系 统 科 学 与 数 学
J. Sys. Sci. & Math. Scis.
29(5) (2009, 5), 608–616

含估计参数的加权经验过程*

张 军 舰

(北京工业大学应用数理学院，北京 100124; 广西师范大学数学科学学院，桂林 541004)

摘要 在讨论 (广义) 非参数似然比拟合优度检验时，加权经验过程理论是一个非常重要的基础. 但对含估计参数的加权经验过程理论，目前文献中很少讨论. 对含估计参数的加权经验过程的上界型和积分型两种统计量的近似分布进行了讨论，给出了较一般的结果. 所得结论可以为进一步讨论复合零假设下 (广义) 非参似然比拟合优度检验提供理论基础.

关键词 加权经验过程，估计参数，拟合优度检验.

MR(2000) 主题分类号 62E20

1 引言及基本记号

设 $X_1, X_2, \cdots, X_n$ 为来自 X 的 iid. 样本， X 的分布函数为 F. 考虑检验问题

$$H_0 : F \in \mathcal{F}_\theta \longleftrightarrow H_1 : F \notin \mathcal{F}_\theta, \tag{1}$$

其中 $\mathcal{F}_\theta = \{F(\cdot, \theta) : \theta \in \Theta, F \text{ 的形式已知 }\}$. 记 F_n 为相应的经验分布函数， $\theta_0 \in \Theta \subset \boldsymbol{R}^p$ 为参数的真值. 当参数 θ_0 已知时，检验问题 (1) 实际上就是简单假设，此时文献中已有的检验较多，例如卡方检验、加权 Kolmogorov-Smirnov 检验、加权 Cramér-von Mises (包括 Anderson-Darling) 检验、 Berk-Jones 检验 [1] 、修正的 Berk-Jones 检验 [2] 以及文献 [3] 中所讨论的广义非参似然比检验. 这些检验中除卡方外都是基于经验分布函数 F_n 来构造的，因此其性质的讨论必然用到加权经验过程的基本理论. 遗憾的是有关加权经验过程的理论探讨基本都集中在简单假设，具体见文献 [4] 及其后的参考文献. 对带有估计参数的复合零假设情况，也只有经验过程有所讨论，具体见文献 [5–7]. 对含估计参数的加权经验过程，至今没有文献详细讨论.

本文打算就带有估计参数的加权经验过程进行较系统的讨论. 首先，就标准化权函数情况，我们对含估计参数的加权经验过程的上界型和积分型两种统计量的近似分布进行讨论；然后给出相应的一般权函数情况的近似分布. 我们的结论为进一步讨论复合假设下较新的拟合优度检验 (例如广义非参似然比检验等) 提供理论基础.

为后面讨论方便，我们先给出本文一些基本条件和通用记号.

设 $\theta = (\theta_1, \theta_2, \cdots, \theta_p)^\tau$, $\widehat{\theta}_n$ 为 θ_0 的一个 (合理) 估计，满足如下条件.

* 广西自然科学基金 (桂科自 0832102), 广西师范大学博士科研启动基金和国家自然科学基金 (10661003) 资助课题.

收稿日期：2007-11-06, 收到修改稿日期：2008-05-15.

条件 1

$$n^{\frac{1}{2}}(\widehat{\theta}_n-\theta_0)=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}l(X_j,\theta_0)+\varepsilon_{1n}, \tag{2}$$

其中 $\varepsilon_{1n}=o_p(1)$; $l(x,\theta)=(l_1(x,\theta),l_2(x,\theta),\cdots,l_p(x,\theta))^\tau$ 满足 $E_F l(X_j,\theta)=0$, $l_i(x,\theta)$ 在任意有限区间内均为 x 的有界变差函数， $M(\theta_0)\hat{=}E_F\left[l(X_j,\theta_0)l^\tau(X_j,\theta_0)\right]$ 有穷且非负定.

记

$$\alpha_n(x)=n^{\frac{1}{2}}\left[F_n(x)-F(x)\right]=n^{\frac{1}{2}}\left[F_n(x)-F(x,\theta_0)\right],$$

$$\widehat{\alpha}_n(x)=n^{\frac{1}{2}}\left[F_n(x)-F(x,\widehat{\theta}_n)\right].$$

$$V_n(x;q)=\frac{\alpha_n(x)}{q(F(x,\theta_0))},\quad \widehat{V}_n(x;q)=\frac{\widehat{\alpha}_n(x)}{q(F(x,\widehat{\theta}_n))}.$$

从纯技术方面讲，条件 1 保证了 $\widehat{\theta}_n$ 的弱相合性，对所有 $\theta_0\in\Theta\subset\boldsymbol{R}^p$, 当 $n\to\infty$ 时, $P\left(\widehat{\theta}_n\in\Theta\right)\to 1$. 因此从渐近意义上说，不考虑 $\widehat{\theta}_n\notin\Theta$ 的情况，在后面讨论中，均假定对整个样本空间， $\widehat{\theta}_n\in\Theta$, 这样 $\widehat{\alpha}_n(x)$ 的定义就不会引起歧义．事实上， $\alpha_n(x)$ 就是经验过程, $\widehat{\alpha}_n(x)$ 就是含估计参数的经验过程． $V_n(x;q)$ 和 $\widehat{V}_n(x;q)$ 为对应的加权经验过程， q 为权函数.

另记 $\{B(t),0\le t\le 1\}$ 为一 Brownian 桥函数,

$$\bigtriangledown_\theta F(x,\theta_0)=\frac{\partial}{\partial\theta}F(x,\theta_0)=\left(\frac{\partial}{\partial\theta_1}F(x,\theta),\frac{\partial}{\partial\theta_2}F(x,\theta),\cdots,\frac{\partial}{\partial\theta_p}F(x,\theta)\right)^\tau\bigg|_{\theta=\theta_0},$$

$$\bigtriangledown_{\theta\theta}F(x,\theta_0)=\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}F(x,\theta)\bigg|_{\theta=\theta_0}\right).$$

下面我们再集中给出其他几个条件.

条件 2 矢量 $\frac{\bigtriangledown_\theta F(x,\theta)}{[F(x,\theta_0)(1-F(x,\theta_0))]^\alpha}(\frac{1}{2}\le\alpha<1)$ 关于 x 和 $\theta\in\varLambda$ 连续且有界，其中 $\varLambda$ 为 θ_0 的一个给定邻域的闭包.

条件 3 分布 F 有密度 f, 且矢量 $\bigtriangledown_\theta f(x,\theta)$ 关于 $x\in\boldsymbol{R}$ 和 $\theta\in\varLambda$ 连续且有界.

条件 4 $\frac{\bigtriangledown_\theta F(x,\theta)}{q^\gamma}(F(x,\theta_0))(0<\gamma<1)$ 关于 x 和 $\theta\in\varLambda$ 连续且有界.

考虑权函数类

$$FC_{0,1}=\left\{q(\cdot):q(t)\text{在 0 附近非降，在 1 附近非增};\forall\delta\in\left(0,\frac{1}{2}\right),\inf_{\delta\le t\le 1-\delta}q(t)>0\right\}.$$

令

$$I(q,c)=\int_0^1\frac{1}{t(1-t)}\exp\left(-\frac{cq^2(t)}{t(1-t)}\right)\mathrm{d}t.$$

条件 5 $q(t)$ 在 $t\in(0,1)$ 内可导， $q(\cdot)\in FC_{0,1}$, 对任意 $c>0, I(q,c)<\infty$.

条件 6 $q(t)$ 在 $t\in(0,1)$ 内可导， $q(\cdot)\in FC_{0,1}$, 且存在 $\gamma<\frac{1}{2}$, 使得

$$\int_0^1\frac{1}{(q(t))^{2\gamma}}\mathrm{d}t<\infty,\quad \int_0^1\frac{[t(1-t)]^{\frac{1}{2}}}{(q(t))^{2-\gamma}}\mathrm{d}t<\infty. \tag{3}$$

2 标准化经验过程的渐近分布

在加权经验过程中，如果 $q(t)=[t(1-t)]^{\frac{1}{2}}$, 则称 $q(t)$ 为标准化权函数. $V_n(x;q)$ 为标准化经验过程, $\widehat{V}_n(x;q)$ 为其对应的带估计参数形式. 如果考虑标准化经验过程上界型或积分型统计量的渐近分布，首先需要研究统计量 $V_n(x;q)$ 的极限分布 $G(F(x,\theta_0);\theta_0)$.

引理 2.1 设 $F^{-1}(t,\theta)=\inf\{x:F(x,\theta)\ge t\}$ 为函数 $F(x,\theta)$ 的反函数;

$$H(t,\theta)=\frac{\bigtriangledown_\theta F(F^{-1}(t,\theta),\theta)}{\sqrt{t(1-t)}},\quad M(\theta)=E\left\{l(X_1,\theta)l^\tau(X_1,\theta)\right\};$$

$$G(t,\theta)=\frac{B(t)}{\sqrt{t(1-t)}}-H^\tau(t,\theta)\int l(F^{-1}(y,\theta),\theta)\mathrm{d}B(y). \tag{4}$$

则 $\forall\theta\in\Theta$, $\{G(t,\theta),0\le t\le 1\}$ 为一中心化 Gaussian 过程，具有连续的样本轨道，其协方差函数 $\mathrm{Cov}(G(t,\theta),G(s,\theta))$ 为

$$\begin{aligned}&\frac{\min\{s,t\}-st}{\sqrt{t(1-t)s(1-s)}}-H^\tau(t,\theta)\frac{\int_0^s l(F^{-1}(y,\theta),\theta)\mathrm{d}y}{\sqrt{s(1-s)}}\\&-H^\tau(s,\theta)\frac{\int_0^t l(F^{-1}(y,\theta),\theta)\mathrm{d}y}{\sqrt{t(1-t)}}+H^\tau(t,\theta)M(\theta)H(s,\theta).\end{aligned} \tag{5}$$

证 由 Brownian 桥的性质经过简单的计算容易得到引理结论，不再细述.

对标准化经验过程的上界型统计量，有如下结论.

定理 2.1 记 $\mu_n=\frac{(\log n)^k}{n}$, $k\ge 3$, $q(t)=[t(1-t)]^{\frac{1}{2}}$, 其它记号沿用前述. 如果条件 1 和条件 2 成立, 则在 (1) 中零假设 H_0 成立时，有

$$\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}|\widehat{V}_n(x;q)|\stackrel{\mathcal{AD}}{=}\sup_{\mu_n\le t\le 1-\mu_n}|G(t,\theta_0)|, \tag{6}$$

其中 “$\stackrel{\mathcal{AD}}{=}$” 表示等号两边的渐近分布是一样的, $G(t,\theta)$ 的含义见 (4) 和 (5) 式.

证 对 $F(x,\widehat{\theta}_n)$ 在 θ_0 处进行 Taylor 展开，结合条件 1 和条件 2 知，在集合 $\{x:\mu_n\le F(x,\theta_0)\le 1-\mu_n\}$ 上,

$$\begin{aligned}\frac{F(x,\widehat{\theta}_n)(1-F(x,\widehat{\theta}_n))}{F(x,\theta_0)(1-F(x,\theta_0))}&=1+\frac{\|\bigtriangledown_\theta^\tau F(x,\theta_n^*)\|}{[F(x,\theta_0)(1-F(x,\theta_0))]^\alpha}o_p(1)+O_p(n^{-1})\\&=1+o_p(1),\end{aligned}$$

其中 θ_n^* 介于 $\widehat{\theta}_n$ 和 θ_0 之间. 因此

$$\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}\frac{F(x,\theta_0)(1-F(x,\theta_0))}{F(x,\widehat{\theta}_n)(1-F(x,\widehat{\theta}_n))}=1+o_p(1). \tag{7}$$

由 (2) 式和条件 2 可得

$$
\begin{aligned}
&\frac{n^{\frac{1}{2}}[F_n(x)-F(x,\widehat{\theta}_n)]}{[F(x,\theta_0)(1-F(x,\theta_0))]^{\frac{1}{2}}}\\
&=\frac{n^{\frac{1}{2}}[F_n(x)-F(x,\theta_0)]}{[F(x,\theta_0)(1-F(x,\theta_0))]^{\frac{1}{2}}}-\frac{n^{\frac{1}{2}}[F(x,\widehat{\theta}_n)-F(x,\theta_0)]}{[F(x,\theta_0)(1-F(x,\theta_0))]^{\frac{1}{2}}}\\
&=V_n(x;q)-\frac{\bigtriangledown_\theta^\tau F(x,\theta_n^*)\left[\frac{1}{\sqrt{n}}\sum_{j=1}^n l(X_j,\theta_0)+\varepsilon_{1n}\right]}{[F(x,\theta_0)(1-F(x,\theta_0))]^{\frac{1}{2}}}\\
&=V_n(x;q)-\frac{\bigtriangledown_\theta^\tau F(x,\theta_0)}{[F(x,\theta_0)(1-F(x,\theta_0))]^{\frac{1}{2}}}\int l(x,\theta_0)\mathrm{d}\alpha_n(x)+o_p(1).
\end{aligned}
\tag{8}
$$

所以在集合 $\{x:\mu_n\le F(x,\theta_0)\le 1-\mu_n\}$ 上，结合 (7) 有

$$
\begin{aligned}
\widehat{V}_n(x;q)&=\left[V_n(x;q)-\frac{\bigtriangledown_\theta^\tau F(x,\theta_0)}{[F(x,\theta_0)(1-F(x,\theta_0))]^{\frac{1}{2}}}\int l(x,\theta_0)\mathrm{d}\alpha_n(x)+o_p(1)\right](1+o_p(1))\\
&=V_n(x;q)-\frac{\bigtriangledown_\theta^\tau F(x,\theta_0)}{[F(x,\theta_0)(1-F(x,\theta_0))]^{\frac{1}{2}}}\int l(x,\theta_0)\mathrm{d}\alpha_n(x)+o_p(1).
\end{aligned}
\tag{9}
$$

另一方面，由 [8] 中定理 2 知，存在一个 Brownian 桥函数 $\{B(t),0\le t\le 1\}$ 使得

$$
\int l(x,\theta_0)\mathrm{d}\alpha_n(x)=\int_0^1 l(F^{-1}(t,\theta_0),\theta_0)\mathrm{d}B(t)+o_p(1).
\tag{10}
$$

由 (5), (9) 和 (10) 式得

$$
\begin{aligned}
&\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}\left|\widehat{V}_n(x;q)-G(F(x,\theta_0),\theta_0)\right|\\
\le&\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}\left|V_n(x;q)-\frac{B(F(x,\theta_0))}{\sqrt{F(x,\theta_0)(1-F(x,\theta_0))}}\right|\\
&+\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}\left|H^\tau(F(x,\theta_0),\theta_0)\left(\int l(t,\theta_0)\mathrm{d}\alpha_n(t)\ -\int l(F^{-1}(y,\theta_0),\theta_0)\mathrm{d}B(y)\right)\right|\\
&+o_p(1)\\
=&\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}\left|V_n(x;q)-\frac{B(F(x,\theta_0))}{\sqrt{F(x,\theta_0)(1-F(x,\theta_0))}}\right|+o_p(1).
\end{aligned}
$$

再利用 [9] 中引理 2 可知

$$
\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}\left|V_n(x;q)-\frac{B(F(x,\theta_0))}{\sqrt{F(x,\theta_0)(1-F(x,\theta_0))}}\right|=o\left((\log_2 n)^{-\frac{1}{2}}\right)=o(1),\ \text{a.s.}
$$

所以

$$
\sup_{\mu_n\le F(x,\theta_0)\le 1-\mu_n}\left|\widehat{V}_n(x;q)-G(F(x,\theta_0),\theta_0)\right|=o_p(1).
$$

因此定理 2.1 证毕.

注 2.1 文献 [9] 已经证明,

$$P\left(\sup_{\mu_n\le t\le 1-\mu_n}\frac{|B(t)|}{[t(1-t)]^{\frac{1}{2}}}\le\frac{t+e_n}{b_n}\right)\longrightarrow\exp(-4\exp(-x)),$$

其中 $b_n=\sqrt{2\log_2 n}$, $e_n=\log_2 n+\frac{\log_3 n}{2}-\frac{\log(4\pi)}{2}$, $\log_2 n=\log\log n,\log_3 n=\log\log\log n$. 对定理 2.1, 若对 $H(t,\theta_0)$ 施加较强的约束, $\sup\limits_{\mu_n\le t\le 1-\mu_n}|G(t,\theta_0)|$ 的渐近分布应该还是极值分布. 然而, 其小样本的真实分布的计算并不是一件容易的事, 目前还不清楚其具体形式. 实际上, 类似于文献 [10], 我们可以借助 Bootstrap 方法通过模拟来计算.

注 2.2 可以验证, 正态分布, 指数分布等常见的参数分布族和相应参数的极大似然估计均满足定理 2.1 的条件. 事实上, 如果总体属于位置尺度族

$$\mathcal{F}_\theta=\left\{G\left(\frac{x-\theta_1}{\theta_2}\right):-\infty<\theta_1<\infty,\theta_2>0\right\},$$

并且其适当阶矩存在时, 可以验证, 此族以及其参数的极大似然估计满足定理 2.1 的条件.

对于积分型统计量, 则有如下结论.

定理 2.2 如果条件 1 到条件 3 成立, 则在零假设 (1) 成立时, 有

$$\widehat{A}_n^2=\int\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{F(x,\widehat{\theta}_n)[1-F(x,\widehat{\theta}_n)]}f(x,\widehat{\theta}_n)\mathrm{d}x\xrightarrow{\mathcal{D}}\int_0^1G^2(t,\theta_0)\mathrm{d}t,$$

其中 "$\xrightarrow{\mathcal{D}}$" 表示依分布收敛, $G(t,\theta)$ 的含义见 (4) 和 (5) 式.

证 由条件 1 到条件 3 和 Taylor 展开, 容易得到

$$\left|\widehat{A}_n^2-\int\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{F(x,\theta_0)[1-F(x,\theta_0)]}f(x,\theta_0)\mathrm{d}x\right|=o_p(1). \tag{11}$$

又由于

$$\begin{aligned}
&\left|\int\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{F(x,\theta_0)[1-F(x,\theta_0)]}f(x,\theta_0)\mathrm{d}x-\int_0^1G^2(t,\theta_0)\mathrm{d}t\right|\\
=&\left|\int\frac{n[F_n(x)-F(x,\theta_0)-\nabla_\theta^\tau F(x,\theta_n^*)(\widehat{\theta}_n-\theta_0)]^2}{F(x,\theta_0)(1-F(x,\theta_0))}f(x,\theta_0)\mathrm{d}x\right.\\
&\left.-\int\frac{[B(F(x,\theta_0))-\nabla_\theta^\tau F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}B(F(t,\theta_0))]^2}{F(x,\theta_0)(1-F(x,\theta_0))}f(x,\theta_0)\mathrm{d}x\right|\\
\le&\left|\int\frac{\alpha_n(x)-B(F(x,\theta_0))+\nabla_\theta^\tau F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}[B(F(t,\theta_0))-\alpha_n(t)]}{[F(x,\theta_0)(1-F(x,\theta_0))]^\alpha}f(x,\theta_0)\mathrm{d}x\right|\\
&\cdot\sup_x\frac{|\alpha_n(x)+B(F(x,\theta_0))-\nabla_\theta^\tau F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}[B(F(t,\theta_0))+\alpha_n(t)]|}{[F(x,\theta_0)(1-F(x,\theta_0))]^{1-\alpha}}+o_p(1)\\
\le&\left[\left|\int\frac{\alpha_n(x)-B(F(x,\theta_0))}{[F(x,\theta_0)(1-F(x,\theta_0))]^\alpha}f(x,\theta_0)\mathrm{d}x\right|\right.\\
&\left.+\left|\int\frac{\nabla_\theta^\tau F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}[B(F(t,\theta_0))-\alpha_n(t)]}{[F(x,\theta_0)(1-F(x,\theta_0))]^\alpha}f(x,\theta_0)\mathrm{d}x\right|\right]\cdot O_p(1)+o_p(1)\\
=&[o_p(1)+o_p(1)]O_p(1)+o_p(1)=o_p(1).
\end{aligned}\tag{12}$$

最后一个等号成立是利用 [4] 中定理 3.5, [8] 中定理 2 以及条件 2.

另一方面，再由 [4] 中定理 3.5 易得 $\int_0^1 G^2(t,\theta_0)\mathrm{d}t<\infty$, a.s. 由此并结合 (11) 和 (12) 式易得定理结论.

3　加权经验过程的渐近分布

对一般的权函数情况，我们先给出两个引理.

引理 3.1 如果 $q\in FC_{0,1}$, 则

$$\sup_{0<t<1}\Big|\frac{n^{\frac{1}{2}}[F_n(F^{-1}(t,\theta_0),\theta_0)-t]-B(t)}{q(t)}\Big|=o_p(1)$$

当且仅当对任意 $c>0, I(q,c)<\infty$.

证　见 [4] 中定理 3.2.

类似引理 2.1, 有

引理 3.2 记

$$G_q(t,\theta)=\frac{B(t)}{q(t)}-\frac{\nabla_\theta F(F^{-1}(t,\theta),\theta)}{q(t)}\int l(F^{-1}(s,\theta),\theta)\mathrm{d}B(s). \tag{13}$$

则 $\forall\theta\in\Theta$, $\{G_q(t,\theta),0\le t\le 1\}$ 为一中心化 Guassian 过程，其协方差 $\mathrm{Cov}(G_q(t,\theta),G_q(s,\theta))$ 为

$$\begin{aligned}&\frac{\min\{s,t\}-st}{q(t)q(s)}-\Big(\frac{\nabla_\theta F(F^{-1}(t,\theta),\theta)}{q(t)}\Big)^\tau\frac{\int_0^s l(F^{-1}(y,\theta),\theta)\mathrm{d}y}{q(s)}\\&-\Big(\frac{\nabla_\theta F(F^{-1}(s,\theta),\theta)}{q(t)}\Big)^\tau\frac{\int_0^t l(F^{-1}(y,\theta),\theta)\mathrm{d}y}{q(t)}\\&+\Big(\frac{\nabla_\theta F(F^{-1}(t,\theta),\theta)}{q(t)}\Big)^\tau M(\theta)\Big(\frac{\nabla_\theta F(F^{-1}(s,\theta),\theta)}{q(t)}\Big).\end{aligned}$$

对加权经验过程的上界型统计量，有如下结论.

定理 3.1 假设条件 1,4,5 成立，则在零假设 (1) 成立时，有

$$\widehat{KS}_n(q)=\sup_{x\in\boldsymbol{R}}\Big|\frac{n^{\frac{1}{2}}[F_n(x)-F(x,\widehat{\theta}_n)]}{q(F(x,\widehat{\theta}_n))}\Big|\xrightarrow{\mathcal{D}}\sup_{x\in\boldsymbol{R}}\left|G_q(F(x,\theta_0),\theta_0)\right|.$$

证　由定理条件并通过 Taylor 展开，容易得到

$$\Big|\sup_{x\in\boldsymbol{R}}\frac{n^{\frac{1}{2}}[F_n(x)-F(x,\widehat{\theta}_n)]}{q(F(x,\widehat{\theta}_n))}-\sup_{x\in\boldsymbol{R}}\frac{n^{\frac{1}{2}}[F_n(x)-F(x,\widehat{\theta}_n)]}{q(F(x,\theta_0))}\Big|=o_p(1). \tag{14}$$

而由条件 1 知

$$\begin{aligned}&\frac{n^{\frac{1}{2}}[F_n(x)-F(x,\widehat{\theta}_n)]}{q(F(x,\theta_0))}\\&=\frac{\alpha_n(x)}{q(F(x,\theta_0))}-\frac{\nabla_\theta F(x,\theta_0)[\sqrt{n}(\widehat{\theta}_n-\theta_0)+O_p(n^{-\frac{1}{2}})]}{q(F(x,\theta_0))}\\&=\frac{\alpha_n(x)}{q(F(x,\theta_0))}-\frac{\nabla_\theta F(x,\theta_0)}{q(F(x,\theta_0))}\int l(t,\theta_0)\mathrm{d}\alpha_n(t)+O_p(n^{-\frac{1}{2}}).\end{aligned}\tag{15}$$

结合 (14) 和 (15) 式可得

$$\begin{aligned}
&\sup_{x\in\boldsymbol{R}}\left|\widehat{V}_n(x;q)-G_q(F(x,\theta_0),\theta_0)\right| \\
=&\sup_{x\in\boldsymbol{R}}\left|\frac{\alpha_n(x)}{q(F(x,\theta_0))}-\frac{\bigtriangledown_\theta F(x,\theta_0)}{q(F(x,\theta_0))}\int l(t,\theta_0)\mathrm{d}\alpha_n(t)+o_p(1)\right.\\
&\left.-\frac{B(F(x,\theta_0))}{q(F(x,\theta_0))}+\frac{\bigtriangledown_\theta F(x,\theta_0)}{q(F(x,\theta_0))}\int l(F^{-1}(s,\theta_0),\theta_0)\mathrm{d}B(s)\right| \\
\le&\sup_{x\in\boldsymbol{R}}\left|\frac{\bigtriangledown_\theta F(x,\theta_0)}{q(F(x,\theta_0))}\Big[\int l(F^{-1}(s,\theta_0),\theta_0)\mathrm{d}B(s)-\int l(t,\theta_0)\mathrm{d}\alpha_n(t)\Big]\right| \\
&+\sup_{x\in\boldsymbol{R}}\left|\frac{\alpha_n(x)}{q(F(x,\theta_0))}-\frac{B(F(x,\theta_0))}{q(F(x,\theta_0))}\right|+o_p(1).
\end{aligned}$$

由于 $\alpha_n(x)\to B(F(x,\theta_0))$, 所以 $\left|\int l(F^{-1}(s,\theta_0),\theta_0)\mathrm{d}B(s)-\int l(t,\theta_0)\mathrm{d}\alpha_n(t)\right|=o_p(1)$, 再结合引理 3.1 和引理 3.2 可得

$$\sup_{x\in\boldsymbol{R}}\left|\widehat{V}_n(x;q)-G_q(F(x,\theta_0),\theta_0)\right|=o_p(1).$$

再由 [4] 中系 1.1 可得 $\sup\limits_{x\in\boldsymbol{R}}|G_q(F(x,\theta_0),\theta_0)|<\infty$. 因此定理结论成立.

注 3.1 定理 2.1 虽然形式上与本定理相似，但不属于本定理的特殊情况. 因为对任意常数 $c>0$,

$$I(q,c)=\int_0^1\frac{1}{t(1-t)}\exp\Big(-\frac{ct(1-t)}{t(1-t)}\Big)\mathrm{d}t=\infty.$$

不满足本定理的条件，故只能在区间 $[\mu_n,1-\mu_n]$ 内讨论，不考虑其尾部情况.

对加权经验过程的积分型统计量，有如下结论.

定理 3.2 假设条件 1, 3, 4, 6 成立，则在零假设 (1) 成立时，有

$$\widehat{\omega}_n^2(q)=\int\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{q^2(F(x,\widehat{\theta}_n))}\mathrm{d}F(x,\widehat{\theta}_n)\xrightarrow{\mathcal{D}}\int_0^1G_q^2(t,\theta_0)\mathrm{d}t.$$

其中 $G_q(t,\theta_0)$ 见 (13) 式.

证 由定理条件并通过 Taylor 展开，容易得到

$$\Big|\int_{-\infty}^{\infty}\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{q^2(F(x,\widehat{\theta}_n))}\mathrm{d}F(x,\widehat{\theta}_n)-\int_{-\infty}^{\infty}\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{q^2(F(x,\theta_0))}\mathrm{d}F(x,\theta_0)\Big|=o_p(1).\tag{16}$$

由本定理的条件易得

$$\begin{aligned}
&\left|\widehat{\omega}_n^2(q)-\int_0^1G_q^2(t,\theta_0)\mathrm{d}t\right| \\
\le&\left|\int_{-\infty}^{\infty}\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{q^2(F(x,\widehat{\theta}_n))}\mathrm{d}F(x,\widehat{\theta}_n)-\int_{-\infty}^{\infty}\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{q^2(F(x,\theta_0))}\mathrm{d}F(x,\theta_0)\right| \\
&+\left|\int_{-\infty}^{\infty}\frac{n[F_n(x)-F(x,\widehat{\theta}_n)]^2}{q^2(F(x,\theta_0))}\mathrm{d}F(x,\theta_0)-\int_0^1G_q^2(t,\theta_0)\mathrm{d}t\right|
\end{aligned}$$

$$
\begin{aligned}
&= \Bigg| \int \frac{\left\{\alpha_n(x) - \nabla_\theta F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}\alpha_n(t)\right\}^2}{q^2(F(x,\theta_0))} f(x,\theta_0)\mathrm{d}x + o_p(1) \\
&\quad - \int \frac{\left\{B(F(x,\theta_0)) - \nabla_\theta F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}B(F(t,\theta_0))\right\}^2}{q^2(F(x,\theta_0))} f(x,\theta_0)\mathrm{d}x \Bigg| + o_p(1) \\
&\le \int \left| \frac{\alpha_n(x) - B(F(x,\theta_0)) + \nabla_\theta F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}\left[B(F(t,\theta_0)) - \alpha_n(t)\right]}{[q(F(x,\theta_0))]^{2-\gamma}} \right| f(x,\theta_0)\mathrm{d}x \\
&\quad \cdot \sup_x \frac{\left|\alpha_n(x) + B(F(x,\theta_0)) - \nabla_\theta F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}\left(B(F(t,\theta_0)) + \alpha_n(t)\right)\right|}{[q(x)]^\gamma} + o_p(1) \\
&\le o_p(1) + \Bigg[\int \left| \frac{\alpha_n(x) - B(F(x,\theta_0))}{[q(F(x,\theta_0))]^{2-\gamma}} \right| f(x,\theta_0)\mathrm{d}x \\
&\qquad + \int \left| \frac{\nabla_\theta F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}\left[B(F(t,\theta_0)) - \alpha_n(t)\right]}{[q(F(x,\theta_0))]^{2-\gamma}} \right| f(x,\theta_0)\mathrm{d}x \Bigg] O_p(1). \qquad (17)
\end{aligned}
$$

又由于 $\frac{\int_0^1 [t(1-t)]^{\frac{1}{2}}}{(q(t))^{2-\gamma}}\mathrm{d}t < \infty$, 利用文献 [4] 中定理 3.5 易得

$$
\int \left| \frac{\alpha_n(x) - B(F(x,\theta_0))}{[q(F(x,\theta_0))]^{2-\gamma}} \right| f(x,\theta_0)\mathrm{d}x = o_p(1). \qquad (18)
$$

由条件 4,6 并结合 [4] 中系 1.1 可得

$$
\int \left| \frac{\nabla_\theta F(x,\theta_0)\int l(t,\theta_0)\mathrm{d}\left[B(F(t,\theta_0)) - \alpha_n(t)\right]}{[q(F(x,\theta_0))]^{2-\gamma}} \right| f(x,\theta_0)\mathrm{d}x = o_p(1). \qquad (19)
$$

结合 (17)–(19) 式易得

$$
\left| \int \widehat{V}_n^2(x;q)\mathrm{d}F(x,\widehat{\theta}_n) - \int_0^1 G_q^2(t,\theta_0)\mathrm{d}t \right| = [o_p(1) + o_p(1)]O_p(1) + o_p(1) = o_p(1).
$$

由此得定理结论.

注 3.2　显然 $q^2(t) = t(1-t)$ 满足本定理的要求，故定理 2.2 为本定理的一个特例.

注 3.3　可以验证，正态分布族，指数分布等常见的参数分布族中参数的极大似然估计均满足本定理的条件.

参　考　文　献

[1] Berk R H, Jones D H. Relatively optimal combinations of test statistics. *Scand. J. Statist.*, 1978, **5**: 158–162.

[2] Jager L, Wellner J A. A new goodness of fit test: The reversed Berk-Jones statistic. Technical report, Department of Statistics, University of Washington. http://www.stat.washington.edu/www/research/reports/2004/tr443.ps, 2004.

[3] Zhang J J. Generalized nonparametric likelihood ratio for goodness-of-fit tests. Ph. D Dissertation, Chinese Academy of Sciences, 2006(in Chinese).

[4] Csörgő M, Horváth L. Weighted Approximations in Prability and Statistics. John Wiley and Sons, England, 1993.

[5] Darling D A. The Cramér-Smirnov test in the parametric case. *Ann. Math. Statist.,* 1955, **26**: 1–20.

[6] Del Barrio E. Empirical and quantile processes in the asymptotic theory of goodness-of-fit test. http://www.eio.uva.es/ems/Goodness_of_fit−Laredo_2004.pdf, 2004.

[7] Genz M, Haeusler E. Empirical processes with estimated parameters under auxiliary information. *J. Com. App. Math.,* 2006, **186**: 191–216.

[8] Shorack G R, Wellner J A. Empirical Processes with Applications to Statistics. New York: Wiley, 1986.

[9] Jaeschke D. The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. *Ann. Statist.*, 1979, **7**(1): 108–115.

[10] Babu G J, Rao C R. Goodness-of-fit tests when parameters are estimated. *Sankhya.,* 2004, **66**: 63–74.

THE WEIGHTED EMPIRICAL PROCESSES WITH ESTIMATED PARAMETERS

ZHANG Junjian

(*College of Applied Sciences, Beijing University of Technology, Beijing* 100124;

College of Mathematics, Guangxi Normal University, Guilin 541004)

Abstract The theory of weighted empirical processes is an important element for (generalized) nonparametric likelihood ratio goodness of fit test. There are few papers to discuss the weighted empirical processes with estimated parameters. This paper is concerned with the approximation distribution of the weighted empirical processes systematically. The results will provide a certain theoretical basis for (generalized) nonparametric likelihood ratio goodness of fit test with estimated parameters.

Key words The weighted empirical processes, estimated parameter, goodness of fit.