УДК 338.27

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ КОЛИЧЕСТВЕННЫХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ

П.В. БУДАЕВ

Рассмотрена практика применения количественных методов прогнозирования с помощью программы SPSS. Предложен алгоритм построения, оценки и тестирования широко распространенных моделей. Приведены примеры прогнозирования продаж с использованием описанных моделей на производственном предприятии.

В современных условиях любое предприятие, независимо от его географического расположения и рынков, на которых оно работает, сталкивается с проблемой принятия оптимального решения. Под оптимальным в данном случае понимается решение, которое

- принимается на основе полной, объективной и достоверной информации, получаемой из разных источников достаточно длительный период времени;
- соответствует стратегии, маркетинговым и финансовым целям предприятия;
- выбирается в результате правильной оценки возможных рисков.

Как правило, обстоятельства вынуждают действовать в условиях ограниченного времени, большой номенклатуры оказываемых услуг, высоких затрат, связанных с подготовкой и динамично меняющейся ситуацией на рынках. Поэтому целесообразно максимально автоматизировать процесс принятия решений.

Базисом принятия решения в рамках выбранной стратегии предприятия является прогноз продаж. Прогнозируя и оценивая возможный сбыт, предприятие может планировать свою деятельность и оценивать затраты. Понятно, что чем более высокая точность прогнозирования, тем больше шансов для принятия оптимальных решений, а, следовательно, повышения конкурентоспособности и прибыльности предприятия.

Инструментом для принятия оптимального решения служит информационно-аналитическая система предприятия, а ее основой — накопленная информация и методы обработки. В свою очередь, накопленная информация делится на формализованную (статистические данные, отраслевые отчеты, исследования рынка и т.д.), слабо формализованную (новости, новые технологии, перспективные разработки и открытия в отрасли) и неформализованную (знания, интуиция и опыт экспертов). Соответственно разнятся методы обработки этой информации, условно разделяемые на **количественные** (для формализованных данных) и **качественные** (для слабо и неформализованных данных).

Существует достаточное количество работ, в которых детально описаны как количественные [5–8] (наиболее используемые из них — прогнозирование числовых рядов [7, 8]), так и качественные методы [1, 2, 4]. Тем не менее, проблема их практического применения в условиях реально действующего производственного предприятия остается актуальной.

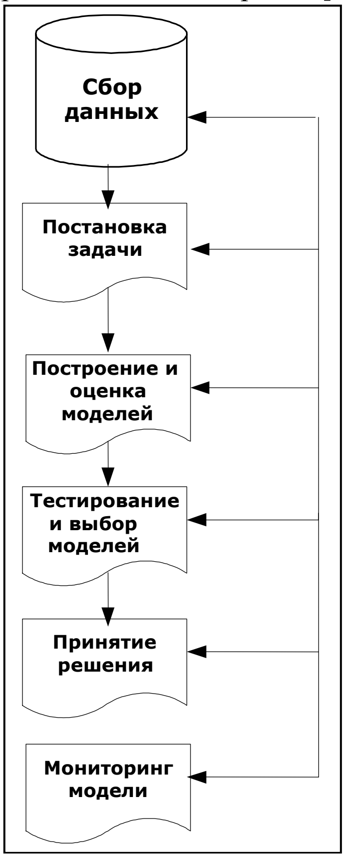

Рис. 1. Структура процесса прогнозирования в виде последовательности формализованных действий

Цель настоящей статьи — систематизация подхода к выбору количественных методов обработки информации, выбор и оценка математических моделей и методов прогнозирования, а также разработка рекомендаций для их практического применения.

На практике процесс прогнозирования представляет собой последовательность действий (рис. 1).

На этапе **сбора данных** используется информационно-аналитическая система предприятия, а ЛПР (лицо, принимающее решение) должно удостовериться в правильности данных и источнике их происхождения.

На этапе **постановки задачи** определяется предмет прогнозирования и выбираются его параметры — горизонт и шаг. Для кратко- и среднесрочных (от одного до пяти лет) прогнозов шаг прогнозирования, как правило, выбирается в один месяц или квартал.

Построение и оценка моделей — ключевой момент прогнозирования. Классифицируются модели в зависимости от горизонта планирования (табл. 1).

Таблица 1. Классификация моделей прогнозирования

Тип	Класс	Горизонт планирования
Модели временных рядов	Сглаживания	Краткосрочный
	Подгонка кривыми	Кратко- и среднесрочный
	Сезонная модель	— // — // —
	Линейная регрессия	— // — // —
Казуальные модели	Парная и множественная регрессии	Кратко-, средне- и долгосрочный

Тестирование и выбор моделей. После проведения оценки моделей необходимо их протестировать. Наилучший и проверенный на практике метод тестирования — испытание модели на заранее известных результатах. Анализ ошибок позволяет корректировать саму модель.

Принятие решения. ЛПР делает выбор в пользу той или иной модели и оценивает возможные последствия своего решения.

Мониторинг модели. Сравнение полученных результатов, оценка ошибки прогнозирования, объяснение причины ее возникновения и поиски возможности использования выбранной модели в дальнейшем.

Рассмотрим использование перечисленных в табл. 1 моделей более подробно на практических примерах.

Модели подгонки кривыми (экстраполяция). Суть модели: найти кривую или группу кривых, с достаточной точностью описывающих историческую информацию [5]. С их помощью найти тренд в исходной информации, т.е. выявить общую тенденцию в изменении данных, на основании которой можно предполагать наличие прогнозируемых данных на найденном тренде. Прогнозирование с использованием этого метода применяется в случае стабильных во времени продаж с ярко выраженным трендом для товаров с неэластичным спросом.

Основные 11 уравнений кривых для прогнозирования приведены в табл. 2.

Таблица 2. Уравнения основных типов кривых

Наименование	**Уравнение**
Линейная	$y = b_1 + b_2 t$
Парабола	$y = b_1 + b_2 t + b_3 t^2$
Логарифмическая	$y = b_1 + b_2 \ln(t)$
Многочлен третьей степени	$y = b_1 + b_2 t + b_3 t^2 + b_4 t^3$
Экспоненциальная	$y = b_1 \exp(b_2 t)$
График степенной функции	$y = b_1 (t^{b_2})$
Гипербола	$y = b_1 + \frac{b_2}{t}$
S-кривая	$y = \exp(b_1 + \frac{b_2}{t})$
Обратная	$y = b_0 + \frac{b_1}{t}$
Составная	$y = b_0 * b_1^t$
Логистическая	$y = \frac{1}{1/b_1 + b_2 (b_3{}^t)}$

Для прогнозирования методом подгонки кривых рекомендуется использовать специализированное программное обеспечение, например, SPSS (12-я версия и выше русифицированы). В процессе автоматизированного построения кривых с помощью этого пакета применен принцип минимизации среднеквадратичной ошибки MSE (mean squared error) [5].

$$MSE = \frac{\sum e_i^2}{n},$$

где $e_i = Y_i - \hat{Y}_i$ — разность между имеющимся значением и значением, находящимся на кривой подгонки (так называемый остаток).

По имеющимся значениям продаж SPSS достаточно легко строит все перечисленные кривые, а также рассчитывает основной критерий оценки выбора — коэффициент детерминации

$$R^2 = 1 - \frac{MSE}{Var(Y)}, \text{ где } Var(Y) = \sum (Y_i - \overline{Y})^2 .$$

Из определения видно, что равенство $R^2 = 1$ возможно в том случае, если процесс поведения данных в точности описывается моделью. Т.е., чем ближе значение R^2 к единице, тем более точно кривая описывает поведение данных.

На практике для оценки выбора той или иной модели применяют также другой показатель — MAPE (mean absolute persent error) [5].

Имея данные продаж ($V_{\text{пр}}$) некой номенклатурной единицы, производимой предприятием, необходимо сделать прогноз продаж на следующий квартал — 21, 22 и 23-й месяцы. Для этого по имеющимся исходным данным (табл. 3) с помощью программного пакета SPSS построены 11 кривых подгонки, рассчитаны значения R^2 для каждой модели и получены прогнозы значения объема продаж на следующий квартал. Задача состоит в выборе модели, прогноз которой наиболее точен.

Остановимся на двух моделях с максимальным значением R^2 — кубической и квадратичной. Рассчитанные SPSS значения этого параметра равны 0,871 у кубической модели и 0,860 — у квадратичной.

Теперь нам понадобится сравнить значения MAPE этих моделей.

Таблица 3. Сравнение квадратичной и кубической моделей

Месяц	$V_{\text{пр}}$, Y	**Квадратичная модель**			**Кубическая модель**		
		Y	e	MAPE, %	Y	e	MAPE,%
1	2698	2200,44	497,56	22,61	2736,14	–38,13789	–1,39
2	2962	2690,12	271,88	10,11	2887,48	74,52076	2,58
3	2720	3171,68	–451,68	–14,24	3124,69	–404,68961	–12,95
4	1948	3645,13	–1697,13	–46,56	3436,71	–1488,7124	–43,32
5	5667	4110,47	1556,53	37,87	3812,49	1854,50901	48,64
6	4619	4567,69	51,31	1,12	4240,97	378,03124	8,91
7	5569	5016,8	552,2	11,01	4711,09	857,91091	18,21
8	5797	5457,81	339,19	6,21	5211,8	585,20463	11,23
9	6023	5890,69	132,31	2,25	5732,03	290,96902	5,08
10	4348	6315,47	–1967,47	–31,15	6260,74	–1912,7393	–30,55
11	6261	6732,13	–471,13	–7,00	6786,86	–525,86371	–7,75
12	6701	7140,69	–439,69	–6,16	7299,35	–598,34761	–8,20
13	7690	7541,12	148,88	1,97	7787,13	–97,13436	–1,25
14	7247	7933,45	–686,45	–8,65	8239,17	–992,16737	–12,04
15	8994	8317,67	676,33	8,13	8644,39	349,61	4,04
16	11041	8693,77	2347,23	27,00	8991,75	2049,25436	22,79
17	9521	9061,76	459,24	5,07	9270,18	250,82232	2,71
18	9740	9421,64	318,36	**3,38**	9468,63	271,3705	**2,87**
19	9212	9773,41	–561,41	**–5,74**	9576,04	–364,04448	**–3,80**
20	9041	10117,06	–1076,06	**–10,64**	9581,37	–540,36601	**–5,64**

Особый интерес представляет поведение модели в последние месяцы. Очевидно, что наши данные более точно описывает кубическая модель. Тем не менее, для окончательного выбора той или иной модели необходимо провести ее тестирование. Предпочтительней с практической точки зрения метод тестирования **ex-post-прогнозирование** [5]. Он применяется следующим образом.

1. Разбиваем исходные данные на две группы (первая — 85% данных, вторая — 15).

2. Производим имитацию процесса прогнозирования на один шаг, чтобы провести оценку работоспособности модели при сопоставлении с исходными данными.

3. Определяем разницу между смоделированным и прогнозным значениями (ошибкой прогнозирования).

4. Оцениваем ошибку прогнозирования.

5. Повторяем пп. 2–4 до последнего значения исходных данных.

6. Делаем выводы и принимаем решение в пользу одной из моделей, среднее значение MAPE которой минимально.

Рассмотрим применение этого метода на предыдущем примере, проведя оценку результатов ex-post-прогнозирования обеих моделей для значений 18, 19 и 20-го месяцев. Для этого по значениям 17-ти месяцев построим прогнозное значение на 18-й месяц, по значениям 18-ти месяцев спрогнозируем 19-й месяц и т.д. (табл. 4).

Таблица 4. Результаты оценки моделей подгонки кривыми

Исторические данные		**Квадратичная модель**			**Кубическая модель**		
Месяц	$V_{пр}$, Y	Y	e	**MAPE, %**	Y	e	**MAPE,%**
18	9740	10768,75	–1028,75	10,56	11556,54	–1816,54	18,65
19	9212	10957,82	–1745,82	18,95	11170,05	–1958,05	21,26
20	9041	10751,15	–1710,15	18,92	10275,5	–1234,5	13,65
				16,14			**17,85**

Из полученных результатов можно сделать вывод: для прогнозирования будущих значений целесообразно выбрать квадратичную модель, среднее значение MAPE которой за последние три месяца меньше, чем у кубической модели.

Модели сглаживания применяются для построения краткосрочных прогнозов для временных рядов, не имеющих тренда. Здесь игнорируются изменения, вызванные случайными факторами, а последним исходным значениям придается больший вес по сравнению с ранними значениями. При этом подразумевается, что все основные факторы, влияющие на исследуемый показатель, продолжают свое действие на горизонте прогнозирования, и определенная тенденция останется в силе на ближайший период. Такие модели целесообразно применять для зрелых рынков со сложившимся балансом сил и низкой вероятностью качественных изменений, которые могли бы повлиять на ситуацию.

Модель скользящего среднего (moving average model) [3] — наиболее простая и часто применяемая модель.

$$\hat{Y}_t = \frac{(Y_{t-1} + Y_{t-2} + Y_{t-3} + ... + Y_{t-k})}{k},$$

где $\hat{Y}$ — прогнозируемое значение; t — шаг прогнозирования; k — порядок скользящего среднего; Y_t — истинные значения.

Термин «скользящее среднее» используется потому, что новое среднее значение вычисляется и служит в качестве прогноза каждый раз, когда мы его производим. При прогнозировании следующего значения метод скользящего среднего придает равные веса каждому из последних значений. Наиболее удобным для практического применения этого метода является MS EXEL. Используя функцию пакета СРЗНАЧ (Ячейка NN:Ячейка MM), где NN и MM — ячейки с историческими данными, можно получить среднее значение для имеющихся данных. Важно отметить, что минимальное значение используемых данных — три (меньшее количество не отражает тенденцию, большое — ее сглаживает). Метод является очень быстрым, но очень неточным способом получения прогноза.

Модель экспоненциально взвешенного скользящего среднего (exponentially weighted moving average (EWMA) [3] — разновидность метода скользящего среднего. Суть метода:

- каждое новое значение определяется совокупностью всех предыдущих значений;
- влияние предыдущих данных ослабевает в геометрической прогрессии

$$\hat{Y}_t = Y_{t-1} + \alpha Y_{t-2} + \alpha^2 Y_{t-3} + ... + \alpha^{n-1} Y_{t-k} \text{ при } 0 \le \alpha \le 1,$$

а в случае, если $n \to \infty$,

$$Y_t = (1-\alpha)Y_{t-1} + \alpha Y_{t-1}.$$

Основной задачей, существенно влияющей на эффективность использования данного метода, является выбор так называемой константы сглаживания α в диапазоне от нуля до единицы. Высокие значения α придают больший вес последним наблюдениям и меньший — более давним. Число $(1-\alpha)$ называется константой сглаживания. В случае использования пакета SPSS значение α предлагается выбрать по сетке с указанным шагом, а наиболее подходящее значение α определяется на этапе тестирования и выбора модели через ex-post-прогнозирование.

Двойное экспоненциальное сглаживание (double exponential smoothing model) [3] получается посредством двойного применения модели экспоненциального сглаживания. Т. е. сначала модель экспоненциального сглаживания применяется к исходным данным, а затем к смоделированным значениям, полученным на первом этапе. При повторном моделировании возможно изменение константы $(1-\alpha)$. Перечисленные модели сглаживания применимы к данным, не имеющим тренда.

Модель Холта [5] применяют, если необходимо учитывать тренд. Основная идея метода: добавление значения тренда к прогнозируемому моделью экспоненциального сглаживания значению

$$Y_t = (1-\alpha)Y_{t-1} + \alpha Y_{t-1} + T_{t-1}, \quad (1)$$

где

$$T_t = (1-\beta)(\hat{Y}_t - \hat{Y}_{t-1}) + \beta T_{t-1} \text{ при } 0 \le \beta \le 1. \quad (2)$$

Чтобы найти прогнозное значение $\hat{Y}_t$ с использованием модели Холта, необходимо выполнить следующие процедуры.

1. Определить величину параметров α и β методом минимизации MAPE.

2. Найти начальные значения $\hat{Y}_1$ и T_1 как $\hat{Y}_1 = \overline{Y}$ или $\hat{Y}_1 = Y_1$, а

$$T_1 = \frac{(Y_2 - Y_1) + (Y_4 - Y_3)}{2}.$$

3. Попеременно найти $\hat{Y}_t$ из уравнения (1) и T_t из уравнения (2).

4. Сделать прогоноз на m шагов вперед, предполагая, что последнее значение тренда останется неизменным для будущих прогнозов.

$$P_{n+m} = \hat{Y}_n + mT_n.$$

Модель Брауна [5] учитывает тренд двойного экспоненциального сглаживания и дает достаточно точный краткосрочный прогноз. К значению модели добавляется простое экспоненциальное сглаживание разницы в значениях между простым и двойным сглаживанием, умноженное на коэффициент, обратно пропорциональный α.

$$\hat{Y}_{t+1} = \hat{Y}'_i + \frac{\hat{Y}'_i - \hat{Y}''_i}{\alpha}.$$

Прогноз на m шагов вперед определяется формулой

$$P_{n+m} = \hat{Y}_n + (\hat{Y}'_n - \hat{Y}''_n)(1 + \frac{m(1-\alpha)}{\alpha}.$$

Значения α при этом определяются минимизацией MAPE.

В качестве примера применения моделей Холта и Брауна рассмотрим поведение продаж другой номенклатурной единицы (рис. 2).

Попытка применить к имеющимся данным метод подгонки кривыми дал очень невысокие значения R^2 — от 0,004 до 0,12, что, очевидно, означает отсутствие тренда.

Применим модели EWMA, Холта и Брауна к имеющимся значениям. Как всегда, наибольший интерес представляет оценка поведения этих моделей методом ex-post-прогнозирования (табл. 5).

Очевидно, что в данном случае следует выбрать модель Холта, как наиболее точно описывающую наши данные. Видно, что среднее значение MAPE этой модели минимально.

Перечисленные модели хорошо работают для данных, не имеющих тренда для получения краткосрочного прогноза. Если значения трендов нам не важны, применяются модели скользящего среднего, экспоненциально взвешенного скользящего среднего, двойного экспоненциального сглаживания. Для учета тренда, определяемого последними значениями, применяются модели Холта и Брауна.

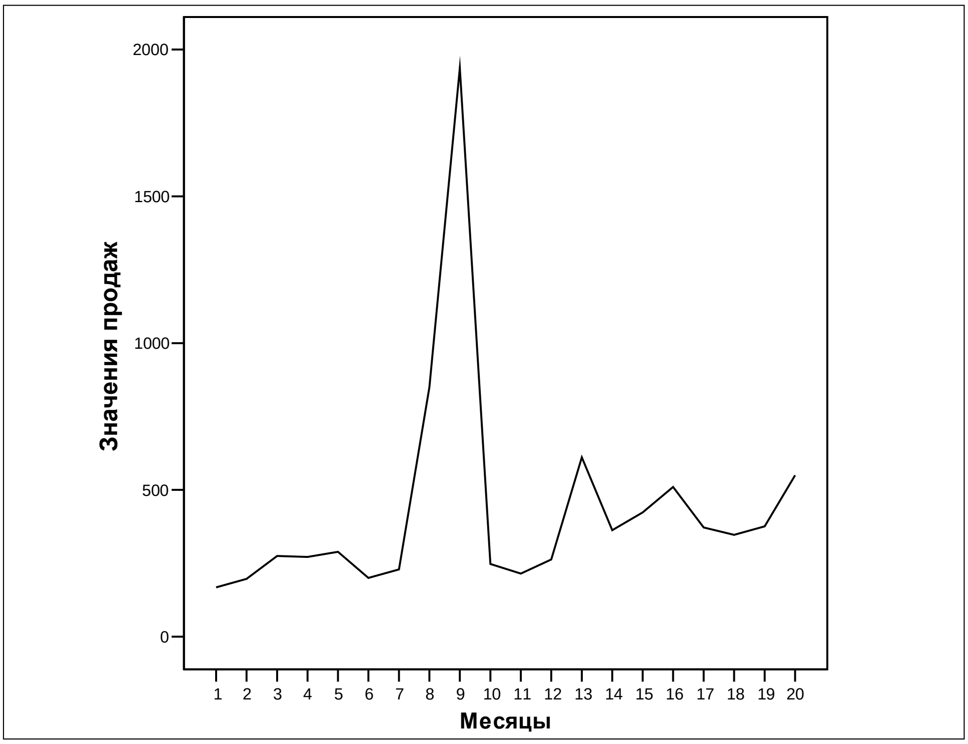

Рис. 2. График продаж номенклатурний единицы

Таблица 5. Результаты оценки моделей сглаживания

Истори-ческие данные		**EWMA**			**Холта**			**Брауна**		
Ме-сяц	$V_{пр}$, Y	Y	e	**MAPE, %**	Y	e	**MAPE, %**	Y	e	**MAPE, %**
18	347	372,18	–25,18	7,26	423,39	–76,39	–22,01	379,49	–32,49	–9,36
19	376	373,34	2,66	0,71	359,49	16,51	4,39	369,04	6,96	1,85
20	550	448,97	101,03	18,37	389,04	160,96	29,27	436,14	113,86	20,70
				8,78			**3,88**			**4,40**

Сезонные модели служат для получения прогнозов при работе с услугами, продажи которых носят сезонный характер. Прогноз в таком случае базируется на факторах тренда и сезонности (повышение или понижение тренда с одинаковыми промежутками времени).

Прогнозное значение $\hat{Y}_t$ определяется как

$$\hat{Y}_t = d_t S_t,$$

где S_t — коэффициент сезонности; d_t — десезонализированное значение.

В случае отсутствия тренда процесс нахождения сезонных коэффициентов и прогнозирования с их помощью основавается на предположении, что исходные данные совершают колебания вокруг своего среднего значения, и в будущем этот процесс не изменится.

Алгоритм прогнозирования в таком случае достаточно прост.

1. Вычисляется среднее значение за исторический период.
2. Вычисляются коэффициенты сезонности.
3. Прогнозируется среднее значение на горизонт прогнозирования.
4. Корректируется среднее значение на коэффициент сезонности.

Покажем это на примере номенклатурной единицы, значения данных продаж которой приведены в табл. 6, а график — на рис. 3. Видно, что среднее значение продаж в 2006 и 2007 гг. практически не изменилось. Применение коэффициентов сезонности 2006 г. в качестве прогнозных на соответствующий интервал 2007 г. дало неплохие результаты (MAPE = 8,86%). Сделав предположение о среднем значении продаж на 2008 г. в количестве 550 штук и применив соответствующие коэффициенты сезонности 2007 г., получим прогноз продаж на первый квартал 2008 г. по месяцам.

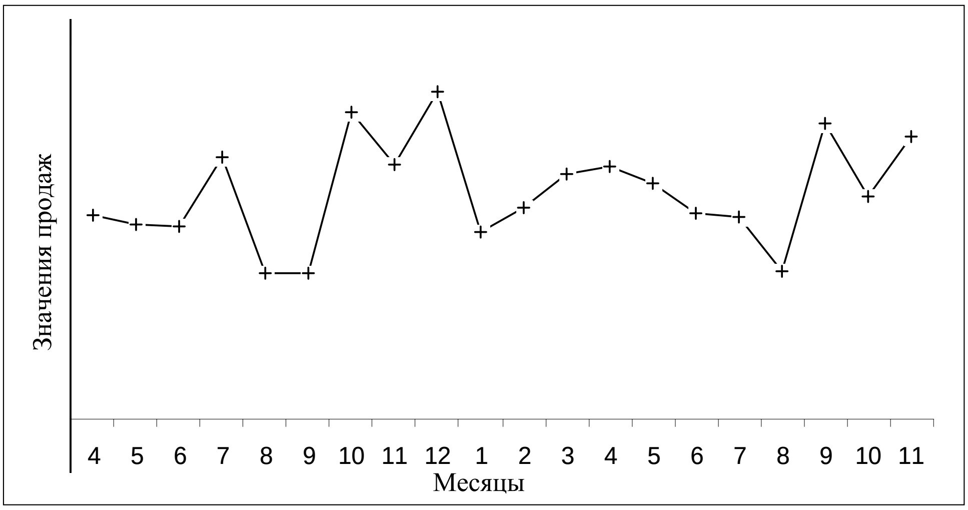

Рис. 3. График продаж номенклатурной единицы

Таблица 6. Результаты прогнозирования с помощью сезонной модели без тренда

Год	Месяц	$V_{пр}$, Y	S	Прогнозное значение	e	MAPE,%
2007	4	491,0	0,90			
	5	467,0	0,86			
	6	464,0	0,85			
	7	628,0	1,16			
	8	352,0	0,65			
	9	353,0	0,65			
	10	739,0	1,36			
	11	611,0	1,12			
	12	786,0	1,45			
	Среднее	**543,4**				

Окончание табл. 6

2008	1	448,0	0,82			
	2	510,0	0,94			
	3	588,0	1,08			
	4	609,0	1,12	491,00	118,00	19,38
	5	569,0	1,04	404,80	164,20	28,86
	6	497,0	0,91	460,82	36,18	7,28
	8	488,0	0,90	531,30	–43,30	–8,87
	9	356,0	0,65	550,27	–194,27	–54,57
	10	710,0	1,30	514,13	195,87	27,59
	11	535,0	0,98	449,07	85,93	16,06
	12	680,0	1,25	440,94	239,06	35,16
	Среднее	**544,5**				**8,86**
2009	1			**452**		
	2			**515**		
	3			**594**		

Более сложным методом определяются коэффициенты сезонности при наличии тренда (рис. 4).

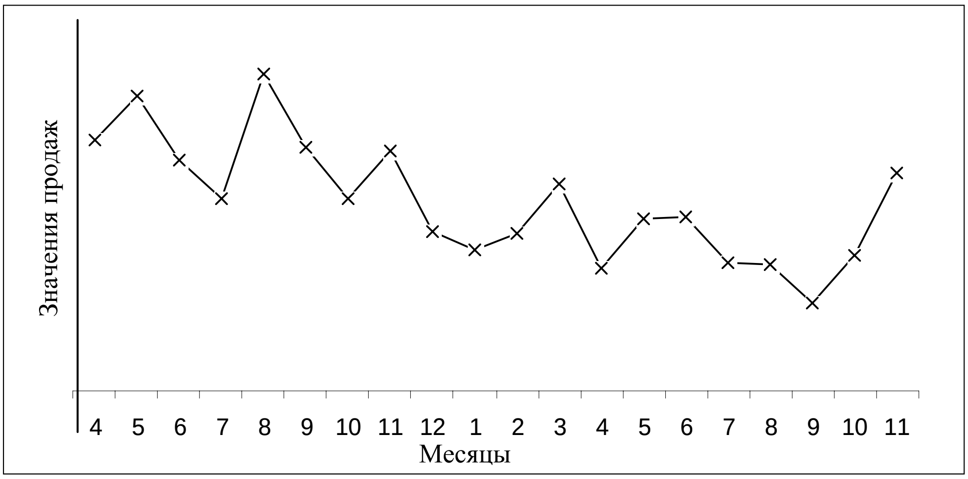

Рис. 4. График продаж номенклатурной единицы

Алгоритм нахождения сезонных коэффициентов и прогноза следующий.

1. Находим скользящее среднее d_t' первого порядка для исходного ряда Y_t.

2. Находим центрированное скользящее среднее d_t'' второго порядка для ряда d_t'.

3. Находим неусредненные значения коэффициентов сезонности $S' = \dfrac{Y_t}{d_t''}$.

4. Находим окончательные десезонализированные значения.

5. Строим прогнозную модель по десезонализированным значениям.

6. Корректируем прогнозные значения на соответствующие коэффициенты сезонности.

Результаты вычислений приведены в табл. 7.

Таблица 7. Результаты прогнозирования с помощью сезонной модели с трендом

Год	**Месяц**	$V_{пр}$, Y	**Скользящее среднее** d'	**Скользящее среднее** d''	S	**Десезонализированное значение** d
2007	4	478,0				
	5	562,0	493,3			
	6	440,0	461,8	477,5	1,09	405,41
	7	367,0	493,3	477,5	1,30	282,07
	8	604,0	469,3	481,3	0,80	758,06
	9	466,0	451,5	460,4	0,99	471,69
	10	369,0	474,8	463,1	1,26	294,00
	11	460,0	399,5	437,1	0,95	484,07
	12	303,0	350,3	374,9	1,24	244,91
2008	1	269,0	333,0	341,6	1,27	211,81
	2	300,0	317,0	325,0	1,08	276,92
	3	396,0	299,5	308,3	0,78	508,73
	4	233,0	314,5	307,0	1,32	176,84
	5	329,0	323,0	318,8	0,97	339,58
	6	334,0	285,5	304,3	0,91	366,66
	7	246,0	287,3	286,4	1,16	211,32
	8	240,0	246,8	267,0	1,11	215,73
	9	167,0	227,8	237,3	1,42	117,55
	10	258,0	270,8	249,3	0,97	267,06
	11	418,0	310,0			
	12	397,0				

Для получения прогноза необходимо на основе десезонализированных значений построить прогнозную модель, например, с помощью одной из кривой подгонки, после чего, выбрав коэффициенты сезонности, скорректировать прогнозные значения на соответствующий коэффициент.

Модель линейной регрессии [4] применяется в том случае, когда стабильность прогнозируемого значения нам гораздо важнее точности, например, при прогнозировании финансовых показателей компании. Если использование предыдущих моделей давало единственное значение прогнозируемой величины, то теперь результатом прогнозирования является числовой интервал. Прогноз продаж в таком случае выглядит так: «Объем продаж с вероятностью 95% будет находиться в диапазоне 3100–3500 шекелей». Единственная сложность заключается в том, чтобы понять, действительно ли наша линейная модель описывает поведение значений. Для этого проводят несколько тестов модели. Наиболее показательным, на наш взгляд, является тест рекурсивной оценки ошибок регрессии.

Модель линейной регрессии выглядит так:

$$Y_i = \beta_i + \beta_2 X_i + \varepsilon_i \,,$$

где β_i, β_2 — константы модели; $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n$ — нормальные случайные величины с одним и тем же математическим ожиданием, равным нулю и стандартным отклонением σ; X_i — константа.

На практике предполагается, что $i = 1, 2, \ldots, n$ соответствуют моментам времени, имеющим одинаковый шаг (неделя, месяц, квартал).

Рассмотрим применение этой модели на практическом примере, взяв за основу данные табл. 2.

Применив для построения уравнения линейной регрессии программный пакет SPSS, мы можем достаточно легко получить прогноз на ближайшие три месяца. Однако наибольший интерес в данном случае представляет собой не точность самого прогноза, а адекватность модели, т.е., насколько параметры модели стабильны и соответствуют реальным значениям за последние шесть месяцев. Для этого необходимо протестировать ошибку регрессии *e* на рекурсивность, проверив ее нахождение в течение интересующего периода с вероятностью 95% в доверительном интервале, последовательно построив модели. Так, по значениям продаж за предыдущие 14 месяцев строится прогноз на 15-й месяц, по результатам 15 месяцев — на 16-й и т.д. (табл. 8, рис. 5).

Таблица 8. Результаты прогнозирования с помощью модели линейной регрессии

Месяц	$V_{пр}, Y$	**Y лин. регр.**	***e* лин. регр.**	**Нижняя граница** 95%	**Верхняя граница** 95%	**Длина интервала**
15 (1)	8994	8264,94	729,06	7651,27	8878,61	613,67
16 (2)	11041	8681,60	2359,40	8013,03	9350,16	668,56
17 (3)	9521	9098,27	422,74	8369,26	9827,27	729,01
18 (4)	9740	9514,93	225,07	8721,20	10308,65	793,72
19 (5)	9212	9931,59	–719,59	9069,84	10793,35	861,76
20 (6)	9041	10348,26	–1307,26	9415,88	11280,63	932,38

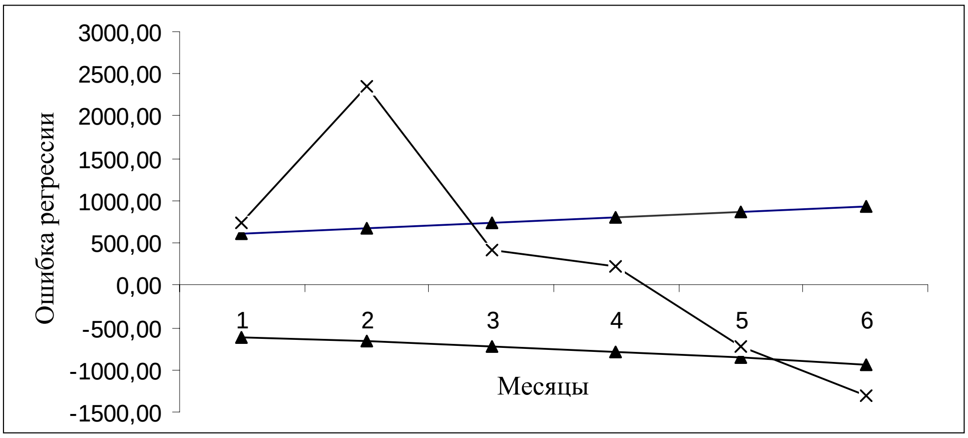

Рис. 5. График нахождения ошибки регрессии в доверительном интервале

Длина доверительного интервала со временем увеличивается, ухудшая тем самым точность прогнозирования. Главное заключается в том, что ошибка прогнозирования из шести наблюдений только трижды попала в доверительный интервал. Следовательно, предложенная модель линейной регрессии для этих данных не является адекватной и ее не следует использовать в таком случае.

Модель множественной регрессии. В предыдущей модели мы использовали два регрессора — время и значения продаж. В модели множественной регрессии таких регрессоров может быть несколько — уровень доходов потребителей, цены на продукты конкурентов, расходы на рекламу и др. Может быть построена регрессионная модель, уравнение которой имеет вид

$$Y_i = \beta_i + \beta_2 X_i^{(2)} + \beta_k X_i^{(k)} + \varepsilon_i ,$$

где β_i, β_2 — константы модели; $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n$ — нормальные случайные величины с одним и тем же математическим ожиданием, равным нулю и стандартным отклонением σ; $i = 1, 2, \ldots, n$.

При $k \geq 2$ модель называется моделью множественной регрессии.

Все регрессоры можно разделить на контролируемые (например, расходы на рекламу) и неконтролируемые (количество рабочих дней в месяце).

С точки зрения практического применения наиболее сложным является нахождение оптимального количества регрессоров — переменных X, описывающих изменения прогнозируемых значений Y. Для этого сначала определяется набор таких регрессоров, а затем оставляются те, влияние которых на прогнозируемые значения существенно.

Рассмотрим применение этого метода на практическом примере, определив в качестве контролируемых регрессоров расходы на рекламу и неконтролируемых — количество рабочих дней в месяце. В свою очередь, понимая, что расходы на рекламу и ее воздействие на потребителя разнесены во времени, добавим в качестве контролируемых регрессоров расходы на рекламу с запаздыванием в один и два месяца.

На следующем этапе необходимо избавиться от лишних регрессоров. С помощью программы SPSS мы можем, поочередно отбрасывая регрессоры, протестировать полученные модели, используя в качестве параметра выбора скорректированный коэффициент детерминации

$$\overline{R}_k^2 = 1 - \frac{(n-1)(1-R_k^2)}{n-k} ,$$

где $R_k^2 = \dfrac{(\hat{Y}_i - \overline{Y})^2}{(Y_i - \overline{Y})^2}$,

и стандартную ошибку регрессии

$$s = \sqrt{\frac{{e_1}^2 + {e_2}^2 + \ldots + {e_n}^2}{n-k}} .$$

Результаты анализа наших моделей приведены в табл. 9.

a Предикторы: (константа) ЗРеклама1, Реклама, ЗРеклама2, Рабдни, Месяц

b Предикторы: (константа) ЗРеклама1, ЗРеклама2, Рабдни, Месяц

c **Предикторы: (константа) 3Реклама1, Рабдни, Месяц**

d Предикторы: (константа) Рабдни, Месяц

e Зависимая переменная: Продажи

Таблица 9. Анализ моделей множественной регрессии

Модель	R	R^2	Скорректированный R^2	Стандартная ошибка оценки
1	0,838(a)	0,702	0,454	1625,42105
2	0,838(b)	0,702	0,532	1505,14004
3	**0,835(c)**	**0,698**	**0,585**	**1417,67654**
4	0,783(d)	0,613	0,526	1513,73175

Очевидно, что оптимальной моделью в нашем случае является модель с регрессорами «Количество рабочих дней в месяце» и «Расходы на рекламу» с запаздыванием в один месяц. Прогноз продаж с использованием такой модели при заданных затратах на рекламу приведен в табл. 10.

Таблица 10. Результаты прогнозирования с помощью модели множественной регрессии

Месяц	Объем продаж, шт.	Количество рабочих дней в месяце	Расходы на рекламу с запаздыванием в месяц	Прогноз продаж, шт.
1	5966	21	0	7928,86
2	8547	19	15394	7184,21
3	7927	21	8571,81	8853,18
4	11119	22	7320,17	9918,75
5	11040	22	12469,79	10380,92
6	9349	20	8445,71	9636,26
7	11543	23	16982,95	11908,65
8	11683	22	4174,02	11767,40
9	12500	19	8641,05	10419,34
10	8845	17	1689,12	9674,69
11	13941	21	0	12550,48
12	10172	20	0	12409,24
13		19	0	12267,99
14		18	0	12126,75
15		20	21000	13795,73
16		22	2000	15464,70

После того как прогноз подготовлен и представлен руководству организации, наступает этап мониторинга и сопоставления полученных прогнозов с реальными данными, что позволяет своевременно выявлять значительные отклонения в ходе развития событий. Если они могут оказать принципиальное влияние на дальнейший ход событий в принятии важных стратегических решений, то прогноз должен быть подвергнут корректировке. В итоге возникает вопрос оценки качества прогноза апостериорно. На практике в качестве такого критерия оценки точности прогноза применяют среднеквадратичное отклонение истинного значения от прогнозируемого. Величину этого параметра каждое предприятие определяет самостоятельно.

В заключение отметим, что список перечисленных моделей (подгонки кривыми, скользящего среднего, экспоненциально взвешенного скользящего среднего, двойного экспоненциального сглаживания, Холта, Брауна, сезонные, линейной и множественной регрессии) не является полным. Приведены самые распространенные и широко применяемые модели. Обычно, чем проще модель, тем шире она применяется в прогнозировании.

Такой подход объясняется следующими причинами.

- Использование более сложных моделей не всегда приводит к повышению точности прогнозов. Имея большой опыт, многие события можно «прочувствовать», но практически невозможно просчитать.
- Чем сложнее модель, тем больше времени требуется на подготовку данных, расчеты, анализ, численные эксперименты. Следовательно, многократно возрастают затраты предприятия. Как результат — чем больше ассортимент, тем проще используемые методы прогнозирования.
- Конкурентная среда, изменяющиеся рынки, проблемы роста компаний и прочие условия при прогнозировании не позволяют рассчитывать на репрезентативные выборки исторических данных. При этом подавляющее большинство моделей прогнозирования так или иначе использует именно эти данные.

Тем не менее, грамотное применение научных методов прогнозирования, умение пользоваться математическим и статистическим аппаратами и прикладным программным обеспечением повышают точность прогнозирования, что в итоге приводит к увеличению конкурентного преимущества предприятия.

ЛИТЕРАТУРА

1. *Беляевский И.К.* Маркетинговое исследование: информация, анализ, прогноз. — М.: Финансы и статистика, 2001. — 320 с.
2. *Круглов В.В., Борисов В.В.* Принятие решений на основе нечетких моделей: примеры использования. — М.: Горячая линия «Телеком», 2001. — 458 с.
3. *Конрад Карлберг.* Бизнес-анализ с помощью Microsoft Excel. — М.: Вильямс, 2006. — 460 с.
4. *Згуровский М.З., Панкратова Н.Д.* Системный анализ. — Киев: Наук. думка, 2005. —744 с.
5. *Слуцкин Л.* Курс MBA по прогнозированию в бизнесе. — М.: Альпина Бизнес Букс, 2006. — 276 с.
6. *Элвин С. Бернс, Рональд Ф. Буш.* Основы маркетинговых исследований с использованием MS EXEL. — М.: Вильямс, 2006. — 694 с.
7. *Бидюк П.И., Половцев О.В.* Анализ и моделирование экономических процессов переходного периода. — Киев: НТУУ «КПИ», 1999. — 230 с.
8. *Бідюк П.І.* Системний підхід до прогнозування на основі моделей часових рядів // Системні дослідження та інформаційні технології. — 2003. — № 3. — С. 88–110.

Поступила 08.12.2008