

# 秩数分级相似预报法

彭 沛 黄丽影

(数学系 78 级)

## 一、 引 言

本文在吴绍敏老师的指导下, 根据吴绍敏老师提出的以《秩距离》为相似度的相似预报法及分级相似预报法的基本思想, 提出一种新的多级相似预报法。

## 二、秩数分级相似预报法

**1. 基本思想** 这个方法的主要思想是根据预报对象与预报因子的秩相关关系找出预报对象的分类规律和各因子的分类规律的相似性, 然后利用这个相似性作预报。假设我们挑选与预报对象  $y$  有秩相关的因子  $x_1, x_2, \dots, x_s$  参加预报。设  $y$  和  $x_i, i=1, 2, \dots, s$  都有  $N$  年的历史资料  $y_1, y_2, \dots, y_N; x_{i1}, x_{i2}, \dots, x_{iN}, (i=1, 2, \dots, s)$ , 将  $y$  的历史资料  $y_1, y_2, \dots, y_N$  按从小到大排列的秩数记为  $b_1, b_2, \dots, b_N$ , 若  $x_i$  与  $y$  正相关, 则用  $a_{ij}$  表示  $x_i$  的第  $j$  个数据  $x_{ij}$  按  $x_i$  的资料从小到大排列的秩数, 若  $x_i$  与  $y$  负相关, 则  $a_{ij}$  表示  $x_i$  按  $x_i$  的资料从大到小排列的秩数。

将  $y$  的历史资料分为  $m$  级, 使各级之间资料的趋势有明显的差别, 记为  $y^{(1)}, y^{(2)}, \dots, y^{(m)}$ , 其中  $y^{(k)} = (y_1^{(k)}, y_2^{(k)}, \dots, y_{N_k}^{(k)}) (k=1, 2, \dots, m)$ , 而  $N_1 + N_2 + \dots + N_m = N$ ; 对应于  $y$  的等级, 每个因子的实测数据  $x_{ij} (i=1, 2, \dots, s; j=1, 2, \dots, N)$  也相应地分为  $m$  级:

$$\begin{aligned}x_i^{(1)} &= (x_{i1}^{(1)}, x_{i2}^{(1)}, \dots, x_{iN_1}^{(1)}) \\x_i^{(2)} &= (x_{i1}^{(2)}, x_{i2}^{(2)}, \dots, x_{iN_2}^{(2)}) \\&\vdots \\x_i^{(m)} &= (x_{i1}^{(m)}, x_{i2}^{(m)}, \dots, x_{iN_m}^{(m)})\end{aligned}$$

此时  $y$  及  $x_i$  的秩数也分别相应地分成了  $m$  级, 记为  $b^{(k)} = (b_1^{(k)}, b_2^{(k)}, \dots, b_{N_k}^{(k)})$ ;  
 $a_i^{(k)} = (x_{i1}^{(k)}, x_{i2}^{(k)}, \dots, x_{iN_k}^{(k)}) (k=1, 2, \dots, m; i=1, 2, \dots, s)$

由于  $y$  与  $x_1, x_2, \dots, x_s$  的秩数相关关系,  $b^{(k)}$  与  $a_i^{(k)}, (i=1, 2, \dots, s)$  应有相同的趋势, 即若记  $\bar{b}^{(k)} = \frac{1}{N_k} \sum_{j=1}^{N_k} b_j^{(k)}, \bar{a}_i^{(k)} = \frac{1}{N_k} \sum_{j=1}^{N_k} a_{ij}^{(k)} (k=1, 2, \dots, m)$  则若  $\bar{b}^{(k-1)}$

小于  $\bar{b}^{(k)}$  那么  $\bar{a}_i^{(k-1)}$  也应小于  $\bar{a}_i^{(k)}$  , 反之亦然,  $(i=1, 2, \cdots, s), (k=2, 3, \cdots, m)$ 。对因子当前值  $(x_{1,0}, x_{2,0}, \cdots, x_{s,0})$  , 将它与历史资料作比较估计, 它的秩数记为  $(\hat{a}_{1,0}, \hat{a}_{2,0}, \cdots, \hat{a}_{s,0})$  , 建立预报测度

$$u(k)=\sum_{i=1}^s\left|\hat{a}_{i,0}-\bar{a}_i^{(k)}\right|, \quad k=1,2, \cdots, m$$

若  $u(l)=\min_{1 \leq k \leq m} u(k)$  ,说明因子当前值属于  $l$  级这一类, 因此我们有理由预报  $y$  应属于  $l$  级。

至此, 我们已经预报了  $y$  的趋势, 再将历史资料中凡属于第  $l$  级的资料列出来, 按相似法, 以《秩距离》为尺度将当前的因子值与之比较, 找出最相似的时段作预报。此方法既可预报  $y$  的趋势, 又可预报  $y$  的数值, 计算量较小并且  $y$  不受模型限制。

## 2. 具体实施步骤

(1) 预报对象  $y$  分成  $m$  级, 挑选与  $y$  秩相关并且使

$$\sigma(i)=\frac{\sum_{k=2}^m\left(\bar{a}_i^{(k)}-\bar{a}_i^{(k-1)}\right)^2}{\sum_{k=1}^m \sum_{j=1}^{N_k}\left(\bar{a}_{ij}^{(k)}-\bar{a}_i^{(k-1)}\right)}$$

达到较大的因子  $x_1, x_2, \cdots, x_s$  参加预报。即选出的因子其秩数在同一级波动较小, 并且因子各级的秩数均值相差较大。将  $y$  和各因子的资料及秩数列成表 1。

表 1

类 别	$y$	$b$	$x_1$	$a_1$	.....	$x_s$	$a_s$
1	$y_1^{(1)}$	$b_1^{(1)}$	$x_{11}^{(1)}$	$a_{11}^{(1)}$	.....	$x_{s1}^{(1)}$	$a_{s1}^{(1)}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	$y_{N1}^{(1)}$	$b_{N1}^{(1)}$	$x_{1N1}^{(1)}$	$a_{1N1}^{(1)}$		$x_{sN1}^{(1)}$	$a_{sN1}^{(1)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	.....	$\vdots$	$\vdots$
$m$	$y_1^{(m)}$	$b_1^{(m)}$	$x_{11}^{(m)}$	$a_{11}^{(m)}$	.....	$x_{s1}^{(m)}$	$a_{s1}^{(m)}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	$y_{Nm}^{(m)}$	$b_{Nm}^{(m)}$	$x_{1Nm}^{(m)}$	$a_{1Nm}^{(m)}$		$x_{sNm}^{(m)}$	$a_{sNm}^{(m)}$

(2) 求  $\bar{b}^{(k)}, \bar{a}_i^{(k)}, (k=1, 2, \cdots, m, i=1, 2, \cdots, s)$  列成表 2。

(3) 建立预报测度:

$$u(k)=\sum_{i=1}^s\left|\hat{a}_{i,0}-\bar{a}_i^{(k)}\right|, \quad (k=1,2, \cdots, m) \tag{1}$$

(4) 预报  $y$  的趋势: 若  $u(l)=\min_{1 \leq k \leq m} u(k)$  , 预报  $y$  属于  $l$  级。

(5) 效果检验: 将往年因子的资料  $x_{ij}$  秩数  $(i=1, 2, \cdots, s) (j=1, 2, \cdots, N)$  代入

(1) 式作回报。

表 2

级 别 \ 均 值	$\bar{\delta}^{(k)}$	$\bar{a}_1^{(k)}$	.....	$\bar{a}_s^{(k)}$
1	$\bar{\delta}^{(1)}$	$\bar{a}_1^{(1)}$	.....	$\bar{a}_s^{(1)}$
$\vdots$	$\vdots$	$\vdots$	.....	$\vdots$
m	$\bar{\delta}^{(m)}$	$\bar{a}_1^{(m)}$	.....	$\bar{a}_s^{(m)}$
当 前		$\delta_{10}$	.....	$\delta_s^{(m)}$

(6) 将在第  $l$  级起主导作用的有关因子属于第  $l$  级的历史资料列出来, 以《秩距离》为相似度作相似定值预报。

**3. 实例** 根据 1957—1981 年的历史资料, 利用《秩数分级相似预报法》预报晋江地区全区七个站 7—9 月份总降水量  $y$ 。

(1) 预报对象: 晋江地区 7—9 月总降水量  $y$ 。资料见表 3。

(2) 预报因子: 用秩相关系数检验法选出五个预报因子:

$x_1$ : 亚洲 500m 6 月平均环流指数  $\left[ \frac{45-65^\circ N}{60-150^\circ E} \right]$  (1 月  $I_m$ ),  $R_{1y} = -0.4577$ 。

$x_2$ : 9 月  $\frac{60-80^\circ E}{60-65^\circ N} \sum_1^5 H$ ,  $R_{2y} = 0.4956$ 。

$x_3$ : 7—9 月亚欧 500m 6 月平均环流指数  $\left[ \frac{0-150^\circ E}{45-65^\circ N} \right]$ ,  $R_{3y} = -0.4221$

$x_4$ : 5 月份北半球 500 m 6 月平均极涡中心强度,  $R_{4y} = 0.6637$ 。

$x_5$ : 4 月份上旬永春降水量。  $R_{5y} = -0.5123$ 。其历史资料见表 3。

(3) 分级:

$y^{(1)}$ :  $0 \leq y < 2600$  称为第一级。

$y^{(2)}$ :  $2600 \leq y < 4000$  称为第二级。

$y^{(3)}$ :  $y \geq 4000$  称为第三级。

分级后将历史资料的秩数列成表 4。

(4) 求各级资料的均值  $\bar{\delta}^{(k)}$ ,  $\bar{a}_i^{(k)}$  ( $k=1, 2, , 3; i=1, 2, 3, 4$ ), 具体计算见表 5。

(5) 建立预报测度并作趋势预报:

$$u(1) = \sum_{j=1}^5 |\bar{a}_{j1} - \delta_{j0}| = 27.44$$

$$u(2) = \sum_{j=1}^5 |\bar{a}_{j2} - \delta_{j0}| = 26.33$$

$$u(3) = \sum_{j=1}^5 |\bar{a}_{j3} - \bar{a}_{j0}| = 30.68$$

$$\therefore u(2) = \min_{1 \leq k \leq 3} u(k)$$

$\therefore$  预报  $\hat{y}_{82}$  属于第二级即  $2600 \leq \hat{y} < 4000(\text{mm})$

(6) 效果检验: 见表 7。从表 7 中易见回报准确率  $P = \frac{20}{25} = 80\%$ 。

(7) 作《秩距离》相似定值预报:

将预报二级时起主导作用的因子属于第二级的有关资料列出来(本例无起突出作用的因子), 求秩距离

$$d_l = \sum_{i=1}^5 a_l^{(i)}, \quad (l=1, 2, \dots, 9), \quad \text{其中 } a_l^{(i)} \text{ 为 } x_{il}' = |x_{il} - x_{i0}|, (l=1, 2, \dots, 9)$$

按从小到大排列的秩数, 具体计算见表 6。

$\therefore d_8 = 14 = \min d_l$  而对应的年号为 80

$\therefore$  作预报  $\hat{y}_{82} = y_{80} = 3371(\text{mm})$ 。

表 3

年 份	y		x <sub>1</sub>		x <sub>2</sub>		x <sub>3</sub>		x <sub>4</sub>		x <sub>5</sub>	
	数据	秩数	数据	秩数	数据	秩数	数据	秩数	数据	秩数	数据	秩数
57	1718	1	0.87	2	196	2	3.17	4	518	4	158.3	9
58	6951	24	0.55	21	306	25	2.38	24.5	528	22	85.7	21
59	4907	22	0.70	9.5	192	1	2.94	11	527	19	117.3	16
60	3559	16	0.94	1	216	4	2.56	20	528	22	158.9	8
61	7194	25	0.56	20	275	21	2.69	19	528	22	78.7	22
62	2522	7	0.77	5	232	5.5	3.11	6.5	523	7.5	153.6	11
63	4418	19	0.64	15	268	17.5	3.11	6.5	526	17.5	105.0	18
64	2606	9	0.67	12.5	265	16	2.88	14	524	11.5	108.2	17
65	2821	12	0.67	12.5	259	14	2.93	12	516	1	65.6	23
66	2540	8	0.75	6	245	11.5	3.04	9	522	2	104.0	19
67	2015	3	0.74	7.5	245	11.5	2.50	21.5	509	3	154.7	10
68	1732	2	0.74	7.5	239	8	3.39	7	524	11.5	189.1	4
69	2884	13	0.64	15	237	7	2.74	18	523	7.5	53.6	24
70	4423	20	0.50	24	278	22	2.47	23	525	15	172.2	6
71	2741	11	0.68	11	301	24	2.75	17	523	7.5	168.5	7
72	5477	23	0.57	19	284	23	2.78	16	524	11.5	50.9	25
73	4543	21	0.62	17	232	5.5	2.38	24.5	525	15	148.3	13
74	2092	4	0.48	25	202	3	2.89	13	520	5	316.5	1
75	4057	18	0.52	23	273	20	2.80	15	526	17.5	129.6	15
76	3793	17	0.61	18	262	15	3.23	3	528	22	141.6	14
77	2366	5	0.78	3.5	240	9	2.50	21.5	524	11.5	181.7	5
78	2655	10	0.53	22	247	13	3.12	5	528	22	86.9	20
79	2500	6	0.70	15	241	10	3.35	2	523	7.5	190.9	3
80	3371	14	0.64	9.5	269	19	3.06	8	525	15	151.7	12
81	3433	15	0.78	3.5	268	17.5	3.00	10	530	25	273.3	2
当前值			0.71	9	272	20	2.54	20	523	7.5	150.9	12

表 4

年 份	$b$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
57	1	2	2	4	4	9
68	2	7.5	8	1	11.5	4
67	3	7.5	11.5	21.5	3	10
74	4	25	3	13	5	1
77	5	3.5	9	21.5	11.5	5
79	6	15	10	2	7.5	3
62	7	5	5.5	6.5	7.5	11
66	8	6	11.5	9	2	19
64	9	12.5	16	14	11.5	17
78	10	22	13	5	22	20
71	11	11	24	17	7.5	7
65	12	12.5	14	12	1	23
69	13	15	7	18	7	24
80	14	9.5	19	8	15	12
81	15	3.5	17.5	10	25	2
60	16	1	4	20	22	8
76	17	18	15	3	22	14
75	18	23	20	15	17.5	15
63	19	15	17.5	6.5	17.5	18
70	20	24	22	23	15	6
73	21	17	5	24.5	15	13
59	22	9.5	1	11	19	16
72	23	19	23	16	11.5	25
58	24	21	25	24.5	22	21
61	25	20	21	19	22	22
当前值		9	20	20	7.5	12

表 5

$i$	$b_i$	$\bar{a}_{1i}$	$\bar{a}_{2i}$	$\bar{a}_{3i}$	$\bar{a}_{4i}$	$\bar{a}_{5i}$
1	4.5	8.94	7.56	9.81	6.50	7.75
2	13	11.67	14.39	11.89	14.83	14.11
3	21.5	18.56	16.88	17.44	17.44	17.00

表 6

$i$	$x'_{1i}$	$a_i^{(1)}$	$x'_{2i}$	$a_i^{(2)}$	$x'_{3i}$	$a_i^{(3)}$	$x'_{4i}$	$a_i^{(4)}$	$x'_{5i}$	$a_i^{(5)}$	$d_i$
1	4	3.5	7	3	0.34	4	1	3	42.7	5	18.5
2	18	8	25	6	0.58	8	5	6	64.0	6	34.0
3	3	2	29	7	0.21	3	0	1.5	17.6	4	17.5
4	4	3.5	13	5	0.39	5	7	8.5	85.3	7	29.0
5	7	5.5	35	8	0.20	2	0	1.5	97.3	8	25.0
6	1	1	3	1	0.52	7	2	4	0.8	1	14.0
7	7	5.5	4	2	0.46	6	7	8.5	122.4	9	31.0
8	23	9	56	9	0.02	1	5	6	8.0	2	27.0
9	10	7	10	4	0.69	9	5	6	9.3	3	29.0

表7 回 报 统 计 表

年份	$\mu(1)$	$\mu(2)$	$\mu(3)$	y的级别	对否
57	22.06	45.89	66.32	1	✓
68	19.44	34.89	55.32	1	✓
67	22.82	32.61	41.94	1	✓
74	32.06	48.77	53.20	1	✓
77	26.32	35.61	44.94	1	✓
79	22.06	36.05	49.82	1	✓
62	13.56	31.39	51.82	1	✓
60	23.44	29.17	43.74	1	✓
64	30.44	10.17	16.32	2	✓
78	51.06	31.67	27.32	3	×
71	27.44	29.83	35.06	1	×
65	32.94	24.05	36.82	2	✓
69	32.06	34.05	30.94	3	×
80	26.56	12.95	28.06	2	✓
81	39.82	35.45	42.56	2	✓
60	37.44	42.45	46.56	1	×
76	45.06	23.11	24.44	2	✓
75	49.94	23.61	12.06	3	✓
63	40.56	18.39	16.18	3	✓
70	52.94	39.33	29.56	3	✓
73	38.56	28.11	26.44	3	✓
59	29.06	22.55	33.94	2	×
72	53.94	34.22	21.94	3	✓
58	72.94	46.66	26.18	3	✓
61	63.44	37.11	16.88	3	✓

## 参 考 文 献

- 【1】吴绍敏, 几种相似预报方法, 福建师大学报(自然科学版), (1979).  
 【2】福建师大数学系统计数学教研组编, 统计预测(讲义), (1977).