第 24 卷 第 6 期
2018年12月

上海大学学报（自然科学版）
JOURNAL OF SHANGHAI UNIVERSITY (NATURAL SCIENCE)

Vol. 24 No. 6
Dec. 2018

DOI: 10.12066/j.issn.1007-2861.1926

非参数回归的贝叶斯估计

苏雅玲， 何幼桦
(上海大学 理学院, 上海 200444)

摘要： 基于因变量 Y 对自变量 X 条件分布的非参数贝叶斯估计, 通过期望计算得到未知回归函数的后验估计表达式, 并计算出估计的均方误差, 证明该估计的均方收敛性. 阐明当先验的选择接近真实的回归函数时, 该估计的均方误差小于局部线性核回归的均方误差. 最后通过实证分析, 表明该非参数贝叶斯回归比非参数局部线性回归具有更好的预测效果.

关键词： 非参数贝叶斯回归; 非参数贝叶斯分布估计; Dirichlet过程; 局部线性回归; 人口预测

中图分类号： O 212.7 **文献标志码：** A **文章编号：** 1007-2861(2018)06-1022-08

Bayesian estimation for nonparametric regression

SU Yaling, HE Youhua
(College of Sciences, Shanghai University, Shanghai 200444, China)

Abstract: Based on nonparametric Bayesian estimation of a conditional distribution, posterior estimation of an unknown regression function is obtained by calculating its expectation. The mean square error of the estimation is calculated. Its convergence in mean square of the estimation is proved. It is shown that the mean square error of the estimation is less than that of the local linear kernel regression when prior regression is chosen to be close to the unknown regression function. Empirical evidence shows that the nonparametric Bayesian regression may be more effective in prediction than local linear regression.

Key words: nonparametric Bayesian regression; nonparametric Bayesian distribution estimation; Dirichlet process; local linear regression; population prediction

在回归分析中, 相对于回归函数形式已知的参数回归, 非参数回归的函数形式自由, 受约束少, 回归模型完全由数据驱动. 对于非参数回归问题, 已经有很多学者进行了研究, 目前回归函数的非参数估计主要集中在核估计、局部多项式估计、样条函数以及小波估计. 如 Devroye[1] 证明了回归函数核估计的强相合性. Greblicki 等[2]和 Devroye[3] 在不同条件下证明了回归函数核估计的逐点相合性. Fan[4] 引入局部线性回归估计并阐述其优越性, 给出了估计量的均方误差(mean-square error, MSE)和积分均方误差(mean integrated square error, MISE), 并研究了估计量的最大最小效. Antoniadis 等[5]引入回归函数的小波估计, 并证明了估计量的相合性和渐进正态性. 上述方法单纯从数据本身出发, 没有充分利用数据以外的信息, 虽然能够达到较好的拟合效果, 但是外推效果较差.

收稿日期: 2017-02-15
基金项目: 国家自然科学基金资助项目(11371242)
通信作者: 何幼桦(1960—), 男, 副教授, 博士, 研究方向为概率统计. E-mail: heyouhua@shu.edu.cn

在参数估计问题中, 贝叶斯方法已经得到广泛的应用, 而在非参数估计问题中, 贝叶斯方法是从 Ferguson[6] 在 1973 年发表了论文 *A Bayesian Analysis of Some Nonparametric Problems*[6]后才开始受到关注. Ferguson 给出了当总体 X 的分布函数 $F(x)$ 的先验为 Dirichlet 过程时, $F(x)$ 的非参数贝叶斯估计为 $F(x)$ 先验分布与经验分布的加权平均. 此外, Jayaram[7]提出了拓展的 Dirichlet 过程先验. 在此基础上姚宗静[8]给出简单 Dirichlet 过程的构造性定义, 讨论了其性质和支撑问题, 求出在该先验下后验分布的具体表达形式. 也有学者将非参数贝叶斯方法应用到回归函数的估计中, 如龙杏芬等[9]在局部线性估计中窗宽h的先验分布为 Gamma 分布的条件下, 基于贝叶斯方法构造了回归函数的局部线性估计, 并给出窗宽和回归函数的后验分布和抽样方法, 通过数值模拟验证了贝叶斯局部线性估计方法的可行性. 卢一强等[10] 对广义非参数模型 B 样条贝叶斯估计进行了研究, 给出了回归函数B 样条贝叶斯估计的马尔科夫链蒙特卡洛(Markov chain Monte Carlo, MCMC)模拟计算方法.

本工作针对非参数回归模型, 在 Ferguson 给出的总体分布函数的贝叶斯估计基础上, 进一步得到一个光滑的条件分布估计. 基于该分布最终构造出回归函数的贝叶斯估计, 并研究该估计的收敛性质, 给出该估计中超参数的合理表达式. 最后, 通过实证分析将非参数贝叶斯方法与局部线性回归进行了比较.

1 非参数回归的贝叶斯估计

文献 [6] 给出了总体 X的分布函数 $F(x)$ 的非参数贝叶斯估计, 在该估计中取 $F(x)$ 的先验分布服从 Dirichlet 过程. Dirichlet 过程定义如下.

定义 1 设 X 为一样本空间, A是X的子集构成的 σ代数, $\alpha > 0$, P_0 为 (X, A) 上的有限非零测度. 如果对 X 的任意可测分割 $A_1, A_2, \cdots, A_m$, $p = (P(A_1), P(A_2), \cdots, P(A_m))$服从参数为$\alpha = (\alpha P_0(A_1), \alpha P_0(A_2), \cdots, \alpha P_0(A_m))$ 的 Dirichlet 分布, 则称p是(X, A)上参数为 α, 基测度为 P_0 的 Dirichlet 过程, 记为$p \sim \mathrm{DP}(\alpha, P_0)$.

文献 [6] 中的 Dirichlet 过程即为 $\mathrm{DP}(\alpha, P_0)$, α 为正实数, 记 $F_0(x) = P_0\{X \leqslant x\}$ 是先验过程的期望(均值函数). 则在样本为 $x_1, x_2, \cdots, x_n$ 时, $F(x)$ 的贝叶斯估计为

$$\widehat{F}_n(x \mid x_1, x_2, \cdots, x_n) = \frac{\alpha}{\alpha + n} F_0(x) + \frac{n}{\alpha + n} F_n(x \mid x_1, x_2, \cdots, x_n),$$

该估计是先验过程的期望分布 $F_0(x)$ 和经验分布估计 $F_n(x \mid x_1, x_2, \cdots, x_n)$ 的加权平均. 由于经验分布函数是阶梯函数, 为得到一个光滑的分布估计, 用核估计代替经验分布函数 $F_n(x \mid x_1, x_2, \cdots, x_n)$, 则总体的密度估计为

$$\widehat{p}_{\mathrm{B}}(x) = \frac{n}{n + \alpha} \widehat{p}_h(x) + \frac{\alpha}{n + \alpha} p_0(x), \tag{1}$$

式中, $\widehat{p}_h(x) = \frac{1}{n} \sum\limits_{i=1}^{n} K_h(x - x_i)$ 是以$K(x)$ 为核函数的核密度估计, $p_0(x)$ 为先验过程期望分布的密度函数.

现考虑 $Y \in R^1$ 对 $X = (X_1, X_2, \cdots, X_d) \in R^d$ 的多元非参数回归模型

$$Y = m(X) + \varepsilon,$$

其中 $m(X)$ 是未知回归函数, ε 是均值为 0 方差为 σ^2 的误差项. 设(x, y) 是变量 (X, Y) 的某个具体取值, $\{(X_i, Y_i), i = 1, 2, \cdots, n\}$为样本数据. 在多维情况下, 式(1)可表示成如下形式:

$$\widehat{p}_{\mathrm{B}}(x, y) = \frac{n}{n + \alpha} \widehat{p}_{h,H}(x, y) + \frac{\alpha}{n + \alpha} p_0(x, y),$$

其中 H 为 X 带宽矩阵, h 为 Y 的带宽. $\widehat{p}_{h,H}(x,y)=\dfrac{1}{n}\sum\limits_{i=1}^{n}K_h(Y_i-y)K_H(X_i-x)$ 为总体(X,Y)的核密度估计, $p_0(x,y)$为联合先验过程的期望概率密度函数.

Y的条件分布的贝叶斯估计为

$$\widehat{p}_{\mathrm{B}}(y\mid x)=\frac{n\widehat{p}_{h,H}(x,y)+\alpha p_0(x,y)}{n\widehat{p}_H(x)+\alpha p_0(x)},\tag{2}$$

式中, $\widehat{p}_H(x)=\dfrac{1}{n}\sum\limits_{i=1}^{n}K_H(X_i-x)$.

在二次损失下, 回归函数 $m(x)$ 的贝叶斯估计为

$$\begin{aligned}\widehat{m}_{\mathrm{B}}(x)&=E_{\widehat{p}_{\mathrm{B}}(y|x)}(Y\mid x)=\int_{-\infty}^{+\infty}y\widehat{p}_{\mathrm{B}}(y\mid x)\mathrm{d}y\\&=\frac{\sum\limits_{i=1}^{n}y_iK_H(X_i-x_i)+\alpha\int_{-\infty}^{+\infty}yp_0(x,y)\mathrm{d}y}{\sum\limits_{i=1}^{n}K_H(X_i-x)+\alpha p_0(x)}\\&=\frac{\sum\limits_{i=1}^{n}y_iK_H(X_i-x)+\alpha p_0(x)m_0(x)}{\sum\limits_{i=1}^{n}K_H(X_i-x)+\alpha p_0(x)}\\&=\frac{\sum\limits_{i=1}^{n}K_H(X_i-x)\widehat{m}_H(x)+\alpha p_0(x)m_0(x)}{\sum\limits_{i=1}^{n}K_H(X_i-x)+\alpha p_0(x)}\\&=\frac{n\widehat{p}_H(x)\widehat{m}_H(x)+\alpha p_0(x)m_0(x)}{n\widehat{p}_H(x)+\alpha p_0(x)},\end{aligned}\tag{3}$$

式中, $m_0(x)$ 为 $m(x)$ 基于 $p_0(x,y)$ 的先验回归函数, $\widehat{m}_H(x)=\dfrac{\sum\limits_{i=1}^{n}y_iK_H(X_i-x)}{\sum\limits_{i=1}^{n}K_H(X_i-x)}$ 是多元 Nadaraya-Watson 核回归估计, 即零阶局部多项式回归. 考虑一阶局部多项式回归(局部线性回归)将会减少边界偏倚, 而不增加方差[11], 因此在计算中可以将式(3)中的多元 Nadaraya-Watson 核回归估计 $\widehat{m}_H(x)$替换成多元局部线性回归估计

$$\widehat{m}_H(x)=e_0^{\mathrm{T}}(X^{\mathrm{T}}WX)^{-1}X^{\mathrm{T}}WY,$$

其中$e_0=(1,0,\cdots,0)^{\mathrm{T}}, X=\begin{pmatrix}1&(X_1-x)^{\mathrm{T}}\\1&(X_2-x)^{\mathrm{T}}\\\vdots&\vdots\\1&(X_n-x)^{\mathrm{T}}\end{pmatrix}, Y=\begin{pmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{pmatrix}, W=\mathrm{diag}(K_H(X_1-x),K_H(X_2-x),\cdots,K_H(X_n-x))$.

另一方面, $m(x)$的贝叶斯估计式(3)中, α 反映 $\widehat{m}_{\mathrm{B}}(x)$ 对先验 $m_0(x)$ 的依赖程度, 如果这种依赖随 x 的变化而有所不同, 则式(3)可写成

$$\widehat{m}_{\mathrm{B}}(x)=\frac{n\widehat{p}_H(x)\widehat{m}_H(x)+\alpha(x)p_0(x)m_0(x)}{n\widehat{p}_H(x)+\alpha(x)p_0(x)}.\tag{4}$$

1.1 估计的均方收敛性

文献 [12] 给出了多元局部线性回归估计 $\widehat{m}_H(x)$ 的方差和偏差.

引理 1 对于样本模型 $Y_i = m(X_i) + \varepsilon_i$, $i = 1, 2, \cdots, n$. $\widehat{m}_H(x)$ 为$m(x)$ 具有带宽矩阵 H的局部线性估计, 并满足文献 [13] 中的正则条件. 设x为一个非边界点, 则在给定$X_1, X_2, \cdots, X_n$ 下$\widehat{m}_H$ 的偏倚为

$$\mathrm{Bias}(\widehat{m}_H(x) \mid X_1, X_2, \cdots, X_n) \approx \frac{1}{2}u_2(K_H)\mathrm{tr}(H^{\mathrm{T}}\mathcal{H}_m(x)H),$$

这里 $\mathcal{H}_m(x)$ 为 $m(x)$ 的 Hessian 矩阵, 而 $u_2(K_H)$ 是核函数 K_H 的二阶矩. $\widehat{m}_H$ 的方差为

$$\mathrm{Var}(\widehat{m}_H(x) \mid X_1, X_2, \cdots, X_n) \approx \frac{1}{n\det(H)}\|K_H\|_2^2\frac{\sigma^2}{p(x)}.$$

记 $\mathrm{Var}(\widehat{m}_H(x)) = \mathrm{V}_n(x)$, $\mathrm{Bias}(\widehat{m}_H(x)) = \mathrm{E}(\widehat{m}_H(x)) - m(x) = \mathrm{B}_n(x)$, 则$m(x)$核估计$\widehat{m}_H(x)$的均方误差为

$$\mathrm{MSE}(\widehat{m}_H(x)) = \mathrm{Var}(\widehat{m}_H(x)) + (\mathrm{Bias}(\widehat{m}_H(x)))^2 = \mathrm{V}_n(x) + \mathrm{B}_n^2(x). \tag{5}$$

定理 1 在$\widehat{m}_H(x)$满足引理 1 的条件下, $m(x)$的贝叶斯估计的均方误差为

$$\mathrm{MSE}(\widehat{m}_{\mathrm{B}}(x)) = \frac{n^2\widehat{p}_H^2(x)\mathrm{V}_n(x) + (n\widehat{p}_H(x)\mathrm{B}_n(x) + \alpha p_0(x)(m_0(x) - m(x)))^2}{(n\widehat{p}_H(x) + \alpha(x)p_0(x))^2}. \tag{6}$$

注意到, 当 $m_0(x) = m(x)$时,

$$\mathrm{MSE}(\widehat{m}_{\mathrm{B}}(x)) = \frac{n^2\widehat{p}_H^2(x)(\mathrm{V}_n(x) + \mathrm{B}_n^2(x))}{(n\widehat{p}_H(x) + \alpha(x)p_0(x))^2} < \mathrm{V}_n(x) + \mathrm{B}_n^2(x) = \mathrm{MSE}(\widehat{m}_H(x)).$$

由此可知, 当 $m(x)$ 先验选择接近 $m(x)$ 时, $m(x)$ 的非参数贝叶斯估计的均方误差将小于其局部线性回归估计的均方误差.

1.2 超参数 $\boldsymbol{\alpha(x)}$ 的确定

式(4)中的超参数 $\alpha(x)$ 反映了分布估计对先验分布的依赖程度, $\alpha(x)$ 越大则这种依赖越强. 在实际计算时, 需要对超参数 $\alpha(x)$ 进行合理地确定.

定理 2 以 MSE 达到最小的 $\alpha(x)$ 可以表示为

$$\widehat{\alpha}(x) = \frac{2nT\det(H)u_2(K_H)p(x)(m_0(x) - m(x)) + nT^2\det(H)u_2^2(K_H)p(x) + 4\|K_H\|_2^2\sigma^2}{2\det(H)p_0(x)(m(x) - m_0(x))(2(m(x) - m_0(x)) - Tu_2(K_H))}, \tag{7}$$

其中 $T = \mathrm{tr}(H^{\mathrm{T}}\mathcal{H}_m(x)H)$.

证明

$$
\begin{aligned}
\mathrm{MSE}(\widehat{m}_{\mathrm{B}}(x)) =& \frac{n^2\widehat{p}_H^2(x)\mathrm{V}_n(x)+(n\widehat{p}_H(x)\mathrm{B}_n(x)+\alpha(x)p_0(x)(m_0(x)-m(x)))^2}{(n\widehat{p}_H(x)+\alpha(x)p_0(x))^2}\\
=& \frac{1}{(n\widehat{p}_H(x)+\alpha(x)p_0(x))^2}\left\{\frac{n\widehat{p}_H^2(x)\|K_H\|_2^2\sigma^2}{\det(H)p(x)}\right.\\
&\left.+\left[\frac{n\widehat{p}_H(x)u_2(K_H)T}{2}+\alpha(x)p_0(x)(m(x)-m_0(x))\right]^2\right\}\\
\approx& \frac{1}{(np(x)+\alpha(x)p_0(x))^2}\left\{\frac{np(x)\|K_H\|_2^2\sigma^2}{\det(H)}+\left[\frac{np(x)u_2(K_H)T}{2}\right.\right.\\
&\left.\left.+\alpha(x)p_0(x)(m(x)-m_0(x))\right]^2\right\}\\
\overset{\text{记}}{=}& \frac{A+[B+\alpha(x)C]^2}{(np(x)+\alpha(x)p_0(x))^2},
\end{aligned}
$$

其中

$$
A=\frac{np(x)\|K_H\|_2^2\sigma^2}{\det(H)},\quad B=\frac{np(x)u_2(K_H)T}{2},\quad C=p_0(x)(m(x)-m_0(x)).
$$

上述约等式中是用 $p(x)$ 代替了核估计$\widehat{p}_H(x)$.

$\mathrm{MSE}(\widehat{m}_{\mathrm{B}}(x))$ 关于 $\alpha(x)$ 的一阶偏导数为

$$
\frac{\partial \mathrm{MSE}(\widehat{m}_{\mathrm{B}}(x))}{\partial\alpha(x)}=-\frac{Ap_0(x)+(p_0(x)B-np(x)C)(B+\alpha(x)C)}{(np(x)+\alpha(x)p_0(x))^3}\overset{\text{令}}{=}0,
$$

得式(7)的结果. 又因 $\alpha(x)$ 为式(7)时

$$
\frac{\partial^2\mathrm{MSE}(\widehat{m}_{\mathrm{B}}(x))}{\partial(\alpha(x))^2}=\frac{2C^4(p_0(x)B-np(x)C)^4}{(Ap_0^2(x)+(p_0(x)B-np(x)C)^2)^3}>0,
$$

所以当 $\alpha(x)$ 取式(7)时, $\mathrm{MSE}(\widehat{m}_{\mathrm{B}}(x))$ 取得极小值.

在定理 2 中, 当 $m(x)$ 对每个分量的二阶偏导接近 0, 即$T\approx 0$时, 则可取

$$
\widehat{\alpha}(x)=\frac{\|K_H\|_2^2\sigma^2}{\det(H)p_0(x)(m(x)-m_0(x))^2}. \tag{8}
$$

根据式(8)可知, 当先验回归函数的选取和真实的回归函数接近时, $(m(x)-m_0(x))^2$较小, $\widehat{\alpha}(x)$较大, 回归函数的非参数贝叶斯估计结果对先验分布依赖度高. 反之, 当先验回归函数的选取和真实的回归函数相差较大时, $\widehat{\alpha}(x)$ 较小, 则估计结果对先验分布依赖度较低. 由于 $m(x)$ 是未知的, 先验 $m_0(x)$ 的选取具有主观性, 因此可以限定$(m(x)-m_0(x))^2\leqslant M$, M 为正实数. 式(8)中$\alpha(x)$ 的确定还依赖方差 σ^2, 其估计可采用文献 [11] 中的方法.

2 实证分析

为了检验所提出算法的有效性, 本工作以人口预测问题作为实证分析, 样本选取 1990—2005 年的中国人口数据, 建立人口数量对时间的回归模型, 运用非参数贝叶斯方法对模型进行估计, 最后以 2006—2010 年的数据检验模型, 并将结果与局部线性回归进行对比分析.

2.1 人口模型的估计

基于 1990—2005 年的中国人口样本建立一元非参数回归模型:

$$y_i = m(x_i) + \varepsilon_i, i = 1, 2, \cdots, 16. \tag{9}$$

选取 Logistics 人口模型[14]作为先验:

$$m_0(x) = \frac{141\,480}{1 + 0.295\,1\exp(-0.071\,3(x - 1987))}.$$

根据式(8), $\widehat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{m}_h(x))^2} = 222.84$, $p_0(x)$ 取均匀分布, 得 $\alpha = 3.358\,7$. 由式(4)得 $m(x)$ 的非参数贝叶斯估计 $\widehat{m}_{\mathrm{B}}(x)$, 并与局部线性回归的估计结果进行比较, 结果如表 1 所示.

表 1 1990—2005 年中国人口数量估计结果

Table 1 Estimation of China's population in 1990—2005 万人

年份(x_i)	实际人口(y_i)	$\widehat{m}_h(x)$	$\widehat{m}_h(x) - y_i$	$\widehat{m}_{\mathrm{B}}(x)$	$\widehat{m}_{\mathrm{B}}(x) - y_i$
1990	114 333	114 496.40	163.40	114 430.44	97.44
1991	115 823	115 833.96	10.96	115 823.20	0.20
1992	117 171	117 149.46	–21.54	117 172.28	1.28
1993	118 517	118 439.05	–77.95	118 481.98	–35.02
1994	119 850	119 698.56	–151.44	119 753.01	–96.99
1995	121 121	120 923.86	–197.14	120 984.13	–136.87
1996	122 389	122 111.23	–277.77	122 173.33	–215.67
1997	123 626	123 257.63	–368.37	123 318.62	–307.38
1998	124 761	124 360.86	–400.14	124 418.49	–342.51
1999	125 786	125 419.60	–366.40	125 472.11	–313.89
2000	126 743	126 433.56	–309.44	126 479.47	–263.53
2001	127 627	127 403.68	–223.32	127 441.47	–185.53
2002	128 453	128 332.26	–120.74	128 359.86	–93.14
2003	129 227	129 222.87	–4.13	129 236.97	9.97
2004	129 988	130 080.08	92.08	130 075.02	87.02
2005	130 756	130 909.02	153.02	130 875.17	119.17

用均方误差$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2$评价估计方法, 计算结果如表 2 所示.

表 2 拟合均方误差 MSE

Table 2 Fitting mean square error MSE 万人

	局部线性回归	非参数贝叶斯估计
$\sqrt{\mathrm{MSE}}$	222.84	182.92

由表 2 可以看出, 非参数贝叶斯估计与局部线性回归方法相比较, 均方误差要小得多, 拟合效果较好.

2.2 人口模型的预测

由 2.1 节得到人口数量的非参数贝叶斯估计结果, 预测 2006—2010 年中国人口数量如表 3 所示. 表 4 为预测均方误差. 表 4 表明在对中国人口数量进行预测时, 非参数贝叶斯估计与局部线性回归相比较, 均方误差较小, 在一定程度上克服了局部线性回归方法在预测外推方面效果较差的问题.

表 3 2006—2010 年中国人口数量预测结果

Table 3 China's population forcast results from 2006 to 2010 万人

年份(x_i)	实际人口(y_i)	$\widehat{m}_h(x)$	$\widehat{m}_h(x)-y_i$	$\widehat{m}_{\mathrm{B}}(x)$	$\widehat{m}_{\mathrm{B}}(x)-y_i$
2006	131 448	131 714.90	266.90	131 636.21	188.21
2007	132 129	132 502.65	373.65	132 353.37	224.37
2008	132 802	133 276.68	474.68	133 017.78	215.78
2009	133 450	134 040.75	590.75	133 618.17	168.17
2010	134 091	134 797.93	706.93	134 146.22	55.22

表 4 预测均方误差 MSE

Table 4 Prediction mean square error MSE 万人

	局部线性回归	非参数贝叶斯估计
$\sqrt{\mathrm{MSE}}$	506.93	180.92

3 结 束 语

本工作利用 Y 对 X 的条件分布的非参数贝叶斯估计来构造回归函数的非参数贝叶斯估计, 在此过程中, 用分布估计的核估计替代 Ferguson 估计的经验分布函数, 用较高阶的局部多项式回归替代原构造中的 Nadaraya-Watson 回归估计, 获得了较为理想的估计效果, 同时还给出了估计的均方误差及其均方收敛性. 实证结果表明, 对于非参数贝叶斯估计, 当先验分布选择较合适时, 在数据拟合和预测方面均表现出了较好的效果.

参考文献:

[1] Devroye L. The uniform convergence of the Nadaraya-Watson regression function estimate [J]. Canadian Journal of Statistics, 1978, 6(2): 179-191.

[2] Greblicki W, Pawlak M. Distribution-free pointwise consistency of kernel regression estimate [J]. J Annals of Statistics, 1984, 12(4): 1570-1575.

[3] Devroye L. On the almost everywhere convergence of nonparametric regression function estimates [J]. J Annals of Statistics, 1981, 9(6): 1310-1319.

[4] Fan J Q. Local linear regression smoothers and their minimax efficiencies [J]. J Annals of Statistics, 1993, 21(1): 196-216.

[5] Antoniadis A, Gregoire G, McKeague I W. Wavelet methods for curve estimation [J]. Journal of the American Statistical Association, 1994, 89(428): 1340-1353.

[6] Ferguson T. A Bayesian analysis of some nonparametric problems [J]. J Annals of Statistics, 1973, 1(2): 209-230.

[7] JAYARAM S. A constructive definition of Dirichlet priors [J]. J Statistics Sinica, 1994, 4(2): 639-650.

[8] 姚宗静. 基于 Dirichlet 过程的非参数贝叶斯分析 [D]. 成都: 西南交通大学, 2007.

[9] 龙杏芬, 张德生, 武新乾. 非参数回归模型的贝叶斯局部线性估计[J]. 山西大学学报(自然科学版), 2010, 33(3): 371-374.

[10] 卢一强, 茆诗松. 广义非参数回归的 B 样本贝叶斯估计 [J]. 应用数学, 2015, 18(1): 8-13.

[11] 沃塞曼. 现代数学译丛 3: 现代非参数统计 [M]. 北京: 科学出版社, 2008: 58-65.

[12] HARDLE W, WERWATZ A, MÜLLER M, et al. Nonparametric and semiparametric models [J]. Springer Series in Statistics, 2006, 62(2): 628-629.

[13] RUPPERT D, WAND M P. Multivariate locally weighted least squares regression [J]. Annals of Statistics, 1994, 22(3): 1346-1370.

[14] 王学保, 蔡果兰. Logistic 模型的参数估计及人口预测 [J]. 北京工商大学学报(自然科学版), 2009, 27(6): 75-78.