МГУ имени М.В. Ломоносова

Пятая международная универсиада по эконометрике

Задания индивидуального тура

Задание № 1 (30 баллов)

Имеется временной ряд: $y_t = \theta * t + \varepsilon_t + \varepsilon_0,\ t = 1, 2, \dots, T$.

Здесь $E(\varepsilon_t) = 0,\ E(\varepsilon_t^2) = \sigma^2,\ E(\varepsilon_t \varepsilon_j) = 0$ при $t \neq j$; $t, j = 0, 1, 2, \dots, T$.

(1) Вычислите дисперсию МНК-оценки параметра θ.

Будет ли эта оценка состоятельной?

Будет ли она эффективной?

(2) Предложите метод для получения эффективной оценки $\hat{\theta}$.

(3) Пусть $T = 4$ и известно, что $y_1 = -1,\ y_2 = 4,\ y_3 = 6,\ y_4 = 8$.

Вычислите эффективную оценку $\hat{\theta}$.

Задание № 2 (30 баллов)

Рассматривается модель:

$$y_i = \theta_0 + \theta_1 x_i + \theta_2 w_i + \theta_3 x_i w_i + \varepsilon_i$$

Здесь $E(\varepsilon_i) = 0,\ E(\varepsilon_i^2) = \sigma^2, i = 1, 2, \dots, n$; $E(\varepsilon_i \varepsilon_j) = 0$ при $i \neq j$, x_i $и$ w_i – бинарные переменные.

Выборка сконструирована так, что она состоит из четырех групп наблюдений равного объема m (общее число наблюдений $n = 4m$). Причем для первой группы $x_i = w_i = 0$, для второй группы $x_i = 0, w_i = 1$, для третьей группы $x_i = 1, w_i = 0$ и для четвертой группы $x_i = w_i = 1$.

Пусть известно, что выборочные средние значения зависимой переменной для первой, второй, третьей и четвертой групп, соответственно, составляют $\bar{y}_{00},\ \bar{y}_{01},\ \bar{y}_{10}$ и $\bar{y}_{11}$, а посчитанные по выборке дисперсии зависимой переменной для этих групп равны $d_{00},\ d_{01},\ d_{10}$ и d_{11}.

(1) Найдите МНК-оценку $\hat{\theta}_3$.

(2) Найдите дисперсию МНК-оценки $\hat{\theta}_3$.

(3) Найдите отношение дисперсий МНК-оценок параметров $\hat{\theta}_3$ и $\hat{\theta}_0$.

Задание №3 (30 баллов)

Представим, что в некоторой стране уровень зарплаты работников данной специальности описывается точным соотношением:

$$Y_i = \theta_0 + \theta_1 X_i + \theta_2 Q_i$$

Y_i — зарплата i-го работника

X_i — бинарная переменная, которая равна 1, если i-ый работник получил высшее образование.

Q_i — уровень гениальности i-го работника (**ненаблюдаемая** переменная)

В вашем распоряжении имеются представленные в таблице данные о 1000 работников данной специальности:

	В родном городе есть университет	**В родном городе нет университета**
Получил высшее образование	400 человек Средняя зарплата \$5000	100 человек Средняя зарплата \$6000
Не получал высшее образование	100 человек Средняя зарплата \$3000	400 человек Средняя зарплата \$4000

(1) Предложите способ состоятельно оценить параметр θ_1 и вычислите его оценку, используя данные, которые есть в вашем распоряжении. Подробно аргументируйте выбор метода оценивания.

(2) Если доступных данных достаточно для того, чтобы осуществить какой-либо тест, характеризующий релевантность данных для применения предложенного вами подхода, то осуществите его.

Задание № 4 (30 баллов)

Ниже представлены результаты МНК-оценивания двух регрессий, часть из которых не сохранилась. Утерянные оценки коэффициентов регрессий заменены символами.

Модель 1 $\quad \hat{y}_i = \hat{\beta}_1 + \underset{(2)}{\hat{\beta}_2} x_i$

Модель 2 $\quad \hat{y}_i = \hat{\alpha}_1 + \hat{\alpha}_2 x_i + \underset{(2)}{10} w_i \qquad R^2 = 0.8$

Оценивание проводилось по 103 наблюдениям. В скобках под оценками коэффициентов указаны их стандартные ошибки.

Восстановите значение оценки коэффициента $\hat{\beta}_2$ первой регрессии?

Задание №5 (40 баллов)

Производственная функция монополиста имеет следующий вид: $Q_t = w_t K_t^{\beta K} L_t^{\beta L}$, где Q_t - объем выпуска монополиста в период времени t, K_t - объем используемого объема капитала, L_t - объем используемого труда.

Функция спроса на продукцию фирмы: $Q_t = A_t p_t^{-\xi}$, где p_t – цена на продукцию монополиста, а ξ – коэффициент, $\xi > 1$. Предложение всех факторов производства совершенно эластично. Цены на ресурсы p_K и p_L не меняются в времени. В распоряжении исследователей имеются следующие временные ряды $(p_t, Q_t, K_t, L_t), t = 1, \dots, T$. Товар монополиста не является товаром длительного пользования. Задача исследования – оценить эластичность выпуска по труду β_L.

Переменная A_t, определяющая объем спроса на рынке, является логорифмически нормальным «белым шумом»: $\ln(A_t) \sim i.i.d.N(0; \sigma_A^2)$

Переменная w_t, определяющая производительность фирмы, описывается авторегрессионным процессом первого порядка вида: $\ln(w_t) = \rho \ln(w_{t-1}) + \varepsilon$, где $\rho < 1$ начальное значение $\ln w_0 \sim N(0, \sigma_w^2)$, а случайная величина $\varepsilon_t \sim i.i.d.N(0, \sigma_\varepsilon^2)$

Монополист максимизирует прибыль в каждом периоде

Представители научной школы «МНК – наш» утверждают, что для решения поставленной задачи достаточно применить метод наименьших квадратов к следующей модели:

$$\ln Q_t = const + \beta_K \ln K_t + \beta_L \ln L_t$$

Полученный коэффициент $\hat{\beta}_L$, по их мнению, будет состоятельной оценкой эластичности выпуска по труду.

Представители конкурирующей научной школы «Все под контроль» утверждают, что предложенная оппонентами оценка страдает от эндогенности, так как спрос на труд и объем выпуска одновременно зависят от ненаблюдаемой реализации случайной величины w_t

Они также утверждают, что включение дополнительной переменной $\ln p_t$ в качестве контрольной устранит эту проблему. Их предложение – оценить методом наименьших квадратов следующую модель:

$$\ln Q_t = const + \beta_K \ln K_t + \beta_L \ln L_t + \beta_p \ln p_t$$

Уважаемые участники универсиады, помогите решить данный спор. Какой из предложенных подходов позволит решить поставленную задачу?

Задание №6 (40 баллов)

Пусть Y и D – две случайные переменные, причем D принимает значение из дискретного множества $\{0,1, \dots, K\}$, где $K \leq \infty$. Нас интересует функция условного математического ожидания $f(x) = E[Y|D = x]$. Чтобы оценить эту функцию, мы делаем предположение о ее линейности:

$$f(x) = \alpha + \beta\, x$$

и оцениваем параметры α и β с помощью МНК.

1. Выразите параметр β в виде функции от некоторых моментов совместного распределения случайных величин (Y, D).
   Обозначая полученное выражение через β_{OLS}, покажите, что оно может быть представлено в терминах совместного распределения величин $(E[Y|D], D)$ в виде
   $$\beta_{OLS} = \frac{E[E[Y|D](D - E[D])]}{E[(D - E[D])^2]}$$
2. Пусть $\mu_j = E[Y|D = j] - E[Y|D = j - 1]$, $j \epsilon \{1, \dots, K\}$. Покажите, что β_{OLS} можно представить в виде $\beta_{OLS} = \sum_{j=1}^{K} \lambda_j\, \mu_j$, где $\lambda_j \geq 0$ и $\sum_{j=1}^{K} \lambda_j = 1$. Что Вы можете сказать о распределении значений весовых коэффициентов λ_j?
   Указание: воспользуйтесь формулами
   $$E[g(D)] = \sum_k g(k)\, P(D = k) \;, \qquad f(k) = f(0) + \sum_{j=1}^{k} (f(j) - f(j-1))$$
3. Дайте интерпретацию результатов, полученных в пункте 2 (в том числе, в случае монотонно возрастающей функции $f(x)$).
4. Рассмотрим следующую ситуацию: Вы оценили модель $ln\ Y = \alpha + \beta\, D + \varepsilon$, где Y – стартовая зарплата работника (назначенная после окончания учебы), D - количество лет обучения. Вы представляете результаты на конференции, и один из коллег высказывает следующее замечание: «Очевидно, что связь между логарифмом заработной платы и числом лет обучения нелинейная. Поэтому Ваши результаты не имеют особого смысла». Что Вы ему ответите?
5. В каких из вышеперечисленных пунктов (1 – 4) нет необходимости предполагать, что «истинная» функция $f(x)$ является линейной (это предположение мы сделали только для того, чтобы оценить $f(x)$)?