

ИССЛЕДОВАНИЕ СОСТОЯНИЯ ЖИЛИЩНОГО ФОНДА В РОССИИ ПУТЕМ ПРИМЕНЕНИЯ ФАКТОРНОГО АНАЛИЗА

Кладовикова Елена Андреевна

магистрант, ФГБОУ ВО «Поволжский государственный технологический университет», РФ, г. Йошкар-Ола

Бакуменко Людмила Петровна,

научный руководитель, д-р экон. наук, профессор, ФГБОУ ВО «Поволжский государственный технологический университет», РФ, г. Йошкар-Ола

Аннотация. В данной статье рассматривается влияние различных факторов на жилищный фонд в регионах Российской Федерации. Проведен регрессионный анализ данных за 2019 год. Анализ проводился средствами прикладного пакета программ «Statistica».

Ключевые слова: анализ, метод, факторный анализ, фактор, статистика, жилищный фонд.

Факторный анализ – многомерный статистический метод, применяемый для изучения взаимосвязей между значениями количественных переменных. Основная идея факторного анализа заключается в том, что имеющиеся зависимости между большим числом исходных наблюдаемых переменных определяются существованием гораздо меньшего числа скрытых или латентных переменных, называемых факторами.

Главными целями факторного анализа являются: сокращение числа переменных и определения структуры взаимосвязей между переменными. Поэтому факторный анализ используется или как метод сокращения данных или как метод классификации. Факторный анализ позволяет исследователю описать объект измерения с одной стороны всесторонне, учитывая множество исходных тесно взаимосвязанных между собой переменных, а с другой стороны компактно с помощью небольшого числа переменных.

Рассмотрим набор данных из 17 переменных:

Показатели, участвующие в классификации, следующие:

Y1 – Жилищный фонд млн. м²;

X1 – Численность рабочей силы, тыс. чел.;

X2 – Среднедушевые денежные доходы населения, руб./мес.;

X3 – Инвестиции в основной капитал, млн. руб.;

X4 – Количество жилищных ипотечных кредитов, единиц;

X5 – Ввод в действие жилых домов, тыс. м²;

X6 – Число действующих строительных организаций;

X7 – Количество зарегистрированных договоров участия в долевом строительстве, единиц;

X8 – Площадь квартир в жилых зданиях, находящихся в незавершенном строительстве, тыс. M^2 ;

X9 – Удельный вес площади жилищного фонда, оборудованной горячим водоснабжением, в общей площади всего жилищного фонда (процент, значение показателя за год);

X10 – Удельный вес площади жилищного фонда, оборудованной водоотведением (канализацией), в общей площади всего жилищного фонда (процент, значение показателя за год);

X11 – Удельный вес площади жилищного фонда, оборудованной водопроводом, в общей площади всего жилищного фонда (процент, значение показателя за год).

Переменная Y (Жилищный фонд) – зависимая переменная, остальные независимые переменные – факторы или предикторы. Таким образом, воспользовавшись методами статистического анализа необходимо провести множественный регрессионный анализ зависимости жилищного фонда в РФ от ряда факторов, провести факторный анализ по исходным данным о регионах РФ, провести кластеризацию списка регионов РФ в зависимости от ряда показателей, провести классификацию регионов РФ по различным классам, используя дискриминантный анализ.

Для решения задачи числа компонент воспользуемся графическим методом – критерием каменистой осыпи. График собственных значений факторов (график каменистой осыпи) позволяет выбрать число значимых факторов для дальнейшего анализа

Так как метод «каменистой осыпи» графический, он является приближенным и не дает точного значения факторов. В данном случае убывание собственных значений слева направо максимально замедляется на 4 факторе. (Рис. 1).

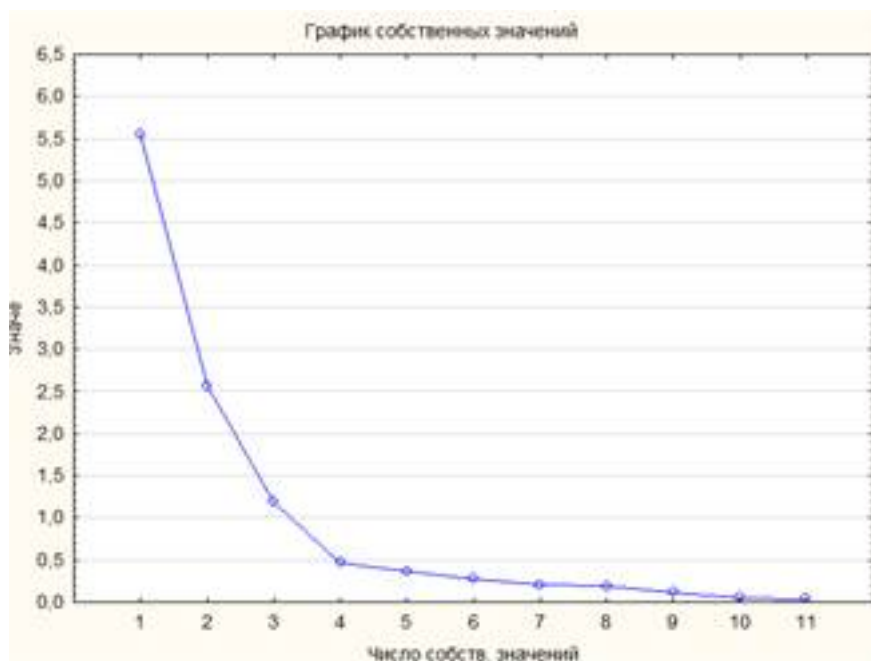


Рисунок 1. График собственных значений

Таблица собственных значений позволяет принять решение о том, сколько факторов нужно оставить в модели. В ней отображены значения в виде абсолютных значений, а также % полной дисперсии, приходящейся на каждый фактор.

Значен.	Собст. значения (Таблица данных1) Выделение: Главные компоненты			
	Собств. Знач.	% общей дисперс.	Кумулятивн. Собств. Знач.	Кумулятивн. %
1	5,559624	50,54204	5,559624	50,54204
2	2,559356	23,26687	8,118981	73,80891
3	1,184313	10,76648	9,303294	84,57540

Рисунок 2. Таблица собственных значений

Как видно из таблицы, собственное значение для первого фактора равно 5,559624, а доля дисперсии, объясненная первым фактором равна приблизительно 50,54%. Собственное значение для второго фактора равно 2,559356, а доля дисперсии, объясненная вторым фактором равна приблизительно 23,27%. Собственное значение для третьего фактора равно 1,184313, а доля дисперсии, объясненная третьим фактором равна приблизительно 10,77%.

Можно сделать вывод, что первый фактор наиболее значимый, т.к. он объясняет почти половину вариаций переменных.

В соответствии с критерием Кейзера, необходимо оставить факторы с собственными значениями больше 1. Из таблицы собственных значений следует, что критерий приводит к выбору полученных трех факторов.

Факторный анализ. Далее следует рассмотреть факторные нагрузки. Чем теснее связь с рассматриваемым фактором, тем выше значение факторной нагрузки. Если коэффициент факторной нагрузки больше 0,7, то он выделяется красным цветом и соответствующая переменная включается в состав фактора

Рассмотрим факторные нагрузки без вращения факторов (Рис.3).

Перемен.	Общности (Таблица данных1) Выделение: Главные компоненты Вращение: Без вращ.			
	Из 1 Фактора	Из 2 Факторов	Из 3 Факторов	Множест. R-квадр.
X1	0,814831	0,885565	0,895570	0,908775
X2	0,040451	0,249699	0,877157	0,453040
X3	0,417756	0,455998	0,816062	0,584042
X4	0,833387	0,890669	0,890670	0,917806
X5	0,740296	0,820448	0,826469	0,817518
X6	0,751012	0,812515	0,814435	0,817048
X7	0,634890	0,730029	0,730555	0,704920
X8	0,613307	0,685866	0,694622	0,607915
X9	0,246531	0,825420	0,857878	0,764243
X10	0,255444	0,924295	0,967848	0,946080
X11	0,211719	0,838477	0,932029	0,921599

Рисунок 3. Факторные нагрузки главных компонент без вращения

Поскольку в этой модели нет переменных, которые можно включить в факторы 1, 2 и 3.

Поэтому модель нельзя считать информативной. Если используем вращение факторов «Варимакс исходных», то получим другие факторные нагрузки (рис. 4)

Перемен.	Фактор.нагрузки (Вар.исходн.) Выделение: Главные компонен (Отмечены нагрузки >,700000)		
	Фактор 1	Фактор 2	Фактор 3
X1	0,933958	0,150639	0,024518
X2	-0,074099	0,181991	0,915721
X3	0,453268	0,196023	0,756429
X4	0,924409	0,141646	0,126786
X5	0,901646	0,111953	0,031139
X6	0,883188	0,100669	0,155821
X7	0,851290	0,046640	0,060709
X8	0,827390	0,100174	0,003546
X9	0,132758	0,901714	0,164816
X10	0,118351	0,963942	0,157023
X11	0,096007	0,959378	0,049056
Общ.дис.	4,980855	2,811858	1,510581
Доля общ.	0,452805	0,255623	0,137326

Рисунок 4. Факторные нагрузки главных компонент с вращением факторов «Варимакс исходных»

Задача вращения обобщающих факторов – главных компонент решается с целью улучшения их интерпретируемости.

Значимость признаков, участвующих в формировании главной компоненты, можно установить коэффициентом информативности. Набор признаков считается удовлетворительным,

если $K_{инф} \in (0,75; 0,95)$. Коэффициент информативности рассчитывается по следующей формуле:

$$K_{инф} = \sum a_i^2 (> 0,7) \sum a_i^2 (все) * 100\%$$

$$K_{инф1}=84\%; K_{инф2}=55\%; K_{инф3}=0\%$$

Фактор 1 «Объем услуг жилищного фонда»

X1 - Численность рабочей силы, тыс. чел.;

X4 - Количество жилищных ипотечных кредитов, единиц;

X5 - Ввод в действие жилых домов, тыс. м²;

X6 - Число действующих строительных организаций;

X7 - Количество зарегистрированных договоров участия в долевом строительстве, единиц;

X8 - Площадь квартир в жилых зданиях, находящихся в незавершенном строительстве, тыс. м².

Кинф1 = 94,95%, что в пределах допустимого уровня.

Фактор 2 «Жилищные условия домохозяйств»

X9 - Удельный вес площади жилищного фонда, оборудованной горячим водоснабжением, в общей площади всего жилищного фонда (процент, значение показателя за год);

X10 - Удельный вес площади жилищного фонда, оборудованной водоотведением (канализацией), в общей площади всего жилищного фонда (процент, значение показателя за год);

X11 - Удельный вес площади жилищного фонда, оборудованной водопроводом, в общей площади всего жилищного фонда (процент, значение показателя за год).

Кинф2 = 94,69%, что в пределах допустимого уровня.

Фактор 3: «Доходы потребителя»

X2 - Среднедушевые денежные доходы населения, руб./мес.

X3 - Инвестиции в основной капитал, млн.руб.

Кинф3 = 93,39%, что в пределах допустимого уровня.

Таким образом, применив вращение факторных нагрузок, значения коэффициентов информативности немного изменились (увеличились).

Регрессионный анализ полученных компонент. Чтобы проверить, являются ли входящие в фактор переменные значимыми, нужно применить регрессионный анализ для каждой компоненты. На рисунке 5-6 представлен регрессионный анализ факторов.

Итоги регрессии для зависимой переменной: Y1 (Таблица да R= ,94762735 R2= ,89799760 Скоррект. R2= ,89397119 F(3,76)=223,03 p<0,0000 Станд. ошибка оценки: 9,0217						
N=80	БЕТА	Ст.Ош. БЕТА	В	Ст.Ош. В	t(76)	p-знач.
Св.член			37,9375	1,008655	37,6120	0,000000
ФАКТОР1	-0,893514	0,036635	-24,7558	1,015019	-24,3895	0,000000
ФАКТОР2	0,291800	0,036635	8,0846	1,015019	7,9650	0,000000
ФАКТОР3	-0,120350	0,036635	-3,3344	1,015019	-3,2851	0,001544

Рисунок 5. Итоги регрессии для компонент фактора 1

Из таблицы видно, что все 3 фактора являются значимым по отношению к жилищному фонду в РФ, так как они являются статистически значимыми.

Уравнение множественной регрессии в обычной форме будет выглядеть следующим образом:

$$Y = 37,94 - 24,76 * \Phi_1 + 8,08 * \Phi_2 - 3,33 * \Phi_3$$

Таким образом, при увеличении второго фактора на 1 млн. м². жилищный фонд увеличится на 8,08 млн. м².

Таким образом, на жилищный фонд в РФ из всех исследуемых характеристик оказывают существенное все 3 главные компоненты: объем услуг жилищного фонда, жилищные условия домохозяйств, доходы потребителя. Для проверки адекватности модели важна гистограмма

остатков, которая должна быть близка к графику нормального распределения, что и наблюдается в данном случае (рисунок 6).

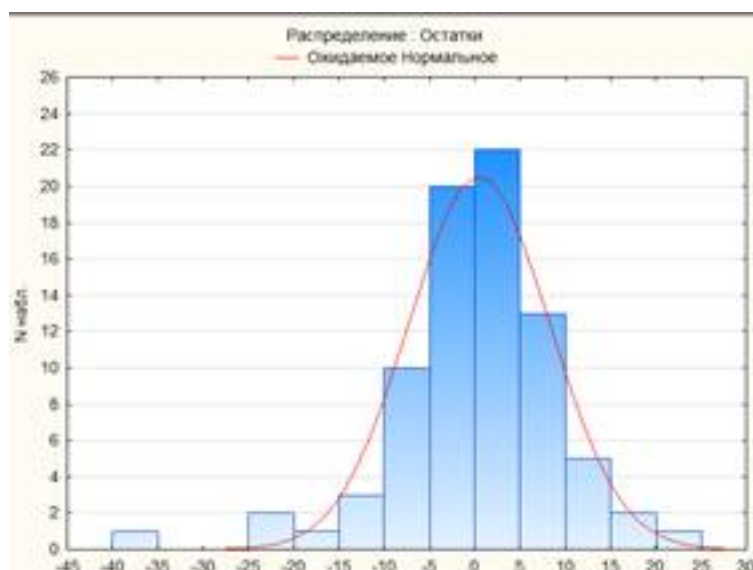


Рисунок 6. Гистограмма остатков

Она должна быть близка к графику нормального распределения, что и наблюдается в данном случае.

Таким образом, можно сделать вывод, что использование факторного анализа позволяет:

- определить основные аспекты различий между факторами;
- обосновать существование доминирующих факторов;
- осуществить позиционирование регионов по уровню региональных различий в разрезе вычисленных главных факторов.

Список литературы:

1. Шеремет Н. Г., Коротков А. В. Экономико-статистический анализ рынка жилья // Молодой ученый. — 2019. — №5. — С. 446-452. — URL <https://moluch.ru/archive/109/26373/> (дата обращения: 30.11.2020).
2. Кремер Н.Ш., Путко Б.А. Эконометрика: Учебник для вузов / Под ред. проф. Н.Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2010. – 311 с.
3. Официальный сайт Федеральной службы государственной статистики – Росстат [Электронный ресурс] // <http://www.gks.ru/> (23.11.2020)