Универсиада по эконометрике

2019

Заочный тур

Задание 1. Рассматривается уравнение $y_i = \beta_1 x_i^{(1)} + \beta_2 x_i^{(2)} + \varepsilon_i$, для которого выполнены все предпосылки классической линейной модели множественной регрессии. Параметры уравнения оцениваются при помощи МНК. Имеются следующие данные о 1000 наблюдениях соответствующих переменных:

$$\sum_{i=1}^{1000} x_i^{(1)} x_i^{(2)} = \sum_{i=1}^{1000} \left(x_i^{(2)}\right)^2 = 1{,}0, \qquad \sum_{i=1}^{1000} \left(x_i^{(1)}\right)^2 = 3{,}0,$$

$$\sum_{i=1}^{1000} x_i^{(1)} y_i = 30, \quad \sum_{i=1}^{1000} x_i^{(2)} y_i = 20.$$

(1.1) Вычислите оценку разности коэффициентов $\beta_1 - \beta_2$. Ответ в виде числа.

(1.2) Пусть также известно, что сумма квадратов остатков в оцененной регрессии оказалась равна 5988. Постройте 99-процентный доверительный интервал для разности $\beta_1 - \beta_2$. Ответ в виде двух чисел через точку с запятой (первое число – начало интервала, второе – конец).

Решение задания 1.

Пункт (1.1)

$$X'X = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$$

$$(X'X)^{-1} = \begin{pmatrix} +0{,}5 & -0{,}5 \\ -0{,}5 & +1{,}5 \end{pmatrix}$$

$$X'y = \begin{pmatrix} 30 \\ 20 \end{pmatrix}$$

$$\hat{\beta} = (X'X)^{-1} X'y = \begin{pmatrix} 5 \\ 15 \end{pmatrix}$$

Ответ (1.1): –10

Пункт (1.2) Оценка ковариационной матрицы вектора оценок коэффициентов:

$$\hat{V}(\hat{\beta}) = (X'X)^{-1} * S^2 = \begin{pmatrix} +0{,}5 & -0{,}5 \\ -0{,}5 & +1{,}5 \end{pmatrix} * \frac{5988}{1000 - 2} = \begin{pmatrix} +3 & -3 \\ -3 & +9 \end{pmatrix}$$

$$\widehat{se}\left(\widehat{\beta_1} - \widehat{\beta_2}\right) = \sqrt{\widehat{var}\left(\widehat{\beta_1} - \widehat{\beta_2}\right)} = \sqrt{\widehat{var}\left(\widehat{\beta_1}\right) + \widehat{var}\left(\widehat{\beta_2}\right) - 2 * \widehat{cov}\left(\widehat{\beta_1}, \widehat{\beta_2}\right)} = \sqrt{3 + 9 - 2 * (-3)}$$
$$= \sqrt{18} = 4{,}24$$

Теперь построим доверительный интервал:

$$\left(\widehat{\beta_1} - \widehat{\beta_2} - \widehat{se}\left(\widehat{\beta_1} - \widehat{\beta_2}\right) * 2{,}58, \qquad \widehat{\beta_1} - \widehat{\beta_2} + \widehat{se}\left(\widehat{\beta_1} - \widehat{\beta_2}\right) * 2{,}58\right)$$

$$(-20{,}946, \qquad +0{,}946)$$

Ответ (1.2): $-20,946$; $0,946$

Задание 2. Используя МНК, исследователь оценил параметры модели заработной платы работника некоторой отрасли:

$$\widehat{\ln w_i} = 10{,}1 + 0{,}20 * male_i + 0{,}90 \ln education_i + 0{,}1 * experience_i, \qquad R^2 = 0{,}6.$$

Здесь $male_i$ — бинарная переменная, равная единице, если i-ый работник мужчина, и равная нулю в противном случае; w_i — заработная плата i-го работника в долларах; $education_i$ — число лет обучения i-го работника; $experience_i$ — опыт работы i-го работника (в годах).

Другой исследователь, используя в точности те же самые данные оценил параметры следующей модели:

$$\widehat{\ln w_i} = \widehat{\alpha_1} + \widehat{\alpha_2} * female_i + \widehat{\alpha_3} \ln education_i + \widehat{\alpha_4} * experience_i.$$

Здесь $female_i$ — бинарная переменная, равная единице, если i-ый работник **женщина**, и равная нулю в противном случае; w_i — заработная плата i-го работника в долларах; $education_i$ — число **месяцев** обучения i-го работника; $experience_i$ — опыт работы i-го работника (в **месяцах**).

(2.1) Вычислите оценку $\widehat{\alpha_1}$. Ответ в виде числа.

(2.2) Вычислите оценку $\widehat{\alpha_3}$. Ответ в виде числа.

(2.3) Вычислите R^2 в новой модели. Ответ в виде числа (округлите до двух знаков после запятой)

Решение задания 2.

$$\widehat{\ln w_i} = 10{,}1 + 0{,}20 * (1 - female_i) + 0{,}90 \ln 12 * education_i + 0{,}1 * 12 * \ experience_i$$

$$\widehat{\ln w_i} = 10{,}1 + 0{,}2 + 0{,}9 \ln 12 - 0{,}20 * female_i + 0{,}90 \ln education_i + 1{,}2 * \ experience_i$$

Ответ (2.1): $\boldsymbol{\widehat{\alpha_1} = 10,1 + 0,2 + 0,9 \ln 12 = 12,54}$

Ответ (2.2): $\boldsymbol{\widehat{\alpha_3} = 0,9}$

Ответ (2.3): $\boldsymbol{R^2 = 0,6}$

Задание 3. Вопросы этого задания основаны на следующем на результатах следующего эксперимента: 600 водителей, выбранных случайным образом, попросили пройти специальный тест на вождение автомобилем. Для каждого водителя были собраны следующие данные: Pass — фиктивная переменная, равная единице, если водитель сдал тест, Male — фиктивная переменная, равная единице, если водитель мужчина, и равная 0, если водитель женщина, Experience — опыт вождения автомобиля (в годах). В таблице представлены результаты двух пробит-моделей, оцененных на основе имеющихся данных.

Зависимая переменная: Pass. Метод оценивания: пробит-модель		
	(1)	(2)
Experience	0,06 (0,01)	0,08 (0,03)
Male	—	–0.17 (0,02)
Male*Experience	—	–0,04 (0,01)
Constant	0,70 (0,12)	0,80 (0,20)

(3.1) Иван — водитель с пятилетним стажем вождения. На основе каждой из двух моделей оцените вероятность сдачи теста для Ивана. Ответ в виде числа (округлите до двух знаков после запятой).

(3.2) На основе второй модели вычислите предельный эффект увеличения опыта вождения для Ивана. Ответ в виде числа (округлите до двух знаков после запятой).

(3.3) Зависит ли от пола то, как опыт влияет на успешность сдачи? (Для ответа на вопрос осуществите соответствующий статистический тест, используя пятипроцентный уровень значимости). Ответ в виде описания последовательности ваших действий (что за тест, какие параметры и цифры были использованы, что получилось, то есть какой вывод из этого можно сделать)

Решение задания 3.

Пункт (3.1) Φ(0,06*5+0,7)=0,841; Φ(0,08*5-0,17-0,04*5+0,80)=0,797

Ответ (3.1): 0,84; 0,8

Пункт (3.2) Φ'(0,08*5-0,17-0,04*5+0,80)*(0,08-0,04)=0,011

Ответ (3.2): 0,01

Пункт (3.3) $-\frac{0,04}{0,01} = -4$. По модулю эта величина больше, чем 1,96. Следовательно, коэффициент при переменной Male*Experience статистически значим на пятипроцентном уровне. Поэтому можно заключить, что влияние опыта вождения на успешность сдачи экзамена зависит от пола.

Ответ (3.3): Да.

Задание 4. В стране X постоянно возникают проблемы с пробками. В связи с этим государство решило хотя бы частично решить эту проблему путем введения новой системы парковки. У данной системы, однако, есть много противников, которые полагают, что она неэффективна и коррумпирована. Для устранения разногласий государство решило провести эксперимент и проверить, помогает ли данная система сократить количество пробок.

В стране есть всего два города: A и B. Путём подбрасывания монеты государство решает, в каком городе оно проведёт реформу системы парковки, а в каком не будет делать ничего. Пусть W_A – показатель того, что был выбран город A. Предположим, что $E[W_A] = \frac{1}{2}$. Предположим также, что верны следующие равенства:

$$Y_i^A(w) = \mu_A + \tau w + \varepsilon_i^A$$

$$Y_i^B(w) = \mu_B + \tau w + \varepsilon_i^B$$

$$E[\varepsilon_i^A] = E[\varepsilon_i^B] = 0$$

$$V[\varepsilon_i^A] = V[\varepsilon_i^B] = \sigma^2$$

Где $Y_i^x(w)$ – потенциальный исход для индивида i, который живёт в городе x, если переменная, отвечающая за проведение политики, равна w. Другими словами, $Y_i^x(0)$ – время, которое индивид i проводит в пробке в городе x, если в нём не произошло никаких изменений, $Y_i^x(1)$ – если новая система платежей на парковках была введена. Поскольку W_A определяется вследствие бросания монеты, данная величина независима от ε_i^x. Основная цель – оценить эффект проведения политики (τ).

Наблюдаемые последствия политики (время, проводимое в пробках каждым индивидом в каждом городе) определяется следующим образом:

$$Y_i^A \coloneqq Y_i^A(W_A)$$

$$Y_i^B \coloneqq Y_i^B(1 - W_A)$$

У нас есть две случайные выборки из каждого города $D_A = \{Y_i^A\}_{i=1}^{n_A}$ и $D_B = \{Y_i^A\}_{i=1}^{n_B}$ и рассматривается следующая оценка τ:

$$\hat{\tau} = W_A\left(\overline{Y^A} - \overline{Y^B}\right) + (1 - W_A)\left(\overline{Y^B} - \overline{Y^A}\right)$$

Где $\overline{Y^x} \coloneqq \frac{1}{n_x}\sum_{i=1}^{n_x} Y_i^x$

(4.1) Предполагая, что $\mu_A = 5, \mu_B = 6, \tau = 2$, посчитайте смещение $\hat{\tau}$:

$$bias \coloneqq E[\hat{\tau}] - \tau$$

Приведите ответ в виде числа.

(4.2) Предполагая, что $\mu_A = 5, \mu_B = 6, \tau = 2, n_A = 100, n_B = 200, \sigma^2 = 1$, найдите оценку дисперсии $\hat{\tau}$. Ответ в виде числа (округлите до двух знаков после запятой).

(4.3) Состоятельная ли оценка $\hat{\tau}$ (при условии, что n_A, n_B стремятся к бесконечности)? Обоснуйте ответ, приведя необходимые вычисления.

Решение задания 4.

Пункт (4.1) Оценка имеет следующую форму:

$$\hat{\tau} = \tau + \{W_A = 1\}(\mu_A - \mu_B) + \{W_A = 0\}(\mu_B - \mu_A) + \{W_A = 1\}\left(\overline{\varepsilon^A} - \overline{\varepsilon^B}\right) + \{W_A = 0\}(\overline{\varepsilon^B} - \overline{\varepsilon^A})$$

Так как $E[\{W_A = 1\}] = \frac{1}{2}$ и ошибки не зависят от W_i и имеют нулевое матожидание, мы получаем, что оценка является несмещённой.

Ответ (4.1): 0

Пункт (4.2) Условная (по W_A) дисперсия всегда равна $\frac{(n_A+n_B)\sigma^2}{n_A n_B}$, в то время как смещение равно $\pm(\mu_A + \mu_B)$. В результате, общая дисперсия оценки равна:

$$V[\hat{\tau}] = (\mu_A + \mu_B)^2 + \frac{(n_A + n_B)\sigma^2}{n_A n_B} = (5 + 6)^2 + \frac{(100 + 200)1}{100 * 200} = 121 + \frac{3}{200} = 121{,}015 \approx 121{,}02$$

Ответ (4.2): 121,02

Пункт (4.3) Дисперсия сходится к $(\mu_A + \mu_B)^2$, когда $n_1, n_2 \to \infty$, b, следовательно, оценка не состоятельна.

Ответ (4.3): Оценка не состоятельна.