

§ 3. Проверка адекватности модели. Исследование остатков

3.1. Адекватность

Адекватная модель должна отражать закономерности, которые имеются в исходных данных, соответствовать реальности. В узком смысле адекватность понимается как постоянство (в среднем) отклонений реальных наблюдений от модели. Отличить адекватную модель от неадекватной можно с помощью проверки гипотезы о постоянстве математического ожидания ошибок. В [1, с.48 - 50] рекомендуется использовать для этого критерий, основанный на дисперсионном анализе. Этот критерий можно использовать, если имеются повторные наблюдения отклика при постоянстве факторов. В противном случае данные нужно сгруппировать, так же как при построении гистограммы: интервалов должно быть не слишком много и не слишком мало, а в каждый интервал должно попадать приблизительно одинаковое количество наблюдений.

Более полное исследование остатков должно включать в себя не только проверку постоянства математического ожидания, а всех предположений регрессионного анализа:

- постоянство дисперсии;
- некоррелированность;
- нормальное распределение остатков.

Эти свойства (кроме нормального распределения) зависят от того, в каком порядке расположены остатки. Всегда можно их отсортировать в порядке возрастания или убывания – тогда ни математическое ожидание, ни дисперсия не будут казаться постоянными, всегда будет автокорреляция. Поэтому перед исследованием необходимо отсортировать данный в порядке возрастания одного из факторов (функция `sortrows()`). Затем следует повторить эту процедуру для всех факторов, которые не зависят монотонно друг от друга.

Исследование каждого предположения должно включать в себя как графические интуитивные методы, так и формальную проверку гипотез. Формальные критерии должны учитывать количество оценённых параметров и зависимость остатков между собой даже для адекватной модели.

Рекомендуется применять следующие методы:

1. для проверки постоянства математического ожидания
 - график остатков (функция `plot()`);
 - дисперсионный анализ ошибок ε_i через остатки e_i (напишите свой вариант функции `anova1()`);
2. для проверки постоянства дисперсии:
 - график остатков;
 - критерий Уайта (напишите свою функцию);
 - критерий Голдфелда-Куандта (напишите свою функцию);
3. для проверки отсутствия автокорреляции:
 - коррелограмма – график автокорреляционной функции (функция `autocorr()`);
 - Q-критерий Льюнга-Бокса (функция `lbqtest()` – укажите в опциях правильное кол-во степеней свободы);

4. для проверки согласия с нормальным распределением:

- гистограмма + график плотности (функция `histfit()`);
- вероятностная бумага (функция `normplot()`);
- χ^2 -критерий Пирсона-Фишера (функция `chi2gof()` – укажите в опциях правильное кол-во степеней свободы).

3.2. Задача

Исследуйте остатки наилучшей модели, построенной в предыдущем задании.

3.3. Литература

1. Сухарев М.Г. Методы прогнозирования. М.: МАКС Пресс, 2010. – 176 с.