Задачи и решения

Решения

Решение 8.1

Рассмотрим регрессию на стационарных временных рядах

$$y_t = \alpha + \beta x_t + e_t,$$

где ошибка e_t имет нулевое среднее и не коррелирует с регрессором x_t. Нас интересует асимптотическая дисперсия V_β МНК-оценки β.

(а) Постройте пример, в котором e_t серийно коррелирована, но тем не менее

$$V_\beta = \frac{\mathbb{V}[e_t]}{\mathbb{V}[x_t]},$$

то есть можно не поправлять стандартные ошибки на серийную корреляцию, хотя таковая имеется.

(б) Постройте пример, в котором e_t серийно некоррелирована, но тем не менее

$$V_\beta \neq \frac{\mathbb{V}[x_t e_t]}{(\mathbb{V}[x_t])^2}.$$

то есть нужно поправлять стандартные ошибки на серийную корреляцию, хотя таковой в ошибках нет.

(в) Какие последствия для предварительного тестирования на серийную корреляцию имеют результаты в (а) и (б)?

Данная задача была опубликована как Anatolyev (2001).

(a) Пусть x_t – НОД-процесс с нулевым средним, независимый от e_t, тогда для $j \neq 0$, хотя и $\mathbb{C}(e_t, e_{t-j}) \neq 0$, выполнено

$$\mathbb{C}[x_t e_t, x_{t-j} e_{t-j}] = \mathbb{C}[x_t, x_{t-j}]\,\mathbb{C}[e_t, e_{t-j}] = 0$$

(где первое равенство следует из того, что x_t – НОД с нулевым средним), в то время как

$$\mathbb{V}[x_t e_t] = \mathbb{V}[x_t]\,\mathbb{V}[e_t],$$

так что

$$V_\beta = \frac{\mathbb{V}[e_t]}{\mathbb{V}[x_t]}.$$

(b) Пусть x_t – процесс с нулевым средним и единичной дисперсией, следующий авторегрессии первого порядка, порожденной строгим белым шумом:

$$x_t = \rho x_{t-1} + \eta_t, \quad \eta_t \sim IID\left(0, 1-\rho^2\right),$$

где $0 < \rho < 1$. Пусть стандартный двумерный нормальный белый шум $(v_t, w_t)'$ независим от процесса x_t. Построим e_t как

$$e_t = v_{t+1} + w_{t+1} x_{t+1} + \theta(v_t - w_t x_t),$$

где $0 < \theta < 1$. Тогда можно обнаружить, что

$$\mathbb{C}\left[x_t, e_t\right] = 0, \quad \mathbb{C}\left[e_t, e_{t-j}\right] = 0 \quad \forall\ j \neq 0$$

и

$$\mathbb{C}\left[x_t e_t, x_{t-1} e_{t-1}\right] \neq 0, \quad \mathbb{C}\left[x_t e_t, x_{t-j} e_{t-j}\right] = 0 \quad \forall\ j > 1.$$

Поэтому

$$V_\beta = \frac{\mathbb{V}\left[x_t e_t\right] + 2\mathbb{C}\left[x_t e_t, x_{t-1} e_{t-1}\right[}{\left(\mathbb{V}\left[x_t\right]\right)^2} \neq \frac{\mathbb{V}\left[x_t e_t\right]}{\left(\mathbb{V}\left[x_t\right]\right)^2}.$$

(c) Следовательно, более разумно тестировать на серийную корреляцию процесс $x_t e_t$, а не процесс e_t. Аналогично, следует тестировать процесс $z_t e_t$, если рассматривается оценивание с инструментальными переменными с вектором инструментов z_t.

Решение 8.2

Рассмотрим стандартную панельную линейную регрессию с индивидуальными эффектами. Пусть для простоты константа отсутствует, а все переменные имеют нулевое среднее. Разработайте тест на случайные эффекты на основе разности внутри-оценки и между-оценки коэффициентов регрессии.

Пусть $P = I_n \otimes \frac{1}{T}\iota\iota\prime$, $Q = I_{nT} - P$, и вспомним, что $\hat\beta_W - \beta = \left(X'QX\right)^{-1} X'QU$ и $\hat\beta_B - \beta = \left(X'PX\right)^{-1} X'PU$, где U – вектор двухкомпонентной ошибки, и что $\mathbb{V}\left[U|X\right] = \alpha_Q Q + \alpha_P P$ для определенных весов α_Q и α_P, зависящих от дисперсий компонент. Тогда, условно на X,

$$\begin{aligned}
\mathbb{C}[\hat\beta_W, \hat\beta_B] &= \left(X'QX\right)^{-1} X'Q\left(\alpha_Q Q + \alpha_P P\right) PX\left(X'PX\right)^{-1} \\
&= \alpha_Q\left(X'QX\right)^{-1} \mathcal{X}'QQPX\left(X'PX\right)^{-1} + \alpha_P\left(X'QX\right)^{-1} X'QPPX\left(X'PX\right)^{-1} \\
&= 0,
\end{aligned}$$

поскольку $QP = PQ = 0$. Используя этот результат, имеем, условно на X,

$$\begin{aligned}
\mathbb{V}[\hat\beta_W - \hat\beta_B] &= \mathbb{V}[\hat\beta_W] + \mathbb{V}[\hat\beta_B] \\
&= \alpha_Q\left(X'QX\right)^{-1} + \alpha_P\left(X'PX\right)^{-1}.
\end{aligned}$$

Известно, что при случайных эффектах $\hat\beta_W$ и $\hat\beta_B$ состоятельны и асимптотически нормальны, поэтому тестовая статистика

$$\mathcal{R} = \left(\hat\beta_W - \hat\beta_B\right)'\left(\hat\alpha_Q\left(X'QX\right)^{-1} + \hat\alpha_P\left(X'PX\right)^{-1}\right)^{-1}\left(\hat\beta_W - \hat\beta_B\right)$$

асимптотически распределена как хи-квадрат с $k = \dim(\beta)$ степенями свободы при случайных эффектах. Здесь $\hat\alpha_Q = RSS_W/\left(nT - n - k\right)$ и $\hat\alpha_P = RSS_B/\left(n - k\right)$. Когда случайные эффекты неприемлемы, оценка $\hat\beta_W$ по-прежнему состоятельна, но $\hat\beta_B$ несостоятельна, так что $\hat\beta_W - \hat\beta_B$ сходится к ненулевомй пределу, в то время как $X'QX$ и $X'PX$ расходятся, из-за чего $\mathcal{R}$ тоже расходится.

Решение 8.3

Выведите асимптотические свойства оценки Надарая–Уотсона регрессионной функции в регрессии с абсолютной подгонкой, т.е. когда регрессионной ошибки нет.

Используем обозначения, принятые в Анатольев (2009). Когда дисперсия ошибки равна нулю,

$$\hat{g}(x) - g(x) = \frac{(nh)^{-1}\sum_{i=1}^{n}(g(x_i) - g(x))\,K\left(h^{-1}(x_i - x)\right)}{(nh)^{-1}\sum_{i=1}^{n}K\left(h^{-1}(x_i - x)\right)}.$$

Обычного источника дисперсии, регрессионной ошибки, в данном случае нет, так что дисперсия оценки порождается дисперсией регрессоров.

Знаменатель сходится к $f(x)$ при $n \to \infty$, $h \to 0$. Рассмотрим числитель, обозначив его как $\hat{q}(x)$. Он представляет собой среднее НОР случайных величин, скажем ζ_i. Следуя Анатольев (2009), получаем, что

$$\mathbb{E}\left[\hat{q}(x)\right] = h^2 B(x) f(x) + o(h^2), \quad \mathbb{V}\left[\hat{q}(x)\right] = o((nh)^{-1}).$$

Взглянем попристальней на дисперсию.

$$\begin{aligned}
\mathbb{E}\left[\zeta_i^2\right] &= h^{-2}\int (g(x_i) - g(x))^2 K\left(\frac{x_i - x}{h}\right)^2 f(x_i)dx_i \\
&= h^{-1}\int (g(x + hu) - g(x))^2 K(u)^2 f(x + hu)du \\
&= h^{-1}\int \left(g'(x)hu + o(h)\right)^2 K(u)^2 (f(x) + o(h))\,du \\
&= hg'(x)^2 f(x)\int u^2 K(u)^2\,du + o(h),
\end{aligned}$$

так что

$$\mathbb{V}\left[\zeta_i\right] = \mathbb{E}\left[\zeta_i^2\right] - \mathbb{E}\left[\zeta_i\right]^2 = hg'(x)^2 f(x)\Psi_K^2 + o(h),$$

где $\Psi_K^2 = \int u^2 K(u)^2\,du$. Поэтому, согласно ЦПТ,

$$\begin{aligned}
&\sqrt{nh^{-1}}\left((nh)^{-1}\sum_{i=1}^{n}(g(x_i) - g(x))\,K\left(\frac{x_i - x}{h}\right) - h^2 B(x) f(x) + o(h^2)\right) \\
&\xrightarrow{d} \mathcal{N}\left(0, g'(x)^2 f(x)\Psi_K^2\right).
\end{aligned}$$

Пусть $\lambda = \lim_{n\to\infty, h\to 0}\sqrt{nh^3}$, и предположим, что $\lambda < \infty$. Тогда

$$\sqrt{nh^{-1}}\left(\hat{g}(x) - g(x)\right) \xrightarrow{d} \mathcal{N}\left(\lambda B(x), \frac{g'(x)^2}{f(x)}\Psi_K^2\right).$$

Список литературы

Анатольев, С. (2009). Непараметрическая регрессия. *Квантиль* 7, 37–52.

Anatolyev, S. (2001). Serial correlation and asymptotic variance. *Econometric Theory* 17, задача 01.5.3, стр. 1026.