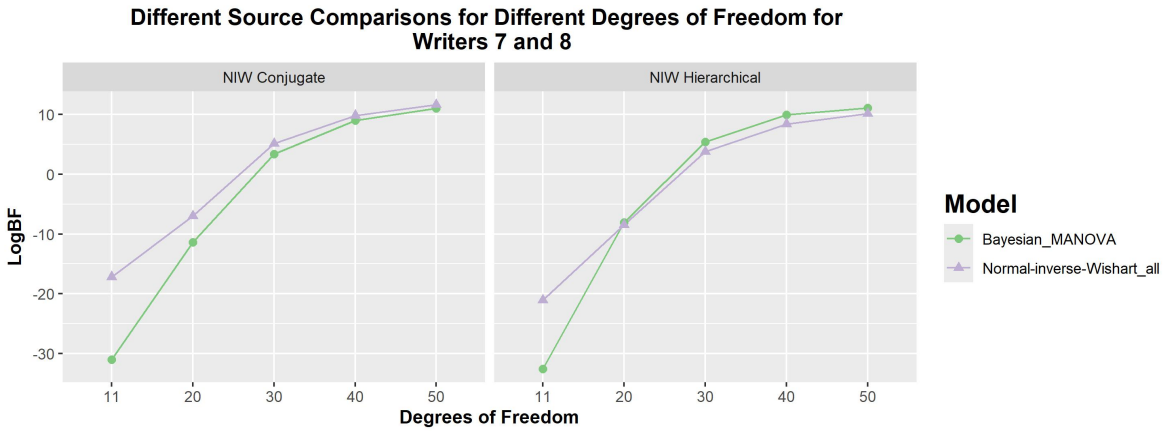


图 G.2：在不同自由度下，对书写者 7 和 8 进行笔迹检验的平均对数贝叶斯因子（log BF）。正态模型采用所有字符类型的联合数据实现。

G.3 LKJ分布参数的敏感性

本节将介绍一系列涉及不同书写者的实验，探讨 LKJ 分布的 η 参数对书写者比较的影响。该参数 η 取值为 1、2、5、10 和 20，以便清楚地观察其对书写者比较的影响。这里只考虑大于 1 的取值，因为这些取值对应于偏好单位矩阵的先验信念。这种做法的依据是傅里叶分析的结果，该结果表明傅里叶系数彼此不相关。因此，使用较大的 η 会突出强调协方差矩阵中变量彼此不相关的先验结构。

图 G.3展示了对数贝叶斯因子（LogBF）如何随 LKJ 参数的不同数值而变化 η （适用于贝叶斯正态模型和多元方差分析模型）。结果显示，随着 LKJ 参数 η 增大，两个模型的 LogBF 值也随之增大。这表明，较高的 η数值对应于更强的、倾向于单位矩阵的先验信念，往往会产生更正的 LogBF 值。这会影响判断，正如图 G.4所示，该图比较了书写者 7 和 8；随着 η 增大，LogBF 值由负转正，而从数值 $\eta = 20$开始，证据支持错误假设。这表明，先验的选择，特别是 LKJ分布中的 η 数值，可能会显著影响书写者比较的结果。

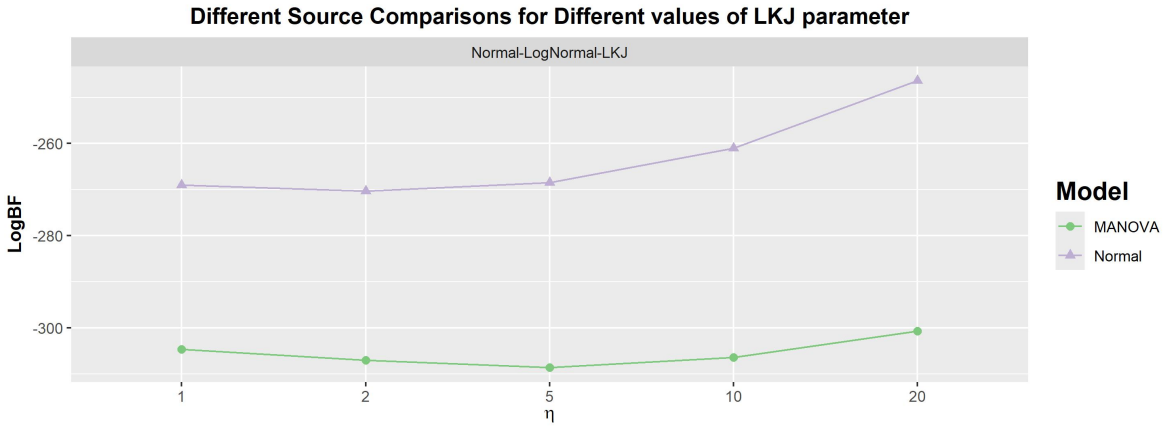


图 G.3：在 LKJ分布的不同 η 数值下，对不同书写者比较进行笔迹检验的平均对数贝叶斯因子（log BF）。正态模型采用所有字符类型的联合数据实现。

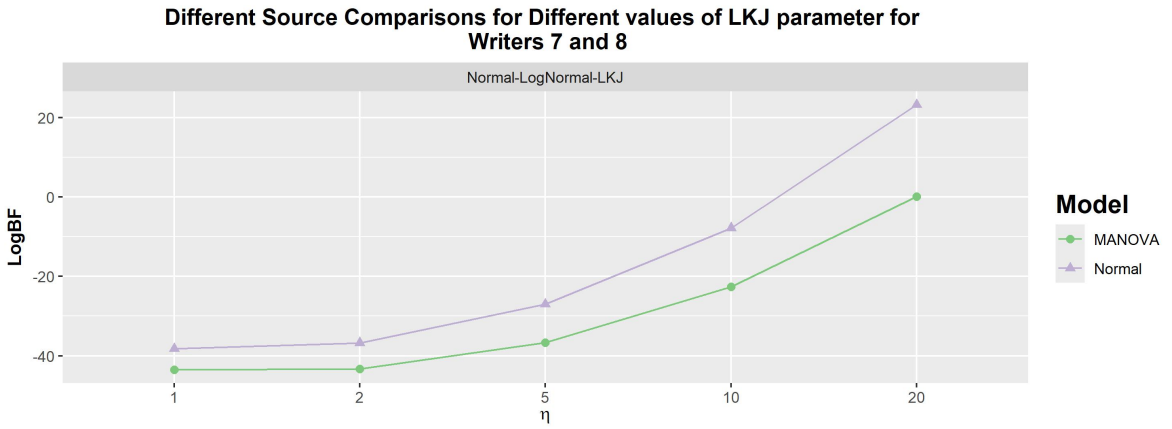


图 G.4：在 LKJ分布的不同 η 值下，对书写者 7 和 8 进行笔迹检验比较所得的平均对数贝叶斯因子（log BF）。正态模型采用所有字符类型联合实现。

H 数据与代码

本文使用的所有数据均由 Raymond Marquis博士慷慨提供给作者。出于保密原因，我们无法公开提供本研究实际数据集的访问权限。因此，我们提供了一个基于 IAM笔迹数据库的示例，可在 Git代码仓库中找到 https://tinyurl.com/lampis-tzai-github-paper。

参考文献

Aitken, C., Taroni, F.& Bozza, S.（2021），法证科学家证据统计与评估，第 3 版，John Wiley & Sons, 奇切斯特。

Ardia, D., Ba¸st¨urk, N., Hoogerheide, L. & VanDijk, H. K.（2012年），‘一项关于蒙特卡罗方法在高效评估边际似然方面的比较研究’，《计算统计与数据分析》**56**(11)，3398–3414。

Bozza, S., Taroni, F.& Biedermann, A.（2022年），《使用*R*进行法证决策分析的贝叶斯因子》，Springer。开放获取：https://link.springer.com/book/10.1007/978-3-031-09839-0。

Bozza, S., Taroni, F., Marquis, R. & Schmittbuhl, M.（2008年），《笔迹证据的概率评估：用于作者身份判定的似然比》，《皇家统计学会杂志：*C*辑（应用统计学）》**57**(3), 329–341。

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P. & Riddell, A.（2017年），《Stan：一种概率编程语言》，《统计软件杂志》**76**。

Crawford, A. M., Ommen, D. M. & Carriquiry, A. L.（2023），‘封闭集合中法证评估笔迹证据的基于旋转的特征与贝叶斯层级模型’，《应用统计年刊》**17**(2), 1127–1151。Gaborini, L.（2021），《可疑笔迹与签名中的贝叶斯模型》，博士论文，Ph.D.´博士论文，犯罪科学学院，洛桑大学。Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B.（2013），贝叶斯数据分析，Chapman and Hall/CRC，纽约。Gronau, Q. F., Heathcote, A. & Matzke, D.（2020），‘使用 Warp-III 桥接抽样计算证据累积模型的贝叶斯因子’，《行为研究方法》**52**(2), 918–937。

Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., Leslie, D. S., Forster, J. J., Wagenmakers, E.-J. & Steingroever, H.（2017），‘桥接抽样教程’，《数学心理学杂志》81, 80–97。Gronau, Q. F., Singmann, H. & Wagenmakers, E.-J.（2020），‘bridgesampling：用于估计归一化常数的 R 软件包’，《统计软件杂志》92, 1–29。Hopwood, A. J., Puch-Solis, R., Tucker, V. C., Curran, J. M., Skerrett, J., Pope, S. & Tully, G.（2012），‘考量单一供体的15重短串联重复序列谱在英国人群中的证明价值及其在英国法院中的呈现’，《科学与司法》52, 185–190。Huang, A. & Wand, M.（2013），‘协方差矩阵的简单边际非信息先验分布’，《贝叶斯分析》8(2), 439–452。Izenman, A. J.（2020），《可疑文件中的笔迹比较》，载于《法证统计学手册》，Chapman and Hall/CRC，纽约，第341–363页。Johnson, M. Q. & Ommen, D. M.（2022），‘使用随机森林和基于分数的似然比进行笔迹鉴定’，《统计分析与数据挖掘：美国统计学会数据科学期刊》15(3), 357–375。Kass, R. E. & Raftery, A. E.（1995），‘贝叶斯因子’，《美国统计学会杂志》90(430), 773–795。Lewandowski, D., Kurowicka, D. & Joe, H.（2009），‘基于藤结构和扩展洋葱法生成随机相关矩阵’，《多元分析杂志》100(9), 1989–2001。Mardia, K. V.（1970），‘多元偏度与峰度的度量及其应用’，《生物统计学》57(3), 519–530。Marquis, R., Schmittbuhl, M., Mazzella, W. D. & Taroni, F.（2005），‘手写字符形状的量化：基于大写字母 o 研究实现书写者客观区分的一步’，《国际法证科学》150(1), 23–32。Marquis, R., Taroni, F., Bozza, S. & Schmittbuhl, M.（2006），‘手写字符环状形态多态性的定量表征’，《国际法证科学》164(2-3), 211–220。Meng, X.-L. & Wong, W. H.（1996），‘通过一个简单恒等式模拟归一化常数比：理论探讨’，《中华统计学志》6(4), 831–860。Murphy, K. P.（2007），《高斯分布的共轭贝叶斯分析》，技术报告 – CS 340/640，不列颠哥伦比亚大学，计算机科学系。最后更新于2007年10月3日。https://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf。Ommen, D. M., Saunders, C. P. & Neumann, C.（2017），‘法证证据价值量化中蒙特卡罗误差的表征’，《统计计算与模拟杂志》87(8), 1608–1643。Overstall, A. M. & Forster, J. J.（2010），‘广义线性混合模型的默认贝叶斯模型判定方法’，《计算统计与数据分析》54(12), 3269–3288。Press, S. J.（1980），‘多元方差分析中的贝叶斯推断’，《统计学手册》，第117–132页。Press, S. J.（2005），《应用多元分析：使用贝叶斯和频率学派推断方法》，Courier Corporation，北切姆斯福德，马萨诸塞州。Rowe, D. B.（2002），《多元贝叶斯统计：源分离与信号解混模型》，Chapman and Hall/CRC，纽约。Schmittbuhl, M., Le Minor, J.-M., Allenbach, B. & Schaaf, A.（1998），‘西部大猩猩、黑猩猩和现代智人的梨状孔形状：表征与多态性分析’，《美国体质人类学杂志：美国体质人类学家协会官方出版物》106(3), 297–310。

Sinharay, S. & Stern, H. S. (2005年), ‘广义线性混合模型中贝叶斯因子计算方法的实证比较’, 计算与图形统计学期刊 **14**(2), 415–435. Soch, J. (2019), ‘统计证明之书’, https://statproofbook.github.io/。Suits, D. B. (1957), ‘回归方程中虚拟变量的使用’, *Journal of the American Statis-tical Association* **52**(280), 548–551. Thiel, Y., Singh, K. & Balakrishnan, R. (2011), Elasticurves：利用笔画动力学和惯性实时修整手绘二维曲线，载于‘第24届ACM用户界面软件与技术年度研讨会论文集’，第383–392页。Willis, S., McKenna, L., McDermott, S., O’Donell, G., Barrett, A., Rasmusson, B., Nordgaard, A., Berger, C., Sjerps, M., Lucena-Molina, J., Zadora, G., Aitken, C., Lovelock, T., Lunt, L., Cham- pod, C., Biedermann, A., Hicks, T. & Taroni, F. (2015年), ‘欧洲法庭科学研究所网络法庭科学评价性报告指南：加强欧洲法庭科学结果评估（STEOFRAE）’，都柏林。https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf。Wydra, J. & Matuszewski, S. (2022年), ‘利用相似度指数评估笔迹证据的似然比’, 法律、概率与风险**21**(1), 21–42.