

УДК 534.014.1, 534.5

Ермолаева А.А., Самсонова Л.В.

**О ПРИМЕНЕНИИ АНАЛИТИЧЕСКИХ ПРОГРАММ STATISTICA И LOGINOM В
ВЫЯВЛЕНИИ ЗАВИСИМОСТИ ПРОДАЖ ЛЕКАРСТВЕННЫХ ПРЕПАРАТОВ ОТ
ВРЕМЕНИ ГОДА, НА ПРИМЕРЕ ОДНОЙ ИЗ АПТЕК Г. УФЫ**

Башкирский государственный медицинский университет, г. Уфа

Проведено численное исследование по выявлению зависимости продажи лекарственных препаратов от времени года с помощью аналитических программ Statistica и Loginom.

Ключевые слова: аналитические программы Loginom, Statistica, интеллектуальный анализ данных.

В научных исследованиях часто возникает необходимость в статистическом анализе имеющихся данных. Нередко факторные признаки способствуют изменению результативных, что может говорить об их взаимосвязи и зависимости одного от другого. Для установления факта наличия или отсутствия зависимости между переменными, применяют корреляционный анализ. Для нахождения количественной зависимости между переменными, выраженной в виде уравнения или графика, — регрессионный анализ [1]. В отличие от корреляционного, регрессионный анализ говорит не только о наличии зависимости между одной или несколькими переменными, но и позволяет определить её количественно [2]. Целью нашего же исследования является выявление зависимости продажи лекарственных препаратов от времени года.

Материал и методы

Для проведения анализа были использованы данные о продажах различных лекарств в течение четырех сезонов 2022 года в одной из аптек города Уфы (Рисунок 1). Статистический анализ проводился с использованием программы : Statistica [5] и Loginom [6]. Количественные показатели оценивались на предмет соответствия нормальному распределению с помощью критерия Шапиро-Уилка. Количественные показатели, имеющие нормальное распределение, описывались с помощью средних арифметических величин (M) и стандартных отклонений (SD), границ 95% доверительного интервала (95% ДИ). Прогностическая модель, характеризующая зависимость количественной переменной от факторов, разрабатывалась с помощью метода линейной регрессии. Различия считались статистически значимыми при $p < 0,05$.

Лекарство	Зима	Весна	Лето	Осень	Всего	Количество поступивших препаратов
Противовирусное, Полиоксидоний 12мг, 10т.	25	7	12	16	60	56
Антибиотик, Сумамед 500 мг, 3т.	31	6	11	13	61	98
Витаминный комплекс от А до Zn, 30т.	19	17	14	24	74	100
Сухой кашель, Панатус форте сироп 7,5 мг\5мл, 200мл	18	2	3	14	37	67
Антигистаминное, Цетрин 10мг, 30т.	40	16	3	34	93	152
Анксиолитическое, Атаракс 25мг, 25т.	45	37	31	42	155	126
ЖКТ, ферменты, Микразим 10кМЕ, 40 к.	13	16	10	14	53	101
ЖКТ, спазмолитическое, Тримедат 200мг, 30т.	51	56	44	56	207	190

Рис. 1 - Данные о продажах лекарств в течение четырех сезонов 2022 г

Часть 1.

Загрузив исходные данные, в аналитической программе Statistica, мы воспользовались методом корреляционного и регрессивного анализа. Программа выдала нам следующую информацию: среднее значение; стандартное отклонение; значение коэффициента корреляции r ; значение коэффициента детерминации r^2 ; t -критерий; p – уровень значимости; число коррелируемых пар; свободный член уравнения регрессии: для зимы – 35,383; для весны – 71,406; для лета – 79,396; для осени – 46,469. Коэффициент при независимой переменной уравнения регрессии: для зимы – 2,508; для весны – 2,03; для лета – 1,99; для осени – 2,433. Очевидно, что взаимосвязь между временами года и проданными лекарствами есть. О связи между взятыми нами данными можно утверждать на основании коэффициента корреляции, который практически в каждый сезон объясняет около 80% значений переменной X относительно среднего значения. Коэффициент корреляции за летний сезон меньше, чем за другие сезоны (объясняет 65%) (Рисунок 2).

		Корреляции (Таблица данных1) Отмеченные корреляции значимы на уровне $p < 0,05000$ (Построчное удаление ПД)									
Пер.Х и Пер.У	Среднее	Стд.откл	$r(X,Y)$	r^2	t	p	N	Св. член завис.У	Наклон завис.У	среднее завис.Х	Наклон завис.Х
зима	30.2500	13.86414									
поступило	111.2500	43.91144	0.791846	0.627021	3.175953	0.019174	8	35.38331	2.507990	2.436470	0.250005
		Корреляции (Таблица данных1) Отмеченные корреляции значимы на уровне $p < 0,05000$ (Построчное удаление ПД)									
Пер.Х и Пер.У	Среднее	Стд.откл	$r(X,Y)$	r^2	t	p	N	Св. член завис.У	Наклон завис.У	среднее завис.Х	Наклон завис.Х
весна	19.6250	18.18113									
поступило	111.2500	43.91144	0.840608	0.706621	3.801495	0.008952	8	71.40630	2.030252	-19.0951	0.348046
		Корреляции (Таблица данных1) Отмеченные корреляции значимы на уровне $p < 0,05000$ (Построчное удаление ПД)									
Пер.Х и Пер.У	Среднее	Стд.откл	$r(X,Y)$	r^2	t	p	N	Св. член завис.У	Наклон завис.У	среднее завис.Х	Наклон завис.Х
лето	16.0000	14.28286									
поступило	111.2500	43.91144	0.647569	0.419346	2.081627	0.082549	8	79.39566	1.990896	-7.43277	0.210632
		Корреляции (Таблица данных1) Отмеченные корреляции значимы на уровне $p < 0,05000$ (Построчное удаление ПД)									
Пер.Х и Пер.У	Среднее	Стд.откл	$r(X,Y)$	r^2	t	p	N	Св. член завис.У	Наклон завис.У	среднее завис.Х	Наклон завис.Х
осень	26.6250	15.93682									
поступило	111.2500	43.91144	0.883048	0.779773	4.609190	0.003657	8	46.46868	2.433101	-9.02899	0.320485

Рис. 2 – Расчет коэффициента корреляции за летний сезон в Statistica.

Используя статистические расчеты программы Statistica, была получена прогностическая модель, характеризующая зависимость количественной переменной от факторов, разрабатывалась с помощью метода линейной регрессии: $Y=2,508X+35,383$; $r = 0,791$. Откуда, делаем вывод, что связь между количественной переменной от факторов (количество проданных лекарств от сезона) высокая. Направление прямое.

Аналогичные модели можно составить для остальных сезонов. Замечено, что в летнее время о корреляционной связи рассматриваемых препаратов от времени года говорить не приходится.

Часть 2.

Используя эти же данные в аналитической программе Loginom, мы воспользовались функцией «линейная регрессия». Были заданы стандартные настройки, предлагаемые программой. В качестве входных данных (независимая переменная) использовались столбцы сезонов и суммарного количества реализованной продукции каждого препарата (столбец «всего»); в качестве выходных данных (зависимая переменная) использовались данные о поступивших препаратах (столбец «количество поступивших препаратов») (Рисунок 3).

Настройка входных столбцов

Метка	Имя	Вид данных	Назначение
ab Лекарство	Lekarstvo	Дискретный	Не задано
9.0 Зима	Zima	Непрерывн...	Входное
9.0 Весна	Vesna	Непрерывн...	Входное
9.0 Лето	Leto	Непрерывн...	Входное
9.0 Осень	Osen	Непрерывн...	Входное
9.0 Всего	Vsego	Непрерывн...	Входное
9.0 Количество поступивших препаратов	Kolichestvo_postupivshikh_preparatov	Непрерывн...	Выходное

Рис. 3 – Оформление входных данных в Loginom

Общее качество регрессионной модели определяется с помощью коэффициента детерминации. В отчете по регрессии (Информация о модели) его величина составляет 0,93. Поскольку максимальное значение $R^2 = 1$, то можно утверждать, что качество регрессионной модели высокое. Здесь же представлены данные дисперсионного анализа, из которых следует статистическая значимость модели в целом при уровне значимости 0,05. В отчете по регрессии приведены регрессионные коэффициенты и уровень их статистической значимости (Рисунок 4).

Информация о модели		Шаги построения	Таблица	Дерево	Нулевые значения	
Модель	Финальная модель					
Показатель	Значение	Атрибут	Коэффициент	Стандартная ошибка	T-статистика	P-значение
Константа	Включена	ЖКТ, спазмолитическо...	6,306486	7,765416	0,812125	0,476177
Логарифм функции правдоподобия	-61,832764	Антибиотик, Сумамед 5...	5,485137	8,610421	0,637035	0,569387
Коэффициент детерминации	0,858425	Антигистаминное, Цет...	4,212908	10,333613	0,407690	0,710850
Коэффициент детерминации (скорр.)	0,480892	Противовирусное, Пол...	2,440675	13,085946	0,186511	0,863944
Стандартное отклонение	83,674891	Витаминный комплекс ...	-2,951250	16,398578	-0,179970	0,868646
		Анксиолитическое, Ата...	7,531081	13,629880	0,552542	0,619079

Рис. 4 – Расчет коэффициента корреляции за летний сезон в Logiном

По данным отчета, регрессионная модель имеет вид:

$$y_x = -45,576 + 6,306x_1 + 5,485x_2 + 4,212x_3 + 2,44x_4 - 2,951x_5 + 7,531x_6 + 4,498x_7 + 0,609x_8.$$

В графическом изображении мы получили удобную и понятную иллюстрацию, показывающую нам результат анализа (Рисунок 5).

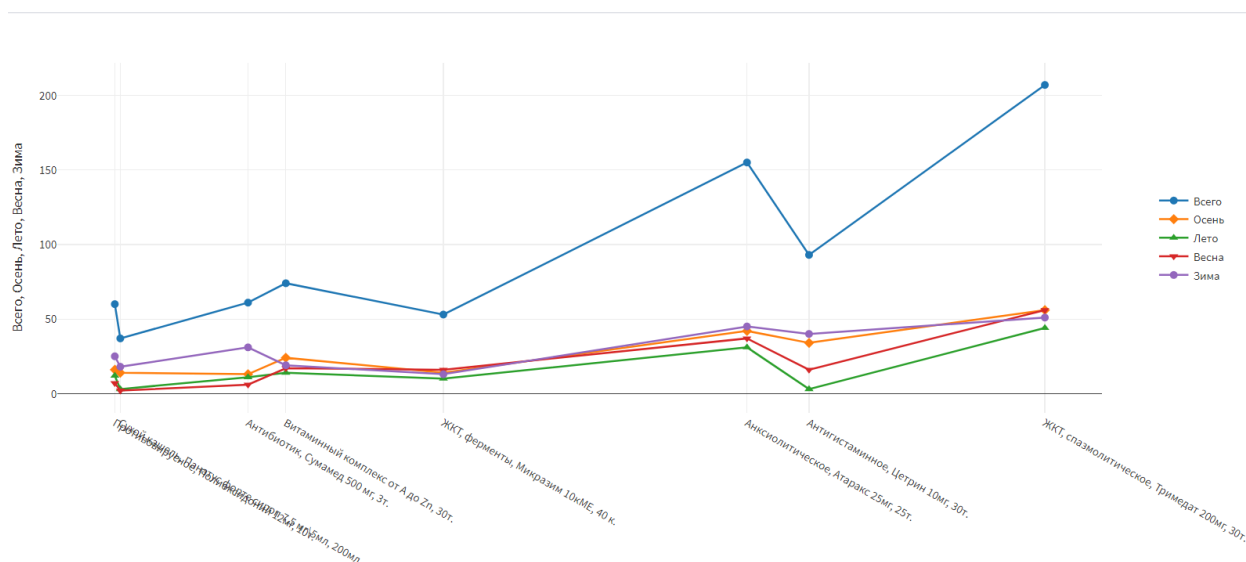


Рис. 5 – Модель в Logiном

Благодаря данному графику можно отслеживать влияние реализации конкретного препарата как в отдельном сезоне, так и за весь год. Например, можно заметить, что антибиотик «Сумамед» пользуется наибольшим спросом зимой и наименьшим — весной. Это означает, что количество поставок данного препарата весной в будущем будет снижено. Приложение Logiном позволяет вычислить точные числовые значения, показывающие примерное количество будущих поставок для каждого лекарства на основе исходных данных, что может существенно повлиять на прибыль.

Результаты и обсуждение

При изучении исходных данных мы выявили, что определенные препараты имеют сезонную динамику продаж. Анализируя объемы продаж за определенный период (например, год или сезон), мы можем предположить необходимое количество лекарств для улучшения

оборота товаров в аптеке и оценить эпидемиологическую обстановку в районе. Этот анализ, проведенный с использованием программ Loginom и Statistica, также может быть полезен для медицинского персонала поликлиник и других учреждений здравоохранения.

ЛИТЕРАТУРА

1. Жученко Ю. М., Ковалев А. А., Игнатенко В. А. Основы статистики Лабораторный практикум: учеб.-метод. Пособие для студентов 2 курса всех факультетов медицинских вузов / Ю. М. Жученко, А. А. Ковалев, В. А. Игнатенко. — Гомель: ГомГМУ, 2016. — 176 с.
2. Кочетов А.Г., Лянг О.В., Масенко В.П., Жиров И.В., Наконечников С.Н. Методы статистической обработки медицинских данных. Методические рекомендации для ординаторов и аспирантов медицинских учебных заведений, научных работников / сост.: А.Г. Кочетов, О.В. Лянг., В.П. Масенко, И.В.Жиров, С.Н.Наконечников, С.Н.Терещенко – М.: РКНПК, 2012. – 42 с.
3. Омельченко В. П. Информатика, медицинская информатика, статистика: учебник / В. П. Омельченко, А. А. Демидова. - Москва : ГЭОТАР-Медиа, 2021. - 608 с
4. Царик Г. Н. Информатика и медицинская статистика / под ред. Г. Н. Царик - Москва: ГЭОТАР-Медиа, 2017. - 304 с.
5. [Статистическое программное обеспечение](http://statsoft.ru) [Электронный ресурс]. URL: <http://statsoft.ru>. (Дата обращения: 02.09.23).
6. Аналитическая программа [Электронный ресурс]. URL.: <https://loginom.ru> (Дата обращения: 10.09.23).