

СТАТИСТИЧЕСКИЙ АНАЛИЗ ПОТРЕБЛЕНИЯ ПРОДУКТОВ ПИТАНИЯ НАСЕЛЕНИЕМ РЕГИОНОВ

Специальность: Экономика и управление народным хозяйством

Направление: Региональная экономика

Авторы: В.М. ИГНАТЬЕВ, доцент кафедры автоматизации и управления технологическими процессами и производствами ФГБОУ ВПО «Южно-Российский государственный политехнический университет (НПИ) имени М.И. Платова»,

М.В. СЕРЕДА, доцент кафедры менеджмента и информатики ФГБОУ ВПО "Донской государственный аграрный университет" Новочеркасский инженерно-мелиоративный институт имени А.К. Кортунова

В статье проводится статистический анализ потребление продуктов питания, оценивается их влияние на валовой региональный продукт, на продолжительность жизни, на коэффициенты рождаемости, смертности и заболеваемости.

In the article a statistical analysis of foodstuffs consumption is given, influence of these products on gross regional product and lifetime as well as on the birth rate, mortality and sickness rate is assessed.

Ключевые слова: валовой региональный продукт, продолжительность жизни, коэффициенты рождаемости, смертности и заболеваемости, потребление продуктов питания, корреляция, регрессионная зависимость, значимость, остатки уравнения.

Keywords: gross regional product, lifetime, birth rate, mortality, sickness rate, foodstuffs consumption, correlation, regression dependence, significance, equation residuals.

Потребляемые продукты питания влияют на валовой региональный продукт (ВРП), который определяется произведённой продукцией в регионе за год на душу населения. При оценивании степени влияние питания населения на ВРП необходимы статистические данные, по 80 регионам России. При проведении статистического анализа рассматриваются: среднедушевой ВРП, годовое среднедушевое потребление мяса, потребление молока, потребление яиц, сахара, растительного масла, хлеба, картофеля, овощей и фруктов. Значения показателей в регионах России ежегодно фиксируется Росстатом [1].

Статистический анализ данных проводился в следующей последовательности: проверка однородности данных, определение законов

распределения величин, корреляционный анализ, регрессионный анализ, кластерный анализ и анализ временных рядов.

Все рассматриваемые данные за 2011 г. однородны на 5 % уровне значимости, два ряда данных однородны на 10 % уровне. Все данные соответствуют нормальному закону распределения. Для данных о продолжительности жизни, для коэффициентов рождаемости и смертности, потребление сахара и растительного масла наилучшим является распределение Лапласа. Для коэффициента заболеваемости, ВРП, потребление мяса, молока, картофеля и х изделий овощей лучшим является распределение Стьюдента. Потребление хлеба и хлебобулочных изделий описывается распределением Парето на уровне значимости 0,15.

С помощью корреляционного анализа данных за 2011 г. по корреляционной матрице [2], по абсолютному значению коэффициента корреляции определили влияние на ВРП продуктов питания (по убыванию линейной взаимосвязи): хлеб, растительное масло, овощи, молоко, мясо, сахар, яйца, картофель. На прямолинейную зависимость указывают коэффициенты корреляции следующих продуктов: растительное масло, мясо, яйца.

В результате применения регрессионного анализа [3] было получено следующее регрессионное экспоненциальное уравнение для среднедушевого валового регионального продукта за 2011 год:

$$y = \exp(3,7456 + 0,06 x_1 + 0,00058 x_2 - 0,0076 x_3 - 0,0078 x_4 + 0,9 x_5 - 0,302 x_6 + 0,019 x_7 + 0,0044 x_8)$$

при коэффициенте корреляции равном 0,9565 (1)

где y – годовой среднедушевой валовой региональный продукт, руб. ;
 x_1 – годовое душевое потребление мяса и мясных продуктов, кг;
 x_2 – годовое душевое потребление молока и молочных продуктов, л;
 x_3 – годовое душевое потребление яиц, шт. ;
 x_4 – годовое душевое потребление сахара, кг;
 x_5 – годовое душевое потребление растительного масла, л;
 x_6 – годовое душевое потребление хлеба и хлебобулочных изделий, кг;
 x_7 – годовое душевое потребление картофеля, кг;
 x_8 – годовое душевое потребление овощей и фруктов, кг.

Коэффициент детерминации значим на 5 % уровне, что указывает на линейную зависимость потребления продуктов питания с экспонентой ВРП. Критерий Фишера для регрессионного уравнения равен $F(8,72) = 221,1$ и указывает, что уравнение значимо, так как теоретический критерий Фишера при уровне значимости 0,05 равен 2,07. Построение уравнения проводилось в среде пакета *STATISTICA* [4]. Остатки регрессионных уравнений – это разности между ретроспективными данными и

значениями, полученными по построенным регрессионным уравнениям. Остатки уравнения (1) гетероскедастичны [3], т. е. не соответствуют равномерному закону распределению по методу Дарбина-Уотсона на уровне значимости 0,05. Значение ВРП принимает значения с 6 разрядами, поэтому в качестве зависимой переменной мы использовали натуральный логарифм от значения ВРП.

Построенные регрессионные уравнения зависимости значения ВРП о потребления продуктов питания за 2000–2010 гг. дают подобные уравнения. Так регрессионное уравнение за 2000 г. имеет вид:

$$y = \exp(1,1435 + 0,0668 x_1 + 0,000544 x_2 - 0,007254 x_3 - 0,0114 x_4 + 0,907 x_5 - 0,271 x_6 + 0,0184 x_7 + 0,00527 x_8) \quad (2)$$

при скорректированном коэффициенте детерминации равном 0,9566.

В уравнении (1) факторы, оказывающие положительное влияния на ВРП это: потребление мяса, молока, растительного масла, картофеля и овощей. Факторы, оказывающие отрицательное влияния на ВРП это: потребление яиц, сахара и хлеба. Если потребление мяса и мясных продуктов в регионе увеличится на 1 кг, то ВРП возрастет на величину в $\exp(3,7456) \cdot \exp(0,06) = 44,95$ рубля; потребление молока и молочных продуктов возрастет на 1 литр – ВРП увеличится на $\exp(3,7456) \cdot \exp(0,00058) = 42,36$ рубля; потребление растительного масла в регионе увеличится на 1 литр – ВРП возрастёт на величину в $\exp(3,7456) \cdot \exp(0,9) = 104,13$ рубля; потребление картофеля увеличится на 1 кг – ВРП возрастет на величину в $\exp(3,7456) \cdot \exp(0,019) = 43,15$ рубля; потребление овощей и фруктов увеличится на 1 кг – ВРП возрастет: $\exp(3,7456) \cdot \exp(0,0044) = 42,52$ рубля. При увеличении потребления яиц в регионе на 1 шт., ВРП, напротив, будет уменьшаться на величину в $\exp(3,7456) \cdot \exp(0,0076) = 42,66$ рубля, так как этот фактор отрицательно влияет на величину ВРП. Аналогично: увеличение потребление сахара в регионе на 1 кг, приведет к уменьшению ВРП на величину в $\exp(3,7456) \cdot \exp(0,0078) = 42,67$ руб.; увеличение потребление хлеба и хлебобулочных изделий в регионе на 1 кг, приведет к снижению ВРП на величину в $\exp(3,7456) \cdot \exp(0,302) = 57,26$ рубля. В регрессионных уравнениях (1) и (2) зависимости ВРП от продуктов питания значимы (робастные) коэффициенты при значениях потребления мяса, растительного масла и картофеля.

Добавим в число независимых переменных влияющих на ВРП в 2011 г. коэффициент заболеваемости в регионе. Заболеваемость в регионе обозначим через x_9 , которое измеряется в количестве заболеваний на 1000 человек: промилле; или ‰. Остальные обозначения прежние.

В результате использования регрессионного анализа получим следующее регрессионное уравнение:

$$y = \exp(6,966 + 0,05901 x_1 - 0,00088 x_2 - 0,00678 x_3 - 0,0163 x_4 + 0,8841 x_5 - 0,313 x_6 + 0,0216 x_7 + 0,0216 x_8 - 0,003 x_9) \quad (3)$$

при скорректированном коэффициенте детерминации равном 0,95786.

Регрессионное уравнение (3) значимо более чем уравнение (1). Это указывает на то, что средняя заболеваемость населения региона влияет на значение ВРП. Уравнение (3) значимо, а остатки уравнения (3) на уровне 0,05 гомоскедастичны. Поэтому уравнение (3) можно использовать в качестве расчётной модели.

Влияние потребления продуктов питания на коэффициент заболеваемости населения (z) можно описать экспоненциальным регрессионным уравнением вида:

$$z = \exp(1,344 + 0,05464 x_1 + 0,0021 x_2 - 0,01 x_3 - 0,00457 x_4 + 0,91 x_5 - 0,0265 x_6 + 0,0177 x_7 + 0,00537 x_8) \quad (4)$$

при скорректированном коэффициенте детерминации, равном 0,955.

Регрессионное уравнение (4) значимо по критерию Фишера на уровне значимости 0,05.

Для продолжительности жизни (p) в регионах России было получено регрессионное уравнение вида (5):

$$p = 63,66 - 0,6145x_1 + 0,21x_2 + 0,0006x_3 + 0,133x_4 - 0,264x_5 - 0,01145x_6 - 0,234x_7 + 0,0184 x_8 \quad (5)$$

при значении коэффициента детерминации, равном 0,4.

Коэффициент детерминации по критерию Стьюдента значим на 5 % уровне, уравнение не значимо даже при уровне 0,1.

Для получения приемлемых суждений о взаимосвязи параметров проводилось разбиение регионов на кластеры (данные Росстата). Число классов, на которое были разбиты регионы, равно 4. При проведении кластеризации регионов с помощью метода k -средних [4] значения исходных статистических данных пересчитываются по следующей формуле: $K_j = (M - I_j) / \sigma$, $j = 1 \dots 80$,

где M – среднее данного; σ – среднеквадратичное отклонение данного; в качестве I последовательно используются удельное потребление каждого продукта питания и средняя продолжительность жизни. Кластерный анализ разбил регионы на четыре кластера с центрами, приведёнными на рис. 1. Список регионов приведён в табл. 1.

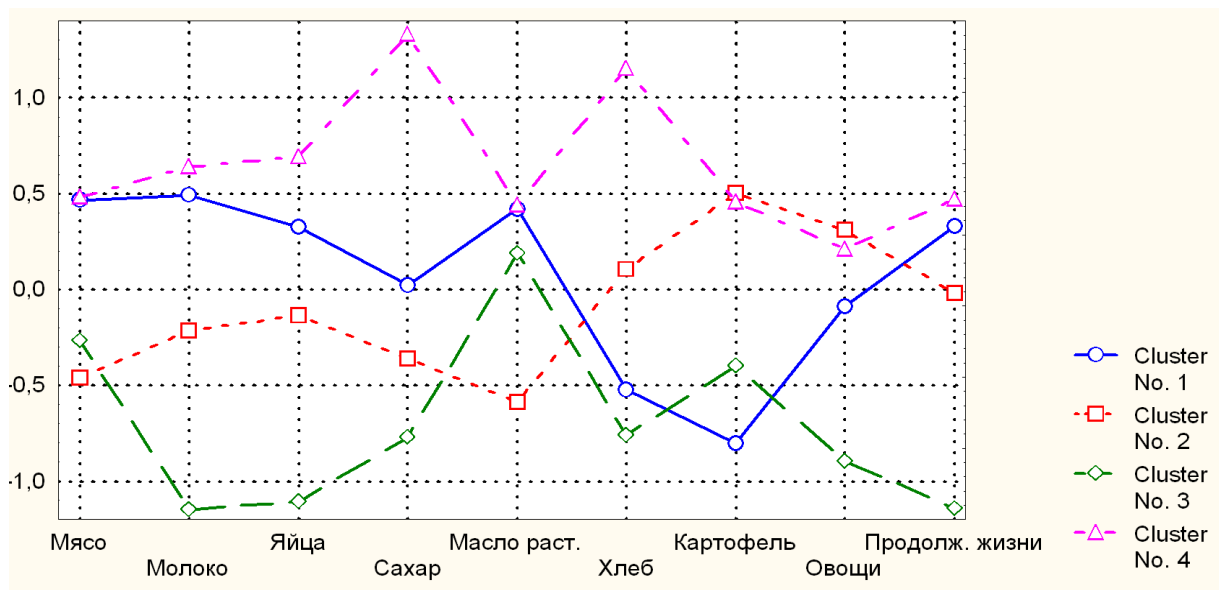


Рисунок 1 – Диаграмма центров кластеров

В первый кластер вошел 22 региона из 80. Это следующие регионы: 17,18,20÷25,27÷30,34,39,42,49,50÷54,57,72. Для этого кластера характерна: низкая доля потребления мяса, молока; высокая доля потребления картофеля; высокая продолжительность жизни. Третий кластер составляют 12 регионов: 5,16,21,61,62,65,67,73,74,77,78, 80. Для третьего кластера характерна: высокая доля потребления молока, яиц и овощей. В регионах кластера наблюдается высокая продолжительность жизни. Четвертый кластер составляют 15 регионов: 1,4,8,9,10,14,15,31,38,41,43, 44, 45,64,70. В 4-м кластере наблюдается низкое потребление мяса, молока, яиц, сахара, растительного масла, хлеба, картофеля и овощей. Продолжительность жизни для данного кластера принимает среднее значение. Во второй кластер вошли остальные регионы. Регионы второго кластера имеют самую низкую долю потребления картофеля и овощей и низкую продолжительность жизни.

Таблица 1 – Номера и список регионов Российской Федерации

№	Регион	№	Регион	№	Регион
1	Белгородская область	28	г. Санкт-Петербург	55	Ульяновская область
2	Брянская область	29	Республика Адыгея	56	Курганская область
3	Владимирская обл.	30	Республика Калмыкия	57	Свердловская область
4	Воронежская область	31	Краснодарский край	58	Тюменская область
5	Ивановская область	32	Астраханская область	59	Челябинская область
6	Калужская область	33	Волгоградская область	60	Республика Алтай
7	Костромская область	34	Ростовская область	61	Республика Бурятия
8	Курская область	35	Республика Дагестан	62	Республика Тыва
9	Липецкая область	36	Республика Ингушетия	63	Республика Хакасия
10	Московская область	37	Кабардино-Балкарская Республика	64	Алтайский край
11	Орловская область	38	Карачаево-Черкесская Республика	65	Забайкальский край
12	Рязанская область	39	Республика Северная Осетия	66	Красноярский край
13	Смоленская область	40	Чеченская Республика	67	Иркутская область
14	Тамбовская область	41	Ставропольский край	68	Кемеровская область
15	Тверская область	42	Республика Башкортостан	69	Новосибирская обл.
16	Тульская область	43	Республика Марий Эл	70	Омская область
17	Ярославская область	44	Республика Мордовия	71	Томская область
18	г. Москва	45	Республика Татарстан	72	Республика Саха (Якутия)
19	Республика Карелия	46	Удмуртская Республика	73	Камчатский край
20	Республика Коми	47	Чувашская Республика	74	Приморский край
21	Архангельская обл.	48	Пермский край	75	Хабаровский край
22	Вологодская область	49	Кировская область	76	Амурская область
23	Калининградская обл.	50	Нижегородская область	77	Магаданская область
24	Ленинградская обл.	51	Оренбургская область	78	Сахалинская область
25	Мурманская область	52	Пензенская область	79	Еврейская авт. обл.
26	Новгородская обл.	53	Самарская область	80	Чукотский автономный округ
27	Псковская область	54	Саратовская область		

По рассматриваемым 8 параметрам, характеризующим потребление продуктов питания в каждом кластере, и продолжительность жизни населения, можно построить линейные значимые регрессионные зависимости.

Для регионов первого кластера регрессионное уравнение имеет вид:
 $p_1 = 75,24 - 0,095x_1 + 0,43x_2 - 0,00022x_3 + 0,1833x_4 - 0,2344x_5 - 0,0926x_6 - 0,28x_7 - 0,0043x_8$ при коэффициенте детерминации 0,54.

Для регионов третьего кластера регрессионное уравнение имеет вид:
 $p_2 = 55,78 - 0,026x_1 + 0,245x_2 - 0,0032x_3 + 0,0672x_4 + 0,1273x_5 - 0,0487x_6 + 0,0082x_7 - 0,062x_8$
 при коэффициенте детерминации 0,773.

Зависимость для регионов четвертого кластера получилась следующая:

$$p_3 = 66,71 - 0,02x_1 + 0,0196x_2 - 0,00734x_3 + 0,101x_4 - 0,208x_5 + 0,0193x_6 - 0,0353x_7 - 0,0294x_8$$

при коэффициенте детерминации 0,342.

Для регионов второго кластера зависимость получилась следующая:

$$p_4 = 67,22 - 0,083x_1 + 0,0138x_2 - 0,0134x_3 + 0,061x_4 - 0,52x_5 + 0,058x_6 - 0,0356x_7 + 0,0109x_8$$

при коэффициенте детерминации 0,63. Коэффициенты детерминации для регрессионных уравнений по критерию Стьюдента значимы для всех кластеров. Построенные линейные регрессионные уравнения значимы для кластеров, за исключением четвертого кластера. Остатки всех полученных регрессионных уравнений для продолжительности жизни гетероскедастичны.

Увеличение потребления мяса в год на душу населения во всех регионах отрицательно сказывается на величину продолжительности жизни, так как коэффициент при x_1 во всех уравнениях имеет отрицательное значение. Влияние потребления продуктов питания на другие демографические показатели в регионе рассматривалась в статье [5]. В качестве этих показателей выступали коэффициенты рождаемости и смертности. Полученные зависимости коэффициента рождаемости и коэффициента смертности имели коэффициенты детерминации $d_{\text{рожд.}} = 0,6036$ и $d_{\text{смерт.}} = 0,3638$. Для получения регрессионных уравнений с гомоскедастичными остатками необходимо провести кластеризацию регионов по 9 показателям (x_1-x_8 и $rogd$) и (x_1-x_8 и $smer$). Затем для каждого кластера построить своё регрессионное уравнение.

Влияние на величину потребления продуктов питания идёт также и со стороны ВРП через среднедушевой доход, заработную плату и пенсию.

В нормативном документе [1] даются рекомендации по рациональным нормам потребления пищевых продуктов. Нормы необходимы для активного и здорового образа жизни населения страны, они разработаны в целях укрепления здоровья населения, профилактики неинфекционных заболеваний и улучшения демографической ситуации в Российской Федерации. Значения фактического потребления продуктов питания в регионах страны должны соответствовать этим нормам и эти нормы соблюдаются.

На основе данных Росстата можно провести прогнозы потребления продуктов питания. Ряды составляют 20 значений по каждому продукту питания для Ростовской области за 1992 ÷ 2011 гг.

Прогноз, выполненный с помощью ряда Фурье, при 12 гармониках, задаётся с помощью формулы:

$$V_i = a \cdot i + b + \sum_{k=1}^6 \left(c_{2k-1} \sin \frac{k\pi(i-1)}{6} + c_{2k} \cos \frac{k\pi(i-1)}{6} \right), \quad (6)$$

где V_i – величина прогнозируемого показателя за i -й год; коэффициенты $a, b, c; r$ – коэффициент корреляции.

В результате построения рядов Фурье были получены коэффициенты a, b, c уравнения (5), приведенные в табл. 2.

Таблица 2 – Коэффициенты уравнения (6) потребления продукта

Обозначения	Значение коэффициентов уравнения потребления продукта							
	Мясо	Молоко	Яйца	Сахар	Картофель	Овощи	Растит, масло	Хлеб
a	6.347	12.668	-3.746	-0.11	6.808	3.755	0.089	9.437
b	0.896	14.872	29.539	0.04	2.902	-1.027	0.588	5.62
c_1	1.765	5.655	11.046	-0.343	-1.699	0.734	0.281	0.805
c_2	3.404	4.135	7.246	-1.01	-1.619	3.226	-0.565	1.616
c_3	0.16	-2.379	1.026	-7e-10	4.964	2.022	0.293	5.806
c_4	0.243	-1.879	3.776	-0.5	-3.036	2.689	-0.107	-1.111
c_5	2.11	4.312	3.077	0.09	-1.32	3.306	-0.076	-0.173
c_6	1.102	3.931	4.127	-0.531	-1.814	-0.076	0.079	0.882
c_7	-0.303	0.428	-1.794	-0.327	0.631	-0.258	0.151	0.491
c_8	0.841	5.121	3.764	-0.79	-0.225	4.405	-0.152	1.408
c_9	1	1	1	1	1	1	1	1
c_{10}	0.413	-0.273	0.721	-0.333	-2.268	-2.114	-0.049	-0.68
c_{11}	6.347	12.668	-3.746	-0.11	6.808	3.755	0.089	9.437
c_{12}	0.896	14.872	29.539	0.04	2.902	-1.027	0.588	5.62
r	0,64	0,545	0,798	0,392	0,388	0,898	0,692	0,74

График изменения годового потребления овощей и фруктов на одного человека в Ростовской области: с 1992 по 2011 гг. приведён на рис. 2.

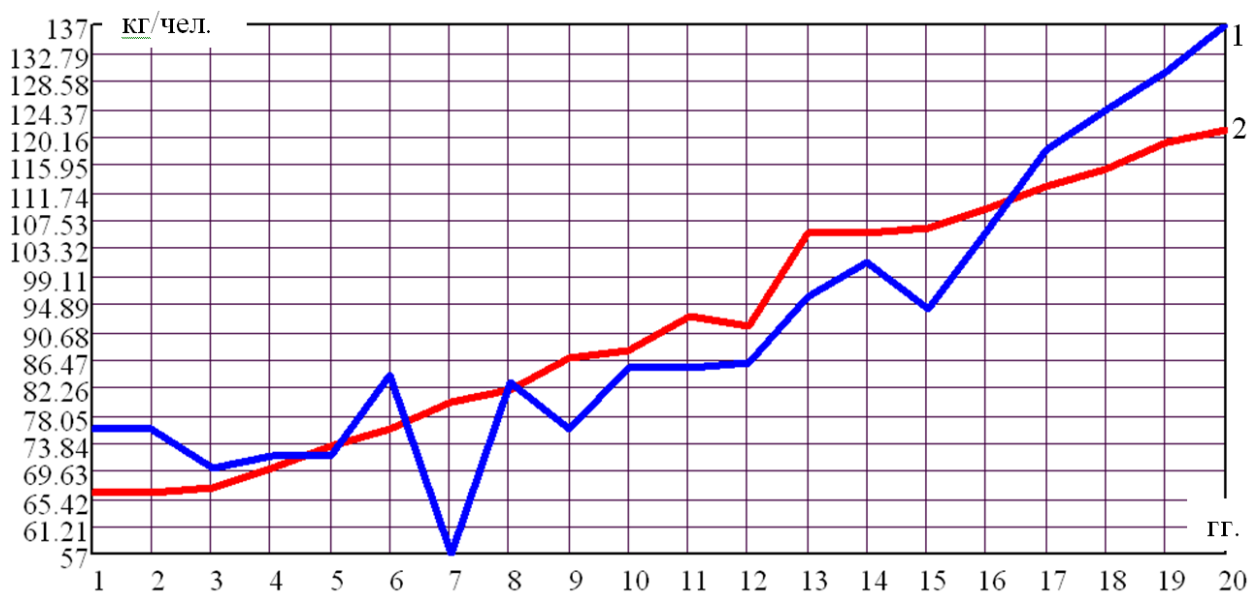


Рисунок 2 – Среднедушевое годовое потребление овощей и фруктов в Ростовской области: с 1992 по 2011 гг: 1 – факт; 2 – прогноз

Коэффициент линейного тренда a в табл. 2 в столбце «овоши», равный 3,755, указывает на ежегодное увеличение потребления овощей и фруктов в кг для жителей Ростовской области.

Список литературы

1. Регионы России. Социально-экономические показатели. 2013: Стат. сб. // Росстат. – М., 2014.
2. Доугерти К. Введение в эконометрику: Учебник. – М.: ИНФРА-М, 2007, 432 с.
3. Дрейпер Н.Р., Смит Г. Прикладной регрессионный анализ. – М.: Диалектика, 2007, 912 с.
4. Боровиков В. *STATISTICA*: искусство анализа данных на компьютере. – М.: Питер, 2001, 650 с.
5. Игнатьев В.М., Потребление продуктов питания населением регионов // Стратегия устойчивого развития регионов России, 2015. – № 25. – С. 132–137.
6. Рекомендации по рациональным нормам потребления пищевых продуктов, отвечающих современным требованиям здорового питания: утверждённые приказом Министерства здравоохранения и социального развития РФ от 2 августа 2010 г. № 593н.