第 1 章　数学建模概述

§1.1　什么是数学模型

通常我们把现实问题的一个模拟称为模型，如交通图、地质图、航空模型和建筑模型等.利用数学的语言、公式、图、表或符号等来模拟现实的模型称为数学模型.我们知道，对一个现实问题的研究，一般不需要甚至不可能直接研究现实问题的本身，而是研究模拟该现实问题的模型.举个简单例子：某司机欲把某货物从甲地运往乙地，应如何选择运输路线使总路程最短？该司机不会开着车去试探，而是利用交通图来确定自己的行车路线.从这个简单的例子中我们可以看到数学建模的重要性.

§1.2　数学建模包含哪些步骤

数学建模主要包含模型建立、求解以及对结果的分析与检验等步骤.

模型建立　模拟现实问题建立数学模型，不仅要有一定的数学知识与技巧，还要有敏锐的洞察力与理解力，善于抓住问题的内在联系，作出合理的假设与简化，找出影响问题的各种因素及其相互关系.例如，2000 年全国大学生数学建模竞赛 A 题 :DNA 系列分类中，给出的 DNA 系列全是由 4 个字符（碱基)A,T,C,G 组成虽然 3 个碱基可以构成一种氨基酸（共有 20 多种),但是 DNA 系列的长度不一,且无法‘断句”,这给分类造成困难.然而，每条 DNA 中 A,T,C,G 的百分含量可体现该条 DNA 的特性,且与系列的长度和‘断句’无关,并且根据 DNA 中 A,T,C,G 的百分含量对 DNA 序列进行分类的方法比较多（此例详见本书第 5 章 §5.6).从这例可看出，建立数学模型，不仅要有一定的数学知识与技巧，还要具备其他学科的一些知识.另外还要有一定的编程能力.

一般来说，模型建立的方法不止一种.如最短路线问题，可以用图论方法，也可以用线性规划方法，有时还可以用动态规划的方法.

模型求解　在建立模型之后，就要求解模型，给出有效的计算方法.例如旅

行推销员问题：一个推销员要到 n 个城市去推销，如何安排行程？如果用简单的组合算法，其计算步数是 $n!$ 的倍数,随着 n 的增大，计算量之大以至无法得到结果．如 $n=30$,即使以每秒 10^{24} 步的速度来计算，也需要 8 年多，况且现在的计算机还没有达到上述速度．顺便说一句，旅行推销员问题的有效算法目前还没有找到，只有一些近似算法．

结果的分析与检验　有些问题需要对解的现实意义作出解释，检验模型的正确性，并对模型的稳定性进行分析．如种群的相互竞争问题（参考第 4 章)需要对解的现实意义作出解释，并对模型的稳定性进行分析．又如 2001 年全国大学生数学建模竞赛 B题—— 公交车的调度，需要对模型结果进行检验．

具体来说，在建立数学模型，撰写论文时，建议包含以下内容：

(1) 问题的提出与分析按你的理解对所给的问题用数学语言作更清晰的表述，根据问题的性质你打算建立什么样的模型．

(2) 必要而合理的假设．现实问题远不像纯粹数学问题那样理想化，因此有必要作一些合理的假设，有条件时对实际问题作一些调查．

(3) 模型设计．对出现的数学符号必须有明确的定义，模型的具体表现形式必须写清楚．

(4) 模型求解与结果将具体的求解过程和结果写清楚．

(5) 模型结果的检验与分析．对模型的结果进行检验与分析，可以通过实际问题来检验，也可以用计算机模拟检验，并指出模型的优缺点及改进方向．

(6) 必要的计算机程序．

§1.3　建模实例

1.3.1　报童订报模型

(1) 问题

报童每天清晨从报社购进报纸零售，晚上将没有卖掉的报纸退回．每份报纸订购价格为 a,零售价格为 b,退回价格为 $c(b>a>c)$．请你为报童制定一个最佳订购方案．

(2) 问题的分析

报童每天卖出报纸的数量 ξ 是一个随机变量，因此报童每天的收入也是一个随机变量，所以作为优化模型的目标函数，不能是报童每天的收入，而应该是他长期卖报的日平均收入．从概率论中大数定律的观点来看，这相当于报童每天收入的期望值．另一方面，如果报纸订得太少，供不应求，报童就会失去一些挣钱

的机会，将会减少收入；但如果订多了，当天卖不完，退回的多余报纸得赔钱，报童也会减少收入

（3）问题的假设

设报社有足够的报纸可供定购；当天卖不出去的报纸只能退回；报童除了订购报纸费用外，其他费用（如交通费、摊位费等）一概不计；报童每天订购 n 份报纸,实际能卖出 r 份报纸,且 $P\{\xi=r\}=p(r)$

（4）模型建立

如果 $0\leqslant r\leqslant n$,则售出 r 份报纸增加收入$(b-a)r$,退回 $n-r$ 份减少收入$(a-c)(n-r)$ 如果 $r>n$,则售出 n 份报纸增加收入 $(b-a)n$. 因此报童每天收入的期望值：

$$f(n)=\sum_{r=0}^{n}[(b-a)r-(a-c)(n-r)]p(r)+\sum_{r=n+1}^{\infty}(b-a)np(r). \tag{1-1}$$

问题归结为在 $a,b,c,p(r)$为已知时,求 n 使 $f(n)$最大.

（5）模型求解与结果

方法 1　用差分法求解令

$$\Delta f(n)=f(n+1)-f(n)=0, \tag{1-2}$$

上式表示多订一份报纸与少订一份报纸的收入期望值相等.由(1-1)式得

$$\begin{aligned}
f(n+1)&=\sum_{r=0}^{n+1}[(b-a)r-(a-c)(n+1-r)]p(r)+\sum_{r=n+2}^{\infty}(b-a)(n+1)p(r)\\
&=\sum_{r=0}^{n}[(b-a)r-(a-c)(n+1-r)]p(r)+\sum_{r=n+1}^{\infty}(b-a)(n+1)p(r)\\
&=\sum_{r=0}^{n}[(b-a)r-(a-c)(n-r)]p(r)-\sum_{r=0}^{n}(a-c)p(r)+\\
&\quad\sum_{r=n+1}^{\infty}(b-a)np(r)+\sum_{r=n+1}^{\infty}(b-a)p(r)\\
&=f(n)-\sum_{r=0}^{n}(a-c)p(r)+\sum_{r=n+1}^{\infty}(b-a)p(r)\\
&=f(n)-\sum_{r=0}^{n}(b-c)p(r)+\sum_{r=0}^{n}(b-a)p(r)+\sum_{r=n+1}^{\infty}(b-a)p(r)\\
&=f(n)-(b-c)\sum_{r=0}^{n}p(r)+(b-a)\sum_{r=0}^{\infty}p(r).
\end{aligned}$$

注意到 $\sum_{r=0}^{\infty}p(r)=1$,得

$$f(n+1)-f(n)=b-a-(b-c)\sum_{r=0}^{n}p(r). \tag{1-3}$$

由(1－2)式和(1－3)式得

$$\sum_{r=0}^{n} p(r)=\frac{b-a}{b-c}, \tag{1-4}$$

从(1－4)式中解出 n，可使报童每天收入的期望值 $f(n)$最大.

方法 2　当 r 和 n 都相当大时，可将 r 视为连续变量更便于分析和计算，这时概率 $p(r)$转化为概率密度函数 $\varphi(r)$，由(1－1)式得

$$\begin{aligned}
f(n)&=\int_0^n[(b-a)r-(a-c)(n-r)]\varphi(r)\mathrm{d}r+\int_n^{\infty}(b-a)n\varphi(r)\mathrm{d}r\\
&=\int_0^n(b-c)r\varphi(r)\mathrm{d}r-(a-c)n\int_0^n\varphi(r)\mathrm{d}r+(b-a)n\int_n^{\infty}\varphi(r)\mathrm{d}r\\
&=(b-c)\int_0^n r\varphi(r)\mathrm{d}r-(b-c)n\int_0^n\varphi(r)\mathrm{d}r+(b-a)n\int_0^{\infty}\varphi(r)\mathrm{d}r\\
&=(b-c)\int_0^n r\varphi(r)\mathrm{d}r-(b-c)n\int_0^n\varphi(r)\mathrm{d}r+(b-a)n.
\end{aligned}$$

计算

$$\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}n}&=(b-c)n\varphi(n)-(b-c)n\varphi(n)-(b-c)\int_0^n\varphi(r)\mathrm{d}r+(b-a)\\
&=b-a-(b-c)\int_0^n\varphi(r)\mathrm{d}r,
\end{aligned}$$

令$\frac{\mathrm{d}f}{\mathrm{d}n}=0$得，

$$\int_0^n\varphi(r)\mathrm{d}r=\frac{b-a}{b-c}. \tag{1-5}$$

从(1－5)式中解出 n，可使报童每天收入的期望值 $f(n)$最大.

(6) 模型结果的模拟检验

报童每天卖出报纸的数量 ξ是一个随机变量，它一般服从泊松（Poisson)分布

$$P\{\xi=r\}=\frac{\lambda^r}{r!}\mathrm{e}^{-\lambda} \tag{1-6}$$

或服从正态分布

$$\xi\sim N(\mu,\sigma^2). \tag{1-7}$$

(1－6)式或(1－7)式中的参数可根据统计报童以前卖出报纸的数量得到

当 $a=35,b=50,c=12,\xi\sim N(80,20^2)$时，从(1－5)式中可解出 $n=75$.用计算机模拟产生服从正态分布 $N(80,20^2)$的随机数 20 个如下：

93	85	103	73	70	53	80	93	90	59
81	97	38	64	86	79	69	87	53	88

假设上述数据为报童 20 天中每天实际卖出报纸的份数，则当报童每天订购 73 份报纸时，20 天的总收入 17910 达到最大，这一结果与本模型中制定报童每天订购 75 份报纸（20 天的总收入 17902）的方案基本相同.

（7）模型的推广

报童订报模型适用于一些季节性强、更新快、不易保存等特点的货物订货模型.但是模型中有一个严格的限制条件：两次订货之间没有联系，这种策略是决策论中的一种定期定量订货策略.

1.3.2 空洞探测模型

本题为 2000 年全国大学生数学建模竞赛 D 题.

作者：黄祥富　吴成　郑玉学　指导老师：谢冰锋

（1）问题

山体、隧洞、坝体等的某些内部结构可用弹性波测量来确定.一个简化问题可描述为，一块均匀介质构成的矩形平板内有一些充满空气的空洞，在平板的两个邻边分别等距地设置若干波源，在它们的对边对等地安放同样多的接收器，记录弹性波由每个波源到达对边上每个接收器的时间，根据弹性波在介质中和在空气中不同的传播速度，来确定板内空洞的位置.现考察如下的具体问题：

一块 240(m)×240(m)的平板（如图 1－1），在 AB 边等距地设置 7 个波源 $P_i(i=1,\cdots,7)$，CD 边对等地安放 7 个接收器 $Q_j(j=1,\cdots,7)$，记录由 P_i 发出的弹性波到达 Q_j 的时间 t_{ij}(t)；在 AD 边等距地设置 7 个波源 $R_i(i=1,\cdots,7)$，BC 边对等地安放 7 个接收器 $S_j(j=1,\cdots,7)$，记录由 R_i 发出的弹性波到达 S_j 的时间 τ_{ij}(t).已知弹性波在介质和空气中的传播速度分别为 2 880(m/s)和 320(m/s)，且弹性波沿板边缘的传播速度与在介质中的传播速度相同.表 1－1，表

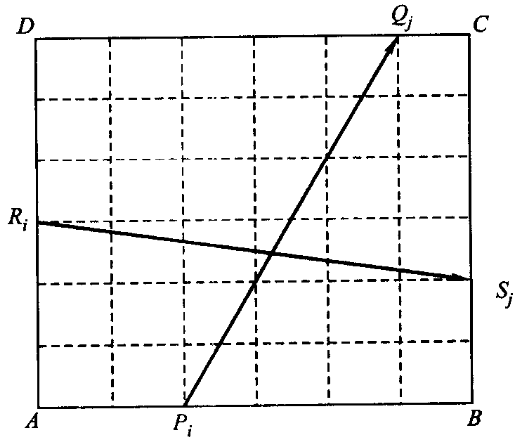

图 1－1

1－2 为观测到的数据.

①确定该平板内空洞的位置.

②只根据由 P_i 发出的弹性波到达 Q_j 的时间 $t_{ij}(i,j=1,\cdots,7)$，能确定空洞的位置吗？讨论在同样能够确定空洞位置的前提下，减少波源和接受器的方法

表 1－1

t_{ij}	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7
P_1	0.061 1	0.089 5	0.199 6	0.203 2	0.418 1	0.492 3	0.564 6
P_2	0.098 9	0.059 2	0.441 3	0.431 8	0.477 0	0.524 2	0.380 5
P_3	0.305 2	0.413 1	0.059 8	0.415 3	0.415 6	0.356 3	0.191 9
P_4	0.322 1	0.445 3	0.404 0	0.073 8	0.178 9	0.074 0	0.212 2
P_5	0.349 0	0.452 9	0.226 3	0.191 7	0.083 9	0.176 8	0.181 0
P_6	0.380 7	0.317 7	0.236 4	0.306 4	0.221 7	0.093 9	0.103 1
P_7	0.431 1	0.339 7	0.356 6	0.195 4	0.076 0	0.068 8	0.104 2

表 1－2

τ_{ij}	S_1	S_2	S_3	S_4	S_5	S_6	S_7
R_1	0.064 5	0.060 2	0.081 3	0.351 6	0.386 7	0.431 4	0.572 1
R_2	0.075 3	0.070 0	0.285 2	0.434 1	0.349 1	0.480 0	0.498 0
R_3	0.345 6	0.320 5	0.097 4	0.409 3	0.424 0	0.454 0	0.311 2
R_4	0.365 5	0.328 9	0.424 7	0.100 7	0.324 9	0.213 4	0.101 7
R_5	0.316 5	0.240 9	0.321 4	0.325 6	0.090 4	0.187 4	0.213 0
R_6	0.274 9	0.389 1	0.589 5	0.301 6	0.205 8	0.084 1	0.070 6
R_7	0.443 4	0.491 9	0.390 4	0.078 6	0.070 9	0.091 4	0.058 3

（2）问题的分析

用弹性波测量山体、隧洞、坝体等的某些内部空洞的位置，应该说只能测出空洞的大致位置，这是由于安放在矩形每一对边的探测器的数量是有限的．因此我们将一块矩形平板划分为若干个小矩形块，判断每个小矩形块内是否有空洞．测量的精度与安放在矩形每一对边的探测器的数量是有关的．

（3）问题的假设

①[注]假设所有的探测均在同一平面上进行，且所有的测量数据是按弹性波走直线得到的．

② 假设所有的测量数据是可靠的，但存在测量误差．若某路径上的弹性波到达时间的实际测量值小于无空洞时的理论计算值，我们认为是由测量误差引起的,并由此确定误差范围．

③若某路径上的弹性波到达时间的实际测量值与无空洞时的理论计算值的差值在误差范围内，我们认为该路径上不存在空洞；否则认为该路径上存在空洞．

(4)符号说明

除了问题中已给的符号外，再设

① 无空洞时弹性波到达时间的理论计算值 $T_{ij}=|P_iQ_j|/2880=|R_iS_j|/2880$,计算结果如表 1－3.

表 1－3

T_{ij}	1	2	3	4	5	6	7
1	0.083 3	0.084 5	0.087 8	0.093 2	0.100 2	0.108 5	0.117 9
2	0.084 5	0.083 3	0.084 5	0.087 8	0.093 2	0.100 2	0.108 5
3	0.087 8	0.084 5	0.083 3	0.084 5	0.087 8	0.093 2	0.100 2
4	0.093 2	0.087 8	0.084 5	0.083 3	0.084 5	0.087 8	0.093 2
5	0.100 2	0.093 2	0.087 8	0.084 5	0.083 3	0.084 5	0.087 8
6	0.108 5	0.100 2	0.093 2	0.087 8	0.084 5	0.083 3	0.084 5
7	0.117 9	0.108 5	0.100 2	0.093 2	0.087 8	0.084 5	0.083 3

②最大时间偏差：$\Delta t=\max\{T_{ij}-t_{ij},T_{ij}-\tau_{ij}\mid T_{ij}>t_{ij},T_{ij}>\tau_{ij},1\leqslant i,j\leqslant 7\}=T_{77}-\tau_{77}=0.083\,3-0.058\,3=0.025$ (s).

③ x_{ij} 表示该块(边长为 40 米的正方形)是否为空洞块的变量（$x_{ij}=1$ 时，表示该块是空洞块；$x_{ij}=0$ 时，表示该块不是空洞块．如图 1－2).

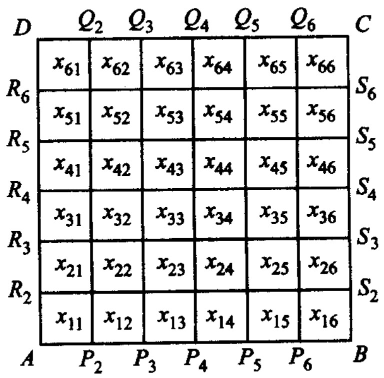

图 1－2

(5) 模型建立

为了确定该平板内空洞的位置，先给出 x_{ij} 满足的条件．在连线 $P_iQ_j(i\neq j)$上,一共有 36 个不等式约束．例如在连线 P_3Q_6 上的不等式约束为

$$-\Delta t\leqslant(|P_3Q_6|/320-|P_3Q_6|/2880)$$

$(x_{13}+x_{23}+x_{34}+x_{44}+x_{55}+x_{65})/6+T_{36}-\tau_{36}\leqslant\Delta t$

在连线 $R_iS_j(i\neq j)$上，一共也有 36 个不等式约束例如在连线 R_2S_4 上的不等式约束为

$-\Delta t\leqslant(|R_2S_4|/320-|R_2S_4|/2880)(x_{21}+x_{22}+x_{23}+x_{34}+x_{35}+x_{36})/6+T_{24}-\tau_{24}\leqslant\Delta t$.

解出满足以上 72 个不等式中的 x_{ij}，即可确定该平板内空洞的位置.

（6）模型求解

根据假设②，③，在误差范围内画出无空洞的连线如图 1－3. 从图中可看出,除了 $x_{22},x_{23},x_{25},x_{32},x_{33},x_{34},x_{44},x_{53}$ 外,其余的 $x_{ij}=0$ 这样大大简化了模型求解过程，只需用简单的穷举法,即将变量 $x_{22},x_{23},x_{25},x_{32},x_{33},x_{34},x_{44},x_{53}$ 分别取 1 和 0(共有 $2^8=256$ 中取法，根据以上分析,先将上述 8个变量取 1),其余的变量都取 0 代入(5)中所指的 72个不等式中，最后求得其解为 $x_{22}=x_{23}=x_{25}=x_{32}=x_{33}=x_{34}=x_{44}=x_{53}=1$,其他 $x_{ij}=0$.

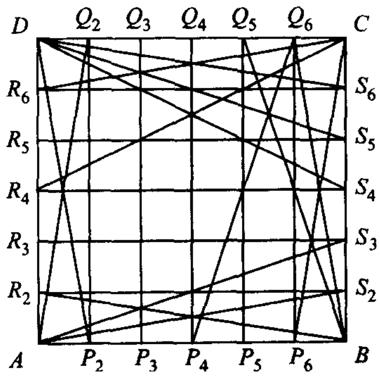

图 1－3

（7）对减少波源和接受器的方法进行讨论

只在 AB 边等距地设置 7 个波源 $P_i(i=1,\cdots,7)$,CD 边对等地安放 7 个接收器 $Q_j(j=1,\cdots,7)$,如图 1－4 此时当测量误差很小时，可以确定空洞的位置.但在本题的测量误差下，不能确定空洞的准确位置，即按（5）中的模型求解，得到的解不唯一（如图 1－5）

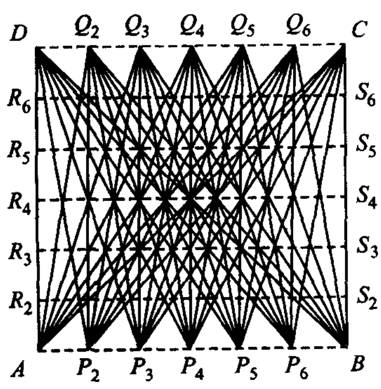

图 1－4

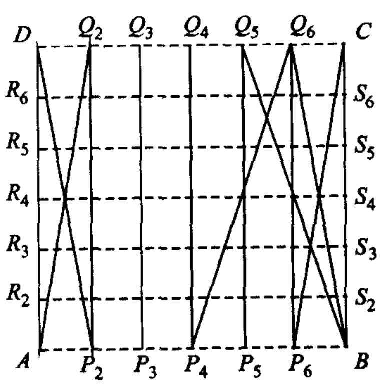

图 1－5

一般来说，减少波源和接受器对空洞的准确定位是有一定的影响，特别是当空洞比较小，测量误差又很大时，不能确定空洞的准确位置本题由于空洞较大，每一对边可减少一对探测器，依照(5）中的模型求解（波源和接受器依然等距设置,变量个数不变,不等式约束减至 50个），可以确定空洞的位置.

（8）模型的优缺点及改进方向

模型可应用于对山体、隧洞、坝体等某些内部空洞的定位.模型的缺点是对空洞的定位不够准确，改进的办法是将矩形平板划分为更多的小矩形块.当探测器的数量不变时，不等式约束的个数不会变，因为变量个数的增加会导致满足不等式约束的解不唯一.

（注：假设①与实际情况不相符这是命题的毛病——编者）

习 题 1

1. 找次品 80个乒乓球中有1个次品，它的质量比正品轻.现有一台无砝码的天平，如何找出该次品乒乓球？

2. 智力游戏1——安全渡河问题 3个商人各带1个随从乘船渡河，一只小船只能容纳2人，由他们自己划行.随从们密约，在河的任一岸，一旦随从的人数比商人多，就杀人越货.但是如何乘船渡河由商人们决定.商人们怎样才能安全渡河？

3. 智力游戏2——抓石子问题 将15颗小石子分为3堆,每堆分别有3,5,7颗,两人依次从中取走小石子，规定每次只能从一堆中取，至少要取走1颗，多取不限，取到最后一颗石子为胜.问先取者是否有必胜取法？

4. 边界点问题 在平面上有若干个点，如何找出其中的边界点？

第 2 章　数理统计方法

数理统计的任务是以概率论为基础，根据试验的数据，对研究对象的客观规律性作出合理的估计与推断.

§2.1　数理统计的基本概念

在数理统计中，我们把研究对象的全体称为总体，而把组成总体的每个基本单元称为个体.要了解总体的规律性，必须对其中的个体进行统计、观测，一是对全部个体逐一进行观测，这样做当然对总体有充分的了解，但实际上这种方法往往是行不通的，而且也很不经济；二是随机抽样观测，即从总体 X 中随机抽取 n 个个体($X_1, X_2, \cdots, X_n$) 进行观测，然后根据样本（$X_1, X_2, \cdots, X_n$）来推断总体的性质或规律性，这在实际中是常用的方法.

由于样本是随机抽样的，可以认为来自总体 X 中一个样本（$X_1, X_2, \cdots, X_n$）是一组相互独立且与总体 X 同分布的随机变量.n 称为样本容量，样本 $(X_1, X_2, \cdots, X_n)$ 的观测值记为$(x_1, x_2, \cdots, x_n)$，样本$(X_1, X_2, \cdots, X_n)$的不含任何未知参数的函数 $f(X_1, X_2, \cdots, X_n)$称为统计量.下述统计量

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

分别称为样本均值、样本方差、样本标准差.

根据样本的观测值（$x_1, x_2, \cdots, x_n$），可以绘出样本的频率直方图和累积频率直方图，方法如下：

①适当选取 $a \leqslant \min\{x_1, x_2, \cdots, x_n\}, b \geqslant \max\{x_1, x_2, \cdots, x_n\}$，将$[a, b]$等分为 m 个小区间，称 $d = (b-a)/m$ 为组距；

②计算 $x_1, x_2, \cdots, x_n$ 在各个小区间出现的频数 m_i，得到 $x_1, x_2, \cdots, x_n$ 在各个小区间出现的频率 $p_i = m_i/n, i = 1, 2, \cdots, m$；

③计算样本的频率函数 $p(x)$和累积频率函数 $F(x)$：

$$p(x)=\begin{cases}0, & x\leqslant a,\\ p_1, & a<x\leqslant a+d,\\ p_2, & a+d<x\leqslant a+2d,\\ \vdots & \vdots\\ p_m, & a+(m-1)d<x\leqslant a+md,\\ 0, & x>b=a+md,\end{cases}$$

$$F(x)=\begin{cases}0, & x\leqslant a,\\ p_1, & a<x\leqslant a+d,\\ p_1+p_2, & a+d<x\leqslant a+2d,\\ \vdots & \vdots\\ p_1+p_2+\cdots+p_{m-1}, & a+(m-2)d<x\leqslant a+(m-1)d,\\ 1, & x>a+(m-1)d.\end{cases}$$

④画出样本的频率函数 $p(x)$和累积频率函数 $F(x)$ 的图形可以得到样本的频率直方图（图 2－1)和累积频率直方图（图 2－2)

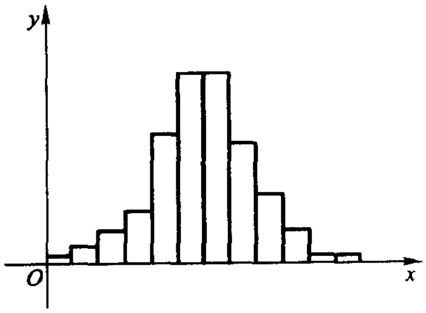

图 2－1

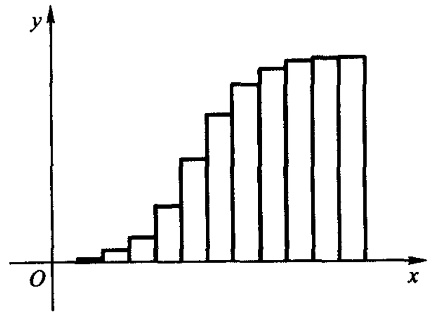

图 2－2

根据样本的频率直方图和累积频率直方图可以近似描绘出总体的分布密度函数曲线(图 2－3)和分布函数曲线（图 2－4)

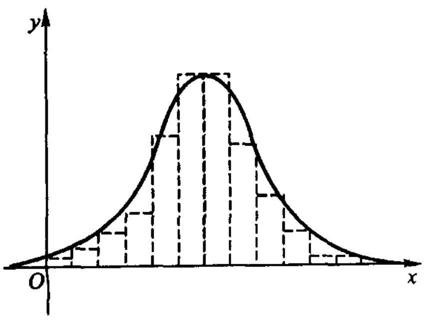

图 2－3

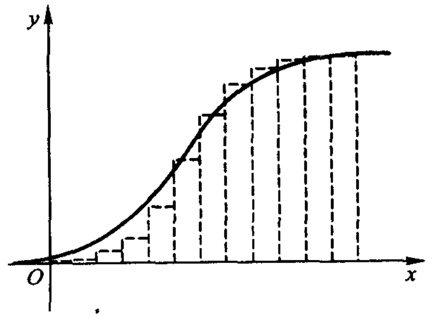

图 2－4

§2.2 参数估计

在对实际问题进行数学建模时，我们常常会遇到下列问题：在确定了问题涉及的关键量和发现了制约问题的基本规律或部分规律之后，可以得到刻画这些关键量之间关系的数学关系式.但是在这些数学关系式中尚包含若干未知参数.我们不能直接从这些关系式中得到这些关键量的定量变化规律，但实际问题往往又提供了某些表征关键量变化的信息（如实验数据等）.如果利用这些信息结合关键量的表达式来估计未知参数，那么实际问题能得到解决.

2.2.1 参数估计的方法

参数估计的方法比较多.一般情况下，参数估计问题可归结为求一个函数的极值点问题下面主要介绍最小二乘法和极大似然法.

(1) 最小二乘法

设 $y=f(x;\lambda)$,其中 x 是自变量（或自变量向量),λ 是未知参数,y 是 x 的函数,x 和 y 都是可观测的由于 λ 是未知的，因此要对 λ 进行估计.

设$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$是 n 组观测值，最小二乘法的基本思想就是求 λ 的一个估计值$\hat{\lambda}$,使函数

$$Q(\lambda)=\sum_{i=1}^{n}[y_i-f(x_i;\lambda)]^2$$

取最小值,即

$$Q(\hat{\lambda})=\min Q(\lambda)=\min\sum_{i=1}^{n}[y_i-f(x_i;\lambda)]^2,$$

称这样求得的 $\hat{\lambda}$ 为参数 λ 最小二乘估计值.

例 1 某种医用薄膜有允许一种物质的分子穿透它从高浓度的溶液向低浓度溶液扩散的功能，在试制时需测定薄膜被这种分子穿透的能力.测定方法如下 用面积为 s 的薄膜将容器分成 A,B 两部分，体积分别为 v_a 和 v_b,在两部分中分别注满该物质的两种不同浓度的溶液.此时该物质分子就会从高浓度溶液穿透薄膜向低浓度溶液中扩散.通过单位面积膜分子扩散的速度与两侧溶液的浓度差成正比，比例系数 k 表征了薄膜被该物质分子穿透的能力，称为渗透率.定时测量容器中某一侧的溶液浓度值以确定 k 的数值.试建立该问题的数学模型.

解 如果 $v_a=v_b=100\ \mathrm{cm}^3, s=10\ \mathrm{cm}^2$,且对容器的 B 部分在时刻 t_i 的溶液浓度 c_i 进行测试，结果如下：

$t_i(s)$:	100	200	300	400	500	600	700	800	900	1 000
$c_i(10^{-3}\mathrm{mg/cm^3})$:	4.54	4.99	5.35	5.65	5.90	6.10	6.26	6.39	6.50	6.59

试确定 k 的值.

假设薄膜两侧的溶液始终是均匀的，即在任何时刻膜两侧的每一处溶液的浓度都是相同的，且薄膜是双向同性的，即物质从薄膜任一侧向另一侧渗透的性能是相同的,设时刻 t 时膜两侧溶液的浓度分别为 $c_a(t)$和 $c_b(t)$,初始时刻膜两侧溶液的浓度分别为 α_0 和 α_b，单位均为 $10^{-3}\mathrm{mg/cm^3}$,又设 B侧在 t_i 时刻测得的浓度为 c_i.

考察时段 $[t,t+\Delta t]$膜两侧容器中该物质质量的变化.在容器的 A 侧,该时段内物质质量的增加为 $v_a c_a(t+\Delta t)-v_a c_a(t)$ 另一方面从 B侧渗透至 A侧的该物质质量为 $sk(c_b-c_a)\Delta t$. 由质量守恒定律，有

$$v_a c_a(t+\Delta t)-v_a c_a(t)=sk(c_b-c_a)\Delta t .$$

两边除以 Δt ,令 $\Delta t\to 0$,得

$$\frac{\mathrm{d}c_a}{\mathrm{d}t}=\frac{sk}{v_a}(c_b-c_a), \tag{2-1}$$

同理有

$$\frac{\mathrm{d}c_b}{\mathrm{d}t}=\frac{sk}{v_b}(c_a-c_b). \tag{2-2}$$

由(2-2)～(2-1)得

$$\frac{\mathrm{d}(c_b-c_a)}{\mathrm{d}t}=-sk\left(\frac{1}{v_a}+\frac{1}{v_b}\right)(c_b-c_a).$$

由于 $c_a(0)=\alpha_a, c_b(0)=\alpha_b$,所以

$$c_b-c_a=(\alpha_b-\alpha_a)\exp\left\{-sk\left(\frac{1}{v_a}+\frac{1}{v_b}\right)t\right\}. \tag{2-3}$$

由质量守恒定律,有

$$v_a c_a(t)+v_b c_b(t)=v_a\alpha_a+v_b\alpha_b. \tag{2-4}$$

联立(2-3),(2-4) 式可得:

$$c_b=\frac{\alpha_a v_a+\alpha_b v_b}{v_a+v_b}+\frac{v_a(\alpha_b-\alpha_a)}{v_a+v_b}\exp\left\{-sk\left(\frac{1}{v_a}+\frac{1}{v_b}\right)t\right\}.$$

至此，问题归结为利用 c_b 在时刻 t_i 的测量数据 $c_i(i=1,2,\cdots,n)$来估计 k, α_a,α_b ,根据使 $c_b(t_i)$与 c_i 的误差平方和最小的原则来求 k,α_a,α_b 的估计值.对应的数学模型为求函数

$$Q(k,\alpha_a,\alpha_b)=\sum_{i=1}^{n}[c_b(t_i)-c_i]^2$$

的最小值.

令

$$x=\frac{\alpha_a v_a+\alpha_b v_b}{v_a+v_b}, y=\frac{v_a(\alpha_b-\alpha_a)}{v_a+v_b},$$

则参数估计问题可转化为求函数

$$Q(k,x,y)=\sum_{i=1}^{n}\left[x+y\exp\left\{-sk\left(\frac{1}{v_a}+\frac{1}{v_b}\right)t_i\right\}-c_i\right]^2$$

的最小值点(k,x,y).将已知测试数据代入，得极小化的函数

$$Q(k,x,y)=\sum_{i=1}^{n}\left[x+y\mathrm{e}^{-kt_i/5}-c_i\right]^2.$$

利用 MATLAB 软件中的 fmins 函数求得 $k=0.010\ 12$ (cm/s) , $x=7(10^{-3}$ mg/cm$^3)$, $y=-3(10^{-3}$ mg/cm$^3)$，进一步求得 $\alpha_a=10(10^{-3}$ mg/cm$^3)$, $\alpha_b=4(10^{-3}$ mg/cm$^3)$.

(2) 极大似然法

设总体 ξ 的概率分布为 $p(x;\lambda)$(当 ξ 为连续型时，$p(x;\lambda)$为 ξ 的分布密度 当 ξ 为离散型时,$p(x;\lambda)$为 ξ 的概率分布,即 $P\{\xi=x\}=p(x;\lambda)$),其中 λ 是未知参数，它在一定范围内取值 $x_1, x_2,\cdots,x$ 是总体的样本观测值.

令

$$L(\lambda)=\prod_{i=1}^{n}p(x_i;\lambda),$$

称 $L(\lambda)$为似然函数极大似然法的基本思想是：在 λ 的取值范围内，挑选使似然函数 $L(\lambda)$取得最大值的 $\hat{\lambda}$ 作为参数 λ 的估计值，由于 $L(\lambda)$与 $\ln L(\lambda)$同时达到最大值,故只需求 $\ln L(\lambda)$的最大值点即可：

$$\ln L(\hat{\lambda})=\max\{\ln L(\lambda)\}.$$

用这种方法求得的 $\hat{\lambda}$ 称为参数 λ 的极大似然估计值.

2.2.2 评价估计量 的优劣标准

在对参数进行估计时，人们总希望估计量 $\hat{\theta}$ 能代表真实参数 θ.根据不同的要求，评价估计量的好坏可以有各种各样的标准.这里只介绍两种最常用的标准.

(1) 无偏估计

根据样本推得的估计值 $\hat{\theta}$ 可能与未知参数的真值 θ 不同,然而,如果有一序列抽样构成各个估计，很合理地会要求这些估计的期望值与未知参数的真值相等它的直观意义是样本估计量的数值在参数的真值附近摆动，而无系统误差

定义 1 如果 $E\hat{\theta}=\theta$ 成立，则称估计 $\hat{\theta}$ 为参数 θ 的无偏估计.

(2)有效估计

对总体的某一参数 θ 的无偏估计量往往不只一个，而且无偏性仅仅表明 $\hat{\theta}$ 所有可能取的值按概率平均等于 θ,可能它取的值大部分与 θ 相差很大. 为保证 $\hat{\theta}$ 的取值能集中于 θ 附近,自然要求 $\hat{\theta}$ 的方差越小越好.

定义 2 设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 都是 θ 的无偏估计,如果 $\hat{\theta}_2$ 的方差小于 $\hat{\theta}_1$ 的方差，则称 $\hat{\theta}_2$ 是比 $\hat{\theta}_1$ 有效的估计量如果在 θ 的一切无偏估计量中，$\hat{\theta}$ 的方差达到最小 ,则称 $\hat{\theta}$ 为θ 的有效估计.

实际上，样本均值是总体期望值的有效估计量.

一个无偏有效估计量取的值是在可能范围内最密集于真值 θ 附近的，也就是说，它以最大的概率保证这估计的观测值在未知参数的真值 θ附近摆动.

§2.3 假 设 检 验

我们常常会假设总体期望值为 μ_0 ,总体方差不大于 σ_0^2 ，总体服从某种分布，上述种种假设是否成立呢？还有某种品种是否比其他品种更优？某种药品是否比其他药品更有效？等等.这些问题需通过概率计算，采用假设检验的方法，才能做出正确的推断

2.3.1 假设检验的基本方法

(1) 提出待检验的假设 H_0.它可能有以下几个来源：① 依据以往的经验或某些实验的结果;② 依据某种理论或某种模型;③ 根据事先所做的某种规定.

(2) 选择检验假设 H_0 的统计量，并确定其分布（其推导过程见参考文献[2]），再根据样本观测值计算出该统计量的值.

(3) 确定拒绝域并作出判断.在给定的检验水平（或显著性水平）α($0<\alpha<1$)下，查所选统计量服从的分布表（一般数理统计书后都附有该表），求出临界值，然后根据小概率事件在一次观测中不可能发生的原理确定检验假设 H_0 的拒绝域并做出判断.

2.3.2 一个正态总体的假设检验

本段假定($X_1, X_2, \cdots, X$) 是来自正态总体 $N(\mu,\sigma^2)$ 中的一个样本. 以下给出实际问题中常用的两种检验假设方法.

(1) 未知方差 σ^2 ,检验假设 $H_0:\mu=\mu_0$(μ_0 为已知)

①提出待检假设 $H_0:\mu=\mu_0$.

②选取样本（ $X_1, X_2, \cdots, X_n$ ）的统计量：

$$T=\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\sim t(n-1).$$

其中 S 是样本标准差，再根据样本观测值计算出统计量 T 的值.

③查表得临界值：$t_\alpha=t_\alpha(n-1)$. 然后根据 $P\{|T|>t_\alpha\}=\alpha$ 下结论：若 $|T|>t_\alpha$,则否定 H_0 否则,一般情况下接受 H_0.

凡涉及到 t 分布的假设检验通常称之为 t 检验.

(2) 未知期望 μ,检验假设 $H_0:\sigma^2\leqslant\sigma_0^2$($\sigma_0^2$ 已知)

①提出待检假设 $H_0:\sigma^2\leqslant\sigma_0^2$.

②选取样本（$X_1,X_2,\cdots,X_n$）的统计量：

$$\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1).$$

再根据样本观测值计算出统计量 χ^2 的值.

③查表得临界值：$\chi_\alpha^2=\chi_\alpha^2(n-1)$. 根据 $P\{\chi^2\geqslant\chi_\alpha^2\}=\alpha$ 下结论：若 $\chi^2\geqslant\chi_\alpha^2$,则否定 H_0 否则,一般情况下接受 H_0.

凡涉及到 χ^2 分布的假设检验通常称之为 χ^2 检验.

例 1 机器包装食盐，假设每袋盐的净重服从正态分布，规定每袋标准重量为 500 g,标准差不能超过 10 g.某天开工后，为检查机器工作是否正常，从装好的食盐中随机抽取 9袋，测其净重（单位:g)为

497, 507, 510, 475, 484, 488, 524, 491, 515

问这天包装机工作是否正常？

解 先检验假设 $H_0:\mu=500$.选取统计量

$$T=\frac{\bar{x}-500}{S/\sqrt{9}},$$

根据样本观测值计算得:$|T|=\left|\frac{499-500}{16.03/\sqrt{9}}\right|=0.187$

取检验水平 $\alpha=0.05$,查表得临界值 $t_{0.05}(9-1)=2.306$.

由于 $|T|=0.187<2.306$,故接受 H_0，即可认为平均每袋食盐净重为 500 g,亦即机器包装没有产生系统误差.

再检验假设 $H_1:\sigma^2\leqslant10^2$. 选取统计量

$$\chi^2=\frac{(9-1)S^2}{10^2},$$

根据样本观测值计算得:$\chi^2=\frac{(9-1)\times16.03^2}{10^2}=20.56$.

查表得临界值 $\chi_{0.05}^2(9-1)=15.5$. 由于 $\chi^2=20.56>15.5$,故拒绝 H_1,可认为其方差超过 10^2. 即包装机虽然没有系统误差，但是不够稳定. 因此认为该

天包装机工作不正常.

2.3.3 两个正态总体的假设检验

样本$(X_1,X_2,\cdots,X_n)$来自正态总体$N(\mu_1,\sigma_1^2)$，样本$(Y_1,Y_2,\cdots,Y_n)$来自正态总体$N(\mu_2,\sigma_2^2)$. 且设$\bar{x}=\frac{1}{m}\sum_{i=1}^{m}x_i$，$S_1^2=\frac{1}{m-1}\sum_{i=1}^{m}(x_i-\bar{x})^2$，$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$，$S_2^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2$.

以下给出实际问题中常用的三种检验假设方法.

(1) 未知期望μ_1,μ_2，检验假设$H_0:\sigma_1^2=\sigma_2^2$

①建立待检假设$H_0:\sigma_1^2=\sigma_2^2$

②若$S_1^2\geqslant S_2^2$，则选取统计量：$F=S_1^2/S_2^2\sim F(m-1,n-1)$，临界值$F_{\alpha/2}=F_{\alpha/2}(m-1,n-1)$；否则选取统计量：$F=S_2^2/S_1^2\sim F(n-1,m-1)$，临界值$F_{\alpha/2}=F_{\alpha/2}(n-1,m-1)$. 再根据样本观测值计算出统计量$F$的值.

③根据$P\{F>F_{\alpha/2}\}=\alpha/2$下结论：若$F>F_{\alpha/2}$，则否定$H_0$否则，一般情况下接受$H_0$.

凡涉及到F分布的假设检验通常称之为F检验.

(2) 未知期望μ_1,μ_2，检验假设$H_0:\sigma_1^2\leqslant\sigma_2^2$

①建立待检假设$H_0:\sigma_1^2\leqslant\sigma_2^2$.

②选取统计量：$F=S_1^2/S_2^2\sim F(m-1,n-1)$，再根据样本观测值计算出统计量$F$的值.

③临界值$F_\alpha=F_\alpha(m-1,n-1)$. 根据$P\{F>F_\alpha\}=\alpha$下结论若$F>F_\alpha$，则否定$H_0$否则，一般情况下接受$H_0$

(3) 未知方差σ_1^2,σ_2^2，但$\sigma_1^2=\sigma_2^2$，检验假设$H_0:\mu_1=\mu_2$

①建立待检假设$H_0:\mu_1=\mu_2$.

②选取统计量：$T=\frac{\bar{x}-\bar{y}}{W}\sim t(m+n-2)$，其中

$$W=\sqrt{\frac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}}\sqrt{\frac{1}{m}+\frac{1}{n}},$$

特别地，当$m=n$时，$W=\sqrt{(S_1^2+S_2^2)/n}$. 再根据样本观测值计算出统计量T的值.

③临界值$t_\alpha=t_\alpha(m+n-2)$. 根据$P\{|T|>t_\alpha\}=\alpha$下结论若$|T|>t_\alpha$，则否定H_0否则，一般情况下接受H_0.

注意作此t检验时，实际问题中的方差一般未知，因此要先检验假设