Универсиада по эконометрике 2023
Решения задач второго тура

Задача 1. (25 баллов)
О градообразующих предприятиях

(1)

В терминах оцененной пробит-модели предсказанная вероятность создания территории опережающего развития (ТОРа) имеет вид[1]

$$\hat{\mathbb{P}}(y = 1 \mid X) = \Phi\left(0.360 + 0.187X - 0.17X^2\right)$$

Поскольку функция $\Phi(x)$ очевидным образом монотонна, она достигает своего максимума при максимальном значении аргумента (2 балла), иными словами нас интересует решение задачи

$$0.360 + 0.187X - 0.17X^2 \to \max_X \implies X_{\max} = -\frac{0.187}{2 \times (-0.17)} = \frac{11}{20} = 0.55 \qquad \text{(2 балла)}$$

(2)

В общем виде рассуждение выше можно представить так

$$\hat{\mathbb{P}}(y = 1 \mid X) = \Phi\left(\hat{\beta}_0 + \hat{\beta}_1 X + \hat{\beta}_2 X^2\right) \implies X_{\max} = -\frac{\hat{\beta}_1}{2\,\hat{\beta}_2}$$

Получается, равенство $-\frac{\hat{\beta}_1}{2\,\hat{\beta}_2} = 0.6$ эквивалентно равенству

$$1.2\,\hat{\beta}_2 + \hat{\beta}_1 = 0$$

[1] Далее $\Phi(x)$ означает функцию распределения для стандартного нормального распределения:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-x^2/2}\,dx$$

Теперь можно воспользоваться стандартным t-тестом (тест на одно линейное ограничение), для которого

$$\begin{array}{l} H_0 \;:\; 1.2\,\beta_2 + \beta_1 = 0 \\ H_1 \;:\; 1.2\,\beta_2 + \beta_1 \neq 0 \end{array} \qquad \text{(1 балл)}$$

$$t_{\text{расч}} = \frac{1.2\,\hat{\beta}_2 + \hat{\beta}_1}{\hat{\text{se}}\left(1.2\,\hat{\beta}_2 + \hat{\beta}_1\right)} = \frac{1.2\,\hat{\beta}_2 + \hat{\beta}_1}{\sqrt{1.44\,\hat{\mathbb{V}}\hat{\beta}_2 + 2.4\,\hat{\text{cov}}\left(\hat{\beta}_2, \hat{\beta}_1\right) + \hat{\mathbb{V}}\hat{\beta}_1}} \qquad \text{(2 балла)}$$

Соответственно, нулевая гипотеза будет отвергаться, если

$$|t_{\text{расч}}| > t^{\alpha}_{\text{крит}}(n-k) \qquad \text{(1 балл)}$$

Выше $t^{\alpha}_{\text{крит}}(n-k)$ – соответствующее критическое значение t-распределения с уровнем значимости α и $n-k$ степенями свободы, где n – число наблюдений, а k – число коэффициентов в модели.

Вычислим расчетную статистику, пользуясь данной в задаче ковариационной матрицей оценок коэффициентов и результатом пункта 1:

$$t_{\text{расч}} = \frac{-1.2 \times 0.17 + 0.187}{\sqrt{1.44 \times 0.051 + 2.4 \times 0.024 + 0.075}} = -\frac{0.017}{\sqrt{0.206}} \approx -0.037 \qquad \text{(2 балла)}$$

Поскольку $n - k = 200 - 3 = 197$ достаточно велико, можно использовать критические значения нормального распределения вместо t-распределения. Поскольку при этом $|t_{\text{расч}}| < 1.69$, то даже на 10% уровне значимости нельзя отвергнуть нулевую гипотезу о том, что $X_{\max} = 0.6$ (1 балл).

3

Пусть угловой коэффициент при $X \leq 0.6$ составляет β_1^L, а при $X > 0.6$ – β_1^R. Тогда нетрудно убедиться в том, что нужная спецификация пробит-модели имеет вид[2]

$$\mathbb{P}(y = 1 \mid X) = \Phi\left(\beta_0 + \beta_1^L (X - 0.6)\,[X \leq 0.6] + \beta_1^R (X - 0.6)\,[X > 0.6]\right) \qquad (1.3.1)$$

[2] Под выражением $[X > \vartheta]$ подразумевается дамми-переменная события $X > \vartheta$, то есть

$$[X > \vartheta] = \begin{cases} 1, & \text{если} \quad X > \vartheta \\ 0, & \text{если} \quad X \leq \vartheta \end{cases}$$

Как легко видеть, спецификация (1.3.1) выше эквивалентна следующей

$$\mathbb{P}(y = 1 \mid X) = \Phi\left(\gamma_0 + \gamma_1 X + \gamma_2 (X - 0.6)\,[X > 0.6]\right) \tag{1.3.2}$$

При этом связь между $\boldsymbol{\beta}$ и $\boldsymbol{\gamma}$ такова

$$\gamma_0 = \beta_0 - 0.6\,\beta_1^L \qquad \gamma_1 = \beta_1^L \qquad \gamma_2 = \beta_1^R - \beta_1^L$$

В этом пункте 1 балл снимался за типичную ошибку, при которой участник забывал, что имеет дело с пробит-моделью (забывал применить $\Phi(\cdot)$ к выражению с дамми-переменными). Далее 1 балл давался за введение переменной для сдвига и еще 1 балл – для наклона. Наконец, оставшиеся 2 балла ставились за то, что предложенная спецификация непрерывна по X.

4

Предложенная модель в явном виде предполагает излом в точке $X = 0.6$, поэтому для ответа участника достаточно написать, что эта точка окажется точкой максимума тогда и только тогда, когда верны неравенства (в терминах уравнений (1.3.1) и (1.3.2) соответственно)

$$\begin{cases} \beta_1^L > 0 \\ \beta_1^R < 0 \end{cases} \qquad \text{или} \qquad \begin{cases} \gamma_1 > 0 \\ \gamma_1 + \gamma_2 < 0 \end{cases} \qquad \text{(4 балла)}$$

И соответствующие коэффициенты статистически значимы. Для полного балла не требовалось предъявить строгий тест такой гипотезы, однако интересующийся читатель может обратиться к статье [Wolak, 1989], в которой такой тест предложен.

Предложения по проверке равенства $X_{\max} = 0.6$ или $\beta_1^L \neq \beta_1^R$ не оценивались.

5

Обозначим $X_{\max} = \vartheta$. Поскольку ϑ неизвестна и не оценивается внутри модели (является так называемым *гипер*параметром), то ее можно найти, например, так:

1. Рассмотреть некоторую решетку $\Theta = \{\vartheta_1, \vartheta_2, \ldots, \vartheta_k\}$
2. Для каждого $\vartheta \in \Theta$ (то есть для каждого $m = 1, 2, \ldots, k$) оценить модель, аналогичную модели из пункта 3
3. Из значений $\vartheta \in \Theta$ выбрать то, которое соответствует оптимальной модели

Оптимальность модели следует понимать в терминах некоторого заранее зафиксированного критерия качества. Например, в качестве такого критерия можно выбрать информационный критерий Шварца (и, соответственно, предпочтение отдается модели с его наименьшим значением). Альтернативой лобовому (или случайному) перебору по решетке Θ может быть, например, более продвинутая байесовская оптимизация или иные методы подбора гиперпараметров.

Более подробно интересующийся читатель может ознакомиться с материалами о пороговых регрессиях Брюса Хансена.

Идея перебора параметров по сетке оценивалась в 3 балла. При этом для получения оставшихся 2 баллов требовалось обоснование

1. Того, что предложенный алгоритм выберет **единственный** порог[3]
2. Того, что такой выбор единственного ϑ рационален (иными словами, максимизирует какой-то критерий качества)

[3] В этом смысле недостаточно выбирать такие ϑ, чтобы они проходили тест из пункта 4, потому что в таком случае «оптимальный» ϑ не будет единственным.

Задача 2. (20 баллов)
О пропавших оценках

Ⓐ

В результатах Маши нет противоречий, поскольку, вообще говоря, среднее арифметическое из коэффициентов парных регрессий, оцененных по отдельным подвыборкам, не совпадает с коэффициентом по всей выборке. Здесь можно сослаться на парадокс Симпсона, на наличие «фиксированных эффектов» стран. (2 балла)

Ⓑ

Сразу разумно оговориться о том, что в модели (5) очевидным образом присутствует чистая мультиколлинеарность, поэтому оценки, соответствующие буквам Т, Ф, Ц, Ч, Ш, Ы, Э и Ю, невозможно вычислить (3 балла).

Теперь обозначим сумму квадратов остатков и коэффициент детерминации j-ой модели как $\text{ESS}_{(j)}$ и $R^2_{(j)}$ соответственно. Тогда для $R^2_{(1)}$ имеет место формула[4]

$$R^2_{(1)} = 1 - \frac{\text{ESS}_{(1)}}{\text{TSS}_{(1)}} = 1 - \frac{200}{\text{TSS}_{(1)}} = 0.6 \iff \text{TSS}_{(1)} = 500$$

Поскольку во всех моделях (1)–(5) использована одна и та же зависимая переменная, то $\text{TSS}_{(j)}$ совпадают для всех j и равны 500. В таком случае, можно вычислить

$$R^2_{(2)} = 1 - \frac{\text{ESS}_2}{\text{TSS}_2} = 1 - \frac{220}{500} = 0.56 \implies \boxed{\text{А} = 0.56} \qquad \text{(1 балл)}$$

Далее, как нетрудно заметить, $\text{X} + \text{IMP} = \text{Sum}$ и $\text{X} - \text{IMP} = \text{Nx}$, поэтому модели (1), (3) и (4) при оценивании с помощью МНК будут проектировать y на одно и то же линейное многообразие (в данном конкретном случае – линейную оболочку X и IMP). Это означает, что для этих моделей неминуемо совпадут ESS и коэффициенты детерминации, поэтому

$\boxed{\text{Ж} = 200}$	(1 балл)	$\boxed{\text{С} = 200}$	(1 балл)
$\boxed{\text{Е} = 0.6}$	(1 балл)	$\boxed{\text{П} = 0.6}$	(1 балл)

[4] Формально в задаче явно не указано, присутствует ли в уравнении регрессия константа. Далее R^2 вычисляются, исходя из формул для регрессии **с константой**, то есть $R^2 = 1 - \frac{\text{ESS}}{\text{TSS}}$. На остальные выводы и значения неизвестных параметров спецификация никак не влияет.

Более того, из-за этого совпадут и предсказанные значения зависимой переменной, то есть три следующих вектора совпадают

$$\hat{y}_{(1)} = \hat{y}_{(3)} = \hat{y}_{(4)}$$

Теперь с учетом этого сравним оцененные уравнения моделей (1) и (3)

$$\underbrace{0.6\,\mathrm{X} + 0.8\,\mathrm{IMP} = \hat{y}_{(1)}}_{\text{Модель (1)}} = \underbrace{\hat{y}_{(3)} = \text{Б} \times \mathrm{Sum} + \Gamma \times \mathrm{Nx}}_{\text{Модель (3)}}$$
$$\Longleftrightarrow \; 0.6\,\mathrm{X} + 0.8\,\mathrm{IMP} = \text{Б}\,(\mathrm{X} + \mathrm{IMP}) + \Gamma\,(\mathrm{X} - \mathrm{IMP})$$

Отмечается, что равенство выше выполняется покоординатно. Раскрывая скобки и группируя слагаемые в правой части последнего уравнения, получаем систему

$$\begin{cases} \text{Б} + \Gamma = 0.6 \\ \text{Б} - \Gamma = 0.8 \end{cases} \Longleftrightarrow \boxed{\begin{array}{l} \text{Б} = 0.7 \\ \Gamma = -0.1 \end{array}} \qquad \text{(2 балла)}$$

Аналогично можно сравнить модели (1) и (4), откуда получается

$$\underbrace{0.6\,\mathrm{X} + 0.8\,\mathrm{IMP} = \hat{y}_{(1)}}_{\text{Модель (1)}} = \underbrace{\hat{y}_{(4)} = \text{И} \times \mathrm{IMP} + \text{Л} \times \mathrm{Sum}}_{\text{Модель (4)}}$$
$$\Longleftrightarrow \; 0.6\,\mathrm{X} + 0.8\,\mathrm{IMP} = \text{И} \times \mathrm{IMP} + \text{Б}\,(\mathrm{X} + \mathrm{IMP})$$
$$\begin{cases} \text{Л} = 0.6 \\ \text{И} + \text{Л} = 0.8 \end{cases} \Longleftrightarrow \boxed{\begin{array}{l} \text{Л} = 0.6 \\ \text{И} = 0.2 \end{array}} \qquad \text{(2 балла)}$$

Наконец, что касается стандартных ошибок, проще всего дело обстоит с Д и К. Как известно, вложенные модели, отличающиеся только одним регрессором, можно сравнить как с помощью F-теста, так и с помощью t-теста, причем соответствующие расчетные статистики связаны соотношением

$$F_{\text{расч}} = t^2_{\text{расч}} \tag{2.Б}$$

Можно заметить, что модели (2) и (3) вложенные: модель (2) является частным случаем модели (3). Тогда гипотезу $H_0 \,:\, \Gamma = 0$ можно проверить двумя тестами, для которых равенство (2.Б) будет иметь вид[5]

$$\frac{\left(R^2_{(3)} - R^2_{(2)}\right)/1}{\left(1 - R^2_{(3)}\right)/(43-2)} = \frac{0.04 \times 41}{0.4} = F_{\text{расч}} = t^2_{\text{расч}} = \left(\frac{\text{Г}}{\text{Д}}\right)^2 = \frac{0.01}{\text{Д}^2}$$

$$\Longrightarrow \boxed{\text{Д} = \sqrt{\frac{0.01}{4.1}} \approx 0.049} \qquad \text{(2 балла)}$$

Аналогичным образом модели (4) и (2) вложенные, поэтому гипотезу $H_0 : \text{И} = 0$ можно вновь проверить с помощью двух тестов

$$\frac{\left(R^2_{(4)} - R^2_{(2)}\right)/1}{\left(1 - R^2_{(4)}\right)/(43-2)} = 4.1 = F_{\text{расч}} = t^2_{\text{расч}} = \left(\frac{\text{И}}{\text{К}}\right)^2 = \frac{0.04}{\text{К}^2}$$

$$\Longrightarrow \boxed{\text{К} = \sqrt{\frac{0.04}{4.1}} \approx 0.099} \qquad \text{(1 балл)}$$

Несколько сложнее обстоит ситуация с В и М. На самом деле, их можно вычислить, что можно показать на примере В. Пусть ковариационная матрица оценок в модели (1) имеет вид[6]

$$\hat{\mathbb{V}}\hat{\boldsymbol{\beta}}_{(1)} = \begin{pmatrix} 0.04^2 & c \\ c & 0.04^2 \end{pmatrix}$$

Здесь c – неизвестная ковариация между оценками в первой модели. При вычислении Б и Г было показано, что

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \hat{\boldsymbol{\beta}}_{(2)} = \hat{\boldsymbol{\beta}}_{(1)} \iff \hat{\boldsymbol{\beta}}_{(2)} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \hat{\boldsymbol{\beta}}_{(1)}$$

Как известно, при линейном преобразовании ковариационная матрица меняется так же, как соответствующая квадратичная форма, поэтому[7]

$$\hat{\mathbb{V}}\hat{\boldsymbol{\beta}}_{(2)} = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \underbrace{\begin{pmatrix} 0.04^2 & c \\ c & 0.04^2 \end{pmatrix}}_{=\hat{\mathbb{V}}\hat{\boldsymbol{\beta}}_{(1)}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

[5] Здесь подразумевается, что регрессии оцениваются без константы, поэтому модель (3) включает 2 коэффициента, а модель (2) – только один.

[6] Здесь и далее $\hat{\boldsymbol{\beta}}_{(j)}$ означает вектор оценок из модели (j).

[7] А именно верно равенство $\mathbb{V}\left(A^\top \boldsymbol{\xi}\right) = A^\top \, \mathbb{V}(\boldsymbol{\xi})\, A$.

С другой стороны, в матрице $\hat{\mathbb{V}}\hat{\boldsymbol{\beta}}_{(2)}$ есть один заведомо известный элемент – Д^2, вычисленный ранее. В таком случае, если выразить его из матричного уравнения выше, получится уравнение с одним неизвестным – ковариацией c. Если найти ее из него и вычислить $\hat{\mathbb{V}}\hat{\boldsymbol{\beta}}_{(2)}$, то получится найти и В. Аналогичным образом с помощью выражения $\hat{\mathbb{V}}\hat{\boldsymbol{\beta}}_{(4)}$ через $\hat{\mathbb{V}}\hat{\boldsymbol{\beta}}_{(1)}$ ищется М (3 балла).

Для получения полного балла на этом этапе участник мог просто очертить план вычисления В и М, показав, что их можно найти. Здесь давать численный ответ было необязательно[8].

[8] Если это интересно, то количественно $c = -\frac{84}{25\,625}$, откуда нетрудно получить искомые величины.

Задача 3. (30 баллов)
О нелинейной регрессии

I

Из уравнения $y_i = \beta x_i + u_i$ можно получить, что $y_i - bx_i = (\beta - b)\, x_i + u_i$ (при верной нулевой гипотезе). Подставляя этот результат в статистику H, получаем

$$H = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (y_i - bx_i) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left((\beta - b)\, x_i + u_i\right) = (\beta - b) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i =$$

$$= -\sqrt{n}(b - \beta) \frac{1}{n} \sum_{i=1}^{n} x_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i \qquad \text{(2 балла)}$$

II

Как известно, оценка коэффициента в парной линейной регрессии без константы $y_i = \beta x_i + u_i$ имеет вид (вновь при верной нулевой гипотезе)

$$b = \frac{\sum\limits_{i=1}^{n} y_i x_i}{\sum\limits_{i=1}^{n} x_i^2} = \frac{\sum\limits_{i=1}^{n} (\beta x_i + u_i) x_i}{\sum\limits_{i=1}^{n} x_i^2} = \beta + \frac{\sum\limits_{i=1}^{n} u_i x_i}{\sum\limits_{i=1}^{n} x_i^2} \iff b - \beta = \frac{\sum\limits_{i=1}^{n} u_i x_i}{\sum\limits_{i=1}^{n} x_i^2} \qquad \text{(1 балл)}$$

Подставляя этот результат в предыдущее равенство и замечая в результате скалярное произведение соответствующих векторов, имеем

$$H = -\sqrt{n} \left(\frac{\sum\limits_{i=1}^{n} u_i x_i}{\sum\limits_{i=1}^{n} x_i^2} \right) \frac{1}{n} \sum_{i=1}^{n} x_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i = -\frac{\frac{1}{n} \sum\limits_{i=1}^{n} x_i}{\frac{1}{n} \sum\limits_{i=1}^{n} x_i^2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i u_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i =$$

$$= \begin{pmatrix} -\frac{\frac{1}{n} \sum_{i=1}^{n} x_i}{\frac{1}{n} \sum_{i=1}^{n} x_i^2} & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} x_i u_i \\ \frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} u_i \end{pmatrix} \qquad \text{(3 балла)}$$

(III)

Закон повторного математического ожидания позволяет понять, что

$$\mathbb{E}\, u_i = \mathbb{E}_x \underbrace{\mathbb{E}\left(u_i \mid x_i\right)}_{=0} = 0 \qquad \mathbb{E}\left(x_i u_i\right) = \mathbb{E}_x\, \mathbb{E}\left(x_i u_i \mid x_i\right) = \mathbb{E}_x\, x_i \underbrace{\mathbb{E}\left(u_i \mid x_i\right)}_{=0} = 0$$

Отсюда мгновенно получается, что[9]

$$\mathbb{V}\, u_i = \mathbb{E}\, u_i^2 \qquad \mathbb{V}\left(x_i u_i\right) = \mathbb{E}\left(x_i^2 u_i^2\right) \qquad \operatorname{cov}\left(u_i, x_i u_i\right) = \mathbb{E}\left(x_i u_i^2\right)$$

Случайные величины u_i независимы, одинаково распределены и, как подразумевается, имеют конечную дисперсию. Аналогичное верно для величин $x_i u_i$. Таким образом, по классической ЦПТ суммы $\frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} u_i$ и $\frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} u_i x_i$ асимптотически нормальны, поэтому из полученного выше имеем[10]

$$\begin{pmatrix} \frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} x_i u_i \\ \frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} u_i \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \underbrace{\begin{pmatrix} \mathbb{E}\left(x_i^2 u_i^2\right) & \mathbb{E}\left(x_i u_i^2\right) \\ \mathbb{E}\left(x_i u_i^2\right) & \mathbb{E}\, u_i^2 \end{pmatrix}}_{\Sigma} \right)$$

Наконец, из соображений ЗБЧ и теоремы Манна–Вальда при $n \to \infty$ имеем

$$\frac{1}{n} \sum_{i=1}^{n} x_i \xrightarrow{p} \mathbb{E}\, x_i \qquad \frac{1}{n} \sum_{i=1}^{n} x_i^2 \xrightarrow{p} \mathbb{E}\, x_i^2 \qquad -\frac{\frac{1}{n} \sum\limits_{i=1}^{n} x_i}{\frac{1}{n} \sum\limits_{i=1}^{n} x_i^2} \xrightarrow{p} -\frac{\mathbb{E}\, x_i}{\mathbb{E}\, x_i^2}$$

По теореме Слуцкого для случайных векторов получаем, что при $n \to \infty$ верно[11]

$$H = \underbrace{\begin{pmatrix} -\frac{\frac{1}{n} \sum\limits_{i=1}^{n} x_i}{\frac{1}{n} \sum\limits_{i=1}^{n} x_i^2} & 1 \end{pmatrix}}_{\xrightarrow{p} \left(-\frac{\mathbb{E}\, x_i}{\mathbb{E}\, x_i^2} \quad 1\right)} \underbrace{\begin{pmatrix} \frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} x_i u_i \\ \frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} u_i \end{pmatrix}}_{\xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma)} \xrightarrow{d} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -\frac{\mathbb{E}\, x_i}{\mathbb{E}\, x_i^2} & 1 \end{pmatrix} \begin{pmatrix} \mathbb{E}\left(x_i^2 u_i^2\right) & \mathbb{E}\left(x_i u_i^2\right) \\ \mathbb{E}\left(x_i u_i^2\right) & \mathbb{E}\, u_i^2 \end{pmatrix} \begin{pmatrix} -\frac{\mathbb{E}\, x_i}{\mathbb{E}\, x_i^2} \\ 1 \end{pmatrix} \right)$$

---

[9] Здесь использованы те факты, что $\mathbb{V}\,\xi = \mathbb{E}\,\left(\xi^2\right) - \left(\mathbb{E}\,\xi\right)^2$ и $\operatorname{cov}(\xi, \eta) = \mathbb{E}\,(\xi\eta) - \mathbb{E}\,\xi\,\mathbb{E}\,\eta$.

[10] Здесь использовано, что $\mathbb{V}\left(\frac{1}{\sqrt{n}} \sum\limits_{i=1}^{n} \xi_i\right) = \frac{1}{n} \sum\limits_{i=1}^{n} \mathbb{V}\,\xi_i = \mathbb{V}\,\xi_i$ для независимых и одинаково распределенных ξ_i (также верен похожий факт для ковариации).

[11] Здесь использован тот факт, что при линейном преобразовании случайного вектора ковариационная матрица меняется как матрица соответствующей квадратичной формы: $\mathbb{V}\left(A^\top \boldsymbol{\xi}\right) = A^\top\, \mathbb{V}\,\left(\boldsymbol{\xi}\right)\, A$.

Таким образом, статистика H в самом деле асимптотически нормальна и имеет соответствующие предельные $\mathbb{E}H$ и $\mathbb{V}H$.

В этом пункте 1 балл ставился за любое верное обоснование асимптотической нормальности (быть может, менее подробное, чем приведенное выше). Также 1 балл ставился за верное асимптотическое математическое ожидание и 2 балла – за дисперсию. При этом вычисление математического ожидания и дисперсии H на конечной выборке не требовалось.

IV

При верной альтернативной гипотезе $H_1 : \alpha \neq 1$ имеем следующее представление для оценки b

$$b = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} (\beta x_i^{\alpha} + u_i) x_i}{\sum_{i=1}^{n} x_i^2} = \beta \frac{\sum_{i=1}^{n} x_i^{\alpha+1}}{\sum_{i=1}^{n} x_i^2} + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \qquad \text{(4 балла)}$$

V

При верной альтернативной гипотезе $y_i = \beta x_i^{\alpha} + u_i$. С учетом этого, а также результата пункта IV статистика H принимает вид

$$H = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (y_i - b x_i) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\beta x_i^{\alpha} + u_i - b x_i) = \frac{\beta}{\sqrt{n}} \sum_{i=1}^{n} x_i^{\alpha} - \frac{b}{\sqrt{n}} \sum_{i=1}^{n} x_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i =$$

$$= \frac{\beta}{\sqrt{n}} \sum_{i=1}^{n} x_i^{\alpha} - \frac{1}{\sqrt{n}} \left(\beta \frac{\sum_{i=1}^{n} x_i^{\alpha+1}}{\sum_{i=1}^{n} x_i^2} + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \right) \sum_{i=1}^{n} x_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i =$$

$$= \beta\sqrt{n} \underbrace{\left[\frac{1}{n} \sum_{i=1}^{n} x_i^{\alpha} - \frac{\frac{1}{n} \sum_{i=1}^{n} x_i^{\alpha+1}}{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \frac{1}{n} \sum_{i=1}^{n} x_i \right]}_{A_n} + \underbrace{\left[- \frac{\frac{1}{n} \sum_{i=1}^{n} x_i}{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i u_i + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} u_i \right]}_{B_n = \text{Правая часть уравнения (5)}} \qquad \text{(4 балла)}$$

VI

Из соображений ЗБЧ и теоремы Манна–Вальда при $n \to \infty$ имеем

$$\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha} \xrightarrow{p} \mathbb{E}\, x_i^{\alpha} \qquad \frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha+1} \xrightarrow{p} \mathbb{E}\, x_i^{\alpha+1} \qquad \frac{1}{n}\sum_{i=1}^{n} x_i^{2} \xrightarrow{p} \mathbb{E}\, x_i^{2}$$

$$\frac{1}{n}\sum_{i=1}^{n} x_i \xrightarrow{p} \mathbb{E}\, x_i \qquad \frac{\frac{1}{n}\sum\limits_{i=1}^{n} x_i^{\alpha+1}}{\frac{1}{n}\sum\limits_{i=1}^{n} x_i^{2}} \xrightarrow{p} \frac{\mathbb{E}\, x_i^{\alpha+1}}{\mathbb{E}\, x_i^{2}}$$

Таким образом, предел A_n по вероятности имеет вид

$$A_n = \frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha} - \frac{\frac{1}{n}\sum\limits_{i=1}^{n} x_i^{\alpha+1}}{\frac{1}{n}\sum\limits_{i=1}^{n} x_i^{2}} \frac{1}{n}\sum_{i=1}^{n} x_i \xrightarrow{p} \mathbb{E}\, x_i^{\alpha} - \frac{\mathbb{E}\, x_i^{\alpha+1}}{\mathbb{E}\, x_i^{2}}\, \mathbb{E}\, x_i \qquad \text{(4 балла)}$$

VII

Поскольку при $\alpha \neq 1$ предложено[12], что $A_n \not\to 0$ при $n \to \infty$, то

$$\operatorname*{plim}_{n\to\infty} \underbrace{\sqrt{n}}_{\to\infty}\ \underbrace{\beta A_n}_{\to\,\mathrm{const}\,\neq 0} = \infty \qquad \text{(2 балла)}$$

Теперь обратим внимание на второй предел

$$\operatorname*{plim}_{n\to\infty} H_n = \operatorname*{plim}_{n\to\infty} \left(\beta\sqrt{n}A_n + B_n\right) = \operatorname*{plim}_{n\to\infty} \sqrt{n}\left(\beta A_n + \frac{1}{\sqrt{n}}B_n\right)$$

Зафиксируем произвольное $\varepsilon > 0$ и рассмотрим[13]

$$P_\varepsilon = \lim_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{\sqrt{n}}B_n - 0\right| > \varepsilon\right) \leq \lim_{n\to\infty} \frac{\mathbb{V}B_n}{n\varepsilon^2} = 0$$

По определению сходимости по вероятности это в точности означает, что по аналогии с первым пределом

$$\operatorname*{plim}_{n\to\infty} \frac{1}{\sqrt{n}}B_n = 0 \implies \operatorname*{plim}_{n\to\infty} H = \infty \qquad \text{(2 балла)}$$

[12] Кстати говоря, показать это можно, например, с помощью неравенства Гельдера или техникой дифференцирования под знаком математического ожидания.

[13] В этот момент используется неравенство Чебышева и тот факт, что $\mathbb{E}B_n = 0$. Также из сходимости B_n по распределению к нормальной случайной величине следует, что последовательность $\mathbb{V}B_n$ ограничена.

VIII

Рассмотрим следующий критерий отверждения гипотезы H_0:

$$\frac{|H|}{\sqrt{\mathbb{V}H}} > z_{1-\alpha/2} \qquad \text{(4 балла)}$$

Выше $z_{1-\alpha/2}$ – критическое значение стандартного нормального распределения, соответствующее уровню значимости α при двустороннем тесте. Как следует из предыдущих пунктов, статистика $\frac{|H|}{\sqrt{\mathbb{V}H}}$ имеет распределение $\mathcal{N}(0,1)$, поэтому такой критерий корректен.

На практике в тесте выше $\mathbb{V}H$ можно заменить на ее состоятельную оценку или использовать бутстрап-распределение для H.

Задача 4. (25 баллов)
О репликации исследования

1

Да, использовать X_2 в качестве несмещенной (условно на $X_1 > c$) оценки для μ можно, поскольку из соображений независимости

$$X_2 \perp\!\!\!\perp X_1 \implies \mathbb{E}(X_2 \mid X_1 > c) = \mathbb{E}X_2 = \mu$$

Интуитивно это так, поскольку, хотя репликация и проводится только в случае публикации оригинальной работы (то есть когда $X_1 > c$), их результаты независимы. Поскольку репликация проводится «честно», то ее результатам можно доверять.

2

Нет, X_1 не годится в качестве несмещенной оценки μ при условии $X_1 > c$. Вычислим плотность распределения X_1 условно на событие $X_1 > c$

$$p_{X_1}(x_1 \mid X_1 > c) = \frac{\phi(x_1 - \mu)\,[x_1 > c]}{\mathbb{P}(X_1 > c)} = \frac{\phi(x_1 - \mu)\,[x_1 > c]}{1 - \Phi(c - \mu)}$$

В таком случае, нетрудно вычислить искомое математическое ожидание

$$\mathbb{E}(X_1 \mid X_1 > c) = \int\limits_{-\infty}^{+\infty} x_1 p_{X_1}(x_1 \mid x_1 > c)\; dx_1 = \frac{1}{1 - \Phi(c-\mu)} \int\limits_{-\infty}^{+\infty} x_1 \phi(x_1 - \mu)\,[x_1 > c]\; dx_1 =$$

$$= \frac{1}{1 - \Phi(c-\mu)} \int\limits_{-\infty}^{+\infty} (\mu + (x_1 - \mu))\,\phi(x_1 - \mu)\,[x_1 > c]\; dx_1 =$$

$$= \mu \times \underbrace{\frac{1}{1 - \Phi(c-\mu)} \int\limits_{-\infty}^{+\infty} \phi(x_1 - \mu)\,[x_1 > c]\; dx_1}_{=1} + \frac{1}{1 - \Phi(c-\mu)} \int\limits_{-\infty}^{+\infty} (x_1 - \mu)\phi(x_1 - \mu)\,[x_1 > c]\; dx_1 =$$

$$= \mu + \frac{1}{1 - \Phi(c-\mu)} \frac{1}{\sqrt{2\pi}} \int\limits_{c}^{+\infty} (x_1 - \mu) e^{-\frac{(x-\mu)^2}{2}}\; dx_1 = \mu + \underbrace{\frac{1}{1 - \Phi(c-\mu)} \frac{1}{\sqrt{2\pi}} e^{-\frac{(c-\mu)^2}{2}}}_{>0} > \mu$$

Получается, X_1 не может служить несмещенной оценкой для μ.

③

Поскольку случайные величины X_1 и X_2 независимы, то их совместная плотность (условно на $X_1 > c$) факторизуется до вида

$$p\left(x_1, x_2 \mid X_1 > c\right) = p_{X_1}(x_1 \mid X_1 > c)\underbrace{p_{X_2}(x_2 \mid X_1 > c)}_{\phi(x_2-\mu)} = \frac{\phi(x_1-\mu)\phi(x_2-\mu)\left[x_1 > c\right]}{1-\Phi(c-\mu)}$$

④

В случае, если работа публикуется только при достаточно большом X_1, математическое ожидание X_1 при публикации $\mathbb{E}(X_1 \mid X_1 > c)$ также оказывается завышено относительно μ. Таким образом, при условии публикации работы заведомо получается, что $\frac{X_1+X_2}{2}$ – смесь несмещенной оценки X_2 и завышенной оценки X_1, поэтому она сама является завышенной.

Формально в этом нетрудно убедиться, используя вычисленные в предыдущих пунктах математические ожидания

$$\mathbb{E}\left(\frac{X_1+X_2}{2} \,\middle|\, X_1 > c\right) = \frac{1}{2}\underbrace{\mathbb{E}\left(X_2 \mid X_1 > c\right)}_{=\mu} + \frac{1}{2}\underbrace{\mathbb{E}\left(X_1 \mid X_1 > c\right)}_{>\mu} > \frac{1}{2}\mu + \frac{1}{2}\mu = \mu$$