



III Студенческая Универсиада по эконометрике  
МГУ имени М. В. Ломоносова  
19 апреля 2014, Москва

**Задание I тура**

**Задача 1.**

При помощи МНК исследователь **А** оценил две регрессии:

$$\widehat{\ln Y}_i = 10,0 + 0,10 * MALE_i \quad , \quad R^2 = 0,16 \quad (1)$$

$$\widehat{\ln Y}_i = 7,0 + 0,05 * EDUCATION_i \quad , \quad R^2 = 0,25 \quad (2)$$

где  $Y_i$  — зарплата  $i$ -го работника (в **рублях**),  
 $EDUCATION_i$  — количество лет обучения  $i$ -го работника,  
 $MALE_i$  — фиктивная переменная, равная единице для мужчин и нулю для женщин.

Известно, что выборочный коэффициент корреляции между переменными  $MALE$  и  $EDUCATION$  равен 0,1, и ровно половину выборки составляют мужчины.

Исследователь **П**, используя в точности ту же самую выборку, планирует оценить регрессию вида:

$$\widehat{\ln Z}_i = \hat{\theta}_1 + \hat{\theta}_2 * EDUCATION_i + \hat{\theta}_3 * FEMALE_i \quad , \quad (3)$$

где  $Z_i$  — зарплата  $i$ -го работника (в **тысячах рублей**),  
 $FEMALE_i$  — фиктивная переменная, равная единице для женщин и нулю для мужчин.

Воспользуйтесь результатами расчетов исследователя **А** и помогите исследователю **П**

1. Вычислить МНК-оценку  $\hat{\theta}_3$  в регрессии (3).
2. Вычислить коэффициент детерминации в регрессии (3).

**Задача 2.**

Регрессия стоимости квадратного метра общей площади однокомнатной квартиры в ЗАО Москвы в январе 2009 г. дала следующий результат (см. таблицы 2.1, 2.2, 2.3)

Здесь PRICE — цена квартиры (\$1000); TOTSP — общая площадь квартиры ( $m^2$ ); KITSP — площадь кухни ( $m^2$ ); METRDIST - расстояние до метро в минутах; DIST - расстояние до центра (км); WALK = 1, если пешком до метро, = 0, если на транспорте; BRICK = 1, если дом

кирпичный или из монолитного железобетона, = 0 иначе; FLOOR1 – индикатор квартиры на 1-м этаже

(а) Оцените эластичность цены «средней» квартиры по расстоянию до центра.

(б) Оцените эластичность цены «средней» квартиры по жилой площади.

(в) Интерпретируйте коэффициенты при  $(1-WALK)*METRDIST$ ,  $WALK*METRDIST$ ,  $LOG(DIST)$ ,  $FLOOR1$ ,  $LOG(TOTSP)$ .

(г) При прочих факторах фиксированных модель показывает на «оптимальное» значение площади кухни  $KITSP_{opt}$ , при котором стоимость квадратного метра квартиры максимальна. Найдите 95%-й доверительный интервал для  $KITSP_{opt}$ .

Таблица 2.1. Deskриптивные статистики (646 наблюдений)

	Mean	Median	Max	Min	Std.Dev.
PRICE	212.7631	195.8405	700	120	73.02876
PRICE/TOTSP	5.458703	5.227316	13.33832	2.583383	1.338476
KITSP	8.863622	9	20	5	2.273847
KITSP^2	83.72618	81	400	25	45.58565
LOG(TOTSP)	3.643188	3.637586	4.379524	3.113515	0.177144
TOTSP	38.85433	38	79.8	22.5	7.623778
LIVESP	19.59319	19.1	45	10	3.028242
LOG(DIST)	2.32146	2.397895	2.833213	1.252763	0.330918
DIST	10.65944	11	17	3.5	2.702496
FLOOR1	0.06192	0	1	0	0.241196
WALK	0.512384	1	1	0	0.500234
METRDIST	8.843653	10	20	1	4.063681
BRICK	0.312693	0	1	0	0.46395

Таблица 2 2. Результаты оценивания регрессии

Dependent Variable: PRICE/TOTSP

Method: Least Squares

Included observations: 646

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	9.774787	1.479911	6.604981	0.0000
KITSP	0.372701	0.105070	3.547174	0.0004
KITSP^2	-0.013295	0.005087	-2.613413	0.0092
LOG(TOTSP)	-0.902797	0.449671	-2.007684	0.0451
LOG(DIST)	-1.352729	0.138383	-9.775235	0.0000
FLOOR1	-0.412041	0.184933	-2.228054	0.0262
WALK	0.583089	0.221292	2.634933	0.0086
WALK*METRDIST	-0.037876	0.014887	-2.544260	0.0112
(1-WALK)*METRDIST	-0.058256	0.017004	-3.426050	0.0007
BRICK	0.259290	0.106271	2.439885	0.0150
R-squared	0.314272	Mean dependent var		5.458703
Adjusted R-squared	0.304568	S.D. dependent var		1.338476
S.E. of regression	1.116190	Akaike info criterion		3.073077
Sum squared resid	792.3792	Schwarz criterion		3.142285
Log likelihood	-982.6040	Durbin-Watson stat		1.534229
F-statistic	32.38679			
Prob(F-statistic)	0.000000			

Таблица 2.3. (начало). Ковариационная матрица оценок коэффициентов регрессии

	C	KITSP	KITSP^2	LOG(TOTSP)	LOG(DIST)
C	2.190138	-0.013031	0.002435	-0.604292	-0.027296
KITSP	-0.013031	0.011040	-0.000505	-0.012937	0.001007
KITSP^2	0.002435	-0.000505	2.59E-05	1.47E-05	-3.69E-05
LOG(TOTSP)	-0.604292	-0.012937	1.47E-05	0.202204	-0.005400
LOG(DIST)	-0.027296	0.001007	-3.69E-05	-0.005400	0.019150
FLOOR1	-0.029341	0.000464	-2.76E-05	0.007795	-0.000921
WALK	-0.004770	0.000752	-7.11E-06	-0.006424	-0.003683
WALK*METRDIST	-0.000713	3.33E-05	-2.83E-06	0.000233	-9.06E-05
(1-WALK)*METRDIST	-0.001280	0.000148	-5.77E-06	-0.000398	-0.000401
BRICK	0.005052	0.003233	-0.000154	-0.008548	0.002863

Таблица 2.3 (продолжение). Ковариационная матрица оценок коэффициентов регрессии

	FLOOR1	WALK	WALK*METRDIST	(1-WALK)*METRDIST	BRICK
C	-0.029341	-0.004770	-0.000713	-0.001280	0.005052
KITSP	0.000464	0.000752	3.33E-05	0.000148	0.003233
KITSP^2	-2.76E-05	-7.11E-06	-2.83E-06	-5.77E-06	-0.000154
LOG(TOTSP)	0.007795	-0.006424	0.000233	-0.000398	-0.008548
LOG(DIST)	-0.000921	-0.003683	-9.06E-05	-0.000401	0.002863
FLOOR1	0.034200	-0.001738	0.000145	-0.000105	-0.000150
WALK	-0.001738	0.048970	-0.001764	0.002775	-0.000640
WALK*METRDIST	0.000145	-0.001764	0.000222	1.07E-06	-2.06E-05
(1-WALK)*METRDIST	-0.000105	0.002775	1.07E-06	0.000289	0.000130
BRICK	-0.000150	-0.000640	-2.06E-05	0.000130	0.011294

### Задача 3.

Преподавателя интересует причинно-следственная связь между посещением лекций и сдачей зачета по математическому анализу. Преподаватель разделил доступную ему выборку студентов на две группы (студенты, посетившие 90% лекций или более, и все остальные студенты). Кроме того, преподаватель знает, что часть студентов живут в корпусе общежития совсем рядом с университетом, а часть — в другом корпусе, от которого до университета нужно 45 минут добираться на общественном транспорте (после поступления в университете студенты были расселены по корпусам случайным образом).

Все доступные преподавателю данные представлены в таблице 3.1.

**Таблица 3.1. Результаты студентов в зависимости от посещения лекций и удаленности общежития от университета**

	Студенты, которые живут рядом с университетом	Студенты, которые живут далеко от университета
Студенты, посетившие 90% и более лекций	Количество студентов — 90 Средний балл за экзамен — 80	Количество студентов — 60 Средний балл за экзамен — 90
Студенты, посетившие меньше 90% лекций	Количество студентов — 70 Средний балл за экзамен — 30	Количество студентов — 100 Средний балл за экзамен — 40

1. Объясните, почему регрессия экзаменационного балла по посещаемости скорее всего не позволит исследователю при помощи обычного МНК состоятельно оценить интересующий его эффект.
2. Объясните, почему удаленность от университета может быть хорошей инструментальной переменной для посещения лекций.
3. Оцените интересующий преподавателя эффект, вычислив 2МНК-оценку коэффициента  $\theta_2$  в регрессии

$$exam_i = \theta_1 + \theta_2 * d_i + \varepsilon_i ,$$

где  $exam_i$  — средний балл  $i$ -го студента за экзамен,  $d_i$  — фиктивная переменная, равная единице для студентов, посетивших 90% и более лекций.

#### **Задача 4.**

Известно, что динамика временного ряда  $x_t$  описывается стационарным процессом авторегрессии второго порядка. Известно, что безусловное математическое ожидание для этого процесса равно 10, а коэффициенты автокорреляции первого и второго порядка равны, соответственно, 0,5 и 0,4.

1. Восстановите уравнение, описывающее динамику временного ряда  $x_t$ .
2. Известно, что  $x_1 = 8$  и  $x_2 = 6$ . Постройте прогноз  $\hat{x}_4$ .