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

（如有印装质量问题，我社负责调换〈新欣〉）

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:

$$\text{信息}\longrightarrow\text{信念}\longrightarrow\text{决策}$$

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 $\ln2$;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 $H(I)=\ln4-\frac{3}{4}\ln3$,这个信息 I 提供的信息量是

$$\begin{aligned} H-H(I) &= \ln2-\ln4+\frac{3}{4}\ln3 \\ &= \frac{3}{4}\ln3-\ln2 \\ &= 0.82396-0.69315 \\ &= 0.12081. \end{aligned}$$

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$\frac{P(A)}{P(\bar{A})}$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$\dfrac{P(A)}{P(\bar{A})}$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$\dfrac{P(A)}{P(\bar{A})}=\dfrac{5}{95}=1/19$.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

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

$$=\frac{5}{95}\frac{2}{98}$$

$$\doteq 0.001.$$

他患肝炎的可能大大降低.从这个例子明显看出贝叶斯公式对我们作出决策有多大的作用.

信念是主观的，然而它可以有客观的依据，例如这个地区肝炎患者占0.05，那么我们在这个地区遇上一个肝炎患者的可能性是0.05. 我们的主观信念是由客观的数据来支持的. 这就如同我们判断一枚硬币掷一次出现正、反面的机会是一样的，各为1/2. 这是因为我们从逻辑上没有理由断定哪一面有更大的概率出现，从而得到的这种信念. 所以信念的得来有许多途径，有些是客观的规律转化的，有的是从别人的知识、经验、信念中演变得来的. 这一转化是单向的，客观的可以转化成主观的，主观的信念并不会转化成客观的规律. 别人的主观信念可以影响某人的信念，许多人的共同信念会影响社会的发展，但不会转变成客观规律. 在经济现象中比较明显，消费者的共同的心态对商品的市场价格会有影响，这种影响也是一种客观规律，但不能认为是主观的信念转化为客观的事实.

要注意的是信息与它的真实性往往是同时出现的,这表示告诉我们一个信息时,伴随着它的真实性有多少,这一点在以后几章的讨论中会不断以各种数学方式来描述,读者要注意如何去描述,是否合适,是否充分.

§3 效用与决策

在人们对事物有了自己的看法之后,决策自然是由信念来支配了.然而,一个好的决策不是单纯凭借信念来确定.决策应对一个目标考虑,这个目标如果是量化的,就可以考虑最优的策略,在社会、经济的讨论中,通常用效用函数来表示量化的目标,因此一个决策问题就成了一个优化的问题.

由于信息具有不确定性，相应的决策就变得复杂起来，一个好的决策能面对不同的结果，不同结果的概率，在整体上有一个“优”的结局. 这样就需要一些数学的概念和符号，把这个问题表述清楚. 可能出现的不同结果用 $S_1, S_2, \cdots, S_m$ 表示，这些 S_i 出现的可能性分别用 π_i 来表示，$i=1,2,\cdots,m$，$\{S_i\}$ 是由于我们采取了某个“措施”x 所引起各种可能结果，不同的 x 影响各个 S_i 出现的概率 π_i，因而 π_i 是措施 x 的函数，记为 $\pi_i(x)$. $\pi_i(x)$ 反映了三个内容：x 是措施，i 是可能出现的结果 S_i，$\pi_i(x)$ 是 S_i 出现的概率.

只知道 $\pi_i(x)$，我们还难以选择，因为 x 的好坏无法度量. 这时要引入效用函数(utility function) $U(S_i)$，它反映了出现结果 S_i 时的效用，效用的大小反映了结果的好坏. 既然 $U(S_i)$ 越大越好，自然希望出现 $U(S_i)$ 达到最大的那个结果，设为 S_*，即有

$$U(S_*) = \max_{1\leqslant i\leqslant m} U(S_i). \tag{3.1}$$

但是 S_* 出现的概率 $\pi_*(x)$ 是随 x 而变的，这就可以导出一种决策，在所有可以选择的措施 x 中使 $\pi_*(x)$ 达到最大的，就是一种好的决策，也就是在全体可能的措施组成的集合 X 中，寻找 x_* 使 $\pi_*(x_*)$ 达到最大，即有

$$\pi_*(x_*) = \max_{x\in X} \pi_*(x). \tag{3.2}$$

这样求得的 x_* 看来似乎不错，然而细想一下，就会感到还存在问题. 因为采取 x_* 之后，$\pi_*(x_*)$ 是所有 $\pi_*(x)$ 中最大的，但它仍然可以是 $\pi_i(x)$ 中很小的，甚至是最小的. 这可以从下面简单的例中直接看出(见表 3.1).

这就启发我们应作更深入的考虑，直接简单地比较效用函数是不合适的.

这使我们进一步考虑期望效用，这是把不确定因素用确定的数值来代替的一种想法，使得决策更便于考虑. 当 x 选定后，我们面对的状况是：

表 3.1

	$i=1$ S_1	$i=2$ S_2	$i=3$ S_3	
$x=1$	0.36	0.04	0.60	$x_*=1$
$x=2$	0.58	0.02	0.40	
$U(i)$	3	5	2	

注表中的数值是 $\pi_i(x)$; * =2

结局	S_1	S_2	$\cdots$	S_m
概率	$\pi_1(x)$	$\pi_2(x)$	$\cdots$	$\pi_m(x)$
效用	$U(S_1)$	$U(S_2)$	$\cdots$	$U(S_m)$

当 x 选定后,$U(S_1),U(S_2),\cdots,U(S_m)$是 m 个可能的效用,它们是数值,但出现这些数值的概率是不同的,按概率的大小来加权平均,就得到期望效用(expected utility)

$$U(x)=\sum_{i=1}^{m}U(S_i)\pi_i(x)$$
$$=U(S_1)\pi_1(x)+\cdots+U(S_m)\pi_m(x).$$

期望效用是一个平均值(加权得来的),比较不同 x 相应的期望效用$U(x)$,就可以获得期望效用最大的决策 $\tilde{x}$,即

$$U(\tilde{x})=\max_{x\in X}U(x).$$

期望效用实际上是一种长期效益的标准.从概率论的知识知道,一个随机变量的期望值,一次观察是很难实现的,但是长期、大量、重复的观察,实际观察数值的算术平均是很接近期望值的,这就是大数定律.

有了期望效用,就可以衡量信息的效用.设想信息 A 的获得,改变了我们的看法,原来 x 措施相应的各个结果出现的概率为 $\pi_i(x)$,现在认为是 $\pi_i(x|A)$,这样 $\pi_i(x|A)$相应的期望效用就是

$$U(x\mid A)=\sum_{i=1}^{m}U(S_i)\pi_i(x\mid A)$$
$$=U(S_1)\pi_1(x\mid A)+\cdots+U(S_m)\pi_m(x\mid A).$$

$U(x|A)$的最大值在 $\bar{x}(A)$达到,这表示极大值点 $\bar{x}$ 与信息A 是有关系的,随 A 而变,有

$$U(\bar{x}(A)) = \max_{x\in X} U(x \mid A).$$

期望效用 $U(\bar{x}(A))$与 $U(\bar{x})$的差别就体现了信息 A 的效用,用通俗的语言来说,它就反映了信息 A 的“价值”.这种想法的进一步细微的讨论在以下几章中逐步进行,它表明了信息的价值是可以度量的.

有了效用、期望效用,使决策问题更明确了,但是决策问题的风险没有得到反映,信息可以影响效用,信息能否减少风险,用什么方式来反映,这就是信息与决策联系起来的重要内容.

§4 风险与决策

有了期望效用,把一个不确定性的决策转化为一个比较期望效用的决策,实质上的不确定性并未排除,因为决策实施后,所得的效用并不是期望效用,可能好,也可能差,这就是风险.伴随着决策是它的效用和风险,效用在上一节已讨论了,这一节讨论风险.

为了便于说明,我们把效用看成“价值”,于是风险就与损失相联系.风险的产生是由于效用的不确定性,因此可以把效用 u 看成是随机变量,一个决策、措施 x,就确定了 u 的一个概率分布,u 的期望效用就是u 对于这个分布的期望值,用概率论中常用的符号来表示,就可以写成 $E_{\pi(x)}(u)$,它表示按概率分布 $\pi(x)$对随机变量 u 来加权平均(取期望).当 u 与 m 个结果$s_1,\cdots,s_m$ 有关时,u 的取值为

$$u(s_1),\cdots,u(s_m).$$

相应于概率 π 的分布是$P(u=u(s_i))=\pi_i(x)$,因此

$$E_{\pi(x)}u = \sum_{i=1}^{m} u(s_i)\pi_i(x).$$

如果 u 不是只取有限个值,$\pi(x)$就可以是一个分布,$E_{\pi(x)}u =$

$\int ud\pi(x)$ 是一般的一个积分，可以概括更多更复杂的情形.

总结一下，我们用随机变量 u 描述效用的数值，$\pi(x)$表示采用措施 x 后 u 的分布，期望效用就是 u 在 $\pi(x)$下的数学期望 $E_{\pi(x)}(u)$. 效用自然是越大越好. 有了这些概念，可以从承担风险的多少来比较措施 x 的好坏.

什么是风险，人们的直觉是“风险反映的是损失的可能性”，因此风险的大小应与损失的大小、损失的可能性相联系. 这一点怎么能用上面提到的效用来表示呢？效用是正面的评价，损失是负面的评价，两者之间的联系实际上是一个函数与另一个函数之间的转换关系.

如果效用 u 反映的是收益，人们预期应得的收益(如长期存款的收益)是 c，$u-c$ 反映实际与预期的差别. 当 $u-c>0$ 时，没有任何损失；当 $u-c<0$，就导致了损失，于是损失 l 可以定义为

$$l=\begin{cases}0\ , & \text{当 } u\geqslant c,\\ c-u, & \text{当 } u<c.\end{cases}$$

l 取什么值的可能性与效用 u 取值的可能性相关，这就明显地看出了它们之间的关系. 因此，我们的讨论仍然是用效用函数的术语，而不是用损失函数的术语.

确立了用效用来表示的术语后，风险与决策的关系就在于措施 x 相应的概率分布影响到 u 的期望值——期望效用. u 是一个随机变量，它的不确定性完全被它的分布函数 F 所描述，F 的定义是：$-\infty<t<\infty$，

$$F(t)=P(u<t)=\text{“}u<t\text{”发生的概率}.$$

当措施 x 不同时，相应的 u 的分布用 $F_x(t)$表示，于是期望效用

$$E_{F_x}(u)=\int t\mathrm{d}F_x(t)=\int tf_x(t)\mathrm{d}t$$

$f_x(t)$称为分布函数 $F_x(t)$相应的密度，$F_x(t)$和 $f_x(t)$总是在$(-\infty,\infty)$上定义的，所以积分号不写上、下限，表示在整个实轴上积分.

现在来比较措施之间的差别.若有措施 x_1 与 x_2,分别对应于分布 $F_1(t)$和 $F_2(t)$,即

$$F_{x_1}(t) = F_1(t), F_{x_2}(t) = F_2(t).$$

当 $F_1(t) \leqslant F_2(t)$对每个 t 都成立时,就有“$u < t$”发生的概率在 x_1 时比 x_2 时小,也就是“$u \geqslant t$”在 x_1 发生的概率比 x_2 大,用符号表示

$$P_{F_1}(u < t) \leqslant P_{F_2}(u < t),$$

或

$$P_{F_1}(u \geqslant t) \geqslant P_{F_2}(u \geqslant t).$$

这表示 x_1 使效用大的可能性总比 x_2 相应的大,因此 x_2 是不可取的.当上述不等号对某个 t 是真不等时,就可以确认 x_1 比 x_2 好.我们就说 x_2 被 x_1 所控,或者说 x_2 是不允许的(因为确有比它好的 x_1 存在).这样,我们应该选取措施 x,它相应的分布 $F_x(t)$是允许的,这是一个很低的要求.

对每个分布 $F_x(t)$,它反映了效用 u 的取值情况,现在来度量它的风险.风险是 u 值的不确定性,如果 u 值是完全确定的,就谈不上风险,因此风险可以用 u 值的不确定程度来度量,在数学上就是随机变量 u 的分散程度,于是概率论中一些概念均可引入来反映风险的大小,常见的有下述几项:

(i)u 的方差、标准差.即

$$\mathrm{Var}(u) = E(u - Eu)^2 \triangleq^{1)} \sigma^2(u),$$

$\sigma^2(u)$是 u 的方差,$\sigma(u)$是 u 的标准差.

(ii)u 的一阶绝对中心矩.即

$$E\mid u - Eu \mid.$$

(iii)用 u_α 表示$P(u < u_\alpha) = \alpha$ 的解,u_α 称为u 的α 分位点(数),$0 < \alpha < 1$,$(u_{0.25}, u_{0.75})$反映了 u 的值有 1/2 的概率在这个区间内,因此

---

1)≜表示“记成”,即左边记成右边,或相反.

$$u_{0.75} - u_{0.25} = d$$

反映了 u 的分散程度.

(iv)可以用 u 相应的熵来表示. u 的分布密度设为 $p(x)$,则 u 的熵 $H(u)$ 为

$$H(u) = -\int p(x)\ln p(x)\mathrm{d}x.$$

在实际问题中,常用的是标准差.

§5 信息、博奕与决策

本世纪中叶,博奕论(theory of games)诞生了.它的影响非常深远,从决策的观点来看,它把一些不确定性的问题作了更深一步的剖析.博奕论也译成对策论,无论从原文的意义和实质性的内容,前者的译名比后者要准确些,所以今后我们采用博奕这一名词.

社会上无论是个人、集团,还是政府,每天都面临着不少必须处理的决策问题,而掌握的信息又不多,真正能够"知己知彼"的,实在很少,在这种情况下作出的决策,后果的不确定性就难以避免.博奕论就是在这种情况下寻求合理的决策,尽可能避免大的损失,争取好的结局.

人们一开始接触博奕时,容易有错觉,认为博奕的各方一定是有利害冲突的.从博奕论的发展来看,也是如此,一开始二人零和博奕(即一方所失就是另一方所得)是研究的中心,后来转向多人博奕,考虑联合,再后来就认识到并不总是冲突的,纳什平衡的概念提出之后,博奕论就到了另一个阶段.

博奕论的讨论基于两条:局中人都是理性的,他的目标非常明确,就是使自己的效用达到最大(尽管在现实生活中这种做法是不实际的,但理论研究必须是这样).其次,博奕论中的例子是简单而很不实际的,但是它比一些具体实际的复杂的例子更能揭示实质,使得很多人即使不去学习博奕论的理论,也能理解这些例子中

提出的问题和分析的方法,这是有指导意义的.所以我们在学习博奕论的知识时,要注意这些简单而典型的例子,学习分析问题,提出概念,解决问题的过程.

要注意使效用最大并不一定是使自己所得的财富最多,因为效用与财富之间的关系是非线性的,考虑到风险之后,问题就复杂化了.冯·诺依曼的贡献是他发现了这个事实:只需对一个理性决策者作一些明确的、一般人都能接受的少数几条假设,那么一定可以将他对事情的偏好、各种情况下的后果用量化的效用函数来表示,也就是说,理性的人应该使用期望效用来分析问题.关于博奕的一些基本概念,在下一章我们会仔细论述.

有人认为博奕论的来源是贝叶斯决策理论.很多人自然会问:凭什么对一个简单的定量模型,可以指望它对人类的行为给出合理的描述呢?决策论正是回答这个问题的理论.它证明了一个决策者如果符合一些直观的公理,那么他的行为总是使他的期望效用最大化,这个期望是与他的信念——主观概率——相联系的.当新的信息使他的信念发生了改变时,他的决策和行为也就随之而变,调整信念的公式就是贝叶斯公式.

不确定的决策模型通常有两类:概率模型和状态变量模型.这两个模型都可用彩票的术语来定义.概率模型适宜于描述一些与事件发生有关的收益,而事件本身有客观的可以确定的概率,尽管这些概率是未知的,但它是客观的,例如轮盘赌、掷硬币这一类博奕.另一类现象是不存在客观的概率,这就适宜用状态变量的模型,这种模型能描述收益依赖于各种不可预测的事件,这些事件是不存在客观概率的,例如赛马、体育竞赛等,这时的决策是由主观的概率来判断的.

当然实际的决策问题既涉及客观的概率,又涉及主观的概率,这时用概率模型可以比较好地统一处理,今后的不少例子将会说明这一点.

信息对博奕的解有相当重要的作用.博奕看来似乎是一个静态的问题,然而实质上往往是一个动态的问题,只是把整个过程都