УДК 519.86:65(075.8)

Ю.И. Паршин
кандидат технических наук, доцент
ГВУЗ "Украинский государственный химико-технологический университет", г. Днепропетровск

РАЗРАБОТКА МНОЖЕСТВЕННОЙ РЕГРЕССИОННОЙ МОДЕЛИ ФУНКЦИОНИРОВАНИЯ ЭКОНОМИКИ ДНЕПРОПЕТРОВСКОЙ ОБЛАСТИ

Предложен алгоритм выявления экономических взаимосвязей и разработана множественная регрессионная модель функционирования экономики Днепропетровской области.

Ключевые слова: *множественная регрессионная модель функционирования экономики, алгоритм выявления экономических взаимосвязей.*

I. Вступление

Для формирования эффективных стратегий, направленных на обеспечение устойчивого экономического развития, необходима аналитическая система поддержки стратегических решений. Одним из составляющих компонентов такой системы является комплекс адекватных математических моделей, описывающих сложный процесс функционирования экономики области.

Проблеме исследования развития региональных экономических систем уделяют внимание многие современные ученые, в частности А.И. Амоша, В.М. Геец, В.П. Вишневский, Р.Н. Лепа, Н.И. Верхоглядова, Т.С. Клебанова, К.Ф. Ковальчук и многие другие.

По результатам проведенных исследований [1] были получены модели двухфакторных производственных функций, на основе которых выявлены наиболее приоритетные виды экономической деятельности Днепропетровской области. Однако процесс функционирования экономики области зависит от множества взаимосвязанных и взаимодействующих факторов, что требует проведения детальных исследований с учетом множества факторных признаков и получения на этой основе многофакторных математических моделей.

II. Постановка задачи

Цель статьи заключается в разработке множественной регрессионной модели для математического описания сложного процесса функционирования экономики Днепропетровской области.

III. Результаты

Для математического описания функционирования экономики Днепропетровской области в качестве результативного признака принят основной экономический показатель – выручка от реализации продукции. Проведенные исследования показывают, что процесс функционирования экономики области по принятому результативному показателю обусловлен влиянием множества факторных признаков.

На основе разработанной методики отбора наиболее влиятельных факторов с использованием пошаговой процедуры сформировано множество из пяти факторных признаков $D = \{ x_1, x_2, x_3, x_4, x_5 \}$: материальные затраты x_1, затраты на оплату труда x_2, затраты на топливо и энергию x_3, амортизация x_4 и другие операционные затраты по реализации продукции x_5.

Сформированное множество составляет основу для построения множественной регрессионной модели функционирования экономики области и установления взаимосвязи между выручкой от реализации и пятью наиболее значимыми факторами.

Экономическая взаимосвязь между признаками для генеральной совокупности может быть описана линейной или нелинейной функциональной зависимостью. Для выявления этой взаимосвязи по данным статистической отчетности за период 2006–2011 гг. [2; 3] произведена выборка объема n:

$$(y_i, x_{1i}, x_{2i}, x_{3i}, x_{4i}, x_{5i}),\ i = 1,2,\ldots,n. \quad (1)$$

Для получения адекватной математической модели и выявления экономических взаимосвязей функционирования экономики области разработан алгоритм, предусматривающий разработку линейной, степенной и экспоненциальной функциональной зависимости (рис. 1). Определение параметров математических моделей осуществляется с использованием метода наименьших квадратов [4].

В частности, для определения параметров линейной функциональной зависимости:

$$y_i = a_0 + a_1 \cdot x_{1i} + a_2 \cdot x_{2i} + \\ + a_3 \cdot x_{3i} + a_4 \cdot x_{4i} + a_5 \cdot x_{5i} \quad (2)$$

$$\begin{cases}
n \cdot a_0 + a_1 \cdot \sum_{i=1}^{n} x_{1i} + a_2 \cdot \sum_{i=1}^{n} x_{2i} + a_3 \cdot \sum_{i=1}^{n} x_{3i} + a_4 \cdot \sum_{i=1}^{n} x_{4i} + a_5 \cdot \sum_{i=1}^{n} x_{5i} = \sum_{i=1}^{n} y_i \\
a_0 \cdot \sum_{i=1}^{n} x_{1i} + a_1 \cdot \sum_{i=1}^{n} x_{1i}^2 + a_2 \cdot \sum_{i=1}^{n} x_{1i} \cdot x_{2i} + a_3 \cdot \sum_{i=1}^{n} x_{1i} \cdot x_{3i} + a_4 \cdot \sum_{i=1}^{n} x_{1i} \cdot x_{4i} + a_5 \cdot \sum_{i=1}^{n} x_{1i} \cdot x_{5i} = \sum_{i=1}^{n} y_i \cdot x_{1i} \\
a_0 \cdot \sum_{i=1}^{n} x_{2i} + a_1 \cdot \sum_{i=1}^{n} x_{2i} \cdot x_{1i} + a_2 \cdot \sum_{i=1}^{n} x_{2i}^2 + a_3 \cdot \sum_{i=1}^{n} x_{2i} \cdot x_{3i} + a_4 \cdot \sum_{i=1}^{n} x_{2i} \cdot x_{4i} + a_5 \cdot \sum_{i=1}^{n} x_{2i} \cdot x_{5i} = \sum_{i=1}^{n} y_i \cdot x_{2i} \\
a_0 \cdot \sum_{i=1}^{n} x_{3i} + a_1 \cdot \sum_{i=1}^{n} x_{3i} \cdot x_{1i} + a_2 \cdot \sum_{i=1}^{n} x_{3i} \cdot x_{2i} + a_3 \cdot \sum_{i=1}^{n} x_{3i}^2 + a_4 \cdot \sum_{i=1}^{n} x_{3i} \cdot x_{4i} + a_5 \cdot \sum_{i=1}^{n} x_{3i} \cdot x_{5i} = \sum_{i=1}^{n} y_i \cdot x_{3i} \\
a_0 \cdot \sum_{i=1}^{n} x_{4i} + a_1 \cdot \sum_{i=1}^{n} x_{4i} \cdot x_{1i} + a_2 \cdot \sum_{i=1}^{n} x_{4i} \cdot x_{2i} + a_3 \cdot \sum_{i=1}^{n} x_{4i} \cdot x_{3i} + a_4 \cdot \sum_{i=1}^{n} x_{4i}^2 + a_5 \cdot \sum_{i=1}^{n} x_{4i} \cdot x_{5i} = \sum_{i=1}^{n} y_i \cdot x_{4i} \\
a_0 \cdot \sum_{i=1}^{n} x_{5i} + a_1 \cdot \sum_{i=1}^{n} x_{5i} \cdot x_{1i} + a_2 \cdot \sum_{i=1}^{n} x_{5i} \cdot x_{2i} + a_3 \cdot \sum_{i=1}^{n} x_{5i} \cdot x_{3i} + a_4 \cdot \sum_{i=1}^{n} x_{5i} \cdot x_{4i} + a_5 \cdot \sum_{i=1}^{n} x_{5i}^2 = \sum_{i=1}^{n} y_i \cdot x_{5i}
\end{cases}.$$

Для нелинейных зависимостей предусмотрены преобразования, позволяющие привести нелинейную зависимость к линейному виду и сформировать соответствующие системы уравнений для определения параметров математических моделей.

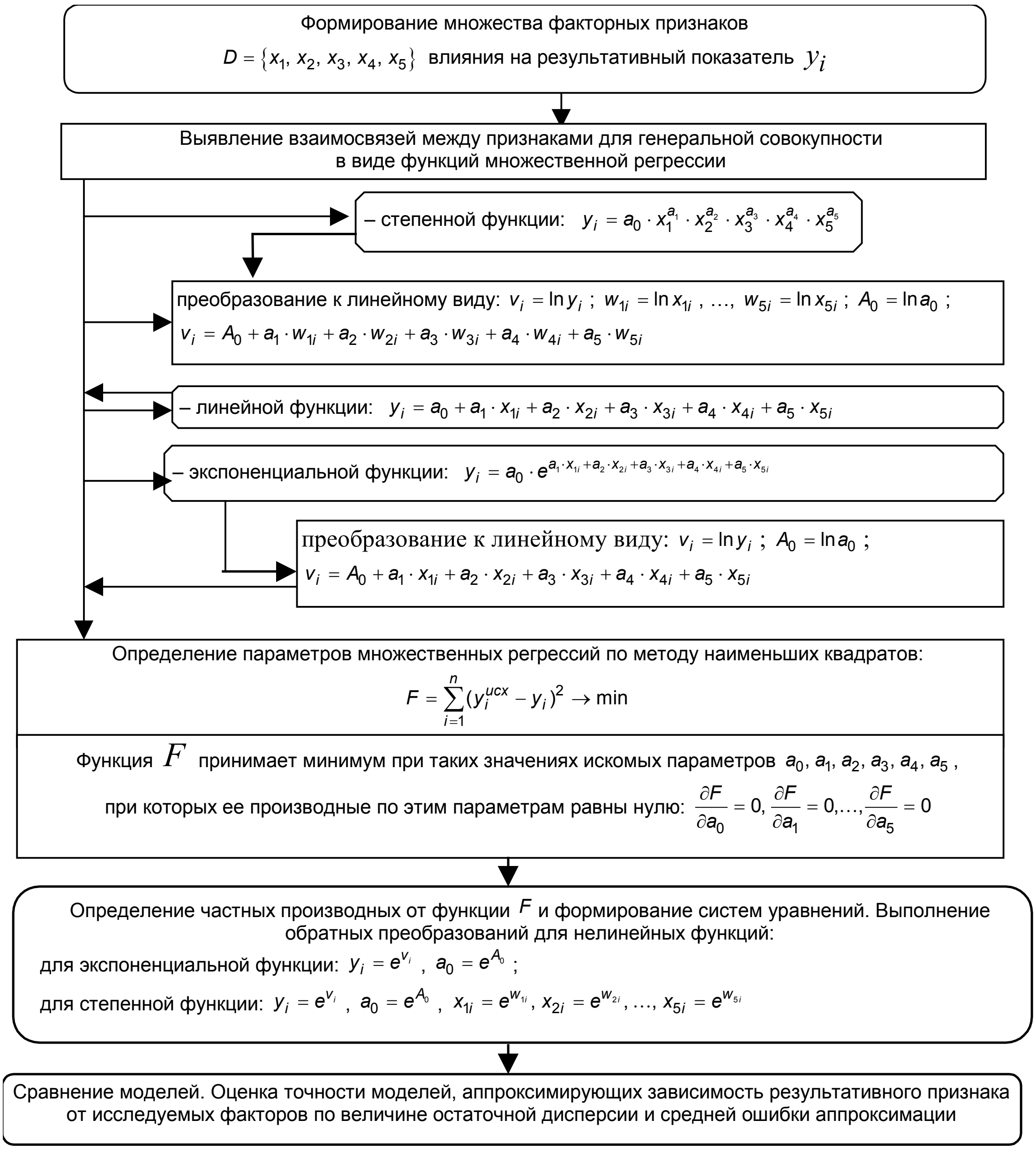

Рис. 1. Алгоритм выявления экономических взаимосвязей

В частности, для степенной зависимости:

$$\begin{cases} n\cdot A_0+a_1\cdot\sum_{i=1}^n w_{1i}+a_2\cdot\sum_{i=1}^n w_{2i}+a_3\cdot\sum_{i=1}^n w_{3i}+a_4\cdot\sum_{i=1}^n xw_{4i}+a_5\cdot\sum_{i=1}^n w_{5i}=\sum_{i=1}^n v_i \\ A_0\cdot\sum_{i=1}^n w_{1i}+a_1\cdot\sum_{i=1}^n w_{1i}^2+a_2\cdot\sum_{i=1}^n w_{1i}\cdot w_{2i}+a_3\cdot\sum_{i=1}^n w_{1i}\cdot w_{3i}+a_4\cdot\sum_{i=1}^n w_{1i}\cdot w_{4i}+a_5\cdot\sum_{i=1}^n w_{1i}\cdot w_{5i}=\sum_{i=1}^n v_i\cdot w_{1i} \\ A_0\cdot\sum_{i=1}^n w_{2i}+a_1\cdot\sum_{i=1}^n w_{2i}\cdot w_{1i}+a_2\cdot\sum_{i=1}^n w_{2i}^2+a_3\cdot\sum_{i=1}^n w_{2i}\cdot w_{3i}+a_4\cdot\sum_{i=1}^n w_{2i}\cdot w_{4i}+a_5\cdot\sum_{i=1}^n w_{2i}\cdot w_{5i}=\sum_{i=1}^n v_i\cdot w_{2i} \\ A_0\cdot\sum_{i=1}^n w_{3i}+a_1\cdot\sum_{i=1}^n w_{3i}\cdot w_{1i}+a_2\cdot\sum_{i=1}^n w_{3i}\cdot w_{2i}+a_3\cdot\sum_{i=1}^n w_{3i}^2+a_4\cdot\sum_{i=1}^n w_{3i}\cdot w_{4i}+a_5\cdot\sum_{i=1}^n w_{3i}\cdot w_{5i}=\sum_{i=1}^n v_i\cdot w_{3i} \\ A_0\cdot\sum_{i=1}^n w_{4i}+a_1\cdot\sum_{i=1}^n w_{4i}\cdot w_{1i}+a_2\cdot\sum_{i=1}^n w_{4i}\cdot w_{2i}+a_3\cdot\sum_{i=1}^n w_{4i}\cdot w_{3i}+a_4\cdot\sum_{i=1}^n w_{4i}^2+a_5\cdot\sum_{i=1}^n w_{4i}\cdot w_{5i}=\sum_{i=1}^n v_i\cdot w_{4i} \\ A_0\cdot\sum_{i=1}^n w_{5i}+a_1\cdot\sum_{i=1}^n w_{5i}\cdot w_{1i}+a_2\cdot\sum_{i=1}^n w_{5i}\cdot w_{2i}+a_3\cdot\sum_{i=1}^n w_{5i}\cdot w_{3i}+a_4\cdot\sum_{i=1}^n w_{5i}\cdot w_{4i}+a_5\cdot\sum_{i=1}^n w_{5i}^2=\sum_{i=1}^n v_i\cdot w_{5i} \end{cases}$$

и для экспоненциальной зависимости:

$$\begin{cases} n\cdot A_0+a_1\cdot\sum_{i=1}^n x_{1i}+a_2\cdot\sum_{i=1}^n x_{2i}+a_3\cdot\sum_{i=1}^n x_{3i}+a_4\cdot\sum_{i=1}^n x_{4i}+a_5\cdot\sum_{i=1}^n x_{5i}=\sum_{i=1}^n v_i \\ A_0\cdot\sum_{i=1}^n x_{1i}+a_1\cdot\sum_{i=1}^n x_{1i}^2+a_2\cdot\sum_{i=1}^n x_{1i}\cdot x_{2i}+a_3\cdot\sum_{i=1}^n x_{1i}\cdot x_{3i}+a_4\cdot\sum_{i=1}^n x_{1i}\cdot x_{4i}+a_5\cdot\sum_{i=1}^n x_{1i}\cdot x_{5i}=\sum_{i=1}^n v_i\cdot x_{1i} \\ A_0\cdot\sum_{i=1}^n x_{2i}+a_1\cdot\sum_{i=1}^n x_{2i}\cdot x_{1i}+a_2\cdot\sum_{i=1}^n x_{2i}^2+a_3\cdot\sum_{i=1}^n x_{2i}\cdot x_{3i}+a_4\cdot\sum_{i=1}^n x_{2i}\cdot x_{4i}+a_5\cdot\sum_{i=1}^n x_{2i}\cdot x_{5i}=\sum_{i=1}^n v_i\cdot x_{2i} \\ A_0\cdot\sum_{i=1}^n x_{3i}+a_1\cdot\sum_{i=1}^n x_{3i}\cdot x_{1i}+a_2\cdot\sum_{i=1}^n x_{3i}\cdot x_{2i}+a_3\cdot\sum_{i=1}^n x_{3i}^2+a_4\cdot\sum_{i=1}^n x_{3i}\cdot x_{4i}+a_5\cdot\sum_{i=1}^n x_{3i}\cdot x_{5i}=\sum_{i=1}^n v_i\cdot x_{3i} \\ A_0\cdot\sum_{i=1}^n x_{4i}+a_1\cdot\sum_{i=1}^n x_{4i}\cdot x_{1i}+a_2\cdot\sum_{i=1}^n x_{4i}\cdot x_{2i}+a_3\cdot\sum_{i=1}^n x_{4i}\cdot x_{3i}+a_4\cdot\sum_{i=1}^n x_{4i}^2+a_5\cdot\sum_{i=1}^n x_{4i}\cdot x_{5i}=\sum_{i=1}^n v_i\cdot x_{4i} \\ A_0\cdot\sum_{i=1}^n x_{5i}+a_1\cdot\sum_{i=1}^n x_{5i}\cdot x_{1i}+a_2\cdot\sum_{i=1}^n x_{5i}\cdot x_{2i}+a_3\cdot\sum_{i=1}^n x_{5i}\cdot x_{3i}+a_4\cdot\sum_{i=1}^n x_{5i}\cdot x_{4i}+a_5\cdot\sum_{i=1}^n x_{5i}^2=\sum_{i=1}^n v_i\cdot x_{5i} \end{cases}$$

Практическая реализация разработанного алгоритма, а также комплексная проверка точности полученных моделей выполнена с использованием пакета анализа программы EXCEL. На рис. 2 схематически представлена методика выбора адекватной математической модели с проверкой значимости и объяснением вариации результативного признака.

Для оценки точности полученных многофакторных моделей – экономических взаимосвязей между результативным признаком и факторами, используется величина остаточной дисперсии [4, с. 33]:

$$S_{ocm}^2=\frac{\sum_{i=1}^n (y_i^{ucx}-y_i)^2}{n-(m+1)}, \qquad (3)$$

где y_i^{ucx}, y_i – соответственно, исходные (эмпирические) и теоретические значения результативного признака;

n – количество данных в выборке;

m – количество факторов;

$m+1$ – количество параметров в уравнении регрессии.

Для сравнения и оценки точности моделей используем также величину средней ошибки аппроксимации:

$$|\bar{e}|=\frac{1}{n}\cdot\sum_{i=1}^n\left|\frac{y_i^{ucx}-y_i}{y_i^{ucx}}\right|\cdot 100\%. \qquad (4)$$

Модель, для которой значения остаточной дисперсии и величина средней ошибки аппроксимации меньше является более точной, так как меньше рассеивание результативного признака относительно условного математического ожидания. В табл. 1 представлены результаты расчетов средней ошибки аппроксимации, которые свидетельствуют о большей точности нелинейных моделей.

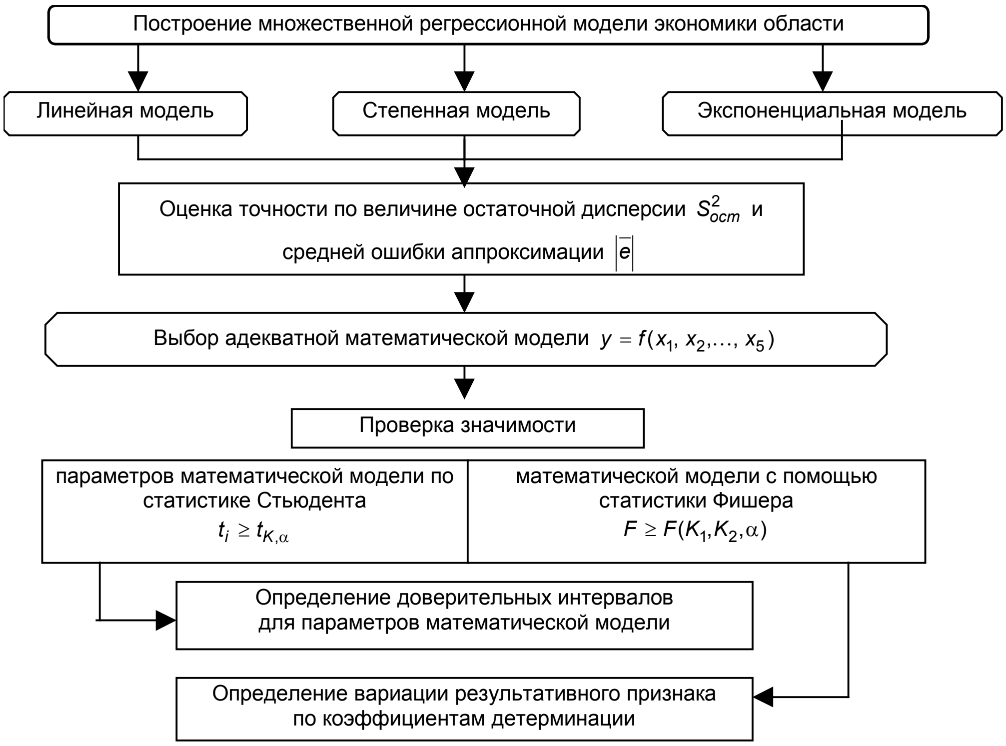

Рис. 2. Выбор адекватной математической модели с проверкой значимости и объяснением вариации результативного признака

Таблица 1

Оценки точности моделирования экономики области

Показатель точности моделей	Линейная функция	Степенная функция	Экспоненциальная функция
Средняя ошибка аппроксимации	18,14%	0,006%	0,0999%

После выявления более точных математических моделей выполняем оценку значимости их параметров (коэффициентов регрессии), а также определяем значимость полученных уравнений множественной регрессии.

Проверка значимости оценок параметров регрессии осуществляется следующим образом. Для каждого коэффициента регрессии a_i вычисляем ошибки коэффициентов регрессии S_{ai} и статистику Стьюдента [4, с. 57]:

$$t_i = \frac{|a_i|}{S_{ai}}, \; i = 0, 1, 2, \ldots, m, \qquad (5)$$

и сравниваем ее с критическим значением $t_{K,\alpha}$, при заданном уровне значимости α и числе степеней свободы K. Если $t_i \geq t_{K,\alpha}$, то предположение о равенстве нулю коэффициента регрессии a_i отвергается, и его считаем значимым. В случае, если $t_i < t_{K,\alpha}$, то оснований отвергать данное предположение нет, и поэтому оценку a_i будем считать незначимой.

При уровне значимости $\alpha = 0{,}05$ и $K = n - m - 1 = 6$ находим критическое значение статистики Стьюдента: $t_{6;0.05} = 2{,}45$ [4, табл. 2.13, с. 47].

Поскольку все расчетные $t_{a,i}$ больше критического значения, гипотезу о равенстве нулю коэффициентов регрессии отвергаем, и оценки $a_0, a_1, \ldots, a_5$ соответствующих параметров признаем значимыми (фрагмент результатов вычислений для экспоненциальной модели представлен в табл. 2). Наряду с проверкой значимости каждого коэффициента регрессии выполним оценку значимости полученной модели – уравнения множественной регрессии.

Таблица 2

Ошибки коэффициентов регрессии

ta_0	ta_1	ta_2	ta_3	ta_4	ta_5
300,7389	29,10	20,78	2,71	23,62	42,05

На основе полученных результатов исследований можно высказать гипотезу о том, что все коэффициенты регрессии, кроме a_0, равны нулю (эта гипотеза называется нулевой и обозначается H_0 [4, с. 57]). Проверка гипотезы H_0 осуществляется с помощью статистики Фишера:

$$F = \frac{(Q - Q_{ост})/K_1}{Q_{ост}/K_2}, \qquad (6)$$

где F – расчетное значение статистики Фишера;

$Q, Q_{ост}$ – сумма квадратов отклонений результативного признака соответственно от среднего значения и от условно среднего y_i;

K_1 и K_2 – степени свободы $K_1 = m$, $K_2 = n - m - 1$.

При заданном уровне значимости $\alpha = 0{,}05$ для степеней свободы $K_1 = 5$ и $K_2 = 6$ по таблице F-распределения Фишера находим критическое значение $F(K_1, K_2, \alpha) = F(5;6;0{,}05) = 4{,}39$ [4, табл. 2.15, с. 58] и сравниваем его с расчетным значением *F*. Для принятия решения о значимости модели используем известные условия [4, с. 57–58]. Если $F \geq F(K_1, K_2, \alpha)$, то гипотезу H_0 об одновременном равенстве нулю всех коэффициентов регрессии отвергаем и уравнение считаем значимым. Если $F < F(K_1, K_2, \alpha)$, то уравнение регрессии считаем незначимым, то есть отвергается влияние факторных признаков $x_1, x_2, \ldots, x_m$ на результативный. Расчетные значения статистики Фишера составляют: для степенной модели – *F = 781,9,* для экспоненциальной – *F = 822,6*. Полученные значения значительно превышают критическое значение $F(5;6;0{,}05) = 4{,}39$, что и является доказательством значимости нелинейных многофакторных моделей. Вероятность того, что гипотеза H_0 справедлива, составляет 0,05. Таким образом, при $F >> F(K_1, K_2, \alpha)$ все коэффициенты могут иметь нулевые значения с вероятностью, меньшей 0,05. На рис. 3 представлена графическая интерпретация результатов моделирования.

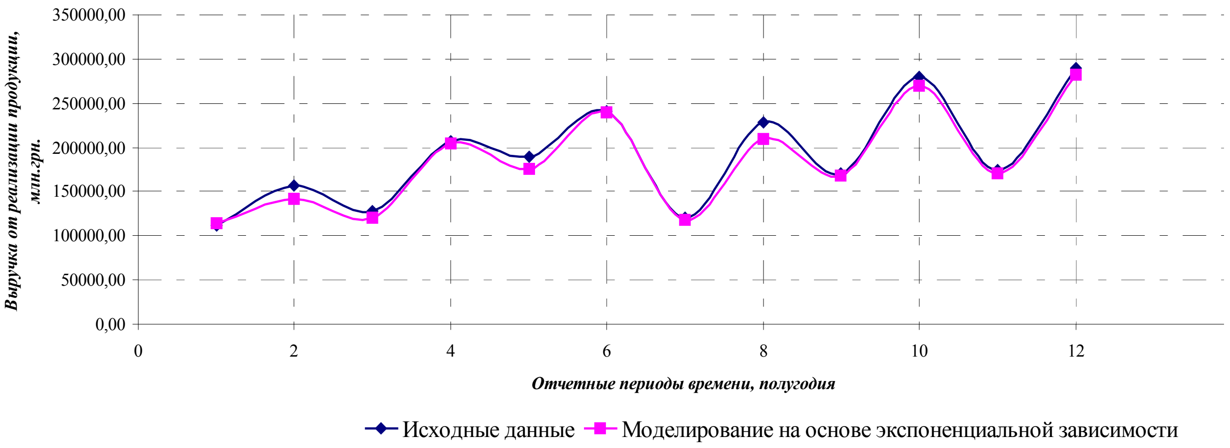

Рис. 3. Графическая интерпретация результатов моделирования на основе экспоненциальной зависимости $y_i = 81122{,}17 \cdot e^{0{,}00002 \cdot x_1 - 0{,}00009 \cdot x_2 - 0{,}00007 \cdot x_3 - 0{,}00012 \cdot x_4 + 0{,}0001 \cdot x_5}$

Дальнейшие исследования направлены на изучение вариации результативного признака. Как известно [4, с. 50], при вычислении парного коэффициента корреляции r_{yx_i} вся объясняемая им вариация результативного признака y_i обусловлена изменением факторного признака x_i. Частный коэффициент корреляции отражает тесноту корреляционной связи только между *y* и x_i, при этом влияние других признаков на результативный показатель *y* исключено. Совокупную связь результативного признака со всеми факторными признаками можно оценить с помощью коэффициента множественной корреляции $R_{yx_1x_2\ldots y_m}$. Тесноту связи оцениваем в долях полной дисперсии с помощью коэффициента детерминации [4, с. 38]. Результаты расчетов, представленные в табл. 3, объясняют вариацию результативного признака с помощью отдельных факторов и их сочетаний.

Результаты исследований показывают, что в действительности корреляционная связь между результативным и факторными признаками сильнее, чем это следует из значений парных коэффициентов корреляции. Коэффициент множественной детерминации $R^2 = 0{,}9854$, то есть факторы x_1, x_2, x_3, x_4, x_5 на 98,54% объясняют вариацию выручки от реализации продукции, а влияние других (неучтенных) факторов составляет 1,46%.

Таблица 3

Объяснение вариации результативного признака

Объяснение вариации результативного признака с помощью отдельных факторов				
X_1	X_2	X_3	X_4	X_5
92,18 %	90,74 %	67,36 %	65,40 %	90,08

Объяснение вариации результативного признака с помощью сочетания отдельных факторов			
X_1 и X_2	X_1 и X_3	X_1 и X_4	X_1 и X_5
92,69 %	92,44 %	93,77 %	92,99 %

IV. Выводы

На основе разработанной множественной регрессионной модели экономики Днепропетровской области установлена взаимосвязь между выручкой от реализации и пятью наиболее значимыми факторами: материальные затраты, затраты на оплату труда, затраты на топливо и энергию, амортизация и другие операционные затраты по реализации продукции. Выявленные взаимосвязи составляют основу принятия решений по формированию стратегии развития экономики области.

Список использованной литературы

1. Паршин Ю.І. Моделювання економічного розвитку Дніпропетровської області / Ю.І. Паршин // Інноваційно-інвестиційна політика сталого розвитку регіонів України: від теорії до практики : колективна монографія : в 2-х т. / під заг. ред. К.Ф. Ковальчука. – Дніпропетровськ : ІМА-прес, 2012. – Т. 2. – С. 141–151.
2. Статистичний щорічник Дніпропетровської області за 2009 р. / Держ. ком. статистики України – Д. : Головне упр. стат. у Дніпропетровській обл., 2010. – 535 с.
3. Статистичний щорічник Дніпропетровської області за 2010 рік / Держ. ком. статистики України – Д. : Головне упр. стат. у Дніпропетровській обл., 2011. – 515 с.
4. Кухарев В.Н. Экономико-математические методы и модели в планировании и управлении : учебник / В.Н. Кухарев, В.И. Салли, А.М. Эрперт. – К. : Выща шк., 1991. – 303 с.

Статья поступила в редакцию 13.12.2012.

Паршин Ю.І. Розробка множинної регресійної моделі функціонування економіки Дніпропетровської області

Запропоновано алгоритм виявлення економічних взаємозв'язків та розроблено множинну регресійну модель функціонування економіки Дніпропетровської області.

Ключові слова: *множинна регресійна модель функціонування економіки, алгоритм виявлення економічних взаємозв'язків.*

Parshin Y. Development of multiple regressive model of functioning of economy of the Dnepropetrovsk area

The algorithm of definition of economic relationship is offered and the multiple regressive model of functioning of economy of the Dnepropetrovsk area is developed.

Key words: *multiple regressive model of functioning of economy, algorithm of definition of economic relationship.*