理学硕士学位论文

相依误差下线性模型的经验似然推断

马 昀 蓓

北京工业大学

2006年5月

分类号： O212.7 单位代码： 10005

学 号： S200306020

北京工业大学硕士学位论文

题 目 相依误差下线性模型的经验似然推断

英文并列题目 Empirical Likelihood Inference for Linear Models Based on Dependent Errors

研究生姓名： 马 昀 蓓

专 业： 概率论与数理统计 **研究方向：** 非参数统计

导师姓名： 薛 留 根 **职 称：** 教 授

论文报告提交日期 2006 年 5 月 学位授予日期

授予单位名称和地址 北京工业大学 北京市朝阳区平乐园 100 号

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表和撰写过的研究成果，也不包含为获得北京工业大学或其他教育机构的学位或证书而使用过的材料，与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签名：_____ 日期：_____

关于论文使用授权的说明

本人完全了解北京工业大学有关保留、使用学位论文的规定，即：学校有权保留送交论文的复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。

（保密的论文在解密后应遵守此规定）

签名：_____ 导师签名：_____ 日期：_____

摘　要

在一些近代科学研究中，如生命科学和信息科学的研究，人们获得的数据往往具有量大、高维和相依的特点．它是目前统计学应用和理论中面临最多、挑战最严峻，也是最有可能取得突破的研究领域之一．于是，关于相依随机变量的研究已引起人们的重视．然而，对各种具体的相依数据的统计模型的研究还不够充分和重视．本文研究了基于两种具体的相依误差的线性模型，一种是鞅差序列，另一种则是弱平稳线性过程．

对于参数分量 β, 文献中一般用最小二乘法来估计，得到该估计的渐近性质．本文则采用了新的推断方法 — 经验似然推断．由于经验似然方法的诸多优点，例如，构造的置信区间具有变换不变性，置信域的形状由数据自行决定以及 Bartlett 纠偏性等等．因此这一方法无论在理论上还是实际应用中都非常有意义．

本文首先针对误差为鞅差序列的线性模型采用了经验似然方法，构造了参数的对数经验似然比统计量，并在给定的基本条件下证明了该统计量具有渐近卡方分布．接下来，本文又针对误差为弱平稳线性过程的线性模型采用了经验似然方法．同样地，我们也构造了参数的对数经验似然比统计量，虽然证明得出了在一定的假设条件下该统计量具有渐近正态分布，但由于渐近方差的未知性，因此本文进一步采用了分组经验似然的方法．，构造了参数的分组对数经验似然比统计量，更证明了该统计量在一定条件下服从渐近卡方分布，这正是我们所期望得到的结果．这就是本文所得到的三个主要结论．这就

为进行大样本参数假设检验和构造参数的置信域奠定了基础. 最后, 本文还对鞅差序列误差情行进行了数据模拟, 分为几种不同情况分别进行了模拟, 并对模拟结果做了解释, 从而说明了利用经验似然方法构造的置信区间与最小二乘方法构造的置信区间相比, 具有较高的覆盖概率以及较短的区间长度, 并且相对于最小二乘方法具有较高的稳定性.

关于相依随机变量的研究已成为一个主要的研究方向, 特别是具有相依数据的统计模型. 而将极具有优势的经验似然方法应用于具有相依误差的线性模型, 这在实际中有着更为广阔的应用背景. 深信随着二者在理论和方法上的不断完善和发展, 它们对经济、通信、生物等各个领域都将起着积极的促进作用.

关键词: 线性模型　鞅差序列　弱平稳线性过程　经验似然　分组经验似然　渐近卡方分布

Abstract

The data gathered from study of some modern sciences such as biology and information science is usually large, high dimensional and dependent. This is one of the most popular, most challenging domains and is most likely to be broken through in applications and theories of statistics. So, the studies for dependent data has aroused people's attention. The studies, however, for statistical models based on dependent data is still insufficient and not much accounted of. This paper investigate a linear model which is based on two kinds dependent errors, one is martingale difference process and another is sub-stable linear process.

For the parameter β, there is usually least square method in literatures.In this paper, we present a new way–empirical likelihood inference to treat the model. Empirical likelihood method has many advantages. For example, the confidence intervals constructed has invariablity, the shapes of confidence regions are all determined by their data and Bartlett-Correctable, and so on. So this method is very significative in both theories and applications.

First of all, we adopt empirical likelihood to the linear model based on martingale difference errors, construct empirical log-likelihood ratio statistic of the parameter and we prove that the empirical log-likelihood ratio statistic is asymptotically chi-squared distributed under quite general conditions. Second,

we adopt this method to the linear model based on sub-stable linear errors. Similarly, we also construct parameter's empirical log-likelihood ratio statistic. Although we prove that it is asymptotically normal distributed under some conditions, we present blockwise empirical likelihood to this model because we don't know the asymptotically variance. Then we prove that the blockwise empirical log-likelihood ratio statistic is sympotically chi-squared distibuted under some conditions and this is our expected. These are three main conclusions in this paper. Therefore, the empirical likelihood confidence regions can be constructed accordingly. Finally, a simulation study is constructed for the linear model based on martingale difference errors. We study it in several different situationsand then we explain the reasons of simulant results. All of these simulant results confirm that the confidence intervals constructed by the empirical likelihood method have higher coverage accuracies and shorter interval lengths.

The researchs for dependent data have already became a main direction, specially of statistic modles based on dependent data. With the development of the model and the empirical likelihood, they will be beneficial to all sorts of fields.

Keywords: Linear model; Martingale difference process; Sub-stable linear process; Empirical likelihood ; Bloclwise empirical likelihood ; Asymptoti-

cally chi-squared distributed .

目　录

第 1 章　绪　论

1.1　引言

现代科学技术和社会经济的许多领域都遇到高维、相依和不完全数据的统计分析问题，它是目前统计学应用和理论中面临最多、挑战最严峻，也是最有可能取得突破的研究领域之一.

在当今的信息和知识经济时代，人类研究的科学和社会问题更加高深、更加复杂、更加庞大，有效地收集和分析数据以提取信息和获得知识变得更加须臾不可离. 而且发达的信息技术和高性能的计算机使收集、储存、传输数据和进行科学计算更加便捷，这就给统计学提出了许多更大、更难、更复杂的问题. 例如，复杂系统建模，经济金融中的预测和决策，高维、定性和不完全数据的统计分析，数据库等巨型复杂数据的信息提取和知识发现，相依变量和动态系统的统计规律，小样本和相关信息的统计推断，信号和图像的统计处理等等. 因此，统计学正面临着前所未有的巨大挑战和机遇，其中高维、相依和不完全等复杂数据的统计分析使一个带有普遍性的突出难题. 例如，雷达网、信息网等接受和传输的信号数据，卫星遥感观测的图像数据，文字语言的记录数据， DNA 和蛋白质结果的测试数据，全国乃至全球的气象和环境数据，人口、企业、教育、科技、医疗卫生等社会调查数据，以及股市、保险、信用卡、房地产等经济金融数据，都是维数很高 (数十，数百乃至上千维) 、结构十分复杂的数据，许多数据还是前后相依 (不独立) 的；在工业、国防、天文和医学等领域存在大量的不完全数据，即人们不能得到所关心的某些变量的数值，而制止到它们所在的区域，甚至完全不知道某些变量在部分场合的数值 (即缺失数据). 而且在实际问题中，往往同时遇到多种情况. 例如，来自复杂现象的数据一般是高维的，许多时候还是相依的，动态的，有时还含有不完全数据；图像和信号可以转化为高维数据处理，而信号一般是前后相依的动态数据；许多高维数据 (例如 DNA 和基因数据), 相对其维数而言，样本量相当

小；许多小样本问题中经常含有不完全数据等等. 对于这些问题，统计学工作者和实际领域中的数据分析工作者都进行了许多研究，有些方面已经形成了一些比较有效的方法，但很多方面还处在探索阶段，缺乏系统有效的方法，更缺乏完整系统的理论. 近些年，特别是在发达国家中，有许多统计学家深入到实际领域，或与实际领域的专家合作，探索解决这些挑战性问题的途径.

关于相依随机变量的研究，已取得一些研究成果. 然而，对各种具体的相依数据的统计模型的研究还不够充分和不够完备，本文利用经验似然和分块经验似然的方法，研究了分别具有两种特定相依误差的线性模型. 我们先来了解相依数据的研究现状和应用情况.

1.2 相依数据简介

在一些近代科学研究中，如生命科学和信息科学的研究，人们获得的数据往往具有量大，高维和相依等特点. 于是，关于相依随机变量的研究，已引起人们的重视，取得一些研究成果，并提出了一些研究和加工相依变量的有效方法，如 Bootstrap 、分块 Bootstrap 、 Jackknife 和分块 Jackknife(Blocks of blocks jackknife) 等方法.

1 、与相依数据的统计分析关系密切的是时间序列分析. 在工农业生产、科学技术和社会经济生活的许多领域，普遍存在着按时间顺序发生的具有概率特征的各种随机现象. 人们通过观测把这些现象记录下来成为可供分析的随机数据. 通常有两种类型：一种是时间上连续的，即按时间顺序连续地、不间断地发生和记录下来；另一种是离散的，即每隔一段时间观测记录一次. 这后一种离散型数据人们就成为时间序列或随机序列. 工厂逐日记录的废品率；农业地区逐年记录的小麦产量；气象工作者记录的某一地区每日最高和最低气温；雷达跟踪目标时不同时刻的测距误差；某种商品逐月的销售数量；等等. 这些都是时间序列的实例. 由于时间序列的随机性，人们必须利用概率和统计学中的一些概念和方法来分析这些数据. 这就是时间序列分析. 其目的是

为了在一列离散型随机数据中提取有用的信息 – 找出客观事物发展的规律，预测其发展趋势并进行必要的控制. 这些工作无疑在经济上、国防上和科学上具有重要的意义.

时间序列最早是应用在经济学上的. 像：全国每月销售总额，政府财政的顺差和赤字，个人消费总额等等. 对于这类数据的分析和预报是时间序列分析的专长. 时间序列还可用于市场预报，价格预报，产量预报等等. 近年来经济学中一门新兴的学科叫做计量经济学，其中最主要的数学方法之一就是时间序列分析. 在自然现象的研究中，时间序列分析被广泛应用到天气预报当中，并且可以用来预报河流流量，预报洪水等自然灾害，还可用于海洋波的分析. 工业自动化中也大量用到时间序列分析方法：可以把测试仪器测到的数据用时间序列分析方法处理并预报下一时刻值，把这预报值和标准值进行比较，将差值输给控制器对系统进行控制，使下一时刻的值接近标准值. 在自动控制中，时间序列分析用于系统识别，用于建立卡尔曼 (Kalman) 滤波器的噪声模型，用于建立陀螺漂移模型，建立雷达测量误差模型等等. 在医学、生物、生态研究中，用尖利时间序列模型的方法可描述人体或其它生物体对外界刺激反应的情况. 这种方法在挑选宇航员和飞行员时是很有用的. 用时间序列分析还可进行疫情的预测，以表彰我传染病率的变化情况. 还可以预报生物群体的中枢，预报渔汛. 用多维序列还可以研究几种有关生物群体所构成生态系统的繁衍规律.

半个多世纪以来，时间序列分析已逐步发展成为概率统计学中的一个重要分支. 如果我们简单回顾一下这一学科的历史发展过程，大体上可以把它分成两个阶段，早期阶段和以现代谱分析为特征的成熟阶段. 早期阶段的工作主要使用一种比较简单的有限参数模型，包括“滑动和”以及由差分方程所定义的自回归模型等. 用这些模型可近似地代替一类相当广泛的平稳随机序列，能解决许多实际问题，其主要缺点是精确度不高. 从五十年代后期开始，时间序列分析进入了一个较高的领域. 这一发展阶段的特点是，一方面它与近代“

随机过程论”密切结合，使其理论也走向严格化；同时在处理方法上，它考虑的是无限维模型，把以前的有限维模型看成为一种特例；而近代谱估计理论和方法称为这一阶段的一个显著特点．近年来，随着现代计算技术的发展，时间序列分析中的线性模型方法引起了广泛注意．

在这方面我国学者也有很好的工作．北京大学江泽培是这方面最资深的学术带头人．改革开放以来，他和他的学生何书元等在平稳序列和随机场的谱估计方面取得了一系列重要成果．北京大学谢衷洁、程乾生等对天文、地质、医学和金融等实际数据进行分析，均获得了有价值的发现．中国科学院应用数学所安鸿志在时间序列分析中的重对数律，自回归的选元方法及其相容性，以及非线性和条件异方差模型的性质等方面有不少重要成果，有些成果在国际上有较多引用，受到好评．

2 、另一类重要的相依数据问题时统计信号处理．赵林城在信号检测和叠加指数信号的估计及其性质方面有不少好的结果，受到国际同行的重视．他提出的一个信号检测准则被外国学者称为有效检测准则，在文献中多次被人引用和仿效．

1.3 经验似然推断的研究简状

经验似然是 Owen[1] 在完全样本下提出的一种非参数统计推断方法，它类似于 Bootstrap 的抽样特性．这一方法与经典的和现代的统计方法比较有很多突出的优点，如：用经验似然方法构造的置信区间有变换不变性及置信域的形状由数据自行决定等诸多优点．正因如此，这一方法引起了许多统计学家的兴趣，她们将这一方法应用到各种统计模型及各种领域，如 Owen[1,2,3] 由对总体均值得推断提出经验似然并随后将其应用到线性回归模型的统计推断； Kolaczyk[10] 将经验似然应用于广义线性模型； Qin[11] 应用经验似然于偏度抽样模型的统计推断； Owen 还将经验似然应用到投影寻综回归的研

究； Zhang[12] 将经验似然应用于分位回归及 M- 泛函的统计推断； Chang 和 Chan[13] 发展了自回归模型的经验似然方法； Zhong 和 Rao[14] 将经验似然应用于抽样调查问题的研究； Kitamura[15] 等将经验似然应用到经济模型的研究等. 更值得注意的是，近年来一些统计学家又将经验似然方法应用到不完全数据的统计分析，发展了所谓的被估计的经验似然、调整经验似然及 Bootstrap 经验似然. 应该指出，经验似然的思想至少可以追溯到 Thomas 和 Gunkemeier(1975) 用 Kaplan-Meier 曲线估计生存概率，他们直观地推测生存概率的经验似然比置信区间有近似正确的覆盖程度. 注意到这一方法的本质是在约束条件下极大化非参数似然比，感兴趣的参数由约束条件带入这一极大化似然比中. Owen 将这一思想方法应用到完全独立同分布样本下总体均值这一简单而重要情形的统计推断. 由于 Owen 使用线性约束条件，从而表明了这一方法有非常一般的应用，这是因为统计中许多估计方程关于感兴趣的参数或参数的某已知函数是线性的或许多统计模型的参数可由关于该参数或它的某已知函数的线性方程决定. 注意到这一特点，许多统计学家将 Owen 在完全样本下的经验似然方法推广到了不完全数据类型的统计推断.

那么什么是经验似然呢？

设 $X_1, X_2, \cdots, X_n \in R^d$ 独立有共同的累积分布 F, 则 F 的非参数似然是

$$L(F) = \prod_{i=1}^{n} F(\{X_i\}),$$

这里 $F(\{X_i\})$ 是分布 F 在 X_i 处的概率质量，其中 $i = 1, 2, \cdots, n$. 大家知道 $X_1, X_2, \cdots, X_n \in R^d$ 的经验累积分布函数 $F_n = n^{-1} \sum_{i=1}^{n} \sigma_{x_i}$ 使上式达到极大，其中 $\sigma_x(A) = I[x \in A]$. 也就是 F_n 是 F 的非参数极大似然估计.

在参数推断中人们利用参数似然比进行假设检验与置信区间估计. 类似的，在分布完全未知的情况下非参数似然比

$$R(F) = \frac{L(F)}{L(F_n)}.$$

也可以用于统计推断. 不象参数似然比, 非参数似然比中不包含未知参数. 一个自然的问题是如何使用它对参数作统计推断. 注意到一些参数 θ 是总体分布的泛函, 即 $\theta = T(F) \in R^p$, 其中 $T(\cdot)$ 是分布 F 的某泛函, F 属于某分布类 ℓ, 如总体均值及分位点等就是有上述形式泛函的例子. 为了对 $T(F) = \theta$ 作检验, Owen[29] 定义如下经验似然比统计量

$$\Re(\theta) = \sup_F \{R(F)|T(F) = \theta, F \in \ell\}.$$

很显然, 经验似然比实际上是一种截面非参数似然比函数, 它要求 F 在满足约束条件 $T(F) = \theta$ 下使非参数似然比达到极大 (在无约束条件时, 极大非参数似然比为 1), 而参数 θ 由这一约束条件引入这一极大似然比中, 从而得到关于参数 θ 的极大截面非参数似然比函数, 用这一非参数似然比作假设检验、区间估计或进行其他统计推断, 这一方法就是所谓的经验似然方法. 如果 $\Re < r_0$, 经验似然假设检验拒绝 $H_0 : T(F) = \theta_0$, 而似然置信域为

$$\{\theta : \Re(\theta) \geq r_0\},$$

其中 r_0 是某临界值.

注意到 $\Re(\theta)$ 是关于非参数似然比在限制条件下关于 F 求极大, 这种形式的定义显然不利于计算. 下面给出与上面定义等价但容易计算的形式. 由 $\Re(\theta)$ 的定义, 注意到只有那些在观察值点有正概率质量的离散分布类中求极大. 我们首先假设数据没有 " 结 "(当有 " 结 " 时, 根据 Owen[1], 所求的经验似然是相同的), 在没有 " 结 " 时, 我们设分布 F 为 F_p, 即 $F_p = \sum_{i=1}^{n} p_i \sigma_{X_i}$. 注意到 $L(F_n) = n^{-n}$, 因而

$$\Re(\theta) = \sup_{\sum_{i=1}^{n} p_i = 1, T(F_p) = \theta} \prod_{i=1}^{n} (np_i),$$

显然 Lagrange 法可应用于计算 $\Re(\theta)$.

经验似然应用于推断的另一个问题是如何确定临界值 r_0, 这一问题实际

上归结为求 $\Re(\theta)$ 的渐近分布. 本文的重要定理之一就是证明了似然比统计量的渐近卡方性质，这一特性就形成了经验似然推断的基础.

经验似然推断在总体均值推断、线性模型推断、分位数推断、估计方程推断及利用辅助信息进行推断等几种重要统计推断中有着广泛的应用，是一种重要的统计推断方法.

1.4　本文主要研究内容

本文考虑线性模型：

$$y_i = x_i^T\beta + \varepsilon_i, \qquad i = 1, 2, \cdots,$$

其中 $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^T$ 为 p 维的已知固定设计点列，β 为 p 维未知回归系数，$y_i \in \mathbb{R}$ 为响应变量，$\varepsilon_i \in \mathbb{R}$ 为随机误差，这里分别假设 $\{\varepsilon_n, \mathcal{F}_n, n \geqslant 1\}$ 为 L_2 中的鞅差序列及弱平稳线性过程，对任意的 n≥1, 有 $\mathbf{X_n} = (x_1\ x_2 \cdots\ x_n)$ 满足 $\mathbf{X_n X_n^T} = \sum_{i=1}^{n} x_i x_i^T$ 满秩.

近年来，对各种具体的相依数据的统计模型的研究越来越引起人们的重视和兴趣. 文献中一般是针对平稳混合序列，例如 φ- 混合序列、α - 混合序列、β - 混合序列等的回归模型. 得到的是参数 β 的最小二乘估计的强相合性. 对于误差为鞅差序列和弱平稳线性相依过程的线性模型，文献中 (胡舒合) 得到了参数 β 的渐近正态性. 本文采用了新的推断方法 - 对数经验似然和分组对数经验似然方法研究上述模型，这种方法为研究回归模型和相依数据提供了新的途径. 定理 2.1 证明了误差是鞅差序列时，线性模型参数 β 的对数似然比统计量的渐近 χ^2 性；定理 2.2 和定理 2.3 分别证明了，误差为弱平稳线性过程时，线性模型参数 β 的对数似然比统计量的渐近正态性和分组对数似然比统计量的渐近 χ^2 性. 这就为构造参数 β 的置信区间提供了理论依据. 在本文的最后，通过 Monte-Carlo 方法进行了数值模拟，不仅比较了经验似然方法与最小二乘方法所构造的置信区间之间的差别，证明了经验似然方法的优良性，还比较了不同的误差方差下用经验似然方法所构造的置信区间

的差别，并做了一定的理论解释.

用经验似然推断方法可来构造模型中参数的置信域或进行假设检验是 Owen[1] 在 1988 年提出的，它主要是由样本经验分布是总体分布的非参数极大似然估计这一思想的启发，从而构造出了一个似然比统计量. 近年来，国内一些学者也作了相关方面的研究. 王启华等人 [19,20] 对右删失数据进行了经验似然线性回归分析，并得出了经验似然置信区间有更精确的区间覆盖程度. 目前，也有很多学者对固定设计下半参数回归模型的经验似然推断进行了深入的研究. 秦永松 [18] 对一类固定设计下的半参数模型进行研究，得到了对数经验似然比统计量的渐近性质. Wang[19] 和 Shi[21] 等人也对这类模型也进行了深入的研究.

本文是针对两种具体的相依误差的线性模型的研究，分别得到对数经验似然比统计量和分组对数经验似然比统计量的渐近 χ^2 性，并且采用经验似然的方法构造参数 β 的置信区间. 通过模拟比较得出这种方法的覆盖率优于利用最小二乘法构造的 β 的置信区间.

经验似然推断是一个历史不长，尚在发展中的领域. 目前取得的成果中，有一些还不是最终的. 本领域的研究主要属大样本性质，所得到的结果也属大样本结果，还有必要对小样本容量时，研究它的性质. 再则，对于相依误差，渐近正态性，渐近 χ^2 性已具有一些深度，但是本文得到的毕竟是两种具体的相依误差序列，不具有较大的普遍性. 由于相依数据的的复杂性，因此，对于一般的相依误差序列的线性模型的统计推断，以至于其它模型，例如非线性回归模型和半参数回归模型，必将遇到不少困难. 用具有诸多优点的经验似然方法对具有很强解释能力的回归模型进行研究，在实际中有着更为广阔的实用背景. 深信随着二者在理论和方法上的成熟，对经济、医药、工农业生产等方面将起着积极的促进作用.

1.5 本文结构

由于本文中研究的误差类型有着广泛的应用背景，因此我们在第一章前半部分，简约介绍了相依数据和近年来的研究成果. 第一章的后半部分，我们介绍了经验似然推断的发展状况，包括了此类推断方法产生的原因，对典型统计模型中推断的影响. 这两部分主要是作为正文的预备知识而出现的，所以我们只是做了简略介绍.

本文的第二章主要是相依误差下线性模型的统计推断理论. 第一节介绍了所要研究的统计模型和经验似然比方法，第二节是针对误差为鞅差序列的线性模型中的参数的统计推断方法以及理论性质的证明，第三节则是针对误差为弱平稳线性相依过程的线性模型中的参数的统计推断方法以及理论性质的证明.

第三章则注重于用数值结果说明本文提出的统计推断方法的优良性. 我们通过抽取随机数，进行 Monte-Carlo 模拟，以此验证新的方法中的优良性质.

附录是模拟的源程序，以便读者参考.

第 2 章　主要结果及其证明

2.1　引言

经验似然是 Owen(1988,1990) 引入的一种非参数推断方法，并随后将其应用到线性模型当中. 此后一些学者又将其应用到半参数模型，讨厌参数，回归函数，密度核估计，有偏样本等等，但是大部分的结论都是在样本独立同分布的情况下讨论的，对样本非独立的情况研究较少. 本文就误差为鞅差序列及弱平稳线性过程的线性模型进行了讨论.

考虑线性模型：

$$y_i = x_i^T\beta + \varepsilon_i, \qquad i = 1, 2, \cdots,$$

其中 $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^T$ 为 p 维的已知固定设计点列，β 为 p 维未知回归系数，$y_i \in \mathbb{R}$ 为响应变量，$\varepsilon_i \in \mathbb{R}$ 为随机误差，这里假设 $\{\varepsilon_n, \mathcal{F}_n, n \geqslant 1\}$ 分别为 L_2 中的鞅差序列及弱平稳线过程，对任意的 n≥1，有 $\mathbf{X_n} = (x_1\ x_2 \cdots\ x_n)$ 满足 $\mathbf{X_n}\mathbf{X_n^T} = \sum_{i=1}^{n} x_i x_i^T$ 满秩.

文献中已有这两种误差下的线性模型的文章，但是参数估计方法是最小二乘法，基于此方法的估计不稳健，但可以利用其渐近正态性得到参数 β 的对数经验似然统计量的渐近分布，从而构造参数 β 的置信域，并可进行假设检验.

经验似然方法是研究半参数模型的一种新途径，它可以用来构造模型中参数的置信域或进行假设检验. 用这种方法构造置信区域是 Owen[1] 在 1988 年提出的，它主要是由样本经验分布是总体分布的非参数极大似然估计这一思想的启发，从而构造出了一个似然比统计量. 张军舰和王成名把分组经验似然的方法引入到了相依数据，得到了强平稳 m 相依数据和强平稳 ϕ 混合数据的分布的经验似然估计，并构造了均值的经验似然置信区间. 由于这种方法的优良性质，它正越来越引起经济学家、生物学家和统计学家的重视. 它对构

造相依误差的线性模型中参数的置信区间提供了一个很有价值的方法.

2.2　误差为鞅差序列

2.2.1　方法和结果

考虑线性模型：

$$y_i = x_i^T\beta + \varepsilon_i \qquad i = 1, 2, \cdots, \tag{2.1}$$

其中 $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^T$ 为 p 维的已知固定设计点列，β 为 p 维未知回归系数，$y_i \in \mathbb{R}$ 为响应变量，$\varepsilon_i \in \mathbb{R}$ 为随机误差, 这里假设 $\{\varepsilon_n, \mathcal{F}_n, n \geqslant 1\}$ 为 L_2 中的鞅差序列，对任意的 n≥1, 有 $\mathbf{X_n} = (x_1\ x_2 \cdots\ x_n)$ 满足 $\mathbf{X_n X_n^T} = \sum_{i=1}^{n} x_i x_i^T$ 满秩.

令 $Z_i = x_i(y_i - x_i'\beta)$, i =1,2,$\cdots$. 则由 Owen[1] 可知 β 的经验似然函数为

$$\mathcal{R}(\beta) = \sup\{\prod_{i=1}^{n} n\omega_i | \sum_{i=1}^{n} \omega_i = 1,\ \sum_{i=1}^{n} \omega_i Z_i = 0,\ \omega_i \geqslant 0\}, \tag{2.2.1}$$

并经 Lagrange 乘子法运算得到：

$$\mathcal{I}_E(\beta) = -2\log \mathcal{R}(\beta) = 2\sum_{i=1}^{n} \log(1 + \lambda' Z_i), \tag{2.2.2}$$

其中，$\lambda \in \mathbb{R}^p$, 满足：

$$0 = g(\lambda) = \frac{1}{n}\sum_{i=1}^{n} \frac{Z_i}{1 + \lambda' Z_i}.$$

类似于 Owen 的对数经验似然比统计量渐近于 χ^2 分布. 因此我们也希望 $\mathcal{I}_E(\beta_0)$ 也是一个渐近 χ^2 分布. 为此做以下的一些假设：

(i) $\{\varepsilon_n, \mathcal{F}_n, n \geqslant 1\}$ 为 L_2 中的鞅差序列，且有

$$\lim_{c\to\infty} \sup_{1\leqslant j<\infty} E(\varepsilon_j^2 I_{\{|\varepsilon_j|>c\}} | \mathcal{F}_{j-1}) = 0, \quad a.s. \tag{2.2.3}$$

(ii) $E(\varepsilon_i^2|\mathcal{F}_{j-1})=\sigma_i^2$, a.s. $i=1,2,\cdots$, 且有 $\inf\limits_i \sigma_i^2>0$ 及

$$\lim_{n\to\infty}\frac{\max\limits_{1\leqslant i\leqslant n}(\parallel x_i\parallel^2)}{\lambda_p(\mathbf{X_n}\mathbf{X_n^T})}=0. \tag{2.2.4}$$

这里, $\lambda_p(\mathbf{X_n}\mathbf{X_n^T})>0$ 为 $\mathbf{X_n}\mathbf{X_n^T}$ 的最小特征根.

(iii) 当 $n\to\infty$ 时, $\exists A_0>0$, 使得

$$\frac{1}{n}\sum_{i=1}^{n}x_ix_i^T\sigma_i^2=A_o+o(1). \tag{2.2.5}$$

(iv)ε_i 存在 4 阶矩, $i=1,2,\cdots$, 且有

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\|x_i\|^4E\varepsilon_i^4=0. \tag{2.2.6}$$

从而得出第一个定理:

定理 2.1: 假设条件 (i) – (iv) 成立, 则有当 β 等于真参数 β_0 时,

$$\mathcal{I}_E(\beta_0)\xrightarrow{L}\chi_p^2\ , \qquad (n\to\infty).$$

2.2.2 几个引理

引理 2.2.1 对于本节中的模型 (2.1), 若条件 (i),(ii) 成立, 则对任意的 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_p)^T$,

$\alpha\in\mathbb{R}^p$, $\parallel\alpha\parallel=1$, 都有

$$\frac{\alpha^T\mathbf{X_n}\varepsilon}{\sqrt{Var(\alpha^T\mathbf{X_n}\varepsilon)}}\xrightarrow{L}\mathbf{N}(0,1) \qquad (n\to\infty).$$

其中 $\varepsilon=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)^T\in\mathbb{R}^p$.

证明: 证明类似于胡舒合 [6], 因为对任意的 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_p)^T$,

$\alpha\in\mathbb{R}^p$, $\parallel\alpha\parallel=1$, 有

$$\alpha^T\mathbf{X_n}\varepsilon=\alpha^T\sum_{j=1}^{n}x_j\varepsilon_j=\sum_{j=1}^{n}(\sum_{i=1}^{p}\alpha_ix_{ji})\varepsilon_j,$$

所以

$$Var(\alpha^T\mathbf{X_n}\varepsilon)=\alpha^T Cov(\mathbf{X_n}\varepsilon)\alpha=\sum_{j=1}^{n}(\sum_{i=1}^{p}\alpha_i x_{ji})^2\sigma_j^2,$$

所以有

$$\frac{\alpha^T\mathbf{X_n}\varepsilon}{\sqrt{Var(\alpha^T\mathbf{X_n}\varepsilon)}}=\sum_{j=1}^{n}\frac{\sum\limits_{i=1}^{p}\alpha_i x_{ji}}{\sqrt{\sum\limits_{j=1}^{n}(\sum\limits_{i=1}^{p}\alpha_i x_{ji})^2\sigma_j^2}}\varepsilon_j.$$

令 $U_k=\dfrac{\sum\limits_{i=1}^{p}\alpha_i x_{ji}}{\sqrt{\sum\limits_{j=1}^{n}(\sum\limits_{i=1}^{p}\alpha_k x_{ki})^2\sigma_j^2}}$，　$k=1,2,\cdots,n$, 则显然有 $\{U_k\varepsilon_k,\mathcal{F}_k,k\geqslant 1\}$ 也为鞅差序列，且对任意的 $1\leqslant k\leqslant n$ 有

$$E(U_k\varepsilon_k)^2=\frac{(\sum\limits_{i=1}^{p}\alpha_i x_{ki})^2\sigma_k^2}{\sum\limits_{j=1}^{n}(\sum\limits_{i=1}^{p}\alpha_i x_{ji})^2\sigma_j^2}>0,$$

以及 $E(\sum\limits_{j=1}^{n}U_j\varepsilon_j)^2=1$. 所以对任意的 $\delta>0$, 有

$$\begin{aligned}
&\sum_{k=1}^{n}E(U_k^2\varepsilon_k^2 I_{|U_k\varepsilon_k|>\delta}\mid\mathcal{F}_{k-1})\\
=\;&\sum_{k=1}^{n}\frac{(\sum\limits_{i=1}^{p}\alpha_i x_{ki})^2\sigma_k^2}{\sum\limits_{j=1}^{n}(\sum\limits_{i=1}^{p}\alpha_i x_{ji})^2\sigma_j^2}\sigma_k^2E(\varepsilon_k^2I(\varepsilon_k^2>\delta\frac{\sum\limits_{j=1}^{n}(\sum\limits_{i=1}^{p}\alpha_i x_{ji})^2\sigma_j^2}{(\sum\limits_{i=1}^{p}\alpha_i x_{ki})^2\sigma_k^2})\mid\mathcal{F}_{k-1})/\sigma_k^2\\
\leqslant\;&\sup_k E(\varepsilon_k^2I(\varepsilon_k^2>\delta\frac{\sum\limits_{j=1}^{n}(\sum\limits_{i=1}^{p}\alpha_i x_{ji})^2\sigma_j^2}{(\sum\limits_{i=1}^{p}\alpha_i x_{ki})^2})\mid\mathcal{F}_{k-1})/\inf_i\sigma_i^2.
\end{aligned}$$

又因为

$$\frac{\sum\limits_{j=1}^{n}(\sum\limits_{i=1}^{p}\alpha_i x_{ji})^2\sigma_j^2}{(\sum\limits_{i=1}^{p}\alpha_i x_{ki})^2}\;\geqslant\;\frac{\alpha^T(\mathbf{X_n}\mathbf{X_n}^T)\alpha}{(\sum\limits_{i=1}^{p}\alpha_i x_{ki})^2}\cdot\inf_i\sigma_i^2$$

$$\geqslant \frac{\lambda_p(\mathbf{X_n}\mathbf{X_n^T})}{(\sum_{i=1}^{p}\alpha_i x_{ki})^2}\cdot\inf_i \sigma_i^2$$
$$\geqslant \frac{\lambda_p(\mathbf{X_n}\mathbf{X_n^T})}{\max\limits_{1\leqslant i\leqslant n}(\|x_i\|^2)}\cdot\inf_i \sigma_i^2.$$

所以由条件 (ii) 式可知， 当 $n\to\infty$ 时， 上式 $\longrightarrow\infty$.

又由条件 (i) 以及 $\inf\limits_i \sigma_i^2>0$ 可知

$$\lim_{n\to\infty}\sum_{k=1}^{n}E(U_k^2\varepsilon_k^2 I_{|U_k\varepsilon_k|>\delta}\mid\mathcal{F}_{k-1})=0,\quad a.s.$$

所以由胡舒合 [6] 可知

$$\frac{\alpha^T\mathbf{X_n}\varepsilon}{\sqrt{Var(\alpha^T\mathbf{X_n}\varepsilon)}}\xrightarrow{L}\mathbf{N}(0,1)\qquad(n\to\infty).$$

证毕.

注: 显然对于该模型，由条件 (iii) 可得

$$Cov(\frac{1}{\sqrt{n}}\mathbf{X_n}\varepsilon)=\frac{1}{n}\sum_{i=1}^{n}x_i x_i^T\sigma_i^2=A_o+o(1).$$

所以由 α 的任意性可知

$$\frac{1}{\sqrt{n}}\mathbf{X_n}\varepsilon\xrightarrow{L}N_p(0,A_o).\tag{2.2.7}$$

引理 2.2.2 对于本节中的模型 (2.1), 令 $Y_n=\max\limits_{1\leqslant i\leqslant n}\|Z_i\|$, 若条件 (iv) 成立， 则有 当 β 等于真参数 β_0 时， $Y_n=o_p(n^{\frac{1}{2}})$.

证明：因为当 β 等于真参数 β_0 时， $Z_i=x_i\varepsilon_i,\quad i=1,2,\cdots$, 所以，由契毕雪夫不等式知，当 β 等于真参数 β_0 时,

$$P(Y_n>n^{\frac{1}{2}}) = P(\max_{1\leqslant i\leqslant n}\|Z_i\|>n^{\frac{1}{2}})$$

$$\leqslant \sum_{i=1}^{n} P(\|Z_i\| > n^{\frac{1}{2}})$$
$$\leqslant \frac{1}{n^2}\sum_{i=1}^{n} \|x_i\|^4 E\varepsilon_i^4.$$

由条件 (iv) 式知： 上式 $\longrightarrow 0$ $(n \to \infty)$. 所以 $Y_n = o_p(n^{\frac{1}{2}})$. 证毕.

引理 2.2.3 对于本节中的模型 (2.1), 令 $S_n = \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i^T$, 若条件 (i)-(iv) 成立， 则有

当 β 等于真参数 β_0 时， $S_n - ES_n \xrightarrow{P} 0$, $(n \to \infty)$.

证明： 因为当 β 等于真参数 β_0 时， $Z_i = x_i\varepsilon_i$, $i = 1, 2, \cdots,$

且由 $\{\varepsilon_n, \mathcal{F}_n, n \geqslant 1\}$ 为 L_2 中的鞅差序列， 则可知当 β 等于真参数 β_0 时,

$$ES_n = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T \sigma_i^2.$$

因为

$$S_n = \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i^T = \frac{1}{n}\sum_{i=1}^{n} x_i x_i^T \varepsilon_i^2,$$

所以只需证

$$\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T(\varepsilon_i^2 - \sigma_i^2) \xrightarrow{P} 0 \qquad (n \to \infty)$$

即可. 取 $\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T(\varepsilon_i^2 - \sigma_i^2)$ 的第 k 行第 l 个元素， 即

$$\frac{1}{n}\sum_{i=1}^{n} x_{ik}x_{il}(\varepsilon_i^2 - \sigma_i^2).$$

则对 $\forall \delta > 0$, 有

$$P(\frac{1}{n}|\sum_{i=1}^{n} x_{ik}x_{il}(\varepsilon_i^2 - \sigma_i^2)| < \delta) \leqslant \frac{1}{n^2\delta^2} E(\sum_{i=1}^{n} x_{ik}x_{il}(\varepsilon_i^2 - \sigma_i^2))^2.$$

由条件 (ii) 可知， $\{\varepsilon_n^2 - \sigma_n^2, \mathcal{F}_n, n \geqslant 1\}$ 也为鞅差序列. 则

$$\text{上式} \quad = \quad \frac{1}{n^2\delta^2}\sum_{i=1}^{n} x_{ik}^2 x_{il}^2 E(\varepsilon_i^2 - \sigma_i^2)^2$$

$$\begin{aligned} &\leqslant \quad \frac{1}{n^2\delta^2}\sum_{i=1}^{n}(x_{ik}^4+x_{il}^4)E\varepsilon_i^4 \\ &\leqslant \quad \frac{1}{n^2\delta^2}\sum_{i=1}^{n}\|x_i\|^4E\varepsilon_i^4 \\ &(\text{由 (2.2.6) 式}) \quad \longrightarrow 0 \quad (n\to\infty). \end{aligned}$$

所以

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}x_{ik}x_{il}(\varepsilon_i^2-\sigma_i^2)\xrightarrow{P}0.$$

证毕.

2.2.3 定理 2.1 的证明

在这里沿用引理 2.2.2 和引理 2.2.3 中的记号： $Y_n=\max\limits_{1\leqslant i\leqslant n}\|Z_i\|$, $S_n=\frac{1}{n}\sum\limits_{i=1}^{n}Z_iZ_i^T$.

则由引理 2.2.3 及 (2.2.5) 可知：当 β 等于真参数 β_0 时，对任意的 $\theta\in\mathbb{R}^p$, $\|\theta\|=1$, 有

$$\theta^TS_n\theta\geqslant\lambda_p(A_o)+o(1)>0.$$

又由引理 2.2.2: 此时， $Y_n=o_p(n^{\frac{1}{2}})$, 则可类似于 Owen[2] 中定理 1 的证明得到 $\|\lambda\|=o_p(n^{-\frac{1}{2}})$. 继而有：

$$\lambda=S_n^{-1}(\frac{1}{n}\sum_{i=1}^{n}Z_i)+\alpha, \tag{2.2.8}$$

其中 $\|\alpha\|=o_p(n^{-\frac{1}{2}})$.

又因为同 Owen[2] 中定理 1 有：

$$\begin{aligned} \mathcal{I}_E(\beta_0) &= 2\sum_{i=1}^{n}\log(1+\lambda^TZ_i) \\ &= 2\sum_{i=1}^{n}[\lambda^TZ_i-\frac{(\lambda^TZ_i)^2}{2}]+2\sum_{i=1}^{n}\eta_i. \end{aligned}$$

其中 $|2\sum_{i=1}^{n}\eta_i| = o_p(1)$.

所以将 (2.2.8) 式带入即得

$$\begin{aligned}\mathcal{I}_E(\beta_0) &= 2\sum_{i=1}^{n}\log(1+\lambda^T Z_i)\\ &= (\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_i)^T S_n^{-1}(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_i)+o_p(1).\end{aligned}$$

所以由 (2.2.6) 式， (2.2.7) 式及引理 2.2.3 即可得到

$$\mathcal{I}_E(\beta_0)\xrightarrow{L}\chi_p^2\,. \qquad (n\to\infty).$$

证毕.

2.3　误差为弱平稳线过程

2.3.1　方法和结果

考虑线性模型：

$$y_i = x_i^T\beta+\varepsilon_i, \qquad i=1,2,\cdots, \tag{(2.1)}$$

其中 $x_i=(x_{i1},x_{i2},\cdots,x_{ip})^T$ 为 p 维的已知固定设计点列， β 为 p 维未知回归系数， $y_i\in\mathbb{R}$ 为响应变量， $\varepsilon_i\in\mathbb{R}$ 为随机误差，对任意的 n≥1, 有 $\mathbf{X_n}=(x_1\ x_2\cdots\ x_n)$ 满足 $\mathbf{X_nX_n^T}=\sum_{i=1}^{n}x_ix_i^T$ 满秩.

这里假设 ε_i 是一个弱平稳线性过程：

$$\varepsilon_i=\sum_{j=-\infty}^{\infty}\psi_jY_{i-j}, \qquad i=1,2,\cdots, \tag{2.3.1}$$

其中 $\{Y_i,\mathcal{F}_i,i\geqslant 1\}$ 是一个 L_2 中的鞅差序列，且存在 $\sigma>0$, 使

$$E(Y_i^2\mid\mathcal{F}_{i-1})=\sigma^2, \quad a.s., \quad i=1,2,\cdots, \quad \sum_{j=-\infty}^{\infty}\mid\psi_j\mid<\infty. \tag{2.3.2}$$

令 $Z_i = x_i(y_i - x_i^T\beta)$, $i = 1,2,\cdots$. 则由 Owen[2] 可知 β 的经验似然函数为

$$\mathcal{R}(\beta) = \sup\{\prod_{i=1}^{n} n\omega_i | \sum_{i=1}^{n} \omega_i = 1,\ \sum_{i=1}^{n} \omega_i Z_i = 0,\ \omega_i \geqslant 0\}, \tag{2.3.3}$$

并经 Lagrange 乘子法运算得到

$$\mathcal{I}_E(\beta) = -2\log \mathcal{R}(\beta) = 2\sum_{i=1}^{n} \log(1 + \lambda^T Z_i). \tag{2.3.4}$$

其中，$\lambda \in \mathbb{R}^p$, 满足 $0 = g(\lambda) = \frac{1}{n}\sum\limits_{i=1}^{n} \frac{Z_i}{1+\lambda^T Z_i}$.

同样的，我们也希望得到对数经验似然统计量 $\mathcal{I}_E(\beta_0)$ 的渐近分布，因此要做以下的假设：

(v)ε_i 为 (2.3.1) 式和 (2.3.2) 式定义的线性过程，且满足

$$\lim_{c\to\infty} \sup_j E(Y_j^2 I_{(|Z_j|>c)} \mid \mathcal{F}_{j-1}) = 0,\ a.s. \tag{2.3.5}$$

(vi)$\forall n \geqslant 1$, 有 $\mathbf{A_n} \triangleq \sum\limits_{|k|\leqslant n-1} C(k) \sum\limits_{i=1}^{n-|k|} x_i x_{i-|k|}^T$ 正定，且有

$$\lim_{n\to\infty} \frac{n \cdot \max\limits_{1\leqslant i\leqslant n}(\|x_i\|^2)}{\lambda_p(\mathbf{A_n})} = 0, \tag{2.3.6}$$

这里，$\lambda_p(\mathbf{A_n}) > 0$ 为 $\mathbf{A_n}$ 的最小特征根，$C(k) = Cov(\varepsilon_i, \varepsilon_{i+k})$.

(vii)$\exists$ 正定阵 A_1 和 A_2, 使得 $n \to \infty$ 时，有

$$\frac{1}{n} Cov(X_n\varepsilon) = A_1 + o(1), \tag{2.3.7}$$

$$\frac{1}{n} C(0) \sum_{i=1}^{n} x_i x_i^T = A_2 + o(1), \tag{2.3.8}$$

(viii)ε_i 存在 4 阶矩，$i = 1,2,\cdots$, 且有

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \|x_i\|^4 E\varepsilon_i^4 = 0. \tag{2.3.9}$$

根据以上的条件就可以得出下面的定理 2.2,

定理 2.2: 假设条件 (v)-(viii) 成立，则可以得到结论：当 β 等于真参数 β_0 时，

$$\mathcal{I}_E(\beta_0) \xrightarrow{L} Z^T A_2^{-1} Z \ , \qquad (n \to \infty).$$

其中 $Z \sim N_p(0, A_1)$.

可以看出虽然该定理得到了 $\mathcal{I}_E(\beta_0)$ 的渐近分布，但这里若 A_1 与 A_2 未知，从而就无法构造 β 的置信区域，也无法进行假设检验. 因此在下面的讨论中我采用了局部经验似然的方法来解决这个问题.

令 $h = [n^\alpha]$, $0 < \alpha < \frac{1}{3}$, $g = [\frac{n}{h}]$. 为表达方便， 取 $\frac{n}{h}$ 为整数. 记

$$\xi_k = \frac{Z_{(k-1)h+1} + \cdots + Z_{kh}}{h}, \qquad k = 1, 2, \cdots, g.$$

显然有

$$\sum_{k=1}^{g} \xi_k = \frac{1}{h} \sum_{i=1}^{n} Z_i = \frac{1}{h} \mathbf{X_n} \varepsilon,$$

$$Cov(\sqrt{\frac{h}{g}} \sum_{k=1}^{g} \xi_k) = Cov(\frac{1}{\sqrt{n}} \mathbf{X_n} \varepsilon).$$

分组后的局部经验似然比函数为

$$\mathcal{R}_1(\beta) = \sup_{1 \leqslant k \leqslant g} \{ \prod_{k=1}^{g} g p_k \mid \sum_{k=1}^{g} p_k = 1, p_k \geqslant 0, \sum_{k=1}^{g} p_k \xi_k = o \}.$$

同样，利用 Lagrange 乘子法可以得到

$$\mathcal{R}_1(\beta) = \prod_{k=1}^{g} \frac{1}{1 + t^T \xi_k}.$$

从而得到局部对数似然比函数：

$$\mathcal{I}_{E_1}(\beta) = -2 \log \mathcal{R}_1(\beta) = 2 \sum_{k=1}^{g} \log(1 + t^T \xi_k),$$

其中， $t \in \mathbb{R}^p$, 满足

$$f(t) = \frac{1}{g}\sum_{k=1}^{g}\frac{\xi_k}{1+t^T\xi_k} = 0.$$

此时我们期望得到的是局部对数经验似然比函数 $\mathcal{I}_{E_1}(\beta)$ 的渐近分布，需要如下的条件：

(ix) ε_i 存在 4 阶矩， $i = 1,2,\cdots,$ 且当 $n \to \infty$ 时，有：

$$\frac{1}{n}\sum_{i=1}^{g-1}\sum_{k=1}^{h}\sum_{j=1}^{n-ih}[Cov(Z_{(i-1)h+k}Z_{ih+j}) + Cov(Z_{ih+j}Z_{(i-1)h+k})] = o(1), \quad (2.3.10)$$

$$\frac{1}{g^2h}\sum_{i=1}^{n}\|x_i\|^4 E\varepsilon_i^4 = o(1). \quad (2.3.11)$$

得到定理 2.3:

定理 2.3: 假设条件 (v)-(vii),(ix) 成立，则有：当 β 等于真参数 β_0 时，

$$\mathcal{I}_{E_1}(\beta_0) \xrightarrow{L} \chi_p^2\ , \qquad (n \to \infty). \quad (2.3.12)$$

2.3.2 几个引理

引理 2.3.1 若假设本节中的模型 (2.1) 中： ε_i 为 (2.3.1) 式和 (2.3.2) 式定义的线性过程，满足

$$\lim_{c\to\infty}\sup_j E(Y_j^2 I(|\ Y_j\ | > c)\ |\ \mathcal{F}_{j-1}) = 0,\ a.s.$$

并且 $\forall n \geqslant 1$, 有 $\mathbf{A_n} \triangleq \sum_{|k|\leqslant n-1} C(k)\sum_{i=1}^{n-|k|} x_i x_{i-|k|}^T$ 正定，且有

$$\lim_{n\to\infty}\frac{n\cdot\max\limits_{1\leqslant i\leqslant n}(\|x_i\|^2)}{\lambda_p(\mathbf{A_n})} = 0, \quad (2.3.13)$$

这里 $\lambda_p(\mathbf{A_n}) > 0$ 为 $\mathbf{A_n}$ 的最小特征根， $C(k) = Cov(\varepsilon_i, \varepsilon_{i+k})$.

则对 $\forall a = (a_1,a_2,\cdots,a_p)^T \in \mathbb{R}^p$, $\|\ \mathbf{a}\ \| = 1$, 都有：

$$\frac{a^T\mathbf{X_n}\varepsilon}{\sqrt{Var(a^T\mathbf{X_n}\varepsilon)}} \xrightarrow{L} \mathbf{N}(0,1) \qquad (n \to \infty).$$

证明：由于后面证明的需要，我们延用文献 [6] 里面的符号标记，并且与该文献中定理 2 的证明类似，令 $U_i = \sum_{j=1}^{p} a_j x_{ij} = a^T x_i$, 即有 $a^T \mathbf{X_n} \epsilon = \sum_{i=1}^{n} U_i \epsilon_i$. 取 $m = m(n) \uparrow \infty$, 记

$$\epsilon_i(m) = \sum_{j=-m}^{m} \psi_j Y_{i-j}, \qquad C(k) = Cov(\epsilon_i, \epsilon_{i+k}),$$

$$C_k(m) = Cov(\epsilon_i(m), \epsilon_{i+k}(m)), \quad \mathbf{s_n^2} = Var(\sum_{i=1}^{n} U_i \epsilon_i) = Var(a^T \mathbf{X_n} \epsilon),$$

$$\mathbf{s_n^2}(m) = Var(\sum_{i=1}^{n} U_i \epsilon_i(m)),$$

则

$$\mathbf{s_n^2} = \sum_{|k| \leqslant n-1} C(k) \sum_{i=1}^{n-|k|} U_i U_{i+|k|},$$

$$\mathbf{s_n^2}(m) = \sum_{|k| \leqslant n-1} C_k(m) \sum_{i=1}^{n-|k|} U_i U_{i+|k|}.$$

由胡舒合 [6] 可知

$$\lim_{n \to \infty} \mid \mathbf{s_n^2} - \mathbf{s_n^2}(m) \mid / (\max_{1 \leqslant j \leqslant n} \mid U_j \mid \sum_{i=1}^{n} \mid U_i \mid) = 0. \tag{2.3.14}$$

又因为

$$\begin{aligned} \frac{\mathbf{s_n^2}}{(\max\limits_{1 \leqslant j \leqslant n} \mid U_j \mid \sum\limits_{i=1}^{n} \mid U_i \mid)} &= \frac{\sum\limits_{|k| \leqslant n-1} C(k) \sum\limits_{i=1}^{n-|k|} U_i U_{i+|k|}}{(\max\limits_{1 \leqslant j \leqslant n} \mid U_j \mid \sum\limits_{i=1}^{n} \mid U_i \mid)} \\ &= \frac{\sum\limits_{|k| \leqslant n-1} C(k) \sum\limits_{i=1}^{n-|k|} a^T x_i x_{i+|k|} a}{(\max\limits_{1 \leqslant j \leqslant n} \mid U_j \mid \sum\limits_{i=1}^{n} \mid U_i \mid)} \\ &\geqslant \frac{a^T [\sum\limits_{|k| \leqslant n-1} C(k) \sum\limits_{i=1}^{n-|k|} x_i x_{i+|k|}] a}{n \cdot \max\limits_{1 \leqslant j \leqslant n} U_j^2} \end{aligned}$$

$$(\parallel a \parallel = 1) \quad \geqslant \quad \frac{\lambda_p(\mathbf{A_n})}{n \cdot \max\limits_{1\leqslant i\leqslant n}(\|x_i\|^2)}.$$

所以由 (2.3.13) (2.3.14) 式可得

$$\lim_{n\to\infty} \mathbf{s_n^2}(m)/\mathbf{s_n^2} = 1. \tag{2.3.15}$$

再记

$$T_n(m) = \sum_{i=1}^{n} U_i\epsilon_i(m)/\mathbf{s_n}(m), \qquad T_n = \sum_{i=1}^{n} U_i\epsilon_i/\mathbf{s_n},$$
$$C(n,m,i) = \sum_{j=\max(1-i,-m)}^{min(m,n-i)} U_{i+j}\psi_j/\mathbf{s_n}(m).$$

则同胡舒合 [6], 由 (2.3.13) (2.3.15) 以及 (2.3.3) 式可得

$$T_n(m) \xrightarrow{L} \mathbf{N}(0,1) \qquad (n\to\infty).$$

同样可由 Cr 不等式及 (2.3.13) 可得

$$\lim_{n\to\infty} E(T_n - T_n(m))^2 = 0,$$

从而得到

$$T_n \xrightarrow{L} \mathbf{N}(0,1) \qquad (n\to\infty),$$

即

$$\frac{a^T\mathbf{X_n}\varepsilon}{\sqrt{Var(a^T\mathbf{X_n}\varepsilon)}} \xrightarrow{L} \mathbf{N}(0,1) \qquad (n\to\infty).$$

证毕.

引理 2.3.2 对于本节中的模型 (2.1), 令 $Y_n = \max\limits_{1\leqslant i\leqslant n}\|Z_i\|$, $V_n = \max\limits_{1\leqslant i\leqslant g}\|\xi_i\|$, 若条件 (viii) 式成立, 则当 β 等于真参数 β_0 时, $Y_n = o_p(n^{\frac{1}{2}})$. 而若条件 (ix) 中第二个式子成立, 则此时, $V_n = o_p(\sqrt{g/h})$.

证明：当 β 等于真参数 β_0 时，因为 $Z_i = x_i\varepsilon_i$，$i = 1, 2, \cdots$，由契毕雪夫不等式得，

$$\begin{aligned} P(Y_n > n^{\frac{1}{2}}) &= P(\max_{1\leqslant i\leqslant n} \|Z_i\| > n^{\frac{1}{2}}) \leqslant \sum_{i=1}^{n} P(\|Z_i\| > n^{\frac{1}{2}}) \\ &\leqslant \frac{1}{n^2}\sum_{i=1}^{n} \|x_i\|^4 E\varepsilon_i^4, \end{aligned}$$

$$\begin{aligned} P(V_n > \sqrt{\frac{g}{h}}) &= P(\max_{1\leqslant i\leqslant g} \|\xi_i\| > \sqrt{\frac{g}{h}}) \leqslant \sum_{i=1}^{g} P(\|\xi_i\| > \sqrt{\frac{g}{h}}) \\ &\leqslant \frac{h^2}{g^2}\sum_{i=1}^{g} E\|x_i\| = \frac{1}{n^2}\sum_{i=1}^{g} E\|\sum_{j=1}^{h} Z_{(i-1)h+j}\|^4 \\ &\leqslant \frac{1}{n^2}\sum_{i=1}^{g} hC_1\sum_{j=1}^{h} E\|Z_{(i-1)h+j}\|^4 = \frac{C_1}{g^2h}\sum_{i=1}^{n} E\|Z_i\|^4 \\ &= \frac{C_1}{g^2h}\sum_{i=1}^{n} \|x_i\|^4 E\varepsilon_i^4. \end{aligned}$$

其中 C_1 为一常数，　则由 (2.3.9) 式知 $Y_n = o_p(n^{\frac{1}{2}})$, 由条件 2.3.12 中第二个式子式知 $V_n = o_p(\sqrt{g/h})$. 证毕.

引理 2.3.3 对于本节中的模型 (2.1), 设 $S_{n_1} = \frac{1}{n}\sum_{i=1}^{n} Z_iZ_i^T$, $S_{n_2} = \frac{h}{g}\sum_{i=1}^{g} \xi_i\xi_i^T$. 若定理 2.2 中的条件成立，　则当 β 等于真参数 β_0 时，

$$S_{n_1} - ES_{n_1} \xrightarrow{P} 0 \quad (n \to \infty). \tag{2.3.16}$$

若定理 2.3 中的条件成立，　则当 β 等于真参数 β_0 时，

$$S_{n_2} - ES_{n_2} \xrightarrow{P} 0 \quad (n \to \infty). \tag{2.3.17}$$

证明：因为当 $\beta = \beta_0$ 时，　$S_{n_1} = \frac{1}{n}\sum_{i=1}^{n} Z_iZ_i^T = \frac{1}{n}\sum_{i=1}^{n} x_ix_i^T\varepsilon_i^2$. 所以，此时

$$S_{n_1} - ES_{n_1} = \frac{1}{n}\sum_{i=1}^{n} x_ix_i^T(\varepsilon_i^2 - E\varepsilon_i^2).$$

由 Stout,W.F.[7] 及 Hölder 不等式则有,

$$
\begin{aligned}
& E[(S_{n_1}-ES_{n_1})_{kl}]^2 \\
= \;& E[\frac{1}{n}\sum_{i=1}^{n}x_{ik}x_{il}(\varepsilon_i^2-E\varepsilon_i^2)]^2 \\
= \;& \frac{1}{n^2}E[\sum_{i=1}^{n}x_{ik}^2x_{il}^2(\varepsilon_i^2-E\varepsilon^2)^2+2\sum_{1\leqslant j<t\leqslant n}x_{jk}x_{jl}x_{tk}x_{tl}(\varepsilon_j^2-E\varepsilon_j^2)(\varepsilon_t^2-E\varepsilon_t^2)] \\
\leqslant \;& \frac{1}{n^2}\Big\{\sum_{i=1}^{n}x_{ik}^2x_{il}^2E(\varepsilon_i^2-E\varepsilon_i^2)^2 \\
+ \;& 4\sum_{1\leqslant j<t\leqslant n}|\psi_{t-j}|x_{jk}x_{jl}x_{tk}x_{tl}[E(\varepsilon_j^2-E\varepsilon_j^2)^2]^{\frac{1}{2}}[E(\varepsilon_t^2-E\varepsilon_t^2)^2]^{\frac{1}{2}}\Big\} \\
\leqslant \;& \frac{1}{n^2}\Big\{\sum_{i=1}^{n}x_{ik}^2x_{il}^2E(\varepsilon_i^2-E\varepsilon_i^2)^2 \\
+ \;& 2\sum_{j=1}^{n-1}\sum_{i=1}^{n-j}\mid\psi_j\mid x_{ik}x_{il}x_{(i+j)k}x_{(i+j)l}[E(\varepsilon_i^2-E\varepsilon_i^2)^2+E(\varepsilon_{i+j}^2-E\varepsilon_{i+j}^2)^2]\Big\} \\
\leqslant \;& \frac{1}{n^2}(1+4\sum_{j=1}^{n}|\psi_j|)\sum_{i=1}^{n}x_{ik}^2x_{il}^2E(\varepsilon_i^2-E\varepsilon_i^2)^2 \\
\leqslant \;& \frac{1}{n^2}(1+4\sum_{j=1}^{n}|\psi_j|)\sum_{i=1}^{n}x_{ik}^2x_{il}^2E\varepsilon_i^4 \\
\leqslant \;& \frac{1}{n^2}(1+4\sum_{j=1}^{n}|\psi_j|)\sum_{i=1}^{n}\|x_i\|^4E\varepsilon_i^4.
\end{aligned}
$$

又因为 $\sum\limits_{j=-\infty}^{+\infty}|\psi_j|<\infty$, 所以由 (2.3.9) 式即可得 $\beta=\beta_0$ 时,

$$E[(S_{n_1}-ES_{n_1})_{kl}]^2=o(1).$$

从而得到 $\beta=\beta_0$ 时:

$$S_{n_1}-ES_{n_1}\xrightarrow{P}0\quad(n\to\infty).$$

同样, $\beta=\beta_0$ 时,

$$E[(S_{n_2}-ES_{n_2})_{kl}]^2$$

$$
\begin{aligned}
&= \frac{h^2}{g^2}E[\sum_{i=1}^{g}(\xi_{ik}\xi_{il}-E\xi_{ik}\xi_{il})]^2 \\
&\leqslant \frac{Ch^2}{g^2}\sum_{i=1}^{g}E(\xi_{ik}\xi_{il}-E\xi_{ik}\xi_{il})^2 \\
&\leqslant \frac{Ch^2}{g^2}\sum_{i=1}^{g}E(\xi_{ik}^2\xi_{il}^2) \\
&\leqslant \frac{Ch^2}{g^2}\sum_{i=1}^{g}\sqrt{E\xi_{ik}^4E\xi_{il}^4} \\
&\leqslant \frac{C}{g^2h}\sum_{i=1}^{g}\sqrt{\sum_{j=1}^{h}x_{(i-1)h+j,k}^4E\varepsilon_{(i-1)h+j}^4\sum_{j=1}^{h}x_{(i-1)h+j,l}^4E\varepsilon_{(i-1)h+j}^4} \\
&\leqslant \frac{C}{g^2h}\sum_{i=1}^{g}\Big[\sum_{j=1}^{h}x_{(i-1)h+j,k}^4E\varepsilon_{(i-1)h+j}^4+\sum_{j=1}^{h}x_{(i-1)h+j,l}^4E\varepsilon_{(i-1)h+j}^4\Big] \\
&= \frac{C}{g^2h}\sum_{i=1}^{gh}(x_{ik}^4+x_{il}^4)E\varepsilon_i^4 \\
&\leqslant \frac{C}{g^2h}\sum_{i=1}^{n}\|x_i\|^4E\varepsilon_i^4,
\end{aligned}
$$

其中 C 为一常数. 则由 (2.3.11) 式即得 $\beta=\beta_0$ 时：

$$E[(S_{n_2}-ES_{n_2})_{kl}]^2=o(1).$$

从而得到 $\beta=\beta_0$ 时：

$$S_{n_2}-ES_{n_2}\xrightarrow{P}0\ ,\quad (n\to\infty).$$

证毕.

2.3.2　定理 2.2 和 2.3 的证明

定理 2.2 的证明：因为

$$\mathcal{I}_E(\beta)=-2\log\mathcal{R}(\beta)=2\sum_{i=1}^{n}\log(1+\lambda^TZ_i),$$

所以类似于 Owen[2] 中定理一的证明，由引理 2.3.2 可知

$$\mathcal{I}_E(\beta_0) = (\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i)^T S_{n_1}^{-1}(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i) + o_p(1). \tag{2.3.18}$$

又因为

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i = \frac{1}{\sqrt{n}}\mathbf{X_n}\varepsilon, \qquad Cov(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i) = \frac{1}{n}Cov(\mathbf{X_n}\varepsilon),$$

所以由引理 2.3.1 及 (2.3.7) 式可得，当 β 等于真参数 β_0 时，

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i \xrightarrow{L} N_p(0, A_o)\ , \quad (n \to \infty). \tag{2.3.19}$$

又因为

$$\begin{aligned} S_{n_1} &= \frac{1}{n}\sum_{i=1}^{n} Z_i Z_i^T &&= \frac{1}{n}\sum_{i=1}^{n} x_i^T x_i \varepsilon_i^2, \\ ES_{n_1} &= \frac{1}{n}\sum_{i=1}^{n} x_i^T x_i E\varepsilon_i^2 &&= \frac{1}{n}C(0)\sum_{i=1}^{n} x_i^T x_i, \end{aligned}$$

所以根据引理 2.3.3 中 (2.3.15) 式，且由 (2.3.8),(2.3.18) 和 (2.3.19) 式可得

$$\mathcal{I}_E(\beta_0) \xrightarrow{L} Z^T A_1^{-1} Z\ , \qquad (n \to \infty).$$

其中 $Z \sim N_p(0, A_o)$. 证毕.

定理 2.3 的证明：

类似于 Owen[2] 中定理一的证明，由引理 2.3.2 可得

$$\mathcal{I}_{E_1}(\beta_0) = (\sqrt{\frac{h}{g}}\sum_{i=1}^{g} \xi_i)^T S_{n_2}^{-1}(\sqrt{\frac{h}{g}}\sum_{i=1}^{g} \xi_i) + o_p(1). \tag{2.3.20}$$

又因为当 β 等于真参数 β_0 时，

$$\sqrt{\frac{h}{g}}\sum_{k=1}^{g} \xi_k = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i = \frac{1}{\sqrt{n}}\mathbf{X_n}\varepsilon,$$

所以由引理 2.3.1 及 (2.3.7) 式得，此时当 $n \to \infty$ 时，

$$\sqrt{\frac{h}{g}} \sum_{i=1}^{g} \xi_i \xrightarrow{L} N_p(0, A_o). \tag{2.3.21}$$

又因为此时

$$\begin{aligned} ES_{n_2} &= E(\frac{h}{g} \sum_{i=1}^{g} \xi_i \xi_i^T) = \frac{h}{g} \sum_{i=1}^{g} E(\xi_i \xi_i^T) \\ &= \frac{1}{n} \sum_{i=1}^{g} E[(Z_{(i-1)h+1} + \cdots + Z_{ih})(Z_{(i-1)h+1} + \cdots + Z_{ih})^T], \\ \frac{1}{n} Cov(X_n \varepsilon) &= \frac{1}{n} E[(Z_1 + \cdots + Z_n)(Z_1 + \cdots + Z_n)^T], \end{aligned}$$

所以有

$$\begin{aligned} & \frac{1}{n} Cov(X_n \varepsilon) - ES_{n_2} \\ = & \frac{1}{n} \sum_{i=1}^{g-1} \sum_{k}^{h} \sum_{j=1}^{n-ih} [Cov(Z_{(i-1)h+k} Z_{ih+j}) + Cov(Z_{ih+j} Z_{(i-1)h+k})]. \end{aligned}$$

所以根据 (2.3.7),(2.3.10) 中第一个式子及引理 2.3.3 中 (2.3.17) 式得，当 β 等于真参数 β_0 时，

$$S_{n_2} \xrightarrow{P} A_o \ , \quad (n \to \infty). \tag{2.3.22}$$

所以由 (2.3.20),(2.3.21) 和 (2.3.22) 式即得

$$\mathcal{I}_{E_1}(\beta_0) = -2 \log \mathcal{R}_1(\beta_0) \xrightarrow{L} \chi_p^2 \ , \qquad (n \to \infty).$$

证毕.

2.4　本章小结

本章，我们首先考虑的是误差为鞅差序列的线性模型，得到了参数 β 的对数经验似然比统计量的渐近 χ^2 性. 其次考虑的是误差为弱平稳线过程的线

性模型，得到了参数 β 的对数经验似然比统计量的渐近正态性，然后采用了分组经验似然的方法，进一步得到参数 β 的分组对数经验似然比统计量的渐近 χ^2 性.

第 3 章　数值模拟

本章将通过抽取随机数, 给出数值模拟的结果, 说明本文提出的推断方法的优良性. 本文仅对该模型给出数值模拟的结果.

$$y_i = x_i^T\beta + \varepsilon_i \qquad i = 1, 2, \cdots, \tag{1}$$

其中 $x_i = (x_{i1}, x_{i2}, \cdots x_{ip})^T$ 为 p 维的已知固定设计点列, β 为 p 维未知回归系数, $y_i \in \mathbb{R}$ 为响应变量, $\varepsilon_i \in \mathbb{R}$ 为随机误差, 这里假设 $\{\varepsilon_n, \mathcal{F}_n, n \geqslant 1\}$ 为 L_2 中的鞅差序列, 对任意的 n≥1, 有 $\mathbf{X_n} = (x_1\ x_2 \cdots\ x_n)$ 满足 $\mathbf{X_n X_n^T} = \sum_{i=1}^{n} x_i x_i^T$ 满秩.

3.1　经验似然与最小二乘法之间的比较

3.1.1　一些假定和计算

所要做的工作是对该模型的一维情况进行模拟, 给定样本容量 N, 取

$$x_i = 1, \quad , i = 1, 2, \cdots, N. \tag{3.1.1}$$

由于 $\{\varepsilon_i, \mathcal{F}_i, i \geqslant 1\}$ 为 L_2 中鞅差序列, 因此在抽取随机数时可以先抽取 L_2 中的鞅序列 $\{\xi_i, \sigma(\xi_i), i \geqslant 1\}$, 再得到 $\varepsilon_i = \xi_{i+1} - \xi_i, \quad i = 1, 2, \cdots, N.$

取 $\xi_i \sim N(o, \eta_i^2)$, $(\xi_i, \xi_{i+1}) \sim N(0, 0, \eta_i^2, \eta_{i+1}^2, \rho_i)$, $i = 1, 2, \cdots, N.$ 因此, 给定 ξ_i 时,

$$\xi_{i+1}|\xi_i \sim N(\rho_i \frac{\eta_{i+1}}{\eta_i}\xi_i, (1-\rho_i^2)\eta_{i+1}^2), \ i = 1, 2, \cdots, N. \tag{3.1.2}$$

由鞅性可知, $\rho_i = \frac{\eta_i}{\eta_{i+1}}$, $i = 1, 2, \cdots, N.$ 又因为 $\varepsilon_i = \xi_{i+1} - \xi_i, \quad \rho_i = \frac{\eta_i}{\eta_{i+1}}$, $i = 1, 2, \cdots, N.$ 所以 $E\varepsilon_i^2 = \sigma_i^2 = \eta_{i+1}^2 - \eta_i^2, \quad i = 1, 2, \cdots, N.$ 代入到 (3.1.2) 式即得到, 给定 ξ_i 时,

$$\xi_{i+1}|\xi_i \sim N(\xi_i, \sigma_i^2), \ i = 1, 2, \cdots, N. \tag{3.1.3}$$

并且

$$E\varepsilon_i^4 = 2\sigma_i^4. \tag{3.1.4}$$

假设

$$\eta_1^2 = 1, \quad \sigma_i^2 = 1, \quad i = 1, 2, \cdots, N.$$

则由 (3.1.1)-(3.1.4) 可知, 定理条件 (i)-(iv) 成立. 由此, 我们可以比较以下两种方法的置信区间.

(A): 根据胡舒合 [6], 由最小二乘法可得真参数 β_0 的估计 $\hat{\beta} = \sum_{i=1}^{n} x_i y_i / \sum_{i=1}^{n} x_i^2$, 并且有

$$(\hat{\beta} - \beta_0)/\sqrt{Var(\hat{\beta})} \xrightarrow{L} N(0,1).$$

因此 β_0 的置信度为 $1-\alpha$ 的置信区间为

$$\hat{\beta} \pm z_{1-\alpha/2}\left(\sqrt{Var(\hat{\beta})}\right). \tag{3.1.5}$$

其中 $Var(\hat{\beta}) = \sum_{i=1}^{n} x_i^2\sigma_i^2/(\sum_{i=1}^{n} x_i^2)^2$.
(B): 由定理 2.1 可构造真参数 β_0 的置信度为 $1-\alpha$ 的置信区域为

$$J_\alpha = \{\beta : \mathcal{I}_E(\beta) \le C_\alpha\}, \tag{3.1.6}$$

其中 C_α 满足 $P(\chi_p^2 \le C_\alpha) = 1-\alpha$. 并且由 Owen[1,2,3] 可知该置信域可近似地为一置信区间.

下面要比较两种方法所得覆盖概率和平均区间长度.

3.1.2 覆盖概率的比较

取样本容量 n=20 、 50 和 100, 对样本量的每种选择模拟重复 M 次. 模拟过程如下:

第一步：抽取第一个随机数 $\xi_1 \sim N(0, \eta_1^2)$. 然后按照条件分布

$$\xi_{i+1}|\xi_i \sim N(\xi_i, \sigma_i^2),\ i = 1, 2, \cdots, N$$

抽取剩下的 N 个随机数： $\xi_2, \xi_3, \cdots, \xi_{N+1}$.

第二步：取参数真值 $\beta_0 = 2$, 再根据 $\varepsilon_i = \xi_{i+1} - \xi_i,\ \ i = 1, 2, \cdots, N$. 及 (3.1.1) 式按照该模型计算得到 $\{y_i, i = 1, 2, \cdots, N\}$

第三步：按照 (3.1.5) 式，得到由方法 A 构造的参数 β 的置信区间. 同样，按照 (3.1.6) 式也可以求得由方法 B 构造的参数的置信区间.

第四步：将上述步骤重复 M 次，分别记录真值 β_0 属于以上两个的区间的次数，然后求出置信区间的覆盖概率 (M=10,000).

结果见表 3.1.1.

表 3.1.1

	$\alpha = 0.05$		$\alpha = 0.1$	
n	方法 A	方法 B	方法 A	方法 B
20	0.9230	0.9420	0.8990	0.9040
50	0.9480	0.9530	0.9110	0.9130
100	0.9520	0.9560	0.9210	0.9207

从模型的模拟结果可以得出下列结论：

(1) 即使在中小样本量的场合，由经验似然方法构造的模型中参数的置信区间比由最小二乘法构造的置信区间有更精确的区间覆盖概率.

(2) 随样本容量的增加，区间覆盖概率有相应的提高.

(3) 置信度越高区间覆盖率也越大，这也是和实际相符合的.

3.1.2　平均区间长度的比较

比较区间平均长度的模拟过程与 3.1.2 节的模拟过程大体相似，将第四步

改为求出区间长度，重复 M 次得区间的平均长度.

具体结果见表 3.1.2.

表 3.1.2

	$\alpha = 0.05$		$\alpha = 0.1$	
n	方法 A	方法 B	方法 A	方法 B
20	0.8765	0.6813	0.7334	0.5488
50	0.5544	0.2603	0.4639	0.1501
100	0.3920	0.1257	0.3280	0.0814

从模型的模拟结果可以得出下列结论：

(1) 在相同的置信水平下，由经验似然方法构造的置信区间平均长度要小于由最小二乘法得到的置信区间.

(2) 置信度越高平均区间长度越长.

(3) 随样本容量的增加，区间的平均长度也相应地减小了.

3.2 误差方差不同的情况

3.2.1 一些假定和计算

在本节中，我们仍对该模型的一维情况进行模拟. 如同 3.1.1, 我们仍做类似的假设，

$$x_i = 1, \quad , i = 1, 2, \cdots, N.$$

$$\xi_i \sim N(o, \eta_i^2),\ (\xi_i, \xi_{i+1}) \sim N(0, 0, \eta_i^2, \eta_{i+1}^2, \rho_i),\ i = 1, 2, \cdots, N.$$

以及 $\varepsilon_i = \xi_{i+1} - \xi_i, \quad i = 1, 2, \cdots, N.$

同样通过计算可得，给定 ξ_i 时，

$$\xi_{i+1}|\xi_i \sim N(\xi_i, \sigma_i^2),\ i = 1, 2, \cdots, N$$

这里我们给予 σ_i^2 不同的假设，并且满足条件 (i)-(iv),

$$B_1: \quad \eta_1^2 = 1, \sigma_i^2 = 1\ ; \qquad B_2: \eta_1^2 = 10, \sigma_i^2 = 1$$

$$B_3: \quad \eta_1^2 = 1, \sigma_i^2 = 10\ ; \qquad B_4: \eta_1^2 = 1, \sigma_i^2 = \frac{1}{10}$$

3.2.2　模拟结果及解释

我们对以上四种情况按 3.1.2 及 3.1.3 中的步骤进行模拟，结果见表 3.2.1-3.2.4 .

表 3.2.1　　B_1, B_2 覆盖概率的比较

	$\alpha = 0.05$				$\alpha = 0.1$			
n	AB_1	AB_2	BB_1	BB_2	AB_1	AB_2	BB_1	BB_2
20	0.9230	0.9142	0.9420	0.9205	0.8990	0.8815	0.9040	0.8937
50	0.9480	0.9367	0.9530	0.9413	0.9110	0.9073	0.9130	0.9001
100	0.9520	0.9489	0.9560	0.9481	0.9210	0.9157	0.9207	0.9163

表 3.2.2　　B_1, B_2 平均区间长度的比较

	$\alpha = 0.05$				$\alpha = 0.1$			
n	AB_1	AB_2	BB_1	BB_2	AB_1	AB_2	BB_1	BB_2
20	0.8765	0.8765	0.6813	0.7437	0.7334	0.7334	0.5488	0.6143
50	0.5544	0.5544	0.2603	0.4152	0.4639	0.4639	0.1501	0.3003
100	0.3920	0.3920	0.1257	0.2109	0.3280	0.3280	0.0814	0.1127

表 3.2.3　　B_3, B_4 覆盖概率的比较

	$\alpha = 0.05$				$\alpha = 0.1$			
n	AB_3	BB_3	AB_4	BB_4	AB_3	BB_3	AB_4	BB_4
20	0.9130	0.9210	0.8830	0.8997	0.8715	0.8957	0.8517	0.8641
50	0.9511	0.9639	0.9001	0.9173	0.9127	0.9239	0.8914	0.9035
100	0.9723	0.9814	0.9132	0.9301	0.9431	0.9513	0.9001	0.9197

表 3.2.4　　B_3, B_4 平均区间长度的比较

	$\alpha = 0.05$				$\alpha = 0.1$			
n	AB_3	BB_3	AB_4	BB_4	AB_3	BB_3	AB_4	BB_4
20	8.7654	6.1374	0.0877	0.0737	7.3343	5.1107	0.0733	0.0694
50	5.5437	4.0139	0.0554	0.0413	4.6386	3.9714	0.0464	0.0371
100	3.9200	2.5750	0.0392	0.0271	3.2800	2.4713	0.0328	0.0219

(1). 从表 3.2.1-3.2.2 可以看出，同样的经验似然方法构造的置信区间比最小二乘方法构造的置信区间相比，仍然覆盖概率较大，平均区间长度较短，同时 B_1 比 B_2 的覆盖概率要大，相应的区间长度也较短，特别是在样本容量较小时，差别比较明显. 区别就在于 η_1^2 根据 (3.1.3) 式可以看到，在给定 ξ_i 时，B_2 与 B_1 中 ξ_{i+1} 的条件分布是一致的，但对于 ξ_1 来说，B_1 中的方差要比 B_2 中的方差小，从而造成了表 3.2.1-3.2.2 中的结果.

(2). 从表 3.2.3-3.2.4 可以看出，B_4 比 B_3 的覆盖概率要小很多，但同时平均区间长度也要下. 这是由于在 B_4 中，σ_i^2 要远远小于在 B_3 中的 σ_i^2, 从而使得在模拟中抽取的随机数过度依赖 ξ_1, 造成表中的模拟结果.

(3). 同时我们还可以看出，在 B_3 情况下，经验似然方法得到的覆盖概率不仅大于最小二乘方法的覆盖概率，平均区间长度也远远小于最小二乘方法的，充

分体现了经验似然方法的优越性.

3.3　本章小结

综合表 (3.1.1)-(3.1.2) 、(3.2.1)-(3.2.4) 可知, 用经验似然法构造基于鞅差误差的线性模型中参数 β 的置信区间比用最小二乘法构造的置信区间有更精确的覆盖概率, 且所得区间的平均长度比用最小二乘法得到的区间长度短, 并且具有较高的稳定性. 因而可以说, 用经验似然方法研究半参数回归模型有着重要的意义.

后　记

本文异于一般的参数估计方法，在面对误差为鞅差序列的线性模型时，首先给出了对数经验似然比统计量，在给定的假设条件下，证明了这个统计量的渐近分布. 其次，针对误差为弱平稳线性过程的线性模型，不仅给出了对数经验似然比统计量，还给出了分组对数经验似然比统计量，并在给定的假设条件下，分别得出了它们的渐近分布. 由此可以构造参数 β 的置信域. 主要结论为：

定理 2.1: 假设条件 (i) – (iv) 成立， 则有当 β 等于真参数 β_0 时，

$$\mathcal{I}_E(\beta_0) \xrightarrow{L} \chi_p^2 \qquad (n \to \infty).$$

定理 2.2: 假设条件 (v)-(viii) 成立，则可以得到结论：当 β 等于真参数 β_0 时，

$$\mathcal{I}_E(\beta_0) \xrightarrow{L} Z^T A_2^{-1} Z \qquad (n \to \infty).$$

其中 $Z \sim N_p(0, A_1)$.

定理 2.3:: 假设条件 (v)-(vii),(ix) 成立，则有：当 β 等于真参数 β_0 时，

$$\mathcal{I}_{E_1}(\beta_0) \xrightarrow{L} \chi_p^2 \qquad (n \to \infty).$$

最后，利用 Monte-carlo 模拟，抽取随机数的方法，同样也发现了我们的结论的良好性质. 即：用经验似然方法构造的置信区间比用最小二乘法构造的置信区间有更高的区间覆盖率和更短的平均区间长度.

虽然我们在一定程度上说明了经验似然方法的优越性质. 但仍有很多问题值得进一步思考. 主要有以下几方面的问题：

问题 1　本文在考虑模型的误差项时是两类特定的相依过程，虽然得到的结果比较理想，但毕竟过于具体，不具有很大的普遍性. 因此我们可以进一步考虑比较一般的相依误差.

问题 2　本文考虑的模型是固定设计的线性模型，我们还可以进一步考虑其它类型的模型，比方说非参数模型和半参数模型等.

可见，本文仅仅是这方面研究中的一个基本情形，仍旧有许多问题值得我们去思考解决。

问题 3　在进行模拟计算时，还可以考虑给出对数经验似然比统计量的若干分位点并和极限分布 χ^2 分布的分位点进行比较，观察对数经验似然比统计量与 χ^2 分布的拟合程度.

参 考 文 献

1 Owen,A.B.Emprical Likelihood Ratio Confidence Intervals for a Single Functional.Biometrika. 75(1988),237–249.

2 Owen,A.B.Empirical Likelihood Confedence Regions.Ann.Statist.18(1990), 90–120.

3 Owen,A.B.Empirical Likelihood for Linear Models.Ann.Statist.19(1991), 1725–1747.

4 张军舰、王成名、王炜烯，相依样本情形是的经验似然比置信区间，高校应用数学学报， 14(1999),63–72.

5 张军舰、王成名、王炜烯， m 相依样本下参数及分布函数的经验似然估计，广西师范大学学报， 17(1999),21–27.

6 胡舒合, 相依误差下线性模型参数估计的渐近正态性, 科学通报, 23(1998), 2489–2493.

7 Stout,W.F.Almost Sure Convergence.Academic Press.New York,1974.

8 Peligrad,M.Invariance Principles for Mixing Sequence of Random Variables. Ann.Probab.10(1982),968–981.

9 Thomas DiCiccio,Peter Hall,Joseph Romano.The Annals of Statistics. 19(1991),1053–1061.

10 Kolaczyk,E.D.Empirical Likelihood for Generalized Linear Models[J]. Statistics Sinica.4(1994),199–218.

11 Jing Qin.Empirical Likelihood in Based Sample Problems.The Annals of Statistics.21(1993),1182–1196.

12 Zhang Biao.Quantile Processes in the Presence of Auxiliary Information[J].

Annals of the Institute of Statistical Mathematics.49(1997),35–55.

13 Chuang C-S & Chen Songxi.Empirical Likelihood Confidence Regions for Parameters in Error-in-variables Models[J].Journal of Multivariate Analysis.

84(2002)101–115.

14 Zhong Bob and Rao,JNK.Empirical likelihood inference under stratified random sampling using auxiliary population information [J].Biometrika, 2000,87:929–938.

15 Kitamura Y.Asymptotic optimality of empirical likelihood for testing moment restrictions [J].Econometrica,2001,69:1661–1672.

16 Gang Li and Qi-Hua,Wang. Emipirical likelihood regression analysis for right censored data. Statistica Sinica 2003,13: 51–68.

17 WANG,Q.H.and RAO,J.N.K. Empirical likelihood for linear error–in–covariables models with validation data.Biomatrika, 2002,89:354–358 .

18 秦永松，Empirical Likelihood Ratio Confidence Regions in a Partly Linear Model, 应用概率统计， 15(1999), 363–369.

19 Qihua Wang, Bingyi Jing. Empirical Likelihood for Partical Linear Models with Fixed Designs. Statistics & Probability Letters, 41(1999), 425–433.

20 Yudi Pawitan. Computing Empirical Likelihood from the Bootstrap. Statistics & Probability Letters, 47(2000), 337–345.

21 Jian Shi. Empirical Likelihood for Partially Linear Models. Journal of Multivariate Analysis, 72(2000), 132–148.

22 Jahua Chen,Jin Qin, Empirical Likelihood Estimation for Finite Popula-

tions and the Effective Usage of Auxiliary Information.Biometrika.80(1993), 107–116.

23 B.W.Silverman, Density Ratios.Empirical Likelihood and Cot Death. Applied Statistics.27(1978), 26–33.

24 Thomas J.DiCiccio, Peter Hall, Joseph P.Romano. Comparison of Parametric and Empirical Likelihood Functions. Biometrika, 76(1989), 465–476.

25 Jin Qin, Jerry Lawless. Empirical Likelihood and General Estimating Equations. Ann. Statist, 22(1994), 300–325.

26 Anna Clara Monti. Empirical Likelihood Confidence Regions in Time Series Models. Biometrika, 84(1997), 395–405.

27 Kerth A. Baggerly. Empirical Likelihood as a Goodness–of–Fit Measure. Biometrika, 85(1998), 535–547.

28 Gengsheng Qin, Bingyi Jing. Censored Partial Linear Models and Empirical Likelihood. Journal of Multivariate Analysis, 78(2001), 37–61.

29 Lu Lin, Runchu Zhang. Blockwise Empirical Euclidean Likelihoood for Weakly Dependent Processes. Statistics & Probability, 53(2001), 143–152.

30 Stephen A. Corcoran. Bartlett Adjustment of Empirical Discrepancy Statistics. Biometrika, 85(1998), 967–972.

31 Nicole Lazar, Per Aslak Mykland. An Evaluation of the Power and Conditionality Properties of Empirical Likelihood. Biometrika, 85(1998), 523–534.

32 S.A.Murphy.Likelihood Ratio–Baesd Confidence Intervals in Survival Analysis.Journal of the American Statistical Association.90(1995),1399–1405.

33 David E.Matthews.Likelihood–Based Confidence Intervals for Functions of Many parameters.Biometrika.75(1988),139–144.

34 Gang Li,Myles Hollander,Jan W.MeKeague,Jie Yang.Nonparametric Likelihood Ratio Confidence Bands for Quantile Functions from Incomplete Survival Data.The Annals of Statistics.24(1996),628–640.

35 Gang Li.Nonparametric Likelihood Ratio Estimation of Probabilities for Truncated Data.Journal of the American Statistical Association.90(1995), 997–1003.

36 Xue Liugen,Zhu Lixing.Emprical Likelihood Confidence Regions of the Parameters in a Partially Linear Single–index Model.Science in China Ser.A Mathematics.48(2005),1333–1348.

37 彭国强, 混合样本的经验似然比检验, 广西师范大学学报, 16(1998), 222–227.

38 胡舒合, ϕ – 混合、α – 混合序列和的强大数律, 工程数学学报, 9(1992),57–63.

39 薛留根, 相依样本下条件密度的双重核估计, 高校应用数学学报, 6(1991), 592–603.

40 林路, 相依数据的若干统计模型及分析, 南开大学博士学位论文, 2001.

41 石坚, 线性相关模型中误差方差的经验似然估计及其 Bootstrap, 数学物理学报, 17(1997), 38–46.

42 高秀红、崔恒建, EV 线性模型参数的经验似然比置信区域, 北京师范大

学学报， 37(2001), 586–591.

43 姜波、秦永松， M– 泛函的经验似然置信区间，广西科学， 11(2004), 4–9.

44 张军舰、王成名，光滑经验似然分布估计，广西师范大学学报， 16(1998), 37–41.

45 石坚， Bootstrapping Regression Models via Empirical Likelihood Resampling, 应用概率统计， 13(1997), 37–44.

46 秦永松、赵林成，有偏模型中一类统计泛函的经验似然估计及其渐近性质，应用数学学报， 21(1998), 428–436.

47 李国英,关于高维、相依和不完全数据的统计分析,数学进展, 31(2002),193–199.

48 王启华, 经验似然统计推断方法发展综述, 数学进展， 33(2004), 141–151.

49 Biao Zhang. Empirical Likelihood Confidence Intervals for M–functionals in the Presence of Auxiliary Information. Statistics & Probability Letters, 32(1997), 87–97.

50 王松桂. 线性模型的理论及其应用. 安徽教育出版社， 1986.

51 茆诗松，王静龙，濮晓龙. 高等数理统计. 高等教育出版社，施普林格出版社. 1998.

52 陈希孺. 高等数理统计，中国科学技术大学出版社. 1999.2.

53 Admari,G.Empirical likelyhood type confidence intervals under random sensorship J. Ann.Statist,Math,1997 ,49(3),447–456.

54 林正炎，陆传荣，苏中根. 概率极限理论基础. 高等教育出版社. 1998.

55 李人厚, 张平安. 精通 MATLAB 综合辅导与指南. 西安交通大学出版社. 1999.

附录　主要算法源程序

模拟的主要程序如下：

```
function [X,Y,a,cdu]=gg(n);
n=20;
X=1:n;
Y=1:n;
F=1:n+1;
for i = 1:n
    X(i)=1/(sqrt(i));
end
F(1)=1;
for i = 1:n+1
    F(i)=i;
end
E=1:n+1
e=1:n
E(1)=normrnd(0,F(1));
for i = 2:(n+1)
    E(i)=normrnd(E(i-1),F(i));
end
for i = 1:n
    e(i)=E(i+1)-E(i)
    Y(i)=2*X(i)+e(i);
    Sigma2(i)=X(i)^2;
    Sigma3(i)=X(i)*Y(i);
```

```
end
    guji=mean(Sigma3)/mean(Sigma2);
for i = 1:n
    Sigma1(i)=X(i)^2*F(i)^2
end
sigma=mean(Sigma1)/(n*((mean(Sigma2))^2));
lowermu=guji-1.64*sqrt(sigma);
uppermu=guji+1.64*sqrt(sigma);
cdu=2*1.64*sqrt(sigma);
if (2>lowermu) & (2<uppermu);
    a=1;
else a=0;end


function b=emp(n);
n=20;
X=1:n;
Y=1:n;
for i = 1:n
    [X,Y]=gg;
    w(i)=X(i)*(Y(i)-X(i)*2);
    h(i)=w(i)*w(i);
end
ww=mean(w);
S=1/[mean(h)];
for j = 1:n
```

```
    r(j)=log(1+S*ww*w(j));
end
R=2*sum(r);
k1=3.841;
k2=0.004;a=0
if (R<k1) & (R>k2);cdu=0;
    b=1b=0;
else b=0;end

a=0
cdu=0;
b=0;
n=20;
X=1:n;
Y=1:n;
m=10000;
t1=0;
t2=0;
t3=0;
for i =1:m
    [X,Y,a,cdu]=gg;
    b=emp;
    t1=t1+a;
    t2=t2+cdu;
    t3=t3+b;
end
```

```
t1=t1/m
t2=t2/m
t3=t3/m

function [R,X,Y,a,cdu] = empirical1(B)
n=20;
X=1:n;
Y=1:n;
a=0;
cdu=0;
[X,Y,a,cdu]=gg;
for i = 1:n;
    w(i)=X(i)*(Y(i)-X(i)*B);
    h(i)=w(i)*w(i);
end
ww=mean(w);
S=1/[mean(h)];
for j = 1:n
    r(j)=log(1+S*ww*w(j));
end
R=2*sum(r)-3.841;

k=0.05;
n=20;
m=10000;
for i = 1:m
```

```
    maxz1(i)=fzero(@empirical1,2);
    k1(i)=fzero(@empirical1,2+k);;
    k2(i)=fzero(@empirical1,2-k);
end
c=mean(maxz1)-mean(k1);
d=mean(maxz1)-mean(k2);
if abs(c)>abs(d);
    answer=abs(c)
else answer=abs(d);end
c
d
```

致　谢

三年的研究生生涯即将结束，这段难忘的时光构成了我人生重要的篇章．当我走出校门时，不禁想起了入校时，那时专业基础薄弱，是这里老师们的辛勤工作和同学们的热情帮助使我有了很大的进步．

首先，要感谢我的导师薛老师三年来对我学业上孜孜不倦的培育，生活上不断的关心．特别是在对我论文撰写过程中给予的悉心指导．导师开拓性的思维方式和谆谆教导带给我许多启迪，使我克服了困难，圆满完成了毕业论文的撰写．此外，薛老师严谨的治学态度，广博的学识，敏捷的思维，以及热情的工作精神都深深地影响着我，是我以后学习的榜样．三年中，使我学到了很多必备的专业知识，无论是从专业学习的角度，还是个人知识与能力的积累方面，都使我受益匪浅．同时，感谢我的师母张辉霞老师对我的热心关怀！

感谢王松桂教授、杨振海教授、张忠占教授、李寿梅教授、程维虎教授以及统计专业各位老师对我的教育和培养！老师们兢兢业业，一丝不苟地从事教学科研工作，使我不仅在学习中丰富了知识，开拓了视野，更重要的是从老师身上学到了如何对待学习、工作和人生的态度．

感谢实验室的兄弟姐妹们三年来对我的关心和帮助，愿我们的友谊天长地久！

感谢我的父母亲，他们一直站在我身后，给了我最大的支持和鼓舞！

此外，向所有曾经帮助和鼓励过我的同学以及亲人表示感谢！