第七章 排列试验

第一节 排序检验法

比较数个样品，按照其某项品质程度（如，某特性的强度或嗜好程度等）的大小进行排序的方法，成为排序检验法。该法只排出样品的次序，表明样品之间的相对大小、强弱、好坏等等，属于程度上的差异，而不评价样品间的差异大小。此法的优点是可利用同一样品，对其各类特征进行检验，排出优劣，且方法较简单，结果可高，即使样品间差别很小，只要评价员很认真，或者具有一定的检验能力，都能在相当精确的程度上排出顺序。

当试验目的是就某一项性质对多个产品进行比较时，比如，甜度、新鲜程度等，使用排序检验法是进行这种比较的最简单的方法。排序法比任何其他方法更节省时间。它常被用在以下几个方面：

1） 确定由于不同原料、加工、处理、包装和储藏等各环节而造成的产品感官特性差异。

2） 当样品需要为下一步的试验预筛或预分类，即对样品进行更精细的感官分析之前，可应用此方法。

3） 对消费者或市场经营者订购的产品的可接受性调查。

4） 企业产品的精选过程。

5） 可用于品评员的选择和培训。

一、方法特点

1. 此法的试验原则是：以均衡随机的顺序将样品呈送给品评员，要求品评员就指定指标将样品进行排序，计算序列和，然后利用 Friedman 法等对数据进行统计分析。
2. 参加试验的人数不得少于 8 人，如果参加人数在 16 以上，区分效果会得到明显效果。根据试验目的，品评人员要有区分样品指标之间细微差别的能力。
3. 当评定少量样品的复杂特性时，选用此法是快速而又高效的。此时的样品数一般小于 6 个。
4. 但样品数量较大（如大于 20 个），且不是比较样品间的差别大小时，选用此法也具有一定优势。但其信息量却不如定级法大，此法可不设对照样，将两组结

果直接进行对比。

5. 进行检验前，应由组织者对检验提出具体的规定，对被评价的指标和准则要有一定的理解。如对那些特性进行排列；排列的顺序是从强到弱还是从弱到强；检验时操作要求如何；评价气味时是否需要摇晃等。
6. 排序检验只能按照一种特性进行，如要求对不同的特性进行排序，则按不同的特性安排不同的顺序。
7. 在检验中，每个评价员以事先确定的顺序检验编码的样品，并安排出一个初步顺序，然后进一步整理调整，最后确定整个系列的强弱顺序，如果实在无法区别两种样品，则应在问答表中注明。

二、问答表设计与做法

在进行问答表的设计时，应明确评价的指标和准则，如对那些特性进行比较，是对产品的一种特性进行排序，还是对一种产品的多种特性进行比较；排列顺序是从强到弱还是从弱到强；要求的检验操作过程如何；是否进行感官刺激的评价，如果是，应使评价员在不同的评价之间使用水、淡茶或无味面包等以恢复原感觉能力。排序检验法问答表的一般形式如表7-1、表7-2。

表 7-1 排序检验法问答表的一般形式示例

姓名
日期
产品
试验指令；品尝样品后，请根据您所感受的甜度，把样品号码填入适当的空格中（每格中必须填一个号码）。
甜味最强　　　　　　　　　　　　　　甜味最弱

表 7-2 排序检验法问答表的一般形式

<table>
<tr><td colspan="5">排序检验法
姓名：________ 日期：__________</td></tr>
<tr><td colspan="5">试验指令：
1．从左到右依次品尝样品 A、B、C、D。
2．品尝之后，就指定的特性方面进行排序。</td></tr>
<tr><td colspan="5">试验结果：</td></tr>
<tr><td>样品 秩次
品评员</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>1</td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td></td><td></td><td></td><td></td></tr>
<tr><td>3</td><td></td><td></td><td></td><td></td></tr>
<tr><td>4</td><td></td><td></td><td></td><td></td></tr>
<tr><td>5</td><td></td><td></td><td></td><td></td></tr>
<tr><td>6</td><td></td><td></td><td></td><td></td></tr>
</table>

三、结果分析与判断

在试验中，尽量同时提供样品，品评员同时收到以均衡、随机顺序排列的样品。其任务就是将样品排序。同一组样品还可以以不同的编号被一次或数次呈送，如果每组样品被评价的次数大于 2，那么试验的准确性会得到最大提高。在倾向性试验中，告诉参评人员，最喜欢的样品排在第一位，第二喜欢的样品排在第二位，依次类推，不要把顺序搞颠倒。如果相邻两个样品的顺序无法确定，鼓励品评员去猜测，如果实在猜不出，可以取中间值，如 4 个样品中，对中间两个的顺序无法确定时，就将它们都排为（2+3）/2=2.5。如果需要排序的感官指标多于一个，则对样品分别进行编号，以免发生相互影响。排出初步顺序后，若发现不妥之处，可以重新核查并调整顺序，确定个样品在尺度线上的相应位置。

下面对实例进行分析，以便理解。

（一） 样品甜味排序

将品评员对每次检验的每一特性的评价汇集在如表 7-3 所示的表格内。表 7-3 是六个品评员对 A、B、C、D 四种样品的甜味排序结果。

表 7-3 品评员的排序结果

品评员 \ 样品 \ 秩次	1		2		3		4
1	A		B		C		D
2	B	=	C		A		D
3	A		B	=	C	=	D
4	A		B		D		C
5	A		B		C		D
6	A		C		B		D

（二） 统计样品秩次和秩和

在每个品评员对每个样品排出的秩次当中有相同秩次时，则取平均秩次。表 7-4 是表 7-3 中的样品秩次与秩和。

表 7-4 样品的秩次与秩和

品评员 \ 秩次 \ 样品	A	B	C	D	秩和
1	1	2	3	4	10
2	3	1.5	1.5	4	10
3	1	3	3	3	10
4	1	2	4	3	10
5	1	2	3	4	10
6	1	3	2	4	10
每种样品的秩和 R	8	13.5	16.5	22	50

（三）统计解释

使用 Friedman 检验和 Page 检验对被检样品之间是否有显著差异作出判定。

1． Friedman 检验

先用下式求出统计量 F：

$$F=\frac{12}{JP\left(P+1\right)}\left[R_1^2+R_2^2+\cdots\ R_P^2-3J\left(P+1\right)\right]$$

式中

J——品评员数

P——样品（或产品）数

R_1、R_2、R_P——每种样品的秩和

查表 8-5，若计算出的 F 值大于或等于表中对应于 P、J、α 的临界值；则可以判定样品之间有显著性差异；若小于相应临界值，则可以判定样品之间没有显著性差异。

当品评员数 J 较大，或当样品数 P 大于 5 时，超出表 7-5 的范围，可查 x^2 分布表(附表 1.)，

F 值近似服从自由度为 P-1 的 x^2 值.

上例中（见表 7-4））的 F 值为：

$$F=\frac{12}{6\times4\times\left(4+1\right)}\left[8^2+13.5^2+16.5^2+22^2-3\times6\times\left(4+1\right)\right]$$
$$=10.25$$

当品评员是在分不出某两种样品之间的差距时，可以允许将这两种样品排定同一秩次，这时用 $F^{'}$ 代替 F：

$$F^{'}=\frac{F}{1-E\Big/\left[JP\left(P^2-1\right)\right]}$$

式中 E 值由如下得出：

令 n_1、n_2、$\cdots$ n_k 为出现相同秩的样品数，若没有相同秩次，$n_k=1$，则

$$E=\left(n_1{}^3-n_1\right)+\left(n_2{}^3-n_2\right)+\cdots+\left(n_k{}^3-n_k\right)$$

表中，出现相同秩次的样品数有：n_2=2，n_3=3，其余均没有相同秩次。所以

$$E=\left(2^3-2\right)+\left(3^3-3\right)+\cdots+\left(1^3-1\right)$$
$$=6+24$$
$$=30$$

故 $$F^{'}=\frac{F}{1-30/6\times4\left[\left(4^{2}-1\right)\right]}=1.09F=11.17$$

用 $F^{'}$ 与表 7-5 或 x^2 分布表(附表 1.)中的临界值比较，从而得出统计结论。

本例中，$F^{'}=11.17$，大于表 7-5 中相应的 J、P、α（6，4，0.01）的临界值 10.20，所以可以判定在 1%显著水平下，样品之间有显著性差异。

表 7-5 Friedman 秩和检验近似临界值表

品评员数目 J	样品（或产品）的数目 P					
	3	4	5	3	4	5
	显著水平 α =0.05			显著水平 α =0.01		
2	—	6.00	7.60	—	—	8.00
3	6.00	7.00	8.53	—	8.20	10.13
4	6.50	7.50	8.80	8.00	9.30	11.10
5	6.40	7.80	8.96	8.40	9.96	11.52
6	6.33	7.60	9.49	9.00	10.20	13.28
7	6.00	7.62	9.49	8.85	10.37	13.28
8	6.25	7.65	9.49	9.00	10.35	13.28
9	6.22	7.81	9.49	8.66	11.34	13.28
10	6.20	7.81	9.49	8.60	11.34	13.28
11	6.54	7.81	9.49	8.90	11.34	13.28
12	6.16	7.81	9.49	8.66	11.34	13.28
13	6.00	7.81	9.49	8.76	11.34	13.28
14	6.14	7.81	9.49	9.00	11.34	13.28
15	6.40	7.81	9.49	8.93	11.34	13.28

2．Page 检验

有时样品有自然的顺序，例如样品成分的比例、温度、不同的贮藏时间等可测因素造成的自然顺序。为了检验该因素的效应，可以使用 Page 检验。该检验也是一种秩和检验，在样品有自然顺序的情况下，Page 检验比 Friedman 检验更有效。

如果r_1、r_2、$\cdots$ r_p是以确定的顺序排列得 P 种样品的理论上的平均秩次，如果两种样品之间没有差别，则应r_1=r_2=$\cdots$ $=r_p$。否则，应$r_1 \le r \le \cdots \le r_p$，其中至少有一个不等式是成立的，也就是原假设不能成立，检验原假设能够成立，用下式计算统计量来确定：

$$L = R_1 + 2R_2 + \cdots + PR_P$$

若计算出的L值大于或等于表 7-6 中的相应的临界值，则拒绝原假设而判定样品之间有显著性差异。

如若品评员人数 J 或样品数 P 超出表 7-5 的范围，可用统计量$L^{'}$作检验，见下式：

$$L^{'} = \frac{12L - 3JP\left(P+1\right)^2}{P\left(P+1\right)\sqrt{J\left(P-1\right)}}$$

当$L^{'} \ge 1.65$，α=0.05

$L^{'} \ge 2.33$，α=0.01

以此判定样品之间有显著性差异。

表 7-6 Page 检验临界值表

品评员数目 J	样品（或产品）数 P											
	3	4	5	6	7	8	3	4	5	6	7	8
	显著水平 α =0.05						显著水平 α =0.01					
2	28	58	103	166	252	362	—	60	106	173	261	376
3	41	84	150	244	370	532	42	87	155	252	382	549
4	54	111	197	321	487	701	55	114	204	331	5041	722
5	66	137	244	397	603	869	68	141	251	409	620	893
6	79	163	291	474	719	1037	81	167	299	486	737	1063
7	91	189	338	550	835	1204	93	193	346	563	855	1232
8	104	214	384	925	950	1371	106	220	393	640	972	1401
9	116	240	431	701	1065	1537	119	246	441	717	1088	1569
10	128	266	477	777	1180	1703	131	272	487	793	1205	1736
11	141	292	523	852	1295	1868	144	298	534	869	1321	1905
12	153	317	570	928	1410	2035	156	324	584	946	1437	2072
13	165	343	615	1003	1525	2201	169	350	628	1022	1553	2240
14	178	368	661	1078	1639	2367	181	376	674	1098	1668	2407
15	190	394	707	1153	1754	2532	194	402	721	1174	1784	2574
16	202	420	754	1228	1868	2697	206	427	767	1249	1899	2740
17	215	445	800	1303	1982	2862	218	453	814	1325	2014	2907
18	227	471	846	1378	2097	3028	231	479	860	1401	2130	3073
19	239	496	891	1453	2217	3193	243	505	906	1476	2245	3240
20	251	522	937	1528	2325	3358	256	531	953	1552	2360	3406

（三） 统计分组

当用 Friedman 检验或 Page 检验确定了样品之间存在显著性差异之后，可采用下述方法进一步确定个样品之间的差异程度。

1． 多重比较和分组

根据各样品的秩和 R_p，从小到大将样品初步排序，上例的排序为：

R_A	R_B	R_C	R_D
8	13.5	16.5	22

计算临界值 r（I，α）：

$$r\left(I,\alpha\right)=q\left(I,\alpha\right)\frac{\sqrt{JP\left(P+1\right)}}{12}$$

式中 $q\left(I,\alpha\right)$ 值可查表 8-7,其中:

$I=2,3,\cdots P$。

本例中，根据表 7-3，临界值 $r\left(I,\alpha\right)$ 为：

$$r\left(I,\alpha\right)=q\left(I,\alpha\right)\frac{\sqrt{6\times4\left(4+1\right)}}{12}$$
$$=3.16q\left(I,\alpha\right)$$

比较与分组：

以下列的顺序检验这些秩和的差数：最大减最小，最大减次小，$\cdots\cdots$ 最大减次大；然后次大减最小，次大减次小 $\cdots\cdots$ 依次减下去，一直到次小减最小；

$R_{AP}-R_{A1}$ 与 $r\left(p,\alpha\right)$ 比较；

$R_{AP}-R_{A2}$ 与 $r\left(p-1,\alpha\right)$ 比较；

$\vdots$

$R_{AP}-R_{AP-1}$ 与 $r\left(2,\alpha\right)$ 比较；

$R_{AP-1}-R_{A1}$ 与 $r\left(P-1,\alpha\right)$ 比较；

$R_{AP-1}-R_{A2}$ 与 $r\left(P-2,\alpha\right)$ 比较；

$\vdots$

$R_{A2}-R_{A1}$ 与 $r\left(2,\alpha\right)$ 比较

若相互比较的两个样品 A_j 与 A_i 的秩和之差 $R_{Aj}-R_{Ai}$（j>i）小于相应的 r 值，则表示这

两个样品以及秩和位于这两个样品之间的所有样品无显著差异，在这些样品以下可用一横线表示，即：

$$\underline{A_i \quad A_{i+1} \quad \cdots \quad A_j}$$

横线内的样品不必再作比较。

若相互比较的两个样品 A_i 与 A_j 的秩和之差 $R_{Aj} - R_{Ai}$ 大于或等于相应的 r 值，则表示这两个样品有显著性差异，其下面不划横线。

不同横线上面的样品表示不同的组，若有样品处于横线重叠处，应单独列为一组。

根据表 7-4，查表 7-7 可得：

$$r(4,0.05) = q(4,0.05) \times 3.16 = 3.63 \times 3.16 = 11.47$$

$$r(3,0.05) = q(3,0.05) \times 3.16 = 3.31 \times 3.16 = 10.46$$

$$r(2,0.05) = q(2,0.05) \times 3.16 = 2.77 \times 3.16 = 8.75$$

由于：

$R_4 - R_1 = 22 - 8 = 14 > r(4,0.05) = 11.47$，不可划线。

$R_4 - R_2 = 22 - 13.5 = 8.5 < r(3,0.05) = 10.46$，可划线。

$R_3 - R_1 = 16.5 - 8 = 8.5 < r(3,0.05) = 10.46$，可划线。

结果如下：

A　B　C　D

最后分为 3 组

A　B　C　D

结论：在 5%的显著水平上，D 样品最甜，C、B 样品次之，A 样品最不甜，C、B 样品在甜度上无显著性差别。

检验表（α =5%，α =1%）中的相应于品评员数 J 和样品数 P 的临界值，从而分析出检验的结果。

表 7-7 $q(I,\alpha)$ 值表

I	α =0.01	α =.05	I	α =0.01	α =0.05
12	3.64	2.77	11	5.23	4.55
3	4.12	3.31	12	5.29	4.62
4	4.40	3.63	13	5.35	4.69
5	4.60	3.86	14	5.40	4.74
6	4.76	4.03	15	5.45	4.80
7	4.88	4.17	16	5.49	4.85
8	4.99	4.29	17	5.54	4.89
9	5.08	4.39	18	5.57	4.93
10	5.16	4.47	19	5.61	4.97
20	5.65	2.15	38	6.06	5.46
22	5.71	2.08	40	6.09	5.50
24	5.77	5.14	50	6.23	5.65
26	5.82	5.20	60	6.34	5.76
28	5.87	5.25	70	6.43	5.86
30	5.91	5.30	80	6.51	5.95
32	5.95	5.35	90	6.58	6.02
34	5.99	5.39	100	6.64	6.09
36	6.03	5.43			

2． Kramer 检定法

同方法 1 一样，首先列出表 7-3 与表 7-4 那样的统计表，查附表 5 与附表 6 顺位检验法。检验表（α =5%，α =1%）中的相应于品评员数 J 和样品数 P 的临界值，从而分析出检验的结果。

查附表 5（α =5%）和附表 6（α =1%），相应于 J=6 和 P=4 的临界值：

	5%显著水平	1%显著水平
上段	9~12	8~22
下段	11~19	9~21

首先通过上段来检验样品间是否有显著差异，把每个样品的位级和与上段的最大值 $R_{i\max}$ 和最小值 $R_{i\min}$ 相比较。若样品位级和的所有数值都在上段的范围内，说明样品间没有显著差异。若样品位级和 $\geq R_{i\max}$ 或 $\leq R_{i\min}$，则样品间有显著差异。据表 8-4，由于最大 $R_{i\max}$ =22= R_D，最小 $R_{i\min}$ =8= R_A，所以说明在 1%显著水平，四个样品之间有显著性差异。再通过下段检查样品间的差异程度，若样品的 R_n 处在下段范围内，则可将其划为一组，表明其间无差异；若样品的位级和 R_n 落在下段的范围内之外，则落在上限之外和落在下限之外的样品就可分别组成一组。由于最大 $R_{i\max}$ =21< R_D =22；最小 $R_{i\min}$ =9> R_A =8；$R_{i\min}$ =9< R_R =13.5< R_C =16.5< $R_{i\max}$ =21，所以 A、B、C、D 四个样品可划分为三个组：

$\underline{D}$ $\underline{B\ \ C}$ $\underline{A}$

结论：在 1%的显著水平上，D 样品最甜，C、B 样品次之，A 样品最不甜，C、B 样品在甜度上无显著性差别。

第二节　分类试验法

评价员品评样品后，划出样品应属的预先定义的类别，这种评价试验的方法称为分类试验法。它是先由专家根据某样品的一个或多个特征，确定出样品的质量或其它特征类别，再将样品归纳入相应类别的方法或等级的办法。此法是使样品按照已有的类别划分，可在任何一种检验方法的基础上进行。

一、方法特点

1．此法是以过去积累的已知结果为根据，在归纳的基础上，进行产品分类。

2．当样品打分有困难时，可用分类法评价出样品的好坏差异，得出样品的级别、好坏、也可以鉴定出样品的缺陷等。

二、问答表设计与做法

把样品以随机的顺序出示给鉴评员，要求鉴评员按顺序鉴评样品后，根据鉴评表中所规定的分类方法对样品进行分类。分类检验法问答表的一般形式入表 7-8 所示。

表 7-8 分类检验法问答表的一般形式示例

<table>
<tr><td colspan="5">分类检验法
姓名：__________ 日期：____________
样品类型：_______________</td></tr>
<tr><td colspan="5">试验指令：
3．从左到右依次品尝样品。
4．品尝后把样品划入你认为应属的预先定义的类别。</td></tr>
<tr><td colspan="5">实验结果：</td></tr>
<tr><td>样品</td><td>一级</td><td>二级</td><td>三级</td><td>合计</td></tr>
<tr><td>A
B
C
D</td><td></td><td></td><td></td><td></td></tr>
<tr><td>合计</td><td></td><td></td><td></td><td></td></tr>
</table>

三、结果分析

统计每一种产品分属每一类别的频数，然后用 x^2 检验比较两种或多种产品落入不同类别的分布，从而得出每一种产品应属的级别。

下面就举例具体分析：

例如，有四种产品，通过检验分成三级，了解它们由于加工工艺的不同对产品质量所造成的影响。

由 30 位评价员进行鉴评分级，各样品被划入各等级的次数统计填入表 7-9 中.

表 7-9 四种产品的分类检验结果

等级 / 次数 / 样品	一级	二级	三级	合计
A	7	21	2	30
B	18	9	3	30
C	19	9	2	30
D	12	11	7	30
合计	56	50	14	120

假设各样品的级别分不相同，则个级别的期待值为：

$$E=\frac{\text{该等级次数}}{120}\times 30=\frac{\text{该等级次数}}{4}\text{，即 }E_1=\frac{56}{4}=14$$

$E_2=\frac{50}{4}=12.5$，$E_3=\frac{14}{4}=3.5$，而实际测定值 Q 与期待值之差 $Q_{ij}-E_{ij}$ 列出如表 7-10。

表 7-10 各级别期待值与实际值之差

i / j	一级	二级	三级	合计
样品 A	-7	8.5	-1.5	0
B	4	-3.5	-0.5	0
C	5	-3.5	-1.5	0
D	-2	-1.5	3.5	0
合计	0	0	0	

$$x_0^2 = \sum_{i=1}^{t}\sum_{j=1}^{m}\frac{\left(Q_{ij}-E_{ij}\right)^2}{E_{ij}}$$

$$=\frac{\left(-7\right)^2}{14}+\frac{4^2}{14}+\frac{5^2}{14}+\cdots+\frac{3.5^2}{3.5}=19.49$$

误差自由度 f =样品自由度×级别自由度

=（m-1）(t-1)

= (4-1) (3-1) =6

查 x^2 分布表(附表 1):

$x^2\left(6,0.05\right)=12.59; x^2\left(6,0.01\right)=16.81$

由于 $x_0^2=19.49>16.81$

所以，这三个级别之间在 1%显著水平是有显著性差异，即这四个样品可以分成三个等级，

其中 C、B 之间相近，可表示为 C、B、 A、D， 即 C、B 为一级，A 为二级，D 为三级。