

现代统计模型

薛留根 著



科学出版社

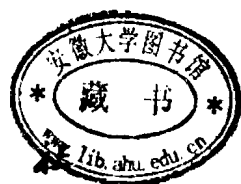
现代数学基础丛书 145

现代统计模型

薛留根 著

科学出版社

北京



内 容 简 介

本书阐述了半参数回归模型的统计理论和方法,所考虑的模型包括部分线性模型、单指标模型、变系数模型和可加模型等.这些模型对复杂数据分析起着重要作用.本书在取材上侧重内容的科学性和应用性,体现学术思想;在写作上注重阐述方法论、模拟计算和实例分析;在结构上安排每个模型为一章.本书的内容不仅为从事该领域的科研人员提供了尽可能全的资料,又为实际应用者提供了一些数据分析的方法,同时也为想全面了解现代统计模型的读者提供参考读物.

本书可以作为高等院校有关专业的教学用书,同时也对高等院校和科研机构的研究人员、工程技术人员和研究生有参考价值.

图书在版编目(CIP)数据

现代统计模型/薛留根著. —北京:科学出版社, 2012

ISBN 978-7-03-035020-6

(现代数学基础丛书; 145)

I. ①现… II. ①薛… III. ①统计模型—研究 IV. ①C8

中国版本图书馆 CIP 数据核字 (2012) 第 133253 号

责任编辑:陈玉琢/责任校对:朱光兰

责任印制:钱玉芬/封面设计:王浩

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

骏立印刷厂印刷

科学出版社发行 各地新华书店经销

*

2012年6月第一版 开本:B5(720×1000)

2012年6月第一次印刷 印张:26 1/4

字数:500 000

定价:88.00 元

(如有印装质量问题,我社负责调换)

《现代数学基础丛书》序

对于数学研究与培养青年数学人才而言，书籍与期刊起着特殊重要的作用。许多成就卓越的数学家在青年时代都曾钻研或参考过一些优秀书籍，从中汲取营养，获得教益。

20世纪70年代后期，我国的数学研究与数学书刊的出版由于文化大革命的浩劫已经破坏与中断了10余年，而在这期间国际上数学研究却在迅猛地发展着。1978年以后，我国青年学子重新获得了学习、钻研与深造的机会。当时他们的参考书籍大多还是50年代甚至更早期的著述。据此，科学出版社陆续推出了多套数学丛书，其中《纯粹数学与应用数学专著》丛书与《现代数学基础丛书》更为突出，前者出版约40卷，后者则逾80卷。它们质量甚高，影响颇大，对我国数学研究、交流与人才培养发挥了显著效用。

《现代数学基础丛书》的宗旨是面向大学数学专业的高年级学生、研究生以及青年学者，针对一些重要的数学领域与研究方向，作较系统的介绍。既注意该领域的基础知识，又反映其新发展，力求深入浅出，简明扼要，注重创新。

近年来，数学在各门科学、高新技术、经济、管理等方面取得了更加广泛与深入的应用，还形成了一些交叉学科。我们希望这套丛书的内容由基础数学拓展到应用数学、计算数学以及数学交叉学科各个领域。

这套丛书得到了许多数学家长期的大力支持，编辑人员也为其付出了艰辛的劳动。它获得了广大读者的喜爱。我们诚挚地希望大家更加关心与支持它的发展，使它越办越好，为我国数学研究与教育水平的进一步提高做出贡献。

杨 乐

2003年8月

前 言

在实际问题中,人们常常遇到各种各样的数据,如独立数据、纵向数据、缺失数据、删失数据、测量误差数据等.统计模型是对数据内在规律性的数学描述.由于所研究的客观现象越来越复杂,要求也越来越精细,回归模型已从线性模型发展到各种非线性回归模型、非参数和半参数回归模型.随着科学技术的快速发展,半参数回归模型对复杂数据的分析起着重要应用,其中包括部分线性模型、单指标模型、变系数模型和可加模型等.这类模型通过引入参数分量和非参数函数,能够真正组合数据的相关信息,从而提高了模型的拟合优度.半参数回归模型是统计学中重要的模型,它们在信息科学、生物医学、金融工程、经济管理、质量控制、交通工程、能源与环境、人文与社会科学等领域都有着相当广泛的应用.近年来,国内外许多学者致力于对这些模型的研究,现已成为统计学的热门研究课题之一.

本书系统阐述了半参数回归模型的统计推断方法,主要内容包括估计方法、模型检验和变量选择等.本书不但介绍了各种模型的发展动态,而且展示了最新的研究成果.本书的内容不仅为从事该领域的科研人员提供了尽可能全的资料,又为实际应用者提供了一些数据分析方法,同时也为想全面了解现代统计模型的读者提供参考读物.

本书根据作者多年从事科研工作所积累的大量资料、并结合本人的研究成果撰写而成.在取材上精心提炼,既考虑到了深度与广度,又顾及了内容的科学性和应用性,体现学术思想.在写作上注重阐述方法论、模拟计算和实例分析,省略或简化理论证明.本书结构是把每个模型作为一章,读者可以独立阅读每一章内容.初学者可以避开阅读理论部分.

本书得到了国家自然科学基金(11171012)、北京市优秀博士学位论文指导教师科技项目(20111000503)和北京市教育委员会科技计划项目(KM201110005029)的资助,作者谨在此表示感谢.

薛留根

2012年3月

符号表

iid	独立同分布
$\mathbf{R}^p$	p 维 Euclid 空间
$\ \cdot\ $	Euclid 模
A^T	向量或矩阵 A 的转置
$\text{tr}(A)$	矩阵 A 的迹
$I_B(\cdot)$ 或 $I(B)$	集合 B 的示性函数
$A \otimes B$	矩阵 A 与 B 的 Kronecker 乘积
$A^{\otimes 2}$	AA^T
$\text{diag}\{a_1, \dots, a_n\}$	由元素 $a_1, \dots, a_n$ 组成的对角矩阵
$\xrightarrow{D}$	依分布收敛
$\xrightarrow{P}$	依概率收敛
a.s.	几乎处处收敛
$\equiv$	“定义为”或“记为”
$y = O(1)$	y 是有界变量
$y = o(1)$	y 是无穷小量
$\xi_n = o_p(\eta_n)$	对任一 $\varepsilon > 0$, 有 $P\{\ \xi_n\ \geq \varepsilon\ \eta_n\ \} \rightarrow 0$
$\xi_n = o_P(1)$	ξ_n 依概率收敛到 0
$\xi_n = O_P(1)$	ξ_n 依概率有界
$f'(t)$	函数 $f(t)$ 的一阶导数
$N(\mu, \Sigma)$	均值为 μ , 协方差阵为 Σ 的正态分布
χ_p^2	自由度为 p 的 χ^2 分布
$\chi_p^2(1 - \alpha)$	自由度为 p 的 χ^2 分布的 $1 - \alpha$ 分位数
$z_{1-\alpha/2}$	标准正态分布的 $1 - \alpha/2$ 分位数
x_+	变量 X 的正部
c	不依赖于 n 的正的常数, 在不同地方可以表示不同的值

目 录

前言

符号表

第 1 章	预备知识	1
1.1	回归模型简介	1
1.1.1	参数回归模型	1
1.1.2	非参数回归模型	2
1.1.3	半参数回归模型	3
1.2	光滑方法	4
1.2.1	核光滑	4
1.2.2	局部多项式光滑	5
1.2.3	样条光滑	7
1.2.4	权函数法与其他估计	10
1.3	选择光滑参数	10
1.3.1	交叉验证	11
1.3.2	广义交叉验证	11
1.4	经验似然	12
1.4.1	经验似然的思想	12
1.4.2	参数的经验似然	13
1.4.3	估计方程与经验似然	14
1.4.4	主要成果与文献注记	14
1.5	惩罚方法	16
1.5.1	惩罚函数	16
1.5.2	惩罚估计与变量选择	18
1.5.3	主要成果与文献注记	19
1.6	复杂数据简介	20
1.6.1	纵向数据	20
1.6.2	缺失数据	22
1.6.3	删失数据	23
1.6.4	测量误差数据	24
	参考文献	24

第 2 章 部分线性模型	31
2.1 引言	31
2.2 估计方法	37
2.2.1 权函数方法	37
2.2.2 偏光滑样条方法	38
2.2.3 一般光滑方法	38
2.2.4 分段多项式方法	40
2.2.5 拟似然方法	41
2.2.6 经验似然方法	43
2.2.7 模拟	44
2.3 删失数据分析	44
2.3.1 估计方法	44
2.3.2 经验似然	46
2.4 测量误差数据分析	48
2.4.1 校正的最小二乘估计	48
2.4.2 广义最小二乘估计	51
2.4.3 分位数回归估计	51
2.4.4 经验似然	53
2.5 缺失数据分析	54
2.5.1 缺失响应下回归系数的估计	54
2.5.2 缺失响应下基准函数的经验似然	59
2.5.3 缺失响应下响应均值的估计	63
2.5.4 缺失协变量下回归系数的估计	64
2.5.5 模拟研究	68
2.5.6 定理的证明	73
2.6 纵向数据分析	86
2.6.1 模型	86
2.6.2 后移算法	87
2.6.3 profile 核估计方程	87
2.6.4 样条逼近	88
2.6.5 经验似然	90
2.6.6 计数过程方法	104
2.7 模型检验	106
2.8 变量选择	110
2.8.1 惩罚最小二乘	110

2.8.2 惩罚分位回归	113
2.8.3 惩罚多项式样条	113
参考文献	114
第 3 章 单指标模型	124
3.1 引言	124
3.2 半参数最小二乘估计	127
3.3 半参数极大似然估计	129
3.4 平均导数估计	131
3.5 切片逆回归估计	135
3.5.1 多指标模型	135
3.5.2 逆回归曲线	136
3.5.3 运算步骤	137
3.6 最小平均方差估计	139
3.6.1 有效降维方向的估计	139
3.6.2 有效降维空间的维数	142
3.6.3 带宽和算法	142
3.7 惩罚样条估计	144
3.8 Bayes 估计	144
3.8.1 建模构想	145
3.8.2 调整	146
3.8.3 MCMC 算法	147
3.9 分位数回归估计	147
3.10 估计方程估计	150
3.10.1 估计方法	150
3.10.2 主要结果	152
3.10.3 模拟研究	155
3.10.4 定理的证明	156
3.11 拟似然估计和估计函数方法估计	170
3.12 经验似然推断	174
3.12.1 估计的经验似然	174
3.12.2 调整的经验似然	176
3.12.3 纠偏的经验似然	177
3.12.4 模拟研究	178
3.12.5 定理的证明	181
3.13 拟合优度检验	185

3.13.1 Cramér-von Mises 检验	185
3.13.2 bootstrap 方法	187
3.13.3 得分检验	189
3.14 变量选择	192
3.14.1 Leave- m -out 交叉验证	192
3.14.2 分离式变量选择	194
参考文献	196
第 4 章 部分线性单指标模型	203
4.1 引言	203
4.2 最小二乘估计和拟似然估计	205
4.3 惩罚样条估计	207
4.4 最小平均方差估计	207
4.5 估计方程估计	211
4.5.1 估计方法	211
4.5.2 主要结果	215
4.5.3 模拟研究	218
4.6 经验似然推断	219
4.6.1 方法论	219
4.6.2 模拟研究	220
4.7 部分线性单指标 EV 模型	222
4.7.1 Pseudo- θ 方法	222
4.7.2 修正的拟似然方法	224
4.8 删失数据分析	225
4.9 假设检验	227
4.9.1 检验参数分量	227
4.9.2 检验非参数分量	228
4.10 变量选择	229
参考文献	230
第 5 章 变系数模型	233
5.1 引言	233
5.2 截面数据分析	235
5.2.1 局部线性回归估计	236
5.2.2 光滑样条估计	238
5.2.3 多项式样条估计	238
5.2.4 拟合优度检验	241

5.2.5 模拟研究	242
5.3 纵向数据分析	243
5.3.1 模型	243
5.3.2 局部核估计	243
5.3.3 局部多项式估计	246
5.3.4 光滑样条估计	247
5.3.5 最小二乘基估计	249
5.4 经验似然推断	254
5.4.1 自然的经验似然	254
5.4.2 两种纠偏方法	255
5.4.3 近似置信域	257
5.4.4 数值计算结果	262
5.4.5 定理的证明	265
参考文献	269
第 6 章 部分线性变系数模型	275
6.1 引言	275
6.2 估计方法	276
6.2.1 profile 最小二乘估计	276
6.2.2 一般序列估计	278
6.3 profile 似然比检验	282
6.3.1 profile 似然比统计量	282
6.3.2 Wilks 现象	283
6.3.3 profile 似然比检验的功效	283
6.3.4 Wald 检验	284
6.3.5 非参数分量的检验	284
6.3.6 带宽选择	285
6.4 变量选择	286
6.4.1 引言	286
6.4.2 方法论和主要结果	286
6.4.3 计算方法	288
6.4.4 模拟研究	290
6.4.5 实例分析	292
6.4.6 定理的证明	293
参考文献	298

第 7 章 单指标变系数模型	301
7.1 引言	301
7.2 估计方法和主要结果	301
7.2.1 估计方法	301
7.2.2 带宽选择	303
7.2.3 算法	304
7.2.4 主要结果	305
7.3 经验似然	307
7.3.1 估计的经验似然	308
7.3.2 调整的经验似然	308
7.4 模拟和应用	309
7.4.1 模拟	309
7.4.2 实例分析	313
7.5 定理的证明	314
参考文献	321
第 8 章 部分变系数单指标模型	323
8.1 引言	323
8.2 估计方法	325
8.3 渐近性质	328
8.4 渐近置信区间/置信域	332
8.5 带宽选择	333
8.6 数值计算结果	334
8.6.1 模拟研究	334
8.6.2 实际例子	335
8.7 定理的证明	337
参考文献	348
第 9 章 单指标混合效应模型	351
9.1 引言	351
9.2 估计方法	352
9.2.1 参数和非参数分量的估计	352
9.2.2 方差分量的估计	354
9.3 主要结果	355
9.4 模拟与应用	358
9.4.1 模拟结果	358
9.4.2 一个实例	363

9.5 定理的证明	365
参考文献	370
第 10 章 可加模型	373
10.1 引言	373
10.2 估计方法	375
10.2.1 局部多项式估计	375
10.2.2 样条后移核估计	379
10.2.3 多项式样条估计	381
10.3 变量选择	384
10.4 可加部分线性模型	386
10.4.1 估计方法	386
10.4.2 经验似然	389
10.4.3 变量选择	389
10.5 广义可加部分线性模型	390
10.5.1 估计方法	391
10.5.2 变量选择	391
参考文献	392
《现代数学基础丛书》已出版书目	397

第1章 预备知识

本章介绍一些预备知识,以便读者阅读后面各个章节.本章的主要内容包括:回归模型简介、光滑方法、自动选择光滑参数、经验似然、惩罚方法、数据集和符号表等.

1.1 回归模型简介

回归模型是统计学中发展较早、理论丰富并且应用性强的统计模型.自 Galton 于 1886 年首次提出回归模型以来,这类模型一直倍受人们关注.在过去的一个多世纪里,回归模型被广泛地应用于工业、农业、气象、地质、经济管理以及医药卫生等各个领域,取得了丰富的理论和应用成果.同时,由于实际应用的需要,为了更加接近现实,更好地解释数据,回归模型也一直处于不断发展进步之中,其模型由初期的参数回归模型发展到非参数回归模型,20 世纪 80 年代以来又兴起了半参数回归模型,其理论研究和应用领域都在不断深入和扩大.

1.1.1 参数回归模型

参数回归模型具有形式

$$Y_i = f(X_i; \beta) + \varepsilon_i, \quad i = 1, 2, \dots, n,$$

其中 f 为已知函数, β 为未知的有限维参数, ε_i 为均值为 0 的随机误差.这类模型包括大家熟知的线性模型、广义线性模型和非线性回归模型.参数回归模型有一个共同的特点:函数形式假定是已知的,未知的仅仅是其中有限个参数.由于参数回归模型容易处理,且对其研究已有相当长的历史,因而已形成一套成熟的理论和方法.

参数回归模型通常由经验和历史资料对回归函数提供了大量的额外信息,因而当模型假定成立时,其推断有较高的精度.例如熟知的线性模型的最小二乘估计,其方差有 $O(n^{-1})$ 的阶.但响应变量与协变量之间的函数关系形式不总是已知的,在许多实际应用中,没有证据表明已知函数关系的存在.因而,参数回归模型在实际应用中往往存在模型的设定误差,当模型假定不成立时,基于假定模型所作的统计推断其表现可能很差,甚至导致错误的结论.

为了减少参数回归的模型偏差,统计学家提出了一个假设更宽松自由的模型——非参数回归模型.

1.1.2 非参数回归模型

设响应变量 Y 是随机变量, 协变量 X 是随机变量或非随机变量. 给定随机样本 $\{(X_i, Y_i), 1 \leq i \leq n\}$, 可建立如下回归模型:

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (1.1.1)$$

其中 $m(\cdot)$ 是未知的回归函数, ε_i 为随机误差. 假定 $\varepsilon_1, \dots, \varepsilon_n$ 相互独立同分布, 并满足:

(1) X 为非随机时: $E(\varepsilon_i) = 0$, $\text{Var}(\varepsilon_i) = \sigma^2 < \infty$, 此时 $E(Y_i) = m(X_i)$.

(2) X 为随机时: $E(\varepsilon_i|X_i) = 0$, $\text{Var}(\varepsilon_i|X_i) = \sigma^2(X_i) < \infty$, 此时 $E(Y_i|X_i) = m(X_i)$.

非参数回归模型的特点是: 回归函数的形式可以任意, 对协变量和响应变量的分布也很少限制, 因而有较大的适应性和稳健性. 非参数回归模型是基于数据的模型, 它假定响应变量和协变量之间的关系形式未知, 由观测数据本身对整个回归函数进行估计, 是较参数回归模型更符合现实的模型.

回归函数解释了 X 与 Y 之间的相互关系. 回归分析的目的在于对于未知函数 $m(x)$, 给出一个合理的近似值, 称之为 $m(x)$ 的估计. 构造 $m(x)$ 的估计量的方法称为非参数回归方法, 又称为光滑方法, 与参数回归方法相反, 不采用现成的数学函数作为模型, 在统计领域中是较新的方法. 用非参数回归估计回归曲线具有以下特点: 第一, 关于两个变量关系的探索是开放式的, 不套用任何现成的数学函数. 第二, 所拟合的曲线可以很好地描述变量之间关系的细微变化. 第三, 非参数回归提供的是万能的拟合曲线, 不管多么复杂的曲线关系都能进行成功拟合. 因此, 非参数估计方法在分析未知回归函数关系时提供了一个灵活的工具.

与传统回归拟合相比, 非参数回归则显得灵活机动, 可以对同一数据进行多次拟合, 以探索数据中可能隐藏的某种关系. 这是普通的回归拟合和散点图绝对做不到的. 而且, 非参数拟合可以对数据中的任何模式或变量间任何一种曲线关系进行拟合, 而传统回归却只能对个别的数据模式 (直线、二次曲线等) 进行拟合. 实际上, 非参数回归拟合往往会带来意想不到的结果, 会极富戏剧性地改变人们对数据进行进一步分析的方向, 得到更深刻的结论.

在非参数回归模型中, 回归函数既可以是一元的, 也可以是多元的. 理论上讲, 一元非参数回归的估计方法可以直接推广到多元非参数回归, 但是有一个很棘手的问题, 即维数问题. 由于非参数回归估计方法本质上讲都是局部估计或局部光滑, 要想使 $m(x)$ 在 x 点得到比较精确的估计, 就必须使得 x 的邻域包含足够多的数据. 当 x 的维数增大时, 局部光滑所需要的数据点个数成指数倍增加. 例如, 如果一个局部邻域沿着每一个坐标轴包含 10 个数据点, 那么在相应的 p 维邻域就需要

10^p 个数据点. 而实际上, 高维数据具有内部稀疏性, 即随着维数 p 的增加, 一个局部邻域所包含的数据点个数在整个样本中所占的比例越来越小. 为了获取非参数光滑所必须的足够多的数据, 可以有两种选择, 或者增大带宽, 或者增加样本数目. 增大带宽必然导致估计的偏差增加, 而要获取数目非常多的样本在许多情况下是不实际的. 由此可见, 当协变量 X 的维数增加时, 多元非参数回归估计的精度下降很快, 人们称这种现象为“维数灾祸”(curse of dimensionality).

为了克服“维数灾祸”问题, 统计学家在寻找既能达到数据降维同时又能保留非参数光滑优点的方法方面已经作了许多研究, 提出了多种降维模型. 主要的降维模型为半参数回归模型.

1.1.3 半参数回归模型

本书仅介绍几种重要的半参数回归模型, 其中包括: 部分线性模型、单指标模型、部分线性单指标模型、变系数模型、部分线性变系数模型、单指标变系数模型、部分变系数单指标模型和可加模型等. 在以后各章节中再对这些模型作一一阐述.

对于半参数回归模型的研究, 主要集中在统计推断方法和大样本性质上. 研究的基本问题大致可分为五个方面.

(1) 估计量的构造方法和统计性质. 包括回归参数、模型误差方差和非参数回归函数的估计方法的研究, 并研究估计量的相合性、渐近正态性及其收敛速度等. 对参数分量的估计方法有: 最小二乘法、广义估计方程法、广义矩方法、M 估计法等. 对非参数分量的估计方法包括: 核光滑、局部多项式光滑、样条光滑、小波光滑等.

(2) 回归参数的置信域以及非参数回归函数置信带的构造方法. 构造置信域的常用的方法有经验似然和 bootstrap 法等. 这类问题的研究给出了估计的精度, 具有重要性.

(3) 模型检验和拟合优度检验. 主要研究各种检验统计量的构造方法及其统计性质, 如广义似然比检验、Wald 检验、经验似然比检验等. 这类问题的研究为实际应用提供了理论依据.

(4) 变量选择或模型选择. 常用的变量选择方法包括: LASSO(least absolute shrinkage and selection operator) 和 SCAD(smoothly clipped absolute deviation) 及其各种变形. 这类问题是目前统计学研究的热点课题之一.

(5) 估计的稳健性. 所谓稳健性, 直观意义是说, 当假设模型与真实模型相符时估计量具有良好的性质; 当假设模型与真实模型稍许偏离时, 其性能所受影响较小; 当假设模型与真实模型发生严重偏离时, 其性能也还过得去. 如参数的 M 估计就是一种稳健估计.

1.2 光滑方法

本节概述回归光滑方法, 主要包括: 核光滑、局部多项式光滑、样条光滑、权函数法等. 此处介绍的光滑方法着眼于单变量回归估计. 这主要是出于如下两种原因: 第一, 对于高维的变量 X 表达式将变得极为复杂; 第二, 在以后各章讨论各种回归模型时, 主要使用单变量光滑方法.

1.2.1 核光滑

Nadaraya 和 Watson 于 1964 年分别提出了非参数回归函数 $m(x)$ 的核估计, 文献中也称之为核光滑 (kernel smoothing). 其思路如下: 选定 $\mathbf{R}$ 上的函数 $K(\cdot)$ 和正的常数列 $h = h_n$, 记 $K_h(\cdot) = h^{-1}K(\cdot/h)$, 定义

$$\hat{m}_{NW}(x) = \sum_{i=1}^n W_{ni}^K(x) Y_i, \quad (1.2.1)$$

其中

$$W_{ni}^K(x) = K_h(X_i - x) / \sum_{j=1}^n K_h(X_j - x), \quad (1.2.2)$$

则称 $\hat{m}_{NW}(x)$ 为 $m(x)$ 的 Nadaraya-Watson 估计, 简称为 N-W 估计或核估计; 称 $K(\cdot)$ 为核函数, h 为带宽. 通常取核函数 $K(\cdot)$ 为概率密度函数. 常用的核函数有:

- (1) 均匀核: $K(u) = \frac{1}{2}I(|u| \leq 1)$,
- (2) Epanechnikov 核 (或抛物线核): $K(u) = \frac{3}{4}(1 - u^2)I(|u| \leq 1)$,
- (3) 四次核: $K(u) = \frac{15}{16}(1 - u^2)^2I(|u| \leq 1)$,
- (4) 高斯核: $K(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$.

通常称带宽 h 为光滑参数. 例如, $K(\cdot)$ 为均匀核时, $\hat{m}_{NW}(x)$ 是集中在 x 附近一个邻域的样本的加权平均, 而 h 正好是这个邻域的宽度. 当 h 大时, 参与平均的样本就多, 会提高估计的精度, 但可能会增大偏差. 当 h 小时则正好相反. 因此, h 的选取至关重要. 关于它的选取问题将在 1.3 节解决.

在一些正则条件下, N-W 估计是相合的, 即当 $h \rightarrow 0$ 且 $nh \rightarrow \infty$ 时,

$$\hat{m}_{NW}(x) \xrightarrow{P} m(x).$$

相对于正确指定下的参数模型而言, 非参数估计是有偏的. 偏差 $\text{bias}(\hat{m}_{NW}) = E[\hat{m}_{NW}(x) - m(x)]$ 可以通过减小带宽而减小, 但 $\hat{m}_{NW}(x)$ 的方差将增大. 非参数回归的技巧是如何平衡偏差和方差的值使得所需要的结果达到基本合理.