

# 夏皮罗—威尔克和李立福正态性检定法\*

俞 钟 行

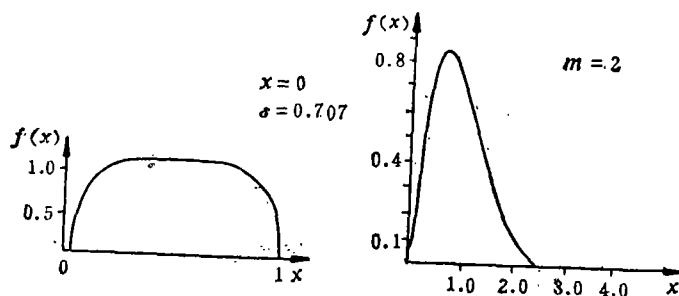
(地矿部物化探研究所)

文中介绍的两种正态性检定法于 60 年代中后期提出, 具有检出能力强, 送检标本少, 使用方便的优点, 十分适合我国工程技术人员使用。

**关键词:** 夏皮罗—威尔克法; 李立福法; 正态性检定

在没有经受次生变化的岩石中, 密度和弹性波传播速度参数的分布遵从正态分布, 磁化率和磁化强度则遵从对数正态分布。为准确判断实测分布是否符合正态分布或对数正态分布, 很多著述都介绍了用正态概率纸检定法及 $\alpha^2$ 检定法。正态概率纸检定法使用方便, 但因为存在作图误差且点对直线的偏离程度没有一个定量的标准, 甚至因人而异, 所

以只能算一个粗略的正态性检定法。古典的 $\alpha^2$ 检定法适用于任意分布, 但对正态性的实际检验效果不太突出。国外在 70 年代中期发表的技术报告表明,  $\alpha^2$ 检定法对典型的非正态分布, 如约翰逊分布及威布尔分布检出能力仅为 27% 及 17% (子样大小  $n = 50$ , 危险率  $\alpha = 10.0\%$ ), 而夏皮罗—威尔克检定法则分别达 85% 及 59% (见图)。此外,  $\alpha^2$  检定法使



## 约翰逊分布

用条件较高。第一,  $\alpha^2$  检定法要求子样标本数  $\geq 50$ 。由于野外采样受到地质体的大小和出露情况限制, 除很大的岩体和很厚的地层且出露良好外, 大都很难采到 50 个样。采样要求的随机性, 规定了即使勉强凑数达到 50 个标本也并无意义。第二,  $\alpha^2$  检定法要求各个区间的理论频数 (有时还要求实际频数) 都不小于 5, 最好在 10 以上, 否则就要重新分组或适当合并区间, 这给利用计算机使用此法增加了难度。

## 威布尔分布

S·S·夏皮罗及 M·B·威尔克于 1965 年提出的正态性检定法, 目前被认为是最有效的; 李立福于 1969 年提出的正态性检定法也相当有效。这两种方法很适合我国工程技术人员、尤其是地质勘探界的工程技术人员使用。下面, 结合实例对这两种方法给出说明。

某地区标本中 15 块花岗岩类标本的密度及磁化率值测定如下:

| 编 号 | 1      | 2     | 3     | 4     | 5      | 6      | 7      | 8      | 9      | 10    | 11     | 12    | 13    | 14     | 15    |
|-----|--------|-------|-------|-------|--------|--------|--------|--------|--------|-------|--------|-------|-------|--------|-------|
| 密 度 | 2.621  | 2.634 | 2.571 | 2.643 | 2.550  | 2.589  | 2.638  | 2.593  | 2.628  | 2.581 | 2.650  | 2.623 | 2.487 | 2.581  | 2.665 |
| 磁化率 | 1425.6 | 904.4 | 523.7 | 8.836 | 863.26 | 164.97 | 17.766 | 11.644 | 12.386 | 4.878 | 531.79 | 1490  | 6.896 | 192.24 | 13.6  |

\* 已用 BASIC 和 FORTRAN77 语言编成程序, 需要者可与作者 (河北省廊坊市) 联系。

试验密度是否符合正态分布、磁化率是否符合对数正态分布。

下面介绍检定密度是否符合正态分布的步骤，对磁化率检定仅给出结果。

### 夏皮罗-威尔克法

(一) 按子样数据的大小，从小到大排列成  $x_{(1)}, x_{(2)}, \dots, x_{(n-1)}, x_{(n)}$ 。

本例中，密度值排成 2.487, 2.550, 2.571, 2.581, 2.581, 2.589, 2.593, 2.621, 2.623, 2.628, 2.634, 2.638, 2.643, 2.650, 2.665。

(二) 查夏皮罗-威尔克的  $a_{in}$  表 (见附表)，查出对应于  $n$  值的诸  $a_{in}$  值共

$$\begin{cases} (n+1)/2 & (n \text{ 为奇数时}) \text{ 个} \\ n/2 & (n \text{ 为偶数时}) \end{cases}$$

这里  $n$  是测定值个数， $i$  是按从小到大排定后的顺序数。本例中  $n=15$ ，查出：

$$a_{1,15} = 0.5150, a_{2,15} = 0.3306, a_{3,15} = 0.2495, a_{4,15} = 0.1878, a_{5,15} = 0.1353, a_{6,15} = 0.0880, a_{7,15} = 0.0433, a_{8,15} = 0.0000。$$

本文附表仅给出  $n \leq 30$  时的检定系数。实际上已编出  $n \leq 50$  时严格算得的检定系数及  $50 \leq n \leq 100$  时近似算得的检定系数。

(三) 计算

$$W = \frac{\left\{ \sum_i a_{in} (x_{(n-i+1)} - x_{(i)}) \right\}^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2}$$

这里分子中的  $\sum_i$ ，当  $n$  为偶数时为  $\sum_{i=1}^{n/2}$ ，当  $n$  为奇数时为  $\sum_{i=1}^{(n+1)/2}$ 。本例中

$$\begin{aligned} \sum_i a_{in} (x_{(n-i+1)} - x_{(i)}) &= 0.5150(2.665 - 2.487) \\ &+ 0.3306(2.650 - 2.550) + 0.2495(2.643 - 2.571) \\ &+ 0.1878(2.638 - 2.581) + 0.1353(2.634 - 2.581) \\ &+ 0.0880(2.628 - 2.589) + 0.0433(2.623 - 2.593) \\ &+ 0.0000(2.621 - 2.621) = 0.1653, \end{aligned}$$

$$\sum_{i=1}^{15} (x_{(i)} - \bar{x})^2 = 0.02974,$$

$$W = -\frac{0.1653^2}{0.02974} = 0.9188。$$

(四) 给定危险率  $\alpha$ ，查夏皮罗-威尔克的  $W(n, \alpha)$  表中相应的  $W(n, \alpha)$  值。本例中， $n=15$ ，取  $\alpha=5.0\%$ ，得  $W(15, 5.0\%) = 0.881$ 。

(五) 如  $W \leq W(n, \alpha)$  则以危险率  $\alpha$  否定原子样母体为正态性的假设。如  $W > W(n, \alpha)$  则不能以危险率  $\alpha$  否定原子样母体为正态性的假设。本例中  $W = 0.9188 > 0.881 = W(15, 5.0\%)$ ，故不能以危险率  $5.0\%$  否定原子样母体是正态的。

### 李立福法

(一) 同上法 (一)。

(二) 计算子样均值  $\bar{x}$  和子样方差  $s^2$  及  $s$ 。本例中  $\bar{x} = 2.6036$ ， $s^2 = 0.002124$ ， $s = 0.046087$ 。

(三) 按下表计算各  $u_{(i)} = \frac{x_{(i)} - \bar{x}}{s}$ ， $\hat{F}(x_{(i)}) = \frac{i}{n}$ ， $F(x_{(i)}) = \int_{-\infty}^{u_{(i)}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$  及  $F(x_{(i)}) - \hat{F}(x_{(i)})$ 。

其中  $\hat{F}(x_{(i)}) = \frac{i}{n}$  中的  $n$  是标本数， $i$  是按从小到大排定后的顺序数。 $F(x_{(i)})$  的值或可查标准正态分布表，或可利用哈斯廷近似式计算 (用电算)。

$$F(x_{(i)}) = \begin{cases} 1 - \frac{1}{2}(1 + a_1 u_{(i)} + \dots + a_6 u_{(i)}^6)^{-16} & u_{(i)} \geq 0 \\ \frac{1}{2}(1 + a_1 u_{(i)} + \dots + a_6 u_{(i)}^6)^{-16} & u_{(i)} < 0 \end{cases}$$

$$u_{(i)} \geq 0$$

$$u_{(i)} < 0。$$

其中  $a_1 = 0.049867347$ ， $a_2 = 0.021141061$ ，

$$a_3 = 0.0032776263, a_4 = 0.0000380036,$$

$$a_5 = 0.0000488966, a_6 = 0.0000053836。$$

李立福法计算表

| 序号<br>$i$ | $x_{(i)}$ | $u_{(i)}$ | $F(x_{(i)})$ | $\hat{F}(x_{(i)})$ | $F(x_{(i)}) - \hat{F}(x_{(i)})$ |
|-----------|-----------|-----------|--------------|--------------------|---------------------------------|
| 1         | 2.487     | -2.5299   | 0.0057       | 0.0667             | -0.0610                         |
| 2         | 2.550     | -1.1630   | 0.1224       | 0.1333             | -0.0109                         |
| 3         | 2.571     | -0.7073   | 0.2397       | 0.2000             | 0.0397                          |
| 4         | 2.581     | -0.4903   | 0.3119       | 0.2667             | 0.0453                          |
| 5         | 2.581     | -0.4903   | 0.3119       | 0.3333             | -0.0214                         |
| 6         | 2.589     | -0.3167   | 0.3757       | 0.4000             | -0.0243                         |
| 7         | 2.593     | -0.2299   | 0.4091       | 0.4667             | -0.0576                         |
| 8         | 2.621     | 0.2775    | 0.6471       | 0.5333             | 0.1138                          |
| 9         | 2.623     | 0.4209    | 0.6631       | 0.6000             | 0.0631                          |
| 10        | 2.628     | 0.5294    | 0.7018       | 0.6667             | 0.0351                          |
| 11        | 2.634     | 0.6596    | 0.7452       | 0.7333             | 0.0119                          |
| 12        | 2.638     | 0.7464    | 0.7723       | 0.8000             | -0.0277                         |
| 13        | 2.643     | 0.8549    | 0.8037       | 0.8667             | -0.0630                         |
| 14        | 2.650     | 1.0067    | 0.8429       | 0.9333             | -0.0903                         |
| 15        | 2.665     | 1.3322    | 0.9086       | 1.0000             | -0.0914                         |

夏皮罗-威尔克的 $a_{in}$ 系数表

| $n$ | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      | 10     |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1   | 0.7071 | 0.7071 | 0.6872 | 0.6646 | 0.6431 | 0.6233 | 0.6052 | 0.5888 | 0.5739 |
| 2   | —      | 0.0000 | 0.1677 | 0.2413 | 0.2806 | 0.3031 | 0.3164 | 0.3244 | 0.3291 |
| 3   | —      | —      | —      | 0.0000 | 0.0875 | 0.1401 | 0.1743 | 0.1976 | 0.2141 |
| 4   | —      | —      | —      | —      | —      | 0.0000 | 0.0561 | 0.0947 | 0.1224 |
| 5   | —      | —      | —      | —      | —      | —      | —      | 0.0000 | 0.0399 |

| $n$ | 11     | 12     | 13     | 14     | 15     | 16     | 17     | 18     | 19     | 20     |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1   | 0.5601 | 0.5475 | 0.5359 | 0.5251 | 0.5150 | 0.5056 | 0.4968 | 0.4886 | 0.4808 | 0.4734 |
| 2   | 0.3315 | 0.3325 | 0.3325 | 0.3318 | 0.3306 | 0.3290 | 0.3273 | 0.3253 | 0.3232 | 0.3211 |
| 3   | 0.2260 | 0.2347 | 0.2412 | 0.2460 | 0.2495 | 0.2521 | 0.2540 | 0.2553 | 0.2561 | 0.2665 |
| 4   | 0.1429 | 0.1586 | 0.1707 | 0.1802 | 0.1878 | 0.1939 | 0.1988 | 0.2027 | 0.2059 | 0.2085 |
| 5   | 0.0695 | 0.0922 | 0.1099 | 0.1240 | 0.1353 | 0.1447 | 0.1524 | 0.1587 | 0.1641 | 0.1686 |
| 6   | 0.0000 | 0.0303 | 0.0539 | 0.0727 | 0.0880 | 0.1005 | 0.1109 | 0.1197 | 0.1271 | 0.1334 |
| 7   | —      | —      | 0.0000 | 0.0240 | 0.0433 | 0.0593 | 0.0725 | 0.0837 | 0.0932 | 0.1013 |
| 8   | —      | —      | —      | —      | 0.0000 | 0.0196 | 0.0359 | 0.0496 | 0.0612 | 0.0711 |
| 9   | —      | —      | —      | —      | —      | —      | 0.0000 | 0.0303 | 0.0303 | 0.0422 |
| 10  | —      | —      | —      | —      | —      | —      | —      | —      | 0.0000 | 0.0140 |

| $n$ | 21     | 22     | 23     | 24     | 25     | 26     | 27     | 28     | 29     | 30     |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1   | 0.4643 | 0.4590 | 0.4542 | 0.4493 | 0.4450 | 0.4407 | 0.4366 | 0.4328 | 0.4291 | 0.4254 |
| 2   | 0.3185 | 0.3156 | 0.3126 | 0.3098 | 0.3069 | 0.3043 | 0.3018 | 0.2992 | 0.2968 | 0.2944 |
| 3   | 0.2578 | 0.2571 | 0.2563 | 0.2554 | 0.2513 | 0.2533 | 0.2522 | 0.2510 | 0.2499 | 0.2487 |
| 4   | 0.2119 | 0.2131 | 0.2139 | 0.2145 | 0.2148 | 0.2151 | 0.2152 | 0.2151 | 0.2150 | 0.2148 |
| 5   | 0.1736 | 0.1764 | 0.1787 | 0.1807 | 0.1822 | 0.1836 | 0.1848 | 0.1857 | 0.1864 | 0.1870 |
| 6   | 0.1399 | 0.1443 | 0.1480 | 0.1512 | 0.1539 | 0.1563 | 0.1584 | 0.1601 | 0.1616 | 0.1630 |
| 7   | 0.1092 | 0.1150 | 0.1201 | 0.1245 | 0.1283 | 0.1316 | 0.1346 | 0.1372 | 0.1395 | 0.1415 |
| 8   | 0.0804 | 0.0878 | 0.0941 | 0.0997 | 0.1046 | 0.1089 | 0.1128 | 0.1162 | 0.1192 | 0.1219 |
| 9   | 0.0530 | 0.0618 | 0.0696 | 0.0764 | 0.0823 | 0.0876 | 0.0923 | 0.0965 | 0.1002 | 0.1036 |
| 10  | 0.0263 | 0.0368 | 0.0459 | 0.0539 | 0.0610 | 0.0672 | 0.0728 | 0.0778 | 0.0822 | 0.0862 |
| 11  | 0.0000 | 0.0122 | 0.0228 | 0.0321 | 0.0403 | 0.0476 | 0.0540 | 0.0598 | 0.0650 | 0.0697 |
| 12  | —      | —      | 0.0000 | 0.0107 | 0.0200 | 0.0284 | 0.0358 | 0.0424 | 0.0483 | 0.0537 |
| 13  | —      | —      | —      | —      | 0.0000 | 0.0094 | 0.0178 | 0.0253 | 0.0320 | 0.0381 |
| 14  | —      | —      | —      | —      | —      | —      | 0.0000 | 0.0084 | 0.0159 | 0.0227 |
| 15  | —      | —      | —      | —      | —      | —      | —      | —      | 0.0000 | 0.0076 |

 $W(n, \alpha)$  表

| $n$               | 3     | 4     | 5     | 6     | 7     | 8     | 9     | 10    | 11    | 12    | 13    | 14    |
|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\alpha = 10.0\%$ | 0.789 | 0.792 | 0.806 | 0.826 | 0.838 | 0.851 | 0.859 | 0.869 | 0.876 | 0.883 | 0.889 | 0.895 |
| $\alpha = 5.0\%$  | 0.767 | 0.748 | 0.762 | 0.788 | 0.803 | 0.818 | 0.829 | 0.842 | 0.850 | 0.859 | 0.866 | 0.874 |
| $\alpha = 1.0\%$  | 0.753 | 0.687 | 0.686 | 0.713 | 0.730 | 0.749 | 0.764 | 0.781 | 0.792 | 0.805 | 0.814 | 0.825 |

续表

| $n$             | 15    | 16    | 17    | 18    | 19    | 20    | 21    | 22    | 23    | 24    | 25    | 26    |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\alpha=10.0\%$ | 0.901 | 0.906 | 0.910 | 0.914 | 0.917 | 0.920 | 0.923 | 0.926 | 0.928 | 0.930 | 0.931 | 0.933 |
| $\alpha=5.0\%$  | 0.881 | 0.887 | 0.892 | 0.897 | 0.901 | 0.905 | 0.908 | 0.911 | 0.914 | 0.916 | 0.918 | 0.920 |
| $\alpha=1.0\%$  | 0.835 | 0.844 | 0.851 | 0.858 | 0.863 | 0.868 | 0.873 | 0.878 | 0.881 | 0.884 | 0.888 | 0.891 |

| $n$             | 27    | 28    | 29    | 30    | 31    | 32    | 33    | 34    | 35    | 36    | 37    | 38    |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\alpha=10.0\%$ | 0.935 | 0.936 | 0.937 | 0.939 | 0.940 | 0.941 | 0.942 | 0.943 | 0.944 | 0.945 | 0.946 | 0.947 |
| $\alpha=5.0\%$  | 0.923 | 0.924 | 0.926 | 0.927 | 0.929 | 0.930 | 0.931 | 0.933 | 0.934 | 0.935 | 0.936 | 0.938 |
| $\alpha=1.0\%$  | 0.894 | 0.896 | 0.898 | 0.900 | 0.902 | 0.904 | 0.906 | 0.908 | 0.910 | 0.912 | 0.914 | 0.916 |

| $n$             | 39    | 40    | 41    | 42    | 43    | 44    | 45    | 46    | 47    | 48    | 49    | 50    |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| $\alpha=10.0\%$ | 0.948 | 0.949 | 0.950 | 0.951 | 0.951 | 0.952 | 0.953 | 0.953 | 0.954 | 0.954 | 0.955 | 0.955 |
| $\alpha=5.0\%$  | 0.939 | 0.940 | 0.941 | 0.942 | 0.943 | 0.944 | 0.945 | 0.945 | 0.946 | 0.947 | 0.947 | 0.947 |
| $\alpha=1.0\%$  | 0.917 | 0.919 | 0.920 | 0.922 | 0.923 | 0.924 | 0.926 | 0.927 | 0.928 | 0.929 | 0.929 | 0.930 |

李立福方法的 $D(n, \alpha)$ 表

| $n$   | $\alpha$ (%)             |                          |                          |
|-------|--------------------------|--------------------------|--------------------------|
|       | 10                       | 5.0                      | 1.8                      |
| 4     | 0.352                    | 0.381                    | 0.417                    |
| 5     | 0.315                    | 0.337                    | 0.405                    |
| 6     | 0.294                    | 0.319                    | 0.364                    |
| 7     | 0.276                    | 0.300                    | 0.348                    |
| 8     | 0.261                    | 0.285                    | 0.331                    |
| 9     | 0.249                    | 0.271                    | 0.311                    |
| 10    | 0.239                    | 0.258                    | 0.294                    |
| 11    | 0.230                    | 0.249                    | 0.284                    |
| 12    | 0.223                    | 0.242                    | 0.275                    |
| 13    | 0.214                    | 0.234                    | 0.268                    |
| 14    | 0.207                    | 0.227                    | 0.261                    |
| 15    | 0.201                    | 0.220                    | 0.257                    |
| 16    | 0.195                    | 0.213                    | 0.250                    |
| 17    | 0.189                    | 0.206                    | 0.245                    |
| 18    | 0.184                    | 0.200                    | 0.239                    |
| 19    | 0.179                    | 0.195                    | 0.235                    |
| 20    | 0.174                    | 0.190                    | 0.231                    |
| 25    | 0.165                    | 0.180                    | 0.203                    |
| 30    | 0.144                    | 0.161                    | 0.187                    |
| $>30$ | $\frac{0.805}{\sqrt{n}}$ | $\frac{0.886}{\sqrt{n}}$ | $\frac{1.031}{\sqrt{n}}$ |

本例计算结果如下:

取 $D = \max |F(x_{(i)}) - \hat{F}(x_{(i)})|$ 。

本例中 $D = 0.1138$ 。

(四) 给定危险率 $\alpha$ , 查李立福方法的 $D(n, \alpha)$ 表中相应的 $D(n, \alpha)$ 值。本例中 $n = 15$ , 取 $\alpha = 5.0\%$ ,

得 $D(15, 5.0\%) = 0.220$ 。

(五) 如 $D \geq D(n, \alpha)$ 则以危险率 $\alpha$ 否定原子样母体为正态性的假设。如 $D < D(n, \alpha)$ 则不能以危险率 $\alpha$ 否定原子样母体为正态性的假设。本例中 $D = 0.1138 < 0.220 = D(15, 5.0\%)$ , 故不能以

危险率5.0%否定原子样母体是正态的。

对磁化率的对数值作正态检验,结果如下,

$$D = 0.2331 > 0.220 = D(15, 5.0\%)$$

$$W = 0.8514 < 0.881 = W(15, 5.0\%)$$

故李立福法及夏皮罗-威尔克法均以5.0%的危险率否定原子样母体是正态的。

使用夏皮罗-威尔克方法及李立福方法有两点需要注意。

(一) 当 $W > W(n, \alpha)$ 时,即从“不能以危险率 $\alpha$ 否定原子样母体为正态性”得出“原子样母体为正态”的结论是不恰当的。当 $D < D(\alpha, n)$ 时同样如此。但如果原子样母体的正态性假设有相当

的物理根据,则可以维持该子样母体为正态分布的假设。

(二) 危险率 $\alpha$ 值不能取太小。因为危险率小时固然减少了原母体为正态时予以否定的概率,但也增大了原母体不为正态时不予以否定的概率。故 $\alpha$ 值一般取10.0%或5.0%。

### 参 考 文 献

[1] 浙江大学:《概率论与数理统计》,高等教育出版社,1985年。

[2] 何国伟:《误差分析方法》,国防工业出版社,1978年。

[3] 何国伟:《计量与测试的分析方法》,国防工业出版社,1985年。

## Shapiro-Wilk and Lilifu Variance Testing

### Techniques for Normality

Yu Zhongxing

In the present paper two variance testing techniques for normality are briefly described: one is the Shapiro-Wilk method and the other is the Lilifu method. They were first introduced in periodicals in the middle-late period of the sixties. High detectivity, small sample size requirement and being easy to use are their strong points. It is believed that they are helpful to geophysicists. Examples are cited for illustrations.

## 地质经济学术讨论会及推广南京白云矿微机在矿山地质中应用会在白云石矿召开

〔本刊讯〕 由中国金属学会冶金地质学会矿山地质及地质经济学术委员会主持召开的“1988年地质经济学术讨论会及推广南京白云矿微机在矿山地质工作中应用会议”,于1988年5月25~29日在南京白云石矿召开。来自全国地质勘探部门、生产矿山、科研院所和大专院校等54个单位的80位代表参加了会议。

会议共收的论文34篇,其中:地质经济25篇,微机在矿山地质工作中的应用7篇,其他2篇。会议评选出优秀论文9篇。

从会议交流的地质经济方面的论文来看,既有

矿床经济评价、矿石合理综合利用、矿床品位指标优化和采选指标优化方面的论文,也有矿石质量管理、三级矿量管理,以及地质勘探等经济管理方面的文章。在微机应用方面,除了介绍南京白云石矿的科研与应用成果和经验外,还介绍了其他一些单位的研究成果。特别是一些新的数学方法和手段在地质经济研究中应用的论文,对与会代表有很大启发。微机在矿山地质工作中的应用,以及一些软件包的现场表演,将对在矿山地质工作中推广应用微机起促进作用。

〔冶金地质学会办公室供稿〕