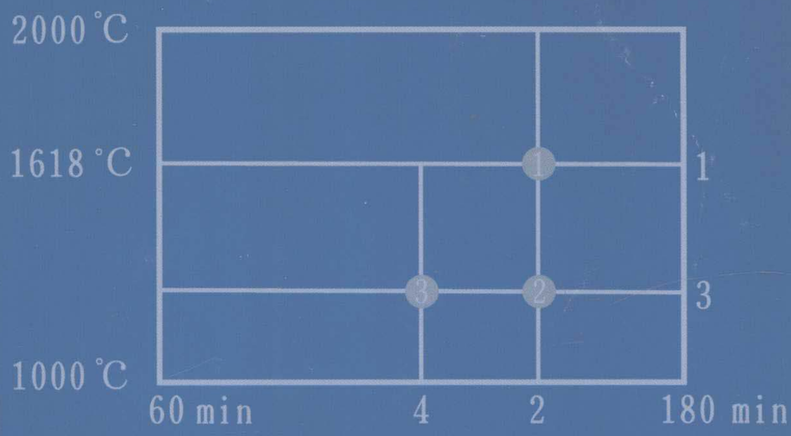
经全国中小学教材审定委员会2005年初审通过

数学

普通高中课程标准实验教科书

选修 4-7

优选法与试验设计初步



凤凰出版传媒集团
凤凰国标教材
江蘇教育出版社
JIANGSU EDUCATION PUBLISHING HOUSE

经全国中小学教材审定委员会2005年初审通过

普通高中课程标准实验教科书

数学

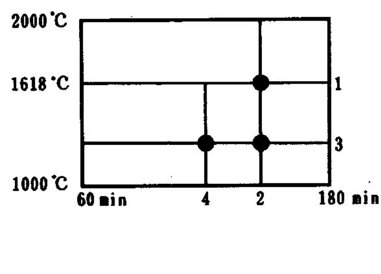
优选法与试验设计初步

youxuanfa yu shiyan sheji chubu

选修4－7

主　编：单　墫

副主编：李善良　陈永高　王巧林



凤凰国标教材

凤凰出版传媒集团

江蘇教育出版社

JIANGSU EDUCATION PUBLISHING HOUSE

书　　名　普通高中课程标准实验教科书·数学
　　　　　优选法与试验设计初步(选修4—7)
作　　者　本书编写组
责任编辑　蔡　立
出版发行　凤凰出版传媒集团
　　　　　江苏教育出版社(南京市马家街31号 210009)
网　　址　http://www.1088.com.cn
集团网址　凤凰出版传媒网 http://www.ppm.cn
经　　销　江苏省新华发行集团有限公司
照　　排　南京理工出版信息技术有限公司
印　　刷　通州市印刷总厂有限公司
厂　　址　通州市交通路55号(邮编 226300)
电　　话　0513—80237871
开　　本　1000×1436毫米　1/32
印　　张　2.5
版　　次　2005年12月第1版
　　　　　~~2006年6月第3次印刷~~
书　　号　ISBN 7-5343-6803-0/G·6488
定　　价　3.07元
盗版举报　025—83204538

苏教版图书若有印装错误可向承印厂调换
提供盗版线索者给予重奖

主　　编　单　墫

副 主 编　李善良　陈永高　王巧林

编写人员　王振羽　汪仁官

参与设计　徐稼红

责任编辑　蔡　立

一枝粉笔多长最好？每枝粉笔用到最后都要丢掉一段一定长的粉笔头，单就这一点来说，愈长愈好.但太长了，使用起来既不方便，又容易折断.每断一次，必然多浪费一个粉笔头，反而不合算.因而就出现了“粉笔多长最合适”的问题.

在某种化工产品的生产中，反应温度、反应时间及催化剂种类是影响质量的三个重要因素，每个因素又有多个等级，如果要对各因素、各等级的所有组合方案进行试验，所花的时间会很多.那么，是否可以考虑在全部方案中选取一部分进行试验，而又能达到较好的效果？

像上面这样的问题，正是本专题所要探讨的.本专题将学习单因素、双因素的优选法和多因素的正交试验设计方法方面的知识.

目　　录

7.1 优选法

7.1.1 什么是优选法

在生产、生活和科学实验中，人们为了达到优质、高产、低消耗等目的，需要对有关因素（如配比、配方、工艺操作条件等）的最佳点进行选择. 所有这些选择最佳点的问题，都称之为选优问题.

选优问题是很常见的. 例如，蒸馒头，为了使蒸出的馒头好吃，就要放碱. 碱放少了，蒸出的馒头就酸；碱放多了，馒头就发黄有碱味. 对一定量的面来说，放多少碱最合适呢？这就是一个选优的问题. 前面我们讲到的“粉笔多长最合适”问题，其实也是一个选优问题. 再如膨胀珍珠岩是一种建筑保温材料，工厂在生产这种材料时，会遇到这样的问题：焙烧温度高了，浪费能源，又损耗焙烧炉；温度低了，会影响珍珠岩的膨胀系数. 在什么样的温度下焙烧珍珠岩才能使其产量提高，并节约能源呢？这是一个有关温度最佳点的选择问题，也是选优问题.

怎样才能找到这类问题的最佳方案呢？有时可以用理论计算的方法求得，但是在许多实际情况下，试验结果与因素的关系要么很难用数学形式来表达，要么表达式很复杂，因此只能通过试验找到最佳方案.

这样就出现了如何合理安排试验的问题，也就是优选方法的问题. 比如上面所说的焙烧珍珠岩问题中，焙烧温度高了不好，低了也不好. 假如已经知道温度的选择范围在 1 000 ℃到 2 000 ℃之间，要求确定它的最佳操作温度. 如果每隔 1 ℃做一次试验，那就要做 1 000 次试验才能找到最好的操作温度（精度为 1 ℃）；如果每隔 10 ℃做一次试验，那也要做 100 次（精度为 10 ℃）. 这种方法叫做“均分法”. 用均分法找最佳点，既浪费精力、时间，又浪费原材料，而且有时还不一定能找到. 如在焙烧珍珠岩问题中，除了温度因素外，还要考虑焙烧时间，设它也有 100 种可能，

因此用均分法找最佳的温度与时间，要做试验的次数为

$$1\,000 \times 100 = 100\,000(\text{次}),$$

而 10 万次试验要全部做完显然是不大可能的.

而**优选法**(optimum seeking method)就是利用数学原理，合理安排试验点，以尽量少的试验次数，找到上述一类问题①的最佳方案的一种科学方法.

就上面的例子来说，如果用优选法做试验，只要做 15 次试验，精度就达到 1 ℃以内，超过均分法做 1 400 次实验的效果，可见优选法是解决此类问题的一种很好的方法.

下面，我们通过一个实例来了解优选法是如何安排试验的.

例 1 焙烧珍珠岩的问题中，已经知道在 1 000 ℃到 2 000 ℃之间有温度 c，使产品的产量 $y = f(c)$ ②最好. 现在限于生产条件，只能做 4 次试验，找出温度 x_0，使 x_0 与 c 的距离尽可能小.

如果知道产量 y 与温度 x 之间的函数表达式 $y = f(x)$，在理论上有许多找 c 点的方法. 但在这个实际问题中，并不知道 $f(x)$ 的具体函数表达式，只能通过试验来找最接近 c 的点. 我们现在的任务是在指定的试验次数 4 内，找到最接近 c 的点.

方法 1　用均分法寻找最佳点

将区间(1 000，2 000)5 等分，在 4 个分点

$$1\,200,\ 1\,400,\ 1\,600,\ 1\,800$$

上做试验，试验得到产量 y 的值分别为

$$63\%,\ 74\%,\ 80\%,\ 81\%,$$

① 这里的“一类问题”指单峰问题，也就是在试验范围内，试验指标只有一个峰，具体说明见 7.1.6 小节.

② 我们在本例中假设产量函数为 $y = f(x) = \sin\left(\frac{x}{1\,000}\right) - e^{-\frac{x}{1\,000}}$. 如果事先知道这个函数形式，可以求导后用两分法求出极大值点 c；如果事先不知道这个函数形式，就只能通过做试验来找较接近 c 的点 x_0.

则估计最佳点 c 在 1 800 附近，在区间(1 600，2 000)内，精度为 200(℃).

方法 2　用优选法的分数法寻找最佳点

第 1 步　将区间(a, b)，即(1 000，2 000)8 等分，取第 5 个分点为第一个试验点 x_1，即

$$x_1 = a + (b - a) \times \frac{5}{8} = 1\,000 + (2\,000 - 1\,000) \times \frac{5}{8} = 1\,625,$$

在 x_1 处进行第一次试验，试验得到产量 $y_1 = f(x_1)$ 的值为 80%(如图 7-1-1).

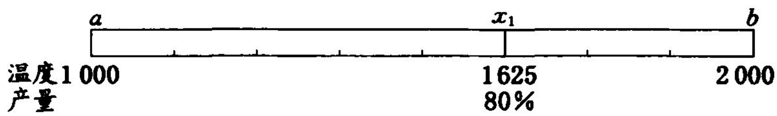


图 7-1-1

第 2 步　取第二个试验点 x_2 为 x_1 在区间(a, b)内的对称点，即

$$x_2 = b + a - x_1 = 2\,000 + 1\,000 - 1\,625 = 1\,375$$

(x_2 是整个区间的第 3 个分点)，在 x_2 处做第 2 次试验，得 $f(x_2) =$ 73%(如图 7-1-2).

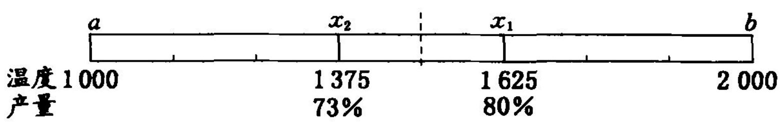


图 7-1-2

第 3 步　指标函数 $f(x_1)$优于 $f(x_2)$，x_1 是好点，x_2 是差点，我们舍去 x_2 的左边的区间$(a, x_2]$. 余下的试验区间为(x_2, b)，即(1 375，2 000)(如图 7-1-3).

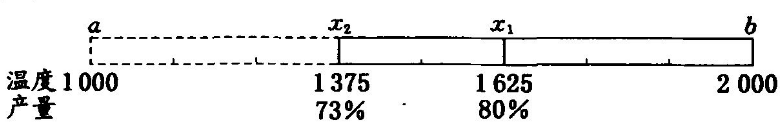


图 7-1-3

第 4 步　取第三个试验点 x_3 为 x_1 在余下的试验区间内的对称点，

$$x_3 = b + x_2 - x_1 = 2\,000 + 1\,375 - 1\,625 = 1\,750$$

(x_3 是整个区间的第 6 个分点)，在 x_3 处做第 3 次试验，得 $f(x_3)=81\%$(如图 7-1-4).

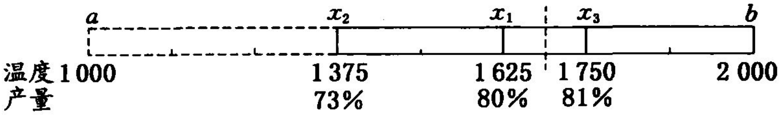


图 7-1-4

第 5 步　指标函数 $f(x_3)$优于 $f(x_1)$，所以舍去 x_1 的左边的区间 $(x_2, x_1]$. 余下的试验区间为(x_1, b)，即(1 625，2 000)(如图 7-1-5).

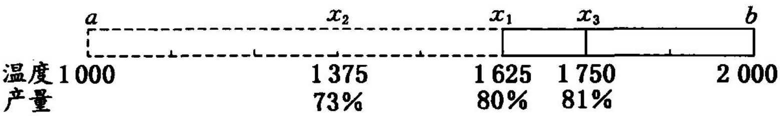


图 7-1-5

第 6 步　取第四个试验点 x_4 为 x_3 在余下的试验区间内的对称点，

$$x_4=b+x_1-x_3=2\,000+1\,625-1\,750=1\,875$$

(x_4 是整个区间的第 7 个分点)，在 x_4 处做第 4 次试验，得 $f(x_4)=80\%$(如图 7-1-6).

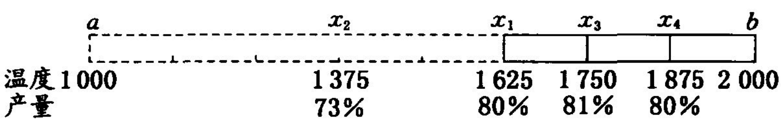


图 7-1-6

第 7 步　指标函数 $f(x_3)$优于 $f(x_4)$，所以舍去 x_4 的右边的区间 $[x_4, b)$. 余下的试验区间为(x_1, x_4)，即(1 625，1 875)(如图 7-1-7).

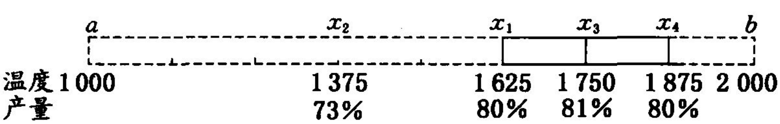


图 7-1-7

现在已做完 4 次试验，估计最佳点 c 在 1 750 附近，在区间(1 625，1 875)内，精度为 125(℃)，也就是区间长度的$\frac{1}{8}$. 我们看到，同样是做 4 次试验，用分数法进行试验的精度(125 ℃)比均分法的(200 ℃)高.

如果做10次试验，分数法的精度可达区间长度的$\frac{1}{144}$；如果做30次试验，精度达区间长度的$\frac{1}{2\ 178\ 309}$；而用均分法要达到这样的精度，则分别需做100，2 000 000以上次数的试验. 由此可见，好的试验方法可以大大减少试验的次数.

7.1.2 分数法

在单因素优选问题中，如果因为某些原因，只允许做指定次数的试验，就像7.1.1小节例1中指定只做4次试验，或者试验只能在试验范围的几个分点上进行，要找出最好的分点，就可以用分数法进行试验.

7.1.1小节的例子抽象起来看，就是：先将试验范围(1 000，2 000)8等分，从$\frac{5}{8}$出发，用对称的方法找出下一个分点，4次就可以找到7个分点中最好的分点.

同样，可以将试验范围13等分，从$\frac{8}{13}$出发，5次找到12个分点中最好的分点；将试验范围21等分，从$\frac{13}{21}$出发，6次找到20个分点中最好的分点……

这种方法叫做“分数法”，因为试验是从分数$\frac{1}{2}$，$\frac{2}{3}$，$\frac{3}{5}$，$\frac{5}{8}$，$\frac{8}{13}$，$\frac{13}{21}$，$\frac{21}{34}$，……出发进行的. 怎样求得这些分数呢？事实上，这些分数是数列

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ \cdots\cdots$$

相邻两数的商. 此数列的构造方法是，前两个数都是1，后面的每个数是它前面两个数的和. 记这个数列为$\{F_n\}$，则

$$F_0 = F_1 = 1,\ F_{n+1} = F_n + F_{n-1},\ n = 1,\ 2,\ \cdots\cdots$$

数列$\{F_n\}$称为斐波那契数列，分数法中所用分数就是数列

$$\frac{F_1}{F_2},\ \frac{F_2}{F_3},\ \frac{F_3}{F_4},\ \cdots\cdots,\ \frac{F_n}{F_{n+1}},\ \cdots\cdots$$

中的数.

用分数法进行优选试验的步骤是：

(1) 明确实际问题的试验范围；

(2) 指定需要试验的次数 n；

(3) 根据斐波那契数列找出分数 $\frac{F_n}{F_{n+1}}$；

(4) 计算第 1 个试验点的位置. 将试验区间 (a, b) F_{n+1} 等分，第 1 个试验点在第 F_n 个分点处.

即第 1 个试验点 x_1 的计算公式是①

$$\boxed{小+(大-小)\times\frac{F_n}{F_{n+1}}.}$$

在 x_1 处进行第 1 次试验，得到结果 y_1；

(5) 计算第 2 个试验点的位置，它是第 1 个实验点在试验范围内的对称点，计算公式是

$$\boxed{大+小-中.}$$

在 x_2 处进行第 2 次试验，得到结果 y_2；

(6) 比较两点的试验结果，保留好点，舍去差点以外的部分；

(7) 在剩下的范围内再取保留点的对称点作为第 3 个试验点，比较两点的试验结果，依上面“保留好点，舍去差点以外的部分”的原则继续下去，共进行 n 次试验，得到离最佳点最近的分点.

例 1 卡那霉素发酵液生物测定培养温度优选.

卡那霉素发酵液生物测定，国内外通常规定培养温度为 37±1 ℃，培养时间在 16 小时以上. 某制药厂为缩短时间，决定优选培养温度，试验范围定为 29～50 ℃，精确度要求 ±1 ℃. 中间试验点共有 20 个，用分

① 公式中的“大”、“小”是试验范围的两个端点，它随着试验次数的不同而取不同的值. 公式中的“中”是指在保留的试验范围内的试验点，不是指中点.

数法安排试验，要将试验范围 21 等分，于是 $F_{n+1}=21$，所以 $n=6$，共需做 6 次试验.

第一个试验点选在第 $(F_n=)13$ 个分点，即

$$x_1=29+(50-29)\times\frac{13}{21}=42(℃)$$

处；第二个试验点是其对称点，即

$$x_2=50+29-42=37(℃)$$

的地方，也就是第 8 个分点处；依次类推，6 次试验的结果见图 7-1-8，也可以将试验过程列成表 7-1.

从图表中可以看到，在 41～43 ℃培养最好，只需 8～9 小时.

试验点 x_i									x_2			x_4	x_6	x_1	x_5		x_3						
试验点编号	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
培养温度 x	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	℃
培养时间 y_i									16			10	9	8	8.5		11						小时

图 7-1-8

表 7-1 卡那霉素发酵液生物测定培养温度优选①

试验范围(℃)	试验点 x_i(℃)	试验结果 y_i(小时)	结果比较	舍去范围(℃)
29～50	$x_1\left(=29+(50-29)\times\frac{13}{21}=42\right)$	8		
29～50	$x_2(=50+29-42=37)$	16	y_1 比 y_2 好	29～37
37～50	$x_3(=50+37-42=45)$	11	y_1 比 y_3 好	45～50
37～45	$x_4(=45+37-42=40)$	10	y_1 比 y_4 好	37～40
40～45	$x_5(=45+40-42=43)$	8.5	y_1 比 y_5 好	43～45
40～43	$x_6(=43+40-42=41)$	9	y_1 比 y_6 好	40～41

① 在以后的例子中，我们都用 x_i 表示第 i 次试验的试验点位置，y_i 表示第 i 次试验的结果.

例 2 在机械加工中，车床的走刀量 mm/转是不均匀地分档的，也就是说试验点只能取在某档上. 设某车床走刀量共有 10 档，试对走刀量（档数）用分数法进行优选.

由于走刀量是不均匀地分档的，因此，不能直接对走刀量进行优选. 现在，我们把对走刀量的优选变为对档的顺序进行优选.

先将各档由小到大排列，记为 1, 2, …, 10（档）. 由于 $10+1=11$ 不是斐波那契数列中的某一项，而 $8<11<13$，于是，我们可以增加两个虚档 11, 12，再在两端增加两个虚档 0, 13（如图 7-1-9），并规定在虚档上的试验结果总是最差的. 这样就可以用分数为 $\frac{8}{13}$ $(n=5)$ 的分数法进行优选试验. 试验过程见图 7-1-9 和表 7-2.

通过用分数法优选车床的走刀量档数，知道最好档是 4 档（走刀量 0.48 mm/转）.

表 7-2 车床走刀量（档数）优选

试验范围	试验点 x_i（档数）	结果比较	舍去范围
0～13	$x_1(=0+(13-0)\times\frac{8}{13}=8)$		
0～13	$x_2(=13+0-8=5)$	y_2 比 y_1 好	8～13
0～8	$x_3(=8+0-5=3)$	y_2 比 y_3 好	0～3
3～8	$x_4(=8+3-5=6)$	y_2 比 y_4 好	6～8
3～6	$x_5(=6+3-5=4)$	y_5 比 y_2 好	5～6

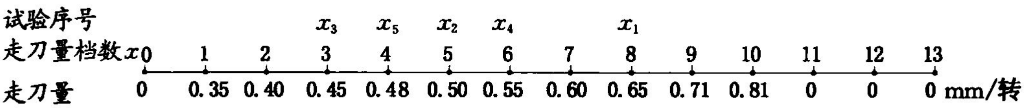


图 7-1-9

我们在分数法中采用了对称取点的方法. 那么，为什么要对称取点呢？

设 x, y $(y<x)$ 是试验范围 (a, b) 内的两个试验点，因为不知道好点在哪里，所以去掉 (a, y) 和去掉 (x, b) 的可能性是相同的. 如果关于

(a, b)中点对称地取x, y, (a, y)与(x, b)就等长，不论去掉哪一段，留下的试验范围长度相同；反之，如果不是对称取点，例如取y' $(y' < y)$，在去掉较短的一段(a, y')时，对于迅速减小试验范围不利(如图 7-1-10).

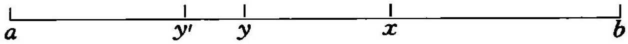


图 7-1-10

在分数法中为什么要用斐波那契数列相邻两数构成的一串分数$\frac{1}{2}$，$\frac{2}{3}$，$\frac{3}{5}$，$\frac{5}{8}$，……作为试验的第一点在整个试验区间内的相对位置呢？这可以参看“阅读:斐波那契数列与优选法”.

阅　读

斐波那契数列与优选法

在“分数法”一节中，我们用斐波那契数列相邻两数构成的一串分数$\frac{1}{2}$，$\frac{2}{3}$，$\frac{3}{5}$，$\frac{5}{8}$，……作为试验的第1点在整个试验区间内的相对位置.这是为什么呢？

下面，我们在对预给的试验次数n $(n=1, 2, 3, 4, \cdots\cdots)$依次进行讨论时，为了方便起见，假定试验范围是区间$(0, 1)$，并且，将试验得到的好点与单峰曲线的真正好点的最大可能距离作为精度.

(1) $n=1$，即预给做1次试验.

这时，试验点x取在$\frac{1}{2}$处，精度是$\frac{1}{2}$(如图 7-1-11).

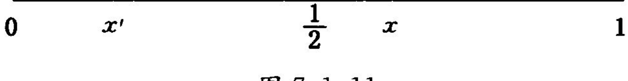


图 7-1-11

我们说明，$\frac{1}{2}$是预给做1次试验情形下的最佳精度.事实上，当$x > \frac{1}{2}$时，区间$(0, x)$的长度大于$\frac{1}{2}$，因而精度大于$\frac{1}{2}$；类似地，$x' < \frac{1}{2}$时，区间$(x', 1)$的长度大于$\frac{1}{2}$，因而精度也大于$\frac{1}{2}$.

因此，预给做 1 次试验的情形下，以从$\frac{1}{2}$出发的分数法最好.

(2) $n=2$，即预给做 2 次试验.

我们这样来考虑，设对称的两个试验点是 x 和 y $(y<x)$，比较后去掉的一段是区间$(x, 1)$，在留下的区间$(0, x)$中只能安排 1 次试验（如图 7-1-12(1)）. 根据上面对 $n=1$ 情形的讨论可知，y 应取在区间$(0, x)$的$\frac{1}{2}$处，也就是说

$$y=\frac{1}{2}x.$$

再由点 x 和点 y 对称，可知

$$y=1-x,$$

因此得出

$$\frac{1}{2}x=1-x,$$

即

$$x=\frac{1}{1+\frac{1}{2}}=\frac{2}{3}$$

$\left(\text{如果去掉的一段是区间}(0, y)\text{，类似地也可得出 } x=\frac{2}{3}\right)$. 这时精度是$\frac{1}{3}$.

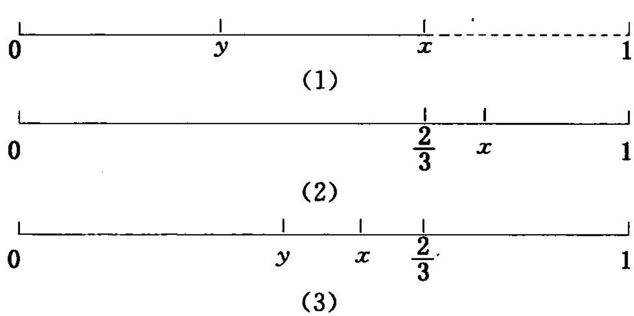


图 7-1-12

我们说明，$\frac{1}{3}$是预给做 2 次试验情形下的最佳精度. 事实上，当 $x>\frac{2}{3}$ 时，区间$(0, x)$的长度大于$\frac{2}{3}$（如图 7-1-12(2)），因而在$(0, x)$内再做

一次试验后，精度必大于$\frac{1}{3}$；当$x<\frac{2}{3}$时，区间$(0, y)$的长度大于$\frac{1}{3}$(如图7-1-12(3))，因而在y处再做一次试验后，精度也必大于$\frac{1}{3}$.

因此，预给做2次试验的情形下，以从$\frac{2}{3}$出发的分数法最好.

(3) $n=3$，即预给做3次试验.

仍然设第一、第二次试验点是x和y $(y<x)$，比较后去掉的一段是区间$(x, 1)$，在留下的区间$(0, x)$中，只能安排2次试验(如图7-1-13(1)). 根据上面对$n=2$的讨论，可知y应取在区间$(0, x)$的$\frac{2}{3}$处，也就是说

$$y=\frac{2}{3}x.$$

再由点x和点y对称，可知

$$y=1-x,$$

因此得出

$$\frac{2}{3}x=1-x,$$

即

$$x=\frac{1}{1+\frac{2}{3}}=\frac{3}{5}$$

$\left(\text{如果去掉的一段是区间}(0, y)\text{，类似地也可得出 } x=\frac{3}{5}\right)$. 这时精度是$\frac{1}{5}$.

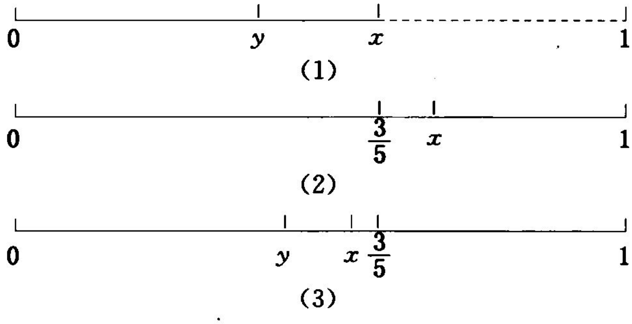


图 7-1-13