



ФАКТОРНЫЙ АНАЛИЗ СТАТИСТИК

КОННИКОВ Е.А.



ФАКТОРНЫЙ АНАЛИЗ СТАТИСТИК

Если значения прогнозируемого параметра зависят не от времени, а от каких-либо других факторов, то используется факторный статистический анализ. Обычно для этого с помощью ПЭВМ по известной статистике подбирается аппроксимирующая функция одной или многих переменных, которая и служит моделью для выработки прогноза. Рассмотрим эту процедуру на примере.

ПРИМЕР

Предприниматель реализует мороженое у станции метро Дыбенко. Он должен сделать заказ на следующую неделю с разбивкой по дням. Каждое утро заказанное количество товара завозится на его точки реализации. При неправильном заказе (прогнозе) в конце дня мороженого может не хватить – тогда имеет место упущенная выгода, либо часть его останется нереализованной; в таком случае возникнут проблемы с его хранением. Требуется выявить факторы, определяющие объем продаж, собрать статистику продаж и значений этих факторов, далее – разработать прогноз продаж мороженого на следующую неделю. Предполагается, что дело происходит в разгар лета.

ПРИМЕР

Среди факторов, влияющих на объем продаж мороженого в это время, отобраны два наиболее существенных: **температура воздуха** и **день недели**. Отметим, что второй фактор имеет логический характер, что создает дополнительные трудности решения. Собранные за три недели статистика представлена в таблице ниже. Будем считать, что к моменту оценки объема продаж известен прогноз погоды (температуры воздуха) на следующую неделю.

День недели	Первая неделя		Вторая неделя		Третья неделя	
	Продажа, кг	Температура воздуха, °С	Продажа, кг	Температура воздуха, °С	Продажа, кг	Температура воздуха, °С
ПН	32,5	8	84,5	28	65	21
ВТ	35	10	84	26	77	23
СР	25	11	32,5	14	55	29
ЧТ	33,7	16	31,5	12	58,5	25
ПТ	95	17	66,5	15	114	27
СБ	216	22	171	18	216	24
ВС	277,5	30	166,5	20	259	19
ИТОГО:	714,7		636,5		844,5	

ПРИМЕР

I. Классический метод решения

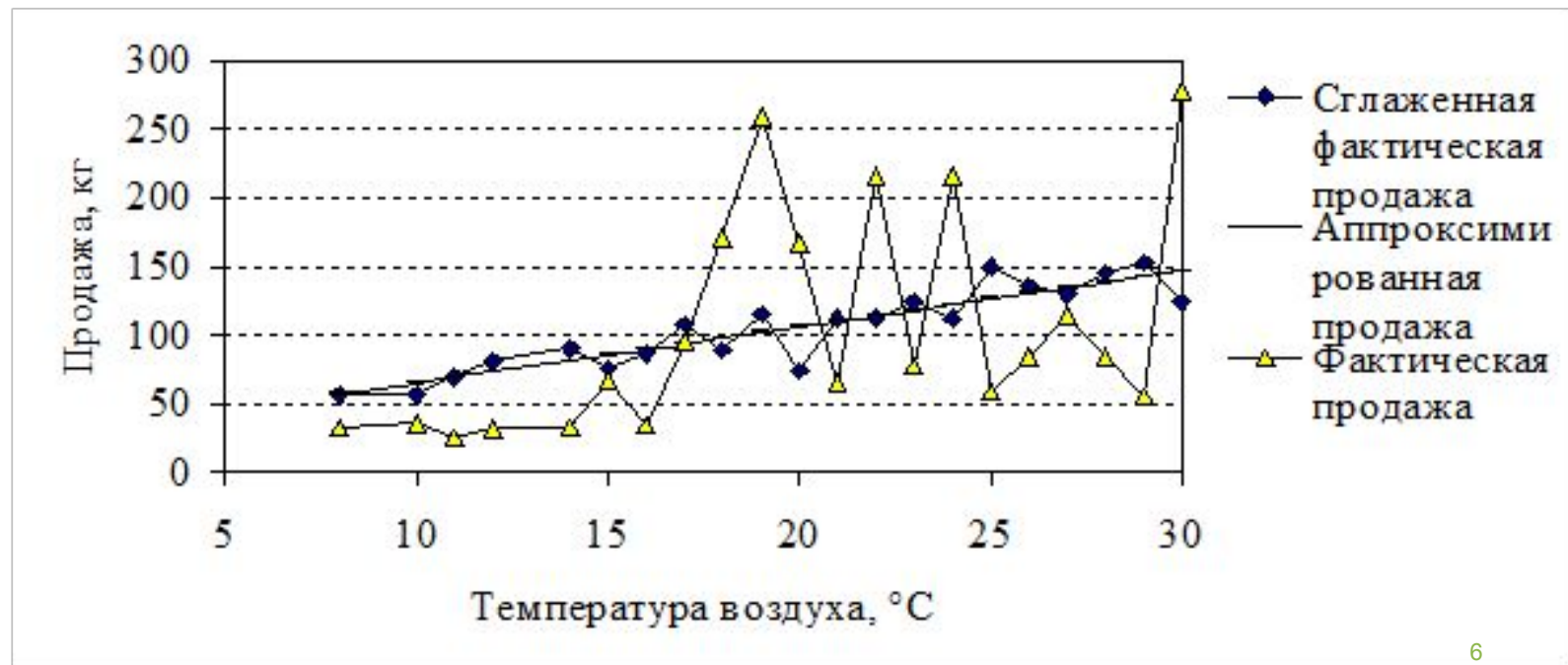
Для устранения влияния на температурную функцию продаж логической переменной – *дня недели* - рассчитаем коэффициенты приведения для каждого дня недели к среднедневной продаже (таблица ниже).

Среднее на день недели, кг	Коэффициент приведения
60,66	$60,66/104,55 = 0,580$
65,33	0,624
37,50	0,358
41,23	0,394
91,83	0,878
201,00	1,922
234,33	2,241
104,55	Среднее

ПРИМЕР

I. Классический метод решения

Затем с помощью этих коэффициентов пересчитаем исходные данные о продажах (получим приведенные фактические продажи, показанные в таблице и на рисунке).



Температура воздуха, °С	Фактическая продажа, кг	Приведенная фактическая продажа, кг	Трендовое значение продажи, кг
8	32,5	56	56,560413
10	35	56,4	64,760667
11	25	69,4	68,860793
12	31,5	80,8	72,960920
14	32,5	90,3	81,161174
15	66,5	75,6	85,261300
16	33,7	86,4	89,361427
17	95	108	93,461554
18	171	89,1	97,561681
19	259	115,6	101,66180
20	166,5	74,3	105,76193
21	65	112,1	109,86206
22	216	112,5	113,96218
23	77	124,2	118,06231
24	216	112,5	122,16244
25	58,5	150	126,26256
26	84	135,5	130,36269
27	114	129,5	134,46282
28	84,5	145,7	138,56294
29	55	152,8	142,66307
30	277,5	123,8	146,76320
Коэффициент корреляции	0,462	0,905	
	Значение коэффициента а		4,10
	Значение коэффициента в		23,76

ПРИМЕР

I. Классический метод решения

Аппроксимация этой зависимости прямой, описываемой уравнением $Q = 4,1*t + 23,76$, дает очень хорошие результаты (коэффициент корреляции **0,9**). В таблице даны также результаты расчетов продаж на основе полученной трендовой линейной зависимости. С использованием этой же модели можно спрогнозировать приведенные продажи на следующую неделю, а затем с помощью коэффициентов приведения пересчитать их в индивидуальные прогнозы на каждый день недели (таблица ниже).

День недели	Прогноз температуры воздуха, °С	Приведенный прогноз продажи, кг	Прогноз продажи, кг
ПН	14	81,16	47,0
ВТ	11	68,86	43,0
СР	9	60,66	21,7
ЧТ	7	52,46	20,7
ПТ	8	56,56	49,7
СБ	13	77,06	148,1
ВС	15	85,26	191,1
		ИТОГО:	521,3

ПРИМЕР

I. Классический метод решения

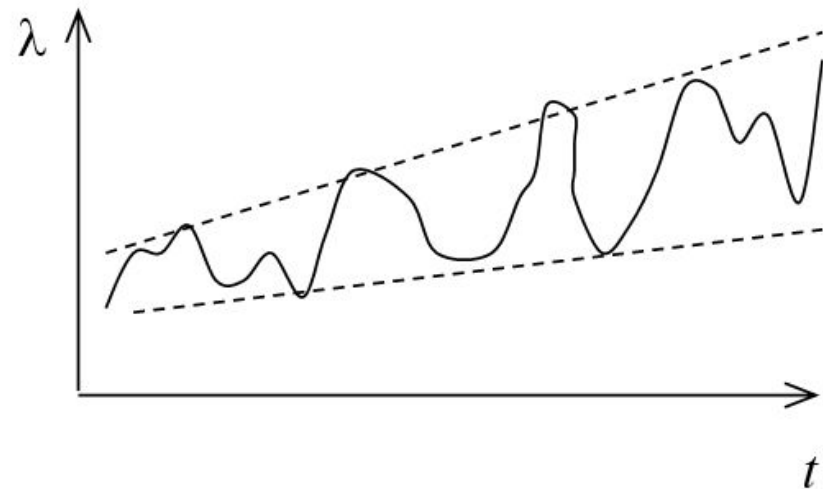
Значения коэффициентов a и b при линейной аппроксимации могут быть рассчитаны как на ПЭВМ, так и вручную по формулам:

$$b = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}; \quad a = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}.$$

ПРИМЕР

I. Классический метод решения

Попытка связать объемы продаж только с температурой, игнорируя влияние на них дня недели, несостоятельна. Это наглядно подтверждают вид графика (см. рисунок) и значение коэффициента корреляции.



ПРИМЕР

II. Неклассический метод решения

- Другой, менее точный подход к решению состоит в том, чтобы в единый статистический массив свести данные с понедельника по четверг без разделения их на дни недели. То же следует проделать с данными пятницы, субботы и воскресенья. Для каждого из массивов надо подобрать аппроксимирующую кривую зависимости объемов продаж от температуры и на ее основе делать прогноз.
- **Ошибка прогноза** определяется как среднеквадратическое отклонение фактических продаж от величин, рассчитанных описанным выше способом для фактических значений параметров (t , день недели). Особенность этого расчета состоит в том, что здесь $f = 2$, так как прогноз продаж строится на прогнозе температуры воздуха. Обычно $f = 1$. Полученное значение ошибки $\sigma = 22,548$.

ЗАДАЧИ

I. Задача №1

- Известна статистика продаж товара за первые 6 месяцев года (см. табл. ниже). Сегодня 1 июля. Спрогнозируйте продажи на последующие 4 месяца методами простого и скользящего среднего, методом экспоненциального сглаживания, если каждый месяц приносит новые данные о продажах, также показанные в таблице. Оцените погрешность прогноза каждым из методов. Постройте графики прогнозов всеми методами и график фактических продаж.

Месяц	1	2	3	4	5	6	7	8	9	10
Продажа, шт.	476	458	439	460	462	444	430	421	450	412

ЗАДАЧИ

II. Задача №2

- Решите задачу 1 с другими исходными данными (см. таблицы ниже). Дополнительно попробуйте подобрать аппроксимирующую кривую на интервале в десять месяцев на ПЭВМ с помощью пакета MS EXCEL. Сравните точность прогнозов разными способами.

Месяц	1	2	3	4	5	6	7	8	9	10
Продажа, шт.	2320	2335	2331	2370	2366	2357	2391	2388	2402	2412

Месяц	1	2	3	4	5	6	7	8	9	10
Продажа, шт.	46	42	39	45	40	48	43	42	40	41

ЗАДАЧИ

III. Задача №3

- Дайте для кафе обоснованный прогноз спроса на газированные напитки в бутылках на следующую неделю, основываясь на статистике продаж за прошедшие три недели и на прогнозе погоды на следующую неделю. Для расчета тенденции используйте пакет MS EXCEL. Постройте графики прогноза и фактических продаж.

Неделя	День недели	ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС
1-я	Продажа, шт.	182	155	163	187	111	63	54
	Температура воздуха, °С	26	24	27	23	24	19	18
2-я	Продажа, шт.	159	149	141	159	101	55	60
	Температура воздуха, °С	19	23	18	16	15	15	15
3-я	Продажа, шт.	151	139	144	163	94	58	57
	Температура воздуха, °С	16	19	21	18	14	15	12
4-я	Прогноз температуры, °С	13	13	13	20	23	23	23