МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Юго-Западный государственный университет»

Кафедра «Машиностроительные технологии и оборудование»

УТВЕРЖДАЮ
Проректор по учебной работе
_____________ О.Г. Локтионова
«_7_» _02_ ___________

ЛИНЕЙНАЯ РЕГРЕССИЯ. ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ ЛИНЕЙНОЙ РЕГРЕССИИ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

Методические указания к проведению практических занятий для студентов по направлению подготовки 15.03.05 Конструкторско-технологическое обеспечение машиностроительных производств профиль «Технология машиностроения»

Курск 2018

УДК 519.6

Составители: В.В. Куц, М.С. Разумов

Рецензент
Кандидат технических наук, доцент *А.О. Гладышкин*

Линейная регрессия. Определение параметров линейной регрессии методом наименьших квадратов : методические указания к проведению практических и лабораторных занятий / Юго-Зап. гос. ун-т; сост.: В.В. Куц, М.С. Разумов. – Курск, 2018. 20 с.: ил. 3.: табл. 3.

Содержат сведения по вопросам определения параметров линейной регрессии методом наименьших квадратов. Указывается порядок выполнения практического занятия, подходы к решению и правила оформления.

Методические рекомендации соответствуют требования программы, утвержденной учебно-методическим объединением по специальности автоматизированного машиностроительного производства (УМОАМ).

Предназначено для студентов направлений 15.03.05 «Конструкторско-технологическое обеспечение машиностроительных производств» профиль «Технология машиностроения» дневной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать 07.02.18 г. Формат 60х84 1/16.

Усл.печ.л. 1,1. Уч.-изд.л. 0,9. Тираж 40 экз. Заказ.874 Бесплатно.
Юго-Западный государственный университет.

1 **Цель работы:** исследовать зависимости между результатами наблюдений и определить их характер.

2 Задание: построить диаграммы рассеивания, рассчитать коэффициенты корреляции для результатов наблюдения трёх случайных величин, а установить значимость этих коэффициентов.

3 Краткие теоретические сведения

3.1 Линейная регрессия

Регрессионный анализ позволяет приближенно определить форму связи между результативным и факторными признаками, а также решить вопрос о том, значима ли эта связь. Вид функции, с помощью которой приближенно выражается форма связи, выбирают заранее, исходя из содержательных соображений или визуального анализа данных. Математическое решение задачи основано на методе наименьших квадратов, с помощью которого определяют степень приближения полученной функции к опытным данным.

В вычислительном аспекте метод наименьших квадратов сводится к составлению и решению системы так называемых **нормальных уравнений.** Исходным этапом для этого является подбор вида функции, отображающей статистическую связь.

Тип функции в каждом конкретном случае можно подобрать путем прикидки на графике исходных данных подходящей, т. е. достаточно хорошо приближающей эти данные, линии, которая может быть записана в виде

$$Y = a_0 + a_1 X , \tag{1}$$

где Y – результативный признак или зависимая переменная; X – факторный признак или независимая переменная; a_0 и a_1 – параметры уравнения, которые могут быть найдены методом наименьших квадратов.

Для нахождения искомых параметров нужно составить систему уравнений, которая в данном случае будет иметь вид

$$\begin{cases} na_0 + a_1\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i \\ a_0\sum_{i=1}^{n} X_i + a_1\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i Y_i \end{cases}. \qquad (2)$$

Полученная система может быть решена методом Гаусса. Искомые параметры системы из двух нормальных уравнений можно вычислить и непосредственно с помощью последовательного использования нижеприведенных формул:

$$a_1 = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2},\ a_0 = \overline{Y} - a_1\overline{X}, \qquad (3)$$

где Y_i, – i-е значение результативного признака; X_i – i-е значение факторного признака; $\overline{Y}$ и $\overline{X}$ – средние арифметические результативного и факторного признаков соответственно; n – число значений признака Y_i, или, что то же самое, число значений признака X_i.

Уравнение регрессии не только определяет форму анализируемой связи, но и показывает, в какой степени изменение одного признака сопровождается изменением другого признака.

Коэффициент при X (a_1), называемый **коэффициентом регрессии,** показывает, на какую величину в среднем изменяется результативный признак Y при изменении факторного признака X на единицу.

3.2 Нелинейная регрессия

Не всегда связь между признаками может быть достаточно хорошо представлена линейной функцией. Иногда для описания существующей связи более пригодными, а порой и единственно возможными являются более сложные нелинейные функции. Ограничимся рассмотрением наиболее простых из них.

Одним из простейших видов нелинейной зависимости является парабола, которая в общем виде может быть представлена функцией

$$Y = a_0 + a_1 X + a_2 X^2. \qquad (4)$$

Неизвестные параметры a_0, a_1, a_2 находятся в результате решения следующей системы уравнений:

$$\begin{cases} na_0 + a_1\sum_{i=1}^{n} X_i + a_2\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} Y_i, \\ a_0\sum_{i=1}^{n} X_i + a_1\sum_{i=1}^{n} X_i^2 + a_2\sum_{i=1}^{n} X_i^3 = \sum_{i=1}^{n} Y_i X_i, \\ a_0\sum_{i=1}^{n} X_i^2 + a_1\sum_{i=1}^{n} X_i^3 + a_2\sum_{i=1}^{n} X_i^4 = \sum_{i=1}^{n} Y_i X_i^2. \end{cases} \tag{5}$$

Систему уравнений (5) можно представить в виде:

$$\begin{cases} g_{00}a_0 + g_{01}a_1 + g_{02}a_2 = h_0; \\ g_{10}a_0 + g_{11}a_1 + g_{12}a_2 = h_1; \\ g_{20}a_0 + g_{21}a_1 + g_{22}a_2 = h_2; \end{cases} \tag{6}$$

где

$$g_{00} = n\ ;\ g_{01} = \sum_{i=1}^{n} X_i\ ;\ g_{02} = \sum_{i=1}^{n} X_i^2;\ h_0 = \sum_{i=1}^{n} Y_i;$$

$$g_{10} = \sum_{i=1}^{n} X_i;\ g_{11} = \sum_{i=1}^{n} X_i^2;\ g_{12} = \sum_{i=1}^{n} X_i^3;\ h_1 = \sum_{i=1}^{n} X_i Y_i; \tag{7}$$

$$g_{20} = \sum_{i=1}^{n} X_i^2;\ g_{21} = \sum_{i=1}^{n} X_i^3;\ g_{22} = \sum_{i=1}^{n} X_i^4;\ h_2 = \sum_{i=1}^{n} X_i^2 Y_i$$

Решение системы (6) также может быть найдено с помощью правила Крамера:

$$a_0 = \frac{\Delta_0}{\Delta}, a_1 = \frac{\Delta_1}{\Delta},\ a_2 = \frac{\Delta_2}{\Delta}, \tag{8}$$

где Δ – определитель системы (6), составленный из коэффициентов при неизвестных:

$$\Delta = \begin{vmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{vmatrix} =$$

$$= g_{00}g_{11}g_{22} + g_{10}g_{21}g_{02} + g_{01}g_{12}g_{20} - \tag{9}$$

$$- g_{20}g_{11}g_{02} - g_{10}g_{01}g_{22} - g_{21}g_{12}g_{00}.$$

Определители Δ_0, Δ_1 и Δ_2 получаются из определителя Δ путем замены столбца коэффициентов при соответствующих переменных столбцом правых частей системы:

$$\Delta_0 = \begin{vmatrix} h_0 & g_{01} & g_{02} \\ h_1 & g_{11} & g_{12} \\ h_2 & g_{21} & g_{22} \end{vmatrix} =$$

$$= h_0 g_{11} g_{22} + h_1 g_{21} g_{02} + g_{01} g_{12} h_2 - \qquad (10)$$

$$- h_2 g_{11} g_{02} - h_1 g_{01} g_{22} - g_{21} g_{12} h_0,$$

$$\Delta_1 = \begin{vmatrix} g_{00} & h_0 & g_{02} \\ g_{10} & h_1 & g_{12} \\ g_{20} & h_2 & g_{22} \end{vmatrix} =$$

$$= g_{00} h_1 g_{22} + g_{10} h_2 g_{02} + h_0 g_{12} g_{20} - \qquad (11)$$

$$- g_{20} h_1 g_{02} - g_{10} h_0 g_{22} - h_2 g_{12} g_{00},$$

$$\Delta_2 = \begin{vmatrix} g_{00} & g_{01} & h_0 \\ g_{10} & g_{11} & h_1 \\ g_{20} & g_{21} & h_2 \end{vmatrix} =$$

$$= g_{00} g_{11} h_2 + g_{10} g_{21} h_0 + g_{01} h_1 g_{20} - \qquad (12)$$

$$- g_{20} g_{11} h_0 - g_{10} g_{01} h_2 - g_{21} h_1 g_{00}.$$

Дает ли преимущества описание связи с помощью параболы по сравнению с описанием, построенным по гипотезе линейности? Ответ на этот вопрос можно получить, рассчитав последовательный F-критерий.

На практике для изучения связей используются полиномы более высоких порядков (3-го и 4-го порядков). Составление системы, ее решение, а также решение вопроса о полезности повышения порядка функции для этих случаев аналогичны описанным. При этом никаких принципиально новых моментов не возникает, но существенно увеличивается объем расчетов.

Кроме класса парабол для анализа нелинейных связей можно применять и другие виды функций. Для расчета неизвестных параметров этих функций рекомендуется использовать метод наименьших квадратов, как наиболее мощный и широко применяемый.

Однако метод наименьших квадратов не универсален, поскольку он может использоваться только при условии, что выбранные для выравнивания функции линейны по отношению к своим параметрам. Не все функции удовлетворяют этому условию, но большинство применяемых на практике с помощью специальных преобразований могут быть приведены к стандартной форме функции с линейными параметрами.

3.3 Коэффициент детерминации

При регрессионном анализе важным является вопрос - в какой степени величина X определяет величину Y? На этот вопрос можно ответить, рассчитав, какая часть вариации результативного признака Y может быть объяснена влиянием факторного признака X.

Рассмотрим отношение

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n} \left(Y_i - \overline{Y}\right)^2}. \qquad (15)$$

Оно показывает долю разброса, учитываемого регрессией, в общем разбросе результативного признака и носит название **коэффициента детерминации.** Этот показатель, равный отношению факторной вариации к полной вариации признака, позволяет судить о том, насколько «удачно» выбран вид функции. Проведя расчеты, основанные на одних и тех же исходных данных, для нескольких типов функций, мы можем из них выбрать такую, которая дает наибольшее значение R^2 и, следовательно, в большей степени, чем другие функции, объясняет вариацию результативного признака. Действительно, при расчете R^2 для одних и тех же данных, но разных функций знаменатель выражения (15) остается неизменным, а числитель показывает ту часть вариации результативного признака, которая учитывается выбранной функцией. Чем больше R^2, т. е. чем больше числитель, тем больше изменение факторного признака объясняет изменение результативного признака и тем, следовательно, лучше уравнение регрессии, лучше выбор функции.

Коэффициент детерминации всегда находится в пределах интервала [0;1]. Если значение R^2 близко к единице, это означает, что построенная модель объясняет почти всю изменчивость соответствующих переменных. И наоборот, значение R^2 близкое к нулю, означает плохое качество построенной модели.

Коэффициент детерминации R^2 показывает, на сколько процентов $R^2 \cdot 100\%$ найденная функция регрессии описывает связь между исходными значениями Y и X. На рис. 1 показана $\left(\hat{Y}_i - \overline{Y}\right)$ – объясненная регрессионной моделью вариация и $\left(Y_i - \overline{Y}\right)$ – общая вариация. Соответственно, величина $\left(1 - R^2\right) \cdot 100\%$ показывает, сколько процентов вариации параметра Y обусловлены факторами, не включенными в регрессионную модель.

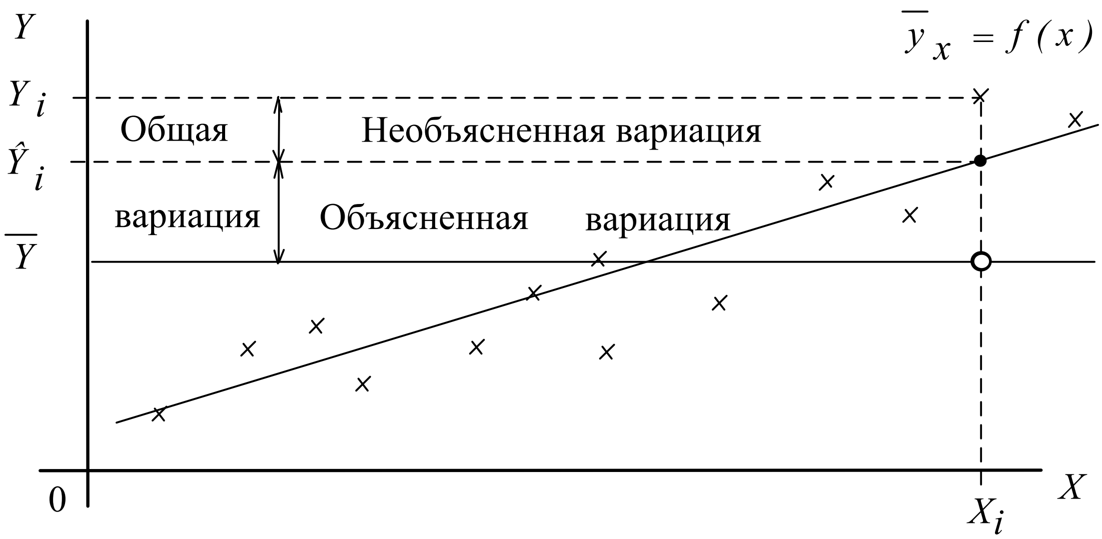

Рис. 1. Графическая интерпретация коэффициента детерминации (линейная регрессия)

При высоком значении коэффициента детерминации $R^2 \geq 75\%$ можно делать прогноз $Y^* = f\left(X^*\right)$ для конкретного значения X^* в пределах диапазона исходных данных. При прогнозах значений, не входящих в диапазон исходных данных, справедливость полученной модели гарантировать нельзя. Это объясняется тем, что может проявиться влияние новых факторов, которые модель не учитывает.

3.4 Критерий значимости регрессии

Критерий значимости регрессии позволяет дать ответы на вопросы, существует ли связь? Значимо ли уравнение регрессии, используемое для отображения предполагаемой связи?

Мерой значимости линии регрессии может служить следующее соотношение

$$F[m, n-m-1] = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}$$

где n – число значений X или, что то же самое, значений Y; m – число факторных признаков (независимых переменных, в нашем случае m=1).

Действительно, связь тем больше, чем значительнее мера рассеяния признака, обусловленная регрессией, превосходит меру рассеяния отклонений фактических значений от выравненных.

Соотношение (13) позволяет решить вопрос о значимости регрессии. Регрессия значима, т. е. между признаками существует связь, если для данного уровня значимости вычисленное значение $F[m, n\text{-}m\text{-}1]$ превышает критическое значение $F_{кр}[m, n\text{-}m\text{-}1]$, стоящее на пересечении m-го столбца и $[n\text{-}m\text{-}1]$-й строки таблицы Фишера (Приложение А).

3.5 Пример расчета

Для 20 парных наблюдений X и Y (табл. 1) построить уравнение линейной регрессии и нелинейной (полином второй степени), сделать вывод о значимости полученных уравнений (при q=0,05) и о том, какое уравнение лучше описывает взаимосвязь между параметрами X и Y.

Таблица 1.

№ наб-я	1	2	3	4	5	6	7	8	9	10
X_i	1,338	1,803	0,844	0,933	0,604	1,649	0,543	0,071	1,349	0,919
Y_i	9,392	13,124	5,932	6,52	4,492	11,945	4,191	1,844	9,477	6,336
№ наб-я	1	2	3	4	5	6	7	8	9	10
X_i	0,794	1,909	0,257	1,817	1,219	1,062	0,046	1,674	0,839	0,026
Y_i	5,577	14,199	2,717	13,332	8,512	7,406	1,763	12,117	6,02	1,687

Для построения линейной регрессии заполним следующую таблицу (табл. 2).

Таблица 2

№	X_i	Y_i	X_i^2	X_iY_i	$\hat{Y}_i$	$Y_i-\hat{Y}_i$	$(Y_i-\hat{Y}_i)^2$	$Y_i-\overline{Y}$	$(Y_i-\overline{Y})^2$
1	1,338	9,392	1,7902	12,566	9,664	-0,272	0,074	2,063	4,255
2	1,803	13,124	3,2508	23,663	12,738	0,386	0,149	5,795	33,580
3	0,844	5,932	0,7123	5,007	6,398	-0,466	0,217	-1,397	1,952
4	0,933	6,52	0,8705	6,083	6,987	-0,467	0,218	-0,809	0,655
5	0,604	4,492	0,3648	2,713	4,812	-0,320	0,102	-2,837	8,049
6	1,649	11,945	2,7192	19,697	11,720	0,225	0,051	4,616	21,306
7	0,543	4,191	0,2948	2,276	4,408	-0,217	0,047	-3,138	9,848
8	0,071	1,844	0,0050	0,131	1,288	0,556	0,309	-5,485	30,087
9	1,349	9,477	1,8198	12,784	9,737	-0,260	0,068	2,148	4,613
10	0,919	6,336	0,8446	5,823	6,894	-0,558	0,312	-0,993	0,986
11	0,794	5,577	0,6304	4,428	6,068	-0,491	0,241	-1,752	3,070
12	1,909	14,199	3,6443	27,106	13,439	0,760	0,577	6,870	47,195
13	0,257	2,717	0,0660	0,698	2,518	0,199	0,040	-4,612	21,272
14	1,817	13,332	3,3015	24,224	12,831	0,501	0,251	6,003	36,034
15	1,219	8,512	1,4860	10,376	8,877	-0,365	0,134	1,183	1,399
16	1,062	7,406	1,1278	7,865	7,840	-0,434	0,188	0,077	0,006
17	0,046	1,763	0,0021	0,081	1,123	0,640	0,410	-5,566	30,982
18	1,674	12,117	2,8023	20,284	11,885	0,232	0,054	4,788	22,924
19	0,839	6,02	0,7039	5,051	6,365	-0,345	0,119	-1,309	1,714
20	0,026	1,687	0,0007	0,044	0,990	0,697	0,485	-5,642	31,834
Σ	19,696	146,58	26,437	190,9			4,045		311,762
	$\overline{X}=0{,}985$	$\overline{Y}=7{,}329$							

Далее рассчитаем коэффициенты линейного уравнения

$$a_1=\frac{n\sum_{i=1}^{n}X_iY_i-\sum_{i=1}^{n}X_i\sum_{i=1}^{n}Y_i}{n\sum_{i=1}^{n}X_i^2-\left(\sum_{i=1}^{n}X_i\right)^2}=\frac{20\cdot190{,}9-19{,}696\cdot146{,}58}{20\cdot26{,}437-19{,}696^2}=6{,}61;$$

$$a_0=\overline{Y}-a_1\overline{X}=7{,}329-6{,}61\cdot0{,}985=0{,}82.$$

Построим диаграмму рассеивания экспериментальных данных и линию полученного уравнения регрессии Y=0,82+6,61X (рис. 2). На рисунке хорошо заметны отклонения экспериментальных точек от линии регрессии.

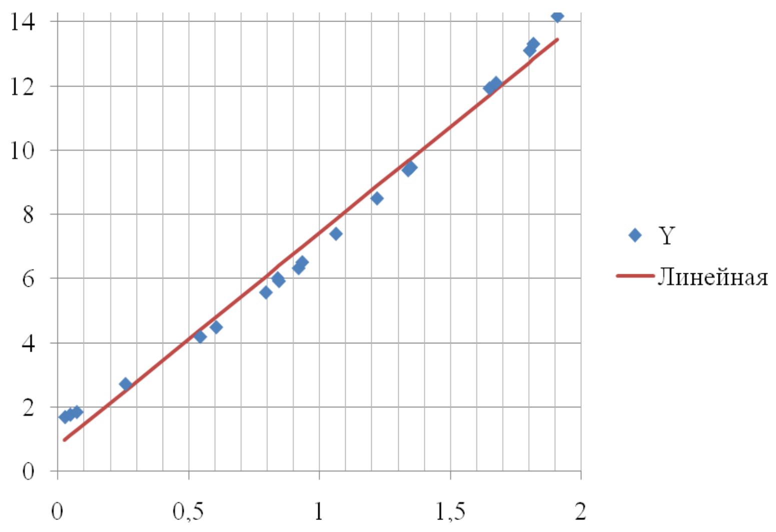

Рис. 2 - Диаграмма рассеивания экспериментальных данных и линия регрессии Y=0,82+6,61X

Выполним оценку значимости линейной модели, расчитаем коэффициент детерминации

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2} = 1 - \frac{4{,}045}{311{,}762} = 0{,}987$$

и значение F-критерия

$$F[1{,}18] = \frac{R^2}{1 - R^2} \cdot \frac{n - m - 1}{m} = \frac{0{,}987}{1 - 0{,}987} \cdot \frac{20 - 1 - 1}{1} \approx 1369$$

Из таблицы приложения А при q=0,05 и k_1=1 и k_2=18 выбираем критическое значение F-критерия

$F_{кр}[1,18]= 4,41.$

Т.к. $F[1,18]$= 1369>$F_{кр}[1,18]$= 4,41, то полученное уравнение можно признать значимым.

Далее выполним построения уравнения регрессии в виде полинома второй степени, для этого заполним следующую таблицу (табл. 3).

Таблица 3

№	X_i	Y_i	X_i^2	X_i^3	X_i^4	$X_i^2 Y$	$\hat{Y}_i$	$Y_i - \hat{Y}_i$	$\left(Y_i - \hat{Y}_i\right)^2$
1	1,338	9,392	1,7902	2,395	3,205	16,814	9,387	0,005	0,00002
2	1,803	13,124	3,2508	5,861	10,568	42,664	13,217	-0,093	0,00862
3	0,844	5,932	0,7123	0,601	0,507	4,226	5,948	-0,016	0,00025
4	0,933	6,52	0,8705	0,812	0,758	5,676	6,520	$3{,}6 \cdot 10^{-4}$	$1{,}3 \cdot 10^{-7}$
5	0,604	4,492	0,3648	0,220	0,133	1,639	4,510	-0,018	0,00034
6	1,649	11,945	2,7192	4,484	7,394	32,481	11,885	0,060	0,00359
7	0,543	4,191	0,2948	0,160	0,087	1,236	4,169	0,022	0,00047
8	0,071	1,844	0,0050	$3{,}6 \cdot 10^{-4}$	$2{,}5 \cdot 10^{-5}$	0,009	1,865	-0,021	0,00042
9	1,349	9,477	1,8198	2,455	3,312	17,246	9,471	0,006	0,00003
10	0,919	6,336	0,8446	0,776	0,713	5,351	6,428	-0,092	0,00852
11	0,794	5,577	0,6304	0,501	0,397	3,516	5,636	-0,059	0,00344
12	1,909	14,199	3,6443	6,957	13,281	51,745	14,170	0,029	0,00084
13	0,257	2,717	0,0660	0,017	0,004	0,179	2,702	0,015	0,00022
14	1,817	13,332	3,3015	5,999	10,900	44,015	13,341	-0,009	0,00008
15	1,219	8,512	1,4860	1,811	2,208	12,649	8,500	0,012	0,00015
16	1,062	7,406	1,1278	1,198	1,272	8,353	7,386	0,020	0,00041
17	0,046	1,763	0,0021	$9{,}7 \cdot 10^{-5}$	$4{,}5 \cdot 10^{-6}$	0,004	1,759	0,004	0,00002
18	1,674	12,117	2,8023	4,691	7,853	33,955	12,097	0,020	0,00040
19	0,839	6,02	0,7039	0,591	0,496	4,238	5,916	0,104	0,01077
20	0,026	1,687	0,0007	$1{,}8 \cdot 10^{-5}$	$4{,}6 \cdot 10^{-7}$	0,001	1,676	0,011	0,00013
Σ	19,696	146,58	26,437	39,530	63,088	285,996			0,039

Рассчитаем значения коэффициентов системы уравнений (6)

$$g_{00} = 20;\ g_{01} = 19{,}696;\ g_{02} = 26{,}437;\ h_0 = 146{,}58;$$

$$g_{10} = 19{,}696;\ g_{11} = 26{,}437;\ g_{12} = 39{,}530;\ h_1 = 190{,}9;$$

$$g_{20} = 26{,}437;\ g_{21} = 39{,}530;\ g_{22} = 63{,}088;\ h_2 = 285{,}996\ ,$$

на основании которых расчитаем значения определителей (10)-(12)

$$\Delta = \begin{vmatrix} 20 & 19{,}696 & 26{,}437 \\ 19{,}696 & 26{,}437 & 39{,}53 \\ 26{,}437 & 39{,}53 & 63{,}088 \end{vmatrix} = 320{,}58;$$

$$\Delta_0 = \begin{vmatrix} 146{,}58 & 19{,}696 & 26{,}437 \\ 190{,}9 & 26{,}437 & 39{,}53 \\ 285{,}996 & 39{,}53 & 63{,}088 \end{vmatrix} = 502{,}99;$$

$$\Delta_1 = \begin{vmatrix} 20 & 146{,}58 & 26{,}437 \\ 19{,}696 & 190{,}9 & 39{,}53 \\ 26{,}437 & 285{,}996 & 63{,}088 \end{vmatrix} = 1304{,}3;$$

$$\Delta_2 = \begin{vmatrix} 20 & 19{,}696 & 146{,}58 \\ 19{,}696 & 26{,}437 & 190{,}9 \\ 26{,}437 & 39{,}53 & 285{,}996 \end{vmatrix} = 425{,}26.$$

По полученным значениям рассчитываем значения коэффициентов уравнения (4)

$$a_0 = \frac{\Delta_0}{\Delta} = \frac{502{,}99}{320{,}58} = 1{,}569;$$

$$a_1 = \frac{\Delta_1}{\Delta} = \frac{1304{,}3}{320{,}58} = 4{,}069;$$

$$a_2 = \frac{\Delta_2}{\Delta} = \frac{425{,}26}{320{,}58} = 1{,}327.$$

Строим диаграмму рассеивания экспериментальных данных и линию полученного уравнения регрессии $Y=1{,}569+4{,}069X+1{,}327X^2$ (рис. 3). Следует отметить, что экспериментальные точки достаточно плотно лежат на линии регрессии.

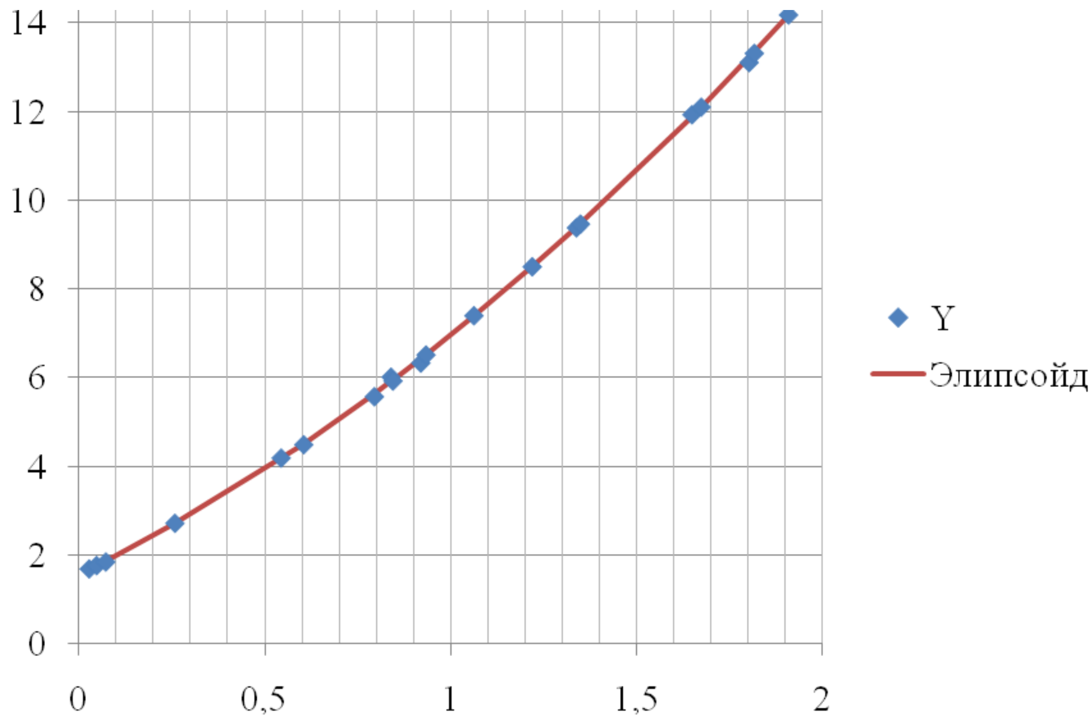

Рис. 3 - Диаграмма рассеивания экспериментальных данных и линия регрессии $Y=1{,}569+4{,}069X+1{,}327X^2$

Выполним оценку значимости нелинейной модели, расчитаем коэффициент детерминации

$$R^2 = 1 - \frac{0{,}039}{311{,}762} = 0{,}999$$

и значение F-критерия

$$F[1{,}18] = \frac{0{,}999}{1-0{,}999} \cdot \frac{18}{1} \approx 144\,985.$$

Т.к. $F[1{,}18]= 144\,985 > F_{кр}[1{,}18]= 4{,}41$, то полученное уравнение можно также признать значимым.

Сравнивая два уравнения, можно сделать вывод, что взаимосвязь между X и Y лучше описывается уравнением $Y=1{,}569+4{,}069X+1{,}327X^2$.

4 Выполнение работы

Получив исходные данные для выполнения практической работы (см. приложение Б), студент изучает теоретические сведения согласно пункту 3. Далее выполняет расчеты аналогичные в рассмотренном примере **с учетом имеющихся особенностей задания.**

В отчёте по практической работе должны найти отражение следующие пункты:

- название работы;
- цель работы;
- индивидуальное задание для выполнения работы;
- краткие теоретические сведения;
- результаты выполнения работы;
- подробные выводы по работе.

Контрольные вопросы

1. Что такое линейная регрессия?
2. Что такое многофакторная регрессия?
3. Что такое нелинейная регрессия?
4. Что рассчитывается при построении регрессионного уравнения?
5. Что такое критерий значимости регрессии?
6. Что такое коэффициент детерминации?
7. Какова интерпретация коэффициентов уравнения множественной регрессии?
8. Чем характеризуется точность уравнения регрессии?
9. Как осуществляется выбор «наилучшего» уравнения регрессии?

Библиографический список

1. Сергеев, А.Г. Метрология [Текст]/ А.Г. Сергеев, В.В. Крохин. Учебное пособие для вузов. М.: Логос, 2001. 488 с.: ил.

2. Алексахин, С.В. Прикладной статистический анализ [Текст]/ С.В. Алексахин, А.В. Балдин, А.Б. Николаев, В.Ю. Строганов. Учебное пособие для вузов. М.: “Издательство ПРИОР”, 2001. 224 с.

ПРИЛОЖЕНИЕ А

F-распределение (при q=0,05)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,85	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,41
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,69	2,51	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,37	2,20	2,01	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,31	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
31	4,16	3,30	2,91	2,68	2,52	2,41	2,25	2,08	1,88	1,61
35	4,12	3,27	2,87	2,64	2,49	2,37	2,22	2,04	1,83	1,56
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,20	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,47
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,53	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,33
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,30
100	3,94	3,09	2,70	2,46	2,31	2,19	2,03	1,85	1,63	1,28
∞	3,84	3,00	2,61	2,37	2,21	2,10	1,94	1,75	1,52	1,03

F-распределение (при q=0,1)

k_2	k_1 для больших дисперсий									
	1	2	3	4	5	6	8	12	24	∞
2	8,53	9,00	9,16	9,24	9,29	9,33	9,37	9,41	9,45	9,49
3	5,54	5,46	5,39	5,34	5,31	5,28	5,25	5,22	5,18	5,13
4	4,54	4,32	4,19	4,11	4,05	4,01	3,95	3,90	3,83	3,76
5	4,06	3,78	3,62	3,52	3,45	3,40	3,34	3,27	3,19	3,11
6	3,78	3,46	3,29	3,18	3,11	3,05	2,98	2,90	2,82	2,72
7	3,59	3,26	3,07	2,96	2,88	2,83	2,75	2,67	2,58	2,47
8	3,46	3,11	2,92	2,81	2,73	2,67	2,59	2,50	2,40	2,29
9	3,36	3,01	2,81	2,69	2,61	2,55	2,47	2,38	2,28	2,16
10	3,29	2,92	2,73	2,61	2,52	2,46	2,38	2,28	2,18	2,06
11	3,23	2,86	2,66	2,54	2,45	2,39	2,30	2,21	2,10	1,97
12	3,18	2,81	2,61	2,48	2,39	2,33	2,24	2,15	2,04	1,90
13	3,14	2,76	2,56	2,43	2,35	2,28	2,20	2,10	1,98	1,85
14	3,10	2,73	2,52	2,39	2,31	2,24	2,15	2,05	1,94	1,80
15	3,07	2,70	2,49	2,36	2,27	2,21	2,12	2,02	1,90	1,76
16	3,05	2,67	2,46	2,33	2,24	2,18	2,09	1,99	1,87	1,72
17	3,03	2,64	2,44	2,31	2,22	2,15	2,06	1,96	1,84	1,69
18	3,01	2,62	2,42	2,29	2,20	2,13	2,04	1,93	1,81	1,66
19	2,99	2,61	2,40	2,27	2,18	2,11	2,02	1,91	1,79	1,63
20	2,97	2,59	2,38	2,25	2,16	2,09	2,00	1,89	1,77	1,61
21	2,96	2,57	2,36	2,23	2,14	2,08	1,98	1,87	1,75	1,59
22	2,95	2,56	2,35	2,22	2,13	2,06	1,97	1,86	1,73	1,57
23	2,94	2,55	2,34	2,21	2,11	2,05	1,95	1,84	1,72	1,55
24	2,93	2,54	2,33	2,19	2,10	2,04	1,94	1,83	1,70	1,53
25	2,92	2,53	2,32	2,18	2,09	2,02	1,93	1,82	1,69	1,52
26	2,91	2,52	2,31	2,17	2,08	2,01	1,92	1,81	1,68	1,50
27	2,90	2,51	2,30	2,17	2,07	2,00	1,91	1,80	1,67	1,49
28	2,89	2,50	2,29	2,16	2,06	2,00	1,90	1,79	1,66	1,48
29	2,89	2,50	2,28	2,15	2,06	1,99	1,89	1,78	1,65	1,47
30	2,88	2,49	2,28	2,14	2,05	1,98	1,88	1,77	1,64	1,46
31	2,87	2,48	2,27	2,14	2,04	1,97	1,88	1,77	1,63	1,45
35	2,85	2,46	2,25	2,11	2,02	1,95	1,85	1,74	1,60	1,41
40	2,84	2,44	2,23	2,09	2,00	1,93	1,83	1,71	1,57	1,38
45	2,82	2,42	2,21	2,07	1,98	1,91	1,81	1,70	1,55	1,35
50	2,81	2,41	2,20	2,06	1,97	1,90	1,80	1,68	1,54	1,33
60	2,79	2,39	2,18	2,04	1,95	1,87	1,77	1,66	1,51	1,29
70	2,78	2,38	2,16	2,03	1,93	1,86	1,76	1,64	1,49	1,27
80	2,77	2,37	2,15	2,02	1,92	1,85	1,75	1,63	1,48	1,25
90	2,76	2,36	2,15	2,01	1,91	1,84	1,74	1,62	1,47	1,23
100	2,76	2,36	2,14	2,00	1,91	1,83	1,73	1,61	1,46	1,22
∞	2,71	2,30	2,08	1,95	1,85	1,77	1,67	1,55	1,38	1,03

ПРИЛОЖЕНИЕ Б

Исходные данные к работе

Таблица 1

	№ варианта														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
q	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05	0,1	0,05

Таблица 2

№ варианта									
№	1			2			3		
1	178,435	202,092	144,037	194,150	14,380	12,011	113,620	138,372	10,443
2	178,416	201,730	143,846	194,104	14,325	12,009	113,616	138,355	10,668
3	178,370	201,814	143,907	194,116	14,357	12,005	113,614	138,368	10,617
4	178,457	201,802	143,892	194,113	14,317	12,013	113,606	138,362	10,622
5	178,398	201,720	143,924	194,122	14,345	11,987	113,624	138,374	10,685
6	178,444	201,868	143,899	194,116	14,343	12,002	113,585	138,346	10,463
7	178,351	202,009	144,018	194,142	14,354	11,978	113,610	138,371	10,485
8	178,465	201,810	143,978	194,135	14,352	11,994	113,631	138,358	10,483
9	178,466	201,773	143,980	194,129	14,394	12,000	113,641	138,381	10,554
10	178,490	201,816	144,029	194,144	14,375	11,991	113,628	138,364	10,458
11	178,442	201,754	143,880	194,111	14,306	11,995	113,634	138,375	10,465
12	178,510	202,089	144,062	194,153	14,426	11,969	113,707	138,418	10,637
13	178,512	201,997	144,099	194,162	14,408	11,994	113,621	138,376	10,469
14	178,360	201,826	144,041	194,144	14,391	11,984	113,619	138,370	10,428
15	178,417	201,812	144,059	194,155	14,393	11,991	113,650	138,380	10,539

№ варианта									
№	4			5			6		
1	49,293	142,859	169,932	197,329	23,619	203,924	61,057	34,567	206,450
2	49,353	142,918	169,724	197,099	23,735	203,934	61,220	34,554	206,483
3	49,336	142,788	169,858	197,098	23,761	203,950	61,146	34,588	206,380
4	49,333	142,860	169,732	197,095	23,826	203,944	61,149	34,564	206,372
5	49,340	142,836	169,747	197,119	23,690	203,923	61,269	34,629	206,326
6	49,318	142,894	169,765	197,029	23,763	203,929	61,147	34,588	206,444
7	49,324	142,829	169,923	197,097	23,702	203,928	61,088	34,592	206,411
8	49,316	142,862	169,766	197,196	23,729	203,935	61,235	34,584	206,376
9	49,325	142,843	169,602	196,985	23,723	203,920	61,233	34,602	206,288
10	49,317	142,816	169,630	197,146	23,699	203,938	61,280	34,565	206,391
11	49,321	142,881	169,856	197,242	23,698	203,925	61,152	34,575	206,378
12	49,338	142,897	169,697	197,101	23,714	203,923	61,189	34,544	206,465
13	49,323	142,934	169,603	197,085	23,725	203,927	61,200	34,602	206,391

№ варианта									
№	4			5			6		
14	49,322	142,859	169,691	197,210	23,677	203,930	61,238	34,591	206,372
15	49,322	142,928	169,860	197,128	23,681	203,931	61,241	34,632	206,318

№ варианта									
№	7			8			9		
1	55,995	80,812	206,307	138,182	183,767	44,262	145,114	154,467	191,922
2	56,028	80,817	206,274	138,133	183,792	44,302	145,091	154,464	191,933
3	55,948	80,732	206,265	138,204	183,777	44,260	145,104	154,479	191,945
4	55,986	80,935	206,242	138,148	183,797	44,263	145,113	154,441	191,904
5	56,096	80,891	206,266	138,156	183,806	44,291	145,104	154,429	191,921
6	56,098	80,818	206,255	138,068	183,798	44,280	145,106	154,425	191,908
7	55,984	80,738	206,281	138,194	183,765	44,257	145,113	154,375	191,910
8	55,846	80,690	206,254	138,217	183,775	44,282	145,104	154,423	191,906
9	56,117	80,934	206,316	138,155	183,788	44,294	145,083	154,365	191,887
10	56,057	80,826	206,266	138,129	183,785	44,274	145,118	154,444	191,891
11	56,062	80,776	206,274	138,165	183,776	44,227	145,127	154,468	191,919
12	55,858	80,735	206,264	138,091	183,814	44,208	145,163	154,463	191,929
13	56,076	80,752	206,283	138,141	183,803	44,235	145,136	154,464	191,929
14	56,025	80,861	206,281	138,170	183,798	44,261	145,103	154,415	191,913
15	55,851	80,832	206,248	138,189	183,771	44,246	145,129	154,339	191,901

№ варианта									
№	10			11			12		
1	198,209	43,903	200,697	152,883	155,297	114,758	50,691	195,826	96,043
2	198,198	43,910	200,691	152,882	155,270	114,765	50,612	195,813	95,961
3	198,191	43,937	200,693	152,868	155,253	114,762	50,498	195,780	95,920
4	198,212	43,910	200,697	152,882	155,172	114,743	50,751	195,837	96,074
5	198,203	43,933	200,694	152,912	155,317	114,742	50,615	195,809	95,988
6	198,212	43,902	200,686	152,840	155,303	114,760	50,651	195,818	96,005
7	198,210	43,902	200,696	152,870	155,177	114,750	50,618	195,807	95,969
8	198,211	43,936	200,683	152,889	155,339	114,759	50,577	195,799	95,934
9	198,227	43,898	200,692	152,898	155,270	114,763	50,613	195,809	95,986
10	198,221	43,911	200,683	152,867	155,273	114,761	50,609	195,802	95,958
11	198,208	43,922	200,684	152,876	155,272	114,744	50,637	195,817	95,997
12	198,200	43,898	200,696	152,888	155,270	114,766	50,527	195,795	95,914
13	198,200	43,920	200,691	152,894	155,340	114,750	50,643	195,817	95,998
14	198,209	43,910	200,686	152,874	155,337	114,746	50,524	195,788	95,925
15	198,218	43,875	200,688	152,898	155,359	114,746	50,576	195,791	95,964

№ варианта									
№	13			14			15		
1	206,732	118,321	182,070	190,116	180,982	83,113	117,240	114,728	139,641
2	206,735	118,275	181,949	190,122	181,088	83,153	117,263	114,733	139,626
3	206,773	118,307	182,053	190,117	180,977	83,117	117,242	114,730	139,637
4	206,675	118,325	182,000	190,117	180,978	83,120	117,249	114,738	139,641
5	206,761	118,298	182,005	190,121	181,039	83,145	117,258	114,735	139,636
6	206,727	118,302	181,996	190,129	181,229	83,223	117,180	114,731	139,636
7	206,748	118,315	181,988	190,118	181,054	83,124	117,281	114,738	139,624
8	206,639	118,283	181,984	190,117	181,039	83,138	117,203	114,728	139,645
9	206,622	118,313	182,020	190,117	181,014	83,136	117,330	114,734	139,649
10	206,688	118,298	181,996	190,114	180,985	83,127	117,321	114,733	139,635
11	206,690	118,290	181,955	190,117	181,042	83,173	117,181	114,731	139,631
12	206,661	118,290	181,968	190,123	181,085	83,161	117,306	114,728	139,649
13	206,749	118,297	181,969	190,119	181,041	83,132	117,329	114,724	139,649
14	206,717	118,318	182,034	190,121	181,038	83,151	117,232	114,732	139,633
15	206,695	118,281	181,924	190,118	181,046	83,150	117,246	114,728	139,643