

## Универсиада по эконометрике 2025, 1 тур

### Задание 1 (7 баллов)

Цифровые экосистемы разрастаются за счёт собственных сервисов и за счёт покупки стартапов...

#### Примерное решение и баллы задачи 1

Показан и прокомментирован один из способов решения.

Пропуск А – 1 балл.

Ищется из формулы t-расчётного для теста на значимость коэффициента. Дана оценка, дано t-расчётное. Значит, стандартная ошибка оценки коэффициента при переменной X в первом уравнении = 0.4

Пропуск Б – 3 балла

Можно использовать тот факт, что квадрат t-расчётного = F расчётному = 4, т.к проверяется одна и та же гипотеза о равенстве нулю коэффициента при X.

То есть эту же гипотезу можно рассматривать как проведение теста на значимость уравнения в целом, и тогда:

$$F_{\text{расчет}} = \frac{R^2/1}{(1 - R^2)/(n - k)}$$
$$4 = \frac{R^2/1}{(1 - R^2)/(38 - 2)}$$

Отсюда  $R^2 = 0.1$

При этом способе за недоделанные мысли в правильном направлении 1 балл, за полностью правильное решение 3 балла, а если есть ошибка в расчётах, то 2 балла.

Пропуск В – 3 балла

Можно сделать, например, через F-тест «короткая против длинной»: две модели отличаются друг от друга наличием переменной W:

$$F_{\text{расчет}} = \frac{(R_{\text{длин}}^2 - R_{\text{корот}}^2)/1}{(1 - R_{\text{длин}}^2)/(n - k_{\text{длин}})}$$
$$F_{\text{расчет}} = \frac{(0.475 - 0.1)/1}{(1 - 0.475)/(38 - 3)} = 25$$

Отсюда t-расчётное для гипотезы о равенстве нулю коэффициента при переменной W равно либо 5, либо -5. Поэтому оценка коэффициента при W либо 0.75, либо -0.75.

При этом способе за недоделанные мысли в правильном направлении 1 балл

За полностью правильное решение 3 балла.

Если есть ошибка, которая возникла из-за АРИФМЕТИЧЕСКОЙ (а не содержательной) ошибки на шаге Б, то тоже 3 балла. Если из-за содержательной, то 2 балла.

Если есть ошибка в расчётах, то 2 балла

Если при взятии корня забывают, что есть плюс и минус, то 2 балла из 3.

## Задание 2 (8 баллов)

Стас занимается организацией Международной молодёжной конференции....

### Примерное решение и баллы за пункт А (2 балла)

Ответ: Стас не смог получить никакие численные оценки, потому что в модели совершенная мультиколлинеарность: есть 4 категории и 4 бинарные переменные и константа. 0.5 балл за этот правильный ответ. Он может звучать как “ловушка бинарных переменных”, или слова про линейную зависимость, или про вырожденную матрицу регрессоров – любые верные по смыслу.

И 1.5 балла за комментарии к остальным вариантам, по 0.5 за каждый:

“Стас получил смещённые оценки” - такая ситуация был бы возможна при пропуске существенных переменных. (Тут могут ещё что-нибудь написать, например, что если бы не было в модели 1 совершенной мультиколлинеарности, то там явно ещё и пропущены важные переменные, и оценки были бы смещёнными).

“Стас получил несмещённые эффективные оценки” – если бы выполнялись все условия теоремы Гаусса-Маркова

“Стас получил неэффективные оценки” – если бы в модели были все существенные переменные, но была бы, например, гетероскедастичность или добавлены несущественные переменные, или несовершенная мультиколлинеарность.

При этом обязательно указание, какой именно вариант – верный.

### Решение и баллы за пункт Б (2 балла)

Здесь принимались три варианта.

Первый вариант: просто тест на значимость уравнения в целом, потому что во всех произведениях есть переменная d.

Нулевая гипотеза:  $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ , альтернативная, что хотя бы одна из этих бет не равна нулю

Вариант 2:

ИЛИ нулевая гипотеза, что все беты равны друг другу (без равенства нулю), а альтернативная: хотя бы одна бета не равна другой

За гипотезы 0.5 балла

Для первого варианта

$$F_{\text{расчет}} = \frac{0,6/1}{(1 - 0,6)/(1005 - 5)} = 375$$

1 балл за расчётную F статистику и 0.5 балла за сравнение с табличной и вывод.

F табличная, например, для 5% уровня значимости = 2.38

Нулевая гипотеза отвергается, наличие новизны по-разному влияет на принятие тезисов в разных подсекциях.

Во втором случае:

Нельзя посчитать так как не хватает информации для проведения F-теста (не дан коэффициент детерминации для регрессии с ограничением) – за это 0,5 балла

Есть формула расчетной статистики, куда подставлено все что доступно, 0,5 балла  
Есть описание процедуры тестирования (рассуждение о сравнении критической и расчетной) еще 0,5 балла

Третий вариант: аргументированные рассуждения, что нельзя проверить, так как случайные ошибки не распределены нормально и тогда не подходят обычные статистики - если есть гипотеза то 2, если нет гипотезы, то 1,5 балла.

### **Решение и баллы за пункт В (2 балла)**

По 1 баллу за каждую из гипотез

Первая: принимались два варианта:  $\beta_1 - \beta_2 = 0$  либо  $\beta_1 = \beta_2$

Вторая:  $\beta_3 = 0$ .

### **Решение и баллы за пункт Г (2 балла)**

2 балла за любой верный ответ. Проблема хотя бы в том, что вероятность от 0 до 1, а область значения у линейной модели от минус бесконечности до плюс бесконечности, и чисто теоретически может получиться, что прогноз по линейной модели вероятности выходит за границы от 0 до 1. И проблема с гетероскедастичностью.

В качестве совета: модели бинарного выбора (логит, пробит, деревья решений, случайный лес и другие алгоритмы классификации).

### **Задание 3 (5 баллов)**

Для линейной регрессии приведите пример оценки, отличной от МНК, которая является несмещённой и линейной по  $y$ , и покажите, что Ваша оценка является несмещённой и линейной по  $y$ .

Вы можете привести пример для парной регрессии или для множественной регрессии в матричном виде.

### **Решение и баллы задачи 3**

Достаточно одного примера для парной либо множественной регрессии. Полные 5 баллов ставятся только в том случае, если: приведена оценка, соответствующая критериям из условия, а также строго по определению доказана линейность по  $y$  (либо оценка представлена в виде  $Ay$ ), и вычислено матожидание оценки, чтобы продемонстрировать её несмещённость. Все обозначения пояснены.

Если просто приведена оценка в ответе и ничего не доказано, то это 1 балл из 5.

ОМНК-оценка – да, принимается, с обоснованием

Тут были разночтения в условии. Если человек находил оценку другим способом, но в итоге она выглядит как обычная МНК, то так тоже можно.

#### Задание 4 (10 баллов)

Вам необходимо разработать дизайн эконометрического исследования.

Лена оценивает, как доходы бюджета стран-экспортёров нефти реагируют на изменение цен на нефть...

#### План решения и баллы задачи 4

За запись моделей «в общем виде» для тестирования гипотез, с пояснениями (расшифровкой) показателей – 2 балла.

Например:

Уравнение регрессии для проверки первой гипотезы

$$Rev_{it} = \beta_0 + \beta_1 Shock_t + \beta_2 Dev_i + \beta_3 Shock_t * Dev_i + u_{it}$$

где:

$Rev_{it}$  – доходы бюджета страны  $i$  в периоде  $t$  (в % ВВП)

$Shock_t$  – нефтяной шок в периоде  $t$  (ниже пояснения для его расчёта)

$Dev_i$  – показатель развитости страны  $i$  -- бинарная переменная, равная 1 для развитых стран и 0 для развивающихся стран

Возможны и другие варианты решения

$u_{it}$  – случайная ошибка для страны  $i$  в периоде  $t$

Это только один из множества вариантов, возможно добавление контрольных переменных и т.д. Принципиально важно, чтобы по выписанному уравнению можно было проверить гипотезу 1. Например, если бы в выписанном выше уравнении не было бы произведения, то это было бы невозможно сделать. Возможно решение, что развитость задаётся не бинарной переменной, а, например, через Индекс человеческого развития.

Уравнение регрессии для проверки второй гипотезы (один из вариантов, самый простой)

$$volat_i = a_0 + a_1 OilShare_i + v_i$$

где:

$volat_i$  – волатильность доходов бюджета страны  $i$  за весь рассматриваемый период, рассчитывается как стандартное отклонение  $Rev_{it}$  (таким образом, сделан переход от панельных к пространственным данным),

$OilShare_i$  – доля нефти в экспорте страны  $i$  средняя за весь рассматриваемый период

$v_i$  – случайная ошибка для страны  $i$

В уравнениях важна также непротиворечивость по индексам и адекватная зависимая переменная.

Проверка гипотез

Гипотеза 1 (2 балла):

Нулевая гипотеза:  $\beta_3 = 0$

Альтернативная гипотеза:  $\beta_3 < 0$

Проверяется односторонним тестом на значимость коэффициента.

При этом важно, чтобы  $\beta_1$  не было бы значимым отрицательным (скорее всего, получится значимым положительным). Тогда реакция доходов бюджета на нефтяной шок в развивающихся странах равна  $\beta_1$ , а в развитых  $\beta_1 + \beta_3 < \beta_1$ .

Гипотеза 2 (1 балл):

Нулевая гипотеза:  $a_1 = 0$

Альтернативная гипотеза:  $a_1 > 0$ .

Проверяется односторонним тестом на значимость коэффициента.

1 балл – за обоснование выборки стран и лет. Просто «5-10-20-50 стран развитых и развивающихся за последние 20-25 лет» не засчитывается.

За расшифровку, что такое нефтяной шок, 1 балл. Расшифровка не должна противоречить обозначениям из уравнения регрессии. Например, выше указано, что  $Shock_t$  – нефтяной шок в периоде  $t$ . В таком случае используем мировую цену на нефть, и шок одинаков для всех стран в один и тот же период. Рассчитывать размер шока можно, например, как максимум из 0 и превышения текущей цены над максимальной за последние 3 года.

Как вариант расшифровки, например, любой способ из статьи Гамильтона, которая приведена в условии, например, с разделением положительных и отрицательных шоков (из-за чего уравнение 1 будет выглядеть иначе, и гипотеза 1 тоже). Либо любой другой расписанный непротиворечивый способ.

Можно указать в уравнении  $Shock_{it}$  с двумя индексами, и тогда, например, привести рассуждения о нефти разных сортов, экспортируемой разными странами, и использовании цен соответствующего сорта. Либо переход к контрактным ценам. Например, с сайта WITS с 2000 года возможно рассчитать цену для каждой страны  $i$  в году  $t$  как отношение экспорта нефти (товарная позиция ТН ВЭД 2709) в стоимостном выражении к физическому объёму экспорта.

Просто ссылка на Гамильтона без пояснений расчётов – 0 баллов.

1 балл за расшифровку прочих показателей.

1 балл за обоснование с использованием литературы со ссылками. Балл не ставится, например, если литература указана просто как «отчёт МВФ 2019 года» без ссылки, либо написаны утверждения, которых нет в цитируемой статье, либо литература просто перечислена в ответе, но неясно из написанного, как именно ею пользовались.

1 балл за верные и аргументированные методы и тесты.

За дополнительные гипотезы, если они полностью обоснованы и формализованы, до 1 балла, но сумма за всю задачу не превышает 10 баллов.