МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Государственное образовательное учреждение высшего профессионального образования
«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ
И РАДИОЭЛЕКТРОНИКИ» (ТУСУР))

Кафедра автоматизации обработки информации (АОИ)

УТВЕРЖДАЮ
Заведующий кафедрой АОИ
____________________Ю.П.Ехлаков
«___» ____________________ 2011 г.

Методические указания
к практическим и самостоятельным работам
по дисциплине

Эконометрика

Направление подготовки: **080700.62 «Бизнес-информатика»**

Форма обучения: **очная**

Факультет систем управления (ФСУ)

Кафедра автоматизации обработки информации (АОИ)

Курс 3 Семестр 5

Разработчик
Старший преподаватель кафедры АОИ И.В. Потахова

Практические занятия 26 часов
Самостоятельная работа 39 часов

2011

СОДЕРЖАНИЕ

Методические указания

Практические работы выполняются в рамках курса «Эконометрика», предусматривающего изучение методов проверки, обоснования, оценивания количественных закономерностей и качественных утверждений на основе анализа статистических данных. Кроме этого рассматриваются возможности применения Excel для решение означенных задач.. В работах предусмотрено выполнение ряда практических заданий.

Работы рекомендуется выполнять в порядке их следования.

По выполненным практическим работам студент отчитывается перед преподавателем. Отчет студента должен быть представлен выполненными заданиями и пояснениями по ходу их выполнения.

Тема 1. ЛИНЕЙНАЯ ПАРНАЯ РЕГРЕССИЯ

Эта тема включает выполнение практических работ, посвященных построению и исследованию уравнения линейной регрессии вида

$$\hat{y}_x = a + b \cdot x \qquad (1.1)$$

Пространственная выборка для построения этого уравнения взята из следующего примера.

Пример 1.1. Для определения зависимости между сменной добычей угля на одного рабочего (переменная *Y*, измеряемая в тоннах) и мощностью угольного пласта (переменная *X*, измеряемая в метрах) на 10 шахтах были проведены исследования, результаты которых представлены таблицей 1.1.

Таблица 1.1

i	1	2	3	4	5	6	7	8	9	10
x_i	8	11	12	9	8	8	9	9	8	12
y_i	5	10	10	7	5	6	6	5	6	8

1.1 Вычисление коэффициентов уравнения линейной регрессии

Цель. Вычисление коэффициентов уравнения линейной регрессии по пространственной выборке таб. 1.1.

Расчетные соотношения. Коэффициенты, определяемые на основе метода наименьших квадратов, являются решением системы уравнений

$$\begin{cases} b_0 + b_1 \cdot \overline{x} = \overline{y}; \\ b_0 \cdot \overline{x} + b_1 \cdot \overline{x^2} = \overline{xy}, \end{cases} \qquad \begin{cases} a + b \cdot \overline{x} = \overline{y} \\ a \cdot \overline{x} + b \cdot \overline{x^2} = \overline{x \cdot y} \end{cases} \qquad (1.2)$$

где

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i; \; \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i; \; \overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i \cdot y_i; \; \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2. \qquad (1.3)$$

Решая эту систему уравнений, получаем

$$b = \frac{\overline{x \cdot y} - \overline{x} \cdot \overline{y}}{\overline{x^2} - (\overline{x})^2} = \frac{\mathrm{cov}(x, y)}{\sigma_x^2} \qquad (1.4)$$

$$a = \overline{y} - b \cdot \overline{x} \quad , \qquad (1.5)$$

где *cov (x, y)*– выборочное значение корреляционного момента (ковариация), определенного по формуле:

$$\mathrm{cov}(x, y) = \overline{x \cdot y} - \overline{x} \cdot \overline{y}, \qquad (1.6)$$

σ_x^2 – выборочное значение дисперсии величины *X*, определяемой по формуле:

$$\sigma_x^2 = \overline{x^2} - (\overline{x})^2 \qquad (1.7)$$

Задание 1. Вычислим эти коэффициенты a, b, используя табличный процессор Excel (версия XP). Для этого выполнить следующие действия:

а) разместить на листе книги данные таблицы 1;

б) запрограммировать вычисление коэффициентов $\overline{x}, \overline{y}$, $\overline{x^2}, \overline{xy}$ системы (1.2);

в) запрограммировать вычисление *a*, *b* по формулам (1.4), (1.5) соответственно.

Заметим, что для вычисления средних значений используется функция Excel СРЗНАЧ(*диапазон ячеек*).

В результате выполнения запрограммированных вычислений получаем a = –2.75; b = 1.016, а само уравнение регрессии (1.1) примет вид

$$\hat{y}(x) = -2.75 + 1.016x. \qquad (1.8)$$

Задание 2. Используя уравнение (1.8), определите производительность труда шахтера, если толщина угольного слоя равна:

а) 8.5 метров (интерполяция данных);

б) 14 метров (экстраполяция данных).

1.2. Вычисление выборочного коэффициента корреляции

Цель. Вычисление выборочного коэффициента корреляции по пространственной выборке таб. 1.1.

Расчетные соотношения. Выборочный коэффициент корреляции определяется соотношением

$$r_{xy} = \frac{\operatorname{cov}(x, y)}{\sigma_x \cdot \sigma_y}, \qquad (1.9)$$

где $\sigma_x^2 = \overline{x^2} - \bar{x}^2$, $\sigma_y^2 = \overline{y^2} - \bar{y}^2$, $\overline{y^2} = \frac{1}{n}\sum_{i=1}^{n} y_i^2$. (1.10)

1.3. Вычисление оценок дисперсий коэффициентов парной линейной регрессии

Цель. Вычислить оценки $S^2_{общ}$, $S^2_{факт}$, $S^2_{ост}$, S^2_a, S^2_b для дисперсий (квадрата стандартных ошибок) коэффициентов a, b.

Расчетные соотношения. Оценки для дисперсий определяются формулами:

$$S^2_{общ} = \frac{\sum_{i=1}^{n}(y - \bar{y})^2}{n-1}, \qquad S^2_{факт} = \frac{\sum_{i=1}^{n}(\hat{y}_x - \bar{y})^2}{m},$$

$$S^2_{ост} = \frac{\sum_{i=1}^{n}(y - \hat{y}_x)^2}{n-m-1},$$

$$S^2_a = S^2_{ост} \cdot \frac{\overline{x^2}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \qquad S^2_b = S^2_{ост} \cdot \frac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}, \qquad (1.11)$$

где m = 1 для парной линейной регрессии

1. 4. Функции Excel для вычисления коэффициентов парной линейной регрессии

Цель работы. Вычислить коэффициенты уравнения линейной регрессии по пространственной выборке таб. 1.1, используя функции Excel.

Функции Excel. Приведем некоторые статистические функции Excel, полезные при построении парной линейной регрессии.

Функция ОТРЕЗОК. Вычисляет коэффициент b_0 и обращение имеет вид

ОТРЕЗОК(*диапазон_значений_y ; диапазон_значений_x*).

Функция НАКЛОН. Вычисляет коэффициент b_1 и обращение имеет вид

НАКЛОН(*диапазон_значений_ y ; диапазон_значений_ x*).

Функция ПРЕДСКАЗ. Вычисляет значение линейной парной регрессии при заданном значении независимой переменной (обозначена через z) и обращение имеет вид

ПРЕДСКАЗ(z ;*диапазон_значений_ y ;диапазон_значений_ x*).

Функция СТОШYX. Вычисляет оценку S для среднеквадратического отклонения σ возмущений ε_i и обращение имеет вид (YX – латинские буквы):

СТОШYX(*диапазон_значений_ y ; диапазон_значений_ x*).

Решение. Фрагмент документа Excel, вычисляющего требуемые величины приведен на рисунке. Обратите внимание на использовании абсолютной адресации при вычислении $\hat{y}_i$.

	A	B	C
1	**Исходные данные**		
2	x_i	y_i	$\hat{y}_i$
3	8	5	5,377
4	11	10	8,428
5	12	10	9,443
6	9	7	6,393
7	8	5	5,377
8	8	6	5,377
9	9	6	6,393
10	9	5	6,393
11	8	6	5,377
12	12	8	9,443
13			
14		a	-2,754
15		b	1,016393
16		$S_{ост}$	1,024295

=ПРЕДСКАЗ(A3;$B$3:$B$12;$A$3:$A$12)

=ОТРЕЗОК(B3:B12;A3:A12)

=НАКЛОН(B3:B12;A3:A12)

=СТОШYX(B3:B12;A3:A12)

Рис. Использование функций Excel

Задание. Сравните вычисленные значения с значениями, полученными на предыдущих шагах выполнения работы.

1.5. Построение интервальной оценки для коэффициентов регрессии, функции парной линейной регрессии

Задание 1. Построить интервальные оценки для коэффициентов регрессии, прогнозного значения $\hat{y}_p$ с надежностью γ = 0.95, используя для этого уравнение регрессии $\hat{y}(x)$, построенное в практической работе № 1.

Расчетные соотношения.

- Интервальная оценка (доверительный интервал) для коэффициента b с надежностью (доверительной вероятностью) равной γ определяется выражением:

$b \pm t_{табл} \cdot m_b$

Фактическое значение t -критерия Стьюдента: $t_b = \dfrac{b}{m_b}$

Стандартная ошибка коэффициента (b) регрессии определяется по формуле:

$$m_b = \frac{S_{ост}}{\sigma_x \cdot \sqrt{n}}, \qquad (1.12)$$

где $S^2_{ост} = \frac{\sum_{i=1}^{n}(y - \hat{y}_x)^2}{n - m - 1}$, – остаточная дисперсия на одну степень свободы,

$\sigma_x^2 = \overline{x^2} - \overline{x}^2$ - дисперсия признака x.

Аналогично строится интервальная оценка параметра a. При этом используются следующая расчетная формула вычисления стандартной ошибки коэффициента a:

$$m_a = \frac{S_{ост} \cdot \sqrt{\sum_{i=1}^{n} x_i^2}}{\sigma_x \cdot n} \qquad (1.13)$$

- Интервальная оценка (доверительный интервал) для прогнозного значения $\hat{y}_p$ при заданном значении x_p с надежностью (доверительной вероятностью) равной $\gamma = 1 - \alpha$ определяется выражением

$$\hat{y}_p \pm t_{табл} \cdot m_{\hat{y}_p} \qquad (1.14)$$

Стандартная ошибка прогнозного значения определяется по формуле:

$$m_{\hat{y}_p} = S_{ост} \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_p - \overline{x})^2}{n \cdot \sigma_x^2}} \qquad (1.15)$$

Таким образом, в (1.14) входят две величины $m_{\hat{y}_p}$ (зависит от x_p) и $t(\gamma, n-2)$, вычисляемая с помощью функции Excel:

$t(\gamma, n-2)$ =СТЬЮДРАСПОБР($1-\gamma; n-2$).

- Аналогично строится интервальная оценка регрессии $f(x) = M(Y \mid x)$ в заданных точках x.

Задание 2. Вычислить значения нижней y_i^H и верхней y_i^B границ интервала для $x = x_i, i = 1,...,10$. Данные расчетов записать в таблицу.

1.6 Проверка значимости уравнения линейной регрессии по критерию Фишера

Задание. Оценить на уровне α = 0.05 значимость уравнения регрессии $\hat{y}(x) = -2.75 + 1.016 \cdot x$, построенного в практической работе № 1

Расчетные соотношения. Уравнение парной регрессии значимо с уровнем значимости α, если выполняется следующее неравенство:

$$F = \frac{S^2_{факт}}{S^2_{ост}} > F_{\alpha;1;n-2} \qquad (1.16)$$

где $F_{\alpha;1;n-2}$ – значения квантиля уровня α F-распределения с числами степеней свободы k_1 = 1 и k_2 = n – 2. Для вычисления квантиля можно использовать следующую функциюExcel

$F_{\alpha;1;n-2}$ = FРАСПОБР($\alpha; 1; n-2$) (1.17)

1.7 Вычисление средней ошибки аппроксимации

Расчетные соотношения. Средняя ошибка аппроксимации вычисляется по формуле:

$$\overline{A} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y-\hat{y}_x}{y}\right| \cdot 100\%$$

1.8. Сделать выводы по проделанной работе

Тема 2. НЕЛИНЕЙНАЯ ПАРНАЯ РЕГРЕССИЯ

Пример 2.1. В таблице 2.1 приведены значения независимой переменной X (доход американской семьи в тысячах долларов) и значения зависимой переменной Y (доля расходов на товары длительного пользования в процентах от общей суммы расходов).

Таблица 2.1

x_i	1	2	3	4	5	6
y_i	10	13.4	15.4	16.5	18.6	19.1

2.1 Построение нелинейной регрессии с использованием команды «Добавить линию тренда»

Цель работы. Используя пространственную выборку таблицы 2.1 необходимо построить уравнение нелинейной регрессии вида $\hat{y} = b_0 \cdot x^{b_1}$ с использованием команды «Добавить линию тренда» и вычислить коэффициент детерминации R^2.

Команда «Добавить линию тренда». Используется для выделения тренда (медленных изменений) при анализе временных рядов. Однако эту команду можно использовать и для построения уравнения нелинейной регрессии, рассматривая в качестве времени t независимую переменную x.

Эта команда позволяет построить следующие уравнения регрессии:

- линейную $\hat{y} = b_0 + b_1 x$
- полиноминальную $\hat{y} = b_0 + b_1 x + \ldots + b_k x^k$ ($k \le 6$);
- логарифмическую $\hat{y} = b_0 + b_1 \cdot \ln x$
- степенную $\hat{y} = b_0 \cdot x^{b_1}$;
- экспоненциальную $\hat{y} = b_0 e^{b_1 x}$.

Для построения одной из перечисленных регрессий необходимо выполнить следующие шаги:

Шаг 1. В выбранном листе Excel ввести по столбцам исходные данные $\{x_i, y_i\}, i = 1,2,\ldots,n$ (см. рис. 2.1).

Шаг 2. По этим данным построить график в декартовый системе координат (см. рис 2.1).

Шаг 3. Установить курсор на построенном графике, сделать щелчок правой кнопкой и в появившемся контекстном меню выполнить команду *Добавить линию тренда* (см. рис. 2.1).

Шаг 4. В появившемся диалоговом окне (см. рис. 2.2) активизировать закладку «Тип» и выбрать нужное уравнение регрессии.

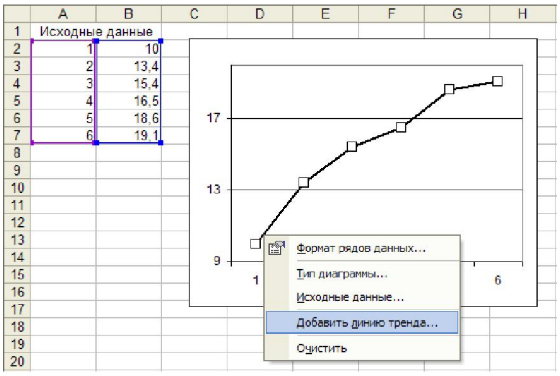


Рис. 2.1. Построение графика по исходным данным

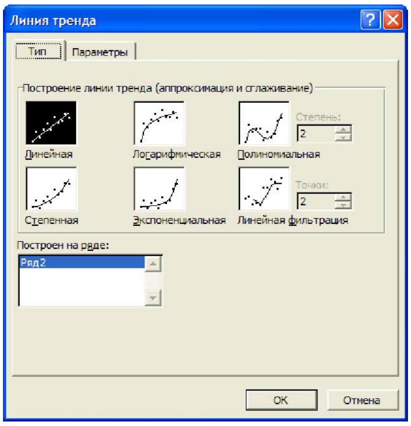


Рис. 2.2. Выбор вида уравнения регрессии

Шаг 5. Активизировать закладку «Параметры» (см. рис. 2.3) и «включить» *необходимые для нас опции:*

- *«Показать уравнение на диаграмме»* - на диаграмме будет показано выбранное уравнение регрессии с вычисленным коэффициентами;

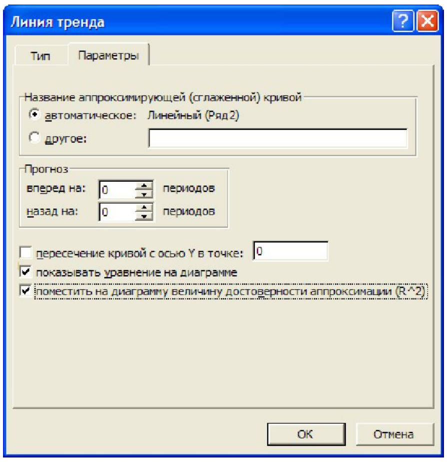


Рис. 2.3. Задание опций вывода информации

- *«Поместить на диаграмму величину достоверности аппроксимации (R^2)»* - на диаграмме будет показано значение коэффициент детерминации ρ^2 (для нелинейной регрессии - индекс детерминации), вычисляемый по формуле $\rho_{xy}^2 = 1 - \dfrac{\sigma_{ост}^2}{\sigma_y^2}$,

где $\sigma_{ост}^2 = \dfrac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_{x_i})^2$, $\quad \sigma_y^2 = \dfrac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$

Если по построенному уравнению регрессии необходимо выполнить прогноз, то нужно указать число периодов прогноза (см. рис. 2.3).

Назначение других опций понятны из своих названий.

Шаг 6. После задания всех перечисленных опций щелкнуть на кнопке «ОК» и на диаграмме появиться формула построенного уравнения регрессии и значение индекса детерминации ρ_{xy}^2 (выделено на рис. 2.4 затемнением).

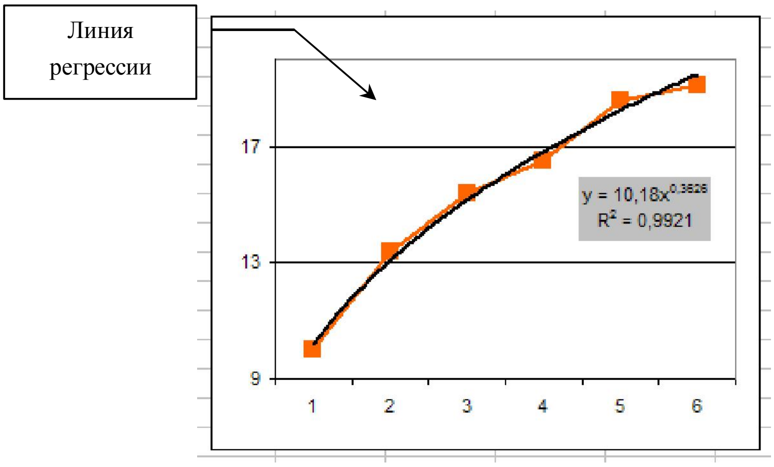


Рис. 2.4. График и уравнение построенной регрессии

Решение. Построение уравнения $\hat{y} = b_0 \cdot x^{b_1}$ осуществляем по описанным выше шагам. Получаем уравнение

$$\hat{y}(x) = 10.18 \cdot x^{0.3626},$$

для которого коэффициент детерминации равен ρ_{xy}^2 =0.9921 (см. рис. 2.4). Такая величина говорит о хорошем соответствии построенного уравнения исходным данным.

2.2 Выбор наилучшей нелинейной регрессии по приведенному индексу детерминации

Цель работы. Используя пространственную выборку таблицы 2.1 и команду «Добавить линию тренда» построить шесть уравнений нелинейной регрессии (полиномиальное уравнение строится при $m = 2$ и $m = 3$), определить для каждого уравнения индекс детерминации ρ_{xy}^2 (значение выводится), приведенный индекс детерминации $\hat{\rho}^2$ (значение вычисляется) и по максимальному значению $\hat{\rho}^2$ найти наилучшее уравнение нелинейной регрессии.

Приведенный индекс детерминации. Индекс детерминации ρ_{xy}^2 характеризует близость построенной регрессии к исходным данным, которые содержат «нежелательную» случайную составляющую ε. Очевидно, что, построив по данным таб. 2.1 полином 5-ого порядка, получаем «идеальное» значение ρ_{xy}^2 =1, но такое уравнение содержит в себе не только независимую переменную x, но составляющую ε и это снижает точность использования построенного уравнения для прогноза. Поэтому при выборе уравнения регрессии надо учитывать не только величину ρ_{xy}^2, но и «сложность» регрессионного уравнения, определяемое количеством коэффициентов уравнения. Такой учет удачно реализован в так называемом *приведенном индексе детерминации:*

$$\hat{\rho}^2 = 1 - \frac{(n-1)\cdot\sigma^2_{ост}}{(n-m)\cdot\sigma^2_y} = 1 - \frac{n-1}{n-m}\cdot(1-\rho^2_{xy}), \qquad (2.1)$$

где m - количество вычисляемых коэффициентов регрессии. Видно, что при неизменных $\sigma^2_{ост}$, σ^2_y увеличение m уменьшает значение $\hat{\rho}^2$. Если количество коэффициентов у сравниваемых уравнений регрессии одинаково (например, $m = 2$), то отбор наилучшей регрессии можно осуществлять по величине ρ^2_{xy}. Если в уравнениях регрессии меняется число коэффициентов, то такой отбор целесообразно по величине $\hat{\rho}^2$.

Решение. Для построения каждого уравнения выполняем шаги 2 – 6 (для первого уравнения еще и шаг 1) и размещаем в одном документе шесть окон, в которых выводятся найденные уравнения регрессии уравнения и величина ρ^2_{xy}. Затем формулу уравнения и ρ^2_{xy} заносим в таблицу 2.2. Далее по формуле (2.1) вычисляем приведенный индекс детерминации $\hat{\rho}^2$ и заносим эти значения также в таблицу (см. таб. 2.2).

Таблица 2.2

№	**Уравнение**	ρ^2_{xy}	$\hat{\rho}^2$
1	$\hat{y} = 9.28 + 1.777x$	0.949	0.938
2	$\hat{y} = 9.8759 + 5.1289\cdot\ln x$	0.9916	0.9895
3	$\hat{y} = 6.93 + 3.5396x - 0.2518x^2$ (полиноминальная, $m = 2$)	0.9896	0.9827
4	$\hat{y} = 5.8333 + 4.9192\cdot x - 0.7087\cdot x^2 - 0.0435\cdot x^3$ (полиноминальная, $m = 3$)	0.9917	0.9792
5	$\hat{y} = 10.18x^{0.3626}$	0.9921	0.9901
6	$\hat{y} = 9.8675\cdot e^{0.1225x}$	0.9029	0,8786

В качестве «наилучшего» уравнения регрессии выбираем уравнение, имеющее наибольшую величину приведенный коэффициент детерминации $\hat{\rho}^2$. Из таб. 2.2 видно, что таким уравнением является степенная функции (в таблице строка с этой функцией выделена серым цветом)

$$\hat{y} = 10.18x^{0.3626},$$

имеющая величину $\hat{\rho}^2 = 0.9901$.

Задание.

1. Используя статистические данные по численности населения России выполнить построение «наилучшей» модели парной регрессии. Оценить численность населения в 2000 году.

Год	1960	1970	1980	1990	1991	1992	1993	1994	1995	2000
Численность стат., млн. чел.	117,5	**130,2**	137,6	147,4	148,5	147,7	148,7	148,4	148,3	?

2. Введя дополнительное данное: значение численности населения России в 1998 году 146,2 млн. человек, уточнить экстраполяцию, используя только данные 90-х годов.

Тема 3. ЛИНЕЙНАЯ МНОЖЕСТВЕННАЯ РЕГРЕССИЯ

Эта тема включает выполнение практических работ, посвященных построению и исследованию уравнения линейной множественной регрессии вида

$$\hat{y}(x_1, x_2) = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 \qquad (3.1)$$

Пространственная выборка для построения этого уравнения взята из следующего примера.

Пример 3.1. Данные о сменной добыче угля на одного рабочего (переменная Y – измеряется в тоннах), мощности пласта (переменная X_1 – измеряется в метрах) и уровнем механизации работ в шахте (переменная X_2 – измеряется в процентах), характеризующие процесс добычи угля в 10 шахтах приведены в таблице 3.1.

Предполагая, что между переменными Y, X_1, X_2 существует линейная зависимость, необходимо найти аналитическое выражение для этой зависимости, т.е. построить уравнение линейной регрессии.

Таблица 3.1

Номер шахты i	x_{i1}	x_{i2}	y_i
1	8	5	5
2	11	8	10
3	12	8	10
4	9	5	7
5	8	7	5
6	8	8	6
7	9	6	6
8	9	4	5
9	8	5	6
10	12	7	8

3.1 Вычисление коэффициентов линейной множественной регрессии

Цель работы. Используя пространственную выборку таблицы 3.1 необходимо вычислить вектор коэффициентов $b = \begin{vmatrix} b_0 \\ b_1 \\ b_2 \end{vmatrix}$ уравнения регрессии (3.1).

Расчетные соотношения. Вектор коэффициентов, найденный методом наименьших квадратов является решением следующей системы уравнений:

$$X^T X b = X^T y,$$

где X - матрица размера 10×3, первый столбец которой составлен из 1, а другие два столбца составлены из значений x_{i1}, x_{i2}, т.е. матрица X имеет следующую структуру (символы …

означают не отображенные элементы)

$$X = \begin{vmatrix} 1 & 8 & 5 \\ 1 & 11 & 8 \\ \cdots & \cdots & \cdots \\ 1 & 12 & 7 \end{vmatrix},$$

а y - вектор, составленный из 10 значений y_i, т.е.

$$y = \begin{vmatrix} 5 \\ 10 \\ \vdots \\ 8 \end{vmatrix}.$$

Матрица $X^T X$ имеет обратную матрицу $\left(X^T X\right)^{-1}$ и тогда вектор коэффициентов вычисляется в виде:

$$b = A^{-1}(X^T y). \qquad (3.2)$$

Матричные функции Excel. Для реализации этой матричной формулы в необходимо выполнить следующие операции: транспонирование; умножение матриц (частный случай – умножение матрицы на вектор); вычисление обратной матрицы. Все эти операции можно реализовать с помощью следующих *матричных функций Excel*. Для работы с этими функциями можно или а) обратиться к ***Мастеру функций*** и выбрать нужную категорию функций, затем указать имя функции и задать соответствующие диапазоны ячеек, или б) ввести с клавиатуры имя функции задать соответствующие диапазоны ячеек.

Транспонирование матрицы осуществляется с помощью функции ТРАНСП (категория функций – *Ссылки и массивы*). Обращение к функции имеет вид:

ТРАНСП (*диапазон ячеек*),

где параметр *диапазон ячеек* задает все элементы транспонируемой матрицы (или вектора).

Умножение матриц осуществляется с помощью функции МУМНОЖ (категория функций – *Математические*). Обращение к функции имеет вид:

МУМНОЖ(*диапазон_1;диапазон_2*),

где параметр *диапазон_1* задает элементы первой из перемножаемых матриц, а параметр *диапазон_2* – элементы второй матрицы. При этом перемножаемые матрицы должны иметь соответствующие размеры (если первая матрица $n \times k$, вторая - $k \times m$, то результатом будет матрица $n \times m$).

Обращение матрицы (вычисление обратной матрицы) осуществляется с помощью функции МОБР (категория функций – *Математические*). Обращение к функции имеет вид:

МОБР (*диапазон ячеек*),

где параметр *диапазон ячеек* задает все элементы обращаемой матрицы, которая должна быть квадратной и невырожденной.

При использовании этих функций ***необходимо соблюдать следующий порядок действий:***

- *выделить фрагмент ячеек*, в которые будет занесен результат выполнения матричных функций (при этом надо учитывать размеры исходных матриц);
- *ввести арифметическое выражение*, содержащее обращение к матричным функциям Excel;

- *одновременно нажать клавиши* [Ctrl], [Shift], [Enter]. Если этого не сделать, то *вычислится только один элемент* результирующей матрицы или вектора.

Решение. Сформируем матрицу X и вектор y (см. рис. 3.1).

	A	B	C	D	E	F
1		***1***	***8***	***5***		***5***
2		***1***	***11***	***8***		***10***
3		***1***	***12***	***8***		***10***
4		***1***	***9***	***5***		***7***
5	$X =$	***1***	***8***	***7***	$y =$	***5***
6		***1***	***8***	***8***		***6***
7		***1***	***9***	***6***		***6***
8		***1***	***9***	***4***		***5***
9		***1***	***8***	***5***		***6***
10		***1***	***12***	***7***		***8***
11						
12		***10***	***94***	***63***		***68***
13	$X^T \cdot X =$	***94***	***908***	***603***	$X^T y =$	***664***
14		***63***	***603***	***417***		***445***
15	=МУМНОЖ(ТРАНСП(B1:D10);B1:D10)					
16		=МУМНОЖ(ТРАНСП(B1:D10);F1:F10)				
17		***4,0201***	***-0,323***	***-0,14***		***-3,5393***
18	$(X^T \cdot X)^{-1} =$	***-0,323***	***0,054***	***-0,029***	$b =$	***0,8539***
19		***-0,14***	***-0,029***	***0,0653***		***0,3670***
20						
21	=МОБР(B12:D14)			=МУМНОЖ(B17:D19;F12:F14)		

Рис. 3.1. Вычисление коэффициентов множественной регрессии

Затем выполним формирование матрицы $X^T X$, вектора $X^T y$ и вычисление вектора $b = |b_0, b_1, b_2|^T$ по формуле (3.2). Все эти вычисления показаны на рис. 3.1.

Получен вектор коэффициентов $b = \begin{vmatrix} -3.5393 \\ 0.8539 \\ 0.3670 \end{vmatrix}$ и тогда уравнение регрессии (3.1) примет вид:

$$\hat{y}(x_1, x_2) = -3.54 + 0.854 x_1 + 0.367 x_2 \ . \qquad (3.3)$$

3.2 Вычисление коэффициентов линейной множественной регрессии и проверка значимости в режиме *Регрессия*

Цель работы. Используя пространственную выборку таблицы 3.1 и используя режим ***Регрессия*** необходимо вычислить вектор коэффициентов уравнения регрессии

$$\hat{y}(x_1, x_2) = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 \ .$$

Режим Регрессия модуля Анализ данных. Табличный процессор Excel содержит модуль ***Анализ данных.*** Этот модуль позволяет выполнить статистический анализ выборочных данных (построение гистограмм, вычисление числовых характеристик и т.д.). Режим работы *Регрессия* этого модуля осуществляет вычисление коэффициентов линейной множественной регрессии с k переменными, построение доверительные интервалы и проверку значимости уравнения регрессии.

Для вызова режима *Регрессия* модуля ***Анализ данных*** необходимо:

- обратиться к пункту меню **Сервис**;
- в появившемся меню выполнить команду ***Анализ данных;***

- в списке режимов работы модуля ***Анализ данных*** выбрать режим *Регрессия* и щелкнуть на кнопке ***Ok***.

После вызова режима *Регрессия* на экране появляется диалоговое окно (см. рис. 3.2), в котором задаются следующие параметры:

1. Входной интервал Y – вводится диапазон адресов ячеек, содержащих значения y_i (ячейки должны составлять один столбец).

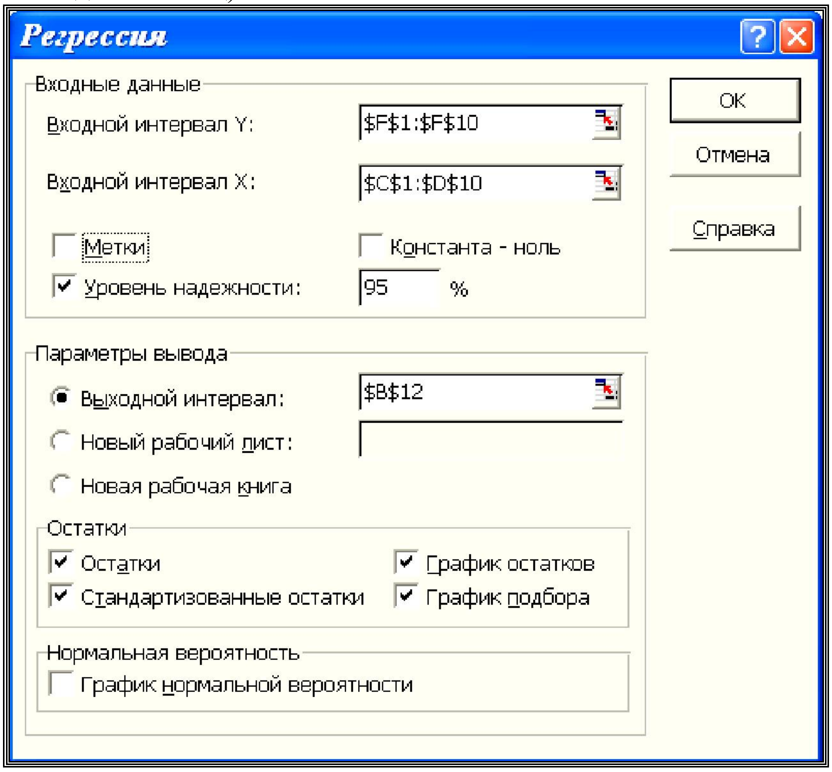


Рис. 3.2. Диалоговое окно режима *Регрессия*

2. Входной интервал X – вводится диапазон адресов ячеек, содержащих значения независимых переменных. Значения каждой переменной представляются одним столбцом. Количество переменных не более 16 (т.е. $k \le 16$).

3. Метки – включается если первая строка во входном диапазоне содержит заголовок. В этом случае автоматически будут созданы стандартные названия.

4. Уровень надежности – при включении этого параметра задается надежность γ при построении доверительных интервалов.

5. Константа-ноль – при включении этого параметра коэффициент $b_0 = 0$.

6. Выходной интервал – при включении активизируется поле, в которое необходимо ввести адрес левой верхней ячейки выходного диапазона, который содержит ячейки с результатами вычислений режима *Регрессия.*

7. Новый рабочий лист – при включении этого параметра открывается новый лист, в который начиная с ячейки A1 вставляются результаты работы режима *Регрессия.*

8. Новая рабочая книга - при включении этого параметра открывается новая книга на первом листе которой начиная с ячейки A1 вставляются результаты работы режима *Регрессия.*

9. Остатки – при включении вычисляется столбец, содержащий невязки $y_i - \hat{y}_i, i = 1,...,n$.

10. Стандартизованные остатки – при включении вычисляется столбец, содержащий стандартизованные остатки.

11. График остатков – при включении выводятся точечные графики невязки $y_i - \hat{y}_i, i = 1,...,n$, в зависимости от значений переменных $x_j, j = 1,...,k$. Количество графиков равно числу k переменных x_j .

12. График подбора – при включении выводятся точечные графики предсказанных по построенной регрессии значений $\hat{y}_i$ от значений переменных $x_j, j = 1,...,k$. Количество графиков равно числу k переменных x_j.

Решение. Первоначально введем в столбец C десять значений первой переменной, в столбец D - десять значений первой переменной (см. рис. 3.2), а в столбец F – десять значений зависимой переменной.

После этого вызовем режим *Регрессия* и в диалоговом окне зададим необходимые параметры (см. рис. 3.2). Результаты работы приводятся рис. 3.3 – 3.5. Заметим, из-за большой «ширины» таблиц, в которых выводятся результаты работы режима *Регрессия,* часть результатов помещены в другие ячейки.

ВЫВОД ИТОГОВ			
Регрессионная статистика			
Множественный R	0,9009		
R-квадрат	0,8116		
Нормированный R-квадрат	0,7578		
Стандартная ошибка	0,9509		
Наблюдения	10		
Дисперсионный анализ			
	df	*SS*	*MS*
Регрессия	2	27,2704	13,635
Остаток	7	6,3296	0,904
Итого	9	33,6000	
		F	*Значимость F*
		15,0794	0,0029

Рис. 3.3. Результаты работы режима *Регрессия*

Дадим краткую интерпретацию показателям, значения которых вычисляются в режиме *Регрессия.* Первоначально рассмотрим показатели, объединенные названием *Регрессионная статистика* (см. рис. 3.3).

Множественный R - корень квадратный из коэффициента детерминации.

R *– квадрат* – коэффициент детерминации R^2.

Нормированный R *– квадрат* – приведенный коэффициент детерминации $\hat{R}^2$.

Стандартная ошибка – оценка s для среднеквадратического отклонения σ.

Наблюдения – число наблюдений n.

Перейдем к показателям, объединенным названием *Дисперсионный анализ* (см. рис. 3.3).

Столбец df - число степеней свободы. Для строки *Регрессия* показатель равен числу независимых переменных $k_r = k = m - 1$; для строки *Остаток* – равен $k_e = n - (k_r + 1) = n - m$; для строки *Итого* – равен $k_r + k_e$.

Столбец SS – сумма квадратов отклонений. Для строки *Регрессия* показатель равен величине факторной суммы квадратов

$$SS_r = \sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2 ;$$

для строки *Остаток* - равен величине остаточной суммы квадратов

$$SS_r = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 ;$$

для строки *Итого* – $SS = SS_r - SS_e$— общая сумма квадратов отклонений переменной y от среднего значения $\overline{y}$.

Столбец MS – дисперсии, вычисленные по формуле

$$MS = \frac{SS}{df},$$

т.е. дисперсия на одну степень свободы.

Столбец F – значение F_c, равное F – критерию Фишера, вычисленного по формуле:

$$F_c = \frac{SS_r / k_r}{SS_e / k_e} .$$

Столбец значимость F - значение уровня значимости, соответствующее вычисленной величине F – критерия и равное вероятности $P(F(k_r, k_e) \geq F_c)$, где $F(k_r, k_e)$- случайная величина, подчиняющаяся распределению Фишера с k_r, k_e степенями свободы. Эту вероятность можно также определить с помощью функции FРАСП($F_c; k_r; k_e$). *Если вероятность меньше уровня значимости* α *(обычно* $\alpha = 0.05$*), то построенная регрессия является значимой..*

Перейдем к следующей группе показателей, объединенных в таблице, показанной на рис. 3.4.

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
Y-пересечение	-3,539	1,907	-1,8564	0,1058	-8,0477	0,9690
Переменная X 1	0,854	0,221	3,8726	0,0061	0,3325	1,3753
Переменная X 2	0,367	0,243	1,5108	0,1746	-0,2074	0,9415

Рис. 3.4. Продолжение результатов работы режима *Регрессия*

Столбец Коэффициенты – вычисленные значения коэффициентов $b_0, b_1, ..., b_k$, расположенных сверху-вниз.

Столбец Стандартная ошибка – значения m_{b_j} , *j=0,...k*, вычисленные по формуле .

Столбец $t-$*статистика* – значения статистик T_{b_j} .

Столбец P *– значение* – содержит вероятности случайных событий $P(t(n-m) \geq T_{b_j})$, где

$t(n-m)$ – случайная величина, подчиняющаяся распределению Стьюдента с $n-m$ степенями свободы.

Если эта вероятность меньше уровня значимости α, *то принимается гипотеза о значимости соответствующего коэффициента регрессии.*

Из рис. 3.4 видно, что значимым коэффициентом является только коэффициент b_1.

Столбцы Нижние 95% и Верхние 95% - соответственно нижние и верхние интервалы для оцениваемых коэффициентов b_j.

Перейдем к следующей группе показателей, объединенных в таблице, показанной на рис. 3.5.

ВЫВОД ОСТАТКА			
Наблюдение	**Предсказанное Y**	**Остатки**	**Стандартные остатки**
1	5,127	-0,127	-0,152
2	8,790	1,210	1,443
3	9,644	0,356	0,424
4	5,981	1,019	1,215
5	5,861	-0,861	-1,027
6	6,228	-0,228	-0,272
7	6,348	-0,348	-0,415
8	5,614	-0,614	-0,732
9	5,127	0,873	1,041
10	9,277	-1,277	-1,523

Рис. 3.5. Продолжение результатов работы режима *Регрессия*

Столбец Наблюдение – содержит номера наблюдений.

Столбец Предсказанное У – значения $\hat{y}_i$, вычисленные по построенному уравнению регрессии.

Столбец Остатки – значения невязок $y_i - \hat{y}_i$

В заключении рассмотрения результатов работы режима *Регрессия* приведем график невязок (на рисунке 3.6 невязки названы остатками) $y_i - \hat{y}_i$ при заданных значениях только второй переменной. Наличие чередующихся положительных и отрицательных значений невязок является косвенным признаком *отсутствия систематической ошибки* (неучтенной независимой переменной) в построенном уравнении регрессии.

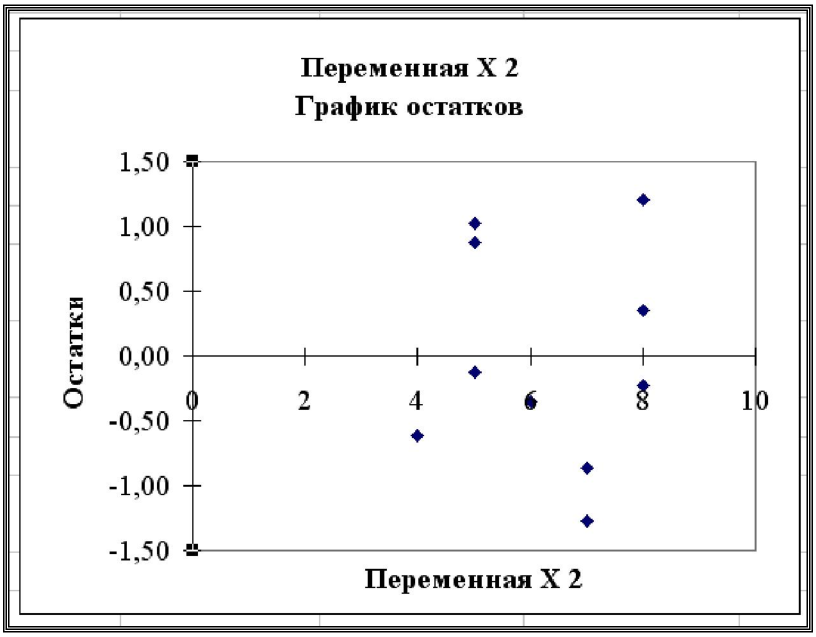


Рис. 3.6. График невязок как функция переменной X_2

3.3 Индивидуальное задание

По 20 предприятиям региона изучается зависимость выработки продукции на одного работника y (тыс. руб.) от ввода в действие новых основных фондов x_1 (% от стоимости фондов на конец года) и от удельного веса рабочих высокой квалификации в общей численности рабочих x_2 (%) (смотри таблицу своего варианта).

Требуется:

1. Построить линейную модель множественной регрессии. Записать стандартизованное уравнение множественной регрессии. На основе стандартизованных коэффициентов регрессии и средних коэффициентов эластичности ранжировать факторы по степени их влияния на результат.

2. Найти коэффициенты парной, частной и множественной корреляции. Проанализировать их.

3. Найти скорректированный коэффициент множественной детерминации. Сравнить его с нескорректированным (общим) коэффициентом детерминации.

4. С помощью F – критерия Фишера оценить статистическую надежность уравнения регрессии и коэффициента детерминации $R^2_{yx_1x_2}$.

5. С помощью частных F – критериев Фишера оценить целесообразность включения в уравнение множественной регрессии фактора x_1 после x_2 и фактора x_2 после x_1.

6. Составить уравнение линейной парной регрессии, оставив лишь один значащий фактор.

Вариант 1

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	6	3,6	9	11	9	6,3	21
2	6	3,6	12	12	11	6,4	22
3	6	3,9	14	13	11	7	24
4	7	4,1	17	14	12	7,5	25
5	7	3,9	18	15	12	7,9	28
6	7	4,5	19	16	13	8,2	30
7	8	5,3	19	17	13	8	30

8	8	5,3	19	18	13	8,6	31
9	9	5,6	20	19	14	9,5	33
10	10	6,8	21	20	14	9	36

Вариант 2

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	6	3,5	10	11	10	6,3	21
2	6	3,6	12	12	11	6,4	22
3	7	3,9	15	13	11	7	23
4	7	4,1	17	14	12	7,5	25
5	7	4,2	18	15	12	7,9	28
6	8	4,5	19	16	13	8,2	30
7	8	5,3	19	17	13	8,4	31
8	9	5,3	20	18	14	8,6	31
9	9	5,6	20	19	14	9,5	35
10	10	6	21	20	15	10	36

Вариант 3

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,7	9	11	11	6,3	22
2	7	3,7	11	12	11	6,4	22
3	7	3,9	11	13	11	7,2	23
4	7	4,1	15	14	12	7,5	25
5	8	4,2	17	15	12	7,9	27
6	8	4,9	19	16	13	8,1	30
7	8	5,3	19	17	13	8,4	31
8	9	5,1	20	18	13	8,6	32
9	10	5,6	20	19	14	9,5	35
10	10	6,1	21	20	15	9,5	36

Вариант 4

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,5	9	11	10	6,3	22
2	7	3,6	10	12	10	6,5	22
3	7	3,9	12	13	11	7,2	24
4	7	4,1	17	14	12	7,5	25
5	8	4,2	18	15	12	7,9	27
6	8	4,5	19	16	13	8,2	30
7	9	5,3	19	17	13	8,4	31
8	9	5,5	20	18	14	8,6	33

9	10	5,6	21	19	14	9,5	35
10	10	6,1	21	20	15	9,6	36

Вариант 5

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,6	9	11	10	6,3	21
2	7	3,6	11	12	11	6,9	23
3	7	3,7	12	13	11	7,2	24
4	8	4,1	16	14	12	7,8	25
5	8	4,3	19	15	13	8,1	27
6	8	4,5	19	16	13	8,2	29
7	9	5,4	20	17	13	8,4	31
8	9	5,5	20	18	14	8,8	33
9	10	5,8	21	19	14	9,5	35
10	10	6,1	21	20	14	9,7	34

Вариант 6

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,5	9	11	10	6,3	21
2	7	3,6	10	12	10	6,8	22
3	7	3,8	14	13	11	7,2	24
4	7	4,2	15	14	12	7,9	25
5	8	4,3	18	15	12	8,1	26
6	8	4,7	19	16	13	8,3	29
7	9	5,4	19	17	13	8,4	31
8	9	5,6	20	18	13	8,8	32
9	10	5,9	20	19	14	9,6	35
10	10	6,1	21	20	14	9,7	36

Вариант 7

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,8	11	11	10	6,8	21
2	7	3,8	12	12	11	7,4	23
3	7	3,9	16	13	11	7,8	24
4	7	4,1	17	14	12	7,5	26
5	7	4,6	18	15	12	7,9	28
6	8	4,5	18	16	12	8,1	30
7	8	5,3	19	17	13	8,4	31
8	9	5,5	20	18	13	8,7	32
9	9	6,1	20	19	13	9,5	33
10	10	6,8	21	20	14	9,7	35

Вариант 8

Номер предприятия	y	x_1	x_2	Номер предприятия	y	x_1	x_2
1	7	3,8	9	11	11	7,1	22
2	7	4,1	14	12	11	7,5	23
3	7	4,3	16	13	12	7,8	25
4	7	4,1	17	14	12	7,6	27
5	8	4,6	17	15	12	7,9	29
6	8	4,7	18	16	13	8,1	30
7	9	5,3	20	17	13	8,5	32
8	9	5,5	20	18	14	8,7	32
9	11	6,9	21	19	14	9,6	33
10	10	6,8	21	20	15	9,8	36

Тема 4. НЕЛИНЕЙНАЯ МНОЖЕСТВЕННАЯ РЕГРЕССИЯ

Эта работа, посвящена построению нелинейной множественной регрессии на примере производственной функции Кобба-Дугласа.

Вычисление коэффициентов нелинейной множественной регрессии для производственной функции Кобба-Дугласа

Цель работы. Используя пространственную выборку таблицы 1 и команду *Поиск решения*, построить нелинейную множественную регрессию для производственной функции Кобба-Дугласа.

Таблица 1

Q	657	1200	2427	4257	8095	9849
L	162	245	452	714	1083	1564
K	279	1167	3069	5585	9119	13989

Производственная функция Кобба-Дугласа имеет вид:

$$Q = A \cdot K^{\beta_1} \cdot L^{\beta_2}, \qquad (1)$$

где Q – объем производства, K – затраты капитала, затраты труда. Показатели β_1, β_2 являются коэффициентами частной эластичности производства Q соответственно по затратам капитала K и труда L. Это означает, что при увеличении одних только затрат капитала (труда) на 1% объем производства увеличивается на β_1 % (β_2 %). При этом имеет место ограничение:

$$\beta_1 + \beta_2 = 1.$$

Решение. Нахождение оценок B, b_1, b_2 для коэффициентов A, β_1, β_2 нелинейной модели (1) будем осуществлять из решения следующей задачи условной минимизации:

$$\min\left[\sum_{i=1}^{n}\left(Q_i - B \cdot K_i^{b_1} \cdot L_i^{b_2}\right)^2\right] \qquad (2)$$

при ограничении

$$b_1 + b_2 = 1. \qquad (3)$$

Для решения этой задачи используем команду *Поиск решения.* Первоначально введем в столбцы A, B, C значения $K_i, L_i, Q_i, i = 1,...,6$ (см. рис. 1). Затем в ячейках B10, B11, B11 зададим начальные («стартовые») значения искомых коэффициентов: $B = 2,\ b_1 = 0.5,\ b_2 = 0.5$.

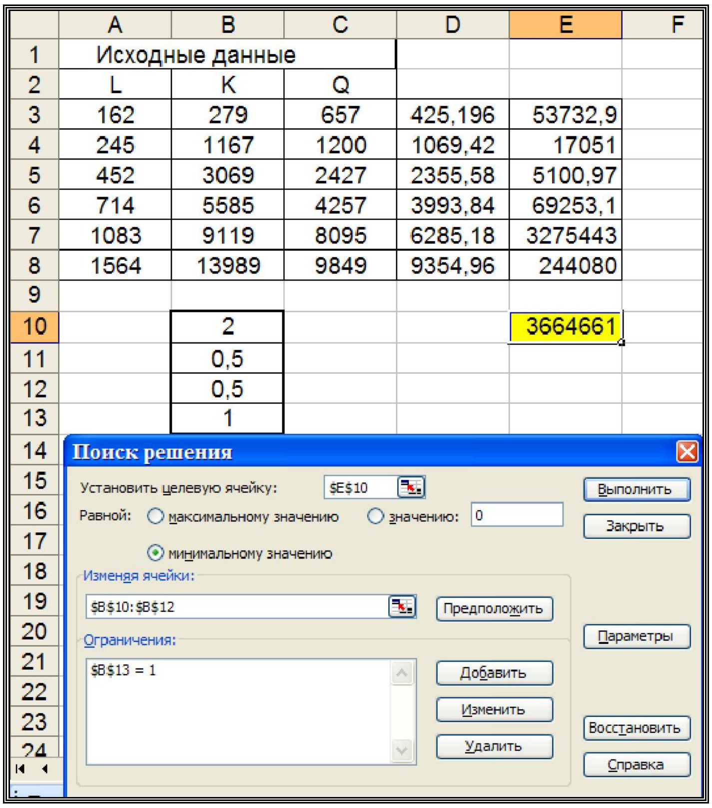
	A	B	C	D	E	F	G
1	Исходные данные						
2	L	K	Q		=$B$10*B3^$B$11*A3^$B$12		
3	162	279	657	425,196	53732,9		
4	245	1167	1200	1069,42	17051		
5	452	3069	2427	2355,58	5100,97		
6	714	5585	4257	3993,84	69253,1		
7	1083	9119	8095	6285,18	3275443		
8	1564	13989	9849	9354,96	244080		
9							
10		2			3664661		
11		0,5					
12		0,5		=СУММ(E3:E8)			
13		1					
14			=B11+B12				

Рис. 1. Подготовительные вычисления
для решения задачи условной минимизации

После этого в соответствующих ячейках столбца D вычислим значения $\hat{Q}_i = B \cdot K_i^{b_1} \cdot L_i^{b_2}$. В столбце E запрограммируем вычисления значений $(Q_i - \hat{Q}_i)^2$, а в ячейке E10 (выделена цветом) вычислим значения функционала

$$F(B, b_1, b_2) = \sum_{i=1}^{6} (Q_i - \hat{Q}_i)^2. \qquad (4)$$

После этих подготовительных вычислений для выполнения команды «Поиск решения» необходимо обратиться к пункту основного меню **Сервис** и в появившемся меню щелкнуть мышью на команде ***Поиск решения***. Затем в появившемся диалоговом окне выполнить следующие действия (см. рис. 2):



	A	B	C	D	E	F
1	Исходные данные					
2	L	K	Q			
3	162	279	657	425,196	53732,9	
4	245	1167	1200	1069,42	17051	
5	452	3069	2427	2355,58	5100,97	
6	714	5585	4257	3993,84	69253,1	
7	1083	9119	8095	6285,18	3275443	
8	1564	13989	9849	9354,96	244080	
9						
10		2			3664661	
11		0,5				
12		0,5				
13		1				

Рис. 2. Задание параметров команды «Поиск решения»

- в поле ввода *Установить целевую ячейку* ввести адрес ячейки, в которой вычисляется значение минимизируемого функционала (в нашем примере – E10);
- включить опцию *Минимальное значение* (ищутся значения коэффициентов, при которых функционал достигает своего минимального значения);
- в поле ввода *Изменяя значения* ввести адреса ячеек, в которых находятся значения искомых коэффициентов (в нашем примере это ячейки B10:B12);
- щелкнув мышью на кнопке *Добавить* формируем ограничения на значения искомых коэффициентов (в нашем примере это условие (3)).

После задания параметров щелкаем на кнопке *Выполнить* и в ячейках B10, B11, B12 выводятся вычисленные значения коэффициентов, а в ячейке E10 – значение функционала (4) при этих значениях коэффициентов. Видно, что вычисленные значения коэффициентов $B = 3.197, b_1 = 0.332$, $b_2 = 0.668$ удовлетворяют ограничению (3)

Таким образом, получено следующее уравнение регрессии:

$$\hat{Q}(K,L) = 3.197 \cdot K^{0.332} \cdot L^{0668}$$

ЗАДАНИЕ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ:

В таблице приведены данные по объемам выпуска ***Q***, затрат капитала ***K*** и труда ***L*** в некоторой отрасли за 10 лет. Используя эти данные, оцените производственную функцию Кобба-Дугласа $Y = AK^{\alpha}L^{\beta}$:

1) сведите данную модель к линейной $y = \beta_0 + \beta_1 k + \beta_2 l$;

2) оцените коэффициенты $\beta_0, \beta_1, \beta_2$, используя метод наименьших квадратов;

3) дайте экономическую интерпретацию коэффициентов $\beta_1, \beta_2, \beta_1 + \beta_2$.

Вариант 1.

Y	**46000**	**5960**	**37500**	**107000**	**130000**	**128000**	**154000**	**226000**	**146500**	**31500**
K	**2**	**5,6**	**2**	**5,6**	**2**	**10,4**	**5,6**	**10,4**	**10,4**	**2**

L	2	2	4	4	6	2	6	4	6	2

Вариант 2.

Y	70500	70500	108000	90500	74000	160000	225000	167500	88500	54000
K	2	5,6	5,6	2	10,4	5,6	10,4	10,4	5,6	2
L	4	2	4	6	2	6	4	6	4	2

Вариант 3.

Y	46000	5960	37500	107000	130000	128000	154000	226000	146500	31500
K	2	5,6	5,6	2	10,4	5,6	10,4	10,4	5,6	2
L	4	2	4	6	2	6	4	6	4	2

Вариант 4.

Y	70500	71000	108000	90500	74000	160000	225000	167500	88500	55500
K	2	5,6	2	5,6	2	10,4	5,6	10,4	10,4	2
L	2	2	4	4	6	2	6	4	6	2

Вариант 5.

Y	46500	59000	37000	47000	135000	122000	154500	226500	146000	30000
K	2	6,5	2	6,5	2	10,4	5,6	10,4	9,4	2
L	2	2	4	2	6	2	6	4	6	2

Вариант 6.

Y	70500	71000	108000	90500	74000	160000	225000	167500	88500	55500
K	2	6,5	2	6,5	2	10,4	5,6	10,4	9,4	2
L	2	2	4	2	6	2	6	4	6	2

Вариант 7.

Y	42500	50960	37500	97000	130000	125000	145000	216000	164500	31500
K	5,6	5,6	2	5,6	2	10,4	2	10,4	2	2
L	2	2	4	4	6	2	6	4	6	2

Вариант 8.

Y	70500	71000	108000	90500	74000	160000	225000	167500	88500	55500
K	2	5,6	2	5,6	2	10,4	5,6	10,4	10,4	2
L	2	2	4	4	6	2	6	4	6	2

Вариант 9.

Y	46000	5960	37500	107000	130000	128000	154000	226000	146500	31500
K	5,6	5,6	2	5,6	2	10,4	2	10,4	2	2
L	2	2	4	4	6	2	6	4	6	2

Тема 5. ГЕТЕРОСКЕДАСТИЧНОСТЬ

Цель: научиться оценивать наличие эффекта гетероскедастичности.

Основные формулы и понятия:

<u>Тест Парка</u>

$$\ln e_i^2 = a + b \cdot \ln x_{ij} + v_i,$$

$$где \quad x_{ij} \; - \; i-е\ значение\ о-го\ \ фактора$$

$$v_i - случайный\ \ остаток$$

Условие принятия гипотезы: $t_b > t_{\alpha,n-2}$

Если данное условие выполняется, то нулевая гипотеза о наличии гетероскедастичности будет принята при уровне значимости α .

<u>Тест ранговой корреляции Спирмена</u>

$$r_{x,e}=1-\frac{6\cdot\sum_{i=1}^{n}d_i^2}{n\cdot\left(n^2-1\right)}$$ — коэффициент ранговой корреляции Спирмена,

где x — одна из объясняющих переменных,

d_i — разность между рангом *i-го* наблюдения *x* и рангом модуля остатка *в i-м* наблюдении.

$$t_r=\frac{r_{x,e}\cdot\sqrt{n-1}}{\sqrt{1-r_{x,e}^2}}$$ — статистика.

Если в модели регрессии имеется более одной объясняющей переменной, то проверка гипотезы может выполняться с использованием каждой из них.

Условие принятия гипотез: $t_r > t_{\alpha,n-2}$.

Если данное условие выполняется, то нулевая гипотеза об отсутствии гетероскедастичности отклоняется при уровне значимости α .

<u>Тест Голдфельда — Кванта</u>

В этом случае все наблюдения необходимо упорядочить по мере возрастания значений x. Затем построить регрессионную модель для первых ***k*** и последних ***k*** наблюдений. Соответственно обозначим через $SS_{ост}^{(1)}$ и $SS_{ост}^{(3)}$ необъясненную сумму квадратов отклонений в каждой регрессии. Тогда статистика имеет вид

$$F=\frac{SS_{ост}^{(3)}}{SS_{ост}^{(1)}}.$$

Если выполняется условие $F > F_\gamma(k-m-1,k-m-1)$*, то гипотеза об отсутствии гетероскедастичности отвергается.*

Для проведения теста ранговой корреляции Спирмена необходимо выполнить следующие действия:

1. Отсортировать данные в таблице по возрастанию значений *x*;
2. Придать каждому наблюдению ранг, для чего необходимо добавить новый столбец, в котором задать числа от 1 до *n*;
3. Вызвать из пакета анализа надстройку ***Регрессия***, указав в диалоговом окне опцию ***Остатки***. После выполнения данной надстройки появится дополнительная таблица, в которой содержатся номера наблюдений, прогнозы и остатки. Тот столбец таблицы, в котором находятся остатки, необходимо перенести к исходным данным. После выполнения этих действий наша таблица будет содержать четыре столбца: ранг наблюдения, упорядоченные значения регрессора *x*, значения *y* и значения остатков;
4. Отсортировать данные по возрастанию модулей остатков и добавить новый столбец рангов остатков, аналогичным образом задав значения от 1 до *n*;
5. В дополнительном столбце вычислить значения разности между двумя полученными рангами (это и будет значение d_i);
6. На основании формул подсчитать коэффициент ранговой корреляции и статистику;
7. Проверить гипотезу.

Вид таблицы для проведения теста ранговой корреляции Спирмена

Ранг по х	Ценах1(р.)	Спрос у (тыс. шт.)	Остатки	Ранг по остаткам	Разность рангов D_i	D_i* D_i
8	15,91р.	117,088	-0,34387	1	7	49
5	15,54р.	119,864	-0,39014	2	3	9
15	16,76р.	110,023	-0,84306	3	12	144
2	15,21р.	123,809	1,019821	4	-2	4
3	15,28р.	121,175	-1,11646	5	-2	4
9	15,92р.	116,17	-1,12322	6	3	9
10	15,95р.	118,344	1,257187	7	3	9
14	16,69р.	110,106	-1,31194	8	6	36
1	15,09р.	125,178	1,426776	9	-8	64
6	15,62р.	118,068	-1,5813	10	-4	16
11	16,31р.	116,201	1,847847	11	0	0
12	16,33р.	111,457	-2,67328	12	0	0
13	16,60р.	115,103	3,003645	13	0	0
4	15,49р.	116,914	-3,7319	14	-10	100
7	15,70р.	123,589	4,559903	15	-8	64
					Сумма	508

Следовательно, значение ранговой корреляции Спирмена будет равно

$$r_{x,e} = 1 - \frac{6 \cdot \sum_{i=1}^{n} D_i^2}{n \cdot \left(n^2 - 1\right)} = 1 - \frac{6 \cdot 508}{15 \cdot (225 - 1)} = 0.0928$$

А значение статистики будет $t = 0.0928 \cdot \sqrt{15-1} = 0.028$

Выбрав уровень значимости 5 %, получаем критическую точку $t_{0.05,13} = 2.16$. Данное значение получено формулой СТЬЮДРАСПОБР(0,05;13).

Поскольку условие $t < t_{\alpha,n-2}$ не выполняется, то гипотеза о наличии гетероскедастичности будет принята.

Для проверки подобной гипотезы на основании теста Гольдфельда — Кванта необходимо подобным образом отсортировать наблюдения по возрастанию значения *х*, а затем отдельно оценить каждую регрессионную модель для первой трети и для последней трети наблюдений. Просчитать соответствующую статистику и проверить гипотезу об отсутствии гетероскедастичности.

Задание для самостоятельной работы

Провести исследование табличных данных практической работы «множественная регрессия» на наличие гетероскедастичности, между значением ***y*** и каждым регрессором отдельно

a) Тестом Парка
b) Тестом ранговой корреляции Спирмена;
c) Тестом Гольдфельда — Кванта.

Сделать выводы.

Тема 6. ОБОБЩЕННЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

1 задание. По данным таблицы построить линейную регрессионную модель, характеризующую зависимость показателя y от факторов x_1 и x_2. Построение модели начать с тестирования на гетероскедатичность. При построении модели применить многоэтапную процедуру оценивания ее коэффициентов с помощью обобщенного МНК.

№	x_1	x_2	y		№	x_1	x_2	y
1.	13	43	79		9.	58	161	207
2.	28	56	110		10.	23	108	152
3.	33	24	97		11.	69	86	199
4.	42	98	171		12.	8	143	144
5.	12	176	204		13.	60	42	140
6.	44	124	174		14.	11	199	183
7.	36	130	184		15.	26	145	178
8.	33	291	311		16.	61	115	185
9.	34	141	206		17.	18	111	152
10.	21	95	128		18.	30	192	204

2 задание. По данным таблицы построить линейную регрессионную модель, характеризующую зависимость показателя y от факторов x_1 и x_2 и x_3. Построение модели начать с тестирования на гетероскедатичность. При построении модели применить многоэтапную процедуру оценивания ее коэффициентов с помощью обобщенного МНК.

№	x_1	x_2	x_3	y		№	x_1	x_2	x_3	y
1.	123	53	538	1882		9.	153	25	782	2565
2.	122	83	734	2006		10.	164	23	627	1757
3.	143	48	605	2083		11.	193	93	945	3055
4.	159	29	864	2388		12.	151	119	590	1636
5.	133	42	703	2334		13.	148	33	770	2529
6.	183	69	457	1310		14.	103	88	574	1563
7.	139	141	565	1983		15.	140	114	344	1389
8.	162	51	390	1117		16.	129	31	449	1254

Тема 7. ФИКТИВНЫЕ ПЕРЕМЕННЫЕ И КАТЕГОРИИ

Цель: научиться использовать в модели фиктивные переменные сдвига и наклона, а также различные категории.

Основные формулы и понятия:

Фиктивная переменная необходима для описания качественного изменения и может принимать два значения 0 и 1.

$y = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot D + u$ — *модель с фиктивной переменной сдвига*;

$y = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot D \cdot x + u$ — *модель с фиктивной переменной наклона*;

$y = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot D \cdot x + \beta_3 \cdot D + u$ — *модель с фиктивной переменной наклона и сдвига*.

Категория — событие, про которое для каждого наблюдения можно определенно сказать, произошло оно в этом наблюдении или нет.

Набор категорий — конечный набор взаимоисключающих событий, полностью исчерпывающий все возможности.

Для описания категорий необходимо ввести совокупность фиктивных переменных.

Электронная таблица Excel

До сих пор нами рассматривался только случай количественных регрессоров, поскольку значение цен и спроса являются числами. Однако может возникнуть ситуация, когда необходимо учесть некоторую специфическую информацию. Рассматривая модель спроса, можно предположить, что продаются два одинаковых продукта по одной цене, но имеющие некоторые различия. Например, наряду с уже давно продающимся чистящим порошком, поступает в продажу такой же порошок, но с новым ароматом. И имеется задача исследовать, насколько большим или меньшим спросом пользуется новая продукция. Конечно, можно построить две различные модели, и посмотреть разницу между ними, однако нас будет интересовать общая модель. В этом случае в модель необходимо вносить качественный регрессор, для чего нужно использовать фиктивную переменную. Данная переменная может принимать только два значение 0 или 1, в зависимости от отсутствия или наличия нового качества. В этом случае можно строить модель с фиктивной переменной наклона и сдвига. Работа с фиктивными переменными ни чем не отличается от построения регрессионной модели.

Поэтому рассмотрим задачу. Значение цены x и спроса y на два различных товара, которые мы условно назовем «обычный» и «новый», представлены в таблице 17.

Таблица 1

Номер наблюдения	Вид	Цена x^1(р.)	Спрос y (тыс. шт.)
1	новый	15,09р.	125,1779
2	новый	15,21р.	123,8094
3	старый	15,28р.	121,175
4	старый	15,49р.	116,9143
5	старый	15,54р.	119,8643
6	старый	15,62р.	118,0681
7	новый	15,70р.	123,5887
8	новый	15,91р.	117,0877
9	старый	15,92р.	116,1699
10	новый	15,95р.	118,3436
11	новый	16,31р.	116,2008
12	старый	16,33р.	111,4565
13	новый	16,60р.	115,1026
14	старый	16,69р.	110,1056
15	старый	16,76р.	110,0231

В электронной таблице Excel имеются возможности для быстрого задания значений фиктивной переменой. Для этого необходимо вставить столбец между колонками с названиями *Вид* и *Цена*. Озаглавим этот столбец как *Фиктивная переменная*, и для определения значений будем использовать логическую функцию ЕСЛИ. Данная функция имеет три аргумента. Первый — это логическое выражение, которое может принимать истинное или ложное значение. Вторым аргументом идет то значение, которое появляется в ячейке при истинности условия, а соответственно в третьем аргументе — значение, которое появляется в противном случае.

Выполнив данные действия, получим первые две строки таблицы 18.

Таблица 2

Номер наблюдения	Вид	Фиктивная переменная	Цена x^1 (р.)	Спрос y (тыс. шт.)
1	новый	=ЕСЛИ(B2="новый";1;0)	15,09р.	125,1779

В столбце фиктивной переменной появится значение 1, если в предыдущем столбце находилось слово «новый», и 0 в противоположном случае. После этого необходимо значение функции, находящейся в столбце *C,* скопировать во все нижние ячейки, а поскольку адресация относительная, то адрес будет меняться. Необходимо отметить, что логическая функция может иметь и другой вид:

ЕСЛИ(B2 = "обычный";0;1).

Теперь наша задача заключается в определении степени влияния фиктивной переменной. А именно, влияет ли это значение на свободный член (в этом случае при изменении качества можно говорить о том, что спрос изменится на какое-то количество) или на наклон линии регрессии (спрос изменится во сколько-то), или на оба эти значения сразу.

Вначале оценим регрессию, при условии, что фиктивная переменная влияет только на значение свободного члена. В этом случае итоговая таблица после выполнения надстройки **Регрессии**, при условии, что *Входной интервал Y* задан в виде ***E1:E16***, а *Входной интервал X* в виде ***C1:D16***, имеет вид, изображенный в таблице 19.

Таблица 3

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественный R	0,963696
R-квадрат	0,928711
Нормированный R-квадрат	0,916830
Стандартная Ошибка	1,363084
Наблюдения	15

Продолжение табл. 4

Дисперсионный анализ

	df	*SS*	*MS*	*F*	*Значимость F*
Регрессия	2	290,4628387	145,231419	78,16547142	1,31E–07
Остаток	12	22,29599593	1,85799966		
Итого	14	312,7588347			

	Коэффи-циенты	*Стандартная ошибка*	*t-статистика*	*P-значение*	*Нижние 95 %*	*Верхние 95 %*
Y-пересечение	232,0028	10,78827	21,5051052	5,9691E-11	208,49	255,508
Фиктивная переменная	3,474500	0,7109700	4,8869856	0,00037407	1,9254	5,02357

Цена x(р.)	-7,30442	0,675558	–10,8124125	1,5303E–07	–8,77634	–5,83251

Регрессионная модель имеет вид: $y = 232 + 3{,}47D - 7{,}304x$

Поскольку значение фиктивной переменной D равно 1 для «нового» вида и 0 для «обычного», то данную модель можно отдельно расписать для каждого случая.

$y = 232 - 7{,}304x$ — обычный вид,

$y = 235{,}47 - 7{,}304x$ — новый вид.

Следовательно, спрос на новый вид продукции приблизительно на 3,47 тыс. ед. больше. Коэффициент детерминации равен 0,928, что намного больше, чем данное значение для парного случая.

Рассмотрим теперь возможность построения модели с фиктивной переменной наклона, для чего в качестве регрессоров значения необходимо использовать переменные x и Dx. Следовательно, необходимо добавить дополнительный столбец между фиктивной переменной и значениями x, в который надо записать их произведения.

Опустим таблицу, которая генерируется надстройкой **Регрессия**. Однако, самостоятельно выполнив данные операции, можно получить следующую модель: $y = 233{,}52 + 0{,}21Dx - 7{,}403x$.

Аналогичным образом интерпретируя значение фиктивной переменной, можно расписать два случая:

$y = 233{,}52 - 7{,}4x$ — для обычного вида продукции;

$y = 233{,}52 - 7{,}19x$ — для нового вида продукции.

Выводы из полученных моделей совершенно очевидны, поскольку видна разница во влиянии цены на спрос для каждого вида продукции. Коэффициент детерминации в этом случае равен 0,929, что не намного больше соответствующего значения для фиктивной переменной сдвига, а следовательно, они обе пригодны для прогнозирования. Однако результаты использования моделей будут во многом различными. В первом случае спрос на «новый» вид продукции на 3,47 тыс. ед. больше, чем на «старый», во втором случае цена сильнее влияет на «старый» вид продукции.

При необходимости можно построить модель, в которой фиктивная переменная влияет как на наклон, так и на сдвиг.

До сих пор нами рассматривался случай, когда имеются всего два значения качества, то есть два вида продукции. Однако нередки случаи, когда необходимо проанализировать спрос для различных продуктов. Тогда необходимо вводить *набор категорий* — как конечный набор взаимоисключающих событий, полностью описывающий все возможности. Предположим, что исследуется влияние цены на спрос при наличии «старой», «обычной», «новой» и «самой новой» продукции.

В этом случае для описания этих категорий необходимо вводить набор фиктивных переменных по следующему правилу.

1. Число фиктивных переменных должно быть на единицу меньше, чем число категорий. В данном случае имеется четыре категории, а следовательно, необходимо ввести три фиктивные переменные, которые мы обозначим D1, D2, D3.
2. Выбрать произвольную категорию в качестве эталонной. Именно с этой категорий в последствии будут сравниваться все остальные. Для эталонной категории необходимо, чтобы значения всех фиктивных переменных равнялись нулю.
3. Для всех остальных категорий необходимо, чтобы одна из фиктивных переменных равнялась 1, в то время как значение всех остальных равно 0.

Достаточно легко можно расставить значения фиктивных переменных, используя ту же условную функцию ЕСЛИ. При наличии четырёх различных видов продукции необходимо вставить три дополнительных столбца, в которых будут находиться фиктивные переменных. Задать логические функции можно так, как показано в таблице 20.

Таблица 5

Номер наблюдения	Вид	Фиктивная переменная D1	Фиктивная переменная D2	Фиктивная переменная D3	Цена x^1 (р.)	Спрос y (тыс.шт.)
1		=ЕСЛИ(B2= «обычный»;1;0)	=ЕСЛИ(B2= «новой»;1;0)	=ЕСЛИ(B2= «самой новый»;1;0)	15,09р.	125,1779

После копирования данных функций вниз для значения старой все фиктивные переменные будут равны нулю, для обычной — только значение первой фиктивной переменной будет равно 1 и т. д.

После этого можно вызвать надстройку **Регрессия**, у которой в качестве входного интервала X, необходимо указать значения всех фиктивных переменных D и нефиктивной переменной *X*, то есть задать *Входной интервал X* в виде ***C1:F16***.

Полученные результаты поддаются достаточно простой интерпретации. Значение, находящееся напротив фиктивной переменной *D*1, показывает, насколько изменился спрос при переходе от эталонной к первой категории, то есть насколько различен спрос между «обычной» и «новой» продукцией. Аналогично интерпретируются значения, стоящие напротив других фиктивных переменных.

Задания для самостоятельной работы

2. Для данных своего варианта подобрать наилучшее воздействие фиктивной переменной (влияние на наклон или сдвиг). При этом категории «старый» и «обычный» воспринимать как одно значение, а категории «новый» и «самый новый» — как другое.
3. Определить, насколько изменяется спрос при переходе от одной категории к другой.

Тема 8. СИСТЕМЫ ЭКНОМЕТРИЧЕСКИХ УРАВНЕНИЙ

Пример решения типовой задачи

Рассмотрим **пример.** Изучается модель вида

$$\begin{cases} C_t = a_1 + b_{11}\cdot Y_t + b_{12}\cdot C_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_{21}\cdot r_t + b_{22}\cdot I_{t-1} + \varepsilon_2, \\ r_t = a_3 + b_{31}\cdot Y_t + b_{32}\cdot M_t + \varepsilon_3, \\ Y_t = C_t + I_t + G_t, \end{cases}$$

где C_t – расходы на потребление в период t, Y_t – совокупный доход в период t, I_t – инвестиции в период t, r_t – процентная ставка в период t, M_t – денежная масса в период t, G_t – государственные расходы в период t, C_{t-1} – расходы на потребление в период $t-1$, I_{t-1} инвестиции в период $t-1$.

Первое уравнение – функция потребления, второе уравнение – функция инвестиций, третье уравнение – функция денежного рынка, четвертое уравнение – тождество дохода.

Модель представляет собой систему одновременных уравнений. Проверим каждое ее уравнение на идентификацию.

Модель включает четыре эндогенные переменные (C_t, I_t, Y_t, r_t) и четыре предопределенные переменные (две экзогенные переменные – M_t и G_t и две лаговые переменные – C_{t-1} и I_{t-1}).

1. Проверим необходимое условие идентификации для каждого из уравнений модели.

Первое уравнение: $C_t = a_1 + b_{11} \cdot Y_t + b_{12} \cdot C_{t-1} + \varepsilon_1$. Это уравнение содержит две эндогенные переменные C_t и Y_t и одну предопределенную переменную C_{t-1}. Таким образом, $H = 2$, а $D = 4 - 1 = 3$, т.е. выполняется условие $D + 1 > H$. Уравнение сверхидентифицируемо.

Второе уравнение: $I_t = a_2 + b_{21} \cdot r_t + b_{22} \cdot I_{t-1} + \varepsilon_2$. Оно включает две эндогенные переменные I_t и r_t и одну экзогенную переменную I_{t-1}. Выполняется условие $D + 1 = 3 + 1 > H = 2$. Уравнение сверхидентифицируемо.

Третье уравнение: $r_t = a_3 + b_{31} \cdot Y_t + b_{32} \cdot M_t + \varepsilon_3$. Оно включает две эндогенные переменные Y_t и r_t и одну экзогенную переменную M_t. Выполняется условие $D + 1 = 3 + 1 > H = 2$. Уравнение сверхидентифицируемо.

Четвертое уравнение: $Y_t = C_t + I_t + G_t$. Оно представляет собой тождество, параметры которого известны. Необходимости в идентификации нет.

2. Проверим для каждого уравнения достаточное условие идентификации. Для этого составим матрицу коэффициентов при переменных модели.

	C_t	I_t	r_t	Y_t	C_{t-1}	I_{t-1}	M_t	G_t
I уравнение	–1	0	0	b_{11}	b_{12}	0	0	0
II уравнение	0	–1	b_{21}	0	0	b_{22}	0	0
III уравнение	0	0	–1	b_{31}	0	0	b_{32}	0
Тождество	1	1	0	–1	0	0	0	1

В соответствии с достаточным условием идентификации ранг матрицы коэффициентов при переменных, не входящих в исследуемое уравнение, должен быть равен числу эндогенных переменных модели без одного.

Первое уравнение. Матрица коэффициентов при переменных, не входящих в уравнение, имеет вид

	I_t	r_t	I_{t-1}	M_t	G_t
II уравнение	–1	b_{21}	b_{22}	0	0
III уравнение	0	–1	0	b_{32}	0
Тождество	1	0	0	0	1

Ранг данной матрицы равен трем, так как определитель квадратной подматрицы 3×3 не равен нулю:

$$\begin{vmatrix} b_{22} & 0 & 0 \\ 0 & b_{32} & 0 \\ 0 & 0 & 1 \end{vmatrix} = b_{22}b_{32} \neq 0.$$

Достаточное условие идентификации для данного уравнения выполняется.

Второе уравнение. Матрица коэффициентов при переменных, не входящих в уравнение, имеет вид

	C_t	Y_t	C_{t-1}	M_t	G_t
I уравнение	–1	b_{11}	b_{12}	0	0
III уравнение	0	b_{31}	0	b_{32}	0
Тождество	1	–1	0	0	1

Ранг данной матрицы равен трем, так как определитель квадратной подматрицы 3×3 не равен нулю:

$$\begin{vmatrix} b_{12} & 0 & 0 \\ 0 & b_{32} & 0 \\ 0 & 0 & 1 \end{vmatrix} = b_{12}b_{32} \neq 0.$$

Достаточное условие идентификации для данного уравнения выполняется.

Третье уравнение. Матрица коэффициентов при переменных, не входящих в уравнение, имеет вид

	C_t	I_t	C_{t-1}	I_{t-1}	G_t
I уравнение	–1	0	b_{12}	0	0
II уравнение	0	–1	0	b_{22}	0
Тождество	1	1	0	0	1

Ранг данной матрицы равен трем, так как определитель квадратной подматрицы 3×3 не равен нулю:

$$\begin{vmatrix} b_{12} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & 1 \end{vmatrix} = b_{12}b_{22} \neq 0.$$

Достаточное условие идентификации для данного уравнения выполняется.

Таким образом, все уравнения модели сверхидентифицируемы. Приведенная форма модели в общем виде будет выглядеть следующим образом:

$$\begin{cases} C_t = A_1 + \delta_{11}C_{t-1} + \delta_{12}I_{t-1} + \delta_{13}M_t + \delta_{14}G_t + u_1, \\ I_t = A_2 + \delta_{21}C_{t-1} + \delta_{22}I_{t-1} + \delta_{23}M_t + \delta_{24}G_t + u_2, \\ r_t = A_3 + \delta_{31}C_{t-1} + \delta_{32}I_{t-1} + \delta_{33}M_t + \delta_{34}G_t + u_3, \\ Y_t = A_4 + \delta_{41}C_{t-1} + \delta_{42}I_{t-1} + \delta_{43}M_t + \delta_{44}G_t + u_1. \end{cases}$$

Варианты индивидуальных заданий

Даны системы эконометрических уравнений.

Требуется

1. Применив необходимое и достаточное условие идентификации, определите, идентифицируемо ли каждое из уравнений модели.

2. Определите метод оценки параметров модели.

3. Запишите в общем виде приведенную форму модели.

Вариант 1

Модель протекционизма Сальватора (упрощенная версия):

$$\begin{cases} M_t = a_1 + b_{12}N_t + b_{13}S_t + b_{14}E_{t-1} + b_{15}M_{t-1} + \varepsilon_1, \\ N_t = a_2 + b_{21}M_t + b_{23}S_t + b_{26}Y_t + \varepsilon_2, \\ S_t = a_3 + b_{31}M_t + b_{32}N_t + b_{36}X_t + \varepsilon_3. \end{cases}$$

где M – доля импорта в ВВП; N – общее число прошений об освобождении от таможенных пошлин; S – число удовлетворенных прошений об освобождении от таможенных пошлин; E – фиктивная переменная, равная 1 для тех лет, в которые курс доллара на международных валютных рынках был искусственно завышен, и 0 – для всех остальных лет; Y – реальный ВВП; X – реальный объем чистого экспорта; t – текущий период; $t-1$ – предыдущий период.

Вариант 2

Макроэкономическая модель (упрощенная версия модели Клейна):

$$\begin{cases} C_t = a_1 + b_{12}Y_t + b_{13}T_t + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + b_{24}K_{t-1} + \varepsilon_2, \\ Y_t = C_t + I_t, \end{cases}$$

где C – потребление; I – инвестиции; Y – доход; T – налоги; K – запас капитала; t – текущий период; $t-1$ – предыдущий период.

Вариант 3

Макроэкономическая модель экономики США (одна из версий):

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}C_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + b_{23}r_t + \varepsilon_2, \\ r_t = a_3 + b_{31}Y_t + b_{34}M_t + b_{35}r_{t-1} + \varepsilon_3, \\ Y_t = C_t + I_t + G_t, \end{cases}$$

где C – потребление; Y – ВВП; I – инвестиции; r – процентная ставка; M – денежная масса; G – государственные расходы; t – текущий период; $t-1$ – предыдущий период.

Вариант 4

Модель Кейнса (одна из версий):

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}Y_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + \varepsilon_2, \\ Y_t = C_t + I_t + G_t, \end{cases}$$

где C – потребление; Y – ВВП; I – валовые инвестиции; G – государственные расходы; t – текущий период; $t-1$ – предыдущий период.

Вариант 5

Модель денежного и товарного рынков:

$$\begin{cases} R_t = a_1 + b_{12}Y_t + b_{14}M_t + \varepsilon_1, \\ Y_t = a_2 + b_{21}R_t + b_{23}I_t + b_{25}G_t + \varepsilon_2, \\ I_t = a_3 + b_{31}R_t + \varepsilon_3, \end{cases}$$

где R – процентные ставки; Y – реальный ВВП; M – денежная масса; I – внутренние инвестиции; G – реальные государственные расходы.

Вариант 6

Модифицированная модель Кейнса:

$$\begin{cases} C_t = a_1 + b_{11}Y_t + \varepsilon_1, \\ I_t = a_2 + b_{21}Y_t + b_{22}Y_{t-1} + \varepsilon_2, \\ Y_t = C_t + I_t + G_t, \end{cases}$$

где C – потребление; Y – доход; I – инвестиции; G – государственные расходы; t – текущий период; $t-1$ – предыдущий период.

Вариант 7

Макроэкономическая модель:

$$\begin{cases} C_t = a_1 + b_{11}D_t + \varepsilon_1, \\ I_t = a_2 + b_{22}Y_t + b_{23}Y_{t-1} + \varepsilon_2, \\ Y_t = D_t + T_t, \\ D_t = C_t + I_t + G_t, \end{cases}$$

где C – расходы на потребление; Y – чистый национальный продукт; D – чистый национальный доход; I – инвестиции; T – косвенные налоги; G – государственные расходы; t – текущий период; $t-1$ – предыдущий период.

Вариант 8

Гипотетическая модель экономики:

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}J_t + \varepsilon_1, \\ J_t = a_2 + b_{21}Y_{t-1} + \varepsilon_2, \\ T_t = a_3 + b_{31}Y_t + \varepsilon_3, \\ Y_t = C_t + J_t + G_t, \end{cases}$$

где C – совокупное потребление в период t; Y – совокупный доход в период t; J – инвестиции в период t; T – налоги в период t; G – государственные доходы в период t.

Вариант 9

Модель денежного рынка:

$$\begin{cases} R_t = a_1 + b_{11}M_t + b_{12}Y_t + +\varepsilon_1, \\ Y_t = a_2 + b_{21}R_t + b_{22}I_t + \varepsilon_2, \\ I_t = a_3 + b_{33}R_t + \varepsilon_3, \end{cases}$$

где R – процентные ставки; Y – ВВП; M – денежная масса; I – внутренние инвестиции.

Вариант 10

Конъюнктурная модель имеет вид:

$$\begin{cases} C_t = a_1 + b_{11}Y_t + b_{12}C_{t-1} + \varepsilon_1, \\ I_t = a_2 + b_{21}r_t + b_{22}I_{t-1} + \varepsilon_2, \\ r_t = a_3 + b_{31}Y_t + b_{32}M_t + \varepsilon_3, \\ Y_t = C_t + I_t + G_t, \end{cases}$$

где C – расходы на потребление; Y – ВВП; I – инвестиции; r – процентная ставка; M – денежная масса; G – государственные расходы; t – текущий период; $t-1$ – предыдущий период.

Методические указания по выполнению самостоятельной работы

Введение

Самостоятельная работа предусмотрена учебным планом. Цель самостоятельной работы студента в рамках курса «Эконометрика» — закрепление и расширение знаний, полученных во время проведения аудиторных занятий.

Содержание самостоятельной работы

1. Проработка лекционного материала осуществляется студентом с использованием конспекта лекций и рекомендуемых учебников. Цель — подготовка к восприятию очередной темы, рассматриваемой на лекции.

2. Подготовка к практическим занятиям. В соответствии с темой практической работы студент должен изучить теоретический материал, подготовить решение задания к реализации на компьютере.
   Темы практических (соответственно, самостоятельных) работ:
   1. Линейная парная регрессия
   2. Нелинейная парная регрессия
   3. Множественная линейная регрессия
   4. Режим РЕГРЕССИЯ в программе Excel
   5. Нелинейная множественная регрессияаса
   6. Обобщенный МНК
   7. Системы эконометрических уравнений
   8. Сглаживание и экстраполяция временных рядов
   9. Авторегрессионные процессы
   10. Простейшие адаптивные модели временных рядов

3. В рамках раздела «Изучение дополнительных тем курса» студент самостоятельно изучает предложенные вопросы, связанные с построением и анализом непараметрических эконометрических моделей. Для достижения этой цели сформулированы следующие задания:

   - Непараметрические оценки функции плотности и функции регрессии, их практическое применение.

- Статистические методы прогнозирования. Задача о моменте встречи (пересечении регрессионных прямых) и непараметрические методы ее решения.
- Экспертные методы прогнозирования. Построение и анализ сценариев.
- Методы проверки однородности для связанных выборок

Список литературы

1. Тихомиров, Николай Петрович. Эконометрика : учебник для вузов / Н. П. Тихомиров, Е. Ю. Дорохина . — М. : ЭКЗАМЕН, 2007 – 510[2] с. : ил., табл. (в библиотеке 11 экз.) (**Гриф**)
2. Яновский, Леонид Петрович. Введение в эконометрику : учебное пособие для вузов / Л. П. Яновский, А. Г. Буховец ; ред. Л. П. Яновский. - 2-е изд., доп. — М. : КноРус, 2009. - 254[2] с. : ил., табл. (в библиотеке 10 экз.)
3. Эконометрика : учебник для вузов / И. И. Елисеева [и др.] ; ред. И. И. Елисеева. - 2-е изд., перераб. и доп. - М. : Финансы и статистика, 2008. - 574[2] с. : ил., табл. (в библиотеке 5 экз.) (**Гриф**)

3.2. Дополнительная литература

1. Орлов, Александр Иванович. Эконометрика: Учебник для вузов/ А. И. Орлов. — 3-е изд., перераб и доп.. — М.: Экзамен, 2004. - 573[3] с.. (в библиотеке 1 экз.)
2. Практикум по эконометрике: Учебное пособие для вузов / Ирина Ильинична Елисеева, Светлана Владимировна Курышева, Нелли Михайловна Гордеенко и др; Ред. И. И. Елисеева. - М.: Финансы и статистика, 2001. - 192 с. (в библиотеке 2 экз.)
3. Бородич, Сергей Аркадьевич. Эконометрика: Учебное пособие для вузов. — Минск: Новое знание, 2001. - 408[8] с. : ил. (в библиотеке 4 экз.) (**Гриф**)
4. Кремер, Наум Шевелевич. Эконометрика: Учебник для вузов. — М.: ЮНИТИ-ДАНА, 2003. - 311 с. : ил. (в библиотеке 2 экз.) (**Гриф**)