Семинар 1. Ответы.

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

   (a) Рассмотрим модель $Y_i = \beta X_i + \varepsilon_i$:

      i. Найдите МНК-оценку β для имеющихся трёх наблюдений.

      ii. Нарисуйте исходные точки и полученную прямую регрессии.

      iii. Выведите формулу для β в общем виде для n наблюдений.

   (b) Рассмотрим модель $Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$:

      i. Найдите МНК-оценки β_1 и β_2 для имеющихся трёх наблюдений.

      ii. Нарисуйте исходные точки и полученную прямую регрессии.

      iii. Выведите формулы для β_1 и β_2 в общем виде для n наблюдений.

   Ответы:
   (a) i. $\hat{\beta} = 13/9$.
   (a) iii. $\hat{\beta} = \frac{\sum X_i Y_i}{\sum X_i^2}$.

   (b) i. $\hat{\beta}_1 = -1, \hat{\beta}_2 = 2$.
   (b) iii. $\hat{\beta}_2 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}, \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$.

2. Упростите выражения:

   (a) $n\bar{X} - \sum X_i$

   (b) $\sum (X_i - \bar{X})\bar{X}$

   (c) $\sum (X_i - \bar{X})\bar{Z}$

   (d) $\sum (X_i - \bar{X})^2 + n\bar{X}^2$

   Ответы:

   (a) 0

   (b) 0

   (c) 0

   (d) $\sum (X_i^2)$

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

   (a) $Y_i = \theta + \theta X_i + \varepsilon_i$;

   (b) $Y_i = 1 + \theta X_i + \varepsilon_i$;

   (c) $Y_i = \theta / X_i + \varepsilon_i$;

   (d) $Y_i = \theta X_i + (1-\theta) Z_i + \varepsilon_i$.

   Ответы:

   (a) $\hat{\theta} = \sum Y_i(1+X_i) / \sum(1+X_i)^2$

   Стандартная процедура МНК:

$$RSS = \sum \varepsilon_i^2 = \sum \left(Y_i - \theta - \theta X_i\right)^2 \to \min_{\theta}$$

$$\frac{\partial RSS}{\partial \theta} = 2 \sum \left(Y_i - \theta - \theta X_i\right)\left(-1 - X_i\right)$$

$$\sum \left(Y_i - \hat{\theta} - \hat{\theta} X_i\right)\left(-1 - X_i\right) = 0$$

$$\sum Y_i(-1 - X_i) + \hat{\theta} \sum (-1 - X_i)^2 = 0$$

$$\hat{\theta} = \frac{\sum Y_i(1+X_i)}{\sum(1+X_i)^2}$$

   (b) $\hat{\theta} = \sum \left((Y_i - 1)X_i\right) / \sum X_i^2$

   (c) $\hat{\theta} = \sum (Y_i/X_i) / \sum(1/X^2)$

   (d) $\hat{\theta} = \sum \left((Y_i - Z_i)(X_i - Z_i)\right) / \sum \left(X_i - Z_i\right)^2$

4. Рассмотрите модели $Y_i = \alpha + \beta(Y_i + Z_i) + \varepsilon_i$, $Z_i = \gamma + \delta(Y_i + Z_i) + \varepsilon_i$.

   (a) Как связаны между собой $\hat{\alpha}$ и $\hat{\gamma}$?

   (b) Как связаны между собой $\hat{\beta}$ и $\hat{\delta}$?

   Решение:

   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 $\hat{\beta} + \hat{\delta} = 1$.

5. 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 $Z_i = 2Y_i$?

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

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

$$Y_i = \frac{\gamma}{2} + \frac{\delta}{2} X_i + \frac{1}{2} \upsilon_i$$

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

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

6. 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.

   Решение:
   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

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

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

7. 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.

   Решение:
   Ане, Насте и Тане нужно оценить модель $Y_i = \beta + \varepsilon_i$. Для этого они должны решить следующую задачу:

   $$RSS = 2(10 - \beta)^2 + (3 - \beta)^2 \to \min_{\beta}$$

   $$\frac{\partial RSS}{\partial \beta} = -4(10 - \beta) - 2(3 - \beta)$$

   Условия первого порядка:

   $$-4(10 - \hat{\beta}) - 2(3 - \hat{\beta}) = 0$$

$$20 - 2\hat{\beta} + 3 - \hat{\beta} = 0$$

$$3\hat{\beta} = 23$$

$$\hat{\beta} = \frac{23}{3}$$

8. 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 $\hat{Y}_i = \hat{\phi}_1 + \hat{\phi}_2 X_i$ по всем наблюдениям.

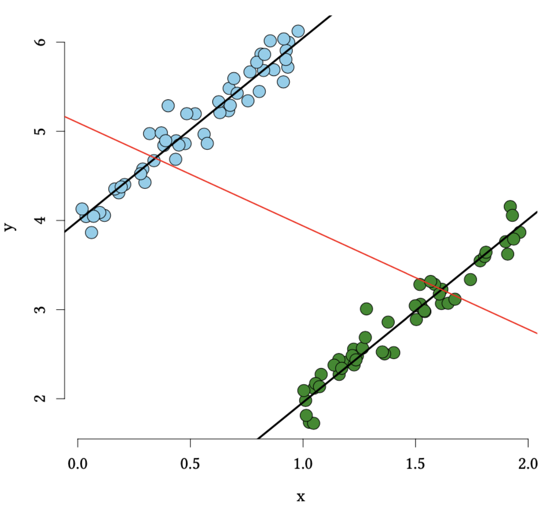
   (a) Возможно ли, что $\beta_2 > 0$, $\hat{\gamma}_2 > 0$, но $\hat{\phi}_2 < 0$?

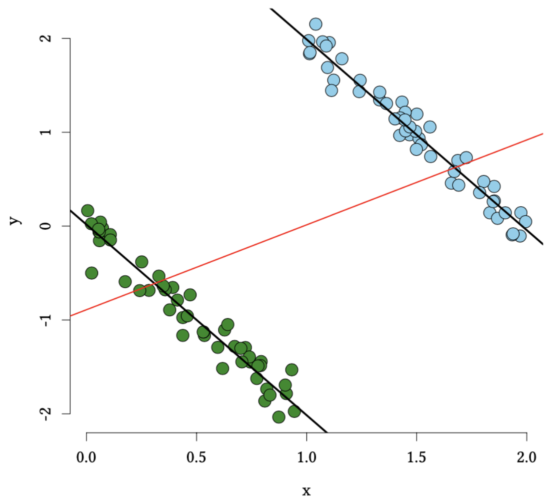
   (b) Возможно ли, что $\beta_1 > 0$, $\hat{\gamma}_1 > 0$, но $\hat{\phi}_1 < 0$?

   (c) Возможно ли одновременное выполнение всех упомянутых условий?

   (d) Возможно ли одновременное выполнение всех упомянутых условий, если в каждой сотне наблюдений $\sum X_i > 0$?

   Подсказка:





9. 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.

Решение:

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 $b = 7$.