

Эконометрика

ПМ-1701

Преподаватель:

КУРЫШЕВА СВЕТЛАНА ВЛАДИМИРОВНА

Санкт-Петербург
2020 г., 6 семестр

Список литературы

- [1] Эконометрика: Учебник/И.И.Елисеева и др.-М.:Проспект, 2009
- [2] Практикум по эконометрике: Учебное пособие/И.И.Елисеева и др.,М.:Финансы и статистика,2006
- [3] Эконометрика: Учебник/В. С.Мхитарян и др.-М.:2008
- [4] Доугерти К. Введение в эконометрику: Учебник. 2-е изд. / Пер. с англ. – М.: ИНФРА – М, 2007
- [5] Берндт Э. Практика эконометрики: классика и современность. М.,2005

Содержание

1	21.02.2020	2
1.1	Homework	3
2	28.02.2020	6

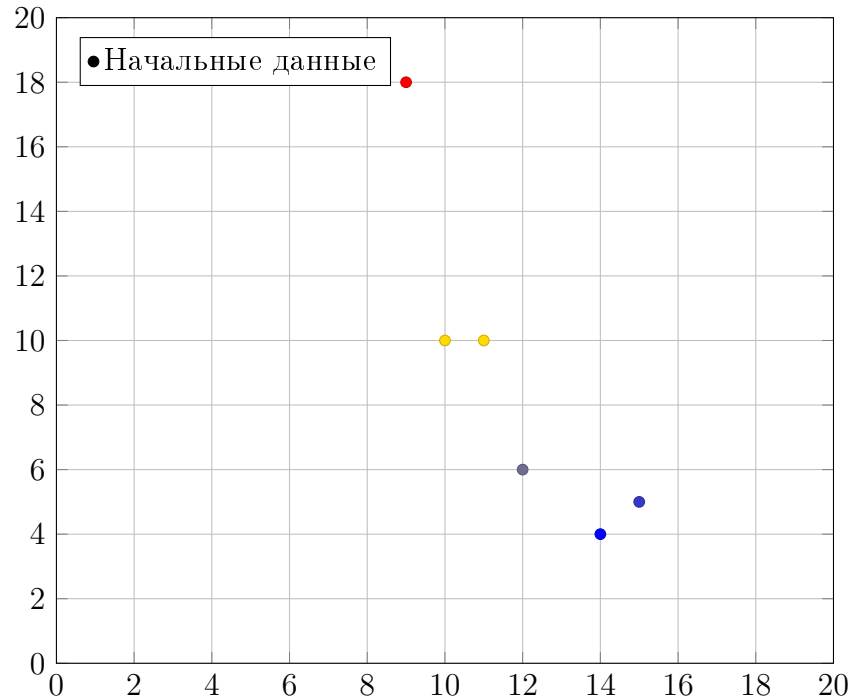
1 21.02.2020

Дана зависимость спроса от цены:

$$X = (15, 14, 12, 11, 9, 10)$$

$$Y = (5, 4, 6, 10, 18, 10)$$

1. Необходимо построить поле корреляции и выбрать математическую функцию.



По данной информации лучшей аппроксимации является нелинейная регрессия - степенная функция.

2. Найти линейное уравнение, используя МНК.

$$y = a + bx$$

Согласно формуле (7) получаем следующую систему уравнений:

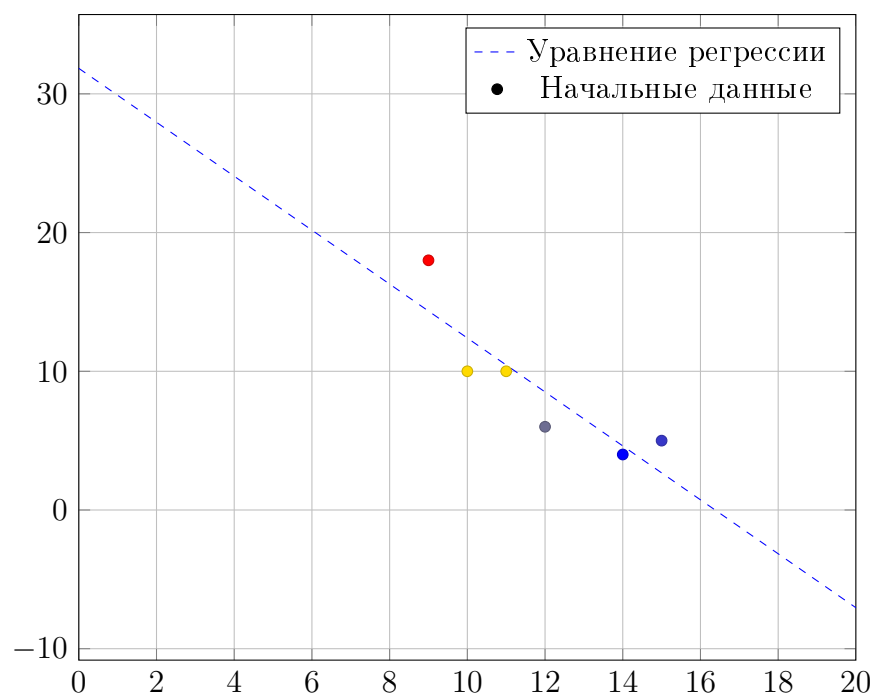
$$\begin{cases} 53 = 6a + 71b \\ 575 = 71a + 867b \end{cases}$$

Из данной системы уравнений находим значения параметров регрессии a и b :

$$a = 31.8385; b = -1.9441$$

Построим график прямой

$$\hat{Y} = 31.8385 - 1.9441X$$



Линейный коэффициент корреляции по формуле (10):

$$r = -0.87378$$

3. Построить таблицу дисперсионного анализа:

Источник вариации	df	SS	MS	F-критерий
Регрессия (r)	1	101.417	101.417	12.9127
Остаток (e)	4	31.4161	7.85404	1
Итого (t)	5	132.833	26.5667	x

Таблица 1: Таблица дисперсионного анализа для примера

Найдем табличное значение распределения Фишера-Снедекора при заданном уровне значимости:

$$F_{1-\alpha}(m, n-1-m) = F_{0.95}(1, 4) = 7.71$$

4. Найти линейное уравнение регрессии, используя программу Excel.

1.1 Homework

5. Дать интервальный прогноз спроса, для $x_p = 9$

Выражение для **стандартной ошибки предсказываемого по линии регрессии значения \hat{y}** :

$$m_{\hat{y}_x} = \sqrt{MS_E} \sqrt{\frac{1}{n} + \frac{(x_k - \bar{X})^2}{\sum (X - \bar{X})^2}} \quad (37)$$

Для прогнозируемого значения \hat{y} доверительный интервал выглядит следующим образом:

$$\begin{aligned} & \hat{y}_{x_k} \pm t_{1-\frac{\alpha}{2}} \cdot m_{\hat{y}_x} \\ & \hat{y}_{x_k} - t_{1-\frac{\alpha}{2}} \cdot m_{\hat{y}_x} \leq \hat{y}_{x_k} \leq \hat{y}_p + t_{1-\frac{\alpha}{2}} \cdot m_{\hat{y}_x} \end{aligned} \quad (38)$$

Средняя ошибка прогнозируемого индивидуального значения составит:

$$m_y = \sqrt{MS_E} \sqrt{1 + \frac{1}{n} + \frac{(x_k - \bar{X})^2}{\sum (X - \bar{X})^2}} \quad (39)$$

Доверительный интервал для y_p - предсказываемого значения регрессии:

$$\hat{y}_p - t_{\alpha} m_y \leq y_p \leq \hat{y}_p + t_{\alpha} m_y \quad (40)$$

Вычислим стандартную ошибку предсказываемого по линии регрессии значения \hat{Y} по формуле (37):

$$m_{\hat{y}_x} = \sqrt{7.85404} \sqrt{\frac{1}{6} + \frac{(x_k - 11.8333)^2}{26.8333}}$$

Подставляя различные значения из выборки X мы можем узнать ошибку предсказываемого значения. Минимальная ошибка будет при подстановке $x_k = \bar{X} = 11.8333$:

$$m_{y_{\bar{X}}} = \sqrt{11.8333} \sqrt{\frac{1}{6}} = 1.14412$$

Построим доверительный интервал для \hat{Y} при каком-то произвольном значении x_k , например $x_k = 9$. Воспользуемся формулой (38).

Сначала вычислим значение линейной регрессии в точке $x_k = 9$:

$$\hat{y}_9 = 31.8385 - 1.9441 \cdot 9 = 14.3416$$

Затем вычислим стандартную ошибку в точке $x_k = 9$:

$$m_{\hat{y}_9} = \sqrt{7.85404} \sqrt{\frac{1}{6} + \frac{(9 - 11.8333)^2}{26.8333}} = 1.91278$$

Теперь можно и построить доверительный интервал для уровня значимости $\alpha = 0.05$:

$$\hat{y}_9 - t_{0.975} \cdot m_{\hat{y}_9} \leq \hat{y}_9 \leq \hat{y}_9 + t_{0.975} \cdot m_{\hat{y}_9}$$

$$9.0309 \leq \hat{y}_9 \leq 19.6523$$

Средняя ошибка прогнозируемого индивидуального значения:

$$\begin{aligned} m_y &= \sqrt{MS_E} \sqrt{1 + \frac{1}{n} + \frac{(x_k - \bar{X})^2}{\sum (X - \bar{X})^2}} \\ &= \sqrt{7.85404} \sqrt{1 + \frac{1}{6} + \frac{(9 - 11.8333)^2}{26.8333}} = 3.39304 \end{aligned}$$

Доверительный интервал для y_p - предсказываемого значения регрессии:

$$4.92101 \leq \hat{y}_9 \leq 23.7622$$

6. Используя Excel найти уравнение регрессии по степенной функции.

1. Степенная функция

Модель:

$$y = a \cdot x^b \cdot \varepsilon$$

Логарифмируем обе части равенства (линеаризация):

$$\ln y = \ln a + b \ln x + \ln \varepsilon$$

Замена переменных:

$$\ln y = z, \alpha_1 = \ln a, t = \ln x, \varepsilon_1 = \ln \varepsilon$$

Линейный вид:

$$z = \alpha_1 + b \cdot t + \varepsilon_1$$

В нашем случае нам нужно прологарифмировать наши ряды x и y , найти коэффициенты a и b линейной функции и перейти обратно к степенной, сделав замену $a = e^{\alpha_1}$

$$X = (15, 14, 12, 11, 9, 10)$$

$$x_{new} = \ln X = (2.70805, 2.63906, 2.48491, 2.3979, 2.19722, 2.30259)$$

$$Y = (5, 4, 6, 10, 18, 10)$$

$$y_{new} = \ln Y = (1.60944, 1.38629, 1.79176, 2.30259, 2.89037, 2.30259)$$

Находим коэффициенты методом МНК:

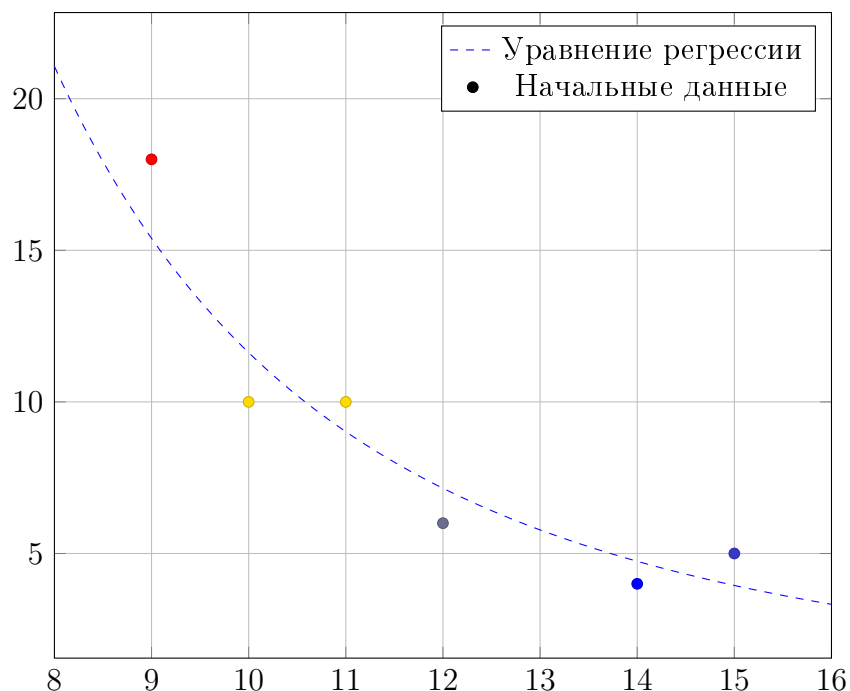
$$\alpha_1 = 8.58854, b = -2.66456$$

Исходная степенная модель имеет вид:

$$y = e^{\alpha_1} x^b = 5369.78 \cdot x^{-2.66456}$$

Построим график степенной функции

$$y = 5369.78 \cdot x^{-2.66456}$$



2 28.02.2020

Задача 1.

Изучается зависимость спроса от цены: $Y(X)$. Имеется выборка из 5 наблюдений и по данным наблюдениям построена модель линейной регрессии: $Y = 12.48 - 1.05X$. Линейный коэффициент корреляции для данной модели равен $r_{xy} = -0.97206$. Дисперсия спроса $\sigma_Y^2 = 2.3336$, а дисперсия цены равна $\sigma_X^2 = 2$. Также известно, что $\bar{X} = 4, \bar{Y} = 8.28$

Найти уравнение регрессии: $X(Y)$.

Решение:

Необходимо найти уравнение $X = A + BY$.

Известна формула $|r_{yx}| = \sqrt{b \cdot B}$, поэтому:

$$B = \frac{r_{yx}^2}{b} = -0.8999$$

$$A = \bar{X} - B\bar{Y} = 4 + 0.8999 \cdot 8.28 = 11.4512$$

Соответственно, уравнение регрессии принимает вид:

$$X = 11.4512 - 0.8999Y$$

Задача 2.

Коэффициент регрессии $a = 4$, его стандартная ошибка $m_a = 0.8$, коэффициент регрессии $b = 0.12$, его стандартная ошибка равна $m_b = 0.045$. Также даны результаты регрессии:

Источник вариации	df	SS
Регрессия (r)	1	1400
Остаток (e)	18	3600
Итого (t)	19	5000

Решение:

1. Найти общее количество наблюдений.

$$df_{SS_E} = n - 1 - m \Rightarrow n = df_{SS_E} + 1 + 1 = 20$$

2. Оценить значимость уравнения регрессии с помощью F -критерия и его параметров ($\alpha = 0.05$).

$$t_b = \frac{b}{m_b} = 2.66667 > t_{0.975}(20 - 2) = 2.10092 - \text{значим}$$

$$t_a = \frac{a}{m_a} = 5 > t_{0.975}(20 - 2) = 2.10092 - \text{значим}$$

$$F_{stat} = \frac{MS_R}{MS_E} = \frac{SS_R}{m} \cdot \frac{n - 1 - m}{SS_E} = \frac{1400}{1} \cdot \frac{18}{3600} = 7 > \\ > F_{1-\alpha}(m, n - 1 - m) = F_{0.95}(1, 18) = 4.41 - \text{связь существенна}$$

3. Найти коэффициент детерминации.

$$R^2 = \frac{SS_R}{SS_T} = \frac{1400}{5000} = 0.28$$

4. Дать интервальную оценку для коэффициента регрессии.

$$\delta_b = \pm t_{table} \cdot m_b$$

$$b - t_{1-\frac{\alpha}{2}}(n - 2) \cdot m_b \leq b \leq b + t_{1-\frac{\alpha}{2}}(n - 2) \cdot m_b \quad (35)$$

$$0.0254585 \leq b \leq 0.214541$$

Домашнее задание: практикум, парная регрессия, задачи 2,8 (стр.32).