

УДК: 338.43

## ФАКТОРНЫЙ АНАЛИЗ ИЗМЕНЧИВОСТИ ДАННЫХ АГРАРНОГО СЕКТОРА ЭКОНОМИКИ КАЗАХСТАНА

*Дагмирзаев О.А.*

*канд. техн. наук, ст. преподаватель кафедры  
информационно-коммуникационных технологий  
Казахского агротехнического университета  
им.С.Сейфуллина, г.Нур-Султан, Казахстан*

## FACTOR ANALYSIS OF VARIABILITY OF AGRARIAN SECTOR DATA ECONOMY OF KAZAKHSTAN

*Dagmirzaev O.A.*

*Cand. tech. sciences, senior lecturer of the department  
information and communication technologies  
Kazakh Agrotechnical University  
named after S. Seifullin, Nur-Sultan, Kazakhstan*

**Аннотация.** Исследование посвящено раскрытию внутренней структуры взаимосвязей в массиве данных с помощью одного из методов снижения размерности как факторный анализ. Методика использована для оценки изменчивости данных аграрного сектора экономики Казахстана. Статистическая обработка исходных данных производилась с использованием пакета программ STATISTICA–10.

**Abstract.** The study is devoted to the disclosure of the internal structure of relationships in the data set using one of the methods of dimensionality reduction as factor analysis. The technique was used to assess the variability of the data of the agricultural sector of the economy of Kazakhstan. Statistical processing of the initial data was carried out using the STATISTICA–10 software package.

**Ключевые слова:** многомерный статистический анализ, корреляционная матрица, дисперсия, факторный анализ.

**Keywords:** multivariate statistical analysis, correlation matrix, variance, factor analysis.

---

### **Введение.**

Исследования в аграрной сфере экономики, как правило, сводятся к анализу динамики показателей с применением корреляционного и регрессионного анализов. Из обычного анализа динамики показателей тоже можно получить важную информацию о природе и закономерностях изменения данных, но подобное исследование не дает полной картины взаимосвязей.

Из методов многомерного статистического анализа применение снижения размерности, как иногда называют факторный анализ, представляет возможность вскрыть логическую структуру сложного явления [1]. Основное предположение факторного анализа заключается в том, что явления в определенной области исследований можно описать относительно небольшим числом факторов. Предполагается, что в исследуемой области существует закономерность, которая и объединяет изучаемые данные в некоторые группы. Причем считается, что за каждой группой данных стоит определенный фактор. Термин «фактор» здесь используется в смысле объективной причины, которая и объединяет разнородные данные.

Т.е., по сути можно создать статистическую модель задачи, в которой изменчивость всех параметров объясняется выявленными факторами. Вес или значимость того или иного фактора можно оценить значением его доли в общей дисперсии переменных. А теснота связи каждого параметра задачи с конкретным фактором оценивается коэффициентом корреляции.

**Результаты исследования.** Аграрный сектор экономики Казахстана рассмотрен как сложная система, оцениваемая с помощью множества переменных и исследуемая с использованием такого метода многомерного статистического анализа, как факторный анализ.

В качестве источника данных использовали «Основные социально-экономические показатели Республики Казахстан за 1991–2019 г.г.» [2] Проанализированы данные 1993–2018 г.г., т.е. за 26 лет. Из всего перечня данных аграрного сектора экономики страны к анализу были включены следующие показатели: а) ВВП аграрного сектора экономики в целом и по отраслям растениеводства и животноводства; б) данные отрасли растениеводства: валовой сбор урожая и урожайность по видам с/х культур.

Переменные, подвергнутые к статистическому анализу:

ВВП, млн. долларов США: – с/х сектора экономики в целом (v1);

– отрасли растениеводства (v2);

– отрасли животноводства (v3);

Урожайность, ц/га:

зерновые (включая рис) и бобовые культуры; семена подсолнечника; хлопок; свекла сахарная; табак; картофель; овощи открытого грунта.

Валовый сбор урожая, тыс. тонн: зерновые (включая рис) и бобовые культуры; семена подсолнечника; хлопок; свекла сахарная; табак; картофель; овощи открытого грунта.

Показатели ВВП по отраслям растениеводства и животноводства, представленные в тенге, пересчитаны в доллары США с использованием среднегодовых обменных курсов доллара. Статистическая обработка исходных данных производилась с использованием пакета программ STATISTICA–10.

Статистические характеристики результатов обработки первичных данных – показателей ВВП приведены в таблице №1.

Таблица №1

#### Результаты обработки первичных данных

Показатели	Среднее, млн. \$	Стандартное отклонение, млн. \$	Коэффициент вариации, %
ВВП с/х в целом, млн. \$	8524.873	5600.198	65.6
ВВП отрасли растениеводства, млн. \$	4705.826	3159.921	67.1
ВВП отрасли животноводства, млн. \$	3760.138	2492.750	66.2

Изменчивость показателей ВВП за рассмотренный период времени, оцениваемая по значению коэффициента вариации, находилась в пределах 65–67%. С целью анализа взаимосвязей между показателями ВВП в таблице №2 приводятся парные коэффициенты корреляции.

Таблица №2

#### Корреляционная матрица

Переменные:		v1	v2	v3
ВВП с/х в целом	v1	1.00	0.99	0.99
ВВП отрасли растениеводства	v2		1.00	0.97
ВВП отрасли животноводства	v3			1.00

Представляет интерес значение парного коэффициента корреляции между v2 и v3: 0.97. Т.е., показатели ВВП отраслей растениеводства и животноводства являются взаимозависимыми переменными с тесной корреляционной связью. Как известно, парный коэффициент корреляции показывает степень тесноты связи только между двумя переменными при опосредованном влиянии других переменных. Т.е., корреляционный анализ не дает полной картины взаимосвязей между переменными.

Стоит подчеркнуть, что при интерпретации результатов факторного анализа используется один из важных показателей математической статистики, как дисперсия (средний квадрат отклонений), отражающая меру разброса данных вокруг средней арифметической.

Как видно из таблицы №3, основная часть изменчивости переменных (которые включены в анализ), т.е. 87.3 % их общей дисперсии, объясняются тремя факторами. При этом на долю первого фактора приходится 65.5% общей дисперсии переменных, в то время как на долю второго фактора – 13.3%, а на долю третьего фактора – всего 8.5%.

Таблица №3

#### Полная объясненная дисперсия переменных

Факторы:	% общей дисперсии переменных	Кумулятивная дисперсия, %
1	65.5	65.5
2	13.3	78.8
3	8.5	87.3

Проще говоря, почти 90% разброса переменных, оцениваемых их общей дисперсией – это следствие влияния трех факторов. Попробуем разобраться в ситуации.

В таблице №4 приведены коэффициенты корреляции между всеми рассмотренными переменными и выявленными факторами.

**Коэффициенты корреляции между всеми переменными  
и выявленными факторами**

Переменные	Фактор 1	Фактор 2	Фактор 3
ВВП: - с/х в целом	-0.85	0.45	-0.13
- отрасли растениеводства	-0.85	0.44	-0.10
- отрасли животноводства	-0.85	0.44	-0.18
Урожайность: - зерновых культур	-0.72	-0.30	0.18
- семян подсолнечника	-0.84	-0.37	0.10
- хлопка	-0.83	-0.06	-0.04
- свеклы сахарной	-0.87	-0.34	-0.06
- табака	-0.90	-0.15	0.02
- картофеля	-0.98	-0.10	-0.04
- овощей из открытого грунта	-0.96	-0.06	-0.15
Валовой сбор: - зерновых культур	-0.64	-0.11	0.48
- семян подсолнечника	-0.91	0.001	0.22
- хлопка	-0.47	-0.62	-0.46
- свеклы сахарной	0.25	-0.51	0.73
- табака	0.29	-0.77	-0.48
- картофеля	-0.95	0.04	0.18
- овощей из открытого грунта	-0.98	0.01	-0.02
% общей дисперсии	65.5	13.3	8.5

Тесная корреляционная связь (коэффициенты корреляции от -0.72 до -0.96) с первым фактором у всех переменных группы «урожайность». Что касается показателей валового сбора урожая, то здесь ситуация неоднозначная: колебания валового сбора хлопка, сахарной свеклы и табака в большей степени продиктованы влиянием второго и третьего факторов (коэффициенты корреляции: -0.62; 0.73; -0.77). У такого важного показателя, как валовой сбор зерновых культур, коэффициент корреляции с первым фактором только -0.64.

Влияние первого фактора ощутимо и в отношении показателей ВВП, об этом можно судить по значениям коэффициентов корреляции: -0.85; -0.85; -0.85.

Первый фактор, объясняющий изменчивость основной части переменных, по нашему мнению, можно принять за «удельный вес» влияния природно-климатических условий.

Попробуем исключить из дальнейшего анализа переменные группы ВВП. Также будем проводить факторный анализ с данными каждой из оставшихся групп («урожайность» и «валовой сбор») в отдельности. Разумеется, теперь значения коэффициентов корреляции между переменными и факторами будут иные, чем в таблице №4. Здесь важно следующее: в результатах новых анализов сохранятся ли выявленные ранее структуры взаимосвязей, если да, то это признак наличия некоторых объективных закономерностей.

Анализ показал, что колебание урожайности всех культур продиктовано только влиянием первого фактора, т.е. 82.3% общей дисперсии переменных группы «урожайность» предопределены упомянутым фактором (таблица №5).

Таблица №5

**Коэффициенты корреляции между переменными  
группы «урожайность» и выявленным первым фактором**

Урожайность:	Фактор 1
зерновых культур	-0.76
семян подсолнечника	-0.92
хлопка	-0.82
свеклы сахарной	-0.94
табака	-0.94
картофеля	-0.97
овощей из открытого грунта	-0.95
% общей дисперсии	82.3

В таблице №6 приведены результаты факторного анализа данных, сформированные только из показателей валового сбора урожая по всем культурам.

Таблица №6

**Коэффициенты корреляции между переменными группы**

### «валовой сбор урожая» и выявленными факторами

Валовой сбор:	Фактор 1	Фактор 2	Фактор 3
зерновых культур	-0.71	-0.01	-0.44
семян подсолнечника	-0.94	0.06	-0.001
хлопка	-0.37	-0.88	0.04
свеклы сахарной	0.12	-0.02	-0.95
табака	0.40	-0.87	-0.03
картофеля	-0.98	0.007	-0.01
овощей из открытого грунта	-0.95	-0.08	0.19
% общей дисперсии	51.6	22.2	16.4

То, что колебания показателей валового сбора хлопка, сахарной свеклы и табака в большей степени зависят от влияния второго и третьего факторов, мы уже знаем из анализа данных таблицы №1. По нашей версии, показатели валового сбора урожая указанных культур, в меньшей степени зависят от природно-климатических условий (т.е. от влияния первого фактора), поскольку эти культуры возделываются на орошаемых площадях.

Наша гипотеза насчет природы выявленных второго и третьего факторов: считаем, что данные факторы показывают меру влияния иных (кроме природно-климатических условий) объективных причин, которые пока не поддаются логическому объяснению.

Изменчивость показателей валового сбора семян подсолнечника, картофеля и овощей связаны только влиянием первого фактора (коэффициенты корреляции выше 0.9).

Что касается статистической зависимости между данными валового сбора зерновых культур и выявленными факторами (коэффициенты корреляции с первым фактором -0.71 и с третьим фактором -0.44, таблица №6), то попробуем высказать свои предположения. Показатель валового сбора зерновых культур, как мы знаем, включает в себя данные по сбору риса и бобовых культур, возделываемых на орошаемых площадях. При возможности анализа изменчивости валового сбора зерновых культур без учета данных по рису и бобовых культур, вероятно абсолютное преобладание влияния первого фактора. Но это только наши догадки.

**Заключение.** Практической ценностью работы является созданная статистическая модель, в которой изменчивость большого набора параметров объясняется небольшим количеством выявленных факторов. Установлено, что наибольшее влияние на изменчивость рассмотренных данных имеет фактор, названный нами фактором природно-климатических условий. Например, более 80% общей дисперсии показателей урожайности всех сельскохозяйственных культур объясняются только данным фактором. Поэтому считаем актуальным изыскание инновационных технологий, способных обеспечивать устойчивое функционирование отрасли растениеводства вопреки неблагоприятным внешним условиям.

#### Литература

1. Иберла К. Факторный анализ. – М.: Мир, 1980. – 398 с.
2. Данные Комитета по статистике Министерства национальной экономики Республики Казахстан – [Электронный ресурс]. Режим доступа: [www.stat.gov.kz](http://www.stat.gov.kz).

#### References

1. Iberla K. Faktornyy analiz. – M.: Mir, 1980. – 398 p.
2. Data of the Committee on Statistics of the Ministry of National Economy of the Republic of Kazakhstan – [Electronic resource]. Access mode : [www.stat.gov.kz](http://www.stat.gov.kz)