

# Эконометрика. Лекция 6. Автокорреляция

11 декабря 2014 г.

# Автокорреляция!

Для проверки гипотез мы предполагали условную некоррелированность ошибок:

$$E(\varepsilon_i \varepsilon_j | X) = 0 \text{ при } i \neq j$$

Что произойдет если эта предпосылка будет нарушена?

# Когда логично ожидать автокорреляцию?

- $\langle \rangle$  наблюдений во времени или в пространстве
- наличие ненаблюдаемого фактора, действующего на  $\langle \rangle$  наблюдения

# Автокорреляцию подробно изучают!

- анализ временных рядов
- пространственная эконометрика

# Автокорреляция бывает небезобидной

- может привести к несостоятельности оценок  $\hat{\beta}$

$$\varepsilon_1 = \varepsilon_2 = \dots = \varepsilon_n = \pm 1$$

отметим, что  $E(\varepsilon_1 \varepsilon_2 | x) = 1$

# Автокорреляция может иметь очень сложную богатую структуру

- AR, MA, ARMA, ARIMA, VAR, VMA, VARMA, VECM, ARCH, GARCH, EGARCH, FIGARCH, TARCH, AVARCH, ZARCH, CCC, DCC, BEKK, VEC, DLM, ...

(тут можно страшными сокращениями заполнить весь экран)

# Мы рассмотрим автокорреляцию порядка $p$

- Начнем с автокорреляции первого порядка,  $p = 1$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$



# Предпосылки

- $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$
- $u_t$  — независимы между собой,
- $u_t$  независимы от регрессоров
- $u_t$  одинаково распределены
- $E(u_t) = 0, \text{Var}(u_t) = \sigma_u^2$

## упражнение у чудо-доски

Как выглядит  $\text{Corr}(\varepsilon_t, \varepsilon_{t-k})$  при автокорреляции первого порядка?

## Автокорреляция порядка $p$ :

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \dots + \phi_p \varepsilon_{t-p} + u_t$$

допускает более богатую структуру  $\text{Corr}(\varepsilon_i, \varepsilon_j)$

Как и в случае автокорреляции первого порядка,

$$\lim_{k \rightarrow \infty} \text{Corr}(\varepsilon_t, \varepsilon_{t-k}) = 0$$

## условная автокорреляция и другие предпосылки

- автоматом нарушена предпосылки о независимости наблюдений  $(x_i, y_i)$
- во временных рядах обычно нарушена предпосылка  $E(\varepsilon_t|X) = 0$

например, использование  $y_{t-1}$  в качестве регрессора нарушает  $E(\varepsilon_t|X) = 0$

(сказать про остальные предпосылки, и более слабые варианты)

## Мы используем прежние формулы:

- Для оценок коэффициентов:  $\hat{\beta} = (X'X)^{-1}X'y$
- Для оценки ковариационной матрицы оценок коэффициентов,  
 $\widehat{Var}(\hat{\beta}|X) = \frac{RSS}{n-k}(X'X)^{-1}$
- В частности,  $\widehat{Var}(\hat{\beta}_j|X) = \frac{\hat{\sigma}^2}{RSS_j}$  и  $se(\hat{\beta}_j) = \sqrt{\widehat{Var}(\hat{\beta}_j)}$

## Три группы свойств:

- конечная выборка без предположения о нормальности  $\varepsilon$
- конечная выборка с предположением о нормальности  $\varepsilon$
- асимптотические свойства (без предположения о нормальности  $\varepsilon$ )

Что происходит в каждом случае?

# Конечная выборка без предположения о нормальности $\varepsilon$

- Линейность по  $y$
- Условная несмещенность,  $E(\hat{\beta}|X) = \beta$
- (—) Оценки неэффективны

# Конечная выборка с предположением о нормальности $\varepsilon$

- $(-)\ \frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} | X \sim t_{n-k}$
- $(-)\ \frac{RSS}{\sigma^2} | X \sim \chi_{n-k}^2$
- $(-)\ \frac{(RSS_R - RSS_{UR})/r}{RSS_{UR}/(n-k)} \sim F_{r, n-k}$



# Асимптотические свойства:

- $\hat{\beta} \rightarrow \beta$
- $\frac{RSS}{n-k} \rightarrow \sigma^2$
- $(-)\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \rightarrow N(0, 1)$
- $(-)\frac{RSS_R - RSS_{UR}}{RSS_{UR}/(n-k)} \rightarrow \chi_r^2$

# Мораль:

- Сами  $\hat{\beta}$  можно интерпретировать и использовать
- Стандартные ошибки  $se(\hat{\beta}_j)$  несостоятельны
- Не можем строить доверительные интервалы для  $\beta_j$  и проверять гипотезы

# Что делать?

- Исправить стандартные ошибки!
- Другая формула для оценки  $\widehat{Var}_{HAC}(\hat{\beta}|X)$
- Следовательно, другие  $se_{HAC}(\hat{\beta}_j)$

## Робастная (устойчивая) к условной гетероскедастичности и автокорреляции оценка ковариационной матрицы

- Вместо  $\widehat{Var}(\hat{\beta}|X) = \frac{RSS}{n-k}(X'X)^{-1}$

ИСПОЛЬЗОВАТЬ

$$\widehat{Var}_{HAC}(\hat{\beta}|X) = (X'X)^{-1}\hat{\Phi}(X'X)^{-1}$$

- Нью-Вест (Newey-West), 1987 (Существует много вариантов)

$$\hat{\Phi} = \sum_{j=-k}^k \frac{k-|j|}{k} \left( \sum_t \hat{\varepsilon}_t \hat{\varepsilon}_{t+j} x'_t \cdot x_{t+j} \right)$$

# Суть корректировки:

Мы меняем  $se(\hat{\beta}_j)$  на  $se_{HAC}(\hat{\beta}_j)$

Какие проблемы решены?

- $\frac{\hat{\beta}_j - \beta_j}{se_{HAC}(\hat{\beta}_j)} \rightarrow N(0, 1)$  (УРА!)

# Какие проблемы не решены?

(—) оценки  $\hat{\beta}$  не меняются и остаются неэффективными даже при предположении о нормальности  $\varepsilon$ :

- (—)  $\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} | X \sim t_{n-k}$
- (—)  $\frac{RSS}{\sigma^2} | X \sim \chi^2_{n-k}$
- (—)  $\frac{(RSS_R - RSS_{UR})/r}{RSS_{UR}/(n-k)} \sim F_{r, n-k}$

## С практической точки зрения:

- Новая формула для  $\widehat{Var}_{HAC}(\hat{\beta}|X)$ , и, следовательно, для  $se_{HAC}(\hat{\beta}_j)$
- ковариационная матрица в R:

```
model <- lm(y~x, data=data)
vcovHAC(model)
```

- С ней жизнь прекрасна!

$$\frac{\hat{\beta}_j - \beta_j}{se_{HAC}(\hat{\beta}_j)} \rightarrow N(0, 1)$$

# Когда следует использовать

- Когда мы подозреваем наличие автокорреляции и не хотим заниматься её моделированием



# Обнаружение автокорреляции

- Оцениваем интересующую нас модель с помощью МНК
- Строим график остатков в осях  $\hat{\varepsilon}_{t-1}$ ,  $\hat{\varepsilon}_t$

/здесь пришлю три графика/

# Формальные тесты на автокорреляцию

- тест Дарбина-Уотсона (Durbin-Watson)
- тест Бройша-Годфри (Breusch-Godfrey)

## Тест Дарбина-Уотсона предпосылки:

- Автокорреляция первого порядка в остатках

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

- нормальность ошибок  $\varepsilon$
- сильная экзогенность,  $E(\varepsilon_t|X) = 0$
- $H_0$  об отсутствии автокорреляции,  $\rho = 0$

## процедура теста Дарбина-Уотсона

- Шаг 1. Оценить основную регрессию, получить  $\hat{\varepsilon}_i$
- Шаг 2. Посчитать статистику

$$DW = \frac{\sum_{i=2}^n (\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2}{\sum_{i=1}^n \hat{\varepsilon}_i^2}$$

# Распределение статистики $DW$

- $H_0$  об отсутствии автокорреляции,  $\rho = 0$
- Точный закон распределения сложным образом зависит от  $X$
- Если  $\hat{\rho}$  — выборочная корреляция остатков, то  $DW = 2(1 - \hat{\rho})$

## Качественные выводы по статистике $DW$

$DW = 2(1 - \hat{\rho})$ , поэтому  $0 < DW < 4$

- $DW \approx 0$  означает положительную автокорреляцию  $\hat{\rho} \approx 1$
- $DW \approx 2$  означает отсутствие автокорреляции  $\hat{\rho} \approx 0$
- $DW \approx 4$  означает отрицательную автокорреляцию  $\hat{\rho} \approx -1$

# иллюстрация (рисунок прилагается: график про Дарбина-Уотсона)

теховские надписи для графиков:

$H_0$  не отвергается  $H_0$  отвергается  $DW_{cr}$   $H_0: \rho = 0$

## С практической точки зрения:

- R рассчитывает точные P-значения для теста  $DW$
- существуют таблицы диапазонов критических значений



# Тест Бройша-Годфрей (Breusch-Godfrey)

- для тестирования автокорреляции порядка  $p$  в ошибках

$$\varepsilon_t = \phi_1 \varepsilon_{t-1} + \dots + \phi_p \varepsilon_{t-p} + u_t$$

- не требуется нормальность остатков
- верен при ряде нарушений предпосылки  $E(\varepsilon_t|X) = 0$
- асимптотический

$$H_0: \phi_1 = \phi_2 = \dots = \phi_p = 0$$

# Процедура теста Бройша-Годфри

- Шаг 1. Оцениваем исходную модель, получаем остатки  $\hat{\varepsilon}_t$
- Шаг 2. Строим вспомогательную регрессию  $\hat{\varepsilon}_t$  на исходные регрессоры,  $\hat{\varepsilon}_{t-1}, \hat{\varepsilon}_{t-2}, \dots, \hat{\varepsilon}_{t-p}$ , находим  $R_{aux}^2$
- Шаг 3. Считаем статистику  $BG = (n - p)R_{aux}^2$

## Тест Бройша-Годфри продолжение

- При верной  $H_0$  об отсутствии автокорреляции

$$H_0: \phi_1 = \phi_2 = \dots = \phi_p = 0$$

$$BG = (n - p)R_{aux}^2 \sim \chi_p^2$$

Здесь график распределения BG (рисунок прилагается) Подписи на графике:

$H_0$  не отвергается  $H_0$  отвергается  $\chi_{cr}^2$   $H_0: \phi_1 = \phi_2 = \dots = \phi_p = 0$

# Тест Бройша-Годфри требует меньше предпосылок

## Вставка с чудо-доской

Тест Дарбина-Уотсона и Бройша-Годфри (уже снят)

здесь в задаче было дано  $DW$ , надо было найти  $\hat{\rho}$  и провести тест Бройша-Годфри

- Мы рассмотрели ситуацию нарушения предпосылки условной некоррелированности ошибок модели
- Нарушена во временных рядах и пространственных данных
- В простейшем случае достаточно использовать специальные стандартные ошибки  $se_{HAC}$
- Большое количество специальных моделей