张尧庭　编著

全国文科数学教育研究会推荐教材

信息与决策

科学出版社

信 息 与 决 策

张尧庭　编著

科 学 出 版 社

2000

内 容 简 介

本书着重分析、介绍信息的概念与度量方法，信息与决策的关系，决策的依据和调整．全书共分四章．第一章从各个侧面来论述一些基本概念的实际背景；第二章是博奕论，重点放在一些基本问题的探讨和生动模型上；第三章是信息与信息决策，从分析信息内容到决策；第四章是一些典型的实例和模型．

本书是“全国文科数学教育研究会”推荐教材中的一本．读者对象为大学文、理、工科的本科生．

图书在版编目(CIP)数据

信息与决策/张尧庭编著．-北京：科学出版社，2000.1
ISBN 7-03-007752-0

Ⅰ.信… Ⅱ.张… Ⅲ.信息-关系-决策 Ⅳ.C934

中国版本图书馆CIP数据核字(1999)第29777号

科学出版社 出版
北京东黄城根北街16号
邮政编码：100717

新蕾印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

2000年1月第一版 开本：850×1168 1/32
2000年1月第一次印刷 印张：5 1/8
印数：1-3 800 字数：128 000

定价：11.00元

序　　言

信息与决策是现代人生活、工作离不开的两个对象，它们是那么重要，又那么难以掌握，似乎还没有一本教科书，把这二者联系在一起，本书就是要弥补这个空缺.

什么是信息？这一名词有多种含意，本书集中于对决策问题有关的信息，对信息的描述、结构、分类，它的价值、价格（从生产者、使用者各自不同的角度来考虑）应如何确定，都给以数学的描述和逐一地讨论. 把信息的概念逐步精确化，当然是一项非常困难的工作. 本书只是集中了一些国外材料上的看法，并加以整理、消化，也许还有作者错误的理解.

决策是与博奕论联系在一起的，没有博奕论的指导，要想作出一个好的决策，事实上是很难的. 博奕论中的各种模型，会对人们考虑问题提供有益的启迪，博奕论对解的寻求与探讨，使我们能不断深化对决策问题的认识，了解到每一种解的概念只是反映了一个方面的需求，要想寻找一个合于各种需求的解，看来是不行的. 而且在博奕论的研讨中，近代数学的公理化方法是非常普通的. 一方面显示了这种方法的力量所在；另一方面，也为我们学习、掌握这种方法提供了另一条途径. 博奕论中一些生动、精彩的例子会使我们大开眼界.

本书第一章是一个导引，从各个侧面来论述一些基本概念的实际背景；第二章介绍博奕论，重点放在一些基本问题的探讨和生动的模型；第三章介绍信息与信息决策，从分析信息内容到决策；第四章介绍一些典型的实例和模型.

希望本书能起到抛砖引玉的作用，作者期待今后会有这一方面更好、更新的教材.

本书的写作得到上海财经大学211工程教改科研项目的资助.

张尧庭

1999年4月

绪论 社会科学中的数学

现在，在社会科学中使用数学已被广泛接受了，只有少数顽固抵制的领域是除外的．原因不能从任何高深哲学的论战中去寻找，而是从许多简单的事实中就可明白．首先是社会科学的许多分支明显地在定量化，甚至可以说是被迫地定量化，人口统计学和经济是这方面明显的例子．其次，社会科学的主题，关于复杂系统的理论是用文字表达的，而它们的分析与比较用数学形式来表示会有很大的帮助．第三，除非这些理论关系可以定量化，否则它们的应用就只能依然是很一般化的．第四，对于一些主题中比较模糊、甚至很难得到确切信息的概念，数学可以提供一种领会的手段．最后，社会科学关注的不只是描述发生了什么，它们之间是如何连系的，还在于隐藏在背后的有效、还是无效的决策；在很大程度上，这些决策过程可以数学地表述和分析，使得我们的决策可以较多地依赖于知识，较少地依赖于推测．

已经证实，一些到目前还未见诸文字的想法，用数学概念来表达的可能性是行的，这就极大地克服了用数学的抵触情绪，越来越多的学者在他们研究的社会科学分支中体现出应用数学方法的价值．就现在情况而言，成功的进展是不平衡的，还未达到人们预期那样的一致．然而数学的运用已经普遍开展了，我们可以看到，许多不同领域的研究在下一代会积聚在一起，而现在还各自保持着差别并担心的专业．

大约75年以前，当伟大的美国经济学家欧文·费希尔撰写他的博士论文时，认为那时在全世界大约只有50本关于数量经济的书和一些论文，称得上是名符其实的．现在的情况完全不同了：不仅在经济学中，而且在所有社会科学中，数学书籍和论文每年要数以千计地出现．

在社会科学中，特别是涉及实证分析的，要大量采用有限数学的方法，尤其是矩阵、矩阵代数与差分方程，因为这些明显地与大部分实证研究基础的离散观测资料相吻合. 但这并不表示传统的工具，如微积分，特别是微分方程就没有用处了，在一些纯理论的分析中是会涉及的.

对于这个大而复杂的问题，最好的处理是先把确定性条件下的决策与不确定性条件下的决策区分开，然后在这两个主类之内，把单个阶段的决策与多个阶段的决策分开. 在这一节我们将论述一些对决策者有用的技术，主要在规划论、博奕论、统计决策论等等不同的领域之内. 这就再一次发现许多新的方法在很大程度上是使用有限数学的，虽然这许多决策问题，至少在理论上，正如我们将会看到的那样，是可以用过去的待定乘子法和变分法来解决的问题属于同一类型的.

(1) 在确定条件下的单阶段决策.

开始的一个好的例子就是消费者行为的理论，它是在 19 世纪末发展起来的. 根据这个理论，消费者有一个明确的偏好系统，他从所购买的商品中得到的效用（或满足）是依赖于各种商品和各种服务的数量. 他有一笔固定的钱，是他的收入，可供他花费，面对的是一组固定的各种价格. 他的目标是在收入和价格体系的约束条件下，使效用达到最大值.

这是一个在约束条件下求极大值的问题，可以用待定乘子法来解，并有一个几何的解释.

然而还需细加说明：偏好是会系统地改变的；人们要花费时间来适应正在改变的环境；人们也受到别人做了些什么的影响. 一种特别要注意的适应的方式是与新商品的介入相连的. 对新商品的反应包含着一个适应的过程，在许多方面与时尚流行相似. 接受的速度部分依赖于已拥有新商品的人数，部分依赖于还未拥有的人数. 这个适应的过程能用一个正向偏态的曲线，如对数正态积分，来很好近似地描述.

常常会遇到一个决策问题，虽然与刚才提到的多少有些类

似，但用经典方法是没有解的，因为求极大值（或极小值）的函数是线性的．这方面的一个例子是费用最小的食谱问题，求解必须用线性规则的现代方法．这个问题可叙述如下．一个合适的食谱是按营养成分的最低含量来确定的，营养成分有：热量、蛋白质、维生素等等．含有这些营养成分的各种食品都是可用的，它们的含量和价格都是已知的、固定的．问题是：每种食品应该买多少既能达到营养的要求又使花费最少．这等同于说：将这些食品的未知的购买量乘以他们的价格求和后达到最小，还需受一些不等式的限制．这些不等式中首先是所买的量要提供足够的每种营养成分，其次购买食品的量不能是负的．

（2）在确定条件下的多阶段决策．

采取决策时，常会考虑某一变量能有一个所要的随时间变化的路径该怎么办．这种情况要极大化（或极小化）的不是一个函数而是泛函．用于这种目的的经典方法是变分法．例如，让我们问这样的问题：一个社会最佳的储蓄是多少？若一个社会储蓄很少，相应的用于消费的就多些．于是短期内，可享受一种相对较高的生活水平；但长期看，资本设备增加很少，不能指望生活水平会有多大提高．因此我们必须问另一个问题：这个社会为了将来能得到更多，现在应该放弃多少？并且我们必须想到这个问题将持续地一直到难以确定多长的未来．

（3）不确定条件下的决策．

让我们考虑以下几点：首先，所谓不确定是由于不可控的事件引起的；其次，不确定是由于在冲突情况下别人的行为引起的；第三是已讨论过的方法的确定性条件被不确定的所代替．

与前面的例子相比，不确定性的作用是把一个假设已知的量代之以一个量的分布，现在我们的问题是找出分布的特性并判断它应如何影响我们的决策．这类问题将我们引入概率和统计的领域．

还应考虑决策的另一方面，避免一项大的损失比之于获得一项大的收益，个人可能会更在意些．于是在某些商业，他愿意承

担额外的费用来保持商品的库存，为了不让顾客等待，这并不是眼前的收益就能相抵，而是不想交货期弄得较长，最终会失去顾客．换句话说，正如决策者想对客观的概率有一个准确的估计一样，他也要对与他有关的主观评价有一个准确的估计．这就引向这一节的第二个内容，从冲突和联合中产生了不确定性．这个内容是另一种新的方法，博奕论的主题．

在一场博奕中，局中人不能确切地知道别的局中人将怎么办，而他的行动却依赖于他期望别人的反应而定的．至少在简单的情况下，我们可以制订出局中人可用的纯策略，并能给出任何特定的策略组合的结果如何．

虽然我们的特定模型是从社会生活中经济方面开始的，它有一种不可逆转的趋势蔓延开来．例如，在讨论生产商品和提供劳务中劳力、资本和制造性的作用时，就需要考虑另一方面，不同的技能将我们引向学习这些技能的教育和培训的系统；再一方面，研究和创新把我们引向社会心理学，最终我们将不得不面对社会——经济系统的全面分析．

目　　录

第一章 引 论

§1 什么是信息

信息这个名词现在是常常可以见到的，它的确切意思是指什么，却往往不易说清楚，而且信息这个名词在不同的场合表示的内容还不一样，所以这一节我们来叙述一下各种不同类型的内容.

信息既可以是一种知识，也可以是一件发生的事实，也可以是一条消息，所以信息这个概念不是那么明确的.另一方面，对讨论、研究问题的情况，由此作出决策的人而言，不同的信息，起的作用是很不相同的；即使是同一个信息，对于不同的人所起的作用也不见得相同.我们这里讨论的信息，总是与某一决策问题相连系来谈，否则就很难对信息的价值作出判断.

信息之所以能对决策有作用，它是通过对管理者(人)的信念产生影响来体现的，所以信息的作用可以这样描述：

信息⟶信念⟶决策

知识是信息的一种，它反映了客观规律或人们的共识，往往是用一种确定性的命题表现出来.例如市场的繁荣与萧条是由供需双方的状况来确定的，因此市场萧条时，应采取促销的方针，研究如何扩大需求量.知识是客观存在的，人们是否接受这种知识，也就是认为这个信息对他能不能有价值，不同的人是不一样的，因而决策也就不同.

消息也是一种信息，它可以是一件事实，也可以是一种传闻，它的真实性与知识相比就有差别.消息的另一个特点就是它的时间性，随着消息的传播面越来越广，它的价值往往也就越来越低.特别在现代社会，传播消息的手段——通讯非常发达，因此它的时效性更为突出.正因为消息可以是不真实的，这样就可以故意发布

一些欺骗性的消息来影响人们的信念，误导一部分人的行为，成为一种手段.正因为消息有时效性，少数人之间彼此沟通的消息与公众都知道的消息差别就很大，因此消息作为一种信息来处理，就必须评估它的真实性、时效性、公众程度.

还有一种信息，它本身就具有不确定性，例如明天下雨的概率是0.60，这种信息对不同的人起的作用不同，在很大程度上是各人对不确定性信息的态度不同.对这种不确定性信息的分析，往往是讨论决策方案的重点.

正因为信息对决策有用，决策的正确与否与企业经营的收益密切相关，这就产生了提供信息服务的机构.以提供信息服务为目的的产业已经成为现代社会不可缺少的部分，它们提供的信息，既可以给这个单位，又可以给那个单位，又具有私有性，又具有公开性.如果它提出的建议都是公众已知的，这种信息是不会有人去购买的，所以信息服务机构提供的信息又是一种类型.

信息还可以就它的来源分为直接信息(第一手信息)还是转送信息(第二手、第三手的信息).直接信息往往是一个事实——如某一货轮沉没、某一战役失败了，等等.转送的信息，是传递直接信息的消息，转送的内容、方式和时间会不同，因而同一个直接信息可以导致不同的转送的间接信息.这两种不同的信息在真实性、时效性上均有不同.

对于一个决策者，他希望得到的是真实的、最早的、直接的第一手信息，而实际上他能得到的往往是不确定的、间接的、具有一定时效的转手信息，这正是决策者的困难所在.

如何将上述各种信息用恰如其分的数学形式来描述，如何去评估各种信息的价值，有了信息怎样使用才能真正发挥它的作用，等等.这些问题都将提成数学的问题给以分析、讨论，这就是本书的主要内容，当然首先是如何用数学去描述这些问题.

也许有人会问：通常的信息论不是以信息作为对象来研究的吗？信息论是否已经回答了上面提到的种种问题呢？这一点也是要说清楚的.信息论泛指数学的信息论与通讯技术中的信息论，它

首先是由通讯技术需要发展起来的.从传送信息的角度来考虑,信息是靠码来传送的,如何编码,如何确定信息量,如何度量信息通道的容量大小,如何安排各种信息的传送,接收信息时如何抗干扰,等等,到现在发展到传送图像时如何压缩,如何接收时还原等等都是通信技术中迫切需要解决的问题,这些是工程信息论的内容.数学信息论是研究上述通讯技术中的数学问题.这两部分的内容与我们日常用语中的信息,还是有相当的不同,它们并不考虑信息的价值(连传送信息的费用也不是讨论的重点),也不讨论信息获得后,人的认识怎样转变.它与决策所需的信息还不是等同的,所以我们专门有一节来讨论信息的结构,这是与决策紧密相连的.尽管如此,在这一方面,还存在着不少问题是值得研究的.

工程信息论中有两个概念对于我们是有用的,一个是熵、一个是信息量.

假设在某种条件下,可能发生的事件是 k 个,用 $A_1, A_2, \cdots, A_k$ 来表示,A_i 发生的概率是 p_i, $i=1,2,\cdots,k$.用什么来度量这种情况的不确定性呢?熵就是这样的概念.当 p_i 很小时,它的不确定性就比较大,当 p_i 接近于 1 时,它的不确定性就小,因此 $\ln\frac{1}{p_i}$ 就是一个很好的量.当 $p_i \to 0$ 时,$\ln\frac{1}{p_i} \to \infty$;当 $p_i \to 1$ 时,$\ln\frac{1}{p_i} \to 0$.于是"平均"的 $\ln\frac{1}{p_i}$,即按出现可能的概率大小 p_i 来加权平均,就得到平均的不确定性,它就是

$$-\sum_{i=1}^{k} p_i \ln p_i.$$

在信息论中,它称为熵(entropy),记为 H.我们看简单的情形,$k=2$时,熵就是

$$-p\ln p-(1-p)\ln(1-p).$$

当 $p=\frac{1}{4}$ 时,$H=\frac{1}{4}\ln 4+\frac{3}{4}\ln\frac{4}{3}=\ln 4-\frac{3}{4}\ln 3$;当 $p=\frac{1}{2}$ 时,熵 $H=\frac{1}{2}\ln 2+\frac{1}{2}\ln 2=\ln 2>\ln 4-\frac{3}{4}\ln 3$.可以证明,$p=\frac{1}{2}$ 时,熵最大.

概率越分散,不确定性就越大,这是符合人们直观的想法的.对任一 k,均匀分布的熵最大,这也是可以证明的.

如果我们获得了信息,此时概率分布就会改变,用 H 表示原来的熵,$H(I)$就是信息 I 已知后相应的熵,很明显,$H(I)\leqslant H$,否则信息 I 是没有意义的.信息 I 的作用就反映在熵的减少上,因此 $H-H(I)$就称为信息 I 提供的信息量(information).例如原来认为一枚硬币是均匀的,正反面的概率都是$\frac{1}{2}$,相应的熵是 $\ln 2$;有人提供信息 I,知道硬币不均匀,正面向上的概率是$\frac{1}{4}$,反面是$\frac{3}{4}$,相应的熵 $H(I)=\ln 4-\frac{3}{4}\ln 3$,这个信息 I 提供的信息量是

$$\begin{aligned} H-H(I) &= \ln 2-\ln 4+\frac{3}{4}\ln 3 \\ &= \frac{3}{4}\ln 3-\ln 2 \\ &= 0.82396-0.69315 \\ &= 0.12081. \end{aligned}$$

这就对信息提供了一种度量大小的方法.

§2 信息与信念

信息对决策的影响是通过信念而起作用的,人们对事物原有一些看法,形成了自己的信念,获得了有关的信息后,信念就发生了改变,这种改变就明确地体现了信息的作用.能否有一种方式把这样一种转变用数学表示出来呢?最容易被大家接受的数学公式就是条件概率和由它导出的贝叶斯公式.

概率与信念是不是一回事,会有各种看法,然而用概率这个数学工具来反映人们对某些事物的信念程度,这是不少人都可以接受的.人们对某一事件发生的可能性大小用它的概率来描述,事件 A 发生的概率用 $P(A)$表示.当人们知道某一事件 B 已经发生后,这时 A 发生的可能性就用条件概率$P(A|B)$来表示,从$P(A)$

转变为 $P(A|B)$,这就是人的认识发生了变化,信念也随着就变了. $P(A|B)$的计算公式

$$P(A \mid B) = P(AB)/P(B). \tag{2.1}$$

这是第一个描述了由于知道"B 已经发生"这个信息后,人们对 A 发生的可能性调整了看法的公式. 从(2.1)不难得到

$$P(AB) = P(B)P(A \mid B), \tag{2.2}$$

因而自然同样有

$$P(AB) = P(A)P(B \mid A). \tag{2.3}$$

把(2.2)、(2.3)的右端相等,就导出贝叶斯公式:

$$P(B \mid A) = \frac{P(B)P(A \mid B)}{P(A)}. \tag{2.4}$$

可见贝叶斯公式实际上是与条件概率等同的一个表达式. 当然(2.4)也可以写成

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)},$$

这两个式子是一样的内容. 实际上,我们可以将(2.4)式改写成另一种形式,它更能显示出"事件 B 已经发生"这个信息对信念的改变是如何起作用的. 将 A 的逆事件 $\bar{A}$(A 不发生)代替(2.4)中的 A,于是对 A 和 $\bar{A}$ 有两个等式

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)},$$

$$P(\bar{A} \mid B) = \frac{P(\bar{A})P(B \mid \bar{A})}{P(B)}.$$

将两式相除,就得到

$$\frac{P(A \mid B)}{P(\bar{A} \mid B)} = \frac{P(A)}{P(\bar{A})} \frac{P(B \mid A)}{P(B \mid \bar{A})}. \tag{2.5}$$

(2.5)式清晰地显示了人的认识是如何调整的. 原来人们认为 A 发生的可能性大小是 $P(A)$,它不发生的可能性大小是 $P(\bar{A})$,注意到 $P(A)+P(\bar{A})=1$,因此 $P(A)/P(\bar{A})$就充分反映了人原来的认识. 在知道"B 已经发生"这一信息后,人的认识成为(2.5)式左端 $P(A|B)/P(\bar{A}|B)$,它与原来的认识$\dfrac{P(A)}{P(\bar{A})}$的差别就反映了

“B 已经发生”这个信息的作用. 调整的方法,就是将$\frac{P(A)}{P(\bar{A})}$乘以(2.5)右端的第二项 $P(B|A)/P(B|\bar{A})$,这一项通常称为贝叶斯因子(Bayes factor). 值得注意的是:无论是 $P(A)/P(\bar{A})$ 还是 $P(A|B)/P(\bar{A}|B)$,它们都是相同条件下的概率比,而且分子与分母的和总是 1;而贝叶斯因子 $P(B|A)/P(B|\bar{A})$是不同条件下的概率比,分子与分母之和不一定是 1.

贝叶斯因子的表达式很合于人们的直觉,如果 A 发生时B 发生的概率$P(B|A)$比 A 不发生时B 发生的概率$P(B|\bar{A})$大,那么 B 发生时A 发生的概率就似乎应该比A 不发生的概率大. 这种考虑问题的方法被著名的数学家波里亚称之为“合情推理”,贝叶斯公式在一定意义上使合情推理严格化,给出了条件和公式.

现在用一个简单的例子来说明这一点. 在一个地区,患肝炎的人占的比例是 0.05,用 A 表示某人有肝炎,此时 $P(A)=0.05$, $P(\bar{A})=0.95$,因此$\frac{P(A)}{P(\bar{A})}=\frac{5}{95}=1/19$. 作体格检查,检查反应为阳性用 B 表示,已知肝炎患者呈阳性的概率为 0.99,即 $P(B|A)=0.99$,而非肝炎患者呈阳性的概率是 0.02,即 $P(B|\bar{A})=0.02$. 若体格检查后,某人呈阳性,那么他患肝炎的概率就是条件概率 $P(A|B)$,用公式(2.5)

$$\begin{aligned}\frac{P(A\mid B)}{P(\bar{A}\mid B)}&=\frac{P(A)}{P(\bar{A})}\,\frac{P(B\mid A)}{P(B\mid \bar{A})}\\&=\frac{5}{95}\,\frac{99}{2}=\frac{495}{190}\\&\doteq 2.65.\end{aligned}$$

这明显看出,此人患肝炎的可能从 0.05 增加到 0.73,体格检查提供的信息作用是很大的.

如果体格检查结果是阴性,此时患肝炎的概率为 $P(A|\bar{B})$,用(2.5)得

$$\frac{P(A\mid \bar{B})}{P(\bar{A}\mid \bar{B})}=\frac{P(A)}{P(\bar{A})}\,\frac{P(\bar{B}\mid A)}{P(\bar{B}\mid \bar{A})}$$