

1 МНК — это...

Минитеория:

1. Истинная модель. Например, $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + u_i$.
2. Формула для прогнозов. Например, $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{\beta}_3 z_i$.
3. Метод наименьших квадратов, $\sum (y_i - \hat{y}_i)^2 \rightarrow \min$.

Задачи:

1.1 Каждый день Маша ест конфеты и решает задачи по эконометрике. Пусть x_i — количество решённых задач, а y_i — количество съеденных конфет.

x_i	y_i
1	1
2	2
2	4

1. Рассмотрим модель $y_i = \beta x_i + u_i$:
 - (a) Найдите МНК-оценку $\hat{\beta}$ для имеющихся трёх наблюдений.
 - (b) Нарисуйте исходные точки и полученную прямую регрессии.
 - (c) Выведите формулу для $\hat{\beta}$ в общем виде для n наблюдений.
2. Рассмотрим модель $y_i = \beta_1 + \beta_2 x_i + u_i$:
 - (a) Найдите МНК-оценки $\hat{\beta}_1$ и $\hat{\beta}_2$ для имеющихся трёх наблюдений.
 - (b) Нарисуйте исходные точки и полученную прямую регрессии.
 - (c) Выведите формулу для $\hat{\beta}_2$ в общем виде для n наблюдений.

1.2 Упростите выражения:

1. $n\bar{x} - \sum x_i$
2. $\sum (x_i - \bar{x})\bar{x}$
3. $\sum (x_i - \bar{x})\bar{z}$
4. $\sum (x_i - \bar{x})^2 + n\bar{x}^2$

1.3 При помощи метода наименьших квадратов найдите оценку неизвестного параметра θ в следующих моделях:

1. $y_i = \theta + \theta x_i + \varepsilon_i$;
2. $y_i = 1 + \theta x_i + \varepsilon_i$;
3. $y_i = \theta/x_i + \varepsilon_i$;
4. $y_i = \theta x_i + (1 - \theta)z_i + \varepsilon_i$.

1.4 Найдите МНК-оценки параметров α и β в модели $y_i = \alpha + \beta y_i + \varepsilon_i$.

1.5 Рассмотрите модели $y_i = \alpha + \beta(y_i + z_i) + \varepsilon_i$, $z_i = \gamma + \delta(y_i + z_i) + \varepsilon_i$.

1. Как связаны между собой $\hat{\alpha}$ и $\hat{\gamma}$?
2. Как связаны между собой $\hat{\beta}$ и $\hat{\delta}$?

1.6 Как связаны МНК-оценки параметров α, β и γ, δ в моделях $y_i = \alpha + \beta x_i + \varepsilon_i$ и $z_i = \gamma + \delta x_i + v_i$, если $z_i = 2y_i$?

1.7 Для модели $y_i = \beta_1 x_i + \beta_2 z_i + \varepsilon_i$ решите условную задачу о наименьших квадратах:

$$Q(\beta_1, \beta_2) := \sum_{i=1}^n (y_i - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i)^2 \rightarrow \min_{\beta_1 + \beta_2 = 1}.$$

1.8 Перед нами два золотых слитка и весы, производящие взвешивания с ошибками. Взвесив первый слиток, мы получили результат 300 грамм, взвесив второй слиток — 200 грамм, взвесив оба слитка — 400 грамм. Оцените вес каждого слитка методом наименьших квадратов.

1.9 Аня и Настя утверждают, что лектор опоздал на 10 минут. Таня считает, что лектор опоздал на 3 минуты. С помощью МНК оцените, на сколько опоздал лектор.

1.10 Есть двести наблюдений. Вовочка оценил модель $\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 x$ по первой сотне наблюдений. Петечка оценил модель $\hat{y} = \hat{\gamma}_1 + \hat{\gamma}_2 x$ по второй сотне наблюдений. Машенька оценила модель $\hat{y} = \hat{m}_1 + \hat{m}_2 x$ по всем наблюдениям.

1. Возможно ли, что $\hat{\beta}_2 > 0$, $\hat{\gamma}_2 > 0$, но $\hat{m}_2 < 0$?
2. Возможно ли, что $\hat{\beta}_1 > 0$, $\hat{\gamma}_1 > 0$, но $\hat{m}_1 < 0$?
3. Возможно ли одновременное выполнение всех упомянутых условий?

1.11 Эконометрист Вовочка собрал интересный набор данных по студентам третьего курса:

- переменная y_i — количество пирожков, съеденных i -ым студентом за прошлый год;
- переменная f_i , которая равна 1, если i -ый человек в выборке — женщина, и 0, если мужчина.
- переменная¹ m_i , которая равна 1, если i -ый человек в выборке — мужчина, и 0, если женщина.

Вовочка попробовал оценить 4 регрессии:

- A: y на константу и f ;
B: y на константу и m ;
C: y на f и m без константы;
D: y на константу, f и m .

1. Какой смысл будут иметь оцениваемые коэффициенты?
2. Как связаны между собой оценки коэффициентов этих регрессий?

1.12 Эконометрист Вовочка оценил методом наименьших квадратов модель 1, $y = \beta_1 + \beta_2 x + \beta_3 z + \varepsilon$, а затем модель 2, $y = \beta_1 + \beta_2 x + \beta_3 z + \beta_4 w + \varepsilon$. Сравните полученные ESS , RSS , TSS и R^2 .

1.13 Что происходит с TSS , RSS , ESS , R^2 при добавлении нового наблюдения? Если величина может изменяться только в одну сторону, то докажите это. Если возможны и рост, и падение, то приведите пример.

1.14 Эконометресса Аглая подглядела, что у эконометрессы Жозефины получился R^2 равный 0.99 по 300 наблюдениям. От чёрной зависти Аглая не может ни есть, ни спать.

1. Аглая добавила в набор данных Жозефины ещё 300 наблюдений с такими же регрессорами, но противоположными по знаку игреками, чем были у Жозефины. Как изменится R^2 ?
2. Жозефина заметила, что Аглая добавила 300 наблюдений и вычеркнула их, вернув в набор данных в исходное состояние. Хитрая Аглая решила тогда добавить всего одно наблюдение так, чтобы R^2 упал до нуля. Удастся ли ей это сделать?

¹Это нетолерантная задача и здесь либо f равно 1, либо m

- 1.15** На работе Феофан построил парную регрессию по трём наблюдениям и посчитал прогнозы \hat{y}_i . Придя домой он отчасти вспомнил результаты:

y_i	\hat{y}_i
0	1
6	?
6	?

Поднапрягшись, Феофан вспомнил, что третий прогноз был больше второго. Помогите Феофану восстановить пропущенные значения.

- 1.16** Вся выборка поделена на две части. Возможны ли такие ситуации:

1. Выборочная корреляция между y и x примерно равна нулю в каждой части, а по всей выборке примерно равна единице;
2. Выборочная корреляция между y и x примерно равна единице в каждой части, а по всей выборке примерно равна нулю?

2 Дифференциал — просто няшка!

Минитеория.

Дифференциал для матриц подчиняется правилам:

1. $da = 0$, $dA = 0$;
2. $d(RS) = dR \cdot S + R \cdot dS$

- 2.1** Вспомним дифференциал :)

1. Известно, что $f(x) = x^2 + 3x$. Найдите $f'(x)$ и df . Чему равен dx в точке $x = 5$ при $df = 0.1$?
2. Известно, что $f(x_1, x_2) = x_1^2 + 3x_1x_2^3$. Найдите df . Чему равен df в точке $x_1 = -2$, $x_2 = 1$ при $dx_1 = 0.1$ и $dx_2 = -0.1$?
3. Известно, что $F = \begin{pmatrix} 5 & 6x_1 \\ x_1x_2 & x_1^2x_2 \end{pmatrix}$. Найдите dF .
4. Известно, что $F = \begin{pmatrix} 7 & 8 & 9 \\ 2 & -1 & -2 \end{pmatrix}$. Найдите dF .
5. Матрица F имеет размер 2×2 , в строке i столбце j у неё находится элемент f_{ij} . Выпишите выражение $\text{tr}(F'dF)$ в явном виде без матриц.

- 2.2** Пусть t — скалярная переменная, r, s — векторные переменные, R, S — матричные переменные. Кроме того, a, b — векторы констант, A, B — матрицы констант.

Применив базовые правила дифференцирования найдите:

1. $d(ARB)$;
2. $d(r'r)$;
3. $d(r'Ar)$;
4. $d(R^{-1})$, воспользовавшись тем, что $R^{-1} \cdot R = I$;
5. $d \cos(r'r)$;
6. $d(r'Ar/r'r)$.

- 2.3** В методе наименьших квадратов минимизируется функция

$$Q(\hat{\beta}) = (y - X\hat{\beta})'(y - X\hat{\beta}).$$

1. Найдите $dQ(\hat{\beta})$ и $d^2Q(\hat{\beta})$;
2. Выпишите условия первого порядка для задачи МНК;
3. Выразите $\hat{\beta}$ предполагая, что $X'X$ обратима.

2.4 В методе LASSO минимизируется функция

$$Q(\hat{\beta}) = (y - X\hat{\beta})'(y - X\hat{\beta}) + \lambda\hat{\beta}'\hat{\beta},$$

где λ — положительный параметр, штрафующий функцию за слишком большие значения $\hat{\beta}$.

1. Найдите $dQ(\hat{\beta})$ и $d^2Q(\hat{\beta})$;
2. Выпишите условия первого порядка для задачи LASSO;
3. Выразите $\hat{\beta}$.

3 МНК в матрицах и геометрия!

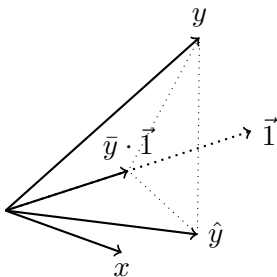
3.1 Пусть $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ — регрессионная модель, где $X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, $y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$,

$\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix}$, ошибки ε_i независимы и нормально распределены с $\mathbb{E}(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2 I$. Для удобства

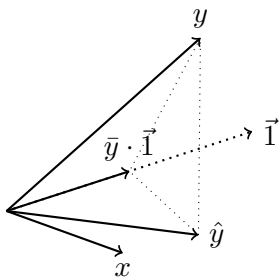
расчётов даны матрицы: $X'X = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ и $(X'X)^{-1} = \begin{pmatrix} 0.3333 & -0.3333 & 0.0000 \\ -0.3333 & 1.3333 & -1.0000 \\ 0.0000 & -1.0000 & 2.0000 \end{pmatrix}$

1. Укажите число наблюдений.
2. Укажите число регрессоров в модели, учитывая свободный член.
3. Найдите $TSS = \sum_{i=1}^n (y_i - \bar{y})^2$.
4. Найдите $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$.
5. Методом МНК найдите оценку для вектора неизвестных коэффициентов.
6. Чему равен R^2 в модели? Прокомментируйте полученное значение с точки зрения качества оценённого уравнения регрессии.

3.2 Найдите на картинке все перпендикулярные векторы. Найдите на картинке все прямоугольные треугольники. Сформулируйте для них теоремы Пифагора.



3.3 Покажите на картинке TSS , ESS , RSS , R^2 , $sCorr(\hat{y}, y)$, $sCov(\hat{y}, y)$



3.4 Предложите аналог R^2 для случая, когда константа среди регрессоров отсутствует. Аналог должен быть всегда в диапазоне $[0; 1]$, совпадать с обычным R^2 , когда среди регрессоров есть константа, равняться единице в случае нулевого $\hat{\varepsilon}$.

3.5 Вася оценил регрессию y на константу, x и z . А затем, делая ему нечего, регрессию y на константу и полученный \hat{y} . Какие оценки коэффициентов у него получатся? Чему будет равна оценка дисперсии коэффициента при \hat{y} ? Почему оценка коэффициента неслучайна, а оценка её дисперсии положительна?

3.6 При каких условиях $TSS = ESS + RSS$?

3.7 Вася построил парную регрессию y на x и получил коэффициент наклона 1.4. Построил парную регрессию x на y и получил коэффициент наклона 0.6. Известно, что $y = x + z$.

1. Найдите выборочные корреляции между x и y , y и z , x и z ;
2. В какой пропорции соотносятся выборочные дисперсии x , y и z ?

4 Распределения, связанные с проецированием

4.1 Компоненты вектора $x = (x_1, x_2)'$ независимы и имеют стандартное нормальное распределение. Вектор y задан формулой $y = (2x_1 + x_2 + 2, x_1 - x_2 - 1)$.

1. Выпишите совместную функцию плотности вектора x ;
2. Нарисуйте на плоскости линии уровня функции плотности вектора x ;
3. Выпишите совместную функцию плотности вектора y ;
4. Найдите собственные векторы и собственные числа ковариационной матрицы вектора y ;
5. Нарисуйте на плоскости линии уровня функции плотности вектора y .

4.2 Компоненты вектора $x = (x_1, x_2, x_3)'$ независимы и имеют стандартное нормальное распределение.

1. Как выглядят в пространстве поверхности уровня совместной функции плотности?
2. Рассмотрим три апельсина с кожурой одинаковой очень маленькой толщины: бэби-апельсин радиуса 0.1, стандартный апельсин радиуса 1 и гранд-апельсин радиуса 10. В кожуру какого апельсина вектор x попадает с наибольшей вероятностью?
3. Мы проецируем случайный вектор на x на плоскость $2x_1 + 3x_2 - 7x_3 = 0$. Какое распределение имеет квадрат длины проекции?
4. Введём вектор y независимый от x и имеющий такое же распределение. Спроецируем вектор x на плоскость проходящую через начало координат и перпендикулярную вектору y . Какое распределение имеет квадрат длины проекции?

5 Ожидания и ковариационные матрицы

5.1 Исследовательница Мишель собрала данные по 20 студентам. Переменная y_i — количество решённых задач по эконометрике i -ым студентом, а x_i — количество просмотренных серий любимого сериала за прошедший год. Оказалось, что $\sum y_i =$, $\sum x_i =$, $\sum x_i^2 =$, $\sum y_i^2 =$, $\sum x_i y_i =$.

- 1.
2. Предположим дополнительно, что $\text{Var}(u_i|X) = \sigma^2$ и u_i при фиксированных X независимы. Найдите $\text{Var}(y_i|X)$, $\text{Var}(y_i(x_i - \bar{x})|X)$, $\text{Var}(\sum y_i(x_i - \bar{x})|X)$, $\text{Var}(\hat{\beta}_2|X)$.

5.2 Пусть регрессионная модель $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$, $i = 1, \dots, n$, задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = (\beta_1 \ \beta_2 \ \beta_3)'$. Известно, что $\mathbb{E}\varepsilon = 0$ и $\text{Var}(\varepsilon) = 4 \cdot I$. Известно также, что:

$$y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

Для удобства расчётов ниже приведены матрицы:

$$X'X = \begin{pmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ и } (X'X)^{-1} = \begin{pmatrix} 0.5 & -0.5 & 0 \\ -0.5 & 1 & -0.5 \\ 0 & -0.5 & 1.5 \end{pmatrix}.$$

1. Найдите $\mathbb{E}(\hat{s}^2)$, \hat{s}^2 .
2. Найдите $\text{Var}(\varepsilon_1)$, $\text{Var}(\beta_1)$, $\text{Var}(\hat{\beta}_1)$, $\widehat{\text{Var}}(\hat{\beta}_1)$, $\mathbb{E}(\hat{\beta}_1^2) - \beta_1^2$;
3. Предполагая нормальность ошибок, постройте 95% доверительный интервал для β_2 .
4. Предполагая нормальность ошибок, проверьте гипотезу $H_0: \beta_2 = 0$;
5. Найдите $\text{Cov}(\hat{\beta}_2, \hat{\beta}_3)$, $\widehat{\text{Cov}}(\hat{\beta}_2, \hat{\beta}_3)$, $\text{Var}(\hat{\beta}_2 - \hat{\beta}_3)$, $\widehat{\text{Var}}(\hat{\beta}_2 - \hat{\beta}_3)$;
6. Найдите $\text{Var}(\beta_2 - \beta_3)$, $\text{Corr}(\hat{\beta}_2, \hat{\beta}_3)$, $\widehat{\text{Corr}}(\hat{\beta}_2, \hat{\beta}_3)$;
7. Предполагая нормальность ошибок, проверьте гипотезу $H_0: \beta_2 = \beta_3$;

6 Гипотезы и интервалы

7 Гамма, бета

7.1 Вася делает эксперименты без усталости со скоростью d экспериментов в минуту. Каждый эксперимент независимо от других может закончиться успехом с вероятностью p или неудачей.

Пусть X — количество успехов за первую минуту, а Y — номер опыта, в котором произошёл первый успех, Z — время, когда случился первый успех.

1. Найдите $\mathbb{P}(X = k)$, $\mathbb{E}(X)$, $\text{Var}(X)$. Как называется закон распределения X ?
2. Найдите $\mathbb{P}(Y = k)$, $\mathbb{E}(Y)$, $\text{Var}(Y)$. Как называется закон распределения Y ?
3. Найдите $\mathbb{P}(Z \leq t)$, $\mathbb{E}(Z)$, $\text{Var}(Z)$.

Теперь Вася ускоряется и устремляет d в бесконечность. Из-за того, что он торопится, p начинает стремиться к нулю :) Причем ожидаемое количество успехов за минуту оказывается постоянно и равно λ .

4. Выразите p через λ и d .
5. Найдите предел $\mathbb{P}(Z \leq t)$. Является ли предельная функция $\mathbb{P}(Z \leq t)$ непрерывной? Какая в предельном случае получается функция плотности у величины Z ? Как называется этот закон распределения Z ? Чему равен предел $\mathbb{E}(Z)$ и $\text{Var}(Z)$?

6. Найдите предел вероятности $\mathbb{P}(X = k)$ и пределы $\mathbb{E}(X)$ и $\text{Var}(X)$. Как называется предельный закон распределения X ?

7.2 Энтомолог Джон Поллак ловит бабочек. На поимку i -ой бабочки у него уходит Y_i минут, величины Y_i независимы. Каждая Y_i имеет экспоненциальное распределение с интенсивностью λ бабочек в минуту. Всего он решил поймать n бабочек. Рассмотрим величины $S = Y_1 + \dots + Y_n$, $X_1 = Y_1/S$, $X_2 = Y_2/S$, \dots , $X_{n-1} = Y_{n-1}/S$.

1. Выпишите совместную функцию плотности Y_1, \dots, Y_n ;
2. Найдите совместную функцию плотности $X_1, X_2, X_3, \dots, X_{n-1}, S$.
3. Зависит ли величина S и вектор X_1, X_2, \dots, X_{n-1} ?
4. С точностью до сомножителя выпишите функцию плотности S . Как называется закон распределения S ?
5. С точностью до сомножителя выпишите совместную функцию плотности для X_1, \dots, X_{n-1} .

Рассмотрим также величины $Z_1 = Y_1/(Y_1 + Y_2)$, $Z_2 = (Y_1 + Y_2)/(Y_1 + Y_2 + Y_3)$, \dots , $Z_{n-1} = (Y_1 + \dots + Y_{n-1})/(Y_1 + \dots + Y_n)$.

6. Найдите совместную функцию плотности $Z_1, Z_2, \dots, Z_{n-1}, S$.
7. Зависимы ли величины $Z_1, Z_2, \dots, Z_{n-1}, S$?
8. С точностью до константы найдите частную функцию плотности S и каждого Z_i в отдельности;

7.3 Быстрый исследователь Вася снова проводит независимые идентичные опыты с очень высокой скоростью. В среднем λ опытов в минуту оказываются успешными. Поэтому время до очередного успеха можно считать экспоненциально распределённым, а время от начала до k -го успеха — имеющим гамма-распределение $\text{Gamma}(k, \lambda)$. На этот раз Вася решил дождаться k_1 успеха, затем быстренько пообедать, а затем дождаться ещё k_2 успехов. Пусть X_1 — время от начала наблюдения до обеда, а X_2 — время от обеда до конца опытов. Также введём $S = X_1 + X_2$ и $Z = X_1/S$ — долю времени до обеда от общего времени на опыты.

1. Найдите совместную функцию плотности S и Z с точностью до константы.
2. Являются ли S и Z независимыми случайными величинами?
3. Найдите частные функции плотности S и Z .
4. Как называется закон распределения S ?
5. Как называется закон распределения Z ?
6. Какой закон распределения имеет величина $W = 1 - Z$?

7.4 ууу

7.5 Вася оценивает регрессию y на регрессоры X , включающие константу, а на самом деле все коэффициенты β_j кроме константы равны нулю. Ошибки u_i распределены нормально $\mathcal{N}(0; \sigma^2)$. Какое распределение имеет R^2 ?

8 Максимально правдоподобно

8.1 Величины y_1, y_2, \dots, y_n независимы и экспоненциально распределены с параметром λ . По выборке из 100 наблюдений оказалось, что $\sum y_i = 200$ и $\sum y_i^2 = 900$. Исследователь Андреас хочет проверить гипотезу $H_0: \mathbb{E}(y_i) = 1$ против альтернативной $\mathbb{E}(y_i) \neq 1$.

1. Выпишите логарифмическую функцию правдоподобия $\ell(\lambda)$;

2. Найдите оценку $\hat{\lambda}$ методом максимального правдоподобия в общем виде и для имеющейся выборки;
3. Найдите теоретическую информацию Фишера $I(\lambda)$ для n наблюдений;
4. Выведите формулы для статистик отношения правдоподобия, множителей Лагранжа и Вальда в общем виде;
5. Найдите значения статистик отношения правдоподобия, множителей Лагранжа и Вальда для имеющейся выборки;
6. Проверьте гипотезу H_0 с помощью трёх статистик.

8.2 Рассмотрим модель простой регрессии $y_i = \beta x_i + u_i$, где ошибки u_i независимы и имеют стандартное нормальное распределение, $u_i \sim \mathcal{N}(0; 1)$. По выборке из 100 наблюдений оказалось, что $\sum x_i^2 = 100$, а $\sum y_i x_i = 250$. Исследователь Рамирес хочет проверить $H_0: \beta = 0$.

1. Выпишите логарифмическую функцию правдоподобия $\ell(\beta)$;
2. Найдите оценку $\hat{\beta}$ методом максимального правдоподобия в общем виде и для имеющейся выборки;
3. Найдите теоретическую информацию Фишера $I(\beta)$ для n наблюдений;
4. Выведите формулы для статистик отношения правдоподобия, множителей Лагранжа и Вальда в общем виде;
5. Найдите значения статистик отношения правдоподобия, множителей Лагранжа и Вальда для имеющейся выборки;
6. Проверьте гипотезу H_0 с помощью трёх статистик.

8.3 Исследовательница Геральдина заглядывает n раз в случайные аудитории бывшей шпульно-катушечной фабрики. В каждой аудитории независимо от других идёт семинар по теории вероятностей, эконометрике, микро или макро. Пусть p_1, p_2, p_3 — это вероятности семинаров по теории вероятностей, эконометрике и микро. Вероятность семинара по макро мы отдельным параметром не вводим, так как иначе параметры будут зависимы и нужно будет искать ограниченный экстремум правдоподобия. Пусть y_1, y_2, y_3 — количество попаданий Геральдины на теорию вероятностей, эконометрику и микро.

По выборке из 100 наблюдений оказалось, что $y_1 = 20, y_2 = 30, y_3 = 20$. Геральдина предполагает, что все четыре дисциплины равновероятны.

1. Выпишите логарифмическую функцию правдоподобия $\ell(\mathbb{P})$;
2. Найдите оценку \hat{p} методом максимального правдоподобия в общем виде и для имеющейся выборки;
3. Найдите теоретическую информацию Фишера $I(p)$ для n наблюдений;
4. Найдите явно $I^{-1}(p)$;
5. Выведите формулы для статистик отношения правдоподобия, множителей Лагранжа и Вальда в общем виде;
6. Найдите значения статистик отношения правдоподобия, множителей Лагранжа и Вальда для имеющейся выборки;
7. Проверьте гипотезу H_0 с помощью трёх статистик на уровне значимости 5%.
8. (*) Обобщите формулы трёх статистик на случай произвольного количества дисциплин и произвольной гипотезы $H_0: p = p^0$.

9 Решения

1.1.

1.2. Ответы: 0, 0, 0, $\sum x_i^2$.

1.3.

1. $\hat{\theta} = \sum ((y_i - z_i)(x_i - z_i)) / \sum (x_i - z_i)^2$

1.4. $\hat{\alpha} = 0, \hat{\beta} = 1$

1.5. Рассмотрим регрессию суммы $(y_i + z_i)$ на саму себя. Естественно, в ней

$$\widehat{y_i + z_i} = 0 + 1 \cdot (y_i + z_i).$$

Отсюда получаем, что $\hat{\alpha} + \hat{\gamma} = 0$ и $\hat{\beta} + \hat{\delta} = 1$.

1.6.

Исходя из условия, нужно оценить методом МНК коэффициенты двух следующих моделей:

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

$$y_i = \frac{\gamma}{2} + \frac{\delta}{2} x_i + \frac{1}{2} v_i$$

Заметим, что на минимизацию суммы квадратов остатков коэффициент $1/2$ не влияет, следовательно:

$$\hat{\gamma} = 2\hat{\alpha}, \hat{\delta} = 2\hat{\beta}$$

1.7. Выпишем задачу:

$$\begin{cases} RSS = \sum_{i=1}^n (y_i - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i)^2 \rightarrow \min_{\hat{\beta}_1, \hat{\beta}_2} \\ \hat{\beta}_1 + \hat{\beta}_2 = 1 \end{cases}$$

Можем превратить ее в задачу минимизации функции одного аргумента:

$$RSS = \sum_{i=1}^n (y_i - x_i - \hat{\beta}_2(z_i - x_i))^2 \rightarrow \min_{\hat{\beta}_2}$$

Выпишем условия первого порядка:

$$\frac{\partial RSS}{\partial \hat{\beta}_2} = \sum_{i=1}^n 2(y_i - x_i - \hat{\beta}_2(z_i - x_i))(x_i - z_i) = 0$$

Отсюда:

$$\sum_{i=1}^n (y_i - x_i)(x_i - z_i) + \hat{\beta}_2 \sum_{i=1}^n (z_i - x_i)^2 = 0 \Rightarrow \hat{\beta}_2 = \frac{\sum_{i=1}^n (y_i - x_i)(z_i - x_i)}{\sum_{i=1}^n (z_i - x_i)^2}$$

А $\hat{\beta}_1$ найдется из соотношения $\hat{\beta}_1 + \hat{\beta}_2 = 1$.

1.8. Обозначив вес первого слитка за β_1 , вес второго слитка за β_2 , а показания весов за y_i , получим, что

$$y_1 = \beta_1 + \varepsilon_1, \quad y_2 = \beta_2 + \varepsilon_2, \quad y_3 = \beta_1 + \beta_2 + \varepsilon_3$$

Тогда

$$(300 - \beta_1)^2 + (200 - \beta_2)^2 + (400 - \beta_1 - \beta_2)^2 \rightarrow \min_{\beta_1, \beta_2}$$

$$\hat{\beta}_1 = \frac{800}{3}, \quad \hat{\beta}_2 = \frac{500}{3}$$

1.9. Можем воспользоваться готовой формулой для регрессии на константу:

$$\hat{\beta} = \bar{y} = \frac{10 + 10 + 3}{3} = \frac{23}{3}$$

(можно решить задачу $2(10 - \beta)^2 + (3 - \beta)^2 \rightarrow \min$)

1.10.

1.11.

1.12.

1.13. Пусть \bar{y} — средний y до добавления нового наблюдения, \bar{y}' — после добавления нового наблюдения. Будем считать, что изначально было n наблюдений. Заметим, что

$$\bar{y}' = \frac{(y_1 + \dots + y_n) + y_{n+1}}{n+1} = \frac{n\bar{y} + y_{n+1}}{n+1} = \frac{n}{n+1}\bar{y} + \frac{1}{n+1}y_{n+1}$$

Покажем, что TSS может только увеличиться при добавлении нового наблюдения (остается неизменным при $y_{n+1} = \bar{y}$):

$$TSS' = \sum_{i=1}^{n+1} (y_i - \bar{y}')^2 = \sum_{i=1}^n (y_i - \bar{y} + \bar{y} - \bar{y}')^2 + (y_{n+1} - \bar{y}')^2 =$$

$$= \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \bar{y}')^2 + (y_{n+1} - \bar{y}')^2 = TSS + \frac{n}{n+1}(y_{n+1} - \bar{y})^2$$

Следовательно, $TSS' \geq TSS$.

Также сумма RSS может только вырасти или остаться постоянной при добавлении нового наблюдения. Действительно, новое $(n+1)$ -ое слагаемое в сумме неотрицательно. А сумма n слагаемых минимальна при старых коэффициентах, а не при новых.

ESS и R^2 могут меняться в обе стороны. Например, рассмотрим ситуацию, где точки лежат симметрично относительно некоторой горизонтальной прямой. При этом $ESS = 0$. Добавим наблюдение — ESS вырастет, удалим наблюдение — ESS вырастет.

1.14.

1. R^2 упал до нуля.
2. Да, можно. Если добавить точку далеко слева внизу от исходного набора данных, то наклон линии регрессии будет положительный. Если далеко справа внизу, то отрицательный. Будем двигать точку так, чтобы поймать нулевой наклон прямой. Получим $ESS = 0$.

1.15. На две неизвестных a и b нужно два уравнения. Эти два уравнения — ортогональность вектора остатков плоскости регрессоров. А именно:

$$\begin{cases} \sum_i (y_i - \hat{y}_i) = 0 \\ \sum_i (y_i - \hat{y}_i) \hat{y}_i = 0 \end{cases}$$

В нашем случае

$$\begin{cases} -1 + (6 - a) + (6 - b) = 0 \\ -1 + (6 - a)a + (6 - b)b = 0 \end{cases}$$

Решаем квадратное уравнение и получаем два решения: $a = 4$ и $a = 7$. Итого: $a = 4$, $b = 7$.

1.16. Обе ситуации возможны.

2.1.

2.2.

2.3.

2.4.

3.1.

1. $n = 5$

2. $k = 3$

3. $TSS = 10$

4. $RSS = 2$

5. $\hat{\beta} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = (X'X)^{-1}X'y = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$

6. $R^2 = 1 - \frac{RSS}{TSS} = 0.8$. R^2 высокий, построенная эконометрическая модель хорошо описывает данные

3.2. $\sum y_i^2 = \sum \hat{y}_i^2 + \sum \varepsilon_i^2$, $TSS = ESS + RSS$,

3.3. $s\text{Corr}(\hat{y}, y) = \frac{s\text{Cov}(\hat{y}, y)}{\sqrt{s\text{Var}(\hat{y}) s\text{Var}(y)}}$

$s\text{Corr}(\hat{y}, y)^2 = \frac{(s\text{Cov}(\hat{y}, y))^2}{s\text{Var}(\hat{y}) s\text{Var}(y)}$

$R^2 \cdot TSS / (n-1) \cdot ESS / (n-1) = (s\text{Cov}(\hat{y}, y))^2 = (s\text{Cov}(\hat{y} - \bar{y}, y - \bar{y}))^2$ Отсюда можно понять, что ковариация для двумерного случая равна произведению длин векторов $\hat{y} - \bar{y}$ и $y - \bar{y} = \sqrt{ESS}$ и \sqrt{TSS} на косинус угла между ними ($\sqrt{R^2}$). Геометрически скалярное произведение можно изобразить как произведение длин одного из векторов на проекцию второго вектора на первый. Если будет проецировать $y - \bar{y}$ на $\hat{y} - \bar{y}$, то получим как раз ESS — тот квадрат на рисунке, что уже построен.

$s\text{Cov}(\hat{y}, y) = \sqrt{ESS^2 / (n-1)^2} = ESS / (n-1)$

3.4. Спроецируем единичный столбец на «плоскость», обозначим его $1'$. Делаем проекцию y на «плоскость» и на $1'$. Далее аналогично.

3.5. Проекция y на \hat{y} это \hat{y} , поэтому оценки коэффициентов будут 0 и 1. Оценка дисперсии $\frac{RSS}{(n-2)ESS}$. Нарушены предпосылки теоремы Гаусса-Маркова, например, ошибки новой модели в сумме дают 0, значит коррелированы.

3.6. Либо в регрессию включена константа, либо единичный столбец (тут была опечатка, столбей) можно получить как линейную комбинацию регрессоров, например, включены дамми-переменные для каждого возможного значения качественной переменной.

3.7.

4.1.

4.2. Сферы с центром в начале координат. Проекция имеет хи-квадрат распределение с тремя степенями свободы. Для нахождения максимальной вероятности максимизируем функцию

$$\exp(-R^2/2) \cdot ((R+t)^3 - R^3) \rightarrow \max_R$$

, где R — радиус мякоти, а t — толщина кожуры апельсина. Оставляем только линейную часть по t и затем максимизируем.

Наибольшая вероятность попасть в апельсин радиуса $R = 1$.

5.1.

5.2.

1. $\text{Var}(\varepsilon_1) = \text{Var}(\varepsilon)_{(1,1)} = 4 \cdot I_{(1,1)} = 4$

2. $\text{Var}(\beta_1) = 0$, так как β_1 — детерминированная величина.

3. $\text{Var}(\hat{\beta}_1) = \sigma^2 (X'X)^{-1}_{(1,1)} = 0.5\sigma^2 = 0.5 \cdot 4 = 2$

4. $\widehat{\text{Var}}(\hat{\beta}_1) = \hat{\sigma}^2 (X'X)^{-1}_{(1,1)} = 0.5\hat{\sigma}_{(1,1)}^2 = 0.5 \frac{RSS}{5-3} = 0.25RSS = 0.25y'(I - X(X'X)^{-1}X')y = 0.25 \cdot 1 = 0.25$
 $\hat{\sigma}^2 = \frac{RSS}{n-k} = \frac{1}{2}$.

5. Так как оценки МНК являются несмещёнными, то $\mathbb{E}(\hat{\beta}) = \beta$, значит:

$$\mathbb{E}(\hat{\beta}_1) - \beta_1^2 = \mathbb{E}(\hat{\beta}_1) - (\mathbb{E}(\hat{\beta}_1))^2 = \widehat{\text{Var}}(\hat{\beta}_1) = 0.25$$

6. $\text{Cov}(\hat{\beta}_2, \hat{\beta}_3) = \sigma^2 (X'X)^{-1}_{(2,3)} = 4 \cdot (-\frac{1}{2}) = -2$

$$7. \widehat{\text{Cov}}(\hat{\beta}_2, \hat{\beta}_3) = \widehat{\text{Var}}(\hat{\beta})_{(2,3)} = \hat{\sigma}^2 (X'X)^{-1}_{(2,3)} = \frac{1}{2} \cdot \left(-\frac{1}{2}\right) = -\frac{1}{4}$$

$$8. \text{Var}(\hat{\beta}_2 - \hat{\beta}_3) = \text{Var}(\hat{\beta}_2) + \text{Var}(\hat{\beta}_3) + 2 \text{Cov}(\hat{\beta}_2, \hat{\beta}_3) = \sigma^2((X'X)^{-1}_{(2,2)} + (X'X)^{-1}_{(3,3)} + 2(X'X)^{-1}_{(2,3)}) = 4(1 + 1.5 + 2 \cdot (-0.5)) = 6$$

$$9. \widehat{\text{Var}}(\hat{\beta}_2 - \hat{\beta}_3) = \widehat{\text{Var}}(\hat{\beta}_2) + \widehat{\text{Var}}(\hat{\beta}_3) + 2\widehat{\text{Cov}}(\hat{\beta}_2, \hat{\beta}_3) = \hat{\sigma}^2((X'X)^{-1}_{(2,2)} + (X'X)^{-1}_{(3,3)} + 2(X'X)^{-1}_{(2,3)}) = \frac{1}{2} \cdot 1.5 = 0.75$$

$$10. \text{Var}(\beta_2 - \beta_3) = 0$$

$$11. \text{Corr}(\hat{\beta}_2, \hat{\beta}_3) = \frac{\text{Cov}(\hat{\beta}_2, \hat{\beta}_3)}{\sqrt{\text{Var}(\hat{\beta}_2) \text{Var}(\hat{\beta}_3)}} = \frac{-2}{\sqrt{4 \cdot 6}} = -\frac{\sqrt{6}}{6}$$

$$12. \widehat{\text{Corr}}(\beta_2, \beta_3) = \frac{\widehat{\text{Cov}}(\hat{\beta}_2, \hat{\beta}_3)}{\sqrt{\widehat{\text{Var}}(\hat{\beta}_2) \widehat{\text{Var}}(\hat{\beta}_3)}} = \frac{-\frac{1}{4}}{\sqrt{\frac{1}{2} \cdot \frac{3}{4}}} = -\frac{\sqrt{6}}{6}$$

$$13. (n - k) \frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k}.$$

$$\mathbb{E} \left((n - k) \frac{\hat{\sigma}^2}{\sigma^2} \right) = n - k$$

$$\mathbb{E} \left(\frac{\hat{\sigma}^2}{\sigma^2} \right) = 1$$

$$\mathbb{E}(\hat{\sigma}^2) = 2$$

$$14. \hat{\sigma}^2 = \frac{RSS}{n-k} = \frac{1}{2}$$

7.1.

7.2.

7.3.

7.4.

7.5.

$$\mathbf{8.1.} \ell'(\lambda) = \frac{n}{\lambda} - \sum y_i; \quad I(\lambda) = n/\lambda^2$$

8.2.

8.3.

10 Источники мудрости