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, $\sum (y_i - \hat{y}_i)^2 \to \min$.

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 $z_i = \gamma + \delta x_i + \upsilon_i$, 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:

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

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 $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$ по первой сотне наблюдений. Петечка оценил модель $\hat{y}_i = \hat{\gamma}_1 + \hat{\gamma}_2 x_i$ по второй сотне наблюдений. Машенька оценила модель $\hat{y}_i = \hat{\varphi}_1 + \hat{\varphi}_2 x_i$ по всем наблюдениям.

1. Возможно ли, что $\hat{\beta}_2 > 0$, $\hat{\gamma}_2 > 0$, но $\hat{\varphi}_2 < 0$?
2. Возможно ли, что $\hat{\beta}_1 > 0$, $\hat{\gamma}_1 > 0$, но $\hat{\varphi}_1 < 0$?
3. Возможно ли одновременное выполнение всех упомянутых условий?
4. Возможно ли одновременное выполнение всех упомянутых условий, если в каждой сотне наблюдений $\sum x_i > 0$?

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.
- переменная[1] 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, $\hat{y}_i = \hat{\alpha}_1 + \hat{\alpha}_2 f_i$;
B: y на константу и m, $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 m_i$;
C: y на f и m без константы, $\hat{y}_i = \hat{\gamma}_1 f_i + \hat{\gamma}_2 m_i$;
D: y на константу, f и m, $\hat{y}_i = \hat{\delta}_1 + \hat{\delta}_2 f_i + \hat{\delta}_3 m_i$;

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, $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$, а затем модель 2, $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \beta_4 w_i + \varepsilon$. 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?

[1]Это 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?

1.17 Бесстрашный исследователь Ипполит оценил парную регрессию. При этом оказалось, что сумма все $x_i > 0$ и обе оценки коэффициентов $\hat{\beta}_1$ и $\hat{\beta}_2$ также положительны.

1. Возможно ли, что среди $\hat{y}_i$ есть отрицательные? Среди y_i есть отрицательные?
2. Возможно ли, что сумма $\sum \hat{y}_i$ отрицательна? Сумма $\sum y_i$ отрицательна?
3. Как изменятся ответы, если известно, что $\sum x_i > 0$?

2. Дифференциал — просто няшка!

Минитеория.
Дифференциал для матриц подчиняется правилам:

1. $d(A+B) = dA + dB$;
2. Если A — матрица констант, то $dA = 0$;
3. $d(AB) = dA \cdot B + A \cdot dB$. Если A — матрица констант, то $d(AB) = AdB$;

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 $dx = 0.1$?
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 $\operatorname{tr}(F'dF)$ в явном виде без матриц.

2.2 Пусть 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. $d(r'Ar/r'r)$.

Упростите ответ для случая симметричной матрицы A.

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 гребневой регрессии (ridge regression) 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 $\hat{\beta}$.

2.5 Исследователь Никодим поймал 100 морских ежей и у каждого измерил длину, a_i, и вес b_i. Вектор измерений, относящихся к одному ежу обозначим $y_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix}$. Никодим считает, что ежи независимы друг от друга, а длина и вес имеют совместное нормальное распределение

$$y_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix} \sim \mathcal{N}(\mu, C)$$

1. Выпишите логарифмическую функцию правдоподобия, $\ell(\mu, C)$;
2. Предполагая ковариационную матрицу известной, $C = \begin{pmatrix} 9 & 4 \\ 4 & 6 \end{pmatrix}$, найдите $d\ell$ и оценку $\hat{\mu}$ методом максимального правдоподобия.
3. Предполагая, вектор ожиданий известным, $\mu = \begin{pmatrix} 10 \\ 5 \end{pmatrix}$, найдите $d\ell$ и оценку $\hat{C}$ методом максимального правдоподобия.
4. Найдите $d\ell(\mu, C)$ и оценки для параметров μ и C, в случае, когда μ и C неизвестны.

3. МНК в матрицах и геометрия!

3.1 Рассмотрим регрессию $\hat{y}_i = \hat{\beta}_1 z_i + \hat{\beta}_2 x_i$. Все исходные данные поместим в матрицу X и вектор y:

$$X = \begin{pmatrix} z_1 & x_1 \\ \vdots & \vdots \\ z_n & x_n \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

1. Выпишите явно матрицы X', $X'y$, $X'X$, $y'X$, $y'z$ и укажите их размер.
2. Выпишите условия первого порядка для оценок $\hat{\beta}_1$ и $\hat{\beta}_2$ по методу наименьших квадратов.
3. Запишите эти же условия в виде линейной системы

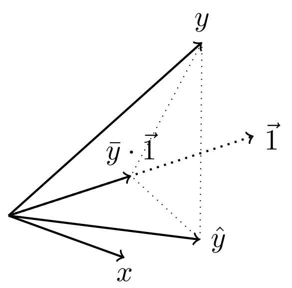
$$\begin{cases} \hat{\beta}_1 \cdot \ldots + \hat{\beta}_2 \cdot \ldots = \ldots \\ \hat{\beta}_1 \cdot \ldots + \hat{\beta}_2 \cdot \ldots = \ldots \end{cases}$$

4. Как упростится данная система для регрессии $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$?
5. Запишите систему условий первого порядка с помощью матрицы X и вектора y;

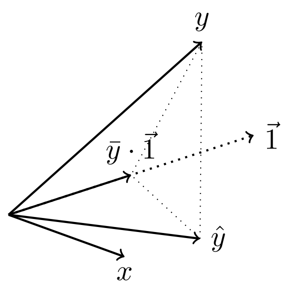
3.2 Пусть $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \varepsilon_i$ — регрессионная модель, где $X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, $y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$, $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix}$, 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.

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



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



3.5 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 $\hat{\varepsilon}$.

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

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

3.8 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?

3.9 Какие матрицы являются положительно полуопределёнными?

1. $X'X$;
2. XX';
3. $H = X(X'X)^{-1}X'$;
4. $I - H$;
5. $A'(I - H)A$;
6. $A'A - G(G'A^{-1}(A')^{-1}G)^{-1}G'$

4. Распределения, связанные с проецированием

4.1 Рассмотрим пространство $\mathbb{R}^3$ и два подпространства в нём, $W = \{(x_1, x_2, x_3) \mid x_1 + 2x_2 + x_3 = 0\}$ и $V = \mathrm{Lin}((1, 2, 3)^T)$.

1. Найдите $\dim V$, $\dim W$, $\dim V \cap W$, $\dim V^{\perp}$, $\dim W^{\perp}$.
2. Найдите проекцию произвольного вектора u на V, W, $V \cap W$, $V^{\perp}$, $W^{\perp}$. Найдите квадрат длины каждой проекции.
3. Как распределён квадрат длины проекции в каждом случае, если дополнительно известно, что вектор u имеет многомерное стандартное нормальное распределение?

4.2 Рассмотрим пространство $\mathbb{R}^n$, где $n > 2$, и два подпространства в нём, $V = \mathrm{Lin}((1, 1, \ldots, 1)^T)$ и $W = \{x \mid x_1 = x_2 + x_3 + \ldots + x_n\}$.

1. Найдите $\dim V$, $\dim W$, $\dim V \cap W$, $\dim V^{\perp}$, $\dim W^{\perp}$, $\dim V \cap W^{\perp}$, $\dim V^{\perp} \cap W$.
2. Найдите проекцию произвольного вектора u на каждое упомянутое подпространство. Найдите квадрат длины каждой проекции.
3. Как распределён квадрат длины проекции в каждом случае, если дополнительно известно, что вектор u имеет многомерное стандартное нормальное распределение?

4.3 Храбрая исследовательница Евлампия оценивает модель множественной регрессии $\hat{y} = X\hat{\beta}$. Однако на самом деле β_0, и $y = u$, где u_i независимы и нормальны $u_i \sim \mathcal{N}(0; \sigma^2)$.

Какое распределение в регрессии Евлампии имеют $\bar{y}$, $\sum y_i^2/\sigma^2$, $\sum \hat{y}_i^2/\sigma^2$, $n\bar{y}^2/\sigma^2$, TSS/σ^2, RSS/σ^2, ESS/σ^2?

4.4 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.

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

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

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

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 $\mathrm{Var}(\varepsilon_1)$, $\mathrm{Var}(\beta_1)$, $\mathrm{Var}(\hat{\beta}_1)$, $\widehat{\mathrm{Var}}(\hat{\beta}_1)$, $\mathbb{E}(\hat{\beta}_1^2) - \beta_1^2$;
3. Предполагая нормальность ошибок, постройте 95% доверительный интервал для β_2.
4. Предполагая нормальность ошибок, проверьте гипотезу H_0: $\beta_2 = 0$;
5. Найдите $\mathrm{Cov}(\hat{\beta}_2, \hat{\beta}_3)$, $\widehat{\mathrm{Cov}}(\hat{\beta}_2, \hat{\beta}_3)$, $\mathrm{Var}(\hat{\beta}_2 - \hat{\beta}_3)$, $\widehat{\mathrm{Var}}(\hat{\beta}_2 - \hat{\beta}_3)$;
6. Найдите $\mathrm{Var}(\beta_2 - \beta_3)$, $\mathrm{Corr}(\hat{\beta}_2, \hat{\beta}_3)$, $\widehat{\mathrm{Corr}}(\hat{\beta}_2, \hat{\beta}_3)$;
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 устали со скоростью d экспериментов в минуту. Каждый эксперимент независимо от других может окончится 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 $\mathbb{P}(X = k)$, $\mathbb{E}(X)$, $\mathrm{Var}(X)$. Как называется закон распределения X?
2. Найдите $\mathbb{P}(Y = k)$, $\mathbb{E}(Y)$, $\mathrm{Var}(Y)$. Как называется закон распределения Y?
3. Найдите $\mathbb{P}(Z \leqslant t)$, $\mathbb{E}(Z)$, $\mathrm{Var}(Z)$.

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

4. Выразите p через λ и d.
5. Найдите предел $\mathbb{P}(Z \leqslant t)$. Является ли предельная функция $\mathbb{P}(Z \leqslant t)$ непрерывной? Какая в предельном случае получается функция плотности у величины Z? Как называется этот закон распределения Z? Чему равен предел $\mathbb{E}(Z)$ и $\mathrm{Var}(Z)$?
6. Найдите предел вероятности $\mathbb{P}(X = k)$ и пределы $\mathbb{E}(X)$ и $\mathrm{Var}(X)$. 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 $S = Y_1 + \ldots + Y_n$, $X_1 = Y_1/S$, $X_2 = Y_2/S$, ..., $X_{n-1} = Y_{n-1}/S$.

1. Выпишите совместную функцию плотности Y_1, ..., Y_n;
2. Найдите совместную функцию плотности X_1, X_2, X_3, ..., X_{n-1}, S.
3. Зависит ли величина S и вектор X_1, X_2, ..., 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, ..., X_{n-1}.

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

6. Найдите совместную функцию плотности Z_1, Z_2, ..., Z_{n-1}, S.
7. Зависимы ли величины Z_1, Z_2, ..., 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 $Gamma(k, \lambda)$. 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 набопытов.

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 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. Блок

8.1 Найдите матрицу M и укажите размеры всех блоков

1. Блок C имеет размер $p \times p$, блок F — размер $q \times q$

$$M = \begin{pmatrix} A & B \end{pmatrix} \cdot \begin{pmatrix} C & D \\ E & F \end{pmatrix}$$

2. Блок C имеет размер $p \times p$, блок F — размер $q \times q$

$$M = \begin{pmatrix} C & D \\ E & F \end{pmatrix} \cdot \begin{pmatrix} A \\ B \end{pmatrix}$$

3. Блок C имеет размер $p \times p$, блок B — размер $q \times q$

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^T$$

8.2 Найдите обратную матрицу M^{-1} для каждого из случаев

1. Блоки $A_{p\times p}$ и $B_{q\times q}$ обратимы,

$$M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

2. Блоки $A_{p\times p}$ и $B_{q\times q}$ обратимы,

$$M = \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$$

3. Блоки $A_{p\times p}$ и $B_{q\times q}$ обратимы,

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

4. Блоки $A_{p\times p}$ и $B_{q\times q}$ обратимы,

$$M = \begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$$

8.3 Блоки $A_{p\times p}$ и $B_{q\times q}$ обратимы, матрица M имеет вид

$$M = \begin{pmatrix} A & C \\ D & B \end{pmatrix}$$

Рассмотрим обратную матрицу M^{-1}

$$M^{-1} = \begin{pmatrix} X & Z \\ Y & W \end{pmatrix}$$

1. Найдите блок X с помощью процедуры Гаусса;
2. Найдите блок X решив систему двух уравнений на блоки X и Y;
3. Докажите тождество Вудберри

$$(A - CB^{-1}D)^{-1} = A^{-1} + A^{-1}C(B - DA^{-1}C)^{-1}DA^{-1}$$

9. Максимально правдоподобно

9.1 Величины y_1, y_2, ..., y_n независимы и экспоненциально распределены с параметром λ. По выборке из 100 наблюдений оказалось, что $\sum y_i = 200$. 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.

9.2 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 $\sum x_i^2 = 100$, $\sum y_i^2 = 900$, 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.

9.3 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 выборки 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. (*) Обобщиет формулы трёх статистик на случай произвольного количества дисциплин и произвольной гипотезы H_0: $p = p^0$.

9.4 Логарифмическая функция правдоподобия имеет вид

$$\ell(\theta) = a - \frac{1}{2}(\theta - h(y))'Q(\theta - h(y)),$$

где Q — постоянная симметричная матрица, а $h(y)$ — функция от выборки. Вектор параметров θ состоит из двух блоков, а матрица Q — из четырёх блоков

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, \quad Q = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$$

Настырный исследователь Никанор хочет проверить гипотезу H_0: $\theta_1 = \theta_1^0$ про часть параметров, входящих в вектор θ;

1. Найдите неограниченную оценку метода максимального правдоподобия $\hat{\theta}^{UR}$;
2. Найдите ограниченную оценку метода максимального правдоподобия $\hat{\theta}^{R}$;
3. Выведите формулу для LR статистики;
4. Выведите формулу для LM статистики;
5. Выведите формулу для W статистики;
6. Какие из указанных формул равны?

9.5 Рассмотрим модель множественной регрессии $y = X\beta + u$, где регрессоры детерминистические, ошибки имеют многомерное нормальное распределение, а ковариационная матрица $\mathrm{Var}(u)$ единичная. Разобьём вектор коэффициентов β на две части

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

1. Докажите, что логарифмическая функция правдоподобия представима в виде

$$\ell(\theta) = a - \frac{1}{2}(\theta - h(y))'Q(\theta - h(y)),$$

2. Явно найдите матрицу Q и функцию $h(y)$;
3. Выведите формулу для LR, LM и теста Вальда для проверки гипотезы H_0: $\beta_1 = \beta_1^0$;
4. Как найденная формула отличается от обычной F статистики?
5. Как найденная формула упрощается для случая проверки гипотезы о незначимости регрессии в целом?

9.6 Рассмотрим модель множественной регрессии $y = X\beta + u$, где регрессоры детерминистические, ошибки имеют многомерное нормальное распределение, а ковариационная матрица $\mathrm{Var}(u) = \sigma^2 I$. Разобьём вектор коэффициентов β на две части

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

1. Докажите, что логарифмическая функция правдоподобия представима в виде

$$\ell(\theta) = a + (\theta - h(y))'Q(\theta - h(y)),$$

2. Явно найдите матрицу Q и функцию $h(y)$;
3. Выведите формулы для LR, LM и теста Вальда для проверки гипотезы H_0: $\beta_1 = \beta_1^0$;
4. Как найденные формулы отличается от обычной F статистики?
5. Как найденные формулы упрощается для случая проверки гипотезы о незначимости регрессии в целом?

10. Гетероскедастичность

10.1 Имеется три наблюдения

x_i	1	2	2
y_i	1	2	3

Эконометр Антоний хочет оценить зависимость $y_i = \beta x_i + u_i$.

1. Найдите оценку $\hat{\beta}$ с помощью МНК;
2. Найдите стандартную ошибку $se(\hat{\beta})$ предполагая гомоскедастичность;
3. Найдите робастные к гетероскедастичности стандартные ошибки $se_{HC0}(\hat{\beta})$ и $se_{HC3}(\hat{\beta})$;
4. Найдите эффективную оценку $\hat{\beta}$, если дополнительно известно, что $\mathrm{Var}(u_i|x_i) = \sigma^2(3x_i - 2)$;
5. Найдите эффективную оценку $\hat{\beta}$, если дополнительно известно, что

$$\mathrm{Var}(u|X) = \begin{pmatrix} 4\sigma^2 & -\sigma^2 & 0 \\ -\sigma^2 & 9\sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}$$

10.2 Известно, что после деления каждого уравнения регрессии $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$ на x_i^2 гетероскедастичность ошибок была устранена. Какой вид имела дисперсия ошибок, $\mathrm{Var}(\varepsilon_i)$?

10.3 Для линейной регрессии $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$ была выполнена сортировка наблюдений по возрастанию переменной x. Исходная модель оценивалась по разным частям выборки:

Выборка	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	RSS
$i = 1, \ldots, 30$	1.21	1.89	2.74	48.69
$i = 1, \ldots, 11$	1.39	2.27	2.36	10.28
$i = 12, \ldots, 19$	0.75	2.23	3.19	5.31
$i = 20, \ldots, 30$	1.56	1.06	2.29	14.51

Известно, что ошибки в модели являются независимыми нормальными случайными величинами с нулевым математическим ожиданием. Протестируйте ошибки на гетероскедастичность на уровне значимости 5%.

10.4 Рассмотрим линейную регрессию $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$ по 50 наблюдениям. При оценивании с помощью МНК были получены результаты: $\hat{\beta}_1 = 1.21$, $\hat{\beta}_2 = 1.11$, $\hat{\beta}_3 = 3.15$, $R^2 = 0.72$.

Оценена также вспомогательная регрессия: $\hat{\varepsilon}_i^2 = \delta_1 + \delta_2 x_i + \delta_3 z_i + \delta_4 x_i^2 + \delta_5 z_i^2 + \delta_6 x_i z_i + u_i$. Результаты оценивания следующие: $\hat{\delta}_1 = 1.50$, $\hat{\delta}_2 = -2.18$, $\hat{\delta}_3 = 0.23$, $\hat{\delta}_4 = 1.87$, $\hat{\delta}_5 = -0.56$, $\hat{\delta}_6 = -0.09$, $R^2_{aux} = 0.36$

Известно, что ошибки в модели являются независимыми нормальными случайными величинами с нулевым математическим ожиданием. Протестируйте ошибки на гетероскедастичность на уровне значимости 5%.

10.5 Найдите число коэффициентов во вспомогательной регрессии, необходимой для выполнения теста Уайта, если число коэффициентов в исходной регрессии равно k, включая свободный член.

10.6 Рассмотрим модель регрессии $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$, в которой ошибки ε_i независимы и имеют нормальное распределение $N(0, \sigma^2)$. Для $n = 200$ наблюдений найдите

1. вероятность того, что статистика Уайта окажется больше 10;
2. ожидаемое значение статистики Уайта;
3. дисперсию статистики Уайта.

10.7 Рассматривается модель $y_t = \beta_1 + \varepsilon_t$, где ошибки ε_t — независимые случайные величины с $\mathbb{E}(\varepsilon_t) = 0$ и $\mathrm{Var}(\varepsilon_t) = t$. Найдите наиболее эффективную оценку неизвестного параметра β_1 в классе линейных по y и несмещённых оценок.

10.8 Эконометр Антоний исследует зависимость надоя коров в литрах в год, y_i, от дамми-переменной x_i, отвечающей за прослушивание коровами ежедневно Девятой симфонии, $y_i = \beta_1 + \beta_2 x_i + u_i$. Антоний раздобыл следующие данные:

Подвыборка	Размер	$\sum y_i$	$\sum y_i^2$
$x_i = 0$	$n_0 = 100$	200	4000
$x_i = 1$	$n_1 = 100$	300	5000

1. Найдите МНК-оценки β_1 и β_2;
2. Постройте 95%-ый доверительный интервал для β_2 предполагая гомоскедастичность u_i;
3. Найдите робастную к гетероскедастичности оценку $\widehat{\mathrm{Var}}_{HC0}(\hat{\beta})$;
4. Найдите робастную к гетероскедастичности оценку $\widehat{\mathrm{Var}}_{HC3}(\hat{\beta})$;
5. Постройте 95%-ый доверительный интервал для β_2 с помощью скорректированной $se_{HC0}(\hat{\beta}_2)$;
6. Дополнительно предположив, что $\mathrm{Var}(u_i|x_i) = \sigma^2(1 + 3x_i)$, найдите эффективную оценку $\hat{\beta}_2$ и постройте доверительный интервал для неё.

10.9 Эконометресса Прасковья использует традиционную оценку ковариационной матрицы, а эконометресса Мелони — оценку Уайта.

Какие оценки дисперсии $\hat{\beta}_1$ и формулы для t-статистики получат Прасковья и Мелони в модели $y_i = \beta_1 + u_i$?

10.10 Эконометресса Прасковья использует традиционную оценку ковариационной матрицы, а эконометресса Мелони — оценку Уайта.

Обе эконометрессы оценивают модель $y_i = \beta_1 + \beta_2 d_i + u_i$, где d_i — дамми-переменная, равна 0 или 1. Дамми-переменная делит выборку на две части. Обозначим количество наблюдений в «нулевой» части как n_0, среднее — как $\bar{y}_0$, и общую сумму квадратов — как TSS_0. Аналогичные величины для «единичной» части выборки — n_1, $\bar{y}_1$ и TSS_1. И для всей выборки — n, $\bar{y}$, TSS.

1. Найдите оценки $\hat{\beta}_1$ и $\hat{\beta}_2$.
2. Найдите оценки $\widehat{\mathrm{Var}}(\hat{\beta}_1)$ и $\widehat{\mathrm{Var}}_W(\hat{\beta}_1)$. Верно ли, что $\widehat{\mathrm{Var}}_W(\hat{\beta}_1) \geqslant \widehat{\mathrm{Var}}(\hat{\beta}_1)$?
3. Найдите оценки $\widehat{\mathrm{Var}}(\hat{\beta}_2)$ и $\widehat{\mathrm{Var}}_W(\hat{\beta}_2)$. Верно ли, что $\widehat{\mathrm{Var}}_W(\hat{\beta}_2) \geqslant \widehat{\mathrm{Var}}(\hat{\beta}_2)$?
4. Найдите оценки $\widehat{\mathrm{Cov}}(\hat{\beta}_1, \hat{\beta}_2)$ и $\widehat{\mathrm{Cov}}_W(\hat{\beta}_1, \hat{\beta}_2)$.
5. Найдите оценки $\widehat{\mathrm{Var}}(\hat{\beta}_1 + \hat{\beta}_2)$ и $\widehat{\mathrm{Var}}_W(\hat{\beta}_1 + \hat{\beta}_2)$.

10.11 В модели $y_i = \beta x_i + \varepsilon_i$ предполагается гетероскедастичность вида $\mathrm{Var}(\varepsilon_i) = \exp(\gamma_1 + \gamma_2 x_i)$ и нормальность ошибок.

1. Сформулируйте гипотезу о гомоскедастичности с помощью коэффициентов.
2. Выведите в явном виде оценку максимального правдоподобия при предположении о гомоскедастичности.
3. Выпишите условия первого порядка для оценки максимального правдоподобия без предположения о гомоскедастичности.
4. Выведите в явном виде формулу для LM теста множителей Лагранжа.

10.12 Для регрессии $y = X\beta + \varepsilon$ с $\mathbb{E}(\varepsilon) = 0$, $\mathrm{Var}(\varepsilon) = \Sigma \neq \sigma^2 I$, оцененной с помощью обобщённого метода наименьших квадратов, найдите ковариационную матрицу $\mathrm{Cov}(\hat{\beta}_{GLS}, \varepsilon)$

10.13 В оригинальном тесте Бройша-Пагана на гетероскедастичность два шага. Сначала строится основная регрессия y_i на некоторые регрессоры и получаются остатки $\hat{u}_i$. На втором шаге строится регрессия квадрата остатков $\hat{u}_i$ на переменные, от которых потенциально зависит условная дисперсия $\mathrm{Var}(u_i|Z)$. Статистика Бройша-Пагана считается как $BP = ESS/2$, где ESS — объяснённая сумма квадратов регрессии второго шага. Оригинальный тест Уайта считается как $W = nR^2$, где R^2 — коэффициент детерминации регрессии второго шага.

1. Найдите отношение $\frac{nR^2}{ESS/2}$;
2. Найдите предел по вероятности $\operatorname{plim} \frac{nR^2}{ESS/2}$;
3. Какое распределение имеют статистики BP и W?
4. Какой вид имеет статистика множителей Лагранжа?

распотрошить статью BP на задачу, статья о похожести BP и W, отдельно Коэнкера про студентизированную версию

11. Логит, пробит и хоббит!

11.1 Бандерлог из Лога оценил логистическую регрессию по четырём наблюдениям и одному признаку с константой, получил $b_i = \hat{\mathbb{P}}(y_i = 1|x_i)$, но потерял последнее наблюдение:

y_i	b_i
1	0.7
-1	0.2
-1	0.3
?	?

1. Выпишите функцию правдоподобия для задачи логистической регрессии.
2. Выпишите условие первого порядка по коэффициенту перед константой.
3. Помогите бандерлогу восстановить пропущенные значения!

11.2 Рассмотрим логистическую функцию $\Lambda(w) = e^w/(1+e^w)$.

1. Как связаны между собой $\Lambda(w)$ и $\Lambda(-w)$?
2. Как связаны между собой $\Lambda'(w)$ и $\Lambda'(-w)$?
3. Постройте графики функций $\Lambda(w)$ и $\Lambda'(w)$.
4. Найдите $\Lambda(0)$, $\Lambda'(0)$, $\ln \Lambda(0)$.
5. Найдите обратную функцию $\Lambda^{-1}(p)$.
6. Как связаны между собой $\frac{d \ln \Lambda(w)}{dw}$ и $\Lambda(-w)$?
7. Как связаны между собой $\frac{d \ln \Lambda(-w)}{dw}$ и $\Lambda(w)$?
8. Разложите $h(\beta_1, \beta_2) = \ln \Lambda(y_i(\beta_1 + \beta_2 x_i))$ в ряд Тейлора до второго порядка в окрестности точки $\beta_1 = 0, \beta_2 = 0$.

11.3 Винни-Пух знает, что мёд бывает правильный, $honey_i = 1$, и неправильный, $honey_i = 0$. Пчёлы также бывают правильные, $bee_i = 1$, и неправильные, $bee_i = 0$. По 100 своим попыткам добыть мёд Винни-Пух составил таблицу сопряженности:

	$honey_i = 1$	$honey_i = 0$
$bee_i = 1$	12	36
$bee_i = 0$	32	20

Винни-Пух использует логистическую регрессию с константой для прогнозирования правильности мёда с помощью правильности пчёл.

1. Какие оценки коэффициентов получит Винни-Пух?
2. Какой прогноз вероятности правильности мёда при встрече с неправильными пчёлами даёт логистическая модель? Как это число можно посчитать без рассчитывания коэффициентов?

11.4 Винни-Пух оценил логистическую регрессию для прогнозирования правильности мёда от высоты дерева (м) x_i и удалённости от дома (км) z_i: $\ln odds_i = 2 + 0.3x_i - 0.5z_i$.

1. Оцените вероятность того, что $y_i = 1$ для $x = 15$, $z = 3.5$.
2. Оцените предельный эффект увеличения x на единицу на вероятность того, что $y_i = 1$ для $x = 15$, $z = 3.5$.
3. При каком значении x предельный эффект увеличения x на единицу в точке $z = 3.5$ будет максимальным?

11.5 Придумайте такие три наблюдения для парной логистической регрессии, чтобы все x_i были разными, не все y_i были одинаковые, а оценки логит-модели не существовали.

Какое решение задачи этой проблемы разумно предложить при большом количестве наблюдений?

11.6 При оценке логит модели $\mathbb{P}(y_i = 1) = \Lambda(\beta_1 + \beta_2 x_i)$ по 500 наблюдениям оказалось, что $\hat{\beta}_1 = 0.7$ и $\hat{\beta}_2 = 3$. Оценка ковариационной матрицы коэффициентов имеет вид

$$\begin{pmatrix} 0.04 & 0.01 \\ 0.01 & 0.09 \end{pmatrix}$$

1. Проверьте гипотезу о незначимости коэффициента $\hat{\beta}_2$.
2. Найдите предельный эффект роста x_i на вероятность $\mathbb{P}(y_i = 1)$ при $x_i = -0.5$.
3. Найдите максимальный предельный эффект роста x_i на вероятность $\mathbb{P}(y_i = 1)$.
4. Постройте точечный прогноз вероятности $\mathbb{P}(y_i = 1)$ если $x_i = -0.5$.
5. Найдите стандартную ошибку построенного прогноза.
6. Постройте 95%-ый доверительный интервал для $\mathbb{P}(y_i = 1)$ двумя способами (через преобразование интервала для $\hat{y}_i^*$ и через дельта-метод).

11.7 Почему в пробит-модели предполагается, что $\varepsilon_i \sim \mathcal{N}(0; 1)$, а не $\varepsilon_i \sim \mathcal{N}(0; \sigma^2)$ как в линейной регрессии?

11.8 Что произойдёт с оценками логит-модели $\mathbb{P}(y_i = 1) = F(\beta_1 + \beta_2 x_i)$, их стандартными ошибками, если у зависимой переменной поменять 0 и 1 местами?

11.9 Исследователь Матвей оценил логит-модель по 10 тысячам наблюдений. $\hat{\mathbb{P}}(y_i = 1) = F(-0.5 + 1.2x_i)$. Переменная x_i — бинарная, 4 тысячи единиц и 6 тысяч нулей.

1. Сколько наблюдений с $y_i = 1$?
2. Сколько наблюдений с $y_i = 1$ и $x_i = 0$?
3. Сколько наблюдений с $y_i = 0$ и $x_i = 1$?

12. Эндогенность

12.1 Величины x_i, z_i и u_i имеют совместное распределение, задаваемое табличкой:

x_i	0	1	0	1
z_i	0	1	0	0
u_i	-1	-1	1	1
Вероятность	0.2	0.3	0.3	0.2

Рассмотрим модель $y_i = \beta x_i + u_i$.

1. Найдите $\operatorname{plim} \hat{\beta}_{OLS}$;
2. Найдите $\operatorname{plim} \hat{\beta}_{IV}$, если в качестве инструмента для x_i используется z_i;

12.2 Рассмотрим три вектора: y, x и z. Проведем гипер-плоскость ортогональную z через конец вектора y. Эта гипер-плоскость пересекает прямую порождаемую вектором x в точке $\hat{\beta}_{IV} x$.

Исходя из данного геометрического определения $\hat{\beta}_{IV}$:

1. Выведите алгоритм двухшагового МНК;
2. Выведите явную формулу для $\hat{\beta}_{IV}$;
3. Докажите, что оценка $\hat{\beta}_{IV}$ показывает, насколько в среднем растёт y при таком росте z, при котором x в среднем растёт на единицу.

сказать смысл IV попроще? на две фразы?

12.3 Исследовательница Мишель строит оценку $\hat{\beta}_{IV}$ в регрессии y на x с инструментом z. Исследовательница Аграфена строит обычную МНК оценку в регрессии $\hat{y} = \hat{\beta}_x x + \hat{\beta}_w w$.

1. Выразите w через x, z и y так, чтобы оценка $\hat{\beta}_{IV}$ Мишель и оценка $\hat{\beta}_x$ Аграфены совпали.
2. Сформулируйте ещё одну интерпретацию оценки $\hat{\beta}_{IV}$;

12.4 Величины x_i, z_i и u_i имеют совместное распределение, параметры которого известны:

$$\operatorname{Var}\left(\begin{pmatrix} x_i \\ z_i \\ u_i \end{pmatrix}\right) = \begin{pmatrix} 5 & 1 & -1 \\ 1 & 9 & 0 \\ -1 & 0 & 4 \end{pmatrix}; \quad \mathbb{E}\left(\begin{pmatrix} x_i \\ z_i \\ u_i \end{pmatrix}\right) = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}$$

Наблюдения с разными номерами независимы и одинаково распределены.

Рассмотрим модель $y_i = \beta_1 + \beta_2 x_i + u_i$.

1. Найдите $\operatorname{plim} \hat{\beta}_2^{LS}$ и $\operatorname{plim} \hat{\beta}_1^{LS}$; Являются ли оценки состоятельными?
2. Храбрый исследователь Афанасий использует двухшаговый МНК. На первом шаге он строит регрессию x_i на константу и z_i, $\hat{x}_i = \hat{\gamma}_1 + \hat{\gamma}_2 z_i$. А на втором регрессию $\hat{y}_i = \hat{\beta}_1^{IV} + \hat{\beta}_2^{IV} \hat{x}_i$. Найдите $\operatorname{plim} \hat{\beta}_2^{IV}$ и $\operatorname{plim} \hat{\beta}_1^{IV}$; Являются ли оценки состоятельными?
3. Как изменятся $\operatorname{plim} \hat{\beta}_2^{IV}$ и $\operatorname{plim} \hat{\beta}_1^{IV}$, если Афанасий забудет включить константу на первом шаге?

12.5 Приведите примеры дискретных случайных величин ε и x, таких, что

1. $\mathbb{E}(\varepsilon) = 0$, $\mathbb{E}(\varepsilon \mid x) = 0$, но величины зависимы. Чему в этом случае равно $\operatorname{Cov}(\varepsilon, x)$?
2. $\mathbb{E}(\varepsilon) = 0$, $\operatorname{Cov}(\varepsilon, x) = 0$, но $\mathbb{E}(\varepsilon \mid x) \neq 0$. Зависимы ли эти случайные величины?

12.6 Эконометресса Агнесса хочет оценить модель $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, но, к сожалению, величина x_i ненаблюдаема. Вместо неё доступна величина x_i^*. Величина x_i^* содержит ошибку измерения, $x_i^* = x_i + a_i$. Известно, что $\operatorname{Var}(x_i) = 9$, $\operatorname{Var}(a_i) = 4$, $\operatorname{Var}(\varepsilon_i) = 1$. Величины x_i, a_i и ε_i независимы.

Агнесса оценивает регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i^*$ с помощью МНК.

1. Найдите $\operatorname{plim} \hat{\beta}_2$.
2. Являются ли оценки, получаемые Агнессой, состоятельными?

12.7 Эконометресса Анжелла хочет оценить модель $y_i = \beta_1 + \beta_2 x_i + \beta_3 w_i + \varepsilon_i$, но, к сожалению, величина w_i ненаблюдаема. Известно, что $\operatorname{Var}(x_i) = 9$, $\operatorname{Var}(w_i) = 4$, $\operatorname{Var}(\varepsilon_i) = 1$ и $\operatorname{Cov}(x_i, w_i) = -2$. Случайная составляющая не коррелированна с регрессорами.

За неимением w_i Анжелла оценивает регрессию $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$ с помощью МНК.

1. Найдите plim $\hat{\beta}_2$.
2. Являются ли оценки, получаемые Агнессой, состоятельными?

12.8 Наблюдения представляют собой случайную выборку. Зависимые переменные y_{t1} и y_{t2} находятся из системы:

$$\begin{cases} y_{t1} = \beta_{11} + \beta_{12} x_t + \varepsilon_{t1} \\ y_{t2} = \beta_{21} + \beta_{22} z_t + \beta_{23} y_{t1} + \varepsilon_{t2} \end{cases},$$

где вектор ошибок ε_t имеет совместное нормальное распределение

$$\varepsilon_t \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

Эконометресса Анжела оценивает с помощью МНК первое уравнение, а эконометресса Эвридика — второе.

1. Найдите пределы по вероятности получаемых ими оценок.
2. Будут ли оценки состоятельными?

12.9 Эконометресса Венера оценивает регрессию

$$y_i = \beta_1 + \beta_2 x_i + u_i,$$

А на самом деле $y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2 + w_i$, причём $x_i \sim \mathcal{N}(0; 1)$ и ошибки $w_i \sim \mathcal{N}(0; \sigma^2)$. Все остальные предпосылки теоремы Гаусса-Маркова выполнены.

1. Будут ли оценки $\hat{\beta}_1$ и $\hat{\beta}_2$, получаемые Венерой, несмещёнными?
2. Будут ли оценки $\hat{\beta}_1$ и $\hat{\beta}_2$, получаемые Венерой, состоятельными?

12.10 Вектора (x_1, u_1), (x_2, u_2), ...независимы и одинаково распределены. Также известно, что $x_i \sim \mathcal{N}(10; 9)$ и $\mathbb{E}(u_i | x_i) = 0$.

Найдите plim $\left(\frac{1}{n} x' x\right)^{-1}$, plim $\frac{1}{n} x' u$ и plim$(x'x)^{-1} x' u$

12.11 Возможно ли, что $\mathbb{E}(x_i | u_i) = 0$ и $\mathbb{E}(u_i | x_i) = 0$, но при этом x_i и u_i зависимы?

12.12 В некотором интституте на некотором факультете задумали провести эксперимент: раздать студентам учебники разных цветов случайным образом и посмотреть на итоговую успеваемость по эконометрике. Учебники есть двух цветов: зелёненькие и красненькие. Поэтому модель имеет вид:

$$y_i = \beta_1 + \beta_2 green_i + u_i$$

Здесь y_i — результат по эконометрике, $green_i$ — дамми-переменная на зелёный учебник и u_i — прочие характеристики студента. Зелёные и красненькие учебники планировалось раздавать равновероятно. Однако библиотекарь всё прошляпил и разрешил студентам самим выбирать учебник, какой понравится. В результате вместо переменной $green_i$ получилась переменная $green_i^*$. Известно, что $\mathbb{E}(green_i^*) = \alpha$ и $\mathrm{Cov}(green_i^*, u_i) = \gamma$.

Де-факто оценивалась модель

$$\hat{y}_i = \hat{\theta}_1 + \hat{\theta}_2 green_i^*$$

1. Найдите plim $\hat{\theta}_1$, plim $\hat{\theta}_2$.
2. Найдите $\mathbb{E}\hat{\theta}_1$, $\mathbb{E}\hat{\theta}_2$.

12.13 Придумайте такие случайные величины x_1, x_2, u_1, что x_1 и x_2 независимы и одинаково распределены, причём $\mathbb{E}(u_1 | x_1) = 0$, а $\mathbb{E}(u_1 | x_1, x_2) \neq 0$.

12.14 Рассмотрим классическую парную регрессию со стохастическим регрессором. Всего три наблюдения:

y_1	x_1
y_2	x_2
y_3	x_3

1. Соедините линиями независимые случайные величины.
2. Соедините линиями одинаково распределённые случайные величины.

13. GMM

Прочесть про обобщённый метод моментов можно, например, [Zso10].

13.1 Исследователь Максимилиан оценивает параметр θ с помощью двух моментных условий, $\mathbb{E}(y_i) = 2\theta$ и $\mathbb{E}(1/y_i) = \theta$. С трудом Максимилиан нашёл 200 наблюдений и оказалось, что $\sum 1/y_i = 1.5$. Сначала Максимилиан оценил θ с помощью простого метода моментов и первого моментного условия, получил $\hat{\theta} = 1$.

Затем Максимилиан решил применить обобщённый метод моментов, чтобы учесть оба момента. В процессе получения GMM оценки Максимилиан обнаружил, что $\mathrm{Cov}(y_i, 1/y_i) = -\theta^2$, $\mathrm{Var}(1/y_i) = 9\theta^2$, $\mathbb{E}(y_i^2) = 20\theta^2$.

1. Найдите ковариационную матрицу моментных условий, $\mathrm{Var}(g)$;
2. Найдите оптимальную теоретическую матрицу весов W;
3. Оцените матрицу весов $\hat{W}$;
4. Найдите оценку GMM с использованием найденной оценки матрицы весов;

13.2 Величины X_i равномерны на отрезке $[-a; 3a]$ и независимы. Есть несколько наблюдений, $X_1 = 0.5$, $X_2 = 0.7$, $X_3 = -0.1$.

1. Найдите $\mathbb{E}(X_i)$ и $\mathbb{E}(|X_i|)$.
2. Постройте оценку метода моментов, используя $\mathbb{E}(X_i)$.
3. Постройте оценку метода моментов, используя $\mathbb{E}(|X_i|)$.
4. Постройте оценку обобщёного метода моментов используя моменты $\mathbb{E}(X_i)$, $\mathbb{E}(|X_i|)$ и взвешивающую матрицу.
$$W = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$
5. Найдите оптимальную теоретическую взвешивающую матрицу для обобщённого метода моментов
6. Постройте двухшаговую оценку обобщённого метода моментов, начав со взвешивающей матрицы W
7. С помощью полученных оценок постройте 95%-ый доверительный интервал для неизвестного параметра a

13.3 Винни-Пух и Пятачок оценивают неизвестный параметр правильности пчёл θ. Когда Винни-Пух проводит очередное измерение параметра правильности, он получает значение X_i нормально распределенное вокруг неизвестного параметра, $X_i \sim \mathcal{N}(\theta, 1)$. Когда Пятачок проводит измерение параметра правильности, он получает значение Y_i, также нормально распределенное вокруг θ, но имеющее большую дисперсию, $Y_i \sim \mathcal{N}(\theta, 4)$. Различные измерения независимы между собой.

1. Найдите $\mathbb{E}(X_i)$ и постройте соответствующую оценку метода моментов.
2. Найдите $\mathbb{E}(Y_i)$ и постройте соответствующую оценку метода моментов.

3. Используя два указанных момента найдите обобщённую оценку метода моментов для взвешивающей матрицы
$$W = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}.$$
4. Найдите оптимальную взвешивающую матрицу W.

13.4 Начинающий футболист делает независимые удары по воротам. С вероятностью θ он попадает левее ворот, с вероятностью 2θ — правее ворот и попадает с вероятностью $1 - 3\theta$. Из n ударов он попал N_L раз левее ворот и N_R раз — правее.

1. Найдите $\mathbb{E}(N_L)$ и постройте соответствующую оценку θ методом моментов.
2. Найдите $\mathbb{E}(N_R)$ и постройте соответствующую оценку θ методом моментов.
3. Используя два указанных момента постройте оценку обобщённого метода моментов со взвешивающей матрицей
$$W = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix}.$$
4. Найдите оптимальную теоретическую взвешивающую матрицу.
5. Для каждой из найденных оценок постройте 95%-ый доверительный интервал, если $N_L = 10$, $N_R = 30$, $n = 200$.

13.5 Можно ли получить МНК-оценки в классической задаче регрессии как оценки обобщённого метода моментов? Можно ли получить оценки метода максимального правдоподобия как оценки обобщённого метода моментов?

13.6 Равшан и Джамшут измеряют длину θ оставшегося куска рулона обоев много раз. Измерения Равшана, X_i, распределены нормально, $\mathcal{N}(2\theta, \theta^2 + 100)$. Измерения Джамшута, Y_i, также распределены нормально $\mathcal{N}(\theta, \theta^2 + 10)$. Поскольку Равшан и Джамшут спорят друг с другом, их измерения зависимы, $\mathrm{Cov}(X_i, Y_i) = -1$.

Оказалось, что по 200 (?проверить?) измерениям $\sum X_i = 300$, $\sum Y_i = 100$.

Насяльника хочет измерить параметр θ.

1. Запишите два моментных условия на $\mathbb{E}(X_i)$ и $\mathbb{E}(Y_i)$ в виде
$$\begin{cases} \mathbb{E}(g_1(X_i, \theta)) = 0 \\ \mathbb{E}(g_2(Y_i, \theta)) = 0 \end{cases}$$
2. Найдите ковариационную матрицу $\mathrm{Var}(g)$ и теоретическую оптимальную матрицу весов W для обобщённого метода моментов;
3. Найдите оценку параметра θ методом моментов с единичной матрицей весов;
4. Найдите оценку параметра θ методом моментов, предварительно оценив оптимальную матрицу с помощью $\hat{\theta}$ из предыдущего пункта;

14. Большая сила о-малых

15. Решения

1.1.

1. (a) $\hat{\beta} = 13/9$

   c) $\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$

2. (a) $\hat{\beta}_1 = -1, \hat{\beta}_2 = 2$

   c) $\hat{\beta}_2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$

1.2.

1. 0
2. 0
3. 0
4. $\sum x_i^2$

1.3.

1. $\hat{\theta} = \frac{\sum y_i(1+x_i)}{\sum (1+x_i)^2}$
2. $\hat{\theta} = \frac{\sum (y_i-1)x_i}{\sum x_i^2}$
3. $\hat{\theta} = \frac{\sum (y_i/x_i^2)}{\sum (1/x_i^3)}$
4. $\hat{\theta} = \sum\left((y_i - z_i)(x_i - z_i)\right) / \sum\left(x_i - z_i\right)^2$

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:

$$\begin{cases} RSS = \sum\limits_{i=1}^{n}(y_i - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i)^2 \to \min\limits_{\hat{\beta}_1, \hat{\beta}_2} \\ \hat{\beta}_1 + \hat{\beta}_2 = 1 \end{cases}$$

Можем превратить ее в задачу минимизации функции одного аргумента:

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

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:

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

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

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

Тогда

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

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

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 $2(10 - \beta)^2 + (3 - \beta)^2 \to \min$)

1.10.

1. Да.
2. Да.
3. Да.
4. Нет. Из условия первого порядка для первой выборки следует, что $\sum_A y_i = \hat{\beta}_1 + \hat{\beta}_2 \sum x_i$. Значит $\sum_A y_i > 0$. Аналогично, $\sum_B y_i > 0$, но $\sum y_i < 0$.

1.11.

1.12. Поскольку значения y остались теми же, $TSS_1 = TSS_2$.

Добавление ещё одного регрессора не уменьшит точность оценки, то есть $RSS_2 \leqslant RSS_1$, $ESS_2 \geqslant ESS_1$.

Тогда и коэффициент детерминации $R^2 = ESS/TSS$ не уменьшится, то есть $R_2^2 \geqslant R_1^2$.

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

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

Покажем, что TSS может только увеличится 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, $TSS' \geqslant TSS$.

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.

1.17.

1. нет, да

2. нет, нет

3. да, да, нет, нет

2.1.

1. $f'(x) = 2x + 3$, $df = 2xdx + 3dx$, $df = 1.3$

2. $df = 2x_1 dx_1 + 3dx_1 \cdot x_2^3 + 3x_1 \cdot 3x_2^2 dx_2$, $df = 1.7$

2.2.

1. $A(dR)B$
2. $2r'dr$
3. $r'(A'+A)dr$
4. $R^{-1} \cdot dR \cdot R^{-1}$
5. $-\sin(r'r) \cdot 2r'dr$
6. $\frac{r'(A'+A)dr \cdot r'r - r'Ar2r'dr}{(r'r)^2}$

2.3.

1. $dQ(\hat{\beta}) = 2(y - X\hat{\beta})^T(-X)d\hat{\beta}$, $d^2Q(\hat{\beta}) = 2d\hat{\beta}^T X^T X d\hat{\beta}$
2. $dQ(\hat{\beta}) = 0$
3. $\hat{\beta} = (X^TX)^{-1}X^Ty$

2.4.

1. $dQ(\hat{\beta}) = -2((y - X\hat{\beta})^T X + \lambda\hat{\beta}^T)d\hat{\beta}$, $d^2Q(\hat{\beta}) = 2d\hat{\beta}^T(X^TX - \lambda I)d\hat{\beta}$
2. $dQ(\hat{\beta}) = 0$
3. $\hat{\beta} = (X^TX - \lambda I)^{-1}X^Ty$

2.5.

1. $\hat{\mu} = \sum y_i/n$
2.
3.

3.1.

$$\begin{cases} \hat{\beta}_1 \sum z_i^2 + \hat{\beta}_2 \sum x_i z_i = \sum z_i y_i \\ \hat{\beta}_1 \sum x_i z_i + \hat{\beta}_2 \sum x_i^2 = \sum x_i y_i \end{cases}$$

$$\begin{cases} \hat{\beta}_1 n + \hat{\beta}_2 \sum x_i = \sum y_i \\ \hat{\beta}_1 \sum x_i + \hat{\beta}_2 \sum x_i^2 = \sum x_i y_i \end{cases}$$

$$X'X\hat{\beta} = X'y$$

3.2.

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

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

3.4. $\mathrm{sCorr}(\hat{y}, y) = \frac{\mathrm{sCov}(\hat{y},y)}{\sqrt{\mathrm{sVar}(\hat{y})\,\mathrm{sVar}\,(y)}}$
$\mathrm{sCorr}(\hat{y}, y)^2 = \frac{(\mathrm{sCov}(\hat{y},y))^2}{\mathrm{sVar}(\hat{y})\,\mathrm{sVar}\,(y)}$
$R^2 \cdot TSS/(n-1) \cdot ESS/(n-1) = (\mathrm{sCov}(\hat{y}, y))^2 = (\mathrm{sCov}(\hat{y} - \bar{y}, y - \bar{y}))^2$ Отсюда можно понять, что ковариация для двухмерного 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 $y - \bar{y}\vec{1}$ на $\hat{y} - \bar{y}\vec{1}$, то получим как раз ESS — тот квадрат на рисунке, что уже построен.
$\mathrm{sCov}(\hat{y}, y) = \sqrt{ESS^2/(n-1)^2} = ESS/(n-1)$

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

3.6. 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.

3.7. 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.

3.8. Для удобства центрируем мысленно все переменные. Это не меняет ни корреляций, ни выборочных дисперсий, ни угловых коэффициентов в регрессиях. В регрессиях при этом оценка коэффициента при константе превращается в ноль, но какое нам до неё дело? :) Зато при нулевом среднем выборочные корреляции превратились в косинус угла между векторами:

$$\mathrm{sCorr}(x, y) = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}} = \cos(x, y)$$

И при нулевом среднем выборочная дисперсия — это длина вектора с точностью до умножения на $(n-1)$:

$$\mathrm{sVar}(x) = \frac{\sum x_i^2}{n-1} = \frac{||x||^2}{n-1}$$

Начать можно с геометрического смысла оценок МНК:

$$\begin{cases} 1.4 = \frac{||y||}{||x||} \cos(x, y) \\ 0.6 = \frac{||x||}{||y||} \cos(x, y) \end{cases}$$

Отсюда находим $||y||/||x|| = \hat{\sigma}_y/\hat{\sigma}_x$ и $\cos(x, y) = \mathrm{sCorr}(x, y)$
Дальше номер решается, например, по теореме косинусов.

1. $\mathrm{sCorr}(x, y) = \sqrt{0.84}$, $\mathrm{sCorr}(y, z) = \frac{\sqrt{70}}{10}$, $\mathrm{sCorr}(x, z) = -\frac{\sqrt{30}}{10}$

2. $\frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} = \frac{3}{7}$, $\frac{\hat{\sigma}_z^2}{\hat{\sigma}_y^2} = \frac{8}{35}$, $\frac{\hat{\sigma}_x^2}{\hat{\sigma}_z^2} = \frac{15}{8}$

3.9. все :)

4.1. $\dim V = 1$, $\dim W = 2$, $\dim V \cap W = 0$, $\dim V^{\perp} = 2$, $\dim W^{\perp} = 1$. Эти же числа и будут степенями свободы хи-квадрат распределения.

4.2. $\dim V = 1$, $\dim W = n-1$, $\dim V \cap W = 0$, $\dim V^{\perp} = n-1$, $\dim W^{\perp} = 1$.

4.3.

4.4.

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

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

, 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. $\mathrm{Var}(\varepsilon_1) = \mathrm{Var}(\varepsilon)_{(1,1)} = 4 \cdot I_{(1,1)} = 4$
2. $\mathrm{Var}(\beta_1) = 0$, так как β_1 — детерминированная величина.
3. $\mathrm{Var}(\hat\beta_1) = \sigma^2 (X'X)^{-1}_{(1,1)} = 0.5\sigma^2 = 0.5 \cdot 4 = 2$
4. $\widehat{\mathrm{Var}}(\hat\beta_1) = \hat\sigma^2 (X'X)^{-1}_{(1,1)} = 0.5\hat\sigma^2_{(1,1)} = 0.5\frac{RSS}{5-3} = 0.25RSS = 0.25y'(I - X(X'X)^{-1}X')y = 0.25 \cdot 1 = 0.25$
   $\hat\sigma^2 = \frac{RSS}{n-k} = \frac{1}{2}$.
5. Так как оценки МНК являются несмещёнными, то $\mathbb{E}(\hat\beta) = \beta$, значит:
   $$\mathbb{E}(\hat\beta_1) - \beta_1^2 = \mathbb{E}(\hat\beta_1) - (\mathbb{E}(\hat\beta_1))^2 = \widehat{\mathrm{Var}}(\hat\beta_1) = 0.25$$
6. $\mathrm{Cov}(\hat\beta_2, \hat\beta_3) = \sigma^2 (X'X)^{-1}_{(2,3)} = 4 \cdot \left(-\frac{1}{2}\right) = -2$
7. $\widehat{\mathrm{Cov}}(\hat\beta_2, \hat\beta_3) = \widehat{\mathrm{Var}}(\hat\beta)_{(2,3)} = \hat\sigma^2 (X'X)^{-1}_{(2,3)} = \frac{1}{2} \cdot \left(-\frac{1}{2}\right) = -\frac{1}{4}$
8. $\mathrm{Var}(\hat\beta_2 - \hat\beta_3) = \mathrm{Var}(\hat\beta_2) + \mathrm{Var}(\hat\beta_3) + 2\,\mathrm{Cov}(\hat\beta_2, \hat\beta_3) = \sigma^2((X'X)^{-1}_{(2,2)} + (X'X)^{-1}_{(3,3)} + 2(X'X)^{-1}_{(2,3)} = 4(1 + 1.5 + 2 \cdot (-0.5)) = 6$
9. $\widehat{\mathrm{Var}}(\hat\beta_2 - \hat\beta_3) = \widehat{\mathrm{Var}}(\hat\beta_2) + \widehat{\mathrm{Var}}(\hat\beta_3) + 2\widehat{\mathrm{Cov}}(\hat\beta_2, \hat\beta_3) = \hat\sigma^2((X'X)^{-1}_{(2,2)} + (X'X)^{-1}_{(3,3)} + 2(X'X)^{-1}_{(2,3)} = \frac{1}{2} \cdot 1.5 = 0.75$
10. $\mathrm{Var}(\beta_2 - \beta_3) = 0$
11. $\mathrm{Corr}(\hat\beta_2, \hat\beta_3) = \frac{\mathrm{Cov}(\hat\beta_2, \hat\beta_3)}{\sqrt{\mathrm{Var}(\hat\beta_2)\,\mathrm{Var}(\hat\beta_3)}} = \frac{-2}{\sqrt{4 \cdot 6}} = -\frac{\sqrt{6}}{6}$
12. $\widehat{\mathrm{Corr}}(\beta_2, \beta_3) = \frac{\widehat{\mathrm{Cov}}(\hat\beta_2, \hat\beta_3)}{\sqrt{\widehat{\mathrm{Var}}(\hat\beta_2)\widehat{\mathrm{Var}}(\hat\beta_3)}} = \frac{-\frac{1}{4}}{\sqrt{\frac{1}{2} \cdot \frac{3}{4}}} = -\frac{\sqrt{6}}{6}$

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

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

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

7.1.

7.2.

7.3.

7.4.

8.1.

1. $M = \begin{pmatrix} AC+BE & AD+BF \end{pmatrix}$

   Предположим, что матрицы A и B имеют m строк. Тогда размерность блока $AC+BE - m \times p$, $AD+BF - m \times q$.

2. $M = \begin{pmatrix} CA+DB \\ EA+FB \end{pmatrix}$

   Предположим, что матрицы A и B имеют n столбцов. Тогда размерность блока $CA+DB - p \times n$, $EA+FB - q \times n$.

3. $M = \begin{pmatrix} A^T & C^T \\ B^T & D^T \end{pmatrix}$

   Блок A^T имеет размерность $p \times q$, $B^T - q \times q$, $C^T - p \times p$, $D^T - q \times p$.

8.2.

1. $M^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}$

2. $M^{-1} = \begin{pmatrix} 0 & B^{-1} \\ A^{-1} & 0 \end{pmatrix}$

3. $M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}CB^{-1} \\ 0 & B^{-1} \end{pmatrix}$

4. $M^{-1} = \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}CA^{-1} & B^{-1} \end{pmatrix}$

8.3.

1.

$$\left(\begin{array}{cc|cc} A & C & I & 0 \\ D & B & 0 & I \end{array}\right) \sim \left(\begin{array}{cc|cc} I & A^{-1}C & A^{-1} & 0 \\ D & B & 0 & I \end{array}\right) \sim \left(\begin{array}{cc|cc} I & A^{-1}C & A^{-1} & 0 \\ 0 & B - DA^{-1}C & -DA^{-1} & I \end{array}\right) \sim$$

$$\left(\begin{array}{cc|cc} I & A^{-1}C & A^{-1} & 0 \\ 0 & I & -(B - DA^{-1}C)^{-1}DA^{-1} & (B - DA^{-1}C)^{-1} \end{array}\right) \sim$$

$$\left(\begin{array}{cc|cc} I & 0 & A^{-1} + A^{-1}C(B - DA^{-1}C)^{-1}DA^{-1} & -A^{-1}C(B - DA^{-1}C)^{-1} \\ 0 & I & -(B - DA^{-1}C)^{-1}DA^{-1} & (B - DA^{-1}C)^{-1} \end{array}\right)$$

То есть $X = A^{-1} + A^{-1}C(B - DA^{-1}C)^{-1}DA^{-1}$.

2. Из равенства

$$\begin{pmatrix} A & C \\ D & B \end{pmatrix} \begin{pmatrix} X & Z \\ Y & W \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

получаем систему:

$$\begin{cases} AX + CY = I \\ DX + BY = 0 \end{cases} \Rightarrow \begin{cases} X = A^{-1}(I - CY) \\ DX + BY = 0 \end{cases}$$

Подставляя первое уравнение во второе, получим:

$$DA^{-1}(I - CY) = -BY \Rightarrow DA^{-1} = (DA^{-1}C - B)Y \Rightarrow I = (C - AD^{-1}B)Y \Rightarrow Y = (C - AD^{-1}B)^{-1}$$

И окончательно из второго уравнения:

$$DX = -B(C - AD^{-1}B)^{-1} \Rightarrow -(C - AD^{-1}B)B^{-1}DX = I \Rightarrow X = (A - CB^{-1}D)^{-1}$$

3.

$$\begin{aligned} &(A - CB^{-1}D)(A^{-1} + A^{-1}C(B - DA^{-1}C)^{-1}DA^{-1}) = \\ &\quad I - CB^{-1}DA^{-1} + (C - CB^{-1}DA^{-1}C)(B - DA^{-1}C)^{-1}DA^{-1} = \\ &\quad I - CB^{-1}DA^{-1} + CB^{-1}(B - DA^{-1}C)(B - DA^{-1}C)^{-1}DA^{-1} = \\ &\quad\quad I - CB^{-1}DA^{-1} + CB^{-1}DA^{-1} = I \end{aligned}$$

9.1.

1. $\ell = n \ln \lambda - \lambda \sum y_i$
2. $\hat{\lambda} = \frac{\sum y_i}{n} = \frac{1}{2}$
3. $I(\lambda) = \frac{n}{\lambda^2}$
4. $LR = 2(n \ln \frac{n}{\sum y_i} - n - n \ln \lambda_R + \lambda_R \sum y_i)$

   $LM = \left(\frac{n}{\lambda} - \sum y_i\right)^2 \frac{\lambda^2}{n}$

   $W = \left(\frac{\sum y_i}{n} - \lambda_R\right)^2 \frac{n}{\lambda^2}$
5. $LR \approx 61.37$, $LM = W = 100$
6. $\chi^2_{1,0.95} = 3.84$, основная гипотеза отвергается.

9.2.

1. $\ell = n \ln \sqrt{2\pi} - \frac{1}{2} \sum (y_i - \beta x_i)^2$
2. $\hat{\beta}_{ML} = \frac{\sum y_i x_i}{\sum x_i^2} = 2.5$
3. $I(\beta) = \sum x_i^2$
4. $LR = -\sum (y_i - \hat{\beta}_{ML} x_i)^2 + \sum (y_i - \beta_R x_i)^2$
   $LM = (\sum (y_i x_i - \beta_R x_i^2))^2 \cdot \frac{1}{\sum x_i^2}$
   $W = (\hat{\beta}_{ML} - \beta_R)^2 \sum x_i^2$
5. $LR = LM = W = 625$
6. $\chi^2_{1,0.95} = 3.84$, основная гипотеза отвергается.

9.3.

1. $\ell = const + y_1 \ln p_1 + y_2 \ln p_2 + y_3 \ln p_3 + (n - y_1 - y_2 - y_3) \ln(1 - y_1 - y_2 - y_3)$
2. $\hat{p} = \begin{pmatrix} y_1/n \\ y_2/n \\ y_3/n \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.3 \\ 0.2 \end{pmatrix}$
3. $I(p) = \begin{pmatrix} \frac{n}{p_1} + \frac{n}{1-p_1-p_2-p_3} & \frac{n}{1-p_1-p_2-p_3} & \frac{n}{1-p_1-p_2-p_3} \\ \frac{n}{1-p_1-p_2-p_3} & \frac{n}{p_2} + \frac{n}{1-p_1-p_2-p_3} & \frac{n}{1-p_1-p_2-p_3} \\ \frac{n}{1-p_1-p_2-p_3} & \frac{n}{1-p_1-p_2-p_3} & \frac{n}{p_3} + \frac{n}{1-p_1-p_2-p_3} \end{pmatrix}$
4. $I^{-1}(p) = \begin{pmatrix} \frac{p_1(1-p_1)}{n} & -\frac{p_1 p_2}{n} & -\frac{p_1 p_3}{n} \\ -\frac{p_1 p_2}{n} & \frac{p_2(1-p_2)}{n} & -\frac{p_2 p_3}{n} \\ -\frac{p_1 p_3}{n} & -\frac{p_2 p_3}{n} & \frac{p_3(1-p_3)}{n} \end{pmatrix}$

9.4.

1. $\hat{\theta}^{UR} = h(y)$
2. $\theta^R = \begin{pmatrix} \theta_1^0 \\ h_2(y) - C^{-1} B^T (\theta_1^0 - h_1(y)) \end{pmatrix}$

3–6. $LR = LM = W = (\theta_1^0 - h_1(y))^T (A - BC^{-1}B^T)(\theta_1^0 - h_1(y))$

9.5.

2. $Q = \frac{1}{\sigma^2} X^T X$, $h(y) = (X^T X)^{-1} X^T y$

9.6.

10.1.

1. $\hat{\beta}_{OLS} = 11/9$
2. $se(\hat{\beta}) = \sqrt{5/162}$
3. $se_{HC0}(\hat{\beta}) = \sqrt{168}/81$, $se_{HC3}(\hat{\beta}) = \sqrt{2649}/180$
4. $\hat{\beta} = 7/6$

10.2. $\text{Var}(\varepsilon_i) = cx_i^4$

10.3. Протестируем гетероскедастичность ошибок при помощи теста Голдфельда- Квандта. $H_0 : \text{Var}(\varepsilon_i) = \sigma^2$, $H_a : \text{Var}(\varepsilon_i) = f(x_i)$

1. Тестовая статистика $GQ = \frac{RSS_3/(n_3-k)}{RSS_1/(n_1-k)}$, где $n_1 = 11$ — число наблюдений в первой подгруппе, $n_3 = 11$ — число наблюдений в последней подгруппе, $k = 3$ — число факторов в модели, считая единичный столбец.
2. Распределение тестовой статистики при верной H_0: $GQ \sim F_{n_3-k,n_1-k}$
3. Наблюдаемое значение $GQ_{obs} = 1.41$
4. Область, в которой H_0 не отвергается: $GQ \in [0; 3.44]$
5. Статистический вывод: поскольку $GQ_{obs} \in [0; 3.44]$, то на основании имеющихся наблюдений на уровне значимости 5% основная гипотеза H_0 не может быть отвергнута. Таким образом, тест Голдфельда-Квандта не выявил гетероскедастичность.

10.4. Протестируем гетероскедастичность ошибок при помощи теста Уайта. $H_0 : \text{Var}(\varepsilon_i) = \sigma^2$, $H_a : \text{Var}(\varepsilon_i) = \delta_1 + \delta_2 x_i + \delta_3 z_i + \delta_4 x_i^2 + \delta_5 z_i^2 + \delta_6 x_i z_i$.

1. Тестовая статистика $W = n \cdot R^2_{aux}$, где n — число наблюдений, R^2_{aux} — коэффициент детерминации для вспомогательной регрессии.
2. Распределение тестовой статистики при верной H_0: $W \sim \chi^2_{k_{aux}-1}$, где $k_{aux} = 6$ — число регрессоров во вспомогательной регрессии, считая константу.
3. Наблюдаемое значение тестовой статистики: $W_{obs} = 18$
4. Область, в которой H_0 не отвергается: $W \in [0; W_{crit}] = [0; 11.07]$
5. Статистический вывод: поскольку $W_{obs} \notin [0; 11.07]$, то на основании имеющихся наблюдений на уровне значимости 5% основная гипотеза H_0 отвергается. Таким образом, тест Уайта выявил гетероскедастичность.

10.5. $k(k+1)/2$

10.6. 0.0752, 5, 10

10.7.

10.8.

10.9. Одинаковые.

10.10.

$$\widehat{\text{Var}}(\hat{\beta}_1) = TSS \frac{1}{n_0} \frac{1}{n-2}$$

$$\widehat{\text{Var}}(\hat{\beta}_1) = TSS_0 \frac{n}{n_0} \frac{1}{n_0} \frac{1}{n-2}$$

10.11. В предположении о гомоскедастичности, $\gamma_2 = 0$, оценка правдоподобия совпадает с МНК-оценкой, значит $\hat{\beta} = \sum y_i x_i / \sum x_i^2$. И $\hat{s}_i^2 = RSS/n$, значит $\hat{\gamma}_1 = \ln(RSS/n)$.

10.12.

$$\begin{aligned}\operatorname{Cov}(\hat{\beta}_{GLS}, \varepsilon) = \operatorname{Cov}\left((X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y, \varepsilon\right) = \\ = \operatorname{Cov}\left((X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\varepsilon, \varepsilon\right) = \\ = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\operatorname{Cov}(\varepsilon, \varepsilon) = \\ = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\Sigma = (X'\Sigma^{-1}X)^{-1}X'\end{aligned}$$

10.13.

11.1. $\hat{\mathbb{P}}(y_i = 1|x_i) = \frac{1}{1+\exp(-\beta_1-\beta_2 x_i)}$

1. $loss(\beta_1, \beta_2) = -\sum_{i=1}^{l}\left([y_i = 1]\ln\frac{1}{1+\exp(-\beta_1-\beta_2 x_i)} + [y_i = -1]\ln\left(1 - \frac{1}{1+\exp(-\beta_1-\beta_2 x_i)}\right)\right)$
2. $\frac{\partial loss}{\partial \beta_1} = -\sum_{i=1}^{l}\left([y_i = 1]\cdot\frac{1}{1+\exp(\beta_1+\beta_2 x_i)} + [y_i = -1]\cdot(-1)\cdot\frac{1}{1+\exp(-\beta_1-\beta_2 x_i)}\right)$
3. $y_4 = 1, x_4 = 0.8$

11.2.

1. $\Lambda(w) + \Lambda(-w) = 1$
2. $\Lambda'(w) = -\Lambda'(-w)$
3.
4. $\Lambda(0) = 0.5, \Lambda'(0) = 0.25, \ln\Lambda(0) = -\ln 2$
5. $\Lambda^{-1}(p) = \ln\frac{p}{1-p}$
6. $\frac{d\ln\Lambda(w)}{dw} = \Lambda(-w)$
7. $\frac{d\ln\Lambda(-w)}{dw} = -\Lambda(w)$
8.

11.3.

1. Выпишем аппроксимацию функции потерь:

   $$loss(\beta_1, \beta_2) \approx 100\ln 2 + 6\beta_1 + 12\beta_2 + \frac{1}{2}(25\beta_1^2 + 2\cdot 12\beta_1\beta_2 + 12\beta_2^2) \to \min_{\beta_1, \beta_2}$$

   Взяв производные по β_1 и β_2, получим $\hat{\beta}_1 = \frac{6}{13}$, $\hat{\beta}_2 = -\frac{19}{13}$.

2. $\hat{P}(honey_i = 1|bee_i = 0) = \frac{1}{1+\exp(-6/13)} \approx 0.615$.

   Это же число можно было получить из таблицы: $\frac{32}{32+20} \approx 0.61$.

11.4. Предельный эффект максимален при максимальной производной $\Lambda'(\hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 z)$, то есть при $\hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 z = 0$.

11.5. Ввести штраф в жанре LASSO или гребневой регрессии.

11.6. $z = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)} = \frac{3}{0.3} = 10$, H_0 отвергается. Предельный эффект равен $\hat{\beta}_2 \Lambda'(-0.8) \approx 0.642$. Для нахождения $se(\hat{\mathbb{P}})$ найдём линейную аппроксимацию для $\Lambda(\hat{\beta}_1 + \hat{\beta}_2 x)$ в окрестности точки $\hat{\beta}_1 = 0.7$, $\hat{\beta}_2 = 3$. Получаем

$$\Lambda(\hat{\beta}_1 + \hat{\beta}_2 x) \approx \Lambda(\beta_1 + \beta_2 x) + \Lambda'(\beta_1 + \beta_2 x)(\hat{\beta}_1 - \beta_1) + \Lambda'(\beta_1 + \beta_2 x) x (\hat{\beta}_2 - \beta_2).$$

11.7. Если в пробит-уравнении ненаблюдаемой переменной домножить все коэффициенты и стандартную ошибку на произвольную константу, то в результате получится ровно та же модель. Следовательно, модель с $\varepsilon_i \sim \mathcal{N}(0; \sigma^2)$ не идентифицируема. Поэтому надо взять какое-то нормировочное условие. Можно взять, например, $\beta_2 = 42$, но традиционно берут $\varepsilon_i \sim \mathcal{N}(0; 1)$.

11.8.

11.9.

12.1.

12.2.

12.3.

12.4.

12.5.

12.6.

12.7.

12.8.

12.9. Оценки будут несмещёнными и состоятельными. Вызвано это тем, что $u_i = \beta_3 x_i^2 + w_i$ некоррелировано с x_i.

12.10. $\operatorname{plim} \left(\frac{1}{n} x'x\right)^{-1} = 109^{-1}$, $\operatorname{plim} \frac{1}{n} x'u = 0$ и $\operatorname{plim}(x'x)^{-1} x'u = 0$

12.11. Да, например, равномерное распределение (u_i, x_i) на круге или на окружности. Или равновероятное на восьми точках, $(\pm 1, \pm 1)$, $(\pm 2, \pm 2)$.

12.12.

12.13. Например, можно взять $u_1 = x_2$, и все величины $\mathcal{N}(0; 1)$.

12.14. Одинаково распределены: $y_1 \sim y_2 \sim y_3$, $x_1 \sim x_2 \sim x_3$. Независимы переменные с разными номерами.

13.1.

$$\mathrm{Var}(g) = \mathrm{Var}\left(\begin{pmatrix} y_i - 2\theta \\ 1/y_i - \theta \end{pmatrix}\right)$$

13.2.

13.3.

13.4.

13.5. да, да

13.6. Оценка при единичной весовой матрице равна $\hat{\theta}_{W=I} = 1.5$. С точностью до деления на определитель ковариационной матрицы оценка матрицы весов имеет вид:

$$\hat{W} = \begin{pmatrix} 12 & 1 \\ 1 & 102 \end{pmatrix}$$

16. Источники мудрости

[Zso10] Peter Zsohar. *Short introduction to the generalized method of moments.* 2010. URL: http://www.ksh.hu/statszemle_archive/2012/2012_K16/2012_K16_150.pdf.