论文标题

**摘要**

丝绸之路是古代中西方文化交流的通道，其中玻璃是早期贸易往来的宝贵物证。早期的玻璃在西亚和埃及地区常被制作成珠形饰品传入我国，并在时光的长河中淬炼出中国玻璃艺术特有的神彩，尽显中华文明博大的包容性和伟大的创造性，成为中华优秀传统文化的有机组成部分。

针对问题一：根据表单一发现纹饰、类型、颜色、表面风化均为定类变量，针对多组定类变量之间的差异性分析我们采用卡方检验，有卡方检验分析结果可得纹饰、颜色与表面风化的显著性P值大于0.05，不存在显著性差异；而类型与表面风化的显著性P值小于0.05，故存在显著性差异，颜色、类型和表面风化相关程度为中等。通过玻璃类型和是否风化对文物样品进行分类，得出不同类型玻璃风化前后的化学成分含量图表，通过对图标信息分析得出高钾玻璃在风化后主要化学成分含量呈下降趋势；铅钡玻璃在风化后主要化学成分含量呈上升趋势。采用独立样本t检验，检验各化学成分和表面风化的差异性显著与否，从而寻找风化前后化学成分含量均值的关系，最后计算各成分有风化均值和无风化均值之间的比值，利用这个比值可以求出风化前的化学成分含量。

针对问题二：根据表单二中数据分析，通过SPSS聚类分析得出通过与氧化铅含量占比的均值的对比可将文物样品分为两类，即高钾玻璃和铅钡玻璃。由相同的方法得出高钾玻璃和铅钡玻璃的亚类划分分别与氧化钾，氧化铅、氧化钡化学成分含量占比的均值有关，从而将高钾玻璃和铅钡玻璃分别划分成a、b、c、d类以及e、f、g、h类。利用蒙特卡罗算法模型对划分结果的敏感性进行分析，从而得出在风化前后高钾玻璃和铅钡玻璃的亚类划分具有可靠性和稳健性。

针对问题三：根据表单三所给出的未知类别文物数据，结合问题二所得出的亚类划分结果，对表单三的化学成分数据进行对比，从而鉴别出未知文物类型，其中A6、A7为a类，A2、A5为b类，A1为c类，A3、A4、A8为h类。运用问题二相同的蒙特卡罗算法模型对未知类型的玻璃文物进行敏感性分析，进而得出对未知类型玻璃文物的分类结果相对合理、可靠。

针对问题四：通过建立灰色关联系数模型，对化学成分之间的关联关系进行分析，利用SPSS求出选择不同的化学成分作为模型的母序列的关联度，将高钾玻璃和铅钡玻璃风化前后某一化学成分与其他化学成分的关联度绘制出折线图，直观对比各化学成分之间的关联关系，同理可得不同种类的化学成分关联关系的差异性，最后得出玻璃文物风化前后各化学成分间的关联关系基本保持不变，高钾玻璃文物的化学成分的关联关系比铅钡玻璃文物的更集中更为稳定。

关键字：卡方检验、独立样本t检验、聚类分析、蒙特卡洛算法模型、灰色模型

一、问题重述

玻璃最早通过丝绸之路从西亚和埃及地区传入我国，二氧化硅作为玻璃的主要化学成分，因熔点较高在玻璃的制作过程中需添加助溶剂降低熔化温度，因助溶剂的种类多样，从而出现了不同种类的玻璃，然而古代玻璃极易受埋藏环境的影响而风化，内部元素与环境元素进行大量的交换，从而导致其成分比例发生变化，影响对玻璃文物的类别判断，通过对文物样品分类规律，表面风化因素分析，化学成分含量的统计与分析，各化学成分之间的关联和差异进行研究，对探索古代玻璃器的来源、技术起源、发展和成分体系以及早期中外玻璃贸易和技术交流有着十分深远的意义和价值。

问题一：根据表单一所提供的信息对玻璃的类型，纹饰颜色与其表面风化的关系进行分析；分别分析高钾玻璃和铅钡玻璃表面有无风化化学成分含量的统计规律，根据表单二所提供的文物采样点以及同一文物采样点不同部位化学成分含量预测玻璃文物样品风化前的化学成分含量。

问题二：依据表单二所给出的各化学成分含量，找出高钾玻璃和铅钡玻璃的分类规律，并且对两类玻璃进行再次分类同时给出划分依据，最后，建立蒙特卡洛算法模型对亚分类结果的合理性及敏感性进行分析。

问题三：分析表单三所给的未知玻璃文物的化学成分含量，结合问题二的分类方法，对玻璃文物进行分类并分析分类结果得敏感性。

问题四：对高钾玻璃和铅钡波玻璃的化学成分的关联关系进行分析，通过对各化学成分之间的关联关系的对比得出其差异性。

二、问题分析

2.1问题一的分析

对表单一所提供的四个非连续变量进行分析，采用斯皮尔曼相关系数分析玻璃文物纹饰类型，颜色，表面风化的相关性，采用卡方检验，对其差异进行分析；首先将表单二所提供的数据进行缺失值处理，其次所给数据进行合成及分类，由此推断出玻璃类型对表面有无风化与其化学成分的差异性及相关性，最后得出文物样品表面有无风化化学成分含量的统计规律公式；根据所得统计规律公式预测玻璃文物风化前的化学成分含量。

2.2问题二的分析

利SPSS对表单二所给出的化学成分含量进行聚类分析，通过聚类分析谱系图得出高钾玻璃与铅钡玻璃的分类规律，分别对风化前后的高钾玻璃和铅钡玻璃的化学成分含量进行聚类分析，将高钾玻璃和铅钡玻璃进行亚分类，最后通过蒙特卡洛算法模型对分类结果的敏感性进行分析。

2.3问题三的分析

将表单三中的数据分为风化和未风化两组，通过问题二得出的分类方法，首先确定风化和未风化组的文物样品的玻璃类型，其次进一步分析文物样品所属的分类类别，利用问题二所建立的蒙特卡洛算法模型对未知文物分类的结果进行敏感性分析。

2.4问题四的分析

首先选择表单三中占比大的化学成分作为分析的因变量，将剩余的合适的化学成分作为自变量，针对不同类别的玻璃，通过建立灰色关联分析模型，计算灰色关联度的情况；其次将不同类别的玻璃所得的灰色关联系数做方差分析处理，通过显著性检验观察分析高钾玻璃和铅钡玻璃之间的差异性。

三、模型假设

（1）假设表面风化情况只与题中所给变量类型、颜色、纹饰及化学成分含量有关，将时间、温度、湿度、气压等视为无关变量。

（2）假设纹饰对分类无影响

四、符号说明

|  |  |
| --- | --- |
| 符号 | 符号含义及说明 |
|  | 基本变量 |
|  | 失效概率 |
|  | 对分布参数 |
|  | 第i个变量分布参数总数的偏导数 |
|  | 可靠性灵敏度 |
|  | 基本随机变量 |
|  | 联合累计分布函数 |
|  | 可靠性指标 |
|  | 可靠概率 |
|  | 随机样本 |
|  | 可靠度 |
|  | 可靠域指示函数 |

五、模型的建立与求解

5.1问题一模型的建立与求解

根据题意将问题一分为三个小问，第一小问首先通过交叉列表的方法粗略判断玻璃类型、纹饰、颜色与表面是否风化的相关关系，然后运用相关系数测量法来测量相关性的大小，最后用卡方检验判断表单一所给样本统计量是否可以推论总体；第二小问分别对高价玻璃和铅钡玻璃分化前后化学成分含量统计及分析，得出玻璃风化前后化学成分频率直方图，而得出风化前后化学成分的统计规律公式；第三小问通过统计规律，预测风化文物在风化前的化学成分含量。

5.1.1 数据预处理

附件表单1中颜色列中有四个空值，通过观察数据变化情况发现颜色的深浅程度与风化程度呈现简单正相关变化，因此我们将四个空值进行填补，填补为“黑色”。

5.1.2 问题一第（1）问的模型建立与求解

通过观察数据发现纹饰、类型、颜色、表面风化均为定类变量，针对多组定类变量之间的差异性分析我们采用卡方检验。卡方检验主要是比较定类变量与定类变量之间的差异性分析。通过统计样本的实际观测值与理论推断值之间的偏离程度，实际观测值与理论推断值之间的偏离程度就决定卡方值的大小，如果卡方值越大，二者偏差程度越大；反之，二者偏差越小；若两个值完全相等时，卡方值就为0，表明理论值完全符合。

下列表为四个定类变量卡方分析的结果：

表格

描述已自动生成

表 变量卡方分析结果



表 表面风化效应量化分析

由上表卡方检验分析结果可得，纹饰、颜色与表面风化的显著性P值大于0.05，不存在显著性差异；而类型与表面风化的显著性P值小于0.05，故存在显著性差异。

由表 表面风化效应量化分析可知，纹饰Phi值为0.292，小于0.3，说明其和表面风化相关性较弱，颜色、类型Phi值均介于0.3和0.6之间，说明颜色、类型和表面风化相关程度为中等。

图表, 条形图

描述已自动生成

图 纹饰和表面风化的交叉图

图表, 树状图

描述已自动生成

图 类型和表面风化的交叉图

图表, 条形图

描述已自动生成

图 颜色和表面风化的交叉图

表面风化对于类型呈现出0.01水平显著性(chi=6.880, *p*=0.009<0.01)，通过百分比对比差异可知，风化选择铅钡的比例82.35%，明显高于无风化的选择比例50.00%。无风化选择高钾的比例50.00%，明显高于风化的选择比例17.65%。

总结可知：不同表面风化样本对于纹饰, 颜色共2项不会表现出显著性差异，另外，表面风化样本对于类型共1项呈现出显著性差异。

5.1.3 问题一第（2）问的模型建立与求解

题目要求分析不同类型玻璃文物样品表面风化化学成分含量的统计规律，所以我们首先将化学成分数据作基于类型的分类，再基于表面是否风化作进一步分类，最终得到以下高钾玻璃样品风化前后化学成分含量表和铅钡玻璃样品风化前后化学成分含量表：

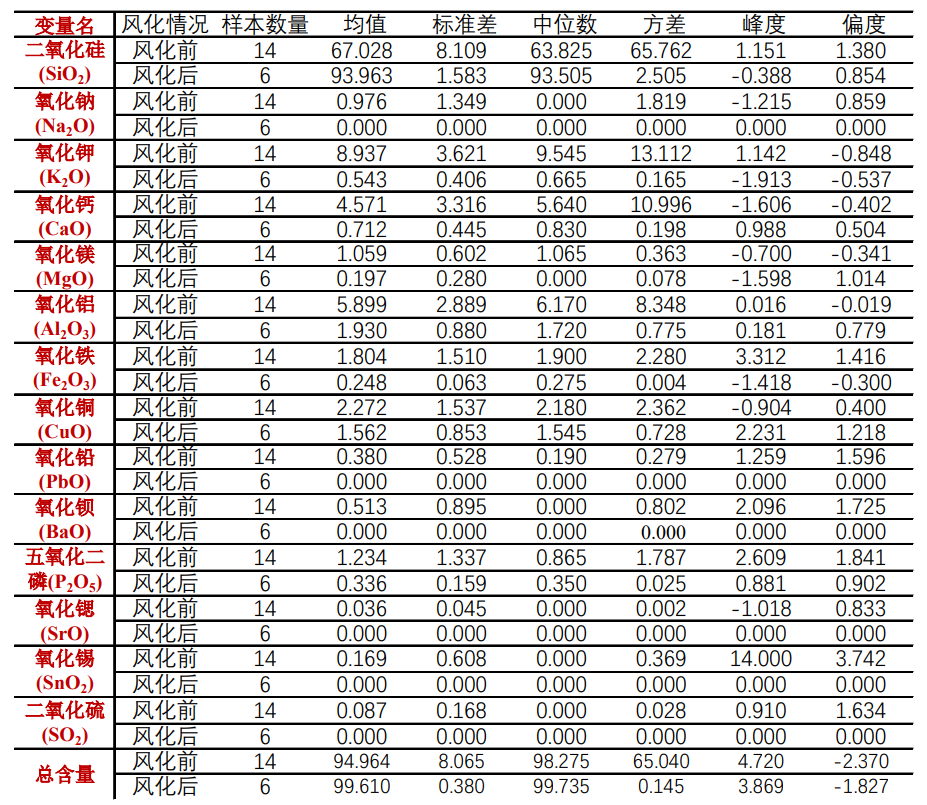


表 高钾玻璃样品风化前后化学成分含量

电脑屏幕的照片

中度可信度描述已自动生成

表 铅钡玻璃样品风化前后化学成分含量

日程表

低可信度描述已自动生成

图 高钾玻璃样品表面风化和所有项分析对比

图表, 条形图

描述已自动生成

图 铅钡玻璃样品表面风化和所有项分析对比

由表面风化和所有项分析对比图可以直观地看出，高钾玻璃在风化后主要化学成分含量呈下降趋势；铅钡类玻璃在风化后主要化学成分含量呈上升趋势。

由此，第（2）问解决。

5.1.4 问题一第（3）问的模型建立与求解

题目要求由玻璃样品风化数据预测其风化前的各化学成分含量，所以我们根据风化前后的数据规律，总结出各个化学成分的变化情况，找到其映射关系并预测风化前的含量。本文采用独立样本t检验，先检验各化学成分和表面风化的差异性显著与否，从而寻找风化前后化学成分含量均值的关系，借助SPSS软件对数据进行处理。

表 高钾玻璃文物样品风化前后数据

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **变量名** | 风化情况 | 均值 | Sig.（双尾） | 均值比值 |
| **二氧化硅(SiO2)** | 风化前 | 67.028 | 0.000 | 0.713 |
| 风化后 | 93.963 |
| **氧化钠(Na2O)** | 风化前 | 0.976 | 0.022 | 无效 |
| 风化后 | 0.000 |
| **氧化钾(K2O)** | 风化前 | 8.937 | 0.000 | 16.459 |
| 风化后 | 0.543 |
| **氧化钙(CaO)** | 风化前 | 4.571 | 0.001 | 6.420 |
| 风化后 | 0.712 |
| **氧化镁(MgO)** | 风化前 | 1.059 | 0.005 | 5.376 |
| 风化后 | 0.197 |
| **氧化铝(Al2O3)** | 风化前 | 5.899 | 0.006 | 3.056 |
| 风化后 | 1.930 |
| **氧化铁(Fe2O3)** | 风化前 | 1.804 | 0.003 | 7.274 |
| 风化后 | 0.248 |
| **氧化铜(CuO)** | 风化前 | 2.272 | 0.326 | 1.455 |
| 风化后 | 1.562 |
| **氧化铅(PbO)** | 风化前 | 0.380 | 0.022 | 无效 |
| 风化后 | 0.000 |
| **氧化钡(BaO)** | 风化前 | 0.513 | 0.059 | 无效 |
| 风化后 | 0.000 |
| **五氧化二磷(P2O5)** | 风化前 | 1.234 | 0.175 | 3.673 |
| 风化后 | 0.336 |
| **氧化锶(SrO)** | 风化前 | 0.036 | 0.140 | 无效 |
| 风化后 | 0.000 |
| **氧化锡(SnO2)** | 风化前 | 0.169 | 0.527 | 无效 |
| 风化后 | 0.000 |
| **二氧化硫(SO2)** | 风化前 | 0.087 | 0.085 | 无效 |
| 风化后 | 0.000 |
| **总含量** | 风化前 | 94.964 | 1.98 | 0.953 |
| 风化后 | 99.610 |

表 铅钡玻璃文物样品风化前后数据

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **变量名** | 风化情况 | 均值 | Sig.（双尾） | 均值比值（前/后） |
| **二氧化硅(SiO2)** | 风化前 | 55.580 | 0.000 | 2.174 |
| 风化后 | 25.562 |
| **氧化钠(Na2O)** | 风化前 | 1.613 | 0.010 | 6.888 |
| 风化后 | 0.234 |
| **氧化钾(K2O)** | 风化前 | 0.210 | 0.423 | 1.449 |
| 风化后 | 0.145 |
| **氧化钙(CaO)** | 风化前 | 1.462 | 0.012 | 0.549 |
| 风化后 | 2.665 |
| **氧化镁(MgO)** | 风化前 | 0.665 | 0.936 | 0.979 |
| 风化后 | 0.680 |
| **氧化铝(Al2O3)** | 风化前 | 4.528 | 0.086 | 1.498 |
| 风化后 | 3.023 |
| **氧化铁(Fe2O3)** | 风化前 | 0.805 | 0.429 | 1.393 |
| 风化后 | 0.578 |
| **氧化铜(CuO)** | 风化前 | 1.509 | 0.349 | 0.690 |
| 风化后 | 2.188 |
| **氧化铅(PbO)** | 风化前 | 21.206 | 0.000 | 0.484 |
| 风化后 | 43.787 |
| **氧化钡(BaO)** | 风化前 | 8.709 | 0.264 | 0.771 |
| 风化后 | 11.293 |
| **五氧化二磷(P2O5)** | 风化前 | 1.051 | 0.000 | 0.209 |
| 风化后 | 5.034 |
| **氧化锶(SrO)** | 风化前 | 0.257 | 0.030 | 0.607 |
| 风化后 | 0.423 |
| **氧化锡(SnO2)** | 风化前 | 0.045 | 0.639 | 0.607 |
| 风化后 | 0.074 |
| **二氧化硫(SO2)** | 风化前 | 0.153 | 0.312 | 0.179 |
| 风化后 | 0.854 |
| **总含量** | 风化前 | 97.788 | 0.107 | 1.013 |
| 风化后 | 96.539 |

由表 高钾玻璃文物样品风化前后数据和表 铅钡玻璃文物样品风化前后数据总结可知：在高钾玻璃文物样品中，不同表面风化样本对于氧化铜, 氧化钡, 五氧化二磷, 氧化锡, 二氧化硫共5项不会表现出显著性差异，另外表面风化样本对于二氧化硅, 氧化钠, 氧化钾, 氧化钙, 氧化镁, 氧化铝, 氧化铁, 氧化铅, 氧化锶共9项呈现出显著性差异。

在铅钡玻璃文物样品中，不同表面风化样本对于氧化钾, 氧化镁, 氧化铝, 氧化铁, 氧化铜, 氧化钡, 氧化锡, 二氧化硫共8项不会表现出显著性差异，另外表面风化样本对于二氧化硅, 氧化钠, 氧化钙, 氧化铅, 五氧化二磷, 氧化锶共6项呈现出显著性差异。

在经过独立样本t检验证明有些化学成分显示出显著性差异后，我们可以计算各成分有风化均值和无风化均值之间的比值，利用这个比值可以求出风化前的化学成分含量。

在表 高钾玻璃文物样品风化前后数据和表 铅钡玻璃文物样品风化前后数据中已经列出各化学成分风化前后均值的比值，在已知风化后化学成分的值后，风化前化学成分为：

在玻璃文物样品中，由于一些化学成分风化后的均值为0，故由计算各成分有风化均值和无风化均值之间的比值的方法，预测不到这几个化学成分风化前的含量，这也是这种办法的局限性。

5.2问题二模型的建立与求解

5.2.1模型建立

利用SPSS对表单2的化学成分含量进行聚类分析

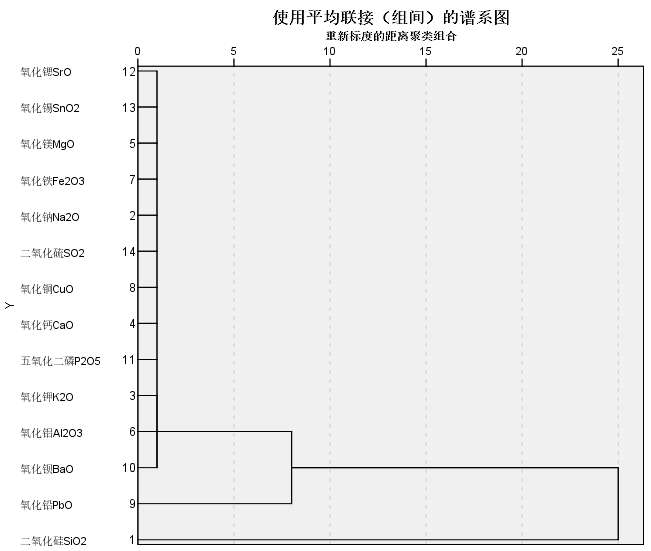


图 化学成分含量聚类分析谱系图

分析谱系图可以发现，将玻璃文物分为高钾玻璃和铅钡玻璃两类的分类方法与氧化铅的含量有关。对铅钡玻璃风化前后的氧化铅成分含量求均值，若玻璃文物风化前氧化铅的含量大于均值21.2061%，则是铅钡玻璃，反之则是高钾玻璃；若玻璃文物风化后氧化铅的含量大于均值43.7871%，则是铅钡玻璃，反之则是高钾玻璃。

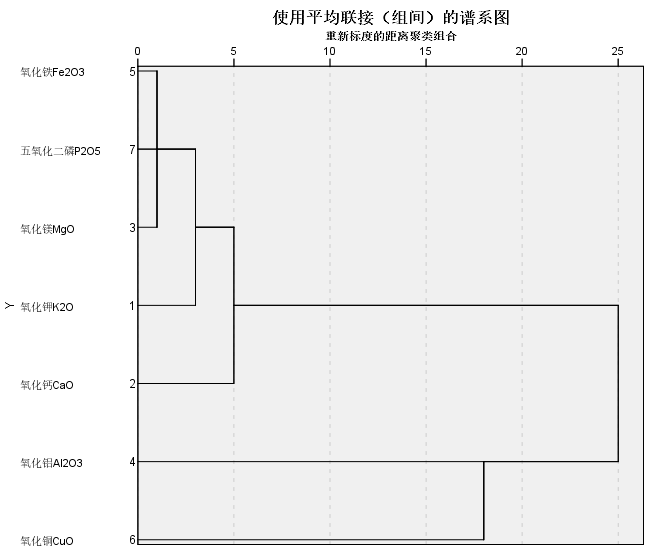


图 高钾风化聚类分析谱系图

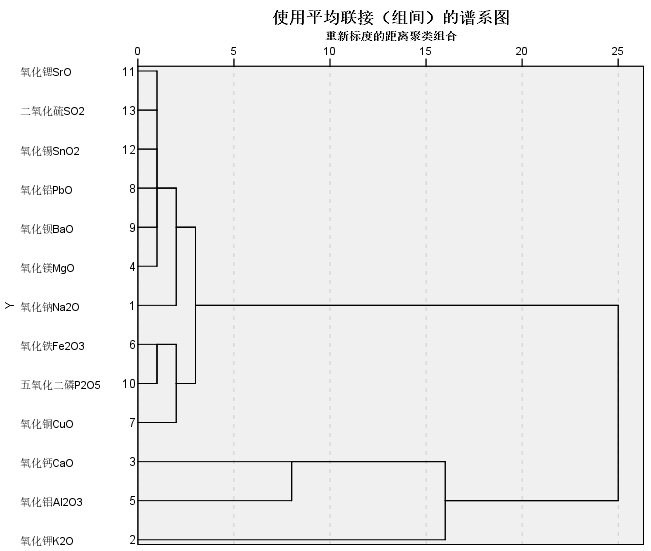


图 高钾未风化聚类分析谱系图

（此处仅提供高钾玻璃风化与未风化的聚类分析谱系图，铅钡玻璃的聚类分析谱系图详见附录）

将风化的和风化前的高钾玻璃和铅钡玻璃的化学成分进行聚类分析，根据得出聚类分析谱线图对两种玻璃类型进行亚分类，将风化的氧化钾含量大于均值0.5433%，氧化铝含量大于均值1.93%的高钾玻璃称为a类，反之则为b类；将风化前不含有氯化钾的高钾玻璃称为c类，反之则为d类；将风化的氧化铜含量大于均值2.1883%，氧化钡含量大于均值11.2925%，氧化铅含量大于均值43.7871%的铅钡玻璃称e类，反之则为f类；将风化前的氧化铝含量大于均值4.5283%的铅钡玻璃称为g类；反之则为h类。

通过建立蒙特卡洛算法模型，对上述亚分类的敏感性进行分析。

灵敏度未失效概率对基本变量的分布参数（i=1,2,...,n;k=1,2,...,m）,其中为第i个变量分布参数总数的偏导数，将失效概率的积分对分布参数求导数，则得到可靠性灵敏度如下：

（1）

Monte Carlo模拟过程中，采用样本均值代替总体均值，数学期望可以用样本函数的均值来估计：

（2）

失效概率对于参数的无量纲灵敏度系数表达式如下：

（3）

同理采用样本点均值作为总体的均，则的Monte Carlo估 （4）

用表示基本随机变量，的来联合概率密度函

数，用表示来联合累计分布函数，那么系统的失效概率可表示为：

（5）

假设功能函数,则功能函数的概率密度函数为：

（6）

通过变化可以将功能函数转换为标准正态分布，那么其他概率密度分布函数为：

（7）

（8）

从而代入分析得到失效概率为：

（9）

可靠性指标、失效概率和可靠概率分别表示为：

（10）

（11）

（12）

Monte Carlo法计算分类可靠度的大致过程如下：由基本随机参数的联合概率密度函数模拟产生N组随机样本，把样本数据值带入分类化学成分含量的极限状态方程里，然后统计落入可靠域里的样本点数，用频率近似代替可靠度，则

（13）

其中：为可靠域的指示函数；表示n维变量空间；表示数学期望算子。

（模型的相关计算结果详见附录）

综上，我们可以确定对于高钾玻璃和铅钡玻璃的亚类划分结果的敏感性较高，对某种化学成分含量占比与其均值进行比较进行划分的方法相对合理。

5.3问题三模型的建立与求解

通过问题二所得出的结论，未风化时，氧化铅的含量大于均值21.2061%，则是铅钡玻璃，反之则是高钾玻璃，故A3、A4、A8所以这三个属于铅钡玻璃，A1属于高钾玻璃。风化后，氧化铅的含量大于均值43.7871%，则是铅钡玻璃，反之则是高钾玻璃，所以A2、A5、A6、A7为高钾玻璃；结合问题二所给出的亚分类方法，表单三中给出的未分类型的元素里所有的铅钡玻璃都是未风化的，而未风化的铅钡玻璃的分亚类的依据是氧化铜、氧化铝、氧化钡，所以只有当三种化学元素的值都大于均值时，才分为g类，所以都不满足，所以他们都属于h类，故推测出，A6、A7为a类，A2、A5为b类，A1为c类，A3、A4、A8为h类。

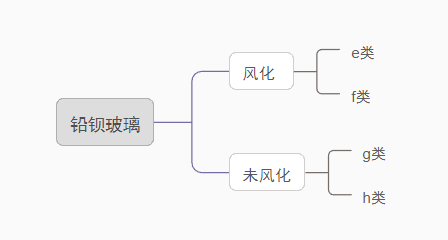
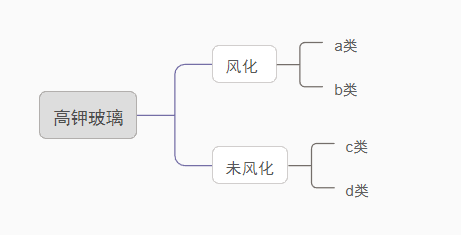
 

图 高钾玻璃和铅钡玻璃的亚类划分

5.3.1建模思路

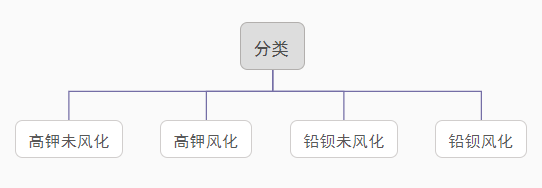
根据问题二所建立的蒙特卡洛算法模型同理可得，将本文对未知玻璃文物分为a、b、c、h类进行敏感度分析，确认未知文物分类的标准具有可靠性和稳健性。

（计算过程详见附录）

5.4问题四模型的建立与求解

5.4.1数据预处理

根据玻璃文物类别的差异，将文物分为四类，如下：



5.4.2模型建立

针对不同类别的玻璃文物样品，分析其化学成分之间的关联关系，我们选择建立灰色关联分析模型，用关联系数和关联度来表征各化学成分间的关联关系。关联系数表示了**各个时刻参考序列和比较序列之间的关联程度，**为了从总体上了解序列之间的关联程度，必须求出它们的时间平均值，即关联度。在不同类别的玻璃文物中，分别选择不同的的化学成分作为模型的母序列，其余变量作为子序列，利用SPSS处理得到以下结果：

图表, 条形图

描述已自动生成

图 高钾玻璃氧化前二氧化硅作为母序列的灰色模型关联度图

图表, 条形图

描述已自动生成图高钾玻璃氧化后二氧化硅作为母序列的灰色模型关联度图

图片包含 表格

描述已自动生成图片包含 表格

描述已自动生成

表 高钾风化前 表 铅钡风化前

图片包含 表格

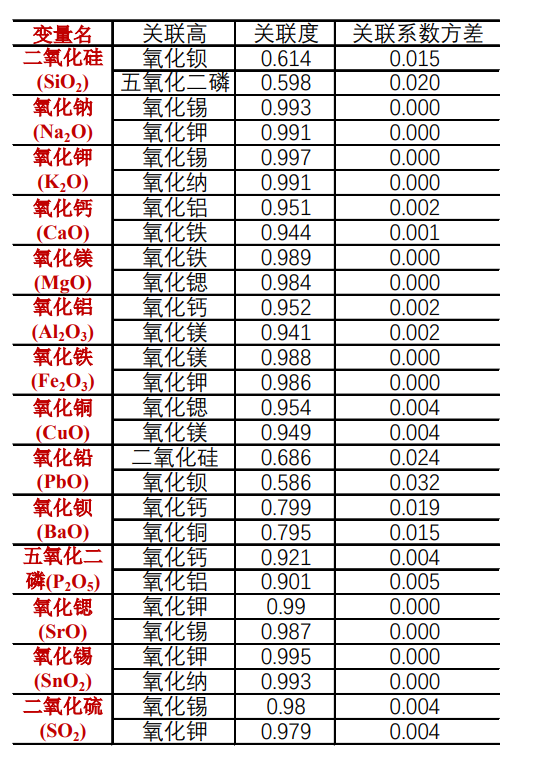
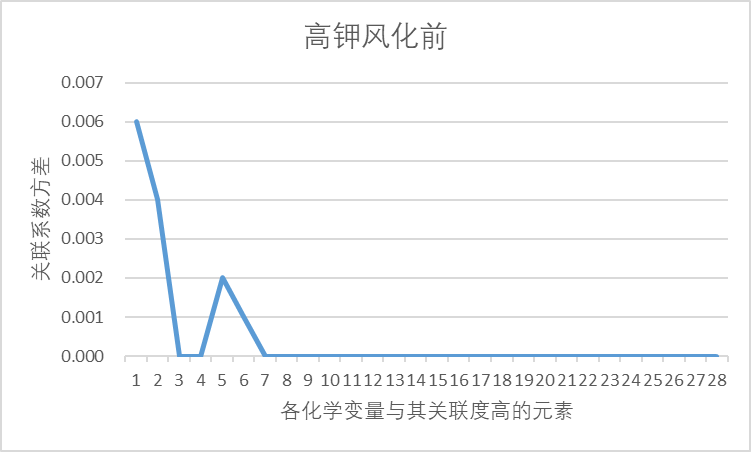
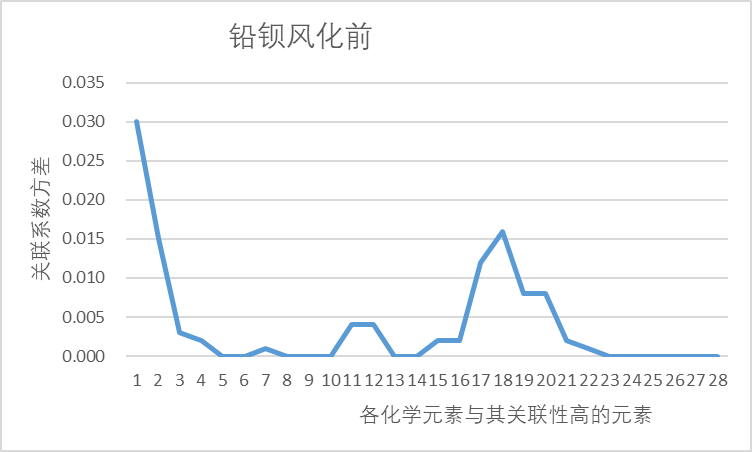
描述已自动生成 

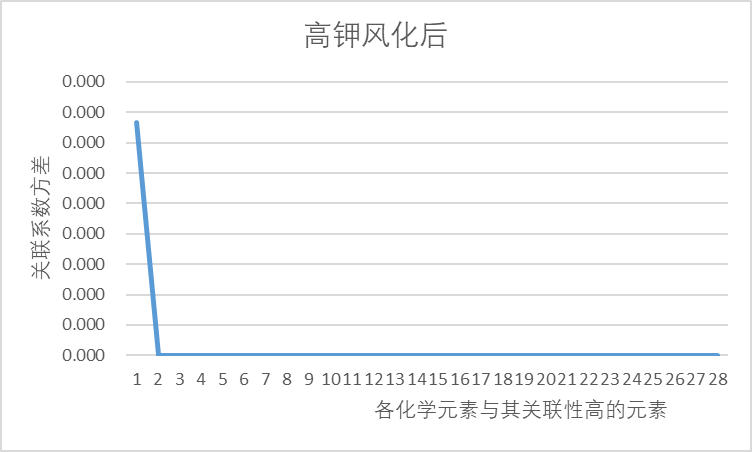
表 高钾风化后 表 铅钡风化后

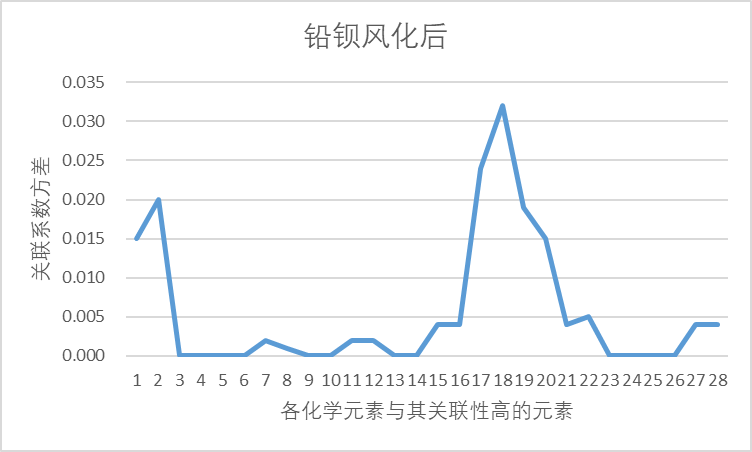
由上列表格可知，在高钾玻璃风化前，与二氧化硅联系最紧密的是氧化钾和氧化铝，其关联度分别为0.894和0.866。同理通过上列表格我们也能读出不同类别的玻璃文物样品其化学成分之间的关联关系。

由此，第一问解决。









对于第二问，比较不同类别之间的化学成分关联关系的差异性。同样在上面的折线图中，我们对所计算出来的灰色关联系数进行了求方差处理，处理结果也在上列折线图

中。通过观察上列表格我们很容易发现，玻璃文物风化前后各化学成分间的关联关系基本保持不变，高钾玻璃文物的化学成分的关联关系比铅钡玻璃文物的更集中更为稳定。

六、模型的评价与推广

6.1模型的优点

蒙特卡洛算法模型方法的误差与问题的维数无关，对于具有统计性质问题可以直接进行解决，对于连续性的问题不必进行离散化处理；聚类分析法原理比较简单,实现也是很容易,收敛速度快，处理大数据集,该算法保持可伸缩性和高效性；灰色模型需要的数据较少,运算方便,是处理小样本预测的有效工具。

6.2模型的缺点

蒙特卡洛算法模型对于确定性问题需要转化成随机性问题，误差是概率误差，通常需要较多的计算步数N. 聚类分析法如果所有点在指派步骤都未分配到某个簇,就会得到空簇，对噪声和异常点比较敏感，K值不好把握，对于不是凸的数据集比较难收敛；灰色模型不考虑系统内在机理,有所会出现较大错误。

6.3模型的推广

很多实际问题要么没有具体求解方法，要么求解方法非常复杂，这时往往可以采取蒙特卡洛方法解决问题.

七、参考文献

书籍的表述方式为：

[1] 刘晓静,基于蒙特卡洛方法的可靠性灵敏度分析,山西工程职业学院：2021(11)-0053-03

参考文献中期刊杂志论文的表述方式为：

[2] 作者，论文名，杂志名，卷期号：起止页码，出版年。

参考文献中网上资源的表述方式为：

[3] 作者，资源标题，网址，访问时间（年月日）

八、附录