УДК 338.27

Г.А. Гадельшина, А.В. Аксянова

ПРОГНОЗИРОВАНИЕ ПРИБЫЛИ ПРЕДПРИЯТИЯ С ПОМОЩЬЮ МУЛЬТИТРЕНДОВОЙ МОДЕЛИ

Ключевые слова: прогнозирование, мультитрендовая модель, мультиномиальная логит-модель, авторегрессионная модель.

Применен метод, предложенный В.В.Давнисом, основанный на совместном использовании авторегрессионной модели с фиктивными переменными и мультиномиальной логит-модели для прогнозирования прибыли предприятия. Полученная в результате комбинированная модель позволяет строить прогноз с помощью вероятностного распределения траекторий мультитрендовой модели. На основе данной модели получены прогнозные значения прибыли ОАО «Казанский вертолетный завод».

Keywords: forecasting, multitrend model, multinomial logit (MNL) model, autoregressive (AR) model.

The method proposed by Davnis V.V., was applied for income forecast of Kazan Helicopters Joint Stock Company. The method is based on the joint use of autoregressive model with dummy variable and the model of the expert preferences. A combined model allows to build the forecast with the help of the probability distribution of the trajectories of multitrend model.

Перспективный анализ финансового состояния предприятия представляет собой изучение его финансово-хозяйственной деятельности с целью определения финансового состояния этого предприятия в будущем. Перечень прогнозируемых показателей может ощутимо варьировать. В качестве объектов прогнозирования могут выступать, например, выручка от продаж, прибыль, себестоимость продукции.

В зависимости от вида используемой модели все методы прогнозирования можно подразделить на группы. Первая группа включает в себя методы *экспертных оценок*, которые предусматривают многоступенчатый опрос экспертов по специальным схемам и обработку полученных результатов с помощью инструментария экономической статистики. Вторая группа использует *стохастические методы*, предполагающие вероятностный характер как прогноза, так и самой связи между исследуемыми показателями. Вероятность получения точного прогноза растет с ростом числа эмпирических данных. Эти методы занимают ведущее место с позиции формализованного прогнозирования и существенно варьируют по сложности используемых алгоритмов.

Стохастические методы можно разделить на три типовые группы. К первой группе относятся динамические и авторегрессионные модели, применяемые для анализа данных, представленных в виде временного ряда. Вторая группа используется для анализа данных, представленных в виде пространственно-временной совокупности, когда ряды динамики недостаточны по своей длине для построения статистически значимых прогнозов и в прогнозе следует учесть влияние факторов, различающиеся по экономической природе и их динамике. В этом случае применяется многофакторный регрессионный анализ. К третьей группе относятся детерминированные методы, предполагающие наличие функциональных связей. /1/

В том случае, когда выборка содержит достаточно большое количество наблюдений, либо известны факторы, оказывающие существенное влияние на наблюдаемую величину, лучше всего работают первая и вторая группы методов. Для них существуют хорошо разработанный математический аппарат, позволяющий строить прогнозы с достаточной степенью точности. /2/ В противоположной ситуации, если исходные данные практически отсутствуют, для оценки исследуемого значения привлекают экспертов. Существенным недостатком данного метода является низкая надежность оценки экспертов.

В.В.Давнис в своей работе /3/ предложил механизм совмещения экспертных оценок и методов экономического прогнозирования. Главная особенность классических методов в том, что при опросе экспертов для числового представления получаемых качественных результатов, в основном, используются номинальные и ранговые шкалы. Новый взгляд на проблему экспертного оценивания заключается в том, чтобы обосновать необходимость и полезность получения результатов опроса в виде моделей, концентрирующих в своей аналитической структуре субъективные предпочтения экспертов.

Таким образом, экспертная информация используется для построения модели, с помощью которой будут получены оценки, а не для непосредственного получения самих оценок. Интуиция и знания экспертов применяются для формирования специальных выборочных совокупностей, по которым оцениваются коэффициенты модели, имеющей, в отличие от непосредственных экспертных оценок, многоплановое применение: анализ, оценка значимости факторов, прогноз ожидаемых событий и т.п. /3/.

При исследовании связей между социально-экономическими явлениями зависимая переменная довольно часто является дискретной, т.е. принимает значения 0, 1, 2, 3,…. При этом зависимая переменная может быть количественной, качественной и порядковой (ранговой).

В классе моделей бинарного выбора зависимая переменная может принимать только два значения: $y_i = 0$ и $y_i = 1$. Модели множественного выбора позволяют моделировать зависимость между переменной, определяющей более двух возможных состояний характеризуемого объекта, и одной или более независимыми переменными. Следовательно, вектор $Y = (y_1, y_2, \ldots, y_n)$ исходных статистических данных будет содержать только дискретные значения признаков. Для исследования зависимости $Y = (y_1, y_2, \ldots, y_n)$ от ряда объясняющих переменных $X = (x_{i1}, x_{i2},..., x_{ik})$ может быть использована модель линейной регрессии:

$$y_i = \sum_{j=0}^{k} x_{ij}\beta_j + \varepsilon_i, \quad \mathrm{i} = 1,\ldots, \mathrm{n} \qquad (1)$$

где i – номер наблюдения; $\beta = (\beta_0, \beta_1, \beta_2,..., \beta_k)$ – набор неизвестных параметров; ε_i – случайная ошибка.

В том случае, когда зависимая переменная бинарная, y_i принимает значения 0 или 1 и математическое ожидание $E(\varepsilon_i) = 0$, то можно записать:

$$E(y_i) = 1 \cdot P(y_i = 1) + 0 \cdot P(y_i = 0) = P(1) = x_i\beta \qquad (2)$$

Таким образом, модель (2) может быть записана в виде

$$P(y_i = 1) = x_i\beta \qquad (3)$$

Для моделирования значений $P(y_i = 1)$ подбирают функции, область значений которых определяется отрезком[0; 1], а $x_i\beta$ играет роль аргумента этой функции, т.е.

$$P(y_i = 1) = F(x_i\beta) \qquad (4)$$

Функция $F(x_i\beta)$ должна быть непрерывной, неубывающей функцией. Известны разные интерпретации модели (4). Один из подходов основан на введении некоторой ненаблюдаемой, или латентной переменной y^*, изменяющейся от $-\infty$ до $+\infty$ и порождающей наблюдаемую зависимую переменную y. Пусть латентная переменная y^* линейно зависит от вектора объясняющих переменных:

$$y_i^* = x_i\beta + \varepsilon_i. \qquad (5)$$

В том случае, когда y_i может принимать некоторое множество значений, латентная переменная y^* связана с переменной y следующей системой уравнений:

$$y_i = \begin{cases} 0, & y^* \le 0 \\ 1, & 0 < y^* \le \mu_1 \\ 2, & \mu_1 < y^* \le \mu_2 \\ \ldots & \\ J, & y^* > \mu_{J-1} \end{cases} \qquad (6)$$

Выбор функции $F(x_i\beta)$ определяет тип модели. Наиболее часто в качестве функции $F(x_i\beta)$ используют:

– стандартное нормальное распределение

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{\frac{z^2}{2}} dz \qquad (7)$$

и соответствующую модель называют пробит-моделью (probit-model);

– логистическую функцию

$$F(u) = \Lambda(u) = \frac{e^u}{1 + e^u} \qquad (8)$$

и соответствующую модель называют логит-моделью (logit-model).

Логистическое распределение имеет тенденцию давать большие, чем нормальное распределение, вероятности для очень малых $x_i\beta$ и меньшие вероятности для очень больших значений $x_i\beta$. Для выборок с небольшим разбросом объясняющих переменных качественные выводы, полученные при использовании пробит- и логит-моделей, совпадают.

Таким образом, вероятность выбора j-й альтернативы $y_i = j$ – это вероятность того, что $\mu_{j-1} < y^*_i < \mu_j$. (1)

Поскольку плотность распределения неотрицательна, направление изменения эффекта зависит только от знака коэффициента. Положительные значения показывают, что вероятность прогноза зависимой переменной увеличится, а отрицательные – что вероятность прогноза понизится.

Процедуры оценивания пробит- и логит-моделей реализованы в большинстве современных эконометрических прикладных программах, например, «STATISTICA» и SPSS.

Для оценки качества модели используются два аналога R^2 для линейной регрессии: pseudo R^2 и McFadden R^2. /4/

Построение комбинированной прогнозной модели осуществляется в несколько этапов. На первом этапе строится трендовая или авторегрессионная модель. Трендовые модели целесообразнее применять в случаях, когда данные можно с достаточной степенью точности описать какой-либо аналитической зависимостью. Авторегрессионные модели, как правило, обеспечивают более высокий уровень адекватности, когда данные можно описать в виде зависимости текущих уровней ряда от предыдущих. Однако они позволяют строить достоверные прогнозы только на один временной интервал.

Экстраполяционная составляющая авторегрессионной модели порядка m может быть записана следующим образом:

$$\tilde{y}_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_m y_{t-m} + \varepsilon_t. \qquad (9)$$

Следующий шаг предусматривает включение в модель механизма, реализующего расщепление траектории на варианты ожидаемого будущего. Предполагается, что возможно несколько вариантов развития процесса в будущем и однозначное описание предшествующего периода (в соответствии с экспериментальными данными).

Существует несколько способов, позволяющих получить формальное описание подобного механизма. Один из них - мультитрендовая модель с pretest-оценками /5/.

Эта модель строится на базе авторегрессионной модели. Сначала оцениваются коэффициенты уравнения (9). Если построенная модель не устраивает по точности, то вычисляются отклонения расчетных значений от фактических

$$y_t - \tilde{y}_t, \; t = \overline{m+1, T}, \qquad (10)$$

которые позволяют выборочное множество наблюдений разделить на две части, соответствующие различным вариантам по следующему правилу:

$$\begin{cases} y_t - \tilde{y}_t \le 0, & t \in I_0, \\ y_t - \tilde{y}_t > 0, & t \in I_1. \end{cases} \qquad (11)$$

На основе полученного деления формируется вектор f_1 (фиктивная переменная) из нулей и единиц таким образом, что на местах с номерами $t \in I_0$ будут стоять нули, а с номерами $t \in I_1$ - единицы. Введение новой переменной приводит к модели вида:

$$\tilde{y}_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_m y_{t-m} + d_1 f_{1t} + \varepsilon_t,\quad t = \overline{m+1,T}, \qquad (12)$$

в которой d_1 - коэффициент при фиктивной переменной; f_{1t} - значение фиктивной переменной в момент t. Модель с одной фиктивной переменной дает расщепление траектории на две, различие между которыми определяется идентифицируемыми эффектами

Далее вновь производится оценка коэффициентов уравнения (12). Если требуемый уровень не достигнут, то после получения расчетных значений $\tilde{y}_t$ вновь осуществляется деление каждого из подмножеств I_0 и I_1 на две части по следующему правилу:

$$\begin{cases} \textit{если } y_t - \tilde{y}_t \le 0, & t \in I_0, \quad \text{то} \quad t \in I_{00} \\ \textit{если } y_t - \tilde{y}_t > 0, & t \in I_0, \quad \text{то} \quad t \in I_{01} \\ \textit{если } y_t - \tilde{y}_t \le 0, & t \in I_1, \quad \text{то} \quad t \in I_{10} \\ \textit{если } y_t - \tilde{y}_t > 0, & t \in I_1, \quad \text{то} \quad t \in I_{11} \end{cases} \qquad (13)$$

Полученная классификация позволяет ввести вторую фиктивную переменную f_2, следуя правилу

$$\begin{cases} \textit{если } t \in I_{00}, & \text{то} \quad f_2 = 0 \\ \textit{если } t \in I_{01}, & \text{то} \quad f_2 = 1 \\ \textit{если } t \in I_{10}, & \text{то} \quad f_2 = 0 \\ \textit{если } t \in I_{11}, & \text{то} \quad f_2 = 1 \end{cases} \qquad (14)$$

Модель с двумя фиктивными переменными

$$\begin{aligned} \tilde{y}_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_m y_{t-m} + \\ + d_1 f_{1t} + d_2 f_{2t}, \qquad t = \overline{m+1,T} \end{aligned} \qquad (15)$$

обеспечивает получение четырех прогнозных траекторий. При достижении требуемого уровня точности оценки коэффициентов d_k ($k = \overline{1,m}$) проверяются на статистическую значимость по t - критерию. Уровень значимости должен быть выше общепринятого 0,95.

Незначимые фиктивные переменные исключаются из модели, и ее окончательный вариант оценивается только со статистически значимыми переменными. Полученная таким образом модель называется мультитрендовой. Вычисленные оценки параметров тренда $a_0, a_1, \ldots, a_m$ называются pretest-оценками.

Далее необходимо дополнить результаты мультитрендового моделирования вероятностными оценками правдоподобности прогнозных траекторий. С этой целью необходимо построить модель множественного выбора /3/.

Всем вариантам траекторий мультитрендовой модели присваиваются произвольные номера 0, 1, 2,…, J. Вероятность возможной реализации того или иного варианта в будущем описывается мультиномиальной логит-моделью

$$\begin{aligned} \mathrm{P}(y_i = j) &= \frac{e^{x_i b_j}}{1 + \sum_{j=0}^{J-1} e^{x_i b_j}}, \quad j = \overline{1, J-1}, \\ \mathrm{P}(y_i = J) &= \frac{1}{1 + \sum_{j=0}^{J-1} e^{x_j b_j}} \end{aligned} \qquad (16)$$

Вектор независимых переменных x_i представляет собой показатели или экспертные оценки, по которым различаются альтернативные варианты. /2/

Определение коэффициентов модели осуществляется путем численного решения уравнений правдоподобия. На практике расчет коэффициентов осуществляется, например, в пакетах STATISTICA, Eviews.

Таким образом, комбинированная модель может быть записана следующим образом:

$$\begin{aligned} &\mathrm{P}\left(\tilde{y}_{t+1}^k\right) = \mathrm{P}_{t+1}^k, \\ &\tilde{y}_{t+1}^k = a_0 + a_1 y_t + a_2 y_{t-1} + \cdots + a_m y_{t-m+1} + \vec{f}^k \vec{d} \\ &\mathrm{P}_{t+1}^k = \frac{\exp\left(x_{t+1} \vec{b}^k\right)}{1 + \sum_{k=0}^{K} \exp\left(x_{t+1} \vec{b}^k\right)}, \end{aligned} \qquad (17)$$

где y_{t+1}^k - k-й вариант прогнозной оценки; y_t - запаздывающие значения зависимой переменной; a_j - оценка j-го коэффициента регрессии; P^k - вероятность реальности k-го варианта прогнозной оценки; $\vec{f}^k$ - вектор значений, которые в k-м варианте приняли фиктивные переменные (например, в случае двух фиктивных переменных f_1 и f_2 имеем набор из четырех векторов $f^0 = (0{,}0)$, $f^1 = (0{,}1)$, $f^2 = (1{,}0)$, $f^0 = (0{,}0)$; $\vec{d}$ - вектор коэффициентов при фиктивных переменных; $\vec{b}^k$ - оценка вектора параметров логит-модели множественного выбора k-го варианта; x_{t+1} - вектор значений, описывающий условия, ожидаемые в упреждающем периоде.

В данной работе комбинированная модель была применена для определения ожидаемой прибыли Открытого акционерного общества «Казанский вертолетный завод» (ОАО «КВЗ») на 2012 год.

Динамика прибыли ОАО «КВЗ» за период с0 2002 по 2011 гг. приведена в таблице 1 /6/.

Таблица 1 – Динамика прибыли ОАО «КВЗ»

Год	Прибыль, тыс.руб.
2002	287 532
2003	348 451
2004	585 018
2005	757 835
2006	-97 092
2007	573 787
2008	1 499 063
2009	1 996 379
2010	5 255 752
2011	9 712 120

В соответствие с этапами разработки модели, изложенными ранее, сначала необходимо определить порядок авторегрессии.

С помощью регрессионного анализа пакета «Анализ данных» MS Excel были построены авторегрессионные модели первого и второго порядка. При оценке значимости модели оказалось, что коэффициенты модели второго порядка незначимы. Следовательно, прогнозные расчеты необходимо осуществлять с использованием авторегрессионной модели первого порядка:

$$y_t = -57939.27 + 1{,}888 y_{t-1} \qquad (18)$$
$$(-0.14) \quad (9.30)$$

Для этой модели коэффициент детерминации R^2=0.93, F=86.4, что позволяет сделать вывод о его адекватности, коэффициент b_1 значим. Используя полученное уравнение, прогнозное значение прибыли составит 18 274 млн. р.

Хотя расчетные и экспериментальные данные согласуются достаточно хорошо, качество модели можно повысить путем введения фиктивной переменной f_1, которая поможет учесть влияние случайных эффектов.

Для получения значений f_1 вычисляются отклонения расчетных значений от фактических по правилу (11). В случае, когда отклонение положительно переменная f_1 принимает значение 1, если нет – 0.

Введение новой переменной привело к следующему регрессионному уравнению:

$$y_t = -657007 + 1472164 f_1 + 1.97 y_{t-1}. \qquad (19)$$
$$(-2.20) \quad (3.55) \quad (15.55)$$

Все коэффициенты значимы, уравнение адекватно (R^2 = 0,976). Таким образом, данное уравнение пригодно для дальнейшего использования.

На рисунке 1 показаны исходные и расчетные данные. Они хорошо согласуются. При получении прогнозных значений на 2012 год траектория расщепляется на 2 возможных варианта: для f_1=0 и f_1=1. Соответствующие прогнозные значения прибыли составят 18 520 и 19 992 млн. руб. То есть мультитрендовая модель дает более высокие прогнозные значения по сравнения с обычной авторегрессионной моделью.

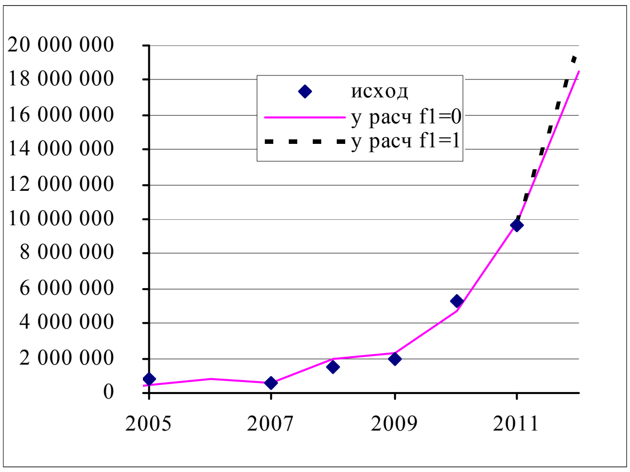

Рис. 1 – Сравнение исходных данных и расчетных значений с одной фиктивной переменной

Для определения вероятностей траекторий следует воспользоваться моделью множественного выбора. Так как в данном случае две траектории, можно использовать бинарную модель.

Экспертам было предложено оценить по стобалльной шкале экономическую ситуацию на заводе за прошедшие годы и на прогнозируемый период, то есть на 2012 год. Для получения экспертных оценок был проведен опрос сотрудников планово-экономического отдела ОАО «КВЗ» и по их оценкам рассчитаны средние баллы. Результаты экспертного оценивания представлены в таблице 2.

Таблица 2 – Оценки экспертов и результаты расчетов

Год	Баллы	f_1	$P(f_1=1)$
2002			
2003	15	0	0,044
2004	24	0	0,082
2005	38	0	0,199
2006	7	0	0,025
2007	46	1	0,310
2008	61	1	0,574
2009	41	0	0,237
2010	74	1	0,778
2011	72	0	0,751
2012	68		0,693

Построение логит-модели по данным таблицы 2 было проведено с помощью пакета «*STATISTICA*».

В результате была получена следующая модель

$$P(f_1 = 1) = \frac{1}{1 + \exp(4.178 - 0.073 \cdot x)}, \qquad (20)$$

где x – оценка экспертов по стобалльной шкале. Статистика Вальда свидетельствуют о значимости коэффициентов. Расчетные значения приведены в последнем столбце таблицы 2. Таким образом, данная модель дает прогнозное значение прибыли 19 992 млн. руб. с вероят6ностью 69,3%. Математическое ожидание прибыли составило 19 539 млн. руб.

Введение второй фиктивной переменной f_2 приводит к следующему уравнению:

$$y_t = -1223259 + 1673324 f_1 + 885159 f_2 + 2.059 y_{t-1} \qquad (21)$$
$$(-5.97) \quad (7.63) \quad (4.19) \quad (30.05)$$

Коэффициент детерминации повысился до 0.995, все коэффициенты уравнения (21) значимы. На рисунке 2 показаны результаты моделирования.

Использование двух фиктивных переменных приводит к расщеплению траектории на 4 линии в зависимости от сочетания значений f_1 и f_2. При f_1=0 и f_2 =0 прогнозное значение прибыли составляет 18 779 млн. руб., если f_1=0 и f_2 =1 расчетная прибыль 19 664 млн. руб., для f_1=1 и f_2 =0 прибыль 20 452 млн. руб., а для f_1=1 и f_2 =1 21 337 млн. руб. Таким образом, введение дополнительной фиктивной переменной привело к еще большему повышению прогнозных значений прибыли.

Для оценки вероятностей траекторий использовалась мультиномиальная логит-модель (16). Номера вариантов определялись в зависимости от комбинаций f_1 и f_2 следующим образом: f=0 для f_1=0 и f_2 =0, f=1 для f_1=0 и f_2 =1, f=2 для f_1=1 и f_2 =0 и f=3 для f_1=1 и f_2 =1. В результате расчетов была получена модель со следующими вероятностями вариантов.

$$\mathrm{P}(A_i)=\exp(1{,}28-0{,}02x_i)+$$
$$+\exp(1{,}94-0{,}05x_i)+\exp(-0{,}55+0{,}01x_i)$$
$$\mathrm{P}(f=0)=\frac{\exp(1{,}28-0{,}02x_i)}{1+\mathrm{P}(A_i)};$$
$$\mathrm{P}(f=1)=\frac{\exp(1{,}94-0{,}05x_i)}{1+\mathrm{P}(A_i)}; \qquad (22)$$
$$\mathrm{P}(f=2)=\frac{\exp(-0{,}55-0{,}01x_i)}{1+\mathrm{P}(A_i)};$$
$$\mathrm{P}(f=3)=\frac{1}{1+\mathrm{P}(A_i)}.$$

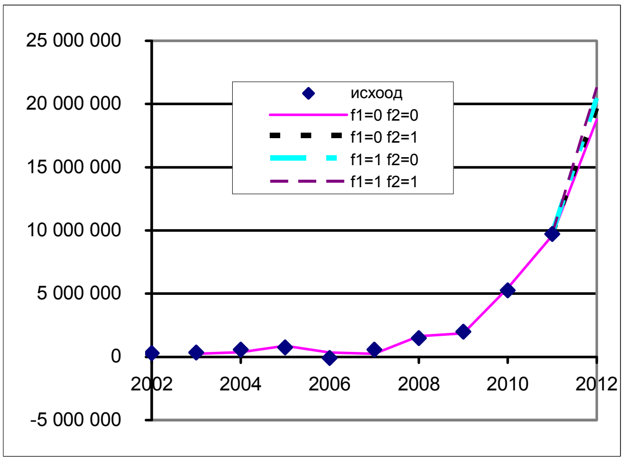

Рис. 2 – Сравнение исходных данных и расчетных значений с двумя фиктивными переменными

Эксперты оценили экономическую обстановку в 2012 году в 68 баллов. Поэтому для оценки вероятностей использовалось $x=68$. В соответствие с моделью (22) были получены следующие данные.

f=0	(f_1=0, f_2=0)	y_{t+1} =18 778 668	P(0)=0,28
f=1	(f_1=0, f_2=1)	y_{t+1} =19 663 826	P(1)=0,06
f=2	(f_1=1, f_2=0)	y_{t+1} =20 451 991	P(2)=0,36
f=3	(f_1=1, f_2=1)	y_{t+1} =21 337 150	P(3)=0,30

Наиболее вероятным вариантом является третий (P=0,36), для которого ожидаемая прибыль предприятия составит 20 452 млн. руб. Однако нельзя исключить первый и четвертый варианты, вероятность которых (0,28 и 0,3) также достаточно большие. В этом случае лучше ориентироваться на математическое ожидание величины прибыли, которое составит

$$M(y_{t+1})=20\ 195\ 158\ \text{тыс.руб.} \qquad (23)$$

Заключение

В данной работе был применен метод, предложенный Давнисом В.В. для оценки возможных значений прибыли ОАО «Казанский вертолетный завод» на 2012 год.

Комбинированная модель позволяет получать прогнозные оценки показателей, представленных динамическим рядом, на одну точку вперед. При необходимости прогноза на более отдаленную перспективу указанная процедура повторяется; при этом полученная на первом шаге прогнозная оценка используется в качестве фактической исходной информации.

Модель позволяет строить многовариантный прогноз. При принятии решений можно руководствоваться не только наиболее вероятным прогнозом, но и использовать наименее благоприятные сценарии развития событий, если этого требуется для оценки худшего варианта.

Литература

1. Первадчук В.П. Математическая модель прогнозирования финансового состояния предприятия / В.П. Первадчук, И.Б. Масенко // Вестник ОГУ. – 2007. - №77. – С. 181-190.
2. Аксянова А.В. Прогнозирование показателей развития социально-экономической сферы региона / А.В. Аксянова, Ю.В. Хайрутдинова // Вестник Казанского технологического университета. – 2011. – № 20. – С. 305-311.
3. Давнис В.В. Прогнозные модели экспертных предпочтений: монография / В.В. Давнис, В.И. Тинякова. – Воронеж: Издательство Воронежского государственного университета, 2005. – 248 с.
4. Валеев Н.Н. Теория и практика эконометрики / Н.Н. Валеев, А.В. Аксянова, Г.А.Гадельшина. – Казань: КГТУ, 2010. – 302 с.: ил.
5. Давнис В.В. Прогноз и адекватный образ будущего / В.В. Давнис, В.И. Тинякова // Вестник ВГУ. – 2005. - №2. – С.183-188.
6. Официальный сайт Казанского вертолетного завода www.kazanhelicopters.ru.

© **Г. А. Гадельшина** - доцент кафедры химической кибернетики КНИТУ; **А. В. Аксянова** – доц. той же кафедры, romanova_rg@mail.ru.