

ПРИМЕНЕНИЕ РЕГРЕССИОННОГО АНАЛИЗА ДЛЯ ВЫЯВЛЕНИЯ ФАКТОРОВ, ВЛИЯЮЩИХ НА ЧИСЛЕННОСТЬ НАСЕЛЕНИЯ

Луцейн Анна Андреевна

магистрант Поволжский государственный технологический университет, РФ, г. Йошкар-Ола

Бакуменко Людмила Петровна

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

Аннотация. Статистика населения – демографическая статистика, или просто демография, – один из важнейших разделов социально-демографической статистики. В статье рассматривается применение регрессионного анализа для выявления влияния социально-экономических показателей на численность населения.

Ключевые слова: регрессионный анализ, численность населения.

Одной из важных потребностей в управлении социально-экономическими процессами является изучение численности и состава населения, так как с одной стороны население выступает в качестве производителя национального богатства, а с другой стороны — является потребителей духовных и материальных ценностей. Широкий круг проблем, прямо или косвенно связанных с населением. Исходя из этого — существует много факторов, напрямую влияющие на численность населения, которые рассмотрены в данной статье.

Цель данной работы - провести анализ по всем субъектам, в которых определяются важные показатели численности населения. Такой анализ позволит субъектам выделить слабые места, в которые стоит вкладывать средства для повышения численности населения.

Для проведения анализа выбрано 76 регионов РФ. Выявлен ряд важных показателей, которые влияют на численность населения.

В анализе зависимой переменной является y - численность населения, тыс. чел., а независимыми переменными являются: X_1 – Коэффициент рожившихся (чел. на 1000 человек населения), X_2 – Коэффициент умерших (чел. на 1000 человек населения), X_3 – Естественный прирост/убыль (чел.), X_4 – Продолжительность жизни (лет), X_5 – Количество заключенных браков (чел. на 1000 человек населения) X_6 – Количество разводов (чел. на 1000 человек населения) X_7 – Коэффициент миграционного прироста (чел. на 1000 человек населения), X_8 – Величина прожиточного минимума (руб.), X_9 – Среднемесячная заработная плата работников (руб.), X_{10} – Удельный вес автомобильных дорог с твердым покрытием (общая протяженность), X_{11} – Ввод в эксплуатацию жилых домов (тыс. м²), X_{12} – Число зарегистрированных преступлений (ед. на 100 000 чел. населения), X_{13} – Численность безработных (тыс. чел.), X_{14} – Индексы потребительских цен (%), X_{15} – Число дошкольных учреждений (ед.), X_{16} – Число образовательных учреждений, реализующих программы среднего профессионального образования (ед.), X_{17} – Число образовательных учреждений, реализующих программы высшего профессионального образования (ед.), X_{18} – Численность населения на одну больничную койку (на конец года; человек).

В связи с небольшим количеством наблюдений – 76, для проведения анализа использовался метод моделирования измеряемых данных и исследования их свойств.

Регрессионный анализ, пожалуй, является наиболее широко используемым статистическим методом для исследования или оценки взаимосвязи между зависимыми и набором независимых объясняющих переменных. Он также используется в качестве общего термина для обозначения различных методов анализа данных, которые используются в качественном методе исследования для моделирования и анализа многочисленных переменных. В методе регрессии зависимая переменная является предиктором или объясняющим элементом, а зависимая переменная - результатом или ответом на конкретный запрос. Суть регрессионного анализа заключается в нахождении наиболее важных факторов, которые влияют на зависимую переменную.

Построив и проанализировав матрицу парных коэффициентов корреляции были выявлены что «Ввод в эксплуатацию жилых домов» (X_{11}) имеют коэффициент корреляции равный 0,87, «Численность безработных» (X_{13}) имеет коэффициент 0,83, «Число дошкольных учреждений» (X_{15}) имеют коэффициент 0,91, «Число образовательных учреждений, реализующих программы среднего профессионального образования» (X_{16}) имеют коэффициент 0,89, «Число образовательных учреждений, реализующих программы высшего профессионального образования» (X_{17}) имеют коэффициент 0,87. Вышеописанные факторы влияют на «Число населения» больше чем все остальные.

Следующим шагом будет сравнение процедуры «пошагово с включением» и «пошагово с исключением», а также сравнение с данными процедурами после стандартизации значений. После сравнения полученных результатов было решено использовать итоги регрессии с использованием процедур «Пошаговая с включением», так как в данной модели стандартная ошибка меньше чем во всех остальных, а R^2 больше.

Построим уравнение регрессии:

$$Y = 0,19X_{15} + 0,37X_{11} + 0,29X_{13} + 0,14X_{16} + 0,14X_{17} - 0,09X_1 - 0,02X_{14}$$

$$F(7,74) = 667,91$$

$$F_{кр}(0,05;7,74) = 2,02$$

$R_2 = 0,98$ – Коэффициент линейной детерминации показывающий зависимость одной переменной от других.

$$\text{скор. } R_2 = 0,98$$

$$\text{Стандартная ошибка оценки (S)} = 166,76$$

$$\text{Статистика Дарбина-Уотсона (Dw)} = 1,84$$

Нормальность распределения переменной оценивается по полученным остаткам. Для этого построен нормальный вероятностный график остатков (Рисунок 1) и распределение остатков (ожидаемое нормальное) (Рисунок 2).

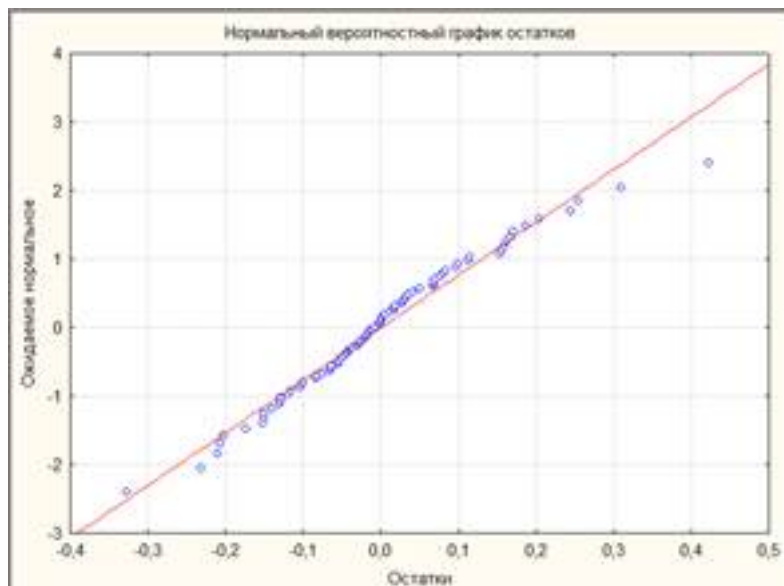


Рисунок 1. Нормальный вероятностный график остатков

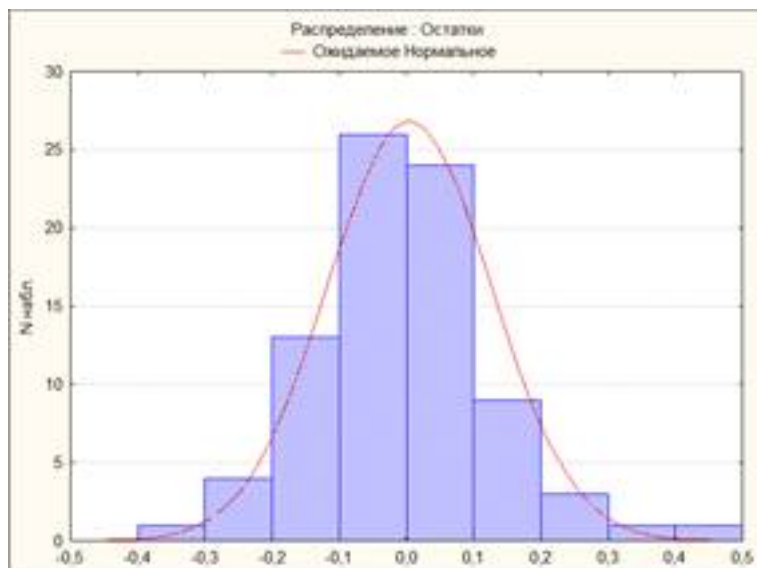


Рисунок 2. Распределение остатков

По графикам видно, что остатки образуют прямую зависимость и подчиняются нормальному закону распределения.

Так как R^2 больше 0,7, то это говорит нам о средней точности аппроксимации, то есть модель хорошо описывает явление. Уравнение значимо так как $F > F_{кр}$ и гипотеза об отсутствии линейной связи отклоняется.

В результате регрессионного анализа выяснилось:

Если «Число дошкольных учреждений» (X_{15}) увеличить на 1 единицу, то «Численность населения» увеличится на 1 тыс. человек при прочих неизменных факторах.

Если «Ввод в эксплуатацию жилых домов» (X_{11}) увеличить на 1 тыс. кв. метров, то «Численность населения» увеличится на 0,37 тыс. человек при прочих неизменных факторах.

Если «Число образовательных учреждений, реализующих программы среднего

профессионального образования» (X_{16}) увеличить на 1 единицу, то «Численность населения» увеличится на 7,55 тыс. человек при прочих неизменных факторах.

Если «Число образовательных учреждений, реализующих программы высшего профессионального образования» (X_{17}) увеличить на 1 единицу, то «Численность населения» увеличится на 22,67 тыс. человек при прочих неизменных факторах.

Список литературы:

1. Крамер, Н.Ш., Путко, Б.А. Эконометрика: Учебник для вузов [Текст]/ Под рек.проф. Н.Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2012. – 311 с.
2. Федеральная служба государственной статистики [Электронный ресурс] – Режим доступа: <https://rosstat.gov.ru/emiss> (дата обращения 20.11.2022).
3. Харисова А.Ф., Бакуменко Л.П. Применение метода главных компонент для анализа производственных показателей на предприятиях // Экономика и менеджмент инновационных технологий. 2017. № 2 [Электронный ресурс]. URL: <https://ekonomika.snauka.ru/2017/02/13907> (дата обращения: 26.11.2022).