最优停止理论

上海科学技术出版社

最 优 停 止 理 论

〔美〕周元燊 H. Robbins D. Siegmund 著
何声武 汪振鹏 译

上海科学技术出版社

Great Expeotation:

The Theory of Optimal Stopping

Y. S. Chow H. Robbins D. Siegmund

Houghton Mifflin Company, Boston

封面设计 甘晓培

最优停止理论

〔美〕周元燊 H. Robbins D. Siegmund 著

何声武 汪振鹏 译

上海科学技术出版社出版

（上海瑞金二路450号）

新华书店上海发行所发行 无锡县人民印刷厂印刷

开本787×1092 1/32 印张5.125 字数112,000

1983年4月第1版 1983年4月第1次印刷

印数 1— 7,400

统一书号：13119·1038 定价：（科五）0.60元

序　　言

这本小册子起源于作者们在过去十年里讲授的最优停止理论的课程．本书试图给出离散时间最优停止理论的一个系统的数学基础，同时也提供一些例子说明一般理论对特殊问题的应用．但本书并非是这一领域的全面概述．例如，读者找不到连续时间问题的叙述(参阅 Shiryaev [3])，也找不到序贯概率比检验的最优性的证明(参阅 Ferguson [1])．

一些特殊的最优停止问题，如第三章中的"秘书问题"，在概率论中已有悠久的历史．然而最初的某些一般性的结果是在本世纪四十年代后期由 Wald, Wolfowitz [1] 和 Arrow, Blackwell, Girshick [1] 在研究序贯统计决策问题时得到的．Snell [1] 在 1952 年提出了处理这些问题的另一个方法——参阅 4.4 节．最优停止理论作为概率论的一个部分和它对统计的特殊应用(但并非只限于统计的应用)，大约自 1960 年以来处于迅速发展之中，这本小册子所总结的主要是作者们对这一理论的发展所作的贡献．

凡掌握测度论水平的概率论(包括对条件期望的抽象概念有一定了解)的读者，对最优停止的一般理论都应是能够理解的．第一章提供了基本的准备知识．第二章简要地叙述了鞅论，鞅论的方法与最优停止理论的方法有相似之处．第三

到第五章的一般理论并不依赖于第二章的结果，虽然在把这个理论应用于一些特殊问题时有时却要用到它们．所以读者可在略读第一章之后直接进入第三章，只是在必要的时候再回过来参考第二章．

我们要感谢许多同事与学生，他们的兴趣和参与促成这一计划具有现在这样的面貌．在最后定稿时 H. Chernoff 作了许多有益的建议．我们还要感谢 M. Lof 小姐的高超的打字工作．

尽管本书还不够全面与完善，但它是我们献给曾使我们迷恋多年的一门学科的一点贡献．

目　　录

引　　言

概率论是从计算机会游戏中的输赢开始的. 在这意义下, 最优停止问题所关注的是赌徒在一系列的赌局中用各种可能的系统去决定何时停止赌博对赌金会有什么影响. 一个较现代的应用领域是统计推断, 这里一个试验者经常要问: 进一步得到的资料中所包含的新增加的信息在价值上是否超过了收集资料所耗的费用.

最优停止理论提供一个一般的数学模型, 使能精确地叙述这些问题, 并在某些情形中完全解决它们. 在这个引言中我们将给出四个简单的例子, 以说明在一般理论中所产生的某些问题.

例 1　"加倍或输掉"的有利修正

重复掷一均匀硬币. 每掷一次后我们可选择停止或再掷一次, 每一步所作的决断取决于迄今为止所掷出的结果. 在掷了有限次(但不是预先定好的次数)后我们必须停止, 并约定, 若掷 n 次后停止, 我们会赢得报酬 x_n. 在"加倍或输掉"的赌博中, 对于以一个单位的赌本开始的赌徒, 他赢得的报酬是 $\prod_{j=1}^{n}(y_j+1)$, 其中 $y_j=1$ 表示在第 j 次掷出正面, $y_j=-1$ 则

为掷出反面. 然而我们将考虑这样的情形, 为鼓励继续赌下去, 把上述报酬乘以一个递增的序列 $\frac{2n}{n+1}$, 使报酬序列为

$$(1)\qquad x_n=\frac{2n}{n+1}\prod_{j=1}^{n}(y_j+1)\quad(n=1,\ 2,\ \cdots).$$

这里, 只要在前 n 次中出现反面它就等于 0, 其它情形就等于 $2^{n+1}n/(n+1)$. 这只是一般情形的一个特例, 在一般情形中, 第 n 步的报酬 x_n 是过去结果的任一给定的函数, $x_n=f(y_1,\ \cdots,\ y_n)$.

一个停止规则是一个随机变量 t, 它取值于 $\{1,\ 2,\ 3,\cdots\}$, 且事件 $\{t=n\}$ 只取决于 $y_1,\ \cdots,\ y_n$ 的值而与将来的值 $y_{n+1},\ \cdots$ 无关. 采用任何停止规则 t, 报酬 x_t 是一个随机变量, 它的期望 Ex_t 平均地度量了停止规则 t 的优劣. 对存在 Ex_t 的停止规则 t 的全体 C, 其上确界 $V=\sup\{Ex_t\}$ 称为序列 $\{x_n\}$ 的值, 若存在停止规则 t 使 $Ex_t=V$, 则称 t 是最优的.

我们的总目标是要得到表征 V 的值及最优停止规则 (如果它存在) 的方法. 略作思考即可证明, 对特殊情形 (1) 我们只要考虑停止规则 $\{t_k\}$, $k=1,\ 2,\ \cdots$, 其中 $t_k=k$, 即不管正反面出现的情况如何, 掷完第 k 次后就停止. 明显地,

$$Ex_{t_k}=\frac{1}{2^k}\cdot\frac{k\,2^{k+1}}{k+1}+\left(1-\frac{1}{2^k}\right)\cdot 0=\frac{2k}{k+1},$$

所以 $V=2$, 但不存在最优停止规则.

我们顺便指出, 在任何时刻 n, 如一直只有正面出现, 即 $x_n=n2^{n+1}/(n+1)$ 时, 则再掷一次而停止的条件期望报酬为

$$E(x_{n+1}|x_n=n2^{n+1}/(n+1))=\frac{1}{2}\cdot\frac{(n+1)2^{n+2}}{(n+2)}$$
$$=2^{n+1}(n+1)/(n+2)>x_n,$$

所以在全是正面时停止总是"愚蠢的", 即从期望的观点看, 再

掷一次总是更有利的．但是如果我们不在某一点“愚蠢地”行动，那就要继续赌下去直到出现第一个反面，这是一个以概率1最终要出现的事件，而最后的报酬将永远是零．所以在这种情形，每一步的行动都是“聪明的”，但从长远来看却是最劣的策略．顺便指出，这个最劣的策略自然可认为是期望报酬收敛于值 $V=2$ 的停止规则 t_k 的极限形式．

这个例子说明了一个问题和两个反常情况．

(a) 最优规则可以不存在．在怎样的条件下存在最优规则及什么时候可简单地描述它？

(b) 一个在每一步都“聪明地”行动的策略却可能是拙劣的．

(c) 一列好的策略的极限形式可能是拙劣的．

在什么条件下我们能肯定上述两种情况不会发生呢？

如果用 $x'_n=1-x_n$ 来代替 x_n，我们得到一列不利的赌局．这时等待出现反面的策略说明

(d) 即使一个赌徒参加一系列不利的赌局，他仍可以保证自己得利．

例 2　报酬等于平均数

诸 y_n 同例 1，设报酬为

(2) $$x_n=\frac{y_1+y_2+\cdots+y_n}{n}\quad(n=1,\ 2,\ \cdots).$$

这个问题比前一问题难得多，因为有庞大的一族停止规则必须考虑．这种规则的一个简单例子是

(3) $$\begin{aligned} t&=1,\ 若\ y_1=1,\\ &=n,\ 在其它情形，其中\ n\ 是第一个整数，使得\\ &\qquad y_1+\cdots+y_n=0. \end{aligned}$$

已知(见习题2.2)对任意整数k,事件$\{y_1+\cdots+y_n=k$, 对某个$n\}$以概率1发生,所以在$P\{t<\infty\}=1$这一意义下,t是合法的停止规则. 对(2)及(3)计算Ex_t:

$$Ex_t=\frac{1}{2}\cdot\frac{1}{1}+\frac{1}{2}\cdot 0=\frac{1}{2},$$

由此得出$V\geqslant\frac{1}{2}$, 并且因为x_n总是$\leqslant 1$,当然$V\leqslant 1$. 用试验调整法我们能造出另一个停止规则t使$Ex_t>\frac{1}{2}$,但是我们却找不到一个t使$Ex_t>0.9$的例子. 然而要证明$V\leqslant 0.9$是不容易的, 更不用说去找到V的精确值和决定是否存在最优的t.

象这样的情形会使人想起"把问题放到计算机上去". 下面就是一个计算机能做的事情. 假设我们限于在$\{1, 2, \cdots, N\}$中取值的一切可能停止规则t的全体O^N, 这里N为某固定的正整数; 换言之, 我们限于至多掷N次硬币后总停止的规则. 以V^N记Ex_t在O^N上取的上确界. 我们看到因为只有有限多个向量$(y_1, \cdots, y_N)$, $y_i=\pm 1$, O^N是一个有限类, 所以O^N中的最优规则必定存在. 虽然如此, 譬如说对$N=1000$, O^N已经大到不能直接分析了, 在这方面最优停止的一般理论能给我们帮助. 只要问题中的步数是有界的(即使所涉及的随机变量是连续的而不是离散的), 总存在最优规则, 同时有一个求最优规则的类似于构造性的算法(见定理3.2). 这个算法的一般思想可用一个术语"后退归纳法"来概括. 在本问题中能写下一个计算程序, 对直到几千的N都能迅速地求得V^N. 由定义, 序列V^N是非降的, 所以$V'=\lim\limits_{N\to\infty}V^N$存在. 但是用计算机对逐个$N$的值计算出的$V^N$的数值显示不出明显的规律, 从计算机所获得的结果不可能猜测V'的精确值,

更谈不上是否 $V=V'$ 或 $V>V'$，是否在 C 中存在最优的 t.

事实上在本问题中能证明(定理 4.3 及 4.11) $V=V'$ 及存在最优的 t，但仍不知道 V 的数值及 t 的确切描述.

一般地，我们要问：

(e) $V=V'$ 总是对的吗？

(f) 有无有效的方法把 V 介于两个彼此接近的可计算的上、下界之间？

例 3 修改的和

诸 y_n 同例 1，设

$$x_n=\min(1,\ y_1+\cdots+y_n)-\frac{n}{n+1}\quad(n=1,\ 2,\ \cdots).$$

考虑特殊的停止规则

$$(4)\qquad t=\text{第一个 } n \text{ 使得 } y_1+\cdots+y_n=1,$$

如同例 2 那样可得 $P\{t<\infty\}=1$，并且由于 $\frac{n}{n+1}<1$，

$$Ex_t=1-E\left(\frac{t}{t+1}\right)>0.$$

(能计算出 Ex_t 的精确值，但这里我们不必关心它.) 因此 $V>0$. 稍微想一想就可证明，在这个例子中 t 事实上是最优的.

另一方面，由于诸 y_n 是独立同分布的，且 $Ey_n=\frac{1}{2}\cdot 1+\frac{1}{2}\cdot(-1)=0$，Wald 引理(参阅定理 2.3 或引理 3.1)证明，如 t 与(4)不一样而是任何一个使 $Et<\infty$ 的停止规则，特别，如果对某个 $N=1,\ 2,\ \cdots,\ t\in C^N$，则 $E(y_1+\cdots+y_t)=0$，从而

$$Ex_t\leqslant E(y_1+\cdots+y_t)-E\left(\frac{t}{t+1}\right)\leqslant-\frac{1}{2},$$

所以 $V'=\lim\limits_{N\to\infty}V^N\leqslant-\frac{1}{2}$，而我们已看到 $V>0$. 所以对 (e) 的回答是"不"，但人们可能问：

(e′) 在什么条件下 $V=V'$？换句话说，什么时候这个值可用预先给定最大步数的计算程序逼近？

把本例略为变化一下便得

例 4 另一个修改的和

设 $y_1, y_2, \cdots$ 为独立(但不同分布)随机变量使得

(5) $$P(y_j=1-a_j)=P(y_j=-1-a_j)=\frac{1}{2},$$

$$a_j=\frac{1}{j(j+1)}.$$

令

(6) $$x_n=y_1+\cdots+y_n \quad (n\geqslant 1)$$

是在第 n 步停止的报酬. 于是 x_n 代表一个赌徒参加一系列不利于他的赌局在赌了 n 局后停止时的净收益,

$$Ey_j=\frac{1}{2}(1-a_j)+\frac{1}{2}(-1-a_j)=-a_j<0.$$

看来似乎不论他用什么样的停止规则 t，停止后的平均净收益 Ex_t 总要 <0. 然而令

(7) $$t=\text{第一个 } n\geqslant 1 \text{ 使得 } \sum_1^n(y_j+a_j)=k,$$

其中 k 是任一预先给定的正整数. 例 2 中的论证又一次证明了 $P(t<\infty)=1$，而由(6)及(7),

$$Ex_t=E\left(\sum_1^t y_j\right)=k-E\left(\sum_1^t a_j\right)=k-E\left(\sum_1^t\left(\frac{1}{j}-\frac{1}{j+1}\right)\right)$$

$$=k-E\left(\frac{t}{t+1}\right)>k-1.$$

由于我们可以要 k 多大就多大，所以对这个例子 $V=+\infty$. (读者可判断这时是否存在一个最优的 t，即使 $Ex_t=\infty$ 的那个 t.) 我们指出，在这个例子中尽管一步的条件期望报酬 $E(x_{n+1}|x_n)=x_n+E(y_{n+1})$ 总是小于现在的报酬 x_n，利用一个适当的停止规则仍可使赌博十分有利.

第 一 章

准 备 知 识

这一章扼要叙述测度论和积分论中某些标准结果，这些结果将在下面各章中用到.

1. 事件代数

集合 Ω 的一个子集类 $\mathscr{F}$ 称为一个代数，如果满足

(1.1) $\Omega \in \mathscr{F}$,

(1.2) $A \in \mathscr{F}$, $B \in \mathscr{F}$ 时, $\bar{A} \equiv \Omega - A \in \mathscr{F}$, $A \cup B \in \mathscr{F}$.

如果另外还满足

(1.3) $A_n \in \mathscr{F}$, $n=1, 2, \cdots$时, $\bigcup_{1}^{\infty} A_n \in \mathscr{F}$,

则称 $\mathscr{F}$ 是一个 σ-代数.

σ-代数 $\mathscr{F}$ 中的集合称为事件. Ω 的所有子集组成的类显然是一个 σ-代数，由空集 $\emptyset$ 和 Ω 组成的类也是一个 σ-代数. 前一个类是 Ω 的子集类中最大的一个 σ-代数，而后一个则是最小的.

给定 Ω 的任意一个子集类 $\mathscr{A}$，包含 $\mathscr{A}$ 的所有 σ-代数的交仍然是一个包含 $\mathscr{A}$ 的 σ-代数，并且这个交包含在每一个包含 $\mathscr{A}$ 的 σ-代数中. 这个 σ-代数称为由 $\mathscr{A}$ 产生的 σ-代数，并记作 $\mathscr{B}(\mathscr{A})$. 如果 $\mathscr{A}$ 是广义数直线 $[-\infty, \infty]$ 上所

有区间组成的类，则 $\mathscr{B}(\mathscr{A})$ 称为(线性) Borel 集的 σ-代数，并记作 $\mathscr{B}$.

2. 随机变量

给定分别具有 σ-代数 $\mathscr{F}$, $\mathscr{X}$ 的集合 Ω, X, 一个 $\Omega\to X$ 的函数 x，如果对每一个 $A\in\mathscr{X}$，有

$$x^{-1}(A)\in\mathscr{F},$$

则称 x 是一个随机变量. 如果同时讨论 Ω 中的几个 σ-代数，为明确起见，有时也称 x 是一个 $\mathscr{F}$-可测的随机变量. 若 $X=[-\infty,\infty]$，且 $\mathscr{X}=\mathscr{B}$，则称 x 是 Ω 上的一个实值随机变量. 若 $B\in\mathscr{F}$，以 I_B(B 的示性函数)记在 B 上为 1，在 B 外为 0 的实值随机变量. 若 $\{x_t,\ t\in T\}$ 是一族随机变量，其中 x_t 是 $\Omega\to X_t$ 的随机变量，这时以 $\mathscr{B}(x_t,\ t\in T)$ 记使得所有的 x_t, $t\in T$ 都是可测的最小的 σ-代数，并称它为 $\{x_t,\ t\in T\}$ 产生的 σ-代数.

3. 概率与期望

定义在集合 Ω 的子集组成的代数 $\mathscr{F}$ 上的一个非负广义实值集函数 μ 称为测度，如果满足 $\mu(\emptyset)=0$ 及

$$\mu\left(\bigcup_1^\infty A_n\right)=\sum_1^\infty\mu(A_n), \tag{1.4}$$

$$A_n\in\mathscr{F}(n\geqslant1),\ A_n\cap A_m=\emptyset(n\neq m),\ \bigcup_{n=1}^\infty A_n\in\mathscr{F}.$$

如果 Ω 是可列个 $\mathscr{F}$ 中的测度为有限的集的并，则称 μ 是 σ-有限的. 如果有

$$\mu(\Omega)=1, \tag{1.5}$$

则称 μ 是概率测度. 通常概率测度用字母 P 表示. 若 $\mathscr{F}$ 是

一个 σ-代数，而 P 是一个概率测度，则称 $(\Omega, \mathscr{F}, P)$ 是一个概率空间.

定理1.1 扩张定理 给定集合 Ω 的子集组成的一个代数 $\mathscr{A}$ 及 $\mathscr{A}$ 上一个概率测度 P，则存在 $\mathscr{B}(\mathscr{A})$ 上的一个唯一的概率测度，使得它在 $\mathscr{A}$ 上与 P 相等.

下面的结果可用来证明扩张定理中的唯一性，但它本身也是有独立意义的.

引理 1.1 给定一个概率空间 $(\Omega, \mathscr{F}, P)$ 和一个代数 $\mathscr{A}$，使得 $\mathscr{F}=\mathscr{B}(\mathscr{A})$，则对每一 $A\in\mathscr{F}$ 和 $\varepsilon>0$，存在 $B\in\mathscr{A}$，使得

$$P(A-B)+P(B-A)<\varepsilon. \tag{1.6}$$

证明 令 $\mathscr{C}$ 表示具有上述性质的 $A\in\mathscr{F}$ 全体组成的类，显然有 $\mathscr{C}\supset\mathscr{A}$，容易验证 $\mathscr{C}$ 是一个 σ-代数，从而 $\mathscr{C}\supset\mathscr{F}$.

令 $(\Omega, \mathscr{F}, P)$ 是一个概率空间，且 $x:\Omega\to X$ 是一个随机变量. 下式定义的 $\mathscr{X}$ 上的概率测度 Q 称为 P 在 x 下的象：

$$Q(A)=P(x^{-1}(A)), \quad A\in\mathscr{X},$$

Q 也称为 x 的分布.

设 $(\Omega, \mathscr{F}, P)$ 是概率空间，且 x 是非负实值随机变量，x 的期望，记作 Ex，定义为

$$Ex=\lim_{n\to\infty}\left(\sum_{k=0}^{n2^n-1}k2^{-n}P\{k2^{-n}<x\leqslant(k+1)2^{-n}\}+nP\{x>n\}\right)$$

(其中 $\{\cdots\}$ 表示 Ω 中满足…的点 ω 全体组成的集). 若 x 是任意的实值随机变量，不一定是非负的，定义 $x^+=\max(x,0)$，$x^-=-\min(x,0)$. 如果有 $Ex^+<\infty$ 或 $Ex^-<\infty$，则称 Ex 存在，并定义

$$Ex=Ex^+-Ex^-.$$

对于 $E(I_Ax)$，我们常常把它写成 $\int_A x$ 或 $\int_A x\,dP$. 期望的基

本性质如下：

(1.7)　若 $x\geqslant 0$，则 $Ex\geqslant 0$。

(1.8)　$E1=1$。

(1.9)　若 Ex 存在，则对任何实数 c，$E(cx)$ 也存在，且

$$E(cx)=c\cdot Ex.$$

(1.10)　若 Ex_1 和 Ex_2 存在，Ex_1+Ex_2 不是 $+\infty-\infty$ 的形式，则 $E(x_1+x_2)$ 存在且等于 Ex_1+Ex_2。

定理 1.2　单调收敛定理　若 $x_n\uparrow x$ 且 $Ex_1^-<\infty$，则 $Ex_n\uparrow Ex$。类似地，若 $x_n\downarrow x$ 且 $Ex_1^+<\infty$，则 $Ex_n\downarrow Ex$。

注　如果我们只假定 $P(x_n\uparrow x)=1$，也就是说以概率 1 或几乎必然 (almost surely, a. s.) 有 $x_n\uparrow x$，则定理的结论不变。今后如果没有特别的说明，在概率空间上实随机变量的收敛总是理解为在几乎必然意义下的收敛。类似地，诸如"$x\leqslant y$"和"$x_n\to x$ 在 A 上"应理解为"$P(x\leqslant y)=1$"和 $P(A\cap\{x_n\not\to x\})=0$"。

4. 一致可积性

一个实值随机变量序列 (x_n) 称为一致可积的，如果有

(1.11)
$$\lim_{a\to\infty}\sup_n\int_{\{|x_n|>a\}}|x_n|=0,$$

或等价地如果有

(1.12)
$$\sup_n E|x_n|<\infty \quad 和 \quad \lim_{P(A)\to 0}\sup_n\int_A|x_n|=0.$$

引理 1.2　Fatou 引理　若 (x_n^+) 是一致可积的，且 $E(\limsup_n x_n)$ 存在，则

$$E(\limsup_n x_n)\geqslant\limsup_n Ex_n.$$

证明　若对一切 $n\geqslant 1$，$x_n\leqslant a<\infty$，我们注意到

$$\sup_{k\geq n} x_k \downarrow \limsup_n x_n, \quad n\to\infty$$

和

$$E\left(\sup_{k\geq 1} x_k\right) \leqslant a < \infty,$$

所以可应用定理 1.2 的第二部分得到

$$E\left(\sup_{k\geq n} x_k\right) \downarrow E\left(\limsup_n x_n\right).$$

再连同$E\left(\sup_{k\geq n} x_k\right) \geqslant Ex_n$ 即完成了特殊情形下的证明. 在一般情形下, 令 $\varepsilon>0$, $x_n(a)=\min(x_n, a)$, 其中 a 取得充分大使

$$Ex_n(a) \geqslant Ex_n - \varepsilon \quad (n\geqslant 1).$$

则

$$E\left(\limsup_n x_n\right) \geqslant E\left(\limsup_n x_n(a)\right) \geqslant \limsup_n Ex_n - \varepsilon.$$

再令 $\varepsilon\to 0$ 即完成了引理的证明.

定理 1.3 设 $0\leqslant x_n \to x$ 和 $Ex_n<\infty (n\geqslant 1)$, 则 $Ex_n\to Ex<\infty$ 当且仅当 (x_n) 一致可积.

证明 先证充分性. 对 (x_n) 和 $(-x_n)$ 应用 Fatou 引理可得

$$(1.13) \qquad Ex \leqslant \liminf_n Ex_n \leqslant \limsup_n Ex_n \leqslant Ex.$$

((1.12) 的第一部分和 (1.13) 表明 $Ex<\infty$). 再证必要性. 首先可以看到集合

$$B=\{b: P(x=b)>0\}$$

是可列的(事实上, 对每一个 $m=1, 2, \cdots$, $\left\{b: P(x=b)\geqslant \frac{1}{m}\right\}$ 是有限的). 对任何 $a\notin B$, 有

$$x_n I_{(x_n<a)} \to x I_{(x<a)},$$

显然, $(x_n I_{(x_n<a)})$满足一致可积性的第一个定义. 从而由上述

充分性知

$$\int_{(x_n<a)} x_n \to \int_{(x<a)} x,$$

于是对每一个 $a\notin B$ 有

$$\int_{(x_n>a)} x_n \to \int_{(x>a)} x.$$

令 $\varepsilon>0$, 选取 $a_0\notin B$ 充分大，使得

$$\int_{(x>a_0)} x<\frac{\varepsilon}{2}.$$

再令 N_0 充分大,使对所有的 $n\geqslant N_0$ 有

$$\int_{(x_n>a_0)} x_n \leqslant \int_{(x>a_0)} x+\frac{\varepsilon}{2}.$$

最后再取充分大的 $a_1\geqslant a_0$, 使对所有的 $n=1, 2, \cdots, N_0$, 有

$$\int_{(x_n>a_1)} x_n<\varepsilon,$$

于是对一切 $a\geqslant a_1$, 我们有

$$\int_{(x_n>a)} x_n<\varepsilon \quad (n=1, 2, \cdots).$$

推论 1 若 $x_n\to x$ 且 (x_n) 一致可积，则 $E|x_n-x|\to 0$.

证明 由 Fatou 引理和 (1.12) 的第一部分知

$$E|x|\leqslant \liminf E|x_n|<\infty.$$

不难从(1.12)看出 $(|x_n-x|)$ 是一致可积的,且由于 $|x_n-x|\to 0$, 从定理 1.3 即得

$$E|x_n-x|\to 0.$$

推论 2 **Lebesgue 控制收敛定理** 若 $x_n\to x$ 且存在一个随机变量 y, 使得

$$|x_n|\leqslant y \quad (n\geqslant 1)$$

及 $Ey<\infty$, 则 $E|x_n-x|\to 0$.

注 我们称 x_n 依概率收敛于 x 并记作 $x_n \xrightarrow{P} x$，如果对每个 $\varepsilon>0$，有

$$\lim_{n\to\infty} P\{|x_n-x|>\varepsilon\}=0.$$

利用等式 $\{x_n\to x\}=\bigcap_{r=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{|x_k-x|<r^{-1}\}$，不难看出，如果 $x_n\to x$，则 $x_n \xrightarrow{P} x$. 反过来是不成立的. 但如果把 $x_n\to x$ 的假设换成 $x_n \xrightarrow{P} x$，定理 1.3 及其推论仍然成立. 因为依概率收敛的概念在以后只起次要的作用，我们就不去仔细讨论它了.

在涉及条件期望(下面将定义)的论证中常用的一个基本结果是

引理 1.3 若 Ex_1 和 Ex_2 存在，且对每个 $A\in\mathscr{F}$ 有

$$\int_A x_1 \leqslant (=) \int_A x_2,$$

则 $x_1\leqslant(=)x_2$.

证明 对每个 $m>0$ 有

$$\int_{A(|x_1|<m)} (x_1-x_2)\leqslant 0.$$

令 $A=\{x_1-x_2>\varepsilon\}$，我们得到

$$P\{|x_1|\leqslant m,\ x_1-x_2>\varepsilon\}=0.$$

令 $\varepsilon\to 0$，$m\to\infty$ 就有

$$P\{|x_1|<\infty,\ x_1>x_2\}=0.$$

用类似的方法可证

$$P\{|x_2|<\infty,\ x_1>x_2\}=0.$$

因为 $Ex_1\leqslant Ex_2$ 包含了 $P\{x_1=+\infty,\ x_2=-\infty\}=0$，我们得到 $P\{x_1\leqslant x_2\}=1$. 等号情形的证明立即由上面的情形得到，因为这时可交换 x_1 和 x_2 的地位.

5. 条件期望

给定 $\mathscr{F}$ 的一个子 σ-代数 $\mathscr{G}$ 和一个非负实值随机变量 x, 则 x 关于 $\mathscr{G}$ 的条件期望是任一个 $\mathscr{G}$-可测的实值随机变量 g, 它使得对每一个 $A\in\mathscr{G}$,

$$\int_A x=\int_A g.$$

由引理 1.3 知, 除在一个概率为 0 的集上外, 这样的 g 是唯一确定的, 而这样的 g 的存在性是由 Radon–Nikodym 定理保证的. 若 x 是任一实值随机变量, Ex 存在, 则 x 关于 $\mathscr{G}$ 的条件期望(用 $E(x|\mathscr{G})$ 表示)定义为

$$E(x|\mathscr{G})=E(x^+|\mathscr{G})-E(x^-|\mathscr{G}).$$

如果 x_0, x_1, $\cdots$ 是一列实值随机变量, 对一切 $n\geqslant 0$, Ex_n 存在, 则(除一个零概率事件外)有

(1.14) 若 $x_0\geqslant 0$, 则 $E(x_0|\mathscr{G})\geqslant 0$.

(1.15) $E(1|\mathscr{G})=1$.

(1.16) 若 Ex_0+Ex_1 不是 $+\infty-\infty$ 的形式, 则

$$E(x_0+x_1|\mathscr{G})=E(x_0|\mathscr{G})+E(x_1|\mathscr{G}).$$

(1.17) 若 x_0 为 $\mathscr{G}$-可测且 $E(x_0x_1)$ 存在, 则

$$E(x_0x_1|\mathscr{G})=x_0E(x_1|\mathscr{G}).$$

(1.18) 若 $\mathscr{G}^*\subset\mathscr{G}$, 则 $E(x_0|\mathscr{G}^*)=E(E(x_0|\mathscr{G})|\mathscr{G}^*)$.

(1.19) **条件期望的单调收敛定理** 若 $x_n\uparrow x$ 且 Ex 存在, 则在 $\{E(x_0|\mathscr{G})>-\infty\}$ 上, $E(x|\mathscr{G})=\lim\limits_{n\to\infty}E(x_n|\mathscr{G})$.

(1.20) **条件期望的 Fatou 引理** 若 $x_n\leqslant x_0$ $(n\geqslant 0)$ 且 $E(\limsup x_n)$ 存在, 则在 $\{E(x_0|\mathscr{G})<\infty\}$ 上,

$$E(\limsup x_n|\mathscr{G})\geqslant\limsup E(x_n|\mathscr{G}).$$

(1.21) **条件期望的控制收敛定理** 若 $|x_n|\leqslant x_0$ $(n\geqslant 1)$, x_n

$\to x$，且 Ex 存在，则在 $\{E(x_0|\mathscr{G})<\infty\}$ 上，$E(|x_n-x||\mathscr{G})\to 0$。

我们将给出 (1.19) 的证明. 采用从单调收敛定理推导 Fatou 引理同样的方法即可从 (1.19) 推得 (1.20)，而 (1.21) 则是 (1.20) 的一个直接推论.

对任何 $m=1, 2, \cdots$，令 $B_m=\{E(x_0|\mathscr{G})>-m\}$，

$$B=\bigcup_1^{\infty}B_m=\{E(x_0|\mathscr{G})>-\infty\},$$

$x_n\uparrow x$ 蕴含了 $E(x_n|\mathscr{G})\uparrow$. 设 $y=\lim\limits_{n\to\infty}E(x_n|\mathscr{G})$. 因为

$$\int_{B_m}x_0=\int_{B_m}E(x_0|\mathscr{G})\geqslant -m>-\infty.$$

对每个 $A\in\mathscr{G}$ 和 $m=1, 2, \cdots$，由单调收敛定理我们有

$$\int_{AB_m}y=\lim_n\int_{AB_m}E(x_n|\mathscr{G})=\lim_n\int_{AB_m}x_n=\int_{AB_m}x$$

$$=\int_{AB_m}E(x|\mathscr{G}).$$

由引理 1.3 即得在 B_m 上

$$y=E(x|\mathscr{G}).$$

由于 m 是任意的，所以在 B 上

$$y=E(x|\mathscr{G}).$$

定理 1.4(**P. Lévy**) 设 $\{\mathscr{F}_n, n\geqslant 1\}$ 是 $\mathscr{F}$ 的单调非降子 σ-代数列，令 $\mathscr{F}_\infty=\mathscr{B}(\bigcup\limits_1^\infty\mathscr{F}_n)$，则对任一实值随机变量 x，$E|x|<\infty$，有

$$\lim_{n\to\infty}E(x|\mathscr{F}_n)=E(x|\mathscr{F}_\infty).$$

证明 令 $x'=E(x|\mathscr{F}_\infty)$. 由 (1.18) 可知，不失一般性我们可以假设 x 是 $\mathscr{F}_\infty$-可测的，故 $x=E(x|\mathscr{F}_\infty)$. 设 $\varepsilon>0$. 因为 $\bigcup\limits_1^\infty\mathscr{F}_n$ 是一个代数，由引理 1.1 以及 (1.6) 式等价于

$E|I_A-I_B|<\varepsilon$，易知存在一个 $n_0=n_0(\varepsilon)$ 和一个 x_{n_0}，x_{n_0} 是 $\mathscr{F}_{n_0}$-可测的，且使得

$$E|x-x_{n_0}|<\frac{\varepsilon^2}{2}.$$

令 $y=|x-x_{n_0}|$，又定义

$$\begin{aligned} t&=\text{第一个 } n\geqslant n_0 \text{ 使得 } E(y|\mathscr{F}_n)>\varepsilon,\\ &=\infty, \text{若不存在这样的 } n, \end{aligned}$$

则对一切 $n\geqslant n_0$，$\{t=n\}\in\mathscr{F}_n$. 我们可写

$$\begin{aligned} P\{E(y|\mathscr{F}_n)>\varepsilon \text{ 对某个 } n\geqslant n_0\}&=\sum_{n=n_0}^{\infty}P\{t=n\}\\ &\leqslant\sum_{n_0}^{\infty}\frac{1}{\varepsilon}\int_{\{t=n\}}E(y|\mathscr{F}_n)\\ &=\sum_{n_0}^{\infty}\frac{1}{\varepsilon}\int_{\{t=n\}}y\leqslant\frac{1}{\varepsilon}Ey<\frac{\varepsilon}{2}. \end{aligned}$$

但是对一切 $n\geqslant n_0$，因为 x_{n_0} 是 $\mathscr{F}_n$-可测的，所以

$$\begin{aligned} |E(x|\mathscr{F}_n)-x|&\leqslant|E(x-x_{n_0}|\mathscr{F}_n)|+|x_{n_0}-x|\\ &\leqslant E(y|\mathscr{F}_n)+y. \end{aligned}$$

于是

$$\begin{aligned} &P\{|E(x|\mathscr{F}_n)-x|>2\varepsilon,\ \text{对某个 } n\geqslant n_0\}\\ &\quad\leqslant P\{E(y|\mathscr{F}_n)>\varepsilon,\ \text{对某个 } n\geqslant n_0\}+P\{y>\varepsilon\}\\ &\quad\leqslant\frac{\varepsilon}{2}+\frac{1}{\varepsilon}Ey\leqslant\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon. \end{aligned}$$

6. 本性上确界

设 $\{x_t, t\in T\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一族实值随机变量. 如果一个实值随机变量 y 满足：

(i) 对每个 $t\in T$，$P\{y\geqslant x_t\}=1$，

(ii) 若 y' 是任一实值随机变量，使得对每个 $t\in T$，

$$P\{y'>x_t\}=1, \text{则 } P\{y'\geqslant y\}=1,$$
则称 y 是 $\{x_t,\ t\in T\}$ 的本性上确界，并记作 $y=\operatorname*{ess\,sup}_{t\in T} x_t$.

定理 1.5 $y=\operatorname*{ess\,sup}_{t\in T} x_t$ 一定存在，而且有 T 的某个可列子集 C 使得

$$y=\sup_{t\in C} x_t.$$

证明 我们可以假设对一切 $t\in T$ 及 $\omega\in\Omega$, $|x_t(\omega)|\leqslant\frac{\pi}{2}$, 因为如有必要可考虑 $\operatorname{arc\,tg} x_t(t\in T)$. 令 b 表示

$$E(\sup_{t\in A} x_t)$$

当 A 取遍 T 的一切有限子集时的上确界. 令 $C=\bigcup\limits_{1}^{\infty}A_n$, 其中 A_n 满足

$$E(\sup_{t\in A_n} x_t)\geqslant b-n^{-1}.$$

令 $y=\sup\limits_{t\in C} x_t$, C 显然是可列的，且容易验证 y 具有性质 (i) 和 (ii).

7. 独立随机变量与强大数定律

$\mathscr{F}$ 的子 σ–代数的有限族 $\{\mathscr{F}_1,\cdots,\mathscr{F}_k\}$ 称为是独立的，若对每一组

$$A_1,\ \cdots,\ A_k,\quad \text{其中 } A_i\in\mathscr{F}_i,\ i=1,\ \cdots,\ k,$$

有
$$P(A_1\cap\cdots\cap A_k)=P(A_1)\cdots P(A_k).$$

对任意的一族子 σ–代数 $\{\mathscr{F}_t,\ t\in T\}$, 如果它的每一个有限子族都是独立的，则称 $\{\mathscr{F}_t,\ t\in T\}$ 是独立的. 一族随机变量 $\{x_t,\ t\in T\}$ 称为是独立的，如果子 σ–代数族 $\{\mathscr{B}(x_t),\ t\in T\}$ 是独立的.

两个子 σ–代数 $\mathscr{F}_1$ 和 $\mathscr{F}_2$ 是独立的，当且仅当对每一个

$\mathscr{F}_2$- 可测的有界实值随机变量 x 有

$$E(x|\mathscr{F}_1)=Ex.$$

一族子 σ–代数 $\{\mathscr{F}_t,\ t\in T\}$ 是独立的当且仅当对 T 的任何两个不相交的有限子集 $\{t_1,\cdots,t_n\}$ 和 $\{s_1,\cdots,s_m\}$, $\mathscr{B}(\mathscr{F}_{t_1}\cup\cdots\cup\mathscr{F}_{t_n})$ 和 $\mathscr{B}(\mathscr{F}_{s_1}\cup\cdots\cup\mathscr{F}_{s_m})$ 是独立的.

定理 1.6 强大数定律 设 $x_1, x_2, \cdots$ 是独立同分布实值随机变量序列, 如果 Ex_1 存在, 则

$$P\{\lim_{n\to\infty} n^{-1}(x_1+\cdots+x_n)=Ex_1\}=1. \tag{1.22}$$

如果存在有限常数 c, 使得

$$P\{\lim_{n\to\infty} n^{-1}(x_1+\cdots+x_n)=c\}=1, \tag{1.23}$$

则 Ex_1 存在且等于 c.

证明 若 Ex_1 存在且有限, 则从 2.2 节 (c) 得到 (1.22). 现在设 $Ex_1=+\infty$ ($Ex_1=-\infty$ 的情形类似地处理), 对每个 $a>0$, 存在 $g=g(a)$, 使得

$$\int_{\{x_1<g\}} x_1\geqslant a.$$

定义 $x_n'=x_nI_{\{x_n<g\}}$, 则 $a\leqslant Ex_n'\leqslant g<+\infty$, 从而

$$\liminf_n n^{-1}\sum_1^n x_k\geqslant \lim_n n^{-1}\sum_1^n x_k'\geqslant a.$$

因为 a 是任意的, 由此可知这时 (1.22) 同样成立. 现在设 (1.23) 成立, 于是 $\frac{x_n}{n}\to 0$, 由 Borel–Cantelli 引理(见 2.4 节 (d)),

$$\sum_1^\infty P\{|x_1|>n\}=\sum_1^\infty P\{|x_n|>n\}<\infty,$$

于是 $E|x_1|<\infty$, 现在由定理的第一部分即得 $Ex_1=c$.

第 二 章

鞅. Wald 引理. 应用

本章建立鞅的几个基本定理，同时这一章也用来使读者熟悉在以后各章里展开的最优停止理论中使用的典型论证方法.

1. 定义. 例子. 收敛定理

设 $(\Omega, \mathscr{F}, P)$ 为一概率空间，I 是有序集

$$\{-\infty, \cdots, -1, 0, 1, \cdots, +\infty\}$$

的任一形为 (a, b), $[a, b)$, $(a, b]$ 或 $[a, b]$ 的区间. 下鞅是指一族 $\{x_n, \mathscr{F}_n, n\in I\}$，它满足：

（i） 对一切 $m<n$, $\mathscr{F}_m\subset\mathscr{F}_n\subset\mathscr{F}$；

（ii） 对一切 n, x_n 是 $\mathscr{F}_n$- 可测实随机变量且 $Ex_n^+<\infty$；

（iii） 对一切 $m<n$, $x_m\leqslant E(x_n|\mathscr{F}_m)$

（或等价地，对一切 $m<n$ 及 $A\in\mathscr{F}_m$, $\int_A x_m\leqslant\int_A x_n$）.

如果 I 既不包含 $+\infty$ 也不包含 $-\infty$，则 (iii) 可代之以：对一切 n, $n+1\in I$, $x_n\leqslant E(x_{n+1}|\mathscr{F}_n)$.

若 $\{-x_n, \mathscr{F}_n, n\in I\}$ 是下鞅，则称 $\{x_n, \mathscr{F}_n, n\in I\}$ 是上鞅. 若 $E|x_n|<\infty(n\in I)$ 且在 (iii) 中成立等号，即 $\{x_n, \mathscr{F}_n, n\in I\}$ 同时是上鞅与下鞅，则称它为鞅.

例子

(a) 设$\{\mathscr{F}_n, n\in I\}$满足(i)，$z$为一实随机变量且$E|z|<\infty$。令$x_n=E(z|\mathscr{F}_n)(n\in I)$，则对一切$m$，$n\in I(m<n)$，由(1.18)我们有

$$E(x_n|\mathscr{F}_m)=E[E(z|\mathscr{F}_n)|\mathscr{F}_m]=E(z|\mathscr{F}_m)=x_m,$$

故$\{x_n, \mathscr{F}_n, n\in I\}$是鞅.

(b) 设$y_1, y_2, \cdots$为一列独立、期望为零的实随机变量. 令$x_n=y_1+\cdots+y_n$，$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$，则$\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$是鞅.

(c) 设$\{\mathscr{F}_n, 1\leqslant n<\infty\}$满足(i). 对$\mathscr{F}_\infty\equiv\mathscr{B}(\bigcup\limits_1^\infty\mathscr{F}_n)$上的任一概率测度$Q$，以$Q_n$记$Q$在$\mathscr{F}_n$上的限制，并设每个$Q_n$对$P$在$\mathscr{F}_n$上的限制$P_n$绝对连续，即对每个$A\in\mathscr{F}_n$，$P_n(A)=0$时有$Q_n(A)=0$. 设$x_n=\dfrac{dQ_n}{dP_n}$为$Q_n$对$P_n$的Radon-Nikodym导数，则对任何$A\in\mathscr{F}_n$及一切$m>n$，

$$\int_A x_n=Q_n(A)=Q_m(A)=\int_A x_m,$$

由此得到$\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$是鞅. 特别，若$y_1, y_2, \cdots$独立同分布，概率密度为$f$，它是对直线的Borel集上的某个$\sigma$-有限测度$\mu$取的，而$g$是另一个关于$\mu$的概率密度，使得对直线上的Borel集$A$，$\int_A fd\mu=0$时有$\int_A gd\mu=0$，那么对$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$，

$$\left\{\frac{g(y_1)\cdots g(y_n)}{f(y_1)\cdots f(y_n)}, \mathscr{F}_n, 1\leqslant n<\infty\right\}$$

是鞅. 若对某个非零(实数)λ，$\varphi(\lambda)\equiv E(e^{\lambda y_1})<\infty$，我们可取$\mu(\cdot)=P\{y_1\in(\cdot)\}$(因此$f\equiv1$)和$g(y)=e^{\lambda y}/\varphi(\lambda)$，则

$\{e^{\lambda(y_1+\cdots+y_n)}/(\varphi(\lambda))^n,\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$ 是鞅，这事实也容易直接予以证明.

注 为简单起见,我们假设对每个 $n\geqslant1$, Q_n 对 P_n绝对连续.更一般地，我们可令 $R=\frac{1}{2}(P+Q)$, $\tilde{x}_n=\frac{dQ_n}{dR_n}\Big/\frac{dP_n}{dR_n}$, 则不难看出 $\{\tilde{x}_n,\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$ 是上鞅，在绝对连续情形它等于 $\{x_n,\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$.

(d) 设 $y_1,\ y_2,\ \cdots$独立同分布且 $E|y_1|<\infty$. 对 $n=1,\ 2,\ \cdots$, 令 $s_n=y_1+\cdots+y_n$, $x_{-n}=\frac{s_n}{n}$, $\mathscr{F}_{-n}=\mathscr{B}(s_n,\ s_{n+1},\ \cdots)$, 则由对称性,对任一 $n\geqslant1$, $k=1,\ 2,\ \cdots,\ n$,

$$E(y_1|\mathscr{F}_{-n})=E(y_k|\mathscr{F}_{-n}).$$

由于 $\sum_{k=1}^{n}E(y_k|\mathscr{F}_{-n})=E(s_n|\mathscr{F}_{-n})=s_n$, 故得 $E(y_1|\mathscr{F}_{-n})=\frac{s_n}{n}=x_{-n}$, 从而由上面的 (a) 可知 $\{x_n,\ \mathscr{F}_n,\ -\infty<n\leqslant-1\}$ 是鞅.

(e) 设 $y_1,\ y_2,\cdots$为一列随机变量,且对一切 $k\geqslant1$, $E|y_k|<\infty$. 令 $\mathscr{F}_0=\{\emptyset,\ \Omega\}$ 及 $\mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n)$, 则 $\left\{\sum_1^n(y_k-E(y_k|\mathscr{F}_{k-1})),\ \mathscr{F}_n,\ 1\leqslant n<\infty\right\}$ 是鞅. 若 $\{z_n,\mathscr{F}_n,1\leqslant n<\infty\}$ 是下鞅, $Ez_1^-<\infty$, 令

$$z_0=0,\qquad y_n=z_n-z_{n-1}\quad(n\geqslant1),$$

我们可记

$$z_n=\sum_1^n(y_k-E(y_k|\mathscr{F}_{k-1}))+\sum_1^n E(y_k|\mathscr{F}_{k-1}).$$

由下鞅性质

$$E(y_k|\mathscr{F}_{k-1})=E(z_k-z_{k-1}|\mathscr{F}_{k-1})\geqslant0\quad(k\geqslant2),$$

我们得到表示式

(2.1) $$z_n = x_n + \alpha_n,$$

其中 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是鞅, $Ez_1\leqslant\alpha_n\uparrow$.

引理 2.1 若 $\{x_n, \mathscr{F}_n, n\in I\}$ 是(下)鞅, φ 是实变量的实值(递增)凸函数, 对某个 $n_0\in I$, $E(\varphi(x_{n_0}))^+<\infty$, 则

$$\{\varphi(x_n), \mathscr{F}_n, n\in I, n\leqslant n_0\}$$

是下鞅.

证明 设 $m, n\in I, m\leqslant n\leqslant n_0$. 由条件期望的 Jensen 不等式(参阅 Loève[1], 348 页),

(2.2) $$\varphi(x_m)\leqslant\varphi(E(x_{n_0}|\mathscr{F}_m))\leqslant E(\varphi(x_{n_0})|\mathscr{F}_m).$$

故对一切 $m\in I$, $m\leqslant n_0$, $E(\varphi(x_m))^+<\infty$. 在不等式 (2.2) 中以 n 换 n_0 即完成了证明.

例子

若 $\{x_n, \mathscr{F}_n, n\in I\}$ 是鞅, 则 $\{|x_n|, \mathscr{F}_n, n\in I\}$ 是下鞅; 若对某个 n_0, $Ex_{n_0}^2<\infty$, 则 $\{x_n^2, \mathscr{F}_n, n\in I, n\leqslant n_0\}$ 也是下鞅. 若 $\{x_n, \mathscr{F}_n, n\in I\}$ 是下鞅, 则对任一实数 a,

$$\{\max(x_n, a), \mathscr{F}_n, n\in I\}$$

是下鞅; 特别, $\{x_n^+, \mathscr{F}_n, n\in I\}$ 是下鞅.

引理 2.2 设 $\{x_n, \mathscr{F}_n, n\in I\}$ 是下鞅, I 包含它的上确界. 则对每个实数 a, $\{\max(x_n, a), \mathscr{F}_n, n\in I\}$ 是一致可积下鞅. 若 $\{x_n, \mathscr{F}_n, n\in I\}$ 又是鞅, 则它本身是一致可积的.

证明 以 m 记 I 的上确界, 记 $x_n(a)=\max(x_n, a)$. 由引理 2.1, $\{x_n(a), \mathscr{F}_n, n\in I\}$ 是下鞅. 由此对一切 $c>0$, $n\in I$,

(2.3) $$cP\{x_n(a)>c\}\leqslant\int_{\{x_n(a)>c\}}x_n(a)$$
$$\leqslant\int_{\{x_n(a)>c\}}x_m(a)\leqslant E[x_m^+(a)].$$

(2.3) 的两个边项证明了当 $c\to\infty$ 时, 对 n 一致地 $P\{x_n(a)$

$>c\}\to 0$. 然后中间两项就建立了要证的一致可积性. 若 $\{x_n, \mathscr{F}_n, n\in I\}$ 是鞅, 把前面的结果应用于下鞅 $\{|x_n|, \mathscr{F}_n, n\in I\}$ 即知它是一致可积的.

设 $\{x_n, \mathscr{F}_n, 0\leqslant n<\infty\}$ 是下鞅, $Ex_0>-\infty$. 令 $y_0=x_0, y_n=x_n-x_{n-1}(n>0)$, 则 $x_n=\sum_{k=0}^{n} y_k$. 设 $u_0=1$, 对每个 $k>0$, u_k 是只取 0 及 1 两个值的 $\mathscr{F}_{k-1}$- 可测随机变量. 令 $\hat{x}_n=\sum_{k=0}^{n} u_k y_k$ $(0\leqslant n<\infty)$. 很清楚, $E\hat{x}_n$ 存在, 且对一切 $n=1, 2, \cdots$,

$$
\begin{aligned}
(2.4)\qquad E\hat{x}_n &= Ey_0+\sum_{k=1}^{n}\int_{\{u_k=1\}} y_k \\
&= Ey_0+\sum_{k=1}^{n}\int_{\{u_k=1\}} E(y_k|\mathscr{F}_{k-1}) \\
&\leqslant Ey_0+\sum_{k=1}^{n} Ey_k=Ex_n.
\end{aligned}
$$

不等式 (2.4) 可给予下列解释. 设 x_0 是一个赌徒的初始赌本, $y_1, y_2, \cdots$ 是进行一系列"有利的" $(E(y_k|\mathscr{F}_{k-1})\geqslant 0)$ 随机赌博后逐次增加的赌金. 序列 (u_k) 是这赌徒不参加某几局赌博的策略. u_k 为 $\mathscr{F}_{k-1}$-可测的要求是说赌徒是否参加第 k 局赌博只依赖于他过去的经历. 不等式 (2.4) 是说在赌了 n 局之后, 采用这一策略的赌徒的平均赌金不超过每局都赌的赌徒的平均赌金.

现在我们能够证明所谓的上穿不等式, 它是用来证明鞅收敛定理的. 设 $a_1, a_2, \cdots, a_n$ 为任何实数, (r, s) 为任一非空的区间. $a_1, \cdots, a_n$ 关于 (r, s) 的上穿数定义为这些数 a_k, $1\leqslant k\leqslant n$, 从 $\leqslant r$ 跑到 $\geqslant s$ 的次数. 正式地, 定义 $t_0=0$, 对每个 $m=1, 2, \cdots$, 定义 t_m 为使得

$$a_k\leqslant r \quad (\text{当 } m \text{ 为奇数时}),$$
$$a_k\geqslant s \quad (\text{当 } m \text{ 为偶数时})$$

的最小的 $k>t_{m+1}$ (如果存在). 上穿数为 β, 这里 2β 是使 t_m 有定义的最大的偶指标 m.

现在设 $\{x_i, \mathscr{F}_i, 1\leqslant i\leqslant n\}$ 是非负下鞅, 取前一段中的 $a_i=x_i$ 及 $r=0$. 令 $x_0\equiv 0$, $u_0\equiv 1$, 定义 $u_1, \cdots, u_n$ 如下:

$$\begin{aligned} u_i&=1, \quad 若\ t_m<i\leqslant t_{m+1}\ (m\ 为奇数),\\ &=0, \quad 若\ t_m<i\leqslant t_{m+1}\ (m\ 为偶数), \end{aligned}$$

这里如果 t_1 原来没有定义就令 $t_1=n$. 很清楚

$$s\beta\leqslant \sum_{i=0}^{n} u_i y_i=\hat{x}_n,$$

如同上面一样, 其中令 $y_i=x_i-x_{i-1}(1\leqslant i\leqslant n)$. 因为对每个 $i=1, \cdots, n$,

$$\{u_i=1\}=\bigcup_{m为奇数}(\{t_m<i\}-\{t_{m+1}<i\})\in\mathscr{F}_{i-1},$$

由 (2.4) 得到

$$sE\beta\leqslant E\hat{x}_n\leqslant Ex_n.$$

注意到下鞅 $\{x_i, \mathscr{F}_i, 1\leqslant i\leqslant n\}$ 关于任一区间 (r,s) 的上穿数与非负下鞅 $\{(x_i-r)^+, \mathscr{F}_i, 1\leqslant i\leqslant n\}$ 关于 $(0, s-r)$ 的上穿数是相同的, 这样我们就证明了下列

引理 2.3 上穿不等式 设 $\{x_i, \mathscr{F}_i, 1\leqslant i\leqslant n\}$ 是下鞅. $x_1, x_2, \cdots, x_n$ 关于任一非空区间 (r, s) 的上穿数 β 满足:

$$E\beta\leqslant(s-r)^{-1}E(x_n-r)^+\leqslant(s-r)^{-1}[Ex_n^++|r|].$$

定理 2.1 鞅收敛定理 设 $\{x_n, \mathscr{F}_n, -\infty<n<\infty\}$ 是下鞅. 令 $\mathscr{F}_{-\infty}=\bigcap_{-\infty}^{0}\mathscr{F}_n$, $\mathscr{F}_\infty=\mathscr{B}(\bigcup_{1}^{\infty}\mathscr{F}_n)$, 则

(a) 存在 $x_{-\infty}=\lim\limits_{n\to-\infty}x_n$ a. s.;

(b) $Ex_{-\infty}^+<\infty$;

(c) $\{x_n, \mathscr{F}_n, -\infty\leqslant n<\infty\}$ 是下鞅.

若 $\sup\limits_n Ex_n^+<\infty$, 则

(d) 存在 $x_\infty=\lim\limits_{n\to+\infty}x_n$ a. s.;

(e) $Ex_\infty^+<\infty$, 且若对某个 $n\geqslant-\infty$, $Ex_n>-\infty$, 则 $E|x_\infty|<\infty$;

(f) $\{x_n, \mathscr{F}_n, -\infty\leqslant n\leqslant\infty\}$ 是下鞅当且仅当 (x_n^+) 一致可积.

证明 (a) 令 $x^*=\limsup\limits_{n\to-\infty}x_n$, $x_*=\liminf\limits_{n\to-\infty}x_n$, 并设 $P\{x^*>x_*\}>0$. 因为 $\{x^*>x_*\}$ 是事件

$$B(r, s)=\{x^*>s>r>x_*\}$$

对全部有理数 $r<s$ 的并集, 故对某两有理数 $r<s$, $P(B(r,s))>0$. 现在, 在集合 $B(r, s)$ 上 $x_{-n}, \cdots, x_0$ 关于区间 (r, s) 的上穿数 β_n 随 n 而单调地趋于 $+\infty$, 因此 $\lim\limits_{n\to-\infty}E\beta_n=+\infty$. 但由引理 2.3,

$$E(\beta_n)\leqslant(s-r)^{-1}E(x_0-r)^+<\infty, \tag{2.5}$$

这个矛盾证明了 (a).

(b) 由引理 2.1, $\{x_n^+, \mathscr{F}_n, -\infty<n<\infty\}$ 是下鞅, 故 Ex_n^+ 随 n 而递增. 所以由 (a) 及 Fatou 引理,

$$Ex_{-\infty}^+\leqslant\lim_{n\to-\infty}Ex_n^+<Ex_0^+<\infty.$$

(c) 令 $x_n(a)=\max(x_n, a)$ $(-\infty\leqslant n<\infty)$. 由引理 2.2, $\{x_n(a), \mathscr{F}_n, -\infty<n\leqslant0\}$ 是一致可积下鞅. 设 $-\infty<m<n<\infty$, 则对任一 $A\in\mathscr{F}_{-\infty}\subset\mathscr{F}_m$,

$$\int_A x_m(a)\leqslant\int_A x_n(a).$$

令 $m\to-\infty$, 由一致可积性及 (a) 我们有

$$\int_A x_{-\infty}(a)\leqslant\int_A x_n(a).$$

再令 $a\to-\infty$, 证得 (c).

(d) 用类似于(a)的证法, 把 (2.5) 的右边换成 $(s-r)^{-1}\cdot E(x_n^+ + |r|)$, 按假设, 这数列当 $n\to\infty$ 时仍有界, 同样可证得(d).

(e) 由 Fatou 引理, $Ex_\infty^+ \leqslant \sup_n Ex_n^+ < \infty$. 为了证明 (e) 的第二部分, 我们指出 $E|x_n| = Ex_n^+ + Ex_n^- = 2Ex_n^+ - Ex_n$. 设整数 n_0 使得 $Ex_{n_0} > -\infty$. 由于Ex_n 递增, 由 Fatou 引理我们有

$$E|x_\infty| \leqslant \sup_{n>n_0} E|x_n| = \sup_{n>n_0} (2Ex_n^+ - Ex_n)$$
$$\leqslant 2\sup_{n>n_0} Ex_n^+ - Ex_{n_0} < \infty.$$

(f) 充分性的证明与 (e) 类似. 现设 $\{x_n, \mathscr{F}_n, -\infty \leqslant n \leqslant \infty\}$是下鞅, 则由引理 2.1, $\{x_n^+, \mathscr{F}_n, -\infty \leqslant n \leqslant \infty\}$ 是下鞅. 引理 2.2 证明了 (x_n^+) 是一致可积的.

2. 鞅收敛定理的应用

a. P. Lévy 定理(定理 1.4)

设 $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \cdots \subset \mathscr{F}$, $\mathscr{F}_\infty = \mathscr{B}\left(\bigcup_1^\infty \mathscr{F}_n\right)$; z 为期望有限的非负 $\mathscr{F}$-可测随机变量, 则 $\{E(z|\mathscr{F}_n), \mathscr{F}_n, 1 \leqslant n < \infty\}$ 是鞅 (2.1 节中的例 (a)), 且由引理 2.2, 是一致可积的. 由定理 2.1, $x_\infty = \lim_{n\to\infty} E(z|\mathscr{F}_n)$ 以概率 1 存在, 且

$$(2.6)\quad \int_A x_\infty = \int_A E(z|\mathscr{F}_n) = \int_A z \quad (A \in \mathscr{F}_n, n=1, 2, \cdots).$$

因此 (2.6) 对一切 $A \in \bigcup_1^\infty \mathscr{F}_n$ 成立. 由于 (2.6) 的两个边项都是 $\mathscr{F}_\infty$ 上的有限测度, 即与概率测度只差条件 (1.4), 由扩张定理(定理 1.1)的唯一性部分, (2.6)对一切 $A \in \mathscr{F}_\infty$ 成立. 所以由条件期望的定义, $x_\infty = E(z|\mathscr{F}_\infty)$, 故 $\lim_{n\to\infty} E(z|\mathscr{F}_n) =$

$E(x|\mathscr{F}_\infty)$.

b. Kolmogorov 0-1 律

设 $y_1, y_2, \cdots$ 为独立随机变量, $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$, A 为"尾事件", 即对一切 $n=1, 2, \cdots$, $A\in\mathscr{B}(y_{n+1}, y_{n+2}, \cdots)$. 于是对一切 n, $P(A|\mathscr{F}_n)=P(A)$. 但由 a, $P(A|\mathscr{F}_n)\to I_A$, 由此得出 $P(A)$ 必须为 0 或 1.

c. 强大数定律

按 2.1 节的例 (d) 的假定, 由定理 2.1(a) 及前面的 b, 我们看到

$$x_{-\infty}=\lim_{n\to\infty}\frac{s_n}{n}$$

几乎必然存在且为常数. 由定理 2.1(c), $Ex_{-\infty}=Ex_{-1}=Ey_1$, 由此得到 $\frac{s_n}{n}\to Ey_1$.

d. 概率比

设 $\{\mathscr{F}_n, n\geqslant1\}$, Q 及 $\{x_n, n\geqslant1\}$ 与 2.1 节的例 (c) 一样. 若 Q 在 $\mathscr{F}_\infty$ 上对 P 绝对连续, 即存在随机变量 x, 使得对一切 $A\in\mathscr{F}_\infty$,

$$Q(A)=\int_A x,$$

则 $x_n=E(x|\mathscr{F}_n)$, 可应用上面的 a. 另一方面, 设 Q 是奇异的, 即存在一个 $S\in\mathscr{F}_\infty$, 使得 $Q(S)=0$ 及 $P(S)=1$, 则从 $Q(A)=\int_A x_n\ (A\in\mathscr{F}_n, n=1, 2, \cdots)$, 定理 2.1 及 Fatou 引理我们有, 对一切 $A\in\bigcup\limits_1^\infty\mathscr{F}_n$,

$$\int_A x_\infty\leqslant Q(A). \tag{2.7}$$

利用扩张定理的唯一性部分，(2.7) 对一切 $A\in\mathscr{B}\left(\bigcup\limits_1^\infty\mathscr{F}_n\right)=\mathscr{F}_\infty$ 成立．令 $A=S$，我们有 $P\{x_\infty=0\}=1$.

一般地，$Q=pQ_1+(1-p)Q_2$ $(0\leqslant p\leqslant 1)$，其中 Q_1 为绝对连续，Q_2 为奇异．若 $0<p<1$，则 $0<P\{x_\infty=0\}<1$．容易看出，$\{x_n^{1/2},\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$一致可积，由此得到 $P\{x_\infty=0\}=1$，即 Q 为奇异的当且仅当 $\lim\limits_{n\to\infty}Ex_n^{1/2}=0$．如果 $y_1,\ y_2,\ \cdots$ 为独立同分布，具有关于某个 σ-有限测度 μ 的概率密度函数 f，而 g 是另一个关于 μ 的密度，$\mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n)$ 及 $x_n=\prod\limits_1^n\dfrac{g(y_k)}{f(y_k)}$ $(n\geqslant 1)$，则由 Cauchy-Schwarz 不等式，

$$E\left(\frac{g(y_k)}{f(y_k)}\right)^{\frac{1}{2}}=\int(gf)^{\frac{1}{2}}d\mu<\left[\int g\,d\mu\int f\,d\mu\right]^{\frac{1}{2}}=1.$$

因此 $Ex_n^{1/2}=\left[E\left(\dfrac{g(y_1)}{f(y_1)}\right)^{\frac{1}{2}}\right]^n\to 0\ (n\to\infty)$，由此得到 $P\{x_n\to 0\}=1$.

3. 停时——定义与基本性质

设 $(\Omega,\ \mathscr{F},P)$ 是一概率空间，$\{\mathscr{F}_n,n\in I\}$ 是一族递增的 $\mathscr{F}$ 的子 σ-代数．一个停时按定义是一个随机变量 t，使得

(i) $P\{t\in I\}=1$;

(ii) 对一切 $n\in I$, $\{t=n\}\in\mathscr{F}_n$.

若对每个 $n\in I$，x_n 是 $\mathscr{F}_n$-可测随机变量，t 是停时，则

$$x_t\equiv\sum_{n\in I}x_nI_{\{t=n\}}$$

是随机变量，因为对任一 Borel 集 B，

$$\{x_t\in B\}=\bigcup_{n\in I}\{x_n\in B,\ t=n\}\in\mathscr{F},$$

使得对一切 $n\in I$，$A\cap\{t=n\}\in\mathscr{F}_n$ 的集合 $A\in\mathscr{F}$ 全体是 $\mathscr{F}$

的一个子 σ-代数，它记作 $\mathscr{F}_t$. 容易看出 t 和 x_t 是 $\mathscr{F}_t$-可测的.

如果我们把一个鞅 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 的值 $x_1, x_2, \cdots$ 解释为一个参加一系列公平的赌博的赌徒的赌金，那么一个停时就是这个赌徒决定何时停止赌博的策略，而 x_t 是他最终的赌金. $\{t=n\}\in\mathscr{F}_n$ 的要求是说赌徒在时刻 n 是否停止赌博的决定只能依赖于他过去的经历而不依赖于尚未看到的将来的情况. 很自然地要问是否有

$$Ex_t=Ex_1,$$

即"公平"的性质是否在任一停时 t 仍保持. 这一节及下面各节就探讨这个问题. 我们指出，回答不可能是绝对肯定的. 例如，x_n 是掷一个均匀硬币 n 次所得正面朝上次数与反面朝上次数的差，t = 第一个 $n\geqslant 1$ 使 $x_n=1$ ($=\infty$, 若不存在这样的 n), 则 $P\{t<\infty\}=1$, 但 $Ex_t=1\neq 0=Ex_1$. 然而有

定理 2.2 设 $\{x_n, \mathscr{F}_n, n\in I\}$ 是下鞅，t 是停时. 令 $t(n)=\min(t, n)$ $(n\in I)$.

(a) 若对某个 $N\in I$, $P\{t\leqslant N\}=1$, 则 $x_t\leqslant E(x_N|\mathscr{F}_t)$.

(b) $\{x_{t(n)}, \mathscr{F}_n, n\in I\}$ 是下鞅. 特别，若 $+\infty\in I$, 则对一切 $n\in I$,

$$E(x_t|\mathscr{F}_n)\geqslant x_n, \quad 在 \{t\geqslant n\} 上. \tag{2.8}$$

(c) 若 $P\{t<\infty\}=1$, Ex_t 存在且

$$\liminf_{n\to\infty}\int_{\{t>n\}} x_n^+=0, \tag{2.9}$$

则 (2.8) 对每个 $n\in I$ 成立.

证明 (a) 由引理 2.1, 对一切 $n\in I$, $n\leqslant N$, $A\in\mathscr{F}_t$,

$$\int_{A(t=n)} x_n^+\leqslant\int_{A(t=n)} x_N^+. \tag{2.10}$$

把 (2.10) 对 $n(n\in I, n\leqslant N)$ 相加就证得

(2.11) $$Ex_t^+ \leqslant Ex_N^+ < \infty.$$

以 $x_n(x_N)$ 代替 $x_n^+(x_N^+)$，用同样的方法可证 (a).

(b) 由 (2.11)，对一切 $n\in I$, $Ex_{t(n)}^+<\infty$. 设 $n\in I$, $-\infty<n<\infty$，则对任一 $A\in\mathscr{F}_n$,

$$\begin{aligned}\int_A x_{t(n)} &= \int_{A\{t<n\}} x_t + \int_{A\{t\geq n\}} x_n \\ &\leqslant \int_{A\{t<n\}} x_t + \int_{A\{t\geq n\}} E(x_{n+1}|\mathscr{F}_n) \\ &= \int_{A\{t<n+1\}} x_t + \int_{A\{t\geq n+1\}} x_{n+1} \\ &= \int_A x_{t(n+1)}.\end{aligned}$$

因此 $\{x_{t(n)}, \mathscr{F}_n, n\in I, -\infty<n<\infty\}$ 是下鞅. 现在设 $+\infty\in I$, 并且 $\mathscr{F}_\infty=\mathscr{B}\left(\bigcup\limits_1^\infty \mathscr{F}_n\right)$, $x_\infty=\lim x_n$. 由 (a), $Ex_t^+\leqslant Ex_\infty^+<\infty$, 由引理 2.2, (x_n^+) 一致可积. 由于 $x_{t(n)}^+\leqslant x_t^++x_n^+$, 可得 $(x_{t(n)}^+)$ 是一致可积的，故由定理 2.1(f), $\{x_{t(n)}, \mathscr{F}_n, -\infty<n\leqslant\infty\}$ 是下鞅. 在一般情形, $\mathscr{F}_\infty\supset\mathscr{B}\left(\bigcup\limits_1^\infty \mathscr{F}_n\right)$, 还需要再加一段简短的论证，但我们省略这些细节了. $-\infty\in I$ 的情形可类似地处理.

(c) 设 $n\in I-\{+\infty\}$ 及 $A\in\mathscr{F}_n$. 由 (b), 对每个整数 $m>n$,

$$\int_{A\{t\geq n\}} x_n \leqslant \int_{A\{t\geq n\}} x_{t(m)} \leqslant \int_{A\{n\leq t\leq m\}} x_t + \int_{A\{t>m\}} x_m^+.$$

设子列 m' 使得

$$\int_{\{t>m'\}} x_{m'}^+ \to \liminf_m \int_{\{t>m\}} x_m^+ = 0.$$

令 m 沿子列 m' 趋于 ∞, 由于 $P\{t<\infty\}=1$, $\int_{A\{n\leq t\leq m\}} x_t \to$

$\int_{\Delta(t>n)} x_t$，因此

$$\int_{\Delta(t>n)} x_n \leqslant \int_{\Delta(t>n)} x_t.$$

这就完成了证明.

下面两个引理中的第二个有时在证明 (2.8) 成立时是有用的.设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是下鞅, $Ex_1\geqslant 0$, 令 $\mathscr{F}_0=\{\emptyset, \Omega\}$, $x_0=0$, $y_n=x_n-x_{n-1}$, $a_n=E(|y_n|\,|\mathscr{F}_{n-1})$, $b_n=E(y_n^+|\mathscr{F}_{n-1})$, $\sigma_n^2=E(y_n^2|\mathscr{F}_{n-1})\,(n=1, 2, \cdots)$.

引理 2.4 设 $y_n\geqslant 0\,(n=1, 2, \cdots)$，则对一切停时 t,

$$E\left(\sum_1^t y_k\right)=E\left(\sum_1^t E(y_k|\mathscr{F}_{k-1})\right).$$

证明

$$E\left(\sum_1^t y_k\right)=E\left(\sum_1^\infty I_{(t>k)}y_k\right)=\sum_1^\infty \int_{(t>k)} y_k$$

$$=\sum_1^\infty \int_{(t>k)} E(y_k|\mathscr{F}_{k-1})$$

$$=E\left(\sum_1^t E(y_k|\mathscr{F}_{k-1})\right),$$

这里用得到第一与第二个等号的相反步骤就得到最后一个等号.

引理 2.5 若停时 t 使得

$$E\left(\sum_1^t b_k\right)<\infty, \tag{2.12}$$

则 (2.8) 成立 $(n=1, 2, \cdots)$.

证明 由定理 2.2(c)，只要证明 Ex_t 存在及(2.9)成立. 由 (2.12) 及引理 2.4, $E\left(\sum_1^t y_k^+\right)=E\left(\sum_1^t b_k\right)<\infty$. 因此 $Ex_t^+ \leqslant E\left(\sum_1^t y_k^+\right)<\infty$, 且

$$\int_{\{t>n\}} x_n^+ \leqslant \int_{\{t>n\}} \left(\sum_1^n y_k^+\right) \leqslant \int_{\{t>n\}} \left(\sum_1^t y_k^+\right) \to 0 \quad (n\to\infty).$$

定理 2.3 设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是鞅, $Ex_1=0$, t 是任意停时.

(a) 若

$$(2.13)\qquad Ex_t \text{ 存在 } \quad \text{和} \quad \liminf \int_{\{t>n\}} |x_n| = 0,$$

则 $Ex_t=0$; 更一般地,

$$E(x_t | \mathscr{F}_n) = x_n, \quad \text{在}\{t\geqslant n\}\text{ 上}, \quad n=1, 2, \cdots.$$

(b) 若 $Ey_n^2<\infty, n=1,2,\cdots$ 及 (2.13) 的第二部分成立, 则 $Ex_t^2=E\left(\sum_1^t \sigma_n^2\right)$.

(c) 为了使 (2.13) 成立, 只要

$$E\left(\sum_1^t a_n\right)<\infty \quad \text{或} \quad E\left(\sum_1^t \sigma_n^2\right)<\infty.$$

证明 (a) 及 (c) 的第一部分立即由定理 2.2(c) 及引理 2.5 的证明得到.

(b) 令 $z_n=x_n^2-\sum_1^n \sigma_k^2$. 容易看出 $\{z_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是鞅. 对 $n=1, 2, \cdots$, 令 $t(n)=\min(t, n)$, 则由 (a),

$$Ez_{t(n)}=0,$$

因此

$$Ex_{t(n)}^2 = E\left(\sum_1^{t(n)} \sigma_k^2\right).$$

当 $n\to\infty$ 时 $t(n)\to t$, 由 Fatou 引理及单调收敛定理,

$$(2.14)\qquad Ex_t^2 \leqslant \lim Ex_{t(n)}^2 = \lim E\left(\sum_1^{t(n)} \sigma_k^2\right) = E\left(\sum_1^t \sigma_k^2\right).$$

为了得到反过来的不等式, 只要假设 $Ex_t^2<\infty$ 并证明

$$Ex_{t(n)}^2 \leqslant Ex_t^2, \quad n=1, 2, \cdots.$$

但对每个 n, 由定理 2.2(c) 及引理 2.2,

$$E(|x_t| \mid \mathscr{F}_n) \geqslant |x_n|, \quad \text{在 } \{t \geqslant n\} \text{ 上,}$$

因此由 Schwarz 不等式,

$$E(x_t^2 \mid \mathscr{F}_n) \geqslant (E(|x_t| \mid \mathscr{F}_n))^2 \geqslant x_n^2, \quad \text{在 } \{t \geqslant n\} \text{ 上.}$$

由此得出

$$Ex_t^2 = \int_{\{t<n\}} x_t^2 + \int_{\{t>n\}} x_t^2 \geqslant \int_{\{t<n\}} x_t^2 + \int_{\{t>n\}} x_n^2 = Ex_{t(n)}^2.$$

(c) 设 $E\left(\sum_1^t \sigma_n^2\right) < \infty$, 还设 $Ey_n^2 < \infty, n \geqslant 1$. 从 (2.14) 及 Schwarz 不等式,

$$E|x_t| \leqslant (Ex_t^2)^{1/2} < \infty \quad \text{和} \quad B = \liminf \int_{\{t>n\}} x_n^2 < \infty.$$

这样,

$$(2.15) \quad \liminf \int_{\{t>n\}} |x_n| = \liminf \int_{\{t>n,\, |x_n|>a\}} |x_n|$$

$$\leqslant \frac{1}{a} \liminf \int_{\{t>n,\, |x_n|>a\}} x_n^2$$

$$\leqslant Ba^{-1} \to 0, \quad \text{当 } a \to \infty \text{ 时.}$$

令 $y_n' = y_n I_{(t \geqslant n)}$ 即可去掉 $Ey_n^2 < \infty$ 的假设.

如果诸 y_n 是独立及零均值的, 则随机变量 a_n 及 σ_n^2 都是常数. 在这特殊情形, 定理 2.3(c) 的条件常常是很容易验证的. 例如, 若诸 y_n 有一致有界的一阶绝对矩, 则对一切使 $Et < \infty$ 的 t 有 $Ex_t = 0$. (若诸 y_n 为同分布, 这结果早已知晓为 Wald 引理.) 若 $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma^2 < \infty$, 则 $Et < \infty$ 蕴含了 $Ex_t = 0$ 及 $Ex_t^2 = \sigma^2 Et$. (关于定理 2.3 在独立情形的进一步的应用, 参阅定理 2.4 及 2.5.)

4. 停时的应用

a. Kolmogorov 不等式

设 n 为一正整数, $\{x_k, \mathscr{F}_k, 1 \leqslant k \leqslant n\}$ 是非负下鞅, 则对任

何 $\varepsilon>0$,

$$\varepsilon P\{\max_{1\leqslant k\leqslant n} x_k>\varepsilon\}\leqslant\int_{\{\max\limits_{1\leqslant k\leqslant n} x_k>\varepsilon\}} x_n\leqslant Ex_n. \tag{2.16}$$

事实上,如果 $A=\{\max\limits_{1\leqslant k\leqslant n} x_k>\varepsilon\}$ 及

$$\begin{aligned} t&=\text{第一个 } k\geqslant 1 \text{ 使得 } x_k>\varepsilon, && \text{在 } A \text{ 上},\\ &=n && \text{在 } A \text{ 外}, \end{aligned}$$

则容易看出 $A\in\mathscr{F}_t$,因此由定理 2.2(a),

$$\varepsilon P(A)\leqslant\int_A x_t\leqslant\int_A x_n\leqslant Ex_n.$$

若 $x_k=s_k^2$, 其中 s_k 是一列独立零均值方差有限的实随机变量的第 k 个部分和,这结果就给出了 Kolmogorov 不等式. 另一个不等式如下:

设 y_1, y_2, $\cdots$ 独立, $Ey_k=0$, $Ey_k^2=\sigma_k^2$ $(1\leqslant k\leqslant n)$. 令 $z_n=\max\limits_{1\leqslant k\leqslant n}|y_k|$, $x_k=\sum\limits_1^k y_i$ $(1\leqslant k\leqslant n)$,则对任何 $\varepsilon>0$,

$$P\{\max_{1\leqslant k\leqslant n}|x_k|>\varepsilon\}\geqslant 1-\frac{E(\varepsilon+z_n)^2}{\sum\limits_1^n\sigma_k^2}. \tag{2.17}$$

证明 令 $\mathscr{F}_k=\mathscr{B}(y_1,\cdots,y_k)$ $(1\leqslant k\leqslant n)$,

$$A=\{\max_{1\leqslant k\leqslant n}|x_k|>\varepsilon\},$$

$$\begin{aligned} t&=\text{第一个 } k\geqslant 1 \text{ 使得 } |x_k|>\varepsilon, && \text{在 } A \text{ 上},\\ &=n, && \text{在 } A \text{ 外}, \end{aligned}$$

则 t 是停时,由定理 2.3,

$$\left(\sum_1^n\sigma_k^2\right)P(\Omega-A)\leqslant E\left(\sum_1^t\sigma_k^2\right)=Ex_t^2\leqslant E(\varepsilon+z_n)^2,$$

这等价于 (2.17).

b. Hájek-Rényi-周元燊不等式

下述结果推广了第一个 Kolmogorov 不等式 (2.16). 设 $\{x_k, \mathscr{F}_k, 0\leqslant k\leqslant n\}$ 是非负下鞅, $Ex_0=0$, (c_k) 是一列非负常

数，则

$$c_k x_k = \sum_{i=1}^{k} (c_i x_i - c_{i-1} x_{i-1}) \leqslant \sum_{i=1}^{k} [c_i(x_i - x_{i-1}) + (c_i - c_{i-1})^+ x_{i-1}].$$

显然 $\left\{\sum_{i=1}^{k} [c_i(x_i - x_{i-1}) + (c_i - c_{i-1})^+ x_{i-1}],\ \mathscr{F}_k,\ 1 \leqslant k \leqslant n\right\}$ 是非负下鞅，因此由 (2.16)，

$$P\{\max_{1 \leqslant k \leqslant n} c_k x_k \geqslant 1\} \leqslant \sum_{k=1}^{n} [c_k E(x_k - x_{k-1}) + (c_k - c_{k-1})^+ E x_{k-1}].$$

c. 投票定理

设 $y_1, y_2, \cdots, y_n$ 为独立同分布（或更一般地，可交换的）期望有限的非负整值随机变量，令 $s_k = \sum_{1}^{k} y_i\ (1 \leqslant k \leqslant n)$，则

$$(2.18) \qquad P\{s_k < k \text{ 对一切 } 1 \leqslant k \leqslant n \mid s_n\} = \left(1 - \frac{s_n}{n}\right)^+.$$

证明 令 $\mathscr{F}_{-k} = \mathscr{B}(s_k, s_{k+1}, \cdots, s_n)$，$x_{-k} = \frac{s_k}{k}(1 \leqslant k \leqslant n)$，则 $\{x_k, \mathscr{F}_k, -n \leqslant k \leqslant -1\}$ 是鞅（参阅 2.1 节例 (d)）。因为 (2.18) 在 $\{s_n \geqslant n\}$ 上显然满足，故设 $s_n < n$. 定义

$$t = \inf\{k: -n \leqslant k \leqslant -1,\ x_k \geqslant 1\},$$

这里空集的 inf 规定为 -1. 因为在 $\left\{\max_{1 \leqslant k \leqslant n} \frac{s_k}{k} \geqslant 1\right\}$ 上 $x_t = 1$，在其它地方 $x_t = 0$，由定理 2.2，在 $\{s_n < n\}$ 上我们有

$$P\{s_k \geqslant k \text{ 对某个 } 1 \leqslant k \leqslant n \mid s_n\} = E(x_t \mid \mathscr{F}_{-n}) = x_{-n} = \frac{s_n}{n},$$

这蕴含了 (2.18).

d. Borel-Cantelli-Lévy 引理

设 $B_1, B_2, \cdots$ 为一列事件，$\{\mathscr{F}_n, 0 \leqslant n < \infty\}$ 为一列非降的 $\mathscr{F}$ 的子 σ-代数，使得 $B_n \in \mathscr{F}_n\ (1 \leqslant n < \infty)$，则除了一个零概率事件外，

$$\sum_1^\infty I_{B_n}<\infty \quad \text{当且仅当} \sum_1^\infty P(B_n|\mathscr{F}_{n-1})<\infty.$$

证明 令 $x_n=\sum_1^n [I_{B_k}-P(B_k|\mathscr{F}_{k-1})]$ $(1\leqslant n<\infty)$，则 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是鞅. 对任何 $c>0$, 令

$$t=\text{第一个 } n\geqslant 1 \text{ 使得 } x_n\geqslant c,$$
$$=\infty, \text{若不存在这样的 } n,$$

及 $t(n)=\min(t, n)$. 由定理 2.2(b), $\{x_{t(n)}, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是鞅. 很明显, 对一切 $n=1, 2, \cdots$, $x_{t(n)}\leqslant c-1$, 因此由定理 2.1, 以概率 1, $\lim\limits_{n\to\infty} x_{t(n)}$ 存在且有限, 即在 $\{t=\infty\}$ 上 $\lim\limits_{n\to\infty} x_n$ 存在且有限. 因为 $\{\sup x_n<\infty\}=\bigcup\limits_{c=1}^\infty\{\sup x_n<c\}$, 我们知道在 $\{\sup x_n<\infty\}$ 上 $\lim\limits_{n\to\infty} x_n$ 存在且有限, 因此

$$\text{在}\left\{\sum_1^\infty I_{B_k}<\infty\right\}\text{上,}\quad \sum_1^\infty P(B_k|\mathscr{F}_{k-1})<\infty.$$

对 $\{-x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 作同样的讨论即可完成证明.

e. 钟开莱-Fuchs 定理

设 $y_1, y_2, \cdots$ 独立同分布, $Ey_1=0$, $P\{y_1=0\}<1$, 令 $s_n=\sum_1^n y_k$. 钟与 Fuchs 的一个定理[1] 包含了 $P\{\limsup\limits_n s_n=+\infty, \liminf\limits_n s_n=-\infty\}=1$. 为了证明这结果, 令 $q=P\{$对一切 n, $s_n<0\}$ 及对每个 $n=1, 2, \cdots$, $A_n=\{$对一切 $k\neq n, s_k<s_n\}$, 则

$$\begin{aligned}P(A_n)&=P\{\text{对一切 } k<n, s_n-s_k>0 \text{ 和对一切 } k>n,\\&\qquad s_k-s_n<0\}\\&=P\{\text{对一切 } k<n,\ s_n-s_k>0\}q.\end{aligned}$$

令

$$t=\text{第一个 } n\geqslant 1 \text{ 使得 } s_n\leqslant 0,$$
$$=\infty, \text{若不存在这样的 } n,$$

则 $P\{$对一切 $k<n$, $s_n-s_k>0\}=P\{$对一切 $1\leqslant i\leqslant n-1$, $s_i>0\}=P\{t\geqslant n\}$. 因此 $P(A_n)=qP(t\geqslant n)$, $1\geqslant\sum_1^\infty P(A_n)=q\sum_1^\infty P(t\geqslant n)=qEt$. 若 $q>0$, 则 $Et<\infty$. 由 Wald 引理, $Es_t=0$, 因此 $P\{s_t=0\}=1$. 这与 $P(y_1=0)<1$ 的假设矛盾, 所以 $q=0$. 现在令 $t_0=s_0=0$ 及对每个 $n=1,2,\cdots$,

$$t_n=\inf\{k:k>t_{n-1},\ s_k\geqslant s_{t_{n-1}}\},$$

则对一切 n, $P(t_n<\infty)=1$, $s_{t_1}-s_{t_0}$, $s_{t_2}-s_{t_1}$, $\cdots$ 是独立同分布非负随机变量(不恒等于零), 因此 $P\{\lim_n s_{t_n}=+\infty\}=1$. 这证明了 $\limsup_n s_n=+\infty$. $\liminf$ 的情形由对称性可得.

f. $Et^{\frac{1}{\alpha}}<\infty$ 时的 Wald 引理

设 $y_1,y_2,\cdots$ 独立, $Ey_k=0(k=1,2,\cdots)$, 使得对某个 $1<\alpha\leqslant 2$,

$$\sup_n n^{-1}\sum_1^n E|y_k|^\alpha=B<\infty,$$

令 $\mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n)$ 及 $s_n=\sum_1^n y_k$, 则对每个使 $Et^{\frac{1}{\alpha}}<\infty$ 的停时 t, $Es_t=0$. ($\alpha=2$ 的情形由 Burkholder 及 Gundy[1] 得到. 我们按照他们的证明.) 由定理 2.2, $Es_{\min(t,n)}=0$, 因此由控制收敛定理, 只要证

$$E[\sup_n|s_{\min(t,n)}|]<\infty.$$

对每个 $u>0$, 由定理 2.2 及 (2.16) 我们有

$$\begin{aligned}(2.19)\qquad P\{\sup_n|s_{\min(t,n)}|\geqslant u\}&\leqslant P\{t\geqslant u^\alpha\}+P\{\sup_{n<u^\alpha}|s_{\min(t,n)}|\geqslant u\}\\&\leqslant P\{t\geqslant u^\alpha\}+u^{-\alpha}E[|s_{\min(t,u^\alpha)}|^\alpha].\end{aligned}$$

现在由 von Bahr 及 Esseen[1] 的一个不等式及引理 2.4 我们有

$$E[|s_{\min(t,u^\alpha)}|^\alpha]\leqslant 2E\left(\sum_1^{\min(t,u^\alpha)}E|y_k|^\alpha\right)$$
$$\leqslant 2Bu^\alpha P\{t\geqslant u^\alpha\}+2B\int_{\{t<u^\alpha\}}tdP.$$
因此由 (2.19),
$$E[\sup_n|s_{\min(t,n)}|]=\int_0^\infty P\{\sup_n|s_{\min(t,n)}|\geqslant u\}du$$
$$\leqslant(1+2B)\int_0^\infty P(t\geqslant u^\alpha)du$$
$$+2B\int_0^\infty u^{-\alpha}\int_{\{t<u^\alpha\}}tdPdu$$
$$=(1+2B)Et^{\alpha^{-1}}+2B\int_\Omega t\int_{t^{\alpha^{-1}}}^\infty u^{-\alpha}du\,dP$$
$$=(1+2B)Et^{\alpha^{-1}}+2B(\alpha-1)^{-1}Et^{\alpha^{-1}}<\infty.$$

5. 某些首次通过问题

设 y_1, y_2, …为独立随机变量,均值为 $\mu_1,\mu_2,\cdots$, 使得对某个 $0<\mu<\infty$,

(2.20) $$n^{-1}\sum_1^n\mu_k\to\mu\quad(n\to\infty).$$

令 $s_n=\sum_1^n y_k,\mathscr{F}_n=\mathscr{B}(y_1,\cdots,y_n)\quad(n\geqslant1)$, 对每个 $c>0$, 定义

$t=t(c)=$第一个 $n\geqslant1$ 使得 $s_n>c$,

$=\infty$,若不存在这样的 n.

在本节中我们要找使 $c\to\infty$ 时,

(2.21) $$Et\sim\frac{c}{\mu},$$

(2.22) $$\operatorname{Var}t\sim(\text{const})\cdot c$$

的条件.((2.22) 中出现的常数在定理2.5中给出.)这些结果推广了更新理论中的著名结果: 若诸 y_n 同分布、非负, 则 (2.21) 成立,若更有 $\sigma^2=Ey_1^2-\mu^2<\infty$, 则 (2.22) 也成立,且 $\text{const}=\frac{\sigma^2}{\mu^3}$.

定理 2.4 若对任一 $\varepsilon>0$,

$$(2.23)\qquad \lim_{n\to\infty} n^{-1}\sum_{k=1}^{n}\int_{(y_k-\mu_k>\varepsilon n)}(y_k-\mu_k)=0,$$

$$(2.24)\qquad \sup_n n^{-1}\sum_1^n E(y_k-\mu_k)<\infty,$$

则对每个 $c>0$, $Et<\infty$, 且 $c\to\infty$ 时 (2.21) 成立.

证明 设 $Et=\infty$, 令 $\tau=\min(t,\ n)\,(n=1,\ 2,\ \cdots)$, 则 $E\tau\to\infty$. 由 (2.20) 及定理 2.3,

$$(2.25)\qquad \mu E\tau+o(E\tau)=E\left(\sum_1^{\tau}\mu_k\right)=Es_\tau=c+E(s_\tau-c).$$

在引理 2.6 中将证明

$$(2.26)\qquad E(s_\tau-c)\leqslant Ey_\tau^+=o(E\tau)\quad (n\to\infty),$$

这连同(2.25)和 $n\to\infty$ 时 $E\tau\to\infty$ 的假设矛盾. 因此 $Et<\infty$, 且由(2.20),

$$E\left(\left|\sum_1^t\mu_k\right|\right)<\infty.$$

由定义, $s_t>c$, 故得 Es_t 及 $\left(Es_t-\sum_1^t\mu_k\right)$ 存在. 从(2.24)得到

$$\sup_n n^{-1}\sum_1^n E|y_k-\mu_k|<\infty,$$

因此 $$E\left(\sum_1^t E|y_k-\mu_k|\right)\leqslant \text{const}\cdot Et<\infty.$$

这样由 (2.20) 及定理 2.3, 以 t 代 τ 时 (2.25) 成立. 再次引用下面的引理 2.6, 我们有 $0<E(s_t-c)\leqslant Ey_t=o(Et)$, 它会同 (2.25) ($\tau$ 被 t 代) 即证明了 (2.21).

下列引理证明了 (2.26) 以及以 t 代 τ 所得之相应命题. 考虑到进一步的应用(定理 2.5), 它叙述得比所需要的更为一般.

引理 2.6 若 (2.20), (2.23) 及 (2.24) 成立, 则对任一族递增的停时 $\{\tau(r),\ r>0\}$, 如果当 $r\to\infty$ 时,

$$\infty > E\tau(r) \to \infty,$$

则对每个 $r>0$, 我们有 $Ey_\tau^+ < \infty$ 及

$$Ey_\tau^+ = o(E\tau) \quad (r\to\infty).$$

令 $\sigma_n^2 = Ey_n^2 - \mu_n^2$, $b_n^2 = \sum_1^n \sigma_k^2 \ (n=1, 2, \cdots)$. 若

(2.27) $$\mu_n = o(n^{1/2}),$$

对某个 $0<\sigma^2<\infty$,

(2.28) $$b_n^2 \sim n\sigma^2,$$

且若对每个 $\varepsilon>0$,

(2.29) $$\lim_{n\to\infty} n^{-1} \sum_1^n \int_{(y_k-\mu_k>\varepsilon n^{1/2})} (y_k-\mu_k)^2 = 0,$$

则对每个 $r>0$, $E(y_\tau^+)^2<\infty$ 及 $E(y_\tau^+)^2 = o(E\tau) \quad (r\to\infty)$.

证明 我们将只证明第一部分; 第二部分的证明是类似的. 设 ε 为任意正数, 令 $\tilde{y}_k = y_k - \mu_k$, 我们有

(2.30) $$E\tilde{y}_\tau^+ \leqslant \varepsilon E\tau + \int_{(\tilde{y}_\tau>\varepsilon\tau)} \tilde{y}_\tau \leqslant \varepsilon E\tau + E\left(\sum_1^\tau I_{(\tilde{y}_k>\varepsilon k)} \tilde{y}_k\right)$$

$$= \varepsilon E\tau + E\left(\sum_1^\tau \int_{(\tilde{y}_k>\varepsilon k)} \tilde{y}_k\right),$$

这里等式由引理 2.4 得到. 由 (2.24) 我们有

$$\sup_n n^{-1} \sum_1^n E|\tilde{y}_k| = B < \infty.$$

因此对一切充分大的 n, 由 (2.23) 得到

(2.31) $$\sum_1^n \int_{(\tilde{y}_k>\varepsilon k)} \tilde{y}_k \leqslant \sum_1^{[\varepsilon n]} E|\tilde{y}_k| + \sum_{[\varepsilon n]+1}^n \int_{(\tilde{y}_k>\varepsilon^2 n)} \tilde{y}_k$$

$$\leqslant \varepsilon Bn + \varepsilon n.$$

由 (2.20), $\mu_n = o(n)$, 从 (2.30) 及 (2.31) 我们有

$$Ey_\tau^+ \leqslant E\tilde{y}_\tau^+ + E|\mu_\tau| \leqslant \varepsilon(3+B)E\tau + O(1),$$

因为 $\varepsilon>0$ 是任意的, 由此我们得到

$$Ey_\tau^+ = o(E\tau).$$

定理 2.5 设

(2.32) $$\sum_1^n \mu_k - n\mu = o(n^{1/2}),$$

又 (2.28), (2.29) 成立, 则 $c\to\infty$ 时,

(2.33) $$Et = \frac{c}{\mu} + o(c^{1/2}),$$

(2.34) $$\operatorname{Var} t \sim \frac{\sigma^2}{\mu^3} c.$$

证明 为说明方便起见我们将设 $\mu_n \equiv \mu$ 及 $\sigma_n^2 \equiv \sigma^2$. 如同定理 2.4 的证明一样, 可证

(2.35) $$Et < \infty \qquad Et \sim \frac{c}{\mu} \quad (c\to\infty).$$

利用定理 2.3(b) 类似地可证

(2.36) $$Et^2 < \infty \quad (c > 0).$$

由引理 2.6 及 (2.35),

(2.37) $$[E(s_t - c)]^2 \leqslant E(s_t - c)^2 \leqslant E(y_t^+)^2 = o(Et) = o(c),$$

这样由定理 2.3,

$$\mu Et = c + E(s_t - c) = c + o(c^{1/2}).$$

这证明了 (2.33).

由 (2.36) 及定理 2.3 我们有

$$\begin{aligned}\frac{\sigma^2}{\mu}(c + E(s_t - c)) &= \frac{\sigma^2}{\mu} Es_t = \sigma^2 Et = E(s_t - \mu t)^2 \\ &= E(s_t - c + c - \mu t)^2 = E(s_t - c)^2 \\ &\quad + 2\mu E\left(\frac{c}{\mu} - t\right)(s_t - c) + \mu^2 E\left(t - \frac{c}{\mu}\right)^2,\end{aligned}$$

所以

(2.38) $$\mu^2 E\left(t - \frac{c}{\mu}\right)^2 = \frac{\sigma^2}{\mu} c + 2\mu E\left(t - \frac{c}{\mu}\right)(s_t - c) + \frac{\sigma^2}{\mu} E(s_t - c) - E(s_t - c)^2.$$

由 (2.37) 及 Schwarz 不等式,

(2.39) $$\left|E\left(t-\frac{c}{\mu}\right)(s_t-c)\right|\leqslant\left[E\left(t-\frac{c}{\mu}\right)^2E(s_t-c)^2\right]^{1/2}$$
$$=o(c^{1/2})\left[E\left(t-\frac{c}{\mu}\right)^2\right]^{1/2}.$$
从 (2.37)—(2.39) 我们有
$$\mu^2E\left(t-\frac{c}{\mu}\right)^2\leqslant\frac{\sigma^2}{\mu}c+o(c^{1/2})\left[E\left(t-\frac{c}{\mu}\right)^2\right]^{1/2}+o(c),$$
由此得到

(2.40) $$E\left(t-\frac{c}{\mu}\right)^2=O(c).$$

所以由 (2.37)—(2.40),

(2.41) $$E\left(t-\frac{c}{\mu}\right)^2=\frac{\sigma^2}{\mu^3}c+o(c).$$

但由 (2.37),
$$\mathrm{Var}\,t=E\left(t-\frac{c}{\mu}\right)^2-\left[E\left(t-\frac{c}{\mu}\right)\right]^2$$
$$=E\left(t-\frac{c}{\mu}\right)^2-\left[\frac{1}{\mu}E(s_t-c)\right]^2=E\left(t-\frac{c}{\mu}\right)^2+o(c).$$
这连同 (2.41) 给出了 (2.34).

6. 鞅 $x_n=\dfrac{dQ_n}{dP_n}$

在本节中我们将较为详细地研究在例 2.1(c) 中介绍的那类鞅 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$. 这些结果的应用在下一节及本章末的习题中给出. 由定理 2.1 (也参看 2.2 节 (d)), $x_\infty=\lim\limits_{n\to\infty}x_n$ 存在和 $\{x_n, \mathscr{F}_n, 1\leqslant n\leqslant\infty\}$ 是上鞅, 其中如通常一样, 令 $\mathscr{F}_\infty=\mathscr{B}\left(\bigcup\limits_1^\infty\mathscr{F}_n\right)$. 由定理 2.2, 对任一有界停时 t,
$$Ex_t=Ex_1=1.$$
可以应用定理 2.2 把这等式推广到更广的一类停时上去, 但

从首次到达原则出发做更容易些.

对任一停时 t 及 $A\in\mathscr{F}_t$($\mathscr{F}_t$ 的定义见 2.3 节),

$$(2.42)\qquad Q(A\{t<\infty\})=\sum_{n=1}^{\infty}Q(A\{t=n\})=\sum_{n=1}^{\infty}\int_{A\{t=n\}}\frac{dQ_n}{dP_n}dP$$

$$=\int_{A\{t<\infty\}}x_t\,dP.$$

置 $A=\Omega$, 由 (2.42) 得到

$$(2.43)\qquad \int_{\{t<\infty\}}x_t\,dP=1$$

当且仅当

$$(2.44)\qquad Q\{t<\infty\}=1.$$

在例 2.1(c) 中已看到, 缺 $\frac{dQ_n}{dP_n}$ 的一个特殊情形是

$$(2.45)\qquad \frac{dQ_n}{dP_n}=\frac{e^{\lambda s_n}}{(\varphi(\lambda))^n},$$

其中 $y_1, y_2, \cdots$ 是独立同分布随机变量, 使得对某(实数) $\lambda\neq0$, $\varphi(\lambda)=E(e^{\lambda y_1})<\infty$, $s_n=\sum_1^n y_k$. 下述结果对形为 (2.45) 的 $\frac{dQ_n}{dP_n}$, 在决定 (2.44) (从而 (2.43)) 是否对一个特定的停时 t 成立时是有用的.

引理 2.7 以 $F(G)$ 记 y_1 在 $P(Q)$ 之下的分布, 则

(a) $D=\{\lambda:\varphi(\lambda)<\infty\}$ 是一个包含 0 的区间.

(b) φ 在 D 上任意次可微(在 D 的端点, 采用通常关于单边与无限导数的规定), 且

$$\varphi^{(k)}(\lambda)=\int_{-\infty}^{\infty}x^k e^{\lambda x}dF(x),$$

因此 $\varphi^{(k)}(\lambda)/\varphi(\lambda)=\int_{-\infty}^{\infty}x^k dG(x)$.

(c) 除非 F 集中于 0 点, φ 在 D 上严格凸.

(d) 至多存在一个非零的 λ_1 使 $\varphi(\lambda_1)=1$. 若存在这样一

个 λ_1, 它与 Ey_1 符号相反. λ_1 存在的一个充分条件是 $Ey_1 \neq 0$ 及 D 是一个开区间.

证明 证明是指数函数的凸性及单调收敛定理的直接的推论.

引理 2.8 设 $y_1, y_2, \cdots$ 独立同分布, $\mu = Ey_1$ 存在. 对任一实数 b, 定义

$$(2.46)\qquad \begin{aligned} t &= \text{第一个 } n \geqslant 1 \text{ 使得 } s_n \geqslant b, \\ &= \infty, \text{若不存在这样的 } n, \end{aligned}$$

若 $P(y_1 < b) > 0$ 及 $P(y_1 = 0) < 1$, 则 $P(t < \infty) = 1$ 当且仅当 $\mu \geqslant 0$.

证明 不失一般性可设 $b \geqslant 0$. 若 $0 < \mu \leqslant \infty$, $P(t < \infty) = 1$ 立即从强大数定律得到 (定理 1.6——也可参阅 2.2 节 (c)). 若 $\mu = 0$, 钟开莱–Fuchs 定理 (2.4 节 (e)) 蕴含着 $\limsup\limits_{n \to \infty} s_n = +\infty$, 因此在这情形 $P(t < \infty) = 1$ 也成立. 其次设 $-\infty \leqslant \mu < 0$, 并用反证法假设 $P(t < \infty) = 1$. 如同 2.4 节 (e) 中那样, 通过定义 $t_1, t_2, \cdots$, 可断定 $P\{\limsup\limits_n s_n = +\infty\} = 1$, 这与强大数定律矛盾, 因强大数定律意味着 $P\{\lim\limits_n s_n = -\infty\} = 1$.

注 若 $Ey_1^2 < \infty$, 对 $\mu = 0$ 的情形可直接应用中心极限定理及 Kolmogorov 0–1 律. 这条证明路线在 4.7 节的不同内容中得到发展.

定理 2.6 设 $y_1, y_2, \cdots$ 独立同分布, 对某个 $\lambda \neq 0$, $\varphi(\lambda) < \infty$. 以 D 记使 $\varphi(\lambda) < \infty$ 的 λ 值的区间, t 由 (2.46) 定义. 若 $P(y_1 < b) > 0$, 则对任一 $\lambda \in D$, (2.43) 当且仅当 $\varphi'(\lambda) \geqslant 0$ 时成立.

证明 首先注意到 $Q\{y_1 < b\} > 0$ 当且仅当 $P\{y_1 < b\} > 0$.

由引理 2.8, $Q\{t<\infty\}=1$ 当且仅当 $\int_{-\infty}^{\infty}xdG(x)\geqslant 0$, 由引理 2.7(b), 这等价于 $\varphi'(\lambda)\geqslant 0$.

推论(参阅习题 6) 设 y_1, y_2, … 独立同分布, 对某个 $\lambda\neq 0$, $\varphi(\lambda)<\infty$. 对任何实数 a, $b(a<b)$, 令 $t=$ 第一个 $n\geqslant 1$ 使得 $s_n\leqslant a$ 或 $s_n\geqslant b$, 则对每个使 $\varphi(\lambda)<\infty$ 的 λ, (2.43) 成立.

证明 令 $t_1=$ 第一个 $n\geqslant 1$ 使得 $s_n\geqslant b$,

$=\infty$, 若不存在这样的 n,

$t_2=$ 第一个 $n\geqslant 1$ 使得 $s_n\leqslant a$,

$=\infty$, 若不存在这样的 n.

则 $t=\min(t_1, t_2)$. 若 $P\{y_1\in(a, b)\}=0$, 结论显然正确. 不是这种情形, 结论也直接由定理得到, 因为

$$Q\{t>n\}\leqslant\min(Q\{t_1>n\}, Q\{t_2>n\}). \tag{2.47}$$

对任一使 $\varphi(\lambda)<\infty$ 的 λ, (2.47) 右边至少有一项在 $n\to\infty$ 时趋于零.

7. 应用于序贯概率比检验

设 f_0 与 f_1 是两个关于某 σ-有限测度 μ 的概率密度函数, 为方便起见, 假定对任一 Borel 集 A, $\int_A f_0d\mu=0$ 当且仅当 $\int_A f_1d\mu=0$. 设 y_1, y_2, … 独立同分布, 密度 f 未知. 为了对备择假设 H_1: $f=f_1$, 检验原假设 H_0: $f=f_0$, A. Wald[1] 提出了下列程序: 令 $f_{in}=\prod_{k=1}^{n}f_i(y_k)$ $(i=0, 1; n=1, 2, \cdots)$. 取正数 $A<1<B$, 并定义 $t=$ 第一个 $n\geqslant 1$ 使得 $\frac{f_{1n}}{f_{0n}}\geqslant B$ 或 $\leqslant A$, 根据 $\frac{f_{1t}}{f_{0t}}\leqslant A$ 或 $\geqslant B$ 接受 H_0 或 H_1.

以 P_i 记概率测度，诸 y_n 在 P_i 之下有密度函数 $f_i, i=0, 1$. 由 2.2 节 (d),

$$P_0\left\{\frac{f_{1m}}{f_{0m}}\to 0\right\}=1=P_1\left\{\frac{f_{1m}}{f_{0m}}\to\infty\right\}.$$

因此 $P_i\{t<\infty\}=1$ $(i=0,1)$. 以 α 与 β 分别记第一类与第二类错误概率，由 (2.42) 我们得到

$$\alpha\equiv P_0\{f_{1t}\geqslant Bf_{0t}\}=\int_{\{f_{1t}\geqslant Bf_{0t}\}}\frac{f_{0t}}{f_{1t}}dP_1 \tag{2.48}$$

$$\leqslant B^{-1}P_1\{f_{1t}\geqslant Bf_{0t}\}=B^{-1}(1-\beta),$$

$$\beta\equiv P_1\{f_{1t}\leqslant Af_{0t}\}=\int_{\{f_{1t}\leqslant Af_{0t}\}}\frac{f_{1t}}{f_{0t}}dP_0 \tag{2.49}$$

$$\leqslant AP_0\{f_{1t}\leqslant Af_{0t}\}=A(1-\alpha).$$

“忽略” $\dfrac{f_{1t}}{f_{0t}}$ 超过边界 A 与 B 的部分，可把不等式 (2.48) 与 (2.49) 看成近似的等式，我们有

$$\alpha\cong\frac{1-A}{B-A},\quad \beta\cong\frac{A(B-1)}{B-A}. \tag{2.50}$$

Wald 还建议把他的序贯概率比检验应用于有关某些复杂假设的问题，因而就有兴趣去研究 t 及 $\dfrac{f_{1t}}{f_{0t}}$ 在概率测度 P 之下的性状，$y_1, y_2, \cdots$ 在 P 之下独立同分布，但 P 不必是 P_0 或 P_1. 由习题 6, $P\{t<\infty\}=1$. 若存在非零的 λ_1, 使

$$E\left[\exp\left(\lambda_1\log\frac{f_1(y_1)}{f_0(y_1)}\right)\right]=1$$

(引理 2.6(d) 给出了存在这样一个 λ_1 的充分条件)，令 $a=\log A$, $b=\log B$, $u_n=\log(f_1(y_n)/f_0(y_n))$, $s_n=\sum_1^n u_k$, 并如同上面一样忽略 s_t 超过边界 a 与 b 的部分，从定理 2.4 的推论我们有

$$1=E(e^{\lambda_1 s_t})\cong e^{\lambda_1 b}P\{s_t\geqslant b\}+e^{\lambda_1 a}P\{s_t\leqslant a\}.$$

因此 $P\{$拒绝 $H_0\}=P\{f_{1t}\geqslant Bf_{0t}\}\cong\dfrac{1-A^{\lambda_1}}{B^{\lambda_1}-A^{\lambda_1}}$. 特别对 $P=P_0$, $\lambda_1=1$, 因此

$$P_0\{f_{1t}\geqslant Bf_{0t}\}\cong\frac{1-A}{B-A},$$

而对 $P=P_1$, $\lambda_1=-1$ 且

$$P_1\{f_{1t}\leqslant Af_{0t}\}\cong 1-\frac{1-A^{-1}}{B^{-1}-A^{-1}}=\frac{A(B-1)}{B-A},$$

与 (2.50) 相符.

类似地把定理 2.3 应用于鞅 s_n-nEu_1 ($s_n^2-nEu_1^2$, 若 $Eu_1=0$), 可得所需观察数的期望 Et 的近似式(参阅习题 1 及 13).

习　　题

1. **赌徒输光问题**　设 $y_1,y_2,\cdots$ 独立同分布, $P\{y_1=1\}=p$, $P\{y_1=-1\}=q=1-p$, 令 $s_n=\sum_1^n y_i$ $(n\geqslant 1)$. 对任意正整数 a, b, 令

$$t=\text{第一个 } n \text{ 使得 } s_n=-a \text{ 或 } +b.$$

证明:

$$P\{s_t=-a\}=\begin{cases}\dfrac{1-(p/q)^b}{1-(p/q)^{a+b}} & (p\neq q),\\[2ex] \dfrac{b}{a+b} & \left(p=q=\dfrac{1}{2}\right),\end{cases}$$

$$Et=\begin{cases}\dfrac{b}{p-q}-\dfrac{a+b}{p-q}\cdot\dfrac{1-(p/q)^b}{1-(p/q)^{a+b}} & (p\neq q),\\[2ex] ab & \left(p=q=\dfrac{1}{2}\right).\end{cases}$$

(提示: 若 $p\neq q$, $(q/p)^{s_n}$ 及 $s_n-n(p-q)$ 对适当的 σ-代数成为鞅, 前者是 $e^{\lambda s_n}/(\varphi(\lambda))^n$ 在 $\lambda=\lambda_1$ 点的特殊情形. 若 $p=q$, 合适的鞅是 s_n 及 s_n^2-n.)

2. (续)现在令

(2.51) $$t=\text{第一个 } n \text{ 使得 } s_n \geqslant b,$$
$$=\infty,\text{ 若不存在这样的 } n,$$

(a) 若 $p<q$，则 $P\{t<\infty\}=(p/q)^b$，

(b) t 的母函数为

(2.52) $$\sum_1^\infty s^n P\{t=n\}=\left(\frac{1-\sqrt{1-4pqs^2}}{2qs}\right)^b \quad (0<s\leqslant 1).$$

(c) 证明：若 $p>q$，则 $Et=\frac{b}{\mu}$；$\operatorname{Var} t=\frac{\sigma^2 b}{\mu^3}$，其中 $\mu=p-q$，$\sigma^2=1-(p-q)^2$。(利用定理 2.4 及 2.5 的思路能给出一个直接的证明．另一个方法是对 (2.52) 求导.)

(d) 若 $p>q$，则 $b\to\infty$ 时 $\frac{t-(b/\mu)}{\sqrt{\sigma^2 b/\mu^3}}$ 渐近正态分布.

(e) (a), (b) 及 (c) 中哪一个可用解习题 1，然后令 $a\to-\infty$ 的办法得到?

3. 设 $y_1, y_2, \cdots$ 独立同分布，分布函数 F 为
$$dF(x)=ce^{-\alpha x}dx \quad (0<x<\infty),$$
其中 $c>0$, $\alpha>0$，则对 (2.51) 定义的 t 有
$$\sum_1^\infty s^n P\{t=n\}=\frac{\alpha-\lambda}{\alpha}e^{-\lambda b},$$
其中对每个 $s(0<s\leqslant 1)$，$\lambda=\lambda(s)$ 是 $[0, \alpha)$ 中唯一满足 $s\varphi(\lambda)=1$ 的数.

4. 设 $y_1, y_2, \cdots$ 独立同分布，具有负均值 μ. 设对一切 λ, $\varphi(\lambda)<\infty$, t 由 (2.51) 定义．证明

(2.53) $$P\{t<\infty\}\leqslant e^{-\lambda_1 b},$$

其中 λ_1 满足 $\varphi(\lambda_1)=1$. 何时 (2.53) 中成立等号? 你能对习题 3 中的分布找到 $P\{t<\infty\}$ 的精确值吗?

5. 设 $y_1, y_2, \cdots$ 独立同分布，具有正均值 μ，对某个 $\lambda<0$, $\varphi(\lambda)<\infty$. t 由 (2.51) 定义．证明：对某个 $\lambda>0$, $Ee^{\lambda t}<\infty$. (提示：首先考虑 $y_i\leqslant a<\infty$ 的特殊情形.)

6. 给定理 2.4 的推论一个直接的证明．更一般地，由首次到达原则证明：若 $y_1, y_2, \cdots$ 独立同分布，$P\{y_1=0\}<1$，$N=$ 第一个 n 使得 $s_n\geqslant b$ 或 $\leqslant a$，则 $P\{N<\infty\}=1$. (提示：若 $P\{|y_1|\geqslant b-a\}>0$，则 $N\leqslant N^*=$

第一个 n 使得 $|y_n|\geqslant b-a$. 一般地, 存在一整数 r 使 $P\{|s_r|\geqslant b-a\}>0$).

习题 7—11 提供了鞅收敛定理的别的证法. 指标为 $\{\cdots, -1, 0\}$ 及 $\{1, 2, \cdots\}$ 的鞅用不同的方法处理. 在习题 8 中假定 Lévy 定理 (定理1.4)已知.

7. 设 $\{x_n, \mathscr{F}_n, -\infty\leqslant n<0\}$ 是鞅.

(a) 对任一 $A\in\mathscr{F}_{-\infty}=\bigcap\limits_{-\infty}^{0}\mathscr{F}_n$ 及 $-\infty<s<\infty$,

$$sP(A\{\max_{n<0} x_n>s\})\leqslant\int_{A\{\max\limits_{n\leqslant 0} x_n>s\}} x_0.$$

(这实质上是鞅的 Kolmogorov 不等式 (2.16).)

(b) 对 x_n 及 $-x_n$ 应用 (a) 得出: 若 r, s 为有理数 $(r<s)$,

$$A=\{\limsup_{n\to-\infty} x_n>s>r>\liminf_{n\to-\infty} x_n\},$$

则 $P(A)=0$. 因此 a. s. $\lim\limits_{n\to-\infty} x_n$ 存在.

8. 设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是鞅且 $\sup\limits_n E|x_n|<\infty$.

(a) 令 $z_n^{(1)}=\sup\limits_{k\geqslant n} E(x_k^+|\mathscr{F}_n)$, $z_n^{(2)}=z_n^{(1)}-x_n$, 则 $\{z_n^{(i)}, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是非负鞅 $(i=1, 2)$, 且 $x_n=z_n^{(1)}-z_n^{(2)}$. 因此为了证明 $\lim\limits_{n\to\infty} x_n$ 存在, 不失一般性我们可设 $x_n\geqslant 0$ $(n=1, 2, \cdots)$. 利用过渡到 $\min(x_n, b)$ 的办法实际上我们可设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是一致有界非负上鞅.

(b) 现在设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是一致有界非负上鞅, 令 $x_\infty=\limsup\limits_{n\to\infty} x_n$, 则对每个停时 $t\geqslant n$,

$$x_n\geqslant E(x_t|\mathscr{F}_n).$$

把上式应用于 $t=$ 第一个 $k\geqslant n$ 使得 $x_k\geqslant E(x_\infty|\mathscr{F}_k)-\varepsilon$, (由 Lévy 定理, $P\{t<\infty\}=1$) 得到

$$x_n\geqslant E(x_\infty|\mathscr{F}_n)\quad(n=1, 2, \cdots).$$

再用 Lévy 定理可得

$$\liminf_{n\to\infty} x_n\geqslant x_\infty=\limsup_{n\to\infty} x_n.$$

9. 利用表示式 (2.1) 把习题 7 与 8 的结果推广到下鞅.

10. 设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是非负上鞅．证明：对一切 $n=1, 2, \cdots$ 及 $r<s$，

$$P\{\max_{k\geqslant n} x_k \geqslant s \mid \mathscr{F}_n\} \leqslant \frac{r}{s}, \quad \text{在}\{x_n \leqslant r\}\text{上}.$$

11. 应用　设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是非负上鞅，对任何 $0\leqslant r<s$，如同证明上穿不等式时那样定义 $t_1, t_2, \cdots$．证明：

$$P\{t_{2m}<\infty\} \leqslant \frac{r}{s} P\{t_{2m-1}<\infty\}.$$

从而断定 a.s. $\lim\limits_{n\to\infty} x_n$ 存在．

12. 举例说明 (2.13) 中的两个条件中的任一个不包含另一个．

13. 用 2.7 节的记号，找一个 $Eu_1=0$ 时成立的 $P\{f_{1t} \geqslant Bf_{0t}\}$ 的近似式．找一个所需观察数的均值 Et 的近似式．（提示：再看看习题 1.）

14. 设 $y_1, y_2, \cdots$ 为独立随机变量，$Ey_k=0$，$Ey_k^2=\sigma_k^2 (k=1, 2, \cdots)$．利用 Hájek-Rényi 不等式 (2.4 节(b)) 证明：若 $\sum\limits_1^\infty \frac{\sigma_k^2}{k^2}<\infty$，则

$$n^{-1}\sum_1^n y_k \to 0.$$

第 三 章

引 论

现在我们能从事研究最优停止问题了．开始先正式地叙述问题．其次给出几个例子．这些例子将不时地重复出现．然后讨论有限情形和单调情形，它们多少能完全地得到解决而不用求助于第四章的一般理论．当然，一般理论确实有助于人们理解这些特殊问题；但是反过来，这章中给出的结果在某种程度上也促成了后面对一般问题的处理方法．

1．问题的叙述与例

假设我们能序贯地观察随机变量 $y_1, y_2, \cdots$，它们的联合分布是已知的．我们必须在某一点停止观察过程，如果我们在第 n 步停止，就得到“报酬” x_n，它是 $y_1, \cdots, y_n$ 的已知函数．我们感兴趣的是寻找使平均报酬达到最大的停止规则．

更正式地，我们假设给出：(i)概率空间$(\Omega, \mathscr{F}, P)$，(ii)一列递增的 $\mathscr{F}$ 的子 σ–代数$(\mathscr{F}_n)$，(iii)一列随机变量$x_1, x_2, \cdots$，使得 x_n 关于 $\mathscr{F}_n$ 可测，$n=1, 2, \cdots$．这样一对序列 $\{x_n, \mathscr{F}_n\}_1^\infty$ 称为随机序列．一个停止规则或停止变量是一个取值 $1, 2, \cdots, +\infty$ 的随机变量 t，它满足条件 $P\{t<\infty\}=1$ 及 $\{t=n\}\in\mathscr{F}_n$，$n=1, 2, \cdots$．(在利用前一段所述的直观背景时，$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$．记住这个直观解释将很便利，虽然

一般的结果并不依赖于这个解释. 采用这个解释, $\{t=n\}\in\mathscr{F}_n$ 的要求就意味着在时刻 n 停止的决定必须在已观察到的 $y_1, \cdots, y_n$ 的基础上作出, 决不涉及将来的值 $y_{n+1}, \cdots$.) 若 $\{x_n, \mathscr{F}_n\}_1^\infty$ 为任一随机序列, t 为任一停止变量, 则由

$$x_t=\sum_{n=1}^{\infty} x_n I_{(t=n)}=\begin{cases}x_n, & \text{在 } \{t=n\} \text{ 上}, \ n=1, 2, \cdots, \\ 0, & \text{在 } \{t=\infty\} \text{ 上},\end{cases}$$

定义的随机变量 x_t 是实际的报酬, 我们定义随机序列 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的值 V 为 $\sup Ex_t$, 这里上确界是对一切使 Ex_t 存在的停止变量 t 取的. 由于对任意的随机序列并不总是搞得清是否存在规则 t 使 Ex_t 存在, 因此假设 $\{x_n, \mathscr{F}_n\}_1^\infty$ 是一个可积随机序列, 即 $E|x_n|<\infty (n\geqslant 1)$, 将是方便的. 这样就可看到 $-\infty<Ex_1\leqslant V\leqslant\infty$. 我们感兴趣的是回答下列问题:

(a) 如何计算 V?

(b) 是否存在最优停止变量 t, 就是使 Ex_t 存在且等于 V 的 t?

(c) 若最优规则存在, 它的性质如何?

有时对某个停止变量的子类讨论最优停止问题是方便的. 若 D 是使 Ex_t 存在的停止变量 t 的任一个类, 我们定义 $V(D)=\sup\limits_{t\in D} Ex_t$, 并称 t 是 D 中最优的, 若 $t\in D$ 且 $Ex_t=V(D)$. 以 C 记一切使 $Ex_t^-<\infty$ 的停止变量 t 组成的类. 很明显, $V=V(C)$.

例子

(a) Ω 只有一个元素; $x_n=1-\dfrac{1}{n}$, $n=1, 2, \cdots$. 很明显, $V=1$, 但不存在最优规则.

(b) 设 (y_n) 独立同分布, $P(y_n=1)=p=1-P(y_n=0)$. 对 $n=1, 2, \cdots$, 令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$, $x_n=(y_1+\cdots+y_n)/n$. 容

易看出 $V\geqslant p(2-p)$. 因为对给定的 $\varepsilon>0$, 令

$$
\begin{aligned}
t&=1, \quad 若\ y_1=1,\\
&=\inf\{n: x_n\geqslant p-\varepsilon\}, \quad 若\ y_1=0.
\end{aligned}
$$

强大数定律表明 t 是一个停止变量, 且

$$V\geqslant Ex_t\geqslant p+(1-p)(p-\varepsilon).$$

令 $\varepsilon\to 0$ 即得上述结论. 然而并不能显然地知道 V 的精确值是多少(事实上还不知道一个把 V 表成 p 的函数公式), 是否存在最优规则(确实存在), 或 V 是否 p 的递增函数(确实是).

(c) 设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是一个上鞅(定义见 2.1 节), $Ex_1<\infty$. 由定理 2.2, 对停止变量 t, 若有 $1\leqslant N<\infty$ 使 $P(t\leqslant N)=1$, 则

$$Ex_t\leqslant Ex_1. \tag{3.1}$$

因此, 若以 D 记有界停止变量类, 则

$$V(D)=Ex_1 \quad 和 \quad t=1\ 是\ D\ 中最优的.$$

若 (x_n^-) 一致可积, 定理 2.2 表明 (3.1) 对每个停止变量 t 成立.

(d) (续) 设 $y_1, y_2, \cdots$ 独立同分布, 均值为 μ, 令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$, $x_n=y_1+\cdots+y_n$. 若 $\mu\leqslant 0$, 则 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是上鞅, 前一段的注记可适用. 若 $P(y_n=1)=\frac{1}{2}=P(y_n=-1)$, 则 $\mu=0$ 但 $V=+\infty$, 故 (3.1) 一般是不成立的. 然而我们用一个相当简单的直接办法能完全解决 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的最优停止问题.

定理 3.1 设 (y_n) 独立同分布, 期望为 μ, 对 $n=1, 2, \cdots$, 令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$, $x_n=y_1+\cdots+y_n$, 则

(i) 若 $\mu<0$, 则 $V=Ex_1$ 且 $t=1$ 是最优的;

(ii) 若 $\mu\geqslant 0$, 则 $V=\infty$, 除非 (y_n) 全恒等于 0, 那时 $V=0$.

证明 (ii)由引理2.8得到. (i)立即由下述Wald引理的推广可得(与定理2.3相比较).

引理3.1 设(y_n)独立同分布,期望为μ,对$n=1, 2, \cdots$,令$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$, $x_n=y_1+\cdots+y_n$, 则对任一使Ex_t存在的停止变量t,只要$\mu=0$与$Et=\infty$不同时发生,就有$Ex_t=\mu Et$. 为使Ex_t存在且$=\mu t$,只要μ与Et同时有限.

证明 不失一般性我们可设Ω是序列$\omega=(y_1, y_2, \cdots)$组成的空间及$y_n(\omega)=y_n(n\geqslant 1)$. 定义一个$\Omega$到$\Omega$上的变换:

$$Q(\omega)=(y_{t+1}(\omega), y_{t+2}(\omega), \cdots),$$

令$t_0=0$, $t_1(\omega)=t(\omega)$, $t_2(\omega)=tQ(\omega)$, $\cdots$, $t_n(\omega)=tQ^{n-1}(\omega)$, $\cdots$,其中$Q^1=Q$, $Q^{n+1}=Q(Q^n)$. 对$i=1, 2, \cdots$,定义z_i为

$$y_{t_1+\cdots+t_{i-1}+1}+\cdots+y_{t_1+\cdots+t_i},$$

容易看到(t_1, z_1), (t_2, z_2), $\cdots$是独立同分布向量. 假设$Ex_t=Ez_1$存在. 由于

$$(3.2)\qquad n^{-1}\left(\sum_1^n z_i\right)=\left(\frac{y_1+\cdots+y_{t_1+\cdots+t_n}}{t_1+\cdots+t_n}\right)\left(\frac{t_1+\cdots+t_n}{n}\right),$$

引理的第一部分由强大数定律可得. 若μ与Et都有限,则由上面的讨论,(3.2)的右边收敛于一个有限数,从而强大数定律断定了$Ez_1=Ex_t$存在且有限.

(e) 设想有个窃贼每天偷某一家人家. 他每天偷窃所得组成一列独立同分布期望有限的随机变量. 每天他被抓住并被迫归还全部赃物的概率为正数p. 我们认为窃贼在第n天被抓住这一事件与过去已发生的一切是独立的. 问题是窃贼如何最明智地选择洗手不干的时间.

更正式地,设$y_1, y_2, \cdots$为一列独立同分布期望有限的非负随机变量. 设$\delta_1, \delta_2, \cdots$为独立随机变量, $p=P\{\delta_n=0\}=1-P\{\delta_n=1\}$, $n=1, 2, \cdots$,并假设诸y_n与诸δ_n是独立的.

令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n, \delta_1, \cdots, \delta_n)$. 窃贼要解$\{x_n, \mathscr{F}_n\}_1^\infty$的最优停止问题，其中

$$x_n=(\delta_1\delta_2\cdots\delta_n)(y_1+\cdots+y_n).$$

(f) 一个经理面临一个从一群 N 个姑娘中雇用一名秘书的问题. 每次他会见一个姑娘，见面之后就决定接受或拒绝该姑娘. 如果他拒绝了一个姑娘，以后他就不能再召见她了. 每一次会见经理能知道现在正与他谈话的姑娘与她之前的姑娘相比可排在什么名次，但不知道把她与尚未见面的姑娘相比情况会怎样. 我们假设他按随机的次序会见姑娘，问他应如何决定雇用哪一位姑娘. 特别，我们要找这样一个规则，它使选中 N 个姑娘中最好的一位的概率最大，或者使选中的姑娘的绝对名次的均值最小(1 是最好的姑娘的名次，2 是次好的名次，…，N 是最差的名次).

正式地，设 $a_1, a_2, \cdots, a_N$ 记整数 $1, 2, \cdots, N$ 的一个排列，全部排列是等可能的. 整数 1 对应最好的姑娘，…，N 对应最差的. 对任一 $n=1, \cdots, N$，令 $y_n=a_1, \cdots, a_n$ 中 $\leqslant a_n$ 的项数 (y_n=第 n 个出现的姑娘的相对名次)，$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$. 若标准是使选中最好的姑娘的概率最大，那么在选中第 n 个姑娘，且 $a_n=1(n=1, \cdots, N)$时报酬为 1，不然的话为 0. 这里有一个小的技术性困难要克服，即这样一个报酬序列不满足对随机序列的可测性要求，也就是在 $a_n=1$ 时取 1、其它情况取 0 的随机变量在 $n<N$ 时并不 $\mathscr{F}_n$-可测. 但若令 $x_n=P(a_n=1|\mathscr{F}_n)$, $n=1, \cdots, N$, 则 x_n 是 $\mathscr{F}_n$-可测的，且对任何停止变量 t,

$$Ex_t=\sum_1^N\int_{\{t=n\}}x_n=\sum_1^N\int_{\{t=n\}}P(a_n=1|\mathscr{F}_n)$$

$$=\sum_1^N\int_{\{t=n\}}I_{\{a_n=1\}}=P\{a_t=1\}.$$

因此选取一个使选中最好的姑娘的概率最大的策略等价于解$\{x_n, \mathscr{F}_n\}_1^N$的最优停止问题. 另一方面,若标准是使选中的姑娘的平均绝对名次最小,我们应设

$$x_n = -E(a_n | \mathscr{F}_n),$$

并解$\{\tilde{x}_n, \mathscr{F}_n\}_1^N$的最优停止问题.

(g) 我们要使汽车停放地点尽可能地靠近一个所希望停放的位置,这位置规定为0点. 我们沿负半轴接近我们的目标. 对$n=\cdots, -2, -1, 0, 1, \cdots$,第$n$个停车点空着的概率为$p$,且与其它位置的情况相独立. 如果一个位置已被占用,我们就不能停在那儿,而如果我们停在一个空位上,就要招致一个"损失",它与该地点到目标0的距离成正比.

为简单起见,我们设可取的停车位置有一个下界$Q<0$. 设$y_Q, y_{Q+1}, \cdots$独立同分布,

$$P(y_n=1)=p=1-P(y_n=0),$$

对$n=Q, Q+1, \cdots$,令

$$\mathscr{F}_n = \mathscr{B}(y_Q, y_{Q+1}, \cdots, y_n),$$

$$\tilde{x}_n = \begin{cases} -\infty, & 若\ y_n=0, \\ -|n|, & 若\ y_n=1. \end{cases}$$

由于对一切n, $\tilde{x}_n \leqslant 0$,对任何停止变量t, $E\tilde{x}_t$存在是没有问题的. 但是随机序列$\{\tilde{x}_n, \mathscr{F}_n\}_Q^\infty$不满足$E|\tilde{x}_n|<\infty (n \geqslant Q)$的正式要求. 暂时不管这一困难,我们注意到在使$E\tilde{x}_t$最大时,可以限于考虑这样的规则$t$,它具有性质

$$t=t_1, \quad 若\ t>0,$$

其中t_1=第一个$n\geqslant 1$使得$y_n=1$. 事实上,对任一停止变量t,

$$t' = t I_{(t\leqslant 0)} + t_1 I_{(t>0)}$$

是一个停止变量且$\tilde{x}_{t'} \geqslant \tilde{x}_t$. 容易看到

$$E\tilde{x}_{t_1}=-\frac{1}{p}.$$

所以，若令

$$x_1=-\frac{1}{p},$$

$$x_n=\begin{cases} n, & 若\ y_n=1, \\ -\dfrac{1}{p}, & 若\ y_n=0, \end{cases} \quad Q<n<0,$$

我们并没有改变原来的问题，而可积随机序列$\{x_n, \mathscr{F}_n\}_0^1$确定的最优停止问题是符合一般的模型的.

(b) 下述例子引起了相当多的理论统计的研究工作. 我们首先给出一个直观的解答，以后在4.3节(2)及5.2节(a)中再给出严格的解答.

设(y_n)为独立同分布随机变量，密度为f，它是对直线上的Borel集上的某个σ-有限测度μ取的. 要求检验(简单)假设H_0: $f=f_0$，备择假设H_1: $f=f_1$，这里f_0及f_1是两个特定的密度. 由于H_1真而接受H_0造成的损失是某常数$b>0$，而由于H_0真而接受H_1造成的损失是$a>0$; 每次观察y_n的费用是一个单位. 一个序贯决策程序(δ, t)规定了样本容量t及最终决断δ; (δ, t)的损失是

$$a\alpha_0+E_0t, \quad 当\ H_0\ 是真的,$$

$$b\alpha_1+E_1t, \quad 当\ H_1\ 是真的,$$

其中$\alpha_0=P_0$(接受H_1)，$\alpha_1=P_1$(接受H_0).

若H_0为真的先验概率为π(因此H_1为真的概率为$1-\pi$)，(δ, t)的总"风险"为

$$r(\pi, \delta, t)=\pi[a\alpha_0+E_0t] +(1-\pi)[b\alpha_1+E_1t].$$

对一个给定的停止变量t，容易确定一个最终决断规则δ,

它使 $r(\pi, \delta, t)$ 在 a, b 及 π 固定时最小. 因为 $r(\pi, \delta, t)$ 依赖于 δ 的部分(省去诸如 $d\mu(y_1)\cdots d\mu(y_n)$ 的符号)

$$
\begin{aligned}
&\pi a\,\alpha_0+(1-\pi)b\,\alpha_1\\
&\quad=\pi a\sum_1^{\infty}\int_{(t=n,\ \text{接受}\ H_1)} f_0(y_1)\cdots f_0(y_n)\\
&\qquad+(1-\pi)b\sum_1^{\infty}\int_{(t=n,\ \text{接受}\ H_0)} f_1(y_1)\cdots f_1(y_n)\\
&\quad\geqslant\sum_1^{\infty}\int_{(t=n)}\min[\pi a f_0(y_1)\cdots f_0(y_n),(1-\pi)b f_1(y_1)\cdots f_1(y_n)]\\
&\quad=\sum_1^{\infty}\int_{(t=n)}\min[a\pi_n,\ b(1-\pi_n)][\pi f_0(y_1)\cdots f_0(y_n)\\
&\qquad\qquad+(1-\pi)f_1(y_1)\cdots f_1(y_n)],
\end{aligned}
$$

其中

$$
\begin{aligned}
\pi_n&=\pi_n(y_1,\ \cdots,\ y_n)\\
&=\frac{\pi f_0(y_1)\cdots f_0(y_n)}{\pi f_0(y_1)\cdots f_0(y_n)+(1-\pi)f_1(y_1)\cdots f_1(y_n)}.
\end{aligned}
$$

对一个给定的停止变量 t, 定义 δ' 为

$$
\begin{cases}
\text{接受 } H_1, & \text{若 } t=n \text{ 及 } \pi_n a\leqslant(1-\pi_n)b,\\
\text{接受 } H_0, & \text{若 } t=n \text{ 及 } \pi_n a>(1-\pi_n)b.
\end{cases}
$$

那么

$$
\begin{aligned}
&\pi a\,\alpha_0(\delta,\ t)+(1-\pi)b\,\alpha_1(\delta,\ t)\\
&\qquad\geqslant\pi a\,\alpha_0(\delta',\ t)+(1-\pi)b\,\alpha_1(\delta',\ t),
\end{aligned}
$$

所以对给定的 π, 找 (δ, t) 使 $r(\pi, \delta, t)$ 最小 ("Bayes" 决策程序)等于解下列问题: 给定 $0<\pi<1$, 设对每个 n, $y_1, y_2, \cdots, y_n$ 的联合密度函数为

$$
\pi f_0(y_1)\cdots f_0(y_n)+(1-\pi)f_1(y_1)\cdots f_1(y_n),
$$

其中 f_0 及 f_1 是给定的一元密度函数. 给定 a, $b>0$, 令

$$
\begin{cases}
h(\lambda)=\min(a\lambda,\ b(1-\lambda)), \quad 0\leqslant\lambda\leqslant 1,\\
\pi_0=\pi,\\
\pi_n=\dfrac{\pi f_0(y_1)\cdots f_0(y_n)}{\pi f_0(y_1)\cdots f_0(y_n)+(1-\pi)f_1(y_1)\cdots f_1(y_n)}, \\
\qquad\qquad n=1,\ 2,\ \cdots,\\
x_n=-h(\pi_n)-n, \quad n=0,\ 1,\ \cdots,\\
\mathscr{F}_0=\{\emptyset,\ \Omega\},\\
\mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n), \quad n=1,\ 2,\ \cdots.
\end{cases}
$$

我们要找停止变量 t 使 Ex_t 最大.（我们指出，这里我们允许对 y_n 不作任何观察而决定接受 H_0 或 H_1 且 $x_0=h(\pi)$；一个取值 0 的停止变量必须以概率 0 或 1 这样做，因为 $\mathscr{F}_0=\{\emptyset,\ \Omega\}$. 这个注记很容易归并到前面关于最优最终决断规则的规定中去.）若 $a\leqslant 1$ 或 $b\leqslant 1$，问题是显然的，因为这时 $h(\lambda)<1$，对一切正的 n，$x_0>x_n$，因此最优规则是 $t=0$. 所以我们假定 $a>1$，$b>1$.

下述不严格的讨论有相当大的直观的吸引力. 首先要问是否应作第一次观察. 计算

(3.3) $$V(\pi)=\inf_{\delta,t}[\pi(\alpha_0 a+E_0t)+(1-\pi)(\alpha_1 b+E_1 t)],$$

其中 inf 是对一切至少作一次观察的停止变量 t 及一切最终决断规则 δ 取的. 若我们至少作一次观察，$V(\pi)$ 表示最小平均损失. 若我们不作观察，损失为 $h(\pi)$；因此当且仅当 $h(\pi)>V(\pi)$ 时应作第一次观察. 现在对报酬序列 (x_n) 作两点评论：(i) x_n 只是通过 π_n 而依赖于 $y_1,\ \cdots,\ y_n$；(ii) 记

$$\pi_n=\frac{\pi_{n-1}f_0(y_n)}{\pi_{n-1}f_0(y_n)+(1-\pi_{n-1})f_1(y_n)}, \quad n=1,\ 2,\ \cdots,$$

我们看到 π_n 在下述意义下成为一个平稳 Markov 序列：对 $n=0,\ 1,\ \cdots$，π_{n+1} 关于 $\mathscr{F}_n$ 的条件分布是 π_n 的函数，而这函

数与n无关. 现在设我们作了首次观察. 如果我们停止了，损失为$h(\pi_1)+1$. 如果我们继续下去，那么除了(a)我们已为作首次观察的权利付了一个单位的费用及(b)先验概率现在是π_1这两点差别之外，将来的发展状况与刚开始时是完全一样的. 因此在至少再作一次观察的规则中最小的平均损失为$V(\pi_1)+1$，所以当且仅当$h(\pi_1)+1>V(\pi_1)+1$时作第二次观察. 按归纳法可重复讨论下去，所以最优停止变量的一个自然的候选者是$\sigma=\inf\{n:h(\pi_n)\leqslant V(\pi_n)\}$. 令$A=\{\pi:h(\pi)\leqslant V(\pi)\}$. 因为$V(\cdot)$是一族线性函数的inf，所以它是凹的；$V(0)=V(1)=1$. 设$A\neq[0,1]$，容易看出存在唯一确定的两个数$\pi'$及$\pi''$使$\pi'<\pi''$，$A=\{\pi:\pi\leqslant\pi'$或$\pi\geqslant\pi''\}$，而$(\delta',\sigma)$是一个Wald序贯概率比检验(参阅Wald，Wolfowitz[1]或Lehmann[1]，104—106页).

2. 有限情形. 后退归纳法

在某些情形，例如3.1节(f)的秘书问题，要观察的只是有限的N个随机变量. 即使在无限情形，去考察一下，用不超过一个定数N的停止规则可得到怎样的结果，也是饶有兴趣的. 所有这些情形在原则上都可用下面描述的后退归纳法提供完全的解答.

设$\{x_n,\mathscr{F}_n\}_1^\infty$为可积随机序列. 以$C^N$记使$t\leqslant N$的停止变量$t$全体，定义

$$V^N=\sup_{t\in C^N}Ex_t.$$

(我们指出，对一切$t\in C^N$，$E|x_t|\leqslant\sum_{n=1}^N E|x_n|<\infty$，因此$Ex_t$存在.) 按照前面的术语，称$s^N$在$C^N$中最优，若$s^N\in C^N$且$Ex_{s^N}=V^N$. 把计算$V^N$及$s^N$的问题按下面的方法放进一族

问题中去是方便的. 对每个 $n=1, 2, \cdots, N$, 以 C_n^N 记使 $n\leqslant t\leqslant N$ 的停止变量 t 全体, 并定义 $v_n^N=\sup\limits_{t\in C_n^N} Ex_t$ (注意 $C^N=C_1^N$ 及 $V^N=v_1^N$). 现在我们来指明如何找 C_n^N 中的最优停止变量. 若 $n=N$, 问题是不足道的, 因为 $t=N$ 是 C_N^N 中唯一的停止规则, 故 $v_N^N=Ex_N$. 对 $n=N-1$, 直观上很明显, 应该比较 x_{N-1} 与 $E(x_N|\mathscr{F}_{N-1})$, 并利用规则

$$t=\begin{cases}N-1, & 若\ x_{N-1}\geqslant E(x_N|\mathscr{F}_{N-1}),\\ N, & 若\ x_{N-1}<E(x_N|\mathscr{F}_{N-1}).\end{cases}$$

这个分析导致下列"动态规划"定理, 它正式地叙述了后退归纳法原则.

定理 3.2 设 N 是固定的正整数. 逐个定义 $\gamma_N^N, \gamma_{N-1}^N, \cdots, \gamma_1^N$ 如下:

$$\gamma_N^N=x_N,$$
$$\gamma_n^N=\max[x_n, E(\gamma_{n+1}^N|\mathscr{F}_n)], \quad n=N-1, \cdots, 1.$$

对每个 $n=1, 2, \cdots, N$, 令

$$s_n^N=第一个\ i\geqslant n\ 使得\ x_i=\gamma_i^N,$$

则 $s_n^N\in C_n^N$, 且

$$E(x_{s_n^N}|\mathscr{F}_n)=\gamma_n^N\geqslant E(x_t|\mathscr{F}_n), \quad t\in C_n^N,$$

所以

$$Ex_{s_n^N}=E\gamma_n^N\geqslant Ex_t, \quad t\in C_n^N \quad 及 \quad v_n^N=E\gamma_n^N.$$

证明 对 $n=N$, 定理是不足道的. 假设它对某个 $n=2, 3, \cdots, N$ 成立. 取任一 $t\in C_{n-1}^N$ 及任意 $A\in\mathscr{F}_{n-1}$, 令 $t'=\max(t, n)\in C_n^N$, 则

$$(3.4)\quad \int_A x_t=\int_{A(t=n-1)} x_{n-1}+\int_{A(t\geqslant n)} x_{t'}$$
$$=\int_{A(t=n-1)} x_{n-1}+\int_{A(t\geqslant n)} E[E(x_{t'}|\mathscr{F}_n)|\mathscr{F}_{n-1}]$$

$$\leqslant \int_{A(t=n-1)} x_{n-1} + \int_{A(t\geqslant n)} E(\gamma_n^N | \mathscr{F}_{n-1}) \leqslant \int_A \gamma_{n-1}^N,$$

所以 $E(x_t|\mathscr{F}_{n-1}) \leqslant \gamma_{n-1}^N$. 对 $t = s_{n-1}^N$, 在 $\{s_{n-1}^N \geqslant n\}$ 上有 $t' = s_n^N$, 故 (3.4) 中的不等式可代之以等式. 事实上, 由归纳法假设 $E(x_{s_N}|\mathscr{F}_n) = \gamma_n^N$, 由 s_{n-1}^N 之定义,

$$\int_A x_{s_{n-1}^N} = \int_{A(x_{n-1}\geqslant E(\gamma_{n-1}^N|\mathscr{F}_{n-1}))} x_{n-1} + \int_{A(x_{n-1}<E(\gamma_{n-1}^N|\mathscr{F}_{n-1}))} E(\gamma_n^N | \mathscr{F}_{n-1})$$

$$= \int_A \max(x_{n-1}, E(\gamma_n^N | \mathscr{F}_{n-1})) = \int_A \gamma_{n-1}^N,$$

所以 $E(x_{s_{n-1}^N}|\mathscr{F}_{n-1}) = \gamma_{n-1}^N$.

3. 一个应用

定理 3.2 中的 γ_n^N 极少能用定义它的后退归纳法解析地计算出来. 然而, 偶或也有这样的问题, 定理 3.2 多少可用来定出最优停止规则 s^N 的明显表达式.

设在例 3.1(f) 的问题中, 我们感兴趣的是使选中最好的姑娘的概率最大. 容易看出 $y_1, \cdots, y_N$ 是独立的且

$$(3.5) \qquad P\{y_n = j\} = \frac{1}{n}, \quad j = 1, 2, \cdots, n,$$

$$(3.6) \quad x_n = P\{a_n = 1 | \mathscr{F}_n\} = P\{a_n = 1 | y_n\} = \begin{cases} n/N, & \text{若 } y_n = 1, \\ 0, & \text{若 } y_n > 1. \end{cases}$$

利用 (3.5), (3.6) 及前面介绍的后退归纳法, 略去固定的上标 N, 令 $v_n = E\gamma_n$, 我们有

$$\begin{aligned} &\gamma_N = x_N, \\ &\gamma_{N-1} = \max(x_{N-1}, E(x_N)) = \max(x_{N-1}, v_N), \\ &\vdots \\ &\gamma_n = \max(x_n, E\gamma_{n+1}) = \max(x_n, v_{n+1}), \\ &\vdots \end{aligned}$$

及 s = 第一个 $n \geqslant 1$ 使得 $x_n \geqslant v_{n+1}$, 这里为统一起见, 令 $v_{N+1} = 0$.

因为 $v_1 \geqslant v_2 \geqslant \cdots \geqslant v_N = Ex_N = \frac{1}{N} > 0$，由(3.6)，$s$ 可描述如下：存在一个数 r，$1 \leqslant r \leqslant N$，使 s = 第一个 $n \geqslant r$ 使得 $y_n = 1$。直接计算表明(除了不足道的 $N=1$ 的情形)

$$(3.7) \qquad Ex_s = [(r-1)/N] \sum_{k=r}^{N} (k-1)^{-1}.$$

我们应用定理 3.2 限制了求最优规则的规则类，这时较方便的是不管通过后退归纳法计算 s 的实际可能性究竟怎样，而直接把(3.7)的右边看作 r 的函数 $\varphi(r)$，使它达到最大。容易看出

$$(3.8) \qquad \varphi(r) - \varphi(r+1) = N^{-1}\left(1 - \sum_{r}^{N-1} k^{-1}\right),$$

因此 $\varphi(r)$ 在

$$(3.9) \qquad \begin{aligned} r^* = r^*(N) &= \inf\{r:\ \varphi(r) - \varphi(r+1) \geqslant 0\} \\ &= \inf\left\{r:\ \frac{1}{r} + \frac{1}{r+1} + \cdots + \frac{1}{N-1} \leqslant 1\right\} \end{aligned}$$

取最大值。往证

$$(3.10) \qquad \lim_{N\to\infty} N^{-1} r^*(N) = \lim_{N\to\infty} v^N = e^{-1}.$$

事实上，从(3.8)及(3.9)得到，对充分大的 N

$$\sum_{r^*}^{N-1} k^{-1} \leqslant 1 < \sum_{r^*-1}^{N-1} k^{-1},$$

因此

$$(3.11) \qquad \int_{r^*}^{N-1} y^{-1} dy \leqslant 1 < \int_{r^*-2}^{N-1} y^{-1} dy.$$

$\lim_{N} N^{-1} r^*(N)$ 存在且等于 e^{-1} 这一事实是(3.11)的直接推论；在(3.7)的右边适当地取极限就得(3.10)的其余的结论。

4. 若干基本引理

在开始讨论最简单的无限情形，所谓单调情形之前，我们

建立一些基本的一般结果. 引理 3.3 是定理 2.2 的一个适用于最优停止问题的推广. 它在整个下文中起中心作用. 事实上, 定理 3.2 可由它容易地推得(见习题 8).

给定可积随机序列 $\{x_n, \mathscr{F}_n\}_1^\infty$. 以 C 记使 $Ex_t^- < \infty$ 的停止变量 t 全体; 由定义, $V = \sup_C Ex_t$.

引理 3.2 若 $s, t \in C$, 对一切 $n \geqslant 1$,

$$(3.12) \qquad E(x_s | \mathscr{F}_n) \geqslant x_n, \quad 在 \{s>n\} 上,$$

$$(3.13) \qquad E(x_t | \mathscr{F}_n) \leqslant x_n, \quad 在 \{s=n,\ t\geqslant n\} 上,$$

则 $Ex_s \geqslant Ex_t$. 作为部分的逆命题, 若 $Ex_s = V < \infty$, 则对每个 $t \in C$, (3.12) 及 (3.13) 成立, $n=1, 2, \cdots$.

证明

$$\begin{aligned}
Ex_s &= \int_{\{s<t\}} x_s + \int_{\{s\geqslant t\}} x_s \\
&= \sum_1^\infty \int_{\{s=n<t\}} x_n + \sum_1^\infty \int_{\{t=n\leqslant s\}} E(x_s | \mathscr{F}_n) \\
&\geqslant \sum_1^\infty \int_{\{s=n<t\}} E(x_t | \mathscr{F}_n) + \sum_1^\infty \int_{\{t=n\leqslant s\}} x_n = Ex_t.
\end{aligned}$$

现在设 $Ex_s = V < \infty$. 为了证明 (3.12), 设 n 为任意的,

$$A = \{s>n,\ E(x_s | \mathscr{F}_n) < x_n\} \in \mathscr{F}_n,$$

则 $t' = nI_A + sI_{\bar{A}} \in C$, 若 $P(A)>0$, 则

$$Ex_{t'} = \int_A x_n + \int_{\bar{A}} x_s > \int_A x_s + \int_{\bar{A}} x_s = Ex_s,$$

得出矛盾. 类似地, 若 $t\in C$, 令 $B = \{s=n,\ t>n,\ E(x_t | \mathscr{F}_n) > x_n\}$, $t' = tI_B + sI_{\bar{B}}$, 我们可看出 (3.13) 成立.

引理 3.3 设 s 是任一停止变量, 使对 $n=1, 2, \cdots$,

$$(3.14) \qquad E(x_{n+1} | \mathscr{F}_n) \geqslant x_n, \quad 在 \{s>n\} 上,$$

则 $\{x_{\min(s,n)}, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是下鞅. 若更有 Ex_s 存在及

$$(3.15) \qquad \liminf_n \int_{(s>n)} x_n^+ = 0,$$

则(3.12)成立.

证明 令 $s(n)=\min(s, n)$. 显然 $Ex_{s(n)}^+ \leqslant \sum_1^n Ex_i^+ < \infty$ 及 $Ex_{s(n)}$ 存在, $n=1, 2, \cdots$. 对一切 $A \in \mathscr{F}_n$,

$$\int_A x_{s(n)} = \int_{A(s\leqslant n)} x_s + \int_{A(s>n)} x_n$$

$$\leqslant \int_{A(s\leqslant n)} x_s + \int_{A(s\geqslant n+1)} x_{n+1} = \int_A x_{s(n+1)}.$$

由此推出 $\{x_{s(n)}, \mathscr{F}_n, 1 \leqslant n < \infty\}$ 是下鞅.

现在设 Ex_s 存在, 则对任何 $A \in \mathscr{F}_n (n=1, 2, \cdots)$, 由下鞅性, 对每个 $m=n+1, n+2, \cdots$, 我们有

$$\int_{A(s>n)} x_n \leqslant \int_{A(s>n)} x_{s(m)} = \int_{A(n<s\leqslant m)} x_s + \int_{A(s>m)} x_m$$

$$\leqslant \int_{A(n<s\leqslant m)} x_s + \int_{A(s>m)} x_m^+.$$

设子列 (m') 使

$$\int_{(s>m')} x_{m'}^+ \longrightarrow \liminf_{m\to\infty} \int_{(s>m)} x_m^+ = 0,$$

沿子列 (m'), 令 $m\to\infty$ 我们得到

$$\int_{A(s>n)} x_n \leqslant \int_{A(s>n)} x_s,$$

由此得(3.12). 如设 $Ex_n^+ < \infty$ 而不是 $E|x_n| < \infty$, 证明仍然成立. 类似的讨论(见习题 1)可证明

引理 3.4 若 $s, t \in C$, 对每个 $n=1, 2, \cdots$,

$$(3.16) \qquad E(x_{n+1} | \mathscr{F}_n) \leqslant x_n, \quad 在 \{s \leqslant n\} 上,$$

$$(3.17) \qquad \liminf_n \int_{(t>n)} x_n^- = 0,$$

则(3.13)成立.

5. 单调情形

有一种情形,存在一个最优停止变量的自然的候选者.这就是单调情形,这时随机序列 $\{x_n, \mathscr{F}_n\}_1^\infty$ 满足某些条件.

令

$$A_n=\{E(x_{n+1}|\mathscr{F}_n)\leqslant x_n\},\quad n=1,2,\cdots.$$

我们说我们处于单调情形,若

$$(3.18)\qquad A_1\subset A_2\subset\cdots;\quad \bigcup_1^\infty A_n=\Omega.$$

当(3.18)成立时,从最优观点来看,停止变量

$$(3.19)\qquad s=\text{第一个 } n\geqslant 1 \text{ 使得 } x_n\geqslant E(x_{n+1}|\mathscr{F}_n).$$

值得特别注意,但在我们被引诱试图去证明,在单调情形 s 总是最优的之前,让我们举两个简单的例子说明有时这并非如此.(它们是引言中的例1及例3.)

(a) 设 $y_1, y_2, \cdots$ 独立同分布, $P(y_n=1)=\frac{1}{2}=P(y_n=-1)$, $\mathscr{F}_n=\mathscr{B}(y_1,\cdots,y_n)$. 令 $x_n=2n/(n+1)\cdot\prod_1^n(y_k+1)$,

则 $Ex_n=\frac{2n}{n+1}$, $E(x_{n+1}|\mathscr{F}_n)=[(n+1)^2/n(n+2)]x_n$, 因此

$E(x_{n+1}|\mathscr{F}_n)\leqslant x_n\Leftrightarrow x_n=0\Rightarrow x_{n+1}=0\Rightarrow E(x_{n+2}|\mathscr{F}_{n+1})\leqslant x_{n+1}.$

所以 $A_n\subset A_{n+1}$, 且 $P\left(\bigcup_1^\infty A_n\right)=P(\text{对某个 } n\geqslant 1, y_n=-1)=1$,

所以是单调情形. 这里(3.19)等于

$$s=\text{第一个 } n \text{ 使得 } y_n=-1,$$

而对一切 t, $0=Ex_s\leqslant Ex_t$, 因此 s 是最劣停止变量. 不难证明,在这情形 $V=2$ 且不存在最优停止变量.

(b) 设 (y_n), $(\mathscr{F}_n)$ 同上例,但令

$$x_n=\min(1,\ y_1+\cdots+y_n)-\frac{n}{n+1}.$$

易见 $E(x_{n+1}|\mathscr{F}_n)<x_n$, $n=1, 2, \cdots$, 故显然是单调情形，且由(3.19)定义之 s 恒等于 1, $Ex_s=-\frac{1}{2}$. 现在令 $t=$第一个 n 使得 $y_1+\cdots+y_n=1$, $P\{t<\infty\}=1$, 所以 t 是停止变量，由于 $0<x_t\leqslant\frac{1}{2}$, $0<Ex_t\leqslant\frac{1}{2}$. 因此 s 不是最优的，易见 t 是最优的.

下面是单调情形的基本的正面结果.

定理 3.3 在单调情形，假设由(3.19)定义的 s 属于 C, 那么若(3.15)成立，则对一切使(3.17)成立的 $t\in C$, $Ex_s\geqslant Ex_t$.

证明 定理直接由引理 3.2, 3.3 及 3.4 得到.

推论 设在单调情形 $s\in C$, (3.15)成立且存在(i)一个期望有限的随机变量 $w\geqslant 0$, (ii)递增的正常数列(c_n)使得

$$x_n^-\leqslant w+c_n\quad(n=1,\ 2,\ \cdots). \tag{3.20}$$

则对每一个使 $Ec_t<\infty$ 的停止变量 t, $Ex_s\geqslant Ex_t>-\infty$.

证明 由(3.20), 若 $Ec_t<\infty$, $Ex_t^-\leqslant Ew+Ec_t<\infty$, 故 $t\in C$; 又

$$\int_{\{t>n\}}x_n^-\leqslant\int_{\{t>n\}}w+\int_{\{t>n\}}c_t\to 0\quad(n\to\infty).$$

故(3.17)成立，从而定理 3.3 可用来证明推论.

6. 应用

(a) 设 $y, y_1, y_2, \cdots$ 独立同分布, $E|y|<\infty$, 令

$$\mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n),$$

$$m_n=\max(y_1,\ \cdots,\ y_n),\ x_n=m_n-c_n\quad(n=1,\ 2,\ \cdots),$$

其中(c_n)是任一严格递增的正常数列，则

$$x_{n+1}-x_n=(y_{n+1}-m_n)^+-b_n,$$

其中记 $b_n=c_{n+1}-c_n(n=1, 2, \cdots)$. 这样,

$$E(x_{n+1}|\mathscr{F}_n)\leqslant x_n \Leftrightarrow E((y_{n+1}-m_n)^+|\mathscr{F}_n)\leqslant b_n.$$

定义 β_n 为方程

$$E(y-\beta_n)^+=b_n \tag{3.21}$$

的(唯一)解,则

$$E(x_{n+1}|\mathscr{F}_n)\leqslant x_n \Leftrightarrow m_n\geqslant \beta_n. \tag{3.22}$$

由于 $m_n\leqslant m_{n+1}$, 若 $\beta_n\geqslant\beta_{n+1}$, 即 $b_{n+1}\geqslant b_n$, 则(3.18)的第一部分成立.

下面的讨论将设(b_n)是递增的. 由(3.22), 定义(3.19)变成

s=第一个 $n\geqslant1$ 使得 $m_n\geqslant\beta_n$($=\infty$, 若不存在这样的 n).

由于 $E(y_1-\beta_1)^+>0$, 可得 $p\equiv P(y_1<\beta_1)<1$. 因此 $n\to\infty$ 时,

$$\begin{aligned}P(s>n)&\leqslant P(y_1<\beta_1, \cdots, y_n<\beta_n)\\&\leqslant P(y_1<\beta_1, \cdots, y_n<\beta_1)=p^n\to0,\end{aligned}$$

故 $P\{s<\infty\}=1$, 所以是单调情形. 又

$$Es^m\leqslant\sum_1^\infty k^mP(s\geqslant k)\leqslant\sum_1^\infty k^mp^{k-1}<\infty \quad (m=1, 2, \cdots),$$

所以 s 具有有限的一切阶矩, 由引理 3.1 (或定理 2.3),

$$Ex_s^+\leqslant E(y_1^++\cdots+y_s^+)=Ey^+Es<\infty.$$

由于

$$\int_{\{s>n\}}x_n^+\leqslant\beta_1^+P(s>n)\to0 \quad (n\to\infty),$$

故(3.15)成立. 取 $w=y_1^-$, 条件(3.20)成立. 由定理 3.3 及其推论得出, 对一切使 $Ec_t<\infty$ 的 $t\in C$, $Ex_s\geqslant Ex_t$.

为了再进一步证明 s 是 C 中最优的, 只要证明对每个停止变量 $t\in C$, $Ec_t<\infty$. 我们不能直接达到这个目的, 而要费番周折. 考虑特殊情形 $c_n=c\cdot n$. 这时 $b_n=c$, $\beta_n=\beta$, 其中 β 是方程

$$E(y-\beta)^+=c$$

的唯一解，$s=$ 第一个 $n\geqslant 1$ 使得 $y_n\geqslant\beta$. 从而有

$$Ex_s=\sum_1^\infty\left[P(s\geqslant n)\int_{(y>\beta)}y-cP(s\geqslant n)\right]$$

$$=\frac{1}{P(y\geqslant\beta)}\left[\int_{(y>\beta)}y-c\right]=\beta.$$

设 $\beta'>\beta$, 定义 $S_n=\sum_1^n[(y_k-\beta')^+-c]$, 则 $x_n\leqslant\beta'+S_n$, 故对任意停止变量 t, $x_t\leqslant\beta'+S_t$. 若 $Ex_t^-<\infty$, 则 ES_t 存在且由引理 3.1,

$$ES_t=(Et)(E(y-\beta')^+-c)<0.$$

从而 $Et<\infty$ 及 $Ex_t\leqslant\beta$. 所以这时 $V=\beta$, s 是 C 中最优的.

现在设(c_n)是任意的(满足$0<c_{n+1}-c_n\uparrow$),且对某个 $t\in C$,

(3.23) $$Ec_t=\infty,$$

因为 $x_n\leqslant m_n-\frac{1}{2}c_n\leqslant m_n-\frac{1}{2}nc_1(n\geqslant 1)$, 从 $c_n=c\cdot n$ 的特殊情形可得

$$-\infty<Ex_t\leqslant E\left(m_t-\frac{1}{2}c_t\right)\leqslant E\left(m_t-\frac{1}{2}tc_1\right)<\infty.$$

故由 (3.23), $Ex_t=E\left(m_t-\frac{1}{2}c_t\right)-\frac{1}{2}Ec_t=-\infty$. 这矛盾证明了,对每个 $t\in C$, $Ec_t<\infty$, 所以 s 是 C 中最优的.

注 值得指出,上面的讨论实际上只用引理 3.1, 不用定理 3.3 证明了,在 $c_n=c\cdot n$ 的重要特殊情形, s 是 C中最优的.

(b) 假设与上例同且 $c_n=c\cdot n$, 取 $x_n'=y_n-c\cdot n$,则 $x_n'\leqslant x_n$, 但 $x_s=x_s'$, 从而 s 对 $\{x_n', \mathscr{F}_n\}_1^\infty$ 也是最优的. 注意到对于这个报酬序列,我们不是处于单调情形.

(c) 设 y_1, y_2, $\cdots$ 独立,在$\left[\theta-\frac{1}{2},\theta+\frac{1}{2}\right]$上均匀分布,其中 θ 是要估计的未知参数. 假设若用 θ^*估计 θ,损失是$(\theta^*-\theta)^2$,

并设每观察一个 y_n, $n=1, 2, \cdots$, 付出的费用是常数 c. 最后，设未知的 θ 有先验分布: $[0, 1]$ 上均匀分布.

易见 θ 在已知 y_1 的条件下的后验分布是 $[u_1, v_1]$ 上的均匀分布，其中 $[u_1, v_1]$ 是 $[0, 1]$ 与 $\left[y_1-\frac{1}{2}, y_1+\frac{1}{2}\right]$ 的公共部分，从而可得 θ 在已知 $y_1,\cdots,y_n$ 的条件下的后验分布是 $[u_n, v_n]$ 上的均匀分布，其中 $[u_n, v_n]$ 是 $[u_{n-1}, v_{n-1}]$ 与 $\left[y_n-\frac{1}{2}, y_n+\frac{1}{2}\right]$ 的公共部分. 记 $A_n=v_n-u_n$, $n=1, 2, \cdots$.

上述问题的决策规则等于一个停止规则连同一个最终决断规则(见例 3.1(h)). 按与例 3.1(h) 类似的推论易见存在一个一致最优的最终决断规则，即这规则与所用的停止变量无关，它包括取样在第 n 步停止时，用 $(u_n+v_n)/2$ 估计 θ, 后验平均损失为 $A_n^2/12$. 继续仿照例 3.1(h) 的讨论，我们的问题可归结为下述问题.

设 (y_n), (u_n), (v_n), (A_n), θ 同上. 令

$$\mathscr{F}_0=\{\emptyset, \Omega\}, \quad x_0=-A/12,$$

对 $n=1, 2, \cdots$, 令

$$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n), \quad x_n=-A_n^2/12-c\cdot n,$$

我们要找 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的最优停止变量. 直接计算可证

$$E(A_{n+1}^2 | A_n=A)=A^2(1-A/2),$$

易见这是单调情形，可用定理 3.2 给出最优规则.

(d) 给定直线上的一个分布 F, 且 $\int_{-\infty}^{\infty} x\,dF(x)=0$, 随机变量 $y_1, \cdots, y_N$ (N 为定数)按下述法则产生: y_1 按 F 分布，在时刻 $2, 3, \cdots, N$, 试验者可选作新的独立观察(按 F 分布)或重复以前的观察值中的最大值. 要求这样产生的随机变量的和有最大可能的期望. 显然，试验者只要考虑具有下述性质

的策略：如果在某时刻 n 重复一个观察值，那么在时刻 $n+1$ $(n=2, 3, \cdots, N-1)$ 也重复那个观察值．事实上，若一个策略在值 $y_1, \cdots, y_{n_0}$ 之后在时刻 $n_0+1, \cdots, n_1$ 重复，而在 n_1+1 作新的观察，那么它必劣于另一个策略，该策略在时刻 n_0+1 作新的观察，然后在时刻 $n_0+2, \cdots, n_1+1$ 重复（它们之前的最大观察值可能更大），而在其它方面与原来的策略相同．原来的问题可归结为 $\{x_n, \mathscr{F}_n\}_1^N$ 的最优停止问题，其中

$$x_n = y_1+\cdots+y_n+(N-n)\max(y_1, \cdots, y_n),$$

$$\mathscr{F}_n = \mathscr{B}(y_1, \cdots, y_n).$$

（若 n 是最后一次作新的观察的时刻，$n=1, 2, \cdots, N$，则 x_n 是得到的报酬．）

由直接计算可证 $E(x_{n+1}-x_n|\mathscr{F}_n) = -m_n+(N-n-1)\cdot E((y_{n+1}-m_n)^+|\mathscr{F}_n)$，其中记 $m_n=\max(y_1, \cdots, y_n)$．所以我们处于单调情形，且 $s=$ 第一个 $n\geqslant 1$ 使得 $m_n \geqslant \alpha_{N-n}$，其中 $\alpha_0 = -\infty$，对 $k=1, 2, \cdots$，α_k 满足方程

$$\alpha_k = (k-1)\int_{\alpha_k}^{\infty}(y-\alpha_k)dF(y). \tag{3.24}$$

易见 (3.24) 有唯一解，$0=\alpha_1<\alpha_2<\cdots$．

看作有限情形，易见 (3.19) 定义的停止变量 s 事实上就是定理 3.2 的 s^N（见习题 3.3）．

习　题

1. 证明引理 3.4．（提示：直接证明或对上鞅 $\{y_n \equiv E(x_{\max(s,n)}|\mathscr{F}_n), \mathscr{F}_n\}_1^\infty$ 应用引理 3.3.）
2. 证明若 $y_k \geqslant 0$，例 3.1(e) 是单调情形，可应用定理 3.2．s 是什么？现在去掉 y_k 为非负的要求，但（假设 $Ey_k>0$）在一个子类中找最优规则，这子类中的规则具有下列性质：存在一个正数 a，使 $t=$ 第一个 n 使得 $y_1+\cdots+y_n \geqslant a$．于是这问题可归结为非负的情形．（直

观上)很清楚,如存在最优规则,它一定包含在这子类之中.(见习题5.1.)

3. 证明:在单调情形 $s^N=\min[N,$ 第一个 n 使得 $x_n\geqslant E(x_{n+1}|\mathscr{F}_n)]$.

4. 证明:s^N 是 C^N 中的最小最优规则,即若 $t\in C^N$ 及 $Ex_t=V^N$,则 $t\geqslant s^N$.

5. 设在单调情形,并令

$$C^-=\{t: t\in C, \liminf_{n\to\infty}\int_{\{t>n\}}x_n^-=0\},$$

$$V^-=\sup_{C^-}Ex_t,$$

证明:若 $V^-<\infty$ 且存在 C^- 中的最优规则,则 $s\in C^-$ 且 $Ex_s=V^-$.(提示:先证若 $t\in C^-$,则 $t'=\min(s,t)\in C^-$ 及 $Ex_{t'}\geqslant Ex_t$. 然后证明:若 $t\in C^-$ 及 $P\{t=n<s\}>0$,t 在 C^- 中不可能最优.)易见,$V^-\geqslant\lim_{N\to\infty}V^N$. 是否可能严格不等?

6. 对停车问题(例 3.1(g)),证明我们能限于考虑下列形式的规则:

$$t_r=\text{第一个 } k\geqslant r \text{ 使得 } y_k=1 \quad (r\geqslant Q).$$

计算 Ex_{t_r},把它看作 r 的函数求最大值.

7. 在秘书问题中假设我们感兴趣的是使选中最好的姑娘的概率最大,但不再设姑娘们按随机的次序出现,而设姑娘们以概率 $\frac{1}{N}$ 按 $(N, N-1, \cdots, 1)$ 的第 n 个 $(n=0, 1, \cdots, N-1)$ 循环排列的次序出现. V 是什么? s^N 是什么? 如果标准是使选中的姑娘的平均名次最小,姑娘们以什么样的分布出现会有类似的效果?(提示:第一个出现的姑娘应以 $\frac{1}{2}$ 的概率是最好的姑娘,以 $\frac{1}{2}$ 的概率是最差的姑娘.)

8. 应用 3.4 节的基本引理于 $\{\gamma_n^N, \mathscr{F}_n\}_1^N$,以证明定理 3.2.

9. 对 3.3 节的秘书问题,直接用后退归纳法计算最优规则.

10. 在秘书问题中,如果标准是使选中最好的两个姑娘之一的概率为最大,求最优规则. 当 $N\to\infty$ 时最大概率的极限值是什么?

第 四 章

一 般 理 论

在涉及 $x_1, \cdots, x_N$ 的有限情形，我们定义了

$$(4.1)\qquad \begin{aligned}&\gamma_N^N = x_N,\\ &\gamma_n^N = \max(x_n,\ E(\gamma_{n+1}^N \mid \mathscr{F}_n)) \quad (n = N-1, \cdots, 1).\end{aligned}$$

定理 3.2 证明了

$$(4.2)\qquad \gamma_n^N = \operatorname*{ess\,sup}_{t\in C_n^N} E(x_t \mid \mathscr{F}_n) \quad (n = 1, 2, \cdots, N)$$

(关于一族随机变量的本性上确界的定义与基本性质见 1.6 节)及

$$(4.3)\qquad s_n^N = \text{第一个 } i \geqslant n \text{ 使得 } x_i = \gamma_i^N$$

在 C_n^N 中最优. 在一般的最优停止问题中涉及一个无限的序列 $x_1, x_2, \cdots$，不存在一个 x_N 可作为出发点定义序列(4.1)，例 3.1(a) 也证明最优规则不一定存在. 然而我们仍能类似于(4.2)，对 $n=1, 2, \cdots$，定义

$$(4.2')\qquad \gamma_n = \operatorname*{ess\,sup}_{C_n} E(x_t \mid \mathscr{F}_n) \quad (n = 1, 2, \cdots),$$

其中 C_n 是使 $t \geqslant n$ 的 $t \in C$ 全体. 然后就可能证明与(4.1)类似有

$$(4.1')\qquad \gamma_n = \max(x_n,\ E(\gamma_{n+1} \mid \mathscr{F}_n)) \quad (n = 1, 2, \cdots),$$

并寻找使

(4.3′) $\sigma_n =$ 第一个 $i \geqslant n$ 使得 $x_i = \gamma_i$ ($=\infty$，若不存在这样的 i)在 C_n 中最优的条件(从而 $\sigma = \sigma_1$ 在 $C = C_1$ 中最优). 最后,

因为(4.2')定义 γ_n 是非构造性的，我们将证明，具有构造性解的有限问题当 $N\to\infty$ 时在某种意义下逼近一般问题．例如，我们将寻找使 $\lim\limits_{N\to\infty}\gamma_n^N=\gamma_n\ (n=1,\ 2,\ \cdots)$ 的条件．

1．定义及引理

基本定义如下．以 C 记使 $Ex_t^-<\infty$ 的停止变量 t 全体，

$$C_n=\{\max(t,\ n):t\in C\},\ v_n=\sup_{C_n}Ex_t,$$

$$\gamma_n=\operatorname*{ess\,sup}_{C_n}E(x_t|\mathscr{F}_n)\quad(n=1,\ 2,\ \cdots).$$

看来，似乎更自然的是代替 C_n，考虑更大的停止变量类 C_n^*——使 $t\geqslant n$ 及 Ex_t 存在的 t 全体．然而这将产生同样的 v_n 及 γ_n．因为若 $t\in C_n^*$，定义

$$t'=\begin{cases}t, & \text{若 } E(x_t|\mathscr{F}_n)\geqslant x_n,\\ n, & \text{若 } E(x_t|\mathscr{F}_n)<x_n.\end{cases}$$

令 $A=\{E(x_t|\mathscr{F}_n)\geqslant x_n\}$，我们有 $Ex_{t'}^-\leqslant Ex_n^-+\int_A x_t^-$．但是 $-\infty<\int_A x_n\leqslant\int_A x_t$，故 $\int_A x_t^-<\infty$，由此得到 $t'\in C_n$．现在 $E(x_{t'}|\mathscr{F}_n)\geqslant E(x_t|\mathscr{F}_n)$ 及 $Ex_{t'}\geqslant Ex_t$，由此可得，若在 v_n 及 γ_n 的定义中用 C_n^* 代替 C_n，它们保持不变．

下列引理将用来证明 4.2 节中的基本的一般定理．

引理 4.1 对每个 $n=1,\ 2,\ \cdots$，存在 C_n 中的一列递增的序列 (t_k) 使得

$$x_n\leqslant E(x_{t_k}|\mathscr{F}_n)\uparrow\gamma_n\quad(k\to\infty).\tag{4.4}$$

证明 按定理 1.5 选取 C_n 中的 (t_k)，使得 $t_1=n$ 及 $\gamma_n=\sup\limits_k E(x_{t_k}|\mathscr{F}_n)$．由下面的引理 4.2 及 4.3 能肯定(4.4)成立．

引理 4.2 对任意的 $n=1,\ 2,\ \cdots$ 及 $t\in C_n$，定义

$t'=$第一个 $k\geqslant n$ 使得 $E(x_t|\mathscr{F}_k)\leqslant x_k$,
则

(a) $t'\leqslant t$, $\quad t'\in C_n$,

(b) $E(x_{t'}|\mathscr{F}_n)\geqslant E(x_t|\mathscr{F}_n)$,

(c) $t'>j\geqslant n\Rightarrow E(x_{t'}|\mathscr{F}_j)>x_j$.

证明 若 $t=j\geqslant n$, 则 $E(x_t|\mathscr{F}_j)=x_j$ 及 $t'\leqslant j$; 因此 $t'\leqslant t$. 现在

$$Ex_{t'}^-=\sum_{k=n}^{\infty}\int_{\{t'=k\}}x_k^-\leqslant\sum_{k=n}^{\infty}\int_{\{t'=k\}}[E(x_t|\mathscr{F}_k)]^-$$

$$\leqslant\sum_{k=n}^{\infty}\int_{\{t'=k\}}E(x_t^-|\mathscr{F}_k)=Ex_t^-<\infty,$$

故 $t'\in C_n$. 因此(a)成立. 对 $j\geqslant n$ 及任意的 $A\in\mathscr{F}_j$,

$$\int_{A\{t'>j\}}x_{t'}=\sum_{j}^{\infty}\int_{A\{t'=k\}}x_k\geqslant\sum_{j}^{\infty}\int_{A\{t'=k\}}E(x_t|\mathscr{F}_k)$$

$$=\int_{A\{t'>j\}}x_t.$$

令 $j=n$ 就得出(b). 在 $\{t'>j\}$ 上我们有

$$E(x_{t'}|\mathscr{F}_j)\geqslant E(x_t|\mathscr{F}_j)>x_j.$$

这给出了(c).

每一个满足上述引理中(c)的 $t\in C_n$ 称为 n-可取的. 每一个 1-可取的停止变量称为可取的.

引理 4.3 设对某固定的 $n\geqslant1$, t_1, $t_2\in C_n$ 是 n-可取的, 令 $\tau=\max(t_1, t_2)$, 则 $\tau\in C_n$, 是 n-可取的且

$$E(x_\tau|\mathscr{F}_n)\geqslant\max_{i=1,2}E(x_{t_i}|\mathscr{F}_n).$$

证明 $\tau\in C_n$ 是明显的. 对任意的 $j\geqslant n$ 及 $A\in\mathscr{F}_j$,

$$\int_{A\{t_1>j\}}x_\tau=\sum_{j}^{\infty}\left[\int_{A\{t_1=k<\tau\}}x_{t_2}+\int_{A\{t_1=k=\tau\}}x_k\right]$$

$$\geqslant\sum_{j}^{\infty}\left[\int_{A\{t_1=k<\tau\}}x_k+\int_{A\{t_1=k=\tau\}}x_k\right]=\int_{A\{t_1>j\}}x_{t_1}.$$

对 $j=n$, 这就给出了 $E(x_\tau|\mathscr{F}_n)\geqslant E(x_{t_1}|\mathscr{F}_n)$, 再由对称性 $E(x_\tau|\mathscr{F}_n)\geqslant\max\limits_{i=1,2}E(x_{t_i}|\mathscr{F}_n)$. 为了证明 τ 是 n-可取的, 我们注意到 $\{\tau>j\}=\{t_1>j\}\cup\{t_2>j\}$, 和在 $\{t_i>j\}$ 上有 $E(x_\tau|\mathscr{F}_j)\geqslant E(x_{t_i}|\mathscr{F}_j)>x_j\quad(i=1,2)$.

引理 4.4 若 $t\in C$, 则对任意的 $n=1,2,\cdots$,

$$t\geqslant n\Rightarrow E(x_t|\mathscr{F}_n)\leqslant\gamma_n\quad\text{和}\quad E(x_t^-|\mathscr{F}_n)\geqslant\gamma_n^-.$$

证明 令 $t'=\max(t,n)$, 则 $t'\in C_n$, 由此

$$E(x_{t'}|\mathscr{F}_n)\leqslant\gamma_n,\qquad E(x_{t'}^-|\mathscr{F}_n)\geqslant[E(x_{t'}|\mathscr{F}_n)]^-\geqslant\gamma_n^-.$$

但在 $\{t\geqslant n\}$ 上 $t'=t$, 故得引理.

对每个 $n=1,2,\cdots$, 定义随机变量 $\sigma_n=$第一个 $k\geqslant n$ 使得 $x_k=\gamma_k(=\infty$, 若不存在这样的 $k)$, 并记 $\sigma_1=\sigma$. 一般来讲, $P(\sigma_n<\infty)<1$, 所以 σ_n 不是停止变量.

引理 4.5 若 $t\in C_n$, 则 $t'=\min(t,\sigma_n)\in C_n$ 及 $E(x_{t'}|\mathscr{F}_n)\geqslant E(x_t|\mathscr{F}_n)$.

证明 记 $\sigma_n=\sigma$, 用引理 4.4, 对任何 $A\in\mathscr{F}_n$,

$$(4.5)\qquad \int_{A(\sigma<t)}x_\sigma^-=\sum_{k=n}^{\infty}\int_{A(\sigma=k<t)}x_k^-=\sum_{k=n}^{\infty}\int_{A(\sigma=k<t)}\gamma_k^-$$
$$\leqslant\sum_{k=n}^{\infty}\int_{A(\sigma=k<t)}x_t^-=\int_{A(\sigma<t)}x_t^-.$$

因此 $Ex_{t'}^-=\int_{(t\leqslant\sigma)}x_t^-+\int_{(\sigma<t)}x_\sigma^-\leqslant Ex_t^-<\infty$. 现在, 在(4.5)中用 $(\cdot)$ 代替 $(\cdot)^-$, 并把不等式反过来, 我们有

$$\int_A x_{t'}=\int_{A(t\leqslant\sigma)}x_t+\int_{A(\sigma<t)}x_\sigma\geqslant\int_A x_t.$$

引理 4.6 设 $V<\infty$, (t_k) 为 C 中一列递增的停止变量使得 $k\to\infty$ 时, $Ex_{t_k}\uparrow V$, 则 $P(\lim t_k\geqslant\sigma)=1$.

证明 对任一 $n=1,2,\cdots$, 令 (s_k) 如引理 4.1 中所定义, 则由单调收敛定理, $V\geqslant Ex_{s_k}=E(E(x_{s_k}|\mathscr{F}_n))\uparrow E\gamma_n$, 由此

得到 $E\gamma_n<\infty\ (n=1, 2, \cdots)$. 令 $t=\lim t_k$ 并设对某个 i, $A=\{t=i<\sigma\}$ 有正概率. 存在一个 $\varepsilon>0$, 使得

$$\int_A \gamma_i-3\varepsilon\geqslant\int_A x_i.$$

对每个 $k\geqslant 1$, 令 $B_k=\{t_k=i<\sigma\}$; 则 $I_{B_k}\to I_A$, 由控制收敛定理, 存在一个 k_0 使得对一切 $k\geqslant k_0$,

$$\int_{B_k} x_i\leqslant\int_{B_k}\gamma_i-2\varepsilon. \tag{4.6}$$

由引理 4.1, 对每个 $k\geqslant k_0$, 存在一个 $t_k'\in C_i$ 使得

$$\int_{B_k} x_{t_k'}\geqslant\int_{B_k}\gamma_i-\varepsilon. \tag{4.6'}$$

令 $\tau_k=t_k'I_{B_k}+t_kI_{\bar B_k}$, 易见 $\tau_k\in C$, 由 (4.6) 及 (4.6′), 对一切 $k\geqslant k_0$ 我们有

$$Ex_{\tau_k}=\int_{B_k}x_{t_k'}+\int_{\bar B_k}x_{t_k}\geqslant\int_{B_k}x_i+\int_{\bar B_k}x_{t_k}+\varepsilon=Ex_{t_k}+\varepsilon.$$

因此 $\sup\limits_k Ex_{\tau_k}\geqslant V+\varepsilon$, 得到矛盾.

2. 若干一般定理

现在我们能证明(4.1′)总是成立的.

定理 4.1

(a) $\gamma_n=\max(x_n, E(\gamma_{n+1}|\mathscr{F}_n))$,

(b) $v_n=E\gamma_n\quad (n=1, 2, \cdots)$.

证明 设 $n=1, 2, \cdots$ 及 $t\in C_n$ 是任意的, 令 $B=\{t=n\}$. 由引理 4.4, 在 $\bar B$ 上 $E(x_t|\mathscr{F}_{n+1})\leqslant\gamma_{n+1}$, 故在 $\bar B$ 上 $E(x_t|\mathscr{F}_n)\leqslant E(\gamma_{n+1}|\mathscr{F}_n)$. 这样,

$$\begin{aligned}E(x_t|\mathscr{F}_n)&=I_Bx_n+I_{\bar B}E(x_t|\mathscr{F}_n)\\&\leqslant\max(x_n, E(\gamma_{n+1}|\mathscr{F}_n)),\end{aligned}$$

所以

$$\gamma_n \leqslant \max(x_n,\ E(\gamma_{n+1}|\mathscr{F}_n)).$$

为完成证明，设(t_k)如引理 4.1 中所定义. 由条件期望的单调收敛定理，当 $k\to\infty$ 时，

$$\gamma_n \geqslant E(x_{t_k}|\mathscr{F}_n) = E[E(x_{t_k}|\mathscr{F}_{n+1})|\mathscr{F}_n] \uparrow E(\gamma_{n+1}|\mathscr{F}_n).$$

因为不等式 $\gamma_n \geqslant x_n$ 显然是满足的，故定理得证.

定理 4.2 (a) 若 $\sigma\in C$ 且 σ 是可取的，则它是最优的. (b) 若 $V<\infty$ 且存在最优停止变量，则 $\sigma\in C$，σ 是最优及可取的，并且

(4.7) $$\sigma \geqslant n \Rightarrow E(x_\sigma|\mathscr{F}_n) = \gamma_n \quad (n=1,\ 2,\ \cdots).$$

证明 (a) 由引理 4.4，条件(3.13)在 $s=\sigma$ 时满足，由假设条件(3.12)满足. 引理 3.2 即完成了证明.

(b) 引理 4.5 及 4.6 包含了 σ 的最优性. 为了证明 σ 是可取的，只要证(4.7). 对任一 $n=1,\ 2,\ \cdots$，令 $A=\{\sigma>n,\ E(x_\sigma|\mathscr{F}_n)<\gamma_n\}$. 若 $P(A)>0$，则 $\int_A x_\sigma < \int_A \gamma_n$，因为 $E\gamma_n \leqslant V<\infty$. 由引理 4.1，存在 $t\in C_n$ 使得 $\int_A x_t > \int_A x_\sigma$. 令 $\tau = tI_A + \sigma I_{\bar{A}}$. 易见 $\tau\in C$，$Ex_\tau > Ex_\sigma$，得出矛盾.

定理 4.2 或许是一般理论的主要结果. 然而随机变量 γ_n 及常数 v_n 一般不可能直接由 4.1 节开头给出的定义去计算. 在许多情形用截尾法可提供一个构造性的方法去得到 γ_n 及 v_n. 为应用这个方法，对任一 $N\geqslant 1$ 及 $n=1,\ \cdots,\ N$，我们定义

$$C_n^N = \{\min(t,\ N): t\in C_n\}, \qquad v_n^N = \sup_{C_n^N} Ex_t,$$

$$\gamma_n^N = \operatorname*{ess\,sup}_{C_n^N} E(x_t|\mathscr{F}_n),$$

则有

$$-\infty < Ex_n = v_n^n \leqslant v_n^{n+1} \leqslant \cdots \leqslant v_n,$$

$$x_n=\gamma_n^n\leqslant\gamma_n^{n+1}\leqslant\cdots\leqslant\gamma_n,$$

所以我们能定义

(4.8) $$v_n'=\lim_{N\to\infty}V_n^N,\qquad \gamma_n'=\lim_{N\to\infty}\gamma_n^N,$$

并有

$$-\infty<Ex_n\leqslant v_n'\leqslant v_n,\qquad x_n\leqslant\gamma_n'\leqslant\gamma_n.$$

把定理 4.1 的论证应用于有限序列 $\{x_n,\ \mathscr{F}_n\}_1^N$，我们有

$$\gamma_N^N=x_N,$$

$$\gamma_n^N=\max(x_n,\ E(\gamma_{n+1}^N|\mathscr{F}_n))\quad(n=N-1,\ \cdots,\ 1),$$

及 $E\gamma_n^N=v_n^N$，所以(γ_n^N)及(v_n^N)可递归地计算，我们现在定义的(γ_n^N)及(v_n^N)与第三章中的相一致．由单调收敛定理，

$$E\gamma_n'=v_n',$$

(4.9) $$\gamma_n'=\max(x_n,\ E(\gamma_{n+1}'|\mathscr{F}_n))\quad(n=1,\ 2,\ \cdots).$$

因此 (γ_n') 与 (γ_n) 满足相同的递归关系，但一般没有 $(\gamma_n')=(\gamma_n)$．(例如，设 $y_1,\ y_2,\ \cdots$ 独立同分布，$P\{y_1=1\}=\frac{1}{2}=P\{y_1=-1\}$．令 $x_n=\sum_1^n y_k$，$\mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n)$ $(n\geqslant1)$．由于 $\{x_n,\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$ 是鞅，由定理 2.2，对每个有界停止变量 t，在 $\{t\geqslant n\}$ 上，$E(x_t|\mathscr{F}_n)=x_n$，因此 $\gamma_n'=x_n(n\geqslant1)$．但 $P\{\limsup x_n=+\infty\}=1$，故对一切 n，$\gamma_n=+\infty$．)然而

定理 4.3 对任一 $n=1,\ 2,\ \cdots$，若对每个 $t\in C_n$，

(4.10) $$\liminf_{N\to\infty}\int_{\{t>N\}}(\gamma_N')^-=0,$$

则 $\gamma_n'=\gamma_n$ 及 $v_n'=v_n$．特别，$v_1'=V$．

证明 方程 (4.9) 蕴含了 $\{\gamma_n',\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$ 是一个上鞅．由引理 3.3(或定理 2.2)，对每个 $t\in C_n$，

$$\gamma_n'\geqslant E(\gamma_t'\mid\mathscr{F}_n).$$

因为 $\gamma'_k \geqslant x_k (k=1, 2, \cdots)$，故知对每个 $t \in C_n$，

$$\gamma'_n \geqslant E(x_t | \mathscr{F}_n),$$

从而

$$\gamma'_n \geqslant \gamma_n.$$

由于反过来的不等式 $\gamma'_n \leqslant \gamma_n$ 一般总是成立的，定理得证。

注　由引理 4.5，若只假设(4.10)对使 $P\{t \leqslant \sigma_n\}=1$ 的那些 $t \in C_n$ 成立，定理 4.3 仍然正确．这点注意将在 5.7 节中用到．

自然希望叙述一些容易验证的加在序列 $\{x_n, \mathscr{F}_n\}_1^\infty$ 上的条件，使得(4.10)从而定理 4.3 的结论成立．为此只要证明

$$\int_{\{t>n\}} x_n^- \to 0 \quad (t \in C),$$

由此也只要给出非负随机变量序列 $(z_{1,n}), \cdots, (z_{K,n})$，使得

$$x_n^- \leqslant \sum_{k=1}^{K} z_{k,n} \quad (n \geqslant 1),$$

$$\lim_{n\to\infty} \int_{\{t>n\}} z_{k,n} = 0 \quad (k=1, \cdots, K;\ t \in C).$$

在应用中有用的两类 (z_n) 序列是

(z_n) 是一致可积的，

(z_n) 递增且对一切 $t \in C$，$Ez_t < \infty$．

下面的定理及其推论就是由这样的考虑和所谓"统计模型"促成的，在该模型中 $-x_n$ 是两个非负 $\mathscr{F}_n$-可测随机变量之和，其中一个是作出错误的最终决断所引起的损失，另一个是取样费用，取样费用随观察次数的增加而趋于 $+\infty$（参阅 3.1 节(h)及 3.6 节(c)）．

定理 4.4　设 $x_n = x'_n - x''_n$，其中 x'_n 与 x''_n 为 $\mathscr{F}_n$-可测 $(n \geqslant 1)$，并且

(a) $((x'_n)^-)$ 一致可积；

(b) 存在非负递增随机序列 $\{z_n, \mathscr{F}_n\}_1^\infty$，使得 $x''_n \leqslant z_n (n \geqslant 1)$，且对一切 $t \in C, Ez_t < \infty$，

则对一切 $n=1, 2, \cdots, \gamma'_n = \gamma_n$.

证明 由定理 4.3，只要证对每个 $t \in C$，

$$\int_{\{t>n\}} x_n^- \to 0 \quad (n \to \infty).$$

但 $x_n^- \leqslant (x'_n)^- + (x''_n)^+ \leqslant (x'_n)^- + z_n$，因此由 (a)，(b) 及定理 1.3，

$$\int_{\{t>n\}} x_n^- \leqslant \int_{\{t>n\}} (x'_n)^- + \int_{\{t>n\}} z_n$$
$$\leqslant \int_{\{t>n\}} (x'_n)^- + \int_{\{t>n\}} z_t \to 0 \quad (n \to \infty).$$

推论 若 $x_n = \hat{x}_n - \hat{\hat{x}}_n = x_n^* - x_n^{**}$，其中各项都为 $\mathscr{F}_n$-可测 $(n \geqslant 1)$，若

(a) $((x_n^*)^-)$ 一致可积，

(b) $E(\sup \hat{x}_n^+) < \infty$，

(c) $0 \leqslant \hat{\hat{x}}_1 \leqslant \hat{\hat{x}}_2 \leqslant \cdots$，

(d) 对某个 $c > 0$，$x_n^{**} \leqslant c\hat{\hat{x}}_n \quad (n \geqslant 1)$，

则对一切 $n=1, 2, \cdots, \gamma'_n = \gamma_n$.

证明 令 $x'_n = x_n^*$，$x''_n = x_n^{**}$，$z_n = c\hat{\hat{x}}_n$. 由 (b) 对一切 $t \in C$，$Ez_t < \infty$，推论立即可得.

注 若我们不再坚持要 $z_n \uparrow$，但假设 $\{z_n, \mathscr{F}_n, 1 \leqslant n < \infty\}$ 是 (4.4 节中定义的) C-正则下鞅，定理 4.4 依然正确. 推论的条件 (c) 可类似地减弱.

统计模型也启发了保证 σ 实际上是最优的充分条件.

定理 4.5 设 $E(\sup x_n^+) < \infty$. 若 $P(\sigma < \infty) = 1$，则 σ 是最优的. 为了 $P(\sigma < \infty) = 1$，只要 $\lim x_n = -\infty$.

证明 选 t_i 为可取及递增的，使得对一切 i，$P(t_i \leqslant \sigma) = 1$

及 $Ex_{t_i}\uparrow V$（由引理 4.2, 4.3 及 4.5, 这总是能做到的）. 由引理 4.6, $\sigma=\lim t_i$. 由 Fatou 引理, V 等于

$$\lim_{i\to\infty} Ex_{t_i}\leqslant E\left(\limsup_{i\to\infty} x_{t_i}\right)$$

$$\leqslant\int_{\{\sigma<\infty\}} x_\sigma+\int_{\{\sigma=\infty\}}\left(\limsup_{n\to\infty} x_n\right).$$

若 $P(\sigma<\infty)=1$, 则 $Ex_\sigma\geqslant V$, σ 是最优的. 若 $x_n\to-\infty$, 则必须有 $P(\sigma<\infty)=1$.

我们指出, 若在上面推论的条件上再加 $\hat{x}_n\to\infty$, 则由定理 4.5, σ 是最优的.

3. 应　用

(a) 定理 4.5 可立即应用于例 3.1(h) 及 3.6(c) 的问题以证明 σ 是最优的. 在每个情形不难应用定理 4.4 在原则上提供一个计算 σ 的方法.

(b) 假设我们处于单调情形, 那么

$$E(x_{n+2}|\mathscr{F}_{n+1})\leqslant x_{n+1},\quad \text{在 } \{E(x_{n+1}|\mathscr{F}_n)\leqslant x_n\} \text{ 上}$$

(见 3.5 节). 由习题 3.3,

$$s^N=\min(N, \text{第一个 } n \text{ 使得 } x_n\geqslant E(x_{n+1}|\mathscr{F}_n)),$$

因此由 (3.19) 定义的 s 是 $\lim\limits_{N\to\infty} s^N$. 引理 4.6 (也见习题 4.3) 说, 为了 $s=\sigma$, 只要 $V^N\to V<\infty$; 定理 4.4 提供了验证 $V^N\to V$ 的一个方法; 定理 4.5 告诉我们 σ 为最优的条件. 现在我们把这些工具应用于 3.6 节(a), (b)的问题(不是所有这些问题都是单调情形).

我们要用到下列引理, 它的证明要推迟一点.

引理 4.7　设 $w, w_1, w_2, \cdots$ 为同分布非负随机变量, 对任一 $\alpha>0$, 令

$$z=\sup_n(\max(w_1, \cdots, w_n)-n^\alpha),$$

则

(i) $Ew^{\alpha^{-1}}<\infty$ 蕴含 $P\{z<\infty\}=1$;

(ii) 对任一 $\beta>0$, 若 $Ew^{\alpha^{-1}+\beta}<\infty$, 则 $E(z^+)^\beta<\infty$.

现在设 $y, y_1, y_2, \cdots$ 同分布, α 为一正常数, 令

$$x_n=\max(y_1, \cdots, y_n)-n^\alpha, \quad \tilde{x}_n=y_n-n^\alpha$$
$$(n=1, 2, \cdots).$$

若 $E|y|<\infty$ 与 $E(y^+)^{1+\alpha^{-1}}<\infty$, 则存在停止变量 σ 与 $\tilde{\sigma}$ 使得 $V=Ex_\sigma$, $\tilde{V}=E\tilde{x}_{\tilde{\sigma}}$ 和 $\sigma=\lim s^N$, $\tilde{\sigma}=\lim \tilde{s}^N$. 若 y_n 又是独立的, 则对 $\alpha\geqslant 1$ 我们有

$$\sigma=\text{第一个 } n \text{ 使得 } \max(y_1, \cdots, y_n)\geqslant\beta_n,$$

其中 β_n 由 $E[(y-\beta_n)^+]=(n+1)^\alpha-n^\alpha$ 定义.

证明 由于引理 4.7 蕴含着

$$E\left(\sup_n\left[\max(y_1, \cdots, y_n)-\frac{n^\alpha}{2}\right]\right)<\infty,$$

令 $\hat{x}_n=\max(y_1, \cdots, y_n)-(1/2)n^\alpha$, $x_n^*=\max(y_1, \cdots, y_n)$, 我们马上看到, 定理 4.4 的推论可应用于 $\{x_n, \mathscr{F}_n\}_1^\infty$, 因此 $\sigma=\lim\limits_{N\to\infty} s^N$. 由定理 4.5, σ 是 C 中最优的. 类似的论证可应用于 $\{\tilde{x}_n, \mathscr{F}_n\}_1^\infty$. 若 y_n 是独立的且 $\alpha\geqslant 1$, 则对 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的最优停止问题是单调情形, 且 σ 是由 (3.19) 定义的停止变量.

引理 4.7 的证明 易见 $z=\sup\limits_n(w_n-n^\alpha)$. 因此由 Borel-Cantelli 引理(2.4 节(d)),

$$\text{(i)}\quad Ew^{\alpha^{-1}}<\infty \Rightarrow \sum_1^\infty P\{2w>n^\alpha\}<\infty$$

$$\Rightarrow \sum_1^\infty P\{2w_n>n^\alpha\}<\infty$$

$$\Rightarrow P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\left\{w_k>\frac{k^\alpha}{2}\right\}\right)=0$$

$$\Rightarrow P\{z<\infty,\ w_n-n^\alpha\to-\infty\}=1.$$

(ii) 假设 $Ew^{\alpha^{-1}+\beta}<\infty$. 为了证明 $E(z^+)^\beta<\infty$, 只要证明

$$\int_0^\infty u^{\beta-1}P\{z>u\}du<\infty.$$

现在对任一 $u>0$,

$$\begin{aligned}P\{z>u\}&\leqslant\sum_{n=1}^{\infty}P\{w>u+n^\alpha\}\\&=\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}P\{k^\alpha<w-u\leqslant(k+1)^\alpha\}\\&=\sum_{n=1}^{\infty}kP\{k^\alpha<w-u\leqslant(k+1)^\alpha\}\\&\leqslant E(w-u)^+)^{\alpha^{-1}}.\end{aligned}$$

令 $F(x)=P\{w\leqslant x\}$ 及 $\gamma=\alpha^{-1}$, 我们有

$$\begin{aligned}E((w-u)^+)^{\alpha^{-1}}&=\int_u^\infty(x-u)^\gamma dF(x)\\&=\gamma\int_u^\infty(x-u)^{\gamma-1}(1-F(x))dx.\end{aligned}$$

因此

$$\begin{aligned}\int_0^\infty u^{\beta-1}P\{z>u\}du&\leqslant\gamma\int_0^\infty u^{\beta-1}\int_u^\infty(x-u)^{\gamma-1}(1-F(x))dxdu\\&=\gamma\int_0^\infty\left(\int_0^x u^{\beta-1}(x-u)^{\gamma-1}du\right)(1-F(x))dx\\&\leqslant\text{const}\int_0^\infty x^{\beta+\gamma-1}(1-F(x))dx<\infty.\end{aligned}$$

(c) 假设我们在观察一个生产过程的产量，它们构成独立同分布随机变量 $y_1, y_2, \cdots$. 生产过程在某个随机时刻 θ 发生变化. 问题是探测这个变化的发生.

为了使问题简化与定形，我们假设 θ 是非负整值随机变

量，使得

$$P\{\theta=0\}=\pi,$$

$$P\{\theta=k\,|\,\theta>0\}=r_k,\quad k=1,\ 2,\ \cdots,\ \left(\sum_1^\infty r_k=1\right).$$

我们观察 $y_1,\ \cdots,\ y_{\theta-1},\ y'_\theta,\ y'_{\theta+1},\ \cdots$，其中 $y_1,\ y_2,\ \cdots(y'_1,\ y'_2,\ \cdots)$ 独立同分布，服从已知分布 $F_0(F_1)$。若我们在时刻 n 停止过程，我们损失

$$\begin{aligned}&c(\text{一个固定的检查费用}), &&\text{若 } \theta>n,\\ &n-\theta, &&\text{若 } \theta\leqslant n.\end{aligned}$$

我们要使平均损失最小.

令 $\mathscr{F}_n$ 是我们观察的前 n 个随机变量产生的 σ-代数，但如同例 3.1(e) 一样，上面所说的在第 n 步停止的损失不是 $\mathscr{F}_n$-可测的．为弥补这点，我们令

$$x_n=-c(1-\pi_n)-\sum_{i=0}^{n-1}(n-i)p_i^n\quad(n=0,\ 1,\ 2,\ \cdots),$$

其中我们令

$$p_i^n=P(\theta=i\,|\,\mathscr{F}_n)\quad(i,\ n\geqslant 0),$$

$$\pi_n=P(\theta\leqslant n\,|\,\mathscr{F}_n)\quad(n\geqslant 0).$$

现在我们证明定理 4.3 及 4.5 的条件是满足的，从而 σ 是最优的，它可用后退归纳法及取极限来计算．显然，$E(\sup x_n^+)<\infty$. 令 $\mathscr{F}_\infty=\mathscr{B}(\bigcup_1^\infty \mathscr{F}_n)$，由定理 1.4，对每个 $n\geqslant N=1,2,\ \cdots$，当 $n\to\infty$ 时，

$$\sum_0^n(n-i)p_i^n\geqslant(n-N)P\{\theta\leqslant N\,|\,\mathscr{F}_n\}\to\infty,$$

$$\text{在 } \{P(\theta\leqslant N\,|\,\mathscr{F}_\infty)>0\} \text{ 上}.$$

但 $\bigcup_N\{P(\theta\leqslant N\,|\,\mathscr{F}_\infty)>0\}=\Omega$，因此 $\sum_{i=0}^n(n-i)p_i^n$ 以概率 $1\to 0$ $(n\to\infty)$．由定理 4.5，σ 是 C 中最优的．由于 $c(1-\pi_n)\leqslant c$

$(n\geqslant 1)$，为了证明 $(\gamma'_n)=(\gamma_n)$，由定理 4.3，只要证明 $\lim\limits_{n\to\infty}\int_{(t>n)}\left(\sum\limits_{i=0}^{n}(n-i)p_i^n\right)=0\ (t\in C)$. 然而，由于 $\{p_i^n, \mathscr{F}_n\}_i^\infty$ 是有界鞅，由引理 3.8 (定理 2.2)，对任一 $t\in C$，我们有

$$\int_{(t>n)}\sum_{i=0}^{n}(n-i)p_i^n=\sum_{i=0}^{n}(n-i)\int_{(t>n)}p_i^t$$
$$\leqslant\int_{(t>n)}\sum_{i=0}^{t}(t-i)p_i^t\to 0\quad(n\to\infty).$$

(d) 定理 4.4 中的思想是：为证明 $V^N\to V$，只要证明对每个 $t\in C$，$\liminf\limits_{N\to\infty}E\,x_{\min(t,N)}\geqslant Ex_t$. 下述例子的内容说明这一点决不是必要的，在这个例子中 $V=\lim\limits_{N\to\infty}V^N=1$，$\sigma$ 是最优的，但对每个 N，$Ex_{\min(\sigma,N)}=0$.

设 $x_1, x_2, \cdots$ 独立且

$$P\left(x_n=\begin{cases}1\\0\\-2n\end{cases}\right)=\begin{cases}\dfrac{1}{n+1},\\[2mm]\dfrac{n}{2(n+1)},\\[2mm]\dfrac{n}{2(n+1)},\end{cases}\qquad \mathscr{F}_n=\mathscr{B}(x_1,\cdots,x_n).$$

很清楚，σ(＝第一个 n 使得 $x_n=1$)是最优的，且 $V=1$. 又对每个 $N\geqslant 1$，

$$Ex_{\min(\sigma,N)}=P(\sigma\leqslant N)-NP(\sigma>N)$$
$$=1-\frac{1}{N+1}-\frac{N}{N+1}=0.$$

现在给定 $n\geqslant 1$，定义

$$t=\begin{cases}\text{第一个 } 1\leqslant k\leqslant n \text{ 使得 } x_k=1,\\ \text{第一个 } n+1\leqslant k\leqslant n^3-1 \text{ 使得 } x_k\neq -2k,\\ \qquad\qquad\text{若不存在前行所述的 } k,\\ n^3,\quad \text{若前述的那些 } k \text{ 都不存在},\end{cases}$$

则 $t\leqslant n^2$，故 $Ex_t\leqslant V^{n^2}$. 然而

$$P\{x_t=1\}\geqslant P\{\sigma\leqslant n\}=\frac{n}{n+1},$$

$$P\{x_t=-2n^2\}=\frac{1}{n+1}\cdot\frac{n+1}{2(n+2)}\cdot\frac{n+2}{2(n+3)}\cdots\frac{n^2}{2(n^2+1)}=\frac{1}{2^{n^2-n}(n^2+1)},$$

故

$$Ex_t\geqslant\frac{n}{n+1}-\frac{2n^2}{2^{n^2-n}(n^2+1)}\to 1,\quad n\to\infty.$$

4. (γ_n) 与 (γ'_n) 的鞅特征

定义 上鞅 $\{y_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 称为关于一个停止变量的类 D 为正则的，若对每个 $t\in D$, Ey_t 存在且在 $\{t\geqslant n\}$ 上，$E(y_t|\mathscr{F}_n)\leqslant y_n$ $(n\geqslant 1)$.

在这一节中我们简扼地指出，系统地考察这个概念将如何提供另一个得到本章许多结果的方法. 例如，若 $\{\beta_n, \mathscr{F}_n\}_1^\infty$ 是 C-正则上鞅使得 $\beta_n\geqslant x_n(n=1, 2, \cdots)$，则对一切 $t\in C_n$，$\beta_n\geqslant E(\beta_t|\mathscr{F}_n)\geqslant E(x_t|\mathscr{F}_n)$，从而 $\beta_n\geqslant\gamma_n(n=1, 2, \cdots)$. 由 (4.9)，$\{\gamma'_n, \mathscr{F}_n\}_1^\infty$ 是上鞅，定理 4.4 可解释为在某种条件下建立 $\{\gamma'_n, \mathscr{F}_n\}_1^\infty$ 的 C-正则性.

定义 若 $Y=\{y_n, \mathscr{F}_n\}_1^\infty$ 与 $Z=\{z_n, \mathscr{F}_n\}_1^\infty$ 是 $(\Omega, \mathscr{F}, P)$ 上的两个随机序列，我们说 Y 控制 Z，若对每个 $n=1, 2, \cdots$, $y_n\geqslant z_n$. 若 Y 与 Z 是随机序列，存在一个随机序列的类 $\mathscr{K}$ 使得

(a) $Y\in\mathscr{K}$,

(b) Y 控制 Z,

(c) 每一个控制 Z 的 $Y'\in\mathscr{K}$ 也控制 Y,

那么我们说 Y 是控制 Z 的类 $\mathscr{K}$ 的最小元.

定理 4.6 (a) $\{\gamma'_n, \mathscr{F}_n\}_1^\infty$ 是控制 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的最小上鞅. (b) $\{\gamma_n, \mathscr{F}_n\}_1^\infty$ 是控制 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的最小 C-正则上鞅.

证明 (a) 由(4.9)很明显 $\{\gamma'_n, \mathscr{F}_n\}_1^\infty$ 是控制 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的上鞅. 若 $\{\beta_n, \mathscr{F}_n\}_1^\infty$ 是任何另一个这样的上鞅, 则

$$\beta_N \geqslant x_N = \gamma_N^N \quad (N=1,\ 2,\ \cdots),$$

由后退归纳法,

$$\beta_n \geqslant \gamma_n^N \quad (N=1,\ 2,\ \cdots;\ n=1,\ 2,\ \cdots,\ N).$$

令 $N\to\infty$, 我们有

$$\beta_n \geqslant \gamma'_n \quad (n=1,\ 2,\ \cdots).$$

(b) 只要建立 $\{\gamma_n, \mathscr{F}_n\}_1^\infty$ 的 C-正则性; 由本节开头的注记, 最小性是明显的. 设 $E(\sup x_n^+)<\infty$ (这条件将用引理 4.12 去掉). 若 $\{\gamma_n, \mathscr{F}_n\}_1^\infty$ 不是 C-正则的, 那么存在一个整数 n, 一个 $\varepsilon>0$ 及一个 $A\in\mathscr{F}_n$, 使得对某个 $t\in C$,

$$\int_{A(t\geqslant n)}\gamma_n+\varepsilon\leqslant\int_{A(t\geqslant n)}\gamma_t.$$

由引理 4.1, 对每个 $k=n,\ n+1,\ \cdots$, 存在一个停止变量 $t_k\in C_k$, 使得

$$\int_{A(t=k)}\gamma_k-\frac{\varepsilon}{2^{k+1}}\leqslant\int_{A(t=k)}x_{t_k}.$$

令 $t^*=\sum_n^\infty t_k I_{A(t=k)}+nI_{\Omega-[A(t\geqslant n)]}$, 我们看到 t^* 是一个停止变量, 且

$$\int_{A(t\geqslant n)}E(x_{t^*}|\mathscr{F}_n)=\sum_{k=n}^\infty\int_{A(t=k)}x_{t_k}\geqslant\sum_{k=n}^\infty\int_{A(t=k)}\gamma_k-\frac{\varepsilon}{2}$$

$$=\int_{A(t\geqslant n)}\gamma_t-\frac{\varepsilon}{2}\geqslant\int_{A(t\geqslant n)}\gamma_n+\frac{\varepsilon}{2}.$$

这样, $t^*\in C_n$ 及 $P\{E(x_{t^*}|\mathscr{F}_n)>\gamma_n\}>0$, 得出矛盾.

作为应用定理 4.6 的一个例子，我们给出引理 4.6 的下述另一证明．设 k_0 使得(4.6)对一切 $k\geqslant k_0$ 成立．因为

$$\int_{B_k} x_{t_k}\leqslant\int_{B_k}\gamma_{t_k},$$

由 $\{\gamma_n,\ \mathscr{F}_n\}_1^\infty$ 的 C-正则性我们有

$$Ex_{t_k}\leqslant E\gamma_{t_k}-2\varepsilon\leqslant E\gamma_1-2\varepsilon=V-2\varepsilon,$$

与引理的假设矛盾．

5. 广义停止变量．三重极限定理

尽管"在实践中发生"的某些问题可能有充分好的性状，允许通过计算 γ_n^N 及取极限的方法计算 γ_n，然而(γ_n)在我们的理论中起到中心的作用，使得找一个总是正确的计算技巧是值得向往的．本节提供这样的技巧；尽管人们会发现用这些技巧去进行计算并不合意，但从这些结果可推导出的序列(γ_n)的定性的性质常常是非常有用的．

在本节下文中(除定理 4.9 外)将使用下列记号．设 $-\infty\leqslant a<0<b\leqslant\infty$，

$$x_n(a,\ b)=\begin{cases} b, & 若\ x_n>b,\\ x_n, & 若\ a\leqslant x_n\leqslant b,\\ a, & 若\ x_n<a,\end{cases}$$

$$x_n(b)=x_n(-\infty,\ b),$$

$$x_n(a)=x_n(a,\ +\infty),$$

以 $\gamma_n(a,b)$ ($\gamma_n^N(b)$ 等等)记与 $\{x_n(a,b),\mathscr{F}_n\}_1^\infty$ ($\{x_n(b),\ \mathscr{F}_n\}_1^N$ 等等)相联系的 γ_n (γ_n^N 等等)．

因为由定理 4.3，对一切 $a<0$，$\lim\limits_{N\to\infty}\gamma_n^N(a)=\gamma_n(a)$，$\lim\limits_{N\to\infty}v_n^N(a)=v_n(a)\ (n=1,\ 2,\ \cdots)$，看来自然地要问是否

(4.11) $\lim_{a\to-\infty}\gamma_n(a)=\gamma_n,\ \lim_{a\to-\infty}v_n(a)=v_n\quad(n=1,\ 2,\ \cdots).$

然而，不作进一步的假设，(4.11)不会成立(见 4.6 节(a))。马虎地讲，其原因是可以存在规则 t_k 使得

$$\sup_k Ex_{t_k}^+=\sup_k Ex_{t_k}^-=\infty\quad 但\ Ex_{t_k}=Ex_{t_k}^+-Ex_{t_k}^-$$

仍然有界。这样就会发生 $V<\infty$ 但对一切 $a>-\infty$，$V(a)=\infty$ 的情形。为了避免这个困难，我们在假设 $E(\sup_n x_n^+)<\infty$ 之下去证(4.11)。然后我们证明

$$\lim_{b\to\infty}\gamma_n(b)=\gamma_n.$$

这与(4.11)相结合就产生三重极限定理(定理 4.8(b))：

$$\lim_{b\to\infty}\lim_{a\to-\infty}\lim_{N\to\infty}\gamma_n^N(a,\ b)=\gamma_n;$$

$$\lim_b\lim_a\lim_N v_n^N(a,\ b)=v_n\quad(n=1,\ 2,\ \cdots).$$

在研究过程中引进广义停止变量的概念是方便的，即去掉 t 以概率 1 有限的要求，同时定义 $x_\infty=\limsup_{n\to\infty}x_n$。较正式地，若 $\{y_n,\ \mathscr{F}_n\}_1^\infty$ 是任一随机序列，称 t 为广义停止变量，若 t 是在 $\{1,\ 2,\ \cdots,\ +\infty\}$ 中取值的随机变量使得 $\{t=n\}\in\mathscr{F}_n$ $(n=1,\ 2,\ \cdots)$。$\{t=\infty\}$是指 $\bigcap_{n=1}^\infty\{t>n\}$。令 $y_\infty=\limsup_{n\to\infty}y_n$ 并约定在 $\{t=\infty\}$ 上 $y_t=y_\infty$。回到基本的可积随机序列 $\{x_n,\mathscr{F}_n\}_1^\infty$，以 $\bar{C}$ 记使 $Ex_t^-<\infty$ 的广义停止变量全体，$\bar{C}_n$ 为使 $t\geqslant n$ 的 $t\in\bar{C}$ 全体；$\bar{\gamma}_n=\operatorname*{ess\,sup}_{t\in\bar{C}_n}E\ (x_t|\mathscr{F}_n)\quad(n=1,\ 2,\ \cdots)$。注意，按我们的定义，$\bar{\gamma}_\infty=\limsup_{n\to\infty}\bar{\gamma}_n$。一般 $\bar{\gamma}_\infty=x_\infty$ 并不成立(参阅引理 4.10)。

下面的讨论会证实，广义停止变量的概念是一个有用的技巧，它启发证明的方法，在通常的框架下统一了结果。4.6

节(也见5.2节(c))包含一个例子，该例表明这样的规则本身也会是有趣的. 引理3.3的下述翻版相应于广义停止变量的研究.

引理4.8 设$\{y_n, \mathscr{F}_n\}_1^\infty$是可积随机序列，$t$是广义停止变量，使得

$$E(y_{n+1}|\mathscr{F}_n)\geqslant y_n, \quad 在\ \{t>n\}\ 上\ (n=1, 2, \cdots),$$

则$\{y_{\min(t,n)}, \mathscr{F}_n, 1\leqslant n<\infty\}$是一个下鞅. 若又有$Ey_t$存在及$(I_{(t>n)}y_n^+)$一致可积，则

$$(4.12)\quad E(y_t|\mathscr{F}_n)\geqslant y_n, \quad 在\ \{t\geqslant n\}上(n=1, 2, \cdots).$$

证明 第一部分如同引理3.3的证明那样可得. 为了证明(4.12)，注意到对任意的$A\in\mathscr{F}_n$, $N=n+1, n+2, \cdots$,

$$\int_{A(t>n)} y_n\leqslant\int_{A(t>n)} y_{\min(t,N)}=\int_{A(n<t\leqslant N)} y_t+\int_{A(t>N)} y_N.$$

令$N\to\infty$，由所设$(I_{(t>n)}y_n^+)$的一致可积性及Fatou引理(引理1.2)，我们有

$$\int_{A(t>n)} y_n\leqslant\int_{A(n<t<\infty)} y_t+\int_{A(t=\infty)} y_\infty=\int_{A(t>n)} y_t,$$

由此得(4.12).

引理4.9 若$\{\beta_n, \mathscr{F}_n\}_1^\infty$是一随机序列，使得对每个$n\geqslant 1$,

(a) 对某个期望有限的非负随机变量u,

$$x_n\leqslant\beta_n\leqslant E(u|\mathscr{F}_n),$$

(b) $\beta_n\leqslant\max(x_n, E(\beta_{n+1}|\mathscr{F}_n))$,

(c) $\beta_\infty=x_\infty$,

则$\beta_n\leqslant\gamma_n\quad(n=1, 2, \cdots)$.

证明 令$n=1, 2, \cdots$为任意正整数. 只要证明

$$\int_A\beta_n\leqslant\int_A\gamma_n\quad(A\in\mathscr{F}_n).$$

反之，设存在 $\varepsilon>0$, $A\in\mathscr{F}_n$, 使得

$$(4.13)\qquad \int_A\beta_n>\int_A\gamma_n+\varepsilon.$$

定义

$$t=\text{第一个 } k\geqslant n \text{ 使得 } x_k\geqslant\beta_k-\varepsilon,$$
$$=\infty,\ \text{若不存在这样的 } k,$$

从(a)及(b)可见引理 4.8 可应用于$\{\beta_n, \mathscr{F}_n\}_1^\infty$ 及 t，故由(c)，对任一 $B\in\mathscr{F}_n$, 我们有

$$(4.14)\qquad -\infty<\int_B\beta_n\leqslant\int_{B(t<\infty)}\beta_t+\int_{B(t=\infty)}x_\infty.$$

置 $B=\Omega$，可见 $P\{t=\infty, x_\infty=-\infty\}=0$. 但由(c)，$P\{t=\infty, x_\infty>-\infty\}=0$. 因此 $P\{t<\infty\}=1$ 及 $t\in C_n$; 从(4.13)，(4.14)(其中的 $B=A$)以及 t 的定义，我们有

$$\int_A\gamma_n<\int_A\beta_n-\varepsilon\leqslant\int_A x_t=\int_A E(x_t|\mathscr{F}_n),$$

与 γ_n 的定义矛盾.

引理 4.10 设 $E(\sup x_n^+)<\infty$, $\{\beta_n, \mathscr{F}_n\}_1^\infty$ 为一随机序列，使得对每个 $n=1, 2, \cdots$, $x_n\leqslant\beta_n\leqslant E(\sup\limits_{k\geqslant n}x_k|\mathscr{F}_n)$，则 $\beta_\infty=x_\infty$.

证明 很清楚 $\beta_\infty\geqslant x_\infty$. 又对每个 $m=1, 2, \cdots, n$,

$$\beta_n\leqslant E(\sup_{k\geqslant m}x_k|\mathscr{F}_n),$$

故由 Lévy 定理(定理 1.4)，

$$\beta_\infty\leqslant\sup_{k\geqslant m}x_k\to x_\infty,\ m\to\infty \text{ 时}.$$

对每个 $N=1, 2, \cdots$, 定义

$$(4.15)\qquad \begin{aligned}&\tilde{\gamma}_N^N=E(\sup_{k\geqslant N}x_k|\mathscr{F}_N),\\&\tilde{\gamma}_n^N=\max(x_n, E(\tilde{\gamma}_{n+1}^N|\mathscr{F}_n))\quad(n=N-1, \cdots, 1).\end{aligned}$$

由归纳法易见，对每个 n，$\tilde{\gamma}_n^n \geqslant \tilde{\gamma}_n^{n+1} \geqslant \cdots$，因此我们可定义 $\tilde{\gamma}_n = \lim\limits_{N\to\infty} \tilde{\gamma}_n^N$。

引理 4.11 若 $E(\sup x_n^+) < \infty$，则对每个 $n=1, 2, \cdots$，

(a) $\tilde{\gamma}_n = \bar{\gamma}_n = \gamma_n$，

(b) $\lim\limits_{a\to-\infty} \gamma_n(a) = \gamma_n$。

证明 (a) 易见

$$\tilde{\gamma}_n \geqslant \bar{\gamma}_n \geqslant \gamma_n \quad (n=1, 2, \cdots).$$

事实上，如同定理 4.1 的证明那样，我们有

$$(4.16) \qquad \bar{\gamma}_n \leqslant \max(x_n, E(\bar{\gamma}_{n+1} | \mathscr{F}_n)).$$

但对一切 N，$\bar{\gamma}_N \leqslant \tilde{\gamma}_N^N$，由 (4.15) 及 (4.16) 可见 $\bar{\gamma}_n \leqslant \tilde{\gamma}_n^N$ ($N=n, n+1, \cdots$; $n=1, 2, \cdots$)，故 $\bar{\gamma}_n \leqslant \tilde{\gamma}_n$ ($n=1, 2, \cdots$)。由 (4.15) 及条件期望的单调收敛定理，

$$\tilde{\gamma}_n = \max(x_n, E(\tilde{\gamma}_{n+1} | \mathscr{F}_n)) \quad (n=1, 2, \cdots).$$

再援用引理 4.9 及 4.10，我们就完成了 (a) 的证明。

(b) 对每个 n，$\gamma_n(a)$ 是 a 的递增函数；令 $\gamma_n^* = \lim\limits_{a\to-\infty} \gamma_n(a)$，则 $\gamma_n^* \geqslant \gamma_n$，且由条件期望的单调收敛定理，

$$\gamma_n^* = \max(x_n, E(\gamma_{n+1}^* | \mathscr{F}_n)) \quad (n=1, 2, \cdots).$$

从引理 4.10 可见对一切 $a > -\infty$，$x_\infty(a) = \gamma_\infty(a)$。但 $\gamma_\infty^* \leqslant \gamma_\infty(a) = x_\infty(a) = \max(x_\infty, a) \downarrow x_\infty$，当 $a \downarrow -\infty$ 时，再援用引理 4.9 即完成证明。

除了其它内容之外，上面这引理说，如果按所规定的方式放宽停止变量的定义，并假定 $E(\sup x_n^+ < \infty)$，那么 (γ_n)（从而 (V_n)）并不增加。没有这样一个限制，这结果仍成立则是下列引理的推论。

引理 4.12

$$\lim_{b\to\infty} \bar{\gamma}_n(b) = \bar{\gamma}_n \quad (n=1, 2, \cdots).$$

证明 令 $n=1, 2, \cdots$ 为任意正整数. $b\uparrow$ 时 $\bar{\gamma}_n(b)\uparrow$; 令 $\gamma_n^*=\lim\limits_{b\to\infty}\gamma_n(b)$, 则 $\gamma_n^*\leqslant\bar{\gamma}_n$. 对每个 $b>0$, $t\in\bar{C}_n$,

$$E(x_t(b)\,|\,\mathscr{F}_n)\leqslant\bar{\gamma}_n(b)\leqslant\gamma_n^*.$$

但是因为对每个 $t\in\bar{C}_n$, $x_t^-(b)=x_t^-$ 及 $Ex_t^-<\infty$, 由条件期望的单调收敛定理我们有

$$E(x_t\,|\,\mathscr{F}_n)=\lim_{b\to\infty}E(x_t(b)\,|\,\mathscr{F}_n)\leqslant\gamma_n^*\quad(t\in\bar{C}_n).$$

从而 $\bar{\gamma}_n\leqslant\gamma_n^*$.

注 引理 4.12 使得完成定理 4.6 的证明很容易.

我们把上面的结果综述为

定理 4.7 对每个 $n=1, 2, \cdots$,

$$\gamma_n=\operatorname*{ess\,sup}_{t\in\bar{C}_n}E(x_t\,|\,\mathscr{F}_n),\qquad v_n=\sup_{\bar{C}_n}Ex_t.$$

定理 4.8 对每个 $n=1, 2, \cdots$,

(a) $\gamma_n=\lim\limits_{b\to\infty}\lim\limits_{N\to\infty}\tilde{\gamma}_n^N(b)$, $\quad v_n=\lim\limits_{b}\lim\limits_{N}E\tilde{\gamma}_n^N(b)$,

(b) 三重极限定理:

$$\gamma_n=\lim_{b\to\infty}\lim_{a\to-\infty}\lim_{N\to\infty}\gamma_n^N(a, b),\qquad v_n=\lim_{b}\lim_{a}\lim_{N}v_n^N(a, b).$$

特别, $\quad V=\lim\limits_{b}\lim\limits_{a}\lim\limits_{N}V^N(a, b)$.

尽管定理 4.8 的样子难看, 但它的极其一般性使它十分有用. 第五章包含有它的几个应用.

定理 4.8 的证明方法也能产生更一般的

定理 4.9 设 $\{x_n(p), \mathscr{F}_n\}_1^\infty$ 为一族可积随机序列, 指标 p 为广义实参数, 在某个闭(可能无限)区间 $[a, b]$ 中取值. 设 $p_0\in[a, b]$, 假定 $p\uparrow\downarrow p_0$ 时 $x_n(p)\uparrow\downarrow x(p_0)$ $(n=1, 2, \cdots)$.

(a) 若存在一个 $p^*>p_0$ 使得

$$E(\sup_n x_n^+(p^*))<\infty,$$

且若 $p\downarrow p_0$ 时 $x_\infty(p)\downarrow x_\infty(p_0)$, 则 $p\downarrow p_0$ 时,

$$\gamma_n(p)\downarrow\gamma_n(p_0),\ v_n(p)\downarrow v_n(p_0)\quad(n=1,\ 2,\ \cdots).$$

(b) 若存在 $p_*<p_0$ 使得

$$\gamma_n(p)=\operatorname*{ess\,sup}_{t\in C_n(p_*)}E(x_t(p)\mid\mathscr{F}_n)\quad(p_*\leqslant p\leqslant p_0),$$

则 $p\uparrow p_0$ 时

$$\gamma_n(p)\uparrow\gamma_n(p_0),\ v_n(p)\uparrow v_n(p_0)\quad(n=1,\ 2,\ \cdots).$$

若对每个 n, $x_n(p)$ 关于 p 单调减少，成立一个类似的命题.

定理 4.7 的一个直接的应用是定理 4.5 的下述推广. 我们要推广下面这个思想去给出 σ 在 $\bar{C}$ 中最优的充要条件.

定理 4.5′ 若 $E(\sup x_n^+)<\infty$, 则 σ 在 $\bar{C}$ 中最优. 若又设 $x_n\to-\infty$, 则 $\sigma\in C$.

证明 定理 4.5 的证明表明 $Ex_\sigma\geqslant V$. 因为由定理 4.7, $Ex_t\leqslant V\,(t\in\bar{C})$, 可见 σ 在 $\bar{C}$ 中最优. 若 $x_n\to-\infty$, 很清楚 σ 不能以正概率取 $+\infty$ 值(及得到 $-\infty$ 的报酬).

定理 4.10 以 $\sigma(n)$ 记 $\min(\sigma,\ n)\,(n\geqslant1)$. 若 $V<\infty$, 则下列命题等价:

(a) σ 在 $\bar{C}$ 中最优,

(b) $(\gamma_{\sigma(n)})$ 一致可积,

(c) $(\gamma^+_{\sigma(n)})$ 一致可积,

(d) $\int_{(\sigma<\infty)}x_\sigma^+<\infty$ 及 $(I_{(\sigma>n)}\gamma_n^+)=(I_{(\sigma>n)}E^+(\gamma_{n+1}\mid\mathscr{F}_n))$ 一致可积.

证明 若 σ 在 $\bar{C}$ 中最优, 则用与定理 4.2 的证明中相同的方法可证(4.7)成立, 从而

$$E(x_\sigma\mid\mathscr{F}_n)=\gamma_{\sigma(n)}\quad(n\geqslant1).$$

由引理 2.2, (b)成立, 而由一致可积性的定义, (c)也成立. (c)与(d)的等价性是易证的. 现在我们证(c)包含(a). 由引理 4.8,

$$\{\gamma_{\sigma(n)},\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$$

是一个鞅，故$\{\gamma_{\sigma(n)}^+,\ \mathscr{F}_n,\ 1\leqslant n<\infty\}$是一个下鞅．令$\tilde{x}_n=\gamma_{\sigma(n)}^+$ $(n\geqslant1)$，则对任一停止变量 t 及 $N=1,\ 2,\ \cdots$，

$$\int_{(t<N)}\tilde{x}_t=\sum_1^N\int_{(t=k)}\tilde{x}_k\leqslant\sum_1^N\int_{(t=k)}\tilde{x}_N\leqslant E\tilde{x}_N,$$

令 $N\to\infty$ 我们有

$$E\tilde{x}_t\leqslant\sup_N E\tilde{x}_N<\infty,$$

故 $\tilde{V}<\infty$． 由定理 4.7，

$$E\gamma_\sigma^+=E\tilde{x}_\infty<\infty,\ \text{从而}\ Ex_\sigma^+<\infty.$$

现在对任一 $n=1,\ 2,\ \cdots,\ A\in\mathscr{F}_n$，

$$\int_A\gamma_{\sigma(n)}^+\leqslant\int_A\gamma_{\sigma(m)}^+\quad(m=n,\ n+1,\ \cdots),$$

故由 Fatou 引理(引理 1.2)，

$$(4.17)\qquad \int_A\gamma_{\sigma(n)}^+\leqslant\varlimsup_{m\to\infty}\int_A\gamma_{\sigma(m)}^+\leqslant\int_A\gamma_\sigma^+.$$

由引理 4.1—4.5，存在一列停止变量 $t_1\leqslant t_2\leqslant\cdots\leqslant\sigma$ 使得 $Ex_{t_i}\uparrow V$． 由引理 4.6，$\lim\limits_{i\to\infty}t_i=\sigma$．对每个 $c>0$，由(4.17)我们有

$$\begin{aligned}cP\{\gamma_{t_i}>c\}&\leqslant\int_{(\gamma_{t_i}>c)}\gamma_{t_i}=\sum_{n=1}^\infty\int_{(t_i=n,\gamma_n>c)}\gamma_n\\&=\sum_{n=1}^\infty\int_{(t_i=n,\gamma_n>c)}\gamma_{\sigma(n)}^+\leqslant\sum_{n=1}^\infty\int_{(t_i=n,\gamma_n>c)}\gamma_\sigma^+\\&=\int_{(\gamma_{t_i}>c)}\gamma_\sigma^+\leqslant E\gamma_\sigma^+<\infty,\end{aligned}$$

由此得出(见引理 2.2 的证明)$(\gamma_{t_i}^+)$一致可积，从而$(x_{t_i}^+)$更是一致可积的．由 Fatou 引理，

$$Ex_\sigma\geqslant\varlimsup_{i\to\infty}Ex_{t_i}=V.$$

因为由定理 4.7，$Ex_t\leqslant V\,(t\in\bar{C})$，由此得到 σ 在 $\bar{C}$ 中最优．

注 适当参考定理 2.1 可缩短上面的 (c) 包含 (a) 的证明. 一个有点不同的证明, 但也利用鞅收敛定理, 如下. 由定理 2.1, $\lim\limits_{n\to\infty}\gamma_{\sigma(n)}$ 存在(即在 $\{\sigma=\infty\}$ 上 $\gamma_n\to\gamma_\infty$) 及

$$(4.18)\qquad E\gamma_\sigma\geqslant\lim_{n\to\infty}E\gamma_{\sigma(n)}=E\gamma_1=V.$$

因此由下面的引理 4.13, 在 $\{\sigma=\infty\}$ 上 $x_\infty=\gamma_\infty$, 因为由 σ 的定义, 在 $\{\sigma<\infty\}$ 上 $x_\sigma=\gamma_\sigma$, 从(4.18)我们有

$$Ex_\sigma\geqslant V.$$

引理 4.13 若 $V<\infty$, 则对每个 $\varepsilon>0$,

$$P\{x_n\geqslant\gamma_n-\varepsilon\ \text{i. o.}\}=1,$$

其中 i. o. 是发生无限多次的缩写(见习题 7).

证明 以 $B=B(m,\varepsilon)$ 记事件: 对一切 $n\geqslant m, x_n\leqslant\gamma_n-\varepsilon$. 用反证法, 设对某个 $\varepsilon>0$, $m=1, 2, \cdots$, $P(B)>0$. 对一切 $t\in C_m$,

$$\int_B x_t\leqslant\int_B\gamma_t-\varepsilon P(B).$$

但 $\int_B x_t\leqslant\int_B\gamma_t$, 故由定理 4.6, 对一切 $t\in C_m$,

$$Ex_t\leqslant E\gamma_t-\varepsilon P(B)\leqslant E\gamma_m-\varepsilon P(B),$$

与定理 4.1 矛盾.

6. 例与反例

(a) 我们的第一个例子表明, 在一般情形下(4.11)不成立. 为了以后引证的需要, 我们十分详尽地构造这个例子. 设 $a_1, a_2, \cdots, b_1, b_2, \cdots$ 是两个非负递增实数列, Ω 是形如

$$\omega_j=(a_1, \cdots, a_j, -b_{j+1}, -b_{j+2}, \cdots)\quad(j=1, 2, \cdots)$$

的序列所组成的空间, $\mathscr{F}$ 是 Ω 的子集全体并且

$$P(\omega_j)=\frac{1}{j}-\frac{1}{j+1}\quad(j=1, 2, \cdots).$$

对每个 $n=2, 3, \cdots$, 令 $x_n(\omega_j)$ 是 ω_j 的第 n 个坐标, $\mathscr{F}_n=\mathscr{B}(x_2, \cdots, x_n)$, 容易看到对每一个 n,

$$P(x_n=a_n)=\frac{1}{n}=1-P(x_n=-b_n),$$

$$E(x_{n+1}|x_n=a_n)=a_{n+1}\frac{n}{n+1}-b_{n+1}\left(1-\frac{n}{n+1}\right),$$

$$Ex_n=\frac{a_n}{n}-b_n\left(1-\frac{1}{n}\right).$$

令 $t=$第一个 $k\geqslant 2$ 使得 $x_k=-b_k$, 且 $t_n=\min(t, n)(n=2, 3, \cdots)$. 显然, 在计算 $\{x_n, \mathscr{F}_n\}_2^\infty$ 的值或寻找最优规则时, 只有 (t_n) 是需要考虑的规则. 直接计算后可得

$$Ex_{t_n}^+=\frac{a_n}{n},\quad Ex_{t_n}^-=\sum_{j=2}^n\frac{b_j}{j(j-1)}.$$

现在令 $a_j=j(j-1)=b_j(j=1, 2, \cdots)$, 则 $Ex_{t_n}^+\to+\infty$, 并且按前一节的记号, 对一切 $a\neq-\infty$, $V(a)=\infty$. 但是 $Ex_{t_n}=n-1-\sum_2^n 1=0, V=0$, 于是 $\lim\limits_{a\to-\infty}V(a)=V$ 不成立.

(b) 我们的第二个例子表明定理 4.5(也见定理 4.10)中的某些条件, 如 $E(\sup x_n^+)<\infty$, 的必要性. 在例(a)的随机序列中, 令 $a_n=n^2-1$, $b_n=n(n-1)$, 则 $a_n\to-\infty$, $Ex_n\to-\infty$, 但是 $Ex_{t_n}=1-\frac{1}{n}$, 故最优规则不存在.

(c) 打折扣的单口硬币赌博机. 我们的最后一个例子是一个 Bayes 决策问题, 对于这个问题, 在 C 中不存在最优规则, 即使 σ 在 $\bar{C}$ 中是最优的.

设在关于 $p(0<p<1)$ 的条件下, $y_1, y_2, \cdots$是独立同分布随机变量序列, $P\{y_1=1\}=p=1-P\{y_1=-1\}$. 对某个 $0<\alpha<1$ 和每个 $n=0, 1, 2, \cdots$, 令 $x_n=\sum_{k=1}^n\alpha^{k-1}y_k$, $\mathscr{F}_n=\mathscr{B}(y_1,$

$\cdots, y_n)$，并假定存在一个已知的 p 的先验分布．由定理 4.5 知 σ 在 $\bar{C}$ 中是最优的．为简单起见，设 p 的先验分布是参数为(r, q)的 Beta 分布．令 $s_n = y_1^+ + \cdots + y_n^+$，众所周知(也容易验证)在观察到 $y_1, \cdots, y_n$ 后 p 的后验分布是参数为$(r+s_n, q+n-s_n)$的 Beta 分布．以 $V(r, q)$记序列$\{x_n, \mathscr{F}_n\}_0^\infty$ 的值，它是(r, q)的函数．有理由认为并且可以从定理 5.2 严格地推得

$$\sigma = \text{第一个 } n \geqslant 0 \text{ 使得 } V(r+s_n, q+n-s_n) \leqslant 0$$
$$(=\infty, \text{若不存在这样的 } n).$$

现在 $E(x_t|p)=2p-1$，从而 $V(r, q) \geqslant \dfrac{r-q}{r+q}$，这里的不等式是从考虑停止规则 $t=1$ 和参数为 (r, q) 的 Beta 分布的期望是$\dfrac{r}{r+q}$这一事实得到的．于是，如果 $r>q$，为了使得 σ 是无限的，只要对一切 $n \geqslant 1$ 有 $s_n > n/2$，但是当 $p>1/2$ 时，这种情形发生的概率大于零，于是如果 $r>q$，则$P\{\sigma=\infty\}>0$．

7．对 s_n/n 的最优停止

设 $y_1, y_2, \cdots$是均值为零，方差为 1 的独立随机变量，令 $s_n = \sum_1^n y_k$，$\mathscr{F}_n = \mathscr{B}(y_1, \cdots, y_n)$ $(n \geqslant 1)$．在这一节中我们将讨论 $\{s_n/n, \mathscr{F}_n\}_1^\infty$ 的最优停止规则的存在性问题，我们证明 σ 在 $\bar{C}$ 中是最优的，并且如果满足下面的条件 (4.19)，则 $\sigma \in C$ (这个条件在独立同分布场合总是满足的——见本节末尾的注(i))．

令 $s_0=0$，$\mathscr{F}_0=[\emptyset, \Omega]$，及 $s_k^{(n)} = s_k - s_n$ $(n \geqslant 0,\ k \geqslant n)$．

定理 4.11 设(y_n)，(s_n)，$(\mathscr{F}_n)$如上所述，则对 $\{s_n/n, \mathscr{F}_n\}_1^\infty$，$V<\infty$，且 σ 在 $\bar{C}$ 中最优．如果

(4.19) $$P\{\limsup_n n^{-1/2}s_n>0\}>0,$$
则 $\sigma\in C$.

在证明定理 4.11 之前，先证明两个引理.

引理 4.14

$$E(\sup|s_n/n|)<\infty,$$

且对每个 $1\leqslant p<2$, $n=1, 2, \cdots$,

(4.20) $$E(\sup_{k>n}|s_k^{(n)}/k|^p)\leqslant\frac{2}{2-p}n^{-p/2}.$$

证明 设 $n=1, 2, \cdots$是取定的，把 2.4 节(b) 应用于下鞅 $\{(s_k^{(n)})^2, \mathscr{F}_k\}_{n+1}^{\infty}$, 我们有

$$\begin{aligned}E(\sup_{k>n}|s_k^{(n)}/k|^p)&\leqslant p\int_0^{\infty}u^{p-1}P\{|s_k^{(n)}|>ku,\ 对某个\ k>n\}du\\&\leqslant n^{-(p/2)}+p\int_{1/\sqrt{n}}^{\infty}-\left(\sum_{n+1}^{\infty}k^{-2}\right)u^{p-3}du\\&\leqslant\frac{2}{2-p}n^{-(p/2)}.\end{aligned}$$

引理 4.15 对每个 $t\in\overline{C}$, $\varepsilon>0$, $1<p<2$, $n=1, 2, \cdots$, 有

(4.21) $$E(s_t/t|\mathscr{F}_n)<s_n/n,\quad 在\ A_n(t)\ 上,$$

其中

$$A_n(t)=\left\{s_n\geqslant\varepsilon n^{1/2},\ P\{t=\infty|\mathscr{F}_n\}>\frac{1}{2},\right.$$
$$\left.P\{n\leqslant t<\infty|\mathscr{F}_n\}<\left(\frac{2-p}{2}\right)^{1/(p-1)}\left(\frac{\varepsilon}{2}\right)^{p/(p-1)}\right\}.$$

证明 由强大数定律 (习题 2.14), 在 $\{t=\infty\}$ 上有 $s_t/t=0$, 于是由引理 4.14 和 Hölder 不等式, 在 $A_n(t)$ 上我们有

$$\begin{aligned}E\left(\frac{s_t}{t}\Big|\mathscr{F}_n\right)&=s_nE(t^{-1}|\mathscr{F}_n)+E\left(\frac{s_t^{(n)}}{t}I_{(n<t<\infty)}\Big|\mathscr{F}_n\right)\\&\leqslant\frac{s_n}{n}(1-P\{t=\infty|\mathscr{F}_n\})+\left[E\left(\sup_{k>n}\left|\frac{s_k^{(n)}}{k}\right|^p\right)\right]^{1/p}\end{aligned}$$

$$\times[P\{n\leqslant t<\infty|\mathscr{F}_n\}]^{(p-1)/p}$$
$$<\frac{s_n}{n}-\frac{\varepsilon n^{1/2}}{n}\cdot\frac{1}{2}+\left(\frac{2}{2-p}\right)^{1/p}n^{-1/2}\left(\frac{2-p}{2}\right)^{1/p}\frac{\varepsilon}{2}$$
$$\leqslant\frac{s_n}{n}.$$

定理 4.11 的证明 由引理 4.14 和定理 4.5′, $V<\infty$ 且 σ 在 $\bar{C}$ 中最优, 剩下要证明 $P\{\sigma<\infty\}=1$. 设结论不成立, 令

$$t=\text{第一个 } n\geqslant1 \text{ 使得 } A_n(\sigma) \text{发生},$$
$$=\infty, \text{ 若对一切 } n,\ \bar{A}_n(\sigma) \text{ 发生},$$

其中 $A_n(\sigma)$ 如引理 4.15 中所定义. 容易由 Lévy 定理(定理 1.4) 看到, 当 $n\to\infty$ 时 $P\{n\leqslant\sigma<\infty|\mathscr{F}_n\}\to0$, 并且在 $\{\sigma=\infty\}$ 上 $P\{\sigma=\infty|\mathscr{F}_n\}\to1$. 再由(4.19)和 Kolmogorov 0-1 律, 存在一个 $\varepsilon>0$ 使得 $P\{\limsup\limits_n n^{-1/2}s_n>\varepsilon\}=1$, 从而在$\{\sigma=\infty\}$上 t 是有限的. 令 $t'=\min(t,\sigma)$, 则由引理 4.15,

$$\int_{\{t<\sigma\}}\frac{s_\sigma}{\sigma}=\sum_1^\infty\int_{\{t=n<\sigma\}}E\left(\frac{s_\sigma}{\sigma}\Big|\mathscr{F}_n\right)$$
$$<\sum_1^\infty\int_{\{t=n<\sigma\}}\frac{s_n}{n}$$
$$=\int_{\{t<\sigma\}}\frac{s_t}{t}.$$

于是 $E(s_{t'}/t')>E(s_\sigma/\sigma)$, 得出矛盾.

上述方法修改以后可以处理形如 $h_n(s_n)$ 的报酬序列. 在下面我们指出一个稍有不同的方法, 它在某些情况下有其优点.

为了说明第二种方法, 我们将把它应用于随机序列 $\left\{\frac{s_n^2}{n^\alpha},\mathscr{F}_n\right\}_1^\infty$, 其中 $\alpha>1$. 因为对一切 n, $s_n^2\geqslant0$, 为了证明 σ 在 $\bar{C}$ 中是最优的, 由定理 4.10, 只要证明

(4.22) $$\lim_{n\to\infty} E\gamma_n=0.$$

下面的引理为序列(γ_n)提供了适当的上界.

引理 4.16 对每个$n=0, 1, 2, \cdots$和每个广义停止变量t,

(4.23) $$E\left(\frac{(s_t^{(n)})^2}{t^\alpha}\middle|\mathscr{F}_n\right)\leqslant\sum_{n+1}^{\infty}k^{-\alpha}<\infty, \text{ 在}\{t>n\}\text{上}.$$

特别,对每个$n=1, 2, \cdots$,

(4.24) $$E\left(\frac{(s_t^{(n)})^2}{t^\alpha}\middle|\mathscr{F}_n\right)\leqslant(\alpha-1)^{-1}n^{1-\alpha}, \text{在}\{t>n\}\text{上}.$$

证明 由定理 4.7 我们可以假设$t\in C$, 由引理 2.1, $\{(s_k^{(n)})^2, \mathscr{F}_k\}_{n+1}^{\infty}$是一个下鞅.对每一个固定的$n=0, 1, 2, \cdots$, $A\in\mathscr{F}_n$, 令$\tilde{A}=A\{t>n\}$, 我们有

(4.25) $$\int_{\tilde{A}}\frac{(s_t^{(n)})^2}{t^\alpha}\leqslant\sum_{n+1}^{\infty}k^{-\alpha}\left[\int_{\tilde{A}\{t\leqslant k\}}(s_k^{(n)})^2-\int_{\tilde{A}\{t\leqslant k-1\}}(s_{k-1}^{(n)})^2\right]$$
$$=\sum_{n+1}^{\infty}k^{-\alpha}c_k,$$

其中,对每个$N=n+1, n+2, \cdots$,

(4.26) $$\sum_{n+1}^{N}c_k\leqslant P(\tilde{A})E(s_N^{(n)})^2=P(\tilde{A})(N-n),$$

因为$k^{-\alpha}$是严格递减的, 如果$c_{n+1}=c_{n+2}=\cdots=P(\tilde{A})$, (4.25)的最后一项是受(4.26)约束之下的最大值, 这就证明了(4.23). 为证明(4.24), 我们以$\int_0^\infty(n+x)^{-\alpha}dx$为(4.23)的右边的上界.

现在我们能证明对序列$\{s_n^2/n^\alpha, \mathscr{F}_n\}_1^\infty$来说, σ在$\tilde{C}$中是最优的. 由(4.23)(其中取$n=0$),我们看到$V<\infty$. 为了证明(4.22),注意到由引理 4.16, 对每个$t\in c_{n+1}$,

$$E\left(\frac{s_t^2}{t^\alpha}\middle|\mathscr{F}_n\right)\leqslant2\left(\frac{s_n^2}{n^\alpha}+(\alpha-1)^{-1}n^{1-\alpha}\right)\quad(n=1, 2, \cdots).$$

于是 $$\gamma_n=\max\left(\frac{s_n^2}{n^\alpha}, E(\gamma_{n+1}|\mathscr{F}_n)\right)$$
$$\leqslant 3\frac{s_n^2}{n^\alpha}+2(\alpha-1)^{-1}n^{1-\alpha},$$
由此即得(4.22)。

如果对每个 $K>0$,

(4.27) $$P\{\limsup_n n^{-(1/2)}|s_n|>K\}=1,$$

则用类似于定理 4.11 的证法我们能够证明 $P\{\sigma<\infty\}=1$.

注 (i) 如果 $y_1, y_2, \cdots$独立同分布, 则条件 (4.19) 和 (4.27)成立, 更一般地, 譬如有

(4.28) $$\lim_{n\to\infty} P\left\{\frac{s_n}{\sqrt{n}}\leqslant x\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-(1/2)y^2}dy$$
$$(-\infty<x<\infty),$$

则对任一 K,
$$p\equiv P\left\{\limsup_{n\to\infty}\frac{s_n}{\sqrt{n}}>K\right\}\geqslant\limsup_{n\to\infty} P\left\{\frac{s_n}{\sqrt{n}}>K\right\}>0,$$
从而由 Kolmogorov 0-1 律(见 2.2 节(b))知 $p=1$.

(ii) 关于$\left|\frac{s_n}{n}\right|\left(\frac{s_n}{n}\right)$的结果可以容易地从相应于$\left(\frac{s_n}{n}\right)^2$ $\left(\left(\frac{s_n^+}{n}\right)^2\right)$的结果中推出. 事实上, 如果 $\{x_n, \mathscr{F}_n\}_1^\infty$ 是任一可积随机序列, 对于它 $V<\infty$ 及 σ 在 $\bar{C}(C)$ 中最优, 又 g 是任一非降凹函数, 则对 $\{g(x_n), \mathscr{F}_n\}_1^\infty$ 有 $V_g<\infty$及 σ_g 在 $\bar{C}(C)$中最优.

8. 条件 $V<\infty$ 和 $E[\sup x_n^+]<\infty$

定理 4.5′ 说, 如果

(4.29) $$E[\sup x_n^+]<\infty,$$

则 σ 在 $\bar{C}$ 中最优．但是，我们证明定理 3.1 的方法，这个定理在 3.6 节(a)中的应用以及上一节中的技巧表明，不用定理 4.5′，在某些情形下我们也能够证明

$$V<\infty, \tag{4.30}$$

及 σ 在 $\bar{C}$ 中最优．显然(4.29)蕴含了(4.30)．在这一节里我们将指出在何种程度上逆命题也成立．

（为了看出一般来说(4.30)不蕴含(4.29)，令 $\{x_n, \mathscr{F}_n\}_1^\infty$ 是一个不一致可积的非负鞅（这样的鞅的例子见 2.2 节(d)或 4.6 节(a)，其中取 $a_n=n$, $b_n=0$)．由定理 2.2，$V=Ex_1<\infty$，但是(4.29)不成立.）

定理 4.12 如果 $Ex_\infty^-<\infty$，且

$$\limsup_{a\to\infty} aP\{\sup_n x_n>a\}=+\infty, \tag{4.31}$$

则 $V=+\infty$.

证明 令 $a_m\to\infty$ 使得

$$a_mP\{\sup_n x_n>a_m\}\to\infty \quad (m\to\infty).$$

对每个 $m=1, 2, \cdots$，令

$$\begin{aligned} t(m) &= \text{第一个 } n\geqslant 1 \text{ 使得 } x_n>a_m, \\ &=\infty, \text{ 若不存在这样的 } n, \end{aligned}$$

则

$$V\geqslant Ex_{t(m)}\geqslant a_mP\{\sup_n x_n>a_m\}-Ex_\infty^-\to\infty, \quad \text{当 } m\to\infty \text{ 时}.$$

推论 若 $Ex_\infty^-<\infty$，且对某个 $0<\alpha<1$，

$$E(\sup_n x_n^+)^\alpha=\infty,$$

则 $V=\infty$.

证明 如果相反有 $V<\infty$，则由定理 4.12，存在常数 $c<\infty$，使得 $\sup_a aP\{\sup_n x_n>a\}\leqslant c$，于是对任何 $0<\alpha<1$，

$$E(\sup_n x_n^+)^\alpha=\alpha\int_0^\infty a^{\alpha-1}P\{\sup_n x_n>a\}da$$

$$\leqslant\alpha\left(1+c\int_1^\infty a^{-2+\alpha}da\right)<\infty,$$

这与推论的假设矛盾.

定理 4.12 及其推论说明, 如果 $Ex_\infty^-<\infty$, 则条件(4.30)并不明显地比(4.29)弱. 定理 4.14 证明, 当 $y_1, y_2, \cdots$独立同分布时, 对于 s_n/n 和 y_n/n 的问题 (见 4.7 节和 5.8 节) 来说, 条件(4.29)和(4.30)是等价的. 另一方面, 下面的定理证明, 对于定理 3.1 和 3.6 节(a)中的随机序列, (4.29)和(4.30)是明显不同的. 特别, 连同定理 4.12 及其推论, 它提供了又一个如同 4.6 节(a)中 $\sup_t Ex_t^+=\infty$, 但 $V<\infty$ 的现象的例子.

引理 4.17 设 $y_1, y_2, \cdots$ 是独立随机变量, 且 $Ey_n=0$, $\sigma_n^2=Ey_n^2<\infty$. 令 $s_n=\sum_1^n y_k(n=1, 2, \cdots)$, 则对任何 $a>0$ 和 $n=1, 2, \cdots$,

$$P\{\max_{1\leqslant k\leqslant n} s_k>a\}\leqslant 2P\left\{s_n>a-\left(2\sum_1^n\sigma_k^2\right)^{1/2}\right\}.$$

证明 以 t 记第一个 $k\geqslant1$ (如有的话) 使得 $s_k>a$, 并且令 $b_n=\left(2\sum_1^n\sigma_k^2\right)^{1/2}$, 则

$$P\{t\leqslant n,\ s_n\leqslant a-b_n\}=\sum_{k=1}^n P\{t=k,\ s_n\leqslant a-b_n\}$$

$$\leqslant\sum_{k=1}^n P\{t=k,\ s_n-s_k\leqslant-b_n\}$$

$$\leqslant\sum_{k=1}^n P\{t=k\}P\{|s_n-s_k|\geqslant b_n\}$$

$$\leqslant\frac{1}{2}P\{t\leqslant n\},$$

其中最后一个不等式是 Chebyshev 不等式和 b_n 的定义的推论. 于是

$$P\{t\leqslant n\}=P\{t\leqslant n,\ s_n>a-b_n\}+P\{t\leqslant n,\ s_n\leqslant a-b_n\}$$
$$\leqslant P\{s_n>a-b_n\}+\frac{1}{2}P\{t\leqslant n\},$$

由此即推得引理.

定理 4.13 设 $y, y_1, y_2, \cdots$是独立同分布随机变量, 且 $Ey=0$. 设 $s_n=\sum_1^n y_k(n\geqslant 1)$及 $\alpha>0$, 若

$$E[(y^+)^{1+\alpha}]<\infty, \tag{4.32}$$

则对每个 $a>0$, 有

$$E[\sup_n (s_n-na)^+]^\alpha<\infty, \tag{4.33}$$

$$E[\sup_n (y_n-na)^+]^\alpha<\infty. \tag{4.34}$$

反之,若对某个 $a>0$ 有(4.33)或(4.34)成立,则(4.32)满足.

证明 我们将给出 $\alpha=1$ 的特殊情形的证明,一般情形下的证明原则上是类似的,虽然在细节上会更复杂一些.

假定 (4.32) 成立, 用过渡到 $y_n(c)=\max(y_n, -c)$ 的办法, 其中 c 取得足够大使 $Ey(c)<\frac{a}{2}$,并注意到

$$s_n-na\leqslant\sum_1^n[y_k(c)-Ey(c)]-\frac{na}{2}\quad(n\geqslant 1),$$

我们看到, 不失一般性可以假定序列 $y_1, y_2, \cdots$是下有界的. 定义

$$y'_n=y_nI_{(y_n<n)},\qquad y''_n=y_n-y'_n,$$
$$s'_n=\sum_1^n y'_k\quad(n\geqslant 1).$$

为证明(4.33)(其 $\alpha=1$), 只要证明

$$E\left(\sum_1^\infty y''_n\right)<\infty, \tag{4.35}$$

(4.36) $$E(\sup[s'_n - na]^+) < \infty.$$

由(4.32)我们有

$$E\left(\sum_1^\infty y''_n\right) = \sum_{n=1}^\infty \sum_{k=n}^\infty \int_{(k<y\leqslant k+1)} y = \sum_{k=1}^\infty k \int_{(k<y\leqslant k+1)} y$$
$$\leqslant E(y^+)^2 < \infty.$$

为完成(4.33)的证明，余下还要证明(4.36)，而这只要证明对某个 $k_0 = 1, 2, \cdots$，有

$$\sum_{k=k_0}^\infty E\{\sup_{n_k \leqslant n < n_{k+1}} (s'_n - na)^+\} < \infty,$$

其中我们以 n_k 记 $\leqslant \exp(k)$ 的最大整数($k=0, 1, \cdots$)。由引理 4.17，对一切 $k=0, 1, \cdots$和 $u>0$，

(4.37) $$P\{\sup_{n_k \leqslant n < n_{k+1}} (s'_n - na)^+ > u\}$$
$$\leqslant P\left(\bigcup_{n=1}^{n_{k+1}} \{s'_n - Es'_n > u + n_k a\}\right)$$
$$\leqslant 2P\left\{s'_{n_{k+1}} - Es'_{n_{k+1}} > u + n_k a - \left(2\sum_1^{n_{k+1}} E(y'_n)^2\right)^{1/2}\right\}.$$

因为$\sum_1^n E(y'_i)^2 \leqslant (\text{const})n$ $(n \geqslant 1)$，由(4.37)，对一切充分大的 k_0 我们有

(4.38) $$\sum_{k=k_0}^\infty E\{\sup_{n_k \leqslant n < n_{k+1}} (s'_n - na)^+\}$$
$$\leqslant \sum_{k_0}^\infty \int_0^\infty P\{\sup_{n_k \leqslant n < n_{k+1}} (s'_n - na) > u\} du$$
$$\leqslant 2\sum_{k=k_0}^\infty \int_0^\infty P\{s'_{n_{k+1}} - Es'_{n_{k+1}} > u + a_1 n_{k+1}\} du,$$

其中我们令 $a_1 = \dfrac{a}{2e}$，现在利用 Markov 和 C_r 不等式，

$$P\{s'_{n_k} - Es'_{n_k} > u + a_1 n_k\}$$
$$\leqslant (u + a_1 n_k)^{-4} E[s'_{n_k} - Es'_{n_k}]^4$$

$$\leqslant (u+a_1n_k)^{-4}\left[\sum_1^{n_k} E(y_j'-Ey_j')^4+O(n_k^2)\right]$$

$$\leqslant 16(u+a_1n_k)^{-4}\left[\sum_1^{n_k} E(y_j')^4+O(n_k^2)\right],$$

于是由(4.38),

$$\sum_{k=k_0}^{\infty} E\{\sup_{n_k\leqslant n<n_{k+1}}(s_n'-na)^+\}$$

$$\leqslant \text{const}\left[\sum_{k=1}^{\infty}\sum_{j=1}^{n_k} n_k^{-3}E(y_j')^4+\sum_{k=1}^{\infty} n_k^{-3}O(n_k^2)\right]$$

$$\leqslant \text{const}\sum_{j=1}^{\infty}\sum_{k=[\log j]}^{\infty} n_k^{-3}E(y_j')^4$$

$$\leqslant \text{const}\sum_{j=1}^{\infty} j^{-3}E(y_j')^4\leqslant \text{const}\sum_{j=1}^{\infty}\sum_{k=1}^{j} j^{-3}\int_{(k-1<|y|\leqslant k)} y^4$$

$$\leqslant \text{const}\sum_{k=1}^{\infty} k^{-2}\int_{(k-1<|y|\leqslant k)} y^4\leqslant \text{const}\, Ey^2<\infty.$$

这就完成了在(4.32)假设下对(4.33)的证明.

现在设 $a>0$, 并对任何 $k=1, 2, \cdots$, 令

$$A_k=\{y_k>2(u+ak)\},\ B_k=\{|s_{k-1}|<u+ak\}.$$

则

$$P\left(\bigcup_1^{\infty} A_k\right)\leqslant \sum_1^{\infty} P(A_k)=\sum_{k=1}^{\infty}\sum_{j=k}^{\infty} P\left\{aj<\frac{y}{2}-u\leqslant a(j+1)\right\}$$

$$=\sum_{j=1}^{\infty} jP\left\{aj<\frac{y}{2}-u\leqslant a(j+1)\right\}$$

$$\leqslant \frac{1}{a}E\left(\frac{y}{2}-u\right)^+\to 0\quad (u\to\infty).$$

因此存在一个数 u_0, 使得对一切 $u\geqslant u_0$, $P\left(\bigcup_{k=1}^{\infty} A_k\right)<\frac{1}{3}$. 由弱大数定律我们还可假定对一切 $u\geqslant u_0$ 和 $k=1, 2, \cdots$, 有 $P(B_k)>\frac{2}{3}$, 从而对一切 $u\geqslant u_0$ 和 $n=1, 2, \cdots$,

$$P\{\max_{1<k<n}(s_k-ak)>u\}\geqslant P\Big(\bigcup_1^n A_kB_k\Big)$$

$$\geqslant\sum_{k=1}^n P(\overline{A}_1\cdots\overline{A}_{k-1}A_kB_k)$$

$$\geqslant\sum_{k=1}^n\Big[P(A_kB_k)-P\Big(A_k\cap\bigcup_{i=1}^{k-1}A_i\Big)\Big]$$

$$\geqslant\sum_{k=1}^n P(A_k)\Big[P(B_k)-P\Big(\bigcup_{i=1}^{k-1}A_i\Big)\Big]$$

$$\geqslant\frac{1}{3}P\Big(\bigcup_{k=1}^n A_k\Big).$$

令 $n\to\infty$，对一切充分大的 n，我们有

$$P\{\sup_n(s_n-na)^+>u\}\geqslant\frac{1}{3}P\{\sup_n(y_n-2na)^+>2u\},$$

因而如果对某一个 $a>0$，(4.33)成立，则(4.34)对 $2a$ 成立。为完成定理的证明，只要再证明如果对某个 $a>0$，(4.34)成立，则(4.32)一定满足．在 $a=1$ 的情形中，(4.34)蕴含着

$$\sum_{k=1}^\infty P\{\sup_n(y_n-na)>k\}<\infty,$$

或等价地有

$$\prod_{k=1}^\infty\prod_{n=1}^\infty F(k+na)>0,$$

这里我们以 F 记 y 的分布函数，并假定 $F(1)>0$，因为我们可以改变尺度．于是

$$\sum_{k=1}^\infty\sum_{n=1}^\infty\log F(k+na)>-\infty,$$

又因为 $\log F(x)\sim-(1-F(x))\ (x\to\infty)$，我们有

$$\int_0^\infty\int_0^\infty(1-F(x+ay))dxdy$$

$$\leqslant\text{const}+\sum_1^\infty\sum_1^\infty(1-F(k+na))<\infty,$$

令 $u=x+ay$，则 $\int_0^\infty\int_{ay}^\infty(1-F(u))\,du\,dy<\infty$，或等价地由 Fubini 定理，

$$\int_0^\infty u(1-F(u))du<\infty;$$

(在 $a=1$ 的情形)它转过来又与(4.32)等价.

定理 4.14 设 $y_1, y_2, \cdots$ 独立同分布，且对某个 $-\infty<\mu<\infty$，$Ey_i=\mu$，令 $s_n=\sum_1^n y_k(n=1, 2, \cdots)$，则下列命题等价：

(a) $E[y_1^+\log^+y_1^+]<\infty$, (b) $E[\sup_n s_n/n]<\infty$,

(c) $\sup_{t\in C} E[s_t/t]<\infty$, (d) $E[\sup y_n/n]<\infty$,

(e) $\sup_{t\in C} E[y_t/t]<\infty$.

证明 (b) $\Rightarrow$ (c)和(d) $\Rightarrow$ (e)是显然的，于是只要证明 (a) $\Rightarrow$ (b)，(a) $\Rightarrow$ (d)，(c) $\Rightarrow$ (a)和(e) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b). 对 2.1 节 (d) 的鞅的正部应用第二章的 (2.16)，我们有

$$(4.39)\qquad P\left\{\sup_{1\leqslant n\leqslant k}\frac{s_n^+}{n}\geqslant s\right\}\leqslant s^{-1}\int_{\{\sup_{1\leqslant n\leqslant k} s_n^+/n\geqslant s\}} y_1^+dP.$$

于是，令 $z_k=\sup_{1\leqslant n\leqslant k}(s_n^+/n)$，我们得到

$$\begin{aligned}Ez_k&=\int_0^\infty P\{z_k\geqslant s\}ds\leqslant 1+\int_1^\infty s^{-1}\int_{\{z_k\geqslant s\}}y_1^+dPds\\&=1+\int_{\{z_k>1\}}y_1^+\left(\int_1^{z_k}s^{-1}ds\right)dP\leqslant 1+E(y_1^+\log^+z_k).\end{aligned}$$

容易验证对一切 $a, b\geqslant 0$，

$$a\log^+b\leqslant a\log^+a+e^{-1}b,$$

从而有

$$Ez_k\leqslant\frac{e}{e-1}[1+E(y_1^+\log^+y_1^+)]<\infty.$$

令 $k\to\infty$，得

$$E\sup_n \frac{s_n^+}{n}\leqslant\frac{e}{e-1}[1+E(y_1^+\log^+y_1^+)]<\infty.$$

(a) $\Rightarrow$ (d)． 把 (a) $\Rightarrow$ (b) 的证明应用到序列 $y_1^+, y_2^+, \cdots$ 即可得

$$E\left[\sup_n\frac{y_n}{n}\right]\leqslant E\left[\sup_n\frac{y_n^+}{n}\right]\leqslant E\left[\sup_n n^{-1}\left(\sum_1^n y_k^+\right)\right]<\infty.$$

(e)$\Rightarrow$(a)　假设(a)不成立，由定理 4.7 知，只要给出一个广义停止变量 t，使得

$$E\left(\frac{y_t}{t}\right)=\infty.$$

对 $c>0$ 令 $t=\inf\{n: y_n\geqslant cn\}$．因为

$$\sum_1^\infty P\{y_n\geqslant cn\}=\sum_1^\infty P\{y_1\geqslant cn\}<\infty.$$

由此得出，对一切充分大的 c，

(4.40)　　$$P\{t=\infty\}=\prod_1^\infty P(y_n<cn)>0,$$

固定 $c>1$ 大得使(4.40)成立，则有

$$\begin{aligned}E\left(\frac{y_t}{t}\right)&\geqslant\sum_{n=1}^\infty n^{-1}\int_{(t=n)}y_n=\sum_{n=1}^\infty n^{-1}P\{t\geqslant n\}\int_{\{y_n\geqslant cn\}}y_n\\&\geqslant P\{t=\infty\}\sum_{n=1}^\infty n^{-1}\sum_{k=n}^\infty\int_{\{ck\leqslant y_1<c(k+1)\}}y_1\\&\geqslant P\{t=\infty\}\text{const}\sum_{k=1}^\infty\log(c(k+1))\int_{\{ck\leqslant y_1<c(k+1)\}}y_1\\&\geqslant P\{t=\infty\}\text{const}\,E(y_1^+\log^+y_1^+)=\infty.\end{aligned}$$

(c)$\Rightarrow$(a)．再次假设(a)不成立，不失一般性我们可以设 $\mu=0$，于是由强大数定律，对任意一个广义停止变量 t，在 $\{t=\infty\}$ 上 $\frac{s_t}{t}=0$，并且在 $\{t<\infty\}$ 上 $\frac{s_t}{t}\geqslant-\frac{s_{t-1}^-}{t}+\frac{y_t}{t}$．于是由前面部分的证明，只要证对那里定义的广义停止变量 t，

$$\int_{\{t<\infty\}}\left(\frac{s_{t-1}^{-}}{t}\right)<\infty.$$

但是

$$\int_{\{t<\infty\}}\frac{s_{t-1}^{-}}{t}\leqslant\int_{\{t<\infty\}}\left[\frac{1}{t}\sum_{k=1}^{t-1}y_k^{-}\right]=\sum_{n=1}^{\infty}n^{-1}\int_{\{t=n\}}\sum_{k=1}^{n-1}y_k^{-}$$

$$=\sum_{n=1}^{\infty}n^{-1}P\{y\geqslant cn\}\sum_{k=1}^{n-1}\int_{\{t\geqslant n\}}y_k^{-}$$

$$\leqslant Ey_1^{-}\sum_{n=1}^{\infty}P\{y_1\geqslant cn\}<\infty,$$

因为 $E|y_1|<\infty$.

注 不等式(4.39)表明，对报酬序列 $\frac{s_n}{n}$，(4.31) 不成立，即使 $V=\infty$.

9. 对鞅论的一个应用

下面叙述的定理 4.7 对鞅论的应用是有趣的. 设 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是一个上鞅，且 $\sup_n E|x_n|<\infty$，定理 2.1 说 $x_\infty=\lim_n x_n$ a.s. 存在且 $E|x_\infty|<\infty$.

用最优停止理论的术语，引理 3.3 (定理 2.2)是说，在某些条件下，

$$v_n=Ex_n\quad(n=1, 2, \cdots).$$

类似地，人们猜测 $\inf_t Ex_t$ 应该是 Ex_∞ (从定理 2.2 可知，对每一停止变量 $t\leqslant n$ 有

$$Ex_t^{-}\leqslant Ex_n^{-},$$

于是由 Fatou 引理，对每个停止变量 t，$Ex_t^{-}\leqslant\liminf_{n\to\infty}Ex_{\min(t,n)}^{-}\leqslant\sup E|x_n|<\infty$.)精确的叙述是

定理 4.15 如果 $\{x_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是一个上鞅，且 $\sup E|x_n|<\infty$，则下列命题等价：

(a) (x_n^-)是一致可积的，

(b) $v_n = Ex_n \quad (n=1, 2, \cdots)$，

(c) $Ex_\infty = \inf\limits_t Ex_t$.

证明 (a)⇒(b). 这可以从定理 2.2 推出.

(b)⇒(c). 对每个 n, $v_n = Ex_n$ 蕴含着对每个 $t \in C_n$ 有 $x_n \geq E(x_t | \mathscr{F}_n)$，于是由定理 4.7, $x_n \geq E(x_\infty | \mathscr{F}_n)$，因此对每个停止变量 t,

$$\int_{\{t=n\}} x_n \geq \int_{\{t=n\}} x_\infty.$$

从而(c)成立.

(c)⇒(a). 由定理 4.7, 令 $\gamma_n^* = \operatorname{ess\,sup}\limits_{t \geq n} E(x_t | \mathscr{F}_n)$，我们有 $\gamma_n^* \leq E(x_\infty | \mathscr{F}_n)(n=1, 2, \cdots)$. 由(c)可得 $\gamma_n^* = E(x_\infty | \mathscr{F}_n)$，从而 $x_n \geq E(x_\infty | \mathscr{F}_n) \quad (n=1, 2, \cdots)$. (x_n^-) 的一致可积性由引理 2.2 推得.

习 题

1. 证明：如果 $E(\sup x_n^+) < \infty$，且 $t \in C$，则存在一个 $\tau \in C$，使得 $\tau \leq t$ 且 $Ex_\tau = \sup\limits_{\tau' \leq t} Ex_{\tau'}$. （提示：考虑 $\{x_{\min(t,n)}, \mathscr{F}_n\}_1^\infty$.）
2. 利用第 6 节(a)的框架证明 σ 可能比其它任一停止变量处处更坏，即对一切停止变量 t, $P\{x_\sigma \leq x_t\} = 1$. 在第三章的什么地方我们遇见过这样的一个例子？
3. 定义 s 为
$$s = \text{第一个 } n \text{ 使得 } x_n = \gamma'_n,$$
证明：$s = \lim\limits_{N \to \infty} s^N$. 在单调情形，这个停止变量就是(3.19)的“$s$”.
4. 设 $\{x_n, \mathscr{F}_n\}_1^\infty$ 是任一随机序列，且 $E(\sup x_n^+) < \infty$（但是不假设 $Ex_n^- < \infty$），从首次到达原则证明定理 4.1. 现在假设 $E(x_n^-) < \infty$ $(n=1, 2, \cdots)$并利用引理 4.12，除去 $E(\sup x_n^+) < \infty$ 这个条件.
5. （续）第四章中有哪些其它结果不需要假设 $E(x_n^-) < \infty$？

6. 我们称一个上鞅 $\{y_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是半正则的，若对每个使得 Ey_t 存在的停止变量 t,

$$E(y_t|\mathscr{F}_n)\leqslant y_n, \quad 在 \{t>n\} 上.$$

证明：如果对每一个停止变量 t, Ex_t 存在，则 $\{\gamma_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是半正则的.（提示：注意到只要证明

$$E(\gamma_t|\mathscr{F}_n)\leqslant\gamma_n, \quad 在 \{t\geqslant n, \gamma_n\leqslant b\} 上,$$

并让 $b\to\infty$. 再适当地修改定理 4.6 的证明.）

7. 设 $\{y_n, \mathscr{F}_n, 1\leqslant n<\infty\}$ 是一个上鞅，且 y_∞ 是任一随机变量.对任一广义停止变量 t, 我们约定在 $\{t=\infty\}$ 上 $y_t=y_\infty$（注意这和我们通常的规定不同）. 证明：如果存在一个广义停止变量 t, 使得 $Ey_t^-<\infty$, 且对一切的 $n=1, 2, \cdots$,

$$(*) \qquad E(y_t|\mathscr{F}_n)\leqslant y_n, \quad 在 \{t\geqslant n\} 上,$$

那么把 t 换成任一停止变量 $\tau\leqslant t$, $(*)$式仍成立.（应用：如果 $V<\infty$, 且 σ 在 $\bar{C}$ 中最优，则对每个 $\varepsilon>0$,

$$\sigma(\varepsilon)=第一个 n 使得 x_n\geqslant\gamma_n-\varepsilon$$

在 C 中是"ε-最优的"（即 $\sigma(\varepsilon)\in C$ 且 $Ex_{\sigma(\varepsilon)}\geqslant V-\varepsilon$）.

8. 设 $y_1, y_2, \cdots$ 独立同分布，且 $Ey_1<0$, $E[(y_1^+)^2]<\infty$, 令

$$t=第一个 n\geqslant1 使得 \sum_1^n y_i>0,$$

$$=\infty, 若不存在这样的 n,$$

证明：$E\left(\sum_1^t y_i\right)^+=\infty$.（提示：应用定理 4.13.）

9. 设 $\mathscr{F}_1\subset\mathscr{F}_2\subset\cdots$ 及 $-x_n=y_n+c_n$, 其中 y_n 和 c_n 都是 $\mathscr{F}_n$-可测的，且 $y_n\geqslant0$, $0\leqslant c_n\uparrow\infty$. 定义

$$\tilde{\gamma}_N^N=-c_N,$$

$$\tilde{\gamma}_n^N=\max(x_n, E(\tilde{\gamma}_{n+1}^N|\mathscr{F}_n)) \quad (n=N-1, \cdots, 1).$$

证明：当 $N\uparrow\infty$ 时有 $\tilde{\gamma}_n^N\downarrow\gamma_n$. 在许多统计中有兴趣的问题里，这提供了另一个计算序列(γ_n)的方法.

10. 在习题 2.1 的记号下，$p=1/2$, 令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$ 及 $x_n=f(s_{\min(t,n)})$, 其中 f 是任一定义在 $\{-a, -a+1, \cdots, b-1, b\}$ 上的实值函数. 以 g 记大于 f 的最小凹函数. 证明：$\gamma_n=g(s_{\min(t,n)})$.（参阅 Dynkin, Yushkevich[1].）

第　五　章

Markov 和独立情形

点注意引导出定理 4.8 的各种可能的应用，在这一章里我们要定义 Markov 情形和独立情形，并表明定理 4.8 怎样使我们能推导出序列 (γ_n) 和规则 σ 的有用的定性性质.

1. Markov 情形——定义与基本定理

定义 设 $\{x_n, \mathscr{F}_n\}_1^\infty$ 是一个随机序列. 假设对每一个 $n=1, 2, \cdots$, 存在一个可测空间 $(Z_n, \mathscr{Z}_n)$ 和一个在 Z_n 中取值的 $\mathscr{F}_n$-可测随机变量 z_n, 使得 x_n 能表示成 z_n 的某个 $\mathscr{Z}_n$-可测函数，比方说 φ_n. 如果

$$(5.5)\qquad P\{z_{n+1}\in B\,|\,\mathscr{F}_n\}=P\{z_{n+1}\in B\,|\,z_n\}$$
$$(B\in\mathscr{Z}_{n+1},\ n=1,\ 2,\ \cdots),$$

我们称这给序列 $\{x_n, \mathscr{F}_n\}_1^\infty$ 提供了一个 Markov 表示. 如果还有 $Z_1=Z_2=\cdots=Z$, $\mathscr{Z}_1=\mathscr{Z}_2=\cdots=\mathscr{Z}$, 以及对每一个 n 和 $B\in\mathscr{Z}$, $P\{z_{n+1}\in B\,|\,z_n=z\}$ 是 Z 上与 n 无关的函数，则我们就有 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的一个平稳的 Markov 表示.

例子

(a) 考察 $x_n=\max(y_1, \cdots, y_n)-c_n$ 的问题 (3.6 节(a)), 令 $z_n=x_n$, 我们得到一个 Markov 表示; 令 $z_n=\max(y_1, \cdots, y_n)$, 我们得到一个平稳 Markov 表示. 类似的结果对报酬序列 $\frac{s_n}{n}$(4.7 节)也成立.

(b) 对于 3.1 节(h)的问题，令 $z_n=\pi_n$, 我们得到一个平稳 Markov 表示. 注意到对这个问题，我们不能令 $z_n=h(x_n)$ 或 $z_n=x_n$.

我们的第一个结果说，在 Markov 情形，不失一般性我们可以假定:

(i) 基本概率空间是 $(\Omega', \mathscr{F}', P')$, 其中 $\Omega'=Z_1\times Z_2$

$\times\cdots$, $\mathscr{F}'=\mathscr{Z}_1\times\mathscr{Z}_2\times\cdots$, 且 P'是P在$(z_1, z_2, \cdots)$之下的象;

(ii) 对每个 n, $\mathscr{F}'_n=\mathscr{B}(z_1, \cdots, z_n)$, 其中 z_n 现在看作为 Ω' 上的第 n 个坐标变量;

(iii) 一个停止规则 t 在下述意义下由一列吸收集(B_n)定义: 对每个 $n=1, 2, \cdots$, $B_n\in\mathscr{Z}_n$, $t=$第一个 $n\geqslant1$ 使得 $z_n\in B_n$.

我们将在下面的定理 5.1 中更精确地叙述这些结果, 它们的证明依靠两个引理.

引理 5.1 在 Markov 情形, γ_n 关于 $\mathscr{B}(z_n)$ 可测 $(n=1, 2, \cdots)$.

证明 直接应用定理 4.8 和 Markov 情形的定义即得.

引理 5.2 对每一个 $b>0$, 令$(x_n(b))$和$(\gamma_n(b))$如 4.5 节中所定义, 且对任一 $n=1, 2, \cdots$, 令 $t(b)=$第一个 $k\geqslant n$ 使得 $x_k\geqslant\gamma_k(b)-b^{-1}$, 则对一切 b, $t(b)\in C_n$, 并且

$$\gamma_n=\sup_b E(x_{t(b)}|\mathscr{F}_n), \qquad v_n=\sup_b Ex_{t(b)}.$$

证明 这引理由引理 4.8, 4.12 和 4.13 即可推得. 引理 4.10, 4.12 及引理 4.9 的证明也提供了另一种证法的基本要点(也可参阅第四章的习题 7).

定义 假设我们处于Markov情形, 用$D_n(\bar{D}_n)$记$C_n(\bar{C}_n)$的具有下述性质的子类: 对每个 $t\in D_n(\bar{D}_n)$, 存在 $B_k\in\mathscr{Z}_k$, $k=n, n+1, \cdots$, 使

$$\begin{aligned} t &= \text{第一个 } k\geqslant n \text{ 使得 } z_k\in B_k, \\ &= \infty, \text{若不存在这样的 } k. \end{aligned}$$

读者可能会认为定理 5.1 在直观上是明显的, 因为在 Markov 情形, 停止的决策只依赖于系统现在的状态, 而与过去所处的状态无关这一点看来是显然的. 在处理最优停止问

题中，这个原则常常被认为是理所当然的．它的严格论证和下面 5.3 节中随机化停止规则的讨论密切有关．

定理 5.1 在 Markov 情形，对每个 $n=1, 2, \cdots$,

$$\sigma_n \in \bar{D}_n,$$

并且

$$\gamma_n = \operatorname*{ess\,sup}_{t\in D_n} E(x_t|\mathscr{F}_n) = \operatorname*{ess\,sup}_{t\in D_n} E(x_t|z_n),$$

$$v_n = \sup_{D_n} Ex_t.$$

证明 从引理 5.1 和引理 5.2 立即可推得定理 5.1.

下面我们对平稳 Markov 情形感兴趣，我们可以并且就假定基本概率空间有上面讨论过的结构．我们还假定存在一个支配序列 (z_n) 进展的转移概率，即存在一个 $Z\times\mathscr{Z}$ 上的实值函数 $q(\cdot, \cdot)$，使得对每个 $B\in\mathscr{Z}$, $q(\cdot, B)$ 是 $\mathscr{Z}$-可测的，而对每个 $z\in Z$, $q(z, \cdot)$ 是 $\mathscr{Z}$ 上的概率测度，且 $q(z_n, B)$ 是 $P\{z_{n+1}\in B|z_n\}$ 的一个版本 $(n=1, 2, \cdots)$．因为由引理 5.1, γ_n（γ_n^N 等等）是 $\mathscr{B}(z_n)$-可测的，我们将以 $\gamma_n(z)$（$\gamma_n^N(z)$ 等等）记那个 Z 上的（可测）函数在 z 点的值，它在 $z_n(\omega)=z$ 时等于 $\gamma_n(\omega)$（$\gamma_n^N(\omega)$ 等等）．于是（对照定理 3.2）

$$\gamma_N^N(z) = \varphi_N(z),$$

(5.6)

$$\gamma_n^N(z) = \max\left(\varphi_n(z), \int_Z q(z, dz')\gamma_{n+1}^N(z')\right),$$

$$(n=N-1, \cdots, 1).$$

我们用 P^z 记支配从初始点 z 出发的 (z_n) 的行为的 $\mathscr{F}$ 上的概率，对每个 $n=0, 1, 2, \cdots$, 我们以 $V_n(z)$ 记 $(\Omega, \mathscr{F}, P^z)$ 上的 $\{\varphi_{n+k}(z_k), \mathscr{F}_k\}_1^\infty$（在 3.1 节意义下）的值，并置

$$\Gamma_n(z) = \max(\varphi_n(z), V_n(z)) \quad (n=1, 2, \cdots).$$

（我们已默认假定了对每个 $n=0, 1, 2, \cdots$ 和 $z\in Z$, $\{\varphi_{n+k}(z_k), \mathscr{F}_k\}_1^\infty$ 是 $(\Omega, \mathscr{F}, P^z)$ 上的可积随机序列．因为

$$E^z|\varphi_{n+k}(z_k)|=E(|\varphi_{n+k}(z_{n+k})|\,|z_n=z),$$

和$\{\varphi_n(z_n), \mathscr{F}_n\}_1^\infty$是$(\Omega, \mathscr{F}, P)$上的可积随机序列，由此推得对每个 $n=1, 2, \cdots$ 和几乎所有的（关于 $P\{z_n\in(\cdot)\}$）z，$E^z|\varphi_{n+k}(z_k)|<\infty$。在实际上，我们的附加假设不会引起任何麻烦.)

定理 5.2 在平稳 Markov 情形，存在(γ_n)的一个版本，使对每个 $n=1, 2, \cdots$，

$$V_n(\cdot)=E(\gamma_{n+1}(z_{n+1})\,|z_n=\cdot),$$

$$\Gamma_n(\cdot)=\gamma_n(\cdot).$$

证明 我们感兴趣的(γ_n)的版本是定理 4.8 中的三重极限. 只要证明关于 V_n 的结论就够了，由定理 4.8 我们可以假定对一切 n，$|\varphi_n|\leqslant B<\infty$. 如果 $V_n^N(\cdot)$ 类似于 $V_n(\cdot)$ 是对至多取 $N-n$ 个观察的规则类所定义的，那么只要证明对 $n=1, 2, \cdots$，有

$$(5.7)\qquad V_n^N(z)=E(\gamma_{n+1}^N(z_{n+1})\,|z_n=z)$$
$$(z\in Z,\ N=n+1, \cdots).$$

利用(5.6)容易看到，对 $N=2, 3, \cdots$，(5.7) 对 $n=N-1, \cdots, 1$ 成立.

推论 在定理5.2的假定下，设 $\varphi_n=\varphi-n(n=1, 2, \cdots)$，且令 $B=\{z:\varphi(z)\geqslant V_0(z)\}$，则对每个 $n=1, 2, \cdots$，

$$\{\sigma=n\}=\{z_1\notin B, \cdots, z_{n-1}\notin B, z_n\in B\}.$$

证明 显然有 $V_n=V_0-n(n=0, 1, \cdots)$，由定理和 σ 的定义即可得推论.

注 如果 $x_n=\sum_1^{n-1}W_k(z_k)+\varphi_n(z_n)$，其中 $z_1, z_2, \cdots$ 有(平稳) Markov 性(5.5)，由直接但冗长的论证我们能够证明一个类似的结果. 令

$$V_n(z)=\sup_t E^z\Big(\sum_1^{t-1} W_{n+k}(z_k)+\varphi_{n+t}(z_t)\Big),$$

现在我们有

$$E(\gamma_{n+1}|z_1, \cdots, z_n)=\sum_1^{n-1} W_k(z_k)+V_n(z_n)$$
$$(n=1, 2, \cdots),$$

从而

$$\begin{aligned}\sigma&=\text{第一个 } n\geqslant 1 \text{ 使得 } \varphi_n(z_n)\geqslant V_n(z_n),\\ &=\infty, \text{若不存在这样的 } n.\end{aligned}$$

这个注将在 5.2 节(b)中用到.

2. Markov 情形——应用

(a) 对一个简单备择假设检验一个简单原假设的问题在费用为常数、观察独立同分布时的 Bayes 解是 Wald 序贯概率比检验(见例 3.1(h)). 由定理 4.5, σ 是最优的, 且定理 5.2 的推论为第三章中引入问题时所作的直观的论证提供了一个严格的基础.

注 虽然定理 5.1 蕴含着, 在这个问题中我们可以不管随机变量 $y_1, y_2, \cdots$, 并且装作我们"观察"到了 $\pi_0, \pi_1, \pi_2, \cdots$, 但在我们的分析中, 诸 y_n 还是起了作用. 在我们利用表示式(见(3.3))

$$V_0(\pi)=\sup_t\{-[\pi(a\alpha_0+E_0t)+(1-\pi)(b\alpha_1+E_1t)]\}$$

时蕴含着假定我们原来处理的停止变量 t 是用 $y_1, y_2, \cdots$ 定义的. 正是这个假定使我们能推断出 α_i 和 $E_i(t)$ $(i=0, 1)$ 不依赖于 $\pi_0=\pi$, 从而 $V_0(\cdot)$是凸的.

(b) 在 4.3 节(c)的假设中, 还假定 $r_k=p^{k-1}q$ $(k=1, 2, \cdots)$, 其中 $0<p, q<1, p+q=1$, 则存在一个数 $\pi^*(0<\pi^*<1)$,

使

$$\sigma=\text{第一个 } n\geqslant 0 \text{ 使得 } \pi_n\geqslant\pi^*.$$

证明 记

$$\begin{aligned}v_n&=-\left[c(1-\pi_n)+\sum_{k=0}^{n-1}(n-k)p_k^n\right]\\&=-\left[c(1-\pi_n)+\sum_{0}^{n-1}\pi_k^n\right]\\&=-\left[c(1-\pi_n)+\sum_{0}^{n-1}\pi_k+\sum_{0}^{n-1}(\pi_k^n-\pi_k)\right],\end{aligned}$$

其中我们置$\pi_k^n=\sum\limits_{i=0}^{k}p_i^n(n=0,\ 1,\ \cdots,\ k=0,\ 1,\ \cdots)$，我们注意到

$$\left\{f_n=\sum_{0}^{n-1}(\pi_k^n-\pi_k),\ \mathscr{F}_n,\ 0\leqslant n<\infty\right\}$$

是一个 C-正则鞅．直接计算表明它是一个 $Ef_0=0$ 的鞅．此外，如果 $t\in C$，把引理 3.3(定理 2.2)应用到有界鞅

$$\{\pi_k^n,\ \mathscr{F}_n,\ k\leqslant n<\infty\},$$

我们就有

$$\begin{aligned}\infty>E\left(\sum_{k=0}^{t-1}\pi_k^t\right)&=\sum_{n=0}^{\infty}\int_{\{t=n\}}\sum_{k=0}^{n-1}\pi_k^n=\sum_{k=0}^{\infty}\sum_{n=k+1}^{\infty}\int_{\{t=n\}}\pi_k^n\\&=\sum_{k=0}^{\infty}\int_{\{t>k\}}\pi_k^t=\sum\int_{\{t>k\}}\pi_k=E\left(\sum_{0}^{t-1}\pi_k\right),\end{aligned}$$

其中最后一个等式是从导出前面三个等式的相反步骤得到的．于是

$$\begin{aligned}\int_{\{t>n\}}|f_n|&\leqslant\int_{\{t>n\}}\sum_{0}^{n-1}(\pi_k^n+\pi_k)=\int_{\{t>n\}}\sum_{0}^{n-1}(\pi_k^t+\pi_k)\\&\leqslant\int_{\{t>n\}}\sum_{0}^{t-1}(\pi_k^t+\pi_k)\to 0,\quad\text{当 } n\to\infty \text{ 时},\end{aligned}$$

再从引理 3.3(定理 2.2)就可得出

$$\{f_n,\ \mathscr{F}_n,\ 0\leqslant n<\infty\}$$

是 C-正则的. 由于 $Ef_t=Ef_0=0(t\in C)$, 我们可以并且就假定

$$x_n=-\left[c(1-\pi_n)+\sum_0^{n-1}\pi_k\right].$$

记 $V_0(\pi)=\sup\limits_{t\geq 1}E^\pi\left[-c(1-\pi_t)-\sum\limits_{k=0}^{t-1}\pi_k\right]$, 其中 E^π 记在假设 $\pi_0=\pi$ 下的期望. 容易从加在 (r_k) 上的条件和 5.1 节末尾的注看出

$$\begin{aligned}\sigma&=\text{第一个 } n\geq 0 \text{ 使得} -c(1-\pi_n)\geq V_0(\pi_n)\\&=\text{第一个 } n\geq 0 \text{ 使得 } \pi_n-1\geq c^{-1}V_0(\pi_n).\end{aligned}$$

从表示式

$$-V_0(\pi)=\inf_{t\geq 1}\left\{\pi E_1t+(1-\pi)q\sum_{n=1}^{\infty}p^{n-1}(cP_0\{t<n\}\right.$$
$$\left.+\hat{E}((t-n)^+|\theta=n))\right\}$$

(见上面(a)之后的注), 其中

$$\hat{E}((t-n)^+|\theta=n)=\sum_{n+1}^{\infty}(k-n)\int_{(t\geq n)}P_1(t=k|\mathscr{F}_{n-1})dP_0.$$

我们看出 $V_0(\cdot)$ 是凸的. 由于 $V_0(1)=-1$, 即得欲证之结果.

(c) 设 $y_1, y_2, \cdots$ 是独立同分布随机变量序列, 诸 y_n 具有关于某个 σ-有限测度 μ 的概率密度函数 f, 且已知 f 是两个特定的密度函数 f_0 或 f_1 中的一个. 如果 $f=f_1$, 则每次观察的费用为一个单位. 如果 $f=f_0$, 则没有抽样费用, 但是如果我们停止抽样, 就要承担一个单位的费用. 所以, 如果看来是从 f_1 中抽样, 我们就要尽快停止, 而只要看来是 $f=f_0$, 我们就要无限地继续抽样下去. 如果存在一个 $f=f_0$ 的先验概率 $\pi(0<\pi<1)$, 我们就要找一个广义停止变量使

(5.8) $$\pi P_0\{t<\infty\}+\bar{\pi}E_1t$$

最小, 其中 $\bar{\pi}=1-\pi$, 且我们用 $P_i(E_i)$ 记无限序列 $(y_1, y_2, \cdots)$

的空间上由 $f_i(i=0,\ 1)$ 决定的概率(期望).

令 $f_{in}=f_i(y_1)\cdots f_i(y_n)\ (i=0,\ 1;\ n=1,\ 2,\ \cdots)$，对于任一使得 $P_0\{t<\infty,\ f_{1t}=0\}=0$ 和 $E_1 t<\infty$ 的停止规则 t (在试图使(5.8)最小时只要考虑这样的规则)，省略掉记号 $d\mu(y_1)\cdots d\mu(y_n)$，我们有

$$P_0\{t<\infty\}=\sum_1^\infty\int_{(t=n)}f_{0n}=\sum_1^\infty\int_{(t=n)}\frac{f_{0n}}{f_{1n}}\cdot f_{1n}=E_1\left(\frac{f_{0t}}{f_{1t}}\right).$$

令 $-x_n=\pi\dfrac{f_{0n}}{f_{1n}}+\bar{\pi}n$, $\mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n)\ (n\geqslant 1)$，我们看到为使 (5.8) 最小，只要使 E_1x_t 最大，即在概率 P_1 下求解 $\{x_n,\ \mathscr{F}_n\}_1^\infty$ 的最优停止问题.

令 $z_n=\dfrac{f_{0n}}{f_{1n}}(n\geqslant 1)$，我们处于平稳 Markov 情形. 由定理 4.5, σ 在 C 中是最优的，并且由定理 5.2,

$$\sigma=\text{第一个 } n\geqslant 1 \text{ 使得} -\pi z_n\geqslant V_0(z_n),$$

容易看到 $V_0(0)=-\bar{\pi}$ 及 $V_0(\cdot)$ 是凸的. 由此得到，对某个常数 c

$$\sigma=\text{第一个 } n\geqslant 1 \text{ 使得 } \frac{f_{0n}}{f_{1n}}\leqslant c, \tag{5.9}$$

然而，因为 $\{z_n,\ \mathscr{F}_n\}_1^\infty$ 在 P_1 下是一个鞅，对一切 N 我们有 $s^N\equiv 1$, 从而 $V^N\nrightarrow V$. 令 $s_n=\sum\limits_1^n\log\dfrac{f_0(y_k)}{f_1(y_k)}$，我们注意到 (5.9)变成

$$\sigma=\text{第一个 } n\geqslant 1 \text{ 使得 } s_n\leqslant\log c,$$

且(假定 $P_0(\sigma=1)<1$)

$$1>P_0(\sigma<\infty)=E_1\left(\frac{f_{0\sigma}}{f_{1\sigma}}\right)=E_1(e^{s_\sigma}),$$

与 2.5 节的结果一致. 为了计算 (γ_n)，我们置

$$\gamma_N^N=-\bar{\pi}N,$$

$$\gamma_n^N=\max(x_n,\ E(\gamma_{n+1}^N|\mathscr{F}_n))\quad(n=N-1,\ \cdots,\ 1),$$

则由引理 4.9, $\lim\limits_{N\to\infty}\gamma_n^N=\gamma_n(n\geqslant1)$.

(d) 下述(c)的另一种叙述方法提供了这一类问题的一个例子，在这类问题里用广义停止规则是有好处的，但是在时刻 $+\infty$ 的报酬 x_∞，适合于这类问题的不是 $\limsup\limits_{n\to\infty}x_n$. 因此第四章的结果不能应用. 例如定理 4.7, 它蕴含着 $V=\sup\limits_t Ex_t$, 不管其中的 sup 是取遍停止规则类还是广义停止规则类都是一样的，就不成立. 但是我们能用下面指出的直接的方式把一般理论作一个适当的推广.

考虑上述(c)中相同的问题，但是令 $P=\pi P_0+\bar{\pi}P_1$. 在 P 下，对每个 n, y_1, $\cdots$, y_n 的联合密度等于 $\pi f_{0n}+\bar{\pi}f_{1n}$. 置 $\pi_n=\dfrac{\pi f_{0n}}{\pi f_{0n}+\bar{\pi}f_{1n}}$, $\bar{\pi}_n=1-\pi_n$, 于是

$$\begin{aligned}(5.10)\qquad \pi P_0(t<\infty)+\bar{\pi}E_1t&=\sum_1^\infty\int_{(t=n)}(\pi f_{0n}+\bar{\pi}nf_{1n})\\&=\sum_1^\infty\int_{(t=n)}(\pi_n+n\bar{\pi}_n)(\pi f_{0n}+\bar{\pi}f_{1n})\\&=\int_{(t<\infty)}(\pi_t+t\bar{\pi}_t)dP,\end{aligned}$$

因为由 2.2 节(d),

$$P_1\left\{\frac{f_{0n}}{f_{1n}}\to0\right\}=1\ \text{及}\ P_0\left\{\frac{f_{1n}}{f_{0n}}\to0\right\}=1,$$

故得 $\pi_\infty=\lim\limits_{n\to\infty}\pi_n$ a.s. P 存在且 $P_0(\pi_\infty=1)=1$, $P_1(\pi_\infty=1)=0$. 令

$$-x_n=\pi_n+n\bar{\pi}_n,\quad \mathscr{F}_n=\mathscr{B}(y_1,\ \cdots,\ y_n)\quad(n\geqslant1),$$

$$-x_\infty=\infty I_{(\pi_\infty=1)},$$

并约定对任一广义停止变量 t,

$$x_t=\begin{cases} x_n, & 若\ t=n \quad (n=1,\ 2,\ \cdots), \\ x_\infty, & 若\ t=\infty, \end{cases}$$

由(5.10)我们有

$$\pi P_0(t<\infty)+\bar{\pi}E_1t=-Ex_t,$$

从而使(5.8)最小就等价于使 Ex_t 最大．但是这后一问题并不属于我们所发展的最优停止理论的范围，因为 $P\{x_\infty\neq \limsup x_n\}>0$．此外，容易看出 $\{\pi_n,\ \mathscr{F}_n\}_1^\infty$ 是一个有界鞅，于是由引理3.3(定理2.2)，对每一个 $P\{t<\infty\}=1$ 的停止变量 t，

$$Ex_t=-\pi-E_1t.$$

更一般地，假设 $\{x_n,\ \mathscr{F}_n\}_1^\infty$ 是任一可积随机序列，x_∞ 是一个任意的 $\mathscr{B}\left(\bigcup_1^\infty\mathscr{F}_n\right)$-可测的随机变量．对任一广义停止变量 t，令 x_t 在 $\{t=n\}$ 上等于 x_n，$n=1,\ 2,\ \cdots$，在 $\{t=\infty\}$ 上等于 x_∞．以 $\bar{C}$ 记所有使 $Ex_t^-<\infty$ 的那种 t 的类，$\bar{V}=\sup\limits_{t\in\bar{C}}Ex_t$ 等等．可以用直接的方式把我们的一般理论作一个适当的推广．例如，基本引理4.11(a)变成

引理 定义

$$\tilde{\gamma}_N^N=E(\max\ (x_\infty,\ \sup_{k>N}x_k)\,|\,\mathscr{F}_N),$$

$$\tilde{\gamma}_n^N=\max(x_n,\ E(\tilde{\gamma}_{n+1}^N\,|\,\mathscr{F}_n))\ (n=N-1,\ \cdots,\ 1),$$

$$\tilde{\gamma}_n=\lim_{N\to\infty}\tilde{\gamma}_n^N.$$

如果 $E(\max[x_\infty,\ \sup\limits_n x_n])^+<\infty$，则 $(\tilde{\gamma}_n)=(\bar{\gamma}_n)$；若又有 $x_\infty\leqslant\limsup\limits_{n\to\infty}x_n$，则 $(\bar{\gamma}_n)=(\gamma_n)$．

细节我们就省略了(见习题8—10)．

3. 随机化停止规则

从一种观点看来，定理5.1是一个关于消去"随机化"停

止规则的结果．在 Markov 情形，在任一时刻 n, x_n, x_{n+1}, $\cdots$ 的条件联合分布仅仅通过 z_n 的值依赖于过去，所以在第 n 步停止的决策只应依赖于 z_n 看来是合理的．允许我们的决策依赖于以前的观察等于把我们的决策基于一个辅助的"不相干的"随机化之上．

现在假设给定一个可积随机序列 $\{x_n, \mathscr{F}_n\}_1^\infty$，它具有性质：存在一个基本的（一般取抽象值的）随机变量序列 y_1, y_2, $\cdots$，使 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)(n=1, 2, \cdots)$．容易看到，不论 $\{x_n, \mathscr{F}_n\}_1^\infty$ 在其上定义的特定的基本概率空间 $(\Omega, \mathscr{F}, P)$ 是怎样的，只要假定 $(\Omega, \mathscr{F}, P)$ 是与 y_1, y_2, $\cdots$ 的（给定的）联合分布一致，$\{x_n, \mathscr{F}_n\}_1^\infty$ 的最优停止问题都是相同的．

如果 $\mathscr{G}_1$, $\mathscr{G}_2$, $\cdots$ 是 $\mathscr{F}$ 的子 σ–代数的非降序列，使对每个 $n=1, 2, \cdots$,

(a) $\mathscr{F}_n\subset\mathscr{G}_n$,

(b) $P(A|\mathscr{G}_n)=P(A|\mathscr{F}_n)\quad\left(A\in\mathscr{B}\left(\bigcup\limits_1^\infty\mathscr{F}_k\right)\right)$,

则我们称任一关于 $(\mathscr{G}_n)$ 的停止变量 t（即对一切 $n=1, 2, \cdots$, $\{t=n\}\in\mathscr{G}_n$）为对 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的随机化停止变量，对 $\{x_n, \mathscr{F}_n\}_1^\infty$ 的随机化停止变量全体组成的类就是当 $(\Omega, \mathscr{F}, P)$ 取遍与 y_1, y_2, $\cdots$ 的特定联合分布一致的概率空间类时能从这样的序列 $(\mathscr{G}_n)$ 得到的停止变量全体组成的类．我们指出，在第 n 步做一个（可测地）依赖于 y_1, $\cdots$, y_n 的随机试验来决定是否停止这种直观的随机化方法是符合上述模型的．事实上，在这种情形，我们有 $\mathscr{G}_n=\mathscr{B}(y_1, \cdots, y_n$ 和一切直到（包括）第 n 步为止的随机试验）．

例子

设 $y_k=k$, $x_k=y_k(k=1, 2, \cdots)$．对这种情形，"自然的"

概率空间是由单个元素ω构成的空间Ω. $V=\infty$,而且不存在最优规则. 如果我们令$(\Omega,\ \mathscr{F},\ P)$是单位区间、Borel集全体和Lebesgue测度, 则$V=\infty$, 再令$\mathscr{G}_1=\mathscr{G}_2=\cdots=\mathscr{F}$, $t(\omega)=2^k$, 若$\frac{1}{2^k}<\omega\leqslant\frac{1}{2^{k-1}}(k=1,\ 2,\ \cdots)$,我们有

$$Ex_t=\sum_{k=1}^{\infty}2^k\cdot\frac{1}{2^k}=\infty,$$

所以存在一个最优的随机化停止变量.

我们的基本结果是：随机化的引入不改变$\{x_n,\ \mathscr{F}_n\}_1^\infty$的值$V$.

定理5.3 设$\{x_n,\ \mathscr{F}_n\}_1^\infty$是任一随机序列, $\mathscr{G}_1\subset\mathscr{G}_2\subset\cdots$是任一具有性质(a)及(b)的$\mathscr{F}$的子$\sigma$-代数序列, 则序列$\{x_n,\ \mathscr{G}_n\}_1^\infty$的诸$\gamma_n$与序列$\{x_n,\ \mathscr{F}_n\}_1^\infty$的诸$\gamma_n$是相同的.

证明 从定理4.8和上述条件(b)立即可得证明.

由于广义停止变量σ是用序列(γ_n)定义的, 故定理5.3连同定理4.2蕴含了：若$V<\infty$,我们可以限于考虑非随机化的规则. 这一节的第一个例子证明：若$V=\infty$,可以存在一个最优的随机化规则, 但是不存在最优的非随机化规则. 更一般地,每当$V=\infty$时,总存在停止变量序列(t_k),使有$\lim\limits_{k\to\infty}Ex_{t_k}=\infty$. 如有必要通过取子列使我们总可假定$Ex_{t_k}\geqslant2^k$; 以概率$2^{-k}$使用停止变量$t_k$, 我们也许可以希望得到一个随机化停止变量$t$使得$Ex_t=\sum\limits_{k=1}^{\infty}2^{-k}E(x_{t_k})\geqslant\sum\limits_{1}^{\infty}1=\infty$. 这条推理的思路一般是不对的,这是下面第一个例子的内容.

例子

对于4.6节(a)建立的例子,设

$$a_k=k^2(k+2),$$

$$b_k = 2k^2(k-1),$$

则 $Ex_{t_n} = n+2 \uparrow \infty$，但是

$$1 \geqslant \frac{n+2}{n^2+n-2} = \frac{Ex_{t_n}}{Ex_{t_n}^-} \to 0.$$

因此规则(t_n)的任一混合(p_n)（它形式地给出$\sum_1^\infty p_n Ex_{t_n} = \infty$）事实上必定义一个规则 t 使得 $Ex_t^+ = Ex_t^- = \infty$.

下面的例子提供了随机化概念的一个有趣的应用. 设 $y, y_1, y_2, \cdots$独立同分布，$P\{y=1\} = p = 1-P\{y=0\}$，$\mathscr{F}_n = \mathscr{B}(y_1, \cdots, y_n)$，$x_n = \frac{y_1+\cdots+y_n}{n}$，我们将证明$V = V(p)$关于$p$是递增且连续的. 设 $Y_1, Y_2, \cdots$ 独立且在$(0, 1)$上均匀分布，

$$\mathscr{G}_n = \mathscr{B}(Y_1, \cdots, Y_n),$$

$x_n(p) = Y_1, \cdots, Y_n$ 中 $\leqslant p$ 的项数的比例数. 由定理 5.3，$\{x_n, \mathscr{F}_n\}_1^\infty$ 的值和 $\{x_n(p), \mathscr{G}_n\}_1^\infty$ 的值是相同的. 因为 $x_n(\cdot)$ 关于 p 递增，故得 $V(\cdot)$关于 p 是递增的，$V(\cdot)$的连续性由定理 4.9 推得.

4. G. Elfving 问题

假定一个人有件商品要出售，例如一幢房子. 他不时地收到出价，对此他必须接收或拒绝. 他推迟出售的时间越长（等待一个大的出价），他在利息、折旧、纳税等方面的损失也越多. 他应怎么办呢？

我们将考虑这一情况的下述简单的模型. 设 $\{N(u), u \geqslant 0\}$是强度函数为 $p(u)$ 的一个非齐次 Poisson 过程.（过程有独立增量且 $N(u+\delta) - N(u) = 1$ 的概率为 $\delta p(u) + o(\delta)$.）以 $\tau_1, \tau_2, \cdots$记$\{N(u), u \geqslant 0\}$的样本轨道逐次跳跃的时刻，即

$$\tau_0=0,$$

$$\tau_k=\inf\{u: u\geqslant\tau_{k-1},\ N(u)\neq N(\tau_{k-1})\}.$$

与 $\tau_1,\ \tau_2,\ \cdots$(逐次出价的时刻)联系着独立非负随机变量 $y_1,\ y_2,\ \cdots$(出价的金额),它们具有共同的分布 F 和有限均值 μ。还给定一个非负、非增、右连续的折扣函数 $r(\cdot)$,$r(0)=1$。对 $n=1,\ 2,\ \cdots$,令 $\mathscr{H}_n=\mathscr{B}(\tau_1,\ \cdots,\ \tau_n,\ y_1,\ \cdots,\ y_n)$,$x_n=y_n r(\tau_n)$,我们将假定

$$\int_0^\infty r(u)p(u)du<\infty,$$

所以

$$E\Big(\sum_1^\infty y_n r(\tau_n)\Big)=\mu\int_0^\infty r(u)p(u)du<\infty, \tag{5.11}$$

从而有 $E(\sup x_n)<\infty$ 和 $x_\infty=0$. 由定理 4.5′,对随机序列 $\{x_n,\ \mathscr{H}_n\}_1^\infty$,$\sigma$ 在 $\bar{C}$ 中是最优的.

下面我们将证明,特别注意到所给问题的"连续时间"这一方面,我们能够消除在计算 σ 中通常遇到的某些困难,甚至在某些特殊情形还能把结果明显计算出来.

为此令

$$\mathscr{F}_u=\mathscr{B}(N(u'),\ u'\leqslant u;\ y_1,\ \cdots,\ y_{N(u)}),\quad u\geqslant0.$$

于是对每个 $n=1,\ 2,\ \cdots$,

$$\mathscr{H}_n=\mathscr{F}_{\tau_n},$$

其中按定义

$$\mathscr{F}_{\tau_n}=\{A: A\{\tau_n\leqslant u\}\in\mathscr{F}_u \text{ 对一切 } u\geqslant0\}.$$

注意到 $\lim\limits_{u\to\infty}EN(u)=\int_0^\infty p(u')du'$ 可能是有限的. 在这种情形,存在第一个指标 n(依赖于特定的样本轨道 $N(\cdot)$)使得 $\tau_n=\infty$. 采用时间尺度变换

$$\tilde{u}=\int_0^u p(u')du',$$

它是正 u-轴到有限或无限区间 $0 \leqslant \tilde{u} \leqslant \tilde{U} = \int_0^{\infty} p(u')du'$ 上的一一映照，我们可以假定 $p(u) \equiv 1$. 今后我们将省略记号"$\sim$"，不失一般性假定 $p(u) \equiv 1$，且存在一个 $U \leqslant \infty$，使得 r 在 $[0, U)$ 上是正的，在 $[U, \infty)$ 上是零.

令 $z_n = (y_n, \tau_n)$，则可看到我们处于平稳 Markov 情形. 而且，$V_n(z) = \sup_t E(y_t r(u+\tau_t))$ 只是 u 的函数，记作 $V(u)$. 于是由定理 5.2，

$$\sigma = \text{第一个 } n \text{ 使得 } y_n r(\tau_n) \geqslant V(\tau_n).$$

对 $0 \leqslant u < U$，令 $x_n(u) = y_n r(u+\tau_n)$ $(n=1, 2, \cdots)$，并且考虑随机序列族 $\{x_n(u), \mathscr{F}_{\tau_n}\}_1^{\infty}$. 因为对任何固定的 u_0，$P\{u_0 + \tau_n \in r(\ )$ 的不连续点集$\} = 0$，所以得出，对任何 $0 < u_0 < U$，当 $u \uparrow u_0$ 时有 $x_n(u) \downarrow x_n(u_0)$，和对任何 $0 \leqslant u_0 < U$，当 $u \downarrow u_0$ 时有 $x_n(u) \uparrow x_n(u_0)$. 定理 4.9 的其余的条件是容易验证的，从而我们看到

$$V(u) = E\gamma_1(u)$$

在 $[0, U)$ 上是连续的. 令

$$(5.12) \qquad y(u) = V(u)/r(u), \qquad 0 \leqslant u < U,$$
$$= 0, \qquad u \geqslant U,$$

可看得出 $y(\cdot)$ 是分段连续的，并且

$$\sigma(u) = \text{第一个 } n \text{ 使得 } y_n \geqslant y(u+\tau_n) \quad (\sigma = \sigma(0)).$$

记

$$G(y) = 1 - F(y), \quad H(y) = \int_y^{\infty} y' dF(y'),$$

对 $0 \leqslant u \leqslant U$，$0 \leqslant v \leqslant U - u$，令

$$f_u(v) = P\{\tau_{\sigma(u)} > v\}.$$

由于 $y(\cdot)$ 是分段连续的，我们可以用通常的微分方法得到

$$P\{\tau_{\sigma(u)}>v+h\,|\,\tau_{\sigma(u)}>v\}$$
$$=\frac{f_u(v+h)}{f_u(v)}=1-G(y(u+v))h+o(h),$$
它连同对 $h<0$ 的类似结果导致
$$\frac{f'_u(v)}{f_u(v)}=-G(y(u+v)).$$
由于 $f_u(0)=1$，我们有
$$f_u(v)=\exp\left[-\int_u^{u+v}G(y(v'))dv'\right].$$
从
$$V(u)=E[y_{\sigma(u)}r(u+\tau_{\sigma(u)})]$$
$$=\int_0^{U-u}E(y_{\sigma(u)}r(u+\tau_{\sigma(u)})\,|\,\tau_{\sigma(u)}=v)(-f'_u(v))dv,$$
经变量变换之后我们可得

(5.13) $$y(u)r(u)=V(u)$$
$$=\int_u^U r(v)H(y(v))\exp\left[-\int_u^v G(y(v'))dv'\right]dv.$$
对 u 求导我们有

(5.14) $$\frac{d}{du}[r(u)y(u)]=-r(u)\varphi(y(u)),$$
其中我们令

(5.15) $$\varphi(y)=H(y)-yG(y).$$
现在我们宣称：

(i) 一个分段连续函数 $\tilde{y}(\cdot)$ 满足 (5.13) 的充要条件是 $\tilde{y}(\cdot)$ 满足 (5.14) 和

(5.16) $$r(u)\tilde{y}(u)\to 0,\text{ 当 } u\to U \text{ 时}.$$
(ii) 若 $\tilde{y}(\cdot)$ 在 $[0,U)$ 上满足 (5.13)，且在 $[U,\infty)$ 上 $\tilde{y}(\cdot)=0$，则
$$\sigma=\text{第一个 } n \text{ 使得 } y_n\geqslant\tilde{y}(\tau_n),$$

即 $\tilde{y}(\cdot)$是(5.12)的函数 $y(\cdot)$.

(i)的证明完全是分析的，将首先处理．(5.14) 的必要性已在上面证明了，当我们注意到 $\exp[\cdot]\leqslant 1$, $H(y)\leqslant\mu$ 并应用(5.10)时，从(5.13)即可得(5.16)的必要性.

现在假设 $\tilde{y}(\cdot)$满足(5.14)和(5.16)，并由

$$r(u)y_1(u)=\int_u^U r(u)H(\tilde{y}(u))\exp\left[-\int_u^v G(\tilde{y}(v'))dv'\right]dv$$

定义 $y_1(\cdot)$．我们希望证明 $y_1(\cdot)=\tilde{y}(\cdot)$．首先注意到，根据(5.16)成立的完全相同的理由，当 $u\to U$ 时，$r(u)y_1(u)\to 0$，微分之后我们求得

$$(5.17)\quad \frac{d}{du}[r(u)y_1(u)] = -r(u)[H(\tilde{y}(u))-y_1(u)G(\tilde{y}(u))].$$

从(5.17)中减去(5.14)我们得到

$$(5.18)\quad \frac{d}{du}[r(u)(y_1(u)-\tilde{y}(u))] =G(\tilde{y}(u))r(u)[y_1(u)-\tilde{y}(u)].$$

假定在某点 u_0, $y_1=\tilde{y}$, 令 $u_1\leqslant U$ 是 u_0 之后使 $r(y_1-\tilde{y})=0$ 的第一个点；如果不存在这样的 u_1，则令 $u_1=U$．对(5.18)积分，对任何 $u_0<u<u_1$，我们有

$$(5.19)\quad \log[r(u)|y_1(u)-\tilde{y}(u)|] =\log[r(u_0)|y_1(u_0)-\tilde{y}(u_0)|]+\int_{u_0}^u G(\tilde{y}(u'))du'.$$

由于 $u\to U$ 时 $r\tilde{y}\to 0$, $ry_1\to 0$，我们看到，令 $u\to u_1$，(5.19)的左边就趋于 $-\infty$，而右边是有界的，远离 $-\infty$．这个矛盾就完成了(i)的证明.

为了证明(ii)，假设存在一个在$[U, \infty)$上等于零的函数 $\tilde{y}(u)$，使得在$[0, U)$上令 $\tilde{V}=r\tilde{y}$ 后，对 $0\leqslant u\leqslant U$ 我们有

(5.20) $\tilde{V}(u)=\int_u^{U} r(v)H(\tilde{y}(v))\exp\left[-\int_u^v G(\tilde{y}(v'))dv'\right]dv.$

定义

$$\tilde{\sigma}(u)=\text{第一个 } n \text{ 使得 } y_n r(u+\tau_n)\geqslant\tilde{V}(u+\tau_n)$$
$$(\tilde{\sigma}=\tilde{\sigma}(0)),$$

同前一样讨论，我们可看到(5.20)右边是$E[y_{\tilde{\sigma}(u)}r(u+\tau_{\tilde{\sigma}(u)})]$. 从指数分布的著名性质得出

$$(y_{N(u)+1},\ \tau_{N(u)+1}),\ (y_{N(u)+2},\ \tau_{N(u)+2}),\ \cdots$$

对给定 $\mathscr{F}_u$ 的条件联合分布是与

$$(y_1,\ u+\tau_1),\ (y_2,\ u+\tau_2),\ \cdots$$

的无条件联合分布相同的．注意到在$\{\tau_{\tilde{\sigma}}>u\}$上,

$$\tilde{\sigma}=\text{第一个 } N(u)+n \text{ 使得 } y_{N(u)+n}\geqslant\tilde{y}(\tau_{N(u)+n}),$$

我们看到在 $\{\tau_{\tilde{\sigma}}>u\}$上, $y_{\tilde{\sigma}}r(\tau_{\tilde{\sigma}})$对给定 $\mathscr{F}_u$ 的条件分布与

$$y_{\tilde{\sigma}(u)}r(u+\tau_{\tilde{\sigma}(u)})$$

的无条件分布是相同的．于是在$\{\tau_{\tilde{\sigma}}>u\}$上$\tilde{V}(u)=E(y_{\tilde{\sigma}}r(\tau_{\tilde{\sigma}})|\mathscr{F}_u)$, 或稍更一般些, 对一切 $0\leqslant u<U$, $0\leqslant v<U-u$,

(5.21) $\quad \tilde{V}(u+v)=E(y_{\tilde{\sigma}(u)}r(\tau_{\tilde{\sigma}(u)})|\mathscr{F}_v)$, 在$\{\tau_{\tilde{\sigma}(u)}>v\}$上.

现在对一切 $0\leqslant y<\infty, 0\leqslant u<U$, 定义 $\tilde{\Gamma}(y+u)=\max(yr(u),\tilde{V}(u))$. 如果对 $\{x_n(u),\ \mathscr{F}_{\tau_n}\}_1^\infty$ 使用 $\tilde{\sigma}(u)$, 那么我们的报酬在条件 $\mathscr{F}_{\tau_1}$ 下的条件期望是

$$y_1r(u+\tau_1),\ \text{若 } y_1r(u+\tau_1)\geqslant\tilde{V}(u+\tau_1),$$
$$E(y_{\tilde{\sigma}(u)}r(\tau_{\tilde{\sigma}(u)})|\mathscr{F}_{\tau_1}),\ \text{若 } y_1r(u+\tau_1)<\tilde{V}(u+\tau_1),$$

由(5.21), 它就是

$$\max(y_1r(u+\tau_1),\ \tilde{V}(u+\tau_1))=\tilde{\Gamma}(y_1,\ u+\tau_1).$$

从而 $\tilde{V}(u)=E\tilde{\Gamma}(y_1, u+\tau_1)$. 现在定义 $\tilde{\gamma}_n=\tilde{\Gamma}(y_n,\ \tau_n)\ (n=1, 2, \cdots)$, 则

$$E(\tilde{\gamma}_{n+1}|\mathscr{F}_{\tau_n})=E(\tilde{\Gamma}(y_{n+1},\ \tau_{n+1})|\mathscr{F}_{\tau_n})=\tilde{V}(\tau_n),$$

且从 $\tilde{\Gamma}$ 的定义得出

(5.22) $\tilde{\gamma}_n = \max(x_n, E(\tilde{\gamma}_{n+1} | \mathscr{F}_{\tau_n})) \quad (n=1, 2, \cdots)$.

从 $\tilde{\gamma}_n \leqslant \sum_1^{\infty} x_k$ 和(5.11)我们得知存在一个随机变量 $X \geqslant 0$, 使得 $EX < \infty$ 及

(5.23) $\tilde{\gamma}_n \leqslant E(X | \mathscr{F}_{\tau_n}) \quad (n=1, 2, \cdots)$.

最后, 因为 $x_n \to 0$ 及 $\tilde{V}(u) \to 0$, 根据(5.16)成立的同样理由我们有

(5.24) $$\tilde{\gamma}_\infty = 0.$$

从(5.22), (5.23), (5.24)和引理 4.9 推得 $\tilde{\gamma}_n \leqslant \gamma'_n$ $(n=1, 2, \cdots)$. 此外, 由定理 4.6, $\{\gamma'_n, \mathscr{F}_{\tau_n}\}_1^\infty$ 是满足(5.22)的最小随机序列; 而由定理 4.4, $(\gamma'_n) = (\gamma_n)$. 由此得到 $(\tilde{\gamma}_n) = (\gamma_n)$. 这就证明了(ii).

例子

(a) 假设 $p(u)=1$, $0 \leqslant u < U = \infty$, $r(u) = e^{-\alpha u}$, 则(5.14)变成

$$\frac{dy}{du} - \alpha y = -\varphi(y),$$

我们求得 $y(u) = y_0$, 其中 y_0 是

$$\alpha y = \varphi(y)$$

的唯一的根. (见习题 5.2.)

(b) 现在让我们说明 $U < \infty$ 的情形, 即在原来的时间尺度下, 出价次数的期望值是有限的. 设 $r=1$. 由于对(5.14)的合适的解 $y(U)=0$, 我们由积分得到

(5.25) $$\int_0^y \frac{dy'}{\varphi(y')} = U - u.$$

当 y 趋于 y_k 的分布的上界(有限或无限)时, 容易看出左边的积分是发散的. 于是存在一个数 η 使得

$$\int_0^\eta \frac{dy}{\varphi(y)} = U,$$

并且对$[0, U)$中的任一u，(5.25)有唯一的解，且它的值在$(0, \eta)$之中．这个解是临界曲线．例如，若按原来的时间尺度，$p(u') = \lambda e^{-\lambda u'}$，且若$dF(y) = \theta e^{-\theta y}dy$，易见

$$y(u') = \theta^{-1}\log(1 + e^{-\lambda u'}).$$

5. 独立情形

如果对每个$n=1, 2, \cdots$，$\mathscr{F}_n$与$\mathscr{B}(x_{n+1})$独立，则称我们处于独立情形．特别，这时$x_1, x_2, \cdots$是独立随机变量．秘书问题、停车问题和3.6节(b)的问题都是独立情形的问题．

在定理5.1中取$z_n = x_n$，我们得到（也可参阅本章引言）

定理 5.4 在独立情形，对每个$n \geqslant 1$，

(a) $\gamma'_n = \max(x_n, v'_{n+1})$，$\quad v'_n = E[\max(x_n, v'_{n+1})]$

(定义见(4.8))，

(b) $\gamma_n = \max(x_n, v_{n+1})$，$\quad v_n = E[\max(x_n, v_{n+1})]$

(定义见(4.2′))，

(c) $v_n = \sup\limits_{t \in D_n} Ex_t$，

(d) $\gamma_1, \gamma_2, \cdots$是独立随机变量．

在本节的下文中将假定我们处于独立情形且$V < \infty$．由定理5.4(b)我们有

$$\begin{aligned}\sigma &= \text{第一个 } n \text{ 使得 } x_n \geqslant v_{n+1}, \\ &= \infty, \text{若不存在这样的 } n.\end{aligned}$$

引理 5.3 $\lim\limits_{n\to\infty} \gamma_n$存在且$x_\infty = \gamma_\infty = v_\infty$．

证明 由关于尾事件的Kolmogorov 0-1律(2.2节(b))，x_∞几乎必然是常数．假如$x_\infty = -\infty$，由引理4.13，对一切

$\varepsilon>0$, $P\{x_n\geqslant v_{n+1}-\varepsilon,\ \text{i. o.}\}=1$, 且由于 $v_n\downarrow$, 我们有 $v_n\to-\infty$, 于是 $\gamma_n=\max(x_n, v_{n+1})\to-\infty$. 现在假定 $x_\infty>-\infty$. 由定理 4.7, $v_n\geqslant x_\infty(n=1, 2, \cdots)$, 从而 $v_\infty\geqslant x_\infty$, 但现在引理 4.13, 证明 $x_\infty=v_\infty$. 因为由定理 5.4, 对一切 $n=1, 2, \cdots$, $\gamma_n=\max(x_n, v_{n+1})$, 我们看出 $\lim\limits_{n\to\infty}\gamma_n$ 存在且 $\gamma_\infty=x_\infty=v_\infty$.

定理 5.5 若 $x_\infty>-\infty$, 则 σ 在 $\bar{C}$ 中最优. 若 $x_\infty=-\infty$, 且 Ex_σ 存在, 则 σ 在 C 中最优.

证明 若 $x_\infty>-\infty$, 我们有 $x_\sigma\geqslant v_{\sigma+1}\geqslant v_\infty>-\infty$, 因为由引理 5.3, $v_\infty=x_\infty$. 这样我们可以假定 Ex_σ 存在, 且由于 $V<\infty$, 故 $Ex_\sigma<\infty$. 从定理 4.10(d) 并注意到由定理 5.4, 在 $\{\sigma>n\}$ 上有 $E(\gamma_{n+1}|\mathscr{F}_n)=v_{n+1}\leqslant V(n\geqslant1)$, 即得到本定理.

定理 5.6 设 $\{\beta_n, \mathscr{F}_n\}_1^\infty$ 是任一随机序列, 使对每个 $n=1, 2, \cdots$,

(a) $E\beta_n$ 存在,

(b) $\beta_n=\max(x_n, E\beta_{n+1})$,

(c) 对一切 $\varepsilon>0$, $P\{x_k\geqslant\beta_k-\varepsilon$, 对某个 $k\geqslant n\}=1$,

还假定

(d) $x_\infty>-\infty$ 或 $E(\sup x_n^+)<\infty$,

则 $\beta_n\leqslant\gamma_n(n=1, 2, \cdots)$. 如果 $\limsup Ex_n>-\infty$, 则恰好存在一个序列 $\{\beta_n, \mathscr{F}_n\}_1^\infty$ 满足 (a), (b) 和 (c), 即 $(\beta_n)=(\gamma_n)=(\gamma_n')$.

证明 设对某个 n, $P\{\beta_n>\gamma_n\}>0$. 由条件 (b) 及定理 5.4(b) 推得 $E\beta_{n+1}>v_{n+1}$, 于是由归纳法有 $\beta_k\geqslant\gamma_k$, $k=n, n+1, \cdots$. 设 $0<\varepsilon<E\beta_n-v_n$, 令 $t=$ 第一个 $k\geqslant n$ 使得 x_k

$\geqslant\beta_k-\varepsilon$. 由(c)知 $P(t<\infty)=1$, 而由(d)知 $E(x_t)$和 $E(\beta_t)$ 存在. (若 $x_\infty>-\infty$, 我们有 $\beta_t\geqslant x_t\geqslant\beta_t-\varepsilon\geqslant\gamma_t-\varepsilon\geqslant v_\infty-\varepsilon=x_\infty-\varepsilon$.) 由引理 3.3 和(b)我们有

$$(5.26)\qquad E\beta_n=\int_{(t\leqslant k)}\beta_t+(E\beta_{k+1})P(t>k)$$

$$\leqslant\int_{(t\leqslant k)}x_t+\varepsilon+E^+(\beta_{k+1})P(t>k)$$

$$\to Ex_t+\varepsilon,\quad \text{当 } k\to\infty \text{ 时},$$

因此 $Ex_t>v_n$, 这是一个矛盾. 现在假定 $\limsup Ex_n>-\infty$. 由于 $v_n\geqslant Ex_n$ 且 $x_\infty=V_\infty$, 我们有 $x_\infty>-\infty$, 从而可应用定理的第一部分. 由定理 4.6 和 5.4 知 $\{\gamma'_n, \mathscr{F}_n\}_1^\infty$ 是满足 (a) 及(b)的最小随机序列, 所以只要证明$(\gamma_n)=(\gamma'_n)$就够了. 但是因为 $v'_n\geqslant Ex_n$ 且 $v'_n\downarrow$, 我们有 $\lim\limits_{n\to\infty} v'_n=v'_\infty>-\infty$. 因此, 若 $t\in C$, 则

$$\int_{(t>n)}(\gamma'_n)^-\leqslant(v'_\infty)^-P(t>n)\to 0,\quad \text{当 } n\to\infty \text{ 时},$$

再由定理 4.3,

$$\gamma'_n=\gamma_n\quad(n=1,\ 2,\ \cdots).$$

下面的例子表明, 如果我们只假定 $x_\infty>-\infty$, 上述定理的第二部分是不成立的.

例子

设 x_1, x_2, $\cdots$是独立的, $\mathscr{F}_n=\mathscr{B}(x_1,\ \cdots,\ x_n)$, 并假定

$$P\left\{x_n=-\prod_{k=1}^n 2^k\right\}=2^{-n}=1-P\{x_n=0\}\quad(n=1,\ 2,\ \cdots).$$

那么 $x_n\to 0$(Borel-Cantelli), 但是 $Ex_n\to-\infty$. 显然 σ=第一个 n 使得 $x_n=0$, 在 C 中是最优的. 然而由后退归纳法容易看出, 对一切 $N=1,\ 2,\ \cdots,\ s^N=1$, $Ex_{s^N}=-1$.

6. 独立情形——应用

(a) 设 $y_1, y_2, \cdots$ 独立同分布，且 $E|y_1|<\infty$. 令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$, $x_n=y_n-c_n$, 其中 (c_n) 是任一正常数序列. 如果 $c_{n+1}-c_n\uparrow$, 则由 3.6 节(a)的结果知 $V<\infty$, 我们将证明 σ 在 C 中最优, $\sigma=\lim\limits_{N\to\infty}s^N$. 事实上, 从假定 $c_{n+1}-c_n\uparrow$ 我们有 $V_{n+1}+c_n\downarrow$. 因此如同 3.6 节(a)中的讨论, σ 有任意阶矩且 $Ex_\sigma^+<\infty$. 从定理 5.5 即得 σ 在 C 中最优, 而简单应用一下定理 4.4 和引理 4.6 就证明了 $\sigma=\lim\limits_{N\to\infty}s^N$.

(b) 设 $y, y_1, y_2, \cdots$ 独立同分布, 且 $E|y|<\infty$, 对每个 $n=1, 2, \cdots$, 令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$, $x_n=y_n/n$. 由 $E|y|<\infty$ 得到: 对每一个 $\varepsilon>0$, $\sum\limits_1^\infty P\{|y|>\varepsilon n\}<\infty$, 从而由 Borel-Cantelli 引理有 $x_n\to 0$. 因此不失一般性我们可以假定 $P\{y\geqslant 0\}>0$ (不然的话, $V=0$ 且 $\sigma=+\infty$ 是最优的), 从而 $P\{y\geqslant 0\}=1$. 如果 $V<\infty$ (或由定理 4.4, 等价地如果 $E(y\log^+y)<\infty$), 则由定理 5.5, σ 在 $\bar{C}$ 中最优. 如果还有 $y\leqslant B<\infty$ 或 $a\to\infty$ 时 $E(y|y>a)=O(a)$, 则 $\sigma\in C$(见 5.8 节对这个问题在不同假定下的讨论).

为了证明最后的结论, 假定 $E(y|y>a)=O(a)$, y 有界的情形是类似的. $\sigma=$第一个 n 使得 $y_n\geqslant nv_{n+1}$. 因为对任一停止变量 t,

$$nE\left(\frac{y_t}{n+t}\right)\leqslant(n+1)E\left(\frac{y_t}{n+1+t}\right),$$

所以 nv_{n+1} 是递增的. 由于 $x_\infty=0$ 且 σ 在 $\bar{C}$ 中最优, 我们有

(5.27) $$V=\int_{\{\sigma\leqslant n\}}y_\sigma/\sigma+\int_{\{n<\sigma<\infty\}}y_\sigma/\sigma \qquad (n=1, 2, \cdots).$$

从(5.26)及(5.27)推得对每个n,

(5.28) $$v_{n+1}P(\sigma>n)=\int_{\{n<\sigma<\infty\}}y_\sigma/\sigma.$$

现在

$$\begin{aligned}\int_{\{n<\sigma<\infty\}}y_\sigma/\sigma&=\sum_{k=n+1}^{\infty}P(\sigma\geqslant k)\int_{\{y_k\geqslant kv_{k+1}\}}y_k/k\\&=\sum_{k=n+1}^{\infty}v_{k+1}P(\sigma=k)\frac{E(y\,|\,y\geqslant kv_{k+1})}{kv_{k+1}}\\&\leqslant v_{n+1}\sup_k\frac{E(y\,|\,y\geqslant kv_{k+1})}{kv_{k+1}}P\quad(n<\sigma<\infty).\end{aligned}$$

因为$kv_{k+1}\uparrow$(从而有不为0的下界),由假定$E(y\,|\,y>a)=O(a)$,我们有

(5.29) $$\int_{\{n<\sigma<\infty\}}y_\sigma/\sigma\leqslant\text{const}\,v_{n+1}P\quad(n<\sigma<\infty).$$

由(5.28)及(5.29)推得$n\to\infty$时$P\{\sigma>n\}\to0$,所以$\sigma\in C$.

7. 均匀博弈

设$y_1, y_2, \cdots$独立,在(0, 1)上均匀分布,$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$,对某个$a>0$,设我们在时刻$n$停止而遭受的损失是$x_n=n^a y_n(n=1, 2, \cdots)$. 我们感兴趣的是找使$Ex_t$最小的停止变量$t$.

定理5.7

(i) 对任何$0<a\leqslant1$, $V=0$且$\sigma=+\infty$在$\bar{C}$中最优.

(ii) 存在一个最小的数$a^*>1$,使对任一$a\geqslant a^*$, $V=1/2$且$\sigma=1$在C中最优.

(iii) 对任何$1<a<a^*$,

(a) $0<V<\frac{1}{2}$, (b) σ在C中最优,

(c) $E\sigma=\infty$, (d) $\sigma=\lim_{N\to\infty}s^N$,

(e) $v_n \sim 2(\alpha-1)n^{1-\alpha} \quad (n\to\infty)$.

(iv) V 是 α 的连续函数.

证明　设 $A>0$ 是任意的,

$$\sum^{\infty} P(n^{\alpha}y_n \leqslant A) = \sum^{\infty} n^{-\alpha}A. \tag{5.30}$$

因此对任一 $\alpha\leqslant 1$, 由 Borel-Cantelli 引理, $P\{x_n\leqslant A, \text{i.o.}\}=1$, 由于 A 是任意的, 所以 $x_\infty=0$. 这就证明了(i).

令 $u_n=\inf\limits_{k>n} x_k$, 对任何 $0\leqslant A<n^{\alpha}$,

$$P\{u_n\geqslant A\}=\prod_{k=n}^{\infty}\left(1-\frac{A}{k^{\alpha}}\right). \tag{5.31}$$

因此 $P\left\{u_1\geqslant \frac{1}{2}\right\}=\prod_{1}^{\infty}\left(1-\frac{1}{2n^{\alpha}}\right)>0 \ (\alpha>1)$, 所以

$$V\geqslant Eu_1>0 \quad (\alpha>1). \tag{5.32}$$

由(5.30)和 Borel-Cantelli 引理, 我们看到, 对一切 $\alpha>1$, $x_\infty=\infty$, 从而由定理 4.5, σ 在 C 中是最优的.

为了完成(ii)—(iv)的证明, 我们从定理 5.4(b)的函数方程得(v_n)的界限, 这时这函数方程变成

$$v_n=\int_0^1 \min(n^{\alpha}y,\ v_{n+1})dy, \tag{5.33}$$

从(5.33)有

$$v_n\leqslant \frac{1}{2}n^{\alpha}, \tag{5.34}$$

因此

$$v_{n+1}n^{-\alpha}\leqslant \frac{1}{2}\left(\frac{n+1}{n}\right)^{\alpha}\leqslant \frac{1}{2}\left(\frac{3}{2}\right)^{\alpha}<1 \tag{5.35}$$

$$\left(n\geqslant 2,\ 1<\alpha<\frac{3}{2}\right).$$

于是(5.33)可以写成

$$(5.36)\qquad v_n=v_{n+1}(1-v_{n+1}/2n^{\alpha})\quad \left(n\geqslant 2,\ 1<\alpha\leqslant\frac{3}{2}\right),$$

因为对一切 $n=1, 2, \cdots, \alpha>1$ 有 $v_n>0$，我们可以把(5.36)改写成

$$(5.37)\qquad \frac{1}{v_n}=\frac{1}{v_{n+1}}+\frac{1}{2n^{\alpha}-v_{n+1}}\quad \left(n\geqslant 2,\ 1<\alpha\leqslant\frac{3}{2}\right),$$

把它对 $k=n, n+1, \cdots, m$ 求和就成为

$$(5.38)\qquad \frac{1}{v_n}=\sum_{k=n}^{m}\frac{1}{2k^{\alpha}-v_{k+1}}+\frac{1}{v_{m+1}}\quad \left(n\geqslant 2,\ 1<\alpha\leqslant\frac{3}{2}\right).$$

因为由引理5.3知对一切 $\alpha>1$ 有 $v_{\infty}=x_{\infty}=\infty$，我们可以在(5.38)中令 $m\to\infty$ 而得到

$$(5.39)\qquad \frac{1}{v_n}=\sum_{k=n}^{\infty}\frac{1}{2k^{\alpha}-v_{k+1}}\quad \left(n\geqslant 2,\ 1<\alpha\leqslant\frac{3}{2}\right).$$

从(5.39)和(5.35)我们得到

$$(5.40)\qquad \frac{1}{2(\alpha-1)(n+1)^{\alpha-1}}\leqslant\frac{1}{2}\int^{\infty} y^{-\alpha}dy\leqslant\frac{1}{2}\sum^{\infty}k^{-\alpha}$$

(5.43) $$v_4 \leqslant 1.52,$$
$$v_5 = 1.7,$$

另一方面由(5.41),

$$\frac{v_{n+1}}{n^{\alpha}} \leqslant \frac{1}{n+1}\left(\frac{n+1}{n}\right)^{3/2} \leqslant \frac{4}{50} \quad (n \geqslant 5).$$

从而由(5.39), 其中 α 取 $\frac{3}{2}$, 得到

$$\frac{1}{v_5} = \sum_5^{\infty} \frac{1}{2k^{\alpha} - v_{k+1}} \leqslant \sum_5^{\infty} \frac{1}{2k^{\alpha}\left(1 - \frac{11}{100}\right)}$$

$$\leqslant \frac{50}{89} \int_{9/2}^{\infty} t^{-\alpha} dt < \frac{1}{1.7},$$

这与(5.43)矛盾. 所以, 若 $\alpha \geqslant 3/2$, 则 $V = 1/2$ 及 $\sigma = 1$.

由(5.41), $\alpha \to 1$ 时 $V \to 0$. 设 α 是任一 >1 的数, 且对这个 α, $V < 1/2$. 由 (5.33), $P\{\sigma > 1\} > 0$, 因为

$$\sigma = \inf\{n : y_n \leqslant v_{n+1}/n^{\alpha}\},$$

由(5.35)得出, 对每一个 N, $P\{\sigma > N\} > 0$. 此外, 如果对某个 $\varepsilon > 0$, $1 < \alpha = \frac{3-\varepsilon}{2}$, 则由(5.41)我们有

$$v_{n+1}/n^{\alpha} \leqslant (1-\varepsilon)\left(\frac{n+1}{n^2}\right) \leqslant \frac{1}{n}, \quad \text{对 } n \geqslant (1-\varepsilon)/\varepsilon.$$

因此若 $V < \frac{1}{2}$, 则 σ 是无界的, 对任一 $n > N \geqslant (1-\varepsilon)/\varepsilon$ 及某个 $K > 0$, 我们有

$$P\{\sigma > n\} \geqslant K(1 - 1/N)\left(1 - \frac{1}{N+1}\right) \cdots (1 - 1/n)$$
$$= K \cdot \frac{N-1}{n},$$

所以 $E\sigma = \sum_0^{\infty} P\{\sigma > n\} = \infty$. 因此对任一 $\alpha > 1$, 或 $V = 1/2$ 且 $\sigma = 1$, 或 $0 < V < 1/2$, $P\{\sigma < \infty\} = 1$ 但 $E\sigma = \infty$.

为了证明(iii)(e)，我们首先注意到对任一使$V<\frac{1}{2}$的$1<\alpha<\frac{3}{2}$，由(5.41)有$\frac{v_{n+1}}{n^\alpha}\to 0(n\to\infty)$．因此对任一$\varepsilon>0$，对一切充分大的$k$，$v_{k+1}\leqslant \varepsilon k^\alpha$，于是由(5.39)，对一切充分大的$n$，我们有

$$\frac{1}{v_n}=\sum_{k=n}^\infty\frac{1}{2k^\alpha-v_{k+1}}\leqslant\frac{1}{2-\varepsilon}\sum_n^\infty k^{-\alpha}\leqslant\frac{1}{2-\varepsilon}\int_{n-1}^\infty x^{-\alpha}dx$$
$$=\frac{1}{(2-\varepsilon)(\alpha-1)(n-1)^{\alpha-1}}.$$

故得$\liminf\limits_{n\to\infty}\frac{v_n}{n^{\alpha-1}}\geqslant(2-\varepsilon)(\alpha-1)$，由于$\varepsilon$是任意的，由此并连同(5.41)就给出了(iii)(e)．

现在我们转向(iii)(d)及(iv)的证明((iv)证明了对$\alpha=\alpha^*$有$V=\frac{1}{2}$，故它转过来也完成了(ii)的证明)．用与推导(5.41)相同的论证可证明

(5.44) $$\frac{\alpha-1}{n}\leqslant\frac{v'_{n+1}}{n^\alpha}\leqslant\frac{2(\alpha-1)}{n+1}\left(\frac{n+1}{n}\right)^\alpha$$
$$\left(n\geqslant 1,\ 1<\alpha\leqslant\frac{3}{2}\right).$$

由定理4.10及(5.41)我们有

$$(\alpha-1)n^{\alpha-1}P\{\sigma>n\}\leqslant v_{n+1}P\{\sigma>n\}$$
$$=\int_{(\sigma>n)}\gamma_n\to 0,\text{当 }n\to\infty\text{ 时}.$$

于是由(5.44)，当$n\to\infty$时，

$$\int_{(\sigma>n)}\gamma'_n=v'_{n+1}P\{\sigma>n\}\leqslant\text{const}\,n^{\alpha-1}P\{\sigma>n\}\to 0,$$

并由定理4.3后面的注和引理4.6我们有$V'=V$，$\sigma=\lim\limits_{N\to\infty}s^N$．

V在$\alpha>1$的右连续性现在可从定理4.9(b)得到(在$\alpha=1$的

右连续性由(5.41)立即可得)，V在任一$\alpha>1$的左连续性从定理 4.9(a)得到．这就完成了证明．

8. y_n/n 的问题

设$y_1, y_2, \cdots$是独立同分布非负随机变量，且$Ey_1=1$. 令$\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$及$x_n=y_n/n(n=1, 2, \cdots)$. 在5.6节(b)中曾经证明了，如果$E(y_1\log^+y_1)<\infty$和$a\to\infty$时$E(y_1|y_1>a)=O(a)$，则$\sigma$在$C$中最优．在这一节中我们要在对某个$\alpha>1$,

$$Ey_1^\alpha<\infty \tag{5.45}$$

的假设下证明同样的结果.

首先注意到(5.45)蕴含着

$$v_n\leqslant E(\sup_{k\geqslant n} x_k)\leqslant (E(\sup_{k\geqslant n} x_k^\alpha))^{1/\alpha}$$
$$\leqslant \text{const}\left(\sum_n^\infty k^{-\alpha}\right)^{1/\alpha}=O(n^{-1+1/\alpha}), \tag{5.46}$$

因此由定理 4.5′，σ在$\bar{C}$中是最优的．为了证明$\sigma\in C$，只要证明$P\{\sigma=\infty\}>0$蕴含

$$\frac{v_{2n}}{v_n}\to 1, \quad 当 n\to\infty 时. \tag{5.47}$$

因为如果(5.47)成立，则对任意的$\varepsilon>0$，由(5.46)对一切充分大的m和一切$n=1, 2, \cdots$，我们有

$$\text{const}(2^n m)^{-1+1/\alpha}\geqslant v_{2^n m}\geqslant (1-\varepsilon)^n v_m,$$

从而

$$\log v_m+n\log(1-\varepsilon)\leqslant(1/\alpha-1)(n\log 2+\log m)+O(1).$$

令$n\to\infty$，我们得到

$$\log(1-\varepsilon)\leqslant(1/\alpha-1)\log 2.$$

对充分小的ε这是一个矛盾.

于是假定 $P\{\sigma=\infty\}>0$. 因为 $x_\infty=0$, 把定理 4.5′ 应用于随机序列 $\{x_k, \mathscr{F}_k\}_n^\infty$ 就证明了对每个 $n=1, 2, \cdots$,

$$(5.48)\qquad v_n=\sum_{i=0}^{\infty}(n+i)^{-1}\int_{\{\sigma_n=n+i\}}y_{n+i}.$$

令

$$(5.49)\qquad p_n=P\{y_n<nv_{n+1}\}.$$

由定理 5.4 和 σ_n 的定义我们有

$$P\{\sigma_n=n+i\}=p_np_{n+1}\cdots p_{n+i-1}(1-p_{n+i}),$$

且(5.48)变成

$$(5.50)\quad v_n=\sum_{i=0}^{\infty}(p_n\cdots p_{n+i-1})(n+i)^{-1}\int_{\{y_{n+i}\geqslant(n+i)v_{n+i+1}\}}y_{n+i}.$$

现在定义一个新的停止规则 t 为

$$(5.51)\qquad t=\inf\{m: m\geqslant 2n,\ y_m\geqslant[m/2]v_{[m/2]+1}\},$$

其中 $[x]$ 记 x 的整数部分, 则 $t\in\bar{C}_{2n}$, 于是

$$(5.52)\qquad v_{2n}\geqslant Ex_t=\sum_{m=2n}^{\infty}m^{-1}\int_{\{t=m\}}y_m.$$

但是由(5.49), (5.51)和诸 y_n 的独立性, 对每个 $i=0, 1, \cdots$ 我们有

$$\int_{\{t=2n+2i\}}y_{2n+2i}=p_n^2\cdots p_{n+i-1}^2\int_{\{y_{2n+2i}\geqslant(n+i)v_{n+i+1}\}}y_{2n+2i},$$

$$\int_{\{t=2n+2i+1\}}y_{2n+2i+1}$$

$$=p_n^2\cdots p_{n+i-1}^2p_{n+i}\int_{\{y_{2n+2i+1}\geqslant(n+i)v_{n+i+1}\}}y_{2n+2i+1}.$$

因为诸 y_n 有共同的分布, 最后两个方程右边的积分每一个都等于

$$\int_{\{y_{n+i}\geqslant(n+i)v_{n+i+1}\}}y_{n+i}.$$

代入(5.52)我们得到

$$(5.53)\quad Ex_t=\sum_{i=0}^{\infty}\left(p_n^2\cdots p_{n+i-1}^2(2n+2i)^{-1}\right.$$
$$\left.+p_n^2\cdots p_{n+i-1}^2 p_{n+i}(2n+2i+1)^{-1}\int_{(y_{n+i}\geqslant(n+i)v_{n+i+1})} y_{n+i}\right).$$

由这个等式以及(5.50)和(5.52)，我们得到

$$(5.54)\quad \frac{v_{2n}}{v_n}\geqslant\inf_{i\geqslant0}\left(p_n\cdots p_{n+i-1}(2n+2i)^{-1}\right.$$
$$\left.+p_n\cdots p_{n+i}(2n+2i+1)^{-1}\right)/(n+i)^{-1}$$
$$\geqslant\left(\prod_{i=0}^{\infty}p_{n+i}\right)\inf_{i\geqslant0}\frac{(n+i)^2}{(n+i+1/2)^2}.$$

但是 $\prod_{n=1}^{\infty}p_n=P\{\sigma=\infty\}$，于是如果 $P\{\sigma=\infty\}>0$，则(5.54)的右边当 $n\to\infty$ 时趋于1. 这就得到(5.47)，完成了证明.

习　题

1. 完成第三章的习题 2，即证明 σ 是最优的并且具有性质：对某个正数 a，σ＝第一个 n 使得 $y_1+\cdots+y_n\geqslant a$.
2. 在 5.4 节的问题中，假定 $p(u)=1$，$0\leqslant u<U=\infty$；$r(u)=e^{-\alpha u}$，寻找并求解一个等价的单调情形的问题.
3. 证明：如果 $c_{n+1}-c_n$ 是递减的，3.6 节(a)和 5.6 节(a)的问题实际上是相同的. 这一点对于用后退归纳法及取极限来计算 σ 这一意图有什么意义？
4. 直接证明：恰好存在一个趋于 ∞ 并满足(5.39)的常数序列. 这给出了 5.7 节的(ii)(d)和(iii)的另一个证明.
5. 考虑 3.1 节(g)和 5.2 节(a)的问题. 描述最优规则 σ，如果抽样费用依赖于真正的密度，即如果
$$-x_n=\min(a\pi_n,\ b\bar{\pi}_n)+c_0n\pi_n+c_1n\bar{\pi}_n\quad(\bar{\pi}_n=1-\pi_n).$$
6. (续)沿着 5.2 节(c)的思路把这个问题表达成一个 $(\Omega,\ \mathscr{F},\ P_1)$ 上的最优停止问题.
7. (续)当 $c_0\to0$，$b\to\infty$ 时这个问题与 5.2 节(c)的问题有什么关系？
8. 证明 5.2 节(d)的引理.

9. 叙述并证明定理 4.5′ 的一个适用于 2(d) 的翻版.

10. 把习题 8 与 9 的结果应用于 5.2 节(d)的最优停止问题.

11. 设 $y_1, y_2, \cdots$ 是独立的，且 $P(y_k=1)=p=1-P(y_k=0)$ $(k=1, 2, \cdots)$. 对某个 $0<c<1$, 令

$$x_n=\sum_1^n I_{(y_k=y_{k+1}=\cdots=y_n=1)}-cn \text{ 和 } \mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n),$$

(出现在 x_n 的定义中的和表示从第 n 个观察向后计算 1 个游程的长度.)对 $\{x_n, \mathscr{F}_n\}_1^\infty$ 找一个最优规则. V 是什么? (这个问题是 N. Starr 提出的.)

12. (见 Taylor[1]) 设 $y_1, y_2, \cdots$ 独立同分布，且 $-\infty<Ey_1<0$, $E[(y_1^+)^2]<\infty$. 令 $\mathscr{F}_n=\mathscr{B}(y_1, \cdots, y_n)$ 和 $x_n=\left(\sum_1^n y_k\right)^+$ (由定理 4.5′ 和定理 4.13, 对$\{x_n, \mathscr{F}_n\}_1^\infty$ σ 在 $\bar{C}$ 中最优). 证明:

$\sigma=$第一个 $n\geqslant 1$ 使得 $x_n\geqslant q^{-1}Ex_\tau$,

$=\infty$, 若不存在这样的 n,

其中

$$\tau=\text{第一个 } n\geqslant 1 \text{ 使得 } x_n>0,$$
$$=\infty, \text{若不存在这样的 } n,$$

及 $q=P(\tau<\infty)$. (提示: 这个问题可以直接利用定理 5.2 来解或者可归结为第三章的窃贼问题——见习题 3.2 及习题 5.1.)

文献注释

第 一 章

这一章包含概率论的测度论基础的扼要描述，并且是按 Kolmogorov[1]所提出的方式叙述的．更详细的处理(例如)可参阅 Loève[1]．引理 1.2 比通常 Fatou 引理的形式略为更一般些．这个推广用在第四章中(例如引理 4.8，定理 4.10)．定理 1.4 属于 Lévy[1]，129 页．

第 二 章

鞅论的基本结果，特别是定理 2.1 及 2.2，属于 Doob[1]，[2]．定理 2.3 的(b) 及(c) 由 Chow, Robbins, Teicher[1]证明．利用鞅得到 L. Takács 的投票定理的想法(4.4 节(c))属于 G. Simons．

2.5 节取自 Siegmund[1]，[4]，虽然基本的思想可在概率与统计的近期文献中找到(例如参阅 Chow, Robbins[6]和 Gundy, Siegmund[1])．

围绕 2.6 及 2.7 节的一系列想法源自 Wald (例如[1])．测度 Q 的使用属于 Bahadur[1]．

第 三 章

定理 3.1 及引理 3.1 为 Robbins, Samuel[1]所证．

序贯概率比检验的 Bayes 性质（例如 3.1 节 (h)）最初由 Wald, Wolfowitz[1] 和 Arrow, Blackwell, Girshick[1] 所讨论，较近则见 Chow, Robbins[3] 和 Siegmund[2]. Arrow 等认出了问题的最优停止的一面，他们用有限情形逼近的方法解决了问题. 特别，定理 3.2 源自他们.

秘书问题（3.1 节例 (f)）已被许多作者研究过. Gilbert 及 Mosteller 的文章[1] 包含一个收罗了这方面许多工作的文献索引以及有关这问题的各种变形的讨论. Chow, Moriguti, Robbins, Samuel[1] 研究了使选中的姑娘的平均名次最小的问题，证明了

$$\lim_{N\to\infty} V^N = \prod_{j=1}^{\infty}\left(\frac{j+2}{j}\right)^{1/(j+1)} \cong 3.8695.$$

3.4 节的基本结果属于 Chow, Robbins[2]，他们分离出并研究了单调情形（特别，3.6 节 (a) 的问题）. Derman, Sacks [1] 独立地用类似的方法研究了一个类似的问题. 把引理 3.1 应用于 3.6 节 (a) 的问题的想法属于 D. Burdick. 对这问题以及它的许多变形作出贡献的还包括 MacQueen, Miller[1], Sakaguchi[1], Bramblett[1], Yahav[1], Cohn[1] 和 DeGroot[1].

3.6 节 (c) 及 (d) 的问题分别为 Wald [2] 和 Mallows, Robbins[1] 所研究.

第　四　章

第四章的一般理论大部取自 Chow, Robbins[3], [5] 和 Siegmund [2]. Arrow 等[1], Snell[1] 和 Haggstrom[1] 也研究了最优停止理论的基础，在这些作者的工作中可找到我们的几个主要定理的描述.

在定理 4.5 的条件下最优规则的存在性为 Snell[1] 所证，他推广了 Arrow 等[1]的一个早期结果. Chow, Robbins[1] 给了这结果一个不同的证明. Haggstrom [1] 和 Chow, Robbins[1] 独立地认出最优规则是 σ. 定理 4.5′ 属于 Siegmund[2]. 引理 4.6 及其系统的应用是新的.

引理 4.7 为 Dvoretzky 所证（参阅 Bramblett[1]）. 这里给出的证明是新的.

4.3 节 (e) 的问题在 Shiryaev（参阅 [1], [2]）的一系列文章中作了研究.

上鞅的正则性概念（4.4 节）属于 Snell[1], 这概念成了他研究最优停止问题的基础.

定理 4.10 是新的.

对 $\frac{s_n}{n}$ 的最优停止问题首先为 Chow, Robbins[4] 所研究，他们证明了当 y_n 各以 $\frac{1}{2}$ 的概率取值 ± 1 时存在最优停止规则. Dvoretzky 把这结果推广到 (y_n) 为独立同分布，具有零均值及有限二阶矩的情形. 这里给出的讨论借用自 Siegmund, Simons, Feder[1] 和 Ruiz-Moncayo[1]. 最近 Shepp[1] 得到了关于最优停止规则 σ 的更明确的信息，他把这问题与 Wiener 过程的类似问题相联系，而对后者可以得到精确解.

定理 4.12（对一切 n, $\alpha_n \geqslant 0$ 的情形）及其推论属于 Samuel[1]. 定理 4.13 可从 Chow, Robbins[2], Kiefer, Wolfowitz[1] 等工作，也可从 Darling, Erdös, Kakutani 的未发表的工作提取出. 这里的证明借用了 Erdös[1] 的方法. 定理 4.14 中 (a) 包含 (b) 的证明属于 Doob[1]. 定理的逆命题部分独立地由 B.Davis[1] 和 McCabe, Shepp[1] 得到.

第 五 章

Markov 情形的各种想法隐含在许多作者的工作中，这里给出的讨论取自 Siegmund[2]，他也证明了密切有关的定理 5.3(也参阅 Dynkin, Yushkevich[1])。

5.4 节的问题为 Elfving[1]所研究，他假设存在一个用分段连续函数 $y(\cdot)$ 定义的最优规则，推导出基本积分方程(5.13)，并证明了 (i)．Siegmund[1]去掉了 Elfving 的假设．

独立情形首先由 Chow, Robbins[5]，后来由 Chow[1]和 Siegmund [3]研究．

5.7 节的结果取自 Chow, Robbins[1]，5.8 节的结果取自 Chow, Dvoretzky[1]。

参考文献

Arrow, K. J., Blackwell, D., Girshick, M. A.

[1] Bayes and minimax solutions of sequential decision problems, Econometrica, **17**(1949), 213～244.

Bahadur, R. R.

[1] A note on the fundamental identity of sequential analysis, Ann. Math. Statist. **29**(1958), 534—543.

Bellman, Richard

[1] A problem in the sequential design of experiments, Sankhya **16** (1956), 221—229.

Bramblett, J. E.

[1] Some approximations to optimal stopping procedures, Ph. D. dissertation (unpublished), New York: Columbia University, 1965.

Breiman, L.

[1] "Stopping-rule problems" in Applied Combinatorial Mathematics, E. F. Beckenbach (Editor), New York: John Wiley & Sons, Inc., 1964.

Burkholder, D. L., Gundy, R. F.

[1] Extrapolation and interpolation of quasi-linear operators on martingales, Acta Mathematica **24**(1970), 249—304.

Chernoff, H.

[1] A note on risk and maximal regular generalized submartingales in stopping problems, Ann. Matt. Statist. **38**(1967), 606—607.

[2] Optimal stochastic control, Sankhyā, Series A, **30**(1968), 221—252.

Chow, Y. S.

[1] On optimal stopping rules for independent random variables, mimeographed, Lafayette, Ind.: Purdue University, 1966.

Chow, Y. S., Dvoretzky, A.

[1] Stopping rules for x_n/n and related problems, Stanford, Calif.: Stanford Univ. Tech. Report, 1969.

Chow, Y. S., Moriguti, S., Robbins, H., Samuels, S. M.

[1] Optimal selection based on relative rank, Israel Journal of Mathematics **2**(1964), 81—90.

Chow, Y. S., Robbins, H.

[1] A class of optimal stopping problems, Proc. Fifth Berkeley Symposium Math. Statist. Prob., Vol. 1, Berkeley and Los Angeles: University of California Press, 1967, 419—426.
[2] A martingale system theorem and applications, Proc. Fourth Berkeley Symposium Math. Statist. Prob., Vol. 1, Berkeley and Los Angeles: University of California Press, 1961, 93—104.
[3] On optimal stopping rules, Z. Wahrscheinlichkeitstheorie **2**(1963), 33—49.
[4] On optimal stopping rules for S_n/n, Illinois J. Math. **9**(1965), 444—454.
[5] On values associated with a stochastic sequence, Proc. Fifth Berkeley Symposium Math. Statist. Prob., Vol. 1, Berkeley and Los Angeles: University of California Press, 1967, 427—440.
[6] A renewal theorem for random variables which are dependent or non-identically distributed, Ann. Math. Statist. **36**(1963), 457—462.

Chow, Y. S., Robbins, H., Teicher, H.
[1] Moments of randomly stopped sums, Ann. Math. Statist. **36**(1965), 789—799.

Chung, K. L., Fuchs, W. H. J.
[1] On the distribution of values of sums of random variables, Memoirs Amer. Math. Soc. **6**(1950).

Cohn, H.
[1] On certain optimal stopping rules, Rev. Roumaine Math. Pures Appl. **12**(1967), 1173—1177.

Davis, B.
[1] Stopping rules for S_n/n and the class $L\log L$, New Brunswick, N. J.: Rutgers Univ. Tech. Report, 1969.

De Groot, M.
[1] Some problems of optimal stopping, Roy. Stat. Soc. B. **30**(1968), 108—122.
[2] Optimal Statistical Decisions, New York: McGraw-Hill, Inc., 1970.

Derman, C., Sacks, J.
[1] Replacement of periodically inspected equipment (an optimal stopping rule), Naval Research Logistics Quarterly **7**(1960), 597—607.

Doob, J. L.

[1] Regularity properties of certain families of chance variables, Trans. Amer. Math. Soc. **47** (1940), 455—486.

[2] Stochastic Processes, New York: John Wiley & Sons, Inc., 1953.

Dubins, L. E., Teicher, H.

[1] Optimal stopping when the future is discounted, Ann. Math. Statist. **38** (1967), 601—605.

Dvoretzky, A.

[1] Existence and properties of certain optimal stopping rules, Proc. Fifth Berkeley Symposium Math. Statist. Prob., Berkeley and Los Angeles: University of California Press, 1967, 441—452.

Dynkin, E. B.

[1] The optimum choice of the instant for stopping a Markov process, Doklady Akad. Nauk SSS **150** (1963), 238—240; Soviet Math. **4** (1963), 627—629.

Dynkin, E.B., Yushkevich, A. A.

[1] Markov Processes, Theorems and Problems, New York: Plenum Press, 1969.

Elfving, G.

[1] A persistency problem connected with a point process, J. Appl. Prob. **4** (1967), 77—89.

Erdös, P.

[1] On a theorem of Hsu and Robbins, Ann. Math. Statist. **20** (1949) 286—291.

Ferguson, T.

[1] Mathematical Statistics, a Decision Theoretic Approach, New York: Academic Press, Inc., 1967.

Gilbert, J., Mosteller, F.

[1] Recognizing the maximum of a sequence, Amer. Stat. Assoc. **61** (1966), 35—73.

Gundy, R. F., Siegmund, D.

[1] On a stopping rule and the central limit theorem, Ann. Math. Statist. **38** (1967), 1915—1917.

Haggstrom, G.

[1] Optimal stopping and experimental design, Ann. Math. Statist. **37** (1966), 7—29.

Kiefer, J., Wolfowitz, J.
[1] On the theory of queues with many servers, Trans. Amer. Math. Soc. 78(1955), 1—18.
Kolmogorov, A. N.
[1] Grundbegriffe der Wahrscheinlichkeitsrechnung, Ergebnisse der Mathematik 2(1933), No. 3. (English translation: Foundations of Probability Theory, New York: Chelsea, 1955.)
Lehmann, E. L.
[1] Testing Statistical Hypotheses, New York: John Wiley & Sons, Inc., 1959, 105.
Lévy, P.
[1] Theorie de l'Addition des Variables Aléatories, Paris: Gauthier-Villars, 1937.
Loéve, M.
[1] Probability Theory, 3rd ed. Princeton, N. J.: D. Van Nostrand Co., Inc., 1963.
McCabe, B. J., Shepp, L. A.
[1] On the supremum of S_n/n, Ann. Math. Statist. 41(1970), 2166—2168.
MacQueen, J. Miller, R. G., Jr.
[1] Optimal persistence policies, Operat. Res. 8(1960), 362—380.
Mallows, C., Robbins, H.
[1] Some problems of optimal sampling strategy, J. Math. Anal. Appl. 8(1964), 90—103.
Neveu, J.
[1] Mathematical Foundations of the Calculus of Probability, San Francisco: Holden-Day, 1965, 44.
Robbins, H. E., Samuel, E.
[1] An extension of a lemma of Wald, J. Appl. Prob. 3(1966), 272—273.
Ruiz-Moncayo, A.
[1] Optimal stopping for functions of Markov chains, Ann. Math. Statist. 39(1968), 1905—1912.
Sakaguchi, M.
[1] Dynamic programming of some sequential sampling designs, J. Math. Anal. Appl. 2(1961), 446—466.
Samuel, E.

[1] On the existence of the expectation of randomly stopped sums, J Appl. Prob. 4(1967), 197—200.

Shepp, L.

[1] Explicit solutions to some problems of optimal stopping, Ann. Math. Statist. 40(1969), 993—1010.

Shiryaev, A. N.

[1] On Markov sufficient statistics in non-additive Bayes problems of sequential analysis, Theory Prob. Appl. 9(1964), 604—618.

[2] On optimal methods in quickest detection problems, Theory Prob. Appl. 8(1963), 22—46.

[3] Statistical Sequential Analysis, Moscow, 1969. (In Russian.)

Siegmund, D.

[1] Some one-sided stopping rules, Ann. Math. Statist. 38(1967), 1641—1646.

[2] Some problems in the theory of optimal stopping rules, Ann. Math. Statist. 38 (1967), 1627—1640.

[3] Some problems in the theory of optimal stopping rules, mimeographed, New York: Columbia University, 1966.

[4] The variance of one-sided stopping rules, Ann. Math. Statist. 40 (1969), 1074—1077.

Siegmund, D., Simons, G., Feder, P.

[1] Existence of optimal stopping rules for rewards to s_n/n, Ann. Math. Statist. 39(1968), 1228—1235.

Snell, J. L.

[1] Application of martingale system theorems, Trans. Amer. Math. Soc. 73(1952), 293—312.

Taylor, H.

[1] Bounds on the expected value of the positive part of a stopped random sum, unpublished manuscript. 1970.

Teicher, H., Wolfowitz, J.

[1] Existence of optimal stopping rules for linear and quadratic rewards, Z. Wahrscheinlichkeitstheorie 5(1966), 316—368.

von Bahr, B., Esseen, C. G.

[1] Inequalities for the rth absolute moment of a sum of random variables, $1\leqslant r\leqslant 2$, Ann. Math. Statist. 36(1965), 299—303.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19**(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75**(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37**(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37** (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19**(1948), 326—339.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75**(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37**(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19 (1948), 326—339.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75 (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37 (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37** (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19 (1948), 326—339.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75 (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37 (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statiatical. **37** (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37** (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37** (1966), 30—35.

Wald, A.
[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.
[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.
Wald, A., Wolfowitz, J.
[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.
Walker, L. H.
[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.
Yahav, J. A.
[1] On optimal stopping, Ann. Math. Statist. **37** (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—339.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37** (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. **19** (1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. **75** (1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. **37** (1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald,A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—339.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.

Wald, A.

[1] Sequential Analysis, New York: John Wiley & Sons, Inc., 1947.

[2] Statistical Decision Functions, New York: John Wiley & Sons, Inc., 1950.

Wald, A., Wolfowitz, J.

[1] Optimum character of the sequential probability ratio test, Ann. Math. Statist. 19(1948), 326—329.

Walker, L. H.

[1] Regarding stopping rules for Brownian motion and random walks, Bull. Amer. Math. Soc. 75(1969), 46—50.

Yahav, J. A.

[1] On optimal stopping, Ann. Math. Statist. 37(1966), 30—35.