

Универсиада по эконометрике Финальный тур

Каждое задание оценивается в 25 баллов.

УДАЧИ!

Задание 1. Нестандартная оценка

В заведении 3 встретились эконометристы В и П. Между ними завязалась дискуссия о том, как оценить параметр β модели $y_i = \beta x_i + \varepsilon_i$ в условиях гетероскедастичности (и при выполнении прочих стандартных предпосылок классической модели со стохастическими регрессорами).

Эксцентричный В предложил вместо МНК использовать необычную оценку:

$$\widehat{\beta}_В = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{x_i}.$$

Его коллега П на это возразил, что такая оценка будет несостоятельной и неэффективной.

(а) (5) Верно ли, что эта оценка будет несостоятельной?

(б) (10) Верно ли, что не существует функционального вида условной дисперсии случайной ошибки модели $\sigma^2(x_i)$, при котором данная оценка все-таки окажется эффективной?

(в) (10) Сам П предлагает использовать оценку доступного обобщенного МНК:

$$\widehat{\beta}_П = \frac{\frac{1}{n} \sum_{i=1}^n \frac{x_i y_i}{\widehat{\sigma}^2(x_i)}}{\frac{1}{n} \sum_{i=1}^n \frac{x_i^2}{\widehat{\sigma}^2(x_i)}}.$$

Его коллега В утверждает, что такая оценка не будет состоятельной из-за того, что вместо истинной дисперсии случайной ошибки модели $\sigma^2(x_i)$ тут придется использовать её оценку $\widehat{\sigma}^2(x_i)$. Прав ли В?

Задание 2. Две ошибки измерения

Пусть переменные y_i^* и x_i^* связаны соотношением $y_i^* = \beta x_i^*, \beta > 0, i = 1, \dots, n$.

Исследователь наблюдает незашумленные значения зависимой переменной y_i^* , однако вместо истинных значений объясняющей переменной наблюдает только измеренные с ошибкой $x_i^{(1)}$ и $x_i^{(2)}$:

$$x_i^{(1)} = x_i^* + \varepsilon_i^{(1)}, \quad \text{где } \text{var}(\varepsilon_i^{(1)}) = \sigma_1^2 > 0, \quad \text{cov}(\varepsilon_i^{(1)}, \varepsilon_j^{(1)}) = 0 \text{ при } i \neq j;$$

$$x_i^{(2)} = x_i^* + \varepsilon_i^{(2)}, \quad \text{где } \text{var}(\varepsilon_i^{(2)}) = \sigma_2^2 > 0, \quad \text{cov}(\varepsilon_i^{(2)}, \varepsilon_j^{(2)}) = 0 \text{ при } i \neq j;$$

$$\text{cov}(\varepsilon_i^{(1)}, \varepsilon_i^{(2)}) = 0.$$

Таким образом, каждое значение объясняющей переменной измерено дважды, но оба раза с ошибкой. Также известно, что второе измерение является более точным, чем первое: $\sigma_1^2 > \sigma_2^2$ (исследователь знает обе дисперсии).

Цель исследователя состоит в том, чтобы состоятельно оценить коэффициент β . Он выбирает между тремя эмпирическими стратегиями (во всех случаях речь идет об оценивании при помощи обычного МНК):

1. **Использовать только самое точное измерение:** просто оценить регрессию y^* по $x^{(2)}$.
2. **Использовать оба измерения одновременно:** оценить регрессию y^* по $\frac{(x^{(1)} + x^{(2)})}{2}$.
3. **Последовательно использовать оба измерения:** оценить регрессию y^* по $x^{(1)}$. Потом оценить регрессию y^* по $x^{(2)}$. И в качестве окончательной оценки коэффициента β взять среднее арифметическое из двух полученных МНК-оценок.

(a) (15) Найдите предел по вероятности для каждой из трех оценок.

(б) (10) Какую из стратегий следует выбрать исследователю? Для ответа на этот вопрос укажите, какая из оценок является состоятельной. А если все они окажутся несостоятельными, то выберите «наименее несостоятельную» (то есть такую, что предел по вероятности для неё будет ближе всего к истинному значению коэффициента β).

Задание 3. Эффективность рынков

Одна из самых горячих дискуссий в области финансов — это спор о том являются ли финансовые рынки эффективными. То есть о том отражают ли цены активов всю информацию об этих активах. В этой задаче вам предлагается проанализировать различные подходы к тестированию гипотезы эффективности рынков.

В эконометрических терминах можно сказать, что если эта гипотеза верна, то доходность актива должна быть непредсказуема.

(а) (5) Shiller and Perron (1984) предложили «наивный» способ тестировать гипотезу эффективности рынка — variance ratio test.

Тестовая статистика, которую они предлагают использовать, имеет вид:

$$V(j) = \frac{\text{var}(r_t - r_{t-j})}{j * \text{var}(r_t - r_{t-1})}$$

где r_t — это логарифм доходности актива. Определите, чему должна быть равна эта величина, если гипотеза об эффективности рынков верна и поэтому динамика логарифма доходности актива описывается процессом случайного блуждания?

Какие несовершенства такого подхода к тестированию вы можете указать?

(б) (5) Принято считать, что отношения цены к дивидендам может предсказать доходности. Когда цена актива падает, то отношение $\frac{P_{t+1}+D_{t+1}}{P_t}$ растет (здесь P — цена актива, D — дивиденд). Чтобы проверить это, Джон Кокрейн оценивает так называемые «прогнозные регрессии»:

$$\sum_{j=1}^k r_{t+j} = \alpha + \beta_k(d_t - p_t) + \varepsilon_{t+k}$$
$$d_{t+k} - d_t = a + b_k(d_t - p_t) + u_{t+k}$$

где $d_t - p_t$ — логарифм отношения дивидендов к цене, $d_{t+k} - d_t$ — логарифмический темп прироста дивидендов за k периодов. β_k известен как прогнозный коэффициент. В такой постановке говорят, что доходности непредсказуемы, если прогнозный коэффициент равен нулю.

МНК-оценки Кокрейна представлены в следующей таблице:

Период	β	$t(\beta)$	b	$t(b)$
1	4.0	2.7	0.07	0.06
3	12.6	3.0	0.16	0.13
5	20.6	2.6	2.42	1.11

Предсказуемы ли доходности в соответствии с результатами Кокрейна? (Считайте, что объем выборки достаточен для использования асимптотического подхода к тестированию гипотез.) Какие несовершенства такого подхода к тестированию вы можете указать?

(в) (5) Чтобы избежать одного из недостатков прогнозной регрессии, Роберт Ходрик предложил использовать следующую модель из двух уравнений для логарифма однолетних доходностей и логарифма отношения цены к дивидендам:

$$r_t = \beta_0 + \beta_1(p_{t-1} - d_{t-1}) + \beta_2 r_{t-1} + \varepsilon_t^{(1)}$$

$$p_t - d_t = \theta_0 + \theta_1(p_{t-1} - d_{t-1}) + \theta_2 r_{t-1} + \varepsilon_t^{(2)}$$

Результаты оценивания представлены в таблице ниже (в скобках под оценками коэффициентов указаны стандартные ошибки).

	r_t	$p_t - d_t$
r_{t-1}	0.72 (0.04)	0.70 (0.19)
$p_{t-1} - d_{t-1}$	-0.01 (0.00)	0.90 (0.02)

R-значения теста Грейнджера на причинно-следственную связь для первого и второго уравнения равны 0.0001 и 0.0004 соответственно.

Предсказуемы ли доходности? Предсказуемы ли отношения цен к дивидендам? Что можно сказать о состоятельности оценок прогнозной регрессии из предыдущего пункта в свете этих новых результатов? Какие потенциальные недостатки модели Ходрика вы можете указать?

(г) (10) Один из вариантов развития подхода Ходрика состоит в том, чтобы добавить в модель прирост дивидендов Δd_{t+1} . Он может оказывать влияние на доходности и отношения цен к дивидендам. Джон Кэмпбелл и Боб Шиллер показали, что эти три переменные связаны друг с другом следующим образом:

$$r_{t+1} \approx \kappa_1(p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t).$$

Предложите способ оценить прогнозный коэффициент и проверить его значимость, используя все три ряда данных. Как бы вы посчитали изменение доходности в ответ на экзогенный шок каждой, из двух других переменных?

Задание 4. Служу Вестеросу!

В королевстве Вестерос военная служба осуществляется на контрактной основе: любой желающий при должном состоянии здоровья может подать заявку, отслужить в армии в течение трех лет, а затем вернуться к работе в гражданском секторе. Исследователь Джон Сноу решил выяснить, как служба в армии влияет на заработные платы граждан королевства?

Для этого он собрал данные о тысяче индивидов, часть из которых в прошлом отслужила в вооруженных силах Вестероса, а часть — нет. Ниже перечислены переменные, информация о которых доступна Джону:

lnWAGE — логарифм зарплаты индивида

EXP — опыт работы индивида в гражданском секторе (в годах)

sq_EXP — квадрат переменно EXP

ED — образование индивида (в годах)

FEMALE — бинарная переменная, равная единице для женщин и нулю для мужчин

SWORN_SWORD — бинарная переменная, равная единице для тех, кто отслужил в армии Вестероса, и равная нулю для всех остальных

DESCEDANT — бинарная переменная, равная единице для тех респондентов, у которых хотя бы один из родителей отслужил в армии, и равная нулю для всех остальных.

(a) (20) В приложении в конце варианта представлено 12 эконометрических уравнений, которые оценил исследователь. Проанализировав доступную информацию, ответьте на исследовательский вопрос Джона Сноу: *влияет ли служба в армии на зарплату граждан Вестероса? И если влияет, то в каком направлении и на сколько процентов служба в армии изменяет будущую зарплату индивида?*

Аргументируйте свой ответ: поясните, какие уравнения и каким образом вы использовали. Приведите необходимые содержательные соображения. Осуществите все необходимые тесты и интерпретируйте их результаты.

(б) (5) Вполне возможно, вы посчитаете нашего исследователя из этой задачи недостаточно квалифицированным¹, так как он не оценил каких-то моделей, которые могли бы уточнить ответ на вопрос о влиянии службы в армии на зарплату. В этом случае укажите, какие еще модели могли бы вам пригодиться и для чего.

Приложение к заданию 4. Таблицы с результатами оценивания уравнений

Таблица 1. Метод оценки — МНК. Зависимая переменная: lnWAGE

	(1)	(2)	(3)	(4)	(5)
const	10,710 (0,084)	10,206 (0,120)	6,043 (0,103)	6,039 (0,134)	6,053 (0,158)
SWORN_SWORD	0,154 (0,146)	-0,142 (0,143)	-0,159 (0,063)	-0,159 (0,063)	-0,156 (0,073)
ED		0,101 (0,017)	0,100 (0,007)	0,100 (0,011)	0,098 (0,017)
EXP			0,508 (0,009)	0,508 (0,009)	0,508 (0,009)
sq_EXP			-0,010 (0,001)	-0,010 (0,001)	-0,010 (0,001)
FEMALE				0,004 (0,083)	
DESCEDANT					-0,013 (0,152)
Количество наблюдений	1000	1000	1000	1000	1000
R ²	0,001	0,035	0,807	0,807	0,807

Примечание: В скобках указаны робастные стандартные ошибки.

¹ — Ничего ты не знаешь, Джон Сноу! (с)

Таблица 2. Метод оценки – МНК. Зависимая переменная: SWORN_SWORD

	(1)	(2)	(3)	(4)	(5)
const	0,249 (0,017)	0,337 (0,019)	0,331 (0,051)	-0,503 (0,047)	0,349 (0,057)
DESCEDANT	0,251 (0,032)			0,999 (0,039)	
FEMALE		-0,007 (0,023)			-0,017 (0,021)
ED			-0,001 (0,004)	0,100 (0,005)	-0,002 (0,005)
EXP			0,001 (0,004)	0,001 (0,004)	0,001 (0,004)
sq_EXP			-0,001 (0,001)	-0,001 (0,001)	-0,001 (0,001)
Количество наблюдений	1000	1000	1000	1000	1000
R ²	0,062	0,001	0,002	0,446	0,003

Примечание: В скобках указаны робастные стандартные ошибки.

Таблица 3. Метод оценки — двухшаговый МНК. Зависимая переменная: lnWAGE

	(1)	(2)
const	12,047 (0,112)	13,121 (1,557)
SWORN_SWORD	0,239 (0,023)	-0,392 (4,688)
ED	0,060 (0,007)	0,059 (0,007)
EXP	0,508 (0,009)	0,508 (0,009)
sq_EXP	-0,010 (0,000)	-0,010 (0,000)
Количество наблюдений	1000	1000
R ²	0,806	0,809
Р-значение теста Хаусмана	0,030	0,020

Примечание: В скобках указаны стандартные ошибки. В модели (1) в качестве инструмента для переменной SWORN_SWORD используется переменная DESCEDANT. В модели (2) в качестве инструмента для переменной SWORN_SWORD используется переменная FEMALE.