

Оценки OLS считаются по формуле: $\hat{\beta} = (X'X)^{-1}X'Y$
 Оценка дисперсии ошибки считается по формуле $\hat{\sigma}_u^2 = \frac{RSS}{n-k}$
 Оценка дисперсии оценок $\hat{Var}(\hat{\beta}) = (X'X)^{-1}\hat{\sigma}_u^2$

Парадигма 1. Неслучайные X .

Предпосылки.

A1. $y_t = \beta_1 + \beta_2 x_{2,t} + \dots + \beta_k x_{k,t} + u_t$

A2. X - константа

A3. $E(u_t) = 0$

A4. Среди X нет линейно зависимых столбцов, $n > k$

A5.1. Гомоскедастичность $Var(u_t) = \sigma_u^2$

A5.2. Отсутствие автокорреляции $Cov(u_t, u_j) = 0$

A6. u нормально распределен

Утверждения:

A5.1 и A5.2 можно заменить на A5. $Var(u) = \sigma_u^2 \cdot I_{n \times n}$

Существование $\hat{\beta}$

Если выполнено A4, тогда $\hat{\beta}$ можно посчитать. Иначе не получится обратить матрицу.

Линейность $\hat{\beta}$

$\hat{\beta}$ - всегда линейны по Y , если они существуют.

Несмещенность $\hat{\beta}$

Если выполнены A1-A4, тогда $\hat{\beta}$ - несмещенные

Формула для расчета дисперсии $\hat{\beta}$

Если A1-A5 выполнены, тогда $Var(\hat{\beta}) = (X'X)^{-1}\sigma_u^2$

Теорема Гаусса-Маркова

Эффективность $\hat{\beta}$ среди линейных несмещенных оценок

Если A1-A5 выполнены, тогда $\hat{\beta}^{OLS}$ более эффективна, чем любая другая оценка $\hat{\beta}^{nonOLS}$, обладающая линейностью и несмещенностью

Точнее:

$Var(\hat{\beta}_i^{OLS}) \leq Var(\hat{\beta}_i^{nonOLS})$

Еще точнее:

Матрица $Var(\hat{\beta}^{nonOLS}) - Var(\hat{\beta}^{OLS})$ является положительно определенной.

Несмещенность $\hat{\sigma}_u^2$

Если выполнены A1-A5, то $\hat{\sigma}_u^2$ несмещенная

Если выполнены A1-A6, то при любом n применимы тесты, в частности:

T1. Тест на значимость отдельного коэффициента.

$\sim t_{n-k}$

T2. Тест на значимость регрессии в целом

$\frac{ESS/(k-1)}{RSS/(n-k)} \sim F_{k-1, n-k}$

T3. Тест на выполнение нескольких линейных ограничений

$\frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(n-k)} \sim F_{q, n-k}$

Нарушения (или кажущиеся нарушения):

Оценки OLS считаются по формуле: $\hat{\beta} = (X'X)^{-1}X'Y$

Оценка дисперсии ошибки считается по формуле $\hat{\sigma}_u^2 = \frac{RSS}{n-k}$

Оценка дисперсии оценок $\hat{Var}(\hat{\beta}) = (X'X)^{-1}\hat{\sigma}_u^2$

Парадигма 2. Случайные X .

Предпосылки.

A1. $y_i = \beta_1 + \beta_2 x_{2,i} + \dots + \beta_k x_{k,i} + u_i$

A2. Вектор $(y_i, x_{2,i}, x_{3,i}, \dots, x_{k,i})$ является случайной выборкой. Т.е. векторы, соответствующие разным наблюдениям независимы и одинаково распределены

Уточнение: существует $Var((y_i, x_{2,i}, x_{3,i}, \dots, x_{k,i}))$

A3. $E(u|X) = 0$ (наилучший прогноз u при известных X - это ноль)

A3'. Ослабленный A3. $E(u) = 0$ и $E(uX) = 0$ (некоррелированность u и X)

A4. Среди регрессоров нет линейно зависимых (матрица X полного ранга), $n > k$

A5.1. Гомоскедастичность $Var(u_t|X) = \sigma_u^2$

A5.2. Отсутствие автокорреляции $Cov(u_i, u_j|X) = 0$

A6. u нормально распределен

Утверждения.

Из A3 следует A3'. Обратное неверно.

A5.1 и A5.2 можно заменить на A5. $Var(u|X) = \sigma_u^2 \cdot I_{n \times n}$

Существование $\hat{\beta}$

Если выполнено A4, тогда $\hat{\beta}$ можно посчитать. Иначе не получится обратить матрицу.

Линейность $\hat{\beta}$

$\hat{\beta}$ - всегда линейны по Y , если они существуют.

Несмещенность.

Если выполнены A1-A4, то $\hat{\beta}$ - несмещенные

Состоятельность.

Если выполнены A1, A2, A3', A4 то $\hat{\beta}$ - состоятельные

Формула для расчета условной дисперсии $\hat{\beta}$

Если A1-A3-A5 выполнены, тогда $Var(\hat{\beta}|X) = (X'X)^{-1}\sigma_u^2$

Теорема Гаусса-Маркова

Эффективность $\hat{\beta}$ среди линейных несмещенных оценок

Если A1-A3-A5 выполнены, тогда $\hat{\beta}^{OLS}$ более эффективна, чем любая другая оценка $\hat{\beta}^{nonOLS}$, обладающая линейностью и несмещенностью

Точнее:

$$Var(\hat{\beta}_i^{OLS}|X) \leq Var(\hat{\beta}_i^{nonOLS}|X)$$

Еще точнее:

Матрица $Var(\hat{\beta}^{nonOLS}|X) - Var(\hat{\beta}^{OLS}|X)$ является положительно определенной.

Асимптотическая т. Гаусса-Маркова

Если A1-A3'-A5 выполнены, тогда при $n \rightarrow \infty \dots$

Несмещенность $\hat{\sigma}_u^2$

Если выполнены A1-A3-A5, то $\hat{\sigma}_u^2$ несмещенная

Тесты разные

Если выполнены A1-A3-A6, то при любом n применимы тесты, в частности:

T1. Тест на значимость отдельного коэффициента.

$$\frac{\hat{\beta}_i - \beta_i}{sd} \sim t_{n-k}$$

T2. Тест на значимость регрессии в целом

$$\frac{ESS/(k-1)}{RSS/(n-k)} \sim F_{k-1, n-k}$$

T3. F-Тест на выполнение нескольких линейных ограничений

$$\frac{(RSS_R - RSS_{UR})/q}{RSS_{UR}/(n-k)} = \frac{(R_{UR}^2 - R_R^2)/q}{(1 - R_{UR}^2)/(n-k)} \sim F_{q, n-k}$$

Estimate both restricted and unrestriced model.

T4. Wald test. Тест на выполнение нескольких линейных ограничений

$$\frac{W}{q} = \frac{1}{q} (R\hat{\beta} - b)' (R' \hat{Var}(\hat{\beta}) R)^{-1} (R\hat{\beta} - b) \sim F_{q, (n-k)}$$

Estimate only unrestricted model.

Если выполнены A1-A3'-A5, то при $n \rightarrow \infty$ можно применять тесты, в частности:

T1. Тест на значимость отдельного коэффициента.

$$\frac{\hat{\beta}_i - \beta_i}{sd} \sim N(0; 1)$$

T2. Тест на значимость регрессии в целом

$$\frac{ESS}{RSS/(n-k)} \sim \chi_{k-1}^2$$

T3. Тест на выполнение нескольких линейных ограничений

$$\frac{(RSS_R - RSS_{UR})}{RSS_{UR}/(n-k)} = \frac{(R_{UR}^2 - R_R^2)}{(1 - R_{UR}^2)/(n-k)} \sim \chi_q^2$$

T4. Wald test. Тест на выполнение нескольких линейных ограничений

$$W = (R\hat{\beta} - b)' (R' \hat{Var}(\hat{\beta}) R)^{-1} (R\hat{\beta} - b) \sim \chi_q^2$$

Estimate only unrestricted model.

Легко запомнить: $t_n \rightarrow N(0; 1)$, $kF_{k,n} \rightarrow \chi_k^2$

Separately:

LM test for omitted variable.

H0: no omitted variables

Ha: at least one omitted variable

1. Run original regression.

2. Regress residuals on included and suspected omitted variables.

$nR^2 \sim \chi_q^2$, where q - number of omitted variables.

LR (likelyhood ratio test):

H0: restrictions

H1: at least one restriction is invalid

1. Estimate unrestricted model using maximum likelyhood. Obtain $\log(L_{UR})$

2. Estimate restricted model using maximum likelyhood. Obtain $\log(L_R)$

$2(\log(L_{UR}) - \log(L_R)) \sim \chi_q^2$, where q - number of restrictions

Maximum likelyhood method:

Maximize probability L of given observations.

Often y_i are independent and $L = P(Y_1 = y_1)P(Y_2 = y_2)...P(Y_n = y_n)$

Нарушения (псевдонарушения)

Ошибка измерения x .

Вместо нужного нам $x_{i,t}$ доступен только $\tilde{x}_{i,t} = x_{i,t} + w_{i,t}$. Т.е. в расчетах оценок вместо $x_{i,t}$ используется $\tilde{x}_{i,t}$.

Означает нарушение АЗ' (и, следовательно, АЗ)

Последствия:

Потеряна смещенность.

Потеряна состоятельность.

Неприменимы тесты.

Ошибка измерения y .

Вместо нужного нам y_t доступен только $\tilde{y}_t = y_t + w_t$. Т.е. в расчетах оценок вместо y_t используется \tilde{y}_t .

Ни одна предпосылка не нарушена.

Последствия:

Несмещенность сохраняется.

Состоятельность сохраняется.

Тесты применимы.

По сравнению с ситуацией доступного y_t потеряна эффективность.

Невключена нужная переменная.

Означает нарушение А1.

Последствия:

Потеряна смещенность.

Потеряна состоятельность.

Неприменимы тесты.

Включена лишняя переменная.

Означает нарушение А1

Последствия:

Несмещенность сохраняется.

Состоятельность сохраняется.

Тесты применимы.

Потеряна эффективность.

В качестве x используется дамми-переменная.

Ничего не нарушено.

В качестве y используется дамми-переменная

Нарушено АЗ' (и АЗ) (ошибка может принимать всего два значения)

Нарушено А5.1 (дисперсия ошибки зависит от регрессора)

Последствия:

Потеряна смещенность.

Потеряна состоятельность.

Неприменимы тесты.

Другие недостатки linear probability model:

1. Predicted probability may lay outside $[0; 1]$
2. Effect of change of regressor on probability is constant

Что делать? Использовать модели бинарного выбора (Logit, Probit)

Гетероскедастичность. Нарушено А5.1.

Причины: как правило, разница «размеров» объектов входящих в выборку. Например, если пытаться проанализировать, от чего зависит прибыль предприятия, то конечно окажется, что на крупных предприятиях сильные колебания прибыли, на маленьких - маленькие.

Последствия:

Несмещенность, состоятельность $\hat{\beta}$ - сохраняются.

Неверна формула расчета дисперсии $\hat{\beta}$. А раз все тесты ее используют, то все тесты (Т1-Т3) неприменимы.

Эффективность $\hat{\beta}$ - потеряна. Т.е. можно придумать альтернативную $\hat{\beta}^{nonOLS}$ с меньшей дисперсией.

Как бороться.

Если цель - получить несмещенные оценки - то бороться не надо.

Если цель - проверить гипотезы - то использовать исправленную формулу для дисперсии.

Если цель - получить эффективных оценки - то использовать GLS или (как правило GLS недоступен) FGLS.

Как обнаружить?

Тест Breush-Pagan. (LM-test, nR^2 test)

1. Run original regression, obtain residuals \hat{u}_i

2. Regress \hat{u}_i^2 on constant and all original regressors.

$nR^2 \sim \chi_q^2$, q - number of regressors in the second regression (excluding constant).

Тест White. (LM-test, nR^2 test)

1. Run original regression, obtain residuals \hat{u}_i

2. Regress \hat{u}_i^2 on constant, all original regressors, all original regressors squared, all pairwise products of original regressors.

$nR^2 \sim \chi_q^2$, q - number of regressors in the second regression (excluding constant).

Сравнение White vs Breush-Pagan: White допускает нелинейную зависимость дисперсии от регрессоров; White требует большего числа наблюдений.

Как бороться (детали)

За применимость тестов:

Оценки дисперсии (White) $Var_{White}(\hat{\beta}) = (X'X)^{-1}(X'uu'X)(X'X)^{-1}$

Если выполнены А1-А3'-А5, то при $n \rightarrow \infty$ можно применять Т1, Т4, заменив в них обычные оценки дисперсий $Var(\hat{\beta})$ на оценки White'a $Var_{White}(\hat{\beta})$.

За асимптотическую эффективность оценок:

1. Run the regression of y_i on constant, $x_{2,i}, \dots, x_{k,i}$ and obtain the residuals, \hat{u}_i .

2. Create $\log(\hat{u}_i^2)$ by first squaring the OLS residuals and then taking the natural log.

3. Run the regression of $\log(\hat{u}_i^2)$ on variables which determine variance. Obtain the fitted values, g_i .

4. Exponentiate the fitted values from $h_i = \exp(g_i)$.

5. Use WLS (weighted least squares), using weights $1/h_i$.

That means:

5.1. Create $y_i^* = \frac{y_i}{\sqrt{h_i}}$, $x_{i,j}^* = \frac{x_{i,j}}{\sqrt{h_i}}$

5.2. Run the regression of y_i^* on constant, $x_{2,i}^*, \dots, x_{k,i}^*$

All asymptotic tests are applicable.

Коррелированность disturbance term u_t and X . Нарушено А3. (и А3').

причины: наличие omitted variable (чаще всего unobservable), которая коррелирована с включенными регрессорами, случайная ошибка в измерении хотя бы одного регрессора

Последствия:

Потеряна смещенность.

Потеряна состоятельность.

Неприменимы тесты.

Как обнаружить?

Hausman-Wu test.

H0: all regressors X are uncorrelated with u

Ha: at least one of X from a given subset of X (not all X !) are correlated with u

Регрессоры поделены на две части. Про одну мы уверены, что она не коррелирована с u , про другую - нет.

Step 1. Строим регрессию каждого подозрительного регрессора на все инструментальные переменные. Obtain fitted values.

Step 2. Regress y on all original regressors and all fitted values. Это UR-regression.

Step 3. Regress y on all original regressors. Это R-regression.

Step 4. If $n \rightarrow \infty$ we may use T3 (comparison of RSSs using χ^2 distribution)

Hausman test.

H0: all regressors X are uncorrelated with u

Ha: at least one of X (maybe all) are correlated with u

$$H = (\hat{\beta}_{IV} - \hat{\beta}_{OLS})'(Var(\hat{\beta}_{IV}) - Var(\hat{\beta}_{OLS}))^{-1}(\hat{\beta}_{IV} - \hat{\beta}_{OLS}) \sim \chi^2$$

For Hausman-Wu and for Hausman:

under H0 both estimators $\hat{\beta}_{OLS}$ and $\hat{\beta}_{IV}$ are consistent, $\hat{\beta}_{OLS}$ is more efficient than $\hat{\beta}_{IV}$

under Ha only $\hat{\beta}_{IV}$ is consistent

Как бороться?

Найти инструментальную переменную, IV:

- коррелирована с регрессором (чем сильнее, тем лучше)

- некоррелирована с ошибкой

Метод IV (instrumental variables)

Для одного регрессора - одна инструментальная переменная

Z - матрица инструментальных переменных

$$\hat{\beta}_{IV} = (Z'X)^{-1}Z'Y$$

$$Var(\hat{\beta}_{IV}) = (Z'X)^{-1}Z'Z(X'Z)^{-1}\hat{\sigma}_u^2$$

Метод TSLS (two stage LS)

Для одного регрессора - несколько инструментальных переменных

Stage 1. Regress correlated variable x_{it} on instruments z_{1t}, \dots, z_{kt} . Obtain fitted values \hat{x}_{it} .

Stage 2. Replace correlated variables x_{it} in original regression by fitted values \hat{x}_{it} .

TSLS дает те же результаты, что IV, если число инструментов = число регрессоров

Conditional heteroskedasticity.

$h_t = \text{Var}_{t-1}(u_t)$ - conditional variance of u_t .

ARCH(1):

$$h_t = a + b_1 u_{t-1}^2$$

GARCH(1,1):

$$h_t = a + b_1 u_{t-1}^2 + c_1 h_{t-1}$$

TARCH(1,1):

$$h_t = a + \gamma u_{t-1}^2 d_{t-1} + b_1 u_{t-1}^2 + c_1 h_{t-1}$$

d_{t-1} - dummy variable. Usually $d_{t-1} = 1$ if $u_{t-1} < 0$ and $d_{t-1} = 0$ otherwise.

EGARCH:

$$\log(h_t) = a + b \log(h_{t-1}) + c \left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| + d \frac{u_{t-1}}{\sqrt{h_{t-1}}}$$

GARCH-M: (GARCH in mean)

$$y_t = \beta_1 + \beta_2 h_t + u_t$$

$$h_t = a + b_1 u_{t-1}^2 + c_1 h_{t-1}$$

Empirical: GARCH(1,1) - one of the most popular

News Impact Curve - зависимость h_t от u_{t-1} .

Оценки моделей получаются методом максимального правдоподобия.

TARCH and EGARCH capture the asymmetry property of

Logit-model

Используется, если Y_i принимает значения 0 и 1.

$$P(Y_i = 1) = f(\beta_1 + \beta_2 x_{2i} + \dots + \beta_k x_{ki}), \text{ где } f(z) = \frac{1}{1 + \exp(-z)}.$$

Такая $f(z)$ возрастает и всегда попадает в промежуток $[0; 1]$

Оценивают logit-model методом максимального правдоподобия.

Т.е. максимизируют функцию: $L = P(Y_1 = y_1)P(Y_2 = y_2) \cdot \dots \cdot P(Y_n = y_n)$

Интерпретация коэффициентов:

$$\frac{dP(Y_i=1)}{dx} = \beta P(Y_i = 1)P(Y_i = 0)$$

Т.е. знак β определяет направление зависимости:

Если $\beta > 0$, то вероятность $P(Y_i = 1)$ положительно зависит от x и наоборот

Для положительного β :

Если $P(Y_i = 1) > 0.5$ то имеет место убывающий предельный эффект:

Чем больше x , тем $P(Y_i = 1)$ выше, однако скорость роста $P(Y_i = 1)$ падает

Если $P(Y_i = 1) < 0.5$ то имеет место возрастающий предельный эффект:

Чем больше x , тем $P(Y_i = 1)$ выше и скорость роста $P(Y_i = 1)$ растет

Probit-model.

Аналогична logit-model. Отличие состоит в том, что используется другая функция $f(z)$ (тоже возрастает и гарантированно попадает в диапазон $[0; 1]$)

Интерпретация коэффициентов качественно такая же

$$\frac{dP(Y_i=1)}{dx} = \beta f'(z)$$