普通高中课程标准实验教科书

# 数学

选修 4–9

## 风险与决策

人民教育出版社 课程教材研究所
中学数学课程教材研究开发中心 编著

人民教育出版社

A版

普通高中课程标准实验教科书 数学 选修 4-9 A版 风险与决策
人民教育出版社 课程教材研究所
中学数学课程教材研究开发中心 编著

出版发行 人民教育出版社
（北京市海淀区中关村南大街17号院1号楼 邮编：100081）
网 址 http://www.pep.com.cn
经 销 全国新华书店
印 刷 北京天宇星印刷厂
版 次 2007 年 8 月第 1 版
印 次 2019 年 8 月第 16 次印刷
开 本 890 毫米 × 1240 毫米 1/16
印 张 3.75
字 数 72 千字
书 号 ISBN 978-7-107-20621-4
定 价 3.85 元
价格依据文件号：京发改规［2016］13号



绿色印刷 保护环境 爱护健康

亲爱的同学们：

你们手中的这本教科书采用绿色印刷标准印制，在它的封底印有“绿色印刷产品”标志。从2013年秋季学期起，北京地区出版并使用的义务教育阶段中小学教科书全部采用绿色印刷。

按照国家环境标准（HJ2503-2011）《环境标志产品技术要求 印刷 第一部分：平版印刷》，绿色印刷选用环保型纸张、油墨、胶水等原辅材料，生产过程注重节能减排，印刷产品符合人体健康要求。

让我们携起手来，支持绿色印刷，选择绿色印刷产品，共同关爱环境，一起健康成长！

北京市绿色印刷工程

主　　编：刘绍学
副 主 编：钱珮玲　章建跃

主要作者：李　勇　张淑梅　宋莉莉　张唯一
责任编辑：宋莉莉　张唯一
美术编辑：王　艾
封面设计：李宏庆

# 主 编 寄 语

同学们，欢迎大家使用这套普通高中数学教科书，希望它能够成为你们学习数学的好朋友.

作为这套教科书的主编，在大家开始用这套书学习数学之前，对于为什么要学数学，如何才能学好数学等问题，我有一些想法与你们交流.

为什么要学数学呢？我想从以下两个方面谈谈认识.

**数学是有用的.** 在生活、生产、科学和技术中，在这套教科书中，我们都会看到数学的许多应用. 实际上，“数量关系与空间形式”，在实践中，在理论中，在物质世界中，在精神世界中，处处都有，因而研究“数量关系与空间形式”的数学，处处都有用场. 数学就在我们身边，她是科学的语言，是一切科学和技术的基础，是我们思考和解决问题的工具.

**学数学能提高能力.** 大家都觉得，数学学得好的人也容易学好其他理论. 实际上，理论之间往往有彼此相通和共同的东西，而“数量关系与空间形式”、逻辑结构及探索思维等正是它们的支架或脉络，因而数学恰在它们的核心处. 这样，在数学中得到的训练和修养会很好地帮助我们学习其他理论，数学素质的提高对于个人能力的发展至关重要.

那么，如何才能学好数学呢？我想首先应当对数学有一个正确的认识.

**数学是自然的.** 在这套教科书中出现的数学内容，是在人类长期的实践中经过千锤百炼的数学精华和基础，其中的数学概念、数学方法与数学思想的起源与发展都是自然的. 如果有人感到某个概念不自然，是强加于人的，那么只要想一下它的背景，它的形成过程，它的应用，以及它与其他概念的联系，你就会发现它实际上是水到渠成、浑然天成的产物，不仅合情合理，甚至很有人情味. 这将有助于大家的学习.

**数学是清楚的.** 清楚的前提，清楚的推理，得出清楚的结论，数学中的命题，对就是对，错就是错，不存在丝毫的含糊. 我们说，数学是易学的，因为它是清楚的，只要大家按照数学规则，按部就班地学，循序渐进地想，绝对可以学懂；我们又说，数学是难学的，

也因为它是清楚的，如果有人不是按照数学规则去学去想，总想把“想当然”的东西强加给数学，在没有学会加法的时候就想学习乘法，那就要处处碰壁，学不下去了.

在对数学有一个正确认识的基础上，还需要讲究一点方法.

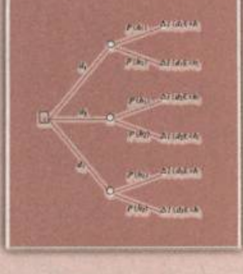
**学数学要摸索自己的学习方法.** 学习、掌握并能灵活应用数学的途径有千万条，每个人都可以有与众不同的数学学习方法. 做习题、用数学解决各种问题是必需的，理解概念、学会证明、领会思想、掌握方法也是必需的，还要充分发挥问题的作用，问题使我们的学习更主动、更生动、更富探索性. 要善于提问，学会提问，“凡事问个为什么”，用自己的问题和别人的问题带动自己的学习. 在这套书中，我们一有机会就提问题，希望“看过问题三百个，不会解题也会问”. 类比地学、联系地学，既要从一般概念中看到它的具体背景，不使概念“空洞”，又要在具体例子中想到它蕴含的一般概念，以使事物有“灵魂”.

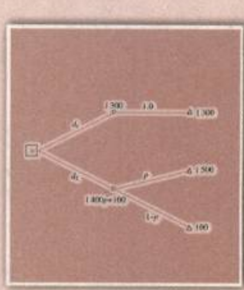
**学数学趁年轻.** 同学们，你们正处在一生中接受数学训练、打好数学基础的最佳时期. 这个时期下点功夫学数学，将会终生受益. 我们构建了这片数学天地，期盼它有益于大家的成长. 你们是这片天地的主人，希望大家在学习的过程中能对它提出宝贵的改进意见. 预祝同学们愉快地生活在这片数学天地中.

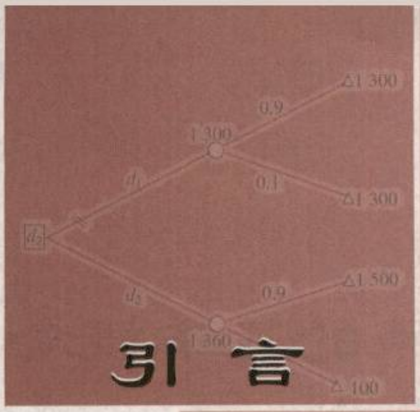
# 目 录











# 引言

在现实生活中，几乎每个人每天都要做出各种各样的决策，如决定早餐吃什么，出门是否带雨伞，上学或上班走哪条路线，出外旅行是否预定火车票，是否购买航空意外险，报考国内大学还是申请国外大学，等等．一个企业也面临着各种类型的决策，如投资、经营、销售、人员聘用等．在医学、军事、工程和行政领域，决策问题都广泛存在．

通常情况下，决策者有很多行动方案可以选择．但由于未来事态的发展往往受随机因素的影响，不能确切预料，因此决策是带有风险的．为了在众多的行动方案中选出最优方案，减少不合理决策造成的损失，应该按照科学的方法进行决策．借助于统计分析方法进行决策，可以大大减少由于随机因素造成的损失，提高决策的有效性．因此，统计决策方法和统计决策分析将会在社会的发展和进步中发挥越来越大的作用．本专题将介绍一些简单的统计决策方面的基本知识和方法．

本专题共分四讲．在第一讲中，首先，同学们将通过现实生活中的例子，理解“风险”与“决策”之间的关系，以及通常情况下对一个问题做出决策的过程；然后，同学们将通过具体例子，理解风险型决策问题的特点，以及解决这类问题的常用方法——风险型决策的基本概念和决策过程．在第二讲中，同学们将学习用决策树表示风险型决策问题的相关信息，以及通过决策树进行决策的方法．在第三讲中，通过敏感性分析，同学们可以清楚地了解状态发生概率的变化对最优决策产生的影响．在第四讲中，同学们将通过实例，了解马尔可夫型决策模型及其决策方法．

| $h$ | $h_1$ | $h_2$ | $h_3$ |
|---|---|---|---|
| $P(h)$ | 0.3 | 0.5 | 0.2 |

# 第一讲 风险与决策的基本概念

## 一 风险与决策的关系

在日常生活和生产实践中，经常需要我们根据一定的条件做出某种决策.

**问题 1** 天气预报说“今天的降水概率是55%”. 你今天上学时，是否会带雨具？

有的同学可能会这样想：“降水概率为55%，降水的可能性比不降水的可能性大，所以应该带上雨具.”也有的同学可能会这样想：“虽然降水的可能性比不降水的可能性大，但还有45%的可能不下雨，况且这只是一个预报，不一定准确，所以没有必要带雨具.”前一种决策，因为存在不下雨的可能性，所以要冒“带了不用”的风险；后一种决策，因为存在下雨的可能性，所以要冒“挨雨淋”的风险.

类似的问题非常多，我们再看两个例子.

**问题 2** 在十一黄金周期间，许多人都会乘火车出外旅行，所以人们往往面临是否提前预定火车票的问题. 如果提前预定火车票，可以顺利成行，但需支付40元的定票费；如果不提前预定火车票，在临时能买到火车票的情况下，能顺利成行，否则就要被迫取消旅行. 如果你正面临这个问题，你会做出什么决策呢？其根据是什么？

你可能会根据自己的实际情况做出不同的决策——提前预定火车票或不提前预定火车票. 但你是否想到：当你决定提前预定火车票时，如果临时能买到火车票，那你就损失了40元的订票费；当你决定不提前预定火车票时，就要冒因买不到火车票而被迫取消旅行的风险. 可见，无论哪一种决策，都要冒一定的风险.

**问题 3** A，B两个大小相同的箱子中装有同样规格的产品，且产品的价格相同. 已知A箱中产品的废品率是1%，B箱中产品的废品率是3%. 若你想购买1件产品，且购买时不能打开产品包装，你会选择哪个箱中的产品？

对于这个问题，通常情况下，我们会选择购买A箱中的产品. 这是因为A，B两个箱

中装有同样数量的产品，而 A 箱中产品的废品率低于 B 箱中产品的废品率，所以从 A 箱中买到废品的概率也低于从 B 箱中买到废品的概率. 但是，这种选择也是带有风险的，因为我们从 A 箱中买到的产品也可能是废品.

上述三个问题有一个共同的特征：需要在多个可供选择的决策中，选择一个决策；并且任何决策所面临的未来情况都具有不确定性. 虽然可以通过一定的方法了解未来各种情况发生的可能性的大小，但因为它们具有随机性，所以无论我们做出什么决策，都要冒一定的风险.

下面是一个利用概率统计知识进行决策的简单案例.

**案例 1** 某同学需要做一项实验，在该项实验过程中出现设备过热现象的概率为 0.2. 请专家指导实验，需支付 50 元指导费，并且出现设备过热现象时会损失 10 元；自己独立完成实验，出现设备过热现象时会损失 100 元. 如果在实验过程中没有出现过热现象，则不会造成损失. 问该同学是否应该请专家指导实验.

这个决策问题的目标是使损失最小. 该同学有两种可以采取的行动方案：

行动一　请专家指导实验；

行动二　独立完成实验.

需要决策的是采取哪种行动方案.

在做出决定之前，我们先分析一下这两种行动方案导致损失的情况.

1. 采取"行动一"，即请专家指导实验：没有出现过热现象的损失为请专家的费用 50 元；出现过热现象的损失为 $50+10=60$(元). 由于是否出现过热现象是一个随机现象，因此采取行动方案一所造成的损失是一个随机变量，其平均损失为

$$0.8\times50+0.2\times60=52(\text{元}).$$

2. 采取"行动二"，即独立完成实验：没有出现过热现象的损失为 0 元；出现过热现象的损失为 100 元. 类似地，采取行动二的平均损失为

$$0.8\times0+0.2\times100=20(\text{元}).$$

由上述分析可知，应该选择行动方案二，即该同学独立完成实验.

分析案例 1 的决策过程，可以看出其中包括下述基本步骤：

> 第一步，明确问题的决策目标；
> 第二步，寻找所有的行动方案；
> 第三步，分析每种行动方案的平均损失；
> 第四步，按照平均损失最小的准则，选择最优决策.

从上面的案例中可以看出，平均损失与决策是密切相关的. 通过对它的分析计算，使我们能根据量化指标进行合理、有效的决策，从而尽量地化解风险，提高效益.

## 习题1.1

1. 中国历史上有许多杰出的思想家、政治家、军事家、外交家，他们博学多才，高瞻远瞩，运筹帷幄之中，决胜千里之外，为我国的决策科学谱写了光辉篇章. 请举出几个实例.
2. 在一个暖和、潮湿的早晨，你离家骑车去学校，路上大约需要 20 min. 你可以采取的行动有两个：带上雨衣和不带雨具. 分析采取每种行动后可能出现的结果，你更希望哪一种结果出现？什么对你的决策影响最大？
3. 甲、乙、丙 3 人参加一场趣味射击比赛. 比赛设置了 3 块目标靶，分别是 1 号靶、2 号靶和 3 号靶. 比赛规定：如果 1 号靶被击中，则甲被淘汰出局；如果 2 号靶被击中，则乙被淘汰出局；如果 3 号靶被击中，则丙被淘汰出局；最后一位未被淘汰者赢得比赛. 3 个人开枪射击的顺序是甲、乙、丙、甲、乙、丙……每人一次只能打一枪；被淘汰出局者不能开枪射击，直接把枪交给下一位选手. 由甲、乙、丙以往的射击记录可知，甲击中目标的概率为$\frac{1}{3}$，乙击中目标的概率为$\frac{2}{3}$，丙百发百中. 那么，甲第 1 枪应该如何射击，才能保证自己在第 2 枪后未被淘汰出局的可能性最大？

## 二　风险与决策的基本概念

### 1. 风险（平均损失）

怎样度量风险的大小呢？下面介绍数学中常用的方法. 为此，我们以案例 1 为背景引入几个基本概念.

在案例 1 中，用 $d_1$ 表示“请专家指导实验”，$d_2$ 表示“独立完成实验”，则可供选择的行动方案仅有 $d_1$ 和 $d_2$. 无论该同学采取哪一行动方案，他都将面临两个可能的结果：设备没有出现过热现象，或设备出现过热现象. 我们把它们看成两个状态，用 $h_1$ 表示“设备没有出现过热现象”，$h_2$ 表示“设备出现过热现象”.

一般地，对于给定的行动方案 $d$ 和状态 $h$，用 $l(d, h)$表示行动方案 $d$ 在状态 $h$ 下的损失大小，并称 $l(d, h)$为**损失函数**.

在案例 1 中，损失函数在不同行动方案和不同状态下的取值可用表 1-1 表示：

表 1-1

| 状态 / 损失 / 行动方案 | $h_1$ | $h_2$ |
|---|---|---|
| $d_1$ | 50 | 60 |
| $d_2$ | 0 | 100 |

即

$$l(d_1, h_1)=50, \ l(d_1, h_2)=60;$$
$$l(d_2, h_1)=0, \ l(d_2, h_2)=100.$$

在通常情况下，对于给定的一个行动方案 $d$，未来的状态具有随机性，因此 $l(d, h)$ 为随机变量. 那么，如何比较各行动方案所造成损失的大小呢？

可以通过随机变量 $l(d, h)$ 的均值来衡量行动方案 $d$ 的平均损失大小. 因此，我们可以通过比较各行动方案 $d$ 所造成的平均损失的大小来作出决策.

为了计算行动方案 $d$ 的平均损失，需要知道各种状态发生的概率. 对于给定的行动方案 $d$，未来各状态发生的概率可以由历史资料的统计分析、过去的经验或人们的主观判断获得.

在案例 1 中，出现状态 $h_1$ 和 $h_2$ 的概率分别是 0.8 和 0.2，用列表的方式可以表示为：

| $h$ | $h_1$ | $h_2$ |
|---|---|---|
| $P(h)$ | 0.8 | 0.2 |

一般地，我们把各个状态出现的概率称为**状态分布列**，并用列表的方式表示.

在案例 1 中，根据状态分布列，行动方案 $d_1$ 和 $d_2$ 对应的损失函数的均值分别为

$$\begin{aligned}E(l(d_1, h))&=l(d_1, h_1)\times P(h_1)+l(d_1, h_2)\times P(h_2)\\&=50\times 0.8+60\times 0.2=52,\end{aligned}$$
$$\begin{aligned}E(l(d_2, h))&=l(d_2, h_1)\times P(h_1)+l(d_2, h_2)\times P(h_2)\\&=0\times 0.8+100\times 0.2=20.\end{aligned}$$

一般地，我们用 $R(d)$ 表示行动方案 $d$ 所对应损失函数的均值，并且称 $R(d)$ 为行动方案 $d$ 的**风险**（**平均损失**）. 显然，我们应该选用风险最小的行动方案，即按照风险最小准则选择行动方案.

由上面的计算可知，$R(d_1)=52$，$R(d_2)=20$. 由于 $R(d_1)>R(d_2)$，即行动方案 $d_2$ 所造成的风险较小，因此选择行动方案 $d_2$，即“独立完成实验”.

## 2. 平均收益

案例 1 是按照风险（平均损失）最小准则选择最优决策的，我们还可以按照平均收益最大准则选择最优决策.

**案例 2** 某位农民打算种植新品种蔬菜，可选择的种植量有 3 种：大量、适量、少量. 他应当如何决策呢？首先，应当做市场预测. 根据收集到的市场信息，可知未来市场出现好、中、差 3 种情况的概率分别为 0.3，0.5，0.2. 然后，这位农民根据过去的经验，得到一个收入表（单位：千元），如表 1-2 所示：

**表 1-2**

| 种植方案 \ 收入 \ 市场情况 | 好 | 中 | 差 |
|---|---|---|---|
| 大量 | 12 | 5 | −4 |
| 适量 | 8 | 6 | −1 |
| 少量 | 3 | 3 | 2 |

表中的“12”表示，在市场“好”的情况下，如果采取了“大量”种植的方案，可以收入 12 000 元，表中其他数字的含义依此类推．如果你是这位农民，你会选择哪种种植方案呢？

显然，这位农民的决策目标是使种植获得最大的收益．他所能采取的全部行动方案包括：

$d_1$：大量种植，$d_2$：适量种植，$d_3$：少量种植.

未来市场的状态包括：

$h_1$：好，$h_2$：中，$h_3$：差.

由于出现 $h_1$，$h_2$，$h_3$ 的概率分别为 0.3，0.5，0.2，所以 $h$ 的分布列为

| $h$ | $h_1$ | $h_2$ | $h_3$ |
|---|---|---|---|
| $P(h)$ | 0.3 | 0.5 | 0.2 |

为了表示收益的大小，我们可以从收益的角度定义一个函数，表示不同行动方案 $d$ 在不同状态 $h$ 下的收益的大小，这样的函数称为**收益函数**，用 $q(d, h)$ 表示．损失函数和收益函数统称为**损益函数**.

由表 1-2 可知，本例的收益函数可以定义为：

$$q(d_1, h_1)=12,\ q(d_1, h_2)=5,\ q(d_1, h_3)=-4;$$
$$q(d_2, h_1)=8,\ q(d_2, h_2)=6,\ q(d_2, h_3)=-1;$$
$$q(d_3, h_1)=3,\ q(d_3, h_2)=3,\ q(d_3, h_3)=2.$$

因此，行动方案 $d_1$，$d_2$，$d_3$ 所对应的收益函数的均值分别为

$$\begin{aligned}Q(d_1)&=q(d_1, h_1)P(h_1)+q(d_1, h_2)P(h_2)+q(d_1, h_3)P(h_3)\\&=12\times0.3+5\times0.5+(-4)\times0.2\\&=5.3;\end{aligned}$$

$$\begin{aligned}Q(d_2)&=q(d_2, h_1)P(h_1)+q(d_2, h_2)P(h_2)+q(d_2, h_3)P(h_3)\\&=8\times0.3+6\times0.5+(-1)\times0.2\\&=5.2;\end{aligned}$$

$$\begin{aligned}Q(d_3)&=q(d_3, h_1)P(h_1)+q(d_3, h_2)P(h_2)+q(d_3, h_3)P(h_3)\\&=3\times0.3+3\times0.5+2\times0.2\\&=2.8.\end{aligned}$$

由于 $Q(d_1)>Q(d_2)>Q(d_3)$，故行动方案 $d_1$ 所带来的平均收益最大，所以应该选择 $d_1$，即“大量种植”这种蔬菜.

**思考**

若执行行动方案 $d_1$，一定能获得最大收益吗？

上述结论是通过比较各行动方案所带来的“平均收益”得到的，这里“平均收益”的含义是如果种植行动多次发生，那么每次都选择行动方案 $d_1$ 所带来的平均收益是最大的，但某一次采取行动方案 $d_1$ 的收益不一定是最大的. 例如，若某年的市场情况为“差”，则采取行动方案 $d_1$ 的收益就是最小的.

**思考**

若从损失的角度考虑，本例的损失函数应如何定义？若按照风险最小准则做出决策，与上述决策结果相同吗？

显然，行动方案 $d_1$，$d_2$ 和 $d_3$ 所造成的损失与它们在同一状态下的收入差异有关. 下面我们根据表 1-2 来构造损失函数. 当 $h_1$ 发生时，选择行动方案 $d_1$ 没有损失，选择行动方案 $d_2$ 的损失为 $12-8=4$（元），选择行动方案 $d_3$ 的损失为 $12-3=9$（元）；当 $h_2$ 发生时，选择行动方案 $d_1$ 的损失为 $6-5=1$（元），选择行动方案 $d_2$ 没有损失，选择行动方案 $d_3$ 的损失为 $6-3=3$（元）；当 $h_3$ 发生时，选择行动方案 $d_1$ 的损失为 $2-(-4)=6$（元），选择行动方案 $d_2$ 的损失为 $2-(-1)=3$（元），选择行动方案 $d_3$ 没有损失. 所以，损失函数 $l(d,h)$ 的全部取值为：

$$l(d_1,h_1)=0,\ l(d_1,h_2)=1,\ l(d_1,h_3)=6;$$
$$l(d_2,h_1)=4,\ l(d_2,h_2)=0,\ l(d_2,h_3)=3;$$
$$l(d_3,h_1)=9,\ l(d_3,h_2)=3,\ l(d_3,h_3)=0.$$

因此，行动方案 $d_1$，$d_2$，$d_3$ 的风险分别为

$$\begin{aligned}R(d_1)&=l(d_1,h_1)P(h_1)+l(d_1,h_2)P(h_2)+l(d_1,h_3)P(h_3)\\&=0\times0.3+1\times0.5+6\times0.2\\&=1.7;\end{aligned}$$

$$\begin{aligned}R(d_2)&=l(d_2,h_1)P(h_1)+l(d_2,h_2)P(h_2)+l(d_2,h_3)P(h_3)\\&=4\times0.3+0\times0.5+3\times0.2\\&=1.8;\end{aligned}$$

$$\begin{aligned}R(d_3)&=l(d_3,h_1)P(h_1)+l(d_3,h_2)P(h_2)+l(d_3,h_3)P(h_3)\\&=9\times0.3+3\times0.5+0\times0.2\\&=4.2.\end{aligned}$$

由 $R(d_1)<R(d_2)<R(d_3)$，知行动方案 $d_1$ 所造成的风险最小，所以应该选择行动方案 $d_1$. 这与按照平均收益最大准则做出的决策是相同的.

### 3. 损益矩阵

通常，人们用一种常用的数学工具——矩阵来表示上述计算损益函数的均值的过程. 下面我们就简单介绍一下这种表示方法. 例如，案例 1 中的损失函数可以用矩阵表示为

$$\boldsymbol{L}^{❶}=\begin{pmatrix}50 & 0\\60 & 100\end{pmatrix},$$

❶ 像这样由圆括号和括号内的数字阵列组成的式子称为**矩阵**. 关于矩阵的概念请参见附录.

案例 2 中的收益函数可以用矩阵表示为

$$\boldsymbol{Q}=\begin{pmatrix}12 & 8 & 3\\5 & 6 & 3\\-4 & -1 & 2\end{pmatrix}.$$

像 $\boldsymbol{L}$ 这样表示损失函数的矩阵称为**损失矩阵**，像 $\boldsymbol{Q}$ 这样表示收益函数的矩阵称为**收益矩阵**. 损失矩阵和收益矩阵统称为**损益矩阵**.

在损失矩阵 $\boldsymbol{L}$ 中，第 1 列表示行动方案 $d_1$ 在不同状态下的损失值，第 2 列表示行动方案 $d_2$ 在不同状态下的损失值；第 1 行表示在状态 $h_1$ 下各行动方案的损失值，第 2 行表示在状态 $h_2$ 下各行动方案的损失值. 类似地，在收益矩阵 $\boldsymbol{Q}$ 中，第 1 列表示行动方案 $d_1$ 在不同状态下的收益值，第 2 列表示行动方案 $d_2$ 在不同状态下的收益值，第 3 列表示行动方案 $d_3$ 在不同状态下的收益值；第 1 行表示在状态 $h_1$ 下各行动方案的收益值，第 2 行表示在状态 $h_2$ 下各行动方案的收益值，第 3 行表示在状态 $h_3$ 下各行动方案的收益值.

为了表示损益矩阵的各行和各列的实际含义，常在行的右边标注该行所对应的状态，在列的上边标注该列所对应的行动方案. 例如，收益矩阵 $\boldsymbol{Q}$ 可以标注为

$$\begin{array}{cc} & \begin{array}{ccc}d_1 & d_2 & d_3\end{array} \\ \boldsymbol{Q}= & \begin{pmatrix}12 & 8 & 3\\5 & 6 & 3\\-4 & -1 & 2\end{pmatrix}\end{array}\begin{array}{c}\\ h_1\\h_2\\h_3\end{array}$$

**思考**

如何由损益函数写出损益矩阵？

以收益矩阵为例. 先将行动方案 $d_1$ 所对应的收益函数 $q(d_1, h)$在不同状态下的取值依次写入矩阵的第 1 列，然后将行动方案 $d_2$ 所对应的收益函数 $q(d_2, h)$在不同状态下的取值依次写入矩阵的第 2 列，将行动方案 $d_3$ 所对应的收益函数 $q(d_3, h)$在不同状态下的取值依次写入矩阵的第 3 列……依此类推，就可以得到收益函数相应的收益矩阵了. 在收益矩

阵中，第 $i$ 行第 $j$ 列位置的元素恰好等于 $q(d_j, h_i)$.

类似地，案例 2 中 $h$ 的分布列也可以用矩阵表示为

$$(0.3 \quad 0.5 \quad 0.2),$$

那么收益函数 $q(d, h)$ 的均值可以通过矩阵的乘法运算❶得到

$$(0.3 \quad 0.5 \quad 0.2)\begin{pmatrix} 12 & 8 & 3 \\ 5 & 6 & 3 \\ -4 & -1 & 2 \end{pmatrix}$$

$$=(0.3\times12+0.5\times5+0.2\times(-4) \quad 0.3\times8+0.5\times6+0.2\times(-1) \quad 0.3\times3+0.5\times3+0.2\times2)$$

$$=(5.3 \quad 5.2 \quad 2.8),$$

则 $Q(d_1)=5.3$，$Q(d_2)=5.2$，$Q(d_3)=2.8$.

❶关于矩阵的运算请参见附录.

**思考**

案例 1 中的风险能用矩阵的乘法计算吗？

下面，我们举例说明如何利用损益矩阵对一个具体问题做出决策.

**案例 3** 根据气象预报，某地区近期有小洪水的概率为 0.25，有大洪水的概率为 0.01. 该地区某工地上有一台大型设备，遇到大洪水时要损失 60 000 元，遇到小洪水时要损失 10 000 元. 为保护设备，有以下 3 种方案可供选择：

方案 1：运走设备，搬运费为 3 800 元.

方案 2：建保护围墙，建设费为 2 000 元，但围墙只能防小洪水.

方案 3：不采取措施，希望不发生洪水.

你会选择哪个方案？为什么？

这个问题的决策目标是使受到的损失最小. 用 $d_1$，$d_2$ 和 $d_3$ 分别表示 3 种可供选择的方案，即

$d_1$：运走设备，$d_2$：建保护围墙，$d_3$：不采取措施.

用 $h_1$，$h_2$ 和 $h_3$ 分别表示 3 种可能发生的状态，即

$h_1$：发生大洪水，$h_2$：发生小洪水，$h_3$：不发生洪水.

由于发生小洪水的概率为 0.25，发生大洪水的概率为 0.01，则不发生洪水的概率为 $1-0.25-0.01=0.74$，那么 $h$ 的分布列为

| $h$ | $h_1$ | $h_2$ | $h_3$ |
|---|---|---|---|
| $P(h)$ | 0.01 | 0.25 | 0.74 |

若采取行动方案 $d_1$，无论有无洪水，都会损失 3 800 元；若采取行动方案 $d_2$，发生大洪水时，损失 2 000+60 000=62 000（元），发生小洪水或不发生洪水时，损失 2 000 元；

若采取行动方案 $d_3$，发生大洪水时，损失 60 000 元，发生小洪水时，损失 10 000 元，不发生洪水时，没有损失. 因此，可以定义损失矩阵为

$$\begin{array}{cc} & \begin{array}{ccc} d_1 & d_2 & d_3 \end{array} \\ \boldsymbol{L}= & \begin{pmatrix} 3\,800 & 62\,000 & 60\,000 \\ 3\,800 & 2\,000 & 10\,000 \\ 3\,800 & 2\,000 & 0 \end{pmatrix} \end{array} \begin{array}{c} h_1 \\ h_2 \\ h_3 \end{array}$$

$h$ 的分布列可以用矩阵表示为

$$(0.01 \quad 0.25 \quad 0.74).$$

因此，行动方案 $d_1$，$d_2$ 和 $d_3$ 的风险为

$$\begin{aligned}(R(d_1) \quad R(d_2) \quad R(d_3)) &= (0.01 \quad 0.25 \quad 0.74)\begin{pmatrix} 3\,800 & 62\,000 & 60\,000 \\ 3\,800 & 2\,000 & 10\,000 \\ 3\,800 & 2\,000 & 0 \end{pmatrix} \\ &= (3\,800 \quad 2\,600 \quad 3\,100).\end{aligned}$$

由于 $R(d_2)<R(d_3)<R(d_1)$，故行动方案 $d_2$ 所造成的风险最小，所以应该选择行动方案 $d_2$，即“建保护围墙”.

在《选修 2-3》中，利用平均损失的方法来解决这个问题，实际上用的就是风险决策的思想.

## 4. 风险型决策

下面我们对上述解决问题的方法做一个小结. 案例 1～案例 3 的共同特点是：未来的情况是不确定的，但各种可能状态发生的概率可以估算出来. 通常，这样的问题可以用三个要素表达出来：

**1. 状态及其分布列** 这里的状态也称为自然状态，即可能发生的客观情况，是人为不可控制的，但各种状态发生的概率是可以通过统计办法获得的，且它们的概率之和等于 1.

**2. 行动方案** 指决策者可以采取的所有策略. 通常，大部分行动方案的后果是不确定的，取决于不同的状态. 例如，案例 3 中，行动方案 $d_1$ 所造成的损失是确定的，但行动方案 $d_2$ 和 $d_3$ 所造成的损失都与状态有关.

**3. 损益函数** 指各行动方案在各种状态下的损失值或收益值，实际上是对所采取的行动方案造成的后果的数量化.

对这样的问题做出决策的过程称为**风险型决策**. 它的一般决策过程是：

> 第一步，明确问题的决策目标；
> 第二步，寻找所有可选择的行动方案；
> 第三步，确定所有可能出现的状态，以及每种状态发生的概率；
> 第四步，确定损益函数或损益矩阵；
> 第五步，计算各行动方案所对应的损益函数的均值；
> 第六步，根据决策目标，按照风险最小准则或平均收益最大准则选择最优决策.

在进行风险型决策时，状态及其分布列、行动方案或损益函数的不同，都会影响最终

的最优决策．下面，我们来看一个损失函数影响最优决策的例子．

案例 4　A，B 两个箱中装有同样规格的 1 件产品．A 箱产品的废品率是 1%，价格是 120 元；B 箱产品的废品率是 3%，价格是 60 元．若不允许开箱验货，购买时你会选择哪箱产品？

一种自然的想法是：把买到正品看成无损失，买到废品看成有损失，其损失量为所买废品的价格，那么决策目标就是使购买产品所造成的损失最小．

我们所能采取的全部行动方案包括两种：

$$d_1\text{：购买 A 箱中的产品，}d_2\text{：购买 B 箱中的产品．}$$

各行动方案所造成的损失与买到的是正品还是废品有关，所以购买中可能出现的状态包括：

$$h_1\text{：买到正品，}h_2\text{：买到废品．}$$

因为各行动方案所造成的损失量就是所买废品的价格，所以本例中的损失函数可以定义为：

$$l(d_1,\ h_1)=0,\ l(d_1,\ h_2)=120;$$
$$l(d_2,\ h_1)=0,\ l(d_2,\ h_2)=60.$$

下面，我们来计算两种行动方案下，不同状态发生的概率．当采取行动方案 $d_1$ 时，出现状态 $h_1$ 的概率为 $1-0.01=0.99$，出现状态 $h_2$ 的概率为 0.01，这时 $h$ 的分布列为

| $h$ | $h_1$ | $h_2$ |
|---|---|---|
| $P(h)$ | 0.99 | 0.01 |

当采取行动方案 $d_2$ 时，出现状态 $h_1$ 的概率为 $1-0.03=0.97$，出现状态 $h_2$ 的概率为 0.03，这时 $h$ 的分布列为

| $h$ | $h_1$ | $h_2$ |
|---|---|---|
| $P(h)$ | 0.97 | 0.03 |

因此，行动方案 $d_1$ 和 $d_2$ 的风险分别为

$$\begin{aligned}R(d_1)&=l(d_1,\ h_1)\times P(\text{采取 }d_1\text{，出现 }h_1)+l(d_1,\ h_2)\times P(\text{采取 }d_1\text{，出现 }h_2)\\&=0\times0.99+120\times0.01\\&=1.2;\end{aligned}$$

$$\begin{aligned}R(d_2)&=l(d_2,\ h_1)\times P(\text{采取 }d_2\text{，出现 }h_1)+l(d_2,\ h_2)\times P(\text{采取 }d_2\text{，出现 }h_2)\\&=0\times0.97+60\times0.03\\&=1.8.\end{aligned}$$

因为 $R(d_1)<R(d_2)$，即行动方案 $d_1$ 所造成的风险较小，所以应该选择行动方案 $d_1$，即“购买 A 箱中的产品”．

从“买到正品看成无损失，买到废品看成有损失”的角度定义损失函数，在特定条件下有其合理性．如当产品是灯泡，且 A 箱中正品灯泡的寿命是 B 箱中正品灯泡寿命的 2 倍时是合理的．但当两个箱子中正品的质量没有差别时，同样买到正品，买 A 箱要比买 B 箱

多花 60 元. 因此我们也可以这样来界定损失的情况：买到 A 箱正品看成损失 60 元，买到 B 箱正品看成无损失.

因此本例的损失函数又可以定义为

$$l(d_1,\ h_1)=60,\ l(d_1,\ h_2)=120;$$
$$l(d_2,\ h_1)=0,\ l(d_2,\ h_2)=60.$$

这时，行动方案 $d_1$ 和 $d_2$ 的风险分别为

$$\begin{aligned}R(d_1)&=l(d_1,\ h_1)\times P(\text{采取 } d_1\text{，出现 } h_1)+l(d_1,\ h_2)\times P(\text{采取 } d_1\text{，出现 } h_2)\\&=60\times 0.99+120\times 0.01\\&=60.6,\end{aligned}$$

$$\begin{aligned}R(d_2)&=l(d_2,\ h_1)\times P(\text{采取 } d_2\text{，出现 } h_1)+l(d_2,\ h_2)\times P(\text{采取 } d_2\text{，出现 } h_2)\\&=0\times 0.97+60\times 0.03\\&=1.8.\end{aligned}$$

因为 $R(d_1)>R(d_2)$，即行动方案 $d_2$ 所造成的风险较小，所以应该选择行动方案 $d_2$，即“购买 B 箱中的产品”.

思考

前面我们基于两种不同的想法，构造了两种不同的损失函数，得到了两个截然相反的最优决策. 面对这两个最优决策，你应该怎么办？

实际上，可以把损失函数（或收益函数）看成评价最优决策的标准，不同的标准可能得到不同的结论. 这与评价一个学生的学习成绩的情形类似，我们可以用数学成绩作为评价标准，也可以用语文成绩作为评价标准，还可以用这两个成绩之和作为评价标准. 不同的评价标准有不同的实际含义. 当我们要选拔数学能力强的学生时，用数学成绩作为评价标准更合理；要选拔语文能力强的学生时，用语文成绩作为评价标准更合理；要选拔数学和语文的综合能力强的学生时，用数学和语文的成绩之和作为评价标准更合理. 也就是说，应该根据不同的实际要求，选择评价标准.

综上所述，在进行风险型决策时，应该注意下列问题：

1. 应该考虑所有可供选择的行动方案和所有可能的状态，并且能够确定各种状态发生的概率，才有可能得到最优决策.

2. 应该根据实际情况合理构造损益函数，以使评价各个行动方案的标准更合理.

## 探究与发现

### 风险相差不大时该如何决策

当一个决策问题中有多个行动方案的风险达到最小风险（或平均收益达到最大平均收益）时，应该如何决策呢？

我们用一个例子来回答这个问题.

有一个投资100万元的零售企业，其发生火灾的概率为0.1%. 若发生火灾，该企业将蒙受100万元的损失；若购买保险，可以弥补所有的损失，但需交纳1 000元的保险费. 如果你是企业的决策者，你会选择购买保险还是不购买保险？

这个问题的决策目标是使企业受到的损失最小. 分别用 $d_1$ 和 $d_2$ 表示两种行动方案，用 $h_1$ 和 $h_2$ 表示两种可能发生的状态，即

$d_1$：购买保险，$d_2$：不购买保险；

$h_1$：发生火灾，$h_2$：不发生火灾.

$h$ 的分布列为

| $h$ | $h_1$ | $h_2$ |
|---|---|---|
| $P(h)$ | 0.001 | 0.999 |

根据问题的背景，损失函数可以定义为

$$l(d_1, h_1)=1\,000,\ l(d_1, h_2)=1\,000;$$
$$l(d_2, h_1)=1\,000\,000,\ l(d_2, h_2)=0.$$

因此，行动方案 $d_1$ 和 $d_2$ 所造成的风险分别为

$$\begin{aligned}R(d_1)&=l(d_1, h_1)\times P(h_1)+l(d_1, h_2)\times P(h_2)\\&=1\,000\times 0.001+1\,000\times 0.999\\&=1\,000;\\R(d_2)&=l(d_2, h_1)\times P(h_1)+l(d_2, h_2)\times P(h_2)\\&=1\,000\,000\times 0.001+0\times 0.999\\&=1\,000.\end{aligned}$$

所以，$R(d_1)=R(d_2)$.

此时该如何决策？可能有的同学认为可以在两个行动方案中随机选择一个. 我们可以进一步这样思考：购买保险仅损失1 000元，而不购买保险，该企业将冒损失100万元的风险，因此两个行动方案所造成的结果是不同的.

我们已经知道在每个行动方案下，损失函数是随机变量，而均值体现了随机变量的中心位置，方差体现了随机变量集中于中心位置的程度. 因此，当两个随机变量的均值相同，方差不同时，方差小的更集中于均值的附近，所以这时应该选择损失方差小的行动方案.

行动方案 $d_1$ 和 $d_2$ 所对应的损失函数的方差分别为

$\mathrm{var}(l(d_1, h))=0$；

$$\mathrm{var}(l(d_2,\ h))=(1\ 000\ 000-1\ 000)^2\times0.001+(0-1\ 000)^2\times0.999$$
$$=999\ 000\ 000.$$

因为 $\mathrm{var}(l(d_1,\ h))=0$，且远远小于 $\mathrm{var}(l(d_2,\ h))$，所以行动方案 $d_1$ 所造成的损失稳定在 1 000 元，行动方案 $d_2$ 所造成的损失很不稳定，可能为 100 万元，也可能为 0. 因此，应该选择行动方案 $d_1$，即“购买保险”.

事实上，只要风险相差不大，我们就可以考虑用方差来帮助分析风险.

## 习题1.2

1. 若烹饪一道菜需要 4 个鸡蛋，已经把 3 个好鸡蛋敲破，倒在一只碗里，第 4 个鸡蛋可能是坏鸡蛋，若把它敲进同一只碗里，将破坏前 3 个鸡蛋. 怎样处理这个未敲开的鸡蛋？写出可选择的行动方案和所有可能发生的状态.

2. 某公司需要决策明年的生产情况，计划部门进行市场调研后得到下面的收益数据表(单位：元)：

| | 市场对产品的需求增加（概率为 0.6） | 市场对产品的需求下降（概率为 0.4） |
|---|---|---|
| 保持现状的收益 | 340 000 | 300 000 |
| 增加员工的收益 | 420 000 | 270 000 |
| 增加新设备的收益 | 440 000 | 260 000 |

(1) 写出可选择的行动方案和所有可能发生的状态.

(2) 构造一个损失函数.

(3) 计算每个行动方案的风险，并给出最优决策.

3. 某商场在采购商品时，需要决策当月的进货数量. 根据市场预测，某种商品的月销售量及相应的概率为

| 销售量/台 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|
| 概率 | 0.05 | 0.10 | 0.20 | 0.40 | 0.20 | 0.05 |

该商品每台成本为 800 元，售价为 1 000 元. 如果当月不能售出，则下个月报废.

(1) 写出可选择的行动方案和所有可能发生的状态.

(2) 构造一个收益函数.

(3) 计算每个行动方案的平均收益，并给出最优决策.

4. 某书店希望订购最新出版的图书出售，计划新书到货后按 6 元一本销售一个月，一个月后降价处理，处理价为 2 元. 已知该书店过去新书在一个月内销售数量的规律如下表所示：

| 销售量/本 | 50 | 100 | 150 | 200 |
|---|---|---|---|---|
| 所占百分比 | 20% | 40% | 30% | 10% |

若每本新书的订购价为 4 元，

(1) 写出可选择的行动方案和所有可能发生的状态.

(2) 写出收益矩阵.

(3) 计算每个行动方案的平均收益，并给出最优决策.

# 第二讲 决策树方法

决策树方法因其运用树状图形来分析和选择决策方案而得名. 通过绘制决策树，我们可以把各种行动方案、可能出现的状态及其概率以及损益函数的值在同一张图上清楚明确地表示出来，使得决策过程形象直观地在图上完成.

下面，我们就先来了解决策树的基本构成. 风险型决策问题是决策者在一些可供选择的行动方案中选出最优方案. 例如，若决策者有 3 个备选行动方案，则可以用图 2-1 表示为

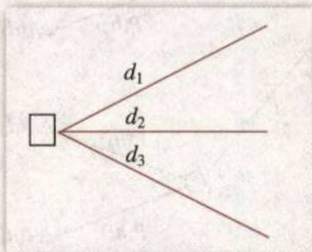



图 2-1

其中方形框叫做决策点，即决策者在这里要对各行动方案进行选择. 由决策点引出的 3 条线段分别代表可供选择的 3 种行动方案，叫做方案枝.

在风险型决策中，行动方案的后果往往是不确定的，取决于不同的状态，不同行动方案要面临的状态也可能不同. 例如，上述问题中，假设每种行动方案都有 2 种可能发生的状态，则可以用图 2-2 表示为

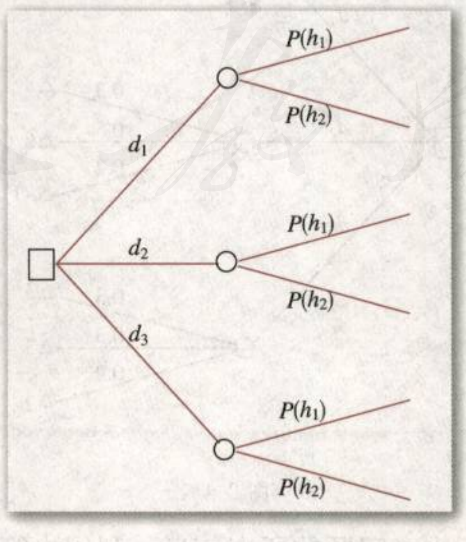



图 2-2

方案枝末端的圆叫做机会点，表示由此开始的情况是决策者无法控制的. 由机会点引出的 2 条线段表示有 2 种可能发生的状态，叫做状态枝（概率枝）. 我们将状态发生的概率标注在相应的概率枝上.

对于同一个行动方案，不同的状态对应不同的损失值（或收益值），为此在状态枝的末端画一个三角形，叫做后果点，并将相应损失函数（或收益函数）的值标注在后果点的旁边. 例如，若上述问题中构造的是损失函数，则可以用图 2-3 表示为

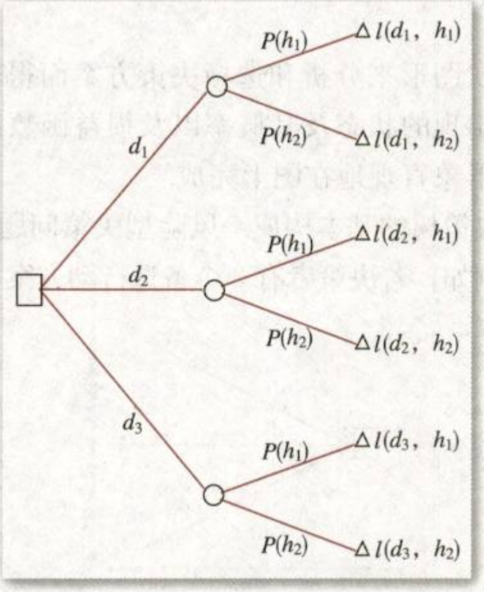



图 2-3

案例 2 的决策树可以画成如图 2-4 的形式.

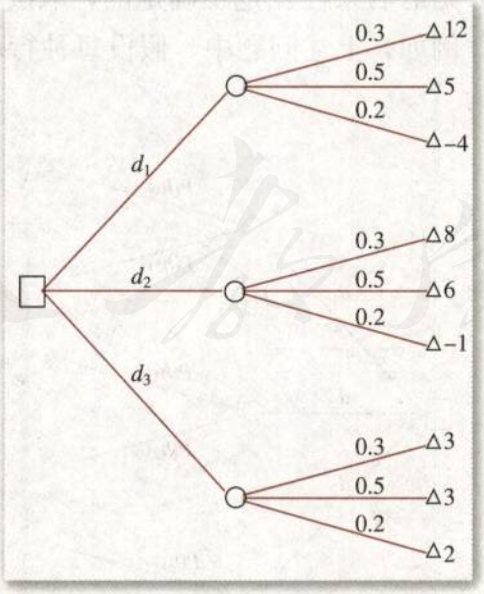



图 2-4

由此可见，决策树包含了决策问题的所有信息. 利用决策树，还能方便直观地将风险型决策过程表达出来. 按照决策树从右到左的顺序，决策过程可按如下步骤进行：

第一步，计算损益函数的均值.

求连接到同一机会点的各个后果点的值 $l(d,h)$（或 $q(d,h)$）与相应的状态枝的标注值 $P(h)$ 的乘积，并将所得的乘积加起来，得到该机会点的均值. 例如，由图 2-4 可得各机会点的均值为

$$12\times0.3+5\times0.5+(-4)\times0.2=5.3,$$
$$8\times0.3+6\times0.5+(-1)\times0.2=5.2,$$
$$3\times0.3+3\times0.5+2\times0.2=2.8.$$

把这 3 个值标注在相应的机会点上（图 2-5）.

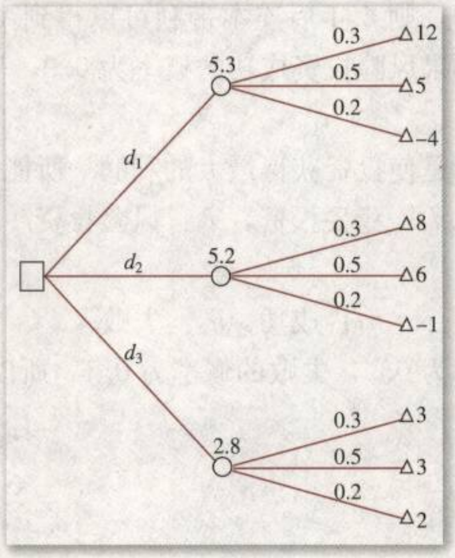



图 2-5

第二步，根据决策目标选择最优决策.

案例 2 的决策目标是使种植获得最大的收益. 在图 2-5 中，方案枝 $d_1$ 末端的机会点的值是 5.3，方案枝 $d_2$ 末端的机会点的值是 5.2，方案枝 $d_3$ 末端的机会点的值是 2.8，所以最优决策为 $d_1$. 在图 2-5 中，把“$d_1$”填入决策点的方框中，并“剪去”方案枝 $d_2$ 和 $d_3$，如图 2-6 所示.

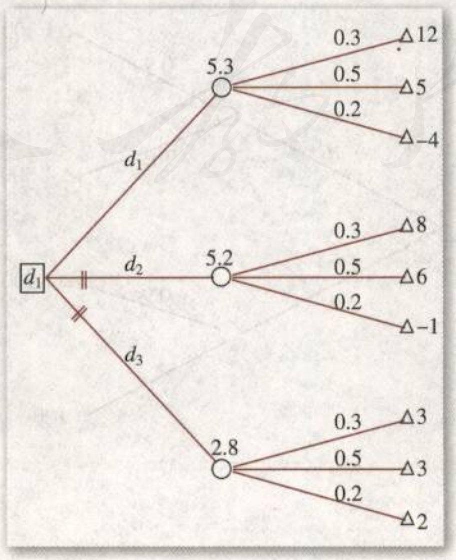



图 2-6

图 2-6 就是最终的决策树，它不仅包括了案例 2 中的所有信息，而且显示了决策的结果，即应该选择行动方案 $d_1$，它所对应的平均收益为 5.3.

下面，我们再来看一个用决策树进行决策的例子.

**例 1** 假如你有 1 000 元钱，准备进行投资. 有两种投资方案可供选择：一种是稳妥的投资，比如储蓄、国债等；一种是有风险的投资，比如经营、股票等. 通常，稳妥投资的利润小，而有风险的投资利润可能很大，也可能很小，甚至赔本.

假设选择稳妥投资，5 年后连本带利可以得到 1 300 元；选择有风险的投资，回报分为成功和失败两种情况：若成功，则 5 年后连本带利可以得到 1 500 元；若失败，则 5 年后连本带利只能得到 100 元. 如果风险投资成功的概率为 0.9，失败的概率为 0.1，应该选择哪种投资方案？

**解：** 这个问题的决策目标是使投资获得最大的利润. 所能采取的全部行动方案包括：

$$d_1\text{：稳妥投资，}d_2\text{：风险投资.}$$

所面临的状态包括：

$$h_1\text{：成功，}h_2\text{：失败.}$$

由于风险投资成功的概率为 0.9，失败的概率为 0.1，所以 $h$ 的分布列为

| $h$ | $h_1$ | $h_2$ |
|---|---|---|
| $P(h)$ | 0.9 | 0.1 |

对于行动方案 $d_1$，无论是状态 $h_1$ 发生还是状态 $h_2$ 发生，5 年后的收益都是 1 300 元；对于行动方案 $d_2$，当状态 $h_1$ 发生时，5 年后的收益是 1 500 元，当状态 $h_2$ 发生时，5 年后的收益是 100 元. 因此，可以定义收益函数如下：

$$q(d_1,\ h_1)=1\ 300,\ q(d_1,\ h_2)=1\ 300;$$

$$q(d_2,\ h_1)=1\ 500,\ q(d_2,\ h_2)=100.$$

综上所述，本例的决策树可以画成如下的形式（图 2-7）.

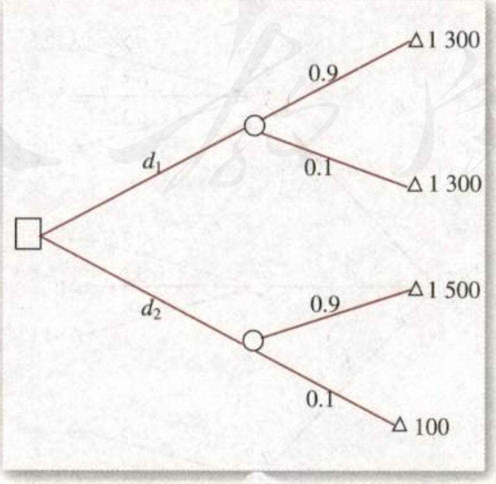



图 2-7

由图 2-7 可得，两个机会点的取值分别为

$$1\ 300\times0.9+1\ 300\times0.1=1\ 300,$$
$$1\ 500\times0.9+100\times0.1=1\ 360.$$

把这两个值标注在相应的机会点上（图 2-8）．由于本例的决策目标是使投资获得最大的利润，所以由方案枝 $d_2$ 末端机会点的值大于方案枝 $d_1$ 末端机会点的值可知，应该选择行动方案 $d_2$．在图 2-7 中，把“$d_2$”填入决策点的方框中，并“剪去”方案枝 $d_1$，如图 2-8 所示．

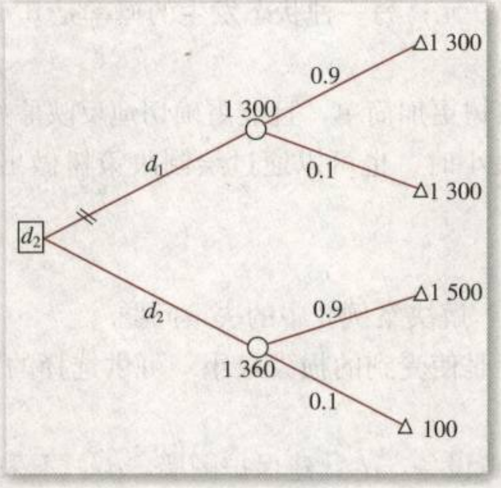



图 2-8

思考

观察图 2-8 中的决策树，方案枝 $d_1$ 末端的机会点所连接的两个后果点的标注值都是 1 300，它所对应的实际含义是什么？例 1 的决策树可以简化吗？

由例 1 的背景可以知道，对应于稳妥投资方案 $d_1$，5 年后的回报只有 1 种，即 1 300 元．因此，采用行动方案 $d_1$ 后，只有一种情况发生，这在决策树上表现为：方案枝 $d_1$ 末端的机会点所连接的各个后果点的标注值都是 1 300．在实际应用中，如果决策树上同一方案枝末端的机会点所连接的后果点的标注值相同，那么该机会点所连接的状态枝可以合并为一枝，以简化决策树的结构．因此，图 2-8 中的决策树可以简化为(图 2-9)：

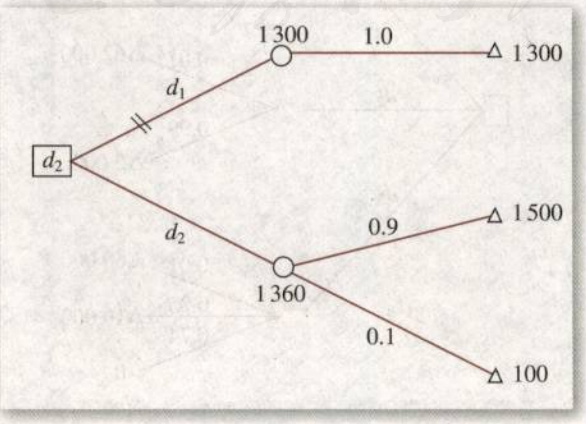



图 2-9

在图 2-9 中，最上面一条状态枝是图 2-8 中最上面两条状态枝合并的结果. 合并后状态枝的标注值，等于合并前的两个状态枝的标注值之和，即 $0.9+0.1=1.0$；而合并后的状态枝末端后果点的标注值保持不变，还是 1 300. 图 2-9 中的决策树可以解释为：

1. 采取行动方案 $d_1$ 后，只有一种状态发生，结果是 5 年后的回报为 1 300 元.

2. 采取行动方案 $d_2$ 后，有两种不同的状态发生. 一种状态发生的概率是 0.9，相应的结果是 5 年后的回报为 1 500 元；另一种状态发生的概率是 0.1，相应的结果是 5 年后的回报为 100 元.

显然，图 2-9 中的决策树更加简单，且能更确切地反映原始信息.

当决策目标是使风险最小时，也可以通过绘制决策树做出决策. 下面，我们以案例 3 为例加以说明.

**例 2** 利用决策树方法，解决案例 3 中的决策问题.

**解：** 案例 3 的决策目标是使受到的损失最小. 可供选择的行动方案和可能发生的状态分别为：

$d_1$：运走设备，$d_2$：建保护围墙，$d_3$：不采取措施；

$h_1$：发生大洪水，$h_2$：发生小洪水，$h_3$：不发生洪水.

$h$ 的分布列为：

| $h$ | $h_1$ | $h_2$ | $h_3$ |
|---|---|---|---|
| $P(h)$ | 0.01 | 0.25 | 0.74 |

相应的损失函数的取值为：

$$l(d_1,\ h_1)=3\ 800,\ l(d_1,\ h_2)=3\ 800,\ l(d_1,\ h_3)=3\ 800;$$
$$l(d_2,\ h_1)=62\ 000,\ l(d_2,\ h_2)=2\ 000,\ l(d_2,\ h_3)=2\ 000;$$
$$l(d_3,\ h_1)=60\ 000,\ l(d_3,\ h_2)=10\ 000,\ l(d_3,\ h_3)=0.$$

把这些信息用决策树表示如下（图 2-10）

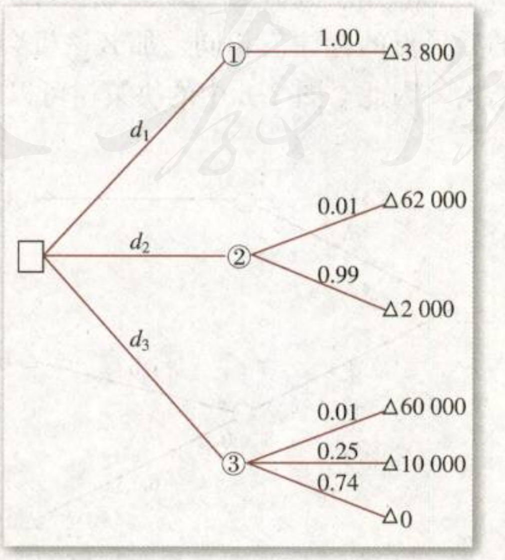



图 2-10

为了方便地表达决策过程，我们给图 2-10 中的机会点编了序号．由图 2-10 可得：

机会点 1 的均值为

$$3\ 800\times 1=3\ 800;$$

机会点 2 的均值为

$$62\ 000\times 0.01+2\ 000\times 0.99=2\ 600;$$

机会点 3 的均值为

$$60\ 000\times 0.01+10\ 000\times 0.25+0\times 0.74=3\ 100.$$

把这些值标注到图 2-10 中相应的机会点上，得到图 2-11．根据案例 3 的决策目标，比较上述 3 个均值可知，应该选择行动方案 $d_2$．因此，在图 2-11 中，把“$d_2$”填入决策点的方框中，并“剪去”方案枝 $d_1$ 和 $d_3$.

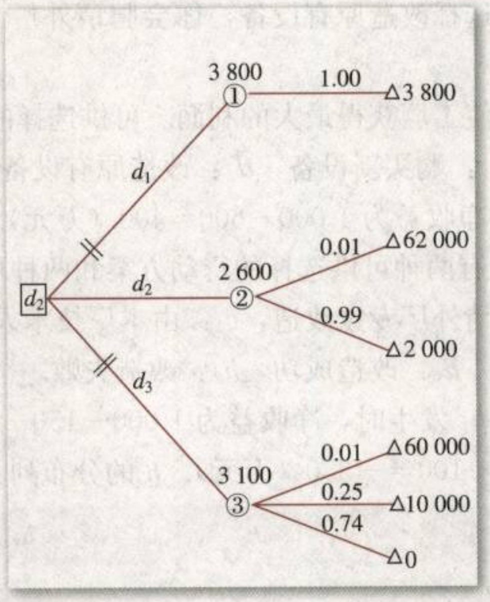



图 2-11

像这样利用决策树进行决策的方法称为**决策树方法**.

思考

根据例 1 和例 2 的决策树的绘制过程，你能总结出决策树方法的一般步骤吗？

一般地，绘制决策树的步骤如下：

第一步，明确问题的决策目标、所有可选择的行动方案、所有可能出现的状态以及每种状态发生的概率、损益函数的所有取值；

第二步，从左到右依次画出决策点、方案枝、机会点、状态枝和后果点，并标注方案枝、状态枝和后果点；

第三步，计算并标注各机会点的值；

第四步，根据决策目标，比较各机会点的值，将最小值或最大值所对应的行动方案填在决策点的方框中，并“剪去”其他的方案枝.

前面我们用决策树方法解决的都是一步决策问题．对于多步决策问题，可以通过在决策树中进一步引进决策点来实现．下面看一个例子．

**例 3** 某工厂打算投资生产一种新产品．现有两种生产方式可供选择：购买新设备或改造原有设备．已知购买新设备需要花费 600 万元，改造原有设备需要花费购置配件费用等 150 万元．而对原有设备进行改造又有两种备选方案：聘请外厂专家改造或由本厂技术人员改造．若聘请外厂专家改造，需要另外支付聘金 100 万元，改造成功的概率为 80%；若由本厂技术人员改造，需要另外支付劳务费 50 万元，改造成功的概率为 70%．若改造失败，除上面的花费外将给工厂另外造成 100 万元的损失．

若这种新产品成功投产，该工厂将赢利 1 000 万元．如果你是这家工厂的决策者，你会选择哪种生产方式？如果选择改造原有设备，你会聘请外厂专家改造还是请本厂技术人员改造？

**解：** 本例的决策目标是使工厂获得最大的利润．可供选择的行动方案包括：

$d_1$：购买新设备，$d_2$：改造原有设备．

若采取行动方案 $d_1$，则净收益为 $1\,000-600=400$（万元）．

若采取行动方案 $d_2$，又有两种可供选择的行动方案和两种可能发生的状态：

$c_1$：聘请外厂专家改造，$c_2$：由本厂技术人员改造；

$h_1$：改造成功，$h_2$：改造失败．

这时若采取行动方案 $c_1$，当 $h_1$ 发生时，净收益为 $1\,000-150-100=750$（万元）；当 $h_2$ 发生时，净收益为 $-150-100-100=-350$（万元）．$h$ 的分布列为：

| $h$ | $h_1$ | $h_2$ |
|---|---|---|
| $P(h)$ | 0.80 | 0.20 |

若采取行动方案 $c_2$，当 $h_1$ 发生时，净收益为 $1\,000-150-50=800$（万元）；当 $h_2$ 发生时，净收益为 $-150-50-100=-300$（万元）．$h$ 的分布列为：

| $h$ | $h_1$ | $h_2$ |
|---|---|---|
| $P(h)$ | 0.70 | 0.30 |

因此，本例的决策树可以表示如下（图 2-12）．

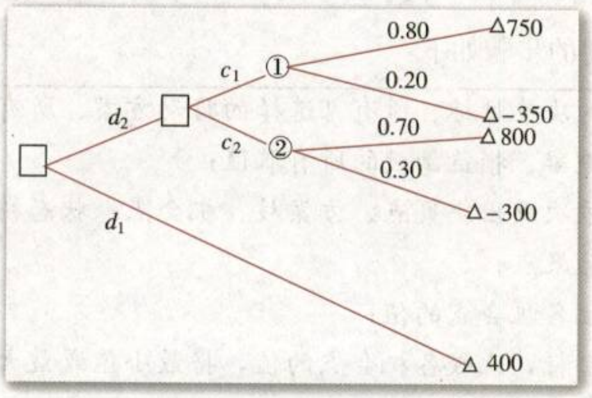



图 2-12

下面，我们按照决策树从右到左的顺序，先对改造原有设备的两个行动方案 $c_1$ 和 $c_2$ 做出选择. 由图 2-12 可得，机会点 1 的净收益的均值为

$$750\times0.80+(-350)\times0.20=530;$$

机会点 2 的净收益的均值为

$$800\times0.70+(-300)\times0.30=470.$$

所以按照平均收益最大准则，应该选择行动方案 $c_1$.

在图 2-12 中，把机会点 1，2 的均值标在相应的机会点上，把“$c_1$”填入相应决策点的方框中，并把行动方案 $c_1$ 的平均收益标在决策点的方框上，同时“剪去”方案枝 $c_2$，得到图 2-13.

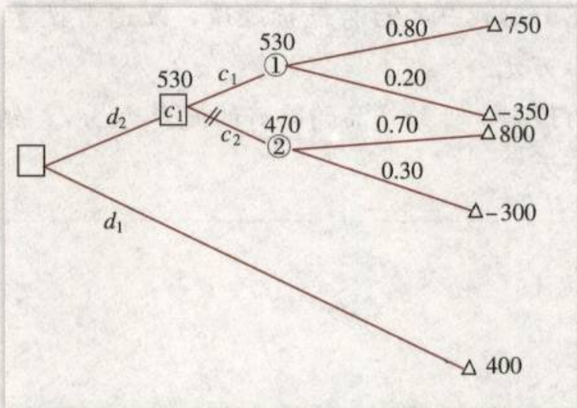



图 2-13

这样，问题就转化为一步决策问题了. 下面，我们对 $d_1$ 和 $d_2$ 做出选择. 由图 2-13 可知，当采取行动方案 $d_1$ 时，净收益为 400 万元，当采取行动方案 $d_2$ 时，净收益的均值为 530 万元，所以应该选择行动方案 $d_2$.

在图 2-13 中，把“$d_2$”填入最左侧决策点的方框中，并“剪去”方案枝 $d_1$，得到图 2-14.

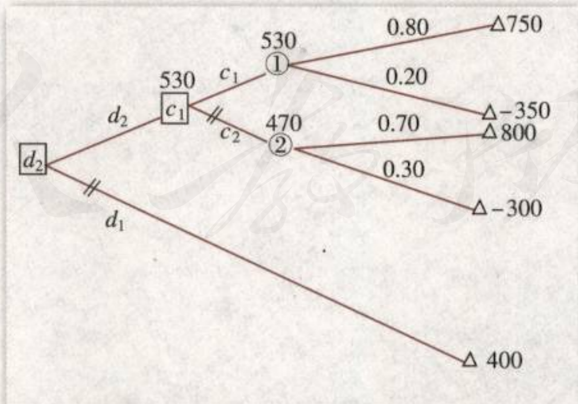



图 2-14

由 2-14 可知，本例的最优决策是通过“聘请外厂专家，改造原有设备”的方式生产新产品.

## 习题 2

1. 用决策树方法解决习题 1.2 第 2 题中的决策问题.

2. 用决策树方法解决习题 1.2 第 3 题中的决策问题.

3. 用决策树方法解决习题 1.2 第 4 题中的决策问题.

4. 某工厂为增加产品的生产量需进行项目投资，现有 3 个可供选择的投资方案，分别为：

   $A_1$：改造设备，需投资 400 万元，使用期限为 10 年. 在此期间销路好时年收益为 120 万元，销路差时为 −30 万元；

   $A_2$：进行扩建，需投资 180 万元，使用期限为 10 年. 在此期间销路好时年收益为 80 万元，销路差时为 20 万元；

   $A_3$：先进行扩建，销路好时，三年后再进行改造，改造需投资 200 万元，使用期限为 7 年，每年收益 110 万元.

   通过对市场销售形势的预测，可知产品销路好的概率为 0.7，销路差的概率为 0.3. 用决策树方法选择最优投资方案.

# 第三讲 风险型决策的敏感性分析

从前面的学习可以发现，在决策问题中，最优决策由状态及其分布列、行动方案以及损益函数所决定，因此任何一个因素的改变都可能影响最优决策.

例如，某一随机事件发生的概率不易获得，需要根据频率近似概率的思想，用样本频率估计各个状态发生的概率，然后进行决策. 虽然随着样本容量的增加，这样的估计会越来越接近于状态分布列，但在实际应用中，样本容量总是有限的，所以样本频率和状态发生的概率之间总会存在一定的差距. 因此，我们必须考虑这种差距会对最优决策产生什么影响，也就是分析未来状态的概率有所变化时，会不会影响最优决策的选择.

**探究**

在第二讲例 1 中，如果风险投资 $d_2$ 的成功概率发生变化，最优决策会发生变化吗？如何评价所得到的最优决策对分布列的依赖性？

沿用例 1 的记号，先考察风险投资 $d_2$ 的成功概率变为 0.8 时，相应的最优决策是否改变. 此时相应的决策树为（图 3-1）：

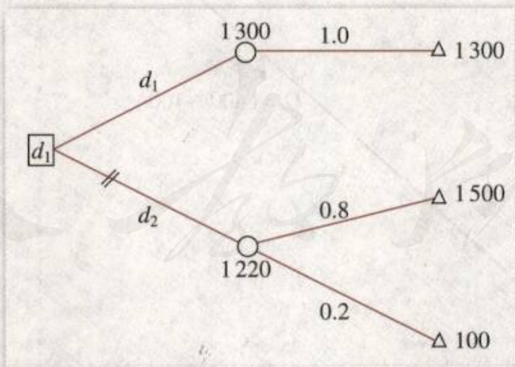


图 3-1

由图 3-1 可知，当风险投资 $d_2$ 的成功概率由 0.9 变为 0.8 时，相应的最优决策由风险投资 $d_2$ 变为稳妥投资 $d_1$. 因此，分布列的变化会引起最优决策的变化.

我们感兴趣的另一个问题是，在风险投资 $d_2$ 的成功概率的变化过程中，什么时候开始出现转折，即成功概率由 0.9 降到多少时，最优决策由风险投资 $d_2$ 变为稳妥投资 $d_1$. 通常，这样的概率叫做转折概率.

假设风险投资 $d_2$ 的成功概率为 $p$，则不成功的概率为 $1-p$，这时稳妥投资 $d_1$ 和风险投资 $d_2$ 的平均收益 $Q_1$，$Q_2$ 与 $p$ 之间的关系为

$$Q_1(p)=1\,300;$$

$$Q_2(p)=1\,500p+100(1-p)=1\,400p+100.$$

相应的决策树为（图 3-2）：

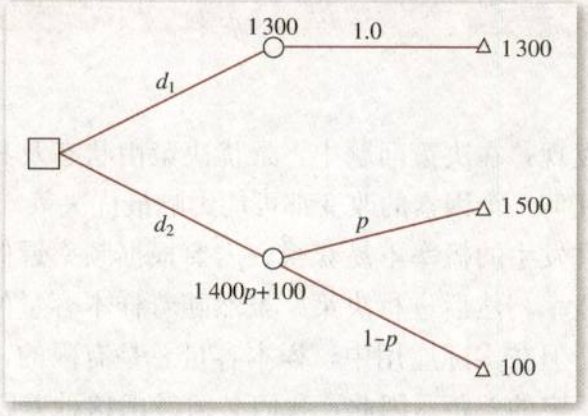



图 3-2

这时，可以用如图 3-3 所示的方法进行分析．以概率 $p$ 为横坐标，平均收益 $Q$ 为纵坐标，把两个行动方案对应的平均收益画在同一个坐标系内．在图 3-3 中，水平的直线代表 $Q_1(p)$，倾斜的直线代表 $Q_2(p)$.

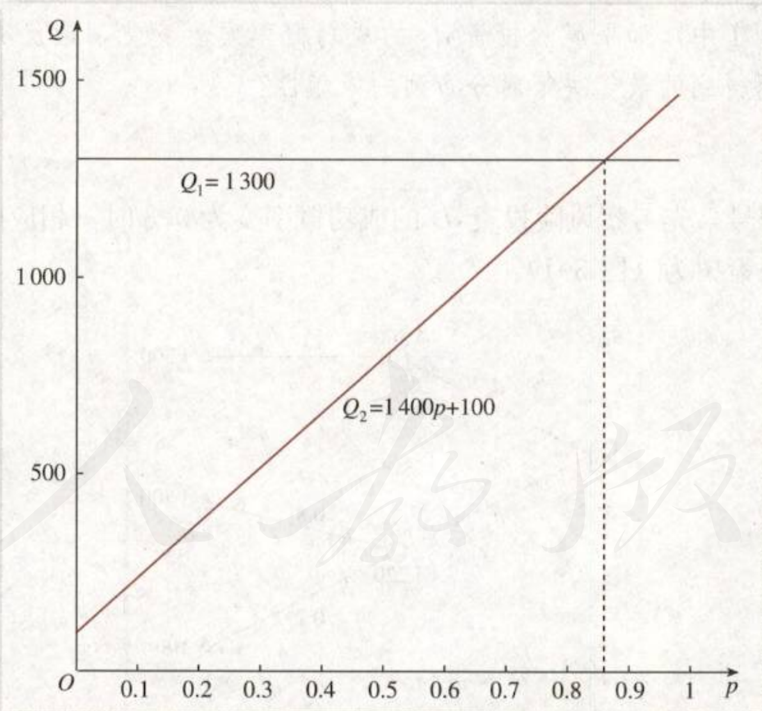



图 3-3

从图中可以清楚地看出，$Q_1(p)$ 和 $Q_2(p)$ 都与 $p$ 呈线性关系．在两条平均收益直线交点处，两种行动方案的平均收益相等，即 $Q_1(p)=Q_2(p)$．解方程

$$1\,400p+100=1\,300,$$

得 $p=\frac{6}{7}$．因此，当 $p>\frac{6}{7}$ 时，行动方案 $d_2$ 为最优决策；当 $p<\frac{6}{7}$ 时，行动方案 $d_1$ 为最优决策．

上面分析了 $p$ 的变化对最优决策的影响，这种分析称为**敏感性分析**．有了敏感性分析的结果，便可以清楚地了解状态发生概率的变化对以均值为标准的最优决策产生的影响．

敏感性分析还可以使我们能够分析自己最初对状态发生概率的估计．例如，如果我们估计 $p$ 的值为 0.7，根据敏感性分析，$0.7<\frac{6}{7}\approx 0.857$，应该选择 $d_1$ 作为最优决策．此时，只要 $p$ 的估计值的误差不超过 $0.857-0.7=0.157$，所得的结论就不会出错．如果我们估计 $p$ 的值为 0.84，与 0.857 很接近，这时最优决策对 $p$ 的变化的反应比较敏感．如果概率估计误差的精度不高，有可能真实的概率值是大于 0.857 的，那么按照 0.84 来决策，就会导致决策错误．

综上所述，若最优决策对概率的变化反应不敏感，则用估计值 $\hat{p}$ 代替 $p$ 所得的最优决策的可靠性就大，决策错误的风险就小；若最优决策对 $p$ 的变化反应敏感，则用估计值 $\hat{p}$ 代替 $p$ 所得的最优决策的可靠性就小，决策错误的风险就大．因此，为了尽量避免这种错误，必须提高概率估计的精度．

**例**　(1) 若 $p$ 的估计值为 $\hat{p}=0.93$，则估计的误差小于多少时，才能保证 $d_2$ 是最优决策？

(2) 若 $p$ 的估计值为 $\hat{p}=0.9$，且估计的误差小于 0.04 的概率为 0.95．那么能有多大的把握认为最优决策是 $d_2$？

**解：**(1) 不妨记误差为 $\delta$．则 $p$ 的值落在区间 $[0.93-\delta,\ 0.93+\delta]$ 内．所以只要区间左端 $0.93-\delta>\frac{6}{7}$，即 $\delta<0.073$，就能保证 $p>\frac{6}{7}$．

因此，估计的误差小于 0.073 时，就能保证 $d_2$ 是最优决策．

(2) 由题意知，$p$ 的值落在区间 $[0.9-0.04,\ 0.9+0.04]$ 内的概率为 0.95，而整个区间 $[0.86,\ 0.94]$ 都在 $\frac{6}{7}$ 的右边．

因此，至少有 95% 的把握认为最优决策是 $d_2$．

### 习题3

1. 假如你有 200 000 元钱，准备进行投资．有两种投资方案可以选择：一种是将 200 000 元存入银行，一年后连本带利可获得 206 000 元；另一种是买股票，如果股票市场是“牛”市，估计一年后连本带利可获得 220 000 元，如果股票市场是“熊”市，估计一年后连本带利只能得到 150 000 元．你会选择哪种投资方案？说明理由．假设未来一年股票市场是“牛”市的概率为 $p$，计算第二种投资方案的平均收益，并讨论 $p$ 在什么范围内取值时，应该选择第二种投资方案．
2. 假设一个资产为 100 万元的家庭，在一年内发生火灾的概率为 0.1%．如果发生火灾，将损失 50 万元；如果购买家庭财产保险，保险公司将赔付损失的 50 万元，但需交纳保险费 800 元．你认为该家庭是否应该购买家庭财产保险？当在一年内发生火灾的概率为多少时，该家庭就应该购买家庭财产保险？

# 第四讲　马尔可夫型决策简介

在实际生活和生产中，我们所研究对象的状态还可能随时间而变化．例如，一台运行的机器，有“正常工作”和“出故障”两种状态．随着时间的推移，机器会从“正常工作”转变到“出故障”状态；经过工人的维修，机器又可以从“出故障”转变到“正常工作”状态．又如，某种产品的销售有“畅销”“平销”和“滞销”三种状态．随着时间的推移，产品的销售情况在这三种状态间相互转换．在状态随时间发生变化的情况下，决策问题变得更加复杂．

马尔可夫链是一个著名的概率模型，它描述随着时间的推移，各个不同状态间的相互转化，在自然科学和社会科学中有着广泛的应用．在实际决策问题中，很多对象的状态随时间变化的规律，都可以用马尔可夫链来刻画．因此，人们将马尔可夫链理论与决策理论相结合，以处理一类状态随时间变化的决策问题，这样就形成了一种独特的决策方法——马尔可夫型决策．

马尔可夫型决策的应用十分广泛，例如，它可以用于研究水文、气象、地质以及市场管理、经营管理、人事管理、项目选址等方面的决策问题．本讲将通过几个案例，介绍一些简单的马尔可夫型决策思想．

## 一　马尔可夫链简介

### 1. 马尔可夫性与马尔可夫链

**案例 5**　考察某工厂一台自动加工机的工作状态．该机器有两种工作状态：正常状态和故障状态．在每个整数钟点的起始时刻检查机器的工作情况，若机器处于正常状态，则让它继续工作；若机器处于故障状态，则对它进行检修．假设处于正常状态的机器，在 1 小时后发生故障的概率为 0.05；处于故障状态的机器，在 1 小时内排除故障的概率为 0.6．

显然在任何时刻，机器只能处于正常状态或故障状态，我们分别用 1 和 2 表示这两种状态．那么，这台机器在第 $n$ 小时的状态 $X_n$ 或者等于 1，或者等于 2，即 $X_n$ 是仅取两个值的离散型随机变量．我们关心的是 $X_n$ 的随机变化规律．

思考

如果已知机器的初始状态 $X_0$ 为正常状态，$X_1$ 的分布列是什么？

由于处于正常状态的机器在 1 小时后发生故障的概率为 0.05，即

$$P(X_1=2|X_0=1)=0.05,$$

所以机器在 1 小时后仍处于正常状态的概率为

$$P(X_1=1|X_0=1)=1-P(X_1=2|X_0=1)=0.95.$$

此时，$X_1$ 的分布列由条件概率构成，这个分布列实际是已知 $X_0=1$ 的情况下 $X_1$ 的条件分布列❶，用表格表示如下

| $X_1$ | 1 | 2 |
|---|---|---|
| $P(\cdot|X_0=1)$ | 0.95 | 0.05 |

❶条件分布指的是在已知某事件发生的情况下随机变量的分布.

思考

如果已知机器在第 $n$ 小时处于故障状态，能够得到 $X_{n+1}$ 的条件分布列吗？如果已知机器在第 $n$ 小时处于正常状态，$X_{n+1}$ 的条件分布列又是什么？

已知机器在第 $n$ 小时处于故障状态，等价于事件$\{X_n=2\}$已经发生．此时在第 $n+1$ 小时机器处于第 $j(j=1,2)$状态的概率就是条件概率 $P(X_{n+1}=j|X_n=2)$．类似于前面的讨论，由故障机器在 1 小时内被修复的概率为 0.6，得

$$\begin{cases}P(X_{n+1}=1|X_n=2)=0.6,\\P(X_{n+1}=2|X_n=2)=1-P(X_{n+1}=1|X_n=2)=0.4.\end{cases}\tag{1}$$

因此，若已知机器在第 $n$ 小时处于故障状态，则 $X_{n+1}$ 的条件分布列为

| $X_{n+1}$ | 1 | 2 |
|---|---|---|
| $P(\cdot|X_n=2)$ | 0.6 | 0.4 |

类似地，若已知机器在第 $n$ 小时处于正常状态，则机器在第 $n+1$ 小时处于第 $j(j=1,2)$状态的概率为

$$\begin{cases}P(X_{n+1}=1|X_n=1)=0.95,\\P(X_{n+1}=2|X_n=1)=0.05.\end{cases}\tag{2}$$

因此，若已知机器在第 $n$ 小时处于正常状态，则 $X_{n+1}$ 的条件分布列为

| $X_{n+1}$ | 1 | 2 |
|---|---|---|
| $P(\cdot|X_n=1)$ | 0.95 | 0.05 |

思考

如果已知机器在第 $n$ 小时处于故障状态，并且机器在 $n-1$（$n\geqslant 1$）小时处于正常状态，$X_{n+1}$ 的条件分布列会发生改变吗？

虽然现在已经知道 $X_n=2$ 和 $X_{n-1}=1$，但由于 $X_{n+1}$ 的概率分布规律完全与机器在第 $n$ 小时所处的状态有关系，即仅由 $X_n$ 的取值所决定，所以

$$P(X_{n+1}=1|\{X_n=2\}\cap\{X_{n-1}=1\})=P(X_{n+1}=1|X_n=2),$$
$$P(X_{n+1}=2|\{X_n=2\}\cap\{X_{n-1}=1\})=P(X_{n+1}=2|X_n=2).$$

事实上，在已知机器在第 $n$ 小时所处状态 $X_n$ 的情况下，$X_{n+1}$ 的随机变化规律与 $X_0$，$X_1$，…，$X_{n-1}$ 的取值都没有关系. 随机变量序列 $\{X_n\}$ 所具有的这类性质称为**马尔可夫性**.

若把 $n$ 看成"现在"，$n+1$ 看成"将来"，小于 $n$ 的整数看成"过去"，则马尔可夫性可以解释为：在已知现在的情况下，将来的随机变化规律与过去发生的事件无关.

在现实生活中，很多随着时间变化的随机变量序列都具有马尔可夫性. 一般地，我们称具有马尔可夫性的随机变量序列 $\{X_n\}$ 为**马尔可夫链**，称序列中随机变量的所有可能取值的集合为该马尔可夫链的状态空间. 为了叙述方便，有时把状态空间的第 $i$ 个元素就用 $i$ 来表示，并称这个元素为该马尔可夫链的第 $i$ 个状态，或状态 $i$. 在本专题中，总是假设状态空间中仅包括有限个不同状态.

例如，表示机器工作状态的随机变量序列 $\{X_n\}$ 就是一个马尔可夫链，其状态空间为 $\{1, 2\}$.

在实际应用中，可以通过问题的背景直接判断随机变量序列是否具有马尔可夫性，而不必计算条件概率. 例如，在表示机器工作状态的随机变量序列 $\{X_n\}$ 中，如果知道了 $X_n=i$（$i=1, 2$），第 $n+1$ 时刻机器状态的随机变化规律与 $n$ 之前机器所处的状态无关，由此可以判断 $\{X_n\}$ 具有马尔可夫性.

## 2. 转移概率与转移概率矩阵

若 $\{X_n\}$ 为马尔可夫链，就可以利用马尔可夫性简化 $X_n$ 的分布列的计算. 下面，我们来看具体的计算方法.

思考

如果已知第 $n$ 小时机器的状态 $X_n$ 的分布列❶为

| $X_n$ | 1 | 2 |
|---|---|---|
| $P(X_n)$ | $p_1^{(n)}$ | $p_2^{(n)}$ |

能求得 $X_{n+1}$ 的分布列吗？

❶注意分布列与条件分布列之间的区别.

可以利用概率的加法公式和条件概率的知识解决这个问题. 首先对于任何给定的 $j$（$j=1, 2$），事件 $\{X_n=1, X_{n+1}=j\}$ 和事件 $\{X_n=2, X_{n+1}=j\}$ 是互斥事件，且

$$\{X_{n+1}=j\}=\{X_n=1, X_{n+1}=j\}\cup\{X_n=2, X_{n+1}=j\},$$

由概率的加法公式得

$$P(X_{n+1}=j)=P(X_n=1, X_{n+1}=j)+P(X_n=2, X_{n+1}=j).$$

另一方面，由条件概率的定义得

$$P(X_n=1, X_{n+1}=j)=P(X_n=1)P(X_{n+1}=j|X_n=1)=p_1^{(n)}P(X_{n+1}=j|X_n=1),$$

$$P(X_n=2, X_{n+1}=j)=P(X_n=2)P(X_{n+1}=j|X_n=2)=p_2^{(n)}P(X_{n+1}=j|X_n=2).$$

所以对于 $j=1, 2$，有

$$P(X_{n+1}=j)=p_1^{(n)}P(X_{n+1}=j|X_n=1)+p_2^{(n)}P(X_{n+1}=j|X_n=2). \tag{3}$$

因此，只要知道

$$p_{ij}=P(X_{n+1}=j|X_n=i), i=1, 2, j=1, 2,$$

就能由 $X_n$ 的分布列推出 $X_{n+1}$ 的分布列. 我们把 $p_{ij}$ 称为**从状态 $i$ 到状态 $j$ 的转移概率**，简称**转移概率**.

> 在一般的马尔可夫链中，条件概率 $P(X_{n+1}=j|X_n=i)$ 可能与 $n$ 有关，但是本专题总是讨论它与 $n$ 无关的简单情况.

例如，由（1）（2）可知，$p_{11}=P(X_{n+1}=1|X_n=1)=0.95$，$p_{12}=P(X_{n+1}=2|X_n=1)=0.05$，$p_{21}=P(X_{n+1}=1|X_n=2)=0.6$，$p_{22}=P(X_{n+1}=2|X_n=2)=0.4$.

对于表示机器工作状态的马尔可夫链 $\{X_n\}$ 而言，可以把 $X_0$，$X_1$，…，$X_n$，…看做机器工作状态的变化过程. 该过程可以分解为一系列的经过单位时间（这里是 1 小时）转变的子过程，而各个转移概率 $p_{ij}$ 恰好刻画了子过程中由状态 $i$ 经单位时间转化为状态 $j$ 的概率规律，这也是称 $p_{ij}$ 为转移概率的原因.

利用矩阵可以直观地表达马尔可夫链的各个状态间的转移概率，并且这种表达方式还为研究马尔可夫链的随机变化规律提供了方便. 例如，在表示机器运行状态的马尔可夫链 $\{X_n\}$ 中，$X_n$ 的分布列可以表示为

$$(p_1^{(n)}\quad p_2^{(n)}),$$

$X_{n+1}$ 的分布列可以表示为

$$(p_1^{(n+1)}\quad p_2^{(n+1)}),$$

$\{X_n\}$ 的两个状态间的转移概率可以用矩阵表示为

$$\boldsymbol{P}=\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}=\begin{bmatrix} 0.95 & 0.05 \\ 0.6 & 0.4 \end{bmatrix}.$$

于是，利用向量与矩阵的乘法运算（详见附录），可以把（3）表达为

$$(p_1^{(n+1)}\quad p_2^{(n+1)})=(p_1^{(n)}\quad p_2^{(n)})\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}. \tag{4}$$

一般地，对于一个马尔可夫链，像 $\boldsymbol{P}$ 这样的矩阵称为该马尔可夫链的**转移概率矩阵**，像 $(p_1^{(n)}\quad p_2^{(n)})$ 这样的行向量称为该马尔可夫链**在时刻 $n$ 的分布**，简称为**分布**.

转移概率矩阵由所有的转移概率构成，它刻画了马尔可夫链各个状态经单位时间相互转化的概率规律. 而时刻 $n$ 的分布则刻画了马尔可夫链 $\{X_n\}$ 中随机变量 $X_n$ 的概率分布规

律，该分布与 $X_n$ 的分布列相互唯一确定.

特别地，马尔可夫链在 0 时刻的分布称为初始分布. 初始分布刻画了马尔可夫链最初时刻($n=0$)的概率分布规律. 例如，在表示机器运行状态的马尔可夫过程中，初始分布 $(1\quad 0)$ 表示机器最初处于正常状态，初始分布 $(0\quad 1)$ 表示机器最初处于故障状态，初始分布 $(0.5\quad 0.5)$ 表示机器最初处于正常状态的概率为 0.5，处于故障状态的概率也为 0.5.

**思考**

一个马尔可夫链的初始分布是一个向量，该向量的各个分量之和等于多少？各个分量是否都是非负数？对于时刻 $n$ 的分布也有同样的结论吗？

**习题4.1**

1. 请举出几个马尔可夫链的例子.

2. 考察某工厂一台自动加工机的工作状态. 该机器的工作状态分为两种，正常状态和故障状态. 已知在正常状态下工作的机器，在 1 小时内出故障的概率为 0.01，而在故障状态下工作的机器，总是处于故障状态. 用 $X_n$ 表示机器在开始工作后第 $n$ 小时的工作状态.

(1) $\{X_n\}$ 是一个马尔可夫链吗？

(2) 如果 $\{X_n\}$ 是马尔可夫链，写出它的转移概率矩阵.

## 二 马尔可夫型决策简介

**案例 6** 某工厂一台自动加工机有两种工作状态：正常状态和故障状态. 在每个整数钟点的起始时刻检查机器的工作情况，若机器处于正常状态，则让它继续工作；若机器处于故障状态，则对它进行检修. 假设处于正常状态的机器，在 1 小时后发生故障的概率为 0.05. 对于故障机器有两种检修方案可供选择，一种是加急检修，在 1 小时内排除故障的概率为 0.9；一种是常规检修，在 1 小时内排除故障的概率为 0.6.

已知这台机器正常工作 1 小时可收益 10 元，加急检修 1 小时费用为 9 元，常规检修 1 小时费用为 6 元. 那么，当机器出现故障时，应选择哪种检修方案排除故障？

这是一个决策问题，决策目标是使机器的生产获得最大的收益. 可供选择的行动方案包括：

$$d_1\text{：加急检修，}d_2\text{：常规检修.}$$

机器在任何整数钟点时可能出现的状态包括：

$$h_1\text{：正常状态，}h_2\text{：故障状态.}$$

**思考**

这台机器可能出现的状态，和前两讲的决策问题中可能出现的状态有什么不同？

在前两讲中，决策问题中所出现的状态都与时间无关，而这里机器在第 $n$ 小时的工作状态 $X_n$ 与时间有关，且 $\{X_n\}$ 是一个马尔可夫链. 因此各个行动方案的收益也会随着时间的改变而改变. 为了获得最优决策，我们需要确定在各个整数钟点时机器的状态情况，进而确定相应的收益情况.

**思考**

上述马尔可夫链 $\{X_n\}$ 的转移概率矩阵与所采用的行动方案有关吗？

为了回答这个问题，只需考察马尔可夫链 $\{X_n\}$ 在行动方案 $d_1$ 和 $d_2$ 下的转移概率矩阵. 行动方案 $d_1$ 的转移概率矩阵为

$$\boldsymbol{P}_1=\begin{bmatrix}0.95 & 0.05\\ 0.9 & 0.1\end{bmatrix},$$

行动方案 $d_2$ 的转移概率矩阵为

$$\boldsymbol{P}_2=\begin{bmatrix}0.95 & 0.05\\ 0.6 & 0.4\end{bmatrix},$$

由此可见，马尔可夫链 $\{X_n\}$ 的转移概率矩阵与所采用的行动方案有关.

下面，我们考虑机器在时间段 $[n,\ n+1)$ $(n\geqslant 0)$ 内的收益情况. 当机器在 $n$ 时刻处于正常状态时，行动方案 $d_1$ 和行动方案 $d_2$ 在该时间段内的收益都是 10 元；当机器在 $n$ 时刻处于故障工作状态时，行动方案 $d_1$ 在该时间段内的收益是 $-9$ 元，行动方案 $d_2$ 在该时间段内的收益是 $-6$ 元. 因此，机器在时间段 $[n,\ n+1)$ 内的收益矩阵为

$$\begin{array}{cc} & \begin{array}{cc} d_1 & d_2 \end{array} \\ Q= & \begin{bmatrix}10 & 10\\ -9 & -6\end{bmatrix} \end{array}\begin{array}{c} \\ h_1\\ h_2\end{array}$$

在行动方案 $d_1$ 之下，利用公式 (4)，可以计算机器各个时刻的概率分布. 例如，当机器最初为正常工作状态时，初始分布为

$$(p_1^{(0)}\quad p_2^{(0)})=(1\quad 0),$$

时刻 1 的分布为

$$(p_1^{(1)} \quad p_2^{(1)})=(p_1^{(0)} \quad p_2^{(0)})\boldsymbol{P}_1=(1 \quad 0)\begin{pmatrix}0.95 & 0.05\\ 0.9 & 0.1\end{pmatrix}=(0.95 \quad 0.05),$$

时刻 2 的分布为

$$(p_1^{(2)} \quad p_2^{(2)})=(p_1^{(1)} \quad p_2^{(1)})\boldsymbol{P}_1=(0.95 \quad 0.05)\begin{pmatrix}0.95 & 0.05\\ 0.9 & 0.1\end{pmatrix}=(0.947\,5 \quad 0.052\,5).$$

对任何初始状态$(p_1^{(0)} \quad p_2^{(0)})$，利用公式（4）和矩阵的乘法运算的结合律（详见附录），可得

$$(p_1^{(n)} \quad p_2^{(n)})=(p_1^{(0)} \quad p_2^{(0)})\boldsymbol{P}_1^n,\ n=1,\ 2,\ \cdots, \tag{5}$$

即马尔可夫链在时刻 $n$ 的分布完全由初始分布和转移概率矩阵所决定.

因此，行动方案 $d_1$ 在时间段$[n,\ n+1)$内的平均收益为

$$Q(d_1,\ n)=(p_1^{(0)} \quad p_2^{(0)})\boldsymbol{P}_1^n\begin{pmatrix}10\\ -9\end{pmatrix}=(p_1^{(0)} \quad p_2^{(0)})\begin{pmatrix}0.95 & 0.05\\ 0.9 & 0.1\end{pmatrix}^n\begin{pmatrix}10\\ -9\end{pmatrix}.$$

类似地，行动方案 $d_2$ 在时间段$[n,\ n+1)$内的平均收益为

$$Q(d_2,\ n)=(p_1^{(0)} \quad p_2^{(0)})\boldsymbol{P}_2^n\begin{pmatrix}10\\ -6\end{pmatrix}=(p_1^{(0)} \quad p_2^{(0)})\begin{pmatrix}0.95 & 0.05\\ 0.6 & 0.4\end{pmatrix}^n\begin{pmatrix}10\\ -6\end{pmatrix}.$$

在第二讲中，我们按照平均收益最大准则选择最优决策．如果我们只关心机器在时间段$[n,\ n+1)$内的收益，就可以通过比较$Q(d_1,\ n)$和$Q(d_2,\ n)$来做出决策，这可以使机器在该单位时间段内获得最大平均收益．但是，这样获得的决策不一定能够保证在$[0,\ n+1)$时间段内获得最大平均收益．例如，当机器最初为正常工作状态时，初始分布为$(1 \quad 0)$，表 4-1 列出了$Q(d_1,\ n)$和$Q(d_2,\ n)$的取值情况：

**表 4-1**

| $n$ | $Q(d_1,\ n)$ | $Q(d_2,\ n)$ |
|---|---|---|
| 1 | 9.050 0 | 9.200 0 |
| 2 | 9.002 5 | 8.920 0 |
| 3 | 9.000 1 | 8.822 0 |
| 4 | 9.000 0 | 8.787 7 |
| 5 | 9.000 0 | 8.775 7 |
| 6 | 9.000 0 | 8.771 5 |
| 7 | 9.000 0 | 8.770 0 |
| 8 | 9.000 0 | 8.769 5 |
| 9 | 9.000 0 | 8.769 3 |
| 10 | 9.000 0 | 8.769 3 |

由表 4-1 可以看出，仅在时间段$[1,\ 2)$内，行动方案 $d_2$ 的平均收益大于行动方案 $d_1$ 的平均收益．按照机器在时间段$[1,\ 2)$内的平均收益最大准则选择行动方案，应该选择行动方案 $d_2$．但是如果机器运行的时间超过 4 小时，行动方案 $d_2$ 就不是平均收益最大的

方案了（此时行动方案 $d_2$ 在各个时间段内的平均收益之和小于行动方案 $d_1$ 的平均收益之和）.

思考

如果我们只关心机器在$[0, n+1)$内的收益，应该按照什么准则选择行动方案？

可以把$[0, n+1)$内各个单位时间的平均收益之和作为行动方案 $d_1$ 和行动方案 $d_2$ 在该时间段内的平均收益，然后按照平均收益最大准则选择行动方案. 行动方案 $d_1$ 和 $d_2$ 在该时间段内的平均收益分别为

$$Q_n(d_1)=\sum_{k=0}^{n}Q(d_1, k)\text{和}Q_n(d_2)=\sum_{k=0}^{n}Q(d_2, k).$$

取初始分布为$(0 \quad 1)$，表 4-2 列出了两个行动方案前 10 个小时的 $Q_n(d_1)$和 $Q_n(d_2)$的取值情况：

**表 4-2**

| $n$ | $Q_n(d_1)$ | $Q_n(d_2)$ |
|---|---|---|
| 1 | 19.050 0 | 19.200 0 |
| 2 | 28.052 5 | 28.120 0 |
| 3 | 37.052 6 | 36.942 0 |
| 4 | 46.052 6 | 45.729 7 |
| 5 | 55.052 6 | 54.505 4 |
| 6 | 64.052 6 | 63.276 9 |
| 7 | 73.052 6 | 72.046 9 |
| 8 | 82.052 6 | 80.816 4 |
| 9 | 91.052 6 | 89.585 7 |
| 10 | 100.052 6 | 98.355 0 |

由表 4-2 可以看出，行动方案 $d_2$ 能使机器在 2 小时或 3 小时内获得最大平均收益；行动方案 $d_1$ 能使机器在运行 4 小时或更长的时间内获得最大平均收益.

一般地，在各行动方案所对应的状态随着时间的推移而形成马尔可夫链的情况下，可以根据情况确定选择最优决策的准则，然后利用马尔可夫链的性质来计算各行动方案的平均收益（风险），以选择最优决策. 这样的决策方法称为**马尔可夫型决策方法**. 在马尔可夫型决策方法中，常用的选择行动方案的准则包括：

1. 短期准则：确定某一时刻 $n$，使得在$[0, n+1)$时间段内获得最大平均收益或最小风险；

2. 长期准则：使得在$[0, +\infty)$时间段内获得最大平均收益或最小风险.

如果考虑短期效益，应该用短期准则选择行动方案；如果考虑长期效益，则应该用长期准则来选择行动方案.

习题4.2

1. 请举一个需要用短期准则进行马尔可夫型决策的例子.

2. 在案例 6 中，假设在开始时刻机器处于正常工作状态. 如果用机器在 0～12 小时内的平均收益作为衡量准则，应该选择哪个行动方案？

## *三 长期准则下的马尔可夫型决策理论[1]

为利用长期准则进行马尔可夫型决策，需要做一些理论上的准备.

### 1. 马尔可夫链的平稳分布

思考

在案例 5 中，对于表示机器运行状态的马尔可夫链 $\{X_n\}$，时刻 $n$ 的分布随着 $n$ 的增加有什么样的变化规律？

由转移概率矩阵和给定的初始分布，可以反复利用公式（5）计算马尔可夫链在各个时刻的分布. 表 4-3 列出了 $\{X_n\}$ 在不同初始分布下，时刻 $n$ 的分布随着时间推移的变化情况.

由表 4-3 可知，当初始分布为(1　0)时，随着时间的推移，各个时刻的分布稳定于(0.923 1　0.076 9)；特别地，当 $n \geqslant 7$ 时，各个时刻的分布近似相等（这里用"近似相等"的原因是表中数据的精确度为小数点后 4 位）. 类似地，当初始分布为(0　1)或(0.5　0.5)时，各个时刻的分布也随着时间的推移越来越稳定于(0.923 1　0.076 9).

[1] 本节内容供选学.

表 4-3

| 初始分布 / 时刻 $n$ 的分布 / $n$ | (1　0) | (0　1) | (0.5　0.5) |
|---|---|---|---|
| 1 | (0.950 0　0.050 0) | (0.600 0　0.400 0) | (0.775 0　0.225 0) |
| 2 | (0.932 5　0.067 5) | (0.810 0　0.190 0) | (0.871 2　0.128 8) |
| 3 | (0.926 4　0.073 6) | (0.883 5　0.116 5) | (0.904 9　0.095 1) |
| 4 | (0.924 2　0.075 8) | (0.909 2　0.090 8) | (0.916 7　0.083 3) |
| 5 | (0.923 5　0.076 5) | (0.918 2　0.081 8) | (0.920 9　0.079 1) |
| 6 | (0.923 2　0.076 8) | (0.921 4　0.078 6) | (0.922 3　0.077 7) |
| 7 | (0.923 1　0.076 9) | (0.922 5　0.077 5) | (0.922 8　0.077 2) |
| 8 | (0.923 1　0.076 9) | (0.922 9　0.077 1) | (0.923 0　0.077 0) |
| 9 | (0.923 1　0.076 9) | (0.923 0　0.077 0) | (0.923 0　0.077 0) |
| 10 | (0.923 1　0.076 9) | (0.923 1　0.076 9) | (0.923 1　0.076 9) |
| …… | …… | …… | …… |
| 50 | (0.923 1　0.076 9) | (0.923 1　0.076 9) | (0.923 1　0.076 9) |
| …… | …… | …… | …… |
| 100 | (0.923 1　0.076 9) | (0.923 1　0.076 9) | (0.923 1　0.076 9) |

根据公式（4），可得

$$(0.923\,1\quad 0.076\,9)\approx(0.923\,1\quad 0.076\,9)\boldsymbol{P}.$$

两边同时乘以矩阵 $\boldsymbol{P}$，得

$$(0.923\,1\quad 0.076\,9)\approx(0.923\,1\quad 0.076\,9)\boldsymbol{P}^2.$$

重复上述过程，得

$$(0.923\,1\quad 0.076\,9)\approx(0.923\,1\quad 0.076\,9)\boldsymbol{P}^n,\ n=1,\ 2,\ \cdots.$$

由(5)可知，上式意味着：若以$(0.923\,1\quad 0.076\,9)$为初始分布，能使各个时刻机器运行状态的分布近似相等.

一般地，如果 $n$ 时刻的分布 $\boldsymbol{w}$ 能使马尔可夫链在 $n$ 以后各个时刻的分布相等，即

$$\boldsymbol{w}=\boldsymbol{w}\boldsymbol{P},\tag{6}$$

其中 $\boldsymbol{P}$ 为该马尔可夫链的转移概率矩阵，则称 $\boldsymbol{w}$ 为该马尔可夫链的一个**平稳分布**.

**思考**

对于表示机器运行状态的马尔可夫链$\{X_n\}$，它的平稳分布是否存在？如何求该马尔可夫链的平稳分布？

假设 $\boldsymbol{w}=(w_1 \quad w_2)$ 为 $\{X_n\}$ 的平稳分布，则由(6)和分布的性质知下面的关系成立

$$\begin{cases}(w_1 \quad w_2)=(w_1 \quad w_2)\begin{pmatrix}0.95 & 0.05\\ 0.6 & 0.4\end{pmatrix},\\ w_1+w_2=1,\end{cases}$$

并且 $w_1\geqslant 0$，$w_2\geqslant 0$. 即 $w_1$ 和 $w_2$ 恰为线性方程组

$$\begin{cases}0.95w_1+0.6w_2=w_1,\\ 0.05w_1+0.4w_2=w_2,\\ w_1+w_2=1\end{cases}$$

的非负解. 由于这个方程组有唯一的非负解 $w_1=\frac{12}{13}$，$w_2=\frac{1}{13}$，并且

$$\left(\frac{12}{13} \quad \frac{1}{13}\right)=\left(\frac{12}{13} \quad \frac{1}{13}\right)\begin{pmatrix}0.95 & 0.05\\ 0.6 & 0.4\end{pmatrix},$$

所以 $\{X_n\}$ 的平稳分布存在且唯一，为

$$\boldsymbol{w}=\left(\frac{12}{13} \quad \frac{1}{13}\right).$$

事实上，表 4-3 中的(0.923 1　0.076 9)是 $\boldsymbol{w}$ 的一个近似. 因此，表 4-3 中的数据可以解释为机器状态的分布随着时间的推移稳定于平稳分布 $\left(\frac{12}{13} \quad \frac{1}{13}\right)$.

一般地，若马尔可夫链的转移概率矩阵为 $\boldsymbol{P}$，则其平稳分布 $\boldsymbol{w}=(w_1 \quad w_2 \quad \cdots \quad w_m)$ 一定是下列方程组

$$\begin{cases}\boldsymbol{wP}=\boldsymbol{w},\\ w_1+w_2+\cdots+w_m=1\end{cases} \tag{7}$$

的非负解，并且反之也是正确的. 特别地，当(7)仅有唯一非负解时，相应马尔可夫链的唯一平稳分布恰为这个解.

**案例 7** 甲、乙两名同学进行一场趣味乒乓球比赛，约定：每打 1 球，给获胜者记 1 分. 当甲同学比乙同学多赢 2 分时，甲同学获胜；当乙同学比甲同学多赢 1 分时，乙同学获胜. 已知每打 1 球，甲同学赢 1 分的概率为 $p=0.75$，乙同学赢 1 分的概率为 $q=0.25$.

在打了 $n$ 个球后，若用 $Y_n$ 表示甲同学的得分与乙同学的得分之差，那么 $\{Y_n\}$❶是马尔可夫链吗？如果是，求其平稳分布.

❶在打了 $n$ 个球分出胜负的情况下，约定 $Y_{n+k}=Y_n$，$k=1, 2, \cdots$.

由题意知，$Y_n$ 的可能取值为 $-1, 0, 1, 2$. 在已知 $Y_n=i$ $(i=-1, 0, 1, 2)$ 的情况下，$Y_{n+1}$ 的分布与 $n$ 个球之前两名同学的得分情况无关，所以 $\{Y_n\}$ 具有马尔可夫性，即 $\{Y_n\}$ 为马尔可夫链，其状态空间为 $\{-1, 0, 1, 2\}$.

进一步地，设 $Y_{n+1}=j$ $(j=-1, 0, 1, 2)$，则从状态 0 到状态 $j$ 的转移概率为

$$p_{0j}=P(Y_{n+1}=j \mid Y_n=0)=\begin{cases}0.25, & j=-1,\\ 0.75, & j=1,\\ 0, & \text{其他}.\end{cases}$$

从状态 1 到状态 $j$ 的转移概率为

$$p_{1j}=P(Y_{n+1}=j|Y_n=1)=\begin{cases}0.25, & j=0,\\ 0.75, & j=2,\\ 0, & \text{其他}.\end{cases}$$

当 $Y_n=-1$ 时，乙同学获胜，比赛结束，这时可以认为 $Y_{n+1}=-1$，所以从状态 $-1$ 到状态 $j$ 的转移概率为

$$p_{-1j}=P(Y_{n+1}=j|Y_n=-1)=\begin{cases}1, & j=-1,\\ 0, & \text{其他}.\end{cases}$$

类似地，从状态 2 到状态 $j$ 的转移概率为

$$p_{2j}=P(Y_{n+1}=j|Y_n=2)=\begin{cases}1, & j=2,\\ 0, & \text{其他}.\end{cases}$$

因此，马尔可夫链 $\{Y_n\}$ 的转移概率矩阵为

$$\boldsymbol{P}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0.25 & 0 & 0.75 & 0\\ 0 & 0.25 & 0 & 0.75\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

下面求 $\{Y_n\}$ 的平稳分布 $\boldsymbol{w}=(w_1\quad w_2\quad w_3\quad w_4)$. 将转移概率矩阵的表达式代入 (7)，得

$$\begin{cases}w_1+0.25w_2=w_1,\\ 0.25w_3=w_2,\\ 0.75w_2=w_3,\\ 0.75w_3+w_4=w_4,\\ w_1+w_2+w_3+w_4=1.\end{cases}$$

解方程组，得其非负解为 $w_1=a$，$w_2=0$，$w_3=0$，$w_4=1-a$，其中 $a\in[0,1]$. 因此 $\{Y_n\}$ 有无穷多个平稳分布

$$(a\quad 0\quad 0\quad 1-a),\ 0\leqslant a\leqslant 1.$$

特别地，$(1\quad 0\quad 0\quad 0)$ 和 $(0\quad 0\quad 0\quad 1)$ 都是 $\{Y_n\}$ 的平稳分布.

由表 4-3 可以发现：对于给定的 3 个不同的初始分布，$\{X_n\}$ 的分布随着 $n$ 的增加迅速接近于平稳分布. 由此可以猜想：无论初始分布是什么，马尔可夫链的分布都会随着 $n$ 的增加而稳定于平稳分布. 下面，我们研究对于任何初始分布，马尔可夫链在时刻 $n$ 的分布是否都随着 $n$ 的增加稳定于同一分布.

先考察案例 7 中的马尔可夫链 $\{Y_n\}$. 如果以平稳分布 $(1\quad 0\quad 0\quad 0)$ 为初始分布，则任何时刻的分布都保持不变，都为 $(1\quad 0\quad 0\quad 0)$；如果以平稳分布 $(0\quad 0\quad 0\quad 1)$ 为初始分布，则任何时刻的分布都为 $(0\quad 0\quad 0\quad 1)$. 因此，对于不同的初始分布，随着时间的推移，$\{Y_n\}$ 在时刻 $n$ 的分布可能稳定于不同的分布. 这说明一个马尔可夫链在时刻 $n$ 的分布随着 $n$ 的增加可能稳定于不同的分布.

> 只有当马尔可夫链有有限个不同的状态时，此结论才成立.

概率统计学家证明了如下结论：如果一个马尔可夫链存在唯一的平稳分布 $\boldsymbol{w}$，则无论初始分布是什么，随着时间的推移，该马尔

可夫链在时刻 $n$ 的分布会稳定于唯一的平稳分布 $\boldsymbol{w}$.

在案例 5 中，表示机器状态的马尔可夫链$\{X_n\}$具有唯一的平稳分布，因此该马尔可夫链在时刻 $n$ 的分布总会稳定于同一平稳分布$\left(\frac{12}{13}\quad\frac{1}{13}\right)$；而在案例 7 中，表示甲乙两名同学乒乓球比赛结果的马尔可夫链$\{Y_n\}$的平稳分布不唯一，因此这个马尔可夫链在时刻 $n$ 的分布会稳定于不同的分布.

给定一个马尔可夫链，它有唯一的平稳分布的充分必要条件是方程组（7）有唯一的非负解. 因此，只要方程组（7）有唯一的非负解，则无论初始分布是什么，随着时间的推移，该马尔可夫链在时刻 $n$ 的分布都会稳定于唯一的平稳分布.

## 2. 平稳分布与马尔可夫型决策的长期准则

**思考**

在案例 6 中，如果利用长期准则，应该选择哪一个行动方案？

直观上，只要选择足够大的 $n$，使得机器在时间段$[0, n+1)$内平均收益最大即可. 但问题是什么样的 $n$ 算是足够大？由表 4-2 知：两个行动方案在时间段$[0, n+1)$内的平均收益 $Q_n(d_1)$和 $Q_n(d_2)$都随着 $n$ 的增大而增大，但是 $Q_n(d_1)$的增加幅度更大. 如果这种猜想是正确的，则只要机器运行时间足够长，行动方案 $d_1$ 的平均收益就更大，从而应该选择行动方案 $d_1$.

现在的问题是 $Q_n(d_1)$的增幅更大的猜想是否正确？在时间段由$[0, n)$增大到$[0, n+1)$的过程中，$Q_n(d_1)$的增幅为

$$Q(d_1, n)=Q_n(d_1)-Q_{n-1}(d_1),$$

$Q_n(d_2)$的增幅为

$$Q(d_2, n)=Q_n(d_2)-Q_{n-1}(d_2).$$

根据表 4-1 中的数据，可以猜想：当 $n$ 充分大时，$Q_n(d_1)$和 $Q_n(d_2)$的增幅应该分别稳定在 9.000 0和8.769 3的附近，即 $Q_n(d_1)$的增幅更大. 这个猜想是否正确，需要利用马尔可夫链的知识来判断.

下面先考虑选择行动方案 $d_1$ 的情况. 此时马尔可夫链$\{X_n\}$的转移概率矩阵为 $\boldsymbol{P}_1$，平稳分布 $\boldsymbol{w}=(w_1\quad w_2)$应该满足(7)，即

$$\begin{cases}0.95w_1+0.9w_2=w_1,\\0.05w_1+0.1w_2=w_2,\\w_1+w_2=1.\end{cases}$$

因为这个方程组存在唯一的非负解 $w_1=\frac{18}{19}$，$w_2=\frac{1}{19}$，所以在行动方案 $d_1$ 之下，$\{X_n\}$存在唯一的平稳分布$\left(\frac{18}{19}\quad\frac{1}{19}\right)$. 从而无论初始分布是什么，$\{X_n\}$在时刻 $n$ 的分布都会随着 $n$ 的

增大而稳定于$\begin{pmatrix}\frac{18}{19} & \frac{1}{19}\end{pmatrix}$. 因此

$$Q(d_1,\ n)=(p_1^{(0)} \quad p_2^{(0)})\boldsymbol{P}_1^n\begin{bmatrix}10\\-9\end{bmatrix}\approx\begin{pmatrix}\frac{18}{19} & \frac{1}{19}\end{pmatrix}\begin{bmatrix}10\\-9\end{bmatrix}=9,$$

即 $Q(d_1,\ n)$随着 $n$ 的增大越来越接近于 9.

类似地，在选择行动方案 $d_2$ 的情况下，马尔可夫链$\{X_n\}$的转移概率矩阵为 $\boldsymbol{P}_2$，前面已经得到此时的唯一平稳分布为$\begin{pmatrix}\frac{12}{13} & \frac{1}{13}\end{pmatrix}$. 因此，

$$Q(d_2,\ n)=(p_1^{(0)} \quad p_2^{(0)})\boldsymbol{P}_2^n\begin{bmatrix}10\\-6\end{bmatrix}\approx\begin{pmatrix}\frac{12}{13} & \frac{1}{13}\end{pmatrix}\begin{bmatrix}10\\-6\end{bmatrix}=\frac{114}{13},$$

即 $Q(d_2,\ n)$随着 $n$ 的增大越来越接近于$\frac{114}{13}\approx 8.769\ 2$.

因此，$Q_n(d_1)$的增幅大于 $Q_n(d_2)$的增幅，当机器运行的时间足够长时，应该选择行动方案 $d_1$，即“加急检修”.

通过上面的分析可以看出，当机器运行的时间足够长时，应该按照收益相对于平稳分布的加权平均最大的准则选择行动方案.

在长期准则中，平稳分布的作用相当于第一讲中介绍的状态分布. 在这种观点之下，用长期准则进行决策的方法就基本等价于第一讲介绍的风险型决策，只是这里的平稳分布需要利用马尔可夫性计算. 由于在长期准则中平稳分布起着重要的作用，因此也把长期准则称为**平稳准则**.

用平稳准则进行马尔可夫型决策的具体实施步骤如下：

> 第一步，根据问题的背景，确定所研究的对象是否可以用马尔可夫链描述. 如果可以，就明确问题的决策目标，然后执行第二步；否则，选用其他决策方法.
>
> 第二步，确定马尔可夫链所有可供选择的行动方案、可能出现的状态、损益函数或损益矩阵.
>
> 第三步，对于给定的行动方案，确定马尔可夫链在该行动方案下的转移概率矩阵，用 (7) 判断相应的平稳分布是否唯一.
>
> 第四步，如果各个行动方案所对应的平稳分布都唯一，就可以用平稳准则选择行动方案.

### 3. 平稳准则的应用案例

**案例 8** （广告方案的选择）某公司为了扩大一项产品的市场，以相同的成本制作了两个不同风格的电视广告，长度都是 20 秒. 为了了解广告的播出效果，该公司在两个地区进行了实验，发现顾客群的转移规律如下：

1. 在这两个地区没有播出广告之前，该公司的顾客群中有 40%转为其他公司的顾客；而其他公司的顾客群中有 20%转为该公司的顾客.

2. 在第 1 个实验区播出第 1 个广告一段时间后，该公司的顾客群中有 30%转为其他公司的顾客；而其他公司的顾客群中有 20%转为该公司的顾客.

3. 在第 2 个实验区播出第 2 个广告一段时间后，该公司的顾客群中有 20%转为其他公司的顾客；而其他公司的顾客群中有 10%转为该公司的顾客.

该公司面临的决策是：是否应该播出广告；如果决定播出，应该选用哪一个广告.

根据这个问题的实际背景，可以认为消费者对产品的选择随时间的变化而变化，具有马尔可夫性，并且我们需要考虑的是广告的长期效益. 共有 3 种行动方案可供选择：

$d_1$：不播出广告，$d_2$：播出第 1 个广告，$d_3$：播出第 2 个广告.

共有 2 种不同的状态：

$h_1$：顾客选择该公司的产品；$h_2$：顾客选择其他公司的产品.

我们可以认为，当顾客选择该公司的产品时，该公司就获得收益，否则该公司的收益为 0. 因此，收益函数可定义为

$$q(d_1,\ h_1)=1,\ q(d_1,\ h_2)=0;$$
$$q(d_2,\ h_1)=1,\ q(d_2,\ h_2)=0;$$
$$q(d_3,\ h_1)=1,\ q(d_3,\ h_2)=0.$$

收益矩阵为

$$\boldsymbol{Q}=\begin{bmatrix}1 & 1 & 1\\ 0 & 0 & 0\end{bmatrix}.$$

由顾客转移规律 1 知：当选择行动方案 $d_1$ 时，马尔可夫链的转移概率矩阵为

$$\boldsymbol{P}_1=\begin{bmatrix}0.6 & 0.4\\ 0.2 & 0.8\end{bmatrix}.$$

由顾客转移规律 2 知：当选择行动方案 $d_2$ 时，马尔可夫链的转移概率矩阵为

$$\boldsymbol{P}_2=\begin{bmatrix}0.7 & 0.3\\ 0.2 & 0.8\end{bmatrix}.$$

由顾客转移规律 3 知：当选择行动方案 $d_3$ 时，马尔可夫链的转移概率矩阵为

$$\boldsymbol{P}_3=\begin{bmatrix}0.8 & 0.2\\ 0.1 & 0.9\end{bmatrix}.$$

当选择行动方案 $d_1$ 时，平稳分布 $\boldsymbol{w}=(w_1\quad w_2)$满足方程组

$$\begin{cases}0.6w_1+0.2w_2=w_1,\\ 0.4w_1+0.8w_2=w_2,\\ w_1+w_2=1,\end{cases}$$

解方程组得唯一的平稳分布为$\left(\frac{1}{3}\quad\frac{2}{3}\right)$. 因此，行动方案 $d_1$ 所对应的平稳平均收益为

$$Q(d_1)=\left(\frac{1}{3}\quad\frac{2}{3}\right)\begin{bmatrix}1\\ 0\end{bmatrix}=\frac{1}{3}.$$

当选择行动方案 $d_2$ 时，平稳分布 $w=(w_1\quad w_2)$满足方程组

$$\begin{cases} 0.7w_1+0.2w_2=w_1, \\ 0.3w_1+0.8w_2=w_2, \\ w_1+w_2=1, \end{cases}$$

解方程组得唯一的平稳分布为$\begin{pmatrix}\frac{2}{5} & \frac{3}{5}\end{pmatrix}$. 因此，行动方案 $d_2$ 所对应的平稳平均收益为

$$Q(d_2)=\begin{pmatrix}\frac{2}{5} & \frac{3}{5}\end{pmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\frac{2}{5}.$$

当选择行动方案 $d_3$ 时，平稳分布 $w=(w_1 \quad w_2)$满足方程组

$$\begin{cases} 0.8w_1+0.1w_2=w_1, \\ 0.2w_1+0.9w_2=w_2, \\ w_1+w_2=1, \end{cases}$$

解方程组得唯一的平稳分布为$\begin{pmatrix}\frac{1}{3} & \frac{2}{3}\end{pmatrix}$. 因此，行动方案 $d_3$ 所对应的平稳平均收益为

$$Q(d_3)=\begin{pmatrix}\frac{1}{3} & \frac{2}{3}\end{pmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\frac{1}{3}.$$

由于 $Q(d_2)>Q(d_1)=Q(d_3)$，所以应该选用行动方案 $d_2$，即“播出第 1 个广告”.

**案例 9** （项目选址）某建筑公司的施工队长期分别在甲、乙、丙三个地区施工. 假设三个地区之间的相互距离相等. 施工所需的大型设备由建筑公司统一调配. 由以往数据可知，大型设备在三地的状态转移规律为：停留在甲地的概率为 0.8；由甲地转到乙地的概率为 0.2；由乙地转到甲地的概率为 0.2；由乙地转到丙地的概率为 0.8；停留在丙地的概率为 0.6；由丙地转到乙地的概率为 0.4.

如果该建筑公司想要建设一个大型设备修理厂，从长远角度考虑，应该在甲、乙、丙中的哪一处建厂？

显然，该公司的决策目标是从长远角度考虑，使运输大型设备的费用最少. 可供选择的行动方案包括：

$d_1$：建在甲地，$d_2$：建在乙地，$d_3$：建在丙地.

大型设备所处的状态包括：

$h_1$：大型设备在甲地，$h_2$：大型设备在乙地，$h_3$：大型设备在丙地.

我们先来考察行动方案 $d_1$ 所造成的损失情况. 当大型设备恰好是在甲地需要维修时，可以就地维修，不需要运输成本；当大型设备在乙地或丙地需要维修时，由于乙地和丙地与甲地的距离相等，因此需要花费同样的运输费. 这样，行动方案 $d_1$ 所对应的损失函数为：

$$l(d_1, h_1)=0, \ l(d_1, h_2)=1, \ l(d_1, h_3)=1.$$

类似地，行动方案 $d_2$ 所对应的损失函数为：

$$l(d_2, h_1)=1, \ l(d_2, h_2)=0, \ l(d_2, h_3)=1,$$

行动方案 $d_3$ 所对应的损失函数为：

$$l(d_3, h_1)=1, \ l(d_3, h_2)=1, \ l(d_3, h_3)=0.$$

该公司的大型设备所处的状态随时间的变化而变化，构成一个马尔可夫链. 又因为需要考虑长期的损失情况，所以应该采用平稳准则选择行动方案.

根据本例的条件，马尔可夫链的转移概率矩阵与所采用的行动方案无关，均为

$$\begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.2 & 0 & 0.8 \\ 0 & 0.4 & 0.6 \end{bmatrix},$$

其平稳分布$(w_1 \quad w_2 \quad w_3)$的各个分量满足方程组

$$\begin{cases} 0.8w_1+0.2w_2=w_1, \\ 0.2w_1+0.4w_3=w_2, \\ 0.8w_2+0.6w_3=w_3, \\ w_1+w_2+w_3=1, \end{cases}$$

解方程组得该马尔可夫链的唯一平稳分布为$\left(\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{2}\right)$. 所以，3 个行动方案所对应的平稳风险分别为

$$R(d_1)=\left(\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{2}\right)\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}=\frac{3}{4},$$

$$R(d_2)=\left(\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{2}\right)\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}=\frac{3}{4},$$

$$R(d_3)=\left(\frac{1}{4} \quad \frac{1}{4} \quad \frac{1}{2}\right)\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}=\frac{1}{2}.$$

因此，应该选用风险最小的行动方案 $d_3$，即“建在两地”.

## 习题4.3

1. 某台设备的运行情况分为 3 个状态：正常状态、小故障状态和大故障状态.

根据历史数据统计，正常状态工作的设备经过一个工作日后，出现小故障的概率为 0.18，出现大故障的概率为 0.02；小故障状态工作的设备经过一个工作日后，出现大故障的概率为 0.5. 对于设备的维修有两种方案可以采纳：

(1) 出现小故障就进行维修，需要花费 3 000 元维修费；

(2) 出现大故障才进行维修，需要花费 6 000 元维修费.

假设两种维修方案都能在一个工作日内使设备恢复到正常状态，仅从节约维修费用的角度考虑，应该选用哪种维修方案？

2. 在第 1 题中，若在一个工作日内，这台设备在正常状态下工作，可以获得 3 000 元的收益；在小故障状态下工作，可以获得 1 000 元的收益；在大故障状态下工作，没有收益. 那么应该选用哪种维修方案？

## 学习总结报告

同学们，学过了《风险与决策》，你都有哪些收获和体会？通过本专题的学习，你能解释下面这些问题吗？

1. “风险”是日常生活中常遇见的名词，它的含义为：遭受损失、伤害、不利或毁灭的可能性. 决策模型中的“风险”和日常生活中的“风险”有区别吗？“决策”也是日常生活中常遇见的一个名词，科学地进行决策需要考虑哪些因素？

2. 风险型决策是按照什么准则选择最优决策的？利用决策树进行决策的一般过程是什么？

3. 为什么要进行灵敏度分析？最优决策对状态概率分布变动的反应是否敏感，会影响最优决策的可靠性吗？

4. 在现实中，我们所研究对象的状态还可能随时间而变化，马尔可夫链是描述其中一类现象的概率模型. 怎样用马尔可夫链来描述这样一类现象，并做出决策呢？

对于现实生活中的一个决策问题，行动方案和状态的确定相对比较容易，而行动方案的评价准则要视问题的背景，依据概率统计的思想灵活确立. 对问题背景的不同理解可以建立不同的决策模型，从而得到不同的决策，哪一种理解更符合实际背景，相应的决策模型所得到的决策就能达到更好的效果.

数学学习中，只有不断地反思、归纳、总结，才能逐步领会数学的思想和研究方法. 风险与决策理论是现代数学应用的重要研究对象，其中蕴涵了丰富的概率统计思想方法，值得我们仔细地品味.

请你写一篇学习总结报告. 建议报告包括下列三方面的内容：

**1. 知识的总结**　如对本专题整体结构和内容的理解，对风险、收益、决策、马尔可夫性、转移概率矩阵、马尔可夫型决策等的认识；

**2. 拓展**　通过查阅资料、调查研究、访问求教、独立思考，进一步探讨决策模型在现实生活中的作用，尝试用该模型解决自己身边的学习、生活中的决策问题；

**3. 学习体会**　通过本专题的学习，谈谈你对研究随机现象的思想方法的认识，你对概率统计方法所得结论的认识.

风险与决策的进一步知识，同学们可以参考下列书籍：

1. 严武，程震源，李海东，《风险统计与决策分析》，经济管理出版社，1999 年；

2. James O. Berger（贾乃光译），《统计决策论及贝叶斯分析》，中国统计出版社，1998 年；

3. 董泽清，刘克，《马氏决策浅说》，辽宁教育出版社，1986 年.

## 附录

### 一、矩阵的概念

**定义** 由 $m\times n$ 个数 $a_{ij}$ 构成的一个用括号起来的 $m$ 行、$n$ 列的矩形数表

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

称为一个 $m\times n$ 矩阵，并称 $m$ 为该矩阵的行数，$n$ 为该矩阵的列数. 在 $m\times n$ 矩阵中，第 $i$ 行的数字组成的矩阵称为该矩阵的第 $i$ 行；第 $j$ 列的数字组成的矩阵称为该矩阵的第 $j$ 列；第 $i$ 行与第 $j$ 列的公共位置上的数字称为该矩阵第 $i$ 行第 $j$ 列的元. 称 $n\times 1$ 矩阵为 $n$ 维列向量，简称为列向量；称 $1\times n$ 矩阵为 $n$ 维行向量，简称为行向量；行向量或列向量统称为向量. 为了方便，称行（列）向量第 $i$ 个位置的元为该行向量的第 $i$ 分量.

例如，

$$\begin{pmatrix} 1 & 6 & 7 & 3 & 4 \\ 2 & 5 & 8 & 4 & 3 \\ 3 & 4 & 9 & 6 & 7 \end{pmatrix}$$

为 $3\times 5$ 矩阵，其第 1 行为 $(1 \quad 6 \quad 7 \quad 3 \quad 4)$，第 3 行第 1 列的元为 3. 又如

$$(1 \quad 6 \quad 7 \quad 3 \quad 4)$$

为 $1\times 5$ 矩阵，或 5 维行向量，其第 2 分量为 6. 而列向量

$$\begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

的第 2 分量为 5.

### 二、矩阵的加法

只有当两个矩阵的行数和列数分别相等时，才能进行加法运算. 设 $\boldsymbol{A}$，$\boldsymbol{B}$ 是两个矩阵，且它们的行数和列数分别相等，那么 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和仍然是一个矩阵，记作 $\boldsymbol{A}+\boldsymbol{B}$. $\boldsymbol{A}+\boldsymbol{B}$ 的行数和列数分别与 $\boldsymbol{A}$，$\boldsymbol{B}$ 的行数和列数相同，且第 $i$ 行第 $j$ 列位置的元等于 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的第 $i$ 行第 $j$ 列位置的元之和.

例如，$\boldsymbol{A}=\begin{pmatrix} 1 & 8 & 5 \\ -4 & 3 & -1 \end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix} 9 & -3 & -2 \\ 1 & 0 & 7 \end{pmatrix}$，则 $\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix} 10 & 5 & 3 \\ -3 & 3 & 6 \end{pmatrix}$.

### 三、矩阵的相等关系与乘法运算

如果矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 有相同的行数与列数，并且对应位置的元也相等，就称矩阵 $\boldsymbol{A}$ 等于矩阵 $\boldsymbol{B}$，记作 $\boldsymbol{A}=\boldsymbol{B}$.

如果矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{pmatrix},$$

则定义矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 之积 $\boldsymbol{AB}$ 为一个 $m\times k$ 矩阵，其第 $i$ 行第 $j$ 列的元为

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}, \quad 1\leqslant i\leqslant m, \quad 1\leqslant j\leqslant k,$$

即

$$\boldsymbol{AB}=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mk} \end{pmatrix}.$$

例如

$$(1 \quad 1)\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}=(1\times1+1\times(-1) \quad 1\times2+1\times3)=(0 \quad 5).$$

矩阵的乘法满足结合律，即若矩阵 $\boldsymbol{A}$ 的列数等于矩阵 $\boldsymbol{B}$ 的行数，而矩阵 $\boldsymbol{B}$ 的列数又等于矩阵 $\boldsymbol{C}$ 的行数，则

$$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC}),$$

并可将 $(\boldsymbol{AB})\boldsymbol{C}$ 简记作 $\boldsymbol{ABC}$.

例如，$\boldsymbol{A}=\begin{pmatrix} 1 & 8 & 5 \\ -4 & 3 & -1 \end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix} 9 & -3 \\ 1 & 0 \\ -2 & 7 \end{pmatrix}$，则

$$\boldsymbol{AB}=\begin{pmatrix} 1\times9+8\times1+5\times(-2) & 1\times(-3)+8\times0+5\times7 \\ (-4)\times9+3\times1+(-1)\times(-2) & (-4)\times(-3)+3\times0+(-1)\times7 \end{pmatrix}=\begin{pmatrix} 7 & 32 \\ -31 & 5 \end{pmatrix}.$$

在案例 3 中，有

$$(0.01 \quad 0.25 \quad 0.74)\begin{pmatrix} 3\,800 & 62\,000 & 60\,000 \\ 3\,800 & 2\,000 & 10\,000 \\ 3\,800 & 2\,000 & 0 \end{pmatrix}$$

$$=(0.01\times3\,800+0.25\times3\,800+0.74\times3\,800 \quad 0.01\times62\,000+0.25\times2\,000+0.74\times2\,000 \quad 0.01\times60\,000+0.25\times10\,000+0.74\times0)$$

$$=(3\,800 \quad 2\,600 \quad 3\,100).$$

如果矩阵 $\boldsymbol{P}$ 的行数与列数相等，则定义

$$\boldsymbol{P}^1=\boldsymbol{P},$$

$$\boldsymbol{P}^2=\boldsymbol{PP},$$

$$\boldsymbol{P}^3=(\boldsymbol{P}^2)\boldsymbol{P},$$

$$\cdots$$

$$\boldsymbol{P}^n=(\boldsymbol{P}^{n-1})\boldsymbol{P},$$

并称 $\boldsymbol{P}^n$ 为矩阵 $\boldsymbol{P}$ 的 $n$ 次方. 特别地，定义 $\boldsymbol{P}^0$ 为与 $\boldsymbol{P}$ 有相同的行数与列数的矩阵，且具有如下的形式

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

并称上述形式的矩阵为单位矩阵.

例 1　假设 $\boldsymbol{P}$ 为 $m\times m$ 矩阵，且

$$(p_1^{(n)} \quad p_2^{(n)} \quad \cdots \quad p_m^{(n)})=(p_1^{(n-1)} \quad p_2^{(n-1)} \quad \cdots \quad p_m^{(n-1)})\boldsymbol{P},\ n=1,\ 2,\ \cdots,$$

试证明

$$(p_1^{(n)} \quad p_2^{(n)} \quad \cdots \quad p_m^{(n)})=(p_1^{(0)} \quad p_2^{(0)} \quad \cdots \quad p_m^{(0)})\boldsymbol{P}^n,\ n=0,\ 1,\ 2,\ \cdots.$$

证明：用数学归纳法证明. 当 $n=0$ 时，有

$$(p_1^{(0)} \quad p_2^{(0)} \quad \cdots \quad p_m^{(0)})=(p_1^{(0)} \quad p_2^{(0)} \quad \cdots \quad p_m^{(0)})\boldsymbol{P}^0.$$

设结论对 $n$ 成立，即

$$(p_1^{(n)} \quad p_2^{(n)} \quad \cdots \quad p_m^{(n)})=(p_1^{(0)} \quad p_2^{(0)} \quad \cdots \quad p_m^{(0)})\boldsymbol{P}^n,$$

则由所给条件和归纳假设得

$$\begin{aligned}(p_1^{(n+1)} \quad p_2^{(n+1)} \quad \cdots \quad p_m^{(n+1)}) &=(p_1^{(n)} \quad p_2^{(n)} \quad \cdots \quad p_m^{(n)})\boldsymbol{P} \\ &=((p_1^{(0)} \quad p_2^{(0)} \quad \cdots \quad p_m^{(0)})\boldsymbol{P}^n)\boldsymbol{P},\end{aligned}$$

再由矩阵乘法的结合律得

$$(p_1^{(n+1)} \quad p_2^{(n+1)} \quad \cdots \quad p_m^{(n+1)})=(p_1^{(0)} \quad p_2^{(0)} \quad \cdots \quad p_m^{(0)})\,\boldsymbol{P}^{n+1}.$$

由数学归纳法原理知结论对任何非负整数 $n$ 成立.

例 2　如果 $w_i\geqslant 0$，$i=1,\ 2,\ \cdots,\ m$，$w_1+w_2+\cdots+w_m=1$，且

$$(w_1 \quad w_2 \quad \cdots \quad w_m)=(w_1 \quad w_2 \quad \cdots \quad w_m)\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix},$$

试证明 $w_1$，$w_2$，$\cdots$，$w_m$ 满足方程组

$$\begin{cases} p_{11}w_1+p_{21}w_2+\cdots+p_{m1}w_m=w_1, \\ p_{12}w_1+p_{22}w_2+\cdots+p_{m2}w_m=w_2, \\ \qquad\vdots \\ p_{1m}w_1+p_{2m}w_2+\cdots+p_{mm}w_m=w_m, \\ w_1+w_2+\cdots+w_m=1. \end{cases}$$

证明：由矩阵乘法的定义知

$$(w_1 \quad w_2 \quad \cdots \quad w_m)\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

为 $m$ 维行向量，其第 $j$ 分量为

$$w_1 p_{1j} + w_2 p_{2j} + \cdots + w_m p_{mj}, \ j=1, 2, \cdots, m.$$

再根据矩阵相等的定义有

$$w_j = w_1 p_{1j} + w_2 p_{2j} + \cdots + w_m p_{mj}, \ j=1, 2, \cdots, m,$$

再由题目所给条件知结论正确.

# 后　　记

为了全面贯彻党的教育方针，适应时代发展的需要，为学生的终身发展奠定基础，根据教育部制订的普通高中各学科课程标准（实验），人民教育出版社课程教材研究所编写的各学科普通高中课程标准实验教科书，得到了诸多教育界前辈和各学科专家学者的热情帮助和大力支持．在各学科教科书终于同课程改革实验区的师生见面时，我们特别感谢担任教科书总顾问的丁石孙、许嘉璐、叶至善、顾明远、吕型伟、王梓坤、梁衡、金冲及、白春礼、陶西平同志，感谢担任教科书编写指导委员会主任委员的柳斌同志和编写指导委员会委员江蓝生、李吉林、杨焕明、顾泠沅、袁行霈等同志．

我们聘请北京师范大学刘绍学教授为主编，与高中数学课程标准研制组的部分成员、大学数学教师、数学教育理论工作者、中学数学教研员和数学教师共同组成编写委员会，根据教育部制订的《普通高中数学课程标准（实验）》，编写了这套数学实验教科书．这里特别要感谢北京师范大学数学科学学院领导对本套教科书编写工作的高度重视和大力支持，同时还要感谢所有对本套教科书提出修改意见，提供过帮助与支持的专家、学者和教师，以及社会各界朋友．

本册教科书是编委会全体成员集体智慧的成果．除已列出的主要编写者外，参加本册教科书讨论的还有田载今等．

我们还要感谢使用本套教材的实验区的师生们．希望你们在使用本套教材的过程中，能够及时把意见和建议反馈给我们，对此，我们将深表谢意．让我们携起手来，共同完成教材建设工作．我们的联系方式如下：

电话：（010）58758316

E-mail：jcfk@pep. com. cn

人民教育出版社　课程教材研究所

中学数学课程教材研究开发中心

# 后　记