

Эконометрика — 2022

ЭКЗАМЕН

Пробный вариант

Декабрь 2022

Дорогие студенты! Экзамен — это 100% самостоятельная письменная проверочная работа, которая проводится в течение 80 минут в формате "*closed-book*". Вам не разрешается использовать какие-либо материалы или устройства, кроме персонального калькулятора, который не находится в телефоне или ином коммуникационном или носимом смарт-устройстве или т.п. Общение с кем-либо в или вне аудитории не разрешено. Если у вас есть вопрос, задайте его преподавателю или ассистенту.

Каждый вопрос задания имеет фиксированную стоимость в баллах, которая указана в квадратных скобках после номера вопроса. Например, "3. [8] ..." означает, что ответ на вопрос номер 3 может принести вам до 8 баллов. Отвечайте на вопросы аккуратно и максимально кратко, по сути. Примите к сведению и помните, что во всех вопросах правильный ответ является кратким и ясным. При этом запутанные ответы или ответы, содержащие не относящуюся к вопросу информацию, приведут к штрафам (будут оценены отрицательными баллами).

Оценка за работу определяется по сумме набранных баллов согласно шкале перевода:

Баллы	до 20	[20;30)	[30;40)	[40;50)	[50;60)	[60;70)	[70;75)	[75;80)	[80;85)	85 и более
Оценка	1	2	3	4	5	6	7	8	9	10

Распределяйте своё время и усилия разумно. Удачи!

Представьтесь

0. [1] Фамилия Имя Группа

Множественный выбор

В этой части нужно ответить на тестовые вопросы. В каждом вопросе есть ровно один правильный ответ. Выберите ответ, который считаете верным, и обведите соответствующую ему букву.

1. [1] В целом какие две задачи решает эконометрика (один ответ!)?

- а Оценка моделей и проверка гипотез
- б Поиск лучших моделей и прогнозирование
- в Совместное и условное распределение величин
- г Анализ взаимосвязей и прогнозирование будущего

2. [2] В модели регрессии на константу $x_i = \mu + u_i, \dots$ с достаточным набором предположений оценкой параметра μ методом аналогии является

- а $\hat{\mu}$ б $\bar{\mu}$ в \hat{x} г \bar{x}

3. [3] Какое предположение гарантирует популяционную идентификацию параметров в модели парной линейной регрессии $y_i = \beta_0 + \beta_1 x_i + u_i, \dots$?

- а $E[u_i | x_i] = 0$
- б $E[y_i | x_i] = \beta_0 + \beta_1 x_i$

в $\sigma^2(x) < \infty \forall x$

г $\text{Var}[x_i] \neq 0$

4. [2] Формула оценки угла наклона в парной линейной регрессии - это

а $\beta_1 = \text{Cov}(y, x) / \text{Var}(x)$

б $\hat{\beta}_1 = \widehat{\text{Cov}}(y, x) / \widehat{\text{Var}}(x)$

в $\hat{\beta}_1 = \beta_1 + \frac{\text{Cov}(u, x)}{\text{Var}(x)}$

г $\sqrt{n}(\hat{\beta}_1 - \beta_1) \rightarrow N(0, 1)$

5. [3] Получены оценки МНК $\hat{\beta}$ параметров ММЛР с двумя регрессорами (+ константа) и оценка матрицы ковариаций оценок параметров $\hat{V}_{\hat{\beta}}$. Что является выражением для стандартной ошибки $SE(\hat{\beta}_1 - \hat{\beta}_2)$?

а $\sqrt{\frac{[0 \ 1 \ -1]\hat{V}_{\hat{\beta}}[0 \ 1 \ -1]'}{n}}$

б $\sqrt{\frac{[0 \ 1 \ 0]\hat{V}_{\hat{\beta}}[0 \ 0 \ -1]'}{n}}$

в $\sqrt{\frac{[1 \ 1 \ 1]\hat{V}_{\hat{\beta}}[1 \ 1 \ 1]'}{n}}$

6. [2] Мы проверили композитную гипотезу $H_o : \{\beta_1 = \beta_2 = 0\}$ и отвергли ее. Какой вывод следует сделать?

а Хотя бы один из коэффициентов не равен нулю

б Оба коэффициента не равны нулю

в Оба коэффициента не равны нулю и не равны между собой

г Оба коэффициента больше нуля

Задача 1. Инструментальные переменные

В этой части приводится пример небольшого эмпирического исследования. Необходимо ответить на вопросы об исследовании, по возможности подкрепляя их расчетами на основе предоставленных данных (Таблица 1). Если расчеты занимают слишком много времени, постарайтесь описать хотя бы их ход и логику.

Исследователей заинтересовал вопрос о том, дают ли частные школы лучшее образование, чем государственные. Поскольку на католические школы приходится около двух третей учащихся частных школ в США, оценка эффективности католических школ является важной частью оценки частного школьного образования.

Для оценки исследователями были доступны данные о 7430 американских школьниках:

- *math12* — результаты стандартизированного теста по математике,
- *cathhs* — индикатор католической школы (=1, если школьник учится в католической школе, и 0 иначе),
- *lfaminc* — логарифм семейного дохода,
- *motheduc* — количество лет образования матери,
- *fatheduc* — количество лет образования отца,
- *parcath* — индикатор католичества (=1, если кто-то из родителей — католик, и 0 иначе).

при этом выборка достаточно сбалансирована по этим переменным, в ней нет пропусков и сильных выбросов.

Исследователи решили, что этих переменных достаточно для постановки и успешного применения двухшагового метода наименьших квадратов для ответа на вопрос **"Действительно ли учащиеся католических школ лучше сдают стандартизированные тесты по математике, чем учащиеся других школ (при прочих равных)?"**

Было предпринято несколько попыток, оценено несколько моделей (см. Таблицу 1 на стр. 8). В таблице представлены следующие результаты:

- (1) МНК-оценки для уравнения с *math12* в качестве зависимой переменной;
- (2) МНК-оценки для уравнения с *cathhs* в качестве зависимой переменной;
- (3) 2МНК-оценки с использованием *parcath* в качестве инструмента для *cathhs*.

Ответьте на следующие вопросы.

7. [4] Если бы мы использовали уравнение (1) из таблицы 1, каков был бы ответ на исследовательский вопрос? Проинтерпретируйте оценку коэффициента при *cathhs*.

8. [4] Поясните, в чем проблема оценки (1). Подходит ли она для качественного ответа на вопрос исследования?

9. [4] Судя по уравнению (3), что можно сказать о качестве оценки уравнения (1)?

10. [2] В чем смысл уравнения (2)?

11. [4] Является ли *parcath* сильным инструментом для *cathhs*? Объясните свой ответ.

12. [4] Является ли *parcath* годным инструментом для *cathhs*? Объясните свой ответ.

13. [6] Как конкретно были получены оценки коэффициентов в колонке (3)?

14. [4] В чем принципиальное отличие оценок в колонке (3) от оценок в колонке (2)?

15. [6] Если бы мы использовали уравнение (3), каков был бы ответ на исследовательский вопрос? Проинтерпретируйте оценку коэффициента при переменной интереса. Как эта оценка соотносится с истинным эффектом от учебы в католической школе?

Задача 2. Бинарный выбор

В этой части приводится пример небольшого эмпирического исследования. Необходимо ответить на вопросы об исследовании, по возможности подкрепляя их расчетами на основе предоставленных данных (Таблица 2). Если расчеты занимают слишком много времени, постарайтесь описать хотя бы их ход и логику.

Исследователи обратили внимание на вопрос о влиянии статуса здоровья пожилых американцев на покупку ими частной медицинской страховки. Был проведен сбор данных о 2748 индивидах:

- *ins* — наличие купленной частной медицинской страховки (=1, если страховка есть, и 0 иначе),
- *hstatusg* — индикатор статуса здоровья (=1, если человек оценивает свое здоровье как хорошее, и 0 иначе),
- *age* — возраст (в выборке индивиды от 52 до 88 лет),
- *retire* — индикатор того, вышел ли человек на пенсию (=1, если уже вышел, и 0 иначе),
- *hhincome* — доход домохозяйства,
- *educyear* — количество лет образования,
- *married* — индикатор семейного положения (=1, если человек женат/замужем, и 0 иначе).

при этом выборка достаточно сбалансирована по этим переменным, в ней нет пропусков и сильных выбросов.

Исследователи предполагают, что здоровые люди с меньшей вероятностью купят страховку. Для проверки они оценили несколько спецификаций с *ins* в качестве зависимой переменной (см. Таблицу 2 на стр. 8). Ответьте на следующие вопросы.

16. [2] Выпишите уравнение регрессии в терминах оцененных коэффициентов, соответствующее спецификации (1) из Таблицы 2

17. [2] Выпишите уравнение регрессии в терминах оцененных коэффициентов, соответствующее спецификации (4) из Таблицы 2

18. [4] В чем общая проблема моделей (1), (3)?

19. [4] В чем общая проблема моделей (1), (2)?

20. [4] Объясните роль переменных *age* и *retire* в спецификациях.

21. [6] Исходя из исследовательского вопроса, сформулируйте основную гипотезу исследования в содержательной, модельной и нормальной формах.

22. [4] Что можно предварительно сказать об ответе на основной вопрос и гипотезе, основываясь только на оценках коэффициентов в Таблице 2?

23. [4] Прокомментируйте оценку коэффициента при *hstatus* в спецификации (4).

24. [6] Опишите общую схему проверки гипотезы об эффекте статуса здоровья на вероятность покупки частной медицинской страховки.

25. [4] Приведите два *разных и конкретных* примера угроз *внутренней* валидности этого исследования.

26. [4] Приведите два *разных и конкретных* примера угроз *внешней* валидности этого исследования.

Таблица 1: Результаты оценивания для задачи 1

	(1)	(2)	(3)
(Intercept)	11.149*** (1.245)	-0.321*** (0.036)	11.915*** (1.303)
cathhs	1.477*** (0.388)		4.117** (1.458)
lfaminc	1.849*** (0.141)	0.018*** (0.004)	1.785*** (0.144)
motheduc	0.716*** (0.062)	0.004* (0.002)	0.713*** (0.062)
fatheduc	0.891*** (0.056)	0.007*** (0.002)	0.875*** (0.057)
parcath		0.143*** (0.007)	

В колонке (2) зависимой переменной является *cathhs*, а в остальных — *math12*.

В колонках (1), (2) представлены МНК-оценки, а в колонке (3) — 2МНК-оценки с использованием *parcath* в качестве инструмента.

В скобках даны робастные стандартные ошибки.

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Таблица 2: Результаты оценивания для задачи 2 (зависимая переменная — *ins*)

	(1)	(2)	(3)	(4)
(Intercept)	0.412* (0.173)	-0.388 (0.755)	0.136 (0.172)	-1.669* (0.796)
hstatusg	0.144*** (0.020)	0.635*** (0.092)	0.064** (0.021)	0.309** (0.098)
retire	0.066** (0.020)	0.284** (0.087)	0.039* (0.020)	0.187* (0.091)
age	-0.003 (0.003)	-0.011 (0.012)	-0.004 (0.003)	-0.019 (0.012)
hhincome			0.0005** (0.0002)	0.002** (0.0008)
educyear			0.027*** (0.003)	0.129*** (0.015)
married			0.134*** (0.021)	0.637*** (0.101)

В колонках (1), (3) представлены МНК-оценки коэффициентов для линейной вероятностной модели, в колонках (2), (4) — ММП-оценки коэффициентов для логит-модели.

В скобках даны стандартные ошибки.

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$