



普通高等教育“十一五”国家级规划教材  
交通规划理论系列教材

# 交通规划理论与方法 习题集

Exercises Collection  
of Theory and Method in  
Transportation Planning

陆化普 石京 编著

Lu Huapu Shi Jing

清华大学出版社



普通高等教育“十一五”国家级规划教材  
交通规划理论系列教材

# 交通规划理论与方法 习题集

**Exercises Collection  
of Theory and Method in  
Transportation Planning**

清华大学出版社  
北京

## 内 容 简 介

全书共分6章,为启发读者思考和把握基本理论,习题分思考题和计算题两类,具体包括交通调查设计、交通网络的抽象与简化、交通发生与吸引、交通分布、交通分配、预测模型的扩展及交通项目的社会经济效益分析等内容,基本覆盖了交通规划理论与方法的各个领域。

本书可供高等院校交通规划及相关专业的教师、本科生、研究生,以及从事交通规划工作的技术人员参考阅读。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

交通规划理论与方法习题集/陆化普,石京编著. —北京:清华大学出版社,2009.12  
(交通规划理论系列教材)

ISBN 978-7-302-21066-5

I. 交… II. ①陆… ②石… III. 交通规划—高等学校—习题 IV. U491.1-44

中国版本图书馆CIP数据核字(2009)第168151号

责任编辑:汪亚丁 李 嫚

责任校对:刘玉霞

责任印制:李红英

出版发行:清华大学出版社

地 址:北京清华大学学研大厦A座

<http://www.tup.com.cn>

邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质 量 反 馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

印 装 者:北京国马印刷厂

经 销:全国新华书店

开 本:175×245 印 张:13 字 数:258千字

版 次:2009年12月第1版 印 次:2009年12月第1次印刷

印 数:1~3000

定 价:26.00元

---

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:(010)62770177 转 3103 产品编号:030937-01

# 前 言

《交通规划理论与方法习题集》是交通规划系列教材之一。本教材系列包括：《交通规划理论与方法》(第二版)、《交通规划理论与方法习题集》、《城市交通规划案例集》和《交通规划理论研究前沿》。

本习题集是配合《交通规划理论与方法》教材编写的。该教材自出版以来，在清华大学和其他院校的研究生课程教学中已使用多年。

为进一步完善教材建设，结合清华大学精品课建设，作者对使用本书的高等院校部分教授进行了问卷调查。从汇总的修改建议看，最多的建议是增加例题和习题，以方便教师备课和学生自学，以及增加案例分析，使学生能够深刻理解如何应用这些理论解决工程实践问题。参考上述修改建议，结合近年教学的实践体会，作者在对主教材进行修改充实的同时，编写了《交通规划理论与方法习题集》。

本书提供了相当数量的习题，有助于读者加深对教材的理解。本书着重结合我国近年的交通规划实践，习题和数据全部采用了我国城市交通的实际案例和数据，以使學生能够更好地理解理论及其在我国城市实际交通规划中的应用。

本书与主教材《交通规划理论与方法》配套使用，希望能够给使用本教材的师生提供教学上的方便。

作者诚恳地希望广大读者对书中的缺点与不足之处提出批评指正。

作 者

2009年春于清华园

# 目 录

<b>第 1 章 交通调查与分析</b> .....	1
1.1 交通调查方法 .....	1
1.2 交通调查与数据处理 .....	2
1.3 部分思考题与习题解答 .....	5
<b>第 2 章 交通需求预测</b> .....	21
2.1 交通与土地利用.....	21
2.2 交通的发生与吸引.....	21
2.3 交通分布.....	23
2.4 交通方式分担.....	24
2.5 非集计模型的理论与方法.....	29
2.6 部分思考题与习题解答.....	30
<b>第 3 章 交通分配</b> .....	52
3.1 交通网络的抽象与简化.....	52
3.2 最短路径的计算方法.....	54
3.3 流量费用函数与路网容量分析.....	54
3.4 交通量分配方法.....	57
3.5 部分思考题与习题解答.....	62
<b>第 4 章 预测模型的扩展</b> .....	100
4.1 交通分配模型的扩展 .....	100
4.2 组合模型 .....	101
4.3 动态交通分析模型 .....	102
4.4 动态需求快速预测模型 .....	102
4.5 部分思考题与习题解答 .....	103
<b>第 5 章 项目评价</b> .....	111
5.1 评价的基础知识 .....	111

5.2	投资效果分析 .....	112
5.3	社会经济效益评价 .....	113
5.4	综合评价 .....	118
5.5	环境影响评价 .....	120
5.6	部分思考题与习题解答 .....	120
<b>第6章</b>	<b>综合练习</b> .....	<b>156</b>
6.1	思考题 .....	156
6.2	习题 .....	157
6.3	部分思考题与习题解答 .....	172
<b>附录</b>	.....	<b>193</b>
<b>参考文献</b>	.....	<b>200</b>

# 第 1 章 交通调查与分析

## 1.1 交通调查方法

### 思考题

**思考题 1.1** 阐述交通调查的意义并论述交通调查的种类与用途。

**思考题 1.2** 试设计一份居民出行调查表并利用设计的调查表说明设计意图和关键点。假定城市规模为 1000 万人,交通方式有地铁、城市铁路、常规公交、私家车、自行车和步行等。

**思考题 1.3** 假设某人一天的交通出行如图 1-1 所示,图中的①~⑨表示出行的顺序。请分别以出行目的、出发地、目的地、包含的所有交通方式及其代表交通方式几项对其交通出行进行分析。

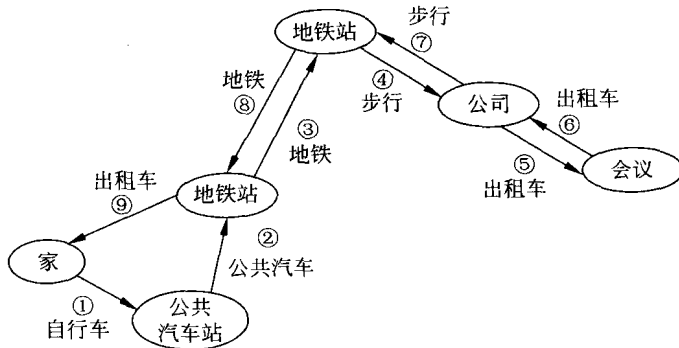


图 1-1 某人一天的交通行为

**思考题 1.4** 东西方向的一条河被当做核查线(screen-line),用于观测从河的南面小区到河的北面小区的所有机动车交通。对由南向北的驾驶员寻找起点和终点,记录为 OD(origin-destination)交通。此时能不能得出正确的表示小区间交通需求量的 OD 矩阵? 如何通过核查线调查检验居民出行调查的精度?

## 1.2 交通调查与数据处理

### 思考题

**思考题 1.5** 对任一城市的居民出行调查数据按个人属性、出行特性等数据进行单独分析(如出行目的构成)和交叉分析(如不同年龄带的出行目的构成、不同交通方式的出行距离、不同出行目的的交通方式构成),并讨论居民出行的主要特征。

### 习题

#### 习题 1.1 相关分析与回归分析

表 1-1 中给出了私人汽车保有量与持有驾驶执照人口的变化情况。

- (1) 判断它们之间的相关关系(计算它们的相关系数,并对其作统计检验)。
- (2) 做线性回归分析。

表 1-1 私人汽车的保有量与驾驶执照持有者数的变化情况

项 目	1989 年	1990 年	1991 年	1992 年	1993 年	1994 年
持有驾照人口/千人	59 159	60 909	62 554	64 172	65 696	67 206
保有量/千辆	28 720	30 625	32 177	33 690	34 974	36 250

#### 习题 1.2 偏相关分析与检验

(偏相关分析是为了去除其他变量对于 2 个变量的影响,找到 2 个变量之间本来的相关关系。通过偏相关分析得到的相关系数称为偏相关系数。)

对私人汽车保有量,持有驾驶执照人口以及道路总长进行相关分析,并对其作统计检验(见表 1-2)。

表 1-2 与私人汽车有关数据的变化情况

项 目	1989 年	1990 年	1991 年	1992 年	1993 年	1994 年
有驾照人口/千人	59 159	60 909	62 554	64 172	65 696	67 206
保有量/千辆	28 720	30 625	32 177	33 690	34 974	36 250
道路总长/ $10^3$ km	1104	1110	1115	1120	1125	1131

#### 习题 1.3 多元回归分析与多元相关系数

表 1-3 中列出了机动车货物周转量,国民生产总值和货车数量。试以机动车货物周转量为目标函数做出回归分析式,并作检验。



表 1-3 机动车货物周转量等的变化情况

项 目	1987 年	1988 年	1989 年	1990 年	1991 年	1992 年
货物周转量 $Y$ /亿 $t \cdot km$	2246	2461	2629	2742	2838	2816
GDP $X_1$ /亿元	350	371	387	407	422	425
货车 $X_2$ /万辆	3745	3886	4097	4273	4439	4572

## 习题 1.4 判别分析

(所谓判别分析是通过将样本所具有的特性进行分析来判别这个样本属于哪个组群的方法。)

表 1-4 列出了利用平行的两种铁路(国有铁路为 1 组,私营铁路为 2 组)的 10 个男性对票价与舒适程度的满意度的调查结果(10 分为满分)。在此调查结果的基础上,判断票价期待度为 3、舒适程度期待度为 6 的男性属于哪一组。判断时使用线性判别函数法。

表 1-4 使用不同交通方式的满意度

样本	票价	舒适程度	组别	样本	票价	舒适程度	组别
1	3	10	1	6	1	4	2
2	5	8	1	7	5	6	2
3	4	7	1	8	3	2	2
4	6	5	1	9	2	1	2
5	2	9	1	10	6	3	2

## 习题 1.5 用聚类(cluster)分析方法对城市的人口规模与交通量进行评价

(聚类分析方法是根据样本或是变量相互近似的程度对其进行分类、分组的方法。)

表 1-5 中表示的是对 5 个城市的人口规模与交通量的评价,请根据表中数据将城市分组。

表 1-5 5 个城市的人口规模与交通量的评价

	1	2	3	4	5
人口规模	5	4	2	1	1
交通量	3	4	4	1	2

## 习题 1.6 用数量化一类方法进行交通量预测

(数量化一类方法是把目标函数与可能对其产生影响的说明变量表现为关系式,并利用该式进行预测的方法,基本上与多元回归分析相同。但是多元回归分析中的说明变量为定量数据,这里使用的是定性数据。)

表 1-6 中所列的是 14 天的白天机动车交通量的调查结果,包括调查的日期以及天气状况。请根据这些数据预测下雨时星期一的交通量。

表 1-6 每日的交通量及其相应的天气情况

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
交通量/千辆	30	77	75	102	81	73	48	74	72	79	96	92	87	55
星期	日	一	二	三	四	五	六	日	一	二	三	四	五	六
天气	晴	晴	晴	雨	晴	晴	云	晴	晴	晴	雨	雨	云	雨

习题 1.7 数量化二类方法

(数量化二类方法是从样本的种种特性中,判别这个样本属于哪一组群的方法,基本上与判别分析相同。但是判别分析中的自变量为定量数据,这里使用的是定性数据。)

表 1-7 中列出了关于拥有 2 辆汽车与不拥有汽车的家庭的年收入、主妇的就业情况、附近公共交通的服务水平等调查结果。请根据调查结果预测收入较多、主妇不工作、附近的公共交通服务状况好的家庭保有 2 辆汽车的可能性。

表 1-7 两种家庭状况抽样调查表

拥有汽车家庭											未拥有汽车家庭												
抽样	家庭收入					主妇就业状况		交通状况				抽样	家庭收入					主妇就业状况		交通状况			
	少	较少	普通	较多	多	有工作	无工作	好	普通	差	少		较少	普通	较多	多	有工作	无工作	好	普通	差		
	11	12	13	14	15	21	22	31	32	33	11		12	13	14	15	21	22	31	32	33		
1		○					○	○			21	○					○	○					
2		○				○			○		22	○					○	○					
3			○				○		○		23	○				○		○					
4			○			○				○	24	○					○	○					
5			○			○			○		25	○					○		○				
6			○			○				○	26	○				○			○				
7			○			○			○		27	○				○				○			
8				○			○			○	28		○				○			○			
9				○			○	○			29		○				○		○				
10				○		○				○	30		○				○		○				
11				○		○				○	31		○				○	○					
12				○			○			○	32		○				○	○					
13				○		○			○		33		○			○		○					
14					○	○			○		34			○			○	○					

续表

拥有汽车家庭											未拥有汽车家庭												
抽样	家庭收入					主妇就业状况		交通状况				抽样	家庭收入					主妇就业状况		交通状况			
	少	较少	普通	较多	多	有工作	无工作	好	普通	差	少		较少	普通	较多	多	有工作	无工作	好	普通	差		
	11	12	13	14	15	21	22	31	32	33	11		12	13	14	15	21	22	31	32	33		
15					○		○			○	35			○			○	○					
16					○		○			○	36			○			○				○		
17					○	○				○	37			○			○		○				
18					○	○				○	38			○			○			○			
19					○	○				○	39				○		○	○					
20					○	○					40				○		○		○				

习题 1.8 主成分分析

(主成分分析是对若干个变量值进行加权合成,做成新的变量(也就是主成分),再把变量汇总的方法。)

表 1-8 中列出了 10 个城市通勤时所使用交通方式的调查结果。请按照这个调查结果,试对这些城市进行分类。

表 1-8 10 个城市通勤时使用不同交通方式的分担率 %

城市	步行	铁路	公共汽车	私人汽车	城市	步行	铁路	公共汽车	私人汽车
1	20.5	5.7	0.8	0.6	6	23.5	7.3	7.5	34.7
2	20.4	12.7	5.9	10.4	7	27.0	28.3	2.0	6.7
3	40.9	30.1	14.0	3.7	8	35.6	27.4	12.5	0.1
4	6.2	37.5	38.6	16.6	9	7.0	52.6	27.8	8.6
5	1.4	11.8	37.2	32.0	10	1.4	6.9	42.8	40.3

1.3 部分思考题与习题解答

思考题 1.3

解答:出行(trip)是计量交通参与者交通行为的单位,是指人为实现某种目的而从特定出发地(起点,origin)到目的地(终点,destination)的单方向的移动。一般来说,交通本身不是本源需求(essential demand),而是为了达到某种目的

而产生的派生需求(derived demand)。这时,为了计量交通出行及预测将来的交通需求,给出行下定义时,最好将产生出行的目的也包括进去。因此,只要交通目的不变,即使改变了交通方式,这个过程也被看做是一个出行。

居民出行调查中的出行目的通常分为6种:①上班(去公司);②上学;③回家;④回公司、学校(返回公司、学校);⑤自由(可细分为日常购物、非日常购物、学习、社交、观光等);⑥业务(可细分为会议、收款、配送、进货、做工、接待、调查等)。

虽然多数的出行采用的交通方式有多种,但是在交通需求预测中有必要将一个出行需求对应一个目的和一种交通方式,因此有了代表交通方式(代表此次出行的交通方式)的概念。在多个交通方式的出行中,决定代表交通方式的方法有最长距离方式法和优先顺序法。最长距离方式法是把该次出行中时间或距离最长交通方式定为代表交通方式的方法;优先顺序法是把所利用的交通方式中公共性最高的方式定为代表交通方式的方法。此时,所采用的公共性顺序可以是:①飞机;②轮船;③铁路;④新交通系统;⑤公共汽车;⑥汽车;⑦摩托车;⑧自行车;⑨步行。在这两种方法中,从交通规划的公共性方面来看,一般采用后者。但是在优先顺序法中利用比公共汽车的公共性低的交通方式的出行有可能被低估,特别是利用自行车或步行的出行往往容易被忽略。所以在做公共汽车线网规划、道路网规划、自行车道规划、人行道规划等专项交通规划时,有必要根据各自情况对出行分别进行求和计算。

末端交通方式分为两种:一种是从出发地到代表交通方式起点的衔接交通方式;另一种是从代表交通方式到目的地的衔接交通方式。

根据以上的出行概念,对图1-1所示的交通行为从起点(家)开始按顺序表示出行目的、出发地、目的地、包含的所有交通方式及其代表交通方式进行分析,结果如下:

- 第1次出行:上班目的、家、公司、①②③④、地铁;
- 第2次出行:会议目的、公司、会议场所、⑤、出租车;
- 第3次出行:回公司、会议场所、公司、⑥、出租车;
- 第4次出行:回家目的、公司、家、⑦⑧⑨、地铁。

#### 习题 1.1 相关分析与回归分析

解答:(1)可利用下式计算私人汽车保有量与持有驾驶执照人口的相关系数  $r$ :

$$r = \frac{\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i / n}{\sqrt{[\sum_i X_i^2 - (\sum_i X_i)^2 / n][\sum_i Y_i^2 - (\sum_i Y_i)^2 / n]}}$$

对照表1-1,并分别对应各年份的持有驾照人口和机动车保有量,列出如表1-9所示的计算过程。

表 1-9 计算过程(可利用 Excel 计算)

	$X_i$	$Y_i$	$X_i^2$	$Y_i^2$	$X_i Y_i$
1	59 159	28 720	3 499 787 281	824 838 400	1 699 046 480
2	60 909	30 625	3 709 906 281	937 890 625	1 865 338 125
3	62 554	32 177	3 913 002 916	1 035 359 329	2 012 800 058
4	64 172	33 690	4 118 045 584	1 135 016 100	2 161 954 680
5	65 696	34 974	4 315 964 416	1 223 180 676	2 297 651 904
6	67 206	36 250	4 516 646 436	1 314 062 500	2 436 217 500
总和	379 696	196 436	4 073 352 914	6 470 347 630	12 473 008 747

相关系数为

$$r = \frac{12\,473\,008\,747 - 379\,696 \times 196\,436/6}{\sqrt{(24\,073\,352\,914 - 379\,696^2/6)(6\,470\,347\,630 - 196\,436^2/6)}}$$

$$= 0.998\,846$$

为了用相关系数来判断两者之间是否有相关关系,必须对相关系数的显著性作  $t$  检验。检验时可用下式计算  $t$  值,然后与检验表进行比较。

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2}$$

$$t = \frac{0.998\,846}{\sqrt{1-0.998\,846^2}} \sqrt{6-2} = 41.594\,504$$

假设显著性水平为  $\alpha=0.01$  时,从  $t$  分布表中可以得到

$$41.594\,504 > t(n-2, \alpha/2) = t(6-2, 0.005) = 5.5976$$

所以可以证明它们之间有相关关系。

(2) 线性回归公式的系数可用下式计算:

$$b = \frac{n \sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n \sum_i X_i^2 - (\sum_i X_i)^2}$$

$$= \frac{6 \times 12\,473\,008\,747 - 379\,696 \times 196\,436}{6 \times 24\,073\,352\,914 - 379\,696^2} = 0.929\,994\,5$$

$$a = (\sum_i Y_i - b \sum_i X_i)/n = (196\,436 - 0.929\,994\,5 \times 379\,696)/6$$

$$= -26\,113.2$$

因此回归公式如下:

$$\hat{Y} = -26\,113.2 + 0.929\,994\,5X$$

式中,  $\hat{Y}$  表示用回归公式得到的估计值。

## 习题 1.2 偏相关分析与检验

解答：偏相关系数用下式计算：

$$r_{xy,z} = \frac{r_{xy} - r_{xy}r_{yz}}{\sqrt{(1-r_{xz}^2)(1-r_{yz}^2)}}$$

首先计算驾照持有量( $X$ )，保有量( $Y$ )，道路总长( $Z$ )之间的单相关系数  $r_{xy}, r_{yz}, r_{xz}$ ，其计算公式和习题 1-1 中的相同，计算结果为

$$r_{yz} = 0.997\ 086 [t = 26.142\ 65 > t(n-2, \alpha/2) = t(6-2, 0.005) = 5.5976]$$

$$r_{xz} = 0.999\ 145 [t = 48.345\ 833 > t(n-2, \alpha/2) = t(6-2, 0.005) = 5.5976]$$

$$r_{xy} = 0.998\ 846$$

可以得到去除了道路总长( $Z$ )影响的汽车保有量( $Y$ )与驾照持有量( $X$ )之间的偏相关系数

$$r_{xy,z} = \frac{0.998\ 846 - 0.999\ 145 \times 0.997\ 086}{\sqrt{1 - 0.828\ 341^2}} = 0.828\ 341$$

对其进行  $t$  检验：

$$t = \frac{r_{xy,z}}{\sqrt{1-r_{xy,z}^2}} \sqrt{n-3} = \frac{0.828\ 341}{\sqrt{1-0.828\ 341^2}} \times \sqrt{6-3} = 2.560\ 99$$

$$> t(n-3, \alpha/2) = t(6-3, 0.05) = 2.353$$

假设显著性水平为 0.1 时，可以认为汽车保有量( $Y$ )与驾照持有量( $X$ )之间存在相关关系。

注：当变量为  $m$  组时的偏相关系数也可以通过单相关行列式来求出。

$$\begin{aligned} r_{12} &= (-1)^{1+2} \begin{vmatrix} 0.998\ 846 & 0.997\ 086 \\ 0.999\ 145 & 1.000\ 000 \end{vmatrix} \\ &= -(0.998\ 846 - 0.997\ 086 \times 0.999\ 145) = -0.002\ 612\ 5 \end{aligned}$$

$$\begin{aligned} r_{11} &= (-1)^{1+1} \begin{vmatrix} 1.000\ 000 & 0.997\ 086 \\ 0.997\ 086 & 1.000\ 000 \end{vmatrix} \\ &= 1.000\ 000^2 - 0.997\ 086^2 = 0.005\ 819\ 5 \end{aligned}$$

$$\begin{aligned} r_{22} &= (-1)^{2+2} \begin{vmatrix} 1.000\ 000 & 0.999\ 145 \\ 0.999\ 145 & 1.000\ 000 \end{vmatrix} \\ &= 1.000\ 000^2 - 0.999\ 145^2 = 0.001\ 709\ 3 \end{aligned}$$

则偏相关系数

$$r_{12,0} = \frac{-r_{12}}{\sqrt{r_{11} \times r_{22}}} = \frac{-(-0.002\ 612\ 5)}{\sqrt{0.005\ 819\ 5 \times 0.001\ 709\ 3}} = 0.828\ 341$$

## 习题 1.3 多元回归分析与多元相关系数

解答：首先列出计算表格，如表 1-10 所示。

表 1-10 计算过程(可用 Excel 计算)

年份	$Y_i$	$X_{1i}$	$X_{2i}$	$X_{1i}^2$	$X_{2i}^2$	$X_{1i}X_{2i}$	$X_{1i}Y_i$	$X_{2i}Y_i$
1987	2264	350	3745	122 500	14 025 025	1 310 750	792 400	8 478 680
1988	2461	371	3886	137 641	15 100 996	1 441 706	913 031	9 563 446
1989	2629	387	4097	149 769	16 785 409	1 585 539	1 017 423	10 771 013
1990	2742	407	4273	165 649	18 258 529	1 739 111	1 115 994	11 716 566
1991	2838	422	4439	178 084	19 704 721	1 873 258	1 197 636	12 597 882
1992	2816	425	4572	180 625	20 903 184	1 943 100	1 196 800	12 874 752
总和	15 750	2362	25 012	934 268	104 777 864	9 893 464	6 233 284	66 002 339

根据表 1-10 计算常数项与回归系数

$$\begin{aligned}
 b &= \frac{\sum_i X_{1i} Y_i \sum_i X_{2i}^2 - \sum_i X_{2i} Y_i \sum_i X_{1i} X_{2i}}{\sum_i X_{1i}^2 \sum_i X_{2i}^2 - (\sum_i X_{1i} X_{2i})^2} \\
 &= \frac{104\,777\,864 \times 6\,233\,284 - 9\,893\,464 \times 66\,002\,339}{934\,268 \times 104\,777\,864 - 9\,893\,464^2} = 11.8709 \\
 c &= \frac{\sum_i X_{2i} Y_i \sum_i X_{1i}^2 - \sum_i X_{1i} Y_i \sum_i X_{1i} X_{2i}}{\sum_i X_{1i}^2 \sum_i X_{2i}^2 - (\sum_i X_{1i} X_{2i})^2} \\
 &= \frac{934\,268 \times 66\,002\,339 - 9\,893\,464 \times 6\,233\,284}{934\,268 \times 104\,777\,864 - 9\,893\,464^2} = -0.49096 \\
 a &= \frac{\sum_i Y_i - b \sum_i X_{1i} - c \sum_i X_{2i}}{n} \\
 &= \frac{15\,750}{6} - \frac{11.8709 \times 2362}{6} - \frac{(-0.49096) \times 25\,012}{6} = -1.51939
 \end{aligned}$$

多元回归式为

$$\hat{Y} = -1.51939 + 11.8709X_1 - 0.49096X_2$$

决定系数(determination coefficient)

$$r^2 = 1 - \frac{\sum_i (Y_i - \hat{Y}_i)^2}{\sum_i (Y_i - \bar{Y})^2} = 1 - \frac{6478.349}{252\,772} = 0.974371$$

式中,  $\sum_i (Y_i - \hat{Y}_i)^2$  为残差平方和(sum of squares for residuals);  $\sum_i (Y_i - \bar{Y})^2$  为偏差平方和。

多元相关系数为

$$r = \sqrt{r^2} = \sqrt{0.974371} = 0.987102$$

重回归分析中当样本数  $n$  较少时,通常采用自由度修正后的决定系数,式中  $i$  为自变量的个数:

$$r^2 = 1 - \frac{\sum_i (Y_i - \hat{Y}_i)^2 / (n - p - 1)}{\sum_i (Y_i - \bar{Y})^2 / (n - 1)} = 1 - \frac{6478.349 / (6 - 2 - 1)}{252772 / (6 - 1)} = 0.957285$$

对于多元回归式的检验需要用  $F$  检验:

$$F = \frac{\sum_i (\hat{Y}_i - \bar{Y})^2 / p}{\sum_i (Y_i - \hat{Y}_i)^2 / (n - p - 1)} = \frac{252772 / 2}{6478.349 / (6 - 2 - 1)} = 58.5269$$

$$> f(2, 6 - 2 - 1, \alpha) = f(2, 3, 0.01) = 30.817$$

所以在显著性水平  $\alpha = 0.01$  的情况下, 它们之间有相关关系。

#### 习题 1.4 判别分析

解答: 当从样本所具有的各种特性来判断样本属于哪个组群时, 可以使用判别分析法。判别分析所使用的线性判别函数如下:

$$Z = a_0 + a_1 X_1 + a_2 X_2 + \cdots + a_m X_m$$

当样本的属性只有  $X_1, X_2$  两个, 分别为票价和舒适度的满意度, 对应系数  $a_0, a_1, a_2$  的计算公式为

$$a_1 \sum_i (X_{1i} - \bar{X}_1)^2 + a_2 \sum_i (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) = \bar{X}_{11} - \bar{X}_{12}$$

$$a_1 \sum_i (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) + a_2 \sum_i (X_{2i} - \bar{X}_2)^2 = \bar{X}_{21} - \bar{X}_{22}$$

$$a_0 = -(a_1 \bar{X}_1 + a_2 \bar{X}_2)$$

$\bar{X}_1$  和  $\bar{X}_2$  分别为票价和舒适度满意度的均值:

$$\bar{X}_1 = (3 + 5 + 4 + 6 + 2 + 1 + 5 + 3 + 2 + 6) / 10 = 3.7$$

$$\bar{X}_2 = (10 + 8 + 7 + 5 + 9 + 4 + 6 + 2 + 1 + 3) / 10 = 5.5$$

$\bar{X}_{11}$  与  $\bar{X}_{12}$  和  $\bar{X}_{21}$  与  $\bar{X}_{22}$  分别为第一和二组的票价与舒适程度满意度的均值

$$\bar{X}_{11} = (3 + 5 + 4 + 6 + 2) / 5 = 4.0$$

$$\bar{X}_{12} = (1 + 5 + 3 + 2 + 6) / 5 = 3.4$$

$$\bar{X}_{21} = (10 + 8 + 7 + 5 + 9) / 5 = 7.8$$

$$\bar{X}_{22} = (4 + 6 + 2 + 1 + 3) / 5 = 3.2$$

代入第一个联立方程组, 可以得到判别函数式的系数

$$a_1 = 0.0164, \quad a_2 = 0.05526, \quad a_0 = -0.3647$$

由此可以得到线性判别函数为

$$Z = -0.3647 + 0.0164X_1 + 0.05526X_2$$

因此判断票价满意度为 3、舒适程度满意度为 6 的男性, 由于其  $Z = 0.0164 > 0$ , 故属于第一组国有铁路。(注: 当  $Z < 0$  时, 属于第二组; 当  $Z = 0$  时, 无法判别。)

#### 习题 1.5 用聚类(cluster)分析方法对城市的人口规模与交通量进行评价

解答: 考虑人口规模和交通量的双重影响, 用下面的公式计算城市  $i$  和城



市  $j$  的虚拟距离  $r_{ij}$  :

$$r_{ij} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}$$

其中  $X_i, X_j$  分别为城市  $i$  和城市  $j$  的人口规模;  $Y_i, Y_j$  分别为城市  $i$  和城市  $j$  的交通量。计算结果列在表 1-11 中, 如

$$r_{12} = \sqrt{(5-4)^2 + (3-4)^2} = 1.4142$$

表 1-11 距离矩阵表

	①	②	③	④	⑤
①		1.4142	3.1623	4.4721	4.1231
②			2.0000	4.2426	3.6056
③				3.1623	2.2361
④					1.0000

分组的步骤为:

首先城市④,⑤的距离为 1.0000 最为接近,将其划分为一组;

其次,①,②的距离为 1.4142 第二接近,划归为另一组;

接下来城市③与①,②的距离为 2.0000,可划分为一组;

最后剩下{④,⑤}与{③与①,②}的组合;

分组情况如图 1-2 所示。如需分为两组的话,在虚线的地方分类即可。

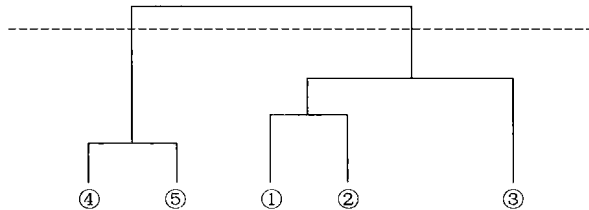


图 1-2 分组情况

习题 1.6 用数量化一类方法进行交通量预测

解答: 首先将项目的交通量与出现的次数进行汇总, 做成表 1-12。

表 1-12 交通量汇总结果(1)

	星 期							天 气		
	日	一	二	三	四	五	六	晴	云	雨
交通量	104	149	154	198	173	160	103	561	135	345
出现次数	2	2	2	2	2	2	2	8	2	4