Лекция №17
Завершение теста Рамзи и дополнительные сюжеты
План

1. Завершение обсуждения теста Рэмзи первой предпосылки теоремы Гаусса-Маркова;
2. Тест Джарки-Бера нормального распределения случайного возмущения в модели;
3. Проблема несовершенной мультиколениарности в линейной модели и отбор объясняющих перменных методом дополнительной регрессии при помощи скорректированного коэффициента детерминации;
4. Вложенные эконометрические модели и тест Вальда ограничений во вложенной модели;

На прошлой лекции мы приступили к обсуждения теста первой предпосылки теоремы Гаусса-Маркова. Этот тест состоит из следующих шагов:

Шаг 1. Методом наименьших квадратов оценивается базовая модель эконометрики и вычисляются прогнозные значения:

$$\widetilde{u}_i = y_i - \left(\widetilde{a}_0 + \widetilde{a}_1 \cdot x_{1,i} + \ldots + \widetilde{a}_k \cdot x_{k,i}\right) \qquad (3.5.8)$$

Прогнозные значения эндогенной переменной в модели; эти прогнозные значения $\left(\widetilde{a}_0 + \widetilde{a}_1 \cdot x_{1,i} + \ldots + \widetilde{a}_k \cdot x_{k,i}\right)$ мы обозначим $\widetilde{y}_i$.

Шаг 2. Задаётся целочисленное значение $(p = 2, 3)$ и оценивается вспомогательная модель (3.2):

$$\begin{cases} y_i = a_1 \cdot x_{1t} + a_2 \cdot x_{2t} + \ldots + a_k \cdot x_{kt} + b_1 \cdot \widetilde{y}_t^{\,2} + \ldots + b_p \cdot \widetilde{y}_t^{\,p} + \varepsilon_t, \\ \text{где } \widetilde{y}_i = \widetilde{a}_1 \cdot x_{1t} + \widetilde{a}_2 \cdot x_{2t} + \ldots + \widetilde{a}_k \cdot x_{kt} \end{cases}$$

По правилу (3.1) вычисляется статистика критерия предпосылки 1 теоремы Гаусса-Маркова.

$$RESET = RAM\left(\widetilde{u}_1, \widetilde{u}_2, \ldots, \widetilde{u}_n\right) =$$
$$= (ESS - ESS_{3.2}) : p - 1 / ESS_{3.2} : (n - (k + p - 1))$$

Если гипотеза H_0 справедлива и случайные возмущения в создоваемой модели нормально распределено, то статистика имеет F распределение Фишера с кол-ом степеней свободы $[(p-1),\ (n + (k + p - 1))]$. Если величина $RESET$ меньше чем квантиль распределения $1 - \alpha$, то принимается, если больше, то отвергается.

$$P_{RESET}(q|H_0) = P_{F_{(p-1),(n-(k+p-1))}}(q). \qquad (3.3)$$
$$RESET_{\text{крит}} = F_{(p-1),(n-(k+p-1))}(1 - \alpha) \qquad (3.4)$$

Тест Джарки-Бера нормального распределения случайного возмущения в модели

В этом тесте исследуется гипотеза о том, что в уравнениях наблюдений объекта случайные возмущения имеют нормальный закон распределения с 0 математическим ожиданием с одной и тойже дисперсией и являются независимыми случайными переменными.

Тест состоит из следующих шагов:

Шаг 1. Модель оценивается методом наименьших квадратов и вычисляются оценки случайных возмущений (остатки).

Шаг 2. По правилу (2.1) расчитывается статистика критерия гипотезы H_0

$$JB = JB\left(\widetilde{u}_1, \widetilde{u}_2, \dots, \widetilde{u}_n\right) = n \cdot \left\{\frac{sk^2}{6} + \frac{(kur-3)^2}{24}\right\}; \tag{2.1}$$

$$sk = m_3 / m_2^{\frac{3}{2}},\ kur = \frac{m_4}{(m_2)^2}, m_k = \frac{1}{n}\sum_{i=1}^{n}\left(\widetilde{(u_i)}\right)^k$$

m_k - оценка начального момента случайного остатка. Где sk обозначена оценка скошенности (ассиметрии). Для нормального закона скошенность равна 0. Символом kur обозначена оценка островершинности распределения (оценка куртозиса или эксцесса). Для нормального распределения истинное значение $kur = 3$. При справедливой гипотезе H_0 величина JB имеет распределение χ^2 с двумя степенями свободы (2.3):

$$P_{JB}(q|H_0) = P_{\chi_2^2}(q)$$

$$P(JB < JB_{\text{крит}}|\text{в}H_0) = \int_0^{JB_{\text{крит}}=\chi_2^2(1-\alpha)} p_{\chi_2^2}(q)dq = 1-\alpha$$

Гипотеза H_0, когда велина $JB < JB_{\text{крит}}$, то принимается, а если больше то отвергается

Проблема несовершенной мультиколениарности в линейной модели и отбор объясняющих перменных методом дополнительной регрессии при помощи скорректированного коэффициента детерминации

Отметили критерий мультиколениарности:

$$rk(X) < (k+1) \Leftrightarrow |Cor(\vec{x}, \vec{x})| = 0$$

На практике часто встречаются ситуации, когда упомянутый выше определитель не равен 0, но очень мал. Тогда говорят, что в исходной модели присутствует несовершенная мультиколениарность. Отметим симптомы несовершенной мультиколениарности.

Симптомы:

Симптом №1. F – тест отвергает гипотезу:

$$H_0 : a_1 = a_2 = \dots = a_k = 0,$$

но t – тест **не отвергает** гипотезу $H_0 : a_j = 0$ о незначимости многих x_j.

Симптом № 2. При добавлении в обучающую выборку одного уравнения (или удаления одного из уравнений), оценки коэффициентов модели резко меняют свои значения, то есть оценки неустойчивые $\widetilde{a}_j$ и значит они не надёжные.

При несовершенной мультиколлениарности часто используется процедура пошаговой регресии отбора в модель объясняющих переменных. Процедура состоит из следующих шагов:

Шаг № 1. Модель оценивается по всем объясняющим переменным и отмечается скорректированный коэффициент детерминации (обозначим символом $\overline{R}_0^2$).

Шаг № 2. Из исходной оценённой модели выбирается коэффициент $\widetilde{a}_j$ с максимальным значением $|t_j| = \max x_j$. Удаляется из модели, переоценивается и запоминается значение скорректированной детерминации. Если $\overline{R}_0^2 \geqslant \overline{R}_1^2$, то удаление x_j нецелесобразно.

Шаг №3. Шаг №1 повторяется пока условие $\overline{R}_j^2 \geqslant \overline{R}_{j+1}^2$ перестанет повторяться