

**Федеральное государственное образовательное бюджетное учреждение  
высшего образования**

**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РФ»**

**Департамент анализа данных и машинного обучения**

**Отчет по практике №3**

**по дисциплине «эконометрика»**

Студента группы ПМ23-1

Факультета  
информационных  
технологий и анализа  
больших данных

Тищенко И.С.

Преподаватель

Михайлова С.С.

Москва 2024

# Задание 1. Дз 3 нелинейная.xlsx. Лист2

Возьмем данные из столбцов таблицы и переведем их в объект класса `pandas.DataFrame`.

Прологарифмируем данные, чтобы создать требуемую модель.

На основе полученных данных создадим линейную регрессионную модель с помощью библиотеки `stats.models.api`:

OLS Regression Results						
Dep. Variable:		lnw		R-squared:	0.682	
Model:		OLS		Adj. R-squared:	0.668	
Method:		Least Squares		F-statistic:	47.26	
Date:	Sat, 19 Oct 2024		Prob (F-statistic):		6.64e-07	
Time:	22:34:37		Log-Likelihood:		-0.056948	
No. Observations:		24		AIC:	4.114	
Df Residuals:		22		BIC:	6.470	
Df Model:		1				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
1	2.2997	0.186	12.366	0.000	1.914	2.685
lnu	-0.7791	0.113	-6.875	0.000	-1.014	-0.544
Omnibus:		4.052	Durbin-Watson:		1.103	
Prob(Omnibus):		0.132	Jarque-Bera (JB):		2.290	
Skew:		-0.477	Prob(JB):		0.318	
Kurtosis:		4.174	Cond. No.		7.97	

Проанализируем информацию о модели:

- Коэффициент Детерминации равен: 68,235%
- Коэффициенты  $b$  при каждом члене равны соответственно: 2,2997037728451235; -0,7790563496663773 .
- Коэффициент 2,2997037728451235 стат. значим, т.к. значение  $t = 12,3658$  больше  $t_{\text{критического}} = 2,0739 \Leftrightarrow p\text{value} = 2,228541583865526e-11 < 0,05$
- Коэффициент -0,7790563496663773 стат. значим, т.к. значение  $t = 6,8746$  больше  $t_{\text{критического}} = 2,0739 \Leftrightarrow p\text{value} = 6,638587637386962e-07 < 0,05$
- Регрессия стат. значима, т.к. F-значение критерия фишера = 47,2597 больше  $F_{\text{критического}} \Leftrightarrow f\text{value} = 6,638587637386955e-07 < 0,05$

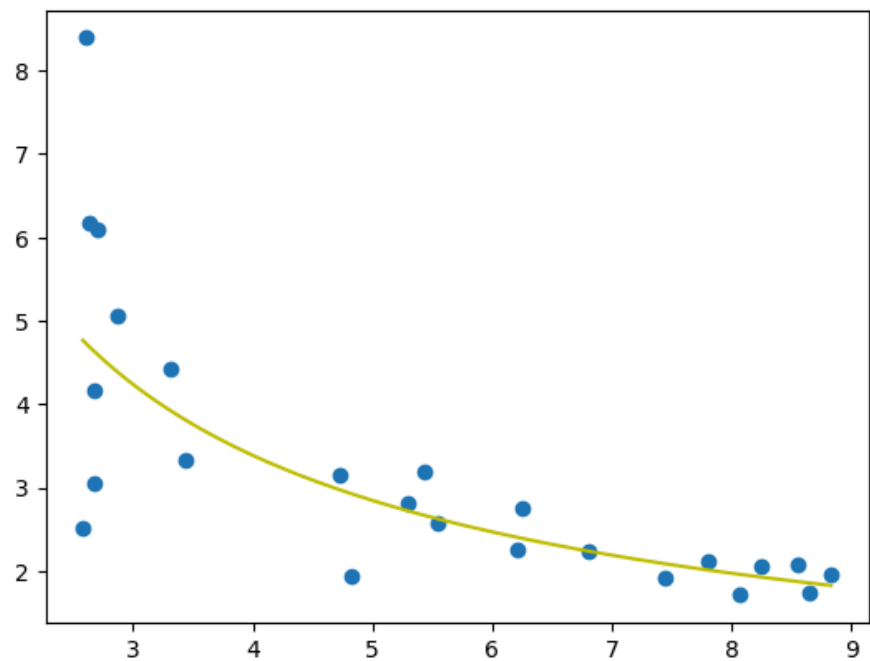
Напишем уравнение нелинейной регрессии.

$$w = Epsilon + u^b e^a$$

$$w = Epsilon + \frac{9.97119064887161}{u^{0.7791}}$$

Коэффициент детерминации этой модели составил уже 57,327%

Построим график регрессии вместе с диаграммой рассеяния:



Теперь стандартизируем данные с помощью инструмента `sklearn.preprocessing.StandardScaler().fit_transform()`, на основе преобразованных данных составим модель и выведем статистическую информацию:

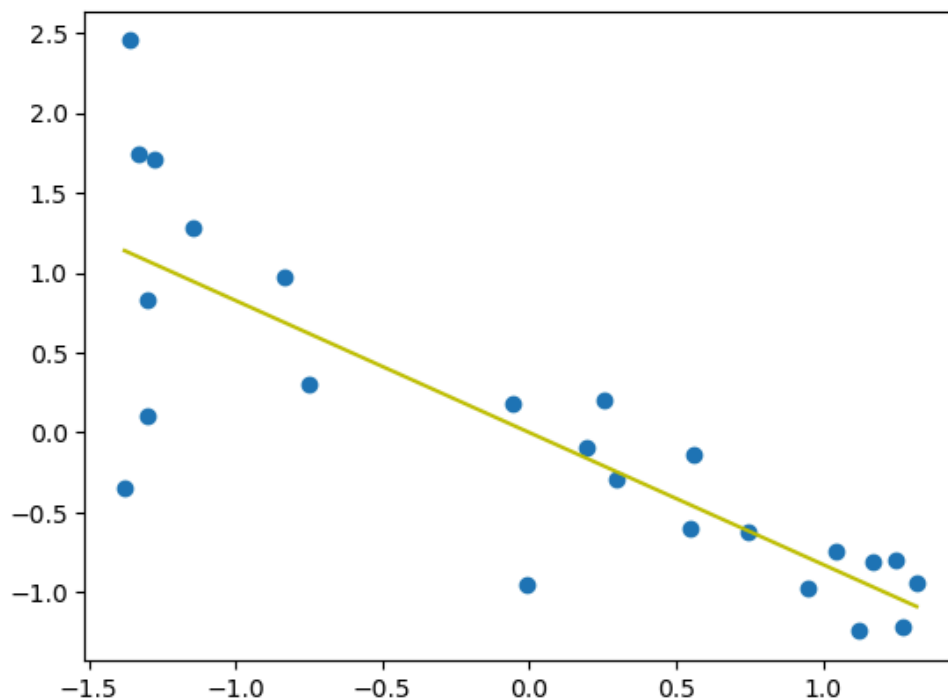
OLS Regression Results						
Dep. Variable:		ln_std_w		R-squared (uncentered):		0.682
Model:		OLS		Adj. R-squared (uncentered):		0.669
Method:		Least Squares		F-statistic:		49.41
Date:		Sat, 19 Oct 2024		Prob (F-statistic):		3.66e-07
Time:		22:34:37		Log-Likelihood:		-20.293
No. Observations:		24		AIC:		42.59
Df Residuals:		23		BIC:		43.76
Df Model:		1				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
ln_std_u	-0.8260	0.118	-7.029	0.000	-1.069	-0.583
Omnibus:		4.052	Durbin-Watson:		1.103	
Prob(Omnibus):		0.132	Jarque-Bera (JB):		2.290	
Skew:		-0.477	Prob(JB):		0.318	
Kurtosis:		4.174	Cond. No.		1.00	

Проанализируем информацию о модели:

- Коэффициент Детерминации равен: 68,235% .
- Коэффициенты  $b$  при каждом члене равны соответственно: -0,8260476877422089 .
- Коэффициент -0,8260476877422089 стат. значим, т.к. значение  $t = 7,0291$  больше  $t_{\text{критического}} = 2,0739 \Leftrightarrow p\text{value} = 3,6637058596786566e-07 < 0,05$
- Регрессия стат. значима, т.к. F-значение критерия фишера = 49,4078 больше  $F_{\text{критического}} \Leftrightarrow f\text{value} = 3,6637058596786704e-07 < 0,05$

Напишем уравнение нелинейной регрессии.

$$w = \text{Epsilon} + u^{-0.826}$$



## Задание 2. Дз 3 нелинейная.xlsx. Лист3

Возьмем данные из столбцов таблицы и переведем их в объект класса `pandas.DataFrame`. Преобразуем уравнение нелинейной регрессии в линейную форму.

$$\Delta K_t = \alpha I_t^\beta \times \epsilon_t$$

$$\ln(\Delta K_t) = \ln(\alpha) + \beta \times \ln(I_t) + \epsilon_t$$

Построим модель на изначальных данных и выведем статистическую информацию.

OLS Regression Results						
Dep. Variable:		y		R-squared:		0.982
Model:		OLS		Adj. R-squared:		0.981
Method:		Least Squares		F-statistic:		1359.
Date:		Sat, 19 Oct 2024		Prob (F-statistic):		2.58e-23
Time:		22:50:18		Log-Likelihood:		41.209
No. Observations:		27		AIC:		-78.42
Df Residuals:		25		BIC:		-75.83
Df Model:		1				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
lnl	0.9726	0.026	36.869	0.000	0.918	1.027
1	0.0363	0.117	0.311	0.759	-0.204	0.277
Omnibus:		18.786	Durbin-Watson:		0.522	
Prob(Omnibus):		0.000	Jarque-Bera (JB):		27.506	
Skew:		-1.489	Prob(JB):		1.06e-06	
Kurtosis:		6.947	Cond. No.		51.6	

Проанализируем информацию о модели:

- Коэффициент Детерминации равен: 98,194% .
- Коэффициенты b при каждом члене равны соответственно: 0,9726013794191091; 0,03626064833517084 .
- Коэффициент 0,9726013794191091 стат. значим, т.к. значение  $t = 36,8692$  больше  $t_{\text{критического}} = 2,0595 \Leftrightarrow p\text{value} = 2,5764166737745987e-23 < 0,05$
- Коэффициент 0,03626064833517084 стат. незначим, т.к. значение  $t = 0,3107$  меньше  $t_{\text{критического}} = 2,0595 \Leftrightarrow p\text{value} = 0,7585750524336108 > 0,05$
- Регрессия стат. значима, т.к. F-значение критерия фишера = 1359,3374 больше  $F_{\text{критического}} \Leftrightarrow f\text{value} = 2,5764166737745728e-23 < 0,05$

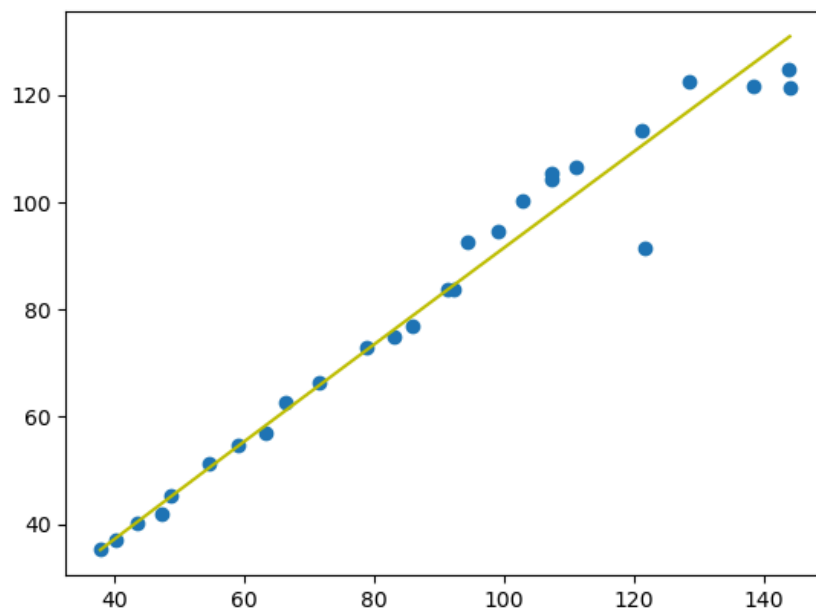
Можно заметить, что коэффициент при свободном члене не проходит проверку по Т-критерию Стьюдента, так что резонно построить модель без этого поля.

OLS Regression Results					
Dep. Variable:	y	R-squared (uncentered):	1.000		
Model:	OLS	Adj. R-squared (uncentered):	1.000		
Method:	Least Squares	F-statistic:	1.763e+05		
Date:	Sat, 19 Oct 2024	Prob (F-statistic):	2.42e-51		
Time:	22:50:19	Log-Likelihood:	41.157		
No. Observations:	27	AIC:	-80.31		
Df Residuals:	26	BIC:	-79.02		
Df Model:	1				
Covariance Type:	nonrobust				
	coef	std err	t	P> t	[0.025 0.975]
lnI	0.9808	0.002	419.829	0.000	0.976 0.986
Omnibus:	20.825	Durbin-Watson:	0.518		
Prob(Omnibus):	0.000	Jarque-Bera (JB):	32.803		
Skew:	-1.629	Prob(JB):	7.53e-08		
Kurtosis:	7.306	Cond. No.	1.00		

- Коэффициент Детерминации равен: 99,985% .
- Коэффициенты b при каждом члене равны соответственно: 0,9807652414335774 .
- Коэффициент 0,9807652414335774 стат. значим, т.к. значение  $t = 419,829$  больше  $t_{\text{критического}} = 2,0595 \Leftrightarrow p\text{value} = 2,4224054494664673e-51 < 0,05$
- Регрессия стат. значима, т.к. F-значение критерия фишера = 176256,354 больше  $F_{\text{критического}} \Leftrightarrow f\text{value} = 2,4224054494664916e-51 < 0,05$

Если использовать эти коэффициенты для нелинейной регрессии, то получим детерминацию, равную 96,428%.

$$K = I_t^{0.9808} + \epsilon$$



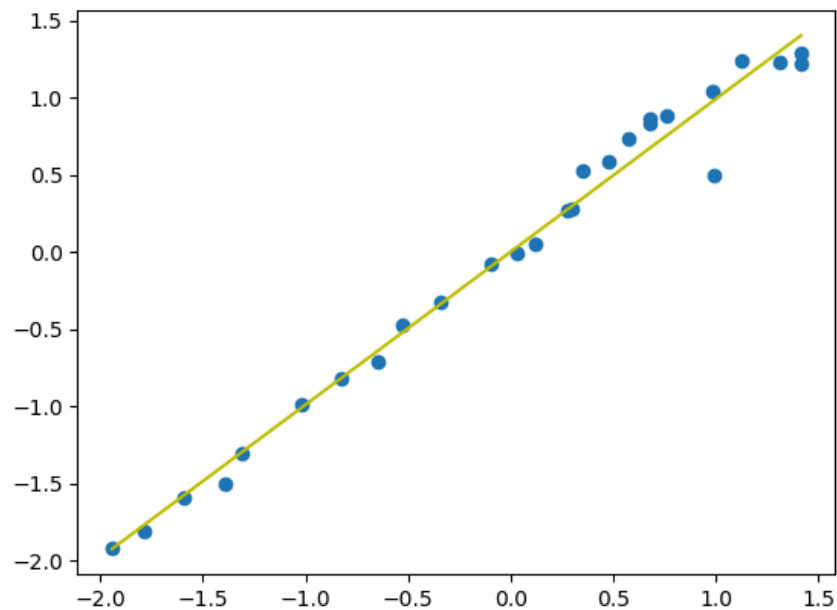
Стандартизуем логарифмированные данные и построим модель линейной регрессии:

OLS Regression Results						
Dep. Variable:		ln_std_K		R-squared (uncentered):		0.982
Model:		OLS		Adj. R-squared (uncentered):		0.981
Method:		Least Squares		F-statistic:		1414.
Date:		Sat, 19 Oct 2024		Prob (F-statistic):		3.40e-24
Time:		23:05:13		Log-Likelihood:		15.879
No. Observations:		27		AIC:		-29.76
Df Residuals:		26		BIC:		-28.46
Df Model:		1				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
ln_std_I	0.9909	0.026	37.599	0.000	0.937	1.045
Omnibus:	18.786	Durbin-Watson:		0.522		
Prob(Omnibus):	0.000	Jarque-Bera (JB):		27.506		
Skew:	-1.489	Prob(JB):		1.06e-06		
Kurtosis:	6.947	Cond. No.		1.00		

- Коэффициент Детерминации равен: 98,194% .
- Коэффициенты b при каждом члене равны соответственно: 0,9909292700664827 .
- Коэффициент 0,9909292700664827 стат. значим, т.к. значение  $t = 37,5993$  больше  $t_{\text{критического}} = 2,0595 \Leftrightarrow p\text{value} = 3,396455335888668\text{e-}24 < 0,05$
- Регрессия стат. значима, т.к. F-значение критерия фишера = 1413,7109 больше  $F_{\text{критического}} \Leftrightarrow f\text{value} = 3,39645533588866\text{e-}24 < 0,05$

$$\ln(K) = \epsilon + 0.9909\ln(I_t)$$

$$K = I_t^{0.9909} + \epsilon$$



### Задание 3. Дз 3 нелинейная.xlsx. Лист4

Возьмем данные из столбцов таблицы и переведем их в объект класса `pandas.DataFrame`.

Аналогично с предыдущими заданиями:



OLS Regression Results						
Dep. Variable:		lnQ		R-squared:		0.965
Model:		OLS		Adj. R-squared:		0.959
Method:		Least Squares		F-statistic:		163.6
Date:		Sat, 19 Oct 2024		Prob (F-statistic):		1.96e-09
Time:		23:08:35		Log-Likelihood:		44.745
No. Observations:		15		AIC:		-83.49
Df Residuals:		12		BIC:		-81.37
Df Model:		2				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
1	-5.3417	0.829	-6.443	0.000	-7.148	-3.535
lnL	1.6809	0.116	14.477	0.000	1.428	1.934
lnK	0.0077	0.019	0.415	0.685	-0.033	0.048
Omnibus:		6.644	Durbin-Watson:		2.148	
Prob(Omnibus):		0.036	Jarque-Bera (JB):		3.999	
Skew:		-1.241	Prob(JB):		0.135	
Kurtosis:		3.483	Cond. No.		2.58e+03	

- Коэффициенты  $b$  при каждом члене равны соответственно: -5,341670376619618; 1,6809432632307555; 0,007704622816963713 .
- Коэффициент -5,341670376619618 стат. значим, т.к. значение  $t = 6,4432$  больше  $t_{\text{критического}} = 2,1788 \Leftrightarrow p\text{value} = 3,1919952969638325e-05 < 0,05$
- Коэффициент 1,6809432632307555 стат. значим, т.к. значение  $t = 14,4774$  больше  $t_{\text{критического}} = 2,1788 \Leftrightarrow p\text{value} = 5,826128816733424e-09 < 0,05$
- Коэффициент 0,007704622816963713 стат. незначим, т.к. значение  $t = 0,4153$  меньше  $t_{\text{критического}} = 2,1788 \Leftrightarrow p\text{value} = 0,6852734964090719 > 0,05$

Третий параметр не подходит по Т-критерию Стьюдента, значит, стоит его убрать из уравнения.

OLS Regression Results						
Dep. Variable:		lnQ		R-squared:		0.964
Model:		OLS		Adj. R-squared:		0.961
Method:		Least Squares		F-statistic:		349.3
Date:		Sat, 19 Oct 2024		Prob (F-statistic):		8.92e-11
Time:		23:08:35		Log-Likelihood:		44.638
No. Observations:		15		AIC:		-85.28
Df Residuals:		13		BIC:		-83.86
Df Model:		1				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
1	-5.5010	0.711	-7.736	0.000	-7.037	-3.965
lnL	1.7090	0.091	18.689	0.000	1.511	1.907
Omnibus:		7.273	Durbin-Watson:		2.073	
Prob(Omnibus):		0.026	Jarque-Bera (JB):		4.414	
Skew:		-1.295	Prob(JB):		0.110	

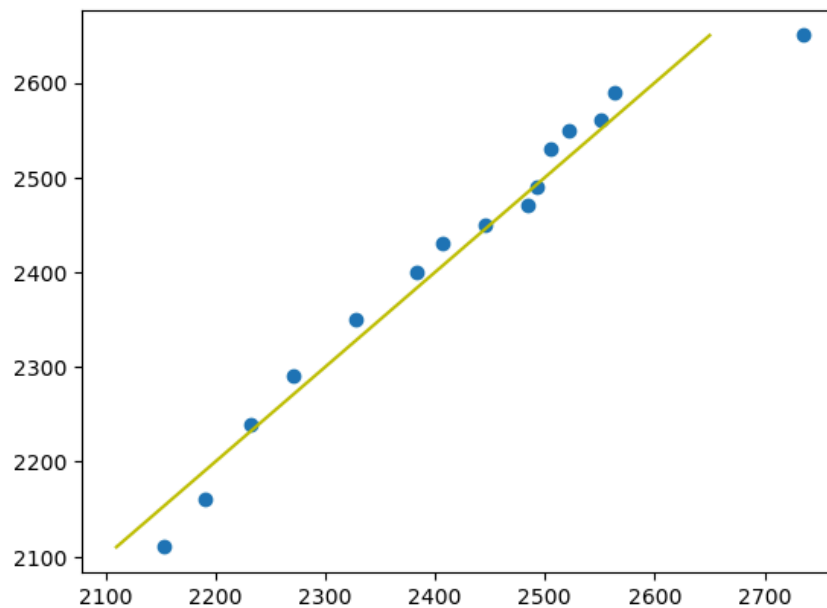
- Коэффициент Детерминации равен: 96,412% .
- Коэффициенты b при каждом члене равны соответственно: -5,501021987061762; 1,7089577240081084 .
- Коэффициент -5,501021987061762 стат. значим, т.к. значение  $t = 7,7359$  больше  $t_{\text{критического}} = 2,1604 \Leftrightarrow p\text{value} = 3,223205444753363\text{e-}06 < 0,05$
- Коэффициент 1,7089577240081084 стат. значим, т.к. значение  $t = 18,6889$  больше  $t_{\text{критического}} = 2,1604 \Leftrightarrow p\text{value} = 8,920073563460382\text{e-}11 < 0,05$
- Регрессия стат. значима, т.к. F-значение критерия фишера = 349,2753 больше  $F_{\text{критического}} \Leftrightarrow f\text{value} = 8,920073563460092\text{e-}11 < 0,05$

Уравнение линейной регрессии:

$$\ln(Q) = \epsilon + 1.709\ln(L) - 5.501$$

Уравнение в стандартной форме:

$$Q = 0.00408268670973036L^{1.709} + \epsilon$$



Стандартизируем логарифмированные данные и построим модель линейной регрессии:

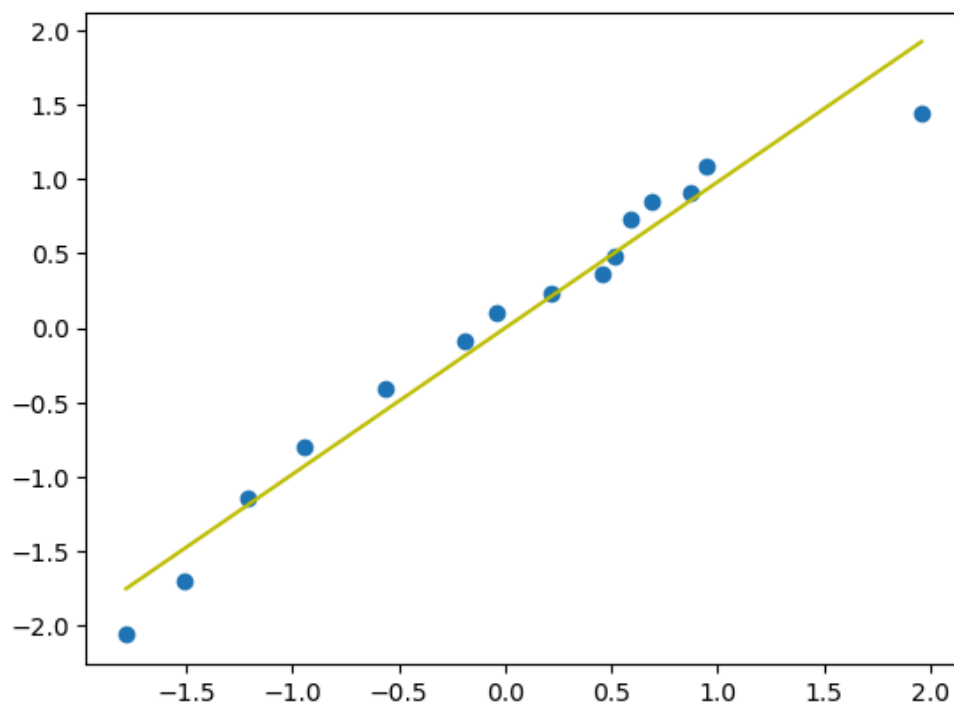
OLS Regression Results						
Dep. Variable:	ln_std_Q		R-squared (uncentered):		0.965	
Model:	OLS		Adj. R-squared (uncentered):		0.959	
Method:	Least Squares		F-statistic:		177.2	
Date:	Sat, 19 Oct 2024		Prob (F-statistic):		3.69e-10	
Time:	23:08:36		Log-Likelihood:		3.7788	
No. Observations:	15		AIC:		-3.558	
Df Residuals:	13		BIC:		-2.142	
Df Model:	2					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
ln_std_L	0.9658	0.064	15.069	0.000	0.827	1.104
ln_std_K	0.0277	0.064	0.432	0.673	-0.111	0.166
Omnibus:	6.644	Durbin-Watson:		2.148		
Prob(Omnibus):	0.036	Jarque-Bera (JB):		3.999		
Skew:	-1.241	Prob(JB):		0.135		
Kurtosis:	3.483	Cond. No.		1.94		

Аналогично с нестандартизированными данными коэффициент при К не проходит проверку по Т-критерию Стьюдента.

OLS Regression Results						
Dep. Variable:		ln_std_Q		R-squared (uncentered):		0.964
Model:		OLS		Adj. R-squared (uncentered):		0.962
Method:		Least Squares		F-statistic:		376.1
Date:		Sat, 19 Oct 2024		Prob (F-statistic):		1.63e-11
Time:		23:08:36		Log-Likelihood:		3.6718
No. Observations:		15		AIC:		-5.344
Df Residuals:		14		BIC:		-4.636
Df Model:		1				
Covariance Type:		nonrobust				
	coef	std err	t	P> t	[0.025	0.975]
ln_std_L	0.9819	0.051	19.394	0.000	0.873	1.090
Omnibus:		7.273	Durbin-Watson:		2.073	
Prob(Omnibus):		0.026	Jarque-Bera (JB):		4.414	
Skew:		-1.295	Prob(JB):		0.110	
Kurtosis:		3.598	Cond. No.		1.00	

- Коэффициент Детерминации равен: 96,412% .
- Коэффициенты b при каждом члене равны соответственно: 0,9818939281520398 .
- Коэффициент 0,9818939281520398 стат. значим, т.к. значение  $t = 19,3944$  больше  $t_{\text{критического}} = 2,1604 \Leftrightarrow p\text{value} = 1,630760045990652e-11 < 0,05$
- Регрессия стат. значима, т.к. F-значение критерия фишера = 376,1426 больше  $F_{\text{критического}} \Leftrightarrow f\text{value} = 1,6307600459906518e-11 < 0,05$

$$\ln(Q) = \epsilon + 0.9819\ln(L)$$



## Задание 4. Задача 18.pdf

Выпишем все данные из условия задачи.

$$\mathbb{E}(Y) = 1000$$

$$\mathbb{E}(X_1) = 420$$

$$\mathbb{E}(X_2) = 41,5$$

$$\sigma_Y = 27$$

$$\sigma_{X_1} = 45$$

$$\sigma_{X_2} = 18$$

$$r_Y$$

$$r_Y$$

$$r_{X_1}$$

### №1

Для нахождения коэффициентов уравнения приведем коэффициенты парной корреляции в матричный вид

$$R = \begin{pmatrix} 1 & r_{X_1 X_2} = 0,38 \\ r_{X_1 X_2} = 0,38 & 1 \end{pmatrix}$$

$$r = \begin{pmatrix} r_{Y X_1} = 0,77 \\ r_{Y X_2} = 0,43 \end{pmatrix}$$

$\beta = R^{-1} \times r$  - Коэффициенты уравнения в стандартизированном масштабе, откуда:

$$y = 0.708976157082749x_1 + 0.16058906030855535x_2$$

Чтобы найти коэффициенты уравнения в натуральном масштабе, полученные коэффициенты  $\beta$  умножим на среднее каждого  $X$ , просуммируем полученные произведения и вычтем из среднего значения  $Y$ . Запишем результаты в виде уравнения в натуральном масштабе.

$$y = 811.3413394109397 + 0.4253856942496494x_1 + 0.24088359046283303x_2$$

### №2

Формула для частной корреляции между переменными  $X_i$  и  $X_j$  с учётом всех остальных переменных в отношении матриц выглядит следующим образом :

$$r_{ij \cdot other} = -\Omega_{ij} / \sqrt{\Omega_{ii}\Omega_{jj}}$$

Здесь  $\Omega$  - обратная матрица к матрице парных корреляций.

Из этого равенства получим следующие коэффициенты.

$$r_{YX_1 \cdot other} = 0.72637614236143$$

$$r_{YX_2 \cdot other} = 0.232809551177815$$

$$r_{X_1X_2 \cdot other} = 0.084889282207865$$

Для множественного коэффициента корреляции формула в матричном виде выглядит так:

$$R_{plur} = r^T \times R^{-1} \times r = 0.7841970013245367$$

### №3

Для нахождения коэффициентов эластичности при каждом  $X$  вспомним про коэффициенты  $b$ . Игнорируя коэффициент свободного члена, поэлементно умножим каждый из  $b_i$  на  $\mathbb{E}(X_i)$  и разделим на  $\mathbb{E}(Y)$  :

Частный коэффициент эластичности при  $X_1 = 0.178661991584853$

Он меньше, чем коэффициент  $\beta_{X_1}$

Частный коэффициент эластичности при  $X_2 = 0.00999666900420757$

Он меньше, чем коэффициент  $\beta_{X_2}$

### №4

С помощью посчитанных в №2 частных коэффициентов корреляции, найдем частные значения F-критериев Фишера для всех  $X_i$ :

$$F_{X_i} = r_{YX_i \cdot other}^2 \times (n - k - 1) / (1 - r_{YX_i \cdot other}^2) \times k$$

$n$  - количество наблюдений

$k$  - количество  $X$

А общий F-критерий Фишера рассчитаем по формуле:

$$F_{all} = R_{plur}^2 \times (n - k - 1) / (1 - R_{plur}^2) \times k$$

Откуда:

$$F_{all} = 21.5617418861327$$

$$F_{X_1} = 15.0788258113867$$

$$F_{X_2} = 0.773635121844607$$