

利用(3.6)和(3.7),可算得 $\hat{\beta}_0 = 0.013$, $\hat{\beta}_1 = -0.25$. 于是有回归方程

$$\ln \frac{p(x)}{1-p(x)} \approx 0.013 - 0.25x$$

即

$$P(Y=1|x) \approx \frac{\exp\{0.013 - 0.25x\}}{1 + \exp\{0.013 - 0.25x\}}$$

不难看出,一个人文化程度越高,害怕生人的概率越低.

以上只对一个自变量的逻辑斯谛回归作了初步介绍. 若要了解更多的知识,请参看张尧庭编著的《定性资料的统计分析》一书(广西师范大学出版社,1991).

第八章 正交试验法

在生产和科研项目中,为了改革旧工艺或试制新产品,经常要做许多多因素试验.如何安排多因素试验,是一个很值得研究的问题.试验安排得好,既可减少试验次数、缩短时间和避免盲目性,又能得到好的结果;试验安排得不好,试验次数既多,结果还不一定满意.“正交试验法”是研究与处理多因素试验的一种科学方法,它在实际经验与理论认识的基础上,利用一种排列整齐的规格化的表——“正交表”,来安排试验.由于正交表具有“均衡分散”的特点,能在考察范围内,选出代表性强的少数次试验条件,做到能均衡抽样.由于是均衡抽样,能够通过少数的试验次数,找到较好的生产条件,即最优或较优的方案.

正交试验法(又称正交设计法、正交法)在国外已得到广泛的应用.特别在日本,对经济的发展起了很好的作用.在我国,正交法的理论工作有了进展,应用效率有了提高,也取得了不少可喜的成果.但是应用的规模,比之国外还有不小差距.今后,值得大力推广普及,使得这种科学方法能更好地为四个现代化服务.

本章从实用的角度介绍正交试验法的基本内容,至于它的数学理论(特别是它与数理统计一般理论的关系)就不去讨论了.

§1 正 交 表

下面先来介绍两张最常用的正交表.

正交表 $L_8(2^7)$

试验号	列号						
	1	2	3	4	5	6	7
1	1	1	1	2	2	1	2
2	2	1	2	2	1	1	1
3	1	2	2	2	2	2	1
4	2	2	1	2	1	2	2
5	1	1	2	1	1	2	2
6	2	1	1	1	2	2	1
7	1	2	1	1	1	1	1
8	2	2	2	1	2	1	2

正交表 $L_9(3^4)$

试验号	列号			
	1	2	3	4
1	1	1	3	2
2	2	1	1	1
3	3	1	2	3
4	1	2	2	1
5	2	2	3	3
6	3	2	1	2
7	1	3	1	3
8	2	3	2	2
9	3	3	3	1

正交表 $L_8(2^7)$ 有 8 个横行 7 个直列, 由字码“1”和“2”组成. 它有两个特点:

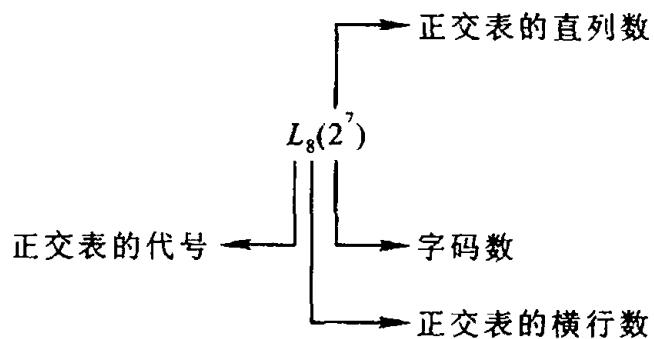
- (1) 每直列恰有四个“1”和四个“2”;
- (2) 任意两个直列, 其横方向形成的八个数字对中, 恰好 $(1, 1), (1, 2), (2, 1)$ 和 $(2, 2)$ 各出现两次. 这就是说对于任意两个直列, 字码“1”, “2”间的搭配是均衡的.

正交表 $L_9(3^4)$ 有 9 个横行、4 个直列, 由字码“1”, “2”和“3”组成. 它也具有两个特点:

- (1) 每直列中, “1”, “2”和“3”出现的次数相同, 都是三次;
- (2) 任意两个直列, 其横方向形成的九个数字对中, $(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2)$ 和 $(3, 3)$ 出现的次数相同, 都是一次; 即任意两列的字码“1”, “2”和“3”间的搭配是均衡的.

其他常用的正交表列于附表中, 它们都具有“搭配均衡”的特性, 这也就是正交表的“正交性”的含义. 至于正交表记号所表示的

意思,如下图所示:



怎样利用正交表来安排与分析多因素试验呢?在第二节中我们将用几个实例来详细说明.

§2 几个实例

例 2.1 2,4-二硝基苯肼的工艺改革

1. 试验目的与考核指标

2,4-二硝基苯肼是一种试剂产品.过去的工艺过程长、工作量大且产品经常不合格.北京化工厂改革了工艺,采用2,4-二硝基氯代苯(以下简称氯代苯)与水合肼在乙醇作溶剂的条件下合成的新工艺.小试验已初步成功,但产率只有45%,希望用正交法找出好生产条件,达到提高生产的目的.考核指标是产率(%)与外观(颜色).

2. 制定因素位级表

影响试验结果的因素是多种多样的.通过分析矛盾,在集思广益的基础上,决定本试验需考察乙醇用量、水合肼用量、反应温度、反应时间、水合肼纯度和搅拌速度六种因素.对于这六个要考察的因素,现分别按具体情况选出要考察、比较的条件——正交法中称之为位级(也称水平).

因素A——乙醇用量

第一位级 $A_1 = 200 \text{ mL}$, 第二位级 $A_2 = 0 \text{ mL}$ (即中途不再加乙醇). (挑选这个因素与相应的位级, 是为了考察一下能否砍掉中途加乙醇这道工序? 从而节约一些乙醇.)

因素 B ——水合肼用量

第一位级 B_1 = 理论量的 2 倍, 第二位级 B_2 = 理论量的 1.2 倍.

(水合肼的用量应超过理论量, 但应超过多少? 心中无数. 经过讨论, 选了 2 倍和 1.2 倍两个位级来试一试.)

因素 C ——反应温度

第一位级 C_1 = 回流温度, 第二位级 $C_2 = 60^\circ\text{C}$. (回流温度容易掌握, 便于操作, 但对反应是否有利呢? 现另选一个 60°C 跟它比较.)

因素 D ——反应时间

第一位级 $D_1 = 4 \text{ h}$, 第二位级 $D_2 = 2 \text{ h}$.

因素 E ——水合肼纯度

第一位级 E_1 = 精品(浓度为 50%), 第二位级 E_2 = 粗品(浓度为 20%).

(考察这个因素是为了看看能否用粗品取代精品, 以降低成本与保障原料的供应.)

因素 F ——搅拌速度

第一位级 = 中快速, 第二位级 = 快速.

(考察本因素及反应时间 D , 是为了看看不同的操作方法对于产率和质量的影响.)

现把以上的讨论, 综合成一张因素位级表:

因素	乙醇用量 A (mL)	水合肼用量 B	温度 C	时间 D (h)	水合肼纯度 E	搅拌速度 F
位级 1	200	理论量的 2 倍	回流	4	精品	中快
位级 2	0	理论量的 1.2 倍	60°C	2	粗品	快速

由表看出,不同的位级可以是不同的原料用量(如 A, B),也可以是不同的操作方法(如 C, D, F),或不同的原料(如 E)等等.至于每个因素要考察几个位级,这可根据需要及可能而定.可以选用二位级,三位级或更多的位级(详见 § 3).

因素、位级选好了,怎么去安排试验呢?如果要把全部搭配都试验一遍,六因素二位级需要做 $2^6 = 64$ 次试验,如果用正交表 $L_8(2^7)$ 来安排,意味着从 64 次试验中挑出 8 个代表先做试验.

3. 确定试验方案

表 $L_8(2^7)$ 最多能安排 7 个 2 位级的因素.本例有 6 个因素,可用该表来安排.具体过程如下:

(1) 因素顺序上列.

按照因素位级表中固定下来的六种因素的次序,A(乙醇用量)、B(水合肼用量)、C(反应温度)、D(反应时间)、E(水合肼纯度)和 F(搅拌速度),顺序地放到 $L_8(2^7)$ 前面的六个直列上,每列上放一种.第 7 列没有放因素,那么,它在安排试验条件上不起作用,我们可抹掉它.

(2) 位级对号入座.

六种因素分别在各列上安置好以后,再来把相应的位级,按因素位级表所确定的关系,对号入座.具体来说:

第 1 列由 A(乙醇用量)所占有,那么,在第 1 列的四个号码“1”的后面,都写上“200 mL”,即因素位级表中因素 A 的位级 1 所对应的具体用量 A_1 ;在第 1 列的四个号码“2”的后面都写上“0 mL”,即 A_2 .其余几列是类似的.综合得表 8.1.

(3) 列出试验条件.

表 8.1 是一张列好的试验方案表.表的每一横行代表要试验的一种条件.每种条件试验一次,该表共 8 个横行,因此要做 8 次试验.8 次试验的具体条件如下:

第 1 号试验: $A_1 B_1 C_1 D_2 E_2 F_1$, 具体内容是

表 8.1

试 验 号	因 素 列	乙醇用量 A	水合肼 用量 B	温度 C	时间 D	水合肼 纯度 E	搅拌速度 F
		1	2	3	4	5	6
1	1(200 mL)	1(2 倍)	1(回流)	2(2 h)	2(粗品)	1(中快)	
2	2(0 mL)	1(2 倍)	2(60℃)	2(2 h)	1(精品)	1(中快)	
3	1(200 mL)	2(1.2 倍)	2(60℃)	2(2 h)	2(粗品)	2(快)	
4	2(0 mL)	2(1.2 倍)	1(回流)	2(2 h)	1(精品)	2(快)	
5	1(200 mL)	1(2 倍)	2(60℃)	1(4 h)	1(精品)	2(快)	
6	2(0 mL)	1(2 倍)	1(回流)	1(4 h)	2(粗品)	2(快)	
7	1(200 mL)	2(1.2 倍)	1(回流)	1(4 h)	1(精品)	1(中快)	
8	2(0 mL)	2(1.2 倍)	2(60℃)	1(4 h)	2(粗品)	1(中快)	

乙醇用量:200 mL;

水合肼用量:理论量的 2 倍;

反应温度:回流温度;

反应时间:2 h;

水合肼纯度:粗品;

搅拌速度:中快.

第 3 号试验: $A_1 B_2 C_2 D_2 E_2 F_2$, 读者不难看出它的具体内容.

同样可以写出另外六个试验条件.

到这里, 就完成了试验方案的制订工作. 我们通过正交表 $L_8(2^7)$, 从全体六十四种搭配中选了有规则的八个来做试验, 这八个试验同时考察了六个因素, 并且满足以下两条: 1°任何一个因素的任何一个位级都作了四次试验. 2°任何两个因素的任何一种位级搭配都作了两次试验. 因此这八个试验条件均衡地分散到全体 64 个搭配条件中, 对全体有较强的代表性.

方案排好了. 随后的任务是按照方案中规定的每号条件严格

操作，并记录下每号条件的试验结果。至于 8 个试验的顺序，并无硬性规定，看看怎么方便而定。对于没有参加正交表的因素，最好让它们保持良好的固定状态；如果试验前已知其中某些因素的影响较小，也可以让它们停留在容易操作的自然状态。

4. 试验结果的分析

本例的考察指标是产品的产率和颜色。八个试验的结果填在表 8.2 的右方。

怎样充分利用这八个试验的结果呢？

(1) 直接看。

直接比较八个试验的产率，容易看出：

第 2 号试验的产率为 65%，最高；其次是第 5 号试验，为 63%。这些好效果，是通过试验的实践直接得到的，比较可靠。

对于另一项指标——外观，开始同时做这八个试验时，第 2 号和第 7 号是紫色，颜色不合格；而第 2 号的产率还是最高。为弄清出现紫色的原因，对这两号条件又各重复做一次试验。结果，产率差别不大；奇怪的是，其颜色却得到桔黄色的合格品。这表明，对于产率，试验是比较准确的；对于颜色，还有重要因素没有考察又没有固定在某个状态。工人师傅对这两号试验的前后两种情况进行具体分析后推测，影响颜色的重要因素可能是加料速度，决定在下一批试验中进一步考察。

(2) 算一算。

对于正交试验的数量结果，通过简单的计算，往往能由此找出更好的条件，也能粗略地估计一下哪些因素比较重要，以及各因素的好位级在什么地方。怎么算呢？

在表 8.2 每一列的下方，分别列出了 I, II 与极差 R，它们的算法如下：

如第 1 列的因素是乙醇用量 A。它的 I = 215，是由这一列四个位级 1(A_1)的产率加在一起得出的。第 1 列的数码“1”所相应的试验号是第 1, 3, 5 和 7 号，所以

表 8.2

		试验设计						试验结果	
试验号	因列号	A	B	C	D	E	F	产率(%)	颜色
1	1	1(200 mL)	1(2 倍)	1(回流)	2(2 h)	2(粗品)	1(中快)	56	桔黄
2	1	2(0 mL)	1	2(60°C)	2	1(精品)	1	65	紫色, 桔黄
3	2	2	2(1.2 倍)	2	2	2	2(快)	54	桔黄
4	1	1	1	2	1	1	2	43	桔黄
5	2	2	1	2	1(4 h)	1	2	63	桔黄
6	1	1	1	1	1	2	2	60	桔黄
7	2	1	2	1	1	1	1	42	紫色, 桔黄
8	2	2	2	2	1	2	1	42	桔黄
I = 位级 1 四次产率之和		215	244	201	207	213	205		
II = 位级 2 四次产率之和		210	181	224	218	212	220		
极差 R = I, II 中, 大数 - 小数		5	63	23	11	1	15		1 + II = 425 = 总和

$$\begin{aligned}
 (\text{产率和数}) I &= ① + ③ + ⑤ + ⑦ \\
 &= 56 + 54 + 63 + 42 = 215
 \end{aligned}$$

同样, $I = 210$, 是由第 1 列中四个位级 2(A_2)的产率加在一起得出的, 即

$$\begin{aligned}
 (\text{产率和数}) II &= ② + ④ + ⑥ + ⑧ \\
 &= 65 + 43 + 60 + 42 = 210
 \end{aligned}$$

其他五列的计算 I , II 的方法, 跟第 1 列相同.

为了检查计算是否正确, 对每列算得的 I 和 II 进行验证:

$$I + II = 425 \text{ (即 8 次试验产率的总和)}$$

至于各列的极差 R , 由各列 I , II 两数中, 用大数减去小数即得.

怎样看待这些计算所得的结果呢?

首先, 对于各列, 比较其产率和数 I 和 II 的大小. 若 I 比 II 大, 则占有该列的因素的位级 1, 在产率上通常比位级 2 好; 若 II 比 I 大, 则占有该列的因素的位级 2 比位级 1 好. 比如第 4 列的 $II = 218$, 它比 $I = 207$ 大, 这大致表明了时间因素以 2 位级为好, 即反应时间 2 h 优于 4 h.

极差 R 的大小用来衡量试验中相应因素作用的大小. 极差大的因素, 意味着它的两个位级对于产率所造成的差别比较大, 通常是重要因素. 而极差小的因素往往是不重要的因素. 在本例中, 第 2 列(水合肼用量 B 所占有)的 $R = 63$, 比其他各列的极差大. 它表明对产率来说, 水合肼用量是重要因素, 理论量的 2 倍比 1.2 倍明显地提高产率. 要想再提高产率, 可对水合肼用量详加考察, 决定在第二批试验中进行. 第 3, 6 和 4 列的 R 分别是 23, 15 和 11, 相对来说居中, 表明反应温度、搅拌速度和反应时间是二等重要的因素, 生产中可采用它们的好位级, 本例中为 C_2 , F_2 和 D_2 . 第 1 列的 $R = 5$, 第 5 列的 $R = 1$, 极差值都很小, 说明两个位级的产率差不多, 因而这两个因素是次要因素. 本着减少工序、节约原料、降低成本和保障供应的要求, 选用了不加乙醇(砍掉这道工序) A_2 和

用粗品水合肼 E_2 这两个位级. 对于次要因素, 选用哪个位级都可以, 应根据节约方便的原则来选用.

(3) 直接看和算一算的关系.

直接看, 第 2 号的产率 65% 和第 5 号的产率 63% 比做正交试验前的 45% 提高了很多. 但我们毕竟只做了八次试验, 仅占六因素二位级搭配完全的 $2^6 = 64$ 个条件的八分之一, 即使不改进位级, 也还有继续提高的可能.“算一算”的目的, 就是为了展望一下更好的条件. 对于大多数项目, “算一算”的好条件(当它不在已做过的 8 个条件中时), 将会超过“直接看”的好条件. 不过, 对于少部分项目, “算一算”的好条件却比不上“直接看”的, 由此可见, “算一算”的好条件(本例中为 $A_2 B_1 C_2 D_2 E_2 F_2$), 还只是一种可能好的配合.

如果生产上急需, 通常应优先补充试验“算一算”的好条件. 经过验证, 如果效果真有提高, 就可将它交付生产上使用. 倘若验证后的效果比不上“直接看”的好条件, 就说明该试验的现象比较复杂. 还有一种情况是, 由于试验的时间较长, 等不到验证试验的结果. 对于这两种情况, 生产上可先使用“直接看”的好条件, 也可结合具体情况做些修改; 而与此同时, 另行安排试验, 寻找更好的条件.

5. 第二批正交试验

在第一批试验的基础上, 为弄清影响颜色的原因及进一步提高产率, 决定再撒个小网, 做第二批正交试验.

(1) 制定因素位级表.

因素	水合肼用量	时间	加料速度
位级 1	1.7 倍	2 h	快
位级 2	2.3 倍	4 h	慢

水合肼是上批试验中最重要的因素, 应该详细考察. 现决定在好用量 2 倍的周围, 再取 1.7 倍与 2.3 倍两个新用量继续试验. 另外, 在追查出现紫色原因的验证试验后, 猜想加料速度可能是影响

颜色的重要原因,因此在这批试验中要着重考察这个猜想.关于反应时间,因为第一线的同志对于用 2 h 代替 4 h 特别重视,所以再比较一次.

对于上批试验的其他因素,为了节约与方便,这一批决定砍掉中途“加乙醇”这道工序;用“快速搅拌”;“反应温度 60℃”虽然比回流好,但 60℃ 难于控制,决定用 60~70℃ 之间.另外,一律采用粗品水合肼.

(2) 利用正交表确定试验方案.

$L_4(2^3)$ 是两位级的表,最多能安排 3 个两位级的因素,本批试验用它来安排是很合适的.

至于填表及确定试验方案的过程,即所谓“因素顺序上列”、“位级对号入座”及列出试验条件的过程已经介绍过,不再细述.现将试验计划与试验结果列于表 8.3.

表 8.3

试验计划				试验结果		
试验号	因素列号	水合肼用量 A	时间 B	加料速度 C	产率 (%)	颜色
		1	2	3		
1		1(1.7 倍)	1(2 h)	1(快)	62	不合格
2		2(2.3 倍)	1	2(慢)	86	合格
3		1	2(4 h)	2	70	合格
4		2	2	1	70	不合格
I = 位级 1 二次产率之和		132	148	132		
II = 位级 2 二次产率之和		156	140	156	I + II = 288 = 总和	
极差 R = I, II 中, 大数 - 小数		24	8	24		

(3) 试验结果的分析.

关于颜色，“快速加料”的第 1,4 号试验都出现紫色不合格品，而“慢速加料”的第 2,3 号试验都出现桔黄色的合格品。另外两个因素的各个位级，紫色和桔黄色各出现一次，这说明它们对于颜色不起决定性的影响。由此看出，加料速度是影响颜色的重要因素，应该慢速加料。

关于产率，从“直接看”与“算一算”，都是第 2 号的最高。

最后顺便提一下投产效果。通过正交试验法，决定用下列工艺投产：用工业 2,4 -二硝基氯代苯与粗品水合肼在乙醇溶剂中合成；水合肼用量为理论量的 2.3 倍，反应时间为 2 h，温度掌握在 60~70℃ 之间，采用慢速加料与快速搅拌。效果是：平均产率超过 80%，从未出现紫色外形，质量达到出口标准。总之，这是一个较优的方案，可以达到优质、高产、低消耗的目的。

例 2.2 晶体退火工艺的改进

1. 试验目的与考核指标

检查癌细胞，用到一种碘化钠晶体 φ40，要求应力越小越好，希望不超过 2 度。退火工艺是影响质量的一个重要环节。国营 261 厂经过 30 多炉试验，其他指标都已合格，只是应力未能低于 7 度。现在通过正交试验，希望能找到降低应力的工艺条件。考核指标是应力(度)。

2. 挑因素、选位级，制定因素位级表

考察升温速度、恒温温度、恒温时间和降温速度共四个因素。每个因素取 3 个位级，因素位级表如下：

因素	升温速度 A ($^{\circ}\text{C} \cdot \text{h}^{-1}$)	恒温温度 B ($^{\circ}\text{C}$)	恒温时间 C (h)	降温速度 D
位级 1	30	600	6	1.5 A 电流降温
位级 2	50	450	2	1.7 A 电流降温
位级 3	100	500	4	$15^{\circ}\text{C} \cdot \text{h}^{-1}$

关于升温速度 A，除了原工艺的 $50^{\circ}\text{C} \cdot \text{h}^{-1}$ ，在它的周围看一个慢速升温

“ $30^{\circ}\text{C} \cdot \text{h}^{-1}$ ”, 和一个快速升温“ $100^{\circ}\text{C} \cdot \text{h}^{-1}$ ”; 关于恒温温度 B , 原工艺的恒温温度 600°C 是从国外资料中借鉴的, 现在增添两个较低的温度 500°C 与 450°C , 看看行不行? 关于恒温时间 C , 原工艺为 6 h , 现在看看缩短些是否更好? 关于降温速度 D , 原工艺是通 1.5 A 的电流降温, 现在加一个 1.7 A 的慢速降温. 另外, 虽然过去的经验表明等速降温不好, 这次还是安排了一个 $15^{\circ}\text{C} \cdot \text{h}^{-1}$ 的等速降温; 三个降温的位级, 都是下降到 250°C 后断电而自然降温.

3. 确定试验方案

表 $L_9(3^4)$ 最多能安排四个三位级的因素, 本例有四个三位级的因素, 因此, 用 $L_9(3^4)$ 来安排试验, 正是恰到好处.

至于填表过程与试验条件的列出手续与上例同, 综合于表 8.4 中.

4. 试验结果的分析

每号条件做一炉试验, 应力结果记在表 8.4 相应条件的右边.

(1) 直接看.

第 5 号试验 $A_2 B_2 C_3 D_3$ 的 0.5 度最好, 第 7 号 $A_1 B_3 C_1 D_3$ 的 1 度次之. 这两号试验的具体条件是:

	升温速度	恒温温度	恒温时间	降温速度
第 5 号	$50^{\circ}\text{C} \cdot \text{h}^{-1}$	450°C	4 h	$15^{\circ}\text{C} \cdot \text{h}^{-1}$
第 7 号	$30^{\circ}\text{C} \cdot \text{h}^{-1}$	500°C	6 h	$15^{\circ}\text{C} \cdot \text{h}^{-1}$

(2) 算一算.

1° 对于各因素列, 算出各个位级相应的三次应力之和. 如第三列恒温时间:

$$\begin{aligned} I &= \text{位级 1 三次应力之和} = \text{第 } 2, 6, 7 \text{ 号应力之和} = 7 + 7 + 1 \\ &= 15 \text{ 度}; \end{aligned}$$

$$II = \text{第 } 3, 4, 8 \text{ 号应力之和} = 15 + 8 + 6 = 29 \text{ 度};$$

$$III = \text{第 } 1, 5, 9 \text{ 号应力之和} = 6 + 0.5 + 13 = 19.5 \text{ 度}.$$

同样, 算出另外三列的 I, II 和 III.

表 8.4

因 素 列 号	试 验 号	升温速度 A			恒温温度 B			恒温时间 C			降温速度 D			应 力 (度)	
		1	2	3	1	2	3	4	5	6	7	8	9		
1	1	1(30 °C·h⁻¹)	1(600 °C)	3(4 h)	2(1.7A 电流降溫)	6									
2	2	2(50 °C·h⁻¹)	1	1(6 h)	1(1.5A 电流降溫)	7									
3	3	3(100 °C·h⁻¹)	1	2(2 h)	3(15 °C·h⁻¹)	15									
4	4	1	2(450 °C)	2	1										
5	5	2	2	3	3										
6	6	3	2	1	2										
7	7	1	3(500 °C)	1	3										
8	8	2	3	2	2										
9	9	3	3	3	1										

对于每列,比较各自 I, II 和 III 的大小,因为应力越小越好,所以应力之和小的位级较好. 第 1 列(A 列) II 小,故 A_2 较好; 第 2 列(B 列) II 小,故 B_2 较好; C 列 I 小,故 C_1 较好, D 列 III 小,故 D_3 较好. 把这四个好位级结合在一起, $A_2 B_2 C_1 D_3$, 称为全体配合(本例有四个因素每个有三个位级,因此全体配合有 81 个)中关于应力的可能好配合.

2° 计算各列的极差 R .

R = 相应列的 I, II, III 中的最大数—相应列的 I, II, III 中的最小数.

如第 3 列恒温时间,它的 I = 15, II = 29, III = 19.5, 它们的最大数是 29, 最小数是 15, 因此按公式

$$R = 29 - 15 = 14 \text{ 度}$$

四个极差记在表 8.4 的最下一行,极差大的因素通常意味着该因素三个位级相应的应力差别大,是重要因素;极差小的因素是不重要的因素.

(3) 画趋势图.

计算完极差后,对于数量性位级的三位级因素,应该画出用量与试验结果之和的关系图,以便从图形上直接看出试验结果随各因素用量变化的大体关系. 具体来说,对于每个因素,以实际用量(而不是位级号码的大小)作为横坐标,试验结果之和作为纵坐标,画出三个点,得出该因素的趋势图.

本例中,前三个定量因素的趋势图合并在图 8.1 中:

(4) 可能好配合与大范围可能好配合.

1° 通过“算一算”得到 $A_2 B_2 C_1 D_3$ 为 81 个配合中的可能好配合.

2° 可能好位级与大范围可能好配合.

对于分两个位级的因素,能看出两个用量谁好谁差,但看不出继续提高效果的好用量的方向;对于分三个位级的因素,情况起了变化,请注意恒温温度 B 的趋势图(见图 8.1).

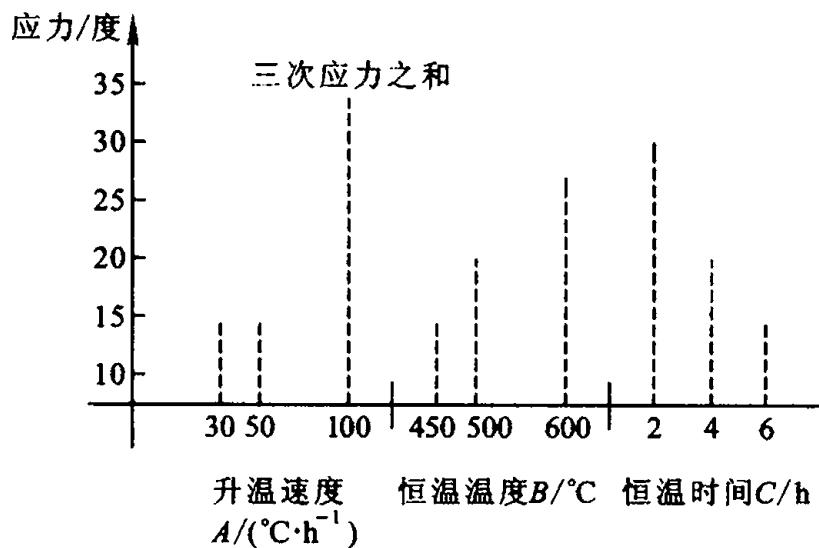


图 8.1

恒温温度的三个高度逐步上升,一个合乎理想的猜测是,如温度继续下降,应力将还能降低.这意味着原来的三个位级都选高了,不仅国外资料中的 600°C 太高,就是最低的 450°C 也偏高了.下批试验中还应降低恒温温度,如能成功,则既省电,又缩短生产时间.因此,展望的好用量不停在 $B_2 = 450^{\circ}\text{C}$,而应取 $B_4 = 400^{\circ}\text{C}$.

恒温时间的三个高度逐步下降.这也证实了过去“时间长应力低”的看法是正确的.但考虑到首批低应力已经降到 0.5 度,更主要的是延长时间不利于节约电力和提高工效,因此展望的好用量取 $C_3 = 4 \text{ h}$.

发现有用量选偏的因素(在本例中是恒温温度 B),是认识上的重大收获.下批试验中把用量选准后,常能取得明显的进展.

升温速度的变化范围没有估偏,好位级是 $A_2 = 50^{\circ}\text{C}$;

降温温度的好位级是 $D_3 = 15^{\circ}\text{C} \cdot \text{h}^{-1}$.

最后,把展望的四个好位级结合到一起,得 $A_2 B_4 C_3 D_3$,它称为大范围可能好配合,或称为“算一算”的好条件.具体是:

升温速度	$50^{\circ}\text{C} \cdot \text{h}^{-1}$
恒温温度	$400^{\circ}\text{C} \cdot \text{h}^{-1}$
恒温时间	4 h
降温速度	$15^{\circ}\text{C} \cdot \text{h}^{-1}$

注:过去的经验认为等速降温不好,这次为什么效果又很好呢?原因在于,过去用的恒温温度 600°C 太高,因而等速降温不好.如第 3 号试验,在恒温 600°C 的情况下,等速降温的应力还是最差.这次突破了国外资料记载,降低了恒温温度,等速降温又变成好办法了.由本例的讨论也清楚地看出,对于较复杂的情况,只有做多因素试验,才能找到产生好效果的条件.

5. 第二批正交试验

由恒温温度的趋势图看出,还有潜力.为进一步降低应力,决定再做第二批正交试验.

(1) 制定因素位级表.

关于升温温度,上批试验中,升温快的 $A_3 = 100^{\circ}\text{C} \cdot \text{h}^{-1}$, 第 3, 6, 9 号试验应力都很坏,不能再用;速度慢的 $A_1 = 30^{\circ}\text{C} \cdot \text{h}^{-1}$, 升温时间过长,不愿再用;所以在这批试验中均用 $50^{\circ}\text{C} \cdot \text{h}^{-1}$, 不再试验其他速度.换句话说,在这批中,升温速度保持在良好的固定状态—— $50^{\circ}\text{C} \cdot \text{h}^{-1}$, 而不再作为要考察的因素.另外三个因素,以算一算的好位级 B_4, C_3, D_3 为主,即以大范围的可能好配合为主,参看“直接看”的好位级,各取两个位级.其因素位级表如下:

因素	恒温温度 ($^{\circ}\text{C}$)	恒温时间 (h)	降温速度 ($^{\circ}\text{C} \cdot \text{h}^{-1}$)
位级 1	450	3	15
位级 2	400	5	25

(降温速度的两个位级都是下降到 250°C 时停电,然后自然降温.)

(2) 用正交表 $L_4(2^3)$, 确定试验方案(见表 8.5).

(3) 试验结果.

由表 8.5 右方看出,这批四个试验基本消除了应力.

表 8.5

试 验 号	因 素 列	恒温温度	恒温时间	等速降温	应 力 (度)
		1	2	3	
1	1(450℃)	1(3 h)	1($15^{\circ}\text{C} \cdot \text{h}^{-1}$)	0	
2	2(400℃)	1	2($25^{\circ}\text{C} \cdot \text{h}^{-1}$)	0.2	
3	1	2(5 h)	2	0.4	
4	2	2	1	0	
I	0.4	0.2	0		
II	0.2	0.4	0.6		I + II = 0.6
R	0.2	0.2	0.6		

用正交表安排试验,试验次数虽然不多,但考察的因素多,每种因素的位级个数也不少.正交表条件均衡地分散在排列完全的位级组合之中,如同一个编织得很好的一个渔网,往往能直接捕捞到“大鱼”(自然,在得到好条件的同时,也会有较差的条件,这是正常现象,不必气馁和惊奇,因为我们关心的是能否找到好条件).通过位级的指标和(I, II, III……)的计算及极差R大小的比较,以及趋势图的变化情况,可以得到重要的信息.本例在第一批正交试验的基础上,看准苗头,又撒了一个小网,结果是硕果累累,更上一层楼.

例 2.3 V_c 的配方试验

1. 试验目的与考核指标

维生素 C(简称 V_c)是一种人体必不可少的营养素,对于多种疾病有治疗作用,是一种常用的药品.为了提高 V_c 的氧化率、降低成本,做了以下的正交试验.考核指标是氧化率.

2. 确定试验方案

(1) 因素位级表.

因素	尿素 (%)	山梨糖 (%)	玉米浆 (%)	K_2HPO_4 (%)	$CaCO_3$ (%)	$MgSO_4$ (%)	葡萄糖 (%)
位级 1	CP0.7	7	1	0.15	0.4	0	0.25
位级 2	CPI.1	9	1.5	0.05	0.2	0.01	0
位级 3	CP1.5	11	2	0.10	0	0.02	0.5
位级 4	工业 0.7						
位级 5	工业 1.1						
位级 6	工业 1.5						

关于尿素,希望能用工业尿素代替 CP(化学纯)尿素,所以把 CP 尿素和工业尿素都安排了三个位级,共六个位级,比较一下它们的效果;关于山梨糖,因为它是主要原料,原生产上的浓度为 7%,做试验的主要目的是要增加它的浓度,以提高生产效率;至于 $CaCO_3$, $MgSO_4$, 葡萄糖这三个因素,我们希望在新配方中能去掉它们中的一个或两个,所以对它们,都有一个加入量为 0 的位数.

(2) 利用正交表,确定试验方案.

在试验中常有这种情况,对某些因素需要详细了解,因此要比其他因素多排位级,这时要用混合位级的正交表.本例要考察一个六位级和六个三位级的因素,我们可用第七个正交表即混合位级正交表 $L_{18}(6^1 \times 3^6)$ 来安排.

正交表 $L_{18}(6^1 \times 3^6)$ 有 18 横行 7 直列. 第一列为“1”, “2”, “3”, “4”, “5”, “6”各三个,后面六列是“1”, “2”, “3”各六个. 第一列和后面任一列组成的 18 个数字对为(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3), (4,1), (4,2), (4,3), (5,1), (5,2), (5,3), (6,1), (6,2), (6,3)各出现一次,后面任两列组成的 18 个数字对为(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)各出现两次.

表

试 验 号	因 素 列 号	尿素	山梨糖	玉米浆
		A(%)	B(%)	C(%)
1	1(CP0.7)	1(7)	3(2)	
2	1	2(9)	1(1)	
3	1	3(11)	2(1.5)	
4	2(CP1.1)	1	2	
5	2	2	3	
6	2	3	1	
7	3(CP1.5)	1	1	
8	3	2	2	
9	3	3	3	
10	4(工业0.7)	1	1	
11	4	2	2	
12	4	3	3	
13	5(工业1.1)	1	3	
14	5	2	1	
15	5	3	2	
16	6(工业1.5)	1	2	
17	6	2	3	
18	6	3	1	
I = 位级 1 氧化率之和	142	380.1	306.3	
II = 位级 2 氧化率之和	179.6	362.6	340.7	
III = 位级 3 氧化率之和	187.3	275.2	370.9	
IV = 位级 4 氧化率之和	144.1			
V = 位级 5 氧化率之和	181.9			
VI = 位级 6 氧化率之和	183			
极差 R = 最大数 - 最小数	45.3	104.9	64.6	

8.6

K_2HPO_4 $D(\%)$	$CaCO_3$ $E(\%)$	$MgSO_4$ $F(\%)$	葡萄糖 $G(\%)$	氧化率 (%)
4	5	6	7	
2(0.05)	2(0.2)	1(0)	2(0)	65.1
1(0.15)	1(0.4)	2(0.01)	1(0.25)	47.8
3(0.1)	3(0)	3(0.02)	3(0.5)	29.1
1	2	3	1	70
3	1	1	3	68.1
2	3	2	2	41.5
3	1	3	2	63
2	3	1	1	65.3
1	2	2	3	59
1	3	1	3	45.7
3	2	2	2	56.4
2	1	3	1	42
3	3	2	1	70
2	2	3	3	58.3
1	1	1	2	53.6
2	1	2	3	66.3
1	3	3	2	66.7
3	2	1	1	50
342.8	340.8	347.8	345.1	总和等于 1 017.9
338.5	358.8	341	346.3	
336.6	318.3	329.1	326.5	
6.2	40.5	18.7	19.8	

现将试验方案、试验结果与计算结果一并列入表 8.6,由于试验误差大,每号条件的氧化率,都是按常规标准,通过三次试验来确定的.

3. 试验结果的分析

(1) 直接看.

第 17 号试验的氧化率为 66.7%. 它用的是工业尿素,山梨糖用量也高于 7%,这是个好条件. 它的具体配方是:

尿素	山梨糖	玉米浆	K_2HPO_4	$CaCO_3$
工业 1.5%	9%	2%	0.15%	0
$MgSO_4$	葡萄糖			
0.02%	0			

(2) 算一算.

1° 计算氧化率之和.

第 1 列的尿素分六个位级, I, II, III, IV, V, VI 分别是位级 1, 2, 3, 4, 5, 6 各三次氧化率之和. 后面的六列因素都分三个位级, I, II, III 分别是位级 1, 2, 3 各六次氧化率之和.

2° 计算极差 R .

第 1 列的 $R =$ 第 1 列的 I, II, III, IV, V, VI 的最大数 - 第 1 列的 I, II, III, IV, V, VI 的最小数 = $187.3 - 142 = 45.3$.

其余各列均为三位级的因素. 极差 R 的计算和例 2 相同.

通常 R 大的因素是重要因素, R 小的不重要,由七列的 R 大致看出:

山梨糖(B)是重要因素,玉米浆(C)、尿素(A)和 $CaCO_3$ (E)也是比较重要的因素,而葡萄糖(G)、 $MgSO_4$ (F)和 K_2HPO_4 (D)是影响较小的因素.

(3) 画趋势图.

化学纯和工业尿素是性质不同的两个品种,应分别画它们的趋势图.

图 8.2 中,由于尿素的氧化率之和是三个数据的和,而其他因素的氧化率之和都是六个数据之和.为了合并在一个图形中,共同使用一根纵坐标轴,把尿素的氧化率之和都乘以二倍.

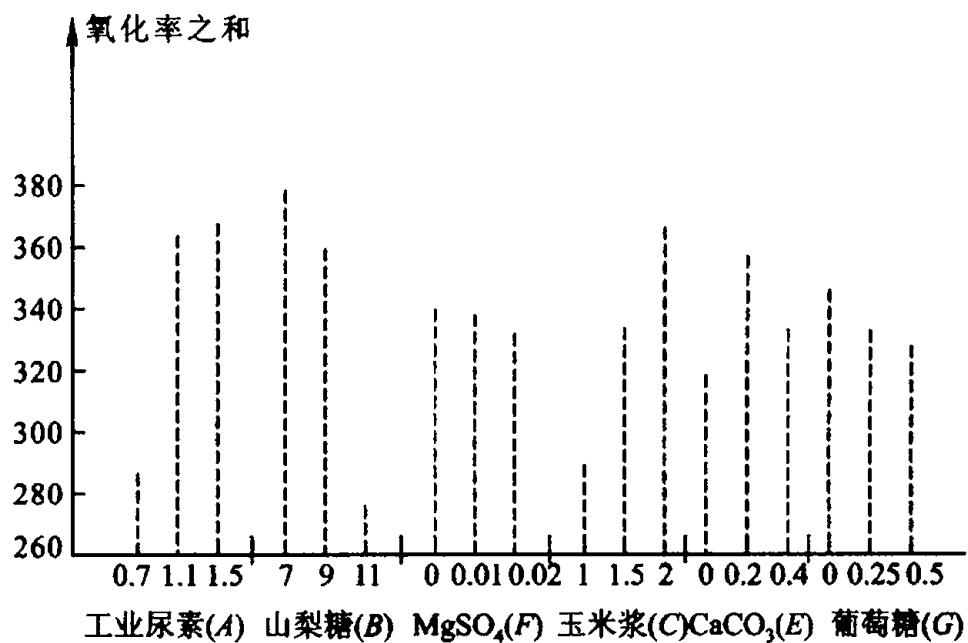


图 8.2

(4) 展望配方.

由上面的分析讨论,我们看出以下几点:

1° 主料山梨糖的用量可由 7% 往上提,但 11% 已知不好,决定提到 9%.

2° 可用工业尿素,由于图形上升,用量不能过低,可取 1.1~1.5 之间,也可再稍提高.

3° 玉米浆图形明显上升,还可以考虑加大用量,但又考虑到浓度超过 2% 后,溶液过稠,因此定为 2%.

以上三个因素,按极差 R ,都是重要与较重要的因素,且与直接看所得的第 17 号试验条件相同.至于另一个较重要的因素 CaCO_3 则不然.

4° 在第 17 号条件下, CaCO_3 加入量为 0; 而趋势图中呈中

间高两边低,0.2%最好.经讨论,大家认为,需要保证Ca盐的供应,算一算的结果是值得重视的.

5° 葡萄糖和MgSO₄的极差很小(图形上体现为三个位级的高度相差不大).这表明它们的位级可在各自的三个位级中随便取定;且加入量为0的位级都是好位级.因此决定不加葡萄糖与MgSO₄.

6° K₂HPO₄的极差最小.它的三个位级(0.05%,0.1%和0.15%)可随意取.考虑到玉米浆要加大到2%,而玉米浆中含有PO₄盐,因此还可以探索一下在培养基中不加K₂HPO₄.

综合以上,从直接看得到的好配方——第17号条件,结合算一算及以上具体分析,下列展望配方是值得一试的:

玉米浆	山梨糖	工业尿素	CaCO ₃
2%	9%	1.1~1.5%	0.2%
MgSO ₄	K ₂ HPO ₄	葡萄糖	
0	0	0	

对于这个配方,重复了十多次小试验,效果都好.北京制药厂在大罐上试生产,氧化率超过80%,与生产上原配方相比,不仅提高了主料山梨糖的浓度,还减少了三种成分,达到了节约成本、简化工艺和提高生产效率的目的.

正交表L₁₈(6¹×3⁶)的特点

L₁₈(6¹×3⁶)是一张颇有特色的正交表.它不仅可安排众多的三位级因素;而且还可以安排一个六位级的因素,使得用它安排的试验既照顾了一般又突出了重点.不少难度较大的项目(如北京化工三厂的抗氧剂1076,北京化工厂的荧光粉K₁₁及成色剂等等),用了L₁₈(6¹×3⁶)之后,都有很大进展.因此,当试验难度较大,试验次数受很大限制,而且要考察的因素又较多时,采用该表来安排是十分相宜的.

§3 小 结

1. 一般步骤

在这里,我们把正交试验法的一般步骤小结一下.

第一步:明确试验目的,确定考核指标.

第二步:挑因素,选择合适的正交表,选位级,制定因素位级表,确定试验方案.

第三步:对试验结果进行分析,其中有

- (1) 直接看;
- (2) 算一算;

(I) 各位级的指标和与极差的计算,对于多于两个位级的因素画出趋势图.

(II) 区分因素的主次及位级的优劣,得出可能好配合或大范围的可能好配合.

(3) 综合直接看和算一算这两步的结果,并参照实际经验与理论上的认识,提出展望.

关于第一步和第三步,前面三个例子,已说得比较清楚.下面只对第二步作些补充.

2. 关于挑选因素

先把试验过程中有关的因素排排队,分一下类.

一类是由于测试技术未臻完善,测不出因素的数值(或者得不到定性的了解).这样就无法看出因素的不同位级的差别,也就是说看不出因素的作用,所以不能列为被考察的因素.

一类是虽然能测出因素的量,但由于控制手段还不具备,不能把因素控制在指定的用量上,那么也不能作为正交表考察的因素,因为正交表每列的位级,都具有指定的用量.当然,所谓用量能否控制也是相对的.一方面,尽可能加强控制的准确性;另一方面,经过努力后,只要大体上能控制得住,误差不是很大,能区分开不同

的位级,还是可以当作正交试验考察的因素.对于这类因素,在试验过程中,应随时记录它的实际观测数据.

除去以上两类正交表难以考察的因素外,在能控制住用量的各因素中,要考察哪些因素呢?这自然由试验工作者决定.但是,考虑到(1)如果漏掉重要因素,可能大大降低试验效果;(2)正交表是安排多因素试验的得力工具,不怕因素多;(3)有时增加一、二个因素,并不增加试验次数.因此,一般倾向于多考察些因素,除了事先能肯定作用很小的因素不安排以外,凡是可能起作用或情况不明或意见有分歧的因素都值得考察.有时,为了减少工作量或简化手续层次,减少些次要因素是可以的;另外,也有些试验,费用很高或单个试验花费的时间过长,不可能多作试验.这时,选一些重要因素先考察也是可以的,但我们不提倡这样办.

3. 选择合用的正交表及其他

(1) 位级个数的确定.

对于某些因素,如“品种”、“次序”等,其位级个数自然形成.例如在§2例1的2,4-二硝基苯肼试验中,关于水合肼的纯度,只有“精品”和“粗品”两个位级.又如某试验要考虑甲、乙两种原料的加入次序,分为“甲先加”、“乙先加”和“同时加”三个位级,它们都是自然形成的.但是,很多因素的用量是连续变化的,如温度、时间、加入量和配比等.对于它们,只有用量范围的限制,并无位级个数的约束.对于这种因素,究竟该选几个位级呢?这由做试验的目的及性质来定.但这也和选用哪个正交表有密切联系,因为一旦正交表选定后,每列因素的位级数就随之而定了.

(2) 正交表的选择.

选择合用的正交表,主要须考虑三个方面的情况:① 考察因素的个数;② 一批允许作试验的次数;③ 有无重点因素要详细考察.

实际上正交表的选择又和考察因素的位级个数有关,也就是说,位级数的确定与正交表的选择这两个问题是互相牵扯的,要经

常放在一起考虑.下面举例说明.

§2例1中2,4-二硝基苯肼及§2例3中 V_c 的试验中,考察因素的个数是6个和7个,它们既可以采用正交表 $L_8(2^7)$,也可以采用正交表 $L_{18}(6^1 \times 3^6)$.但在2,4-二硝基苯肼的试验中,为了缩短做试验的时间、加快试验进度,又由于该试验很容易分批做,所以第一批就采用了 $L_8(2^7)$,而 V_c 发酵的试验则不同.由于这种生物化学试验的误差较大,分两个位级不易看清好坏,又由于尿素这个因素要详加考察,更重要的是该试验一批多作几个(例如18个)没有困难,所以就选用了 $L_{18}(6^1 \times 3^6)$.在§2例2晶体退火试验中,考察的因素有四个,各分三个位级,选用了正交表 $L_9(3^4)$ 安排试验,试验效果很好.这种安排是有一般性的.如果某试验要考察的因素只有三、四个,只要允许做九个试验时,建议采用三位级的正交表 $L_9(3^4)$ 来安排.因为用它可以看出不同位级的变化趋势,常常是通过做一批试验就解决了问题.当然,当试验费用高或很容易分批做试验时,三因素的情况,也可以各分两个位级用最小的正交表 $L_4(2^3)$ 来安排.

正交表的选择虽然是比较灵活的,并且常用的正交表也有几十张(本章的附表列出了23张).但是,某一项试验该采用哪一张表,只要综合考虑以上三方面的情况,对具体问题进行具体分析,还是不难确定的.

(3) 位级用量的选取.

随着正交表的选定,各因素的位级个数就随之而定,对于连续变化的因素,应该把它们对应上哪些用量呢?首先要估计一下因素取值的变化范围,其次用均分法或类似的方法确定位级的用量.一般来说,范围可以放宽些,以不漏掉合理值为原则.例如在§2例2的晶体退火试验中,考虑恒温时间这个因素,它的最小值可定为0,经商议最大值定为8小时(否则时间太长,生产效率过低),所以范围是(0,8).范围确定以后,如果采用两位级,可将范围三等

分；如果采用三位级，可将范围四等分，取中间的分点作为位级的用量。晶体退火的试验采用了三位级，所以将范围(0,8)四等分，中间的三个分点是2,4和6，于是恒温时间的三个位级定为2小时，4小时和6小时。

当然，也可以不采用均分法，如§2例1的2,4-二硝基苯肼的试验中，水合肼用量的范围原订为(1倍,2.5倍)，但我们采用了1.2倍和2倍两个位级，是为了考察一下能否节约一些水合肼又能保证产率。由此可知，位级的选择同样是由实际问题决定的，它也不是按照什么一定的法则来安排的。但是，必须注意不同位级要适当拉开，才能看出差别。

(4) 制定因素位级表。

因素及其不同位级的用量都选定以后，下一步就是制定因素位级表。对于各个因素，用哪个位级号码对应上哪个用量，这可以任意规定。但是，一经选定以后，试验过程中就不许再变了。一般来说，最好是打乱次序来安排。如§2例3的 V_c 发酵的试验中， K_2HPO_4 的用量：位级1是0.15，位级2是0.05，位级3是0.10。

(5) 确定试验方案。

因素位级表制定好以后，就可以在预先选好的正交表上填表，确定试验方案了。这在前面三个实例中已讲得很多，不赘述了。

4. 均衡分散性和整齐可比性

在正交表的任意两列中，各种位级号码搭配的出现次数都一样多。这叫作正交性，也是正交表名称的来由。由于正交性：(1)保证了试验条件均衡地分散在配合完全的位级组合之中，因而代表性强，容易出现好条件。这是均衡分散性。(2)对于每列因素，在各个位级的结果之和中，其他因素各个位级的出现次数都是相同的。这保证了在各个位级的效果中，最大限度地排除了其他因素的干扰，因而能很有效地进行比较，为我们提供有参考价值的展望。这

是整齐可比性. 正交试验法效率高的原因主要在于这两种特性. 本书限于篇幅, 不详细叙述了.

以上只介绍了正交试验法的基本内容, 没有去谈各种灵活运用的方法. 例如, 活动位级法和综合评分法就没有介绍. 有关的详细内容可参看参考书目[20].

关于正交试验法与一般数理统计理论的联系, 可参看[10].

正交试验法是各种试验设计方法中实用价值很高的一种. 关于试验设计的一般方法和理论, 读者可参看 Montgomery D C 的著作《实验设计与分析》(汪仁官等译. 北京: 中国统计出版社, 1998).

第八章附表 常用正交表

一、 $L_4(2^3)$

试验号	列号		
	1	2	3
1	1	1	1
2	2	1	2
3	1	2	2
4	2	2	1

二、 $L_8(2^7)$

试验号	列号						
	1	2	3	4	5	6	7
1	1	1	1	2	2	1	2
2	2	1	2	2	1	1	1
3	1	2	2	1	2	2	1
4	2	2	1	2	1	2	2
5	1	1	2	1	1	2	2
6	2	1	1	1	2	2	1
7	1	2	1	1	1	1	1
8	2	2	2	1	2	1	2

三、 $L_{16}(2^{15})$

试验号 \ 列号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1	1	1	2	2	1	2	1	2	2	1	1	1	2	2
2	2	1	2	2	1	1	1	1	1	2	2	1	2	2	1
3	1	2	2	2	2	2	1	1	2	1	2	1	1	1	1
4	2	2	1	2	1	2	2	1	1	1	1	1	2	1	2
5	1	1	2	1	1	2	2	1	2	2	2	2	1	2	
6	2	1	1	1	2	2	1	1	1	2	1	2	1	1	1
7	1	2	1	1	1	1	1	1	2	1	1	2	2	2	1
8	2	2	2	1	2	1	2	1	1	1	2	2	1	2	2
9	1	1	1	1	2	2	1	2	1	1	2	1	2	2	2
10	2	1	2	1	1	2	2	2	2	1	1	1	1	2	1
11	1	2	2	1	2	1	2	2	1	2	1	1	1	2	1
12	2	2	1	1	1	1	1	2	2	2	2	1	1	1	2
13	1	1	2	2	1	1	1	2	1	1	1	2	1	1	2
14	2	1	1	2	2	1	2	2	2	1	2	2	2	1	1
15	1	2	1	2	1	2	2	2	1	2	2	2	1	2	1
16	2	2	2	2	2	2	1	2	2	2	1	2	2	2	2

四、 $L_{12}(2^{11})$

试验号 \ 列号	1	2	3	4	5	6	7	8	9	10	11
1	1	1	1	2	2	1	2	1	2	2	1
2	2	1	2	1	2	1	1	2	2	2	2
3	1	2	2	2	2	2	1	2	2	1	1
4	2	2	1	1	2	2	2	2	1	2	1
5	1	1	2	2	1	2	2	2	1	2	2
6	2	1	2	1	1	2	2	1	2	1	1
7	1	2	1	1	1	1	2	2	2	1	2
8	2	2	1	2	1	2	1	1	2	2	2
9	1	1	1	1	2	2	1	1	1	1	2
10	2	1	1	2	1	1	1	2	1	1	1
11	1	2	2	1	1	1	1	1	2	1	1
12	2	2	2	2	2	1	2	1	1	1	2

五、 $L_9(3^4)$

试验号	列号				
		1	2	3	4
1		1	1	3	2
2		2	1	1	1
3		3	1	2	3
4		1	2	2	1
5		2	2	3	3
6		3	2	1	2
7		1	3	1	3
8		2	3	2	2
9		3	3	3	1

六、 $L_{27}(3^{13})$

试验号	列号	1	2	3	4	5	6	7	8	9	10	11	12	13
		1	1	3	2	1	2	2	3	1	2	1	3	2
1		1	1	1	1	1	1	3	3	2	1	1	2	1
2		2	1	1	1	1	1	3	3	2	1	1	2	1
3		3	1	2	3	1	3	1	3	3	3	1	1	2
4		1	2	2	1	1	2	2	2	3	1	3	1	1
5		2	2	3	3	1	1	3	2	1	3	3	3	2
6		3	2	1	2	1	3	1	2	2	2	3	2	3
7		1	3	1	3	1	2	2	1	2	3	2	2	2
8		2	3	2	2	1	1	3	1	3	2	2	1	3
9		3	3	3	1	1	3	1	1	1	1	2	3	1
10		1	1	1	1	2	3	3	1	3	2	3	3	2
11		2	1	2	3	2	2	1	1	1	1	3	2	3
12		3	1	3	2	2	1	2	1	2	3	3	1	1

续表

试验号 \ 列号	1	2	3	4	5	6	7	8	9	10	11	12	13
13	1	2	3	3	2	3	3	3	2	1	2	1	3
14	2	2	1	2	2	2	1	3	3	3	2	3	1
15	3	2	2	1	2	1	2	3	1	2	2	2	2
16	1	3	2	2	2	3	3	2	1	3	1	2	1
17	2	3	3	1	2	2	1	2	2	2	1	1	2
18	3	3	1	3	2	1	2	2	3	1	1	3	3
19	1	1	2	3	3	1	1	2	2	2	2	3	1
20	2	1	3	2	3	3	2	2	3	1	2	2	2
21	3	1	1	1	3	2	3	2	1	3	2	1	3
22	1	2	1	2	3	1	1	1	1	1	1	1	2
23	2	2	2	1	3	3	2	1	2	3	1	3	3
24	3	2	3	3	3	2	3	1	3	2	1	2	1
25	1	3	3	1	3	1	1	3	3	3	3	2	3
26	2	3	1	3	3	3	2	3	1	2	3	1	1
27	3	3	2	2	3	2	3	3	2	1	3	3	2

七、 $L_{18}(6^1 \times 3^6)$

试验号 \ 列号	1	2	3	4	5	6	7
1	1	1	3	2	2	1	2
2	1	2	1	1	1	2	1
3	1	3	2	3	3	3	3
4	2	1	2	1	2	3	1
5	2	2	3	3	1	1	3
6	2	3	1	2	3	2	2
7	3	1	1	3	1	3	2

续表

试验号 \ 列号	1	2	3	4	5	6	7
8	3	2	2	2	3	1	1
9	3	3	3	1	2	2	3
10	4	1	1	1	3	1	3
11	4	2	2	3	2	2	2
12	4	3	3	2	1	3	1
13	5	1	3	3	3	2	1
14	5	2	1	2	2	3	3
15	5	3	2	1	1	1	2
16	6	1	2	2	1	2	3
17	6	2	3	1	3	3	2
18	6	3	1	3	2	1	1

八、 $L_{18}(2^1 \times 3^7)$

试验号 \ 列号	1	2	3	4	5	6	7	8
1	1	1	1	3	2	2	1	2
2	1	2	1	1	1	1	2	1
3	1	3	1	2	3	3	3	3
4	1	1	2	2	1	2	3	1
5	1	2	2	3	3	1	1	3
6	1	3	2	1	2	3	2	2
7	1	1	3	1	3	1	3	2
8	1	2	3	2	2	3	1	1
9	1	3	3	3	1	2	2	3
10	2	1	1	1	1	3	1	3
11	2	2	1	2	3	2	2	2
12	2	3	1	3	2	1	3	1
13	2	1	2	3	3	3	2	1
14	2	2	2	1	2	2	3	3
15	2	3	2	2	1	1	1	2
16	2	1	3	2	2	1	2	3
17	2	2	3	3	1	3	3	2
18	2	3	3	1	3	2	1	1

九、 $L_8(4^1 \times 2^4)$

试验号	列号					
		1	2	3	4	5
1		1	1	2	2	1
2		3	2	2	1	1
3		2	2	2	2	2
4		4	1	2	1	2
5		1	2	1	1	2
6		3	1	1	2	2
7		2	1	1	1	1
8		4	2	1	2	1

十、 $L_{16}(4^5)$

试验号	列号					
		1	2	3	4	5
1		1	2	3	2	3
2		3	4	1	2	2
3		2	4	3	3	4
4		4	2	1	3	1
5		1	3	1	4	4
6		3	1	3	4	1
7		2	1	1	1	3
8		4	3	3	1	2
9		1	1	4	3	2
10		3	3	2	3	3
11		2	3	4	2	1
12		4	1	2	2	4
13		1	4	2	1	1
14		3	2	4	1	4
15		2	2	2	4	2
16		4	4	4	4	3

十一、 $L_{16}(4^4 \times 2^3)$

试验号 \ 列号	1	2	3	4	5	6	7
1	1	2	3	2	2	1	2
2	3	4	1	2	1	2	2
3	2	4	3	3	2	2	1
4	4	2	1	3	1	1	1
5	1	3	1	4	2	2	1
6	3	1	3	4	1	1	1
7	2	1	1	1	2	1	2
8	4	3	3	1	1	2	2
9	1	1	4	3	1	2	2
10	3	3	2	3	2	1	2
11	2	3	4	2	1	1	1
12	4	1	2	2	2	2	1
13	1	4	2	1	1	1	1
14	3	2	4	1	2	2	1
15	2	2	2	4	1	2	2
16	4	4	4	4	2	1	2

十二、 $L_{16}(4^3 \times 2^6)$

试验号 \ 列号	1	2	3	4	5	6	7	8	9
1	1	2	3	1	2	2	1	1	2
2	3	4	1	1	1	2	2	1	2
3	2	4	3	2	2	1	2	1	1
4	4	2	1	2	1	1	1	1	1
5	1	3	1	2	2	2	2	2	1
6	3	1	3	2	1	2	1	2	1
7	2	1	1	1	2	1	1	2	2
8	4	3	3	1	1	1	2	2	2
9	1	1	4	2	1	1	2	1	2
10	3	3	2	2	2	1	1	1	2
11	2	3	4	1	1	2	1	1	1
12	4	1	2	1	2	2	2	1	1
13	1	4	2	1	1	1	1	2	1
14	3	2	4	1	2	1	2	2	1
15	2	2	2	2	1	2	2	2	2
16	4	4	4	2	2	2	1	2	2

十三、 $L_{16}(4^2 \times 2^9)$

试验号 \ 列号	1	2	3	4	5	6	7	8	9	10	11
1	1	2	2	1	1	2	2	1	1	1	2
2	3	4	1	1	1	1	2	2	1	2	2
3	2	4	2	2	1	2	1	2	1	1	1
4	4	2	1	2	1	1	1	1	1	2	1
5	1	3	1	2	1	2	2	2	2	2	1
6	3	1	2	2	1	1	2	1	2	1	1
7	2	1	1	1	1	2	1	1	2	2	2
8	4	3	2	1	1	1	1	2	2	1	2
9	1	1	2	2	2	1	1	2	1	2	2
10	3	3	1	2	2	2	1	1	1	1	2
11	2	3	2	1	2	1	2	1	1	2	1
12	4	1	1	1	2	2	2	2	1	1	1
13	1	4	1	1	2	1	1	1	2	1	1
14	3	2	2	1	2	2	1	2	2	2	1
15	2	2	1	2	2	1	2	2	2	1	2
16	4	4	2	2	2	2	2	1	2	2	2

十四、 $L_{16}(4^1 \times 2^{12})$

试验号 \ 列号	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	1	2	2	1	2	1	2	2	1	1	1	2
2	3	2	2	1	1	1	1	1	2	2	1	2	2
3	2	2	2	2	2	1	1	2	1	2	1	1	1
4	4	1	2	1	2	2	1	1	1	1	1	2	1
5	1	2	1	1	2	2	1	2	2	2	2	2	1
6	3	1	1	2	2	1	1	1	2	1	2	1	1
7	2	1	1	1	1	1	1	2	1	1	2	2	2
8	4	2	1	2	1	2	1	1	1	2	2	1	2
9	1	1	1	2	2	1	2	1	1	2	1	2	2
10	3	2	1	1	2	2	2	2	1	1	1	1	2
11	2	2	1	2	1	2	2	1	2	1	1	2	1
12	4	1	1	1	1	1	2	2	2	2	1	1	1
13	1	2	2	1	1	1	2	1	1	1	2	1	1
14	3	1	2	2	1	2	2	2	1	2	2	2	1
15	2	1	2	1	2	2	2	1	2	2	2	1	2
16	4	2	2	2	2	1	2	2	2	1	2	2	2

十五、 $L_{25}(5^6)$

试验号	列号					
	1	2	3	4	5	6
1	1	1	2	4	3	2
2	2	1	5	5	5	4
3	3	1	4	1	4	1
4	4	1	1	3	1	3
5	5	1	3	2	2	5
6	1	2	3	3	4	4
7	2	2	2	2	1	1
8	3	2	5	4	2	3
9	4	2	4	5	3	5
10	5	2	1	1	5	2
11	1	3	1	5	2	1
12	2	3	3	1	3	3
13	3	3	2	3	5	5
14	4	3	5	2	4	2
15	5	3	4	4	1	4
16	1	4	4	2	5	3
17	2	4	1	4	4	5
18	3	4	3	5	1	2
19	4	4	2	1	2	4
20	5	4	5	3	3	1
21	1	5	5	1	1	5
22	2	5	4	3	2	2
23	3	5	1	2	3	4
24	4	5	3	4	5	1
25	5	5	2	5	4	3

十六、 $L_{12}(3^1 \times 2^4)$

试验号	列号	1	2	3	4	5
1		2	1	1	1	2
2		2	2	1	2	1
3		2	1	2	2	2
4		2	2	2	1	1
5		1	1	1	2	2
6		1	2	1	2	1
7		1	1	2	1	1
8		1	2	2	1	2
9		3	1	1	1	1
10		3	2	1	1	2
11		3	1	2	2	1
12		3	2	2	2	2

十七、 $L_{12}(6^1 \times 2^2)$

试验号	列号	1	2	3
1		1	1	1
2		2	1	2
3		1	2	2
4		2	2	1
5		3	1	2
6		4	1	1
7		3	2	1
8		4	2	2
9		5	1	1
10		6	1	2
11		5	2	2
12		6	2	1

十八、 $L_{24}(3^1 \times 2^{16})$

试验号	列号																
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	17
1		2	1	1	1	2	2	1	2	1	2	2	1	2	1	1	2
2		2	2	1	2	1	2	1	1	2	2	2	2	2	1	2	1
3		2	1	2	2	2	2	2	1	2	2	1	1	2	1	2	2
4		2	2	2	1	1	2	2	2	2	1	2	1	2	2	2	1
5		1	1	1	2	2	1	2	2	2	1	2	2	2	1	1	2
6		1	2	1	2	1	1	2	2	1	2	1	1	2	2	1	2
7		1	1	2	1	1	1	1	2	2	2	1	2	2	1	2	1
8		1	2	2	1	2	1	2	1	1	2	2	2	2	2	1	2
9		3	1	1	1	1	2	2	1	1	1	1	2	2	1	1	1
10		3	2	1	1	2	1	1	1	2	1	1	1	2	2	1	1
11		3	1	2	2	1	1	1	1	1	1	2	1	2	1	2	2
12		3	2	2	2	2	2	1	2	1	1	1	2	2	2	2	2
13		2	2	2	2	1	1	2	1	2	1	1	2	1	1	1	2
14		2	1	2	1	2	1	2	2	1	1	1	1	1	2	1	2
15		2	2	1	1	1	1	1	2	1	1	2	2	1	1	2	2
16		2	1	1	2	2	1	1	1	1	2	1	2	1	2	2	1
17		1	2	2	1	1	2	1	1	1	2	1	1	1	1	2	2
18		1	1	2	1	2	2	1	1	2	1	2	2	1	2	1	2
19		1	2	1	2	2	2	1	1	1	2	1	1	1	2	1	1
20		1	1	1	2	1	2	1	2	2	1	1	1	2	2	1	2
21		3	2	2	2	2	1	1	2	2	2	2	1	1	1	1	1
22		3	1	2	2	1	2	2	2	1	2	2	2	1	2	1	2
23		3	2	1	1	2	2	2	2	2	1	2	1	1	2	2	1
24		3	1	1	1	1	1	2	1	2	2	2	1	1	2	2	2

十九、关于 $L_{24}(3^1 \times 4^1 \times 2^{13})$, $L_{24}(6^1 \times 2^{14})$, $L_{24}(6^1 \times 4^1 \times 2^{11})$

在上表 $L_{24}(3^1 \times 2^{16})$ 中, 把第 13 列和第 14 列的位级配合 11, 12, 21, 22 顺次换成 1, 2, 3, 4, 再取消第 2 列, 可得 $L_{24}(3^1 \times 4^1 \times 2^{13})$.

在上面的 $L_{24}(3^1 \times 2^{16})$ 或 $L_{24}(3^1 \times 4^1 \times 2^{13})$ 中, 把第 1 列和第 15 列的位级配合 11, 12, 21, 22, 31, 32 顺次换成 1, 2, 3, 4, 5, 6, 再取消第 16 列, 可得 $L_{24}(6^1 \times 2^{14})$ 或 $L_{24}(6^1 \times 4^1 \times 2^{11})$.

二十、 $L_{20}(5^1 \times 2^8)$

试验号	列号	1	2	3	4	5	6	7	8	9
1		4	2	1	2	2	1	2	1	2
2		4	1	1	1	1	2	1	2	2
3		4	2	2	2	2	2	1	2	1
4		4	1	2	1	1	1	2	1	1
5		2	1	1	2	2	2	1	1	2
6		2	2	1	1	1	1	1	1	1
7		2	1	2	1	2	1	2	2	2
8		2	2	2	2	1	2	2	2	1
9		5	1	1	1	2	2	1	2	1
10		5	2	1	1	1	1	2	2	2
11		5	2	2	2	2	1	1	1	2
12		5	1	2	2	1	2	2	1	1
13		3	1	1	2	2	1	2	2	1
14		3	2	1	1	2	2	2	1	1
15		3	1	2	2	1	1	1	2	2
16		3	2	2	1	1	2	1	1	2
17		1	1	1	2	1	1	1	1	1
18		1	2	1	2	1	2	2	2	2
19		1	2	2	1	2	1	1	2	1
20		1	1	2	1	2	2	2	1	2

二十一、 $L_{16}(8^1 \times 2^8)$

试验号 \ 列号	1	2	3	4	5	6	7	8	9
1	1	2	1	2	1	2	2	1	1
2	2	2	1	1	1	1	2	2	2
3	3	2	2	1	1	2	1	2	1
4	4	2	2	2	1	1	1	1	2
5	5	1	2	2	1	2	2	2	2
6	6	1	2	1	1	1	2	1	1
7	7	1	1	1	1	2	1	1	2
8	8	1	1	2	1	1	1	2	1
9	1	1	2	1	2	1	1	2	2
10	2	1	2	2	2	2	1	1	1
11	3	1	1	2	2	1	2	1	2
12	4	1	1	1	2	2	2	2	1
13	5	2	1	1	2	1	1	1	1
14	6	2	1	2	2	2	1	2	2
15	7	2	2	2	2	1	2	2	1
16	8	2	2	1	2	2	2	1	2

习题十九

1. 安排试验方案：

(1) 某轴承厂为了提高轴承圈退火的质量, 制定因素位级表如下:

因 素	上升温度 (℃)	保温时间 (h)	出炉温度 (℃)
位级 1	800	6	400
位级 2	820	8	500

用正交表 $L_4(2^3)$ 安排试验, 并写出第 3 号试验条件.

(2) 高州县良种繁殖场为了提高水稻产量, 制定因素位级表如下:

因 素	品 种	密度(万棵/亩)	施肥量(斤/亩)
位级 1	窄叶青 8 号	30	纯氮 10
位级 2	南二矮 5 号	25	纯氮 5
位级 3	珍珠矮 11 号	20	纯氮 15

用正交表 $L_9(3^4)$ 安排试验，并写出第 8 号试验条件。

(3) 北京化工厂为了处理含有毒性物质锌和镉的废水，摸索沉淀法条件，制定因素位级表如下：

因 素	pH 值	凝聚剂	沉淀剂	CaCl ₂	废水浓度
位级 1	7~8	加	NaOH	不加	稀 浓
位级 2	8~9	不加	Na ₂ CO ₃	加	
位级 3	9~10				
位级 4	10~11				

用正交表 $L_8(4^1 \times 2^4)$ 安排试验，并写出第 6 号试验条件。

2. 试验结果的计算：

现将第 1 题的三组试验结果列在下表中：

试验号	(1)硬度合格率(%)	(2)亩产(斤)	(3)综合评分
1	100	839	45
2	45	761	70
3	85	688	55
4	70	734	65
5		774	85
6		754	95
7		843	90
8		676	100
9		726	

此处(3)中的综合评分,是根据八个试验去锌去镉的效果,综合给出的,实用中常利用综合评分的方法来处理多指标的试验.在每个小题所列出的试验方案右边,填上试验结果,通过表格化计算,求出可能好的配合及各因素的极差 R ,对于位级个数 ≥ 3 的定因素作出试验结果之和与位级用量之间的趋势图.

* 第九章 统计决策与贝叶斯统计大意^①

§ 1 统计决策问题概述

我们已经比较详细地讨论了估计和假设检验. 这二者可以看成是更一般的“统计决策”的特殊情形. 统计决策由四个要素组成.

设随机变量 X 的分布函数是 $F(x, \theta)$, θ 是未知参数, $\theta \in \Theta$, 这个 Θ 叫做参数空间, $X = (X_1, X_2, \dots, X_n)$ 是 X 的样本. 又设 A 是某项实际工作中可能采取的各种行动所组成的非空集合, A 叫做行动空间. $L(\theta, a)$ 是二元函数 ($\theta \in \Theta, a \in A$), 它表示参数是 θ 时采取行动 a 引起的损失, $L(\theta, a)$ 叫做损失函数. $(\Theta, X, A, L(\theta, a))$ 叫做统计决策问题的四个要素.

统计决策问题是: 如何根据样本 X 的值恰当地选取行动 a 使得引起的损失尽可能的小.

下面要对这句话的确切含义进行论述.

定义 1.1 称样本空间(即样本所有可能值组成的集合)到行动空间 A 的映射 $\delta = \delta(x_1, x_2, \dots, x_n)$ 为决策函数, 简称决策.

换句话说, 决策 $\delta = \delta(x_1, x_2, \dots, x_n)$ 乃是一个规则: 当样本值是 (x_1, x_2, \dots, x_n) 时采取的行动是 $\delta(x_1, x_2, \dots, x_n)$.

定义 1.2 设 $\delta = \delta(x_1, x_2, \dots, x_n)$ 是一个决策, 称平均损失

$$R(\theta, \delta) = E[L(\theta, \delta(X_1, X_2, \dots, X_n))]$$

① 本章内容取自陈家鼎、孙山泽、李东风合著的《数理统计学讲义》(北京: 高等教育出版社, 1993)

为 δ 的风险.

自然想到要找风险最小的决策,但风险是 θ 的函数,对一切 θ 风险最小的决策(所谓一致最优决策)难得存在.

在估计问题中通常取 $A = \Theta$, 损失函数 $L(\theta, a)$ 的类型很多, 当 $A = \Theta = (\underline{\theta}, \bar{\theta})$ (有限或无限区间)时, 比较常用的有

$$L(\theta, a) = (\theta - a)^2 \quad (1.1)$$

$$L(\theta, a) = |\theta - a| \quad (1.2)$$

(1.1) 叫做平方损失,(1.2) 叫绝对偏差损失. 在估计问题中的决策就是估计量, 在平方损失下, 一致最优决策就是均方误差最小的估计量.

在假设检验问题中, 参数空间 $\Theta = \Theta_0 \cup \Theta_1$, 这里 Θ_0 与 Θ_1 不相交, 二者都非空. 假设 H_i 是 " $\theta \in \Theta_i$ " ($i = 0, 1$), 行动空间 $A = \{a_0, a_1\}$, 其中 a_i 表示接受 H_i ($i = 0, 1$). 损失函数的类型很多, 比较常见的有

$$L(\theta, a_i) = \begin{cases} 0 & \theta \in \Theta_i \\ 1 & \theta \notin \Theta_i \end{cases} \quad (i = 0, 1), \quad (1.3)$$

这时 $L(\theta, a)$ 叫做 0-1 损失.

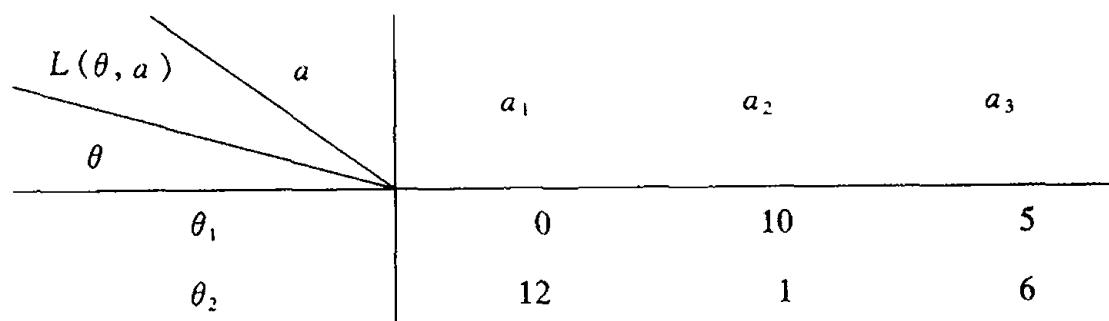
不难看出, 对于假设检验问题, 给出决策等价于给出对 H_0 的否定域.

例 1.1 检查某设备零件:

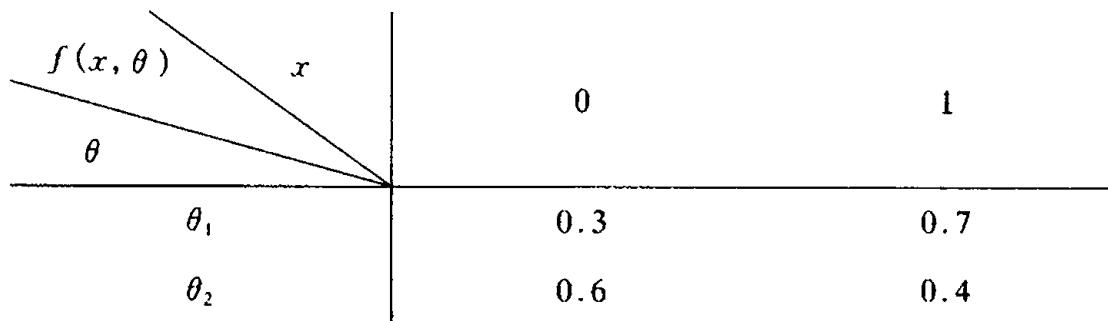
零件的可能状态: θ_1 (好), θ_2 (坏).

可能采取的行动: a_1 (保留), a_2 (更换), a_3 (修理).

损失函数为:



样本: X 的取值为 0 或 1(为判断零件的状态, 在设备工作时用手摸零件, 温度正常时则记 $X = 1$, 发烫则记 $X = 0$). X 的概率函数为



问: 应采取怎样的决策使得损失尽可能小?

可能的决策函数有 9 个, 列表如下:

X	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
0	a_1	a_1	a_1	a_2	a_2	a_2	a_3	a_3	a_3
1	a_1	a_2	a_3	a_1	a_2	a_3	a_1	a_2	a_3

风险为

$$\begin{aligned}
 R(\theta, \delta) &= E[L(\theta, \delta(X))] \\
 &= L(\theta, a_1)P[\delta(X)=a_1] + L(\theta, a_2)P[\delta(X)=a_2] \\
 &\quad + L(\theta, a_3)P[\delta(X)=a_3]
 \end{aligned}$$

例如

$$R(\theta_1, \delta_2) = 0 \times 0.3 + 10 \times 0.7 + 5 \times 0 = 7$$

$$R(\theta_2, \delta_2) = 12 \times 0.6 + 1 \times 0.4 + 6 \times 0 = 7.6$$

9 个决策的风险如下

δ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
$R(\theta_1, \delta)$	0	7	3.5	3	10	6.5	1.5	8.5	5
$R(\theta_2, \delta)$	12	7.6	9.6	5.4	1	3	8.4	4.0	6

哪个决策好呢？有些是明显的，如 δ_4 优于 $\delta_2, \delta_3, \delta_9$. 但 δ_4 与 δ_1 哪个好呢？这就不好回答了.

定义 1.3 称决策 $\delta = \delta(x_1, x_2, \dots, x_n)$ 是容许的，如果不存在另一决策 δ' 使得

$$R(\theta, \delta') \leq R(\theta, \delta) \quad (\text{一切 } \theta \in \Theta)$$

且对至少一个 θ , 严格的不等式成立.

不难看出，在前面的例 1.1 中， $\delta_1, \delta_4, \delta_5, \delta_6, \delta_7$ 是容许的， $\delta_2, \delta_8, \delta_9$ 是不容许的，而且不存在一致最优的决策.

如果一致最优决策不存在，只好去找比较“优良”的决策了. 有两个途径探讨这个问题. 一是对决策 δ 作一定的限制，缩小选择的范围，在较小的范围内找最优的，例如在估计问题中要求 δ 是无偏的估计，在所有无偏估计中找最优的. 另一是放宽比较的要求，不要求风险函数对一切 θ 都最小.

定义 1.4 称决策 δ^* 是 minimax 决策，若对一切决策 δ 成立

$$\sup_{\theta} R(\theta, \delta^*) \leq \sup_{\theta} R(\theta, \delta)$$

minimax 决策是一种保守的决策. 不难看出，在例 1.1 中 δ_4 是 minimax 决策.

当然，在许多情形下应尽量避免采用保守的决策. 为此，应对参数 θ 有所了解，知道哪些 θ 值出现机会较多，哪些 θ 值出现机会较少. 换句话说，应将 θ 看成一个随机变量，它有确定的概率分布. 在这种情形下，风险 $R(\theta, \delta)$ 关于 θ 的分布的平均值 $\rho(\delta) = E[R(\theta, \delta)]$ 是评价决策 δ 的优良性指标，数值越小越好. 把未知参数 θ 看成是随机变量，这是一种重要的观点，对统计学有重大影响，下面要作进一步介绍.

还应指出，统计决策问题的四个要素里损失函数 $L(\theta, a)$ 较难确定. 在实际问题里要给出恰当的损失函数，需要进行很多研究工作，在涉及几方的利益时应共同讨论以便确定各方都能接受的损失函数. 上面的(1.1)、(1.2)和(1.3)只是三个有代表性的比较

简单的损失函数.在有些决策问题里,不提损失函数 $L(\theta, a)$ 而是考虑报酬函数(或效用函数) $M(\theta, a)$, $M(\theta, a)$ 是参数为 θ 时采取行动 a 而产生的报酬.决策问题是:寻找这样的决策函数,使得平均报酬达到最大.很明显,只要引入损失函数 $L(\theta, a) = -M(\theta, a)$, 则问题化为:如何选取行动使得平均损失尽可能的小.

§2 什么是贝叶斯统计

公元 1763 年,贝叶斯(T. Bayes, 1702—1761, 英国长老会的牧师)的著作《论有关机遇问题的求解》发表了.这篇著作里提出了一种归纳推理的理论,以后被一些统计学者发展为一种系统的统计推断方法,称为贝叶斯方法.采用这种方法作统计推断所得的全部结果,构成贝叶斯统计的内容.认为贝叶斯方法是惟一合理的统计推断方法的统计学者,组成统计学中的贝叶斯学派.这个学派始自 20 世纪 30 年代,到 50、60 年代引起人们广泛的注意.时至今日,其影响有日益扩大之势.

设总体 X 有分布密度(或概率函数) $f(x, \theta)$, 其中 θ 是未知的,但知 θ 的变化范围是 Θ , 经典方法把参数 θ 看作是客观常数,通过对样本 $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n)$ 的研究对 θ 给出估计值或者推断 θ 属于某个给定的范围.贝叶斯学派的根本观点,是认为在关于 θ 的任何统计推断问题中,除了使用样本 \tilde{X} 提供的信息外,还必须对 θ 规定一个先验分布,它是进行推断时不可缺少的要素.说得更明确些,应把 θ 看成随机变量(为了强调这一点,有时把 θ 写成 $\bar{\theta}$),它服从某个概率分布(叫做先验分布),总体 X 的分布实际上是 θ 给定时 X 的条件分布.贝叶斯学派把先验分布看成是在抽样(或观测)前就有的关于 θ 的先验信息的概率表述,先验分布可能有客观的依据,也可以部分地或是全部地基于主观信念.根据 X 的分布密度(或概率函数) $f(x, \theta)$ 及 θ 的先验分布密度(或先验概

率函数) $\xi(\theta)$, 可以算出在样本 $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_n) = \tilde{x} = (x_1, x_2, \dots, x_n)$ 的条件下, θ 的条件分布密度(或条件概率函数) $\xi(\theta | \tilde{x})$. 因为这个分布是在抽样(或观测)之后得到的, 故称为后验分布. 贝叶斯学派认为, 这个分布综合了样本 \tilde{X} 及先验分布密度(或概率函数) $\xi(\theta)$ 所提供的信息. 抽样(或观测)的全部目的就在于完成由先验分布到后验分布的转换. 贝叶斯方法的关键在于所作出的任何推断都只须根据后验分布 $\xi(\theta | \tilde{x})$, 而不再涉及样本 \tilde{X} 的分布. 但在如何使用 $\xi(\theta | \tilde{x})$ 上还有较大的灵活性, 涉及到行动空间和损失函数的类型.

沿用 § 1 中的记号, 用 A 表示行动空间, 用 $L(\theta, a)$ 表示损失函数. 若参数 θ 的先验分布是 ξ , 决策是 $\delta = \delta(x_1, x_2, \dots, x_n)$, 则平均风险为 $\rho(\delta) = E_\xi[R(\theta, \delta)]$, 这里 $R(\theta, \delta)$ 是风险, E_ξ 是关于 θ 的概率分布 ξ 取平均值.

定义 2.1 称 $\delta^* = \delta^*(x_1, x_2, \dots, x_n)$ 是贝叶斯决策, 若

$$\rho(\delta^*) = \inf_{\delta} \rho(\delta)$$

此时 $\rho(\delta^*)$ 叫做贝叶斯风险.

要注意的是, 贝叶斯决策依赖于先验分布. 先验分布变了, 贝叶斯决策一般也要变.

例 2.1 在例 1.1 中, 若 θ 的先验分布(概率因数)是 $\xi(\theta_1) = 0.7, \xi(\theta_2) = 0.3$, 则

$$\rho(\delta) = 0.7 \times R(\theta_1, \delta) + 0.3 \times R(\theta_2, \delta)$$

各个决策的 $\rho(\delta)$ 值如下

δ	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
$\rho(\delta)$	3.6	5.13	12.48	3.72	7.20	5.45	3.57	7.15	5.3

从上表可看出

δ_1 : 发烫($X=0$), 就修理(a_3)
 δ_2 : 正常($X=1$), 就保留(a_1)

是贝叶斯决策.

若先验分布改为 $\xi(\theta_1) = 0.9, \xi(\theta_2) = 0.1$, 则贝叶斯决策是 δ_1 .

贝叶斯统计的基本任务就是针对先验分布 $\xi(\theta)$, 找出贝叶斯决策.

怎样寻找贝叶斯决策呢? 例 2.1 所处理的是很简单的情形, 穷举所有的决策, 加以比较后找出贝叶斯决策. 一般情形下不能这样做, 而要利用后验分布. 设 $\xi(\theta | x_1, x_2, \dots, x_n)$ 是在样本 $(X_1, X_2, \dots, X_n) = (x_1, x_2, \dots, x_n)$ 下 θ 的条件分布密度(后验分布^①). 令

$$\rho(x_1, x_2, \dots, x_n) = \inf_{a \in A} \int_{\Theta} L(\theta, a) \xi(\theta | x_1, x_2, \dots, x_n) d\theta \quad (2.1)$$

这个 $\rho(x_1, x_2, \dots, x_n)$ 乃是样本值为 x_1, x_2, \dots, x_n 时的最小后验平均损失.

若决策 $\delta^* = \delta^*(x_1, x_2, \dots, x_n)$ 满足

$$\begin{aligned} & \int_{\Theta} L(\theta, \delta^*(x_1, x_2, \dots, x_n)) \xi(\theta | x_1, x_2, \dots, x_n) d\theta \\ &= \rho(x_1, x_2, \dots, x_n) \end{aligned} \quad (2.2)$$

(对一切 x_1, x_2, \dots, x_n)

则 δ^* 就是贝叶斯决策. 这个结论的数学证明用到较深的数学知识, 从略. 许多书上干脆把满足(2.2)的 δ^* 定义为贝叶斯决策. (2.2)的意义是: 采用决策 δ^* 引起的后验平均损失达到最小值.

(2.2)是寻找贝叶斯决策的基本出发点. 为此首先需要求出后验分布.

^① 当 θ 是离散型随机变量时, 可进行类似的讨论, 下面公式里的积分要改为求和.

怎样计算后验分布呢? 可以证明, 当 θ 的先验分布密度是 $\xi(\theta)$ 时, 在 $X = (x_1, x_2, \dots, x_n)$ 的条件下 θ 的后验分布密度是

$$\xi(\theta | x_1, x_2, \dots, x_n) = \frac{\prod_{i=1}^n f(x_i, \theta) \xi(\theta)}{\int_{\Theta} \prod_{i=1}^n f(x_i, \theta) \xi(\theta) d\theta} \quad (2.3)$$

当 $\Theta = \{\theta_1, \theta_2, \dots\}$ 时, (2.3) 要用下式代替

$$\xi(\theta | x_1, x_2, \dots, x_n) = \frac{\prod_{i=1}^n f(x_i, \theta) \xi(\theta)}{\sum_j \prod_{i=1}^n f(x_i, \theta_j) \xi(\theta_j)} \quad (2.4)$$

这时(2.4)是 θ 的后验概率函数.

(2.3)的严格证明涉及到零概率事件发生下如何计算条件概率的问题, 要用测度论才能彻底处理. 但是, 由于 $\prod_{i=1}^n f(x_i, \theta) \xi(\theta)$ 是随机变量 $X_1, X_2, \dots, X_n, \bar{\theta}$ 的联合密度, 利用第四章中关于条件分布的论述, 在一定条件下(在实际工作中这些条件常常是满足的), 条件分布密度确由(2.3)来表达. 当 X 及 $\bar{\theta}$ 是离散型随机变量时, 很易证明(2.4)成立. 实际上, 这时 $\xi(\theta_j) = P(\bar{\theta} = \theta_j)$, $f(x_i; \theta_j) = P(X = x_i | \bar{\theta} = \theta_j)$, 于是

$$\begin{aligned} & P(\bar{\theta} = \theta_j | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ &= \frac{P(\bar{\theta} = \theta_j, X_1 = x_1, \dots, X_n = x_n)}{P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)} \\ &= \frac{P(\bar{\theta} = \theta_j) \prod_{i=1}^n P(X_i = x_i | \bar{\theta} = \theta_j)}{\sum_j P(\bar{\theta} = \theta_j, X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)} \\ &= \frac{\xi(\theta_j) \prod_{i=1}^n f(x_i, \theta_j)}{\sum_j \prod_{i=1}^n f(x_i, \theta_j) \xi(\theta_j)} \end{aligned}$$

这就证明了(2.4). (2.4)就是初等概率论中的贝叶斯公式.

例 2.2 设 X 服从伯努利分布, 参数是 p , 即

$$P(X=1)=p=1-P(X=0), \quad 0 < p < 1$$

设 p 的先验分布 ξ 是 $(0,1)$ 上的均匀分布. 从(2.3)知在样本 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ 下 p 的后验分布密度是

$$\xi(p|x_1, x_2, \dots, x_n) = \frac{p^{\sum_{i=1}^n x_i} \cdot (1-p)^{n - \sum_{i=1}^n x_i}}{\int_0^1 p^{\sum_{i=1}^n x_i} \cdot (1-p)^{n - \sum_{i=1}^n x_i} dp} \quad (2.5)$$

这是参数为 $\sum_{i=1}^n x_i + 1, n + 1 - \sum_{i=1}^n x_i$ 的 β 分布.

设行动空间 $A = (0, 1)$, 损失函数是(1.1)(平方损失), 则(2.2)式化为

$$\begin{aligned} & \int_0^1 (p - \delta^*(x_1, x_2, \dots, x_n))^2 \xi(p|x_1, x_2, \dots, x_n) dp \\ &= \inf_{a \in (0,1)} \int_0^1 (p - a)^2 \xi(p|x_1, x_2, \dots, x_n) dp \end{aligned} \quad (2.6)$$

从(2.6)式出发可以证明: $\delta^*(x_1, x_2, \dots, x_n)$ 恰好是后验分布 $\xi(p|x_1, x_2, \dots, x_n)$ 的均值. 利用(2.5)知

$$\begin{aligned} \delta^*(x_1, x_2, \dots, x_n) &= \int_0^1 p \xi(p|x_1, x_2, \dots, x_n) dp \\ &= \frac{1}{n+2} \left(\sum_{i=1}^n x_i + 1 \right) \end{aligned}$$

这就是 p 的贝叶斯估计.

例 2.3 设随机变量 X 的密度函数是 $f(x, \theta), \theta \in \Theta, \Theta = \Theta_0 \cup \Theta_1, \Theta_1$ 与 Θ_0 不相交, θ 的先验分布密度是 $\xi(\theta)$. 检验问题是:

$$H_0: \theta \in \Theta_0 \leftrightarrow H_1: \theta \in \Theta_1$$

设 $A = \{a_0, a_1\}$ (a_i 表示接受假设 $H_i, i = 0, 1$). 若 X 的样本 (X_1, X_2, \dots, X_n) 取值 (x_1, x_2, \dots, x_n) , θ 的后验分布为 $\xi(\theta|x_1, x_2, \dots, x_n)$, 又 $L(\theta, a)$ 是 0-1 损失(见(1.3)). 则

$$\begin{aligned}
& \int_{\Theta} L(\theta, a_i) \xi(\theta | x_1, \dots, x_n) d\theta \\
&= P(\tilde{\theta} \in \Theta_i | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\
&= 1 - P(\tilde{\theta} \in \Theta_i^c | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \quad (i = 0, 1)
\end{aligned}$$

从(2.2)知贝叶斯决策为

$$\delta^*(x_1, x_2, \dots, x_n) = \begin{cases} a_0 & \text{当 } P(\tilde{\theta} \in \Theta_0 | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \geq \frac{1}{2} \\ a_1 & \text{否则} \end{cases}$$

换句话说,贝叶斯检验是:当且仅当条件概率

$$P(\tilde{\theta} \in \Theta_0 | X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

不小于 $\frac{1}{2}$ 时接受假设 H_0 .

对于一般的决策问题(包含估计、假设检验在内),可以证明在相当广泛的条件下,贝叶斯决策是存在的,读者可参看[19].

作为本节的末尾,我们还要简略地介绍贝叶斯序贯统计决策.

设 X 有分布密度(或概率函数) $f(x, \theta)$, θ 取值属于 Θ , 有先验分布 $\xi(\theta)$, A 是行动空间, $L(\theta, a)$ 是损失函数. 设对 X 进行逐次观测, 得 X_1, X_2, \dots , 每次观测的费用是 $C > 0$.

序贯统计决策由停止法则 τ 及判决法则 δ 组成. 停止法则 τ 告诉我们何时停止观测, 判决法则 δ 告诉我们, 如何根据序贯样本 X_1, X_2, \dots, X_τ 选取行动 $\delta(X_1, X_2, \dots, X_\tau)$ (δ 的值属于 A).

序贯统计决策 (τ, δ) 的总风险为

$$\rho(\tau, \delta) = E [L(\theta, \delta(X_1, X_2, \dots, X_\tau)) + \tau C]$$

(损失和观测费用之和的平均值). 这里 E 表示 θ 之先验分布为 ξ 及给定 θ 时诸 X_i 的分布密度(或概率函数)为 $f(x, \theta)$ 时计算期望.

定义 2.2 称 (τ^*, δ^*) 是贝叶斯序贯决策, 若

$$\rho(\tau^*, \delta^*) = \inf_{(\tau, \delta)} \rho(\tau, \delta)$$

可以证明, 在相当广泛的条件下, 贝叶斯序贯决策是存在的.

在一些特殊情形下,可以具体找出贝叶斯序贯决策(包括停止法则 τ^* 和判决法则 δ^*).读者可参看[19].

§ 3 先验分布的确定

如何确定先验分布,这是贝叶斯统计中最困难,也是使用贝叶斯方法必须解决但又最易引起争议的问题.这个问题现代有很多研究成果,但还没有圆满的理论与普遍有效的方法.从实用角度出发,应充分利用专家的经验或者对历史上积累的数据进行分析和拟合,以确定先验分布.在确定先验分布时,许多人利用下列的协调性假说.

协调性假说:若总体 X 的分布密度(或概率函数)是 $f(x, \theta)$,则 θ 的先验分布与由它和 X 的样本确定的后验分布应属于同一类型.这时先验分布叫做是与 $f(x, \theta)$ 共轭的分布.

这里未对“同一类型”四字给出精确的定义,也很难给出恰当的定义.通常的理解是,将概率性质相似的所有分布算作同一类型.例如,所有正态分布构成一类;所有 Γ 分布构成一类;所有 β 分布构成一类.

这个假说指示我们,先验分布应取何种类型,然后再利用历史数据来确定先验分布中的未知部分.许多实践表明,这个假说是符合实际的.

以下我们要对一些常见的分布找出其共轭的先验分布.

定理 3.1 设 X_1, X_2, \dots, X_n 是来自伯努利分布(参数是 $p, 0 < p < 1$)的样本.若 p 的先验分布是 β 分布,参数是 α, β ,则在 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ 下 p 的后验分布是参数为 $\alpha + y, \beta + n - y$ 的 β 分布,这里 $y = \sum_{i=1}^n x_i$.

证 X_1, X_2, \dots, X_n 的联合概率函数为

$$f_n(x_1, x_2, \dots, x_n; p) = \binom{n}{\sum_{i=1}^n x_i} p^{\sum_{i=1}^n x_i} \cdot (1-p)^{n - \sum_{i=1}^n x_i}$$

p 的先验分布密度

$$\xi(p) = C(\alpha, \beta) p^{\alpha-1} (1-p)^{\beta-1}$$

(这里 $C(\alpha, \beta)$ 是常数). 从公式(2.3)知 p 的后验分布密度为

$$\begin{aligned} \xi(p | x_1, x_2, \dots, x_n) &= \frac{f_n(x_1, x_2, \dots, x_n; p) \xi(p)}{\int_0^1 f_n(x_1, x_2, \dots, x_n; p) \xi(p) dp} \\ &= A p^{\alpha + \sum_{i=1}^n x_i - 1} \cdot (1-p)^{\beta + n - \sum_{i=1}^n x_i - 1} \end{aligned}$$

这里 A 是与 p 无关的数, 可见后验分布是参数为 $\alpha + y, \beta + n - y$ 的 β 分布. 证毕.

定理 3.1 告诉我们, β 分布族是伯努利分布族的共轭分布族.

定理 3.2 设 X_1, X_2, \dots, X_n 是来自泊松分布的样本. 泊松分布的参数是 λ , λ 的先验分布是 Γ 分布(参数是 α, β), 则在 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ 下 λ 的后验分布是参数为 $\alpha + \sum_{i=1}^n x_i, \beta + n$ 的 Γ 分布.

证 X_1, X_2, \dots, X_n 的联合概率函数为

$$f_n(x_1, x_2, \dots, x_n; \lambda) = \frac{\lambda^{x_1 + x_2 + \dots + x_n}}{x_1! x_2! \dots x_n!} e^{-n\lambda}$$

λ 的先验分布密度是

$$\xi(\lambda) = C \lambda^{\alpha-1} e^{-\beta\lambda} \quad (\lambda > 0)$$

根据(2.3)知 λ 的后验分布密度为

$$\xi(\lambda | x_1, x_2, \dots, x_n) = A \lambda^{\alpha + \sum_{i=1}^n x_i - 1} \cdot e^{-(\beta + n)\lambda} \quad (\lambda > 0)$$

其中 A 是与 λ 无关的常数. 这表明后验分布是参数为 $\alpha + \sum_{i=1}^n x_i, \beta + n$ 的 Γ 分布. 证毕.

定理 3.3 设 X_1, X_2, \dots, X_n 是来自指数分布 $F(x) = 1 - e^{-\lambda x}$ ($\lambda > 0, x > 0$) 的样本. 设参数 λ 的先验分布是参数为 α, β 的 Γ 分布, 则在 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ 下 λ 的后验分布是参数为 $\alpha + n, \beta + \sum_{i=1}^n x_i$ 的 Γ 分布.

证 X_1, X_2, \dots, X_n 的联合密度为

$$f_n(x_1, x_2, \dots, x_n; \lambda) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

λ 的先验分布密度为

$$\xi(\lambda) = C \lambda^{\alpha-1} e^{-\beta \lambda} \quad (\lambda > 0)$$

从公式(2.3)知 λ 的后验分布密度

$$\xi(\lambda | x_1, x_2, \dots, x_n) = A \lambda^{\alpha+n-1} \cdot e^{-(\beta + \sum_{i=1}^n x_i) \lambda}$$

可见后验分布是参数为 $\alpha + n, \beta + \sum_{i=1}^n x_i$ 的 Γ 分布. 证毕.

定理 3.4 设 X_1, X_2, \dots, X_n 是来自 $N(\mu, \sigma^2)$ 的样本, 其中 σ^2 已知, μ 未知, 设 μ 的先验分布为 $N(\mu_0, \sigma_0^2)$, 则在 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ 下, μ 的后验分布为 $N(\mu^*, (\sigma^*)^2)$, 其中

$$\mu^* = \frac{\mu_0 \sigma^2 + n \sigma_0^2 (\bar{x})}{\sigma^2 + n \sigma_0^2} \left(\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \right) \quad (3.1)$$

$$(\sigma^*)^2 = \frac{\sigma^2 \sigma_0^2}{\sigma^2 + n \sigma_0^2} \quad (3.2)$$

证 X_1, X_2, \dots, X_n 的联合密度为

$$f_n(x_1, x_2, \dots, x_n; \mu) = C \exp \left\{ - \sum_{i=1}^n (x_i - \mu)^2 / (2\sigma^2) \right\}$$

这里 C 与 μ 无关. 又 μ 的先验分布密度为

$$\xi(\mu) = C_1 \exp \left\{ - \frac{(\mu - \mu_0)^2}{2\sigma_0^2} \right\}$$

从公式(2.3)知后验分布密度为

$$\xi(\mu | x_1, x_2, \dots, x_n) = A \exp \left\{ -\frac{n}{2\sigma^2} (\mu - \bar{x})^2 - \frac{1}{2\sigma_0^2} (\mu - \mu_0)^2 \right\}$$

其中 A 与 μ 无关. 但是

$$\begin{aligned} & \frac{n}{\sigma^2} (\mu - \bar{x})^2 + \frac{1}{\sigma_0^2} (\mu - \mu_0)^2 \\ &= \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \right) \mu^2 - 2\mu \left(\frac{n(\bar{x})^2}{\sigma^2} + \frac{\mu_0}{\sigma_0^2} \right) + \frac{n}{\sigma^2} (\bar{x})^2 + \frac{\mu_0^2}{\sigma_0^2} \\ &= \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \right) (\mu - \mu^*)^2 + B \end{aligned}$$

其中 μ^* 由 (3.1) 确定, B 与 μ 无关. 于是

$$\xi(\mu | x_1, x_2, \dots, x_n) = \tilde{A} \exp \left\{ -\frac{1}{2} \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \right) (\mu - \mu^*)^2 \right\}$$

可见后验分布是 $N(\mu^*, (\sigma^*)^2)$. 证毕.

从定理 3.4 看出, μ^* 是 \bar{x} 与 μ_0 的加权平均. 这与直观相符, σ_0^2 越大, μ_0 的值越不重要.

定理 3.5 设 X_1, X_2, \dots, X_n 是来自 $N\left(\mu, \frac{1}{R}\right)$ 的样本, μ 已知, R 未知, 设 R 的先验分布是参数为 α, β 的 Γ 分布. 则在 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ 下, R 的后验分布是参数为 α', β' 的 Γ 分布, 这里 $\alpha' = \alpha + \frac{n}{2}, \beta' = \beta + \frac{1}{2} \sum_1^n (x_i - \mu)^2$.

证 X_1, X_2, \dots, X_n 的联合密度为

$$f_n(x_1, x_2, \dots, x_n; R) = \left(\sqrt{\frac{R}{2\pi}} \right)^n e^{-\frac{1}{2} R \sum_1^n (x_i - \mu)^2}$$

R 的先验分布密度为

$$\xi(R) = C_1 R^{\alpha-1} e^{-\beta R}$$

从公式(2.3)知 R 的后验分布密度为

$$CR^{\alpha+\frac{n}{2}-1} \cdot e^{-\left(\beta + \frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2\right) R}$$

可见后验分布是参数为 α', β' 的 Γ 分布. 证毕.

为了研究均匀分布的共轭分布,引进定义 3.1.

定义 3.1 称 X 服从参数为 $x_0 > 0, \alpha > 0$ 的 Pareto 分布, 若它的分布密度是

$$\varphi(x; x_0, \alpha) = \begin{cases} 0 & x < x_0 \\ \alpha x_0^\alpha x^{-\alpha-1} & x \geq x_0 \end{cases}$$

定理 3.6 设 X_1, X_2, \dots, X_n 是来自 $[0, \theta]$ 上均匀分布的样本. 若 θ 的先验分布是参数为 θ_0, α 的 Pareto 分布. 则在 $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ 下 θ 的后验分布是参数为 $\theta'_0, \alpha + n$ 的 Pareto 分布, 这里 $\theta'_0 = \max\{\theta_0, x_1, x_2, \dots, x_n\}$.

证 X_1, X_2, \dots, X_n 的联合密度是

$$\begin{aligned} f_n(x_1, x_2, \dots, x_n; \theta) &= \frac{1}{\theta^n} \prod_{i=1}^n I_{[0, \theta]}(x_i) \\ &= \frac{1}{\theta^n} I_{[0, \theta]}(\max_{1 \leq i \leq n} x_i) \end{aligned}$$

而 θ 的先验分布密度是

$$\varphi(\theta; \theta_0, \alpha) = \frac{\alpha \theta_0^\alpha}{\theta^{\alpha+1}} I_{[\theta_0, +\infty)}(\theta)$$

根据(2.3)知 θ 的后验分布密度是

$$\begin{aligned} \xi(\theta | x_1, x_2, \dots, x_n) &= A \theta^{-n} I_{[0, \theta]}(\max x_i) \cdot \alpha \theta_0^{-\alpha-1} I_{[\theta_0, +\infty)}(\theta) \\ &= C I_{[\theta'_0, +\infty)}(\theta) \cdot \theta^{-\alpha-n-1} \end{aligned}$$

这里 C 是与 θ 无关的常数, $\theta'_0 = \max(\theta_0, x_1, x_2, \dots, x_n)$. 可见, 后验分布是参数为 $\theta'_0, \alpha + n$ 的 Pareto 分布. 证毕.

以上都是讨论单个实参数的先验分布. 至于两个实参数的二维先验分布, 由于表达式比较复杂, 这里从略.

最后, 我们介绍广义先验分布的概念: Θ 上任何非负函数 $g(\theta)$ 称为 θ 的广义先验分布密度(又叫权函数). 值得注意的是, 虽然 $g(\theta)$ 可能不是概率密度函数, 但是公式

$$\xi(\theta | x_1, x_2, \dots, x_n) \stackrel{\Delta}{=} \frac{f_n(x_1, x_2, \dots, x_n; \theta) g(\theta)}{\int_{\Theta} f_n(x_1, x_2, \dots, x_n; \theta) g(\theta) d\theta}$$

(当分母有限时)仍然确定一分布密度,它仍叫做 θ 的后验分布密度.

在实际使用贝叶斯方法时,总要充分利用以往积累的知识,以确定先验分布.若以往一点知识也没有,怎么办?此时应如何确定先验分布?

这个重要问题迄今尚未解决,此时没有统一的、公认的方法确定先验分布.有些人愿意采用下列假设:

1) 若 θ 在 $(-\infty, \infty)$ 上取值,且无先验知识,则广义先验分布(权函数)应取 $g(\theta) \equiv C (C > 0)$.

2) 若 θ 在 Θ 上取值, $\varphi = \varphi(\theta)$ 是 Θ 到 Ψ 上的一一映射,则 θ 无先验知识时 φ 也无先验知识.

3) θ 在 (a, b) 上取值 ($-\infty \leq a < b \leq +\infty$), 权函数是 $g(\theta)$, φ 是 (a, b) 上有定义的函数, $\varphi'(x) > 0$ (一切 x) 或者 $\varphi'(x) < 0$ (一切 x), 则 $\psi = \varphi(\theta)$ 的权函数为

$$h(\psi) = g(\varphi^{-1}(\psi)) \frac{1}{|\varphi'(\varphi^{-1}(\psi))|} \quad (3.3)$$

这里 φ^{-1} 是 φ 的反函数.

根据这三条假设可以推出:① $(0, +\infty)$ 上对应无先验知识的权函数应取 $\frac{C}{\theta}$; ② $(0, 1)$ 上对应无先验知识的权函数可取

$$\frac{C}{\theta(1-\theta)}.$$

应该指出,上面提到的假设中的第一条和第三条缺乏科学依据,从这些假设所导出的结论也就不足为信了.

若要了解先验分布和贝叶斯统计的更多知识,请看 S. Kotz 和吴喜之合著的《现代贝叶斯统计学》(北京:中国统计出版社, 2000);至于统计判决的一般理论和方法,叙述全面又避免使用高

深数学的专著则首推 J. Berger 的《统计决策论及贝叶斯分析》(第二版, 贾乃光译, 中国统计出版社, 1998).

§ 4 应用实例——电视机寿命验证 试验的贝叶斯方法

电视机的使用寿命是很重要的质量指标, 所谓使用寿命是指从开始使用到规定的功能丧失为止所经历的时间, 当然希望寿命越长越好. 怎样验证一批电视机的寿命是否达到合格标准呢? 当然只能采取抽样检查的方法. 这实质上是个假设检验问题, 数学上常描述如下. 根据国内外的大量数据, 可以认为电视机的寿命 X 服从指数分布, 其密度函数如下

$$f(x, \theta) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{\theta} e^{-\frac{x}{\theta}} & x > 0 \end{cases}$$

其中 $\theta \in (0, +\infty)$ 是未知参数. 因为 $E(X) = \theta$, θ 叫做平均无故障工作时间(记作 MTBF), 即平均寿命. 通常取定两个值 θ_1, θ_0 , $\theta_1 < \theta_0$, 这个 θ_1 叫检验的下限值(在验收问题中, 平均寿命低于 θ_1 时不应接受, θ_1 又叫做极限寿命水平); θ_0 叫做检验的上限值(平均寿命高于 θ_0 时应该接受, θ_0 又叫合格寿命水平), 比值 $d = \theta_0/\theta_1$ 叫做鉴别比. 传统的试验方案是这样的: 从一大批电视机中随机抽取 n 台, 同时进行寿命试验, 试验 t_0 h. 在这 $n \times t_0$ 个台时的试验中, 电视机的失效数为 r . 当 $r \leq C$ 时判定整批合格, 使用方接收; 当 $r > C$ 时判定整批不合格, 使用方拒收. 这个 C 叫做合格判定数, 是一个非负整数. 通常称 $T = n \times t_0$ 为总台时数. 这个验证试验方案由三个参数 n, t_0, C 组成. 参数的选择依赖于两类错误带来的风险的允许值. 用 $L(\theta) = P_\theta(r \leq C)$ 表示平均寿命是 θ 时接收整批的概率, $L(\theta)$ 叫做抽样特性函数. $\alpha = \sup_{\theta \geq \theta_0} P_\theta(r > C)$ 叫做

生产方风险,即产品寿命合格却被拒收的概率; $\beta = \sup_{\theta < \theta_1} P_\theta(r \leq C)$

叫做使用方风险,即产品不合格却被接收的概率.可以证明
 $\alpha = P_{\theta_0}(r > C), \beta = P_{\theta_1}(r \leq C)$.有下列公式:

$$\alpha = 1 - \sum_{i=0}^C \frac{1}{i!} \left(\frac{T}{\theta_0} \right)^i e^{-\frac{T}{\theta_0}}$$

$$\beta = \sum_{i=0}^C \frac{1}{i!} \left(\frac{T}{\theta_1} \right)^i e^{-\frac{T}{\theta_1}}$$

(理由见(4.2)).利用这两个公式就可对给定的 α, β 确定出 T 和 C .

在以往通行的彩色电视机试验中,取 $\theta_1 = 15\ 000$ h, $\theta_0 = 45\ 000$ h, $n = 100$, 试验时间 $t_0 = 1\ 000$ h, 在 $T = n \times t_0 = 100\ 000$ 台时试验中失效数 $r \leq 3$ 时判为合格,这时 $\alpha = 0.20, \beta = 0.10$.可以看出,在多批交付验收过程中,仍有 20% 的质量合格批被判为质量不合格而遭拒收.

这 100 000 台时的试验实行起来很不容易,费时费电,随着彩电质量的提高,若要验证 MTBF 下限 θ_1 为 20 000 h, 则要进行 134 000 台时的试验(鉴别比仍为 3,下同). 若要验证 θ_1 为 30 000 h, 则需要 201 000 台时的试验. 上百台彩电同时进行试验需要数百立方米的高温房,仅一次试验就耗电 26 000 kW·h. 可见,这种既费时又费钱的试验方法,显然不适应彩电生产发展的形势.

由上述可见,对于高可靠性产品,用古典统计方法建立的抽样验收试验是不能令人满意的. 我们必须利用已往积累起来的关于平均寿命 θ 的信息. 换句话说,应该另辟捷径,从贝叶斯方法找出路. 我国部分统计工作者于 80 年代中期采用贝叶斯方法对国产彩电制定出新的验证方案,形成了“国家标准”,下面从数学角度对此进行介绍.

(一) 首先找出在固定总台时 T 的试验中失效数 r 服从的概率分布.

我们考虑有替换的试验. 设开始时有 n 台电视机投入试验, 一旦出现失效, 就换一台新的继续试验, 试验进行到总台时数为 T 时停止. 设 r 是停止时的累计失效数, 则有下列基本定理^①.

定理 4.1

$$P_\theta(r=k) = \frac{1}{k!} \left(\frac{T}{\theta} \right)^k e^{-\frac{T}{\theta}} \quad (k=0,1,\dots) \quad (4.1)$$

这里 P_θ 表示平均寿命是 θ 时相应的概率.

证 我们首先指出, 若 $\xi_1, \xi_2, \dots, \xi_k$ 相互独立, 都服从均值是 θ 的指数分布, 则对一切 $t > 0$, 有

$$P\left(\sum_{i=1}^k \xi_i \leq t\right) = \int_0^t \frac{1}{\theta^k} e^{-\frac{u}{\theta}} \cdot \frac{u^{k-1}}{(k-1)!} du \quad (4.2)$$

可用归纳法证明. 当 $k=1$ 时, 显然成立. 设 $k=l$ 时(4.2)成立, 则

$\sum_{i=1}^{l+1} \xi_i$ 的密度函数

$$p(x) = \int_0^x \frac{1}{\theta^l} e^{-\frac{u}{\theta}} \frac{u^{l-1}}{(l-1)!} \cdot \frac{1}{\theta} e^{-\frac{1}{\theta}(x-u)} du$$

① 若试验是无替换的, 则可以证明下列公式:

$$P_\theta(r=k) = \frac{1}{k!} \left(\frac{T}{\theta} \right)^k e^{-\frac{T}{\theta}}, k=0,1,2,\dots,n-1 \quad (\text{注 4.1})$$

$$P_\theta(r=n) = 1 - \sum_{k=0}^{n-1} \frac{1}{k!} \left(\frac{T}{\theta} \right)^k e^{-\frac{T}{\theta}} \quad (\text{注 4.2})$$

实际上, 设 n 个产品的寿命分别为 X_1, X_2, \dots, X_n , 从小到大排列为 $X_{(1)} \leq X_{(2)} \leq \dots \leq$

X 若 n 个产品同时投入试验, 则在时刻 $X_{(1)}$, 累计的试验时间为 $S_1 = \sum_{k=1}^1 X_{(k)} + (n-1)X_{(1)}$. 于是对 $k \leq n-1$ 有 $P_\theta(r=k) = P_\theta(S_k \leq T < S_{k+1})$. 令 $u_k = X_{(k)} - X_{(k-1)}$ ($X_{(0)} = 0$), 可以证明 u_1, u_2, \dots, u_n 相互独立, 且 $V_k = (n-k+1)u_k$ 服从均值是 θ 的指数分布

$$P(r=k) = P_\theta\left(\sum_{j=1}^k V_j \leq T < \sum_{i=1}^{k+1} V_i\right)$$

利用下面的(4.5)即知(4.2)成立, (4.3)从(4.2)直接得到.

$$= \frac{1}{\theta^{l+1}} e^{-\frac{t}{\theta}} \cdot \frac{1}{l!} x^l$$

可见(4.2)对 $k = l + 1$ 也成立. 所以对一切 $k \geq 1$, (4.2) 总成立.

从(4.2)推知

$$\begin{aligned} P\left(\sum_{i=1}^k \xi_i \leq t < \sum_{i=1}^{k+1} \xi_i\right) \\ = \frac{1}{k!} \left(\frac{t}{\theta}\right)^k e^{-\frac{t}{\theta}} \end{aligned} \quad (4.3)$$

设有 n 个产品同时投入试验, 若 $\{\xi_{ij} : i = 1, 2, \dots, n; j = 1, 2, \dots\}$ 是一族相互独立同分布的随机变量, 共同分布是均值为 θ 的指数分布. 因为试验为有替换的, 总试验时间为 T , 故试验持续时间为 $t_0 = \frac{T}{n}$. 显然

$$\begin{aligned} P(r=k) &= \sum_{r_1+r_2+\dots+r_n=k} \prod_{i=1}^n P_\theta\left(\sum_{j=1}^{r_i} \xi_{ij} \leq t_0 < \sum_{j=1}^{r_i+1} \xi_{ij}\right) \\ &= \sum_{r_1+r_2+\dots+r_n=k} \prod_{i=1}^n \left[\frac{1}{r_i!} \left(\frac{t_0}{\theta}\right)^{r_i} e^{-\frac{t_0}{\theta}} \right] \\ &= \frac{1}{k!} e^{-\frac{T}{\theta}} \sum_{r_1+r_2+\dots+r_n=k} \frac{k!}{r_1! r_2! \dots r_n!} \left(\frac{t_0}{\theta}\right)^{r_1} \left(\frac{t_0}{\theta}\right)^{r_2} \dots \left(\frac{t_0}{\theta}\right)^{r_n} \\ &= \frac{1}{k!} e^{-\frac{T}{\theta}} \left(\sum_{i=1}^n \frac{t_0}{\theta}\right)^k = \frac{1}{k!} e^{-\frac{T}{\theta}} \cdot \left(\frac{T}{\theta}\right)^k \end{aligned}$$

这就证明了(4.1). 证毕.

从(4.1)看出, 失效数 r 服从泊松分布. 从(注4.1)和(注4.2)知, 对无替换的试验, 失效数近似服从泊松分布.

(二) 先验分布的确定.

这是采用贝叶斯方法最关键的一步. 从定理3.2知下列逆 Γ 分布^①:

^① 称 η 服从逆 Γ 分布, 若 η^{-1} 服从 Γ 分布.

$$\xi(\theta) = \frac{b^a}{\Gamma(a)} \theta^{-a-1} e^{-\frac{b}{\theta}} \quad (\theta > 0) \quad (4.4)$$

是 Poisson 分布(4.1)的共轭分布,其中 a, b 是正数.

实际上,若取(4.4)为 θ 的先验分布密度,则在失效数 $r = k$ 的条件下, θ 的后验分布密度为

$$\xi(\theta | k) = \frac{(b + T)^{a+k}}{\Gamma(a+k)} \theta^{-(a+k+1)} e^{-\frac{b+T}{\theta}} \quad (4.5)$$

故取(4.4)为 θ 的先验分布.怎样确定 a, b 呢? a, b 的直观意义不明显.但先验分布的均值 $E(\bar{\theta})$ 及 10% 分位数 θ_L (即有 $P(\bar{\theta} \leq \theta_L) = 0.10$) 有明确的直观意义.有下列关系式:

$$E(\bar{\theta}) = \frac{b}{a-1} \quad (4.6)$$

$$\int_0^{\theta_L} \frac{b^a}{\Gamma(a)} \theta^{-(a+1)} e^{-\frac{b}{\theta}} d\theta = 0.1 \quad (4.7)$$

若能根据历史数据近似估计出 $E(\bar{\theta})$ 及 θ_L , 则从(4.6)和(4.7)就可求出 a, b .

为了利用历史数据,对全国 15 个企业生产的 22 种型号彩电的寿命数据进行了调查(包括实验室数据和现场使用数据),经过多次研究,将产品按寿命分成三级,分别记为 P, Q, R ;然后根据管理经验和历史数据确定每级产品的 θ_1, θ_0 的取值,鉴别比 $d = 3$. P, Q, R 三级的 θ_1 分别定为 15 000, 20 000 和 30 000 h, 根据历史数据,取 $E\bar{\theta} = 3\theta_1$. P, Q, R 三级的 θ_L 分别为 $0.75\theta_1, 0.90\theta_1, 0.95\theta_1$. 以 P 级为例,取 $E(\bar{\theta}) = 3\theta_1 = 45 000, \theta_L = 0.75\theta_1 = 11 250$, 再利用(4.8)、(4.9)经计算机算得先验分布中的两个参数如下:

$$a = 1.956, b = 2.868$$

对 Q 级, R 级可进行同样的计算, 总之得各级所需的参数如下表所示:

参 数 级 别	MTBF(平均寿命)/h		先验分布参数	
	θ_1	θ_0	a	b
P	15 000	45 000	1.956	2.868
Q	20 000	60 000	2.298	3.895
R	30 000	90 000	2.431	4.294

(三) 验证方案的制定.

验证方案由(T, C)两个参数构成。 T 是总试验时间(台时数), C 是合格判定数.当试验停止时累计失效数 $r \leq C$,则判定批合格;当 $r > C$ 时,则判定批不合格。 T, C 如何确定? 应使后验生产方风险 α^* 及后验使用方风险 β^* 都达到指定的水平,这里

$$\alpha^* = P(\bar{\theta} \geq \theta_0 \mid r > C)$$

$$\beta^* = P(\bar{\theta} \leq \theta_1 \mid r \leq C)$$

可以证明有下列计算公式:

定理 4.2

$$\alpha^* = \frac{\Gamma(a, \frac{b}{\theta_0}) - \left(\frac{b}{b+T}\right)^a \sum_{k=0}^C \frac{1}{k!} \left(\frac{T}{b+T}\right)^k \Gamma(a+k, \frac{b+T}{\theta_0})}{\Gamma(a)[1 - P(r \leq C)]} \quad (4.8)$$

$$\beta^* = \frac{\left(\frac{b}{b+T}\right)^a \sum_{k=0}^C \frac{1}{k!} \left(\frac{T}{b+T}\right)^k [\Gamma(a+k) - \Gamma(a+k, \frac{b+T}{\theta_1})]}{P(r \leq C)} \quad (4.9)$$

其中 $\Gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$ 是不完全 Γ 函数,而

$$\begin{aligned} P(r \leq C) &= \int_0^{+\infty} P_\theta(r \leq C) \xi(\theta) d\theta \\ &= \left(\frac{b}{b+T}\right)^a \sum_{k=0}^C \frac{\Gamma(a+k)}{\Gamma(a)k!} \left(\frac{T}{b+T}\right)^k \end{aligned}$$

是接收概率,这是一个无条件概率,表示 θ 的先验分布为(4.4)时产品被接收的概率.

证

$$\begin{aligned}\alpha^* &= \frac{P(\tilde{\theta} \geq \theta_0, r > C)}{P(r > C)} \\ &= \frac{P(\tilde{\theta} \geq \theta_0) - P(\tilde{\theta} \geq \theta_0, r \leq C)}{P(r > C)}\end{aligned}$$

由于

$$\begin{aligned}P(\tilde{\theta} \geq \theta_0) &= \int_{\theta_0}^{+\infty} \xi(\theta) d\theta = \int_{\theta_0}^{+\infty} \frac{b^a}{\Gamma(a)} \theta^{-(a+1)} e^{-\frac{b}{\theta}} d\theta \\ &= \int_0^{\frac{b}{\theta_0}} \frac{1}{\Gamma(a)} t^{a-1} e^{-t} dt \\ &= \frac{1}{\Gamma(a)} \Gamma\left(a, \frac{b}{\theta_0}\right)\end{aligned}$$

另一方面

$$\begin{aligned}P(\tilde{\theta} \geq \theta_0, r \leq C) &= \int_{\theta_0}^{+\infty} P_r(r \leq C) \xi(\theta) d\theta \\ &= \int_{\theta_0}^{+\infty} \sum_{k=0}^C \frac{1}{k!} \left(\frac{T}{\theta}\right)^k e^{-\frac{T}{\theta}} \cdot \frac{b^a}{\Gamma(a)} \theta^{-(a+1)} e^{-\frac{b}{\theta}} d\theta \\ &= \frac{1}{\Gamma(a)} \sum_{k=0}^C \frac{1}{k!} \frac{T^k b^a}{(T+b)^{a+k}} \int_0^{\frac{T+b}{\theta_0}} t^{a+k-1} e^{-t} dt \\ &= \frac{1}{\Gamma(a)} \left(\frac{b}{T+b}\right)^a \sum_{k=0}^C \frac{1}{k!} \left(\frac{T}{T+b}\right)^k \Gamma\left(a+k, \frac{T+b}{\theta_0}\right)\end{aligned}$$

由此知(4.8)成立. 同理知(4.9)成立. 证毕.

对于 P 级产品, $\theta_1 = 15000$, $a = 1.956$, $b = 2.868$, 取 $\beta^* = 0.10$, 由公式(4.9)可算出几组 (T, C) 的值, 再由公式(4.8)计算相应的后验生产方风险 α^* , 得下表:

P 级产品的贝叶斯方案

C(允许失效数)	1	2	3
T(总试验时间)	25 389	36 421	47 748
α^* (后验生产方风险)	0.044	0.019	0.008

同样可以制定 Q 级和 R 级(优质品或名牌产品)的验证试验方案.

经过反复讨论和对历史数据的分析,兼顾使用方和生产方的利益,最后在验收标准中对 P 级产品选用了 $C=2$ 的方案,这时试验总台时数 $T = 36 421 \approx 37 000$,是原来的经典验收方案台时数 $T = 100 000$ 的 37%,这就大大减少了试验工作量. 对 Q 级、R 级产品也选定了验收方案,总结在下表中:

验收试验方案表

级 别 $\backslash (T, C)$ 方案	T(台时)	C(合格判定数)
P	37 000	2
Q	32 000	2
R	20 000	1

以上通过实例介绍了如何运用贝叶斯方法,读者可联系 §2 中的内容加深对贝叶斯方法的了解.

在上面的实例中,只是利用先验分布和后验分布去确定电视机的验证试验方案,使得后验生产方风险 α^* 及后验使用方风险 β^* 均达到指定的水平 ($\beta^* = 0.10, \alpha^* \leq 0.05$),并没有去追求“贝叶斯决策”——“风险”最小的验收方案. 要获得贝叶斯决策,只知道先验分布还不够,还必须知道损失函数是怎样的. 在制定验收方案时,选取生产方和使用方都同意的损失函数并不是很简单的. 由于较难确定“客观”的损失函数,人们常常不去追求贝叶斯决策了. 当然,一旦选好损失函数,则应该追求贝叶斯决策. 以上面的电视机验证

试验为例,如果损失函数是0-1损失,则不难求出相应的贝叶斯决策.具体介绍如下:

沿用正文中的各种记号.检验问题是:

$$H_0: \theta \geq \theta_0 \leftrightarrow H_1: \theta \leq \theta_1 \quad (\theta_0 > \theta_1)$$

设行动空间 $A = \{a_1, a_2\}$, a_1 表示“接收”(即接受假设 H_0), a_2 表示“拒收”(即拒绝假设 H_0). 损失函数 $L(\theta, a)$ 如下:

$$L(\theta, a_1) = \begin{cases} 0 & \theta > \theta_1 \\ 1 & \theta \leq \theta_1 \end{cases}$$

$$L(\theta, a_2) = \begin{cases} 1 & \theta \geq \theta_0 \\ 0 & \theta < \theta_0 \end{cases}$$

仍取(4.4)为 θ 的先验分布密度,其中参数 a, b 的确定方法见正文.

设 T 是试验的总台时数(有替换的试验), γ 是试验截止时的累计失效数.从(4.5)推知

$$P(\bar{\theta} \geq \theta_0 | \gamma = k) = \frac{1}{\Gamma(a+k)} \int_0^{\frac{T+b}{\theta_0}} t^{a+k-1} e^{-t} dt$$

$$P(\bar{\theta} \leq \theta_1 | \gamma = k) = 1 - \frac{1}{\Gamma(a+k)} \int_0^{\frac{T+b}{\theta_1}} t^{a+k-1} e^{-t} dt$$

$$(k = 0, 1, 2, \dots)$$

利用微分法可以证明 $P(\bar{\theta} \geq \theta_0 | \gamma = k)$ 是 k 的减函数,从而 $P(\bar{\theta} \leq \theta_1 | \gamma = k)$ 是 k 的增函数.故对任何非负整数 C 有

$$\inf_{0 \leq k \leq C} P(\bar{\theta} \geq \theta_0 | \gamma = k) = P(\bar{\theta} \geq \theta_0 | \gamma = C) \quad (4.10)$$

$$\sup_{k > C} P(\bar{\theta} \geq \theta_0 | \gamma = k) = P(\bar{\theta} \geq \theta_0 | \gamma = C + 1) \quad (4.11)$$

$$\sup_{0 \leq k \leq C} P(\bar{\theta} \leq \theta_1 | \gamma = k) = P(\bar{\theta} \leq \theta_1 | \gamma = C) \quad (4.12)$$

$$\inf_{k > C} P(\bar{\theta} \leq \theta_1 | \gamma = k) = P(\bar{\theta} \leq \theta_1 | \gamma = C + 1) \quad (4.13)$$

我们指出,对任何非负整数 C ,必有 $T = T(C)$ 使得

$$P(\bar{\theta} \geq \theta_0 | \gamma = C) = P(\bar{\theta} \leq \theta_1 | \gamma = C) \quad (4.14)$$

实际上,令

$$\varphi(T) \stackrel{\Delta}{=} P(\bar{\theta} \geq \theta_0 | \gamma = C) - P(\bar{\theta} \leq \theta_1 | \gamma = C)$$

$$= \frac{1}{\Gamma(a+C)} \int_0^{\frac{T+b}{\theta_0}} t^{a+C-1} e^{-t} dt$$

$$+ \frac{1}{\Gamma(a+C)} \int_0^{\frac{T+b}{\theta_1}} t^{a+C-1} e^{-t} dt - 1$$

显然 $\varphi(T)$ 是 T 的严格增连续函数, 且 $\varphi(0+) = -\infty$ (当 b 足够小), $\varphi(+\infty) = 1$, 故有 $T = T(C)$ 使得(4.14)成立.

给定 C , 设试验的总台时数为 $T = T(C)$, γ 是累计失效数. 令

$$\delta^*(\gamma) = \begin{cases} a_1 & \text{当 } \gamma \leq C \\ a_2 & \text{当 } \gamma > C \end{cases}$$

我们来证明, δ^* 是贝叶斯决策.

设 $\xi(\theta|k)$ 是失效数为 k 时 θ 的后验分布密度, 具体表达式见(4.5). 显然有

$$P(\bar{\theta} \geq \theta_0 | \gamma = k) = \int_{\theta_0}^{\infty} \xi(\theta|k) d\theta$$

$$P(\bar{\theta} \leq \theta_1 | \gamma = k) = \int_0^{\theta_1} \xi(\theta|k) d\theta$$

从(4.10)~(4.14)知

$$\inf_{a \in A} \int_0^{\infty} L(\theta, a) \xi(\theta|k) d\theta$$

$$= \min\{P(\bar{\theta} \leq \theta_1 | \gamma = k), P(\bar{\theta} \geq \theta_0 | \gamma = k)\}$$

$$= \begin{cases} P(\bar{\theta} \leq \theta_1 | \gamma = k) & \text{当 } k \leq C \\ P(\bar{\theta} \geq \theta_0 | \gamma = k) & \text{当 } k > C \end{cases}$$

另一方面, 当 $\gamma = k$ 时 δ^* 的后验平均损失

$$\int_0^{\infty} L(\theta, \delta^*(k)) \xi(\theta|k) d\theta = \begin{cases} P(\bar{\theta} \leq \theta_1 | \gamma = k) & k \leq C \\ P(\bar{\theta} \geq \theta_0 | \gamma = k) & k > C \end{cases}$$

$$= \inf_{a \in A} \int_0^{\infty} L(\theta, a) \xi(\theta|k) d\theta$$

这表明 δ^* 满足(2.2), 故 δ^* 是贝叶斯决策.

在实际工作中要求试验的总台时数尽可能地小, 因而 C 应该取得很小, 通常取 $C = 0, 1, 2$, 相应的总台时数 $T = T(C)$ 由(4.14)确定.

习题二十

1. 设 X 是离散或连续型随机变量, $E(X^2)$ 存在. 试证明: 为了使 $E(X-a)^2$ (a 是实数) 达到最小值, 必须且只须 $a = E(X)$.

2. 设 X 是离散型或连续型随机变量, $E(X)$ 存在. 试证明: 为了使 $E|X - a|$ (a 是实数) 达到最小值, 必须且只须 a 是 X 的中位数.

3. 设 X_1, X_2, \dots, X_n 是来自 Poisson 分布的样本,

$$P(X_i = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (\lambda > 0, k = 0, 1, \dots)$$

设 λ 的先验分布是参数为 α, β 的 Γ 分布, 损失函数是 $(\lambda - a)^2$, 试求 λ 的贝叶斯估计.

4. 设 X_1, X_2, \dots, X_n 是来自伯努利分布的样本, ($n \geq 3$),

$$P(X_i = 1) = \theta = 1 - P(X_i = 0), \theta \in (0, 1)$$

设 θ 的先验分布是 $(0, 1)$ 上的均匀分布, 损失函数是

$$L(\theta, a) = \left[\frac{\theta - a}{\theta(1 - \theta)} \right]^2$$

试求 θ 的贝叶斯估计.

5. 设 X_1, X_2, \dots, X_n 是来自正态分布 $N(\theta, 1)$ 的样本, θ 是未知参数. 给定检验问题:

$$H_1: \theta \leq \theta_0 \leftrightarrow H_2: \theta > \theta_0$$

用 a_i 表示“接受假设 H_i ”($i = 1, 2$). 设损失函数如下:

$$L(\theta, a_1) = \begin{cases} 0 & \theta \leq \theta_0 \\ k(\theta - \theta_0) & \theta > \theta_0 \end{cases}$$

$$L(\theta, a_2) = \begin{cases} k|\theta - \theta_0| & \theta \leq \theta_0 \\ 0 & \theta > \theta_0 \end{cases}$$

其中 k 是正常数. 若 θ 的先验分布是 $N(\mu_0, \sigma_0^2)$, 试求出贝叶斯检验.

6. 地质学家要根据某地区的地层结构来判断该地是否蕴藏石油. 地层结构总是 0, 1 两种状态之一; 用 θ_0 表示该地无油, θ_1 表示该地有油. 已知有下列概率分布规律(其中 x 表示地层结构的状态, θ 表示石油的状态):

θ	x	0	1
θ_0 (无油)		0.6	0.4
θ_1 (有油)		0.3	0.7

它表示如果该地区无油,那么地层结构呈现状态 0 的概率为 0.6,呈现状态 1 的概率为 0.4;如果该地蕴藏石油,那么地层结构呈现状态 0 的概率是 0.3,呈现状态 1 的概率为 0.7. 土地所有者希望根据地质学家对地层结构的分析决定自己投资钻探石油,还是出卖土地所有权或者在该地区开辟旅游点,分别记这些行动为 a_1, a_2, a_3 . 行动空间 $A = \{a_1, a_2, a_3\}$, 土地所有者权衡利弊之后取损失函数 $L(\theta, a)$ 为

$L(\theta, a)$	a	a_1 (自己投资钻探)	a_2 (出卖所有权)	a_3 (开辟旅游点)
θ				
θ_0 (无油)		12	1	6
θ_1 (有油)		0	7	5

试写出可供土地所有者选择的全部决策(函数)及其风险. 求出 minimax 决策. 若 θ 的先验分布是: $\xi(\theta_0) = 0.2, \xi(\theta_1) = 0.8$, 试求出贝叶斯决策.

第十章 随机过程初步

§ 1 随机过程的概念

到现在为止,我们的研究对象主要是一个或几个随机变量(随机向量).但是在自然现象、社会现象及实际工作中,我们还会碰到无穷多个随机变量在一起需要当作一个整体来对待的情形,这就需要引进随机过程的概念.

定义 1.1 给定参数集 $T \subset (-\infty, +\infty)$, 如果对每个 $t \in T$, 对应一个随机变量 X_t , 则称随机变量族 $\{X_t, t \in T\}$ 为随机过程^① (简称“过程”).

例 1.1 用 X_t 表示某电话机从时刻 0 开始到时刻 t 为止所接到的呼唤次数, 则 $\{X_t, t \in [0, +\infty)\}$ 便是一随机过程.

例 1.2 对晶体管热噪声电压进行测量, 每隔一微秒测一次. 测量时刻记作 1, 2, …, 在时刻 t 的测量值记作 X_t , 则 $\{X_t, t = 1, 2, \dots\}$ 便是一个随机过程.

例 1.3 1826 年布朗(Brown)发现水中花粉(或其他液体中的微粒)在不停地运动, 这种现象后来称为布朗运动. 由于花粉受到水中分子的碰撞, 每秒钟所受碰撞次数多到 10^{21} 次, 这些随机的微小的碰撞力的总和使得花粉作随机运动, 以 X_t 表示花粉在 t 时刻所在位置的一个坐标(例如横坐标), 则 $\{X_t, t \in [0, \infty)\}$ 便是一

^① 随机过程的定义可以更广泛些, 例如 X_t 可以是多维随机向量. 参数集 T 也可用任一非空集 Λ 代替. 这时 $\{X_t, t \in \Lambda\}$ 常叫做随机函数. 本章不涉及这些复杂的情形.

个随机过程.

例 1.4 考察纺织机所纺出的一根棉纱, 以 X_t 表示 t 时刻纺出的纱的横截面的直径, 由于工作条件随 t 变化而不能恒定, $\{X_t, t \in [0, \infty)\}$ 便是一个随机过程.

随机过程的例子太多了, 只要考察随机现象如何随着时间而变, 就会遇到随机过程.

我们常用 E 表示诸 X_t 所可能取的值所组成的集合, E 叫做状态空间(或相空间). 如果 $X_t = x$, 则说过程 $\{X_t, t \in T\}$ 在时刻 t 处于状态 x .

当 T 是一个有限集或可列集(即可排成一个序列)时, $\{X_t, t \in T\}$ 叫做离散时间的随机过程(随机序列). 最常见的情况是 $T = \{0, 1, 2, \dots\}$ 或 $T = \{\dots, -1, 0, 1, \dots\}$.

当 T 是一个区间(可以是无穷区间)时, $\{X_t, t \in T\}$ 叫做连续时间的随机过程. 这时最常见的情况是 $T = [0, +\infty)$ 或 $T = (-\infty, +\infty)$.

给定 T 中 n 个数 t_1, t_2, \dots, t_n , 记 $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ 的分布函数为 $F_{t_1 t_2 \dots t_n}(x_1, x_2, \dots, x_n)$. 这种分布函数的全体 $\{F_{t_1 t_2 \dots t_n}(x_1, x_2, \dots, x_n) : n \geq 1, t_1, t_2, \dots, t_n \in T\}$ 叫做 $\{X_t, t \in T\}$ 的有限维分布函数族. 这个族描写了随机过程的概率特性.

随机过程 $\{X_t, t \in T\}$ 也可从另外一个角度进行考察. 每个随机变量 X_t 乃是某条件组 S 下可能结果 ω 的函数. 条件组 S 下所有可能的结果组成的集合记作 Ω . X_t 就是 Ω 上的函数 $X_t(\omega)$ (参看第二章的开始部分). 所以固定 ω 后, $X_t(\omega)$ 便是 t 的函数. 这个函数叫做随机过程的一个“实现”, 或叫“现实”, 或叫“轨道”, 或叫“样本函数”. 我们对一个随机过程进行观察, 所得的记录就是这个过程的“实现”.

例如在例 1.2 中对晶体管热噪声电压进行测量. 在时刻 1, 2, … 测得的具体数据:

x_1, x_2, \dots

就是 $\{X_t, t = 1, 2, \dots\}$ 的一个“实现”.

在连续时间的情形, 过程的“实现”常用曲线表示.

如何根据一个“实现”去推断随机过程的性质, 是随机过程论的一个重要问题, 属于过程统计的范围.

怎样去研究随机过程呢? 通常是按照随机过程的概率特性划分成几个大类进行研究. 每类过程都有专门的名称, 最重要的有三类:

- ① 独立增量过程;
- ② 马尔可夫(Марков)过程;
- ③ 平稳过程.

以下三节分别介绍这三种过程的最基本知识. 特别强调一下, 这三类过程并不互相排斥. 实际上, 独立增量过程是特殊的马尔可夫过程, 而有些马尔可夫过程又是平稳过程. 除这些外, 理论上和实践上都很重要的还有“鞅”. 现代随机过程论内容丰富多彩, 应用相当广泛, 读者如有兴趣, 请参阅有关的专著.

§2 独立增量过程

1. 定义

定义 2.1 称 $\{X_t, t \in T\}$ 为独立增量过程, 如果对任何 $t_1 < t_2 < \dots < t_n$ ($t_i \in T, i = 1, \dots, n$), 随机变量

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}}$$

是相互独立的.

如果此时 $X_{t+\tau} - X_t$ ($\tau > 0$) 的分布函数只依赖于 τ 而不依赖于 t , 则称 $\{X_t, t \in T\}$ 为时齐的独立增量过程.

从定义看出, 若 $\{X_t, t \in T\}$ 是独立增量过程, Y 是随机变量, 则 $\{X_t + Y, t \in T\}$ 也是独立增量过程. 故在研究工作中常设

$X_0 \equiv 0$ (当 $0 \in T$ 时).

例 2.1 设 X_1, X_2, \dots 是相互独立的随机变量列. $S_n = X_1 + X_2 + \dots + X_n$ ($n \geq 1$). 则 $\{S_n, n \geq 1\}$ 便是独立增量过程(详细证明从略). 若所有的 X_i 服从相同的分布, 则这过程是时齐的. 第四章中的大数定律与中心极限定理就是讨论这个特殊过程的性质.

限于篇幅, 我们不去讨论一般的独立增量过程, 只介绍两个最基本最典型的例子: 泊松过程和维纳(Wiener)过程.

2. 泊松过程

定义 2.2 称 $\{X_t, t \geq 0\}$ 是泊松过程, 若它是独立增量的, 而且 X_t 取值是非负整数, 增量 $X_t - X_s$ ($0 \leq s < t$) 服从泊松分布:

$$P\{X_t - X_s = k\} = e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^k}{k!} \quad (k = 0, 1, \dots)$$

其中 λ 是与 t, s 无关的正常数.

什么情况下会出现泊松过程呢? 有下列定理.

定理 2.1 设 $\{X_t, t \geq 0\}$ 是取非负整数值的时齐的独立增量过程, 满足

$$P\{X_0 = 0\} = 1$$

$$P\{X_{t+\Delta t} - X_t = 1\} = \lambda \Delta t + o(\Delta t) \quad (\Delta t \rightarrow 0+)$$

$$P\{X_{t+\Delta t} - X_t \geq 2\} = o(\Delta t)$$

(这里 $\lambda > 0$). 则 $\{X_t, t \geq 0\}$ 就是泊松过程.

证 设 $\Delta t > 0$, 从所给条件知 $X_{t+\Delta t} - X_t$ 与 $X_{\Delta t}$ 的概率分布相同, 从而 $P\{X_{t+\Delta t} - X_t \geq 0\} = 1$, 且

$$P\{X_{t+\Delta t} - X_t = 0\} = 1 - \lambda \Delta t + o(\Delta t) \quad (\Delta t \rightarrow 0)$$

令 $W_i(t) = P\{X_t = i\}$ ($t \geq 0$), 则

$$\begin{aligned} W_i(t + \Delta t) \\ = P\{X_t = i, X_{t+\Delta t} = i\} + P\{X_t = i-1, X_{t+\Delta t} = i\} \\ + P\{X_t \leq i-2, X_{t+\Delta t} = i\} \end{aligned}$$

$$\begin{aligned}
&= P\{X_t = i\} \cdot P\{X_{t+\Delta t} - X_t = 0\} + P\{X_t = i-1\} \\
&\quad \cdot P\{X_{t+\Delta t} - X_t = 1\} + P\{X_{t+\Delta t} = i, X_{t+\Delta t} - X_t \geq 2\} \\
&= W_i(t)[1 - \lambda \Delta t + o(\Delta t)] + W_{i-1}(t)[\lambda \Delta t + o(\Delta t)] \\
&\quad + o(\Delta t)
\end{aligned}$$

故 $W_i(t + \Delta t) - W_i(t)$

$$\begin{aligned}
&= -\lambda W_i(t) \Delta t + \lambda W_{i-1}(t) \Delta t + W_i(t) \cdot o(\Delta t) \\
&\quad + W_{i-1}(t) \cdot o(\Delta t) + o(\Delta t)
\end{aligned}$$

于是 $\lim_{\Delta t \rightarrow 0^+} \frac{W_i(t + \Delta t) - W_i(t)}{\Delta t} = -\lambda W_i(t)$

$$\begin{aligned}
&\quad + \lambda W_{i-1}(t)
\end{aligned}$$

即有 $W'_i(t) = -\lambda W_i(t) + \lambda W_{i-1}(t) \quad (i \geq 0) \quad (2.1)$

(这里 $W_{-1}(t)$ 约定为 0.)

易知 $W_0(0) = P\{X_0 = 0\} = 1$, 故 $W_0(t) = e^{-\lambda t}$.

令 $W_i(t) = u_i(t) e^{-\lambda t} \quad (i \geq 0)$. 则 $W'_i(t) = u'_i(t) e^{-\lambda t} - \lambda W_i(t)$, 故 $W'_i(t) + \lambda W_i(t) = u'_i(t) e^{-\lambda t}$, 但从(2.1)知

$$W'_i(t) + \lambda W_i(t) = \lambda W_{i-1}(t) = \lambda u_{i-1}(t) e^{-\lambda t}$$

于是得方程组

$$\begin{aligned}
u'_i(t) &= \lambda u_{i-1}(t) \quad (i \geq 1) \\
u_0(t) &\equiv 1
\end{aligned}$$

注意 $u_i(0) = W_i(0) = 0 \quad (i \geq 1)$, 故不难推得 $u_i(t) = \frac{1}{i!} (\lambda t)^i$,

所以 $W_i(t) = \frac{(\lambda t)^i}{i!} e^{-\lambda t} \quad (i \geq 0)$. 于是 $P\{X_t - X_s = i\} = P\{X_{t-s} = i\} = W_i(t-s) = \frac{[\lambda(t-s)]^i}{i!} e^{-\lambda(t-s)} \quad (s < t)$

这就证明了 $\{X_t, t \geq 0\}$ 是泊松过程.

3. 维纳过程

定义 2.3 称独立增量过程 $\{X_t, t \geq 0\}$ 是维纳过程, 如果对任何 $s < t$,

$$X_t - X_s \sim N(0, (t-s)\sigma^2)$$

这里 σ 是固定的正数(与 s, t 无关).

通过物理学的研究知道, 维纳过程是描述布朗运动的概率模型. X_t 表示液体中运动的微粒在时刻 t 的位置的横坐标.

经过数学研究知道, 只要在数学上加点条件, 维纳过程的几乎所有的轨道(或“实现”)是 t 的连续函数, 但这些连续函数几乎处处没有导数(这里“几乎”二字是有精确的数学含义的, 我们不具体说了). 对维纳过程以及更一般的扩散过程, 现代有大量研究.

§ 3 马尔可夫过程

1. 定义

设 $\{X_t, t \in T\}$ 是一个随机过程, 状态空间是 E , 我们可以把这个随机过程看成某系统的“状态”的演变过程.“ $X_t = x$ ”表示该系统在时刻 t 处于状态 x .

定义 3.1 称 $\{X_t, t \in T\}$ 是马尔可夫过程, 如果对于 T 中任何 n 个数 $t_1 < t_2 < \dots < t_n$, E 中任何 n 个状态 x_1, x_2, \dots, x_n 及任何实数 x 均成立:

$$\begin{aligned} & P\{X_{t_n} \leq x | X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_{n-1}} = x_{n-1}\}^{\textcircled{1}} \\ &= P\{X_{t_n} \leq x | X_{t_{n-1}} = x_{n-1}\} \end{aligned} \quad (3.1)$$

换句话说, 马尔可夫过程的特征是, 如已知“现在: $X_{t_{n-1}} = x_{n-1}$ ”, 则“将来: $X_{t_n} \leq x$ ”不依赖于“过去: $X_{t_1} = x_1, X_{t_2} = x_2, \dots,$

① 这是条件概率. 当 $P\{X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_{n-1}} = x_{n-1}\} \neq 0$ 时, 这种条件概率在第一章已讨论过. 当 $P\{X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_{n-1}} = x_{n-1}\} = 0$ 时, 以前未处理过, 这要用测度论的知识才能进行比较严密的讨论. 这里只要求读者对条件概率有一个大致的、朴素的理解.

$X_{t_{n-2}} = x_{n-2}$ ”。这表达了过程的“无后效性”。

(3.1)式所表达的性质称为马尔可夫性,简称马氏性。马尔可夫过程简称马氏过程。马尔可夫是俄罗斯数学家,他在本世纪初研究过现在称之为马尔可夫过程的一种特殊情形。

马尔可夫过程论的内容十分丰富,但很多讨论都涉及到较深的数学知识。以下只就最简单的情况,即 $T = \{0, 1, 2, \dots\}$, E 为至多可列集的情形,进行初步的讨论。(这里所谓 E 至多可列,是指 E 是有限集或者 E 的全体元素可排成一无穷序列。)

2. 马尔可夫链

定义 3.2 设 $\{X_n, n \geq 0\}$ 是随机序列,状态空间 E 至多可列,若对任何 $i_0, i_1, \dots, i_n \in E$,只要 $P\{X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}\} \neq 0$,就成立:

$$\begin{aligned} & P\{X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}\} \\ &= P\{X_n = i_n | X_{n-1} = i_{n-1}\} \end{aligned} \quad (3.2)$$

则称 $\{X_n, n \geq 0\}$ 为马尔可夫链,简称马氏链。可以验证,马尔可夫链是一种特殊的马氏过程。

通常称条件概率 $P\{X_t = j | X_s = i\}$ 为转移概率,记作 $p_{ij}(s, t)$ ^①(这里 $s \leq t$)。

定理 3.1 设 $\{X_n, n \geq 0\}$ 是马氏链,则对 $s < t < u$,有

$$p_{ij}(s, u) = \sum_{k \in E} p_{ik}(s, t) p_{kj}(t, u) \quad (3.3)$$

证 $p_{ij}(s, u) = P\{X_u = j | X_s = i\}$

$$\begin{aligned} &= \sum_{k \in E} P\{X_t = k, X_u = j | X_s = i\} \\ &= \sum_{k \in E} P\{X_t = k | X_s = i\} P\{X_u = j | X_s = i, X_t = k\} \\ &= \sum_{k \in E} P\{X_t = k | X_s = i\} P\{X_u = j | X_t = k\} \end{aligned}$$

① 为简单计,以下假设 $p_{ij}(s, t)$ 对所有 $i, j \in E$ 都有定义。

$$= \sum_{k \in E} p_{ik}(s, t) p_{kj}(t, u)$$

(3.3) 叫做 Chapman-柯尔莫哥洛夫 方程.

若任意固定 i, j 后, $p_{ij}(s, t) = p_{ij}(s + \tau, t + \tau)$ (对一切 $\tau \geq 0$), 则称马氏链 $\{X_n, n \geq 0\}$ 是时齐的, 也叫齐次的. 本节往下只讨论齐次马氏链, 简称马氏链. 此时记 $p_{ij} = P\{X_{t+1} = j | X_t = i\}$, 称矩阵 $P = (p_{ij}, i, j \in E)$ 为一步转移概率矩阵. P 有下列性质.

定理 3.2

$$p_{ij} \geq 0$$

$$\sum_{j \in E} p_{ij} = 1 (i \in E)$$

证 显然.

我们说, 只要一步转移概率矩阵 $P = (p_{ij})$ 知道了, 则马氏链的转移概率特性就完全确定了. 实际上,

$$\begin{aligned} & P\{X_{s+n} = j | X_s = i\} \\ &= \sum_{i_1, i_2, \dots, i_{n-1} \in E} P\{X_{s+n} = j, X_{s+1} = i_1, \dots, X_{s+n-1} = i_{n-1} | X_s = i\} \\ &= \sum_{i_1, i_2, \dots, i_{n-1} \in E} P\{X_{s+1} = i_1 | X_s = i\} \cdot P\{X_{s+2} = i_2 | X_{s+1} = i_1\} \\ &\quad \cdots P\{X_{s+n} = j | X_{s+n-1} = i_{n-1}\} \\ &= \sum_{i_1, i_2, \dots, i_{n-1} \in E} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} \end{aligned}$$

以下记 $p_{ij}^{(n)} = P\{X_{s+n} = j | X_s = i\}$ (n 步转移概率). 不难看出, 矩阵 $(p_{ij}^{(n)}, i, j \in E) = P^n$.

例 3.1 (自由随机游动) 某质点在整数点集 $\{\dots, -1, 0, 1, \dots\}$ 上随机游动. 设开始时质点在位置 0, 以后每经过一个单位时间按下列概率规则改变一次位置: 如果它在某时刻位于点 i , 则它以概率 p ($0 < p < 1$) 转移到 $i + 1$, 以概率 $1 - p$ 转移到 $i - 1$. 用 X_n 表示质点在时刻 n 所在的位置, 则 $\{X_n, n \geq 0\}$ 便是一个马氏链, 其一步转移概率矩阵是