诺贝尔经济学奖获奖者著作丛书

信息经济学

〔美〕肯尼思·阿罗 著

北京经济学院出版社

·诺贝尔经济学奖获奖者著作丛书·

信　息　经　济　学

〔美〕　肯尼思·J· 阿罗　著

何宝玉　姜忠孝　刘永强　译

北京经济学院出版社

1989年·北京

Kenneth J. Arrow
The Economics of Information
Basil Blackwell Limited, 1984

信 息 经 济 学
Xinxi Jingjixue

〔美〕 肯尼思·J·阿罗 著
何宝玉等 译

北京经济学院出版社出版
（北京市朝阳区红庙）
北京经济学院出版社永乐印刷厂印刷
新华书店北京发行所发行

850×1168毫米 32开本 10.75印张 280千字
1989年3月第1版 1989年3月第1次印刷
印数：00 001—7 000
ISBN7—5638—0083—2/F·56

定价：4.35 元

《诺贝尔经济学奖获奖者著作丛书》
编 辑 委 员 会

编者献词

近年来，我国经济学界介绍了不少现代西方经济学家的著作，对我国经济学理论的研究起了积极的作用。为了使这种介绍更具有全面性和代表性，我们决定系统地翻译出版获诺贝尔经济学奖的学者的主要著作，作为翻译丛书奉献给广大读者。

诺贝尔经济学奖是当今世界上最有影响的经济学奖。它是1968年瑞典中央银行在其建立300周年时为纪念诺贝尔奖提供者而设立的基金，全称是“纪念诺贝尔经济学奖”。该奖由瑞典皇家科学院委任组成的“经济科学委员会”评定。从1969年起在每年的10月中旬与其他各项诺贝尔奖同时公布，到1987年为止已颁发19届，共有欧美的25位经济学家获奖。获奖的经济学家的经济理论基本上代表了西方经济学的主要成就，对西方经济学的发展有较大的影响。读者通过这些著作能全面、深入、系统地了解和研究西方经济学，从中吸取对我国经济改革和经济学发展有价值的东西。

当代西方经济学最显著的特点之一是数学方法的应用。经济学要反映经济中的数量关系，因此，经济学与数学早就结下了不解之缘。从17世纪起经济学家就用数学公式来表述经济理论。本世纪以后经济学家又把经济理论、数学和统计学结合起来建立了经济计量学，用于解决实际问题，并取得了重要进展。从某种意义上可以说经济学的数学化是经济学精密化、实用化的标志。在获奖的经济学家中，有三分之一以上在这方面作出了开创性的贡献。获得第一届诺贝尔经济学奖的挪威经济学家R.费瑞布和荷兰经济学家J.丁伯根是经济计量学的创立者。以后获奖的美国经济学家P.萨缪尔森、K.阿罗、W.列昂惕夫、T.库普

曼、L.克莱因、G.德布鲁，英国经济学家J.希克斯、苏联经济学家L.康托洛维奇都在经济学数学化方面作出了重要贡献。

在流派林立的当代西方经济学中，新古典综合派是公认的“主流派”。这一派继承并发展了凯恩斯主义，对西方各国的经济政策和经济理论都有重大的影响。因此，在获奖的经济学家中属于这一派的有相当比例，如美国经济学家P.萨缪尔森、L.克莱因、J.托宾、F.莫迪利亚尼、R.索洛，都是这一派的代表人物。

70年代后西方经济政策与经济理论的显著变化是自由主义的影响日益增强，代表自由主义的奥地利经济学家F.哈耶克，美国经济学家M.费里德曼、G.斯蒂格勒、J.布坎南先后获得诺贝尔经济学奖，便对映了这种趋势。

经济学是一门实用性很强的科学。因此，获奖者中还有一批对解决各种实际问题作出贡献的经济学家，如：美国经济学家S.库兹涅茨和英国经济学家R.斯通对国民收入统计作出开创性贡献；美国经济学家A.刘易斯和T.舒尔茨对研究发展中国家经济发展问题有独特贡献；瑞典经济学家B.奥林和英国经济学家J.米德对国际贸易和国际经济关系问题作出重要贡献；美国经济学家H.西蒙对管理科学的发展起了重要作用。

本丛书的选编以获奖者的代表作，特别是在西方经济中影响重大的名著以及对我国经济改革和社会主义经济理论有启发和借鉴的著作为主，每位获奖者的著作至少选一本，以保持其系统性与代表性。

还应该说明的是，西方经济学是资本主义的意识形态，获奖的经济学家是资产阶级学者，因而在这些著作必然有这样那样的问题，在阅读时我们应注意这一点。但是，西方经济学作为人类文明发展的成果之一，有其不可忽视的精华，资产阶级学者对经济问题的研究有其可取之处。从这种意义上说，只要我们以分析的

态度认真阅读这些著作，必将得到许多有益的收获。

本丛书的编辑出版得到了老一代学者与中青年学者的关心与帮助，谨向关心与帮助这套丛书编辑出版工作的所有朋友表示感谢。

编　委　会

1988年4月

前　言

奈曼和皮尔逊的研究①为统计方法的基本原则提供了一种经济计算，他们提出了进行此种计算标准，并力图使其最优化。1950年，沃尔德给出了一种可为明确的经济计算方法。②甚至在更早的时期，在票据承兑审查和质量控制方面，统计方法论就已建立在明确的经济概念（生产者的和消费者的风险）基础上了。统计方法就是获取信息的一个例子。在一个不确定性的世界里信息的确很有价值，认识到这一点并非难事。但是，事实已经证明，把信息作为一种经济商品去构造它的一般理论却困难重重，因为不同种类的信息并没有达成统一的单位。本书中的论文试图从不同的角度搞清这个问题的范围难度，并且，在某些特定的情形中，还提出了一些解决问题的方法。但遗憾的是，一般化的处理方法仍然是难以捉摸的。

玛丽·埃伦十分仔细地通篇编辑了这本书，迈克尔·巴克利和罗伯特·伍德为本书准备了词条索引，对于他们的工作，我深表感谢。

---

① 见J·奈曼和E·皮尔逊合写的文章"论统计假设的最有效检验问题"，载《伦敦皇家学会哲学学报》A辑，第231期，第289—337页，1933年。

② 见A·沃尔德《统计决策函数》，纽约：威利出版公司，1950年。

中 译 本 序 言

或许，从来没有一个经济学家会否认，大多数经济决策都是在具有相当的不确定性的条件下作出的，但只是在最近几十年，大约从1950年开始，明确地分析不确定性下的经济行为的工具才得以运用。一旦不确定性的存在在形式上是可以分析的，信息的经济作用就变得十分重要了。人们可以花费人力及财力来改变经济领域（以及社会生活的其他方面）所面临的不确定性，这种改变恰好就是信息的获得。不确定性具有经济成本，因而，不确定性的减少就是一项收益。所以，把信息作为一种经济物品来加以分析，既是可能的，也是非常重要的。

数理统计学家早就认识到了复杂的抉择问题，信息经济学已经吸收了、并且还将继续利用他们的研究成果。本书中的论文是在一段很长的时期内分别撰写的，它们的演变过程为读者揭示了信息经济领域里研究重点的转变。书中主题引出了经济和社会组织中最深奥复杂的问题以及它的意义和作用，并且能够使人们在不同的社会制度下深思这个问题所带来的利益。

因而，我确信，正象在其他国家那样，这些论文对中国的经济大众也将是有所裨益和启发的。

肯尼思·J·阿罗

1988.1.于斯坦福

目　录

1 序列化决策问题的贝叶斯解和最小最大解*

亚伯拉罕·沃尔德在“统计研究小组”发展了用序列化分析来检验统计假设的思想和方法。“统计研究小组”是第二次世界大战期间为开发用于国防的统计方法而成立的。毫无疑问，象其他类似的研究成果一样，在需引起这些研究的紧急状态结束之前，许多成果是得不到利用的。最初标有“机密”字样的备忘录（统计研究小组，1945年）已经公开，我们中的几个人在空军司令部气象部门使用了这种方法，在几个月内，我们检验了一下用它作出的长期天气预报是否显然比随机结果更好（结果并不显著）。我逐渐对沃尔德统计决策论的更一般化形式产生了特殊兴趣，包括非序列化的以及序列化的决策理论（沃尔德，1947年b）。

1948年夏，在兰德公司工作期间，我和梅叶·A·格

---

*该章是同D·布莱克韦尔和M·A·格希克合写的。这次是重印《计量经济学》杂志第17卷(1949年)上第213—244页的内客。本论文所进行的研究工作是为同美国空军签订了合同的非盈利性研究机构—兰德公司而做的。1948年9月9日，计量经济学会和数理统计学会在威斯康星州麦迪逊城联合召开了题为“统计学与对策论”的讨论会，我们向大会提交了这篇论文。本论文中的许多结论与沃尔德和沃尔弗威茨1948年以前的研究结果以及1948年6月22日数理统计学会在加利福尼亚州伯克利市召开的会议上由沃尔德所宣读的他们二人尚未公开发表的论文有重迭之处。尽管由于一般方法的不同，我们的论文与沃尔德等二人的表述与证明都不相同，但是本文第三节和第六节的内容，却与沃尔德与沃尔弗威茨1948年论文中的引理1—4相似。本文第五节中的结论：“序列概率比的二分法检验在任一个假设下都使观察次数的期望值最小化”是从第三节的推导中得出来的，所采用的证明方法，与沃尔德等二人的论文中由引理1—8推论出相同结论时所用的证明方法完全一样。前面提到的沃尔德等二人尚未发表的论文成果，都包括在本文第二节的主要内容当中，即“有限多级决策问题的优化序列过程的结构”。

希克进行了特别合作。大概就在那时，格希克参加了数理统计学会的一个会议，会上沃尔德和雅各布、沃尔弗威茨报告了有关“在最多两个备择假设的条件下序列分析结构”的一些新的研究成果。格希克回来之后非常兴奋，鼓动大卫·布莱克韦尔和我同他一起重新建立形式更清楚的结论；因为那些初步的结果难以理解且基本逻辑不清。我们三人紧紧抓住这样一个基本思想，即：尽管每一步的参数值不同，但决策情况是重复的，因此，决策规则只存在于参数空间的特定区域，且不随时间而改变。当然，在沃尔德（1947年）和沃尔德与沃尔弗威茨（1948年）的研究中都包含有这种观点，但却没有成为研究的中心内容。

由于没有充分承认沃尔德和沃尔弗威茨的工作就出版了一个版本，他们认为这是对其优先权的侵犯，这是这篇论文的一段不愉快插曲。这次出版的版本公正地揭示了这些论文之间的关系。

本文明确提出了递归优化的概念。它为我提供了应用于最优库存决策的模型。最优库存决策的工作是由我和西奥多·E·哈里斯、雅各布·马尔萨克共同完成的。更为重要的是，它帮助理查德·贝尔曼（1951年）提出了动态规划的一般原理，这一原理已被应用到许多领域。

沃尔德（1947年b)把统计决策问题形式化为：要求统计学家从可能行动集A中选择某个行动a，此时他所受损失是L(u,a)，这是行动a和未知的“自然”状态u的一个已知的有界函数。那么，统计学家采取什么行动最好呢？

如果u是一个随机变量，不一定是数值的，已知其先验分布，那么$\Sigma L(u,a)=R(a)$就是采取行动a时损失的期望值，而且，任何行动或者是行动的随机组合，若能使R(a)极小化，沃尔

德就称它为决策问题的相应于给定先验分布u的贝叶斯解。

现设一个随机变量序列x：$x_1, x_2, x_3 \cdots$，其联合分布由u决定。统计学家可能不是立即选择一个行动，而决定选取x的一个样本，因为由此可以得到关于u的部分信息，从而使他做一个更明智的选择。获得样本x_1，x_2，$\cdots X_N$要付出成本$C_N(x)$，在取样过程中，统计学家必须权衡预期成本和预期可得到的信息量。

形式上，观测的可能性并不改变局势，除非统计学家的可能行动集A扩展了。现在他的行动由取样程序T和决策函数D组成，D指定了对每个可能的实验结果将采取什么行动。现期预期损失为$R(T, D) = l(T, D) + o(T)$，此处$l(T, D)$是$L(u, a)$对指定的取样程序和决策规则的期望值，$C(T)$是取样程序的预期成本。一个贝叶斯解就是一个偶对（T，D）或偶对的随机组合，使R（T，D）达到它的极小值。

沃尔德隐约刻划了极小化过程$T = T^*$，它可以描述为下述规则：在每一步上，当且仅当存在某后续观察能使预期风险降低到目前水平之下时，再作另一次观测。这里的主要困难是，出现了许多不可测度的量：例如，如果第一次观测是x_1，那么我们必须比较一下现在的风险水平比如说$w_1(X_1)$和$Z(X_1) = \inf w(X_1, T, D)$谁大谁小，此处，$w(X_1, T, D)$是任意一个可能后续偶对(T，D)的预期风险；当且仅当$w_1 > Z$时，我们才做另一次观测。由于事先并不清楚Z是不是X_1的可测函数，所以停止点X_1的集合可能是不可测的。实际上，正如我们将要指出的Z总是可测的。①

在随机取样条件下，对于涉及有限数目的备择假设，可得出极小化$T = T^*$的一个特性。我们已知K种假设$H_i(i = 1, 2, \cdots, k)$每种H_i有一个先验的发生概率g_i，已知风险矩阵$W = (w_{ij})$，此处w_{ij}表示当Hi为真时选择H_j要受的损失，还知道代表观察n

① 沃尔德（1947年a）或沃尔德和沃尔弗威茨（1948年）没有考虑不可测这种可能性。

次的成本函数C（n）。可以看出，对每个样本容量N，在由K维欧氏空间的单位向量生成的（k—1）维单纯形中，存在K个凸区域S^*_i，其边界取决于假设H_i、风险矩阵W及成本函数$C_N(n)=C(N+n)-C(n)$。这些区域有这样的性质：如果向量g（N）的分量代表S^*_i中k个假设的后验概率分布，那么，最佳程序就是接受H_i不作进一步试验。然而，如果g（N）落在$\bigcup_{j=1}^{k}S^*_j$的补集中，那么最佳程序是继续观测。在每一阶段上，是继续还是终止取样，由k个区域的序列唯一地决定，而且，该区域序列完全刻划了T*。

当成本函数是线性时，决定这些凸区域边界的方法对K=2（二分法）是已知的。在这种情况下，T*与沃尔德的序列化概率比检验是一致的。

我们将考虑多值决策问题的最小最大解，并给出二分情况下获得该解的方法。应该指出，一般情况下，最小最大策略对统计学家来说是完美的，除非假设中包含了离散变量。在后一种情况下，混合策略应成为规则而不是例外。双二分法，二项式二分法，三分法的例子将说明如何构造T*和最小最大解的概念。

有必要强调指出，几个行动中最优的连续选择问题，同不确定情况下企业家合理行为的经济问题紧密相关。在每个时点上，企业家都有两种选择：是缔结某个不完全流动的协约呢，还是以现金持有部分或全部资金等待得到另外的信息？后者将是代价昂贵的，因为要损失掉在此之前的利润。

1. 贝叶斯解的构造

决策函数

我们已经知道统计学家必须选择一个偶对（T，D），这就

说明D的选择不依赖于T的选择。

引理·存在着一个确定的决策函数序列D_m，使R（T，D_m）$\longrightarrow$ infR（T，D）= W（T）对所有T。（1—1）

这是本节的主要结论。由此可推论出：程序T带来的预期损失可取为W（T），因为这个损失可以通过适当地选择D_m以任意精确度逼近；对给定的选择集，最佳的序列程序T^*是使w（T^*）= infw(T)的那一个，其中下确界是对所考察的集合中所有程序T而取的。

接下去我们考察一下随机变量u和随机变量序列x：x_1，x_2 …。一个序列化过程T是互不相交集合的序列S_0，S_1，S_2，…，S_N，…，其中S_N仅依赖于X_1, X_2, X_n，它是在样本$X_1, X_2, X_3, \cdots, X_n$中取样程序结束的事件集；我们要求$\sum\limits_{N=0}^{\infty} P(S_N)=1$。$S_0$是我们不作任何取样而直接采取行动的事件，它发生的概率是0或1。

决策函数D是函数序列d_0，$d_1(x_1)$，…$d_N(x_1, \cdots x_N)$，…，每个d_N在A中取值、确定取样在x_1，…，x_n中结束时所采取的行动。我们仅承认那些使L（u，$d_N(x)$）对每个N都是可测函数的决策函数D。

引理的证明：(T,D)引起的损失是G（u,x;T,D = L(u,d_N(x)〕+ C_N（x）对于$x \in S_N$，$\in G$ = R（T，D)。这里$C_N(x)$仅依赖于x_1，x_2，…，x_N。记$\in_N$为已知x_1，…，x_N时的条件期望，我们有$\in_N G = \in_N L[u, d_N(x)] + C_N(x)$对$x \in S_N$，且

$$R(T,D)=\sum_{N=0}^{\infty}\int_{S_N}\in_N L(u,d_N)\,dp+c(T) \quad (1—2)$$

现在固定N，我们将证明能够选择函数序列$d_{Nm}(x)$，m = 1，2，…，使

(a) $\in_N L(u, d_{Nm}) \geqslant \in_N L(u, d_{N,m+1})$ 对所有x，

(b) $\mathcal{E}_N L(u, d_N) \geqq r_N$ 对所有 d_N 和 x，此处

$$r_N(x) = \lim_{m\to\infty} \mathcal{E}_N L(u, d_{Nm})。$$

(c) $r_N \geqq \mathcal{E}_N r_n$ 如果 $n \geqslant N$。

首先选择一个序列 d'_{Nm} 使

$$\mathcal{E}L(u, d'_{Nm}) \longrightarrow \inf_{dN} \mathcal{E}L(u, d_N) = r$$

现归纳定义 d_{Nm} 如下：$d_{N1} = d'_{N1}$；对那些取值使 $\mathcal{E}_N L(u, d'_{Nm}) \leqslant \mathcal{E}_N L(u, d_{N, m-1})$ 的 x，令 $d_{Nm} = d_{Nm}$，否则 $d_{Nm} = d_{N, m-1}$。这样（a）当然成立，所以 $\lim_{m\to\infty} \mathcal{E}_N L(u, d_{Nm}) = r_N(x)$ 存在。又 $\mathcal{E}_N L(u, d_{Nm}) \leqslant \mathcal{E}_N L(u, d'_{Nm})$，所以 $\mathcal{E}r_N = r$。任取 d_N 和 $\delta > 0$，令S为事件 $\{\mathcal{E}_N L(u, d_N) < r_N(x) - \delta\}$，然后，在S上定义 $d^*_{Nm} = d_N$，其它地方令 $d^*_{Nm} = d_{Nm}$，我们有

$$\mathcal{E}L(u, d^*_{Nm} \leqq \int r_N(x)dp + \int_{CS} \mathcal{E}_N L(u, d_{Nm})dp - \delta P(S),$$

所以 $\lim_{m\to\infty} \mathcal{E}L(u, d^*_{Nm}) \leqslant r - \delta P(S), P(S) = 0$。这就证明了(b)。

最后，（c）可以从下面的事实推导出来：如果 $n > N$，每个 $d_N(x)$ 也可能是 $d_n(x)$；这就是说，定义 $d^*_n = d_N$，我们有 $\mathcal{E}_N L(u, d_N) = \mathcal{E}_N(\mathcal{E}_n L(u, d^*_n)) \geqslant \mathcal{E}_N r_n$ 对所有 d_N 成立，因此（c）成立。

现在来定义 $D_m = \{d_{Nm}\}$。因为 $\mathcal{E}_N L(u, d_{Nm})$ 随m递减，从（1-2）可得

$$R(T, D_m) \to \sum_{N=0}^{\infty} \int_{SN} r_N(x)dp + c(T) = w(T),$$

再利用一下（b），即知 $R(T, D) \geqslant W(T)$ 对所有D成立。这样，我们就把求贝叶斯解的问题简化为如下问题：已知随机变量序列x：x_1，x_2，…，和非负的预期损失函数序列：w_0，w_1，

…，这里$w_N = r_N(x_1, \cdots, x_N) + C_N(x_1, \cdots, x_N)$，$C_N$是前N次观测的成本，$r_N$是信息不完备带来的损失；对每个序列化过程$T = \{S_N\}$，相应地有风险$WT(\ ) = \sum_N \int_{S_N} w_N(x)dP$，怎样选择T才能使W（T）尽可能小呢？

最佳截短程序

在所有的序列化程序中，N次观测之内就会出现一个最好的，即预期风险不超过其他任一个的序列化程序。而且，该程序可以用倒推归纳明确地刻划出来，在这个过程中可以明显地看出其可测性。经过(N－1)次观测$x_1, \cdots, x_{N-1}$之后，我们就可以把现在的风险$\epsilon_{N-1} w_N$和假如作最后一次观测的条件预期风险$\epsilon_{N-1} w_N$比较一下，通过选择一个较好的步骤，我们可以把损失限定在$\alpha_{N-1} = \min(w_{N-1}, \epsilon_{N-1} w_N)$，可以认为这一风险是（N－1）次观测$x_1, \cdots\cdots, x_{N-1}$后的可达到的风险。接下去，在（N－2）次观测的基础上，我们可以通过比较现在的风险w_{N-1}和如果作了观测x_{N-1}的可达风险$\epsilon_{N-2}\alpha_{N-1}$，决定是否值得作第(N－1次观测)。象这样继续倒推下去，我们就得到每一阶段上对观测$x_1, \cdots\cdots, x_k$的预期可达风险α_k，以及如何达到这些风险，即何时作下次观测。下面的定理是其形式化的结果。

【定理1】 令$x_1, \cdots\cdots, x_N$; $w_0, \cdots\cdots, w_N$为任意随机变量，$w_i = w_i(x_1 \cdots\cdots, x_i)$。定义$\alpha_N = w_N, \alpha_j = \min(w_j, \epsilon_j \alpha_{j+1})$对$j<N$，$S_j = \{wi>\alpha j$对$i<j, w_j = \alpha_j\}$那么，对任意互斥事件$B_0, \cdots\cdots, B_N$，$B_i$仅依赖于$x_1, \cdots\cdots, x_i$，且$\sum_{i=1}^{N} P(Bi) = 1$，我们有

$$\sum_{j=0}^{N} \int_{S_k} w_j dP \leq \sum_{i=0}^{N} \int_{B_i} w_i dP$$