

Семинар №5

Линейная модель множественной регрессии и оценивание её параметров при помощи функции line

1. Проверка адекватности модели Самуэльсона-Хигса для гос. Расходов России
2. Модифицированная модель Самуэльсона-Хигса расходов домохозяйств в России и её оценивание при помощи функции line excel
3. Проверка ДЗ

На занятии вычислили значения случайных возмущений по правилу:

$$\vec{w}_t = G_t - \tilde{g} G_{t-1} \quad (4)$$

$$\tilde{\sigma}_w = \sqrt{\frac{\sum \tilde{w}_t^2}{n - k}} \approx 140 \quad (5)$$

Таким образом **третий этап** схемы завершается записью оценённой модели:

$$\begin{cases} G_t = 1.008 \cdot \tilde{g} \cdot G_{t-1} + w_t \\ \tilde{\sigma}_w = 140 \end{cases} \quad (6)$$

На четвёртом этапе осуществляется прогноз по оценённой модели значений эндогенных переменных из контролирующей выборки. В нашем примере рассчитывается расход на 2018 год. После расчёта прогноза вычисляется относительная ошибка прогноза.

$$\tilde{G}_{2018} = \tilde{g} (= 1.008) \cdot G_{2017} \quad (7)$$

Модель является адекватной, если относительная ошибка прогноза не превосходит 15% от прогнозируемых величин.

$$\delta = 100 \cdot |\tilde{G}_{2018} - G_{2018}| \leq 15\% \quad (8)$$

Задача.

Вычислить по правилу (7) прогноз и проверить адекватность модели:

$$\tilde{G}_{2018} = \tilde{g} \cdot G_{2017} = 1.008 \cdot 7264.2719268 = 7322.386$$

$$\delta = 100 \cdot |7320 - 7322.386| = 0.09\%$$

Ещё один вариант суждения об адекватности модели базируется на правиле 2-3 сигм (σ): модель признаётся адекватной, если абсолютные ошибки прогноза

не превосходя $2-3 \sigma$ в нашем примере мы бы признали модель (6) адекватной, если абсолютная ошибка прогноза:

$$e(\tilde{G}_{2018}) = |\tilde{G}_{2018} - G_{2018}| \leq 2 \cdot \tilde{\sigma}_w = 280 \text{ млрд. руб.}$$

Модифицируем модель Самуэльсона-Хикса при помощи более глубокого обсуждения диаграммы рассеивания. Вернёмся к диаграмме рассеивания ”лаговое ВВП России - текущее потребление домохозяйств” и внимательно изучим эту диаграмму, обращая внимание на наличие явных выбросов. Рассматривая диаграмму можем сделать следующий вывод: первый очевидный выброс датируется 2009 годом C это значит, что его причиной является мировой финансовый кризис. Остальные выбросы датируются следующими годами: 2015, 2016, 2017. И их причиной являются санкции западных стран.

Итог: модель Самуэльсона-Хикса нужно модифицировать отразив в ней воздействие мирового кризиса и санкции западных стран. Вот модифицированных фрагментов модели Самуэльсона-Хикса расхода домохозяйств России.

Обозначим фиктивную переменную связанную с мировым финансовым кризисом Gr_t она равна 0, если в период t кризис отсутствует и 1, если существует

$$Gr_t = \begin{cases} 0, & \text{если в период } t \text{ кризис отсутствует} \\ 1, & \text{если существует} \end{cases} \quad (9)$$

Gr_t — это *индикатор кризиса*. Аналогично индикатор кризиса:

$$San_t = \begin{cases} 1, & \text{если существует} \\ 0, & \text{если отсутствует} \end{cases} \quad (9)$$

С помощью этих величин модифицируем модель Самуэльсона-Хикса:

$$\begin{cases} C_t = a_0 + a_1 \cdot Y_{t-1} + a_2 Gr_t + a_3 \cdot San_t + u_t \\ E(u_t) = 0; Var(u_t) = \sigma^2 \end{cases} \quad (10)$$

Спецификация (10) включает в себя 5 параметров: $(a_0, a_1, a_2, a_3, \sigma_u)$ (11). Обратим внимание, что спецификация (10) служит конкретным примером базовой модели эконометрики, которая носит название *линейной модели множественной регрессии*. Добавим, что при определённых свойствах случайного возмущения u_t параметры модели (11) оптимально оцениваются методом наименьших квадратов и на сегодняшнем занятии мы познакомимся с функцией ЛИНЕЙН в которой запрограммирована процедура наименьших квадратов с которой мы познакомились на прошлом занятии при оценивании модели (1).

Оценивание параметров (11) при помощи функции ЛИНЕЙН

1. На листе Excel занесём символ $date$ времени t (2003 - 2017). Символом C_t - ВВП, Y_{t-1} - лаговый доход. Cr_t — кризис. San_t — санкции. $n = 15$. Ввели

заголовки и заполнили значениями переменных из обучающей модели. *Замечание.* Результат первого шага можно интерпретировать, как запись уравнений наблюдений в рамках модели (10)

2. Размещаем курсор со значениями эндогенных переменных (C_t) и кликаем по символу формул. В столбце категория выбираем статистические далее выбираем ЛИНЕЙН и кликаем ОК. В первую строчку заносим адрес массива эндогенной переменной. Во вторую строчку заносим объясняющие. В третьей и четвертой строчке следует вывести 1.
3. Запишем модель точно также как модель (1).

$$\begin{cases} C_t = -12107 + 0.84 Y_{t-1} - 3606 Gr_t - 2009 \cdot San_t + u_t \\ \tilde{\sigma}_w = 406.5 .. \end{cases} \quad (12)$$

Комментарий. В первой строчке выделенного массива (протокола функции линейн) расположены в обратном порядке оценки коэффициентов. Мера точности во второй строке. Величина $\tilde{\sigma}_w$ всегда содержится в 3 строке 2 столбца протокола. Остальное содержимое мы обсудим позже.

ДЗ Проанализировать диаграмму рассеивания остальных элементов диаграммы Самуэльсона-Хикса и если есть основания модифицировать отразив на низ влияние кризиса и санкций. Воспользоваться функцией линейн и оценить параметры двух остальных параметров модели Самуэльсона-Хикса.