

Моделирование рынка культивируемых грибов методами дискриминантного и факторного анализа

А.А. Набоких

В современных условиях в нашей стране актуальной является задача продовольственной безопасности [1]. Отечественный продовольственный рынок насыщен импортной, зачастую некачественной продукцией, в том числе продукцией грибоводства. Российским производителям все труднее сохранять достигнутое ими положение. Применение маркетингового инструментария, предоставляющего возможность исследования рынка культивируемых грибов, становится сугубо злободневным. Значительное влияние на результативность деятельности грибоводческих предприятий оказывает выбор способов извлечения и анализа маркетинговых данных [2 - 4].

В настоящее время по-прежнему важна роль анализа маркетинговой информации с целью выявления основных тенденций развития в эволюционировании различных сегментов экономики [5]. Достаточно оперативно полученные следствия позволят компетентно инвестировать финансовые средства в экономику региона. Для реализации такого подхода возможно применение математического аппарата дискриминантного и факторного анализов, которые реализованы в некоторых программных системах [6, 7]. Для преобразования, позволяющего анализировать переменные, была выбрана программная система SPSS Statistics [8].

На потребительское поведение влияет множество обстоятельств, среди которых выделяют социальные, культурные, психологические и индивидуальные факторы. В обстановке жесткого обострения уровня конкурентоспособности залогом перспективного успеха любой фирмы прежде всего остается направленность на потребителя. Поведение отдельного покупателя имеет конкретные основания.

Цель исследователя – определение общего шаблона потребительского поведения в обширном вербальном и невербальном информационном потоке. Признание фирмой тезиса удовлетворения интересов потребителей определяет тот факт, что фирма функционирует по следующим направлениям: извлечение требующихся данных о рынке и потребностях, намерениях, перцепциях, предпочтениях и желаниях отдельного индивидуума осуществить приобретения; сопровождение плана и стратегии, сосредоточенных на удовлетворение спроса покупателей и пр.

Дискриминантная модель рынка культивируемых грибов Кировской области

Основной задачей анализа регионального рынка культивируемых грибов является выявление и оценка наиболее значимых факторов, влияющих на принятие решения о покупке культивируемых грибов. Указанная задача может быть решена методами дискриминантного анализа.

Дискриминантный анализ является многомерным статистическим методом, позволяющим исследовать различия между несколькими объектами по целому ряду признаков одновременно. Каждому из них приписывается некоторый коэффициент, выведенный опытным путем на основе обследования значительной группы респондентов. При дальнейшем прогнозировании полученные соответствия позволяют классифицировать потребителей на тех, кто приобретает культивируемые грибы, и тех, кто не относится к этой категории.

Моделью дискриминантного анализа является функция

$$D = b_0 + b_1x_1 + \dots + b_kx_k,$$

где D – дискриминантный показатель, b_i – дискриминантный коэффициент, x_i – предиктор ($i=1,2,\dots,k; k \in \mathbb{N}$).

Для составления модели был проведено пилотажное исследование летом 2013 года. Генеральная совокупность – 100 жителей Кировской области старше 18 лет, осознано приобретающие (не приобретающие) культивируемые грибы. Выборка формировалась случайным образом. Независимые переменные – пол, возраст, доход и социальный статус

опрашиваемого. Зависимая переменная – вероятность покупки культивируемых грибов.

Моделирование производилось с использованием реализованного в системе SPSS модуля «Дискриминантный анализ» [9, 10].

Для проведения теста на равенство средних значений в группах опрашиваемых, приобретающих и не приобретающих культивируемые грибы, в качестве статистического критерия в SPSS используется лямбда Уилкса. В качестве дискриминационных переменных могут выступать лишь доход и социальный статус опрашиваемых (таблица №1). Переменные «пол», «возраст», «семейное положение» не обладают разделительными свойствами и должны быть исключены из дискриминантной модели.

Таблица №1

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Ваш пол?	,994	,609	1	98	,437
Ваш возраст?	,980	2,041	1	98	,156
Ваш доход?	,951	5,020	1	98	,027
Ваш социальный статус?	,924	8,010	1	98	,006
Семейное положение	,987	1,324	1	98	,253

Кроме того, переменные «доход» и «социальный статус» являются некоррелированными (коэффициент корреляции между ними составляет всего $0,123 << 0,5$), т.е. указанные переменные не обуславливают друг друга.

Показатель «Лямбда Уилкса» (0,044) свидетельствует о значимости различий средних значений дискриминантной функции в исследуемых группах (таблица №2).

Таблица №2

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,894	10,894	2	,004

При помощи стандартизированных коэффициентов дискриминантной функции был оценен относительный вклад каждой дискриминационной переменной в различие двух исследуемых групп: доход респондентов в 1,35

(0,760/0,563) раза меньше влияет на желание приобретать культивируемые грибы, чем их социальный статус.

Дискриминантная модель, построенная в результате проведения анализа, имеет следующий вид:

$$D = -3,422 + 0,757x_1 + 0,547x_2,$$

где x_1 – доход, x_2 – социальный статус.

Четкость разделения исследуемых групп характеризуется расстоянием между средними значениями дискриминантной функции в исследуемых группах.

Таблица №3

Functions at Group Centroids

	Function
Покупаете ли Вы культивируемые грибы?	1
Не покупаю	-1,024
Покупаю	,114

Как видно из таблицы №3, среднее значение дискриминантной функции для респондентов, покупающих культивируемые грибы, составляет -1,024, не покупающих – 0,114. Чем больше расстояние между средними значениями дискриминантной функции в исследуемых группах, тем более четко прослеживается различие между исследуемыми группами.

Четкость различия между исследуемыми группами зависит также от рассеяния значений дискриминантной функции в исследуемых группах. Это рассеяние показано на графиках распределения дискриминантной функции (рисунки 1, 2).

Canonical Discriminant Function 1

Покупаете ли Вы культивируемые грибы? = Не покупаю

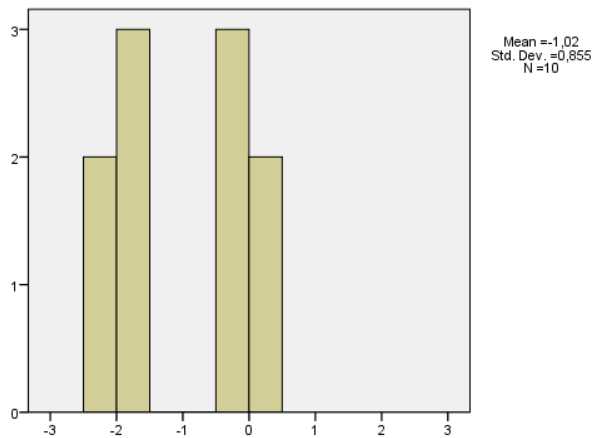


Рис. 1. — Распределение дискриминантной функции в группе респондентов, не приобретающих культивируемые грибы

Canonical Discriminant Function 1

Покупаете ли Вы культивируемые грибы? = Покупаю

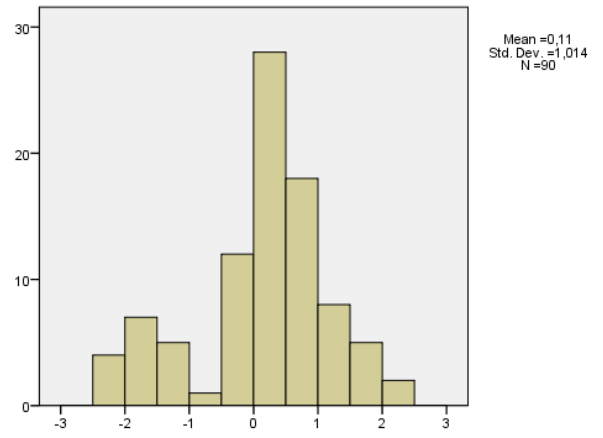


Рис. 2. — Распределение дискриминантной функции в группе респондентов, приобретающих культивируемые грибы

На основе построенной дискриминантной модели можно осуществить прогнозирование того, приобретает данный человек культивируемые грибы или нет, исходя из его возраста и социального статуса. Например, для человека, принадлежащего к социальной группе «рабочие», чей доход средний, значение дискриминантной функции составит:

$$D = -3,422 + 0,757 \cdot 2 + 0,547 \cdot 4 = 0,28.$$

Согласно данным, представленным на рисунке 1, исследуемая группа «респонденты, не покупающие культивируемые грибы» включает 10 человек. Значение дискриминантной функции, близкое к 0,28 имеют 2 человека. По данным рисунка 2, исследуемая группа «респонденты, покупающие культивируемые грибы» включает 90 человек. Значение дискриминантной функции, близкое к 0,28 имеют около 18 человек. На основании вышеизложенного можно сделать вывод, что человек, принадлежащий к социальной группе «рабочие», со средним доходом, скорее всего, будет приобретать культивируемые грибы.

Точность прогнозов на основе построенной дискриминантной модели оценивается по результатам классификации. В таблице №4 представлены

результаты классификации отдельно по каждому респонденту, принявшему участие в опросе по первым пяти наблюдениям.

Таблица №4

Casewise Statistics

Case Number	Actual Group	Highest Group					Second Highest Group			Discriminant Scores
		Predicted Group	P(D>d G=g)		P(G=g D=d)	Squared Mahalanobis Distance to Centroid	Group	P(G=g D=d)	Squared Mahalanobis Distance to Centroid	Function 1
			p	df						
Original	1	1	,476	1	,811	,509	0	,189	3,426	,827
	2	1	,868	1	,698	,028	0	,302	1,700	,280
	3	1	,965	1	,645	,002	0	,355	1,196	,070
	4	1	,965	1	,645	,002	0	,355	1,196	,070
	5	1	0**	,584	1	,781	,299	1	,219	2,838

В столбце «Actual Group» указывается фактическая принадлежность респондента к одной из исследуемых групп. В столбце «Predicted Group» указывается прогнозируемая принадлежность респондента к одной из групп, определяемая на основе построенной модели. Если прогнозируемая принадлежность к группе не совпадает с фактической, ее значение отмечается **. В столбце «P(G=g | D=d)» указывается вероятность, с которой данный респондент может быть причислен к прогнозируемой группе. В столбце «Discriminant Scores» указывается значение дискриминантной функции.

Точность прогнозов на основе построенной модели определяется из данных сводной таблицы результатов классификации (таблица №5).

Таблица № 5

Classification Results^a

		Predicted Group Membership		Total
		Не покупаю	Покупаю	
Original Count 1	Не покупаю	8	2	10
	Покупаю	27	63	90

%	Не покупаю	80,0	20,0	100,0
	Покупаю	30,0	70,0	100,0

a. 71,0% of original grouped cases correctly classified.

Из данных таблицы №5 видно, что корректные результаты классификации составили 71%, данная величина и составляет точность прогнозов.

Факторная модель рынка культивируемых грибов Кировской области

Целью проведения факторного анализа является разбиение таблицы данных на группы (факторы). В один фактор включается группа переменных, имеющих высокую корреляцию друг с другом и слабую корреляцию с переменными, агрегированными другими факторами.

Для оценки пригодности исходных данных к применению факторного анализа были использованы тесты КМО и Bartlett.

Таблица №6

КМО и Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,644
Bartlett's Test of Sphericity	Approx. Chi-Square	3,635E3
	df	300
	Sig.	,000

Из данных таблицы 6 видно, что результаты факторного анализа могут считаться адекватными, поскольку значение КМО=0,644>0,5. Значимость теста Bartlett (Sig.) составляет 0, что свидетельствует о существовании корреляционных связей между переменными исходного массива и возможности их группировки.

Число групп факторной модели было определено при помощи расчета характеристических чисел (Eigenvalues).

Таблица №7

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,862	19,448	19,448	4,862	19,448	19,448
2	2,852	11,407	30,855	2,852	11,407	30,855

3	2,319	9,278	40,133	2,319	9,278	40,133
4	2,031	8,124	48,258	2,031	8,124	48,258
5	1,524	6,097	54,355	1,524	6,097	54,355
6	1,352	5,408	59,762	1,352	5,408	59,762
7	1,195	4,781	64,544	1,195	4,781	64,544
8	1,010	4,038	68,582	1,010	4,038	68,582
9	1,002	4,008	72,590	1,002	4,008	72,590
10	,861	3,444	76,034			
11	,739	2,955	78,990			
12	,671	2,686	81,675			
13	,654	2,615	84,291			
14	,542	2,168	86,459			
15	,511	2,042	88,501			
16	,484	1,935	90,436			
17	,436	1,742	92,178			
18	,395	1,579	93,757			
19	,312	1,248	95,006			
20	,289	1,157	96,163			
21	,253	1,014	97,177			
22	,216	,865	98,041			
23	,191	,765	98,806			
24	,159	,637	99,444			
25	,139	,556	100,000			

Extraction Method: Principal Component Analysis.

На основании таблицы №7 можно утверждать, что оптимальное число групп в модели составляет 9. Факторная модель, состоящая из указанного числа групп, сохраняет 72,59% исходной информации. Результаты таблицы 7 подтверждает рисунок 3.

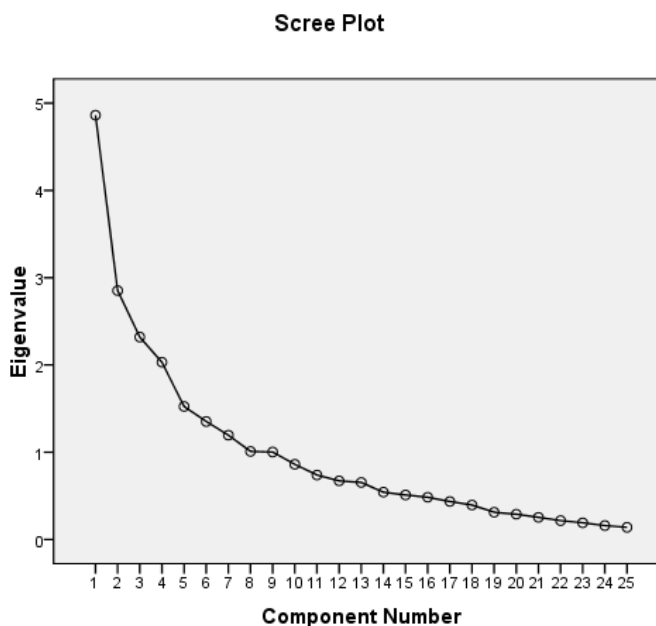


Рис. 3. – Графическое определение количества компонентов факторной модели.

Рисунок 3 отображает зависимость между характеристическими числами и числом компонентов факторной модели. При колебаниях количества факторов с 10 до 25 график представляет собой прямую, что еще раз подтверждает выводы таблицы 7.

В таблице №8 представлены коэффициенты корреляции, описывающие связи между переменными исходного массива данных и факторами построенной модели.

Таблица №8

Rotated Component Matrix

	Component								
	1	2	3	4	5	6	7	8	9
Качество культивируемых грибов	,210	,692	,085	,139	,209	,079	-,205	-,278	-,006
Цена	-,153	-,332	-,192	,084	-,251	-,134	,097	,106	,596
Сортировка культивируемых грибов по форме и размеру	,566	,031	-,014	,047	,340	,408	,035	,056	,074
Упаковка культивируемых грибов	,256	,246	,190	,046	,599	,223	,203	,139	,008
Торговый знак – бренд	,462	,085	,114	-,451	,436	,269	,194	-,262	,025
Наличие свежих (не переработанных) грибов	,170	,191	-,202	-,531	-,018	,450	,162	-,004	,034
Известность производителя	,058	,036	-,053	-,077	,852	,071	,002	-,020	-,015

Вид торговой организации	,011	,434	,311	,289	,242	,460	,204	,222	,036
Место продажи в торговом зале магазина	,043	,293	,704	-,057	-,059	,308	,127	,099	-,037
Уровень знаний продавца о культивируемых грибах	,071	,102	,029	-,001	,170	,875	-,011	,077	-,061
Уровень развития маркетинга культивируемых грибов	-,011	,013	,160	,165	-,059	,332	,032	,705	,036
Уровень развития кулинарии из культивируемых грибов	-,040	,099	,027	-,250	-,033	,006	,543	,494	,084
Уровень позиционирования культивируемых грибов	,037	,018	,047	,129	,026	,123	,865	-,035	,099
Уровень знаний покупателя о грибах	,117	,118	-,004	,016	,366	,219	,349	,100	,684
Уровень знаний покупателя о полезных свойствах культивируемых грибов	,390	,030	,011	,665	,209	,051	,151	,197	-,042
Уровень знаний покупателя о технологии производства культивируемых грибов	-,082	,008	-,039	,803	-,201	,097	-,082	-,121	,100
Уровень культуры питания покупателя	-,025	-,349	,693	,331	,003	,055	-,032	-,110	-,273
Уровень здорового образа жизни покупателя	-,022	,222	,796	-,123	,113	-,265	-,212	,054	,020
Уровень предложений блюд из культивируемых грибов в местах общественного питания	-,637	-,420	-,122	-,009	-,146	,136	,150	,181	,093
Социальный статус покупателя	,237	,017	,191	,104	-,412	,309	,190	-,642	,006
Уровень дохода покупателя	,110	-,644	,031	,214	-,135	-,091	-,126	-,060	,522
Уровень культуры и образования покупателя	,087	,141	,374	,280	-,203	,113	-,655	,064	,075
Наличие доступной для восприятия информации о культивируемых грибах	-,502	-,205	-,064	-,121	,014	-,096	-,159	-,117	,577
Наличие дней грибной культуры	-,871	-,023	,072	,033	,002	-,050	,029	,070	,048
Время года	-,193	-,650	-,175	,109	,069	-,147	-,090	-,270	,132

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser

Normalization.

На основе этих данных производится группировка переменных исходного массива.

Таблица №9

Группировка переменных исходного массива данных

Факторы модели	Переменные исходного массива	Коэффициенты корреляции
Фактор 1 (торговый сервис)	Сортировка культивируемых грибов по форме и размеру	0,566
	Упаковка культивируемых грибов	0,462
	Уровень знаний продавца о культивируемых грибах	0,637
	Уровень позиционирования культивируемых грибов	0,871
Фактор 2 (Равновесие спроса)	Качество культивируемых грибов	0,692
	Цена	0,644
	Уровень дохода покупателя	0,65
Фактор 3 (Культура питания)	Уровень развития кулинарии из культивируемых грибов	0,704
	Уровень культуры питания покупателя	0,693
	Наличие свежих (не переработанных) грибов	0,796
Фактор 4 (Знания покупателя)	Уровень знаний покупателя о грибах	0,531
	Уровень знаний покупателя о полезных свойствах культивируемых грибов	0,665
	Уровень знаний покупателя о технологии производства культивируемых грибов	0,803
Фактор 5 (Имидж производителя)	Торговый знак – бренд	0,599
	Известность производителя	0,852
Фактор 6 (Имидж продавца)	Вид торговой организации	0,46
	Место продажи в торговом зале магазина	0,875
Фактор 7 (Маркетинг)	Уровень развития маркетинга культивируемых грибов	0,543
	Наличие доступной для восприятия информации о культивируемых грибах	0,865
	Время года	0,655
Фактор 8 (Имидж покупателя)	Социальный статус покупателя	0,705
	Уровень культуры и образования покупателя	0,642
Фактор 9 (Образ товара)	Уровень здорового образа жизни покупателя	0,596
	Уровень предложений блюд из культивируемых грибов в местах общественного питания	0,684
	Наличие дней грибной	0,577

Согласно отобранным факторам SPSS были получены девять новых переменных, содержащих вычисленные значения факторов. Эти переменные можно кратко охарактеризовать при помощи следующих категорий: «торговый сервис», «равновесие спроса», «культура питания», «знания покупателя», «имидж производителя», «имидж продавца», «маркетинг», «имидж покупателя», «образ товара». Полученные факторные рейтинги в дальнейшем могут использоваться для построения разрезов, проведения кластерного анализа, пр.

Литература:

1. Хамукова, М.А. Направление развития системы продовольственного обеспечения на уровне региона [Электронный ресурс] // «Инженерный вестник Дона», 2012, №4. – Режим доступа: <http://www.ivdon.ru/magazine/archive/n4p2y2012/1419> (доступ свободный) – Загл. с экрана. – Яз. рус.
2. Kotler Ph., Armstrong G. Principles of Marketing, 14/E. Pearson Prentice Hall. Upper Saddle River, 2012. – 744 p.
3. Malhotra N.K., Birks D.F. Marketing Research: An Applied Approach Trans-Atlantic Publications, Inc., 2008. – 835 с.
4. Хан, Р.С. Оценка эффективности маркетинговых исследований с позиции экономических показателей. [Электронный ресурс] // «Инженерный вестник Дона», 2013, №3. – Режим доступа: <http://www.ivdon.ru/magazine/archive/n3y2013/1936> (доступ свободный) – Загл. с экрана. – Яз. рус.
5. Boone L. E., Kurtz D. L. Contemporary Marketing. 7th ed. The Dryden Press, 1992. – 682 p.
6. Айвазян, С.А. Прикладная статистика: Классификация и снижение размерности [Текст] / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1989. – 607 с.

7. Ким, Дж.-О. Факторный, дискриминантный и кластерный анализ [Текст] / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка. – М.: Финансы и статистика, 1989. – 215 с.
8. Gustav Levine, Sanford L. Braver, David P. Mackinnon, Melanie C. Page, Gustav. Levine's Guide to SPSS for Analysis of Variance. – 2nd ed. – Lawrence Erlbaum Associates; 2nd edition, 2005. – 200 p.
9. Бююль, А. SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей [Текст] / Ахим Бююль, Петр Цёфель. – Спб.: ДиаСофтЮП, 2005. – 608 с.
10. Наследов, А.Д. SPSS: Компьютерный анализ данных в психологии и социальных науках [Текст]. – Изд. 2-е, перераб. и доп. — Спб.: Питер, 2006. – 416 с.