

统计分析软件 SPSS 在化探数据处理中的应用

蔡朝阳, 韩秀梅, 吴国学, 孙德有

(吉林大学地球科学学院, 吉林 长春 130061)

社会科学统计软件包 (SPSS-Statistical Package for the Social Science) 是公认的最优秀的统计分析软件包之一。该软件以易于使用的操作界面、强大而成熟的资料处理和统计分析功能、人性化的图形操作界面以及美观的结果输出等优点, 被广泛应用于自然科学、社会学、经济学、教育学、心理学、医学等各个领域 (黄燕波, 2007)。

近年来, 运用 SPSS 对东北地区多个勘查区的水系沉积物测量、土壤地球化学测量等化探数据进行了统计分析。结果表明, 该软件不仅能够提高数据处理的效率, 而且有助于从大量测试数据中提取有效的找矿信息, 为化探异常的圈定、区域成矿规律总结及成矿远景预测等奠定基础。本文以 SPSS Statistics 17.0 为蓝本, 以吉林珲春地区的相关资料为实例, 简要介绍该软件在化探数据处理中的具体使用方法。

1 地球化学背景值与异常下限

利用 SPSS 软件, 可通过直方图解法和计算法综合确定不同元素的地球化学背景值和异常下限。具体步骤包括:

- (1) 在 SPSS 主窗口, 点击“文件”主菜单, 打开单元素测试结果原始数据文件 (.xls);
- (2) 点击“分析”菜单, 在下拉菜单中选择“描述统计”中的“频率”选项;
- (3) 选择要计算或统计的变量 (元素), 点击对话框右侧“统计量”菜单, 根据需要逐项选择“均值”、“众值”、“最小值”、“最大值”、“标准离差”等参数;
- (4) 在“图表”对话框中选择“直方图”和“带正态曲线”等选项;
- (5) 点击“确定”, 程序自动生成所选变量 (元素) 的描述统计结果文件 (.spv)。

在直方图解法中, 根据 SPSS 生成的元素含量 (对数含量) 直方图, 确定众数值 (或对数众数值) (C_o), 用其代表统计单元内的元素含量背景值, 用图解法或计算结果确定标准离差 (δ)。一般将背景值与 2 倍标准离差之和作为该元素的异常下限值 (C_A), 即: $C_A = C_o + 2\delta$ 。

计算法则是利用 SPSS 统计出的“均值”和“标准离差”, 采用“迭代剔除高值法”消除高值点的影响。即: 以均值加 3 倍标准离差为高值点的界限, 逐次剔除, 直至无高值点 (一般剔除三次), 根据最后一次统计结果中的“均值”和“标准离差”, 按前述公式计算该元素的背景值和异常下限。

利用 SPSS 软件对吉林珲春地区 15482 件土壤地球化学样品的 12 种元素测试数据进行统计分析, 结果表明: Au、Cu、Pb、Zn、Co、As、Sb 等元素含量的对数直方图具有明显的正态分布特点, 两种方法确定的异常下限值基本相同 (相对误差 3.1%~5.8%), 对化探数据处理及异常圈定无明显影响; Ag、W、Mo、Bi、Hg 等元素含量的对数直方图呈偏态分布, 两种方法确定的异常下限的相对误差较大 (15.4%~41.7%)。根据该区成矿地质条件、矿化特征和找矿目标矿种, 采用 SPSS 统计分析结果, 作为本区土壤地球化学测量不同元素的背景值、异常下限。

2 元素的相关性分析

为深入分析主成矿元素和成矿指示元素间的地球化学行为, 进一步划分元素的共生组合, 可利用 SPSS 软件对化探样品不同元素间的相关性进行分析。具体方法如下:

- (1) 打开 SPSS 主窗口, 点击“文件”, 打开不同元素测试结果的原始数据文件 (.xls);
- (2) 点击“分析”菜单, 在下拉菜单中选择“相关”中的“距离”选项;
- (3) 将对话框左侧的变量 (12 种元素) 依次选入右侧的“变量”框内, 首先点击“计算距离”选项中的“变量间”, 然后

点击“度量标准”选项中的“相似性”，程序自动选择“Pearson相似性”；

(4) 点击“确定”后程序自动生成不同元素间的相似性矩阵，即相关系数矩阵(.spv)。

表1为利用上述方法求得的吉林珲春地区土壤地球化学测量12种元素间的相关系数矩阵，从中可以看出：12种元素间相关性均较弱，相对而言，Cu与Bi、Au、Ag间具较弱的正相关性，W与As的相关性相对较强。该分析结果与该区小西南岔铜金矿床中富含铋矿物(刘向友等, 2006)、杨金沟钨矿石中含有毒砂(李峰等, 2008)等地质现象相符。

表 1 大六道沟勘查区土壤地球化学测量12种元素相关系数矩阵 (N=15842)

元素	Cu	Pb	Zn	Ag	Co	W	Mo	As	Sb	Bi	Hg	Au
Cu	1.000											
Pb	0.049	1.000										
Zn	0.068	0.405	1.000									
Ag	0.230	0.222	0.177	1.000								
Co	0.216	0.087	0.173	-0.099	1.000							
W	0.135	0.063	0.059	0.155	-0.056	1.000						
Mo	0.070	0.222	0.136	0.231	-0.085	0.124	1.000					
As	0.147	0.260	0.140	0.317	-0.012	0.333	0.173	1.000				
Sb	0.103	0.136	0.092	0.194	0.001	0.124	0.110	0.298	1.000			
Bi	0.326	0.205	0.107	0.345	-0.066	0.247	0.162	0.251	0.096	1.000		
Hg	0.001	0.290	0.203	0.164	-0.032	0.118	0.284	0.128	0.111	0.143	1.000	
Au	0.214	0.054	0.026	0.083	-0.011	0.125	0.033	0.104	0.047	0.263	0.038	1.000

3 元素的聚类分析

SPSS软件提供了两步聚类、K-均值聚类、系统聚类、决策树、判别分析等不同的数据分类方法。根据地球化学找矿的实际需要，本文以“系统聚类”为例，介绍利用SPSS软件对化探分析数据进行聚类分析的主要步骤和方法：

- (1) 打开SPSS主窗口，点击“文件”，打开不同元素测试结果的原始数据文件(.xls)；
- (2) 点击“分析”菜单，在下拉菜单中选择“分类”中的“系统聚类分析”选项；
- (3) 将对话框左侧的变量(元素)依次选入右侧的“变量”框内，在“分群”选项中选择“变量”，在“统计量”选项中可根据需要选择“相似性矩阵”；
- (4) 在“绘制”选项中选择“树状图”及其输出的方向(垂直或水平)等，而在“方法”选项中选择“Pearson相似性”等不同的系统聚类分析方法；
- (5) 点击“确定”，程序自动生成不同元素间的相似性矩阵、R型聚类分析元素地球化学谱系图(树状图)(.spv)。

在珲春地区土壤地球化学测量12种元素的聚类分析所得的元素地球化学谱系图中，截取距离系数10为指标，仅As、Bi、Pb、Cu、Ag、Hg为一独立小群，而其它个元素各自构成一个独立小群，与相关性分析结果基本一致。

除上述功能外，SPSS软件还可对化探数据进行因子分析，绘制折线图、饼图、面积图、散点图等不同图件，在此不再赘述。总之，通过地质技术人员的不断使用和功能开发，SPSS软件必将在地质数据的统计分析中发挥更广泛的作用。

参 考 文 献

- 黄燕波. 2007. SPSS在气象统计分析中的应用[J]. 气象研究与应用, 28(增刊1): 76-78.
- 刘向友, 康旭, 刘连登, 等. 2006. 小西南岔铜金矿床北山矿段控矿构造系统的新认识[J]. 黄金, 27(1): 7-9.
- 李峰, 姜建军. 2008. 吉林省珲春市杨金沟白钨矿床地质特征及成因机制探讨[J]. 吉林地质, 27(2): 22-25.